

## 植物天门冬酰胺连接酶在大环肽类药物应用中的研究进展

申 欣, 刘恣之, 杨 燕, 王 伟\*

(中国医学科学院、北京协和医学院药物研究所, 天然药物活性物质与功能国家重点实验室/国家卫生健康委员会天然药物生物合成重点实验室, 北京 100050)

**摘要:** 植物中的天门冬酰胺内肽酶 (asparaginyl endopeptidases, AEPs) 属于半胱氨酸蛋白酶家族, 以酶原的形式在酸性液泡中发生自激活, 在种子储存蛋白成熟、蛋白质降解、细胞程序性死亡和宿主防御中发挥着重要的生理作用。最近从几种富含环肽的植物中分离鉴定了促进环肽成熟的生物加工酶-天门冬酰胺连接酶 (peptidyl Asx-specific ligases, PALs), 由 AEPs 进化而来, 可位点特异性催化天冬酰胺或天冬氨酸肽键的生成, 由于相对无痕和宽泛的底物谱以及催化效率高等优势, 已经在多肽和蛋白质的环化和修饰方面起到重要作用, 是提高多肽类药物稳定性的有力工具。本文介绍了 AEPs 在植物中的生理功能, 综述了 PALs 的发现、结构特点、催化机制和蛋白质工程, 以及限制其应用的原因和未来的发展趋势。此外, 还重点阐述了 PALs 在环肽生物合成以及在大环肽类药物开发中的应用, 旨在为生物催化合成环多肽类药物提供一种新思路。

**关键词:** 天门冬酰胺内肽酶; 天门冬酰胺连接酶; 大环肽; 酶促环化; 位点特异性连接

**中图分类号:** R931      **文献标识码:** A      **文章编号:** 0513-4870(2023)09-2656-13

## Advances in peptidyl Asx-specific ligases for the application of cyclic peptides

SHEN Xin, LIU Min-zhi, YANG Yan, WANG Wei\*

(State Key Laboratory of Bioactive Substance and Function of Natural Medicines, NHC Key Laboratory of Natural Drug Biosynthesis, Institute of Materia Medica, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100050, China)

**Abstract:** Asparaginyl endopeptidases (AEPs) in plants belong to the family of cysteine protease that undergo self-activation in the form of zymogen in acidic vacuole and play important physiological roles in maturation of seed storage proteins, protein degradation, programmed cell death and host defense. Bioprocessing enzymes (peptidyl Asx-specific ligases, PALs) that promote the maturation of cyclotides have recently been isolated and identified from several cyclotide-rich plants. PALs derived from AEPs can site-specifically catalyze the formation of asparagine or aspartate peptide bonds. Due to the advantages of relatively traceless peptide bonds and broad substrate spectrum and high catalytic efficiency, they have been playing important roles in the cyclization and modification of peptides and proteins, and are powerful tools for improving the stability of peptide drugs. This review describes the physiological functions of AEPs in plants and summarizes the discoveries, structural characteristics, catalytic mechanism and protein engineering of PALs, as well as the limitation of their applications and future trends. In addition, the applications of PALs in cyclotides biosynthesis and the development of macrocyclic peptides are highlighted, with the aim of providing a new idea for the biocatalytic synthesis of cyclic

收稿日期: 2023-03-22; 修回日期: 2023-05-05.

基金项目: 国家自然科学基金项目 (82073757, 81903487); 中国医学科学院医学与健康科技创新工程项目 (2021-I2M-1-029); 北京协和医学院合成生物学学科建设专项资助 (201920100801).

\*通讯作者 Tel / Fax: 86-10-63165196, E-mail: wwang@imm.ac.cn

DOI: 10.16438/j.0513-4870.2023-0341

peptides.

**Key words:** asparaginyl endopeptidase; peptidyl Asx-specific ligase; cyclic peptide; enzymatic cyclization; site-specific ligation

多肽类药物被用于治疗多种人类疾病,包括代谢性疾病、肿瘤、传染病、中枢神经系统疾病、自身免疫性疾病等,其中代谢性疾病占比较多,目前美国FDA已批准超过80种多肽药物上市,还有数百种多肽处于临床前或临床试验阶段<sup>[1]</sup>。多肽具有对靶标的高选择性和高亲和力且副作用低,逐渐成为比小分子药物更具吸引力的治疗分子,但却受限于稳定性差和口服生物利用度低,临床应用的治疗性肽类药物的给药途径多以皮下和静脉注射为主,使患者的用药依从性降低<sup>[2,3]</sup>。

目前获批的多肽类药物有40种都是大环肽类(cyclic peptides)药物,例如环孢菌素A(cyclosporine A, CsA)是环状亲脂性的11肽,其胶囊剂(sandimmune)于1990年被美国FDA批准上市,用于预防器官移植排斥的免疫反应,口服生物利用度高达25%~30%。因此,源于天然大环肽结构稳定的启示,多肽环化是一个提高其分子稳定性的技术策略<sup>[4]</sup>。天然多肽环状骨架的形成降低了分子的灵活性、约束了分子构象、还提高了与其靶蛋白结合的亲和力和减少了非特异性结合。此外,大环肽的刚性构象不仅耐受蛋白酶降解,还具有与脂质膜结合相互作用的熵优势,并且增加的稳定性还可以改善多肽在生物体内活性的持久性,提高其药代动力学和药效动力学<sup>[5]</sup>。

临床上批准的绝大多数大环肽均源于天然多肽,基于合理设计和优化从头开发大环肽类药物,解决胞内“不可成药”的靶点或者将具有生物活性的多肽表位嫁接到刚性的环肽骨架上,可以增加线性肽对人体内消化液的抵抗力,从而提高多肽的稳定性,这使大环肽类药物作为治疗剂或是作为多肽分子支架的开发充满了机遇和挑战,但是却受限于环化酶的发现和应用<sup>[6,7]</sup>。蛋白水解酶普遍存在于自然界,而肽连接酶却是稀有品种,目前已知的肽连接酶有PATG<sup>[8]</sup>、PCY1<sup>[9]</sup>、POBP<sup>[10]</sup>、sortase A<sup>[11]</sup>、subtiligase<sup>[12]</sup>等。前3种酶是底物特异性的环化酶,环化肽的大小在5~11个氨基酸,且线性肽前体来源于蓝藻、植物和蘑菇,天然底物谱非常有限。Sortase A的识别序列较长,也在一定程度上限制了其应用。Subtiligase的突变体peptiligase克服了对缓冲液中钙离子的依赖性,更高效的omniligase-1突变体已经可商业购买,高效地应用于制备克级的环肽,但依然受限于较窄的底物谱<sup>[13]</sup>。从植物中发现的天门冬酰胺连接酶(PALs)具有无可比拟的优势,不仅具有

高的催化效率且具有宽泛的底物谱,是目前最有潜力的多肽环化工具酶之一,可用于多肽的环化以及蛋白质修饰等生物与医药技术领域。

## 1 AEPs

### 1.1 天门冬酰胺内肽酶(AEPs)在植物中的生理功能

早在1960年就在植物中发现广泛存在着一类豆类蛋白酶(legumains),而后被证实与20世纪80年代后期发现的液泡加工酶(vacuolar processing enzymes, VPEs)属于同一家族,它们具有相似的结构,是组织蛋白酶家族(cathepsin D)半胱氨酸蛋白酶C13亚家族(EC 3.4.22.34)的成员。而从功能上讲,它们属于AEPs,可水解Asn/Asp(Asx)后的肽键<sup>[14]</sup>。

植物AEPs前体基因编码N末端信号肽(signal peptide, SP)、N末端前肽(N-terminal propeptide, NTPP)、核心结构域(core domain, CD)、柔性连接区域(linker)及C末端前肽(C-terminal propeptide, CTPP)。翻译后的新生多肽由SP引导到内质网(endoplasmic reticulum, ER),在ER中经历折叠处理步骤-信号肽断裂和二硫键形成,在高尔基体上进行糖基化修饰和进一步的折叠和加工形成正确折叠的AEPs酶原前体,无活性的酶原前体被靶向运输到液泡,在液泡中经酸性条件(pH值为4.0~5.5)诱导的构象变化,进而暴露活性催化位点后连续水解去除较大的CTPP和短的NTPP<sup>[15]</sup>。

成熟后的AEPs负责液泡中常驻蛋白(如蛋白酶、转化酶、储存蛋白和植物防御肽)的蛋白水解和成熟。在植物程序性细胞死亡期间,液泡会塌陷并释放AEPs,从而激活非液泡驻留底物而发挥作用<sup>[16]</sup>。

### 1.2 AEPs介导植物中环肽的生物合成

1970年Gran<sup>[17]</sup>从非洲民族药茶(*Oldenlandia affinis*)中分离了第一个由29个氨基酸组成且具有生物活性的原型kalata B1环肽分子,非洲当地人口服用于分娩催产,表明这种环肽对高温、酸性条件和消化酶具有耐受性。Craik等<sup>[18]</sup>证实了异常稳定的kalata B1具有独特的三维拓扑结构,于1999年环肽(cyclotides)首次被定义为独特的植物防御肽家族,一般由28~37个氨基酸组成,主要结构特征为从头到尾的环肽骨架和三个保守的二硫键组成的胱氨酸结,其中两个二硫键及其连接的主链在结构中形成了一个嵌入的环,该环被第三个二硫键穿过,这种结构被称为环状胱氨

酸结 (cyclic cystine knot, CCK) 基序。

植物中的环肽主要分布在堇菜科 (Violaceae)、茜草科 (Rubiaceae)、豆科 (Fabaceae)、茄科 (Solanaceae)、葫芦科 (Cucurbitaceae) 和禾本科 (Poaceae), 并显示出多种生物活性特性, 例如抗菌、杀虫、蛋白酶抑制、细胞毒性、抗 HIV 和溶血等。环肽大小 (~3 000 Da) 处于小分子药物 (<500 Da) 和生物制剂药物 (>5 000 Da) 之间, 兼有小分子的稳定性和细胞膜渗透能力, 又具有生物制剂药物的高靶向特异性、良好的结合亲和力、低毒性及潜在的口服生物利用度, 还具有靶向胞内“不可成药”靶点介导蛋白质相互作用的优势且生产成本适中。环肽由于其出色的稳定性、序列修饰的耐受性和易于嫁接的特性, 可被用作治疗性肽工程的支架, 在药物设计和医学领域上具有广泛的潜在应用, 然而植物中尚未完全阐明的环肽生物合成途径, 限制了环肽的异源生产<sup>[19-21]</sup>。

环肽是核糖体合成和翻译后修饰肽 (ribosomally synthesized and posttranslationally modified peptides, RiPPs), 环肽前体基因通常包含 SP、NTPP、N 末端重复序列 (N-terminal repeat, NTR)、环肽结构域 (cyclotide domain, CD)、C 末端重复序列 (C-terminal repeat, CTR) 和 CTPP, 其中 NTR 和 CTR 仅在部分环肽前体蛋白中出现<sup>[22]</sup>。植物细胞中环肽的生物加工步骤 (图 1), 包括: ① 环肽前体基因由 DNA 转录为 RNA 并翻译成为前体肽; ② 由 SP 靶向运输到 ER 中, 去除 SP 并由蛋白质二硫键异构酶 (protein-disulfide isomerases, PDI) 辅助折叠再分泌到高尔基体中进行进一步加工; ③ 由 NTR 或 NTPP 介导靶向运输到液泡中; ④ 液泡中未知的 N 末端加工酶切割以释放 NTR 并暴露环肽 N 末端

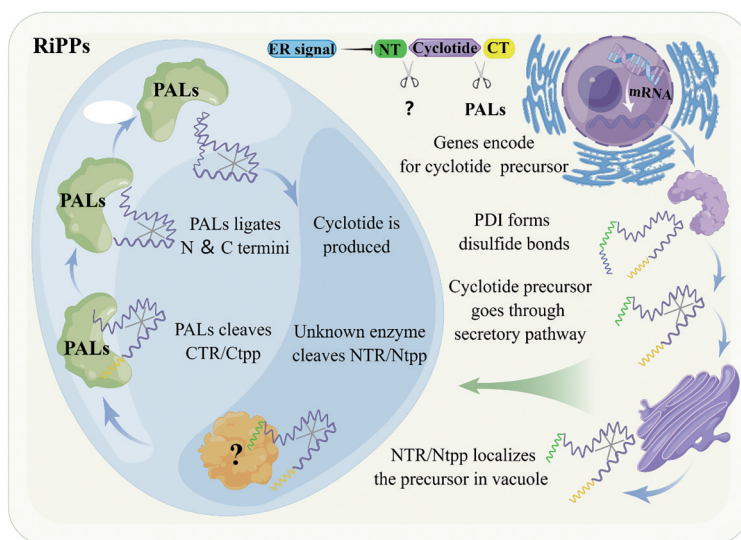
的小氨基酸残基; ⑤ 液泡中自催化成熟的 AEPs 加工切割环肽 C 末端的 CTR 或 CTPP 形成酰基酶中间体; ⑥ N 末端暴露的小氨基酸残基作为亲核试剂攻击酰基酶中间体, AEPs 再生, 形成环肽终产物<sup>[23-26]</sup>。

由于环肽前体序列 N 末端的加工位点不同, 植物中可能存在不同的蛋白酶。最近的一项研究表征了参与 kalata B1 环肽前体 N 末端裂解的生物加工酶 OaRD21A, 属于木瓜蛋白酶样半胱氨酸蛋白酶家族的特定成员 kalatase A。由于胰蛋白酶抑制剂家族环肽在 N 末端位点具有保守的 Asn, 因此 AEPs 已被证实参与了胰蛋白酶抑制剂家族环肽的 N 末端加工过程<sup>[27]</sup>。

### 1.3 AEPs 的分类

在植物中, AEPs 主要发挥蛋白酶的作用, 在液泡中水解储存蛋白使其成熟。少数 AEPs 进化为 PALs, 在环肽的生物合成途径中起到环化作用, 催化肽键的形成。部分 AEPs 还参与了豆科植物刀豆凝集素 A (concanavalin A, ConA) 和胰蛋白酶抑制剂家族环肽的成熟过程, 涉及到剪接 (切割和连接) 机制<sup>[28]</sup>。

Hemu 等<sup>[29]</sup>最新的研究表明, 通过对 1 500 种豆类植物 (Legumains) 综合生物信息学和功能表征分析, 可以将 AEPs 按功能细分为 3 种类型, 分别是占比 88% 的 AEP-like 蛋白酶, 在较为宽泛的 pH 值范围内都起到高效的水解作用; 占比 11% 的 hybrid 剪接酶, 又称非典型 AEP 连接酶, 在酸性 pH 值条件下起水解作用, 在中性 pH 值条件下起连接作用; 占比 1% 的 PAL-like 连接酶, 又称典型 AEP 连接酶, 在较宽泛的 pH 值条件下都起到高效的连接作用。值得注意的是 PALs 在酸性条件下仍然具有十分缓慢的水解作用, 这既是分子进化的结果, 也是其通过自激活形成成熟的 PALs 所必需的。



**Figure 1** Schematic representation of the major steps thought to occur in cyclotide processing inside the plant cell (Figure 1 was created with Figdraw)<sup>[23]</sup>

## 2 PALs

### 2.1 PALs发现

植物AEPs的蛋白酶功能早在20世纪就被表征,随着环肽结构于20世纪末被表征后,直到21世纪初才被推测其参与环肽的生物合成<sup>[23]</sup>。PALs具有宽泛的底物特异性和高催化效率,一般在接近中性pH值条件下可以有效地环化或连接大小从10到>200个氨基酸残基且C端含Asx(Asn/Asp)的各种肽或蛋白质。PALs是特异性地识别C端三肽基序Asx-Xaa-Yaa(P1-P1'-P2',基于Schechter和Berger命名法<sup>[30]</sup>),通常在P1'位置,底物的氨基酸残基要求并不严格,通常偏好小氨基酸残基,例如His、Gly或Ser;在P2'处,通常偏好具有较大侧链的疏水氨基酸残基,例如Val、Leu或Ile。天然环肽前体C端序列与PALs对底物C端识别序列的偏好性一致,表明这是酶和环肽底物共同进化的结果。PALs识别N端氨基酸残基序列具有宽泛性,除Pro外的所有残基都在P1"位置被接受,但对相邻的P2"残基表现出一定的选择性,对于butelase-1优选Cys、Ile、Leu或Val。虽然没有经过全面评估,但这些规则一般也适用于其他PALs<sup>[14]</sup>。总之,PAL型连接酶具有不同偏好性的C端三肽识别序列、底物特异性和最适的连接反应条件,下文主要选取具有代表性且催化效率较高的PALs进行综述。

2014年,Tam团队<sup>[31]</sup>率先从豆科产Cter M环肽的蝴蝶豌豆(*Clitoria ternatea*)豆荚中<sup>[32]</sup>,利用活性导向和四步色谱法的策略分离得到了第一个可以催化肽键形成的PALs,由于*C. ternatea*在当地又被称为“*Bunga telang*”,所以被命名为butelase-1。Butelase-1的底物C端识别序列偏好NHV基序,具有极高的催化效率,通常使用底物与酶的比例为100~1 000:1,典型的环化反应在5~30 min内完成,连接反应的温度为25~42 °C,pH值在5.0~7.0内,暴露于pH > 7.5的条件下会使酶变性。Butelase-1具有pH依赖性的热稳定性,这与其pH依赖性的催化活性相关,即pH影响反应速率和环化/水解比率,所以存在最佳条件从25 °C的pH = 7.0到42 °C的pH = 6.0的变化。当以酶原前体形式存在时将比活性形式更稳定且对pH的敏感性更低。通常当缓冲液中的盐浓度>200 mmol·L<sup>-1</sup>时,butelase-1分子不稳定导致活性降低,但通过去除NaCl可以恢复酶活性,所以最终反应缓冲液中的总盐离子浓度应<100 mmol·L<sup>-1</sup><sup>[33]</sup>。

2015年Harris等<sup>[34]</sup>从茜草科产kalata B1环肽的植物*O. affinis*中鉴定了OaAEP1b,它与OaAEP1相比有一个氨基酸的突变。OaAEP1b(E371V)底物C端识别序列偏好NGL基序,但催化活性比butelase-1慢500

倍,多肽环化时间长达22 h,而butelase-1催化仅需要5 min。2017年Yang等<sup>[35]</sup>通过蛋白质工程改造获得了效率提高160倍的OaAEP1-C247A,其催化效率接近butelase-1。OaAEP1-C247A的反应温度为37 °C,最佳pH值在6.0~6.5内。

2019年Hemu等<sup>[36]</sup>从盛产环肽的堇菜科<sup>[37,38]</sup>植物的*Viola yedoensis*(Vy)和*V. canadensis*(Vc)中鉴定了VyPAL2和VcAEP,其中VcAEP是蛋白酶。VyPAL2的最佳连接酶活性温度为37 °C,且在pH值为5.5~8.0时不产生任何水解产物,在最佳pH = 6.5时,观察到超过80%的环化产率。VyPAL2在P1'位置偏好小氨基酸残基,尤其是Gly或Ser,但Pro除外,在P2'位置偏好疏水性或芳香族残基的存在,如Leu、Ile或Phe。VyPAL2环化效率中等,是butelase-1的28.5%。2021年Chen等<sup>[39]</sup>通过蛋白质工程改造,获得了高催化效率的VcAEP双突变体,起到高效的连接酶作用,与butelase-1等效。

剪接酶在酸性条件下具有水解功能,在中性具有连接酶作用,在酸性到中性的合适pH条件下具有剪接作用,可以催化环肽前体生成环肽。2020年Du等<sup>[40]</sup>从葫芦科产MCoTI-II(*Momordica cochinchinensis* trypsin inhibitor)环肽的植物*Momordica cochinchinensis*中发现了双功能的MCoAEP2,其发挥剪接作用。MCoAEP2是MCoTI-II天然环肽底物的生物加工酶,首先催化MCoTI-II环肽前体完成N端Asn的肽键水解,再介导C端Asp的水解从而发生转肽反应,生成MCOTI-II环肽产物,是目前催化环肽生产最快的剪接酶。MCoAEP2催化环化反应生成的最佳pH值为5.0~6.0,反应温度为22 °C,对于C端是Asp的底物最佳pH值为5,而对C端是Asn的底物最佳pH值是6.0,且在中性pH值条件下对于C端是Asn的底物催化效率更高。2021年Liew等<sup>[41]</sup>发现的McPAL1起到同样的剪接功能,研究表明重组表达的McPAL1在pH = 5~6的条件下,首先催化MCoTI-II环肽前体N端Asn位点的水解,然后完成C端Asp位点的转肽反应,介导MCoTI-I/II的生物合成。McPAL1具有pH依赖性的三峰酶谱特征,该机制涉及pH的条件以及P1残基是Asn还是Asp,赋予了其作为剪接酶,还是水解酶或连接酶的多功能性。在pH值为4~6时,McPAL1选择性催化Asp连接和Asn水解,但在pH值为6.5~8时,Asn连接占主导地位,连接反应最佳pH值为7。大多数PALs在中性至碱性的pH条件下不稳定,在pH > 8时会不可逆失活。而McPAL1是一种异常稳定的酶,耐热和高pH值,在pH = 9的极端条件下仍然具有催化活性,与butelase-1(pI = 4.6)和VyPAL2(pI = 4.5)相比,活性形式的McPAL1具有相对较高的pI为6.0,存在更多带正

电荷的残基和更少带负电荷的残基,即降低了酶表面残基的负/正电荷比就可以提高酶在碱性pH条件下的耐受性。

2023年Hemu等<sup>[42]</sup>通过生物信息学分析了来自大量植物超过1500种legumains的序列,通过序列比对获得了legumains保守的共有序列,人工设计了conLEG蛋白酶,再通过与PALs连接酶的氨基酸序列和结构进行对比分析,利用蛋白质工程获得了conLEG的双突变体conPAL3。conPAL3介导连接反应的最适pH值为7.0,最适反应温度为25℃,这与其他PALs略有不同。conPAL3是高效的连接酶,催化效率是butelase-1的2倍且其更具热稳定性,对于pH的耐受性也更强,环化反应的pH值可以高达8,这就极大拓宽了环化反应的底物谱,对于一些在低pH值条件下不稳定的多肽和蛋白来说,conPALs催化其环化具有了可行性。

表1列举了具有代表性且催化效率较高的PALs。

## 2.2 PALs结构

PALs的催化中心是Cys,属于半胱氨酸蛋白酶C13亚家族,具有保守的“Asn-His-Cys”催化三联体特征<sup>[43]</sup>。PALs酶表面通常包括四个底物结合域,即S1、S2、S1'和S2',它们可以特异性地与多肽底物残基(Pn)相互作用<sup>[30]</sup>。

尽管PALs是一种有效的环化酶,但在自然系统中却仍然无法克服自我加工进化的限制,尚未完全特化为大环化酶,在酸性pH条件下仍然具有低水平的蛋白酶活性,以发生自激活的功能。PALs在植物体内首先表达为非活性酶原,结构含有内质网SP、NTPP、CD和CTPP。CTPP由5个 $\alpha$ 螺旋形成帽状结构(cap domain)并通过长而灵活的链接肽段(linker)与CD相连接,CD由6个 $\beta$ 折叠片层构成,其外围有5个 $\alpha$ 螺旋包围。CTPP完全覆盖了催化裂隙,需要释放CTPP催化肽或蛋白质底物反应,活性位点附近的底物结合域由2个反向平行的 $\beta$ 折叠组成(图2)<sup>[29]</sup>。

CTPP具有稳定和调节核心结构域活性的重要作用,可增加酶原的稳定性和溶解度,其与CD的残基之间产生多处相互作用使得酶原处于非活性状态,其中

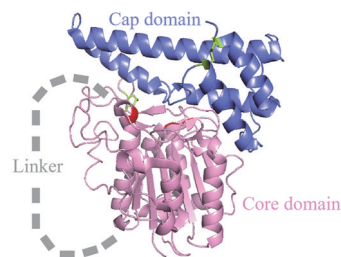


Figure 2 Crystal structure of OaAEP1 in its zymogenic form (PDB code:5H0I)

$\alpha$ 6螺旋中的一个保守残基Gln深入到S1口袋与周围的活性位点残基建立了极性相互作用。晶体结构表明,酶原分子的两个亚基之间的CTPP通过氢键和盐桥等分子间相互作用起到稳定作用,即PALs在中性条件下会形成无活性的同源二聚体或四聚体。在酸性条件下,CTPP的残基质子化,破坏了与核心区域之间的静电相互作用,暴露酶活性催化位点,使得PALs发生顺式水解作用达到自激活的目的,依次去除CTPP和NTPP。活性PALs在酸性条件下还可以介导其他酶原发生反式激活,从而在植物体内起到一些重要的生理调节作用。这种自动激活步骤发生在植物液泡中,pH值约为4.0~5.5,在体外AEP活化通常在pH4~5时进行<sup>[44]</sup>。值得注意的是,当pH值高于4.5时其激活步骤是可逆的,即两个结构域保持完整并且切割后在空间上接近时,就可以发生CTPP和CD的重新结合。大量研究表明,PALs核心结构域两端的水解激活位点存在异质性,NTPP和linker区域存在多处Asn/Asp切割位点,在多个位点处进行自加工,即产生了活性酶的混合物<sup>[45]</sup>。

## 2.3 PALs催化机制

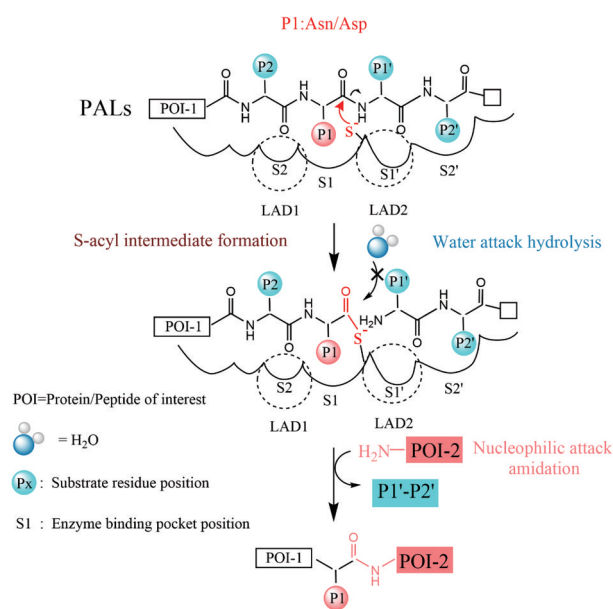
PALs连接酶由AEPs蛋白酶进化而来,二者具有相似的氨基酸序列和结构特征,以催化植物产生环肽所需的转肽或连接反应。PALs是植物中天然存在的强力胶,被宿主用于防御病原体的侵入,其机制是ATP非依赖性的,催化的连接反应不受遗传密码或机制的限制。

研究表明,<sup>18</sup>O同位素标记的H<sub>2</sub>O在PALs介导的水解反应产物中存在,但在连接或环化反应产物中不

Table 1 Important features of peptidyl Asx-specific ligases (PALs) for peptides backbone cyclization. N.R.- Not reported

Enzyme	Category	Feature	Enzyme PDB	Origin	Recognition motif	Ref.
Butelase-1	PAL	Efficiency	6DHI	<i>C. ternatea</i>	GI-Xn-NHV	[31]
OaAEP1b-C247A	PAL	Efficiency	5H0I (OaAEP1)	<i>O. affinis</i>	GL-Xn-NGL	[35]
VyPAL2	PAL	Middle efficiency	6IDV	<i>V. yedoensis</i>	GI-Xn-NSL	[36]
VcAEP-V238A-Y168A	PAL	Efficiency	5ZBI (VcAEP)	<i>V. canadensis</i>	GI-Xn-NGI	[39]
ConPAL3-G225V-G155A	PAL	High efficiency Stability against heat and basic pH	N.R.	Legumains	GF-Xn-NXL	[42]
MCoAEP2	Hybrid	Effectively cyclize engineered MCoTI-II scaffolds	N.R.	<i>M. cochinchinensis</i>	DING-Xn-DAL	[40]
McPAL1	Hybrid	High stability against heat and basic pH	N.R.	<i>M. cochinchinensis</i>	DING-Xn-DAL	[41]

存在, 这揭示了转肽反应是水解非依赖性的<sup>[46]</sup>。PALs 的催化机制可以分为两个步骤 (图 3)<sup>[47]</sup>: ① PALs 催化中心的 Cys 巯基亲核攻击底物 P1 羰基, 导致硫酯酶中间体的形成以及 P1-P1' 肽键的裂解; ② 亲核试剂对该酰基酶中间体进行亲核攻击破坏了 Cys 和 P1-Asx 残基之间形成的瞬时硫酯键。水解反应和转肽反应的不同在于水解反应进入的亲核基团为水分子, 从而起到了蛋白酶的水解作用; 而连接反应进入的亲核基团为导入的多肽残基, 发生氨解反应, 促进了肽键的生成, 完成了转肽反应<sup>[29]</sup>。



**Figure 3** Mechanism of ligation by PALs, and amino-acid composition of the S2 and S1' pockets<sup>[47]</sup>

PALs 介导较大的多肽和蛋白质环化时需要底物具有天然的正确折叠构象, 且研究表明 N 末端和 C 末端的空间接近程度也会影响到底物的环化效率, 即 PALs 的催化不涉及辅助因子或其他外部能量来源, 是通过环化多肽产生的稳定构象以及裂解识别多肽基序获得的熵增益促使转肽反应的发生。最新的一项研究表明, PALs 的自动激活机制不同于上述的水解机制, 是通过催化三联体中的 His 和 Asn 残基完成自我加工成熟的步骤<sup>[47]</sup>。

## 2.4 PALs 蛋白质工程

鉴于 PALs 在生物技术和蛋白质工程应用中的潜力, 了解 PALs 催化肽键形成的分子基础, 将有望实现提高其催化效率、逆转 AEPs 水解功能以及拓宽底物谱等目标。基于 AEPs 和 PALs 具有相似的空间结构, 假设酶活性由 Cys 催化中心附近关键位置处氨基酸残基的细微差异来控制, 即此局部改变可能控制水分子 (导

致水解) 或亲核试剂 (导致连接) 对 S-酰基酶中间体的亲核攻击。通过使用生物信息学和体内外功能表征验证分析 PALs 的序列和结构, 大量研究结果表明: LAD1 (连接酶活性决定簇-ligase activity determinant, 包含门控氨基酸残基)、LAD2、聚脯氨酸环 (poly-proline loop, PPL) 和  $\alpha 5\text{-}\beta 6$  环 [也称为连接酶活性标志物 (marker of ligase activity, MLA)] 在 PALs 连接酶中起到重要作用<sup>[48]</sup>。

Yang 等<sup>[35]</sup>于 2017 年研究报道了氨基酸残基 Cys247 变化影响 PALs 连接效率的第一个决定因素, 命名为门控氨基酸残基区 (gate-keeper residue), 后续研究将此区域称为 LAD1。通过比较 legumain 蛋白酶与 OaAEP1 连接酶晶体结构的底物结合口袋区域, 发现连接酶相比蛋白酶具有更加平坦的酶表面。Cys247 被预测为亲核过滤器, 因为它位于 S2 底物口袋深处, Cys247 突变成 Ala 后局部空间位阻变小且疏水, 有利于 N 末端亲核试剂的进攻促进转肽反应的进行, 连接反应效率提高了 160 倍。而用 Gly 取代则会严重影响连接效率, Cys247Gly 突变体被证明可以增强多肽水解的效率, 更小且亲水的 Gly 作为门控氨基酸残基, 将允许水分子进入活性位点从而产生蛋白酶活性。

Jackson 等<sup>[49]</sup>于 2018 年证实来源于牵牛花 (*Petunia*) 的 PxAEP3b 与其他亚型酶具有超过 90% 的一致性, 但在催化功能上也存在差异; 通过比较分析其他合成环肽的 PALs 连接酶, 推测了 16 个可能支持 PALs 连接酶功能的残基, 确定了 3 个主要结构特征 (门控区、聚脯氨酸环和 MLA) 协同作用以改变其催化方向。由于 MLA 或 PPL 都远离活性位点, 这些单个特征是如何促进植物中 PALs 的功能偏好性的机制还尚未确定。因此, MLA 和 PPL 是 PALs 连接酶活性的相关区域。

Hemu 等<sup>[36]</sup>于 2019 年研究发现了连接酶活性的第二个决定因素, 命名为 LAD2。VyPAL3 是一个催化效率很低的 PAL, VcAEP 是蛋白酶。通过与 VyPAL2 进行对比发现, 它们的晶体结构和氨基酸序列在 LAD1 区域是相似的, 都具有非 Gly 的 Gate keeper 特征, 故推测是位于酶催化活性中心对侧 S1' 附近的 LAD2 区域中两个不保守的氨基酸残基起到重要作用, 影响了酶催化特征。通过 LAD2 区域的单个氨基酸残基的突变研究证实, VyPAL3 (Y175G) 显著提高了催化效率和环化水解比率; 也表明较大的 Tyr175 残基侧链可能会阻碍肽亲核基团攻击酰基酶中间体。因此, 在 LAD2 的第一个位置插入较小或疏水氨基酸残基如 Gly (或 Ala) 可显著提高连接效率, 加速离去基团的解离以及降低水分子的导入, 即有利于连接而不是水解反应。同样, VcAEP-Y168A 也改变了酶催化的偏好性, 使其从蛋

白酶变成连接酶且提高了酶的催化效率和环化产率。

Hemu 等<sup>[50]</sup>于 2020 年研究表明基于 LAD 假设可以成功地将天冬酰胺内肽酶 AEPs 逆转为 PALs 连接酶。Butelase-2 在 pH 值为 4~6.5 条件下具有显著的蛋白酶活性, 但通过突变 LAD 处的关键残基获得的 butelase-2 双突变体, 逆转了蛋白酶活性, 在酸性 pH 值下也表现出显著的连接酶活性。Chen 等<sup>[39]</sup>于 2021 年基于 LAD 假说, 针对 VcAEP 的 S2 和 S1' 附近口袋单独和组合地进行突变, 研究表明针对 LAD1 和 LAD2 区域的 VcAEP 双突变体是一种高效的连接酶, 可用作蛋白质环化和细胞表面标记的工具。

Hemu 等<sup>[42]</sup>于 2023 年人工设计了 conLEG 蛋白酶, 其 LAD1 和 LAD2 区域关键残基分别是 Gly225 和 Gly155, 具有类似 AEPs 的特征。为了将 conLEG 逆转为 PAL 样连接酶, 基于 LAD 假设针对 LAD1 进行了 Gly225Val 突变以产生类 Butelase-1 样的 conPAL1, 针对 LAD2 进行了 Gly155Ala 突变以产生类 VyPAL2 样的 conPAL2。针对 LAD1 和 LAD2 区域设计了 G155A/G225V 双突变体 conPAL3, 它结合了两个单突变体的特征, 研究表明, conPAL3 不仅逆转了水解作用, 其催化环化效率是 butelase-1 的 2 倍, 且此突变体还具有对 pH 和热的双重稳定性, 到目前为止即此突变体是一个最高效有力的多肽环化工具酶。

这些微妙的结构变化通过稳定硫酯酶中间体以及改变多肽底物的方向性从而实现改变酶功能的偏好性。除此之外, 提高酶表面的局部疏水性, 也有助于排斥水分子进入活性位点, 从而促进亲核基团的导入和转肽反应的进行。即控制 AEPs 和 PALs 活性的分子决定簇主要存在于 S1 口袋两侧的底物结合槽中, 特别是以 S2 口袋为中心的 LAD1 区域 (包含门控区) 和 S1' 旁边的 LAD2 区域。结合结构分析和诱变研究表明, 对

于高效的 PALs, LAD1 的第一个位置优选体积大且芳香的氨基酸残基如 Trp 或 Tyr; 第二个位置是疏水的氨基酸残基如 Val、Ile、Cys、或 Ala 而不是 Gly。对于 LAD2, 更受青睐的 GlyAla、AlaAla 或 AlaPro 二肽中第一个位置的小氨基酸残基 (如 Gly 或 Ala) 是连接酶活性的必要但不充分条件。在满足 LAD2 的条件下, AEPs 通过联合 LAD1 或 MLA 等更远区域的突变成为 PALs。即 LAD1、LAD2 和 MLA 中的氨基酸残基的联合作用决定了 AEPs 蛋白酶和连接酶活性的偏好性<sup>[29]</sup>。

此外, 底物和 pH 值也会对特定 PALs 功能的偏好性产生显著影响, 例如 MCoAEP2 不具有典型连接酶的特征, 但对于 MCoTI-II 底物的催化效率远优于其他 PALs, 这是酶与其天然底物共同进化的结果<sup>[40]</sup>。

### 3 PALs 异源重组表达策略

PALs 的发现是大环肽类药物研究过程中的一个重要里程碑, 只要能够获得足够量的活性 PALs 就可将所期望的活性多肽进行环化, 为提高多肽药物的稳定性提供强有力的工具酶 (表 2)<sup>[36,39-42,51-54]</sup>。

Nguyen 等<sup>[31]</sup>从热带植物蝴蝶豌豆的豆荚中通过 4 步色谱法从 1 kg 新鲜植物材料中分离获得了 5 mg 的 butelase-1。最新研究表明通过分析比较植物中不同组织的环化活性, 选取活性最高的幼嫩组织 (芽) 可以获得 15 mg 天然的 butelase-1, 较之前提高了 3 倍产量, 但其繁琐的提取分离步骤还是限制了 PALs 的大规模应用<sup>[55]</sup>。

PALs 在植物细胞中以酶原的形式通过分泌途径表达并运输到液泡, 在酸性的环境下激活成熟为活性 PALs, 由于较为复杂的酶原结构和必需的酸性条件诱导自激活的成熟机制在一定程度上也限制了其异源重组表达。目前通过使用不同的异源表达宿主和采取不同的优化表达策略以期实现 PALs 的高水平表达, 但又

**Table 2** Recombinant expression of PALs. c: Core domain of PALs. f: PALs without ER signal sequence. t: Truncated PALs. N.R. - Not reported

Enzyme	<i>E. coli</i>	Construction	Others	Construction	High yield /mg·L <sup>-1</sup>	Ref.
Butelase-1	<i>Rosetta</i> (DE3)	pMAL-c5x-MBP-f Butelase-1	<i>P. pastoris</i> X-33	pPICZαA-His <sub>6</sub> -TEVs-t Butelase-1	20	[51,54]
OaAEP1b-C247A	SHuffle T7	pET-28b-His <sub>6</sub> -Ubiquitin-c OaAEP1	BL21(T1R)	pET-28-His <sub>6</sub> -TEVs-t OaAEP1b-C247A-Δ351	9	[52,53]
VyPAL2	SHuffle T7	pET-28a-f VyPAL2 pMJS9 (Erv1p)	<i>Sf9</i>	pFB-Sec-NH-His <sub>6</sub> -TEVs-f VyPAL2, P3 virus	5-10	[36]
VcAEP-V238A-Y168A	SHuffle T7	pET-28a-His <sub>6</sub> -Ubiquitin-f VcAEP-V238A-Y168A	N.R.	N.R.	10	[39]
ConPAL3-G225V-G155A	SHuffle T7	pET-28a-His <sub>6</sub> -f conPAL3-His <sub>6</sub>	N.R.	N.R.	12	[42]
MCoAEP2	SHuffle T7	pHUE-His <sub>6</sub> -Ubiquitin-f MCoAEP2	N.R.	N.R.	N.R.	[40]
McPAL1	SHuffle T7	pET-28-His <sub>6</sub> -Ubiquitin-f McPAL1	<i>Sf9</i>	pDP1381-gp64-His <sub>6</sub> -TEVs-f McPAL1, P2 virus	5-10	[41]

受限制于繁琐的纯化步骤, 寻找一种高效又便捷的重组表达策略是未来需要克服的技术难题。

### 3.1 大肠杆菌表达体系

大肠杆菌 (*Escherichia coli*) 原核表达系统在表达外源重组蛋白方面具有很多优势, 包括遗传背景清晰、基因操作简单、生长快速和培养成本低廉以及重组蛋白产量高等优势, 被用于生产各种重组蛋白, 但由于不具备翻译后修饰的加工步骤, 对于真核蛋白的重组表达能力有限<sup>[56,57]</sup>。

Nguyen 等<sup>[31]</sup>在首次发现 butelase-1 后, 为了简化重组表达, 将去除信号肽的酶原编码框直接亚克隆到 pNIC28-Bsa4, 重组蛋白表达量高达  $200 \text{ mg} \cdot \text{L}^{-1}$ , 但却以不可溶的包涵体形式存在。这可能与 butelase-1 酶原中含有比较复杂的 3 对二硫键结构有关, 所以在大肠杆菌中表达经常会形成不可溶的包涵体。针对这个难点可以使用 SHuffle-T7 表达菌株, 此菌株将大大促进 PALs 的折叠和二硫键的形成。该菌株是通过突变硫氧还蛋白还原酶 (*trx*B) 和谷胱甘肽还原酶 (*gor*) 基因, 阻断大肠杆菌胞质还原途径使胞质形成氧化环境。同时将二硫键异构酶 *Dsb*c 基因整合到大肠杆菌染色体中, 使胞质组成型表达伴侣蛋白 *Dsb*c, 以纠正错误氧化配对的胱氨酸对, 促进了胞质蛋白二硫键的形成和折叠<sup>[58]</sup>。James 等<sup>[44]</sup>受自然界中 PALs 成熟机制的启发, 在体外酸性环境下模拟植物液泡的低 pH 环境, 将重组表达的酶原在缓冲液中进行孵育实现了 butelase-1 的自激活。但活性酶的获取还是受限于较低的产量和复杂的纯化激活步骤, 首先利用镍离子亲和色谱和离子交换以及分子尺寸排阻色谱纯化酶原, 酸性激活后, 再利用分子排阻色谱分离获得活性酶。随后有学者<sup>[51]</sup>采取融合表达 MBP 标签的策略来提高酶原的表达水平和溶解性, 酶原产量可以达到  $20 \text{ mg} \cdot \text{L}^{-1}$ 。除此之外, 使用 PALs 天然编码的密码子以及去除包含内质网信号肽在内的前 25 个氨基酸残基的策略也被证实有助于提高其产量。大肠杆菌重组表达的 butelase-1 和从植物中提取的天然活性酶相比, 存在 Asn94-N 连接糖基化的差别, 但被证实并不影响酶的催化效率。冷冻干燥中通过添加 12%~20% 的蔗糖和 0.1% Tween-20 将有助于 butelase-1 的储存, 防止冷冻过程中冰晶的形成和变性应力的产生<sup>[55]</sup>。

OaAEP1b-C247A 的重组表达策略包括酶原激活和直接构建核心区两种方式。酶原的 CTPP 除了抑制酶的活性以外, 还具有增强酶原溶解性和稳定性的作用, 所以大多数 PALs 都需要构建酶原的表达策略。酶原激活的构建体使用 6 个组氨酸和泛素融合标签替代酶原信号肽, 这种构建方式也常被用于其它 PALs

的表达, 包括 VcAEP-V238A-Y168A<sup>[39]</sup>、MCoAEP2<sup>[40]</sup>和 McPAL1<sup>[41]</sup>, 但此程序涉及多个纯化分离和低 pH 孵育激活步骤, 耗时长且产量低限制了其应用。最新的研究表明, OaAEP1b 较其他 PALs 的核心结构域在中性 pH 下更具稳定性, 即可以在没有前体肽的存在下直接表达活性区, 融合 6 个组氨酸和泛素融合标签绕过了酸性激活的步骤, 且只需要两步色谱法即可分离获得活性酶, 但产量为  $1.8 \sim 2.0 \text{ mg} \cdot \text{L}^{-1}$ <sup>[52]</sup>。还有学者<sup>[53]</sup>考虑到酶原自激活过程中存在末端切割异质性的问题并且随后通过色谱法难以分离不同的剪切体, 直接构建了 OaAEP1b-C247A- $\Delta$ 351 的活性区, 在 BL21 (T1R) 中表达, 获得不可溶的包涵体, 通过体外变性和复性的方法获得了  $9 \text{ mg} \cdot \text{L}^{-1}$  的活性酶, 但是又受限于繁琐耗时的梯度透析步骤以及对大量重折叠缓冲液的需求。

VyPAL2 和 VcAEP 以不融合标签的酶原前体形式进行表达, 但需要共表达一个 *Erv*1p 的 pMJS9 质粒以促进蛋白质的折叠和二硫键的形成从而获得可溶性的 PALs<sup>[36]</sup>。

最新研究<sup>[41]</sup>表明, 基于生物信息分析通过使用 AEPs 同源蛋白中最常出现的残基, 可以避免系统发育过程中的偏差。人工设计的 conPAL3 具有显著改善的生化特性, 由于提高了酶 mRNA 的稳定性以及促进了翻译和折叠过程中多肽的可溶性, 激活后活性酶的产量可高达  $12 \text{ mg} \cdot \text{L}^{-1}$ 。

### 3.2 昆虫杆状病毒表达体系

昆虫杆状病毒表达系统 (baculovirus expression vector system, BEVS) 是一种利用杆状病毒作为外源基因的载体, 通过感染昆虫细胞来高效表达外源蛋白的系统, 具有识别和处理信号肽的能力, 还可以进行翻译后修饰, 例如糖基化、酰化、磷酸化和二硫键的形成, 与原核表达系统相比, 能够正确形成真正的真核蛋白<sup>[59]</sup>。

PALs 以酶原策略在昆虫杆状病毒系统中进行表达, 包括 VyPAL2<sup>[36]</sup>和 McPAL1<sup>[41]</sup>, 其构建体具有 gp64 分泌信号肽序列、N 末端 6 个组氨酸标签和 TEV 切割位点特征的供体质粒。通过提取杆状病毒穿梭载体 bacmid, 感染 *Sf*9 昆虫细胞, 其分泌表达产量通常可以达到  $5 \sim 20 \text{ mg} \cdot \text{L}^{-1}$ , 但这种方法除了复杂的纯化和激活步骤的限制, 还受限于较高的培养成本。昆虫杆状病毒表达系统会产生不同糖基化的酶, 但是活性没有任何显著的差异, 糖基化修饰可以增加分子的稳定性, 因此会提高蛋白表达的产量。

### 3.3 毕赤酵母表达体系

毕赤酵母 (*Pichia pastoris*) 属于甲醇酵母, 作为外源基因真核表达系统具有许多优势, 包括具有调控机

制最严格的醇氧化酶基因 AOX1 强启动子、糖基化和蛋白磷酸化等翻译后修饰机制, 外源蛋白表达水平高且可以外泌表达等优势<sup>[60-62]</sup>。

Pi 等<sup>[54]</sup>利用毕赤酵母 X33 作为表达宿主菌, 构建了 pPICZαA-rBTase, 其特征包括外泌信号肽 α-Factor、6 个组氨酸纯化标签和 TEV 切割位点以及截短部分 CTPP 的 butelase-1 表达框, 实现了外泌表达, 通过镍离子亲和和色谱纯化获得了 16 mg·L<sup>-1</sup> 的产量, 但是此策略在其他 PALs 并未成功应用。

#### 4 PALs 应用

PALs 是一种多功能且高效的生物催化剂, 能够针对复杂的生物制剂进行精确修饰, 具有在温和和水性条件下催化肽键形成的天然优势, 在生物制造领域可以实现绿色可持续发展目标, 用于开发预防、诊断和治疗人类疾病的下一代生物制剂, 为生物药物设计和医学研究领域提供了强有力的酶工具<sup>[63-65]</sup>。

##### 4.1 PALs 单独催化

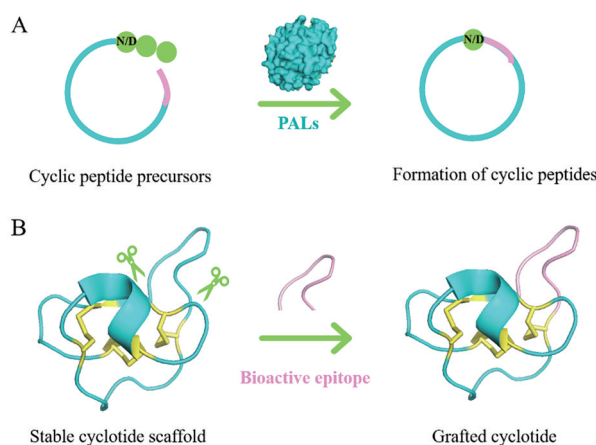
**4.1.1 分子内环化** PALs 介导分子内连接导致主链环化, 环化通常会降低构象柔性并赋予分子更高的热稳定性和化学稳定性, 多肽末端的连接可以抵抗外肽酶(氨肽酶和羧肽酶)的降解并提高多肽的口服生物利用度。理想的口服候选多肽应该具有低分子量、高效能、适当的酶或化学稳定性、高治疗指数和相对较低的开发成本特点。环肽稳定的拓扑结构, 成为了目前最具吸引力的分子骨架<sup>[66]</sup>。

药物从头合理设计中的嫁接概念来源于植物的人工繁殖, 即把一株植物的枝或芽, 嫁接到另一株植物的茎或根上, 使接在一起的两个部分长成一个完整的植株, 获得更具优良性状的品种, 嫁接成活的因素是两者的亲和力, 越相近越易成活。在药物设计中, 支架(类似于砧木)卓越的化学和热力学稳定性可以与小肽(类似于接穗)的有益特性相结合。将活性多肽分子直接进行环化或者将稳定的环肽骨架(CCK 框架)作为多肽药物的分子骨架进行蛋白质工程, 将生物活性表位嫁接到环肽的支架上, 可提高活性表位抵御人体内消化液降解的稳定性<sup>[67]</sup>。Craik 等<sup>[68]</sup>研发了一种治疗神经性疼痛的天然芋螺毒素口服活性肽, 就是设计了芋螺毒素的环状类似物。Wong 等<sup>[69]</sup>利用 kalata B1 作为天然支架, 将 DALK 或 DAK 移植到分子支架上设计出了用于治疗炎症性疼痛的口服活性肽。

在环肽支架的独特功能中, 最令人兴奋的是一些环肽可以穿透细胞并进入细胞胞质, 胰蛋白酶抑制剂亚家族的环肽可以靶向细胞内的蛋白质相互作用。研究表明环肽 MCoTI-I 可以设计成一种靶向 p53 与蛋白质 Hdm2/HdmX 之间相互作用的抑制剂<sup>[70]</sup>。这种基于

MCoTI 的环肽能够以非常高的亲和力与 Hdm2/HdmX 的 p53 结合结构域结合, 并且在血清中表现出很高的稳定性。经人结直肠癌动物模型研究表明该环肽在体外和体内均激活了 p53 肿瘤的抑制通路, 对野生型 p53 癌细胞系具有细胞毒性。该研究是一个代表着工程化环肽能够靶向细胞内蛋白质相互作用的成功示例, 也显示了基于 MCoTI 的环肽靶向细胞内蛋白质相互作用的真正治疗潜力<sup>[71]</sup>。MCoAEP2 可以介导环肽前体的合成, 即可支持胰蛋白酶抑制剂中的环肽移植, 使胰蛋白酶抑制剂作为分子骨架的嫁接肽能够得到更广泛的应用<sup>[40,72]</sup>。最新研究表明可以利用毕赤酵母表达宿主菌外泌生产环肽前体, 在体外利用 PALs 介导环肽的生产, 这一技术平台打通了环肽生物生产的技术壁垒, 加速了环肽在药物设计领域中的应用<sup>[73]</sup>。

目前已经开发了用于多种疾病嫁接环肽的例子, 包含疼痛、炎症、伤口愈合、心血管疾病、癌症和多发性硬化症等, 虽然这些实例均未到达临床试验, 但体外实验已充分证实了环肽巨大的药用潜力和口服生物利用度的优势, 这使环肽成为目前最有开发前景的多肽分子(图4)<sup>[67,74,75]</sup>。



**Figure 4** Schematic diagram illustrating the concept of PALs (A) and grafting an epitope into a cyclotide scaffold (B)<sup>[66]</sup>

除了在药物设计领域中的应用, PALs 还可用于改良工业酶或者不稳定的蛋白, 使其环化提高稳定性, 但对于蛋白质的分子内环化, N 末端和 C 末端必须在空间结构上靠近才能进行有效的转肽反应<sup>[76-78]</sup>。

**4.1.2 分子间连接** PALs 还可用于分子间连接介导蛋白质特异性位点的标记, 包括阐明细胞内的代谢途径、定义蛋白质-蛋白质相互作用、开发抗体药物偶联物和创新的医学成像方法以及治疗方法等用途。例如活细菌表面蛋白的位点特异性标记, 在大肠杆菌锚定蛋白 OmpA 的 C 末端设计 NHV 识别基序。使用 PALs

孵育活细胞时,可以将一系列目标分子成功地与工程细菌表面蛋白 OmpA 连接,包括荧光素探针,可用于病原体反应的细胞跟踪等<sup>[79-81]</sup>。

**4.1.3 固定化酶** 固定化酶可以克服可溶性酶的局限性,包括聚集和自溶成活性较低的形式。Hemu 等<sup>[82]</sup>使用 3 种方法实现了两种 PALs 的固定化,研究表明固定化的连接酶具有更高的稳定性,并且在重复使用 100 次后其活性损失小于 10%。在大规模工业应用中,还可以通过提高酶浓度来加速连接反应,进而简化纯化过程<sup>[83]</sup>。

## 4.2 PALs 正交策略

蛋白质治疗学是将诊断和治疗功能整合到靶向单一疾病的蛋白质上。然而,这些多功能药物的制备仍然是一个重大挑战。PALs 可以通过改变反应 pH 条件或者与其他多肽连接酶联用的生物正交方法,在一锅法条件下通过串联反应协调有序地精确控制蛋白质的修饰,扩展了生化工具包。

**4.2.1 单酶正交策略** PALs 对天冬酰胺和天冬氨酸位点的催化活性具有 pH 依赖性,在酸性 pH 值条件下对 Asp 底物表现出显著的连接活性,在接近中性的 pH 值条件下对 Asn 底物具有偏好性。利用 PALs 对 P1-Asn/Asp 底物催化活性的特点,即可介导蛋白质的顺序连接,精密制造具有多种功能的复杂生物偶联物<sup>[84]</sup>。

**4.2.2 双酶正交策略** Butelase-1 和 VyPAL2 具有可区分的底物 C 端识别序列偏好性,这为蛋白质的双重修饰提供了一个正交策略,在接近中性 pH 值条件下,可以实现 N 末端到 C 末端和 C 末端到 N 末端两个方向上的蛋白质顺序串联连接<sup>[85]</sup>。

OaAEP1b 对 N 端导入的亲核试剂具有杂泛性,当生成的产物 C 端识别序列再次被识别时不具有酶的偏好性,即可降低水解反应的速率,从而提高了产物的产量,通过与 TEV 酶的联合使用即可对底物进行顺序蛋白质连接或修饰<sup>[86]</sup>。

## 5 小结与展望

PALs 的发现和在大环肽类药物设计以及生物技术领域中无疑是一个助推器,但是却受限于 PALs 高效简便的异源重组表达,组合应用提高蛋白表达水平的策略实现规模化生产是目前努力的方向。

综合生物信息学和体内外重组表达的功能表征策略,从植物中挖掘和鉴定更多的环肽和 PALs,通过阐明进化关系以及对酶催化机制进行更深入的探索,总结规律系统地阐明决定连接酶活性的因素或标志物,这将有助于对其进行人工改造,提高 PALs 的催化效率以及拓宽其在生物技术应用中的底物谱。

参与环肽生物合成途径中的翻译后修饰关键酶基

因还尚未完全解析,鉴定更多的 N 端剪切酶,才可能实现与 PALs 的组合应用,在其他异源宿主主体内重构环肽生物合成途径从而实现环肽类药物的大规模生产,这将为环肽类嫁接药物的开发奠定坚实的基础<sup>[87-90]</sup>。

大环肽类的药物开发属于结构修饰,头尾环化的骨架对人体内消化液的降解具有一定的耐受性。通过联合应用口服递送的制剂技术,例如添加 SNAC [sodium N-[8-(2-hydroxybenzoyl)amino] caprylate] 渗透促进剂或者使用纳米递送载体,即可真正地完从分子结构到制剂技术的联合应用,改善多肽口服生物利用度低的问题,实现多肽类药物的口服给药。

在多肽环化酶的工具箱中,虽然 PALs 因为无痕的肽键形成和高效率的优势成为有力的多肽环化工具酶,但是其他类型的环化酶以及商用的 omniligase-1 也具有自身的独特优势。对于想要环化的多肽或蛋白底物,可以综合分析进行选择或者是组合使用,以达到最终目的<sup>[91]</sup>。

**作者贡献:** 申欣撰写了全文,对内容进行修改和补充;王伟、刘恣之和杨燕对文章进行审阅并提出修改意见。

**利益冲突:** 所有作者均声明不存在利益冲突。

## References

- [1] Craik DJ, Kan MW. How can we improve peptide drug discovery? learning from the past [J]. *Expert Opin Drug Discov*, 2021, 16: 1399-1402.
- [2] Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions [J]. *Drug Discov Today*, 2015, 20: 122-128.
- [3] Muttenthaler M, King GF, Adams DJ, et al. Trends in peptide drug discovery [J]. *Nat Rev Drug Discov*, 2021, 20: 309-325.
- [4] Zhang H, Chen S. Cyclic peptide drugs approved in the last two decades (2001 - 2021) [J]. *RSC Chem Biol*, 2022, 3: 18-31.
- [5] Jing X, Jin K. A gold mine for drug discovery: strategies to develop cyclic peptides into therapies [J]. *Med Res Rev*, 2020, 40: 753-810.
- [6] de Veer SJ, Kan MW, Craik DJ. Cyclotides: from structure to function [J]. *Chem Rev*, 2019, 119: 12375-12421.
- [7] Schmidt M, Huang Y, Teixeira de Oliveira EF, et al. Efficient enzymatic cyclization of disulfide-rich peptides by using peptide ligases [J]. *Chembiochem*, 2019, 20: 1524-1529.
- [8] Lee J, McIntosh J, Hathaway BJ, et al. Using marine natural products to discover a protease that catalyzes peptide macrocyclization of diverse substrates [J]. *J Am Chem Soc*, 2009, 131: 2122-2124.
- [9] Chekan JR, Estrada P, Covello PS, et al. Characterization of the macrocyclase involved in the biosynthesis of RiPP cyclic peptides in plants [J]. *Proc Natl Acad Sci U S A*, 2017, 114: 6551-6556.

- [10] Luo H, Hong SY, Sgambelluri RM, et al. Peptide macrocyclization catalyzed by a prolyl oligopeptidase involved in  $\alpha$ -amanitin biosynthesis [J]. *Chem Biol*, 2014, 21: 1610-1617.
- [11] Mazmanian SK, Liu G, Ton-That H, et al. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall [J]. *Science*, 1999, 285: 760-763.
- [12] Chang TK, Jackson DY, Burnier JP, et al. Subtiligase: a tool for semisynthesis of proteins [J]. *Proc Natl Acad Sci U S A*, 1994, 91: 12544-12548.
- [13] Weeks AM, Wells JA. Subtiligase-catalyzed peptide ligation [J]. *Chem Rev*, 2020, 120: 3127-3160.
- [14] Jackson MA, Nguyen LTT, Gilding EK, et al. Make it or break it: plant AEPs on stage in biotechnology [J]. *Biotechnol Adv*, 2020, 45: 107651.
- [15] Vorster BJ, Cullis CA, Kunert KJ. Plant vacuolar processing enzymes [J]. *Front Plant Sci*, 2019, 10: 479.
- [16] Yamada K, Basak AK, Goto-Yamada S, et al. Vacuolar processing enzymes in the plant life cycle [J]. *New Phytol*, 2020, 226: 21-31.
- [17] Gran L. An oxytocic principle found in *Oldenlandia affinis* DC [J]. *Med Nor Farm Selsk*, 1970, 12: 80.
- [18] Craik DJ, Daly NL, Bond T, et al. Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif [J]. *J Mol Biol*, 1999, 294: 1327-1336.
- [19] Rosengren KJ, Daly NL, Plan MR, et al. Twists, knots, and rings in proteins [J]. *J Biol Chem*, 2003, 278: 8606-8616.
- [20] Colgrave ML, Craik DJ. Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot [J]. *Biochemistry*, 2004, 43: 5965-5975.
- [21] Craik DJ. Circling the enemy: cyclic proteins in plant defence [J]. *Trends Plant Sci*, 2009, 14: 328-335.
- [22] Cascales L, Craik DJ. Naturally occurring circular proteins: distribution, biosynthesis and evolution [J]. *Org Biomol Chem*, 2010, 8: 5035-5047.
- [23] Craik DJ, Malik U. Cyclotide biosynthesis [J]. *Curr Opin Chem Biol*, 2013, 17: 546-554.
- [24] Qu H. *Plant Derived Cyclic Peptides: From Discovery to Biotechnological Applications* [D]. Queensland: The University of Queensland, 2019.
- [25] Kalmankar NV, Venkatesan R, Balaran P, et al. Transcriptomic profiling of the medicinal plant *Clitoria ternatea*: identification of potential genes in cyclotide biosynthesis [J]. *Sci Rep*, 2020, 10: 12658.
- [26] Narayani M, Babu R, Chadha A, et al. Production of bioactive cyclotides: a comprehensive overview [J]. *Phytochem Rev*, 2020, 19: 787-825.
- [27] Rehm FBH, Jackson MA, De Geyter E, et al. Papain-like cysteine proteases prepare plant cyclic peptide precursors for cyclization [J]. *Proc Natl Acad Sci U S A*, 2019, 116: 7831-7836.
- [28] Nonis SG, Haywood J, Schmidberger JW, et al. Structural and biochemical analyses of concanavalin a circular permutation by jack bean asparaginyl endopeptidase [J]. *Plant Cell*, 2021, 33: 2794-2811.
- [29] Hemu X, Chan NY, Liew HT, et al. Substrate-binding glycine residues are major determinants for hydrolase and ligase activity of plant legumains [J]. *New Phytol*, 2023, 238: 1534-1545.
- [30] Schechter I, Berger A. On the active site of proteases. III. Mapping the active site of papain; specific peptide inhibitors of papain [J]. *Biochem Biophys Res Commun*, 1968, 32: 898-902.
- [31] Nguyen GKT, Wang S, Qiu Y, et al. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis [J]. *Nat Chem Biol*, 2014, 10: 732-738.
- [32] Oguis GK, Gilding EK, Jackson MA, et al. Butterfly Pea (*Clitoria ternatea*), a cyclotide-bearing plant with applications in agriculture and medicine [J]. *Front Plant Sci*, 2019, 10: 645.
- [33] Nguyen GKT, Qiu Y, Cao Y, et al. Butelase-mediated cyclization and ligation of peptides and proteins [J]. *Nat Protoc*, 2016, 11: 1977-1988.
- [34] Harris KS, Durek T, Kaas Q, et al. Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase [J]. *Nat Commun*, 2015, 6: 10199.
- [35] Yang R, Wong YH, Nguyen GKT, et al. Engineering a catalytically efficient recombinant protein ligase [J]. *J Am Chem Soc*, 2017, 139: 5351-5358.
- [36] Hemu X, El Sahili A, Hu S, et al. Structural determinants for peptide-bond formation by asparaginyl ligases [J]. *Proc Natl Acad Sci U S A*, 2019, 116: 11737-11746.
- [37] Wang CKL, Colgrave ML, Gustafson KR, et al. Anti-HIV cyclotides from the Chinese medicinal herb *Viola yedoensis* [J]. *J Nat Prod*, 2008, 71: 47-52.
- [38] Park S, Yoo KO, Marcussen T, et al. Cyclotide evolution: insights from the analyses of their precursor sequences, structures and distribution in Violets (*Viola*) [J]. *Front Plant Sci*, 2017, 8: 2058.
- [39] Chen Y, Zhang D, Zhang X, et al. Site-specific protein modifications by an engineered asparaginyl endopeptidase from *Viola canadensis* [J]. *Front Chem*, 2021, 9: 768854.
- [40] Du J, Yap K, Chan LY, et al. A bifunctional asparaginyl endopeptidase efficiently catalyzes both cleavage and cyclization of cyclic trypsin inhibitors [J]. *Nat Commun*, 2020, 11: 1575.
- [41] Liew HT, To J, Zhang X, et al. The Legumain McPAL1 from *Momordica cochinchinensis* is a highly stable Asx-specific splicing enzyme [J]. *J Biol Chem*, 2021, 297: 101325.
- [42] Hemu X, Zhang X, Chang HY, et al. Consensus design and engineering of an efficient and high-yield peptide asparaginyl ligase for protein cyclization and ligation [J]. *J Biol Chem*, 2023, 299: 102997.
- [43] Nonis SG, Haywood J, Mylne JS. Plant asparaginyl endopeptidases and their structural determinants of function [J]. *Biochem Soc Trans*, 2021, 49: 965-976.

- [44] James AM, Haywood J, Leroux J, et al. The macrocyclizing protease butelase 1 remains autocatalytic and reveals the structural basis for ligase activity [J]. *Plant J*, 2019, 98: 988-999.
- [45] Zhao J, Ge G, Huang Y, et al. Study on activation mechanism and cleavage sites of recombinant butelase-1 zymogen derived from *Clitoria ternatea* [J]. *Biochimie*, 2022, 199: 12-22.
- [46] James AM, Haywood J, Mylne JS. Macrocyclization by asparaginyl endopeptidases [J]. *New Phytol*, 2018, 218: 923-928.
- [47] Hu S, El Sahili A, Kishore S, et al. Structural basis for proenzyme maturation, substrate recognition, and ligation by a hyperactive peptide asparaginyl ligase [J]. *Plant Cell*, 2022, 34: 4936-4949.
- [48] Rehm FBH, Tyler TJ, Xie J, et al. Asparaginyl ligases: new enzymes for the protein engineer's toolbox [J]. *Chembiochem*, 2021, 22: 2079-2086.
- [49] Jackson MA, Gilding EK, Shafee T, et al. Molecular basis for the production of cyclic peptides by plant asparaginyl endopeptidases [J]. *Nat Commun*, 2018, 9: 2411.
- [50] Hemu X, El Sahili A, Hu S, et al. Turning an asparaginyl endopeptidase into a peptide ligase [J]. *ACS Catal*, 2020, 10: 8825-8834.
- [51] Zhao J, Fan R, Jia F, et al. Enzymatic properties of recombinant ligase butelase-1 and its application in cyclizing food-derived angiotensin I-converting enzyme inhibitory peptides [J]. *J Agric Food Chem*, 2021, 69: 5976-5985.
- [52] Tang TMS, Cardella D, Lander AJ, et al. Use of an asparaginyl endopeptidase for chemo-enzymatic peptide and protein labeling [J]. *Chem Sci*, 2020, 11: 5881-5888.
- [53] Chua N, Wong YH, El Sahili A, et al. On the design of a constitutively active peptide asparaginyl ligase for facile protein conjugation [J]. *FEBS Open Bio*, 2023. DOI: 10.1002/2211-5463.13575.
- [54] Pi N, Gao M, Cheng X, et al. Recombinant butelase-mediated cyclization of the p53-binding domain of the oncoprotein MdmX-stabilized protein conformation as a promising model for structural investigation [J]. *Biochemistry*, 2019, 58: 3005-3015.
- [55] Hemu X, Zhang X, Nguyen GKT, et al. Characterization and application of natural and recombinant butelase-1 to improve industrial enzymes by end-to-end circularization [J]. *RSC Adv*, 2021, 11: 23105-23112.
- [56] Francis DM, Page R. Strategies to optimize protein expression in *E. coli* [J]. *Curr Protoc Protein Sci*, 2010, 5: 5.24.1-5.24.29.
- [57] Baeshen MN, Al-Hejjin AM, Bora RS, et al. Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives [J]. *J Microbiol Biotechnol*, 2015, 25: 953-962.
- [58] Lobstein J, Emrich CA, Jeans C, et al. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm [J]. *Microb Cell Fact*, 2012, 11: 56.
- [59] Chambers AC, Aksular M, Graves LP, et al. Overview of the baculovirus expression system [J]. *Curr Protoc Protein Sci*, 2018, 91: 5.4.1-5.4.6.
- [60] Damasceno LM, Huang CJ, Batt CA. Protein secretion in *Pichia pastoris* and advances in protein production [J]. *Appl Microbiol Biotechnol*, 2012, 93: 31-39.
- [61] Ahmad M, Hirz M, Pichler H, et al. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production [J]. *Appl Microbiol Biotechnol*, 2014, 98: 5301-5317.
- [62] Mohammadzadeh R, Karbalaee M, Soleimanpour S, et al. Practical methods for expression of recombinant protein in the *Pichia pastoris* system [J]. *Curr Protoc*, 2021, 1: e155.
- [63] Nguyen GKT, Kam A, Loo S, et al. Butelase 1: a versatile ligase for peptide and protein macrocyclization [J]. *J Am Chem Soc*, 2015, 137: 15398-15401.
- [64] Tam JP, Chan NY, Liew HT, et al. Peptide asparaginyl ligases—renegade peptide bond makers [J]. *Sci China Chem*, 2020, 63: 296-307.
- [65] Zhang D, Wang Z, Hu S, et al. Vypal2: a versatile peptide ligase for precision tailoring of proteins [J]. *Int J Mol Sci*, 2021, 23: 458.
- [66] Poth AG, Chan LY, Craik DJ. Cyclotides as grafting frameworks for protein engineering and drug design applications: cyclotides as grafting frameworks [J]. *Biopolymers*, 2013, 100: 480-491.
- [67] Craik DJ, Du J. Cyclotides as drug design scaffolds [J]. *Curr Opin Chem Biol*, 2017, 38: 8-16.
- [68] Clark RJ, Jensen J, Nevin ST, et al. The engineering of an orally active conotoxin for the treatment of neuropathic pain [J]. *Angew Chem Int Ed Engl*, 2010, 49: 6545-6548.
- [69] Wong CTT, Rowlands DK, Wong CH, et al. Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment [J]. *Angew Chem Int Ed Engl*, 2012, 51: 5620-5624.
- [70] Ji Y, Majumder S, Millard M, et al. *In vivo* activation of the p53 tumor suppressor pathway by an engineered cyclotide [J]. *J Am Chem Soc*, 2013, 135: 11623-11633.
- [71] Philippe GJB, Craik DJ, Henriques ST. Converting peptides into drugs targeting intracellular protein-protein interactions [J]. *Drug Discov Today*, 2021, 26: 1521-1531.
- [72] Bernath-Levin K, Nelson C, Elliott AG, et al. Peptide macrocyclization by a bifunctional endopeptidase [J]. *Chem Biol*, 2015, 22: 571-582.
- [73] Yap K, Du J, Looi FY, et al. An environmentally sustainable biomimetic production of cyclic disulfide-rich peptides [J]. *Green Chem*, 2020, 22: 5002-5016.
- [74] Aboye T, Meeks C, Majumder S, et al. Design of a MCoTI-based cyclotide with angiotensin (1-7)-like activity [J]. *Molecules*, 2016, 21: 152.
- [75] Mehta L, Dhankhar R, Gulati P, et al. Natural and grafted cyclotides in cancer therapy: an insight [J]. *J Pept Sci*, 2020, 26: e3246.

- [76] Hemu X, Qiu Y, Nguyen GKT, et al. Total synthesis of circular bacteriocins by butelase 1 [J]. *J Am Chem Soc*, 2016, 138: 6968-6971.
- [77] Nguyen GKT, Hemu X, Quek JP, et al. Butelase-mediated macrocyclization of *D*-amino-acid-containing peptides [J]. *Angew Chem Int Ed Engl*, 2016, 128: 12994-12998.
- [78] Harris KS, Guarino RF, Dissanayake RS, et al. A suite of kinetically superior AEP ligases can cyclise an intrinsically disordered protein [J]. *Sci Rep*, 2019, 9: 10820.
- [79] Hemu X, Zhang X, Bi X, et al. Butelase 1-mediated ligation of peptides and proteins [M]//Nuijens T, Schmidt M. *Enzyme-mediated Ligation Methods*. New York: Springer New York, 2019: 83-109.
- [80] Hemu X, Zhang X, Tam JP. Ligase-controlled cyclo-oligomerization of peptides [J]. *Org Lett*, 2019, 21: 2029-2032.
- [81] Bi X, Yin J, Nguyen GKT, et al. Enzymatic engineering of live bacterial cell surfaces using butelase-1 [J]. *Angew Chem Int Ed Engl*, 2017, 56: 7822-7825.
- [82] Hemu X, To J, Zhang X, et al. Immobilized peptide asparaginyl ligases enhance stability and facilitate macrocyclization and site-specific ligation [J]. *J Org Chem*, 2020, 85: 1504-1512.
- [83] Ma Q, He B, Tang G, et al. Enzymatic protein immobilization on amino-functionalized nanoparticles [J]. *Molecules*, 2023, 28: 379.
- [84] Zhang D, Wang Z, Hu S, et al. pH-Controlled protein orthogonal ligation using asparaginyl peptide ligases [J]. *J Am Chem Soc*, 2021, 143: 8704-8712.
- [85] Wang Z, Zhang D, Hemu X, et al. Engineering protein theranostics using bio-orthogonal asparaginyl peptide ligases [J]. *Theranostics*, 2021, 11: 5863-5875.
- [86] Rehm FBH, Harmand TJ, Yap K, et al. Site-specific sequential protein labeling catalyzed by a single recombinant ligase [J]. *J Am Chem Soc*, 2019, 141: 17388-17393.
- [87] Poon S, Harris KS, Jackson MA, et al. Co-expression of a cyclizing asparaginyl endopeptidase enables efficient production of cyclic peptides in planta [J]. *J Exp Bot*, 2018, 69: 633-641.
- [88] Wang CK, Craik DJ. Designing macrocyclic disulfide-rich peptides for biotechnological applications [J]. *Nat Chem Biol*, 2018, 14: 417-427.
- [89] Camarero JA, Campbell MJ. The potential of the cyclotide scaffold for drug development [J]. *Biomedicines*, 2019, 7: 31.
- [90] Lee MH. *Elucidating Cyclotide Biosynthetic Pathways for Plant-based Recombinant Expression* [D]. Queensland: The University of Queensland, 2022.
- [91] Morgan HE, Turnbull WB, Webb ME. Challenges in the use of sortase and other peptide ligases for site-specific protein modification [J]. *Chem Soc Rev*, 2022, 51: 4121-4145.