**Journal of Basrah Researches ((Sciences)) Vol. 35, No.3, 15 June ((2009))**

# Theoretical Study of the Ground State Intramolecular Proton Transfer in Cytidines

**Bahjat A. Saeed**

*Department of Chemistry, College of Education, University of Basrah, Basrah, Iraq.*

*e-mail: bas_chem_dep@yahoo.com*

**ISSN -1817 -2695**

**((Received 22/3/2009 , Accepted 25/5/2009))**

## Abstract

Four isomers of 1-methyl-N-methoxycytosine and their mutual interconversion were studied theoretically at the B3LYP level using variety of basis sets and a different number of polarization and diffuse functions. It was found that the imino form of the studied molecule is the most stable form within the tautomeric mixture. Transition state of interconversion were studied at the B3LYP/6-31G(d,p) level of theory . The study has shown the exclusion the possibility of direct proton transfer in the gas phase due to the strain in the four-centered transition state, in which the proton being transferred is forced to come close to the positively charged carbon atom at the opposite corner of the four-membered ring.

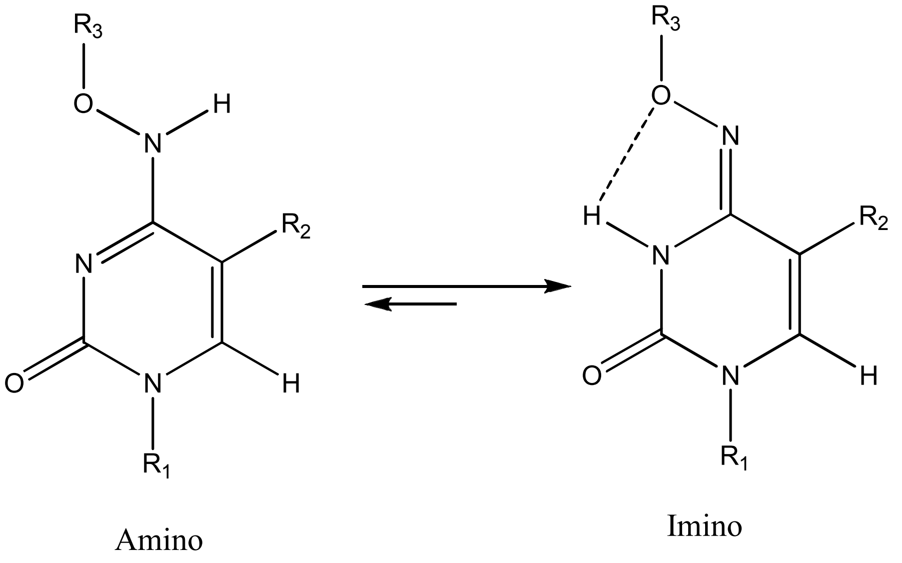
**Keywords**: cytidine; tautomerism; proton transfer; density functional study.

## Introduction

Cytidine, the pyrimidine nucleoside that is composed of cytosine linked to D-ribose is known to show deamination. Cytidine deamination was discovered as a mechanism of editing mRNA transcripts in eukaryotes[1-5]. N-hydroxy(*1*) and N-methoxycytidines(*2*) are the most interesting of promutagenic base

analogues. They are the products of reaction of cytidine with the potent mutagen $NH_2OH$. *In vivo* these analogues lead to point transition, C-T (U) [6, 7]. *In vitro* transcription they exhibit dual functional activity, reflecting the tautomerism of these promutagenic analogues and their base pairing properties [8-10]. They exhibit two tautomeric forms, amino = imino, the equilibrium being dependent on the solvent with predominance of the imino form in aqueous medium [8,11,12]. In less polar medium the equilibrium shifts

towards the imino species. Spectroscopic methods [13-15] point to predominance, and probably exclusively, of the imino species in weakly polar solvents as well as in the gas phase. In the solid state the imino form is stabilized by intramolecular hydrogen bond. The promutagenic activity of these is reflected in the tautomerism of these promutagenic analogues [16 ] (Scheme 1).



*1*, $R_1$= ribose, $R_2$= $R_3$= H
*2*, $R_1$= ribose, $R_2$= H, $R_3$= $CH_3$
*3*, $R_1$= $CH_3$, $R_2$= H, $R_3$= $CH_3$

Scheme 1. Tautomerism in cimitidines.

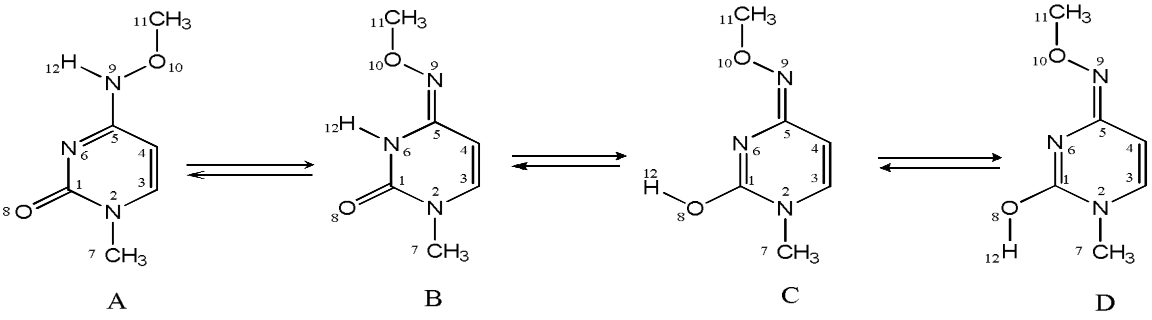
The aim of this paper is to theoretically study the mechanism of tautomeric equilibrium using the DFT hybrid functional B3LYP. Their structurally analogue1-methyl-N-methoxycytosine(*3*) ($R_1$=$CH_3$, $R_2$=H and $R_3$=$CH_3$) will be taken as representative and the calculations will be done on this molecule for saving computation time.

## Computational method

The geometries of the four isomers were optimized at the B3LYP level using several basis sets with no symmetry restrictions. Subsequent frequency calculations were carried out in order to establish the structures as local or global minima (with no imaginary frequencies in their vibration spectra) at B3LYP/6-31G(d,p) level of theory. The QST2 was used to locate each transition state using B3LYP/6-31G(d,p) level of theory. All calculations were performed with Gaussian 03 program package [17].

## Results and discussion

Four isomers and their mutual interconversions were studied. They are illustrated in Scheme 2.



Scheme 2. Tautomers and tautomeric interconversions of 1-methyl-$N^4$-methoxycytocine (the numeration does not follow the IUPAC nomenclature).

The structural properties of the isomers are shown in Fig. 1 and in Table 1.

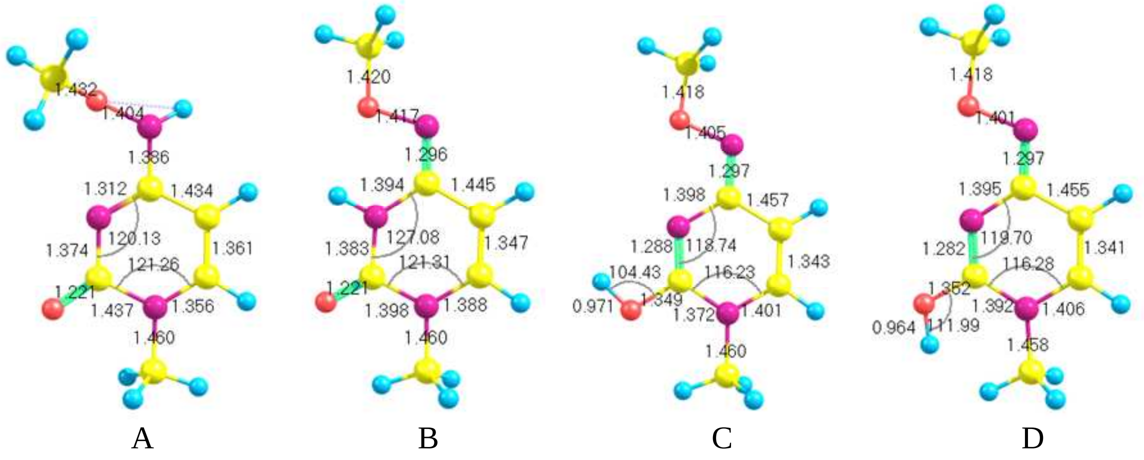


**Fig. 1. The structural properties of the isomers.**

Comparing the bond lengths of the tautomers and the transition states supports the general rout for the tautomerization. It is clear from Fig. 1 that the calculated C(5)-N(9) bond lengths in tautomers A and B are 1.386 and 1.296 Å respectively reflecting their single and double bond characters, while the length of the same bond in $TS_{AB}$ is 1.328 Å (Table 3) which represents an intermediate length between the bond lengths in tautomers A and B. This is also obvious for the bond C(1)-N(6) in the tautomers B and C and $TS_{BC}$ since the calculated bond lengths are 1.383, 1.288 and 1.331 Å for B, C and $TS_{BC}$ respectively. The same situation is also clear for the C(1)-O(8) on going from tautomer B to C through $TS_{BC}$ since the calculated bond lengths in foregoing species are 1.221, 1.349 and 1.288 Å respectively reflecting the change from double to single bond through a partial double bond.

Except for the isomer A all other isomers have planar structures as could be seen in Table 1.

**Table 1. Dihedral angles (°) of the four tautomers.**

| Dihedral angle | A | B | C | D |
|---|---|---|---|---|
| C(11)O(10)N(9)C(5) | 95.2 | 180.0 | 180.0 | 180.0 |
| O(10)N(9)C(5)N(6) | -25.3 | 0.0 | 0.0 | 0.0 |
| O(10)N(9)C(5)C(4) | 158.8 | 180.0 | 180.0 | 180.0 |
| N(9)C(4)C(5)C(3) | 176.1 | -180.0 | -180.0 | 180.0 |
| C(5)C(4)C(3)N(2) | -0.2 | 0.0 | 0.0 | 0.0 |
| C(4)C(3)N(2)C(7) | -179.8 | 180.0 | 180.0 | 180.0 |
| O(8)C(1)N(2)C(7) | -0.2 | 0.0 | 0.0 | 0.0 |
| H(12)O(10)N(9)C(5) | 29.9 | 0.0 | 0.0 | 0.0 |

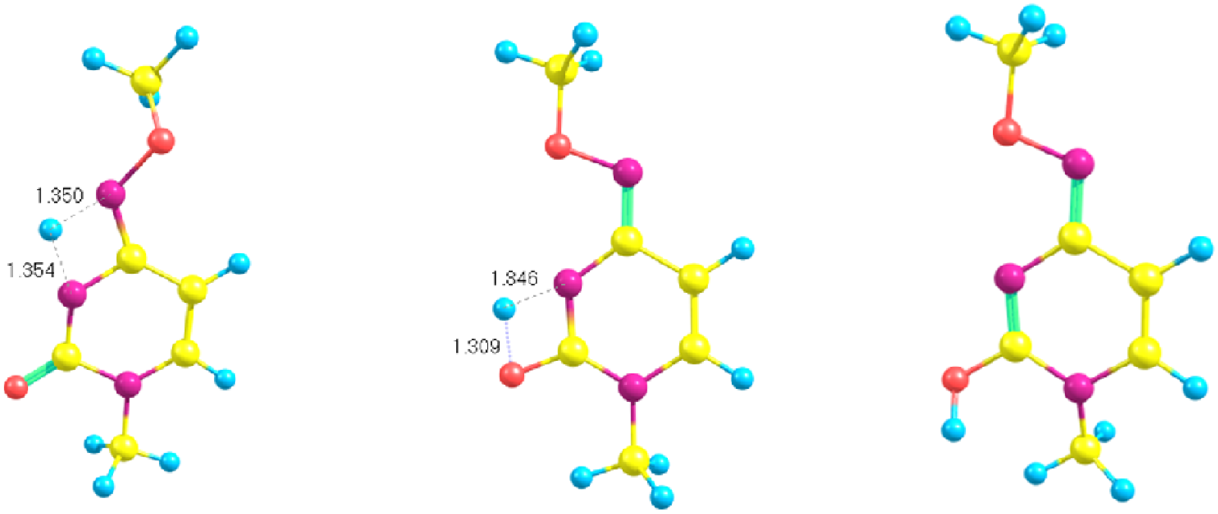
The calculated energies and the relative energies with respect to the stable isomer are given in Table 2. From Table 2 the isomer B was found to be the most stable isomer while the isomer D was found to be the most unstable one (B3LYP/6-31G(d,p)) The stability of isomer B may be expected to rise in part due to the intramolecular hydrogen bond N(6)-H….O-N(9). In isomer D the steric hindrance which occurs as a result of the repulsion between the OH hydrogen and the hydrogens of the methyl group may be responsible for its higher energy.

**Table 2. Energies (E, a.u.) and relative energies ($E_{rel}$,kJ $mol^{-1}$) of isomers.**

| Method/basis set | E | | | |
|---|---|---|---|---|
| | A | B | C | D |
| B3LYP/6-31G(d) | -548.71062272 | -548.73493243 | -548.68449044 | -548.66465329 |
| B3LYP/6-31G(d,p) | -548.7160533 | -548.7349924 | -548.7011827 | -548.6816891 |
| B3LYP/6-31+G(d,p) | -548.73965562 | -548.75870334 | -548.70653866 | -548.7065276 |
| B3LYP/6-31++G(d,p) | -548.73988262 | -548.75892178 | -548.70681890 | -548.70681860 |
| B3LYP/6-311G(d,p) | -548.84995458 | -548.86961878- | -548.83517616 | -548.81567789 |
| B3LYP/6-311+G(d,p) | -548.86287582 | -548.8820422 | -548.84857037 | -548.82972942 |
| B3LYP/6-311++G(d,p) | -548.8690018 | -548.88221178 | -548.82961150 | -548.82911250 |
| | | $E_{rel}$ | | |
| B3LYP/6-31G(d) | 87.5 | 0.0 | 132.4 | 184.5 |
| B3LYP/6-31G(d,p) | 49.6 | 0.0 | 88.6 | 139.8 |
| B3LYP/6-31+G(d,p) | 50.0 | 0.0 | 137.0 | 137.0 |
| B3LYP/6-31++G(d,p) | 50.0 | 0.0 | 136.8 | 136.8 |
| B3LYP/6-311G(d,p) | 51.6 | 0.0 | 90.4 | 142.6 |
| B3LYP/6-311+G(d,p) | 50.3 | 0.0 | 78.9 | 137.3 |
| B3LYP/6-311++G(d,p) | 50.7 | 0.0 | 138.1 | 139.4 |

As could be seen from Scheme 2 the direct transformation of isomer A to C and D is impossible, because it includes simultaneous proton exchange between atoms N(9) and N(6) as well as between N(6) and O(8), since the probability of simultaneous two events is very small [18]. In addition, the isomer transformations A-B and B-C are tautomeric conversions, whereas the transformation C-D is a conformation conversion.

The transition states are shown in Fig. 2 and their structural properties are listed in Table 3.



TS$_{AB}$ TS$_{BC}$ TS$_{CD}$

**Fig. 2. Transition states of the isomer transformation (B3LYP/6-31G(d,p).**

**Table 3. Calculated structures of the transition states calculated by B3LYP/6-31G(d,p).**

| Parameters | TSAB | TSBC | TSCD |
|---|---|---|---|
| Bonds (Å ) | | | |
| C(1)C(3) | 1.443 | 1.351 | 1.393 |
| N(2)C(3) | 1.363 | 1.408 | 1.406 |
| C(3)C(4) | 1.365 | 1.346 | 1.342 |
| C(4)C(5) | 1.420 | 1.462 | 1.456 |
| C(5)N(6) | 1.363 | 1.392 | 1.399 |
| N(6)C(1) | 1.392 | 1.331 | 1.280 |
| C(1)O(8) | 1.219 | 1.288 | 1.329 |
| N(2)C(7) | 1.462 | 1.429 | 1.297 |
| C(5)N(9) | 1.328 | 1.293 | 1.297 |
| N(6)H(12) | 1.354 | 1.346 | ------ |
| N(9)H(12) | 1.350 | ------ | ------ |
| N(9)O(10) | 1.400 | 1.407 | 1.399 |
| O(10)C(11) | 1.426 | 1.419 | 1.419 |
| Bond angles (°) | | | |
| C(1)N(2)C(3) | 122.3 | 116.1 | 116.0 |
| N(2)C(3)C(4) | 123.2 | 121.8 | 121.3 |
| C(3)C(4)C(5) | 114.5 | 120.9 | 119.1 |
| C(4)C(4)N6) | 122.0 | 114.5 | 118.2 |
| C(5)N(6)C(1) | 123.6 | 121.1 | 119.1 |
| N(6)C(1)N(2) | 113.3 | 122.1 | 126.1 |
| C(1)N(2)C(7) | 115.7 | 124.7 | 123.5 |
| N(6)C(1)O(8) | 127.2 | 107.7 | 117.7 |
| N(6)C(5)N(9) | 101.6 | 126.5 | 124.5 |
| C(5)N(9)O(10) | 118.5 | 110.4 | 112.0 |
| N(9)O(10)C(11) | 109.5 | 108.3 | 108.3 |
| N(6)H(12N(9) | 100.8 | 105.6 | ------ |
| C(5)N(6)H(12) | 78.0 | 72.0 | ------ |
| C(5)N(9)H(12) | 79.4 | 74.6 | ------ |
| O(8)C(1)N(2) | ------ | 127.6 | ------ |
| C(1)O(8)H(12) | ------ | ------ | 110.1 |

The imaginary frequencies in the transition states were -1814 (671), -1861(760)and -110(2) for TSAB, TSBC and TSCD respectively. Imaginary frequencies in TSAB and TSBC correspond to parallel mode (the N-H-O stretching) along the reaction coordinates. In TSCD it corresponds to both O(8)-H(12) and N(2)-C(7) out-of-plane bending. The tautomeric conversions TSAB and TSBC are passing through planar transition states (with respect to the cytosine ring) - the proton transfer occurs in the molecular plane. Mullikan charges are listed in Table 4. It is found that the nitrogen atoms N(10) and N(6) from which the hydrogen migrates in the tautomeric conversion states TSAB and TSBC respectively, become more negative in the transition states (their charges are -0.321 and -0.636 in the isomers A and B). Meanwhile the migrating hydrogen (H12) becomes more positively charged, due to its state of simultaneously bonding to two electronegative atoms, N(9) and N(6) in TSAB and N(6) and O(8) in TSBC. In these cases their charges changed from 0.274 and 0.298 in the isomers A and B to 0.313 and 0.350. The opposite atoms H(12) and C(5) and H(12) and C(1) in the four- membered rings in the transition states $TS_{AB}$ and $TS_{BC}$ are positively charged which means that the high energies in the first two transition states are not only due to the strain in the four-membered ring, but also because the interatomic repulsion between the two positively charged atoms which are forced to approach within 1.710 and 1.574 Å of each other in TSAB and TSBC respectively.

**Table 4. Mullikan charges for the transition states calculated at B3LYP/6-31G(d,p) level.**

| Atom | $TS_{AB}$ | $TS_{BC}$ | $TS_{CD}$ |
|---|---|---|---|
| C(1) | 0.792 | 0.810 | 0.682 |
| N(2) | -0.491 | -0.497 | -0.496 |
| C(3) | 0.191 | 0.106 | 0.104 |
| C(4) | -0.139 | -0.131 | -0.112 |
| C(5) | 0.570 | 0.525 | 0.469 |
| N(6) | -0.642 | -0.644 | -0.484 |
| C(7) | -0.166 | -0194 | -0.184 |
| O(8) | -0.515 | -0.586 | -0.507 |
| N(9) | -0.380 | -0.302 | -0.299 |
| O(10) | -0.380 | -0.391 | -0.373 |
| C(11) | -0.050 | -0.031 | -0.036 |
| H(12) | 0.313 | 0.350 | 0.321 |

Fig.3 shows the energy diagrams of the isomer conversions which contains also the values of some thermodynamic parameters of conversions showing TSBC has the highest activation energy (191 kJ/mol) which may be the result of possible intermolecular hydrogen bond and thus for the transition to be occurred it needs to break this bond. As it is seen the equilibrium shifted towards isomer B and the transformation of D into B passes through isomer C through an energy barrier of 47.2 kJ/mol, then isomer C transforms into B through an energy barrier of 174 kJ/mol. (Fig. 4).

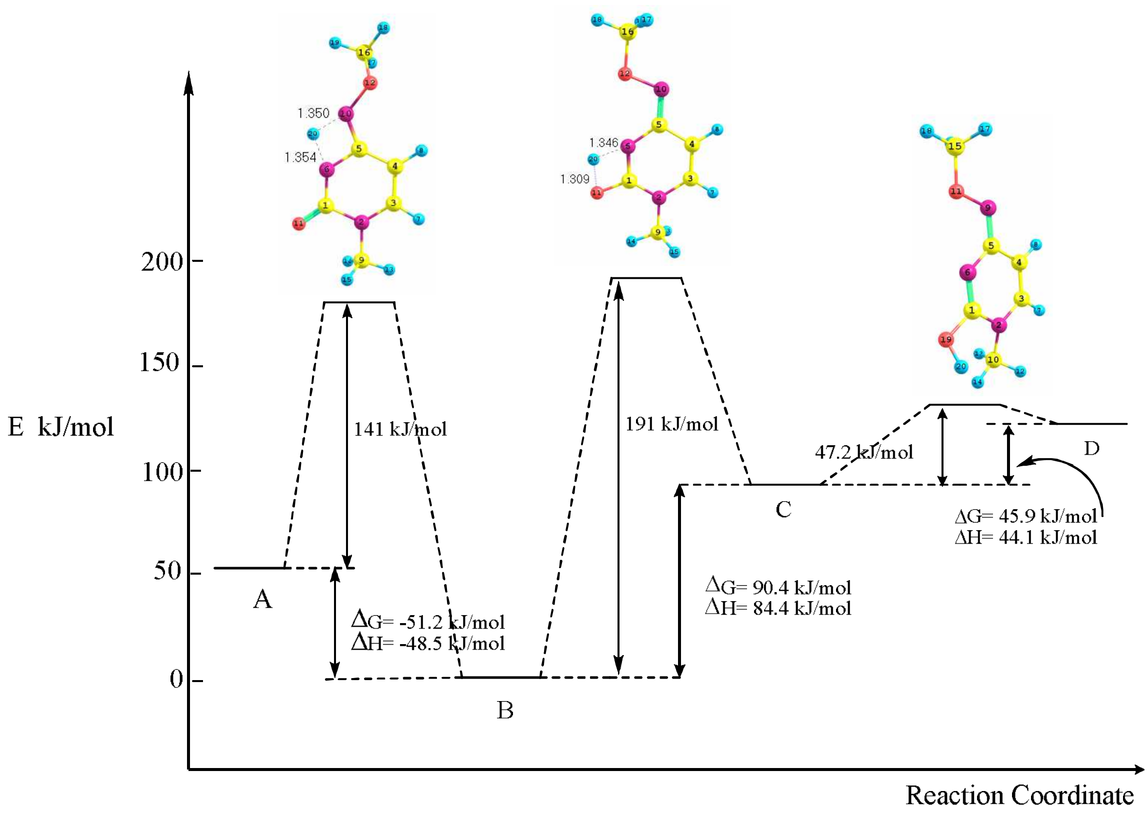


Fig. 3. Potential energy surfaces curves of the isomer transformations.

## References

[1] M. *J.* Cascalho, *Immunol*. 172 (2004) 6518.

[2] X. Wu,; P. Geraldes,; J. L Platt,; Cascalho, M. *J. Immunol*. 174 (2005) 934.

[3] K. N. Bishop, R. K. Holmes, M. H. Malim *J. Virol.* , 80 (2006**)** 8450.

[4] G. S. C. Dance, M. P. Sowden, Y. Yang and H. C. Smith *Nucleic Acid Res*. 28 (2000) 424.

[5] A. Durandy, E. J. Immunol. 33 (2003) 2069.

[6] E. I. Budowsky *Prog. Nucleic Acid Res. Mol. Biol.* 16 (1976) 125.

[7] F. Marfey and E. Robinson *Mutation Res*. 86 (1981) 155.

[8] E. I. Budowsky, E. D. Sverdlov and T. N. Spasokukotskaya *FEBS Lett*, 17 (1971) 336.

[9] R. A. Flavell, D. L. Sabo, E. F. Bandle and C. Weissman *J. Mol. Biol*. 255 **(**1974) 89.

[10] B. Singer, H. Fraenkel-Conrat, L. Abbott and S. J. Spengler *Nucleic Acid Res*. 12 **(**1984) 4609.

[11] D. M. Brown, M. J. E. Hewlins and P. Schell *J. Chem. Soc. C,* (1968) 1925.

[12] A. Psoda, B. Kierdaszuk, A. Pohorille, M. Geller, J. T. Kusmierek and D. Shugar *Int. J. Quant. Chem*. 20 (1981) 543.

[13] B. Kierdaszuk and D. Shugar *Biophys. Chem.* 17 **(**1983) 285.

[14] K. Kulinska, A. Psoda and D. Shugar *Acta Biochim. Polon.* 26 **(**1979) 145.

[15] I. G. Birnbaum, T. Kulikowski and D. Shugar *Can. J. Biochem.* 57 **(**1979) 308.

[16] D. Shugar and B. Kierdaszuk *Proc. Int. Sump. Biol. Struct. Interactions, Suppl. J. Biosci*., 8 (1985) 657.

[17] Gaussian 03. Gaussian, Inc. Carnegie Office Park, Building 6, Pittsburgh, PA 15106 USA.

[18] V. B. delchev and N. V. Nenkova *Acta Chim. Slov*. 55 **(**2008) 132.

# دراسة نظرية للانتقال الظمني للبروتون في الحالة المستقرة للسايتيدينات

**بهجت علي سعيد**

***جامعة البصرة – كلية التربية – قسم الكيمياء***

## الخلاصة

تمت في هذا البحث دراسة نظرية لاربعة ايزوميرات للمركب 1- مثيل – ن – هيدروكسي سايتوسين باستخدام نظرية الكثافة الدالية و عند مستويات مختلفة من الدوال الاساسية و باستخدام دوال انتشار و استقطاب مختلفة.

لقد بينت الدراسة ان توتومر الايمينو هو الشكل الاكثر استقرارا من بين جميع التوتومرات التي يمكن ان تتواجد بها الجزيئة. لقد درست الحالة الانتقالية للتحول التوتومري باستخدام نظرية الكثافة الدالية و بينت الدراسة استبعاد امكانية حصول انتقال بروتوني مباشر في الحالة الغازية كنتيجة للتوتر العالي في الحالة الانتقالية رباعية الحلقة و التي يجبر فيها البروتون المنتقل للاقتراب من ذرة الكاربون سالبة الشحنة الواقعة عند الزاوية المقابلة من الحلقة الرباعية.