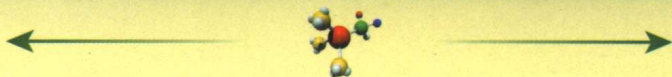


核酸碱基与无机小分子间的相互作用

© 张士国 著



国防工业出版社

National Defense Industry Press

本书由滨州学院学术著作出版基金资助出版

核酸碱基与无机小分子间的相互作用

张士国 著

国防工业出版社

·北京·

图书在版编目(CIP)数据

核酸碱基与无机小分子间的相互作用/张士国著.
北京:国防工业出版社,2006.12
ISBN 978-7-118-04952-7

I.核... II.张... III.核酸碱基-研究 IV.
Q526

中国版本图书馆 CIP 数据核字(2007)第 003124 号

※

国防工业出版社 出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100044)

天利华印刷装订有限公司印刷

新华书店经售

*

开本 850×1168 1/32 印张 6% 字数 166 千字
2006 年 12 月第 1 版第 1 次印刷 定价 30.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)68428422
发行传真:(010)68411535

发行邮购:(010)68414474
发行业务:(010)68472764

目 录

第 1 章 绪论	1
1.1 核酸碱基的正常结构与异构体研究	4
1.2 碱基与电子作用的研究	6
1.3 碱基质子化的研究	9
1.4 碱基与溶剂水分子作用的研究	11
1.5 碱基与金属离子作用的研究	13
1.6 碱基与其它客体分子作用的研究	17
1.7 小结	19
1.8 我们的工作	19
参考文献	21
第 2 章 研究方法及其理论基础	30
2.1 第一性原理与 Hartree - Fock 近似	31
2.2 从头计算方法(ab initio Method)	33
2.3 密度泛函方法(DFT)	35
2.4 分子力学方法和分子动力学方法	38
2.5 Gaussian98(03)程序	40
2.5.1 程序功能	41
2.5.2 结构优化与频率分析	41
2.5.3 分子轨道分析	48
2.5.4 自然键轨道分析	48
2.5.5 分子中的原子(AIM)分析	50
2.5.6 分子间相互作用的研究方法	52
2.5.7 过渡态的寻找计算	56
参考文献	56

第3章 碱基-BH₃复合物的研究	59
3.1 引言	59
3.2 胸腺嘧啶-BH ₃ 复合物	60
3.2.1 计算方法	60
3.2.2 结果与讨论	61
3.2.3 小结	70
3.3 鸟嘌呤-BH ₃ 复合物	70
3.3.1 计算方法	70
3.3.2 结果与讨论	71
3.3.3 小结	81
3.4 腺嘌呤-BH ₃ 复合物	82
3.4.1 计算方法	82
3.4.2 结果与讨论	83
3.4.3 小结	90
3.5 胞嘧啶-BH ₃ 复合物	91
3.5.1 计算方法	93
3.5.2 结果与讨论	93
3.5.3 小结	99
3.6 尿嘧啶-BH ₃ 复合物	100
3.6.1 计算方法	101
3.6.2 结果与讨论	101
3.6.3 小结	106
参考文献	107
第4章 碱基-BX₃(X=F,Cl)复合物的研究	111
4.1 引言	111
4.2 鸟嘌呤-BX ₃ (X=F,Cl)复合物	112
4.2.1 计算方法	112
4.2.2 结果与讨论	113
4.2.3 小结	125
4.3 腺嘌呤-BX ₃ (X=F,Cl)复合物	126

4.3.1	计算方法	126
4.3.2	结果与讨论	126
4.3.3	小结	136
4.4	胞嘧啶-BX ₃ (X=F,Cl)复合物	136
4.4.1	计算方法	136
4.4.2	结果与讨论	137
4.4.3	小结	149
4.5	尿嘧啶-BX ₃ (X=F,Cl)复合物	150
4.5.1	计算方法	150
4.5.2	结构与讨论	151
4.5.3	小结	163
	参考文献	163
第5章	碱基-CO复合物的研究	165
5.1	引言	165
5.2	计算方法	166
5.3	结果与讨论	166
5.3.1	复合物的几何结构	172
5.3.2	复合物的结合能与稳定性	173
5.3.3	复合物的振动分析	174
5.3.4	自然键轨道分析	175
5.3.5	复合物的AIM分析	176
5.4	小结	176
	参考文献	177
第6章	碱基-NO复合物的研究	179
6.1	引言	179
6.2	计算方法	180
6.3	结果与讨论	181
6.3.1	复合物的几何结构	186
6.3.2	复合物的结合能与稳定性	187
6.3.3	复合物的振动分析	188

6.3.4	自然键轨道分析	189
6.3.5	复合物的 AIM 分析	190
6.4	小结	190
	参考文献	191
第 7 章	主要结论与展望	193
7.1	主要结论	193
7.2	展望	196

第 1 章 绪 论

从 20 世纪 20 年代末量子力学原理应用于化学问题的研究开始，化学领域的面貌发生了十分重大的变化。今天，我们对化学键本质、分子光谱现象、分子反应活性、分子反应动力学以及许多其它基本问题的认识深度都紧密地和我们对于原子、分子以及其中电子运动规律的了解深度关联着。不仅静态的分子结构理论，而且动态的化学反应理论也主要基于量子化学来解释，量子化学正在改变着化学从一门传统实验性学科向着实验和理论研究并重的方向发展。

现代化学既重视合成新化合物或提取天然存在的物质以满足人类生活的需要，又极为注重研究化合物的几何结构、电子结构及其结构与性质之间的联系。借助于现代实验仪器与实验方法已经确立了许多经验规律，随着计算机技术和量子化学理论的发展，应用量子力学原理计算化合物的电子结构、反应能力等已经得到广泛的应用。量子化学方法使人们能够仅从实验数据与经验规律的描述进一步向从微观水平上描述相互作用的方向发展。通过电子结构特性的一些理论计算结果与实验值的对比，人们不仅可以判断量子化学对已知化合物几何结构、电子结构、能量以及其它性质计算和描述的正确程度，更重要的是，随着量子化学计算方法的发展和计算机计算能力的发展，现在的量子化学计算完全可以用来预测未知的化合物的结构和性质。正因为如此，量子化学计算已经广泛地用于生命科学、药物设计和材料科学的研究中，并且不仅被理论工作者掌握而且被广大的实验工作者所利用。

当今自然科学各领域中最引人入胜的恐怕要数研究生命现象

和本质的生命科学了。近 20 年来，世界上许多著名的科学家纷纷预言：“生物学对 21 世纪的影响就等于物理、化学对 20 世纪的影响一样大”。“今后将不再是‘矿物时代’，而是‘生物时代’。谁抓住了生物学，谁就是时代的霸主”^[1]。在分子生物学急剧发展并取得重大成就以及物理、化学、计算机科学、实验技术飞速进步的推动下，产生了量子力学与分子生物学相结合的产物——量子生物学，它是一门运用量子力学原理，通过数学运算来研究生物分子的几何结构、电子结构以及电子结构与生物功能之间的联系、生物分子间的相互作用、电磁性质、化学反应、能量转移和电荷转移等的学科。

生物学与化学的密切联系是众所周知的，尤为突出的是量子生物学与量子化学的紧密联系，事实上量子生物学的理论方法都是沿袭了量子化学的研究方法。和量子化学一样，量子生物学研究的分子、原子、电子体系可用体系相应的 Schrodinger 方程解的波函数来描述。原则上 Schrodinger 方程的全部解保证了多电子体系中电子结构与相互作用的全面描述。然而由于数学处理的复杂性，在实践中总希望发展和运用量子力学的近似方法，从而无需进行很复杂的计算就可以说明复杂分子体系的主要特性。这就必须在原始量子力学方程中引进一些重要的简化，以便得到一定准确度的近似解。在某些情况下，则还需要把量子力学方法和经典力学方法结合起来，运用分子力学方法(Molecular Mechanics, MM)以及分子动力学(Molecular Dynamics)、Monte Carlo 等经典力学方法，计算出体系的能量，进行模拟研究^[2]。

一切物质都是由原子构成的，生物分子也不例外。站在原子和电子的角度看，就无所谓生物与非生物之间的区别了。构成生物有机体的基本化学元素——氢、碳、氮、氧元素的原子，由于它们具有的特殊性质，使它们成为构成“生命机器”中的最基本的“元件”，这些原子的行为在“生命机器”中是至关重要的。生命现象的某些缺陷，如突变、致癌、免疫缺乏等，如果仅从群体、个体、器官、组织、细胞乃至分子水平上加以认识，往往是很不

够的。只有在原子、电子的水平上才能真正弄清楚其本质，这可以看成是对生物大分子进行量子力学处理和计算的必要性。

生物大小分子之间的相互作用的规律是生命过程，也是生物应答的普遍基础^[3]。它们不但普遍存在，而且具有共同的本质和作用规律。生物应答取决于构成它的分子的应答；而分子应答实际上是生物体系与外来物质之间相互作用的化学过程。例如，与单电子转移有关的自由基链反应是正常的和异常的生物氧化化学过程的关键环节，它在动植物和生物分子的氧化性损伤中是共同的基本环节。现在已经积累了相当多的实验结果，尚有待于在分子层次上来研究。这些分子间相互作用的研究、化学过程的机理和控制规律的研究将有助于解决许多疾病的诊断与防治、环境因素对人和生物的胁迫、提高农牧业产率和产品质量、安全合理利用化学品和天然物质的问题。

核酸是重要的生物大分子之一，在生命活动中起着及其关键的作用。作为遗传信息的载体，它参与遗传信息在细胞内的传递与表达，从而促成代谢过程及其控制。尽管 DNA 美丽的双螺旋结构在 20 世纪 50 年代就已被人们认识，同时其它类型的生物高分子的结构也不断地被发现和研究，从而对遗传、变异、免疫、发育和繁殖在分子水平上有所了解。我们已经知道碱基通过氢键相互配对(Watson-Crick 对)以及碱基之间的堆积作用，还有磷酸部分带的电荷构成了 DNA 双螺旋的稳定因素，且为其生物功能的发挥提供了支持，但碱基本身在生理条件下会存在各种变体，以及碱基在生理条件下会与各种粒子(电子、质子、无机或有机分子)作用，而发生各种所谓 DNA 的损伤、碱基的错配^[4, 5]。上述过程均与生物体的变异、疾病等过程有关。因此研究碱基的结构变化以及其与各种生物体内存在的分子，包括小分子的作用规律是一项极有意义的工作。而在该领域的研究中，虽然有多种实验手段(如 X-射线分析、核磁共振、红外光谱、微波等技术)可以利用，但采用以量子力学为基本原理的研究化学问题的计算方法——量子化学，也是不可缺少的手段，而且越来越受到化学家、生物学家等

的重视并且取得了广泛的进展。以下将对近年来采用量子化学方法对核酸碱基(NAB)的结构变化及其与电子、质子、溶剂分子、金属离子和其它分子(离子)作用的研究做一综述。

1.1 核酸碱基的正常结构与异构体研究

业已知道,在核酸中的碱基有两类,即嘌呤碱基和嘧啶碱基,其中嘌呤碱基由两个杂环构成。从碱基的基本组成原子看,它们有明显的多样性特点,其中包括氢、碳、氮和氧。这种特点决定了核酸碱基具有多方面的反应性能。事实上,在碱基的不同部位的确能够发生诸如烷化、卤化、脱氮、氰乙基化和氧的加成等反应。这不仅是核酸储存信息和传递信息的必要基础,而且也可能是造成突变使物种淘汰和进化、衰老或致癌的原因。另一方面,由于碱基种类的多样化(碱基是嘧啶或嘌呤的衍生物),经过适当的组合会得到足够多的信息,因此也是这类生命物质形成高分子的可能的和必要的条件。含有氧的鸟嘌呤、胸腺嘧啶、胞嘧啶和尿嘧啶,它们在一定条件下存在酮式和醇式构型的互变。不含氧的腺嘌呤碱基也存在胺型和亚胺型互变。胞嘧啶除了存在酮式和醇式互变外,还存在胺型和亚胺型互变。然而,研究表明,在正常生理条件下,酮式(内酰胺)是优势结构,含量比醇式(内亚酰胺)要多得多,故一般把前者视为正常型,将后者视为异常型。常见形式的碱基的结构由图 1-1 示出。

除了图 1-1 所列的 5 种常见碱基外,在活性细胞的核酸组成中,常常还包括一些修饰碱基(有时也称稀有碱基)。例如在小麦的胚、细菌和高等生物中发现有 5-甲基胞嘧啶(m^5C),在某些细菌和噬菌体中则包含有 N^6 -甲基腺嘌呤等等。迄今为止,在 DNA 和 RNA 中已经检测到的修饰碱基约 70 多种。碱基与碱基之间的作用或碱基与其它分子或离子的作用是由碱基的几何结构和电子结构决定的,因此关于碱基的几何结构和电子结构的研究自 20 世纪 50 年代就开展了各种分子轨道方法的计算研究^[1]。前期的工作主要是使

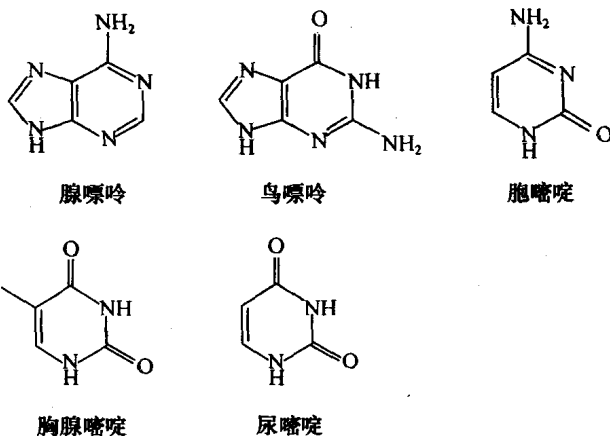


图 1-1 核酸中正常碱基的结构图

用各种半经验分子轨道方法的计算，如 HMO、EHMO、CNDO 和 INDO 等方法或使用基于 Hartree Fock 方法的 *ab initio* 计算研究。

但是由于方法本身的缺陷和计算机计算能力的限制，在 20 世纪 90 年代初以前，对于类似于碱基这样的中等大小的分子的计算没有得到可靠的结果^[7]。随着量子化学计算方法的发展，尤其是以惊人的速度增加的计算机计算能力的提高，使现代高水平的从头算方法(*ab initio*)能够提供精确的、可信的、可与实验数据相比拟的研究结果。从头算的计算精度取决于两个因素：一是构成分子轨道的基组的大小，二是包括电子相关计算的计算方法的使用。对于研究的体系如何选取基组和计算方法需要平衡这两个因素的影响。毫无疑问，目前高水平的从头算方法的计算结果可以达到精确的生理实验结果的精度。并且高水平量子化学计算可以用于尚没有实验研究信息的体系^[7]。

量子化学计算方法的原理是基于 Hartree-Fock (HF) 近似，而在含有电子相关作用的后 Hartree-Fock (Post-HF) 中，由于明显地考虑到了电子的相关作用，使计算结果得到了很好的改善。例如，在研究分子间作用时 London 色散作用是非常重要的，而这种作用

是电子相关作用的结果，另外电子相关作用的考虑对于电荷分布的计算、分子偶极矩的计算都是非常有益的。目前，最实用的含有电子相关作用的 Post-HF 计算是二阶 Moller-Plesset 微扰方法和密度泛函理论方法，其计算量较小但能达到较高的精度。而其它的精度更高的方法如含有三重电子激发态的耦合簇方法 (CCSD(T))，虽然在精度方面有优势，但在现有的条件下对于中等大小的分子的计算量是令人难以接受的。

对于碱基电子结构的研究，常常计算其原子的净电荷、偶极矩和极化率，用于讨论碱基与碱基之间或碱基与其它分子或离子的作用。但有一个问题值得注意，由于量子力学中没有与原子电荷对应的算符，因而也没有实验值去衡量原子电荷的大小。因此量子化学计算中不管是哪种方法给出的原子电荷均是主观任意的。尽管文献中大量使用 Mulliken 布局数分析作为讨论原子净电荷的基础，但这一数据对计算使用的方法和基组有很强的依赖性^[8]，因此应尽量不用。比较好的计算原子净电荷的方法是由拟合分子静电势(ESP)得到的原子电荷，因为分子的静电势、偶极矩和高阶电矩均是可测量的^[9]。ESP 电荷分布在研究静电作用能时是很有用的，在普遍使用的 AMBER 力场中就是使用分子静电势来拟合得到原子净电荷的。ESP 电荷的使用也有缺点，就是不便于使用与类似于氨基酸一样的柔性分子的电荷计算，因为由一种构象得到的电荷对另一种构象是不适用的。对于比较刚性的核酸中碱基来说，用 ESP 电荷研究其分子间作用是非常有效的。另外两种得到原子电荷的较好方法是自然键轨道方法(Natural Bond Orbital method, NBO)^[10]和基于电子密度拓扑分析的“分子中的原子”理论(Atom In Molecules, AIM)产生的原子电荷^[11]，上述两种电荷对于研究分子间的静电作用都是非常有效的。

1.2 碱基与电子作用的研究

DNA 和 RNA 辐射受损的初始阶段是与其结构基团吸收电荷

有关的，尤其是在 DNA 的损伤和修复中，电子被嘧啶和嘌呤部分捕捉起着重要的作用^[12, 13]。由于这一过程的内在复杂性，每一个碱基与电子的结合与它们的绝热电子亲合能(Adiabatic Electron Affinity, AEA)密切相关。所以，准确地了解相关的热力学数据对于理解可能引起突变的部分与电子结合的过程是重要的。另外，理解沿 DNA 链的电子流动或电子转移这一过程也依赖于对碱基与电子形成的阴离子的研究。尽管有许多实验研究和理论研究成果，但对核酸碱基的绝热电子亲合能的研究仍是值得关注的课题。Henry F. Scharfer III 小组^[14]在近期使用密度泛函理论的多种不同方法计算了 DNA 和 RNA 中碱基的绝热电子亲合能。研究表明，尿嘧啶和胸腺嘧啶形成的价型阴离子时放出的能量在 0.05eV~ 0.25eV 之间，而实验测得的 U 和 T 的 AEA 是 0.4eV 和 0.3eV。气态的腺嘌呤不能形成稳定的价型阴离子，而对胞嘧啶和鸟嘌呤的计算没有得出能否形成价型阴离子的明确结论，不过这对于理解有关辐射生物学问题是有帮助的。所谓价型阴离子 (Valence Anion)，是额外的电子填充在中性分子的最低空轨道 (LUMO) 中，并且阴离子的构象与中性分子相比有较大的变化。而与之对应的是偶极结合阴离子 (Dipole-Bound Anion)，在这种阴离子中，外来电子占据的轨道是中性分子轨道的混合(组合)，电子与中性分子的结合较弱，一般不能引起分子构象的变化^[15]。Maciej Guoto-wski 等人使用从 B3LYP 到 CCSD(T) 等不同精度的计算方法对鸟嘌呤的 4 种较为稳定的异构体与电子的作用进行了研究^[16](图 1-2)。计算中对于气态和极性溶剂中的情况均进行了研究。在极性溶剂中的计算结果作为初始猜测函数和初始构型，再进行气相状态的计算。结果表明，典型构型 G 和 GNH7、GHN7H 变体与电子的结合在气相是价型阴离子，在 CCSD(T) 水平上的电子亲合能分别是 0.58eV、0.21eV 和 0.39eV。另外，在这些价型阴离子的形成过程中，其分子骨架发生了严重的扭曲，G 主要发生在六员环部分而 GNH7 和 GHN7H 主要发生在五员环部分。在标准状况下气相鸟嘌呤的主要构型是 GNH7，而典型的构型 G 是

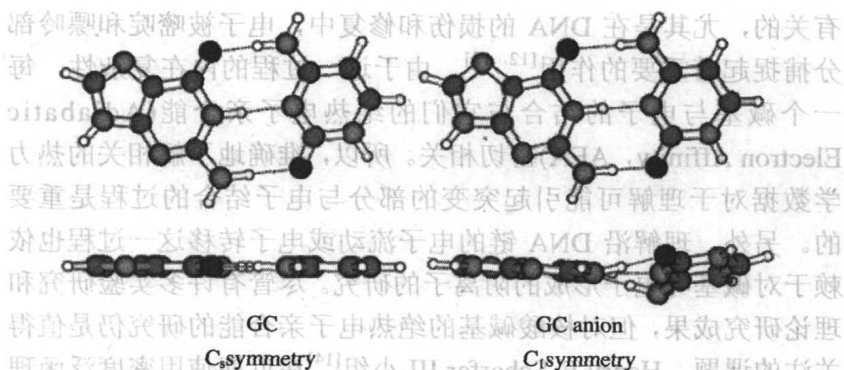


图 1-2 GC 和 GC^- 的侧面图，阴离子中 C 部分的构象发生了变化^[17]。人们最感兴趣的。但 GNH7 构型与电子结合的能力强于 G，那么对于阴离子而言，后者的能量较低，在气态的平衡构型中占多数。Henry F. Scharfer III 小组对鸟嘌呤-胞嘧啶 Watson-Crick 碱基对的电子亲和能进行了理论研究^[17]。结果表明，在 GC^- 对中，G 部分的几何参数与 G 的中性分子几乎相同，碱基对阴离子中 C 部分的几何构型更像阴离子 C^- 。对电荷和几何构型的分析可以得到如下的结论：外来的电子分布于胞嘧啶部分。另外，在 GC^- 中，G 和 C 之间的氢键也由于电荷分布而发生变化。由于体系带有更多的负电荷，O-N 键长增长，并且变得类似于 GpC 晶体中的长度(图 1-3)。在 GpC 晶体中，似乎是有负电荷定域于 G-C 片段而使得 O-N 键长增长。

由于电子亲和能和电离能可以由分子束实验得到^[18, 19]，其实实验结果可以与从头算得到的结果直接比较^[20, 21]。关于碱基的 EA 和 IP 数据在文献中有很多报道，这些数据之间也存在差异，因此对碱基的这一性质做出评价是困难的。但对核酸碱基的 EA 和 IP 数据进行评价是非常有意义的，因为对这些数据的分析，有助于理解沿 DNA 链的电子转移过程以及这一过程是如何被碱基对内的质子转移影响和终止的，也有助于理解与 DNA 的辐射损伤过程有关的碱基与电子的结合^[22, 23]。

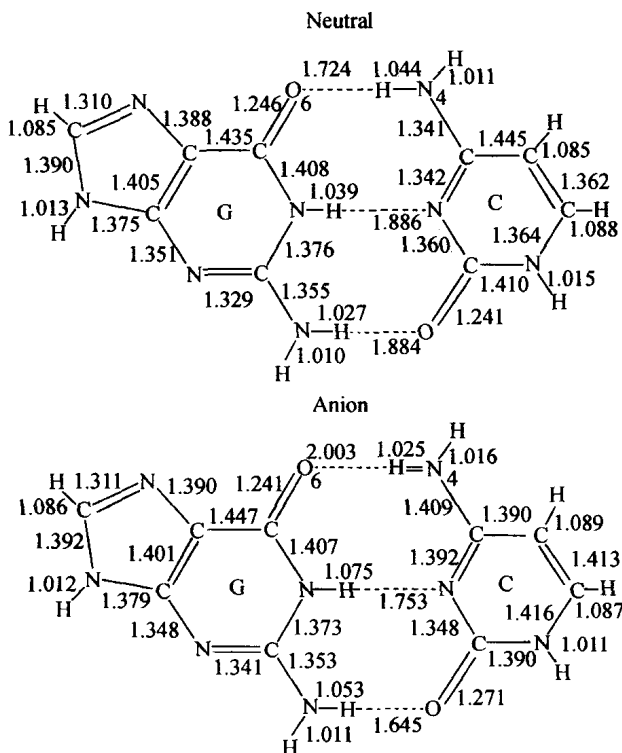


图 1-3 B3LYP/DZP++计算的 GC 中性对和阴离子的键长^[17]

1.3 碱基质子化的研究

在核酸中碱基的质子化是经常出现的^[24-28]。在四链 DNA 中，胞嘧啶的质子化对三维结构的存在起了稳定作用，在嘧啶-嘌呤-嘧啶类三链 DNA 中也存在质子化的胞嘧啶，但若存在连续质子化的碱基则形成去稳定化作用。四链 DNA 是令人兴奋的研究分子间作用的体系之一^[29]，在其富含胞嘧啶的区域，存在连续的质子化胞嘧啶-胞嘧啶对。在该碱基对之间存在快速的内部质子传递，因为 NMR 和 X-射线的实验结果表明在该体系中胞嘧啶是等价的^[30-32]。碱基在气相的质子化能是比较容易计算的，尽管这些数据尚不能

准确表示在核酸内部碱基的质子化情况。虽然碱基在中性条件下是不倾向于质子化的，但核酸内部的分子间相互环境存在的结果却大大地改变了质子化作用的自由能，使其向有利于质子化的方向发生了变化^[33]。核酸中的碱基在接近中性的条件下能够质子化，是因为有较高极化率的碱基之间存在异常强的堆积作用。质子化的胞嘧啶倾向于使腺嘌呤质子化结合^[33]，然而当质子化的腺嘌呤-胞嘧啶碱基对一旦形成，质子的向腺嘌呤的转移就使分子离子-分子偶极的作用得到加强。

表 1-1 给出了计算得到的部分核酸碱基的质子亲和能，关于这些数据的计算可参见文献^[34, 35]。质子亲和能定义为质子化的碱基的总能量与未质子化碱基总能量之差，即 $P=E(BH^+)-E(B)-E(H^+)$ ，当然 $E(H^+)=0$ 。在 20 多年前，Del Bene 用比较精确的方法计算得到各碱基的质子化能顺序如下：鸟嘌呤>胞嘧啶>腺嘌呤>>尿嘧啶~胸腺嘧啶。但在表 1-1 列出的数据则是由以下顺序：胞嘧啶>硫代

表 1-1 基于 MP2/6-31G**方法计算的部分核酸碱基的质子亲和能

Cytosine	P _{N3} -241.4	P _{O2} -241.9 ^a	P _{O2} -232.8 ^b	P _{N4} -212.7	
Adenine	P _{N1} -234.8	P _{N3} -237.1	P _{N7} -228.6	P _{N6} -218.6	
Guanine	P _{N3} -223.5	P _{N7} -239.8	P _{O6} -233.5 ^c	P _{O6} -224.3 ^d	P _{N2} -205.3
Thymine	P _{O4} -217.5 ^b	P _{O4} -214.9 ^a	P _{O2} -211.1 ^b	P _{O2} -209.8 ^a	
Uracil	P _{O4} -216.4 ^b	P _{O4} -213.4 ^a	P _{O2} -208.0 ^b	P _{O2} -206.6 ^a	
6-Oxopurine	P _{N3} -217.7	P _{N7} -232.8	P _{O6} -228.3 ^c	P _{O6} -220.1 ^b	
2-Amino adenine	P _{N1} -238.8	P _{N3} -239.2	P _{N7} -235.2	P _{N6} -223.3	P _{N2} -230.8
6-Thioguanine	P _{N3} -221.9	P _{N7} -240.1	P _{S6} -235.1 ^a	P _{S6} -231.3 ^b	P _{N2} -203.8
2-Thiouracil	P _{O4} -215.5 ^b	P _{O4} -212.5 ^a	P _{S2} -210.9 ^b	P _{S2} -210.3 ^a	
Purine	P _{N1} -230.9	P _{N3} -219.4	P _{N7} -225.2		
^a c is with respect to N3; ^b trans with respect the N3; ^c trans with respect to N1; ^d c is with respect to N1					