



内源性蛋白质交联的生物质谱分析进展

韦秋实¹, 宫鹏云², 李诺敏³, 刘超^{2*}, 张永谦^{1*}, 邓玉林¹

1. 北京理工大学生命学院, 北京 100081;

2. 北京航空航天大学大数据精准医疗高精尖创新中心, 北京 100191;

3. 北京理工大学医工融合研究院, 北京 100081

* 联系人, E-mail: liuchaobuaa@buaa.edu.cn; zyq@bit.edu.cn

2021-07-15 收稿, 2021-09-07 修回, 2021-09-08 接受, 2021-09-09 网络版发表

北京理工大学科技创新计划前沿交叉与学科创新专项计划(2021CX11017)资助

摘要 内源性蛋白质交联通常指细胞内蛋白质分子间或蛋白质分子内中的两个氨基酸被肽键之外的共价键相连, 其形成会对蛋白质构象和功能造成影响。质谱具有高通量、高灵敏度等特点, 是大规模鉴定内源性蛋白质交联的工具之一。结合质谱在组学水平上对内源性蛋白质交联的研究工作层出不穷, 但在样品处理、数据采集、定性分析、定量分析和生信分析等环节还有诸多问题仍待解决, 构建一套完整的内源性蛋白质交联技术体系将极大推动交联组学研究的开展。因此, 本文对内源性蛋白质交联相关研究工作进行了概述, 总结了目前内源性蛋白质交联研究中取得的重要成果, 梳理了基于质谱的内源性蛋白质交联分析工作流程, 最后对内源性蛋白质交联在组学水平上的发展提出了展望。

关键词 内源性交联组学, 生物质谱, 分子内交联, 分子间交联

内源性蛋白质交联是一种特殊的蛋白质翻译后修饰(post-translational modifications, PTMs)^[1], 它的形成会改变蛋白质结构或蛋白质相互作用并影响蛋白质功能^[2]。交联蛋白质组学旨在阐明细胞内的蛋白质和网络结构特征, 从分子水平上理解生命机制, 最终用于药物的设计^[3]。交联质谱技术具有高通量、高灵敏度等特点^[4], 是目前内源性蛋白质交联研究的重要工具之一。基于质谱的内源性蛋白质交联组学研究也逐渐成为生命科学研究的重要方向。

本文介绍了目前内源性蛋白质交联研究取得的重要成果, 并对大规模鉴定内源性蛋白质交联方法进行总结, 最后对内源性蛋白质交联的研究发展方向进行了展望。

1 蛋白质交联

蛋白质交联(简称交联)通常指蛋白质中的两个氨

基酸被肽键之外的共价键连接起来^[5]。蛋白质交联包括分子间和分子内的交联, 这两者的区别在于交联的两端连接的氨基酸是否来自同一个蛋白。连接的部分按来源可分为外源性蛋白质交联(简称外源性交联)和内源性蛋白质交联(简称内源性交联)。

1.1 外源性交联

外源性交联指通过化学交联试剂将蛋白质中的氨基酸连接起来的化学交联, 即在蛋白质溶液中加入化学试剂, 连接蛋白质或蛋白质复合物中氨基酸侧链的两个官能团。这类交联是研究蛋白质空间结构和相互作用的工具, 不影响蛋白质本身的结构和功能^[6]。外源性交联的形成能将蛋白质拓扑结构、构象状态和蛋白质相互作用(protein-protein interactions, PPIs)信息通过交联肽段的形式保留下来^[7~9]。通常能形成交联的氨基

引用格式: 韦秋实, 宫鹏云, 李诺敏, 等. 内源性蛋白质交联的生物质谱分析进展. 科学通报, 2021, 66: 4633–4648

Wei Q S, Gong P Y, Li N M, et al. Mass spectrometry-based endogenous protein crosslinking omics (in Chinese). Chin Sci Bull, 2021, 66: 4633–4648, doi: 10.1360/TB-2021-0700

酸残基在空间上的距离相对较近,这个距离取决于交联剂的臂长,因此可通过交联位点信息来推测蛋白质的三级结构.蛋白质相互作用通常是瞬时的,可利用交联剂将它们固定下来,使其在后续的处理过程中将始终被保留下来.由于外源性交联是由外加的化学交联剂结合氨基酸形成,因此在利用质谱进行外源性交联的鉴定时,可以专门针对化学交联剂添加各种功能基团使其碎裂具有特殊规律,简化对交联的鉴定^[10].并且外源性交联的化学性质相对稳定,形成交联肽段的类型也相对单一.

1.2 内源性交联

不同于外源性交联,内源性交联的形成依靠的是细胞中存在的小分子或者化学键将氨基酸残基连接起来.内源性交联在蛋白质内部天然存在,是蛋白质的一种特殊的翻译后修饰.它的形成包括酶催化和非酶催化,对蛋白质的结构和功能有重要影响^[11].相较于外源性交联中依赖的化学试剂,在利用质谱进行数据采集时,内源性交联的化学性质不够稳定,容易发生异构和断裂.这使得在使用质谱对内源性交联研究时通常需要先探究碎裂规律,才能实现对内源性交联的准确鉴定.并且在一个蛋白质中,会存在多种内源性交联.有报道发现,30%的枯草芽孢杆菌孢子外壳蛋白中存在蛋白质交联,包括异肽键、二硫键等^[12].另外,内源性交联有可逆的也有不可逆的,它们在维持蛋白质结构稳定、氧化还原调节等方面起关键作用^[13].现有研究表明,交联与衰老和病变组织中的蛋白质聚集密切相关^[14],若有一种能直接识别蛋白质交联和连接类型的方法将能有助于理解这些过程.

2 内源性交联研究进展

目前已发现,内源性蛋白质交联包括二硫键、异肽键、色氨酸交联、酪氨酸交联、甲醛交联、丙酮醛(methylglyoxal, MG)交联、NOS桥等(图1),它们均被证明在生物过程中有着重要影响.

2.1 二硫键

连接不同肽链或同一肽链中两个不同半胱氨酸残基的巯基的化学键称为二硫键.二硫键的形成是在酶的催化下进行,涉及的酶包括蛋白质二硫键异构酶(protein disulfide isomerase, PDI)和内质网氧化还原蛋白1蛋白(endoplasmic reticulum oxidoreductin 1 protein,

Ero1p).最终结果是同一蛋白质或不同蛋白质中的两个半胱氨酸侧链巯基被共价结合,所形成的交联包括蛋白质分子内和分子间的交联^[15].作为结构的重要组成部分,二硫键的形成能保证蛋白质或蛋白质复合物具有正确的构象,并提高它们的稳定性.许多分泌蛋白,如抗体、生长因子、细胞外基质蛋白和细胞表面受体或转运体,都富含二硫键,它们对治疗具有重要价值^[16].二硫键的可逆性使其能够作为一个分子开关来调节酶或转录因子的活性,以应对环境的变化.另外,在氧化应激状态下,蛋白质中的二硫键也会发生改变.要充分了解含有二硫键的蛋白质生物学功能和规律,必须准确地定位参与二硫化形式或自由硫醇的两个半胱氨酸残基位置.在确定形成二硫键的两个半胱氨酸位点会面临两个问题:一是二硫键广泛分布在蛋白质中,想要确定形成二硫键的半胱氨酸位点,就需要利用蛋白质组学的方法进行;二是在样品制备过程中,二硫键容易发生重排,这将导致错误定位蛋白质中形成二硫键的半胱氨酸位点^[17].

二硫键是最常见的蛋白质内源性交联,在蛋白质中广泛地分布,具有重要的生物学功能,与之相关的研究也比较丰富. Butera等人^[18]对817种血浆蛋白的三级结构进行分析,推测蛋白质与二硫键的比例近似为1:5. Farrah等人^[19]鉴定到2000个血浆蛋白,并估计存在有10000个二硫键. Lu等人^[20]在蛋白质组学尺度上鉴定出大肠杆菌周质部分存在199个二硫键,在人脐静脉内皮细胞分泌的蛋白质中鉴定出568个二硫键. Liu等人^[21]利用质谱技术解析拟南芥PCP-Bs(pollen coat protein B-class peptides)结构中的二硫键配对情况,揭示了花粉和柱头相互识别的分子机理.从结构上讲,二硫键保证了蛋白质的适当折叠并形成结构亚型,维持蛋白质执行其生物学功能所必需的天然构象.非天然的二硫键加入到蛋白质中可以增加蛋白质的稳定性.例如,最初一些缺乏二硫键的蛋白质被证明在人为引入二硫键后更稳定,而具有天然二硫键的蛋白质的稳定性也随着引入额外的二硫键而增加.在某些情况下,人为引入二硫键可以增加蛋白质的半衰期,降低蛋白质的自聚集性和免疫原性.还有一些二硫键被称为变构二硫键,负责蛋白质的有效生物学功能,这种键的断裂将导致蛋白质活性发生变化. Hogg^[22]对变构二硫键在血液和癌细胞中的功能作用进行了总结.基于IgG1设计的药物利妥昔单抗(rituximab)和曲妥珠单抗(trastuzumab),部分二硫键被还原并烷基化修饰后,可以增加与某些

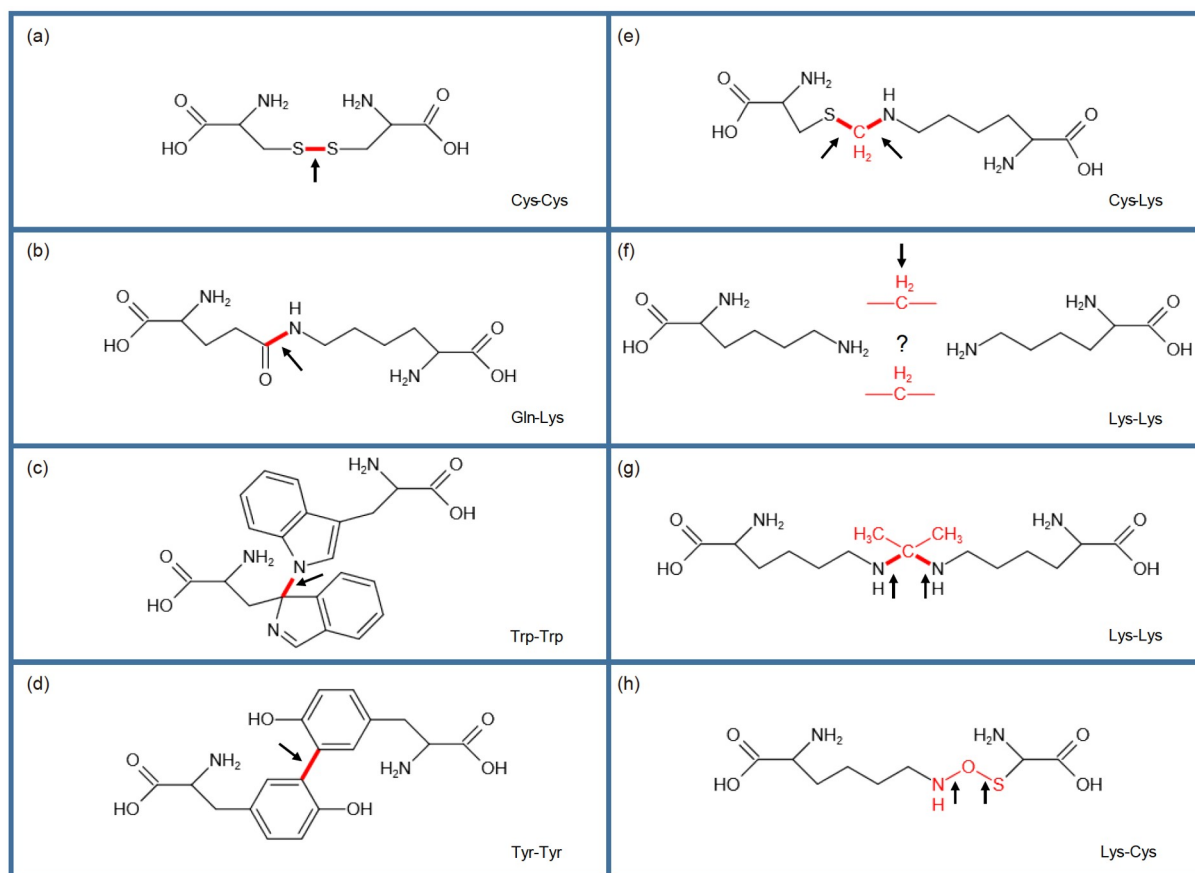


图1 (网络版彩色)内源性蛋白质交联。(a) 二硫键; (b) 异肽键; (c) 色氨酸交联; (d) 酪氨酸交联; (e) 甲醛交联(12 Da); (f) 甲醛交联(24 Da); (g) 丙酮醛交联; (h) NOS桥。箭头标注的是交联形成的共价键位置。甲醛交联(24 Da)的连接机制尚未清楚

Figure 1 (Color online) Endogenous protein crosslinking. (a) Disulfide bond; (b) isopeptide bond; (c) dityryptophan; (d) dityrosine; (e) formaldehyde crosslinking (12 Da); (f) formaldehyde crosslinking (24 Da); (g) methylglyoxal crosslinking; (h) NOS bridge. The arrows indicate the positions of covalent bonds formed by crosslinking. Mechanism of formaldehyde crosslinking (24 Da) is not clear

同型Fc伽马受体的结合亲和力,但也导致与其他Fc伽马受体的结合减少^[22~24]。绘制蛋白质中的二硫键连接模式,将为蛋白质稳定性、结构-功能关系以及二硫键介导的蛋白质异构体的研究提供重要信息。此外,二硫键的表征在生物制药的开发过程中具有重要意义,它可以确保生物制剂的安全性和药效。因此,越来越需要高效的分析方法来准确表征蛋白质中的二硫键,特别是那些具有治疗作用的蛋白质^[17]。

2.2 异肽键

蛋白质中的氨基酸通过侧链上的羧基或氨基形成的酰胺键被称为异肽键。与异肽键形成密切相关的生物过程包括泛素化(ubiquitylation)^[25]、类泛素化(su-moylation)^[26]、转谷氨酰胺(transglutamination)^[27]、分拆酶介导的细胞表面蛋白锚定(sortase-mediated cell

surface protein anchoring)^[28]和纤毛的形成(pilus formation)^[27,29,30]。这些过程有两个共同的特点: (1) 除了转谷氨酰胺过程中赖氨酸的目标是谷氨酰胺侧链酰胺基,其余都是涉及一个蛋白质上的赖氨酸 α -氨基与另一个蛋白质上的 ϵ -羧基的反应; (2) 所有这些过程都由酶介导,其活性位点上的半胱氨酸会形成一种瞬时硫酯中间体,然后这个中间体在赖氨酸 ϵ -氨基的亲核攻击下分解形成异肽键^[27]。

目前的研究发现,转谷氨酰胺过程与很多疾病有密切的联系。这一过程是在谷氨酰胺转氨酶的作用下进行的。谷氨酰胺转氨酶(transglutaminase)家族由TG1、TG2(tTG)、TG3、TG4~TG7、血浆凝血因子XIIIa和红细胞膜蛋白4.2等9个成员组成^[31,32]。谷氨酰胺转氨酶参与蛋白质多肽中谷氨酰胺残基和赖氨酸残基之间的交联反应等蛋白质翻译后修饰,并能稳定蛋白

质结构,催化蛋白质聚集体的形成.前5个成员在人体中以活跃的形式存在,但只有TG1、TG2、TG3、TG6和TG7在大脑中被检测到,特别是在额叶皮质胞浆中,TG2表达最高.TG2在神经元和胶质细胞膜、细胞外部分,甚至在神经母细胞瘤细胞的细胞核中都有表达.谷氨酰胺转氨酶催化形成的异肽键交联既可以是分子内交联,也可以是分子间交联^[33,34].分子内交联会改变蛋白质的构象,而分子间交联的形成会导致刚性、稳定和高度不溶性的蛋白质复合物产生.许多生理过程如血液凝固、细胞程序性死亡过程中衰老小体的形成、头发和指甲的形成以及白内障产生等均与异肽键的形成密切相关.另外,很多神经退行性疾病也与异肽键的形成有关联.如患有阿尔茨海默病(Alzheimer's disease, AD)患者的大脑通常由于大脑皮层大量神经元的丢失而萎缩,体积和重量有明显的减少^[35].在组织学上其显著的特征是神经性老年斑与神经原纤维缠结.神经性老年斑的主要成分是具有大量交联的 β -淀粉样蛋白^[36].尽管这些交联背后的机制仍缺乏具体的认识,但大量的蛋白质交联是该病的一些关键特征.鉴于TG2催化的交联可导致AD相关脑区乙酰胆碱或Tau蛋白的积累和聚集,绘制蛋白质中异肽键的连接模式或许可以有助于对这些疾病机制的理解^[37].Schmitt等人^[38]利用质谱研究血块样本,将已知的异肽键数目从1个增加到了100多个,并由此提出了血凝块的溶解是由纤维蛋白原的结合以及丝氨酸蛋白酶抑制剂复合物的形成来共同调节的.

2.3 色氨酸交联

色氨酸交联是由两个色氨酸残基丢失了两个氢原子而形成的交联.色氨酸交联是一种不可逆的蛋白质交联^[39].目前,对生物样本中色氨酸交联的检测主要基于荧光和抗原抗体反应,也有少量基于质谱的色氨酸交联研究.

色氨酸交联的发现来自对牛 α -晶体蛋白的研究.牛 α -晶体蛋白在犬尿喹啉酸(kynurenine acid)存在的情况下接受厌氧照射(330~390 nm),导致色氨酸-色氨酸的形成.牛 α -晶体蛋白含有3个色氨酸残基(α A-晶体蛋白中的Trp9、 α B-晶体蛋白中的Trp9和Trp60).在这3个色氨酸残基中都发现了色氨酸-色氨酸交联(即检测到2.017 Da质量修饰的多肽),研究者由此提出犬尿喹啉酸的三重态是从蛋白质的色氨酸和酪氨酸残基中提取电子,产生相应的自由基,导致蛋白质聚集的假说^[40].

Paviani等人^[39,41]利用质谱研究证明,人源超氧化物歧化酶1(human superoxide dismutase 1, hSOD1)的共价二聚体是由两个hSOD1单体通过色氨酸-色氨酸交联结合而成的(Trp32-Trp32).hSOD1依赖于碳酸氢盐的过氧化物酶活性的文章发表之后,许多报道描述了在肽和蛋白质氧化过程中,通过各种氧化系统形成色氨酸交联.尽管色氨酸在蛋白质中含量较低,但它们具有与其他蛋白质和细胞结构相互作用的独特潜力,使色氨酸-色氨酸交联在蛋白质功能障碍和蛋白质聚集中具有重要作用.这些相互作用可能会在色氨酸氧化后发生改变,并对细胞内稳态产生显著影响^[39].此外,色氨酸残基的氧化还可能导致蛋白质二聚、寡聚化和聚集,破坏细胞的蛋白酶抑制作用.最近的研究还表明,硫氧还蛋白还原酶(thioredoxin reductase)含有进化保守和表面暴露的色氨酸残基(Trp114),对氧化非常敏感,可能在氧化应激条件下通过齐聚和交联发挥氧化还原传感器功能.硫氧还蛋白还原酶在自然状态下以同二聚体的形式存在,但在氧化过程中,两种不同的酶二聚体通过Trp114-Trp114之间的交联形成四聚体.这些共价连接的酶在用抗氧化剂抗癌药物处理的细胞提取物中显著增加,同时也发现细胞死亡的增加.这些结果表明,硫氧还蛋白还原酶中的Trp114-Trp114交联可能是一种诱导细胞死亡的调节机制^[42].

2.4 酪氨酸交联

氧化应激可以导致两个酪氨酸(Tyr)残基之间形成共价交联,形成酪氨酸交联(dityrosine, DT).这种翻译后修饰也被认为是氧化应激、衰老和各种病理条件的临床生物标志物^[43].在阿尔茨海默病^[44]、帕金森病(Parkinson's disease, PD)^[45]患者脑的脂褐质颗粒^[46]和脑脊液中^[47]均发现存在酪氨酸交联.据报道,在人类动脉粥样硬化病变的斑块中也富集到了酪氨酸交联的低密度脂蛋白(low density lipoproteins, LDL)^[48].

Mukherjee等人^[49]研究常用的质谱裂解技术确定合成的双酪氨酸交联A β 二聚体在质谱中的碎裂情况.据此他们建立了一套碎片离子规则,并提出了识别酪氨酸交联肽的通用工作流程,成功地鉴定了人类血红蛋白和 α -突触核蛋白中的酪氨酸交联位点,为内源性酪氨酸交联所涉及的蛋白质鉴定提供了基础.Ren等人^[50]和Kumar等人^[51]在研究氧气对含有酪氨酸的生物内源分子自进化中的调控作用发现,分子中所含有的酪氨酸基团,在富氧条件下优先形成黑色素类似物,而

在乏氧条件下优先形成酪氨酸交联。该研究也为药物递送、光热治疗以及新型生物器件的设计提供了新思路。另外, Kumar等人^[51]发现, 冈比亚按蚊(*Anopheles gambiae*)中肠分泌的过氧化物酶可催化酪氨酸形成交联。由酪氨酸交联组成的网络导致冈比亚按蚊的肠道对免疫激发因子的通透性降低, 阻止上皮免疫的激活, 为疟疾寄生虫在中肠腔内发育提供了适宜的环境。

2.5 甲醛交联

甲醛交联是通过甲醛分子将蛋白质分子中的氨基酸连接起来而形成的交联。甲醛是生活中的常见致癌物, 会对人体健康造成危害。生物体本身也会产生一定量的甲醛。这些内源甲醛是多条生理代谢途径的原料或产物, 也与细胞的存活与生长密切相关。有研究表明, 甲醛作为一碳化合物能够被纳入“一碳循环”(one-carbon-cycle), 用于DNA和必需氨基酸的合成^[52]。甲醛也能够与蛋白质上的亲核性氨基酸侧链发生加成反应; 甲醛还能作为蛋白质交联试剂, 在蛋白质的亲核性残基(如赖氨酸)之间形成交联, 使蛋白质发生变性而失活^[53]。

枯草芽孢杆菌中的HxIR蛋白是一种转录激活因子, 能够在甲醛刺激下激活下游基因转录, 进行功能获得性(gain-of-function)的生理响应。Zhu等人^[53]发现, 位于HxIR蛋白N端第一个 α 螺旋的半胱氨酸(Cys11)与赖氨酸(Lys13)残基之间形成了位点特异的“亚甲基桥”, 并通过位点突变来确认该交联反应与HxIR蛋白功能的联系。结果显示, 在Cys11或Lys13上的突变, HxIR都失去响应甲醛的能力, 表明HxIR蛋白会通过Cys11和Lys13特异性识别甲醛, 形成亚甲基桥这一独特的化学修饰, 实现转录功能的激活。另外, Tayri-Wilk等人^[54]将甲醛作为一种交联剂用于蛋白质鉴定从而开展结构研究, 提出24 Da的甲醛交联结构代替以往的12 Da亚甲基桥结构。虽然交联部分真实的化学结构尚未证实, 但是该研究为甲醛交联鉴定提供了新的思路。

2.6 丙酮醛交联

丙酮醛交联是通过丙酮醛分子将两个氨基酸连接起来而形成的交联。MG是一种醛类化合物, 易聚合形成各种环状和非环状化合物。MG也是胞内酶促和非酶促反应过程中的产物^[55]。许多研究表明, MG可以修饰很多生物大分子, 特别是蛋白质。MG可与精氨酸残基反应生成咪唑啉酮类衍生物(MG-H1、MG-H2、MG-

H3)^[56-58]、argpyrimidine^[59]和 N_δ -(4-carboxy-4,6-dimethyl-5,6-dithydroxy-1,4,5,6-tetrahydropyrimidine-2-yl)-L-ornithine(THP)^[60], 也可以与赖氨酸残基反应生成 N_ϵ -(1-carboxyethyl)lysine(CEL)^[61]和赖氨酸二聚体(MG-derived lysine dimer, MOLD)^[62]。另外, MG可以同时与赖氨酸和精氨酸残基反应形成MODIC^[63]。这些由MG衍生的糖基化产物(advanced glycation end product, AGE)在体内实验中被证明与糖尿病和其他疾病的并发症相关^[64-69]。

MG在A β 淀粉斑块的形成中发挥着重要作用。Shuster等人^[70]研究发现, 人和鼠A β 蛋白在序列上存在的区别(Arg5-Gly5、Tyr10-Phe10、His13-Arg13), 导致mA β 蛋白中的Lys16可能不像hA β 蛋白一样被丙酮醛交联形成聚集, 也可以解释鼠脑在氧化应激状态下的聚集比人脑少且没有出现AD样症状的现象。Bollong等人^[71]发现, 当KEAP1接触葡萄糖降解过程中产生的代谢分子丙酮醛时会选择性地修饰KEAP1, 在近端半胱氨酸和精氨酸残基之间形成甲基咪唑交联(methylimidazole crosslink between proximal cysteine and arginine residues, MICA)。这种翻译后修饰导致KEAP1二聚体的积累, NRF2聚集和NRF2转录程序激活, 从而开始启动解毒机制。

2.7 NOS桥

NOS桥是在半胱氨酸和赖氨酸残基之间形成的一种共价交联。NOS桥可在酶中充当变构氧化还原开关, 控制酶的活性, 有望被推广到蛋白质设计和小分子抑制剂开发等领域。

NOS桥是由Wensien等人^[72]对淋病奈瑟氏球菌(*neisseria gonorrhoeae*)转醛酶(NgTAL)的研究中发现的。淋病奈瑟氏球菌代谢过程中的关键酶NgTAL其酶活性可以通过调节其氧化还原环境发生可逆性激活-抑制。最初认为这是由于蛋白质内两个半胱氨酸间形成二硫键引起的, 但对该蛋白序列内3个半胱氨酸残基位点的单独突变后, 仅Cys38的突变体能使氧化还原激活无效。在氧化和还原状态下对NgTAL进行结晶, 可获得两种状态的结构。在氧化状态下, Cys38与Lys8之间发生交联且存在一个额外的氧原子, 形成带有NOS桥的共价交联键, 蛋白中其他两个半胱氨酸残基(Cys87和Cys90)保持硫醇形式; 还原态下不存在Cys38和Lys8的交联, 两个残基均不含化学修饰, 且有多个空间构象可选择^[72]。

3 催化蛋白质交联形成的酶

酶在蛋白质交联的形成过程中发挥着重要作用。表1对参与蛋白质交联形成的酶的生理功能和生化特性进行了总结。对于酶催化形成的蛋白质交联,可以分为两种类型:一是由转移酶(EC 2)和水解酶(EC 3)通过蛋白质-酶-硫酯中间体直接催化形成共价键;二是由氧化还原酶(EC 1)酶介导的活性物质(EC 1)与蛋白质进一步反应形成蛋白质网络^[5]。这些能催化蛋白质交联形成的酶中,有一部分能在体内发挥作用直接催化蛋白质形成内源性交联,如Protein disulfide isomerase、Transferases transglutaminase等;另一部分是在体外用于催化特定蛋白质形成交联,改变蛋白质的特性,如Subtilisin、Laccase、Peroxidase,具有催化效率高,绿色环保等特点。在食品工业、皮革和纺织加工等中已经被应用于体外催化蛋白质交联,具有重要应用价值。但是目前酶的蛋白质交联技术仍处于起步阶段^[5],后续需要在生物体内发现更多具有催化蛋白质交联形成的酶,同时对催化机理进行深入研究。

4 交联质谱技术在内源性交联中的应用

化学交联结合质谱技术(chemical cross-linking coupled with mass spectrometry, 简称交联质谱技术或CXMS),是20世纪末、21世纪初发展起来的方法^[92]。它

利用化学交联剂(chemical cross-linker)处理蛋白质样品,将空间距离足够接近、可以与交联剂反应的两个氨基酸以共价键形式连接起来,然后利用基于质谱的蛋白质组学方法分析交联产物^[93]。交联质谱技术主要用于蛋白质结构和蛋白质相互作用的研究。在交联剂的固定作用下,使许多结构松散、易解离的复合物也能得以保留下来。而交联质谱相较核磁共振(nuclear magnetic resonance, NMR)、X射线晶体衍射(X-ray crystallography)等技术又具有高通量、高灵敏度、低检出限等优势,使交联质谱被广泛应用^[94]。

交联质谱的工作流程与蛋白质组学中的bottom-up方式类似。整个过程基本可分为6个阶段(图2)^[95]。第一阶段,待分析的蛋白质样品通过生化分离或亲和选择从细胞裂解物或组织中分离出来。第二阶段,向蛋白质样品中加入化学交联剂进行交联反应,形成蛋白质交联。第三阶段,将上一阶段得到的样本在蛋白酶的作用下降解为肽段,为后续的肽测序作准备。第四阶段,得到的肽段通过液相色谱仪的洗脱后,进入质谱仪的电喷雾室,在强电场作用下,洗脱的肽段能够带上电荷,使质谱仪能够对肽段进行分析。第五阶段,质谱仪对这些带上电荷的洗脱肽段进行质谱分析得到一级谱图(MS1谱)。第六阶段,计算机生成这些肽段的优先级列表,然后进行串联质谱分析得到二级谱图(MS2谱)。经过上述阶段后,再通过对上述过程产生的谱图进行分析,即可

表 1 催化蛋白质交联形成的酶

Table 1 Overview of enzymes used for protein crosslinking

酶	主要功能	辅助因子/离子依赖 ^{a)}	目标基团	形成机理	文献
Protein disulfide isomerase (EC 5.3.4.1)	催化蛋白质二硫键的形成和异构化	—	半胱氨酸	形成二硫键	[73]
Transferases transglutaminase (EC 2.3.2.13)	催化不同蛋白质交联	钙离子(仅在真核生物中)	谷氨酰胺侧链	形成异肽键	[74~76]
Hydrolases Sortase A (EC 3.4.22.70)	细胞表面蛋白质附着	钙离子	LPXTG基序中苏氨酸甘氨酸间的肽键	形成苏氨酸和甘氨酸间的肽键	[77,78]
Subtilisin (EC 3.4.21.62)	胞外蛋白质水解	—	多肽C端的氨基酸	形成肽键	[79]
Oxidoreductases tyrosinase (EC 1.14.18.1)	抑制黑色素形成	铜离子	酪氨酸侧链; 酚类化合物	形成醌; 1,4-加成	[80~83]
Laccase (EC 1.10.3.2)	木质素降解与生物合成	铜离子	酪氨酸侧链; 酚类化合物	自由基耦合	[84~87]
Peroxidase (EC 1.11.1.x)	多种化合物氧化	血红素	酪氨酸侧链; 酚类化合物	自由基耦合	[88,89]
Lysyl oxidase (EC 1.4.3.13) /amine oxidase (EC 1.4.3.6)	胶原蛋白和弹性蛋白的交联	Lysine tyrosinyl quinone (LTQ); 铜离子	赖氨酸侧链	醛的形成; 自发形成醛醇缩合和席夫碱产物	[90,91]

a) “—”表示不需要依赖辅助因子或金属离子的激活

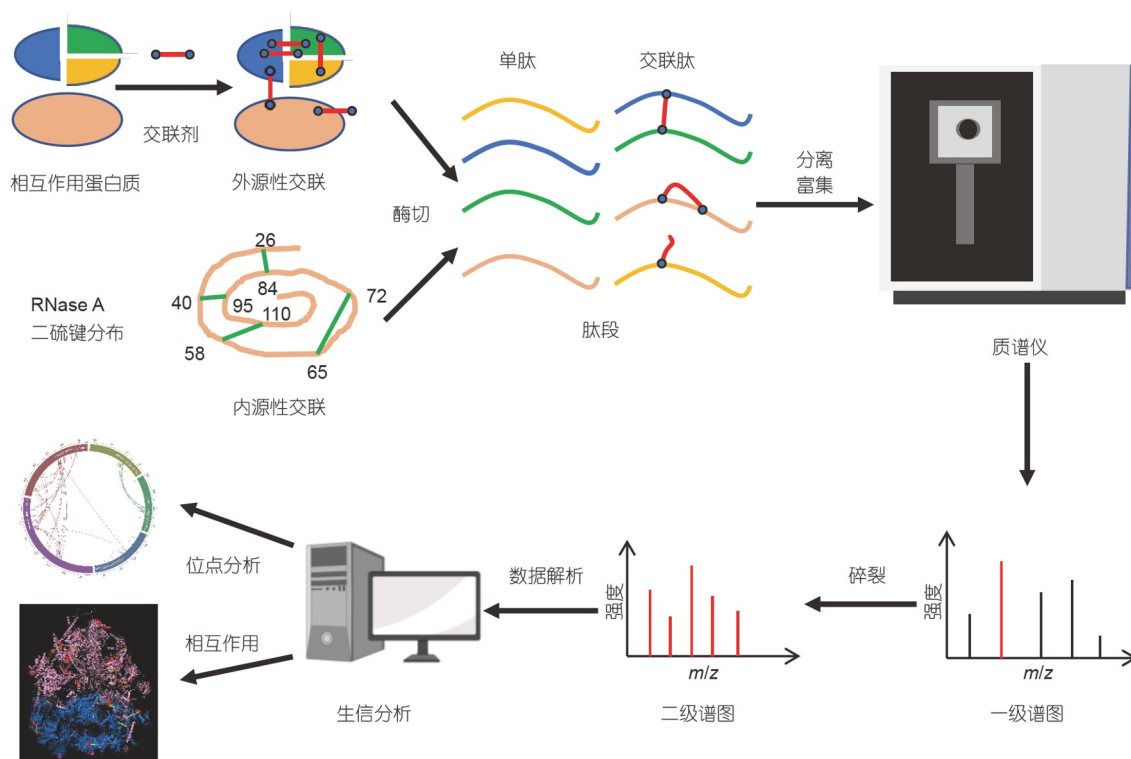


图2 (网络版彩色) 基于质谱分析蛋白质交联的流程. 部分图片由biorender.com网站制作

Figure 2 (Color online) Flowchart of protein crosslinking analysis by mass spectrometry. Some images are created by biorender.com

实现对肽段和蛋白质的定性定量分析. 利用质谱数据分析软件对采集到的谱图进行分析, 给出交联肽段鉴定结果^[6,96,97]. 根据这些结果中的交联位点位置信息, 还可以再使用蛋白质空间结构分析软件如PyMOL^[98]等对交联位点空间距离进行分析, 判断两个交联位点的可及性来获取蛋白质的空间结构信息以及相互作用关系.

对内源性交联位点网络的绘制是探究内源性蛋白质交联对生理过程影响的基础. 二硫键作为研究较早的内源性蛋白质交联, 在20世纪末就已利用NMR、X射线晶体衍射技术对其在蛋白质中的排列进行测定, 并用来理解蛋白质分子的折叠. 但是上述两种方法对样品含量要求较高且不能进行高通量的测量, 限制了内源性蛋白质交联的研究开展. 由于内源性交联与外源性交联在结构上类似, 也逐渐将交联质谱技术应用到内源性交联的研究上. 略微不同的是, 内源性交联在蛋白质中天然存在, 因此可以跳过上述步骤中的第二阶段. 交联质谱技术在内源性交联中的应用可分为5个环节: 样品预处理、数据采集、定性分析、定量分析和生物信息学分析.

4.1 样品预处理

样品预处理指样品在送入质谱采集数据前所作的处理, 主要目的是让内源性交联能更容易通过质谱采集与准确鉴定到, 包括酶解、分离富集和氨基酸封闭等. 具体处理方式需要依据内源性交联的特点来设计. 基于质谱的内源性蛋白质交联研究通过bottom-up的分析方法来获取氨基酸位点信息, 因此蛋白质混合物酶解成肽段混合物是样品预处理中的重要一步. 实际研究中, 对肽段混合物进行分离富集也能一定程度提高内源性交联的解析率. 另外, 内源性交联涉及维持蛋白质功能稳定, 容易出现断裂和重排情况, 这也是样品预处理过程中需要考虑的.

酶解的目的是将发生交联的蛋白质消化为交联肽段并用于质谱的数据采集. 交联肽段按结构可分为3种基本形式: mono-link、loop-link和cross-link(图3). 然而, 受到酶切位点和交联位点具有特异性, 以及氨基酸在蛋白质中分布不均的影响, 复杂样品酶解后得到的交联肽段并不是都能得到上述3种简单交联形式. 如血液中D-dimer含量是评价凝血功能异常的重要指标, 酶解

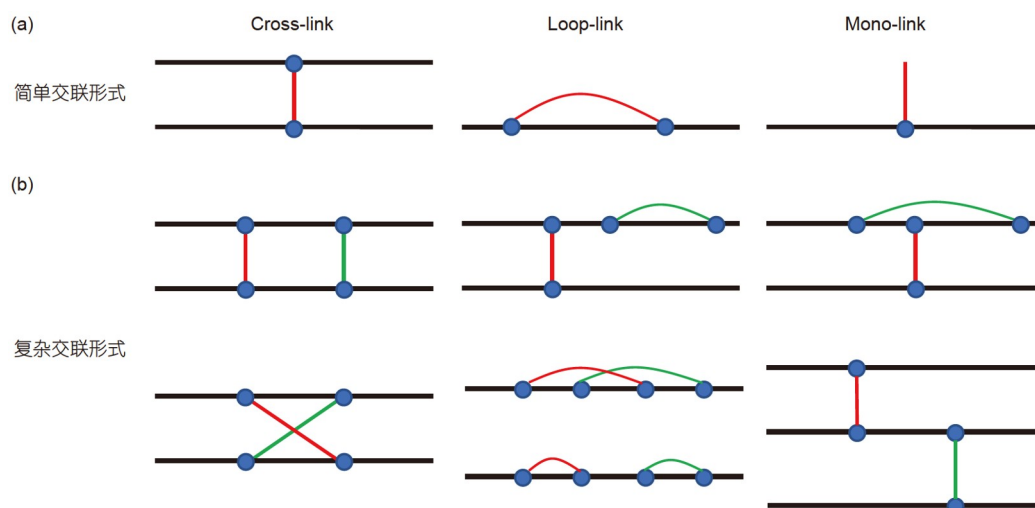


图3 (网络版彩色)交联肽段形式。(a)简单交联形式包括cross-link、loop-link和mono-link。(b)复杂交联形式是简单交联形式的组合
Figure 3 (Color online) Structure of crosslinking peptides. (a) Simple crosslinking includes cross-link, loop-link and mono-link. (b) Complex crosslinking is considered as the combinations of single crosslinking

后会产生一个含有两个异肽键的交联肽段^[74]。Lu等人^[20]在研究10个常见蛋白质中的二硫键分布时发现,即使使用多种酶进行酶解,在得到的交联肽段中仍含有存在两个二硫键的复杂交联形式肽段。而Ursem等人^[12]对枯草芽孢杆菌孢子外壳蛋白质进行分析时发现,这些蛋白质中存在多种类型的内源性交联。内源性交联在形式和组成上有着多种组合情况,所以大规模鉴定内源性交联时需要考虑更多样的复杂交联形式(图3)。对于这种复杂交联形式的肽段,也可以采用混合酶切的方式处理为简单交联肽段形式,如Zhao等人^[99]提出的Smart Cutter方法可有效提高交联肽段鉴定的覆盖率。

对于复杂样品,发生内源性交联的蛋白质相对总体是少数的,这些蛋白质在酶解处理后得到的肽段中,具有内源性交联的肽段相较单肽也是少数。交联肽段的含量和丰度偏低将导致很多内源性交联无法被采集和鉴定出来。因此,进行分离富集在样品预处理中也是十分必要的。在分离过程中常采用亲和色谱技术、强阳离子交换技术(strong cation exchange, SCX)和分子筛色谱技术(size exclusion chromatography, SEC)。在富集过程中,内源性交联可以分为蛋白质和肽段水平。发生内源性交联的蛋白质由于酶解后蛋白质已经被消化成肽段,所以蛋白质水平的富集在酶解前一步进行。而肽段水平的富集可通过对交联肽段上的氨基酸或者内源性交联独有的特性进行,如Zhang等人^[100]依据D-dimer上含有4个组氨酸提出使用镍固定化金属亲和

层析(nickel immobilized metal affinity chromatography, Ni(II)-IMAC))进行富集。另外,在数据采集过程中也可以使用四极杆和离子阱等方法去除高峰期蛋白质实现对低丰度蛋白质的富集^[101]。

氨基酸封闭主要是避免在数据采集过程中内源性交联发生断裂和重排。内源性蛋白质交联多数靠化学共价键将蛋白质上的氨基酸直接连接起来,部分内源性交联的形成过程是可逆的。受到加入的试剂及体系的环境影响,内源性交联很容易发生改变。这些改变后的交联如果被错误认为成天然的交联将会导致后续的研究朝着错误方向进行。这样的情况在对二硫键的研究中经常遇到。在还原条件下,二硫键会发生断裂。在碱性pH或高温条件下^[102~104],二硫键会与蛋白质中含有游离巯基的半胱氨酸之间发生重排,即使使用常规的烷基化试剂如iodoacetamide(IAM)或iodoacetic acid (IAA)将具有游离巯基的半胱氨酸封闭起来,仍无法避免这样的情况出现。Lu等人^[20]在对二硫键进行研究时证明,使用N-ethylmaleimide(NEM)进行烷基化可解决二硫键的重排问题。因此二硫键的分析样品一般在室温和微酸性pH下进行制备与酶解来避免上述情况的发生。

4.2 数据采集

内源性交联通常在酶的催化下形成,这类交联虽然也是共价交联,但是它们的性质却不完全相同。在交联肽段鉴定中,目前主要是依据肽谱匹配进行,即通过

对候选肽段产生的理论碎片离子与谱图中采集到的实际碎片离子比较。这一过程中, 肽段的理论碎片离子正确产生才能使匹配正确展开。肽段在进行碎裂时有多种方式可以采用, 不同的碎裂方式将使肽段产生不同的碎片离子。碎裂过程中破坏了肽段中的化学键, 由于部分内源性交联是通过化学键直接将两个氨基酸侧链进行连接, 交联部分在碎裂过程中是否被破坏需要实验来证明。通常将合成的标准肽段进行交联实验, 再将其利用质谱进行数据采集, 根据谱图中的碎片离子探究碎裂过程中碎片离子的规律。

在二硫键研究中发现, 使用高能碰撞解离(higher-energy collisional dissociation, HCD)优于电子转移解离(electron-transfer dissociation, ETD), 并且二硫键断裂后也能形成一些特殊形式的碎片离子^[20]。而Tayri-Wilk等人^[54]在研究甲醛交联3种纯化蛋白(牛血清白蛋白、卵转铁蛋白和 α -淀粉酶)组成的混合物过程中发现, 在HCD的碎裂模式下, 甲醛交联呈现对称碎裂, 即交联肽段会碎裂成两条单肽, 每条肽段会加上12 Da的质量, 甲醛交联肽段的谱图将会呈现为两条mono-link肽段的碎片离子分布。这与其他内源性交联碎裂规律完全不同。

4.3 定性分析

交联质谱的定性分析主要通过数据库搜索的方法进行, 即对数据库中的蛋白质序列进行理论酶切得到候选肽段, 并产生理论谱图与实际谱图进行肽谱匹配(peptide spectrum match, PSM)。交联质谱的定性分析主要有两个难点: 一是候选空间巨大; 二是交联肽段的碎片离子多样。对单肽而言, 每张谱图可直接与所有候选肽段进行肽谱匹配, 所面临的检索空间为 $O(n)$ 。对交联肽段而言, 每张谱图应与交联肽段的理论谱图进行肽谱匹配, 交联肽段是根据酶切的候选肽段进行两两组合得到的, 因此面临的检索空间为 $O(n^2)$ ^[93]。交联肽段的碎片离子相较于单肽的碎片离子, 其质量发生了较大的改变。另外, 交联肽段是由两条单肽组成, 其碎片离子数目也比单肽离子数目要多。因此在肽谱匹配打分时, 交联肽段的打分需要考虑更多的因素。对于上述两个问题, 在特殊情况下可以通过交联的特性来设计特殊的检索策略来解决。目前主流交联肽段鉴定软件虽然其具体的检索策略略有不同, 但是根据面向的样本特点可分为三大类(图4)。

第一大类鉴定策略是基于非标记的不可碎裂交联

设计, 即认为交联在整个质谱数据采集过程中不会发生改变。这类软件在鉴定时需要直面 $O(n^2)$ 问题, 因为无法判断谱图是否为交联肽段的谱图, 所以需要对所有谱图都与候选肽段进行匹配。这类交联在送入质谱进行数据采集时, 在采集二级谱图前需要对肽段进行碎裂, 但碎裂只会发生在肽键上, 因此采集到的二级谱图在理想条件下会存在两条肽段的所有碎片离子, 其中一部分碎片离子连着另一条肽段。对于这类鉴定需要在 $O(n^2)$ 的空间内检索, 目前也有几种不同的实现方法。如Kojak^[105]使用开放式搜索方法缩减候选肽段范围, 在小范围内进行枚举所有肽段进行匹配确定最优结果。pLink^[106,107]使用两步检索策略, 即使用粗打分确定其中一条候选肽段后, 再对所有能与它构成交联肽段的候选肽段使用细打分对候选肽段与谱图进行匹配, 确定最优结果。Xolik^[108]采用双端队列算法实现快速对所有候选交联肽段与谱图进行匹配并获取最优结果。

第二大类鉴定策略是基于同位素标记的不可碎裂交联设计。由于交联是将两条肽段连接起来, 因此可以考虑在交联上引入轻重同位素标记来缩小搜索范围。在同位素标记的作用下, 发生交联后一级谱图上会出现两个峰簇, 这两个峰簇会满足轻重同位素标记的交联剂的质量差, 进而可以判断含有交联肽段的谱图, 并且由于不同同位素标记的交联具有质量差, 因此二级谱图上的碎片离子也同样具有质量关系。基于这种方式而设计的鉴定软件有xQuest^[109]等。

第三大类鉴定策略是基于可碎裂交联设计。可以通过调整碎裂方式和碎裂能量, 使交联发生断裂。交联上若存在两处可断裂位点, 将能与两条肽段分别组合产生4种情况, 这样谱图上就会出现两组峰。每组峰包括两个峰, 这两个峰的差值是相同的, 等于交联剂两处断裂位点间的质量。因此在鉴定时只需要搜寻这两组峰来判断组成交联肽段的两条肽段的质量, 极大地缩小搜索范围, 再根据这一搜索范围来与肽段的碎片离子进行匹配获取最优结果, 这类软件有XlinkX^[110]等。

由于缺乏对具体内源性交联在质谱数据采集过程中碎裂特性的研究, 目前主要采用第一种检索策略的交联软件来鉴定内源性交联, 即认为内源性交联不会发生断裂(表2)。2015年, 针对研究较多的二硫键开发了pLink-SS^[20], 专门用于二硫键的质谱数据解析。这是目前唯一一款针对内源性交联中的二硫键而开发的鉴定软件。通过对二硫键交联肽段标准谱图中的碎片离子进行统计, 发现了多种二硫键交联肽段的特征碎片

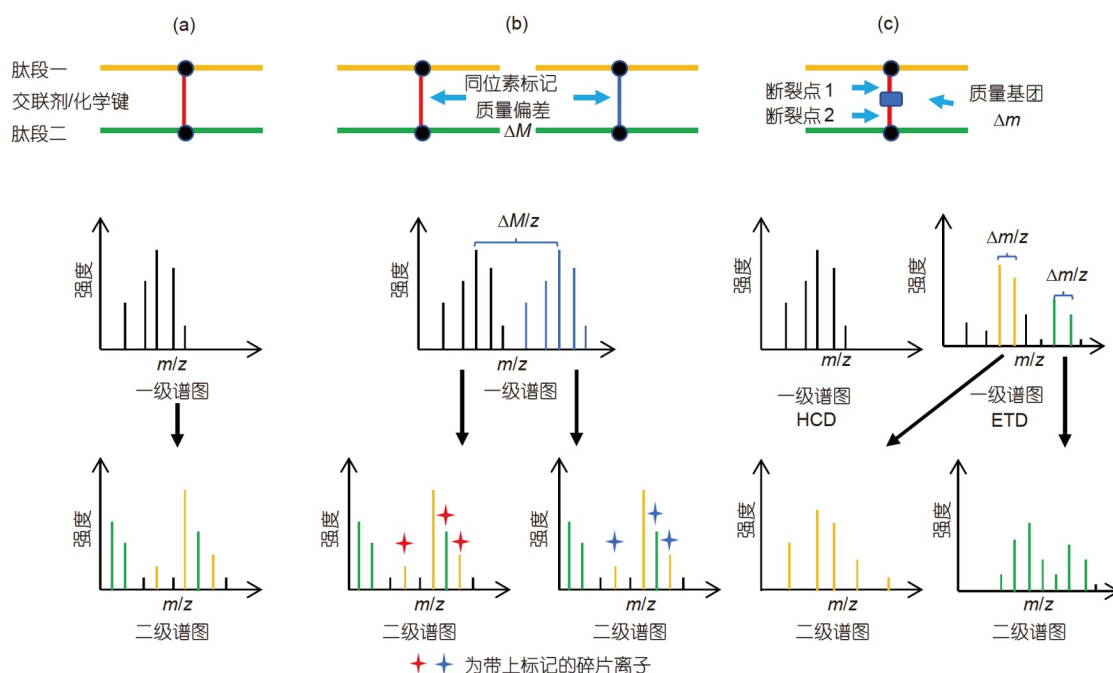


图 4 (网络版彩色) 交联肽段鉴定策略. (a) 基于非标记不可碎裂交联的鉴定策略. 采集一级谱图和二级谱图, 通过二级谱图与所有候选肽段(包括单肽和交联肽段)的碎片离子进行匹配. (b) 基于同位素标记不可碎裂交联的鉴定策略. 非标记和同位素标记的交联肽段在一级谱图上会成对出现, 根据对应的二级谱图与候选交联肽段的碎片离子进行匹配. (c) 基于可碎裂交联的鉴定策略. 交联肽段上的交联剂在采集一级谱图时发生断裂, 得到两根谱峰, 根据断裂点之间的质量基团差值来确定组成交联肽段的两条肽段质荷比, 再根据二级谱图与候选肽段的碎片离子进行匹配. **Figure 4** (Color online) Identification of crosslinking peptides is achieved by different strategies. (a) Identification of label-free and non-cleavable cross-linking peptides. MS1 and MS2 are acquired for all peptides from the sample including single peptides and cross-linking peptides. The MS2 spectra are matched to the fragment ions of all candidate single peptides and cross-linking peptides. (b) Identification of isotope-labeling and non-cleavable cross-linking peptides. The label-free and isotope-labeling crosslinking peptides have similar isotope peak cluster in the same MS1 spectrum. The MS2 spectra are matched to the fragment ions of all candidate cross-linking peptides. (c) Identification of cleavable crosslinking peptides. The cross-linker on the crosslinking peptide breaks in MS1 spectrum and obtains two spectral peaks group. The precursors of the two peptides that constitute the cross-linking peptide were calculated based on the mass difference between the breaking points. The MS2 spectra of two peptides are matched to the fragment ions of the candidate peptides

表 2 用于内源性交联研究的软件

Table 2 Software used for endogenous cross-linking studies

软件	获取方法	相关研究
pLink-SS ^[20]	http://pfnd.ict.ac.cn/software/pLink/index.html	二硫键
pLink ^[106,107]	http://pfnd.ict.ac.cn/software/pLink/index.html	异肽键
Protein Prospector ^[111]	http://prospector.ucsf.edu	色氨酸交联
StavroX ^[112]	Send email to michael.goetze@biochemtech.uni-halle.de	酪氨酸交联
Formaldehyde_XL_Analyser ^[54]	https://github.com/Kaliman-Lab/Search_Formaldehyde_Cross-links	甲醛交联
MassLynx ^[70]	https://www.waters.com/	丙酮醛交联
Integrated Proteomics Pipeline ^[113]	https://www.bruker.com/	丙酮醛交联
MaxQuant ^[114]	https://www.maxquant.org/	NOS桥

离子, 其中也包括一些二硫键发生断裂后的碎片离子. 将这些特征离子引入到pLink-SS的肽谱匹配打分中可以大幅提升二硫键交联肽段的鉴定数目, 研究者发现了很多以往未发现的二硫键. 由二硫键的研究可知, 想

要对内源性交联进行完整鉴定, 先探究其碎裂规律是十分必要的, 否则可能会导致很多内源性交联不能被正确鉴定出来, 对不同的内源性交联或许需要开发与之配套的鉴定软件才能作到更完整的鉴定^[41].

4.4 定量分析

内源性交联的定量分析主要是对交联肽段进行, 包括标记定量和非标记定量, 具体原理与单肽的定量分析流程类似. 在标记定量中, 通过报告离子的强度来反映内源性交联含量. 在非标记定量中, 依据内源性蛋白质交联的定性结果, 重构对应交联肽段的母离子色谱曲线, 色谱曲线所围成的峰面积反映内源性交联含量^[115]. 目前内源性交联的定量研究较少, 主要针对单个蛋白或者简单样本进行. 对内源性交联进行定量分析, 将有助于深入了解内源性交联在蛋白质功能发挥中的独特机制. 另外, 通过使用同位素标记相关蛋白质, 在体外进行内源性交联实验可一定程度确定内源性交联优先发生在蛋白质分子内还是蛋白质分子间. 由于同位素标记质量差异, 分子间和分子内交联所对应的母离子质量是不同的, 它们的色谱曲线峰面积比例可反映出分子间和分子内交联的含量情况^[49].

4.5 生物信息学分析

内源性交联的生物信息学分析依据交联肽段的定性和定量结果展开. 生物信息学分析能阐明内源性交联与蛋白质结构和功能可能存在的关系. 通常根据得到的鉴定结果, 先将交联位点映射到蛋白质序列上, 绘制出蛋白质交联网络图, 以此反映出内源性交联在蛋白质中的分布. 之后可对鉴定到的交联肽段位点进行 motif 分析, 发现形成交联的热点序列^[20]. 也可对实验组和对照组中的蛋白质交联定量结果进行差异分析, 具有显著性差异的蛋白质交联可能是导致蛋白质功能改变的潜在因素. 在生理病理分析过程中, 这些交联很有可能成为重要的生物标志物^[49]. 另外, 由于内源性交联的形成也受到空间距离的限制, 对距离不满足条件限制的交联肽段, 可能是蛋白质构象发生了改变的结果, 这类交联肽段结果将有助于蛋白质结构的确定^[96].

5 总结和讨论

内源性交联目前被认为是一种蛋白质的翻译后修饰, 它的产生对蛋白质的结构和功能发挥着重要作用. 二硫键在组学水平上的研究表明, 蛋白质中的二硫键可分为结构型和功能型两大类, 其中功能型二硫键涉及了催化和调节等功能. 对蛋白质中的其他内源性交联, 也需要进行相类似的分析工作. 这些内源性交联位点的发现将有助于对蛋白质的调节功能、细胞凋亡、

疾病机制等方面的研究. 经过十几年发展, 初步形成了一套基于质谱的内源性交联分析方法, 包括样品预处理、数据采集、定性分析、定量分析和生物信息学分析等环节.

然而, 内源性交联的鉴定目前还存在很多挑战. 首先, 内源性交联虽然在蛋白质中分布广泛, 但是其含量在蛋白质样品中并不高, 难以采集到交联肽段的谱图, 即使采集到交联肽段的谱图, 谱峰强度也偏低, 不易进行后续定性定量分析. 其次, 内源性交联是一个大类, 不同类型内源性交联的化学性质差异会导致上述环节的具体实现有很大不同. 对内源性交联进行完整的组学分析需要先进行很多的前期工作, 这包括对内源性交联的化学性质和碎裂规律的探究. 这些规律是定性定量分析的理论依据, 可从交联标准肽段的质谱数据中发掘. 但是, 很多内源性交联是在酶参与下形成, 而酶的功能发挥需要与受体蛋白结合才能实现. 这种情况下很难得到交联的标准肽段, 后续碎片离子规律的探究也将无法进行. 定性上, 要作到内源性交联的完整鉴定, 还需要考虑复杂交联肽段形式. 定量上, 在组学水平的研究中, 通常会进行定量和差异分析, 但是目前的研究几乎不涉及内源性交联的定量分析. 这是因为发生交联的蛋白质, 可分别在蛋白质和交联肽段的层次上进行定量, 不同层次上的定量结果如何比较目前还有待讨论. 另外, 对于内源性交联, 还要保证从样品制备到数据采集过程中不会发生重构, 否则将会导致错误的鉴定结果产生.

6 展望

内源性交联的种类、性质和功能值得深入研究. 质谱技术和鉴定方法的发展将推动内源性交联的研究向更深层次发展. 从早期只能支持单一纯化样品, 到略微复杂样品的小规模鉴定, 再到现在可以实现较大规模的复杂样品鉴定. 不可否认, 内源性交联的特点也使利用质谱技术进行分析还有很多困难亟待解决, 一套针对细胞内源性交联组学的技术体系的建立或许是解决这些问题的最佳方案.

在组学水平上对内源性交联的定性定量分析将是内源性交联研究的发展方向. 现阶段, 内源性交联富集方法的开发和碎裂规律的探究仍是重点, 它们的优化能有效地提高最终的鉴定效果. 对内源性交联肽的富集可结合分离色谱、免疫亲和纯化、离子阱选择性富集等方法来实现, 而碎裂规律的探究需要大量的交联

肽段标准谱图作为数据支撑。随着实验方法流程和鉴定算法的完善, 内源性交联的研究将变得简单有效并

得到广泛应用, 成为蛋白质组学研究中的一个重要组成部分, 并推动蛋白质组研究向更深层次发展。

参考文献

- Maes E, Dyer J M, McKerchar H J, et al. Protein-protein cross-linking and human health: The challenge of elucidating with mass spectrometry. *Expert Rev Proteomics*, 2017, 14: 917–929
- Zhu W G, Hou T Y, Lu X P. Importance of protein post-translational modifications in finding partners (in Chinese). *Chin Sci Bull*, 2017, 62: 759–769 [朱卫国, 侯天云, 陆小鹏. 蛋白质翻译后修饰在蛋白质-蛋白质相互作用中的调控作用. 科学通报, 2017, 62: 759–769]
- Gupta M N, Perwez M, Sardar M. Protein crosslinking: Uses in chemistry, biology and biotechnology. *Biocatal Biotransform*, 2020, 38: 178–201
- Zhang Y K, Song C X, Qin H Q, et al. The new method and technology for the characterization and identification of post-translational modifications of proteins (in Chinese). *Chin Sci Bull*, 2013, 58: 3007–3016 [张玉奎, 宋春侠, 秦洪强, 等. 修饰蛋白质组学分离鉴定新技术新方法. 科学通报, 2013, 58: 3007–3016]
- Heck T, Faccio G, Richter M, et al. Enzyme-catalyzed protein crosslinking. *Appl Microbiol Biotechnol*, 2013, 97: 461–475
- Sinz A. Cross-linking/mass spectrometry for studying protein structures and protein-protein interactions: Where are we now and where should we go from here? *Angew Chem Int Edit*, 2018, 57: 6390–6396
- Lee Y J. Mass spectrometric analysis of cross-linking sites for the structure of proteins and protein complexes. *Mol Biosyst*, 2008, 4: 816–823
- Politis A, Stengel F, Hall Z, et al. A mass spectrometry-based hybrid method for structural modeling of protein complexes. *Nat Methods*, 2014, 11: 403–406
- Tran B Q, Goodlett D R, Goo Y A. Advances in protein complex analysis by chemical cross-linking coupled with mass spectrometry (CXMS) and bioinformatics. *BBA-Proteins Proteomics*, 2016, 1864: 123–129
- Reddy N C, Kumar M, Molla R, et al. Chemical methods for modification of proteins. *Org Biomol Chem*, 2020, 18: 4669–4691
- Gaar J, Naffa R, Brimble M. Enzymatic and non-enzymatic crosslinks found in collagen and elastin and their chemical synthesis. *Org Chem Front*, 2020, 7: 2789–2814
- Ursem R, Swarge B, Abhyankar W R, et al. Identification of native cross-links in *Bacillus subtilis* spore coat proteins. *J Proteome Res*, 2021, 20: 1809–1816
- Bechtel T J, Weerapana E. From structure to redox: The diverse functional roles of disulfides and implications in disease. *Proteomics*, 2017, 17: 1600391
- Jiang S, Häggglund P, Carroll L, et al. Crosslinking of human plasma C-reactive protein to human serum albumin via disulfide bond oxidation. *Redox Biol*, 2021, 41: 101925
- Woycechowsky K J, Raines R T. Native disulfide bond formation in proteins. *Curr Opin Chem Biol*, 2000, 4: 533–539
- Cramer C N, Kelstrup C D, Olsen J V, et al. Generic workflow for mapping of complex disulfide bonds using in-source reduction and extracted ion chromatograms from data-dependent mass spectrometry. *Anal Chem*, 2018, 90: 8202–8210
- Lakub J C, Shipman J T, Desaire H. Recent mass spectrometry-based techniques and considerations for disulfide bond characterization in proteins. *Anal Bioanal Chem*, 2018, 410: 2467–2484
- Butera D, Cook K M, Chiu J, et al. Control of blood proteins by functional disulfide bonds. *Blood*, 2014, 123: 2000–2007
- Farrah T, Deutsch E W, Omenn G S, et al. A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Mol Cell Proteomics*, 2011, 10: M110.006353
- Lu S, Fan S B, Yang B, et al. Mapping native disulfide bonds at a proteome scale. *Nat Methods*, 2015, 12: 329–331
- Liu C, Shen L, Xiao Y, et al. Pollen PCP-B peptides unlock a stigma peptide-receptor kinase gating mechanism for pollination. *Science*, 2021, 372: 171–175
- Hogg P J. Targeting allosteric disulphide bonds in cancer. *Nat Rev Cancer*, 2013, 13: 425–431
- Suzuki M, Yamanoi A, Machino Y, et al. Effect of trastuzumab interchain disulfide bond cleavage on Fcγ receptor binding and antibody-dependent tumour cell phagocytosis. *J Biochem*, 2016, 159: 67–76
- Suzuki M, Yamanoi A, Machino Y, et al. Cleavage of the interchain disulfide bonds in rituximab increases its affinity for FcγRIIIA. *Biochem Biophys Res Commun*, 2013, 436: 519–524
- Pickart C M. Mechanisms underlying ubiquitination. *Annu Rev Biochem*, 2001, 70: 503–533
- Hay R T. SUMO: A history of modification. *Mol Cell*, 2005, 18: 1–12
- Kang H J, Baker E N. Intramolecular isopeptide bonds: Protein crosslinks built for stress? *Trends Biochem Sci*, 2011, 36: 229–237
- Mandlik A, Swierczynski A, Das A, et al. Pili in Gram-positive bacteria: Assembly, involvement in colonization and biofilm development. *Trends*

- [Microbiol](#), 2008, 16: 33–40
- 29 Ton-That H, Schneewind O. Assembly of pili on the surface of *Corynebacterium diphtheriae*. [Mol Microbiol](#), 2003, 50: 1429–1438
 - 30 Kang H J, Coulibaly F, Clow F, et al. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. [Science](#), 2007, 318: 1625–1628
 - 31 Griffin M, Casadio R, Bergamini C M. Transglutaminases: Nature’s biological glues. [Biochem J](#), 2002, 368: 377–396
 - 32 Grenard P, Bates M K, Aeschlimann D. Evolution of transglutaminase genes: Identification of a transglutaminase gene cluster on human chromosome 15q15. [J Biol Chem](#), 2001, 276: 33066–33078
 - 33 Sivadó É, El Alaoui M, Kiraly R, et al. Optimised methods (SDS/PAGE and LC-MS) reveal deamidation in all examined transglutaminase-mediated reactions. [FEBS Open Bio](#), 2019, 9: 396–404
 - 34 D’Eletto M, Farrace M G, Piacentini M, et al. Assessing the catalytic activity of transglutaminases in the context of autophagic responses. [Methods Enzymol](#), 2017, 587: 511–520
 - 35 Lerner A, Matthias T. Microbial transglutaminase should be considered as an environmental inducer of celiac disease. [World J Clin Cases](#), 2019, 7: 3912–3914
 - 36 Zhang J, Ding Y R, Wang R. Inhibition of tissue transglutaminase promotes A β -induced apoptosis in SH-SY5Y cells. [Acta Pharmacol Sin](#), 2016, 37: 1534–1542
 - 37 Noor A, Zafar S, Zerr I. Neurodegenerative proteinopathies in the proteoform spectrum—Tools and challenges. [Int J Mol Sci](#), 2021, 22: 1085
 - 38 Schmitt L R, Henderson R, Barrett A, et al. Mass spectrometry-based molecular mapping of native FXIIIa cross-links in insoluble fibrin clots. [J Biol Chem](#), 2019, 294: 8773–8778
 - 39 Paviani V, Galdino G T, dos Prazeres J N, et al. Dityryptophan cross-links as novel products of protein oxidation. [J Braz Chem Soc](#), 2018, 29: 925–933
 - 40 Sherin P S, Zelentsova E A, Sormacheva E D, et al. Aggregation of α -crystallins in kynurenic acid-sensitized UVA photolysis under anaerobic conditions. [Phys Chem Chem Phys](#), 2016, 18: 8827–8839
 - 41 Paviani V, Queiroz R F, Marques E F, et al. Production of lysozyme and lysozyme-superoxide dismutase dimers bound by a dityryptophan cross-link in carbonate radical-treated lysozyme. [Free Radic Biol Med](#), 2015, 89: 72–82
 - 42 Xu J, Eriksson S E, Cebula M, et al. The conserved Trp114 residue of thioredoxin reductase 1 has a redox sensor-like function triggering oligomerization and crosslinking upon oxidative stress related to cell death. [Cell Death Dis](#), 2015, 6: e1616
 - 43 Giulivi C, Traaseth N J, Davies K J A. Tyrosine oxidation products: Analysis and biological relevance. [Amino Acids](#), 2003, 25: 227–232
 - 44 Al-Hilaly Y K, Williams T L, Stewart-Parker M, et al. A central role for dityrosine crosslinking of Amyloid- β in Alzheimer’s disease. [Acta Neuropathol Commun](#), 2013, 1: 83
 - 45 Pennathur S, Jackson-Lewis V, Przedborski S, et al. Mass spectrometric quantification of 3-nitrotyrosine, ortho-tyrosine, and o,o’-dityrosine in brain tissue of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-treated mice, a model of oxidative stress in Parkinson’s disease. [J Biol Chem](#), 1999, 274: 34621–34628
 - 46 Kato Y, Maruyama W, Naoi M, et al. Immunohistochemical detection of dityrosine in lipofuscin pigments in the aged human brain. [FEBS Lett](#), 1998, 439: 231–234
 - 47 Abdelrahim M, Morris E, Carver J, et al. Liquid chromatographic assay of dityrosine in human cerebrospinal fluid. [J Chromatogr B](#), 1997, 696: 175–182
 - 48 Leeuwenburgh C, Rasmussen J E, Hsu F F, et al. Mass spectrometric quantification of markers for protein oxidation by tyrosyl radical, copper, and hydroxyl radical in low density lipoprotein isolated from human atherosclerotic plaques. [J Biol Chem](#), 1997, 272: 3520–3526
 - 49 Mukherjee S, Kapp E A, Lothian A, et al. Characterization and identification of dityrosine cross-linked peptides using tandem mass spectrometry. [Anal Chem](#), 2017, 89: 6136–6145
 - 50 Ren X, Zou Q, Yuan C, et al. The dominant role of oxygen in modulating the chemical evolution pathways of tyrosine in peptides: Dityrosine or melanin. [Angew Chem Int Edit](#), 2019, 58: 5872–5876
 - 51 Kumar S, Molina-Cruz A, Gupta L, et al. A peroxidase/dual oxidase system modulates midgut epithelial immunity in *Anopheles gambiae*. [Science](#), 2010, 327: 1644–1648
 - 52 Burgos-Barragan G, Wit N, Meiser J, et al. Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. [Nature](#), 2017, 548: 549–554
 - 53 Zhu R, Zhang G, Jing M, et al. Genetically encoded formaldehyde sensors inspired by a protein intra-helical crosslinking reaction. [Nat Commun](#), 2021, 12: 581
 - 54 Tayri-Wilk T, Slavin M, Zamel J, et al. Mass spectrometry reveals the chemistry of formaldehyde cross-linking in structured proteins. [Nat Commun](#), 2020, 11: 3128
 - 55 Nemet I, Varga-Defterdarović L, Turk Z. Methylglyoxal in food and living organisms. [Mol Nutr Food Res](#), 2006, 50: 1105–1117
 - 56 Westwood M E, Thornalley P J. Molecular characteristics of methylglyoxal-modified bovine and human serum albumins. Comparison with glucose-derived advanced glycation endproduct-modified serum albumins. [J Protein Chem](#), 1995, 14: 359–372

- 57 Ahmed N, Argirov O K, Minhas H S, et al. Assay of advanced glycation endproducts (AGEs): Surveying AGEs by chromatographic assay with derivatization by 6-aminoquinolyl-*N*-hydroxysuccinimidyl-carbamate and application to *N*^ε-carboxymethyl-lysine- and *N*^ε-(1-carboxyethyl)lysine-modified albumin. *Biochem J*, 2002, 364: 1–14
- 58 Ahmed N, Thornalley P J. Chromatographic assay of glycation adducts in human serum albumin glycated *in vitro* by derivatization with 6-aminoquinolyl-*N*-hydroxysuccinimidyl-carbamate and intrinsic fluorescence. *Biochem J*, 2002, 364: 15–24
- 59 Shipanova I N, Glomb M A, Nagaraj R H. Protein modification by methylglyoxal: Chemical nature and synthetic mechanism of a major fluorescent adduct. *Arch Biochem Biophys*, 1997, 344: 29–36
- 60 Oya T, Hattori N, Mizuno Y, et al. Methylglyoxal modification of protein. *J Biol Chem*, 1999, 274: 18492–18502
- 61 Ahmed M U, Brinkmann Frye E, Degenhardt T P, et al. *N*^ε-(carboxyethyl)lysine, a product of the chemical modification of proteins by methylglyoxal, increases with age in human lens proteins. *Biochem J*, 1997, 324: 565–570
- 62 Brinkmann E, Wells-Knecht K J, Thorpe S R, et al. Characterization of an imidazolium compound formed by reaction of methylglyoxal and *N*^ε-hippuryllysine. *J Chem Soc-Perkin Trans 1*, 1995, (22): 2817–2818
- 63 Lederer M O, Klaiber R G. Cross-linking of proteins by maillard processes: Characterization and detection of lysine-arginine cross-links derived from glyoxal and methylglyoxal. *Bioorg Med Chem*, 1999, 7: 2499–2507
- 64 Odani H, Shinzato T, Usami J, et al. Imidazolium crosslinks derived from reaction of lysine with glyoxal and methylglyoxal are increased in serum proteins of uremic patients: Evidence for increased oxidative stress in uremia. *FEBS Lett*, 1998, 427: 381–385
- 65 Wilker S C, Chellan P, Arnold B M, et al. Chromatographic quantification of argpyrimidine, a methylglyoxal-derived product in tissue proteins: Comparison with pentosidine. *Anal Biochem*, 2001, 290: 353–358
- 66 Nagai R, Araki T, Hayashi C M, et al. Identification of *N*^ε-(carboxyethyl)lysine, one of the methylglyoxal-derived AGE structures, in glucose-modified protein: Mechanism for protein modification by reactive aldehydes. *J Chromatogr B*, 2003, 788: 75–84
- 67 Bourajjaj M, Stehouwer C D A, van Hinsbergh V W M, et al. Role of methylglyoxal adducts in the development of vascular complications in diabetes mellitus. *Biochem Soc Trans*, 2003, 31: 1400–1402
- 68 Fosmark D S, Torjesen P A, Kilhovd B K, et al. Increased serum levels of the specific advanced glycation end product methylglyoxal-derived hydroimidazolone are associated with retinopathy in patients with type 2 diabetes mellitus. *Metabolism*, 2006, 55: 232–236
- 69 Maynard S E, Min J Y, Merchan J, et al. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J Clin Invest*, 2003, 111: 649–658
- 70 Shuster S O, Fica-Contreras S M, Hedges J S, et al. Comparison of the reaction of methylglyoxal (MGO) with murine and human amyloid beta (Aβ): Insights into a mechanism of Alzheimer's disease (AD). *Biochem Biophys Res Commun*, 2020, 533: 1298–1302
- 71 Bollong M J, Lee G, Coukos J S, et al. A metabolite-derived protein modification integrates glycolysis with KEAP1-NRF2 signalling. *Nature*, 2018, 562: 600–604
- 72 Wensien M, von Pappenheim F R, Funk L M, et al. A lysine-cysteine redox switch with an NOS bridge regulates enzyme function. *Nature*, 2021, 593: 460–464
- 73 Matsusaki M, Kanemura S, Kinoshita M, et al. The protein disulfide isomerase family: From proteostasis to pathogenesis. *Biochim Biophys Acta-Gen Subj*, 2020, 1864: 129338
- 74 Lorand L, Graham R M. Transglutaminases: Crosslinking enzymes with pleiotropic functions. *Nat Rev Mol Cell Biol*, 2003, 4: 140–156
- 75 Serafini-Fracassini D, Del Duca S. Transglutaminases: Widespread cross-linking enzymes in plants. *Ann Bot*, 2008, 102: 145–152
- 76 Yokoyama K, Nio N, Kikuchi Y. Properties and applications of microbial transglutaminase. *Appl Microbiol Biotechnol*, 2004, 64: 447–454
- 77 Popp M W L, Ploegh H L. Making and breaking peptide bonds: Protein engineering using sortase. *Angew Chem Int Edit*, 2011, 50: 5024–5032
- 78 Spirig T, Weiner E M, Clubb R T. Sortase enzymes in Gram-positive bacteria. *Mol Microbiol*, 2011, 82: 1044–1059
- 79 Gupta R, Beg Q K, Lorenz P. Bacterial alkaline proteases: Molecular approaches and industrial applications. *Appl Microbiol Biotechnol*, 2002, 59: 15–32
- 80 Claus H, Decker H. Bacterial tyrosinases. *Syst Appl Microbiol*, 2006, 29: 3–14
- 81 Faccio G, Kruus K, Saloheimo M, et al. Bacterial tyrosinases and their applications. *Process Biochem*, 2012, 47: 1749–1760
- 82 Fairhead M, Thöny-Meyer L. Bacterial tyrosinases: Old enzymes with new relevance to biotechnology. *New Biotechnol*, 2012, 29: 183–191
- 83 Halaouli S, Asther M, Sigoillot J C, et al. Fungal tyrosinases: New prospects in molecular characteristics, bioengineering and biotechnological applications. *J Appl Microbiol*, 2006, 100: 219–232
- 84 Mayer A. Laccase: New functions for an old enzyme. *Phytochemistry*, 2002, 60: 551–565
- 85 Quintanar L, Stoj C, Taylor A B, et al. Shall we dance? How a multicopper oxidase chooses its electron transfer partner. *Acc Chem Res*, 2007, 40: 445–452
- 86 Santhanam N, Vivanco J M, Decker S R, et al. Expression of industrially relevant laccases: Prokaryotic style. *Trends Biotechnol*, 2011, 29: 480–489
- 87 Witayakran S, Ragauskas A. Synthetic applications of laccase in green chemistry. *Adv Synth Catal*, 2009, 351: 1187–1209

- 88 Koua D, Cerutti L, Falquet L, et al. PeroxiBase: A database with new tools for peroxidase family classification. *Nucleic Acids Res*, 2009, 37: D261–D266
- 89 Veitch N C. Horseradish peroxidase: A modern view of a classic enzyme. *Phytochemistry*, 2004, 65: 249–259
- 90 Lucero H A, Kagan H M. Lysyl oxidase: An oxidative enzyme and effector of cell function. *Cell Mol Life Sci*, 2006, 63: 2304–2316
- 91 Schmelzer C E H, Heinz A, Troilo H, et al. Lysyl oxidase-like 2 (LOXL2)-mediated cross-linking of tropoelastin. *FASEB J*, 2019, 33: 5468–5481
- 92 Gorman J J, Wallis T P, Pitt J J. Protein disulfide bond determination by mass spectrometry. *Mass Spectrom Rev*, 2002, 21: 183–216
- 93 Fan S B, Wu Y J, Yang B, et al. A new approach to protein structure and interaction research: Chemical cross-linking in combination with mass spectrometry (in Chinese). *Progr Biochem Biophys*, 2014, 41: 1109–1125 [樊盛博, 吴妍洁, 杨兵, 等. 蛋白质结构与相互作用研究新方法——交联质谱技术. *生物化学与生物物理进展*, 2014, 41: 1109–1125]
- 94 Postic G, Marcoux J, Reys V, et al. Probing protein interaction networks by combining MS-based proteomics and structural data integration. *J Proteome Res*, 2020, 19: 2807–2820
- 95 Zhang J, Liu C L, Li X, et al. Research progress in cross-linking mass spectrometry (in Chinese). *Chem Bioeng*, 2020, 37: 1–11 [张洁, 刘春丽, 李欣, 等. 交联质谱技术研究进展. *化学与生物工程*, 2020, 37: 1–11]
- 96 O'Reilly F J, Rappsilber J. Cross-linking mass spectrometry: Methods and applications in structural, molecular and systems biology. *Nat Struct Mol Biol*, 2018, 25: 1000–1008
- 97 Yu C, Huang L. Cross-linking mass spectrometry: An emerging technology for interactomics and structural biology. *Anal Chem*, 2018, 90: 144–165
- 98 Baugh E H, Lyskov S, Weitzner B D, et al. Real-time PyMOL visualization for Rosetta and PyRosetta. *PLoS One*, 2011, 6: e21931
- 99 Zhao L, Zhao Q, Shan Y, et al. Smart cutter: An efficient strategy for increasing the coverage of chemical cross-linking analysis. *Anal Chem*, 2020, 92: 1097–1105
- 100 Zhang N, An J, Qin H, et al. A mass-spectrometry-based antibody-free approach enables the quantification of D-dimer in plasma. *J Proteome Res*, 2020, 19: 3143–3152
- 101 Chen T C, Ibrahim Y M, Webb I K, et al. Mobility-selected ion trapping and enrichment using structures for lossless ion manipulations. *Anal Chem*, 2016, 88: 1728–1733
- 102 Monahan F J, German J B, Kinsella J E. Effect of pH and temperature on protein unfolding and thiol/disulfide interchange reactions during heat-induced gelation of whey proteins. *J Agric Food Chem*, 1995, 43: 46–52
- 103 Kerr J, Schlosser J L, Griffin D R, et al. Steric effects in peptide and protein exchange with activated disulfides. *Biomacromolecules*, 2013, 14: 2822–2829
- 104 Wu S L, Hühmer A F R, Hao Z, et al. On-line LC-MS approach combining collision-induced dissociation (CID), electron-transfer dissociation (ETD), and CID of an isolated charge-reduced species for the trace-level characterization of proteins with post-translational modifications. *J Proteome Res*, 2007, 6: 4230–4244
- 105 Hoopmann M R, Zelter A, Johnson R S, et al. Kojak: Efficient analysis of chemically cross-linked protein complexes. *J Proteome Res*, 2015, 14: 2190–2198
- 106 Yang B, Wu Y J, Zhu M, et al. Identification of cross-linked peptides from complex samples. *Nat Methods*, 2012, 9: 904–906
- 107 Chen Z L, Meng J M, Cao Y, et al. A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat Commun*, 2019, 10: 3404
- 108 Dai J, Jiang W, Yu F, et al. Xolik: Finding cross-linked peptides with maximum paired scores in linear time. *Bioinformatics*, 2019, 35: 251–257
- 109 Leitner A, Walzthoeni T, Aebersold R. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc*, 2014, 9: 120–137
- 110 Liu F, Rijkers D T S, Post H, et al. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat Methods*, 2015, 12: 1179–1184
- 111 Chu F, Baker P R, Burlingame A L, et al. Finding chimeras: A bioinformatics strategy for identification of cross-linked peptides. *Mol Cell Proteomics*, 2010, 9: 25–31
- 112 Götze M, Pettelkau J, Schaks S, et al. StavroX—A software for analyzing crosslinked products in protein interaction studies. *J Am Soc Mass Spectrom*, 2012, 23: 76–87
- 113 Carvalho P C, Lima D B, Leprevost F V, et al. Integrated analysis of shotgun proteomic data with PatternLab for proteomics 4.0. *Nat Protoc*, 2016, 11: 102–117
- 114 Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*, 2016, 11: 2301–2319
- 115 Liu C, Song C Q, Yuan Z F, et al. pQuant improves quantitation by keeping out interfering signals and evaluating the accuracy of calculated ratios. *Anal Chem*, 2014, 86: 5286–5294

Summary for “内源性蛋白质交联的生物质谱分析进展”

Mass spectrometry-based endogenous protein crosslinking omics

Qiushi Wei¹, Pengyun Gong², Nuomin Li³, Chao Liu^{2*}, Yongqian Zhang^{1*} & Yulin Deng¹

¹ School of Life Science, Beijing Institute of Technology, Beijing 100081, China;

² Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing 100191, China;

³ Institute of Engineering Medicine, Beijing Institute of Technology, Beijing 100081, China

* Corresponding authors, E-mail: liuchaobuaa@buaa.edu.cn; zyq@bit.edu.cn

Endogenous protein crosslinking usually means the connection of two amino acids in one or two proteins by covalent bonds, which includes intermolecular and intramolecular crosslinks. As a special posttranslational modification of proteins, endogenous protein crosslinking affects the structure and function of the proteins. The main endogenous protein crosslinks currently discovered in prokaryotes and eukaryotes are disulfide bonds, isopeptide bonds, ditryptophan, dityrosine, formaldehyde, methylglyoxal, NOS bridges and so on. All these crosslinks have been shown to have important effects on biological processes and molecular functions. The formation of endogenous protein crosslinking comprises the enzyme-catalyzed and non-enzyme-catalyzed. It is reported that several enzymes are utilized to crosslink the proteins in the literature, such as protein disulfide isomerase, transglutaminase, sortase A, tyrosinase, laccase and peroxidase. Some enzymes are already applied for protein crosslinking in the field of the food industry. However, the endogenous substrates of these enzymes are not fully discovered *in vivo* so far. In addition, the mechanism of protein crosslinks and the function of crosslink protein require further exploration. Mass spectrometry is a high-throughput and high-sensitivity tool that is reliable for the large-scale identification of endogenous protein crosslinking. In recent years, there have been an ever-increasing number of studies on endogenous protein crosslinking by mass spectrometry at the omics level, and mass spectrometry has become one important tool to study the endogenous protein crosslinking. For example, it is reported that for the first time, 199 disulfide bonds are identified in *Escherichia coli* and 568 disulfide bonds are identified in secreted proteins of endothelial cells. The results show that disulfide bonds can be divided into two categories: structural and allosteric disulfide bonds. Allosteric disulfide bonds are involved in catalysis and regulation in the cell. Besides the disulfide bond, other types of endogenous protein crosslinking also need to be identified and quantified, which will better understand their functions. With the development of mass spectrometry technology and identification methods, the study of endogenous protein crosslinking will be enhanced to a deeper level. From the early stage, the crosslinking can be identified from single purified protein, followed by the standard protein mixtures, and now, large-scale identification can be achieved in complex samples, such as in the prokaryotic and eukaryotic cells. However, there are still many challenges in the process of sample preparation, data acquisition, qualitative and quantitative analysis and bioinformatics software. It is urgent to establish a unique workflow to study the endogenous protein crosslinking technology, which will greatly promote the development of crosslinking omics. In the present paper, we summarize the important achievements in the recent studies of endogenous protein crosslinking, sort out problems on endogenous protein crosslinking analysis based on mass spectrometry, and finally put forward the prospect of developing endogenous protein crosslinking at the omics level.

endogenous crosslinking omics, mass spectrometry, intra-protein crosslinking, inter-protein crosslinking

doi: [10.1360/TB-2021-0700](https://doi.org/10.1360/TB-2021-0700)