Statistical Modeling and Application of Gramicidin A Ion Channel

by

Kai Yiu Luk

B. A. Sc., The University of British Columbia, 2005

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

The Faculty of Graduate Studies

(Electrical and Computer Engineering)

The University Of British Columbia

July, 2007

Abstract

Ion channels are aqueous pores in the cell membrane for selected ions to flow down their electrochemical gradient. These channels play a prominent role in a variety of biological processes in the human body. Determining the structure and function of ion channels is of fundamental importance in biology. Also, the selective conductivity and specific gating mechanism of ion channels have attracted much interest in the area of artificial molecular detectors. Ion channel based biosensors are developed to detect molecular species of interest in medical diagnostics, environmental monitoring and general bio-hazard detection. This thesis is concerned with statistical techniques used to describe ion channel permeation and to develop ion channel based biosensors. Brownian dynamics is a popular technique to simulate ion channel permeation but is too computationally expensive to run when ionic concentration is high. By fitting binding site statistics of BD simulation to a semi-Markov chain, we obtain a simpler model with conduction properties that are statistically the same as the simulations. This approach enables the use of extrapolation techniques to predict channel conduction when performing the actual simulation is computationally infeasible. Numerical studies on the simulation of gramicidin A channels are presented. In a separate study, we show the use of statistical modeling and detection techniques as

part of a sensitive biosensing platform. A nano-scale biosensor is built by incorporating dimeric gramicidin A channels into bilayer membranes of giant unilamellar liposomes. The presence of specific target molecules changes the statistics of the biosensor's conduction. By capturing the change in real time, we devise a maximum likelihood detector to detect the presence of target molecules. The performance of the biosensor is tested with the addition of various target molecules known to inhibit conduction of gramicidin A channels. Experimental results show that the detection performed well even when the conductance change was difficult to visualize. The detection algorithm provides a sensitive detection system for ongoing development of membrane-based biosensors.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor, Dr. Vikram Krishnamurthy, for his guidence. Thanks to my collaborator, Bruce Cornell, for providing figures for this thesis. Thanks to my sister Candy for helping me with the biology and Phoebe for her support and encouragement. Thanks to my parents for their unconditional support. This thesis is dedicated to them.

Chapter 1

Introduction

Ion channels are aqueous pores in the cell membrane for selected ions to flow down their electrochemical gradient. The class of ion channels is very diverse. In many channels, ion conduction is controlled by a gate, which opens and closes in response to electrical and chemical signals. In addition, ion channels are highly selective, allowing only desired ions to pass through. Thus ion channels are used to regulate membrane potential as well as intracellular and extracellular concentrations of ions.

Ion channels play a prominent role in a variety of biological processes in the human body. In the nervous system, voltage-gated sodium and potassium channels open and close co-operatively to propagate nerve impulses, also known as action potentials, along the axon. In cardiac muscles, voltage-gated calcium channels release calcium ions to stimulate contraction of the muscles. Not surprisingly, the malfunctioning of ion channels is also the cause of many diseases, collectively known as channelopathies. These diseases include epilepsy, muscular disorders, cystic fibrosis and diabetes. [3] Knowledge of the structure and function of ion channels can help biologists understand the cause of and ultimately treat these diseases. Therefore, the study of ion channels is of fundamental importance in biology. In 2003, Rod-

erick MacKinnon's work in the determination of structure and operation of ion channels with crystallographic analyses won him a share of the Nobel Prize in Chemistry. [17, 18]

Ion channels have also attracted much interest in the area of artificial molecular detectors. Their selective conductivity and specific gating mechanisms make ion channels an ideal biological recognition component of a biosensor. In addition, ion channels are also highly sensitive; the binding of a single target molecule to an ion channel can lead to macroscopic level of conduction. As a result, ion channel based biosensors are designed to detect molecular species of interest across a wide range of applications that include medical diagnostics, environmental monitoring and general biohazard detection.

This thesis is concerned with statistical techniques used to describe ion channel permeation and to develop ion channel based biosensors. First we study the statistics of computer-simulated ion channel permeation and predict the channel behaviour in conditions where the simulation approach is computationally infeasible. In a separate study, we apply a statistical model to describe macroscopic output current recorded from an ion channel membrane-based biosensor. We show that the model, together with our detection algorithm, provides a sensitive and accurate biosensing platform. The reminder of this chapter is organized as follows. Sec. 1.1 introduces the computer simulation approach used to study ion channel permeation. Sec. 1.2 discusses some considerations for using gated ion channels to develop a biosensor. Finally, Sec. 1.3 discusses the organization for the rest of this thesis.

1.1 Permeation Model of Ion Channel

A permeation model describes the propagation of individual ions through an ion channel. With recent advances in computational power, one popular approach is to study the permeation of ion channels with computer simulations of channel interactions at an atomic level. These techniques include molecular dynamics (MD) and Brownian dynamics (BD). The general philosophy is to observe permeation properties of the channel in the simulations and to match with electrophysiological measurements.

In molecular dynamics, each atom in the channel protein, channel water and the environment near the channel mouth are modeled explicitly. At each discrete time step of the simulation, the forces on each atom are calculated and the location of each atom is adjusted accordingly. This approach computes the exact model of the channel, but the computational cost is very high, making it infeasible to use molecular dynamics to simulate ion conductions. However, molecular dynamics remains a useful tool to study different properties of ion channels. [4, 6]

Brownian dynamics is similar to molecular dynamics in that the movement of individual ions are calculated; however, it reduces computational complexity by treating water molecules in the channel as a continuum. The interaction between ions and the water molecules can be modeled as a random force and a frictional force. The charges of channel protein is assumed to be rigid. Also, BD simulations can operate at a much slower timescale than MD simulations. [39] Typical BD simulations are performed in the order of 100 fs per time step while MD simulations operate at 1 fs per time step. As

a result, BD has computational complexity several magnitudes lower than that of MD and is the preferred technique for simulating conduction events to estimate channel current.

Other theoretical approaches are used to study ion channel permeation. Reaction rate theory is based on partitioning of the channel into specific energy wells and estimating conductance from the transition probabilities between the wells. Poisson-Nernst-Planck theory model both ions and water molecules as continua and compute the dynamics using electrodiffusion theory and macroscopic electrostatics. These two techniques are not discussed in this thesis. For a detailed review of these techniques to model ion channel permeation, see [24, 31].

1.2 Gated Ion Channel Based Biosensor

Triggered by recent threats of biological warfare, there is a high demand for fast, economical and portable methods to detect the presence of harzardous biochemicals. Traditional laboratory techniques such as mass spectrometry and polymerase chain reaction (PCR) are not suited for this type of usage. One particular type of detector that is gaining popularity is a receptor-based sensor that performs biological recognition at the molecular level. Receptor-based sensor uses complementary molecules that bind to a specific type of target molecules and then indicate the result in the form of fluorescence [35], surface plasmon resonance [38] or other observable signals. In this part of the thesis, we focus on the ion channel based biosensor.

An ion channel based biosensor is typically built by inserting ion chan-

nels into artificial lipid membranes. The biosensor consists of a receptor element engineered to bind to a specific type of target molecules. The successful binding of the target molecule to the receptor element distorts the channel pore. This results in an alteration in the channel's conductance or gating mechanism. Detection of this change can be acheived by monitoring membrane conductivity, pH levels or intracellular concentration of ions. In addition, ion channels are highly sensitive devices and are suitable tools for signal amplification. The binding of a single target molecule to the receptor system can lead to an open channel current in the order of pico-Amperes.

1.3 Organization

The rest of this thesis is organized as follows. Chapter 2 studies some background information in ion channel permeation model and ion channel based biosensor. In particular, the technique of simulating ion channel conduction with Brownian dynamics is discussed in detail to identify the difficulty in running these simulations at high ionic concentration. Also, we summarize the development efforts in ion channel based biosensing and use a specific example of ion channel based biosensor to illustrate its functional properties. In Chapter 3, we introduce a statistical model to capture the binding site kinetics of an ion channel in Brownian dynamics simulation. The modeling, combined with extrapolation techniques, are used to predict the permeation behaviour of Brownian dynamics at high ionic concentration. Finally, in Chapter 4, we introduce the use of stochastic modeling and maximum likelihood detection results as part of a sensitive biosensing platform.

Chapter 2

Background

In this chapter, we study Brownian dynamics, a practical technique used to simulate ion channel conduction events, and investigate previous development efforts in the area of ion channel based biosensor. Through the discussion of theories and examples, we hope to motivate the use of statistical techniques to improve our understanding in ion channel permeation and to assist the development of sensitive and reliable biosensors.

Brownian dynamics have been widely used to determine the structure-function relationship of various ion channels, such as sodium channels [28, 40], potassium channels [26], calcium channels [14] and gramicidin channels [27]. These papers utlize different parameterization and optimization techniques to determine various structural properties of ion channels. In [27], the authors computed the potential of mean force experienced by a cation inside gramicidin channel. In [28], the authors parameterized and adaptively determined the molecular structure of sodium channels. The general approach in these papers is similar. The structural properties are determined by optimizing the fit between experimental current measured patch clamp techniques and simulated current estimated from Brownian dynamics.

Ion channels can perform chemical recognition at a molecular level and

the opening and closing of channels results in a macroscopic change in the conduction level. Thus the use of artificial gated ion channels for biosensing is a promising strategy. The majority of the development focuses on improving the sensitivity and stability of the sensor. That has prompted the use of a variety of ion channels and membranes to construct artificial biosensors. Multiple studies have shown the use of engineered pores to detect a variety of ions [9], proteins [32] and DNA [21].

The rest of this chapter is organized as follows. First, in Sec. 2.1 we discuss the permeation and conduction properties of gramicidin A (gA) channels. The algorithms presented in Ch. 3 and Ch. 4 are based on experimental and simulation results of these channels. In Sec. 2.2, we provide an overview of Brownian dynamics simulation of an ion channel. In particular, we describe the simulation setup, dynamic equations to determine ion movement, and estimation of current using Brownian dynamics. In Sec. 2.3, we provide a case study of a biosensor developed using gramicidin A channels. The specific example is used to identify the major components of an ion channel based biosensor and to highlight the development efforts in this area.

2.1 Gramicidin A Channel

Gramicidin A channels are used in antibotics to increase cation permeability across the bacterial cell wall, thus destroying the cell's ion gradient. In addition to its pharmaceutical use, gramicidin A is important as a model ion channel and is often used to test molecular theories of ion channel per-

meation. [1] For a recent review of gA channels, see [41].

Gramicidin A is a linear polypeptide consisting of fifteen amino acids. In its conducting form, two gramicidin A peptides are linked transiently by six hydrogen bonds to form a channel, which is 4 Å in diameter and 25 Å in length. The chemical structure of gramicidin A dimer is shown in Figure 2.1. The gating mechanism of gA channels results from the formation and breakdown of peptide bonds. Inside the gA channel exists six water molecules. During the permeation of an ion through the gA channel, the ion drags the water molecules in single file. The structure of dimerized gA channels in a lipid bilayer has been resolved by NMR studies. [25]

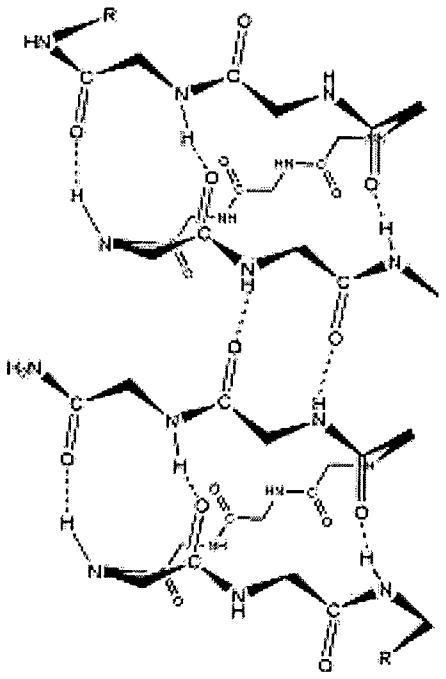


Figure 2.1: Chemical structure of gramicidin A dimers

Gramicidin A channels are selective for monovalent cations and impermeable to divalent cations and anions. NMR experiments [33] and MD simulations [5] show that there exists two monovalent binding sites at approximately ± 9Å from the center of the gA channel. The cation conductivity of gA channels varies linearly as a function of the bias voltage. The single channel current i_s depends on the cation concentration $[S]$ in the form of the Michaelis-Menten curve:

$$i_s = \frac{i_{\max}[S]}{[S] + K_S} \tag{2.1}$$

where $i_{\max}$ and K_S are parameter constants that depend on the cation type [20].

2.2 Brownian Dynamics

In this section, we summarize the use of Brownian dynamics to estimate channel current. Poisson's equations and *Langevin's* equations are solved iteratively to compute the movement of ions. Channel current can be estimated by counting ion crossing events. At the end of the chapter, we discuss some computational difficulties faced in BD simulations and show a simpler statistical approach to study Brownian dynamics.

The actual Brownian dynamics computer program belongs to the Computational Biophysics Group of Research School of Biological Sciences at the Australian National University (http://langevin.anu.edu.au/). For more detailed analysis on Brownian dynamics simulations, see [22, 23].

2.2.1 Simulation Setup

Our simulation setup consists of two reservoirs $\mathcal{R}_1$ and $\mathcal{R}_2$ connected by an ion channel $\mathcal{C}$. An external potential is applied across the channel. At the start of BD simulation, each reservoir contains N K^+ ions and N Cl^- ions indexed by $i = 1, 2, ..., N$ and $i = N+1, N+2, ..., 2N$ respectively. Let Λ denote the set of different experimental conditions, which represents varying external potentials across the channel and ion concentrations in the reservoirs. Let $\mathbf{x}_t^{(i)} \in \mathcal{R}^3$ and $\mathbf{v}_t^{(i)} \in \mathcal{R}^3$ denote the position and velocity of ion i at time t. $\mathbf{x}_t^{(i)} = (x_t^{(i)\prime}, y_t^{(i)\prime}, z_t^{(i)\prime})$, where $x_t^{(i)}$, $y_t^{(i)}$ and $z_t^{(i)}$ are the Cartesian components of the position vector. The velocity vector $\mathbf{v}_t^{(i)}$ follows a similar construct.

2.2.2 Solving Langevin Equations for Ion Dynamics

The position and velocity of each ion evolve according to the *Langevin* equations. Let $\mathbf{X}_t = \left(\mathbf{x}_t^{(1)\prime}, \mathbf{x}_t^{(2)\prime}, \ldots, \mathbf{x}_t^{(2N)\prime}\right)'$ denote the positions and $\mathbf{V}_t = \left(\mathbf{v}_t^{(1)\prime}, \mathbf{v}_t^{(2)\prime}, \ldots, \mathbf{v}_t^{(2N)\prime}\right)'$ denote the velocities of all $2N$ ions in our setup.

$$\mathbf{x}_t^{(i)} = \mathbf{x}_0^{(i)} + \int_0^t \mathbf{v}_s^{(i)} ds, \tag{2.2}$$

For $i = 1, 2, \ldots, N$,

$$\begin{aligned} m^+\mathbf{v}_t^{(i)} = m^+\mathbf{v}_0^{(i)} &- \int_0^t m^+\gamma^+(\mathbf{x}_s^{(i)})\mathbf{v}_s^{(i)} ds \\ &+ \int_0^t F_\lambda^{(i)}(\mathbf{X}_s) ds + 2m^+\gamma^+(\mathbf{x}_t^{(i)})\mathbf{w}_t^{(i)} \end{aligned} \tag{2.3}$$

For $i = N+1, N+2, \ldots, 2N$,

$$m^- \mathbf{v}_t^{(i)} = m^- \mathbf{v}_0^{(i)} - \int_0^t m^- \gamma^-(\mathbf{x}_s^{(i)}) \mathbf{v}_s^{(i)} ds + \int_0^t F_\lambda^{(i)}(\mathbf{X}_s) ds + 2m^- \gamma^-(\mathbf{x}_t^{(i)}) \mathbf{w}_t^{(i)} \tag{2.4}$$

$F_\lambda^{(i)}(\mathbf{X}_t)$ is the systematic forces acting on ion i for experimental condition $\lambda \in \Lambda$. Details about the systematic forces are given in Sec. 2.2.3. The frictional coefficient $m^\pm \gamma^\pm(\mathbf{x}_s^{(i)}) = m^\pm \gamma^\pm = \frac{kT}{D^\pm}$, if $\mathbf{x}_s^{(i)} \in \mathcal{R}_1 \cup \mathcal{R}_2$ and the diffusion coefficient of K^+ is $D^+ = 1.96 \times 10^{-9}$ m^2/s in a bulk solution, while the diffusion coefficient for Cl^-ions is $D^- = 2.03 \times 10^{-9}$ m^2/s. The process $\mathbf{w}_t^{(i)}$ is a 3-dimensional Brownian motion with component-wise independence and can be written as a zero-mean Gaussian random variable with a 3×3 diagonal covariance matrix.

2.2.3 Solving Poisson's Equation for Systematic Forces

The systematic forces in (2.3) and (2.4) can be rewritten as:

$$F_\lambda^{(i)}(\mathbf{X}_t) = -q^{(i)} \nabla_{\mathbf{x}_t^{(i)}} \Phi_\lambda^{(i)}(\mathbf{X}_t) \tag{2.5}$$

where the scalar valued process $\Phi_\lambda^{(i)}(\mathbf{X}_t)$ is the total electric potential experienced by ion i given the position $\mathbf{X}_t$ of the $2N$ ions. The potential $\Phi_\lambda^{(i)}(\mathbf{X}_t)$ experienced by each ion i consists of five components:

$$\Phi_\lambda^{(i)}(\mathbf{X}_t) = \Phi_\lambda^{C,i}(\mathbf{X}_t) + \Phi_\lambda^{\text{ext}}(\mathbf{x}_t^{(i)}) + \Phi_\lambda^{IW}(\mathbf{x}_t^{(i)}) + \Phi_\lambda^{SR,i}(\mathbf{X}_t) + U(\mathbf{x}_t^{(i)}) \tag{2.6}$$

$\Phi_\lambda^{\text{ext}}(\mathbf{x}_t^{(i)})$ denotes the external potential applied along the z axis of the ion channel. $\Phi_\lambda^{IW}(\mathbf{x}_t^{(i)})$ denotes the ion-wall interaction potential, also known as the σ/r^9 potential. $\Phi_\lambda^{C,i}(\mathbf{X}_t)$ denotes the inter-ion Coulomb potential and $\Phi_\lambda^{SR,i}(\mathbf{X}_t)$ denotes the short range ion-ion potential. $U(\mathbf{x}_t^{(i)})$ denotes the potential of mean force (PMF), which is a smooth function used to model the Coulomb interaction between the ion and channel protein as well as induction of surface charges at the water-protein interface when the ion is near the protein wall. For the simulations discussed in this thesis, we have used the PMF that has the best fit to experimental results of gramicidin A current measurements. Details of this result can be found in [29].

The systematic forces $F_\lambda^{(i)}(\mathbf{X}_t)$ are calculated as a function of the ion positions. This is done by solving Poisson's equation. By assuming that the simulation space can be partitioned into several regions with different dielectric constants. Let ϵ_j be the dielectric constant of region j and we can write the modified Poisson's equation as:

$$\nabla^2 \Phi_j = \frac{-\rho}{\epsilon_j \epsilon_0} \tag{2.7}$$

2.2.4 Estimating Brownian Dynamics Current

The computer simulation of ion channel implements a discretized version of the dynamics described in Sec. 2.2.2 and Sec. 2.2.3. Suppose the channel model for experimental condition λ is simulated for L seconds. At each discrete time step, the *Langevin's* equations in (2.2), (2.3) and (2.4) and modified Poisson's equation in (2.7) are solved to compute the movement of ions. Two crossing events are of importance to the estimation of channel

conductance: 1) forward crossing event, during which an ion crosses from $\mathcal{R}_1$ to $\mathcal{R}_2$, and 2) backward crossing event, during which an ion crosses from $\mathcal{R}_2$ to $\mathcal{R}_1$. Let $N_{\mathcal{R}_1,\mathcal{R}_2}$ and $N_{\mathcal{R}_2,\mathcal{R}_1}$ be the number of forward and backward crossing events respectively. We can compute the Brownian dynamics estimate of the single channel current:

$$\hat{i}_\lambda = \frac{q^+(N_{\mathcal{R}_1,\mathcal{R}_2} - N_{\mathcal{R}_2,\mathcal{R}_1})}{L} \tag{2.8}$$

where q^+ is the electric charge of a proton, 1.6×10^{-19} C. The current estimate in (2.8) is based on the random samples observed in the simulation and therefore cannot be represented by a closed form expression.

2.2.5 Motivation for Statistical Modeling

Brownian dynamics is a very useful tool to study the structure-function relationship of ion channels. The technique simulates ion channel interaction at an atomic level and is computationally tractable to use BD to simulate channel crossing events because of a few simplifying assumptions discussed in Sec. 1.1. As a result, biophysicists often use BD to estimate single channel current at various applied voltages and concentrations and compare with experimental patch-clamp measurements.

Although BD is generally computationally tractable, obtaining an accurate current estimates at multiple voltages and concentrations can take a very long time. Increasing ionic concentration involves placing additional ions in both reservoirs. That has a quadratic effect on the computational

complexity of the BD simulations. Fig. 2.2 shows the time required to finish 1 μs of BD simulation of gramicidin A channel. The simulation is performed on the supercomputer cluster on Westgrid (glacier.westgrid.ca).

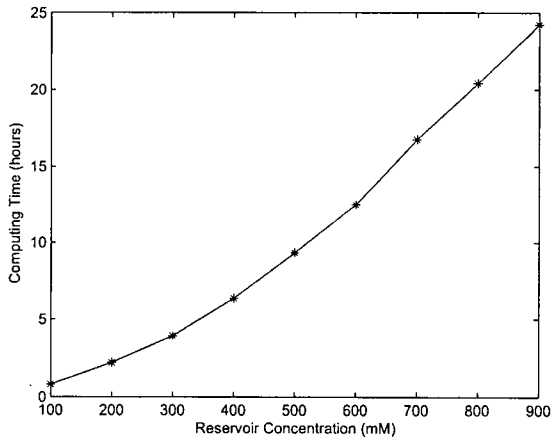


Figure 2.2: Computing time of 1 μs of Brownian dynamics simulation of gramicidin A channel.

It can be seen that the computing time increases dramatically as more ions are included in the simulation. To obtain an accurate estimate of the channel current, typical BD simulations of an ion channel last 10 μs to 100 μs [11]. Generally it is rare to see BD simulations with reservoir concentrations above $1M$.

In Chapter 3, we construct a statistical fully-parameterized model to characterize ion channel permeation observed in Brownian dynamics. Instead of modeling movements of all ions in the system, we focus on the kinetics of ion activities inside the channel. As a result, the model is a much simpler representation of Brownian dynamics and we show that conduction

statistics are completely recoverable. One useful application of this model is model prediction. We can apply extrapolation techniques to predict channel current at conditions where BD simulation is computationally intractable.

2.3 Previous Work on Ion Channel Biosensor

To better illustrate the biosensing problem, we investigate previous work in the area of ion channel based biosensor. In particular, we describe the composition and detection mechanism of the Ion Channel Switch (ICS^{TM}) biosensor. [2, 13] The gating mechanism of the novel biosensor exploits the association and disassociation probabilities of the gramicidin A dimers. This flexible and adaptive sensor can detect a variety of molecules such as growth factors, glucose, and DNA strands. [11]

2.3.1 Composition of ICS^{TM} Biosensor

Figure 2.3 shows a diagram of the major components in the ICS^{TM} sensor. The ICS^{TM} biosensor is built by incorporating gramicidin A monomers into a tethered lipid bilayer. Because of their chemical and structural stability, gA channels are often used in the development of biosensors. The gA monomers in the inner leaflet of the bilayer are immobile while the monomers in the outer leaflet are free to diffuse. As discussed earlier in the chapter, gA molecules can only conduct when two gA monomers dimerize to form a channel. Therefore, the conductance of the lipid bilayer arises from random diffusion of free-moving monomers that dimerize to tethered monomers in the inner leaflet.

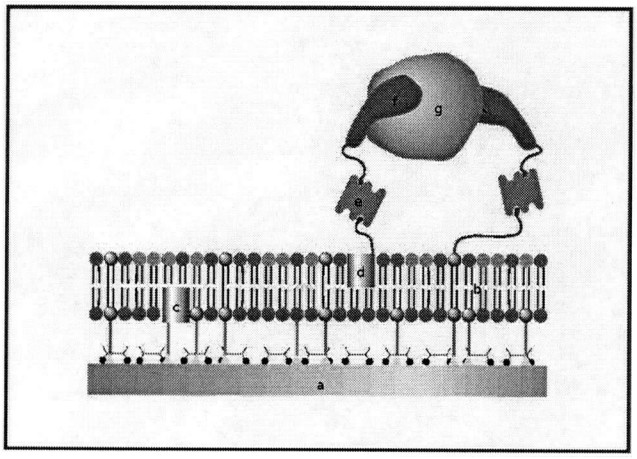
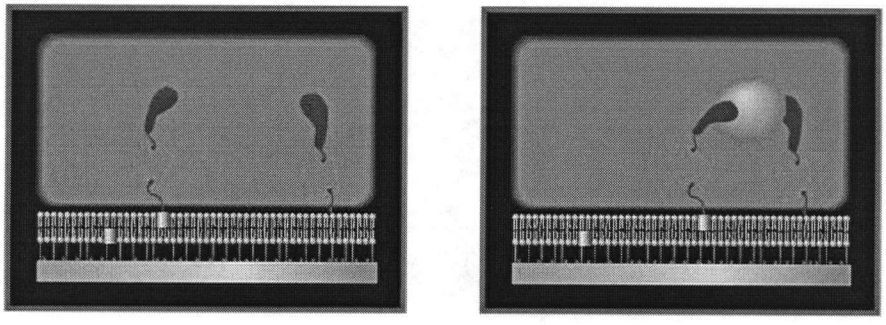


Figure 2.3: Components of the ICS^{TM}biosensor consists of (a) gold electrode and (b) tethered lipid bilayer. (c) Immobile gA monomer is embedded in the inner leaflet of the bilayer and (d) free-moving gA monomer embedded in the outer leaflet. In this setup, the receptor system, made of (e) streptavidin and (f) biotinylated receptor, is used to detect a specific (g) target molecule. [11]

molecules with matching antibodies.



(a) (b)

Figure 2.4: Detection mechanism of ICS^{TM}biosensor. (a) When no target molecules are present, the gA monomer in the outer leaflet is free to diffuse and couple with immobile monomers in the inner leaflet. (b) When target molecule is present, the receptor system binds to the target and the attaching gA monomer is no longer mobile. [11]

2.3.3 Motivation for Statistical Detection

The above example of a biosensor illustrates how an artificial ion channel based system can be designed to detect the presence of target molecules. Other ion channel biosensors have employed different gating mechanisms, such as channel blockage by molecules or channel distortion using antichannel antibodies. In general, the majority of the research focuses on achieving large differences in the biosensor conductance with and without the presence of target molecules. Little research has been done on extracting information from the biosensor's response using signal processing techniques.

In Ch. 4, we propose the use of statistical modeling and maximum likelihood detection algorithms as part of a biosensing platfom. The algorithm

offers an automatic approach to analyze the results from the biosensor in real-time. This approach eliminates the majority of electrical and mechanical interfering effects and offers the ability to enhance the specificity and affinity of the biosensor.

Chapter 3

A Simple Statistical Model to Characterize Brownian Dynamics Permeation of Gramicidin A Channels

Brownian dynamics have been widely used to determine structure and function of various ion channels. The general approach is to optimize the fit between the channel current measured with patch clamp techniques and the one estimated from BD simulations at various external voltages and ionic concentrations. However, as discussed in Sec. 2.2.5, the computational complexity of Brownian dynamics makes it computationally intractable to perform the simulation at high concentrations.

We propose a simpler statistical model to characterize ion permeation simulated in Brownian dynamics. In particular, we use the simple gramicidin A channel to demonstrate the feasibility of this approach. We model only the binding site statistics of gA channel with a finite-state semi-Markov process. To evaluate the performance of this model, we compare conduction statistics

of the proposed model with that obtained from Brownian dynamics simulation. The model captures dwell time and transition statistics with a small number of parameters and enables us to apply extrapolation techniques to predict the behaviour of Brownian Dynamics at high concentrations.

The rest of this chapter is organized as follows. In Sec. 3.1, we map the continuous simulation space into finite states of binding site occupation and fit it to a three-state semi-Markov chain (SMC). In Sec. 3.2, we apply extrapolation techniques to predict channel current at high concentrations where performing the actual simulation is computationally intractable. Finally, numerical results of the modeling and prediction of gA channels are presented in Sec. 3.3.

3.1 Statistical Modeling of Ion Channel Permeation

In this section, we use statistical modeling techniques to capture ion dynamics inside the channel. The aim is to develop a simple model that describes only the binding site statistics. By mapping the continuous space into discrete states of the channel, we construct and estimate a finite-state semi-Markov chain that is statistically indistinguishable from sample paths of actual BD simulations. Statistical validation methods are carefully applied at various stages of the modeling to verify goodness-of-fit.

3.1.1 Finite State Representation

A Brownian dynamics simulation is extremely complex, consisting of $2N$ ions each with its own position and velocity vectors. To reduce the dimensionality of the space, we choose to only model the occupancy of the binding sites in the gramicidin A channel. Figure 3.1 shows the shape of the channel. The shaded regions $-z_2 \leq z \leq -z_1$ and $z_1 \leq z \leq z_2$ are the left and right binding sites respectively. $z = -z_3$ and $z = z_3$ denote the reservoir-channel boundaries of the simulation space. "Transition regions" are defined in the regions $-z_3 \leq z \leq -z_2$, $-z_1 \leq z \leq z_1$ and $z_2 \leq z \leq z_3$ to smooth out unimportant events caused by ions bouncing around near the region boundaries. Only the z-coordinates are considered because K^+ions can only pass through the narrow gA channel in single-file.

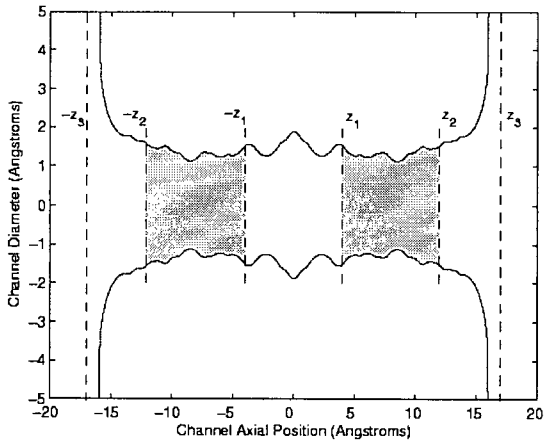


Figure 3.1: Shape of gramicidin A channel and location of binding sites.

Figure 3.2 shows a sample path of BD simulated ion movements. Ions

inside the channel spend the majority of the time in the monovalent binding sites centered at around ± 9Å. The existence of the binding sites are discussed in Sec. 2.1. We argue that since the ions spend negligible time elsewhere in the channel, it is sufficient to model the dwell time and transition statistics in the binding site regions.

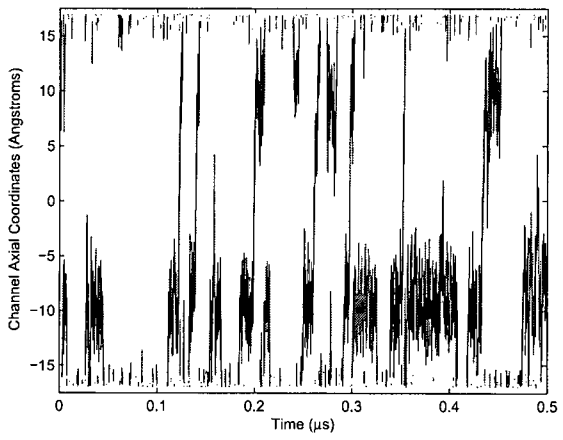


Figure 3.2: Sample path of BD-simulated ion movement inside gramicidin A channel. The sequence clearly shows the ions spending most of the time in the binding sites. The color of the trace is used to differentiate between ions currently in the channel and does not belong to the same ion for the entire sequence.

Let $\mathcal{L} \in \{0, 1\}$ and $\mathcal{R} \in \{0, 1\}$ denote the occupancy of the left and right binding site respectively, where 0 indicates an empty binding site and 1 an occupied binding site. When an ion is within a transition region, $\mathcal{L}$ and $\mathcal{R}$ simply assume their previous values.

Due to the close proximity of the two binding sites, it is extremely rare to have both sites occupied for a long period of time. For example, in Figure

3.2, an ion (in red) enters the left binding site at 0.21 μs and both binding sites are occupied concurrently. However, the presence of the left ion forces the right ion (in blue) out of the channel soon after.

To further illustrate this scenario, we observe the change in potential energy when a K^+ion is at various locations near the left binding site. In Figure 3.3 we plot only the energy due to the PMF and inter-ion Coulomb potential. The other three components in (2.6) are ignored for simplicity. The presence of an ion at the left binding site reduces the energy well from 6 kT to 2 kT.

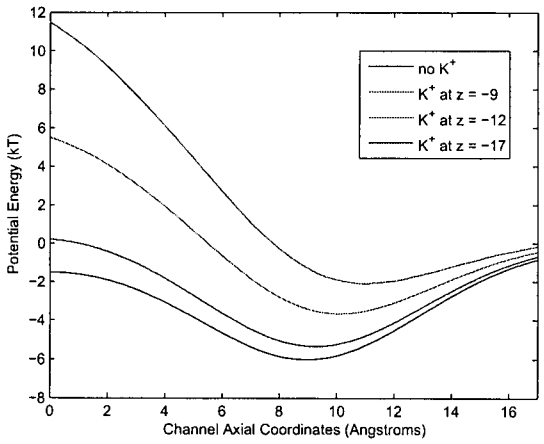


Figure 3.3: Effect of binding K^+ion on the potential energy of gramicidin A channel.

The energy E is related to the occupation probability p as follows:

$$p \propto \exp(-E) \tag{3.1}$$

Based on the above relationship, the probability of an ion binding to the right site is 55 times smaller when the left site is occupied than when it is empty. We conclude that the state $\mathcal{L} = \mathcal{R} = 1$ is very rare and treat it a transition state. Let the state space be $\mathcal{S} \in \{00, 10, 01\}$, where the first and second digit represent $\mathcal{L}$ and $\mathcal{R}$ respectively.

3.1.2 Semi Markov Model of Ion Channel Permeation

In Sec. 3.1.1, we showed the mapping from all the ion positions $\mathbf{X}_t$ to three finite states, representing the occupancy of the binding sites in the channel. In Figure 3.4, we plot the finite state mapping of the ion sample paths observed in Figure 3.2.

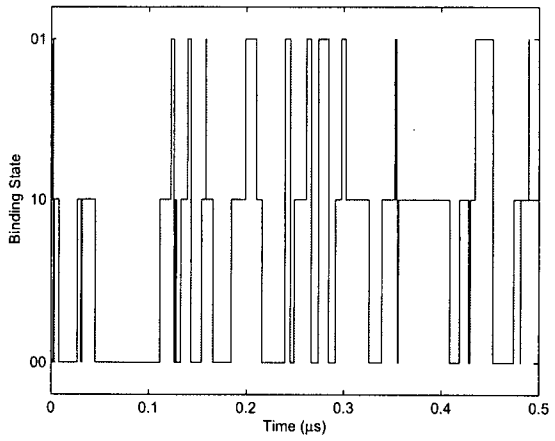


Figure 3.4: Finite state mapping of ion sample paths observed in Fig. 3.2. The smoothing technique results in crisp transitions that adequately describes the state of the binding sites.

Modeling and Estimation of SMC

In this section, we introduce the semi-Markov chain, a probabilistic model used to capture the transition probabilities and rates of a finite state model. SMC, a popular technique used in reliability engineering and lifetime analysis, is parameterized by state space $\mathcal{S}$, transition matrix $\mathbf{P}$ and state holding time matrix $\mathbf{Q}$. Here is a summary of the binding site kinetics that result in a state transition:

- $00 \rightarrow 10$

  An ion from the left reservoir binds to the left site.

- $10 \rightarrow 00$

  An ion from the left site enters the left reservoir.

- $10 \rightarrow 01$

  (a) An ion from the left site jumps to the right site.

  (b) An ion from the right reservoir binds to the right site and pushes the ion at the left site out of the channel.

- $00 \rightarrow 01$

  An ion from the right reservoir binds to the right site.

- $01 \rightarrow 00$

  An ion from the right site enters the right reservoir.

- $01 \rightarrow 10$

  (a) An ion from the right site jumps to the left site.

(b) An ion from the left reservoir binds to the left site and pushes the ion at the right site out of the channel.

Note that both $10 \rightarrow 01$ and $01 \rightarrow 10$ can result from one of two possible processes. For the rest of this chapter, we refer to process (a) as the jumping process and process (b) as the pushing process. Let β^{23} denote the probability that $10 \rightarrow 01$ event is due to the jumping process and $(1 - \beta^{23})$ the probability of the pushing process. Similarly, let β^{32} denote the probability that the $01 \rightarrow 10$ event is due to the jumping process and $(1 - \beta^{32})$ the probability of the pushing process. To simplify the notations, we rewrite state 00, 10 and 01 to be state 1, 2 and 3 respectively. Let p_{ij} denote the transition probability from state i to state j, $i, j = 1, 2$ and 3. The transition matrix $\mathbf{P}$ can be written as:

$$\mathbf{P} = \begin{bmatrix} 0 & p_{12} & p_{13} \\ p_{21} & 0 & p_{23} \\ p_{31} & p_{32} & 0 \end{bmatrix} \tag{3.2}$$

Similarly we can write $\mathbf{Q}$ as a matrix of state holding time distributions.

$$\mathbf{Q} = \begin{bmatrix} -- & q_{12}(t) & q_{13}(t) \\ q_{21}(t) & -- & q_{23}^{\beta^{23}}(t) \\ q_{31}(t) & q_{32}^{\beta^{32}}(t) & -- \end{bmatrix} \tag{3.3}$$

where $q_{ij}(t)$ denote the probability distribution function (PDF) of the state holding times from state i to state j, $i, j = 1, 2$ and 3. The distributions $q_{11}(t)$, $q_{22}(t)$ and $q_{33}(t)$ are undefined because state transitions in a semi-Markov model must have different start and end states. The model is complicated by the fact that the transitions $10 \rightarrow 01$ and $01 \rightarrow 10$ can arise from two different processes and therefore possess different PDFs. For the transition $10 \rightarrow 01$, let $q_{23,\text{jump}}(t)$ denote the PDF for the jumping process and $q_{23,\text{push}}(t)$ the PDF for the pushing process. We can write $q_{23}^{\beta_{23}}(t)$ as:

$$q_{23}^{\beta^{23}}(t) = \begin{cases} q_{23,\text{jump}}(t) & \text{with probability } \beta^{23} \\ q_{23,\text{push}}(t) & \text{with probability } (1-\beta^{23}) \end{cases} \tag{3.4}$$

$q_{32,\text{jump}}(t)$, $q_{32,\text{push}}(t)$ and $q_{32}^{\beta_{32}}(t)$ can be defined similarly. As a result, $\mathbf{Q}$ is a 3×3 matrix of probability distributions, of which two distributions are Bernoulli-modulated. Experimentally we found that the gamma distribution provided the best fit to the state holding time distributions. Other PDFs considered include exponential, normal, log-normal and Weibull distribution. The gamma distribution is defined as:

$$q(t|a,b) = \frac{1}{b^a\,\Gamma(a)} t^{a-1} e^{-\frac{t}{b}}, \tag{3.5}$$

where a and b are known as the shape and scale parameter and $\Gamma(\cdot)$ is the gamma function. The estimation of the holding times as a gamma distribution allows us to model each state holding time distribution with only two parameters. Suppose n different transition times of a particular transition

are observed in a BD simulation. Let t_i be the i^{th} largest state holding time. The ordering of the holding times is irrelevant for the estimation but is important for the statistical validation method to be introduced later. The maximum-likelihood estimates for the gamma distribution can be computed by solving the following set of equations:

$$\frac{\sum_i^n \log t_i}{n} - \log \bar{t} = \psi(\hat{a}) - \log(\hat{a}) \tag{3.6}$$

$$\hat{b} = \frac{\bar{t}}{\hat{a}} \tag{3.7}$$

where $\bar{t}$ is the mean of the random samples and $\hat{a}$ and $\hat{b}$ are the maximum likelihood estimated shape and scale parameter.

Statistical Validation of SMC

The Anderson-Darling statistics is the standard technique used to evaluate the goodness of fit in distribution fitting. The Anderson-Darling statistics is part of a large class known as the empirical distribution function (EDF) statistics, which measures the difference between the estimated distribution and the empirical distribution of the random samples:

$$A^2 = -n - \frac{1}{n}\sum_i^n (2i-1)\log z_i + (2n+1-2i)\log(1-z_i) \tag{3.8}$$

where $z_i = Q(t_i|\hat{a},\hat{b})$ is the cumulative distribution function (CDF) of the estimated gamma distribution. Table 3.1 shows the Anderson-Darling statis-

tics computed at concentration 100 mM, 400 mM and 800 mM.

(a)

–	0.154	0.669
0.994	–	0.608, 0.916
0.779	0.256, 0.343	–

(b)

–	0.250	0.506
1.237	–	0.898, 0.513
2.054	0.483, 0.219	–

(c)

–	1.277	0.381
1.537	–	0.636, 1.717
0.940	0.727, 0.456	–

Table 3.1: Anderson Darling statistics of holding time distributions at ionic concentration (a) 100 mM, (b) 400 mM and (c) 800 mM. The statistics for $q_{ij}(t)$ are shown in cell (i,j) of each table. Multiple entries exist in cell (2,3) and (3,2) because separate statistics are computed for the jumping (left) and pushing (right) process.

It can be seen that the Anderson Darling statistics for most state holding time distributions fall below the critical value of 1.092 found in [15]. In fact, 81% of the transition times satisfied the Anderson-Darling critical value for the gamma distribution at 0.01 significance level. Out of all transitions in Table 3.1, the holding time distribution $q_{31}(t)$ at concentration of 400 mM showed the biggest deviation from the gamma distribution. The empirical CDF of $q_{31}(t)$ and the estimated gamma CDF are compared in Figure. 3.5. It can be seen that the estimated gamma CDF followed the empirical CDF of the holding times closely even though it failed the Anderson Darling test. Thus it is a reasonable approximation to estimate all the holding times as

gamma distributions.

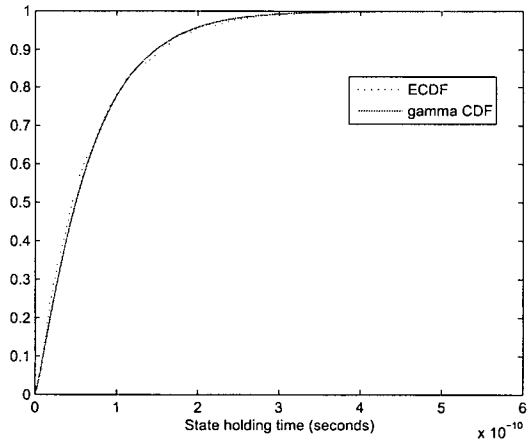


Figure 3.5: Comparision between ECDF and estimated gamma CDF for $q_{31}(t)$ at ionic concentration 400 mM.

Once the transition matrix $\mathbf{P}$ and the parameters in the state holding time matrix $\mathbf{Q}$ are estimated, a Monte-Carlo simulation of the SMC produces an estimate of the channel current.

3.2 Extrapolation Technique for Model Prediction

In this section, we devise an extrapolation procedure that can predict the behaviour of Brownian dynamics. In particular, we are interested in predicting the channel current at high ionic concentration because BD is intractable at that region. It is important to choose a reasonable functional form for the extrapolation to reduce the effect of outliers and estimation errors. General

extrapolation such as cubic spline leads to overfitting. Here we make two reasonable assumptions about the effect of increasing ionic concentrations:

1. Increasing ionic concentration has a monotonic effect on some transition probabilities and dwell times. For example, an empty binding site is likely to be filled quickly when more ions are in the reservoirs. Other transitions may be unaffected by concentration.

2. The effect of increasing ionic concentration saturates at very high concentrations.

Linear extrapolation is monotonic but its simple functional form does not model saturation very well. Polynomial extrapolation of second order or higher is much more flexible but does not guarantee monotonicity. To satisfy the above assumptions, we choose the four-parameter logistic function:

$$f(x) = A_0 + \frac{A_\infty - A_0}{1 + k\, e^{-\frac{x}{d}}} \tag{3.9}$$

where A_0, A_∞, k and d are the function parameters. Fig. 3.6 plots an example of the logistic function, commonly used in biology and economics. In addition to its monotonicity, the logistic function is bounded from above or below as $x \to \infty$. The function is flexible enough to model increasing, decreasing and constant functions. The effect of increasing ionic concentration on the transition probabilities and state holding times is unknown and it seems reasonable to use the logistic function as the basis to extrapolate these parameters.

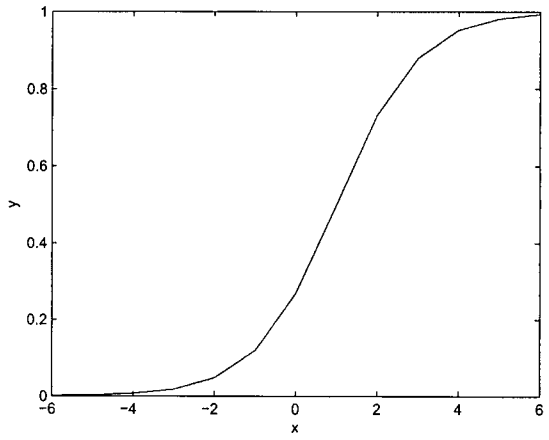


Figure 3.6: Example of logistic function, with $A_0 = 0$, $A_\infty = 1$, $k = 1$ and $d = 1$.

To perform the model prediction, permeation properties of the gA channel at multiple reservoir concentrations are simulated using Brownian dynamics. Each set of simulation result is mapped to the three-state semi-Markov process described in Sec. 3.1.2. The estimated parameters in transition probabilities P and holding time matrix Q are then fitted to the logistic function using nonlinear least square methods such as the Gauss-Newton method. Finally, the parameters of the predicted model can be extrapolated from the fitted logistic function.

3.3 Numerical Results

In this section, we present some numerical results based on BD simulations of the gA channel and evaluate the performance of our modeling and prediction techniques. We simulated the ion channel at an external voltage of 100 mV. A total of nine independent simuations were performed with the reservoir concentration varying from 100 mM to 900 mM, in increments of 100 mM. The duration of each simulation is 8 μs.

3.3.1 Parameter Estimation for Binding Site Kinetics

Table 3.2 shows the estimated transition probabilities, which nicely captures statistical changes of the the binding site kinetics at various ionic concentrations.

For example, P(1,2) and P(1,3) does not vary with increasing concentration. With concentration in both reservoirs being equal, it is reasonable to see that the ratio of ion entry from the left and right reservoir is constant.

(a)

0	0.639	0.361
0.811	0	0.189
0.879	0.121	0

(b)

0	0.639	0.361
0.715	0	0.285
0.723	0.277	0

(c)

0	0.638	0.361
0.622	0	0.378
0.616	0.384	0

(d)

0	0.633	0.367
0.607	0	0.393
0.581	0.419	0

(e)

0	0.647	0.353
0.582	0	0.418
0.527	0.473	0

Table 3.2: Estimated transition matrices at ionic concentration (a) 100 mM, (b) 400 mM, (c) 700 mM, (d) 800 mM and (e) 900 mM. The probability for p_{ij} are shown in cell (i,j) of each table.

Furthermore, as the concentration increases, it is less likely for an ion in the binding site to leave and go back into the reservoir. Consequently, both P(2,1) and P(3,1) are monotonically decreasing.

The effect of increasing concentration on several holding time distribution can be seen in Figure 3.7. In particular, the biggest effect of increasing concentration is the decrease in state holding time $q_{12}(t)$. An empty left binding site is filled much more quickly as the left reservoir becomes more crowded with ions.

3.3.2 Model Prediction at High Concentration

To demonstrate the feasibility of our extrapolation technique, we predict the channel current at 800 mM and 900 mM, using simulation results from

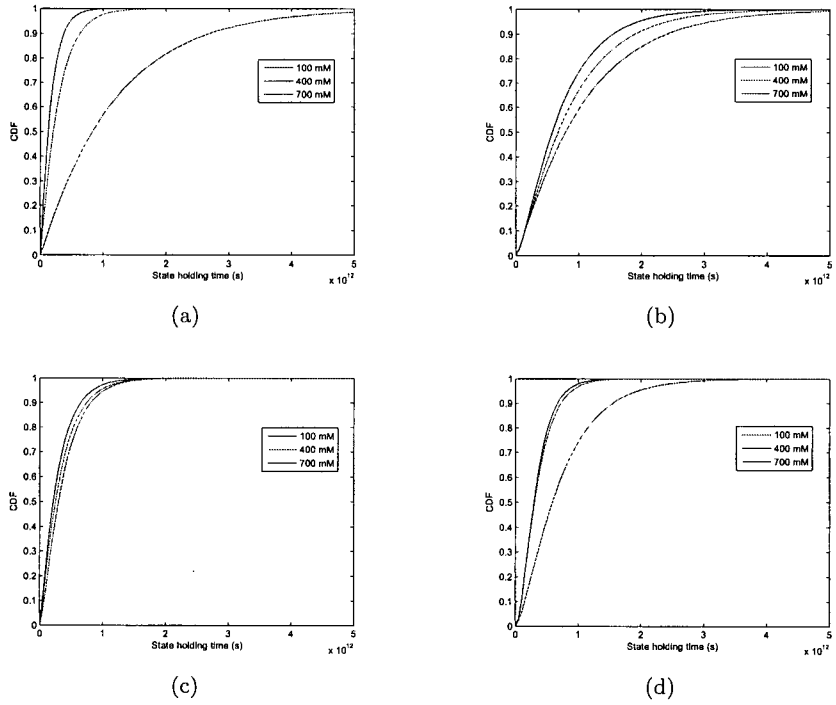


Figure 3.7: Estimated state holding time distributions (a) $q_{12}(t)$, (b) $q_{23,\text{jump}}(t)$, (c) $q_{31}(t)$ and (d) $q_{32,\text{push}}(t)$ at ionic concentrations of 100 mM, 400 mM and 700 mM.

(a)

0	0.636	0.364
0.591	0	0.409
0.601	0.399	0

(b)

0	0.636	0.364
0.568	0	0.432
0.576	0.424	0

Table 3.3: Extrapolated transition matrices at ionic concentration (a) 800 mM and (b) 900 mM. The probability for p_{ij} are shown in cell (i,j) of each table.

100 mM to 700 mM. This allows us to match BD-simulated and SMC-extrapolated results at 800 mM and 900 mM and evaluate the performance of our model. The first seven BD simulations, at ionic concentration 100 mM to 700 mM, are estimated and the SMC parameters at each concentration are fitted to the logistic function as a function of concentration. The extrapolation is used to determine the SMC parameters at 800 mM and 900 mM. The extraporated transition matrix is shown in Table 3.3.

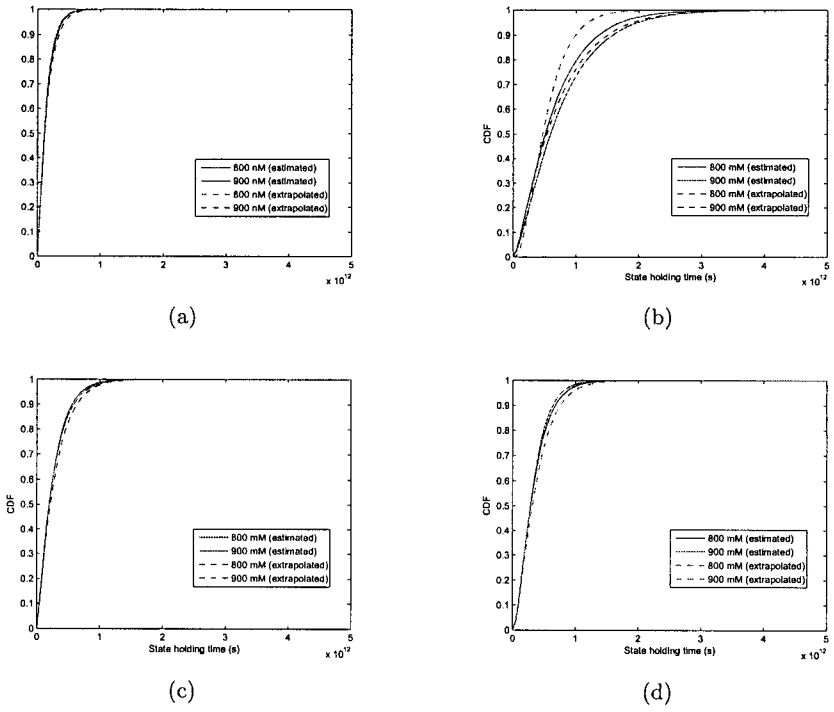


Figure 3.8: Comparison between estimated and extrapolated state holding time distributions (a) $q_{12}(t)$, (b) $q_{23,\text{jump}}(t)$, (c) $q_{31}(t)$ and (d) $q_{32,\text{push}}(t)$ at ionic concentrations of 800 mM and 900 mM.

In Figure 3.8, we compare the extrapolated distribution with results

estimated from BD simulations. With the exception of $q_{23,\text{jump}}(t)$, the extrapolated distributions are very similar to the estimated distributions. One possible explanation of the poor performance in extrapolating $q_{23,\text{jump}}(t)$ is the lack of observations of that transition. The $10 \rightarrow 01$ jumping process is one of the rarest transitions in BD simulations of gA channels; therefore, the estimation of this process may have higher variance than other transitions.

The extrapolated SMC model at 800 mM and 900 mM are simulated for 1000 independent runs. During each 8 μs run, the ion movements into and out of the binding sites are simulated not with Poisson's and *Langevin's* equations, but instead with the extrapolated SMC models. Figure 3.9 shows the result of the simulations.

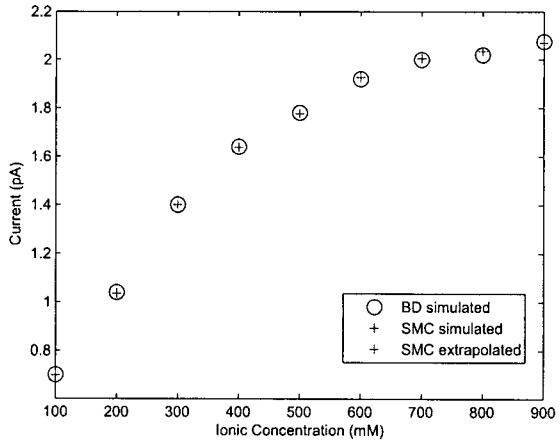


Figure 3.9: Comparison between SMC-simulated and BD estimated currents show similar conduction levels.

At the concentrations 100nM to 700nM, it is not surprising that the BD-simulated and SMC-simulated current have similar conduction level. The

SMC model was estimated directly from random paths observed in Brownian dynamics. The extrapolation at 800 mM and 900 mM also performed extremely well. Clearly the SMC-extrapolated exhibits very similar levels as the ones estimated with Brownian dynamics. Using the two-sample Kolmogorov-Smirnov test, we compare the similarities of the conduction distribution simulated in BD and SMC. The Kolmogorov-Smirnov test is another type of EDF testing techniques and can be used to test general distributions. The tradeoff for this generality is that the test is less effective than other EDF testing techniques designed for specific distributions. [10] Out of 1000 independent simuations at the concentration 800 nM, 96.5% of the time the Komogorov-Smirnov test requirement were satisfied, with an average P-value of 0.5295. At 900 nM, 97.8% of the test passed with an average P-value of 0.5069. Thus the hypothesis that the BD-simulated and SMC-simulated conduction events share the same distribution cannot be rejected.

At this point, we do not have simulation results beyond ionic concentration of 900 mM and therefore cannot test our modeling and prediction technique at higher concentrations. We conclude this section with a prediction of channel current up to 5 M. These estimates are also obtained by simulating the extrapolated model for 1000 independent runs. The extrapolation results are plotted in Figure 3.10.

Since there are no BD results to compare with our extrapolated current, we compare our results with the Michaelis-Menten kinetics described in 2.1. The plot shows that the SMC technique predicts that BD estimated current saturates at a lower conduction level than that of the Michaelis-Menten

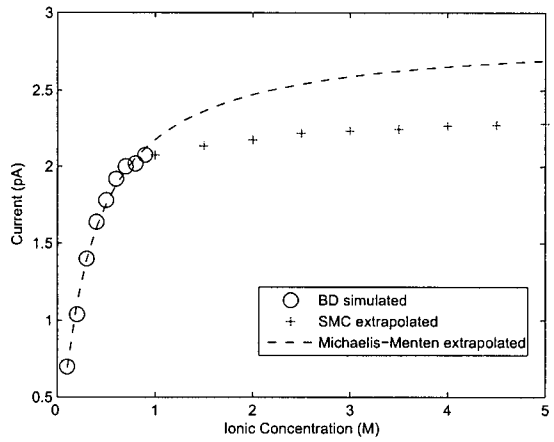


Figure 3.10: Comparison between SMC-extrapolated results and best-fit Michaelis-Menten curve. The comparison shows that the SMC extrapolation expects the current to saturate at a lower conduction level than that of the Michaelis-Menten curve.

curve.

Chapter 4

Bis-gA Channel Biosensor

In this chapter, we discuss the construction, statistical modeling and detection algorithm of a new type of ion channel based biosensor. The biosensor constructed comprises engineered dimeric gramicidin A (bis-gA) ion channels incorporated into a lipid bilayer membrane, supported over a 1-micron diameter opening of a micropipette, which was excised from a giant lipid vesicle. In a giant lipid vesicle, covalent dimeric gramicidin A ion channels were incorporated by codispersion with the vesicle forming lipids. This type of artificially constructed biosensor mimics the naturally occurring ion transport processes of a living cell. The biosensor is constructed by Bruce Cornell in AMBRI Ltd. and Donald Martin and his students in the Department of Medical and Molecular Biosciences in the University of Technology, Sydney.

Having constructed the biosensor, we formulate a stochastic dynamical model to capture its experimental behaviour. The engineered dimeric gA channels provided the conducting pore for the vesicle membrane, but not with the kinetics of the ICS^{TM}sensor discussed in Sec. 2.3. Instead, the gating mechanism of the dimeric gA ion channels in this biosensor is thought to arise from random movement of excess lipid lenses in the liposome that

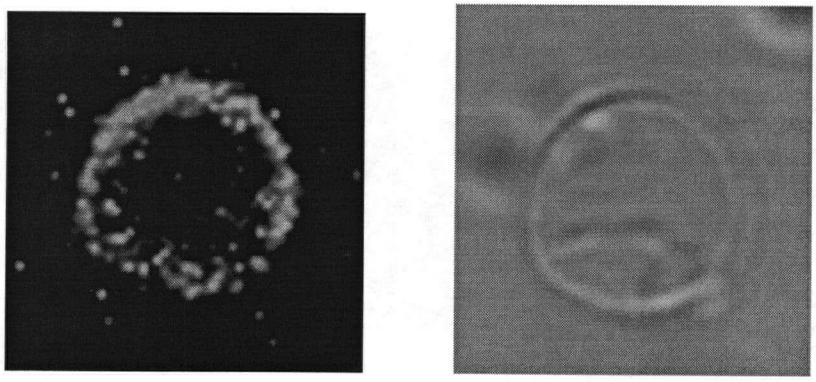
diffuse over the membrane surface and block the conducting channels. In experiments we found that a hidden Markov model (HMM), which takes into account of the 1/f noise in the biosensor's response, is an adequate model for the biosensor currents.

In the presence of target molecules, the stochastic behaviour of the biosensor current changes. By using a sequential maximum likelihood detector, we show that the biosensor can be used in real-time target molecule detection. We illustrate the use of the biosensor in detecting two types of target molecules, methylbenzthonium chloride (MBC) and 2-methyl-4-tert.-octylphenol (MTOP). The experimental data shows that the detection algorithm performed remarkably well even when it was difficult to visually identify the model change. Thus the system we present in this chapter provides a sensitive platform for the development of artificially constructed biosensors that better mimic the function of living cells.

The remainder of this chapter is organized as follows. Sec. 4.1 describes the construction of the biosensor. Sec. 4.2 presents a hidden Markov model to describe the dynamical behaviour of the biosensor. We also present model valididation methods to verify the goodness-of-fit of the HMM to the biosensor response. In Sec. 4.3, we discuss sequential detection algorithms for detecting target molecules in real-time; and finally, in Sec. 4.4 we demonstrate the experimental detection of two target moecules, MBC and MTOP, both of which are known to inhibit the conduction of bis-gA channels.

4.1 Experimental Construction of bis-gA Channel Biosensor

The biosensor considered in this chapter was constructed by incorporating bis-gA ion channels into the lipid bilayer membrane of giant unilamellar liposomes and then excising small patches (1 μm in diameter) of the lipid membrane using a patch-clamp micropipette. The bis-gA was synthesized at the Ambri laboratories. Figure 4.1 shows the fluorescence and phase-contrast image of the optical section through the diameter of the biosensor. The solutions and chemicals used for the model biosensor included DL-alpha-phosphatidylcholine from soybean, cholesterol, chloroform, sucrose, glucose, sodium chloride, potassium chloride and 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid.



(a) (b)

Figure 4.1: (a) Fluorescence image of biosensor's horizontal optical section shows the bis-gA channels labelled using fluorescein and identified by the green color. (b) Phase-contrast image of the same horizontal slice shows the overall shape of the biosensor.

Giant unilamellar liposomes were prepared from a standard hydration procedure with minor modifications. Bis-gA channels were incorporated into lipid membrane patches excised from the giant unilamellar liposomes. A patch-clamp pipette with a tip-opening between 0.9 and 1.5 μm was used to perform the excision and to record the ionic currents from the model biosensor that resulted from the permeation of ions through the bis-gA ion channels. The recording from these membrane patches was amplified and filtered at 1 kHz (4-pole Bessel) using an Axopatch 200B amplifier (Axon Instruments) and sampled online at 10 kHz. Figure 4.2 shows a schematic of the experimental setup.

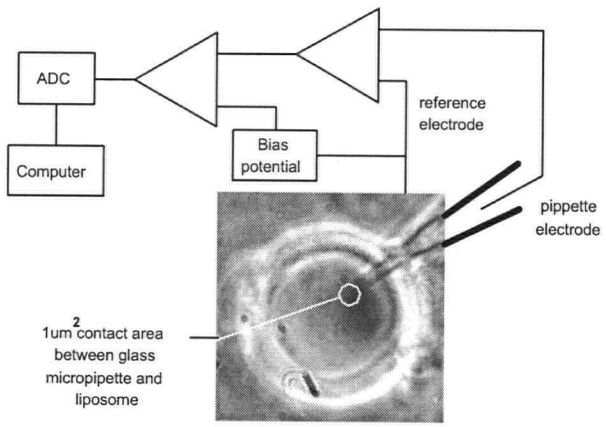


Figure 4.2: Block diagram of experimental setup with photo of glass micropipette and liposome

4.2 Statistical Modeling and Validation of Biosensor Current

In this section we discuss the use of hidden Markov modeling to describe the biosensor response and introduce estimation techniques to extract the model parameters. Statistical model validation tests are used to verify the goodness-of-fit of the model.

Suppose a patch clamp experiment is conducted with $(N-1)$ gramicidin A channels in the biosensor. At discrete time k, each bis-gA channel can be either in the "open" or "closed" state and each open channel conducts a fixed current. Thus the total current due to all $(N-1)$ ion channels at any given time can take on one of N possible levels $\{\mu_1, \ldots, \mu_N\}$ and can be modeled as a N-state Markov chain. Write as $\mu = (\mu_1, \ldots, \mu_N)$. Let I_k denote the total channel current at discrete time k. Let

$$a_{ij} = P(I_k = \mu_j | I_{k-1} = \mu_i),\ i, j \in \{1, \ldots, N\} \tag{4.1}$$

denote the transition probabilities of the Markov chain. Also let

$$\pi_0(i) = P(I_1 = \mu_i), i \in \{1, \ldots, N\} \tag{4.2}$$

denote the initial distribution of the Markov chain. Write $A = [a_{ij}]_{N \times N}$ and $\pi_0 = [\pi_0(i)]_{N \times 1}$. The measured current from the biosensor is a distorted version of the signal I_k. The distortion arises from thermal noise as well as open channel noise that has its power proportional to the inverse of frequency. Thus this is also known as 1/f noise and is discussed in other

studies of bis-gA ion channels. [36, 37]. Figure 4.3 shows the power spectral density of a typical sequence of biosensor recordings.

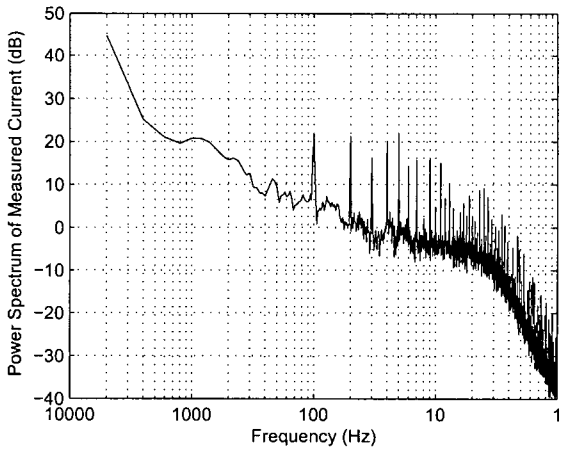


Figure 4.3: Power spectral density of biosensor response clearly shows the 1/f noise and anti-aliasing effect.

At low frequencies, the 1/f noise process is evident as the power spectrum decreases at a rate of -10 dB/dec. The effect of the anti-aliasing filter causes a sharp cutoff at approximately 1 kHz. To model this correlated noise process, we can use an auto-regressive (AR) Gaussian process that comprises white Gaussian noise process W_k filtered by an all-pole filter. We represent the filter with transfer function $H(q^{-1})$, where q^{-1} denotes the unit delay operator:

$$Y_k = I_k + \frac{W_k}{H(q^{-1})} \tag{4.3}$$

or equivalently

$$H(q^{-1})Y_k = H(q^{-1})I_k + W_k \tag{4.4}$$

where

$$H(q^{-1}) = 1 + h_1 q^{-1} + \ldots + h_M q^{-M} \tag{4.5}$$

The model in (4.4) can be rewritten as:

$$\mathbf{h}^T \mathbf{Y}_k = \mathbf{h}^T \mathbf{I}_k + W_k \tag{4.6}$$

where $\mathbf{Y_k} = (Y_k, Y_{k-1}, \ldots, Y_{k-M})^T$ and $\mathbf{I_k} = (I_k, I_{k-1}, \ldots, I_{k-M})^T$. Let $\mathbf{h} = (1, h_1, \ldots, h_M)^T$. The standard algorithm to estimate the model parameters in (4.6) is the expectation-maximization (EM) algorithm, which will be explained in detail in Sec. 4.2.1. The key point is that the EM algorithm involves computing the marginal probabilities of the state space $P(\mathbf{I_k})$. The size of the state space increases exponentially with M. This leads to huge complexity issues because 1/f noise is a long memory process and modeling it requires choosing the filter order, M, to be large.

To alleviate the computational complexity, we adjust the model so that the size of the state space is indepedent of the filter order M. We observe that in our current setup, the gating mechanism of the bis-gA channels is a much slower process relative to the sampling rate and most of the power spectrum of the channel resides in low frequencies. Therefore, we argue that

we can redefine the response of the noisy observation as:

$$Y_k = \frac{I_k + W_k}{H(q^{-1})} \tag{4.7}$$

and closely approximate the behaviour of the gramicidin channels with our model in (4.7), provided that the filter $H(q^{-1})$ has unity DC gain. We apply model validation methods in Sec. 4.2.1 to evaluate the performance of our model.

It is convenient to model the noise corrupting the state of the biosensor as state dependent noise - that is the noise variance at any given time instant is dependent on the state of the biosensor at that time instant. Let σ_i^2 be the variance of state i, $i = (1, \ldots, N)$. Write $\sigma^2 = \left(\sigma_1^2, \ldots, \sigma_N^2\right)$. As a result, the observations can be formulated as a hidden Markov model sequence. Let $\theta = \left(A, \pi, \mu, \sigma^2, \mathbf{h}\right)$ be the HMM that characterizes the output measured current from the biosensor.

4.2.1 Parameter Estimation for Biosensor Current

Given an observation sequence $\{Y_k\}$ of length T, we define $L_k(\theta)$ as the log-likelihood of our model at discrete time k. The estimation of the model θ involves processing $\{Y_k\}$ through a HMM maximum likelihood estimator (MLE). The system in (4.7) can be rewritten as:

$$\mathbf{h}^T \mathbf{Y}_k = I_k + W_k \tag{4.8}$$

This formulation is analogous to a standard HMM, except that the ob-

servation sequence is modulated by a finite impulse response (FIR) filter. As mentioned earlier, the EM algorithm is an iterative procedure that solves for local maximum of the likelihood function. The E-step uses the forward-backward recursions to compute the forward probability α_k and backward probability β_k:

$$\alpha_k(i) = P(Y_1, Y_2, \ldots, Y_k, I_k = \mu_i|\theta)$$
$$\beta_k(i) = P(Y_{k+1}, Y_{k+2}, \ldots, Y_T, I_k = \mu_i|\theta)$$

and evaluates the log-likelihood

$$L_k(\theta) = \sum_{t=1}^{k}\sum_{i=1}^{N}\gamma_t(i)\log\left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\exp\left(\frac{-(\mathbf{h}^T\mathbf{Y_k}-\mu_i)^2}{2\sigma_i^2}\right)\right) + \sum_{t=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\zeta_t(i,j)\log(a_{ij}) \tag{4.9}$$

where the probabilities γ_k and ζ_k can be computed from the forward and backward probabilities. The M-step maximizes the log-likelihood with respect to the model parameters as follows:

$$\begin{aligned} a_{ij} &= \frac{\sum_{k=1}^{T}\zeta_k(i,j)}{\sum_{k=1}^{T}\gamma_k(i)} \\ \sigma_i^2 &= \frac{\sum_{k=1}^{T}\gamma_k(i)(\mathbf{h}^T\mathbf{Y_k}-\mu_i)^2}{\sum_{k=1}^{T}\gamma_k(i)} \\ \mu_i &= \frac{\sum_{k=1}^{T}\gamma_k(i)\,\mathbf{h}^T\mathbf{Y_k}}{\sum_{k=1}^{T}\gamma_k(i)} \end{aligned}$$

The coefficients of the filter $H(q^{-1})$ can also be estimated by taking first-order derivatives of the likelihood function. Alternatively, the estimation can be implemented with the Yule-Walker type equations. For details about HMM estimation and the EM algorithm, please see [12, 16, 30, 34].

4.2.2 Statistical Validation of Biosensor Model

Statistical model validation is of key importance for the biosensor since, once a satisfactory stochastic model is determined, an appropriate molecular detection algorithm can be constructed. In Sec. 4.2.1, we discussed model estimation techniques with the assumption that the N-state hidden Markov model with an AR filter is an approriate model for our biosensor measurements. In this section, we need to statistically validate that assumption. Model validation can be done by analyzing the autocorrelations of the residuals, which are generated via a HMM one-step predictor:

$$e_{k|k-1} = Y_k - \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}\alpha_{k-1}(i)\mu_j - \sum_{n=1}^{M} h_n Y_{k-n} \tag{4.10}$$

The lag-l autocorrelation function of the residual is defined as:

$$r(l) = \frac{\sum_{k=l+1}^{T}(e_k - \bar{e})(e_{k-l} - \bar{e})}{\sum_{k=1}^{T}(e_k - \bar{e})^2} \tag{4.11}$$

where $\bar{e}$ is the mean of the residual process. The residuals of an adequately fitted model should be uncorrelated and the autocorrelation should approach 0 as $T \to \infty$. Rather than examining the autocorrelation at each lag l,

the standard procedure is to compute the Ljung-Box Q-statistic defined in (4.12), which computes the cumulative sum of autocorrelations at the first L lags. The Ljung-Box test is used as a portmanteau lack of fit test for model adequacy. [7]

$$Q_{LB} = T(T+2)\sum_{l=1}^{L}\frac{r^2(l)}{(T-l)} \tag{4.12}$$

It is shown in [8] that for an adequate model, the Q-statistics of the residual is approximately distributed as $\chi^2(L)$. We report on the validity of the model on the biosensor experimental data in Sec. 4.4.

4.3 Target Molecule Detection Algorithm

After the design and validation of the dynamical models, in this section we discuss algorithms for the detection of analytes that are known to change the statistics of bis-gA channel conductance. MBC and MTOP are two analytes that are known to inhibit conductance of bis-gA channels. The chemical structure of MBC and MTOP are shown in Figure 4.4 and 4.5 respectively. The interaction of MBC and MTOP with gA channels is described in the manuscript [19].

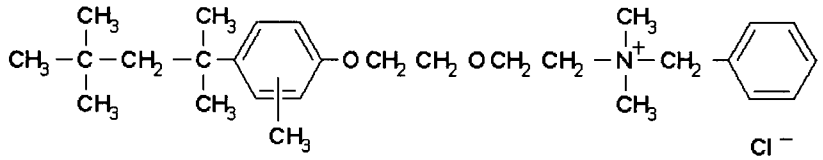


Figure 4.4: Chemical structure of Methylbenzthonium Chloride

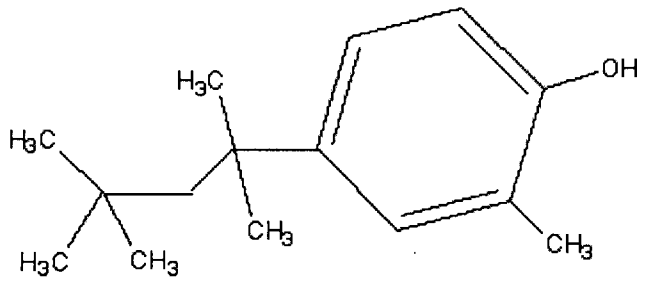


Figure 4.5: Chemical structure of 2-Methyl-4-tert.-octylphenol

Given a measured sequence observed in an unknown condition, the detection problem involves the identification of the condition that most likely contributes to the biosensor's response. This is a model classification problem and can be solved by comparing the likelihood of each known model. Let $Y = (Y_1, \ldots, Y_T)$ be a sequence of observed responses of the biosensor and let $\Theta = (\theta_1, \ldots, \theta_M)$ denote the set of model parameters that characterizes the biosensor's response for known conditions. It is assumed that at each time point k, the sequence Y behaves according to one of M possible models in Θ.

The model parameters Θ are estimated and the log-likelihood at each time point L_k can be computed from (4.9). To make the detection more robust to nonstationary disturbances and outliers in the measurements, we apply a

geometric moving-average filter to the log-likelihood. Let ρ be the forgetting factor, $0 < \rho < 1$. Define the filtered likelihood of model θ at time k:

$$S_k(\theta) = \begin{cases} L_1(\theta) & \text{for } k = 1 \\ (1-\rho)S_{k-1} + \rho L_k(\theta) & \text{for } 2 \leq k \leq T \end{cases} \tag{4.13}$$

The filtered likelihood is a weighted sum of the likelihood of the entire sequence $\{Y_1, \ldots, Y_k\}$, with higher weights on the recent observations. The apriori probabilities of each model is generally unknown, so we devise a maximum-likelihood detector which picks the most likely model at time k given the measured sequence $\{Y_1, \ldots, Y_k\}$:

$$\hat{\theta}_k = \arg\max_{\theta \in \Theta}(S_k(\theta)) \tag{4.14}$$

where $\hat{\theta}_k$ is the maximum likelihood detection of the model at time k.

4.4 Experimental Results of Biosensor

Here we report on the goodness-of-fit of the dynamical model and the performance of the detection algorithm on experimental data. We recorded output from the biosensor by measuring the activity of the bis-gA ion channels that were incorporated into the small lipid membrane patches excised from unilamellar giant liposomes. Bis-gA ion channels were incorporated into the unilamellar giant liposomes at a concentration of 1/100 from a 66 nM stock solution. We used 0.5 M KCl solution in the recording pipette

with the microparticles suspended in a 0.5 M NaCl solution.

4.4.1 Model Estimation and Validation for Biosensor

We recorded the biosensor response at an applied voltage of 50 mV and an amplifer gain of 200. Figure 4.6 shows a three-second recording of biosensor response.

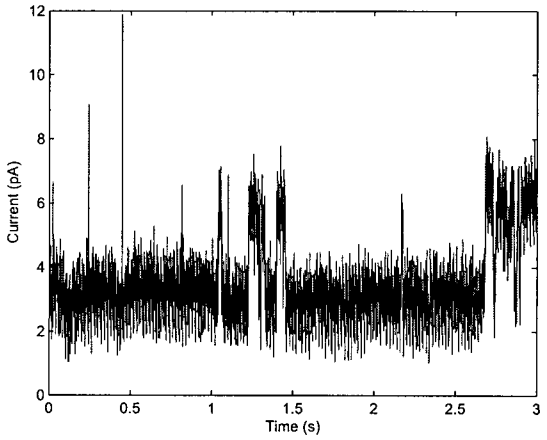


Figure 4.6: Experimental measurement of biosensor response was recorded at an applied voltage of 50 mV and an amplifer gain of 200. The giant liposome was constructed from PC with 10% cholesterol in chloroform.

Notice that the bis-gA current measurements obtained here are more noisy compared to the ones obtained from a solvent sealed black lipid membrane (BLM). The quality is typical of patch clamp recordings in which a membrane patch is captured on the tip of a 1 μm pipette. The seal obtained using the patch clamp technique is generally poorer than that obtained from a BLM, resulting a worse signal-to-noise ratio for the measurements.

Transition Probabilities		**Conductance Level (pA)**	**Variance $(pA)^2$**
0.9964	0.0036	5.831	0.513
0.0006	0.9994	3.057	0.371

Table 4.1: Maximum Likelihood Estimation of HMM parameters of the biosensor recording in Figure 4.6.

A further potential difficulty is the possibility of multiple lipid layers being present in the patch as compared to the BLM geometry which is thinned to a single bilayer. When recording from multiple lipid membrane stacks the signal-to-noise ratio of the individual conduction events is reduced by noise arising from the complex impedance of the series elements.

We fitted the sequence with two-state HMM and a 12^{-th} order AR filter. The model parameters were estimated with the MLE and listed in Table 4.1. The most likely conductance level sequence is extracted from the HMM procedure and plotted in Figure 4.7.

To verify that the two-state HMM outlined in Table 4.1 provides sufficient statistics to model the response of the biosensor, we computed the Q-statistics of the residuals, using (4.10) through (4.12). It can be seen in Figure 4.8 that for the first 13 lags, the Q-statistics are below the critical values of the chi-square distribution at 0.05 significance level. Therefore, the hypothesis that the residual is a white process cannot be rejected. It is interesting to note that the estimated state levels, at approximately 3 pA and 6 pA, are rather large compared to single channel measurement of channel current at a low applied voltage. One plausible explanation for this observation is that the activity of some of the bis-gA channels are correlated. In other words, the level switch from 3 pA to 6 pA may be contributed by

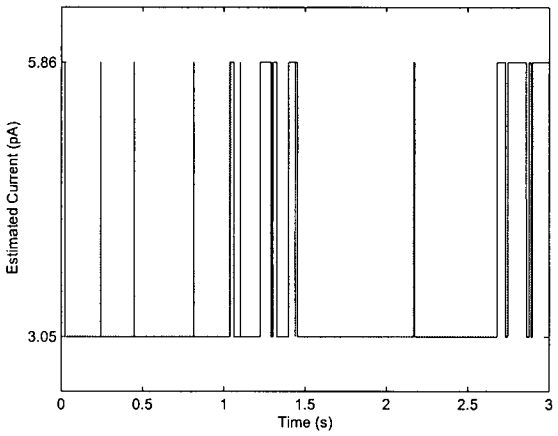


Figure 4.7: Maximum likelihood estimate of biosensor conductance was extracted from the forward-backward procedure in the HMM estimation algorithm.

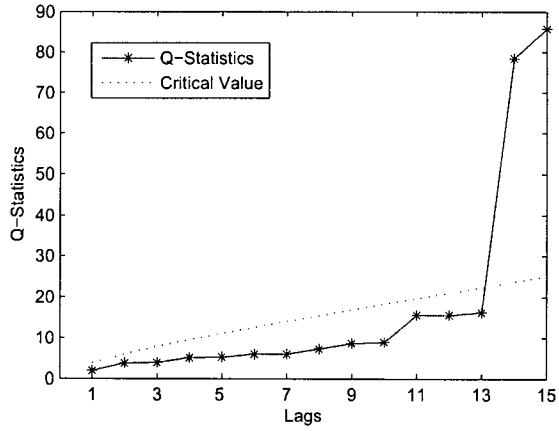


Figure 4.8: Q-statistics of residuals are plotted against the critical values of the Ljung-Box test at significance level of 0.05. For the first 13 lags, the Q-statistics are below the critical values and the hypothesis that the residual is a white process cannot be rejected. Thus the proposed two-state HMM is an adequate model for the biosensor current.

several coupled channels opening at the same time.

4.4.2 Real-Time Detection of Target Molecules

In this section we illustrate the performance of the algorithm in detecting the presence of target molecules in real-time. We present experimental results of the biosensor in detecting two analytes, MBC and MTOP. Both compounds are known to inhibit conduction of the bis-gA channels. Patch-clamp experiments are conducted with and without MBC in the bath solution. Let

θ_1 = with MBC in the bath solution

θ_2 = with no MBC in the bath solution

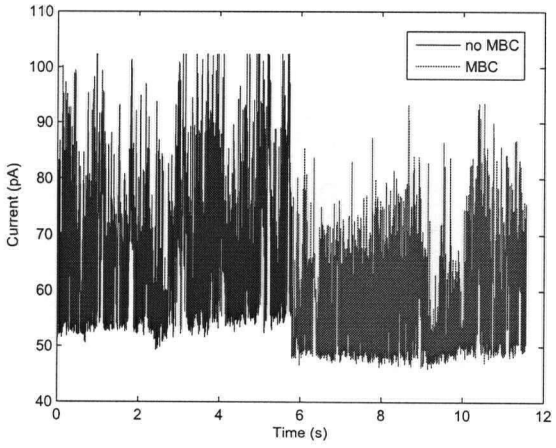


Figure 4.9: Biosensor's response to addition of MBC at $k = 57.828$ seconds.

To simulate the addition of MBC into the bath solution, we merged together sequences recorded with and without MBC. The merged sequence is plotted in Figure 4.9. Assuming a four-state HMM, the parameters of θ_1

and θ_2 are estimated with the EM algorithm.

The filtered likelihood of each model is shown in Figure 4.10. Since the stochastic model successfully captures the dependencies in the biosensor's response, the detection algorithm quickly and accurately estimates the model switching point. The detection trace in Figure 4.11 indicates a switch in the most likely model from θ_1 to θ_2 at $k = 59.865$ seconds, approximately 0.2 seconds after MBC is added.

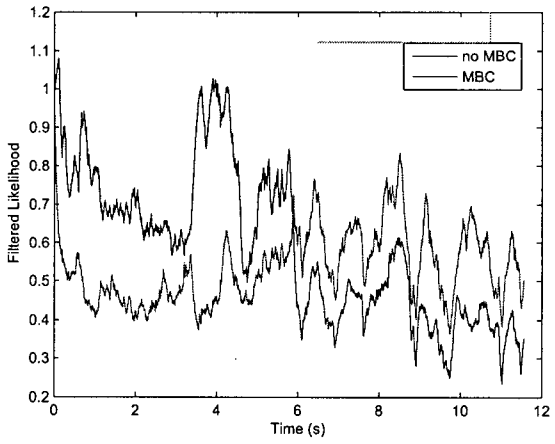


Figure 4.10: Filtered likelihoods for estimated models of MBC, with $\rho = 0.0001$.

In the second example, we test the algorithm's performance in identifying the concentration of MTOP in the bath solution. Let

θ_1 = no MTOP in the bath solution

θ_2 = 50 uM MTOP in the bath solution

θ_3 = 100 uM MTOP in the bath solution

We simulated the change in concentration of MTOP by merging channel

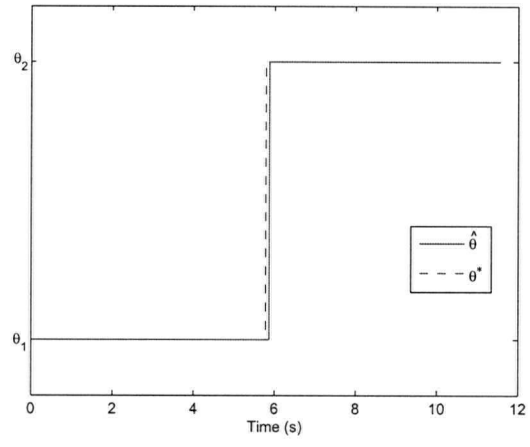


Figure 4.11: Maximum likelihood detection of MBC is generated using the likelihoods in Figure 4.10. While it is difficult to visually identity the transition from the recordings, the detection algorithm detected a model change shortly after MBC was added to the solution.

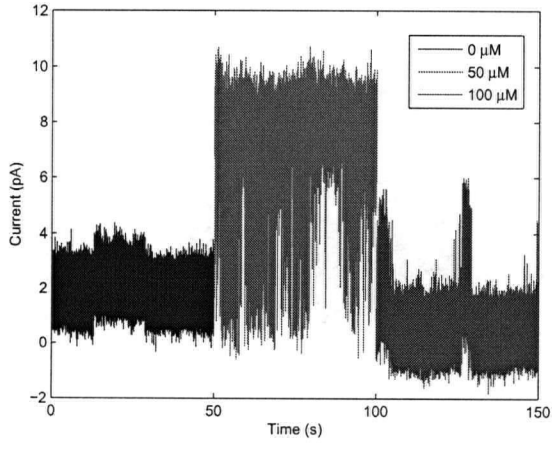


Figure 4.12: Biosensor's response to various concentrations of MTOP. 50 μM of MTOP was added to the bath solution every 50 seconds.

recordings independent from the training sequences. The merged sequence is plotted in Figure 4.12. The parameters of θ_1, θ_2 and θ_3 were estimated offline with the EM algorithm. The filtered likelihood and the detection trace are plotted in Figure 4.13 and Figure 4.14 respectively. The algorithm identified the correct concentration of MTOP 98.1% of the time.

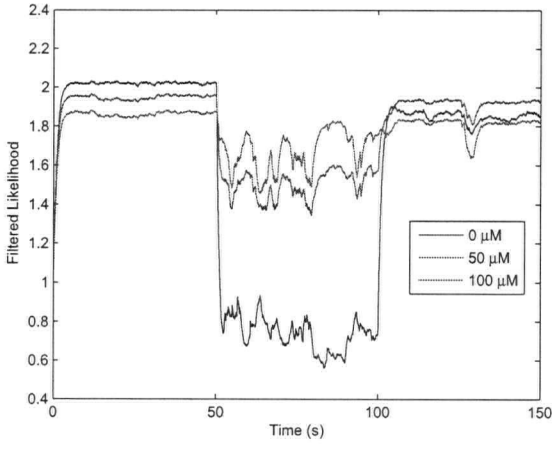


Figure 4.13: Filtered likelihoods for estimated models of MTOP, with $\rho = 0.0001$.

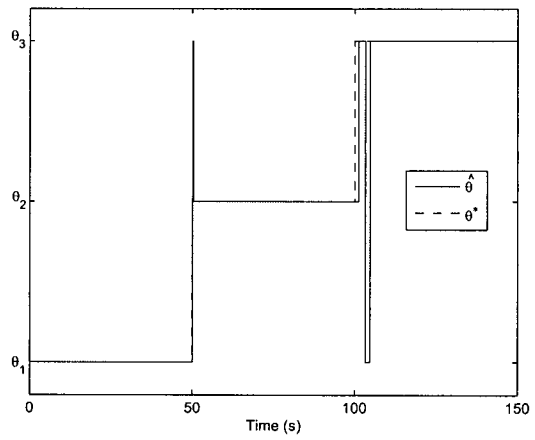


Figure 4.14: Maximum likelihood detection of MTOP generated using the likelihoods in Figure 4.13, showing the most likely condition. The algorithm detected the correct model 98.1% of the time.

Chapter 5

Conclusion and Extensions

This thesis is concerned with the statistical modeling and application of gramicidin A ion channels. In particular, we have addressed two important issues in the area of ion channel permeation and ion channel based biosensor. Brownian dynamics is a practical technique to simulate ion channel conduction but it suffers from huge computational complexity at high ionic concentration. We proposed the use of a finite-state semi-Markov model to describe the binding site kinetics of gramicidin A channels. We extrapolated the model parameters with the logistic function and used the results to predict the channel current at high ionic concentrations. Numerical results showed that the technique correctly predicted the channel current and conduction event distribution. In future work, we will use the same modeling and extrapolation technique to study single channel current as a function of applied voltage.

In a separate study, we developed stochastic modeling and detection algorithms as part of an ion channel based bionsensing platform that can automatically detect a variety of target molecules. The conduction of an ion channel biosensor can be modeled by an AR-modulated hidden Markov model. The presence of specific target molecules distorts the channels in

the biosensor, thus changing the statistics of the biosensor's conduction. We devise a real-time maximum likelihood detector to capture the change and detect the presence of target molecules. The algorithm is tested on a biosensor built by incorporating dimeric gramicidin A channels into bilayer membranes of giant unilamellar liposomes. The performance of the biosensor is tested with addition of MBC and MTOP, both known to inhibit conduction of gramicidin A channels. Experimental results show that the detection performed well even when the change in the conduction level was difficult to visualize. The detection algorithm provides the sensitive detection system for ongoing development of membrane-based biosensors. In future work, we will construct a more robust biosensor by incorporating bis-gA channels into a BLM, where we can observe the activity of a single bis-gA channel and investigate in more detail biological properties of the bis-gA channels.

Bibliography

[1] T. W. Allen, T. Bastug, S. Kuyucak, and S. H. Chung. Gramicidin A channel as a test ground for molecular dynamics force fields. *Biophysical Journal*, 84:2159–2168, 2003.

[2] A. Anastasiadis and F. Separovic. Solid-state NMR structural determination of components in an ion channel switch biosensor. *Australian Journal of Chemistry*, 56:163–166, 2003.

[3] F. M. Ashcroft. *Ion Channels and Disease: Channelopathies.* Academic Press, 2000.

[4] T. Bastug, S. M. Patra, and S. Kuyucak. Molecular dynamics simulations of gramicidin A in a lipid bilayer: From structure-function relations to force fields. *Chemistry and Physics of Lipids*, 141:197–204, 2006.

[5] Turgut Bastug and Serdar Kuyucak. Molecular dynamics simulation of calcium binding in gramicidin. *Chemical Physics Letter*, 424:82–85, 2006.

[6] S. Berneche and B. Roux. Molecular dynamics of the KcsA K^+ channel in a bilayer membrane. *Biophysical Journal*, 78:2900–2917, 2000.

[7] G. E. P. Box and G. M. Jenkins. *Time Series Analysis - Forecasting and Control.* Holden-Day, 1976.

[8] G. E. P. Box and D. A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of American Statistical Association*, 65:1509–1526, December 1970.

[9] O. Braha, L. Gu, L. Zhou, X. Lu, S. Cheley, and H. Bayley. Simultaneous stochastic sensing of divalent metal ions. *Nature Biotechnology*, 18:1005–1007, 2000.

[10] M. Chakravarti, R. G. Laha, and J. Roy. *Handbook of Methods of Applied Statistics*, volume 1. John Wiley and Sons, 1967.

[11] S. H. Chung, O. S. Andersen, and V. Krishnamurthy. *Handbook of Ion Channels: Dynamics, Structure, and Applications.* Springer Verlag, 2006.

[12] S. H. Chung, V. Krishnamurthy, and J. B. Moore. Adaptive processing techniques based on hidden Markov models for characterising very small channel currents buried in noise and deterministic interferences. *Proc. Phil. Trans. Roy. Soc. Lond. B*, 334:357–384, 1991.

[13] B. A. Cornell, V. L. Braach-Maksvytis, L. G. King, P. D. Osman, B. Raguse, L. Wieczorek, and R.J. Pace. A biosensor that uses ion-channel switches. *Nature*, 387:580–583, 1997.

[14] B. Corry, T. Allen, S. Kuyucak, and S. H. Chung. Mechanisms of permeation and selectivity in calcium channels. *Biophysical Journal*, 80:195–214, 2001.

[15] R. B. D'Agostino and M. A. Stephens. *Goodness-of-fit techniques.* CRC, 1986.

[16] A. P. Dempster, M. M. Laird, and D. B. Rubin. Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, 39:1–38, 1977.

[17] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. Mackinnon. The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science*, 280:69–77, 1998.

[18] R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, and R. MacKinnon. X-ray structure of a CIC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. *Nature*, 415:287–294, 2002.

[19] B. A. Cornell et al. Making lipid membranes even tougher. *Journal of Material Research*, 2007. In Press.

[20] B. Hille. *Ionic Channels of Excitable Membranes.* Sinauer Associates, 2001.

[21] S. Howorka, S. Cheley, and H. Bayley. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnology*, 19:636–639, 2001.

[22] Matthew Hoyles. Computer simulation of biological ion channels. *PhD Thesis*, 2000.

[23] Matthew Hoyles, Serdar Kuyucak, and S. H. Chung. Computer simulation of ion conductance in membrane channels. *Phys. Rev. E. of Computational and Theoretical Nanoscience*, 58:3654–3661, 1998.

[24] E. Jakobsson. Using theory and simulation to understand permeation and selectivity in ion channels. *Methods*, 14:342–351, 1998.

[25] R. R. Ketchem, W. Hu, and T. A. Cross. High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science*, 261:1457–1460, 1993.

[26] V. Krishnamurthy and S. H. Chung. Brownian dynamics simulation for modeling ion permeation across bionanotubes. *IEEE Transactions on Nanobioscience*, 4:102–111, 2005.

[27] V. Krishnamurthy, M. Hoyles, R. Saab, and S. H. Chung. Permeation in gramicidin ion channels by directly estimating the potential of mean force using Brownian dynamics simulations. *Journal of Computational and Theoretical Nanoscience*, 3:702–711, 2006.

[28] V. Krishnamurthy, T. Vora, and S. H. Chung. Adaptive Brownian dynamics algorithms for estimating the shape of sodium ion channels. *Journal of Nanoscience and Nanotechnology*, 7:2273–2282, 2007.

[29] Vikram Krishnamurthy, Matthew Hoyles, Rayan Saab, and S. H. Chung. Permeation in gramicidin ion channels by directly estimating

the potential of mean force using Brownian dynamics simulations. *Journal of Computational and Theoretical Nanoscience*, 3:702–711, 2006.

[30] Vikram Krishnamurthy and A. Logothetis. Iterative and recursive estimators for hidden Markov errors-in-variables models. *IEEE Trans. Signal Processing*, 44:629–639, March 1996.

[31] David G. Levitt. Modeling of ion channels. *Journal of General Physiology*, 113:789–794, 1999.

[32] L. Movileanu, S. Howorka, O. Braha, and H. Bayley. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nature Biotechnology*, 18:1091–1095, 2000.

[33] G. A. Olah, H. W. Huang, W. H. Liu, and Y. L Wu. Location of ion-binding sites in the gramicidin channel by x-ray diffraction. *Journal of Molecular Biology*, 218:847–858, 1991.

[34] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*, 77:257–285, 1989.

[35] K. E. Sapsford, Y. S. Shubin, J. B. Delehanty, J. P. Golden, C. R. Taitt, L. Cl Shriver-Lake, and F. S. Ligler. Fluorescence-based array biosensors for detection of biohazards. *Journal of Applied Microbiology*, 96:47–58, 2004.

[36] R. Sauve and E. Bamberg. 1/f noise in black lipid membranes induced by ionic channels formed by chemically dimerized gramicidin A. *Journal of Membrane Biology*, 43:317–333, November 1978.

[37] R. Sauve and G. Szabo. Interpretation of 1/f fluctuations in ion conducting membranes. *Journal of Theoretical Biology*, 113:501–516, April 1985.

[38] E. A. Smith and R. M. Corn. Surface plasmon resonance imaging as a tool to monitor biomolecular interactions in an array based format. *Appl. Spectroscopy*, 57:320–332, 2003.

[39] W. F. van Gunsteren and H. J. C. Berendsen. Algorithms for Brownian dynamics. *Molecular Physics*, 45:637–647, 1982.

[40] T. Vora, B. Corry, and S. H. Chung. A model of sodium channels. *Biochimica et Biophysica Acta - Biomembranes*, 1668:106–116, 2004.

[41] B. A. Wallace. Recent advances in the high resolution structures of bacterial channels: Gramicidin A. *Journal of Struct. Biol.*, 121:123–141, 1968.

Appendix A

List of Acronyms

AR	Auto-regressive
BD	Brownian dynamics
bis-gA	dimeric gramicidin A
BLM	black lipid membrane
CDF	cumulative distribution function
EDF	empirical distribution function
EM	expectation-maximization
FIR	finite impulse response
gA	gramicidin A
HMM	hidden Markov model
ICS^{TM}	Ion Channel Switch
MBC	methylbenzthonium chloride
MD	Molecular dynamics
MLE	maximum likelihood estimator
MTOP	2-methyl-4-tert.-octylphenol
NMR	nuclear magnetic resonance
PCR	polymerase chain reaction
PDF	probability distribution function
PMF	potential of mean force
SMC	semi-Markov chain