Binding Free Energy and Ligand Orientation Calculations using A Monte Carlo Method with Markov Sate Analysis

Daniel Lecina-Casas[1,*], Ryoji Takahashi[1], and Victor Guallar[1,2]
1- Barcelona Supercomputing Center, c/Jordi Girona 29, 08034 Barcelona, Spain
2- Instituciò Catalana de Recerac i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Spain
daniel.lecina@bsc.es

Abstract *- Computing binding free energies has great implications in drug design. Using PELE technique, it has been shown that one can get quick and accurate estimations by means of a Markov state model[3]. We improved our methodology to compute faster binding relative free energy differences, mainly by analysis reducing the sampled region. This possibility opens a way in all-atom drug lead optimization by efficiently scoring a list of potential candidates in terms of binding affinities (approximately in 24hours), while still modeling accurately the protein-drug induced fit. Furthermore, we added information of the ligand orientation allowing us to obtain a better insight of the entrance mechanism. First, we show benchmark results - a series of benzamidine-like inhibitors in trypsin. Then, we apply our method to a more realistic scenario: the binding to a glucocorticoid receptor, and we show the performance for a new benchmark with a larger range of binding free energies (~14 kcal/mol). Simulations are obtained with our new in-house code PELE++, an improvement over the technique presented in references [1,2], (paper in preparation).*

I. METHODS

We first run simulations using PELE++. This allows us to obtain more data in a given computational time compared to regular MD.

Pele++ combines a stochastic approach with protein structure prediction algorithms. After every iteration, a step is accepted based on the Metropolis criterion. Pele++ can be used for free in our webserver[2] in https://pele.bsc.es.

We obtain the centroid position and quaternion of rotation from a reference structure for each snapshot.

For the analysis, we build a Markov State Model. To do so we can use for example EMMA (https://simtk.org/home/emma), or MSMBuilder (https://simtk.org/home/msmbuilder).

Quantitative binding free energy difference is obtained by means of:

$$\Delta G = -k_B T \ln(V_b/V_o) - \Delta W$$

To get a better mesoscopic insight, cluster-cluster analysis (e.g. PCCA+) is used.

We use the quaternion information to get quantitative results of the entrance and orientation in the binding site.

Fig. 1. PELE scheme. First, perturbation is performed by means of ligand perturbation followed by protein perturbation (ANM). Afterwards, the structure is refined by means of a side chain prediction and minimization.

PELE Scheme

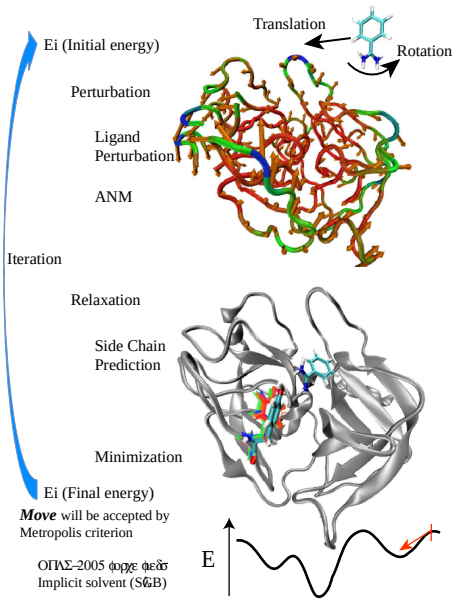

Fig. 1. PELE scheme. First, perturbation is performed by means of ligand perturbation followed by protein perturbation (ANM). Afterwards, the structure is refined by means of a side chain prediction and minimization.

II. RESULTS

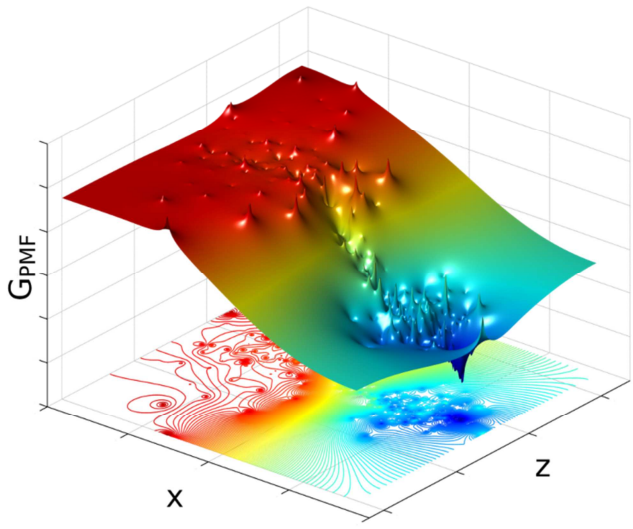

Fig. 2. Potential of mean field for a hormone receptor.

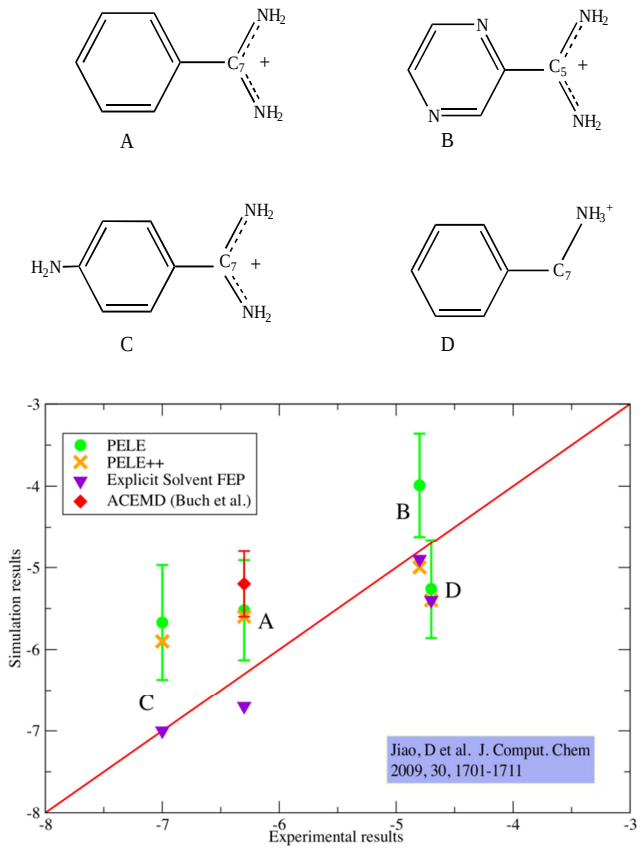

Fig. 2. Different trypsin inhibitors with their predicted and experimental binding free energies.

III. CONCLUSIONS AND FUTURE WORK

We are able to compute satisfactorily differences in binding free energy. Besides, we gained a better insight into the binding mechanism by means of the orientation analysis.

But some questions arose: What approach should be used if the ligand is completely flexible? How do we analyze the rotatable bonds? How much sampling do we need?

REFERENCES

[1] Borrelli, K. W., Vitalis, A., Alcantara, R. & Guallar, V. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. J. Chem. Theory Comput. 1, 1304-1311 (2005).

[2] Madadkar-Sobhani, A. & Guallar, V. PELE web server: atomistic study of biom. Nucl. Acids Res. 41, W322-W328 (2013).

[3] Takahashi, R., Gil, V. A. & Guallar, V. Monte Carlo Free Ligand Diffusion with Markov State Model Analysis and Absolute Binding Free Energy Calculations. Journal of Chemical Theory and Computation 10, 282-288, doi:10.1021/ct400678g (2014).