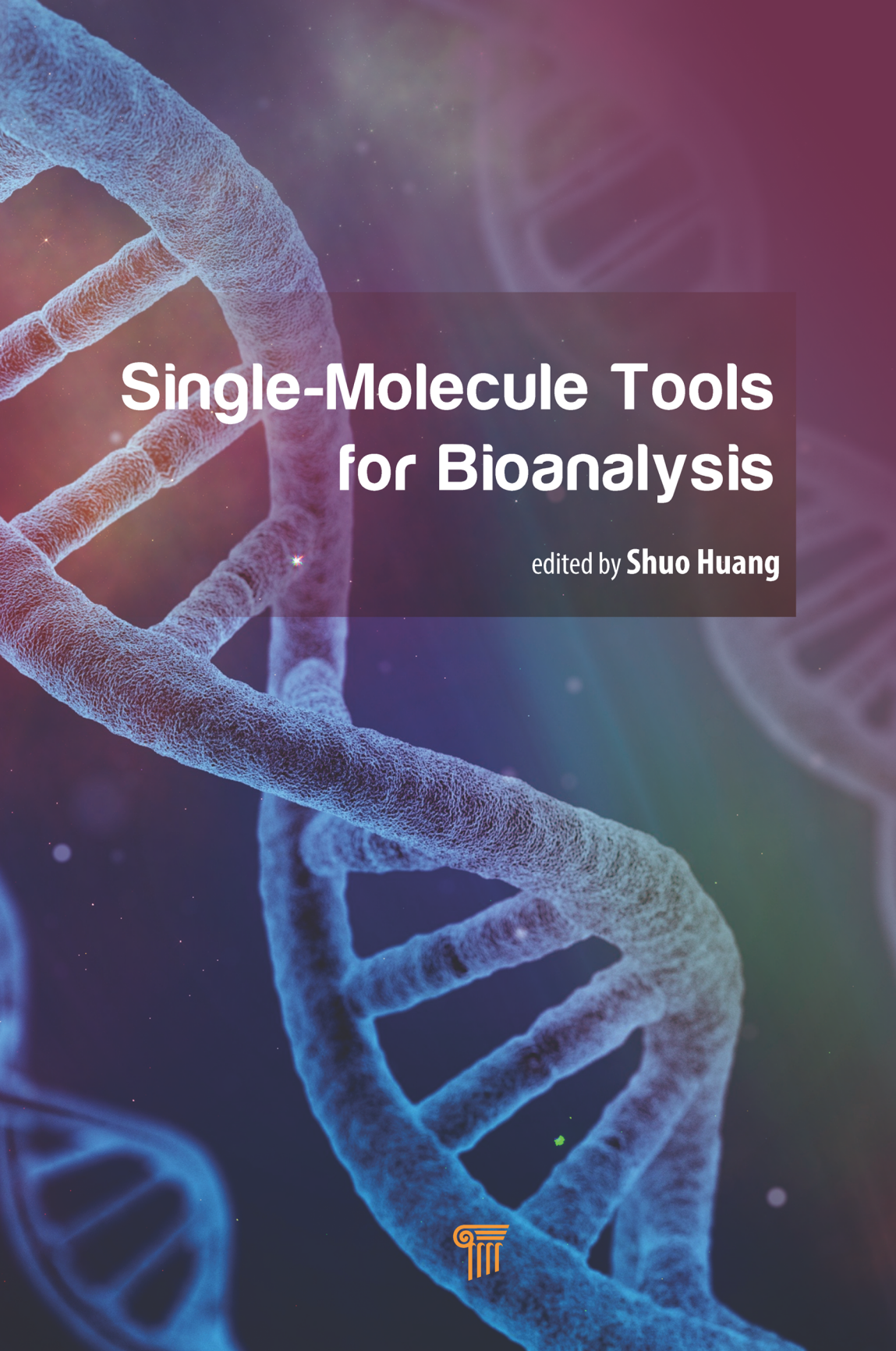
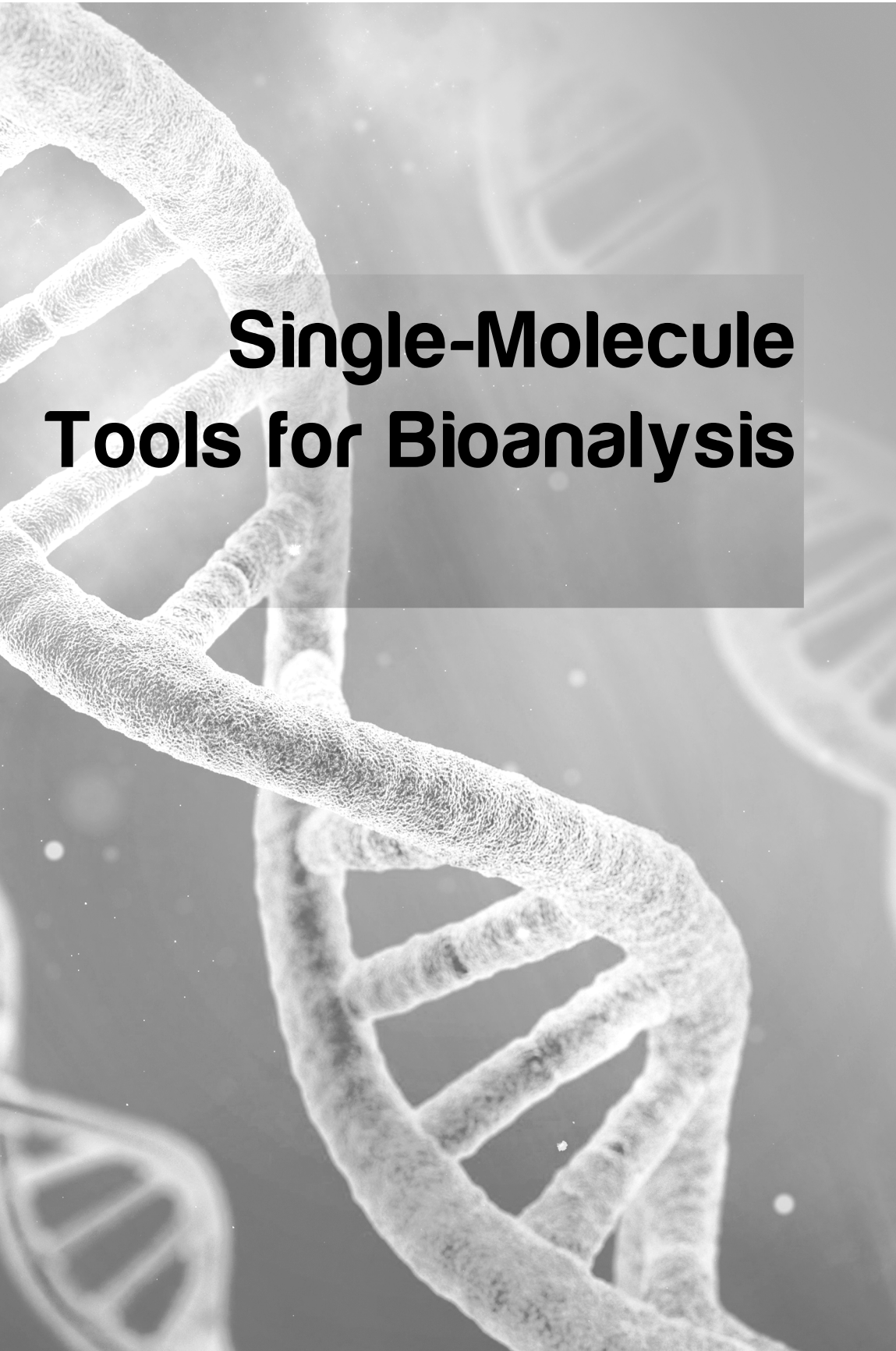

Single-Molecule Tools
for Bioanalysis
edited by Shuo Huang



Single-Molecule Tools for Bioanalysis

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Single-Molecule Tools for Bioanalysis

edited by

Shuo Huang

JENNY STANFORD
PUBLISHING

Published by

Jenny Stanford Publishing Pte. Ltd.
Level 34, Centennial Tower
3 Temasek Avenue
Singapore 039190

Email: editorial@jennystanford.com
Web: www.jennystanford.com

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Single-Molecule Tools for Bioanalysis

ISBN 978-981-4800-44-0 (Hardcover)
ISBN 978-1-003-18913-8 (eBook)

Contents

Preface

Single-molecule biophysics research is a highly interdisciplinary study that requires diverse expertise in biology, chemistry, physics, and engineering, aiming to understand biological processes at single-molecule level against ensemble averaging. Investigations of single-molecule biophysics have enabled direct measurement of single-molecule properties that were not even previously feasible by any ensemble methods. These achievements include, but are not limited to, direct measurement of the elastic property of an individual strand of nucleic acids, direct manipulation of nucleic acids or protein molecules, optical imaging of cellular processes in a nanometer resolution, direct torque measurement of a supercoiled DNA, and several others. The fast development of the field has also stimulated the invention and evolution of a large variety of emerging single-molecule tools, which have enabled new concepts and applications of bioanalysis.

Though there is an urgent need to systematically summarize these achievements, it is too much for any individual review article to achieve a full coverage with sufficient details. Written by young experts in the field of single-molecule research, this book aims to provide a systematic and in-depth recap of representative topics of single-molecule bioanalysis. The book contains six chapters that cover topics on nanopores, optical tweezers, single-molecule FRET, DNA origami sensors, magnetic tweezers, and ABEL trap. Each chapter provides the general concept and a brief history of the methods, technical fundamentals, diversified forms of the methods, and the representative applications of the methods. This makes the book ideal as a textbook for a graduate-level course. In fact, the materials in this book were indeed summarized from the lecture notes of a graduate course supervised by the book editor at Nanjing University since 2015.

To make each chapter appealing to entry-level readers, a highly simplified tutorial protocol is also provided at the end of each chapter. The design of these protocols may be far from the actual measurement performed in the relevant scientific publications. They were, however, designed so that it can be finished by even undergraduate students or most graduate students with highly accessible scientific materials. Owing to the deficit of time, a few interesting topics such as electron microscopy, super-resolution microscopy, and scanning probe microscopy could not be covered. However, they are likely to be included in the future editions of this book.

Dr. Shuo Huang
School of Chemistry and Chemical Engineering
State Key Laboratory of Analytical Chemistry for Life Sciences
Chemistry and Biomedicine Innovation Center (ChemBIC)
Nanjing University

Acknowledgements

I would like to thank Dr. Mark Ian Wallace (King's College of London) for his suggestion to write this book. I would also like to thank the whole editorial team for their support during the preparation of this book. In particular, I would also like to thank Jenny Rompas and Stanford Chong, the publishers, for their support and encouragement during the book writing. I would also like to acknowledge Arvind Kanswal for the editorial support. The efforts of all chapter authors are also highly acknowledged.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Chapter 1

Single-Molecule Analysis by Biological Nanopores

Yuqin Wang and Shuo Huang

School of Chemistry and Chemical Engineering, Nanjing University, China
State Key Laboratory of Analytical Chemistry for Life Science, Nanjing, China
Chemistry and Biomedicine Innovation Center (ChemBIC), Nanjing, China
shuo.huang@nju.edu.cn

Biological nanopores are a type of proteins which form pores and were developed for in vitro single-molecule sensing. In view of their ease of use, consistency and precision of preparation, biological nanopores can be further engineered or modified for highly specialized sensing applications. Using single-channel recording, the identity of an analyte is reported from its interaction with the pore restriction during its translocation. Being geometrically compatible with single-stranded DNA or single-stranded RNA, biological nanopores, such as α-hemolysin (α-HL) or *Mycobacterium smegmatis* porin A (MspA), have long been considered the most promising candidates for third-generation sequencing. After research of ~3 decades, the prototype of a nanopore sequencer was first demonstrated in 2012 and is now widely used in a variety of genomics research programs. Sustained research of nanopore sequencing has also stimulated its

Single-Molecule Tools for Bioanalysis
Edited by Shuo Huang

ISBN 978-981-4800-44-0 (Hardcover), 978-1-003-18913-8 (eBook)
www.jennystanford.com

other applications, such as sensing of single ions, small molecules, macromolecules, biomacromolecules, or their combinations. In this chapter, we introduce the mechanism and the methodology of the biological nanopore techniques along with a tutorial protocol. We hope the reader will benefit from reading this chapter by successfully carrying out a highly simplified nanopore measurement or becoming inspired for their own research.

1.1 Introduction

1.1.1 Single-Molecule Biophysics and Nanopore

The field of single-molecule biophysics lies at the interface of physics, biology, and chemistry; its aim is to understand the mechanism of biological phenomena on a single-molecule scale. Propelled by the dramatic improvement of modern technologies such as patch-clamp technology [1], electron microscopy (EM) [2], scanning probe microscopy (SPM) [3], optical tweezers [4], magnetic tweezers [5], single-molecule fluorescence [6], super resolution microscopy [7], Förster resonance energy transfer [8], and total internal reflection fluorescence microscopy [9], researchers can now easily sense or even visualize unprecedented insights into enzyme kinetics [10], conformational dynamics [11], protein folding kinetics [12], breaking of chemical bonds [13], and ligand-binding activities [14] in a single molecule.

Each of the single-molecule methods mentioned above is irreplaceable, but not universal. For particular single-molecule applications, the proper selection of the methodology becomes critical for success. Nanopore technology is a unique method which is particularly suitable as a sensor for chain-shaped and electrically charged polymers such as nucleic acids [15]. The predecessor of the nanopore method could be traced back to the invention of the patch-clamp technology in the 1970s, which described the first instrument enabling human beings to monitor single ion channel activities on cell membranes [1]. This invention gained Neher and Sakmann the Nobel Prize in physiology or medicine in 1991 and is now widely used as a tool for electrophysiology studies of transmembrane porins or ion channels. The nanopore method, which conceptually originated from patch-clamp measurements, determines molecular identities

by probing pore blockage events caused by molecular interactions of an analyte with a nanopore sensor (Figure 1.1A) [16]. In a typical nanopore measurement, a strand of a single-molecule analyte (e.g. a piece of single-stranded DNA [ssDNA]) electrophoreses through the nanoscopic aperture, generating a transient resistive pulse signal containing the molecular identity information (Figure 1.1B,C). This molecular transport process through a nanopore is termed a "translocation" event. Molecular identities are recognized by analyzing the trace fluctuations caused by the molecular translocation. The nanopore method is so sensitive that detection of subtle differences between analyte molecules is possible, making it an efficient single-molecule sensor like a miniaturized Coulter counter [17].

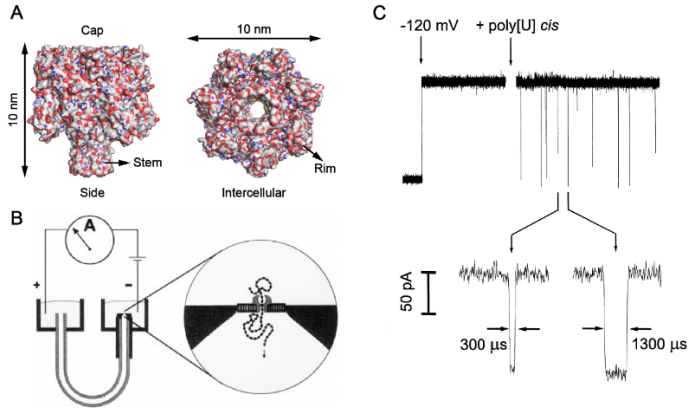


Figure 1.1 The origin of nanopore and its DNA-sensing applications. (A) The crystal structure of a heptameric α-HL nanopore. The heptameric pore appears with a mushroom shape, consisted of a wider cap (vestibule) and a narrower stem (β-barrel). The stem, which is composed of 14 antiparallel β strands, forms a cylindrical channel with a 2.6 nm diameter, permitting translocation of only ssDNA instead of dsDNA. (B) Traditional nanopore apparatus. A single α-HL nanopore can spontaneously insert into a freestanding lipid bilayer forming the only conducting path across the membrane. Analytes such as ssDNA are electrophoretically driven to pass through the pore, giving rise to the appearance of resistive pulses caused by the analyte. Reproduced with permission from ref. [18], Copyright (1999) The Biophysical Society. (C) Characteristic blockades of poly(U) translocation through an α-HL nanopore. Reproduced with permission from ref. [19], Copyright (1996) National Academy of Sciences, USA.

The nanopore measurement has advantages over other optics-based single-molecule methods because it monitors ionic current

instead of photon counts. Due to limited photon emissions from fluorophores, single-molecule methods based on fluorescent microscopy such as single-molecule fluorescence resonance energy transfer (smFRET), normally produce noisy data which may limit its sensing resolution [8]. For a nanopore device based on natural ion channels, the measurement range is between 1 pA and 200 pA, which is equivalent to acquisition of 6.25–1250 million ions per second.

$$\left(1\times10^{-12}\,\mathrm{A}\right)\times\frac{1\mathrm{s}}{1.60\times10^{-19}\,\mathrm{C}}=6.25\times10^{6}\ \text{ions}$$

This amount of charge transport can be reliably amplified and measured by a patch-clamp amplifier with a satisfactory signal-to-noise ratio, but prolonged excitation causes severe photo bleaching of the fluorophore, limiting the duration of the measurement. On the other hand, the nanopore device can withstand hours, or days of continuous measurement.

1.1.2 Nanopore Methods

A nanopore sensor could be generally defined as a nanoscale aperture in an impermeable membrane connecting two chambers containing electrolyte solution. A wide range of materials and methods can be utilized to make nanopore devices with different geometries and properties. An ideal nanopore sensor has to be structurally stable and geometrically consistent. To fit the cross-sectional area of a single biomacromolecule, the size of a useful nanopore sensor is normally between 1 nm and 10 nm in diameter [20]. However, nanofabrication techniques in the 1990s cannot yet reliably produce such delicate a structure over an artificial material. Until 1996, the structural determination of the *Staphylococcus aureus* α-HL by X-ray crystallography [16] suggested a biomimetic strategy to produce nanopores, and this later became an initiator of all subsequent nanopore researches.

In general, nanopore devices can be further classified into "biological nanopores" and "solid-state nanopores." All biological nanopores originate from natural transmembrane porins or their mimics, which can spontaneously penetrate a natural biomembrane or an artificial lipid bilayer and generate ion or molecular passages across the insulating membrane for biological sensing.

Biological nanopores could be massively prepared on a large scale by standard molecular biology protocols such as prokaryotic or in vitro protein expression followed with the appropriate purification steps. Though naturally composed of amino acids, biological nanopores when stored properly can stay active for a few years with no noticeable difference during measurements. It has also been experimentally verified that a biological nanopore such as an α-HL can survive an extreme of salt concentration [21], temperature [22], pH [23], and denaturants [24–26] during measurements. However, it is the fragile lipid bilayer or biomembrane which is unable to withstand harsh measurement conditions such as a high applied electrical bias, violent mechanical vibrations, or the presence of strong detergents.

Alternatively, solid-state nanopores, which are porin mimics artificially fabricated on solid-state thin materials, were developed later and aimed to provide a more durable, silicon industry-compatible solution with a complete freedom of design flexibility. Various methods such as focused ion beam [27], electron beam [28], track-guided chemical etching [29], and dielectric breakdown [30] could be used for pore drilling. Solid-state nanopore techniques offer advantages of a more flexible pore geometry and patterning along with a variety of surface property modifications but suffers from a poor biosensing performance due to the inconsistency of pore manufacturing at the nanometer scale.

Other emerging nanopore technologies, which cannot be classified in either of the types mentioned above, have also been investigated. For instance, the fusion of biological nanopores and solid-state nanopores, termed hybrid nanopores, is expected to overcome the limitations on both sides [31, 32], and the emerging DNA origami nanopores possess the ability for precise control over the geometry and the surface functionality [33–35].

Among different types of nanopores, the biological nanopore method is the first reported method in the field and is currently the only nanopore method which can sequence DNA or RNA. With long-term debates over the pros and cons of different pore types, biological nanopores currently still outperform their solid-state counterparts in the aspects of the signal-to-noise ratio, spatial resolution, and manufacturing consistency. Due to the limitations of space, the remainder of the discussion will be focused on biological nanopores.

1.1.3 Biological Nanopores

Biological nanopores are a category of transmembrane porins with a considerably large pore lumen measuring from 1 to 5 nm in diameter, which are utilized for single-molecule sensing. Reported biological nanopores including *S. aureus* α-HL [19], MspA [36], *Escherichia coli* ferric hydroxamate uptake protein A (FhuA) [37], bacteriophage phi29 connector [38], *E. coli* cytolysin A (ClyA) [39], *E. coli* outer membrane protein G (OmpG) [40], *E. coli* outer membrane protein F (OmpF) [41], *Actinia fragacea* fragaceatoxin C (Frac) [42], *E. coli* curli production assembly/transport component CsgG (CsgG) [43], and human specificity protein 1 (Sp1) [44] could be similarly utilized as single-molecule sensors (Figure 1.2). The α-HL, which is the most studied biological nanopore, is the most robust nanopore sensor used to date. In nature, it is an exotoxin secreted by the human pathogen *S. aureus* bacterium. In its heptameric form, α-HL appears as a mushroom-shaped protein (with a cap domain and a stem domain) and a molecular weight of 232.4 kDa. The stem domain, which is embedded in the lipid membrane, is composed of 14 antiparallel β strands that form a cylindrical channel for molecular transportation. The narrowest spot of the cylindrical channel is ~1.4 nm in diameter. It serves as the recognition site for molecular identity discrimination [16] and permits passage only of ssDNA. The cap domain, which has an inner diameter of ~4.5 nm, is capable of accommodating a short fragment of dsDNA.

The MspA nanopore, which is a funnel-shaped octameric pore with a molecular weight of 157 kDa, is more useful in nanopore sequencing. Benefiting from its short and narrow recognition site, which is ~0.6 nm long and ~1.2 nm wide, the MspA pore is capable of reading a frame of only 4 nucleotides at a time over a piece of ssDNA during its translocation.

Both α-HL and MspA permit translocation of ssDNA but not dsDNA. The phi29 connector protein, which is a dodecamer of the GP10 protein, opens up a ~3.6 nm diameter channel capable of translocating dsDNA or proteins of a low molecular weight. However, it is reported that unlike α-HL and MspA, a phi-29 connector protein cannot spontaneously insert into the lipid bilayer but requires assistance from vesicle fusion [38]. However, vesicle fusion may lead to multichannel insertion or an unknown channel insertion orientation.

As transmembrane proteins with known sequence and structure, the geometry or charged residues of biological nanopores could be slightly but precisely modulated by site-directed mutagenesis [45] for particular biosensing applications [46]. However, site-directed mutagenesis is not always successful and normally leads to significant efforts in the screening of desired mutants.

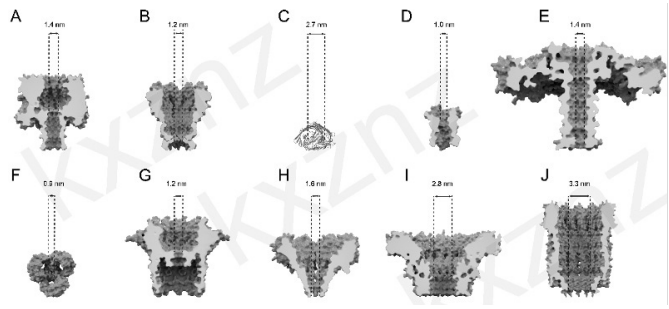


Figure 1.2 Crystal structures of different biological nanopores. (A) α-HL (PDB: 7AHL) [16]. (B) MspA (PDB: 1UUN) [47]. (C) FhuA (PDB: 1BY3) [48]. (D) OmpG (PDB: 2F1C) [49]. (E) Aerolysin (PDB: 5JZT) [50]. (F) OmpF (PDB: 2OMF) [51]. (G) CsgG (PDB: 4Q79) [43]. (H) FraC (PDB: 4TSY) [42]. (I) phi 29 connector (PDB: 1FOU) [52]. (J) ClyA (PDB: 6MRT) [53].

1.2 Methodology

1.2.1 Preparation and Engineering of Biological Nanopores

A key advantage of a biological nanopore over its solid-state counterpart is that it is extremely easy to prepare in a large quantity but with a low cost. The chemical nature of a biological nanopore can be manipulated with atomic precision and consistency. As has been reported, biological nanopores can be made in vitro using standard protein expression protocols [46] or in a prokaryotic expression system [36].

Taking α-HL as an example, the plasmid DNA containing the target gene encoding the protein monomer is custom designed and synthesized by commercial services (Genescript, New Jersey). This plasmid DNA could be further copied to develop the quantity and purity needed by a routine plasmid mini, midi, or maxi prep kit (Qiagen, QIAprep Spin Miniprep Kit).

A quick method of protein expression is by in vitro transcription and translation (IVTT). Following the standard protocol provided by the IVTT kit (TnT Quick Coupled Transcription/Translation System, Promega), ~500 ng of protein monomer can be biosynthesized in 2 h. The prepared protein monomer is then incubated with rabbit red blood cell membrane fragments at 37 °C for a further 2 h to assist protein oligomerization. The oligomerized protein sample is then purified and recycled by polyacrylamide gel electrophoresis and radioactive gel exposure. The whole process takes around 1.5–2 days to completion which is convenient for protein-screening purposes. However, the IVTT method has disadvantages of a low yield and the requirement of radioactive labeling during purification.

An alternative means of preparing biological nanopores is by prokaryotic expression followed with purification by fast protein liquid chromatography. In this approach, a tag protein such as His-tag [54] or Strep-tag [36] is normally placed on either terminus of the target protein for later purification purposes. The prokaryotic protein expression system can efficiently produce several mg of protein in 3–4 days. However, if the target protein is cytotoxic to the host cell during expression, difficulties may ensue. This method also requires more complicated purification procedures to eliminate interferences from the background proteins generated by the host cell. Site-directed mutagenesis is sometimes needed to optimize the performance of the protein nanopore [55]. This can be accomplished by direct synthesis of the mutated gene or with a site-directed mutagenesis kit.

The prepared protein nanopore is dissolved in Tris-EDTA buffer at pH 7.0 along with detergent to avoid protein precipitation. If properly stored at -80°C, the biological nanopores can stay active for up to 10 years and it has been suggested that protein nanopore samples should divided into small aliquots (1 μL each in PCR tubes) for long-term storage so that freezing and thawing of the sample is minimized.

1.2.2 The Instrument and the Device

A complete nanopore measurement platform includes a data acquisition module, a noise-insulating module, and a measurement module. For the data acquisition, most laboratories rely on

commercial patch-clamp amplifiers (Axopatch 200B, Heka EPC 800, Elements Eone, or Chimera VC100) along with a compatible digitizer for analog/digital conversion and a computer for instrument coordination and trace recording. Some laboratories prefer custom-made amplifiers for high bandwidth measurement applications or ease of device integration [56].

In a nanopore measurement, the ionic current through a single nanopore is approximately 0–200 pA for biological nanopores or 0–10 nA for solid-state nanopores. The patch-clamp amplifier serves to convert analog current signal into analog voltage signal, which is then acquired by the digitizer in a binary format. The digitizer, which is either integrated with the patch-clamp amplifier (EPC 800, Heka) or provided as a separate component (Digidata 1550A1, Molecular Devices), may also be replaced with a data acquisition card (National Instrument USB-6210) or even a digital oscilloscope for cost saving or expandability purposes.

In our laboratory, we use an Axopatch 200B patch-clamp amplifier paired with a 1550A1 16-bit digitizer (Molecular Devices) with a dynamic input range of ±10 V. The 16-bit digitizer serves to divide the signal with 2^{16}-1 (65535) digitized bits, generating a resolution of 0.305 mV/bit ($20V/2^{16}$). The gain of the amplifier could be set higher to further improve the measurement resolution but at the expense of the dynamic range or vice versa. For biological nanopore measurements, the gain is set to 10× or higher, which provides a dynamic range of ±1 nA and an improved resolution of 0.0305 pA/bit, which is ideal for discriminating between minimally different molecular analytes.

The noise-insulating module is critical for isolation of external electrical and mechanical noises from a sensitive nanopore experiment. The radiative electrical noise can be effectively shielded if the measurement is carried out in a conducting enclosure (a shielding box or Faraday cage) electrically connected to the common ground of the patch-clamp amplifier. Mechanical vibrations from external sources such as music, walking, talking, and keyboard typing can also be transmitted into the recorded signal and appearing as low frequency fluctuations of the signal baseline. For optimum performance, in our laboratory, the Faraday cage is bolted on an optical table with air floating supports, which effectively isolates it from most mechanical vibrations above 2 Hz.

A minimal measurement module is composed of a pair of Ag/AgCl electrodes, a measurement chamber and other accessories. Pure silver wire 1.5–2 mm in diameter, 2 cm in length is first soldered to an electrical jump cable (maximum 10 cm) at one end, then surface polished to eliminate the oxidized layer before being immersed in NaOCl solution (1–5% (w/v) solution in water) for 3–5 h until a thick, dark gray colored layer forms over the silver surface. The other end of the jump cable is soldered to a male connector pin for an electrical connection to the head stage of the patch-clamp amplifier. The measurement chamber is normally made of electrical insulating Teflon or polyformaldehyde polymers with 2 chambers (*Cis* and *Trans*) of 50–1000 μL on each side to accommodate the electrolyte solution, for example, 1 M KCl. The two chambers are geometrically connected with an aperture ~1 cm in diameter and hold a chip containing a solid-state nanopore or a Teflon film for lipid bilayer formation in biological nanopore experiments. During the measurement, the electrically connected Ag/AgCl electrodes are separately immersed in the electrolyte solution on each side of the chamber to form a closed circuit and the data acquisition for a nanopore measurement can then commence.

There are also other accessories such as peltier stages (temperature controls), magnetic stirrers (analyte mixing), or syringe pumps (automated liquid exchange) to facilitate the measurement. However, turning off these instruments during the recording is suggested as they could introduce noise that can also be picked up by the preamplifier.

1.2.3 The Electrochemistry Mechanism

For any type of nanopore, two chambers filled with electrolyte solution are separated by an impermeable membrane containing a single nanopore, which is the only pathway by which ions and molecular analytes can be transported across the membrane. A pair of Ag/AgCl electrodes electrically extended from the patch-clamp amplifier are placed on each side of the chamber, in contact with the electrolyte solution in either side. A potential bias is applied from the patch-clamp amplifier to promote directional flowing of ionic species through the pore. The corresponding electrochemical reaction on the electrode surface is: $\mathrm{Ag(s)} + \mathrm{Cl^-} \leftrightarrow \mathrm{AgCl(s)} + \mathrm{e}$.

The gain of an electron on the cathode electrode results in the release of a Cl^- ion to the solution. On the other side, the gain of a Cl^- on the anode electrode results in the release of an electron to the electrical circuit and thus generates a sustainable flow of ion across the membrane. Conventionally, the chamber which is electrically grounded is defined as the *Cis* side, while the other one which is electrically biased is defined as the *Trans* side.

Although the pair of Ag/AgCl electrodes are relatively far (>2 cm) from each other, the potential drop between the electrodes is mostly in the vicinity and on the scale of the thickness of the membrane of the nanopore due to the large resistance (10^8 ohm/cm^2) of the membrane, the lipid bilayer for biological pores or synthetic membrane for solid-state pores. Thus, the electric field, which is opposed to the gradient of the electrical potential ($E = -\nabla\varphi$), becomes stronger as it moves closer to the nanopore restriction. The electrophoretic force acting on the analyte is proportional to the electric field ($F = Eq$) and is strong enough against the Brownian motion of the molecule only within a hemispherical area in front of the nanopore. Thus the analyte molecule, for example, ssDNA, has to achieve a high enough final concentration (>200 nM) so that a reasonable amount of the event may be observed.

1.2.4 The Nanopore Measurement

In a typical nanopore measurement, a voltage is applied across the membrane by a pair of Ag/AgCl electrodes. The two chambers filled with electrolyte solutions (i.e., 1 M KCl, 10 mM HEPES, pH 7.0) are separated by a ~30 μm thick Teflon film. An aperture ~100 mm in diameter, coated with hexadecane, is located in the center of the film and serves as a solid support for the lipid bilayer assembly. The lipid bilayer can be formed by pipetting the solution in either chamber "up and down" by forming Langmuir–Blodgett lipid monolayers on each side of the aperture. A triangular wave with a slope of 1 V/s is often applied to differentiate between a successfully formed bilayer and a permanent blockage caused by air bubbles or contaminants. Free nanopores in solution can spontaneously insert into the lipid bilayer forming the only electrical connection between the two electrolyte solutions. After the first pore insertion, the chamber should be exchanged, or perfused with fresh, nanopore free buffer to "wash away" excess biological nanopores and so avoid further pore

insertions. The buffer exchange process should be accomplished with care and patience since abrupt mechanical vibration could potentially cause bilayer rupture and rapid exchange of the electrolyte solution in the chamber could sometimes assist further pore insertions. A voltage applied across the nanopore causes ions to flow through the nanopore, establishing a measurable ion current. The whole process is detailed in the tutorial protocol, below.

1.2.5 Measurement Noise and Bandwidth

In the time domain, the acquired ion flow through an open nanopore generates a root-mean squared noise centered about a steady DC signal. The power spectrum density analysis shows a 1/f noise in the lower frequency regime (below ~300 Hz or depending on the pore materials) and a positive power dependence on frequency in the higher frequency regime (above ~300 Hz or depending on the pore materials) [57]. This increased noise contribution from the high frequency end limits the bandwidth of the nanopore measurement, and this means that at some point in the frequency domain, the noise contribution will be too significant for the signal to appear in the time trace. The main source of the noise in the high-frequency region is the membrane capacitance of the nanopore device. This can be optimized by reducing the membrane capacitance of a solid-state nanopore [58] or minimizing the lipid bilayer area of the biological nanopore. More effectively, the high-frequency noise can be eliminated by a low-pass Bessel filter if the event of interest is significantly slower than the cut-off frequency of the filter. For this reason, much effort has been paid by researchers to slow down the speed of analyte translocation in order to resolve the signals of interest.

Normally, the cut-off frequency should be greater than five times the inverse of the mean event time [59]. According to the Nyquist–Shannon sampling theorem, the sampling rate should be more than twice the frequency of interest. In practice, it is suggested to over sample 10 times to preserve the signal shape.

1.2.6 Data Analysis

Digitally acquired time traces are normally produced in an .abf binary format. During a typical DNA translocation experiment, there could

be more than thousands of translocation events to be analyzed for a single experiment condition, and it is necessary to perform the data analysis by computer. The Clampfit software accompanying the Axon 200B patch-clamp system is a general purpose data analysis program with some integrated statistical analysis functions. Routine data analysis such as I–V curve analysis, step recognition, and analysis of all-points histograms can be performed directly in Clampfit.

For single-channel traces containing pore blockage events, the key parameters to be extracted are the residual current I_b, the dwell time t_d, and the interevent interval t_{on}. An automatic search can be done with the "single-channel search" function of Clampfit and when the results are exported to professional plotting software such as Origin or Igor, they can be eventually plotted as event histograms. The event histogram of I_b normally fits well with a Gaussian function and t_d fits a Gaussian function with an exponential tail.

However, Clampfit is not always sufficient for complicated event extractions. Data from some nanopore blockage events such as sequencing [60], structural unzipping [61], or unfolding data [62] appear as a continuous step-like signal which can be difficult for Clampfit to distinguish. In these cases, writing a custom program (e.g., with Matlab, Python, R, or LabView) according to the characteristic shape of the events of interest for automated event extraction is recommended. There are also many useful online resources on www.thenanoporesite.com/nanopore-software.html available for free download.

1.3 Applications

1.3.1 DNA Sensing and Sequencing

1.3.1.1 Free translocation of polynucleotide molecules

In 1996, Kasianowicz et al. first demonstrated the phenomenon of single-molecule transport of polyuridine oligomers through an α-HL nanopore with a continuously applied potential [19]. Subsequently, the crystal structure of the α-HL nanopore was reported. ssDNA translocation through the pore in a strictly single-file and sequential order was clearly indicated on a structural basis [16], and Kasianowicz and a few others proposed the concept of nanopore sequencing,

from which the sequence of current modulations induced by the DNA translocation could be aligned with its base composition. However, DNA sequencing still remains a challenge due to a lack of spatiotemporal resolution from the pore and the acquisition system.

First, the translocation speed of DNA, at 1–22 μs/nt [18, 63], is always too fast, and individual bases are hardly resolvable due to the limited bandwidth of existing patch-clamp amplifiers. Second, a single-nucleotide resolution cannot be achieved as the current blockades are found to be the consequence of ~10–15 nucleotides that are simultaneously accommodated within the long β-barrel of an α-HL nanopore [20]. Therefore, it has proved to be impossible to directly sequence DNA simply by analyzing translocation events acquired with α-HL nanopores.

1.3.1.2 Discrimination of immobilized nucleic acids in the biological nanopores

To enable prolonged measurements, ssDNA can be directionally immobilized and electrophoretically stretched within a nanopore by forming either a terminal hairpin [64] or a biotin-streptavidin complex [65]. Using this strategy, clear discrimination of single base substitutions of A, C, G in a homopolymeric strand of thymidine, was demonstrated using an α-HL nanopore [65]. Notably, the common epigenetic modifications in the DNA strand, such as 5-methylcytosines and 5-hydroxymethylcytosines, can also be distinguished [66], as can epigenetic ribobases [67]. However, the recognition sites of the α-HL nanopore are still insufficiently sharp to distinguish all bases in a diverse context.

MspA, which was pioneered by Gundlach et al., possesses a short constriction 1.2 nm in width and 0.6 nm in length, indicating an ideal geometry as a high-resolution DNA reader for nanopore sequencing. By performing static measurement with DNA homopolymers with a tethered streptavidin stopper [68], Gundlach et al. observed a signal discrimination with an order of magnitude larger than those found with the α-HL nanopore, which confirmed that MspA is superior to α-HL as the nanopore sensor for sequencing.

1.3.1.3 DNA strand sequencing

Many attempts were made to slow down the speed of DNA translocation, including control over the temperature [63, 69],

manipulation of the voltage [70], modification of analytes [71], and many others [15, 72]. Though one to two orders of magnitude decrement in the translocation speed were achieved, this was at the cost of either a largely suppressed channel current or a much reduced signal-to-noise ratio [15, 69, 73]. A desired DNA translocation modulation was recently enabled by coupling the DNA template with an enzyme motor adjacent to the pore orifice. With this strategy, a highly processive enzyme can ratchet along the DNA template one nucleotide at a time with a spacing of milliseconds and can replicate up to tens of thousands of nucleotides, promising a single-nucleotide pace of reading and a long read length during extended measurements.

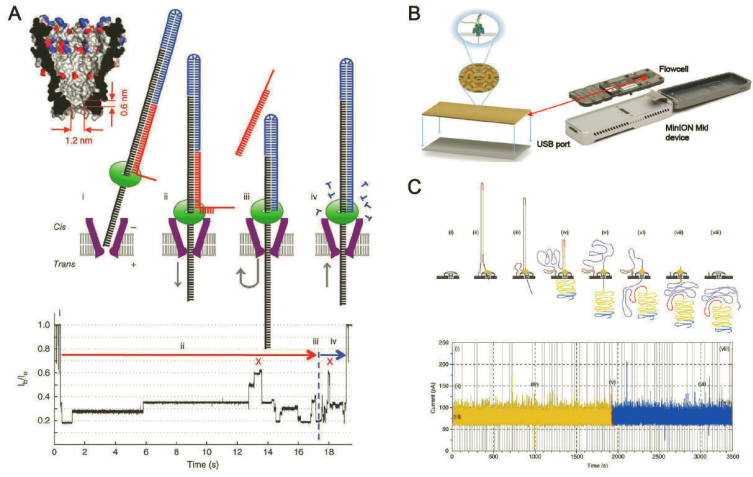


Figure 1.3 Nanopore sequencing. (A) A prototype of nanopore sequencing using an MspA nanopore. A phi29 DNAP, which serves as a motor protein, was employed to control the forward and reverse ratcheting of DNA templates through MspA. Sequence-dependent current levels were observed when the DNA template was translocating through the pore. Reproduced with permission from ref. [60], Copyright (2012) Nature Publishing Group. (B) The Minion nanopore sequencer. Each consumable flow cell generates as much as 512 nanopores for parallel DNA sequencing. Reproduced with permission from ref. [74], Copyright (1969) Elsevier. (C) Sequencing data reported from the MinION nanopore sequencer. Reproduced with permission from ref. [75], Copyright (2016) The Author(s).

Among the selection of the motor enzymes, the bacteriophage phi29 polymerase (phi29 DNAP), as one of the B-family DNAPs, has been shown to be superior for its remarkable processivity, which is ≥70