中国科学: 化学 2026 年 第 56 卷 第 4 期: 1303 ~ 1320
SCIENTIA SINICA Chimica chemcn.scichina.com
《中国科学》杂志社 SCIENCE CHINA PRESS
评 述 清华大学化学系建系100周年专刊

蛋白质合成方法学新进展: 从化学合成到主链编辑

周洋, 魏立臻, 陈永湘*

清华大学化学系, 生命有机磷化学及化学生物学教育部重点实验室, 北京 100084
*通讯作者, E-mail: chen-yx@mail.tsinghua.edu.cn

收稿日期: 2025-12-22; 接受日期: 2026-02-11; 网络版发表日期: 2026-03-23
国家自然科学基金(编号: 22177059, W2412026)资助项目

摘要 蛋白质是生命活动的主要执行者. 精确解析和操控其结构与功能具有重要科学意义和应用价值. 虽然生物重组表达蛋白质技术应用广泛, 但在制备具有均一翻译后修饰、特定外源标记或非天然功能单元的蛋白质时仍面临许多挑战. 自液相和固相多肽合成方法建立以来, 随着自然化学连接、酶促连接等多肽片段连接方法的发展, 研究者实现了许多中等至大分子量蛋白质的化学全合成或半合成. 同时, 利用遗传密码子扩展技术定点将非天然氨基酸引入蛋白质序列, 进一步拓展了蛋白质的化学空间. 近年来, 随着多肽骨架重构反应和分裂内含肽体系的发展, 研究者进一步在体外及细胞内实现了蛋白质主链原位编辑. 本文以“从化学合成到主链编辑”为主线, 总结了固相多肽合成、多肽片段连接、遗传密码子扩展及蛋白质编辑等方法学的最新进展, 为蛋白质的精确构建提供参考, 以推动蛋白质结构和功能研究.

关键词 蛋白质化学合成, 固相多肽合成, 多肽片段连接, 蛋白质编辑, 骨架重构, 片段替换

1 引言

蛋白质化学合成旨在通过人工手段精确组装氨基酸(amino acid, AA)序列, 构筑具有复杂天然或非天然结构的功能性蛋白质分子. 早期化学合成肽键主要依赖液相合成. 我国科学家上世纪60年代利用液相多肽合成(solution-phase peptide synthesis)实现了结晶牛胰岛素的全合成, 是多肽与蛋白质化学合成的里程碑[1,2]. 随后, 固相多肽合成技术(solid-phase peptide synthesis, SPPS)的建立与不断完善, 推动了多肽合成的自动化与标准化, 为蛋白质的全合成与半合成的广泛开展奠定了坚实的技术基础[3~5].

天然提取和生物重组表达是获取蛋白质的主要途径, 且在制备常规蛋白质方面效率极高, 但这些传统路径难以获得具有定点翻译后修饰的均一蛋白质样品, 也无法引入非天然结构单元. 此外, 部分具有高毒性、强疏水性或结构不稳定的天然蛋白质在细胞内也难以高水平表达[5~8]. 这些局限性迫切要求发展蛋白质化学合成与半合成新策略, 以便在可控条件下高效精准构筑通过传统路径难以制备的含特定修饰或表达困难的蛋白质.

在此背景下, 固相多肽合成, 自然化学连接(native chemical ligation, NCL)、酶促多肽连接(enzymatic peptide ligation)等多肽片段连接技术, 遗传密码子扩展(genetic code expansion, GCE)技术等共同构成了现代蛋白质合成的核心技术框架[6~12]. 近年来, 随着多肽

引用格式: Zhou Y, Wei L, Chen YX. Recent advances in protein synthesis: from chemical synthesis to backbone editing. *Sci Sin Chim*, 2026, 56: 1303–1320, doi: 10.1360/SSC-2025-0308

© 2026 《中国科学》杂志社

骨架重构(backbone reconstitution)方法和分裂内含肽(split intein)体系的发展, 在蛋白质主链层面进行精准的骨架重构与功能片段替换(segment replacement)已成为可能, 且该技术已从体外实验逐步拓展至细胞内应用[13~16]. 本文将围绕"从化学合成到主链编辑"这一主线, 系统梳理上述方法学的最新研究进展, 并探讨其在蛋白质科学中的应用.

2 蛋白质合成方法简要回顾

本节将主要回顾固相多肽合成、多肽片段连接技术和遗传密码子扩展等蛋白质合成相关核心技术. 关于这些技术详细的机理和应用, 已有若干优秀综述[5~12,17~23]作了系统梳理. 本文不拟展开详细介绍, 而是选择性地概述关键部分和最新进展, 以此为过渡, 引出"3 蛋白质编辑"部分重点介绍的蛋白质编辑新方法.

2.1 固相多肽合成

固相多肽合成自Merrifield[3~5]提出以来, 一直是蛋白质化学合成的基础核心技术. 固相多肽合成以树脂为固相载体, 将保护的氨基酸单体按顺序(一般从多肽序列的C端往N端)依次偶联生成酰胺键, 从而构建出具有特定序列的多肽(图1a). 保护基设计与脱保护条件的演变一直是SPPS方法学发展的核心. 早期的Boc/Bzl保护基以叔丁氧羰基(*tert*-butoxycarbonyl, Boc)保护N端氨基. 苄基(benzyl, Bzl)类保护基保护侧链, 需使用氢氟酸(hydrogen fluoride, HF)进行树脂切割与脱保护, 操作安全性和环境友好性受限. 目前, 沿用至今的9-芴甲氧羰基(9-fluorenylmethoxycarbonyl, Fmoc)/叔丁基(*tert*-butyl, *t*Bu)以Fmoc保护基保护N端氨基, 以*t*Bu类保护基保护酸性或亲核侧链, 在温和碱性条件下去除Fmoc, 在三氟乙酸(trifluoroacetic acid, TFA)条件下同时完成侧链脱保护与树脂切割. 该方法操作简便、易于自动化, 能高效制备长度约50个氨基酸的多肽[24,25]. 利用SPPS可在肽链的主链与侧链上引入精确修饰, 已经被广泛用于各类修饰多肽和小型蛋白质的合成[7]. 例如, Chen等[26,27]通过合成具有特定修饰模式的多肽, 系统研究了磷酸化、法尼基化对Rnd3与14-3-3ζ蛋白质互作的影响, 阐明了特定修饰在细胞信号调控中的关键作用. Dong等[28]利用Fmoc-SPPS实现了未修饰及含定点单糖、二糖以及三糖基化的β-淀粉样蛋白(amyloid-β, Aβ)全合成, 发现该修饰可诱导Aβ形成一类对降解更敏感的新型淀粉样纤维多型体.

随着对多肽及蛋白质药物需求的增长, 传统SPPS中大量使用保护基、缩合剂及有机溶剂所引发的环境问题日益凸显, 其原子经济性有待提高[29]. 因此, 围绕多肽偶联、保护与脱保护环节的简化与绿色化, 已成为该领域近年来的重要发展方向. 例如, 在偶联环节, Wang等[30]发展了由5-硝基-4,6-二硫氰基嘧啶(NDTP)介导的直接构建酰胺键的方法, 可在温和条件下实现快速偶联并有效避免消旋, 为SPPS提供了新的缩合试剂选择. 在保护与脱保护方面, Zhou等[31]使用2,4-二硝基-6-苯基亚硫酰基苯(DNPBS)构建含S–N键的氨基保护基, 其在近中性的条件下即可快速脱保护. 该条件能有效抑制传统SPPS中的典型副反应(如α-C消旋和天冬酰胺亚胺形成), 不仅提高了合成产率, 也保持了产物立体构型. 下文选取了最新的几例代表性研究进行简要介绍.

Zhao等[29]建立了逆向多肽合成(inverse peptide synthesis)策略. 如图1b, 首先利用炔酰胺(ynamide)作为活化剂活化羧基, 随即投入未保护的氨基酸与硅烷化试剂(silylating reagent), 原位生成硅酯保护的氨基酸. 而暴露的氨基会与活化羧基快速反应, 然后经酸碱处理, 硅酯会水解脱保护暴露出下一个羧基, 重复这个流程便可实现从N端到C端的顺序偶联. 该策略直接以未保护氨基酸为原料, 循环过程中只需一次操作便可实现临时保护、活化、偶联和原位脱保护, 避免了繁琐的保护基操作, 显著提高了原子利用率, 并与液相片段缩合具有良好的兼容性.

Wang等[32]从侧链保护基入手, 开发了温和的光脱保护SPPS体系. 如图1c, 该体系利用吡啶甲基(Pic)替代传统的*t*Bu类保护基, 在肽链组装完成后, 通过可见光催化触发Pic基团上C–X (X=O, N, S)键的选择性断裂, 从而在近中性条件下"一锅"实现侧链脱保护与树脂切割. 在此基础上, 他们进一步将该光触发策略拓展至*N*α-氨基保护, 报道了温和脱保护的吡啶甲氧羰基-固相多肽合成(Picoc-SPPS)[33]. 他们还将不同的Pic类保护基用于半胱氨酸(Cys)的笼化与脱笼化(caging/uncaging), 通过调控pH与光照条件实现正交的脱笼化, 便于与化学连接流程衔接[34]. 总体而言, 该系列工作以可见光为触发手段实现多位点保护基的可控去除与后续组装, 与传统的Fmoc/*t*Bu体系相比, 其在温和条件

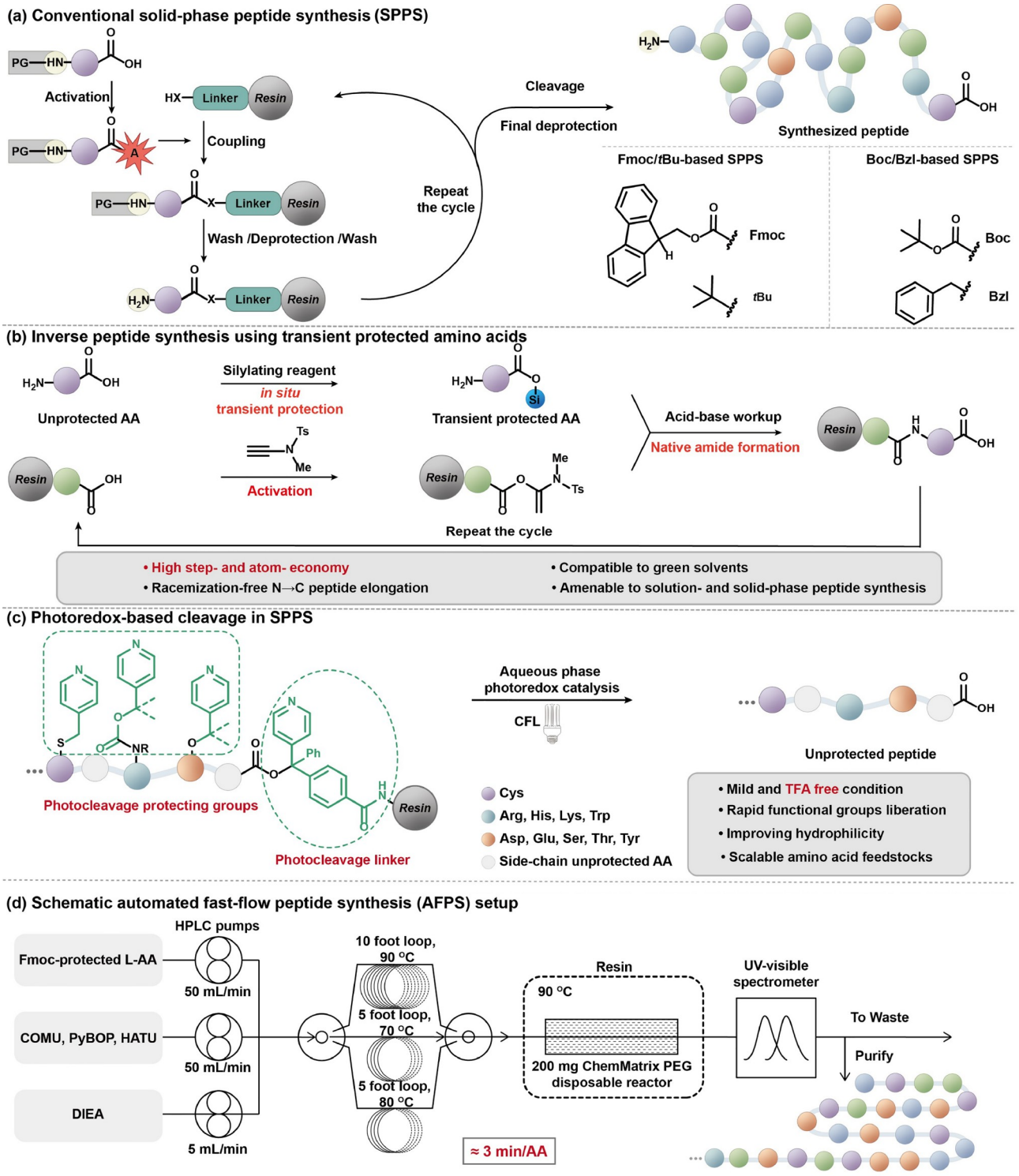


图 1　(网络版彩图)固相多肽合成及相关进展. (a) 固相多肽合成基本流程以及传统的保护基策略; (b) 炔酰胺介导的逆向多肽合成策略[29]; (c) 光催化多肽脱保护策略[32]; (d) 自动化快速流体多肽合成[35]

Figure 1　(Color online) Solid-phase peptide synthesis and recent advances. (a) Basic workflow of solid-phase peptide synthesis and traditional protecting-group strategies; (b) ynamide-mediated inverse peptide synthesis strategy [29]; (c) photoredox-based peptide deprotection strategy [32]; (d) automated fast-flow peptide synthesis [35].

与减少强酸相关废液等方面更具优势[32~34].

除了在保护/脱保护的改进, 合成平台本身也在经历重要升级. Pentelute等[35~37]报道了自动化流体多肽合成(automated fast-flow peptide synthesis, AFPS)体系, 显著提升了SPPS的效率和可扩展性. 在AFPS中, 脱保护、氨基酸活化、偶联以及树脂洗涤等步骤都在连续流固定床反应器中完成. 通过优化反应温度、试剂浓度等参数, 单个氨基酸残基的合成循环可缩短至约2~3 min, 使得几十到上百个氨基酸组成的长肽乃至单域蛋白质链能够在数小时内化学合成完成(图1d). 与传统的批式固相多肽合成(batch SPPS)相比, AFPS不仅缩短了合成时间, 还降低了树脂上肽链聚集和扩散受限带来的效率损失, 为大尺寸蛋白质的全合成提供了强有力的技术基础.

总体而言, 利用SPPS可以有效制备中短长度的多肽片段, 并通过方法学创新, 在合成绿色化与简捷化方面取得了进步. 然而, 在合成长链或结构复杂的蛋白质时, 每步反应收率的累积效应、树脂上的聚集现象、副反应等会导致产率和纯度随链长增加而急剧下降[5], 因此, 收敛式合成(convergent synthesis)代替线性合成(linear synthesis)策略成为必然选择, 即先通过SPPS或生物表达获多肽片段, 再借助高效的连接反应组装成目标蛋白质. 多肽片段连接反应也因此成为连接固相合成与完整蛋白质构建的关键策略.

2.2 多肽片段连接

多肽片段连接反应通常是指在水相条件下, 将两个或多个无保护多肽(或蛋白质)片段通过共价键拼接成完整主链的过程, 是收敛式蛋白质全合成与半合成的核心技术[9~11]. 根据反应机理, 主要可分为化学连接与酶促连接两大类. 前者依赖特异性的末端官能团或辅助基团, 在近中性条件下形成酰胺键; 后者则利用酶或内含肽在温和生理条件下催化酰胺键的形成.

2.2.1 化学连接技术

Kent等[9]发展的自然化学连接(native chemical ligation, NCL)是无保护多肽段片段缩合技术的里程碑. 该方法基于C端硫酯片段与N端Cys片段之间的可逆硫酯交换及随后的S→N酰基转移, 在近中性水溶液中生成天然肽键(图2a)[9]. 相较于早期依赖侧链全保护的缩合方法, NCL具有极高的化学选择性, 可兼容各类未保护侧链, 极大地简化了复杂蛋白质的合成路线. 随后, 借助临时添加的辅助官能团(如巯基等)、硒代半胱氨酸及脱硫技术, NCL的适用位点已从Cys扩展至丙氨酸(Ala)、缬氨酸(Val)、赖氨酸(Lys)等多种氨基酸残基[10,11,38].

在此基础上, Muir等[39,40]将NCL与内含肽(intein)介导的蛋白质剪接(protein splicing)相结合, 发展了表达蛋白连接(expressed protein ligation, EPL)技术, 实现了重组表达与化学合成技术的有效衔接[18,41]. EPL的核心原理与NCL一致, 但在硫酯片段的获取上, EPL是将目标蛋白质片段与内含肽融合表达, 通过内含肽自剪接在体外生成带有C端硫酯的表达片段, 再与合成的N端为Cys的多肽段发生NCL反应, 从而实现重组片段与合成片段的无缝连接(图2b)[12].

多肽酰肼连接技术(peptide hydrazide ligation)是由Liu等[42]发展的一种高效、稳健的肽段连接策略, 其在蛋白质化学合成中具有重要地位. 如图2c, 该方法以C端酰肼作为稳定的硫酯前体, 在水相中通过亚硝酸氧化将酰肼转化为高活性的酰基叠氮, 随后原位转化为硫酯, 进而与N端Cys肽段进行NCL反应. 这一“酰肼→叠氮→硫酯”的转化路径, 有效规避了在Fmoc-SPPS中直接制备肽硫酯的不稳定性问题, 尤其在长肽与疏水肽段的合成中展现出显著优势, 并兼容从N端到C端的多肽片段缩合. 多肽酰肼连接技术也已成为重要化学连接方法, 在复杂蛋白质的制备与修饰中发挥着关键作用[7,43,44].

针对序列中缺乏Cys或不宜在Cys处断开的问题, 研究者开发了一系列与NCL互补的连接方法. 例如, Li等[45,46]开发的丝氨酸/苏氨酸连接(Ser/Thr Ligation, STL)通过在C端引入水杨醛酯, 在酸性条件下与N端Ser/Thr形成N,O-缩醛中间体并重排生成肽键(图2d). 该策略以Ser/Thr为连接位点特征, 为Cys非依赖的片段拼接提供了重要选择, 并在相应序列设计中展现出良好的实用性[47].

与此同时, 为进一步提升断点选择的通用性, 近年来逐步形成了一类不依赖连接位点氨基酸侧链官能团参与的片段连接思路, 通常通过片段端基引入可编程反应基团来实现主链肽键构建[48]. 例如, Bode等[49~51]开发的α-酮酸-羟胺连接(α-ketoacid-hydroxylamine (KAHA) ligation)利用α-酮酸与*N*-烷基羟胺在水相中直接缩合生成酰胺键, 具有良好的官能团兼容性(图2e).

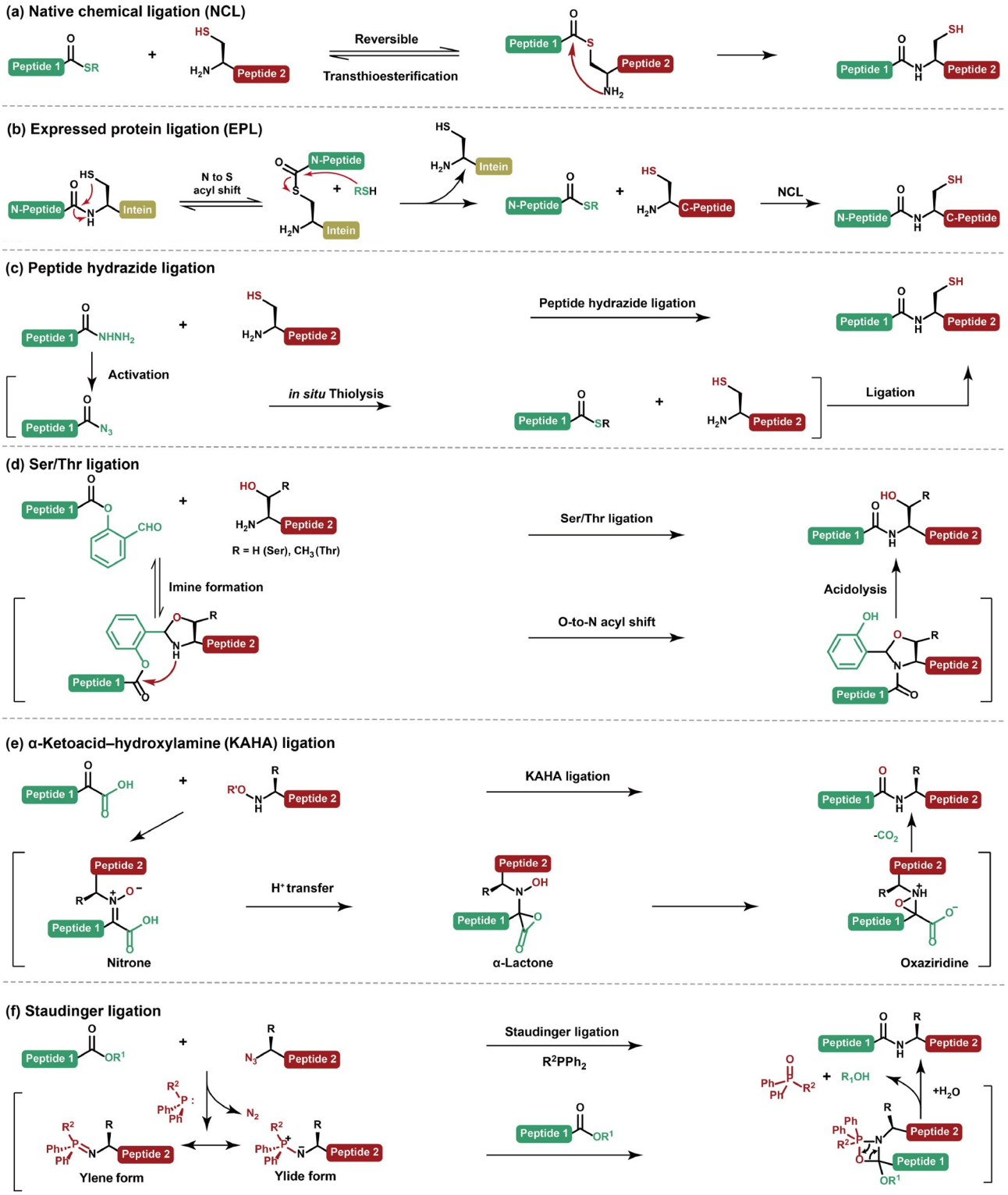


图 2　(网络版彩图)多肽片段化学连接方法的原理. (a) 自然化学连接[9]; (b) 表达蛋白连接技术[41]; (c) 多肽酰肼连接[42]; (d) 丝氨酸/苏氨酸连接[45]; (e) α-酮酸/羟胺脱羧缩合[49]; (f) Staudinger连接[53]

Figure 2　(Color online) Principles of peptide chemical ligation methods. (a) Native chemical ligation [9]; (b) expressed protein ligation [41]; (c) peptide hydrazide ligation [42]; (d) serine/threonine ligation [45]; (e) KAHA ligation [49]; (f) Staudinger ligation [53].

而最常用的羟胺构建模块(S)-5-氧杂脯氨酸((S)-5-oxaproline), 会在连接位点生成高丝氨酸残基. Bode等[52]则是继续改良, 开发了环肽衍生的羟胺构建模块, 实现了生成天然残基的连接. Staudinger连接和点击化学辅助连接也被用于构建主链或主链-侧链桥接, 特别适用于引入可追踪标记或特殊修饰单元(图2f)[53~55]. 此外, Liu等[56]报道了以氨酰基-*N*-羟基肽(aminoacyl-*N*-hydroxy peptide, AAHO peptide)为特征的硫酯连接策略, 通过N-OH参与实现片段高效拼接. Zhao等[57]则提出醛捕获连接(aldehyde capture ligation, ACL), 同样不涉及连接位点侧链官能团的参与. 他们利用醛基对N端氨基的捕获并在N-OH辅助下形成1,2,5-氧杂二氮杂环己烷(1,2,5-oxadiazinane)中间体, 通过酰基转移和除去辅助基团即可构建酰胺键. 总之, 这些方法为在更广泛的断点上实现长肽与蛋白质的片段化学连接提供了有效补充.

在应用层面, 化学连接与SPPS结合已被用于构建多种具有各类定点修饰的蛋白质[7,8]. 例如, 针对α-突触核蛋白质(α-synuclein), Li等[58~60]通过SPPS合成修饰片段, 利用NCL组装获得具有磷酸化、氧连接的*N*-乙酰葡糖胺(O-GlcNAc)糖基化等特定修饰的全长蛋白质, 揭示了不同修饰对蛋白质纤维结构、细胞间传播及构象特征的影响. 此外, Liu等[43,61]将多肽酰肼连接的策略用于合成镜像蛋白质和d-多肽配体, 为镜像生物学研究及抗蛋白质酶解的d-多肽药物开发提供了有效途径. 他们还合成了多种非典型多聚泛素链, 将其作为研究E3泛素连接酶介导蛋白降解的重要化学探针[62,63].

2.2.2 酶促连接技术

酶促多肽连接体系类型多样, 通常具有水相温和条件、反应选择性高、对复杂官能团和生物体系兼容性好等优势, 可作为化学连接的重要补充, 尤其适用于蛋白质的定点功能化与复杂体系中的构建[64]. 鉴于酶促连接体系类型多样, 目前已有专门的综述[7,64,65]进行系统总结, 以下仅选取较为成熟的转肽酶(transpeptidase)作简要介绍. 总体而言, 转肽酶主要可分为两大类: Sortase与天冬酰胺内切酶(asparaginyl endopeptidase, AEP).

以细菌来源的Sortase A为代表的Sortase类转肽酶, 能够特异识别底物C端LPXTG基序, 并在切割位点形成硫酯中间体, 随后由带N端寡聚甘氨酸的受体底物进攻完成转肽反应, 从而实现蛋白质N/C端的定点标记与功能化(图3a)[65~67].

AEP类转肽酶中, Tam等[68]报道的Butelase-1是典型代表. Butelase-1通过识别底物中Asx (天冬酰胺(Asn)/天冬氨酸(Asp))相关序列并形成硫酯中间体, 随后由特定序列的N端氨基进攻完成转肽, 从而实现片段连接或位点特异性修饰(图3b). 在此基础上, Tam与Liu等[69~71]进一步发展了pH调控的正交连接、级联酶促策略以增强转肽驱动力等方法学, 并将该类连接拓展至如非Asx位点连接等复杂的修饰与应用场景.

此外, 来源于*Oldenlandia affinis*的OaAEP-1也是一种重要的AEP类转肽酶, 其催化过程与Butelase-1基本相似(图3c)[64]. 在应用研究方面, Craik等基于OaAEP-1建立了多种蛋白质偶联与修饰策略, 包括蛋白质的位点特异性标记[72]、C端到C端(C to C)的非传统蛋白连接构型[73]以及将反应进一步拓展至赖氨酸ε-氨基的靶向酰化等位点类型拓展[74].

2.3 遗传密码子扩展

Schultz等[75~77]发展的遗传密码子扩展技术通过引入正交的氨酰-tRNA合成酶/tRNA对(aminoacyl-tRNA synthetase/tRNA pair, aaRS/tRNA pair), 并重新编码终止密码子(通常为琥珀密码子UAG), 使非天然氨基酸(noncanonical amino acid, ncAA)能够在翻译过程中被定点插入蛋白质多肽链中(图4). 与常规重组表达相比, 该策略能在活细胞内、不干扰内源蛋白质合成的前提下, 在目标蛋白质的特定位点精准引入携带叠氮、炔烃、金属螯合基团或光敏基团的功能性残基. Wang等[78]将含二苯甲酮发色团的光敏型ncAA插入脱卤酶获得了可在整体细胞中催化芳基卤代物氢化/氘化的人工光脱卤酶. Liu等[79]设计了分子转子型荧光氨基酸, 通过遗传编码, 可将多种目标蛋白直接“转化”为荧光蛋白和生物传感器; Chen等[80]则发展了笼化色氨酸及其四嗪触发的逆电子需求Diels-Alder反应脱笼化, 实现了在活细胞内对关键色氨酸残基的位点特异性激活与功能调控. 近期我们团队[81]也利用GCE生物合成了含膦基蛋白质, 可在蛋白质特定位点构建非天然金属中心, 为人工金属酶的设计提供了新思路.

遗传密码子扩展不仅极大丰富了蛋白质的侧链化

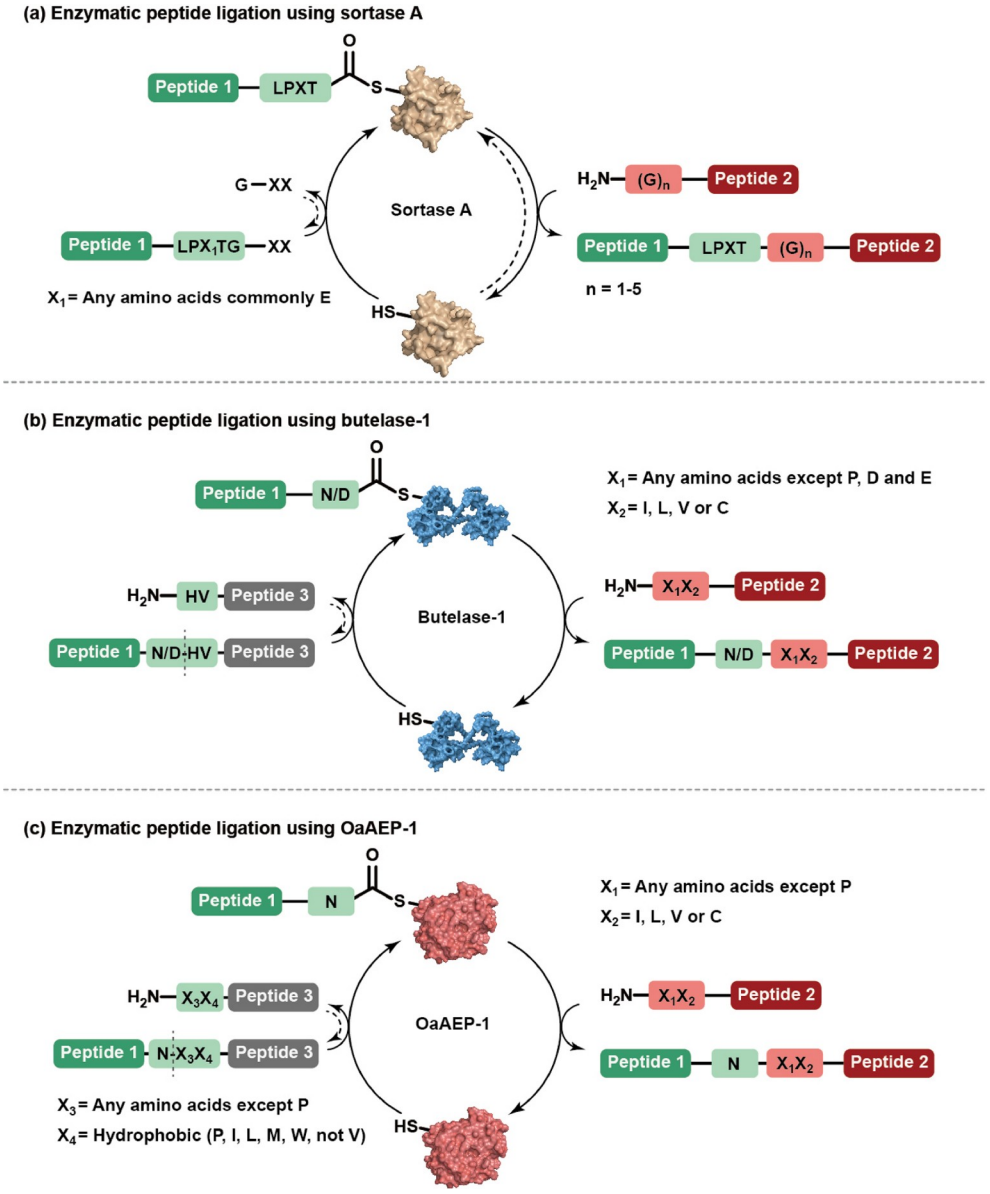


图 3　(网络版彩图)酶促多肽连接技术. (a) Sortase A; (b) Butelase-1; (c) OaAEP-1[64]

Figure 3　(Color online) Enzymatic peptide ligation. (a) Sortase A; (b) Butelase-1; (c) OaAEP-1 [64].

学空间, 更为在主链层面进行精确操作埋下了伏笔. 从方法学演化的角度看, 遗传密码子扩展技术使在核糖体翻译阶段将“可编辑前体单元”精准嵌入蛋白质成为可能. 这些预置的单元可作为特定位点的化学锚点, 进而在翻译后阶段, 通过化学或酶学方法进行精确触发, 从而介导蛋白质主链的骨架重构或功能片段替换. 此方法构成了后续章节将重点深入探讨的蛋白质编辑技术的核心基础之一.

3　蛋白质编辑

近年来, 基因编辑(gene editing)技术的发展实现了对生命遗传物质(主要是DNA)的精确插入、替换与敲除[82,83]. 受此通过原位编辑直接调控生物功能的思路启发, 相关研究自然延伸至蛋白质领域. 与直接改写遗传密码的基因编辑不同, 蛋白质编辑着眼于对已翻译生成的蛋白质进行原位修饰. 广义上, 蛋白质编辑

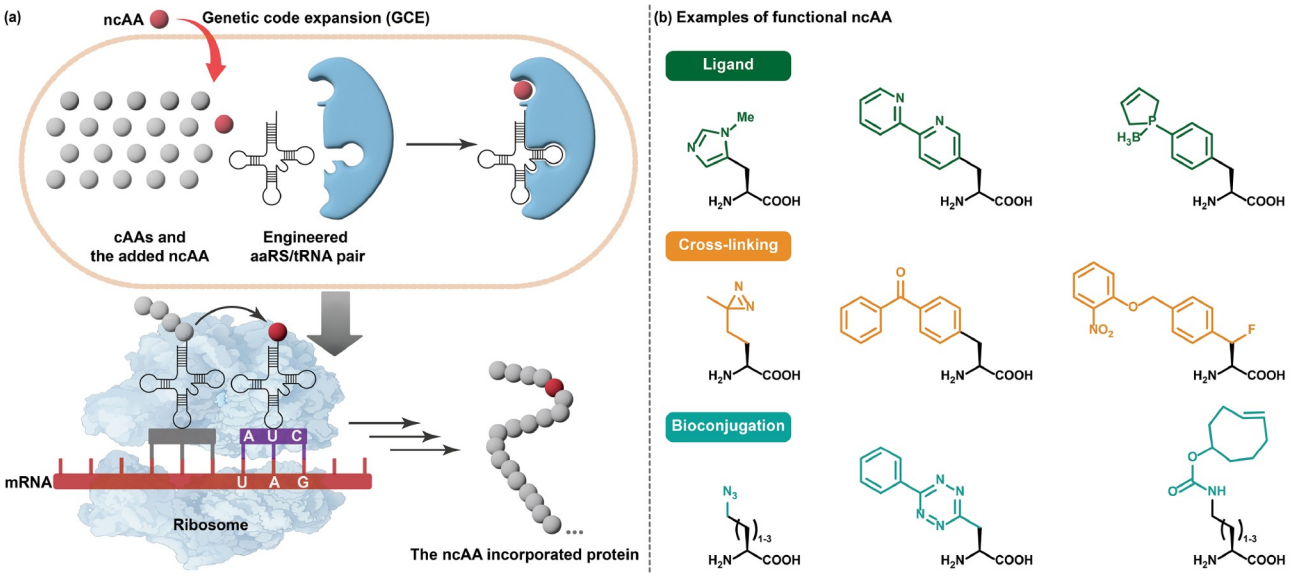


图 4 (网络版彩图) (a) 遗传密码子扩展技术基本原理以及(b) 基于该技术插入蛋白质不同功能化的非天然氨基酸实例[75~77]

Figure 4 (Color online) (a) Basic principle of genetic code expansion technique and (b) examples of functional ncAAs incorporated into protein via GCE [75–77].

既包括侧链编辑, 即改变特定氨基酸侧链的化学结构或翻译后修饰状态[84~87]; 也包括主链编辑, 即直接改写蛋白质骨架, 涉及主链原子或键型的改变以及序列片段的替换[14~16].

近几年, 侧链层面的蛋白质编辑已经出现了许多代表性工作. 例如, Wu等[88]通过将肽段与*N*-糖苷酶(PNGase)融合, 在活体哺乳动物细胞中实现了膜蛋白*N*-糖基化位点的选择性去糖基化, 并将相应天冬酰胺残基原位转化为天冬氨酸. Davis等[89]则发展了立体保持的翻译后蛋白质编辑方法, 通过在温和条件下构建蛋白质上的丙氨酰Cβ·自由基, 实现多种侧链构建, 并保持立体构型, 从而在化学层面编辑侧链结构并系统探索其结构-功能关系. Wang等[90]利用黄素依赖半胱氨酸脱羧酶, 在温和水相条件下对底物肽段C端Cys残基进行氧化脱羧, 原位生成高反应性的C端硫醛(thioaldehyde), 并将其作为化学手柄, 实现多肽和蛋白的生物偶联(bioconjugation)与大环化(macrocyclization). Chen等[91]提出了基于吡啶三唑甲醛(pyridotriazole carbaldehyde)的选择性生物偶联策略, 可在温和水相条件下选择性地高效修饰蛋白质N端及赖氨酸侧链伯胺, 为精确构建蛋白质-小分子偶联物和侧链层面的功能编辑提供了通用工具. 关于侧链修饰或者侧链编辑的工作已有专门的综述[84~87]进行了系统总结, 本文将不对侧链编辑技术再做论述, 而是将讨论范围限定在主链层面的蛋白质编辑.

相较于侧链修饰, 主链层面的改动更易影响蛋白质的稳定性、动力学性质及生理功能, 因此对位点选择性与结构完整性的控制要求极高. 蛋白质主链编辑主要包括两类策略: (1) 对主链本身的可逆修饰; (2) 对骨架原子、键型或序列片段的直接改写. 典型的主链骨架可逆修饰为疏水性蛋白骨架的可逆修饰, 其通过在主链酰胺上暂时引入可去除的亲水基团或临时结构支撑单元, 实现辅助增溶、抑制聚集, 并促进肽段纯化与片段的高效连接; 随后脱除修饰以恢复天然骨架, 从而显著促进膜蛋白以及易聚集的高疏水性蛋白的化学合成[20,92].

关于主链修饰, 已有综述[93,94]从保护基与临时结构支撑等角度进行了系统归纳. 而本文重点关注对骨架原子、键型或序列片段的直接改写. 随着多肽骨架直接转化反应、遗传密码子扩展技术及分裂内含肽体系的发展, 此类蛋白质主链编辑的策略逐渐清晰, 主要可分为两条路线: 一是骨架重构(backbone reconstitution), 通过酶催化或者成环、重排等化学反应实现局部主链原子或键型的变化[16,95~99]; 二是片段替换(segment replacement), 借助分裂内含肽等工具实现序列模块的整体置换[14,15].

3.1　骨架重构

骨架重构旨在不改变蛋白质整体序列长度的前提下, 直接改造主链连接方式. 这种主链上的细微改动往往能显著调控多肽的构象倾向、亲脂性及代谢稳定性等[13]. 从技术发展路径来看, 骨架重构最初主要在多肽水平实现, 而对完整蛋白质主链进行定点编辑则是近期才取得的重要突破[16]. 在多肽骨架重构中, 存在两种重要的形式, 一种是通过酶催化对天然的多肽进行重构[95~97], 另一种是通过特殊的化学反应对多肽骨架进行重构[98,99].

关于酶催化的骨架重构, 主要指对核糖体合成并经翻译后修饰的多肽(ribosomally synthesized and post-translationally modified peptide, RiPPs)的酶催化骨架重构, 直接作用于–(NH–C_α–CO)$_n$–重复骨架单元, 包括但不限于脱水、加成、环化、硫酰胺化、主链断裂与再连接以及肽剪接反应等[100]. 例如, Piel等[95,96]对自由基S-腺苷甲硫氨酸(radical SAM)肽剪接酶(splicease)的进行相关机理研究. 这类酶可以在RiPP前体肽内部的酪氨酸(Tyr)位点上生成α-酮-β-酰胺(α-keto-β-amide), 实现了在核糖体框架下无法直接完成的α-酮-β-酰胺形成(图5a), 其本质是一种由肽剪切酶PcpXY催化的主链剪接与重排.

在骨架重构类型中, 硫酰胺(thioamide)取代是骨架改变最小的一类主链编辑策略, 其结构特征是将单个肽键羰基的O替换为S, 在尽量保持主链几何构型的前提下, 改变该肽键的电子结构与非共价相互作用. Raines等[101]从相邻羰基之间的n→π*相互作用出发, 讨论了酰胺与硫酰胺在电子效应层面的差异及其对蛋白稳定性的意义. Mitchell等[97]通过体外酶学重构, 系统解析了YcaO酶将肽主链酰胺转变为硫酰胺的机理. YcaO通过ATP依赖的O-磷酸化活化肽键, 再由外源硫化物(如Na_2S)进攻并取代酰胺氧, 生成主链硫酰胺单元(图5b), 并显著提高肽链对蛋白酶的抗降解能力. 除了RiPPs的酶促改写, Hecht等[102]还证实, 在体外体系中核糖体能够直接并入二肽及其类似物. 这项工作为利用预先构建的骨架单元来拓展翻译体系、使其兼容非天然主链结构奠定了基础. 在此基础上, Suga等[103]进一步利用核酶Flexizyme, 制备了硫酰胺取代的氨酰tRNA, 并成功在翻译过程中, 通过核糖体合成了含有硫酰胺键的多肽, 为硫酰胺主链的编辑提供了一条可行的生物合成路径.

关于环化、重排等化学反应实现的多肽骨架重构近期也涌现出了许多工作[13]. 在硫酰胺取代方面, 化学合成方法在底物普适性与编辑位点选择性上表现出

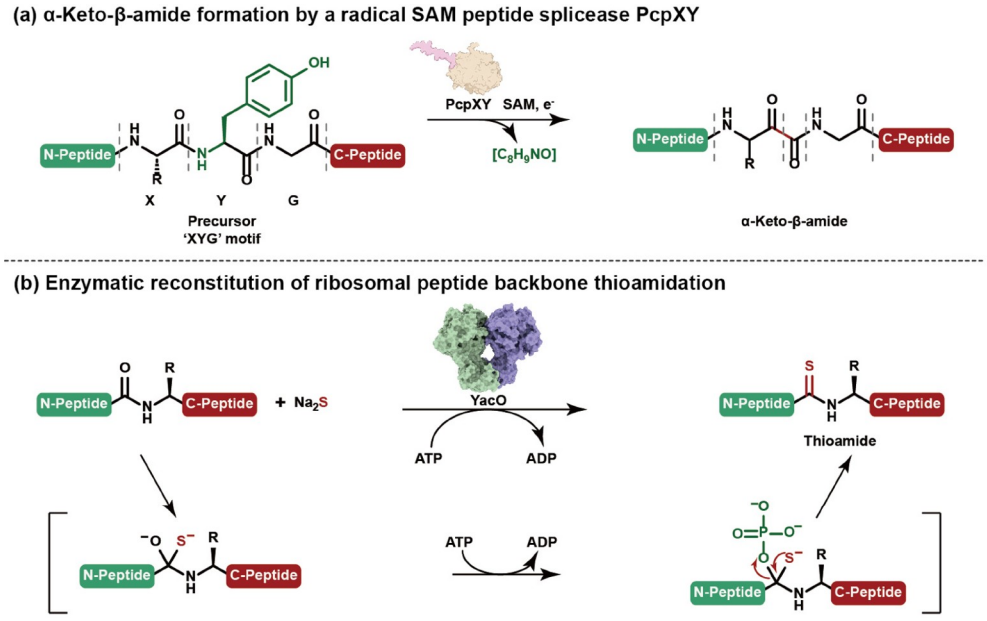


图 5　(网络版彩图)酶催化的骨架重构. (a) PcpXY催化形成α-酮-β-酰胺骨架[95,96]; (b) YacO催化实现骨架酰胺转变为硫酰胺[97]

Figure 5　(Color online) Enzyme-catalyzed backbone reconstitution. (a) PcpXY-catalyzed formation of an α-keto-β-amide backbone [95,96]; (b) YacO-catalyzed conversion of a backbone amide to a thioamide [97].

显著优势. Petersson等[104]通过化学合成, 在典型蛋白质二级结构中定点引入硫酰胺键, 并系统评估了其对蛋白质稳定性的影响. 随后, 他们[105]进一步将这一策略应用于治疗性肽激素体系, 证明硫酰胺替换不仅能选择性调控蛋白酶切割过程, 还能影响受体相关活性, 为优化肽类分子的性质提供了直接依据. 在方法学上, Zhao等[106]发展了一种基于炔酰胺的新型硫酰胺键构建反应. 如图6a, 该方法利用单硫代羧酸先与炔酰胺反应, 生成稳定、可分离的硫酰化活化中间体. 该中间体随后再与肽段的胺基发生硫酰化反应, 从而实现了硫酰胺肽键的精准定点构建.

针对其他化学反应实现的骨架重构, Yang等[98]报道了一种无需过渡金属催化的N端脂肪族侧链环化骨架编辑反应. 该方法先将多肽N端氨基氯化, 在光照射下诱导N–Cl键均裂产生氮自由基, 经1,5-氢原子转移(1,5-hydrogen atom transfer, 1,5-HAT)生成碳自由基, 并被氯原子捕获形成C–Cl键, 最后在碱性条件下发生分子内亲核取代, 构建出类似脯氨酸的环状骨架(图6b). Francis与Schepartz等[99]则利用翻译后O→C酰基迁移反应实现了骨架重构. 他们利用核糖体或SPPS在多肽中引入α-羟基苯基硒代半胱氨酸, 通过氧化原位生成脱氢乳酸(dehydrolactic acid, DHL)单元, 构建含O-酰基片段的前体. 在温和条件下, 该片段自发发生O→C酰基迁移, 生成嵌入主链的α,γ-二酮酰胺(α,γ-diketoamide)结构(图6c). 这一重排过程将一段常规肽键改写为α,γ-二酮酰胺, 不仅引入了新的亲核反应位点, 还改变了N–H键的氢键供体性质, 有望对多肽的稳定性、膜通透性及后续修饰潜力产生影响.

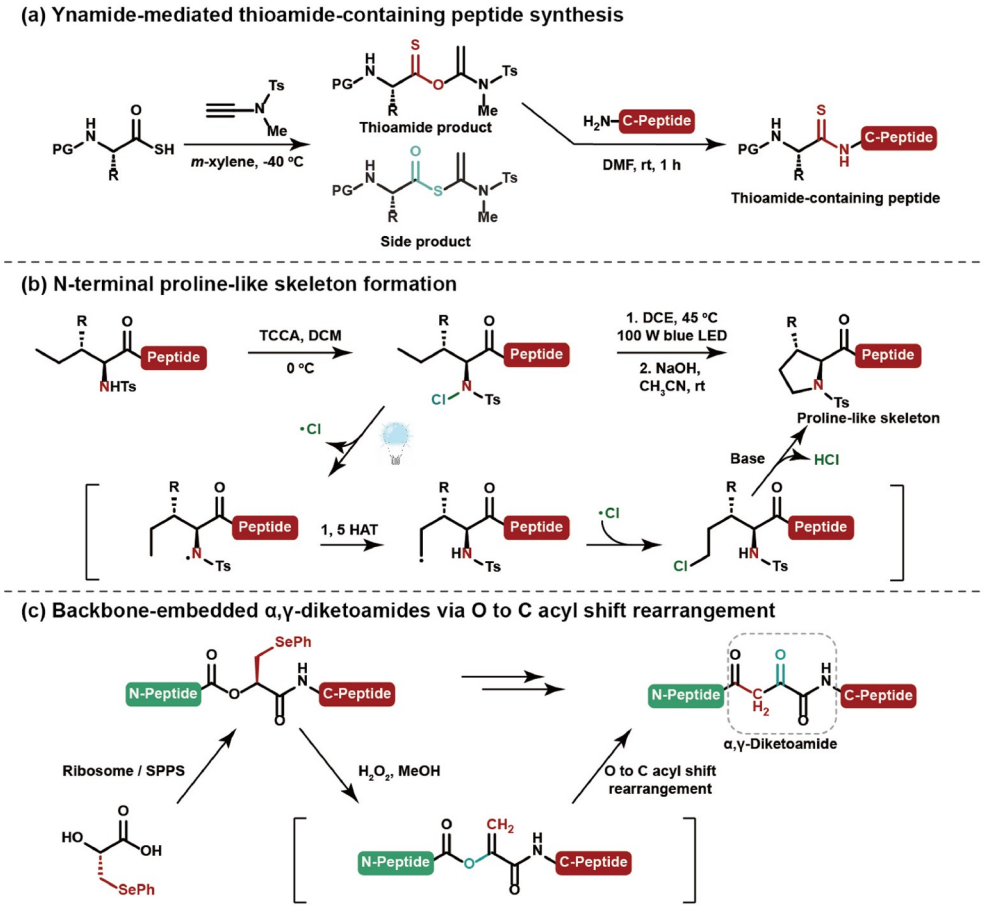


图 6 (网络版彩图)化学反应实现骨架重构. (a) 炔酰胺介导的硫酰胺多肽合成[106]; (b) 光触发的N端类脯氨酸结构形成[98]; (c) 重排实现α,γ-二酮酰胺骨架形成[99]

Figure 6 (Color online) Backbone reconstitution enabled by chemical reactions. (a) Ynamide-mediated thioamide-containing peptide synthesis [106]; (b) light-promoted formation of an N-terminal proline-like skeleton [98]; (c) rearrangement to form an α,γ-diketoamide backbone [99].

总体而言，现有策略的验证与应用范围主要集中于短肽或前体肽层面. 一方面，相关反应往往依赖于有机溶剂、高当量试剂或光化学活化等条件，难以直接推广至折叠状态下的完整蛋白质或活细胞环境；另一方面，RiPP相关酶通常对底物序列、前导肽结构及局部构象有较高要求，且其催化活性多在线性或近线性肽段上表现最优，因而在常规蛋白质复杂三维结构中的普适性与位点可及性受到显著限制. 这些因素共同制约了现有骨架重构方法向常规蛋白质体系的直接应用，使其在实现普适性蛋白质主链编辑方面仍面临明显挑战.

而真正在蛋白质层面上实现骨架重构的方法则是由Francis、Miller和Schepartz合作[16]提出的基于骨架延伸酰基重排(backbone extension acyl rearrangements, BEAR)反应. 如图7所示, 该方法首先利用GCE技术, 在细胞内将携带掩蔽(mask)亲核基团的α-羟基酸定点引入目标蛋白质主链；随后通过光照或化学手段去除掩蔽基团，暴露侧链亲核中心；侧链亲核中心进攻邻近的酰基，形成五、六或七元的环状酰胺中间体，诱发邻近酰基发生分子内转移(O→N)，将局部的α-酰胺键原位转化为β-、γ-或δ-主链片段. 密度泛函理论(DFT)计算结果表明，这些酰基迁移过程在热力学上相对常规α-酰胺键更为有利且符合Baldwin规则所预测的有利环尺寸，使得BEAR能够在水相、接近中性的缓冲

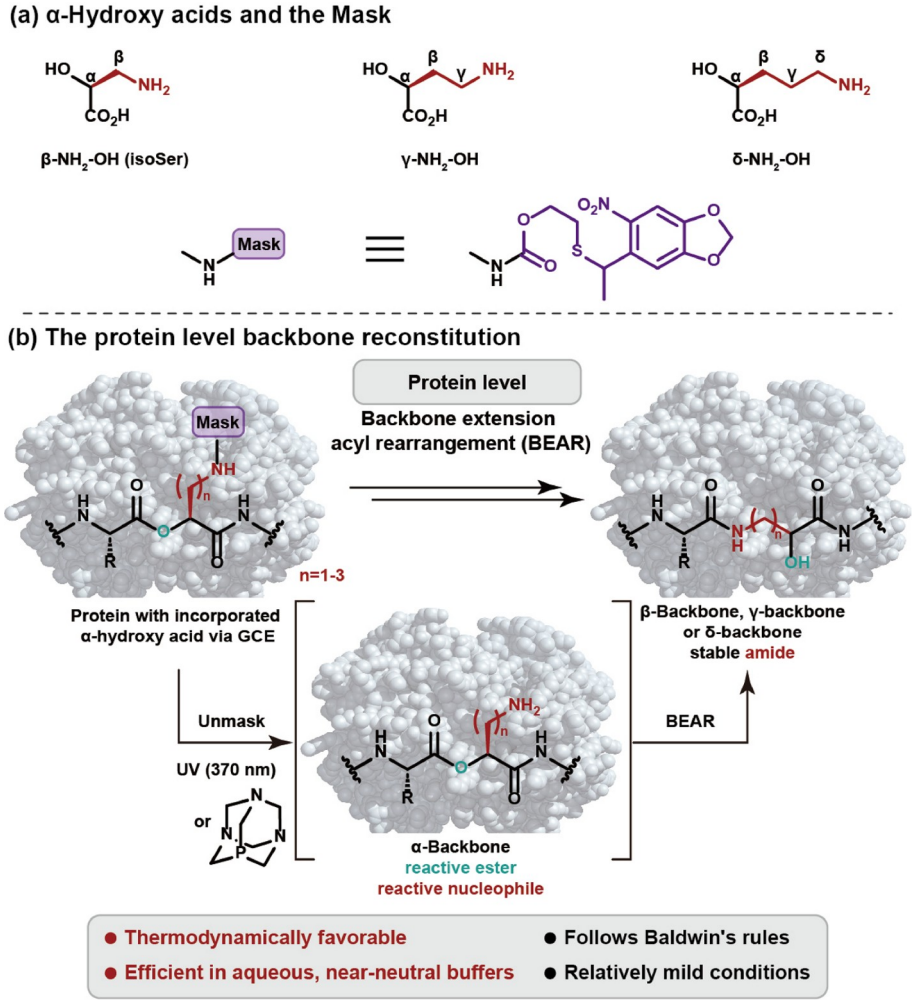


图 7　(网络版彩图)蛋白质骨架延伸酰基重排反应. (a) 所使用的α-羟基酸及掩蔽基团; (b) 骨架延伸酰基重排实现了α-酰胺键原位转化为β-、γ-或δ-主链片段[16]

Figure 7　(Color online) Backbone extension acyl rearrangement of protein. (a) α-Hydroxy acids and the mask group used; (b) BEAR enables *in situ* conversion of an α-amide bond into β-, γ-, or δ-backbone segments [16].

液中高效进行且无需大量有机溶剂、强氧化/还原剂或极端pH的参与. 这一过程完全在翻译后进行, 无需对核糖体进行复杂的工程化改造, 仅依赖现有的GCE体系即可在活细胞表达的全长蛋白质中定点插入非天然结构单元.

无论是基于化学合成的多肽/蛋白质骨架重构, 还是源自RiPP天然产物的酶催化骨架重构, 都遵循着统一的逻辑, 即预置可编辑前体单元或可识别序列, 再通过特定条件诱导反应完成骨架重构. 与传统的全合成策略相比, 这种方法能够在更接近天然生理环境的背景下, 对局部主链结构进行最小干预式的精准调整, 为系统解析骨架原子改变对蛋白质结构与功能的影响提供了强有力的工具.

3.2 片段替换

片段替换主要依赖分裂内含肽介导的蛋白质反式剪接(protein trans-splicing, PTS). 该策略旨在保持蛋白质整体折叠状态的前提下, 将目标蛋白质内部的一段序列切除并替换为携带特定修饰或非天然残基的外源多肽片段. 早期的研究表明, 内含肽(intein)能够在无需辅因子、近生理条件的溶液中实现主链的断裂与再连接, 完成两段外显肽(extein)的无痕拼接(图8a)[39,40]. 不同来源及工程化改造的分裂内含肽在剪接动力学、外肽序列兼容性及无痕特性上各具优势, 为蛋白质片段替换提供了多样化的工具选择[12,17,19].

在此基础上, Pless等[107]率先发展了串联蛋白质反式剪接(tandem protein trans-splicing, tPTS)策略, 利用两对正交(biorthogonal)的分裂内含肽在真核细胞内实现了合成肽片段的插入. 该策略的基本设计原理是: 将目标蛋白质拆分为N-端、中间、C-端三个片段; 其中, N端片段的C末端融合一个分裂肽Int^N-A重组表达, 而C端片段的N末端融合另一个正交分裂肽Int^C-B重组表达; 中间片段的N端和C-端则分别连接剩余的互补内含肽片段Int^C-A和Int^N-B, 体外化学合成后, 通过显微注射等方式导入细胞. 在细胞内, 两对内含肽依次发生转剪接, 将合成肽与重组表达的N-端和C-端片段无痕连接, 完成了全长蛋白的合成(图8b). 该方法成功实现了将含复杂翻译后修饰或非天然氨基酸多肽片段, 在活细胞内插入到了目标N-端和C-端片段之间, 证明了利用正交内含肽进行原位片段替换的可能性. 然而, 该策略仍存在剪接效率有限、半剪接中间体易积累等问题; 且依赖于非洲爪蟾卵母细胞的显微注射操作, 通量较低, 应用范围受到一定限制.

针对效率与中间体问题, Muir等[14]提出了蛋白质转座(protein transposition)反应, 用于在保持蛋白质天然折叠的前提下, 实现目标蛋白内部片段的直接替换. 该策略是在待编辑的折叠蛋白中, 先选定替换区域, 在其编码含Int^N-A与Int^C-B的片段, 而将待替换片段两侧分别嵌入与之互补的Int^C-A和Int^N-B, 构建出“转座子(transposon)”. 当待编辑蛋白与“转座子”相遇后, 两对分裂内含肽先特异性结合并发生蛋白质反式剪接, 从而实现特定片段的剪切到粘贴, 并在无需经历整体解折叠和再折叠的条件下, 直接得到携带替换片段的编辑产物(图8c). 该研究的关键在于其筛选出了剪接动力学匹配的正交内含肽, 可实现更接近协同的转座过程, 从而减少半剪接中间体积累并提升对折叠底物的适用性. 整体而言, 蛋白质转座说明在折叠蛋白乃至大分子复合物内部进行高效片段替换是可行的, 为在近天然条件下对折叠蛋白进行内部片段替换提供了新路线. 在此基础上, Muir等[108]进一步将tPTS用于细胞表面可编程蛋白连接, 提出了基于靶向识别的剪接调控执行器(splicing-modulated actuation upon recognition of targets, SMART). 与蛋白质转座反应不同的是, 他们设计了条件触发式蛋白质剪切(conditional protein splicing), 笼化的分裂内含肽只能在近邻时才能发生剪切. SMART系统则是将两段彼此失活的蛋白片段分别与被笼化的分裂内含肽片段及不同的细胞表面靶向模块融合, 使其仅在目标细胞表面因近邻会聚而触发反式剪接, 从而在细胞表面原位拼接生成活性蛋白. 该体系可按预设的细胞表面抗原组合执行类似与门(and gate)和非门(not gate)的逻辑选择性, 并且生成的蛋白质产物既可保持膜表面关联, 也可按设计释放至胞外环境[108].

Burslem和Shalem等[15]则进一步将片段替换拓展至活细胞内的内源蛋白质编辑. 如图8d, 他们使用CRISPR/Cas9基因编辑技术, 首先在基因层面将Int^N-A和Int^C-B的DNA序列嵌入目标蛋白(protein of interest, POI)基因的内含子中, 使细胞表达出带有Int^N-A和Int^C-B的内源蛋白质, 将其作为目标蛋白-分裂肽受体(POI-intein acceptor). 随后, 通过电穿孔(electroporation)或脂质体递送等方式将携带替代片段的重组内含肽供体(recombinant intein donor bearing cargo)导入细胞, 该供

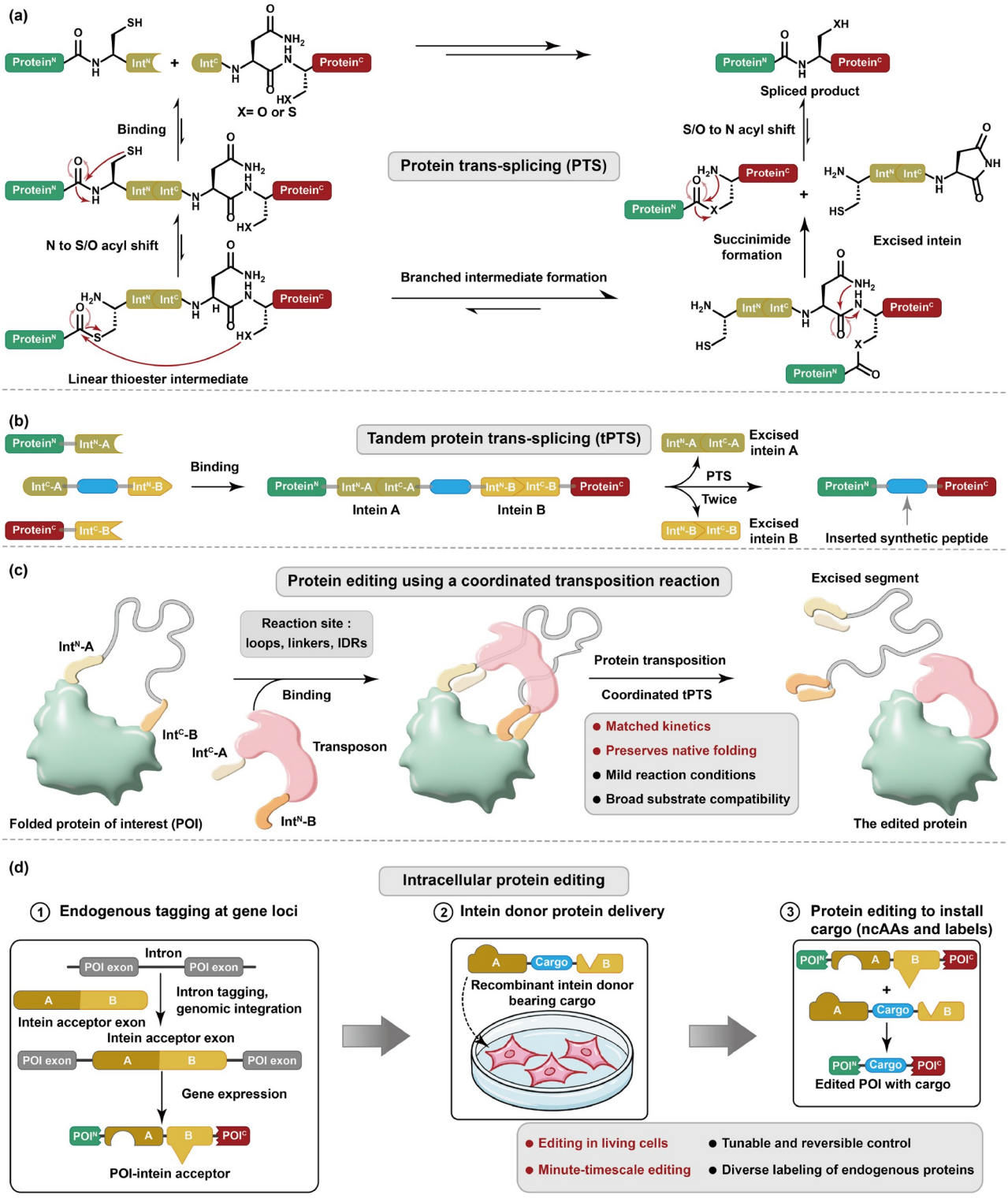


图 8　(网络版彩图) (a) 分裂内含肽及蛋白质反式剪切的基本原理[12]; (b) 串联蛋白质反式剪接[107]; (c) 蛋白质转座反应[14]; (d) 细胞内实现多肽片段替换[15]

Figure 8　(Color online) (a) Basic principle of split inteins and protein trans-splicing [12]; (b) tandem protein trans-splicing [107]; (c) protein transposition reaction [14]; (d) peptide segment replacement in cells [15].

体两端携带与受体互补的Int^C-A和Int^N-B的和修饰模块(如含非天然氨基酸、有用标签等). 当供体进入细胞后, 受体与供体自发识别并发生蛋白质反式剪接, 实现对目标内源蛋白特定位点片段的无痕替换. 经过实验验证, 该编辑技术能在10 min甚至更短的时间内完成. 这一技术不仅能在多种细胞内源蛋白质上安装标签, 更能引入难以通过遗传编码获得的化学修饰, 且编辑过程通常在分钟量级内完成. 与不可逆的DNA水平编辑相比, 这种蛋白水平的编辑具有时间尺度短、可逆性更高和靶蛋白特异性更强等特点, 为在生理环境下精细操控内源蛋白一级结构及其修饰状态提供了新思路.

整体而言, 基于分裂内含肽的片段替换策略与骨架重构在功能上实现互补, 前者通过模块化“剪切-粘贴”在较大尺度上替换蛋白质序列片段, 适于引入复杂修饰或非天然残基片段; 后者则多在局部范围内通过成环、重排或酰基迁移改写主链原子连接方式. 随着高效的工程化分裂内含肽的不断出现以及与CRISPR、蛋白递送和遗传密码子扩展等技术的进一步整合, 片段替换有望从当前主要依赖体外, 特定细胞系统或特定结构区域的应用, 逐步走向在内源靶标、更广泛细胞类型和空间构象的普适蛋白质主链编辑.

4 总结与展望

蛋白质合成与编辑方法的发展, 呈现出从体外化学全合成/半合成向细胞内原位编辑逐步演进的趋势. 回顾蛋白质合成的技术发展脉络, 大体归纳为三个相互衔接的阶段: 第一阶段是由液相多肽合成到固相多肽合成的演化, 以此为基础实现了多种含修饰或困难多肽片段的精准制备[3~5,24]; 第二阶段是利用化学与酶促多肽片段连接技术, 解决了片段间的精准拼接问题, 实现了中等至大尺寸蛋白质的收敛式合成[9~11,25]; 第三阶段是基于骨架反应、遗传密码子扩展技术及分裂内含肽体系的蛋白质编辑技术, 使操控层面推进至翻译后层面, 从构建新蛋白质转变到改造既有蛋白质, 逐步实现了复杂生物环境中蛋白质主链的原位定点编辑. 就本文着重论述的蛋白质编辑而言, 其核心是在尽量保持整体折叠与功能框架的前提下, 直接对主链原子连接方式或序列模块进行定点改造, 包括两种重要策略: (1) 骨架重构, 即不改变整体序列长度的前提下, 通过化学或酶促反应实现局部主链原子连接或键型的改变; (2) 片段替换, 即借助正交分裂内含肽体系介导的蛋白质反式剪接, 在蛋白质内部实现目标片段的整体替换, 从而引入非天然元件、外源标记或复杂修饰模块.

作为一种编辑技术, 蛋白质编辑的概念受到基因编辑的启发[15]. 基因编辑技术经历了多轮迭代, 特别是随着CRISPR/Cas9体系的发现[109,110], 基因编辑才迅速演变为一项简便、高效且通用的技术, 极大地加速了从基础遗传学研究向临床治疗应用的转化[82,83]. 相比之下, 蛋白质编辑仍处于发展的初期阶段. 现有的策略中, 骨架重构往往受限于苛刻的反应条件[98,99]或单一的改造形式[16], 难以在复杂的生命体系中广泛应用. 而对于片段替换策略, 其适用位点往往局限于表面暴露环(surface exposed loop)、连接结构域(linker domain)或内在无序区域(intrinsically disordered region), 尚未实现任意位点及任意结构的精确替换[14,15], 但其在功能拓展性上展现出了独特的优势. 片段替换策略实现了在活细胞内实现时空可控的蛋白质操控, 并在不造成基因组永久改变的前提下进行高时间分辨率的动态调控[15]. 更重要的是, 该策略能够突破天然遗传编码限制, 引入非天然元件以扩展蛋白质的化学空间[14,15]. 由此可见, 突破片段替换策略的位点限制, 构建通用的片段替换工具, 将是未来蛋白质编辑技术发展的核心趋势, 有望真正实现对蛋白质的任意编辑.

展望未来, 蛋白质编辑走向普适化应用的核心挑战, 在于构建兼容性更高的通用型编辑体系. 这意味着需要更丰富且正交性更强的内含肽对, 使片段替换策略能够适配更多蛋白质对象, 甚至有望在同一细胞、不同蛋白内实现多位点、可编程的组合编辑. 与此同时, 随着人工智能在蛋白质结构预测与设计领域的突破[111~116], 其有望用于快速设计更高活性、正交性更好的内含肽与供体模块, 从而加速蛋白质编辑迈向通用工具体系. 未来或许会出现突破现有分裂内含肽框架的全新编辑体系, 发现或改造出更多可用于蛋白质骨架精准编辑的酶, 以更简便高效的生物工具拓宽应用边界. 长远来看, 通过深度融合人工智能驱动的分子设计、化学合成与原位编辑技术, 我们将有望实现从功能设想到精准构建的完整技术闭环, 最终在分子层面实现对生命大分子结构与功能的精准调控.

参考文献

1 Gong YT, Du YC, Huang WD, Chen CQ, Ge LJ, Hu SQ, Jiang RQ, Zhu SQ, Niu JY, Xu JC, Zhang WJ, Chen LL, Li HX, Wang Y, Lu DP, Ji AX, Li CX, Shi PT, Ye YH, Tang KL, Xing QY. *Sci Bull*, 1965, 16: 941–945 (in Chinese) [龚岳亭, 杜雨苍, 黄惟德, 陈常庆, 葛麟俊, 胡世全, 蒋荣庆, 朱尚权, 钮经义, 徐杰诚, 张伟君, 陈玲玲, 李鸿绪, 汪猷, 陆德培, 季爱雪, 李崇熙, 施溥涛, 叶蕴华, 汤卡罗, 邢其毅. 科学通报, 1965, 16: 941–945]

2 Masui H, Fuse S. *Org Process Res Dev*, 2022, 26: 1751–1765

3 Merrifield RB. *J Am Chem Soc*, 1963, 85: 2149–2154

4 Merrifield B. *Science*, 1986, 232: 341–347

5 Kent SBH. *Chem Soc Rev*, 2009, 38: 338–351

6 Tan Y, Wu H, Wei T, Li X. *J Am Chem Soc*, 2020, 142: 20288–20298

7 Dong S, Zheng JS, Li Y, Wang H, Chen G, Chen Y, Fang G, Guo J, He C, Hu H, Li X, Li Y, Li Z, Pan M, Tang S, Tian C, Wang P, Wu B, Wu C, Zhao J, Liu L. *Sci China Chem*, 2024, 67: 1060–1096

8 Sun Z, Liu H, Li X. *Chem*, 2024, 10: 767–799

9 Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. *Science*, 1994, 266: 776–779

10 Conibear AC, Watson EE, Payne RJ, Becker CFW. *Chem Soc Rev*, 2018, 47: 9046–9068

11 Agouridas V, El Mahdi O, Diemer V, Cargoët M, Monbaliu JCM, Melnyk O. *Chem Rev*, 2019, 119: 7328–7443

12 Thompson RE, Muir TW. *Chem Rev*, 2019, 120: 3051–3126

13 Shi J, Sun T, Yang M. *Org Chem Front*, 2024, 11: 1623–1640

14 Hua Y, Tay NES, Ye X, Owen JA, Liu H, Thompson RE, Muir TW. *Science*, 2025, 388: 68–74

15 Beyer JN, Serebrenik YV, Toy K, Najar MA, Feierman E, Raniszewski NR, Korb E, Shalem O, Burslem GM. *Science*, 2025, 388: eadr5499

16 Roe LT, Piper IM, Schissel CK, Dover TL, Shah B, Hamlish NX, Garapaty AM, Zheng S, Dilworth DA, Wong N, Zhang Z, Chatterjee A, Francis MB, Miller SJ, Schepartz A. *Nat Chem Biol*, 2025, 21: 1621–1630

17 Noren CJ, Wang J, Perler FB. *Angew Chem Int Ed*, 2000, 39: 450–466

18 Muir TW. *Annu Rev Biochem*, 2003, 72: 249–289

19 Mootz HD. *ChemBioChem*, 2009, 10: 2579–2589

20 Li JB, Tang S, Zheng JS, Tian CL, Liu L. *Acc Chem Res*, 2017, 50: 1143–1153

21 Kulkarni SS, Sayers J, Premdjee B, Payne RJ. *Nat Rev Chem*, 2018, 2: 0122

22 Abboud SA, Cisse E, Doudeau M, Bénédetti H, Aucagne V. *Chem Sci*, 2021, 12: 3194–3201

23 Harel O, Jbara M. *Angew Chem Int Ed*, 2023, 62: e202217716

24 Nilsson BL, Soellner MB, Raines RT. *Annu Rev Biophys Biomol Struct*, 2005, 34: 91–118

25 Malins LR, Payne RJ. *Curr Opin Chem Biol*, 2014, 22: 70–78

26 Li FY, Zhang ZF, Voss S, Wu YW, Zhao YF, Li YM, Chen YX. *Chem Sci*, 2020, 11: 826–832

27 Hu J, Sun XM, Su JY, Zhao YF, Chen YX. *Chem Sci*, 2021, 12: 4432–4442

28 Liu D, Wei Q, Xia W, He C, Zhang Q, Huang L, Wang X, Sun Y, Ma Y, Zhang X, Wang Y, Shi X, Liu C, Dong S. *J Am Chem Soc*, 2021, 143: 20216–20223

29 Liu T, Peng Z, Lai M, Hu L, Zhao J. *J Am Chem Soc*, 2024, 146: 4270–4280

30 Li Y, Li J, Bao G, Yu C, Liu Y, He Z, Wang P, Ma W, Xie J, Sun W, Wang R. *Org Lett*, 2022, 24: 1169–1174

31 Zhou Y, Li H, Huang Y, Li J, Deng G, Chen G, Xi Z, Zhou C. *Nat Commun*, 2023, 14: 5324

32 Bai H, Ye F, Purnachandar D, Liu X, Huang P, Wang P. *J Am Chem Soc*, 2025, 147: 34011–34018

33 Sun X, Ye F, Hu D, Wang P. *J Am Chem Soc*, 2025, 147: 48244–48253

34 Ye F, Bai H, Liu X, Xu P, Ding G, Huang P, Zhang X, Yu B, Wang P. *Angew Chem Int Ed*, 2025, 64: e202518002

35 Hartrampf N, Saebi A, Poskus M, Gates ZP, Callahan AJ, Cowfer AE, Hanna S, Antilla S, Schissel CK, Quartararo AJ, Ye X, Mijalis AJ, Simon MD, Loas A, Liu S, Jessen C, Nielsen TE, Pentelute BL. *Science*, 2020, 368: 980–987

36 Simon MD, Heider PL, Adamo A, Vinogradov AA, Mong SK, Li X, Berger T, Policarpo RL, Zhang C, Zou Y, Liao X, Spokoyny AM, Jensen

KF, Pentelute BL. *ChemBioChem*, 2014, 15: 713–720
37 Fittolani G, Callahan AJ, Loas A, Pentelute BL. *Chem Commun*, 2025, 61: 5471–5474
38 Durek T, Alewood PF. *Angew Chem Int Ed*, 2011, 50: 12042–12045
39 Hirata R, Ohsumk Y, Nakano A, Kawasaki H, Suzuki K, Anraku Y. *J Biol Chem*, 1990, 265: 6726–6733
40 Kane PM, Yamashiro CT, Wolczyk DF, Neff N, Goebl M, Stevens TH. *Science*, 1990, 250: 651–657
41 Muir TW, Sondhi D, Cole PA. *Proc Natl Acad Sci USA*, 1998, 95: 6705–6710
42 Fang GM, Li YM, Shen F, Huang YC, Li JB, Lin Y, Cui HK, Liu L. *Angew Chem Int Ed*, 2011, 50: 7645–7649
43 Huang YC, Fang GM, Liu L. *Natl Sci Rev*, 2016, 3: 107–116
44 Yang Z, Xiao Y, Shi Y, Liu L. *Sci China Life Sci*, 2025, 68: 2515–2549
45 Li X, Lam HY, Zhang Y, Chan CK. *Org Lett*, 2010, 12: 1724–1727
46 Lee CL, Li X. *Curr Opin Chem Biol*, 2014, 22: 108–114
47 Liu H, Li X. *Acc Chem Res*, 2018, 51: 1643–1655
48 Yang J, Zhao J. *Sci China Chem*, 2018, 61: 97–112
49 Bode JW, Fox RM, Baucom KD. *Angew Chem Int Ed*, 2006, 45: 1248–1252
50 He C, Kulkarni SS, Thuaud F, Bode JW. *Angew Chem Int Ed*, 2015, 54: 12996–13001
51 Bode JW. *Acc Chem Res*, 2017, 50: 2104–2115
52 Han J, Hirao K, Mikami T, Nötel NY, Seidl LL, Bode JW. *J Am Chem Soc*, 2025, 147: 34238–34243
53 Saxon E, Bertozzi CR. *Science*, 2000, 287: 2007–2010
54 Kofoed C, Riesenberg S, Šmolíková J, Meldal M, Schoffelen S. *Bioconjugate Chem*, 2019, 30: 1169–1174
55 Bednarek C, Wehl I, Jung N, Schepers U, Bräse S. *Chem Rev*, 2020, 120: 4301–4354
56 Pasunooti KK, Balamkundu S, Liu J, Zhang Z, Li F, Zhao J, Liu CF. *Org Lett*, 2025, 27: 12425–12428
57 Zhao J, Zhang Z, Yang Y, Zhang X, Balamkundu S, Liu CF. *Org Lett*, 2025, 27: 12629–12633
58 Zhao K, Lim YJ, Liu Z, Long H, Sun Y, Hu JJ, Zhao C, Tao Y, Zhang X, Li D, Li YM, Liu C. *Proc Natl Acad Sci USA*, 2020, 117: 20305–20315
59 Zhang S, Liu YQ, Jia C, Lim YJ, Feng G, Xu E, Long H, Kimura Y, Tao Y, Zhao C, Wang C, Liu Z, Hu JJ, Ma MR, Liu Z, Jiang L, Li D, Wang R, Dawson VL, Dawson TM, Li YM, Mao X, Liu C. *Proc Natl Acad Sci USA*, 2021, 118: e2011196118
60 Hu J, Xia W, Zeng S, Lim YJ, Tao Y, Sun Y, Zhao L, Wang H, Le W, Li D, Zhang S, Liu C, Li YM. *Nat Commun*, 2024, 15: 2677
61 Qi YK, Zheng JS, Liu L. *Chem*, 2024, 10: 2390–2407
62 Tang S, Liang LJ, Si YY, Gao S, Wang JX, Liang J, Mei Z, Zheng JS, Liu L. *Angew Chem Int Ed*, 2017, 56: 13333–13337
63 Pan M, Zheng Q, Wang T, Liang L, Mao J, Zuo C, Ding R, Ai H, Xie Y, Si D, Yu Y, Liu L, Zhao M. *Nature*, 2021, 600: 334–338
64 Cui Y, Han D, Bai X, Shi W. *J Peptide Sci*, 2024, 31: e3657
65 Nuijens T, Toplak A, Schmidt M, Ricci A, Cabri W. *Front Chem*, 2019, 7: 829
66 Mazmanian SK, Liu G, Ton-That H, Schneewind O. *Science*, 1999, 285: 760–763
67 Mazmanian SK, Liu G, Jensen ER, Lenoy E, Schneewind O. *Proc Natl Acad Sci USA*, 2000, 97: 5510–5515
68 Nguyen GKT, Wang S, Qiu Y, Hemu X, Lian Y, Tam JP. *Nat Chem Biol*, 2014, 10: 732–738
69 Zhang D, Wang Z, Hu S, Balamkundu S, To J, Zhang X, Lescar J, Tam JP, Liu CF. *J Am Chem Soc*, 2021, 143: 8704–8712
70 Xia Y, Li F, Zhang X, Balamkundu S, Tang F, Hu S, Lescar J, Tam JP, Liu CF. *J Am Chem Soc*, 2023, 145: 6838–6844
71 Li H, Xia Y, Balamkundu S, Liu J, Li F, Zhang X, Lescar J, Tam JP, Preiser PR, Liu CF. *J Am Chem Soc*, 2025, 148: 1989–1997
72 Rehm FBH, Harmand TJ, Yap K, Durek T, Craik DJ, Ploegh HL. *J Am Chem Soc*, 2019, 141: 17388–17393
73 Rehm FBH, Tyler TJ, de Veer SJ, Craik DJ, Durek T. *Angew Chem Int Ed*, 2022, 61: e202116672
74 Rehm FBH, Tyler TJ, Zhou Y, Huang YH, Wang CK, Lawrence N, Craik DJ, Durek T. *Nat Chem*, 2024, 16: 1481–1489
75 Birch-Price Z, Hardy FJ, Lister TM, Kohn AR, Green AP. *Chem Rev*, 2024, 124: 8740–8786
76 Brouwer B, Della-Felice F, Illies JH, Iglesias-Moncayo E, Roelfes G, Drienovská I. *Chem Rev*, 2024, 124: 10877–10923
77 Huang Y, Zhang P, Wang H, Chen Y, Liu T, Luo X. *Chem Rev*, 2025, 125: 523–598
78 Fu Y, Liu X, Xia Y, Guo X, Guo J, Zhang J, Zhao W, Wu Y, Wang J, Zhong F. *Chem*, 2023, 9: 1897–1909
79 Hu L, Cao W, Jiang Y, Cai W, Lou X, Liu T. *Nat Chem*, 2024, 16: 1960–1971

80 Zhu Y, Ding W, Chen Y, Shan Y, Liu C, Fan X, Lin S, Chen PR. *Nat Chem*, 2024, 16: 533–542
81 Duan HZ, Hu C, Li YL, Wang SH, Xia Y, Liu X, Wang J, Chen YX. *J Am Chem Soc*, 2022, 144: 22831–22837
82 Pacesa M, Pelea O, Jinek M. *Cell*, 2024, 187: 1076–1100
83 Villiger L, Joung J, Koblan L, Weissman J, Abudayyeh OO, Gootenberg JS. *Nat Rev Mol Cell Biol*, 2024, 25: 464–487
84 Gunnoo SB, Madder A. *ChemBioChem*, 2016, 17: 529–553
85 Hoyt EA, Cal PMSD, Oliveira BL, Bernardes GJL. *Nat Rev Chem*, 2019, 3: 147–171
86 Kim Y, Yi HB, Seo K, Lee HS, Shin I. *Trends Chem*, 2025, 7: 240–254
87 Feng GH, Wang TY, Li YM. *Chem Commun*, 2025, 61: 11095–11117
88 Wu M, Bai G, Zhang Z, Xiao H, Sun W, Tian C. *hLife*, 2024, 2: 576–591
89 Fu XP, Yuan Y, Jha A, Levin N, Giltrap AM, Ren J, Mamalis D, Mohammed S, Davis BG. *ACS Cent Sci*, 2023, 9: 405–416
90 Xia Y, Yi Y, Shi Y, Long T, Wang H. *Org Lett*, 2023, 25: 6035–6039
91 Xu HD, Jin P, Chu X, Wei CC, Zhong H, Shen MH, Chen G. *J Am Chem Soc*, 2025, 147: 42964–42971
92 Zheng JS, He Y, Zuo C, Cai XY, Tang S, Wang ZA, Zhang LH, Tian CL, Liu L. *J Am Chem Soc*, 2016, 138: 3553–3561
93 Liu Y, Han D, Liu L. *Angew Chem Int Ed*, 2025, 64: e202504405
94 Paravizzini SJ, Haugaard-Kedström LM, Hutton CA, Karas JA. *Angew Chem Int Ed*, 2025, 64: e202509939
95 Morinaka BI, Lakis E, Verest M, Helf MJ, Scalvenzi T, Vagstad AL, Sims J, Sunagawa S, Gugger M, Piel J. *Science*, 2018, 359: 779–782
96 Vagstad AL, Lakis E, Csizi KS, Walls W, Richter D, Soo Lee K, Stocker R, Gugger M, Broderick WE, Broderick JB, Reiher M, Piel J. *Angew Chem Int Ed*, 2025, 64: e202418054
97 Mahanta N, Liu A, Dong S, Nair SK, Mitchell DA. *Proc Natl Acad Sci USA*, 2018, 115: 3030–3035
98 Han Y, Shi J, Li S, Dan T, Yang W, Yang M. *Chem Sci*, 2022, 13: 14382–14386
99 Schissel CK, Roberts-Mataric H, Garcia IJ, Kang H, Mowzoon-Mogharrabi R, Francis MB, Schepartz A. *J Am Chem Soc*, 2025, 147: 6503–6513
100 Xue Y, Xiong Y, Huang W, Liu J, Liu W. *Nat Prod Rep*, 2025, 42: 1276–1302
101 Newberry RW, VanVeller B, Guzei IA, Raines RT. *J Am Chem Soc*, 2013, 135: 7843–7846
102 Maini R, Dedkova LM, Paul R, Madathil MM, Chowdhury SR, Chen S, Hecht SM. *J Am Chem Soc*, 2015, 137: 11206–11209
103 Maini R, Kimura H, Takatsuji R, Katoh T, Goto Y, Suga H. *J Am Chem Soc*, 2019, 141: 20004–20008
104 Chen X, Mietlicki-Baase EG, Barrett TM, McGrath LE, Koch-Laskowski K, Ferrie JJ, Hayes MR, Petersson EJ. *J Am Chem Soc*, 2017, 139: 16688–16695
105 Walters CR, Szantai-Kis DM, Zhang Y, Reinert ZE, Horne WS, Chenoweth DM, Petersson EJ. *Chem Sci*, 2017, 8: 2868–2877
106 Yang J, Wang C, Xu S, Zhao J. *Angew Chem Int Ed*, 2019, 58: 1382–1386
107 Khoo KK, Galleano I, Gasparri F, Wieneke R, Harms H, Poulsen MH, Chua HC, Wulf M, Tampé R, Pless SA. *Nat Commun*, 2020, 11: 2284
108 Kofoed C, Erkalo G, Tay NES, Ye X, Lin Y, Muir TW. *Nature*, 2025, 645: 793–800
109 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. *Science*, 2012, 337: 816–821
110 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. *Science*, 2013, 339: 819–823
111 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, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. *Nature*, 2021, 596: 583–589
112 Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, Wicky BIM, Courbet A, de Haas RJ, Bethel N, Leung PJY, Huddy TF, Pellock S, Tischer D, Chan F, Koepnick B, Nguyen H, Kang A, Sankaran B, Bera AK, King NP, Baker D. *Science*, 2022, 378: 49–56
113 Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, Ahern W, Borst AJ, Ragotte RJ, Milles LF, Wicky BIM, Hanikel N, Pellock SJ, Courbet A, Sheffler W, Wang J, Venkatesh P, Sappington I, Torres SV, Lauko A, De Bortoli V, Mathieu E, Ovchinnikov S, Barzilay R, Jaakkola TS, DiMaio F, Baek M, Baker D. *Nature*, 2023, 620: 1089–1100
114 Sun K, Li S, Zheng B, Zhu Y, Wang T, Liang M, Yao Y, Zhang K, Zhang J, Li H, Han D, Zheng J, Coventry B, Cao L, Baker D, Liu L, Lu P. *Cell Res*, 2024, 34: 846–858

115 Yin M, Zhou H, Zhu Y, Lin M, Wu Y, Wu J, Xu H, Hsieh CY, Hou T, Chen J, Wu J. *Health Data Sci*, 2024, 4: 0211
116 Ge W, De Silva R, Fan Y, Sisson SA, Stenzel MH. *Adv Mater*, 2025, 37: 2413695

Recent advances in protein synthesis: from chemical synthesis to backbone editing

Yang Zhou, Lizhen Wei, Yong-Xiang Chen*

Key Laboratory of Bioorganic Phosphorus Chemistry & Chemical Biology of Ministry of Education, Department of Chemistry, Tsinghua University, Beijing 100084, China
**Corresponding author (email: chen-yx@mail.tsinghua.edu.cn)*

Abstract: Proteins are the primary executors of life processes. Precise analysis and manipulation of their structure and function are fundamental to life science research. Although traditional recombinant expression systems are widely used, they face challenges in preparing proteins with uniform post-translational modifications, specific labels, or non-canonical structural units. Since the advent of liquid- and solid-phase peptide synthesis, coupled with ligation methods like native chemical ligation and enzymatic ligation, has enabled the chemical total synthesis of medium-to-large proteins. Meanwhile, genetic code expansion has broadened the chemical space of proteins by incorporating noncanonical amino acids. In recent years, with the development of peptide backbone reconstitution and split-intein systems, researchers can now perform *in situ* editing of protein backbones both *in vitro* and in cellular environments. Following the theme "from chemical synthesis to backbone editing," this review summarizes recent methodological progress in solid-phase peptide synthesis, ligation reactions, genetic code expansion, and protein editing. It aims to provide a reference for the precise construction and functional study of proteins.

Keywords: protein chemical synthesis, solid-phase peptide synthesis, peptide ligation, protein editing, backbone reconstitution, segment replacement

doi: 10.1360/SSC-2025-0308