# Time-series segmentation and latent representation of musical instruments

Gregory Szep

*King's College London*

August 20, 2018

**Abstract**

dd

## 1 scientific publication

Select a scientific publication about non-equilibrium molecular dynamics of biomolecules and discuss its relevance respect to the course you attended and this research assignment.

## 2 Unfolding proteins with umbrella sampling

Every student has access to the raw data of umbrella sampling simulations for two systems. These umbrella sampling simulations have been ran having the aim to calculate the Potential of Mean Force (PMF) of the mechanical unfolding of two peptides: one being in the initial conformational state of an a-helix; the other being in the initial conformational state of a b-hairpin. Every student is supposed to calculate the PMF for both the systems

## 2.1 Describe the simulated systems and the type of performed simulations you analyze

## 2.2 Calculate and display (plot) the PMF for the two systems. Explore the relevance of the bootstrap parameter respect to evaluation of the PMF's error using the script $\mathrm{g}_w ham_s cript_n ew.x$

Bootstrapping is a resampling method that quantifies statistical uncertainty by dividing the data into $N$ subsets, hence without additional measruements uncertainty is obtained by *pulling the data up by its own bootstraps*. By default the `g_wham` code forms subsets with replacement over the complete histograms along the reaction coordinate [1], leaving us with $N$ as a hyperparameter.

This hyperparameter is subject to the bias-variance trade-off. Chooseing a large $N$ results in relatively unbiased and uncertain estimates, while a small number of subsets give rise to more certain yet biased estimates. The former is preferred but comes with a computational cost. It appears that the uncertainty bounds converge to a fixed value after around $N \geq 20$ so there is little sense in setting $N$ at any other value than $N = 20$.

## 2.3 Quantify and discuss the histograms' overlap as obtained from the analysis of the umbrella sampling simulations

# 3 Conductance of a porin

## 3.1 Structure and function

Porins are $\beta$-barrel transmembrane proteins that act as a pore through which molecules can diffuse. Outer membrane protein F (ompF) is a nonspecific weakly cationic selective porin found in *E. Coli* [2]. It is nonspecific in the sense that there is no preference for a particular chemical species, however it does exhibit weak selective permeability for cations.

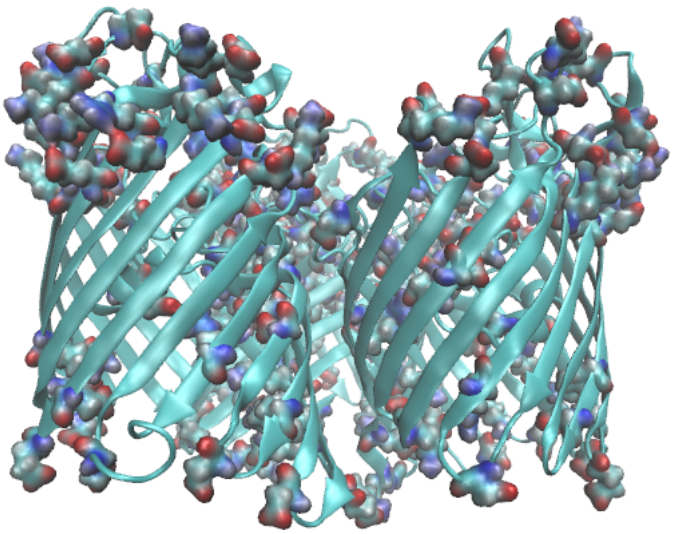

Figure 1: Side-view of ompF: $\beta$-barrel homotrimer forms transmembrane channel. Highlighted are polar residues, which dominanate the outer membrane side.

Figure 1 shows the structure of ompF: antiparallel $\beta$-strands forming three barrels that interact locally with neighboring $\beta$-strands, a long loop inserts into each barrel resulting in channel narrowing, charged residues form clusters at the entrance to a pore on the outer membrane side as well as within the narrow channel zone.

The narrowing geometry prevents larger potentially harmful molecules such as detergents from diffusing across the membrane, while allowing small hydrophilic molecules involved in bacterial metabolism to pass. Charged residues in the narrowing region of each barrel influence the ion selectivity and permeability of the channel. Two clusters of oppositely charged residues as shown in Figure 2 generate a screw-like electric field twisted along the channel axis with a strong transverse component in the pore narrowing region [2]. This field favors cations to pass through the channel.

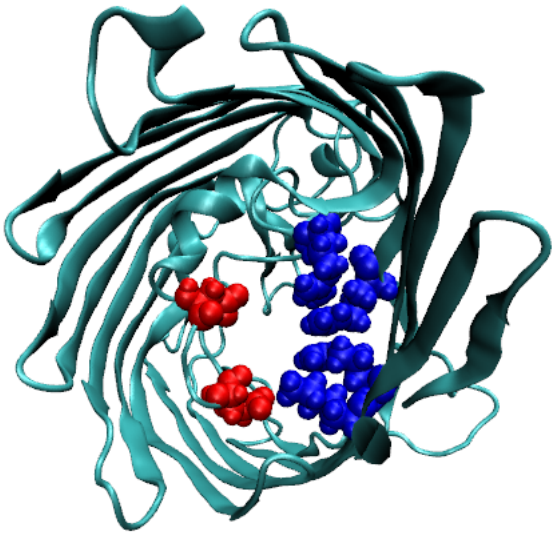

Figure 2: ompF $\beta$-barrel from inner membrane side having residues with positive COOH and negative $NH_2$ partial charged groups that facilitate cationic selectivity

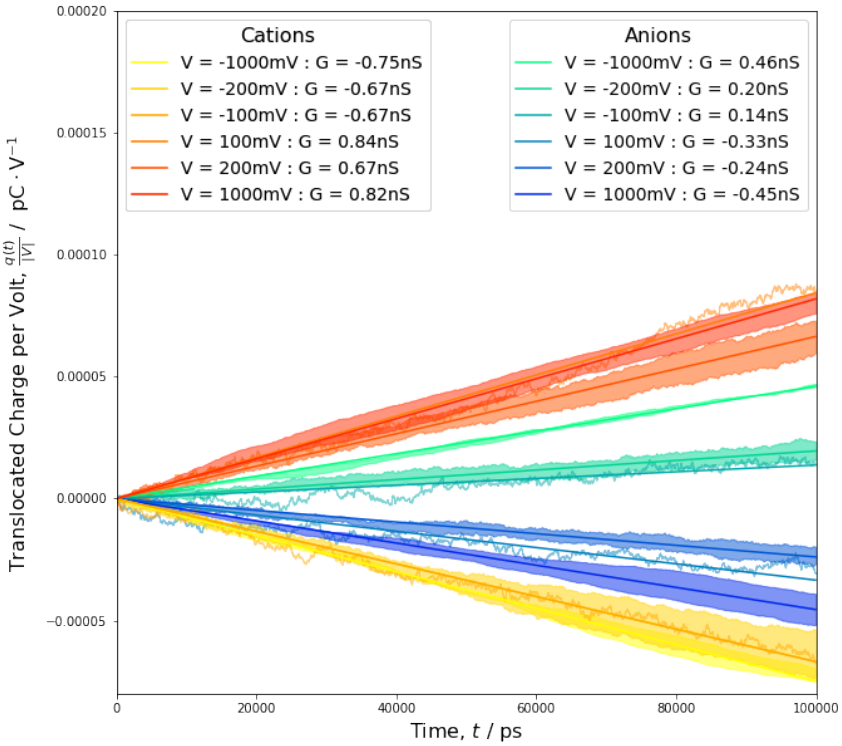


Figure 3: Conductance estimates $G$ from ensemble trajectories for anions and cations across ompF under various applied voltages $V$

### 3.2 Conductance estimates from molecular dynamics

Classical force field molecular dynamics with an extental electric field was used to estimate the conductance of ompF. The structure of ompF was embedded in a lipid bilayer, surrounded by a water, Potassium $K^+$ and Chloride $Cl^-$ ions. Simulations were run for up to 100ns with external fields given by $V = \pm 100\text{mV}, \pm 200\text{mV}, \pm 1\text{V}$. The translocated anionic and cationic charge $q(t)$ was recorded as a function of time. Several trajectories for a given paramerer set where run to obtain a mean and uncertainty. The cationic/anionic conductance $G_\pm$ was estimated via least squares according to the formula

$$q(t) = G_\pm |V| t \tag{3.1}$$

Results summarised in Figure 3 show mean and standard deviation for each trajectory ensemble. The claim of cationic selectivity [3] is supported as $|G_+| > |G_-|$ for all estimates and applied voltages. Estimated conductances also share the same order of magnitude as those reported in literature [3].

# References


[1] J. S. Hub, B. L. de Groot, and D. van der Spoel, "g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates," *Journal of Chemical Theory and Computation*, vol. 6, pp. 3713–3720, dec 2010.

[2] O. D. Novikova and T. F. Solovyeva, "Nonspecific porins of the outer membrane of Gram-negative bacteria: Structure and functions," *Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology*, vol. 3, pp. 3–15, mar 2009.

[3] R. Benz, A. Schmid, and R. E. Hancock, "Ion selectivity of gram-negative bacterial porins.," *Journal of bacteriology*, vol. 162, pp. 722–7, may 1985.