酶催化过程的量子-经典力学研究进展

樊芳芳¹, 付玉状², 曹泽星^{2*}

1. 浙江科技学院生物与化学工程学院, 杭州 310023;

2. 厦门大学化学化工学院, 固体表面物理化学国家重点实验室, 福建省理论与计算化学重点实验室, 厦门 361005

* 联系人, E-mail: zxcao@xmu.edu.cn

2020-10-15 收稿, 2020-11-18 修回, 2020-12-01 接受, 2020-12-07 网络版发表

国家自然科学基金(21873078, 21673185, 21933009)资助

摘要 酶催化主要包括底物输运及其与活性位的结合、蛋白质限域环境中的化学反应以及产物释放等复杂的物理和化学过程, 其中任一化学和非化学步都有可能决定酶的活性. 通过量子/分子力学组合方法(QM/MM)计算和分子动力学模拟(MD), 结合伞形采样, 可以揭示酶催化过程的微观机理和底物输运的通道机制, 分析底物进出口袋过程的热力学和动力学性质, 探明整个酶催化过程中的决速步骤以及关键残基的作用. 本文概述了近几年对几类生物体系量子-经典力学模拟研究的进展, 并对QM/MM多尺度模拟方法的发展进行了展望.

关键词 磷酸三酯酶, 核苷水解酶, 酶催化, QM/MM和MM MD模拟

酶作为一类重要的蛋白质生物大分子, 具有很高的生物催化活性, 且因其高效环保的特点一直受到人们的广泛关注. 在酶催化过程中, 起关键作用的往往局限在一定的区域内, 也就是活性中心^[1]. 活性中心通常位于酶分子的表面或缝穴处, 虽然所占整个酶体系的体积比较小, 其结构特征和关键残基的组成直接影响酶的催化活性与底物选择性. 酶催化通常包括化学和非化学等复杂的多步物理化学过程, 首先, 底物通过输运通道进入到活性中心, 并与活性中心结合形成稳定的酶-底物复合物, 进而触发专一的酶促反应, 生成产物, 酶促反应完成后, 产物释放, 酶恢复其活性态结构^[2], 整个酶催化过程如图1所示.

由于蛋白结构及酶环境的复杂性, 实验上很难获得酶活性态的准确三维结构及其动态信息, 为认识酶催化活性与其结构的关联及其催化机制带来了困难. 此外, 通过实验虽然可以获得催化反应的米氏常数、最大反应速率等表观数据, 但酶催化功能的结构本质、活性区的动态结构特征、决速步的确定与调控,

以及酶催化过程的中间体和过渡态的结构等都很难通过实验手段直接表征.

随着计算机技术和科学计算的快速发展, 高性能计算的能力不断提高, 多尺度计算模拟已经成为当前辅助实验研究必不可少的重要手段. 它可以从原子/分子水平上揭示酶催化过程的化学反应机理以及底物输运的通道机制. 通过对酶催化过程的全程模拟, 可以获得酶催化全过程的自由能变化趋势, 预测整个催化过程中的决速步骤, 深入认识催化反应的微观机制, 探明关键残基的作用. 这将有利于通过酶工程, 进一步有选择性地改善和提高酶催化的反应效率和高效生物催化剂的设计. 目前, 生物体系模拟的理论计算方法主要包括量子力学(QM)^[3]、分子动力学(MD)、量子力学/分子力学组合方法(QM/MM)^[4]以及QM/MM-based MD方法^[5]等. QM/MM组合方法利用量子力学处理化学键变化的可靠性和分子力学处理非化学反应区的计算能力, 可以合理描述复杂生物体系中的化学反应, 已广泛应用于酶催化过程的计算模拟研究.

引用格式: 樊芳芳, 付玉状, 曹泽星. 酶催化过程的量子-经典力学研究进展. 科学通报, 2021, 66: 1131–1143

Fan F F, Fu Y Z, Cao Z X. QM/MM study of the enzymatic catalysis (in Chinese). Chin Sci Bull, 2021, 66: 1131–1143, doi: [10.1360/TB-2020-1321](https://doi.org/10.1360/TB-2020-1321)

图1 酶催化全过程(底物运输与结合、催化反应、产物释放)

Figure 1 The whole enzymatic process (the substrate delivery and binding, chemical reaction, and product release)

1 酶催化机理

生命体中的大多数反应都是在酶催化下完成的, 酶在细胞生长、信号传导、能量传递以及遗传翻译等生命过程中都扮演着不可或缺的角色. 与一般的化学催化不同, 酶催化在接近中性的常温常压条件下就可以反应, 催化效率高, 同时还具有高度的专一性. 同时, 酶的活性是可以调控的, 影响酶催化功能的不仅仅是催化过程中的化学反应步, 也有可能是底物运输和产物释放的非化学步, 需要通过对酶催化过程进行全程模拟才能获得对其催化机制的深入理解. 实验表征结合理论计算模拟, 可以确定酶体系的三维结构特征, 在此基础上, 进一步揭示酶催化的作用机制^[6]. 了解酶催化微观机制, 对寻找最有利的反应条件, 充分发挥酶催化的高效性, 探索药物分子在生物体内的作用机制等都至关重要^[7].

2 通道运输机制

酶催化反应开始之前, 底物需要从外部环境进入到活性口袋内, 同时催化反应结束之后, 产物需要释放离开活性中心. 反应物种运输到活性位和产物释放过程中, 底物与蛋白质侧链残基存在复杂的氢键、范德华等非键相互作用, 同时伴随着活性区水分子的溶剂化和去溶剂化, 这些将决定底物运输路径及其热力学和动力学性质, 也就是通道运输机制. 尽管底物运输主要包括非化学步, 近几年研究发现, 产物的释放也可能是整个酶催化过程的决速步骤^[8]. 因此, 底物的运输通道机制也是影响催化效率的重要因素之一, 在酶催化计算模拟研究中, 从原子/分子水平上, 探明底物进出活性位过程中与蛋白环境之间的相互作用, 有助于深入了解酶催化反应机制与活性^[9]. 随机加速分子动力学模拟(random acceleration molecular dynamics,

RAMD)可以快速获得底物进出活性位的可能通道, 预测底物进出活性口袋通道结构的动态变化, 结合伞形采样技术, 确定底物运输过程中的自由能变化信息^[10].

3 几类生物体系量子-经典力学研究进展

3.1 磷酸三酯酶降解G型神经毒剂

磷酸三酯酶(PTE)作为酰胺水解酶超家族的成员之一, 可以催化降解杀虫剂和神经毒剂等有机磷化合物. 如图2所示, 该酶的活性位有3个口袋: 大口袋、小口袋和离去口袋^[11]. 小口袋深埋于口袋内部, 与外部环境相对隔离, 大口袋与部分溶剂水分子接触, 离去口袋则暴露于溶剂中, 其在协调底物的运输过程中起关键作用. 有机磷酸盐(OPs)通常具有高毒性, 它们已被广泛用作杀虫剂和神经毒剂, 一些高毒性的有机磷化合物, 如G型、V型神经毒剂, 已被用作恐怖袭击和战争中的化学武器^[12]. 神经毒剂主要通过活性位丝氨酸残基的共价磷酸化, 导致生物体中乙酰胆碱酯酶(AChE)不可逆的抑制, 使得神经元-神经元连接处乙酰胆碱过度积累, 导致生物体许多生理功能失效, 甚至导致生物死亡^[13].

在处理该类神经毒剂时, 酶催化降解与传统的化学处理和焚烧相比, 对环境友好, 一直受到关注. PTE酶可以降解有机磷化合物, 并表现出广泛的底物特异性, 但水解效率很大程度上取决于有机磷底物的化学结构特征. 例如, PTE酶催化杀虫剂对氧磷水解具有非常高的效率($k_{\text{cat}}/K_{\text{m}} \approx 10^8 \text{ mol}/(\text{L s})$), 其降解速率几乎达到了扩散主导的极限^[14]. 但是PTE酶降解G型神经毒剂, 催化效率降低到 $10^4 \sim 10^5 \text{ mol}/(\text{L s})$, 对V型神经毒剂的水解效率甚至更低^[15]. 为了更好地认识这种酶促过程的机制, 过去几十年来人们一直在努力. 然而, 由于神经毒剂的高毒性和对酶活性评估实验的高成本, 截

图2 PTE酶活性位的结构示意图, 包括3个口袋以及催化反应的反应坐标(RC2)

Figure 2 The structure of the active site from the PTE-substrate complex, including the three pockets and the reaction coordinate (RC2) for the chemical reaction

至目前, 很少有类似的底物用于酶降解机制研究. 最近, 本课题组^[16]通过QM/MM和MM MD模拟, 系统地研究了PTE酶降解神经毒剂沙林的催化机制.

底物输运及其在活性位的结合, 是酶催化过程的第一步. 如图3和4所示, 沙林输运到活性位的整个过程可分为连续的3个阶段: 初始阶段($RC1 \geq 16.8$ Å), 沙林完全游离在口袋外, 然后通过和loop7上残基的相互作用逐渐接近离去口袋所在的活性位. 当 $RC1=20.2$ Å时, 螺旋解旋为柔性的loop7, 这种状态持续到底物穿过该

loop并进入口袋中. 当底物进入口袋时, 残基Leu271和Phe132在PTE酶的整个通道中起到了门控的作用, 门控机制主要来源于沙林和loop7之间的相互作用, 并使得离去口袋进一步被loop7覆盖. 中间阶段(16.8 Å $> RC1 \geq 13.6$ Å), 在磷酸基和带正电荷的锌离子之间相互作用的引导下, 底物逐渐深入到口袋内同时排挤出里面的部分溶剂, 这种相互作用提供了辅助底物输运和结合的驱动力. 在 $RC1=14.4$ Å时, 磷酸基和 Zn_{β}^{2+} 之间的距离减小到 2.2 Å, loop7部分恢复为螺旋状结构, 入口

图3 PTE酶催化降解沙林的自由能变化全景图

Figure 3 The free energy profile for the degradation of sarin by PTE

图 4 底物运输过程中代表性的通道结构及其构象变化

Figure 4 Representative structures and conformational changes of the delivery channel in the sarin transportation

的直径增加到了18.35 Å, 达到了它在整个底物运输过程中的最大距离, 表明在此阶段底物已经穿过了口袋的最窄部分. 最后阶段($13.6 \text{ Å} > \text{RC} \geq 9.6 \text{ Å}$), 一旦底物成功通过入口门控区域, 将调整其结构以适应活性口袋的内部环境. 由于底物具有体积小, 柔韧性高的特

点, 中心磷基团首先进入对应的口袋, 然后是体积较大的基团, 在 $\text{RC} = 9.6 \text{ Å}$ 时调整到了最佳的结合模式. 随着底物移入口袋中, 运输通道的门控区经历了“半闭合”的构象变化, 以及由Leu271和Phe132之间距离表征的通道直径明显减小. 值得注意的是, 口袋中的水分子数

量在RC1=13.0和10.0 Å时经历了两次急剧下降,而之前保持动态稳定.水分子数量的演变特征表明,在底物与口袋完全结合之前,活性位点结构松弛,并且有相当多的水分子滞留在口袋中.这些水分子在底物结合过程中起重要作用,当底物进入活性口袋时,活性位点的水分子逐渐被挤出口袋,这种去溶剂化作用对底物的结合自由能有显著的贡献.如图3所示,底物与活性位的结合将放热16.8 kcal/mol.更重要的是,当沙林进入活性位时,输运通道的入口直径经历了持续波动,在整个过程中,Loop7作为“门控”,调节底物与通道入口的动态结合和相互影响,对底物的运输起着至关重要的作用.

野生型PTE酶活性位的一个关键特征是双核的锌离子中心,在200 ns长时间的MD模拟过程中,有两个水分子进入了活性口袋并分别与两个锌离子稳定配位,导致两个锌离子形成稳定的六配位,其中 $\text{Zn}\cdots\text{OH}_2$ 平均距离分别为 1.95 ± 0.07 和 1.98 ± 0.09 Å,这两个水分子的存在对催化活性具有重要影响.OP水解通常是从磷中心桥接 OH^- 的亲核攻击开始的,因此选择 r_2 与 r_1 的距离差作为反应坐标(RC2,图2),描述水解反应中的P-F键裂解和P-O $_{\mu}$ 键形成.

如图3所示,PTE催化沙林水解的化学反应包含两步,反应进行到第一步过渡态(TS1)时经历9.8 kcal/mol的自由能垒,其中 $\text{P}\cdots\text{O}_{\mu}$ 距离从4.03 Å明显缩短至2.16 Å,而P-F键长度变化比较小.一旦反应通过TS1态,形成中间体(IM1),将释放2.2 kcal/mol的能量.反应进行到TS2态,相对于初始反应物的自由能垒为12.3 kcal/mol,其中 $\text{P}\cdots\text{F}$ 距离为2.06 Å,Asp301 $\cdots\text{H}$ 距离连续缩短至1.39 Å.随着水解进一步进行, F^- 将离开有机磷酸盐底物, $\mu\text{-OH}$ 上的氢原子转移到附近的残基Asp301并与之结合(P_C/R_F).两步水解化学过程的自由能跨度为12.3 kcal/mol,因此,该化学步在动力学和热力学上都是非常有利的.降解产物通过 $\text{O}_{\mu}\cdots\text{Zn}_{\alpha}$ (2.00 Å)和 $\text{O}_L\cdots\text{Zn}_{\beta}$ (1.95 Å)键合相互作用桥连两个金属锌离子,这与Kim等人^[17]通过X射线观察到的PTE酶与磷酸二乙酯复合物的晶体结构一致.

如图3所示,该酶催化反应步的能垒相对较低,应具有较高的活性.但实验上,该酶催化沙林降解效率却很低.计算模拟揭示,随着氟离子的释放(P_F/R_D),降解产物磷酸酯以双齿配位模式桥连到两个锌位点上,此时 Zn_{α}^{2+} 保持五配位,而 Zn_{β}^{2+} 结合两个水分子后恢复稳定的六配位结构.在产物结合状态下,即使离去基团

暴露于溶剂,双齿螯合结合的降解产物解离和释放也是很困难的.考虑到降解产物稳定的双齿状态和复杂的局部蛋白环境,通过QM/MM MD模拟结合伞形采样的方法,探讨了降解产物解离和释放的多种可能性(模型A~D),最后获得了一个相对合理的离去机理(模型D).

模型D提供了释放最终降解产物的可能途径.首先,质子从一个水分子中转移到带负电的降解产物(重组,对应的过渡态TS_D).然后,新生的 OH^- 桥连配位双核金属锌(稳定化,对应的中间体IM_D).最后,活性位中性产物的解离(P_D).为了描述多步骤解离过程,基于几个选定距离组合定义反应坐标,采用QM/MM MD模拟,预测了降解产物解离过程的自由能变化.如图3所示,重组过程包括了质子转移、从双齿配位到单齿配位的转化,以及桥连氢氧化根基团的形成,经历的自由能垒为21.0 kcal/mol,该反应步是整个沙林降解的决速步.随后的一个水分子取代降解产物与 Zn_{β}^{2+} 配位,具有14.8 kcal/mol的自由能垒,并且该配体交换遵循 $\text{S}_{\text{N}}2$ 机制,整个解离过程是吸热的.随着降解产物的释放,磷酸酯的溶剂化可以产生约25.3 kcal/mol的稳定化能,可以进一步驱动产物从活性位释放.这些全程模拟结果深化了PTE降解沙林机制的认识,对进一步的实验研究和降解有毒有机磷化合物的酶工程具有积极意义.

3.2 质子化对核苷水解酶催化反应的影响

核苷水解酶(nucleoside hydrolases, NHs)是一种含有钙离子的金属酶,能够催化核糖核苷中的糖苷键(C-N)水解,产生游离的核糖和含氮碱基^[18].这种酶对一些生物的存活至关重要,特别是自身缺乏合成嘌呤和嘧啶功能的微生物.该酶广泛存在于原生动物、细菌、昆虫、酵母、线虫等生物体内中,但在哺乳动物中从未发现,因此它们是设计抗寄生虫药物的主要靶标.基于底物特异性的差异,核苷水解酶可细分为4个亚类:第一类是嘧啶特异性的胞苷-尿苷核苷水解酶^[19](CU-NH);第二类是嘌呤特异性的肌苷-腺苷-鸟苷核苷水解酶^[20](IAG-NH);第三类是6-氧取代嘌呤特异性的肌苷-鸟苷核苷水解酶^[21](IG-NH);第四类是非特异性肌苷-尿苷核苷水解酶(IU-NH),可以同时水解嘌呤和嘧啶类核苷.

先前MM MD和QM/MM MD模拟研究表明,底物肌苷上N7位的质子化可显著地降低IAG-NH酶水解N-

糖苷键化学步的自由能垒^[22]. 为此, 对CU-NH酶的水解活性开展了计算模拟研究, 探明了不同底物特异性的本质及其质子化效应^[23]. 通过QM(B3LYP/6-31G*)/MM MD模拟, 结合伞形采样, 预测的反应自由能垒约为22.1 kcal/mol(图5). 基于双精度QM/MM方法, QM区域大基组6-311+G**的使用, 能垒降低至约20 kcal/mol, 该结果与实验上动力学数据估算值17.4 kcal/mol接近^[19]. 如图5所示, CU-NH酶水解尿苷过程中, *N*-糖苷键断裂的同时水分子上的一个质子转移到Asp11, 二者是一个协同但不同步的过程. 在初始状态下, 反应物水分子的氧原子距离底物的C1'原子3.2 Å, 其对C1'原子的亲核攻击, 触发了*N*-糖苷键的水解反应. 随着水分子逐渐接近C1'位点, N1-C1'的糖苷键长度同步增加(图5), 然后形成过渡态, C1'原子的杂化轨道从反应物状态的sp³转换为sp². 随着反应的进行, 碱基进一步远离核糖基团, 水的氧原子逐渐靠近C1'原子. 最后羟基与C1'键合, 同时水分子上的一个质子转移到Asp11. 如图5所示, CU-NH催化*N*-糖苷键的水解断裂是协同的, 这与IAG-NH的酶促水解反应类似.

通常, 碱基的质子化对于核苷水解酶的水解反应是非常重要的. 对于来自*C. fasciculata*的IU-NH, His241作为质子源可以激活次黄嘌呤的离去基团^[24]. 大肠杆菌蛋白质对尿苷水解的动力学同位素效应研究表明嘧啶核碱基可能通过碱基和酶之间的氢键相互作用或底物的质子化而被激活. IAG-NH酶水解肌苷的QM/MM MD模拟也显示, 底物N7的质子化可降低反应能垒, 促进C1'-N键的裂解^[22]. 如图6所示, 在CU-NH酶的初始构型中, Gln227、Tyr231和His239残基都可能

充当底物质子化的质子供体. 基于最后5 ns MD模拟的氢键分析, Tyr231与底物的O4始终保持着强氢键相互作用, 两者的O...O距离为(2.67 ± 0.11) Å; Gln227与O2也有氢键相互作用, 概率高达89%. 然而, His239和尿嘧啶环的N3之间形成稳定氢键的概率只有2.66%, 基本上不可能充当质子源, 这主要因为N3上的孤对电子参与了环共轭成键作用, 此处的质子化比O2和O4位点的质子化需要更多的能量(分别高18.8和37.6 kcal/mol), 因此排除了N3质子化的可能. 为了评估质子化对*N*-糖苷键断裂的影响, 建立了几种可能的质子化模型, 并对水解过程进行了QM/MM能量扫描. 图6的相对能量变化结果表明, 在O2和O4处, 质子化底物的*N*-糖苷键断裂能垒分别为26.0和22.0 kcal/mol. 显然, 对于CU-NH水解尿苷, 底物质子化并不能有效降低化学步C-N键断裂的能垒, 这与IGA-NH酶完全不同.

这些结果表明, 底物不同质子化状态对同一家族酶催化体系的影响也是不一样的, 底物肌苷N7位质子化可明显降低IAG-NH酶水解反应的能垒, 导致水解产物的释放是其水解过程的决速步^[22]. 然而, 对于CU-NH酶催化水解反应, 底物质子化并不能有效降低离去基团水解的能垒, 其水解化学步是该酶反应的决速步, 这与IAG-NH酶完全不同. 因此, 同家族的IAG-NH和IU-NH在*N*-糖苷键的酶促水解中依然具有不同的pH依赖性和水解机制.

3.3 核苷水解酶的底物特异性

尽管嘧啶特异性水解酶(CU-NH)和非特异性核苷水解酶(IU-NH)具有高度同源性, 然而它们对底物的选

图5 水解过程自由能变化、代表性结构(a)及沿着反应方向C-N和C-O原子间距离的变化(b)

Figure 5 The relative free energy profiles, the selected representative structures (a) and the evolution of selected C-N and C-O interatomic distances along the hydrolysis reaction (b)

图6 QM/MM计算预测的质子化尿苷模型中N-糖苷键裂解的相对能量变化以及尿苷周围氢键存在的形成概率

Figure 6 QM/MM-predicted relative energy profiles for the N-glycosidic bond cleavage in various protonated models of uridine, and the probability of hydrogen bonds around the uridine

择性截然不同, 这表明虽然具有3个羟基的核糖部分对底物结合和水解是至关重要的, 然而活性位点残基与碱基之间的选择性结合在底物特异性中起关键作用^[25]. 对于嘧啶特异性的CU-NH, 肌苷和尿苷在活性位点的结合自由能分别是-28.3和-45.7 kcal/mol. 肌苷结合自由能的减少就是因为酶活性区残基与碱基之间强氢键相互作用的消失, 由此可见嘧啶特异性水解酶CU-NH具有的底物特异性, 主要体现在对碱基的选择性.

嘧啶特异性CU-NH 酶的动力学实验数据显示, 其对肌苷、腺苷或鸟苷的催化效率比尿苷至少慢1000倍^[26]. 由于这些底物都具有共同的核糖体部分, 所以底物的特异性主要还是来源于对碱基部分的选择性. 图7(a)给出了主要残基对底物-酶结合能的贡献, 如图7(a)所示, Gln227和Asp15对于YeiK底物特异性具有显著重要性. 残基突变计算分析表明, YeiK的Gln227/Ala突变显著降低了尿苷的结合稳定性, 结合自由能从-45.7降低至-35.3 kcal/mol, 同时也增强肌苷的结合稳定性, 结合自由能从-28.3增加至-34.3 kcal/mol. 此外, Asp15/Ala突变体导致尿苷的结合自由能为-27.5 kcal/mol, 肌苷的结合自由能为-29.4 kcal/mol. 显然, 与底物肌苷相比, 残基Gln227和Asp15在尿苷结合中发挥更加重要的作用. 如图7(d)所示, 残基Gln227和Asp15与底物尿苷可以形成直接的氢键相互作用. 相反, 对于底物肌苷没有发现类似的相互作用, 并且肌苷和YeiK的活性位点残

基之间的结合相互作用受Gln227/Ala和Asp15/Ala突变的影响较小. 由此可以推测, CU-NH YeiK的这些活性位点残基可通过特定的酶-底物相互作用调控底物的特异性选择.

图7(b)给出了CU-NH YeiK酶apo态及其分别结合底物肌苷和尿苷复合态的结构波动. 图中结果显示, 尿苷与酶的活性位点结合时, RMSF值明显降低, 而肌苷结合在蛋白质的RMSF值相对于apo态部分区域中变化较小或略微增加. 与肌苷相比, 尿苷与CU-NH YeiK的活性位点残基之间的相互作用更强. 我们还估算了最后5 ns MD模拟中NHs活性位点口袋的体积, 在结合底物尿苷后, CU-NH的活性位口袋略微膨胀, 体积从apo态的77.6 Å³扩张到115.2±8.8 Å³; 与CU-NH相比, 来自Crithidia的IU-NH的apo态具有更大的结合口袋, 体积为99.9 Å³. 同时, 尿苷与IU-NH酶的结合导致其活性位口袋从99.9 Å³扩张到193.4±16.5 Å³, 对比CU-NH有显著的扩展, 这表明IU-NH具有更灵活的活性口袋以容纳相对大的底物, 如肌苷. 正如实验所观察到的, IU-NH酶对尿苷和肌苷的底物都显示出高活性. QM/MM计算显示, 底物尿苷和肌苷的能垒分别为22.1和30.3 kcal/mol, 表明CU-NH催化水解肌苷比尿苷更困难. 总之, 活性位局部结构环境及其与底物的相互作用不仅影响底物的结合, 还将进一步影响化学步中键的断裂和形成, 这种非化学和化学步的协同作用, 决定了CU-NH的底物特

图7 不同底物与CU-NH酶相互作用及其水解化学步的差异。(a) YeiK活性位点残基与底物尿苷(红色)和肌苷(蓝色)相互作用对结合自由能的贡献;(b) 蛋白质骨架的RMSF: apo态(灰色), 酶-肌苷复合物(蓝色)和酶-尿苷复合物(粉红色);(c) QM/MM计算预测的底物尿苷和肌苷N-糖苷键水解过程的相对能量变化;(d) CU-NH酶活性位点中尿苷周围的氢键网络示意图

Figure 7 The interaction between CU-NH and different substrates and the difference in the chemical step of hydrolysis: (a) Comparison of the energy contribution from active-site residues of YeiK to the binding of uridine (red) and inosine (blue); (b) the RMSF of the protein backbone for the apo state (gray), enzyme-inosine complex (blue), and enzyme-uridine complex (pink); (c) QM/MM predicted relative energy profiles for the *N*-glycosidic bond cleavage of substrates uridine and inosine; (d) the hydrogen bond network around inosine in the active site of the CU-NH enzyme

异性。

3.4 基于选择特异性机制的结构修饰

萘酰亚胺化合物是一种细胞荧光成像剂,同时也是一种潜在的优良小分子抗癌药物,长期以来一直受到药物化学家的广泛关注^[27]。但是由于其不溶于水且存在一些副作用,导致其应用受到了很大限制,人们一直想要通过在萘酰亚胺环上引入官能团对其进行修饰。多胺的出现,正好可以解决这一问题,多胺是生物体内普遍存在的细胞阳离子,它们可以选择性地进入肿瘤细胞,从而显著地降低抗癌药物对正常细胞的副作用^[28]。因此,萘酰亚胺-多胺化合物可以很好地发挥双重药效的作用,这直接推动了萘酰亚胺化合物的发展。这些萘酰亚胺化合物需要通过血清白蛋白(SA)才能顺利到达药物靶点。血清白蛋白,是一种多功能蛋白质,具有良好的选择性,能够与多种化合物结合并将其输送到生物体内的目标靶点,这些化合物包括脂肪酸,氨基酸和药物小分子等。药物小分子通常在其中的一个

或两个药物结合位点(DSI和DSII)与血清白蛋白结合,这两个药物位点分别位于IIA和IIIA亚区域,这最初是由Sudlow等人^[29]提出的。

探究萘酰亚胺-多胺化合物如何在血清白蛋白的帮助下到达可能的疾病靶标,对于该类化合物的研究设计具有重要意义。实验上,近期已成功合成了4种萘酰亚胺-多胺化合物,并对它们进行了初步的光谱研究^[30]。在此基础上,我们建立了7种底物计算模型,如图8所示,应用分子动力学模拟,研究了药物位点的选择性,关键残基的作用及取代基效应。该研究将有助于揭示类似药物与血清白蛋白结合及其转运机制,为抗癌药物的进一步研究和开发提供理论依据^[31]。

为了研究牛血清白蛋白对这些化合物的相对亲和力,使用了MM-GBSA方法来评估它们的结合自由能^[32]。结果显示,在DSI中,萘酰亚胺环上具有吸电子基团的化合物比具有供电子基团的化合物能够更加紧密地与牛血清白蛋白结合。结合作用强烈地依赖于整个化合物的吸电子效应。具有两个强吸电子基团(NO_2)

图8 牛血清白蛋白的结构、主要结合位点以及MM-GBSA预测的各个模型化合物与两个药物位点的结合自由能(kcal/mol)
 Figure 8 The structure of BSA, the primary binding sites, and the binding free energies (kcal/mol) for two drug sites predicted by MM-PBSA

的模型A显示出最强的结合能力,这与光谱实验(紫外、荧光和圆二色光谱)中观察到的结论相一致。对于DSII,它在底物的结构修饰方面与DSI具有完全不同的结合特征,具有供电子基团的化合物显示出比具有吸电子基团的化合物更强的结合作用。此外,还发现结合能与取代基的数目相关,取代基数目越多,结合能力反而越弱,这可能是口袋大小引起的空间位阻效应。

总而言之,对底物结合能力的强弱很大程度上取决于取代基的供电子和吸电子能力以及取代基的数量。值得注意的是,对于DSI,模型A-2NO₂的结合相互作用比模型B-NO₂强12.8 kcal/mol。相反,对于DSII,模型E-2NH₂的结合相互作用弱于模型D-NH₂ 8.8 kcal/mol。这种差异可归因于它们的口袋大小,DSII的口袋不像DSI那样大(分别是427.6±79.2与564.3±87.8 Å³),其空间位阻会削弱它们的相互作用。我们注意到这些化合物在DSII和DSI中结合自由能的最大差异分别是11.9和22.2 kcal/mol,这表明与DSII相比,DSI对底物结合的取代基效应更加明显。为了进一步阐明两个药物位点的结合特征,分别对其进行了能量分解分析。两个药物位点的结合过程在热力学上通常是有利的,其中非极性溶剂

化(ΔG_{SA})对结合具有一定有利的影响,但它在很大程度上被极性溶剂化(ΔG_{GB})抵消,所以总溶剂化效应(ΔG_{sol})反而是不利于结合的。相反,范德华相互作用(ΔG_{vdw})和静电(ΔG_{ele})都有利于化合物的结合,因此包括这两部分的 ΔG_{gas} 对底物结合起着重要的驱动作用。我们还注意到DSI气相中的结合相互作用(ΔG_{gas})比DSII强得多,这主要归因于DSI中底物与周围残基之间的强氢键相互作用。

基于对配体中取代基影响的计算分析发现,DSI更倾向于带有吸电子基团的配体,而DSII更倾向于带有供电子基团的配体。尽管DSI对于底物结合比DSII更加有益,但是两个药物位点都可能参与并协助将萘酰亚胺-多胺化合物输运到它们的生物靶标。所以我们推测,吸电子基团和供电子基团的组合效应可以同时增强两个药物部位对底物结合作用,因此又构建了同时具有吸电子的硝基和供电子的氨基的新萘酰亚胺-多胺化合物,命名为Model F(图8)。理论计算模拟表明,该化合物在DSI和DSII中与牛血清白蛋白的结合表现出更好的平衡。具有不同电子效应的取代基组合可以增加酶结合模式和功能的多样性,这可以用于设计基于萘酰

亚胺-多胺的抗肿瘤药物。

4 总结与展望

本文总结了我们在近些年来在酶催化和底物输运多尺度模拟研究方面的进展,通过系统的MM MD和QM/MM MD计算模拟,可以揭示酶催化的微观机理和底物输运的通道机制,确定催化过程的决速步和关键残基的作用,深入全面地理解酶生物活性的结构本质^[33]。基于全程模拟,构造了几类酶催化的自由能变化全景图,明晰酶催化过程中化学和非化学步的重要性。我们的计算模拟研究表明,在底物输运与结合过程中,活性位、通道口及其侧链关键残基的构象会发生显著变化,它们的动态结构特征及其与底物相互作用对底物特异性具有重要影响^[34]。目前,多尺度理论计算模拟,已广泛用于复杂的金属酶及相关生物体系^[35-37],从原子/分子水平上揭示酶催化机理和底物输运通道机制,弥补了实验方法和实验手段的局限性,为新型药物小分子设计和酶工程的应用研究提供了理论依据。

1976年,Warshel和Levitt^[38]提出了杂化QM/MM的概念,并将其用于研究溶解酵素的反应机理研究。QM/MM方法通过对参与化学反应区域用量子力学(QM)处理,可以合理描述化学键的变化,对剩下的蛋白质和溶剂部分用分子力学(MM)描述,可以利用分子力学处理大体系的高计算效率描述活性区外的环境效应,二者的有机结合实现了酶体系的催化全过程模拟。该方法是目前生物体系计算模拟研究的主要工具^[39]。随着密度泛函理论的发展,量子力学(QM)方法已经能够计算数百个原子的体系,但对生物体系的直接QM计算仍存在困难。在QM/MM水平上,基于统计力学采样方法,预测酶催化过程的自由能变化非常耗时,使得大体系的QM/MM MD模拟研究具有很大的挑战性。此外,目前的QM/MM方法,大多MM力场是非极化的,为了合理描述QM和MM区及QM/MM边界电荷转移与极化作用,

可极化的MM力场和QM/MM边界模型还有待不断完善与发展。

随着结构生物学的拓展以及冷冻电子显微镜技术的发展,让我们获得了大量生物分子的结构信息,但是想要全面理解酶的通道机制和催化机理,除了静态结构,动态信息也不可或缺。展望未来,在生物酶催化过程机理和底物通道机制的研究中,提高理论计算方法的准确性和效率仍是关键。QM/MM方法作为重要的多尺度模拟工具之一,其计算效率需进一步提升,才能实现对复杂生物体系更长时间更加可靠的计算模拟^[40]。此外,在底物输运和产物释放过程中,其局域结构环境不断变化,要提高全程模拟的精度,合理平衡处理不同区域电荷转移及介质变化,多尺度方法本身还存在诸多挑战。近些年来,基于分块模型的量子力学方法取得了一定的发展,通过计算分子片段的能量去获得整个大分子体系的能量,这种分块方法的使用可以帮助我们实现更大范围的量子力学计算^[41,42]。

随着人工智能神经网络方法的发展及其在理论化学中的应用,有望减小QM/MM方法中QM区域的计算量,提高QM/MM计算与模拟效率^[43,44]。如Smith等人^[45]发展的ANI神经网络模型,基于高精度从头算方法得到的能量数据库训练而成的神经网络模型,可以改进力场描述的准确性。除此之外,粗粒化模型的发展也是近些年来多尺度模拟研究的重要方向之一。虽然在分子力学水平上的全原子分子动力学模拟具有较高的计算效率,但对于蛋白质复合物或膜蛋白等较大体系复杂结构变化的动力学模拟,其计算成本相当高昂,采用合适的粗粒化模型可以极大地减小计算量,拓展计算模拟的空间尺度和时间尺度。多尺度模拟方法的不断完善和发展,可以更好地兼顾力场或半经验方法的计算效率和高水平量子力学方法的计算准确性,逐渐突破计算模拟体系大小以及时间尺度的限制,构筑计算模拟与实验研究之间的桥梁^[46,47]。

参考文献

- 1 Zhang X, Houk K N. Why enzymes are proficient catalysts: Beyond the Pauling paradigm. *Acc Chem Res*, 2005, 38: 379–385
- 2 Warshel A, Sharma P K, Kato M, et al. Electrostatic basis for enzyme catalysis. *Chem Rev*, 2006, 106: 3210–3235
- 3 Duster A W, Wang C H, Garza C M, et al. Adaptive quantum/molecular mechanics: What have we learned, where are we, and where do we go from here? *Wiley Interdiscip Rev Computat Mol Sci*, 2017, 7: 1–21
- 4 Pezeshki S, Lin H. Recent developments in QM/MM methods towards open-boundary multi-scale simulations. *Mol Simul*, 2015, 41: 168–189
- 5 Wang Y H, Zhang F, Diao H, et al. Covalent inhibition mechanism of antidiabetic drugs—Vildagliptin vs Saxagliptin. *ACS Catal*, 2019, 9: 2292–2302

- 6 Schomburg I, Chang A, Placzek S, et al. BRENDA in 2013: Integrated reactions, kinetic data, enzyme function data, improved disease classification: New options and contents in BRENDA. *Nucleic Acids Res*, 2013, 41: 764–772
- 7 Wang B, Cao Z, Rovira C, et al. Fenton-derived OH radicals enable the MPnS enzyme to convert 2-hydroxyethylphosphonate to methylphosphonate: Insights from *ab initio* QM/MM MD simulations. *J Am Chem Soc*, 2019, 141: 9284–9291
- 8 Chen N, Zhao Y, Lu J, et al. Mechanistic insights into the rate-limiting step in purine-specific nucleoside hydrolase. *J Chem Theor Comput*, 2015, 11: 3180–3188
- 9 She N, Zhao Y, Hao J, et al. Uridine diphosphate release mechanism in O-N-acetylglucosamine (O-GlcNAc) transferase catalysis. *Biochim Biophys Acta*, 2019, 1863: 609–622
- 10 Peräkylä M. Ligand unbinding pathways from the vitamin D receptor studied by molecular dynamics simulations. *Eur Biophys J*, 2009, 38: 185–198
- 11 Tsai P C, Fan Y, Kim J, et al. Structural determinants for the stereoselective hydrolysis of chiral substrates by phosphotriesterase. *Biochemistry*, 2010, 49: 7988–7997
- 12 Manco G, Porzio E, Suzumoto Y. Enzymatic detoxification: A sustainable means of degrading toxic organophosphate pesticides and chemical warfare nerve agents. *J Chem Technol Biotechnol*, 2018, 93: 2064–2082
- 13 Stone R. How to defeat a nerve agent. *Science*, 2018, 359: 23
- 14 Tsai P C, Fox N, Bigley A N, et al. Enzymes for the homeland defense: Optimizing phosphotriesterase for the hydrolysis of organophosphate nerve agents. *Biochemistry*, 2012, 51: 6463–6475
- 15 Cherny I, Greisen Jr P, Ashani Y, et al. Engineering V-type nerve agents detoxifying enzymes using computationally focused libraries. *ACS Chem Biol*, 2013, 8: 2394–2403
- 16 Fan F, Zheng Y, Zhang Y, et al. A comprehensive understanding of enzymatic degradation of the G-type nerve agent by phosphotriesterase: Revised role of water molecules and rate-limiting product release. *ACS Catal*, 2019, 9: 7038–7051
- 17 Kim J, Tsai P C, Chen S L, et al. Structure of diethyl phosphate bound to the binuclear metal center of phosphotriesterase. *Biochemistry*, 2008, 47: 9497–9504
- 18 Porcelli M, Concilio L, Peluso I, et al. Pyrimidine-specific ribonucleoside hydrolase from the archaeon *Sulfolobus solfataricus*-biochemical characterization and homology modeling. *FEBS J*, 2008, 275: 1900–1914
- 19 Giabbai B, Degano M. Crystal structure to 1.7 Å of the *Escherichia coli* pyrimidine nucleoside hydrolase yeiK, a novel candidate for cancer gene therapy. *Structure*, 2004, 12: 739–749
- 20 Versées W, Decanniere K, Van Holsbeke E, et al. Enzyme-substrate interactions in the purine-specific nucleoside hydrolase from *Trypanosoma vivax*. *J Biol Chem*, 2002, 277: 15938–15946
- 21 Vandemeulebroucke A, Minici C, Bruno I, et al. Structure and mechanism of the 6-oxopurine nucleosidase from *Trypanosoma brucei brucei*. *Biochemistry*, 2010, 49: 8999–9010
- 22 Wu R, Gong W, Liu T, et al. QM/MM molecular dynamics study of purine-specific nucleoside hydrolase. *J Phys Chem B*, 2012, 116: 1984–1991
- 23 Fan F, Chen N, Wang Y, et al. QM/MM and MM MD simulations on the pyrimidine-specific nucleoside hydrolase: A comprehensive understanding of enzymatic hydrolysis of uridine. *J Phys Chem B*, 2018, 122: 1121–1131
- 24 Gopaul D N, Meyer S L, Degano M, et al. Inosine-uridine nucleoside hydrolase from *Crithidia fasciculata*. Genetic characterization, crystallization, and identification of histidine 241 as a catalytic site residue. *Biochemistry*, 1996, 35: 5963–5970
- 25 Przybylski J L, Wetmore S D. Modeling the dissociative hydrolysis of the natural DNA nucleosides. *J Phys Chem B*, 2010, 114: 1104–1113
- 26 Horenstein B A, Parkin D W, Estupifiii B, et al. Transition-state analysis of nucleoside hydrolase from *Crithidia fasciculata*. *Biochemistry*, 1991, 33: 10788–10795
- 27 Wang Y, Liu C, Yang W, et al. Naphthalimide derivatives as multifunctional molecules for detecting 5-formylpyrimidine by both PAGE analysis and dot-blot assays. *Chem Commun*, 2018, 54: 1497–1500
- 28 Dubeau S, Bourassa P, Thomas T J, et al. Biogenic and synthetic polyamines bind bovine serum albumin. *Biomacromolecules*, 2010, 11: 1507–1515
- 29 Sudlow G D, Birkett D J, Wade D N. The characterization of two specific drug binding sites on HSA. *Mol Pharmacol*, 1975, 11: 824–832
- 30 Tian Z, Zhao Z, Zang F, et al. Spectroscopic study on the interaction between naphthalimide-polyamine conjugates and DNA. *J Photochem Photobiol B-Biol*, 2014, 138: 202–210
- 31 Fan F, Zhao Y, Cao Z. Insight into the delivery channel and selectivity of multiple binding sites in bovine serum albumin towards naphthalimide-polyamine derivatives. *Phys Chem Chem Phys*, 2019, 21: 7429–7439
- 32 Misini Ignjatović M, Mikulskis P, Söderhjelm P, et al. Can MM/GBSA calculations be sped up by system truncation? *J Comput Chem*, 2018, 39: 361–372
- 33 Zhao Y, Chen N, Wang C, et al. A comprehensive understanding of enzymatic catalysis by hydroxynitrile lyases with *S* stereoselectivity from the *α*

- β -hydrolase superfamily: Revised role of the active-site lysine and kinetic behavior of substrate delivery and sequential product release. *ACS Catal*, 2016, 6: 2145–2157
- 34 Zhao Y, She N, Ma Y, et al. A description of enzymatic catalysis in *N*-acetylhexosamine 1-kinase: Concerted mechanism of two-magnesium-ion-assisted GlcNAc phosphorylation, flexibility behavior of lid motif upon substrate recognition, and water-assisted GlcNAc-1-P release. *ACS Catal*, 2018, 8: 4143–4159
- 35 Wu P, Fan F, Song J, et al. Theory demonstrated a “coupled” mechanism for O_2 activation and substrate hydroxylation by binuclear copper monooxygenases. *J Am Chem Soc*, 2020, 141: 19776–19789
- 36 Zhao Y, Cao Z X. Global simulations of enzymatic catalysis (in Chinese). *Acta Phys-Chim Sin*, 2017, 33: 691–708 [赵媛, 曹泽星. 酶催化过程的全程模拟. 物理化学学报, 2017, 33: 691–708]
- 37 Diao H, Chen N, Wang K, et al. Biosynthetic mechanism of lanosterol: A completed story. *ACS Catal*, 2020, 10: 2157–2168
- 38 Warshel A, Levitt M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol*, 1976, 103: 227–249
- 39 Xie H J, Lei Q F, Fang W J. Combined quantum mechanics and molecular mechanics (in Chinese). *Univ Chem*, 2015, 30: 44–49 [谢湖均, 雷群芳, 方文军. 量子力学和分子力学组方法. 大学化学, 2015, 30: 44–49]
- 40 Bondanza M, Nottoli M, Cupellini L, et al. Polarizable embedding QM/MM: The future gold standard for complex (bio)systems? *Phys Chem Chem Phys*, 2020, 22: 14433–14448
- 41 Li S, Li W, Ma J. Generalized energy-based fragmentation approach and its applications to macromolecules and molecular aggregates. *Acc Chem Res*, 2014, 47: 2712–2720
- 42 He X, Zhu T, Wang X, et al. Fragment quantum mechanical calculation of proteins and its applications. *Acc Chem Res*, 2014, 47: 2748–2757
- 43 Lahey S L J, Rowley C N. Simulating protein-ligand binding with neural network potentials. *Chem Sci*, 2020, 11: 2362–2368
- 44 Zheng S, Rao J, Zhang Z, et al. Predicting retrosynthetic reactions using self-corrected transformer neural networks. *J Chem Inf Model*, 2019, 60: 47–55
- 45 Smith J S, Isayev O, Roitberg A E. ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chem Sci*, 2017, 8: 3192–3203
- 46 Sadeghi M, Noé F. Large-scale simulation of biomembranes incorporating realistic kinetics into coarse-grained models. *Nat Commun*, 2020, 11: 2951
- 47 Zhang Z, Sanbonmatsu K Y, Voth G A. Key intermolecular interactions in the *E. coli* 70S ribosome revealed by coarse-grained analysis. *J Am Chem Soc*, 2011, 133: 16828–16838

Summary for “酶催化过程的量子-经典力学研究进展”

QM/MM study of the enzymatic catalysis

Fangfang Fan¹, Yuzhuang Fu² & Zexing Cao^{2*}

¹ School of Biological and Chemical Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China;

² State Key Laboratory of Physical Chemistry of Solid Surfaces and Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

* Corresponding author, E-mail: zxcao@xmu.edu.cn

Enzyme catalysis generally includes multiple physical and chemical processes, such as the substrate transportation and binding to the active site, the chemical reaction in the protein confined environment, and the product release, where any chemical or nonchemical step is probably to dominate the activity of enzyme. In order to understand the enzymatic efficiency, the classical molecular dynamics (MM MD) and combined quantum and classical molecular dynamics (QM/MM MD) simulations, in combination with the umbrella sampling, have been used to explore mechanisms for the enzymatic catalysis and the substrate transportation in several biological systems, including the detoxification of the G-type nerve agent sarin by phosphotriesterase (PTE), the hydrolysis of uridine by the pyrimidine-specific nucleoside hydrolase Yeik (CU-NH), and the transportation of naphthalimide-polyamine compounds into the drug sites by the bovine serum albumin (BSA).

Extensive QM/MM MD and MM MD simulations show that the wild-type PTE enzyme is capable of hydrolyzing sarin but with a low catalytic efficiency. Dissociation of the degraded product from the binuclear zinc center is predicted to be the rate-determining step with the free-energy barrier of 21.0 kcal/mol, and the multiple chemical steps are involved in the dissociation process. The P-F cleavage of the substrate sarin follows a two-step mechanism, in which the nucleophilic attack of the bridging hydroxide leads to the P-O_μ formation with the free-energy barrier of 9.8 kcal/mol, and the subsequent P-F cleavage and the proton transfer from the μ-OH to the residue Asp301 yields the degraded product with the free-energy span of 12.3 kcal/mol. The relatively strong hydrogen-bond interactions between the substrate uridine and key residues in the active site of CU-NH play an important role in the substrate specificity. The enzymatic *N*-glycosidic bond cleavage of uridine by CU-NH is the rate-limiting step with the free energy barrier of about 20 kcal/mol, and the substrate protonation less influences the activity of CU-NH, completely different from other nucleoside hydrolases. The combined random acceleration molecular dynamics and MD simulations (RAMD MD) have been used to explore possible transportation channels and the most favorable pathways for the delivery of the naphthalimide-polyamine complex to two drug sites of BSA, as well as their thermodynamic and dynamic properties. MM MD and RAMD MD simulations show that the selective specificity of these two drug sites strongly depends on the electron-withdrawing and electron-donating substituents of the substrate, which can be used for the design of new naphthalimide drugs. Such site-selectivity-dependence design strategy may open up an avenue in drug discovery.

With the fast development of low-scaling QM methodologies and the application of artificial intelligence neural network approaches to theoretical and computational chemistry, it can be expected to reduce the computational costs of QM treatment and improve the efficiency of QM/MM, which makes the direct QM/MM study of larger and more complex systems become possible. In addition, the improvement and development of coarse-grained models can surmount the spatial and temporal limitation of multiscale calculations and simulations. Accordingly, global simulations of extremely large and complicated biosystems can be realized in the future, which may provide a theoretical basis for further expanding the application of enzyme engineering.

phosphotriesterase, nucleoside hydrolases, enzyme catalysis, QM/MM and MM MD simulations

doi: [10.1360/TB-2020-1321](https://doi.org/10.1360/TB-2020-1321)