

Molecular Simulations Studies of the Effect of Ligand Architecture on DNA Binding

by

Alex VanFosson

B.S.E., University of Iowa, 2011

A thesis submitted to the Faculty of the Chemical
and Biological Engineering department of the
University of Colorado at Boulder in partial
fulfillment of the requirement for the degree of
Master of Science

2013

Signature Page

This thesis entitled:

Molecular Simulations Studies of the Effect of Ligand Architecture on DNA Binding

written by: Alex VanFosson

has been approved for the department of Chemical and Biological Engineering

Arthi Jayaraman - Chair

Joel Kaar

Charles Musgrave

Date: _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

VanFosson, Alex (M.S. Chemical and Biological Engineering)

Molecular Simulations Studies of the Effect of Ligand Architecture on DNA Binding

Thesis directed by Arthi Jayaraman

Abstract

In this paper we use atomistic molecular dynamics simulations to study the structural reasons underlying the DNA binding efficacy of several polyamine-aminoglycoside compounds. We calculate the free energy of binding to DNA and conformational entropy loss upon binding for spermine-aminoglycoside compounds and dilysine-aminoglycoside compounds. We also calculate the structural features of the ligands and DNA before and after binding through the radius of gyration, the width of the ligand, the end-to-end distance of the grafts, the center-of-mass distance between the DNA and the ligand, and the distance between the ligands' amine groups and the DNA's phosphate groups. In order to understand the trends in DNA binding efficacy of spermine-aminoglycoside and dilysine-aminoglycoside compounds, we compare results from polyamine-aminoglycoside systems with ungrafted polyamines (spermine and dilysine) to isolate the effects of grafting spermine and dilysine to an aminoglycoside compound on binding behavior. First, we find that grafting spermine to an aminoglycoside compound improves binding efficiency to DNA over spermine alone in agreement with the experimental results of DNA binding. We discover that the improved binding is due to a decrease in the rate of the grafts

unbinding from the DNA. We also find the spermine-aminoglycoside compounds bound to DNA with a greater efficacy than the dilysine-aminoglycoside compounds. Spermine is a longer, unbranched molecule which is more flexible and adept at optimizing its binding location on the DNA strand.

Dedication
To Sarah

Acknowledgments

I would like to thank Janus supercomputing and Teragrid supercomputing for the use of their facilities. I would also like to thank our funding sources: the NIH Biotech training grant and the NSF CBET 1066998 (PI Jayaraman & Rege) grant. I would like to thank my adviser, Prof. Arthi Jayaraman, for her help, advice, and leadership. Finally, I would like to thank members of our research group: Mr. Xiao Ba, Mr. Robert Elder, Mr. Tyler Martin, Dr. Eric Jankowski, and Ms. Arezou Seifpor for various ideas, help with code, and other contributions to the work presented here.

Table of Contents

Table of Contents

Signature Page	ii
Abstract.....	iii
Dedication	iv
Acknowledgments.....	vi
Table of Contents.....	vii
List of Tables	ix
List of Figures	x
I. Introduction.....	1
II. Methods	4
A. Atomistic Model	4
B. Simulation Protocol	5
C. Systems Studied	6
D. Thermodynamic Analyses	7
1. MMPBSA	7
2. Conformational Entropy	8
E. Structural Analyses	9
1. Maximum Diameter of the Ligand	9
2. Radius of Gyration	9
3. End-to-End Distance of the Grafts	10
F. Ligand to DNA Distance Measurements.....	10
1. Center-of-Mass Distance	11
2. Average distance of Ligand Nitrogen Atoms to DNA Phosphorus Atoms	11
3. Number of Ligand Nitrogen Atoms within Four Angstroms of a DNA Phosphorus atom	11
III. Results and Discussion	13
A. Spermine	13
B. Comparison of results so far with spermine-DNA binding results from experiment:.....	20
C. Dilysine	21
IV. Conclusion	27
V. References.....	30
VI. Appendix.....	33

A. Protocol for Using Antechamber to Derive Ligand Parameters.....	33
B. Determination of Equilibrium.....	33
C. Supplementary Figures and Movies	35
D. Other Investigated Metrics	41
1. Radius of Gyration of the Core	41
2. Phosphorus-to-nitrogen Distance.....	42
3. Number of Ligand Nitrogen Atoms within Four Angstroms of a DNA Phosphorus atom.....	43

List of Tables

Table 1 Legend for Abbreviated Ligand Names	4
---	---

List of Figures

Figure 1	13
Figure 2	15
Figure 3	19
Figure 4	21
Figure 5	23
Figure 6	24
Appendix Figures	
Figure A7	36
Figure A8	37
Figure A9	38
Figure A10	38
Figure A11	39
Figure A12	39
Figure A13	39
Figure A14	40
Figure A15	40
Figure A16	41
Figure A17	41
Figure A18	41
Figure A19	42
Figure A20	42
Figure A21	43
Figure A22	44
Figure A23	45

I. Introduction

DNA-ligand binding is an important phenomena occurring in many chemical and biological applications. Gene therapy in particular has been the focus many in the scientific community as means for curing many genetic diseases. Currently, the most difficult hurdle for gene therapy is the delivery of the therapeutic DNA strands to the target cell. Many delivery vectors are cationic ligands^[1] that bind to DNA to form positively charged polyplexes and deliver therapeutic DNA with lower immunogenic responses as compared to viral vectors^{[2][3][4]}. In particular, polyethylenimine (PEI)^{[5][6]} is a widely used polycation for DNA binding. PEI-DNA polyplexes have a high transfection efficiency; however, PEI is also cytotoxic^[6], but it has a low toxicity^[7] relative to other polycations. A polycation with high transfection efficiency can still be slightly cytotoxic because less of the drug is needed for effective gene delivery. Spermine, on the other hand, is significantly less toxic than PEI and is naturally found in bacteria and eukaryotic cells for the express purpose of condensing DNA^[8]. Spermine is known to be the most effective at DNA condensation of the many naturally occurring polyamines^{[9][10]}. Therefore, spermine is a good candidate as a non-toxic alternative to PEI. However, its transfection efficiency is lower than that of PEI, so improving spermine's transfection efficiency without increasing its toxicity is desirable.

Recent experimental work focused on finding compounds that could replace PEI found that when spermine is grafted to an aminoglycoside compound the DNA binding efficacy of the spermine grafts is improved over ungrafted, or free, spermine. We use atomistic simulations to elucidate the molecular-level phenomena that explain why the grafting of spermine improves its DNA binding efficacy over free spermine. Atomistic simulations using the AMBER force field are useful for studying biomacromolecular binding interactions as the simulations provide details of non-covalent bonds (e.g. H-bonds) and atomistically detailed physical organization of the molecules with a high amount of

accuracy due to decades of testing^[11]. This detailed information allows us to know why and how things bind to DNA. It also sheds light on why certain compounds bind better to DNA than other compounds.

Atomistic simulations have shed light on why spermine-aminoglycoside compounds bind stronger and tighter than spermine alone. To the best of our knowledge there are no prior studies investigating the binding behavior of polyamine-aminoglycoside compounds; however, spermine has been studied extensively due to its biological role in condensing DNA^{[12][13][14]}. From these studies it is known that spermine binds primarily in the major groove of B-DNA strands. Spermine does not disrupt the B-form structure of the DNA even after condensation. The binding is driven by electrostatics as spermine attempts to neutralize as many of the negative charges on the DNA strand.

Polyamine-aminoglycoside compounds differ from spermine in that they have a core-graft architecture. The aminoglycoside molecule serves as the core and the polyamines are grafted onto the amine groups of the aminoglycoside molecule via a carboxylic acid linker. In this work we studied the effect of changing the graft chemistry (spermine and dilysine grafts) and the effect of core chemistry (neamine, kanamycin, and neomycin) on the thermodynamics and structure of DNA-ligand binding. The number of ligands in the simulation box was varied to analyze the effect of ligands crowding each other when bound to the DNA. We also investigated the effects of increasing salt concentration on binding. The results from the core-grafted compounds were compared against simulations that contained only spermine or only dilysine (free grafts) in order to elucidate the effect of grafting compounds to aminoglycoside cores. We discovered that the binding of the polyamines is improved by grafting them to a core regardless of the graft chemistry. This is in accordance with the experiments by Rege and coworkers^{[9][15]} who found that spermine grafted compounds, neamine tetraspermine (NTS) and kanamycin tetraspermine (KTS) bound more strongly to DNA than the ungrafted free spermine. The underlying reason for these trends lies in the way the ligand's physical features facilitate binding and

unbinding. If a free graft becomes unbound from the DNA it may or may not rebind; however, if a graft that is attached to a core molecule becomes unbound, then it quickly binds again because it is held in close proximity to the DNA by the core which is in turn held by the other grafts. We also discovered that the chemistry of the grafts affects the binding behavior of the ligand with spermine grafted ligands binding to DNA with greater efficacy than dilysine grafted ligands because Spermine is a longer, unbranched molecule and is more flexible than dilysine. The flexibility of spermine makes it more adept at optimizing its binding location on the DNA strand. Knowing why compounds bind helps to guide synthesis of new compounds and improves the efficiency of finding new compounds. We believe the findings in this paper will help experimentalists develop better polyamine-aminoglycoside compounds.

II. Methods

A. Atomistic Model

We use atomistic molecular dynamics simulations to study systems of a single DNA duplex binding with various ligands which are described in Table 1 and whose structures are shown in Figure A8. Each ligand is constructed using the VegaZZ software package^[16]. The generalized AMBER force field (GAFF)^[17] is used to model bonded and non-bonded interactions of the ligands' atoms. Appendix section A contains details on how the antechamber program assigns force field parameters and charges to the ligands. The AMBER *ff99 bsc0*^[18] force field is used to model bonded and non-bonded interactions involving the DNA's atoms. The B-form DNA strand has the sequence 5'-d(GCGTCCAGGTCTACC)-3' and is the same sequence for all of the simulations. It was constructed using nucleic acid builder (NAB) from the AMBER 11 software package^[19]. The simulations use explicit Na⁺ and Cl⁻ counter ions parameterized with GAFF. The TIP3P model is used to represent water explicitly.

Table 1 Legend for Abbreviated Ligand Names

Short Name	Core(s)	Graft(s)	Number of atoms in the system
4xS1	no core	4 spermine	180
NTS	neamine	4 spermine	220
2xNTS	2 neamine	8 spermine	440
KTS	kanamycin	4 spermine	241
2xKTS	2 kanamycin	8 spermine	482
NHS	neomycin	6 spermine	346
4xL2	no core	4 dilysine	188
NL2	neamine	4 dilysine	228
2xNL2	2 neamine	8 dilysine	456
KL2	kanamycin	4 dilysine	249
2xKL2	2 kanamycin	8 dilysine	498
NHL2	neomycin	6 dilysine	352
12-PLL	A 12-mer linear chain of lysine amino acids		267

B. Simulation Protocol

The DNA and ligand(s) are placed in a $\sim 80 \text{ \AA} \times 80 \text{ \AA} \times 80 \text{ \AA}$ box far enough apart that the system can sample several possible binding states but close enough that the DNA and ligand(s) bind quickly. This distance varies depending on the size of the ligand added to the box. Solvent ions are added next. Two different concentrations of counterions were tested: 0mM and 160mM. The 0mM concentration simulations only contain enough Na^+ ions to neutralize the simulation box. The 160mM concentration simulations contain additional ions beyond those needed to neutralize the system. It is from these additional ions that the salt concentration of the system is calculated; i.e. the values 0mM and 160mM represent the value of the chloride ion concentration. The final addition to the simulation box is the explicit solvent, water.

We perform simulations using the NAMD simulation software package^[20]. Each system is first minimized using the conjugate gradient method and then the simulation box is allowed to shrink as pressure is applied slowly for no more than 10,000 time steps (20ps). The systems are then heated from 100K to 300K for 5,500 steps (1.1 ps) to initialize the system and avoid unstable overlaps. Finally, the simulations are run for 10 million time steps (20 ns). Coordinates are written to the trajectory files every 1000 time steps. Long range electrostatic interactions were handled using the particle-mesh Ewald (PME) method^[21]. An NPT ensemble is used for all of the systems with temperature set to 300K and pressure set to 1 atm. Pressure and temperature are maintained using Langevin dynamics. Each trial is checked to ensure that equilibrium has been reached after 20 ns of simulation. If equilibrium has not been reached, then the trial is run for additional time, usually 20 ns more. Appendix section B contains information on how equilibrium was determined.

We use a “multitrajectory”^[22] approach for each system in this work. This means that for each system there are three trajectories. One trajectory contains DNA and ligand(s) together, one trajectory

contains the DNA alone, and one trajectory contains the ligand(s) alone. Simulating the bound and unbound states separately enables us to compare the two states and obtain more accurate thermodynamic data. To clarify, the ligand and DNA are not bound at the start of the trajectory in which they are together. The DNA and ligand(s) start from an unbound state, approach each other and reach a bound state. Once the simulations are finished, the bound portions of the trajectories are analyzed to obtain data for several different metrics.

C. Systems Studied

The library of molecules studied by Rege et al.^[9] consists of polyamines or polyamine grafted aminoglycoside compounds. We refer to these compounds as free grafts and core-grafted architectures respectively. We investigate three aminoglycoside cores: kanamycin, neamine, and neomycin, which are shown in Figure A8. Each core is grafted with four or six grafts. The grafts are either spermine or dilysine (two L-lysine amino acids joined by a peptide bond). Lysine was studied in order to have a different chemistry with which to compare spermine. In order to investigate a different architecture than a core-grafted architecture, we look at a polymer of twelve L-lysine amino acids (12-PLL). Table 1 summarizes the different ligand abbreviations, number of grafts, and number of atoms.

Most of the systems we study contain a single ligand in the simulation box, and all of the systems contain a single DNA strand. Some systems contain more than one ligand in the simulation box for one of two reasons: to provide a consistent amount of electrostatic attraction or to investigate the effects of ligands crowding each other around the DNA as they bind. Each spermine and dilysine graft has a charge of +3. Kanamycin and neamine core systems have an overall charge of +12 because of their four grafts. Ligands with a neomycin core have an overall charge of +18 because they have 6 grafts. In order to keep the electrostatic attraction between the ligand and DNA constant between the free graft systems and the core-grafted systems, 4 or 6 grafts were used in the free graft systems. The systems

with 4 free grafts (4xS1 and 4xL2) are comparable to the kanamycin and neamine core systems. The systems with 6 free grafts (6xS1 and 6xL2) are comparable to the neomycin core systems. Systems with a 2x in front of their name in Table 1 have two core-grafted ligands in the simulation box. These are the systems used to study crowding around the DNA. They do not have any corresponding free graft systems because they have an overall ligand charge of +24.

D. Thermodynamic Analyses

1. MMPBSA

The free energy of each system is calculated using a method called MMPBSA^{[23][24]}. The configurational entropy of each system is calculated separately and will be discussed in detail later. The MMPBSA method is implemented with the software program MMPBSA.py^[25] found in the AmberTools software suite. MMPBSA.py calculates the different contributions of the free energy and then combines them to determine the free energy of binding as shown in equations 1-3:

$$\Delta G_{bind} = (E_{MM}^{complex} - E_{MM}^{receptor} - E_{MM}^{ligand}) + (\Delta G_{solv}^{complex} - \Delta G_{solv}^{receptor} - \Delta G_{solv}^{ligand}) - T\Delta S_{bind} \quad (1)$$

$$E_{MM} = E_{vdw} + E_{elec} + E_{bond} + E_{angle} + E_{dihedral} \quad (2)$$

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{nonpolar} \quad (3)$$

Specifically, MMPBSA.py calculates the non-bonded interactions, van der Waals potential (E_{vdw}) and electrostatic (E_{elec}) potential; the bonded interaction potentials: bond (E_{bond}), angle (E_{angle}), and dihedral ($E_{dihedral}$); and the solvation free energy (ΔG_{solv}). The solvation free energy is divided into two terms: polar and non-polar. The polar term (ΔG_{polar}) is calculated using the Poisson-Boltzmann equation. The non-polar term ($\Delta G_{nonpolar}$) can be calculated using a number of different methods, but we calculate it by finding the energy required to create a cavity in water the same size and shape as the molecule or

complex of interest i.e. the bound ligand and DNA. MMPBSA.py then combines each of the aforementioned terms with the (normal mode) entropy of the system (ΔS_{bind}) to get the free energy. The separation of the binding free energy into several terms is useful in determining what the primary driving force of binding is. There is an unfortunate side effect though. The error in the largest terms of the binding free energy is on the same order of magnitude as the sum of the individual terms. This makes it very difficult to use the binding free energy as a means to compare the different DNA-ligand systems. This will be discussed in more detail later.

2. Conformational Entropy

We also look into the conformational entropy loss of the ligand and DNA upon binding. We are interested in this term for two reasons. First, normal mode analysis as calculated by MMPBSA is known to underestimate the entropy^[26] change upon binding. Second, the ligands' grafts are very mobile when unbound, but the grafts lose much of their mobility when bound. The loss in mobility directly correlates to a negative change in conformational entropy of the system. Conformational entropy loss of macromolecules can be estimated by measuring the change in entropy of the dihedral angles using the method detailed in reference^[27].

$$S_{conf} = \sum_{k=1}^K (S_k - S_0) \quad (4)$$

$$S_k = -k_b \int_0^{2\pi} P_k(q) \ln(P_k(q)) dq \quad (5)$$

$$S_0 = -k_b \int_0^{2\pi} P_0(q) \ln(P_0(q)) dq \quad (6)$$

This method looks at each of the dihedral angles in the system. The value of each angle k is measured over time. From those values the probability distribution $P_k(q)$ is obtained. Equation 5 is then applied to the probability distribution of that angle. The result, S_k , is then compared to a reference value, S_0 , obtained from a von Mises distribution (a Gaussian distribution for a circle). The difference of each angle's value and the von Mises value is summed across all of the angles to obtain a value for the entropy of that system according to equation 4. Then the ligand system's and the receptor system's values are subtracted from the bound complex's value to obtain the entropy change upon binding. This method is an approximation of the entropy, but the relative values of each ligand-DNA system are useful for comparison.

E. Structural Analyses

In addition to the thermodynamic metrics, several structural metrics are investigated in order to fully understand the structural changes in the ligand and the DNA upon binding, and to determine if those metrics can help explain the thermodynamic trends. These metrics include the maximum diameter of each ligand, the radius of gyration of the ligand, the radius of gyration of the core molecule, the end-to-end distance of the grafts, and finally the radius of gyration of the grafts.

1. Maximum Diameter of the Ligand

Measuring the size the ligands is a simple and effective way to quantitatively describe differences between different core and graft chemistries because different chemistries result in differently sized ligands. The maximum diameter of the ligand is defined as the distance between the two atoms in the ligand which are the farthest apart. The identity of these atoms may change throughout the course of the simulation as the ligand moves.

2. Radius of Gyration

We used two different means for measuring ligand size: the maximum diameter of the ligand and its radius of gyration. The radius of gyration of the ligand gives an average distance of the atoms in

the ligand from the center-of-geometry, not the center-of-mass. It gives insight into the overall size of the ligand. The maximum diameter of the ligand is the measure of the largest distance. In other words, the radius of gyration gives us an average size while the maximum diameter gives us the largest dimension. The radius gyration is calculated according to equation 7:

$$R_G^2 = \frac{1}{N} \sum_{i=1}^N r_i^2 \quad (7)$$

where R_G is the radius of gyration, N is the number of atoms, and r_i is the distance between the atom i and the center of geometry. The radius of gyration of the core of the ligand and the radius of gyration of the grafts of the ligand are also calculated according to equation 7. The only difference is which atoms are selected.

3. End-to-End Distance of the Grafts

The end-to-end distance is defined as the distance between the carboxyl carbon atom on the graft closest to the core and the nitrogen atom of the amine group farthest (in terms of the number of bonds) from the core as shown in Figure A16. Unlike the maximum diameter metric, the atoms used to measure the end-to-end distance do not change throughout the simulation. The relationship between the end-to-end distance and the radius of gyration of the grafts is similar to the relationship between the maximum diameter of the ligand and the radius of gyration of the ligand. The end-to-end distance gives the length of the graft whereas the radius of gyration is the average size of the graft.

F. Ligand to DNA Distance Measurements

The final category of metrics we study is the separation distance between the DNA and ligand. Intuitively it makes sense to hypothesize that the DNA and ligand would be close together if bound strongly and further apart if bound weakly. We chose three metrics to investigate: the center-of-mass

distance between the ligand and the DNA, the average distance between the nitrogen atoms in the amine groups of the grafts and the phosphorus atoms of the DNA, and the number of nitrogen atoms within four angstroms of a DNA phosphorus atom. We also investigated the average distance between the phosphorus atoms in the DNA and the nitrogen atoms in the grafts of the ligand, but the trends we saw were the same as the average distance between the nitrogen atoms in the amine groups of the grafts and the phosphorus atoms of the DNA, so the results for that metric have been placed in the Appendix.

1. Center-of-Mass Distance

The center-of-mass distance between the ligand and DNA double helix is defined as the distance between the center-of-mass of the ligand and the center-of-mass of the DNA base pair nearest to the ligand, i.e. the central axis of the DNA.

2. Average distance of Ligand Nitrogen Atoms to DNA Phosphorus Atoms

We also looked into the distance between the nitrogen atoms in the amine groups of the ligands' grafts and the phosphate groups of the DNA because it is the electrostatic interaction between the amine groups on the grafts and the phosphate groups on the DNA that are primarily responsible for the electrostatic attraction that causes DNA-ligand binding. The nitrogen-to-phosphorus distance is defined as the distance between each nitrogen atom in the grafts and the phosphorus atom nearest to that nitrogen atom.

3. Number of Ligand Nitrogen Atoms within Four Angstroms of a DNA Phosphorus atom

The nitrogen-to-phosphorus distance gives us information about all of the nitrogen atoms in the system. If we want to look at only the nitrogen atoms that are bound to the DNA, then we have to investigate the number of nitrogen atoms within a certain distance of the DNA phosphorus atoms. We arbitrarily chose 4 angstroms to be this distance. This metric is useful for comparing the binding

behavior of larger ligands to smaller ligands because larger ligands typically have more nitrogen atoms than smaller ligands and the nitrogen atoms predisposed to being farther away due to steric hindrance.

III. Results and Discussion

In general, we observe that the core-grafted ligands bind to the DNA within the first few nanoseconds of the simulation, which indicates a strong attraction between the two molecules. In the free graft systems that were at 160mM, some of the grafts do not bind to the DNA throughout the length of the 20ns simulation or they bind and unbind from the DNA during the course of the simulation. This unbinding occurs for both spermine and dilysine systems. Second, the core-grafted ligands and DNA stay bound for the duration of the simulation and the core-grafted ligands do not typically move around the DNA, but instead remain in the same location they initially bind. Third, the majority of the core-grafted ligands bind in or around the major groove of the DNA (Figure A9), and a minority of the core-grafted ligands bind near the minor groove of the DNA (Figure A10). Fourth, all of the spermine and dilysine grafts prefer to minimize the distance between their amine groups and the DNA's phosphate groups.

A. Spermine

In Figure 1a it can be seen that the core-grafted systems have a larger magnitude free energy of binding than their free graft analogs. In other words KTS and NTS bind more strongly than four free spermine grafts and NHS binds more strongly than six free spermine grafts. These trends are not statistically significant (with a 95% confidence interval), but NTS and KTS binding better than free spermine by a statistically significant amount was observed in experiment^[9]. NHS also has a greater magnitude of free energy of binding than the single ligand systems of NTS and KTS. This is because ligands with a neomycin core have an overall charge of +18, whereas NTS and KTS have a charge +12. The higher ligand charge leads to a larger electrostatic attraction and a stronger binding energy. This is why 2xNTS and 2xKTS bind with approximately twice the change in free energy of NTS and KTS respectively. If the free energy of binding is normalized to a per charge basis, then the neomycin core ligands bind with the same efficacy as the neamine and kanamycin core ligands.

We note a few issues with the free energy of binding calculations. While the electrostatic interaction between the ligand and the DNA is a large component of the molecular mechanics energy, the polar portion of the solvation energy (energy due to electrostatic interactions between the solvent and the solute) is approximately equal in size to this electrostatic interaction term, and opposite in sign. This is expected because it is known that water screens electrostatic interactions due to its large dielectric constant. The standard deviations in these two large terms, the electrostatic attraction and the polar solvation energy, are two orders of magnitude smaller than the values for those two terms; however, after adding them together the standard deviation of the free energy is the same order of magnitude as the sum of the two terms. Therefore, we focus next to see if there are statistically significant trends in molecular mechanics energy or E_{MM} , which is the sum of the bonded interactions and non-bonded interactions (no solvation energy and no

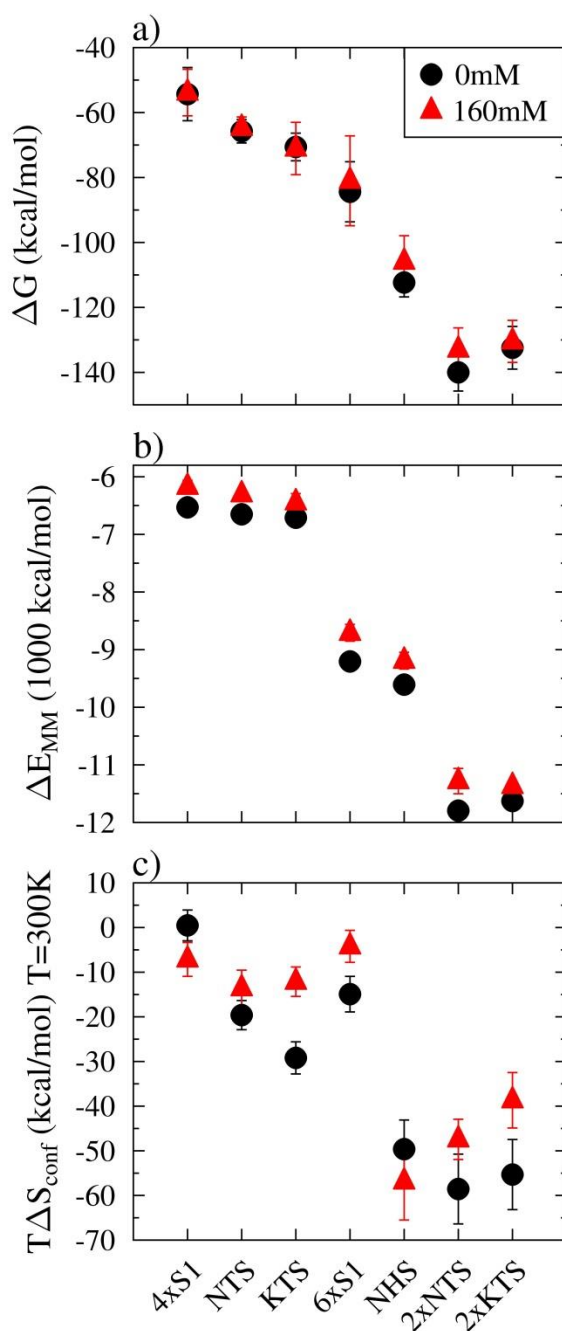


Figure 1 a) Mean free energy of binding for spermine grafted ligands and spermine free grafts. b) Mean change in molecular mechanics energy upon binding measured in Mcal/mol. c) Mean change in conformational entropy upon binding at a temperature of 300K Error bars represent standard error.

normal mode entropy), shown in Figure 1b.

From Figure 1b it is apparent that salt concentration has an impact on the electrostatic interactions with the molecular mechanics energy or E_{MM} being less in the 160mM salt concentration case than the 0mM salt concentration case. Higher salt concentrations screen the electrostatic attraction between the ligand and DNA, thus reducing the electrostatic contribution to the E_{MM} term. This screening effect has been observed previously^[28] and will also be evident in many of the trends seen in the structural data discussed later.

The dependence of the binding free energy on overall ligand charge is even more apparent in Figure 1b than in Figure 1a. The differences in E_{MM} between the charge +12 ligands, the charge +18 ligands, and the charge +24 ligands are clear. However, there does not appear to be a strong difference between the E_{MM} of the core-graft ligands NTS and KTS and the free grafts with the same ligand charge (4xS1). There is a slight difference between the NHS system and the 6xS1 system, but it is not enough to conclusively say that core-grafted ligands bind better than the free grafts. We suspect that the E_{MM} values are unable to show that the core-grafted ligands bind stronger than the free grafts because the free graft have the same overall charge as their core-grafted counterparts. Given that the difference between the core-grafted ligands and the free grafts is mostly structural, we look into the conformational entropy of the ligands since that portion of the free energy would be most affected by structural differences.

The conformation entropy losses of spermine grafted systems are presented in Figure 1c. Ligands lose conformational entropy upon binding to DNA because the grafts' motions are restricted. This loss of conformational entropy is seen in the narrowing of the distribution of probable dihedral angle values that ligand samples. From Figure 1c it can be seen that the spermine grafted ligands lose more entropy than the free spermine grafts. The free grafts are not bound as tightly to the DNA as the

core grafted ligands; occasionally one of the grafts will unbind from the DNA, so the calculated value includes some grafts that are not bound. This shifts the loss in conformational entropy towards zero.

To understand structural reasons behind the thermodynamic trends observed in Figure 1 we present in Figure 2 average size of the ligands before and after binding (Figure 2a,2b,2d and 2e) and the tightness in binding through the average distance between the ligand and DNA (Figure 2c and 2f). From Figure 2a and 2d it is apparent that as the number of atoms in the core molecule increases, the size of

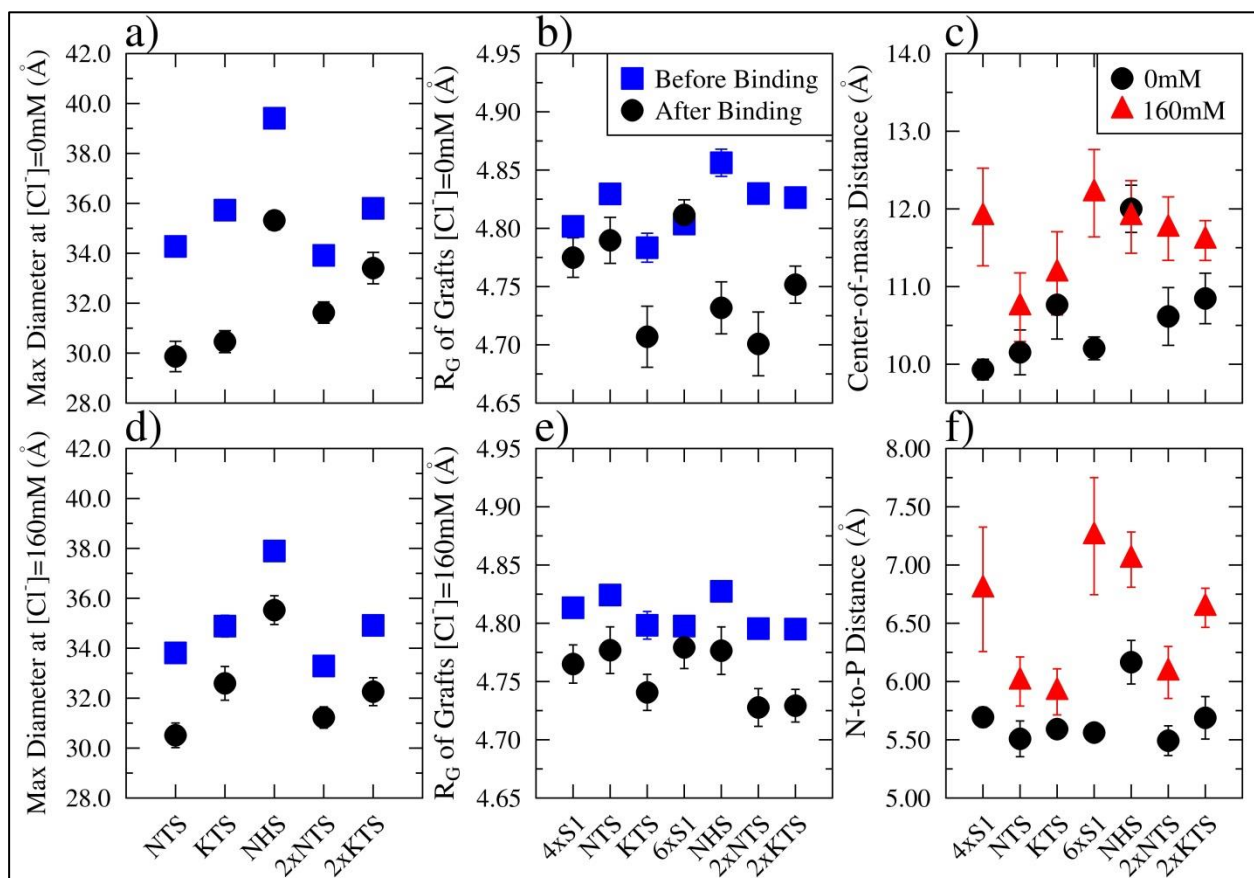


Figure 2 **a)** Average maximum diameter of spermine grafted ligands before and after binding to DNA at 0mM salt concentration. **b)** Average radius of gyration of the spermine grafts before and after binding at 0mM salt concentration. **c)** Average distance between the center-of-mass of the spermine ligands and the center-of-mass of the nearest DNA base pair after binding. **d)** Average maximum diameter of spermine grafted ligands before and after binding to DNA at 160mM salt concentration. **e)** Average radius of gyration of the spermine grafts before and after binding at 160mM salt concentration. **f)** Average distance between the nitrogen atoms of the spermine grafts and the nearest DNA phosphorus atom after binding. Error bars represent standard error.

the ligand (quantified as the larger end-end distance or maximum diameter of the ligand) as a whole increases. Additionally, the size of the whole ligand decreases upon binding in all cases (Figure 2a and 2d). The decrease in ligand size upon binding is due to the behavior of the spermine grafts, rather than the core. From Figure A19 and Figure A20 it can be seen that the size of the core of the ligand does not change upon binding for any system. Figure 2b and Figure 2e show that on average the spermine grafts decrease in size, evident in the radius of gyration decreasing by about 0.5-1.0 angstrom. Being a linear molecule spermine grafts bend and fold a certain extent upon binding to DNA. As a spermine grafted ligand moves towards the DNA, the grafts fold back on themselves or wrap around each other in order to move the whole ligand closer to the DNA (Figure A10). Such structural changes in the spermine grafts lead to overall decrease in the ligand size upon binding. There is a notable change upon increasing the number of spermine grafted molecules. The size of the ligands does not decrease as much upon binding when there are two ligands in the system. This is because in the two ligand case the spermine grafts from the two molecules hinder one another (sterically and electrostatically due to similar positive charges) in such a way as to cause the grafts on both ligands to spread out more on the DNA relative to the single ligand case. This effect is small and only apparent at 0mM salt concentration because higher salt concentrations screen the electrostatic repulsions.

Next, we compare the aminoglycoside grafted spermines to free spermines and find that the free spermine grafts do not change in size as much as the grafted spermine especially for the six graft systems. This indicates that the free grafts are behaving much like they would behave if unbound. On the other hand, grafting spermine to a core molecule causes some restrictions on the bound state of the graft. Additionally, Figure A17 and Figure A18 show that the spermine free grafts are consistently shorter than the grafts on the aminoglycoside cores. This could be explained by crowding of the grafts on the core-grafted ligands. As the grafts become crowded, they stretch out and increase their average

end-to-end distance. It would also explain why the before binding end-to-end distance of the grafts attached to NHS are longer than the before binding distances of NTS and KTS. NHS is the most crowded with six grafts, so its grafts experience the most stretching.

To understand how tightly or loosely the ligands bind to the DNA we plot the distance between the center-of-mass of the ligand and the center of mass of the nearest DNA base pair for varying ligands and varying salt concentration in Figure 2c. As mentioned previously, the salt ions screen the electrostatic attraction between the DNA and the ligand which is responsible for most of the attraction between the two molecules. Since the electrostatic attraction between the ligand and DNA decreases, the distance between the two increases. However, the NHS does not get further away from the DNA as salt concentration increases. This might be because it is already further away from the DNA than NTS and KTS at 0mM salt, so the effect of charge screening on center-of-mass distance is not as significant as it is with NTS and KTS. NHS binds further away than NTS and KTS because of its increased size. NHS is simply too big to get as close as the other core-grafted systems, i.e. it is unable to bind in the major groove of the DNA as effectively.

From Figure 2c, it can be plainly seen that the free graft systems were closer than or as close to the DNA as the core-grafted ligands at 0mM salt concentration. The trend occurs because the center-of-mass is measured from each of the grafts individually in the free grafts systems, while in the core-grafted systems it is measured to the center-of-mass of the ligand, so the shorter distance for the free grafts is an artifact of how the center-of mass distance is measured. However, when the electrostatic attraction between the free grafts and the DNA is screened by either increased salt concentration or additional grafts, some of the grafts do not stay bound to the DNA throughout the whole simulation in the free graft systems. This unbinding contributes to a large standard error and dramatically increased center-of-mass distances. The core-graft ligands do not unbind; instead, if one of the grafts unbind, the

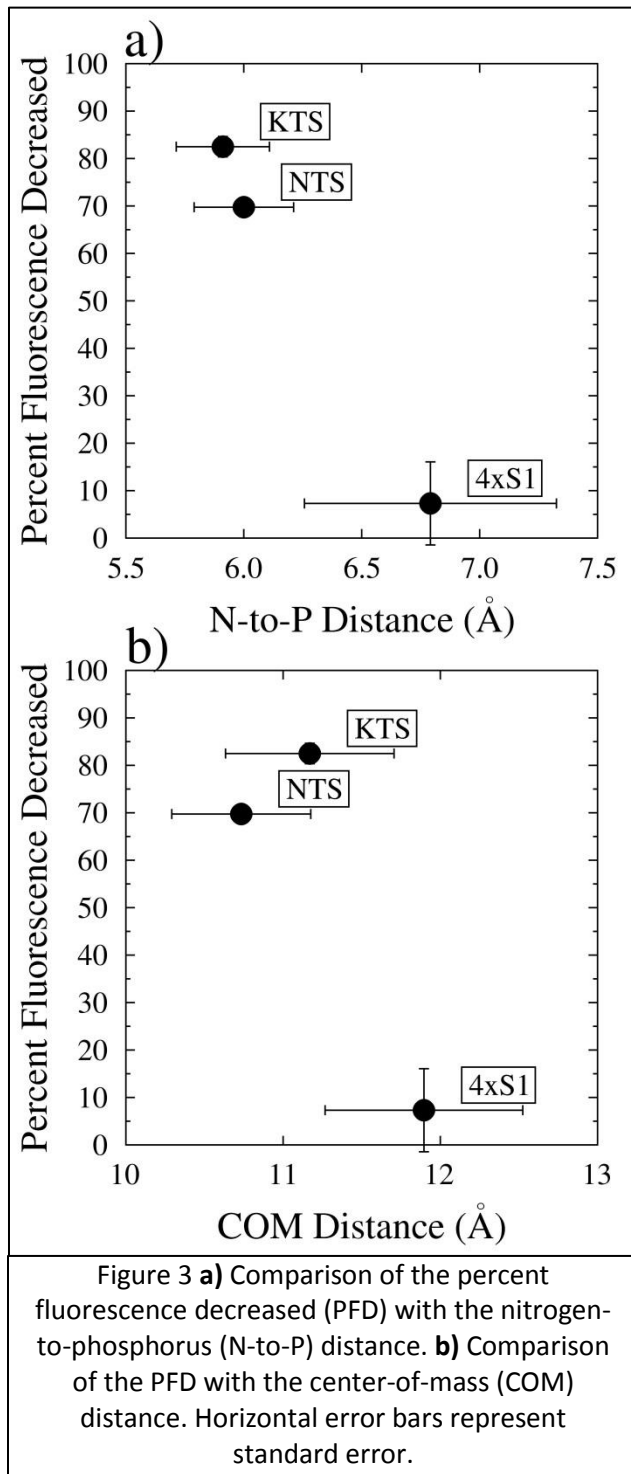
other grafts are still attached to the DNA. The attached grafts hold the unattached graft within close proximity to the DNA until it binds again. This phenomenon has been observed visually. In essence, all of the grafts on the core-grafted ligands would have to become unbound in order for the ligand to unbind from the DNA. This is further evidence that the core-grafted ligands bind more tightly than the free grafts.

It is intuitive that the closer the nitrogen atoms of a ligand are to the phosphorus atoms of the DNA the tighter the binding of a ligand because it is the amine groups of the ligands and the phosphate groups of the DNA that are responsible for the large electrostatic attraction between the two. It is also expected that the closer the amine groups are to the phosphate groups, the closer the center-of-mass of the ligand to the DNA's central axis. Therefore it is unsurprising that the distances between the nitrogen atoms on the grafts and the phosphorus atoms on the DNA closely mirror the trends seen in the center-of-mass data. For example, increasing salt concentration increases the average distance between the two elements, which can be seen in Figure 2f, and NHS is still further away than the NTS and KTS.

The nitrogen-to-phosphorus distance of the free spermine grafts is similar to the nitrogen-to-phosphorus distance of the core-grafted systems with four spermine grafts at 0mM salt concentration, but the spermine core grafted ligands have a shorter nitrogen-to-phosphorus distance than the free grafts at 160mM salt concentration. The core-grafted ligands are able to place their amine groups near the DNA phosphate groups just as well as the free graft systems, but when the salt concentration is increased the core-grafted architectures bind more robustly to the DNA strand than the free grafts which keeps them closer to the DNA relative to the free grafts.

B. Comparison of results so far with spermine-DNA binding results from experiment:

From the thermodynamic data in Figure 1 we saw that the core-grafted ligands bound stronger to DNA than the free spermine grafts, however for the free energy of binding the differences between the ligands were not statistically significant. The structural data presented in Figure 2 had statistically significant trends. During binding the spermine grafts on the core-grafted ligands contort and shrink. Visual evidence showed us that the spermine grafts fold around the DNA, or stretch out along the backbone or major groove. The change in graft shape takes place in order to put the amine groups of the grafts at an optimal distance from the DNA phosphate groups. Therefore, the binding metric used in an experiment should have some correlation to the nitrogen-to-phosphorus distance. The experimental results we use for comparison are from Rege et al.^[9] A



higher percentage fluorescence decreased (PFD) indicates stronger DNA binding. Therefore, if there is a good correlation between the experimental binding and our structural characterization, as the PFD decreases the nitrogen-to-phosphorus distance should increase. That is what we see in Figure 3a. The

average nitrogen to phosphorus distances of KTS and NTS are not statistically different by a meaningful amount. However, the 4xS1 system's nitrogen to phosphorus distance is statistically significant from the NTS system's distance with an alpha of 16.84%. This is in agreement with the conclusions made by Rege et al. We note that the correlation was not as strong with the center of mass distance because the center-of-mass distance is affected by ligand size more than the nitrogen-to-phosphorus distance.

It is important to note that the simulation conditions do not match the experimental conditions exactly. For example, the salt and ligand concentrations do not match exactly and the DNA strand length used in experiment is significantly longer than the length used in the simulations. By studying single to two molecules, we are in a regime where the ligand-ligand interactions are minimal. Despite these differences it is apparent that the higher the PFD the closer the ligand is to the DNA and that we can use these structural trends to predict binding efficacies for dilysine based ligands presented next.

C. Dilysine

We wanted to compare our results with spermine grafts to another chemistry to better understand how the architecture of spermine affects its DNA binding efficacy. Dilysine grafts were chosen for several reasons. First, lysine is a well understood amino acid. Second, dilysine has a branched architecture; whereas, spermine has a linear architecture. Third, both dilysine and spermine, once grafted, have a charge of +3, and the charges are localized on their amine groups. Due to their similarity in charge there are some similarities between the trends seen in free energy of the spermine grafted ligands and the dilysine grafted ligands. For example the core-grafted architectures, NTL2, KTL2, and NHL2, bind more strongly than the free dilysine grafts as can be seen in Figure 4a. Similar to the spermine systems, a larger overall ligand charge leads to a greater magnitude change in the free energy of binding, but the similarities in the free energy of spermine and dilysine ligands ends there.

Comparing **Error! Not a valid bookmark self-reference.a** to Figure 4a shows that the spermine grafted ligands have a larger magnitude binding free energy than the dilysine grafted ligands for a constant core chemistry, which means that the spermine grafted ligands bind stronger than the dilysine grafted ligands. This is easier to observe in the double ligand systems (2x) than the single ligand systems. The double ligand systems have a larger magnitude binding free energy than the single ligand systems because the ligand concentration is higher. The subtle differences between graft chemistries in the single ligand case were virtually unnoticeable, but the increased ligand concentration revealed them. Even in the double ligand case, the difference between spermine and dilysine in the free energy of binding is still not statistically significant with a 95% confidence interval. However the conformational entropy trends further support the claim that spermine binds stronger than dilysine.

When comparing Figure1c to Figure 4c it

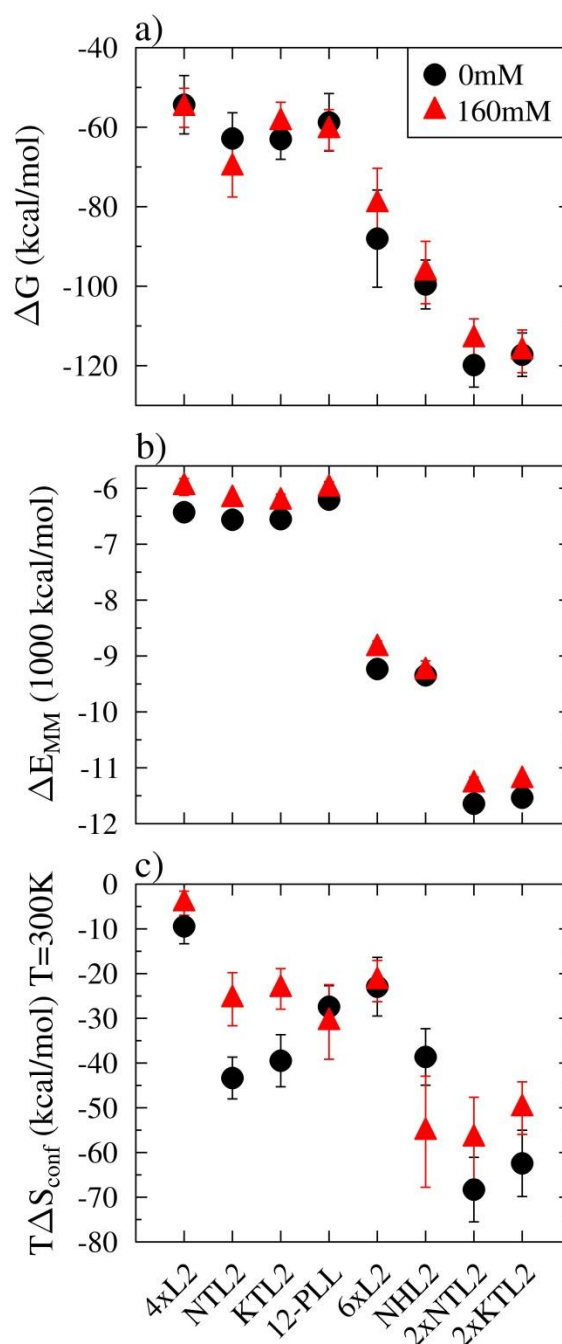


Figure 4 **a)** Mean free energy of binding for dilysine grafted ligands and dilysine free grafts. **b)** Mean change in molecular mechanics energy upon binding measured in Mcal/mol. **c)** Mean change in conformational entropy upon binding at a temperature of 300K Error bars represent standard error.

is apparent from the core-grafted architectures that dilysine grafted ligands lose more dihedral entropy than spermine grafted ligands upon binding. This is true for both the single ligand and double ligand cases; this effect is less pronounced for double ligand systems. The dilysine-grafted ligands must bind with DNA in a way that restricts their movement, specifically the movement of the grafts, more than the spermine grafted ligands. This further supports the conclusion that spermine ligands bind stronger than dilysine grafted ligands.

The data for the neomycin core ligands does not support this trend. Instead, the data for NHL2 is

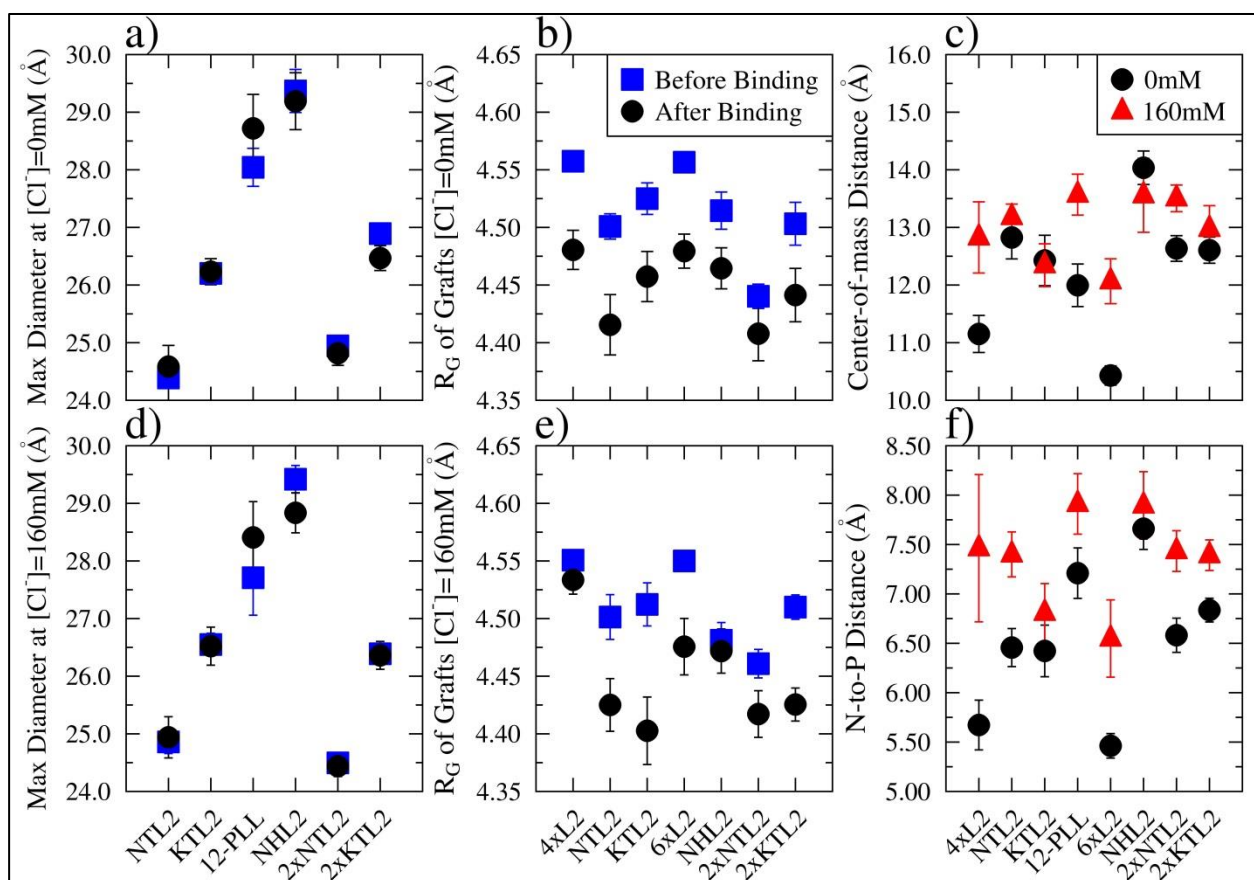


Figure 5 **a)** Average maximum diameter of spermine grafted ligands before and after binding to DNA at 0mM salt concentration. **b)** Average radius of gyration of the spermine grafts before and after binding at 0mM salt concentration. **c)** Average distance between the center-of-mass of the spermine ligands and the center-of-mass of the nearest DNA base pair after binding. **d)** Average maximum diameter of spermine grafted ligands before and after binding to DNA at 160mM salt concentration. **e)** Average radius of gyration of the spermine grafts before and after binding at 160mM salt concentration. **f)** Average distance between the nitrogen atoms of the spermine grafts and the nearest DNA phosphorus atom after binding. Error bars represent standard error.

statistically indistinguishable from the data for NHS. Both ligands have six grafts instead of four, which may be causing the discrepancy, but it is unclear how. Our current theory is that the increased number of grafts increases the crowding of the grafts. This prevents some of the grafts from interacting with the DNA strand, so they behave as if unbound. This behavior increases the variability in the entropy values. The variability in the conformational entropy and free energy motivate comparisons between the structure of spermine and dilysine.

For a given core chemistry, the spermine grafted ligands (Figure 2a & 2d) are larger than the dilysine (Figure 5a & 5d) grafted ligands because spermine is a long chain of nitrogen atoms and carbon atoms, whereas dilysine has two distinct branches that make the grafts shorter (Figure A17 and Figure A18) and bulkier than spermine. The bulkiness of dilysine explains why the maximum diameters of the dilysine grafted ligands do not shrink upon binding while the spermine grafted ligands do shrink. Spermine is able to fold back on itself, and the dilysine grafts do not have the same flexibility as the spermine grafts, so they cannot distort as much as the dilysine grafted ligands. The shorter dilysine grafts lead to different ligand-ligand interactions as well. The NTL2 and KTL2 do not exhibit a change in size between the double and single ligand cases the way NTS and KTS do because of their smaller size, i.e. negligible steric hindrance. The grafts are simply too short to interact with grafts from the other ligand in the double ligand case. Though ligand-ligand interactions decrease because of dilysine's shorter architecture, the intramolecular interactions increase due to crowding.

Since the dilysine grafts are bulkier, it would be expected that the crowding effect be more substantial for them. This can be seen in Figure A17 and Figure A18. The difference in end-to-end distance between the dilysine grafted ligands and the free dilysine grafts is much bigger than the corresponding difference in the spermine systems. NHL2, being the most crowded, exhibits the largest deviation away from the free graft average length. One final observation of the end-to-end distance of

dilysine shows that the end-to-end distance of the dilysine grafts did not change significantly on binding, but according to Figure 5b&e the radius of gyration did. This difference can be explained as follows: when the dilysine binds to the DNA, it can sometimes “pinch”; in other words the two branches move closer together. This happens for one of two reasons: either the dilysine fits into one of the grooves of the DNA (Figure 6b), or it aligns its two amine groups with the phosphate groups on the backbone of the DNA (Figure 6c). This lowers the radius of gyration of the graft without changing the end-to-end distance. Spermine and dilysine undergo different changes in their shape upon binding to the DNA so it is expected that the changes in the grafts would change the overall shape of the ligand.

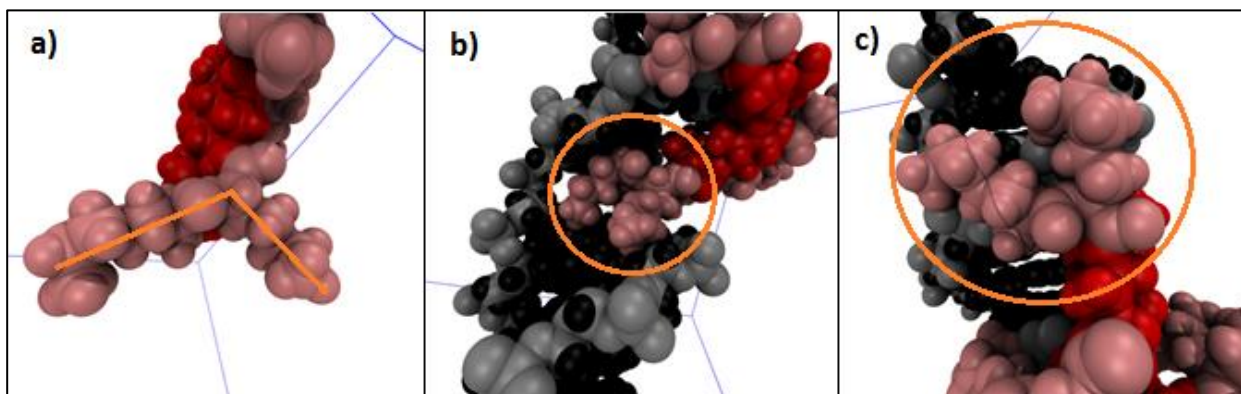


Figure 6 **a)** Unbound, splayed dilysine graft **b)** Bound dilysine graft exhibiting pinching in the major groove of the DNA **c)** Bound dilysine graft exhibiting pinching along the backbone (grey) of the DNA

We can obtain some insight into the overall shape of the ligand by comparing the radius of gyration of the ligand to the maximum diameter of the ligand. The trends in the radius of gyration of the ligand (Figure A14 and Figure A15) are exceedingly similar to the trends seen in the maximum diameter. If all of the trends seen in the maximum diameter match the trends seen in the radius of gyration, then all of the ligands must be roughly the same shape. An example is helpful in explaining why this is; if ligand A had a greater maximum diameter value than ligand B, and ligand A had the same radius of gyration as ligand B, then that would mean that on average they were the same size, but ligand A would be more elongated than ligand B in at least one direction.

It is understandable that the core-grafted architectures are similarly shaped because they are all structurally similar, but it is surprising that the 12-PLL ligand is also shaped like the core-graft ligands. 12-PLL is 12 lysine amino acids attached together serially. Therefore, it is a long, branched chain. 12-PLL's radius of gyration and maximum diameter both increase upon binding at a salt concentration of 0mM. This increase can be explained by viewing the trajectories of the simulation. The 12-PLL goes from a globular state when unbound (Figure A11) to a stretched out state when bound to the DNA (Figure A13). It lengthens itself along either the major or minor groove of the DNA, but it maintains its globular shape due to the curvature of the grooves.

12-PLL also has a large difference between the center-of-mass distance at 0mM salt concentration and the distance at 160mM as can be seen in Figure 6c. This is because 12-PLL is not bound as well at 160mM salt concentration as at 0mM salt concentration. Visualization of the trajectories at 160mM salt concentration shows 12-PLL peeling away from the DNA as if it was only partly bound (Figure A12) in many of the trials. This would suggest that 12-PLL does not bind as well as the core-grafted ligands at higher salt concentrations and that the architecture of a grafted ligand is superior to stringing several grafts together in a linear fashion.

Graft architecture definitely plays an important role as well. For a given core chemistry the dilysine grafted ligands are farther away compared to their spermine grafted analogs. This trend must be due to the tighter binding of spermine compared to dilysine, since it is known that larger ligands tend to be further from the DNA than smaller ones, and dilysine grafted ligands are smaller than spermine grafted ligands. From visual evidence it is known that the spermine grafts are flexible enough and long enough that they can bend around the core of the ligand to reach the DNA. If they are close to the DNA, then they can stretch along the backbone or a groove and bring the center-of-mass of the ligand closer

to the DNA. This implies that the reason the spermine grafted ligands bind tighter than the dilysine grafted ligands is because they are more flexible.

Flexibility is critical in allowing the grafts to optimize their amine groups relative to the phosphate groups of the DNA. Unsurprisingly, the strongest evidence that spermine binds better than dilysine is the comparison of the nitrogen-to-phosphorus distance. The nitrogen atoms of the dilysine free graft systems are still closer to the DNA than their core-grafted counterparts (Figure 5f) at 0mM. This would indicate that the core grafted architecture is not able to place the amine groups as close as the free grafts for the dilysine chemistry. At 160mM the free grafts and the core-grafted architectures are about the same distance. This indicates that the core-grafted architectures are binding more robustly than the free grafts, and the effect is enough to make up for the fact that the dilysine grafted systems are not able to place their amine groups as close as the free grafts. The fact that the spermine grafted ligands are able to optimize their nitrogen-to-phosphorus distance and the dilysine grafted ligands are not is further evidence that the spermine grafted ligands have a greater binding affinity for the DNA than the dilysine grafted ligands due to their flexibility.

IV. Conclusion

The ligands studied in this paper bind strongly to the DNA due to strong electrostatic attraction. The main contributor to this electrostatic attraction is the positively charged amine groups on each graft of the ligand binding to the negatively charged phosphates along the backbone of the DNA. We see this visually in the trajectories and describe it quantitatively with the nitrogen-to-phosphorus distance and number of nitrogen atoms within four angstroms of the DNA phosphorus atoms.

The core-grafted ligands have approximately the same electrostatic attraction to the DNA as the free grafts, but the core-grafted ligands bind with higher affinity than the free grafts. This firstly is indicated in the thermodynamic data from MMPBSA. The reason for this is elucidated from the center-

of-mass data. The center-of-mass data indicates that as the electrostatic attraction becomes screened the free grafts begin to unbind from the DNA, but the core-grafted ligands remain bound because it is more difficult to remove all four or six grafts at once than one graft at a time. However, there is a trade-off. Grafting compounds to aminoglycoside cores increases the crowding of the grafts and increases the steric hindrance between the grafts as well. The neomycin core ligands were consistently farther away from the DNA and had a higher nitrogen-to-phosphorus distance, which indicated that the grafts were being hindered while trying to bind to the DNA.

Grafting polyamines or amino acids improve the binding behavior of the graft, but certain grafts bind DNA better than others. It is clear from this work that spermine grafted ligands bind more strongly to DNA than dilysine grafted ligands. The thermodynamic data supports this. For a constant graft chemistry, spermine grafted ligands have a greater magnitude binding free energy and a smaller conformational entropy loss. But the thermodynamic data does not show statistically significant results. However, the center of mass data and nitrogen to phosphate data do show statistically significant results. The spermine grafted ligands bind closer to the DNA and have a shorter nitrogen-to-phosphorus distance than dilysine. Both of which indicate tighter binding.

The stronger binding of the spermine grafted ligands compared to the dilysine grafted ligands can be explained from the different architectures of the two chemistries. The spermine grafts are linear in architecture this allows them to be more flexible than the dilysine grafted ligands. Upon binding the ligands shrink both in terms of radius of gyration of the ligand and maximum diameter. The grafts end-to-end distance shrinks, and their radius of gyration decreases. From visualizing the trajectories it can be seen that they stretch out along the backbone or grooves of the DNA and can match their nitrogen-to-phosphorus distance with the spermine free grafts nitrogen-to-phosphorus distance.

The dilysine grafts are branched and shorter than the spermine grafts. The dilysine grafted ligands do not shrink upon binding according to the radius of gyration data and the maximum diameter data. The end-to-end distance of the dilysine grafts does not change upon binding, but the radius of gyration does, which indicates that the grafts exhibit “pinching” movements. The grafts are not static, but are not nearly as flexible as the spermine grafts. Furthermore, the grafts are more crowded due to their bulky architecture. This impedes binding by preventing the ligands from getting as close to the DNA as the spermine grafted ligands. The grafts farther away from the DNA cannot bind and this increases the nitrogen-to phosphate distance. The crowding is so bad on NHL2 that the number of nitrogen atoms within four angstroms of the DNA phosphorus atoms matches that for KTL2 and NTL2, which indicates that only four grafts were able to bind to the DNA.

From this data some general conclusions can be made. First, grafting polyamines or other positively charged compounds to a core improves the binding efficacy of the grafts. Second, long, flexible, non-branched grafts are better at binding than short branched grafts. This information should help guide experimentalists in synthesizing new core-grafted molecules for DNA binding

V. References

- [1] C. Sun, T. Tang, and H. Uludağ, "Molecular dynamics simulations of PEI mediated DNA aggregation.," *Biomacromolecules*, vol. 12, no. 10, pp. 3698–707, Oct. 2011.
- [2] R. M. Elder, T. Emrick, and A. Jayaraman, "Understanding the effect of polylysine architecture on DNA binding using molecular dynamics simulations.," *Biomacromolecules*, vol. 12, no. 11, pp. 3870–9, Nov. 2011.
- [3] D. J. Glover, H. J. Lipps, and D. A. Jans, "Towards safe, non-viral therapeutic gene expression in humans.," *Nature reviews. Genetics*, vol. 6, no. 4, pp. 299–310, Apr. 2005.
- [4] M. J. Stevens, "Simple simulations of DNA condensation.," *Biophysical journal*, vol. 80, no. 1, pp. 130–9, Jan. 2001.
- [5] a C. Hunter, "Molecular hurdles in polyfectin design and mechanistic background to polycation induced cytotoxicity.," *Advanced drug delivery reviews*, vol. 58, no. 14, pp. 1523–31, Dec. 2006.
- [6] S. M. Moghimi, P. Symonds, J. C. Murray, a C. Hunter, G. Debska, and A. Szewczyk, "A two-stage poly(ethylenimine)-mediated cytotoxicity: implications for gene transfer/therapy.," *Molecular therapy : the journal of the American Society of Gene Therapy*, vol. 11, no. 6, pp. 990–5, Jun. 2005.
- [7] J. Ziebarth and Y. Wang, "Molecular dynamics simulations of DNA-polycation complex formation.," *Biophysical journal*, vol. 97, no. 7, pp. 1971–83, Oct. 2009.
- [8] B. Sung, "Condensation of DNA by spermine in the bulk and in the bacteriophage capsid: a cryo-electron microscopy study," 2011.
- [9] K. Rege, A. Ladiwala, S. Hu, C. M. Breneman, J. S. Dordick, and S. M. Cramer, "Investigation of DNA-binding properties of an aminoglycoside-polyamine library using quantitative structure-activity relationship (QSAR) models.," *Journal of chemical information and modeling*, vol. 45, no. 6, pp. 1854–63, 2005.
- [10] O. Boussif, F. Lezoualc'h, M. A. Zanta, M. D. Mergny, D. Scherman, B. Demeneix, and J. P. Behr, "A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine.," *Proceedings of the National Academy of Sciences*, vol. 92, no. 16, pp. 7297–7301, Aug. 1995.
- [11] S. Weiner, P. Kollman, and D. Case, "A new force field for molecular mechanical simulation of nucleic acids and proteins," *Journal of the American Chemical Society*, vol. 106, no. 3, pp. 765–784, 1984.
- [12] B. G. Feuerstein, N. Pattabiraman, and L. J. Marton, "Spermine-DNA interactions: a theoretical study.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 16, pp. 5948–52, Aug. 1986.

- [13] L. W. Tari and A. S. Secco, "Base-pair opening and spermine binding—B-DNA features displayed in the crystal structure of a gal operon fragment: implications for protein-DNA recognition," *Nucleic Acids Research*, vol. 23, no. 11, pp. 2065–2073, 1995.
- [14] H. Deng, "Structural basis of polyamine-DNA recognition: spermidine and spermine interactions with genomic B-DNAs of different GC content probed by Raman spectroscopy," *Nucleic Acids Research*, vol. 28, no. 17, pp. 3379–3385, Sep. 2000.
- [15] K. Rege, S. Hu, J. a Moore, J. S. Dordick, and S. M. Cramer, "Chemoenzymatic synthesis and high-throughput screening of an aminoglycoside-polyamine library: identification of high-affinity displacers and DNA-binding ligands.," *Journal of the American Chemical Society*, vol. 126, no. 39, pp. 12306–15, Oct. 2004.
- [16] A. Pedretti, L. Villa, and G. Vistoli, "VEGA - An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming," *Journal of Computer-Aided Molecular Design*, vol. 18, pp. 167–163, 2004.
- [17] J. Wang, R. M. Wolf, J. W. Caldwell, P. a Kollman, and D. a Case, "Development and testing of a general amber force field.," *Journal of computational chemistry*, vol. 25, no. 9, pp. 1157–74, Jul. 2004.
- [18] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. a Loughton, and M. Orozco, "Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers.," *Biophysical journal*, vol. 92, no. 11, pp. 3817–29, Jun. 2007.
- [19] D. Pearlman, D. Case, and J. Caldwell, "AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural," *Computer Physics ...*, vol. 91, no. 1–3, pp. 1–41, Sep. 1995.
- [20] 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.
- [21] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems," *The Journal of Chemical Physics*, vol. 98, no. 12, pp. 10089–10092, 1993.
- [22] H. Gohlke and D. A. Case, "Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf.," *Journal of computational chemistry*, vol. 25, no. 2, pp. 238–50, Jan. 2004.
- [23] N. Homeyer and H. Gohlke, "Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method," *Molecular Informatics*, vol. 31, no. 2, pp. 114–122, Feb. 2012.
- [24] S. Genheden and U. Ryde, "Comparison of end-point continuum-solvation methods for the calculation of protein-ligand binding free energies.," *Proteins*, pp. 1–17, Jan. 2012.

- [25] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, "MMPBSA.py : An Efficient Program for End-State Free Energy Calculations," *Journal of Chemical Theory and Computation*, vol. 8, no. 9, pp. 3314–3321, Sep. 2012.
- [26] D. a. Case, "Normal mode analysis of protein dynamics," *Current Opinion in Structural Biology*, vol. 4, no. 2, pp. 285–290, Jan. 1994.
- [27] D.-W. Li and R. Brüschweiler, "In silico Relationship between Configurational Entropy and Soft Degrees of Freedom in Proteins and Peptides," *Physical Review Letters*, vol. 102, no. 11, pp. 20–23, Mar. 2009.
- [28] Z. Ou and M. Muthukumar, "Entropy and enthalpy of polyelectrolyte complexation: Langevin dynamics simulations.," *The Journal of chemical physics*, vol. 124, no. 15, p. 154902, Apr. 2006.
- [29] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. San Diego: ACADEMIC PRESS LTD, 2002.

VI. Appendix

For Molecular Simulations Studies of the Effect of Ligand Architecture on DNA Binding

A. Protocol for Using Antechamber to Derive Ligand Parameters

Antechamber is a program provided with the AmberTools software package^[19]. It assigns atom types and parameters based on an atom's environment and element. The charges are assigned either using SQM, an AMBER program, or MOPAC. We used MOPAC to calculate charges with the keywords AM1, MMOK, PULAY, ITRY=800, and PRECISE. AM1 means we used the am1-bcc charge method. MMOK means we used the molecular mechanics correction to CONH bonds of 14 kcal/mol (used for peptide linkages). PULAY is the converger that was used. Our maximum number of iterations was 800 and was set with the ITRY keyword. The PRECISE keyword increases the criteria for electronic and geometric optimizations by a factor of 100. If a molecule was too large to allow the calculations to converge in a reasonable amount of time, then the PRECISE keyword was removed.

B. Determination of Equilibrium

In molecular dynamics simulations equilibrium is determined by measuring the system properties. Equilibrium is reached when the system properties are no longer fluctuating. However, if the system becomes kinetically trapped, then this is considered "equilibrium." This problem is endemic to all simulations with small timescales^[29].

In order to determine equilibrium in our systems the potential energy of the system was plotted vs. time. Then the system was split into two blocks. The average was calculated for each block. If the difference between the two averages was no greater than 2.5 % of the all the data considered, then the system was deemed to be at equilibrium. If the difference was greater than 2.5 %, then the system was split into four blocks and the average of the last two were compared. If their difference was less than 2.5%, then the average of the second block was compared to the average of the last two. If the second

block was rejected then the process ended and the last two blocks were assumed to be at equilibrium. If equilibrium was not achieved in the last two blocks then the system was split into 8 blocks and the process repeated until the system found when equilibrium occurred or the block size became smaller than 100 time steps.

All of the potential energy plots were analyzed visually in order to catch errors the automated process missed. For example, if the system reached equilibrium right at the beginning of the simulation, then the potential energy plot might drop suddenly and not affect the average of the first block significantly. If this occurred, then the part of the trajectory was eliminated from consideration and the automated process was rerun.

C. Supplementary Figures and Movies

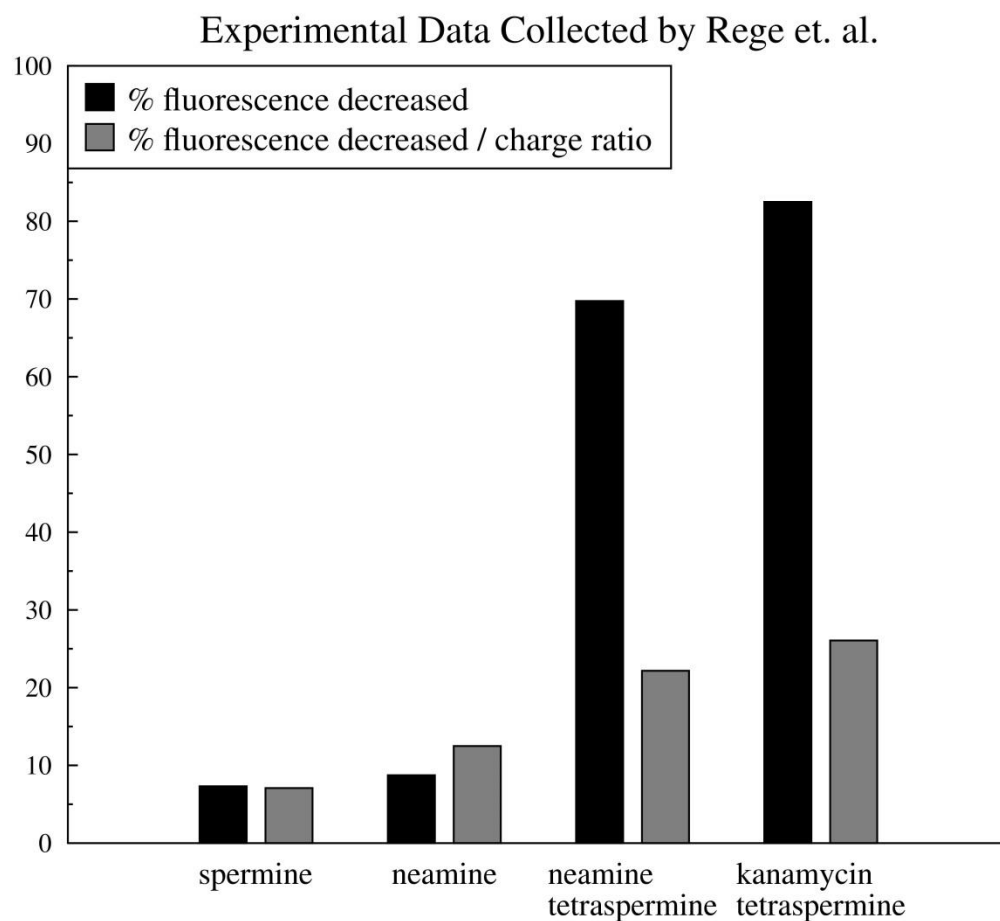


Figure A7: Abridged Data from Rege et al. The original data can be found in reference^[9]. A higher percentage fluorescence decreased implies a better binding affinity with DNA. The charge ratio referenced here is the charge ratio between the polycation and the DNA.

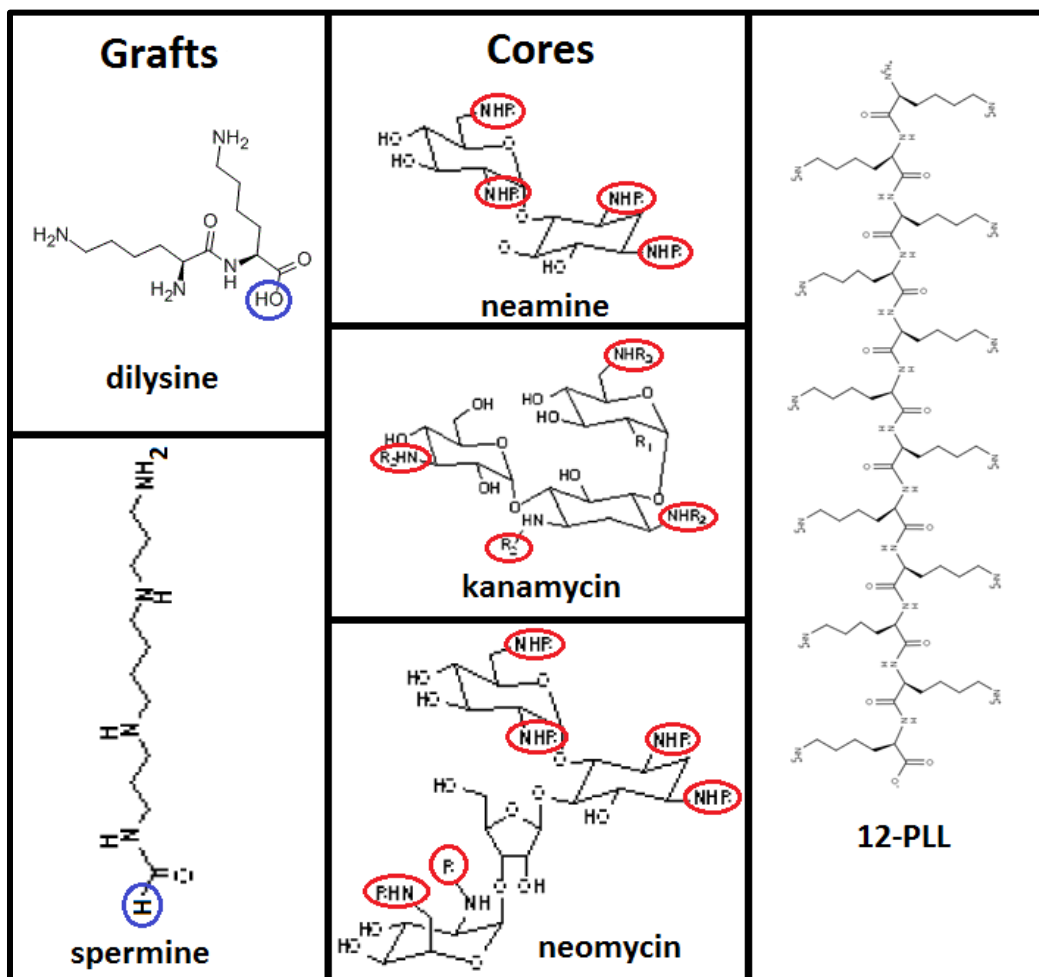


Figure A8: Structures of grafts and cores used in this work. Grafts (left) have their grafting point circled in blue. The hydrogen in spermine is removed when it is attached to the aminoglycoside core. The hydroxyl group on dilysine is removed when it is attached to the aminoglycoside core. The cores (middle) have their grafting points circled in red. The R represents the graft. 12-PLL is shown on the right. The diagram of dilysine comes from http://www.chemicalbook.com/ChemicalProductProperty_EN_CB2324131.htm. The diagram of 12-PLL was provided by Robert Elder. All other diagrams were obtained from^[15].

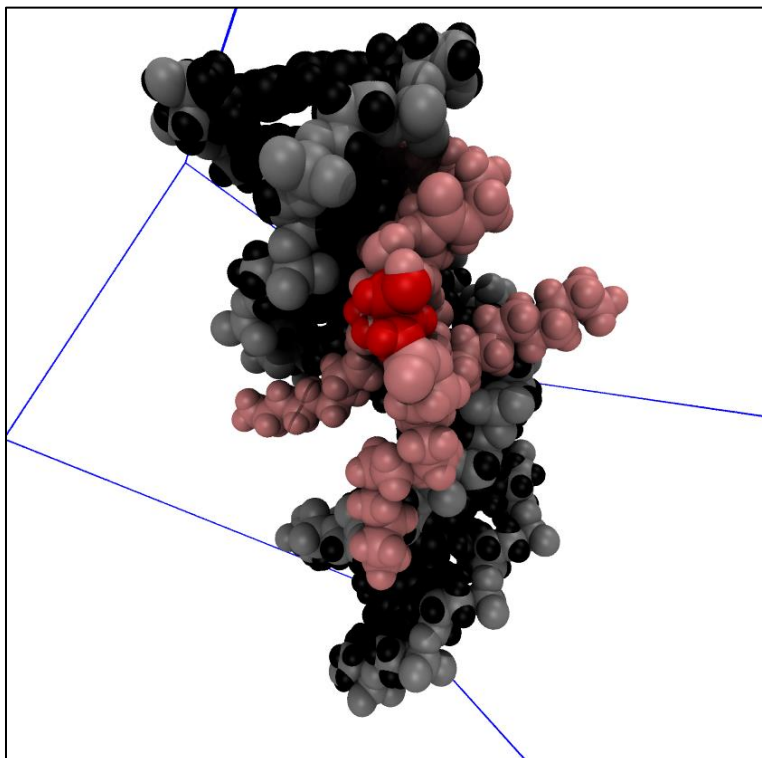


Figure A9: Example of a ligand binding to the major groove of the DNA double strand. The ligand shown here is neamine tetraspermine (NTS).

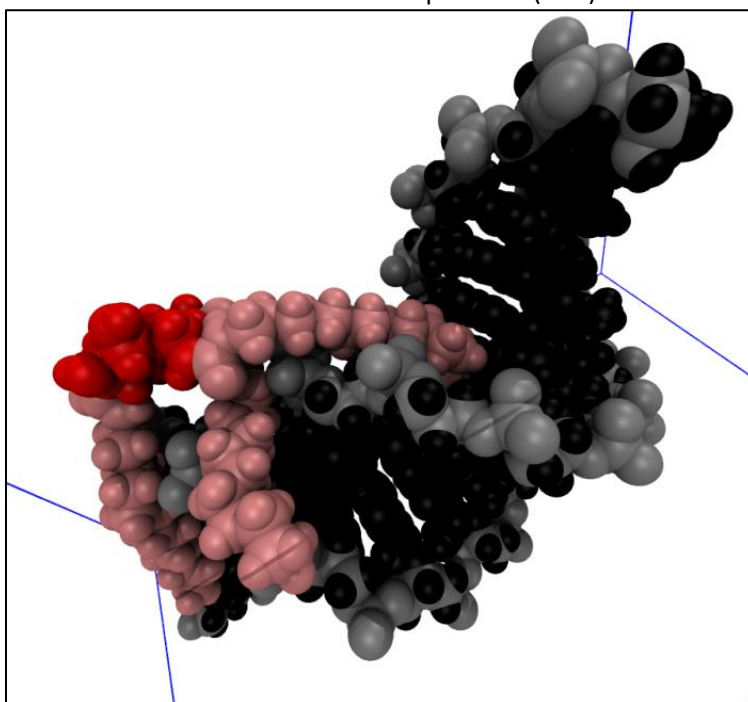


Figure A10: Example of a ligand binding to the minor groove of the DNA double strand. The ligand shown here is neamine tetraspermine (NTS). The middle graft is folding back on its grafting point to reach the DNA backbone.

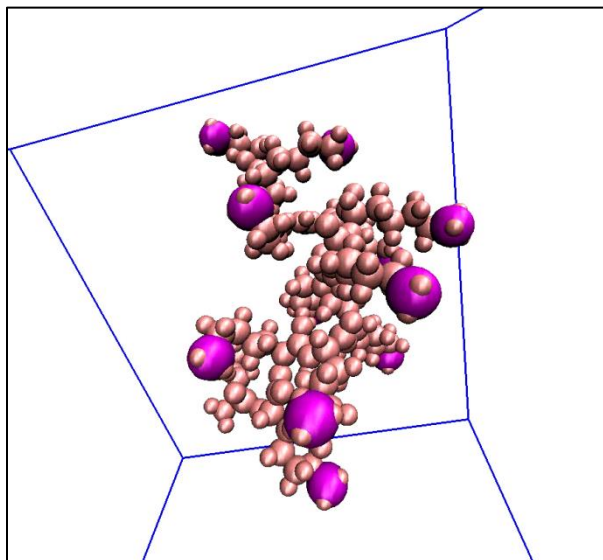


Figure A11: Unbound 12-PLL with the nitrogen atoms of the side chain amine groups labeled in magenta. Note the disorganized nature of the molecule.

Figure A13: 12-PLL bound to DNA at 0mM salt concentration. The nitrogen atoms of the side chain amine groups are labeled in magenta. Note how the amine groups line up along the backbone of the DNA strand.

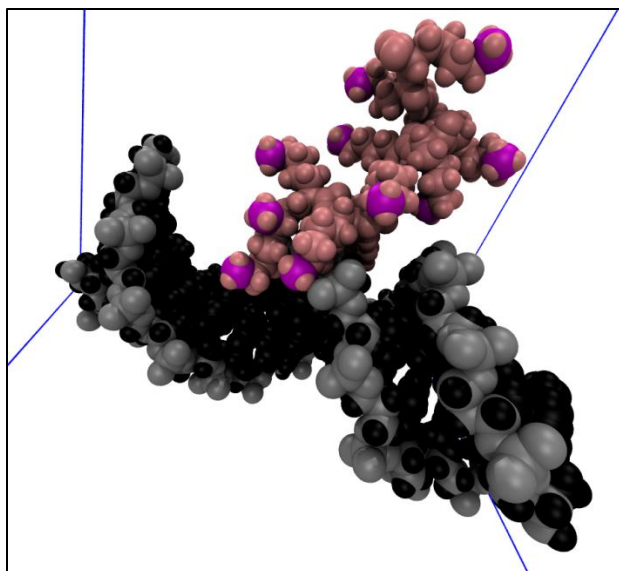
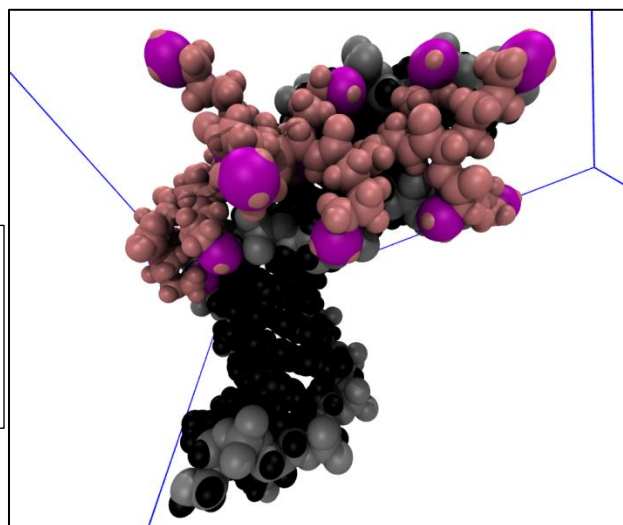


Figure A12: 12-PLL "bound" to DNA at 160mM salt concentration. The nitrogen atoms of the side chain amine groups are labeled in magenta. Note how over half of the 12-PLL is unbound from the DNA strand. This suggests the 12-PLL does not bind to DNA very well at high salt concentrations.

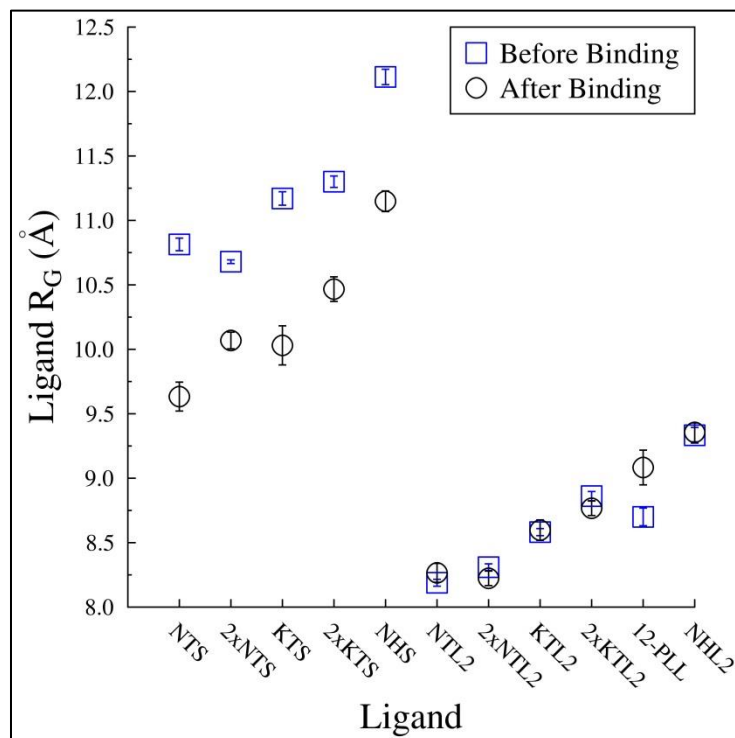


Figure A14: Average radius of gyration of the ligands before and after binding to DNA at 0mM salt concentration. Error bars represent standard error.

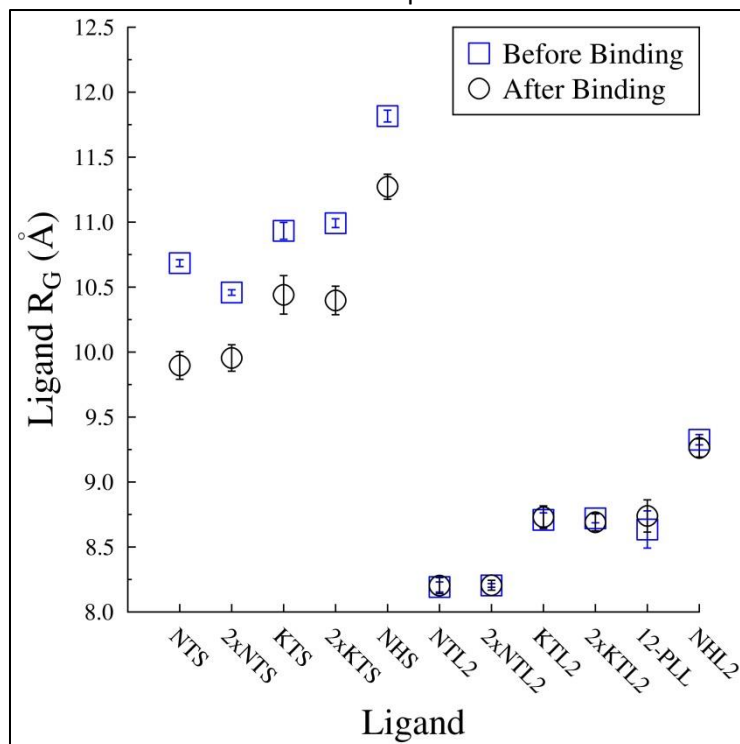


Figure A15: Average radius of gyration of the ligands before and after binding to DNA at 160mM salt concentration. Error bars represent standard error.

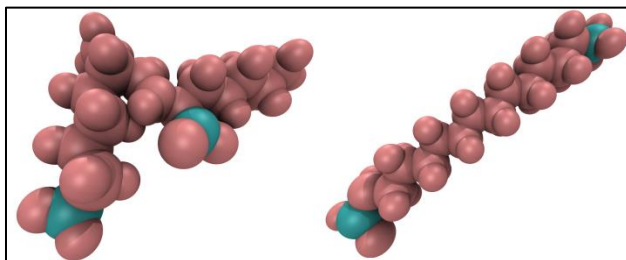


Figure A16: Diagram of the end-to-end distance of dilysine (left) and spermine (right). The atoms used to calculate the end-to-end distance are shown in teal.

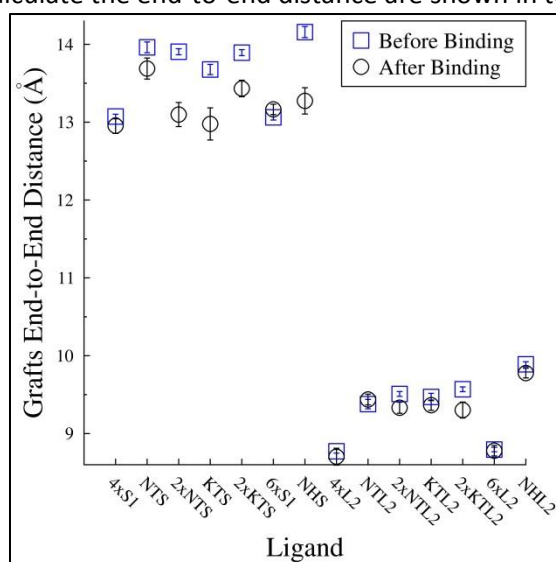


Figure A17: Average end-to-end distance of the grafts before and after binding to DNA at 0mM salt concentration. Error bars represent standard error.

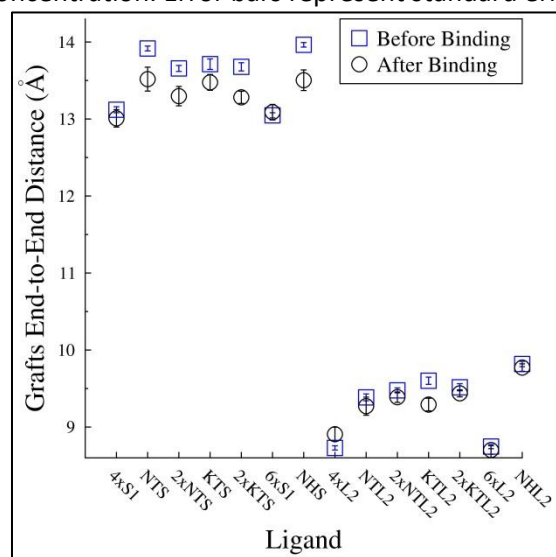


Figure A18: Average end-to-end distance of the grafts before and after binding to DNA at 160mM salt concentration. Error bars represent standard error.

D. Other Investigated Metrics

1. Radius of Gyration of the Core

Core size is predominantly dependent on the number of atoms in the core molecule. There is not much variation in core size between the unbound and bound states at either 0mM or 160mM salt concentration. It also appears that graft chemistry (spermine or dilysine) does not affect core sizes significantly. Figure A19 and Figure A20 show the average radius of gyration of the core for each ligand.

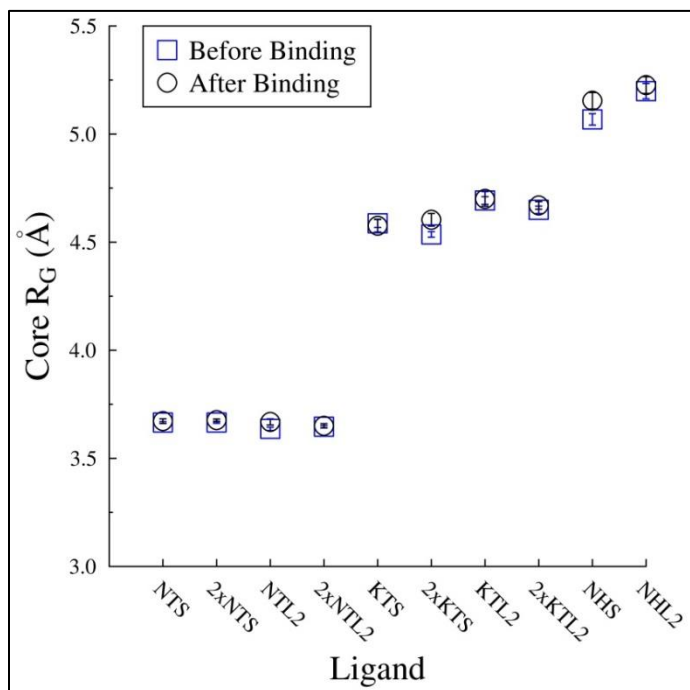


Figure A19: Average radius of gyration of the core before and after binding to DNA at 0mM salt concentration. Error bars represent standard error.

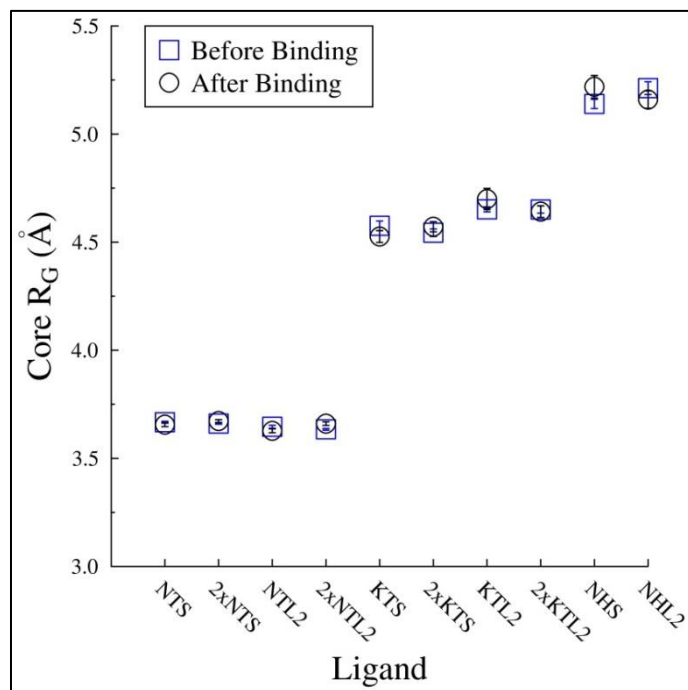


Figure A20: Average radius of gyration of the core before and after binding to DNA at 160mM salt concentration. Error bars represent standard error.

2. Phosphorus-to-nitrogen Distance

The nitrogen-to-phosphorus distance is defined for each nitrogen atom, but there are fewer nitrogen atoms than phosphorus atoms, so we looked at the phosphate-to-nitrogen distance as well. The phosphate-to-nitrogen distance is defined as the average of the distance between each phosphorus atom and the nitrogen atom nearest to that phosphorus atom.

Overall the information from this metric is very similar to the information from the nitrogen-to-phosphorus metric. However, we note from Figure A21 that the phosphate-to-nitrogen distance for the double ligand systems is shorter than for the single ligand systems. This is because there are more nitrogen atoms for the double ligand systems, and the ligands typically bind on opposite sides of the DNA, so there is no unbound side of the DNA in the double ligand system.

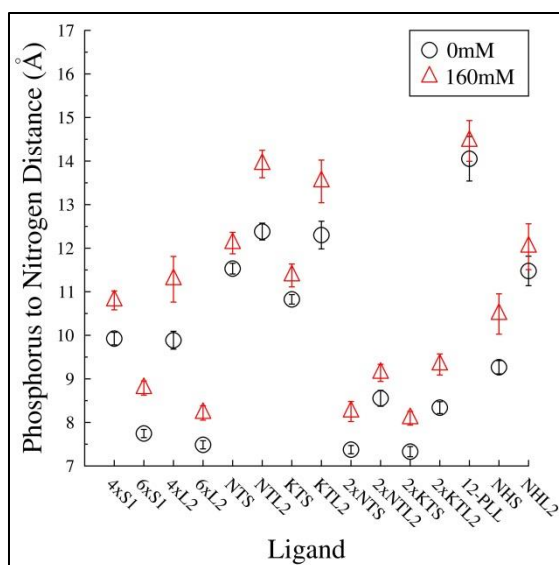


Figure A21: Average DNA phosphate to ligand nitrogen distance. Error bars represent standard error.

3. Number of Ligand Nitrogen Atoms within Four Angstroms of a DNA Phosphorus atom

As previously mentioned, the nitrogen-to-phosphorus distance accounts for all the nitrogen atoms in a ligand, but number of nitrogen atoms within four angstroms of the DNA phosphorus atoms focuses on the bound amine groups only. From Figure A22 it is apparent that the systems with more nitrogen atoms in total have more nitrogen atoms within four angstroms of the phosphorus atoms.

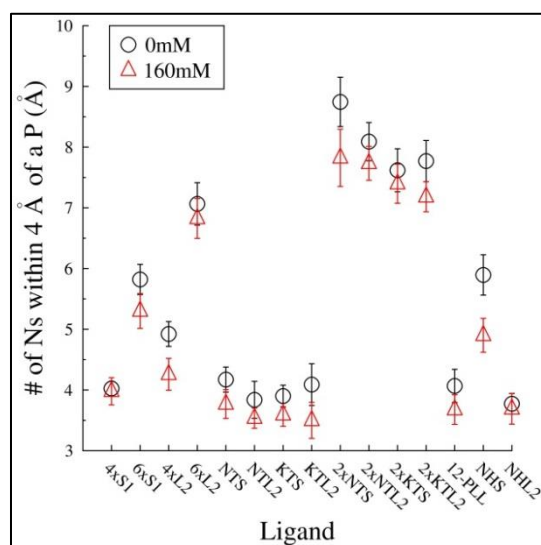


Figure A22: Average number of ligand nitrogen atoms within four angstroms of a DNA phosphorus atom. Error bars represent standard error.

The exception to this trend is NHL2, which has the same number of nitrogen atoms within four angstroms of the DNA as the lysine core-grafted systems with only four grafts. This is due to steric hindrance. NHL2 has very crowded, short grafts which prevents some of the grafts from reaching the DNA. Therefore, approximately the same number of amine groups binds to the DNA as the lysine grafted ligands with four grafts.

NHS is able to get more nitrogen atoms within four angstroms of the DNA strand than the spermine grafted ligands with four grafts because its grafts are flexible enough to reach around the core

and bind to the DNA. Also NHS is closer to the DNA than NHL2 (Figure 2c and Figure 4c), and the grafts are less crowded because they are not branched.

We also compared the percent fluorescence decreased (PFD) to the number of nitrogen atoms within 4 Å of a phosphorus atom in a manner similar to Figure 3. The correlation between the PFD and the number of nitrogen atoms was better than the correlation between the PFD and the center of mass data, but the number of nitrogen atoms near the DNA for NTS, KTS, and 4xS1 were too similar to say that there was a statistically significant correlation.

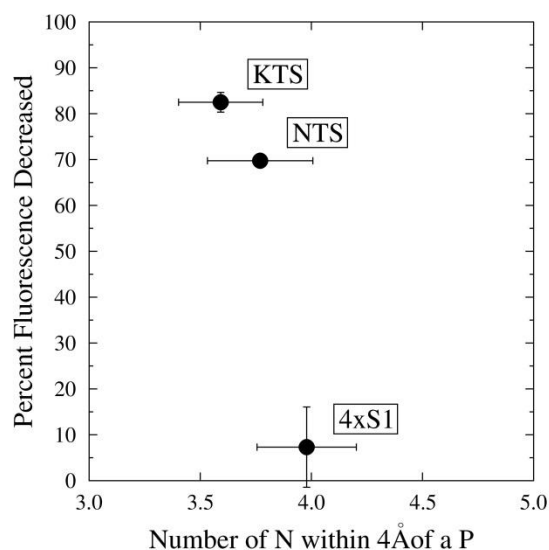


Figure A23: Comparison of the percent fluorescence decreased to the number of nitrogen atoms within 4 Å of a phosphorus atom.