Physical DNA sequencing: codon thermoelectric signature

Enrique Maciá

Dpto. Fisica de Materiales, Facultad CC. Fisicas, Universidad Complutense de Madrid,
E-28040, Madrid (Spain)
emaciaba@fis.ucm.es

The proliferation of large-scale DNA-sequencing projects for applications in clinical medicine, health care and criminal research has driven the quest for alternative approaches to the commonly used Sanger sequencing method in order to reduce time, error rate and cost. Since Sanger's method relies on chemistry to read the bases guanine (G), adenine (A), cytosine (C), and thymine (T) in DNA, this quest has spurred new perspectives in nanotechnology looking for sequencing methods entirely based on physical differences between the bases for non invasive detection of nucleotides along the DNA strands [1]. Measuring transverse (i.e., perpendicular to the helix axis) tunnel currents through single-stranded DNA as it translocates through a nanopore has been proposed as a suitable physical method enabling single-base resolution [2,3], and it has experimentally been shown that the four bases provide a distinguishable transverse electronic signature when measured with a scanning tunnel microscope which directly detects the molecular levels of single DNA bases [4]. On the other hand, improvements of current nanotechnology allow us to confidently measure thermoelectricity at the molecular level as well. Thus, the thermoelectric properties of molecular junctions containing different benzene related moieties chemically bond to gold nanocontacts has been investigated in a series of experiments with a suitably modified scanning tunneling microscope [5,6]. Positive values of the Seebeck coefficient (indicating that the Fermi level is closer to the highest occupied molecular orbital (HOMO) level) were obtained for all considered molecules when contacted through thiol groups, indicating that the charge carriers are holes in this case. On the contrary, a negative value is obtained for a benzene molecule contacted to gold electrodes with cyanide-groups. Thus, end-groups are key to controlling the very nature of charge carriers and by properly varying end groups and molecular junction constituents, one can engineer metal-molecule heterostructures with targeted thermoelectric properties.

Motivated by these experimental results in this talk we will review previous works proposing the possibility of sequencing short DNA fragments by employing thermoelectric measurements [7-11]. The working hypothesis inspiring our proposal is the following. The basic unit of information in the genetic code is the so-called codon. A codon is an ordered sequence of three consecutive nucleotides that specifies a particular amino acid in a protein or initiation (stops) sites where translation into protein synthesis begins (ends). From the viewpoint of condensed matter physics each codon is characterized by its electronic structure, stemming from orbital overlapping among neighboring nucleobases. Therefore, the resulting electronic structure provides a characteristic spectral portrait, also determining the codon transport properties. Then, one may regard charge carrier propagating through the oligonucleotide as a physical probe sensing its electronic structure as a whole. In this way, rather than a one-by-one nucleotide reading, typical of chemistry based techniques, we will be able to directly sensing triplet nucleobases associations (including codons in codifying regions) at once. In order to substantiate this approach, we shall analyze the thermoelectric spectral curves of codons trapped between appropriate contacts at different temperatures, determining their characteristic thermoelectric signature. Due to the extreme sensitivity of thermopower to finer details in the codon-electrode electronic structure the thermoelectric response of trimer nucleobases forming a molecular junction exhibits several narrow resonant features where the Seebeck coefficient attains very large values (200-2000 μV/K at room

temperature). The position of these peaks sensitively depends on the characteristic electronic structure of the considered trimer, hence providing a very accurate method to properly identify different codons of biological interest [12].

References

[1] M. Zwolak and M. Di Ventra, Rev. Mod. Phys. **80** (2008) 141.

[2] J. Lagerqvist, M. Zwolak, and M. Di Ventra, Nano Lett. **6** (2006) 779; M. Zwolak and M. Di Ventra, Nano Lett. **5** (2005) 421.

[3] H. W. Ch. Postman, Nano Lett. **10** (2010) 420.

[4] M. Xu, R. G. Endres, and Y. Arakawa, Small **3** (2007) 1539.

[5] P. Reddy, S. Y. Jang, R. A. Segalman, and A. Majumdar, Science **315** (2007) 1568; K. Baheti, J. A. Malen, P. Doak, P. Reddy, S. Y. Jang, T. D. Tilley, A. Majumdar, and R. A. Segalman, Nano Lett. **8** (2008) 715.

[6] A. Tan, S. Sadat, and P. Reddy, Appl. Phys. Lett. **96** (2010) 013110.

[7] S. Roche and E. Maciá, Mod. Phys. Lett. B **18** (2004) 847.

[8] E. Maciá, Nanotechnology **16** (2005) S254; E. Maciá, Rev. Adv. Mater. Sci. **10** (2005) 166.

[9] E. Maciá, Phys. Rev. B **75** (2007) 035130.

[10] E. Maciá, “Aperiodic Structures in Condensed Matter: Fundamentals and Applications” (CRC Press, Boca Raton, FL, 2009).

[11] E. Maciá, Phys. Rev. B **82** (2010) 945431.

[12] Y. Dubi and M. Di Ventra, Rev. Mod. Phys. **83** (2011)