

Determining the pKa of Simple Molecules Using Gaussian 2009

Edition 1, 16 September 2019

Sydney A. McKee & Taras V. Pogorelov

Department of Chemistry, School of Chemical Sciences,

University of Illinois at Urbana-Champaign

This is a tutorial designed to introduce users to calculating pKa values of simple molecules utilizing *Gaussian 09*. Users are expected to have a foundation in Unix/Linux (see Linux primer tutorial).

Expected tutorial completion time: 2 hrs

Outline: we will use the SCS computer cluster, triton, to perform *Gaussian* calculations and *GaussView* to construct phenol and calculate its pKa in water. Please see the appendix for further information regarding the basis of these calculations, which are based on Chem. Phys. Lett. 367, 145-149, 2003.

1. **Software.** Triton: *Gaussian*, *GaussView*. Personal computer: xserver (Xming for a Windows) or Terminal (Mac).
2. **Starting up.** Open Windows PowerShell (or Terminal) window. Connect to Triton with the following command using your SCS login. If connecting outside of campus, connect to the VPN first.

```
>ssh -YC myscs-login@triton.scs.uiuc.edu
```

Make a directory for this tutorial:

```
>mkdir calculate-pka-tutorial
```

Change your directory to the new directory you just made:

```
>cd calculate-pka-tutorial
```

Start up a session of *Gaussian* with the following command:

```
>source /share/apps/gaussian/g09_login.csh
```

Or (subject to the Linux shell used by your account):

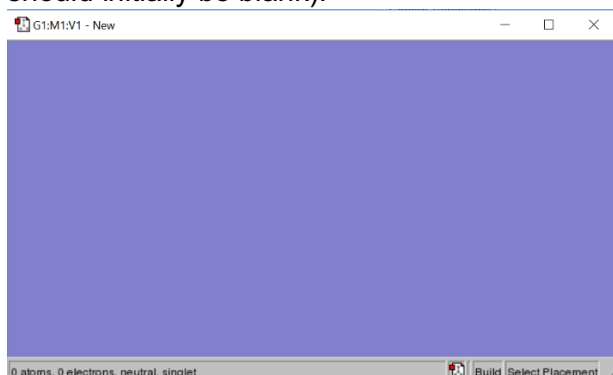
```
>source /share/apps/gaussian/g09_login.sh
```

Open *GaussView* with the following command:

```
>gv
```

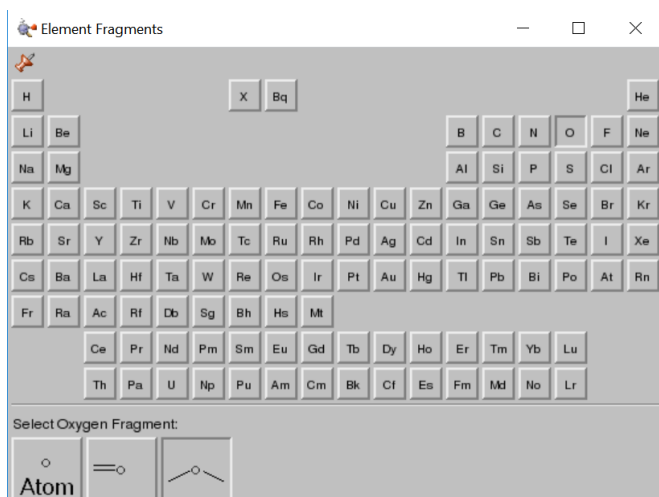
Note: you will likely get an error that says, "failed to locate Gaussian scratch directory." Ignore this error by simply hitting the "X". *GaussView* should load.

3. **Constructing a molecule.** Construct your molecule of interest in the blue window (which should initially be blank):

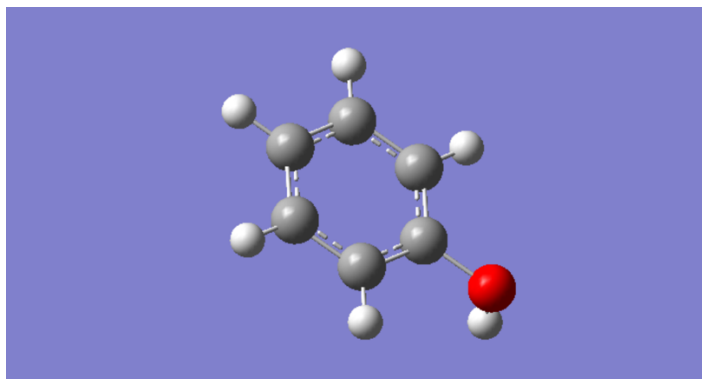


Right click anywhere on the blue window and select “Builder.” Click “Ring fragment.” Right click, select “Builder” again, and then click “Fragment” and select benzene.

Now, right click again and select “Builder”. Select “Element Fragment” which should pull up a periodic table. Select the oxygen fragment on the far right. Click on any hydrogen atom of the ring to substitute one of the hydrogens for a hydroxyl group.

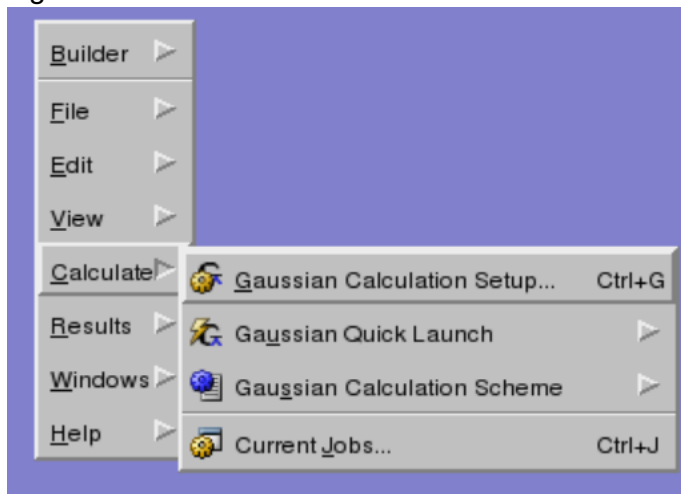


With the structure now built:



Select “Edit” and “Clean” to clean up the structure—this will perform a short geometry optimization.

4. **Optimization and Frequency Calculations in the gas phase.** Now set up a calculation with the built phenol. Right click and select “Calculate” then “Gaussian Calculation Setup...”



Under “Job Type” select “Opt+Freq” to optimize the geometry of the molecule and then calculate energy values of interest. It is important to always optimize first prior to running a frequency calculation.

Under “Method”, select “Ground State” “DFT...” “Default Spin” “B3LYP.” This is the level of theory for the calculation. For basis set, select “6-31G++d,p”. The basis set chosen for this calculation is pretty robust. **Note:** The lowest level of theory and basis sets necessary should be utilized when running calculations to reduce the expense of them. In this case, each calculation will take approximately 10 minutes to run with the level of theory and basis sets chosen.

This will be phenol without solvation (gaseous phase), so now hit “Submit.” Save the Gaussian input file phenol.com. Submit the file to *Gaussian* through *GaussView*—you will immediately get a window describing an error with the input file. Ignore this—we will be submitting though the cluster queueing software.

5. **Submission of the Gaussian input file.** Go back to your terminal window with Triton signed in. In order to submit a job, you must kill GaussView. Execute this with “ctrl+C”. Submit the calculation to Gaussian with the following command:

```
>submit-g09 -n # -q ib1 phenol.com
```

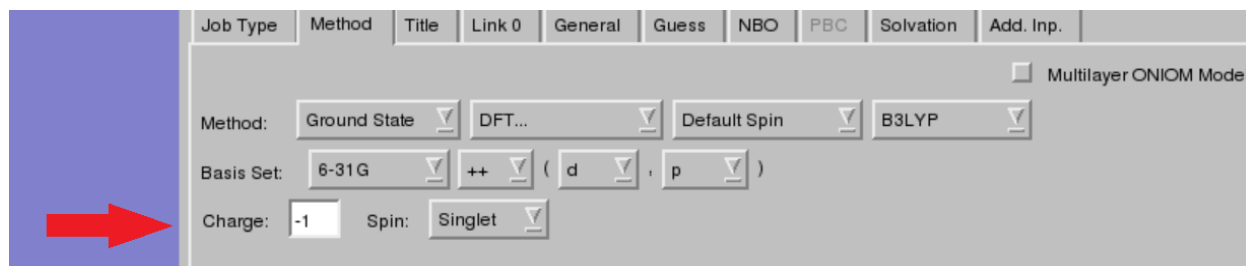
Replace the “#” in the command with the number of cores desired to be utilized for the calculation. Do not exceed 16.

After submission, check the progress of the calculation with

```
>qstat -u myscs-login
```

Once your SCS username no longer appears, the calculation has completed.

6. **Calculation of phenoxide in the gas phase.** Now, repeat the same process for phenoxide as was done for phenol. The only difference is that the hydrogen from oxygen needs to be deleted, and the charge of the molecule needs to be changed to -1. Delete a hydrogen atom by right clicking and selecting “Builder” then “Select Atoms by Clicking”. Click on the hydrogen and then press delete on the keyboard or the X button on the main window. Change the charge in the Method tab for the *Gaussian* calculation set up:



7. **Calculating for solvated molecules.** Now the calculations have been completed for phenol and phenoxide in their gaseous phases. Now we need to obtain the same calculations but for their solvated counterparts. Construct the molecules and set up the calculations just as before. Select the “Solvation” tab, then select the IPCM solvation model. Choose water as the solvent.

8. **Submission of solvated molecules.** Submit the solvated molecules, just as it was done for the gaseous counterparts.

9. **Submission for water.** Based on using thermodynamic cycle 1 (below), the energies of water and hydronium also need to be calculated. Construct and submit the jobs in the same way as before, ensuring that the charge denoted for hydronium is now +1.

10. **Retrieving relevant values.** After all submitted jobs completed, extraction of necessary information is done with the Linux “grep” command. Our values of interest are the sum of thermal and electronic energies (the Gibbs free energy). Input the following command:

```
>grep thermal energies *log
```

This command will search every log file (using the wildcard “*”) in the current directory for the keywords “thermal energies” in the Gaussian calculation output.

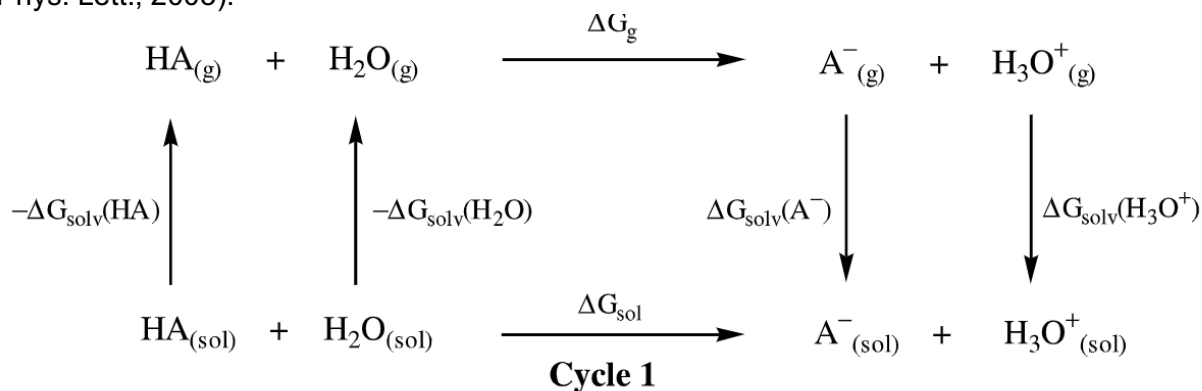
```
hydronium99.log: Sum of electronic and thermal Energies= -76.670644
hydronium99.log: Sum of electronic and thermal Enthalpies= -76.669700
hydronium99.log: Sum of electronic and thermal Free Energies= -76.692718
phenolsol99.log: Sum of electronic and thermal Energies= -307.391252
phenolsol99.log: Sum of electronic and thermal Enthalpies= -307.390308
phenolsol99.log: Sum of electronic and thermal Free Energies= -307.425836
phenoxide99.log: Sum of electronic and thermal Energies= -306.833763
phenoxide99.log: Sum of electronic and thermal Enthalpies= -306.832819
phenoxide99.log: Sum of electronic and thermal Free Energies= -306.867998
phenoxidesol99.log: Sum of electronic and thermal Energies= -306.924685
phenoxidesol99.log: Sum of electronic and thermal Enthalpies= -306.923741
phenoxidesol99.log: Sum of electronic and thermal Free Energies= -306.958762
water99.log: Sum of electronic and thermal Energies= -76.410010
water99.log: Sum of electronic and thermal Enthalpies= -76.409066
water99.log: Sum of electronic and thermal Free Energies= -76.430495
```

The values of interest are the sum of electronic and thermal energies.

11. **Calculation of pKa.** Open Microsoft Excel and format the sheet in the following fashion:

Sum of..	Phenol	Phenoxide	Phenol solvent	Phenoxide solvent	Water	Hydronium
Electronic + Thermal E	-307.384162	-306.833763	-307.391252	-306.929919	-76.41001	-76.670644
	Hartree	kcal/mole				
Delta Gg	0.289765	181.8301454				
Delta G S H3O+ kcal/mol	-110.3					
Delta Gs H2O kcal/mol	-6.32					
	Hartree	kcal/mol		Delta G soln		
Delta Gs Phenol	-0.00709	-4.44903881		21.96042879		
Delta Gs Phenoxide	-0.096156	-60.3387554				
kcal/mole conversion	627.509					
				pKa	9.81580638	
log[H2O]	1.744214725					
Correction Value	4.54					
Level of Theory	Basis Set	Solvation Model				
DFT/B3LYP	6-31G++d,p	IPCM				

The sum of electron and thermal energies are the values extracted utilizing the grep command. ΔG_s for H_3O^+ and H_2O are experimentally determined values. The ΔG_s for phenol and phenoxide are found by subtracting the Gibbs free energy for its gaseous phase from its Gibbs free energy for its solvation. ΔG_g is calculated following the equation from the thermodynamic cycle 1 (Chem. Phys. Lett., 2003):



Add ΔG_s for H_3O^+ and G_s for phenoxide and subtract ΔG_s for H_2O + G_s for phenol. Convert to kcal/mole by multiplying by 627.509. This is now the value for ΔG_g . Now, all the values to find ΔG_{aq} have been acquired:

$$\begin{aligned}
 \Delta G_{sol} = & \Delta G_g + \Delta G_{solv}(A^-) + \Delta G_{solv}(H_3O^+) \\
 & - \Delta G_{solv}(HA) - \Delta G_{solv}(H_2O),
 \end{aligned}$$

Now ΔG_{sol} can be plugged into the following pKa equation to get the uncorrected value:

$$pK_a = \frac{\Delta G_{sol}}{1.364} - \log[H_2O].$$

The concentration of bulk water is 55.49 M.

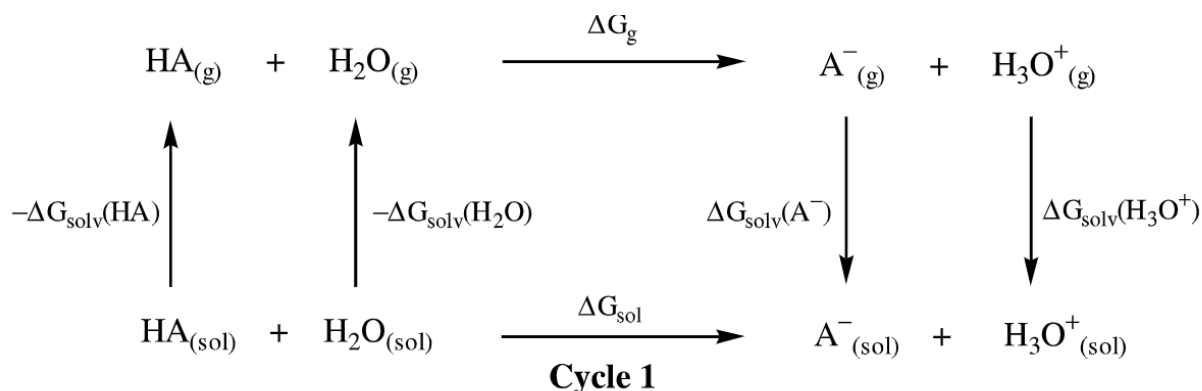
To find the corrected pKa, simply subtract 4.54 (see Appendix) from the initial estimate to get the final pKa value:

$$pK_a(\text{corrected}) = pK_a(\text{corrected}) - 4.54$$

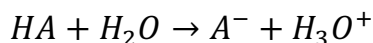
The final pKa for phenol was found to be 9.81. The experimental value of phenol's pKa in water is reported to be 9.95. Not bad!

Appendix

Calculating pKas in *Gaussian* is not straightforward in the sense that it will predict the pKa with a one simple command. Rather, calculations for the Gibbs free energy for the acid, its conjugate base, and water/hydronium need to be obtained in both gaseous phases and solution (solvation). The pKa can then be determined by employing the following thermodynamic cycle:



The equilibrium between the acid and water:



This presents the equation of ΔG_{sol} then as the following:

$$\Delta G_{\text{sol}} = \Delta G_g + \Delta G_{\text{solv}}(\text{A}^{-}) + \Delta G_{\text{solv}}(\text{H}_3\text{O}^{+}) - \Delta G_{\text{solv}}(\text{HA}) - \Delta G_{\text{solv}}(\text{H}_2\text{O})$$

The solved ΔG_{sol} then is plugged into the following pKa equation, where the bulk concentration of water is 55.49 M:

$$pK_a = \frac{\Delta G_{\text{sol}}}{1.364} - \log[\text{H}_2\text{O}]$$

A more detailed derivation for the pKa equation can be found in *Chemical Physics Letters* 367, 145–149, 2003. A source of error in the thermodynamic cycle arises with the solvation free energy of hydronium. To mitigate this, the following equation is used:

$$pK_a(\text{corrected}) = pK_a(\text{corrected}) - 4.54$$