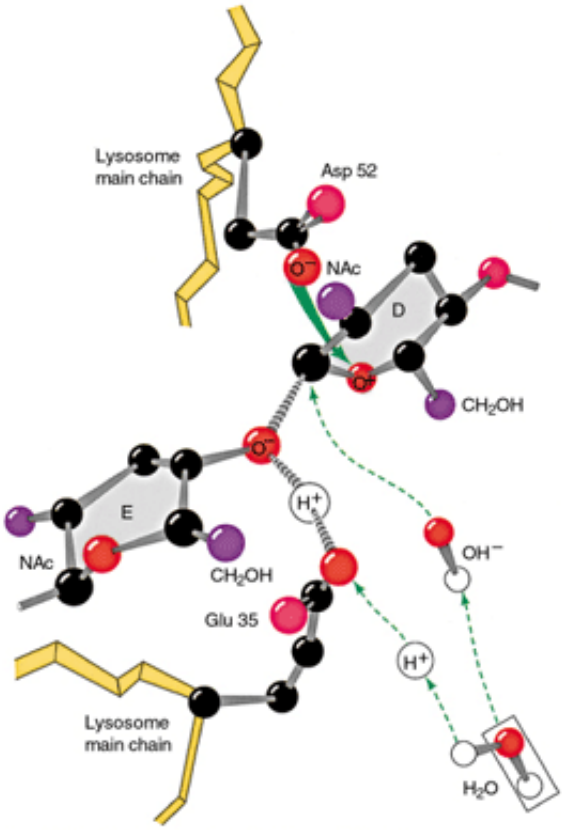
Lysozyme pKa example

Background

Hen egg white lysozyme (HEWL) is a very popular system for pKa calculations as it has a number of interesting values for its titratable residues. Early pKa work on this enzyme is presented in Tanford C and Roxby R (Interpretation of protein titration curves. Application to lysozyme. Biochemistry. 11 (11), 2192-8, 1972) which also contains the pKa values used in this example. More recent pKa calculations and a review of some of the methodology can be found in Nielsen JE and Vriend G (Optimizing the hydrogen-bond network in Poisson-Boltzmann equation-based pKa calculations. Proteins. 43 (4), 403-12, 2001). Finally, the biological relevance of lysozyme is briefly reviewed at Wikipedia.

HEWL has two active site residues GLU 35 and ASP 52 whose titration states determine the catalytic competency of the enzyme:



In particular, the enzyme is only active when ASP 52 is ionized (pKa ≈ 1.2) and GLU 35 is neutral (pKa = 6.3). Additionally, lysozyme has ASP 66 with a pKa of 2.0 and HIS 15 with a pKa of 5.7.

Overview and disclaimer

In what follows, we'll calculate the *intrinsic* pKa of ASP 66. This is the pKa calculated without regard for the titration state changes in other lysozyme residues. Calculating actual pKa changes requires analysis of the coupled titration state energetics of the entire system. This is an extremely labor-intensive process which is best left to computer programs such as WHATIF, PDB2PQR, MCCE, and related software.

Preparing the PDB file

Download the PDB entry 2LZT from the PDB; save it as `2LZT-ASP66.pdb`. If you have time, you should also visit the PDBSum

analysis page as well for additional information about the structure.

Warning: Problems with explicit water! It is very important that you remove all explicit water from the PDB file before proceeding. (Why?)

We're going to need to generate a protonated form of ASP 66 to perform our pKa calculations. We will do this with the PDB2PQR web server. Unless otherwise directed, PDB2PQR adds hydrogens to residues based on model pKa values. Therefore, we will need to specify the titration state of ASP 66 for our pKa calculation by changing the residue name from ASP to ASH using your favorite text editor. Save the result as `2LZT-ASH66.pdb`.

We're now ready to run PDB2PQR to generate protonated versions of our PDB files. Use the command line version of PDB2PQR or one of the web servers listed on the PDB2PQR homepage to generate protonoated PQR files for for `2LZT-ASP66.pdb` and `2LZT-ASH66.pdb`. Name your results `2LZT-ASP66.pqr` and `2LZT-ASH66.pqr`, respectively. Although it is always important to test sensitivity to various force fields, we recommend starting with PARSE.

Note: You can use PDB2PQR to assign titration states with PROPKA but don't do it for the above steps since we need to set the titration states explicitly for our calculations.

Recall that we're also going to need the isolated residue for our electrostatics calculations of intrinsic pKa. Use your favorite text editor to extract the entire ASP 66 and ASH 66 residue from `2LZT-ASP66.pqr` and `2LZT-ASH66.pqr`, respectively. Save the results as separate PQR files containing only the residue of interest: `ASP66.pqr` and `ASH66.pqr`, respectively.

Finally, we'll also need to perform electrostatics calculations on HEWL with an uncharged residue 66. Use your favorite text editor to zero out the charges in `2LZT-ASP66.pqr` and `2LZT-ASH66.pqr` to create `2LZT-noASP66.pqr` and `2LZT-noASH66.pqr`. This can be done by setting the second-to-last column in the PQR file to zero; e.g.

```
    ATOM    1008  N   ASP     66      -1.147
 0.169  33.201 -0.4000 1.5000
```

```
   ATOM   1009  CA  ASP    66      -0.225  -
0.753  32.484 -0.0000 2.0000
   ATOM   1010  C   ASP    66       0.047  -
1.937  33.399  0.5500 1.7000
   ATOM   1011  O   ASP    66       0.881  -
2.784  33.058 -0.5500 1.4000
   ATOM   1012  CB  ASP    66       1.017  -
0.081  31.957  0.0000 2.0000
   ATOM   1013  CG  ASP    66       1.991
0.445  32.943  0.1000 1.7000
   ATOM   1014  OD1 ASP    66       1.939
0.308  34.173 -0.5500 1.4000
   ATOM   1015  OD2 ASP    66       2.951
1.077  32.481 -0.5500 1.4000
   ATOM   1016  H   ASP    66      -0.854
0.593  34.052  0.4000 1.0000
   ATOM   1017  HA  ASP    66      -0.721  -
1.091  31.694  0.0000 0.0000
   ATOM   1018  HB2 ASP    66       1.494  -
0.751  31.351  0.0000 0.0000
   ATOM   1019  HB3 ASP    66       0.718
0.685  31.351  0.0000 0.0000
```

in the lysozyme PQR file would become

```
   ATOM   1008  N   ASP    66      -1.147
0.169  33.201  0.0000 1.5000
   ATOM   1009  CA  ASP    66      -0.225  -
0.753  32.484  0.0000 2.0000
   ATOM   1010  C   ASP    66       0.047  -
1.937  33.399  0.0000 1.7000
   ATOM   1011  O   ASP    66       0.881  -
2.784  33.058  0.0000 1.4000
   ATOM   1012  CB  ASP    66       1.017  -
0.081  31.957  0.0000 2.0000
   ATOM   1013  CG  ASP    66       1.991
0.445  32.943  0.0000 1.7000
   ATOM   1014  OD1 ASP    66       1.939
0.308  34.173  0.0000 1.4000
   ATOM   1015  OD2 ASP    66       2.951
1.077  32.481  0.0000 1.4000
   ATOM   1016  H   ASP    66      -0.854
0.593  34.052  0.0000 1.0000
   ATOM   1017  HA  ASP    66      -0.721  -
1.091  31.694  0.0000 0.0000
   ATOM   1018  HB2 ASP    66       1.494  -
0.751  31.351  0.0000 0.0000
   ATOM   1019  HB3 ASP    66       0.718
0.685  31.351  0.0000 0.0000
```

Setting up the total electrostatic energy calculations

We will be using focusing calculations to calculate the electrostatic potential and free energies for the systems of interest.

Warning: In what follows, we are evaluating total electrostatic free

energies -- e.g., energies which contain charge self-interaction terms. We will cancel these self-interaction terms in subsequent steps when we calculate solvation or transfer free energies. Therefore, it is very important that you use the same grid parameters (grid centers, dimensions, spacings, etc.) for every calculation.

Here is a template input that we will use for each of the solvation energy calculations:

```
   read
       mol pqr compound.pqr # This is the
compound for which we will calculate solvation
energies
       mol pqr ref.pqr      # This is a
compound used as a reference for grid centering
   end
   elec name inhom
       mg-auto              # Focusing
calculations
       dime 129 129 129     # This is a good
grid spacing for this system
       cglen 52.0 66.0 79.0 # These are
reasonable coarse grid settings for this system
(PDB2PQR-recommended)
       fglen 51.0 59.0 67.0 # These are
reasonable fine grid settings for this system
(PDB2PQR-recommended)
       cgcent mol 2         # Center the grid
on the reference molecule
       fgcent mol 2         # Center the grid
on the reference molecule
       mol 1
       lpbe
       bcfl sdh
       pdie 20.00
       sdie 78.54
       srfm smol
       sdens 40.0
       chgm spl2
       srad 1.40
       swin 0.30
       temp 298.15
       calcenergy total
       calcforce no
   end
   # Print the final energy
   print energy inhom end
   quit
```

There are a number of aspects to this input file which are worth noting:

- In general, compound.pqr will change for each calculation but `ref.pqr` will not. Choose one molecule to be `ref.pqr` (`2LZT-ASP66.pqr` is a good choice) and use it in every

calculation.

- We are using a solute dielectric constant (ε_p) of 20 (see pdie). This is a common choice for pKa since <begin vigorous waving of hands> it is thought to implicitly represent internal relaxation and rearrangement of the solute </end vigorous waving of hands>.
- We are using a molecular surface (srfm smol) with a reasonably high density of surface discretization points (sdens 40.0]). pKa and other electrostatics results can be very sensitive to surface choice.

You now have enough information to calculate total electrostatic energies for all of the relevant molecules so far: `2LZT-ASP66.pqr`, `2LZT-ASH66.pqr`, `2LZT-noASP66.pqr`, `2LZT-noASH66.pqr`, `ASP66.pqr`, and `ASH66.pqr`. You should be able to construct APBS input files for each of these systems by modifying the template above. Once these input files are constructed, you can run the PB calculation by

```
$ apbs foo.in | tee foo.out
```

where `foo.in` is the input file of interest and the output is saved in `foo.out`.

Setting up the transfer free energy calculations

Recall the transfer free energies can be evaluated by direct subtraction of total electrostatic energies for the different dielectric environments and components. This is usually the most stable route, assuming identical grids and solute conformations are used for all calculations. You should always check for a lack of convergence in the calculations and can be resolved by decreasing the grid spacing (e.g., increasing the number of grid points).

Putting it all together

At this point, you should have everything you need to calculate the intrinsic pKa of interest. However, if you get stuck, I've attached some example files below that might be helpful:

- 2LZT-ASP66.pqr and 2LZT-ASP66.in
  PQR and input files for ASP 66 in the protein

- ASP66.pqr and ASP66.in
  PQR and input files for isolated ASP 66 in solution
- 2LZT-noASP66.pqr and 2LZT-noASP66.in
  PQR and input files for the protein with an uncharged ASP 66
- 2LZT-ASH66.pqr and 2LZT-ASH66.in
  PQR and input files for ASH 66 in the protein
- ASH66.pqr and ASH66.in
  PQR and input files for isolated ASH 66 in solution
- 2LZT-noASH66.pqr and 2LZT-noASH66.in
  PQR and input files for the protein with an uncharged ASH 66
- run-apbs.sh
  A Bash shell script for running the various APBS calculations
- process.sh
  A Bash shell script for transforming the output from run-apbs.sh into a pKa shift

What's next?

Based on this brief introduction, you should be in a good position to go back and try to evaluate intrinsic pKa's for HIS 15 and GLU 35. How well do your results for those residues agree with experiment? What's different with those residues?

So far, we've only examined intrinsic pKas in this example and have ignored coupling between titratable groups. Jens Nielsen has developed a very nice software package called pKaTool which allows you to explore coupling between titratable sites and its impact on titration events in a protein system. He has provided a tutorial (PDF) which you can use to explore coupled titration states and use to familiarize yourself with pKaTool.