

UNDERSTANDING THE MECHANISM OF ION-MEDIATED INTER-DNA AND
INTER-RNA INTERACTIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Li LI

August 2011

© 2011 Li LI
ALL RIGHTS RESERVED

UNDERSTANDING THE MECHANISM OF ION-MEDIATED INTER-DNA AND INTER-RNA INTERACTIONS

Li LI, Ph. D.

Cornell University 2011

DNA and RNA are very important biological molecules. Both are highly negatively-charged. Positively-charged ions can bind to them through electrostatic interactions. The interaction between DNA (or RNA) molecules is affected by these ions. In this research, we investigate the physical mechanism of ion-mediated inter-DNA and inter-RNA interactions using a variety of X-ray techniques and other analytical tools. In our experiments, we focus on the property of short strand (20 or 25 base pair) DNA and RNA due to their great potential in novel therapeutic applications. The major experimental technique involved in this research is Small Angle X-ray Scattering (SAXS). We explore this method to measure interactions between freely-moving nucleic acids in solution. We demonstrate the impact of positively-charged ions on the interactions between short double-stranded DNA (or RNA) molecules through a series of experiments and theoretical models. The valence of the ion ranges from divalent (Mg^{2+}), trivalent (such as cobalt hexamine and cobalt sepulchrates) to tetravalent (spermine). The results show that not only ionic strength but also the valence, size and hydration level of the ions as well as geometry of nucleic acids play important roles in the inter-nucleic acid interaction. This research provides insight into the physical mechanism of like-charge attraction and establishes the physical basis of DNA (or RNA) packaging achieved by small ions.

BIOGRAPHICAL SKETCH

Li LI was born in Nanjing, Jiangsu, China. He attended Nanjing No. 1 middle school, where he accumulated huge amount of interest in mathematics and physics. He got a bachelor's degree in physics in Nanjing University and Master's degree in physics in Rensselaer Polytechnic Institute. After that, he entered Cornell to persue the Ph.D. degree of biophysics and has been doing biophysical research in Lois Pollack group since July 2007. Besides physics, he is also very interested in music and a fan of hard rock and power metal.

I dedicate this to everyone who helped me throughout Ph.D. study period.

ACKNOWLEDGMENTS

I would like to sincerely acknowledge my advisor, Lois Pollack, for her guidance in biophysical research, and life. Thanks to members of the Pollack Group, Jessica Lamb, Xiangyun Qiu, Hye Yoon Park, Suzette Pabit, Steve Meisburger, Joshua Blose, Christopher Jones, Julie Sutton, and Huimin Chen for the help during the experiments and insightful discussions. Special thanks go to Suzette, who has been helping me both in experiment and all other aspects since I joined the group. I really enjoyed the time studying and working in Pollack lab for the past couple of years.

Facilities and support for X-ray scattering experiments were provided by the Cornell High Energy Synchrotron Source. Thanks to Ken Finkelstein, Arthur Woll, and all of the CHESS staff, especially the operators who helped during beam time.

Facilities for microfabrication were provided by the Cornell Nanoscience and Technology Facility (CNF).

Facilities for running gels were provided by Nanobiotechnology Center (NBTC).

Thanks to Gerald Feigenson and Diane Colf for the initial guide in graduate school.

Thanks to Serena DeBeer for very useful discussions.

Thanks to Nathan Baker and Dennis Thomas for collaborations.

Thanks to Renee King and other staff in Applied and Engineering Physics Department for the help in the administrative affairs.

Thanks to lots of my friends especially my roommates at Cornell who helped me go through difficult times.

Finally, thanks very much to Lois Pollack, Joel Brock and Holger Sondermann for serving as my special committee members for my Ph.D. Study. Thanks to my parents for their support.

This research was supported by the National Science Foundation and the National Institute of Health.

TABLE OF CONTENTS

1. Introduction	1
1.1 Biological roles of DNA and RNA.....	2
1.2 Secondary structures of DNA and RNA.....	3
1.3 DNA condensation.....	7
1.4 Goal of this research.....	9
1.5 Thesis overview.....	12
References.....	13
2. Experimental techniques and computation methods.....	14
2.1 X-ray Scattering Methods—SAXS, ASAXS, mE-ASAXS.....	14
2.1.1 Introduction.....	14
2.1.2 Theoretical background.....	15
2.1.3 Computational background—form factor computation.....	25
2.1.4 Experimental setup.....	29
2.2 X-ray fluorescence Methods—X-ray Absorption Fine Structure.....	31
2.2.1 Theory.....	31
2.2.2 Applications.....	35
2.3 Conclusion.....	36
References.....	37
3. Theoretical modeling.....	38
3.1 Counterion condensation theory.....	38
3.2 Poisson Boltzmann approach.....	42
3.3 Electrostatic Zipper Model.....	44
3.4 Bridges-and-clinches model.....	47

3.5 Conclusion.....	49
References.....	50
4. End-to-end stacking measurement.....	52
Abstract.....	52
Introduction.....	52
Experimental techniques.....	54
Sample preparation.....	54
Results.....	57
Discussion.....	61
Conclusion.....	62
References.....	63
5. Double-stranded RNA resists condensation.....	65
Abstract.....	65
Introduction.....	65
Sample preparation.....	67
Results and discussion.....	69
Conclusion.....	77
References.....	79
6. Counting trivalent ions and measuring hydration.....	82
Abstract.....	82
Introduction.....	82
Experimental techniques.....	84
Sample preparation.....	89
Results and discussion.....	90

Conclusion.....	97
References.....	98
7. Future work.....	100
7.1 DNA and RNA condensation by spermine.....	101
7.2 Wide Angle X-ray Scattering.....	105
7.3 End-to-end stacking: RNA dumbbell.....	110
References.....	113
8. Conclusion.....	114
References.....	118
Appendix.....	119

LIST OF FIGURES

1.1 Watson-Crick base-pairing in double-stranded DNA	4
1.2 structure of 25bp A-RNA and B-DNA	6
1.3 dsDNA packaging into nucleus	8
1.4 dsDNA/dsRNA condensation induced by trivalent ions	11
2.1 form factor and total scattered profile	18
2.2 a sketch of principle of ASAXS	21
2.3 energy selection by CHOOCH for mE-ASAXS	24
2.4 computed versus experimental determined form factor	28
2.5 a sketch of SAXS setup	30
2.6 principle of XAFS	32
2.7 a sketch of the XAFS spectrum	34
3.1 a sketch of the basic framework of manning's condensation theory	40
3.2 electrostatic zipper model	46
4.1 "dumbbell" DNA construction	56
4.2 scattering profile of dsDNA, "semi-dumbbell" DNA and "dumbbell" DNA	59
4.3 structure factor versus DNA concentration to obtain A_2	60
5.1 dsDNA/dsRNA condensation induced by co-hex	70
5.2 SAXS profile of dsDNA/dsRNA in co-hex buffer	72
5.3 SAXS profile of RNA, both computed and experimental	75
6.1 conversion of fluorescence scan by "CHOOCH"	88
6.2 SAXS profile of DNA in co-hex, co-sep, NaCl buffers	91
6.3 5-energy SAXS profiles of DNA in co-hex and co-sep buffers	92
6.4 $I(Q=0)$ versus ion scattering factor	93

6.5 XAFS spectra of co-hex, co-sep and co-en associated with DNA	95
7.1 structure of spermine molecule	101
7.2 dsDNA/dsRNA condensation induced by spermine	103
7.3 SAXS profile of dsDNA/dsRNA in spermine buffer.	104
7.4 a picture of WAXS experimental setup at G-line CHESS.	106
7.5 WAXS profile of dsDNA/dsRNA in co-hex buffer	108
7.6 a picture of co-hex molecule in the major groove of RNA	109
7.7 plot of A_2 of RNA- K^+ and RNA- Mg^{2+} versus ion concentration	112

CHAPTER 1

Introduction

The research presented in this thesis focuses on the mechanism of ion-mediated DNA-DNA (or RNA-RNA) attraction with the goal of establishing the basis for potential therapeutic applications that rely on DNA (or RNA) packaging. DNA and RNA are negatively-charged chains of varied length. A key challenge in developing modern therapeutics is to package these chains into a small volume for delivery. To optimize packaging requires a thorough understanding of how these charged chains interact with each other. The interactions are mediated by external agents which usually are certain ligands. Unfortunately, the underlying physical mechanism of like-charge attraction is still not well understood at the level of physics. Here we provide insights into this topic. In particular we are most interested in short (20~30 base pair) DNA (and RNA) molecules primarily because of great potential of RNA interference pathway [1, 2, 3] in which 21 ~ 25 base pair (bp) RNA molecule plays an important role in gene expression (described in 1.1). These short double-stranded RNA molecules may be used for promising therapeutics due to their ability to target and silence specific genes. In this research, we use both experimental and theoretical tools to investigate how small ions, from monovalent to tetravalent, affect the interactions between short double-stranded DNA (or RNA) molecules.

1.1 Biological roles of DNA and RNA

DNA or Deoxyribonucleic acid carries the required genetic information that allows living organisms to function and replicate. DNA is most commonly a linear

polymer in eukaryotes (in human cells, DNA in the nucleus is about 2~3 meters long), and a circular polymer in prokaryotes. The information carrying units of DNA are called genes. The exact number of genes in human DNA is still uncertain. The most widely recognized information that is stored in the genes is the amino acid sequence for proteins. The relevant DNA sequence is copied (transcribed) into a complementary RNA sequence (called messenger RNA). This RNA copy is translated into chains of amino acids that fold into proteins to perform cellular functions.

Clearly, the primary function of DNA is in storage of genetic information. In contrast, RNA participates in many cellular functions that surpass transcription and translation. RNA plays an active role by catalyzing biological reactions, controlling gene expression, or sensing and communicating responses to cellular signals. To be more specific, as a typical example, there is a lot of interest in the growing field of small interfering RNA (siRNA) research. In biological systems, siRNA, sometimes also known as silencing RNA, is a class of double-stranded RNA molecules, 20-25 nucleotides in length. Those short double-stranded (ds) RNA molecules are involved in the RNA interference (RNAi) pathway that mediates sequence-specific silencing of gene expression. RNAi was initially discovered in *Caenorhabditis elegans* [1] then in mammalian cells [2, 3]. The proposed process of RNAi pathway begins with long dsRNA (usually >100bp) that triggers the whole process. This long dsRNA is chopped by RNase enzyme into 21-23 nucleotide (nt) RNA (the siRNA) molecules. These short siRNAs are incorporated into a specific biological complex and guide it to “silence” target genes by cleaving the corresponding mRNA and stopping protein production. The detailed mechanism is much more complicated than described above and is not well understood.

1.2 Secondary structures of DNA and RNA

Most functional DNA occurs in a double-stranded (ds) form which is of main interest to this research. dsDNA consists of two helical chains (strands) coiled around each other. Each unit of this backbone, known as a nucleotide, consists of a phosphate, base and sugar residue. The nucleotides are connected by ester bonds to form the backbone. There are four types of bases—adenine (A), cytosine (C), guanine (G) and thymine (T). In dsDNA, each base on one strand forms hydrogen bonds with one specific type of base on the other strand as shown in figure 1.1. For example, A pairs with T, and G pairs with C. Exceptions to this Watson Crick base pairing will not be discussed in this thesis. Genetic information is encoded in the sequence of bases in a particular chain.

dsDNA exists in many possible conformations including A-DNA, B-DNA, and Z-DNA forms. Only B-DNA (right-handed) and Z-DNA (left-handed) have been observed in functional organisms; B-DNA is the most common form found under normal cellular conditions. Figure 1.2 (right) shows a model of a 25 base pair (bp) B-DNA. B-DNA has a helical pitch of 34 \AA and a radius of 10 \AA . Because each phosphate group carries one electron charge, DNA is highly negatively charged. The linear surface charge density is $-2e/3.4 \text{ \AA}$, resulting in a very strong electric field on the surface. Another important structural feature of the molecule is the formation of grooves in the base paired structures. The bigger 22 \AA wide groove, is called the major groove. The 12 \AA wide groove is called the minor groove. Since the edges of the bases are more accessible in the major groove it usually serves as an important protein binding site.

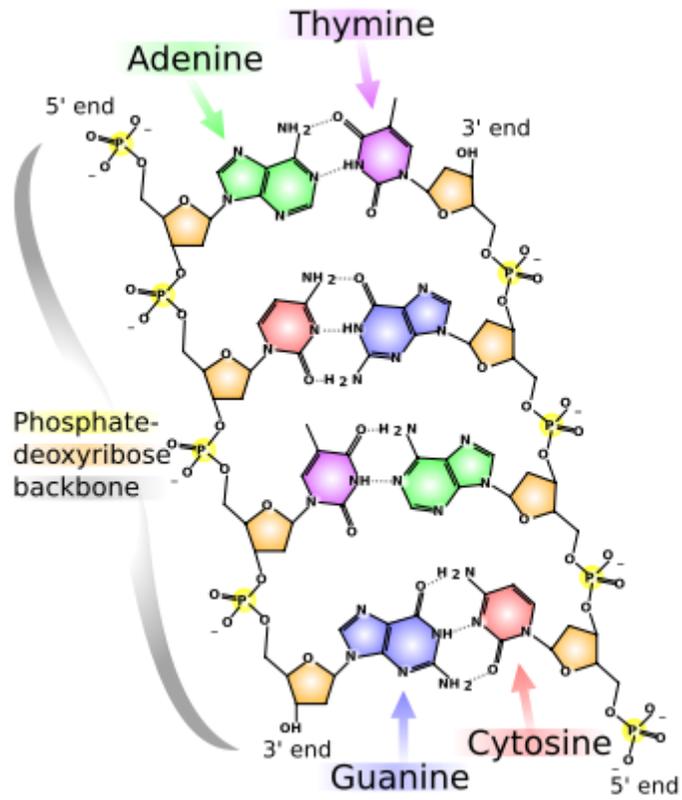


Figure 1.1 Watson-Crick base-pairing in double-stranded DNA

Source: <http://www.dna-sequencing-service.com/wp-content/uploads/2010/09>

Like DNA, RNA is also a chain of nucleotides. However there are major differences in the chemical structure of DNA and RNA. First, the 2-deoxyribose sugar in DNA is replaced in RNA by the alternative pentose sugar ribose. Second, RNA uses the base uracil (U) in place of DNA's thymine (T). Both uracil and thymine base pair with A. Unlike DNA, most naturally occurring RNA molecules are single-stranded. Single-stranded RNA molecules can fold into complex three-dimensional structures which enable specific functions in the cell, including sensing and interacting with small ions [4]. However, for the purposes of this thesis, we are primarily interested in the properties of double-stranded RNA molecules, because of their utility to the RNAi process discussed in 1.1). In comparison to dsDNA, dsRNA is usually in the A-form which is a wider right-handed spiral with a shallow wider minor groove and a narrower deeper major groove (see figure 1.2, left side). The surface charge density thus is $-2e/2.8\text{\AA}$ which is greater than that of dsDNA. As a result, the counterion binding pattern differs from that of DNA, an important feature discussed in chapter 5.

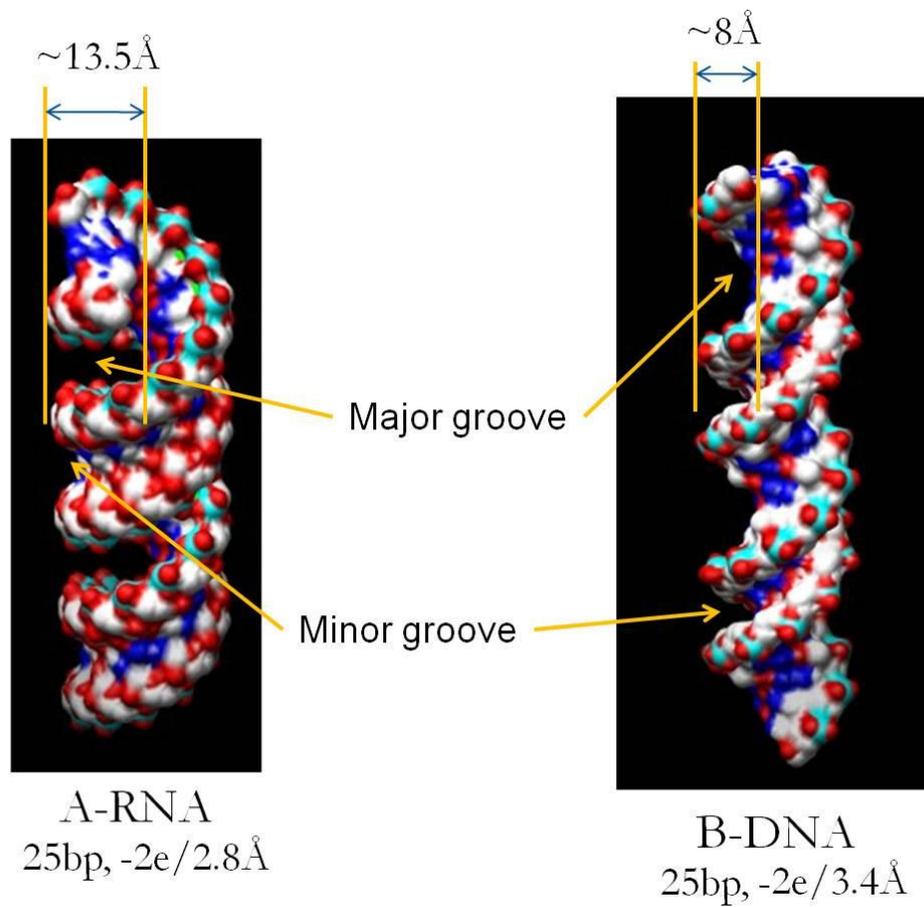
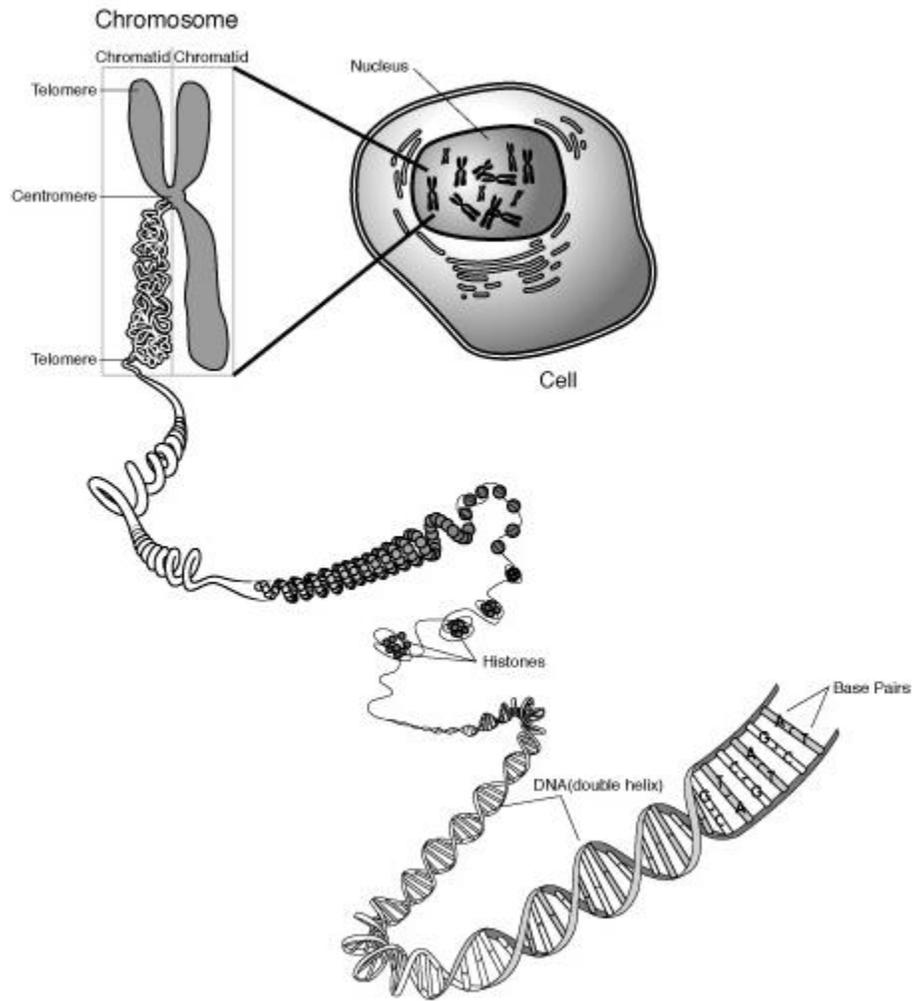


Figure 1.2 structure of 25bp A-RNA and B-DNA

25bp A-form RNA is shown in the left panel and B-DNA is on the right. Note the difference in configuration and charge density. A-RNA is a shorter and fatter cylinder with deeper major groove compared to B-DNA.

1.3 DNA condensation

DNA condensation is the process of compacting DNA molecules *in vitro* or *in vivo*. The mechanism of DNA condensation is of great interest and importance to biology, physics, and therapeutics. Biologically, DNA molecules in cells are tightly compacted in a highly concentrated state. The volume fraction of DNA in viruses and chromatin is very high (15% ~ 60%) [5]. In eukaryotic cells, long double-stranded DNA molecules (of the order of magnitude of meters in length) wrap around positively charged proteins to neutralize the negative charge along the DNA backbone. This complex is subsequently organized into long but compact structures called chromatin (See figure 1.3). However, DNA condensation can also be induced by the addition of small quantities of multivalent ions, raising interesting (and unresolved) questions about ion-induced attraction between like-charged objects [6, 7, 8]. Clearly ions can play a major role in compacting the polyelectrolyte chain. Although it is likely that correlations between condensed counterions may ultimately be responsible for attraction, the exact mechanism has not yet been determined. In addition to understanding the physics of this process, we note its importance to therapeutic processes, such as gene delivery, which requires therapeutic DNA to be packaged in a very small volume. To manipulate the DNA in such a controllable manner, it is important to first obtain a fundamental understanding of how counterions mediate interactions between DNA molecules.



<http://www.accessexcellence.org/AB/GG/chromosome.html>

Figure 1.3 dsDNA packaged into nucleus

Long dsDNA is packaged into higher order structure chromosome and then fit into small nucleus. For human being, the total length of DNA is around 2~3 meters while the diameter of nucleus is only several micrometers.

Source:

<http://employees.csbsju.edu/hjakubowski/classes/ch331/dna/oldnastructure.html>

DNA condensation can be induced *in vitro* either by entropy, applying an external force to bring the DNA molecules together or by enthalpy, inducing attractive interactions between the DNA molecules. The former can be achieved by increasing the osmotic pressure by ‘crowding’, e.g. decreasing the available volume by introducing neutral polymers in the presence of monovalent salts. In this case, the forces pushing the double helices together result from entropic random collisions with the crowding polymers surrounding DNA condensates. Salt is required to neutralize DNA charges and decrease DNA-DNA repulsion. The second possibility is realized by adding multivalent cationic charged ligands (multivalent cations, polyamines, or proteins, etc) to induce attractive interactions between DNA molecules. A comprehensive review of condensation of DNA by multivalent cations is provided in ref 9.

Many experimental studies have probed small ion induced DNA condensation [10-12] *in vitro*. Most experimental investigations use DNA molecules whose length exceeds the DNA persistence length (~150bp). Since bending or twisting of the DNA strand must be considered to derive a complete solution to the problem, the use of long DNA increases the complexity of the system, and limits comparison with theoretical models. Unraveling the effects of mechanical degrees of freedom and electrostatics can be quite challenging. Therefore, in order to focus on electrostatics and avoid mechanical complication, we use 25bp DNA for this research. These DNA duplexes are significantly shorter than the persistence length and allow us to focus on electrostatic contributions to DNA condensation.

1.4 Goal of this research

To obtain a better physical picture of DNA condensation, it is important to understand the underlying mechanism of ion-induced compaction. Using small ions instead of using proteins to condense DNA for genetic material delivery has the advantage of avoiding possible side effects such as immune rejection from cells of the living body. In this research we investigated DNA-DNA interactions in solutions containing ions with a variety of valences, ionic strengths, sizes and hydration levels (discussed in chapter 4, 5, 6, 7).

We also explored the similar process of RNA condensation, focusing on short double-stranded RNA due to its important role in RNAi pathway as described in section 1.1. We would like to understand how electrostatic interactions between short dsRNA molecules are mediated by the small cations that condense DNA very effectively (see figure 1.4). Surprisingly, although dsDNA and dsRNA have similar structures and charge densities they behave quite differently in the presence of multivalent ions. Our research suggests that therapeutic strategies developed to manipulate dsDNA may not be effective for dsRNA (discussed in detail in chapter 5). Results of parallel experiments on dsRNA and dsDNA provide insight into the underlying mechanism of ion-induced DNA condensation and reveal the distinct properties of dsRNA molecules.

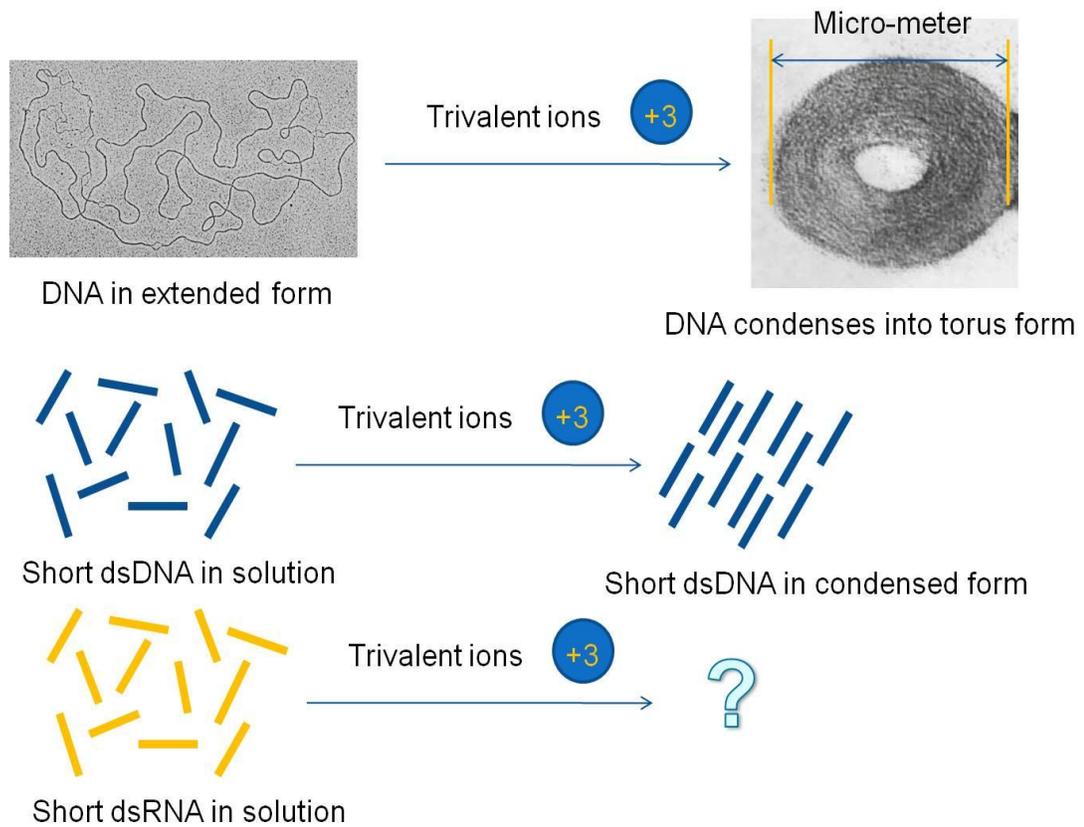


Figure 1.4 dsDNA/dsRNA condensation induced by trivalent ions

Long strand DNA in ionic solutions can condense into torus form by adding certain trivalent ions [13]. Short dsDNA has similar property but how about short dsRNA? We would like to know under the same ionic condition, do RNAs and DNAs behave similarly?

1.5 Thesis overview

In chapter 2, I introduce all the X-ray scattering and fluorescence techniques employed in this research. A computational method to calculate scattering profiles (or form factors) of 25bp DNA is also demonstrated in chapter 2. Chapter 3 covers the relevant theoretical models of DNA-ion interactions and DNA condensation. Chapter 4 presents experimental results illustrating interactions between DNA duplexes in solutions containing divalent ions, specifically illustrating that end-to-end stacking is the preferred interaction mode. Chapter 5 presents experimental results on multivalent ion induced interactions between RNA duplexes, demonstrating the importance of nucleic acid geometry to the condensation process. The next chapter outlines a strategy for counting the number of trivalent ions present as inter-duplex interactions result in attractive forces. We discuss methods for measuring important properties of these ions, including hydration in chapter 6. Results show that in addition to electrostatic interactions, the hydration of highly charged ions plays an important role in DNA condensation. Finally, I propose some interesting follow-up projects for this research in chapter 7. Some preliminary results are shown in chapter 7 as well. A conclusion of this research is drawn in chapter 8.

REFERENCES

- [1] A. Fire, S. Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, *Nature* **391**, 806 (1998)
- [2] M. T. McManus, and P. A. Sharp, *Nat. Rev. Genet.* **3**, 737 (2002)
- [3] S. M. Hammond, A. A. Caudy, and G. J. Hannon, *Nat. Rev. Genet.* **2**, 110 (2001)
- [4] D. E. Draper, *RNA* **10**, 335 (2004)
- [5] E. Kellenberger, *Biophys. Chem.* **29**, 51 (1987)
- [6] B. -Y. Ha, and A. J. Liu, *Phys. Rev. Lett.* **79**, 1289 (1997)
- [7] L. Dai, Y. Mu, L. Nordenskiöld, and J. R. C. Van der Maarel, *Phys. Rev. Lett.* **100**, 118301 (2008)
- [8] R. Zhang, and B. I. Shklovskii, *Phys. Rev. E* **72**, 021405 (2005)
- [9] V. A. Bloomfield, *Biopolymers*, **31**, 1471 (1991)
- [10] B. A. Todd, and D. C. Rau, *Nucleic Acids Res.* **36**, 501 (2007).
- [11] E. Raspaud, M. O. Delacruz, J. L. Silkorav, and F. Livolant, *Biophys. J.* **74**, 381 (1998)
- [12] Y. Mamasakhlisov, B. A. Todd, A. V. Badasyan, V. F. Morozov, and V. A. Parsegian, *Phys. Rev. E* **80**, 031915 (2009)
- [13] W. M. Gelbart, R. F. Bruinsma, P. A. Pincus, and V. A. Parsegian, *Physics Today*, **53**, 38 (2000)

CHAPTER 2

Experimental techniques and computation methods

The major experimental techniques involved in this research include Small Angle X-ray Scattering (SAXS), multiple-Energy Anomalous Small Angle X-ray Scattering (mE-ASAXS), and X-ray Absorption Fine Structure (XAFS). The basic principles of these techniques are outlined in this chapter. A numerical approach to compute the scattering profile (form factor) of DNA molecule in solution using the Debye formula is also introduced in this chapter. Results of these computations can be directly compared with experimentally obtained scattering profiles.

2.1 X-ray Scattering Methods—SAXS, ASAXS, mE-ASAXS

2.1.1 Introduction

Small-angle X-ray scattering (SAXS) is a well-established X-ray technique. The elastic scattering of X-rays (wavelength 0.1~0.2 nm) by the sample is recorded at very low angles (typically 0.1 - 10°). Since this angular range corresponds to real space dimensions of order $2\pi/q$, the scattering profile contains information about the shape and size of macromolecules or characteristic distances associated with partially ordered materials. Typically SAXS is sensitive to macromolecular features on length scales between 5 and 25 nm, and of repeat distances in partially ordered systems of up to 150 nm [1]. This method is complementary to X-ray crystallography or Nuclear magnetic resonance (NMR) spectroscopy which produce higher resolution (on the order of magnitude of Å) macromolecular structures. The greatest advantage of SAXS is its use in solution and for all sizes of macromolecules. In the case of biological

macromolecules such as proteins, crystallization and freezing are usually required for X-ray crystallography but not for SAXS. When the molecular weight is greater than ~30,000Da, NMR methods encounter problems but this is not an issue for SAXS. On the other hand, SAXS has its own limitations. For example, the spatial averaging due to the random orientation of dissolved or partially ordered molecules leads to a loss of information in SAXS measurements.

2.1.2 Theoretical background

Small Angle X-ray Scattering (SAXS)

In this research SAXS is primarily used to report interactions between short double-stranded DNA (or RNA) molecules. The structures of DNA and RNA are already well-known and were briefly described in chapter 1. SAXS experiments measure the total scattering profile [2, 3]

$$I(Q) = NP(Q)\left(1 + \frac{\langle F \rangle^2}{\langle F^2 \rangle} [S(Q) - 1]\right)$$

where N is a scale factor or number density of the molecules, $Q = 4\pi\sin\theta/\lambda$ is the momentum transfer and 2θ is the scattering angle. The form factor $P(Q) = \langle F^2 \rangle$ represents the scattering of a single macromolecule, $F = \iiint \rho(\vec{r}) \exp(iQr) d\vec{r}$ is the scattering amplitude or mathematically, the Fourier transform of the electron density distribution of the molecule, $\langle \rangle$ indicates an average over all DNA orientations (spherical approximation), $S(Q) = 1 + N \iiint \exp(iQr) [g(r) - 1] dr$ is called structure factor which reports inter-molecule ordering. For a system of monodisperse spherical or nearly spherical molecules, the formula above reduces to [2, 4]

$$I(Q) = NP(Q)S(Q)$$

Within this approximation the total intensity equals the product of the form factor and structure factor. The effect of the structure factor is most pronounced in the low Q regime and can become prominent when strong interaction between molecules is present.

The form factor $P(Q)$ reports the scattering when interaction between the molecules is neutralized or can be ignored. It can be determined either experimentally or theoretically. Experimentally, as described in chapter 4, we use scattering profile of a dilute solution (0.1mM) of DNA as the form factor. Theoretically, $P(Q)$ can be calculated from the Debye formula by summing over the scattering from all the atom pairs involved. A detailed demonstration of this approach is described in section 2.1.3, including a description of a program written to carry out this computation. More information about implementing this program can be found in the appendix. The form factor can also be computed from publically available software packages, such as CRY SOL [5]. However, these packages were designed primarily for proteins which have differing solvent shells than nucleic acids which have differing solvent shells.

Once total scattered intensity $I(Q)$ and form factor $P(Q)$ are obtained, it is straightforward to extract the structure factor $S(Q)$ to determine the strength of inter-particle interaction. Qualitatively, we obtain the interaction information by comparing $I(Q)$ and $P(Q)$. This approach is illustrated in figure 2.1: the profiles $I(Q)$ and $P(Q)$ are scaled to match at mid- Q and high- Q regimes where the structure factor does not contribute to the measured scattering. Strong inter-molecule repulsion is indicated by the sharp peak in $I(Q)$ in the leftmost panel. The “Bragg distance” equals $2\pi/Q_{\max}$ where Q_{\max} represents the peak position corresponding to the mean inter-molecular distance. Weak repulsion is reflected by the low Q downturn of $I(Q)$ with

respect to $P(Q)$ in the second panel while weak attraction manifests itself as the low Q upturn in the fourth panel. When attraction and repulsion are almost “neutralized”, $I(Q)$ and $P(Q)$ match at low Q . One of the most convenient ways to quantify the interaction between molecules is to calculate the second virial coefficient A_2 from the experimentally-determined structure factor $S(Q)$. $A_2 > 0$ signifies repulsive interactions while $A_2 < 0$ implies intermolecular association.

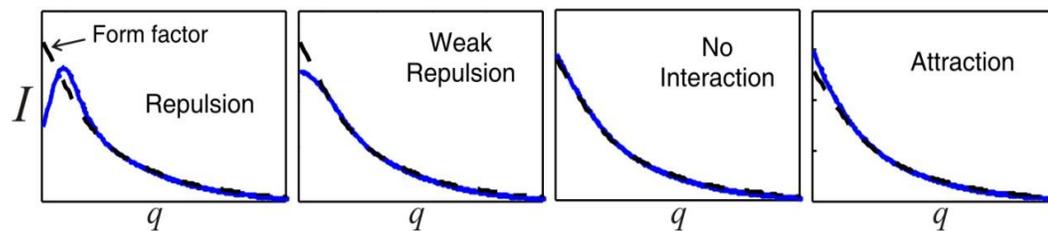


Figure 2.1 form factor and total scattered profile

The form factor (black dash line) and total scattering profile (blue line) are matched at mid-to-high Q regime. The deviation at low Q regime can be used to inform about inter-molecule interactions as indicated in each of the four panels above.

Anomalous Small Angle X-ray Scattering (ASAXS)

ASAXS reports the distribution of counterions around dsDNA/dsRNA [6, 7]. Since the scattering from the ions cannot be neglected relative to the scattering from the DNA, the total scattering intensity is expressed as follows:

$$I(Q, E) = |f_{NA}F_{NA}(Q) + Nf_{ion}F_{ion}(Q)|^2$$

where Q is momentum transfer defined above, N is the excess number of ions, E is the energy of incident X-ray beam, $F(Q)$ is the Fourier transform of the electron density distribution of nucleic acid or ion, f is the scattering power or scattering factor of nucleic acid or ion. Note that the total scattered intensity depends on both Q and E . This dependence is critical if ions are selected with accessible absorption edges; in this case the ion scattering factor f is energy- dependent. Near the ion absorption edge, the ion scattering factor f can be decomposed into three terms—a real part $f'(E)$, an imaginary part $f''(E)$ and the energy-independent solvent-corrected scattering factor f_0 :

$$f_{ion}(E) = f_0 + f'(E) + if''(E)$$

The real part reflects the changes in scattering intensity close to resonant edge and the imaginary part represents the changes in sample absorption. Both factors can be obtained by measuring X-ray fluorescence from a buffer solution containing the energy-dependent scattering element of interest. ASAXS exploits the energy dependence of scattering to extract spatial information about the ion distribution, as illustrated in figure 2.2. First, two energies are chosen with E_{off} well below the absorption edge and E_{on} close to, but below the edge. Second, subtraction of two scattering profiles acquired at E_{on} and E_{off} removes the energy-independent term. The energy-dependent anomalous signal $\Delta I(Q, E)$ which contains information about counterion distribution around DNA/RNA has the form:

$$\Delta I(Q, E) = I(Q, E_{off}) - I(Q, E_{on}) \approx 2f_{NA} F_{NA}(Q) N F_{ion}(Q) |f'(E_{off}) - f'(E_{on})|$$

This anomalous signal can be loosely interpreted as the Fourier transform of the electron density distribution of the ions modified by the form factor of nucleic acid.

The reason to have \approx in the formula above is because the F_{ion} term is much smaller than the DNA (or RNA) scattering term F_{NA} and thus ignored. Likewise, all the scattering from non-resonant ions is small comparing to DNA (or RNA) scattering and only contributes to energy-independent term which is cancelled out by the subtraction.

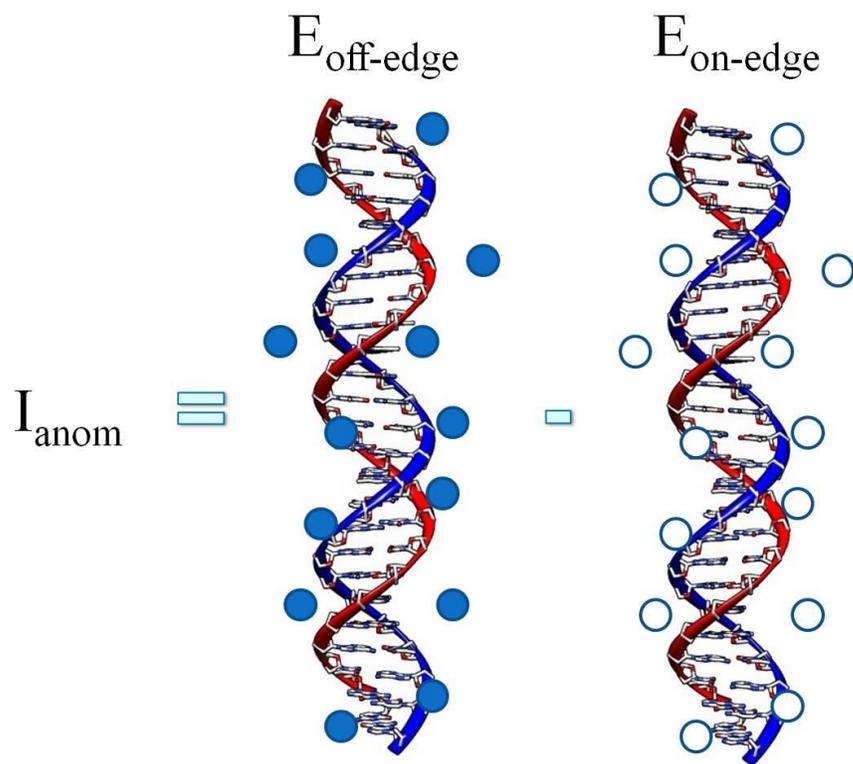


Figure 2.2 a sketch of principle of ASAXS

Scattering profiles measured at two different energies are subtracted to yield anomalous signal. The anomalous signal contains the information about the counterion distribution around dsDNA (see text for details).

Multiple-Energy Anomalous Small Angle X-ray Scattering (mE-ASAXS)

The multiple-Energy Anomalous Small Angle X-ray Scattering (mE-ASAXS) is a newly-developed technique that can be used to directly compute the excess number of ions around macromolecules [8]. The procedure is summarized below.

According to basic scattering theory [1 and references within] we have $f'' \ll f'$ when the X-ray energy is close to but below the ion absorption edge. Therefore the total scattered intensity $I(Q, E)$ defined in previous sub-section can be expanded into quadratic form of $f'(E)$ while ignoring $f''(E)$:

$$I(Q, E) = a(Q) \cdot (f'(E))^2 + b(Q) \cdot f'(E) + c(Q, E)$$

where

$$a(Q) = N^2 F_{ion}^2$$

$$b(Q) = N \cdot [2f_{NA}F_{NA}F_{ion} + 2f_0NF_{ion}^2]$$

$$c(Q, E) = (f_{NA}F_{NA})^2 + 2f_{NA}f_0NF_{NA}F_{ion} + f_0^2N^2F_{ion}^2$$

We measured the scattering profile at multiple (5~7) energies below the absorption edge. We used the program CHOOCH [9] (see chapter 6, illustrated in figure 2.3) to determine the dependence of f' on energy. In principle, at each fixed Q , a quadratic fit of $I(Q, E)$ with respect to $f'(E)$ yields $a(Q)$, $b(Q)$ and $c(Q)$. Note that F_{ion}^2 is much smaller than all other terms, therefore $a(Q)$ is negligible. The excess number of ions is then derived by observing the relationship such that

$$N = b(Q = 0) / 2\sqrt{c(Q = 0)}$$

$F_{ion}(Q = 0) = 1$ is used to obtain the relationship above. The values of $b(Q=0)$ and $c(Q=0)$ can be computed by two approaches. First, the full scattering profile can be decomposed into the sum of two curves: $b(Q)$ and $c(Q)$. Once the latter curves have

been obtained, either GNOM [10] fit or Guinier fit is used to extrapolate the values of $b(Q=0)$ and $c(Q=0)$. Alternatively, GNOM fit (or Guinier fit) applied to the total scattering profile $I(Q,E)$ provides $I(Q=0, E)$. Both $b(Q=0)$ and $c(Q=0)$ can be derived from a linear fit of $I(Q=0, E)$ with respect to $f^*(E)$. Consistent results are achieved using either method. Furthermore, to get the number of ions, an absolute calibration of the scattering intensity is required. Scattering from water can be used as a calibrant, as described in ref [11]. For monovalent and divalent counterions, the results are in good agreement with the number of ions predicted by Nonlinear Poisson Boltzmann approach.

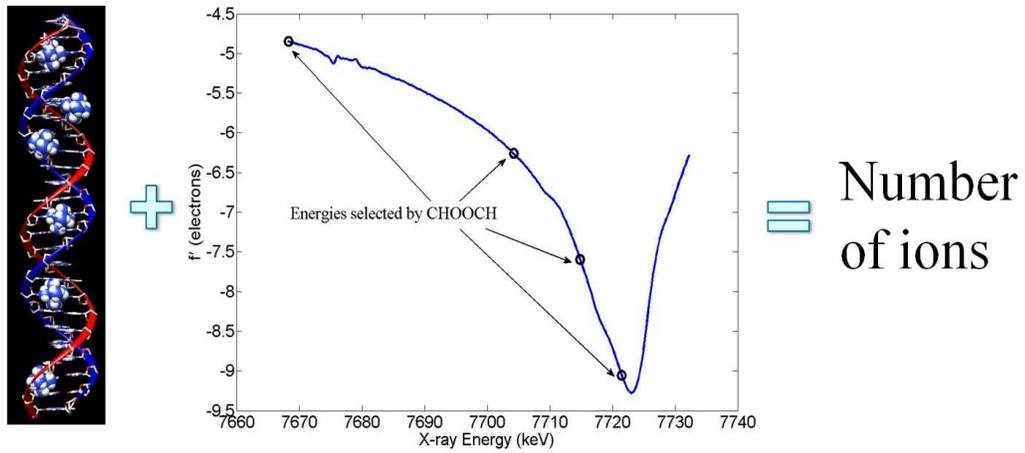


Figure 2.3 energy selection by CHOOCH for mE-ASAXS

A program called “CHOOCH” is used to select multiple energies below the absorption edge of the ion of interest for mE-ASAXS experiment. The example above is the plot of anomalous scattering factor versus X-ray energy around cobalt K-edge.

A faster implementation of this method can be applied when only two distinct energies E_1 and E_2 are selected. Under the assumption that f'' can be neglected, the number of ions N is derived directly from the original expression of $I(Q,E)$.

$$N = [\sqrt{I(Q=0, E_1)} - \sqrt{I(Q=0, E_2)}] / [f'(E_1) - f'(E_2)]$$

Even if water calibration is unavailable, this method can report the relative number of ions and can be used to compute changes of ion numbers quickly.

2.1.3 Computational background—form factor computation

Derivation of Debye formula

The form factor of a macromolecule (defined above as scattering profile of an isolated molecule) can be determined by computation using the Debye formula, as illustrated in detail in ref [1]. The implementation is as follows: For a system with atoms at positions r_1, r_2, \dots, r_N , the total scattering amplitude takes the form

$$F(Q) = \sum_{j=1}^N F_j(Q) \cdot \exp(-iQr_j)$$

where $F_j(Q)$ is the scattering amplitude of each atom and the exponential term reflects phase modification due to differing locations of the atoms. The scattering intensity is the absolute square of the amplitude averaged over all possible orientations:

$$I(Q) = \left\langle \sum_{j=1}^N \sum_{k=1}^N F_j F_k^* \cdot \exp[-iQ(r_j - r_k)] \right\rangle$$

This expression can be separated into two terms—the first has no phase factor when $j = k$, the second one, an interference term, depends on the phase which is a function of $r_{jk} = r_j - r_k$. The scattering intensity can be written as

$$I(Q) = \sum_1^N I_i(Q) + \langle \sum_{j \neq k}^N \sum_{k \neq j}^N |F_j| |F_k| \cdot \cos [Qr_{jk} + \delta] \rangle$$

where δ is the phase that depends on orientation of the atom. When the atom is approximated as a sphere, δ will vanish and the scattering intensity will be independent of orientation. The spherical average yields:

$$I(Q) = \sum_1^N I_i(Q) + 2 \sum_{j \neq k}^N \sum_{k \neq j}^N |F_j(Q)| |F_k(Q)| \cdot \frac{\sin(Qr_{jk})}{Qr_{jk}}$$

When N is very large, the interference term dominates the single term. This was first derived by Debye [1] and the method is named after him. It is widely used for computing the scattering profile of a composite particle system.

Form factor calculation

As described above, accurate computation of the form factor requires that we include a contribution from the nucleic acid as well as from the counterions that surround it. In this project, calculation of form factor of 25bp DNA in monovalent salt solution is performed based on counterion distribution around DNA determined by Nonlinear Poisson Boltzmann model (discussed in chapter 3). The result is then compared with the measured form factor (at low [DNA]) to validate the model. To compute the theoretical form factor of DNA in solution, solvent effects must also be included. The excluded volume prescribed by CRY SOL [5] is assigned to each atom of DNA and counterions for the computation. An extra volume of hydration around DNA is also predicted according to ref [5]. Therefore, hundreds of dummy hydration shell atoms are placed randomly into the solvent accessible region around the DNA molecule. Once the concentrations and coordinates of all DNA atoms, counter-ions and dummy hydration shell atoms are available, the form factor can be computed

using the Debye formula. This computation is repeated 30-40 times to randomize the placement of hydration shell atoms. These results are averaged to generate the final scattering profile. The theoretical form factors calculated based on nonlinear Poisson Boltzmann approach are plotted together with the experimental form factor in figure 2.4. With properly chosen parameters of ion radius r and dielectric constant d , a good match is obtained. A detailed description of how to use the software which computes these curves is provided in the appendix.

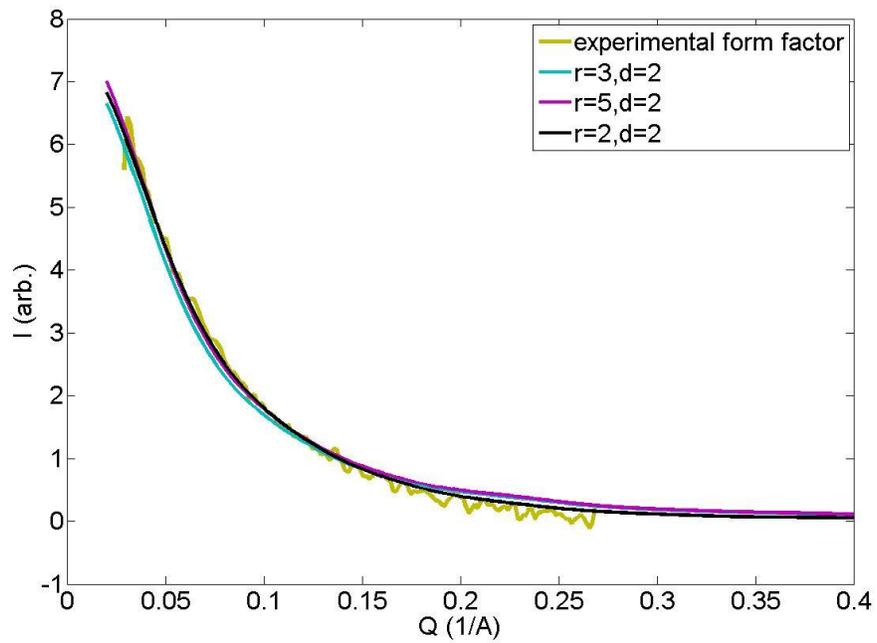


Figure 2.4 computed versus experimental determined form factor

Theoretical scattering profiles obtained based on APBS for different ion radius are compared with experimental form factor. All three calculated curves match the experimental form factor well. R=2 seems to give the best match.

2.1.4 Experimental setup

All SAXS experiments that contributed to this research were carried out at C1 or G1 station at Cornell High Energy Synchrotron Source (CHESS). Details of the experimental setup are provided in ref [6, 12, 13]. ASAXS/mE-ASAXS experiments were carried out at the C1 station where X-ray energy is tunable. A sketch of the experimental setup for SAXS (as well as ASAXS and mE-ASAXS) is shown in figure 2.5. The X-ray beam size and position are adjusted by the upstream slits and optimized to avoid parasitic scattering from guard slits that are closest to the sample. The beam size at the sample is ~1.5 mm wide and ~0.7 mm high at C-line (~0.5mm diameter at G-line). For static SAXS measurement, the DNA/RNA sample is contained in an acrylic sample cell, with sample volume 25 ~ 40 μL . This cell is sealed with ultra-thin (~0.7 μm) silicon nitride windows [14] fabricated at Cornell Nanoscale facility (CNF). After the sample, the scattered X-ray beam passes through an evacuated flight tube which is ~1m long and the scattering profile is captured by a home-made CCD camera behind the flight tube. An example of the scattering geometry is shown in figure 2.5 as well. The direct beam downstream is blocked by a motorized beamstop at the back of the flight tube to protect the CCD detector. For all the measurements, radiation damage of the sample caused by X-ray exposure is checked at the very beginning and the optimum exposure time is chosen accordingly to maximize the signal to noise ratio.

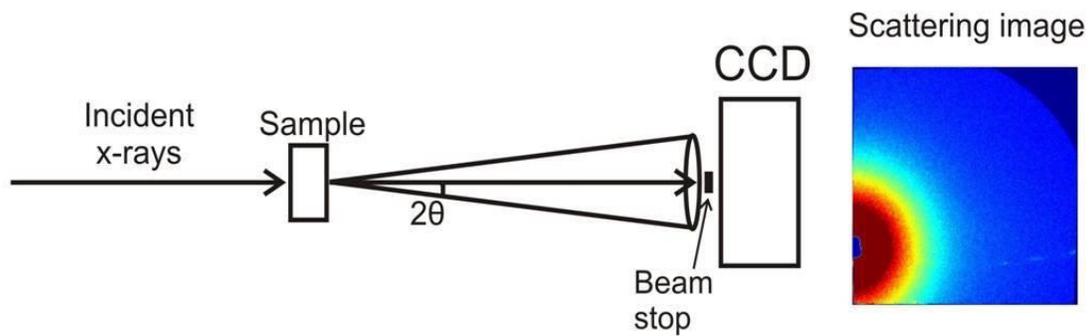


Figure 2.5 a sketch of SAXS setup

X-ray beam is incident and then scattered by the sample. The scattering angle is 2θ . A beam stop is placed right in front of CCD camera. The beam stop protects the CCD, which captures scattering image, from direct incidence of the beam. An example of scattering image is shown on the right side.

2.2 X-ray fluorescence Methods—X-ray Absorption Fine Structure (XAFS)

A brief introduction of X-ray Absorption Fine structure (XAFS) technique [15] is provided in this section. To avoid duplication, please refer to chapter 6 for the experimental setup for XAFS.

2.2.1 Theory

X-ray Absorption Fine structure (XAFS) includes both Extended X-Ray Absorption Fine Structure (EXAFS) and X-ray Absorption Near Edge Structure (XANES). XAFS measures absorption coefficient $\mu(E)$ of a given material as a function of incident beam energy. An x-ray beam with high energy resolution is incident on the sample and both the incident and transmitted X-ray intensity denoted by I_0 and I_T are recorded. The incident X-ray energy is incremented gradually in steps of 1~4eV. The relationship between I_T and I_0 is in a very simple form:

$$I_T = I_0 \exp(-\mu(E) \cdot d)$$

where d is the thickness of the sample. The absorption coefficient $\mu(E)$ is easily obtained from the formula above. If the incident X-ray energy matches the binding energy of electron of the atom of interest (e.g. the X-ray energy is at an absorption edge) within the sample, X-ray absorption by the sample increases sharply. A significant drop of the transmitted X-ray intensity will be measured. On the other hand, when the incident X-ray energy is not close to absorption edge, the absorption coefficient $\mu(E)$ varies with $1/E^3$. A plot of $\mu(E)$ versus E around the absorption edge generates the XAFS spectrum.

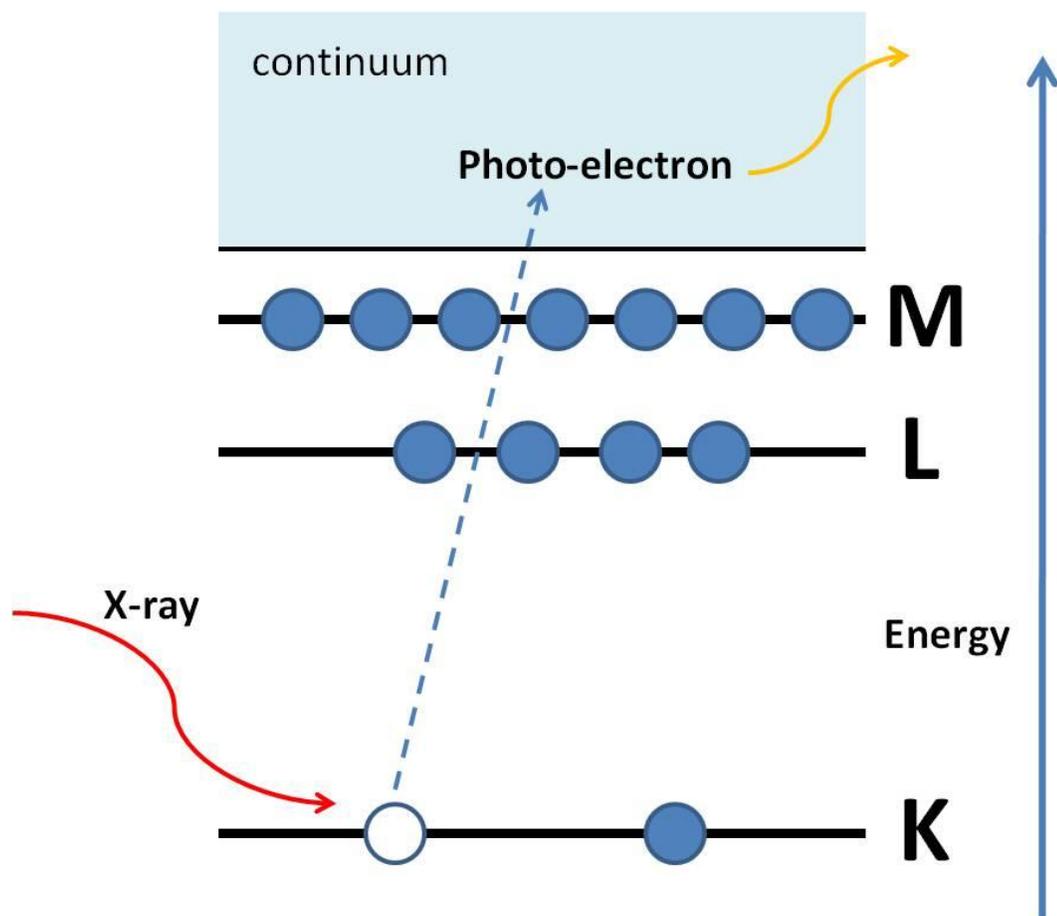


Figure 2.6 principle of XAFS

Atoms (and molecules) can absorb X-ray radiation. At absorption edge, the energy of the photon is used to generate a photoelectron. The presence of these photoelectrons in the material has an effect on the overall pattern of absorption of X-rays as described in the text.

The EXAFS spectrum covers the energy range from 300eV to more than 1000eV above absorption edge while the XANES spectrum is within 300eV of (above) the edge. In the XANES regime, the mean free path of the photoelectron is high, which produces multiple scattering effects. In the EXAFS regime, the mean free path is limited thus single scattering is the major process. The physical process that occurs at an absorption edge is illustrated in figure 2.6 using K-edge absorption as an example. The absorbed photon ejects a core photoelectron from the absorbing atom, leaving behind a core hole. The energy of the ejected photoelectron should be equal to that of the absorbed photon minus the binding energy of the initial core state. The ejected photoelectron, which can be treated as a forward-propagating wave, interacts with electrons in the surrounding non-excited atoms and generates backward-propagating electron waves. The two waves interfere with each other and the interference pattern shows up as a modulation of the measured absorption coefficient causing the oscillation in the XAFS spectra. A simplified plane-wave single-scattering theory has been used for interpretation of XAFS spectra while modern methods demonstrate that curved-wave corrections and multiple-scattering effects cannot be neglected [15 and references within]. An example of XAFS curve is sketched in figure 2.7.

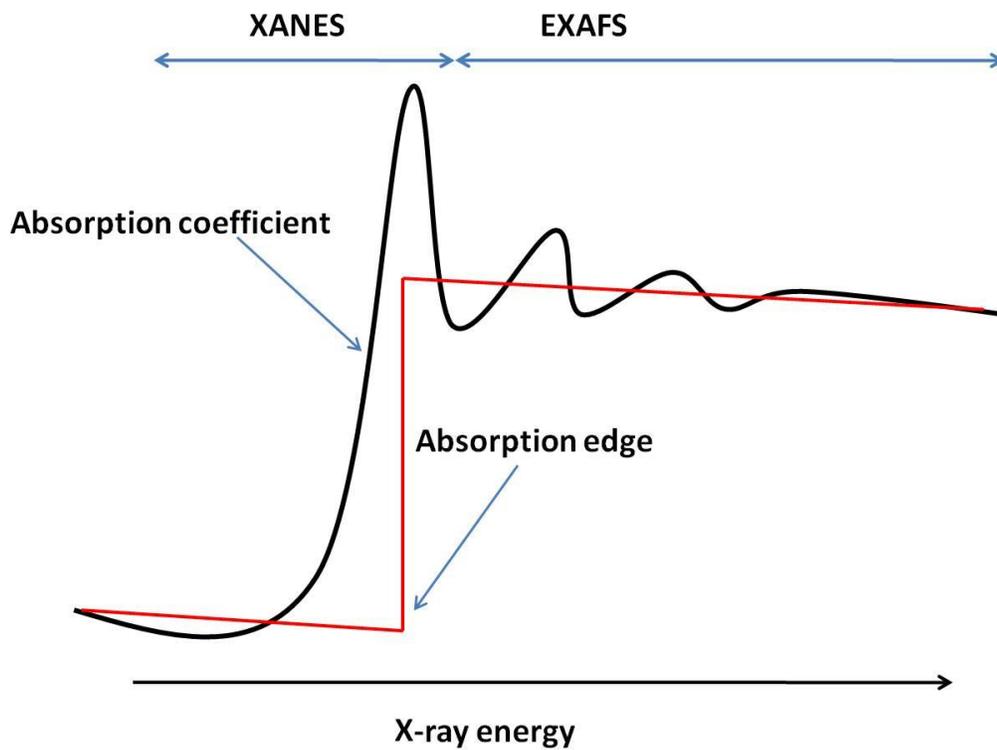


Figure 2.7 a sketch of the XAFS spectrum

The black curve represents the pattern of XAFS spectrum which contains both the XANES and EXAFS regimes. The absorption edge is shown as the strongest peak.

The red lines indicate how the data are normalized as described in text.

2.2.2 Applications

XAFS spectra are especially sensitive to the chemical state, and the distances, coordination number and species of the atoms surrounding the selected element of interest. Therefore, XAFS provides a practical, and relatively simple, way to determine the chemical state and local atomic structure for a selected atomic species. XAFS samples can be gases, solids or liquids since the synchrotron X-ray beam is intense enough to penetrate through the sample. XAFS is routinely used in a wide range of scientific fields, including biology, environmental science, and materials science [15 and references within]. In this research, XAFS is used to measure the ion solvation. The technique is capable of probing the first few solvent shells around the central solute atom [16]. This hydration measurement by XAFS is discussed in chapter 6. The XANES region of XAFS reports information about hydration.

2.3 Conclusion

Experimental X-ray techniques including SAXS, ASAXS, mE-ASAXS, XAFS are summarized in this chapter. These techniques are employed to measure DNA-DNA (or RNA-RNA) interaction, counterion distribution around DNA/RNA and excess number of ions associated with DNA. This chapter also discussed a method for computing the form factor using the Debye formula. Agreement of predicted and measured scattering profiles provides a direct validation of the models that predict the ion distribution around DNA. In this research, Nonlinear Poisson Boltzmann models are tested and a very good agreement between experiment and theory is achieved.

REFERENCES

- [1] O. Glatter, and O. Kratky, *Small Angle X-ray Scattering*. Academic Press, London (1982)
- [2] C. F. Wu and S. H. Chen, *Biopolymers* **27**, 1065 (1988)
- [3] X. Qiu, L. W. Kwok, H. Y. park, J. S. Lamb, K. Andresen, and L. Pollack, *Phys. Rev. Lett.* **96**, 138101 (2006)
- [4] X. Qiu, K. Andresen, L. W. Kwok, J. S. Lamb, H. Y. park, and L. Pollack, *Phys. Rev. Lett.* **99**, 038104 (2007)
- [5] D. Svergun, C. Barberato, and M. H. J. Koch, *J. Appl. Cryst.* **28**, 768 (1995); D. I. Svergun, et al., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2267 (1998)
- [6] K. Andresen, X. Y. Qiu, S. A. Pabit, J. S. Lamb, H. Y. Park, L. W. Kwok, and L. Pollack, *Biophys. J.* **95**, 287 (2008)
- [7] S. A. Pabit, X. Y. Qiu, J. S. Lamb, L. Li, S. P. Meisburger, and L. Pollack, *Nucleic Acids Res.* **37**, 3887 (2009)
- [8] S. A. Pabit, S. P. Meisburger, L. Li, J. M. Blose, C. D. Jones, and L. Pollack, *J. Am. Chem. Soc.* *132*, 16334 (2010)
- [9] G. Evans, and R.F. Pettifer, *J. Appl. Crystallogr.* **34**, 82 (2001)
- [10] D. I. Svergun, *J. Appl. Cryst.* **24**, 485 (1991)
- [11] D. Orthaber, A. Bergmann, and O. Glatter, *J. Appl. Crystallogr.* **33**, 218 (2000)
- [12] S. A. Pabit, K. D. Finkelstein, and L. Pollack, *Methods Enzymol.* **469**, 391 (2009)
- [13] R. Das, T. T. Mills, L. W. Kwok, G. S. Maskel, I. S. Millett, D. Doniach, K. D. Finkelstein, D. Herschlag and L. Pollack, *Phys. Rev. Lett.* **90**, 188103 (2003).
- [14] V. Tereshko, and J. A. Subirana, *Acta Cryst.* **D55**, 810 (1999).
- [15] D. C. Koningsberger and R. Prins, *Xray Absorption: Principles, Applications,*

Techniques of EXAFS, SEXAFS, and XANES, in Chemical Analysis **92**, ed., John Wiley & Sons, 1988

[16] J. L. Fulton, D. M. Pfund, and S. L. Wallen; M. Newville, E. A. Stern, and Y. Ma, J. Chem. Phys. **105**, 2161 (1996)

CHAPTER 3

Theoretical modeling

DNA and RNA helices produce strong electric field near their surfaces that are important for biological functions. It is very challenging to model the electrostatics of DNA (or RNA) because this very high surface charge density can also alter the properties of the associated ions. For example, when multivalent ions are introduced into solutions containing DNA or RNA, correlations between ions can occur, and must be accounted for, further increasing the difficulty of modeling. However, there have been many theoretical methods proposed since early last century, to model the electrostatic interaction between polyelectrolytes mediated by small ions in solution and association between polyelectrolytes and counterions. Some models include counterion condensation theory pioneered by Manning [1], Poisson Boltzmann theory [2 and references within], ion-bridging model [3] and electrostatic zipper model [4, 5]. In this chapter the principles of those models are reviewed.

3.1 Counterion condensation theory

For long, highly-charged, rod-like polyelectrolytes in salt solution, the ions condensed at the surface of the polyelectrolytes lead to a reduction or renormalization of the surface charge. This effect, as well as a description of the more dilute counterion atmosphere at greater distances from the surface, using Debye-Huckel (DH) theory [6], was first pointed out by Manning [1]. This theory assumes that the counterions “condense” onto the polyelectrolytes until the linear surface charge density is reduced below certain critical value. In this model the polyelectrolyte is

assumed to be an idealized line charge with zero radius, infinite length, and uniform charge density. The condensed counterion layer is assumed to be in physical equilibrium with the ionic atmosphere surrounding the polyelectrolyte. The uncondensed mobile ions in the ionic atmosphere are treated within the Debye-Huckel approximation. A sketch illustrating the assumptions of this Manning condensation is shown in figure 3.1.

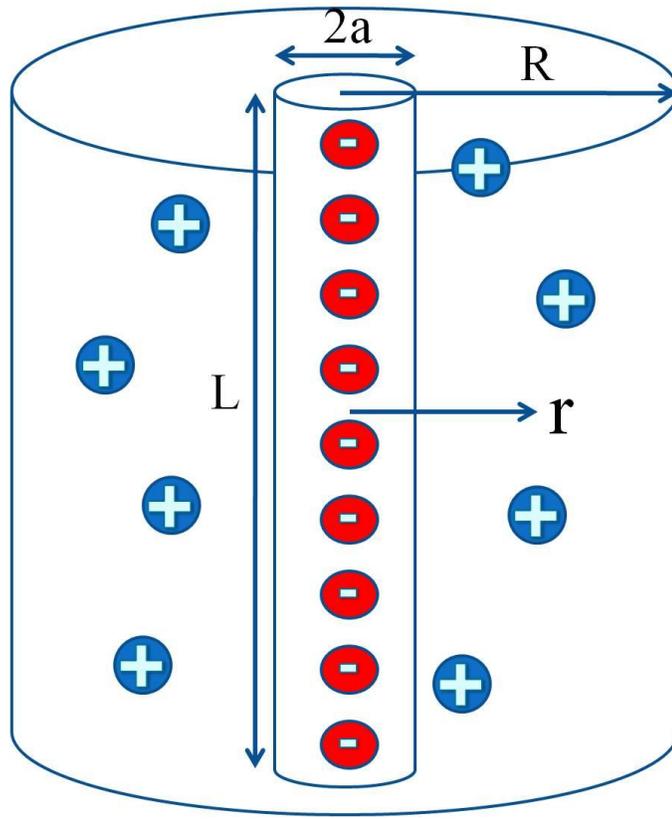


Figure 3.1 a sketch of the basic framework of Manning's condensation theory

L is the length of the polyelectrolyte and assumed to be infinitely long. The radius a of the polyelectrolyte is assumed to be infinitely small. The condensed region is within the cylinder of radius R . Counterions outside of this region are treated by Debye-Huckel theory.

To fully understand this counterion condensation theory, the Debye-Huckel approximation is introduced briefly below. These concepts are also helpful in understanding the Poisson-Boltzmann theory introduced in the next section. In a more general setting, consider an electrolyte solution with dielectric constant ϵ . The molecules with positive charge $+q$ and negative charge $-q$ are treated as spheres with diameter a . The charge density profile is $\rho(r)$ and the electrostatic potential $\phi(r)$ satisfies the Laplace equation $\nabla^2\phi(r) = 0$ in the region $0 < r < a$, and Poisson equation $\nabla^2\phi(r) = -4\pi\rho(r)/\epsilon$ when $r > a$. If the Boltzmann distribution $\rho(r) = \exp[\pm\beta q\phi(r)]$ where $\beta = 1/(k_B T)$ is used, the Poisson equation can be rewritten in the non-linear Poisson Boltzmann form

$$\nabla^2\phi(r) = \frac{4\pi\rho q}{\epsilon} \sinh(\beta q\phi(r))$$

When $\beta q\phi(r) \ll 1$, the equation above is linearized to the Helmholtz equation (the Debye-Huckel approximation) with the inverse of Debye length defined as $\kappa = \sqrt{4\pi q^2 \rho / k_B T \epsilon}$. This equation can be solved analytically and the solution of the electrostatic potential falls exponentially in the regime $r > a$ where nonlinear screening takes effect.

Under the assumption and approximation at the beginning of this section, Manning predicted that counterion condensation is triggered when $z l_b / b > 1$, where z is the valence of the counterion, $l_b = e^2 / \epsilon k_B T$ is called Bjerrum length and b is length per charge (e/b is the backbone charge density). Note that ionic strength is almost always low comparing to the surface charge density of dsDNA/dsRNA. The surface counterion concentration can be calculated as [7]

$$n_s = 2\pi \left(\frac{\sigma}{e}\right)^2 l_b \left(1 - \frac{b}{zl_b}\right)$$

in the excess low salt limit, where $\sigma = e/2\pi ab$. For dsDNA, with $a = 10\text{\AA}$, $b = 1.7\text{\AA}$, $l_b = 7.14\text{\AA}$, the surface counterion concentration can be calculated explicitly using the formula above with ion valence z plugged in. The complete ion distribution profile can also be derived as [7]

$$n(r) = n_s \left(\frac{a}{r}\right)^2 \left[\left(\frac{zl_b}{b} - 1\right) \ln\left(\frac{r}{a}\right) + 1\right]^{-2}$$

Manning's original approach gives a simple picture of counterion condensation. It holds well for monovalent ions but not for divalent or trivalent ions where ion-ion correlations cannot be neglected. Additionally some of its assumptions such as treating dsDNA/dsRNA as infinitely-long cylinder are not quite appropriate. However, though this approach seems to be oversimplified and not rigorous, counterion condensation can be demonstrated using more complicated methods [7, 8, 9].

3.2 Poisson Boltzmann approach

Basic form

The Poisson Boltzmann equation is a nonlinear partial differential equation that describes electrostatic interactions between molecules in ionic solutions. In the field of biology, PB theory has been widely used to analyze fundamental nucleic acid processes [10], RNA folding [11], ligand binding and protein association to nucleic acids [12]. This is the most conceptually simple and popular method used to calculate the electrostatic potential in the electrolyte solution. The PB equation is written as follows

$$\nabla^2 \varphi(r) = -\frac{4\pi\rho(r)}{\varepsilon} - \frac{4\pi e}{\varepsilon} \sum n_i q_i \exp\left(-\frac{e q_i \varphi(r)}{k_B T}\right)$$

where $\rho(r)$ is the density of fixed, external charges (charges on DNA/RNA molecule in our case) and $e q_i$ and n_i are the charge and average number density of electrolyte ions of each kind, respectively. This equation can be solved analytically only in its linearized form for low electrostatic potentials as mentioned in the previous section when discussing the Debye-Huckel approximation. Numerical approaches such as Monte Carlo or Finite Difference can be applied to solve the equation in its nonlinear form with appropriate boundary conditions. The Poisson Boltzmann theory has been successful in predicting ionic profiles close to planar and curved surfaces along with the resulting forces and it works very well for DNA in monovalent salt (Na^+ , K^+ , Rb^+) solution. Adaptive Poisson Boltzmann Solver (APBS) [13], a software package for evaluating electrostatic properties of nanoscale biomolecular systems, was implemented based on Poisson Boltzmann equation (PBE).

Limitations

The Poisson Boltzmann theory is a mean-field theory in which the density of electrolyte ions depends on the mean-field potential $\varphi(r)$ through the Boltzmann distribution. Its application is limited to cases where ion density fluctuations [14] and ion-ion correlations [15, 16] are not significant and thus can be neglected [17]. The long-wavelength fluctuation dominates the high temperature regime while short range correlation usually governs the low temperature regime. It has long been argued that the PB theory should not be used for the calculation of the free energy of a small, spherical ion in an electrolyte solution, which is influenced predominantly by ion-ion correlations. Therefore mean field theory fails to explain the attraction effect between

like-charged objects. In particular, DNA condensation mediated by trivalent ions cannot be simply modeled using PB theory. Two theories that are more suitable for explaining DNA condensation are summarized in 3.3 and 3.4.

Extensions

A successful modification of traditional Poisson Boltzmann theory is called size-modified Poisson Boltzmann theory. One of the biggest limitations of PB theory is that it assumes point-like ions in thermodynamic equilibrium and neglects statistical correlation and steric effect of the ions. As a result, it strongly overestimates the ionic concentrations close to charged surface, which can easily exceed the maximal allowed coverage by orders of magnitude. Many efforts have been made to model size effects in electrolyte solutions. Those theories use a variety of different strategies. One simple and straightforward size-modified Poisson Boltzmann approach was proposed in ref [18]. The effect of steric repulsion is included, by adding a steric correction term to the expression of the entropy and at low ionic concentration the original PB equation is recovered. The modified Poisson Boltzmann for a symmetric z:z electrolyte system is written as

$$\nabla^2 \varphi(r) = \frac{8\pi z e c}{\varepsilon} \cdot \frac{\sinh(z\beta e \varphi)}{1 - \varphi_0 + \varphi_0 \cosh(z\beta e \varphi)}$$

where $\varphi_0 = 2a^3 c$ is the total bulk volume fraction of the positive and negative ions. In the limit of low ionic concentration, $\varphi_0 \rightarrow 0$, the equation above is reduced to the standard PB equation and when $|\beta e \varphi| \ll 1$ (Debye-Huckel limit) it reduces to the linearized PB equation as expected. The key improvement here is in the case of large electrostatic potential $|\beta e \varphi| \gg 1$, the surface ionic concentration is always bounded by $1/a^3$ similar to the case of close packing while it is unbound in the standard PB

framework. This is particularly important when considering the absorption of large ions on charged surfaces. Analytical expressions for the potential and ion concentrations can be obtained from the size-modified equation. Another size-modified PB theory to model DNA with monovalent ion competition was developed by Doniach group at Stanford University, CA [19].

3.3 Electrostatic Zipper Model

This model can be applied to explain the trivalent ion-induced DNA condensation and resistance of dsRNA to condensation by trivalent ions [4, 5]. It has to be first pointed out that divalent ions such as Mg^{2+} , Ca^{2+} do not induce condensation of dsDNA despite their high affinity to the phosphate backbone of DNA. In contrast, trivalent ions such as cobalt compounds, or the polyamine spermidine, are widely used for condensing DNA in the lab. These trivalent ions can form distinct surface charge patterns by binding into DNA grooves [20 and references within]. The surface charge pattern is proposed to determine the specificity and energetics of DNA condensation. As shown in figure 3.2, negatively-charged helical lines of phosphates and positively-charged counterions absorbed in the grooves form stripes of positive and negative charges. Two DNA molecules close to each other can align such that closely opposing stripes have complementary charges along the length of DNA-DNA contact. This creates a “zipper” which pulls the molecules together via electrostatic attraction. This type of interaction is highly dependent on the geometry of the macromolecule. The attraction occurs only if the counterions inside the groove are still “accessible” from the outside. The strength of the attraction depends on the distribution of counterions between two grooves, the helical pitch H and the axial shift z .

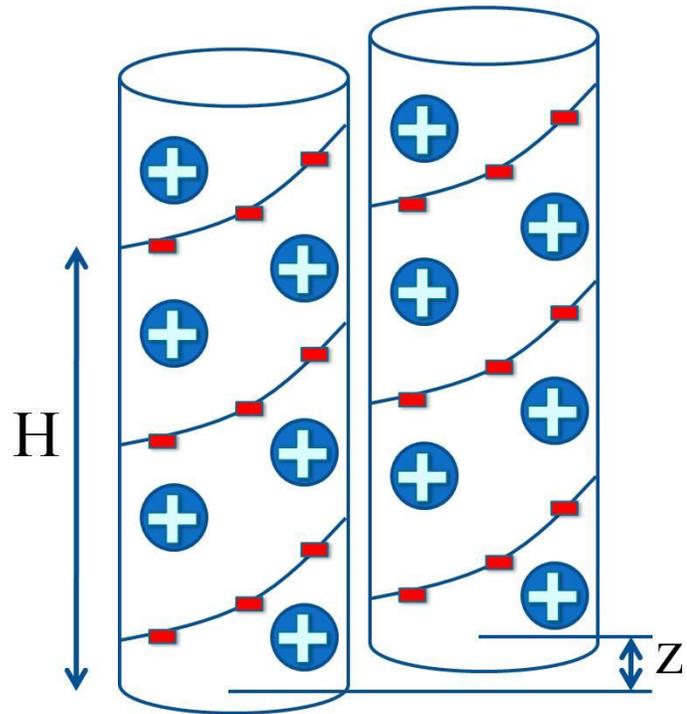


Figure 3.2 electrostatic zipper model

The dsDNA molecules are modeled as cylinders. Two DNA molecules associate with each other facilitated by favorable surface pattern of charges.

Mathematically, in this model, double-stranded DNA is assumed to have a cylindrical shape. The phosphate strands are approximated by two helical lines of charges and absorbed counterions are described by a three-state model, described below. In the theory of counterion condensation, as well as many other models that claim to explain attraction between like-charge objects, it is assumed that all the counterions are freely mobile even when they are close to the highly-charged macromolecule. However, this does not work for the DNA-trivalent ion system. In the electrostatic zipper model the fixed, absorbed and condensed ions were treated separately from those freely-diffusing ions which are modeled using Debye-Huckel theory. Based on the assumptions above, the free energy of interaction can be written explicitly. The results predict that though the charge of DNA molecule is not fully neutralized by counterions bound in the groove, the helices can still attract in the “electrostatic zipper” fashion (see figure 3.2). On the other hand, mathematically, this model proves that counterion associated with phosphate backbone reduces the attraction between DNA molecules due to weaker charge separation, which is consistent with the observation that Mg^{2+} or Ca^{2+} ions do not induce DNA condensation. Experiments on DNA condensation driven by cobalt hexamine or spermine carried out in our group are described in detail in chapter 5, 6, 7.

3.4 Bridges-and-clinches model

This model can also be used to explain the results of DNA (or RNA) condensation experiments discussed in the following chapters. Even though the originally published bridges-and-clinches model deals with DNA condensation by small molecules with specific structure [3], it provides some insights on DNA condensation mechanism in

another perspective.

Two assumptions are made in this model. The first is that there exists stronger binding force between DNA strand and counterions than the association force that Manning condensation theory predicts. More specifically, DNA charge neutralization, which is a prerequisite for DNA condensation, can be induced by multivalent cations bound to the DNA double-helix. While such association is often regarded as the Manning counterion condensation driven by classical Coulomb interactions, other types of interactions often play a major role. In particular, formation of reversible salt and hydrogen bonds could occur between counterions and negatively-charged oxygens on the DNA strand. These bonds can be quite strong, much stronger than what Manning's counterion condensation theory would predict, since the local dielectric constant (on the bond-length scale) is much lower than the bulk macroscopic dielectric constant of water. The strong distance dependence of the dielectric constant defines the short-range character of the DNA-counterion interactions. The second assumption is that the DNA molecule is semi-rigid, e.g. its contour length is longer than its persistent length (~50nm) and much longer than its thickness. Under this assumption DNA strand can be treated as a sequence of nearly straight fragments of length smaller than persistent length but greater than the thickness of the strand.

Based on the two assumptions above, counterions can associate either with one DNA strand (serve as "clinches" inside the groove) or with two DNA strands (serve as "bridges" grabbing two DNA strands together when they are close enough to each other). Those "bridging" counterions are the major driving force of DNA condensation. Note that the negative charge of DNA is neutralized by the "clinching" ions, but condensation requires a fair amount of counterions to "bridge" DNA

molecules. There is competition between serving as clinches and serving as bridges. In other words, charge neutralization and ion condensation are anti-correlated. The association constants of both can be calculated under reasonable assumptions according to ref 3.

3.5 Conclusion

Great efforts have been devoted to investigating inter-DNA interactions mediated by small ions with different valence, ranging from +1 (such as Na^+ , K^+) to +4 (such as spermine). Counterion condensation theory and Poisson Boltzmann theory are two of the most important and widely used theories that model the ion distribution around polyelectrolytes in solution. Both models describe DNA association with low valence counterions but are unsuccessful when considering multivalent (valence >2) counterions. For these more highly charged ions, statistical ion-ion correlations and steric effects cannot be ignored. Electrostatic zipper model and bridges-and-clinches model are introduced in this chapter. These two theories are more suitable for the DNA-trivalent ion system.

REFERENCES

- [1] G. S. Manning, *J. Chem. Phys.* **51**, 924 (1969)
- [2] C. F. Anderson, and M. T. Record, *Ann. Rev. of Biophys. and Biophys. Chem.* **19**, 423 (1990)
- [3] I. A. Nyrkova, and A. N. Semenov, *Soft Matter* **5**, 979 (2009)
- [4] A. A. Kornyshev, and S. Leikin, *Proc. Natl. Acad. Sci.* **95**, 13579 (1998)
- [5] A. A. Kornyshev, and A. Wynveen, *Proc. Natl. Acad. Sci.* **106**, 4683 (2009)
- [6] P. W. Debye, and E. Huckel, *Phys. Z.* **24**, 185 (1923)
- [7] I. Rouzina, and V. A. Bloomfield, *J. Phys. Chem.* **100**, 9977 (1996)
- [8] A. Yu. Grosberg, T. T. Nguyen, and B. I. Shklovskii, *Rev. Mod. Phys.* **74**, 329 (2002)
- [9] B. O'Shaughnessy, and Q. Yang, *Phys. Rev. Lett.* **94**, 048302 (2005)
- [10] N. Korolev, A. P. Lyubartsev, and L. Nordenskiöld, *J. Biomol. Struct. Dyn.* **20**, 275 (2002)
- [11] D. E. Draper, D. Grilley, and A. M. Soto, *Annu. Rev. Biophys. Biomol. Struct.* **34**, 221 (2005)
- [12] S. W. Chen, and B. Honig, *J. Phys. Chem. B.* **101**, 9113 (1997)
- [13] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, *Proc. Natl. Acad. Sci.* **98**, 10037 (2001)
- [14] F. Oosawa, *Biopolymers* **6**, 1633 (1968)
- [15] B. -Y. Ha, and A. J. Liu, *Phys. Rev. Lett.* **79**, 1289 (1997)
- [16] N. Gronbech-Jensen, R. J. Mashl, R. F. Bruinsma, and W. M. Gelbart, *Phys. Rev. Lett.* **78**, 2477 (1997)
- [17] J. G. Kirkwood, *J. Chem. Phys.* **2**, 767 (1934)

- [18] I. Borukhov, D. Andelman, and H. Orland, *Phys. Rev. Lett.* **79**, 435 (1997)
- [19] V. B. Chu, Y. Bai, J. Lipfert, D. Herschlag, and S. Doniach, *Biophys. J.* **93**, 3202 (2007)
- [20] X. Shui, L. McFail-Isom, G. G. Hu, and L. D. Williams, *Biochemistry* **37**, 8341 (1998)

CHAPTER 4

End-to-end stacking measurement

This chapter is published.

L. Li, S. A. Pabit, J. S. Lamb, H. Y. Park, and L. Pollack, *Appl. Phys. Lett.* **92**, 223901 (2008).

Abstract

Recent experiments suggest that short DNA strands associate by end-to-end stacking. Here, we report interactions between DNAs with modified ends. DNA duplexes, 20 bp long, were capped with short T₄ loops at 2, 1 or 0 ends, and were placed in solutions containing 20mM Mg²⁺. Association was observed only in constructs with non-capped ends. DNA-DNA interactions were characterized by measuring variations in small angle X-ray scattering (SAXS) curves at the lowest scattering angles. Second virial coefficients were computed from the SAXS data. Our results confirm that end-to-end stacking plays an important role in short strand DNA-DNA interactions.

Introduction

Biological helices, such as nucleic acid duplexes or α -helical segments of proteins, self assemble into precisely designed structures that regulate life. Since the phosphate backbones of DNA and RNA are highly negatively charged, charge compensation must be provided by counterions, ranging from small cations like K⁺ or Mg²⁺, to basic polyamines or proteins. Counterions play a critical role in modulating interhelical interactions, providing electrostatic screening for these highly charged

polymers, and even facilitating the attraction of like charged strands under certain ionic conditions [1,2,3]. A recent review [4] traces the evolution of theories of counterion localization around charged cylinders, beginning with the pioneering mean field theory of Derjaguin, Landau, Verwey, and Overbeek (DLVO) and leading to sophisticated computational models [Ref. 4 and references within]. Most of these models take into account the breakdown of mean field theories in the presence of the large charge density ($\sim 2e/3.4\text{\AA}$) or high surface electrostatic potential of DNA molecules. Different origins for attractive forces are predicted, ranging from hydration through ionic correlations.

In an effort to provide experimental data for comparison to the numerous and often conflicting theories, we have undertaken studies of interactions between short DNA strands as a function of valence and concentration of ions in solution [3,5]. The use of short, rigid (far less than a persistence length) helices enables comparisons with models that account for the atomically detailed structure of DNA. In the past, such studies have provided new information about the distribution of mono-, di- or tri-valent ions around DNA strands. Additional experimental studies of the small angle scattering of solutions containing short duplexes, revealed an unexpected inter-helical attraction in solutions containing more than a threshold level of divalent counterions [5]. Consideration of end effects [6-9] and previous study of B-DNA crystal formation [10] led us to conjecture that short DNA helices were able to ‘stack up’ in an end-to-end configuration, though the exact mechanism was not revealed. The notion of favorable end-to-end stacking [3] of short DNA strands, facilitated by base stacking of hydrophobic ends, was also highlighted in a recent publication [11], suggesting an intriguing biological origin for these forces. Computation of the magnitude of the

base stacking energies validated the proposal of end-to-end stacking. Here, we describe experimental evidence of end-to-end stacking of short DNA strands, obtained by measurement on DNA helices that are ‘capped’ at one or both ends to partially or completely block end-to-end association.

Experimental techniques

The strength of forces between DNA strands was assessed from small angle X-ray scattering (SAXS) profiles of solutions containing DNA. All measurements were carried out at the C1 station of the Cornell High Energy Synchrotron Source (CHESS) using the experimental setup described in Ref. 12. The measured scattering intensity $I(Q)$ (in which $Q = 4\pi \sin\theta / \lambda$, 2θ is the scattering angle and λ is the x-ray wavelength) is the product of a form factor $P(Q)$, reflecting the electron density distribution within each scattering element (each DNA strand) and an inter-molecule structure factor $S(Q)$, which reports on interactions between DNAs. The form factor for each construct was extracted from measurements carried out in very dilute solutions where the intermolecular interactions are negligible, as explained in Ref. 3. Analysis of DNA-DNA interactions using SAXS was discussed extensively in Ref. 5.

Sample preparation

The presence of end-to-end stacking of short DNAs was tested by studying DNA-DNA interactions in solutions containing 20 base-pair long DNAs (~ 76 Å), terminated with T₄ loops at 2, 1, or 0 ends, DNA “dumbbells”, DNA “semi-dumbbells” or short dsDNA with unmodified ends (figure 4.1). All single-strand DNA oligomers for this experiment were purchased from Integrated DNA Technologies (IDT), Coralville, IA.

The “dumbbell” DNA with 20bp stem “capped” at both ends was constructed [13-15] by base-pairing two DNA hairpins at equal molar amounts with T₄ loop at one end of the stem and a 4-nt complementary overhang at the other end. Double-strand 20bp DNA and 20bp duplex DNA with T₄ loop at only one end were used as control. Standard annealing procedures provided by IDT were applied. Each DNA sample was hydrated and dialyzed against 20mM MgCl₂ solution buffered at 1mM NaMOPS pH 7.



Figure 4.1 (a) "Dumbbell" DNA, consisting of a 20bp duplex capped by T₄ loop at both ends, was constructed by inter-molecular base pairing of two hairpins. (b) "Semi-dumbbell" DNA, consisting of a 20bp duplex capped by a T₄ loop at one end, was constructed by intra-molecular base pairing of a 44 nt long single-strand of DNA. (c) 20bp double-stranded DNA was constructed by annealing together two complementary strands.

Results

Shape differences at the lowest angles of the scattering profiles are easily assessed by matching the curve amplitudes in the high Q regime [5] where scattering is nearly identical for all constructs. The presence of the terminal loops leads to small differences between scattering profiles for the constructs. However, direct comparison of the scattering profiles acquired at the lowest DNA concentration of 0.1 mM shown in figure 4.2, which correspond to molecular form factors, indicates that the differences are nearly negligible. As the DNA concentration increases a clear difference between scattering from “dumbbells”, “semi-dumbbells” and “double-stranded” DNA is measured in the low Q regime, characteristic of S(Q) measurements (figure 4.2). The strength of inter-molecular interactions can be quantified by extracting the second virial coefficient (A_2) of each DNA system. Based on the equation below [16],

$$\frac{P(Q = 0)}{I(c, Q = 0)} = \frac{1}{S(c, Q = 0)} = 1 + 2MA_2c$$

where c represents the concentration of DNA in unit of g/ml and M is the molecular weight of DNA respectively, A_2 can be obtained by linear fit once the structure factor profile S(Q) is extracted from the experimental data. As described above, the scattering profiles collected at the lowest DNA concentration, 0.1mM, were used as form factors P(Q). S(c, Q=0) was calculated by extrapolation of the low Q region of S(Q). The linear fit of $\frac{1}{S(c, Q=0)}$ versus c (figure 4.3) yields the A_2 values for “dumbbell”, “semi-dumbbell” and “double-strand” DNA solutions: 1.4e-4, -2.2e-4, -5.9e-4 (mol ml g⁻²) respectively. The latter two, negative values of A_2 suggest that the “semi-dumbbell” and “double-strand” exhibit weak attraction at 20mM Mg²⁺,

consistent with the previous observation that the low Q upturn occurs when $[\text{Mg}^{2+}] > 10\text{mM}$ [5]. The smaller negative (but larger absolute) value of A_2 of dsDNA solution reflects the fact that the attraction between dsDNA molecules is stronger than that between “semi-dumbbell” DNA molecules under the same ionic conditions. Interestingly, the slightly positive A_2 indicating marginal repulsion is observed in solutions of “dumbbell” DNAs, in which the end-to-end stacking effect is expected to be suppressed by the T_4 loops capping both ends of DNA. To summarize our experimental data, both the magnitude and the sign of the attraction vary for the different constructs. The long range correlations in three types of DNA model systems are distinct presumably due to molecular structural differences.

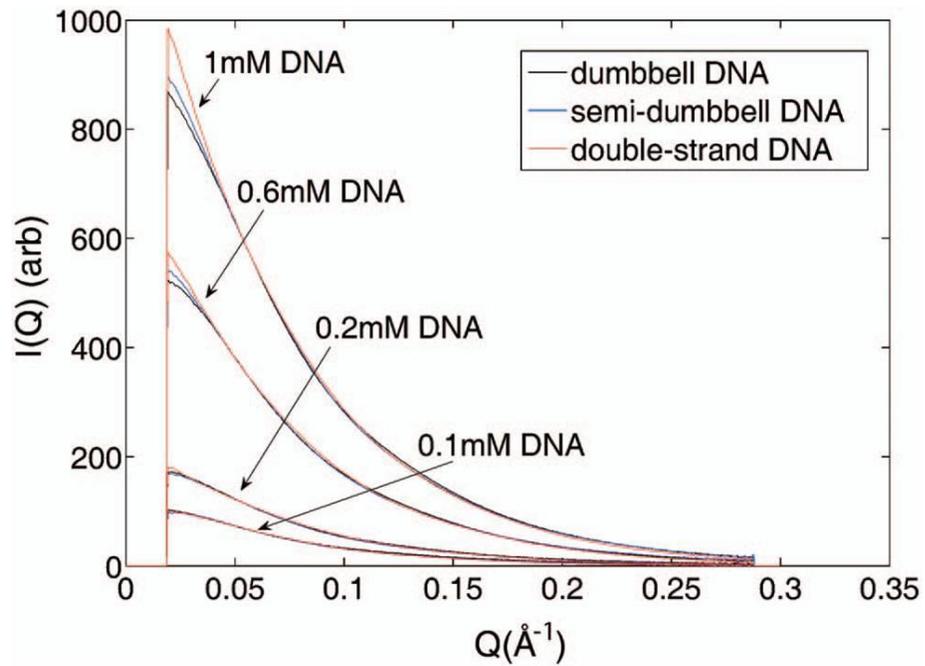


Figure 4.2 Scattering profiles $I(Q)$ were measured for each sample at four different DNA concentrations: 0.1mM, 0.2mM, 0.6mM, and 1.0mM. For each concentration, curves were normalized in the high Q regime to highlight differences in curve shape in the low Q region (see text).

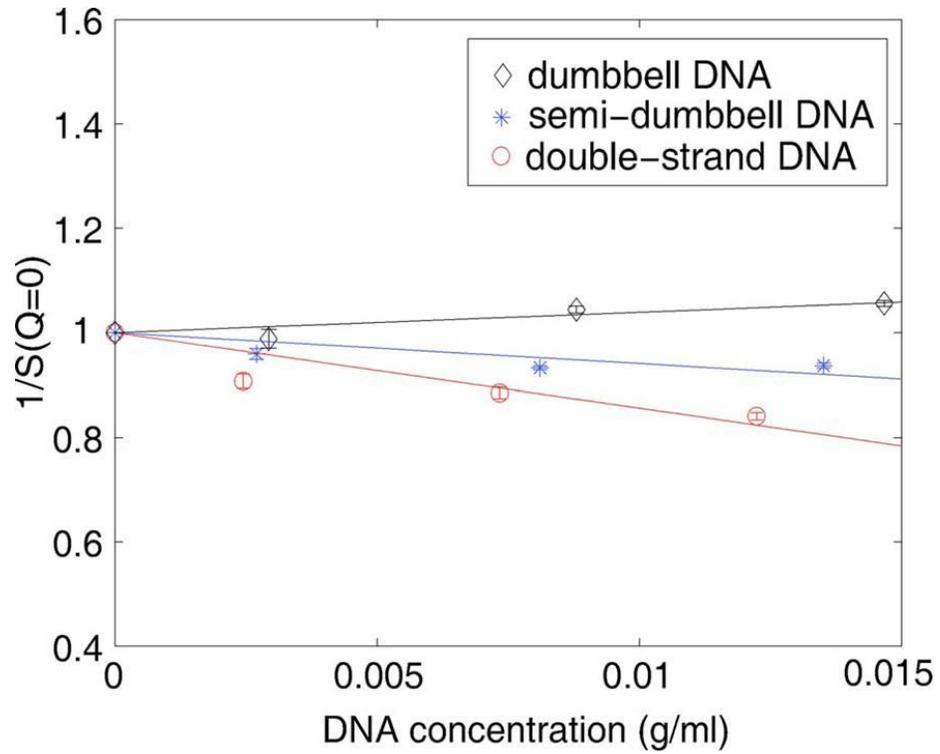


Figure 4.3 The inverse of the structure factor $S(Q=0)$ is plotted as a function of DNA concentration (g/ml). A linear fit was performed to derive the second virial coefficient from these data. “Dumbbell” DNA shows marginal repulsion ($A_2 > 0$) while weak attraction ($A_2 < 0$) is observed in “semi-dumbbell” and double-strand DNA.

Discussion

It is more accurate to interpret the result of A_2 analysis by exploring the definition and physical meaning of the second virial coefficient. A_2 is defined [16] as

$$A_2 \equiv \frac{N_a}{2M^2} \int_V \left[1 - e^{-\frac{u(\vec{r})}{k_B T}} \right] d^3\vec{r}$$

where N_a is Avogadro's number, M is the molecular weight, V represents the 3-D space in which the integral is carried out, and $u(\vec{r})$ is the intermolecular potential, which depends on the relative position of two molecules. It is evident from the above expression that, at a fixed temperature T , the only factor affecting the sign of A_2 is the intermolecular potential. If $u(\vec{r})$ is, on balance, more positive than negative over all space, the factor $1 - e^{-\frac{u(\vec{r})}{k_B T}}$ also exhibits more positive values and A_2 tends to be positive. A more positive $u(\vec{r})$ means two molecules will experience stronger repulsion as the interparticle distance decreases below the equilibrium distance which minimizes the intermolecular potential. A similar argument can be made for the opposite case of attraction. Interpretation of both the sign and magnitude of A_2 is a simple way to quantify and compare intermolecular potentials. Our results suggest that, even in the presence of 20 mM Mg^{2+} , the interactions between “dumbbell” DNAs are slightly repulsive, while interactions between dsDNAs at the same concentration are attractive. Modeling of the potential using APBS [17] does not indicate a buildup of charge around the T_4 loop; we therefore propose that the observed attraction results from the favorable free energy provided by the interaction between hydrophobic ends. These base pairs are exposed only in uncapped constructs; the T_4 loop blocks this type of contact between adjacent strands. This result is consistent with measurements on “semi-dumbbell” DNA, which indicate weaker attraction (smaller A_2) than dsDNA,

due to the presence of one cap. It is important to point out that at (fixed) room temperature it is the potential profile that influences the sign and magnitude of A_2 . The molecular weight also plays a role in determining the magnitude of A_2 but it is much less significant in our case.

Conclusion

In conclusion, we have investigated the behavior of “dumbbell”, “semi-dumbbell” and “double-strand” 20bp DNA molecules mediated by divalent counterions through SAXS experiments. Measurements of the second virial coefficient show distinct interaction modes among these three DNA model systems. Both the structure of the DNA molecule and the electrostatic screening of Mg^{2+} result in modification of the intermolecular potential, which along with temperature, sets the value of A_2 . Our study also highlights the importance of end-to-end base stacking in short strand DNA-DNA interaction, while lateral (side by side) attraction is not ruled out, it may in fact depend on achieving a critical strand length [11].

We thank K. Finkelstein for experimental assistance and D. Burke for suggesting the “dumbbell” construct. This research is funded by the NSF through MCB-0347220 and the Cornell NBTC, and by NIH through P01-GM066275. CHESS is supported by NSF and NIH/NIGMS under Grant No. DMR-9713424.

REFERENCES

- [1] M. H. J. Koch, Z. Sayers, P. Sicre, and D. Svergun, *Macromolecules* **28**, 4904 (1995).
- [2] V. A. Bloomfield, *Biopolymers* **44**, 269 (1997).
- [3] X. Qiu, K. Andresen, L. W. Kwok, J. S. Lamb, H. Y. Park, and L. Pollack, *Phys. Rev. Lett.* **99**, 038104 (2007).
- [4] A. A. Kornyshev, D. J. Lee, S. Leikin, and A. Wynveen, *Rev. Mod. Phys.* **79**, 943 (2007).
- [5] X. Qiu, L. W. Kwok, H. Y. Park, J. S. Lamb, K. Andresen, and L. Pollack, *Phys. Rev. Lett* **96**, 138101 (2006).
- [6] S. A. Allison, *J. Phys. Chem.* **98**, 12091 (1994).
- [7] M. C. Olmsted, C. F. Anderson, and M. T. Record, Jr., *Proc. Natl. Acad. Sci.* **86**, 7766 (1989).
- [8] A. E. Walter, D. H. Turner, J. Kim, M. H. Lyttle, P. Muller, D. H. Mathews, and M. Zuker, *Proc. Natl. Acad. Sci.* **91**, 9218 (1994).
- [9] D. Genest, K. Mazeau and M. Ptak, *J. of Biomol. Struct. Dyn.* **5**, 67 (1987) .
- [10] V. Tereshko, and J. A. Subirana, *Acta Cryst.* **D55**, 810 (1999).
- [11] M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, N. A. Clark, *Science* **318**, 1276 (2007).
- [12] R. Das, T. T. Mills, L. W. Kwok, G. S. Maskel, I. S. Millett, D. Doniach, K. D. Finkelstein, D. Herschlag and L. Pollack, *Phys. Rev. Lett.* **90**, 188103 (2003).
- [13] D. E. Wemmer, and A. S. Benight, *Nucl. Acids Res.* **13**, 8611 (1985).
- [14] R. Owczarzy, P. M. Vallone, R. F. Goldstein, A. S. Benight, *Biopolymers* **52**, 29 (1999).

- [15] M. J. Doktycz, R. F. Goldstein, T. M. Paner, F. J. Gallo, and A. S. Benight, *Biopolymers* **32**, 849 (1992).
- [16] F. Bonnete, and D. Vivares, *Acta Cryst.* **D58**, 1571 (2002).
- [17] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037 (2001).

CHAPTER 5

Double-stranded RNA resists condensation

This chapter is published.

L. Li, S. Pabit, S. Meisburger, and L. Pollack, *Phys. Rev. Lett.* 106, 108101 (2011)

Abstract

Much attention has focused on DNA condensation because of its fundamental biological importance. The recent discovery of new roles for RNA duplexes demands efficient packaging of dsRNA for therapeutics. Here we report results of UV spectroscopic and SAXS studies of short DNA and RNA duplexes in the presence of trivalent ions. Under conditions where UV studies find significant condensation of DNA duplexes into (insoluble) precipitates, RNA duplexes remain soluble. Although complementary SAXS experiments clearly show multivalent ion association to both RNA and DNA, we conclude that the differing surface topologies of RNA and DNA may be crucial in generating the attractive forces that result in precipitation.

Introduction

The attraction of like-charged objects is an important theme in polymer physics, biology and biotechnology. It is remarkable that, despite its uniform, large negative charge, double stranded DNA precipitates from dilute solution when even small numbers of multivalent ions are introduced [1]. Much effort has been expended investigating the nature of this multivalent ion-induced attraction because of its relevance to DNA packaging, either in viruses [2] or for applications in non-viral gene delivery [3]. Here, we extend these studies to duplex RNA. Recent attention has

focused on dsRNA because of its role in RNA interference (RNAi) [4]. In this process, low quantities of short RNA duplexes set into motion a molecular machine that exerts powerful control over gene expression. The nucleotide sequence encoded by the short duplex is used to target and destroy mRNA containing a complementary sequence. RNAi is an ideal vehicle for novel therapeutic applications by targeting and silencing specific genes. The ability to tightly package (condense) numerous, short RNA duplexes is an important prerequisite for optimal design of these next-generation therapeutics [5].

There is no universally accepted explanation of the physical origin of like-charge attraction. Because mean field theories, such as those based on the Poisson Boltzmann equation, do not predict attractive forces, there is intense theoretical interest in developing more sophisticated models. Numerous mechanisms (see recent reviews [6] and references within) have been proposed to explain ion-induced attraction of dsDNA strands, including counter-ion correlations, models that account for ion-bridging, attraction resulting from hydration forces, or from precisely coordinated patterns of charge distributions which form an electrostatic zipper.

With the sudden interest in short dsRNA, we have undertaken a biophysical study comparing DNA condensation to RNA condensation. Using ionic conditions that cause short DNA helices to aggregate, we have attempted to identify similar phases in RNA. Surprisingly, **dsRNA resists condensation**. The strikingly different behavior of these identically charged systems provides important clues about the physics of like charge attraction.

Although complementary strands of RNA or DNA easily combine to form stable, double helices, chemical differences between RNA and DNA drive RNA helices into

the A-form, while DNA helices assume the B-form. The former is shorter and wider, with a deep major groove that is of the order of the radius of the helix. By comparison, the B-form helix is more cylindrical, and the depth of the major and minor grooves is more uniform. Finally, the linear charge density of the A-form helix is 2 e per 2.8 Å, while that of the B-form helix is 2 e per 3.4 Å. Previous work shows that monovalent and divalent counterions penetrate into the major grooves of the RNA, more fully compensating the overall negative charge of the molecule [7]. As a result, screening of the large negative charge of the RNA duplex occurs at lower bulk ionic strength than for DNA. If attraction were solely determined by the degree of charge compensation, one might expect RNA to condense more readily than DNA; we observe the opposite.

Sample preparation

The small trivalent ion Cobalt hexammine (Co-hex, $\text{Co}(\text{NH}_3)_6^{3+}$) is one of the most powerful condensing agents of DNA [8, 9]. This ion has radius of 3 Å and a nearly spherically distributed surface charge. Even small quantities can precipitate dsDNA from dilute solution at room temperature. Co-hex has also been used in RNA folding studies [10].

Here, we examine the role of trivalent Co-hex ions in RNA and DNA charge screening efficiency and condensation. We use two established experimental techniques—UV absorption and Small Angle X-ray Scattering (SAXS)—to probe the condensation of nucleic acid duplexes from dilute solution, driven by the addition of small quantities of Co-hex. Identical experiments were carried out on both DNA and RNA duplexes.

To measure the condensing power of Co-hex, 25bp DNA and RNA were dialyzed against buffer containing 20mM NaCl and 1mM NaMOPS at pH 7, respectively. This low monovalent salt background ensures that added Co-hex will be maximally effective in condensing the sample [8]. Calibrated amounts of Co-hex, between 0 ~ 6mM in solution, were added to each tube, and the mixed solutions were stored at 4°C for 2 hours. Subsequently, each tube was centrifuged at 10,000 rpm (~8000 g) for 10min e.g. [8]. The supernatant was collected and absorbance at 260 nm was measured for each sample in a UV spectrophotometer (Cary 50 Bio, Varian, Inc., Walnut Creek, CA). The fraction of precipitated nucleic acid can be calculated by direct measurement of the change in UV absorption of the supernatant. Since the A-helical form is shorter than the B-helical form, we compared 25bp RNA with both 25 and 16bp DNA, to control for any length-dependent effects of condensation.

Synchrotron small angle x-ray scattering (SAXS) reports the strength of intermolecular interactions. The presence of either repulsive or attractive forces between particles results in distinctive modulation of the scattering profiles at the lowest angles [11]. These experiments were carried out at C1 station at Cornell High Energy Synchrotron Source (CHESS). The experimental setup is described in Ref. [12]. Single strand DNA and RNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA) and Dharmacon Inc. (Chicago, IL), respectively. DNA samples for SAXS studies were prepared as described in Refs. [11, 13] using equilibrium dialysis to establish a fixed bulk ion concentration. All SAXS buffers contain 100 mM NaCl in addition to varying amounts of Co-hex. Identical protocols were employed in preparing RNA samples. Because the competitive association of Co-hex to DNA is a strong function of NaCl concentration [8], the increased

concentration of monovalent ions relative to the UV studies allows measurements of soluble DNA over a broader range of Co-hex concentrations, and enables a more detailed comparison of Co-hex interactions with DNA as opposed to RNA.

Results and discussion

UV spectroscopy

Figure 5.1 shows results of UV absorption measurements of RNA and DNA as a function of (added) Co-hex. These curves clearly indicate that dsRNA precipitation by Co-hex is much less favorable than dsDNA precipitation. For trivalent ion concentrations below 1 mM, almost no RNA precipitates, in contrast to DNA. In the presence of 4mM Co-hex, 85% of 25bp DNA molecules condense while ~80% of RNA molecules remain in the supernatant. Only for Co-hex concentrations in excess of 10 mM do we measure significant RNA condensation.

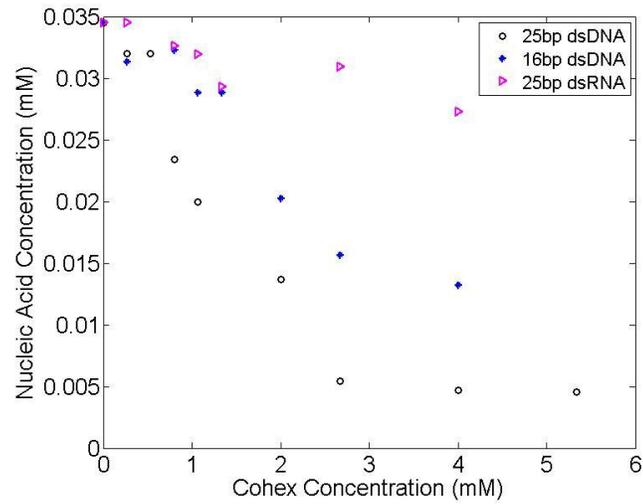


Figure 5.1 The concentration of DNA and RNA molecules in the supernatant as a function of [Co-hex] is calculated from UV absorption. Short, 16 and 25 bp DNA molecules are more easily condensed than 25 bp RNA. The 16 bp DNA, used as a control, indicates that the changing length of the double stranded nucleic acid has a smaller effect in generating condensation than the type of nucleic acid used.

SAXS

One possible explanation of these data is that many fewer Co-hex ions bind to RNA than to DNA: the negative charge of RNA is not effectively screened by these counterions. To test for this eventuality, SAXS was used to measure inter-duplex interactions [11]. SAXS profiles of solutions containing DNA and RNA were measured under carefully controlled ionic conditions, beginning at 100 mM Na⁺, where weak repulsion between nucleic acids is measured [7]. For these studies, the duplex concentration is maintained at 0.6 mM, more than an order of magnitude below the regime where liquid crystalline behavior is expected e.g. [14]. Since Co-hex is a trivalent ion, it very effectively competes with monovalent Na [12]. Its strong electrostatic attraction to the DNA enhances localized screening of the duplex charge [15] and dramatically reduces electrostatic repulsion between neighboring duplexes. Because Co-hex condenses DNA so efficiently, the higher NaCl concentration present in the SAXS studies ensures that samples can be prepared without aggregation. SAXS studies carried out on mixed phase samples, containing both soluble DNA and aggregates, are difficult to interpret [13]. In 100 mM NaCl, the onset of aggregation occurs in DNA with about 1 mM (free) Co-hex [12]. Figure 5.2 shows the results of SAXS studies on DNA (Fig. 5.2a) and RNA (Fig 5.2b) as a function of free Co-hex concentration [12]. To evaluate the magnitude of inter-particle interference, the high angle (or Q) regions of scattering profiles are matched with the form factor [11, 16], which represents the scattering of an isolated duplex. Repulsion/attraction between particles is indicated by a decrease/increase of the scattering profile relative to the form factor at the lowest angle [11].

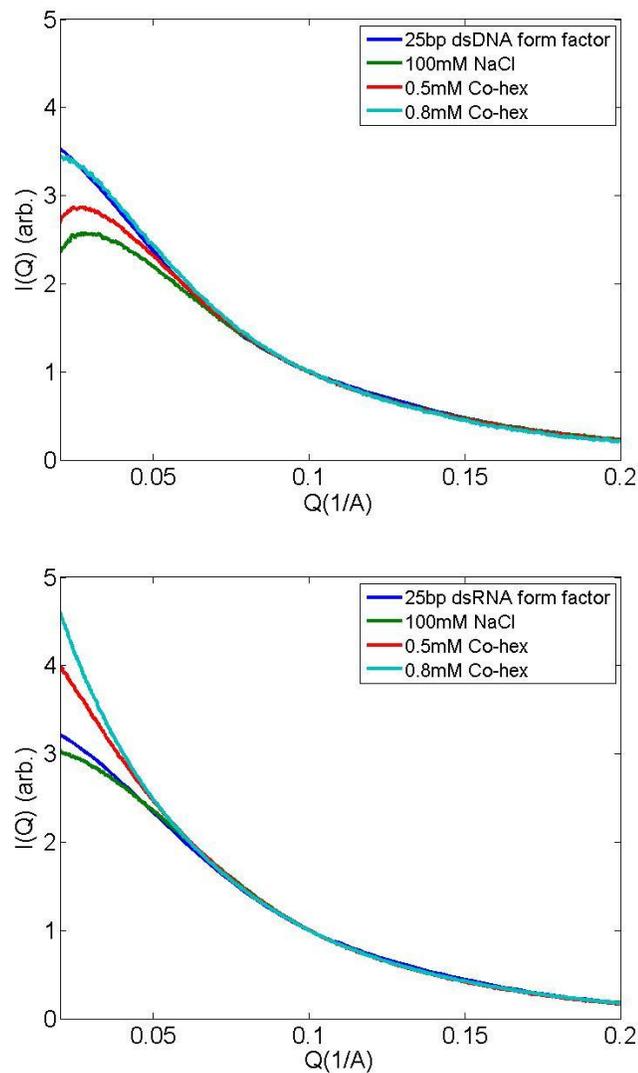


Figure 5.2 To assess inter-particle interactions, solution SAXS profiles of (a) DNA and (b) RNA samples in Cobalt hexamine are compared with the form factor, the scattering from non-interacting molecules (at infinite dilution). Only marginal repulsion between DNA molecules is measured when $[\text{Co-hex}] \sim 0.8\text{mM}$. Comparatively, the sharp and smooth low Q upturn in (b) can be interpreted as end-to-end stacking of RNA molecules when Co-hex is present at concentrations above 0.5mM (see supplementary material, Ref. [7] for a detailed discussion of this effect).

In the absence of Co-hex (100 mM NaCl), SAXS profiles of both nucleic acids indicate clear repulsion, consistent with previous work [7, 11]. The diverging behavior of DNA (Fig 5.2a) and RNA (Fig 5.2b) becomes apparent when small amounts of Co-hex are introduced. Scattering profiles for DNA in 100 mM NaCl plus 0.5 mM Co-hex still decrease at the lowest angles, consistent with repulsive forces. Around 0.8 mM Co-hex electrostatic interactions between duplexes are nearly neutralized. (Small amounts of Co-hex have dramatic effects on DNA. Near ~ 0.8 mM Co-hex, inter-DNA interactions rapidly change from repulsive to attractive. One concern might be potential changes to the amplitude of the scattering resulting from the closer association of co-hex (replacing Na) in the ion cloud. The effects of ion cloud scattering are almost always dominated by the hydration shell of the nucleic acid. The replacement of a few Na ions with several co-hex ions, will lead to a negligible change in scattering amplitude, compared to that from the nucleic acid itself and its hydration shell.)

Figure 5.2b shows an identical experimental series, performed with dsRNA instead of dsDNA. As expected, RNA duplexes repel in the presence of 100 mM NaCl. In 100 mM NaCl plus 0.5 mM Co-hex, where repulsion between DNAs is still evident, the SAXS profiles of RNA duplexes display a strong increase in low angle scattering, consistent with end-to-end stacking. To stack, two helices must come into close contact, thus the appearance of end-to-end stacking signals significant reduction in electrostatic repulsion, consistent with our previous observation that ions more effectively screen dsRNA than dsDNA [7, 17]. However, as opposed to precipitates, which are insoluble, these end-to-end stacked molecules remain soluble and are readily detected by solution SAXS. (A more detailed analysis of the energetics can be

carried out [7, 16].) Validation of end-to-end stacking (as opposed to close packing of DNA into hexagonal arrays described in Ref. [13] and its supplemental information) is achieved by comparison of scattering profiles of Figure 5.2 to computations of pairs of RNA duplexes arranged either end-to-end, or in a side-by-side manner that would be more consistent with DNA packing in hexagonal arrays (Figure 5.3). End-to-end stacking of short nucleic acid duplexes has also been observed in work carried out by others [14, 17]. The seemingly conflicting results of UV absorption and SAXS measurement lead to the well-defined fundamental question—if, as the SAXS data suggest, RNA's charge is more effectively screened than DNA's, why is DNA more susceptible to precipitation by Co-hex than RNA? To answer this question, we must consider how the ions bind to the nucleic acid.

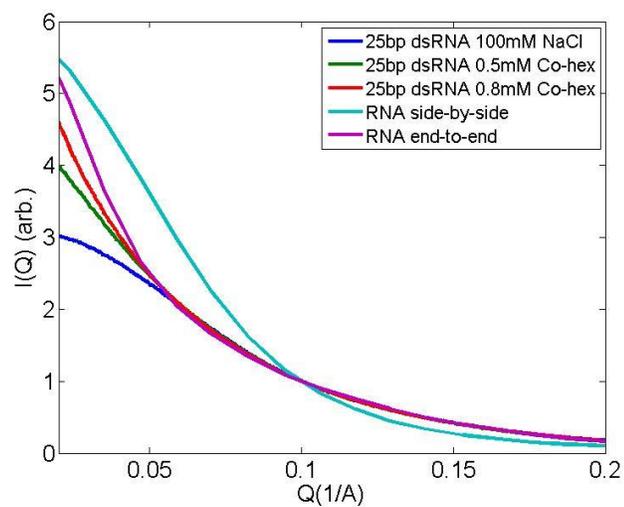


Figure 5.3 Solution SAXS data of RNA samples in NaCl or Co-hex are displayed along with scattering profiles that simulate RNA duplex association in different modes. The purple and cyan curves represent the configurations of RNAs, stacked either end-to-end or placed side-by-side to mimic relative placement in close hexagonal arrays, respectively. The end-to-end stacking model is in better agreement with experimental SAXS profiles, though clearly not all duplexes participate in stacking interactions.

Models

Computations of the potential around DNA or RNA duplexes [18] show that the major groove of A-RNA has a higher negative potential than the minor groove while the opposite is true for B-DNA. These different potentials have a profound impact on the spatial distribution of ions around RNA or DNA. Comparison of experimentally determined ion-distributions with models (based on the non-linear Poisson-Boltzmann equation, which are valid for monovalent ions) confirms that counterion distributions reflect these differences. Monovalent ions are localized to the RNA major groove, while they are more uniformly distributed around DNA [7]. In contrast to monovalent ions, experimental studies of Co-hex suggest that this trivalent ion prefers to bind in the major groove of both nucleic acids. For B-DNA, support for this binding pattern comes from NMR [19], capillary electrophoresis [20] and x-ray crystallographic [21] studies. Notably, the Guanines in the DNA major groove provide a preferential binding site for Co-hex [19]. Other important factors in determining binding sites may include the observed dehydration of Co-hex ions around DNA [22]. Although fewer studies have focused on Co-hex binding to RNA, solution NMR studies [23] find Co-hex ions buried deep within the major groove of a short stem loop. We therefore propose that the observed differences in condensation arise from the dramatically different geometries of the underlying nucleic acid structures. This picture is consistent with two models for condensation. In the first, competition between inter and intra molecular ion bridging explains why condensation forces (inter molecular bridging [24, 25]) and charge screening efficiency (intra molecular binding [25]) should be anti-correlated. Therefore, one possible explanation of the resistance of RNA to condensation arises from the more favorable binding of Co-hex to the RNA

major groove compared to DNA. Second, in the electrostatic zipper model [26], the surface of the nucleic acid presents a pattern of alternating positive and negative charges. This alternating pattern can result in attraction if adjacent molecules pack so that opposing charged surfaces are in contact. There may be a significant difference in RNA⁺ion and DNA⁺ion surface charge due to differences in geometry. For example, the depth of the DNA major groove is $\sim 8 \text{ \AA}$, comparable to the 6 \AA diameter of Co-hex molecule: Co-hex molecules bound in the DNA major groove remain “accessible” from outside. The DNA molecules can be condensed when the surface charge patterns are electrostatically in register with each other. In contrast, although RNA's identical negative charge is also strongly screened, the trivalent ions have the potential to bury themselves too deep within the major groove to be “visible” at the surface.

Either geometric model of nucleic acid association is consistent with results from both absorption and scattering experiments. We note that the hydration structure of DNA and counterions also play an important role in DNA condensation [27], and could lead to a measurable difference in RNA and DNA condensation behavior since the RNA surface is more polar [28] and hence more hydrated, than the DNA surface.

Conclusion

Both SAXS and absorption measurements lead us to propose that the interaction modes of nucleic acids depend on the geometric details of charge arrangement in each system, highlighting the important role of molecular structure in condensation. Under conditions where DNA precipitates readily, the well-buried Co-hex ions inside the major groove of RNA contribute to charge neutralization but ultimately lend

aggregation resistance to RNA duplexes. These readily testable results should provide the basis for further molecular dynamic (MD) simulations of multivalent-ion mediated interactions between like-charged nucleic acids and may provide guidance to overcome the many challenges associated with packaging duplex RNA for therapeutic applications.

We thank K. Finkelstein for experimental assistance. This research was supported by the NSF, the NIH and the NBTC at Cornell. This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS) which is supported by the NSF and the NIH/NIGMS. We also made use of the Cornell NanoScale Facility, a member of the National Nanotechnology Infrastructure Network, which is supported by the NSF. and CNF at Cornell. CHESS is supported by NSF and NIH/NIGMS.

REFERENCES

- [1] V. A. Bloomfield, *Biopolymers* **44**, 269 (1997).
- [2] N. V. Hud and I. D. Vilfan, *Annu. Rev. Biophys. Biomol. Struct.* **34**, 295 (2005).
- [3] O. Boussif, F. Lezoualch, M. A. Zanta, M. D. Mergny, D. Scherman, B. emeneix, and J. P. Behr, *Proc. Natl. Acad. Sci. USA* **92**, 7297 (1995).
- [4] A. Fire, S. Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, *Nature* **391**, 806 (1998).
- [5] N. F. Boussein, C. S. McAllister, K. K. Ewert, C. E. Samuel, and C. R. Safinya, *Biochemistry* **46**, 4785 (2007); D. Grimm and M. A. Kay, *J. Clin. Invest.* **117**, 3633 (2007).
- [6] A. A. Kornyshev, D. J. Lee, S. Leikin, and A. Wynveen, *Rev. Mod. Phys.* **79**, 943 (2007); G. C. L. Wong and L. Pollack, *Annu. Rev. Phys. Chem.* **61**, 171 (2010).
- [7] S. A. Pabit, X. Y. Qiu, J. S. Lamb, L. Li, S. P. Meisburger, and L. Pollack, *Nucleic Acids Res.* **37**, 3887 (2009).
- [8] J. Pelta, F. Livolant, and J. L. Sikorav, *J. Biol. Chem.* **271**, 5656 (1996).
- [9] J. Widom and R. L. Baldwin, *J. Mol. Biol.* **144**, 431 (1980).
- [10] S. L. Heilman-Miller, D. Thirumalai, and S. A. Woodson, *J. Mol. Biol.* **306**, 1157 (2001).
- [11] X. Y. Qiu, L. W. Kwok, H. Y. Park, J. S. Lamb, K. Andresen, and L. Pollack, *Phys. Rev. Lett.* **96**, 138101 (2006); X. Y. Qiu, K. Andresen, L. W. Kwok, J. S. Lamb, H. Y. Park, and L. Pollack, *Phys. Rev. Lett.* **99**, 038104 (2007).
- [12] K. Andresen, X. Y. Qiu, S. A. Pabit, J. S. Lamb, H. Y. Park, L. W. Kwok, and L. Pollack, *Biophys. J.* **95**, 287 (2008).
- [13] X. Y. Qiu, K. Andresen, J. S. Lamb, L. W. Kwok, and L. Pollack, *Phys. Rev.*

Lett. **101**, 228101 (2008).

[14] M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, and N. A. Clark, *Science* **318**, 1276 (2007).

[15] I. Rouzina and V. A. Bloomfield, *Journal of Physical Chemistry* **100**, 4292 (1996).

[16] L. Li, S. A. Pabit, J. S. Lamb, H. Y. Park, and L. Pollack, *Appl. Phys. Lett.* **92**, 223901 (2008).

[17] G. Zanchetta, T. Bellini, M. Nakata, and N. A. Clark, *J. Am. Chem. Soc.* **130**, 12864 (2008).

[18] K. Chin, K. A. Sharp, B. Honig, and A. M. Pyle, *Nat. Struct. Biol.* **6**, 1055 (1999).

[19] H. Robinson and A. H. J. Wang, *Nucleic Acids Res.* **24**, 676 (1996).

[20] A. A. Ouameur and H. A. Tajmir-Riahi, *J. Biol. Chem.* **279**, 42041 (2004).

[21] C. M. Nunn and S. Neidle, *J. Mol. Biol.* **256**, 340 (1996).

[22] B. I. Kankia, V. Buckin, and V. A. Bloomfield, *Nucleic Acids Res.* **29**, 2795 (2001).

[23] J. S. Kieft and I. Tinoco, *Structure* **5**, 713 (1997).

[24] L. Dai, Y. G. Mu, L. Nordenskiöld, and J. R. C. van der Maarel, *Phys. Rev. Lett.* **100**, 118301 (2008); M. O. Delacruz, L. Belloni, M. Delsanti, J. P. Dalbiez, O. Spalla, and M. Drifford, *J. Chem. Phys.* **103**, 5781 (1995).

[25] I. A. Nyrkova and A. N. Semenov, *Soft Matter* **5**, 979 (2009).

[26] A. A. Kornyshev and S. Leikin, *Phys. Rev. Lett.* **82**, 4138 (1999).

[27] B. A. Todd, V. A. Parsegian, A. Shirahata, T. J. Thomas, and D. C. Rau, *Biophys. J.* **94**, 4775 (2008).

[28] M. Egli, S. Portmann, and N. Usman, *Biochemistry* **35**, 8489 (1996).

CHAPTER 6

Counting trivalent ions and measuring hydration

Abstract

Using Multiple-Energy Anomalous Small Angle X-ray Scattering (mE-ASAXS), we count the number of trivalent ions associated with 25 bp double-stranded DNA during the early stages of DNA condensation. Two different trivalent cobalt compounds were employed: cobalt hexamine chloride (co-hex), and cobalt sepulchrane chloride (co-sep). Although both compounds effectively condense DNA, the condensing power of co-sep is almost twice as much as that of co-hex: more DNA precipitates from solution when a fixed amount of co-sep is added, relative to addition of the same amount of co-hex. Our measurements reveal that, under comparable experimental conditions, the number of excess counterions is similar for the two compounds, thus differences in condensing power do not appear to be related to electrostatic charge compensation. Previously, changes in condensing power were explained by ion hydration differences. In complementary studies, we probed the hydration level of the counterions using X-ray Absorption Fine Structure (XAFS). The combined results indicate that, despite the similar electrostatic response of the systems, the more hydrated counterion species (co-hex) has a smaller condensing power,

Introduction

It has long been recognized that DNA condensation can be induced by the addition of even small number of trivalent ions [1]. Clues to the underlying mechanism of this

like-charge attraction lie in observing differences in condensation behavior when ions with different physical properties are employed. Thus studies of DNA condensation have been carried out in the presence of different small trivalent cobalt-amine compounds, including cobalt hexamine (co-hex) and cobalt sepulchrate (co-sep). Both are powerful condensing agents of double-stranded DNA (dsDNA) [2]. But co-sep has more hydrophobic CH₂ groups than co-hex. The hydration levels between the two compounds are expected to be different. Differences in the chemical and physical properties of those molecules were extensively discussed in reference [3], in an effort to explain their differing condensing powers. Light scattering experiment suggests that co-sep molecule condense high-molecular-weight calf thymus DNA about twice as effectively as co-hex despite the same valence of the two cobalt counterions [3]. However, a number of important parameters have evaded detection, such as the number of excess trivalent ions resulting from the presence of the DNA. Here, using different SAXS techniques [4], we count excess ions and measure the strength of attraction they mediate between short 25-bp DNA duplexes. Surprisingly, we find that the excess numbers of co-sep and co-hex ions are very similar when the bulk ionic strength of the solution is fixed. This observation raises the question why dsDNA would be more prone to condensation by co-sep than co-hex?

In addition to the prevailing electrostatic arguments, it has been proposed that the hydration structure of DNA and counterions may have a strong impact on ion-DNA association and inter-DNA attraction [5]. Hydration forces between DNA molecules, either attractive or repulsive, were measured directly by the osmotic stress method [5, 6]. The release of ordered water solvent around DNA and/or ions into the bulk solution is one of the major sources of entropy increase of the system and has been

linked with hydration forces. Likewise, it is proposed that the effectiveness of ion-induced DNA condensation is highly affected by the hydration pattern of the DNA-ion surface. Co-hex in solution is highly hydrated while co-sep has a much lower hydration level [3]. Changes in hydration that accompany ion binding to DNA have been reported [7]. Here, we explore the possibility of probing the difference of hydration structure between co-hex and co-sep ions that are bound with DNA via the well-established technique of X-ray Absorption Fine Structure (XAFS) [8, 9].

Experimental Techniques

mE-ASAXS

Solution Anomalous Small Angle X-ray Scattering (ASAXS) can investigate the spatial distribution of counterions around macro-molecules [10] when the counterions have electron binding energies that are accessible to synchrotron x-rays. For ASAXS experiments SAXS profiles are measured at two different energies; the first is far below the ion absorption edge and the second is close to, but just below the absorption edge. The energy-dependent term which represents the counterion spatial distribution can be obtained by subtraction of the two scattering profiles with proper fluorescence correction. All energy independent terms cancel in this subtraction.

Multiple-Energy Anomalous Small Angle X-ray Scattering or mE-ASAXS is a natural extension of ASAXS. SAXS profiles are measured at multiple different energies (5 in this experiment) carefully selected below the absorption edge. The total scattered intensity $I(Q, E)$ is thus dependent on both X-ray energy E and momentum transfer $Q = (4\pi/\lambda)\sin\theta$, where λ is X-ray wavelength and 2θ is scattering angle. The ion scattering factor consists of one energy-independent term f_0 and two energy-

dependent terms $f''(E)$ and $f'''(E)$. Ignoring the latter term f''' which accounts for absorption, thus is negligibly small below the edge, the total scattered intensity $I(Q, E)$ can be expressed in terms of $f''(E)$ in standard quadratic form

$$I(Q, E) = a(Q) \cdot (f''(E))^2 + b(Q) \cdot f''(E) + c(Q, E)$$

where

$$a(Q) = N^2 F_{ion}^2$$

$$b(Q) = N \cdot [2f_{NA}F_{NA}F_{ion} + 2f_0NF_{ion}^2]$$

$$c(Q, E) = (f_{NA}F_{NA})^2 + 2f_{NA}f_0NF_{NA}F_{ion} + f_0^2N^2F_{ion}^2$$

Because changes in the scattering factor are so small compared to the absolute values of the scattering factors, the coefficient of the second order term is small and this term is negligible. A linear fit of $I(Q, E)$ with respect to $f''(E)$ yields the coefficients of the linear and constant terms which then can be used to derive the excess number of associated ions as shown below

$$N = b(Q = 0) / 2\sqrt{c(Q = 0)}$$

The procedure was described extensively in ref [4]

XAFS

We used X-ray Absorption Fine Structure (XAFS) spectroscopy to measure the hydration level of counterions both in solution and adsorbed to the DNA. XAFS probes the first few hydration shells around selected atom by measuring the absorption coefficient profile over a range of X-ray energies that are very close to the absorption edge [11]. The absorption coefficient of an atom usually decreases as the X-ray energy increases ($\sim 1/E^3$). However, at an absorption edge, this coefficient increases sharply

because the incident photon is absorbed by the atom resulting in an electronic transition to a higher energy state. The photoelectron generated by this process can be treated as an outgoing wave which is reflected by the local structure to produce ingoing waves. The interference between the ingoing and outgoing wave generates an oscillation pattern above the absorption edge. The XANES (X-ray absorption near edge structure) region, in which multiple scattering dominates, covers the energy range 5~200eV above edge while the EXAFS (extended X-ray absorption fine structure) region ranges from ~200eV to 1000eV above the edge. From the XAFS pattern we extract information about the structure of nearest neighbor atoms to the absorbing atom of interest. Spectra from different samples can be compared if they are normalized to have the same step height and zero background (below the edge). This normalization is carried out by regressing a line to the region below and above the absorption edge, subtracting the pre-edge line from the entire data set and dividing by the absorption step height.

Here, we use XAFS to monitor the hydration levels of co-hex and co-sep ions associated with dsDNA [11]. We also measured the absorption spectrum of co-en (cobalt ethylenediamine) associated with DNA since it was previously reported that this ion has the hydration level in between that of co-hex and co-sep [3].

mE-ASAXS setup

The mE-ASAXS measurements were carried out at the C1 station of the Cornell High Energy Synchrotron Source (CHESS). Full details about the experimental setup are provided in Ref. [12]. Briefly, the beam energy is carefully selected using a double-crystal monochromator. The energy can be scanned over a large range and the

energy resolution of the experiment is 2 eV. The beam size was determined by slits to be ~1.5mm wide and ~0.7mm high. All DNA samples were contained in acrylic sample cells, with volume of 25 ~ 40 uL, sealed with ultra-thin (~0.7 um) silicon nitride windows. After interaction with the sample, the scattered X-ray beam passed through a 1-m long, evacuated flight tube. A CCD camera was placed at the exit of the flight tube and used to record scattering profiles. The direct beam was blocked by a motorized beamstop to prevent damage to the CCD detector. Energies for measurements were determined by transmission scans of Co-containing buffer solutions. The program CHOOCH [13] was used to determine values of f' and f'' from transmission scans as described in Ref [4]. Based on results from CHOOCH, we selected five different energies (7.614keV, 7.664keV, 7.705keV, 7.714keV, 7.722keV) for this study (see figure 6.1), which give us within an order of magnitude variation in f' . SAXS profiles were acquired both on DNA-containing samples, as well as matching buffers for background subtraction. An additional fluorescence correction was applied to the data, as described [12]. Finally, due to inter-DNA interactions, low angle data were modulated by a structure factor $S(Q)$. Thus, great care was taken to avoid this low Q region. We analyzed data over Q range between 0.044Å to 0.055Å, including as much information as possible while at the same time avoiding the Q region affected by structure factor. Finally, to obtain the ion number in absolute units, the scattering intensity at zero-angle $I(Q=0)$ was scaled to number of electrons squared using the calibrated scattering signal from water [14].

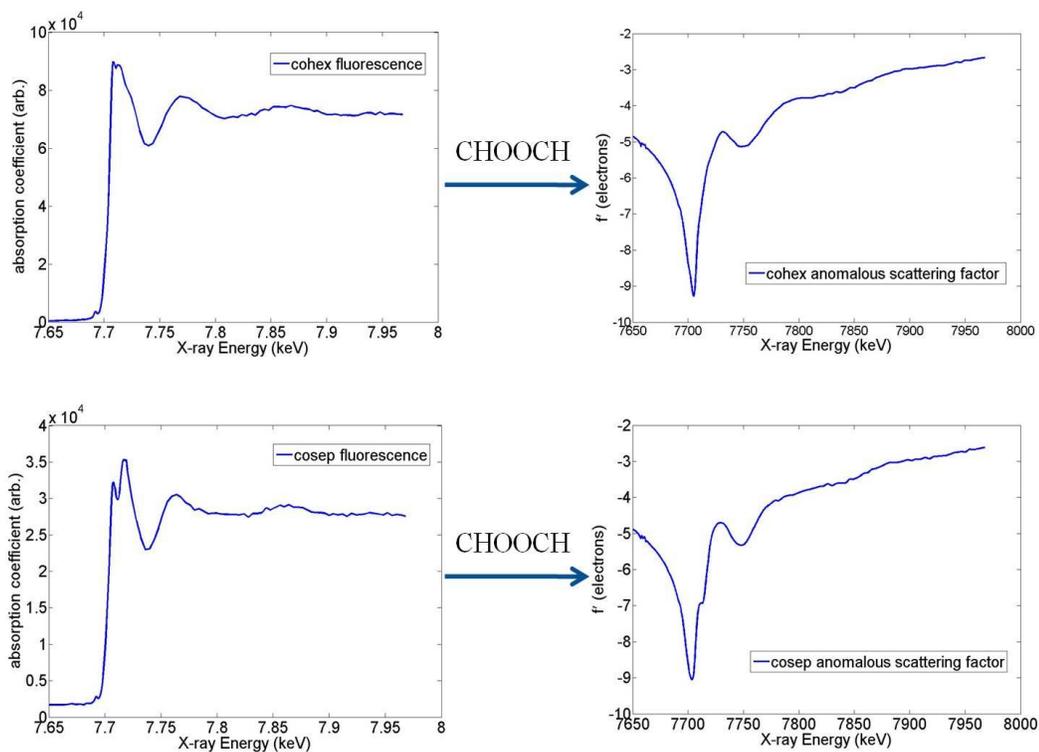


Figure 6.1 conversion of fluorescence scan by “CHOCH”

These plots show the real part of the anomalous scattering factor f' for Cobalt ions obtained by the program CHOCH. Five energies are chosen for the mE-ASAXS measurement.

XAFS setup

X-ray absorption spectra were also acquired at the CHESS C1 station using a Si (111) double-crystal monochromator. Higher harmonics were reduced by detuning the monochromator by 50%~60%. The Cobalt absorption at K-edge was found at 7.726keV using both 1mM cobalt solution contained in mylar capillary and a cobalt foil. Spectra were measured by scanning the X-ray energy from 7.65keV to 8.1keV. Step sizes of 5eV were used between 7.65 and 7.71keV. Higher resolution 2eV steps were employed in the critical region between 7.71 and 7.74keV. Above 7.74 keV and up to 8.1 keV, the step size was increased to 4eV. For these measurements, the beam diameter was set to less than 1 mm. The K-edge XAFS spectra were acquired in transmission mode and 1mm-diameter mylar tubes were used as DNA/cobalt sample container. The X-ray fluorescence was collected using 4 vortex detectors.

Sample preparation

Single-stranded 25mer DNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA) and then annealed to form double-stranded DNA using standard annealing protocol as previous described [14]. All 25bp dsDNA samples for mE-ASAXS experiments were dialyzed against co-hex or co-sep buffers (0.5mM co-hex, and 0.5mM co-sep, respectively where no condensation of DNA was observed) containing 100mM NaCl and 1mM NaMOPS at pH=7 except for NaCl control samples which contain no cobalt compounds. The concentration of DNA is maintained at 0.6mM which was determined by UV Spectroscopy (Cary 50 Bio, Varian, Inc., Walnut Creek, CA). The solutions were brought to a total volume ~30uL for mE-ASAXS measurement. The dialyzed DNA samples containing 0.5mM bulk

concentration of co-hex, co-sep as well as co-en (cobalt ethylenediamine, another trivalent ion, DNA condensing agent) as a control was used for hydration measurement by XAFS. The hydration level of freely-moving co-en is higher than co-sep but lower than co-hex.

Results and discussion

SAXS profiles of DNA + co-hex (0.5mM bulk concentration of co-hex), DNA + co-sep (0.5mM bulk concentration of co-sep) and DNA + NaCl (100mM bulk concentration of NaCl) are shown in figure 6.2. All curves were acquired at X-ray energy = 7.514keV which is below the Co absorption edge. For comparison the curves are intensity normalized at $Q = 0.1(\text{\AA}^{-1})$. The effect of inter-particle repulsion is easily seen in the low Q downturns of the sample. A Guinier fit is performed to obtain $I(Q=0)$ for each of the 5 energies over a specific Q range. This range was carefully selected to lie above the region where inter-particle interactions modify the scattering profile (in figure 6.2 where the 3 curves start to deviate from each other). The upper bound for this fit is consistent with measurements on non-interacting particles. Guinier fits for DNA/co-hex and DNA/co-sep samples at 5 energies are shown in figure 6.3 (a) and (b), respectively. From these fits, the value of $I(Q=0)$ was obtained. This value is plotted as a function of X-ray energy in figure 6.4. The number of excess ions associated with DNA can be extracted from linear fits to these points as described in Ref [4]. We find 6.7 ± 0.8 excess ions for co-hex, in agreement with our previous findings using ICP [10]. A comparable number, 7.0 ± 1.3 ions is measured for co-sep.

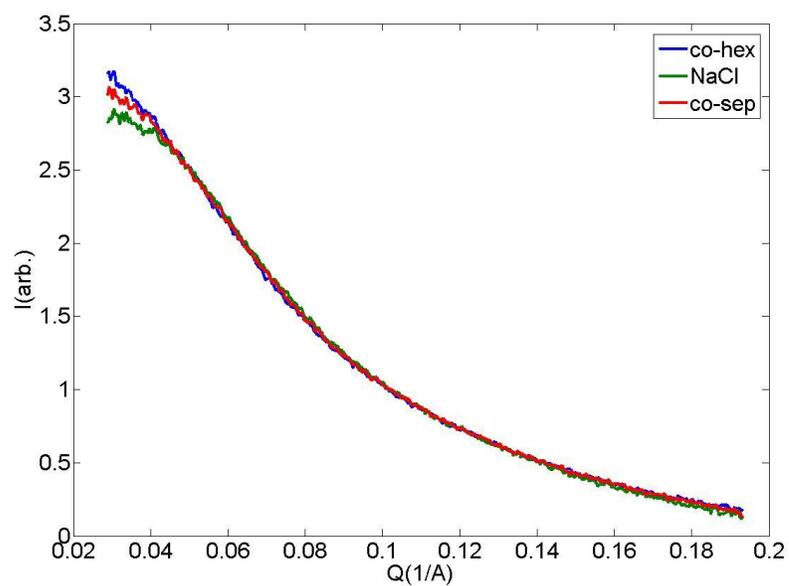


Figure 6.2

SAXS profiles of DNA dialyzed against 100mM NaCl, 0.5mM co-hex, and 0.5mM co-sep solutions (buffered at 100mM NaCl, 1mM NaMOPS, pH = 7), respectively, are shown in this figure. They are normalized at $Q = 0.1 \text{ \AA}^{-1}$. DNA molecules repel each other in all three samples with the strongest repulsion observed in DNA + NaCl. The structure factors at $Q = 0$ are very close for DNA + co-hex and DNA + co-sep sample (see text).

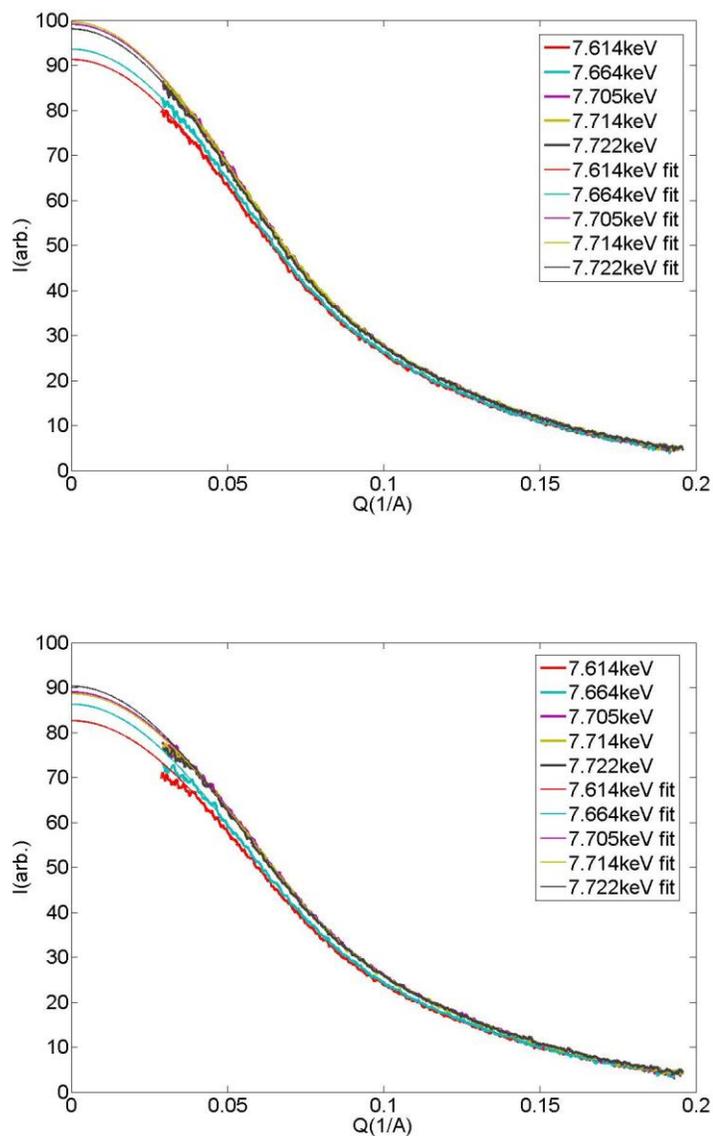


Figure 6.3

Solution SAXS profiles of DNA dialyzed against 0.5mM Co-hex and 0.5mM Co-sep buffer in 100mM NaCl and 1mM NaMOPS at pH 7, respectively. SAXS curves were measured at 5 different energies below the absorption edge. Guinier fit was performed over the Q range within $0.044 \sim 0.055 \text{ \AA}^{-1}$ optimized from theoretical form factor.

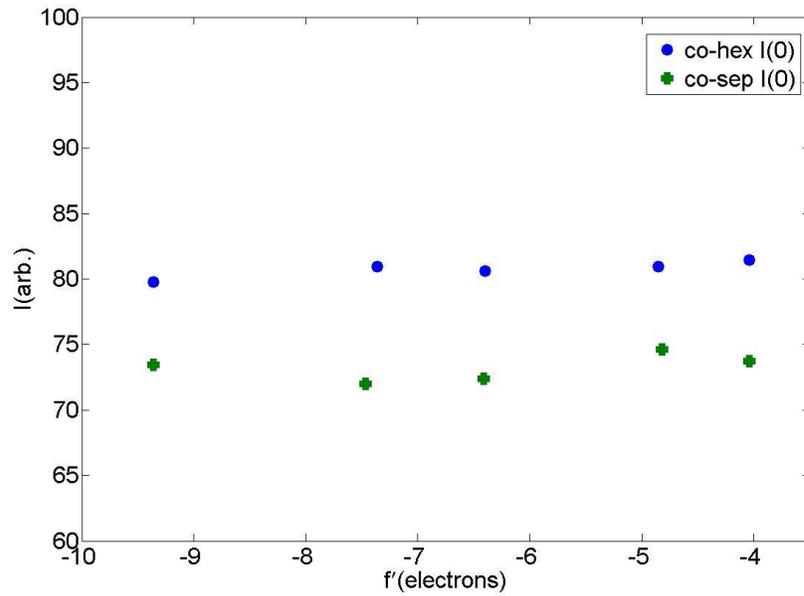


Figure 6.4

$I(Q=0)$ for 5 different X-ray energies were obtained by Guinier fit for DNA + co-hex and DNA + co-sep samples. A plot of $I(Q=0)$ versus the real part of anomalous scattering factor is shown above. A linear fit is then performed to determine the magnitudes of b-term and c-term at $Q = 0$ which are used to compute ion numbers.

We measured the X-ray absorption spectra of DNA samples dialyzed against co-hex, co-en and co-sep. To prepare these samples, 0.6mM DNA duplex was dialyzed against 0.5mM co-hex (co-en and co-sep) buffer containing 100mM NaCl and 1mM NaMOPS at pH 7. These concentrations were selected so that the majority of cobalt ions in the sample would be associated with DNA molecules. We measured 8~15 absorption curves for each sample and the averaged XAFS spectra are shown in figure 6.5. The shape difference in the XANES regime (inset of figure 6.5) suggests that the hydration patterns of those 3 ions are very different [11, 16] (private communication with Prof. Serena DeBeer in Chemistry Department at Cornell University). Co-hex is still hydrated, which is consistent with the findings in ref [17, 5, 6]. The dramatic shape changes in this regime of the spectrum suggest that Co-sep is the least hydrated of the three. A similar trend is observed for ions freely moving in the buffer [3]. Thus, since no differences in the spectrum of DNA-absent vs. DNA-present samples were detected, we deduce that the binding of the cobalt ion to DNA does not change ion hydration.

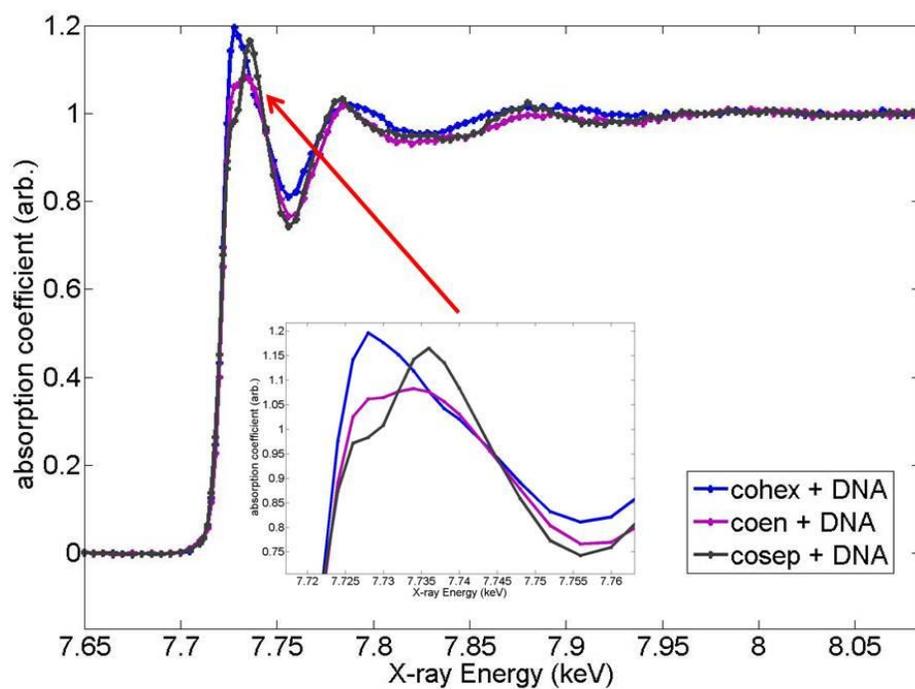


Figure 6.5

Solution XAFS spectra at Cobalt edge for DNA-bound co-hex, co-en, and co-sep ions are plotted together. The spectra are normalized and background subtracted for comparison. The inset represents the XANES regime of XAFS spectrum which provides information about hydration level of these three ions.

Taken together, these results suggest that the hydration level of the DNA-counterion interface has an impact on the effectiveness of DNA condensation. Co-sep and co-hex have identical charge (+3), very similar size and chemical structure. Moreover, according to mE-ASAXS measurements, the number of co-sep bound to DNA is very close to that of co-hex. The only biggest difference between the two cobalt compounds is the hydration level. Co-sep is more hydrophobic probably due to the extra CH₂ groups. It has been found that the freely-moving co-hex is more hydrated than co-sep in the buffer [3]. The XANES pattern further suggests that co-hex is still highly hydrated when it's bound to DNA but co-sep is much less. Other experimental evidence shows that there is a hexagonal packing with $\sim 8\text{\AA}$ separation between DNA surfaces in the condensed DNA phase mediated by co-hex [17]. Both findings indicate that co-hex induced DNA condensation is more affected by hydration structure at DNA/counterion surfaces than co-sep. Therefore we propose that the higher condensing power of co-sep than co-hex is from their distinct hydration structures primarily for two reasons. First, with fewer extra water layers around co-sep, the cost of rearranging ordered water molecules is lower during the short-range association between DNA molecules. Second, when the counterion is more hydrated, the effective size of the counterion is increased hence the effective charge density of the ion is decreased. The effective local dielectric constant is increased due to more surrounding polarized water molecules. Consequently the charge screening effect of less hydrated co-sep is stronger than co-hex. Therefore, the long-range attraction between DNA strands induced by co-sep is stronger even though there are about the same excess numbers of co-sep and co-hex associated with DNA. In summary, the effects of both entropy and enthalpy support our proposal above—DNA condensing

power of co-sep is enhanced relative to co-hex because of its reduced hydration level.

Conclusion

In this paper, we report measurements of excess number of trivalent ions around DNA, just prior to the condensation threshold. Despite differences in hydration of co-hex and co-sep ions, similar numbers are associated to DNA. Thus, it appears that electrostatic contributions to DNA condensation are similar for these two ions. To understand why the condensing power of co-sep on DNA is almost twice as much as co-hex, we then investigated the hydration level of these cobalt-amine compounds as they bind to DNA molecules. Co-hex is more hydrated than co-sep as is the case when both are freely mobile in solution. This confirms the proposal that hydration structure at DNA and counterion surface should be taken into account throughout the DNA condensation process [18].

We thank K. Finkelstein for experimental assistance. This research was supported by the NSF, the NIH and the NBTC at Cornell. This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS) which is supported by the NSF and the NIH/NIGMS. We also made use of the Cornell NanoScale Facility, a member of the National Nanotechnology Infrastructure Network, which is supported by the NSF.

REFERENCES

- [1] V. A. Bloomfield, *Biopolymers* **44**, 269 (1997)
- [2] W. Gelbart, R. Bruinsma, P. Pincus, and V. Parsegian, *Physics Today* **53**, 38 (2000)
- [3] H. Deng, and V. A. Bloomfield, *Biophys. J.* **77**, 1556 (1999)
- [4] S. A. Pabit, S. P. Meisburger, L. Li, J. M. Blose, C. D. Jones, and L. Pollack, *J. Am. Chem. Soc.* **132**, 16334 (2010)
- [5] D. C. Rau, and V. A. Parsegian, *Biophys. J.* **61**, 246 (1992)
- [6] D. C. Rau, and V. A. Parsegian, *Biophys. J.* **61**, 260 (1992)
- [7] B. I. Kankia, V. Buckin, and V. A. Bloomfield, *Nucleic Acids Res.* **29**, 2795 (2001).
- [8] E. A. Stern, and S. M. Heald, in *Handbook of Synchrotron Radiation*, edited by E. E. Koch (North Holland, New York, 1983), Chap. 10, pp. 995-1014
- [9] *X-ray absorption: Principles, Applications, Techniques of EXAFS, SEXAFS, and XANES*, Vol. 92 of *Chemical Analysis*, edited by D. C. Koningsberger and R. Prins (Wiley, New York, 1988)
- [10] K. Andresen, X. Qiu, S. A. Pabit, J. S. Lamb, H. Y. Park, L. W. Kwok, and L. Pollack, *Biophys. J.* **95**, 287 (2008)
- [11] J. L. Fulton, D. M. Pfund, and S. L. Wallen; M. Newville, E. A. Stern, and Y. Ma, *J. Chem. Phys.* **105**, 2161 (1996)
- [12] S. A. Pabit, K. D. Finkelstein, and L. Pollack, *Methods Enzymol.* **469**, 391 (2009)
- [13] G. Evans, and R.F. Pettifer, *J. Appl. Crystallogr.* **34**, 82 (2001)
- [14] D. Orthaber, A. Bergmann, and O. Glatter, *J. Appl. Crystallogr.* **33**, 218 (2000)
- [15] X. Qiu, K. Anderson, J. S. Lamb, L. W. Kwok, L. Pollack, *Phys. Rev. Lett.* **101**,

228101 (2008)

[16] D. M. Pfund, J. C. Darab, J. L. Fulton, and Y, Ma, *J. Phys. Chem.* **98**, 13102 (1994)

[17] S. Leikin, D. C. Rau, and V. A. Parsegian, *Phys. Rev. A.* **44**, 5272 (1991)

[18] S. Leikin, V. A. Parsegian, D. C. Rau, and R. P. Rand, *Annu. Rev. Phys. Chem.* **44**, 369 (1993)

CHAPTER 7

Future work

In this chapter, I summarize some of the experiments and research projects that might be suited for a continuation of this research. Some preliminary results are listed in this chapter as well.

7.1 DNA and RNA condensation by spermine

Different behavior of dsDNA and dsRNA in co-hex solution highlights the importance of nucleic acid geometry to condensation. It would be worthwhile to investigate RNA condensation using other known powerful condensing agents for DNA. Spermine, +4 positively-charged, is one perfect candidate with known structure (PDB ID: SPM). A sketch of the structure of spermine is shown in figure 7.1. The length of spermine is $\sim 17\text{\AA}$ and the width is $3 \sim 4\text{\AA}$. It's been shown that spermine is a very effective condensing agent for DNA [1] (even more effective than co-hex). Similar UV-spec and SAXS experiments as described in chapter 5 were carried out on DNA and RNA with spermine. One key difference in SAXS measurement is that the DNA/RNA samples were prepared by adding spermine to the solution directly rather than equilibrium dialysis as in the case of co-hex. Solution containing 20mM NaCl + 1mM NaMOPS with pH = 7 was used as buffer for both UV and SAXS experiments. This solution is used because the condensing power of spermine at DNA is too large and the dialysis will precipitate all the DNA molecules out of the supernatant.

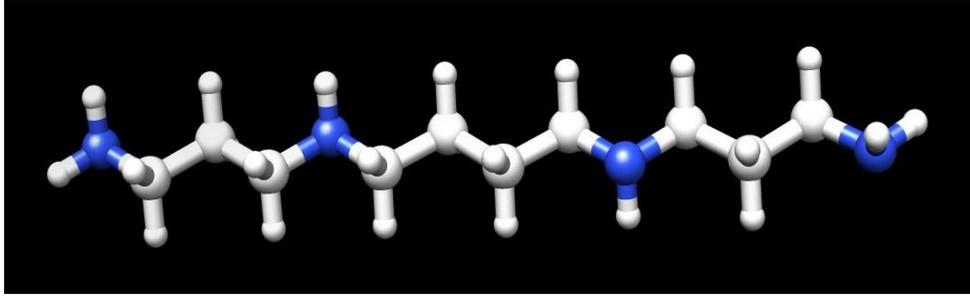
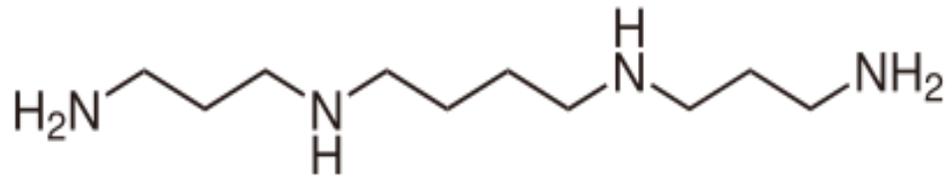


Figure 7.1 structure of spermine molecule

The result of UV spectroscopy experiment is shown in figure 7.2. SAXS curves are presented in figure 7.3. The powerful DNA condensing agent does not condense RNA much even at concentration as high as 30mM. Both DNA and RNA molecules strongly repel each other when spermine concentration is between 0.5mM and 2mM. The behavior of DNA molecules is understandable since some of them have been precipitated out of solution by spermine and what's left in the supernatant repel each other. However, there's no precipitation of RNA observed, they still repel each other in spermine solution in contrast to the observation that RNAs attract each other significantly in co-hex solution. It was proposed that spermine ions reside in DNA major groove [2] and thus neutralize the negative charge of DNA. However, it is likely that spermine is not strongly associated with RNA since both the major and minor grooves are too narrow for spermine. Therefore this suggests that the negative charge of RNA is even not neutralized by spermine. It would be interesting to repeat this experiment and use RNA strands with different lengths to measure any length-dependence of the interaction between RNA molecules.

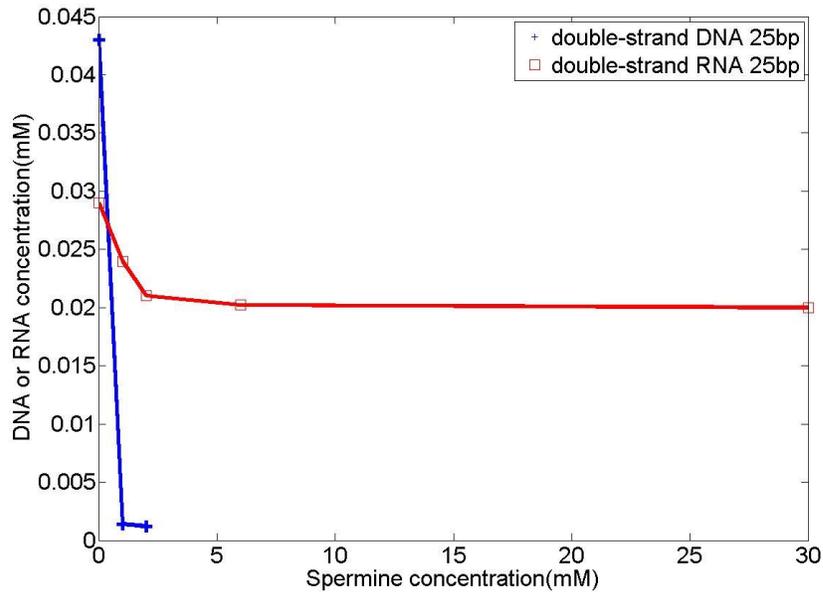


Figure 7.2 dsDNA/dsRNA condensation induced by spermine

Most of the 25bp DNA molecules are precipitated out of the solution when spermine concentration reaches about 1mM. In contrast, most 25bp RNA molecules remain in the supernatant even with extremely high spermine concentration.

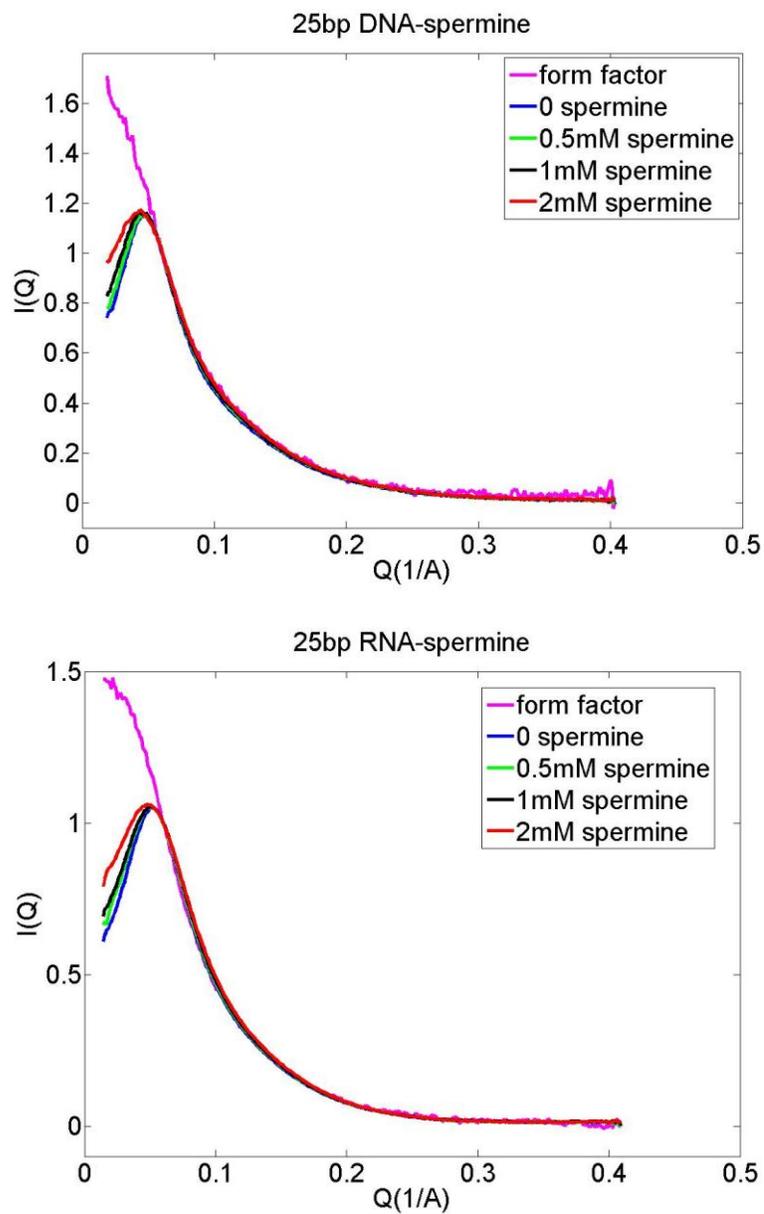


Figure 7.3

SAXS profiles of DNA and RNA are measured in solutions with different spermine concentrations. Both DNA and RNA molecules strongly repel each other in the supernatant.

7.2 Wide Angle X-ray Scattering

Wide Angle X-ray Scattering (WAXS) is an X-ray diffraction technique that can be used to provide higher resolution structural information of polymers than SAXS. The experimental setup of WAXS technique is very similar to that of SAXS except that the distance from sample to the detector is usually much shorter (~30cm in WAXS shown in figure 7.4 versus ~100cm in SAXS) so that the diffraction pattern can be recorded at large angles. The resolution can reach about ~20Å. Solution WAXS provides a direct measure of macromolecular conformation in solution and one-dimensional (1D) “fingerprints” of 3D structure that are directly relatable to atomic configuration by Fourier transform [3].

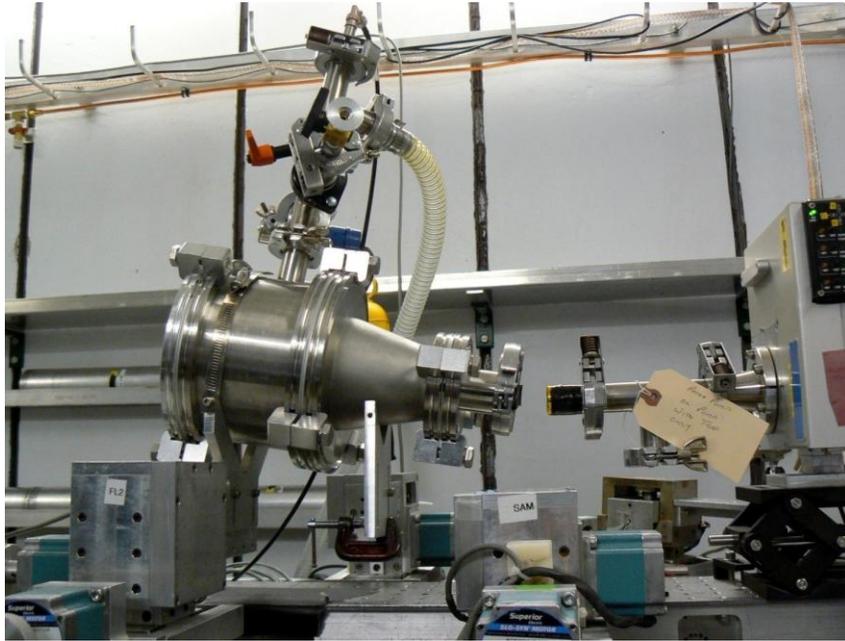
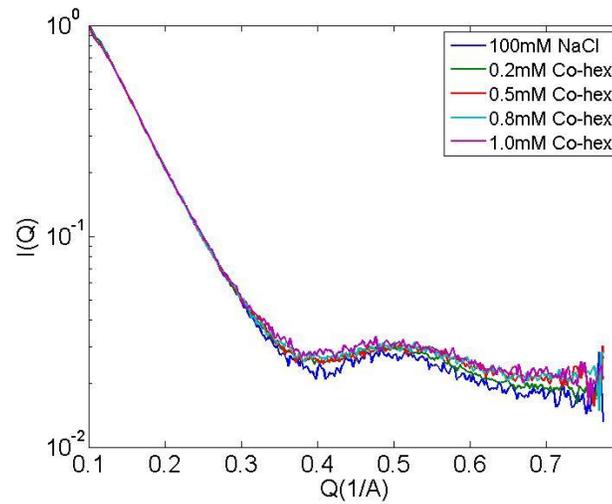


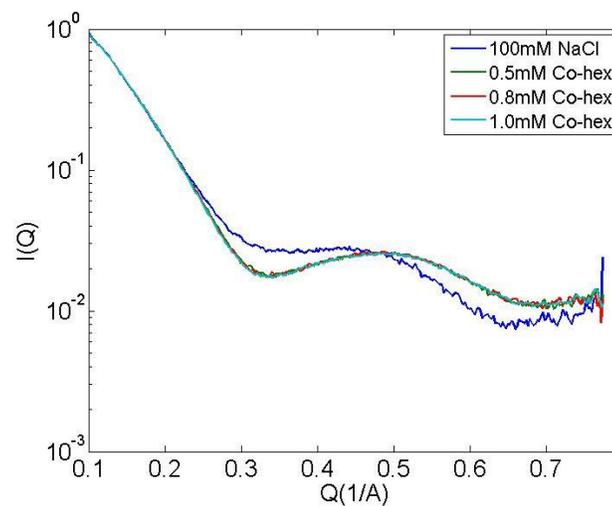
Figure 7.4 a picture of WAXS experimental setup at G-line CHESS

Comparing to SAXS, the flight tube is much shorter. The length of flight tube is ~30cm in WAXS experiment.

WAXS scattering profiles provide additional conformational information about DNA (or RNA)-ion structures, specifically reporting on length scales shorter than those accessible by SAXS. Figure 7.5 shows the extended angular range of these experiments, accessing Q ranging from $0.1 \sim 0.8$ ($1/\text{\AA}$). The WAXS curves of DNA, shown in figure 7.5 display several interesting features, notably a peak at $Q = 0.5$ ($1/\text{\AA}$). When compared with the results of ref 3 we attribute this peak to helical inter-strand pair distance correlations. The peak broadens slightly as the co-hex concentration is increased, likely caused by the structural difference between DNA-ion system and DNA helix itself. Figure 7.5 also shows similar curves acquired on RNA. In striking contrast to DNA in 100 mM NaCl, the scattering curve of RNA-NaCl is almost flat from $Q = 0.32(1/\text{\AA})$ to $Q = 0.45(1/\text{\AA})$. These shape changes arise from the fact that A-RNA helix is “broader” and more compact than B-DNA. When dsRNA is dialyzed against co-hex, a noticeable peak at around $Q = 0.48(1/\text{\AA})$ is generated. We speculate that this is due to the constructive interference between “effective” helix radius modified by co-hex and major/minor groove spacing. The peak shift from $Q \sim 0.4$ ($1/\text{\AA}$) (though it is hard to assign given the small amplitude variation) to $Q \sim 0.48$ ($1/\text{\AA}$) indicates a reduction of the diameter of the RNA-ion structure comparing to RNA helix itself as illustrated by figure 7.6.



(a)



(b)

Figure 7.5

WAXS curves for DNA (a) and RNA (a) samples containing different amount of co-hex. For DNA samples, as the co-hex concentration increases, the shoulder between $Q = 0.4$ ($1/\text{\AA}$) and $Q = 0.6$ ($1/\text{\AA}$) becomes flatter. For RNA, the change is much more significant. The underlying cause of this dramatic change is worth investigating.

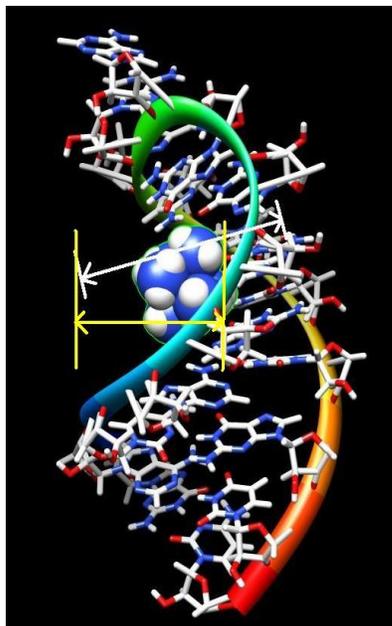


Figure 7.6

When co-hex binds into the major groove of dsRNA, we propose that the effective radius of the RNA molecule is reduced as indicated by the white and yellow double-arrows.

The underlying mechanism of this big change in RNA WAXS profile is still not clear. A simple, qualitative and speculative proposal is provided above but more convincing evidence is required. It might also need some efforts in modeling and computation power may also be necessary to solve this puzzle quantitatively.

7.3 End-to-end stacking: RNA dumbbell

Similar to the DNA dumbbell experiment in chapter 4, RNA dumbbells were made with duplex length of 6bp, 8bp and 16bp. SAXS experiments were carried out on RNA dumbbells to investigate the length-dependence of interactions. The reason to use dumbbell RNA rather than normal double-stranded RNA is to avoid end-to-end stacking which is known to be prominent between short dsRNA molecules (see chapter 5). The protocol for sample preparation is the same as that described in chapter 4 except that 6bp and 8bp RNA dumbbell molecules were made by directly self-folding from one single strand with 16 nucleotides and 20 nucleotides, respectively. To account for the difference of number of phosphate charges, RNA concentrations for 6bp, 8bp, 16bp are maintained at 1.0mM, 0.9mM, 0.6mM, respectively. Second virial coefficients of RNA dumbbells were calculated in 100mM K^+ , 0.5mM Mg^{2+} , 1mM Mg^{2+} and 10mM Mg^{2+} solutions and shown as a function of both the ion concentration and length of RNA duplex in figure 7.7. Two features can be easily observed. First, RNA dumbbells repel at 100mM K^+ and at the low concentrations of Mg^{2+} . The strongest repulsion is observed between the smallest (6bp duplex) RNA dumbbell molecules at the low ionic strength of Mg^{2+} . Shorter RNA molecules also have greater changes in the second virial coefficient when ionic strength of Mg^{2+} changes. Second, 6bp (and 8bp) RNA dumbbells aggregate when $[Mg^{2+}]$ reaches

10mM. This is unexpected since end-to-end stacking is blocked. Moreover, this type of interaction is also length-dependent and favors short strands since 16bp RNA dumbbell molecules still repel each other (remember that 20bp DNA dumbbells repel each other when $[Mg^{2+}] = 10mM$ according to chapter 4). It is also possible that Mg^{2+} may disrupt the confirmation of short RNA dumbbells, which could potentially be investigated by WAXS experiment.

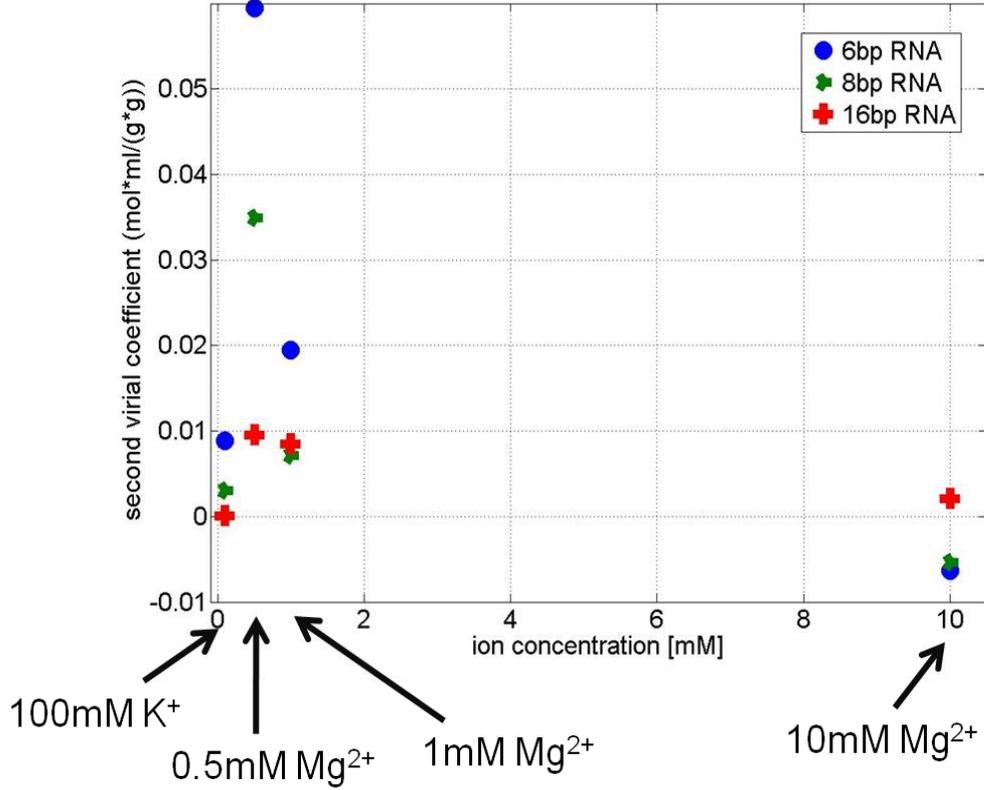


Figure 7.7 plot of second virial coefficients of RNA-K⁺ and RNA-Mg²⁺ versus ion concentration

Shorter RNA dumbbells repel strongly with each other at lower Mg²⁺ concentration compared to longer RNA dumbbells. However, in the solution containing 10mM Mg²⁺, shorter RNA dumbbells aggregate. In contrast longer RNA dumbbells still repel each other weakly.

REFERENCES

- [1] X. Qiu, K. Anderson, J. S. Lamb, L. W. Kwok, L. Pollack, *Phys. Rev. Lett.* **101**, 228101 (2008)
- [2] B. G. Reuerstein, N. Pattabiraman, and L. J. Marton, *Proc. Natl. Acad. Sci.* **83**, 5948 (1986)
- [3] X. Zuo, G. Cui, K. M. Merz, Jr., L. Zhang, F. D. Lewis, and D. M. Tiede, *Proc. Natl. Acad. Sci.* **103**, 3534 (2006)

CHAPTER 8

Conclusion

In this research we used a well-characterized 25bp DNA (or RNA) as the model system to study DNA (or RNA) condensation for the reasons described in chapter 1. We first studied the mode of inter-DNA interaction in Mg^{2+} solution since we've already accumulated knowledge about DNA-DNA interaction mediated by monovalent counterions (Na^+ , specifically) and had some preliminary experimental data with Mg^{2+} [1]. We tried to examine whether DNA molecules attract each other side-by-side or end-to-end. We made DNA “dumbbells” by capping both ends of dsDNA using a 4T loop and measured interactions between these dumbbells in Mg^{2+} and compare the result with that of regular dsDNA. Interestingly, we observed that under the same ionic conditions, when regular DNAs attract, “dumbbell” DNAs repel. Considering that end-to-end stacking should be blocked when both ends are capped, we proposed that dsDNA molecules attract each other in the fashion of end-to-end stacking rather than side-by-side alignment mediated by Mg^{2+} . Mg^{2+} doesn't condense DNA in normal conditions (room temperature, 1 atm) since in order to achieve DNA condensation, side-by-side attraction is usually a requirement. This proposal is consistent with evidences from other experiments [2, 3] and it establishes the basis for further experiments.

Cobalt hexamine (co-hex), a small trivalent ion, is known to be a very effective DNA condensing agent. We are interested in indentifying the driven force of co-hex mediated DNA condensation. We carried out the parallel experiments on both dsDNA

and dsRNA to provide a clearer and more complete picture of nucleic acid condensation. One of the key differences between these two molecules is that dsDNA is usually in B-form while dsRNA adopts A-form. We propose that the configuration difference causes different behavior in condensation process. An unexpected result is observed—under conditions when DNAs condense, RNAs resist condensation while prior to the onset of condensation RNA stack end-to-end but DNA molecules do not. These results, together with theoretical models described in chapter 3, suggest that configuration does play a key role in condensation as discussed in detail in chapter 5. Similar experiments were performed on spermine-induced condensation. The preliminary data are shown in chapter 7—DNA condenses but RNA does not. According to SAXS measurements, RNAs repel strongly in spermine solution which indicates that RNA charges are not even neutralized by spermine. It is very likely that spermine cannot bind into the major groove of RNA. The research on DNA (and RNA) condensation further implies that biophysical methods to package DNA may not be directly applied in packaging RNA for therapeutic applications.

Following up with the condensation experiment, we would like to know how many excess counterions associated with DNA are needed to precipitate DNA from solution. The problem was solved by using the technique of multiple-energy ASAXS (mE-ASAXS). The scattered intensity is a function of anomalous scattering factor which in turn is a function of X-ray energy. The coefficients obtained by linear regression of the scattered intensity with respect to scattering factor are used to derive the excess number of ions bound to DNA. The excess number of co-hex bound with DNA is around 6. Under the same ionic condition, the number of co-sep, another DNA condensing agent with valence +3, is also around 6 according to chapter 6. In terms of

electrostatics, co-hex and co-sep-driven DNA condensation is very similar but the condensing power of co-sep is twice as great as co-hex. Then we measure the hydration level of bounded co-sep and co-hex and observe that co-hex is more hydrated than co-sep, consistent with the case when they are mobile in solution. Consequently we propose that the hydration patterns affect the condensing power since hydration structure has been known as an important driven force for inter-DNA attraction. It is likely that the cost of water restructuring during the condensation process is reduced when co-sep is the condensing agent.

The key observations and conclusions are summarized below:

1. dsDNA molecules can attract side-by-side or stack end-to-end in solution. Short dsDNA molecules tend to stack end-to-end in solution mediated by small monovalent and divalent ions (Na^+ , Mg^{2+} , etc). The strength of this interaction is length-dependent.
2. dsDNA can be precipitated by co-hex, co-sep and spermine. This effect is stronger for longer DNA molecules and is associated with side-by-side attraction.
3. Co-hex binds into major grooves of both DNA and RNA. The stronger attraction between RNAs than DNAs in co-hex solution is due to end-to-end stacking. The configuration of nucleic acid affects the effectiveness of ion-induced condensation. This is confirmed by the results of spermine-induced condensation experiments.
4. The number of ions bound to DNA is similar for co-sep and co-hex but the condensing power of co-sep is much stronger. Therefore, both electrostatics

and hydration level of DNA-ion system have impact on the effectiveness of ion-induced DNA condensation.

5. Nonlinear Poisson Boltzmann theory can be used to model monovalent and divalent ions distribution around DNA. But more complicated models are needed to solve trivalent ion distribution due to the strong ion-DNA binding and ion-ion correlation.

REFERENCES

- [1] X. Qiu, L. W. Kwok, H. Y. Park, J. S. Lamb, K. Andresen, and L. Pollack, *Phys. Rev. Lett* **96**, 138101 (2006).
- [2] H. J. Limbach, and C. Holm, *J. Chem. Phys.* **114**, 9674 (2001)
- [3] J. Liu, A.-C. Declais, and DMJ Lilley, *J. Mol. Bio.* **343**, 851 (2004)

APPENDIX

How to run the code to obtain form factor (collaboration project with Baker group)

This is a readme file about how to use the program to generate the scattering profile (form factor) of 25bp DNA in salt solution with monovalent counterions using Debye formula. The program consists of three executive files—“read_5columndata.exe”, “write_solvent.exe”, and “saxs_calculation.exe”. All three files together with other example files (see below) are saved in the folder “code” (There are 7 files in total in this folder with three of them executive files and four example input/output files for test). A step-by-step protocol to compute form factor is provided below:

1. Get counterion concentration map from Nathan Baker’s group generated by APBS. This should be in the format of txt file with data of counterion in 5 columns. The first three columns are x, y, z, coordinates of the ion. The fourth column is concentration number and the fifth is the element of the ion.
2. This is a step depending on the computation power of the computer. To save running time I usually do it—manually delete those ions that are too far away from DNA molecule in the txt file. Those ions contribute very little in the scattering profile but will cause the running time of calculation much longer. One example of the number-reduced ion concentration map generated using radius 2 and dielectric constant 2 is “ion_r2d2.txt” which is about 116kB. This is the input for “read_5columndata.exe” for further number reduction.
3. Again this is another number reduction step similar to the previous one. You don’t have to worry about this huge computation power is provided. “read_5columndata.exe” will take the ion concentration map (for example, ion_r2d2.txt) as the input file and output a file of ions with concentrations

higher than certain predefined threshold. Those are the ions close to DNA molecule and will contribute most to the scattering profile. This threshold should be estimated from the original ion concentration map to include enough counterions around DNA but minimize the future computation time. Try and error works here and the threshold I use for the specific example provided in this appendix is 5. Run “read_5columndata.exe” with “ion_r2d2.txt” in the same folder and follow the instruction in the program and choose a name for the output file. One example is “output_r2d2.txt” in the “code” folder.

4. Build dummy hydration shell atoms around DNA molecule. Dummy hydration shell atoms are placed around DNA within certain predefined thickness. The region containing the hydration shell atoms will be a rectangular block eliminating the original cuboid defined by the DNA molecule. The default thickness is 2Å and step of the grid is 1Å. The random placement is implemented later in the Debye formula calculation (see step 6 below). Right now, the whole region with predefined thickness will be filled with those hydration shell atoms with neighboring distance 1Å by running the executive file “write_solvation.exe”. This program will ask you for the border of the DNA molecule in x, y, z coordinates. The default values provided in the questions asked after the program runs are for 25bp DNA molecule of which the 5-column data are in the file “BDNA25_5c.txt”. There’s no input file for “write_solvation.exe”. The output file will contains 5-column data of hydration shell atoms placed in the region defined by the border of DNA and thickness input.

5. Combine the output files from step 3 and step 4 as well as the original DNA 5-column data file “BDNA25_5c.txt” into one single file. This will be the input file for the form factor calculation by Debye formula. One example file is “opt_r2d2_sol.txt” in the “code” folder.
6. Run the file “SAXS_calculation.exe” using the combined file from step 5 as the input file, “opt_r2d2_sol.txt” for example. A series questions will be asked after the program runs. Most of them are rather straightforward and the default values work very well. You’re welcome to use values other than the defaults and test the results. One of them asks for the coverage ratio of dummy hydration shell atom placement and the default is 0.8. The use of this value is to define how much percentage of the region will be filled by the HS atoms. The position will be selected randomly for each run. Choose a name for the output file which will contain two columns—the first column is the Q-value defined according to your answers to the questions in the program and the second column contains the corresponding unnormalized scattering intensity. The output file can be loaded into MATLAB and plotted out. One example is shown in figure 2.7.