Investigation on the pKa of Acids by Computational Simulation using the Density Functional Theory

Youze Fang
Shanghai United International School, Qingpu Campus, 201704 Shanghai, China

Abstract

We present a Density Functional Theory (DFT) study to investigate the pKa values of various acids. A brief background on quantum mechanics is provided to introduce the fundamental principles of Kohn–Sham DFT. We then apply DFT to calculate the proton dissociation energies of chlorine-substituted ethanoic acid and aspartic acid. While the computed trends for chlorine-substituted ethanoic acid agree well with experimental observations, noticeable discrepancies arise in the case of aspartic acid. We further analyze and rationalize the underlying factors contributing to this divergence between theoretical predictions and experimental results.

Keywords

Density Functional Theory (DFT); Kohn–Sham formalism; pKa prediction; Aspartic acid; Chlorine-substituted ethanoic acid; Proton dissociation energy; Solvation modelling.

1. INTRODUCTION

Computational chemistry has a wide range of applications, including virtual screening for drug discovery, simulating spectra and molecular behavior for substance identification, and modeling biological processes involving macromolecules. It can also be employed to predict reaction kinetics, improve the efficiency of chemical manufacturing, and simulate the electronic properties of semiconductors. In this study, we utilize computational chemistry to simulate molecular systems and determine their relative energies in aqueous solution. By comparing these energies, we gain insight into the stabilities of different acid states, which enables us to establish a trend in relative acidities.

2. THEORY

To describe any microscopic particle, such as an electron, mechanically, the Schrödinger equation is required, which exists in such form:

$$\hat{H}\psi = E\psi \tag{I}$$

For means of simplicity, we will begin with a basic model, assuming that the particle is in a one-dimensional box with length a. In equation I, $\hat{H}$ is the Hamiltonian operator, which is equal to $-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$, with $\hbar$ being the reduced Planck constant, and m being the mass of the particle. E is the total energy of the particle. The solution to the equation, ψ, is known as the eigenfunction of the Schrödinger equation, and E being the eigenvalue. By setting the equation:

$$\hat{H} = \hat{T} + \hat{V}$$

Where T is the kinetic energy of the particle and V is the potential energy of the particle, equation I may be transformed as such:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+\hat{V}\psi=E\psi \tag{II}$$

In this particular case, V is defined as a piecewise function:

$$V=\begin{cases}0 & 0\leq x\leq a\\ +\infty & otherwise\end{cases}$$

Since the total energy of the particle cannot be +∞, we will only consider cases for which the particle takes on values of x that are within [0, a]. Thus the $\hat{V}\psi$ part of equation II can be considered to be zero. By solving the remaining equation, which is a second order ordinary differential equation, we can obtain the following solution for ψ:

$$\psi=Asin\left(\frac{\sqrt{2mE}}{\hbar}x\right)+Bcos\left(\frac{\sqrt{2mE}}{\hbar}x\right)$$

We then use the property of wavefunctions, such that any wavefunction must have a numerical value of 0, corresponding to wave nodes in classical physics. Hence it is straightforward to obtain the following:

$$\psi(0)=\psi(a)=0\ \rightarrow B=0,\frac{a\sqrt{2mE}}{\hbar}=n,n\in z^+$$

By substituting the solution back into Equation I, we can obtain:

$$E=\frac{\hbar^2n^2\pi^2}{2ma^2}$$

With the successful derivation of this equation, we have confirmed that energy of a particle is quantised. To interpret the ψ(x), we use Born's interpretation, where p(x) is the probability distribution of the particle:

$$\int_a^b p(x)dx=\int_a^b|\psi(x)|^2dx$$

$$\int_{-\infty}^{+\infty}p(x)dx=\int_{-\infty}^{+\infty}|\psi(x)|^2dx=1 \tag{III}$$

Using equation III, we can normalise the wavefunction to obtain such results:

$$A=\sqrt{\frac{2}{a}}$$

$$\psi(x)=\sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)$$

In the DFT, the aim is to find the electron distribution function such that the energy of the system is the lowest, utilising the Schrödinger equation in much more complex ways. Compounds in reality have more than one electron, and the interactions between the various particles may become overly sophisticated to compute. Thus, we find using the Kohn-Sham theory crucial, it enables us to ignore the interactions between electrons for simplicity and ensure the electron density distribution found is the same as reality. Specifically, this approximation is known as the mean field approximation, electrons other than the designated one is assumed to create a field, instead of acting as a particle. The computed solution to the Schrödinger equation, and the probability distribution derived from the wavefunction in an actual molecule is in the form of:

$$\psi(r) = Ae^{-r}$$

$$p(r) = 4\pi r^2|\psi(r)|^2 = 4A\pi r^2 e^{-r}$$

To efficiently evaluate the integral of p(r), we approximate it with the Gaussian integral:

$$\int r^n e^{-r} dr \rightarrow \int e^{(r+b)^2} dr$$

In the Kohn-Sham DFT, the energy functional is defined as:

$$E[n] = T^e[n] + \frac{1}{2}\int d^3\vec{r}\, d^3\vec{r}\, \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|} + V^{n-e}[n] + E^{xc}[n]$$

n is the energy function in terms of distance from the nucleus. To fully utilise the Kohn-Sham DFT [1, 2] in practical calculations, a self-consistent field [3] is required, which is a cycle that can approximate the true distribution of energy through iterations shown in Figure 1 below:

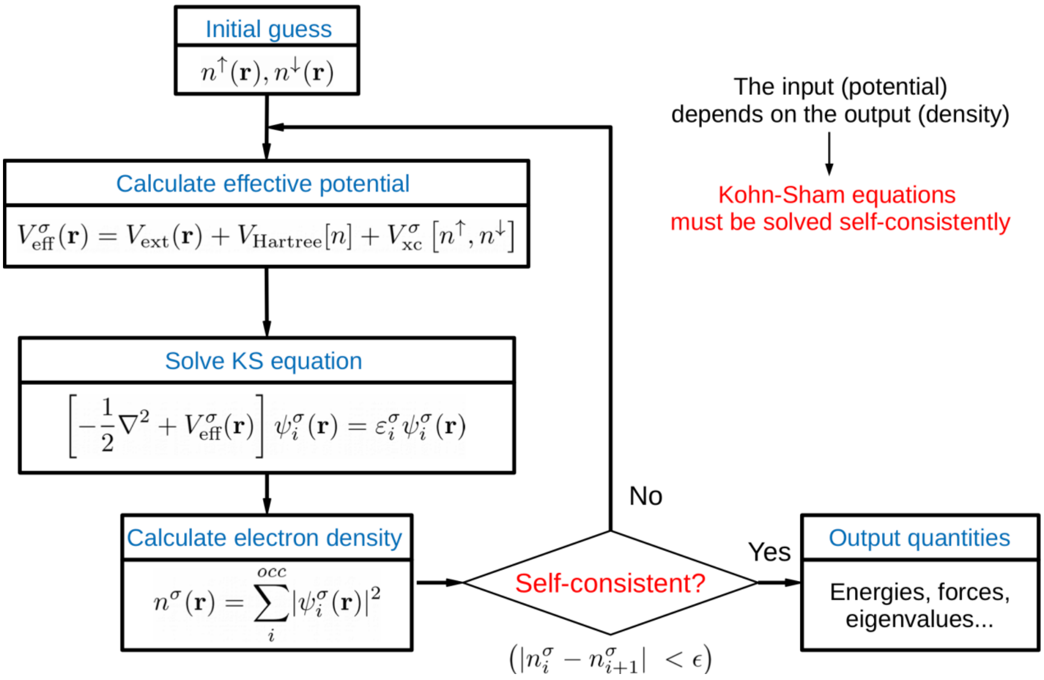


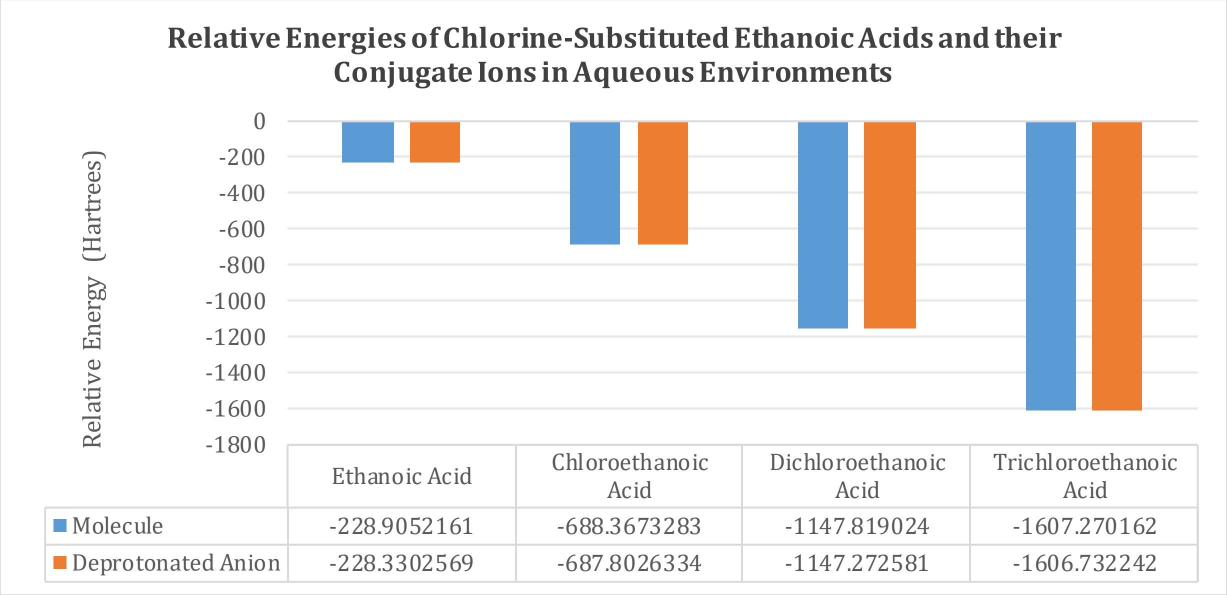
Figure 1. DFT Iterations

The resulting quantities computed can serve as an approximation to those of the desired molecule in reality.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202512_11(12).0002

3. RESULTS

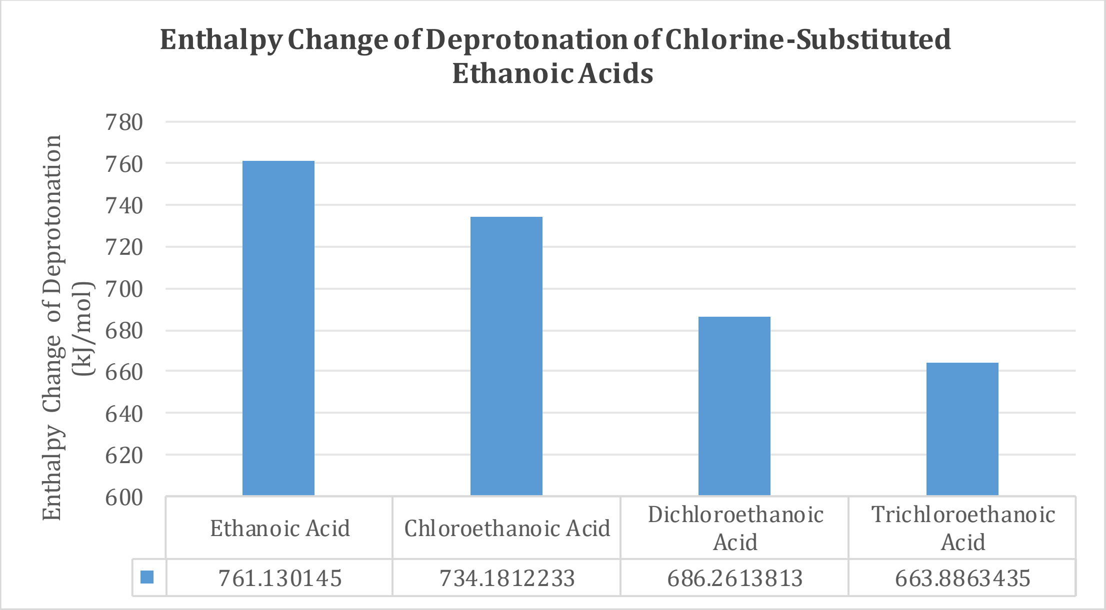
We employed the Psi4 quantum chemistry package to compute the relative energies of various forms of chlorine-substituted ethanoic acid, considering both its neutral and deprotonated states. The 6-31G*[4, 5] basis set was used, which involves solving four Gaussian functions for each hydrogen atom and eleven Gaussian functions for elements ranging from lithium through neon. In the simulation, we first optimized the geometry of each molecule or ion in the absence of solvent to reduce computational complexity. Subsequently, we incorporated water as the solvent to better approximate experimental conditions and obtained the corresponding energetic data that are shown in Figure 2.



	Ethanoic Acid	Chloroethanoic Acid	Dichloroethanoic Acid	Trichloroethanoic Acid
🟦 Molecule	-228.9052161	-688.3673283	-1147.819024	-1607.270162
🟧 Deprotonated Anion	-228.3302569	-687.8026334	-1147.272581	-1606.732242

Figure 2. Relative energies of chlorine substituted acids.

Using Hess' law, we computed the enthalpy change of deprotonation for the acid molecules, incorporating the calculated relative energies of water and the hydronium ion that are shown in Figure 3.



	Ethanoic Acid	Chloroethanoic Acid	Dichloroethanoic Acid	Trichloroethanoic Acid
🟦	761.130145	734.1812233	686.2613813	663.8863435

Figure 3. Enthalpy changes of deprotonation of Ethanoic Acids

Our results indicate that the enthalpy change of deprotonation becomes progressively less endothermic as the number of chlorine substituents increases. From this, we conclude that each additional chlorine atom enhances the ease of deprotonation, thereby increasing the acidity of the molecule. This trend shows good consistency with experimental data.

We then carried on to compute these data for aspartic acid that are shown in Figure 4.

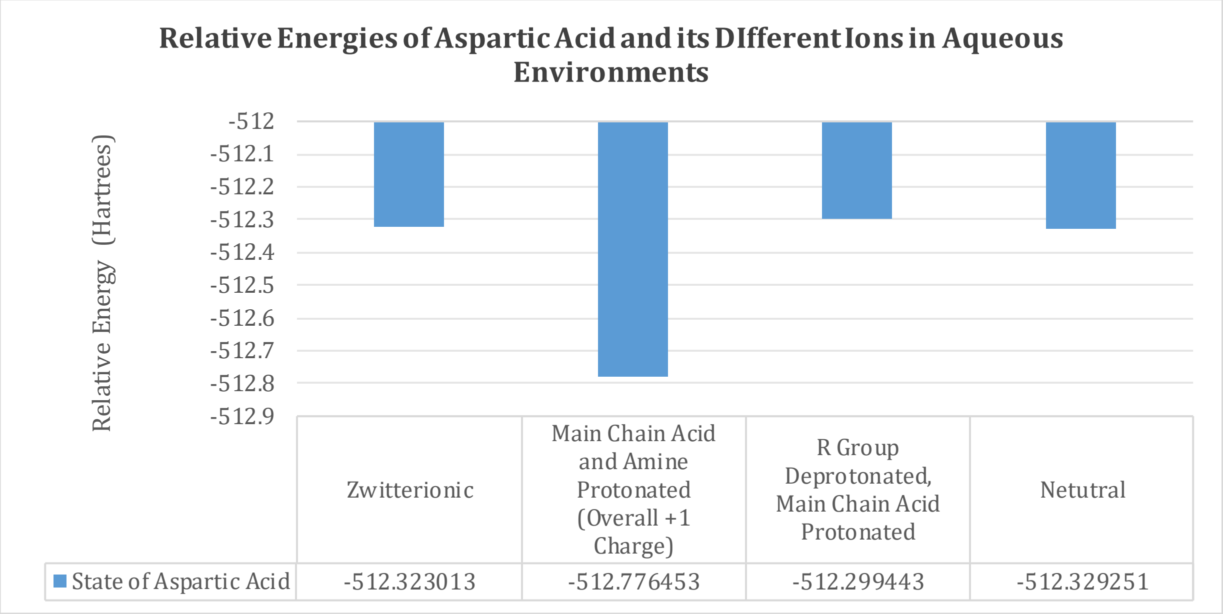


Figure 4. Enthalpy changes of deprotonation of Aspartic Acids

We then extended our analysis to aspartic acid. Using the computed data, we evaluated the enthalpy changes of deprotonation for each acidic hydrogen in the molecule. According to our model, the hydrogen on the NH_3^+ group exhibited the highest acidity (lowest enthalpy change of deprotonation), followed by the main-chain carboxyl hydrogen, and finally, the R-group carboxyl hydrogen as the least acidic.

However, this theoretical ranking contradicts experimental observations, where the main-chain carboxyl group is experimentally determined to be the most acidic, followed by the R-group carboxyl group, with the protonated amine group being the least acidic.

We propose several possible explanations for this discrepancy:

(1) Limitations of the chosen density functional — The functional used in our calculations may exhibit reduced accuracy when treating ionic species or solute–solvent interactions in aqueous environments.

(2) Lack of explicit solvation modeling — Our simulations treated the solvent as a continuous dielectric field rather than incorporating explicit water molecules. This simplification neglects specific hydrogen-bonding and ion–solvent interactions, potentially leading to inaccuracies.

(3)ΔH vs. ΔG distinction — The pKa is fundamentally related to the Gibbs free energy change (ΔG), whereas our calculations focus on enthalpy changes (ΔH). Consequently, our results provide only qualitative insights rather than fully quantitative predictions.

(4) Coupling between protonation states — In polyprotic systems like aspartic acid, the deprotonation energetics of one site can depend strongly on the protonation state of other sites. Our simplified model, which considered only a limited set of states, cannot fully capture the complexity of the real aqueous environment.

(5) These factors collectively suggest that a more sophisticated computational protocol — potentially involving higher-level functionals, explicit solvent modeling, and free-energy calculations — would be required to achieve closer agreement with experimental pKa values.

REFERENCES

[1] Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A), A1133–A1138. https://doi.org/10.1103/PhysRev.140.A1133

[2] Parr, R. G., & Yang, W. (1994). *Density-Functional Theory of Atoms and Molecules*. Oxford University Press.

[3] Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7), 5648–5652. https://doi.org/10.1063/1.464913

[4] Hehre, W. J., Ditchfield, R., & Pople, J. A. (1972). Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics*, 56(5), 2257–2261. https://doi.org/10.1063/1.1677527

[5] Ditchfield, R., Hehre, W. J., & Pople, J. A. (1971). Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *The Journal of Chemical Physics*, 54(2), 724–728. https://doi.org/10.1063/1.1674902