UNIVERSITY OF CALIFORNIA, IRVINE

Microsecond Dynamics of Enzymes: A Single-Molecule Study Using Carbon Nanotube Transistors

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Maxim V. Akhterov

Dissertation Committee:
Professor Philip Collins, Chair
Professor Ilya Krivorotov
Professor Peter Burke

2015
# TABLE OF CONTENTS

LIST OF FIGURES iv

LIST OF TABLES vi

ACKNOWLEDGMENTS vii

CURRICULUM VITAE ix

ABSTRACT OF THE DISSERTATION xi

1 Introduction 1

2 Experimental Design 6
   2.1 Measurement Stability 7
   2.2 Reducing Parasitic Capacitance 9
   2.3 Low-Noise Design 11
      2.3.1 Grounding 14
      2.3.2 Low-Noise Voltage Sources 15
   2.4 Noise in SWCNT-FET Devices 16
   2.5 Apparatus Overview 19
      2.5.1 Flow Cell 20
      2.5.2 Microfluidics 22
      2.5.3 Fluid Delivery 23
      2.5.4 Complete Apparatus 24
   2.6 Capacitive Discharge Problem 24
   2.7 Future Flow Cell Improvements 26

3 Device Fabrication 27
   3.1 SWCNT Synthesis and Photolithography 27
   3.2 Device Passivation 28
   3.3 Lysozyme Conjugation to SWCNT Devices 29
   3.4 DNA Polymerase Conjugation to SWCNT Devices 30
   3.5 Protein Kinase A Conjugation to SWCNT Devices 31
   3.6 Wire Bonding and PDMS alignment 32
## Table of Contents

### 4 Data Analysis

4.1 Analysis Interface ........................................ 36  
4.2 Flattening .................................................. 37  
4.3 Two-level Analysis ......................................... 38  
4.4 Transition Time Analysis .................................. 42  
4.5 Forward-Backward Non-Linear Filtering .................. 44  

### 5 High-Resolution Study of Lysozyme Motions

5.1 T4 Lysozyme ............................................... 50  
5.2 Electrical Measurements ................................... 52  
5.3 Results and Discussion .................................... 53  

### 6 Study of DNA Polymerase and Protein Kinase A

6.1 DNA Polymerase I .......................................... 65  
  6.1.1 Electrical Measurements ............................... 67  
  6.1.2 Results and Discussion ............................... 68  
6.2 Protein Kinase A ........................................... 74  
  6.2.1 Electrical Measurements ............................... 75  
  6.2.2 Results and Discussion ............................... 75  
6.3 Comparing Transition Durations of Three Enzymes ...... 77  

### 7 Conclusion .................................................. 81  

Bibliography ..................................................... 84
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Molecular Motion Timescale</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>SWCNT-FET signals from three molecules</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Two-State Energy Landscape</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Effects of Input Capacitance in Op-Amp Circuits</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Input Capacitance Effect in DLPCA-200</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>Power Spectrum Example</td>
<td>12</td>
</tr>
<tr>
<td>2.4</td>
<td>Aliasing</td>
<td>13</td>
</tr>
<tr>
<td>2.5</td>
<td>Apparatus Grounding</td>
<td>15</td>
</tr>
<tr>
<td>2.6</td>
<td>Two-level Switching in Oxidized SWCNT</td>
<td>17</td>
</tr>
<tr>
<td>2.7</td>
<td>Estimating Signal-to-Noise Ratio</td>
<td>19</td>
</tr>
<tr>
<td>2.8</td>
<td>Flow Cell Diagram</td>
<td>20</td>
</tr>
<tr>
<td>2.9</td>
<td>Flow Cell Body</td>
<td>21</td>
</tr>
<tr>
<td>2.10</td>
<td>Flow Cell Lid</td>
<td>21</td>
</tr>
<tr>
<td>2.11</td>
<td>DB-25 Connector Pinout</td>
<td>22</td>
</tr>
<tr>
<td>2.12</td>
<td>Microfluidics Mold</td>
<td>22</td>
</tr>
<tr>
<td>2.13</td>
<td>Fluid Delivery Schematic</td>
<td>23</td>
</tr>
<tr>
<td>2.14</td>
<td>Flow Cell Apparatus</td>
<td>24</td>
</tr>
<tr>
<td>2.15</td>
<td>Flow Cell Noise Profile</td>
<td>25</td>
</tr>
<tr>
<td>2.16</td>
<td>Capacitive Discharge</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Biofunctionalization</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>Atomic Force Micrographs of SWCNT-FET Biofunctionalized with Lysozyme</td>
<td>30</td>
</tr>
<tr>
<td>3.3</td>
<td>Atomic Force Micrographs of SWCNT-FET Biofunctionalized with KF</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Atomic Force Micrographs of SWCNT-FET Biofunctionalized with PKA</td>
<td>32</td>
</tr>
<tr>
<td>3.5</td>
<td>PDMS Alignment</td>
<td>33</td>
</tr>
<tr>
<td>3.6</td>
<td>Wire bonded SWCNT-FET Device</td>
<td>34</td>
</tr>
<tr>
<td>4.1</td>
<td>TDMS Data Explorer Interface</td>
<td>36</td>
</tr>
<tr>
<td>4.2</td>
<td>Analysis Chain Interface</td>
<td>37</td>
</tr>
<tr>
<td>4.3</td>
<td>Flattening Data</td>
<td>38</td>
</tr>
<tr>
<td>4.4</td>
<td>Thresholding</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>Two-Level Analysis</td>
<td>41</td>
</tr>
<tr>
<td>4.6</td>
<td>Transition Analysis</td>
<td>43</td>
</tr>
<tr>
<td>4.7</td>
<td>Transition Analysis Artifacts</td>
<td>44</td>
</tr>
</tbody>
</table>
4.8  Edge Blurring .................................................. 45
4.9  Forward-Backward Predictors ................................. 46
4.10  Adaptive Filtering ............................................. 49

5.1  Lysozyme Structure and Function ............................ 51
5.2  \( I(t) \) in buffer ............................................. 52
5.3  SWCNT-FET with Lysozyme .................................. 54
5.4  Continuous Closing Examples ................................ 55
5.5  Continuous Opening Examples ................................ 55
5.6  Probability Distributions of Hinge-Bending Durations .... 56
5.7  Examples of Closing with Pauses ............................. 58
5.8  Examples of Opening with Pauses ............................. 58
5.9  Probability Distributions of Hinge-Bending Durations with Pauses 59
5.10  Free-Energy Landscape of T4 lysozyme ....................... 62

6.1  DNA Polymerase Mechanism .................................. 66
6.2  Structure of Klenow Fragment ................................. 67
6.3  KF Signal ..................................................... 69
6.4  KF Excursion Examples ....................................... 70
6.5  KF High and Low Distributions ............................... 71
6.6  KF motifs .................................................... 71
6.7  Distributions of Three Event Types ......................... 73
6.8  PKA Signal .................................................. 75
6.9  Reproducing Previous PKA Results ........................ 76
6.10  Fast PKA Dynamics ......................................... 77
6.11  Distribution of Transition Durations of Three Enzymes .... 78
# LIST OF TABLES

2.1 Capacitance Contributions .................................................. 11

5.1 Timing and Energetics of T4 Lysozyme Closing and Opening .............. 64
5.2 Lysozyme SWCNT-FET Devices that Generated Data ........................ 64

6.1 Motif Discovery Logic ......................................................... 72
6.2 Timing of KF ................................................................. 73
6.3 Transition Durations of Three Enzymes ................................... 79
6.4 KF and PKA SWCNT-FET Devices that Generated Data ................. 80
ACKNOWLEDGMENTS

First, I would like to thank Professor Philip Collins for his guidance and mentorship. I feel very fortunate to have had an opportunity to work with such an outstanding scientist and mentor. His regular feedback and encouragement has been essential in helping me develop the necessary scientific thinking and skill set that made me well-equipped for my next career step.

I would also like to thank Professor Gregory Weiss for advising me on the biochemical facets of this work. His insights into the nature of complex biochemical processes helped me better understand some counterintuitive findings. Furthermore, his positive attitude and excitement about my research was inspiring.

I would like to offer my special thanks to the members of the Collins group. Dr. Steve Hunt, Dr. Israel Perez, Dr. Danny Wan, and Dr. Tatyana Sheps helped me get started in the lab and shared with me their expertise and their skills. A big thanks to Dr. Tolga Gul for growing carbon nanotubes and helping with a DNA Polymerase measurement; Dr. Brad Corso for the cleanroom work, helping with LabVIEW, and workout tips; Dr. Deng Pan for his support and the recent cleanroom work; Dr. Patrick Sims for sharing the highs and lows of single-molecule research and for his feedback on my writing; Dr. Elliot Fuller for being a good friend, supportive colleague, and teaching me how to cook a steak; Dr. Tetyana Ignatova for useful discussions and support; and Arith Rajapaske for growing carbon nanotubes on the recent wafers. I am particularly grateful for Dr. Yongki Choi’s assistance. His work established the foundation for my own work in this dissertation, and his mentorship helped me on both a professional and a personal level.

This research would not be possible without the help of our collaborators in the Gregory Weiss Laboratory. I would like to thank Dr. Tivoli Olsen for the synthesis and purification of T4 Lysozyme; Kaitlin Pugliese for bringing up DNA polymerase and being a great colleague; Mariam Iftikhar for the synthesis of protein kinase A, as well as her hard work, perseverance, and support.

At University of California, Irvine, I was also fortunate to have met some great people outside the lab. Dr. Goran Matijasevic exposed me to the world of business and entrepreneurship. Through his mentorship I learned first-hand about technology transfer, start-ups and venture capital. Professor Eli Simon and Bri McWhorter helped me to dramatically improve my public speaking skills, which greatly improved the quality of my scientific talks.

Most importantly, I want to thank my family and close friends. My parents have been incredibly supportive throughout the years and I owe a lot to them for deciding to emigrate to the US. My brother and his family have always been a source of encouragement. Finally, I want to thank Anna Gelman for making me a better person.
I would like to acknowledge the following funding sources that made this work possible: UCI LifeChips IGERT program, the National Science Foundation, and the National Institutes of Health.
CURRICULUM VITAE

Maxim V. Akhterov

EDUCATION

Doctor of Philosophy in Physics
University of California, Irvine

Bachelor of Science in Applied Physics
University of California, Santa Cruz

EXPERIENCE

Algorithm Engineer
Scanadu Inc.

Graduate Research Assistant
University of California, Irvine

Undergraduate Research Assistant
University of California, Santa Cruz

Research Assistant
Air Force Research Laboratory / Wright State University

Research Assistant
Naval Postgraduate School

Research Assistant
University of California, Berkeley

HONORS & AWARDS

National Science Foundation’s IGERT Fellowship
University of California, Irvine

Chancellors Award for Best Presentation
University of California, Irvine

Second Place in the Business Plan Competition
The Paul Merage School of Business, University of California, Irvine
TEACHING & OUTREACH

Volunteer Scientist at annual Ask-A-Scientist Night 2011 - 2014
Rancho San Joaquin Middle School, Irvine

Judge at Irvine Unified School District Science Fair 2013
Woodbridge High School, Irvine

Teaching Assistant 2010 - 2011
University of California, Irvine

PUBLICATION


CONFERENCE PRESENTATIONS

American Physical Society March Meeting (2014) Denver, CO
Biophysics Society Annual Meeting (2015) Baltimore, MD
Molecular motions of proteins and their flexibility determine conformational states required for enzyme catalysis, signal transduction, and protein-protein interactions. However, the mechanisms for protein transitions between conformational states are often poorly understood, especially in the milli- to microsecond ranges where conventional optical techniques and computational modeling are most limited. The goal of this work was to use single-walled carbon nanotube field-effect transistors (SWCNT-FET) as single-molecule biosensors to investigate microsecond dynamics of three enzymes. Low-noise electrical transport measurements were performed in a home-built electrochemical flow cell. To ensure that the output signal was free from external noise across a 200 kHz measurement bandwidth, parasitic capacitance was minimized by tightly integrating electrical components and constraining the liquid within a microfluidic channel. After attaching a single enzyme molecule to the SWCNT sidewall, dynamic changes of conductance through a SWCNT-FET reported conformational motions of an enzyme with a 1.6 μs resolution. This technique was used to study microsecond dynamics of T4 Lysozyme, Klenow Fragment of DNA polymerase I (KF), and protein kinase A (PKA).
Lysozyme closing and opening took on average 37 µs. The distribution of transition durations was independent of the lysozyme state: either catalytic or nonproductive. The observed symmetry in enzyme opening and closing suggests that substrate translocation occurs while lysozyme is closed. For KF and PKA, the microsecond resolution revealed that the transition duration was 3-5 µs for both enzymes, about ten times shorter than lysozyme. Additionally, the high-bandwidth recording of KF resolved fast non-catalytic closures. When KF was processing poly(dT)$_{42}$ template KF stayed 2-3 times longer in a closed conformation than when it was replicating poly(dA)$_{42}$. However, the open events of KF were 2-5 times shorter when it was processing poly(dT)$_{42}$ compared to poly(dA)$_{42}$. The results demonstrate SWCNT-FET as a sensitive technique for studying conformational dynamics of biological molecules in the microsecond range.
Chapter 1

Introduction

Regulatory and catalytic functions of bio-molecules, such as enzymes, require dynamic changes of protein structure. The interplay between molecular motions and functions has been extensively studied and reveals a wide range of motion timescales [1]. Elucidating the mechanisms of various biological processes requires a broad set of experimental techniques (Figure 1.1). For example, nuclear magnetic resonance (NMR) provides insights into localized sub-nanosecond dynamics of proteins, such as loops motion and bond vibrations [2]. Large-scale structural rearrangements that take seconds to complete have been observed using high-speed atomic force microscopy (HS-AFM) [3]. Fluorescent methods, such as Forster resonance energy transfer (FRET), allow direct observation of molecular movements on the timescale of milliseconds, but photo-bleaching of fluorophores limits the technique’s ability to record shorter events [4, 5, 6]. Therefore, conventional experimental techniques fall short in observing catalytically important collective domain motions in the microsecond range. Also, conformational changes over this time scale are too long for current all-atom molecular-dynamics simulation methods [1]. For these reasons, mechanisms underlying conformational changes in the microsecond range remain poorly understood.
In principle, electronic techniques using nanoscale devices have the bandwidth to resolve fast molecular events and overcome the limitations of fluorescent-based methods. For example, field-effect transistors (FET) made of single-walled carbon nanotubes (SWCNT) can operate at GHz-THz frequencies corresponding to a response time on picosecond timescales [7, 8, 9]. Over the past 15 years, SWCNT-FET devices have been actively researched as a platform for chemical and biological sensors [10, 11, 12]. With diameters of 1-2 nm, and lengths ranging from 50 nm up to 1 cm, SWCNTs are quasi one-dimensional "wires" that are directly comparable to the size of single biomolecules. Since SWCNTs have low charge-carrier density and all carbon atoms are exposed on the surface, SWCNTs are extremely sensitive to its local environment [13, 14]. The inelastic mean free path of a pristine SWCNT at room
temperature approaches 1 µm [15], so resistance of a SWCNT can be a sensitive indicator of environmental interactions. Kong et al. were the first to demonstrate this principle by building a molecular SWCNT sensor for the detection of NO₂ and NH₃ gas [16]. Later, Besteman et al. built the first SWCNT-FET biosensor using a single SWCNT coated with the redox enzyme glucose oxidase (GOx). The conductance of a GOx-coated SWCNT increased in the presence of glucose, the substrate of GOx, suggesting that SWCNTs could be used to monitor enzyme kinetics at a single-molecule level [17].

First experiments that resolved single-molecule kinetics with SWCNT-FETs used point-functionalized carbon nanotubes to covalently bind molecules at the scattering site [18, 19, 20]. For example, Goldsmith et al. electrochemically introduced a single carboxylate group into the SWCNT sidewall and used it to observe attachments and detachments of a carbodiimide hydrochloride (EDC) [19]. Sorgenfrei et al. used the same functionalization technique to attach a single-stranded probe DNA and monitor DNA-hybridization kinetics in the presence of a complimentary DNA target [20].

Recently, Choi et al. accomplished a single-molecule transduction using a pyrene-maleimide linker to tether a single lysozyme molecule and study its enzymatic motion [21]. This non-covalent attachment relied on the π-π interaction between pyrene and the nanotube surface and provided a reliable method for functionalizing carbon nanotubes [22]. Unlike point-functionalization, this conjugation strategy had a high device fabrication yield and did not introduce defects. The underlying signal transduction was driven by an electrostatic mechanism in which the motions of charged residues near the SWCNT sidewall changed local electrostatic potential and modulated the source-drain current \( I(t) \) [23]. The approach was generalized to study molecular motions of two other enzymes - DNA polymerase I, and protein kinase A and Figure 1.2 shows examples of electronic signals generated by each molecule [21, 24, 25].
Each excursion $\Delta I(t)$ above or below the baseline corresponded to a conformational motion of a protein [23]. For example, in Figure 1.2a the baseline current corresponded to an open conformation of an enzyme, while excursions to 4 nA represented enzyme transitions to a closed state [20]. This dynamics can be schematically represented by a two-state energy landscape. In the absence of a substrate enzyme stays in its open conformation Figure 1.3a. In this dissertation, the duration enzyme spends in an open state is called $\tau_{\text{open}}$. Substrate binding changes the energy landscape making the closed state more accessible (Figure 1.3b). The transition time from an open to a closed state is referred to as $\tau_{\text{closing}}$. The time enzyme spends in a closed state is $\tau_{\text{closed}}$. Eventually, the substrate unbinding reverts energy
landscape and the open state becomes more energetically favorable. The time it takes the enzyme to reopen is referred to as $\tau_{\text{opening}}$.

In principle, every excursion in Figure 1.2 carried the information about $\tau_{\text{closing}}$ and $\tau_{\text{opening}}$. However, within the experimental resolution of 100 $\mu$s transitions between open and closed states appeared instantaneous, indicating that the higher bandwidth was needed to resolve the lysozyme motion between the two conformations.

The work in this dissertation was focused on designing a high bandwidth system with a resolution of 2 $\mu$s and using it to study microsecond dynamics of three enzymes. First, Chapter 2 describes the apparatus for a low-level, high-bandwidth electronic measurements. Then, Chapter 3 talks about the fabrication of SWCNT-FET devices and conjugation of a single molecule to the SWCNT sidewalls. Next, Chapter 4 describes how the data were analyzed. Chapter 5 shows how the high bandwidth apparatus was used to study closing and opening motions of T4 lysozyme. Then, Chapter 6 describes a study of microsecond dynamics of DNA polymerase I and protein kinase A.

![Diagram of energy landscape with open and closed states and time to opening and closing](image)

**Figure 1.3**
Chapter 2

Experimental Design

The goal of this work was to use SWCNT-FET to study single-molecule kinetics. The conformational motions of a molecule tethered to a carbon nanotube sidewall modulated a dynamic conductance response in the underlying SWCNT-FET. To successfully resolve these motions the nanotube conductance needed to be recorded with a temporal resolution comparable to the timescale of the underlying kinetic processes. As it was discussed in Chapter 1, the microsecond-to-millisecond dynamics of enzymes plays an important role during catalysis. This implies that the conductance of SWCNT-FET needs to be measured with a kHz-MHz bandwidth.

Typically, the conductance of MOSFET devices is characterized using a three-terminal electrical transport measurement. A small bias voltage is applied across a source-drain channel and a gate electrode modulates the channel conductance. However, a different approach is needed when a biological molecule serves as a functioning element of the circuit. To preserve biochemical functions of enzymes electrical measurements must be performed in a physiological environment, such as a saline solution. This requires that the transport measurement is performed in a three-electrode electrochemical cell. Counter, reference, and working
electrodes should be used together to control the chemical potential of the electrolyte. Performing high-bandwidth electronic measurements in liquid present unique challenges. This chapter describes how to overcome these challenges and build an apparatus for a low-noise electronic measurements with 200 kHz bandwidth.

First, section §2.1 talks about the stability of the electronic measurements in the presence of a parasitic capacitance, such as coupling between the electrolyte and the electronic components and section §2.2 outlines the ways to reduce it. Sections §2.3 and §2.4 discusses the instrumental noise as well as sources of noise specific to SWCNT-FET devices. Section §2.5 provides an overview of the flow cell apparatus. Section §2.6 discusses how to prevent capacitive discharge from damaging the devices. Finally, Section §2.7 proposes a design improvement that would increase the device throughput of the flow cell.

### 2.1 Measurement Stability

Performing the high-bandwidth electronic measurement in a conducting electrolyte presents a number of challenges. To prevent a short-circuit the wiring needs to be isolated from the liquid. However, even without the direct contact the capacitive coupling between the liquid and the interconnects adds to the total parasitic capacitance of the circuit and can cause stability problems if the signals require amplification. This is a well-known problem in the stability theory of the operational amplifiers.

Consider an inverting amplifier circuit that is typically used for a current-to-voltage conversion (Figure 2.2a). The transfer function of this circuit is

\[
\frac{V_{\text{OUT}}}{V_{\text{IN}}} = \beta = \frac{R_1}{(R_{\text{OUT}} + R_2)(1 + sR_1C_{\text{IN}}) + R_1}
\]  

(2.1)
and has a pole located at

$$f_p = \frac{R_1 + R_2 + R_{OUT}}{2\pi R_1 C_{IN}(R_2 + R_{OUT})}$$

If the break frequency, $f_p$, is below the open-loop unity-gain frequency, the system becomes unstable. This causes the rate of closure between the slopes of the open-loop gain (-20 dB/decade at most frequencies) and the noise gain (defined as $\frac{1}{\beta}$) to be around -40 dB/decade. This situation is unstable and will cause oscillations. Alternatively, one can think of a phase shift of the loop-gain, $A\beta$, where $A$ is an open-loop gain and $\beta$ is a feedback factor. If while traversing a feedback loop, a phase of $A\beta$ exhibits a shift of -180° and an amplifier gain is equal or greater than 1, the signal will reinforce itself causing the circuit to oscillate.

Since the location of the break frequency is inversely proportional to the input capacitance $C_{IN}$, the smaller values of the input capacitance give a more stable system. This effect was modeled using an LTSPICE simulation. The value of $C_{IN}$ was varied from 1 uF to 1 pF while other values were held constant ($R_1 = R_2 = 100$ kΩ, Open-loop gain= $10^6$, GBW = 100 kHz). Figure 2.1b shows gain (solid lines) and phase (dash lines) for three values of $C_{IN}$. A gain peaking, -40 dB/decade roll-off, and a sharp phase shift of -180° clearly indicate instability at 400 Hz and 13 kHz for the input capacitance of 1 uF and 1 nF respectively. As the capacitance decreases to 1 pF the system stabilizes: the gain is no longer peaking, the roll-off returns to -20 dB/decade, and a phase margin is 115° at 100 kHz. Reducing the input capacitance is the most straightforward way to prevent instability. Alternatively, it is possible to compensate for $C_{IN}$ by adding a capacitor in parallel with $R_2$. However, this would increase the overall RC constant of the circuit and might filter out important high-frequency information from the SWCNT-FET biosensor.

The simulation in Figure 2.1b shows the response of a simplified amplifier. In the work described here DLPCA-200 (Femto) was used to amplify the source-drain current. It had
The maximum amount of the input capacitance this amplifier could tolerate was determined experimentally by looking at the power spectrum of the amplified signal at different values of the input capacitance. In the power spectrum, the op-amp instability manifests itself as a hump or a peak. Figure 2.2 shows the output power spectra of DLPCA-200 operating at gain $10^8$ V/A for the three values of the input capacitance. To obtain a stable response across 200 kHz the input capacitance should not exceed 120 pF. The flow cell was designed with this "budget" in mind.

### 2.2 Reducing Parasitic Capacitance

The total input capacitance includes the capacitances from the flow-cell wiring $C_W$, the internal components of the amplifier $C_I$, and an electrolyte coupling to the electrodes $C_L$. A
careful choice of the electrical components and their close integration dramatically reduced $C_W$.

The SWCNT-FET devices were plugged into a 44 pin socket. Typically, such sockets are mounted onto a printed circuit board, however, stray capacitance between the traces and the ground plane of the PCB would add up to the input capacitance. Instead, a socket was mounted onto a PTFE pedestal, and its pins were individually soldered to a short coaxial cable connected to the amplifier input.

When a SWCNT-FET is exposed to an electrolyte solution its ions capacitively couple to the electrodes. This capacitance $C_L$ can be modeled by the parallel-plate capacitor, $C_L = \epsilon_0 \epsilon_r A/d$, where $\epsilon_0$ is the permittivity of free space, $A$ is the contact area of the liquid, $\epsilon_r$ and $d$ are the relative permittivity and the thickness of the dielectric respectively. First, to reduce $C_L$, $d$ was increased. Poly(methyl methacrylate) (PMMA) is a commonly used positive resist for an electron-beam lithography. The e-beam writing exposed only a portion of the SWCNT sidewall while keeping device electrodes passivated. Coating the chip with a 200 nm PMMA film reduced the capacitance by 130 pF. To further reduce the capacitive
coupling the contact area $A$ was minimized by confining the liquid within a 100 $\mu$m wide microfluidic channel. The details of microfluidic fabrication are outlined in §2.5.

Combining a close integration of electrical components, passivating chip with PMMA, and reducing the liquid contact resulted in the total input capacitance of 110 pF - well within the allowed capacitance ”budget”. An estimated breakdown of the capacitance contributions is shown in the table below.

<table>
<thead>
<tr>
<th>$C_W$</th>
<th>$C_L$</th>
<th>$C_I$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>76 pF</td>
<td>29 pF</td>
<td>5 pF</td>
<td>110 pF</td>
</tr>
</tbody>
</table>

Table 2.1: Capacitance contributions from wiring $C_W$, liquid $C_L$, and amplifier input $C_I$

### 2.3 Low-Noise Design

High-bandwidth transport measurements of nano-scale circuits could be easily distorted by noise. Some noise is intrinsic to a specific electronic device, and some is the result of the measurement electronics. In any case, it is useful to have a general approach to analyzing noise that can be used to determine the source of noise and reduce its interference with a signal.

For a time-varying signal, like a recording of a source-drain current in SWCNT-FET versus time, the spectral analysis provides information about the frequency components of the signal and their magnitude. Also known as a power spectral density (PSD) it is commonly used to measure and characterize noise profile of the system. Specifically, a PSD describes how the power of the signal is distributed across the frequency spectrum. An example of a PSD is shown in Figure 2.3. The y-axis is the power per unit frequency $S_I = (\Delta I)^2/\Delta f$, where $\Delta I$ is the current fluctuation and $\Delta f$ is the bandwidth. Alternatively, the contribution of each frequency in a PSD can be expressed as a magnitude in decibels $\text{Mag(dB)} = 10 \log_{10} S_I$. 

11
The x-axis is a range of frequencies up to the Nyquist frequency $f_N = f_S/2$, where $f_S$ is the sampling frequency. Finally, the frequency resolution $df$ is determined by the measurement time ($T$) or the total number of samples acquired ($N$) $df = 1/T = f_S/N$.

![Power Spectral Density of the source-drain current in the SWCNT-FET.](image)

Figure 2.3: Power Spectral Density of the source-drain current in the SWCNT-FET.

The goal of the noise analysis is to characterize the magnitude and the location of peaks on a PSD graph. The noise that lies outside of the signal frequencies can be easily filtered out using traditional linear filtering. However, if the noise contaminates the signal, it needs to be reduced by eliminating the noise source or shielding the signal from it. Unfortunately, the position of a peak on a PSD graph might not always represent the real frequency of the noise source. If the signal is digitized at a sampling rate such that the Nyquist frequency is less than the signal’s original frequency, the signal will be aliased back creating a false peak at

$$f_A = |f_s \times n - f_o|$$

(2.3)

where $n$ is the closest integer multiple of the sampling rate $f_s$ to the signal being aliased $f_o$.

For example, if a signal that contains 100 Hz noise is measured at a sampling rate of 70 Hz,
\( n = 1 \) and on a PSD graph the peak will appear at \( f_A = |70 \times 1 - 100| = 30 \text{ Hz} \), instead of 100 Hz (Figure 2.4a). If the sampling frequency changes to 150 Hz, the aliased peak will move to 50 Hz (Figure 2.4b). The peak will continue to shift around and misrepresent the original signal until \( f_S \) exceeds 200 Hz. This problem is a premise of the Nyquist-Shannon theorem that states that a signal must be sampled at a rate that is at least twice that of its highest frequency component. For sampling frequencies above 200 Hz the peak on a PSD graph will stay at 100 Hz representing the original signal (Figure 2.4c,d).

![Figure 2.4: Power spectra of a signal containing 100 Hz noise sampled at different rates. If sampling frequency, \( f_s \), is below 200 Hz, the peak appears at a wrong frequency.](image)

Practically, it means that when acquiring a signal at two different sampling frequencies\(^1\) causes the peaks in the frequency spectra to shift, there is an aliasing error. The sampling frequency needs to be continually increased until the peaks in the PSD spectra do not change,

\(^1\)It is important that the sampling frequencies are not integer multiples of each other (e.g. 30 Hz and 90 Hz). Otherwise, the aliasing peaks will appear at the same frequencies.
and the aliased peaks can be explained. This approach was extensively used in this work to analyze and reduce noise due to grounding and interference issues explained below.

2.3.1 Grounding

Grounding is one of the major techniques to reduce unwanted noise pickup. There are two basic objectives involved in designing a grounding system. The first is to minimize the noise voltage created by current from multiple circuits flowing through a common ground impedance. The second is to avoid creating ground loops that are susceptible to an electromagnetic radiation and differences in potential [26].

The components of the flow cell apparatus were grounded to the equipment rack that serves as a ground bus. The ground bus implies a tree configuration (Figure 2.5). Within each subsystem of the apparatus, the individual components were single-point grounded. Then, the ground points of subsystems were connected to the ground bus by an insulated conductor. For example in the cell, the cable shields, the conducting surfaces, and a metal enclosure were connected to a floating ground of the current amplifier (DLPCA-200). When it was connected to the data acquisition electronics in a ground referenced single-ended configuration (GRSE), the entire flow-cell became grounded to the rest of the apparatus in the electronics rack.

Within the rack, instruments that contain transformers and fans (computer, linear power supply, digital multimeter) were placed at the bottom since they radiate a lot of electromagnetic noise. More sensitive components of the apparatus (breakout box, buffer box, BNC connector block) were placed at the top of the rack.
2.3.2 Low-Noise Voltage Sources

While physical separation helps reduce noise due to EMI, it will not help if a noisy voltage source is used to power sensitive components of the apparatus. The biosensing experiments required two low-noise voltage sources - one for the source-drain bias and one for controlling the liquid potential. The source-drain voltage $V_{sd}$ was supplied using a low-noise bias voltage source built into the DLPCA-200 amplifier. $V_{sd}$ was set to 100 mV by a potentiometer and was held constant for the duration of the experiment.

The liquid potential $V_{lg}$ also required a low noise voltage source, but it must be computer-controlled to accomplish ramping and biasing. Also, the desired potential $V_{lg}$ was sensed on the reference electrode but sourced using a lower impedance counter electrode in a conventional, three-electrode electrochemical setup. This combination of tasks could be performed with a Keithley 2400 source-measurement unit. However, high-frequency noise found in Keithley 2400 output coupled to SWCNT-FET circuit corrupting the signal. Instead, the same the source-sense functionality was achieved using one analog input (AI) and one analog
output (AO) of the data acquisition hardware. An AO channel with a 1 kHz analog low-pass filter provided a low-noise DC output to the counter electrode. A closed-loop controller (PID) implemented in LabVIEW read the potential of the reference electrode once a second and changed the counter electrode to keep the reference at the desired value. The feedback parameters $P, I, D$ were tuned experimentally in a PBS buffer with value $P = 0.5$, $I = 0.1$, and $D = 0$ giving the shortest settling time without overshoot.

Additionally, a high-impedance voltage follower provided impedance matching between the acquisition hardware and the circuit controlling the electrolyte potential. The follower was powered by a linear power supply (Agilent, E3630A). Unlike a switching-mode power supply, linear regulators have a very little ripple, output noise, and radiated EMI.

### 2.4 Noise in SWCNT-FET Devices

So far, the discussion was focused on the instrumental noise. However, the SWCNT-FET devices themselves have a number of possible noise sources such as defects in the SWCNT sidewall, thermal diffusion of metal atoms at the SWCNT-metal interface, adsorbates on the surface of the metal or on the SWCNT, or charge traps in the SiO$_2$. It is important to understand which of these mechanisms are the most prevalent and how they could affect signals in the single-molecule experiments.

The defects in SWCNT lattice could form randomly during the synthesis stage, but could also be purposefully introduced using multiple non-chemical techniques [27]. In either case, the defects could generate significant electronic noise and corrupt the data from a SWCNT-FET biosensor. Figure 2.6 shows an example of two-level switching for a SWCNT damaged by a partial sidewall oxidation and measured in a PBS buffer. Depending on the bias conditions, the amplitude of such switching could have a signal-to-noise ratio comparable to the signals
generated by some biomolecules. Despite having limited control over the defect formation during the synthesis, it was previously found that with 2 µm separation between the source and the drain electrodes most SWCNT-FET devices did not have defects [28].

Figure 2.6: A representative series of stochastic two-level noise in a SWCNT transistor that had been covalently damaged by sidewall oxidation.

The quality of the SWCNT-metal interface depends on a type of metal and the fabrication method. In this work, most of the devices were made with platinum and palladium electrodes with chromium used as a sticking layer for a better metal adhesion. Some wafers were annealed at high temperatures to improve metal contact to the SWCNT and reduce noise at the interface. Additionally, all devices used for this work were passivated with PMMA to protect contacts from the liquid electrolyte.

Another major noise source in SWCNT-FET devices is the contaminants and charge traps on the silicon dioxide. A random trapping and detrapping of charges in the nanotube/oxide interface has been suggested as a primary origin of the $1/f$ noise in the SWCNT devices [29]. When gated, an electric field around the nanotube could cause an injection of electrons from the nanotube into and out of the surface contaminants or an oxide charge traps in SiO$_2$ causing the source-drain current to fluctuate. The field emission is less pronounced in a liquid electrolyte since all the charge traps that are further than a Debye distance away from the SWCNT sidewall are screened by ions in the electrical double layer. The biosensing work described here was carried out in the buffer solutions (pH=7.0-8.0) that have a Debye
length of 1 nm. This implies that only contaminants and charge traps located within 1 nm from the SWCNT could contribute to the electronic noise.

Since a variety of surface effects can generate noise of different magnitudes, it is helpful to define a single metric that would be used to evaluate if a given SWCNT-FET device could be used to monitor microsecond kinetics of proteins tethered to a SWCNT sidewall. The previous work found that the motion of charged residues on the protein surface modulate current fluctuations through the SWTN-FET \[21, 23\]. Since the magnitude of these fluctuations depends on the amount of charge at the residues, the transduction mechanism relies on the additional localized electrostatic gating of the SWCNT-FET \[23\]. The amount of this expected "effective" gating provides a straightforward way to calculate a most-likely signal-to-noise ratio (SNR) for a given SWCNT device prior to a biosensing experiment. For example, consider a SWCNT-FET device that responses to a liquid-gating potential \(V_{lg}\) as shown in Figure 2.7a. If this device is used for a biosensing experiment, one would conduct a measurement at \(V_{lg}=0\) V, since the signal of SWCNT-FET biosensors is maximized when devices are operating where transconductance \(dI/dV_{lg}\) is the largest \[30\]. If a protein molecule modulates fluctuations equivalent to -100 mV of liquid gating, the expected fluctuations would have a magnitude of approximately 24 nA. However, when the device is measured in a buffer solution, the noise magnitude is 20-30 nA relative to the baseline at 50 nA (Figure 2.7b). The resulting SNR is less than 1 and such device would make a poor biosensor. Alternatively, consider a different device with the \(IV_{lg}\) response shown in Figure 2.7c. When the device is measured at \(V_{lg}=-0.1\) V the noise magnitude is bellow 15 nA (Figure 2.7d), resulting in SNR of 2:1. Unlike the other device, this one would make it possible to separate the biological signal from a SWCNT noise and, therefore, can be used for a biosensing experiment.
2.5 Apparatus Overview

To perform an electrical measurement in liquid, a flow cell was built to allow delivery of various electrolytes to a region of the SWCNT-FET circuit with a tethered biomolecule. Figure 2.8 shows the major electronic components of the apparatus. The source-drain current was amplified by a low-noise trans-impedance amplifier (DLPCA-200) and was digitized by a data acquisition board (DAQ). The electrolyte potential was controlled using a combination of a platinum counter electrode (CE), Ag/AgCl reference electrode (RE1) and a platinum pseudo-reference electrode (RE2). The electrodes were controlled by a DAQ and a digital multimeter (Keithley 2000). The voltage follower (buffer) prevented the low-impedance DAQ from loading the high-impedance circuit that controlled the counter and the reference electrodes.
2.5.1 Flow Cell

The flow cell body was machined out of chemically inert PTFE plastic and accommodated two amplifiers (named "1" and "2"). A SWCNT-FET device with a PDMS gasket plugged into a 44-pin receptacle supported by a pedestal mounted onto a positioning stage. The amplifier inputs were connected to the pins on the chip socket via two BNC connectors. The inner conductors carrying the source-drain current signals were wired to pins 5 and 9 for amplifier "1" and "2" respectively. The grounded BNC shield of amplifier "1" was wired to pins 4 and 6. Similarly, the BNC shield of amplifier "2" was wired to pins 8 and 10.

The gasket containing the microfluidic channel was pressure sealed to the chip surface by a lid attached to a hinge and fixed by a clamping screw. The positioning stage helped align the inlet and outlet holes in the gasket directly under the liquid ports in the lid (Figure 2.9).

The lid of the flow cell housed the tubing and the counter and the reference electrodes. The aluminum frame held a transparent window made of PVC and a set of pin receptacles. When the lid was closed the liquid ports in the PVC window contacted the inlet and outlet ports in the PDMS gasket and allowed the liquid to flow through the microfluidic channel. The window was designed to have no dead volume that might lead to formation of bubbles. Two
pairs of inlet and exhaust tubing make it easy to purge air or liquid out of the cell. To control the liquid potential, the counter electrode platinum wire extends about one inch into the inlet tubing to maintain good contact with the liquid. An additional counter electrode on the other side helped maintain the same potential on both sides of the microfluidic channel. The platinum wires were soldered directly to the pin receptacles. All tubing had a grounded copper shield to minimize pickup noise (Figure 2.10).

The wires controlling the back gate, counter-electrode and the two reference electrodes were soldered to a 25-pin D-sub connector mounted into the side of the flow cell. To minimize
cross-talk between the wires in the cable, they were spatially separated, and all the unused pins were grounded. The pin-out diagram is shown in (Figure 2.11).

![DB-25 connector pinout](image)

**Figure 2.11: DB-25 connector pinout**

2.5.2 Microfluidics

Reducing the contact area of the chip exposed to the electrolyte minimized capacitive coupling at the amplifier input and increased the circuit stability. The polydimethylsiloxane (PDMS) gasket was designed to confine liquid within a 100 µm channel. The gasket was fabricated using standard soft lithography techniques. First, a master mold was fabricated using SU-8 photoresist spun on a silicon wafer to produce 100 µm thick film. It was then exposed to UV light through a photomask to transfer the pattern of a microfluidic channel. In addition to the channel, the pattern included two rectangular features on both sides to provide clearance for the wire bonds (Figure 2.12).

![Microfluidics mold](image)

**Figure 2.12: Microfluidics mold**

The PDMS gasket was prepared by mixing 15 g of polymer base and 1.5 g of curing agent and then pouring the mixture over the master mold in a Petri dish. The PDMS was degassed in a
vacuum desiccator and cured overnight. After peeling it off from the master mold, individual PDMS gaskets were cut for each device. Each gasket was at least 2-3 mm thick so that it could be pressure sealed in the flow cell. The inlet and outlet holes were punched on both ends of the microfluidic channel using a biopsy punch.

### 2.5.3 Fluid Delivery

In a typical biosensing experiment, a SWCNT-FET device was exposed to multiple solutions. The fluid delivery system was designed to easily purge solutions through the flow cell without introducing bubbles that can lead to poor electrostatic coupling between the platinum electrodes. Figure 2.13 shows the schematic of the fluid delivery system. The main inlet (Inlet 1) was connected to a computer-controlled perfusion system (AutoMate Scientific Inc.) that pushed liquid out of 2.5 mL cryo-tube reservoirs at a constant flow rate. At the beginning of the experiment, Inlet 1 and Exhaust 1 were used to fill the right U-shaped branch of the flow cell with the buffer solution. Similarly, the lines Inlet 2 and Exhaust 2 were used to fill the left U-shaped branch of the cell. To push the liquid through the microfluidic channel valves on Exhaust 1 and Inlet 2 were closed letting liquid flow from Inlet 1 into the channel and out through Exhaust 2. To introduce a new solution, an old solution was first purged out of both branches.

![Fluid Delivery Schematic](image)

*Figure 2.13: Fluid Delivery Schematic*
2.5.4 Complete Apparatus

The complete flow cell apparatus is shown in Figure 2.14a,b. Implementing the aforementioned design principles reduced the instrumental noise and enabled a high-bandwidth electronic measurement. Figure 2.15 shows the noise profile of the flow cell before (red) and after (green) modifications. Reducing parasitic capacitance by constraining the liquid within a microfluidic channel got rid of an amplifier instability around 10 kHz. Implementing a proper grounding scheme eliminated ground loops that can be seen as noise peaks at harmonics of 60 Hz. Finally, using low-noise voltage sources and shielding sensitive electronic components minimized the high-frequency noise pickup.

![Flow cell apparatus](image)

Figure 2.14: Flow cell apparatus

2.6 Capacitive Discharge Problem

Nanoelectronic circuits could be easily damaged by electrostatic discharge[31]. This problem was observed multiple times during the design of the flow cell. When the liquid was purged through the microfluidic channel the SWCNT-FET devices became open circuits. The open circuit was confirmed by measuring IVg in the cell and on the probe station. The open circuits were observed on every SWCNT device that was exposed to liquid independent of
Figure 2.15: Flow cell noise profile: initial (red) and after implementing low-noise techniques (green).

whether or not its pads were electrically grounded. The devices did not have any damage visible under the AFM.

Most likely, the problem was caused by a capacitive voltage spike between the nanotube and one of the platinum electrodes used to control liquid potential. However, the transient voltage was too fast to be measured directly. Figure 2.16 shows an example of a SWCNT-FET device open-circuiting upon being exposed to the deionized water.

Figure 2.16: Capacitive discharge problem: The SWCNT-FET becomes open circuit upon exposure to liquid at 91.50 s.

To avoid a capacitive discharge, the potential of the liquid and the platinum electrodes needs to be controlled at all times. To accomplish that the counter electrode platinum wire was
extended about one inch into the inlet tubing (Inlet 1) to maintain good contact with the incoming liquid. Additionally, an extra counter electrode was installed on the other side of the microfluidic channel to maintain the same potential in both reservoirs (Figure 2.10). To ensure that all three platinum electrodes were coupled to each other, prior to each measurement the flow cell was filled with a DI water using a "dummy" device (carrier + chip + PDMS gasket). Then, the counter electrodes were ramped between -0.6 and 0.4 V. At the end of the ramp all platinum electrodes were at 0 V and "dummy" device was replaced with an actual SWCNT-FET device.

### 2.7 Future Flow Cell Improvements

One of the major limitations of the current flow cell design is the device throughput. In this work, many devices were lost because wire bonds came off when the PDMS clearance pocket collapsed as the gasket was sealed in the flow cell. Additionally, some measurements failed due to liquid leaking underneath the PDMS gasket causing a short circuit. Frequently, that was caused by the inlet and outlet ports being too close to the edge of the chip. Combined these factors made about half of the prepared devices unusable. The device throughput can be significantly increased if the chips are made larger, and the electrodes are made long enough so that the wire bonding pads are located outside of the PDMS gasket. While the larger electrodes would result in fewer SWCNT-FET devices on each wafer, the overall yield of the flow cell would be increased.
Chapter 3

Device Fabrication

In this project, T4 Lysozyme, the Klenow Fragment of DNA Polymerase I, and protein kinase A were individually conjugated to SWCNT devices and measured in the flow cell. The design, expression, and purification of all three enzymes was developed and executed by collaborators in Prof. G.A. Weiss’ lab at UC Irvine. The protocols can be found in the earlier publications [21, 24, 25]. This chapter describes SWCNT-FET fabrication, protein conjugation to SWCNT, and PDMS alignment.

3.1 SWCNT Synthesis and Photolithography

Single-walled carbon nanotubes were grown on 4” Si wafers by chemical vapor deposition (CVD) using a monodisperse nanoparticle Fe$_{30}$Mo$_{84}$ as the catalytic seed. Nanoparticles were diluted 1:1000 in ethanol and then spin-coated onto a clean wafer surface at 150 rpm. SWCNT growth proceeded by a three-step CVD process. The catalyst-coated wafer was oxidized in air at 700 °C, then reduced at 940 °C (520 sccm H$_2$ in 3000 sccm Ar), and finally exposed to a carbon source at 940 °C (1000 sccm CH$_4$ + 520 sccm H$_2$ in 3000 sccm Ar).
After CVD synthesis, SWCNT randomly coated the entire wafer at an areal density of approximately 0.01 $\mu m^{-2}$. SWCNTs were characterized by noncontact atomic force microscopy (AFM, Pacific Nanotechnology Nano-R) and by scanning electron microscopy (SEM, JEOL 6060 at 1 kV). The CVD growth protocol yielded SWCNTs with a diameter range of 1.1-1.6 nm.

After CVD synthesis, Pd electrodes (30 nm) with a Cr sticking layer (1 nm) were deposited on top of the SWCNTs using wafer-scale optical lithography. Source and drain electrodes were separated by 2 $\mu m$. An undercut bilayer resist (S1808 on top of LOR-A1, MicroChem) was used for a clean liftoff. The p++ Si wafer, separated from the SWCNT by a 500 nm thermal oxide, defined a back-gate electrode. Individual devices were electrically probed to identify semiconducting SWCNT devices for further use. AFM was used to confirm that only one SWCNT formed each source-drain connection.

### 3.2 Device Passivation

After initial characterization, each device underwent electrode passivation followed by an electron beam lithography. Devices were coated with an electron beam resist (A3 PMMA, MicroChem) and then patterned with an electron beam to expose the middle portion of SWCNT channel and the source and drain electrode pads for wire bonding to the ceramic chip carrier. This patterning was designed to protect the majority of the surface, especially source and drain electrodes, from contacting the test solutions. After e-beam lithography, devices were re-imaged by AFM to confirm that a 1.0 $\mu m$ portion of the SWCNT channel was properly exposed and that it remained free of particulates.
3.3 Lysozyme Conjugation to SWCNT Devices

Individual T4 lysozyme molecules were conjugated to SWCNT devices using the bifunctional linker molecule N-(1-pyrenyl)maleimide (Sigma-Aldrich). The pyrene functionality adhered strongly to SWCNT sidewalls via $\pi-\pi$ stacking\cite{22}. The maleimide group formed a stable thiol-ether bond with any free thiol of a surface cysteine in the target protein. Expression of lysozyme variants containing only a single surface cysteine enabled a reproducible attachment orientation (Figure 3.1). Devices were soaked in a 1 mM solution of pyrene-maleimide in ethanol for 30 min without agitation. Afterward, excess pyrene-maleimide was removed by rinsing the devices for 5 minutes with 0.1% Tween-20 (Acros Organics) in ethanol and 5 minutes in de-ionized water.

![Diagram of pyrenemaleimide conjugation to SWCNT](image)

Figure 3.1: Functionalization of SWCNT sidewall with pyrenemaleimide

Next, devices were soaked in a lysozyme solution (54 $\mu$M in PBS) for 120 min without agitation. The devices were then rinsed with wash buffer (5 mM KCl, 10 mM Na$_2$HPO$_4$, 0.05% Tween-20, pH 7) to remove non-specifically adsorbed lysozyme. Figure 3.2b shows examples of SWCNT-FETs imaged by AFM, in which the single molecule attachment is clearly resolved. Typically, lysozyme bioconjugation yield was about 1 attachment per micron of a SWCNT sidewall.
3.4 DNA Polymerase Conjugation to SWCNT Devices

Individual molecules of KF were conjugated to SWCNT devices using the same attachment chemistry used to attach a lysozyme. Devices were soaked in a 1 mM solution of pyrene-maleimide in ethanol for 30 min without agitation. Afterward, excess pyrene-maleimide was removed by rinsing the devices with 0.1% Tween-20 in ethanol for 30 seconds. Then the devices were rinsed under flowing de-ionized water for 1 min.

Next, a 20 µL drop of 0.02-0.2 mg/mL KF in a buffer a solution (10 mM Tris, 50 mM NaCl, 10 mM DTT, 10 mM MgCl₂, pH 8) was put on the device for 30-90 min without agitation. To prevent evaporation during the attachment the device was kept in a covered cell. Then, the device was rinsed with 0.1% Tween-20 in wash buffer (10 mM Tris, 50 mM NaCl, 10 mM MgCl₂, 100 µM TCEP, pH 8) and kept in a standard buffer (10 mM Tris, 50 mM NaCl, 10 mM DTT, 10 mM MgCl₂, pH 8).

The steps of the conjugation protocol were designed to yield a single attachment to a SWCNT device. However, variability in KF synthesis and purification required a series of attachment tests for each new KF batch to determine the concentration and the incubation time needed.
for an acceptable yield. The results of a typical attachment test are shown in Figure 3.3. The AFM image on the left shows a pristine nanotube before the conjugation. The right image taken after the conjugation shows three attachments, each approximately 5-8 nm high. This height is consistent with the physical structure of a KF molecule. Non-specific attachments on the SiO$_2$ surface had a similar height suggesting that the rinsing did not remove all the adsorbed KF from the surface. While it is possible to achieve cleaner surfaces by using harsher rinsing protocols, these were avoided because KF was non-covalently attached to the SWCNT. The attachment density shown in Figure 3.3 of 3 KF molecules per 3 µm of CNT sidewall ensured that devices with 1 µm of exposed CNT sidewall would have on average only one attached molecule.

![Figure 3.3: AFM topography images of a SWCNT device before (a) and after (b) incubation with 0.2 mg/mL KF for 80 min.](image)

3.5 Protein Kinase A Conjugation to SWCNT Devices

Individual molecules of PKA were conjugated to SWCNT devices using pyrene-maleimide linker molecules. The devices were soaked in a 1 mM solution of pyrene-maleimide in ethanol for 30 min without agitation. Then, the excess pyrene-maleimide was removed by rinsing.
the devices with 0.1% Tween-20 in ethanol for 30 seconds. Then the devices were rinsed under flowing de-ionized water for 1 min.

Next, a 20 µL solution of 5 µM PKA in a buffer solution (10 mM Tris, 50 mM NaCl, 10 mM DTT, 10 mM MgCl₂, pH 8) was dropped on the device for 5 minutes without agitation for PKA attachment. Finally, the devices were rinsed with 0.1% Tween-20 in wash buffer (10 mM Tris, 50 mM NaCl, 10 mM MgCl₂, 100 µM TCEP, pH 8) and kept in a standard buffer (10 mM Tris, 50 mM NaCl, 10 mM DTT, 10 mM MgCl₂, pH 8). Figure 3.4 shows an AFM image of a SWCNT device functionalized with a single PKA molecule. The attachment density of 1 PKA molecule per 1 µm of CNT sidewall was obtained using a fresh PKA solution that was purified within 48 hours from the measurements. Also, note that Figure 3.4 is a result of an attachment test. The actual experiment was performed on a different device that went through the same attachment protocol.

![AFM image of Protein Kinase A](image)

Figure 3.4: AFM image of Protein Kinase A

### 3.6 Wire Bonding and PDMS Alignment

After a protein conjugation, a SWCNT device was glued to a ceramic chip carrier using a conductive silver paint. Then, the source and drain pads of the SWCNT device were wire bonded to the pads on a chip carrier. Then, the gasket was aligned over the SWCNT
chip by placing a 100 µm wide microfluidic channel directly over the exposed region of the SWCNT-FET. Specifically, the wedge of the wire bonder held a glass slide with a PDMS gasket. By maneuvering the stage, the device was aligned so that the SWCNT-FET region was positioned under the microfluidic channel (Figure 3.5a). The PDMS gasket was placed by lowering the wedge using a z-lever (Figure 3.5b). The adhesion between the PDMS and SiO$_2$ surface held the gasket in place as the wedge was lifted up (Figure 3.5c).

![Figure 3.5: PDMS alignment process](image)

To enable a leak-free seal the PDMS gasket was cut large enough to cover an entire chip and the inlet and outlet ports were at least 1-2 mm away from the edge of the chip. To prevent wire bonds from coming off the electrode pads, the PDMS gasket should lie flat on the silicon chip. An example of a well-aligned gasket is shown in Figure 3.6.
Figure 3.6: a, PDMS gasket on a device wire bonded to a ceramic chip carrier. b, Proper alignment of the gasket on a chip can be seen under the microscope.
Chapter 4

Data Analysis

In a typical biosensing measurement performed in a flow cell, the source-drain current through a SWCNT-FET device was acquired at 600 kHz or 2 MHz for 10-20 minutes. The goal of the analysis was to extract short segments of data that were characteristic of molecular events, such as the conformational motion of enzymes. Considering the stochastic nature of these events, their short duration (few to tens of microseconds), the $1/f$ noise of the SWCNT-FET devices, and the size of the data set, the analysis was a challenging computational task. The analysis methods were designed with these challenges in mind. Section §4.1 outlines the analysis interface and shows a typical sequence of processing steps. First, the data was flattened to remove low-frequency baseline drift (§4.2). Then, two-level analysis separated the data into two states corresponding to different molecular conformations (§4.3). Finally, a transition analysis algorithm calculated the time between the two states to provide insight into the dynamics of the conformational changes (§4.4). The limited signal-to-noise ratio can corrupt calculations of transition duration. Typically, the signal-to-noise can be improved by filtering. However, traditional filtering methods distort fast transition events. Section §4.5 describes a filtering method that does not introduce artifacts. The algorithm was implemented in LabVIEW.
4.1 Analysis Interface

TDMS Data Explorer is the data processing software developed in LabVIEW by Prof. Philip Collins and Dr. Steve Hunt to playback and analyze data from SWCNT devices (Figure 4.1). Since high-speed conduction data sets are too big to be processed all at once, the software loads data in short segments called frames. Each frame is plotted and analyzed sequentially. To avoid running out of memory, it is recommended to keep the frame size at 1 second or less for data acquired at 100 kHz - 2 MHz.

Figure 4.1: TDMS Data Explorer
Each stage of analysis (Flattening, Power Spectrum, Switching Analysis, etc.) is an individual LabVIEW VI with input and output channels. The analysis sequence is configured through a graphic interface by selecting functions, indicating which channels to process, and choosing the output. A typical analysis chain used in this work is plotted as a flowchart in Figure 4.2.

![Analysis Configuration Interface](image)

Figure 4.2: Analysis Configuration Interface

### 4.2 Flattening

SWCNT devices exhibit $1/f$ noise that is most significant at low frequencies. This low frequency noise appears as a time-dependent, drifting baseline of the DC current. To eliminate the baseline drift, the signal was flattened in three steps. First, the data was down-sampled to 100 points. Second, the down-sampled signal was fit to a third-order polynomial. The resulting fit traced slow baseline fluctuations. Finally, the signal was de-trended by subtract-
ing the polynomial fit from the raw data. An example of a signal before and after flattening is shown in Figure 4.3.

\[ I(t) \]

Figure 4.3: (a) Raw \( I(t) \) signal (black) and its decimated version (yellow). Polynomial fit (red) traces slow baseline fluctuations. (b) Same segment after flattening.

This simple flattening routine does a good job of filtering out the lowest frequencies. The residual baseline drift does not present a significant problem when the magnitude of fast excursions corresponding to molecular activity is at least 3 times larger than the magnitude of the baseline noise. However, when the molecular events have a smaller magnitude, the residual baseline fluctuations may cause artifacts at later stages of analysis (see §4.4).

4.3 Two-level Analysis

Biomolecules tethered to a SWCNT-FET devices generate stochastic pulse trains of excursions \( \Delta I(t) \) excursions. To find a threshold that separates excursions from the baseline, a two-peak curve fitting was performed on histograms of 0.1-1 second data sets. When a
molecule was active and generated $\Delta I(t)$ excursions, curve fitting the histogram would locate two or more Gaussian peaks. The location of the intersect of two Gaussian functions defined a threshold that was used to separate excursions from the baseline current.

Figure 4.4: Determining the signal threshold for three biomolecules (a) lysozyme (b) DNA polymerase I (c) protein kinase A.

Figure 4.4 illustrates the thresholding process for three types of signals generated by different biomolecules. For example, the left part of Figure 4.4a shows a 0.5 s segment of a flattened signal $\Delta I(t)$ generated by an active lysozyme molecule. The right part shows the histogram of that segment ($\Delta I(t)$ vs a number of counts). The major peak of the histogram is at a current baseline level, $\Delta I(t) = 0$ nA. The second peak corresponding to the excursions to the upper level is around $\Delta I(t) = 2$ nA. The intersect of the two Gaussian fits (red)
at 1 nA determine the threshold (purple) used for separating excursions from the baseline. The events from DNA polymerase (Figure 4.4b) and protein kinase A (Figure 4.4c) were extracted using the same thresholding procedure.

Once a threshold was determined, individual events were detected by assigning every point in the data to one of the two states based on whether it was above or below the threshold. This assignment generated asynchronized, binary signal indicating the molecule’s state. An “event” was defined to be a sequence of 10 or more points assigned to the same state. Events that had fewer points were rejected because these events could not be distinguished from extraneous noise. At 600 kHz sampling rate, this rejection threshold was equivalent to a 16 \( \mu s \) cutoff.

Figure 4.5a illustrates the two-level analysis a lysozyme signal. The green signal was generated using a threshold value of 1 nA to separate lysozyme excursions from the baseline current. Figure 4.5b is a magnification of a highlighted part in (a) and illustrates that the binary output carries the information about the number of events, their location, and duration.
Figure 4.5: (a), (b) Binary signal (green) carries the information about the number of events, their location, and duration.
4.4 Transition Time Analysis

Discrete levels identified by the two-level analysis were interpreted as corresponding to different molecular conformations of enzymes. Then, the time it takes an enzyme to change its conformational state is represented by the time it takes the signal to transition between the two $\Delta I(t)$ levels.

Transition analysis was performed on the flattened data after all the switching events within a frame had been identified. The binary signal was used to locate a time stamp of each excursion. Then, the algorithm calculated the low and high $\Delta I(t)$ levels based on the histogram of the raw data in the vicinity of each time stamp. These levels defined the average height $\Delta I(t)$ surrounding a single event. Next, the algorithm calculated the rising and falling transitions duration using 20% and 80% of this local average height as the reference levels (Figure 4.6a). The 20%-80% reference levels captured most of the transition events. However, expanding the thresholds to 10% and 90% which is the traditional criteria for defining rise and fall times in signal processing lead to very long values corrupted by the limited signal-to-noise ratio of the $\Delta I(t)$ signal. Similar measurements using 30% and 70% thresholds lead to proportionally shorter transition duration values. Figures 4.6b,c demonstrate typical outputs of the transition measurement algorithm for the two types of transitions.

Occasionally, the transition analysis algorithm failed to correctly locate a start or an end of a transition resulting in a longer apparent duration. This error happened when the algorithm placed one of the reference levels within the baseline noise (Figure 4.7a). Another common artifact was due to a residual $1/f$ noise. It occurred when the baseline current $\Delta I(t)$ crossed the threshold value used for the two-level analysis. The crossing registered an event in the binary signal for which the transition analysis outputted a duration (Figure 4.7b).

\[1\text{the vicinity was defined as the interval } (t_{stamp} - 1000 \text{ pts}; t_{stamp} + t_{event} + 1000 \text{ pts}) \text{ of each excursion, where } t_{stamp} \text{ was the time stamp of the beginning of each excursion and } t_{event} \text{ was the duration of each excursion.} \]
Figure 4.6: (a) Transition durations are calculated based on the local average high and low states and the reference levels at 20\% and 80\%. (b) An example of a closing transition (red) identified in the vicinity of a time-stamp for an upward transition (green). (c) An example of an opening transition (red) identified in the vicinity of a time-stamp for a downward transition (green).

In this work, to prevent artifacts from obscuring statistics of the underlying conformational dynamics, detected transition events were manually screened, and the erroneous results were excluded from further analysis. A more robust algorithm would require designing a set of rejection criteria and training a classifier on a data set of artifacts using these features.
Figure 4.7: Common analysis artifacts. (a) The transition region is too long because the determined reference levels are buried in the baseline noise. (b) When 10 or more points from the baseline drift over the threshold used in the two-level analysis it generates an event in the binary signal. Every artifact in the binary signal is analyzed, leading to artifacts in the distribution of transition durations.

4.5 Forward-Backward Non-Linear Filtering

Frequently, the 1/f noise of the SWCNT device made it difficult to separate excursions from the signal baseline without introducing artifacts. The signal-to-noise ratio could be improved by low-pass filtering, such as a moving average or a low-pass Butterworth filter. However, this type of linear filtering distorted fast excursions. A signal that originally had a step change showed a continuous ramp across the edge. Figure 4.8 demonstrates this effect known as “edge blurring” on a signal from a SWCNT device. Applying a moving average or a low pass filter distorts the transition between the high and low states by 20-40 µs.

An alternative non-linear digital filtering technique proposed by S.H. Chung and R.A. Kennedy was designed to preserve fast transient events [32]. It assumes no explicit model of the signal and is computationally inexpensive. Conceptually, the way to avoid the edge blurr-
Figure 4.8: Distortion of transitions by a 15 points moving average filter (red) and a first-order 10 kHz Butterworth low-pass filter (blue).

ring is to avoid filtering or averaging at the precise location of transition edges. To achieve that for each data point the algorithm calculates the average value of points that precede it and the average value of points that follow. By dynamically weighting the averages running backward and forward in time the artifacts that distort transition times can be avoided.

As an illustration, consider a signal with a step edge shown in Figure 4.9. A forward predictor of size 3 for the last point before the jump the is given by

\[
\hat{I}(t) = \frac{1}{3}(I(t - 3) + I(t - 2) + I(t - 1)) = 11.83 \text{ nA}
\]

and a backward predictor is

\[
\hat{I}(t) = \frac{1}{3}(I(t + 3) + I(t + 2) + I(t + 1)) = 9.02 \text{ nA}.
\]
Both predictors estimate the desired signal $I(t)$ but are based on different non-intersecting segments of noisy data. The filtered signal combines these two estimates by weighting them

$$\hat{I}(t) = f(t)\hat{I}^f(t) + b(t)\hat{I}^b(t),$$

where $0 \leq f(t) \leq 1$ and $0 \leq b(t) \leq 1$ are the normalized forward and backward weights such that

$$f(t) + b(t) = 1.$$

Using the approach of Chung and Kennedy one can estimate the weights using the following expressions based on the Bayesian framework:
\[ f(t) = \frac{1}{2} \left( \left[ I(t) - \hat{I}^f(t) \right]^2 \right)^{\frac{1}{40}} = 5.66 \times 10^{748} \]
\[ b(t) = \frac{1}{2} \left( \left[ I(t) - \hat{I}^b(t) \right]^2 \right)^{\frac{1}{40}} = 3.34 \times 10^{689} \]

after normalizing the weights, the filtered signal at the event edge

\[ \hat{I}(t) = 1.00 \times \hat{I}^f(t) + 5.90 \times 10^{-60} \times \hat{I}^b(t) = 11.83 \text{ nA}, \]

In this example the jump occurred in the data set \( \{I(t+3) + I(t+2) + I(t+1)\} \), and so the forward predictor was selected, \( f(t) = 1 \). Alternatively, the backward predictor would be selected, \( b(t) = 1 \), if a jump occurred in \( \{I(t-3) + I(t-2) + I(t-1)\} \).

Instead of using a single forward/backward predictor, multiple predictors of different length enable an effective artifact-free filtering of fine signal structures as well as broad signal features. For a bank of \( K \) predictors an estimate of the signal takes the form

\[ \hat{I}(t) = \sum_{i=1}^{K} \left[ f_i(t) \hat{I}^f_i(t) + b_i(t) \hat{I}^b_i(t) \right], \]

and the weights are estimated as following
\[ f_i(t) = \frac{1}{2K} \left( \sum_{j=0}^{M-1} \left[ I(t - j) - \hat{I}_i(t - j) \right]^2 \right)^{-p} \]

\[ b_i(t) = \frac{1}{2K} \left( \sum_{j=0}^{M-1} \left[ I(t + j) - \hat{I}_i(t + j) \right]^2 \right)^{-p} \]

where \( M \) is the analysis window, and \( p \) is the parameter that controls the sharpness of the transitions. A detailed description of the algorithm can be found in [32] and its Matlab code is available online [33]. To improve signal-to-noise of SWCNT devices, the algorithm was implemented in LabVIEW as a VI named "NoRSE filter" that integrates into TDMS analysis software. The best performance and computational efficiency was achieved using a bank of five predictors \( N = [2, 4, 8, 16, 32] \), analysis window \( M = 10 \) and sharpness parameter \( p = 40 \). Using a computer with a single 2.8 GHz CPU, 64-bit and 6 GB of RAM, it took on average 5 s to process 1 s of data acquired at 600 kHz and 15 s to process 1 s of data acquired at 2 MHz. \(^2\) Figure 4.6 shows that the output of the filtering method reduced the noise magnitude and preserved sharp features.

The original work by Chung and Kennedy addressed only the filtering. However, the concept of predictors might be used to design a more robust flattening algorithm described in section (§4.2). We speculate that the regions of the signal that do not contain rapid excursions could be decimated by a smaller factor (more local features would be preserved in the output) and the regions that have excursions could be decimated by a larger factor (more global features would be preserved). This way the output could have fewer baseline fluctuations and the rapid excursions would not be affected.

The adaptive filtering algorithm was developed at a later stage of the Ph.D. work and was not used for analyzing data in this dissertation.

\(^2\)The complexity of this filtering algorithm is approximately \( O(N^3) \)
Figure 4.10: Adaptive filtering reduces noise magnitude and preserves sharp transitions.
Chapter 5

High-Resolution Study of Lysozyme Motions

High-bandwidth transport measurements in the SWCNT-FETs enable real-time recording of the fast molecular events. Specifically, the technique is well suited for studying biochemical events at a single-molecule level. This chapter demonstrates how SWCNT-FETs was used to study the motion of a single lysozyme molecule.

5.1 T4 Lysozyme

A globular protein T4 lysozyme is a key component of the innate immune system. It catalyzes hydrolysis of glycosidic bonds between N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) subunits that make up bacteria cell walls. The mechanical motions of the two sub-domains around the binding site enable structural rearrangements essential for lysozyme catalytic activity (Figure 5.1a). First, a hinge-like mechanical closing of the sub-domains helps distort and strain the NAM ring of the glycan substrate. Then, lysozyme residues
Glu11, Asp20, and Thr26 interact with peptidoglycan and cleave the glycosidic bond between C-1 atom of NAM and O-4 of NAG. Finally, the sub-domains reopen to release the reaction product and prepare the enzyme for the next catalytic cycle (Figure 5.1b).

The hinge-bending motion can occur at two different rates [34]. When the domains open and close at rates 10-80 s\(^{-1}\) glycosidic bonds are being broken. Surprisingly, when the same motion occurs at a rate of 200-400 s\(^{-1}\) it does not result in glycosidic bond hydrolysis, even though peptidoglycan may be bound to the active site. A single lysozyme molecule can remain in one of these states for several seconds before switching a state or becoming inactive [35, 36]. The exact dynamics of the hinge-bending motion and its role in catalysis remains poorly understood.

Unfortunately, the hinge-bending motion is too transient to have been studied using fluorescence [34] or X-ray crystallography [37]. It is also too long to be addressed by either NMR or molecular dynamics simulations. In principle, electronic single-molecule sensing using SWCNT-FETs can overcome these limitations, and the next section describes how this can be done.
5.2 Electrical Measurements

After confirming electrical connectivity of the SWCNT-FET device, PBS was flushed into the microfluidic channel using a SmartSquirt perfusion system (AutoMate Scientific, Inc.). Once the PBS was in direct contact with the SWCNT, source-drain current of SWCNT-FET was measured as the potential between the source electrode and the Ag/AgCl reference electrode $V_{tg}$ was ramped from -0.9V to 0.1V. During the $I(t)$ measurements, $V_{tg}$ was held constant at the value where transconductance $dI/dV_{tg}$ was the largest, typically at -0.6 or -0.7 V with respect to the Ag/AgCl reference electrode.

In every experiment, $I(t)$ data were collected in PBS to confirm that the FET had a steady and featureless baseline current. Satisfactory devices exhibited DC resistances of 0.5 to 5.0 M with no $I(t)$ features outside the 1/f noise spectrum that is normal for SWCNTs (Figure 5.2). Data was collected for 600 s in PBS immediately before the introduction of peptidoglycan, and then for at least 600 s after flushing 2 mL of peptidoglycan solution through the fluid cell. At the conclusion of experiments, $I(t)$ was rechecked after flushing the peptidoglycan with 2 mL of PBS. No data was collected during fluid exchanges or flushing of the cell.

![Figure 5.2: Source-drain current $I(t)$ is featureless in PBS.](image)
5.3 Results and Discussion

Without peptidoglycan substrate in the surrounding electrolyte buffer, lysozyme remained in its open conformation [34, 36, 38]. The SWCNT current $I(t)$ for a lysozyme-conjugated nanocircuit was featureless around a steady baseline (Figure 5.2). After the addition of peptidoglycan, $I(t)$ exhibited stochastic pulse trains of brief excursions to a higher current level (Figure 4.9a). These excursions had all of the characteristics described previously [21]. Specifically, excursions occurred at one of two average rates: either 20-50 s$^{-1}$ corresponding to lysozyme’s processive hydrolysis of substrate, or else 200-400 s$^{-1}$ corresponding to nonproductive motions in which lysozyme processivity and catalysis becomes blocked by peptidoglycan cross-links. These two possible ranges of activity were defined as the catalytic state or the nonproductive state of lysozyme, respectively.

In either state, each transition of lysozyme to its closed conformation moved the positively charged sidechains Lys83 and Arg119 away from the SWCNT. Immersed in an aqueous environment, the two charged sidechains, their appropriate counterions, and the liquid rearranged in sync with the proteins motion. The SWCNT-FET recorded motions of this protein-ion system within the 1-nm Debye screening radius [23]. Each time the enzyme closed, an 8 Å hinge motion brought the two lobes of the enzyme closer, swung the positively charged residues away from the SWCNT, and electrostatically gated the electrical channel to produce the $\Delta I(t)$ signals analyzed here (Figure 5.3).

On average, such $\Delta I(t)$ increases were identical in magnitude, suggesting the enzyme sidechains moved the same distance each time it opened or closed upon substrate. Re-opening of lysozyme reversed the motions and returned $I(t)$ to its baseline value. The mean duration of each excursion $\Delta I(t)$ was interpreted as the duration of the closed configuration, $\tau_{\text{closed}}$, and the pause between excursions was interpreted as the duration of the open configuration, $\tau_{\text{open}}$. Single-molecule distributions of $\tau_{\text{closed}}$ and $\tau_{\text{open}}$ followed Poisson statistics. The mean
durations are reported in the Table 5.1 and are in excellent agreement with previous FRET studies of T4 lysozyme [34, 36, 37, 38].

An experimental resolution of $2 \mu s$ resolved the finite duration of transitions between open and closed conformations. These durations had a broad distribution that included transitions that occurred faster than $2 \mu s$, transitions lasting tens of microseconds, and transitions interrupted by long pauses in the middle of conformational changes.

Examples of closure events of various durations are shown in Figure 5.4 and opening examples are shown in Figure 5.5. In general, the motion from one state to another was observed to have a single, deterministic direction. None of the transitions contained evidence of bistability between the two conformations, indicating that once the conformational change had been initiated the enzyme was committed to the final state.

Figure 5.6a shows the probability distribution for the duration $\tau_{\text{closing}}$ of transitions from the open to the closed conformation. The $\tau_{\text{closing}}$ distribution shown here results from the analysis of over 600 seconds of $\Delta I(t)$ records collected with three different lysozyme-SWCNT devices, each of which gave the same results when analyzed independently. The $\tau_{\text{closing}}$ dis-
Figure 5.4: Three examples of continuous $\tau_{\text{closing}}$ events.

Figure 5.5: Three examples of continuous $\tau_{\text{opening}}$ events.

distribution during catalytically productive closures (green) was statistically identical to the
distribution for events during catalytically nonproductive, higher frequency closures (blue). Both distributions were semilogarithmic with mean durations of approximately 37 $\mu$s. Table 5.1 gives precise values of each slope $\tau$ with its uncertainty (one standard deviation) as determined from a least-squares fit.

Figure 5.6b depicts the two probability distributions for the duration $\tau_{\text{opening}}$ of transitions from the closed conformation back to the open conformation. As with $\tau_{\text{closing}}$, the $\tau_{\text{opening}}$ distributions for the two enzyme conformations were statistically identical and did not depend on the enzymes state, either catalytic or non-productive. Furthermore, the mean value
of $\tau_{opening}$ was the same as for $\tau_{closing}$. Inspection of the example transitions shown in Figure 5.4 and Figure 5.5 indicate that the appearance and timing of lysozymes opening events were virtually identical to its closing events.

Thus, both opening and closing motions are driven by a common energy scale, and the speed of lysozymes motions is almost identical during high-frequency nonproductive motions and less frequent catalytic ones. These findings are consistent with Kramers theory of transition kinetics [39] which expresses the duration of transition in terms of fundamental energy scales and implies that different transition paths should proceed at the same rate once initiated over a reaction barrier [40]. In fact, proteins with vastly different folding rates can exhibit identical transition durations, a prediction experimentally verified by Chung et.al. using FRET [41, 42]. In both protein folding and lysozyme closure, the duration is determined by dissipative interactions such as friction between sidechains surrounding lysozymes active site and the peptidoglycan substrate. These interactions manifest themselves in the frequency distributions of $\Delta I(t)$, but the approach described by Chung et. al.[41] provides a simple method of estimating transition times directly from the values in Table 5.1. Taking $\tau_{open}^{-1}$ as

![Figure 5.6: (a) Probability distributions for the closing durations recorded during periods of catalytic processing (green) or nonproductive binding (blue). (b) Probability distributions for the opening durations. Solid lines depict single-exponential fits to each distribution.](image)
an upper bound for the Kramers rate of barrier crossing and requiring the energy barrier to be no less than the energy difference $\Delta E$ between open and closed conformations, theory predicts transition durations in the range of 10 to 200 $\mu$s, in good agreement with the experimental data (Appendix A).

Unlike protein folding, the duration of lysozymes motions is too transient to have been studied by ensemble methods like fluorescence or X-ray crystallography, and it is too long to be addressed by either NMR or typical first principles simulations. For example, multi-scale molecular dynamics modeling of lysozyme has proposed a <1 ns transition path through six protein conformations [34, 43], but this duration may reflect the limitations of simulation time more than true protein dynamics. The best comparable data comes from FCS, which arguably has the best time resolution among single-molecule techniques. Yirdaw and McHaourab applied FCS to T4 lysozyme and detected a 15 $\mu$s relaxation time between its open and closed conformations [44]. This FCS measurement is in good agreement with the 37 $\mu$s average reported here, especially if one considers that FCS cannot resolve the frequent, short events or the very rare, long-lived events both contained in the full distributions (Figure 5.6). By directly observing both extremes, our electronic technique provides a much fuller representation of the range of lysozymes speeds. Nevertheless, the most probable $\tau_{\text{closing}}$ and $\tau_{\text{opening}}$ durations approach our experimental resolution of 2 $\mu$s. While the distributions in Figure 5.6 show hints of a roll-off at the shortest durations, example events in the Figures 5.4 and 5.5 demonstrate that further improvements are necessary to accurately resolve the shortest events and the true maximum speed of lysozymes motions.

The rare occurrence of long-duration events in both distributions suggested possible correlations between $\tau_{\text{closing}}$ and $\tau_{\text{opening}}$, or between these transition durations and the times $\tau_{\text{closed}}$ and $\tau_{\text{open}}$ spent immediately before or after a transition. However, the four $\tau$ values were found to be independent of each other and uncorrelated in every way, at least within the statistical limits provided by a few thousand events. FCS suggests that much larger data
sets, such as those obtained by monitoring many molecules, might be necessary to reveal weak correlations amongst rare events [44].

The smooth and continuous transitions analyzed in Figure 2 accounted for 90% of the transitions between open and closed conformations. The remaining 10% of events contained a single pause of 40 to 140 μs in the middle of the transition, leading to total transition times of 80 to 180 μs. Example transitions and their probability distributions are shown for closing (Figure 5.7) and opening (Figure 5.8).

![Figure 5.7: Examples of lysozyme closing transitions interrupted by a pause.](image)

![Figure 5.8: Examples of lysozyme opening transitions interrupted by a pause.](image)

The transitions shown here are representative, but the ΔI(t) levels of different pauses were observed to be randomly distributed throughout the entire 20-80% range of enzyme opening.
and closing. Because of the relative rarity of these events, their probability distributions had limited statistics. For clarity, the data sets in Figure 5.9 have not been separated into catalytic and nonproductive states, but an analysis of the catalytic and the nonproductive subsets is included in Table 5.1. When the closing transition was interrupted by a pause, the average duration increased from 37 $\mu$s to about 110 $\mu$s. The opening transitions paused even longer, with an average duration of 135 $\mu$s. Catalytically active cycles occurred 10 to 20 times less frequently than nonproductive ones, so the typical data set of catalytic motions with pauses contained fewer than 100 events. Within these limitations, catalytic activity did not significantly affect the pause duration compared to nonproductive motions.

![Figure 5.9](a) Probability distributions for the durations of closing durations with long-lived pauses (b) Probability distributions for the durations of opening durations with long-lived pauses. Solid lines depict fits to gamma distributions.

Because the SWCNT FET converts the motion of Lys83 and Arg119 sidechains to an electrical signal, the pauses indicate an interruption of these sidechains motions. Since these motions are allosterically linked to the entire protein structure, the interruptions are interpreted as an indication that lysozyme is pausing in some conformation between its fully-open and fully-closed conformations [45]. Furthermore, the durations of these pauses imply complex kinetics among multiple possible conformations. Solid lines in Figure 5.9 depict fits to
the gamma probability distribution for \( N \) consecutive processes

\[
P(t, k, N) = \frac{k^N t^{N-1}}{\Gamma(N)} e^{-kt}
\]

where \( k \) is the shared rate of all \( N \) processes and \( \Gamma \) is the gamma function [46]. Fitting the data to the gamma distribution set an upper bound on \( N \), whereas the statistical variance \( \sigma^2 \) of the same data determined a lower bound \( \sigma^2 = \langle \tau \rangle^2 / N \) [6, 47, 48, 49, 50]. In this manner, it was determined that the closing involved at least 2, but no more than 4 processes during those events which included an intermediate pause. Opening required 3 to 5 processes in series, an increase by one step that is consistent with the longer pauses during \( \tau_{opening} \).

Bhabha et al. have proposed [51] that intermediate steps can play crucial roles in enzyme catalysis, but the intermediate step seen here seems to be independent of chemical activity. For comparison, variances of the smooth and continuous transitions analyzed in Figure 5.9 gave \( N=1.0\pm0.1 \), indicative of a single step or concerted process with no evidence of rate-limiting intermediates.

This statistical analysis is consistent with FCS observations of T4 lysozyme, which noted a range of brightnesses attributed to multiple intermediate states [44]. However, the \( \Delta I(t) \) signals corresponding to \( N > 1 \) never contained two or more distinct levels. The long-lived pause in lysozymes motion was always associated with a single, intermediate configuration, even though statistical analysis indicated multiple consecutive processes. The single plateau during the pause may represent a non-concerted mechanistic pathway during which protein is stuck in a partial conformation while waiting for the correct alignment of functional groups to complete a conformational change.

Figure 5.10 depicts a free-energy landscape for lysozymes primary hinge motions estimated from Boltzman statistics by comparing the durations of \( \tau_{open} \) to \( \tau_{closed} \) [6]. Lysozymes nonproductive cycles occur during closures upon nonhydrolyzable cross-links in the peptidoglycan
(blue), whereas catalytically active motions occur when lysozyme closes upon a hydrolyzable glycosidic bond between $N$-acetylmuramic acid (NAM) and $N$-acetylglucosamine (NAG) (green). Assuming that the open conformation provides a reference energy level that is insensitive to the peptidoglycan structure at the active site, the relative energy level $\Delta E$ of the closed conformation can be estimated using the Boltzmann distribution of a two-state system, $\Delta E = k_B T \ln (\tau_{\text{closed}} / \tau_{\text{open}})$. As in the previous work, the closed configurations are at 1.6 and 2.6 kcal/mol for nonproductive and catalytic closures, respectively [21]. The two closed conformations differ by 1.0 kcal/mol. Energetically, this difference is within the range expected for the formation of an additional hydrogen bond at the active site as a result of successful hydrolysis. The two-state model does not apply if lysozyme’s motions are limited by more substantial energy barriers, but FRET experiments observe lysozyme accessing a range of conformations, including the closed conformation, at high frequency and without any apparent stabilization when substrate is absent [38].

An estimate for the energy level of intermediate configurations was added to this basic diagram using the mean pause durations observed in each type of cycle. Again, assumptions have been made that the observed waiting times are limited by thermodynamics rather than other, rate-limiting barriers. During nonproductive motions, the intermediate configuration was only slightly higher in energy than the closed configuration, by either $0.28 \pm 0.09$ kcal/mol (closing) or $0.16 \pm 0.13$ kcal/mol (opening). Essentially, the energy levels for these closing and opening pauses were statistically identical, suggesting a single unique conformation that was nearly isoenergetic with the nonproductive closed conformation. During catalytically productive cycles, the intermediate pauses occurred at much higher energies (3.4 to 3.8 kcal/mol) than the closed conformation.

The identical distributions of $\tau_{\text{opening}}$ and $\tau_{\text{closing}}$, independent of the catalytic state, is unlikely to be coincidental. Taken together with a 10% propensity to form intermediates during opening and closing in both states, the results suggest the presence of symmetry in
lysozymes hinge-bending motions. Such symmetry of motions has been invoked to explain the movements of Brownian ratchets, which rely on constrained motions and thermal energy to power their movement along a substrate [52, 53]. In fact, lysozyme processively steps along a polysaccharide substrate to consecutively hydrolyze hundreds of glycosidic bonds [21], analogously to a Brownian motor or similar molecular machine.

For this symmetry of motions to apply to lysozyme, the enzyme must open and close upon identical substrates. This condition is clearly satisfied by non-productive motions, during which the enzyme opens and closes upon a peptdioglycan cross-link without catalyzing any change in the substrate [20]. Productive motions, on the other hand, hydrolyze one glycosidic bond during each closure. For the enzyme to open upon identical substrates after this chemical modification, the enzyme needs to move along the polysaccharide to the next glycosidic

Figure 5.10: Free-energy landscape of T4 lysozyme.
bond before reopening. Thus, the kinetic evidence supports a model in which the enzyme closes, hydrolyzes the substrate, translates to the next bond, and then opens. Consequently, it was concluded that lysozyme operates as a Brownian motor. The $1.0 \text{kcal/mol}$ liberated in each catalytic cycle provides the energy for translocation along the polysaccharide. Thermal energy could be powering enzyme closing and opening along a constrained pathway. The fact that motions are disorganized until the enzyme binds its substrate [34, 36] indicates that the substrate provides some essential structural feature, which could be identified by further modeling and structural studies. The catalytic mechanism of lysozyme is well established [54] and is depicted in Figure 5.1b. The data described here suggests the enzyme closes and opens upon an identical glycosidic bond, which allows assignment of the enzyme open and closed conformations to this mechanism.

The formation of a rare intermediate conformation comprised $10\%$ of enzyme openings and closings in both catalytic and non-productive states. Assuming this intermediate does result from heterogeneity in the polysaccharide substrate, the intermediate need not require an alternate conformational trajectory. Instead, observations of the intermediate might illustrate cases where the concerted mechanism of Figure 5.1b is converted into a multi-step, non-concerted mechanism with higher transition state energies. The variable $N$ matches the number of arrows drawn for the non-concerted reaction in the upper left of Figure 5.1b. After any arrow, a pause to allow repositioning of sidechains or other reaction participants would result in the intermediate conformation being observed. The range of $\Delta I(t)$ heights for the observed intermediates suggests that no single step in the mechanism presents a special bottleneck. Furthermore, the symmetry of motions observed here could result from similar conformational pathways for enzyme closing and opening; thus, if the enzyme forms an intermediate $10\%$ of the time while closing, the intermediate is also likely to occur during the constrained opening events. As noted above, the intermediates were observed in both directions during both catalysis and non-productive
motions, demonstrating the extent of conformational constraints for this Brownian motor.

<table>
<thead>
<tr>
<th>Conformation</th>
<th>Catalytically active</th>
<th>Nonproductive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau$(µs)</td>
<td>$\Delta E$(kcal/mol)</td>
</tr>
<tr>
<td>open</td>
<td>52,330 ± 2,290</td>
<td>0</td>
</tr>
<tr>
<td>closing (no pause)</td>
<td>38 ± 2</td>
<td></td>
</tr>
<tr>
<td>(with pause)</td>
<td>81 ± 27</td>
<td>3.84 ± 0.20</td>
</tr>
<tr>
<td>closed</td>
<td>589 ± 61</td>
<td>2.66 ± 0.07</td>
</tr>
<tr>
<td>opening (no pause)</td>
<td>41 ± 5</td>
<td></td>
</tr>
<tr>
<td>(with pause)</td>
<td>181 ± 54</td>
<td>3.36 ± 0.18</td>
</tr>
</tbody>
</table>

Table 5.1: Timing and energetics of T4 lysozyme closing and opening

<table>
<thead>
<tr>
<th>Device</th>
<th>Measurement Date (mm-dd-yy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKC43-GD, 22eg</td>
<td>09/09/13</td>
</tr>
<tr>
<td>AKC43-GB, 32ce</td>
<td>10/09/13</td>
</tr>
<tr>
<td>AKC43-JG, 11df</td>
<td>10/15/13</td>
</tr>
</tbody>
</table>

Table 5.2: Lysozyme SWCNT-FET devices that generated data
Chapter 6

Study of DNA Polymerase and Protein Kinase A

The high-bandwidth recording of SWCNT-FET devices functionalized with a single molecule successfully resolved conformational motions of lysozyme. In this chapter, the same technique was applied to two other enzymes. The chapter describes high-resolution measurements of DNA polymerase I and cAMP-dependent protein kinase A. These two molecules have catalytic cycles that involve multiple binding partners and conformational steps, making them somewhat more complex than lysozyme. Sections §6.1 and §6.2 describe conformational dynamics of each enzyme. Section §6.3 compares the kinetics of transitions between the conformational states for all three enzymes.

6.1 DNA Polymerase I

DNA Polymerase I plays a key role in DNA replication. This fundamental biological process has at least seven steps. Once the DNA double-helix is unzipped by DNA helicase, the
individual strands (also referred to as "templates") are ready to be replicated. First, DNA polymerase I binds a template and a primer, which is usually a short strand of RNA, resulting in a DNA complex, $E \bullet \text{DNA}_n$ (Figure 6.1c, step 1). Additional binding of the correct nucleotide and a metal ion, such as $\text{Mg}^{2+}$, promotes the formation of the ternary complex, $E \bullet \text{DNA}_n \bullet \text{dNTP}$ (steps 2,3). During this step, DNA polymerase selects a nucleotide complementary to the next unpaired base in the template and rejects possible mismatches. After binding the correct dNTP, polymerase transitions from an open conformation to a closed conformation (step 4). Then, the phosphodiester bond is formed between the 3’ OH group and the α-phosphate of the incoming dNTP, adding a new nucleotide to the primer, DNA$_{n+1}$. The reaction generates a pyrophosphate (PP$_i$) and releases a hydrogen (step 5). Then, DNA polymerase returns to an open conformation, releases the pyrophosphate product, and translocates to the next adjacent base (steps 6,7)[55].

![Figure 6.1: Mechanism of DNA replication (a,b) and its catalytic cycle (c)](image-url)

The particular polymerase used for this work was the Klenow Fragment (KF) of DNA Polymerase I. KF has four sub-domains: a template binding domain (thumb), a domain for dNTP binding (fingers), a catalytic domain (palm), and an exonuclease domain for proofreading. The 10 Å swinging motion of the fingers domain around the active site results in two conformations (Figure 6.2) that have been studied using fluorescence techniques [56]. However, the
exact relationship between the motions of the subdomains and the catalytic steps remains an active area of research [57]. In principle, the high-resolution recording of KF motions during catalysis might elucidate the individual steps such as molecular recognition, nucleotide incorporation, or proofreading.

Figure 6.2: Structure of Klenow Fragment and its two conformational states.

6.1.1 Electrical Measurements

The electrical measurements of the SWCNT-FET device with a tethered KF molecule were performed similarly to lysozyme. The device was biased at 100 mV and the liquid potential was held at a constant value between -0.5 and -0.2 V with respect to Ag/AgCl reference electrode. In every experiment, source-drain current, $I(t)$, data were collected in a standard buffer solution to confirm a featureless baseline current (Figure 6.3a). Then, a substrate containing a homopolymeric template (100 nM) and a complimentary dNTP (10 µM) was flushed into the microfluidic channel. The flow cell was flushed with a buffer solution before the introduction of each new template to wash out the previous solutions. Devices that were measured over multiple days were kept in the flow cell filled with a buffer solution.
6.1.2 Results and Discussion

Previous work demonstrated that in the presence of both template and complimentary dNTP, KF generated two-level fluctuations in $\Delta I(t)$. The baseline current $\Delta I(t)$ corresponded to an open KF conformation while the low current level $\Delta I(t)$ represented a closed KF conformation. Excursions with an average duration $\tau_{\text{closed}}$ of 0.3 ms indicated the formation of one base pair. The rate of these $\Delta I(t)$ excursions reproduced the KF incorporation rates measured by single-molecule FRET and ensemble experiments [24].

In this work, similar two-level fluctuations were observed. Characteristic signals for substrates containing poly(dA)$_{42}$ with dTTP and poly(dT)$_{42}$ with dATP are shown in Figure 6.3b and Figure 6.3c, respectively. The excursions were not observed in a buffer solution or in the absence of either a dNTP or a DNA template, suggesting that the additional fluctuations correspond to the catalytic activity of KF.

For both poly(dA)$_{42}$ and poly(dT)$_{42}$ substrates the duration of excursions varied from tens of microseconds to milliseconds. Figure 6.4a shows a 50 ms segment of $\Delta I(t)$ with individual excursions of various durations highlighted in red, green, and yellow. The magnified view of these events is shown in Figure 6.4b,c,d.

The short events like the one shown in Figure 6.4b constituted more than 80% of all closed events. A smaller fraction of events had a time scale of a few hundred microseconds (Figure 6.4c) and, based on the previous research [24], corresponded to the incorporation of a complementary dNTP. Finally, there were long events that lasted up to a millisecond (Figure 6.4d).

Figure 6.5a shows a histogram of durations KF spent in a closed conformation, $\tau_{\text{closed}}$. The shape of the distributions for both substrates suggested that events shorter than 200 $\mu$s had a different time constant than longer events. The double exponential fits showed in pink
and green revealed that the population of short events had a time constant, $\tau_{\text{short}}$, of 36 $\mu$s and 78 $\mu$s for poly(dA)$_{42}$ and poly(dT)$_{42}$ respectively. Longer events had time constants, $\tau_{\text{long}}$, of 107 $\mu$s and 276 $\mu$s for poly(dA)$_{42}$ and poly(dT)$_{42}$ respectively. Similar analysis was performed for the distributions of open durations, $\tau_{\text{open}}$, shown in Figure 6.5b. The double exponential fits revealed that more than 70% of events had a mean time, $\tau_{\text{short}}$, of 196 $\mu$s and 119 $\mu$s for poly(dA)$_{42}$ and poly(dT)$_{42}$ respectively. Longer events had a mean time, $\tau_{\text{long}}$, of 8.3 ms and 1.7 ms for poly(dA)$_{42}$ and poly(dT)$_{42}$ respectively. The time constants are summarized in Table 6.2.

The presence of the fast and slow populations in $\tau_{\text{closed}}$ distributions implied that KF activity involved at least two kinetic processes. In order to understand which events contributed
to each process, the events were categorized into three categories: events with a duration between 16 and 100 \( \mu s \) (Figure 6.4b), events longer than 100 \( \mu s \) (Figure 6.4c), and events like Figure 6.4d that had interruptions in a closed state. The automatic classification was performed by comparing the binary outputs of the two-level analysis using two rejection thresholds (see Chapter 4, §4.3) - 16 \( \mu s \) and 100 \( \mu s \). Figure 6.6 shows the two binary outputs for a 16 \( \mu s \) rejection threshold (red) and a 100 \( \mu s \) rejection threshold (green).

In the binary signals, the high state value corresponded to an open conformation of KF and was labeled "HI", while the low state corresponded to a closed conformation and was labeled "LO". Then, the binary outputs were used to generate a signal with four numeric states (0-3) using the logic Table 6.1. The output signal is shown in blue in Figure 6.6. It encoded three categories of events shown in Figure 6.4b,c,d as following motifs: "323" represented short events, "303" - events longer than 100 \( \mu s \), and "010" represented events with one or more interruptions.
Figure 6.5: Distributions of open and closed durations for poly(dA)$_{42}$ (green) and poly(dT)$_{42}$ (blue) DNA templates in the presence of a complementary dNTP. (a) Histograms of $\tau_{\text{closed}}$ durations (b) Histograms of $\tau_{\text{open}}$ durations. Double exponential fits are shown as solid lines.

Figure 6.6: Binary signals obtained using two rejection thresholds in two-level analysis: 16 $\mu$s (red) and 100 $\mu$s (green). The two signals were compared using the logic summarized in Table 6.1, and the output is shown in blue.

The distributions $\tau_{\text{closed}}$ for each motif are shown in Figure 6.7. The distributions of "303" events are shown blue. The mean time determined by the single exponential fits was 185 $\mu$s and 128 $\mu$s for poly(dT)$_{42}$ and poly(dA)$_{42}$ respectively. The distributions of "010" events are shown in green. Solid green lines are the fits to the gamma probability distribution for
Table 6.1: Logic used to generate a four-state signal

<table>
<thead>
<tr>
<th>Red</th>
<th>Green</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO</td>
<td>LO</td>
<td>0</td>
</tr>
<tr>
<td>HI</td>
<td>LO</td>
<td>1</td>
</tr>
<tr>
<td>LO</td>
<td>HI</td>
<td>2</td>
</tr>
<tr>
<td>HI</td>
<td>HI</td>
<td>3</td>
</tr>
</tbody>
</table>

N consecutive processes. For poly(dT)$_{42}$, the fitting parameter $N = 2.0 \pm 0.3$ suggesting 2-3 consecutive Poisson processes. The mean duration was 940 $\mu$s. For poly(dA)$_{42}$ $N = 2.4 \pm 0.3$ implying up to 3 consecutive processes. The mean duration was 458 $\mu$s. Combined, events ”323” and ”010” represented the slower KF dynamics. The mean time scale of a few hundred microseconds was consistent with the $\tau_{closed}$ distributions reported previously. Therefore, it was hypothesized that events ”303” and ”010” are associated with the incorporation of a complementary base pair. The interruptions observed in long ”010” events could result from KF attempts to re-position the nascent base pair if it had been in the wrong orientation when KF closed.

Motif ”323” shown in red had two orders of magnitude more events than ”303” or ”010” suggesting that most KF events were short. The single exponential fits determined the mean time of 86 $\mu$s and 50 $\mu$s for poly(dT)$_{42}$ and poly(dA)$_{42}$ respectively. The mean time scale of ”323” events and their distribution suggested that these events represented fast KF dynamics observed in Figure 6.5. This rapid KF dynamics could not represent base pair incorporation which only occurs at a rate of 15-20 s$^{-1}$. The fast events could be a result of a different process such as fast O-helix motions during stability checking of a nascent base pair. Such O-helix motions have been recently invoked to explain KF tolerance for dNTPs analogs [58].

\footnote{Due to the artifacts described in Chapter 4 the motif assignment algorithm misclassified some events. Without the artifacts, the distribution of ”323” events should not extend beyond 100 $\mu$s.}
We conclude this discussion by noting that the distributions in Figure 6.5 suggest that KF kinetics depended on the template it was processing. When KF was processing poly(dT)$_{42}$ template $\tau_{\text{closed}}$ was 2-3 times longer than when it was replicating poly(dA)$_{42}$. However, $\tau_{\text{open}}$ during poly(dT)$_{42}$ processing was 2-5 times shorter compared to poly(dA)$_{42}$. Combined, these results imply that the high-bandwidth recordings of KF activity could provide information essential for distinguishing between different base pairs.

<table>
<thead>
<tr>
<th>Conformation</th>
<th>poly(dA)$_{42}$</th>
<th>poly(dT)$_{42}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_{\text{short}}$ ($\mu$s)</td>
<td>$\tau_{\text{long}}$ ($\mu$s)</td>
</tr>
<tr>
<td>open</td>
<td>196 ± 2</td>
<td>8319 ± 205</td>
</tr>
<tr>
<td>closed</td>
<td>36 ± 1</td>
<td>107 ± 2</td>
</tr>
</tbody>
</table>

Table 6.2: Events in the distributions shown in Figure 6.5 belonged to either a population of short events with a mean time $\tau_{\text{short}}$ or a population of long events with a mean time $\tau_{\text{long}}$. 
6.2 Protein Kinase A

The enzyme protein kinase A (PKA) plays an important role in extracellular signaling. It activates other proteins by adding a phosphate group. Normally, PKA itself is inactive until a cyclic adenosine monophosphate (cAMP) binds to its regulatory subunit. Once activated, PKA phosphorylates target proteins by binding simultaneously two ligands: an ATP and a peptide substrate, such as a Kemptide. Next, the enzyme undergoes a conformational change from an open state to an intermediate state, and can then transition to a fully closed conformation [59]. Therefore, unlike lysozyme and polymerase, PKA has three stable conformational states. The exact dynamics of transitions between the states remains mostly unknown. That motivated a high-bandwidth study of PKA motions.

Previous work showed that when PKA interacted with its cofactor ATP or its peptide substrate Kemptide, $I(t)$ of the PKA-functionalized SWCNT device rapidly fluctuated between the two levels [25]. The high state corresponded to an open PKA conformation and the low state represented an intermediate state. In the presence of both ATP and Kemptide in an electrolyte solution, PKA can access a fully closed state that completes formation of the catalytically functional ternary complex [59]. This new conformation manifested itself as a third level in $I(t)$ signal.

In this work, the SWCNT-FET device functionalized with PKA was measured in a buffer solution and then in a solution containing a mixture of ATP and Kemptide. Unfortunately, over the course of the measurement the device became inactive, and no activity was observed in solutions containing just ATP or Kemptide. Therefore, the data presented in this section are preliminary.
6.2.1 Electrical Measurements

The electrical measurements of the SWCNT-FET device with a tethered PKA molecule were performed at a source-drain bias of 100 mV. The liquid potential was held constant at -0.6 V with respect to Ag/AgCl reference electrode. The source-drain current, $I(t)$, data were collected in standard buffer solution (10 mM Tris, 50 mM NaCl, 10 mM DTT, 10 mM MgCl$_2$, pH 8) to confirm a featureless baseline current (Figure 6.8a). Then, a substrate containing a solution of ATP (2 mM) and Kemptide (100 µM) was flushed into the microfluidic channel.

![Signal of a PKA-labeled SWCNT circuit (a) and shows rapid two-level switching in the presence of ATP and Kemptide (b-c).](image)

Figure 6.8: Signal of a PKA-labeled SWCNT circuit is featureless in a buffer solution (a) and shows rapid two-level switching in the presence of ATP and Kemptide (b-c).

6.2.2 Results and Discussion

When the device was exposed to a solution of ATP and Kemptide $I(t)$ exhibited rapid two-level fluctuations shown in Figure 6.8b. Considering the previous work [25], the lack of the third level in the presence of both ATP and Kemptide was surprising.
In order to identify the two levels, the mean duration spent in a high and a low state, $\tau_{hi}$ and $\tau_{lo}$ was compared to the time spent in each of the three states discovered in the previous work [25]. The mean duration of $\tau_{lo}$ was $0.58 \pm 0.02$ ms and matched the duration of the intermediate state $0.53 \pm 0.03$ ms, while the closed state was three times longer $1.45 \pm 0.06$ ms. The duration of the high state, $6.41 \pm 0.67$ ms, was comparable to the duration of the open state, $5.04 \pm 0.40$ ms.

Figure 6.9 shows the distributions of $\tau_{hi}$ and $\tau_{lo}$ plotted alongside the distributions of the time spent in an open and an intermediate state from the previous work. To match the resolution limitation of the previous measurement events shorter than 100 $\mu$s were excluded from the analysis.

![Figure 6.9: Distributions of PKA’s intermediate (a) and open (b) durations from the previous work are shown in black. The distributions of $\tau_{lo}$ (a) and $\tau_{hi}$ (b) from the flow cell measurement are shown in red.](image)

After identifying the high and low states as PKA’s open and an intermediate conformation, the high-resolution signal was investigated in more details. Figure 6.10 shows the distributions of $\tau_{inter}$ and $\tau_{open}$ after including events shorter than 100 $\mu$s into the two-level analysis. More than 90% of the intermediate events followed a single-exponential distribution with a mean time of $50 \pm 2$ $\mu$s. The rest of the events belonged to one of the two populations.
with the mean values of $105 \pm 2 \mu s$ and $203 \pm 7 \mu s$. The fits are shown as solid lines in Figure 6.10a. Similarly, the $\tau_{\text{open}}$ distribution was fitted to three single-exponential distributions. About 80% of the open events had a mean time of $139 \pm 3 \mu s$, 15% had a mean time of $369 \pm 6 \mu s$, and the rest had a mean time of $654 \pm 13 \mu s$.

![Figure 6.10: Distributions of PKA's intermediate (a) and open (b) duration. The distributions were fitted to three exponentials shown in solid lines.](image)

### 6.3 Comparing Transition Durations of Three Enzymes

The three enzymes measured in this work: T4 lysozyme, DNA polymerase I, and protein kinase A have different structures and functions. Lysozyme has a molecular mass of 15 kDa, while PKA has a mass of 40 kDa, and a mass of KF is 68 kDa. In the absence of a binding partner, all three enzymes adopt an open conformation. During catalysis, enzymes interact with a substrate and/or a co-factor and transition to a closed conformation, sometimes through quasi-stable intermediate states. Comparing transition durations $\tau_{\text{closing}}$ and $\tau_{\text{opening}}$ can lead to new insights into the dynamics of conformational changes, determine which enzymes move faster and what determines the speed.
Figure 6.11: Distribution of closing (a) and opening durations (b) for T4 Lysozyme (yellow), PKA (green), KF (blue). For clarity, the insets show the area of the graph between 0 and 50 µs.

Lysozyme closing and opening motion required on average 37 µs. For KF, the mean closing duration $\tau_{\text{closing}}$ was 4 µs and the mean opening duration $\tau_{\text{opening}}$ was 3 µs. Finally, PKA transitioned between an open state and an intermediate conformation in approximately 5 µs both ways. The distributions are shown in Figure 6.11 and the mean times are summarized in Table 6.3. Though PKA had a slightly longer transition duration than KF, the time scale of transitions for both enzymes approached our experimental resolution of 1.6 µs. The difference of 1-2 µs is too small to make a meaningful comparison. However, the transition durations of T4 lysozyme is at least 10 times longer, implying a very different conformational dynamics.

The observation that lysozyme is the slowest of three enzymes is surprising, considering that it is a smaller molecule. A more intuitive model would assume that a protein with fewer residues moves faster. Such models have been proposed for a protein folding dynamics. For example, Kubelka et al. proposed that the protein folding ”speed limit” scales linearly
with the number of residues [60]. Naganathan et al. proposed a semi-empirical relation for a folding time $\tau \sim exp(\sqrt{N})$, where $N$ is the number of residues [61]. Just like folding or unfolding of a protein, conformational motions of enzymes during catalysis involve structural re-arrangement of many amino acids and cannot be instantaneous. Therefore, it is reasonable to assume that the transition time between the two conformations is also proportional to the number of moving residues. Assuming that the entire lysozyme moves during a transition, it takes on average $37 \, \mu s$ to move 164 amino acids. Then, the faster transitions of KF and PKA imply that the conformational motions of these larger molecules involve only a fraction of their residues. At the attachment site, the motion of the two charged residues could be allosterically coupled to a motion of a few actively moving residues. For KF, this hypothesis is consistent with the recent finding that each $\Delta I(t)$ excursion corresponded to a motion of O-helix residues [58].

Finally, the distributions of transition times for all three proteins include events longer than $1.6 \, \mu s$ proving that the transition durations of enzymes can be resolved using the SWCNT-FET. Further reduction of the parasitic capacitances in the SWCNT circuit will push the resolution into a sub-microsecond timescale and enable recording of conformation dynamics that approaches a hypothesized $1 \, \mu s$ speed limit for large proteins [62].

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Closing Duration ($\mu s$)</th>
<th>Opening Duration ($\mu s$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T4 Lysozyme</td>
<td>38 ± 2</td>
<td>41 ± 5</td>
</tr>
<tr>
<td>Protein Kinase A</td>
<td>5 ± 2</td>
<td>5 ± 2</td>
</tr>
<tr>
<td>DNA Polymerase I (KF)</td>
<td>4 ± 2</td>
<td>3 ± 2</td>
</tr>
</tbody>
</table>

Table 6.3: Transition Durations of Three Enzymes
<table>
<thead>
<tr>
<th>Device</th>
<th>Measurement Date (mm-dd-yy)</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKC42-GM, 13fh</td>
<td>06/10/14</td>
<td>KF</td>
</tr>
<tr>
<td>AKC55-HB, 20eg</td>
<td>03/29/15</td>
<td>KF</td>
</tr>
<tr>
<td>AKC43-HC, 31fh</td>
<td>04/15/15</td>
<td>KF</td>
</tr>
<tr>
<td>AKC55-LE, 03bd</td>
<td>05/31/15</td>
<td>PKA</td>
</tr>
</tbody>
</table>

Table 6.4: KF and PKA SWCNT-FET devices that generated data
Chapter 7

Conclusion

In summary, this dissertation investigated the microsecond dynamics of three enzymes using a high-bandwidth transport measurement through SWCNT-FET devices. To ensure that the output signal was free from external noise across a 200 kHz measurement bandwidth, parasitic capacitance was minimized by tightly integrating electrical components and constraining the liquid within a microfluidic channel in an electrochemical flow cell. Eventually, the temporal resolution improved from 100 µs to 1.6 µs.

The faster time resolution allowed the direct recording of lysozymes opening and closing transitions. Both motions required 37 µs, on average. The distribution of transition durations was also independent of the enzymes state: either catalytic or nonproductive. The observation of smooth, continuous transitions suggested a concerted mechanism for glycoside hydrolysis with lysozymes two domains closing upon the polysaccharide substrate in its active site. These smooth motions were distinguished from a nonconcerted mechanism, observed in approximately 10% of lysozyme openings and closings, in which the enzyme paused for an additional 40-140 µs in an intermediate, partially closed conformation. During intermediate forming events, the number of rate-limiting steps observed increased to four, consistent with
four steps required in the stepwise, arrow-pushing mechanism. The formation of such intermediate conformations was again independent of the enzymes state. The observed symmetry in enzyme opening and closing thus suggests that substrate translocation occurs while the enzyme is closed.

Next, the high-bandwidth technique was applied to study the motions of Klenow Fragment of DNA Polymerase I. In the presence of a homopolymeric template and a complimentary dNTP, KF generated two-level fluctuations $\Delta I(t)$. The distributions of $\tau_{\text{closed}}$ revealed two populations of events: one with a mean time of 36 $\mu$s (poly(dA)$_{42}$) and 78 $\mu$s (poly(dT)$_{42}$), and another population with a mean time of 107 $\mu$s (poly(dA)$_{42}$) and 276 $\mu$s (poly(dT)$_{42}$). The rapid KF dynamics could not represent base pair incorporation which only occurs at a rate of 15-20 s$^{-1}$ and could result from a different process. The mean duration of the longer KF events was consistent with the $\tau_{\text{closed}}$ distributions of the base pair incorporation events reported previously. However, the improved resolution of the flow cell revealed that KF closures that lasted more than 400 $\mu$s were interrupted by one or more excursion to an open state. The $\tau_{\text{closed}}$ distributions of events with interruptions followed a gamma distribution for 2-3 consecutive Poisson processes. Overall, when KF was processing poly(dT)$_{42}$ template $\tau_{\text{closed}}$ was 2-3 times longer than when it was replicating poly(dA)$_{42}$. However, $\tau_{\text{open}}$ during poly(dT)$_{42}$ processing was 2-5 times shorter compared to poly(dA)$_{42}$.

Preliminary results were shown for the high-bandwidth measurement of protein kinase A. In the presence of both ATP and Kemptide, $I(t)$ exhibited rapid two-level fluctuations. The lack of the third current level observed in the previous work suggested that within the 2 $\mu$s resolution PKA sampled only two conformations. Comparing the mean durations of $\tau_{\text{hi}}$ and $\tau_{\text{lo}}$ with the previous results revealed that the two states corresponded to the open and the intermediate conformations.

The three enzymes were compared based on their closing and opening durations. Lysozyme motions required on average 37 $\mu$s, while the mean closing and opening durations of KF and
PKA were between 3 and 5 µs. Though PKA had a slightly longer transition duration than KF, the time scale of transitions for both enzymes approached our experimental resolution of 1.6 µs.

Finally, the apparatus and the findings described in this dissertation can be used in future work to provide new insights into the dynamics of conformational changes. For example, performing the electronic measurements described in Chapter 5 at various temperatures would determine how temperature affects the speed of lysozyme hinge-bending motions. For DNA Polymerase, performing experiments described in Chapter 6 with substrates containing a mixture of nucleotides and the mismatched nucleotide alone could further elucidate the nature of different excursions observed in the SWTN-FET signals. For protein Kinase A, performing high-bandwidth measurements in ATP, Kemptide, and a mixture of ATP and Kemptide on a single SWNT-FET device would help explain rapid two-level switching observed in the presence of both ATP and Kemptide.
Bibliography


87


Appendices

A Estimation of Transition Path Time

The following derivation estimates the diffusive transition times for a barrier crossing, \( \langle \tau_{TP} \rangle \).

From Eqn. [3] from [63]:

\[
\langle \tau_{TP} \rangle \approx \frac{\ln[2e^\gamma \ln(k_o/k)]}{2\pi k_o}
\]

where \( \gamma = 0.577 \) is Euler’s constant, \( k_o \) is the Kramers pre-exponential factor, and \( k \) is the rate coefficient.

The expression can be rearranged using the relationship between \( k_o \) and \( k \) so that it only depends on the barrier height \( G^\dagger \) and \( k \). This avoids the problem of estimating \( k_o \), which depends on the curvature of the barrier and the well.

\[
\langle \tau_{TP} \rangle \approx \frac{e^{-G^\dagger \beta} \ln(2G^\dagger e^\gamma)}{2\pi k}
\]

where \( \beta = 1/k_B T \)

Finally, \( \langle \tau_{TP} \rangle \) is estimated by guessing values for \( G^\dagger \) and \( k \). The upper limit for \( k \) is determined by \( 1/t_{open} \) since \( k \) can’t be any faster than the duration measured experimentally.
Assuming that the barrier crossing is the main rate-limiting component, the rate can be estimated as $k = 1/t_{\text{open}}$. For lysozyme, the measured mean durations were 2.67 ms (non-productive) and 52.33 ms (catalytically productive).

The barrier $G^\dagger$ must be higher than $2 \, k_B T$ and the energy of the final closed conformation. From the switching times, it can be assumed that the closed conformation is at an energy of at least $2.7 \, k_B T$ (nonproductive) or $4.5 \, k_B T$ (catalytically productive). If the barrier $G^\dagger$ is the same for both activities, then it must be at least 4.5-5.0 $k_B T$. However, there is a possibility that $G^\dagger$ is actually a different barrier in the nonproductive state, so perhaps it drops as low as $3 \, k_B T$, but that would be an absolute minimum and wouldn’t apply to the catalytic state.

Plugging these values in, the following likely values of the transition time $\langle \tau_{TP} \rangle$ (all given in $\mu s$) are obtained:

$$\langle \tau_{TP} \rangle \approx 13 \, \mu s, \text{ when } \gamma = 0.577, k = 1/0.00267, \beta = 1, \text{ and } G^\dagger = 4.5$$

$$\langle \tau_{TP} \rangle \approx 257 \, \mu s, \text{ when } \gamma = 0.577, k = 1/0.05233, \beta = 1, \text{ and } G^\dagger = 4.5$$

$$\langle \tau_{TP} \rangle \approx 21 \, \mu s, \text{ when } \gamma = 0.577, k = 1/0.00267, \beta = 1, \text{ and } G^\dagger = 4.0$$

$$\langle \tau_{TP} \rangle \approx 50 \, \mu s, \text{ when } \gamma = 0.577, k = 1/0.00267, \beta = 1, \text{ and } G^\dagger = 3.0$$