

文章编号: 1007-4627(2007)02-0098-05

氢键质子的运动与胞嘧啶的变异几率*

赵振民¹, 张启仁¹, 高春媛¹, 卓益忠²

(1 北京大学物理学院, 北京 100871;

2 中国原子能科学研究院, 北京 102413)

摘要: 运用 Gaussian98 MP2/6-311G(d, p) 级的从头计算法计算了 DNA 分子中胞嘧啶碱基中 H13 质子的势能曲面。发现有两个势阱: 其中一个对应正常的胞嘧啶, 而另一个则对应它的顺式亚胺式互变异构体。质子在这两个势阱的束缚对热扰动是稳定的。质子在势阱中像单摆一样绕着离它最近的氮原子摆动, 还可以远离氮原子与其它碱基形成氢键。估算出的 H13 质子从一个势阱到另一个势阱的穿透几率表明质子呆在其中一个势阱中的寿命大约是 6×10^2 a, 这使胞嘧啶和它的顺式亚胺式互变异构体在室温气相实验中远不能达到热平衡。对这些结果的生物学意义作了讨论。

关键词: 氢键质子; 势能曲面; 隧道效应

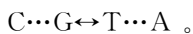
中图分类号: Q319⁺.3; Q526; O641.1⁺1

文献标识码: A

1 引言

氢键质子穿透分隔正常态和病态的势垒的隧道效应会引起 DNA 碱基改变, Lowdin 于 40 多年前据此提出了一个研究突变、衰老和癌症形成的方案^[1]。其后一段时间内, 限于理论化学方法和计算能力的不足, 这个诱人的方案未能实现。而今从头计算理论以及它的软件包都日益成熟, 也有了高性能的计算机, 为实现这个方案提供了条件。沿着这个方向已经有了许多工作。文献[2]—[15]是一些早期的例子。然而, 我们还未看到关于氢键质子的势能曲面及其相应运动方面从头计算结果的报道。

DNA 分子中的 4 个碱基(A, T, C, G)通常处于它们的正常型, 但是偶尔它们也会以小几率突变到它们的不正常的顺式亚胺式/烯醇式构型。也就是 $A \leftrightarrow A_{\text{imino}}$, $T \leftrightarrow T_{\text{enol}}$, $G \leftrightarrow G_{\text{enol}}$ 和 $C \leftrightarrow C_{\text{imino}}$ 突变。这些顺式亚胺式/烯醇式互变异构体的配对方式是 $A_{\text{imino}} \cdots C$, $A \cdots C_{\text{imino}}$, $G_{\text{enol}} \cdots T$ 和 $G \cdots T_{\text{enol}}$ 。这就使得 DNA 在它的两次复制后可能改变碱基配对



例如, 转变 $C \rightarrow C_{\text{imino}}$ 会把配对 $C \cdots G$ 变成 $C_{\text{imino}} G$, 而配对 $A \cdots C_{\text{imino}}$ 在第一次复制中产生 $A \cdots C_{\text{imino}}$ 和 G

$\cdots C$ 配对。进一步复制又会导致 $T \cdots A$, $C_{\text{imino}} \cdots A$, $C \cdots G$ 和 $C \cdots G$ 配对。这样一来 $T \cdots A$ 对就出现在了原来 $C \cdots G$ 对的位置上。这是遗传密码的改变, 因此后果严重。它可能影响生物系统的发育和繁殖, 例如会导致衰老、癌症和变异。在定量理论物理水平上理解和研究这些过程是非常有意义的。

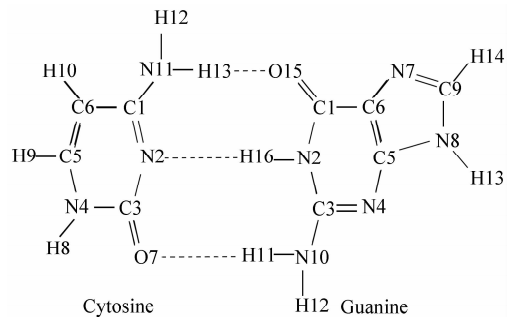


图 1 胞嘧啶和鸟嘌呤碱基以及它们的配对氢键(虚线)数字用来标记碱基中的原子。

图 1 和图 2 给出了胞嘧啶碱基及其顺式亚胺式互变异构体的结构式以及它们不同的配对方式。胞嘧啶的这两个互变异构体的不同主要来源于质子 H13 位置的不同。质子的不同位置使得氢键的分布不同, 最终导致不同的配对。因而, 了解质子 H13 的势能曲面以及它在这个势能面上的运动是理解胞

* 收稿日期: 2007 - 01 - 12

* 基金项目: 国家自然科学基金资助项目(10305001, 10475008); 国家高性能计算中心(北京)资助项目

作者简介: 赵振民(1978-), 男(汉族), 浙江义乌人, 博士, 从事核物理在生物物理中的应用研究; E-mail: pkuzzm@sina.com

嘧啶的互变异构体以及 DNA 中碱基对的改变 C...G → T...A 和 C → C_{imino} 转变之间的关系之关键所在。

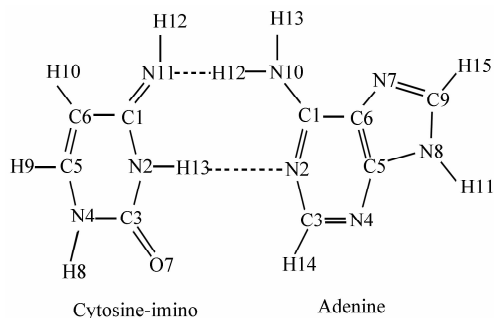


图 2 胞嘧啶顺式亚胺式互变异构体和腺嘌呤碱基以及它们之间的配对氢键

第 2 节中用从头计算法算得胞嘧啶碱基中质子 H13(氢键质子)的势能面。第 3 节中证明质子 H13 在正常型及其顺式亚胺式的构型中都是可以束缚的, 而且能够形成正确的氢键配对。第 4 节中计算质子 H13 在这两个构型之间势垒穿透的几率并讨论了它的生物学意义。第 5 节是结论以及对结果可信度的讨论。

2 胞嘧啶中质子 H13 的势能曲面

我们应用 Gaussian98 软件包中取 6-311G(d, p)基底的 MP2 方法从头计算了质子 H13 的单点能。对 H13 的每一个给定位置, 胞嘧啶中其它原子的位置由要求电子系统的能量最小优化得到。这些原子的位置因而随 H13 的位置变化有一些小的改变。使用固定在胞嘧啶已知结构式上的坐标系, 计算得到的质子 H13 的势能曲面画在图 3 中。图 3 中

的坐标原点位于图 1 中 N11 的位置, x 轴沿着图 1 N11-C1 的方向, y 轴垂直于 x 轴并指向右边。图 3 中颜色越黑的点表示那个位置的势能越低。曲线上的数值是等高线所在位置的单点能数值。可清楚地看到有两个势阱。左边对应胞嘧啶, 右边那个对应胞嘧啶的顺式亚胺式互变异构体。此图解释了图 1 和图 2 中胞嘧啶的两个互变异构体的结构。

3 胞嘧啶及其顺式亚胺式互变异构体的碱基中质子 H13 的结合能

在势阱底部, 我们用坐标的二次式拟合计算得到单点能, 并求得单点能的极小值。左边势阱的单点能极小值是 $-393.953\ 526\ 8$ a. u., 而右边的是 $-393.952\ 654$ a. u.。把坐标原点平移到单点能极小值的位置, 则坐标 1 次项前的系数为 0。用标准方法找到曲面的主轴, 就可以把二次式变换成坐标平方之和。对左边的势阱, 取坐标系 (ξ, η) , 其中 ξ 轴沿着图 1 中胞嘧啶里 N11-H13 的方向, η 轴垂直于 ξ 轴。主轴坐标系 (X, Y) 与 (ξ, η) 相差一转动:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 0.94 & -0.34 \\ 0.34 & 0.94 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}. \quad (1)$$

单点能用 (X, Y) 表示为

$$E = -393.953\ 5 + 0.088\ 937Y^2, \quad (2)$$

其中单点能 E 的单位是 a. u., 长度单位是 Å。我们发现在这个势阱底部 E 与坐标 X 无关。这里的势能面好像是沿着 X 方向的一条谷。坐标变换(1)式表明, 此谷靠近 N11-H13 方向。沿着这个方向质子容易远离原子 N11 而与鸟嘌呤中的 O15 形成一个氢键。这便解释了图 1 中表明的配对情况。质子 H13 还在胞嘧啶中绕着 N11 摆动, 好像‘悬挂’在 N11 上的一个单摆。摆动频率是

$$\nu = \sqrt{\frac{2 \times 0.088\ 937 \times 27.21 \times 1.602\ 177\ 33 \times 10^{-19}}{1.672\ 31 \times 10^{-27} \times (10^{-10})^2}} \\ = 3.427 \times 10^{13} \text{ Hz}, \quad (3)$$

零点能是

$$\frac{1}{2}h\nu = 0.002\ 6 \text{ a. u.}. \quad (4)$$

把这个零点能加在势阱底部的单点能上, 发现质子 H13 的能量是 $-393.950\ 9$ a. u., 低于该势阱附近单点能最小的极大值 $-393.889\ 5$ a. u., 因此质子

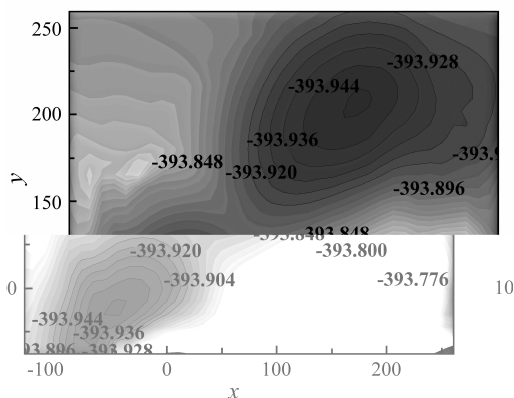


图 3 胞嘧啶中质子 H13 的势能面等高线图
能量单位是原子单位(a. u.), 长度单位是皮米(pm)。

H13 被束缚在正常胞嘧啶内, 结合能 $b = 0.0614$ a. u.。

对图 3 中右面的势阱取坐标架 (x', y') , 其中 x' 轴平行于图 3 中的 x 轴, 把坐标原点移到势阱底部。坐标架 (X, Y) 与 (x', y') 也差一个转动

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 0.30 & -0.95 \\ 0.95 & 0.30 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (5)$$

单点能用 (X, Y) 表示为

$$E = -393.9527 + 0.12878Y^2. \quad (6)$$

在势阱底部 E 又与坐标 X 无关, 势能面再次像沿着 X 方向的山谷。变换(5)式表明, 这个方向几乎垂直于 N11-C1。沿着这个方向质子容易远离图 2 中胞嘧啶顺式亚胺式互变异构体中的原子 N2 而与腺嘌呤中的 N2 原子形成一个氢键。这就解释了图 2 中表明的配对情况。摆动几乎沿着 N11-C1 方向。氢键质子在胞嘧啶顺式亚胺式互变异构体中绕着 N2 原子的摆动, 好像‘悬挂’在 N2 上的一个单摆。摆动频率是

$$\nu = \frac{\sqrt{\frac{2 \times 0.12878 \times 27.21 \times 1.60217733 \times 10^{-19}}{1.67231 \times 10^{-27} \times (10^{-10})^2}}}{2\pi} = 4.12 \times 10^{13} \text{ Hz}, \quad (7)$$

零点能是

$$\frac{1}{2}h\nu = 0.00313 \text{ a. u.} \quad (8)$$

两个势阱附近单点能最小的极大值在同一点, 这个点位于这两个势阱之间。这点的单点能是 -393.8895 a. u., 被称为过渡态。质子的零点能加上在这个势阱底部的单点能得质子在这个势阱中的能量 -393.9496 a. u., 它也比过渡态的能量低, 所以质子 H13 在胞嘧啶顺式亚胺式互变异构体中也是束缚的。这里的结合能 $b' = 0.0601$ a. u.。

生物系统内的热能在室温 $T \simeq 300$ K 下可以估算为 $k_B T = 0.001$ a. u.。我们发现 b 和 b' 都远远大于热能。这表明胞嘧啶及其顺式亚胺式互变异构体对热扰动都是稳定的。

4 两种构形间的势垒穿透及其生物学意义

按照量子力学, 质子还可穿透势垒, 从一个势阱到另一个势阱。我们首先找到连接两个势阱底部

的最低势能路径。用图 3 中坐标, 这条路径可用下面的方程描述:

$$y(x) = 1.1417352 + 0.71244583x + 0.3508047x^2 + 0.04422011x^3 + 0.06419875x^4 - 0.56534255x^5 + 0.25248558x^6 \quad (9)$$

其中长度单位为 Å。采用原子单位的势能可以用公式

$$V(x) = -393.908877 + 0.103691x - 0.102525x^2 - 0.3063312x^3 + 0.102250x^4 + 0.355250x^5 - 0.195806x^6 \quad (10)$$

拟合。图 4 中画出了质子 H13 在两个势阱之间的势垒。按照 WKB 公式, 质子穿透这个势垒的几率为

$$W = \exp\left\{-\frac{2}{\hbar} \int_s \sqrt{2m_p[V(x) - E]} dl\right\} = \exp\left\{-\frac{2}{\hbar} \int_a^b \sqrt{2m_p[V(x) - E] \left[1 + \left(\frac{dy}{dx}\right)^2\right]} dx\right\} \approx \exp[-54.85] \approx 1.50929 \times 10^{-24}. \quad (11)$$

其中 $E = -393.9496$ a. u. 是胞嘧啶顺式亚胺式互变异构体中质子 H13 的总能量, a 和 b 是由条件 $E = V(x)$ 定出的质子的 x 坐标, s 是曲线端点分别为 $x = a$ 和 $x = b$ 的最低势能路径。

单位时间内的穿透几率是

$$P = \nu W = 3.427 \times 10^{13} \times 1.50929 \times 10^{-24} \text{ Hz} = 5.168 \times 10^{-11} \text{ s}^{-1}, \quad (12)$$

因此, 质子处于这些势阱之一内的平均寿命是 6×10^2 a. 这意味着达到热平衡所需的时间大于 600 a. 人类在自己的 100 a 的生命内不可能看到气态胞嘧啶及其顺式亚胺式互变异构体之间的平衡。我们因此认为自然界中胞嘧啶及其顺式亚胺式互变异构体之间的巨大的丰度差别不是热平衡的结果, 而是生物进化中自然选择的结果。从(12)式还可以看到, 在人类 100 a 的寿命内, 大约 16% 的胞嘧啶碱基可能自发转变成顺式亚胺式互变异构体构型。它们中的一些可能又转变回正常型, 然而另一些, 按照 Lowdin 的假设, 在 DNA 的进一步复制中可能引起衰老、突变或癌症。当然胞嘧啶和 DNA 分子中其它的碱基在生物体内不是孤立的。它们到‘病态’互变异构体的转变会受环境的影响。为了阐明这个影

响, 需要针对碱基-环境系统做更复杂的计算。沿着这个方向的更深入的研究将是很有意义的。

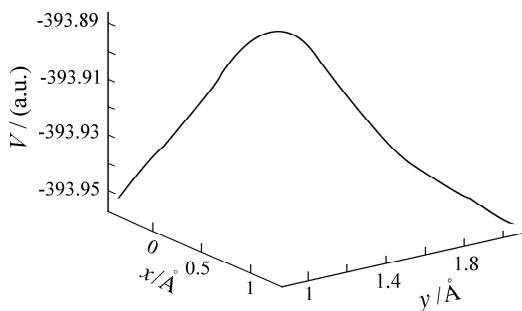


图 4 两个势阱之间 H13 质子的势垒

5 结论和讨论

量子力学从头计算给出了胞嘧啶及其顺式亚胺式互变异构体的正确结构, 包括可能的氢键位置和方向。这使我们相信这个方法在生物物理研究中是可行的、有效的。计算结果还预言了胞嘧啶的正常型及其顺式亚胺式互变异构体之间的转变几率, 结果合理。上面已经讨论了它的生物学意义, 还值得进一步探讨。

生物分子中有大量的原子。分子的能量由其中各原子的能量和它们之间的相互作用能组成。在物理、化学或生物反应中分子能量的改变量通常只占分子总能量的很小一部分。为了研究这些反应中生物分子的行为以及能量改变, 我们需要非常精确的计算以保证足够多的有效数字。例如, 在第 3 节的开始我们已经看到, 为了比较质子在两个势阱底部的单点能, 计算值中至少需要 6 位有效数字是精确

的。在如此大量的数值计算中这很不容易做到。对本文的结果, 势能曲面的一般的趋势是可信的, 但是两个势阱底部的单点能之差就不那么可信了, 因为前 5 位最可信的数字都被减掉了。另一方面, 势垒穿透几率是可信的, 因为部分统计误差在积分中可能相互抵消。总之, 沿着这个方向进一步的研究中提高计算精度是非常重要的。

参考文献 (References):

- [1] Lowdin Per-Olov. *Rev Mod Phys.* 1963, **35**: 724.
- [2] Drake J W, Allens E F, Forsberg S A. *Nature.* 1969, **221**: 1 128.
- [3] Wolfenrn R V. *J Mol Biol.* 1969, **40**: 307.
- [4] Nowak M J, Szczepaniak K, Barski A, *et al.* *Natreforsch.* 1978, **C33**: 876.
- [5] Raszka M, Kaplan N O. *Proc Nat Acad Sci USA.* 1972, **69**: 2 025.
- [6] Shoup R R, Miles H T, Becker E D. *J Phys Chem.* 1972, **76**: 64.
- [7] Topal M D, Fresco J R. *Nature (London).* 1976, **263**: 285.
- [8] Topal M D, Fresco J R. *Nature (London).* 1976, **263**: 289.
- [9] Saenger W. *Principles of Nucleic Acid Structure.* New York: Springer-Verlag, 1984, 149—158.
- [10] Rein R, Harris F E. *J Chem Phys.* 1965, **43**: 4 415.
- [11] Colominas C, Luque F J, Orozco M. *J Am Chem Soc.* 1996, **118**: 6 811.
- [12] Rao C N R. *Jerus Symp Quant Chem Biochem.* 1973, **5**: 107.
- [13] Geller M, Lesyng B. *Biochem Biophys Acta.* 1975, **417**: 407.
- [14] Kulakowska I, Geller M, Lesyng B, *et al.* *Biochem Biophys Acta.* 1975, **417**: 420.
- [15] Rao K G, Rao C N R. *J Chem Soc Perkin Trans.* 1973, **2**: 889.

Motion of Hydrogen Bond Proton in Cytosine and Transition between Its Normal and Imino States*

ZHAO Zhen-min^{1, 1)}, ZHANG Qi-ren¹, GAO Chun-yuan¹, ZHUO Yi-zhong²

(1 *School of Physics, Peking University, Beijing 100871, China;*

2 *China Institute of Atomic Energy, Beijing 102413, China)*

Abstract: The potential energy surface of the H13 proton in base cytosine of the DNA molecules is calcu-

* Received date: 12 Jan. 2007

* Foundation item: National Natural Science Foundation of China(10305001, 10475008); High Performance Computing Center of China (Beijing)

1) E-mail: pkuzzm@sina.com

lated ab initio at the Gaussian98 MP2/6-311G(d, p) level. Two potential wells are found. One corresponds to the normal cytosine, while the other corresponds to its imino tautomer. The bindings of the proton in these wells are stable enough against the thermo-disturbance. The motions of the proton in these wells are oscillations around the nearest nitrogen atom like the pendulum, and may move far away from the nitrogen atom to form the hydrogen bond with other bases. The estimated tunneling probability of the H13 proton from one well to another well shows that the life time of the proton staying in one of these wells is about 6×10^2 a. It is too long to let tautomers of cytosine be in thermodynamical equilibrium in a room temperature gas phase experiment. The biological significance of these results is discussed.

Key words: hydrogen bond proton; potential energy surface; tunneling effect