

• 研究论文 •

## SET 区域 Rubisco 赖氨酸甲基化转移酶催化赖氨酸多重甲基化的量子力学研究

蒋彦可<sup>a</sup> 邹建卫<sup>\*,b</sup> 卢运祥<sup>a</sup> 蒋勇军<sup>b</sup> 俞庆森<sup>a</sup>

(<sup>a</sup>浙江大学化学系 杭州 310027)

(<sup>b</sup>浙江大学宁波理工学院分子设计与营养工程重点实验室 宁波 315104)

**摘要** 用从头计算和 DFT 方法计算 *S*-腺苷甲硫氨酸多重甲基化赖氨酸的反应机理。气相中非催化反应的能垒比较低, 反应势能剖面类似于 SN2 反应, 而在电介质中, 反应能垒很高。在 B3LYP/6-31G\* 和 ONIOM (MP2: B3LYP) 水平, 研究了 Rubisco 大亚基甲基转移酶活性位点的三个残基催化转移甲基到赖氨酸反应的过程。在 ONIOM (MP2: B3LYP) 水平, 预测的反应能垒为 82.8 kJ/mol, 与从实验得到的 87.4 kJ/mol 接近。计算结果表明, 酶对过渡态并没有特殊的稳定作用, 主要通过为反应物提供一个没有溶剂的环境, 用临近的残基把亲电和亲核的底物连接在一起, 进而促使甲基化反应的进行。

**关键词** 赖氨酸甲基化; 甲基化转移酶; *S*-腺苷甲硫氨酸; 洋葱算法; 从头计算

## Quantum Mechanical Study on Multiple Lysine Methylation Catalyzed by the SET Domain Enzyme Rubisco LSMT

Jiang, Yanke<sup>a</sup> Zou, Jianwei<sup>\*,b</sup> Lu, Yunxiang<sup>a</sup> Jiang, Yongjun<sup>b</sup> Yu, Qingsen<sup>a</sup>

(<sup>a</sup> Department of Chemistry, Zhejiang University, Hangzhou 310027)

(<sup>b</sup> Key Laboratory for Molecular Design and Nutrition Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315104)

**Abstract** The reaction mechanism of multiple lysine methylation by *S*-adenosylmethionine (as modeled by sulfonium ion) has been studied by means of *ab initio* and DFT calculations. The barriers of uncatalyzed reaction in gas phase are somewhat lower and the reaction profile resembles typical SN2 reactions, whereas the reactions present significantly higher barriers in dielectric media. The hydrogen bond contribution of three residues in the active site of the Rubisco large subunit methyltransferases to monomethylation reaction has also been evaluated by B3LYP/6-31G\* and ONIOM MP2: B3LYP calculations. At the ONIOM (MP2:B3LYP) level, the reaction barrier was predicted to be 82.8 kJ/mol, which is in excellent agreement with the activation barrier of 87.4 kJ/mol estimated from the experimental value. The enzymes can not yield special stabilization on the transition state, however they provide a solvent-free room for the substrate (sulfonium ion and amine) and can bring the nucleophile and the electrophile together through the nearby residues, therefore facilitating the methylation processes.

**Keywords** lysine methylation; methyltransferase; adoMet; ONIOM; *ab initio*

乙酰化、甲基化、磷酸化和泛素化属不同类型的组蛋白共价修饰, 这些共价修饰调节多种 DNA 相关事件<sup>[1-3]</sup>, 这在 10 年前是难以想象的。人们日益明确地认

识到甲基转移酶的反常调节能够导致各种各样的人类疾病<sup>[4,5]</sup>。组蛋白在赖氨酸和精氨酸的甲基化是由一类通常具有共同保守的、被称作 SET 区域(Suvar 3-9 en-

\* E-mail: jwzou@nit.net.cn; Tel.: 0574-88229517; Fax: 0574-88229516

Received April 1, 2008; revised September 26, 2008; accepted November 3, 2008.

hancer of zeste trithorax domain)的蛋白所执行的, 这个区域对于酶的活性是至关重要的. 最近, SET 区域酶催化甲基化组蛋白赖氨酸受到广泛关注, 因为越来越多的证据表明组蛋白赖氨酸常规甲基化的失败会导致类似癌症等的发展失调类疾病<sup>[6]</sup>.

SET 区域甲基转移酶从 *S*-腺苷甲硫氨酸(AdoMet)中催化转移甲基到蛋白质底物特定的赖氨酸残基, 如组蛋白 H3 和 H4 和双氧全酶复合物(Rubisco holoenzyme complex)的大亚基上. 尽管每一个赖氨酸都能够被一、二、三甲基化, 一个组蛋白赖氨酸甲基转移酶只能转移特定数目的甲基到它对应的赖氨酸残基上<sup>[7,8]</sup>. 不同的赖氨酸甲基转移酶对一个给定的赖氨酸能够进行不同水平的甲基化, 赖氨酸甲基化水平在基因调节中有不同的作用, 在基因活性中具有不同的空间分布<sup>[9~11]</sup>. 组蛋白赖氨酸 9 的一、二、三甲基化与不同的外遗传状态直接相关, 而且能够结合酶和染色质区域关联蛋白进而被引入到相关的染色质区域<sup>[12]</sup>.

Rubisco 甲基转移酶[Rubisco large subunit methyltransferase (LSMT)]是第一个酶域被特征化的 SET 区域甲基转移酶. 菠菜的 Rubisco 甲基转移酶(pLSMT)是一个非组蛋白特定植物酶, 在 Rubisco 全酶大亚基, 特定的用来三甲基化 Lys14, 甲基化方式与组蛋白赖氨酸的甲基化在许多方面类似<sup>[13]</sup>. 为了解 SET 酶的催化机理, Trievel 等<sup>[14,15]</sup>报道了 pLSMT 与一/多赖氨酸和 *S*-腺苷甲硫氨酸相结合的多个晶体结构. 在全部 SET 区域甲基转移酶中, 被转移的甲基在碳原子中心发生构象翻转, 表明甲基转移是一个类似 SN2 的过程. 更重要的是, 合因子 *S*-腺苷甲硫氨酸和包含赖氨酸残基的底物肽酶分别从两侧接近 pLSMT, 并且被酶中间一条狭窄的疏水通道所连接<sup>[16~18]</sup>. 结构研究已经确认三个活性残基 Ser221, Asp239 和 Tyr287 在调节酶催化功能中起重要作用. 这三个活性残基主要通过 Ser221, Asp239 的羰基和 TYR287 的羟基同时提供三个 CH—O 氢键(CH···O)分别从三个方向与 AdoMet 被转移的甲基作用(图 1)<sup>[19]</sup>.

然而, 甲基转移的详细反应路径, 尤其是 SET 区域甲基转移酶如何利用活性位点的氢键催化反应进行, 仍然还不是很清楚. 本文中用理论计算的方法来研究 SET 区域酶 pLSMT 催化赖氨酸多重甲基化的机理, 并考虑溶剂化效应和酶环境对甲基转移反应的能垒的影响. 我们期望这一研究能够使我们对于甲基转移酶的反应机理有更进一步的了解.

## 1 体系和计算方法

本文涉及的所有从头量子(*ab initio* quantum me-

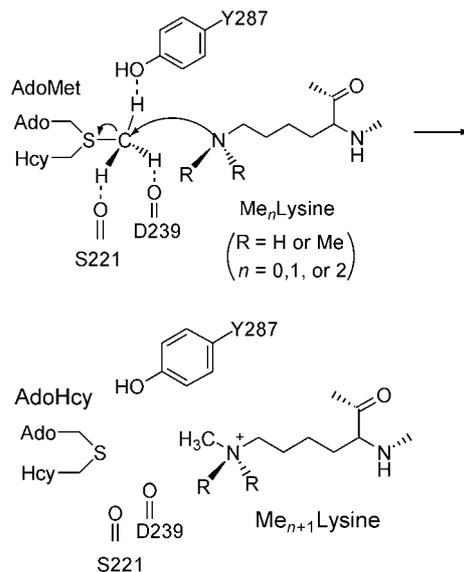


图 1 SET 区域甲基转移酶的三个活性残基 Ser221, Asp239 和 Tyr287 在催化转移甲基中的作用

Figure 1 Catalytic role of the methyl transfer pore in the SET domain active site

chanical)化学计算和密度泛函(DFT)计算都是在高斯 98 (Gaussian 98)程序中进行的<sup>[20]</sup>. 对气相反应, 分别对反应物(R), 离子对复合物中间体(CP), 过渡态(TS), 和产物(P)在 MP2/6-31G\*和 B3LYP/6-31G\*理论水平进行结构优化. 势能面的局部能量最低点和第一鞍点由振动频率的分析方法所识别, 并观察过渡态虚频振动的运动模式. 过渡态进一步用内在反应坐标(IRC, the intrinsic reaction coordinate)程序确认, 也就是在势能面上得到的结构能够顺利的连接反应物和产物. 用 Truhlar 等<sup>[21,22]</sup>发展的 SM5.42R 溶剂化模型来参数化复制中性和带电荷分子的实验水合自由能, 用于估计溶剂化效应.

以 pLSMT 与赖氨酸、AdoMet 所组成的复合物的晶体结构(PDB 登录号 1OZV)作为结合酶催化反应的起始结构. 活性位点的三个重要残基在计算中作为甲基转移酶催化效应的来源. 在目前的理论水平下, 为了使计算体系更加合理可行, 在计算模型中, 包含侧链羧基的 Ser221, Asp239 主要用乙酰胺代替, 而 Tyr287 用 phenol 代替. 酶催化反应模型体系在计算中首先用 B3LYP/6-31G\*方法进行优化. 由于 DFT 在处理诸如 CH—O 氢键之类的弱相互作用时具有缺陷, 对这些体系进一步使用 ONIOM (Our *N*-layered Integrated molecular Orbital + molecular Mechanics)方法进行分层计算. ONIOM 方法又叫洋葱算法, 是由 Morokuma 等<sup>[23~28]</sup>提出来的一个类似洋葱剥皮的方法, 允许量子化学和分子模拟分层对计算体系进行处理. 在 ONIOM 计算中, 计算体系被分成两层: (1)第一层使用“高水平”的计算方法(MP2/6-31G\*); (2)第二层使用“低水平”的方法(B3LYP/6-31G\*).

真实体系的能量由三个互相依赖的计算得到:

$$E(\text{ONIOM, real}) = E(\text{high, model}) + E(\text{low, real}) - E(\text{low, model})$$

## 2 结果和讨论

图 2 显示了由 pLSMT 催化多重赖氨酸甲基化的一般步骤. 目标组蛋白赖氨酸, 它的甲基化产物和甲基供体分别用乙胺, *N*-甲基乙胺, *N,N*-二甲基乙胺和三甲基硫表示. 甲基化步骤如下: 乙胺去质子化产生乙胺, 乙胺进攻三甲基硫完成一甲基化过程. 接下来, 去质子化和多重甲基化以同样的方式进行.

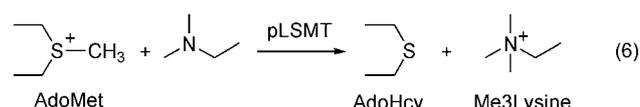
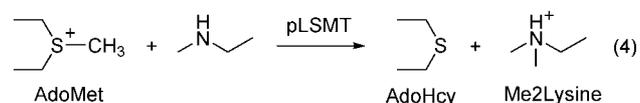
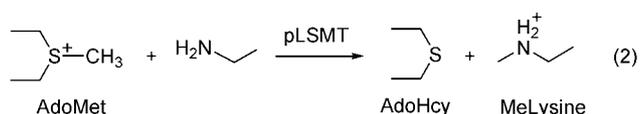


图 2 pLSMT 催化多重赖氨酸甲基化的一般步骤

Figure 2 The general step for methylation catalyzed by pLSMT

对于去质子化的问题, 已有研究表明 SET 区域 HKMTs (Histone lysine methyltransferases), 如 SET7/9 和 DIM-5 在 pH=8 或者更高 pH 值时具有较高的活性<sup>[29,30]</sup>. 因为在 HKMTs 催化过程中一直没有检测到催化碱受体, 目前一致认为底物赖氨酸侧链在 AdoMet 存在时以去甲基化的方式连接到 HKMTs 的 SET 区域<sup>[7,31]</sup>. 最近的 QM/MM 研究表明水通道的出现对去质子化步骤来说是必须的<sup>[32]</sup>. 另外, 许多实验证据表明甲基转移过程对许多甲基转移酶, 如苯酚-O-甲基转移酶 (Catechol *O*-Methyltransferase) 和 N6-腺嘌呤 DNA 甲基转移酶 (N6-Adenine DNA Methyltransferase) 等来说都是

速控步骤<sup>[33-36]</sup>. 因此本文研究主要集中在甲基转移步骤上.

### 2.1 气相反应

对所有复合反应物(R), 中间体(CP), 过渡态(TS), 和产物(P)的起始结构进行优化. 所有在 B3LYP/6-31G\* 和 MP2/6-31G\* 理论水平计算得到的能量值都总结在表 1 中. 对每一个甲基化步骤, 当三甲基硫和胺彼此靠近的时候, 在势能面上存在两个稳定的复合物. 一个通过分别来自三个不同甲基的氢在三甲基硫和胺间形成三个 C—H···N 氢键 (CP $n$ ,  $n=1, 2, 3$  分别指一、二、三甲基化); 另外一个则是胺与其中一个甲基靠近 (CP $n'$ ,  $n=1, 2, 3$ , 同上). 前者是甲基化反应的预反应复合物, 能量比后者要高 20.9 kJ/mol. 甲基化过程中, 稳定的复合物和对应的过渡态优化过的几何结构显示在图 3 中. 在复合物 CP $n'$  中, 三个 C—H···N 氢键在 B3LYP 水平计算得到的氢和氮原子间距离对 CP1', CP2', CP3' 分别为 0.2596, 0.2695, 0.2915 nm (在 MP2 水平分别为 0.1808, 0.1809, 0.1810 nm). C···N 间距离在 B3LYP 水平对 CP1', CP2' 和 CP3' 分别为 0.2994, 0.3023, 0.3041 nm (在 MP2 水平分别为 0.1808, 0.1809, 0.1810 nm). 在过渡态中, 即将断开的 S···C 距离对 TS1, TS2, TS3 在 B3LYP 水平和 MP2 水平分别为 0.2299, 0.2272, 0.2263 nm 和 0.2252, 0.2216, 0.2191 nm. 即将形成的 C···N 距离在 B3LYP 水

表 1 气相中计算得到的能量值(hartrees, 1 hartree = 2625.5 kJ/mol)

Table 1 Calculated electronic energies (hartrees, 1 hartree = 2625.5 kJ/mol) for each species in the gas phase

	B3LYP/6-31G*	MP2/6-31G*
Ethanaminium	-135.53740	-135.04152
Trimethylsulfonium	-517.67033	-516.62484
Ethanamine	-135.17073	-134.67553
CP1	-652.86242	-651.32519
CP1'	-652.85534	-651.31617
TS1	-652.83889	-651.28988
Dimethylsulfane	-478.01381	-477.12111
<i>N</i> -Methylethanaminium	-174.85574	-174.20963
<i>N</i> -Methylethanamine	-174.47953	-173.83402
CP2	-692.16812	-690.48264
CP2'	-692.16227	-690.47380
TS2	-692.14772	-690.45114
<i>N,N</i> -Dimethylethanaminium	-214.17015	-213.37661
<i>N,N</i> -Dimethylethanamine	-213.78859	-212.99542
CP3	-731.47488	-729.64309
CP3'	-731.46984	-729.63459
TS3	-731.45597	-729.61455
<i>N,N,N</i> -Trimethylethanaminium	-253.48152	-252.54337

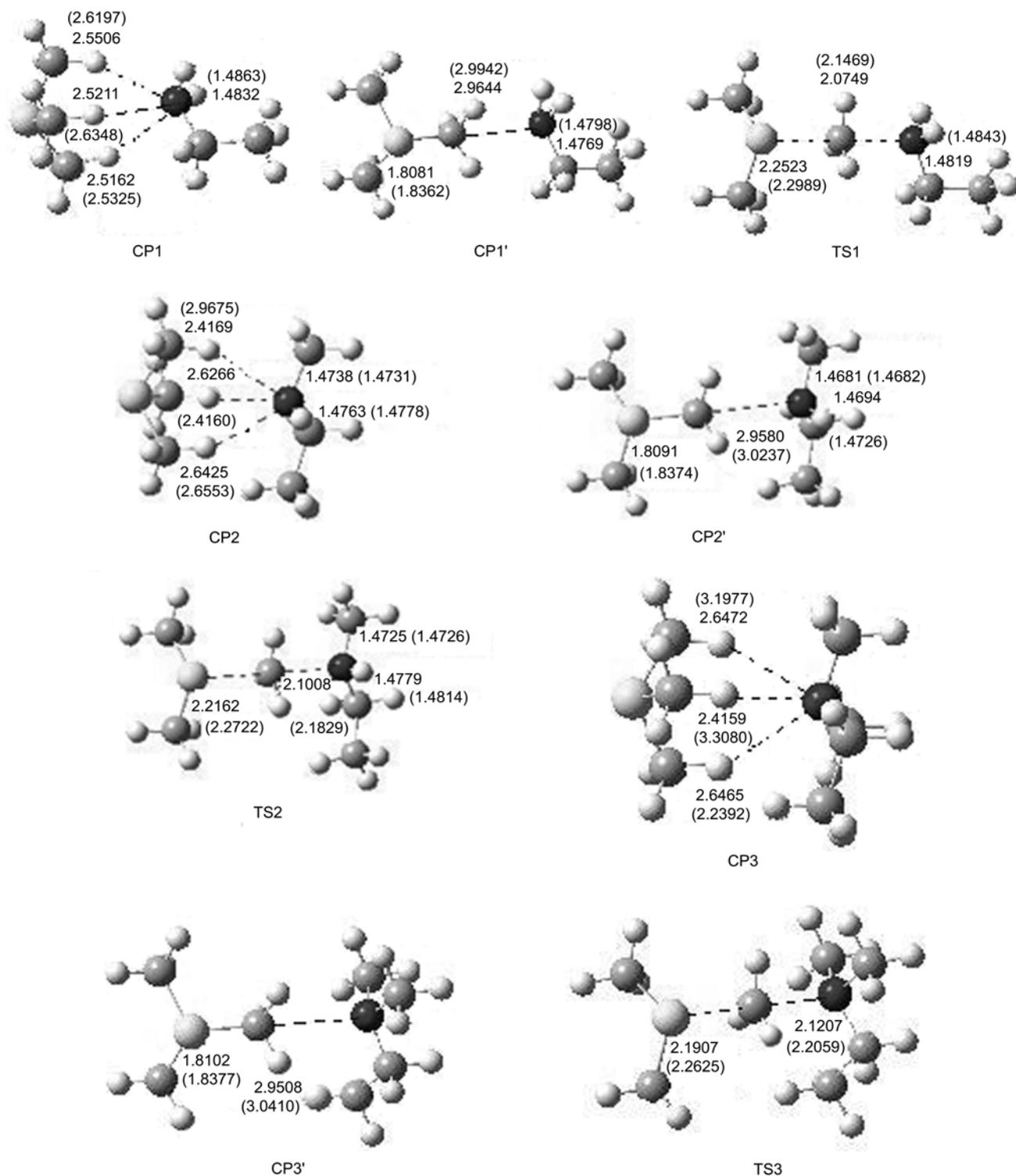


图3 在 B3LYP/6-31G\* 和 MP2/6-31G\* 理论水平优化得到的几何结构和几何参数 ( $10^{-1}$  nm)

Figure 3 Optimized geometrical parameters ( $10^{-1}$  nm) at the B3LYP/6-31G\* (in brackets) and MP2/6-31G\* levels of theory

平对 TS1, TS2, TS3 分别为 0.2147, 0.2183, 0.2206 nm (在 MP2 水平分别为 0.2075, 0.2101, 0.2121 nm).

计算得到的势能剖面显示在图 4 中, 三甲基铈和胺之间的甲基转移明显是放热过程. 计算反应能垒  $[\Delta E = E_{\text{tot}}(\text{CP}n) - E_{\text{tot}}(\text{R}n)]$  在 MP2/6-31G\* 水平对一、二、三甲

基化分别是 -65.3, -60.7, -59.8 kJ/mol, 随着甲基化程度依次升高. 从一甲基化到三甲基化, 反应能垒  $[\Delta E_a = E_{\text{tot}}(\text{TS}n) - E_{\text{tot}}(\text{CP}n)]$  对一、二、三甲基化分别为 92.9, 82.4 和 74.9 kJ/mol, 随反应过程逐渐降低. 这能够从甲基的供电子效应得到解释: 从三甲基铈到胺的甲基

转移过程从某种程度上可以被看作是 SN2 类型的反应, 胺是一个亲核试剂, 而三甲基硫中的二甲基硫醚部分可以看作是离去基团<sup>[37]</sup>. 随着反应的进行, 亲核试剂从最初的伯胺(ethanamine), 仲胺(*N*-methylethanamine)到叔胺(*N,N*-dimethylethanamine), 亲核能力逐渐增加, 从而有利于甲基转移.

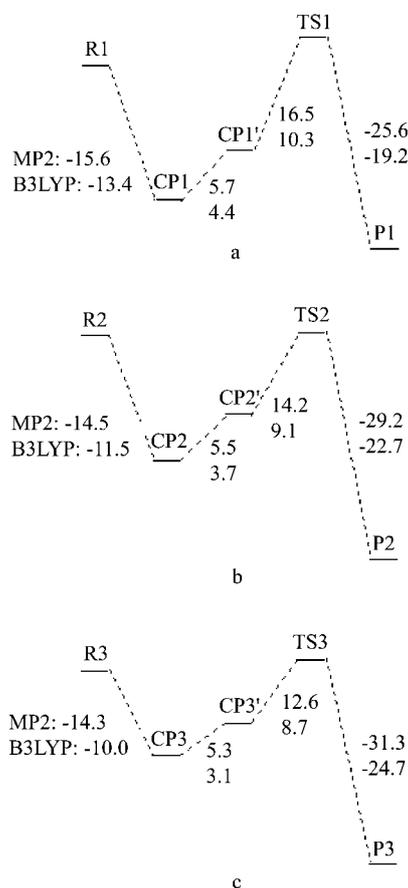


图4 气相中计算得到的势能剖面图: 三甲基硫与(a)乙胺; (b) *N*-甲基乙胺; (c) *N,N*-二甲基乙胺

**Figure 4** Calculated schematic potential energy profiles (kcal/mol, 1 kcal=4.18 kJ) in the gas phase for the reaction between trimethylsulfonium and: (a) ethanamine; (b) *N*-methylethanamine; (c) *N,N*-dimethylethanamine

密度泛函(DFT)在 B3LYP/6-31G\*水平也同样用来计算体系. 计算反应能量分布对一、二、三甲基化分别为-56.1, -48.1 和-41.8 kJ/mol, 比 MP2 结果分别高 9.2, 12.6 和 18.0 kJ/mol. 不过, 这两个不同的计算方法得到结果具有相同的变化趋势: 随着甲基化的进行, 反应放热越来越少. 反应能垒也有类似的变化, 尽管 DFT 计算得到的能垒大都比 MP2 得到的低 16.7~25.1 kJ/mol.

按照 Peralkyala<sup>[38]</sup>对胞核嘧啶甲基化的研究表明, MP2 计算稍微有些高估 C—S 键的稳定性和 C—S 键形成的能垒, 而 B3LYP 计算对这些值可能低估. 因此, 对

我们的体系来说实际反应能垒很可能位于 MP2/6-31G\* 和 B3LYP/6-31G\*结果之间.

## 2.2 溶液电介质中的反应

溶剂化效应在有机反应中起着重要的作用. 我们考虑两种具有不同介电常数的溶剂, 水( $\epsilon=78.5$ )和氯仿( $\epsilon=4.8$ )对反应体系的影响. 图 5 显示了由 SM5.42R 溶剂化模型得到的反应能量剖面, 十分不同于气相中得到的图. 在水溶液中, 两种稳定复合物的能量差异仅仅在 1.7~9.6 kJ/mol 之间, 明显低于气相中的差异. 由此可以看出, 溶剂化效应导致复合物的去稳定化.

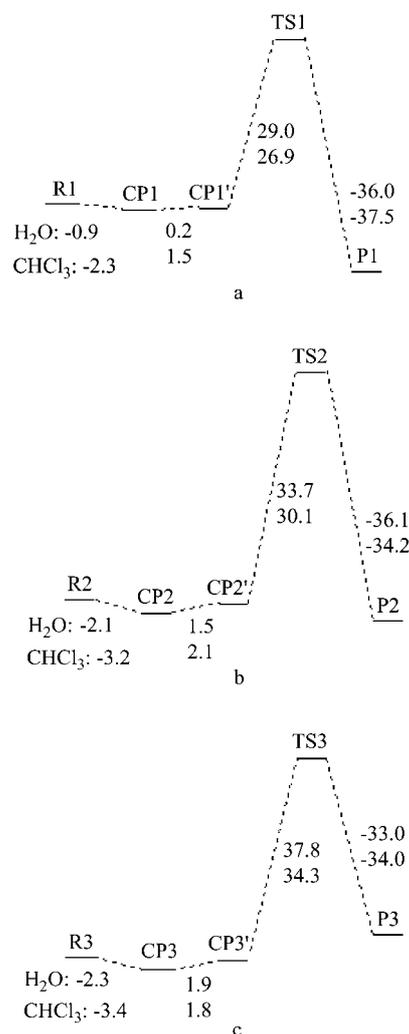


图5 使用 SM5.42R 方法在溶液(水和氯仿)中得到的反应势能剖面(kcal/mol, 1 kcal=4.18 kJ): 三甲基硫与(a)乙胺; (b) *N*-甲基乙胺; (c) *N,N*-二甲基乙胺

**Figure 5** Calculated reaction profiles using the SM5.42R method in the solution (water and chloroforms) for the reaction between trimethylsulfonium and: (a) ethanamine; (b) *N*-methylethanamine; (c) *N,N*-dimethylethanamine (kcal/mol, 1 kcal=4.18 kJ)

复合物 CP<sub>*n*</sub>计算总能量只是稍微的高于(0.8~7.9

kJ/mol)复合物 CP<sub>n</sub> 的能量. 此外, 水溶液对一、二、三甲甲基化的反应能量分别是 3.8, 8.8, 9.6 kJ/mol, 这与气相的高放热反应相比, 也发生了很大的变化. 反应的热动力学性质表明溶剂化效应对多重甲基化来说是不利的. 相似的变化趋势也反映在反应能垒方面, 在水溶液中对一、二、三甲甲基化分别是 121.3, 141.0, 158.2 kJ/mol. 这些能垒值明显高于气相中对应的数值, 但是随着甲基化的进行逐步升高, 与气相反应的变化趋势相反.

对氯仿溶液来说, 我们的计算结果表明反应能量和反应能垒对溶剂介电常数的改变并不敏感. 在氯仿溶液中反应能垒对一、二、三甲甲基化分别是 112.5, 125.9, 143.5 kJ/mol, 明显比水溶液中对应的值低大约 8.8~15.1 kJ/mol (图 3). 这个结果表明溶剂极化率降低, 有利于甲基转移的进行.

**表 2** 使用 SM5.42R 模型在水(78.3)和氯仿(4.0)中计算得到的能量(kcal/mol, 1 kcal=4.18 kJ)

**Table 2** Calculated energies using the SM5.42R with the dielectric constant of the water (78.3) and chloroforms (4.0) (kcal/mol, 1 kcal=4.18 kJ)

	Water	CHCl <sub>3</sub>
Ethanaminium	68.153	83.580
Trimethylsulfonium	93.497	101.852
Ethanamine	-16.377	-15.755
CP1	76.180	83.830
CP1'	76.422	85.301
TS1	105.411	112.191
Dimethylsulfane	-11.195	-13.969
N-Methylethanaminium	76.648	88.631
N-Methylethanamine	-17.220	-17.620
CP2	74.193	81.006
CP2'	75.663	83.144
TS2	109.364	113.293
N,N-Dimethylethanaminium	84.477	93.065
N,N-Dimethylethanamine	-18.215	-19.870
CP3	73.031	78.546
CP3'	74.888	80.323
TS3	112.700	114.602
N,N,N-Trimethylethanaminium	90.878	96.531

### 2.3 CH—O 氢键(CH···O)的作用

在 B3LYP/6-31G\* 和 ONIOM (MP2/6-31G\*: B3LYP/6-31G\*) 基组水平, 对 pLSMT 所催化的转移一个甲基的反应过程使用上述的模型体系进行研究(看计算方法部分). 曾有报道指出, 在甲基化进行前, Tyr 残基作为碱基负责 ε-ammonium 的去质子化作用<sup>[15,29,30,39]</sup>. 然而, 在生理条件下(pka=9.95), phenol 主要以中性形式存在. 在一系列酶反应中, 都是赖氨酸的 NH<sub>2</sub> 基作为碱通过夺

取一个质子进而促进反应进行的<sup>[40]</sup>. 因此, 在目前情况下, Tyr 不可能作为去质子化角色促进反应的进行的, 这与 Raymond 等<sup>[14]</sup>的观点一致. 因此, 中性的 phenol 作为 Tyr-287 残基的模型.

图 6(a)中, 在反应物状态中, 优化得到三甲基铈离子和去质子化的赖氨酸相对于临近酶残基的位置. 从图中可以看出, 在酶反应复合物(ECP)中, 三个活性残基通过 CH—O 氢键从三个不同的方向分别接近三甲基铈离子. 在 B3LYP 水平, 被转移甲基上的氢和两个氨基酸残基的羰基(Ser-221 和 Asp-239)氧之间的距离分别为 0.2321 和 0.2317 nm. Tyr-287 与三甲基铈的两个甲基作用, 形成两个 CH—O 氢键(CH···O), 在 B3LYP 水平长度分别为 0.2444 和 0.2518 nm. 在 ONIOM (MP2/6-31G\*: B3LYP/6-31G\*)水平优化得到相似的优化几何结构(图 6(a)). 对应上述四个 CH—O 氢键(CH···O)优化后对应的长度分别为 0.2320, 0.2320, 0.2405, 0.2573 nm. 酶中的反应复合物类似于在气相中观察得到的 CP1'构象, 表 3 中的重要参数的数据和气相中得到的数据基本一致, 只有角 S—C—N 的键角稍微偏移 180°. 非常明显, Tyr-287 残基占据三甲基铈的下方, 通过位阻作用禁止气相中得到的稳定复合物(CP1)在模型体系中存在. 在实际反应中, 当 AdoMet 和赖氨酸都是大基团离子的时候, 位阻效应会更加明显.

图 6(b)中显示了模型酶体系的过渡态结构(ETS). 除了另外附加的残基, 过渡态结构非常类似于气相过渡态的结果. 在 B3LYP/6-31G\*水平, Tyr-287, Asp-239, Ser-221 主要通过三个氢键与三甲基铈离子正在转移的甲基作用, 这三个氢键的长度分别为 0.2494, 0.2250, 0.2428 nm, 比铈离子与其它甲基的作用要强. 使用 ONIOM(MP2/6-31G\*: B3LYP/6-31G\*)方法优化过渡态后, 对应的氢键键长缩短(0.2494→0.2493, 0.2250→0.2187, 0.2428→0.2225 nm).

因为两个羰基氧原子和三个活性残基所提供的羟基与三甲基铈离子具有相反的电荷分布, 这三个残基与反应单元(铈离子和胺)有着很强的相互作用. 实际上, B3LYP/6-31G\*计算表明 ECP 和 ETS 比 CP1'和 TS1 能量上分别要低 170.7, 159.4 kJ/mol. 氢键对反应物和过渡态的作用基本上是对等的. 在 ONIOM(MP2/6-31G\*: B3LYP/6-31G\*)水平, 模型体系(ECP→ETS)的反应能垒大约是 82.8 kJ/mol, 与实验上由单过渡态理论 $[k(T) = (k_B T/h) \exp(-G^\ddagger/RT)]$ 估算得到的数据 87.4 kJ/mol 非常的接近<sup>[15]</sup>. 如果以 CP1'作为甲基化反应的起始点, 与水溶液中的情况相比, 酶环境将导致活化能降低大约 37.7 kJ/mol. 然而, 活化能垒 ECP→ETS (82.8 kJ/mol)仍然比气相中 CP1'→TS1 (69.0 kJ/mol)的活化能垒高, 因

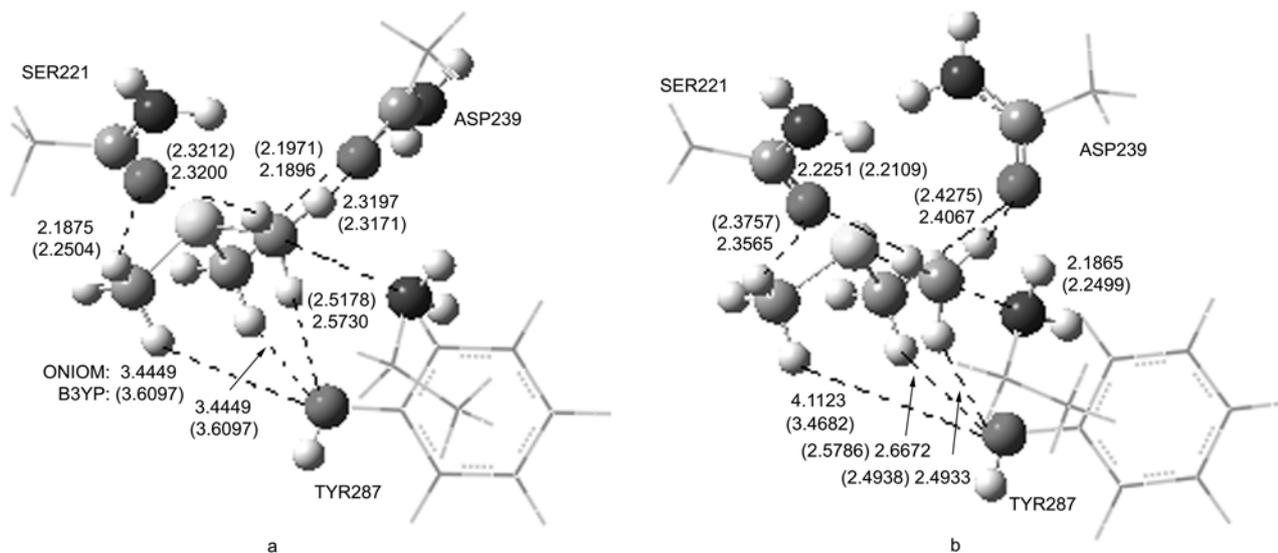


图6 pLSMT 活性残基相对于基态(a)和过渡态(b)在 B3LYP/6-31G\*和 ONIOM(MP2:B3LYP)水平优化得到的相对位置. 由三个活性残基所提供的 CH—O 氢键被标为黑线(以  $10^{-1}$  nm 为单位). B3LYP 结果显示在括号中. 在 ONIOM 水平, 低基组区域标记为线状; 高基组区域被显示为球棍模型

Figure 6 Positions of three active-site residue models relative to: (a) the ground state; (b) the transition state optimized at B3LYP/6-31G\* and ONIOM(MP2:B3LYP) in pLSMT. The CH—O hydrogen bonds provided by the three active residues were indicated by broken black lines (units  $10^{-1}$  nm). The B3LYP results are shown in brackets. At the ONIOM level, low level region is shown by line; high level region is shown by ball and stick

表3 单甲基化在气相和酶反应中在不同基组下优化基态反应复合物和过渡态得到的能量值

Table 3 Structural and energetic characteristics of ground-state reactant complexes and transition states for the monomethylation in the gas phase and enzymatic orientation at different levels

	System	$r_{S-C}/(10^{-1} \text{ nm})$	$r_{C-N}/(10^{-1} \text{ nm})$	$\theta_{S-C-N}/(^{\circ})$	Energy/Hartrees
B3LYP/6-31G*	CP1'	1.836	2.994	179.156	-652.85534
	TS1	2.299	2.147	179.245	-652.83889
	ECP	1.826	3.210	165.483	-652.92036
	ETS	2.237	2.194	179.232	-652.89961
	phol				-307.46487
	ace				-209.21220
	CP1'	1.806	3.009	179.932	-651.61664
ONIOM(MP2/6-31G*: B3LYP/6-31G*)	TS1	2.273	2.070	179.753	-651.58883
	ECP	1.802	3.126	171.186	-653.68580
	ETS	2.260	2.066	177.286	-653.65420
	phol				-307.25280
	ace				-208.71802

此, 活性残基不能带来活化能的降低或者为过渡态提供额外的稳定作用.

因此, 对我们的甲基化反应来说, 酶的角色能够被描述如下: 一方面, 它为反应物(三甲基铈离子和胺)提供了一个没有溶剂的疏水环境. 正像 Bruice 等<sup>[32]</sup>指出的那样, 水分子只能定向地、直接地接近即将甲基化的赖氨酸. 结果, 活性残基催化的反应, 无论是从结构还是能量方面来说, 都非常类似于气相中的反应(如上). 另

一方面, 酶能够把亲核和亲电的底物通过临近活性残基的作用以一种非常完美的构象安排在一起, 几乎不需要重排就可以形成过渡态, 同时使得不利于反应过程的构象(例如 CPn)等通过位阻作用被禁止.

## 2.4 CH—O 氢键(CH...O)广泛存在于 HKMTs 中

SET 区域 HKMTs 序列相似性很低, 活性位点的残基不是完全保守的, 尽管保守残基对 HKMTs 的活性非常重要, 几乎是必不可少的<sup>[41]</sup>. 在病毒 SET 中, 保守残

基的突变导致酶与天然蛋白相比, 活性几乎完全丧失(90%)<sup>[42]</sup>. Tyr-287 就是存在于 pLSMT 中的这样一个保守残基. 通过酶功能的分析推测, 由活性位点残基提供的负电荷环境很可能对所有 SET 区域 HKMTs 来说都是必不可少的. 根据这条线索, 我们对所有已知的 SET 区域 HKMTs 结合 AdoMet 的体系在蛋白质数据库中通过识别相似的功能和结构进行了系统的搜索. 在所有搜索得到的结构中, Tyr 是活性位点附近唯一不变的残基.

图 7 显示了几个 SET 区域 HKMTs 与 AdoMet 结合的几个例子(Tyr 是粉色的). 图 7(a)显示的就是前面提到的菠菜 Rubisco LSMT. 图 7(b)显示一个单甲基酶 SET7/9<sup>[29,43]</sup> 通过络合 *S*-腺苷-*L*-甲硫氨酸(SAH, *S*-adenosyl-*L*-methionine)转移甲基到组蛋白 H3 赖氨酸 4 的特征. 其它例子包括脉孢菌 DIM-5, 一个组蛋白 H3 赖氨酸 9 甲基转移酶(H3-K9me, 图 7c)<sup>[30]</sup>; hSET8 络合 SAH 和组蛋白 lys-20 的三级结构(H4-K20me, Figure 7d)<sup>[44]</sup>; 二甲基化组蛋白 H3 赖氨酸 27 的病毒 SET (vSET) (H3-K27me, Figure 7e)<sup>[42]</sup>; 在异染色质组蛋白 H3 赖氨酸 9 甲基化中起重要作用的 G9a 人类组蛋白甲基转移酶(H3-K9me, Figure 7f)<sup>[45]</sup>. 在图 7 所示的所有结构中, SAH 的 S 中心和酶中活性残基的 O 之间的距离都在 3~5 Å 之间, 和在 pLSMT 中找到的类似. 尽管位点附近的残基不是绝对保守的, 但是起活化作用的由两个残基提供羰基, 一个保守的 Tyr 残基提供羟基所形成的三个 CH—O 氢键(CH...O)对于所有已知的 SET 区域来说都是必不可少的.

### 3 结论

用量子化学从头计算和密度泛函的方法研究了 pLSMT 催化的多重赖氨酸甲基化的机理. 我们的结果表明反应在气相中从能量上来说是可以直接进行的, 然而在溶液中因为溶剂化效应导致较高的能垒使反应难以发生.

在酶体系中, 三个活性残基在活性位点对反应底物有很强的诱导作用. 因为它们对反应物和过渡态有几乎对等的作用, 所以它们不能导致活化能的降低. 酶可能通过两条路径发挥它的催化作用. 一方面, 提供一个无溶剂的环境容纳反应底物, 溶剂只能定向地接近赖氨酸进行去质子化; 另外一方面, 酶能够把亲核和亲电的底物用高效的方式通过活性位点的残基组织成一个合理的构象空间, 这种构象使得从起始态到过渡态几乎不需要结构上的修正.

最后, 应该指明的是, 尽管 SET 区域 HKMTs 序列相似性很低、活性位点的残基不是完全保守的, 但由两

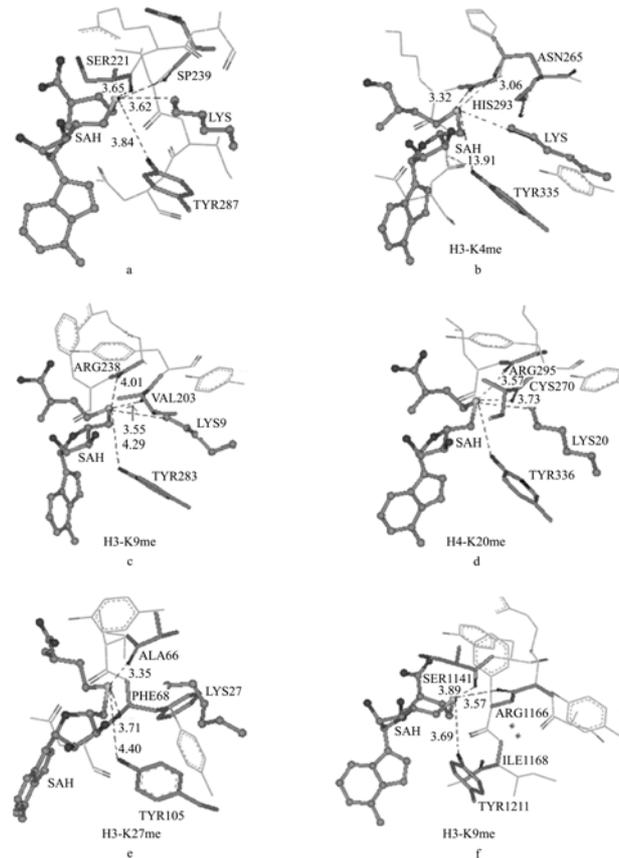


图 7 在 SET 区域中, SAH 和赖氨酸接近通道的几何结构和分子环境示例: (a) pLSMT, PDB 登录号 1ZOV; (b) SET7/9, PDB 1O9S; (c) DIM-5, PDB 登录号 1PEG; (d) hSET8, PDB 登录号 1ZKK; (e) vSET, PDB 登录号 2G46; (f) G9a, PDB 登录号 2O8J. 由活性残基所提供的 CH—O 氢键由断开的黑线标示. 距离以 Å 为单位; 红色 O, 蓝色 N, 灰色 C, 黄色 S. 不变量 Tyr 被标记为紫色

**Figure 7** Examples showing the geometry and molecular environment of the lysine access channel and *S*-adenosyl-*L*-methionine (SAH) in SET domains is shown for: (a) pLSMT, calculated in this text, PDB entry 1ZOV; (b) SET7/9, PDB entry 1O9S; (c) DIM-5, PDB entry 1PEG; (d) hSET8, PDB entry 1ZKK; (e) vSET, PDB entry 2G46 (the 17th structure of NMR); (f) G9a, PDB entry 2O8J. The probable CH—O hydrogen bonds provided by the active-site residues were indicated by broken black lines. The distances are indicated in Å; red O, blue N, gray C, yellow S. The invariant Tyr is indicated by purple labels

个活性位点残基提供羰基、一个保守的 Tyr 残基提供羟基所形成的 CH—O 氢键(CH...O)广泛存在于所有已知的 HKMTs 中. 因此, 我们对 pLSMT 催化机理的研究具有广泛的意义.

### References

- 1 Fischle, W.; Wang, Y.; Allis, C. D. *Curr. Opin. Cell Biol.* **2003**, *2*, 172.

- 2 Strahl, B. D.; Allis, C. D. *Nature* **2002**, *403*, 41.
- 3 Turner, B. M. *Cell* **1993**, *75*, 5.
- 4 Hake, S. B.; Xiao, A.; Allis, C. D. *Br. J. Cancer* **2004**, *90*, 761.
- 5 Schneider, R.; Bannister, A. J.; Kouzarides, T. *Trends Biochem. Sci.* **2002**, *27*, 396.
- 6 Rayasam, G. V.; Wendling, O.; Angrand, P. O.; Mark, M.; Niederreither, K.; Song, L.; Lerouge, T.; Hager, G. L.; Chambon, P.; Losson, R. *EMBO J.* **2003**, *22*, 3153.
- 7 Xiao, B.; Wilson, J. R.; Gamblin, S. J. *Curr. Opin. Struct. Biol.* **2003**, *13*, 699.
- 8 Cheng, X.; Collins, R. E.; Zhang, X. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 267.
- 9 Bannister, A. J.; Schneider, R.; Myers, F. A.; Thorne, A. W.; Robinson, C. C.; Kouzarides, T. *J. Biol. Chem.* **2005**, *280*, 17732.
- 10 Bernstein, B. E.; Humphrey, E. L.; Erlich, R. L.; Schneider, R.; Bouman, P.; Liu, J. S.; Kouzarides, T.; Schreiber, S. L. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 8695.
- 11 Santos-Rosa, H.; Schneider, R.; Bannister, A. J.; Sherriff, J.; Bernstein, B. E.; Emre, N. C.; Schreiber, S. L.; Mellor, J.; Kouzarides, T. *Nature* **2002**, *419*, 407.
- 12 Stancheva I. *Biochem. Cell Biol.* **2005**, *83*, 385.
- 13 Houtz, R. L.; Royer, M.; Salvucci, M. E. *Plant Physiol.* **1991**, *97*, 913.
- 14 Trievel, R. C.; Flynn, E. M.; Houtz, R. L.; Hurley, J. H. *Nat. Struct. Biol.* **2003**, *10*, 545.
- 15 Trievel, R. C.; Beach, B. M.; Dirk, L. M. A.; Houtz, R. L.; Hurley, J. H. *Cell* **2002**, *111*, 91.
- 16 Yeates, T. O. *Cell* **2002**, *111*, 5.
- 17 Xiao, B.; Jing, C.; Wilson, J. R.; Walker, P. A.; Vasisht, N.; Kelly, G.; Howell, S.; Taylor, I. A.; Blackburn, G. M.; Gamblin, S. J. *Nature* **2003**, *421*, 652.
- 18 Kwon, T.; Chang, J. H.; Kwak, E.; Lee, C. W.; Joachimiak, A.; Kim, Y. C.; Lee, J. W.; Cho, Y. *EMBO J.* **2003**, *22*, 292.
- 19 Couture, J.-F.; Hauk, G.; Thompson, M. J.; Blackburn, G. M.; Trievel, R. C. *J. Biol. Chem.* **2006**, *281*, 19280.
- 20 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Peterson, G. A.; Montgomery, J. A.; Rachavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzales, C.; Pople, J. A. *Gaussian 98*, Gaussian, Inc.: Pittsburgh, PA, **1998**.
- 21 Giesen, D. J.; Michael, Z. G.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1996**, *61*, 8720.
- 22 Zheng, Y. J.; Bruice, T. C. *J. Am. Chem. Soc.* **1997**, *119*, 8137.
- 23 Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 16, 1959.
- 24 Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- 25 Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (Theochem)* **1999**, *461*, 1.
- 26 Maseras, F. *Top. Organomet. Chem.* **1999**, *4*, 165.
- 27 Froese, R. D. J.; Morokuma, K. In *The Encyclopedia of Computational Chemistry*; ISBN: 0-471-96588-X. Eds.: Schleyer, P. V. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer III, H. F.; Schreiner, P. R., John Wiley, Chichester, **1998**, p. 1245.
- 28 Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419.
- 29 Wilson, J. R.; Jing, C.; Walker, P. A.; Martin, S. R.; Howell, S. A.; Blackburn, G. M.; Gamblin, S. J.; Xiao, B. *Cell* **2002**, *111*, 105.
- 30 Zhang, X.; Tamaru, H.; Khan, S. I.; Horton, J. R.; Keefe, L. J.; Selker, E. U.; Cheng, X. *Cell* **2002**, *111*, 117.
- 31 Cheng, X.; Collins, R. E. Zhang, X. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 267.
- 32 Zhang, X.-D.; Bruice, T. C. *Biochem.* **2007**, *46*, 5505.
- 33 Pues, H.; Bleimling, N.; Holz, B.; Wolcke, J.; Weinhold, E. *Biochem.* **1999**, *38*, 1426.
- 34 Hegazi, M. F.; Borchardt, R. T.; Schowen, R. L. *J. Am. Chem. Soc.* **1979**, *101*, 4359.
- 35 Gray, C. H.; Coward, J. K.; Schowen, K. B.; Schowen, R. L. *J. Am. Chem. Soc.* **1979**, *101*, 4351.
- 36 Rodgers, J.; Femec, D. A.; Schowen, R. L. *J. Am. Chem. Soc.* **1982**, *104*, 3263.
- 37 (a) Olmstead, W. N.; Brauman, J. I. *J. Am. Chem. Soc.* **1979**, *101*, 3715.  
(b) Shaik, S. S.; Schlegel, H. B.; Wolfe, S. *Theoretical Aspects of Physical Organic Chemistry, The SN2 Mechanism*, Wiley, New York, **1992**.  
(c) Minkin, V. I.; Simkin, B. Y.; Minyaev, R. M. *Quantum Chemistry of Organic Compounds-Mechanisms of Reactions*, Springer-Verlag, Berlin, **1990**.
- 38 Peralkyla, M. *J. Am. Chem. Soc.* **1998**, *120*, 12895.
- 39 Jacobs, S. A.; Harp, J. M.; Devarakonda, S.; Kim, Y.; Rastinejad, F.; Khorasanizadeh, S. *Nat. Struct. Biol.* **2002**, *9*, 833.
- 40 (a) Paetzel, M.; Dalbey, R. E. *Trends Biochem. Sci.* **1997**, *22*, 28.  
(b) Kenyon, G. L.; Gerlt, J. A.; Petsko, G. A.; Kozarich, J. W. *Acc. Chem. Res.* **1995**, *28*, 178.
- 41 Qian, C.; Wang, X.; Manzur, K.; Farooq, S. A.; Zeng, L.; Wang, R.; Zhou, M. M. *J. Mol. Biol.* **2006**, *359*, 86.
- 42 Manzur, K. L.; Farooq, A.; Zeng, L.; Plotnikova, O.; Koch, A. W.; Sachchidanand.; Zhou, M.-M. Karishma, L. *Nat. Struct. Biol.* **2003**, *10*, 187.
- 43 Xiao, B.; Jing, C.; Wilson, J. R.; Walker, P. A.; Vasisht, N.; Kelly, G.; Howell, S.; Taylor, I. A.; Blackburn, G. M.; Gamblin, S. J. *Nature* **2003**, *421*, 652.
- 44 Couture, J. F.; Collazo, E.; Brunzelle, J. S.; Trievel, R. C. *Genes Dev.* **2005**, *19*, 1455.
- 45 Wu, H.; Min, J.; Antoshenko, T.; Loppnau, P.; Weigelt, J.; Sundstrom, M.; Arrowsmith, C. H.; Edwards, A. M.; Bochkarev, A.; Plotnikov, A. N. *To be Published*.