

# 人工智能助力蛋白质翻译后修饰的质谱解析及其临床应用

谢欣然<sup>1</sup>, 杨晓霞<sup>2</sup>, 申华莉<sup>2</sup>, 林 灵<sup>3</sup>

(1. 复旦大学上海医学院, 上海 200032; 2. 复旦大学上海医学院生物医学研究院, 上海 200032;

3. 复旦大学附属中山医院上海心血管病研究所, 上海 200032)

**摘要:** 蛋白质翻译后修饰 (PTM) 作为调控蛋白质功能、拓展生命复杂性的核心分子开关, 在细胞信号转导、代谢调控及疾病发生发展中扮演着重要角色。然而, PTM 固有的低丰度、高动态范围以及化学结构的异质性 (尤其是高度复杂的糖基化修饰), 严重制约了对其进行全面、深度的解析。尽管基于质谱 (MS) 的蛋白质组学已成为该领域的主流技术, 但传统分析流程在富集特异性、修饰位点精准定位以及对数据库依赖性等方面仍面临显著的技术瓶颈。本综述系统梳理了磷酸化、糖基化及赖氨酸酰化等关键 PTM 的生物学功能、经典分析策略及其局限性。在此基础上, 重点探讨了人工智能 (AI), 尤其是深度学习算法在突破上述技术壁垒中的应用潜力; 详细论述了 AI 驱动的计算范式如何通过构建高精度的谱图库预测模型、优化位点定位算法 (如降低假阳性率) 以及开发不依赖数据库的复杂糖肽鉴定策略, 从而显著提升 PTM 分析的覆盖度与准确性。本文进一步阐述了这些新兴技术在临床转化中的实际价值, 重点列举了其在膀胱癌尿液外泌体诊断标志物筛选、KRAS 突变肺癌磷酸化信号网络解析及胰腺癌耐药机制研究中的突破性进展。展望未来, 随着 AI 算法与单细胞蛋白质组学及多组学数据的深度融合, 以及新材料、新富集策略的不断发展, 不仅将有助于系统揭示 PTM 的时空动态调控规律, 更将为精准医学背景下的疾病早期诊断与靶向治疗提供强有力的技术支撑。

**关键词:** 蛋白质翻译后修饰 (PTM); 质谱 (MS); 人工智能 (AI); 谱图解析; 临床应用

中图分类号: O657.63 文献标志码: A 文章编号: 1004-2997(2026)03-0361-19

DOI: 10.7538/zpxb.2025.0119 CSTR: 32365.14.zpxb.2025.0119

## AI-Driven Mass Spectrometry-Based Post-Translational Modification Analysis and Clinical Application

XIE Xin-ran<sup>1</sup>, YANG Xiao-xia<sup>2</sup>, SHEN Hua-li<sup>2</sup>, LIN Ling<sup>3</sup>

(1. Shanghai Medical College of Fudan University, Shanghai 200032, China; 2. Institutes of Biomedical Sciences of Shanghai Medical School, Fudan University, Shanghai 200032, China; 3. Shanghai Institute of Cardiovascular Diseases, Zhongshan Hospital, Fudan University, Shanghai 200032, China)

**Abstract:** Post-translational modifications (PTMs) function as fundamental molecular switches that govern protein activity and exponentially expand the functional complexity of the proteome far beyond the genomic blueprint. These modifications play pivotal and multifaceted roles in regulating essential biological processes, including cellular signal transduction, metabolic homeostasis, and the initiation and progression of various pathogenic conditions. However, the comprehensive and

in-depth characterization of the PTM landscape remains a formidable analytical challenge. This difficulty stems primarily from the inherent physicochemical characteristics of PTMs, such as their substoichiometric abundance, the wide dynamic range in biological samples, and significant structural heterogeneity—a complexity that is particularly pronounced in the analysis of intricate modifications like glycosylation. While mass spectrometry (MS)-based proteomics has emerged as the dominant technology in this field, traditional analytical workflows are frequently hampered by significant technical bottlenecks. These limitations are most evident in the lack of enrichment specificity, ambiguous site localization, and a heavy dependency on incomplete reference databases. This review systematically summarized the critical biological functions of key PTMs, including phosphorylation, glycosylation, and lysine acylation, while also providing a critical evaluation of classical analytical strategies and their inherent limitations. Building on this foundation, the transformative potential of artificial intelligence (AI)—specifically advanced deep learning algorithms—as a revolutionary force capable of breaking through these longstanding technical barriers was focused on in our discussion. The ways in which AI-driven computational paradigms are reshaping the field were elaborated in detail by enabling the construction of high-precision virtual spectral libraries, optimizing site localization algorithms to rigorously control false localization rates, and facilitating the database-independent identification of complex glycopeptides. Collectively, these advancements result in a significant enhancement in the depth, coverage, and accuracy of PTM analysis. Furthermore, this article elucidated the practical utility of these emerging technologies in clinical translation. Breakthrough progress in specific applications was highlighted, such as the screening of diagnostic glycosignatures in urinary extracellular vesicles for bladder cancer, the decoding of dysregulated phosphorylation signaling networks in KRAS-mutant lung cancer, and the elucidation of molecular mechanisms underlying drug resistance in pancreatic cancer. Looking forward, the deep integration of AI algorithms with single-cell proteomics and multi-omics data, coupled with the continuous development of novel enrichment materials, will be instrumental in decoding PTM regulatory networks. This convergence promises to systematically reveal the spatiotemporal dynamics of PTMs, thereby providing robust technical support for early diagnosis and the development of targeted therapies in the era of precision medicine.

**Key words:** post-translational modification (PTM); mass spectrometry (MS); artificial intelligence (AI); spectrum interpretation; clinical application

## 1 研究背景

### 1.1 蛋白质翻译后修饰是解码生命复杂性的分子开关

蛋白质是生命活动中的分子机器, 在生命过程中发挥关键作用。然而, 仅依靠人类基因组中约2万个编码基因所表达的蛋白质, 远不足以支撑人体复杂的生理功能和动态调控网络。翻译后修饰 (post-translational modifications, PTMs) 是在蛋白质翻译完成后, 通过酶促或非酶促反应对特定氨基酸残基进行共价修饰的过程, 也是拓展遗传密码功能的最高效方式之一。PTM通过重塑蛋白质的空间构象、电荷分布、亲疏水性及相

互作用界面等理化性质, 将有限的基因蓝图扩展为功能浩瀚的蛋白质功能网络, 进而实现对生命活动的动态与精准调控。

目前已知的PTM种类已超过700种, 包括磷酸化、乙酰化、甲基化、泛素化、SUMO化、糖基化、棕榈酰化、肉豆蔻酰化、异戊二烯化、胆固醇化、谷胱甘肽化、瓜氨酸化等。这些修饰通过精确调控蛋白质的活性、稳定性、相互作用及亚细胞定位, 构成了细胞信号网络的基础, 赋予了蛋白质在合成后应对环境变化的功能可塑性<sup>[1]</sup>。PTM调控机制主要体现在以下几方面: 1) 调控酶活性与信号转导: 通过动态调控蛋白质催化活

性进而激活或抑制信号通路,如激酶级联反应中的磷酸化开关<sup>[2]</sup>; 2)决定蛋白质的降解命运:如泛素化介导的蛋白酶体降解途径直接调控蛋白质的稳定性与半衰期<sup>[3]</sup>; 3)构筑蛋白动态互作网络:通过改变蛋白质的相互作用界面,介导蛋白复合物的组装与解离,如磷酸化和SUMO化对蛋白复合物组装的影响<sup>[4-5]</sup>; 4)指导蛋白质的亚细胞内定位与运输:如DHHC家族酶(蛋白质棕榈酰转移酶)通过催化外周膜蛋白(如Gao、MGNC)的S-棕榈酸化增强疏水性,从而稳定其在质膜上的锚定<sup>[6]</sup>; 5)参与表观遗传调控:如组蛋白H3K4me3甲基化与H3K27ac乙酰化,通过调控染色质结构、维持增强子功能,直接参与基因表达的表观调控<sup>[7]</sup>。

## 1.2 PTM的功能重要性及临床应用价值

相对于基因组与转录组,蛋白质组的复杂性因PTMs的广泛存在而呈指数级放大。据估计,PTMs使得人体蛋白质种类超过百万种,其丰度差异可跨越7个数量级以上,这为蛋白质组学研究带来了独特的挑战与机遇。在多种重大疾病中,对PTMs的深入研究已成为阐明其发病机制、发掘新型生物标志物和治疗策略的核心<sup>[8-9]</sup>。

在神经退行性疾病中,蛋白质的异常PTM是其错误折叠与致病的关键始动因素。以阿尔茨海默病(Alzheimer's disease, AD)为例, Tau蛋白发生异常的磷酸化、乙酰化、泛素化等多种PTMs,其构象发生改变并错误聚集,形成神经纤维缠结(neurofibrillary tangles, NFTs),继而破坏神经元微管网络稳定性,引发线粒体功能障碍和突触损伤,最终驱动疾病的进展。值得注意的是,特定磷酸化位点(如p-Tau202、p-Tau231等)已被证实与AD严重程度高度相关,可作为临床诊断和疾病分期的关键生物标志物<sup>[10-11]</sup>。

在代谢性疾病中,PTM调控网络同样贯穿始终。PTMs在糖异生、三羧酸(TCA)循环和糖酵解等主要代谢通路间建立了精密的联系,直接影响胰岛β细胞的活力与功能。胰岛素受体(insulin receptor, IR)激活会启动1个由PTM精密调控的信号级联反应,其中蛋白激酶B(PKB)、丝裂原活化蛋白激酶(MAPK3/1)、胞外信号调节激酶(ERK1/2)以及AMP活化蛋白激酶(AMPK)等关键分子被磷酸化修饰,从而精准调控机体能量稳态。相反,在已缺失2个*Ins1*等位基因的小

鼠中,再删除2个*Ins2*等位基因后,成年小鼠β细胞胰岛素生成量急剧下降,同时伴随不可避免的高血糖以及β细胞显著增殖。这与PKB的过度磷酸化有关<sup>[12]</sup>,提示PTMs失衡与代谢性疾病的发生发展密切相关。

在肿瘤免疫领域,PTM的动态调控决定了肿瘤细胞能否成功逃逸免疫监视。例如,程序性死亡配体1(PD-L1)的稳定性受到不同PTM的对话调控(crosstalk)。一方面,糖原合成酶激酶3β(GSK3β)可磷酸化PD-L1的T180和S184位点;另一方面,发生在N192、N200和N219位点的N-糖基化修饰则会竞争性干扰这一磷酸化过程,从而稳定PD-L1的表达,显著促进肿瘤细胞的免疫逃逸,示于图1。这一机制的阐明为基于PD-L1肿瘤免疫治疗提供了全新的干预靶点与策略<sup>[13]</sup>。

综上所述,PTM作为调控蛋白质功能的关键机制,在疾病发生发展及临床转化方面均具有不可替代的科学价值与临床意义。

## 2 PTM常见类型及其机制与功能

根据催化方式的不同,PTMs主要分为酶促修饰(ePTMs)与非酶促修饰(nPTMs)两大类,二者在特异性、可逆性和生物学意义上存在本质区别。ePTMs由特异性生物酶催化,由于人体中修饰酶和去修饰酶常成对存在,因此ePTMs大多具有高度可逆性和底物特异性,从而能够对细胞的生命活动进行精准调控。nPTMs则是在没有酶参与情况下发生的自发化学反应,其过程易受细胞内代谢产物浓度及环境因素(如氧化应激)的影响,这类修饰通常不可逆且随机性较强,因而其异常累积常与代谢紊乱及多种疾病的病理过程密切相关。

### 2.1 磷酸化

磷酸化(phosphorylation)是最普遍、研究最深入的PTM之一,是细胞信号转导网络的核心机制。绝大多数磷酸化属于酶促可逆反应。在蛋白激酶(protein kinase)的催化下,三磷酸腺苷(ATP)中的γ-磷酸基团被共价转移至靶蛋白的丝氨酸(Ser)、苏氨酸(Thr)或酪氨酸(Tyr)残基的羟基上<sup>[14]</sup>,这一过程可被蛋白磷酸酶(protein phosphatase)逆转,从而实现了对蛋白质功能的动态“开/关”调控<sup>[15]</sup>。此外,在氧化应激条件下,活性氧(ROS)也可诱导发生非酶促的磷酸化修饰<sup>[16]</sup>。

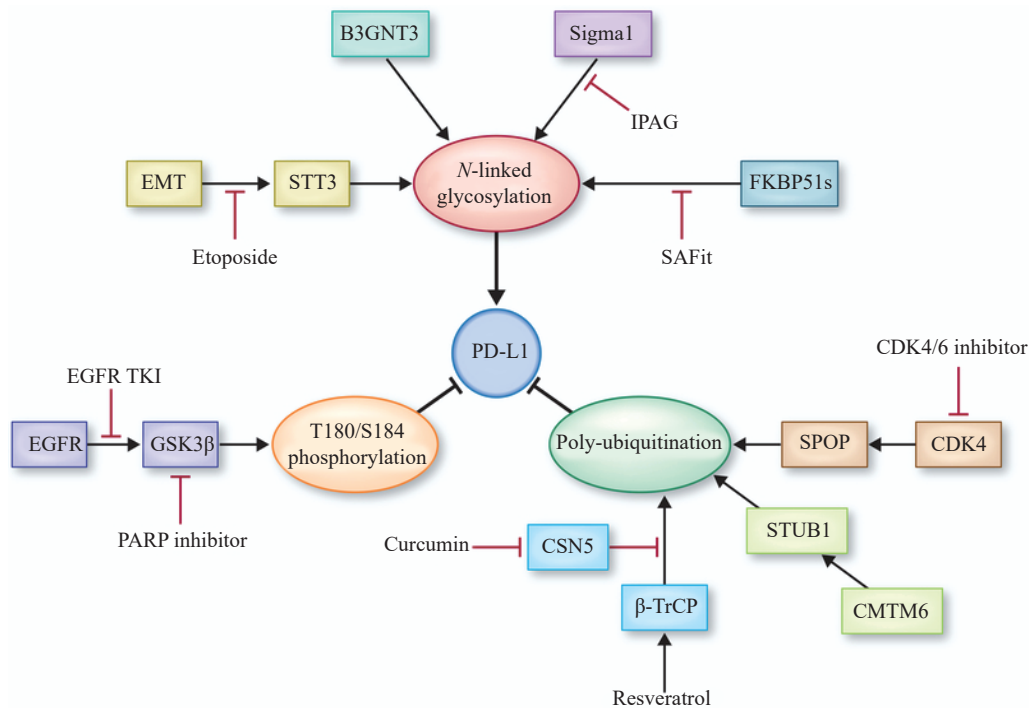


图1 PTM串扰对免疫检查点PD-L1的精细调控<sup>[13]</sup>

Fig. 1 Crosstalk of PTM in the fine regulation of immune checkpoint PD-L1<sup>[13]</sup>

蛋白激酶是最大的蛋白超家族,磷酸化修饰通过引入负电荷基因改变蛋白质构象及相互作用界面,进而改变底物活性,参与细胞内几乎所有重要的生命活动<sup>[17]</sup>。磷酸化是细胞信号传递的主要方式:表皮生长因子受体(EGFR)自身磷酸化修饰(尤其是酪氨酸残基的磷酸化)是激活下游信号通路的核心机制,而酪氨酸 Y1068 位是 EGFR 中 1 个高度保守且功能重要的磷酸化位点。磷酸化的 Y1068 通过招募含有 SH2 或 PTB 结构域的下流信号分子,激活 RAS-RAF-MEK-ERK 通路,调控细胞增殖、分化及存活<sup>[18-19]</sup>。磷酸化调控细胞周期进程:定量磷酸化蛋白质组学揭示有丝分裂期间磷酸化位点的广泛富集,即超过一半已鉴定的受调控磷酸化位点在有丝分裂中达到峰值,提示细胞周期中涉及的调控过程远多于单一信号通路<sup>[20]</sup>,示于图 2。细胞周期蛋白依赖性激酶(CDKs)通过磷酸化 Rb 蛋白,促使其与转录因子 E2F 解离,从而驱动细胞周期从 G1 期向 S 期转换,实现对细胞生长的精确调控<sup>[21]</sup>。1)磷酸化参与表观遗传调控:组蛋白的磷酸化直接影响染色质结构(疏松或紧密),从而调控基因转录。例如,组蛋白 H3 第 10 位丝氨酸(H3S10)的磷酸化由多种激酶催化,其磷酸化会增加组蛋白

尾部的负电荷,削弱组蛋白与 DNA 及核层蛋白相关蛋白之间的相互作用,导致染色质结构疏松,使转录因子更易结合 DNA 启动子区域,从而促进基因转录<sup>[22]</sup>。2)磷酸化参与代谢调控:细胞代谢(如糖酵解和 TCA 循环)的关键步骤由代谢酶催化,而这些酶的活性可通过磷酸化快速调节,从而适应细胞能量需求或环境变化。以糖酵解途径中的关键限速酶丙酮酸激酶(PKM2)为例,其活性受到精密的磷酸化调控:成纤维细胞生长因子受体 1(FGFR1)能够直接磷酸化 PKM2 的 105 位酪氨酸残基(Y105),这一磷酸化事件通过阻碍变构激活剂果糖-1,6-二磷酸(FBP)的结合,抑制 PKM2 形成其高活性的四聚体形态。这种由酪氨酸磷酸化驱动的 PKM2 活性抑制,使得糖酵解中间代谢产物更多地流向生物合成途径,最终赋予了肿瘤细胞独特的代谢优势,从而加速其生长与增殖<sup>[23]</sup>。

鉴于磷酸化在细胞信号网络中的核心地位,其调控失常是促使肿瘤恶性表型形成的关键驱动因素。一个经典的例子是慢性粒细胞白血病(CML),其标志性的 BCR-ABL 融合蛋白由于染色体易位所致,导致 ABL 酪氨酸激酶处于持续、失控的组成性激活状态。这种异常的磷酸化级

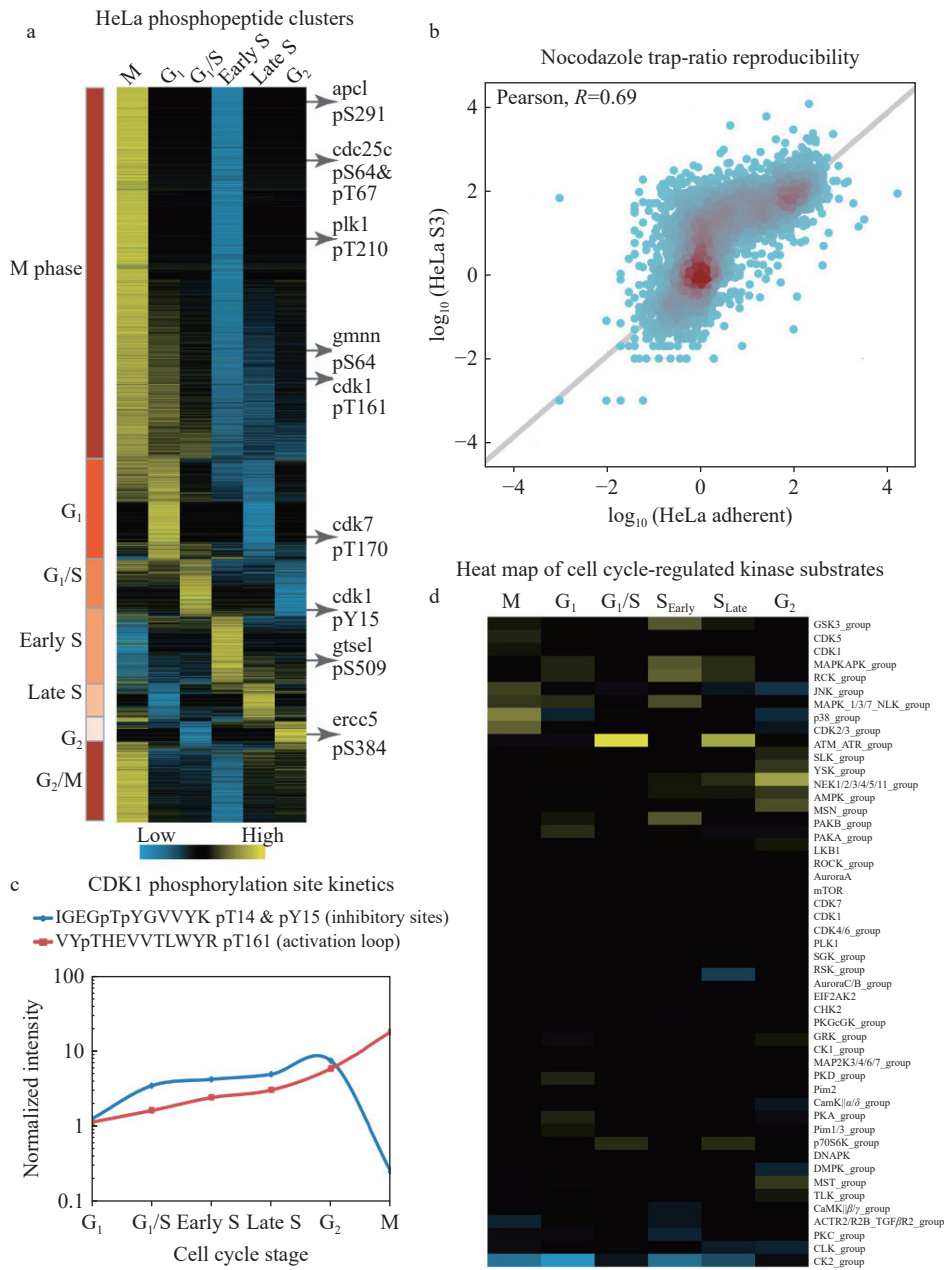


图 2 细胞周期中磷酸化蛋白质组的动态变化<sup>[20]</sup>

Fig. 2 Dynamic changes of the phosphoproteome during the cell cycle<sup>[20]</sup>

联反应持续激活下游多条增殖与生存通路(如 STAT、PI3K/AKT、RAS/MAPK 通路), 导致骨髓中粒细胞系的克隆性异常扩增, 最终驱动了 CML 的发生<sup>[24-25]</sup>。与这种因染色体易位导致的融合蛋白不同, 在实体瘤中, 激酶基因自身的突变是更为常见的致癌机制, 非小细胞肺癌中的 EGFR 突变便是其中的典范。EGFR 激酶结构域的功能获得性突变 (gain-of-function mutations), 如外显子 19 缺失或 L858R 点突变, 使其能够自发地、持续地发生自身磷酸化, 无需配体激活。

这种失控的磷酸化信号同样通过激活 PI3K-AKT 和 RAS/MAPK 等通路, 赋予肿瘤细胞无限增殖的能力。值得一提的是, 正是基于对 BCR-ABL 和突变型 EGFR 这些异常磷酸化事件的深入解析, 直接催生了伊马替尼<sup>[26]</sup>和 EGFR-TKIs<sup>[27]</sup>等革命性的靶向药物。

### 2.2 糖基化

糖基化 (glycosylation) 是指在酶的催化下, 将预先合成的糖链共价连接至蛋白质等大分子底物上的过程, 是目前最复杂、最多样的 PTM 之

一。糖链在单糖组成、支链结构、连接异构体和化学键类型等方面存在显著差异,这种多维度的结构多样性使得糖基化修饰呈现出巨大的组合可能性和高度复杂性,从而极大地拓展了蛋白质的潜在功能空间。根据糖链与底物的连接方式,糖基化可分为以下几种主要类型:1) *N*-连接糖基化(*N*-linked glycosylation): 在内质网中,寡糖链通过糖苷键共价连接到新生肽链的天冬酰胺(Asn)残基上,该修饰对蛋白质的正确折叠、稳定性及膜定位至关重要。2) *O*-连接糖基化(*O*-linked glycosylation): 主要发生在高尔基体(Golgi)中,单糖被逐一添加到 Ser 或 Thr 残基的羟基上。3) *O*-GlcNAc 糖基化(*O*-GlcNAcylation): 这是一种特殊的 *O*-连接糖基化类型,指单个 *N*-乙酰葡萄糖胺(GlcNAc)分子修饰核质蛋白的 Ser/Thr 残基,在信号转导和转录调控中扮演类似磷酸化的分子开关角色。4) *C*-连接糖基化(*C*-linked glycosylation): 是一种较罕见的修饰类型,通过碳-碳键将甘露糖连接到色氨酸(Trp)残基上<sup>[28]</sup>。

糖基化的生物合成是一个多步骤过程。首先,单糖需被核苷酸(如尿苷二磷酸-*N*-乙酰葡萄糖胺,UDP-GlcNAc)激活,形成高能的糖基供体;随后,这些活化的单糖在内质网或高尔基体等细胞器中,经由不同的合成与转移途径逐步组装成复杂的糖链。正是这些合成途径的多样性,决定了糖基化修饰的高度复杂性。

糖基化通过直接改变蛋白质构象与相互作用界面,在细胞内外发挥着不可或缺的生物功能,示于图3<sup>[29]</sup>。糖基化修饰参与蛋白质的质量控制与定位;在内质网中,*N*-糖基化是蛋白质正确折叠的“质量控制”系统;同时,它也是一种“定位信号”,例如高尔基体中 *N*-连接聚糖上的甘露糖-6-磷酸(M6P)标记,可引导其修饰的蛋白质靶向递送到溶酶体<sup>[30]</sup>。此外,糖基化修饰还参与细胞间的识别与通讯:细胞表面的糖链构成了细胞的“身份标识”,在炎症反应中,E-选择素(E-selectin)通过其 N 端凝集素结构域识别白细胞表面糖蛋白(如 PSGL-1、CD44)上的唾液酸化

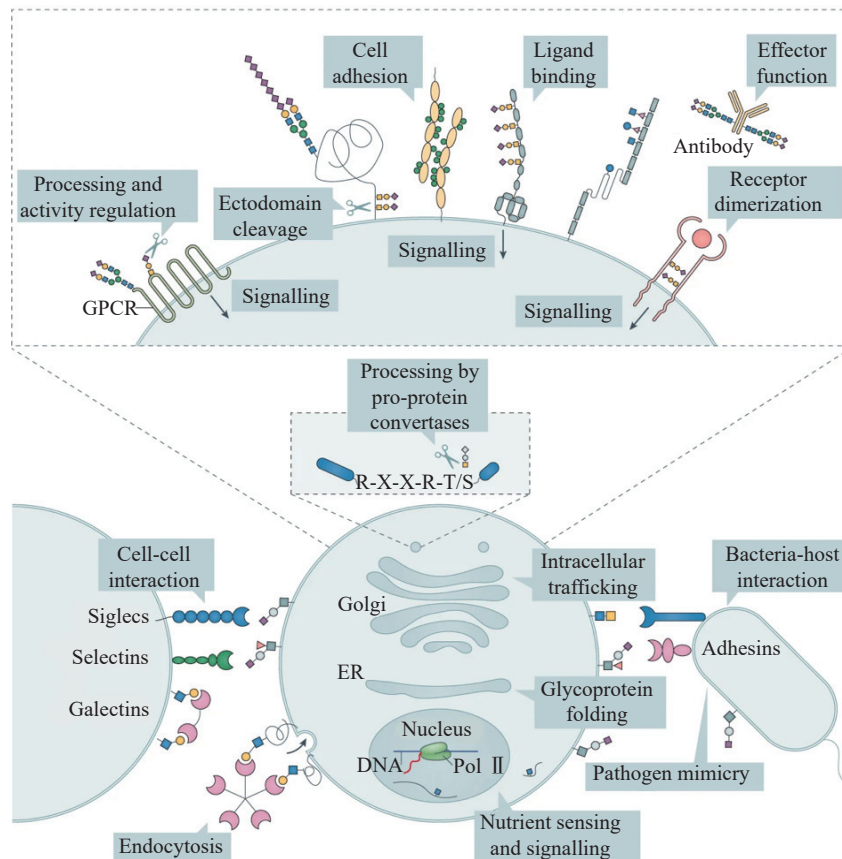


图3 蛋白质糖基化的多效性与核心生物学功能<sup>[29]</sup>

Fig. 3 Pleiotropy and core biological functions of protein glycosylation<sup>[29]</sup>

路易斯 X(sLeX)糖基化修饰抗原,进而介导白细胞在血管壁的稳定粘附与滚动<sup>[31]</sup>。

糖基化途径的遗传缺陷会导致一类被称为先天糖基化缺陷病(congenital disorders of glycosylation, CDG)的严重人类遗传病<sup>[31]</sup>,例如与早衰样综合征相关的糖胺聚糖核心半乳糖转移酶缺乏症<sup>[32]</sup>。在获得性疾病中,糖基化异常同样是推动病理进程的关键因素。例如,在射血分数保留型心力衰竭(HFpEF)中,己糖激酶1(HK1)的异常定位会导致蛋白N-糖基化水平下降,使得糖胺代谢中间产物更多地流向O-GlcNAc修饰途径,进而引发血管新生障碍,并最终驱动HFpEF的发展<sup>[33]</sup>。

RNA糖基化(glycoRNA)的发现是近年来生命科学领域的重大突破之一,打破了蛋白质为糖基化唯一大分子底物的传统观念。这一现象由诺贝尔奖得主Carolyn Bertozzi团队于2021年首次报道,该研究发现特定小RNA(如Y RNA、miRNA)通过N-糖基化修饰核苷酸acp3U位点,形成glycoRNA<sup>[34-35]</sup>。功能研究表明,这种新颖的修饰不仅能够调控RNA的稳定性、介导其在细胞表面的定位,还可影响其与免疫受体的相互作用。GlycoRNA的发现极大地拓展了糖生物学与RNA生物学的功能版图,其在物种间的保守性也暗示了其具有重要的基础性生物学意义。

### 2.3 赖氨酸酰化

赖氨酸酰化(lysine acylation)是调控细胞功能的核心PTM之一,泛指将来源于细胞代谢中间产物的不同酰基(acyl group)连接至蛋白质赖氨酸残基 $\epsilon$ -氨基的过程。这类修饰的核心分子机制是通过中和赖氨酸侧链的正电荷,改变蛋白质的构象、稳定性及相互作用界面,进而在表观遗传调控、代谢调控和免疫应答等多个层面发挥关键作用。在众多赖氨酸酰化修饰中,乙酰化是研究最为系统且深入的经典类型;而近年来,乳酸化等一系列与细胞代谢状态直接关联的新型酰化修饰也不断被发现,它们共同揭示了细胞通过感知代谢物水平精准调控基因表达和细胞功能的内在机制。

**2.3.1 乙酰化(acetylation)** 乙酰化是调控细胞功能的核心PTM之一,在表观遗传调控、代谢调控和免疫应答等多个层面发挥关键作用。乙酰化是指在乙酰转移酶的催化下,将乙酰辅酶

A(acetyl-CoA)的乙酰基( $-\text{CO}-\text{CH}_3$ )共价连接到蛋白质底物上的过程<sup>[36-37]</sup>。赖氨酸乙酰化(lysine acetylation, Kac)是研究最为广泛的功能性乙酰化类型,其修饰靶点为赖氨酸残基侧链的 $\epsilon$ -氨基,是一个高度动态且可逆的过程。该修饰的可逆性由赖氨酸乙酰转移酶(KATs)和赖氨酸脱乙酰酶(KDACs)这对“书写器和擦除器”(writer & eraser)系统共同精密调控<sup>[21,38]</sup>。由于组蛋白是乙酰化最重要的底物之一,KATs和KDACs被分别称为组蛋白乙酰转移酶(HATs)和去乙酰酶(HDACs)。值得注意的是,赖氨酸乙酰化的供体是核心代谢物乙酰辅酶A,这一特征直接将细胞的代谢状态与蛋白质乙酰化调控网络紧密关联。N-末端乙酰化(N-terminal acetylation)则是最普遍的乙酰化形式之一,在超过80%的人类蛋白中均可发生<sup>[39]</sup>。该过程由N-末端乙酰转移酶(NAT)复合物催化,修饰新生肽链N-末端的 $\alpha$ -氨基,且通常是不可逆的<sup>[40]</sup>。此外,O-乙酰化(O-acetylation)是一种相对少见的修饰类型,主要发生在丝氨酸和苏氨酸的羟基侧链上。

乙酰化通过中和赖氨酸的正电荷或诱导构象改变,实现对下游生物学过程的调控。表观遗传调控是乙酰化最经典的功能。例如,组蛋白H4K16位点的乙酰化(H4K16ac)能中和其正电荷,减弱其与DNA及相邻核小体(如H2A)的静电相互作用,从而降低染色质的压缩程度,促进染色质结构开放,最终有利于转录因子结合及基因表达激活,示于图4<sup>[41-44]</sup>。此外,BET家族蛋白(如BRD4)可作为乙酰化信号的阅读器(reader),其溴结构域能够特异性识别组蛋白乙酰化位点(如H3K9ac、H4K16ac),并将RNA聚合酶II(RNAPII)和转录延伸因子(如P-TEFb)招募至染色质,进而高效启动基因表达<sup>[45-46]</sup>。乙酰化亦可直接调控代谢酶活性:例如,磷酸戊糖途径的限速酶——葡萄糖-6-磷酸脱氢酶(G6PD),其关键位点(K403/K488)的乙酰化会抑制其酶活性,从而对细胞的氧化还原状态和生物合成产生影响<sup>[47]</sup>。乙酰化在抗病毒天然免疫中扮演着重要角色:干扰素(IFN)通路的多种关键蛋白(包括cGAS、RIG-I、MAVS和IRF3等)的活性和稳定性均受到乙酰化的精密调控,从而影响宿主对病毒感染的响应<sup>[48]</sup>。

乙酰化调控网络的失衡与肿瘤、代谢疾病

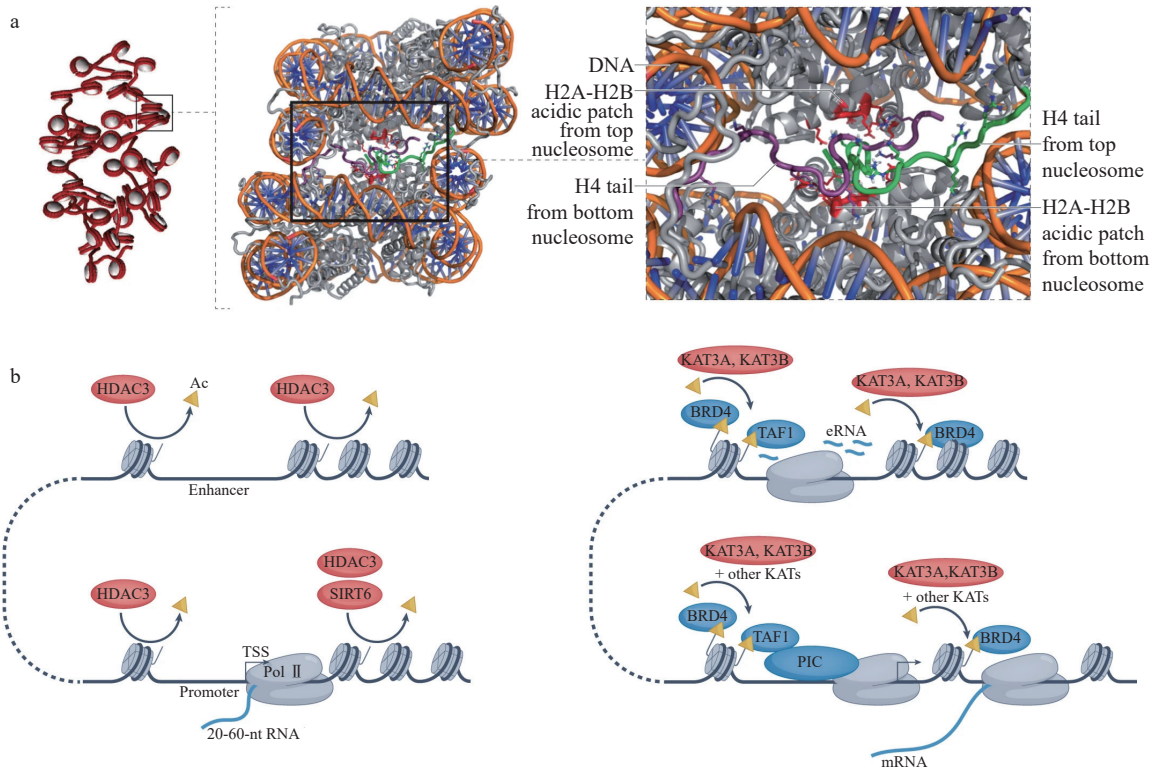


图4 组蛋白乙酰化对染色质的分子效应<sup>[44]</sup>

Fig. 4 Molecular effects of histone acetylation on chromatin<sup>[44]</sup>

及神经退行性疾病的发生发展进程密切相关<sup>[49]</sup>。例如,抑癌基因 p53 的乙酰化缺陷会削弱其功能,进而促进肿瘤的发生<sup>[50]</sup>; G6PD 的异常乙酰化与代谢疾病的病情恶化密切相关<sup>[47]</sup>; HDAC 的过度激活则与阿尔茨海默病的病理进程密切相关<sup>[50]</sup>。

**2.3.2 乳酸化 (lactylation)** 乳酸化是一种新发现的 PTM, 可直接将细胞代谢状态与蛋白质功能关联起来。乳酸化通过将糖酵解产物——乳酸, 以乳酰基的形式共价连接到靶蛋白的赖氨酸残基上, 从而实现对细胞功能的精准调控。乳酸化的发生途径兼具酶促和非酶促 2 种: 1) 酶促途径主要由组蛋白乙酰转移酶 (如 p300) 催化, 利用乳酰辅酶 A (lactyl-CoA) 作为供体, 将乳酰基转移至赖氨酸  $\epsilon$ -氨基上<sup>[51]</sup>; 部分丙氨酰-tRNA 合成酶 (AARS) 也能在 ATP 的协助下, 直接利用乳酸分子对目标蛋白质的赖氨酸残基进行修饰。2) 非酶促途径是在高乳酸环境 (如肿瘤微环境或缺氧组织) 中, 乳酸化可通过非酶促方式自发发生, 例如乳基谷胱甘肽介导乳酸分子直接与赖氨酸  $\epsilon$ -氨基发生酯化反应<sup>[52]</sup>。值得注意的是, 乳酸化与乙酰化可在同一赖氨酸残基位点发生竞争性修饰, 富含组蛋白乳酸化修饰的基因组区域的

乙酰化水平通常较低<sup>[53]</sup>, 这种动态竞争使细胞能够根据代谢状态 (如糖酵解水平) 在不同的基因时空表达程序之间进行切换。

乳酸化最核心的功能是作为一种表观遗传修饰, 可直接将细胞代谢状态与基因转录关联起来。乳酸作为糖酵解的终产物, 其水平波动会直接影响乳酰辅酶 A 的供给, 进而调控组蛋白乳酸化的程度<sup>[54]</sup>。例如, 在免疫应答过程中, 巨噬细胞中的组蛋白乳酸化 (如 H3K18la) 能够响应高糖酵解环境, 促进细胞向免疫抑制性 M2 表型极化, 以此调控巨噬细胞的命运决策<sup>[55]</sup>。

鉴于乳酸化在代谢-表观遗传偶联中的关键作用, 其调控失调与多种重大疾病的病理进程密切相关。在肿瘤发生发展进程中, 乳酸化可通过调控基因表达为肿瘤细胞提供代谢优势。例如, 在眼部黑色素瘤中, 组蛋白乳酸化通过上调 YTHDF2 的转录, 加速抑癌基因 PER1 mRNA 的降解, 进而驱动肿瘤的发生<sup>[56-57]</sup>。类似的促癌机制也在非小细胞肺癌<sup>[57]</sup>和软组织肉瘤<sup>[58]</sup>中得到证实。在心血管疾病中, 乳酸化对维持心肌细胞的生理功能至关重要。在心力衰竭进程中, 心肌肌球蛋白  $\alpha$ -MHC 关键位点 K1897 的乳酸化水平降

低,破坏了其与 Titin 蛋白的相互作用,导致心肌收缩功能障碍<sup>[59]</sup>。在主动脉夹层发病过程中, March2 的表达缺乏会通过 PKM2 二聚体乳酸化促进 H3K18la 水平升高,进而增强 p53 驱动的细胞凋亡,加剧疾病的进程<sup>[60]</sup>。此外,组蛋白乳酸化修饰在心肌梗死后的修复基因激活中也发挥着重要作用<sup>[61]</sup>。

### 3 经典质谱技术在 PTM 解析中的应用与挑战

基于质谱 (mass spectrometry, MS) 的蛋白质组学技术是系统性解析 PTM 的核心手段。在传统的分析流程中,首先将从细胞或组织裂解液中提取的蛋白质,经特异性蛋白酶(如胰蛋白酶)酶切为较短肽段,以满足质谱分析的要求。然而,由于 PTM 修饰肽段的丰度通常远低于未修饰肽段,因此 PTM 信号易被高丰度背景肽段淹没。特异性富集能够较好地解决这一问题,即利用目标修饰的独特性质(如亲和力、电荷或化学标签),将修饰肽段从复杂的肽段混合物中分离出来。富集后的修饰肽段经液相色谱(LC)分离后,进入串联质谱(MS/MS)仪检测,最终通过数据库检索等生物信息学方法,对肽段序列、修饰类型及精确位点进行鉴定与确证。这一“富集-鉴定”的通用流程构成了 PTM 分析传统流程的基础,但针对不同类型的 PTM 修饰开展分析时,仍需应对各具特异性的技术挑战。

#### 3.1 磷酸化的传统质谱分析策略与技术瓶颈

全细胞裂解液经胰蛋白酶(trypsin)酶切后得到的复杂肽段混合物,通常是开展磷酸化修饰质谱分析的起点。获得肽段混合物后,需对其进行特异性富集,常用的富集方法包括固定化金属离子亲和层析(immobilized metal ion affinity chromatography, IMAC)和金属氧化物亲和层析(metal oxide affinity chromatography, MOAC)。

IMAC 的原理是带负电荷的磷酸基团与带正电荷的金属离子会发生高亲和力的相互作用。通过将金属离子(如三价铁离子( $\text{Fe}^{3+}$ )、镓离子( $\text{Ga}^{3+}$ )等)螯合(chelated)于层析填料上,即可制备得到固相亲和介质。当肽段混物流经该介质时,磷酸化修饰的肽段会被特异性捕获,而大部分未修饰肽段则直接流走,从而实现磷酸化肽段的高效富集<sup>[62]</sup>。

MOAC 的原理是将金属氧化物(代表性材料为二氧化钛( $\text{TiO}_2$ )和四氧化三铁( $\text{Fe}_3\text{O}_4$ ))表面作为路易斯酸,进而与作为路易斯碱的磷酸基团形成强相互作用。 $\text{TiO}_2$  材料对水溶液中的磷酸根离子具有极强的亲和力,同时,通过使用含取代有机酸(如二羟基苯甲酸(DHB)、邻苯二甲酸或乙醇酸)的高度酸性上样缓冲液,可高效抑制酸性肽段的非特异性结合,以此实现对磷酸化肽段的高选择性富集<sup>[63-64]</sup>。

不同富集方法对于不同类型磷酸化肽段的选择性存在差异。IMAC 对多磷酸化肽段具有更高的亲和力,但在处理复杂样本时,其结合容量与选择性存在一定局限;相比之下, $\text{TiO}_2$  对单磷酸化肽段表现出更高的富集效率,但因多磷酸化肽段与  $\text{TiO}_2$  结合过强而往往难以被高效洗脱。上述技术限制也推动了将 IMAC 与 MOAC 相结合的富集策略的发展,例如顺序洗脱固定化金属离子亲和色谱(sequential elution from IMAC, SIMAC)法<sup>[65]</sup>。

在完成对磷酸化肽段的高效富集后,样品随即进入质谱分析的核心环节。首先,富集所得的肽段混合物经由反相液相色谱(reverse-phase liquid chromatography, RP-LC)进行在线分离,即利用肽段的疏水性差异,使其进入质谱仪的时间产生先后顺序,以降低进入质谱仪样品的复杂性。随后,分离后的肽段通过电喷雾电离(electrospray ionization, ESI)技术进入串联质谱仪。质谱仪采用数据依赖采集模式(data-dependent acquisition, DDA):首先在一级质谱( $\text{MS}^1$ )中对洗脱的肽段进行扫描,快速测定其质荷比( $m/z$ )和信号强度;然后,智能筛选出信号强度最高的若干个母离子(precursor ions),对其逐一进行碎裂,并检测产生的碎片离子,最终获得对应的二级质谱图<sup>[66]</sup>。

在数据解析层面,采集到的二级质谱图是鉴定肽段序列的关键。传统的生物信息学分析流程是借助自动化搜库软件(如 Mascot、Sequest 等),将实验获得的成千上万张谱图与理论上蛋白质酶切、修饰后产生的肽段谱图进行比对和打分。通过这种方式,不仅能够鉴定出肽段的氨基酸序列,还可初步判断其是否携带磷酸化修饰。

尽管上述分析流程已被广泛应用,但解析磷酸化这类高度动态且复杂的 PTM 时,仍面临诸

多技术瓶颈。其中,磷酸化位点定位的模糊性是核心挑战之一。在常用的碰撞诱导解离(CID)或高能碰撞解离(HCD)碎裂模式下,连接在丝氨酸和苏氨酸上的磷酸基团化学性质极不稳定,易以中性丢失的形式从肽段上脱落。这会导致二级质谱图中磷酸化相关特征信息严重缺失,同时造成可用于确认氨基酸序列的碎片离子信号极弱。因此,当单个肽段内存在多个潜在磷酸化位点时,仅依靠传统搜库算法往往难以精准判断修饰发生的具体位置,进而导致假定位率(false localization rate, FLR)居高不下。电子转移解离(ETD)等基于电子的互补碎裂技术能够完整保留肽段上的磷酸基团,显著提升位点定位的准确性,在很大程度上解决了这一难题。然而,受限于仪器购置成本与扫描速度,ETD技术通常仅配置于高端质谱机型,尚未在常规实验室中普及。因此,当前绝大多数磷酸化蛋白质组学数据仍来源于基于CID/HCD碎裂模式的质谱仪。如何突破硬件设备的限制,借助先进的数据处理工具从发生中性丢失的质谱图中挖掘出精准的位点信息,已成为提升仪器普适性与磷酸化修饰分析深度的关键切入点。

此外,为解决这一问题,传统研究中引入了Ascore等专门后处理算法进行概率评估,计算每个潜在修饰位点是真正位点的概率,仅将概率高于阈值的结果纳入最终报道<sup>[67]</sup>。但这一方面增加了数据分析的复杂性,另一方面其判定准确性仍受限于原始质谱图的质量。

其次,富集策略的偏好性导致覆盖度不全。如前所述,IMAC和MOAC等经典富集方法均存在选择性偏好,单一富集策略无法捕获细胞内全部磷酸化事件,导致最终获得的磷酸化图谱不完整,甚至可能遗漏关键信息。

最后,低丰度磷酸化信号的检测灵敏度仍面临技术瓶颈。许多在信号通路中位于上游,起到开关作用的关键磷酸化事件的修饰丰度极低。在DDA模式下,质谱仪优先选择高丰度离子进行碎裂,微弱而重要的磷酸化信号往往因强度低于检测阈值而被忽略,这一问题在样本量有限的临床研究中表现得尤为突出<sup>[68]</sup>。

### 3.2 糖基化的传统质谱分析策略与技术瓶颈

与磷酸化修饰类似,糖基化蛋白质组学研究同样始于对修饰肽段的特异性富集。常用

的富集策略包括亲水相互作用层析(hydrophilic interaction chromatography, HILIC)和凝集素亲和层析(lectin affinity chromatography),它们分别利用糖链的强亲水性实现糖肽的分离<sup>[69]</sup>和利用不同凝集素对特定链结构的特异性识别与结合能力实现目标糖肽的富集<sup>[70]</sup>。

富集后的糖肽同样采用液相色谱-串联质谱(LC-MS/MS)进行分析,但其数据解析的复杂性远高于磷酸化等结构简单的翻译后修饰。这是因为在二级质谱碎裂过程中,肽段骨架与连接的复杂糖链会同时发生断裂,生成的谱图中包含肽段碎片离子、糖链碎片离子及丢失部分糖基的肽段离子。

因此,在进入数据库检索环节时,算法不仅需要从前谱图中解析出肽段的氨基酸序列,还需同时鉴定出其所连接糖链的精确组成与结构。这通常需要依托Byonic、pGlyco等专用的搜库软件,并结合预先构建的糖链结构数据库开展综合判定<sup>[71]</sup>。同时,利用现有研究总结出的特异性诊断离子,例如N-乙酰葡萄糖胺会产生的 $m/z$  204.087离子,也可以帮助判断糖肽结构<sup>[72]</sup>。

糖肽鉴定过程的复杂性为糖基化修饰分析带来了与磷酸化修饰分析有所区别的核心技术瓶颈。首先,糖基化修饰的种类繁杂。同一蛋白质的同一糖基化位点上,往往连接着数十种甚至上百种结构各异的糖链。糖链的结构异质性本身已使分析难度大幅增加,同时该特征还会将单个位点对应的质谱信号极大地分散至多种糖肽形式上,导致每种糖肽的信号强度都较低。正如前文所述,低丰度信号极易导致修饰信号被忽略,这也为糖基化修饰的精准分析带来挑战。其次,混合型二级谱图使得传统方法难以通过一张谱图就同时获得高质量的肽段序列和糖链结构信息,这对相关解析算法的性能提出了严苛挑战。最后,鉴定高度依赖糖链数据库,而数据库本身是由已知的糖链结构构成的,对于生物样本中新型、稀有或非典型的糖基化修饰完全无法鉴定,限制了糖基化修饰的相关研究进展<sup>[73]</sup>。

### 3.3 赖氨酸酰化的传统质谱分析策略与技术瓶颈

对于赖氨酸酰化,以当前研究最深入的乙酰化为例,其传统分析流程的首个核心环节同样是特异性富集。与前文所述的富集方法不同,乙酰

化修饰分析最常用的方法是免疫亲和富集法。特异性抗体能够广谱性识别并结合乙酰化修饰的赖氨酸残基,且不受其周围氨基酸序列的影响。通过将抗体固定在琼脂糖或磁性微珠上,即可实现对所有乙酰化修饰肽段的分离。富集后的酰化肽段将经历与磷酸化分析高度相似的LC-MS/MS分析及数据库检索流程<sup>[74]</sup>。

然而,这种高度依赖抗体的富集策略同样也是制约赖氨酸酰化分析的核心瓶颈。首先,抗体的特异性与亲和力可能存在偏差。不同批次抗体的质量参差不齐,可能存在脱靶效应或是在不同序列背景下乙酰化位点亲和力不均一的问题。富集过程也可能存在偏差,使最终获得的乙酰化图谱难以真实、全面地反映实际修饰情况。其次,近年来各类新型的赖氨酸酰化修饰被不断发现,例如乳酸化、琥珀酰化、巴豆酰化等。然而,相对应的抗体开发速度远滞后于新型酰化修饰的发现速度,针对这些新兴修饰,往往长期缺乏可靠的富集工具<sup>[75]</sup>。开发兼具高特异性与高亲和力的抗体,需要经历漫长且极具挑战性的过程,严重阻碍了对这些新型修饰进行全蛋白质组规模的系统性研究。最后,与其他PTM类似,赖氨酸酰化修饰的丰度通常极低<sup>[76]</sup>,这对富集效率提出了极高要求,同时也对抗体质量提出了更严格的要求。

### 3.4 传统方法的挑战与新范式的需求

综上所述,针对不同类型PTM的分析流程

既具有各自的特点,也面临着普遍存在的共性技术瓶颈。首先是灵敏度与覆盖度的局限,PTM信号的低丰度以及富集方法的选择性共同导致了单一的传统工作流程难以获得完整且无偏倚的PTM图谱。其次是数据解析的复杂性与不确定性,化学修饰基团自身的化学特性(如磷酸基团易丢失、糖链复杂且易碎裂等)导致位点定位错误率高、鉴定通量低。最后是对先验知识的过度依赖,无论是糖基化分析高度依赖已知糖链组成的数据库,还是新兴酰化修饰因缺乏抗体而导致研究受阻,均体现出现有研究方法在探索全新、未知修饰事件时存在的不足。因此,亟需能够超越传统算法局限,从复杂而不完整的数据中学习深层规律并做出精准预测的分析新范式。

## 4 AI算法用于PTMs谱图分析

质谱技术是解析PTM谱图的核心手段。然而,PTM的低丰度、修饰位点的不确定性,以及复杂修饰(如糖基化)带来的谱图异质性问题,均为实现高通量、高精度的PTM组学分析带来了巨大挑战。近年来,人工智能(artificial intelligence, AI)算法,特别是深度学习,通过谱图预测和质量控制等策略,从根本上革新PTM谱图的解析范式,示于图5。

### 4.1 AI助力磷酸化蛋白质组的精准位点定位与定量

磷酸化谱图解析的核心挑战在于修饰位点

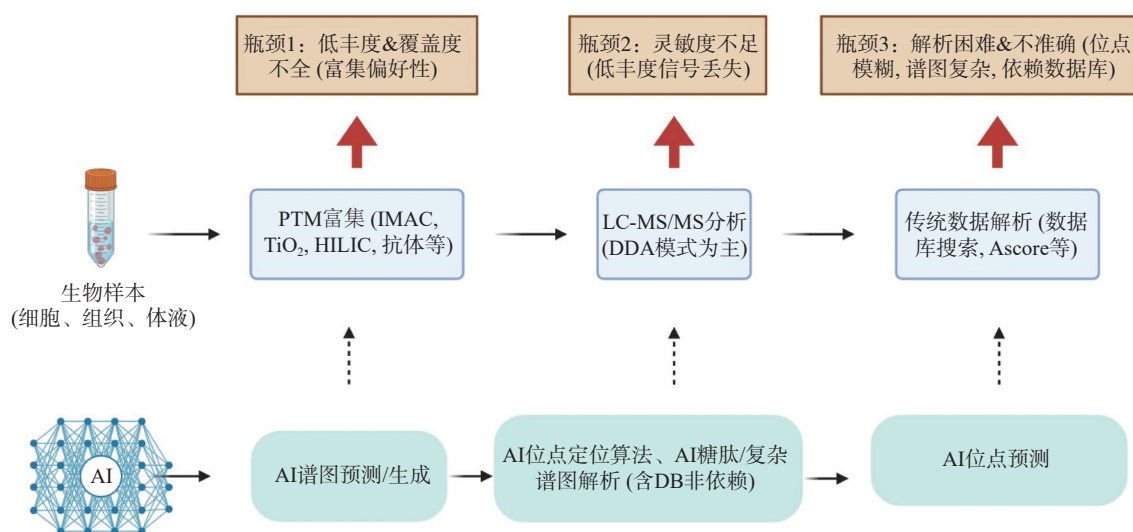


图5 AI赋能PTM谱图的深度解析

Fig. 5 AI-powered deep learning for PTM spectral analysis

的精确定位与假阳性结果的有效控制。策略一为基于深度学习的磷酸化肽段碎片谱图预测及虚拟谱图库构建: DeepPhospho 依托该策略, 显著提升了数据非依赖采集(DIA)模式下磷酸化肽段的鉴定深度与定量准确性<sup>[77]</sup>。策略二为基于神经网络的位点定位与质量控制: 通过训练神经网络识别真实的修饰位点特征, 可有效控制假定位率(FLR)。例如, DeepFLR 基于卷积神经网络(CNN), 借助预训练结合迁移学习, 实现了蛋白质磷酸化修饰位点的精准定位<sup>[78]</sup>; 而 Phosformer 则进一步整合了 Transformer 模型与物种进化信息, 兼容 DIA 和 TMT 数据, 显著增强了对复杂样本中磷酸化位点解析的能力<sup>[79]</sup>。

#### 4.2 AI 应对完整糖肽水平的糖蛋白质组高通量精准解析

糖基化分析的核心瓶颈源于糖链结构的高度异质性和复杂的糖肽碎裂模式。AI 算法则通过精准谱图预测及数据库非依赖匹配策略来应对这一挑战。复旦大学杨芑原教授与中国科学院计算技术研究所贺思敏研究员联合开发了 pGlyco 系列软件(如 pGlyco 2.0), 通过质谱组合碎裂模式联合糖链优先检索策略, 显著提升了糖蛋白质组的鉴定通量、准确性及检索速度, 实现了完整糖肽的一步式鉴定, 并构建了全面且精准的 *N*-糖基化蛋白质组数据库<sup>[71]</sup>; pGlyco3 则进一步拓展了 *O*-糖基化修饰的检测能力, 挖掘出大量蛋白质定向运输的关键信号, 助力糖基化修饰的生物学发现<sup>[80]</sup>。中国科学院大连化学物理研究所叶明亮研究员团队开发了 Glyco-Decipher 软件, 基于不同糖肽的相同肽段骨架具有相似碎裂规律的特征, 建立了基于“模式识别”的肽段序列鉴定新方法, 实现了无需依赖糖库即可鉴定未知糖肽及其修饰基团的功能, 并能够深度解析位点特异性糖型的异质性特征<sup>[73]</sup>。复旦大学乔亮教授联合曹伟倩副研究员研发了 GproDIA 软件, 采用宽窗口二级谱图采集结合深度学习的糖肽信号识别策略, 并结合肽链-糖链二维错误率质控方法, 成功突破了蛋白质糖基化结构解析的难题<sup>[81]</sup>。乔亮教授团队开发了 DeepGP 工具, 首次将图神经网络用于糖结构嵌入, 通过多轮训练策略克服了糖肽质谱数据稀缺的问题, 实现了糖基化肽段质谱图与保留时间的精准预测; 进一步基于完整谱图相似性结合传统搜库打分, 成功区

分了组成相同但结构不同的糖基化修饰<sup>[82]</sup>。此外, 浙江大学方群教授团队开发的 DeepGlyco 算法专注于提升糖肽碎片谱的预测精度, 提高了 *N*-和 *O*-糖肽的鉴定覆盖度<sup>[83]</sup>。

#### 4.3 AI 驱动赖氨酸酰基化谱图解析

在赖氨酸酰基化等新兴 PTM 的研究中, 因缺乏高效富集手段和特异性识别抗体, 其修饰位点的发现与鉴定面临显著的技术瓶颈。在此背景下, AI 驱动的修饰位点预测技术成为突破该瓶颈的关键路径。以研究相对成熟的赖氨酸乙酰化为例, 基于深度学习的预测工具(如 Deep-PLA)能够整合蛋白质序列特征与 HAT/HDAC 等调控酶的特异性信息, 并借助深度神经网络模型, 实现对潜在乙酰化位点的高精度预测。进一步将这些高可信度的预测结果与后续的质谱实验数据相结合, 不仅可显著提升乙酰化位点鉴定的准确性与覆盖度, 还有助于构建系统性的蛋白质乙酰化修饰网络及调控图谱<sup>[84]</sup>。该策略也为其他赖氨酸酰基化修饰(如丙二酰化、巴豆酰化等)的研究提供了可迁移的分析框架。

#### 4.4 AI 赋能 PTM 蛋白质组学驱动的临床转化研究

在突破传统质谱数据解析瓶颈方面, AI 算法, 尤其是深度学习, 展现出巨大的应用潜力。而将 AI 赋能的新型 PTM 分析范式应用于临床研究中, 则是将此类科学潜力转化为实际医疗手段的可靠路径。近年来, 我国学者在这一前沿交叉领域取得了一系列令人瞩目的成果, 充分展示了其在精准医学时代的广阔应用前景。

膀胱癌(bladder cancer, BC)的早期诊断, 尤其是无创筛查仍面临诸多挑战。复旦大学陆豪杰教授团队针对这一临床痛点, 将研究聚焦于尿液中的细胞外囊泡(EVs, 富含外泌体)。这类囊泡直接来源于泌尿系统, 作为液体活检的目标具有极大的临床应用潜力。该团队对大规模临床尿液样本进行高通量尿液 EV *N*-糖基化蛋白质组学分析, 面对高度复杂的糖链结构和组成数据, 借助机器学习算法开展深度挖掘与模式识别。通过对 AI 模型的训练和筛选, 成功构建了由 8 种特定的 *N*-糖链结构组成的 BC 诊断模型。该模型在独立的验证队列中展现出优异的诊断效能, 可精准区分膀胱癌患者与非膀胱癌对照人群。这项工作不仅开创性地建立了尿液 EV *N*-糖组学作为膀胱癌无创诊断的新方法, 更凸显

了AI算法在从复杂组学数据中挖掘高价值临床标志物组合方面的关键作用<sup>[85]</sup>。

KRAS突变是肺腺癌中最常见且最具挑战性的治疗靶点之一。中国科学院上海药物研究所谭敏佳教授团队利用定量磷酸化蛋白质组学技术,对KRAS突变型肺癌细胞及临床样本进行系统研究。面对高度复杂的磷酸化信号网络,该团队借助先进的生物信息学分析方法,成功解析FGF/FGFR信号通路在维持肿瘤恶性表型中的核心调控作用。研究进一步揭示,该通路可通过异常磷酸化下游关键节点蛋白(如GAB1、SHP2),持续驱动癌细胞的增殖与转移潜能。基于上述关键信号轴,该研究团队提出了FGFR抑制剂与MEK抑制剂联合使用的治疗策略,并在临床前模型中验证了该联合方案具有显著的协同抗癌效应。这项研究不仅为KRAS突变肺癌的精准治疗提供了新思路,也展示了AI辅助的磷酸化蛋白质组学在解析复杂疾病机制及指导合理药物联用方面的重要潜力<sup>[86]</sup>。

胰腺癌恶性程度极高,且对铁死亡等治疗方式存在显著的抵抗性。为揭示其潜在的分子机制并发掘新的治疗靶点,北京协和医院陈杰教授团队运用位点特异性N-糖基化蛋白质组学技术,对包括胰腺癌在内的多种癌细胞系进行系统性分析。由于糖基化分析需同时实现肽段和糖链的鉴定,数据高度复杂,因此构建高效、精准的计算分析流程成为确保研究可靠的关键。通过对胰腺导管腺癌(pancreatic ductal adenocarcinoma, PDAC)细胞的糖基化蛋白质组学分析,该研究团队成功鉴定出一个关键的糖基化调控轴:糖基转移酶B3GNT3可介导细胞膜蛋白4F2hc的特异性N-糖基化修饰。进一步研究表明,该异常糖基化修饰能够稳定4F2hc蛋白,进而抑制细胞铁死亡通路,最终赋予胰腺癌细胞抵抗铁死亡的能力。这项工作不仅揭示了糖基化在胰腺癌耐药中的关键调控作用,也表明通过计算蛋白质组学手段筛选得到的异常糖基化轴或相关蛋白,有望成为逆转耐药的新靶点或预测疗效的生物标志物,为攻克胰腺癌治疗耐药难题提供了新方向<sup>[87]</sup>。

综上,AI算法正在从3个层面重塑PTM蛋白质组学的分析范式:1)从“实验驱动”到“计算驱动”:通过精准的谱图预测,AI能够构建覆盖度远超实验谱库的虚拟谱库,为DIA等质谱采

集模式赋能;2)从“模糊匹配”到“精准定位”:通过复杂的神经网络模型,AI能够实现对修饰位点、糖链组成等信息的精准解析,并进行可靠的质量控制;3)从“定性发现”到“动态定量”:AI算法与DIA等技术的深度融合,使得对PTM动态变化进行高精度、高通量的定量成为可能,为揭示疾病机制提供了更强大的工具。未来,将AI模型与多组学数据进行深度整合,是揭示PTM复杂调控网络、推动精准医学发展的必然趋势。

## 5 挑战与展望

尽管PTM研究已取得长足进展,但该领域仍面临着从技术方法到临床转化的多重挑战。其中,分析技术的检测灵敏度与数据覆盖度是当前最核心的技术瓶颈,主要源于两大难题:首先,绝大多数PTM的修饰占有率(PTM site stoichiometry)极低。例如,乙酰化修饰率超过20%的蛋白仅占全蛋白质组的4%<sup>[76]</sup>,导致微弱的PTM信号极易被高丰度的非修饰蛋白信号所掩盖。其次,临床样本(尤其是体液样本)中蛋白质丰度的动态范围极宽(可达 $10^6 \sim 10^{10}$ 倍),使得传统质谱技术难以同时捕获高、低丰度的PTM事件。

为突破上述瓶颈,未来技术发展的关键方向在于微量样本制备的自动化与一体化。微流控芯片技术是实现这一目标的核心平台,其能够将细胞捕获、成像、蛋白质高效提取及特定PTM肽段富集等复杂流程集成于方寸芯片之上,从而有效解决临床样本量稀少和处理过程中目标物样品损失的难题。近期,中国台湾中央研究院化学研究所Yu-Ju Chen教授与Hsiung-Lin Tu教授团队开发的iPhosChip一体化平台正是该技术方向的典型范例<sup>[88]</sup>。该技术创新性地将多路复用微流控芯片与数据非依赖采集质谱(DIA-MS)技术联用,并通过构建与样本输入量相匹配的谱图库,成功将磷酸化蛋白质组学的检测灵敏度提升至单细胞水平。该技术已被用于绘制肺癌EGFR靶向治疗耐药的单细胞磷酸化图谱,不仅发现了耐药细胞中细胞骨架重塑异质性等新型分子特征,还实现了对混合谱系肿瘤亚型的精准分层,为指导肿瘤精准治疗策略的制定提供了直接的分子依据。

在硬件系统集成之外,新材料与新试剂的革

新是突破上述瓶颈的另一关键方向。通过优化富集材料(如磁珠、色谱柱填料)的表面化学性质,或研发兼具更高亲和力与特异性的反应试剂,可进一步降低样本的起始量,并提升目标PTM肽段的回收率与检测选择性。

PTM组学产生的海量高维数据对生物信息学分析提出了极高要求。AI算法为此提供了强有力的解决方案,主要体现在以下2个前沿方向:1)多模态PTM数据整合与临床应用。AI算法能够高效整合多种PTM数据及其他组学数据,从中挖掘出单一组学无法揭示的复杂疾病规律。例如,加州大学洛杉矶分校Peipei Ping教授团队开发的MIND-S<sup>[88]</sup>。该平台创新性地结合了多头注意力机制(用于捕获序列中的长距离依赖关系)与图神经网络(用于挖掘蛋白质互作网络信息),不仅实现了对多种PTM的同时精准预测,还能够系统性阐明这些PTM在健康与疾病状态下参与的关键生物学过程,为解析复杂的PTM调控网络提供了强大的计算工具。2)单细胞PTM组学分析是揭示细胞异质性的新维度。鉴于细胞异质性是肿瘤耐药、免疫逃逸和疾病进展的核心驱动力,仅依靠宏观组织样本分析难以解答关键问题。因此,未来将新兴的单细胞质谱技术与AI算法相结合,将为揭示驱动疾病的稀有细胞亚群(如肿瘤干细胞、耐药细胞株)内部独特的PTM调控网络提供不可或缺的关键技术支持。

PTM研究领域仍在持续拓展。除琥珀酰化、巴豆酰化等新兴修饰外,必然还存在大量功能未知的PTM类型等待被发现和表征。而大量已被鉴定的PTM位点,仍需要借助高通量功能筛选平台(如CRISPR修饰文库、磷酸化模拟突变体库)进行系统性地功能验证。此外,将研究对象从蛋白质拓展至RNA(如glycoRNA)等其他生物大分子,将是未来极具潜力的研究方向。PTM与基因突变、表观遗传调控之间亦存在复杂的交互作用,未来也需要构建整合多维度信息的大型数据库与知识图谱,以系统性解析PTM交互的调控网络。

## 6 结语

蛋白质翻译后修饰在调控蛋白质功能、介导信号转导及维持代谢稳态中发挥关键作用,但

其低丰度性、高动态性及结构复杂性对传统分析方法带来巨大挑战,制约了对PTM调控网络的深入解析。本文系统梳理了磷酸化、糖基化、赖氨酸酰化等关键PTM的生物学功能及经典质谱分析策略,并指出其在检测灵敏度、修饰位点准确定位及覆盖广度方面的技术局限。本文还进一步探讨了AI驱动的深度学习方法在PTM谱图预测、修饰位点鉴定及非依赖数据库检索方面的应用潜力,凸显了其在提升分析精度与通量方面的优势。未来,通过整合多组学数据、单细胞技术及新型富集材料,PTM研究有望实现从机制探索向临床转化的跨越式发展,为疾病的精准诊断与靶向治疗提供新路径。

**致谢:** 谨以此文,深切怀念引领我踏上学术征途的恩师杨芑原教授。作为科研战略先行者,先生以深邃视野洞察学科脉络,推动公共技术平台赋能多学科发展;作为育人者,他怀海纳百川之胸襟,鼓励学生大胆探索、试错创新。惟愿不负先生昔日教诲与殷切期许。

## 参考文献:

- [1] EBERT T, TRAN N, SCHURGERS L, STENVINKEL P, SHIELS P G. Ageing-oxidative stress, PTMs and disease[J]. *Molecular Aspects of Medicine*, 2022, 86: 101-099.
- [2] BEKKER-JENSEN D B, BERNHARDT O M, HOGREBE A, MARTINEZ-VAL A, VERBEKE L, GANDHI T, KELSTRUP C D, REITER L, OLSEN J V. Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries[J]. *Nature Communications*, 2020, 11(1): 787.
- [3] ZHU F, LI L, CHEN Y, PAN Y, ZHANG W, LI L, CAI L, ZHAO X, ZHAO H, WANG S, JIA L. CRL3<sup>Keap1</sup> E3 ligase facilitates ubiquitin-mediated degradation of oncogenic SRX to suppress colorectal cancer progression[J]. *Nature Communications*, 2024, 15(1): 10 536.
- [4] RIBOLDI G M, FARAVELLI I, KUWAJIMA T, DELESTRÉE N, DERMENTZAKI G, de PLANELL-SAGUER M, RINCHETTI P, HAO L T, BEATTIE C C, CORTI S, PRZEDBORSKI S, MENTIS G Z, LOTTI F. Sumoylation regulates the assembly and activity of the SMN complex[J]. *Nature Communications*, 2021, 12(1):

- 5 040.
- [5] NISHI H, HASHIMOTO K, PANCHENKO A R. Phosphorylation in protein-protein binding: effect on stability and function[J]. *Structure*, 2011, 19(12): 1 807-1 815.
- [6] SOLIS G P, KAZEMZADEH A, ABRAMI L, VALNOHOVA J, ALVAREZ C, GISOU van der GOOT F, KATANAEV V L. Local and substrate-specific S-palmitoylation determines subcellular localization of Gao[J]. *Nature Communications*, 2022, 13(1): 2 072.
- [7] CHARLET J, DUYMICH C E, LAY F D, MUNDJERG K, SØRENSEN K D, LIANG G, JONES P A. Bivalent regions of cytosine methylation and H3K27 acetylation suggest an active role for DNA methylation at enhancers[J]. *Molecular Cell*, 2016, 62(3): 422-431.
- [8] WU Z. Impact of post-translational modification on the genesis and progression of diseases[J]. *Molecular Aspects of Medicine*, 2022, 86: 101 105.
- [9] WU Z, JANKOWSKI V, JANKOWSKI J. Irreversible post-translational modifications-emerging cardiovascular risk factors[J]. *Molecular Aspects of Medicine*, 2022, 86: 101 010.
- [10] YE J, WAN H, CHEN S, LIU G P. Targeting Tau in Alzheimer's disease: from mechanisms to clinical therapy[J]. *Neural Regeneration Research*, 2024, 19(7): 1 489-1 498.
- [11] PARRA BRAVO C, NAGUIB S A, GAN L. Cellular and pathological functions of Tau[J]. *Nature Reviews Molecular Cell Biology*, 2024, 25(11): 845-864.
- [12] SZABAT M, PAGE M M, PANZHINSKIY E, SKOVSO S, MOJIBIAN M, FERNANDEZ-TAJES J, BRUIN J E, BROUND M J, LEE J T C, XU E E, TAGHIZADEH F, O'DWYER S, van de BUNT M, MOON K M, SINHA S, HAN J, FAN Y, LYNN F C, TRUCCO M, BORCHERS C H, FOSTER L J, NISLOW C, KIEFFER T J, JOHNSON J D. Reduced insulin production relieves endoplasmic reticulum stress and induces  $\beta$  cell proliferation[J]. *Cell Metabolism*, 2016, 23(1): 179-193.
- [13] HSU J M, LI C W, LAI Y J, HUNG M C. Posttranslational modifications of PD-L1 and their applications in cancer therapy[J]. *Cancer Research*, 2018, 78(22): 6 349-6 353.
- [14] JENNINGS E Q, FRITZ K S, GALLIGAN J J. Biochemical genesis of enzymatic and non-enzymatic post-translational modifications[J]. *Molecular Aspects of Medicine*, 2022, 86: 101 053.
- [15] SINGH V, RAM M, KUMAR R, PRASAD R, ROY B K, SINGH K K. Phosphorylation: implications in cancer[J]. *The Protein Journal*, 2017, 36(1): 1-6.
- [16] VASANT C, SANKARAMANIVEL S, JANA M, RAJARAM R, RAMASAMI T. Non-enzymatic phosphorylation of bovine serum albumin by Cr(V) complexes: role in Cr(VI)-induced phosphorylation and toxicity[J]. *Molecular and Cellular Biochemistry*, 2005, 275(1): 153-164.
- [17] MANNING G, WHYTE D B, MARTINEZ R, HUNTER T, SUDARSANAM S. The protein kinase complement of the human genome[J]. *Science*, 2002, 298(5 600): 1 912-1 934.
- [18] KIM M J, CHOI W G, AHN K J, CHAE I G, YU R, BACK S H. Reduced EGFR level in eIF2 $\alpha$  phosphorylation-deficient hepatocytes is responsible for susceptibility to oxidative stress[J]. *Molecules and Cells*, 2020, 43(3): 264-275.
- [19] LIN C W, NOCKA L M, STINGER B L, DeGRANDCHAMP J B, NUGENT LEW L J, ALVAREZ S, PHAN H T, KONDO Y, KURIYAN J, GROVES J T. A two-component protein condensate of the EGFR cytoplasmic tail and Grb2 regulates Ras activation by SOS at the membrane[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, 119(19): e2122531119.
- [20] OLSEN J V, VERMEULEN M, SANTAMARIA A, KUMAR C, MILLER M L, JENSEN L J, GNAD F, COX J, JENSEN T S, NIGG E A, BRUNAK S, MANN M. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis[J]. *Science Signaling*, 2010, 3(104): ra3.
- [21] ARORA M, MOSER J, HOFFMAN T E, WATTS L P, MIN M, MUSTEANU M, RONG Y, ILL C R, NANGIA V, SCHNEIDER J, SANCLEMENTE M, LAPEK J, NGUYEN L, NIESSEN S, DANN S, VanARSDALE T, BARBACID M, MILLER N, SPENCER S L. Rapid adaptation to CDK2 inhibition exposes intrinsic cell-cycle plasticity[J]. *Cell*, 2023, 186(12): 2 628-2 643. e21.
- [22] CHEN C C L, GOYAL P, KARIMI M M, ABILDGAARD M H, KIMURA H, LORINCZ M C. H3S10ph broadly marks early-replicating domains in interphase ESCs and shows reciprocal antagonism with H3K9me2[J]. *Genome Research*, 2018, 28(1): 37-51.
- [23] HITOSUGI T, KANG S, VANDER HEIDEN M G, CHUNG T W, ELF S, LYTHGOE K, DONG S, LONIAL S, WANG X, CHEN G Z, XIE J, GU T L, POLAKIEWICZ R D, ROESEL J L, BOGGON T J, KHURI F R, GILLILAND D G, CANTLEY L C,

- KAUFMAN J, CHEN J. Tyrosine phosphorylation inhibits PKM2 to promote the Warburg effect and tumor growth[J]. *Science Signaling*, 2009, 2(97): ra73.
- [24] LUGO T G, PENDERGAST A M, MULLER A J, WITTE O N. Tyrosine kinase activity and transformation potency of bcr-abl oncogene products[J]. *Science*, 1990, 247(4 946): 1 079-1 082.
- [25] QUINTÁS-CARDAMA A, CORTES J. Molecular biology of BCR-ABL1-positive chronic myeloid leukemia [J]. *Blood*, 2009, 113(8): 1 619-1 630.
- [26] RIX U, HANTSCHHEL O, DÜRNBERGER G, REMSING RIX L L, PLANAVSKY M, FERNBACH N V, KAUPE I, BENNETT K L, VALENT P, COLINGE J, KÖCHER T, SUPERTI-FURGA G. Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets[J]. *Blood*, 2007, 110(12): 4 055-4 063.
- [27] PARK K, YU C J, KIM S W, LIN M C, SRIURANPONG V, TSAI C M, LEE J S, KANG J H, ALLEN CHAN K C, PEREZ-MORENO P, BUTTON P, AHN M J, MOK T. First-line erlotinib therapy until and beyond response evaluation criteria in solid tumors progression in Asian patients with epidermal growth factor receptor mutation-positive non-small-cell lung cancer: the ASPIRATION study[J]. *JAMA Oncology*, 2016, 2(3): 305-312.
- [28] EICHLER J. Protein glycosylation[J]. *Current Biology*, 2019, 29(7): R229-R231.
- [29] VARKI A. Biological roles of glycans[J]. *Glycobiology*, 2017, 27(1): 3-49.
- [30] JUBELI E, MOINE L, VERGNAUD-GAUDUCHON J, BARRATT G. E-selectin as a target for drug delivery and molecular imaging[J]. *Journal of Controlled Release*, 2012, 158(2): 194-206.
- [31] SCHJOLDAGER K T, NARIMATSU Y, JOSHI H J, CLAUSEN H. Global view of human protein glycosylation pathways and functions[J]. *Nature Reviews Molecular Cell Biology*, 2020, 21(12): 729-749.
- [32] NG B G, FREEZE H H. Perspectives on glycosylation and its congenital disorders[J]. *Trends in Genetics*, 2018, 34(6): 466-476.
- [33] TATEKOSHI Y, MAHMOODZADEH A, SHAPIRO J S, LIU M, BIANCO G M, TATEKOSHI A, CAMP S D, de JESUS A, KOLEINI N, deLa TORRE S, WASSERSTROM J A, DILLMANN W H, THOMSON B R, BEDI K C, MARGULIES K B, WEINBERG S E, ARDEHALI H. Protein *O*-glcNAcylation and hexokinase mitochondrial dissociation drive heart failure with preserved ejection fraction[J]. *Cell Metabolism*, 2025, 37(7): 1 584-1 600. e10.
- [34] FLYNN R A, PEDRAM K, MALAKER S A, BATISTA P J, SMITH B A H, JOHNSON A G, GEORGE B M, MAJZOUB K, VILLALTA P W, CARETTE J E, BERTOZZI C R. Small RNAs are modified with *N*-glycans and displayed on the surface of living cells[J]. *Cell*, 2021, 188(16): 4 470.
- [35] XIE Y, CHAI P, TILL N A, HEMBERGER H, LEBEDENKO C G, PORAT J, WATKINS C P, CALDWELL R M, GEORGE B M, PERR J, BERTOZZI C R, GARCIA B A, FLYNN R A. The modified RNA base acp3U is an attachment site for *N*-glycans in glycoRNA[J]. *Cell*, 2024, 187(19): 5 228-5 237. e12.
- [36] POLEVODA B, SHERMAN F. The diversity of acetylated proteins[J]. *Genome Biology*, 2002, 3(5): reviews0006.1.
- [37] DRIESSEN H P C, de JONG W W, TESSER G I, BLOEMENDAL H. The mechanism of *N*-terminal acetylation of protein[J]. *Critical Reviews in Biochemistry*, 1985, 18(4): 281-325.
- [38] SADOUL K, WANG J, DIAGOURAGA B, KHOCHBIN S. The tale of protein lysine acetylation in the cytoplasm[J]. *Journal of Biomedicine & Biotechnology*, 2011, 2011: 970 382.
- [39] ARNESEN T, van DAMME P, POLEVODA B, HELSENS K, EVJENTH R, COLAERT N, VARHAUG J E, VANDEKERCKHOVE J, LILLEHAUG J R, SHERMAN F, GEVAERT K. Proteomics analyses reveal the evolutionary conservation and divergence of *N*-terminal acetyltransferases from yeast and humans[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(20): 8 157-8 162.
- [40] STARHEIM K K, GEVAERT K, ARNESEN T. Protein *N*-terminal acetyltransferases: when the start matters[J]. *Trends in Biochemical Sciences*, 2012, 37(4): 152-161.
- [41] COLLEPARDO-GUEVARA R, PORTELLA G, VENDRUSCOLO M, FRENKEL D, SCHLICK T, OROZCO M. Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: a multiscale computational study[J]. *Journal of the American Chemical Society*, 2015, 137(32): 10 205-10 215.
- [42] SHOGREN-KNAAK M, ISHII H, SUN J M, PAZIN M J, DAVIE J R, PETERSON C L. Histone H4-K16 acetylation controls chromatin structure and protein interac-

- tions[J]. *Science*, 2006, 311(5 762): 844-847.
- [43] MCGINTY R K, TAN S. Nucleosome structure and function[J]. *Chemical Reviews*, 2015, 115(6): 2 255-2 273.
- [44] SHVEDUNOVA M, AKHTAR A. Modulation of cellular processes by histone and non-histone protein acetylation[J]. *Nature Reviews Molecular Cell Biology*, 2022, 23(5): 329-349.
- [45] DELMORE J E, ISSA G C, LEMIEUX M E, RAHL P B, SHI J, JACOBS H M, KASTRITIS E, GILPATRICK T, PARANAL R M, QI J, CHESI M, SCHINZEL A C, McKEOWN M R, HEFFERNAN T P, VAKOC C R, LEIF BERGSAGEL P, GHOBRIAL I M, RICHARDSON P G, YOUNG R A, HAHN W C, ANDERSON K C, KUNG A L, BRADNER J E, MITSIADES C S. BET bromodomain inhibition as a therapeutic strategy to target c-Myc[J]. *Cell*, 2011, 146(6): 904-917.
- [46] GUTBROD M J, ROCHE B, STEINBERG J I, LAKHANI A A, CHANG K, SCHORN A J, MARTIENSSEN R A. Dicer promotes genome stability *via* the bromodomain transcriptional co-activator BRD4[J]. *Nature Communications*, 2022, 13(1): 1 001.
- [47] WU F, MUSKAT N H, DVILANSKY I, KOREN O, SHAHAR A, GAZIT R, ELIA N, ARBELY E. Acetylation-dependent coupling between G6PD activity and apoptotic signaling[J]. *Nature Communications*, 2023, 14(1): 6 208.
- [48] XUE M, FENG T, CHEN Z, YAN Y, CHEN Z, DAI J. Protein acetylation going viral: implications in antiviral immunity and viral infection[J]. *International Journal of Molecular Sciences*, 2022, 23(19): 11 308.
- [49] LIN Y, LIN A, CAI L, HUANG W, YAN S, WEI Y, RUAN X, FANG W, DAI X, CHENG J, ZHANG J, CHEN W, YE Q, CHEN X, ZHANG J. ACSS2-dependent histone acetylation improves cognition in mouse model of Alzheimer's disease[J]. *Molecular Neurodegeneration*, 2023, 18(1): 47.
- [50] GRÄFF J, REI D, GUAN J S, WANG W Y, SEO J, HENNIG K M, NIELAND T J F, FASS D M, KAO P F, KAHN M, SU S C, SAMIEI A, JOSEPH N, HAGGARTY S J, DELALLE I, TSAI L H. An epigenetic blockade of cognitive functions in the neurodegenerating brain[J]. *Nature*, 2012, 483(7 388): 222-226.
- [51] BROOKS G A. Lactate as a fulcrum of metabolism[J]. *Redox Biology*, 2020, 35: 101 454.
- [52] GAFFNEY D O, JENNINGS E Q, ANDERSON C C, MARENTETTE J O, SHI T, SCHOU OXVIG A M, STREETER M D, JOHANNSEN M, SPIEGEL D A, CHAPMAN E, ROEDE J R, GALLIGAN J J. Non-enzymatic lysine lactoylation of glycolytic enzymes[J]. *Cell Chemical Biology*, 2020, 27(2): 206-213. e6.
- [53] CHEN L, HUANG L, GU Y, CANG W, SUN P, XIANG Y. Lactate-lactylation hands between metabolic reprogramming and immunosuppression[J]. *International Journal of Molecular Sciences*, 2022, 23(19): 11 943.
- [54] ZHANG D, TANG Z, HUANG H, ZHOU G, CUI C, WENG Y, LIU W, KIM S, LEE S, PEREZ-NEUT M, DING J, CZYZ D, HU R, YE Z, HE M, GEORGE ZHENG Y, SHUMAN H A, DAI L, REN B, ROEDER R G, BECKER L, ZHAO Y. Metabolic regulation of gene expression by histone lactylation[J]. *Nature*, 2019, 574 (7 779): 575-580.
- [55] CAI J, SONG L, ZHANG F, WU S, ZHU G, ZHANG P, CHEN S, DU J, WANG B, CAI Y, YANG Y, WAN J, ZHOU J, FAN J, DAI Z. Targeting SRSF10 might inhibit M2 macrophage polarization and potentiate anti-PD-1 therapy in hepatocellular carcinoma[J]. *Cancer Communications*, 2024, 44(11): 1 231-1 260.
- [56] YU J, CHAI P, XIE M, GE S, RUAN J, FAN X, JIA R. Histone lactylation drives oncogenesis by facilitating m6A reader protein YTHDF2 expression in ocular melanoma[J]. *Genome Biology*, 2021, 22(1): 85.
- [57] JIANG J, HUANG D, JIANG Y, HOU J, TIAN M, LI J, SUN L, ZHANG Y, ZHANG T, LI Z, LI Z, TONG S, MA Y. Lactate modulates cellular metabolism through histone lactylation-mediated gene expression in non-small cell lung cancer[J]. *Frontiers in Oncology*, 2021, 11: 647 559.
- [58] HUA G, LIU Y, LI X, XU P, LUO Y. Targeting glucose metabolism in chondrosarcoma cells enhances the sensitivity to doxorubicin through the inhibition of lactate dehydrogenase-A[J]. *Oncology Reports*, 2014, 31(6): 2 727-2 734.
- [59] ZHANG N, ZHANG Y, XU J, WANG P, WU B, LU S, LU X, YOU S, HUANG X, LI M, ZOU Y, LIU M, ZHAO Y, SUN G, WANG W, GENG D, LIU J, CAO L, SUN Y.  $\alpha$ -Myosin heavy chain lactylation maintains sarcomeric structure and function and alleviates the development of heart failure[J]. *Cell Research*, 2023, 33(9): 679-698.
- [60] LI Y E, LIU S, WANG L, DU Y, WU L, CHEN H, ZHU T, LIN J, XIONG S, WANG Y, ZHENG Q, ZOU R, LIN L, LI Z, WANG L, GE J, REN J, ZHANG Y. March2 alleviates aortic aneurysm/dissection by regulating PKM2 polymerization[J]. *Circulation Research*, 2025,

- 136(8): e73-e93.
- [61] WANG N, WANG W, WANG X, MANG G, CHEN J, YAN X, TONG Z, YANG Q, WANG M, CHEN L, SUN P, YANG Y, CUI J, YANG M, ZHANG Y, WANG D, WU J, ZHANG M, YU B. Histone lactylation boosts reparative gene activation post-myocardial infarction[J]. *Circulation Research*, 2022, 131(11): 893-908.
- [62] ANDERSSON L, PORATH J. Isolation of phosphoproteins by immobilized metal (Fe<sup>3+</sup>) affinity chromatography[J]. *Analytical Biochemistry*, 1986, 154(1): 250-254.
- [63] THINGHOLM T E, JØRGENSEN T J D, JENSEN O N, LARSEN M R. Highly selective enrichment of phosphorylated peptides using titanium dioxide[J]. *Nature Protocols*, 2006, 1(4): 1 929-1 935.
- [64] LARSEN M R, THINGHOLM T E, JENSEN O N, ROEPSTORFF P, JØRGENSEN T J D. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns[J]. *Molecular & Cellular Proteomics*, 2005, 4(7): 873-886.
- [65] JENSEN S S, LARSEN M R. Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques[J]. *Rapid Communications in Mass Spectrometry*, 2007, 21(22): 3 635-3 645.
- [66] AEBERSOLD R, MANN M. Mass spectrometry-based proteomics[J]. *Nature*, 2003, 422(6 928): 198-207.
- [67] BEAUSOLEIL S A, VILLÉN J, GERBER S A, RUSH J, GYGI S P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization[J]. *Nature Biotechnology*, 2006, 24(10): 1 285-1 292.
- [68] OLSEN J V, BLAGOEV B, GNAD F, MACEK B, KUMAR C, MORTENSEN P, MANN M. Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks[J]. *Cell*, 2006, 127(3): 635-648.
- [69] HÄGGLUND P, BUNKENBORG J, ELORTZA F, JENSEN O N, ROEPSTORFF P. A new strategy for identification of *N*-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation[J]. *Journal of Proteome Research*, 2004, 3(3): 556-566.
- [70] WU C C, LU Y T, YEH T S, CHAN Y H, DASH S, YU J S. Identification of fucosylated SERPINA1 as a novel plasma marker for pancreatic cancer using lectin affinity capture coupled with iTRAQ-based quantitative glycoproteomics[J]. *International Journal of Molecular Sciences*, 2021, 22(11): 6 079.
- [71] LIU M, ZENG W, FANG P, CAO W, LIU C, YAN G, ZHANG Y, PENG C, WU J, ZHANG X, TU H, CHI H, SUN R, CAO Y, DONG M, JIANG B, HUANG J, SHEN H, WONG C, HE S, YANG P. pGlyco 2.0 enables precision *N*-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification[J]. *Nature Communications*, 2017, 8(1): 438.
- [72] ROUSHAN A, WILSON G M, KLETTER D, SEN K I, TANG W, KIL Y J, CARLSON E, BERN M. Peak filtering, peak annotation, and wildcard search for glycoproteomics[J]. *Molecular & Cellular Proteomics*, 2021, 20: 100 011.
- [73] FANG Z, QIN H, MAO J, WANG Z, ZHANG N, WANG Y, LIU L, NIE Y, DONG M, YE M. Glyco-Decipher enables glycan database-independent peptide matching and in-depth characterization of site-specific *N*-glycosylation[J]. *Nature Communications*, 2022, 13(1): 1 900.
- [74] CHOUDHARY C, KUMAR C, GNAD F, NIELSEN M L, REHMAN M, WALTHER T C, OLSEN J V, MANN M. Lysine acetylation targets protein complexes and co-regulates major cellular functions[J]. *Science*, 2009, 325(5 942): 834-840.
- [75] ZHANG Z, TAN M, XIE Z, DAI L, CHEN Y, ZHAO Y. Identification of lysine succinylation as a new post-translational modification[J]. *Nature Chemical Biology*, 2011, 7(1): 58-63.
- [76] BAEZA J, DOWELL J A, SMALLEGAN M J, FAN J, AMADOR-NOGUEZ D, KHAN Z, DENU J M. Stoichiometry of site-specific lysine acetylation in an entire proteome[J]. *Journal of Biological Chemistry*, 2014, 289(31): 21 326-21 338.
- [77] LOU R, LIU W, LI R, LI S, HE X, SHUI W. DeepPhospho accelerates DIA phosphoproteome profiling through *in silico* library generation[J]. *Nature Communications*, 2021, 12(1): 6 685.
- [78] ZONG Y, WANG Y, YANG Y, ZHAO D, WANG X, SHEN C, QIAO L. DeepFLR facilitates false localization rate control in phosphoproteomics[J]. *Nature Communications*, 2023, 14(1): 2 269.
- [79] ZHOU Z, YEUNG W, GRAVEL N, SALCEDO M, SOLEYMANI S, LI S, KANNAN N. Phosformer: an explainable transformer model for protein kinase-specific phosphorylation predictions[J]. *Bioinformatics*, 2023, 39(2): btad046.
- [80] ZENG W F, CAO W Q, LIU M Q, HE S M, YANG P Y. Precise, fast and comprehensive analysis of intact gly-

- copeptides and modified glycans with pGlyco3[J]. *Nature Methods*, 2021, 18(12): 1 515-1 523.
- [81] YANG Y, YAN G, KONG S, WU M, YANG P, CAO W, QIAO L. GproDIA enables data-independent acquisition glycoproteomics with comprehensive statistical control[J]. *Nature Communications*, 2021, 12(1): 6 073.
- [82] ZONG Y, WANG Y, QIU X, HUANG X, QIAO L. Deep learning prediction of glycopeptide tandem mass spectra powers glycoproteomics[J]. *Nature Machine Intelligence*, 2024, 6(8): 950-961.
- [83] YANG Y, FANG Q. Prediction of glycopeptide fragment mass spectra by deep learning[J]. *Nature Communications*, 2024, 15(1): 2 448.
- [84] YU K, ZHANG Q, LIU Z, DU Y, GAO X, ZHAO Q, CHENG H, LI X, LIU Z. Deep learning based prediction of reversible HAT/HDAC-specific lysine acetylation[J]. *Briefings in Bioinformatics*, 2020, 21(5): 1 798-1 805.
- [85] LI Y, FU B, WANG M, CHEN W, FAN J, LI Y, LIU X, WANG J, ZHANG Z, LU H, ZHANG Y. Urinary extracellular vesicle *N*-glycomics identifies diagnostic glycosignatures for bladder cancer[J]. *Nature Communications*, 2025, 16(1): 2 292.
- [86] LIU Z, LIU Y, QIAN L, JIANG S, GAI X, YE S, CHEN Y, WANG X, ZHAI L, XU J, PU C, LI J, HE F, HUANG M, TAN M. A proteomic and phosphoproteomic landscape of KRAS mutant cancers identifies combination therapies[J]. *Molecular Cell*, 2021, 81(19): 4 076-4 090. e8.
- [87] MA H, CHEN X, MO S, ZHANG Y, MAO X, CHEN J, LIU Y, TONG W, LU Z, YU S, CHEN J. Targeting *N*-glycosylation of 4F2hc mediated by glycosyltransferase B3GNT3 sensitizes ferroptosis of pancreatic ductal adenocarcinoma[J]. *Cell Death and Differentiation*, 2023, 30(8): 1 988-2 004.
- [88] MUNEEER G, GEBREYESUS S T, CHEN C S, LEE T T, YU F, LIN C A, HSIEH M S, NESVIZHSHKII A I, HO C C, YU S L, TU H L, CHEN Y J. Mapping nanoscale-to-single-cell phosphoproteomic landscape by chip-DIA[J]. *Advanced Science*, 2025, 12(1): e2402421.

(收稿日期: 2025-09-06; 修回日期: 2025-12-01)