Molecular dynamics at constant pH in AWSEM coarse-grained model

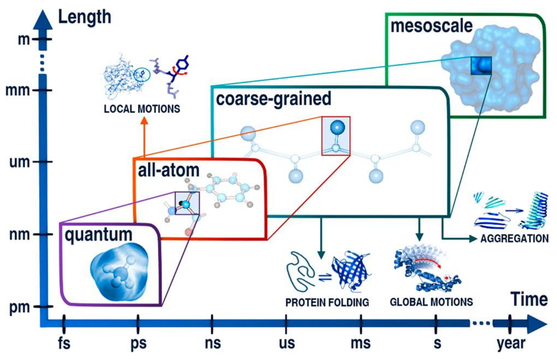
Franco Tavella[1] , E. A. Román[2] .

[1] Instituto de química biológica - Facultad de Farmacia y Bioquímica
[2] Instituto de Química y Fisicoquímica Biológicas (IQUIFIB)

e-mail: tavellafran@gmail.com

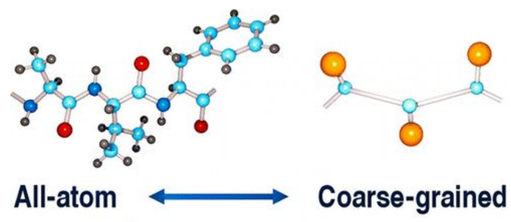
Coupling between global motions and electrostatics can be studied through coarse grained force fields

Solution pH is a critical factor for many biological processes. Lately there has been a development of techniques for including it into **molecular dynamics** simulations. Most methods are based on an **all atom representation** of the protein. This constraints the **time scales** that can be explored because of the **computational cost**



Sebastian Kmiecik et al.

We aim to study phenomena that occur in timescales about **milliseconds and seconds** involving **large conformational changes**. Specifically we want to study the **coupling** of this motions with the protein's **electrostatic interactions.**
Coarse grained force fields offer a suitable option for this purpose.



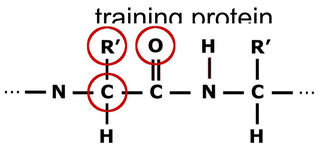
Our objective is to develop a constant pH molecular dynamics scheme in a coarse-grained force field. Throughout the simulation residues' discrete protonation states will be changing to model the effect of solution pH.

Associative memory, water mediated, structure and energy model (**AWSEM**) is a coarse grained force field that contains **physically motivated** energy terms, such as hydrogen bonding, as well as a **bioinformatically based local structure biasing** term, which efficiently takes into account many-body effects that are modulated by the local sequence.

$$V_{total} = V_{backbone} + V_{contact} + V_{burial} + V_{HB} + V_{AM} + V_{DSB}$$

AWSEM uses a **three bead representation** per amino acid. Beads are placed in the alpha carbon, beta carbon and oxygen

The **key idea** behind AWSEM is to simultaneously sculpt deep folding funnels for multiple unrelated proteins, using the same set of parameters, which then produces a **transferrable** protein folding force field. The physical principle from **landscape theory** that drives the **optimization of the parameters** is the maximization of the ratio of folding temperature over glass transition temperature for each training protein



Debye Huckel electrostatics and Monte Carlo sampling for titration

Run classical molecular dynamics

Stop dynamics and **select** a random **ionizable residue**

Change its **charge** and calculate **new** electrostatic **energy**

If the **new state** has **less energy** the change **stays**.
If not accept the new state with **Monte Carlo probability**

Continue molecular dynamics

To evaluate the **change in energy** between protonation states we use a **compound model** approach of the residues behaviour

$$\Delta E = \underbrace{k_B T ln(10)(pH - pKa)}_{\text{Solvation behaviour of the residue alone in solution}} + \underbrace{\Delta E_{elec}}_{\text{Interactions with other ionizable residues}}$$

The residue has a **reference pKa** alone in solution and it is **perturbed** because of the **electrostatic interaction** with the other ionizable residues in the protein

We used the following reference pKas:

ASP	4.0	**ARG**	12.0		
GLU	4.5	**HIS**	6.3		
LYS	10.6	**Nter**	7.5	**Cter**	3.6

And pdbs:

NTL9	1cqu	**HEWL**	2lzt	**BBL**	1w4h

Electrostatic interaction was considered between beta carbons in the AWSEM representation using a **Debye Huckel** potential

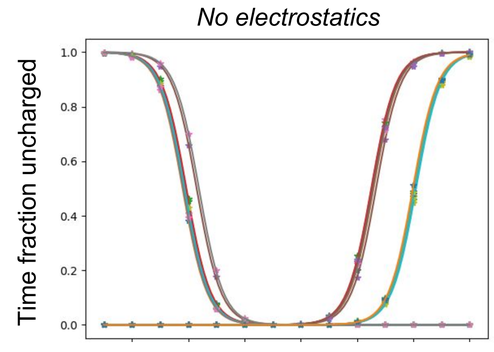
$$V_{DH} = K_{elec} \sum_{i<j} \frac{q_i q_j}{\varepsilon_r r_{ij}} e^{-r_{ij}/l_D}$$

This takes into account in a simplified way the **screening** of the Coulomb's law due to the presence of ions in solution. We used a 10 Å screening length corresponding to a 100 mM salt concentration

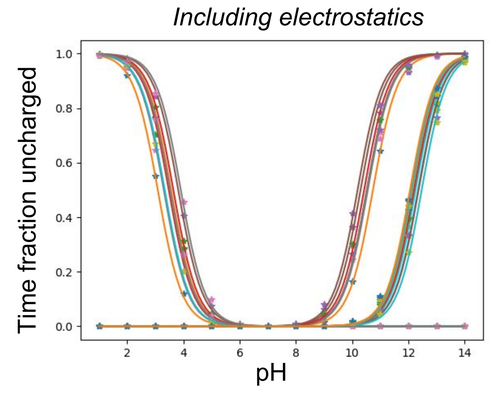
$$l_D = \sqrt{\varepsilon_r \varepsilon_o k_B T / 2e^2 I}$$

To test the implementation of the constant pH simulation scheme we focus on obtaining the correct shifted pKas of three model proteins: NTL9, HEWL and BBL. For this purpose, we calculate the simulation time a residue stays uncharged and fit a Henderson Hasselbalch titration behaviour

No electrostatics



Including electrostatics

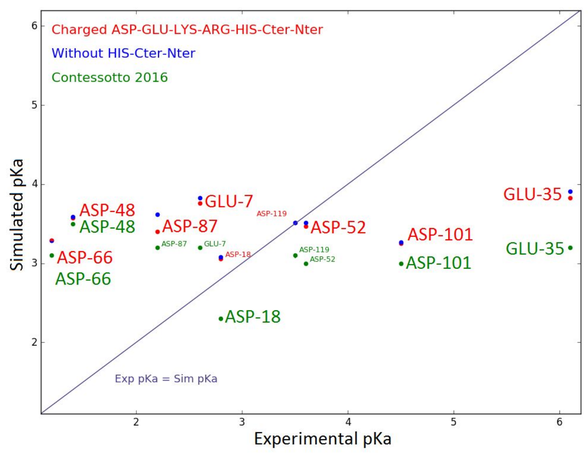


Comparison of two runs of 500.000 steps with and without electrostatic interactions and protonation state sampling

Results

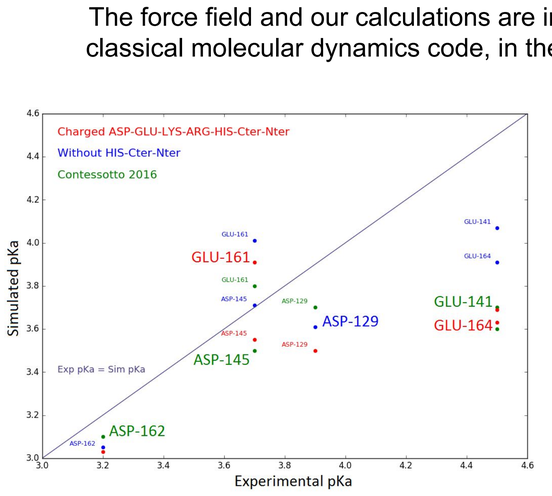
We compare our results with those of **Contessoto et al. 2016** who used a **single bead** representation, an analogous **Monte Carlo sampling** and a **native state guided** force field.
In our dynamics we used a **native structure biasing** term of the entire protein (**single memory**) to get the correct folding.

The force field and our calculations are implemented in **LAMMPS**, the classical molecular dynamics code, in the C++ programming language



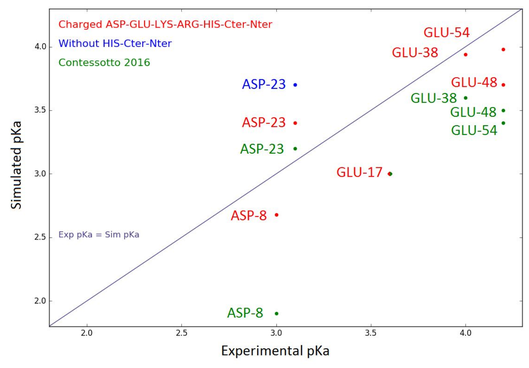
Results for HEWL protein

As in BBL the general trend of the residues is to stay in a pKa around 3.5. This doesn't agree with the experimental shifts observed in HEWL. Specifically GLU-35 and ASP-66 are buried in the protein. This suggests that to improve our model we should take the information of surrounding non-polar residues into account



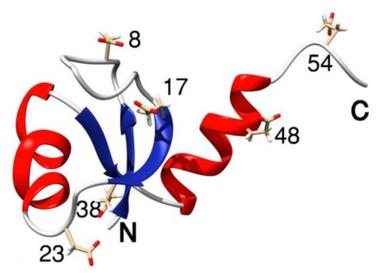
Results for BBL protein

We can see that our results don't capture the essential pKa shifts and get better if we remove HIS-Cter-Nter in some cases. This indicates that we are not treating properly their contribution and other effects aside electrostatics should be considered



For **NTL9** the pKa shifts are better obtained with 3 bead representation against a single bead one. Contrary to BBL and HEWL we have a good agreement. We think this is mainly because NTL9 residues are very exposed and not suffer shifts caused by hydrophobic environments.

NTL9 representation from Contessoto et al. 2016.
Most residues are exposed and shift their pKas mainly because of the protonation state of close charged groups



Conclusions and future

We implemented a constant pH molecular dynamics algorithm using the coarse-grained AWSEM force field. It has the advantage of being transferrable and also taking into account bioinformatic biasing. Results show that a simplified model considering only electrostatic perturbations captures the behaviour of non buried residues. For the buried ones another term in the Monte Carlo evaluation should be considered to account for the effects of hydrophobic environments.