Generating Effective Charge (ECM) sites for small molecules

0.1 Introduction

Effective Charges for Macromolecules in solvent (ECM) are fitted charges in a uniform dielectric that can reproduce the electrostatic potential of the molecule computed with the use of all atomic charges in a heterogeneous dielectric. It is computationally very expensive to calculate electrostatic interaction free energy between a pair of solutes (protein-protein or protein-ligand association) at each time-step of the BD simulation, for realistic systems with thousands of atomic charges in an environment with a non-uniform dielectric permittivity and a solvent of non-zero ionic strength. SDA uses the Effective Charge Model (ECM)[1] to approximate the Poisson-Boltzmann (PB) theory derived electrostatic interaction. In this model, electrostatic interaction energy between a solute pair (ΔG_{el}^{1-2}) is calculated as the interaction between PB derived electrostatic potential (Φ_{el}) of one solute and a set of effective charges (q_i) on the other solute, and vice-versa. The total interaction energy is multiplied by a factor of 1/2 to prevent double counting of the interaction. These effective charges are fitted in such a way that, they reproduce the electrostatic potential derived with PB in in-homogeneous dielectric medium, when they are placed in an uniform dielectric medium. Determining the effective charge sites is relatively simple for proteins where the test charges are placed on the carboxylate oxygens of Asp, Glu, and the C-terminus, and the amine nitrogens of Lys, Arg, and the N-terminus. However, this approach does not work for chemical compounds and cofactors etc. For these small molecules, these charges are assigned on hydrogen bond donor-acceptor atoms (N,O,F,S), halogen atoms (Cl,Br,I) and on P and Fe atoms in case of co-factors using the *ECM_ ligand.py* python script.

$$\Delta G_{el}^{1-2} = \frac{1}{2}\sum_{i_1} q_{i_1}\Phi_{el_2}(r_{i_1}) + \frac{1}{2}\sum_{i_2} q_{i_2}\Phi_{el_1}(r_{i_2}) \tag{1}$$

Here, q_{i_n} is an effective charge on solute n and $\Phi_{el_m}(r_{i_n})$ is the electrostatic potential of solute m, at the position of effective charge q_{i_n} on solute n.

0.2 *ECM_ ligand.py*

This script picks effective charge sites for small molecules and assign appropriate test charges to them (see Figure 1 for the protocol). The script reads in the PQR file and

determines the net charge on the molecule. Then, the N, O, S, F, Cl, Br, I , P and Fe atoms present in drug like compounds and organic cofactors are marked as effective charge sites. The partial atomic charges of hydrogen atoms covalently bonded to these effective charge sites are added to the respective partial charge of effective charge site atom. Then, the charge difference between net charge of molecule and cumulative sum of partial charges of effective charge sites is calculated. This charge difference is then redistributed equally to all the effective charge sites, so that the total test charge on all effective charge sites is equal to the net charge of molecule. These updated test charges are then written to an output file with a specific format. This output file should be used as an input used for *ecm_ expand tool* program of SDA to generate ECM charges.

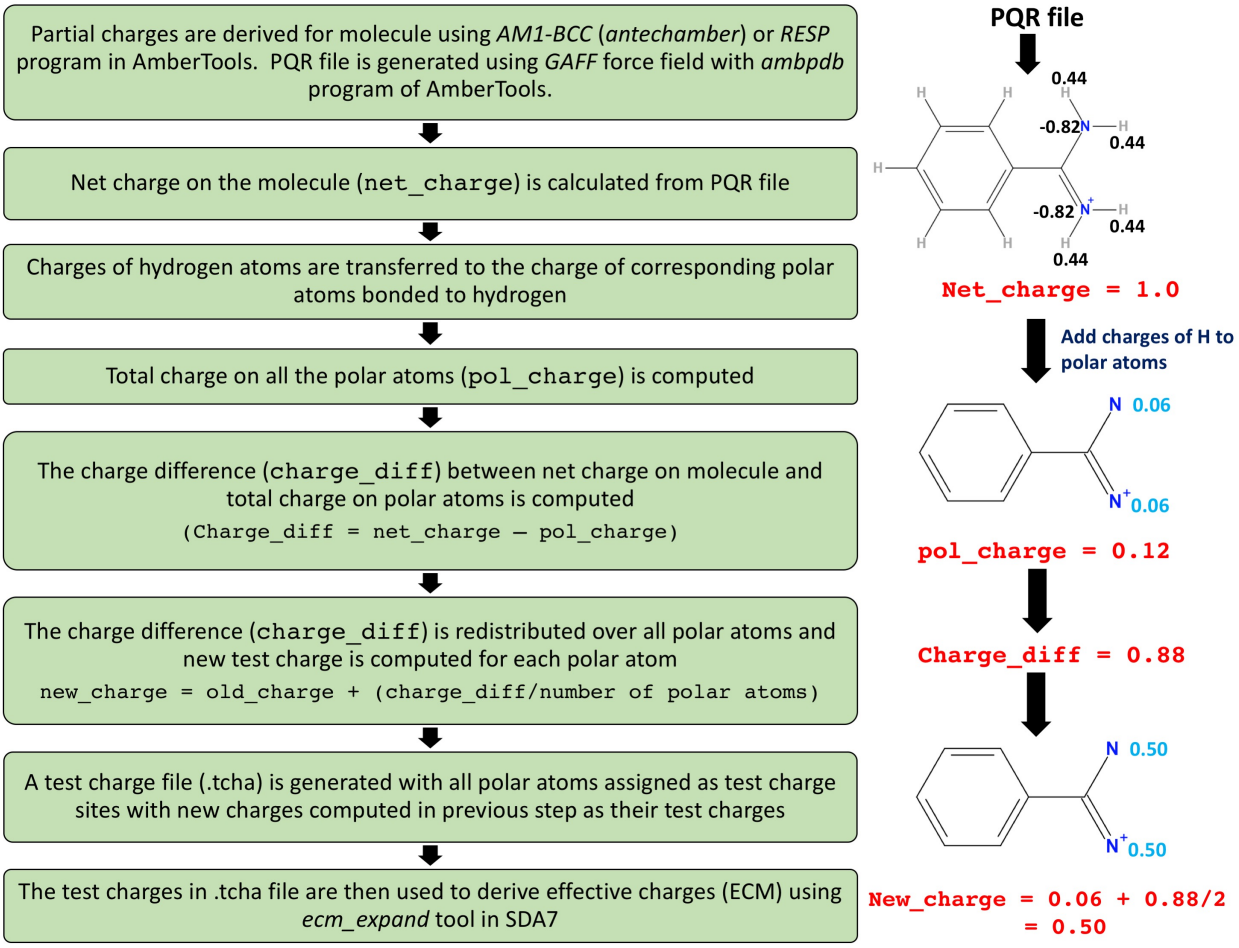

Figure 1: Algorithm to assign test charges for computing effective charges for small molecules.

0.2.1 Usage

```
python ECM_ligand.py ligand_pqr
```

Where, *ligand_pqr* is the Name of PQR file for the ligand with .pqr extension

0.2.2 Example Case

ECM sites and test charges for Rivaroxaban (an inhibitor of Human Coagulation Factor Xa) can be generated from its PQR file (*RIV.pqr*) using *ECM_ligand.py* as follows:

```
python ECM_ligand.py RIV.pqr
```

(**** Output file is generated in the same directory with name : RIV.tcha ****)

```
cat RIV.tcha
```

```
ATOM      1   Cl RIV     1      13.544   7.302  26.531         0.242
ATOM      3   N2 RIV     1       8.303   2.377  18.109         0.217
ATOM      6   O5 RIV     1       9.768   1.124  16.111        -0.130
ATOM     14  O13 RIV     1       6.583   2.613  16.630        -0.297
ATOM     15  N14 RIV     1       5.906   5.863  21.947         0.197
ATOM     18  O17 RIV     1       4.690   6.614  23.757        -0.091
ATOM     20  O19 RIV     1       4.634   4.419  23.228        -0.279
ATOM     22  N21 RIV     1       7.078   8.136  24.932         0.033
ATOM     24  S23 RIV     1      11.131   8.264  25.063         0.346
ATOM     29  O28 RIV     1       8.611   9.379  23.797        -0.239
```

Bibliography

[1] R R Gabdoulline and R C Wade. Effective Charges for Macromolecules in Solvent. *The Journal of Physical Chemistry*, 100(9):3868–3878, 1996.