

PM3 and DFT Computational Studies of the Reaction Mechanism of Formaldehyde and Isoleucine

Umar Hamza¹, Adamu Uzairu², Muhammad Sani Sallau², Stephen E. Abechi²,
Gideon A. Shallangwa²

¹Chemistry Department, Federal University of Education, Kano, Nigeria.

²Chemistry Department, Ahmadu Bello University, Zaria, Nigeria.

*Corresponding author email: huzango@yahoo.com

Received : 1 February 2016

Accepted : 20 April 2016

Online : 30 April 2016

Abstract – The reaction of formaldehyde and Isoleucine was studied using semi-empirical and density functional theory methods. Possible reactive sites are proposed and reaction mechanism postulated. It was found that the Isoleucine nitrogen attacks the carbonyl carbon of formaldehyde and forms a methylol intermediate that undergoes a condensation with another Isoleucine to produce a Methylenediisoleucine through a methylene bridge (cross-linking). The enthalpies of the reaction are -78.79 kJ/mol and -39.14 kJ/mol for PM3 and DFT respectively also ΔS° and ΔG° , for the PM3 and DFT studies predicted. The reaction was found to be exothermic and second order.

Keywords: Formaldehyde; Isoleucine; Transition states; Reaction mechanism; Enthalpies

Introduction

Formaldehyde is an organic compound with the formula CH_2O . It is a colourless gas with a characteristic pungent smell. Formaldehyde is more complicated than many simple carbon compounds because it adopts different forms. It is a gas at room temperature, but the gas readily converts to a variety of derivatives. It has smaller molecular weight and it reacts with a variety of chemical compounds. It is ubiquitous and natural constituent of all living systems, from bacteria and fish to rodents and humans (Robert, 2005). It is one of the most abundant compounds in the universe (Robert, 2005). It is present in indoor and outdoor air (Zhang *et al.*, 2014) and cosmetics, toothpaste, ink and paper (Shabnam, 2014). Formaldehyde is widely used in industries in manufacture of numerous products such as resins, adhesives, plastics, plywood insulators, fertilizers, fungicides, corrosion inhibitors, embalming fluids, disinfectants, preservatives, sterilizing agents and many household products (Robert, 2005; Zhang *et al.*, 2014; Shabnam, 2014; TEACH, 2007).

Formaldehyde possess many potentially detrimental effects to body system, studies have shown that exposure to formaldehyde cause harms on respiratory system and long term exposure inhalation at dose of 15ppm was induced squamous cell carcinoma in nasal cavity of rats and mice and cause headache, irritation and depression in humans (Robert, 2005; Kerns *et al.*, 1983 and Hussein *et al.*, 2014). It rapidly reacts with DNA, RNA and proteins to cause detrimental effects on human health (Zhang *et al.*, 2014; Gurel *et al.*, 2005). Formaldehyde was found to cause cytotoxicity through the formation of strong DNA-protein cross-links as, well as cross-links with other molecules (Gurel *et al.*, 2005). It has also have an ability to cause neurotoxicity by the reaction DNA-protein cross-linking and unsaturated fatty acids (Thrasher *et al.*, 1990).

Formaldehyde is used in the synthesis of biomolecules and participates in synthesis of amino acid glycine and alanine in aqueous medium (Miller, 1953). Formaldehyde is highly reactive and a well-known cross-linking agent which forms a stable methylene bridges between nucleic acids and amines or proteins and their polymerization and also forms methylol compounds (Conaway *et al.*, 1996; Toews *et al.*, 2008; Mertz *et al.*, 2004).

Materials and Methods

The computations processing were made using the standard version of PM3 semi-empirical MO methods and Density Functional Theory (DFT) method in the Spartan 08 software package as run on a Microsoft windows 7 Home Premium version 2009. SPI computer system with Intel(R) Celeron(R) Dual CPU, B830 @ 1.80GHz !.80GHz, 6.00GB RAM.

Geometry optimization of reactants, intermediates, transition states and products

The electronic structure calculations in this work were performed using the Spartan'08 program packages. The geometries included in the reaction of Isoleucine (Ile) with formaldehyde were fully optimised using the semi-empirical method (Dewar *et al.*, 1985; Stewart, 1989). The structures of the reactants, transition state, intermediates and products were built and minimized with the MM2 method in the Spartan 08. All the geometries were optimized using AM1, followed by parametric method 3 (PM3) and MNDO and finally by DFT Becke's three parameter non-local exchange functional with nonlocal correlation function of Lee, Yang and Parr (B3LYP) with 6-31G (d) basis set calculations were employed. There was a little change in comparing the geometrical parameters obtained by the semi-empirical method (Kalkanis and Shield, 1991) and there is marked differences between semi-empirical and DFT methods. The optimized geometries of all species in the reaction were confirmed in terms of vibration analysis (Mshelia *et al.*, 2010). Transition state for each step was located and confirmed by animating the vibration corresponding the coordinate by selecting the imaginary frequency at the top of the IR tab. Heat of formation (ΔH_f) of the reactants, transition states, intermediate and products were all calculated. Infra-red (IR), thermodynamic and kinetics parameters were also calculated according to the expression in the literature (Thomas and Engel, 2006).

Calculation of Molecular information

Molecular information such as molecular orbital energies, electron density, net charges were calculated by selecting the orbitals and energy, thermodynamics, charges, bond orders and vibrational modes in the calculation menu before submission of the species. The information obtained was used for discussing the properties of the species optimized.

Results and Discussion

In general the most likely interaction site for the two molecules is either the two atoms that have highest and opposite charges or two atoms that have a highest electron densities in their highest occupied molecular orbital/ lowest unoccupied atomic molecular orbital, (HOMO/LUMO) interaction (Elder and Worley, 1984). In this reaction we considered the reaction between the HOMO of isoleucine and the LUMO of formaldehyde. For both the PM3 and DFT success was achieved when the interaction of isoleucine N lone pair with the carbonyl center of formaldehyde was performed. When considering both the charge interaction and most likely the HOMO/LUMO interactions led to the conclusion that isoleucine react with formaldehyde by nucleophilic attack of isoleucine nitrogen (N) lone pair on the carbonyl carbon C1 of formaldehyde.

Geometry optimization of intermediate and transition states

The geometry of the reacting species, transition states, intermediates and products in the reaction of Isoleucine with Formaldehyde was performed. The optimization of all the reacting species was completed successfully and heat of formation at 25°C(298.15K) was evaluated and presented in Table 1 together with other activation parameters. If considered the heat of formation of the optimised transition state as proposed in the reaction, TS2 has the least enthalpy of formation (-2619369Kj/mol) then TS1 (-1460121Kj/mol) according to the DFT method. Likewise, the PM3 method followed the same trend in which the TS2 was favoured also. The potential energy diagram showed the energy profiles of the optimised reacting species, where R1 were reactants of step 1, with TS1 as the possible transition state and P1 the product of that particular step, whereas R2, TS2 and P2 are reactants, transition state and products of the step 2 respectively (Figure 2 and Figure 3).

Figure 2 showed that TS1-2 occupied a saddle point as expected for the semi-empirical method. Figure 3 showed that TS1-2 also occupied a saddle point as expected for the DFT method due to its accuracy. For the DFT the heat of formation for the various species R2 = -1460296Kj/mol, TS2 = -2619369Kj/mol and P2 = -2619441.8 Kj/mol were very close but TS2 was still slightly higher as shown in Figure 3.

Table 1. Heat of formation and other activated parameters of reacting species for isoleucine-formaldehyde reaction system according to DFT and PM3 methods

S/ N	Reacting species	Thermodynamic parameters at 298.15K using DFT			Thermodynamic parameters at 298.15K using PM3		
		ΔH_f° (kJ/mol)	$\Delta S^\#$ (J/molK)	$\Delta G^\#$ (kJ/mol)	ΔH_f° (kJ/mol)	$\Delta S^\#$ (J/molK)	$\Delta G^\#$ (kJ/mol)
1	R1	-1460227.2			-610.57		
2	TS1	-1460121			-460.79		
3	Step1		-1.71×10^2	1.56×10^2		-1.63×10^2	1.86×10^2
4	E _a	2.583×10^3			2.616×10^3		
5	R2	-1460296			-650.21		
6	TS2	-2619369			-809.37		
7	Step2		1.74×10^2	-11.59×10^5		-1.85×10^2	-2.45×10^1
8	E _a	-11.56×10^5			2.788×10^3		

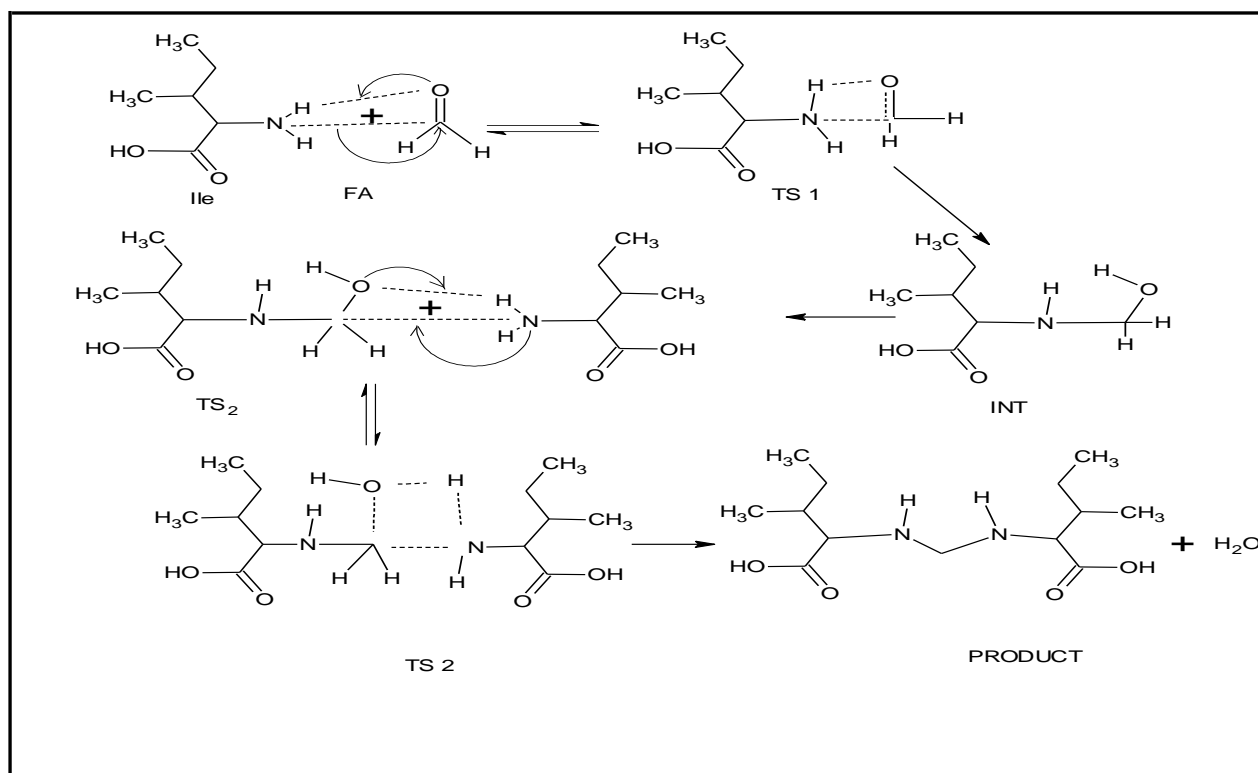


Figure 1. Proposed reaction mechanism

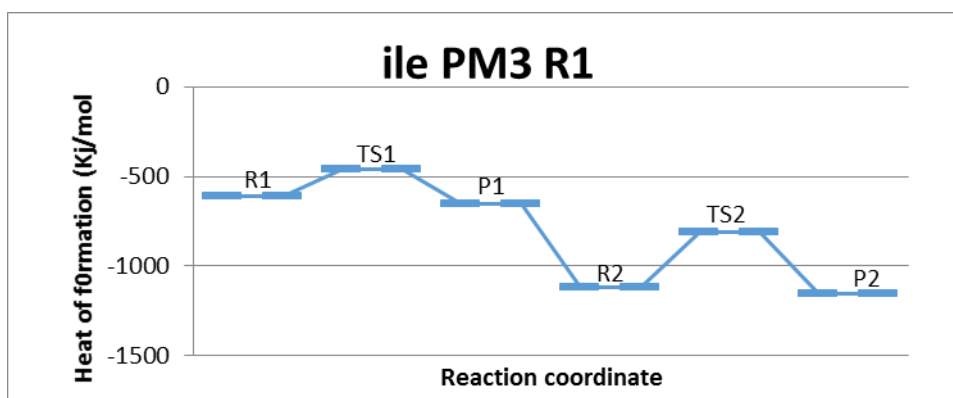


Figure 2. Energy profile of reaction of Isoleucine-Formaldehyde system PM3 method

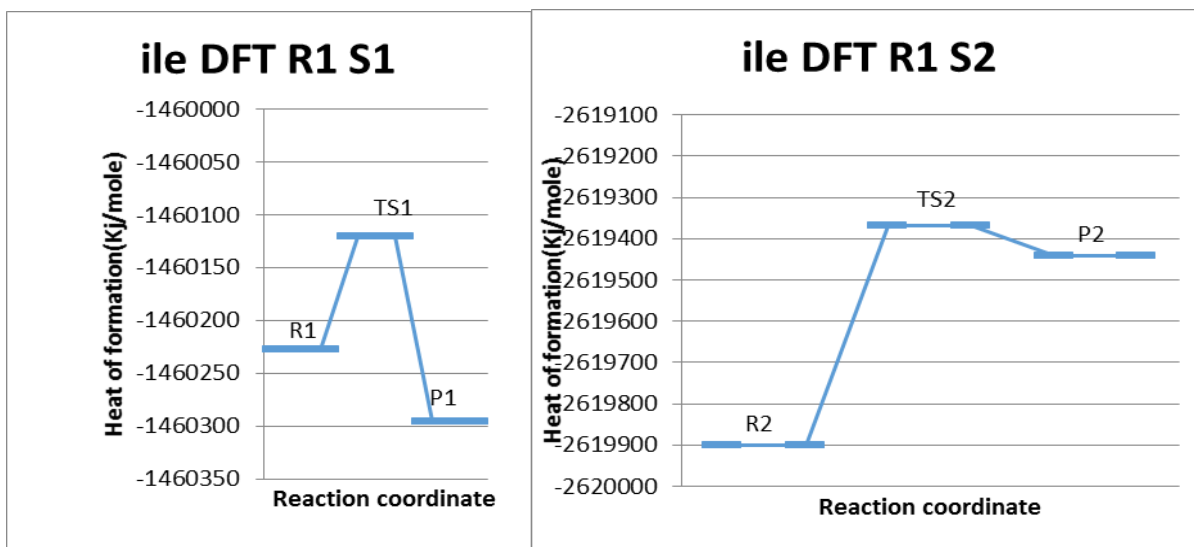
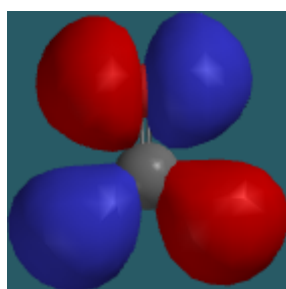


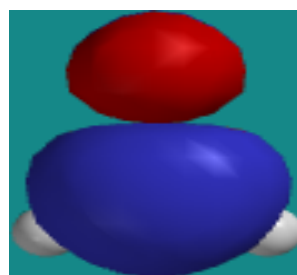
Figure 3. Energy profile of reaction of Isoleucine-Formaldehyde system DFT method

Molecular orbitals and frontier electron density

The plot of the HOMO and LUMO of formaldehyde were shown in Figure 4. It could be seen that the frontier electron densities of both HOMO and LUMO orbital were on the four atoms that formed the formaldehyde molecule. The HOMO and LUMO level determine how a chemical molecule shares its valence electron in the occupied molecular orbital and donates (or accept) electrons to ligands, and thus the HOMO-LUMO energy gap has been suggested as an indicator of the stability of molecules (Lynam *et al.*, 1998). The HOMO-LUMO energy gap, ($\Delta E = -6.16\text{eV}$) is higher than the gaps in the L-isoleucine molecules substrate reactants. This showed that compared to the other reactants, the formaldehyde molecule would be a better electron acceptor than the others, according to Fukui's frontier orbital theory (Fukui, 1982; Ya-Yin *et al.*, 2002). The plot of the HOMO and LUMO of isoleucine were also shown in Figure 5. The HOMO-LUMO energy gap, ($\Delta E = -6.92$), of the isoleucine was smaller than that of formaldehyde molecules. This would make isoleucine as the electron donor.

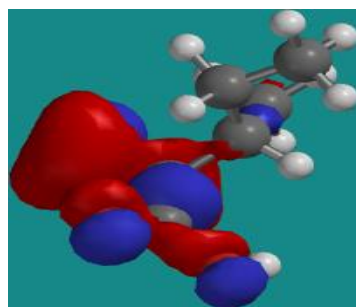


HOMO (-7.31eV)

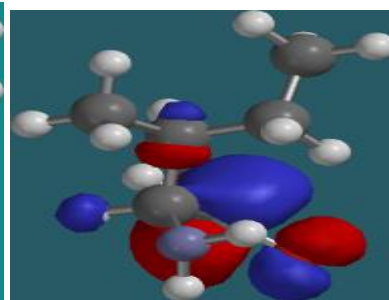


LUMO(-1.15eV)

Figure 4. HOMO and LUMO of Formaldehyde molecule



HOMO (-6.77eV)



LUMO (0.15eV)

Figure 5. HOMO and LUMO of L-Isoleucine molecule

Charge distribution

The optimized isoleucine molecule was shown in figure 6. The conformation with the lowest energy was optimized and calculated using the PM3 and DFT methods. The molecular information such as net charges, electron density, and molecular orbital energies obtained. The identity of all hydrogen atoms of the conformer compound were labeled as in Figure 6. The isoleucine molecule was optimized. The lowest conformer was optimized and calculated by using the PM3 and DFT methods. Molecular information such as identity of hydrogen, bond type, bond length and exposed surface area were obtained and presented in Table 2.

There are thirteen hydrogen atoms present in the molecule, ten (HA, 1HB, 1HD1, 2HD1, 3HD1, 1HG1, 2HG1, 1HG2, 2HG2 and 3HG2) were bonded to carbon atoms while H and H1 were bonded to nitrogen atoms and H2 bonded to oxygen atom. The C-H bond lengths are all 1.096 Å, while the N-H bond lengths are 1.012 Å and that of O-H is 0.958 Å. This indicates that C-H is longer than that of N-H and O-H is the shortest among the bond type. However, the hydrogen atoms bonded to the nitrogen has more exposed surface area (7.89-8.86 Å²) than that attached to carbon (5.45 Å²) and oxygen (6.63 Å²). The two hydrogen's attached to the nitrogen have large exposed surface area and are more active and can react more easily.

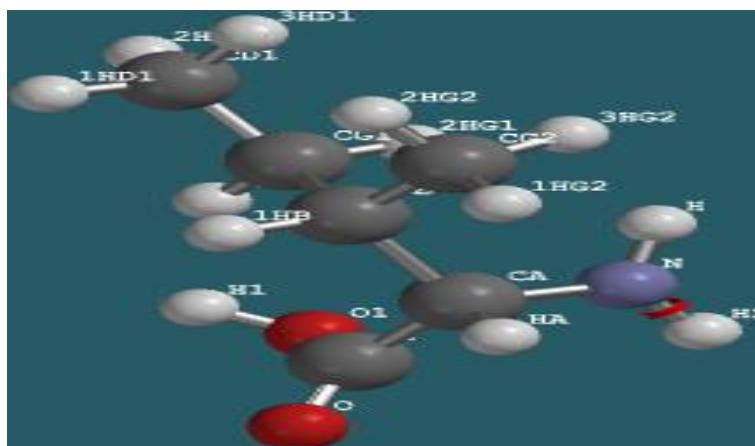


Figure 6. Labeled structure of L-Isoleucine

Table 2. Molecular information of the optimized isoleucine molecule

Hydrogen identity	Bond type	Bond length (Å)	Exposed surface area (Å ²)
H	N-H	1.012	8.86
H1	N-H	1.012	7.89
H2	O-H	0.958	6.63
HA	C-H	1.096	5.45
1HB	C-H	1.096	5.45
1HD1	C-H	1.096	5.45
2HD1	C-H	1.096	5.45
3HD1	C-H	1.096	5.45
1HG1	C-H	1.096	5.22
2HG1	C-H	1.096	3.62
1HG2	C-H	1.096	5.45
2HG2	C-H	1.096	5.40
3HG2	C-H	1.096	5.45

Proposal of plausible mechanism of reaction

The reaction took place in two steps, in the step 1 amino acid reacts with the formaldehyde via the terminal amino group to form a methylol intermediate. In the step 2 the amino acid (either same or different) then reacts with the intermediate methylol to form a methylene bridge (Conaway, 1996) and water as products which confirms the work of Gustavson (1956). In the present work we restrict ourselves to use one amino acid to model the mechanism. As shown in Figure1 isoleucine N attacks the C1 carbonyl

centre of formaldehyde and a bonding occurred between the N and C1 with a concomitant transfer of hydrogen atom from N of isoleucine to a carbonyl O1 of the formaldehyde during the geometry optimization/energy minimization procedures to form intermediate1 which later react with another isoleucine to form the product which is a methylene cross-linkage. The isoleucine N attack the methylol carbon to form a bond between N and the methylol carbon with a transfer of hydrogen from N to the oxygen attach to the methylol carbon to form a transition state (TS2) in which a molecules of water is lost to form the Methylenediisoleucine as a final product.

Heat of formation and rate constant calculation

The enthalpy of reaction was calculated from the Spartan'08 software package default to calculate the heat of formation at standard temperature of 25° C (298.15K) and the appropriate sums and differences taken as given in equation 1 below.

$$\Delta H_f^\circ (2985.1K) = \Sigma \Delta H_f^\circ \text{prod} (298.15K) - \Sigma \Delta H_f^\circ \text{react} (298.15K) \quad (1)$$

The computed enthalpies of the reaction of formaldehyde and L-isoleucine at standard temperature and pressure were -78.79 kJ/mol and -39.14kJ/mol for PM3 and DFT respectively.

Other thermodynamic activated parameters were calculated at standard conditions ΔS° and ΔG° , for the PM3 and DFT studies.

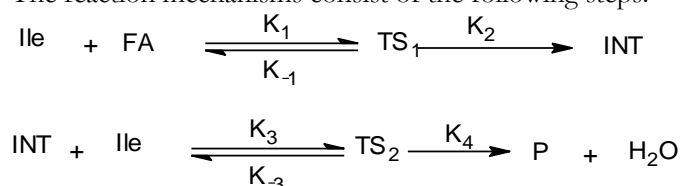
The rate constant (k) was computed by using the equation 2 below.

$$k (298.15K) = \frac{k_B T}{h c^\circ} e^{-\frac{\Delta G^\ddagger}{RT}} \quad (2)$$

The rate constant of the proposed mechanism in this study are $k_1 = 2.659 \times 10^{11} \text{ dm}^3 \text{ mol}^{-1} \text{ sec}^{-1}$ for PM3 and for the DFT the value were found to be $k_1 = 2.474 \times 10^{210} \text{ dm}^3 \text{ mol}^{-1} \text{ sec}^{-1}$. This shows that the DFT method is more favoured kinetically.

Kinetics of the reaction

The reaction mechanisms consist of the following steps:



For step 1 the rate equation is

$$R_f = k_1 [\text{Ile}] [\text{FA}]$$

$$R_b = k_{-1} [\text{TS}_1]$$

At equilibrium $R_f = R_b$

$$\text{Therefore, } k_1 [\text{Ile}] [\text{FA}] = k_{-1} [\text{TS}_1]$$

For step 2

$$R_2 = k_2 [\text{TS}_1]$$

For step 3

$$R_f = k_3 [\text{INT}] [\text{Ile}]$$

$$R_b = k_{-3} [\text{TS}_2]$$

At equilibrium, $R_f = R_b$

$$k_3 [\text{INT}] [\text{Ile}] = k_{-3} [\text{TS}_2]$$

For step 4

$$R_4 = k_4 [\text{TS}_2]$$

Applying steady-state approximation

For [TS1]

$$k_1 [\text{Ile}] [\text{FA}] - k_{-1} [\text{TS}_1] - k_2 [\text{TS}_1]$$

$$[\text{TS}_1] = \frac{k_1 [\text{A}] [\text{F}]}{k_{-1} + k_2}$$

For [INT]

$$k_2 [\text{TS}_1] - k_3 [\text{A}] [\text{INT}] = 0$$

$$[\text{INT}] = \frac{k_2 [\text{TS}_1]}{k_3 [\text{Ile}]}$$

$$= \frac{K_2 K_1 [\text{A}] [\text{F}]}{K_3 [\text{A}] (K_{-1} + K_2)}$$

For [TS2]

$$k_3 [\text{Ile}] [\text{INT}] - k_4 [\text{TS}_2] = 0$$

$$[\text{TS}_2] = \frac{k_3 [\text{Ile}] [\text{I}]}{k_4}$$

$$[\text{TS}_2] = \frac{k_1 [\text{Ile}]}{k_4} \left(\frac{K_2 K_1 [\text{Ile}] [\text{FA}]}{K_3 [\text{Ile}] (K_{-1} + K_2)} \right)$$

But step 4 is the slowest and hence rate determining step

$$R_4 = k_4 [\text{TS}_2]$$

Substitute the value of [TS2]

$$R_4 = k_4 \frac{K_2 K_1 [\text{A}] [\text{FA}]}{K_3 [\text{Ile}] (K_{-1} + K_2)} [\text{Ile}] [\text{FA}]$$

But $\frac{K_2}{K_{-1}}$ is the equilibrium constant for step 1, so

$$\frac{K_2}{K_{-1}} = K_{eq}$$

Therefore, $R_4 = K_{eq} k_2 [\text{Ile}] [\text{FA}]$

Let $k = k_{eq} k_2$

$$R_4 = k [\text{Ile}] [\text{FA}] = 0.22263 \times 10^3 \text{ dm}^3 \text{ mol}^{-1} \text{ sec}^{-1} [\text{Ile}] [\text{FA}]$$

$$k = 0.22263 \times 10^3 \text{ dm}^3 \text{ mol}^{-1} \text{ sec}^{-1}$$

Conclusions

The mechanism of reaction of formaldehyde with L-isoleucine was investigated by using computational method using Spartan'08 V1.2.0. Software Package PM3 and DFT calculation results. The reaction is found to be nucleophilic in which the lone pair of nitrogen attacks the carbonyl carbon of formaldehyde to form isoleucine-formaldehyde adduct that combines with another isoleucine to form a cross-linking (methylene bridge) compound and water as products. The reaction was found to be exothermic and second order overall.

References

- Conaway, C.C., Whysner, J., Verna, L.K. and William, G.M. 1996. Formaldehyde mechanistic data and risk assessment: Endogenous protections from DNA adduct formation. *Pharmacology and Therapeutics*, 71(1-2): 9-55.
- Dewar, M.J.S., Zeobisch, E.G., Healy, E.F. and Stewart, J.J.P. 1985. AM1: A New General purpose quantum mechanical molecular model. *Journal of the American Chemical Society*, 107: 3902-3909.
- Elder, T.J., and Worley, S.D. 1984. In: Wu, C.S., Neely, W.C. and Worley, S.D. 1991. Semi empirical Theoretical Study of the Molecular Interaction of Cocaine with Biological Substrate Glycine. *Journal of Computational Chemistry*, 12(7): 862-867.
- Fukui, K. 1982. Chemical reactivity theory-its pragmatism and beyond. *Pure and Applied Chemistry*, 54(10): 1825-1836.
- Gurel, A., Coskun, O., Armutcu, F., Kanter, M. and Ozen, O.A. 2005. Vitamin E against oxidative damaged caused by formaldehyde in frontal cortex and hippocampus: Biochemical and histological studies. *Journal of Chemical Neuroanatomy*, 29: 173-178.
- Hussein, A., Yamrot, K. and Asrat, A. 2014. Toxic effect of formaldehyde on the nervous system. *International Journal of Anatomy and Physiology*, 3(3): 50-59.
- Kalkanis, G. H. and Shields, G.C. 1991. AM1 and PM3 calculations of the potential energy surface for CH₂OH reactions with NO and NO₂. *The Journal of Physical Chemistry*, 95(13): 5085-5089.
- Toews, J., Rogalski, J.C., Clark, T.J. and Kast, J. 2008. Mass spectroscopic identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions. *Analytica Chimica Acta* 618(2): 168-183.
- Kerns, W.D., Pavkov, K.,L., Donofrio, D.J., Gralla, E.J. and Swenberg, J.A. 1983. Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Research*, 43(9): 4382.-4392.
- Lynam, M., Kutty, M., Damborsky, J., Koca, J. and Adriaens, P. 1998. Molecular orbital calculations to describe microbial reductive dechlorination of polychlorinated dioxins. *Environmental Toxicology and Chemistry*, 17: 988-997.
- Metz, B., Kersten, G.F.A. and Hoogerhout, P. 2004. Identification of formaldehyde-induced modification in proteins: reaction with model peptides. *Journal of Biological Chemistry*, 279: 6235-43.
- Miller, S.L. 1953. A production of amino acids under possible primitive earth conditions. *Science*, 117(3046): 528-529.
- Mshelia, M.S., Iyun, J.F, Uzairu, A. and Idris, S.O. 2010. Kinetics and mechanism of the oxidation of hydrazine dihydrochloride by aqueous iodine. *Journal of American Science*, 6(9): 293-296.
- Mohammadi, S. 2014. Protective Effects of N-Acetyl cysteine against formaldehyde-induced neuronal damage in cerebellum of mice. *Pharmaceutical Sciences*, 20: 61-65.
- Robert, G. 2005. Formaldehyde: overview of current issues and challenges for the future. *International Nonwoven Technical Conference*, September 19-22, 2005, St. Louis, Missouri.
- Stewart, J.J.P. 1989. Optimization of parameters for semi empirical method II-Applications. *Journal of Computational Chemistry*, 10(2): 221-264.
- Thomas, E. and Philip, R. 2006. *Physical chemistry*. Pearson Education, Inc., (Benjamin Cummings) Sansome St. San Francisco, pp 924-925.
- Thrasher, J.D., Broughton, A. and Madison, R. 1990. Immune activation and auto antibodies in humans with long term inhalation exposure to formaldehyde. *Archives of Environmental Health*, 45(4): 217-223.

- Toxic and Exposure Assessment for Children's Health (TEACH), 2007. Formaldehyde; TEACH Chemical Summary. Last accessed March, 15, 2011 at <<http://www.epa.gov/teach>.
- Ya-Yin, F., Shen, R., Sun, H., Yuan, M., Liu, J., Li, Z.M. and Lai, C.M. 2002. Mechanism analysis of catalytic hydrogenation of 3-anilinomethylidene-6-alkyl-5,6-2H-dihydropyran-2,4-diones. *Journal of Molecular Structure (Theochem)*, 578: 71-78.
- Zhang, P., Li, X., Ren, Y., Wang, C., Zou, W. and Tang, X., 2014. Paraoxonase-1 mediates hydrogen sulfide-induced protection against formaldehyde neurotoxicity. *Progress in Biochemistry and Biophysics*, 41(5): 487-496.