

Acquiring Biological Information of Individuals Using Quantum Mechanics

Masateru Taniguchi

The Institute of Scientific and Industrial Research, Osaka Univ.
8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
Phone: +81-6-6879-8445, E-mail: taniguti@sanken.osaka-u.ac.jp

Abstract

We developed a single-molecule technology that measures the tunneling currents induced by single-base molecules of DNA, RNA, and peptides passing between two nanoelectrodes. Unlike existing sequencing methods, our technique can potentially sequence chemically modified bases and amino acids, many of which are important disease markers.

1. Introduction

Many people expected that the end of the Human Genome Project would mark the starting point of personalized medicine and therapeutics based on genomic information. However, the cost and throughput to read out a human genome have been a substantial barrier to the realization of this technology. Our single-molecule electrical sequencing technology is based on semiconductor technologies and is expected to become an ultimate sequencer with ultra-low cost and ultra-high speed. Unlike existing sequencing technologies, single-molecule sequencing technology is expected to identify sequences of base molecules in RNA and sequences of amino acid molecules in peptides because it can identify small differences in molecular electronic structures through differences in single-molecule conductances. In addition, single-molecule sequencing technology may be able to detect chemically modified base molecules and amino acid molecules that are known markers of diseases such as cancer; there are no existing technologies detecting the chemical species directly. We hope to identify base molecules and chemically modified base molecules in DNA and RNA as well as amino acid molecules with and without chemical modifications using single-molecule sequencing technology.

2. Methods

All single-molecule electrical measurements were performed using nanofabricated mechanically controllable break junction (MCBJ) electrodes, in which a metal wire produced on an insulating substrate is fractured by pushing a piezo actuator up, forming gaps of less than several nanometers [1]. The nanoelectrode spacing was configured to be 0.8 nm for DNA and RNA and 0.5 nm and 0.7 nm for peptides because, unlike base molecules of DNA and RNA, the molecular sizes of amino acid molecules differ. Current-time profiles were measured at a sampling rate of 10 kHz under an applied potential of 0.4–0.7 V when a solution containing analytes was dropped onto the nanogap

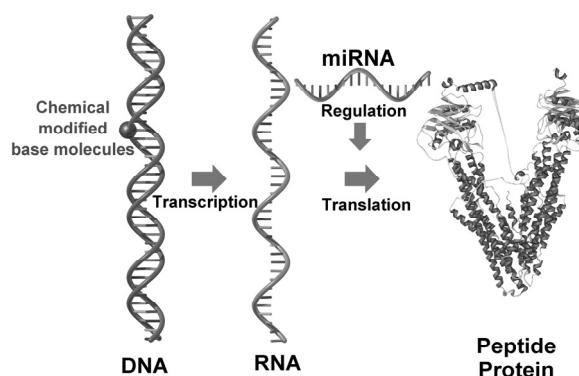


Fig. 1 The central dogma of molecular biology explains that DNA codes for RNA, which codes for proteins. Single-molecule sequencing technology can sequence DNA, RNA, and peptides and can identify chemically modified base and amino acid molecules.

electrodes. Sequences of base molecules and amino acid molecules were determined on the basis of current-time profiles using in-house analysis algorithms.

3. Results and Discussion

We first determined the single-molecule conductances of base molecules of DNA and RNA. When current-time profiles were measured, spike-like signals were observed and were characterized on the basis of their maximum currents and current duration. We explored the electron transport mechanism through single-base molecules to determine the voltage dependence of the electrical currents. We observed that the electrical currents increased linearly with increasing applied voltage, indicating that the electrical currents originated from tunneling currents between nanoelectrodes. The single-molecule conductance histograms showed single peaks, and the single-molecule conductance order in DNA was guanine (G) > adenine (A) > cytosine (C) > thymine (T), whereas that in RNA was G > A > C > U [2]. This result indicates that electric currents allow us to identify single-base molecules.

Because single-molecule conductances strongly depend on the highest occupied molecular orbital (HOMO) energies, single molecules with different HOMO energies can be distinguished even though the differences in HOMO energies are small. Although methylated cytosine (mC) is a well-known disease marker, no existing sequencing technologies can identify it directly. The single-molecule conductance of mC is larger than that of C because of a small difference of 0.1 V in their HOMO energies [3]. In addition

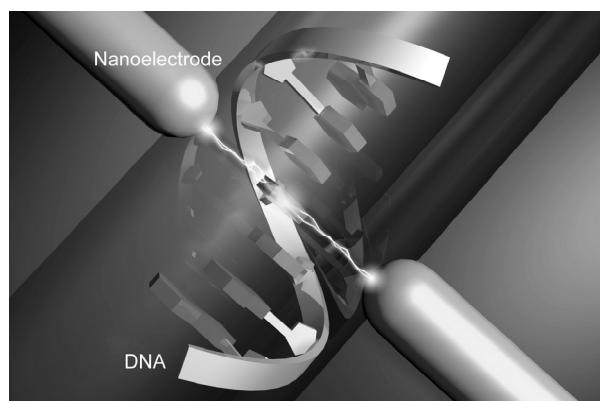


Fig. 2 Sequences of base molecules in DNA and RNA and sequences of amino acid molecules in peptides can be identified on the basis of single-molecule conductances of base and amino acid molecules, where tunneling currents between nanoelectrodes are measured.

to mC, oxo-G, which is a known cancer marker, can also be identified on the basis of single-molecule conductance. Therefore, our single-molecule identification method can detect chemically modified base molecules.

Next, we applied our single-molecule identification method to short DNA molecules such as TGT [4]. When the current–time profiles were measured, spike-like signals were observed in addition to the signals of single-base molecules. Two single-molecule conductance peaks corresponding to G and T were observed in the single-molecule conductance histograms. In the case of TGT, two lower-conductance plateaus were observed in the first and third lower positions, and the plateau associated with guanine was observed in the middle higher plateau. Therefore, we achieved single-molecule electrical sequencing based on single-molecule conductances. However, not only TGT but also G, T, and TG were observed because of stochastic traps caused by Brownian motions.

Using stochastic traps, we can expect to achieve a random sequencing method that is applicable to existing DNA sequencers. We attempted to perform single-molecule random sequencing of microRNA molecules, which are formed by 22–25 base molecules and are well known as cancer markers. As expected, different randomly fragmented sequences were obtained and assembled into sequence contigs, and sequence contigs were assembled into continuous sequences. Therefore, we successfully identified sequences of microRNA, which cannot be sequenced directly by existing DNA sequencing technologies.

Can single-molecule electrical sequencing technology be applied to identify sequences of amino acid molecules in peptides? Peptides are biomolecules formed by 20 types of amino acid molecules and contain approximately 50 amino acid molecules held together by peptide bonds. Sequences of amino acid molecules are very important in the determination of peptide functions, although peptides do not exhibit these functions until one or more amino acid molecules are modified chemically. This chemical modification is

referred to as post-translational modification, and it enables peptides to perform primary biological functions. The phosphorylation of tyrosine is a particularly well documented post-translational modification and is closely related to cancers and metabolic diseases. Here we focused on sequencing peptides and identifying the phosphorylation of tyrosine.

We measured current–time profiles of the 20 amino acids using MCBJ electrodes to obtain single-molecule conductances [5]. In addition to DNA and microRNA signals, spike-like signals were also observed; the single-molecule conductance histograms corresponding to the different amino acid molecules showed easily distinguishable peaks. Consequently, 12 amino acid molecules and phosphotyrosine were identified on the basis of single-molecule conductances. Next, we performed electrical measurements of the original peptide and the phosphorylated peptide. On the basis of the single-molecule conductance measurements of individual amino acids, we assigned the three measured peaks to tyrosine, phenylalanine, and phosphotyrosine. The partial sequences of amino acids comprising the peptides could be read from their conductance–time profiles.

4. Conclusions

Single-base molecules of DNA and RNA and single amino acid molecules of peptides can be identified using tunneling currents between nanoelectrodes. Single-molecule sequencing technology can determine sequences of base molecules in DNA and RNA and partial sequences of amino acid molecules in peptides. In addition, chemically modified base molecules and amino acid molecules can be distinguished via single-molecule conductances. The results of our experiments make us hopeful of using this platform to develop a high-throughput, ultra-fast, and inexpensive system for driving DNA sequencing beyond the limitations of current technologies.

Acknowledgments

This research is partially supported by the Japan Society for the Promotion of Science (JSPS) through its Funding Program for World-Leading Innovative R&D on Science and Technology, and KAKENHI Grant No. 26220603.

References

- [1] M. Tsutsui, K. Shoji, M. Taniguchi, and T. Kawai, *Nano Lett.* **8** (2008) 345.
- [2] M. Tsutsui, M. Taniguchi, K. Yokota, and T. Kawai, *Nat. Nanotechnol.* **5** (2010) 286.
- [3] M. Tsutsui, K. Matsubara, T. Ohshiro, M. Furuhashi, M. Taniguchi, and T. Kawai, *J. Am. Chem. Soc.* **133** (2011) 9124.
- [4] T. Ohshiro, K. Matsubara, M. Tsutsui, M. Furuhashi, M. Taniguchi, and T. Kawai, *Sci. Rep.* **2** (2012) 00501.
- [5] T. Ohshiro, M. Tsutsui, K. Yokota, M. Furuhashi, M. Taniguchi, and T. Kawai, *Nat. Nanotechnol.* **9** (2014) 835.