

从计算生物学分析工业蛋白质构效关系

宁超

天津医科大学总医院精准医学中心, 天津 300052

摘要: 工业催化用酶属于现代生物制造技术主要芯片, 新型高效酶催化剂的研发为工业生物技术的发展重点。深入分析酶和底物作用、功能关系、结构为工业催化剂的设计基础, 在生物信息学与智能计算技术不断发展的过程中, 能够利用计算方法对酶催化反应机理进行计算, 从而重构结构特定区域, 对酶催化性能进行定向设计和改造。重视工业酶结构-功能关系解析的解析的理性设计和计算模拟, 为现代工业酶高效创制改造中的主要技术。以此, 本文就基于计算生物学角度对工业蛋白酶构效关系进行分析。

关键词: 计算生物学; 工业蛋白质; 分子模拟

酶催化为生物催化剂, 主要优势就是具有较高专一性、反应条件吻合、环境友好等。天然酶具备大量错综复杂结构, 从而实现多种类型的催化反应。酶结构和其功能密切相关, 针对大部分酶来说, 催化功能与酶催化口袋中氨基酸残基相关。在晶体结构不断被解析和利用半理性设计、随即突变等定向技术得到大量催化特性优化工业酶来说, 人们能够更好的理解酶结构和功能的关系^[1]。目前, 针对晶体结构解析和计算分析, 科学家能够深入掌握酶催化机制和相应结构特征, 从而促进了工业应用。

1 蛋白质结构预测

在蛋白质构效关系分析中, 蛋白质结构为分析基础, 重视研究实验数据精准性。蛋白质序列数据库 UniProt 与结构数据库 PDB 中的数据条目在不断的增加, 但是只有 0.2% 的蛋白质序列具备实验测定三维结构。所以, 蛋白质结构预测精准度为构造关系研究基础。

美国科学家安分森在 1973 年表示, 在指定环境中, 蛋白质折叠需要的所有信息都在氨基酸序列中。基于物理学角度分析, 氨基酸序列决定蛋白质基本分析构成, 最终折叠结构对应制定环境下此分子热力学稳定的构象态。因为大分子比较复杂, 无法得出精准物理描述, 并且没有计算条件实现计算。所以, 研究者开始研究粗粒化物理模型和优化构象搜索算法^[2]。比较建模为蛋白质三维结构预测主要方法, 其也称之为同源建模, 主要原理为将进化相关序列

通讯作者: 宁超, 天津医科大学总医院精准医学中心。

Copyright © Sanderman Publishing House

This work is licensed under a Commons Attribution-Non Commercial 4.0 International License.

<http://creativecommons.org/licenses/by/4.0/>



具备相同三维结构，并且进化中三维结构保守，通过进化模板结构信息建模，所以要和目标序列相似度高的已知结构模板。比较建模的重点就是针对模板实现插入与空隙建模，目标蛋白和模板结构保守性程序对预测模型精准性具有一定的影响。

在缺少高同源性模板的时候，要求使用复杂方法得到合适模板从而进行对比，此过程为穿线法。穿线方法其实就是利用某策略对比结构和序列，评估使序列通过各匹配方法安装到三位结构中的舒服程度，所以也称之为折叠辨识。此方法的有效性是由于蛋白质三维结构的保守性比较强，蛋白质折叠子总数量有限。在三维结构数据库不断扩充的背景下，积累了大量数据对 folds 覆盖，从而促进了穿线法的发展，尤其只能够寻找同源性在 30% 以下的模板中使用^[3]。

2 蛋白质的可塑性

在温和条件下，酶能够催化特异化学反应，是否和底物选择性结合为酶执行催化功能的主要因素。酶结构特定空间排列为构成酶-底物过渡态复合物的基础，氨基酸是利用酰胺键相互连接的多肽链骨架，能够卷曲折叠构成规则二级结构，比如 α -螺旋、 β -折叠和 β -转角等。二级结构根据非共价力，比如疏水作用、离子键、氢键进行折叠组装构成有功能三级结构。因为非共价键比较弱，外界环境会对三级结构造成影响，从而出现构象变化。所以，酶结构具备可塑性，从而对分子特性与催化功能造成影响^[4]。

酶骨架构象并不是静止、刚性的，而是动态、柔性的。酶和底物分子作用过程中，酶结构会出现轻微改变，有的时候还会出现剧烈变构，酶骨架可塑性对功能实施进行保障。比如羧酸还原酶，其能够催化羧酸到醛还原，赛格尼氏杆菌中羧酸还原酶 SrCAR 晶体结构表示，腺苷化结构与硫酸酯化结构存在变构效应，从而通过同个催化口袋对硫酸酯化与腺苷化的反应进行催化。曲戈等人研究表示，SrCAR 中的 K528 与 K629 在硫酸酯化与腺苷化中具有重要作用。K629 处于腺苷化结构域催化口袋的中心，硫酸酯化阶段出现变化，催化口袋移除 K29，使用 K528 代替催化硫酸酯化反应。使此位点突变成为丙氨酸之后，两者均无法实现模式底物催化能力。另外，结构域或者拓扑结构出现大尺度变构具有重要作用，比如卤代链烷酸脱卤素酶 HAD 超家族以不同的功能划分成为 ATP 酶、磷酸酯酶、糖磷酸酶与脱卤素酶，不同成员的 Cap 结构域的差别导致底物与功能各有不同^[5]。

酶骨架可塑性使催化功能具备混杂性，南极假丝酵母脂肪酶 B 具备甘油三酯水解与催化加成反应的功能，并且骨架可塑性会对酶热稳定性造成影响，利用海栖热袍菌种纤维素酶对比同家族其他晶体结构表示，此酶具有长且高度扭转的 β -链 B8 与 B9，此氨基酸利用氢键与堆叠作用力对底物稳定，此结构特征对酶高热稳定性具有一定的影响^[6]。

3 基于计算生物学的蛋白质构效关系

3.1 解析酶的催化机理

多尺度分子动力学模拟与 QM/MM 模拟能够在多类型酶分子催化反应机理解析中使用, 对于二氢叶酸还原酶 ecDHFR 催化循环中的问题, 也就是 Met20 环在化学氢化物转移步骤中是否使用闭合构象, Major 等利用 OM/MM 模拟对突变型与野生 ecDHFR 酶中 Met20 环在反应中构象进行研究。研究结果表明, 虽然 Met20 环构象对于氢化物转移速率与供体-受体距离并没有明显相关性, 但是会影响到活性位点水合有序性。刘勇军课题组通过 QM/MM 对热带假丝酵母中烯酰基硫脂还原酶 Etr1p 催化烯酰硫脂催化机制, 分析结果表明稳定共价烯加合物满足迈克尔加成机理。另外, 虽然结晶水分子不进行催化反应, 但是利用构成氢键网络在氢化物转移与质子转移中具有重要作用^[7]。

3.2 提高酶的催化效率

将催化反应机理作为基础的催化机理计算解析能够将蛋白质结构-功能关系充分的反应出来, 针对性的设计蛋白质结合活性位点、口袋位点。比如, 大位阻底物对于结合位点构象动力学的依赖性比较高, 利用突变对口袋灵活性调节结合, 从而构成合适的底物识别与稳定构象状态。将单突变应用到活性位点附近位置中, 孔旭东等人使源于巨大芽孢杆菌环氧水解酶 BmEH 底物范围在具备空间要求比较大的环氧化物底物中使用。Klinman 课题组利用计算手段对大豆脂氧合酶 SLO-1 通过溶液到酶活性位点氧分子通道中推测, 与定点突变方法结合, 从而对氧通道的关键氨基酸调控机制进行确定, 并且对突变体氧分子通道模型进行计算。Shin 等进行(S)-构型选择性转氨酶 OATA 晶体结构和计算模拟的分析, 基于分子机制角度对酶催化反应进行解释, 得到双电突变体能够提高三种酮底物催化活力。分析底物结合模式, Mattevi 对于 5-羟甲基糠醛氧化酶 HMFO 氧化 5-甲酰基-2-呋喃甲酸的生成进行理性设计, 其中的关键突变能够使底物亲和力得到增加, 从而使酶和底物醛基通过正确的姿势相互结合^[8]。

3.3 立体选择性设计

基于对映选择性酶作用在前手性底物方面, 能够精准生成所需光学纯对映体产物。生物催化剂立体选择性是通过反映能垒、底物结合等因素相互结合构成, 利用研究酶分子构象表示, 酶底物结合口袋具有稳定构象, 从而使适应生成指定对映体。另外, 利用突变改变酶能量图景, 使其优先构成需要的对映体, 逆转立体选择性。QM/MM 与 MD 模拟能够通过分子水平方面分析调控酶相应选择性关键因素, 大部分计算研究都是对整体 MD 模拟的分析过程, 底物分子与关键催化氨基酸的距离与角度, 对底物分子在催化中心构成可以催化 pro- (S) 频率。Van 等人对酮还原酶 II 型巨酮合成酶 actKP 与关键突变体中对立体选择性造成影响的潜在因素进行分析: 通过经典分子动力学对 pro-R 反应姿势频率进行评估; 使用 MM/PBSA 自由能的方法对比; 利用 QM/MM MD 模拟和伞形采样对生成的 trans-1 反映能垒差异进行评估。

Janssen 等人结合 MD 模拟手段实现环氧化物水解酶活性位点的结合, 得到对映选择性增强有益突变。对于环氧水解酶 VtEH1 缺乏完善的水解过程立体选择性, 从而导致环氧对于映体汇聚反应无法达到 100%理论收率等问题。利用小巧突变体酶库成功靶向到不同调控酶催化到映体区域选择性反映关键位点, 表示此位点调控作用良好, 为同源环氧水解酶区域选择性的改造提供全新思路^[9]。

4 结束语

目前, 计算分析方法在酶催化反应机理解析中使用, 深入分析酶分子结构功能的关系, 重构结构中特定区域, 能够实现高效定向设计和改造, 促进酶工业使用。理性设计方法利用计算代替实验室工作, 使实验成本与工作量降低, 具有较强的目的性。对于计算自身的局限性, 发展高效、快速且精准的计算方法为技术未来发展趋势。合理使用计算方法, 能够使工业酶开发效率得到提高, 促进生物制造产业发展。

参考文献

- [1]余晓, 涂艺声, 叶丽楠, et al. TMT 蛋白质组学与生物信息学分析不同石杉碱甲含量的蛇足石杉叶状体的蛋白差异[J]. 中草药 2019 (12).
- [2]杨强强, 刘军娜, 熊越, et al. 新型黄杨属生物碱衍生物结构与性能的研究[C]// 河南省化学会 2018 年学术年会. 2018.
- [3]杨璐, 胡成进. 副溶血弧菌 CalR 蛋白结构与功能的生物信息学分析[J]. 中国病原生物学杂志, 2019 (10) 1135-1140.
- [4]张银. 工业酶结构与功能的构效关系[J]. 生物工程学报 2019 35 (10) 1806-1818.
- [5]程少文, 陈扬平, 张安强, et al. MBL2 与 MASPs 家族蛋白质相互作用的生物信息学分析[J]. 海南医学院学报 2019 (15) 1121-1124.
- [6]敖竹兰, 郭永辉, 江琛, et al. 水稻抗逆性基因 OsGATA23 可变剪切的生物信息学分析[J]. 湖北大学学报:自然科学版 2019 (6) 584-591.
- [7]王浩, 王春晴, 陈瑞冰. 结合亲和质谱与生物信息学分析构建 TP53BP1 的蛋白相互作用网络[J]. 中国科学:生命科学 2018 (2).
- [8]张一菡, 李卿, 张磊, et al. 西红柿 MADS-box 基因家族的生物信息学分析[J]. 基因组学与应用生物学 2019 (7) 3140-3154.
- [9]肖容, 杜夏夏, 李传峰, et al. NDRV σ C 蛋白互作的宿主因子筛选及生物信息学分析[J]. 中国动物传染病学报 2019 (4).