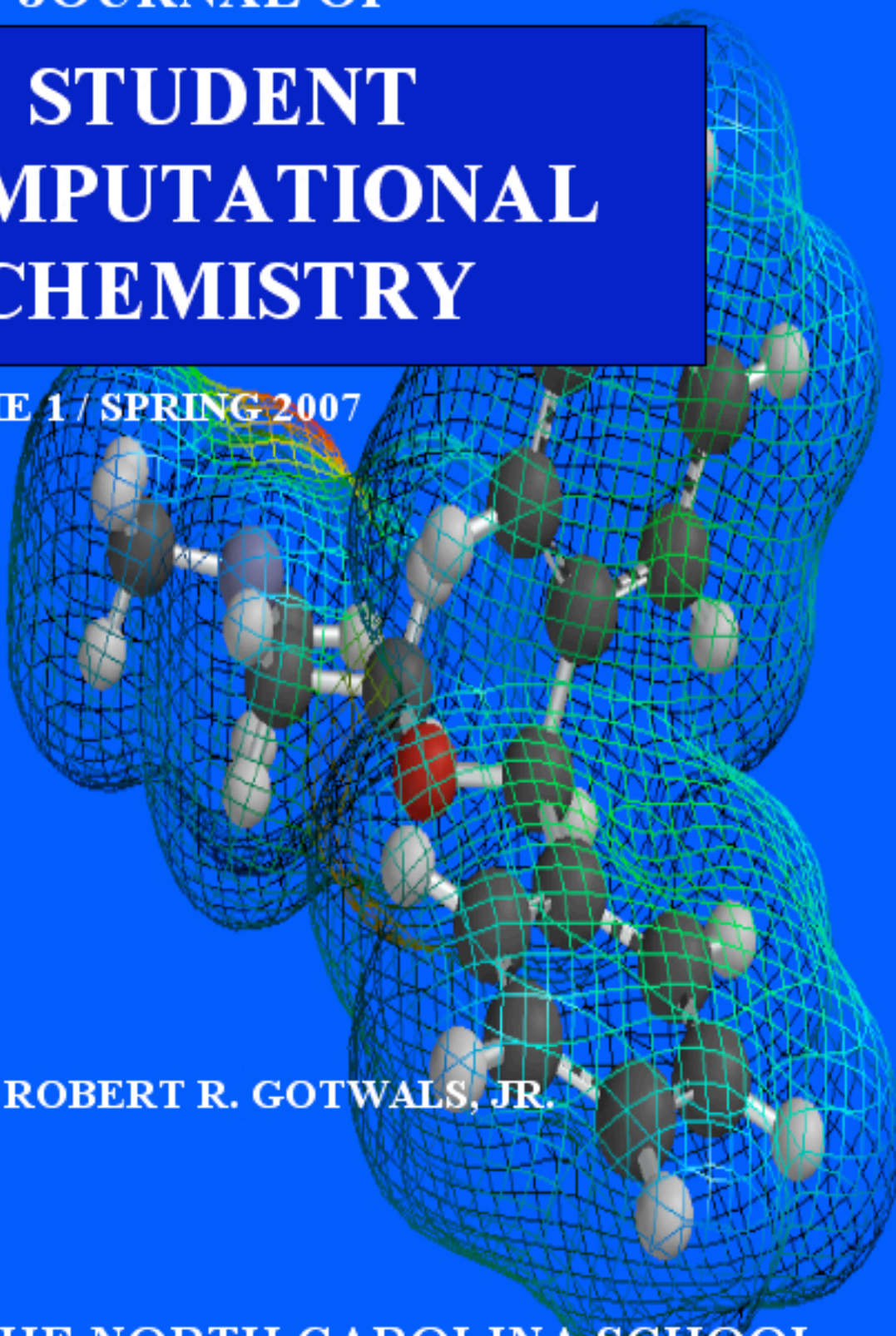


**JOURNAL OF
STUDENT
COMPUTATIONAL
CHEMISTRY**



VOLUME 1 / SPRING 2007

EDITOR: ROBERT R. GOTWALS, JR.

**THE NORTH CAROLINA SCHOOL
OF SCIENCE AND MATHEMATICS**

Modeling Excited States of Fluorescent Compounds with UV-Vis Spectra Calculations

Eric Li, Amy Kim, Lisa Zhang

North Carolina School of Science and Mathematics, Durham, NC

Received 6 June, 2007; Accepted 6 June, 2007

Published online on Comp Chem Moodle (moodle.ncssm.edu)

Abstract: Fluorescence is the property where molecules absorb electromagnetic radiation usually in the ultraviolet range and re-emit light in greater wavelengths, usually in the visible light range. We investigated the fluorescent molecules DMABN, ABN, MABN, MMD, caffeine, anthracene, salicylamide, and acetyl anthranilic acid through excited state computational chemistry calculations and observed the excited state quantum descriptors with a focus on peak absorbance wavelength, excitation energy, and oscillator strength. Calculations were run with DFT theory TD-B3LYP and a 3-21G basis set. The excited state that contributes most to the absorption peak of a molecule is shown by the largest oscillator strength, and the difference between the corresponding excitation energy and the LUMO was found. This value was compared to the HOMO-LUMO gap to determine the amount of extra energy required for electrons to move to the excited state.

Key words: excited states, DMABN, UV-Vis spectrum, oscillator strength, HOMO-LUMO gap

Introduction

The phenomenon of fluorescence is a commonplace one that has been seen in a wide variety of cases, from minerals to live organisms. Fluorescence occurs when a molecule absorbs light (composed of photons) usually in the ultraviolet range. The reason we are able to see a different color of emitted light when the molecules are exposed to ultraviolet light is because the molecule emits light at a longer wavelength that is commonly in the visible spectrum.

Since simple fluorescence is the basis for all other more complicated phenomenon including triboluminescence, which is the emission of light in response to mechanical stress, we hope to apply computational results and relationships found in the course of this research and extrapolate them to larger and more complex triboluminescent molecules.

We began our investigation with the 4-(N,N-dimethylamino)benzonitrile (DMABN) molecule, which is known to exhibit dual fluorescence at 350nm 475nm. In the gas phase, only the 350nm absorbance peak can be seen. This state is the local-excited (LE) state which is formed from the π to π^* orbital transition in the benzene ring of the molecule. The second state at 475nm is the charge-transfer state and it is due to

the transfer of electrons from the nitrogen to the CN group.⁵ In solution, especially polar solvents such as acetonitrile and methanol, both excited states are visible; a pilot calculation was done on DMABN in acetonitrile to see the two fluorescent peaks but this calculation was not included in our data for the purpose of consistency. Afterwards, we proceeded to performing calculations on related molecules MMD, 4-aminobenzonitrile (ABN) and N-methyl-4-aminobenzonitrile (MABN), the latter two of which are not dual-fluorescent because of their amino groups with one or no methyl substituents⁶ (Figure 1). These and the rest of the calculations were done in the gas phase because they do not exhibit the dual fluorescence property and are therefore not affected by solvents in this regard. The molecules MMD, caffeine, anthracene, salicylamide, and acetyl anthranilic acid were found through more background research on fluorescence. Acetyl anthranilic acid (Figure 2) is a known triboluminescent compound we have produced and worked with in the laboratory that also exhibits a green fluorescence under a UV lamp. This molecule represents the bridge between the two phenomena we are attempting to connect and understand in greater depth.

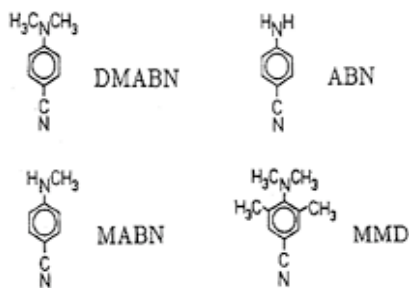


Figure 1. DMABN and related molecules⁶

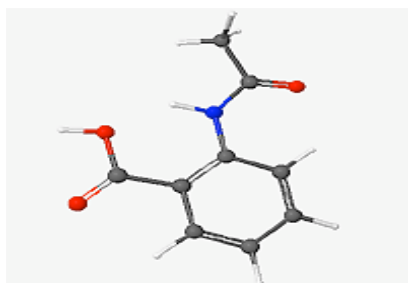


Figure 2. Acetylanthranilic acid

While researching excited states, we found the important factors to be the peak absorbance wavelength shown in the UV-Vis spectrum, the excitation energy of particular excited states, and the oscillator frequency (or oscillator strength). The oscillator strength is the intensity of the band; also, it is the most likely electronic transition of the promotion of an electron in excited states. These excitation energies will be compared to the highest occupied molecular orbital (HOMO), the lowest unoccupied molecular orbital (LUMO), and the HOMO-LUMO gap energy of the respective molecules. This gap energy is very important in chemical reactions because this determines the excitability of the molecule. Thus, small gaps are characteristic of easily excited molecules because only minimal energy needs to be overcome during electron transitions.

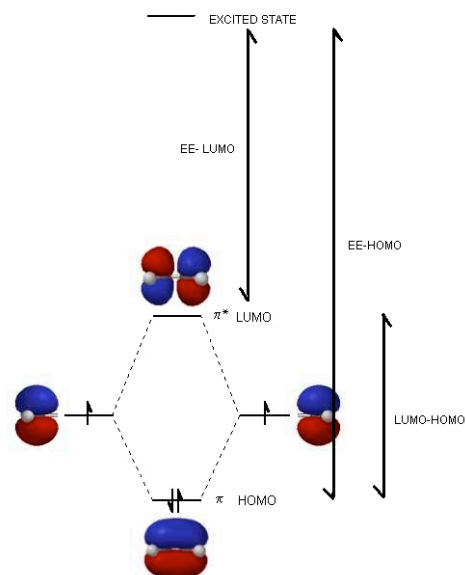


Figure 3. HOMO-LUMO gap and excited energy state

By analyzing these properties and finding various relationships between them, we looked for correlations and trends relating to excited states in fluorescence.

Computational Approach

For DMABN, MABN, ABN, MMD, and salicylamide, geometry optimizations were run using the MOPAC program with AM1 basis set calculations; anthracene and caffeine geometry optimizations were run using the MOPAC program with PM3 basis set calculations. The Gaussian 03 program with B3LYP theory and 6-31G(d) basis set was used for geometry optimization of acetyl anthranilic acid. At first, attempts were made to alter and fix the position of the methyl groups attached to the nitrogen in DMABN so that they remained out of the same spatial plane of the rest of the molecule by altering the bond angle of the amine nitrogen connected to the methyl groups by manually changing the bond angles until they fit a typical trigonal pyramidal molecular shape. Other attempts that were made included changing the bond order of atoms throughout DMABN and also running coordinate scan calculations rather than geometry optimizations in order to try and obtain the correct spatial configuration and shape. However, several of these calculations failed due to fixed values in the geometry that were not compatible in the calculations. When UV-Vis spectrum calculations were run on the DMABN molecules that did undergo successful

geometry optimizations, many of them failed because the geometry of the molecule was not compatible for the type of calculation. Finally, it was determined that a Comprehensive Mechanics cleanup needed to be performed on each molecule in the WEBMO molecule editor, and the dihedral angles of the methyl groups needed to be fixed through manual configuration of the Z-matrix, followed by a geometry optimization. All UV-Vis spectrum calculations for the in gas phase were run on the Gaussian 03 program with TD (Time-dependent) B3LYP theory and 3-21G basis set. For these calculations, it was necessary to manually change the text in the Preview box to execute the calculation with TD-B3LYP theory in the model chemistry. Finally, molecular orbital calculations for all the geometry optimized molecules were run using the MOPAC program with PM3 basis set.

For analysis of excited state quantum descriptors, raw output data was analyzed in conjunction with UV-Vis spectra absorption graphs to determine the excited with the highest oscillator strength, indicating that it was the excited state most responsible for the highest peak on the absorption graph. From the molecular orbital calculations, the energy between the HOMO and the LUMO was calculated by subtracting the energy of the HOMO from the LUMO. The probable energy required to reach the excited state most responsible for the highest peak in the UV-Vis absorption graph was calculated by subtracting the energy of the LUMO from the excitation energy of the excited state being studied.

Results and Discussion

From comparisons of the raw output data of the excited state quantum descriptors with the graph of the UV-Vis spectrum graph, it was determined that the excited state with the greatest corresponding oscillator strength was very close to the wavelength of highest intensity in the absorption peak. Figure 4 shows the UV-Vis spectra of salicylamide with its highest peak at 187.24 nm. The collected oscillator strengths and wavelengths for all ten excited states calculated for each molecule can be found in Table 1.

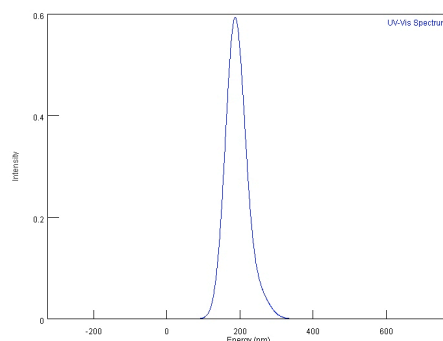


Figure 4. UV-Vis spectra of salicylamide with peak wavelength of 187.24 nm.

It can be deduced that the oscillator strength value of the excited state is correlated with the intensity of the peak at its wavelength. The relationship between the largest oscillator strength value and absorption intensity value was determined to be linear as shown in Figure 5.

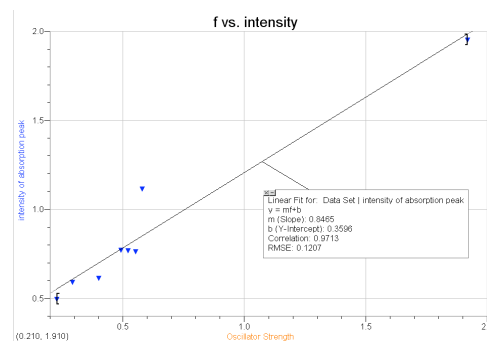


Figure 5. A fairly strong linear relationship ($r=0.9713$) between oscillator strength and the peak intensity was found.

The strong correlation indicated by this set of data points supports the hypothesis that the intensity of the absorption peaks in the spectra are directly proportional to the oscillator strength of the excited state.

From looking at the largest oscillator strength values, it was also determined that higher level excited states, rather than the first excited state, contributed most to the peak intensity of the absorption peaks in the UV-Vis spectrum. Common excited state levels with the largest oscillator strengths ranged from $n=5$ to $n=9$. All the peak wavelengths ranged from approximately 172 nm to 224 nm with excitation energies (EE) that ranged from 5.477 eV to 7.1936 eV. The values for oscillator strengths ranged from 0.2271 to 1.9189. The HOMO-LUMO gap energies ranged from 8.176 eV to

		Excited States										_ of largest absorbance
		1	2	3	4	5	6	7	8	9	10	
DMABN	OS	0.0221	0.5197	0.0655	0	0.0017	0	0.0139	0.5509	0.165	0.0047	176.29
	_(nm)	265.14	256.44	204.23	198.77	185.89	183.08	177.7	174.58	173.7	166.18	
Caffeine	OS	0.098	0.0033	0.005	0.0531	0.024	0.0361	0.0139	0.1633	0.2271	0.0065	200.92
	_(nm)	270.75	262.6	230.78	225.89	212.24	208.68	206.5	200.88	196.19	183.22	
Anthracene	OS	0.0657	0.0017	0	0	1.9189	0	0	0.0615	0	0	223.35
	_(nm)	355.37	299.81	260.15	238.67	223.49	216.55	212.41	199.29	199.04	191.9	
ABN	OS	0.0232	0.4318	0.1387	0.0105	0	0.0648	0.4917	0.0029	0.0818	0.0059	178.68
	_(nm)	255.6	237.02	195.18	194.65	182.14	176.36	172.35	169.92	168.24	164.63	
MMD	OS	0.0773	0.0005	0.0412	0.1875	0.0011	0.005	0.3408	0.5786	0.089	0.0253	186.68
	_(nm)	306.97	262.63	238.42	217.89	190.71	188.14	187.24	183.48	178.64	172.64	
MABN	OS	0.0405	0.4777	0.1042	0.0002	0	0.001	0.0385	0.5202	0.1344	0.0058	176.74
	_(nm)	261.17	246.61	199.58	197.33	182.41	180.98	176.18	171.25	171.25	165.27	
Acet.Anth.Acids	OS	0.1344	0.0002	0.0003	0.1621	0.0006	0.4	0.08	0.0592	0.0001	0.0012	216.63
	_(nm)	291.98	271.28	245.59	235.26	231.27	214.27	198.66	187.89	183.16	182.64	
Salicylamide	OS	0.0103	0.0354	0.0214	0.025	0.0244	0.0079	0.0119	0.2908	0.164	0.0906	187.24
	_(nm)	273.48	250.28	217.51	213.17	210.12	195.6	191.98	188.05	184.06	173.69	

OS: Oscillator Strength

_: wavelength

Table 1. Oscillator strengths and wavelength of each 10 excited states of 8 different fluorescent molecules. Wavelength of largest absorption peak was determined from UV-Vis Spectra

9.273 eV, the EE-LUMO gap energies ranged from 6.4404 eV to 7.6766 eV, and the EE-HOMO gap energies ranged from 13.7957 eV to 16.1952 eV. The summary of these results is shown in Table 2.

These calculation results indicated various trends and relationships within the data. First, a trend was found in which the HOMO-LUMO gap energy for any molecule was always larger than its corresponding EE-LUMO gap energies. This shows that the energy gap required for an electron to cross to reach the next highest orbital level (HOMO-LUMO gap) is always larger than the gap of the succeeding energy orbital jumps that the electron must cross to reach higher excited states.

Also, several correlations were found between quantum descriptors. A fairly strong negative linear relationship was found between the intensity values of the highest absorption peak and the HOMO-LUMO gap energy, as shown in Figure 6.

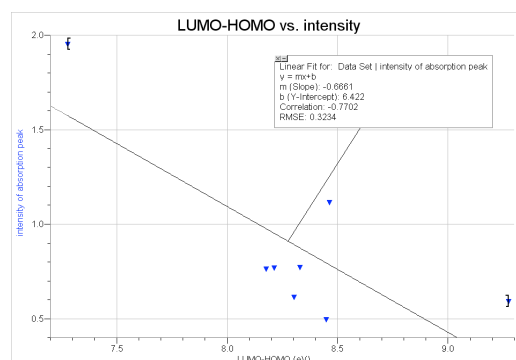


Figure 6. A moderately strong linear relationship between intensity of highest absorption peaks and the HOMO-LUMO energy gap was found ($r = -0.7702$) with regression equation of $y (\text{intensity}) = -0.6661x (\text{LUMO-HOMO}) + 6.422$

Considering the relationship between the intensity of highest absorption peak and the oscillator strength of excited state, a fairly strong relationship between oscillator strengths and HOMO-LUMO energy gaps was also found (Fig. 7).

Molecules	ES #	Oscillator Frequency	_(nm)	E HOMO (eV)	E LUMO (eV)	E (LUMO-HOMO) (eV)	EE-LUMO (eV)	EE-HOMO (eV)	EE-LUMO/HOMO	EE-HOMO/HOMO	Intensity of highest absorption peak
DMABN	8	0.5509	174.58	-8.511	-0.375	8.176	7.4769	15.6129	0.91449364	1.909601272	0.7663
Caffeine	9	0.2271	196.19	-9.002	-0.554	8.448	6.8736	15.3216	0.813636364	1.813636364	0.4975
Anthracene	5	1.9189	223.49	-8.248	-0.97	7.278	6.5177	13.7957	0.895534487	1.895534488	1.9536
ABN	7	0.4917	172.35	-8.813	-0.483	8.33	7.6766	16.0066	0.921560624	1.921560624	0.7735
MMD	8	0.5786	183.48	-9.035	-0.572	8.463	7.3294	15.7924	0.866052227	1.866052227	1.1185
MABN	8	0.5202	173.52	-8.619	-0.407	8.212	7.5522	15.7642	0.919654165	1.919654165	0.7708
Acetylanthranilic Acid	6	0.400	214.27	-8.955	-0.654	8.301	6.4404	14.7414	0.77585833	1.77585833	0.6171
Salicylamide	8	0.2908	188.05	-9.602	-0.329	9.273	6.9222	16.1952	0.746489809	1.746489809	0.5934

** ES: Excited State

**EE: Excited State Energy

Table 2. Summary of computational data

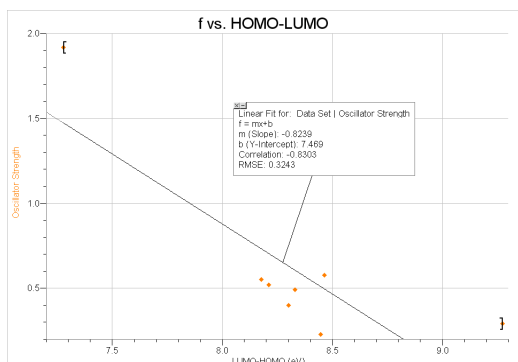


Figure 7. A fairly strong linear relationship between oscillator strength and HOMO-LUMO energy gap was found ($r = -0.8303$) with regression equation of $y(\text{oscillator strength}) = -0.8239 x(\text{LUMO-HOMO}) + 7.469$

Another strong linear regression was derived from plotting the wavelength of the excited state against its corresponding EE-LUMO gap energy (Figure 8).

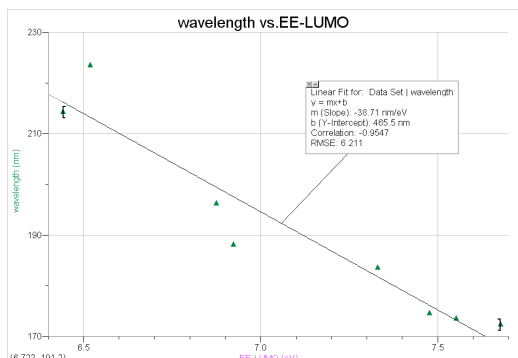


Figure 8. A strong linear relationship between wavelength of highest absorption and energy gap between excited state and LUMO was found ($r = -0.9547$) with regression equation of $y(\text{wavelength}) = -38.71 x(\text{EE-LUMO}) + 465.5$

This relationship suggests that the energy required to promote an electron after the first jump (of HOMO-LUMO gap) is directly proportional to the peak absorbance wavelength of a molecule.

Overall, there were many quantum descriptor matches between which no correlation was found. First, no evident correlation was found between the EE-LUMO gap and the HOMO-LUMO energy gap. This meant that the energy required to get to the excited state was independent of the energy of HOMO-LUMO gap. Additionally, the ratio of the EE-LUMO gap energies to the HOMO-LUMO gap energies was calculated and then plotted against oscillator strength, wavelength, and intensity of absorption peaks; none of these plots showed a reasonable relationship. This result was not a surprise given the fact that EE-LUMO gap and the HOMO-

LUMO energy gap produced no evident relationship.

Conclusions

Various properties of excited state luminescent molecules were calculated including HOMO, LUMO, HOMO-LUMO gap, oscillator strength, and peak wavelength and these quantum descriptors were paired with one another in order to find correlations. Strong linear regressions were found for oscillator strength vs. intensity, LUMO-HOMO gap vs. intensity, oscillator strength and LUMO-HOMO gap, and peak wavelength vs. LUMO.

These relationships help to estimate any of these quantum descriptors when the other is known, which can be helpful in giving a rough prediction of these properties of other unknown luminescent compounds.

Acknowledgement

The author thanks Mr. Robert Gotwals of NCSSM, Dr. Clyde Metz of the College of Charleston, SC, and Dr. Shawn Sendlinger of North Carolina Central University for assistance with this work. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

- Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed May 2007).
- The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed May 2007).
- MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
- Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick,

- A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
5. J. Foresman, R. Fogle, J. Beck. Solvent Stabilization of Charge Transfer Excited States. *Teaching with Cache Workbook*, (2002).
 6. Zachariasse, Klass A. et al. "Intramolecular charge transfer in aminobenzonitriles: Requirements for dual fluorescence." Pure & Appl. Chem. 65(1993): 1745-1750.

QSAR Analysis of Artificial Sweeteners

Kate Lee, Katie Cooper, Caroline Kennedy
North Carolina School of Science and Mathematics, Durham, NC

5 June, 2007

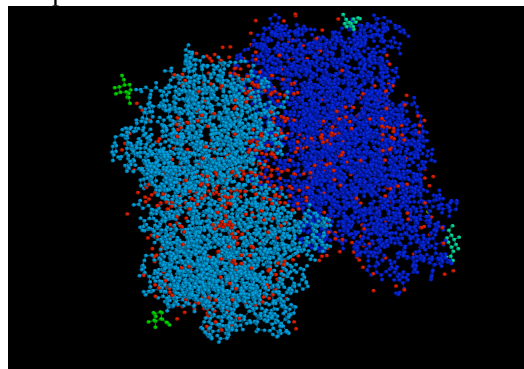
Abstract: This study analyzes the effect of different measured variables on the binding of artificial sweeteners to the T1R2 and T1R3 glutamate receptors on the tongue. A QSAR analysis was used to determine which variables contributed to the goodness of the fit of the molecule into the receptor's active sites. The artificial sweeteners that were studied were saccharin (Sweet N' Low), sucralose (Splenda), and aspartame (Equal). Sucrose, or table sugar, was used as a constant for comparison, and was also included in all of the calculations. Regression equations were used as a means to measure the significance of the different variables in affecting their sweet index (a measure of how well the molecule binds to the receptor). A higher sweet index indicates a tighter, or better, bond to the receptor protein, thus causing a sweeter taste. This study attempted to find the correct combination of variables that most affected the sweet index by using regression equations as a quantitative measure.

Key words: saccharin, sucralose, aspartame, sucrose, QSAR, T1R2, T1R3, sugars, Gaussian, 3-21G

Introduction

Many factors contribute to the overall sweetness of a sugar. Sucrose, or table sugar, is the most common of the sugars used in culinary practices. However, because of various needs such as decreased caloric content, a variety of artificial sweeteners have also come out onto the market. The most common of these sweeteners are Equal (aspartame), Splenda (sucralose), and Sweet N' Low (saccharin). The goal of our project is to perform a QSAR, or Quantitative Structure-Activity Relationship, analysis of these molecules to determine the relationship of their properties to their values on the sweet index. The sweet index is a quantitative system that describes the relative sweetness of various sugars. This index uses sucrose, or table sugar, as a base with a sugar index value of 100. This means, for example, that another sweetener with a sugar index value of 300 is 3 times as sweet as sucrose. A QSAR analysis is a quantitative comparison of chemical structures that correlates chemical properties with another characteristic. In our study, we are correlating chemical properties such as HOMO/LUMO gap, dipole moment, RHF energy, and electrostatic potential with the base characteristic, the sugars quantitative value on the sweet index. The sweetness of a sugar depends on the "tightness" of its fit onto the tongue's sugar receptors. These receptors, called the T1R2 and T1R3 receptors,

are located at the tip of the tongue and contain active sites which bond to the sugar molecules, or glutamates. The better the bond between receptor and molecule, the sweeter that the body perceives the sugar to be. Here is a visualization of the glutamate receptor and the various places to which sugar molecules can bond. The sugar molecules are depicted as green, and the protein receptor is blue and red.



In order for artificial sweeteners to produce the same effect as natural sugars, they must bond similarly to the active site within the receptor. For this reason, we expect the structures of these molecules to be similar in nature. Another goal of the project is to compare the structures of these molecules to investigate similarities and differences that may affect how they bond to the receptor, and thus their sweetness.

Computational Approach

Using the molecular editor builder of WebMO¹ on the North Carolina Computational Chemistry Server², saccharin (C₇H₅O₃N₂S), sucralose (C₁₂H₁₉Cl₃O₈), aspartame (C₁₄H₁₈N₂O₅), and sucrose (C₁₂H₂₂O₁₁) were built. Each was optimized using Gaussian 03³ with a geometry optimization using the PM3 level of theory. Following the optimization, each molecule was run using Gaussian 03³ software with a molecular orbital calculation using HF/3-21G. Dipole moment, RHF energy, HOMO and LUMO energy values, HOMO/LUMO gap, and electrostatic potential minimums and maximums were computed using these calculations. These values were entered into an Excel file and regression analysis was performed on the data in order to calculate correlation.

Results and Discussion

A structural comparison was performed on the molecules to determine similarities and differences between their structures. All of the molecules contain 6-membered rings. However, the nature of these rings differs from molecule to molecule. Saccharin and aspartame both contain 6-membered carbon rings, while sucralose and sucrose have rings consisting of 5 carbons and one oxygen. However, this does not appear to have any significant effect on their sweet index values.

Regression analysis was performed on all of the data collected. This data can be found in Tables 1 and 2, located below.

Molecule	Sweet Index	RHF Energy	Dipole Moment	HOMO	LUMO	HOMOLUMO GAP	ESP max	ESP min
Sucrose	100	1283.43	2.7107	-0.40342	0.18956	0.59298	0.11998	-0.09516
Aspartame	180	-1018	5.8168	-0.27049	0.09733	0.36782	0.12108	-0.12002
Sucralose	600	2430.25	3.1037	-0.41543	0.17077	0.5862	0.13343	-0.08857
Saccharin	300	-939.29	5.9977	-0.39785	0.04295	0.4408	0.11458	-0.08965

Table 1. Calculated Values of Artificial Sweeteners and Sucrose

Relationship	X	Y	R ²	Equation
A	RHF Energy	Sweet Index	0.28993	y= .069343x +264.551
B	Dipole Moment	Sweet Index	0.037186	y= -24.2912x+402.0568
C	RHF Energy (x1), Dipole Moment (x2)	Sweet Index	0.994039	y= .343176x1+287.9481x2-1124.74
D	HOMO	Sweet Index	0.179209	y= -1366.64x-213.114
E	LUMO	Sweet Index	0.008642	y= 301.0365x+257.3245
F	HOMO/LUMO Gap	Sweet Index	0.099598	y= 623.12x-14.6595
G	HOMO (x1), LUMO (x2), HOMO/LUMO Gap (x3)	Sweet Index	0.182188	y= 0x1-1618.51x2+1430.516x3-213.334
H	RHF Energy (x1), Dipole Moment (x2), HOMO (x3), LUMO (x4), HOMO/LUMO Gap (x5)	Sweet Index	1	y= .334094x1+287.3202x2-327.724x3+0x4+0x5-1239.83
I	ESP max	Sweet Index	0.5677	y= 20745.57x-2241.51
J	ESP min	Sweet Index	0.255175	y= 7520.157x+1034.607
K	ESP max (x1), ESP min (x2)	Sweet Index	0.712778	y= 18909.05x1+5756.638x2-1450.8
L	RHF Energy (x1), Dipole Moment (x2), HOMO (x3), LUMO (x4), HOMO/LUMO Gap (x5), ESP max (x6), ESP min (x7)	Sweet Index	1	y= .334094x1+287.3202x2-327.724x3+0x4+0x5+0x6+0x7-1239.83

Table 2. Regression Analysis on Calculated Data

These values show which variables are significant in determining the fit of the molecules into the receptor, and thus their sweetness. R² values, found in Table 2 along with the regression equations of each relationship, are a measure of whether or not the relationship is significant and can be represented linearly. This also measures to what extent the variables being graphed correlate.

Relationship A has an R² value of 0.28993. This shows that the RHF Energy of the molecule alone does not affect the fit of the molecule into the receptor.

Relationship B has an R² value of 0.037186. This means that there is virtually no correlation between the dipole moment and the sweet index of the molecule. This, however, does not necessarily mean that this variable, or any others

that do not have a significant correlation alone, do not have a strong correlation when combined with other variables.

Relationship C has an R² value of 0.994039. While dipole moment and RHF energy taken alone had no effect on the sweet index, together they do have a significant correlation with sweet index.

Relationship D has an R² value of 0.179209. The HOMO value of the molecules does not appear to have any correlation with the sweet index of the molecule.

Relationship E has an R² value of 0.008642. Alone, the LUMO value of the molecules has virtually no correlation with their sweet index.

Relationship F has an R² value of 0.099598. Again, this means that the HOMO/LUMO gap of

each molecule alone does not affect the way that the molecule fits into the receptor protein.

Relationship G has an R^2 value of 0.182188. The HOMO, LUMO, and HOMO/LUMO gap values of the molecules, when combined, still do not produce a significant correlation with the sweet index.

Relationship H has an R^2 value of 1.0. This indicates a perfect correlation between the combinations of the variables RHF Energy, Dipole Moment, HOMO, LUMO, and HOMO/LUMO Gap. However, in the equation, LUMO and HOMO/LUMO Gap have a

coefficient of zero, meaning that they have no significance in the correlation between these values and sweet index.

Relationship I has an R^2 value of 0.5677. This reveals a slight correlation between the ESP max and the sweet index. However, this could be simply a result of chance rather than an actual relationship between these two variables.

Relationship J has an R^2 value of 0.255175. This indicates no real correlation between ESP min and the sweet index of the molecules.

Relationship K has an R^2 value of 0.712778. When combined, ESP max and ESP min have a slight correlation with sweet index.

Relationship L has an R^2 value of 1.0. This again indicates a perfect correlation between all of the variables calculated. However, the regression equation for relationship L is the same as relationship H. ESP max and ESP min, when added to the variables correlated to relationship H, have no effect on the sweet index of the molecules. The only values that have any effect on the sweet index are RHF Energy, Dipole Moment, and HOMO. All other variables have no effect.

The only variables measured that affect the binding of the molecule to the sweet receptor on the tongue are the RHF energy, the Dipole moment, and the highest occupied molecular orbital of the molecule. The coefficient for RHF energy in the regression equation is .334094. This means that there is a positive correlation between RHF energy and the sweet index. The coefficient for dipole moment is 287.3202. Again, this indicates a positive correlation between dipole moment and sweet index value.

There is a negative coefficient associated with the HOMO values of the molecules. This value is -327.724.

Conclusions

The different values that were found to be associated with each of the significant variables can help to explain how the receptor might work in the body and what kind of molecules that it is most likely to bind to.

Because the correlation between RHF energy and sweet index is positive, we can assume that as the RHF energy goes up, the sweet index increases, which indicates a better bond with the receptor. Therefore, molecules with higher RHF energies are more likely to bond well with the receptor and taste sweeter.

Dipole moment also has a positive correlation with sweet index. The higher the dipole moment, the more polar a molecule is. This could mean that the receptor is more likely to accept polar molecules into its active site. The receptor's active sites may be home to atoms that have very high electron affinities that attract the negatively charged end of a polar molecule.

HOMO values, however, have a negative correlation. The smaller the HOMO energy value of a molecule, the more likely it is to react. This means that if the electrons in the highest occupied molecular orbital of the atom are loosely attracted to their nuclei, they are more likely to react. This makes sense because in order for the molecule to bind with the receptor, the electrons in the highest orbital must be very reactive. The lower the HOMO value of the molecule, the more reactive it is, and therefore the more likely it is to bind with the receptor and cause the perception of a sweet taste.

The other variables might not have been significant simply because they do not have any major effect on the binding of the molecule to this specific receptor.

These conclusions were based upon our computational methods and the data collected and not verified by any data on the receptor itself.

Acknowledgement

The authors thank Mr. Gotwals of the North Carolina School of Science and Mathematics for the encouragement and motivation he extends to his students regarding computational chemistry.

The authors would also like to state that Mr. Robert Gotwals is actually younger than dirt, despite popular belief. The authors would also like to extend an offer to Mr. Gotwals of an honorary doctorate from the University of Intro to Comp Chem for his magnificent work in the field of PR for Apple. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

- Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed May 2007).
- The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed May 2007).
- Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
- JC Traeger, RG McLoughlin *J.Am.Chem.Soc.*, **103**, 3647-3652 (1981)

Geometry Optimization of Adsorbed Methanol Vapors on Fluorinated Surfaces of Silicon 100

M. Pham

North Carolina School of Science and Mathematics, Durham, NC

Received May, 2007; Accepted 05 June, 2007

Published online on Comp Chem Moodle (moodle.ncssm.edu)

Abstract: Studies on Silicon 100 have shown that the adsorption of methanol vapors on Silicon 100 or the production of methanol vapors in silica-sol gel matrices increases the amount of methanol that can be produced. These studies indicate that there is promise for Silicon 100 to be incorporated into a mechanism to enhance the efficiency of Direct Methanol Fuel Cells. In this study, an attempt was made to geometry optimize adsorbed methanol vapors on fluorinated surfaces of Silicon 100 to compare experimental data with experimental data. It was found that the geometry optimization portion of the calculations was more difficult than thought to be and only one quick, dry optimization using Tinker software was successful. The other geometry optimizations did not succeed, but the Tinker geometry optimization indicates that the molecule could be potentially geometry optimized if more premature optimizations using redundant checkpoint files in the calculations are performed.

Key words: Silicon 100, Geometry Optimization, 3-21G, 6-31G(d), redundant, checkpoint file

Introduction

In research, alternative energy sources have been a primary focus for many researchers who wish to resolve the petroleum crisis. Although there is still time for an abundant or abundance of resources to be found, researchers know that if the problem is not resolved within the next few decades, the planet could have problems with sustaining the current comfortable standard of living that many nations have.

One form of energy that has been valuable in many portable devices, machines, households, and other materials is electricity. The production of electricity can be produced through various mechanisms, but one particular semiconductor used in many electronic devices is silicon. Silicon technology has proven to be effective for small portable electrical devices, such as CD players, radios, and laptop batteries.

A specific, potential application using silicon is the synthesis of Silicon-100 (Si-100) gel to adsorb methanol vapors. There have been experimental studies that dealt with the adsorption of methanol vapors on fluorinated silicon surfaces and the production of methanol in silica-sol gel matrices. These experiments have demonstrated that methanol production is increased when produced in the porous areas of

Si-100. Thus, the direct application would be to find a way to incorporate methanol production in Si-100 in Direct Methanol Fuel Cells (DMFCs). By doing so, DMFCs would produce more methanol vapors to generate more current, which could be used to power vehicles or electronic devices with longer sustained electricity power.

The use of liquid methanol poses two disadvantages in a Direct Methanol Fuel Cell. One of which pertains to the flammability of the substance and the other of which deals with it being less efficient than methanol in the vapor state. However, utilizing methanol in its vapor state would require the use of a mechanism that would be able to store and release the vapor in the fuel cell. Although there have not been many direct studies for this application in fuel cells, silica gel is a substance that has the ability to absorb methanol. Since silica gel is a non-flammable polymer, it also reduces the flammability issue with methanol fuel cells.¹

Silica gel is a porous, granular amorphous form of silica (SiO_2) and is synthesized from sodium silicate.⁵

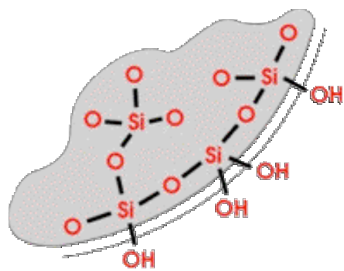


Figure 1: Silica Gel⁵

The gel itself is not actually a gel, but a solid and the gel comes in a variety of names depending on its mesh size. The mesh size difference between various silica gels distinguishes between the differences in the surface area and density of the particle and indicates the type of mesh screen filter the silica gel can be filtered through. Due to the high surface area of silica gel and its many tiny pores, silica gel can adsorb water readily, which is done when preparing silica gel to absorb methanol vapors.¹ This adsorption to silica gel is due to the van der Waals interactions between silica gel and other molecules and capillary condensation, the ability for pores in the silica gel to be filled with liquid that separates from the gas phase.³

Due to silicon's prevalent role as a semiconductor, the surface phases of silicon have been a focus for many researchers for the past thirty years. The surface phases of silicon include the thin layers situated near the surface of silicon that are in equilibrium with crystalline silicon.⁴ One such study that has been conducted involving the surface of silica gel involves the adsorption of methanol vapors on Si-100 silica gel that had different concentrations of fluorine ions bonded to the surface. In this study by R. Nasuto from the Department of Physical Chemistry of Solid Surfaces in 2000, it was seen that as the concentration of bonded fluoride ions increased, the surface of the silica gel began to deactivate and absorbed less methanol vapor. In order to prepare the silica gel for fluorination, the surface of the silica gel was hydroxylated by saturating it with water vapor. Afterwards, the hydroxyl groups were removed by exposing the gel to fluorine ions. Once fluorinated, the adsorbents of silica gel were placed in columns that were exposed to certain amounts of methanol vapors until equilibrium was established between the adsorbed methanol and the silica gel at pressures of 10^{-3} millibars. By

varying the fluorine concentrations, it was observed that the fluoride ions have the ability to form complexes with multiple hydroxyl groups on the silica gel's surface, which causes the pores of the gel to be blocked due to the high electronegative charges. Thus, methanol vapor could not be adsorbed in higher amounts.⁴

Although the mentioned idea of incorporating Si-100 in DMFCs seems reasonable, due to limited resources and time, this study did not focus on engineering a mechanism to increase methanol production in DMFCs via Si-100 gel.

Instead, this study attempts to analyze the structural properties of adsorbed methanol vapor on fluorinated Si-100 through computational methods. Understanding the structural components and properties of a molecule is just as important as knowing that a molecule can be used in practical applications, such as DMFCs. By attempting to understand the structural properties of adsorbed methanol vapor on fluorinated Si-100, certain similar phenomena and/or properties can be looked for in other molecules to find more reliable or cheaper molecules to use as scientists attempt to produce more energy sources.

In computational chemistry, computing resources and compute time are very valuable and limited to researchers. Due to the complexity of the Si-100 molecule, all of the proposed methods could not be carried out within the allotted time span. Thus, the focus of this study in terms of understanding the structural properties of Si-100 involved attempting to geometry optimize fluorinated Si-100 surfaces with methanol vapors adsorbed to the surface..

When studying molecules computationally, one of the first calculations performed on the molecule is the geometry optimization. A geometry optimization calculation done using a specific basis set and theory level is done to ensure that the molecule's quantum descriptors are optimal to place the molecule is at its lowest, most stable energy state. Although a geometry optimization does not appear to be a complex calculation, certain molecules can be very difficult to optimize due to the structure of the molecule.¹⁰ In the event that a molecule is too difficult to geometry optimize using standard protocol, other types of geometry optimizations must be used. Some of these optimizations include natural, redundant, and delocalized

internal coordinate optimizations. In the present study, a redundant optimization was required on the adsorption of methanol vapors on the fluorinated Si-100 gel.¹² This type of optimization will be further explained in the results and discussions section. Finally, when dealing with geometry optimizations, if all of the aforementioned methods for optimizing a molecule do not succeed, single point energy scans are performed instead. In a single point energy scan, the molecule has other calculations (i.e. molecular energy, molecular orbital, transition state, etc) done on it at a specific molecular geometry. Single point energy scans are done when a researcher would like to study specific aspects of the molecule, but does not have enough time to study these aspects when the molecule is at its lowest, most stable energy state.

Computational Approach

On the North Carolina High School Computational Chemistry Server⁸, the molecular editor builder of WebMO⁷ was used to build adsorbed methanol vapors on fluorinated Si-100.

After building the molecule, the molecules were optimized using the “geometry cleanup” molecular mechanics package found in the WebMO molecular editor. With the adsorbed methanol vapors to the fluorinated Si-100 surface attempts to geometry optimize the molecule were made using Tinker, MOPAC, and Gaussian software using Hartree-Fock, B3LYP, AM1, and PM3 levels of theory along with 3-21G, 3-21G(d), and 6-31G(d) basis set types.

Results and Discussion

When performing calculations on molecules, there are two methods that are usually used to optimize molecules. These two methods are *ab initio* and semi-empirical calculations. *Ab initio* calculations are completely theory based, but require much computing time and resources. Since these two factors played a major key in the computational analysis, this study used semi-empirical calculations to geometry optimize the fluorinated Si-100 gel with methanol vapors adsorbed. With semi-empirical methods, half of the calculations are theory based and half of the calculations are based from empirically derived data (experimental data).

Due to the complex structure of methanol vapors adsorbed to fluorinated Si-100 gel along with few experimental studies on the structure, it was found that geometry optimizations using Tinker, MOPAC, and Gaussian at 3-21G, 3-21G(d), 6-31G(d) basis sets and PM3, AM1, Hartree Fock, and B3LYP levels of theory would not succeed in optimizing the molecule. Thus, an attempt to perform redundant checkpoint optimizations was done on the structure. The idea behind the redundant checkpoint optimization involves the notion that the molecule is able to be optimized to a certain extent during each calculation and that the calculation fails when it cannot interpret certain aspects of the molecule. For instance, a common error that caused the geometry optimization to fail in the adsorbed methanol vapor on fluorinated Si-100 structure was the Z-matrix error. In this error, the software could not interpret the internal coordinates for the structure, which means it could not accurately determine the build of the structure.

Thus, with redundant checkpoint optimizations, once a geometry optimization failed, the failed job was used as a starting point to continue the geometry optimization from the point at which the previous job did not succeed. In order to run a checkpoint file, the save checkpoint file tab must be checked and the proper checkpoint file must be chosen when using the geometry optimization.

The idea of redundant checkpoint optimizations can be thought of as premature washing before actual washing. With this concept in mind, attempts are made to clean the molecule with little soap and more water to get the molecule somewhat stable. After the molecule is a bit stable, more soap with water is used in order to get the molecule’s quantum descriptors to be optimal for the molecule’s lowest, most stable energy state.

Although none of the jobs were able to successfully geometry optimize except for one, it is interesting to note that as more optimizations were done and carried out, the longer the jobs took before they failed. With this increase in time before the jobs failed (unless the jobs had to be ended quickly due to the limited computing resources and time available), there is an indication that the molecule was becoming more optimized as more washings occurred. One job that has used the most computing time on the North Carolina Chemistry Server is job 14630, in

which the geometry optimization of the molecule took 17 hours and 49 minutes before it failed.

The one job that did geometry optimize used a molecular mechanics software known as Tinker. This software does a quick, dry optimization on molecules and does not output many parameters that can be used for interpretation or analysis. Instead, the success of this geometry optimization signifies that if this study were to continue, it would be possible to eventually geometry optimize adsorbed methanol vapors on fluorinated Si-100 surfaces with more advanced levels of theory and basis set types for structural analysis of the structure. The geometry optimization using the Tinker software is shown below.

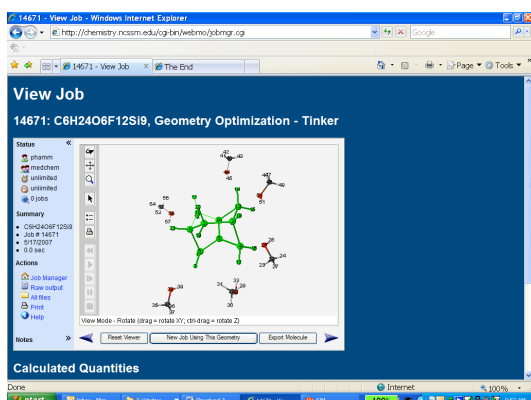


Figure 2: Adsorbed Methanol Vapors on Fluorinated Silicon 100 Surface Structure

Conclusions

The original intent of this computational study was to analyze the quantum descriptors of the Silicon 100 molecule to compare computational data with experimental data. However, it was found that the computational challenge in itself was to successfully geometry optimize the molecule. The adsorbed methanol vapor on fluorinated silicon 100 surface was successfully geometry optimized using Tinker, but was unsuccessful using all other software along with varying degrees of basis set types and levels of theory. Although the geometry optimizations were unsuccessful, the idea behind performing redundant checkpoint files appears to be the route that should be taken to successfully geometry optimize the molecule. This can be noted from the successful Tinker geometry optimization, which indicates that many more

premature optimizations are needed until the molecule can be optimized to its lowest, energy state. From this research, it has been seen that computational studies dealing with complex structures using semi-empirical methods are quite challenging, even with only a geometry optimization and no other calculation available to interpret the molecule.

Acknowledgement

The author thanks Dr. Myra Halpin of the North Carolina School of Science and Mathematics for providing the resources that were used when carrying out the proposed experimental ideas that were not able to be completed.

The author thanks Mr. Robert Gotwals and Steven Lin of the North Carolina School of Science and Mathematics, NC, for providing the author with computational background in this work. Appreciation is also given to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

1. "The Future of Fuel Cells." The Future of Fuel Cells. Ben Wiens Energy Science Inc.. 27 Apr 2007 <<http://www.benwiens.com/energy4.html>>.
2. Lifshits, V.G., A. A. Saranin, and A. V. Zotov. Surface Phases on Silicon Preparation, Structures, and Properties. West Sussex: John Wiley & Sons Ltd. ,1994.
3. "Monolayer and multilayer adsorption, micropore filling and capillary condensation ." Monolayer and multilayer adsorption, micropore filling and capillary condensation . 05 Sept 2002. 27 Apr 2007 <http://www.iupac.org/reports/2001/colloid_2001/manual_of_s_and_t/node17.html>.
4. Nasuto, R. "The Adsorption of Methanol Vapors on Silica Gel Si-100 and Its Surface Containing Different Concentrations of Chemically Bonded Fluoride Ions." Journal of Thermal Analysis and Calorimetry 62(200): 581-585.
5. "Silica Gel." Silica Gel. Grace Davidson. 25 Apr 2007 <http://www.gracedavison.com/eusilica/adsorbents/product/silica_gel.htm>.
6. "What is silica gel and why do I find little packets of it in everything I buy?". April 01, 2000 <<http://science.howstuffworks.com/question206.htm>> (April 26, 2007)
7. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed May 2007).
8. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed May 2007).
9. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N.

- Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
10. Gotwals, Robert, and Shawn Sendlinger. A Chemistry Education's Guide to Molecular Modeling. Durham: 2007.
 11. MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
 12. Reveles, Ulises, and Andreas M. Koster. "Geometry Optimization in Density Functional Methods." Journal of Computational Chemistry 25(2004): 1109-1116.

Potential Energy Scans of 1,1-Dichloroethene

MEREDITH COLE and CLAY DUNWELL

Intro to Computational Chemistry Class, North Carolina School of Science and Mathematics, 1219 Broad Street, Durham, USA 27705

Abstract: A computational examination of the potential energy of 1,1-dichloroethene (trans) was undertaken to compare the output differences between *ab initio*, Hartree-Fock 3-21G, and two semi-empirical methods, AM1 and PM3. The geometric optimization of the molecule, used in all potential energy scans, was computed at the Hartree-Fock level of theory using a 6-31G(d) basis set. We found the AM1 semi-empirical method was closer to the *ab initio* results than the PM3 semi-empirical method, but both semi-empirical methods were very inaccurate in comparison to the *ab initio* method with percent errors ranging from 99.87589% to 161.37303%.

Key words: potential energy scan; AM1; PM3;

Introduction

The purpose of our investigation was to compare the AM1 and PM3 semi-empirical methods for potential energy scans. We sought to find which semi-empirical method was close to the *ab initio* method because computational time and CPU power needed to use semi-empirical method is significantly less, which is extremely useful in the study of larger molecules.

The two semi-empirical methods that we used, AM1 and PM3, are both based on the Neglect of Differential Diatomic Overlap integral approximation. The only differences between the PM3 and AM1 methods are: 1) PM3 employs two Gaussian functions for the core repulsion functions, instead of the variable number used by AM1; 2) the numerical values of the parameters are different in that AM1 takes some of the parameter values from

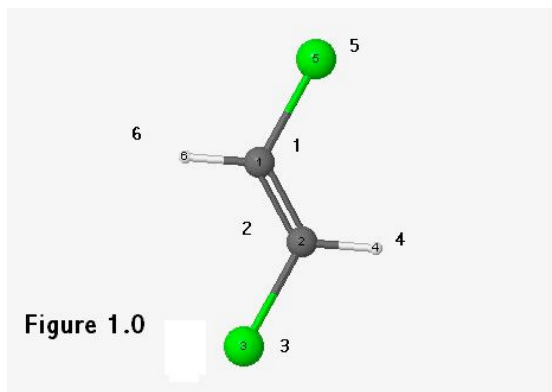
spectroscopic measures while PM3 treats them as values that can be optimized. The potential energy of a molecule is the single point energy of a molecule in a specific molecular geometry. Often known as molecular energies, single point energies in our study are given in units of Hartrees, which can be converted to many other common energy units. A potential energy scan calculates the single point energy at multiple angles and bond lengths as set by the user, returning values of the molecular energy of the molecule in each of the geometries. Usually these values are compiled into a graph for the user to observe the change in energy, allowing the user to find the global minimum or the angle or bond length the molecule is most likely naturally occur in.

Computational Methods

All the *ab initio* and semi-empirical calculations, geometric optimization and potential energy scan, were conducted using Gaussian 03 suite of programs. The 1,1-dichloroethene was

geometrically optimized using a 6-31G(d) basis set at the HF level of theory. Following the optimization, we attempted to conduct potential energy scans of each dihedral angle in the molecule using HF/3-21G. Using Figure 1.0, the potential energy scan of the 6-1-

2-3 dihedral angle failed due to a new curvilinear step not converging. Next, the 4-1-2-3 dihedral angle potential energy scan also failed due to a convergence failure. Lastly, the 5-1-2-3 dihedral angle ran complete with a run time of 00:02:59 with thirty-degree angle increments for a total of twelve geometries. Because of our success with this dihedral angle, we decided to use it in our comparison of *ab initio* and semi-empirical methods. We ran potential energy scans of the 5-1-2-3 using AM1 and PM3 semi-empirical methods with run times of 00:01:45 and 00:01:49 respectively, again with thirty-degree angle increments.



Results and Discussion

We took our *ab initio* results as the base for our comparison of the semi-empirical methods because *ab initio* methods are more accurate due to their pure mathematical approach. All the methods found 180 degrees as the lowest geometry with the lowest molecular energy, which means that configuration is the naturally occurring configuration. The *ab initio* run returned molecular energy results ranging from -991.01559

to -989.6922 Hartrees as shown in Figure 1.1 below.

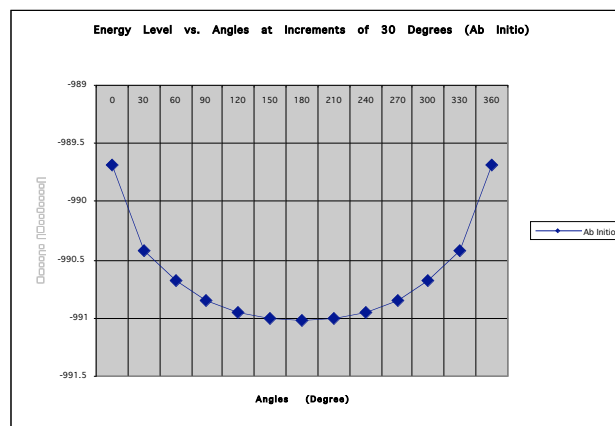


Figure 1.1

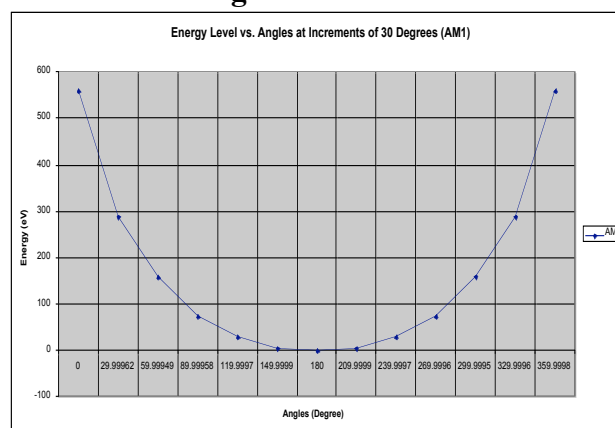


Figure 1.2

The AM1 run returned molecular energy results ranging from -1.22991862 to 559.72592381 Hartrees as shown in Figure 1.2 above. Additionally, the PM3 run returned molecular energy results ranging from 5.30873037 to 607.4041 Hartrees as shown in Figure 1.3 below.

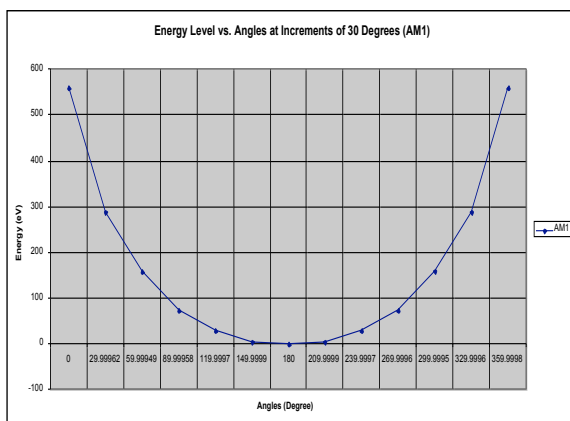


Figure 1.3

All of these graphs share the same parabolic shape with a global minimum at 180 degrees, indicating that every

method found the lowest energy at this geometry. However, in comparison to the *ab initio*, the semi-empirical methods produced significantly larger energies. A chart showing the molecular energy values for each angle along with the percent error in comparison to the *ab initio* is shown below in Table 1.0. It should be noted that percent errors get significantly less the as the angles get closer to 180 degrees due to the overall scale of each method. This can be viewed in the graph in Figure 1.4, which superimposes Figures 1.1 through 1.3 on the same graph and scale.

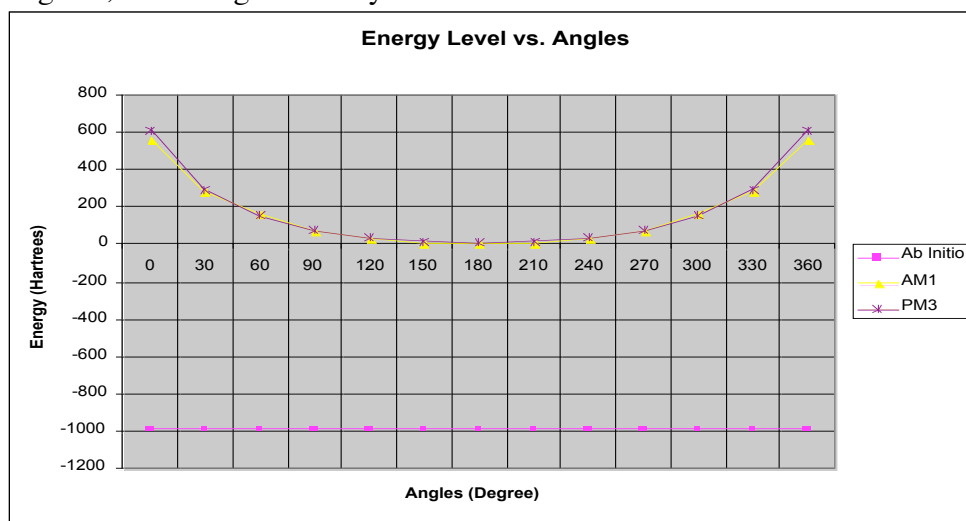


Figure 1.4

Angle	Ab Initio	AM1	AM1-% diff	PM3	PM3-% diff
0	-989.6922	559.7259238	156.55556	607.4041	161.37303
29.99962	-990.42316	287.9076337	129.06915	291.5033	129.43219
59.99949	-990.6812	157.1472041	115.86254	151.5373	115.29627
89.99958	-990.85245	73.7951172	107.44764	74.1779	107.48627
119.9997	-990.95185	28.30067845	102.85591	34.81423	103.51321
149.9999	-991.0003	5.74798702	100.58002	12.32429	101.24362
180	-991.01559	-1.22991862	99.875893	5.30873	100.53569
209.9999	-991.0003	5.74798702	100.58002	12.32429	101.24362
239.9997	-990.95185	28.30067845	102.85591	34.80795	103.51258
269.9996	-990.85247	73.78884211	107.44701	74.1779	107.48627
299.9995	-990.6812	159.4815394	116.09817	150.4391	115.18542
329.9996	-990.42326	287.8699831	129.06535	291.4656	129.42839
359.9998	-989.69229	559.7196487	156.55492	607.3978	161.37239

Table 1.0

Conclusion

From this investigation, we determined that AM1 is a closer semi-empirical

approximation to the HF/3-21G level of theory for potential energy scans of this molecule, 1,1-dichloroethene (trans). Based on this data, we would continue this investigation with molecules of a similar structure in increasing size to determine whether if AM1 is more accurate for all of these molecules.

Acknowledgements

The authors thank Robert Gotwals for assistance and inspiration for this work. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

11. Dewar, M. J. S., Zoebisch, E. G., Healy, E. F. and Stewart, J. J. P., *Journal of the American Chemical Society*, **107**, 3902, (1985).
12. Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
13. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
14. Gotwals, R., Sendlinger, S. A *Chemistry Educator's Guide to Molecular Modeling*, manuscript in press, 2007.
15. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
16. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

Computational Comparison of Pseudoephedrine and Methamphetamine

A. Garza & D. Frick

North Carolina School of Science and Mathematics, Durham, NC

Received 19 May, 2007; Accepted 5 June, 2007

Abstract: There are many similarities between the illegal drug Methamphetamine and legal nasal decongestant medicine Pseudoephedrine. If Pseudoephedrine were to become easily accessible, as it was when the medicine called Sudafed was made OTC, it would be easy to determine a recipe to create Meth. The goal of this project was to determine exactly how similar these drugs are and why it was so easy for drug dealers to create Meth. In addition to this, the replacement for Sudafed, Phenylephrine, was also studied. Our results showed that the structural similarities of Pseudoephedrine and Methamphetamine were shockingly similar. The bond angles of the two drugs varied by at most by 2.949 degrees at critical points connecting the common benzene ring. However, the replacement Phenylephrine has very different bond angles and therefore structures. The energies of all of the drugs were also determined and evaluated. It was found the HOMO and LUMO energies of Sudafed and Meth were most similar. This implies that the reactivity of Sudafed and Meth will be very similar because reactions will occur with the HOMO and LUMO's in order to stabilize the compound. Phenylephrine's HOMO and LUMO energies are similar, yet are not as close to Meth as Sudafed's are. This is reasonable due to the fact that Phenylephrine must be similar to Pseudoephedrine in order to have the same effects in the body. It appears to be that the very similar structures and energies of Pseudoephedrine and Methamphetamine are the main reasons why it was so easy for the common citizen to create Meth from Sudafed.

Key words: pseudoephedrine, phenylephrine, methamphetamine, bond angle, dihedral angle, HOMO energy, LUMO energy, electron to electron repulsion, molecular orbitals

Introduction

It is a well known fact that the extremely harmful and illegal drug Methamphetamine - also known as Crystal Meth when it is produced in its crystallized form - can be manufactured from using a common nasal decongestant, Sudafed, and other household products. It is the main ingredient in Sudafed - Pseudoephedrine - that makes this reaction possible. The other reactants used in this reaction are used only to slightly alter the Pseudo ephedrine's structure and properties.

How exactly the Meth was made is not important; why it was possible to make Meth is important in order to determine a replacement drug to Pseudoephedrine. Determining this replacement medicinal drug that will not be able

to be made into Methamphetamine can be done safely using Computational Software such as the North Carolina High School Computational Chemistry Server¹. Unfortunately, a replacement drug is already being produced and sold. This drug is called Phenylephrine. However, there have been some reports from credible sources such as the British Journal of Clinical Pharmacology and the British Medical Journal² that Phenylephrine is an illogical response to the Meth issue because Phenylephrine is very ineffective when compared to Pseudoephedrine.

Therefore, in this research, the three drugs mentioned - Pseudoephedrine, Methamphetamine, and Phenylephrine - were studied on the basis of structure and various important energies. From the energies of the drugs one was able to determine its reactivity.

The goal of this research was to determine why Meth was so easily manufactured in terms of similarities between Methamphetamine and Pseudoephedrine and to identify main locations for further studies to look at in order to determine more effective replacement drugs.

Computational Approach

Before the molecules were ever run, one had to determine the known arrangement of the atoms in the drugs; these were found on the very useful database searching website ChemFinder.com³. After the structures were found they were then put into the molecular editor builder of WebMO⁴ on the North Carolina High School Computational Chemistry Server. All of the molecules were modified using a Comprehensive Clean-Up in the WebMO builder.

Once ready for Geometry Optimizations, which theory and basis sets to use were being debated. The argument was to use a platform that ran molecules to give the most reliable data, yet it still ran the job in a reasonable amount of time so that the many other researchers could be able to run their jobs as well. Knowing that the Geometry Optimization was the longest calculation that needed to be run we decided to run this one early to determine the best method to use. The main focus of our research, Pseudoephedrine was optimized using Gaussian⁵ with a 6-31G (d) basis set. We did not expect this run to be short by any means, but knowing that most of our research dealt with structure we wanted to get the best results and decided to experiment with this run. As expected, the run time was enormously large at 1 hour 22 minutes and 9 seconds. Where this time would not be so bad if the entire server was dedicated to our project, we had to be reasonable with our run times. We decided that it would be best to optimize everything using the MOPAC⁶ software with a PM3 basis set. The results of this calculation would not be as accurate as the previous job, yet it would still give reasonable values within a reasonable amount of CPU Usage Time. The fact that there were still other calculations to run on the molecules was also taken into consideration.

The next issue that arose was to determine which calculations will put out the results that we desired. It was determined that in order to successfully compare the molecules that Molecular Orbitals must be observed. This is

because molecular orbitals are where reactions will occur and ultimately can determine the behavior the molecule itself. So it was decided that the job we should run would be a Molecular Orbitals job. An ab initio method was then determined to be the best approach because it would give the most accurate results when compared to other theories. Therefore Gaussian was used to run the Molecular Orbital Calculations. To determine the basis set that needed to be run we decided to experiment again. Using the main focus of our research, Pseudoephedrine, we used two different model chemistries were used to determine molecular orbitals; PM3//HF/6-31G (d), and PM3//HF/6-31G (d,p). Although the diffuse and doubly polarized split valence basis set did produce better results, its run time was 19 minutes and 47 seconds. However, the polarized split valence basis set only took 1 minute and 32 seconds. Reasoning that the 6-31G (d) basis set will still give respectable results that are publishable, and that there were still many more jobs to be run – both on our research and other groups – it was determined that we would use the 6-31G (d) basis set. All three Molecular Orbital calculations were run using the following model chemistry: PM3//HF/6-31G (d).

Finally it was decided that a Natural Bond Orbitals calculation needed to be run. This was because we wanted to see how strongly the bonding orbitals were holding on in each of the molecules to determine weak points that could be isolated to create methamphetamine or weak points needed to react in order to be an effective nasal decongestant. Given the knowledge that the Molecular Orbitals calculations worked best under the 6-31G (d) basis set, it was used again in this calculation. All three molecules were run for Natural Bond Orbital calculations using the following model chemistry: PM3//HF/6-31G (d).

It came to our attention after running all of our calculations that a Natural Bond Orbital calculation will produce the same results as a Molecular Orbital calculation with some extra data included. Therefore all of our data was collected from the Natural Bond Orbitals under the idea that all the data in the other calculations were included in the data we were observing. A table of the CPU usage times and all of the calculations can be seen in Figure 1.

Molecule	Calculation	Theory	Basis	Run Time
Pseudoephedrine	Geo Opt	MOPAC	PM3	38.3
Pseudoephedrine	Geo Opt	HF	6-31G (d)	1:22:09
Pseudoephedrine	MO	HF	6-31G (d)	1:32
Pseudoephedrine	MO	HF	6-31G (d,p)	19:47
Pseudoephedrine	NBO	HF	6-31G (d)	1:36
Phenylephrine	Geo Opt	MOPAC	PM3	3.1
Phenylephrine	MO	HF	6-31G (d)	1:19
Phenylephrine	NBO	HF	6-31G (d)	1:29
Methamphetamine	Geo Opt	MOPAC	PM3	27.8
Methamphetamine	MO	HF	6-31G (d)	1:21
Methamphetamine	NBO	HF	6-31G (d)	1:25
		Total Runtime		*1:51:42.7

Figure 1

Results and Discussion

Basic Structures

The data produced suggested that the similarities between methamphetamine and pseudoephedrine were extremely great in regards to their structural geometries and arrangements (all visual references can be seen in figures 2, 3 and 4 which are the molecules methamphetamine, phenylephrine, and pseudoephedrine respectively).

Methamphetamine and Pseudoephedrine were the most similar in structures. Both molecules included a benzene ring and a “tail” which included one nitrogen atom and four other carbon atoms. The major difference in the two molecules was that the pseudoephedrine contained one oxygen atom and the methamphetamine did not; because of this the “tails” were also different with the pseudoephedrine “tail” circling around the benzene ring and the methamphetamine “tail” being more stretched out. The differences in the geometries can be explained because of the lone pairs on the oxygen atom on the pseudoephedrine; it is assumed that the electrons of the carbons on the “tail” repel the lone pairs creating the different bond angles on the pseudoephedrine.

On the contrary, phenylephrine has a fairly linear geometry as well as a considerably less number of carbon atoms compared to the other two molecules. It also has two oxygen atoms whereas pseudoephedrine only has one and methamphetamine does not have any. The pseudoephedrine and methamphetamine have a distinct bond angle that give it more of a “corner” shape; yet, the phenylephrine does not have this characteristic. This is one possible reason why phenylephrine can not be made into methamphetamine, a considerable amount of

elements would have to be added and removed in order to get the arrangement of the methamphetamine molecule.

Bond Angles

The bond angle connecting the three carbons at beginning of the “tail” of the pseudoephedrine with one of the carbons in the benzene ring was 111.344 degrees. This bond angle differs by 2.949 degrees from the methamphetamine bond angle that is also the angle connecting the “tail” to the benzene ring, which is 114.293 degrees. These close bond angles at the critical point connecting the “tail” to the benzene indicate that the structures are extremely similar. It should be noted that due to the oxygen’s lone pairs the dihedral angle between the “tail” and the ring will adjust it. For pseudoephedrine, the dihedral angle between the connecting carbons to the benzene ring and the only nitrogen is 58.625 degrees. The dihedral angle for methamphetamine in the same respective location as previously mentioned is -57.239. These dihedral angles are almost exactly the same in terms of magnitude; however the signs attached to them are opposite indicating that the “tail” as a separate entity itself is simply rotated around the ring to “tail” bond due to the electron to electron repulsions in pseudoephedrine. Because there are no oxygen atoms in methamphetamine, the torsion of the “tail” does not exist. A PES scan was considered, yet due to the fact that the contortions of the molecules were because we were in fact observing two separate molecules themselves, NOT CONFORMERS of a single molecule, the jobs already ran were considered suitable for this situation.

Phenylephrine however is completely different than the other two molecules. The dihedral angle connecting the nitrogen and the benzene ring is 149.459 degrees. This is nearly linear. This extreme difference can be explained due to the extra oxygen and – more so – the significant decrease in number of carbon atoms located on the “tail”. One specific characteristic that can be observed between pseudoephedrine and phenylephrine is the bond angles between the benzene ring, the connecting carbon and the oxygen atom. This characteristic can not be observed on methamphetamine because there are no oxygen atoms in methamphetamine. However, the bond angle mentioned above on pseudoephedrine is 111.155 degrees. The

SUD-NBO

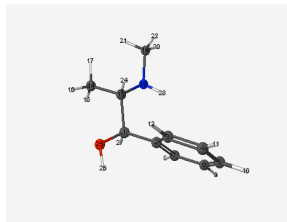


Figure 4

HOMO & LUMO Energies

In the previous sections, the main focus has been on the molecular structures. The properties that will be evaluated in this section can be used to determine the molecules reactivity. The HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital) are very important aspects to consider for this type of observations. This is because the HOMO and LUMO are the most likely locations where a reaction will occur. The reaction is likely to occur there because the electrons in the HOMO have the most energy and therefore are the electrons most willing to react. The LUMO is a likely location for a bond to occur as well because any invading electrons from another molecule will fill into the LUMO. This is why comparing the energies of those orbitals can create an idea of how reactive a molecule is.

Our results showed that the most similar HOMO and LUMO energies were between pseudoephedrine and methamphetamine. The HOMO energy of pseudoephedrine was -0.33189 Hartrees, which is slightly more negative than methamphetamines HOMO energy of -0.3215 Hartrees. The LUMO energy of pseudoephedrine is 0.13303 Hartrees and the energy of the methamphetamine is 0.14107 Hartrees. This indicates that pseudoephedrine is slightly more stable than methamphetamine because the lower the energy, the more stable the molecule. This assumption is in agreement with the overall energy of the molecule which was included in the calculation. The energy of pseudoephedrine was -516.6978717 Hartrees, while the overall energy of methamphetamine was -441.8491096. This proves that the energy of the HOMO and LUMO can predict the stability of the molecule, and the data here suggests that methamphetamine is more reactive than pseudoephedrine.

The most stable molecule studied was phenylephrine. This molecule had HOMO energy of -0.31576 Hartrees and LUMO energy of 0.12751 Hartrees. Suggesting that phenylephrine is less reactive comparatively. This is in agreement with the overall energy of the -552.5004989 Hartrees. This data suggests the phenylephrine is the most stable and least reactive, followed by pseudoephedrine and then methamphetamine. A table of all of the energies can be viewed in Figure 5.

Molecule	RHF Energy (H)	Homo Energy (H)	LUMO Energy (H)
Pseudoephedrine	-516.6978717	-0.33189	0.13303
Phenylephrine	-552.5004989	-0.31576	0.12751
Methamphetamine	-441.8491096	-0.3215	0.14107

Figure 5

When comparing the locations of the HOMO and LUMO of each molecule (which can be viewed in figure 6). It can be seen that both methamphetamine and pseudoephedrine include some of their HOMO on the carbon-nitrogen “tail”. However, phenylephrine’s HOMO is only on the benzene and the unique oxygen atom connected to the ring. This further supports our theory that the purpose of that oxygen is to hinder the ability for one to create meth from phenylephrine.

All three molecules LUMOs are mainly on the benzene ring; for methamphetamine that is the only location of the LUMO. However both pseudoephedrine and phenylephrine have a piece of their LUMO extending to the common oxygen on the “tail”. This suggests that there must be some reaction that occurs around that oxygen in order for the nasal decongestant characteristic of the drugs to work.

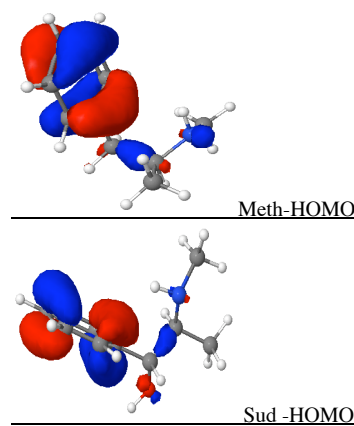


Figure 6

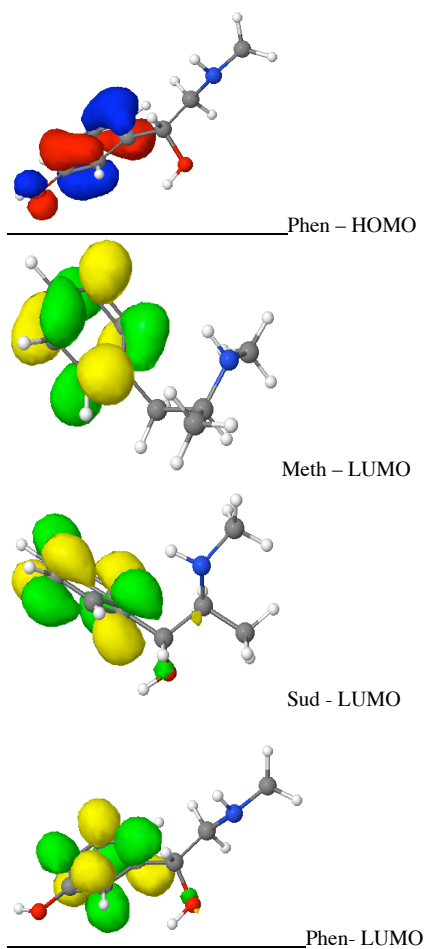


Figure 6

Conclusions

The many similarities between methamphetamine and pseudoephedrine explain why it was possible for criminals to easily manufacture the illegal drug. The main likenesses between the two drugs were based on their structures. Meth and pseudoephedrine differed by only one oxygen atom and an oppositely positioned dihedral angle. The oppositely positioned dihedral angle is assumed to be the cause of the major difference in the two molecules. By removing the oxygen from pseudoephedrine and relieving the stress of electron to electron repulsion, methamphetamine is made into a more unstable and reactive molecule.

The similarity between pseudoephedrine and phenylephrine was the oxygen atom connected to the “tail”. On both molecules the LUMOs were located at this atom and it can be assumed that

this atom plays a vital role in the nasal decongestant abilities of the two drugs.

Phenylephrine is ultimately impossible to be made into meth because of its stability. This stability is caused by the reducing the number of carbon atoms and adding one oxygen atom. Because the arrangement of phenylephrine makes it so stable, it is reluctant to be rearranged into meth, and is therefore successful in that aspect of its intended purpose.

Unfortunately, there are reports that phenylephrine is ineffective when compared to pseudoephedrine. On the premise of this research, a better replacement drug must contain a benzene ring, a connecting “tail” that is composed of carbons, one nitrogen atom, and one oxygen atom that is more similar to pseudoephedrine (not linear as in phenylephrine). The lone oxygen atom should be the location of the LUMO, and the HOMO should be diverted AWAY from the “tail”. In this way, the created molecule should be relatively stable while still having as many similarities to pseudoephedrine.

Acknowledgements

Thanks to Mr. Robert Gotwals for his teaching and guidance in the Computational Chemistry course at the North Carolina School of Science and Math. Without him, the authors of this paper would not have been able to write this paper.

Appreciation is also extended to the Burroughs Welcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

- The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed May 2007).
- Ronald Eccles; British Medical Journal, available from <http://www.bmj.com/cgi/eletters/332/7538/382-b> (accessed May 2007)
- ChemFinder.com - Database and Internet Searching, <http://chemfinder.cambridgesoft.com/> (accessed May 2007)
- Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed May 2007).

21. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
22. MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.

Comparative Analysis of Common Explosive Compounds

J. Hamilton, N. Gallinger

North Carolina School of Science and Mathematics, Durham, NC

Received, Accepted 5 June, 2007

Abstract: Three explosive compounds – trinitrotoluene, nitroglycerin, and RDX – were computationally analyzed for values of molecular energy and dipole moment. These values were compared to major detonative features of the explosives, including explosive velocity and autoignition temperature. Computational efforts made extensive use of the Gaussian engine and the B3LYP theory level. Due to limited data availability, some results were ambiguous.

Key words: TNT, trinitrotoluene, RDX, nitroglycerin, PETN, explosives, explosive velocity, velocity of detonation, auto-ignition temperature, molecular energy, dipole moment

Introduction

The detonative character of an explosive compound can be quantified experimentally both by the kinetic energy they deposit in their surroundings – measured by explosive velocity or velocity of detonation – and by their chemical stability – measured in one way by their autoignition temperature. That is, the temperature at or above which the compound may spontaneously decompose. However, you can also tell how much energy a molecule has computationally by performing molecular energy calculations on them.

Trinitrotoluene is an extremely explosive chemical compound with the formula $C_6H_2(NO_2)_3CH_3$. What makes TNT so explosive is the NO_2 groups attached on it. These NO_2 groups are in a very unstable position and when even the slightest amount of energy is acted upon TNT, these NO_2 groups violently respond by flying off and releasing an extremely large amount of energy. At STP, TNT is a yellow colored solid and when it explodes, it takes on a sooty appearance because of its creation of carbon in the explosion. TNT has a high activation energy, and is difficult to detonate relative to primary explosives like nitroglycerin.

Nitroglycerin is also a very explosive chemical. Its chemical formula is $C_3H_5(NO_3)_3$. Nitroglycerin appears as colorless, oily heavy liquid at STP and can detonate merely on contact. Nitroglycerin is what is used to make dynamite and is an extremely explosive material. In its pure form, nitroglycerin is one of the most powerful explosives, but pure, it is extremely hard to transport because of its explosive nature and it quickly breaks down into its components. Comparable to TNT, nitroglycerin has NO_3 groups which are its main cause for its explosion. TNT has NO_2 groups and it stands to reason that nitroglycerin's NO_3 groups would put off more energy. TNT, however, has a solid central benzene ring which nitroglycerin does not, this may account for nitroglycerin's instability, and for TNT's extremely high activation energy.

RDX, also called Cyclotrimethylenetrinitramine, is found in its pure form as a crystalline solid. Its chemical formula is $C_3H_6N_6O_6$ and like TNT and has NO_2 groups that are most likely at fault for its explosive nature. RDX is mainly a

military explosive and is perfectly stable and easily storable. However, when used as an explosive it is usually mixed with other chemicals to supplement it. RDX has a carbon-nitrogen ring instead of a benzene ring like TNT. These two chemicals with rings are much more stable than their nitroglycerin counterpart, which has no ring in its makeup and proves to be a highly unstable chemical.

Pentaerythritol tetranitrate, more commonly known as penthrite or PETN, while due to computational issues not actually present in the results and analysis portion of this paper, is still worth a brief discussion. It, like nitroglycerin, lacks a central ring structure, benzene or otherwise. Since it and nitroglycerin are the most sensitive to shock and friction of the group, the trend is obvious. PETN is a white crystalline solid used in small caliber ammunition, upper charges in land mines, and the explosive core of detonation cords. It is also considered the benchmark compound between primary and secondary explosives – any compound more explosive than it is considered primary, while any less is secondary.

What all these explosives have in common is some sort of repeated NO_x group that is unstably bonded to the molecule, and a tendency to respond violently to a range of activation energies with substantial releases of kinetic energy. All of these chemicals except nitroglycerin appear as solids at STP, and given its similarly unique relatively high sensitivity, these two features may share a common chemical cause.

Computational Approach

Using the molecular editor builder of WebMO¹ on the North Carolina High School Computational Chemistry Server², the molecules $C_3H_6N_6O_6$ (RDX), $C_3H_5(NO_3)_3$ (nitroglycerin), and $C_6H_2(NO_2)_3CH_3$ (TNT) were created. The Tinker and Gaussian engines were used to successively optimize their geometries, with the latter optimization occurring using the B3LYP//STO-3G level of theory and basis set. Initial attempts at these optimizations were done using the Hartree-Fock theory level and 3-21G basis set, and failed. To remedy these failures, the pre-optimization using Tinker was run to minimize the work required of the Gaussian optimization, the theory level was

changed to B3LYP, and the basis set minimized to STO-3G.

A fourth molecule, C₅H₈N₄O₁₂ (PETN), was also attempted, but failed repeatedly and was unresponsive to changes in engine, theory level, and basis set size. It was attempted at B3LYP//STO-3G (Gaussian), PM3 (MOPAC), and AM1 (MOPAC).

Molecular energy calculations were then run on the optimized molecules, using the Gaussian engine and again using B3LYP level of theory, this time with a basis (3-21G) basis set. Values for velocity of detonation and autoignition temperature were retrieved from an online data base, though one molecule (TNT) lacked an autoignition temperature due to its relative insensitivity to heat.

These values were plotted separately against both the molecular energy and dipole moment values taken from the Gaussian molecular energy calculations mentioned earlier. This was done graphically, using Vernier Software's Graphical Analysis software.

Results and Discussion

Using Vernier Software's Graphical Analysis software, molecular energy and dipole moment were plotted against velocity of detonation and autoignition temperature. The last property is of somewhat limited use, because trinitrotoluene lacks such a feature, and penthrite proved unable to be computationally analyzed, reducing us to only two data points for the analyses regarding autoignition temperature: RDX and nitroglycerin. The four graphs follow:

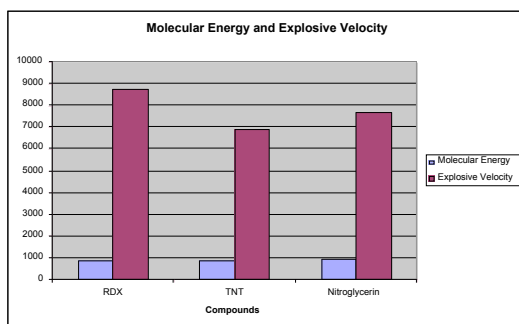


Table 1.1

Table 1.1 shows values of molecular energy and explosive velocity for TNT, RDX, and nitroglycerin. A lack of any clear correlation given the limited data points made a line graph or linear regression pointless.

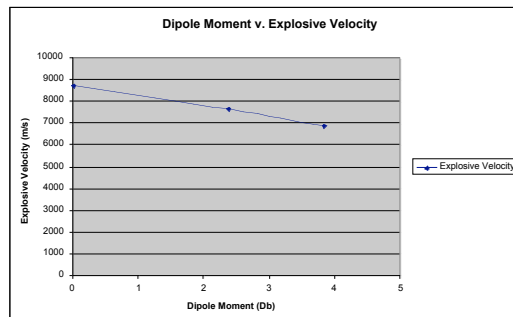


Table 1.2

Table 1.2 graphs the dipole moment of TNT, RDX, and nitroglycerin against their explosive velocities. It shows a fairly strong negative linear correlation, which is an unusual result given that one would expect a highly polar molecule to be less stable.

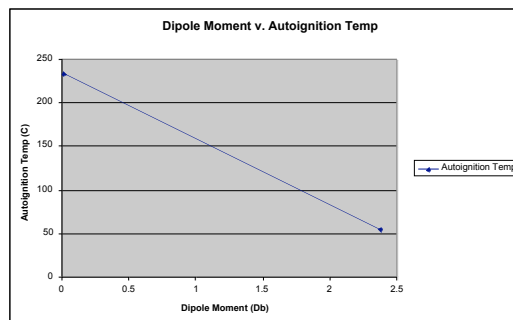


Table 1.3

Table 1.3 shows the relationship between dipole moment and autoignition temperature for the data points RDX and nitroglycerin. Though the literally minimal number of data points makes any analysis tentative, the negative correlation is an expected result – it indicates that as explosive compounds grow more polar, the ambient energies at which they can spontaneously decompose are lower and lower, meaning they are more unstable.

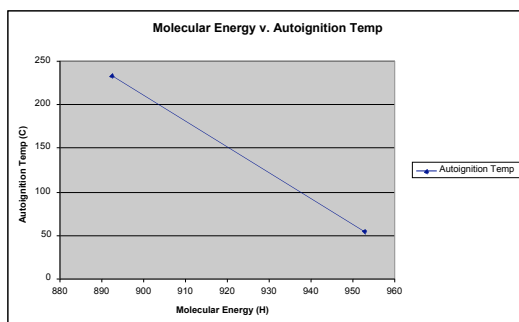


Table 1.4

Table 1.4 plots molecular energy and autoignition temperature values for RDX and nitroglycerin against each other, demonstrating a negative linear relationship. This indicates that higher-energy molecules are more unstable and spontaneously and explosively decompose as successively lower external temperatures.

In addition to plotting these values against each other, in the case of trinitrotoluene and nitroglycerin, the molecular energy values were compared to established values found on an online database. Nitroglycerin showed only an extremely minor difference of ~ 0.33 Hartrees, while trinitrotoluene showed a more substantial difference of just over 10 Hartrees. The database's trinitrotoluene energy value was calculated using a significantly larger basis set (6-31G(d)) than ours, the probable cause of the difference. The database's nitroglycerin energy value, only slightly different, was also calculated with a larger basis set but also with a less accurate level of theory (Hartree-Fock), leading the two effects to perhaps cancel each other out.

Conclusions

The results stated above are generally consistent with expectations regarding molecular energy and the stability and detonative force of explosive compounds.

The lack of a strong and direct relationship between the energy of the molecules and the force of their compound's detonation was surprising, as was the *inverse* correlation between dipole moment and velocity of detonation. Though it was initially hypothesized that more polar molecules would contribute more to detonative character, an alternative theory would be that breaking free of ionic bonding due to polarity would absorb some of the energy of

detonation, decreasing the kinetic energy of the released particles.

Autoignition temperature, though the reliability of its small data set is low, nevertheless showed a more expected relationship with both dipole moment and molecular energy, displaying an inverse correlation with each and therefore their significance as indicators of instability.

The addition of PETN, and if possible other common explosive compounds, would have added to the certainty and legitimacy of these results.

Acknowledgement

The authors extend their appreciation to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server. They also acknowledge the substantial aid of Mr. Robert Gotwals of the North Carolina School of Science and Mathematics in the production of this project.

References

- Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
- The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
- MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
- Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

Effects of Orientation on the Heat of Formation for Substituted Functional Groups on Amphetamine

Marty Goldsmith, Aniqah Shahrier, and Alexandra Fish

North Carolina School of Science and Mathematics
1219 Broad Street, Durham 27715 NC

Abstract: The amphetamine molecule was used to study the effects of the orientation of functional groups on the heat of formation because it is relatively simple and used in common medications, such as the drugs Adderall and Dexedrine. To do this, eleven functional groups were selected, all of which only had one R group. The group was then substituted into amphetamine's carbon ring in place of hydrogen in the ortho, meta-, and para orientations. Using the North Carolina High School Computational Chemistry Server, the molecules were built and geometry optimizations were run using MOPAC with the PM3 theory. Using these geometries, a coordinate scan calculation was run using MOPAC. From the results of this calculation, the heat of formation for each molecule was recorded, as a measure of the stability of the molecule. Amphetamine without any functional groups was used for purposes of comparison. Using these results, it was determined that of the orientations, meta was the most stable and ortho was the least stable orientation. There was also a correlation between the total number of atoms in the functional group added and the increase on the heat of formation.

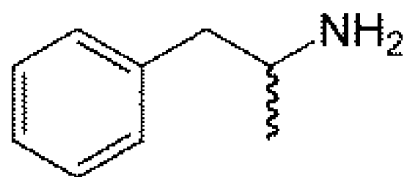
Key Words: Organic chemistry, Orientation, Functional groups, Amphetamine

Introduction

Amphetamine is a drug which affects the central nervous system, causing increased amounts of norepinephrine, serotonin, and dopamine to be released into the synaptic cleft. In the short term, it causes hyperactivity, decreased appetite, hypertension and headaches, among other things. Over time, however, amphetamine can cause disrupted sleep patterns, damage to internal organs, fatigue, and depression. It is also associated with psychological disorders such as insomnia and altered mental states. Amphetamine is highly addictive, and addiction is linked to paranoia, depression, and lethargy.

Amphetamine is legal in the United States as a prescription medication, but it is strictly regulated. It is used to treat attention-deficit disorder (ADD), attention-deficit hyperactivity disorder (ADHD), narcolepsy, and unresponsive depression. It is present in Adderall, a drug used to treat ADHD, due to its effects, such as improved impulse control.

Amphetamine is also a relatively small molecule, with a chemical formula of $C_9H_{13}N$. The following diagram illustrates its structure:



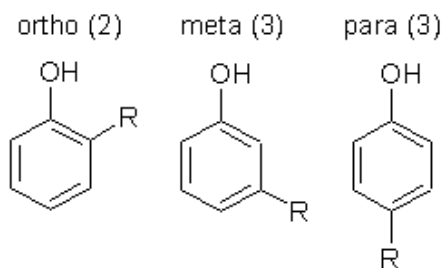
Amphetamine was chosen because of its usage in prescription medications and because of its smaller size and simplicity.

For drug companies, it is an important factor for drugs to be stable, and thereby reduce potential harm to patients. A measure of stability comes from the heat of formation, H_f , for a compound. A decreased H_f means that a compound is less likely to undergo a spontaneous reaction, since its base energy is lower. This means that the molecule is more stable. Spontaneous reactions are not desired because they may have adverse side effects on

those who are on the medication. Therefore, in this experiment, the H_f is used as the gauge for the stability of the molecule.

When chemists find a compound that has beneficial effects, they attempt to build similar compounds, called analogues. Researchers then test all the analogues to find the most effective one so they can use the compound to its full potential. In this experiment, each functional group was tested three times, the difference being the orientation of the functional group in relation to the amine structure on the carbon ring. Only functional groups with one R group were selected. One aspect of this experiment is to investigate the stability of different functional groups, and the reasons behind their respective stabilities.

In organic chemistry, there are three main orientations for functional groups on amphetamine's carbon ring. If the functional group is next to the R group, the orientation is termed ortho. If there is a single carbon in between the functional group and the R group, the molecule has meta orientation. Lastly, if the functional group has two carbons separating it from the R group, the molecule has para symmetry. A visual representation of these orientations is seen below, with an alcohol group used in place of a theoretical functional group.



The different orientations do have an effect on the properties of the molecule, and should therefore be investigated. The second goal of this experiment was to determine which, if any, of the orientations are more stable.

Computational Methods

In this experiment the North Carolina High School Computational Chemistry Server was used to run all calculations. MOPAC computational software was used to do PM3 geometry optimization calculations on amphetamine with different functional groups in different orientations. The key word GEO-OK

was necessary to keep these jobs from failing. After running each molecule, we used the optimized geometry to run coordinate scan calculations using MOPAC PM3 once again, with the same aforementioned keyword required. We then recorded the heat of formation of each molecule and looked for a correlation to this and different properties of the molecules.

Results and Discussion

To determine the stability of the orientations, the stability of each orientation per functional group was ranked according to the value of its H_f . As a low value for H_f indicates stability, the orientation with the lowest H_f was ranked as best, and the highest as worst. We then tallied the number of times each orientation was best, worst, or medium for each functional group. This resulted in the following results:

Orientation of functional group	Lowest H_f	Middle H_f	Highest H_f
Meta	5	5	0
Para	5	3	2
Ortho	0	2	8

This data shows that the meta position had the best stability overall and that ortho had the works overall stability, leaving the para orientation in the middle.

Another factor which correlated to the heat of formation was the total number of atoms in the molecule. As the number of atoms increased, the heat of formation increased as well. Also, as the number of pi bonds within the functional group increased, the heat of formation increased. However, these conclusions might be misleading because all of the functional groups with a large number of total atoms also had the highest number of pi bonds and were ring groups. Because these groups are structurally similar, other characteristics might be the deciding factor on their heat of formation.

The number of nitrogen atoms in the molecule might have had an effect as well. The molecules with the 7th and 10th highest heats of formation both had two nitrogen atoms, while all the other molecules had only one. However, more functional groups with multiple nitrogen atoms would be required to confirm this trend.

The following table shows all the functional groups and their heat of formation in the different orientations. They were compared to amphetamine (C₉H₁₃N) itself, which has a heat of formation of 12.72 kcal/mol.

Group Chemical Formula	Ortho	Meta	Para
Alcohol C ₉ H ₁₃ ON	-31.90318 kcal/mol	-33.99745 kcal/mol	-32.38345 kcal/mol
Benzene C ₁₅ H ₁₇ N	37.78881 kcal/mol	35.24875 kcal/mol	35.24875 kcal/mol
Toluene C ₁₆ H ₁₉ N	30.32908 kcal/mol	30.02471 kcal/mol	31.62197 kcal/mol
Pyridine C ₁₄ H ₁₆ N ₂	43.91472 kcal/mol	40.38100 kcal/mol	40.21254 kcal/mol
Amine C ₉ H ₁₅ N	8.79534 kcal/mol	7.41391 kcal/mol	7.47016 kcal/mol
Aldehyde C ₁₀ H ₁₃ ON	-23.04262 kcal/mol	-24.38050 kcal/mol	24.59987 kcal/mol
Carboxyl C ₁₀ H ₁₃ O ₂ N	-76.22312 kcal/mol	-78.69984 kcal/mol	-80.15367 kcal/mol
Nitrite C ₉ H ₁₂ O ₂ N ₂	13.91701 kcal/mol	11.58380 kcal/mol	11.52858 kcal/mol
Methyl C ₁₀ H ₁₄ N	0.44531 kcal/mol	-0.10587 kcal/mol	0.7124 kcal/mol
Ethyl C ₁₁ H ₁₆ N	-2.70395 kcal/mol	-3.9907 kcal/mol	-3.27619 kcal/mol

The phosphate group was also run; however, its geometry optimization failed. When looking at the raw output of the run, the message "YOU MUST SPECIFY "NOMM" OR "MMOK" REGARDING MOLECULAR MECHANICS CORRECTION" was seen. Since the meaning of this message was unknown and there was already a significant amount of data from the other ten functional groups, it was decided that phosphate was not required for this experiment.

The heat of formation for amphetamine was higher than the heat of formation for several of its derivatives that had functional groups attached. This was surprising because it had been assumed that amphetamine would be one of the more stable compounds in this experiment. Enough scientific background was not available

to interpret this result, and further research would be necessary to fully analyze this trend.

Conclusion

The results of our data and data analysis suggest that orientation of the functional groups has a correlation with the heat of formation, and therefore the stability of the molecule. Meta is the most stable orientation while ortho is the least stable orientation. The size of the functional group also has a correlation. As the functional group increases in size, so does the heat of formation of the molecule, which means the molecule becomes less stable. The stability also decreases with an increase in the number of carbon atoms in the molecule, but that can most likely be attributed to the correlation with size. The number of nitrogen atoms also seems to have a correlation, but further research would have to be conducted to confirm this. The number of oxygen atoms did not seem to have a correlation. Adding a carboxyl group to amphetamine made it the most stable, while adding a pyridine made it the least stable. Amphetamine was surprisingly unstable compared its analogues, ranking 8th overall in stability out of 11 molecules. Finding the impact of this discovery would require further research.

Acknowledgement

The authors thank Mr. Gotwals for his assistance with this paper. Appreciation is also extended to the Burroughs Welcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

- Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed May 2007).
- The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed May 2007).
- MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
- BA Clement, EAB Goff, CM Goff, TDA Forbes *Phytochemistry*, **46**, 249-254 (1997)
- BA Clement, CM Goff, TDA Forbes *Phytochemistry*, **49**, 1377-1380 (1998)

6. LN Kuczenski *Pharmacol Biochem Behav*, **15**, 399-404 (1981)
7. I Chaudary, S Turkanis, R Karler *Neuropharmacology*, **27**, 777-781 (1988)
8. Nagle, Hinter, and Barbara Nagle. Pharmacology: An Introduction. 5th ed. Boston: McGraw Hill, 2005.
9. Newton, David, Alasdair Thopre, and Chris Otter. Revise A2 Chemistry. 2nd ed. London: Heinemann Educational Publishers, 2004.

Determination of the Correlation between Several Properties of Legal and Illegal Drugs on the North Carolina High School Computational Chemistry Server

J. Richardson, J. Barnes, D. Tingen

North Carolina School of Science and Mathematics, Durham, NC

Received 5 June, 2007

Abstract: In this experiment we wanted to determine if there was a correlation between dipole moment and number of atoms and heat of formation and the number of carbon rings of four different legal and illegal drugs. We did this for Hydrocodone, Amphetamine, Cocaine, and Morphine. To do this we used the North Carolina High School Computational Chemistry Server to determine the heat of formation of those drugs as well as dipole moment of each molecule. We found that there was a correlation between dipole moment and number of atoms, and there was a correlation between number of carbon rings and heat of formation.

Key words: carbon rings, dipole moment, heat of formation, correlation, morphine, cocaine, amphetamine, hydrocodone

Introduction

Many legal and illegal drugs have the same properties and effects on the body just in different amounts. Cocaine is a very powerfully addictive stimulant drug that can be used in many forms. The most common form, powdered, hydrochloride salt, can be snorted or dissolved in water and injected. "Crack" cocaine is cocaine that has not been made in hydrochloride salt by being dissolved in an acid. It comes in a rock crystal and when heated, its vapors can be smoked. Cocaine is a strong central nervous system stimulant and interferes with the process of dopamine, a chemical messenger associated with pleasure and movement. Cocaine use causes a buildup of dopamine which in turn causes a continuous stimulation of receiving neurons and thus creates the euphoria reported by cocaine users. There are many physical effects of cocaine use including constricted blood vessels, dilated pupils, and increased temperature, heart rate, and blood pressure. The length of the effects depends on the method of administration. The high is more intense the faster the absorption. Increased use can reduce the period of time the user feels a high and increase the risk of addiction. Some complication that can result from cocaine use can include disturbances in heart rhythm, heart attacks, chest pain and respiratory failure, strokes, seizures, headaches, abdominal pain, nausea, loss of appetite and malnourishment. The chemical formula of cocaine is $C_{17}H_{21}NO_4$ and the scientific name is methyl.

Morphine is used to relieve moderate to severe pain. It is only used but patients who are expected to have constant pain for more than a few days. Morphine is a member of the class of drugs known as opiate (narcotic) analgesics.

These types of medicines work by changing the way the body senses pain. Morphine comes as a tablet or a liquid solution. Morphine should be taken strictly according to the prescription directions that accompany the drug. Some doctors like to start a patient on a low dose of morphine and increase the dose as necessary to control their pain to try to decrease the risk of dependency. Morphine can be very habit-forming if not taken exactly as directed. If a patient takes a larger dose, takes it more often, or in a different way than that described by their doctor then they can become addicted. Morphine can cause dehydration so when taking the medication a patient should drink lots of fluids. Morphine can cause many side effects including dizziness, lightheadedness, drowsiness, upset stomach, vomiting, constipation, diarrhea, loss of appetite, weight loss, changes in ability to taste food, dry mouth, sweating, weakness, headache, agitation, nervousness, mood changes, confusion, difficulty sleeping, stiff muscles, double vision, red eyes, uncontrollable eye movements, chills, flu like symptoms, and difficulty urinating. There are some very serious side effects of morphine including irregular breathing, blue or purple skin, irregular heartbeat, seizures, fainting, rash, tightness in the throat, and swelling of the extremities. Morphine is also an opiate analgesic drug like cocaine. The chemical formula of Morphine is $C_{34}H_{40}N_2O_{10}S$.

Hydrocodone is prescribed as a pain reliever, sometimes in conjunction with other drugs, to relieve moderate to moderately severe pain. It can be used as a tablet, capsule, or liquid. It is generally taken every four to six hours to control pain. Hydrocodone can be addicting and habit-forming so a patient should never take a larger dose, take a dose more often, or for a

longer period of time than it is prescribed. Hydrocodone can cause lightheadedness, dizziness, drowsiness, upset stomach, vomiting, constipation, stomach pain, rash, difficulty urinating, difficulty breathing, and mood changes. The chemical formula for hydrocodone is $C_{18}H_{21}NO_3$.

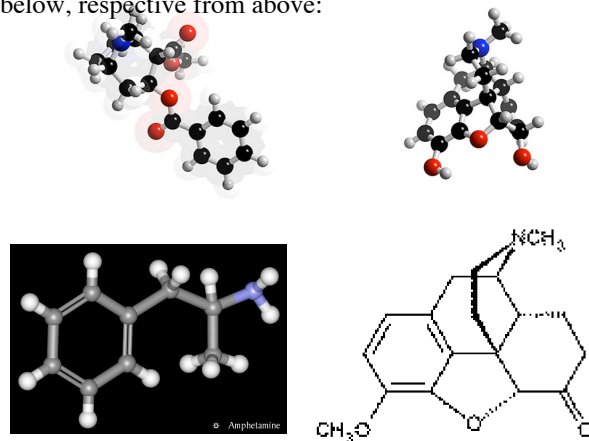
Amphetamine is used as a central nervous system stimulant in the treatment of many conditions such as attention deficit hyperactivity disorder, depression, and narcolepsy but it is also used illegally as a stimulant. Amphetamine is a colorless, volatile liquid that is used for many things both legal and illegal. It is used as an effective nasal decongestant, an analeptic, and an anorexic. It is also used to treat depression because it caused an enhanced sense of well-being and euphoria. Amphetamine cause side effects such as dry mouth, heart palpitations, hypertension, stomach cramps, decreases urinary frequency, dizziness, dysphoria, headache, tremors, restlessness, insomnia, decreases appetite, increased aggressiveness, anxiety, and paranoid panic states. Extreme amounts of amphetamine can cause convulsions, cerebral hemorrhaging, coma or death. The euphoric side effects have made amphetamine a widely abused drug. Tolerance occurs after taking the drug for long periods of time so abusers tend to increase the amount that they consume increases their risks of serious side effects and long term dependency. Amphetamine has the chemical formula $C_9H_{13}N$.

For each of these molecules we wanted to determine the dipole moment and heat of formation using The North Carolina High School Computational Chemistry Server¹. We want to determine if heat of formation is influenced by the number of carbon rings. Heat of formation is the enthalpy change during the formation of a pure substance from its elements, at constant pressure. We also wanted to determine if there was a correlation between dipole moment and number of atoms in the molecule. Dipole moment of a molecule is the product of either charge in an electric dipole with the distance separating them.

Computational Approach

Before performing any calculations for the project we had to determine what illegal and legal drugs would be used. Once determining that cocaine, morphine, amphetamine, and hydrocodone would be used the calculations started. Our calculations were performed on the

WebMO² North Carolina High School Computational Chemistry Server. We used this server to build the molecules and then ran a geometry optimization using MOPAC PM3. After the molecules were done running through the server they were ran again but this time under a molecular energy calculation using MOPAC. The configuration of each molecule can be seen below, respective from above:



Drug	Compound	#of atoms	# of carbon rings	Heat of formation	Dipole moment
Cocaine	C17H22NO4	44	2	-50.50265	3.566
Morphine	C17H19NO3	40	3	-76.60793	2.03
Hydrocodone	C18H21NO3	43	2	-43.47238	4.185
Amphetamine	C9H13N	23	1	14.75653	1.677

Each drug and its heat of formation, dipole moment, number of carbon rings and number of atoms were then recorded.

Our goal was to find a correlation between the number of carbon rings formed and the heat of formation and between the number of atoms and the dipole moment. The structures of each molecule were also viewed as a factor in the correlation. In order to perform the correlations we used Graphical Analysis software to graph and find the correlations stated above.

Results and Discussion

From our data, we were able to compare different properties of the four molecules. We first wanted to compare the dipole moment to the number of atoms found in each molecule. We found the number of atoms in each compound and graphed them against the dipole moment calculated in each run. After looking at our graphs, we then fit a linear line to the data set. For atom number vs. dipole moment, we got an

equation of $y = 0.0993x - 0.633$, where y is the dipole moment and x is the number of atoms. Our data is graphed below in Figure 1.1:

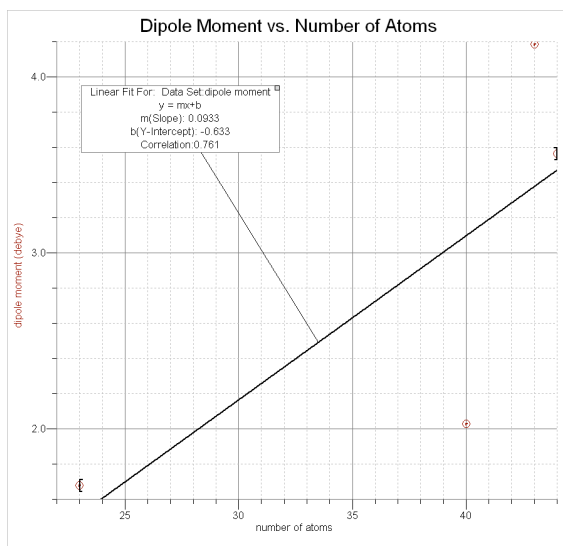


Figure 1.1

For heat of formation and carbon rings, we used the same method of comparison. We found an equation of $y = -34.8x + 39.2$, where y represents heat of formation and x represents the number of carbon rings. Our data for the data is as follows in Figure 1.2:

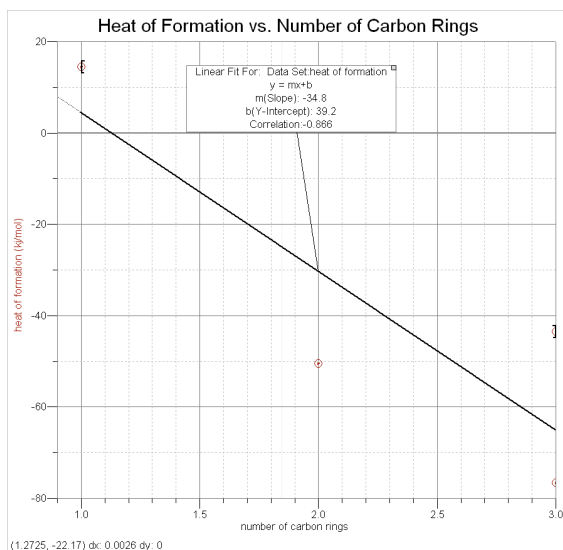


Figure 1.2

Conclusions

During this project, we discovered many things about our four molecules. Between number of atoms and dipole moment, there was a higher dipole moment for larger molecules,

meaning there is a positive linear relationship between the two. A larger amount of carbon atoms in a molecule usually require more hydrogen atoms and will therefore be more polar.

With the number of carbon rings and heat of formation, the opposite was true. The higher the number of carbon rings, the lower the heat of formation. We think this might be because as the number of carbon rings increase, the number of total atoms increases. From basic chemistry knowledge, we know that a molecule with more atoms will have a harder time holding the molecules together, so they will break apart with less energy.

Acknowledgement

The authors thank Mr. Robert Gotwals, the best teacher in the world, of the North Carolina School of Science and Mathematics for his assistance with this research, and his amazing knowledge in the field of computational chemistry. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

27. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
28. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
29. We tried to find the reference for MOPAC but we didn't succeed because we are incompetent.
30. <http://chemistry.about.com/od/dictionari esglossaries/g/bldehform.htm>
31. www.chemindustry.com/chemicals

Heat of Formation: Traditional Energy Sources v. Alternative Energy Sources

M. Call and L. Evans

North Carolina School of Science and Mathematics, Durham, NC

Received 5 June, 2007; Accepted 5 June, 2007

Published online on Comp Chem Moodle (moodle.ncssm.edu)

Abstract: Our purpose was to investigate the relationship between bond length, number of bonds broken, and heat of formation in eight different molecules, four of which were traditional energy sources (gasoline, propane, diesel, methanol) and four of which were alternative energy sources (liquid nitrogen, ethanol, butanol, hydrogen gas). All of our calculations were done using the NC High School Computational Chemistry server (chemistry.ncssm.edu), which supports MOPAC. We built the eight molecules and ran a geometry optimization on all of them using MOPAC at the PM3 level. Using the values we calculated during the geometry optimization, we created a spreadsheet of the heat of formation, number of bonds broken, and average bond length, then created a series of linear regression graphs using the independent variables (number of bonds broken and average bond length) as the x values and using the heat of formation as the y value. The factor we observed was the calculated R square value, which indicates a good correlation if it is close to 100%, and a poor correlation if it is close to 0%. Based on the data and data analysis, the results suggest that the regressions for the number of bonds broken graphed against the heat of formation were much better fits for the data than those of the average bond lengths broken graphed against the heat of formation, and it is therefore possible to assume that while there is a strong correlation between the number of bonds broken and the heat of formation of a molecule the average length of all of the bonds broken has little to do with the heat of formation of a molecule. Bond lengths individually, however, may have much to do with the heat of formation, although this study did not investigate that topic. Any errors in our calculations were caused by the inherent flaw of molecular modeling, that all models are incorrect, as well as the possible human error of building the models incorrectly.

Key words: heat of formation, geometry optimization, bonds broken, MOPAC, PM3

Introduction

Throughout the world many different fuel sources are used for energy. Two major categories of fuels exist at this time. There are the traditional fuel sources, which are non-renewable resources and supposedly contribute to the depletion of the ozone layer, and then there are the alternative fuels, which are less popular but are usually renewable.

In this experiment, four fuel sources from each group of traditional and alternative were chosen. The traditional fuel sources are gasoline,

propane, diesel, and methanol. The alternative fuel sources are liquid nitrogen, ethanol, butanol, and hydrogen gas.

Most of the traditional fuel sources chosen are made from petroleum. Gasoline comes from a petroleum mixture containing hydrocarbons and benzene or iso-octane and is mainly used as fuel in internal combustion engines. Propane comes from petroleum products and is a three-carbon alkane. Propane is commonly seen being used for fuel around the home and for engines. Diesel is a fractional distillate from fuel gas that is mostly petroleum. Diesel is used as fuel for the

diesel engine, created by a man by the same name. Methanol is now made synthetically but once was made from the distillation of wood. Methanol is used as fuel for internal combustion engines.

Liquid nitrogen is made from the fractional distillation of liquid air. It could be used to fuel cars. Ethanol is commonly made from petrochemical feedstocks. It is used for fuel for automobiles and has been used as fuel for rockets. Butanol is created from fossil fuels and can be used in cars and produces more energy than ethanol. Hydrogen gas is found naturally and is a potential alternative as fuel for automobiles.

In this experiment bond length, number of bonds broken, and heat of formation were investigated. According to previous experiments, bond length and the heat of formation are directly correlated - the shorter the bond length, the higher the heat of formation. The number of bonds broken has also related to the heat of formation in the past. As the number of bonds between atoms increase, the length of the bonds decreases and the strength of the bonds increase. Therefore, the heat of formation increases with an increase of bonds broken. This experiment investigated whether past experiments were correct in their results.

Computational Approach

This experiment used the North Carolina Computational Chemistry Server to run all calculations through the WebMO molecular editor. The molecules that make up eight different traditional and alternative energy sources were built. Once each was built, it was optimized using the "comprehensive cleanup" molecular mechanics package found in the WebMO molecular editor. The rough optimization was used to fix the geometry of the chains as well as to add hydrogen atoms in the appropriate places. Finally, using the software package MOPAC, geometry optimizations were calculated using the PM3 basis set.

Results and Discussion

The heat of formation, number of bonds broken and average bond length for each molecule was found in the raw output of the calculation and recorded.

After recording these data the molecules were graphed against each other in Microsoft Excel, with the number of bonds broken as the x variable and the heat of formation being the y variable. The goal of creating this graph was to calculate a regression line that could be used to extrapolate and determine the amount of energy released when a certain number of bonds are broken in a molecule.

The regression for the graph of the data of all eight of the molecules was determined to be:

.....

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the number of bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.600138, which means that the line's fit is ~60.0% accurate. The linear regression graphed onto the data plot can be seen in Figure 1 below.

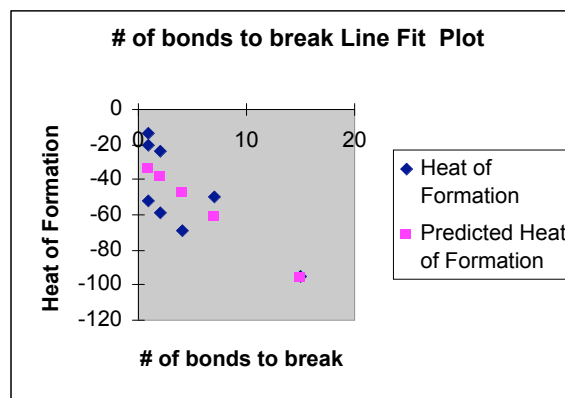


Figure 1

Next, the four molecules of the traditional energy sources (gasoline, diesel, propane, and methanol) were graphed against each other in Microsoft Excel, with the number of bonds broken as the x variable and the heat of formation being the y variable. The goal of creating this graph was to calculate a regression line that could be used to extrapolate and determine the amount of energy released when a certain number of bonds are broken in a traditional energy sources molecule.

The regression for the graph of the data of the traditional energy source molecules was determined to be:

.....

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the number of bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.767365, which means that the line's fit is fairly accurate. The linear regression graphed onto the data plot can be seen in Figure 2 below.

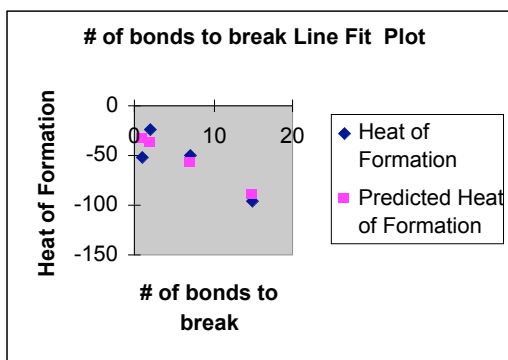


Figure 2

We observed that the out of our four traditional energy sources, gasoline, diesel and propane were all hydrocarbons, and methanol was not. We therefore created a subsequent graph of the three hydrocarbons, excluding methanol which has an OH group. The x variable is the number of bonds broken and the y variable is the heat of formation. The goal of creating this graph was to calculate a regression line that could be used to extrapolate and determine the amount of energy released when a certain number of bonds are broken in a traditional hydrocarbon energy sources molecule

The regression for the graph of the data of the traditional energy source molecules was determined to be:

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the number of bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.999636, which means that the line's fit is extremely accurate. The linear regression graphed onto the data plot can be seen in Figure 3 below.

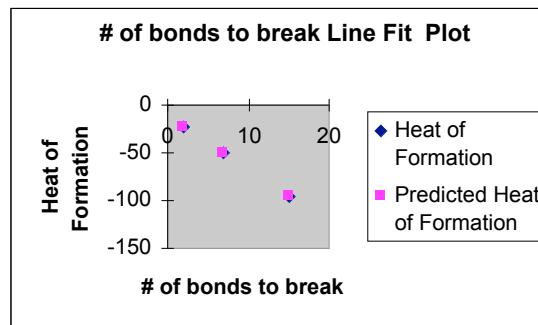


Figure 3

Just as the molecules of the traditional energy sources were graphed against each other in Microsoft Excel, so were the molecules of the alternative energy sources (liquid nitrogen, ethanol, butanol, and hydrogen gas). The number of bonds broken is the x variable and the heat of formation is the y variable. The goal of creating this graph was to calculate a regression line that could be used to extrapolate and determine the amount of energy released when a certain number of bonds are broken in an alternative energy sources molecule.

The regression for the graph of the data of the alternative energy source molecules was determined to be:

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the number of bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.797189, which means that the line's fit is fairly accurate. The linear regression graphed onto the data plot can be seen in Figure 4 below.

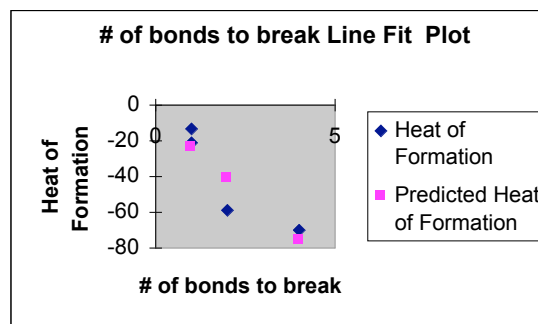


Figure 4

As the linear fits for these four data sets are all fairly decent, the linear regression equations found for each data set can be used together to extrapolate and estimate how much energy will be released in a molecule when a given number of bonds is broken. These values are only estimates, however, because the linear fits for the graphs are not perfect, although all are at least 60% accurate according to the R square value.

Once the number of bonds broken and heat of formation were graphed against each other in Microsoft Excel we graphed the average length of the bonds broken and the heat of formation for each molecule against each other in Microsoft Excel. The goal of creating these graphs was to calculate regression lines that could be used to extrapolate and determine the amount of energy released when a certain number of bonds are broken in a molecule.

The regression for the graph of the data of all eight of the molecules was determined to be:

$$y = -1.25x - 50.00$$

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the average length of the bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.286292, which means that the line's fit is fairly poor. The linear regression graphed onto the data plot can be seen in Figure 5 below.

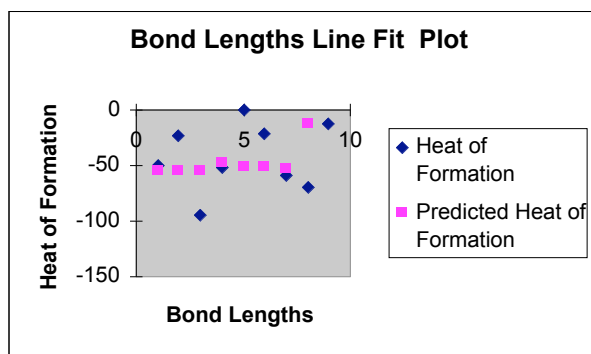


Figure 5

The data set was then divided into traditional energy sources (gasoline, propane, diesel and methanol) and alternative energy sources (liquid nitrogen, hydrogen gas, ethanol, butanol). Using the traditional energy source data, we graphed the average length of the bonds being broken as

the x value against the heat of formation as the y value.

The regression for the graph of the data of the molecules of the traditional energy sources was determined to be:

$$y = -0.015293x - 50.00$$

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the average length of the bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.015293, which means that the line's fit is extremely bad. There is no real correlation in these data. The linear regression graphed onto the data plot can be seen in Figure 6 below.

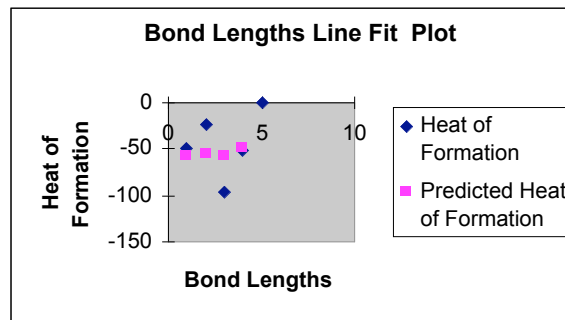


Figure 6

The alternative energy sources were graphed against each other in the same way, with the average length of the bonds broken as the x value and the heat of formation as the y value.

The regression for the graph of the data of the molecules of the alternative energy sources was determined to be:

$$y = -0.477236x - 50.00$$

where y is the dependent variable, or the heat of formation, and x is the independent variable, or the average length of the bonds broken. The R Square value, which shows how well a linear regression line fits a graph, was found to be 0.477236, which means that the line's fit is moderately poor. The linear regression graphed onto the data plot can be seen in Figure 7 below.

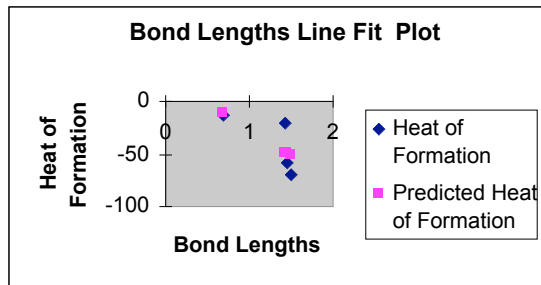


Figure 7

As the linear fits for these 3 data sets are all moderately to extremely poor, it would be unwise to use any of the R Square values to interpolate or extrapolate on the data.

Conclusions

Based on the data and data analysis, the results suggest that the regressions for the number of bonds broken graphed against the heat of formation were much better fits for the data than those of the average bond lengths broken graphed against the heat of formation, and it is therefore possible to assume that while there is a strong correlation between the number of bonds broken and the heat of formation of a molecule the average length of all of the bonds broken has little to do with the heat of formation of a molecule. Bond lengths individually, however, may have much to do with the heat of formation, although this study did not investigate that topic.

It is also interesting to note that although these data is not discussed in this paper, the geometry optimization calculation revealed that the amount of energy per gram of the alternative energy sources was on a whole much larger than the amount of energy per gram of the traditional energy sources. Further studies should be conducted to increase our knowledge of these topics.

Acknowledgements

The authors thank Mr. Robert Gotwals of the NCSSM Chemistry Dept. Durham, NC for his teachings and guidance. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

32. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
33. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
34. MOPAC, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
35. "Diesel." *Wikipedia*. 04 June 2007. 04 June 2007 <<http://en.wikipedia.org/wiki/Diesel>>.
36. "Methanol." *Wikipedia*. 04 June 2007. 2 June 2007 <<http://en.wikipedia.org/wiki/Methanol>>.
37. "Liquid Nitrogen." *Wikipedia*. 04 June 2007. 04 June 2007 <http://en.wikipedia.org/wiki/liquid_nitrogen>.
38. "Ethanol." *Wikipedia*. 04 June 2007. 04 June 2007 <<http://en.wikipedia.org/wiki/ethanol>>.
39. "Butanol." *Wikipedia*. 04 June 2007. 04 June 2007 <<http://en.wikipedia.org/wiki/butanol>>.
40. "Hydrogen Fuel." *Wikipedia*. 19 May 2007. 04 June 2007 <http://en.wikipedia.org/wiki/hydrogen_fuel>.
41. "Cool Stuff." *Riverdeep*. 17 September 2001. 04 June 2007 <http://www.riverdeep.net/current/2001/09/091701_liquidn.jhtml>.

Reactive Activities of Nitrogen-Containing Pollutants in the Environment

T. Hicks, D. Britt, & C. Nance

North Carolina School of Science and Mathematics, Durham, NC

June 5, 2007

Published online on Comp Chem Moodle (moodle.ncssm.edu)

Abstract: We often use computational chemistry methods to solve real world problems. One of the many problems we face today is global warming and air pollution. We decided to look at different nitrogen-containing compounds that are known to have harmful effects on the environment and determine why some are more harmful than others, specifically nitrogen dioxide (NO₂). Using the molecular editor builder of WebMO on the North Carolina High School Computational Chemistry Server allowed us to have a greater understanding of why NO₂ is more harmful and reactive in the environment. Using Gaussian software package at the Hartree-Fock level with a 3-21G basis set we were able to retrieve useful data relating to the EHOMO and ELUMO correlation between the several nitrogen molecules being tested. We began researching each compound by running geometry optimization calculations followed by molecular orbital calculations. We used the molecular orbital calculations to retrieve the HOMO-LUMO gap and to compare the different molecules, which allowed us to draw reasonable conclusions. In the end it became clear that NO₂ was an outlier when compared to the other nitrogen molecules. This difference could be a major factor in how it reacts in the environment when compared to the others.

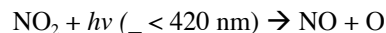
Key words: Nitrogen Dioxide, Gaussian, Basis Set, EHOMO, ELUMO, HOMO-LUMO Gap

Introduction

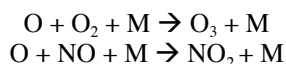
As each day passes by, pollution is accumulated in every aspect of our Earth. From land to water and even in the air, pollution continues to be one of the most researched topics in the modern academic world. Pollution hasn't always been a major problem for our world but as many resources begin to become exhausted and limited due to population increase, pollution rates have skyrocketed. In this research paper we will specifically focus on air pollution and furthermore target the activities of nitrogen containing gaseous compounds in the atmosphere.

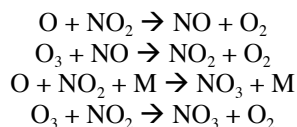
Currently Earth's atmosphere is composed of 78.1% nitrogen (N as N₂), 20.9% oxygen (O as O₂), 0.03% carbon dioxide (CO₂), and a number of trace gases. This clearly shows a significant portion of the Earth's atmosphere is made up of nitrogen which means that nitrogen has an abundant realm activities present in the atmosphere. When natural atmospheric chemicals mix with anthropogenic emissions mainly from fossil fuel burning in solar radiation, the result is photochemical smog. These reactions are aided by the help of the sun which gives it the photochemical part in the proper name. The chemical mixtures produced by reactions such as these lead to the generation

yellow brown, reddish, and gray hazes in the sky that can be very detrimental to living organisms and ecosystems. A few contributors of smog can be carbon monoxide (CO), sulfur and nitrogen oxides, toxic hydrocarbons such as benzene and toluene, lead, and particulates. These harmful chemicals are most abundant in cities where there are a lot of vehicles used for transportation which are important sources of carbon monoxide, hydrocarbons, NO_x's, SO_x's, and other particulates found in urban areas. The primary photochemical reaction producing oxygen atoms is:



Ozone is also an important product of the reactions involved in the generation of smog. Ozone is beneficial in the upper atmosphere (Stratosphere) because it helps protect the Earth's surface from harmful UV radiation. However, when ozone is present in the lower atmosphere it can be deleterious to the health of plants and animals. Reactions involving oxygen species (M is an energy-absorbing third body) and atomic oxygen leads to several reactions involving oxygen and nitrogen oxide species:





Many different compounds of nitrogen are present in the atmosphere but NO_2 is the one chemical which has the most negative effect on the Earth's atmosphere. Varying from NO (nitrous oxide) to N_2O_5 , NO_2 when reacted with a certain wavelength of sunlight begins a chain of reactions which all contribute to photochemical smog and ozone buildup. We believe that this may be due to the EHOMO-ELUMO gap of NO_2 compared to the other nitrogen containing compounds.

In each compound, there are a specific number of occupied orbitals and of unoccupied orbitals, according to how many atoms are in each compound, their charges, etc. Within these orbitals there are numbers that correspond with each occupied and unoccupied orbital: the energy of the highest occupied orbital is the EHOMO, as explained in its title, (Energy of the Highest Occupied Molecular Orbital) and the energy of the lowest unoccupied orbital is the ELUMO (Energy of the Lowest Unoccupied Molecular Orbital). The difference between these two values is the EHOMO-ELUMO gap which determines the speed of the reaction and if the reaction occurs at all. The value of the EHOMO-ELUMO gap in each compound determines its compatibility with other compounds. This is why we decided to investigate the EHOMO-ELUMO gap of each nitrogen-containing compound in our research.

This research project attempts to analyze and explain why NO_2 reacts as it does and why other nitrogen containing compounds act differently. Our goal in performing this research project is to find some relationship between the EHOMO-ELUMO gap of each different compound and the compound itself (# of atoms, # of N atoms, # of orbitals, etc.)

Computational Approach

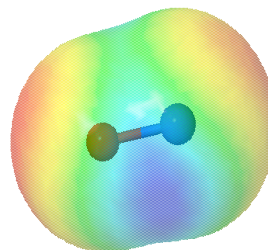
In calculating the data needed to perform our experiment, we relied on several computational methods including computational software, spreadsheet applications, and computers. Using the molecular editor builder of WebMO on the North Carolina High School Computational Chemistry Server we were able to run several

different calculations on the different nitrogen-containing compounds. The particular computational software used for the calculations in our experiment was Gaussian. We used the Gaussian software at the Hartree-Fock level of theory and 3-21G basis set in order to run Geometry Optimization and Molecular Orbital calculations. The NC HS Comp. Chem. Server allowed us to run numerous molecular models in a reasonable amount of time and also generate reasonably accurate results. We were running these numerous calculations in order to compare the E HOMO and E LUMO values for several nitrogen-containing compounds that have known harmful effects in the environment. Throughout our computational procedures we discovered that using the Gaussian software to run geometry optimization calculations with HF 3-21G took a significantly larger amount of time than running molecular orbital calculations with the same software. These run times range from one minute and twenty one seconds to four minutes and thirty seven seconds. This provides a reference for future use of this software package.

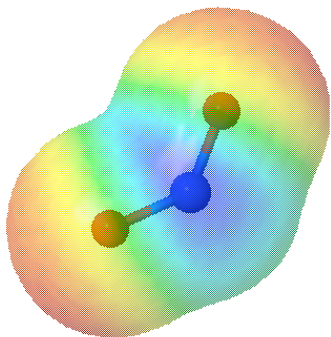
Results and Discussion

Before analyzing any data, we observed the electrostatic potential of each compound. These results displayed to us the positive and negative areas of the compound, which are possible reaction sites for each compound. Using these electrostatic potentials we were able to see where each compound could possibly react with other compounds such as O_2 . Below are the electrostatic potentials for each of the 4 nitrogen-containing compounds.

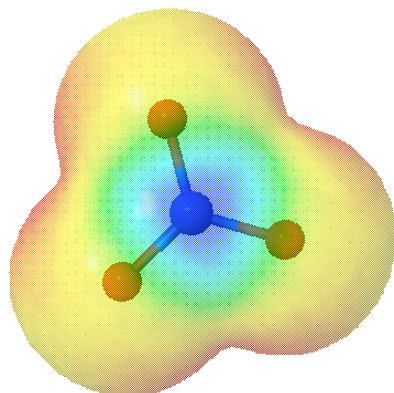
Nitrogen Oxide (NO) Electrostatic Potential



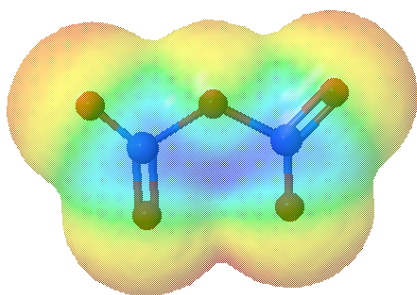
Nitrogen Dioxide (NO₂) Electrostatic Potential



Nitrate (NO₃⁻¹) Electrostatic Potential



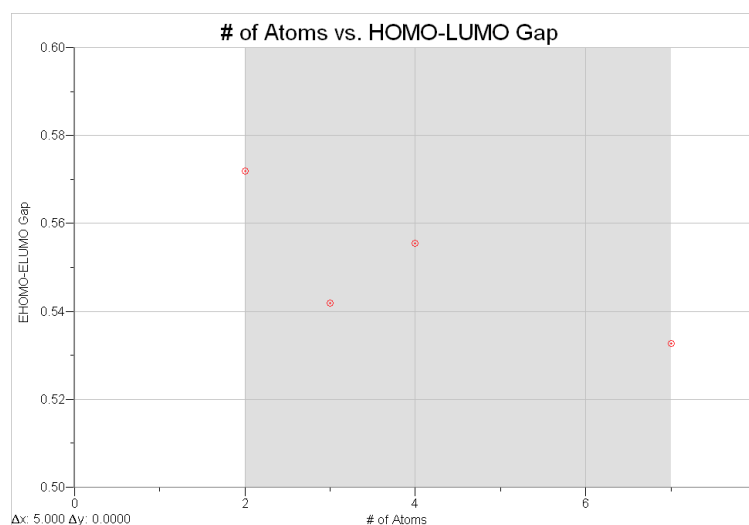
Dinitrogen Pentoxide (N₂O₅) Electrostatic Potential



Our results were organized using Microsoft Excel Spreadsheet in which we recorded each compound and their respective CPU times (secs), EHOMO (Hartrees), ELUMO (Hartrees) and HOMO-LUMO gap (Hartrees), as shown below:

Molecules	CPU Time (secs)	EHOMO (Hartrees)	ELUMO (Hartrees)	HOMO-LUMO Gap
NO	10.5	-0.44079	0.13111	0.5719
NO ₂	10.5	-0.4892	0.05264	0.54184
NO ₃ ⁻¹	10.6	-0.19062	0.36486	0.55548
N ₂ O ₅	10.7	-0.5077	0.0249	0.5326

We also used Logger Pro Version 3.4.1 to graph the number of atoms in each compound vs. the EHOMO-ELUMO gap of each compound. This data resulted in a scatter plot displayed below:



As shown in the graph above, all except one of the nitrogen-containing compounds follow a linear regression line. This one outlier is NO₂ which, as we researched, is the most compound that has the most negative effects on the environment. Nitrogen dioxide's outlier presence in this graph gives us some explanation why NO₂ behaves so much differently than the other nitrogen-containing gaseous compounds in the atmosphere.

Conclusions

Based on the data and the data analysis of nitrogen-containing compounds, the values of the EHOMO-ELUMO gap in each of the compounds follow a linear regression line path with NO₂ being an outlier in the data. The relationship of the EHOMO-ELUMO gap of NO₂ in comparison to the other nitrogen containing compounds (NO, NO₃⁻¹, and N₂O₅) can further be evaluated to determine what

components of the NO₂ compound make it different. Understanding its reactivity and interactions with other compounds in the atmosphere can assist in determining the severity of pollution and its effects on life on Earth. Finding this outlying value of the EHOMO-ELUMO gap of NO₂, shows that there is a difference in the speed of the reaction and the compatibility of NO₂ with other compounds in the Earth's atmosphere.

Acknowledgement

The authors thank Mr. Robert Gotwals of the North Carolina School of Science and Mathematics for assistance with this work. The authors also thank Dr. Marion Brisk for her assistance and Steven Lin, who served as the teacher assistant for Mr. Robert Gotwals' Intro to Computational Chemistry course, for his assistance as well. Appreciation is also extended to the Burroughs Welcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

42. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
43. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
44. MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
45. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
46. Manahan, Stanley E. *Environmental Chemistry*. 5th. Chelsea, Michigan: Lewis Publishers, Inc., 1991.
47. Mackenzie, Fred T. *Our Changing Planet: An Introduction to Earth System Science and Global Environmental Change*. 3rd. Upper Saddle River, NJ: Pearson Education, Inc., 2003.

Comparison of Morphine and Morphine Analog Structures in Relation to Responses of the Central Nervous System

A. Michelson and E. Willis

North Carolina School of Science and Mathematics, Durham, NC

June 5, 2007

Abstract:

Pain in the body can be suppressed by various painkillers such as opioids. Morphine and morphine analogue structures are opioid agonists that bind to the same receptor site. Also, Naltrexone is an opioid antagonist that reverses the effects morphine and morphine analogues on the body; however, it binds to the same receptor site. The structures of these molecules vary somewhat, but this variance affects the potency of each molecule greatly. We hypothesized that the electrostatic potential would show further explanation of the potency differences among each of the molecules. After running a "Molecular Orbital" calculation using MOPAC AM1 on each of the molecules we found that our data shows no correlation between electrostatic potential and potency. Proving our hypothesis incorrect, we did a further analysis and found that the LUMO of the molecules had a correlation with the potency. Also confirmed was the length of the protruding carbon chain on each molecule versus potency was accurate.

Keywords: computational chemistry, medicinal chemistry, morphine, codeine, etorphine, naltrexone, heroin, electrostatic potential, molecular structures, molecular geometry

Introduction

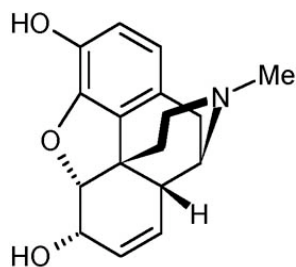
Pain is an uncomfortable sensation that occurs when sensory fibers are exposed to extremes. There are two types of sensory fibers that signal pain: A-delta fibers and C fibers. Both of which are part of the sensory-somatic branch of the peripheral nervous system. From these nerve fibers, the pain signal is sent to the spinal cord where it is transferred to another nerve called the interneuron. From the interneuron, the signal is carried to the spinothalamic tract which proceeds up into the brain. When the signal reaches the brain, nerves within the brain interpret the signal as pain.

Our body produces natural painkillers that suppress the pain signals from reaching the brain. One of several natural painkillers is a ligand called enkephalin. This chemical is released along the interneuron and binds to receptors that are in close proximity to where the pain signals are passed onto the spinothalamic tract. When this ligand binds to the receptor, pain signals become suppressed, thus stopping the signal from reaching the brain. Other than enkephalin, there are synthetic painkillers that can also bind to the same receptors. These receptors include mu, sigma, delta, and kappa

receptors which play an important role in the regulation of pain in the spinal cord and in the central nervous system.

A synthetic chemical that can bind to the same receptors as enkephalin are opioids. Opioids are highly addictive drugs that are known for analgesic effects. The most common opioid agonist used is morphine. Because of morphine's structure, it binds to mu receptors quite easily. As shown in the figure 1 below, morphine has 5 rings: an aromatic ring, a cyclohexane ring, a cyclohexene ring, a piperidine ring, and tetrahydrofuran ring.

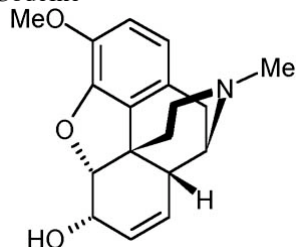
Figure 1: Morphine



The amine nitrogen in the morphine molecule binds to an anionic site of the receptor. Also, the receptor has a cavity that accommodates the projecting piperidine ring, which is the ring containing the amine nitrogen. Lastly, there is a flat surface of the receptor that binds the aromatic ring by Van der Waal's forces. Because of the receptor's shape, morphine is able to fit easily with the receptor.

There are other molecules that have the same basic shape and can be derived from the morphine structure. These molecules are called morphine analogues. An example of a morphine analogue is codeine. If we were to substitute the hydroxide group on the aromatic ring of morphine to $-OCH_3$ by the process of O-methylation, we would yield codeine (Figure 2). Because of this process, when codeine binds to the receptor site, its potency is reduced by 15%.

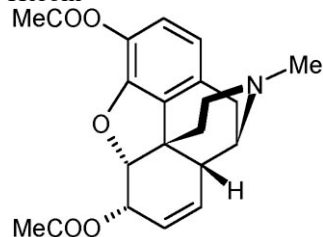
Figure 2: Codeine



Codeine's molecular structure of where it binds to the receptor is identical to that of morphine; however, the $-OCH_3$ makes the protruding carbon chain branch, thus changing the potency. The potency of the molecule seems to be dependent on the length of the protruding carbon chain. This protruding carbon chain decreases lipid solubility, thus a decrease in potency.

Also known as 3,6-diacetylmorphine, heroin is the most well known of morphine analogue structures. If the hydroxide groups were to be substituted with $-OCOCH_3$ we would yield heroin, another opioid agonist that is 2-3 times more potent than morphine (Figure 3).

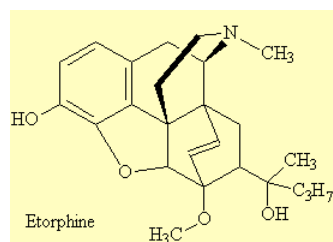
Figure 3: Heroin



Most of the increase in potency is due to the protruding carbon chain. The length of this protruding chain of heroin is two carbons and this causes increased lipid solubility. This leads to enhanced rapid penetration into the central nervous system.

A more powerful morphine analogue and opioid agonist is etorphine. Etorphine has a much larger formation than morphine's structure (Figure 4). The structure of etorphine is identical to morphine in the region where it binds to the receptor site. However, it is thought that there is a lipophilic pocket into which the $C(OH)MePr$ side-chain might fit on the receptor, thus increasing the lipid solubility in the receptor site. This $C(OH)MePr$ side-chain is an eight carbon chain. This increase in carbon chain length makes etorphine 1000 times more potent than the side-chain-lacking-structure morphine.

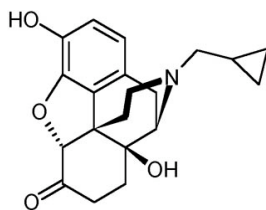
Figure 4: Etorphine



Currently this chemical is used to immobilize large animals, such as elephants, and is too dangerous to use for human therapy.

It is possible to have a chemical structure that is derived from the opioid-agonist morphine but not be an opioid agonist themselves. These structures are called opioid-antagonist morphine-analogue structures. An example of an opioid antagonist is naltrexone (Figure 5). Naltrexone is used to treat patients' narcotics addictions.

Figure 5: Naltrexone



This chemical again has a similar structure to morphine except for several substitutions. When the N-methyl group from morphine is substituted with larger alkyl groups, analgesic effects reduce. In Naltrexone, specifically, the new group is an N-cyclopropyl group thus the chemical name is N-cyclopropyl-noroxycodone.

Morphine, Codeine, Heroin, Etorphine, and Naltrexone all bind to the same receptor. Structurally, these molecules are similar in the region that binds to the receptor site, but different in the protruding carbon chain. Also, the electrostatic potential will be observed to see if there is any support in the binding of the molecules to the receptor site. The purpose of this experiment is to explore morphine and morphine analogue structures and each of the molecules potencies for the same receptor site.

Our hypothesis is that there will be a correlation between the molecular structure and electrostatic potential with potency of the molecule. Hopefully from this experiment, there will be an explanation to why simple changes in the molecular structure and electrostatic potential charges yield such different results in the potency effects of each molecule.

Computational approach

The Molecules were built using *WebMO pro* and then run through the North Carolina High School Computational Chemistry Server. Molecular Orbitals calculations were run using model chemistry AM1//HF/6-31G(d). The Geometry Optimization was done using the software package MOPAC and the Molecular Orbitals calculations were run using Gaussian94. After all calculations were run, regressions were made relating the potency to: LUMO, HOMO, max electrostatic potential, and min electrostatic potential. No significant correlations could be discerned from the HOMO, max or min electrostatic potentials. However a strong correlation was found between the LUMO

energy and the potency. The resultant regression is as follows:

$$y = 10^{-33}x^{-38.379}$$

$$R^2 = .8651$$

Where y is the potency and x is the LUMO energy in Hartrees.

Refer to Index: Chart 1

Results and Discussion

The data showed no correlation for maximum or minimum electrostatic potential versus the potency of the molecules. Even though this could explain the binding potential of the molecules, the electrostatic potential is not a decisive factor of the potency of the molecule. However, further analysis of the molecules reveals that the LUMO energies in relation to the potency of the molecule have a correlation. The fitted regression line for the LUMO energies vs. potency was $y = 1 * 10^{-33}x^{-38.379}$ with an R^2 value of .8651. The graph shows that the lower the LUMO energy, the higher the potency the molecule (Chart 1, Index). This makes sense because when the LUMO energy is lower, it becomes easier for valence electrons to jump into a higher unoccupied molecular orbital. Thus, jumping into the LUMO and making the bound molecule more reactive.

Also, the length of the carbon chain of each molecule in relation to potency was confirmed. As the protruding carbon chain became longer, the potency became greater. This proves to be true up until eight carbons in the carbon chain. After eight carbons, the molecule becomes too large. At this point, the potency starts to decrease due to problems in solubility.

Conclusions

1. Our initial hypothesis that there was a correlation between either the maximum or minimum electrostatic potential and the potency of opioids was incorrect.
2. Following further investigation a strong correlation ($R^2 = .8651$) was found between the LUMO energies and the potency in humans.

3. We also found that there is a very strong relationship between the length of the carbon chain on the cyclohexene ring of the opioid and the potency. The most potent length for the carbon chain is 8 carbons. This is presumably caused by increased lipid solubility.
 4. In addition we found that branching of the carbon chain decreases the potency, i.e. the carbon branched off the oxygen on the cyclohexene ring of codeine. This is also presumably caused by changing the lipid solubility.
52. Zhang, N. et al. "Enkaphalins." *J. Biol. Chem.* 278, 12729. 2003
 53. Booth, Martin. "Heroin and other Opioids." *Opium: A History*. 1996. BLTC. 4 Jun 2007 <<http://opioids.com/refs/index.html>>.
 54. Messer, William. "Opioid peptides." *Opioid Systems*. 23 Feb 2000. MBC 3320. 4 Jun 2007 <<http://www.neurosci.pharm.utoledo.edu/MBC3320/opioids.htm>>.
 55. Kimball, J. "Heat, Cold, and Pain." *Biology Pages*. 25 Feb 2007. 4 Jun 2007 <<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Pain.html#Pain>>.
 56. Wayne, "Chemistry of Opioid Analgesics." *PHA 422: Neurobiology Pharmacotherapeutics*. ACS Medical Chemistry. 4 Jun 2007 <<http://www.acsmedchem.org/module/opioid.html>>.

Acknowledgement

The authors thank Mr. Gotwals for assistance with this work. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

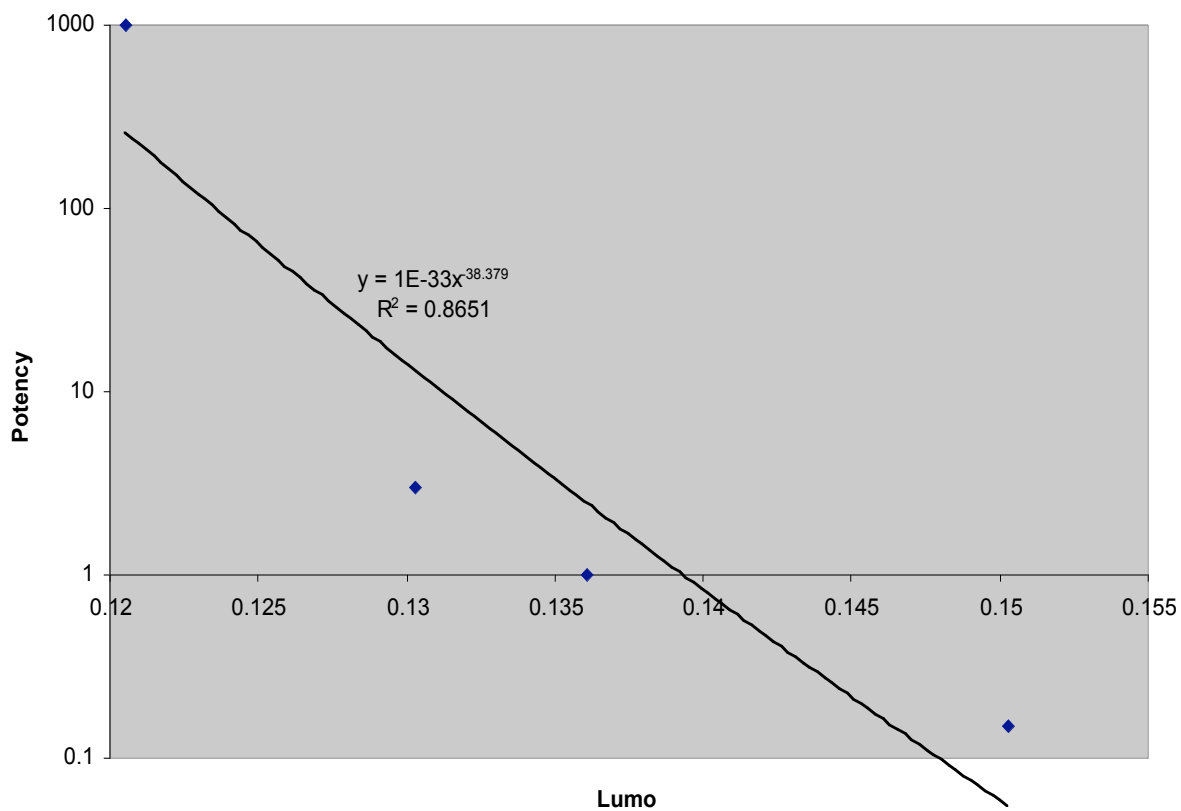
48. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
49. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
50. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
51. Cotton, Simon. "Etorphine." *MOTM 2001*. 2003. 4 Jun 2007 <<http://www.chm.bris.ac.uk/motm/etorphine/etorphine.htm>>.

Index:

Chart 1:

Name	Stoichiometry	HOMO (Hartree)	LUMO (Hartree)	HOMO Orbital	LUMO Orbital	Max El. Pot.	Min El. Pot.	Potency
Morphine	C ₁₇ H ₁₉ NO ₃	-0.29637	0.13607	76	77	0.16617	-0.11362	1
Etorphine	C ₂₉ H ₄₁ NO ₄	-0.29254	0.12054	127	128	0.17521	-0.12075	1000
Heroin	C ₂₁ H ₂₃ NO ₅	-0.31561	0.1303	98	99	0.10673	-0.13317	3
Naltrexone	C ₂₀ H ₂₃ NO ₄	-0.30243	0.12682	91	92	0.17588	-0.1226	
Codeine	C ₁₈ H ₂₁ NO ₃	-0.28006	0.15028	80	81	0.14899	-0.13222	0.15

Lumo vs. Potency



Computational Study of Relationship between Dihedral Angle and Various Chemical Properties of Thermoplastics

Ko, S. and Liu, A.

North Carolina School of Science and Mathematics, Durham, NC

Received 5 June, 2007

Abstract: Thermoplastics, whose properties can be significantly altered with the addition or the removal of heat, are some of the most widely commercially manufactured polymers in the chemical industry. Potential energy scans (PES) were run on the five most common thermoplastics – polyethylene, polystyrene, polyvinyl chloride, and polyethylene terephthalate – to determine their dihedral angles and most stable structural configuration. Various other properties, including the lowest energy, dipole moment, and heat of formation, of the thermoplastics were also computationally calculated, and they were compared to the results of the PES. Relatively high positive correlations were observed between the dihedral angle and both the lowest energy and the heat of formation of the thermoplastic, while a high negative correlation was observed in the graph comparing dihedral angle with dipole moment. From analysis of these results, it was concluded that the various chemical properties of the thermoplastics were closely related to their overall physical structures, which were determined partially by tacticity and dihedral angles.

Key words: Thermoplastic, polymer, potential energy scan, polyethylene, polystyrene, polypropylene, polyvinyl chloride, polyethylene terephthalate, dihedral angle, lowest energy, dipole moment, heat of formation.

Introduction

Thermoplastics

Thermoplastics are plastic materials with a wide range of deformability, melting to a fluid state when heated adequately and freezing to a glassy, inflexible material when cooled. Most thermoplastics are polymers with high molecular weight whose chains can link through Van der Waals forces, dipole-dipole interactions, hydrogen bonding, or stacking of aromatic rings. Currently, there are five main types of thermoplastics extensively manufactured by industries worldwide: polyethylene, polystyrene, polypropylene, polyvinyl chloride, and polyethylene terephthalate.

Polyethylene, as its name indicates, is a polymer consisting of long chains of the monomer ethene. Used in grocery bags, children's toys, and even bulletproof vests, polyethylene is the simplest, as well as the most popular, of all commercial plastics. Polystyrene is formed from the polymerization of the monomer styrene, a liquid hydrocarbon commercially manufactured from petroleum by the chemical industry. Polystyrene materials are usually very strong and able to withstand much stress and wear. Polypropylene

consists of repeating units of propene, and is used in food packaging, ropes, textiles, laboratory equipments, and various automotive components. Polyvinyl chloride (PVC) is one of the most valuable products of the chemical industry, as it is cheap and easy to assemble. Today, over 50% of manufactured PVC is used in construction, replacing traditionally used wood, concrete, and clay. Lastly, polyethylene terephthalate (PET) is used in synthetic fibers, food and liquid containers, and thermoforming applications.

Like many other polymers, thermoplastics are characterized by a certain tacticity, the relative stereochemistry of adjacent chiral centers within a macromolecule chain. There are three types of tacticity – isotacticism, syndiotacticism, and atacticism. As tacticity is closely linked to the physical characteristics of a polymer, the precise knowledge of tacticity is helpful in determining various properties of the polymer, including its melting point, solubility, and other mechanical properties. Isotactic polymers have all their side groups arranged on the same side of the polymer chain, while syndiotactic polymers have alternating orientation of the substituent side groups. Atactic polymers have random

orientation of the side groups, placed at either side of the backbone.

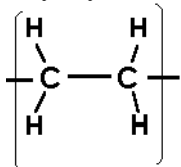
Potential Energy Scan/ Surface (PES)

The Potential Energy Scan is used in association to the Born-Oppenheimer approximation in quantum mechanics to model chemical reactions or interactions. PES functions under the implication that the total molecular wavefunction is written as a product of an electronic and a nuclear wavefunction. The results of the PES can be visualized through a curve or surface, which represents the total energy of an atom arrangement, with atomic positions as variables. Several interesting features can be determined from the analysis of the curve, the most important of which is the global minimum for the energy value. This numerically found value corresponds to the most stable nuclear configuration. Other points of interest may include reaction coordinate, saddle points, local maxima, and local minima.

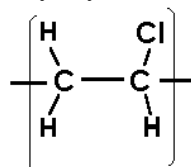
Computational Approach

Using the molecular editor builder of WebMO¹ on the North Carolina High School Computational Chemistry Server², short molecules of polyethylene, polystyrene, polypropylene, polyvinyl chloride, and polyethylene terephthalate (PET) were built. The structures are shown below:

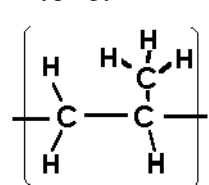
Polyethylene:



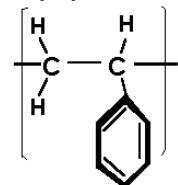
Polyvinylchloride:



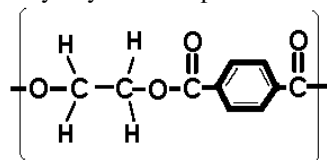
Polypropylene:



Polystyrene:



Polyethylene Terephthalate:



The molecules built consisted only of two repeat units, as run time was an issue.

Each individual molecule was then optimized with the “comprehensive cleanup” option in the WebMO molecular editor. Next, geometry optimizations were run with the semi-empirical software package, MOPAC³, using a PM3 basis set. Geometry optimizations yielded heat of formation and dipole moment data. Using MOPAC again, potential energy (coordinate) scans were run on the bond joining the two repeat units. This calculation gave the dihedral angle in which the molecule was most stable.

All data was placed into an Excel file for analysis and the following were graphed with Graphical Analysis: lowest energy vs. dihedral angle, dipole moment vs. dihedral angle, and heat of formation vs. dihedral angle. Linear regressions were calculated for each data set to determine any correlation.

Results and Discussion

For MOPAC calculations, all energies are reported as heats of formation with units of kilocalories per mole (kcal/mol) and all dipole moments are reported with units of debye. The data results are shown in Table 1.1.

Polymer	Heat of Formation	Dipole Moment	Dihedral Angle	Lowest Energy
Polyethylene	-28.53771	0.005	290	-28.07443
Polystyrene	29.68608	0.249	140	29.74458
Polypropylene	-39.59924	0.051	170	-39.36046
Polyvinyl chloride	-38.39923	1.557	170	-38.26529
PET	-223.7152	2.293	0	-259.34237

Table 1.1 Computational Results

Based on the potential energy scans, the lowest energy states determined by dihedral angle results for each polymer were as followed: polyethylene 290°, polystyrene 140°, polypropylene and polyvinyl chloride 170°, and PET 0°. Polystyrene, polypropylene, and polyvinyl chloride all had dihedral angles close to 180° due to the repulsion of the side groups (i.e. the methyl group, chlorine, and benzene

ring). For polyethylene, the placement of the hydrogens is such that at 290° , there is the least amount of repulsion between the hydrogens. As for PET, the dihedral angle was 0° , causing the double-bonded oxygens of the two repeat units to be farthest away as possible.

The following diagrams are the screen-shots of the optimal structure of each polymer determined by the potential energy scan. Refer to Table 1.1 for the dihedral angle that corresponds to each optimal structure.

Polyethylene-2, Coordinate Scan - Mopac

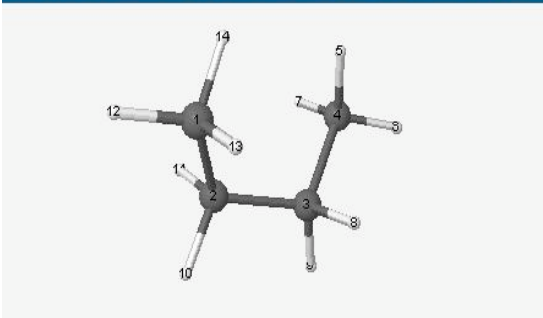


Figure 1.1: Optimal Structure of Polyethylene

Polystyrene-2, Coordinate Scan - Mopac

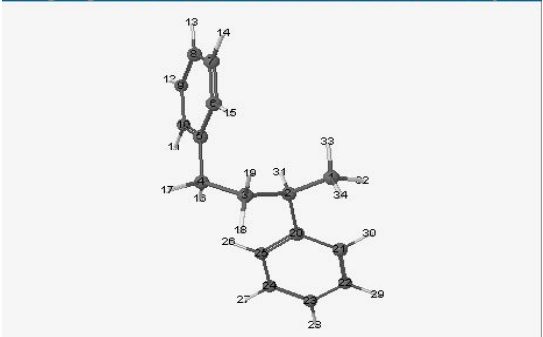
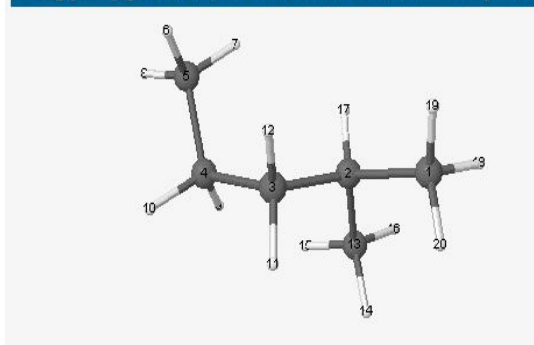


Figure 1.2: Optimal Structure of Polystyrene

Polypropylene-2, Coordinate Scan - Mopac



**Figure 1.3: Optimal Structure of Polypropylene
Polyvinylchloride, Coordinate Scan - Mopac**

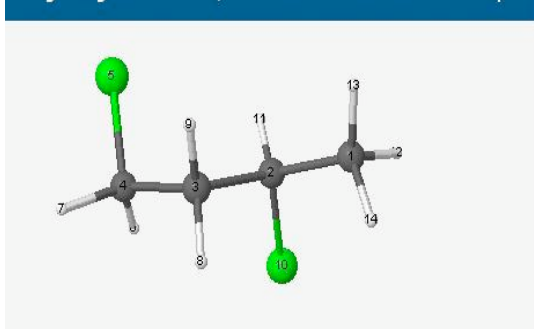


Figure 1.4: Optimal Structure of Polyvinylchloride

PET-2, Coordinate Scan - Mopac

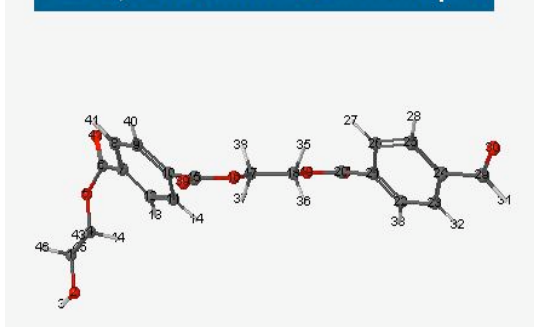


Figure 1.5: Optimal Structure of PET

As seen in these screen-shots, the syndiotactic form was most favorable for polystyrene, polypropylene, and polyvinyl chloride. On the other hand, the isotactic form was most favorable for polyethylene and PET, as seen in Figure 1.1 and 1.5 respectively. In general, the syndiotactic forms had dihedral angles close to 180° while the isotactic forms had dihedral angles close to 0° or 360° .

On the following page are graphs that provide a visual comparison between the dihedral angle and lowest energy (Figure 2.1), dipole moment (Figure 2.2), and heat of formation (Figure 2.3) results.

These data sets had the following correlation coefficients: Lowest Energy vs. Dihedral Angle, 0.748; Dipole Moment vs. Dihedral Angle, -0.769; and Heat of Formation vs. Dihedral Angle, 0.728. All data sets had relatively good correlations (>0.7).

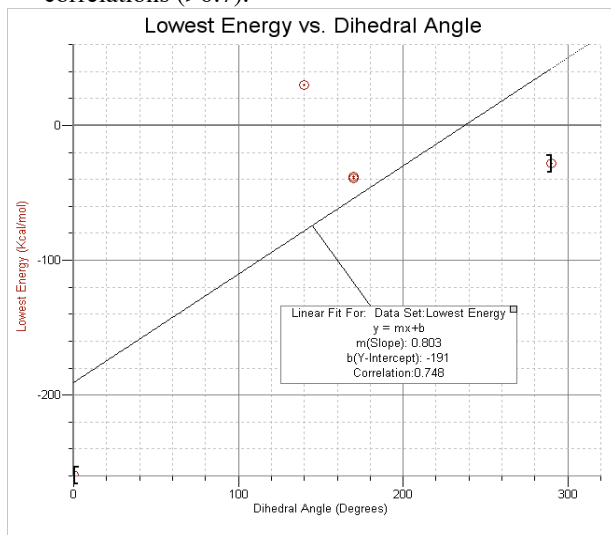


Figure 2.1: Lowest Energy vs. Dihedral Angle

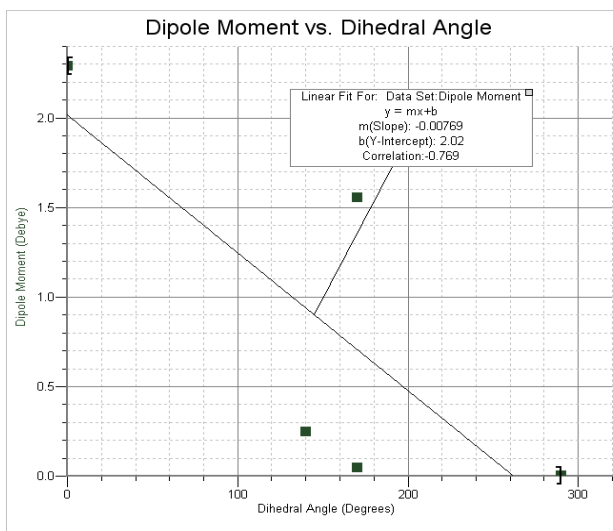


Figure 2.2: Dipole Moment vs. Dihedral Angle

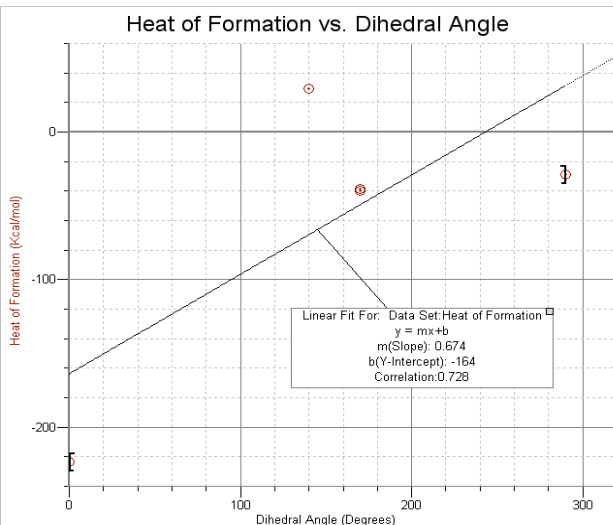


Figure 2.3: Heat of Formation vs. Dihedral Angle
Conclusion

It wasn't surprising that a high negative correlation was observed between dipole moment and dihedral angle ($r = -0.769$). Functional groups are generally the factors that determine polarity, and since the closer the dihedral angles are to 180° , the farther away the functional groups are located, so an inverse relationship was expected between dipole moment and dihedral angle.

The other two graphs (Figure 2.1 and 2.3) were unexpected. They suggest that the lower the dihedral angle, the lower the energy and heat of formation. The opposite phenomenon had been predicted, as it seemed that a dihedral angle approaching 180° would cause the repelling side groups to move farther away from one another, thereby yielding a lower energy and heat of formation. These unexpected results could have been caused by other physical properties or factors that were not taken into account in this project. Despite some unexpected results, the high correlation between dihedral angle, a physical property, and various chemical properties confirmed that structures have a significant impact on a molecule's chemical reactions.

In the future, instead of examining dihedral angles from 0° to 360° , only angles from 0° to 180° should be examined. Another alternative would be to use quadratic fits for the graphs instead of linear fits. These two options would both account for the fact that 170° is really the

same distance as 190° . Without this step, the graphs may be misleading.

Acknowledgement

The authors thank Mr. Robert Gotwals for his assistance in this project. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

57. Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
58. The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
59. MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.

Analysis of the Conformational Structure, Molecular Enthalpy, and Potential Energy Configuration of ethylenediamine tetraacetic acid (EDTA)

T. Stegall and J. Donovan

06/04/07

North Carolina School of Science and Mathematics, Durham, NC

Abstract: Understanding unique characteristics of molecules is crucial for the progress of heavy metals industry and molecular science. Chelating molecules are poorly understood, but they're commonly used in mining, heavy metal recovery, and environmental cleanup. Analysis of ethylenediamine tetraacetic acid, also known as EDTA; the most frequently used chelating molecule, will shed new light into this subcategory of heavy metal chemistry. We performed geometry optimizations using MOPAC AM1 and PM3, along with Gaussian B3LYP. The second set of calculations was a Molecular Energy scan with Gaussian 3-21G. Finally, we ran Potential energy scans utilizing MOPAC PM3 and Gaussian B3LYP, as well as a GAMESS RHF and DFT with B3LYP functional theory with 3-21G basis set. The desired result is a clearer picture of the molecular structure and bonding properties of EDTA. One principle point of interest about the chelated configuration is the forced proximity of strongly electronegative elements: Nitrogen and Oxygen. Under standard conditions, these atoms desire to be as far apart as possible, yet the chelated molecule brings them closer than normal. Logically, the chelated state should therefore possess unfavorable heats of formation. One of the primary goals of this research is to calculate enthalpies of formation for both the chelated and non chelated states of EDTA. Potential Energy and Molecular Energy calculations will help explain why the chelated configuration exists in nature.

Key words: EDTA, Potential energy scan, chelate, ligand

Introduction:

Chelating molecules possess flexible functional groups, forming complexes in which the metal ion is bound to two or more atoms of the chelating agent. A structure composed of primarily single bonds enables the entire molecule to twist and bend into many conformations. These variable configuration states hold great practical applications in medicine, industry, and environmental cleanup. EDTA, or ethylenediamine tetraacetic acid, has four acetic acid functional groups bound to two nitrogen atoms. The nitrogen atoms are connected by carbons. The molecule as a whole has a charge of -4, because the four bonded acetic acid groups have -1 charge. EDTA bonds with metal ions in solution readily

because metal atoms generally have a positive charge ranging from +1 to +4. The unusual characteristic of this bond is that EDTA itself bends around the metal, forming an octahedral set of six bonds, four in the plane, and two perpendicular to the plane. In general, metal ions in solution are toxic to human and animal life, particularly cobalt, mercury, cadmium, and lead. EDTA's cage-like bond is extremely stable, and the bonded system can easily be filtered from water or other solutes for reprocessing and metal recovery. Exact characteristics of these unique bonds are currently indeterminate.

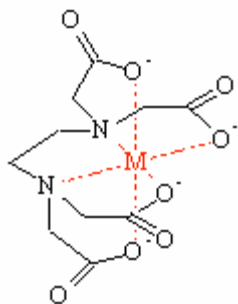


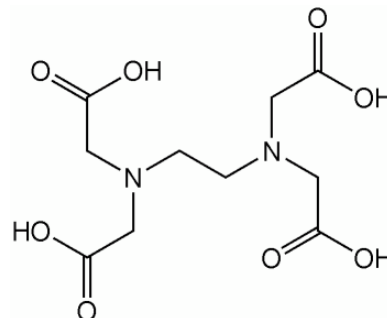
Image of Chelated EDTA, copyright <http://upload.wikimedia.org/wikipedia/commons/5/57/Metal-EDTA.png>

Our primary goal is to configure this molecule using MOPAC PM3 basis sets, calculate the molecular energy and enthalpy of formation, and eventually calculate a 2-dimensional potential energy scan (PES) of the bonds responsible for chelation. If the chelated configuration represents a low potential energy, it might explain why EDTA bonds so readily with the metal ions, because when EDTA chelates, the electronegative oxygen and nitrogen atoms are forced closer together. Forcing electronegative elements together represents an unstable high energy state. Unstable states do not occur naturally in nature, so some other force must lower the potential energy from endothermic values to the exothermic, stable region. Therefore, as the molecule itself does not break bonds during its reconfiguration chelation must reduce the potential energy of the system as a whole through currently unknown means. Ab initio methods are appropriate for a molecule of EDTA's size because it contains only 36 atoms, including the hydrogens. B3LYP theory will provide accurate data. Analyzing the heat of formation change between chelated and unchelated EDTA will indicate how much energy is released

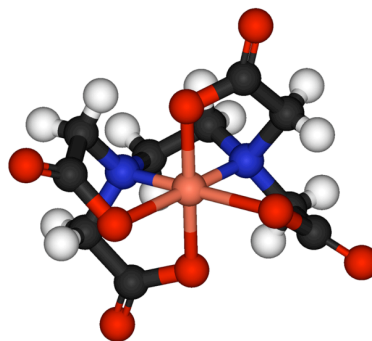
during chelation, an important property to understand if a company wants to extract the metal ion after chelation. The potential energy scan will provide information on energy gain or loss as the molecule reconfigures itself. If the change in potential energy is greater than the repulsion of the electronegative groups, we will understand more about why EDTA exists in nature.

Computational Approach

Using the molecular editor builder of WebMO¹ on the North Carolina High School Computational Chemistry Server², the molecule Ethylenediaminetetraacetic acid, EDTA and Ethylenediaminetetraacetic chelating ligand with Cu²⁺ was built. The initial unchelated structure is shown below.



Unchelated EDTA



Chelated EDTA Cu²⁺

Source:

<http://en.wikipedia.org/wiki/EDTA>

After building the molecule we optimized it with a comprehensive cleanup molecular mechanics package found in WebMO molecular editor. This comprehensive cleanup is very rough and is not used for data acquisition. After running this comprehensive cleanup in WebMO we optimized Ethylenediaminetetraacetic and the EDTA Cu 2+ chelated ligand using many different computational methods. The first optimization used was the computational software MOPAC with the AM1 basis set. Second, we performed an additional MOPAC geometry optimization using the PM3 basis set. Next we performed molecular energy calculations on both geometry optimizations using the computational chemistry software MOPAC with the AM1 basis set along with the PM3 basis set. We also ran molecular energy calculations on both molecules with the computational chemistry software Gaussian. On those calculations we used the Hartree-Fock theory and the B3LYP theory along with the 3-21G basis set.

The MOPAC computational chemistry software along with the 3-21G basis sets provided a rough guide of molecular enthalpy and molecular configuration. These tests also indicated a base line for necessary computational server resources and CPU time. To acquire more accurate data we also used computational methods with greater numbers of basis set calculations. The calculations were run utilizing Gaussian software with B3LYP theory and 6-31G(d) basis sets. The unchelated molecule required seven hours of CPU time. Calculations for the Cu2+ chelated ligand were forcibly halted after sixteen

and half hours due to insufficient available resources.

With the above Geometry Optimizations we performed potential energy scans using the computational chemistry software package MOPAC along with the 3-21G basis set. We also ran a Gaussian potential energy scan on the unchelated molecule with B3LYP theory and 3-21G basis set. We attempted to analyze a potential energy scan of the chelated form but the software encountered unexpected errors and failed to complete its calculation.

Results and Discussion

For MOPAC calculations, all energies are reported as heats of formation with units of kilocalories per mole (kcal/mol). DFT calculations resulted in energies with units of Hartrees (E_h).

The Geometric Optimizations succeeded and provided a valid foundation for the Molecular Energy and Enthalpy calculations. The Potential Energy Scans refused to run to completion for reasons that have yet to be determined. A new method is needed to adequately explore EDTA's molecular configurations.

All data is represented in the following tables.

Gaussian Geometry Optimization: 6-31G(d)

Stoichiometry $C_{10}H_{16}N_2O_8$

Symmetry C1

Basis 6-31G(d)

B3LYP Energy -1101.98387306 Hartree

Gaussian Molecular Energy calculation:
3-21G

Stoichiometry C₁₀H₁₆N₂O₈
Symmetry C1
Basis 3-21G
B3LYP Energy -1095.8888153 Hartree
Dipole Moment
5.7178 Debye

Gaussian Molecular Energy Calculation:
STO-3G

Stoichiometry C₁₀H₁₆N₂O₈
Symmetry C1
Basis STO-3G
RHF Energy -1081.53118947 Hartree
Dipole Moment
4.1917 Debye

Conclusions

Based on the data and analysis, our results indicate the need for more precise, flexible molecular software. Unchelated EDTA has a large negative enthalpy, indicating that it is a stable molecule, but the software does not accept a chelated structure. It is possible that due to the rarity of chelating molecules and the large negative charges of the chelated ligand that the MOPAC, Gaussian and TINKER software were not programmed to perform calculations on EDTA-like molecules. For future study, researchers must either find different software that can deal with molecular charges, or a keyword that forces the calculation to run despite presence of same. One interesting detail meriting further study is our discovery that EDTA is in fact strongly polar, despite its symmetrical structure. The overall polarity of nearly 6 Debye may directly correlate with EDTA's ability to force a chelated state. Perhaps the electric potential between the negatively charged molecule and a positively

charged metal ion is greater than electrostatic Oxygen-Nitrogen repulsion. Again, this is a matter worthy of further detailed study.

Acknowledgement

The author thanks Cmdr. Robert Gotwals of the North Carolina School of Science and Mathematics for assistance with this work. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

- Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).
- The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).
- MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
- Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

Comparison of Atomization Energies on Hydrocarbons Using Theoretical Methods

A. Ellis and M. Scott

North Carolina School of Science and Mathematics, Durham, NC

Received 5 June, 2007; Accepted 5 June, 2007

Published

Abstract: In this study, the accuracies of the PM3 and the B3LYP methods were observed in thermochemical properties. Thermochemical properties are the one of the most difficult fields of chemistry to accurately model, so if accurate data is calculated then that method is effective at calculating areas of thermochemistry. Atomization energy, the chosen thermochemical property, was calculated for a small group a hydrocarbons, optimized at PM3 and B3LYP/6-311+G(d,p) levels of theory. This data was then compared to bench-marked data found in an experimental chemistry database at srdata.nist.gov/cccbdb. Results showed that when calculating energies for molecules, higher levels of theory must always be used in order to obtain accurate data. The PM3 method should never be used when looking for accurate data and only for very general trends, whereas the B3LYP/6-31G(d) method for frequency calculations gave fairly accurate data, with 83.91 and 83.89 Mean Absolute Difference(MAD). Also, it was observed that using a more computationally expensive geometry optimization did not necessarily produce more accurate calculated energies.

Key words: thermochemistry, atomization energy, hydrocarbons, bench-mark, MAD

Table 1: Energies yielded by PM3 optimization

	PM3//B3LYP Energy	PM3//B3LYP ZPE	PM3//PM3 Energy	PM3//PM3 Zpe
methane	-40.51828859	0.045638	-0.020766239	0.045383
ethane	-79.83007986	0.07544	-0.02895405	0.074047
propane	-119.1436194	0.104307	-0.037727922	-23.67435264
butane	-158.4568982	0.132981	-0.046423987	0.130189
pentane	-197.7685024	0.161791	-0.05426388	0.157944
ethylene	-78.58707533	0.051394	0.026455563	0.049805
propene	-117.9067093	0.080264	0.01012594	0.0785
Z-2-butene	-157.2214861	0.109336	-0.005770023	0.107462
Z-2-pentene	-196.5341042	0.13794	-0.013807856	0.135574

Introduction

There are two main reasons for doing computational chemistry research: one, to use the ease and effectiveness of computational chemistry in order to model certain areas of chemistry and the other, which is looked into by this article, to improve or determine accuracy of the theoretical methods making computational chemistry more efficient. Such possible areas of research in computational chemistry include organic, environmental, medicinal, reaction, etc., which can all be used to check efficiency. In this article the area of thermochemistry was used as a means of determining and supporting certain methods' accuracies.

Thermochemical properties are very difficult to model accurately, so good results imply good methods. One of the easiest, while still useful, areas of thermochemistry to compare is atomization energies. Atomization energy is the difference between the molecular energy and the combined energy of its component atoms. For example, the atomization energy for H₂O is

There is already very accurate experimental data for atomization energies already obtained from the computational chemistry comparison data base at srdata.nist.gov/cccbdb.

It is important to look not only at the very accurate methods, but to also look at the less accurate but less computationally expensive methods. The cheaper, traditional methods are more important to maximize efficiency than the computationally demanding ones or complicated compound methods, because the traditional methods are used on a more regular basis than the expensive methods such as the Full CI method, which has full electron correlation.

The atomization energies were found for a small molecule group of hydrocarbons. This list of hydrocarbons includes methane, ethane, propane, butane, pentane, ethylene, butane, propene, Z-2-butene, Z-2-pentene. Z-2-butene and Z-2 pentene were chosen because of the accessibility to very accurate experimental data. A smaller sized group of molecules helps to predict the accuracies of calculated data for specific molecules using specific methods, which produce specific conclusions. These conclusions can then be compared to more general views at method accuracies.

	PM3//B3LYP	Percent Diff.
methane	421.755201	0.061019376
ethane	726.575535	0.07593001
propane	1033.07964	0.080966463
butane	1339.54124	0.083508248
pentane	1644.86664	0.084432123
ethylene	591.884090	0.099134802
propene	902.210565	0.09784688
Z-2-butene	1209.36239	0.095635439
Z-2-pentene	1515.45333	0.095138988

Computational Approach

Gaussian was used for all of the calculations in this article. In order to test the

	B3LYP//B3LYP	Percent Diff
methane	421.9689048	0.06156
ethane	724.1941189	0.07240
propane	1030.518779	0.07829
butane	1337.424324	0.08180
pentane	1647.043272	0.08587
ethylene	592.1316481	0.09959
propene	902.7784379	0.09854
Z-2-butene	1210.181402	0.09638
Z-2-pentene	1518.22602	0.09714

Table 3: AE and percent difference

accuracies of the energy calculations of the PM3 and B3LYP/6-31G(d) methods, we ran the hydrocarbons that we selected through a different combination of geometry optimizations, using the PM3 and B3LYP/6-311+G(d,p) theories, and vibrational frequency calculations, which produces the single-point-energy and zero-point-energy (ZPE). Essentially every molecule was optimized two ways and each of the optimized geometries was used for the two types of vibrational frequency calculations with both PM3 and B3LYP/6-31G(d). As was expected the calculations using PM3 took significantly less CPU run-time than the more expensive B3LYP/6-31G(d).

For the Atomization Energy calculations we needed the single-point-energy of the Carbon and Hydrogen atoms. They were calculated with the molecular energy calculation with the most accurate basis set available, B3LYP/6-311+G(d,p), because we want to smallest source of out-side error. To find the Atomization energies of the hydrocarbons this formula was used.

Table 2: AE and percent difference

	B3LYP//B3LYP Energy	B3LYP//B3LYP ZPE	B3LYP//PM3 Energy	B3LYP//PM3 ZPE
methane	-40.51837515	0.045384	-0.020720887	0.045041
ethane	-79.82590379	0.075059	-0.025873934	0.073764
propane	-119.1392724	0.104041	-0.034115635	0.101823
butane	-158.4532837	0.13274	-0.042327367	0.129679
pentane	-197.7717341	0.161554	-0.05317183	0.157457
ethylene	-78.58743285	0.051357	0.026789989	0.049519
propene	-117.9075253	0.080175	0.010125927	0.0785
Z-2-butene	-157.2215492	0.108094	-0.002398074	0.106236
Z-2-pentene	-196.5380648	0.137482	-0.009994096	0.133563

Table 4: Energies yielded by B3LYP/6-311+G(d,p) optimization

The total energy of the molecule can be found by adding the energy from the frequency calculation and the ZPE that the calculation yielded.

Results and Discussion

PM3//B3LYP/6-31G(d)

Optimizing the hydrocarbons with PM3 level of theory is significantly cheaper than other such as B3LYP, which is why we are testing its accuracy in this experiment. The energy results can be seen in table 1. The energies yielded by this calculation were consistent and followed a pattern. As the molecules increased in size, the energy became increasingly negative. Also, ethylene is less negative than ethane which can be expected. The Atomization energies yielded by this optimization and calculation were very close to the expected values. For example the atomization energy of methane had a 6.1% difference with the expected value. The mean absolute difference(MAD) was 83.91, but as can be seen from table 2 the difference increases as the molecule's size increases. On the other hand percent difference shows a different trend. The percent difference does not drastically increase as the molecules get larger, which shows that the increasing difference is relatively proportional to the size. The percent difference shows that the difference between the calculated and

expected increases for the –enes. The complexity of the double bonds seems to increase the error of the calculations.

B3LYP/6-311+G(d,p)//B3LYP/6-31G(d)

These molecules were optimized using the most expensive method and basis set available, in order to see if it would be significantly different from the PM3 optimizations. The runtimes were much longer than those optimized at the PM3 level. For example it took an hour and seven minutes to optimize C₅H₁₀ with B3LYP/6-311+G(d,p) and a minute and fifty-eight seconds at the PM3 level. The result of these energy calculations can be seen in table 4. These energies follow the same pattern of the ones calculated with PM3//B3LYP/6-31G(d). They increase in size as the complexity of the molecule increases but only proportionally to the total energy.

The atomization energies found using PM3 vibrational frequencies were very inaccurate so we have disregarded them. We compared the calculations using the MAD and largest errors, to see how effective the methods of optimizing the molecule are. The MAD for the B3LYP/6-31G(d) calculations optimized by PM3 was 83.91 and for B3LYP/6-311+G(d,p) it was 83.88. The difference is statistically insignificant, so based on that data, both methods yield relatively the same error. This data is solidified by the similarities in the largest errors which are -131.65 and -134.43 for PM3 and B3LYP/6-311G(d,p) respectively. The methods also yield similar percent differences which serves to support the theory that a less expensive basis set can be used to optimize the geometry and it will yield accurate results for atomization energies.

Conclusions

This study on computational calculations on thermochemical properties has shown how the accuracy of thermochemical calculations are related to the basis set used in the geometry optimization. It can be inferred that without a significant loss of accuracy, it is more efficient to optimize the geometry of hydrocarbons with a lower level of theory and then run a higher

level of theory when doing an energy calculation. The MAD of the two sets ran with the B3LYP/6-31G(d) frequency calculations had a difference of only .02 kcal/mol, whereas the difference between the MAD of the PM3 frequency calculations was 562.31 kcal/mol.

Using a low level theory, such as PM3, will produce highly inaccurate results when performing frequency calculations on hydrocarbons. It should never be used for calculating any type of precision thermochemical properties, only for general trends when computational cost is very high.

It can also be concluded that hydrocarbons are difficult molecules for Semi-empirical and DFT methods to model. There was large error present in every calculated atomization energy; the smallest difference between a calculated value and the corresponding experimental value is 24.23 kcal/mol, which occurred in the PM3//B3LYP/6-31G(d) calculation of methane (note this was not the most computationally expensive job).

As the complexity of the molecule increased, the accuracy of the calculated data stayed very consistent. It was determined that hydrocarbons are difficult to model, and increased complexity of the molecule even worsens the accuracy of the atomization energies. Further studies may be done on thermochemical properties of hydrocarbons by choosing different hydrocarbons, levels of theory, and basis sets.

Acknowledgements

The authors thank Mr. Robert Gotwals for giving us this project. Appreciation is also extended to the Burroughs Wellcome Fund and the North Carolina Science, Mathematics and Technology Center for their funding support for the North Carolina High School Computational Chemistry Server.

References

Schmidt, J.R.; Polik, W.F. *WebMO Pro*, version 7.0; WebMO LLC: Holland, MI, USA, 2007; available from <http://www.webmo.net> (accessed April 2007).

The North Carolina High School Computational Chemistry
Server, <http://chemistry.ncssm.edu> (accessed
April 2007).

MOPAC Version 7.00, J. J. P. Stewart, Fujitsu Limited,
Tokyo, Japan.
Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H.
B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J.
A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J.
M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B.
Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson,
H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J.
Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H.
Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B.
Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.
E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C.
Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A.
Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S.
Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K.
Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J.
V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B.
B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi,
R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y.
Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B.
Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A.
Pople, Gaussian, Inc., Wallingford CT, 2004.

Foresman, James B., and Æleen Frisch.

Exploring Chemistry with Electronic Structure
Methods. 2nd ed. Pittsburgh: Gaussian, Inc.,
1996.

National Institute of Standards and Technology,
Computational Chemistry Comparison and
Benchmark Database. Aug 2005. NIST. 3 Jun
2007 <<http://srdata.nist.gov/cccbdb/>>.

Examining Active Sites in Reactions between O₃ (O-zone) and Leading Chlorofluorocarbons (CFCs)

M. Dutra

North Carolina School of Science and Mathematics, Durham, NC

Received 3 April, 2007; Accepted 5 May, 2007

Published online on Comp Chem Moodle (moodle.ncssm.edu)

Abstract: One of the prominent problems in the world today is the issue of o-zone deterioration. The layer of o-zone in the stratosphere, the layer that protects Earth from harmful ultraviolet rays, is deteriorating due to the abundance of chlorofluorocarbons (CFCs) that pollute it. Because o-zone (O₃) is a highly reactive molecule, it reacts easily with these CFCs to produce O₂ and the reactive CFC molecule (usually a halogen), meaning that the CFC molecule is not “used up,” and is still free to react with other molecules of o-zone. This experiment attempts to find the active site of the reaction between o-zone and five different CFCs: CFCl₃, CF₂Cl₂, C₂F₃Cl₃, CF₂ClBr, and CH₃Br. This was done computationally by using Gaussian 6-31(d,p) calculations of the molecular orbitals of each molecule, and the results were analyzed to find the active sites in each reaction.

Key words: O-zone, chlorofluorocarbons (CFCs), active site, Gaussian, 6-31(d,p), molecular orbitals

Introduction

Perhaps one of the more important problems plaguing the world today is that of the depletion of the o-zone layer. The o-zone layer is a layer of O₃ molecules that sits atop the highest portion of the stratosphere and filters out harmful ultraviolet (UV) radiation, so that only safe amounts reach the earth. However, because of the production and release of chemical compounds known as chlorofluorocarbons, or CFCs, the o-zone layer is developing “holes,” or places in the layer where there is a lack of O₃ molecules. Obviously this is of considerable concern, because these “holes” allow for unsafe levels of UV radiation to reach the earth.

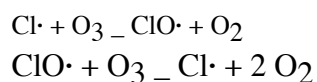
The o-zone layer is in a constant reaction cycle with itself. Since it is located in the upper stratosphere, it receives an abundance of heat from the sun, which provides the energy for the reactions to take place. It is this consumption of energy that actually filters out the harmful ultraviolet rays radiated from the sun, which is how the o-zone layer provides its usefulness to all life on Earth. The energy radiated from the sun causes a bond to split between two of the oxygen molecules, which yields the products of O₂ (the breathable oxygen gas) and O. These two

chemicals then quickly react with each other to once again form O₃, or o-zone. This reaction is once again fueled by the energy radiated by the sun. This cycle is in continuous motion, meaning that it continues to absorb the unsafe levels of radiation from the sun with no net loss of o-zone. The cause for the loss of o-zone occurs when an alien molecule disrupts this cycle.

In the case of the o-zone layer, these foreign molecules are comprised of many different chlorofluorocarbons, often called CFCs. These CFCs interrupt the continuous oxygen-ozone cycle, causing the second reaction in the cycle (the reaction that reforms the o-zone molecule) to cease to occur. This lowers the levels of o-zone in certain parts of the stratosphere, causing for the “holes” in the o-zone layer.

The chlorofluorocarbons disrupt the o-zone-oxygen cycle by means of reacting with the o-zone so that the individual O atom is never produced, leaving the free O₂ molecules with nothing to react with to reform the o-zone molecule. The halogen atom(s) attached to each CFC are the atoms that react with the o-zone, causing it to degrade into products other than O₂ and O. However, the halogen atom is still free at the end of the reaction, so it is capable of

reacting with more o-zone molecules. This reaction process is seen below, with the halogen chlorine (Cl), a halogen often found in CFCs:



Notice that the chlorine atom is reproduced at the end of the reactions, and the process can be repeated.

The reactive property of o-zone is a major factor in its depletion. Quite possibly linked to this property is the molecular structure of both o-zone and the CFCs that react with it. The structure of a molecule plays an important role in determining what reacts with it, and how often the reaction occurs. It is the goal of this journal article to find a connection between the molecular structure of the o-zone molecule, the molecular structure of chlorofluorocarbons, and the reactive tendencies of both.

Computational Approach

Using the molecular editor builder of WebMO¹ on the North Carolina High School Computational Chemistry Server², the o-zone molecule was built. The molecule was then optimized using the “comprehensive cleanup” function in the WebMO package. Following this, the molecule was further optimized using a Hartree/Fock Geometry Optimization at a 6-311(d,p) basis set in the Gaussian '03 package.

After the o-zone molecule was fully optimized, a molecular orbitals calculation was run on it, again using a 6-311(d,p) basis set in the Hartree/Fock level of theory. The goal of these calculations was to obtain an accurate mapping of electrostatic potential for the o-zone molecule.

Following the construction of the o-zone molecule was the construction of four of the leading chlorofluorocarbons. These four CFCs were: trichlorofluoromethane (CFCl₃), dichlorofluoromethane (CF₂Cl₂), 1,2,2-trichloro-1,1,2-trifluoroethane (C₂F₃Cl₃), and chloropentafluoroethane (C₂F₅Cl). The same comprehensive cleanup was performed on these molecules as was o-zone. They were then all optimized further by means of a Geometry Optimization calculation at the Hartree/Fock

level with a 6-311(d,p) basis set. Once these calculations were run, the molecular orbitals of each were calculated, once again using a Molecular Orbitals calculation on the Hartree/Fock level of theory with a 6-311(d,p) basis set.

From the Molecular Orbitals calculations, an electrostatic potential map was constructed for each of the chlorofluorocarbons, as well as the o-zone molecule.

Results and Discussion

The electrostatic potential mappings of each molecule are shown below in Figs. 1.1 – 1.5. The red portions of the maps indicate negative regions of the molecule, or regions of the molecule where the electrons are mainly located. The blue portions of the maps are the positively charged regions of the molecule, or where there is a lack of electrons. The green portions are neutrally charged, and play little purpose in the context of this research. The red regions of each molecule, being negatively charged, are likely to bond with a positively charged pole on any nearby molecules. The blue regions, contrary to the red, are positively charged, so are most likely to bond with any negatively charged poles on nearby molecules. This extreme polarity on the o-zone molecule accounts for its high reactivity, because the two distinct poles allow for either positively charged or negatively charged molecules to interact with it.

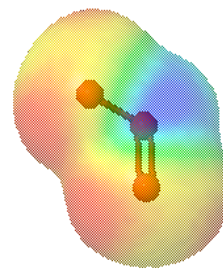


Fig. 1.1 O-zone Electrostatic Potential

Jmol

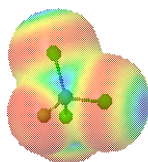


Fig. 1.2 CFCl_3 Electrostatic Potential

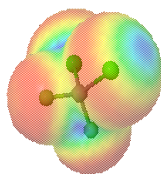


Fig. 1.3 CF_2Cl_2 Electrostatic Potential

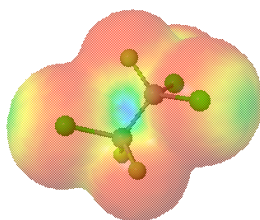


Fig. 1.4 $\text{C}_2\text{F}_3\text{Cl}_3$ Electrostatic Potential

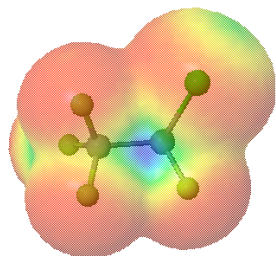


Fig. 1.5 $\text{C}_2\text{F}_5\text{Cl}$ Electrostatic Potential

As can be seen from the electrostatic potential mappings, the o-zone has easily accessible negative and positive poles, leaving it easily susceptible to reacting with chlorofluorocarbons in the surrounding atmosphere. The four CFCs are dominantly negative, with positive regions located near the center of the molecule. The largely negative exterior of the CFCs allows for them to react mainly with positively charged molecules. The HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital) projections for the o-zone molecule can also be useful in determining the reaction site of the o-zone-CFC reaction, because these projections show where lone pair electrons are likely to be (HOMOs), as well as where any lone pair electrons would be likely to bond (LUMOs). These projections contribute to helping find where the active site on the o-zone molecule is by showing where electrons are likely to be, and thereby narrowing the number of potential reaction sites.

Conclusions

From the above data and experimentation, it can be determined that the molecular structure of o-zone, along with the structures of the chlorofluorocarbons that interact with it, plays a significant role in its high reactivity. From analyzing the electrostatic potential projections of all of the molecules, the active site was found to be on the positive pole of the o-zone molecule, on the middle oxygen atom. This is due to the high negative charge commonly found on all of the CFC molecules. These highly negatively charged molecules interact readily with the positively charged region of o-zone, causing a site where the chemical reaction would most likely occur. This conclusion can further be supported by the fact that the negative poles of the o-zone molecule are repulsed by the negatively charged CFCs, leaving only the positive pole as a place for potential reactions.

For further experimentation, the reaction pathway for the interaction between o-zone and the CFCs could be experimentally determined. This data would further knowledge about the reaction between the o-zone molecule and a given CFC.

Acknowledgement

The author thanks Mr. Robert Gotwals for access to the North Carolina High School Computational Chemistry Server, as well as the North Carolina High School Computational Chemistry Server itself.

References

The North Carolina High School Computational Chemistry Server, <http://chemistry.ncssm.edu> (accessed April 2007).

Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.