Alan Yu
Professor Wang
Computational Chemistry
5 May 2021

Final Project Written Report

I. Motivation of Project

For my final project, I decided to look at the racemization of thalidomide. This topic was introduced to me by Professor Taylor from Organic Laboratory 310. As our final exam essay question, he tasked us to research about what the thalidomide epidemic was and what it did to the population. After answering the essay question on the exam, I was interested to learn more about this drug.

II. Introduction

Thalidomide was first marketed in West Germany in 1957 as an incredibly safe sedative for the population. Shortly, it also became a solution for pregnant mothers with morning sickness. However, it was found that it had teratogenic effects as shown in the babies from the mothers who took this remarkable drug.

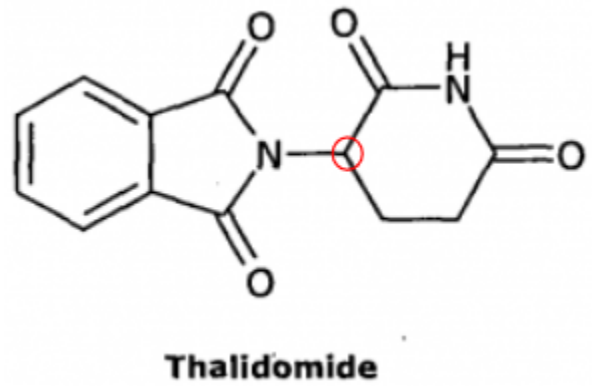

Figure 1. Thalidomide with stereocenter circled in red

Above is the structure of thalidomide where it can be seen that thalidomide can exist in not just one form because of the stereocenter.

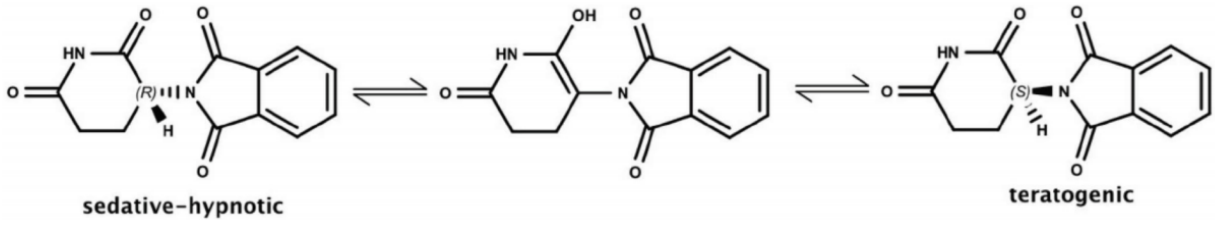

Figure 2. Thalidomide enantiomers and conversion scheme[1]

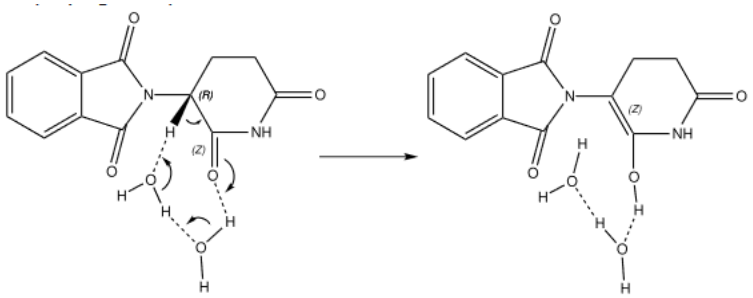

Figure 3. Tautomerization mechanism[2]

In Figure 2, it can be seen that there are two enantiomers where the R-enantiomer acts as the sedative drug while the S-enantiomer acts as the teratogenic drug. It was found that within the body, the sedative form, the R-enantiomer, interconverted to the S-enantiomer which had a better binding to cereblon, a protein on human chromosome 3, created these birth defects seen within the babies[4,5]. Figure 3 provides the mechanism of the formation of the enol from the R-enantiomer. This process begins where two water molecules approach the enantiomer and take the hydrogen from the stereocenter. Once the hydrogen is removed, it forces the electron pair towards the ketone and the electron pair from the ketone is forced to grab a hydrogen from the other water molecule and creates the enol product. Once the enol product forms, to convert to the other enantiomer, two water molecules again approach the molecule and take the hydrogen off of the enol where the ketone used to be. Taking the hydrogen off forces the electron pair down to where the double bond is and forces this electron pair into grabbing the hydrogen of a nearby water molecule thus converting into the other enantiomer. With all this in place, this brings up the question of how this safe drug caused this many issues, specifically how did the R-enantiomer end up becoming the S-enantiomer?

III. Computation Setup

So my idea was to consider the free energy of conversion between the R and S enantiomers based on what Figure 2 depicted. I would be calculating the free energies of each reaction of the conversion as well as the free energy of the total reaction. First, I created the R-enantiomer, S-enantiomer, and the enol within Gaussian. After creating the initial structures, I ran the input files by first finding the optimized structures then creating the thermochemistry calculations. I did these calculations at 298.15K at first with the Hartree Fock method and a basis set of 6-31G to see if any errors would appear. After confirming that there were no errors, I used the BLYP method with the same basis set to get the thermochemistry calculations. Afterwards, I decided to use BLYP and B3LYP, both with a 6-31G basis set, to calculate the thermochemistry at 310.15K, or body temperature. Below are the optimized structures of each molecule of study within this topic.

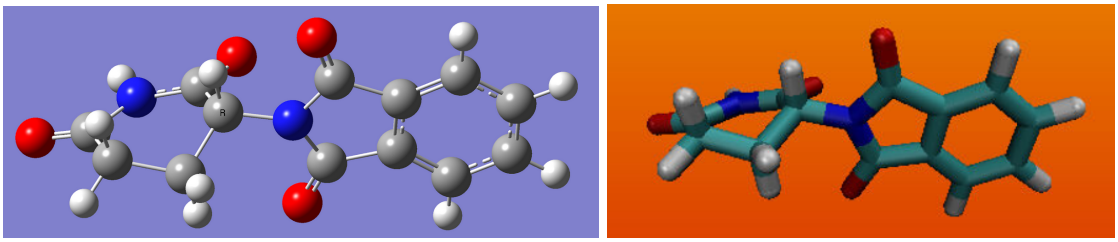

Figure 4. Optimized R-enantiomer in GaussView(left) and VMD(right).

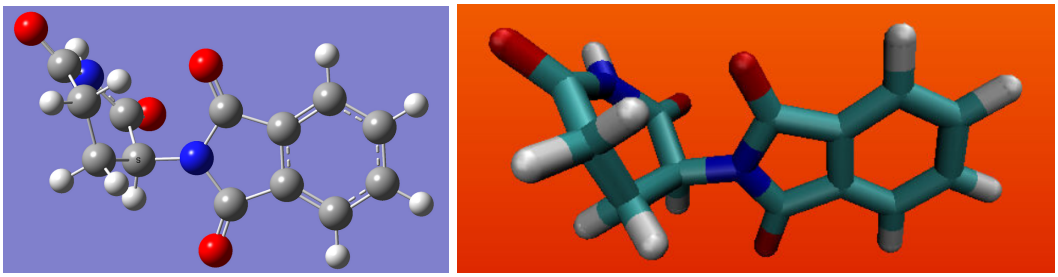

Figure 5. Optimized S-enantiomer in GaussView(left) and VMD(right).

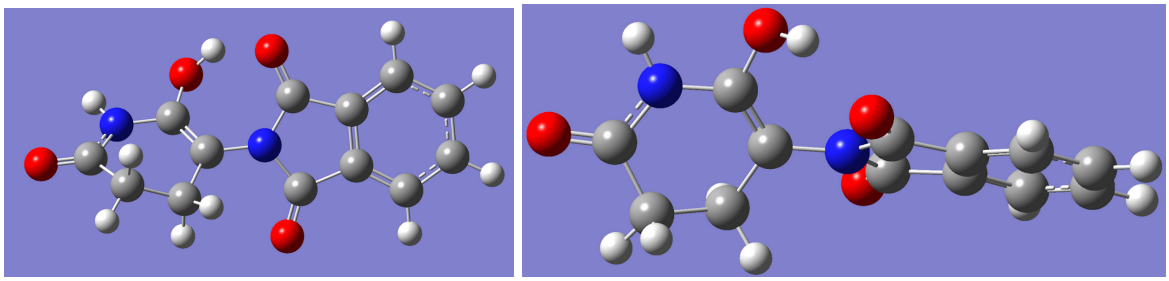

Figure 6. Optimized Enol with head on view(left) and angled view(right).

With these structures and the calculated thermochemistry, I paired it up with their respective equations to find the free energy of the single equation and with the combined equations. Below are the equations used to calculate the free energies.

$$C_{13}H_{10}N_2O_4\ (R) + 2H_2O \rightleftharpoons C_{13}H_{10}N_2O_4\ (I) + 2H_2O \quad (1)$$

$$C_{13}H_{10}N_2O_4\ (I) + 2H_2O \rightleftharpoons C_{13}H_{10}N_2O_4\ (S) + 2H_2O \quad (2)$$

IV. Results and Discussions

After acquiring the data from the .log files and using the equations of interest, the $\Delta G_{298.15K}$ and $\Delta G^{Total}_{298.15K}$ can be found and are found tabulated below.

Table I. Equation (1) free energy calculations at 298.15K

	R-Thalidomide	**Enol**	**$\Delta G_{298.15K}$**
Using HF	-568.4859053 kcal/mol	-568.4673859 kcal/mol	0.0185194 kcal/mol
Using BLYP	-571.7674436 kcal/mol	-571.7503385 kcal/mol	0.0171051 kcal/mol

Table II. Equation (2) free energy calculations at 298.15K

	Enol	**S-Thalidomide**	**$\Delta G_{298.15K}$**
Using HF	-568.4673859 kcal/mol	-568.4830879 kcal/mol	-0.015702 kcal/mol
Using BLYP	-571.7503385 kcal/mol	-571.7656552 kcal/mol	-0.0153167 kcal/mol

Table III. Total free energy and equilibrium constant at 298.15K

	$\Delta G^{Total}_{298.15K}$	**K_{eq}**
Using HF	0.0028174 kcal/mol	0.995
Using BLYP	0.0017884 kcal/mol	0.997102

From table I, it can be seen that it requires some sort of energy to activate the reaction, but as soon as the reaction begins, it can be seen that in table II, the enol readily wants to convert into the other enantiomer. Using the $\Delta G^{Total}_{298.15K}$ value, we can find the equilibrium constant in table III, it shows that the the constant is close to 1, but shifts towards the R-enantiomer meaning that there is almost an equal ratio of R to S enantiomers. After seeing these results, I wondered if the B3LYP method with the same 6-31G basis set would give me a lower value that would make the reaction require a little less energy aside from the already very low energy value. Now calculating the same as before but at the temperature of 310.15K with only BLYP and B3LYP, the $\Delta G_{310.15K}$ and $\Delta G^{Total}_{310.15K}$ can be found and are found tabulated below.

Table IV. Equation (1) free energy calculations at 310.15K

	R-Thalidomide	**Enol**	**$\Delta G_{310.15K}$**
Using BLYP	-571.7689602 kcal/mol	-571.7518245 kcal/mol	0.0171357 kcal/mol
Using B3LYP	-571.9182644 kcal/mol	-571.9011688 kcal/mol	0.0170956 kcal/mol

Table V. Equation (2) free energy calculations at 310.15K

	Enol	**S-Thalidomide**	**$\Delta G_{310.15K}$**
Using BLYP	-571.7518245 kcal/mol	-571.7671781 kcal/mol	-0.0153536 kcal/mol
Using B3LYP	-571.9011688 kcal/mol	-571.9161686 kcal/mol	-0.0149998 kcal/mol

Table VI. Total free energy and equilibrium constant at 310.15K

	$\Delta G^{Total}_{310.15K}$	K_{eq}
Using BLYP	0.0017821 kcal/mol	0.997113
Using B3LYP	0.0020958 kcal/mol	0.996605

Again, we can see that the first step requires some energy before it can readily convert into the other enantiomer. I'd assume that this energy can come from some sort of molecular collision in the bloodstream. Along with finding a smaller energy required, we can also see that the equilibrium constant hovers around the 0.996 and 0.997 values which means that the system prefers slightly more R-enantiomers than S-enantiomers. After calculating the values, I acquired the times of each calculation and created a table that has the minimum, maximum, and the average time per job on the Amarel server.

Table VII. Times for each type of calculation

Method	**Basis Set**	**Shortest Time**	**Longest Time**	**Average Time**
Hartree Fock @ 298.15K	6-31G	5 min 6 sec	5 min 18 sec	5 min 12 sec
BLYP @ 298.15K	6-31G	23 min 50 sec	25 min 59 sec	25 min 9 sec
BLYP @ 310.15K	6-31G	25 min 12 sec	28 min 10 sec	26 min 30 sec
B3LYP @ 310.15K	6-31G	26 min 28 sec	37 min 31 sec	30 min 55 sec

It can be seen that as the more accurate method is chosen to calculate the structures, the time increases. Because the basis sets are all the same, there is no correlation between time and basis set, but a more expensive basis set like 6-31G** would cost longer than what was acquired here.

V. Conclusions and Future Directions

So from the calculations in IV, it can be said that because of the small and positive ΔG for the first step, it is not spontaneous and requires a little bit of energy to begin the tautomerization of the R-enantiomer of thalidomide. Once the enantiomer has the energy to tautomerize, the enol product will readily convert to the S-enantiomer because of the negative ΔG. With the total ΔG value, it can be concluded that because of the very small energy value, this full reaction is not spontaneous, but can be initiated with a slight upset in the system such as a molecular collision. With the calculations I have done now, I would like to explore more expense methods and basis sets to calculate the thermochemistry of the three structures in question and see if the free energy reaches a value even lower than what I have with B3LYP. After that I would also like to explore simulations of each enantiomer of thalidomide binding to cereblon and find the binding energies.

After exploring the simulations of thalidomide binding, I would also like to consider derivatives of thalidomide such as fluorinated thalidomide that Shibata et al. also explored but in a computational sense rather than an experimental sense. Below are figures pertaining to what would be interesting to study in the future.

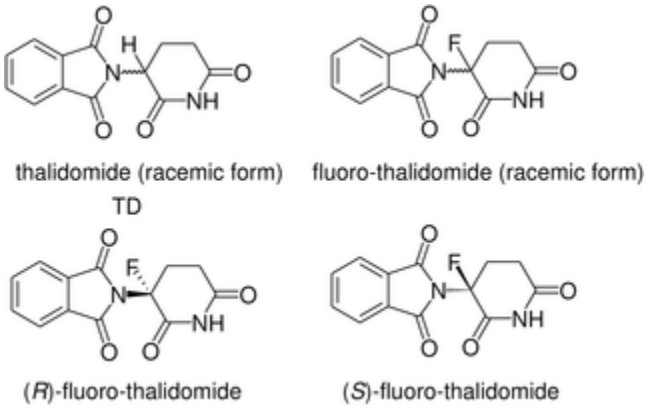

Figure 7. Structures of study from Shibata et al.[3]

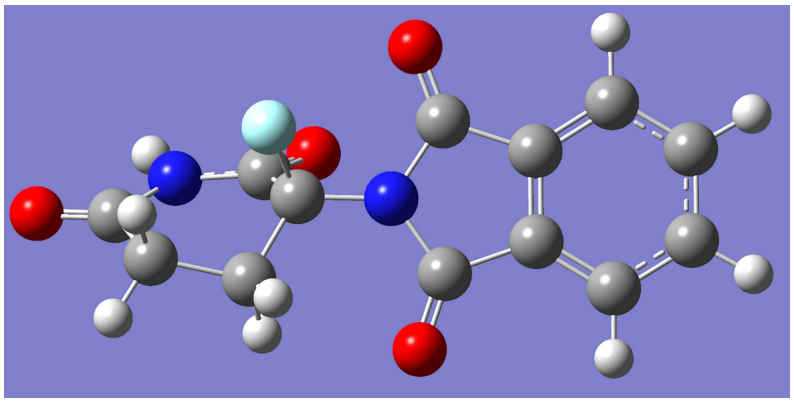

Figure 8. Optimized fluorinated thalidomide from Gaussian.

References

1. Braga, Rodolpho & Alves, Vinicius & Silva, Arthur & Nascimento, Marilia & Silva, Flavia & Lião, Luciano & Andrade, Carolina. (2014). Virtual Screening Strategies in Medicinal Chemistry: The State of the Art and Current Challenges. Current Topics in Medicinal Chemistry. 14. 1899-1912. 10.2174/1568026614666140929120749.
2. https://www.ch.imperial.ac.uk/rzepa/blog/?p=8246
3. https://doi.org/10.1371/journal.pone.0182152
4. Tokunaga, E., Yamamoto, T., Ito, E. et al. Understanding the Thalidomide Chirality in Biological Processes by the Self-disproportionation of Enantiomers. Sci Rep 8, 17131 (2018). https://doi.org/10.1038/s41598-018-35457-6
5. https://en.wikipedia.org/wiki/Cereblon