

基于多重计算设计策略提高奇异变形杆菌脂肪酶的热稳定性

张碧飞, 吕成, 张萌, 许菲

江南大学 生物工程学院 工业生物技术教育部重点实验室, 江苏 无锡 214122

张碧飞, 吕成, 张萌, 许菲. 基于多重计算设计策略提高奇异变形杆菌脂肪酶的热稳定性. 生物工程学报, 2022, 38(4): 1537-1553.

ZHANG BF, LÜ C, ZHANG M, XU F. Improving the thermal stability of *Proteus mirabilis* lipase based on multiple computational design strategies. Chin J Biotech, 2022, 38(4): 1537-1553.

摘要: 来自奇异变形杆菌的脂肪酶 (*Proteus mirabilis* lipase, PML) 具有较好的有机溶剂耐受性, 在生产生物柴油等领域有广泛的应用潜力。然而, 其热稳定性仍有待提高以更好地适应工业生产。近年来, 基于计算机辅助的多种计算设计策略的发展为酶的热稳定性改造提供了更多的思路。文中首先选择了 3 种基于互补算法的软件 ABACUS、PROSS 以及 FoldX 为正筛对 PML 进行计算设计, 选择其两两交集, 利用 PSSM 序列保守性分析和 G_{REMLIN} 共进化分析两种负筛进一步缩小突变文库, 得到 18 个单点突变设计。经实验验证, 有 7 个 T_m 提高的突变体, 其中 K208G 和 G206D 的 ΔT_m 最高, 分别为 3.75 °C 和 3.21 °C。选择其中 5 个酶活高于野生型的突变体利用贪婪累积策略进行组合, 最终五点组合突变体 M10 的 T_m 提高了 10.63 °C, 相对酶活为野生型的 140%。组合过程中 K208G 和 G206D 展现出了一定的上位性, 为 M10 热稳定性的提高作出了主要贡献。分子动力学模拟表明在突变位点及附近区域有新的作用力产生, G206D/K208G 附近作用力的重排可能稳定了对 PML 稳定性有关键作用的 Ca^{2+} 结合位点。本研究为天然酶的热稳定性改造提供了高效且兼顾用户友好性的计算设计方案, 并为 PML 的改造及其工业应用的拓展奠定了基础。

关键词: 奇异变形杆菌; 脂肪酶; 热稳定性; 蛋白质计算设计; 分子动力学模拟

Received: September 7, 2021; **Accepted:** February 22, 2022; **Published online:** February 24, 2022

Supported by: National Natural Science Foundation of China (31800671); the Fundamental Research Funds for the Central Universities, China (JUSRP121014)

Corresponding authors: ZHANG Meng. E-mail: mzhang@jiangnan.edu.cn
XU Fei. E-mail: feixu@jiangnan.edu.cn

基金项目: 国家自然科学基金 (31800671); 中央高校基本科研业务费专项资金 (JUSRP121014)

Improving the thermal stability of *Proteus mirabilis* lipase based on multiple computational design strategies

ZHANG Bifei, LÜ Cheng, ZHANG Meng, XU Fei

The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China

Abstract: *Proteus mirabilis* lipase (PML) features tolerance to organic solvents and great potential for biodiesel synthesis. However, the thermal stability of the enzyme needs to be improved before it can be used industrially. Various computational design strategies are emerging methods for the modification of enzyme thermal stability. In this paper, the complementary algorithm-based ABACUS, PROSS, and FoldX were employed for positive selection of PML mutations, and their pairwise intersections were further subjected to negative selection by PSSM and G_{REMLIN} to narrow the mutation library. Thereby, 18 potential single-point mutants were screened out. According to experimental verification, 7 mutants had melting temperature (T_m) improved, and the ΔT_m of K208G and G206D was the highest, which was 3.75 °C and 3.21 °C, respectively. Five mutants with activity higher than the wild type (WT) were selected for combination by greedy accumulation. Finally, the T_m of the five-point combination mutant M10 increased by 10.63 °C, and the relative activity was 140% that of the WT. K208G and G206D exhibited certain epistasis during the combination, which made a major contribution to the improvement of the thermal stability of M10. Molecular dynamics simulation indicated that new forces were generated at and around the mutation sites, and the rearrangement of forces near G206D/K208G might stabilize the Ca^{2+} binding site which played a key role in the stabilization of PML. This study provides an efficient and user-friendly computational design scheme for the thermal stability modification of natural enzymes and lays a foundation for the modification of PML and the expansion of its industrial applications.

Keywords: *Proteus mirabilis*; lipase; thermal stability; computational protein design; molecular dynamics simulation

脂肪酶 (三酰基甘油酯水解酶, EC 3.1.1.3) 可以介导甘油三酯在油水界面上水解成单甘油酯、脂肪酸、双甘油酯和甘油^[1], 也能催化非水环境中的酯交换、酯化、醇解、酸解和氨解^[2]。由于其独特的催化性能, 脂肪酶被广泛应用于食品工业、烘焙、洗涤剂 and 生物燃料等多种行业^[3-4]。2020 年, 微生物脂肪酶市场规模已达到 4.4 亿美元, 预计到 2027 年将达到 6.5 亿美元^[5]。近年, 来自多种变形杆菌的脂肪

酶因具有较高的酶活、耐受有机溶剂以及可在大肠杆菌中可溶性表达的特点引起了人们的关注^[6-11]。其中, 来自奇异变形杆菌的脂肪酶 (*Proteus mirabilis* lipase, PML)^[12]属于 I.1 家族脂肪酶, 由 287 个氨基酸组成, 其晶体结构显示: PML 是带有盖子结构域的 α/β 水解酶, S79、H254 和 D232 组成其催化三联体, E80 的主链酰胺以及两个水分子组成其氧阴离子洞, 此外, PML 的 α_8 末端附近存在 Ca^{2+} 结合位

点, 该 Ca^{2+} 的结合很可能是正确定位 H254 以激活亲核残基 S79 所必需的^[12], 对脂肪酶的稳定性和活性至关重要^[13-14]。近年, PML 在生物柴油的生产等方面展示出了极大的潜力^[15-17], 然而, 其热稳定性并不突出, 这限制了其在工业上的进一步发展。

提高酶的热稳定性一直是蛋白质工程中研究的热点之一。传统的蛋白质工程如定向进化往往以低筛选效率为代价得到符合工业需求的突变体。近年, 得益于结构生物学及生物信息学的发展, 基于蛋白质结构和序列信息来提高蛋白质热稳定性的计算设计方法得到了广泛应用。在蛋白质结构信息的指导下, 基于不同打分策略的能量函数对蛋白能量进行计算评分可以得到自由能更小、更有利于蛋白质稳定的突变设计, 如利用经验力场建立打分函数的 FoldX^[18-19], 应用统计能量函数进行计算设计的 ABACUS^[20-22], 基于物理能量函数的 Rosetta ddg^[23-24]等; 以及在此基础上开发的对蛋白质空腔进行填充设计的 Rosetta-VIP^[25], 它通过将蛋白空腔周围的氨基酸突变成侧链基团更大或非极性的氨基酸来提高酶包装的紧密性, 进而提高酶的稳定性。此外, 基于多序列比对 (multiple sequence alignment, MSA) 的序列分析可以为蛋白质的计算设计提供有益信息, 如位置特异性评分矩阵 (position specific scoring matrix, PSSM)^[26-27]可以筛选出同一位点上自然界中出现概率更大、更可能稳定的氨基酸种类; 共进化分析 (Generative REgularized ModelS of proteINs, GREMLIN)^[28-29]可以提供蛋白质序列中与功能相关的保守氨基酸对; 通过祖先序列重建 (ancestral sequence reconstruction, ASR)^[30]的策略也有望得到热稳定性较高的突变体设计。

然而, 单独的某一种计算策略可能有一定的局限性。结合多种计算方法, 可以减少单一

算法中的采样偏差, 从而降低假阳性结果的概率^[31]。其中, Goldenzweig 等将序列保守性分析 PSSM 与蛋白质构象自由能计算 Rosetta ddg 相结合, 使人源乙酰胆碱酯酶 (hAChE) 的热熔温度 (melting temperature, T_m) 提高了 20 °C^[32]; Wijma 等把分子动力学模拟及 FoldX、Rosetta ddg 能量分析及二硫键设计相结合, 将环氧化物水解酶 (LEH) 的 T_m 提高了 35 °C^[33]。此外, 吴边课题组^[34]在多种计算策略的基础上, 进一步提出 GRAPE 策略, 对有益突变进行聚类分析和贪婪累积, 实现了高效的组合路径。

本研究以 PML 作为模式蛋白, 利用 PML 的三维结构和序列信息, 通过 ABACUS、PROSS 以及 FoldX 三种正向筛选以及 GREMLIN 和 PSSM 两种负向筛选, 得到潜在的热稳定性提高的单点突变。通过贪婪累积策略进行组合, 以期提高 PML 的热稳定性, 进一步拓展其在工业上的应用。

1 材料与方法

1.1 材料

1.1.1 菌株与质粒

大肠杆菌 (*Escherichia coli*) JM109、*E. coli* BL21(DE3) 由本实验室保存; 表达载体 pET-28a(+) 购自 Novagen 公司; 奇异变形杆菌 (*Proteus mirabilis*) 脂肪酶基因由苏州金唯智生物科技有限公司合成。

1.1.2 主要试剂

Primer StarTM HS DNA 聚合酶、DL-10 000 DNA marker 及 Premixed Protein marker (Low) 均购于 TaKaRa 公司; 琼脂糖凝胶 DNA 回收试剂盒和质粒小提试剂盒购于 Axygen 公司; BCA 蛋白浓度测定试剂盒 (增强型) 和聚丙烯酰胺凝胶电泳 (SDS-PAGE) 试剂盒购于碧云天公司; 卡那霉素 (kanamycin, Kan)、异丙

基- β -D-硫代半乳糖苷 (IPTG) 购于生工生物工程 (上海) 股份有限公司; HisTrap™ HP 1 mL Ni^{2+} 柱以及 HiTrap™ 5 mL Desalting 柱购于美国 GE 医疗集团; PCR 引物均由上海亦欣公司合成; ρ -nitrophenyl laurate (ρ NPL) 和 ρ -nitrophenol (ρ NP) 购于 Sigma-Aldrich 公司; 其余试剂无特殊说明, 均为市售分析纯。

1.2 方法

1.2.1 奇异变形杆菌脂肪酶的计算设计

如图 1 所示, 本研究基于奇异变形杆菌脂肪酶 PML 的三维结构和序列信息, 将 ABACUS、PROSS 以及 FoldX 三种策略作为正筛, 将 G_{REMLIN} 和 PSSM 作为负筛, 通过多重计算设计筛选能提高热稳定性的突变体, 针对阳性突变利用贪婪累积算法进行组合。

首先, 分别利用 ABACUS 2 在线服务器 (<http://biocomp.ustc.edu.cn/servers/abacus-design.php>) 和 PROSS 在线服务器 (<https://pross.weizmann.ac.il/step/pross-terms/>) 对野生型 PML (WT) 进行计算设计; 利用 FoldX 中提供的 PositionScan, 对 PML 的 287 个残基进行虚拟饱和突变得到所有突变的 DDG ($\text{DDG} = \text{DG}_{\text{Mut}} - \text{DG}_{\text{WT}}$), 将 $\text{DDG} < -0.45 \text{ kcal/mol}$ 的突变保留。选择在上述两种及两种以上的计算

中同时出现的设计作为本轮的正向筛选结果。

其次, 利用 PML 序列信息对上一轮的单点突变进行负向筛选。通过在 G_{REMLIN} 服务器 (<http://openseq.org/submit.php>) 上提交 PML-WT 的序列, 可以得到其中共进化的残基对其归一性打分, 剔除 G_{REMLIN} 中得分 > 0.7 的各保守残基对中相关位点的突变设计; PSSM 得分以 PROSS 的比对结果为准, 保留在 PSSM 中得分 > 0 的突变设计。

对计算设计得到的单点突变, 测定其稳定性和酶活, 将稳定性提高的单点通过贪婪累积策略进行组合。以单点突变中热稳定性最好的突变作为下一轮突变的出发点, 分别叠加其他阳性突变得到两点组合突变体, 再以该轮突变中热稳定性最好的两点组合突变体为出发点, 分别叠加其他单点阳性突变, 得到三点组合突变, 以此类推, 直到组合突变的热稳定性降低或所有阳性突变的叠加均已完成为止。

1.2.2 定点突变

以含有奇异变形杆菌脂肪酶基因的 pET28a(+)-PML 质粒作为模板, 根据突变位点设计引物进行全质粒 PCR 扩增, 琼脂糖凝胶电泳验证后回收 PCR 产物, 利用热激转化法将其

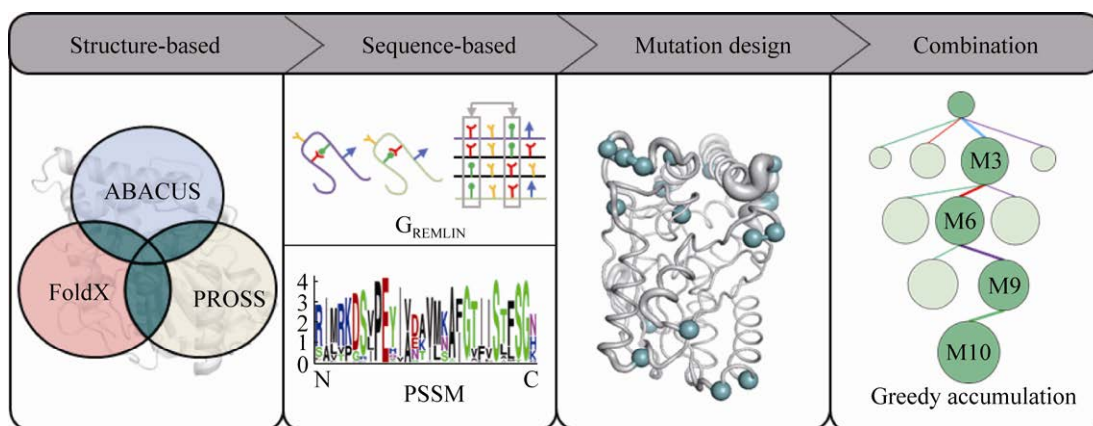


图 1 计算设计策略示意图

Figure 1 Schematic diagram of computational design.

转化至 *E. coli* BL21(DE3) 感受态细胞中。筛选得到的阳性转化子送至无锡天霖生物科技有限公司进行 DNA 测序验证, 获得目标菌株。本研究所使用的引物如表 1 所示。

表 1 本研究使用的引物
Table 1 Primers used in this study

Primer name	Primer sequence (5'→3')	Size (bp)
N17V-F	CTGGCGGGCTTTGTGGAAATTGTG	24
N17V-R	CACAAAGCCCGCCAGGCCATGCAC	24
E18R-F	CGGGCTTTAACCGTATTGTGGGCT	24
E18R-R	CGGTTAAAGCCCGCCAGGCCATGC	24
Q38R-F	AAGATGGCCATCGTGTGTTTACCG	24
Q38R-R	ATGGCCATCTTGGCGCAGCGCATC	24
Q66A-F	AGACCCTGCTGGCAGAAACCCAAG	24
Q66A-R	AGCAGGGTCTGCACAACTGCCAC	24
V132I-F	TGGAAAAAATTCTGAACGCGTTTG	24
V132I-R	TCCACAATATATTCCGGAATGC	22
G148D-F	CGGCCATAGAGACGATCCGCAAG	23
G148D-R	CTCTATGGCCGCTAAAGGTGC	21
D152N-F	CCGCAACATGCGATTGCGGCG	21
D152N-R	CGCATGTTGCGGATCGCCTC	20
A175G-F	CAAATATCCGCAAGGCCTGCCG	22
A175G-R	TTGCGGATATTGTTGTAAATTC	24
G206N-F	GGCCTGATTGCGAACGAAAAAGGC	24
G206N-R	TCGCAATCAGGCCTTGAAT	19
G206D-F	GGCCTGATTGCGACGAAAAAGGC	24
G206D-R	CCGCAATCAGGCCTTGAATAT	21
K208G-F	GCGGGCGAAGGCGGCAACCTGCTG	24
K208G-R	GCCCCGAATCAGGCCTTGAATA	22
G209L-F	ATTGCGGGCGAAAACTGAACC	22
G209L-R	TCGCCCGAATCAGGCCTTGA	21
A217R-F	GATCCGACCCATCGTGCATGCG	23
A217R-R	ATGGGTCGGATCCAGCAGGTTG	22
N223A-F	GCGCGTGCTGGCCACCTTTTTTAC	24
N223A-R	CCAGCACGCGCATCGCCGCATG	22
A251P-F	AGATGATTATCCGCAAGATCATA	23
A251P-R	ATAATCATCTTTAATCAGTTTGC	23
I255L-F	TCATCTTTAATCAGTTTGCCAG	24
I255L-R	ATAATCATCTTTAATCAGTTTGC	24
Q286G-F	TGGCGAGCAAAGGTCTGTAAGCG	23
Q286G-R	TGCTCGCCAGATATTCGCATG	22

1.2.3 野生型及突变体蛋白的表达纯化

将保存于甘油管的 *E. coli* BL21(DE3)/pET28a(+)-PML 野生型及突变型菌株涂布于含 25 μg/mL Kan 的 LB 固体平板, 于 37 ℃ 恒温培养箱中培养 12–16 h。挑取单菌落于 20 mL 含 25 μg/mL Kan 的 LB 液体培养基中, 37 ℃、200 r/min 振荡培养 12 h。以 1% 接种量接入 100 mL 含 25 μg/mL Kan 的 TB 液体培养基, 37 ℃、200 r/min 振荡培养至 $OD_{600}=0.6$ 时, 加入终浓度为 0.5 mmol/L 的 IPTG 诱导, 16 ℃、200 r/min 继续培养 16 h 后, 8 000 r/min 离心 5 min 收集发酵菌体。

用 25 mL 上样缓冲液 (50 mmol/L Tris-HCl, 300 mmol/L NaCl, 10 mmol/L 咪唑, pH 8.0) 重悬菌体, 在冰浴条件下超声破碎菌体 (超声 2 s, 间歇 3 s, 20 min) 后, 10 000 r/min 离心 20 min 收集上清, 并用 0.45 μm 滤膜过滤, 利用镍柱亲和层析纯化目的蛋白。纯化条件: 首先用上样缓冲液平衡 HisTrap™ HP 1 mL Ni²⁺ 亲和层析柱 15 个柱体积, 将过滤后的上清液以 1 mL/min 的流速进样, 等待进样结束后, 使用 5 个柱体积的上样缓冲液平衡; 用 10% 的洗脱缓冲液 (50 mmol/L Tris-HCl, 300 mmol/L NaCl, 250 mmol/L 咪唑, pH 8.0) 进行梯度洗脱除去杂蛋白, 再用 30% 的洗脱缓冲液梯度洗脱 10 个柱体积, 对目的蛋白进行收集。对经过亲和层析的目的蛋白用 HiTrip™ 5 mL Desalting 凝胶柱进行脱盐处理, 并更换至 20 mmol/L Tris-HCl 缓冲液 (pH 8.0) 中, 测量其浓度并用 SDS-PAGE 对其纯度进行验证。于 0 ℃ 条件下储存酶液备用。

1.2.4 野生型及突变体蛋白的酶活测定

奇异变形杆菌脂肪酶活力分析是通过水解 ρNPL 所释放的 ρNP 的量来测定。总体积 2 mL 的底物体系包括: 1 880 μL 反应缓冲液 (20 mmol/L

Tris-HCl, 5 mmol/L CaCl₂, pH 8.0); 20 μ L Triton X-100 及 100 μ L 浓度为 50 mmol/L 的溶于乙腈的 ρ NPL 溶液。充分混匀后以每孔 190 μ L 分装至 96 孔板中, 35 $^{\circ}$ C 下孵育后加入适当稀释的酶液 10 μ L, 用酶标仪测定 2 min 内 410 nm 处的吸光值变化。以相同条件下缓冲液的测定作为空白对照, 每次设 3 组平行试验。

上述条件下, 将每分钟水解 ρ NPL 生成 1 μ mol ρ NP 所需的酶量定义为 1 个酶活单位 (U)。总酶活力除以酶浓度得到比活, 单位为 U/mg。

1.2.5 野生型及突变体的热稳定性测定

利用圆二色光谱 (CD) 测定蛋白质二级结构及 T_m 。将纯化得到的酶液取适量稀释至 0.2 mg/mL, 取 200 μ L 加入到光程为 0.1 cm 的石英比色皿中, 用圆二色光谱仪测定其在 4 $^{\circ}$ C 条件下, 190–260 nm 的摩尔椭圆率。根据全波长扫描来确定蛋白折叠情况。选择最低吸收值处的波长 220 nm, 将温度从 4 $^{\circ}$ C 升温至 80 $^{\circ}$ C (升温速率 1 $^{\circ}$ C/min), 测定其摩尔椭圆率变化得到其溶解曲线。将得到的曲线通过波兹曼拟合, 取蛋白解折叠至一半的温度作为 T_m 。

将纯化后的酶液分别在 30–70 $^{\circ}$ C (间隔 5 $^{\circ}$ C) 下孵育 10 min, 放置于冰上冷却 10 min 后, 采用上述 1.2.4 方法测定不同孵育温度下的酶活力。将未孵育条件下的酶活力设为 100%, 计算各温度条件下的相对比活。相对剩余酶活力降为 50% 时所对应的温度, 即半失活温度 (T_{50})。

将纯化后的酶液在 50 $^{\circ}$ C 下孵育 0–90 min (间隔 10 min), 放置于冰上冷却 10 min 后, 采用上述 1.2.4 方法测定不同孵育时间下的酶活力。将未孵育条件下的酶活力设为 100%, 计算各孵育时间下的相对比活。相对剩余酶活力降为 50% 时所对应的时间, 即为半失活时间 ($t_{1/2}$)。

1.2.6 突变体的分子动力学模拟

基于 PML 的晶体结构 (PDB ID: 4GW3), 在 Pymol (<http://www.pymol.org>) 中进行氨基酸替换得到突变体的三维结构。以该结构作为初始结构, 基于 AMBER99SB-ILDN 力场利用 Gromacs 5.0.4 进行分子动力学模拟^[35]。将蛋白分子放置在正方体的水盒子中, 用 TIP3P 模型^[36]的水分子填充溶剂及离子使系统呈中性后, 用 steepest descent 方法^[37]进行能量最小化。利用 velocity rescaling 算法保持体系温度恒定为 25 $^{\circ}$ C, 并且使用弱耦合方法控制压力为 1 bar^[38], 50 ns NPT 的模型优化后进行 150 ns NVT 的平衡模拟, 其中模拟时间步长设置为 2 fs。

氢键定义: 当氢供体氮原子与氢受体氧原子之间的距离小于 3.5 \AA , 且 $\angle\text{HNO}$ 小于 30 $^{\circ}$ 时氢键存在^[39–40]。盐桥定义: 由带正电残基羧酸酯基团的氧原子与带负电残基酰胺基团的氮原子之间的距离确定, 当两原子之间的距离小于 4.0 \AA 时盐桥存在^[40–41]。分子动力学模拟的最后 50 ns 中, 成键的时间占比为对应成键的占有率。

2 结果与分析

2.1 奇异变形杆菌脂肪酶的计算设计

利用两轮筛选进行奇异变形杆菌脂肪酶的计算设计。第一轮筛选中结合了 ABACUS、PROSS 和 FoldX 三种设计策略, 选取三者两两相交的部分作为正向筛选的结果, 共计得到了 67 个突变设计。随后, 利用 GREMLIN 和 PSSM 对以上设计进行负向筛选, 仅保留 GREMLIN 归一性打分 < 0.7 的位点及 PSSM > 0 的设计, 最终获得了 18 个单点突变设计 (表 2)。18 个设计中, Q69G、A217R 等 5 个设计来自 3 种正向筛选器的交集, N17V 等 6 种设计来自 ABACUS 和 FoldX 的交集, E18R 等 5 种设计来自 PROSS

表 2 突变设计的筛选参数

Table 2 Parameters for mutation designs based on the design by ABACUS and PROSS, the DDG calculation of FoldX, PSSM scores from conservation analysis, and the co-evolution data from G_{REMLIN}

Mutant	ABACUS	PROSS	FoldX (kcal/mol)	Intersection	PSSM	G _{REMLIN}
N17V	√		-0.71	ABACUS+FoldX	3	/
E18R		√	-0.50	PROSS+FoldX	6	/
Q38R	√		-0.52	ABACUS+FoldX	4	/
Q66A	√	√	-0.13	ABACUS+PROSS	5	/
Q69G	√	√	-1.86	ABACUS+PROSS+FoldX	7	/
V132I	√		-0.61	ABACUS+FoldX	3	/
G148D	√		-1.38	ABACUS+FoldX	1	/
D152N		√	-0.64	PROSS+FoldX	5	/
A175G		√	-1.81	PROSS+FoldX	7	/
G206N	√		-0.48	ABACUS+FoldX	3	/
G206D		√	-0.48	PROSS+FoldX	6	/
K208G	√	√	-0.26	ABACUS+PROSS	7	/
G209L		√	-1.03	PROSS+FoldX	3	/
A217R	√	√	-0.46	ABACUS+PROSS+FoldX	2	/
N223A	√	√	-1.05	ABACUS+PROSS+FoldX	2	/
A251P	√	√	-1.35	ABACUS+PROSS+FoldX	8	/
I255L	√	√	-0.68	ABACUS+PROSS+FoldX	5	/
Q286G	√		-2.01	ABACUS+FoldX	7	/

Note: “/” indicates that there was no co-evolution or G_{REMLIN} score was less than 0.7 at the site where the design was located.

和FoldX的交集, Q66A和K208G来自ABACUS和PROSS的交集。

2.2 野生型脂肪酶及突变体在大肠杆菌中的表达纯化

分别收集IPTG诱导后的PML-WT的全细胞、上清以及纯化后获得的酶液进行SDS-PAGE鉴定。SDS-PAGE中目的蛋白条带大小约为33 kDa(图2A), 与理论值相符。测定野生型奇异变形杆菌脂肪酶的比活为(33.48±0.68) U/mg。用表1中的引物进行全质粒PCR突变, 将突变的质粒测序验证后, 进行蛋白的表达纯化。所有的单点突变均能在大肠杆菌中可溶性表达, 且条带对应分子量大小正确, 纯度均达到了电泳纯(图2B)。

2.3 单点突变体的酶学性质

为表征WT与突变体在二级结构上的差

异, 用圆二色谱仪对WT及单点突变体在190–260 nm进行全波长扫描(图2C)。根据全波长扫描结果可知, 所有突变蛋白的扫描结果与野生型变化趋势相同, 在210 nm和220 nm附近均有典型负峰, 说明突变体全都折叠形成了与WT一致的二级结构。

为了表征WT和突变体热稳定性的差异, 用圆二色谱仪测定其熔融曲线(图2D), 各突变体的 T_m 如表3所示。根据 T_m 测定结果可知, T_m 值上升的突变体有7个, ΔT_m 为0.78–3.75 °C不等。其中提高3 °C以上的单点突变体有K208G(3.75 °C)和G206D(3.21 °C)。选择 T_m 高于野生型的单点突变测定其 T_{50} 和 $t_{1/2}$, 其变化趋势同 T_m 基本保持一致。7个单点突变体的 T_{50} 较WT增加了0.66–3.83 °C不等, $t_{1/2}$ 均略高于WT。利用酶标仪进一步测定了单

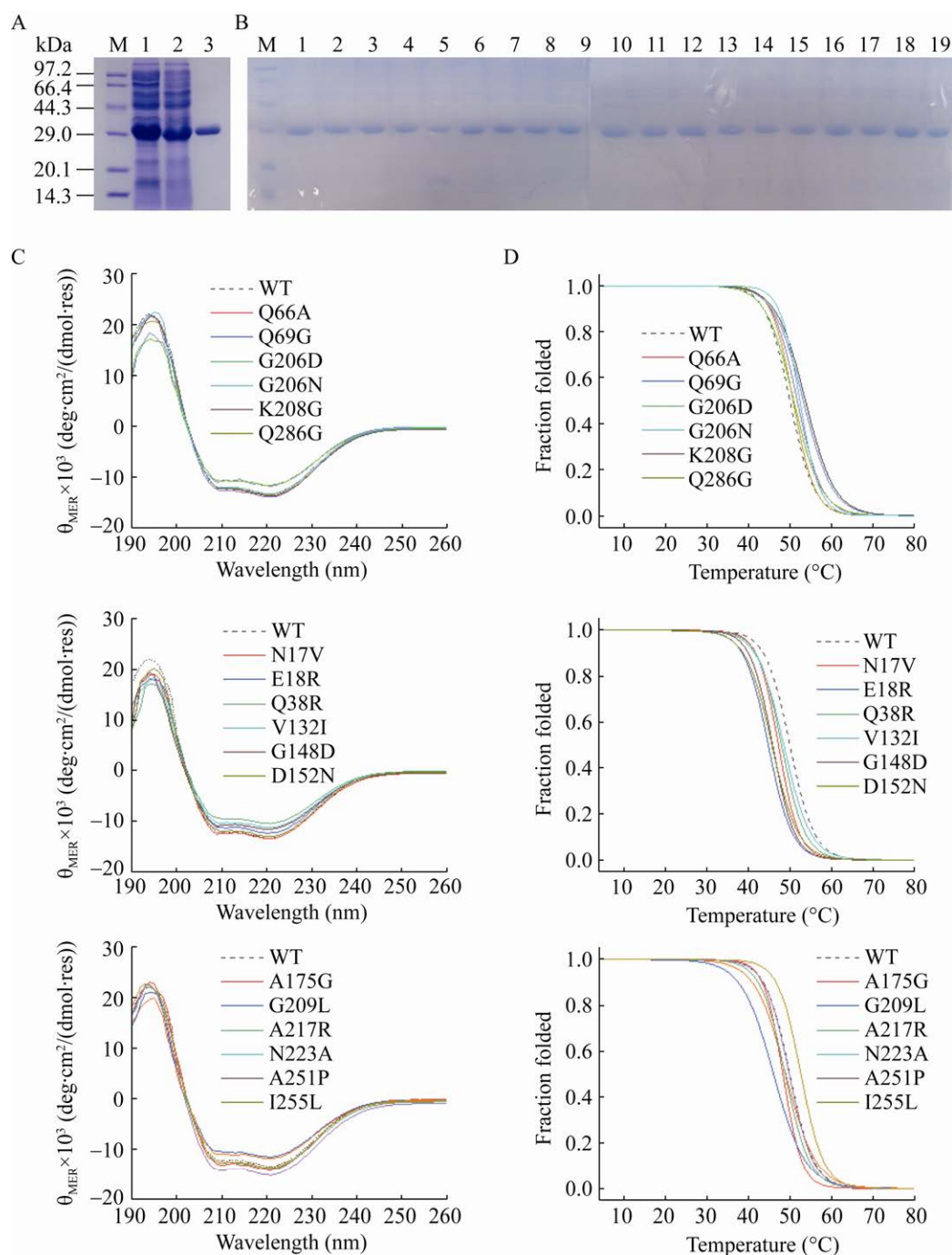


图 2 野生型 PML 及单点突变体的表达纯化及其圆二色谱分析

Figure 2 Expression, purification, and circular dichromatography of wild-type PML and single-point mutants. (A) M: premixed protein marker; 1–2: total and soluble protein of *E. coli* BL21(DE3) cells harboring pET28a(+)-PML, respectively; 3: purified wild type PML. (B) M: premixed protein marker; 1: purified wild type PML; 2–19: purified protein of mutant N17V, E18R, Q38R, Q66A, Q69G, V132I, G148D, D152N, A175G, G206N, G206D, K208G, G209L, A217R, N223A, A251P, I255L, Q286G. (C) Full-wavelength scans. res: residue number. (D) Melting curves.

表3 单点突变体的 T_m 、 ΔT_m 、 T_{50} 、 $t_{1/2}$ 和相对酶活Table 3 T_m , ΔT_m , T_{50} , $t_{1/2}$ and relative activity of single-point mutants

Mutant	T_m (°C)	ΔT_m (°C)	T_{50} (°C)	$t_{1/2}$ (min)	Relative activity (%)
WT	49.96	0.00	56.70	59.30	100.00
N17V	47.20	-2.76	—	—	66.92
E18R	44.76	-5.20	—	—	80.93
Q38R	47.97	-1.99	—	—	128.87
Q66A	51.74	1.78	58.63	61.36	118.08
Q69G	50.74	0.78	57.36	59.75	110.48
V132I	48.50	-1.46	—	—	112.94
G148D	45.79	-4.17	—	—	127.56
D152N	45.64	-4.32	—	—	71.93
A175G	48.96	-1.00	—	—	133.33
G206N	50.87	0.91	57.92	59.25	135.56
G206D	53.17	3.21	59.23	63.89	137.28
K208G	53.71	3.75	60.53	65.68	130.18
G209L	48.29	-1.67	—	—	112.31
A217R	46.05	-3.91	—	—	83.55
N223A	48.82	-1.14	—	—	117.82
A251P	49.83	-0.13	—	—	76.17
I255L	52.63	2.67	58.06	62.85	75.03
Q286G	52.36	2.40	57.87	60.52	124.54

Note: “—” indicates that the data has not been determined by experiment.

点突变体的相对比活。18 个单点突变中, 有 12 个单点突变的比活高于 WT。A175G、G206N、G206D 和 K208G 四个突变体的比活超过了 WT 的 130%, 其中 G206D 比活最高, 达到了 WT 的 137%。

根据 T_m 及比活的大小, 18 个单点突变体被分成了 4 大类 (图 3A)。第一大类中, T_m 及比活均高于 WT, 包括 K208G、G206D、G206N、Q286G、Q66A 及 Q69G 六个突变体。其中, K208G 的热稳定性最强 ($\Delta T_m=3.75$ °C)、相对比活也达到了 WT 的 130%。206 位相关的两个设计 (G206N 及 G206D) 比活力相似, 均为 WT 的 135%左右, 但 G206D 的 T_m 较 G206N 高出 2 °C 以上。其他 3 个突变体的 T_m 及比活的大小关系均为 Q286G>Q66A>Q69G。K208G 及 G206D 是所有突变体中热稳定性及比活力最高的两个

突变体, 这可能与两个位点在 PML 结构中的特殊位置有关 (图 3B)。两个位点均位于盖子区域附近的 loop 上, 该区域柔性较大, 且靠近 Ca^{2+} 结合位点。两个位点的突变引入可能会对 Ca^{2+} 结合位点的稳定起到积极作用。第二大类中, T_m 低于 WT, 比活高于 WT, 包含 A175G 等 6 个突变体。第三大类中, T_m 高于 WT, 比活低于 WT, 仅包含 I255L 一个突变体, 其比活仅有 WT 的 75%左右, 猜测这可能是由于该位点过于靠近活性中心, 突变可能引起活性中心构象的变化 (图 3B)。第四类中的突变体有 5 个, 其 T_m 和比活均低于 WT。

通过以上酶学性质的表征发现, 在 18 个筛选出的单点突变设计中, 有 7 个设计热稳定性提升。这 7 个设计分别涵盖了来自三者共同交集的 2 个 (Q69G 及 I255L); 来自 ABACUS

和 FoldX 交集集中的 2 个 (G206N 及 Q286G); 来自 PROSS 和 FoldX 交集集中的 1 个 (G206D); 来自 ABACUS 和 PROSS 交集集中的 2 个 (Q66A 及 K208G)。这与其他利用多重计算设计取交集以富集阳性突变的研究^[31,42]一致, 也说明不同设计策略之间具有一定的互补性。吴边等^[34]的研究中, 通过将多种算法取并集的方式, 从 85 个候选设计中得到了 21 个 T_m 提高超过 1.5 °C 的塑料降解酶 (PETase) 突变体, 本研究取交集的方式在保证筛选效率相当的情况下提供了一种工作量更小的手段, 当然也承担了错过其他更好突变体的可能性。此外, 7 个较优突变体的 PSSM 得分除 G206N 外均在 5–7 之间, 说明 PML 的序列保守性分析可以一定程度辅助其热稳定性的计算设计。

2.4 组合突变体的酶学性质

为了得到热稳定性更好的突变体, 同时考察突变位点之间的上位性, 用贪婪累积策略对 5 个热稳定性及酶活均优于 WT 的单点突变 (K208G、G206D、Q286G、Q66A 及 Q69G) 进行组合 (由于 G206D 在热稳定性及酶活上均优于 G206N, 因此选择 G206D 进行组合)。经过 4 轮的贪婪累积后得到了 10 个组合突变体。所有的组合突变体均能在大肠杆菌中可溶性表达, 且纯化后纯度均达到了电泳纯 (图 4A)。全波长扫描结果显示, 所有组合突变体均折叠形成了正确的二级结构 (图 4B)。

组合突变体的熔融曲线如图 4C 所示, 其 T_m 、 T_{50} 、 $t_{1/2}$ 和相对比活结果如表 4 所示。第一轮两点组合突变体中, M3 (K208G+G206D)

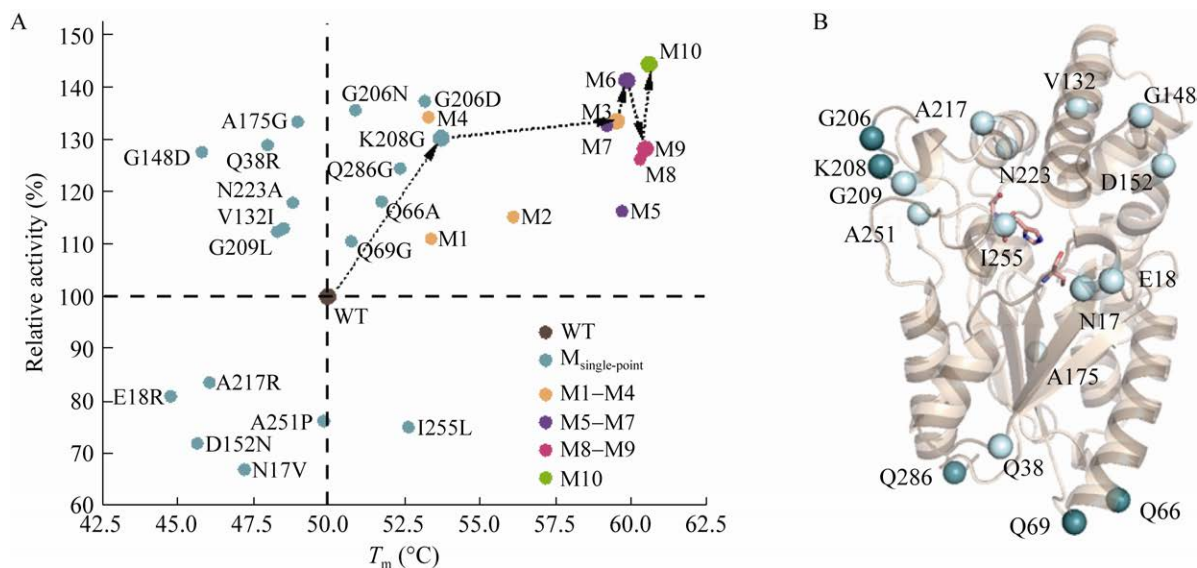


图 3 突变体的酶活、 T_m 及各位点在 PML 晶体结构中的位置

Figure 3 Activity and T_m of mutants and mutation sites in PML crystal structure. (A) The dots of the same color contain mutations from the same parent. The dotted arrow indicates the combined path from WT (brown) to the optimal mutant M10 (green), which contains the best mutant in each round of stacking: K208G (dark cyan), M3 (orange), M6 (purple), M9 (pink). The horizontal and vertical dotted lines represent the enzyme activity and T_m of WT, respectively. (B) Mutational sites shown in wild type PML structure. The mutants with higher T_m and enzyme activity than WT have a darker color. The catalytic triad of the enzyme are shown as stick.

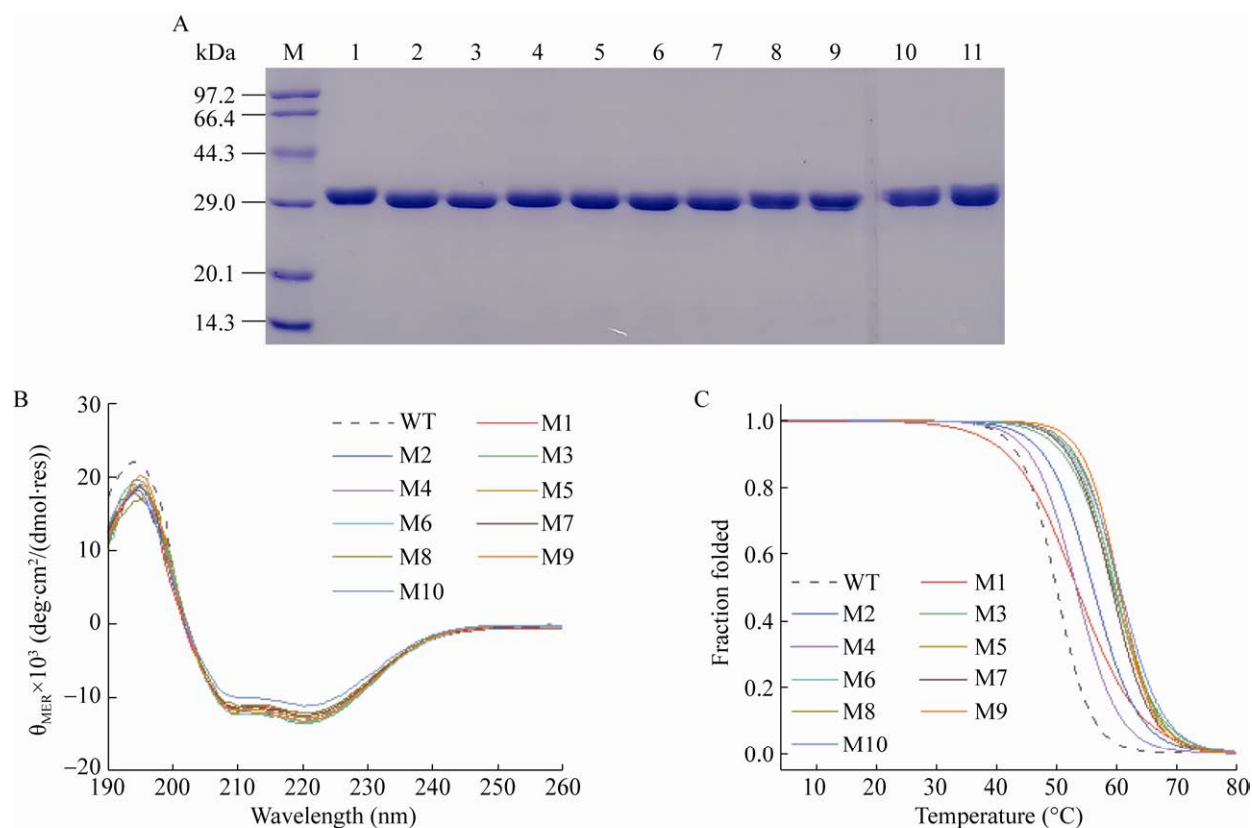


图4 野生型 PML 及组合突变体的表达纯化及其圆二色谱分析

Figure 4 Expression, purification, and circular dichromatography of wild-type PML and its combination mutants. (A) Wild type PML and combined mutants expression in *E. coli* BL21(DE3). M: premixed protein marker; 1: purified wild type PML; 2–11: purified protein of mutant M1, M2, M3, M4, M5, M6, M7, M8, M9, M10. (B) Full-wavelength scans. res: residue number. (C) Melting curves.

表4 组合突变体的 T_m 、 ΔT_m 、 T_{50} 、 $t_{1/2}$ 和相对酶活

Table 4 T_m , ΔT_m , T_{50} , $t_{1/2}$, and relative activity of combination mutants

Mutant	T_m (°C)	ΔT_m (°C)	T_{50} (°C)	$t_{1/2}$ (min)	Relative activity (%)
WT	49.96	0.00	56.70	59.30	100.00
M1 (K208G+Q66A)	53.39	3.43	60.35	66.59	110.96
M2 (K208G+Q69G)	56.10	6.14	59.34	79.46	115.15
M3 (K208G+G206D)	59.51	9.55	64.56	86.97	133.45
M4 (K208G+Q286G)	53.29	3.33	59.52	64.58	134.20
M5 (M3+Q66A)	59.70	9.74	63.96	90.52	116.20
M6 (M3+Q69G)	59.86	9.90	65.36	91.53	141.21
M7 (M3+Q286G)	59.20	9.24	64.58	88.65	132.50
M8 (M6+Q66A)	60.29	10.33	66.32	90.36	126.14
M9 (M6+Q286G)	60.46	10.50	66.85	95.66	128.17
M10 (M9+Q66A)	60.59	10.63	67.73	98.11	144.36

Note: bold numbers indicate the maximum value of ΔT_m in each round of greed accumulation, and the mutant is the parent of next round of accumulation.

和 M2 (K208G+Q69G) 的 ΔT_m 分别为 9.55 °C 和 6.14 °C, 均高于其单点突变的 ΔT_m 之和 (6.96 °C 及 4.53 °C), 说明 K208G 与 G206D、K208G 与 Q69G 之间存在一定的协同作用。M1 (K208G+Q66A) 及 M4 (K208G+Q286G) 的 ΔT_m 较 K208G 分别下降 0.32 °C 和 0.42 °C, 说明 Q66A 及 Q286G 与 K208G 的叠加均出现了负的上位性。四个突变体的 T_{50} 较 WT 高出 2.64–7.86 °C, 其中 M3 最高, 比 K208G 高出 4.03 °C。 $t_{1/2}$ 提高的组合突变体有 3 个, 其中 M3 较 K208G 提高了 21.29 min。四个突变体的比活均高于 WT, 其中, M3 及 M4 的比活与 K208G 接近, 均达到了 WT 的 130% 以上。

选择 M3 作为起点进行了第二轮贪婪累积。三个组合突变体的 T_m 均在 59–60 °C 之间。M6 (M3+Q69G) 及 M5 (M3+Q66A) 的 ΔT_m 分别为 9.90 °C 和 9.74 °C, 较 M3 分别提高了 0.35 °C 和 0.16 °C, M7 (M3+Q286G) 则比 M3 低 0.31 °C, 说明 Q69G、Q66A 两个突变与 M3 之间的协同作用并不明显, Q286G 与 M3 之间可能存在一定负上位性。M6 及 M7 的 T_{50} 均高于 M3, 其中 M6 比 M3 高 0.80 °C。M5 的 T_{50} 较 M3 有所下降。三个组合突变体的 $t_{1/2}$ 均高于 M3, 其中 M6 最高, 比 M3 提高了 4.56 min, 此外, M6 的比活在 M3 基础上进一步上升, 达到了 WT 比活的 140%, M7 的比活与 M3 相似, M5 则低于 M3。

选择 M6 作为第 3 轮贪婪累积的起点得到 M8 (M6+Q66A) 及 M9 (M6+Q286G)。两个组合突变体的 ΔT_m 分别为 10.33 °C 和 10.50 °C, 较 M6 提高了 0.63 °C 及 0.43 °C。M8 及 M9 的 T_{50} 较 M6 分别提高了 0.96 °C 和 1.49 °C。M9 的 $t_{1/2}$ 较 M6 提高了 4.13 min, M8 较 M6 下降了 1.17 min。两个突变体的比活相近, 但均低于 M6, 仅为 WT 的 127% 左右。

最后一轮贪婪累积得到了五点组合突变体 M10 (M9+Q66A), 其 T_m 达到了 60.59 °C, 较 WT 提高了 10.63 °C, T_{50} 为 67.73 °C, $t_{1/2}$ 为 98.11 min, 分别比 WT 提高了 11.03 °C 和 38.81 min。此外, 其比活也是所有组合突变体中最高的, 为 WT 的 144.36%。

以 WT 为出发点, 通过单点突变和四轮贪婪累积构建出了 WT-K208G-M3-M6-M9-M10 组合路径, 突变体的热稳定性逐步提高 (ΔT_m : 3.75-9.55-9.90-10.50-10.63 °C), 且酶活均稳定到了 WT 的 130% 左右 (图 3A, 表 4)。其中需要注意的是, ΔT_m 提高幅度最大的一步是 K208G-M3, $\Delta\Delta T_m$ 为 5.80 °C, 而 M3-M10 的 $\Delta\Delta T_m$ 仅为 1.08 °C; K208G、G206D 单点突变和各相关组合突变体酶活均在 130% 左右。从热稳定性与酶活两方面推测, K208G 与 G206D 这两个位置相近的突变体之间的协同作用为 M10 性质的改善提供了主要贡献。另外, 组合路径中, 前两轮组合突变的效果较好, 随后的叠加没有发生较大变化, 这种现象与以往的发现^[34,43]一致, 表现较好的突变体一般在组合的早期出现, 由于上位性效应, 对两个或三个以上的突变体进行组合不一定能得到持续改进酶性质的突变体。

2.5 突变体的分子动力学模拟

为了进一步研究突变对蛋白质结构的影响, 对 WT 以及五点组合突变体 M10 进行了 150 ns 的分子动力学模拟。利用主链原子的均方根偏差 (root mean square deviation, RMSD) 表征系统整体的稳定性 (图 5A)。通常一个稳定的体系其 RMSD 应小于 3.2 Å^[44]。从 RMSD 的变化趋势可以看出, WT 及 M10 在 50 ns 后均基本达到平衡状态。50–150 ns 中, WT 和 M10 的平均 RMSD 分别为 (1.30±0.12) Å 和 (1.61±0.13) Å, 均远小于 3.2 Å, 这意味着 50–150 ns 的模

拟是稳定且具有代表性的, 可以进行后续的分析。

为了表征单个残基波动对蛋白质整体动力学的贡献, 选择模拟的后 50 ns 计算得到均方根涨落 (root mean square fluctuation, RMSF) (图 5B)。除了少数特殊区域 (黑色方框) 外, WT 和 M10 之间的大多数残基波动大小相似。在 3 个黑框区域中, M10 的波动均小于 WT, 这 3 个区域均是盖子附近的柔性区域, 突变的引入, 减小了局部构象的波动性, 更有利于 PML 结构稳定性的提高。

为了进一步从分子层面探析引入突变对酶学性质的影响, 分别在 WT 和 M10 达到平衡后, 针对突变位点附近的作用力进行分析, 结果如图 6 所示, 所涉及的氢键及盐桥占有率如表 5 所示。5 个突变位点可以分为 3 个区域。G206/K208 处于盖子 α_8 附近的 loop 区, 该区域靠近 Ca^{2+} 结合位点, G206D/K208G 两个突变位点的引入, 在残基 205–209 附近区域引入了 6 个新的作用力 (图 6A, 6B), 实现了这一区域作用力的重排。G206D 突变在该区域引入了一个携带负电荷的酸性氨基酸残基, K208G 在该区域消除了一个带正电的碱性氨基酸残基, 这种

电荷的改变可能会一方面对 Ca^{2+} 的结合产生促进作用, 另一方面改变所在 loop 的原有的运动模式, 使 E207 发生了朝向盖子方向的偏移。这种偏移使 E207 与 R220 形成稳定的盐桥, 同时与 A205 主链之间形成氢键。这两个相对稳定的作用力进一步促成了 D206 与 G208 主链之间、A205 与 E207 的侧链之间新的氢键形成。G209 与 V265、N268 形成的氢键紧靠 Ca^{2+} 结合位点, 可能会进一步提升 Ca^{2+} 的稳定性, 从而使热稳定性得到提高。这也解释了实验中发现的 G206D 与 K208G 之间的协同作用。虽然从氢键占有率看来, 这种作用力重排的构象不是持续稳定存在的, 但是考虑到该区域本身柔性较大, 这种较 WT 而言出现的新构象可能为其功能的提升做出了一定贡献。Q66/Q69 位于 α_2 末端, M10 中观察到 Q66A 与 L64 以及 L65 与 T68 之间形成了新的氢键 (图 6C, 6D), 可能进一步增强了该区域 α 螺旋的稳定性; Q286 位于 C 末端, 在 M10 中, Q286G 与 L282、S284 形成了新的氢键 (图 6E, 6F), 增强了 C 末端的刚性。相较于 G206/K208 区域, 后两区域新形成的氢键占有率较低, 仅在 5% 左右, 这与组合路径中观察到 M3–M10 的热稳定性的小幅度上

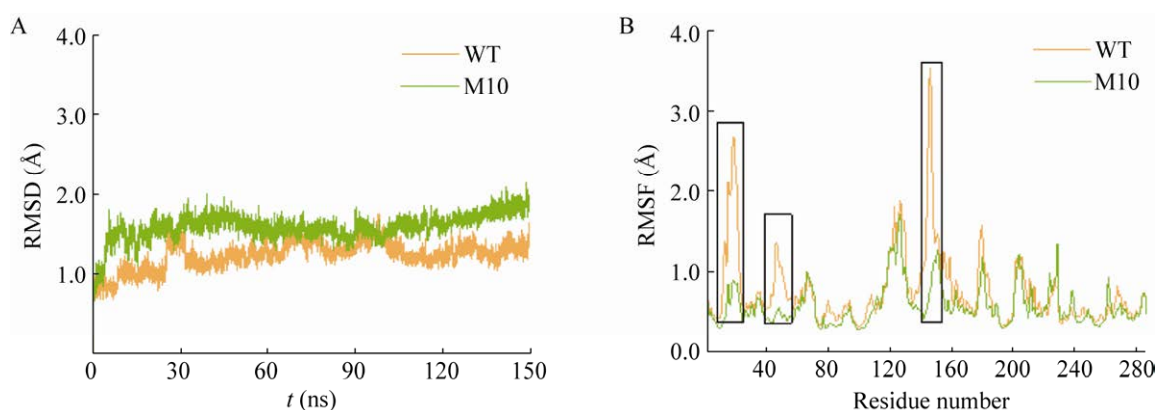


图 5 野生型 PML 及突变体 M10 的分子动力学模拟分析

Figure 5 Molecular dynamics simulation for wild type (brown) and mutant M10 (green). (A) The root mean square deviation (RMSD). (B) The root mean square fluctuation (RMSF).

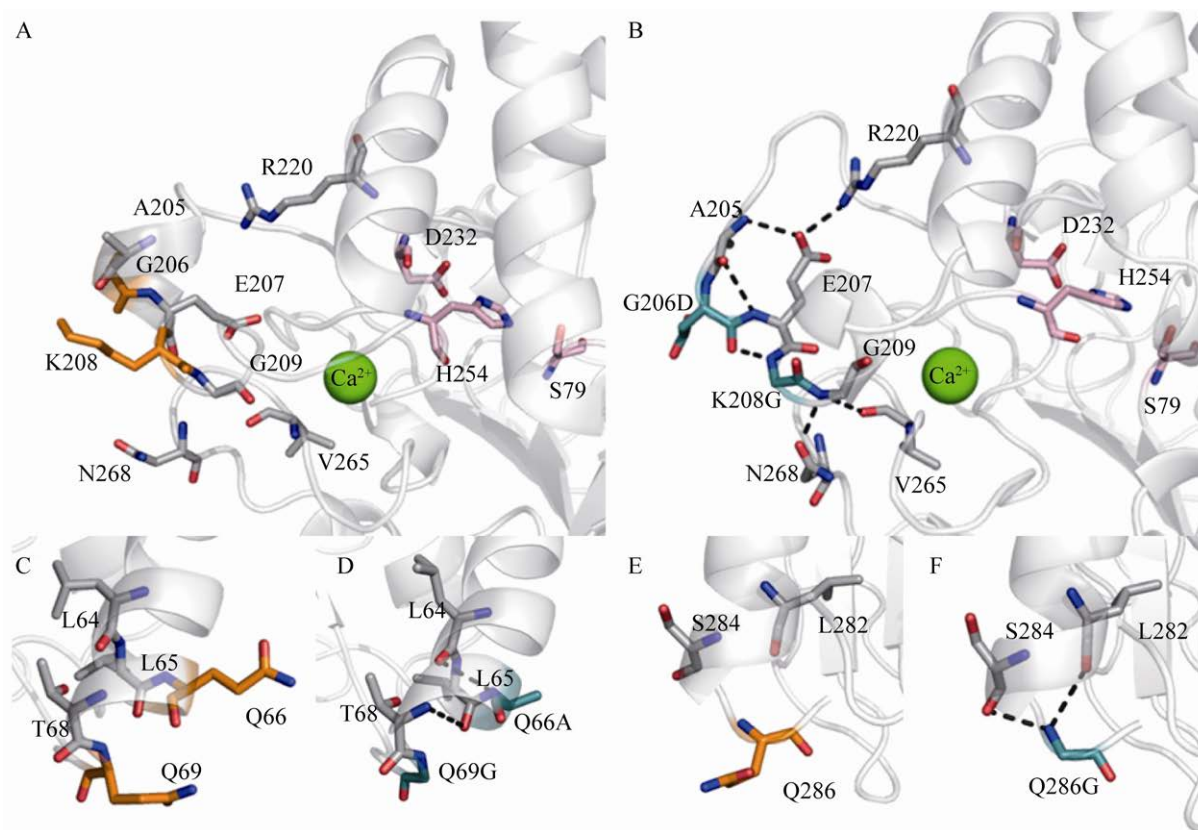


图 6 M10 热稳定性提高的机理分析

Figure 6 Mechanism for improved thermal stability of M10. (A) G206/K208 area of WT. (B) G206D/K208G area of M10. (C) Q66/Q69 area of WT. (D) Q66A/Q69G area of M10. (E) Q286 area of WT. (F) Q286G area of M10. The catalytic triad is represented by a pink stick model, and Ca^{2+} is represented by a green ball. The amino acids before and after the mutation are represented by orange and dark green respectively, the nearby amino acids are represented by silver, and the newly formed hydrogen bonds and salt bridges are represented by black dotted lines.

表 5 突变区域残基间的氢键及盐桥占有率

Table 5 Hydrogen bond and salt bridge occupancy between residues in mutation regions

Atom 1	Atom 2	Occupancy (%)		Bonding type
		WT	M10	
K/G208.N	G/D206.O	0.0	23.5	Hydrogen bond
E207.N	A205.O	0.1	78.0	Hydrogen bond
A205.N	E207.OE2	0.0	24.6	Hydrogen bond
R220.NH2	E207.OE2	7.9	99.0	Salt bridge
G209.N	V265.O	0.0	19.1	Hydrogen bond
G209.N	N268.OD1	0.0	17.9	Hydrogen bond
Q/A66.N	L64.O	0.0	6.8	Hydrogen bond
T68.N	L65.O	0.3	5.0	Hydrogen bond
Q/G286.N	L282.O	0.0	3.9	Hydrogen bond
Q/G286.N	S284.O	0.2	4.7	Hydrogen bond

升相吻合。以上 3 个区域中这些新的作用力共同提高了 PML 整体的稳定性。

类似 G206D/K208G 区域作用力重排的现象在 PML 的其他研究中也被发现^[15]：利用高通量筛选得到的高甲醇耐受性的 PML 变体包含 G202E 及 K208N 突变，通过 X 射线晶体衍射观察到聚集在 Ca^{2+} 结合位点附近的相同 loop 被重塑，且 E207 与盖子 α_8 上的 R220 也形成了盐桥。结合本研究结果推测，PML 相较铜绿假单胞菌脂肪酶 (*Pseudomonas aeruginosa* lipase, PAL) 在 Ca^{2+} 结合位点附近多出的 5 个残基组

成的 loop^[12]是与其生物学性质紧密相关的热点区域,它不仅会影响酶活、热稳定性,还会影响有机溶剂耐受性,这可以为后续 PML 的改造提供参考。

3 讨论与结论

基于结构和序列信息构建起来的多种计算方法为天然蛋白质的热稳定性改造提供了更精确且省时省力的途径。本研究中,我们综合了多重计算设计,建立了以 ABACUS、PROSS 以及 FoldX 三种互补算法的两两交集为基准,利用 G_{REMLIN} 和 PSSM 进一步减小突变文库的计算设计方案。以 PML 作为模式蛋白,利用此筛选方案得到的 18 个单点突变设计中,有 7 个热稳定性提高的突变体,突变效率达到了 38.9%,且其中大部分的突变体较好地保存了酶活力,证明了这种综合已有通用计算方法和程序的计算策略在兼顾用户友好性的同时具有较高的筛选效率,可以用于其他酶的热稳定性改造。选择热稳定性和酶活均高于 WT 的 5 个突变体,采用贪婪累积策略进行组合突变,获得 10 个组合突变体。在组合路径中 K208G 和 G206D 的组合展现出较好的上位性,其两点组合突变体的热稳定性提高了 9.55 °C;最终五点组合突变体 M10 的热稳定性比 WT 高出 10.63 °C,同时其酶活达到了 WT 的 140%。利用分子动力学模拟对突变引起的结构变化进行初步研究,发现 M10 的部分柔性区域波动小于 WT;作用力分析表明突变位点附近有新的作用力生成,其中 G206D 与 K208G 突变不仅增强了所在 loop 的刚性,还产生了新的稳定相互作用,并潜在地稳定了对稳定性和活性至关重要的 Ca²⁺结合位点,为 M10 整体热稳定性的提高提供了主要贡献,其所在 loop 也是经定向进化筛选出的对热

稳定性有重要影响的热点区域,可以成为 PML 后续计算改造的重点关注区域。本研究为提高天然酶热稳定性提供一种高效精巧的计算设计方案,同时为 PML 的进一步工业应用提供了基础。

REFERENCES

- [1] Naik S, Basu A, Saikia R, et al. Lipases for use in industrial biocatalysis: specificity of selected structural groups of lipases. *J Mol Catal B Enzym*, 2010, 65(1/2/3/4): 18-23.
- [2] Yu XW, Zhu SS, Xiao R, et al. Conversion of a *Rhizopus chinensis* lipase into an esterase by lid swapping. *J Lipid Res*, 2014, 55(6): 1044-1051.
- [3] Chandra P, Enespa, Singh R, et al. Microbial lipases and their industrial applications: a comprehensive review. *Microb Cell Fact*, 2020, 19(1): 169.
- [4] Filho DG, Silva AG, Guidini CZ. Lipases: sources, immobilization methods, and industrial applications. *Appl Microbiol Biotechnol*, 2019, 103(18): 7399-7423.
- [5] Global Microbial Lipase Industry[EB/OL]. [2021-10-27]. https://www.reportlinker.com/p05899911/Global-Microbial-LipaseIndustry.html?utm_source=GNW.
- [6] Alquati C, de Gioia L, Santarossa G, et al. The cold-active lipase of *Pseudomonas fragi*. Heterologous expression, biochemical characterization and molecular modeling. *Eur J Biochem*, 2002, 269(13): 3321-3328.
- [7] Gao B, Su EZ, Lin JP, et al. Development of recombinant *Escherichia coli* whole-cell biocatalyst expressing a novel alkaline lipase-coding gene from *Proteus* sp. for biodiesel production. *J Biotechnol*, 2009, 139(2): 169-175.
- [8] Lu YP, Lin Q, Wang J, et al. Overexpression and characterization in *Bacillus subtilis* of a positionally nonspecific lipase from *Proteus vulgaris*. *J Ind Microbiol Biotechnol*, 2010, 37(9): 919-925.
- [9] Whangsuk W, Sungkeeree P, Thiengmag S, et al. Gene cloning and characterization of a novel highly organic solvent tolerant lipase from *Proteus* sp. SW₁ and its application for biodiesel production. *Mol Biotechnol*, 2013, 53(1): 55-62.
- [10] Shao H, Hu XM, Sun LP, et al. Gene cloning, expression in *E. coli*, and *in vitro* refolding of a lipase from *Proteus* sp. NH 2-2 and its application for

- biodiesel production. *Biotechnol Lett*, 2019, 41(1): 159-169.
- [11] Misbah A, Koraichi SI, Jouti MAT. Purification, characterization and immobilization of lipase from *Proteus vulgaris* OR34 for synthesis of methyl oleate. *Microbiol Biotechnol Lett*, 2020, 48(4): 491-505.
- [12] Korman TP, Bowie JU. Crystal structure of *Proteus mirabilis* lipase, a novel lipase from the *Proteus*/psychrophilic subfamily of lipase family I.1. *PLoS One*, 2012, 7(12): e52890.
- [13] Invernizzi G, Papaleo E, Grandori R, et al. Relevance of metal ions for lipase stability: structural rearrangements induced in the *Burkholderia glumae* lipase by calcium depletion. *J Struct Biol*, 2009, 168(3): 562-570.
- [14] El Khattabi M, van Gelder P, Bitter W, et al. Role of the calcium ion and the disulfide bond in the *Burkholderia glumae* lipase. *J Mol Catal B Enzym*, 2003, 22(5/6): 329-338.
- [15] Korman TP, Sahachartsiri B, Charbonneau DM, et al. Dieselzymes: development of a stable and methanol tolerant lipase for biodiesel production by directed evolution. *Biotechnol Biofuels*, 2013, 6(1): 70.
- [16] Heater BS, Yang ZF, Lee MM, et al. *In vivo* enzyme entrapment in a protein crystal. *J Am Chem Soc*, 2020, 142(22): 9879-9883.
- [17] Heater BS, Chan WS, Lee MM, et al. Directed evolution of a genetically encoded immobilized lipase for the efficient production of biodiesel from waste cooking oil. *Biotechnol Biofuels*, 2019, 12: 165.
- [18] Schymkowitz J, Borg J, Stricher F, et al. The FoldX web server: an online force field. *Nucleic Acids Res*, 2005, 33(Web Server issue): W382-W388.
- [19] Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol*, 2002, 320(2): 369-387.
- [20] Xiong P, Hu XH, Huang B, et al. Increasing the efficiency and accuracy of the ABACUS protein sequence design method. *Bioinformatics*, 2020, 36(1): 136-144.
- [21] Zhou XQ, Xiong P, Wang M, et al. Proteins of well-defined structures can be designed without backbone readjustment by a statistical model. *J Struct Biol*, 2016, 196(3): 350-357.
- [22] Xiong P, Wang M, Zhou XQ, et al. Protein design with a comprehensive statistical energy function and boosted by experimental selection for foldability. *Nat Commun*, 2014, 5: 5330.
- [23] Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins*, 2011, 79(3): 830-838.
- [24] Park H, Bradley P, Greisen PJ, et al. Simultaneous optimization of biomolecular energy functions on features from small molecules and macromolecules. *J Chem Theory Comput*, 2016, 12(12): 6201-6212.
- [25] 向玉, 张萌, 许菲. 基于多重计算设计策略提高枯草芽孢杆菌脂肪酶的热稳定性. *生物工程学报*, 2020, 36(8): 1556-1567.
- Xiang Y, Zhang M, Xu F. Improving the thermal stability of *Bacillus subtilis* lipase based on multiple computational design strategies. *Chin J Biotech*, 2020, 36(8): 1556-1567 (in Chinese).
- [26] Altschul SF, Gertz EM, Agarwala R, et al. PSI-BLAST pseudocounts and the minimum description length principle. *Nucleic Acids Res*, 2009, 37(3): 815-824.
- [27] Magliery TJ. Protein stability: computation, sequence statistics, and new experimental methods. *Curr Opin Struct Biol*, 2015, 33: 161-168.
- [28] Balakrishnan S, Kamisetty H, Carbonell JG, et al. Learning generative models for protein fold families. *Proteins*, 2011, 79(4): 1061-1078.
- [29] Ovchinnikov S, Kamisetty H, Baker D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife*, 2014, 3: e02030.
- [30] Thomas A, Cutlan R, Finnigan W, et al. Highly thermostable carboxylic acid reductases generated by ancestral sequence reconstruction. *Commun Biol*, 2019, 2: 429.
- [31] Bi JH, Chen SH, Zhao XH, et al. Computation-aided engineering of starch-debranching pullulanase from *Bacillus thermoleovorans* for enhanced thermostability. *Appl Microbiol Biotechnol*, 2020, 104(17): 7551-7562.
- [32] Goldenzweig A, Goldsmith M, Hill SE, et al. Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Mol Cell*, 2016, 63(2): 337-346.
- [33] Wijma HJ, Floor RJ, Jekel PA, et al. Computationally designed libraries for rapid enzyme stabilization. *Protein Eng Des Sel*, 2014, 27(2): 49-58.
- [34] Cui YL, Chen YC, Liu XY, et al. Computational redesign of a PETase for plastic biodegradation under ambient condition by the GRAPE strategy. *ACS Catal*, 2021, 11(3): 1340-1350.

- [35] Lindorff-Larsen K, Piana S, Palmo K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, 2010, 78(8): 1950-1958.
- [36] Jorgensen WL, Chandrasekhar J, Madura JD, et al. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*, 1983, 79(2): 926-935.
- [37] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*, 2007, 126(1): 014101.
- [38] Berendsen HJC, Postma JPM, Van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. *J Chem Phys*, 1984, 81(8): 3684-3690.
- [39] Luzar A, Chandler D. Effect of environment on hydrogen bond dynamics in liquid water. *Phys Rev Lett*, 1996, 76(6): 928-931.
- [40] Sun TT, Qiang SM, Lu C, et al. Composition-dependent energetic contribution of complex salt bridges to collagen stability. *Biophys J*, 2021, 120(16): 3429-3436.
- [41] Barlow DJ, Thornton JM. Ion-pairs in proteins. *J Mol Biol*, 1983, 168(4): 867-885.
- [42] Li GL, Chen Y, Fang XR, et al. Identification of a hot-spot to enhance *Candida rugosa* lipase thermostability by rational design methods. *RSC Adv*, 2018, 8(4): 1948-1957.
- [43] Cui HY, Cao H, Cai HY, et al. Computer-assisted recombination (CompassR) teaches us how to recombine beneficial substitutions from directed evolution campaigns. *Chemistry*, 2020, 26(3): 643-649.
- [44] Carugo O, Pongor S. A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci*, 2001, 10(7): 1470-1473.

(本文责编 陈宏宇)