



基于 AlphaFold 2 和分子对接探讨非还原型聚酮合酶的碳甲基化程序

廖世玉^{1,2}, 刘庆培³, 陈福生^{1,2*}

1 华中农业大学农业微生物资源发掘与利用全国重点实验室, 湖北 武汉 430070

2 华中农业大学食品科学与技术学院, 湖北 武汉 430070

3 中南民族大学药学院, 湖北 武汉 430070

廖世玉, 刘庆培, 陈福生. 基于 AlphaFold 2 和分子对接探讨非还原型聚酮合酶的碳甲基化程序[J]. 微生物学报, 2024, 64(1): 143-160.

LIAO Shiyu, LIU Qingpei, CHEN Fusheng. C-methylation programming of non-reducing polyketide synthases: based on AlphaFold 2 and molecular docking[J]. Acta Microbiologica Sinica, 2024, 64(1): 143-160.

摘要: 【目的】探讨非还原型聚酮合酶(non-reducing polyketide synthase, NR-Pks)的碳甲基化程序差异的原因。【方法】以红色红曲菌(*Monascus ruber*) M7 中红曲色素和桔霉素的 NR-Pks 为研究对象, 采用生物信息学方法和 AlphaFold 2 软件, 分析了这两种 NR-Pks 及其各结构域的序列和结构差异。再基于分子对接等技术, 比较了它们的碳甲基转移酶结构域(C-methyltransferase domain, CMeT)分别与其他结构域及其中间产物的结合特征。【结果】两种 NR-Pks 各结构域的序列和结构相似性高, 但其整体结构差异大, 表明碳甲基化差异可能源于结构域互作差异。进一步分析发现, 桔霉素 Pks 的 CMeT 比红曲色素 Pks 的更容易结合携带底物的酰基载体蛋白结构域(acyl carrier protein, ACP), 使其中间产物更容易受到 CMeT 催化。CMeT 和 β -酮酰基合成酶结构域(β -ketosynthase domain, KS)相比, 与甲基受体底物的结合自由能更低。【结论】NR-Pks 中的 CMeT 能通过与 KS 竞争, 从而影响其产物的碳甲基化程度。研究结果为 Pks 的碳甲基化程序研究提供了新思路。

关键词: 非还原型聚酮合酶; 红曲菌; 碳甲基转移酶; AlphaFold 2; 分子对接

资助项目: 国家自然科学基金(31730068, 31330059)

This work was supported by the National Natural Science Foundation of China (31730068, 31330059).

*Corresponding author. E-mail: chenfs@mail.hzau.edu.cn

Received: 2023-05-13; Accepted: 2023-08-28; Published online: 2023-09-01

C-methylation programming of non-reducing polyketide synthases: based on AlphaFold 2 and molecular docking

LIAO Shiyu^{1,2}, LIU Qingpei³, CHEN Fusheng^{1,2*}

1 National Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, Hubei, China

2 College of Food Science and Technology, Huazhong Agricultural University, Wuhan 430070, Hubei, China

3 School of Pharmaceutical Sciences, South-Central Minzu University, Wuhan 430070, Hubei, China

Abstract: [Objective] To explore the reasons for differences in the C-methylation programming of non-reducing polyketide synthases (NR-Pkss). [Methods] We used bioinformatics tools and AlphaFold 2 to compare the domain sequences and structures of the NR-Pkss involved in the synthesis of *Monascus* pigment and citrinin in *Monascus ruber* M7, i.e., Mr-PksPT and Mr-PksCT. Furthermore, we employed molecular docking to compare the binding of C-methyltransferase domains (CMeTs) with other domains and the intermediates of the two NR-Pkss. [Results] The large differences of the overall structure and the high similarity of domain sequence and structure between the two NR-Pkss suggested that the differences of C-methylation programming between NR-Pkss may be resulted from domain interactions. The CMeT of Mr-PksCT was more likely to bind to the acyl carrier protein (ACP) carrying the substrate than that of Mr-PksPT, making the intermediate more easily catalyzed by CMeT. Moreover, CMeT had lower binding free energy to methyl receptor substrate than the β -ketosynthase domain (KS). [Conclusion] The CMeTs of NR-Pkss can affect the C-methylation of the products by competing with KS. The findings provide a new idea for the study of C-methylation programming of Pkss.

Keywords: non-reducing polyketide synthase; *Monascus* spp.; C-methyltransferase; AlphaFold 2; molecular docking

真菌迭代I型聚酮合酶(iterative type I polyketide synthase, Pks)是包含多个结构域的复合酶^[1]。它以乙酰或/和丙二酰等为底物,通过重复利用其不同结构域,逐步延伸碳链长度,形成聚酮化合物(polyketide, Pk)^[2]。根据 Pks 结构域的不同,可分为非还原型 Pks (non-reduced Pks, NR-Pks)、部分还原型 Pks (partially reduced Pks)和高度还原型 Pks (highly reduced Pks)^[3-4]。其中, NR-Pks 的基本结构域通常包括起始单元酰基载体蛋白转酰酶(starter unit: acyl-carrier protein transacylase, SAT)、 β -酮酰基合成酶

(β -ketosynthase, KS)、酰基转移酶(acyl transferase, AT)、产物模板(product template, PT)和酰基载体蛋白(acyl carrier protein, ACP)^[3,5]。有些 NR-Pks 还包括碳甲基转移酶(C-methyltransferase, CMeT),以及释放结构域——硫酯酶(thioesterase)、克莱森环化酶(Claisen cyclase)或者还原酶(reductase, R)^[6-7]。

红曲色素(*Monascus* pigments, MPs)和桔霉素(Citrinin, CIT)是由红曲菌(*Monascus* spp.)和青霉(*Penicillium* spp.)等丝状真菌产生的典型PK^[8-11]。合成 MPs 的 Pks (Pks of MPs, PksPT)和

CIT 的 Pks (Pks of CIT, PksCT)均属于 NR-Pks, 它们的结构域组成完全一致, 且产生的第一个游离 PK 均为苯甲醛类(benzaldehyde, BA)化合物^[12-13], 但 BA 甲基化程度不同, PksPT 产生的 BA 仅甲基化 1 次而 PksCT 的甲基化 3 次[PksPT 和 PksCT 的结构域组成和产物结构等数据已提交到国家微生物科学数据中心(National Microbial Data Center, NMDC), 见 NMDCX0000225 中的图 S1]^[10,14]。

天然产物的甲基化是其结构和功能等多样性形成的重要机理之一^[15]。根据甲基受体原子的不同, 甲基化可分为氧甲基化、氮甲基化、碳甲基化和硫甲基化^[16]。PksPT 和 PksCT 产生的 BA 的甲基化均发生在碳原子上, 所以属于碳甲基化。关于 Pks 产物(指 Pks 产生的第一个游离产物, 下同)的碳甲基化过程目前认为是发生在碳链的延长过程中, 而不是发生在 PK 产物从 Pks 上释放之后^[17-18]。Storm 等^[17]通过对 PksCT 结构域的解构和重组实验发现, PksCT 产物 BA 的碳甲基化仅发生在碳链延伸过程中, 且碳链延伸离不开正确的碳甲基化。随后, 他们通过对来自紫色红曲菌 (*M. purpureus*) 的 PksCT (Q65Z23)、短密青霉(*Penicillium brevicompactum*) 的 MpaC (F1DBA9)、黑曲霉(*Aspergillus niger*) 的 DtbA^[7], 以及构巢曲霉(*A. nidulans*)的 PkeA (Q5AUX7)、AusA (Q5ATJ7)和 PkbA (Q5AZ32), 共 6 种 NR-Pks 的 CMeT 结构域交换和体外酶学等实验发现, CMeT 浓度高低可调控产物的碳甲基化程度, 并发现产物甲基化程度可能与 Pks 中间产物碳链延伸过程中 CMeT 与 KS 结构域的动力学平衡(竞争性)相关^[18]。在洛伐他汀(lovastatin) Pks (一种高度还原型 Pks)产物的甲基化研究中也发现, Pks 产物的甲基化程度可能也与 CMeT 和酮基还原酶(ketoreductase, KR)之间竞争性相关^[19]。但是甲基化具体机理目前仍不清楚。

以上关于 Pks 产物甲基化的研究都是通过对 CMeT 的蛋白晶体结构分析等获得的。由于 CMeT 等蛋白晶体及其结构信息等的获得非常繁琐且困难, 所以到目前为止(2023 年 4 月底)在蛋白质数据库(Protein Data Bank, PDB)中已公布的 Pks 的 CMeT 晶体结构信息仅为 11 个, 其中来自丝状真菌的只有 4 个(见 NMDCX0000225 中的表 S1), 这在很大程度上限制了关于 Pks 的 CMeT 在 Pks 产物甲基化过程中作用的深入研究。非常幸运的是, 最近采用 AlphaFold 2 软件预测蛋白质的三维结构取得了长足进步, 采用 AlphaFold 2 预测蛋白质结构的技术位也被列 2021 年“十大科学突破进展”之首^[20]。目前 AlphaFold 2 预测的蛋白质结构数据库(AlphaFold protein structure database, AlphaFold DB)已包含了超过 2 亿个蛋白质的预测结构^[21-22]。此外, 分子对接(molecular docking)技术是研究分子间(如配体和受体)相互作用, 并预测其结合自由能的一种理论模拟方法, 常用于计算机辅助药物设计领域^[23]。采用 AlphaFold 2 软件结合分子对接技术, 可以很好地分析蛋白质与底物、蛋白质与蛋白质的结合特征等^[24-27]。

尽管关于 PksCT 产物 BA 的碳甲基化程序(只有延伸碳链的特定位置才可以发生碳甲基化)已有相关的研究报道^[17], 但与 PksCT 同属于 NR-Pks, 结构域组成完全相同(图 S1), 产物也为 BA 的 PksPT 的碳甲基化程序, 以及 PksCT 和 PksPT 产物的甲基化差异等目前还未见相关研究报道。因此, 本研究首先以生物信息学方法, 对真菌中的 PksPT 和 PksCT 分别进行了聚类分析, 并对 PksPT 和 PksCT 的结构域组成和氨基酸序列的保守性进行了比较, 然后以 AlphaFold 2 预测和分析了红色红曲菌 M7 的 PksPT 和 PksCT (Mr-PksPT 和 Mr-PksCT)及其各结构域的三维结构。在此基础上, 以分子对接

与自由能评估, 比较了 Mr-PksPT 和 Mr-PksCT 在碳链延伸过程中的中间产物与 CMeT 和 KS 结构域等的结合效率, 探讨了 Mr-PksPT 和 Mr-PksCT 产物的碳甲基化差异的形成机理, 并提出了 Mr-PksPT 和 Mr-PksCT 产物 BA 的形成过程。

1 材料与amp;方法

Pks 系统发育树的构建: 采用 MEGA 7.0 软件^[28], 基于邻接法构建来自真菌且分别与 Mr-PksPT 和 Mr-PksCT 同源的 PksPT 和 PksCT 的系统发育树。建树的检验方法为步长检验, 检验次数设定为 1 000 次。

Pks 保守结构域的预测: 利用 InterPro 92.0^[29] 对文中涉及的 PksPT 和 PksCT 进行保守结构域的预测。

结构域的多序列比对: 采用 Geneious 10.2.2 软件^[30]对 PksPT 和 PksCT 中各结构域进行多序列比对。

蛋白质三维结构预测: 利用 AlphaFold 2^[21] 对 PksPT 和 PksCT 及其各结构域进行蛋白质三维结构预测。

分子对接: 采用 Schrödinger 12.8 软件中的 Glide XP 进行分子对接, 再进一步通过 Prime MM/GBSA 进行结合自由能计算。

蛋白质对接: 利用 ZDOCK 3.0.2^[31]进行蛋白质对接, 再进一步通过 Prodigy^[32]对蛋白质复合物进行特性分析。

蛋白质三维结构可视化: Pymol 2.5.4 (<https://pymol.org/#download>)。

2 结果与分析

2.1 PksPT 和 PksCT 的比较分析

2.1.1 PksPT 和 PksCT 的聚类分析

基于红色红曲菌 M7 中 Mr-PksPT 和 Mr-PksCT 的氨基酸序列, 以 BLASTp 对 NCBI

数据库中截至 2023 年 4 月已发布的 4 415 株真菌基因组(包括 16 株已公布的红曲菌基因组)和本实验室已完成测序还未公布的 5 株红曲菌基因组(表 S2, 见国家微生物科学数据中心, 编号 NMDCX0000225), 进行同源性比对, 分别发现了 106 条和 73 条 Mr-PksPT 与 Mr-PksCT 的同源序列 (identity>50%, coverage>90%, E-value=0) (表 S3 和 S4, 见国家微生物科学数据中心, 编号 NMDCX0000225)。它们的系统发育树见图 1。

由图 1A 可知, PksPT 主要存在于红曲菌、篮状菌(*Talaromece* spp.)、炭团菌(*Hypoxylon* spp.)和球孢子菌(*Coccidioides* spp.)等真菌中。其系统发育树包括 3 个分支, 第 I 分支包含 MPs 主要生产菌株红曲菌、篮状菌和曲霉(*Aspergillus* spp.)^[9,12,33]; 第 II 分支和第 III 分支主要包括球炭团菌、孢子菌和木霉菌(*Trichoderma* spp.), 与 MPs 主要生产菌株在生物分类学中关系较远^[12,34]。每个分支中又可进一步分为不同的亚分支。其中, 21 株红曲菌以 100% 的置信度聚集于 I₁ 亚分支, 而谢瓦氏曲霉(*A. chevalieri*)一个菌株单独位于 I₃ 亚分支。由图 1B 可知, PksCT 主要存在于红曲菌、青霉菌和曲霉等真菌中。其系统发育树包括 4 个分支, 第 I 分支包括 74 条 PksCT 序列中的 64 条, 红曲菌、青霉菌和曲霉等 CIT 的主要生产菌株均位于第 I 分支中^[10,35-36]。第 I 分支又可下分为不同亚分支。其中, 13 株红曲菌以 95% 的置信度聚集于 I₁ 亚分支, 而其他 8 株红曲菌中不存在 PksCT 的同源基因(表 S5, 见国家微生物科学数据中心, 编号 NMDCX0000225)。第 II、III 和 VI 分支包含的 PksCT 序列均较少, 菌株主要为木霉菌、假裸囊菌(*Pseudogymnoascus* spp.)等真菌, 与 CIT 主要生产菌株的亲缘关系较远。

上述聚类分析结果表明, 不同红曲菌 PksPT 的起源(来源)可能相同, 而红曲菌 PksCT 的起源可能存在差异。

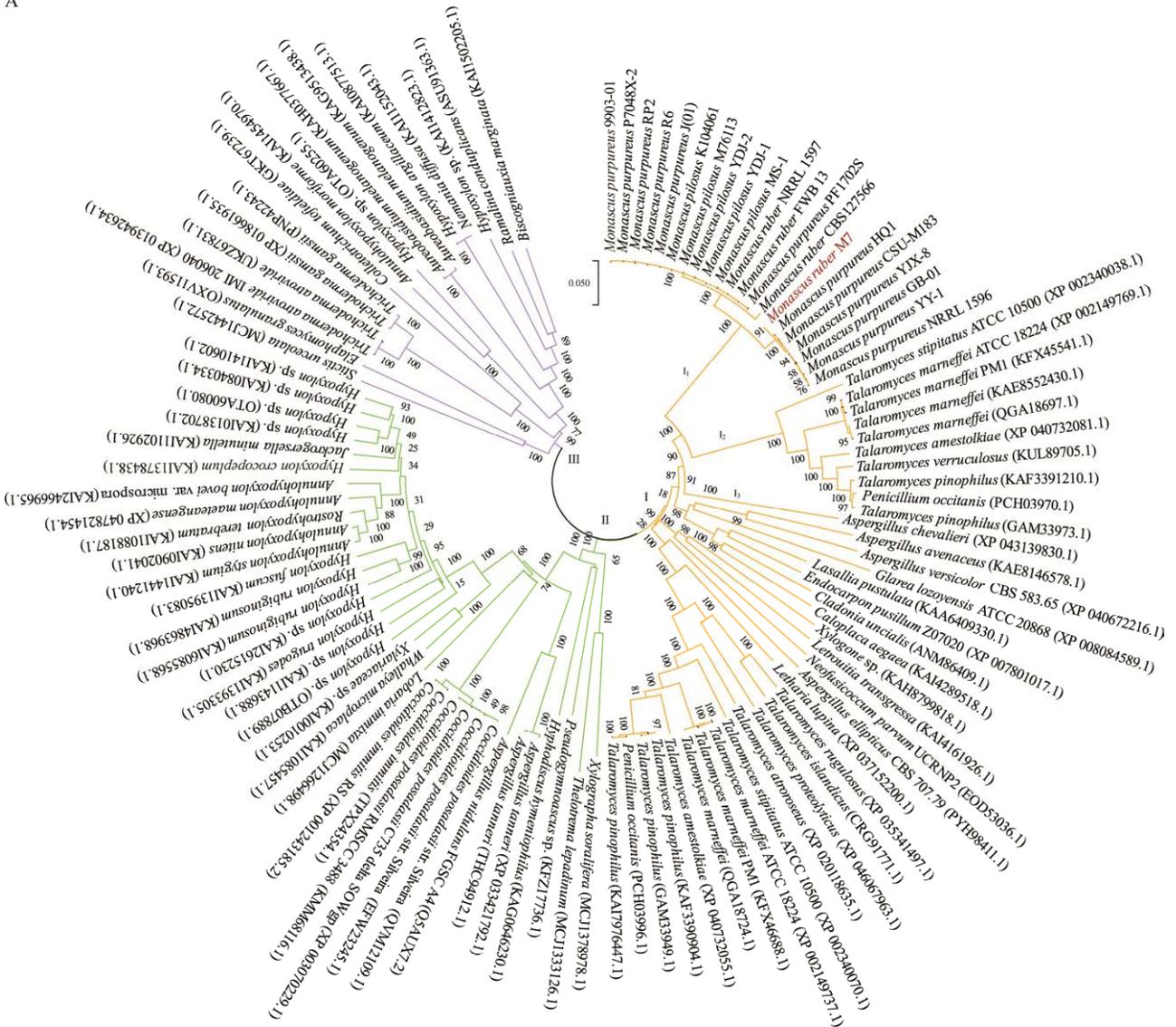
2.1.2 红曲菌中 PksPT 和 PksCT 结构域和氨基酸序列的保守性分析

根据红曲菌的 PksPT 和 PksCT (Ms-PksPT 和 Ms-PksCT), 分别与红色红曲菌 M7 中 Mr-PksPT 和 Mr-PksCT 序列相比较的覆盖度 (coverage)、一致性(identity)和氨基酸数量(length)的异同, 将 21 条 Ms-PksPT 和 13 条 Ms-PksCT 分别分为 5 类(Ms-PksPT 1-5)和 5 类(Ms-PksCT 1-5), 其中, Mr-PksPT 和 Mr-PksCT 分别包含在 Ms-PksPT1 和 Ms-PksCT1 中(见 NMDCX0000225 中的表 S6 和 S7)。各类 Ms-PksPT 和 Ms-PksCT

的结构域见图 2。

由图 2 可知, Ms-PksPT 和 Ms-PksCT 的结构域组成均为 SAT-KS-AT-PT-(ACP)_n-CMeT-R (Ms-PksPT 中 n=2, PksCT 中 n=1)。但在 Ms-PksPT 中, Ms-PksPT5 缺少 SAT 结构域, 理论上其不能合成 MPs^[5], 值得进一步研究。另外, Ms-PksPT 均含有 2 个串联 ACP, 串联 ACP 可能具有更高的 MPs 合成效率^[37]。对于 Ms-PksCT, 除 Ms-PksCT2、Ms-PksCT3 和 Ms-PksCT5 没有 SAT 结构域外, 其他 Ms-PksCT 的结构域组成相同。

A



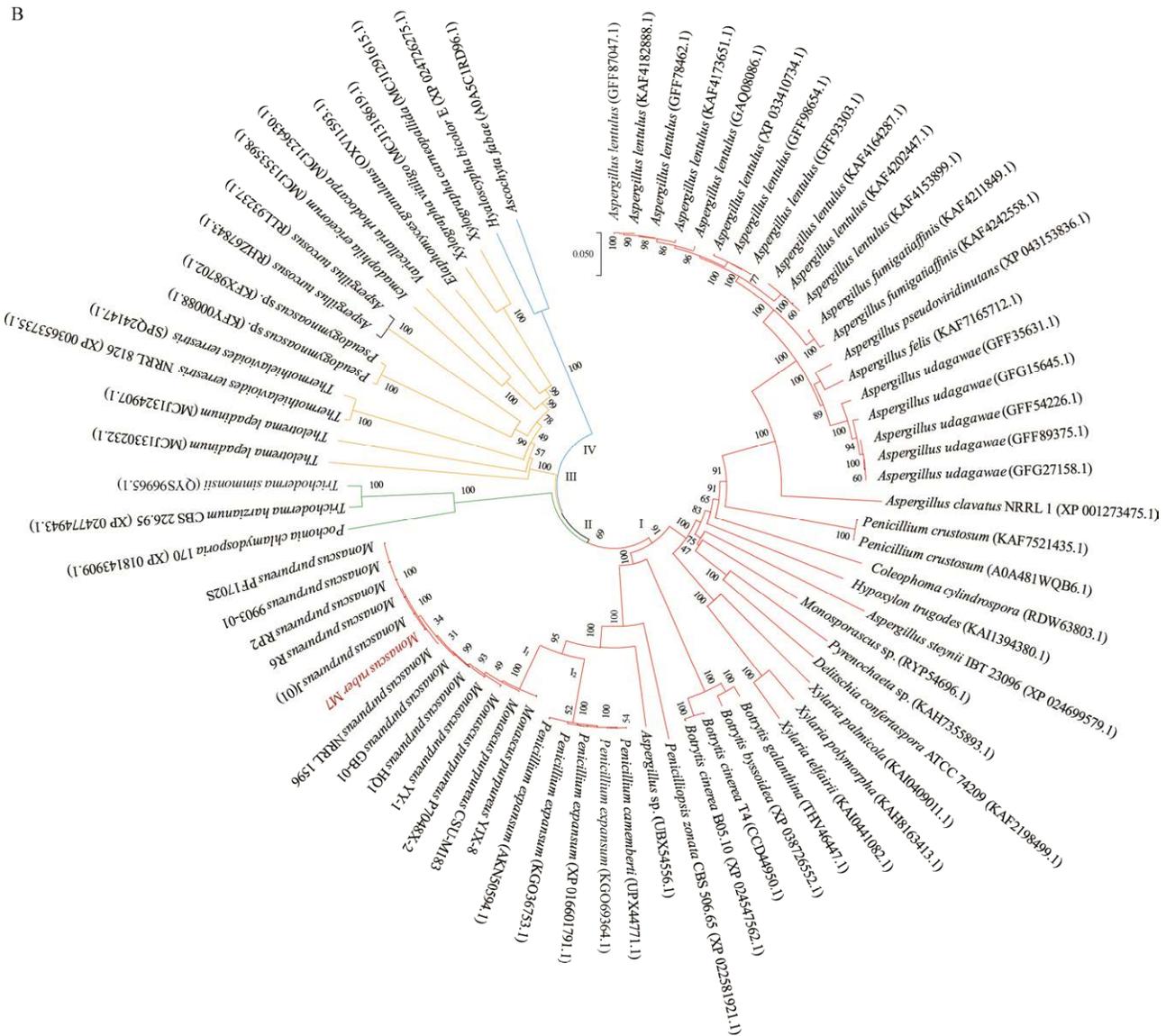


图 1 Mr-PksPT 和 Mr-PksCT 及其同源蛋白的系统发育分析

Figure 1 Phylogenetic analysis of Mr-PksPT and Mr-PksCT and their homologous proteins. A: Phylogenetic analysis of Mr-PksPT and its homologous proteins. B: Phylogenetic analysis of Mr-PksCT and its homologous proteins. All sequences are indicated by “strain name (NCBI number)”.

由表 1 可知，进一步对 Ms-PksPT 1-5 和 Ms-PksCT 1-5 中各结构域的氨基酸数量(表 S8, 见国家微生物科学数据中心, 编号 NMDCX0000225)、保守氨基酸序列(基序)和活性位点(图 S2, 见国家微生物科学数据中心, 编号 NMDCX0000225)进行比较。其中, 只有 Ms-PksCT3 的释放结构域 R 的氨基酸数量偏小,

其余所有的 Ms-PksPT 和 Ms-PksCT 各结构域之间的氨基酸数量差异不明显。此外, Ms-PksPT 和 Ms-PksCT 中 SAT 结构域的保守氨基酸序列均为 G₁₂₉XC₁₃₁XG₁₃₃, 活性位点为半胱氨酸, 氨基酸侧链为巯基; KS 结构域的保守氨基酸序列均为 motif I-VI, 活性位点为催化三联体 Cys₁₆₃-His₂₉₈-His₃₃₇; AT 结构域的保守氨基酸序

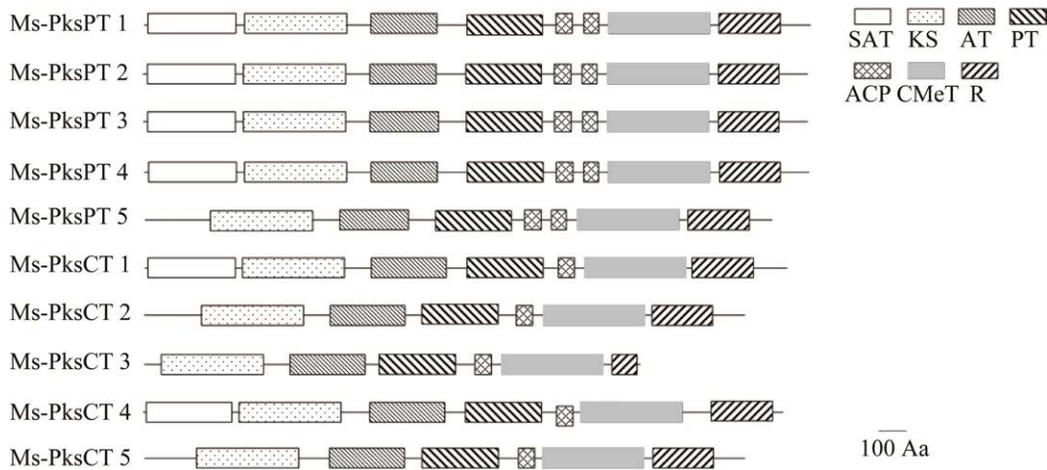


图 2 红曲菌中 PksPT 和 PksCT 的结构域组成

Figure 2 Domain composition of PksPT and PksCT in *Monascus* spp.

表 1 Ms-PksPT 和 Ms-PksCT 各结构域的功能、保守基序及其活性位点

Table 1 Functions, conserved motifs and active sites of domains in Ms-PksPT and Ms-PksCT

Domains	Functions	Conserved motifs	Active sites	References
SAT	Choose acetyl-CoA as the starting unit and load it into KS via ACP	G ₁₂₉ XC ₁₃₁ XG ₁₃₃	Cys ₁₃₁	[38]
KS	Extend polyketide intermediates by decarboxylation condensation	D ₈₀ PQXRXML ₈₇ , G ₁₅₄ P(G/S) ₁₅₆ , D ₁₆₀ TACSSS ₁₆₆ , G ₂₂₃ YCRGE ₂₂₈ , E ₂₉₆ AHGTGTXVGD ₃₀₇ , G ₃₂₉ SVKGLXGHTE ₃₃₉	Cys ₁₆₃ , His ₂₉₈ , His ₃₃₇	[39]
AT	Choose malonyl-CoA as the extension unit and load it into ACP	G ₈₄ H SXG ₈₈	Ser ₈₆	[40]
PT	Catalyze C2–C7 regiospecific monocyclization	H ₂₆ X9P ₃₆ , D ₂₀₉ X3QX6N ₂₂₀	His ₂₆ , Asp ₂₀₉	[41]
ACP	Activate and shuttle substrates	G ₂₇ XDSL ₃₁	Ser ₃₀	[42-43]
CMET	Add one or more methyl groups from S-adenosylmethionine to α-C of the acceptor substrate	I ₂₁₁ LEMGXGXGG ₂₂₀ , E ₂₃₇ YTXDL(A/S) ₂₄₄ , M ₂₆₁ KFX3(D/N)IE ₂₆₉ , Q ₂₈₀ HX3AXNXVH ₂₉₀ , S ₂₉₈ TX6LR ₃₀₇ , D ₃₀₉ GFLXXLEM ₃₁₇	Tyr ₁₇₈ , His ₂₉₀ , Glu ₃₁₆	[15,44]
R	Catalyze the NAD(P)H-dependent reductive release	G ₃ ATGSLG ₉ , V ₂₅ VCXNR ₃₀ , V ₁₄₂ GYYP ₁₄₇ , G ₁₆₈ YXXAK ₁₇₃ , F ₁₉₁ XPM ₁₉₄ , G ₂₃₄ XLSWXP ₂₄₀	Tyr ₁₆₉ , Lys ₁₇₃	[45-46]

列均为 G₈₄H SXG₈₈, 活性位点为丝氨酸, 氨基酸侧链为羟基; PT 结构域的保守氨基酸序列均为 motif I–II, 活性位点为 His₂₆ 和 Asp₂₀₉; ACP 结构域的保守氨基酸序列均为 G₂₇XDSL₃₁, 是磷酸泛酰巯基乙胺的结合位点, 活性位点是丝氨酸; CMET 结构域的保守氨基酸序列均为 motif I–VI, 是 S-腺苷甲硫氨酸(S-adenosylmethionine, SAM)的结合位点, 活性位点为催化三联体 Tyr₁₇₈-His₂₉₀-Glu₃₁₆; R 结构域的保守氨基酸序列

均为 motif I–VI, 是 NAD(P)H 结合位点, 活性位点为 Tyr₁₆₉ 和 Lys₁₇₃。

总之, Ms-PksPT 和 Ms-PksCT 各结构域之间的保守氨基酸序列和活性位点相同。

2.1.3 Mr-PksPT 和 Mr-PksCT 及其结构域的三维结构预测与相似性比较分析

蛋白质三维结构的相似性常采用均方根偏差(root mean square deviation, RMSD)值来衡量, RMSD 值越小则相似性越高。通常如果 2 个蛋

白质结构之间的 RMSD 值小于 2 Å, 则可认为 2 个蛋白质的差异不明显, 相似性好^[47]。采用 AlphaFold 2 软件分别对红色红曲菌 M7 的 Mr-PksPT 和 Mr-PksCT 及其结构域的三维结构进行预测(图 S3, 见国家微生物科学数据中心, 编号 NMDCX0000225), 并通过 RMSD 值对三维结构相似性进行了比较(表 2)。

由表 2 可知, Mr-PksPT 和 Mr-PksCT 中各结构域比较的 RMSD 值均小于 2 Å, 说明 PksPT 和 PksCT 中各结构域的三维结构差异不明显, 但是 Mr-PksPT 和 Mr-PksCT 整体三维结构比较的 RMSD 值为 12.426 Å, 远远大于 2 Å, 表明 Mr-PksPT 和 Mr-PksCT 的三维结构之间差异明显。这也说明 Mr-PksPT 和 Mr-PksCT 三维结构之间的差异主要可能源于结构域之间的位置分布和相互作用等, 而不是结构域本身的三维结构^[48]。这表明 Mr-PksPT 和 Mr-PksCT 产物的碳甲基化差异, 可能与它们的 CMeT 和 KS 对 ACP 上中间产物(底物)的竞争力大小以及各结构域之间的互作相关。

2.2 基于结构域互作和分子对接等分析 Mr-PksPT 和 Mr-PksCT 的碳甲基化程序

2.2.1 基于结构域互作比较 Mr-PksPT 和 Mr-PksCT 的碳甲基化程序

蛋白质间的静电相互作用、互作界面的残基

连接(inter-residue contacts, ICs)以及非互作表面(non-interacting surface, NIS)残基性质等均是影响蛋白质相互作用的主要因素^[49-51], 而蛋白质互作的强弱常用结合亲和力(binding affinity, BA)或者平衡解离常数(equilibrium dissociation constant, K_D)来描述^[32]。影响 Mr-PksPT 和 Mr-PksCT 产物甲基化程序的结构域(蛋白质)主要包括 ACP、KS 和 CMeT^[18], 所以首先通过静电势能图, 分析比较了 Mr-PksPT 和 Mr-PksCT 的 ACP、KS 和 CMeT 结构域的表面电荷差异, 结果见图 3。

由图 3 可知, Mr-PksCT 的 CMeT 结构域的底物结合口袋入口表面(图 3B)比 Mr-PksPT 的(图 3A)带有更多正电荷, 且 Mr-PksCT 的 ACP 结构域底物结合口袋的入口表面(图 3B)比 Mr-PksPT 的(图 3A)带有更多带负电荷, 根据正负电荷相互吸引的原理, Mr-PksCT 的 CMeT 与其 ACP 更容易结合。也就是说从 ACP 和 CMeT 结构域底物结合口袋入口处所带的电荷看, Mr-PksCT 产物的甲基化比 Mr-PksPT 的具有更高的效率。但是, PksPT 和 PksCT 的 KS 结构域底物结合口袋入口表面均带电性弱, 所以基于静电相互作用的 Mr-PksCT 和 Mr-PksPT 的 ACP 与其对应的 KS 结构域的结合情况, 还有待进一步研究。

表 2 Mr-PksPT 和 Mr-PksCT 及其各对应结构域间的相似性比对

Table 2 Similarity comparison between Mr-PksPT and Mr-PksCT and their respective domains

Mr-PksPT/Mr-PksCT	Coverage (%)	Identity (%)	RMSD (Å)
SAT	99	41.90	0.811
KS	100	70.53	0.270
AT	100	54.31	0.513
PT	100	41.75	0.676
ACP ₁ /ACP	98	48.48	1.299
ACP ₂ /ACP	76	33.33	1.214
CMeT	96	51.77	0.725
R	100	59.18	0.450
PksPT/PksCT	98	50.23	12.426

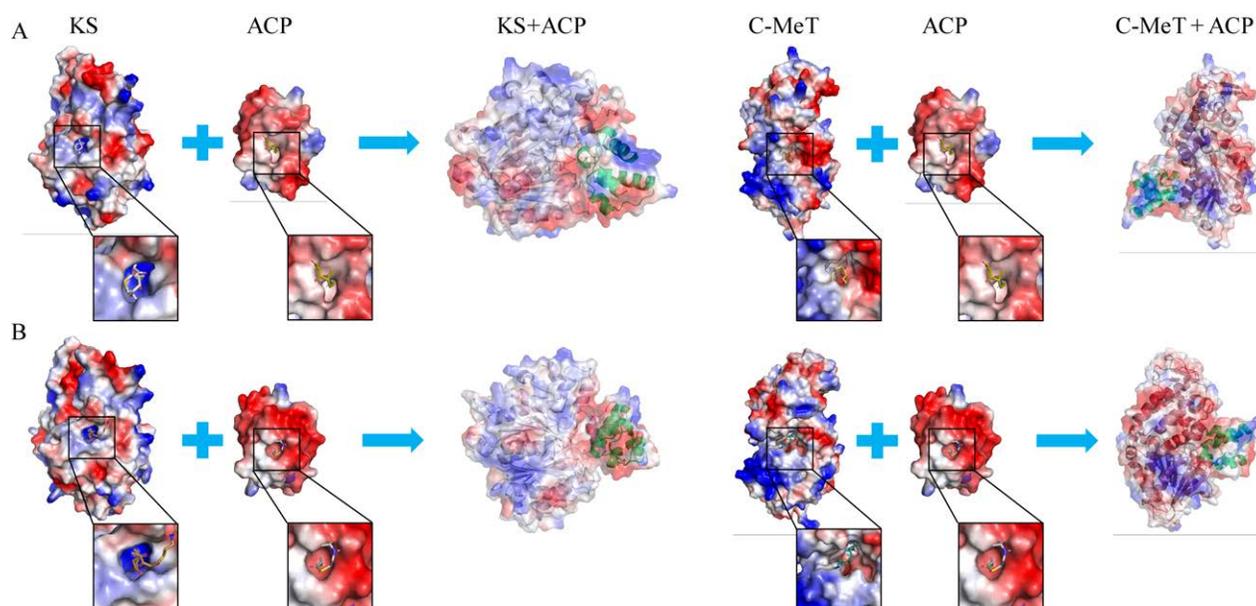


图 3 Mr-PksPT 和 Mr-PksCT 结构域的表面电荷分析

Figure 3 Surface charge analysis of domains in Mr-PksPT and Mr-PksCT. A: Charged surfaces of KS, ACP, CMeT, KS+ACP and CMeT+ACP in Mr-PksPT. B: Charged surfaces of KS, ACP, CMeT, KS+ACP and CMeT+ACP in Mr-PksCT. KS, CMeT and ACP are shown in surface view. KS+ACP and CMeT+ACP are presented in semi-transparent surface view mixed with cartoon view. Red represents negative potentials, blue represents positive potentials, and white indicates zero.

进一步以蛋白质-蛋白质复合物的结合亲和力预测工具 Prodigy^[32], 分别对 Mr-PksPT 和 Mr-PksCT 中 KS-ACP 和 CMeT-ACP 相互作用的 ICs、NIS 残基性质、BA 与 K_D 等蛋白复合物的结合特性进行了分析(表 3)。其中, ICs 类型主要

包括带电/带电、带电/极性、带电/非极性、极性/极性、非极性/极性和非极性/非极性连接等。

由表 3 可知, Mr-PksPT 中 KS-ACP 互作比 Mr-PksCT 中互作的 BA 和 K_D 都更小, 表明 Mr-PksPT 中 KS-ACP 互作时形成的蛋白质复合

表 3 结构域相互作用的特性

Table 3 Properties of domain interactions

Properties	PksPT		PksCT	
	KS-ACP	CMeT-ACP	KS-ACP	CMeT-ACP
ICs_total	99	84	68	117
ICs_charged/charged	3	7	5	8
ICs_charged/polar	11	7	9	8
ICs_charged/apolar	14	23	16	21
ICs_polar/polar	5	0	2	13
ICs_apolar/polar	29	12	18	36
ICs_apolar/apolar	37	35	18	31
NIS_apolar (%)	40.37	38.85	38.61	41.72
BA (kcal/mol)	-12.2	-10.0	-10.7	-12.2
K_D (mol/L)	1.2×10^{-9}	4.9×10^{-8}	1.4×10^{-8}	1.2×10^{-9}

物更稳定,也就说明 Mr-PksPT 中 ACP 比 Mr-PksCT 中的更容易与其对应的 KS 发生相互作用。其原因包括 Mr-PksPT 中 KS-ACP 复合物的残基间连接数更多,以及非互作表面的非极性残基百分比更大。同理,Mr-PksCT 中 CMeT-ACP 互作比 Mr-PksCT 中互作的 BA 和 K_D 都更小,表明 Mr-PksCT 中 CMeT-ACP 互作时形成的蛋白质复合物更稳定,也就说明 Mr-PksCT 中 ACP 比 Mr-PksPT 中的更容易与其对应的 CMeT 发生相互作用。其原因包括 Mr-PksCT 中 CMeT-ACP 复合物的残基间连接数更多,以及非互作表面的非极性残基百分比更大。其中,ACP 与 KS、CMeT 互作的主要连接残基类型为非极性/非极性和极性/非极性连接,而 Mr-PksPT 中 CMeT-ACP 的极性/非极性连接较少,带电/非极性连接较多。此外,CMeT-ACP 互作的确比 KS-ACP 互作有更多的带电残基参与,与图 3 中对 KS 底物结合口袋入口处表面电荷少故带电残基连接少的推

测一致。

2.2.2 基于分子对接分析 Mr-PksPT 和 Mr-PksCT 的碳甲基化程序

分子对接是研究蛋白质-底物相互作用的有力工具^[52]。影响 Mr-PksPT 和 Mr-PksCT 产物碳甲基化程序的分子对接包括 CMeT 与甲基供体底物 SAM 的互作,以及 CMeT 与 KS 竞争性互作^[18-19]。红色红曲菌 M7 中 Mr-PksPT 和 Mr-PksCT 的 CMeT 结构域的三维结构均可被分为 N-端连接、N-端亚结构域和 C-端亚结构域(图 4)。C-端亚结构域有一个罗曼斯折叠,用于结合 S-腺苷甲硫氨酸(S-adenosyl methionine, SAM),包括 7 个标准的 β -折叠和分布于两侧的 α -螺旋,属于 I 类甲基转移酶^[15]。位于 C-端亚结构域的核心插入部分构成底物连接口袋的底部。N-端亚结构域为一个未表征的螺旋折叠,像一个盖子覆盖在底物结合口袋上方,发挥底物选择作用^[17,53]。

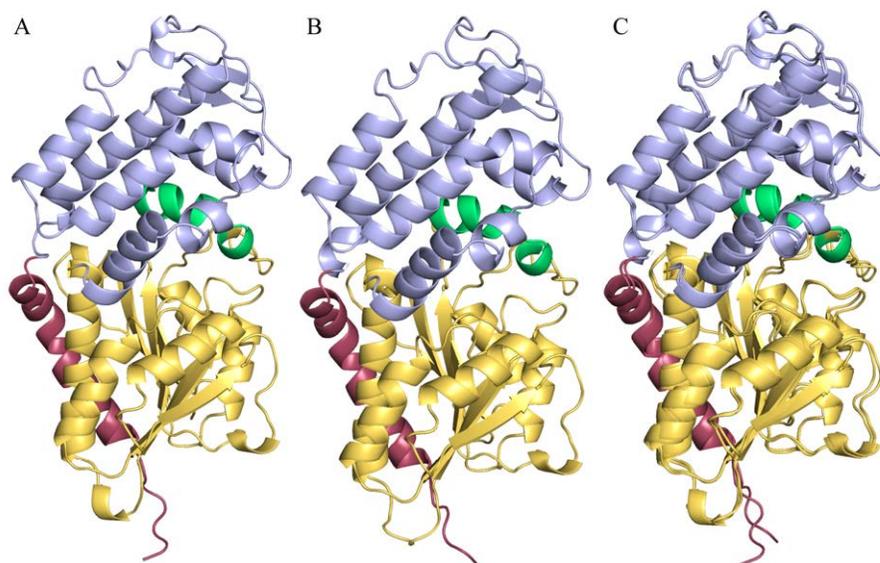


图 4 Mr-PksPT 和 Mr-PksC 的 CMeTs 的结构比较

Figure 4 Structural comparison of Mr-PksPT and Mr-PksCT CMeTs. A: Structure of Mr-PksPT CMeT. B: Structure of Mr-PksCT CMeT. C: Superimposed view of Mr-PksPT and Mr-PksCT CMeT. Magenta: N-terminal linkage; Purple: N-terminal subdomain; Yellow: C-terminal subdomain; Green: Core insertion.

将 Mr-PksPT 和 Mr-PksCT 的 CMeTs 结构域分别与甲基供体底物 SAM 进行分子对接。Mr-PksPT CMeT 中的 Asp₂₆₂ 对应 Mr-PksCT CMeT 中的 Asn₂₆₀, 与 SAM 的腺嘌呤基团形成氢键。Mr-PksPT CMeT 中的 Asp₂₃₇ 对应 Mr-PksCT CMeT 中的 Asp₂₃₅, 与 SAM 的核糖羟基部分形成氢键^[15]。Mr-PksPT CMeT 中的 Thr₂₁₆ 对应 Mr-PksCT CMeT 中的 Thr₂₁₄, 与 SAM 的甲硫氨酸中的氧原子和氮原子形成氢键且后者能形成更多(图 5A、5B)。由于 Mr-PksCT CMeT 能与 SAM 形成更多的氢键, 推测其更容易与 SAM 结合。

再进一步通过分子力学/广义波恩表面积 (molecular mechanics/generalized Born surface

area, MM/GBSA)自由能计算方法预测结合自由能^[54], Mr-PksCT CMeT 与 SAM 的结合自由能为 -45.66 kcal/mol, 小于 PksPT CMeT 与 SAM 的结合自由能 -37.53 kcal/mol, 与上述依据氢键数量推测 Mr-PksCT CMeT 更容易结合 SAM 的结论相符, 也就说明其更容易将 SAM 上的甲基转移到聚酮化合物中间体的 α -C 上。此外, Mr-PksPT CMeT 的底物空腔体积被预测为 809.79 \AA^3 , 深度为 24.01 \AA , 而 Mr-PksCT CMeT 的底物空腔体积为 683.39 \AA^3 , 深度为 21.95 \AA ^[55], Mr-PksPT 的 CMeT 有一个更大和更深的底物空腔, 与能够催化比 PksCT CMeT 更长的聚酮链的事实一致。

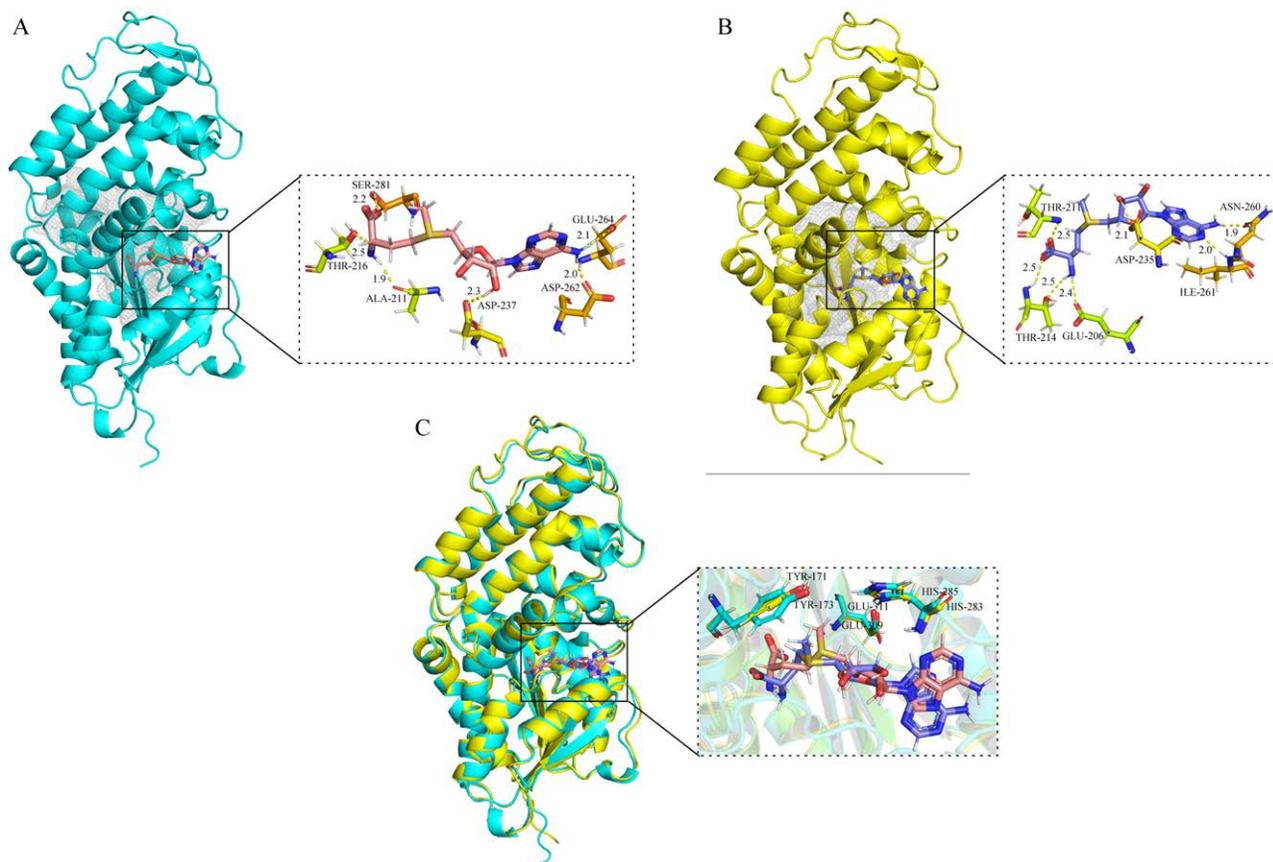


图 5 Mr-PksPT 和 Mr-PksCT 的 CMeTs 与 SAM 的分子对接分析

Figure 5 Molecular docking analysis of Mr-PksPT and Mr-PksCT CMeTs with SAM. A: Molecular docking of Mr-PksPT CMeT to SAM. B: Molecular docking of Mr-PksCT CMeT to SAM. C: Superimposed view of Mr-PksPT and Mr-PksCT CMeTs docked with SAM. SAM are modeled as pink sticks and purple sticks in Mr-PksPT and Mr-PksCT CMeTs, respectively. Substrate cavities are represented by a gray mesh.

进一步将 Mr-PksPT 和 Mr-PksCT 的 CMeT 和 KS 结构域分别与甲基受体底物即 Pks 中间产物进行分子对接, 再通过 MM/GBSA 自由能计算方法预测结合自由能(图 6)。

由图 6B 可知, Mr-PksCT 的 CMeT 与第一轮延伸底物、第二轮延伸底物和第三轮底物的结合自由能分别为 -40.74 、 -46.90 、 -49.77 kcal/mol, 均小于 Mr-PksCT 的 KS 与对应底物的结合自由能。假设若底物与某一结构域的结合自由能更小, 便会优先选择与其结合并发生反应, 则根据假设所得的 Mr-PksCT 的碳甲基化程序如图 6B 所示, 与 Storm

等^[17]提出的 PksCT 的碳甲基化程序一致, 说明此假设较为合理。于是再将此假设用于分析 Mr-PksPT 的碳甲基化程序。由图 6A 可知, Mr-PksPT 的 CMeT 与第四轮延伸后底物的结合自由能为 -47.86 kcal/mol, 小于 Mr-PksPT 的 KS 与对应底物的结合自由能, 而与其余底物的结合自由能均大于 Mr-PksPT 的 KS 与对应底物的结合自由能, 故推测 Mr-PksCT 的唯一一次碳甲基化发生在第四轮延伸后(图 6A), 推测结果与已知的产物结构一致。

基于上述分析得到的碳甲基化程序, 提出了 PksPT 和 PksCT 产生第一个游离产物的过程(图 7)。

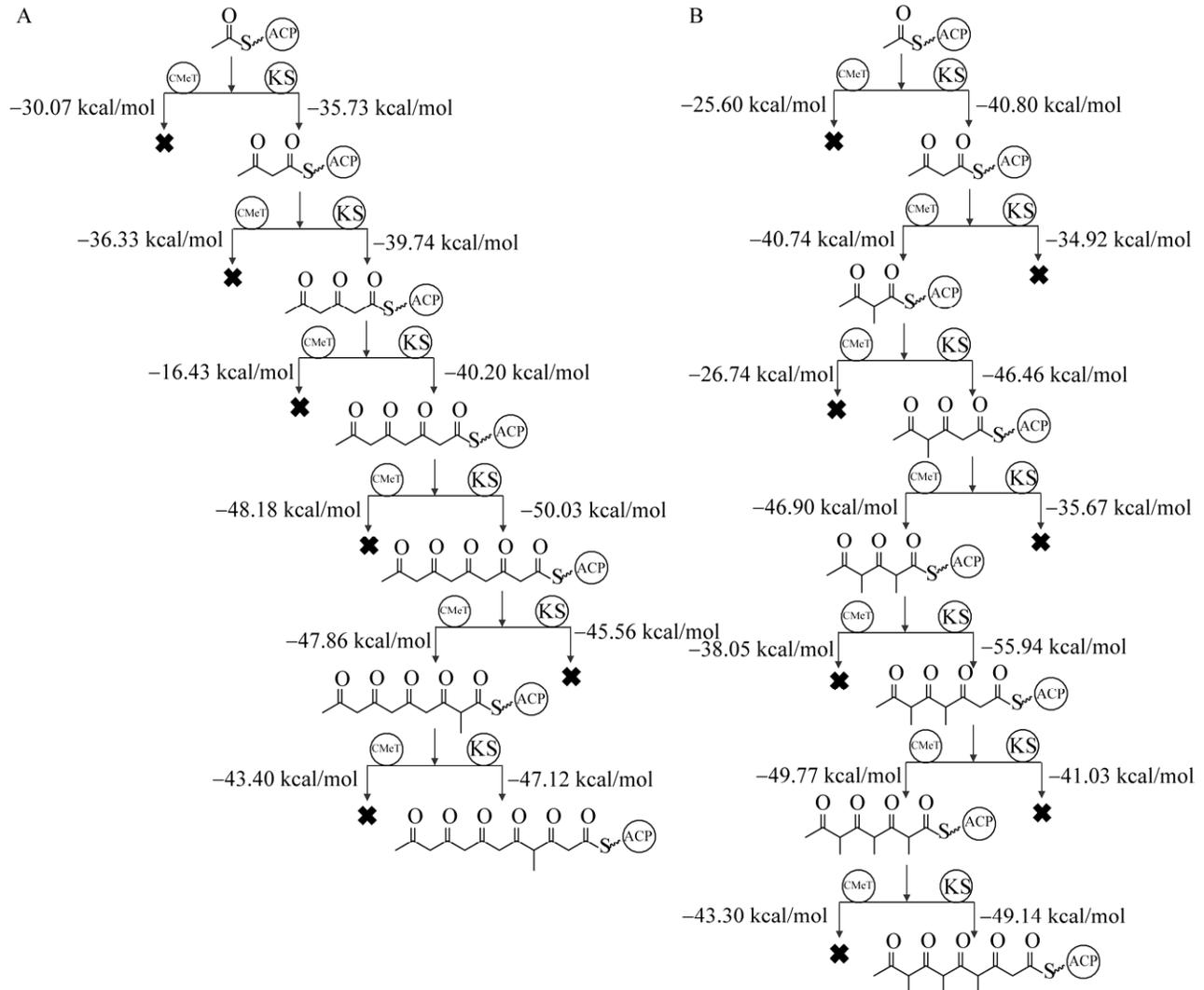


图 6 基于结合自由能的 Mr-PksPT (A)和 Mr-PksCT (B)的碳甲基化程序分析

Figure 6 Analysis of C-methylation programming of Mr-PksPT (A) and Mr-PksCT (B) based on binding free energies.

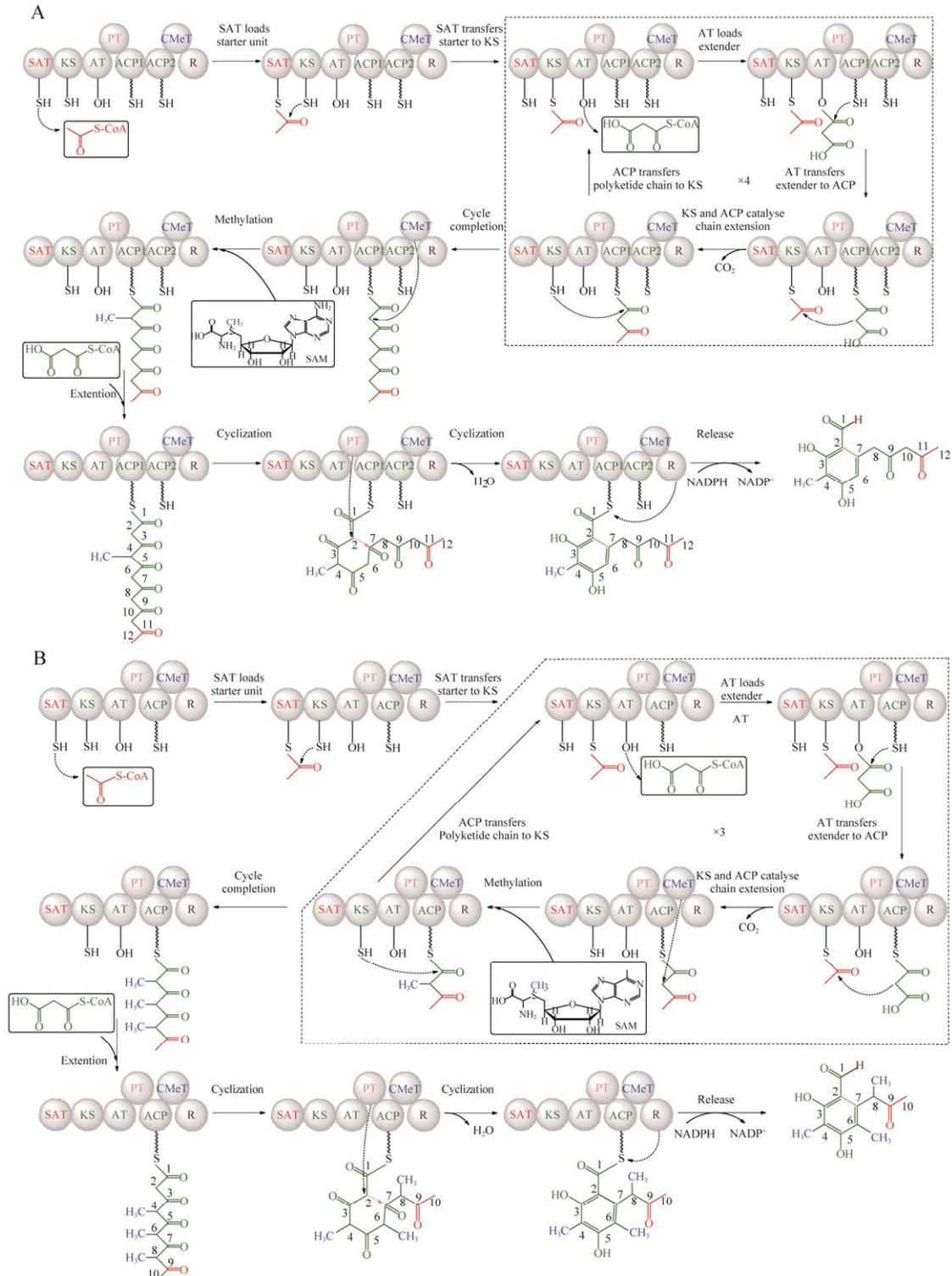


图 7 PksPT (A)和 PksCT (B)产生第一个游离产物苯甲醛化合物的过程

Figure 7 The process of PksPT (A) and PksCT (B) producing the first free product benzaldehyde compound. The solid arrows indicate the specific process and direction of the synthetic process, and the dashed arrows indicate the binding trends and sites of the domains to the substrates. Solid boxes indicate the substrates involved in the reaction, and dashed boxes indicate the cycles of the reaction.

SAT 选择乙酰辅酶 A 作为起始单元, 通过 ACP 转移乙酰基底物到 KS。AT 选择丙二酰辅酶 A 作为延伸单元, 通过 ACP 转移丙二酰底物与 KS 上的乙酰基发生克莱森缩合反应以进行碳链延伸。之后, PksPT 会在 4 轮延伸后进行碳甲基化修饰; 而 PksCT 在前 3 轮延伸后均发生碳甲基化修饰。再分别进行一次延伸, PT 环化, 最后由 R 还原释放。

此外, 由于 Mr-PksPT 和 Mr-PksCT 的 CMeTs 分别对特定底物有显著较小的结合自由能, 推测 CMeT 能特异性识别底物, 而 PksPT 和 PksCT 产物碳甲基化差异可能就是源于各自 CMeT 能特异性识别的底物存在差异^[19,56]。Mr-PksPT 和 Mr-PksCT KS 与合成过程中所有酰基底物的结合自由能都始终稳定在较低水平, CMeT 只有在识别特异性底物时才具有较小的结合自由能且比 KS 与该底物的结合自由能更小, 以方便底物选择与 CMeT 结合并发生反应, 进一步体现了 CMeT 与 KS 在产物合成过程中的竞争关系^[18-19]。

3 讨论与结论

甲基化是天然产物中常见的现象, 但其机理仍不清楚。本文以红色红曲菌 M7 基因组中结构域组成相同, 产物均为 BA 但甲基化程度不同的 2 个 NR-Pks——Mr-PksPT 和 Mr-PksCT 为研究对象, 通过 AlphaFold 2 预测 Pks 及其结构域的三维结构, 并结合分子对接技术和结合自由能分析, 表明 Pks 的 CMeT 结构域通过与 KS 结构域的竞争, 从而影响 Mr-PksPT 和 Mr-PksCT 产物 BA 的甲基化程度。

基于 Mr-PksPT 和 Mr-PksCT 的聚类、结构域及其氨基酸序列保守性及其结构域的三维结构的相似性比较分析发现, 尽管它们的聚类(图 1)、结构域组成(图 2)及其氨基酸序列保守性(表 1)与结构域三维结构的相似性(表 2)等非常相

似, 但是 Mr-PksPT 和 Mr-PksCT 三维结构的 RMSD 值为 12.426 Å (表 2), 远远大于 2 Å, 相似性很差, 表明 Mr-PksPT 和 Mr-PksCT 产生 BA 的甲基化程度差异应该是由 CMeT 参与的结构域互作的差异决定的。随后, 采用分子对接结合自由能分析表明, 影响 Mr-PksPT 和 Mr-PksCT 产物 BA 甲基化程度的是由 Pks 的 CMeT 与 KS 结构域对 Pks 中间产物的竞争性决定的, 并提出了 Mr-PksPT 和 Mr-PksCT 产生 BA 的形成过程(图 7)。

目前, 关于 Pks 的 CMeT 结构域的三维结构信息及其功能都是通过对 CMeT 结构域晶体的信息获得的, 由于蛋白晶体及其结构信息的获得繁琐且困难, 所以截至 2023 年 1 月底在 PDB 蛋白质数据库 PDB 中已公布的 Pks 的 CMeT 晶体结构信息仅为 11 个(表 S1, 见国家微生物科学数据中心, 编号 NMDCX0000225)。最近, 采用 AlphaFold 2 软件预测蛋白质及其结构域的三维结构信息取得了长足进步, 截至 2022 年 12 月采用 AlphaFold 2 已成功预测超过 2 亿个蛋白质三维结构^[21-22], 包括至少 29 240 个甲基转移酶结构域(methyltransferase, MT)的三维结构, 其中许多通过 AlphaFold 2 预测的结构与实验结构的一致性良好且常被用于科学研究(表 S9, 见国家微生物科学数据中心, 编号 NMDCX0000225)^[57-58]。本文在研究过程中, 比较了通过蛋白质结晶获得的来自红曲菌 PksCT 的 CMeT 的三维结构(PDB ID: 5MPT)^[17]与采用 AlphaFold 2 预测的红曲菌 PksCT 的 CMeT 的三维结构结果(见 NMDCX0000225 中的图 S4), 发现两者的 RMSD 值仅为 0.391 Å, 远远小于 2 Å, 表明二者的一致性良好, 这也表明本文采用 AlphaFold 2 预测的 Pks 及其结构域的结果是可靠的。

通过分析 Mr-PksPT 和 Mr-PksCT 的 ACP、KS 和 CMeT 结构域的表面电荷差异发现, 这些

结构域的表面电荷,特别是底物结合口袋入口表面电荷的差异(图 3)影响着 KS 和 CMeT 对 Pks 中间产物的竞争性,从而决定着 Pks 产物 BA 的甲基化程度。为此,尝试这改变 KS 和 CMeT 结构域的底物结合口袋入口处的活性氨基酸组成,以调整表面电荷,从而有望产生不同甲基化程度的 Pks 产物。例如, Mr-PksCT 的 CMeT 中主要参与底物结合口袋入口表面的带电连接的残基包括 Lys₇₈、Lys₂₄₆、Arg₁₆₄, 将这 3 个主要的正电荷残基全部突变为性质相似且不带电残基 Gln。突变后的 CMeT (CMeT_Q)理论上能由于口袋入口表面的正电荷减少,将减少与入口表面带负电荷的 ACP 相互作用的机会,从而改变 C 甲基化程度。于是,进一步对 CMeT_Q 进行分析, Mr-PksCT 的 CMeT_Q 与 CMeT 三维结构的 RMSD 值为 0.434, 表明突变前后的三维结构相似。且 Mr-PksCT 中 CMeT_Q-ACP 互作的结合亲和力为 -10.3 kcal/mol, 与 Mr-PksPT 中 CMeT-ACP 互作的结合亲和力相似, 说明该突变的确能降低 Mr-PksCT 中 CMeT 与 KS 竞争性与 ACP 互作的优势, 以至于降低产物的甲基化效率, 从而可能改变 Mr-PksCT 产物的甲基化程度。

综上所述, 红色红曲菌 M7 中 Mr-PksPT 和 Mr-PksCT 产物 BA 的甲基化程度差异是由 CMeT 结构域与 KS 结构域的竞争造成的。并基于结合自由能分析, 提出了 Mr-PksPT 和 Mr-PksCT 产物 BA 的合成过程。最后提出了通过改造 Mr-PksCT 的 CMeT 从而改变 Mr-PksCT 产物 BA 甲基化程度的建议。本研究结果希望能为 Pks 的碳甲基化研究提供新思路并为设计 Pks 产物的碳甲基化提供方向。

参考文献

- [1] COX RJ. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets[J]. *Organic & Biomolecular Chemistry*, 2007, 5(13): 2010-2026.
- [2] CRAWFORD JM, TOWNSEND CA. New insights into the formation of fungal aromatic polyketides[J]. *Nature Reviews Microbiology*, 2010, 8(12): 879-889.
- [3] CRAWFORD JM, KORMAN TP, LABONTE JW, VAGSTAD AL, HILL EA, KAMARI-BIDKORPEH O, TSAI SC, TOWNSEND CA. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization[J]. *Nature*, 2009, 461(7267): 1139-1143.
- [4] NICHOLSON TP, RUDD BAM, DAWSON M, LAZARUS CM, SIMPSON TJ, COX RJ. Design and utility of oligonucleotide gene probes for fungal polyketide synthases[J]. *Chemistry & Biology*, 2001, 8(2): 157-178.
- [5] CRAWFORD JM, DANCY BCR, HILL EA, UDWARY DW, TOWNSEND CA. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(45): 16728-16733.
- [6] FISCH KM, SKELLAM E, IVISON D, COX RJ, BAILEY AM, LAZARUS CM, SIMPSON TJ. Catalytic role of the C-terminal domains of a fungal non-reducing polyketide synthase[J]. *Chemical Communications*, 2010, 46(29): 5331-5333.
- [7] YEH HH, CHANG SL, CHIANG YM, BRUNO KS, OAKLEY BR, WU TK, WANG CCC. Engineering fungal nonreducing polyketide synthase by heterologous expression and domain swapping[J]. *Organic Letters*, 2013, 15(4): 756-759.
- [8] BALLESTER AR, MARCET-HOUBEN M, LEVIN E, SELA N, SELMA-LÁZARO C, CARMONA L, WISNIEWSKI M, DROBY S, GONZÁLEZ-CANDELAS L, GABALDÓN T. Genome, transcriptome, and functional analyses of *Penicillium expansum* provide new insights into secondary metabolism and pathogenicity[J]. *Molecular Plant-Microbe Interactions: MPMI*, 2015, 28(3): 232-248.
- [9] CHEN WP, CHEN RF, LIU QP, HE Y, HE K, DING XL, KANG LJ, GUO XX, XIE NN, ZHOU YX, LU YY, COX RJ, MOLNÁR I, LI M, SHAO YC, CHEN FS. Orange, red, yellow: biosynthesis of azaphilone pigments in *Monascus* fungi[J]. *Chemical Science*, 2017, 8(7): 4917-4925.

- [10] HE Y, COX RJ. The molecular steps of citrinin biosynthesis in fungi[J]. *Chemical Science*, 2016, 7(3): 2119-2127.
- [11] MAPARI SAS, HANSEN ME, MEYER AS, THRANE U. Computerized screening for novel producers of *Monascus*-like food pigments in *Penicillium* species[J]. *Journal of Agricultural and Food Chemistry*, 2008, 56(21): 9981-9989.
- [12] PAVESI C, FLON V, MANN S, LELEU S, PRADO S, FRANCK X. Biosynthesis of azaphilones: a review[J]. *Natural Product Reports*, 2021, 38(6): 1058-1071.
- [13] WILLIAMS K, GRECO C, BAILEY AM, WILLIS CL. Core steps to the azaphilone family of fungal natural products[J]. *ChemBioChem*, 2021, 22(21): 3027-3036.
- [14] CHEN WP, FENG YL, MOLNÁR I, CHEN FS. Nature and nurture: confluence of pathway determinism with metabolic and chemical serendipity diversifies *Monascus azaphilone* pigments[J]. *Natural Product Reports*, 2019, 36(4): 561-572.
- [15] LISCOMBE DK, LOUIE GV, NOEL JP. Architectures, mechanisms and molecular evolution of natural product methyltransferases[J]. *Natural Product Reports*, 2012, 29(10): 1238-1250.
- [16] LI PW, CHEN M, TANG W, GUO ZY, ZHANG YW, WANG M, HORSMAN GP, ZHONG J, LU ZX, CHEN YH. Initiating polyketide biosynthesis by on-line methyl esterification[J]. *Nature Communications*, 2021, 12: 4499.
- [17] STORM PA, HERBST DA, MAIER T, TOWNSEND CA. Functional and structural analysis of programmed C-methylation in the biosynthesis of the fungal polyketide citrinin[J]. *Cell Chemical Biology*, 2017, 24(3): 316-325.
- [18] STORM PA, PAL P, HUITT-ROEHL CR, TOWNSEND CA. Exploring fungal polyketide C-methylation through combinatorial domain swaps[J]. *ACS Chemical Biology*, 2018, 13(11): 3043-3048.
- [19] CACHO RA, THUSS J, XU W, SANICHAR R, GAO ZZ, NGUYEN A, VEDERAS JC, TANG Y. Understanding programming of fungal iterative polyketide synthases: the biochemical basis for regioselectivity by the methyltransferase domain in the lovastatin megasynthase[J]. *Journal of the American Chemical Society*, 2015, 137(50): 15688-15691.
- [20] THORP HH. Proteins, proteins everywhere[J]. *Science*, 2021, 374(6574): 1415.
- [21] JUMPER J, EVANS R, PRITZEL A, GREEN T, FIGURNOV M, RONNEBERGER O, TUNYASUVUNAKOOL K, BATES R, ŽÍDEK A, POTAPENKO A, BRIDGLAND A, MEYER C, KOHL SAA, BALLARD AJ, COWIE A, ROMERA-PAREDES B, NIKOLOV S, JAIN R, ADLER J, BACK T, et al. Highly accurate protein structure prediction with AlphaFold[J]. *Nature*, 2021, 596(7873): 583-589.
- [22] VARADI M, ANYANGO S, DESHPANDE M, NAIR S, NATASSIA C, YORDANOVA G, YUAN D, STROE O, WOOD G, LAYDON A, ŽÍDEK A, GREEN T, TUNYASUVUNAKOOL K, PETERSEN S, JUMPER J, CLANCY E, GREEN R, VORA A, LUTFI M, FIGURNOV M, et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models[J]. *Nucleic Acids Research*, 2022, 50(D1): D439-D444.
- [23] ALONSO H, BLIZNYUK AA, GREASY JE. Combining docking and molecular dynamic simulations in drug design[J]. *Medicinal Research Reviews*, 2006, 26(5): 531-568.
- [24] FU X, ZAN XY, SUN L, TAN M, CUI FJ, LIANG YY, MENG LJ, SUN WJ. Functional characterization and structural basis of the β -1,3-glucan synthase CMGLS from mushroom *Cordyceps militaris*[J]. *Journal of Agricultural and Food Chemistry*, 2022, 70(28): 8725-8737.
- [25] GAO W, ZHANG Y, CHEN L, LIU XY, LI K, HAN LJ, YU ZW, REN JZ, TANG LF, FAN ZJ. Novel[1,2,4]-triazolo[3,4-*b*]-[1,3,4]thiadizoles as potent pyruvate kinase inhibitors for fungal control[J]. *Journal of Agricultural and Food Chemistry*, 2022, 70(33): 10170-10181.
- [26] KRZEMIŃSKA A, KWIATOS N, ARENHART SOARES F, STEINBÜCHEL A. Theoretical studies of cyanophycin dipeptides as inhibitors of tyrosinases[J]. *International Journal of Molecular Sciences*, 2022, 23(6): 3335.
- [27] ZHAI GQ, ZHANG ZY, DONG CJ. Mutagenesis and functional analysis of SotB: a multidrug transporter of the major facilitator superfamily from *Escherichia coli*[J]. *Frontiers in Microbiology*, 2022, 13: 1024639.
- [28] KUMAR S, STECHER G, TAMURA K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets[J]. *Molecular Biology and Evolution*, 2016, 33(7): 1870-1874.

- [29] PAYSAN-LAFOSSE T, BLUM M, CHUGURANSKY S, GREGO T, PINTO BL, SALAZAR GA, BILESCHI ML, BORK P, BRIDGE A, COLWELL L, GOUGH J, HAFT DH, LETUNIĆ I, MARCHLER-BAUER A, MI HY, NATALE DA, ORENGO CA, PANDURANGAN AP, RIVOIRE C, SIGRIST CJA, et al. InterPro in 2022[J]. *Nucleic Acids Research*, 2023, 51(D1): D418-D427.
- [30] KEARSE M, MOIR R, WILSON A, STONES-HAVAS S, CHEUNG M, STURROCK S, BUXTON S, COOPER A, MARKOWITZ S, DURAN C, THIERER T, ASHTON B, MEINTJES P, DRUMMOND A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data[J]. *Bioinformatics*, 2012, 28(12): 1647-1649.
- [31] PIERCE BG, WIEHE K, HWANG H, KIM BH, VREVEN T, WENG ZP. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers[J]. *Bioinformatics*, 2014, 30(12): 1771-1773.
- [32] VANGONE A, BONVIN AM. Contacts-based prediction of binding affinity in protein-protein complexes[J]. *eLife*, 2015, 4: e07454.
- [33] WOO PCY, LAM CW, TAM EWT, LEE KC, YUNG KKY, LEUNG CKF, SZE KH, LAU SKP, YUEN KY. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in *Penicillium marneffei*[J]. *Scientific Reports*, 2014, 4: 6728.
- [34] SHARMA D, GUPTA C, AGGARWAL S, NAGPAL N. Pigment extraction from fungus for textile dyeing[J]. *Indian Journal of Fibre & Textile Research*, 2012, 37(1): 68-73.
- [35] CAMARDO LEGGIERI M, DECONTARDI S, BERTUZZI T, PIETRI A, BATTILANI P. Modeling growth and toxin production of toxigenic fungi signaled in cheese under different temperature and water activity regimes[J]. *Toxins*, 2016, 9(1): 4.
- [36] SHIMIZU T, KINOSHITA H, ISHIHARA S, SAKAI K, NAGAI S, NIHIRA T. Polyketide synthase gene responsible for citrinin biosynthesis in *Monascus purpureus*[J]. *Applied and Environmental Microbiology*, 2005, 71(7): 3453-3457.
- [37] NEWMAN AG, VAGSTAD AL, STORM PA, TOWNSEND CA. Systematic domain swaps of iterative, nonreducing polyketide synthases provide a mechanistic understanding and rationale for catalytic reprogramming[J]. *Journal of the American Chemical Society*, 2014, 136(20): 7348-7362.
- [38] WINTER JM, CASCIO D, DIETRICH D, SATO M, WATANABE K, SAWAYA MR, VEDERAS JC, TANG Y. Biochemical and structural basis for controlling chemical modularity in fungal polyketide biosynthesis[J]. *Journal of the American Chemical Society*, 2015, 137(31): 9885-9893.
- [39] BHETARIYA PJ, PRAJAPATI M, BHADURI A, MANDAL RS, VARMA A, MADAN T, SINGH Y, SARMA PU. Phylogenetic and structural analysis of polyketide synthases in *Aspergilli*[J]. *Evolutionary Bioinformatics Online*, 2016, 12: 109-119.
- [40] POUST S, YOON I, ADAMS PD, KATZ L, PETZOLD CJ, KEASLING JD. Understanding the role of histidine in the GHSxG acyltransferase active site motif: evidence for histidine stabilization of the malonyl-enzyme intermediate[J]. *PLoS One*, 2014, 9(10): e109421.
- [41] LIU L, ZHANG Z, SHAO CL, WANG JL, BAI H, WANG CY. Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains[J]. *Scientific Reports*, 2015, 5: 10463.
- [42] WATTANA-AMORN P, WILLIAMS C, PŁOSKOŃ E, COX RJ, SIMPSON TJ, CROSBY J, CRUMP MP. Solution structure of an acyl carrier protein domain from a fungal type I polyketide synthase[J]. *Biochemistry*, 2010, 49(10): 2186-2193.
- [43] WEISSMAN KJ, HONG H, POPOVIC B, MEERSMAN F. Evidence for a protein-protein interaction motif on an acyl carrier protein domain from a modular polyketide synthase[J]. *Chemistry & Biology*, 2006, 13(6): 625-636.
- [44] KOZBIAL PZ, MUSHEGIAN AR. Natural history of S-adenosylmethionine-binding proteins[J]. *BMC Structural Biology*, 2005, 5(1): 1-26.
- [45] LIU H, BEGLEY T. *Comprehensive Natural Products III: Chemistry and Biology*[M]. Third Edition. United States of America: Elsevier, 2020.
- [46] GAITATZIS N, KUNZE B, MÜLLER R. *In vitro* reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2001, 98(20): 11136-11141.

- [47] STARK A, SUNYAEV S, RUSSELL RB. A model for statistical significance of local similarities in structure[J]. *Journal of Molecular Biology*, 2003, 326(5): 1307-1316.
- [48] YANG XL, FRIEDRICH S, YIN S, PIECH O, WILLIAMS K, SIMPSON TJ, COX RJ. Molecular basis of methylation and chain-length programming in a fungal iterative highly reducing polyketide synthase[J]. *Chemical Science*, 2019, 10(36): 8478-8489.
- [49] KASTRITIS PL, RODRIGUES JPGLM, FOLKERS GE, BOELENS R, BONVIN AMJJ. Proteins feel more than they see: fine-tuning of binding affinity by properties of the non-interacting surface[J]. *Journal of Molecular Biology*, 2014, 426(14): 2632-2652.
- [50] JANIN J, HENRICK K, MOULT J, TEN EYCK L, STERNBERG MJE, VAJDA S, VAKSER I, WODAK SJ. CAPRI: a critical assessment of predicted interactions[J]. *Proteins: Structure, Function, and Genetics*, 2003, 52(1): 2-9.
- [51] BLUNDELL TL, FERNÁNDEZ-RECIO J. Brief encounters bolster contacts[J]. *Nature*, 2006, 444(7117): 279-280.
- [52] MORRIS CJ, DELLA CORTE D. Using molecular docking and molecular dynamics to investigate protein-ligand interactions[J]. *Modern Physics Letters B*, 2021, 35(8): 2130002.
- [53] PENG Y, SARTINI D, POZZI V, WILK D, EMANUELLI M, YEE VC. Structural basis of substrate recognition in human nicotinamide N-methyltransferase[J]. *Biochemistry*, 2011, 50(36): 7800-7808.
- [54] GENHEDEN S, RYDE U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities[J]. *Expert Opinion on Drug Discovery*, 2015, 10(5): 449-461.
- [55] SCHÖNING-STIERAND K, DIEDRICH K, FÄHRROLFES R, FLACHSENBERG F, MEYDER A, NITTINGER E, STEINEGGER R, RAREY M. ProteinsPlus: interactive analysis of protein-ligand binding interfaces[J]. *Nucleic Acids Research*, 2020, 48(W1): W48-W53.
- [56] KISHIMOTO S, TSUNEMATSU Y, MATSUSHITA T, HARA K, HASHIMOTO H, TANG Y, WATANABE K. Functional and structural analyses of *trans* C-methyltransferase in fungal polyketide biosynthesis[J]. *Biochemistry*, 2019, 58(38): 3933-3937.
- [57] YAN JJ, LIU FQ, GUAN ZY, YAN XH, JIN XH, WANG Q, WANG ZC, YAN JJ, ZHANG DL, LIU Z, WU S, YIN P. Structural insights into DNA N6-adenine methylation by the MTA1 complex[J]. *Cell Discovery*, 2023, 9: 8.
- [58] ZHANG YX, ALSHAMMARI E, SOBOTA J, YANG A, LI CY, YANG Z. Unique SMYD5 structure revealed by AlphaFold correlates with its functional divergence[J]. *Biomolecules*, 2022, 12(6): 783.