

· 工业蛋白科学 ·

许建和 教授、博士生导师，华东理工大学长江学者特聘教授。2006–2017 年担任生物反应器工程国家重点实验室主任，现任上海生物制造产业技术协同创新中心主任。长期从事生物催化、酶工程及合成生物技术研究。在 *Proc Natl Acad Sci USA*、*Angew Chem*、*ACS Catal*、*Metab Eng*、*Biotech Bioeng* 等期刊发表 SCI 论文 300 余篇。授权中国发明专利 50 项、美国发明专利 1 项。国际期刊 *Bioresources and Bioprocessing* 主编。荣获全国优秀教师、谈家桢生命科学创新奖、杜邦杰能科中国酶工程杰出贡献奖、2018 年上海市技术发明一等奖（第一完成人）。



工业蛋白质构效关系的计算生物学解析

陈琦，李春秀，郑高伟，郁惠蕾，许建和

华东理工大学 生物工程学院，上海 200237

陈琦，李春秀，郑高伟，等. 工业蛋白质构效关系的计算生物学解析. 生物工程学报, 2019, 35(10): 1829–1842.

Chen Q, Li CX, Zheng GW, et al. Computational analysis of structure-activity relationship of industrial enzymes. Chin J Biotech, 2019, 35(10): 1829–1842.

摘要：工业催化用酶已经成为现代生物制造技术的核心“芯片”。不断设计和研发新型高效的酶催化剂是发展工业生物技术的关键。工业催化剂创新设计的科学基础是对酶与底物的相互作用、结构与功能关系及其调控机制的深入剖析。随着生物信息学和智能计算技术的发展，可以通过计算的方法解析酶的催化反应机理，进而对其结构的特定区域进行理性重构，实现酶催化性能的定向设计与改造，促进其工业应用。聚焦工业酶结构-功能关系解析的计算模拟和理性设计，已成为工业酶高效创制改造不可或缺的关键技术。本文就各种计算方法和设计策略以及未来发展趋势进行简要介绍和讨论。

关键词：工业蛋白质，酶，结构-功能关系，分子模拟，理性设计

Received: July 23, 2019; **Accepted:** September 16, 2019

Supported by: Natural Science Foundation of Shanghai China (No. 19ZR1472900), National Natural Science Foundation of China (Nos. 31971380, 21536004, 21672063, 21776085).

Corresponding author: Jianhe Xu. Tel/Fax: +86-21-54250840; E-mail: jianhexu@ecust.edu.cn

上海市自然科学基金 (No. 19ZR1472900), 国家自然科学基金 (Nos. 31971380, 21536004, 21672063, 21776085) 资助。

Computational analysis of structure-activity relationship of industrial enzymes

Qi Chen, Chunxiu Li, Gaowei Zheng, Huilei Yu, and Jianhe Xu

School of Biotechnology, East China University of Science and Technology, Shanghai 200237, China

Abstract: Industrial enzymes have become the core "chip" for bio-manufacturing technology. Design and development of novel and efficient enzymes is the key to the development of industrial biotechnology. The scientific basis for the innovative design of industrial catalysts is an in-depth analysis of the structure-activity relationship between enzymes and substrates, as well as their regulatory mechanisms. With the development of bioinformatics and computational technology, the catalytic mechanism of the enzyme can be solved by various calculation methods. Subsequently, the specific regions of the structure can be rationally reconstructed to improve the catalytic performance, which will further promote the industrial application of the target enzyme. Computational simulation and rational design based on the analysis of the structure-activity relationship have become the crucial technology for the preparation of high-efficiency industrial enzymes. This review provides a brief introduction and discussion on various calculation methods and design strategies as well as future trends.

Keywords: industrial protein, enzyme, structure-activity relationship, molecular simulation, rational design

酶催化为大宗化学品、精细化学品和新型材料等产品提供了一种可持续的绿色合成策略, 已经成为现代生物制造产业的核心“芯片”, 也是国际知识产权争夺的焦点。天然环境中获得的酶催化性能一般较差, 难以满足工业应用对其催化活力、稳定性、溶剂耐受性、底物谱等各个方面的性能要求, 需要进行进一步分子改造。因此, 不断设计和研发新型、高效的酶催化剂是发展工业生物技术的关键所在。

工业催化剂创新设计的科学基础是对酶与底物的相互作用、结构与功能关系及其调控机制的深入剖析。对蛋白质特定区域或位点特征的计算研究主要集中在: 活性位点的构象、动态变化规律、突变效应和反应环境的影响; 底物结合和产物释放; 底物传递; 辅因子结合和释放; 催化机制和过渡态途径等。相关的计算手段主要有分子动力学模拟 (Molecular dynamics simulation, MD)^[1]、量子力学模拟 (Quantum mechanics, QM)、混合量子力学和分子力学 (Quantum mechanics/molecular mechanics, QM/MM) 模拟等^[2]。

分子动力学模拟以经典牛顿力学为理论基础, 模拟获取体系内所有原子随时间运动的轨迹,

通过分析轨迹文件中整个体系的结构或原子之间相互作用变化, 来解析整个体系的动力学性质。通过分子动力学模拟或其他采样方法分析酶的构象灵活性, 可以阐明活性位点和突变位点对酶催化活性的调控作用。混合量子力学和分子力学用量子力学方法计算生物大分子活性部位, 用分子力学、分子动力学或自由能微扰理论计算体系中其他原子, 这样就可得到生物大分子活性部位的电子结构, 通过能量分析, 研究底物的结合模式及酶的催化机理等性质^[3]。QM/MM 能够在计算能量分布和能垒、获得催化机理信息、反应途径、反应速率等方面提供准确性更高的解析。

除了常规的计算方法, 研究者不断开发其他或有针对性的计算策略深入研究蛋白质的结构功能关系。结合计算模拟和实验分析, 不仅可以揭示酶分子催化反应背后的结构机理, 同时可以指导酶的分子改造, 促进其工业应用^[4-5]。Baker 课题组开发了 RosettaDesign 和 RosettaReModel 工具来进行蛋白质的结构和灵活性设计^[6]。JANUS 通过分析多序列比对的结果来预测结构相关但功能不同的酶改造所需的突变^[7]。基于 Web 服务器的 HotSpot Wizard 可以用于自动识别酶结构中的“热

点”氨基酸, 指导针对底物特异性、催化活力或对映选择性的设计改造, 以及蛋白质结构的注释^[8]。POCKETOPTIMIZER 针对活性口袋进行设计, 可用于对酶的活性位点残基进行替换, 提高其对小分子底物的结合能力^[9]。CAVER 和 POVME2 可用于分析通过 MD 模拟所获得的轨迹中底物口袋和通道的形态和大小, 有针对性地指导活性口袋和底物通道的结构改造^[10-11]。荷兰 Janssen 课题组开发的 CASCO 工具可以用于突变体的设计, 通过高通量的 MD 模拟对设计的突变进行排序, 以指导酶的立体选择性改造实验^[12]。FRESCO 方法通过预测折叠自由能、增加二硫键、剔除不合理突变等步骤来设计小巧型突变体文库, 理性指导酶的热稳定性改造^[13]。

1 基于计算生物学的酶分子理性设计

1.1 解析酶的催化机理

多尺度的分子动力学模拟和 QM/MM 模拟已被广泛地用于多种类型的酶分子催化反应机理的解析^[14-18]。

针对二氢叶酸还原酶 ecDHFR 催化循环中的关键问题, 即 Met²⁰ 环在化学氢化物转移步骤中是否采用闭合构象, Major 等使用 QM/MM 模拟研究了野生和突变型 ecDHFR 酶的 Met²⁰ 环在反应中的构象。研究表明, 尽管 Met²⁰ 环的构象对氢化物转移速率和供体-受体距离的相关性不明显, 但它对活性位点水合的有序性具有十分显著的影响^[19]。刘永军课题组利用 QM/MM 探索来源于热带假丝酵母的烯酰基硫酯还原酶 Etr1p 催化烯酰基硫酯为酰基硫酯的催化机制。计算结果表明稳定的共价烯加合物的形成遵循迈克尔加成机理。同时, 结晶水分子虽然不参与催化反应, 但通过形成氢键网络在氢化物转移和质子转移过程中起关键作用^[20]。Loll 等解析了典型的二萜合酶 CotB2 催化反应的化学机制。随着阳离子沿着反应路径的变化, CotB2 活性中的三维结构也随之

改变, 作者确定了活性中心通用的结构特征, 可用于其他来源的萜烯合酶生成大环化合物的理性设计^[21]。Leys 获得了合成单萜的芳樟醇合成酶和 1,8-桉树脑合成酶与氟化底物类似物的复合物晶体结构。模拟表明这些单萜合酶在反应中不会发生显著的构象变化, 这样有助于利用基于结构的蛋白质工程手段, 将这些酶的催化范围扩展到其他单萜底物^[22]。

细胞色素 P450 酶的底物谱非常广泛, Houk 课题组综合使用 DFT 计算和 MD 模拟分析酶分子的构象变化和结合机制, 解析酶催化机制和过渡态^[23]。另外, P450 过氧化酶 OleTJE 使用 H₂O₂ 作为氧化剂, 催化脂肪酸转化为烯烃而无需额外的电子传递。Faponle 等同样综合利用 DFT 和 QM/MM 研究, 针对 OleTJE 进行蛋白质工程改造以提高产品的选择性^[24]。德国 Reetz 课题组获得细胞色素 P450-BM3 单加氧酶可以有效生物催化生成对苯二酚。计算研究证实 P450-BM3 突变体对苯酚选择性氧化, 导致在环氧化物中间体 C3=C4 双键上产生区域选择性^[25]。

Moliner 等开展了关于甘氨酸 N-甲基转移酶和 3 种突变体催化 S_N2 反应实现甲基转移的 QM/MM 理论研究, 验证实验观察到的速率恒定还原和反向次级 α -3H 动力学同位素效应的变化是否应该归因于甲基供体与受体间距离的变化, 计算获得的活化自由能和实验数据非常一致^[26]。同样的, Damborsky 等对卤代烷脱卤素酶 DhaA31 催化水解 1,2,3-三氯丙烷所涉及的主要步骤进行了全面的计算研究。研究发现底物从通道进入活性中心并结合是一个相对快速的过程, 碳-卤键的裂解为反应中的限速步骤。突变体中可催化的构象数目增加且 S_N2 的反应能垒降低, 使得突变体中限速步骤的效率显著增强。最终通过对 DhaA31 的突变实验和对中间突变体的动力学分析证实了理论观察的结果^[27]。

分子氧的活化机制是金属酶研究中的核心问

题, Shaik 课题组针对非血红素羟乙基磷酸酯双加氧酶 HEPD 和甲基磷酸合酶 MPnS 这两种特殊 C-H 加氧酶进行 QM/MM 模拟研究, 结果显示 MPnS 中的 O_2 活化是由活性位点中的相邻赖氨酸 Lys28 介导的, 证实了在金属酶的氧活化机制中存在不寻常的质子穿梭机制^[28]。Faber 等解析了金属依赖型异乳清酸脱羧酶 IDCase 催化 *N*-芳香杂环化合物区域选择性羧化的反应机制和底物特异性机理。结果表明天然底物 5-羧基尿嘧啶的脱羧是通过亲电芳香取代反应进行的, 且杂环底物与周围的氨基酸残基形成独特的氢键作用, 使得 IDCase 显示出较高的底物特异性^[29]。

1.2 提高酶的催化效率

基于催化反应机理的计算解析的催化机理, 可以全面揭示蛋白质的结构-功能关系, 并针对性地对蛋白质结合口袋和活性位点进行理性设计改造。例如, 大位阻的底物对结合位点的构象动力学具有更高的依赖性, 通过突变可以调节结合口袋的灵活性 (如结合口袋的体积), 以形成更适合特定底物识别和稳定的构象状态^[30-31]。

通过在活性位点附近的位置引入单突变, 孔旭东等将来源于巨大芽孢杆菌的环氧水解酶 BmEH 底物范围扩大到更具空间要求的环氧化物底物^[32]。英国 Turner 课题组对来自黑曲霉的单胺氧化酶 MAO-N 野生酶进行研究, 设计获得了一系列突变体能够催化各种大位阻的手性胺底物。这些突变位点不仅位于酶的疏水口袋, 其中一部分还位于远离酶催化活性中心的远端位置^[33-34]。Klinman 课题组通过计算手段理性推测得到大豆脂氧合酶 SLO-1 从溶液到酶活性位点的氧分子通道, 结合定点突变方法, 最终确定氧通道中两个关键氨基酸的调控机制, 并计算出相应突变体的氧分子通道模型^[35]。Shin 等开展 (*S*)-构型选择性转氨酶 OATA 的晶体结构及计算模拟分析, 从分子机制角度解释了酶催化的反应机理, 所获得的

双点突变体对 3 种酮底物的催化活力均有显著提高^[36]。Kaila 等对 FeII/ α -酮戊二酸依赖性的双加氧酶 AsqJ 催化喹诺酮生物碱的多步合成进行研究。发现底物的甲基化对活性位点内 π -堆积相互作用进行微调, 对 AsqJ 的催化反应必不可少。为了对 AsqJ 进行理性设计, 作者扩大了活性位点内的相互作用, 最终获得的工程酶对非甲基化替代物具有显著增强的催化活性^[37]。根据对底物结合模式的分析, Mattevi 对 5-羟甲基糠醛氧化酶 HMFO 氧化 5-甲酰基-2-咪喃甲酸生成 2,5-咪喃二甲酸进行理性设计改造, 两个关键突变增加底物结合亲和力并促进酶与底物醛基以正确结合姿势结合^[38]。类似的策略同样被用于设计改造禾谷镰刀菌壳寡糖氧化酶的底物谱, 使其对 *N*-乙酰-D-葡萄糖胺的底物特异性发生改变, 成功实现对其他非天然底物的高效催化^[39]。

针对来源于光滑假丝酵母 *Candida glabrata* 的野生型羰基还原酶 CgKR1 的底物适用性不够广泛、催化活性不够高的缺陷, 郑高伟等通过分子动力学模拟分析发现位于该酶底物通道 loop 上的 92 位苯丙氨酸与 175 位催化酪氨酸形成 π - π 堆积作用, 阻碍底物与催化残基的结合。基于结构分析寻找到了两个位点进行突变验证, 最终获得一种活性显著提高的突变体对 28 种结构多样性的羰基底物的催化活性均得到显著提高, 可成功用于多种药物手性中间体的扩大制备 (图 1)^[40]。除此之外, 课题组分析了羧酸酯酶 rPPE 的野生型和有益突变体中底物结合模式的差别, 揭示了突变体酶的催化活力、热稳定性和对映选择性提高的结构机理。综合分析酶的底物结合模式、氨基酸位置及序列保守性来寻找潜在敏感的关键氨基酸, 预测候选氨基酸突变对蛋白质稳定性和酶活力的影响程度。虚拟筛选出两个潜在的突变体经试验验证催化活力较母本均有显著提高, 大幅度改善了酶促合成药物关键中间体的催化效

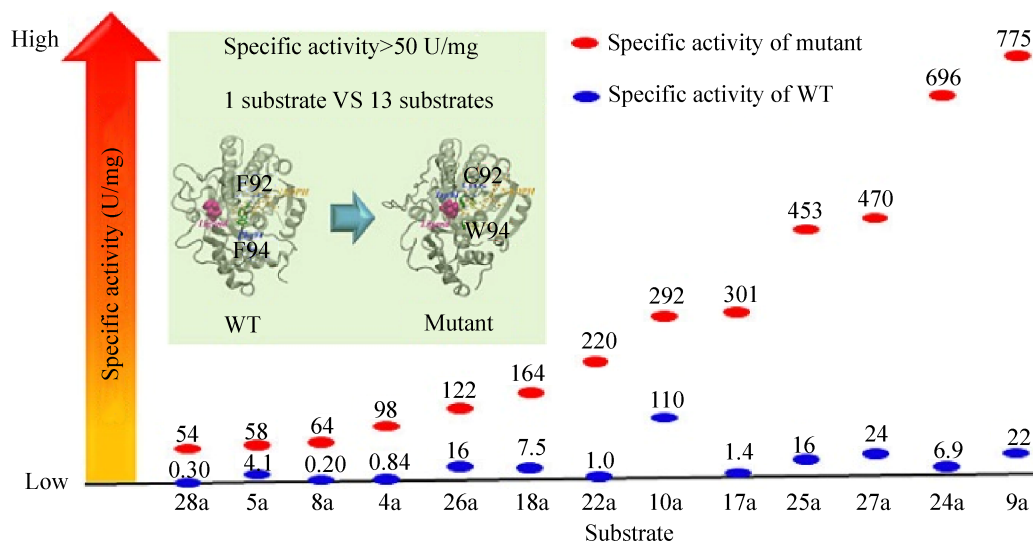


图 1 羧基还原酶 CgKR1 对多种羧基底物的催化活性均显著提高^[40]

Fig. 1 CgKR1 exhibited higher activity toward structurally diverse substrates^[40].

率^[41]。针对胺脱氢酶底物谱较窄的问题，许建和课题组采用计算机辅助的蛋白质工程技术，确定了酶活性口袋中影响酶与大位阻底物结合的关键残基，通过将其突变成分子最小的甘氨酸，实现了对口袋的“裁剪”，拓展了活性口袋的“体积”。将酶催化的底物范围由最长 6 个碳链的脂肪酮拓展至长达 10 个碳链的脂肪酮（如 2-庚酮、2-辛酮、2-壬酮），显著地拓宽了该酶催化的底物范围。此外，这两个关键氨基酸残基的突变在另外两个同源胺脱氢酶的改造中也具有类似的效果，为其他胺脱氢酶的底物谱拓展提供了有益的

参考（图 2）^[42]。

1.3 立体选择性设计

具有对映选择性的酶作用于一些前手性底物上，可以精确地生成所需的光学纯的对映体产物。生物催化剂的立体选择性是由底物结合、反应能垒等多种因素组合而产生的。通过对酶分子构象的研究可以发现，酶的底物结合口袋呈现出一种稳定构象，使得其仅适应生成一种特定的对映体。另一方面，可以通过引入突变改变酶的能量图景，使其能更优先形成所需的对映体，实现立体选择性的逆转^[43]。

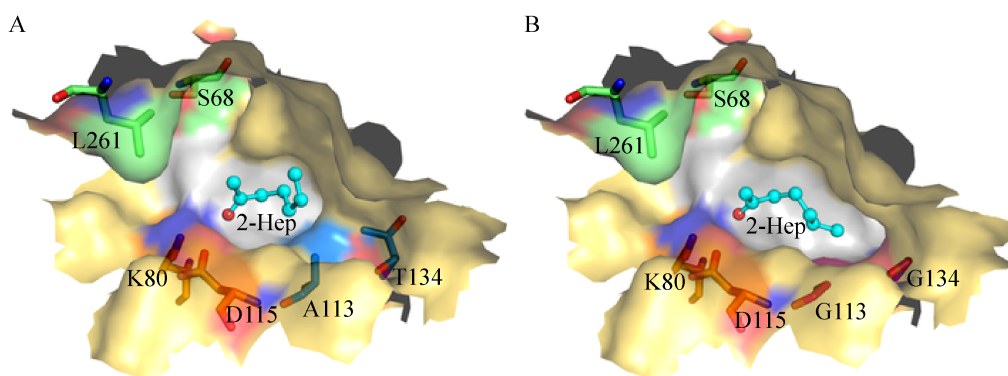


图 2 胺脱氢酶的底物结合口袋改造前后结构对比^[42]

Fig. 2 Structural comparison of the substrate binding pocket for amine dehydrogenase^[42].

QM/MM 和 MD 模拟可以从分子水平上分析调控酶对映选择性的关键因素^[44]。大多数的计算研究都是分析在整个 MD 模拟过程中, 底物分子和关键催化氨基酸之间的距离和角度, 从而定量分析底物分子在催化中心形成可催化的 pro-(S) 或者 pro-(R) 构象的频率^[45-46]。van der Kamp 等模拟解析了酮还原酶 II 型聚酮合成酶 actKR 和两个关键突变体中影响立体选择性的 3 种潜在因素: 利用经典分子动力学评估形成 pro-R 或 pro-S 反应姿势的频率; 采用 MM/PBSA 自由能方法比较 pro-R 或 pro-S 朝向的结合亲和力; 通过 QM/MM MD 模拟与伞形采样评估它们生成 *trans*-1-decalol 反应能垒的差异^[47]。柠檬烯环氧水解酶 LEH 突变体可以催化环戊烯氧化物和环己烯氧化物等环氧氧化物的水解不对称化, 不同突变体具有不同的 (R,R)-或(S,S)-立体选择性。Reetz 等解析了相关突变体对不同底物的特异选择性和立体选择性催化的分子机制。根据计算出的各种途径的活化能, 作者发现了有利反应的潜在立体化学机制。由于活性口袋具有显著的灵活性, 亲核水分子和环戊烯氧化物之间无法形成催化构象。相反, 在催化环己烯氧化物的立体选择性反应时却容易形成反应构象^[48]。相关的研究证明了酶催化中立体选择

性的复杂性以及基于原子的模拟方法在分析这类问题中的重要作用, 同时研究结果也为立体选择性的理性设计提供了理论基础。

设计具有立体选择性的酶的一种途径是降低它们的活性位点体积, 使底物只能以特定的方向结合。亦有研究表明, 利用 MD 模拟进行酶结构灵活性分析可用于识别底物结合口袋附近的关键环结构^[49]。通过调节这些环结构的构象动力学也可以实现酶分子对映选择性的逆转^[50]。Janssen 和 Baker 通过计算与 MD 模拟相结合的手段设计了环氧化物水解酶的活性位点, 成功地获得对映选择性增强的有益突变^[12]。针对环氧水解酶 VrEH1 和 VrEH2 在催化对硝基苯乙烯类氧化物的水解过程中立体选择性不够完善, 造成环氧对映体汇聚反应没有达到 100% 的理论收率的问题。郁惠蕾等在解析 VrEH2 高分辨的酶分子及其与底物复合物晶体结构的基础上, 借助于计算机模拟准确锁定了酶分子中控制底物结合构象的关键氨基酸位点。通过建立了一个小巧型突变体酶库成功靶向到调控酶催化不同对映体区域选择性反应的关键位点。并且发现该位点的调控作用具有一定程度的普适性, 这为同源环氧水解酶的区域选择性改造提供了新思路和新方向 (图 3)^[51]。

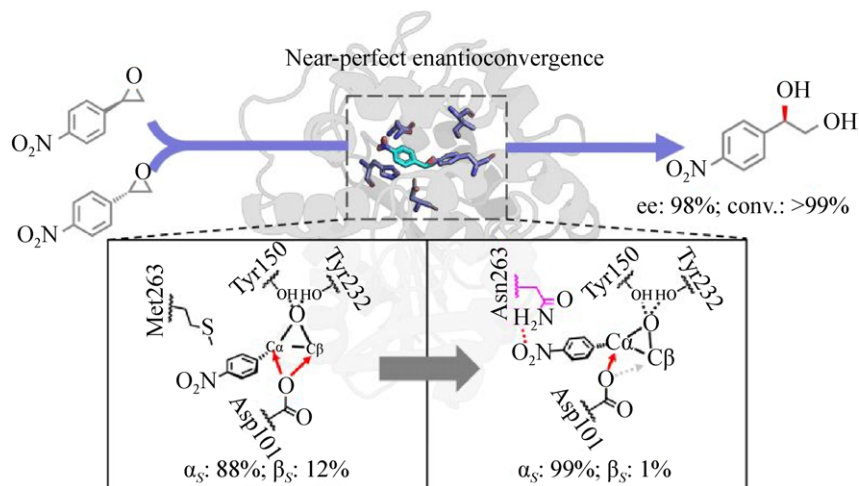


图 3 绿豆环氧水解酶区域选择性改造后实现近乎完美的对映汇聚^[51]

Fig. 3 Near-perfect enantioconvergence of epoxide hydrolase VrEH2^[51].

1.4 提高酶的热稳定性

工业过程所需的酶催化剂需要承受在其自然环境中通常不会遇到的反应条件，例如高温、极端 pH、高盐、高底物浓度、存在有机溶剂等。酶的稳定性依赖于通过一定的分子内相互作用使其能维持特定功能结构。提高蛋白质的内部结构稳定性可以提高蛋白质的热稳定性、溶剂耐受性等。

B-FIT 方法针对蛋白质晶体结构中 B 因子较大的灵活氨基酸残基，通过稳定这些灵活的氨基酸来稳定整个蛋白质的结构。利用 B-FIT 方法，Reetz 等获得了来自枯草芽孢杆菌的脂肪酶 A (LipA) 的突变体，其解链温度 T_m 增加了 45 °C，进一步研究发现稳定化提高与整个蛋白分子上增加的氢键网络以及新的带电氨基酸残基相关^[52]。Bommarius 将基于结构的共义性分析方法和 B-FIT 方法结合起来，成功地将来源于野油菜黄单胞菌的氨基酯水解酶的 T_m 值提高了 8 °C， $k_{d,obs}$ 值提高了 13 倍^[53]。FRESCO 方法将可能的单点突变库减少到 1 634 个，对突变库进行人工检查后，将突变减少到 64 个进行后续实验筛选。最

终获得红色红球菌的柠檬烯环氧化物水解酶的最优突变体在 35 °C 时 T_m 为 85 °C，同时动力学稳定性增加了 250 倍^[13]。许建和课题组利用突变作用的加和性和协同性同时提高了羧基还原酶 *LbCR* 的热稳定性和催化活性。突变残基主要位于靠近活性中心的灵活 loop 环上，热稳定性提高的突变体中引入了新的氢键，并提高了蛋白内部的疏水作用力，从而增强了活性口袋周围灵活 loop 环的刚性。分子对接结果表明，活性提高的突变点虽然与底物没有直接接触，却导致酶和底物之间氢键作用的增加，从而提高了酶对底物的亲和性，稳定了蛋白与底物结合的作用力。此外，热稳定性提高的突变和活性提高的突变位于两个独立的区域，证实突变作用的加和性为热稳定性和活性的同时提高提供了可能。最终作者成功将反应放大，实现了阿托伐他汀前体 6-氰基-3*R*,5*R*-二羟基己酸叔丁酯的规模生物制备 (图 4)^[54]。

通过定向进化发现的热稳定的酶也具有增加的有机溶剂稳定性，如果糖二磷酸醛缩酶^[55]、枯草芽孢杆菌脂肪酶 A^[56]、葡萄糖脱氢酶等^[57]。绘制 T^{50} 与 C^{50} 的关系图发现温度稳定性和有机溶剂

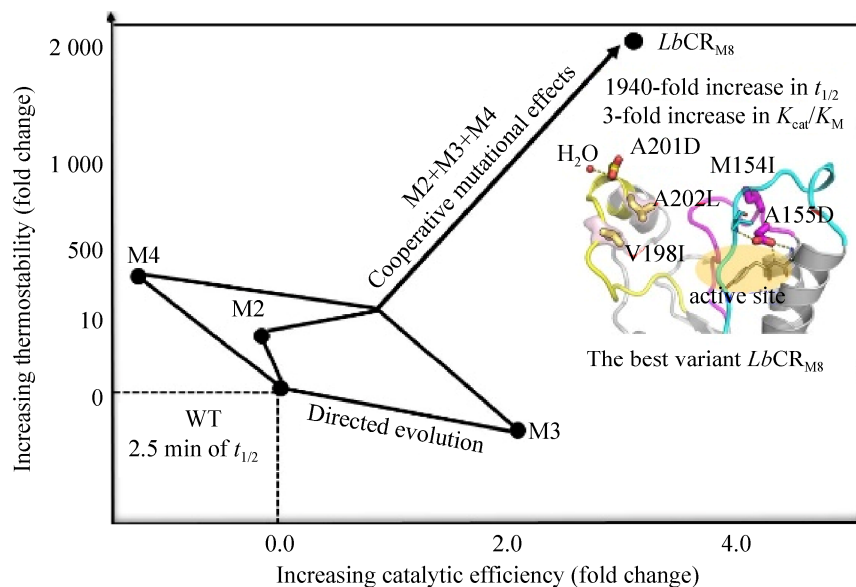


图 4 羧基还原酶 *LbCR* 的热稳定性和催化活性同时提高^[54]

Fig. 4 Engineered *LbCR* with improved thermostability and activity^[54].

稳定性之间存在较强的相关性^[58]。方柏山等利用基因工程手段构建了甲酸脱氢酶超分子自组装体系, 组装的功能片段和融合蛋白形成三维层状结构。将其用于辅酶再生系统时表现出比未组装结构更高的结构稳定性和 NAD(H)再循环效率。利用分子动力学模拟研究超分子组装的构型演变, 发现 FDH 的 C 末端的结构在蛋白质聚合物的形成中起着重要作用^[59]。另外一方面, 热容变化是解释酶催化反应速率的温度依赖性的重要参数, 对酶动力学、热适应性和进化等方面研究都具有重要意义。Arcus 等结合计算和实验手段对类固醇异构酶二聚体和 α -葡糖苷酶分别进行热容的计算分析, 研究发现活性位点周围 loop 环结构的固定、酶的能量波动的变化以及远离活性位点的远端区域对热容变化都具有重要的影响。相关研究证实可以通过原子水平的分子动力学模拟预测催化反应的活化热容变化, 为理性改造酶的最适反应温度提供可行的思路^[60]。

1.5 辅酶依赖性反转

还原酶的催化需要辅酶 NADPH 的参与, 然而额外添加昂贵的辅酶 NADPH, 会导致生产成本过高, 限制了工业化应用。因此, 基于将昂贵 NADPH 替换成相对廉价 NADH 的目的, 研究者们试图通过蛋白质工程手段改造还原酶的辅酶依赖性。

目前有多种方法被成功用于反转还原酶的辅因子特异性, 如: 使用结构信息或使用不同辅酶依赖性的酶同源序列进行多序列比对, 来寻找有益突变。Turner 课题组基于结构和序列比对的方法, 在 *TesADH* 中成功构建了突变体 G198D 可以将辅助因子由 NADP⁺到 NAD⁺约 1 353 倍的反转。最终利用 NADH 依赖型的最优突变体成功实现了将外消旋仲醇高效地转化生成手性胺^[61]。针对辅酶依赖性反转的改造通常会导致酶的催化效率整体下降。2018 年诺贝尔化学奖获得者 Arnold 团队开发了一种计算工具 (CSR-SALAD) 来寻找辅

酶结合关键位点, 可用于简化酶的辅酶 NAD⁺或 NADP⁺依赖性逆转的过程, 该策略同时可以设计出恢复催化活力的半理性突变库。该方法被成功用于 4 种结构不同的 NADP 依赖性的酶 (乙醛酸还原酶、肉桂醇脱氢酶、木糖还原酶和含铁醇脱氢酶) 的辅酶特异性反转, 辅酶依赖性的反变化达 33–4 900 倍, 证实了该策略的有效性^[62]。

针对改变羟甾脱氢酶 7β -HSDHs 辅酶偏好性的研究, 李春秀等提出了一种辅酶特异性反转-小巧突变库的设计策略 (CSR-SaSLiD), 理性地设计了一个包含 5 个突变子的小巧定点突变库, 快速地筛选到辅酶偏好性反转且活性较高的突变体。相较于母本, 突变体实现了辅酶偏好性由 NADPH 到 NADH 约 953 000 倍的反转, 活性提高 223 倍。最后, 作者利用分子动力学模拟揭示了辅酶偏好性反转及活性提高的分子机制, 并将 CSR-SaSLiD 策略成功应用于另外两个 7β -HSDHs 的辅酶偏好性反转和活性提升, 表明 CSR-SaSLiD 是一种有效的改造策略 (图 5)^[63]。

2 新的研究方法

2.1 基于定量构效关系的研究

定量构效关系 (QSAR) 方法被广泛应用于药物设计领域, 目前利用 QSAR 研究底物结构-功能关系的方法已经逐渐延伸到生物催化领域, 成为新的研究热点。酶催化不同底物时将具有不同的催化活力, 可以利用定量构效关系方法对酶催化底物谱进行分析, 构建合理的酶催化活力的定量构效关系模型, 揭示不同底物结构与相应催化活力的关系。

Tomic 课题组提出根据实验数据利用 3D-QSAR 方法构建描述符, 实现了底物专一性常数 k_{cat}/K_m 值的定量预测, 并利用这一方法成功实现了来源于洋葱伯克霍尔德菌 *Bulkholderia cepacia* 酯酶立体选择性的量化预测^[64]。Paier 课题组构建 3D-QSAR

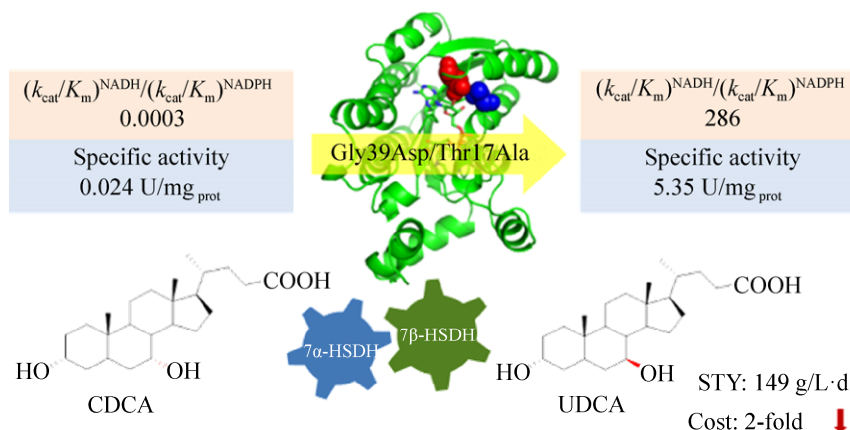


图5 CSR-SaSLiD策略改变羟甾脱氢酶 7β -HSDHs的辅酶偏好性^[63]

Fig. 5 Switching cofactor dependence of 7β -HSDHs using CSR-SaSLiD strategy^[63].

模型分析的环氧化物酶催化反应的立体选择性,分别采用了 CoMFA 和 CoMSIA 两种统计算法,两种方法得到的预测结果显示它们都具有较好的预测准确性^[65]。Braiuca 小组利用实验测得的 k_{cat}/K_m 数据构建回归模型,模型能够准确快速地预测出青霉素酰胺酶的选择性^[66]。于洪巍课题组利用分子对接和定量构效关系相结合的方法,构建了具有较高准确性的南极假丝酵母脂肪酶 B (CALB) 立体选择性的量化预测模型,分别考虑了酰基供体以及 5 个因素(立体、静电、疏水、氢键供体和受体)对模型的影响,证实了酰基供体的重要性^[67]。郁惠蕾等采用 3D-QSAR 方法构建了羰基还原酶 YtbE 的定量构效关系模型。通过分子力场轮廓图考察了底物分子取代基团对活力的影响,证实静电作用和疏水作用在分子力场中起到重要作用。进一步地利用该模型对 ZINC15 小分子数据库进行虚拟筛选,最终成功地筛选到了具有催化活力的化合物,拓展了酶的底物谱^[68]。

2.2 基于氨基酸残基网络的研究

氨基酸残基网络理论将每个氨基酸定义为一个结点、残基间的互相关系数为边。结合网络拓扑参数、氨基酸相互作用、底物结合能等分析,对酶分子整体结构和动态变化进行分析,可以实

现酶分子结构-功能关系的解析。该方法已被成功用于疾病相关蛋白的特异性识别、蛋白质热稳定性等方面的研究。

蒲雪梅等发现单点突变导致 $\beta 2AR$ 蛋白结构中部分螺旋、loop 区及整体结构运动相关性的强弱发生变化,同时也改变了底物结合口袋和信号传导通路的密度,特别是催化氨基酸所在的通路,进而影响药物作用^[69]。Osuna 课题组对一逆醛缩酶活性位点附近和远离活性中心的氨基酸突变所引起的构象变化进行了系统分析和评估。研究表明,酶的构象状态在不同突变体中发生了显著改变,一些较远距离的氨基酸突变会显著影响酶的催化活力^[70]。利用氨基酸残基相互作用网络分析酯酶 *RhEst1* 母本和突变体结构稳定性。许建和课题组研究发现突变体中形成更多的氢键、范德华、 π - π 堆积等相互作用,氨基酸间相互作用的增加稳定了酶的结构稳定性。基于机理分析,结合氨基酸保守性、氨基酸残基相关性和氨基酸间相互作用,对酶的活力、选择性、热稳定性进行了一系列的理性设计和改造研究^[71]。付伟课题组采用动态网络分析方法解析了纤维素酶 *TmCel12A* 突变体热稳定显著提高的结构机理,并找到与突变点具有强相关性的关键氨基酸残基,提出了酶热稳定性改造的新策略^[72]。

2.3 基于机器学习的方法

机器学习技术可以用来识别所有特征变量中的关键特征量,并构建预测模型精确地预测目标参数。研究表明机器学习方法能显著提高蛋白质工程中的筛选通量及有益突变的序列空间的质量和多样性^[73]。

Umetsu 等提出将分子进化与机器学习相结合的蛋白质改造方法。利用第一轮的突变文库构建机器学习模型来指导第二轮的突变。该方法被成功用于绿色荧光蛋白的改造^[74]。Arnold 课题组利用已发表的 GB1 结合蛋白的大量数据分析,证实基于机器学习的定性进化方法可以比其他方法发现更优的突变体。利用这一策略,作者成功地对双加氧酶 *RmaNOD* 催化新型卡宾 Si-H 插入反应生成两种对映体产物进行改造,仅通过两轮突变就找到了优势突变体^[75]。Tidor 等提出了将机器学习与路径采样相结合的新策略来理性寻找影响催化性能的重要区域,并成功地用于酮酸还原异构酶 *KARI* 的研究。作者利用计算统计方法,对过渡态采样来模拟反应的动力学,并将其与机器学习方法相结合。确定了与反应性能相关的重要特征并构建了反应轨迹的预测模型。研究发现,利用单一的描述符就可以准确地预测催化活力,准确度超过 85%。除此以外,研究发现酶-底物复合物中会存在一个重要区域,当这个区域的构象增多时反应性能大幅增加。相关研究不仅证实了在酶-底物复合物的构象空间中存在能促进反应性能的区域,同时开发了用于鉴定和验证这些特别区域的方法,这将大大提高酶分子设计改造的速率^[76]。

2.4 从头设计及人工智能的方法

从头设计的思路是针对一些天然酶无法催化的化学反应,设计全新的蛋白质,使之具有特定的空间结构和预期的功能。从头设计的原则是通过计算获得具有特定空间结构中最低自由能状态

下的氨基酸序列。

美国华盛顿大学 David Baker 团队开发的 Rosetta 算法平台能够从期望的功能反向推测出合适的结构,再计算出最合理的氨基酸序列。利用这一方法, Baker 课题组已成功地实现从头创制逆-醛缩酶^[77]以及催化 Kemp 消除反应^[78]和 Diels-Alder 反应的酶^[79]。近期该课题组从头设计了荧光蛋白 β 桶结构^[80]、自组装螺旋状蛋白纤维^[81]、IL2 和 IL15 受体激动蛋白^[82]、抗癌蛋白 Neo-2/15^[83]、pH 环境变化响应蛋白^[84]、通过诱导构象改变来调节功能的开关蛋白质^[85]等多个复杂的功能蛋白。利用从头设计获得全新的蛋白质为合成生物学和细胞工程的研究提供了新颖的催化元件。

除此之外,谷歌公司 DeepMind 推出新的人工智能项目 AlphaFold,利用经过训练的深度神经网络,可以从基因序列出发预测蛋白质的物理性质和三维结构^[86]。基于深度学习的 AlphaFold 大大缩短了传统计算机模拟计算蛋白质折叠的时间。随着深度学习和人工智能技术的进步,这些技术未来可用于更高效的新酶发掘、更深入的蛋白质结构-功能关系分析及更快速的新药虚拟筛选。

3 总结

越来越多的计算分析方法被用于酶的催化反应机理解析,基于酶分子的结构功能关系的深入解析,对结构中特定区域进行理性重构,可以实现高效的定向设计与改造,推动酶的工业应用。理性设计方法用计算替代部分实验室工作,降低了实验的成本和工作量,目的性更强,更加迅速有效。

针对计算本身具有的局限性,发展更快、更高效、更准确的计算方法是未来的技术发展趋势。合理地应用这些计算方法,将有助于提高工业酶的开发效率,推动生物制造产业的发展。

REFERENCES

- [1] Privett HK, Kiss G, Lee TM, et al. Iterative approach to computational enzyme design. *Proc Natl Acad Sci USA*, 2012, 109(10): 3790–3795.
- [2] Verma R, Mitchell-Koch K. *In silico* studies of small molecule interactions with enzymes reveal aspects of catalytic function. *Catalysts*, 2017, 7(7): 212.
- [3] van der Kamp MW, Mulholland AJ. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry*, 2013, 52(16): 2708–2728.
- [4] Grajciar L, Heard CJ, Bondarenko AA, et al. Towards operando computational modeling in heterogeneous catalysis. *Chem Soc Rev*, 2018, 47(22): 8307–8348.
- [5] Zhu CR, Gao DQ, Ding J, et al. TMD-based highly efficient electrocatalysts developed by combined computational and experimental approaches. *Chem Soc Rev*, 2018, 47(12): 4332–4356.
- [6] Huang PS, Ban YEA, Richter F, et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS ONE*, 2011, 6(8): e24109.
- [7] Addington TA, Mertz RW, Siegel JB, et al. Janus: prediction and ranking of mutations required for functional interconversion of enzymes. *J Mol Biol*, 2013, 425(8): 1378–1389.
- [8] Bendl J, Stourac J, Sebestova E, et al. HotSpot Wizard 2.0: automated design of site-specific mutations and smart libraries in protein engineering. *Nucleic Acids Res*, 2016, 44(W1): W479–W487.
- [9] Malisi C, Schumann M, Toussaint NC, et al. Binding pocket optimization by computational protein design. *PLoS ONE*, 2012, 7(12): e52505.
- [10] Chovancova E, Pavelka A, Benes P, et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol*, 2012, 8(10): e1002708.
- [11] Durrant JD, Votapka L, Sørensen J, et al. POVME 2.0: an enhanced tool for determining pocket shape and volume characteristics. *J Chem Theory Comput*, 2014, 10(11): 5047–5056.
- [12] Wijma HJ, Floor RJ, Bjelic S, et al. Enantioselective enzymes by computational design and *in silico* screening. *Angew Chem Int Ed*, 2015, 54(12): 3726–3730.
- [13] Wijma HJ, Floor RJ, Jekel PA, et al. Computationally designed libraries for rapid enzyme stabilization. *Protein Eng Des Sel*, 2014, 27(2): 49–58.
- [14] Park SB, Bae DW, Clavio NAB, et al. Structural and biochemical characterization of the curcumin-reducing activity of CurA from *Vibrio vulnificus*. *J Agric Food Chem*, 2018, 66(40): 10608–10616.
- [15] Ren WW, Pengelly R, Farren-Dai M, et al. Revealing the mechanism for covalent inhibition of glycoside hydrolases by carbasugars at an atomic level. *Nat Commun*, 2018, 9(1): 3243.
- [16] Fan FF, Chen NH, Wang YH, 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(3): 1121–1131.
- [17] Kulkarni YS, Liao QH, Petrović D, et al. Enzyme architecture: modeling the operation of a hydrophobic clamp in catalysis by triosephosphate isomerase. *J Am Chem Soc*, 2017, 139(30): 10514–10525.
- [18] Schrepfer P, Buettner A, Goerner C, et al. Identification of amino acid networks governing catalysis in the closed complex of class I terpene synthases. *Proc Natl Acad Sci USA*, 2016, 113(8): E958–E967.
- [19] Mhashal AR, Vardi-Kilshtain A, Kohen A, et al. The role of the Met²⁰ loop in the hydride transfer in *Escherichia coli* dihydrofolate reductase. *J Biol Chem*, 2017, 292(34): 14229–14239.
- [20] Ling BP, Li H, Yan LJ, et al. Conversion mechanism of enoyl thioesters into acyl thioesters catalyzed by 2-enoyl-thioester reductases from *Candida Tropicalis*. *Phys Chem Chem Phys*, 2019, 21(19): 10105–10113.
- [21] Driller R, Janke S, Fuchs M, et al. Towards a comprehensive understanding of the structural dynamics of a bacterial diterpene synthase during catalysis. *Nat Commun*, 2018, 9: 3971.
- [22] Karuppiah V, Ranaghan KE, Leferink NGH, et al. Structural basis of catalysis in the bacterial monoterpene synthases linalool synthase and 1,8-Cineole synthase. *ACS Catal*, 2017, 7(9): 6268–6282.
- [23] Haatveit KC, Garcia-Borràs M, Houk KN. Computational protocol to understand P450 mechanisms and design of efficient and selective biocatalysts. *Front Chem*, 2019, 6: 663.
- [24] Faponle AS, Quesne MG, De Visser SP. Origin of the

- regioselective fatty-acid hydroxylation versus decarboxylation by a cytochrome P450 peroxygenase: what drives the reaction to biofuel production? *Chemistry*, 2016, 22(16): 5478–5483.
- [25] Zhou HY, Wang BJ, Wang F, et al. Chemo- and regioselective dihydroxylation of benzene to hydroquinone enabled by engineered cytochrome P450 monooxygenase. *Angew Chem Int Ed*, 2019, 58(3): 764–768.
- [26] Świderek K, Tuñón I, Williams IH, et al. Insights on the origin of catalysis on glycine *N*-methyltransferase from computational modeling. *J Am Chem Soc*, 2018, 140(12): 4327–4334.
- [27] Marques SM, Dunajova Z, Prokop Z, et al. Catalytic cycle of haloalkane dehalogenases toward unnatural substrates explored by computational modeling. *J Chem Inf Model*, 2017, 57(8): 1970–1989.
- [28] Wang BJ, Cao ZX, 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(23): 9284–9291.
- [29] Sheng X, Plasch K, Payer SE, et al. Reaction mechanism and substrate specificity of *Iso*-orotate decarboxylase: a combined theoretical and experimental study. *Front Chem*, 2018, 6: 608.
- [30] Behiry EM, Ruiz-Pernia JJ, Luk L, et al. Isotope substitution of promiscuous alcohol dehydrogenase reveals the origin of substrate preference in the transition state. *Angew Chem Int Ed*, 2018, 57(12): 3128–3131.
- [31] Liu J, Wu SK, Li Z. Recent advances in enzymatic oxidation of alcohols. *Curr Opin Chem Biol*, 2018, 43: 77–86.
- [32] Kong XD, Yuan SG, Li L, et al. Engineering of an epoxide hydrolase for efficient bioresolution of bulky pharmaco substrates. *Proc Natl Acad Sci USA*, 2014, 111(44): 15717–15722.
- [33] Herter S, Medina F, Wagschal S, et al. Mapping the substrate scope of monoamine oxidase (MAO-N) as a synthetic tool for the enantioselective synthesis of chiral amines. *Bioorg Med Chem*, 2018, 26(7): 1338–1346.
- [34] Li GY, Yao PY, Gong R, et al. Simultaneous engineering of an enzyme's entrance tunnel and active site: the case of monoamine oxidase MAO-N. *Chem Sci*, 2017, 8(5): 4093–4099.
- [35] Collazo L, Klinman JP. Control of the position of oxygen delivery in soybean lipoxygenase-1 by amino acid side chains within a gas migration channel. *J Biol Chem*, 2016, 291(17): 9052–9059.
- [36] Han SW, Kim J, Cho HS, et al. Active site engineering of ω -transaminase guided by docking orientation analysis and virtual activity screening. *ACS Catal*, 2017, 7(6): 3752–3762.
- [37] Mader SL, Bräuer A, Groll M, et al. Catalytic mechanism and molecular engineering of quinolone biosynthesis in dioxxygenase AsqJ. *Nat Commun*, 2018, 9(1): 1168.
- [38] Dijkman WP, Binda C, Fraaije MW, et al. Structure-based enzyme tailoring of 5-hydroxymethylfurfural oxidase. *ACS Catal*, 2015, 5(3): 1833–1839.
- [39] Ferrari AR, Lee M, Fraaije MW. Expanding the substrate scope of chitoooligosaccharide oxidase from *Fusarium graminearum* by structure-inspired mutagenesis. *Biotechnol Bioeng*, 2015, 112(6): 1074–1080.
- [40] Zheng GW, Liu YY, Chen Q, et al. Preparation of structurally diverse chiral alcohols by engineering ketoreductase CgKR1. *ACS Catal*, 2017, 7(10): 7174–7181.
- [41] Chen Q, Luan ZJ, Cheng XL, et al. Molecular dynamics investigation of the substrate binding mechanism in carboxylesterase. *Biochemistry*, 2015, 54(9): 1841–1848.
- [42] Chen FF, Zheng GW, Liu L, et al. Reshaping the active pocket of amine dehydrogenases for asymmetric synthesis of bulky aliphatic amines. *ACS Catal*, 2018, 8(3): 2622–2628.
- [43] Maria-Solano MA, Serrano-Hervás E, Romero-Rivera A, et al. Role of conformational dynamics in the evolution of novel enzyme function. *Chem Commun (Camb)*, 2018, 54(50): 6622–6634.
- [44] Romero-Rivera A, Garcia-Borràs M, Osuna S. Computational tools for the evaluation of laboratory-engineered biocatalysts. *Chem Commun (Camb)*, 2016, 53(2): 284–297.
- [45] Wijma HJ, Marrink SJ, Janssen DB. Computationally efficient and accurate enantioselectivity modeling by clusters of molecular dynamics simulations. *J Chem Inf Model*, 2014, 54(7): 2079–2092.
- [46] Maria-Solano MA, Romero-Rivera A, Osuna S.

- Exploring the reversal of enantioselectivity on a zinc-dependent alcohol dehydrogenase. *Org Biomol Chem*, 2017, 15(19): 4122–4129.
- [47] Serapian SA, van der Kamp MW. Unpicking the cause of stereoselectivity in actinorhodin ketoreductase variants with atomistic simulations. *ACS Catal*, 2019, 9(3): 2381–2394.
- [48] Sun ZT, Salas PT, Siirola E, et al. Exploring productive sequence space in directed evolution using binary patterning versus conventional mutagenesis strategies. *Bioresour Bioprocess*, 2016, 3: 44.
- [49] Sun ZT, Wu L, Bocola M, et al. Structural and computational insight into the catalytic mechanism of limonene epoxide hydrolase mutants in stereoselective transformations. *J Am Chem Soc*, 2018, 140(1): 310–318.
- [50] Yang B, Wang HJ, Song W, et al. Engineering of the conformational dynamics of lipase to increase enantioselectivity. *ACS Catal*, 2017, 7(11): 7593–7599.
- [51] Li FL, Kong XD, Chen Q, et al. Regioselectivity engineering of epoxide hydrolase: near-perfect enantioconvergence through a single site mutation. *ACS Catal*, 2018, 8(9): 8314–8317.
- [52] Reetz MT, Soni P, Acevedo JP, et al. Creation of an amino acid network of structurally coupled residues in the directed evolution of a thermostable enzyme. *Angew Chem Int Ed Engl*, 2009, 48(44): 8268–8272.
- [53] Blum JK, Ricketts MD, Bommarius AS. Improved thermostability of AEH by combining B-FIT analysis and structure-guided consensus method. *J Biotechnol*, 2012, 160(3/4): 214–221.
- [54] Gong XM, Qin Z, Li FL, et al. Development of an engineered ketoreductase with simultaneously improved thermostability and activity for making a bulky Atorvastatin precursor. *ACS Catal*, 2019, 9(1): 147–153.
- [55] Hao JJ, Berry A. A thermostable variant of fructose bisphosphate aldolase constructed by directed evolution also shows increased stability in organic solvents. *Protein Eng Des Sel*, 2004, 17(9): 689–697.
- [56] Reetz MT, Soni P, Fernández L, et al. Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem Commun*, 2010, 46(45): 8657–8658.
- [57] Vazquez-Figueroa E, Yeh V, Broering JM, et al. Thermostable variants constructed via the structure-guided consensus method also show increased stability in salts solutions and homogeneous aqueous-organic media. *Protein Eng Des Sel*, 2008, 21(11): 673–680.
- [58] Bommarius AS. Biocatalysis: a status report. *Annu Rev Chem Biomol Eng*, 2015, 6: 319–345.
- [59] Jiang W, Yang RN, Lin P, et al. Bioinspired genetic engineering of supramolecular assembled formate dehydrogenase with enhanced biocatalysis activities. *J Biotechnol*, 2019, 292: 50–56.
- [60] van der Kamp MW, Prentice EJ, Kraakman KL, et al. Dynamical origins of heat capacity changes in enzyme-catalysed reactions. *Nat Commun*, 2018, 9: 1177.
- [61] Thompson MP, Turner NJ. Two-enzyme hydrogen-borrowing amination of alcohols enabled by a cofactor-switched alcohol dehydrogenase. *ChemCatChem*, 2017, 9(20): 3833–3836.
- [62] Cahn JKB, Werlang CA, Baumschlager A, et al. A general tool for engineering the NAD/NADP cofactor preference of oxidoreductases. *ACS Synth Biol*, 2017, 6(2): 326–333.
- [63] You ZN, Chen Q, Shi SC, et al. Switching cofactor dependence of 7 β -Hydroxysteroid dehydrogenase for cost-effective production of ursodeoxycholic acid. *ACS Catal*, 2019, 9(1): 466–473.
- [64] Tomić S, Kojić-Prodić BA. Quantitative model for predicting enzyme enantioselectivity: application to *Burkholderia cepacia* lipase and 3-(aryloxy)-1, 2-propanediol derivatives. *J Mol Graph Model*, 2001, 21(3): 241–252.
- [65] Paier J, Stockner T, Steinreiber A, et al. Enantioselectivity of epoxide hydrolase catalysed oxirane ring opening: a 3D QSAR study. *J Comput Aided Mol Des*, 2003, 17(1): 1–11.
- [66] Braiuca P, Boscarol L, Ebert C, et al. 3D-QSAR applied to the quantitative prediction of penicillin G amidase selectivity. *Adv Synth Catal*, 2006, 348(6): 773–780.
- [67] Gu JL, Liu J, Yu HW. Quantitative prediction of enantioselectivity of *Candida antarctica* lipase B by combining docking simulations and quantitative structure-activity relationship (QSAR) analysis. *J Mol Catal B: Enzym*, 2011, 72(3/4): 238–247.
- [68] Cheng Q, Chen Q, Xu JH, et al. A 3D-QSAR assisted

- activity prediction strategy for expanding substrate spectra of an aldehyde ketone reductase. *Mol Catalys*, 2018, 455: 224–232.
- [69] Gao N, Liang T, Yuan Y, et al. Exploring the mechanism of F282L mutation-caused constitutive activity of GPCR by a computational study. *Phys Chem Chem Phys*, 2016, 18(42): 29412–29422.
- [70] Romero-Rivera A, Garcia-Borràs M, Osuna S. Role of conformational dynamics in the evolution of retro-aldolase activity. *ACS Catal*, 2017, 7(12): 8524–8532.
- [71] Chen Q, Yu HL, Cheng XL, et al. Structural investigation of the enantioselectivity and thermostability mechanisms of esterase *RhEst1*. *J Mol Graph Model*, 2018, 85: 182–189.
- [72] Lian P, Yuan CM, Xu Q, et al. Thermostability mechanism for the hyperthermophilicity of extremophile cellulase *TmCel12A*: Implied from molecular dynamics simulation. *J Phys Chem B*, 2016, 120(30): 7346–7352.
- [73] Yang KK, Wu Z, Arnold FH. Machine-learning-guided directed evolution for protein engineering. *Nat Methods*, 2019, 16(8): 687–694.
- [74] Saito Y, Oikawa M, Nakazawa H, et al. Machine-learning-guided mutagenesis for directed evolution of fluorescent proteins. *ACS Synth Biol*, 2018, 7(9): 2014–2022.
- [75] Wu Z, Kan SBJ, Lewis RD, et al. Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc Natl Acad Sci USA*, 2019, 116(18): 8852–8858.
- [76] Bonk BM, Weis JW, Tidor B. Machine learning identifies chemical characteristics that promote enzyme catalysis. *J Am Chem Soc*, 2019, 141(9): 4108–4118.
- [77] Jiang L, Althoff EA, Clemente FR, et al. De novo computational design of retro-aldol enzymes. *Science*, 2008, 319(5868): 1387–1391.
- [78] Röthlisberger D, Khersonsky O, Wollacott AM, et al. Kemp elimination catalysts by computational enzyme design. *Nature*, 2008, 453(7192): 190–195.
- [79] Siegel JB, Zanghellini A, Lovick HM, et al. Computational design of an enzyme catalyst for a stereo selective bimolecular Diels-Alder reaction. *Science*, 2010, 329(5989): 309–313.
- [80] Lu PL, Min DY, DiMaio F et al. Accurate computational design of multipass transmembrane proteins. *Science*, 2018, 359(6379): 1042–1046.
- [81] Dou JY, Vorobieva AA, Sheffler W, et al. *De novo* design of a fluorescence-activating β -barrel. *Nature*, 2018, 561(7724): 485–491.
- [82] Shen H, Fallas JA, Lynch E, et al. *De novo* design of self-assembling helical protein filaments. *Science*, 2018, 362(6415): 705–709.
- [83] Silva DA, Yu S, Ulge UY, et al. *De novo* design of potent and selective mimics of IL-2 and IL-15. *Nature*, 2019, 565(7738): 186–191.
- [84] Boyken SE, Benhaim MA, Busch F, et al. *De novo* design of tunable, pH-driven conformational changes. *Science*, 2019, 364(6441): 658–664.
- [85] Langan RA, Boyken SE, Ng AH, et al. *De novo* design of bioactive protein switches. *Nature*, 2019, 572(7768): 205–210.
- [86] Evans R, Jumper J, Kirkpatrick J, et al. *De novo* structure prediction with deep-learning based scoring//Thirteenth Critical Assessment of Techniques for Protein Structure Prediction (Abstracts). Iberostar, Paraiso, 2018.

(本文责编 陈宏宇)