

Practical on binding desolvation free energy calculation

Ting Wang

European Media Laboratory, Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg, Germany
24-05-2002

In this practical, we will show how to use the program UHBD (<http://adrik.bchs.uh.edu/ukbd.html>) to calculate the electrostatic component (ΔG_{ele}^{bind}) of the protein-ligand binding free energy, which is defined as in equation 1.

$$\Delta G_{ele}^{bind} = E_{ele}^{P-L} - \Delta G_{ele}^{desol-P} - \Delta G_{ele}^{desol-L} \quad (1)$$

E_{ele}^{P-L} is the electrostatic interaction energy between the protein and the ligand with binding formation. $\Delta G_{ele}^{desol-P}$ and $\Delta G_{ele}^{desol-L}$ are the electrostatic desolvation free energies of the protein and the ligand, respectively, which are defined as the loss of the electrostatic interaction energy between the solvent and the protein (or the ligand) upon binding.

We will use a two-stage procedure to calculate $\Delta G_{ele}^{desol-P}$, $\Delta G_{ele}^{desol-L}$ and E_{ele}^{P-L} , as shown in Figure 1.

- 1) calculate the electrostatic interactions between the protein (or the ligand) and the surrounding solvent in the absence of the ligand (or the protein). See the left two boxes in Figure 1.
- 2) calculate the electrostatic interactions between the protein (or the ligand) and the surrounding solvent with ligand (or protein) bound but in the absence of the partial charges of the ligand (or the protein). See the middle two boxes in Figure 1.

The electrostatic desolvation energy ($\Delta G_{ele}^{desol-P}$ or $\Delta G_{ele}^{desol-L}$) is the difference between the electrostatic energies computed from the two stages. The electrostatic potential generated by the charges of the ligand are assigned to each protein atom and the electrostatic interaction energy E_{ele}^{P-L} is the sum of the electrostatic interaction energy between the ligand and each protein residue (see equation 2).

$$E_{ele}^{P-L} = \sum^{Nres} \sum^{Natom} q\phi \quad (2)$$

q and ϕ are the atomic partial charge and electrostatic potential of each protein atom, respectively. $Nres$ and $Natom$ are the number of protein residues and the number of atoms in each protein residue, respectively. In COMBINE analysis, both $\Delta G_{ele}^{desol-P}$ and $\Delta G_{ele}^{desol-L}$ can be used as additional energy terms for QSAR model building. E_{ele}^{P-L} can

be decomposed on each protein residue to replace the electrostatic energies calculated from the Coulombic formula.

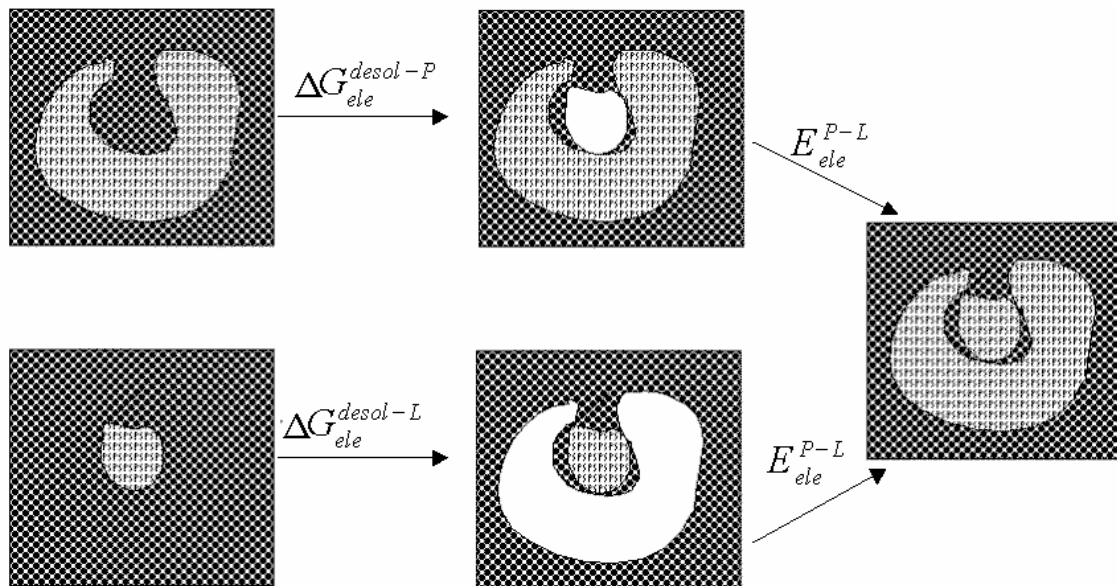


Figure 1. Schematic presentation of the steps of the binding free energy calculation for two solute molecules. The surrounding solvent is darkly shaded. The solutes are unshaded or lightly shaded to represent, respectively, the absence or presence of charges.

Here, we start with the minimized structure of the neuraminidase-inhibitor complex 2qwe with all water molecules removed (Refer to “Practical on Comparative Binding Energy (COMBINE) analysis”). The topology and coordinate files are generated by Xleap in AMBER7. You just copy them into your working directory.

```
% cp $COMBINE/2qwe_nw.tpp . /*copy the topology file */
% cp $COMBINE/2qwe_nw.rst . /* copy the coordinate file */
```

The input for UHBD is CHARMM qcd files of both protein and ligand. A qcd file is very similar to a pdb file but with two more columns of atomic charges and radii. Here, we use AMBER charges and radii by running the following command:

```
% $COMBINE/amber7toqcd 2qwe_nw.tpp 2qwe_nw.rst 2qwe_nw.qcd
/* generate CHARMM qcd for use in UHBD */

% grep 'GNA' 2qwe_nw.qcd >! GNA.qcd /* extract the ligand from the
complex file */
% egrep -v 'GNA|CAA' < 2qwe_nw.qcd >! 2qwe_prot.qcd /* extract
the protein from the complex file */
```

Now, you need to write a protocol input file telling UHBD how to do the calculations.

```
% cp $COMBINE/desol_110.inp . /* copy the input file */
% cp $COMBINE/step4_110.inp . /* copy the input file */
```

Take a look at both desol_110.inp and step4_110.inp.

The program UHBD6.1 is used to implement the continuum electrostatic calculations by solving the Poisson-Boltzmann equation using a finite difference method. The interior dielectric constant of both protein and ligand is set to 2 and the solvent dielectric constant is set to 78 with an ionic strength of 150 mM and ionic radius of 1.5 Å. A probe of radius 1.4 Å and a surface with 400 points per Å² is used to calculate the solvent accessible surface and define the dielectric boundaries. The coarse grid spacing is set to 0.80 Å and the fine grid spacing was set to 0.225 Å. Both the coarse grid and the fine grid are dimensioned to 110×110×110.

The two-stage procedure shown in Figure 1 will be implemented by the 4 steps defined in these two files. The first 3 steps will be run by desol_110.inp to get the desolvation energy of the protein ($\epsilon_{solv} = \epsilon_1 - \epsilon_2$) and the electrostatic interactions between the ligand and the surrounding solvent in the absence of the protein (ϵ_3). Step4_110.inp will run the 4th step to get the electrostatic interactions between the ligand and the surrounding solvent with protein bound but in the absence of the partial charges of the protein (ϵ_4), and the electrostatic potential of each protein atom will be written down.

Run with the first input file :

```
% $UHBD/uhbd110 < desol_110.inp >! 2qwe.desol_110
/* This takes ca. 14 min on R10000 SGI */
```

Check 2qwe.desol_110, you should get:

$$\Delta G_{ele}^{desol-P} = \epsilon_{solv} = \epsilon_1 - \epsilon_2 = -38.968 \text{ kcal/mol}$$

$$\epsilon_3 = .341517E+04 \text{ kcal/mol}$$

Run with the second input file:

```
% $UHBD/uhbd110 < step4_110.inp >! 2qwe.step4_110
/* This takes ca. 4 min on R10000 SGI */
```

Check 2qwe.step4_110, you should get :

$$\epsilon_4 = .3447371E+04 \text{ kcal/mol}$$

then:

$$\Delta G_{ele}^{desol-L} = \epsilon_3 - \epsilon_4 = 32.2 \text{ kcal/mol}$$

Now, we calculate the electrostatic interaction energy between the protein and the ligand.

```
% grep `PHI` 2qwe.step4_110 >! 2qwe.phi
/* extract the electrostatic potential of each protein atom from 2qwe.step4_110 */

% $COMBINE/elect 2qwe.phi 2qwe_prot.qcd 2qwe.elect
/* This step calculates the electrostatic energy between the inhibitor and each protein
residue and the total electrostatic energy between the inhibitor and the protein.*/
```

Take a look at 2qwe.elect, you should get :

```
Total elect energy is -74.5375 kcal/mol.
```

Then, the electrostatic component of binding free energy

$$\begin{aligned}\Delta G_{ele}^{bind} &= -74.5373 - (-38.968) - (-32.2) \\ &= -3.367 \text{ kcal/mol}\end{aligned}$$

Note:

In COMBINE analysis, it is sometimes appropriate to define the dielectric boundary as the van der Waals surface in the UHBD calculations. To do this, the line “nmap 1.4” should be commented out by changing to “!nmap 1.4”.