

中心科学实验 III- 理论化学部分

I. Caffeine 计算实验

一、 实验目的

从茶叶中提取咖啡因让我们更容易了解了咖啡因。咖啡因是目前世界上使用最为广泛的刺激剂，像是咖啡，可乐，茶，巧克力等。而跟我们接触到它的容易程度相比，大部分人对咖啡因的了解都是非常不充分的。咖啡因具有多种有益的药理活性，例如用于神经退行性疾病的治疗、镇痛和心血管健康。尽管是一种广泛研究的活性天然产物，但对咖啡因药理作用的机制理解尚不完整。虽然咖啡因的几个分子靶标（如腺苷受体和磷酸二酯酶）几十年来已经为人所知，开展了许多药物化学项目，但不断发现黄嘌呤的新蛋白质相互作用，为未来的药物化学提供了潜在的改进药理学理解和分子基础。更详细内容参见 J. Med. Chem. 2021, 64, 7156 – 7178。

二、 实验原理

咖啡因的代谢在肝脏中发生，由细胞色素氧化酶P450（特别是CYP1A2）酶系统氧化，形成三种不同的二甲基黄嘌呤，这三种二甲基黄嘌呤对身体有不同的作用（图1）。副黄嘌呤PX（1, 7-二甲基黄嘌呤，80%）能够加速脂解，导致血浆中的甘油及自由脂肪酸的含量增加。可可碱TB（10%）-能够扩张血管，增加尿量。茶碱TP（5%）-舒缓支气管平滑肌，被用作治疗哮喘。过去由于代谢位点、代谢反应路径不明，结合模式存在很多争议。优化可能反应的稳定点，计算结合自由能，探讨CYP1A2 酶的催化活性，显然对它的人体代谢认知具有指导意义。本实验对其结合位点进行计算模拟，以期获得最稳定构象，绘制3D结构图。由于研究体系很大，采用ONIOM模型。

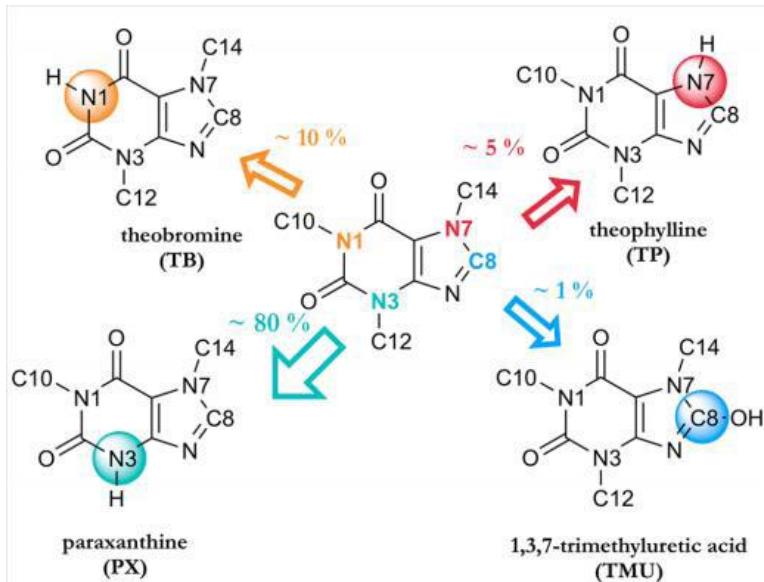


图 1. 咖啡因代谢示意图

在实际的研究中，真实的分子体系可能包含大量的原子，对于这种体系而言，由于计算资源和计算时间的限制，若使用较高精度的模型方法进行模拟往往是不实际的，ONIOM 方法的思想在于把一个分子体系分为若干个区域（也称为层，通常可以分成 2 层和 3 层），不同的层能够使用不同精度的方法进行计算。若划分为 2 层：则通常划分为 High Layer（高层）和 Low Layer（低层）。其中高层包含感兴趣的研究区域（对于化学反应而言，则是断键和成键区域），这部分需要使用最昂贵的计算方法；而低层则是其余的部分，这部分则使用便宜的计算方法。

由于在 ONIOM 计算中，不同的层使用不同精度的方法进行计算，因此相比于传统计算，对于体系的电子能，ONIOM 有更复杂的计算步骤，中每一个圆表示相应的方法在相应的层系统下计算得到的电子能，加号表示在 ONIOM 能量表达式中加上这个能量项，减号则表示减去这个能量项。

对于 2-layer 体系：

$$E(\text{ONIOM}) = E(\text{Low, Real}) + E(\text{High, Model}) - E(\text{Low, Model})$$

为了进行一个 ONIOM 计算需要进行如下三个步骤：

利用 GaussView 对分子体系进行分层

构建 ONIOM 的输入文件

结果分析

三、计算软件与方法

Gaussian 程序是目前最普遍使用的量子化学计算软件，Gaussview 是一个 gaussian 专门配套使用的软件，用于建模和显示图形结果。VMD 程序帮助 ONIOM 分层。

四、实验步骤

1. 以 4-联苯甲醛为例，学习如何使用构建 ONIOM 分层。

打开 Gaussview 程序，如图 2 所示，在 GaussView 中利用建模工具。

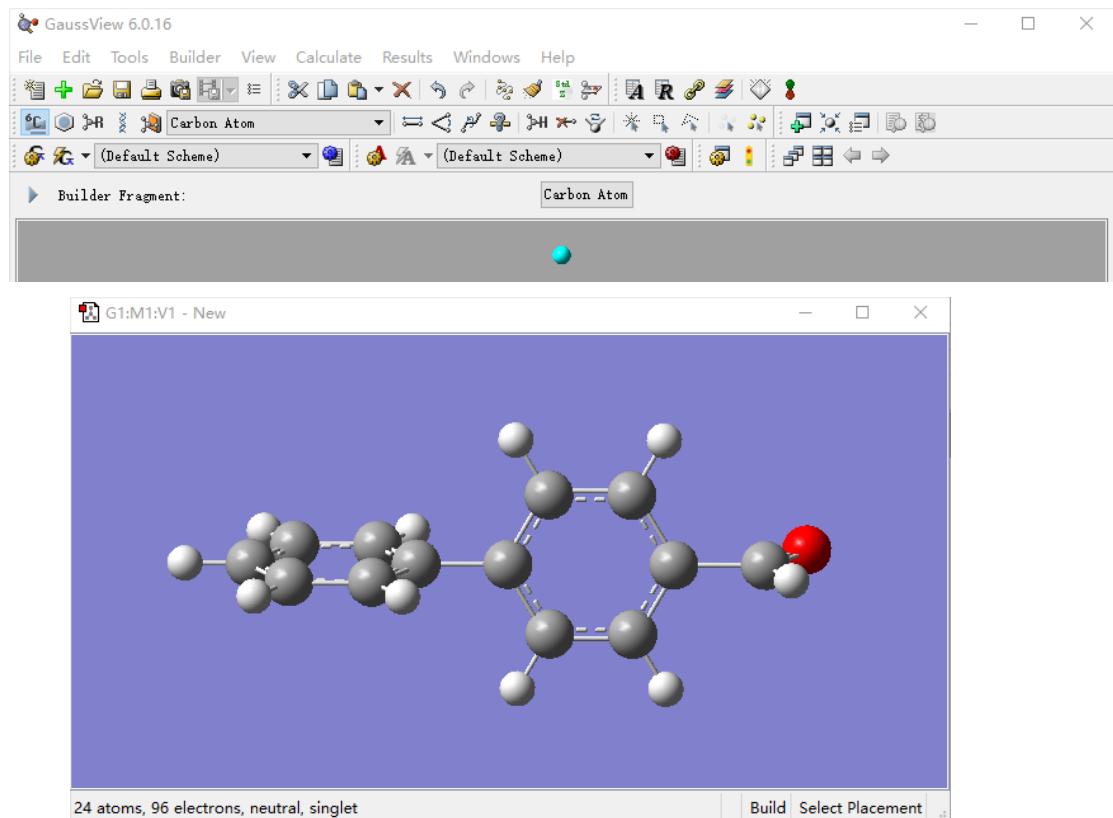


图 2 通过苯环工具和功能团工具构建 4-联苯甲醛图示

2. 进行 ONIOM 编辑。在 GaussView 中菜单栏 Tools 中下拉到 Atom Groups。此时可以看出所有的原子都处于高层。

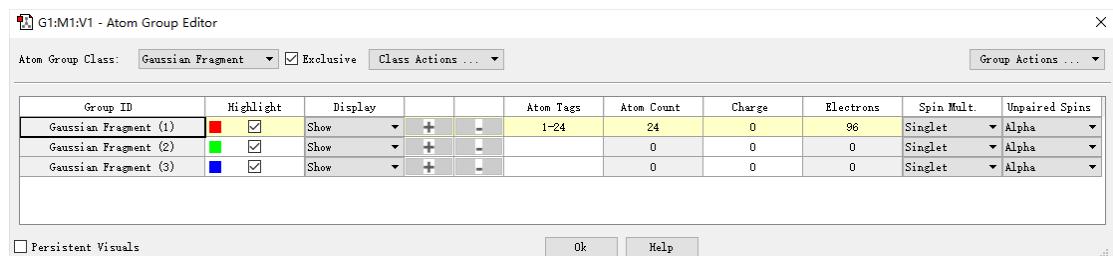


图 3 Atom Group Editor 设置窗口

点击圈选工具，选中处于想要设置为低层（Low）的原子（也就是 1-21 号原子，原子标号可以通过点击 View 下拉菜单中 Labels 选项显示），然后点击“ONIOM layer (Low)”所在行的“+”号，此时显示 1-21 号原子处于低层，即完成我们想要的分层。

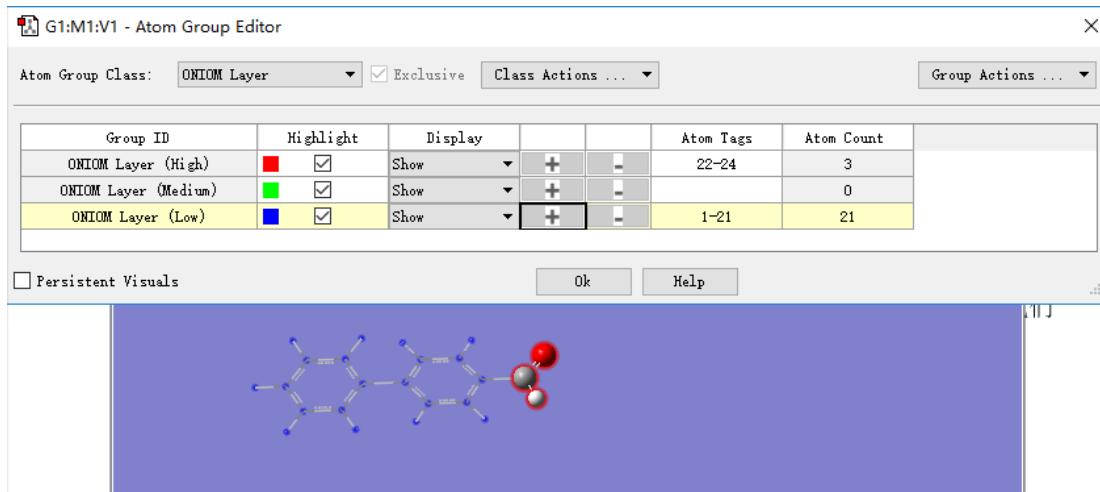


图 4 分层结构示意图

3. 点击原子团编辑对话框中下面的“OK”关闭对话框，然后点击工具栏中 Calculate 下拉菜单中的 Gaussian Calculation Setup 选项，调出 ONIOM 计算路径设置对话框。在 Method 中勾选 Multilayer ONIOM model.

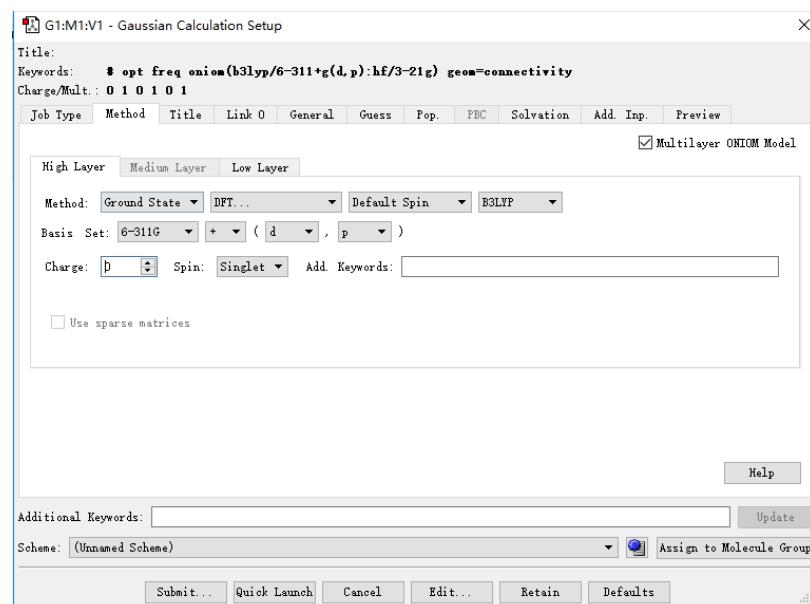


图 5, ONIOM 计算设置

4. 设置 Low Layer 计算方法与基组(HF/3-21G)，High Layer (B3LYP/6-

311+g**)生成的 gjf 文件。

在这里，我们注意到出现了三对电荷/自旋多重度，它们分别表示用不同精度方法计算不同层时采用的设置，依次为：1. 低水平方法计算真实体系；2. 高水平方法计算高层部分；3. 低水平方法计算高层部分。

生成 Gaussian 输入文件，并编辑保存输入文件 ONION-1.gjf。

提交输入文件，进行计算

结果分析：

给出了输出文件中电子能量的结果，根据上面的公式，我们能够得到推测的 ONIOM 电子能。

```
ONIOM: saving gridpoint 1
ONIOM: restoring gridpoint 3
ONIOM: calculating energy.
ONIOM: gridpoint 1 method: low system: model energy: -113.221187371910
ONIOM: gridpoint 2 method: high system: model energy: -114.541163407344
ONIOM: gridpoint 3 method: low system: real energy: -569.780239052889
ONIOM: extrapolated energy = -571.100215088322
ONIOM: calculating first derivatives.
ONIOM: calculating second derivatives.
ONIOM: calculating electric field derivatives.
ONIOM: Integrating ONIOM file 4 number 619
ONIOM: Dipole = -1.16277035D+00 6.84533347D-01 2.91991431D-01
ONIOM: Polarizability= 1.94433086D+02-3.83752087D+00 8.44242177D+01
          -2.14198669D+00 3.05377228D+01 8.62001098D+01
ONIOM: Dipole moment (Debye):
      X=    -2.9555    Y=     1.7399    Z=     0.7422    Tot=     3.5090
ONIOM: Integrating ONIOM file 5 number 695
```

咖啡因实验步骤：

1. 从RCSB 蛋白质数据库中获取2HI4，这是个蛋白受体与奈黄酮结合的复合体，删除结晶水和奈黄酮，构建原始CYP1A2 初始结构。
2. 咖啡因为配体与CYP1A2 受体进行分子对接。
3. 对最稳定的构象进行能量计算，获得能量最低结合位点。开展数据处理，完成结构显示。

五、实验讨论

1. 三层(High, Medium, low Layer)电子能量应如何获得？

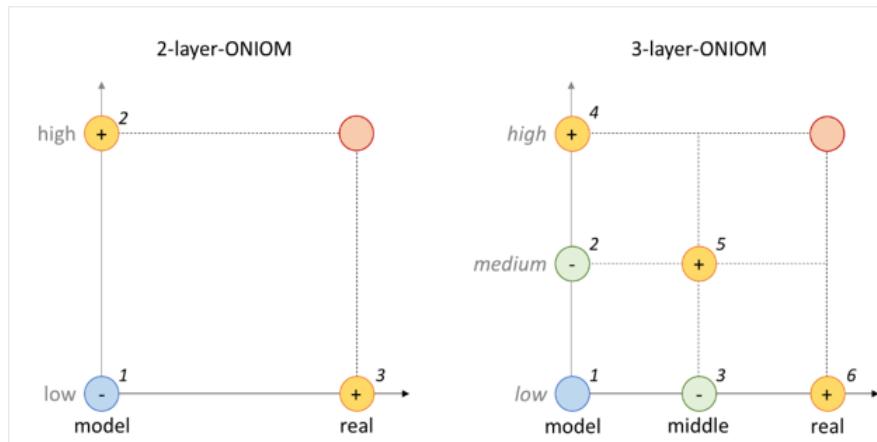


图 6. ONIOM 电子能量图示

六、思考题：

II. 鲁米诺计算实验

一、实验目的

鲁米诺(Luminol)是一种著名的化学发光试剂，因法医学上用以检验犯罪现场的痕量血迹而闻名。同时也被广泛用于多种微量无机物和有机物的检验。但机理推断过程中的中间产物以及生成物并未被表征或验证，又如鲁米诺化学发光反应中，发光的物质究竟是什么？都有什么结构存在？是什么样的空间排布？发光试剂鲁米诺有没有其他的不同结构？而在整个反应机理中，是否有其他作用的影响？这些问题的存在，会对化学发光反应的深入研究造成一定的困难。解决这些问题，也就是本实验研究内容的意义所在，这将会为鲁米诺化学发光反应机理的研究奠定扎实的基础，对机理的深入分析提供重要的理论依据和实用价值。另外实验中证明鲁米诺的化学发光反应量子产率是 1-2%，这就表明仍然有可能提高它的化学发光量子产率以及提高这种方法的灵敏度。在提高鲁米诺化学发光效率的方法中，设计和合成鲁米诺衍生物吸引了众多研究。修饰的方法包括氨基的取代和烷基化以及苯环的上氢原子的取代，取代基环的增多使得分子的共轭性增强。酰肼衍生物是一类具有较高反应活性的有机化合物，到目前为止已有大量的非取代、单取代和双取代酰肼衍生物被合成研究。鉴于实验中为了获得一

一个化学发光强度较高的化合物，首先要合成，然后再进行化学发光检测，这个过程是一个复杂的过程。在分子设计以及其性质的预测中理论计算由于高的准确性和低的成本被认为是一个非常有用的方法。

二、实验原理

密度泛函方法是一种研究多电子体系电子结构的方法，在 Kohn-Sham DFT 的框架中，最难处理的多体问题简化成了一个没有相互作用的电子在有效势场中运动的问题，将电子能量分为几个部分：动能、电子与核相互作用、库仑排斥以及其余部分的交换或相关项。目前已经有大量的泛函，通常根据它处理相关和交换项的方法分为局域交换和相关项。对于交换相关能，采用局域密度近似并进行梯度校正。梯度修正泛函（Gradient-Corrected Functionals）包含了电子自旋密度及其梯度，其中的一些泛函也称为非局域泛函。应用最为普遍的梯度修正交换泛函是 1988 年 Becke 提出的泛函，例如 B3LYP（Becke 型 3 参数杂化密度泛函）。我们一般研究的体系都是气相体系，但是它对于在溶液中所进行的反应就不合适了。实际上，在气相和溶液中的分子的性质等其他现象都有区别。举例说，在溶液中所产生的静电场的作用会给分子的性质带来很大的影响，同时溶剂也会影响反应的热力学和动力学的性能，由此而改变反应的机理和反应的速率。非水溶液体系的一个常用的理论模型为自洽反应场模型 (Self-Consistent Reaction Field, SCRF)。这个模型把溶剂描述为连续的、介电常数 (Eps) 均一的反应场，并且溶质处于溶剂内的空穴中。通常按照对空穴和反应场的定义而出现了不同的理论模型。目前主要有以下四类模型：(1) Onsager 模型，最简单的模型，它把溶质放在溶剂反应场内的球穴中；(2) IPCM 模型，等密度 PCM 模型 (Isodensity PCM) 将空穴定义为分子的等密度面，这个等密度面通过迭代产生直到其形状不再发生变化为止；(3) SCI-PCM 模型，是自洽等密度极化统一场模型 (Self-Consistent Isodensity Polarized Continuum Model) 考虑了耦合的问题，并且它在等密度面的迭代过程中考虑了溶剂的影响，包含了溶剂能，而且溶剂能本身又与等密度面相关；(4) PCM 模型，Tomasi 的极化统一模型 (Polarized Continuum Model) 定义的空穴，由一系列互相连接的原子球组成。

三、计算软件与方法

使用 Gaussian, GaussView 软件。使用的长度单位为埃 (\AA)， $1 \text{\AA} = 10^{-10} \text{ m}$ ；能量单位为 Hartree， $1 \text{ Hartree} = 2 \text{ Ry} = 27.21 \text{ eV} = 4.360 \times 10^{-18} \text{ J} = 2625 \text{ kJ}\cdot\text{mol}^{-1}$ 。本实验在计算鲁米诺的异构体中，所有涉及到的理论计算均采用密度泛函的理论计算方法，并在 B3LYP/6-311G(d, p) 的理论水平上，对鲁米诺的异构体在气相 (Gas) 环境中进行了几何构型的全优化和能量计算，通过对振动频率的计算，得到了所有异构体的振动频率，并且得到的所有异构体结构的频率均无虚频；因为鲁米诺化学发光反应是在溶液中进行，所以要考虑到溶剂效应对异构体的影响，在 B3LYP/6-311G(d, p) 的理论水平上，对鲁米诺的异构体进行了溶剂化效应的计算。

四、实验步骤

(1)

用GaussView对鲁米诺建模 (LH2)，在此基础上构建两个重要的异构体 (LH3 和 LH4)。

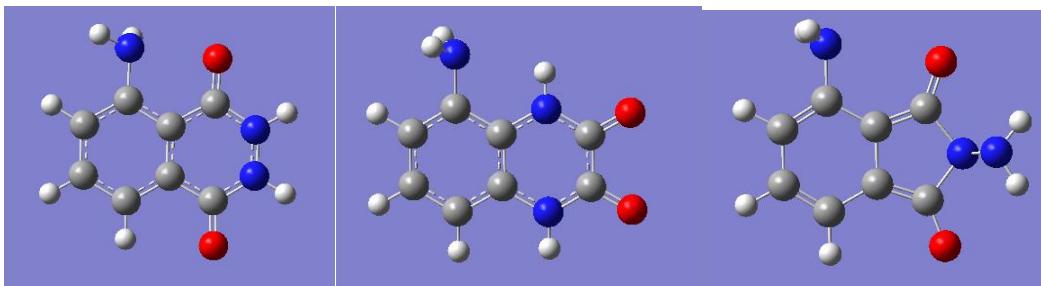


图7. 鲁米诺LH2及其异构体LH3, LH4

将此分子结构提交 Gaussian 程序计算分子能量+优化。单击 Calculate | Gaussian Calculation Setup...，弹出 Gaussian 计算对话框

实验记录：B3LYP/6-311G(d, p) 理论水平上，计算所得的气相和水相中鲁米诺 LH2、异构体 LH3 和异构体 LH4 的优化几何数据、电荷分子偶极矩，电子能量。依据电子能量数值，指出气相和水相中哪种结构更稳定？

(2) 二甲氧基鲁米诺衍生物

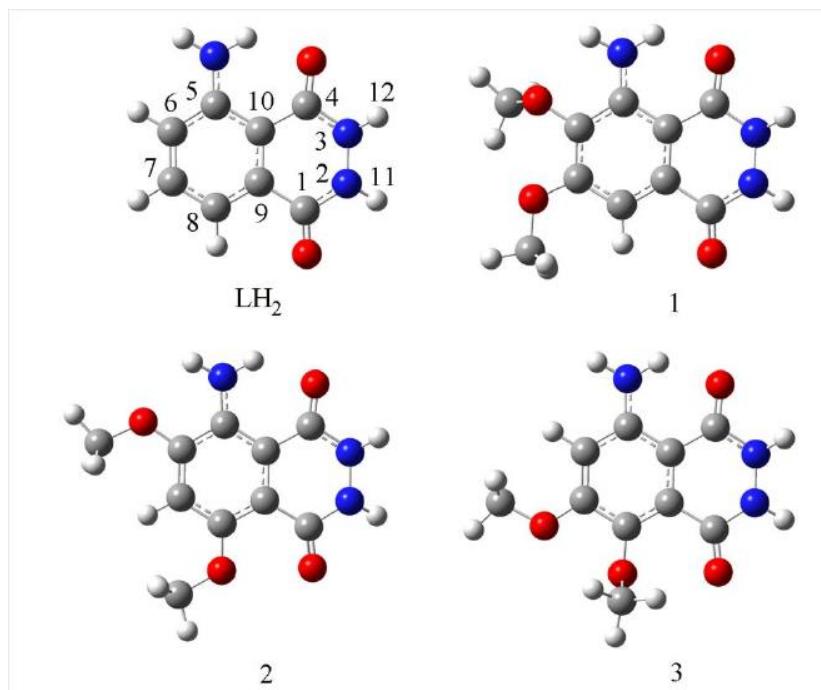


表1：二甲氧基鲁米诺衍生物分别在 B3LYP 和 MP2 方法下计算的相对能量（单位：Hartree）

isomers	B3LYP		
	Gas	water	DMSO
1			
2			
3			
MP2			
	MP2		
	Gas	water	DMSO
1			
2			
3			

绘制它们衍生物1，2，3的前线分子轨道。

五. 实验结果讨论

甲氧基的供电子性对衍生物的共轭效应有何影响？在三种衍生物中，哪个是更好的化学发光试剂？

六. 延伸实验研究

模拟以上紫外吸收光谱图，是否得出一致性的结论？

III. 格式试剂的计算实验 (Grignard)

一、 实验目的

卤代烃 ($R-X$) 与金属镁在醚类溶剂中反应，就得到了有机镁试剂 $RMgX$ (格式试剂)，这样卤代烃 $R-X$ 到烃基卤化镁 $RMgX$ ，烃基中心碳原子的电性发生了反转，由原来的缺电性变为富电性。卤化镁 $RMgX$ 可以与多种亲电试剂反应，从而高效构筑新的 C-C 键，这类反应也被称为格氏反应，其是有机合成反应中延长碳链的重要方法。尽管格氏试剂的使用已有 100 多年的历史，但由于其结构的复杂性致使人们对其反应机理的认识仍不够清晰。对格氏试剂与醛酮的反应的介绍，仅是简单地进行定性描述。利用量子化学计算获得关于格氏试剂反应机理的结果，对实验有指导意义。

二、 实验原理

前一个实验我们学会了分子构型优化及计算分子的谱学性质。本实验进一步增加了对势能面的搜索，获得过渡态。确定过渡态有助于了解反应机理，以及通过势垒高度计算反应速率可进行化学反应变化的预测。本实验涉及的理论与操作技术有：

- (1) 过渡态的定义。过渡态理论认为，反应物生成产物要经过一个活化的配合物，即过渡态，它是反应势能面上的一个鞍点。
- (2) 过渡态能量二阶导数矩阵有唯一负本征值。

(3) 在振动分析中有唯一的一个负频率（虚频）。

(4) 过渡态需要分析验证确定，该方法为内禀反应坐标方法计算。

势能面上，过渡态结构的能量对坐标的一阶导数为0，只有在反应坐标方向上曲率(对坐标二阶导数)为负，而其它方向上皆为正，是能量面上的一阶鞍点。过渡态结构的能量二阶导数矩阵(Hessian矩阵)的本征值仅有一个负值，这个负值也就是过渡态拥有唯一虚频的来源。若将分子振动简化成谐振子模型，这个负值便是频率公式中的力常数，开根号后即得虚数。

过渡态搜索算法：

牛顿-拉弗森法(Newton-Raphson, NR)与准牛顿法(quasi-Newton, QN)。

NR法是寻找函数一阶导数为0(驻点)位置的方法。通过对能量函数的泰勒级数的二阶近似展开，然后使用稳态条件 $dE/dr=0$ ，可导出步进公式：下一步的坐标向量 = 当前坐标向量 - 能量一阶导数向量 * Hessian矩阵的逆矩阵。在势能面上以NR法最终找到的结果是与初猜位置Hessian矩阵本征值正负号一致、离初猜结构最近的驻点，由于能量极小点、过渡态和高阶鞍点的能量一阶导数皆为0，故都可以用NR法寻找。对于纯二次形函数NR法仅需一步即可找到正确位置，而势能面远比之复杂，所以需要反复走步直至收敛。也因为势能面这个特点，为了改进优化，实际应用中NR法一般还结合线搜索步(line search)，对于优化至极小点，就是找当前点与NR法算出来的下一点的连线上的能量极小点作为实际下一步结构；若优化至过渡态，且连线方向主要指向过渡态，则找的是连线上能量极大点，若主要指向其它方向则找连线的能量极小点，若指向二者程度均等则一般不做线搜索。由于精确的线搜索很花时间，所以一般只是在线的当前位置附近计算几个点的能量，以高阶多项式拟合后取其最小/最大点。QN法与NR法的走步原理一样，但Hessian矩阵最初是用低级或经验方法猜出来的，每一步优化中通过当前及前一步的梯度和坐标对Hessian矩阵逆矩阵逐渐修正。由于只需计算一阶导数，即便Hessian矩阵不准确造成所需收敛步数增加，但一般仍比NR法速度快得多。QN法泛指基于此原理的一类方法，常用的是BFGS(Broyden Fletcher Goldfarb Shanno)，此法对Hessian的修正保持其对称性和正定性，最适合几何优化，但显然不能用于找过渡态。还有DFP(Davidon-Fletcher-Powell)，MS(Murtagh-Sargent，亦称symmetric rank 1, SR1)，

PSB(Powell-symmetric-Broyden)。也有混合方法，如Bofill法是PSB和MS法对Hessian修正量的权重线性组合，比二者独立使用时更好，权重系数通过位移、梯度改变量和当前Hessian计算得到，它对Hessian的修正不强制正定，很适宜搜索过渡态。将NR步进公式放到Hessian本征向量为坐标的空间下其意义更为明显，可看出在每个方向上的位移就是这个方向势能的负梯度除以对应的本征值，比如在*i*方向上的位移可写为 $\Delta X(i) = -g(i)/h(i)$ ，在受力越大、曲率越小的方向位移越大。每一步实际位移就是这些方向上位移的矢量和。对于寻找过渡态，因为虚频方向对应Hessian本征值为负，使位移为受力相反方向，所以NR法在过渡态附近每一步都是使虚频方向能量升高，而在其它正交的方向朝着能量降低的方向位移，通过这个原理步进到过渡态。若有n个虚频，则NR法就在n个方向升高能量而其它方向降低能量找到n阶鞍点。由于NR法的这个特点，为找到正确类型的驻点，初猜结构必须在目标结构的二次区域(quadratic region)内。所谓的二次区域，是指驻点附近保持Hessian矩阵本征值符号不变的区域，它的形状可以用多变量的二次函数准确描述，例如二维势能面情况下这样的区域可以用 $F(x, y) = A*x^2 + B*y^2 + C*x + D*y + E*x*y$ 来近似描述。对于能量极小点，就是指初猜点在目标结构附近Hessian矩阵为正定矩阵的范围；对于找过渡态，就需要初猜点在它附近含有且仅含有一个负本征值的范围内。并且这个范围内不能有其它同类驻点比目标结构距离初猜结构更近。NR法方便之处是只需要提供一个初猜结构即可，但是由于过渡态二次区域很小（相对于能量极小点来讲），复杂反应过渡态又不容易估计，故对使用者的直觉和经验有一定要求，即便是老手，也往往需要反复尝试。NR法对初猜结构比较敏感，离过渡态越近所需收敛步数越少，成功机率越高。模版法可以帮助给出合理的初猜，也就是如果已经知道其它机理相同的反应的过渡态结构，可以保持反应位点部分的结构不变而替换周围的原子，使之变成自己要研究的化合物反应的初猜结构。

基于反应物与产物结构的算法：

同步转变方法(synchronous transit, ST) ST 分为两种模型，最简单的就是LST 模型(Linear synchronous transit, 线性同步转变)，这个方法假设反应过程中，反应物结构的每个坐标都是同步、线性地变化到产物结构。如果反应物、产物的坐标分别以向量 A、B 表示，则反应过程中的结构坐标可表示为(1-

$x)*A+x*B$, x 由 0 逐渐变到 1 代表反应进度。以 LST 的假设, 反应路径在其所用坐标下的势能面图上可描述为一条直线, LST 给出的过渡态就是这条直线上能量最高点。LST 的问题也很显著, 其假设的坐标线性变化多数是错误的, 绘制在势能面图上也多数不会是直线, 故给出的过渡态也有较大偏差, 容易带两个或多个虚频。比 LST 更合理的是 QST (quadratic synchronous transit, 二次同步转变), 它假设反应路径在势能面上是一条二次曲线。QST 在 LST 得到的过渡态位置上, 对 LST 直线路径的垂直方向进行线搜索找到能量极小点 A。QST 给出的反应路径可以用经过反应物、A、产物的二次曲线来表示, 如果这条路径上能量最大点的位置恰为 A, 则 A 就是 QST 方法给出的过渡态; 如果不是, 则以最大点作为过渡态。若想结果更精确, 可以再对这个最大点向垂直于路径的方向优化, 再次得到 A 并检验, 反复重复这个步骤, 逐步找到能量更低、更准确的过渡态。

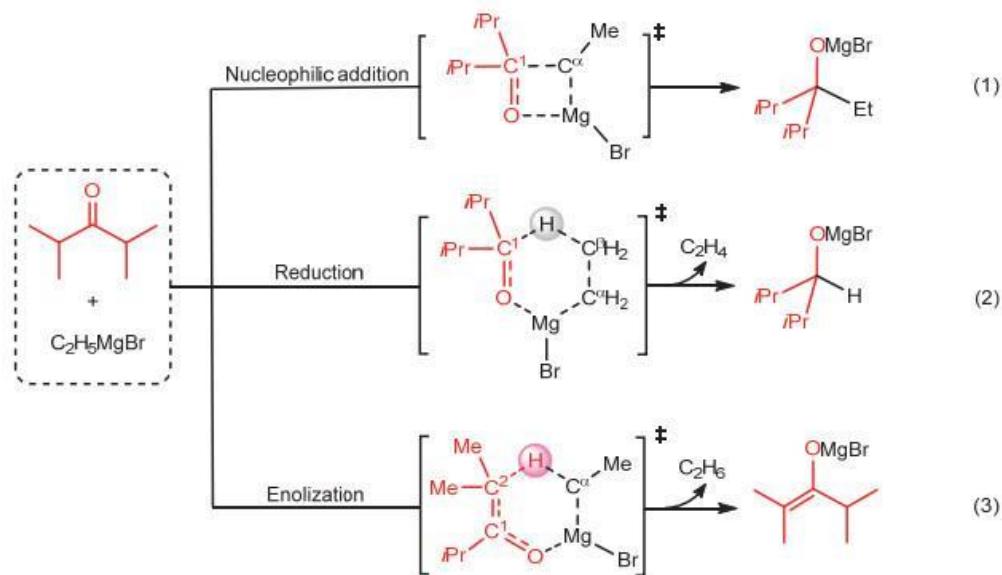
三、计算软件与方法

研究对象为乙基溴化镁(EtMgBr)与二异丙基酮(CO(iPr)₂)的反应, 溶剂为乙醚。根据 Schlenk 平衡, EtMgBr 和 MgEt₂ 在乙醚溶剂中平衡共存, 因此, 计算考虑了 EtMgBr 和 MgEt₂ 两种组分的有机镁与 CO(iPr)₂ 的反应。为便于描述, 将 CO(iPr)₂ 中羰基 C 原子和参与反应的三级 C 原子分别标记为 C₁ 和 C₂, EtMgBr 中亚甲基 C 原子和甲基 C 原子分别标记为 C_α 和 C_β。

本研究基于密度泛函理论, 使用 M062X 泛函和 6-311G(d, p)标准基组对全部稳定点进行了结构优化和频率计算, 确保局域最小点没有虚频, 过渡态(一级鞍点)有且仅有一个虚频, 并通过内禀反应坐标(IRC)分析, 确认了每个过渡态连接的反应物和产物。应用 PCM 连续介质模型评估了溶剂化效应, 在乙醚溶剂中对稳定点进行了全构型优化。全部计算使用 Gaussian 程序完成。图形显示用 GaussView。使用的长度单位为埃 (Å), $1 \text{ Å} = 10^{-10} \text{ m}$; 能量单位为 Hartree, $1 \text{ Hartree} = 2 \text{ Ry} = 27.21 \text{ eV} = 4.360 \times 10^{-18} \text{ J} = 2625 \text{ kJ} \cdot \text{mol}^{-1}$ 。

四、实验步骤

1. 通过GaussView构建



2. 设置计算参数，如下所示。

反应物优化 opt

过渡态搜索 opt=(ts, noeigentest, nrsscale, calcfc) iop(1/8=6) freq

3. 计算完毕后，打开log文件。

4. 在GaussView菜单Results中选择vibrations，观察记录虚频。

5. 记录能量，进行分析作图。

6. 在显含溶剂下重复以上过程。

五. 实验结果讨论

烷基取代基和溶剂对反应性能的影响？

六. 实验拓展