A differentiable representation of solvent-solute interface

Reporter: Zhenyu Wei

Mentor: Yunfei Chen

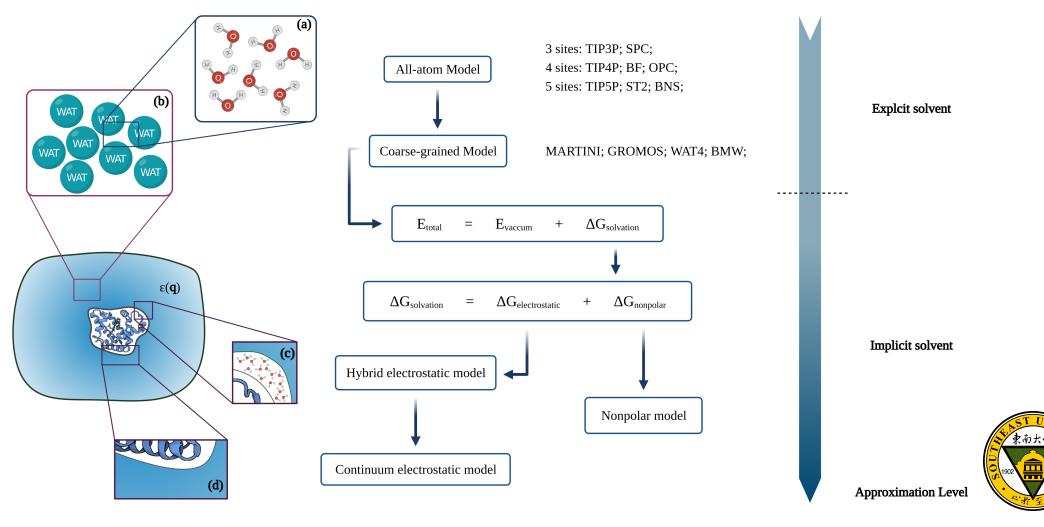
School of mechanical engineering Southeast University



Inspiration



Implicit solvent model



Zhenyu Wei, Sep 2021 3 / 14

Poisson-Boltzmann Equation

$$abla arepsilon(\mathbf{r})
abla \phi(\mathbf{r}) = -
ho_{mol}(\mathbf{r}) - \lambda(\mathbf{r}) \kappa^2 \mathrm{sinh}\left(-rac{z_+ e \phi(\mathbf{r})}{kT}
ight)$$

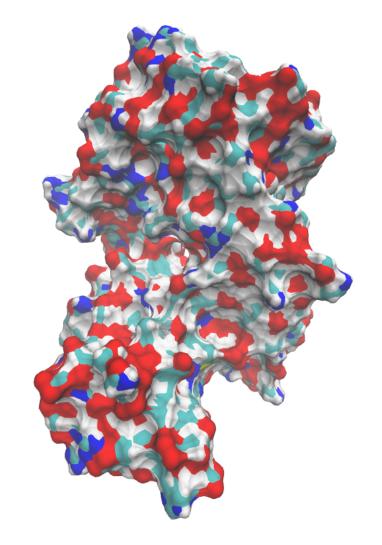
A **differentiable** expression of ε_r is vital for a precise solution of **Poisson-Boltzmann Equation** (PBE)



PBE for globular protein

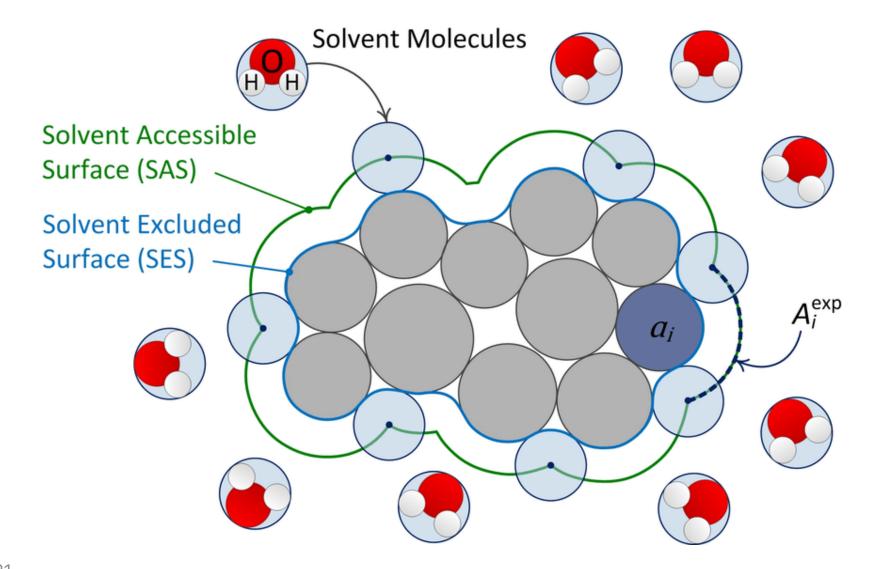
- The surface can be interpreted as a complex spatial distribution
- The spatial distribution is highly nonlinear
- The distribution depends on all of the atom's position and type

PDB id of protein shown right: 1A1N





Current solution: Van der Waals surface





Current obstacles

- Representation is not smooth
- Representation is not expressed explicitly
- Representation based on Hard sphere model
 - Highly approximated
 - Hyper parameter required

Target

A smooth, differentiable, interface representation for solution of PBE



Method



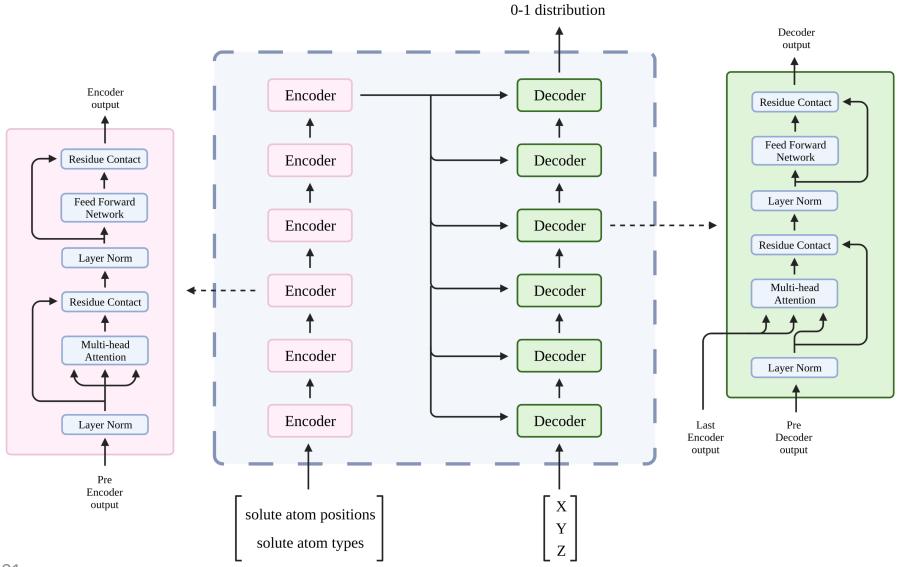
Basic Idea

- Using deep neural network to represent the non-linear interface
- Deep neural network are naturally differentiable

Obstacle

Handling length-varied input: the positions and types of protein's atom

Architecture





10 / 14

Dataset

- Downloaded 12748 structures from PDB website
- Patched and solvated **10962** structures in $60 \times 60 \times 60 \times 60$ A box
- Currently, sampled 4458 structures
 - Label solution atom as 1
 - Label solvent atom as 0



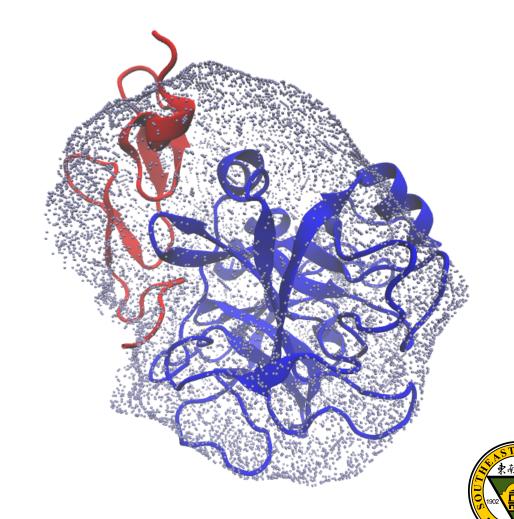
Result and discussion



Feasibility test

- Train network on tiny dataset
 - 28 protein structures
 - Each with 25 samples
- Average accuracy: 92%
- Isosurface with output value of TSSIR between [0.495, 0.505]

PDB id of protein shown right: 1IQG



Thanks for your attention

Q&A

