Insights into protein sequencing with an α-Hemolysin nanopore by atomistic simulations

G. Di Muccio[a], A.E. Rossini[b], D. Di Marino[c,d], G. Zollo[b], M. Chinappi[a]

[a] Dipartimento di Ingegneria Industriale, Università di Roma Tor Vergata,
Via del Politecnico 1, 00133, Roma, Italia
[b] Dipartimento di Base e Applicate per l'Ingegneria, Università di Roma La Sapienza,
Via A. Scarpa 14–16, 00161, Rome, Italy
[c] Faculty of Biomedical Sciences, Institute of Computational Science - Center for Computational Medicine in Cardiology Università della Svizzera Italiana (USI)
[d] Polytechnic University of Marche, Department of Life and Environmental Sciences,
Via Brecce Bianche, 60131, Ancona, Italy.

e-mail: giovanni.dimuccio@students.uniroma2.eu
mauro.chinappi@uniroma2.it
daniele.di.marino@usi.ch

Keywords: (nanopore, protein sequencing, nanofluidics, molecular dynamics)

The number of proteins expressed by a human cell is about 100.000. This expression is very specific for each cell type and is related to the functioning and the state of health of the cell. Molecular diagnostics and biomarker discovery should benefit particularly from the single-molecule characterization of proteomes.
In this framework, over the past two decades nanofluidic devices and nanopore technology have emerged as cheap, fast, single-molecule, label-free alternative to standard analysis techniques. A widely used techniques exploit the ionic current flowing through the pore (resistive pulse), which is affected by the steric hindrance of the molecules inside the nanopore. This technique was already successfully employed for DNA sequencing [1]. The development of nanopore based fluidic devices for protein characterization would have relevant consequences in biosciences and diagnostics [2]. However, the extension of the resistive pulse approach to polypeptide sequencing is proving to be a challenging task, due to the intrinsic chemical complexity of polypeptides, and the concomitant electric and nanofluidics transport phenomena involved in the capture and translocation control inside the nanopore [3]. A first mandatory requirement for an effective nanopore based sequencing system is that the single residues have to be associated to different current levels. Here, we report an unprecedented systematic computational study based on all-atom simulations aimed at answer to this question for the α-hemolysin pore, the most widely employed pore for nanopore sensing. Our results show that pore clogging is not only affected by amino acid volume (as widely expected in the community), but also by hydrophobicity and, in particular, by amino acid net charge. Together with signal analysis from non-equilibrium runs, we also proposed a less computationally demanding approach based on the estimation of the equilibrium electrolyte occupancy profile inside the nanopore, that, in our opinion, given the increase in the computational power, can become a standard procedure for preliminary evaluation of the capability of several nanopores to distinguishing among different amino acids. Moreover, this method can be easily generalized to any other molecule to be analysed. Finally, we discuss the possibility to modify the α-Hemolysin nanopore, cutting a portion of the barrel region close to the trans side, to reduce spurious signals and, hence, to enhance the sensitivity of the nanopore.

[1] Branton, Deamer, Marziali, Bayley, Benner, Butler, …, & Jovanovich, Nature biotechnology 2010. 261-268.
[2] Restrepo-Pérez, Joo, and Dekker, Nature nanotechnology 13.9 (2018): 786
[3] Chinappi, and Cecconi, Journal of Physics: Condensed Matter 30.20 (2018): 204002

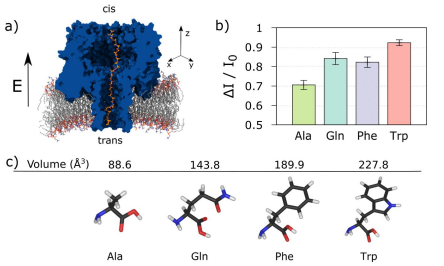

Figure 1. Ionic current measurements.
a) The system is constituted by an α−Hemolysin (blue) nanopore embedded into a lipid membrane (gray), with a 35-residues homopeptide (orange chain) is imported into the nanopore. A constant and homogeneous external electric field E = (0, 0, E_z) parallel to the pore axis is applied. b) Relative current blockage $\Delta I/I_0$, of the ionic current measured with the homopeptide inside the pore with respect to the I_0 empty pore value, for four different homopeptides, Ala, Phe, Gln, Trp. c) Molecular structure and Van der Waals volume of the four amino acids

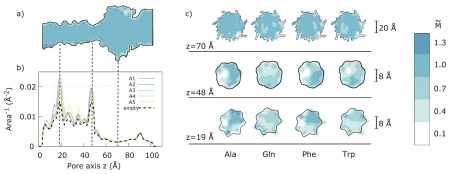

Figure 2. Electrolyte accessible volume estimation.
a) The panel report a cut of the 3D averaged electrolyte occupancy map for the empty pore. Blue areas corresponds to region that are fully accessible by the electrolyte while white ones do not contribute to the volume useful for the ions transport between the two side of the membrane.
b) Inverse of the accessible area, A_z , along the pore. The integral of this profile give an estimation of the electrical resistance of the pore in the different cases. c) Slices of the occupancy map normal to the pore z-axis passing through the two constrictions (z=19Å and z=48Å) and the vestibule (z = 70Å) for the four homopeptides.

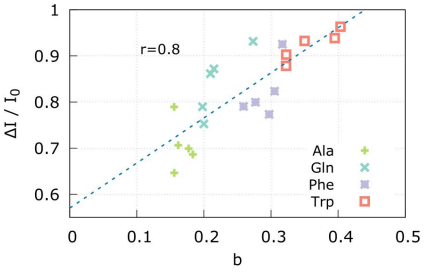

Figure 3. Pore clogging estimator b Vs measured current blockage $\Delta I/I_0$. Linear regression curve is reported in dashed blue, Pearson correlation coefficient r = 0.8.

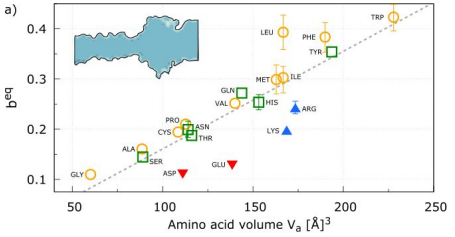

Figure 4. Pore clogging estimator b_{eq} for all residues Vs the amino acid volume V_a. Yellow circles corresponds to hydrophobic residues, green squares to polar, blue up-triangles to positively charged residues and red down-triangles to negatively charged ones. The dashed line is the minimum square fit.