APTAMER-BIOSPECIES INTERACTIONS AT THE SINGLE-MOLECULE LEVEL

by

BIN WANG

(Under the direction of Bingqian Xu)

ABSTRACT

Nucleic acid aptamers are relatively easy to be produced with low cost, and it maintains the activity with various chemical modifications. The studies on the structure-function relationship of aptamers provide deep understanding of the reaction mechanisms between aptamers and their target molecules.

We have investigated the specific single-molecule interactions between aptamers and toxic protein ricin. The ricin molecules were immobilized on Au(111) surface using a series of optimized chemical reactions to specifically expose the active domains of ricin molecule to the aptamers attached to atomic force microscopy (AFM) tips. The AFM topography images of ricin on Au(111) surface showed the morphology of this toxic protein at single-molecule level. The AFM recognition images revealed the binding areas on ricin A chain. It was also observed that the binding sites for aptamers were not interfered by the immobilization reactions for most of ricin molecules. The different locations of the ricin binding sites to aptamer and to antibody were also distinguished by AFM recognition images and interpreted by simulations. The affinity of aptamer to ricin proved to be comparable to the one of antibody.

The anti-ricin aptamer with its PCR primers was also used for the studies on the structure-function relationship of aptamer. This aptamer with longer sequence has more complex folding

structure and multiple binding sites. A Markov transition model was developed to illustrate the multiple binding sites and multiple unbinding energy barriers of this aptamer-ricin complex. The elastic properties of this aptamer during the unbinding reactions were studied with the dynamic force spectroscopy (DFS). The complex relationships among different unbinding pathways were connected by the mechanism model based on Hook's law and warm like chain (WLC) theory.

The aptamers without primers were further used in the biosensor developments with two label-free platforms: AFM and surface plasmon resonance (SPR). The affinity of aptamers to ricin measured by AFM and SPR were compared. The research findings can be used as a reference for the further development of label-free biosensor for pathogen and toxin detection in food matrices. The *Salmonella* Typhimurium sample solution was also tested in the AFM and SPR methods by an aptamer with special interaction with the outer membrane proteins of the bacteria, providing the potential application of detecting certain bacteria serotype.

INDEX WORDS: Atomic force microscope (AFM), Dynamic force spectroscopy (DFS), Topographical and recognition image (TREC), Aptamer, Ricin, Molecular dynamics, Docking, Bell's model, Markov chain model, Elasticity, Warm like chain (WLC), Polyethylene glycol (PEG), Surface plasmon resonance (SPR), *Salmonella*, biosensor.

APTAMER-BIOSPECIES INTERACTIONS AT THE SINGLE-MOLECULE LEVEL

by

BIN WANG

B.S. Peking University, China, 2001

M.S. Georgia State University, USA, 2008

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

APTAMER-BIOSPECIES INTERACTIONS AT THE SINGLE-MOLECULE LEVEL

by

BIN WANG

Major Professor:	Bingqian Xu
Committee:	Bosoon Park
	William Kisaalita
	Zhengwei Pan

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2015

DEDICATION

I dedicate this dissertation to my parents.

ACKNOWLEDGEMENTS

First, I would like to express my greatest gratitude to my major professor, Dr. Bingqian Xu, who has supervised my Ph.D. study at UGA with both financial support and academic advice. His support helped me successfully complete my Ph.D. study through many obstacles and difficult times. I would not have been able to achieve any of my research work and the Ph.D. degree without his help. The strict training and research skills I have obtained under Dr. Xu's direction will continue to benefit my future career.

Next, my thanks go to the supportive and knowledgeable committee members, Dr. Bosoon Park, Dr. William Kisaalita, and Dr. Zhengwei Pan. I appreciate all of them for their generous teaching and advising throughout my Ph.D. study and research work. Dr. Park generously provided me financial support and gave me considerable helpful advice for our collaborative projects. Dr. Kisaalita and Dr. Pan also kindly and continuously gave me very helpful advice and suggestions for my classes, research, and my life.

Moreover, I want to thank my former teammates, Dr. Anna Jagielska, Dr. Cunlan Guo, Dr. Jianfeng Zhou, Dr. Guojun Chen, Dr. Mengmeng Zhang, and Mr. Joseph Hamill from whom I have learned a lot. I also want to thank my current colleagues: Dr. Yangang Pang, Mr. Kun Wang, Mr. Zhichao Lou, and Ms. Yanan Zhang. My research benefited greatly from their friendly collaboration. Our teamwork are very productive, and I have learned from everybody. I also appreciate many in other research groups, from either the University of Georgia or other universities who generously gave me suggestions and offered help for my Ph.D. studies and personal life.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Single-Molecule Studies Using Atomic Force Microscope

Single-molecule studies include various methods and techniques developed for the measurements of individual molecules, which show different properties and behaviors from the bulk ensemble of molecules. The single-molecule methods can be based on the signal of either fluorescence or force.[1] The major fluorescence-based techniques used in single-molecule studies include total internal reflection fluorescence microscope (TIRFM) and confocal microscope, or the combination of these two, according to different time resolution of the biological process observed.[1] However, all of fluorescence methods require the labelling of the sample molecules, which may introduce various interferences to the properties of the molecules studied. The force-based methods include the atomic force microscope (AFM), optical trap (OT), and magnetic trap (MT). These techniques are label-free; therefore they have the advantage for the studies of conformational changes and structure-function relationships of biomolecules. Among these techniques, AFM is the method that can measure the corresponding morphologies and interactions of the sample molecule using one integrated platform.

AFM is one of the techniques that belong to scanning probe microscopes (SPMs). The scanning tunneling microscope (STM) is the basic of these techniques, and was first demonstrated in 1981 to provide atomic resolution of a conductive sample surface.[2] The second technique based on SPMs is the near-field scanning optical microscope.[3] The invention of AFM in 1986 expanded the SPM high-resolution imaging technique to electrically insulating surfaces. In the late 1980s,

AFM imaging had been already conducted in fluids. Subsequently AFM was used for measuring various biological systems, in either ambient or solution environments.[3]

All SPMs include the similar components. First, a probe is used to scan near the surface and detects certain physical quantity. A piezoelectric actuator controls the tip movement above sample surface. A feedback system will control the vertical separation distance between tip and sample surface. In the AFM, the probe is a sharp tip attached to a cantilever beam. It is usually made from Si, SiO_2, or Si_3N_4 monolithic structures by photolithographic process.[3] In order to realize certain functions, conductive or magnetic coatings can be applied to the tip. When the tip approaches the sample surface, the cantilever starts to bend according to the force applied on it. A laser beam and a position-sensitive photodiode are usually used to detect the bending of the cantilever, and transduce the photoelectron signal to electronic signal, and in turn into the AFM image (**Figure 1.1**).[3,4]

The AFM imaging can be operated in either contact mode or dynamic mode. In contact mode, the tip works like a stylus profilometer, where the tip continuously contacts with the sample surface.[2] The repulsive forces between the tip and sample surface causes the bending of the cantilever. The piezoelectric actuator can maintain a constant cantilever deflection using a feedback system when the tip scans the surface. Consequently, the detailed surface topography can be illustrated in AFM topography images.

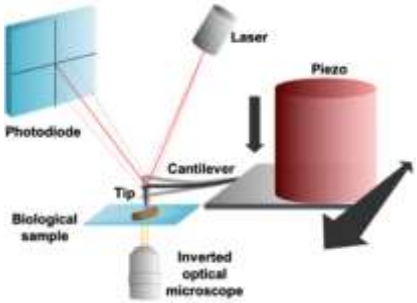

Figure 1.1 Schematics of the AFM instrument.

The dynamic mode was developed to improve two issues: (1) scanning under contact mode generates considerable shear force between tip and surface, which leads to damage of the tip or sample. (2) The increasing of sample-tip distance will be able to distinguish long-range forces such as electrostatic and magnetic forces from shorter ranged van der Waals and repulsive forces.[3] The amplitude and tip-surface distance in this mode can be adjusted to achieve either noncontact imaging, or to a slightly more aggressive intermittent contact mode, in which the tip contacts the surface briefly during each oscillation cycle. Therefore, shear forces are avoided and high resolution images can be obtained from soft biological materials. Recent developments have improved the resolution in noncontact AFM in solution.[5] The ability to acquire molecularly-resolved images in air and in fluid provides more opportunities for the AFM studies of biologically-relevant systems.

One of the widely used AFM modes is called Top MAC mode (magnetic alternating contact mode, Agilent Technologies). In this mode an oscillating magnetic field is applied above the magnetically coated cantilever and drives the cantilever to oscillate. Therefore, the tip only slightly interacts with sample surface in an intermittent contact (repulsive) or non-contact (attractive) regime.

1.2 Molecular Recognition Technique

In MAC mode AFM, one special technique is developed for "simultaneous Topography and RECognition imaging" (PicoTREC, Agilent Technologies), which gives researchers the ability to detect and map molecular binding events while simultaneously generating topographic images in real time.[6,7] Antibody or other biomolecules can be attached to AFM tips, so that other biological species on sample surface will interact with the modified tip and generate so called recognition signals in real-time AFM recognition images. The PicoTREC technique simultaneously provides

both topography and recognition images of the sample surface.[8] This technique has been widely used in the studies of single-molecule interactions, especially in the molecular recognitions of biomolecules.

During the scanning in PicoTREC experiment, the tip-surface interaction causes the amplitude of the oscillation to be reduced, as shown in **Figure 1.2** (a) (lower part of the oscillating signals), and generates the topographic image.[8] As the probe molecule on the AFM tip binds to and then unbinds from its targets on the sample surface, the amplitude in the top part of the oscillation is reduced and the unbinding force is detected.

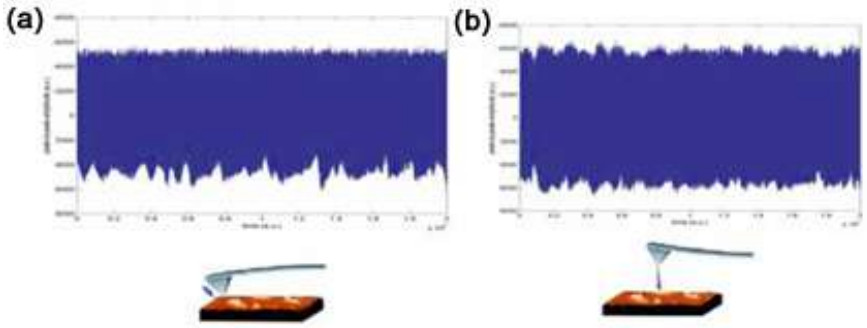

Figure 1.2 The generation of topography and recognition signals in AFM tip scanning.

Figure 1.3 shows how the topography and recognition signals are obtained from the data processing of tip oscillating during the PicoTREC experiments, and the example of corresponding topography and recognition images from published data.[8,9] The PicoTREC control box splits the upper and lower parts of tip oscillating signals and reconstruct the recognition and topography images, respectively.[9] The recognition signals in the image are the dark spots, which were generated by the interactions between ricin on sample surface and the aptamer on AFM tip.[9]

PicoTREC gives AFM the ability to chemically distinguish biomolecular at the single-molecule level. The technique can be used to analyze the components of heterogeneous samples and provides chemical information from the samples that can not be resolved from topographic images alone.[9]

1.3 Unbinding Force Measurements and Bell's Model

Another important AFM technique is the dynamic force spectroscopy (DFS). In DFS experiments, the tip is moving at a constant velocity (loading rate) in the z direction (perpendicular to the sample surface), so an external force will be applied to the tip (modified with probe molecule) when it approaches and withdraws from the sample surface. The technique can be used to measure the unbinding force between the probe molecule on the AFM tip and the sample molecule on the surface. The tip movements and their corresponding force-distance relationships are shown in **Figure 1.4**. The panel 1 in **Figure 1.4** (b) shows the state of tip withdrawing from sample surface with the protein-ssDNA binding with each other. **Figure 1.4** (b) panel 2 corresponds to state 2 where the first bond (non-covalent bond) between protein and ssDNA breaks, and the panel 3 illustrates state 3 where the second bond breaks.[10] The panel 4 in **Figure 1.4** (b) corresponds to the last withdrawing process where the protein on tip and ssDNA on sample surface don't have interaction any more.[10]

The entire withdrawing process will generate a force-distance curve illustrated in **Figure 1.4**. The force value under a certain loading rate can be directly measured by this method. However, the unbinding process measured here is a dynamic process under non-equilibrium state.[11] In order to extract the interaction and binding energy under equilibrium state, specific physical models have been developed, and one of the most widely accepted model is the Bell-Evans model.[12]

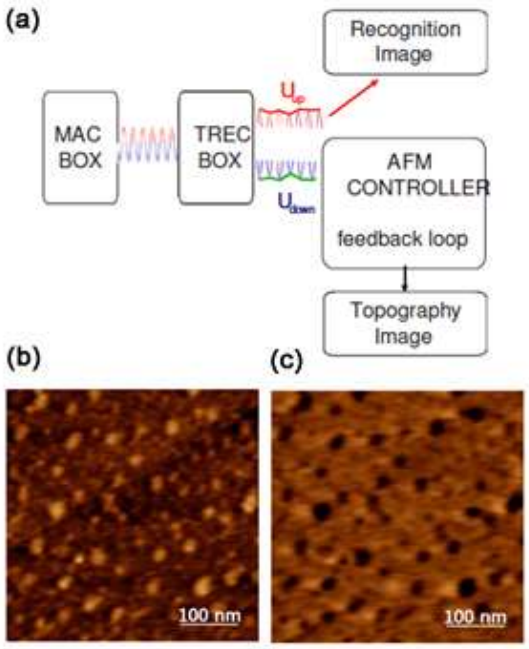

Figure 1.3 PicoTREC simultaneously records topography and recognition images.[8]

According to Boltzmann Ansatz, Bell assumed that the decrease of activation free energy for a reaction at zero force, ΔG*(0) is related to the applied force F:

$$\Delta G^*(F) = \Delta G^*(0) - F x_\beta \qquad (1.1)$$

where ΔG*(F) is the activation free energy under the application of a force F, x_β is the reaction coordinate corresponding to the separation between the bound and the transition state, projected along the direction of the applied force (**Figure 1.5**); x_β is assumed to be not affected by the external force.[11] Consequently, the dissociation rate $k_{off}(F)$, for an unbinding process under the application of a force *F*, increases with *F*:

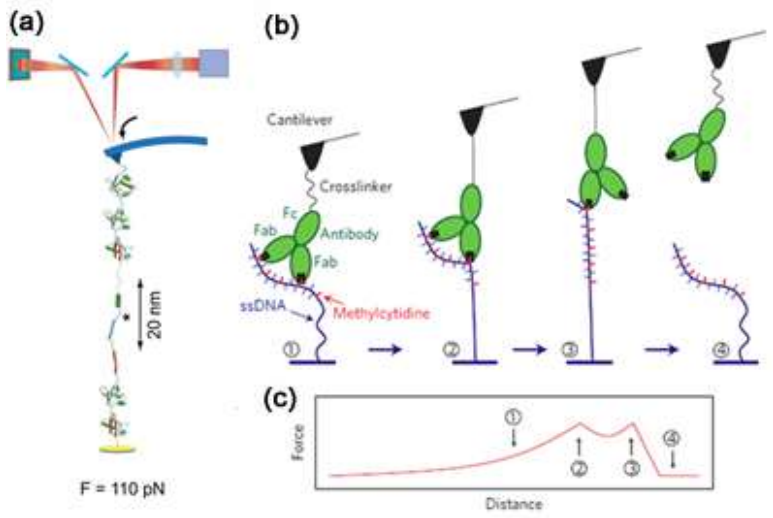

Figure 1.4 The schematics of DFS experiments.

$$k_{off}(F) = k_{off} e^{\frac{F x_\beta}{k_B T}} \tag{1.2}$$

where k_{off} is the dissociation rate at zero force (at equilibrium condition $k_{off} = k_{off}(0)$). Moreover, the lifetime of the complex is given by

$$\tau_{off}(F) = \frac{1}{k_{off}(F)} \tag{1.3}$$

Therefore, $\tau_{off}(F)$ will be shorter than the one of a spontaneous dissociation $\tau_{off}(0)$.

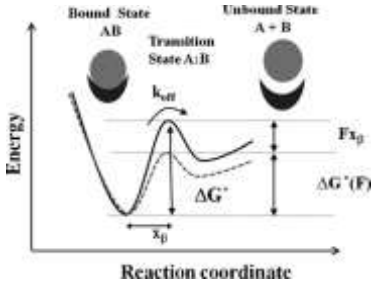

Figure 1.5 The schematic diagram of Bell's model.[11]

Based on Bell's work, Evans and Ritchie derived a model of the unbinding process in terms of overcoming a single barrier through the application of a time-dependent force $F(t)$, providing

the relationship between the unbinding force on the loading rate. Their assumptions include (i) the loading rate during a measurement is constant, and then the force increases linearly with time; (ii) a single couple of interacting biomolecules is investigated during the process; (iii) the rupture time is longer than the diffusional relaxation time and the occurrence of a rebinding process is neglected; (iv) the pulling coordinate is implicitly coincide with the reaction coordinate.[11] In general, the survival probability $S(t)$ of the unbinding process occurring at the time t has to satisfy the first-order rate equation with a time-dependent dissociation rate $k_{off}(F(t))$:[13]

$$\frac{dS(t)}{dt} = -k_{off}\big(F(t)\big)S(t) \tag{1.4}$$

and thus

$$S(t) = e^{-\int_0^t k_{off}(t')dt'} \tag{1.5}$$

Then, the probability distribution P(F) is:[13]

$$P(F) = \frac{k_{off}(F)}{R} e^{-\int_0^F \frac{k_{off}(F')}{R(F')}dF'} \tag{1.6}$$

The integration of equation (1.6) results in:[13]

$$P(F) = \frac{k_{off}}{R} e^{[\frac{Fx_\beta}{k_BT} + \frac{k_{off}k_BT}{x_\beta R}\left(1-e^{\frac{Fx_\beta}{k_BT}}\right)]} \tag{1.7}$$

The most probable unbinding force, F^*, at a fixed value of the loading rate, can be then obtained by calculating the maximum of the $P(F)$ distribution:

$$F^* = \frac{k_BT}{x_\beta}\ln(\frac{Rx_\beta}{k_{off}k_BT}) \tag{1.8}$$

or in this form for data treatment

$$F^* = \frac{k_BT}{x_\beta}\ln R - \frac{k_BT}{x_\beta}\ln\left(k_{off}\frac{k_BT}{x_\beta}\right) \tag{1.9}$$

Here, the most probable force F^* at a given loading rate R, the dissociation rate k_{off} at zero force, and the position of the energy barrier along the reaction coordinate x_β, can be obtained from the F^* vs. $\ln R$ plot, as shown below in **Figure 1.6**. Here (a) showed the single-barrier model, (b) is the

double-barrier model, and (c) is the multiple-barrier model. Column in the right: the corresponding F^* vs. LnR plots for their energy profiles. The F^* values are obtained from histograms of large amount of force-distant curves under different loading rates, then x_β can be calculated from the y-intercept, and k_{off} can be calculated from the slope.

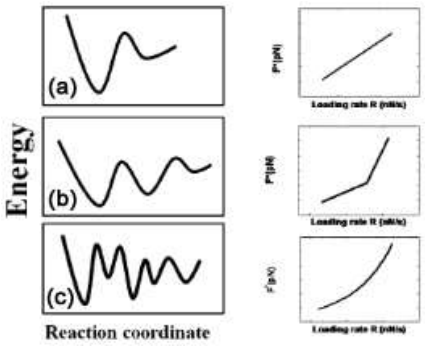

Figure 1.6 Schematic diagrams of the energy profile for a dissociation process of a biomolecular complex. [11]

1.4 Theories Involved in Molecular Modeling Methods

Molecular modeling refers to the simplified or idealized mathematical descriptions of the behavior of molecules and molecular systems, and currently this term is implicitly associated with the computer modeling.[14] Two widely used simulation methods in molecular modeling of biological molecules are molecular docking and molecular dynamics (MD). MD was first developed in late 1950s to study how small molecules reach equilibrium with time.[15] Since then MD has become an increasingly used tool in the structure and dynamics study of biological molecules.[15,16] Docking simulations focus on predicting the binding sites and complex conformations between two or multiple molecules.[14] Various algorithms have been developed to generate more accurate predictions of binding complexes, but docking methods still face lots of challenges their applications in real biological systems.[17,18]

According to physical theories, both quantum mechanics (QM) and classical mechanics can be applied to the simulated system. The QM or *ab initio* methods focus on simulations based on the first principles of QM.[15] Due to the large amount of atoms involved, QM methods are not practical for biological molecules. Under most circumstances, the molecular dynamics simulations of biological molecules are based on classical mechanics, a.k.a. molecular mechanism (MM). In this method, the motion of each atom is highly simplified to follow the Newtonian laws.[19] Some alternative approaches have been developed in order to embed QM into MM and called QM/MM hybrid methods. In general, the MD simulations of biological molecules such as proteins, nucleic acids, carbohydrates, and lipids mentioned later all refer to MM simulations.

In MM simulations, the first step is to prepare the sample system, which mimics the real atomic positions and movements in the real experiments, but at microscopic level.[19] Therefore, statistical mechanics are used to connect the microscopic simulations with the macroscopic properties of the system, such as pressure, temperature, and energy.[19] The microscopic state of a system is defined by the atomic positions and their momenta, which can be represented as coordinates in a multidimensional space called phase space. An ensemble is the collection of all possible points in phase space which satisfy same macroscopic or thermodynamic state but have different microscopic states. Some important ensembles include microcanonical ensemble (NVE), canonical ensemble (NVT), isobario-isothermal ensemble (NPT), and grand canonical ensemble (μVT). The NVE refers to an isolated system with a fixed number of atoms N, a fixed volume V, and a fixed energy E. The NVT refers to a system with a fixed number of atoms N, a fixed volume V, and a fixed temperature T. The NPT refers to a system with a fixed number of atoms N, a fixed pressure P, and a fixed temperature T. The μVT refers to a system with a fixed chemical potential μ, a fixed volume V, and a fixed temperature T. Most binding/unbinding reactions of biological

molecules can be treated as the systems of NPT ensemble or μVT ensembles. According to Ergodic hypothesis, the time average of thermodynamic states equals to the ensemble average of those thermodynamic states in macroscopic level.[19] Therefore, the MD simulation with a long time will eventually include all possible states, and the thermodynamic properties at macroscopic level, which are defined as the ensemble averages, can be calculated from the time averages of those states obtained from the MD simulation.

After a simulated system reaches its thermal equilibrium, the energy of this system is expected to be a potential function of its atomic coordinates.[20] The forces on individual atoms are related to the gradient of this function, so such functions are commonly referred to as "force fields."[20] A typical force filed can be expressed as the equation below:[20]

$$V(r) = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{torsions} k_\phi[\cos(n\phi + \delta) + 1] + \sum_{\substack{nonbond \\ pairs}} \left[\frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{AC_{ij}}{r_{ij}^{6}}\right] \quad (1.10)$$

Here the first term is the summary of energy in covalent bonds, the second term in bond angles, the third term in torsions, the fourth term in nonbond pairs (non-covalent interactions), which include electrostatic interactions (follow Coulomb' law) and Lennard-Jones potential (also called van der Waals term here). **Figure 1.7** illustrates the interactions corresponding to those terms.[20] The interactions forces applied to an atom in the force field are the derivative of the potential function to the position of that atom. The solid lines represent the covalent bonds connecting atoms 1 through 4, and dashed line represents the non-covalent interactions between atoms 3 and 5. Here b is the bond length for bond potential between atoms1 and 2; θ is the bond angle for the potential among atoms 2, 3, and 4; ϕ is the dihedral angle used for the torsion potential among atoms 1, 2, 3, and 4. The rij is the distant used for the nonbond pair potential between atoms 3 and 5.

Most force fields have different sets of parameters (k_b, b_0, k_θ, θ_0, etc.) for different types of molecules such as proteins, nucleic acids, and carbohydrates. These parameters are tested from real experimental data and some empirical methods. So far, many force fields and their corresponding software have been developed with their own specificities. Some widely used force fields include Amber, CHARMM, OPLS, and GROMOS.[20]

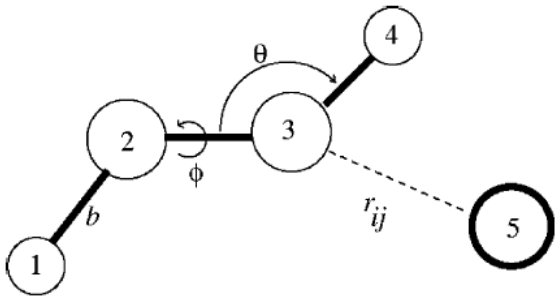

Figure 1.7 The schematic of interactions included in force field equation (1.10).

With the fast development of computer science and technology, the docking simulation has become an emerging method for the studies of structure-function relationships for biomolecules. Especially, the protein-ligand docking simulations have been used to facilitate drug discovery.[21] The methodologies and algorithms for the docking simulations are based on extensive studies on structure-function relationships of biomolecules and accurate 3-dimentional structural information of all molecules involved.

Early docking methods were based on "rigid-body" body assumption and highly simplified the proteins to rigid solid bodies. The fitness of geometric surfaces from the measured data of two molecules were used to rank predicted candidate complexes.[22] However, the rigid-body docking methods do not include the induced fit mechanism observed from experiments. In order to mimic the small but critical conformational changes in the induced fit binding processes, the structural flexibility of proteins have to be simulated. Various algorithms and software have been developed

to obtain more realistic results for the docking simulations, but more work is required in this area. Most docking software or webservers focus on the simulation of one type of binding reactions, such as protein-small ligand docking, or protein-protein docking, or protein-nucleic acid docking. The protein-ligand docking has been more extensively studied because more structures of proteins and their complexes are already known to the research comminutes. On the contrary, the studies on protein-DNA/RNA and DNA/RNA-ligand dockings require more investigations.

Currently docking simulations face lots of challenges from the very first step of docking. Before the docking simulation starts, the conformation and orientation of the ligand need to be determined, as well as the binding sites in the receptor molecule. During the docking simulation, different approaches have been developed for the challenge of molecular flexibility. Some of them include (1) systematic methods, which are based on incremental constructions, conformational search, and databases from experimental data; (2) stochastic methods include Monte Carlo methods and genetic algorithms; (3) simulation methods based on MD and energy minimization.[21]

At the end of docking simulation, the results are ranked by certain scoring schemes to obtain the most stable binding conformation, but the scoring functions are still far from reliable.[21] Currently, three types of scoring functions are used, force-field based functions, empirical functions, and knowledge-based functions. Evaluating different scoring schemes has become an important issue in the field of docking simulations. Overall, the structural-based scoring is not sufficient, more experimental data and human intuition have to be involved in order to predict the successful docking results. One special docking method is called “high ambiguity driven docking approach” (HADDOCK), which is based on the biochemical/biophysical data measured in NMR or mutagenesis experiments.[23] This docking method has the ability to conduct docking simulations among large proteins and nucleic acids molecules.

Amber is a suite of programs for MD simulations on biomolecules. It also refers to the empirical force fields that implemented in the Amber computer package.[24] The Amber force field can also be implemented in other molecular modeling programs, and other force fields can be implemented in the Amber programs. The current development version is Amber 14. Some information are required for the all Amber simulations: (1) Cartesian coordinates from X-ray crystallography or NMR spectroscopy; (2) topology of standard amino acids, DNA, RNA, carbohydrates and lipids, which are stored in Amber standard databases; (3) force field for all of atomic bonds, angles, dihedrals, and atom types in the system; (4) commands to control the simulation constraints, such as simulation time, temperature, pressure, etc. Some non-standard parameters can be generated with the help of other programs.[24] Moreover, explicit solvent such as water can be added into the biomolecule environment in order to simulate the system in solution.

Amber 11 has been used in the simulation of peptide folding, aptamer folding, and the binding energy estimation of carbohydrate-peptide interactions in the explicit water solvent and in the NPT ensemble.[9,25] The MD simulations provide helpful information for the mechanism studies of single-molecule interactions.

The main limit of conventional docking methods is that the search through various conformations of the complex geometry makes the calculation not effective, and may result in combinatorial explosion.[21,23] The advantage of HADDOCK is that the information on the interacting residues is based on experimental data and used as ambiguous interaction restraints (AIRs) to drive the docking.[23] The current NMR techniques have the ability to accurately define the interface between two residues by Nuclear Overhauser effect (NOE).[23] Therefore, the AIRs parameters have been derived from NMR data as an empirical algorithm, and provide the docking

scores according to the intermolecular energy calculated by related parameters, including electrostatic, van der Waals, and AIR energy terms.[23]

The AIRs is defined as an ambiguous intermolecular distance (d_{iAB}) between an atom m of an active residue i in protein A (m_{iA}) and an atom n of residue k (can be either active or passive residues) in protein B (n_{kB}):

$$d_{iAB}^{eff} = (\sum_{m_{iA}=1}^{N_{atoms}} \sum_{k=1}^{N_{res}B} \sum_{N_{kB}=1}^{N_{atoms}} \frac{1}{d_{m_{iA}m_{kB}}^{6}})^{-\frac{1}{6}} \quad (1.11)$$

where N_{atoms} indicates all atoms of a given residue and N_{res} the sum of active and passive residues for a given protein. Here the passive residues do not have direct AIRs to the partner protein but can satisfy the partner protein active restraints. A $1/r^6$ sum averaging is used to mimics the attractive part of a Lennard−Jones potential and ensures that the AIRs are satisfied as soon as any two atoms of the two proteins are in contact. A maximum value of 3 Å is set for the value of d^{eff}_{iAB} as a compromise between hydrogen−hydrogen and heavy atom−heavy atom minimum van der Waals distances for any atom-atom interaction in the docking simulation.[23]

Following equation (1.11), the HADDOCK program will search through all possible configurations around the interacting site defined by the biochemical and biophysical data (such as NMR or mutagenesis data) and to find the most favorable pair of interacting amino acids among the active and passive residues.[23]

HADDOCK webserver allows users to upload protein or nucleic acid structures, and manually set the active residues, the flexibility, and the other paramerters.[26] Multi-body docking with more than two molecules can be conducted in the webserver, which provide a powerful tool to study protein aggregations and complex protein-DNA/RNA interactions.[9] However, the initial determination of active residues and final selection of docking results still need human evaluations based on experimental data and basic biological knowledge. So far HADDOCK is the most

versatile tool for docking of large biomolecules, and the combination of Amber MD and HADDOCK docking have greatly helped the interpretations of AFM single-molecule experimental data.

1.5 Aptamers as Binding Agents

The advantages of aptamer over antibody and the examples are shown as follow. First, aptamers can be selected through relatively fast *in vitro* process against various target biospecies, such as organic biomolecules,[27] antibiotics,[28] human growth hormones,[29] toxic proteins,[30] cancer cells,[31] and infiltrating ductal carcinomas tissues.[32] This feature of aptamer can avoid the limit of using animals or cell lines for antibodies. Secondly, the relatively low-cost and simple way of producing large amount of aptamer by polymerase chain reaction.[33,34] Thirdly, aptamers are more stable than antibodies in harsh conditions such as high temperature or extreme pH, and have shown the reversible denaturation during the temperature and pH changes.[35] Therefore, aptamer is stable for long-term storage.

Moreover, the relatively small molecular weight (usually 10 to 15 kDa) make aptamers non-immunogenic, and have fast tissue penetration and rapid systemic clearance.[36] Literatures reported that aptamers had been modified with various biochemical groups in order to obtain certain clearance rate *in vivo* in the range of a few minutes to several hours.[37,38] This feature make aptamer a versatile binding agents for the *in vivo* studies. These four major advantages of aptamer over antibodies as binding agents have been reported in many review papers.[3,39,40]

The major disadvantage of aptamer is that the complex structure-function relationships of aptamer folding structures make it difficult to predict the aptamer binding affinity and specificity. Some early studies have been focused on aptamer folding structures and their structure-function relationships.[41,42] However, many challenges still exist for further improving the specificity of

aptamer to a certain biospecies. The aptamer with highest specificity is generated from the pool of random sequence, which may not include the sequence that has the most specific interaction with the target molecule. Therefore, the aptamer selected from this pool will have high affinity to the target, but may also have high affinity to other biospecies in the sample matrix of detection measurements, so the aptamer with higher affinity may be less specific to the target biospecies.[42] Another disadvantage is that the selection and purification of the aptamer need to be carefully conducted under optimized conditions. The current aptamer design and selection are based on the polymerase chain reaction (PCR), which is more dependent on the training than the clear understanding of aptamer binding behaviors.[41,43] Currently, PCR and purification techniques are in fast developments, but the reproducibility is still an issue for aptamer selection and amplification.

1.6 The Objectives and Organizations of This Dissertation

1.6.1 The Objectives of This Dissertation

This project will combine single-molecule techniques such as AFM with computer simulations to study the structure-function relationships of aptamer molecules to their target biospecies. The more specific objectives include the AFM single-molecule recognition of toxic species (i.e. ricin), the investigation of reaction mechanisms of aptamer-protein reactions with the help of computer simulation, and the development of aptamer biosensor for the label-free detection of ricin.

We believe that the combination of AFM-DFS and computer simulations should provide more detailed information for the complex unbinding interactions involved in the biomolecules. After the analysis of the experimental results, it is possible to find some general trends or relationships that can contribute to the theories in the single-molecule studies, and improve the

reaction models and data treatments methods to obtain more consistent results compared with experiments. Besides, the approaches used in aptamer-ricin study can be expanded to other systems such as aptamer-cell membrane interactions. This work will also help the rational designs and selections of aptamers for biosensors and drugs.

1.6.2 The Organizations of This Dissertation

The beginning of this dissertation is the introduction of basic theories and techniques used in the following chapters, and the objectives of this dissertation, which are all included in Chapter 1. Especially, the background of single-molecule theories and experimental methods are introduced, aptamers and their applications as the probe molecules in detection of proteins and bacteria will be studied in later chapters. In Chapter 2, the surface modification methods used to immobilize ricin on Au(111) surface are discussed, and the high resolution images of ricin molecules are studied with the help of molecular modeling methods. In chapter 3, the specificity and affinity of aptamer-ricin interactions are further tested using DFS and Bell's model. Chapter 4 will demonstrate the reaction mechanisms of aptamer binding to ricin. The detailed reaction models were revealed by the biophysical model we developed for the aptamer-ricin interaction at single-molecule level. Chapter 5 will focus on the elastic properties of aptamers during their unbinding reactions. Therefore, Chapter 1 to 5 cover the fundamental biophysical studies on the complex interactions between aptamer and toxic protein ricin. In Chapter 6, the research work is focused on the developments of label-free SPR biosensors for label-free and real-time detection of ricin sample in solution. After these two chapters on the applications of aptamers, the summary of this dissertation and the outlook of the future developments in this research field are presented in Chapter 8. Overall, this dissertation is written to present the wide range of research work on the structure-function of aptamers from their reaction mechanisms to biosensor applications.

CHAPTER 2

HIGH-RESOLUTION SINGLE-MOLECULE RECOGNITION IMAGING OF THE MOLECULAR DETAILS OF RICIN-APTAMER INTERACTION[1]

[1]B. Wang, C. Guo, M. Zhang, B. Park, B. Xu. 2012. *J. Phys. Chem. B*, 116:5316-5322.

2.1 Introduction

Protein-nucleic acid interactions play important roles in the physiological behaviors of living cells and have drawn increasingly attentions in recent years.[44, 45] The structure-function relationships of biomolecules (such as proteins, DNAs, and RNAs) are the keys to understand their properties and interactions. Ricin is one of the most potent naturally occurring toxic proteins and 500 μg of ricin can kill an adult.[46] According to the study on ricin toxicology, ricin A chain acts as a glycosidase that removes one adenine residue from the loop of the 28S rRNA, causing this rRNA to lose its function during protein synthesis.[47] Therefore, the study of the interactions between ricin protein and other biospecies has potential impacts on both fundamental research and biomedicine applications. Various techniques have been used to investigate function of ricin molecule and its interactions with antibody or nucleic acids. X-ray crystallography were used to study the active residues of ricin to the ribosome.[48,49] Other techniques including competition assay,[50] electrophoresis,[30] antibody array,[51] surface plasmon resonance (SPR),[52] surface-enhanced Raman scattering (SERS),[53] nanopore,[54] nanoparticle,[55] molecular imprinting,[56] and atomic force microscopy (AFM)[57, 58] have also been used for the study of the ricin toxicology or ricin detection. AFM is one of the techniques that is capable to provide high resolution single-molecule images of proteins and nucleic acids in aqueous solution. Therefore, AFM is widely used in biological science to study the biophysical properties of different biospecies.[11] For most protein-nucleic acids complexes, the conventional experimental methods, such as X-ray crystallography and nuclear magnetic resonance (NMR), require numerous efforts and cost to obtain their detailed structural information. However, the AFM images can provide detailed information for this type of biomolecular interactions with relatively low cost and minute sample consuming. Although AFM single-molecule recognition techniques have been used on the studies of protein-ligand and

protein-protein interactions, the detailed information of the molecule structures and interactions is still in great need and deserve more careful investigations.[57, 59, 60]

Recently, aptamer has been used to successfully study ricin-nucleic acid interactions.[30,33,34,50,60,61] Aptamers are versatile binding reagents that show interactions with wild range of targets such as small molecules, lipids, proteins, carbohydrates, viruses, and cells.[62] Some aptamer researches have even led to medical applications and FDA approved clinical trials.[62,63] The DNA aptamers have relatively lower cost and are more stable than antibody. Therefore, the biosensing and molecular recognition using DNA aptamers have become an emerging research area. Especially, for the study of interactions between proteins and nucleic acids, aptamers can be used as analog models and probes.

In order to keep the biospecies active and mimic the physiological conditions of their interactions, the probe molecule in AFM experiments needs to be attached on the tip surface by a string of linker molecule, mostly heterobifunctional polyethylene glycol molecule. The target molecule is immobilized on sample substrate, such as Au (111) coated on mica, with the help of crosslinker molecules.[64,65] The surface immobilization methods play important roles in this type of study because the chemical reactions used to immobilize the target molecules determine their conformations on the substrate and in turn influence the interactions between probe and target molecules.

Here we use the DNA aptamer with specific affinity to ricin as a probe to study the ricin-nucleic acids interactions. The linker molecule used here is the lipoic acid-*N*-hydroxysuccinimide (LA-NHS) ester. It was used to bond covalently with the lysine residues on ricin surface.[9,58] We then immobilized ricin molecules on Au(111) surface, and tested the activities of their binding sites to aptamer.[58] The AFM molecular recognition experiments were conducted to compare the

recognition signals generated by anti-ricin aptamer and anti-ricin antibody, respectively. The results are shown in Results and Discussion section. When the ricin molecule is immobilized on Au(111) surface with the lysine-NHS reaction, the reactions of the nine different lysine residues on ricin surface are supposed to determine which side of the ricin molecule will be contacting the substrate and which side will be exposed to bulk solution (**Figure 2.1**). We used this NHS ester reaction to control the conformations and orientations of ricin molecule on the gold substrate, and the high resolution AFM images showed the detailed morphologies of these ricin molecules reacting with different lysine residues. The analysis of these nine conformations with AFM topography images are shown later in the section of Results and Discussion.

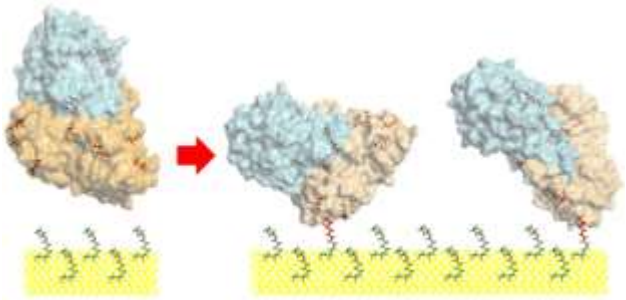

Figure 2.1 Schematics of ricin immobilization on Au(111) surface.

Compared to the high resolution structural information obtained by X-ray crystallography and NMR, current AFM techniques have difficulty to show the atomic details for biomolecules due to the resolution limits. Therefore, the AFM experimental data need help from biological theories and computer simulations to explain the observed phenomenon. On the other hand, simulations also need experimental data to improve their models so that the simulation results will have more meaningful impacts to the real world. Recently, computational simulations were increasingly used to study aptamers and their interactions with other molecules.[66-68] Here we used computer simulations to verify the phenomenon and proposed mechanisms of our AFM

recognition experiments. The binding sites on ricin and aptamer were predicted by molecular modeling and tested by blocking experiments. The molecular modeling methods used in this study include molecular dynamics and molecular docking. They provide helpful guides to study structure-function relationship of aptamer binding because the structural information of the aptamer and its binding conformation with ricin are not available. Especially, the docking and prediction of binding sites helped explain AFM molecular recognition signal and the different ricin conformations observed on Au(111) surface.

2.2 Experimental Details

The aptamer sequence was obtained from literature and purchased from Integrated DNA Technologies (Coralville, IW, USA), and the 5' terminal was modified with a short hydrocarbon linker and amine group.[30] The ricin sample molecule was purchased from Vector Laboratories (Burlingame, CA). The linker molecule HS-PEG-COOH (M.W. 2000) was purchased from Creative PEGWorks (Winston Salem, NC, USA). N-hydroxysuccinimide (NHS) was purchased from Sigma-Aldrich (St. Louis, MO, USA). The 1-(3-dimethylaminopropyl)-3-ethylcarbodimide hydrochloride (EDC) was purchased from Flucka Chemicals (Sigma-Aldrich, St. Louis, MO, USA). LA-NHS was synthesized in our lab. Phosphate buffer saline (PBS, pH 7.2) was purchased from Pierce (Thermo Scientific, Waltham, MA, USA). Triplet DI water was further purified by a Barnstead Nanopure Diamond Laboratory Water System (18 MΩ).

The gold-coated AFM tip was chemically modified with the HS-PEG-COOH via Au-S bond. Consequently, the amine-modified aptamer was attached to the carboxyl group of the PEG molecule with the help of NHS/EDC solution. The reactions involved in this process are shown in **Figure 2.2**. The thiol group of the bifunctionalized PEG linker molecule formed gold-sulfur bond with the gold film on the tip surface, and the carboxyl group at the other end of PEG reacted with

the amine modified aptamer to form the amide bond. The amine reaction was activated by EDC/NHS solution at pH 7.2. The PEG polymer was used here to provide enough freedom for the attached aptamer molecule so that the possible binding sites of the aptamer will not be interfered during the scanning.[69] The ricin molecules were immobilized on Au(111) surface by the reaction between lysine residues on the ricin surface and LA-NHS molecules on the gold surface. Reactions involved in this process were reported in a previous publication.[58, 70]

AFM imaging was conducted in TopMAC mode with the additional module of PicoTREC (Agilent Technologies, Santa Clara, CA) to obtain the topography images of ricin in PBS pH 7.2 buffer solution. In addition, the interaction between aptamer and ricin during the AFM scanning generates recognition signals are shown in the recognition images.

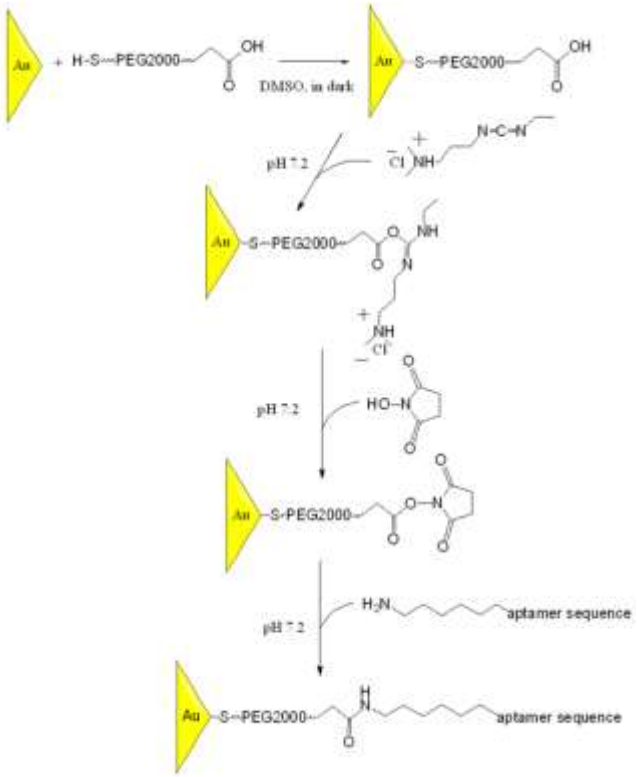

Figure 2.2 Reaction schemes used for AFM tip functionalization.

To test the interaction/binding specificity of the aptamer with ricin, blocking experiments were also conducted. The 1.0 μM aptamer solution in PBS pH 7.2 (blocking solution) was injected into the flow-through liquid cell during the AFM imaging, and the scanning was focused in the fixed area. Therefore, the images on the same scan area before and after the injection of blocking solution were compared. The fading of recognition signal showed that the ricin molecules on the Au(111) surface interacted with aptamer molecules in the blocking solution instead of the one on the AFM tip.[9] We also conducted the blocking of recognition signal of ricin by injection of 1.3 μM anti-ricin antibody solution in PBS pH 7.2 into the flow-through liquid cell. In this antibody blocking experiment, the AFM tip was still modified by the aptamer. Therefore, the aptamer and antibody should not interfere with each other and they should both be able to bind to ricin. The AFM recognition images showed the continuous recognition signals during the antibody blocking experiment, which proved that the antibody and aptamer have different binding sites. The details of the experimental results are shown in the section of Results and Discussions.

The secondary structure of the anti-ricin aptamer (**Figure 2.3** (a)) was predicted by Mfold webserver.[71] The 3-dimensional folding structure of the aptamer was generated using nucleic acids builder (NAB), which is included in Amber 10 molecular dynamics package.[72,73] The aptamer molecule was considered as an isobaric-isothermal ensemble (NTP), and the folding dynamics was conducted in generalized Born (GB) implicit solvation model to predict a structure with lowest energy (**Figure 2.3** (a) cartoon representation). The entire simulation time for aptamer folding is 20 ps. The ricin structure was obtained from PDB 2AAI.[74] The human IGG1 was modified according to literatures as a homological model to obtain ricin-antibody complex structure (**Figure 2.3** (b)).[75,76]

The predictions of binding sites were based on literatures and conducted by molecular docking simulation (HADDOCK web server).[26,77] The binding conformations of ricin-aptamer and ricin-antibody complexes are shown in **Figure 2.3**. The docking results demonstrated that ricin has different binding sites for the aptamer and antibody, although they both locate on ricin A chain.[9] Molecular modeling method is also used to analyze the lysine-NHS ester reaction sites. The results (in the next section) show that the binding sites of ricin to aptamer and antibody were still available after the immobilization. The method of combining AFM experimental images and simulation structures is used in this paper to explain the different ricin conformations observed on Au(111) surface and the high efficiency of molecular recognition. The results from AFM experiments and simulation are coherent with each other at the single-molecule level. However, we notice that the predicted ricin-aptamer and ricin-antibody complexes were based on their structures without other constraints, such as chemical modifications, geometry and surface properties of the Au(111) substrate, and the interference of the AFM tip. Therefore, the simulation results only represent the most stable conformations of these molecules that freely exist in solutions, and their real conformations after chemical modifications on the gold surface may be different from the results shown here. The simulation of this type of protein-DNA interactions on certain surface/interface is a very interesting topic that is worth further investigation.

2.3 Results and Discussion

In order to probe the aptamer-ricin and antibody-ricin interactions, we first immobilized ricin molecules on the Au(111) surface via one of ricin's nine lysine residues reacting with LA-NHS. To probe the details of interactions, it is important to know where the binding sites to aptamer or antibody locate on the ricin molecule surface. These locations of binding sites are deter-

mined by ricin conformations and orientations on substrate. Inside the ricin molecule, we established three axes to help locate the nine lysine residues, the predicted binding site to aptamer, and the predicted binding site to antibody (**Figure 2.3** (a)). The x-axis (black), y-axis (purple), and z-axis (red) were established inside the ricin molecule. Three axes and their relative positions to A chain (cyan) and B chain (orange) are labelled. The nine lysine residues are in red, the predicted binding residues to aptamer are in blue, the predicted binding residues to antibody are in yellow. These axes form right angles with each other in the 3 dimensions, approximately. The longest axis (z-axis, red) is around 9 nm across the molecule, the second longest axis (x-axis, black) is around 6 nm, and the shortest axis (y-axis, purple) is around 5 nm, which are measured in molecular visualization software PyMOL (version 0.99rc6). These axes are not used to show strict coordination but to label the 3-dimentional orientation of the ricin molecule. All the nine lysine residues were found to locate on the ricin surface and are highlighted in red. The binding sites to aptamer (blue) and antibody (yellow) both locate on ricin A chain. The relative positions of all these reacting residues and binding sites to the three axes are shown in **Figure 2.3** (b). We will use these three orientation axes as guiding lines to label ricin molecules and compare them with the ricin AFM images.

Figure 2.4 shows the representative AFM topography (**Figure 2.4** (a)) and recognition (**Figure 2.4** (b)) images. The image size is 1000 nm by 1000 nm. Nine individual ricin molecules (#1 through #9) were selected to represent different ricin conformations predicted by molecular modeling (**Figure 2.5**). The average size of ricin molecule in this figure is determined to be around 20 nm, which is consistent with previous study using antibody.[58] The recognition image shows the corresponding recognition images of those ricin molecules, which demonstrated that the

aptamer-modified AFM tip interact with ricin molecules with the concentration as low as 20 pg/mL.

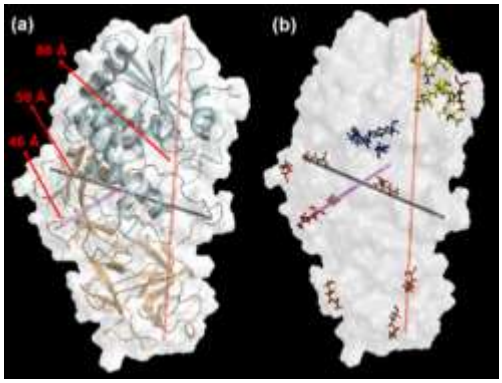

Figure 2.3 Three axes used to illustrate the locations of different active residues on the ricin molecules.

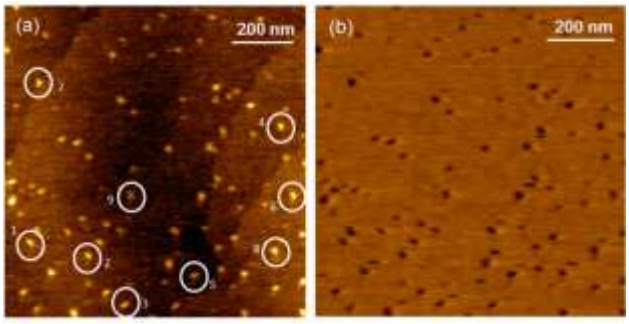

Figure 2.4 AFM topography image and recognition image of ricin molecules on an Au (111) surface.

As shown in **Figure 2.4**, nine individual ricin molecules were highlighted and labeled by 1 to 9 to represent the nine predicted conformations generated by NHS-lysine reaction. This reaction between NHS ester molecules and lysine residues on the protein surface has been widely used for the immobilization and detections of proteins.[78] The lysine residues tend to locate on the outer surface of proteins because lysine is hydrophilic. This property of lysine plays an important role in protein tertiary structures in aqueous solutions and in turn influences the protein functions. Therefore, the different reacting lysine residues on the surface of individual ricin molecule plays

important role on the ricin conformation on the gold substrate. The nine ricin molecules shown in **Figure 2.4** (a) correspond to nine probable binding conformations of ricin to the LA-NHS linker molecules on Au (111) surface. The recognition image in **Figure 2.4** (b) shows that almost all ricin molecules in the topography image have the corresponding recognition image in **Figure 2.4** (b). The recognition efficiency is 98% (100 out of 102), which is also similar to the high efficiency of recognition images generated by antibody.[58] These high efficiencies of recognition signals prove that the ricin A chains were able to directly expose their binding sites to aptamer and to antibody in the bulk solution, otherwise the aptamer or antibody on the AFM tips cannot approach the ricin binding sites during the scanning. Although ricin can bind to NHS ester with different lysine residues and form different conformations and orientations on Au(111) surface, the binding sites to the aptamers were available under most circumstances.

The detailed relationship between each ricin conformation and its LA-NHS reaction can be seen in **Figure 2.5**, which shows the comparison of zoom-in topography images and molecular modeling structures of ricin molecules on Au(111) surface. The conformations #1 through #9 correspond with the ricin molecules #1 through #9 in **Figure 2.4** (a). The image sizes are 40 nm × 40 nm for all conformations. The **Figure 2.4** (a) showed the AFM topography images and predicted conformations shown in the top view, and **Figure 2.4** (b) are the 3-D topography images and predicted conformations of ricin #1, #2, and #3 shown in the side view. The white dashed line crossing each ricin image represents the approximate position of the z-axis inside the ricin molecule (explained in **Figure 2.3**). The simulation structures have been rotated to mimic both the top view and side view of ricin molecules in the AFM topography images.

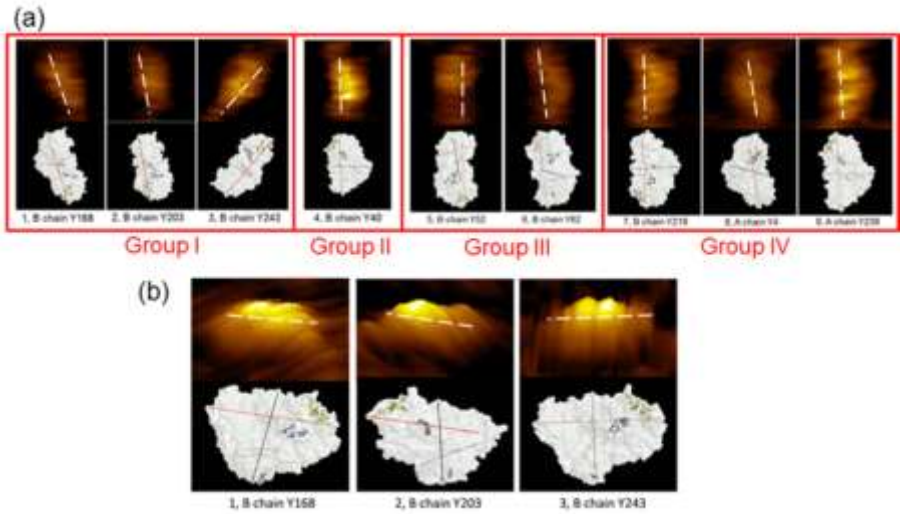

Figure 2.5 The zoom-in images of nine ricin conformations on Au(111) surfaces.

To check the detailed structural information of these nine ricin conformations, we compared the 3D topography images of these nine conformations with the side view of the simulation structures. In **Figure 2.5** (b), the ricin simulation structures have been adjusted to show the predicted side view of each conformation above the gold substrate and LA-NHS layer. The 3D topography images show the relative positions of ricin A chain and B chain along the AFM z-axis (perpendicular to the substrate surface). The simulation structure of the same ricin conformation is put next to the 3D topography images, so that the predicted binding sites to aptamer and ricin can be seen in the comparison. When NHS ester bond with ricin B chain Y168 and Y243, the ricin A chain and B chain were at approximately same height level, while A chain was at higher level than B chain if ricin bond with B chain Y203. This minor change is very difficult to be distinguished in the 2D topography images.

We classified the nine ricin conformations in **Figure 2.4** into four groups since some of the topography images of the ricin molecules have very similar outlines. The simulation structures show that the similarity of the conformations in each group is based on the locations of the reacting

lysine residues. In group I, the conformation 1, 2, and 3 react with lysine B chain Y168, Y203, and Y243, respectively. These lysine residues are close to each other on the ricin surface. In group II, the reacting residue is B chain Y40 for conformation 4. In group III, the reacting residues of conformation 5 and 6 are B chain Y52 and Y62, respectively. These two lysine residues also locate very close to each other. In group IV, the reacting residues of conformation 7, 8, and 9 are B chain Y219, A chain Y4, and A chain Y239, respectively. They are not close to each other, but their relative locations to the three axes are similar. Therefore, it is very difficult to distinguish their topography images among each other in the same group.

From **Figure 2.3**, we also found that most of the lysine residues on the ricin molecule are located in ricin B chain, so the B chain is the most probably region bond to the NHS ester molecule on gold substrate. Therefore, ricin A chain is supposed to contact with the bulk solution and in turn interact with either aptamer or antibody on the AFM tip.[9, 58] For all conformations in group I through III, the ricin A chain is above the ricin B chain. For conformation IV, A chain and B chain are at the same height level. Therefore, almost all of ricin showed recognition signals in **Figure 2.4** (b). For these conformations not showing recognition signals, the possible reason could be that the binding between the aptamer on AFM tip and ricin on the substrate might be disturbed by the aqueous environment or the thermal noise.

To determine the distributions of the four conformations, we also did statistical analysis on the ratio of these conformation groups (I to IV) and the recognition efficiency of these groups. The counting of different groups from four AFM topography images is shown in **Figure 2.6**. Here the ricin conformations determined by nine possible lysine reactions were classified into four groups (group I through IV). The “unknown” group represents the remaining ricin conformations that cannot be distinguished and classified into any other groups. The abundance of group I, II, III, and

IV were 33.7%, 17.5%, 9.5%, and 7.9%, respectively. The relative high abundance of unknown group (31.4%) represents the experimental errors and limitations. The unknown group represents those images that cannot be distinguished or classified due to molecules overlapping, tip broadening, and compression effect.[79] If we neglect the group of "unknown", the most abundant conformations were from group I, which includes three conformations with ricin B chain attached to LA-NHS. This trend is consistent with the proposed binding mechanism. The relatively large amount of conformations in the group of "unknown" represents the limitation of the AFM technique we were using in the recognition experiments[9]. However, the overwhelming majority of the recognition of different ricin conformations show the capability of this AFM experiments and theoretical simulation approach in determining the molecular details of structures and interactions.

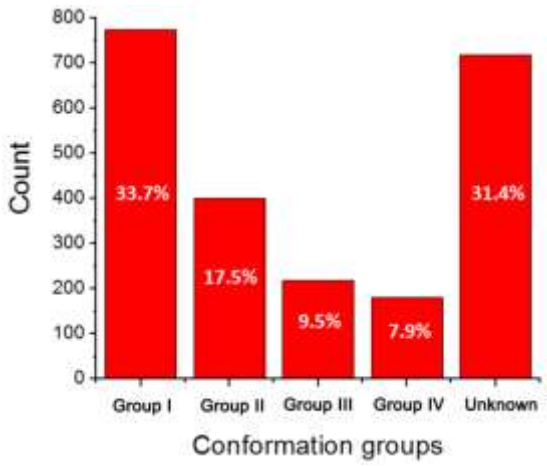

Figure 2.6 The distribution of different ricin conformations on Au(111) surfaces using NHS ester as linker molecule.

When comparing this experimental result with the molecular docking result, the AFM recognition experiments provide very helpful information on the binding sites of ricin to aptamer and antibody. Although we cannot distinguish the detailed binding residues and sequences using

AFM, the recognition images and blocking test help us confirm the availability of binding sites to aptamer and antibody on ricin A chain after the immobilization. Moreover, the difference of these two binding sites was verified by blocking experiments. This experimental result is also very useful for the validation of simulation results, especially, when other experimental method is not available now to directly test the ricin-aptamer and ricin-antibody binding conformation.

2.4 Conclusion

We combined computational simulations and AFM recognition imaging to study the ricin conformations and its interactions with aptamer and antibody. The method used to immobilize ricin on Au(111) surface makes it possible to obtain different ricin conformations due to specific covalent bonding between the LA-NHS ester and the lysine residues on ricin surface. The ricin molecule conformations when reacting with different lysine residues were distinguished by AFM topography images. The simulated ricin structures confirmed that the different ricin conformations on Au(111) surface were caused by different lysine residues reacting with the NHS ester. In addition, the ricin binding sites to aptamer and antibody were predicted by simulation methods and tested by blocking experiments. The binding sites for aptamer and antibody were not interfered by the immobilization method. This study demonstrated that the combination of experimental methods and molecular simulation methods could be used in the binding studies of similar biospecies.

CHAPTER 3

FOLLOWING APTAMER-RICIN SPECIFIC BINDING BY SINGLE MOLECULE RECOGNITION AND FORCE SPECTROSCOPY MEASUREMENTS[1]

[1]B. Wang, C. Guo, M. Zhang, B. Park, B. Xu. 2012. *Chem. Commun.* 48:1644-1646.
Reprinted here with permission of publisher.

3.1 Introduction

The immobilization of aptamers to the gold substrate is the basic step for any further studies. After this step is tested as reliable, next step is to determine the affinity of aptamer to ricin and investigate its reaction mechanism. In this mechanism study, single-molecule techniques are powerful tools to reveal detailed structure-function relationships and provide critical information that may be buried in the average value based on the behaviours of large amount of molecules in bulk solution measurements. This chapter will focus on the AFM experiments and data analysis of aptamer-ricin interactions. Moreover, molecular dynamics and docking simulations provide useful structural information for the deep understanding of aptamer reactions.

AFM with sharp tip and fine cantilever can provide detailed morphology and mechanical measurement for aptamer-ricin complex at single-molecule level. In our AFM molecular recognition and dynamic force spectroscopy (DFS) experiments, aptamer was attached on AFM tip as probe molecule to bind the target molecule of ricin on Au(111) surface (**Figure 3.1**). The target molecule ricin is one of the most cytotoxic proteins and need to be monitored strictly.[14,15] We used the same methods as described in our previous published papers to synthesize the reaction reagents and immobilize ricin molecules on Au(111) surface with the linker molecule lipoic acid-*N*-hydroxysuccinimide ester (LA-NHS) .[58, 70]

3.2 Experimental Details

3.2.1 AFM Measurements

The first step was to modify Au(111) surface with LA-NHS. Then ricin molecules were put on the surface so that the primary amine on lysine residues reacted with the LA-NHS to form stable amide bond. Most of the lysine residues are located on ricin B chain, so for most ricin conformations, the ricin B chain is in close contact with gold surface, while A chain is on the top

of B chain and available to contact to the probe molecules. Therefore, the lysine-LA-NHS reaction, which effectively constrained the ricin conformation, led to the always available binding sites in ricin A chain to its probe molecule.

3.2.2 Simulations of Aptamer and Proteins

It is very difficult to obtain the crystal structures of aptamer and ricin-aptamer complex, so we used molecular modeling and simulation to help us understand their structures and possible conformations during the AFM recognition process. The ricin protein structure we are using is obtained from protein databank (PDB), with the PDB code of 2AAI.[74] The small sugar ligands and water molecules were removed from the structure. Then this ricin structure was minimized using molecular dynamics software AMBER 10 for 500 steps of steepest descent minimization and followed by 500 steps of conjugate gradient minimization in vacuum.[72] The sequences and secondary structure of aptamers was obtained from the reference.[30] Then a NAB program (included in AMBER 10 package) was used to generate possible 3D folding structures of this aptamers (**Figure 3.1**).[80] The antibody used in the AFM experiment is from goats, but its PDB is not available, so we use one PDB file of human IgG1 as a homological structure. The initial PDB file was obtained from the literature, and then the amino acid residues in complementarity determining regions (CDRs) were replaced according to literatures so that the CDRs are specific to ricin.[75,76,81] This replacement is done by Sirius visualization software (version 1.2, University of California San Diego). Then the new model was minimized using AMBER 10 for 500 steps of steepest descent minimization and followed by 500 steps of conjugate gradient minimization in vacuum (**Figure 3.1** (c)). In **Figure 3.1**, the aptamer folding structure is colored in blue, and the active residues to ricin are colored in red, the antibody structure with two heavy chains are in yellow and blue, the two light chains are in orange and black, and the active residues are colored in purple.

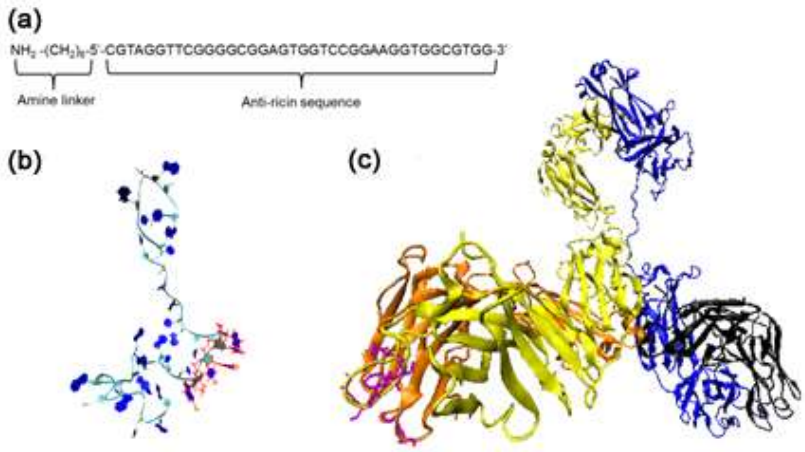

Figure 3.1 The molecular structures of the aptamer and antibody.

In order to design aptamers with high affinities to the target molecules, we need better understanding of the mechanism of the aptamer binding. The function of the DNA aptamer is directly related to its sequence and folding structure. Here we use MD and molecular docking methods to predict the folding structure of the aptamer and the possible binding sties of the binding complexes.

Figure 3.2 shows the predicted binding conformation of ricin-aptamer complex and ricin-antibody complex using AMBER and HADDOCK webserver. Here ricin A chain is in green, the DNA in blue, antibody heavy chain in yellow, and antibody light chain in orange. The left side (a) shows the ricin-aptamer complex, the probable binding residues are highlighted in red and labelled with residue name and number. The right side (b) is for the ricin-antibody complex, and the binding residues are highlighted and labelled in purple.

The ricin A chain, aptamer, antibody heavy chain, and antibody light chain is colored in green, blue, yellow, and orange, respectively. The predicted binding residues of these two complexes are highlighted from the backbone structures. For the ricin-aptamer complex (**Figure**

3.2 (a)), the ricin residues Tyr80, Tyr123, Glu177, and Arg180, and aptamer residues G18, T19, G20, and T22 were the probable binding residues. For ricin-antibody complex (**Figure 3.2** (b)), ricin residues Glu102, Ile104, Thr105, His106, Thr109, Asp110, and Arg114, antibody heavy chain Asp104, Glu105, Arg106, Phe107, Ala108, and light chain Thr33, Lys67, Thr94, and Ser95 were the probable binding residues. These simulations predicted that aptamer and antibody have different binding sites on ricin and both of them are located on ricin A chain. When using aptamer as blocking reagent, the AFM recognition signal reduced significantly. While using anti-ricin antibody as the blocking reagent showed no influence on the recognition signal, because the aptamer on the AFM tip can still bind to the corresponding binding sites of ricin.

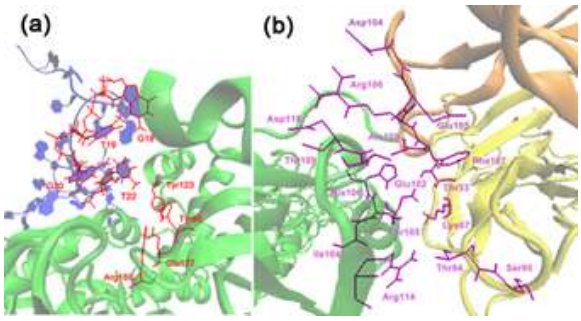

Figure 3.2 The predicted binding sites of ricin-aptamer complex and ricin-antibody complex.

3.3 Results and Discussion

To prove the specificity of aptamer-ricin binding, we did another blocking experiment using anti-ricin antibody as blocking reagent. After 1.3 μM anti-ricin antibody was injected into the liquid cell, the topography and recognition images were monitored. **Figure 3.3** showed the blocking of ricin recognition signals using 1 μM aptamer solution. The upper part is the images before injection of aptamer; and the lower part is 46 min after injection. The aptamer solution significantly reduced the recognition signals. When using 1.3 μM antibody solution, the recognition signals didn't fade away (image size is 450 nm by 450 nm). The antibody didn't block

the recognition signal because the antibody binding site on ricin is different from the aptamer binding site on ricin. Therefore, the aptamer on AFM tip can still bind to ricin with the existence of antibody and show recognition signals.

Next, the affinity of the binding reaction between ricin and aptamer was estimated using the AFM-DFS technique. Based on AFM-DFS experimental data, we compared the interactions of ricin to aptamer with the interaction of ricin to antibody. This ricin-aptamer interaction system is assumed to follow the rules of Bell-Evans single-barrier model, so that the dissociation constant (k_{off}) can be derived from the force distribution. In the AFM-DFS experiments, for both aptamer-ricin and antibody-ricin DFS experiments, five relatively slow loading rates (0.24 nN s^{-1}, 2.72 nN s^{-1}, 15.37 nN s^{-1}, 38.02 nN s^{-1}, 51.64 nN s^{-1}) were used to measure the unbinding forces. Under each loading rate, 250 force-distance curves were measured and the histograms of the force distributions under these loading rates were obtained.

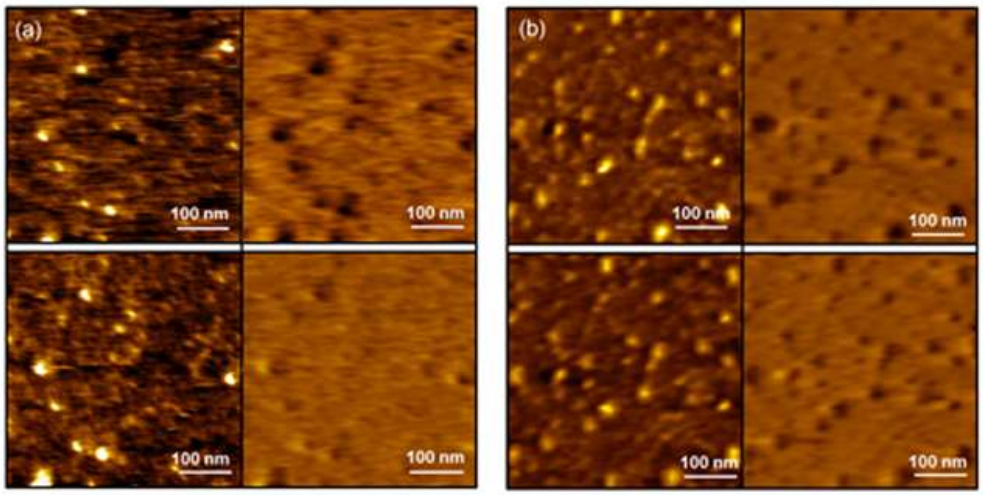

Figure 3.3 Blocking of the recognition signals generated by the aptamer functionalized AFM tip.

The representative force-distance curves of aptamer-ricin binding and antibody-ricin binding were shown under corresponding loading rates in **Figure 3.4** (a). The most probable

unbinding force F* under each loading rate R was obtained from the fitting of the histograms of these force distributions. The distributions of aptamer-ricin unbinding forces at these five loading rates were shown in **Figure 3.4** (b). The most probable unbinding force F* under each loading rate R was obtained by the Gaussian fitting of the histogram of each force distribution. **Figrue 3.4** (c) shows the linear fitting plots of F* vs. the natural logarithm of loading rates (lnR) for the aptamer-ricin interaction (colored in red) and antibody-ricin interaction (colored in blue). The F* values of aptamer were slightly larger than the ones of antibody under all five loading rates. The values of k_{off} for aptamer-ricin and antibody-ricin reactions were both derived by Bell's model described in Chapter 1. The k_{off}(aptamer) was calculated as $6.8\times10^{-4}\cdot s^{-1}$ while the k_{off}(antibody) was $1.5\times10^{-3}\cdot s^{-1}$.[70] Therefore, the k_{off}(aptamer) obtained from our AFM-DFS experiments is around half of the k_{off}(antibody), although there is no significantly difference when it comes to the affinities of biomolecules. The comparison of k_{off} values demonstrates that the aptamer have slightly higher affinity to ricin.

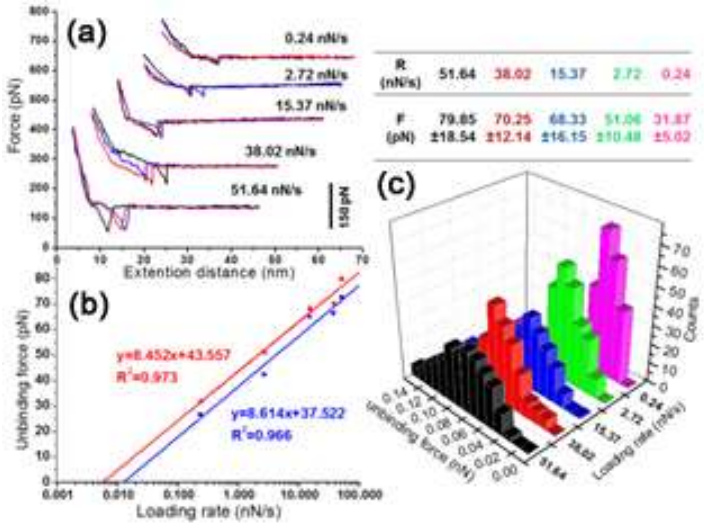

Figure 3.4 The representative data obtained from DFS experiments.

3.4 Conclusion

In conclusion, the AFM recognition and DFS experiments provide both morphology and interaction information of the aptamer and protein, which can be used for the future study on the thermodynamics and kinetics properties of ricin-aptamer/antibody interactions. The immobilization method is proved to be able to constrain the ricin conformations on modified Au(111) surface. The molecular simulations provide detailed structural information of the aptamer, antibody, and ricin, which are critical when we try to understand the functions of these biomolecules and the thermodynamic processes of the binding events. The combination of AFM techniques and molecular simulations provides a new approach for the single-molecule detection and mechanism study of aptamer.

CHAPTER 4

TRANSITION MODEL FOR RICIN-APTAMER INTERACTIONS WITH MULTIPLE PATHWAYS AND ENERGY BARRIERS[1]

[1]B. Wang and B. Xu. 2014. *Phys. Rev. E*. 89:022720.
Reprinted here with permission of publisher.

4.1 Introduction

During the study of aptmer-ricin interactions, one of the essential issues is the influence of aptamer structures to their affinities. The 3-dimentinal structural information of aptamer is difficult to be obtained by experiments such as X-ray crystallography and NMR due to their highly flexible nature. Therefore, single-molecule study has become an important complimentary method in order to understand the complex structure-function relationship of aptamers. Moreover, aptamers are short nucleic acids sequences that show specific interactions with their target molecules. The studies on the interactions between aptamers and their target proteins can help us understand the functions and reactions mechanisms of nucleic acids in general. This chapter will discuss the reaction models we proposed for the complex aptamer reactions.

Interactions between proteins and nucleic acids are critical to many biophysical processes in living cells, such as gene expression, DNA repair, and cell replication.[82,83] For most of these interactions the nucleic acids show complex behaviors, such as flexible folding structures, multiple energy barriers, and multiple reaction pathways for the same ligand-receptor system. Single molecule measurements provide structural and functional information of biomolecules participating in these reactions.[84-86] Biophysical theories developed by Bell, Evans, and Jarzynski have been used to estimate the binding free energy at the equilibrium state from non-equilibrium measurements.[12,87,88] However, while dealing with complex systems, including the unbinding processes of protein and nucleic acids, issues related to multiple energy barriers and reaction pathways have to be considered.

Recently, the biophysical research community has given increasing attention to multiple-bond folding/unfolding and binding/unbinding reactions.[89-95] Nonetheless, detailed binding/unbinding mechanisms for such complex systems require much further investigation.[85,96]

To quantitatively study the specific behaviors of reaction networks, practical biophysical models revealing the relationship between the molecular structures and their behaviors are crucial.

Here we develop a practical transition model for the reaction network of ricin-aptamer interactions. This model combined the Bell-Evans model and Markov-type transition matrices to explain complex relationships among the molecular structures, unbinding forces, and activation energies, using data obtained by AFM-DFS. These target molecule ricin involved in essential processes of living cells, and can cause diseases or cell death.[47,96] We have previously reported the AFM and DFS measurements of ricin conformations and its interactions with another aptamer of simpler folding structure .[9]

In this study, multiple unbinding forces have been observed under the similar extension distance in force-distance curves. This phenomenon indicated that more than one unbinding pathways may exist for the aptamer-ricin interactions (**Figure 4.1**). Shifting of force distributions in the histograms under different loading rates suggested that these multiple unbinding pathways follow the Bell-Evans model. An appropriate quantitative model to explain these observed experimental results is the Markov chain, which relates the peak areas to the likelihoods of unbinding reactions. Therefore, the changes of peak areas in the force histograms under different loading rates reflect the re-distributions of individual measured unbinding force, which connects to each reaction pathway.

A Markov-type transition matrix, a tool to connect multiple unbinding pathways in an overall mathematical expression. A specialized transition matrix can be built according to the reaction network formed by the multiple pathways and was used to build a transition model thus explaining the complex experimental data very well.

4.2 Experimental Details

4.2.1 Two Aptamers with High Affinity to Ricin

The aptamer we used is a sequence of DNA predicted to have two major hairpin-loop structures: loop(I) and loop(III).[96] We found ricin-aptamer interactions happened in three different binding sites. The aptamer loop(I) or loop(III) can bind to its own binding site on ricin surface. The loop(I) and loop(III) can also simultaneously bind to ricin on a different binding site.[24,26] In our DFS experiments, the aptamer is attached to a long PEG linker, while the ricin molecules were in a very low surface concentration. Therefore, the aptamer can easily reach the individual ricin molecule with all of their binding sites/loops. This experimental design increased the probability of multiple-bond binding and unbinding. The force values of this third peak varied from 69.6 pN to 101.2 pN when loading rate changed form 0.24 nN s^{-1} to 51.64 nN s^{-1}. These loading rates and forces values are in the same ranges people used in other DFS studies of multiple-bond reactions

4.2.2 AFM DFS Measurements and Curve Fitting

AFM experiments were conducted in TopMAC mode (Agilent Technologies, Santa Clara, CA). The immobilization of ricin sample and the force measurements were conducted inside the liquid cell. The tip spring constant was 0.1 N/m. In PBS pH 7.2, the measured resonant frequency under TopMAC mode was around 10 kHz. DFS force measurements were done in this liquid cell, with 300 force-distance curves collected under each loading rate. Loading rates were adjusted to certain values by changing the sweep duration and z-position of the tip. Measured ricin molecules were selected from different areas of the sample surface.

Figure 4.1 shows the force histograms obtained under all ten loading rates, from 0.24 nN s^{-1} to 600 nN s^{-1}, with (a) 0.24 nN s^{-1}, (b) 2.72 nN s^{-1}, (c) 15.37 nN s^{-1}, (d) 38.02 nN s^{-1}, (e) 51.64 nN s^{-1}, (f) 82.06 nN s^{-1}, (g) 100.0 nN s^{-1}, (h) 200.0 nN s^{-1}, (i) 400.0 nN s^{-1}, and (j) 600.0 nN s^{-1}. For

the inset from (a) to (e), the f-d curves for peak 1, pathway 1, and loop(III) unbinding, are in red (smallest unbinding force in the inset figure). The f-d curves for peak 2, pathway 2, and loop(I) unbinding, are in green (middle-size unbinding force). The f-d curves for peak 3, pathway 3, and loop(I)-(III) simultaneous unbinding, are in blue (largest unbinding force).In this range the multiple energy barriers are distinguishable in the later F^* vs. lnR plots.

Under the low loading rates (0.24 nN s^{-1} to 51.64 nN s^{-1}), the change of the peak areas of the three pathways is in a clear trend, which showed the redistribution processes of the unbinding reactions among the three different pathways. Peak 3 almost diminished under 51.64 nN s^{-1}. When the loading rate continued to increase, the third peak (peak 3) completely disappeared, and the other two peaks gradually merged to one peak. Under high loading rates (200.0 nN s^{-1} to 600.0 nN s^{-1}), only one peak existed in the histogram, which indicated that those unbinding pathways showed certain relationships with the force applied to the binding complex. Under the medium range of loading rates (82.06 nN s^{-1} to 100 nN s^{-1}), the force histograms showed two peaks, as shown in **Figure 4.1** (f) and (g).

Based on the molecular docking and dynamics simulations, pathway 3 involved the most binding residues and largest binding area compared to the other two pathways, as shown in **Figure 4.1** (e). Therefore, the binding/unbinding conformation for pathway 3 is supposed to require more time to be stabilized and become increasingly rare with the increase of loading rates. Therefore, the two peaks under medium loading rates were assumed to represent pathways 1 and 2. Under medium and high loading rates, the force histograms all significantly broadened toward large force values, which is a common phenomenon in force spectroscopy.[97] Therefore, we consider all the small "tails" of the large force parts in those force histograms as the deviation caused by the thermal noise and measurement errors. The Gaussian fitting also neglected those parts of the force

distribution under high loading rates, and automatically fitted two peaks under the loading rates of 82.06 nN s^{-1} and 100.0 nN s^{-1}. Overall, both three-peak fittings under low loading rates (from 0.24 nN s^{-1} to 51.64 nN s^{-1}) and two-peak fittings under medium loading rates (82.06 nN s^{-1} and 100.0 nN s^{-1}) were determined by the same standards, not due to arbitrary selections. The in-depth explanation of these multiple peaks in force histograms under the low and medium loading rates will be described in section IV with two force transition models.

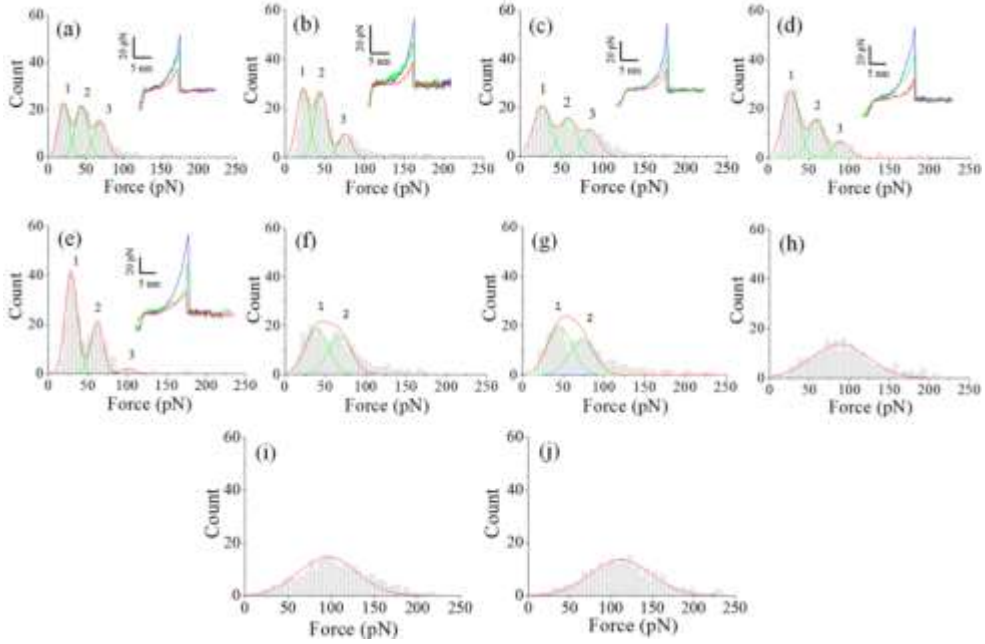

Figure 4.1 The force histograms and representative f-d curves.

4.2.3 Molecular Docking and MD simulations

The aptamer sequence was obtained from literature,[30] and attached with a thiol group at the 5'-end so that it can be attached to the gold coated AFM tip.[9] The entire aptamer sequence including the 5'-end amine functional group is NH_2-$(CH_2)_6$-5'-ATAGG-AGTCA-CGACG-ACCAG-AACCG-TAGGT-TCGGG-GCGGA-GTGGT-CCGGA-AGGTG-GCGTG-GTATG-TGCGT-CTACC-TCTTG-ACTAAT-3'.

The secondary structure of the aptamer was predicted by Mfold webserver,[98] and the tertiary folding structure was generated by Amber 11 molecular dynamic package.[24] Specifically, the program Nucleic Acid Builder in the Amber package was used here to build the constraints on A-T and G-C base pairs, simulate equilibrium, and obtain the most stable folding structure.

The ricin structure was obtained from PDB entry 2AAI.[74] Before docking simulation, both aptamer and ricin structures were neutralized and equilibrated in TIP3P water solvatebox for 2 ns to obtain their most stable conformations in solution. Next, the ricin and aptamer structures were uploaded to the High Ambiguity Driven protein-protein DOCKing (HADDOCK) webserver,[26,77] and the two loops of the aptamer were assigned as the active residues, respectively, in the first two docking simulations. In the third docking simulation, both loops were assigned as the active residues. The prediction of binding residues on aptamer and ricin were based on literature and molecular docking simulation. The HADDOCK webserver was used to search the active residues and form the most stable binding conformation. Then the loop(III) docking and loop(I)-(III) simultaneously docking were tested and all of them showed stable binding conformations.

These docking simulations predicted that some ricin residues were involved in more than one pathway. For example, R125 is involved in all three pathways. Residues Y80, Y123, and R180 are involved in both loop(I) binding and loop(III) binding, but not in loop(I)-loop(III) simultaneous binding. Residue Q128 is involved in both loop(III) binding and loop(I)-loop(III) simultaneous binding, but not in the loop(I) binding. Moreover, the aptamer loop(I) and loop(III) also showed different binding residues for different binding conformations. For loop(I) binding conformation, aptamer residues C23 to G28 bind to ricin. For the two-loop binding conformation, loop(I) residues A21 to T26 bind to ricin. For loop(III) binding conformation, T66 to T70 bind to ricin, while G65 to C71 bind to ricin in the two-loop binding conformation. These overlaps and mismatches of ricin

residues for aptamer loop(I) and loop(III) in different unbinding pathways indicate that the loop(I)-loop(III) simultaneous binding residues/area is not simply the addition of the ones from individual loop(I) and loop(III) binding. Therefore, these simulation results support a transition model that can connect the three pathways together in a reaction network. Especially, based on the experimental data (**Figure 4.1**), the change of loading rates influenced the distributions of those three pathways, which started from their corresponding binding conformations shown in **Figure 4.2**. Here, (a) is the Loop(I)-ricin binding, (b) is the Loop(III)-ricin binding, (c) is the Loop(I)-(III) binding simultaneously. The binding residues from ricin are in blue (stick representation in red, dark), the binding residues from aptamer are in green (stick representation in yellow, grey). The backbone structures of ricin and aptamer are in gray and translucent (light grey). The initial aptamer-ricin intermediate (revealed by high loading rates) is assumed to connect all of the three binding conformations. It will convert to one of these binding conformations under medium and low loading rates. These complex relationships among the different binding residues in different pathways are mainly due to the spatial constraints of ricin and aptamer folding structures.

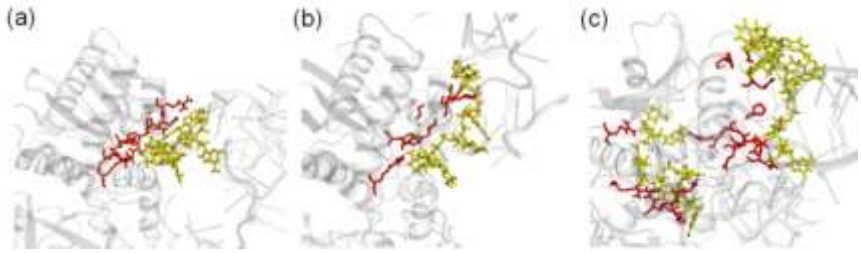

Figure 4.2 The predicted binding residues for different aptamer-ricin conformations.

4.2.4 Markov Model for Force Transition Processes

The development of the transition model was based on Markov analysis, which is a stochastic method used to study the time evolution of a system and describe its transition probabilities.[99-101] Markov models have been used to determine reaction properties by analyzing

data obtained using fluorescence measurements.[102-105] However, previous research focused on stochastic properties of biomolecules evolving along the reaction time and was not specially developed for single-molecule experiments.[101,106] Our work focused on the changing force distributions and their transition rates under different loading rates. The analysis was based on firsthand single-molecule experimental data and simulations of the binding conformations. When an external force was applied to the ricin-aptamer system under low loading rates (0.24 nN s^{-1} to 51.64 nN s^{-1}), it reduced the activation energy barriers of pathways 1, 2, and 3 according to the Bell-Evans model. This transition of the most probable unbinding force (F^*) starts with the loop(I)-loop(III) simultaneous unbinding pathway, and connects all three pathways under low loading rates. Therefore, the change in force histograms revealed how individual unbinding reactions distributed among the three unbinding pathways. Based on simulations, the initial aptamer-ricin intermediate (revealed by high loading rates) is not like any one of the three binding conformations. Therefore it will convert to one of these binding conformations under medium and low loading rates. These complex relationships among the different binding residues in different pathways are mainly due to the spatial constraints of ricin and aptamer folding structures. Based on the measured force data, pathway 3 has the strongest unbinding force and its energy barrier is expected to be higher than the other two pathways, so the unbinding process can easily transit to pathway 1 or pathway 2. However, the redistribution of the unbinding process (or the change of probability) of pathway 1 or 2 to pathway 3 is much more difficult. We defined S as the likelihood of the unbinding state of the molecular pair, and K as the force transition rate of a certain pathway between two loading rates. We assume that pathway 1 and 2 cannot transit to pathway 3, so the K values of pathway 1 and 2 into themselves are unity (100 %).

Under different loading rates, the changes of energy landscapes under the applied forces caused the redistribution of the likelihood (S_1, S_2, and S_3) for each pathway. Therefore, the observed shifts of peak area values between two force histograms under different loading rates were related to the unbinding pathways. A single-molecule unbinding reaction can be simply put as $S_0 \rightleftarrows S_1$, where S_0 is the likelihood of the binding state of the molecular pair, and S_1 is the likelihood of the unbinding state. Using k_1 (unit %) as the forward rate constant and k_0 (unit %) as the backward rate constant, the master equation of this reaction can be written as

$$\frac{dS_0}{dt} = -k_1(t)S_0(t) + k_0(t)S_1(t) \tag{4.1}$$

In AFM DFS experiments the two molecules move apart much faster than their diffusion in the solution, so the rebinding term k_0S_1 goes to zero. We define the constant tip velocity as v, the distance as vdt, and the overall spring constant of the AFM tip is simplified as κ_s. The force and distance follow Hook's law:

$$df = \kappa_s v dt \tag{4.2}$$

Now df can be substituted into the simplified master equation

$$\frac{dS_0}{dt} = -k_1(t)S_0(t) \tag{4.3}$$

to obtain

$$\frac{dS_0}{df} = -\frac{k_1}{\kappa_s v}S_0 \tag{4.4}$$

Here we assume that Hook's law holds for every loading rate.[11] Generally, a higher loading rate will generate a bigger force to break the bond, and this bigger force should be equivalent to the stochastic process if the f-d curve occurs in continuous time. Therefore, we use equation (4.1) to

connect the unbinding force values from two loading rates. When the most probable unbinding force changed during the transition process, we define the force transition rate to be $k_{f1}=\frac{k_1}{K_s v}$ (unit % N^{-1}), and substituting it into equation (4.4) yields

$$\frac{dS_0}{df} = -k_{f1}S_0 \tag{4.5}$$

Equation (4.5) describes the relationship between the force transition rate and the likelihood of the binding state caused by the changes of the applied force from the AFM tip. The integral of equation (4.5) from one force value to another (new) force value yields

$$\Delta S_0 = e^{-k_{f1}\Delta f} \tag{4.6}$$

The force changes in DFS are usually in pN scale, so we assume that the linear approximation of equation (4.6) is valid under the DFS experimental conditions:

$$\Delta S_0 = 1 - k_{f1}\Delta f \tag{4.7}$$

Because the change of DFS loading rates resulted in the corresponding change of unbinding forces, the likelihood of one unbinding reaction through its pathway can be considered a function of the loading rates. The unbinding energy barrier of that certain reaction was lowered by the applied force, so the k_{f1} value can be considered as the mechanical property of a certain unbinding reaction with the change of loading rates. For a system with three pathways, a transition matrix T can be used to connect all k_f of the different pathways together, as shown below:

$$\frac{d[S_0]}{df} = -[T][S_0] \tag{4.8}$$

Using the similar approximation for equation (4.6) and (4.7), the equation (4.8) can be transferred to

$$\left[S_{f1}\right]\times\left[T\right]=\left[S_{f2}\right] \tag{4.9}$$

Here the matrix [T] includes parameters such as k_{f1} and Δf. The matrices [S_{f1}] and [S_{f2}] represent the likelihood of the binding states of the three pathways under two applied forces $f1$ and $f2$, which correspond to two different loading rates. The change of DFS loading rate changes the applied force value. This change of the applied force in turn changes the energy landscapes of those pathways and caused the redistribution of the likelihood (S_1, S_2, and S_3) for each pathway under the applied force. Therefore, the observed shifts of peak area values between two force histograms under different loading rates were related to the unbinding pathways. The transition diagram of this process is shown in **Figure 4.3** (a), and the matrix analysis of this transition model can be written as equation (4.10):

$$\left[S_{1,i},S_{2,i},S_{3,i}\right]\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ K_{o1} & K_{o2} & K_{o3}\end{bmatrix}=\left[S_{1,j},S_{2,j},S_{3,j}\right] \tag{4.10}$$

Here S is the normalized peak area value from the force histogram, the subscripts 1, 2, and 3 represent pathway 1, 2, and 3, while i and j denote the initial and final loading rates from the DFS experiments, respectively. The middle part of equation (4.10) is the matrix [T] in equation (4.9), where K_{o1}, K_{o2}, and K_{o3} are the matrix transition rates for their corresponding unbinding pathways. Here o represents the outer energy barriers revealed by the force-distance measurements under relatively low loading rates, according to the Bell-Evans multiple-barrier model. The K_{o1}, K_{o2}, and K_{o3} are defined as $k_f \Delta f$. Here k_f is the force transition rate per unit of force. The Δf is the change of the most probable unbinding force of each pathway (pathway 1, 2, or 3) between the starting loading rate and the ending loading rate in equation (4.10). Therefore, the K_{o1}, K_{o2}, and K_{o3} values

obtained in equation (4.10) are not from the analytical solutions of the derivative equations for each reaction pathway. Instead, these values are estimated from the linear approximation (using the transition matrix) of the force transition process between the starting loading rate and the ending loading rate. Equation (4.10) connects the values of corresponding peak areas in the two force histograms under two different loading rates. **Figure 4.3** (a) shows the transition diagram under low loading rates (0.24 nN s^{-1} to 51.64 nN s^{-1}). *S*o1, *S*o2, and *S*o3 represent the likelihood of going through each unbinding pathway under a certain loading rate. **Figure 4.3** (b) shows the transition diagram under medium loading rates (from 82.06 nN s^{-1} to 100 nN s^{-1}).

Under the medium range of loading rates (82.06 nN s^{-1} to 100 nN s^{-1}), the force histograms only showed two peaks for pathway 1 and 2. A new transition matrix was necessary to describe the redistribution of the likelihood (S_1 and S_2) for each pathway. This transition process is shown in **Figure 4.3** (b), where S_{m1} and S_{m2} represent the two likelihood states of S_1 and S_2 in the matrix analysis. Therefore, the equation for the Markov process under medium loading rate can be written as

$$\begin{bmatrix} S_{1,i}, S_{2,i} \end{bmatrix} \begin{bmatrix} 1-K_{m2} & K_{m2} \\ K_{m1} & 1-K_{m1} \end{bmatrix} = \begin{bmatrix} S_{1,j}, S_{2,j} \end{bmatrix} \tag{4.11}$$

Here *S* is the normalized peak area value from the force histogram, the subscripts 1 and 2 represent pathways 1 and 2. The subscripts *i* and *j* represent the initial and final loading rates, respectively. The parameters K_{m1} and K_{m2} are the matrix transition rates for their corresponding unbinding pathways under medium loading rates.

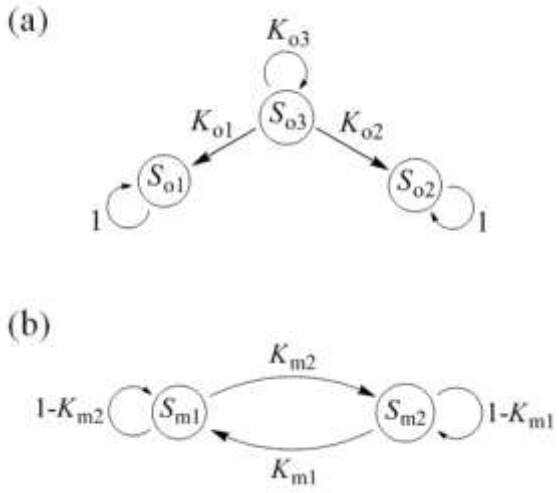

Figure 4.3 Transition diagrams for multiple unbinding pathways.

Figure 4.4 shows the overall reaction network for the ricin-aptamer interactions under different loading rates. This model is based on the histograms of unbinding forces at loading rates ranging from 0.24 nN s^{-1} to 600 nN s^{-1}. Under high loading rates, the inner energy barrier dominated each unbinding pathway, while the outer energy barrier dominated the pathway when the force measurements were under low loading rates. The transition mode changed under the medium loading rates. This change of transition mode, from **Figure 4.3** (a) to (b), was caused by the change of relative positions of energy profiles for each pathway. The transition models in **Figure 4.3** form the reaction network in **Figure 4.4**, and the kinetic parameters such as the reaction rate constants in this reaction network can be determined by Bell's model, which connects the applied force with the energy profile of each pathway. Therefore, the force transition processes among the three pathways are connected with the energy profiles of those pathways. We did further analysis of the energy profiles involved in this multiple-pathway system in order to quantitatively explain the mechanism behind the relative energy levels of those pathways and the transitions among them.

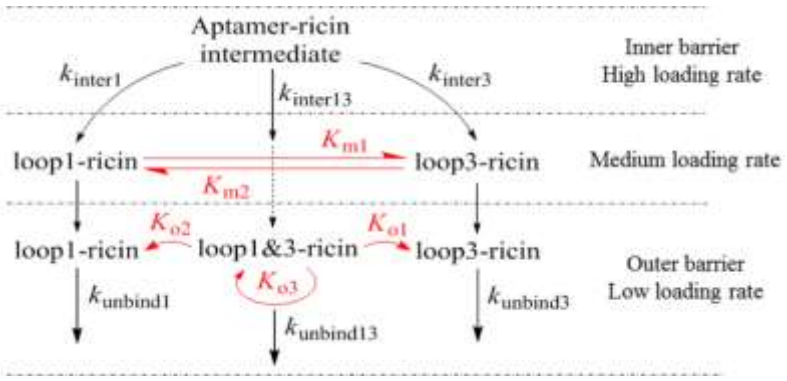

Figure 4.4 Overall reaction network showing the unbinding pathways under different ranges of loading rates.

4.3 Results and Discussion

4.3.1 Estimate of Reaction Off Rates and Energy Profiles

The fittings of F^* vs. ln$R$ plots were based on Bell-Evans multiple-barrier model.[107] Force peak values were obtained by Gaussian fitting from the histograms. In the two-barrier fitting, the plots and data points for loop(I) pathway are labeled as 2 (filled circle in green), the ones for loop(III) pathway are labeled as 1 (inverted triangle in red), and the ones for loop(I)-loop(III) pathway are labeled as 3 (triangle in blue). The shared data points in high loading rates 200.0, 400.0, and 600 nN s^{-1} are marked as crossed circles. Under the low loading rates from 0.24 nN s^{-1} to 51.64 nN s^{-1}, each pathway has three peak values in the force histograms. So the fitted plots were clearly separated from each other in **Figure 4.4** (a), (b), and (c). Under medium loading rates 86.02 and 100.0 nN s^{-1}, the loop(I)-loop(III) pathway didn't show in the force histograms anymore, so no data point is available for this pathway in this range. For the other two pathways, the data points clearly indicated new slopes for their fitted plots. Under high loading rates 200.0 nN s^{-1} to 600 nN s^{-1}, only one peak showed in each force histogram, which is assumed to represent the initial aptamer-ricin intermediate that generated these three pathways, as shown in **Figure 4.4**. In order

to estimate the second slope of individual pathway, the fitting started from the end of the first slope (51.64 nN s^{-1}), included the medium range and high range of the loading rates. The changing of slopes around 51.64 nN s^{-1} indicated the change of reaction mechanisms. The new mechanism is for the inner barrier reaction.

The first fitting range is from 0.24 nN s^{-1} to 51.64 nN s^{-1}, totalling five points for each pathway. The second fitting range is from 51.64 nN s^{-1} to 600 nN s^{-1}, totalling four points for loop(I)-(III) pathway, and in total six points for each of the other two pathways. The off rates of the outer barriers were calculated as $(6.6\pm1.7)\times10^{-6}$ s^{-1}, $(2.6\pm0.6)\times10^{-7}$ s^{-1}, and $(5.8\pm0.6)\times10^{-10}$ s^{-1} for pathway 1, 2, and 3, respectively. The off rates of the inner barriers were calculated as $(6.9\pm2.3)*10^{2}$ s^{-1}, $(1.6\pm0.08)*10^{2}$ s^{-1}, and 2.6 ± 0.2 s^{-1} for pathway 1, 2, and 3, respectively. For each unbinding pathway, the off rates of inner energy barriers and outer energy barriers have significant difference, which indicates that the unbinding process to overcome the inner energy barrier is happening much easier and more frequently. These energy estimates help us understand the detailed structural-function relationships of the aptamer-ricin complexes.

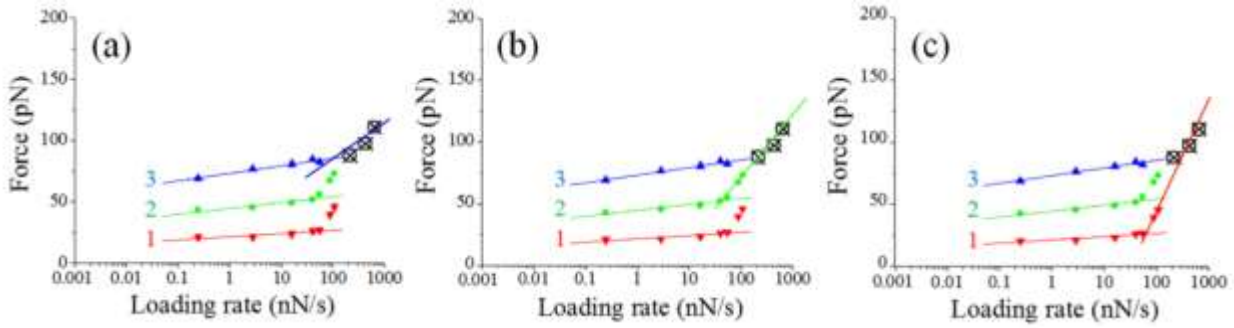

Figure 4.5 The F* vs. lnR plots for the three unbinding pathways.

Figure 4.6 (a) shows the reconstructed energy landscape of each unbinding pathway using the Bell-Evans model.[107,108] Although the ricin-aptamer binding complex has three conformations, the energy levels at their initial binding states were assumed to be very similar, as the black dash curve at x_β=0 shows. The initial energy level of the binding state was estimated using a microscopic diffusive relaxation time of 4.2×10^{-9} s according to the literature.[107, 109] The binding state energy value was estimated to be around -22 k_BT. [108,110]

Next, the outer and inner energy barriers (E) along the reaction course (x_β) of these three unbinding pathways were calculated using the Bell-Evans multiple-barrier model. For outer energy barriers, pathway 1 showed the lowest barrier energy (red curve, 11.9±1.1 k_BT above initial binding state) and longest barrier locations (3.6±0.9 nm), while pathway 3 showed the highest energy barrier (blue curve, 21.3±0.2 k_BT above initial binding state) and shortest barrier position (1.5±0.2 nm), and pathway 2 has the medium energy barrier (green curve, 15.1±1.1 k_BT above initial binding state) and medium barrier position (2.0±0.5 nm).

Based on the experimental data, pathway 1 (i.e., loop(III) binding) showed relatively large errors for x_β values, which were the major contribution to the errors of energy levels. Especially for the inner barrier of pathway 1, the error ±1.2 k_BT is even greater than the energy level -0.9 k_BT. The main reason was that the external applied forces under large loading rates generated bigger noise during the measurements of weak unbinding forces of pathway 1. For the other two pathways, their stronger unbinding forces reduced the errors on the estimates of barrier width x_β and energy level ΔE.

For inner energy barriers, pathway 1 showed highest energy level (red, 0.9±1.2 k_BT below initial binding state, barrier position at 0.1±0.04 nm), while pathway 3 showed lowest energy level (blue, 6.5±1.0 k_BT below the initial binding state, barrier position at 0.4±0.04 nm), pathway 2 is

still in the middle (green, 5.0±1.0 k_BT below initial binding state, barrier position at 0.2±0.01 nm). The relative energy value and barrier positions of these pathways are comparable with other biomolecular interactions.[107] However, the three inner barrier energy levels are slightly higher than the estimated -22 k_BT, and the real value cannot be directly determined by single-molecule measurements.[110] Previous literature investigated the transition path times and diffusion of nucleic acids and proteins across their folding/unfolding barriers, but we cannot find a practical method to experimentally obtain the initial binding energy level.[84,111] Therefore, we used the value of -22 k_BT as a rough reference for this ricin-aptamer system.

Based on the binding conformations of the three pathways, the conformation for pathway 3 (**Figure 4.2** (c)) showed the most binding residues and two binding sites for ricin and aptamer. Therefore, the unbinding process is expected to generate a bigger force and shorter length of the barrier. On the other hand, the conformation for pathway 1 (**Figure 4.2** (a)) has the least binding residues and smallest binding area, which lead to the smallest unbinding force and most flexible structural change, and the longest barrier location. For the inner barriers, we assumed that pathway 1 (red) had the most similar binding conformation to the initial conformation of the binding intermediate, so its barrier level is also the closest to the initial binding state. Therefore, the initial binding complex needs to move a longer distance (and x_β) to reach the binding/unbinding conformations of the other two pathways (green and blue), so the inner energy barriers of these two pathways differ greater still from the initial binding state.

Figure 4.6 shows the schematic energy profiles for all of the three unbinding pathways, and the transitions of unbinding reactions from the starting loading rate to the ending loading rate caused by the applied forces. The reconstructed energy landscapes of ricin-aptamer unbinding pathway are shown in **Figure 4.6** (a). Pathway 1 (labeled as 1 in red) has the highest energy level

among all inner barriers, and lowest energy level among outer barriers. Pathway 3 (labeled as 3 in blue) has the lowest energy level among all inner barriers, and highest energy level among outer barriers. Pathway 2 (labeled as 2, green) has the medium energy level between pathway 1 and 3. The "initial" labeled in black at xβ=0.0 represents the estimated value of initial energy level of the binding state -22 kBT. **Figure 4.6** (b) shows the schematic energy profiles of ricin-aptamer unbinding pathways with the transitions among them. **Figure 4.6** (c) shows the transitions among pathway 1 and 2 for the inner energy barrier, with their corresponding transition constant *K*b1 and *K*b2, respectively. **Figure 4.6** (d) shows the transitions from pathway 3 to pathway 1, 2, and 3 for the outer energy barrier, with their corresponding transition constant *K*a1, *K*a2, and *K*a3, respectively. In (b) and (c), the dashed curves show the energy profiles of starting loading rates, and solid curves show the energy profiles of ending loading rates. The transition inside each pathway is not shown. The force transition rates marked in **Figure 4.6** (c) and (d) are the same ones used in **Figure 4.3**. These schematic energy profiles summarize the two transition models (for inner and outer energy barriers) in the perspective of reaction energy, and help understand the transitions in **Figure 4.3**.

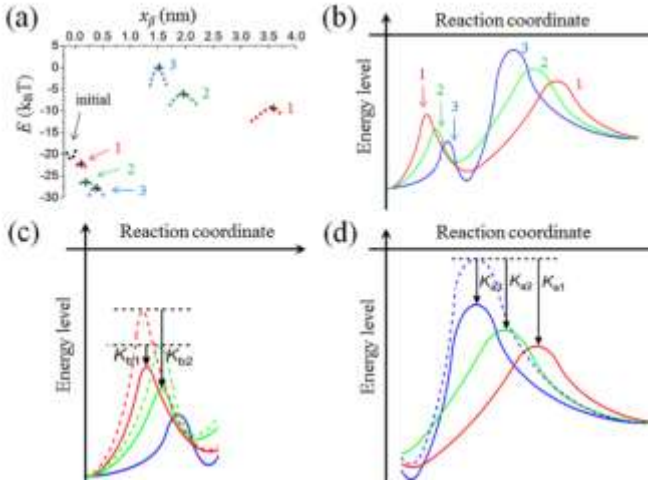

Figure 4.6 Estimate of energy levels for each pathway.

4.3.2 Energy Levels during Transition Process

Finally, based on the peak areas obtained by the DSF experiments, we calculated the matrix transition rates of the three pathways (K_1, K_2, and K_3) for the four transition processes which happened under loading rates from 0.24 nN s^{-1} to 51.64 nN s^{-1}, according to equation (4.10). The K values represent the percentage of unbinding events when pathway 3 transited to pathway 1 or 2, or remained in pathway 3, between two loading rates where different external forces were applied. The unbinding events in pathway 1 and 2 always stayed in their own pathways, as previously discussed. The energy difference caused by the transition between two pathways ($\Delta\Delta E$), from a starting loading rate i to an ending loading rate i+1, can be estimated from the K values. According to the Arrhenius relationship, the off rate under the applied force $k_{off}(F)$ and the off rate under zero force $k_{off}(0)$ can be put as:

$$\frac{k_{off}(F)}{k_{off}(0)} = e^{\frac{Fx_\beta}{k_B T}} \tag{4.12}$$

For the transition between pathway 3 and 1, from loading rates i to loading rate i+1,

$$\ln[\frac{k_{off,1}(F_{i+1})k_{off,3}(0)}{k_{off,1}(0)k_{off,3}(F_i)}] = \frac{(F_{i+1}x_{\beta,1,i+1} - F_i x_{\beta,3,i})}{k_B T} = \frac{\Delta\Delta E_{3\to 1}}{k_B T} \tag{4.13}$$

Here the $k_{off,3}(0)$ and $k_{off,1}(0)$ are the off rates of pathway 3 and 1, respectively, under zero force. The $k_{off,3}(F_i)$ is the off rates value of pathway 3 under applied force F_i, while $k_{off,1}(F_{i+1})$ is the off rate value of pathway 1 under applied force F_{i+1}. These off rate values under certain applied force (loading rate) were calculated using Bell-Evans model. $\Delta\Delta E_{3\to 1}$ is the energy difference (in the unit of $k_B T$) caused by the transition process. The same relationships are also available for the

transition from pathway 3 to pathway 2, and pathway 3 to 3, under two continuous loading rates (**Figure 4.7**). Pathway 1 is in red (lower left part), pathway 2 is in green (lower right part), pathway 3 is in blue (upper part as the starting state and lower middle part as the ending state).

The calculated $\Delta\Delta E$ values for three pathways under different loading rates (K_1, K_2, and K_3) are shown in **Table 4.1**. Here K_1, K_2, and K_3 represent force transition rate values of pathway 1, 2, and 3, respectively. **Table 4.1** further provides quantitative information for the change of energy barriers. For example, **Table 4.1** shows that $\Delta\Delta E_{3\text{-}1}$ and $\Delta\Delta E_{3\text{-}2}$ are always negative, so these transitions under increasing forces (or loading rates from 0.24 nN s^{-1} to 51.64 nN s^{-1}) released energy. However, $\Delta\Delta E_{3\text{-}3}$ was first positive at low loading rate, changing to negative at high loading rate, so this transition process consumed small amounts of energy under low loading rates, but released energy under high loading rates.

Table 4.1 Force transition rate (K) and energy changes caused by transitions among different pathways ($\Delta\Delta E$).

Transition	Force transition rate (%)			$\Delta\Delta E$ (k_BT)		
	K_1	K_2	K_3	Pathway 3 to 1	Pathway 3 to 2	Pathway 3 to 3
0.24 to 2.72 nN s^{-1}	0.096	0.056	0.848	N/A[a]	N/A[a]	N/A[a]
2.72 to 15.37 nN s^{-1}	0.129	0.054	0.817	-7.3	-4.4	1.7
15.37 to 38.02 nN s^{-1}	0.238	0.240	0.522	-7.8	-3.7	0.46
38.02 to 51.64 nN s^{-1}	0.379	0.272	0.349	-8.6	-5.7	-0.10
51.64 to 81.02 nN s^{-1}	0.502	0.553	N/A[b]			
81.02 to 100.0 nN s^{-1}	0.502	0.553	N/A[b]			

[a] Energy comparison not applicable for the first transition process.
[b] K_3 not applicable in the transition model under the medium loading rates.

The applied force changed the energy landscapes of all three unbinding pathways. **Table 4.1** shows that $\Delta\Delta E_{3\text{-}1}$ and $\Delta\Delta E_{3\text{-}2}$ are always negative, so these transitions under increasing forces (or loading rates from 0.24 nN s^{-1} to 51.64 nN s^{-1}) released energy. However, $\Delta\Delta E_{3\text{-}3}$ was first

positive at low loading rate, changing to negative at high loading rate, so this transition process consumed small amounts of energy under low loading rates, but released energy under high loading rates.

4.4 Conclusion

In summary, we present a transition model combining the conventional Bell-Evans model and Markov-type transition matrix to analyze single molecular ricin-aptamer interactions measured by AFM-DFS. Based on the distributions of unbinding forces and molecular dynamics, multiple unbinding pathways were connected in an overall reaction model with their corresponding force transition constants. This model quantitatively described the complex behaviors of single protein and aptamer molecules under external forces, and transition relationships among multiple unbinding pathways. This approach provides detailed information hidden in the DFS experimental data of complex protein-DNA interactions. This approach can be used to study other single molecule interactions, especially those with multiple reaction pathways.

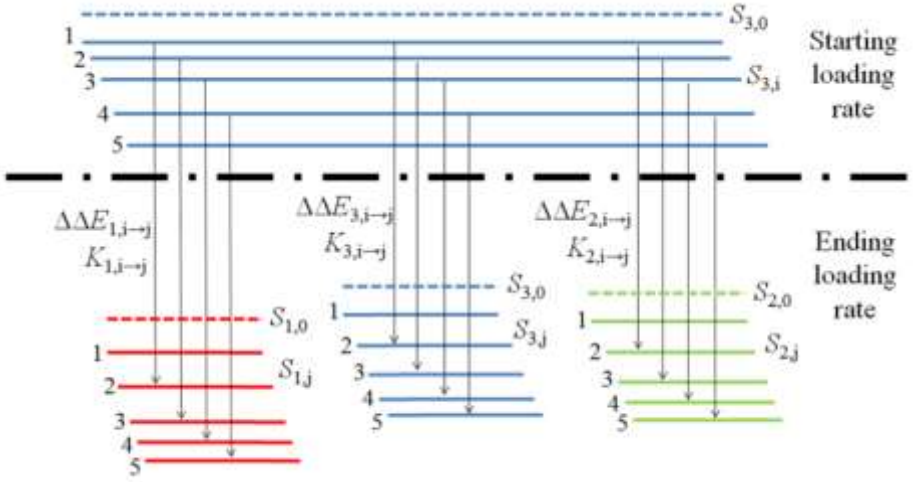

Figure 4.7 Relationships among different transitions following the changes of loading rates.

CHAPTER 5

DETERMINING THE ELASTIC PROPERTIES OF APTAMER-RICIN SINGLE MOLECULE MULTIPLE PATHWAY INTERACTIONS[1]

[1]B. Wang, B. Park, Y. Kwon, B. Xu. 2014. *Appl. Phys. Lett.* 104:193702.
Reprinted here with permission of the publisher.

5.1 Introduction

The previous chapter is focused on the detailed reaction mechanisms of the aptamer interactions, which are directly related to the folding structures of the aptamers. Another factor that influences the binding/unbinding force of the aptamer is the elasticity of its binding conformation during the reaction. This property can be revealed by the single-molecule DFS measurements, which can monitor the individual unbinding process with the aptamer modified AFM tip. However, the current research on aptamer reactions hasn't provided any specific reaction model for the elastic properties of the aptamers during their unbinding reactions. In this chapter, a biophysical model will be provided to estimate the elastic properties of aptamers, based on the reaction mechanisms and structural simulation studied in previous chapter.

Elasticity of biomolecules plays important roles in many cellular activities, and most of these reactions and pathways depend on the molecular motions and fluid behaviors of biomolecules with their specific elastic properties.[112,113] Although some theoretical and experimental methods have been developed to investigate the elasticity of nucleic acids and proteins, most DNAs and some proteins inherently act as flexible and elastic ligand and receptor with complex behaviors, which make it difficult to study their reaction mechanisms.[82,114,115] Among all types of DNA molecules currently studied, aptamer has been considered as a versatile tool for biomedical and biosensing applications.[62] Aptamer refers to short single strand DNA or RNA sequences that have specific interactions with other biomolecules.[62] The specific biochemical properties and relatively low cost have made aptamer a widely used regent in both fundamental research and practical applications.[116,117,118,115,62] However, aptamers have shown complex biophysical and biochemical behaviors, such as multiple binding sites and multiple reaction pathways. Therefore, in order to understand the complex elastic properties of aptamers during their

reactions with other biomolecules, a practical yet detailed investigation method is highly required.[85,119,120,121]

Atomic force microscopy (AFM) and dynamic force spectroscopy (DFS) have been used to experimentally determine mechanical and elastic properties of biospecies such as DNA and protein in nanometer level.[122,123,124] In theoretical studies, Bell-Evans model, warm-like chain (WLC) model, and Hook's law have been used to explain the elastic behaviors of DNAs or proteins.[12,125,126,121] Great efforts have been made to study the behaviors of single DNA molecule under stretching.[127,128,129,130] However, the information about elastic behaviors of DNAs during their unbinding processes is missing because of the complexity of the DNA binding structures and biophysical properties.[131] Our study revealed the single-molecule elastic properties from the AFM experimental data, and developed a relatively simple and practical elasticity model to quantitatively analyze the elastic behaviors of ssDNA during the single-molecule binding/unbinding reaction.

This practical elasticity model is based on the measured extension distributions of each unbinding pathway under a wide range of force loading rates. The DFS experimental setup and series-spring model are schematically shown in **Figure 5.1**, with the series-spring model (a) and the simulated aptamer-ricin binding conformations (b), where the aptamer conformations are labeled as (I) for pathway 1, (II) for pathway 2, and (III) for pathway 3. **Figure 5.1** (c) shows the sequence of the aptamer modified with amino functional group.

A polyethylene glycol chain with molecular weight of 2000, PEG2000, was used as linker and spacer to connect aptamer to the gold coated AFM tip surface.[9] The PEG2000 was treated as an elastic polymer molecule in good solvent, and its physical behaviors follow worm-like-chain (WLC) model.[132] The aptamer formed a stable folding structure in PBS buffer, and the applied

force will distort and break the connection between aptamer and ricin during the unbinding process. Therefore, their elastic behaviors cannot be studied by a simple WLC model. The ricin protein is simplified as a rigid receptor and molecular docking simulations predicted that it has more than one binding site for the aptamer (**Figure 5.1** (b)). Overall, a series-spring model was established to describe this multiple-pathway system (**Figure 5.1** (a)).

The molecular basis of the elastic properties in different pathways is that each reaction pathway has its corresponding binding residues on aptamer and ricin surface. The HADDOCK docking simulation and Amber MD simulation showed that the aptamer and ricin can form three stable binding conformations. Although their binding areas on ricin surface have overlap, the aptamer binding conformations and ricin active residues of these three complexes have significant differences. Moreover, for the two-loop binding pathway, the predicted ricin binding residues are not simply the sum of the ones for individual loop binding. Indeed, the simulations results indicate that three reaction pathways should have different chemomechanical properties, which is confirmed by the DFS experimental data.

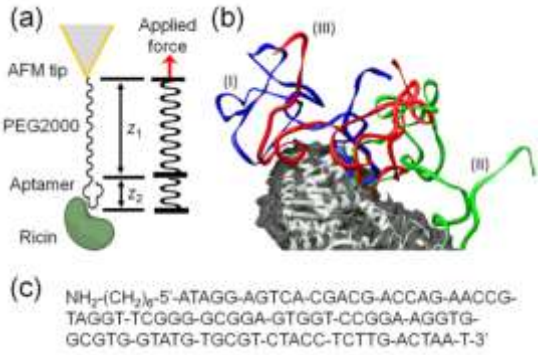

Figure 5.1 The force-extension model for the aptamer-ricin interactions.

Force histograms (**Figure 5.2**) and extension histograms (**Figure 5.3**) suggested that the unbinding reactions can happen through three pathways. The pathway 1 (labeled as binding

conformation (I) in **Figure 5.1** (b)) happens between ricin and aptamer loop#3, with the small "most probable unbinding extension value" in each extension histogram (peak 1 in **Figure 5.3**). The pathway 2 (labeled as binding conformation (II) in **Figure 5.1** (b)) happens between ricin and loop#1. The pathway 3 (labeled as binding conformation (III) in **Figure 5.1** (b)) happens between ricin and loop#1-loop#3 simultaneously. The pathways 1 and 2 have different most probable unbinding force values (**Figure 5.2** (a) to (e)), but the most probable extension values are very similar and merged to one peak (peak 2 in **Figure 5.3** (a) and (b)) in the extension histogram. These experimental data indicate that PEG2000 linker and aptamer have their specific elastic properties during the binding-unbinding processes through different pathways.

5.2 Experimental Details

Under low loading rate 0.24 nN s^{-1}, 2.72 nN s^{-1}, 15.37 nN s^{-1}, 38.02 nN s^{-1}, and 51.64 nN s^{-1}, the force histograms showed three peaks (**Figure 5.2** (a) to (e)), and the extension histograms showed two peaks (**Figure 5.3** (a) to (e)). Under medium loading rates 82.06 nN s^{-1} and 100 nN s^{-1}, force histograms showed two peaks (Under high loading rates 200 nN s^{-1}, 400 nN s^{-1}, and 600 nN s^{-1} (**Figure 5.2** (h) to (j), **Figure 5.3** (h) to (j)), only one peak was observed in both the force histograms and extension histograms.

For the PEG2000 linker molecule (as part of the series-spring model in **Figure 5.1** (a)), a modified WLC model was used to fit the stretching of PEG2000 in the force-extension curves.[132] The contour length of 15 nm, persistence length of 0.38 nm, and enthalpic correction of 1561 pN were used for the extended WLC model of PEG2000[132] and directly adopted in our calculations. The simulations of PEG extension, spring constant (k_{PEG}) calculations, and experimentally measured extension values under different loading rates will be discussed later.

5.3 Force and Extension Histograms

In both **Figure 5.2** and **5.3**, (a) is 0.24 nN s^{-1}, (b) is 2.72 nN s^{-1}, (c) is 15.37 nN s^{-1}, (d) is 38.02 nN s^{-1}, (e) is 51.64 nN s^{-1}, (f) is 82.06 nN s^{-1}, (g) is 100.00 nN s^{-1}, (h) is 200.00 nN s^{-1}, (i) is 400.00 nN s^{-1}, and (j) is 600.00 nN s^{-1}. The Gaussian fitting was used to obtain each peak in these histograms.

For the extension histogram under each loading rate, 250 f-d curves were collected. In each f-d curve, the baseline, aka zero force, was determined from the right part of the f-d curve (the flat part in the right side of individual force-extension curve). Next, the point of zero force in the left part of the f-d curve was assigned as the starting point to measure the unbinding extension values (x-axis zero). The ending point of the extension was the extension value where the force has highest value (x-axis in various values). Therefore, the 250 extension values measured under one loading rate were put together to establish the extension histogram. The two highest peak values from **Figure 5.3** (a) to (g) were determined by the Gaussian fitting of two peaks in each histogram. The single highest peak values from **Figure 5.3** (h) to (j) were determined by Gaussian fitting of one peak in each histogram.

5.4 Biophysical Model for the Elasticity of Aptamer

Based on the series-spring model in **Figure 5.1** (a), the total extension z is the sum of PEG extension (z_1) and aptamer extension (z_2) under the same unbinding force. Therefore, we can estimate the apparent spring constant of aptamer (k_{apt}) during the entire stretching and unbinding process, and this parameter is directly related to the three unbinding pathways. If we ignore any detailed structural and environmental effects, the k_{PEG}, k_{apt}, unbinding force f, and total extension z_{total} should follow the relationships in the series springs:

$$z_1 = f_{PEG} \times k_{PEG} \tag{5.1}$$

$$z_2 = f_{apt} \times k_{apt} \tag{5.2}$$

$$f_{PEG} = f_{apt} \tag{5.3}$$

$$z_{total} = z_1 + z_2 \tag{5.4}$$

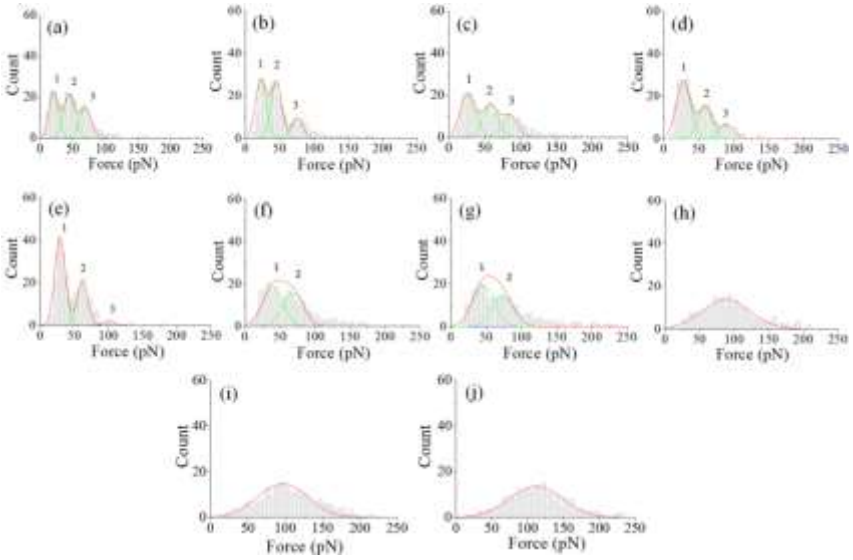

Figure 5.2 The force histograms under different loading rates.

The PEG2000 is a typical long chain polymer molecule. According to MLC model, the spring constant of PEG2000 (k_{PEG}) is dependent on its extension $z_1$during the extension process. Their relationship can be described as the equation below:[132]

$$k_{PEG}(z_1) = \frac{df_{PEG}(z_1)}{dz_1} = \frac{k_B T}{L_P L_0}\left[0.5\left(1-\frac{z_1}{L_0}\right)^{-3}+1\right] \tag{5.5}$$

Here *Lp* is the persistence length, L_0 is the counter length, k_B is the Boltzmann constant, whose values are all obtained from literature.[132] All of the k_{PEG} values were calculated according to this model.

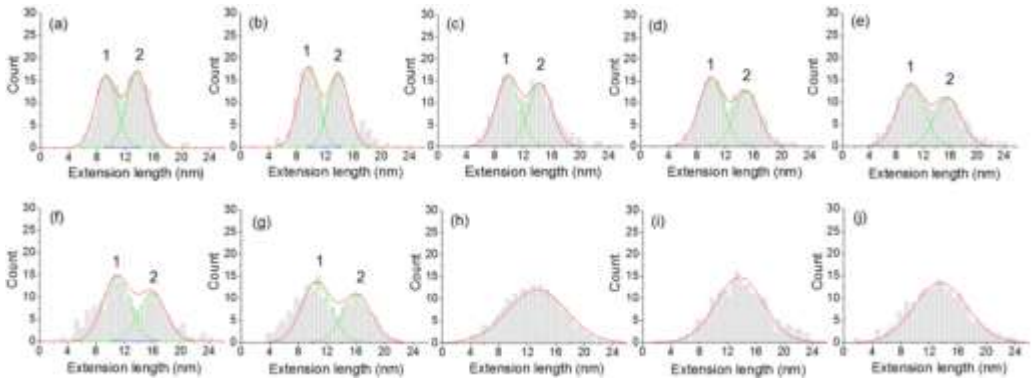

Figure 5.3 The extension histograms under different loading rates.

The details of these calculations were shown in **Table 5.1** and **Table 5.2**. The k_{PEG} and k_{apt} values estimated by this series-spring model are listed in **Table 5.3**. Under low and medium loading rates, the PEG in WLC model showed increasing k_{PEG} values with the increase of force.[132] For the aptamer, different unbinding pathways showed their corresponding k_{apt} values, which were not affected by the changes of loading rates. This phenomenon is especially very clear for pathway 2 through a wide range of loading rate, which indicates that the aptamer binding conformation in this pathway showed very consistent elastic properties during the unbinding reaction.

The data from **Table 5.3** were also used to study the influence of different unbinding pathways on the elastic properties of PEG and aptamer. Pathway 1 can be clearly distinguished with the force histograms (**Figure 5.2**) and extension histograms (**Figure 5.3**), but the aptamer didn't show a consistent k_{apt} value under different loading rates. Under lowest loading rates 0.24 nN s^{-1}, 2.72 nN s^{-1}, and 15.37 nN s^{-1}, the estimated aptamer extensions continued to decrease until it is even smaller than the WLC extension of PEG2000. One reason is that the errors of the measured unbinding forces will significantly increase from a combination effect of lower loading rates (e.g. 10 nN s^{-1}) and longer linker molecule.[129] Here the linker PEG2000 is about 15 nm, which can cause significant errors in the measured force values, and in turn lead to the large errors in the

estimated PEG extension values. Therefore, the simulated PEG extension values could be overestimated. Another reason for the large error is that the interaction through pathway 1 was so weak that it was easily interfered by the errors from the experimental measurements.

The aptamer in pathway 1 always showed the smallest unbinding force value under each loading rate. Therefore, the energy barrier for the unbinding process through pathway 1 is always the lowest compared to other pathways under the same loading rate. The PEG extension to overcome this low energy barrier (z_1) dominated the elastic process in pathway 1, and the aptamer extension (z_2) became comparable to the thermal noise in DFS measurements. According to the calculations in **Table 5.1**, the z_2 values for pathway 1 became invalid under relatively larger loading rates, which are mainly caused by the two reasons mentioned above. On the contrary, based on molecular dynamics and docking simulations, the aptamer in pathway 2 and 3 show more contact areas to ricin surface (**Figure 5.1** (b)), which indicate that the unbinding forces generated from these two pathways were large enough to dominate the extension measurements. Consequently, these two pathways showed very consistent z_2 values under different loading rates, and in turn maintained very consistent k_{apt} values in a large range of loading rates (**Table 5.1** and **5.2**).

Overall, the calculations of k_{apt} values from the extension distributions are consistent with our multiple-barrier multiple-pathway model and can be used to distinguish the three pathways. The disagreement between the aptamer spring constants calculations for Pathway 1 and the ones for other pathways showed the limitation of the series-spring elasticity model, which works well for certain experimental conditions and aptamer conformations. However, the multiple-pathway model is still valid for all of the experimental histograms in **Figure 5.2** and **Figure 5.3**. The k_{apt} for pathway 2 fitted the series-spring model best in a large range of loading rates. The k_{apt} for pathway 3 showed the same consistency as long as it existed in force histograms under low loading

rates. Under the high loading rates, the unbinding processes happened through a new pathway with different unbinding energy barrier,[133] and the three-pathway model will not be valid any more.

Overall, these consistent k_{apt} values for their corresponding pathways proved that the aptamer-ricin complex has its specific elastic properties for different unbinding pathways. The k_{apt} values can be used as a special elastic parameter to distinguish different unbinding pathways under proper experimental conditions.

5.5 Results and Discussion

To test the reliability of these estimated spring constants for different pathways, we simulated the force-extension curves using the series-spring model in **Figure 5.1** (a). The unbinding force value was set by the measured force-extension curve, and the estimated k_{apt} values of the three pathways were used to construct their corresponding total extension values. Finally, the experimental extension value was compared with the simulated ones, and classified into one of the three pathways. The comparison of experimental force-extension curves and the simulated ones under loading rate 51.64 nN s^{-1} are shown in **Figure 5.4**, where (a) is the simulation for pathway 1, (b) for pathway 2, and (c) for pathway 3. The arrows I, II, and III labeled the simulated curves for Pathway 1, pathway 2, and pathway 3, respectively.

Under each loading rate, the experimental curve (black) was compared with three simulated curves constructed for their corresponding pathways. When k_{apt} value was not available, the simulated PEG extension curve was used to represent the total extension. These experimental curves can be called “the most probable curves” for their corresponding pathways. For example, the experimental curve in **Figure 5.4** (b) (black) fits both force and extension values of the simulated curve for pathway 2 (green, labeled as II), which indicates that this experimental curve

is the most probable unbinding curve for pathway 2 under 51.64 nN s^{-1}. The simulated curves can help determine the experimental "most probable curves".

Table 5.1 The experimental and simulation data for PEG and aptamer extensions under low loading rates.

Loading rate (nN s^{-1})	0.24			2.72			15.37			38.02			51.64		
Pathway	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Experiment force (pN)	20.2	44.6	69.6	22.5	45.0	75.6	26.8,	57.7	85.4	27.9	58.8	89.7	29.0	61.7	101.2
Experiment z_{total} (nm)	9.3	13.7	13.7	9.5	13.9	13.9	9.8	14.4	14.4	9.9	14.8	14.8	10.12	15.6	15.6
Simulated PEG z_1 (nm)	9.10	11.47	12.56	9.45	11.49	12.75	10.01	12.11	13.03	10.16	12.16	13.15	10.27	12.28	13.46
PEG spring constant k_{PEG} (N/m)	0.007	0.029	0.085	0.008	0.029	0.108	0.011	0.052	0.161	0.012	0.054	0.194	0.012	0.062	0.336
Aptamer z_2 (nm)	0.20	2.23	1.14	0.05	2.41	1.15	-0.21	2.29	1.37	-0.26	2.64	1.65	-0.15	3.32	2.14
Aptamer spring constant k_{apt} (N/m)	0.101	0.020	0.061	0.450	0.019	0.066	N/A	0.025	0.062	N/A	0.021	0.054	N/A	0.019	0.047

Table 5.2 The experimental and simulation data for PEG and aptamer extensions under medium and high loading rates.

Loading rate (nN s^{-1})	82.06		100		200	400	600
Pathway	1	2	1	2			
Experiment force (pN)	39.6	67.5	45.9	73.1	87.6	97.2	110.6
Experiment z_{total} (nm)	10.91	16.01	11.42	16.22	13.31	13.56	13.84
Simulated PEG z_1 (nm)	11.10	12.50	11.51	12.64	13.10	13.32	13.59
k_{PEG} (N/m)	0.021	0.079	0.029	0.089	0.179	0.250	0.447
Estimated aptamer extension z_2 (nm)	-0.19	3.51	-0.09	3.58	0.21	0.24	0.25
k_{apt} (N/m)	N/A	0.019	N/A	0.020	0.417	0.405	0.442

Table 5.3 The estimated spring constants for PEG2000 and aptamer molecules.

Loading rate (nN s^{-1})	PEG spring constant k_{PEG} (N/m)			Aptamer spring constant k_{apt} (N/m)		
Outer barrier pathway	1	2	3	1	2	3
0.24	0.007	0.029	0.085	0.101	0.020	0.061
2.72	0.008	0.029	0.108	0.450	0.019	0.066
15.37	0.011	0.052	0.161	N/A[a]	0.025	0.062
38.02	0.012	0.054	0.194	N/A[a]	0.021	0.054
51.64	0.012	0.062	0.336	N/A[a]	0.019	0.047
82.06	0.021	0.079	N/A[b]	N/A[a]	0.019	N/A[b]
100.00	0.029	0.089	N/A[b]	N/A[a]	0.020	N/A[b]
Inner barrier pathway						
200.00	0.179			0.417		
400.00	0.250			0.405		
600.00	0.447			0.442		

[a] The estimated z_1 is longer than the measured z_{total}, therefore the series-spring model is not valid here.
[b] The force histograms under loading rates 82.06 nN s^{-1} and 100 nN s^{-1} didn't show the peak for pathway 3.

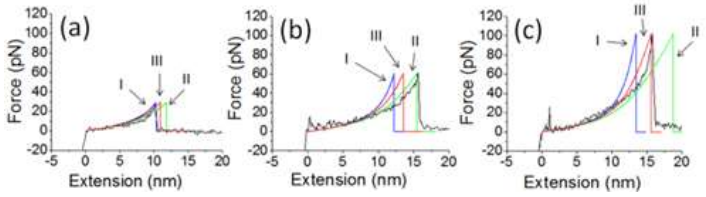

Figure 5.4 Representative experimental and simulated force-extension curves under 51.64 nN s^{-1}.

The standard deviation (±1σ) of each Gaussian fitting peak in **Figure 5.2** and **5.3** was used to classify each experimental force-extension curve into a certain pathway or into the "unknown" group. Some experimental curves have the extension values drifting so far away that the difference between the measured extension and any of the three simulated extension is larger than 1σ. Those curves were classified as "unknown" for each loading rate. After each individual force-extension curve was checked, a new histogram was made to show the curves distributions for the three

pathways and unknown groups under the same loading rate (**Figure 5.5).** For example, under loading rate 0.24 nN s^{-1}, 97 out of 250 force-extension curves can be classified into one of the three unbinding pathways. Among them, 43 curves were classified into pathway 1 (labeled as I), 30 curves were classified into pathway 2 (labeled as II), and 24 curves were classified into pathway 3 (labeled as III). The rest 153 curves were classified into unknown group, labled as "Unknown" in **Figure 5.5** (a). Therefore, 61% (153) of the f-d curves can not be determined to a certain pathway by this analysis. That ratio is larger than the one simply determined by 1σ deviation of the extension histogram (1 - 68% = 32%).

Overall, the combination of force and extension as the criteria can put more curves to the unknown group, which indicates those curves may not be specific for any ricin-aptamer unbinding pathway. With the combined restrictions of both force and extension values, the number of non-specific force-extension curves in the histograms decreased significantly. The uncertainty for the remaining curves was greatly reduced and can be used more confidently as specific f-d curves for ricin-aptamer interactions.

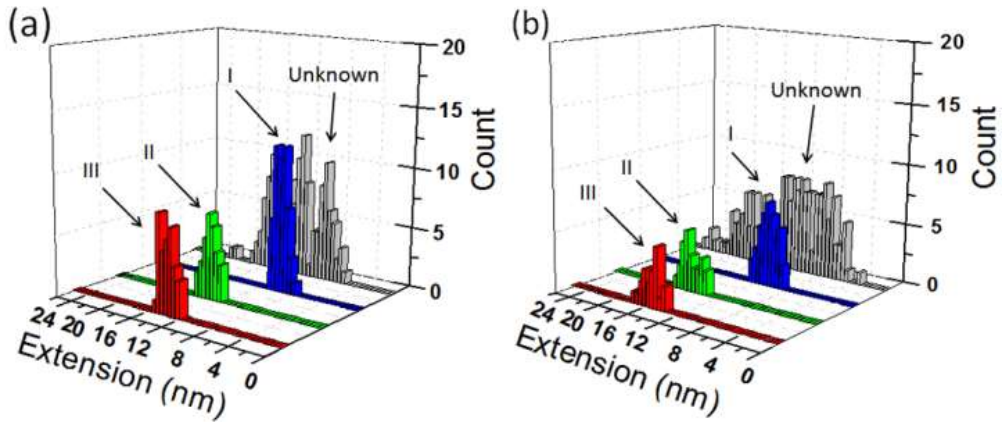

Figure 5.5 The reconstitutions of original extension histograms under two representative loading rates.

5.6 Conclusion

The method we developed here can be used to distinguish the specific unbinding pathways for individual force-extension curves measured in a multiple-pathway system. This analysis is limited by the thermal noise, errors caused by the linker molecule, and certain binding conformations. However, the experimental data and theoretical simulations agreed with each other in a relatively simple series-spring elasticity model. The principles used in this model can be applied to the studies of other systems with multiple unbinding pathways, such as the interactions of biomolecules on cell surface or protein-nucleic acid interactions involved in cellular activities, where the non-specific interaction is always a serious issue for data analysis. With the help of this approach, DFS force-extension curves for a specific biophysical process or reaction pathway can be selected by it corresponding elastic parameters, while the uncertainty in the results of single-molecule experiments can be significantly reduced.

CHAPTER 6

SURFACE CONFORMATIONS OF AN ANTI-RICIN APTAMER AND ITS AFFINITY FOR RICIN DETERMINED BY ATOMIC FORCE MICROSCOPY AND SURFACE PLASMON RESONANCE[1]

[1]B. Wang, Z. Lou, B. Park, Y. Kwon, H. Zhang, B. Xu. 2015. *Phys. Chem. Chem. Phys.* 17:307-314.

6.1 Introduction

This chapter will focus on the application of aptamer in the biosensor development. The fundamental research done in previous chapters provide the methods for the fabrications of SPR biosensor, and the test for the specificity of aptamers to ricin. Consequently, the aptamers studied before have been used as the probe molecules in this SPR biosensor for the label-free detection of ricin in bulk solution. The single-molecule studies were conducted to test the activity of the anti-ricin aptamer after it was immobilized onto the surface of the biosensor. The bulk solution measurements by SPR did show different sensitivity and affinity from the single-molecular DFS measurements. This comparison will help us understand the behaviours of aptamer in different environments, and guide the future biosensor design. Other practical issues such as reducing the non-specific interaction of the biosensor are also discussed in this chapter.

The specific 3-dimentional conformation is directly related to the functions and interactions of proteins and nucleic acids.[134,135] The surface immobilization of these biomolecules may lead to changes of their conformations and orientations on substrate, and in turn cause different affinities to their target species in biosensor measurements.[136,137] Therefore, the morphology and interaction information are both important for the studies of biomolecules on biosensor surfaces. AFM has the ability to detect single biomolecule on substrate surface with miniature amount of sample or at single-molecule level, and provide high resolution images and interaction information.[9,138] SPR is another widely used surface detection technique with advantages of easy to operate and rapid detections, so it become widely used as biosensors.[139,140] The surface modification is an important issue in SPR detections in order to reduce the non-specific interaction between probe molecule and substrate.[141,142] It has been widely accepted that the linker or spacer molecules on the substrate will improve the binding affinity of the probe molecule. However, the surface modification may

also have negative effect on the properties of probe molecule, and SPR data cannot show these detailed changes because of the limitation from bulk measurements. AFM single-molecule measurements have the ability to reveal detailed properties of the biomolecules with less external interference from bulk solution environment.[11] Therefore, the comparison of the experimental results from SPR and AFM should be an important criterion for the developments of future SPR biosensors using probes with complex properties such as aptamer.[143,144]

Nucleic acid aptamers have been widely used as binding reagents for the label free detections of biomolecules.[145,146] Compared to antibodies, aptamers showed advantages such as easy synthesis, low cost, and better stability.[140,137] Therefore, aptamers can be integrated into various detection platforms and provide versatile applications in the fast detection and analysis related to biosafety and biosecurity, such as the detection of toxins in food products and public environment.[147,148] Based on the fast developments of aptamer research and its integration into many nanotechnologies, the aptamer based biosensors will provide more versatile methods and devices for the high throughput and label-free detection in biosafety.[149,150]

Ricin is a type-2 ribosome-inactivating protein (RIP) with two chains connected by a disulfide bond.[74] The relatively easy access of ricin from the caster-oil manufacturing makes it a more dangerous and frequently used toxin species. Ricin has been used as the poisoning agent for military warfare, in criminals, and in recent terrorism cases, but it also has great potential in the development of immunotoxins to treat cancers.[151,152] Here we used a certain anti-ricin aptamer to detect ricin in two platforms: AFM and SPR. AFM provided high resolution images of aptamer on Au(111) surface with and without carboxymethylated-dextran (CD) modification. The possible negative influence of the surface modification method on the aptamer specificity and sensitivity to ricin can be directly studied from the AFM topography and recognition images. The AFM dynamic

force spectroscopy (DFS) can be used to estimate off rate of the aptamer-ricin interactions4 and be compared with the off rate obtained from SPR measurement. Currently, only AFM can provide the information about morphologies and interactions of aptamer and ricin at single-molecule level in a single platform. The comparison of measured off rate values from AFM-DFS and SPR is also a good support for the possible influence of surface modification to aptamers on Au(111) surface (**Figure 6.1** (a) and (b)). The results will be used as reference for the further development of label-free biosensor devices in micrometer and nanometer scales.

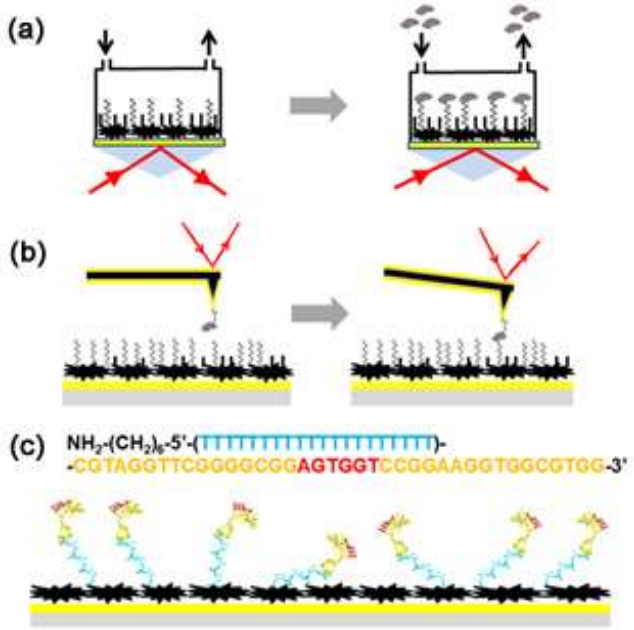

Figure 6.1 The experimental set-ups and aptamer immobilization for SPR and AFM.

6.2 Experimental Details

The anti-ricin aptamer sequence was obtained from literature.[30] In both AFM and SPR experiments, the anti-ricin sequence was modified by twenty thymine bases as the spacer and the amine group as the linker at 5' end.[153,154] The modified sequence was purchased from Integrated DNA Technologies (Coralville, IA, USA).

The ricin sample was provided by Vector Laboratories (Burlingame, CA). In AFM experiment, the ricin was attached by its amine group to the linker molecule and in turn to the gold coated tip. The polymer linker thiol-(polyethylene glycol)-acid (HS-PEG-COOH, M.W. 2000) was purchased from Creative PEGWorks (Winston Salem, NC, USA). The cystamine, carboxymethylated-dextran (CD), N-hydroxysuccinimide (NHS), and ethanolamine were all purchased from Sigma-Aldrich (St. Louis, MO, USA). The 1-(3-dimethylaminopropyl)-3-ethylcarbodimide hydrochloride (EDC) was purchased from Flucka Chemicals (Sigma-Aldrich, St. Louis, MO, USA). Phosphate buffer saline (PBS, pH 7.2) was purchased from Pierce (Thermo Scientific, Waltham, MA, USA). Triplet deionized water was provided by a Barnstead Nanopure Diamond Laboratory Water System.

In the SPR experiment, the sensor film was coated with Cys self-assembled monolayers (SAMs) via S-Au linkage by spreading a droplet (20 μL) consisting of cystamine hydrochloride (20 mM) onto the Au film overnight at room temperature. Then, the mixture solution of 15 mM EDC, 75 mM NHS, and 10 mg/mL CD was dropped on the cystamine modified Au surface for 2.5 h and the Au surface was cleaned with deionized water. After The sensor chip was mounted on the SPR prism with the matching oil, 100 μL of the mixture of 15 mM NHS and 75 mM EDC in deionized water was injected into the cuvette for 10 minutes to activate the carboxyl group of CD. Afterward, the aptamer solution (1 μM in PBS) was injected into the flow cell for its immobilization on the CD modified Au surface. Finally, to deactivate the remaining active sites on the sensor chip surface, 100 μL of 1 M EA in deionized water was injected. With the continuous flushing of the flow cell with deionized water, the ricin solutions with different concentrations were injected into the BI-2000 SPR system (Biosensing Instrument, Tempe, AZ) for the SPR measurements (**Figure 6.1** (a) right part). The off rate of aptamer to ricin in the SPR measurements

was estimated by the data analysis and kinetics analysis software shipped with the BI-2000 system. Flow rate was controlled at 20 μL per minute and all the detection solution was injected at the rate of 20μL per minute with a unified volume of 100 μL.

In both platforms in **Figure 6.1** (a) and (b), Au(111) surface was first blocked by carboxymethylated-dextran (large shapes in black). Next the aptamer molecules (coils in black) were immobilized. Finally, the surface was blocked with ethanolamine (EA, short sticks in black). The ricin molecules (in grey) were detected in bulk solution SPR experiments (a) and in single-molecule AFM experiments (b). The aptamer sequence is shown in (c), with the predicted conformations of aptamer molecules on gold surface. The 20 (dT) spacer is in cyan, the simulated anti-ricin folding structure in yellow, and the predicted binding residues in red. The folding structure of aptamer was predicted by AMBER 11.[24]

For AFM experiments, two surface modification methods were used and their AFM images were compared to investigate the possible influence of CD modification to the aptamer conformations on Au surface. In the first method (M1), the thiol modified aptamer was directly immobilized on gold substrate. In the AFM imaging and DFS experiments, TopMAC mode and PicoTREC module (Agilent Technologies, Santa Clara, CA, USA) were used to obtain the topography images and the recognition images, respectively. The ricin was attached to the gold coated AFM tip with HS-PEG-COOH linker molecule and EDC/NHS coupling reaction.[9] The same tip was used to measure the DFS force-distance curves. Five loading rates were used for the estimate of off rate, and 300 curves were collected for each loading rate. In the second method (M2), the amine modified aptamer was attached to the thin film of CD on Au surface. Therefore, the aptamer molecules were immobilized on Au(111) surface by the same method as in SPR measurement.

The off rate calculation in SPR was conducted using the commercial program "BI-KA" shipped with the SPR instrument. The off rate calculation in AFM-DFS was calculated by the same method mentioned in previous publications.[9]

6.3 Results and Discussion

6.3.1 High Resolution AFM Images

The high resolution topography and recognition images showed the conformations of individual aptamer molecules on Au(111) surface modified by M1 (**Figure 6.2** (a) and (b)) and by M2 (**Figure 6.2** (c) and (d)). The representative topography image of aptamer on Au(111) surface and the corresponding recognition image on Au(111) surface are shown in (a) and (b), (c) is the representative topography image of aptamer on CD modified Au(111) surface, and (d) is the corresponding recognition image on CD modified Au(111) surface. The scan size of all images is 500 by 500 nm^2. Eleven small areas were circled in (a) and (b) with their corresponding topography and recognition images. These highlighted areas showed the difference between the topography images and recognition images caused by different aptamer conformations. The aptamer conformations similar to number 1, 3, 4, 7, and 8 in (a) and (b) can be found in (c) and (d).

The aptamer has its stable folding conformation, but the topography images revealed different conformations of aptamer on Au(111) surface. In the topography images, the length of single aptamer molecule varies from 10 nm to 20 nm, and the width varies from 5 nm to 15 nm. Some aptamers maintained in their folding structures and had shorter shapes, while others extended their single strands on the gold surface with longer shapes. These various round-shaped and stripe-shaped images were caused by the flexibility of aptamer. Especially, the non-binding parts of the aptamer including the 20 (dT) spacer have more degree of freedom to change their conformations in the buffer and show various extensions or orientations. However, most of the

binding parts of the aptamer were still active and showed corresponding recognition signals in the recognition image (**Figure 6.2** (d)). Therefore, the binding conformations and specificity of ricin-aptamer interactions have been directly shown in the topography and recognition images by the single platform of AFM.

The AFM images on CD modified Au surface showed very similar gold terraces morphologies, although the thin film of CD made the edges of gold terraces blurred and added some film features on the gold surface (**Figure 6.2** (c)). The aptamers still showed clear folding conformations and recognition signals (**Figure 6.2** (d)). The numbered aptamers 1, 3, 4, 7, and 8 from **Figure 6.2** (a) and (b) can match similar aptamers in **Figure 6.2** (c) and (d). However, more aptamer molecules stayed very close to each other on the CD modified Au surface than on the one without CD modification, so the surface distribution of aptamers may be influenced by the CD modification. The additional CD thin film was certainly not as flat as the original Au(111) surface, so some aptamer molecules were buried inside the small holes generated from the CD modification, and some aptamers stretched their folding structures and changed their conformations. Therefore, less number of aptamers on CD modified Au(111) surface maintained their activity to ricin, and consequently showed less recognition images with their corresponding topography images. The AFM images proved that the CD modification, M2, did influence the surface morphology and aptamer activities to ricin.

The degree of this influence on the aptamer activity can be quantitatively studied by the AFM topography and recognition images. When one aptamer molecule showed its topography image but not its corresponding recognition image, that aptamer lost its activity to ricin. Therefore, large amount of AFM images of individual aptamer were counted to obtain the recognition ratios of aptamer on Au(111) surface with and without CD modification. The recognition ratio of aptamer

on CD modified Au(111) surface is calculated as 75.8%, while the recognition ratio of aptamer on Au(111) surface is calculated as 93.6%. This significant decrease of recognition ratio is mainly due to the changes of aptamer conformations on CD film and the less exposing of aptamer active residues caused by CD modification. So far, only AFM has the capability to investigate these detailed changes of surface properties and reveal the different activities of individual aptamer molecules on modified substrates.

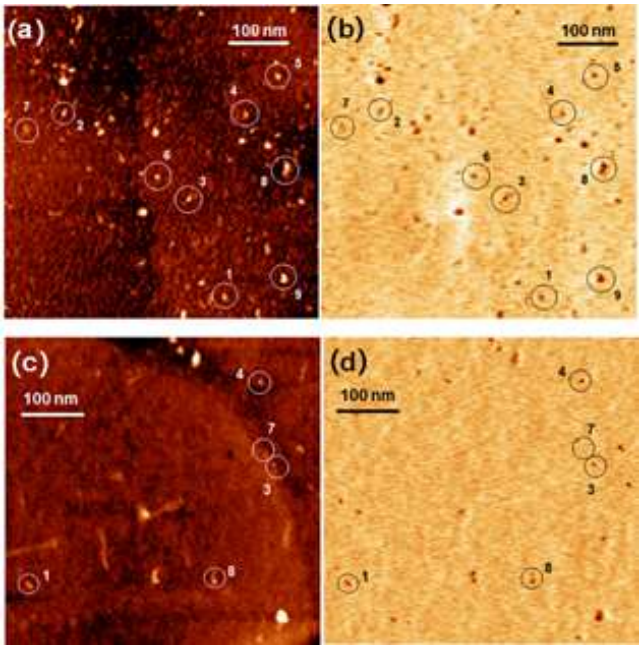

Figure 6.2 The AFM topography and recognition images of aptamer on Au(111) surface with and without CD modification.

In order to further investigate the aptamer conformations on Au(111) surface, eleven small areas were selected from **Figure 6.2** as the representative images of single aptamer molecule. They were compared with the simulated aptamer structure in different orientations. The details of topography and recognition images in **Figure 6.3** revealed the significant influence of different conformations on their binding activities to ricin. For example, the aptamer 1 in **Figure 6.2** (a) and (b) showed rod-shaped topography images. However, the recognition images showed strong

signals in the middle of each structure and weak signals on the two ends of each structure. These round-shaped recognition signals indicated that the ricin binding site located in the middle of each aptamer conformation. The two ends of each aptamer conformation just had very weak interactions with ricin, but the topography image still showed the structure outline. The top-view and side-view (rotated 90 degree from top-view) of aptamer 1 in **Figure 6.3** showed that the position of ricin binding site is pointing to the solution. Therefore, it can easily contact ricin molecule on the AFM tip. In **Figure 6.2** (a), the recognition signals of aptamer 3 is in one end of the structure, the other end showed clear topography image but very weak recognition structure. The other predicted binding sites for ricin were highlighted in the simulated aptamer structures. The aggregation of two or three aptamer molecules may exist in the selected areas 8 and 9. For each numbered area, topography image is in the left and recognition image is in the right. The AFM images in each row show similar alignment of the aptamer active residues (in red). The image size for each area from 1 to 7 is 25 by 25 nm^2. The image size for the area 8 is 30 by 30 nm^2. The molecular simulations of the aptamer folding conformations are listed in the middle column, non-active aptamer sequence in yellow, spacer sequence in blue. The simulated structures in top-view are shown in the left, the side-view (rotated 90 °) are in the right.

The AFM topography images cannot provide resolutions high enough to distinguish the possible aggregations of molecules in such small size. However, the recognition images showed clear signals of the small binding sites and in turn proved the aggregation. In the comparison of the experimental images and simulated folding structures, the aggregation can be reconstructed with the simulated aptamer structures in different orientations on gold surface. The non-binding parts of the two aptamers separated the smaller recognition signals, and provide enough resolution in recognition images to tell the number of molecules in the aggregations, which is corresponding

to the number of recognition spots. This analysis also confirmed the specificity of aptamer binding site for the ricin on AFM tip.

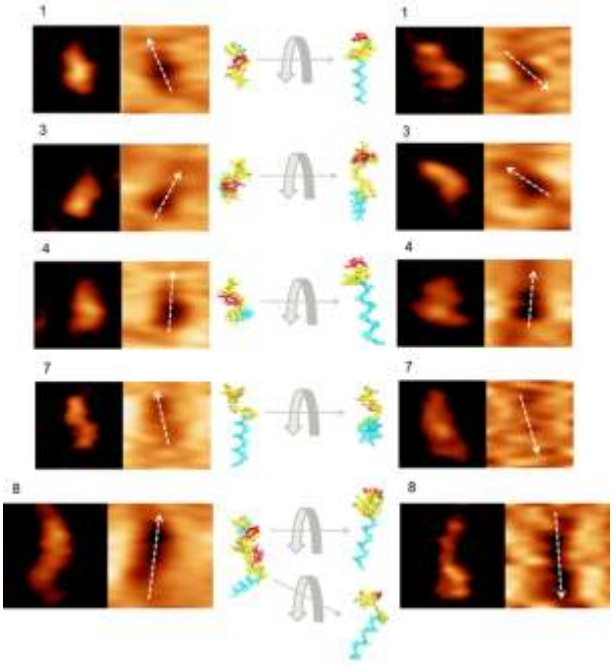

Figure 6.3 The zoom-in AFM topography and recognition images of aptamer.

The aptamer folding conformations simulated from molecular dynamics may not be accurate to represent all of the aptmaers observed on Au(111) surface due to the limits of the simulation methods. For example, some aptamers were distorted and randomly oriented since they are inherently more flexible than proteins. However, the simulated molecular outlines and the orientations of active sites are very helpful for the interpretation of different topography and recognition images in **Figure 6.2** and **6.3**. Here AFM successfully distinguished different aptamer conformations on Au(111) surface with and without CD modification. The detailed activity information from the recognition images revealed the influence of CD modification to the aptamer-ricin interactions.

6.3.2 Ricin and Aptamer Attached to AFM Tips

The affinity of aptamer to ricin can be evaluated by the off rate of the binding reaction (k_{off}). In single-molecule level, DFS was used to measure the unbinding force between aptamer on the Au(111) surface and ricin on the AFM tip. In previous DFS study, we immobilized ricin on Au(111) surface and attached the same anti-ricin sequence to AFM tip.[9] The same five loading rates were used in this DFS experiment and 300 force-distance curves were measured under each loading rate. The force histograms obtained in these five loading rates were compared with the previous ones using aptamer modified AFM tip. In **Figure 6.4**, (a) is 0.24 nN s^{-1}, (b) is 2.72 nN s^{-1}, (c) is 15.37 nN s^{-1}, (d) is 38.02 nN s^{-1}, and (e) is 51.64 nN s^{-1}. In each column, the top one was obtained by the ricin modified tip, the bottom one was obtained by the aptamer modified tip. Each histogram was constructed by 300 force-distance The most probable unbinding force (F*) under each loading rate (R) was obtained by Gaussian fitting. No significant difference was observed from the histograms obtained by aptamer and ricin modified tips.

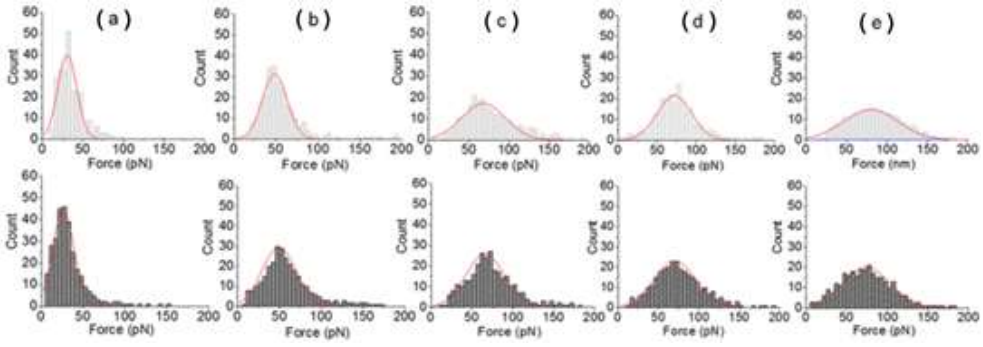

Figure 6.4 The force histograms obtained by DFS single-molecule measurements.

The Bell-Evans single-barrier model was used to estimate k_{off} value from the F* vs. lnR plot. To test the possible difference between the k_{off} using ricin modified tip and the one using aptamer modified tip, two F* vs. lnR plots were shown in **Figure 6.5** with their corresponding

error bars. The data points and fitting line obtained by aptamer modified tip is in black. The ones obtained by ricin modified tip is in red. The error bars showed the standard errors for each data point. The error bars for the data points under same loading rate showed large portions of overlaying, which indicates that the deviations between the data from two types of tip modifications were not significant. The off rate value obtained by aptamer-modified tip ($k_{off}1$) is $(6.8\pm0.9)\times10^{-4}$ s^{-1}. The off rate value obtained by ricin-modified tip ($k_{off}2$) is $(7.3\pm0.4)\times10^{-4}$ s^{-1}.

Therefore, the affinities estimated by these two single-molecule experiments are considered as in good agreement. These two off rate values are also close to the one obtained by antibody modified AFM tip.[70] The low off rate values indicate high affinity of aptamer to ricin in the AFM single-molecule measurements. The CD modification should not significantly influence the DFS measurements, because the force-distance curves were obtained from aptamers with their specific activity to ricin. The aptamers lost their activity did not show specific force-distance curves.

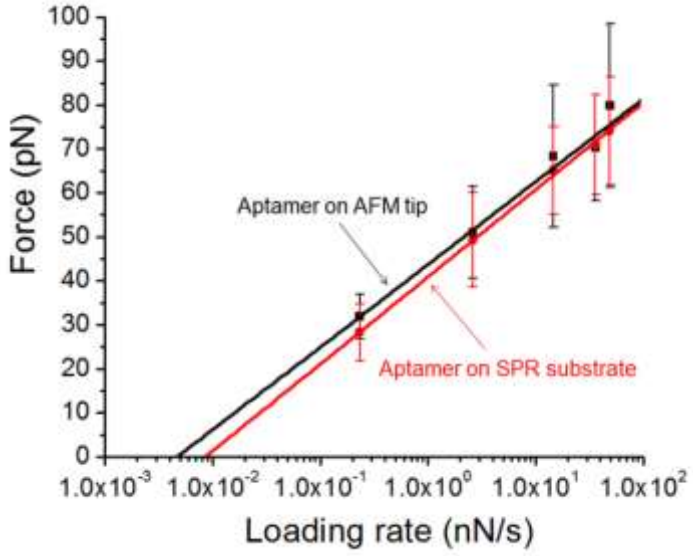

Figure 6.5 The two F* vs. lnR plots.

6.3.3 SPR Measurements

In the real-time SPR detection, six ricin concentrations 0.083 nM (5 ng/mL), 0.167 nM (10 ng/mL), 0.83 nM (50 ng/mL), 1.67 nM (100 ng/mL), 4.17 nM (250 ng/mL), and 8.33 nM (500 ng/mL) were used to test the limitation of detection and estimate k_{off} value. The angle change values increased with the increasing of ricin concentrations. The representative SPR sensorgrams from these six ricin concentrations are shown in **Figure 6.6**.

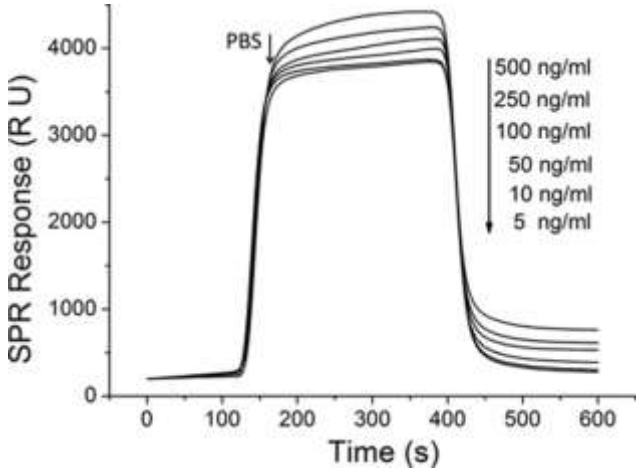

Figure 6.6 The SPR data used in the fitting of ricin-aptamer 1:1 binding mode.

The off rate of aptamer to ricin in the SPR measurements was estimated by aptamer-ricin 1:1 binding mode in the kinetics analysis software. **Figure 6.7** (a) showed the plot of ricin concentration vs. SPR signal obtained from the averaged angle shift values. The error bars represent the standard errors of three repeatable injections for each concentration. High ricin concentrations showed large errors, which are the major sources of the error in the estimate of k_{off} value. **Figure 6.7** (b) was used to estimate the limit of detection for the SPR measurement of ricin. The extrapolate of the fitting straight line at x-axis showed that the limit of detection for ricin should be around 1.5 ng/mL (25 pM). The final calculated k_{off} in SPR bulk solution measurement ($k_{off}3$) is $(1.82\pm0.067)\times10^{-2}$ s^{-1}. It is more than one magnitude larger than the ones obtained in AFM

single-molecule measurement. The deviation is also much bigger than the ones in AFM experiments. Therefore, the difference of off rate values between bulk solution and single-molecule measurements has been clearly revealed from the experimental data in AFM and SPR. Moreover, the AFM experimental results at single-molecule level can be used to connect the difference of off rate values to aptamer conformations on Au(111) surface. The details are discussed in next section.

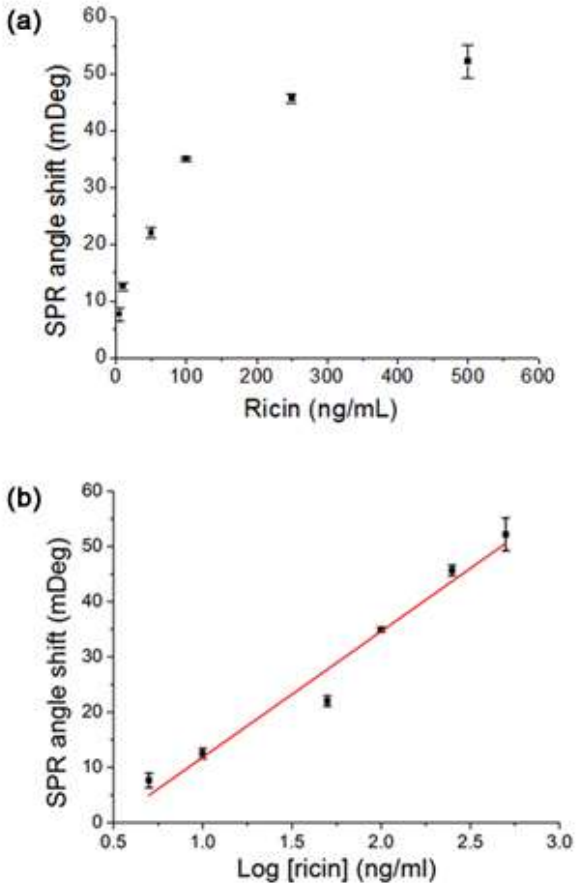

Figure 6.7 The fitting plot and the estimate of limit of detection.

6.3.4 Comparison of AFM and SPR Measurements

Figure 6.8 compared the k_{off} values and their standard errors in the same scale. The $k_{off}1$ is the value obtained by aptamer modified tip in DFS. The $k_{off}2$ is the value obtained by ricin

modified tip in DFS. The $k_{off}3$ is the value obtained by SPR. Error bars showed their standard errors. The k_{off} values ($k_{off}1$ and $k_{off}2$) obtained from two AFM single-molecule measurements don't show significant difference in statistics, but they are much smaller than the one obtained by SPR bulk solution measurements ($k_{off}3$). Moreover, the error for $k_{off}3$ is larger than the values of $k_{off}1$ and $k_{off}2$. Based on this comparison, the SPR measurements showed relatively high sensitivity and specificity for aptamer-ricin detection. However, the off rate estimate based on this bulk solution measurements was different from AFM single-molecule measurements. This difference will show greater impact to the measured parameters when the detection devices reach nanometer scale, such as the detections in nanoarrays or microfluidic channels. In the nanometer scale, the major factors that can cause the deviation include the nonspecific interactions, the influence from solvent, and interface properties.[155],[156] When the developments of biosensors move to the nanometer scale, the thorough study on the difference between bulk solution measurements and single-molecule measurements will provide important information for the further developments of surface modifications and signal analysis.

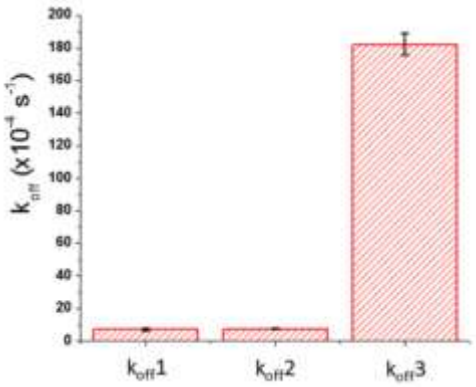

Figure 6.8 Comparison of off rate values obtained by AFM and SPR measurements.

For this study of ricin-aptamer interactions in SPR, the large deviation of k_{off} is even comparable to the k_{off} values in single-molecule measurements. This large error is the common

problem for bulk solution measurement, although SPR can detect ricin in pM level. One source of this large error came from different aptamer conformations on Au surface, partly caused by the CD modification. **Figure 6.2** clearly showed that some aptamer conformations were changed by the CD modification, which caused different recognition ratio for aptamer-ricin interactions. The recognition signals in these images indicated different binding/unbinding interactions during the tip scanning.

In the bulk solution measurements, all of these conformations have the opportunities to interact with ricin molecules with specific or non-specific forces, but the apparent values of the off rate ($k_{off}3$) and affinity are the averaged ones based on large amount of various aptamers. During the surface modification and aptamer immobilization, different chemical and physical environments on the substrate and in the solution could influence the flexible aptamer conformations on gold surface and cause the change of conformations and orientations. Therefore, the averaged off rate value ($k_{off}3$) has a much larger experimental error than the single-molecule measurements.

The standard errors of the two single molecule measurements are also different. In **Figure 6.8**, the aptamer modified AFM tip gave a relatively large error bar for $k_{off}1$, and the ricin modified tip gave a relatively small error bar for $k_{off}2$. When the ricin modified tip interacted with the aptamers on Au(111) surface, the PEG linker provide enough degrees of freedom for the ricin, and the flexible aptamer structure also improve the probability of aptamer binding site to interact with ricin binding site.

On the contrary, the ricin molecules immobilized on Au(111) surface has less degrees of freedom because of the short linker molecule used in that study.[157] Although the aptamer modified tip had enough degrees of freedom, the more rigid ricin conformations on the substrate had more

interference on aptamer-ricin binding events. Therefore, the error bar for $k_{off}1$ is larger than the one for $k_{off}2$.

6.4 Conclusion

We investigated the specific conformations and interactions of anti-ricin aptamer in two different platforms, the AFM and SPR instruments. The topography and recognition images obtained by AFM illustrated the various binding conformations and specific binding site of aptamer molecules on Au(111) surface with and without CD modification. The off rate of ricin-aptamer single-molecule interactions were measured by AFM-DFS as $(7.3\pm0.4)\times10^{-4}$ s^{-1}. This off rate value obtained by ricin modified AFM tip is statistically consistent with the one obtained by aptamer modified tip.[9] This comparison shows that the off rate value in DFS measurement is not effected by which molecule is attached on the tip and which one is immobilized on the substrate for the ricin-aptamer interactions.

In SPR experiment, the ricin sample solutions were detected by aptamer immobilized on CD modified gold substrate as low as 83.3 pM (5 ng/mL). The off rate estimated by SPR sensorgrams is much larger, as $(1.82\pm0.067)\times10^{-2}$ s^{-1}. The very large standard error in the SPR result indicates that the bulk solution measurements were influenced by the various aptamer conformations at single-molecule level. Additionally, the bulk solution measurements included more interference by non-specific interactions from modified environment conditions. The difference between single-molecule measurements and bulk solution measurements are significant, and the single-molecule

AFM images provide more detailed information for individual aptamer binding conformations and activities on CD modified Au(111) surface. The AFM force measurements also provide more precisely techniques to control the ricin-aptamer binding events. This work provides

fundamental information for the aptamer as probe molecule on biosensor surface. The results will facilitate the understanding of aptamer binding mechanism and the developments of aptamer based nanoarrays and other label-free detection devices using nanotechnologies.

CHAPTER 7

LABEL-FREE DETECTION OF SALMONELLA TYPHIMURIUM WITH APTAMERS

7.1 Introduction

The detection of certain bacteria serotype is a challenging for current biosensors. However, many toxic pathogens are certain serotypes, not all of that bacteria species. Therefore, it is important to develop detection methods that can distinguish the serotypes. Here aptamers are used as probe molecules for the detection of *Salmonella* Typhimurium by both single-molecule (AFM) and bulk solution (SPR) methods. Many protocols and methods used in previous studies on aptamer-ricin interactions have been applied into this study. The comparison of the quantitative measurements from AFM and SPR is important for the development of biosensors, because the probe molecules (aptamers) may show different behaviors in different environments. When the dimensions of biosensors go down to nanometer scale, the biosensor components, such as the surface modification, linker molecules, and probe molecules, will show their unique physical and chemical properties, which may be different from the ones in large scale measurements. Therefore, the combination of SPR and AFM will provide critical reference for the future development of bacteria biosensors in nanometer scale.

The label-free detection techniques have been widely used in biosensors. Compared to label-based techniques such as fluorescence and radiolabelling, the label-free techniques have several critical advantages. The labelling step may introduce heterogeneity of the sample, and in turn require further purification process and sample losses.[158] On the contrary, label-free methods avoid the labelling steps that consume extra time and resources. Moreover, some markers used in

the labeling may affect the activity or functionality of the sample and reduce the reliability of the detection. The label-free methods employ the inherent properties of the samples for the signaling, and can maintain the samples' activities.[158] Currently, the most important challenges in label-free detection are to improve the sensitivity and specificity. Therefore, various sensing strategies have been applied into label-free methods, such as quartz crystal microbalance (QCM), colorimetry, electrochemical methods, mass spectrum, field-effect transistors (FET), magnetoelastic (ME) sensors, and optical-based biosensors such as SPR.[158,159] The emerging nanotechnologies have provided more opportunities for the development of label-free biosensors. AFM is the single-molecule method that has been widely used in the detection of biospecies such as DNA, proteins, and cells. The high resolution images and single-molecule force measurements provide the highest space resolution compared to other label-free methods. On the other hand, SPR biosensors are one-dimensional nanoscale devices for the label-free and real-time detection of biospecies. However, the signals acquired from SPR are generated by the bulk sample solution. Therefore, the previous studies in Chapter 6 have been extended into the label-free detections of bacteria, which have much larger size than proteins and should show their unique properties in both single-molecule detections and bulk solution detections. The methods and protocols developed in previous chapters have been used in this study.

7.2 Single-Molecule Detection of Bacteria Using AFM

7.2.1 Chemical Modification of Au(111) Surface and AFM Tip

In order to reduce the non-specific interactions for biosensor development, the Au(111) surface used in AFM and SPR detection is always blocked with CD. The detailed chemical reactions have been shown in Chapter 6. The aptamers used here have specific interactions with outer membrane proteins (OMPs) on *Salmonella* Typhimurium surface.[160] The two sequences

APT33 and APT45 can be seen in **Figure 7.1** (a) and (b). Both of them were selected with high affinity to the OMPs of *S*. Typhimurium in literature, but their folding structures are different, which were predicted by Mfold webserver and Amber molecular dynamics, and shown below their sequences, respectively.[160] The loops predicted by the Mfold simulations are highlighted in red. The different folding structures of these two aptamers should lead to different binding sites and affinities to OMPs. However, no detailed structural information about OMPs is available. Especially, when those proteins are embedded in the outer membrane of *S*. Typhimurium, it is very difficult to measure or simulate their conformation and affinities. Therefore, AFM imaging and DFS measurements have become the critical methods to investigate the *in situ* interactions between the aptamer and OMPs.

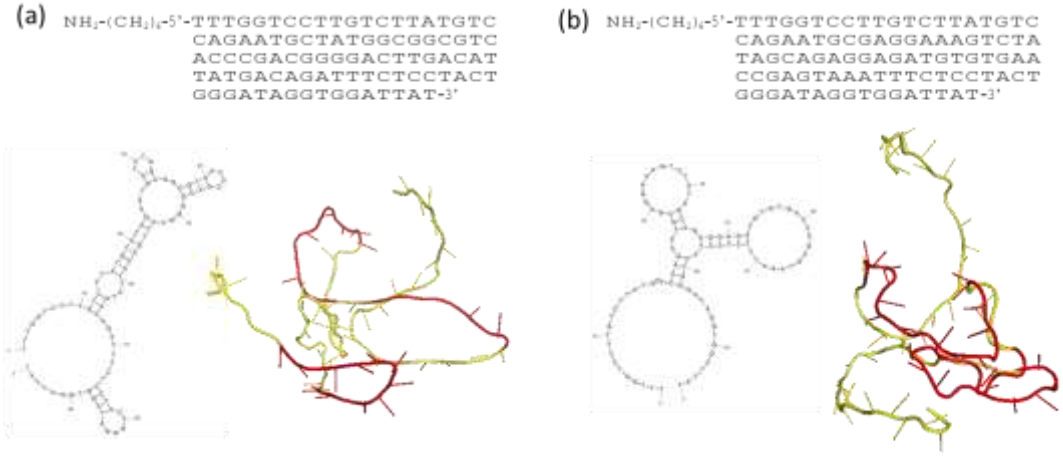

Figure 7.1 The sequences of two aptamers and their predicted folding structures.

For AFM imaging and force measurements, the CD was used to form a thin film on the Au(111) surface so that entire surface was blocked. Consequently, the aptamer was immobilized on the CD film with the help of EDC-NHS coupling reaction so that the *S*. Typhimurium bacteria body can be caught by the aptamer. Next, other unoccupied carboxyl groups on the CD film were

blocked with EA. The *S*. Typhimurium sample was provided by USDA, ARS, US National Poultry Research Center in Athens, GA. The bacteria were re-suspended in D.I. water to make the *S*. Typhimurium sample solution, which was directly dropped into the liquid cell, and cultivated for 3 hours. The liquid cell was washed with DI water three times and ready for the imaging process. The AFM tip was modified with the aptamer so that both topography and recognition images can be simultaneously acquired. All experiments from bacteria preparation to immobilization on Au(111) were conducted under the biosafety cabinet in the BSL-II USDA laboratory.

7.2.2 The AFM Images

The AFM images were all obtained in PBS pH 7.2 buffer. The topography images of *S*. Typhimurium and *E. coli* in PBS are shown in **Figure 7.2** (a) and (b), respectively. The surface morphologies of these two bacteria are similar, with the length of around 1.5 to 2 μm, the width in the range of 600 nm to 1 μm, and the height varying from 200 to 300 nm. These images have relatively low resolution, which is caused by the large thermal noise and hydraulic disturbance during the tip scanning in the buffer solution. Moreover, each individual bacteria body might approach to the aptamers on gold surface via different orientations, and in turn maintained its drifting state during the tip scanning. The AFM images in PBS have revealed more details about the states of bacteria and their interactions with aptamers on substrate.

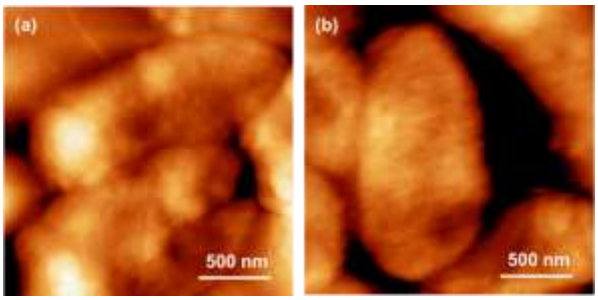

Figure 7.2 AFM topography images of *S*. Typhimurium in (a) and *E. coli* in (b).

The recognition images of *S.* Typhimurium were also obtained on CD-modified Au(111) surface with the APT33 modified tip. **Figure 7.3** (a) shows the representative recognition image of an bacteria body, image size 2400 nm x 2400 nm, while **Figure 7.3** (b) shows the enlarged image of *S.* Typhimurium surface from the area highlighted by red frame in (a). The size of this image is 600 nm x 600 nm. The enlarged AFM recognition image shown in **Figure 7.3** (b) indicates the multiple binding areas on *S.* Typhimurium surface for the APT33-OMPs interactions, shown as the dark spots in the recognition image. They are not evenly distributed and have different sizes ranging from 15 nm to 35 nm in diameter. These recognition signals correspond to the locations and sizes of several OMP concentrated areas on bacteria surface. The dark spots with larger size may indicate aggregation of more OMP molecules, and the one with smaller size could be the single molecule of OMP. However, this result should be carefully investigated with the help of structural information of OMPs. Further study should be focused on the *in situ* structure-function relationship of aptamer and OMP, which will provide essential information for the development of biosensors for *S.* Typhimurium detections. The selectivity of APT33 and APT45 should also be tested by other bacteria such as *E. coli*.

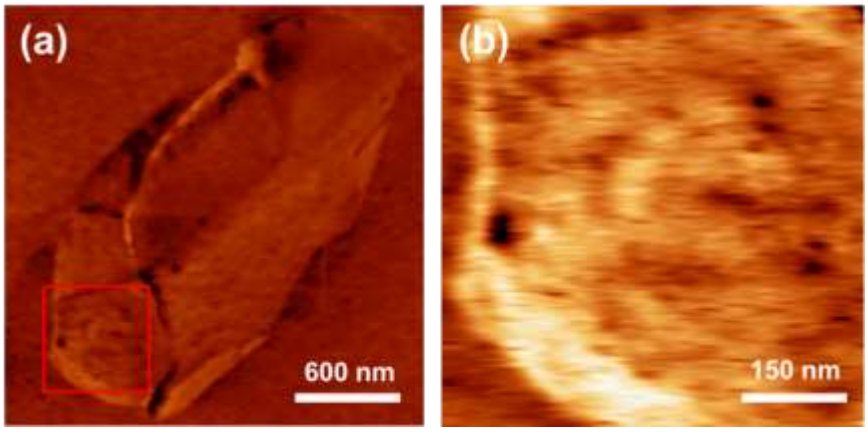

Figure 7.3 The AFM recognition images of *S.* Typhimurium to reveal OMPs on bacterial surface, obtained by the aptamer #33 attached to AFM tip.

7.2.3 DFS Analysis

The DFS technique has been proved to be a powerful tool for the study of single-molecule structure-function relationship in previous chapters. Here the similar methods have been used to investigate the *in situ* interactions between aptamers and OMPs. The experimental setup is shown in **Figure 7.4** (a). The aptamer modified tip was moved into the OMP concentrated area, and the tip will move in the direction perpendicular to the bacteria surface and obtain the force-extension curve. Most curves obtained in the measurements have shown more than one unbinding peak, as shown in **Figure 7.4** (b).

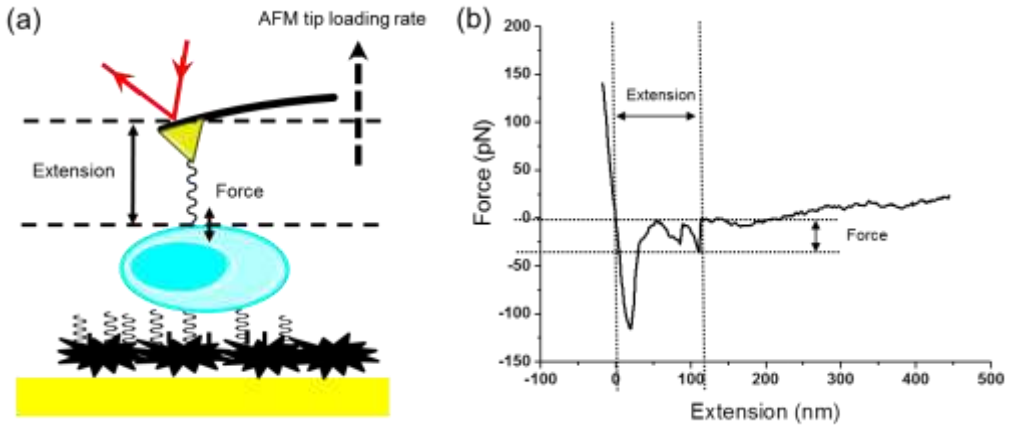

Figure 7.4 The schematics of DFS measurements on bacteria surface.

The force histograms and extension histograms have been constructed under different loading rates, ranging from 1 nN s^{-1} to 400 nN s^{-1}. Under each loading rate, 300 force-extension curves were collected to construct the force histogram and the extension histogram. **Figure 7.5** shows the force histograms under these seven loading rates (R): 1 nN s^{-1}, 10 nN s^{-1}, 20 nN s^{-1}, 50 nN s^{-1}, 100 nN s^{-1}, 200 nN s^{-1}, and 400 nN s^{-1}, form the first row to the seventh row, respectively. The extension histograms were constructed and shown in **Figure 7.6**. The left column shows the force histograms for APT45 interactions, the right column the APT33 interactions. Comparing the

two aptamers, the force histograms under the same loading rate did not demonstrate any significant differe nce for the peak that represents the most probable unbinding force (F*). According to Bell's model, the affinity of APT33 to OMPs (k_{off33}) was estimated from the F* vs. LnR plot, and the (k_{off45}) was obtained by the same method (**Figure 7.7**). The two off rate values are very close, with 5.3×10^{-3} s^{-1} for k_{off33} and 7.1×10^{-3} s^{-1} for k_{off45}. The result indicate that the affinity measurement may not be able to distinguish these two aptamers for their interactions with *S.* Typhimurium.

For the extension histograms in **Figure 7.6**, the seven loading rates (R) 1 nN s^{-1}, 10 nN s^{-1}, 20 nN s^{-1}, 50 nN s^{-1}, 100 nN s^{-1}, 200 nN s^{-1}, and 400 nN s^{-1} are shown form the first row to the seventh row, respectively. The left column shows the extension histograms for APT45 interactions, the right column the APT33 interactions. Here the extension histograms show clearly different distributions between APT45 and APT33 under the same loading rate. For both aptamers, more peaks have shown under higher loading rates (**Table 7.1**). The major reason is that high loading rates will cause more non-specific interactions between the aptamers and the bacteria surface. For example, the skeleton structure of the bacteria membrane will be moved when a large external force was applied to its surface. Another possible reason is that the aptamer may have higher probability to interact with more than one OMP protein under high loading rates, because of the elasticity of the membrane surface. More investigates are required if we want to have better understa nding of the complex relationships among the extension histograms and loading rates.

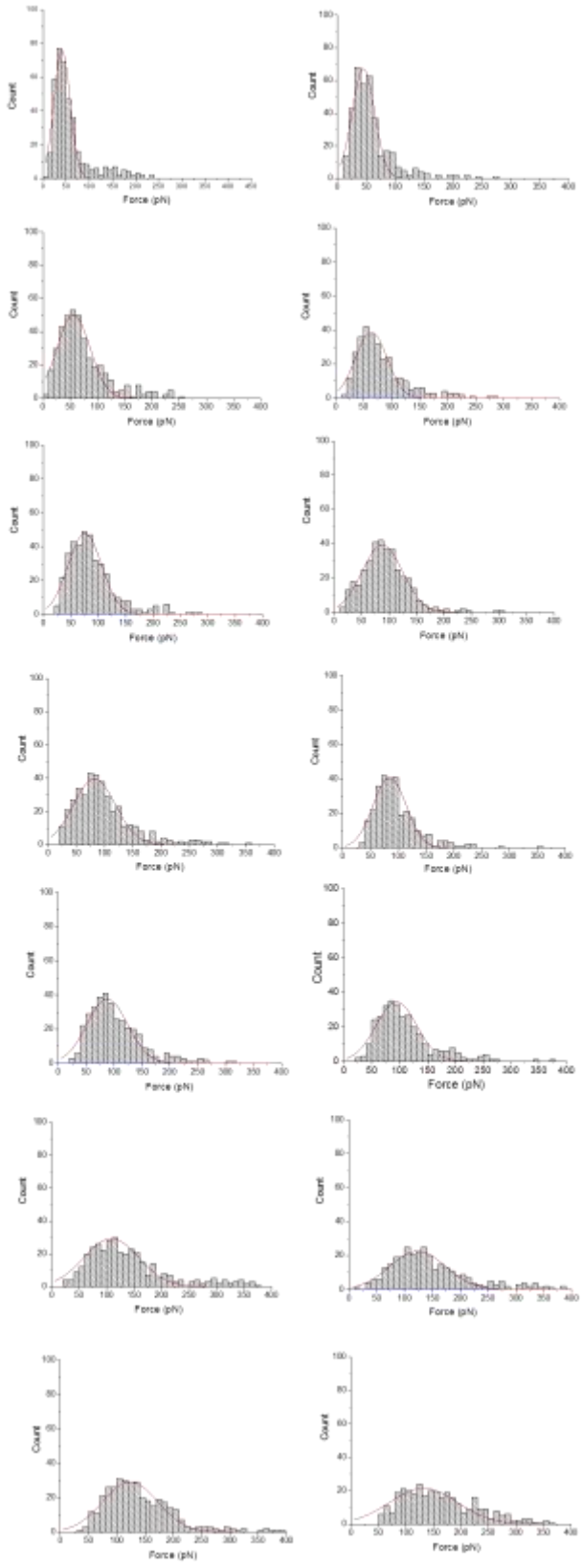

Figure 7.5 The force histograms under seven loading rates.

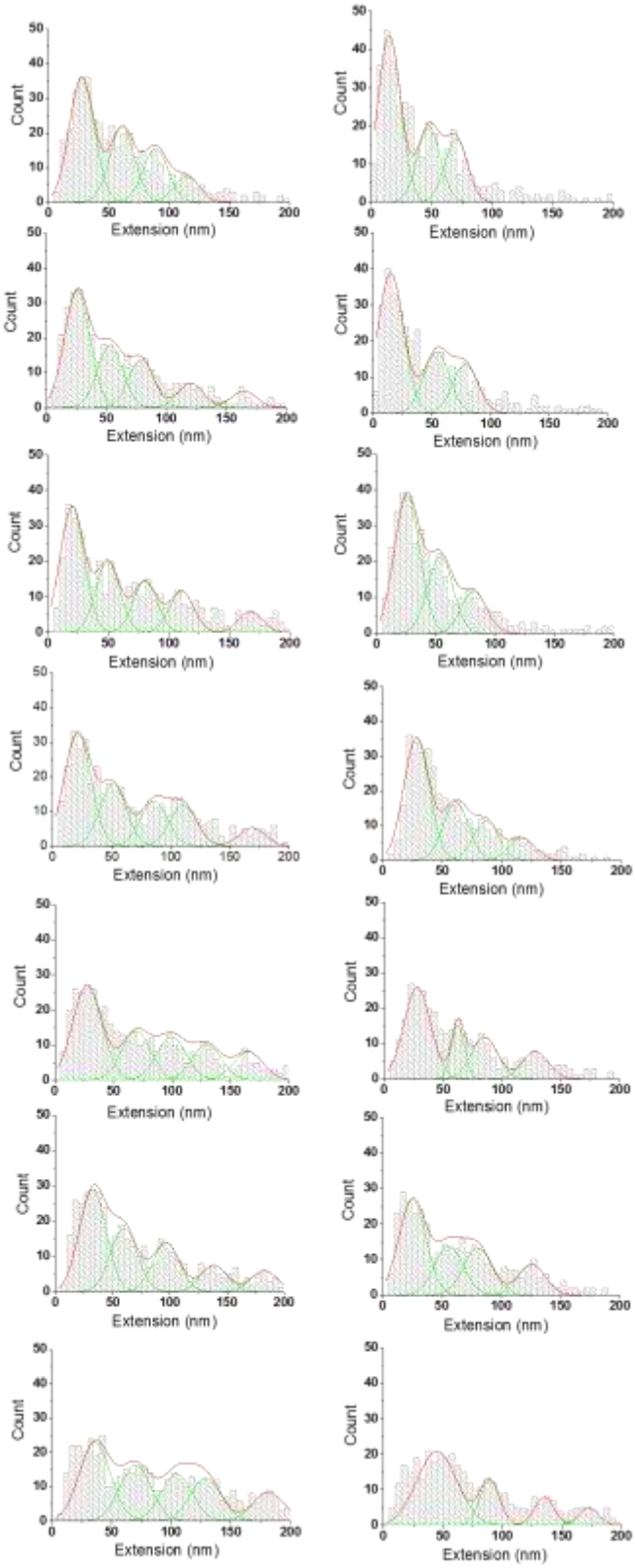

Figure 7.6 The extension histograms under seven loading rates.

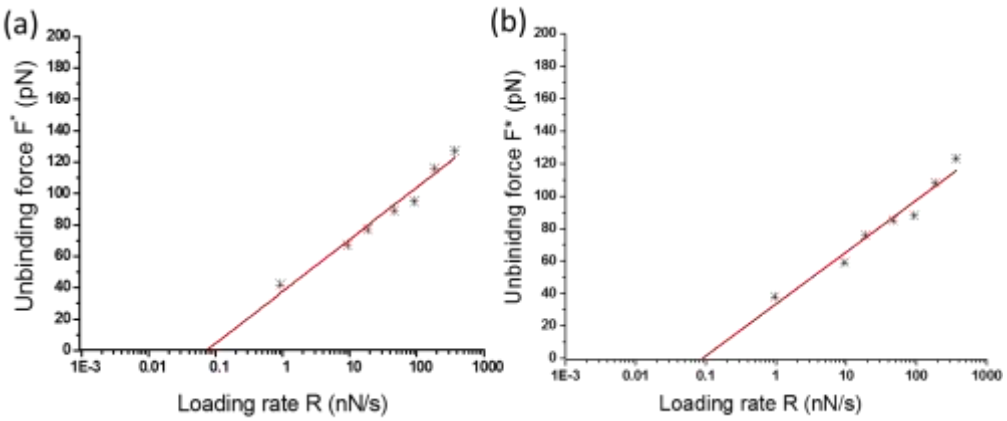

Figure 7.7 The F* vs. LnR plots for APT33 (a) and APT45 (b).

Table 7.1 The peak values in extension histograms of APT33 and APT45 under seven loading rates.

Loading rate (nN s^{-1})	Peak value of extension z (nm)	
	APT33	APT45
1.0	14, 48, 70	25, 58, 87, 114
10	14, 53, 78	25, 58, 88, 120, 164
20	24, 55, 82	26, 59, 90, 121, 167
50	25, 61, 85, 116	27, 59, 90, 121, 168
100	25, 63, 86, 128	28, 67, 97, 126, 168
200	26, 65, 89, 129	32, 67, 98, 135, 181
400	44, 89, 135, 172	35, 69, 104, 135, 182

7.3 SPR Detection

The SPR measurements have been conducted according to the similar protocol of ricin detection in Chapter 6 (**Figure 7.8**). The SPR angle shift signals in the sensorgram indicate the surface modification with the aptamer, the blocking of unoccupied carboxyl groups on CD surface with EA, the binding between aptamer and *S.* Typhimurium, and the surface regeneration with

NaOH. The SPR angle shift here was generated by the *S.* Typhimurium PBS solution with the concentration of 7.6×10^4 CFU/mL. Different concentrations can be used in future experiments to test the limit of detection (LOD) and estimate the affinity values. This preliminary experiment indicates that SPR is a promising technique for the detection of *S.* Typhimurium in solution, such as food matrix.

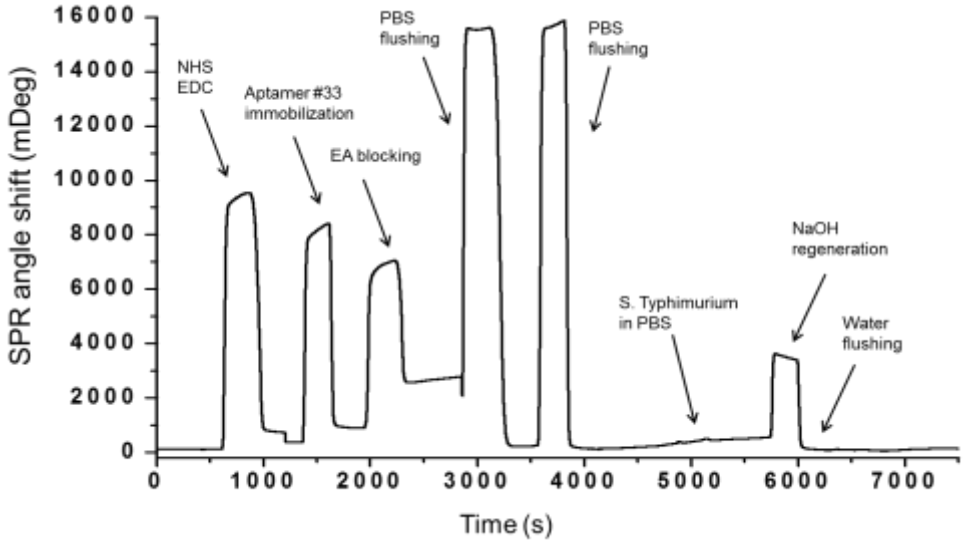

Figure 7.8 The representative SPR sensorgram for the *S.* Typhimurium detection.

7.4 Conclusion

The AFM topography and recognition images have proved that the aptamer has specific interactions to *S.* Typhimurium, However, the DFS measurements indicate that the unbinding force and extension for the interactions between aptamer and *S.* Typhimurium may have very complex relationship. Currently, the Bell's model is used to estimate the off rate value for the single-molecule interaction between aptamer and S. Typhimurium, but more detailed and specific reaction kinetic model is required to explain this *in situ* reaction mechanism. This study is very fundamental

but essential for the better understanding of the bacteria activities on biosensor surface. The detailed information about the mechanism of aptamer binding to the *S.* Typhimurium OMPs at single-molecule level will contribute to the development of biosensors for bulk solution detection of bacteria, such as SPR, as well as other related research fields such as biophysics and bionanotechnology.

The SPR measurements have shown the capability of this technique in the detection of *S.* Typhimurium samples. After the detection limit is determined, real food samples containing *Salmonella* can be tested. Moreover, SPRi technique can be used to conduct high-throughput detection of different bacteria using their corresponding aptamers. The future goal is to develop the label-free, high-throughput biosensors for the food matrix obtained from real industry samples.

CHAPTER 8

SUMMARY AND OUTLOOK

8.1 Summary

The scientific research and medicinal application of aptamer have been developing exponentially since the concepts of aptamer and in vitro selection were established. The studies on aptamers require fundamental understanding of the structure-function relationship of their secondary and tertiary structures. The special physical and chemical properties of nucleic acids have showed advantages over other biospecies such as antibody. The studies on the molecular recognition of aptamers and their reaction mechanisms at single-molecule level can reveal the properties of aptamers that may not be able to be detected by bulk solution measurements. Here in this dissertation both fundamental structure-function relationships and biosensing applications of aptamers were studied at the single-molecule level, as well as the bulk solution level.

For the single-molecule studies, the AFM and AFM-DFS techniques combine high-resolution imaging and force measurement together to study the conformations and interactions of biomolecules. Especially, the computer simulation widely used in biophysics studies, such as protein folding, nucleic acid folding, and docking simulations have been conducted in order to have deeper understanding on the binding sites and binding conformations of aptamer-ricin interactions. The combination of high resolution AFM techniques and the structural simulations have proved to be a successful method for the studies of structure-function relationships of aptamers. Moreover, these techniques used in here can be applied into other biological systems to study the protein-nucleic acid interactions. For example, the similar methods can be used to

investigate the important jet complex problems related to many essential biological processes of living cells, such as gene transcription and translation, cell surface signaling, and diseases.

We have used AFM and DFS techniques to investigate the interactions between two aptamers and their target, toxic protein ricin. These studies have achieved the goal of quantitative analysis of the thermodynamic and kinetic behaviors of aptamers during their unbinding reactions to toxic protein ricin. The biophysical parameters such as the dissociation rate constants and unbinding free energies for aptamer-ricin interactions have been estimated at single-molecule level.

Moreover, we focused on the multiple reaction pathways during the unbinding process of the aptamer-protein complex. The Markov transition models we developed have provided new information and understanding for the single-molecule studies on the aptamer binding event. We also developed SPR bulk solution detection methods and aptamer biosensors for fast detection of ricin. The difference on the measured affinity between single-molecule and bulk solution techniques provides helpful information for the future design of biosensor at nanometer scale.

Moreover, the AFM and SPR methods developed for ricin detection have been extended into the detection of bacteria serotype such as *S*. Typhimurium. The single-molecule study helps us understand the detailed reaction mechanism, and the SPR biosensor is the first step for the future development of label-free and high-throughput SPRi nanoarrays that can detect certain bacteria serotype.

8.2 Outlook

The approaches we use in this dissertation have provided new methods and biophysical models for the fundamental understanding of the structure-function relationship of aptamers. The combination of experimental techniques and simulation methods will greatly stimulate the future

biophysics research. The better understanding of the structure-function relationship will provide guides for the application of aptamers. The molecular engineering of aptamers and their integration into nanoscale devices provide promising tools for biosensors and biomedical applications. Aptamers have already been a fast developing market. The single-molecule study of aptamers will provide special contribution to the aptamer development and their applications, which will benefit both economy and human's quality of life.

REFERENCES

1. W. Colomb and S. K. Sarkar, *Phys. Life Rev.*, 2015. Epub ahead of print, doi: 10.1016/j.plrev.2015.04.019
2. G. Binning and H. Rohrer, *Appl. Phys. Lett.*, 1981, **40**, 178-180.
3. S. R. Cohen and A. Bitler, *Curr. Opin. Colloid. In.*, 2008, **13**, 316-325.
4. F. Pillet, L. Chopinet, C. Formosa and E. Daque, *Biochim. Biophys. Acta*, 2014, **1840**, 1028-1050.
5. T. Fukuma, K. Kobayashi, K. Matsushige and H. Yamada, *Appl. Phys. Lett.*, 2005, **86**, 193108-193110.
6. C. Stroch, H. Wang, R. Bash, B. Ashcroft, J. Nelson, H. Gruber, D. Lohr, S. M. Lindsay and P. Hinterdorfer, *P. Natl. Acad. Sci. USA*, 2004, **101**, 12503-12507.
7. P. Hinterdorfer and Y. F. Dufrene, *Nat. Methods*, 2006, **3**, 347-355.
8. W. T. Johnson, G. K. Kada, H. Gruber, H. Wang, F. Kienberger, S. Lindsay and P. Hinterdorfer, *NSTI-Nanotech*, 2005, **2**, 679-682.
9. B. Wang, C. Guo, G. Chen, B. Park and B. Xu, *Chem. Commun.*, 2012, **48**, 1644-1646.
10. R. Zhu, S. Howorka, J. Proll, F. Kienberger, J. Preiner, J. Hesse, A. Ebner, V. P. Pastushenko, H. J. Gruber and P. Hinterdorfer, *Nat. Nanotech.*, 2010, **5**, 788-791.
11. A. R. Bizzarri and S. Cannistraro, *Chem. Soc. Rev.*, 2010, **39**, 734-749.
12. G. I. Bell, *Science*, 1978, **200**, 618-627.
13. G. Hummer and A. Szabo, *Biophys. J.*, 2003, **85**, 5-15.

14. A. Leach, *Molecular Modelling Principles and Applications*, Pearson Education EMA, United Kingdom, 2001.

15. R. Petrenko and J. Meller, in *eLS*, John Wiley & Sons Ltd, Chichester, 2010.

16. M. Karplus, *Acc. Chem. Res*, 2002, **35**, 321-323.

17. A. R. Leach, B. K. Shoichet and C. E. Peishoff, *J. Med. Chem.*, 2006, **49**, 5851-5855.

18. K. M. Elokely and R. J. Doerksen, *J. Chem. Inf. Model.*, 2013, **53**, 1934-1945.

19. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, San Diego, 2002.

20. J. W. Ponder and D. A. Case, *Adv. Protein Chem.*, 2003, **66**, 27-85.

21. D. B. Kitchen, H. Decornez, J. R. Furr and J. Bajorath, *Nat. Rev. Drug Discov.*, 2004, **3**, 935-949.

22. T. Lengauer and M. Rarey, *Curr. Opin. Struc. Biol.*, 1996, **6**, 402-406.

23. R. B. Cyril Dominguez, Alexandre M J J Bonvin, *J. Am. Chem. Soc.*, 2003, **125**, 1731-1737.

24. D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, Unversity of California, San Francisco., 2010.

25. C. Guo, B. Wang, L. Wang and B. Xu, *Chem. Commun.*, 2012, **48**, 12222-12224.

26. S. J. de Vries, M. van Dijk and A. M. J. J. Bonvin, *Nat. Protoc.*, 2010, **5**, 883-897.

27. S. M. Lato, N. D. S. Ozerova, K. He, Z. Sergueeva, B. R. Shaw and D. H. Burke, *Nucl. Acids Res.*, 2002, **30**, 1401-1407.

28. H. Schürer, K. Stembera, D. Knoll, G. Mayer, M. Blind, H.-H. Förster, M. Famulok, P. Welzel and U. Hahn, *Bioorgan. Med. Chem.*, 2001, **9**, 2557-2563.

29. P. Çalık, O. Balcı and T. H. Özdamar, *Protein Expres. Purif.* , 2010, **69**, 21-28.

30. J. Tang, J. Xie, N. Shao and Y. Yan, *Electrophoresis*, 2006, **27**, 1303-1311.

31. X. Fang and W. Tan, *Acc. Chem. Res*, 2010, **43**, 48-57.

32. S. Li, H. Xu, H. Ding, Y. Huang, X. Cao, G. Yang, J. Li, Z. Xie, Y. Meng, X. Li, Q. Zhao, B. Shen and N. Shao, 2009, **218**, 327-336.

33. C. Tuerk and L. Gold, *Science*, 1990, **249**, 505-510.

34. A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818-822.

35. R. Kirby, E. J. Cho, B. Gehrke, T. Bayer, Y. S. Park, D. P. Neikirk, J. T. McDevitt and A. D. Ellington, *Anal. Chem.*, 2004, **76**, 4066-4075.

36. R. R. White, B. A. Sullenger and C. P. Rusconi, *J. Clin. Invest.*, 2000, **106**, 929-934.

37. M. C. Willis, B. D. Collins, T. Zhang, L. S. Green, D. P. Sebesta, C. Bell, E. Kellogg, S. C. Gill, A. Magallanez, S. Knauer, R. A. Bendele, P. S. Gill and N. Janjić, *Bioconjug. Chem.*, 1998, **9**, 573-582.

38. C. E. Tucker, L.-S. Chen, M. B. Judkins, J. A. Farmer, S. C. Gill and D. W. Drolet, *J. Chromatogr. B Biomed. Sci. Appl.* , 1999, **732**, 203-212.

39. N. O. Fischer, T. M. Tarasow and J. B.-H. Tok, *Curr. Opin. Chem. Biol.*, 2007, **11**, 316-328.

40. A. A. Nery, C. Wrenger and H. Ulrich, *J. Sep. Sci.*, 2009, **32**, 1523-1530.

41. M. Mckeague and M. C. DeRosa, *J. Nucleic Acids*, 2012, **2012**, 748913.

42. J. M. Carothers, S. C. Oestreich and J. W. Szostak, *Journal of the American Chemical Society*, 2006, **128**, 7929-7937.

43. M. U. Musheev and S. N. Krylov, *Anal. Chim. Acta*, 2006, **564**, 91-96.

44. T. Hermann and D. J. Patel, *Science*, 2000, **287**, 820-825.

45. A. A. Deniz, S. Mukhopadhyay and E. A. Lemke, *J. R. Soc. Interface*, 2008, **5**, 15-45.

46. F. Musshoff and B. Madea, *Drug Test. Anal.*, 2009, **1**, 184-191.

47. J. Audi, M. Belson, M. Patel, J. Schier and J. Osterloh, *JAMA*, 2005, **294**, 2342-2351.

48. Y. Kim and J. D. Robertus, *Protein Eng.*, 1992, **8**, 775-779.

49. X. Yan, P. Day, A. F. Monzingo, E. Schelp, J. D. Robertus, G. W. Milne and S. Wang, *Proteins: structure, function, and genetics*, 1998, **31**, 33-41

50. J. R. Hesselberth, D. Miller, J. Robertus and A. D. Ellington, *J. Biol. Chem.*, 2000, **275**, 4937-4942.

51. G. P. Anderson, R. Matney, J. L. Liu, A. Hayhurst and E. R. Goldman, *BioTechniques*, 2007, **43**, 806-811.

52. M. Fais, R. Karamanska, S. Allman, S. A. Fairhurst, P. Innocenti, A. J. Fairbanks, T. J. Donohoe, B. G. Davis, D. A. Russell and R. A. Field, *Chem. Sci.*, 2011, **2**, 1952-1959.

53. L. He, E. Lamont, B. Veeregowda, S. Sreevatsan, C. L. Haynes, F. Diez-Gonzalez and T. P. Labuza, *Chem. Sci.*, 2011, **2**, 1579-1582.

54. S. Ding, C. Gao and L. Gu, *Anal. Chem.*, 2009, **81**, 6649-6655.

55. C. Quan and J. Liu, *Chinese Journal of Analytical Chemistry*, 2010, **38**, 627-631.

56. S. Pradhan, M. Boopathi, O. Kumar, A. Baghel, P. Pandey, T. H. Mahato, B. Singh and R. Vijayaraghavan, *Biosens. Bioelectron.*, 2009, **25**, 592-598.

57. L. Lin, H. Wang, Y. Liu, H. Yan and S. Lindsay, *Biophys. J.*, 2006, **90**, 4236-4238.

58. G. Chen, N. X., B. Park, G. Boons and B. Xu, *Langmuir*, 2009, **25**, 2860-2864.

59. J. Yu, Y. Jiang, X. Ma, Y. Lin and X. Fang, *Chem. Asian J.*, 2007, **2**, 284-289.

60. L. Lin, Q. Fu, B. A. R. Williams, A. M. Azzaz, M. A. Shogren-Knaak, J. C. Chaput and S. Lindsay, *Biophys. J.*, 2009, **97**, 1804-1807.

61. C. Forster, D. Oberthuer, J. Gao, A. Eichert, F. G. Quast, C. Betzel, A. Nitsche, V. A. Erdmann and J. P. Furste, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 2009, **65**, 881-885.

62. Y. Nakamura, *Adv. Polym. Sci.*, 2012, **249**, 135-152.

63. M. Khati, *J. Clin. Pathol.*, 2010, **63**, 480-487.

64. M. Radmacher, M. Fritz, H. G. Hansma and P. K. Hansma, *Science*, 1994, **265**, 1577-1579.

65. P. Hinterdorfer, W. Baumgartner, H. J. Gruber, K. Schilcher and H. Schindler, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 3477-3481.

66. C. M. Reyes and P. A. Kollman, *J. Mol. Biol.*, 2000, **297**, 1145-1158.

67. B. Pagano, L. Martino, A. Randazzo and C. Giancola, *Biophys. J.*, 2008, **94**, 562-569.

68. H. Saneyoshi, S. Mazzini, A. Avino, G. Portella, C. Gonzalez, M. Orozco, V. E. Marquez and R. Eritja, *Nucleic Acid Res.*, 2009, 1-13.

69. P. Hinterdorfer, F. Kienberger, A. Raab, H. J. Gruber, W. Baumgartner, G. Kada, C. Riener, S. Wielert-Badt, C. Borken and H. Schindler, *Single Mol.*, 2000, **1**, 99-103.

70. G. Chen, J. Zhou, B. Park and B. Xu, *Appl. Phys. Lett.*, 2009, **95**, 043103.

71. M. Zuker, *Nucleic Acid Res.*, 2003, **31**, 3406-3415.

72. D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T.

Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews , M. G. Seetin, C. Sagui, V. Babin and P. A. Kollman, Amber 11, University of California, San Francisco, 2008.

73. T. J. Macke and D. A. Case, in *Molecular Modeling of Nucleic Acids*, eds. N. B. Leontes and J. J. SantaLucia, American Chemical Society, Washington, DC, 1998, pp. 379-393.

74. E. Rutenber, B. J. Katzin, S. Ernst, E. J. Collins, D. Mesna, M. P. Ready and J. D. Robertus, *Proteins*, 1991, **10**, 240-250.

75. E. A. Padlan, *Mol. Immunol.*, 1994, **31**, 169-217.

76. L. M. Neal, J. O'Hara, R. N. Brey III and N. J. Mantis, *Infect. Immun.*, 2010, 552-561.

77. M. v. Dijk, A. D. J. v. Dijk, V. Hsu, R. Boelens and M. J. J. Bonvin, *Nucleic Acid Res.*, 2006, **34**, 3317-3325.

78. M. Frasconi, F. Mazzei and T. Ferri, *Anal. Bioanal. Chem.*, 2010, **398**, 1545-1564.

79. A. Raab, W. B. Han, D., S. J. Smith-Gill, S. M. Lindsay, H. Schindler and P. Hinterdorfer, *Nat. Biotechnol.*, 1999, **17**, 901-905.

80. T. Macke, Case, D. A., ed., *Modeling unusual nucleic acid structures*, American Chemical Society, Washington, DC, 1998.

81. M. R. Clark, in *Chemical Immunology*, ed. L. Adorini, Arai, K., Berek, C., Capra, J. D., Schmitt-Verhulst, A-M., Waksman., B. H., Karger, Basel, 1997, pp. 88-110.

82. L. Saiz, *J. Phys.: Condens. Matter*, 2012, **24**, 193102.

83. J. H. J. Hoeijmakers, *N. Engl. J. Med.*, 2009, **361**, 1475-1485.

84. H. Yu, X. Liu, K. Neupane, A. N. Gupta, A. M. Brigley, A. Solanki, I. Sosova and M. T. Woodside, *P. Natl. Acad. Sci. USA*, 2012, **109**, 5283-5288.

85. D. H. Paik and T. T. Perkins, *Angew. Chem. Int. Ed.*, 2012, **51**, 1811-1815.

86. M. Held and F. Noe, *Eur. J. Cell Biol.*, 2012, **91**, 357-364.

87. N. C. Harris, Y. Song and C. Kiang, *Phys. Rev. Lett.*, 2007, **99**, 068101.

88. C. Jarzynski, *Phys. Rev. Lett.*, 1997, **78**, 2690-2693.

89. S. Guo, C. Ray, A. Kirkpatrick, N. Lad and B. B. Akhremitchev, *Biophys. J.*, 2008, **95**, 3964-3976.

90. T. Erdmann, S. Pierrat, P. Nassoy and U. S. Schwarz, *Europhys. Lett.*, 2008, **81**, 48001.

91. D. Leckband, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 1-26.

92. A. M. Berezhkovskii, *J. Chem. Phys.*, 2011, **134**, 074114.

93. U. Seifert, *Phys. Rev. Lett.*, 2000, **84**, 2750-2753.

94. M. Manosas, I. Junier and F. Ritort, *Phys. Rev. E*, 2008, **78**, 061925.

95. Z. Yu, V. Gaerig, Y. Cui, H. Kang, V. Gokhale, Y. Zhao, L. H. Hurley and H. Mao, *J. Am. Chem. Soc.*, 2012, **134**, 5157-5164.

96. Y. Zhao and G. D. Stormo, *Nat. Rev. Genet.*, 2010, **11**, 751-760.

97. E. Evans and K. Ritchie, *Biophys. J.*, 1997, **72**, 1541-1555.

98. M. Zuker, *Nucleic Acid Res.*, 2003, **31**, 3406-3415.

99. D. T. Gillespie, *Physica A*, 1992, **188**, 404-425.

100. M. R. Hoare, *Nature*, 1970, **226**, 599-603.

101. N. T. Schmandt and R. F. Galan, *Phys. Rev. Lett.*, 2012, **109**, 118101.

102. D. A. Smith, W. Steffen, R. M. Simmons and J. Sleep, *Biophys. J.*, 2001, **81**, 2795-2816.

103. J. Shuai and P. Jung, *Biophys. J.*, 2002, **83**, 87-97.

104. S. C. Kou, X. S. Xie and J. S. Liu, *J. Roy. Stat. Soc. C (Appl. Statist.)*, 2005, **54**, 469-506.

105. J. Stigler, F. Ziegler, A. Gieseke, J. C. Gebhardt and M. Rief, *Science*, 2011, **334**, 512-516.

106. T. E. Turner, S. Schnell and K. Burrage, *Comput. Biol. Chem.*, 2004, **28**, 165-178.

107. E. Evans, *Biophys. Chem.*, 1999, **82**, 83-97.

108. R. Merkel, P. Nassoy, A. Leung, K. Ritchie and E. Evans, *Nature*, 1999, **397**, 50-53.

109. B. Tinland, A. Pluen, J. Sturm and G. Weill, *Macromolecules*, 1997, **30**, 5763-5765.

110. H. Grubmuller, B. Heymann and P. Tavan, *Science*, 1996, **271**, 997-999.

111. K. Neupane, D. B. Ritchie, H. Yu, D. A. N. Foster, F. Wang and M. T. Woodside, *Phys. Rev. Lett.*, 2012, **109**, 068102.

112. E. M. Puchner and H. E. Gaub, *Annu. Rev. Biophys.*, 2012, **41**, 497-518.

113. K. A. Melzak, S. Moreno-Flores, A. E. Lopez and J. L. Toca-Herrera, *Soft Matter*, 2011, **7**, 332-342.

114. C.-Y. Chen, T.-Y. Chien, C.-K. Lin, C.-W. Lin, Y.-Z. Weng and D. T.-H. Chang, *PLoS ONE*, 2012, **7**, e30446.

115. Z. Xie, S. Hu, J. Qian, S. Blackshaw and H. Zhu, *Cell. Mol. Life Sci.*, 2011, **68**, 1657-1668.

116. B. Dey, S. Thukral, S. Krishnan, M. Chakrobarty, S. Gupta, C. Manghani and V. Rani, *Mol. Cell. Biochem.*, 2012, **365**, 279-299.

117. A. D. Keefe, S. Pai and A. Ellington, *Nat. Rev. Drug Discov.*, 2010, **9**, 537-550.

118. L. A. Liu and P. Bradley, *Curr. Opin. Struc. Biol.*, 2012, **22**, 397-405.

119. C. Hyeon and D. Thirumalai, *J. Chem. Phys.*, 2012, **137**, 055103.

120. A. V. Dobrynin, J.-M. Y. Carrillo and M. Rubinstein, *Macromolecules*, 2010, **43**, 9181-9190.

121. S. Kumar and G. Mishra, *Soft Matter*, 2011, **7**, 4595-4605.

122. S. Kasas and G. Dietler, *Eur. J. Physiol.*, 2008, **456**, 13-27.

123. N. E. Kurland, Z. Drira and V. K. Yadavalli, *Micron*, 2012, **43**, 116-128.

124. S. Kasas, G. Longo and G. Dietler, *J. Phys. D: Appl. Phys.*, 2013, **46**, 133001.

125. J. F. Marko and E. D. Siggia, *Macromolecules*, 1995, **28**, 8759-8770.

126. R. G. Haverkamp, A. T. Marshall and M. A. K. Williams, *Phys. Rev. E*, 2007, **75**, 021907.

127. S. Cui, Y. Yu and Z. Lin, *Polymer*, 2009, **50**, 930-935.

128. M. Manghia, N. Destainville and J. Palmerib, *Eur. Phys. J. E.*, 2012, **35**, 110.

129. C. Ray, J. R. Brown and B. B. Akhremitchev, *J. Phys. Chem. B*, 2007, **111**, 1963-1974.

130. S. Guo, N. Lad, C. Ray and B. B. Akhremitchev, *Biophys. J.*, 2009, **96**, 3412-3422.

131. M. Ritzefeld, V. Walhorn, D. Anselmetti and N. Sewald, *Amino Acids*, 2013, **44**, 1457-1475.

132. F. Kienberger, V. P. Pastushenko, G. Kada, H. J. Gruber, C. Riener, H. Schindler and P. Hinterdorfer, *Single Mol.*, 2000, **1**, 123-128.

133. B. Wang and B. Xu, *Phys. Rev. E*, 2014, **89**, 022720.

134. Y.-H. Cai and H. Huang, *Amino Acids*, 2012, **43**, 1141-1146.

135. E. A. Dethoff, J. Chugh, A. M. Mustoe and H. M. Al-Hashimi, *Nature*, 2012, **482**, 322-330.

136. Y. Chen and R. M. Corn, *J. Am. Chem. Soc.*, 2013, **135**, 2072-2075.

137. A. Hayat, S. Andreescu and J.-L. Marty, *Biosens. Bioelectron.*, 2013, **45**, 168-173.

138. Y. Lyubchenko, L. Shlyakhtenko, R. Harrington, P. Oden and S. Lindsay, *P. Natl. Acad. Sci. USA*, 1993, **90**, 2137-2140.

139. B. Liedberg, C. Nylander and I. Lundstrom, *Sensor. Actuator.*, 1983, **4**, 299-304.

140. J. Homola, *Chem. Rev.*, 2008, **108**, 462-493.

141. M. J. Linman, A. Abbas and Q. Cheng, *Analyst*, 2010, **135**, 2759-2767.

142. J.-F. Masson, T. M. Battaglia, J. Cramer, S. Beaudoin, M. Sierks and K. S. Booksh, *Anal. Bioanal. Chem.*, 2006, **386**, 1951-1959.

143. S. Ray, G. Mehta and S. Srivastava, *Proteomics*, 2010, **10**, 731-748.

144. Y.-H. Lao, K. Peck and L.-C. Chen, *Anal. Chem.*, 2009, **81**, 1741-1754.

145. A. B. Iliuk , L. Hu and W. A. Tao, *Anal. Chem.*, 2011, **83**, 4440-4452.

146. B. Basnar, R. Elnathan and I. Willner, *Anal. Chem.*, 2006, **78**, 3638-3642.

147. A. Sassolas, L. J. Blum and B. D. Leca-Bouvier, *Biosens. Bioelectron.*, 2011, **26**, 3725-3736.

148. N. de-los-Santos-Alvarez, M. J. Lobo-Castanon, A. J. Miranda-Ordieres and P. Tunon-Blanco, *TrAC Trend Anal. Chem.*, 2008, **27**, 437-446.

149. M. Piliarik, L. Parova and J. Homola, *Biosens. Bioelectron.*, 2008, **24**, 1399-1404.

150. G. Lautner, Z. Balogh, V. Bardóczy, T. Mészáros and R. E. Gyurcsányi, *Analyst*, 2010, **135**, 918-926.

151. S. Worbs, K. Kohler, D. Pauly, M.-A. Avondet, M. Schaer, M. B. Dorner and B. G. Dorner, *Toxins*, 2011, **3**, 1332-1372.

152. M. Friedman and R. Rasooly, *Toxins*, 2013, **5**, 743-775.

153. S. Centi, S. Tombelli, M. Minunni and M. Mascini, *Anal. Chem.*, 2007, **79**, 1466-1473.

154. C. Polonschii, S. David, S. Tombelli, M. Mascini and M. Gheorghiu, *Talanta*, 2010, **80**, 2157-2164.

155. J. Liu, M. A. Eddings, A. R. Miles, R. Bukasov, B. K. Gale and J. S. Shumaker-Parry, *Anal. Chem.*, 2009, **81**, 4296-4301.

156. L. R. Brewer and P. R. Bianco, *Nat. Methods*, 2008, **5**, 517-525.

157. B. Wang, C. Guo, M. Zhang, B. Park and B. Xu, *J. Phys. Chem. B*, 2012, **116**, 5316-5322.

158. S. Sang, Y. Wang, Q. Feng, Y. Wei, J. Ji and W. Zhang, *Crit. Rev. Biotechnol.*, 2015, **Early Online.**

159. Y. Du, B. Li and E. Wang, *Acc. Chem. Res*, 2013, **46**, 203-213.

160. R. Joshi, H. Janagama, H. P. Dwivedi, T. M. A. S. Kumar, L.-A. Jaykus, J. Schefers and S. Sreevatsan, *Mol. Cell. Probe.*, 2009, **23**, 20-28.