

SUMO化蛋白质组的富集策略研究进展

李洋^{1,2,3}, 单亦初¹, 梁振^{1*}, 张丽华¹, 张玉奎¹

(1. 中国科学院分离分析化学重点实验室 中国科学院大连化学物理研究所, 辽宁 大连 116023;
2. 中国科学院大学, 北京 100049; 3. 宁波大学 新药技术研究院, 浙江 宁波 315211)

摘要: 蛋白质的SUMO(Small ubiquitin-like modifier)化修饰是生物体内一类重要的翻译后修饰, 与核转运、转录调控、基因完整性维持和细胞周期调控等重要生物学过程密切相关。然而由于SUMO本身天然丰度低以及氨基酸序列冗长, SUMO化修饰的分析鉴定一直是研究的热点和难点。为实现SUMO化蛋白质组的深度覆盖分析, 发展高效的SUMO化蛋白质/肽段的富集方法十分必要。该文对SUMO化蛋白质组不同富集方法的原理、特点以及最新研究进展进行了综述, 并对其发展前景进行了展望。

关键词: SUMO化修饰; 深度覆盖; 富集; 检测

中图分类号: O629.7; Q493.2 **文献标识码:** A **文章编号:** 1004-4957(2022)01-0058-05

Research Advances of Enrichment Approaches for SUMOylated Proteome

LI Yang^{1,2,3}, SHAN Yi-chu¹, LIANG Zhen^{1*}, ZHANG Li-hua¹, ZHANG Yu-kui¹

(1. Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China; 3. Institute of Drug Discovery Technology, Ningbo University, Ningbo 315211, China)

Abstract: Protein SUMOylation is one of the most important post-translational modifications, which involved in various biological processes, including nuclear transport, transcriptional regulation, maintenance of gene integrity and cell cycle. However, identification of SUMOylation has always been the hotspot and difficulty in this field due to the low abundance and lengthy peptide sequence of endogenous SUMOylation. Therefore, the efficient enrichment of SUMOylated proteins/peptides is necessary to achieve deep coverage analysis of SUMOylome. The basic principles, characteristics and latest progresses of different enrichment methods for SUMOylated proteome are reviewed, and the development prospects are prospected.

Key words: SUMOylation; deep coverage; enrichment; detection

小泛素相关修饰物(Small ubiquitin-like modifier, SUMO)是一类分子量约为11 kDa, 高度保守的类泛素蛋白^[1-3]。在哺乳动物中, SUMO存在5种旁系同源物^[4-6], 研究较多的是SUMO1、SUMO2和SUMO3。其中, SUMO2与SUMO3的氨基酸序列有97%的相似性, 而与SUMO1只有50%的相似性, 因此通常将其合并为SUMO2/3来进行研究^[7]。SUMO与底物蛋白的结合依赖于E1活化酶、E2结合酶和E3连接酶的多步催化反应, 而生物体内多种SUMO特异性蛋白酶(SUMO/sentrin-specific proteases, SENPs)的存在使SUMO化修饰处于高度动态变化的可逆过程^[8-10]。许多被SUMO修饰的底物蛋白都具有一致性序列(Consensus motif) $\Psi K_x D/E$ (Ψ 代表疏水性氨基酸, K代表赖氨酸, D/E代表天冬氨酸/谷氨酸, x代表任何氨基酸)^[11], 还有些修饰位点遵循反向的motif(E/D-x-K)^[12]。蛋白质的SUMO化修饰影响了诸多生理和病理过程, 包括转录调控、DNA损伤修复、自身免疫性疾病以及肿瘤的发生发展等^[13-16], 因此, 参与SUMO化过程的相关底物和蛋白酶是临床治疗疾病和癌症的重要药物靶标^[17-19]。

为了更全面地揭示SUMO化修饰发挥的生理和病理功能, 在蛋白质组水平对SUMO化蛋白及其位点进行深度覆盖分析至关重要^[20]。然而, 规模化分析哺乳动物细胞中的SUMO化位点仍存在巨大挑战。一方面是由于SUMO化蛋白质在生物体内的存在丰度较低, 且高活性的去SUMO化酶使SUMO化修饰

收稿日期: 2021-09-24; 修回日期: 2021-10-20

基金项目: 国家重点研发计划项目(2017YFA0505003)

* 通讯作者: 梁振, 博士, 研究员, 研究方向: 蛋白质组定性定量及相互作用新方法研究, E-mail: liangzhen@dicp.ac.cn

极不稳定;另一方面是在采用自下而上(Bottom-up)的蛋白质组策略时, SUMO1和SUMO2/3的C端由于缺少精氨酸和赖氨酸,当被胰蛋白酶酶切后,残留在底物肽段上的SUMO1和SUMO2/3修饰分别含有19和32个氨基酸,如此冗长的序列会导致其修饰位点难以被质谱检测分析,从而影响SUMO化蛋白质组的分析覆盖度。因此,发展高效的富集方法对SUMO化蛋白质组的深度覆盖分析至关重要。本文针对外源性和内源性的SUMO化蛋白质组富集方法的原理、特点及最新进展进行了综述。

1 外源性SUMO化蛋白质组的富集方法

为了提高对低丰度SUMO化蛋白质的富集效率,最有效的方法是发展针对外源性SUMO化蛋白质组的富集方法,即构建可以稳定表达带亲和和富集标签的外源性SUMO蛋白的细胞系。常用的富集标签有组氨酸(His)^[12, 21-24]、血球凝集素(HA)^[25]和Flag标签^[26]。其中,His标签因为分子量小,在SUMO上表达几乎不影响SUMO与底物蛋白的结合特异性,且在变性环境下,携带His标签的SUMO化蛋白仍能被Ni²⁺亲和材料高效捕获,因而被广泛使用。

1.1 蛋白质水平的富集

Galisson等^[21]利用Ni²⁺亲和色谱富集HEK293细胞中表达的外源性His₆-SUMO1/2/3修饰的底物蛋白。同时,为了使胰蛋白酶酶解后SUMO化肽段的长度更利于质谱检测,分别对His₆-SUMO1^{Q92R}、His₆-SUMO2^{Q88R}和His₆-SUMO3^{Q87R/Q88N}的特定位点进行突变,使酶解后残留在底物肽段上的氨基酸分别减少为EQTGG-、QQTGG-、NQTGG-,再结合无标定量蛋白质组分析技术,在经As₂O₃处理的细胞中共鉴定到17个精确的SUMO化位点,对应于12个SUMO修饰的底物蛋白质。

为了避免对氨基酸进行突变,Hendriks等^[27]发展了一种基于去SUMO化酶的外源性SUMO化位点鉴定方法(简称PRISM)。该方法在封闭其余非SUMO化赖氨酸侧链氨基的前提下,利用去SUMO化酶(SENP2)选择性切除底物蛋白质上的SUMO修饰,使去SUMO化的赖氨酸位点暴露特征性的氨基基团,实现对SUMO化位点的准确鉴定。作者首先在蛋白质水平利用Ni²⁺亲和色谱对His₁₀-SUMO2修饰的蛋白质进行纯化,再对富集所得蛋白质进行乙酰化封闭、去SUMO化、胰蛋白酶酶解和质谱鉴定,最终从HeLa细胞中鉴定到751个SUMO化位点。该方法避免了冗长的SUMO修饰给质谱鉴定造成的困难,但被乙酰化封闭后的赖氨酸无法被胰蛋白酶酶切将导致肽段最终长度较长,并且需要严格控制蛋白质水平的乙酰化反应不完全给SUMO位点鉴定带来的假阳性。

1.2 肽段水平的富集

在蛋白质水平进行富集,当蛋白质被酶解成肽段后会产生大量非SUMO化肽段,对SUMO化肽段的分析造成干扰,从而影响SUMO化位点的鉴定覆盖度。因此,研究人员进一步开发了可以特异性识别酶切后留在底物肽段上的SUMO残基序列的抗体,以实现SUMO化位点在肽段水平的二次富集。例如,Impensa等^[28]将SUMO1和SUMO2的基因分别改造为His₆-SUMO1^{T95R}和His₆-SUMO2^{T91R},先在蛋白水平富集His₆-SUMO修饰底物,再使胰蛋白酶酶切后的SUMO位点产生双甘氨酸残基GG-,接着利用泛素化抗体(antiK-ε-GG)对其进行特异性富集,最终鉴定到295个SUMO1位点和167个SUMO2位点。但是胰蛋白酶酶切后的泛素化位点也有GG-修饰,给鉴定结果带来了假阳性。Lamoliatte等^[22]开发了一种可以特异性识别His₆-SUMO3^{Q87R/Q88N}被胰蛋白酶酶切后留在底物肽段上的氨基酸残基(NQTGG-)的抗体,通过蛋白和肽段水平的两次富集在HEK293细胞中鉴定到了954个SUMO3位点,有效避免了泛素化位点带来的假阳性问题。随后作者在此基础上进一步优化了免疫沉淀的孵育和洗脱条件,采用更灵敏的高分辨率质谱并优化了质谱参数,结合SCX/RP二维色谱分级策略,从HEK293细胞中鉴定到一万个SUMO3位点^[29]。类似的方法还有将SUMO2基因改造为His₆-SUMO2^{T90K},分别利用Ni-NTA和K-ε-GG抗体对Lys-C或Lys-C+Glu-C酶切前后的SUMO修饰底物进行两次富集,鉴定到1000个SUMO位点^[24],该方法用Glu-C取代胰蛋白酶,避免了胰蛋白酶酶切后泛素化位点造成的假阳性。

除了上述方法,Vertegaal课题组还开发了基于K0方法的His-SUMO2位点富集方法,主要原理是将SUMO分子上的所有赖氨酸(K)突变为精氨酸(R),由此经过Lys-C酶切后的SUMO分子依然具有完整序列且携带His标签,利用Ni柱分别在蛋白和肽段水平对His-SUMO2修饰底物进行两次富集,即

可提高对SUMO位点的富集效率。该课题组最初将SUMO2基因改造为His₆-SUMO2^{K0T90R}，利用Ni柱对Lys-C酶切前后的SUMO化蛋白/肽段进行二次富集，再用胰蛋白酶将SUMO从底物肽段上切除，只留下GG-残基，质谱分析后共鉴定到103个SUMO2位点^[12]。然而，该方法同样存在泛素化位点造成的假阳性问题。因此，该课题组进一步将SUMO2基因改造为His₁₀-SUMO2^{K0Q87R}，同样利用Ni柱对Lys-C酶切前后的His₁₀-SUMO修饰底物进行两次富集，再利用胰蛋白酶酶切后留在底物肽段上的SUMO残基序列(QQTGG-)对SUMO位点进行质谱鉴定和数据库检索，从人源细胞中鉴定到了4300个SUMO化位点^[23]，优化后的方法不仅避免了泛素化位点带来的假阳性干扰，而且表达的His₁₀-标签进一步提高了富集效率。在此之后，该课题组进一步将K0方法用于对多种人源细胞系多个细胞刺激条件下His₁₀-SUMO2^{K0Q87R}修饰位点的大规模富集，并结合高pH C₁₈反相分级策略，共鉴定到40765个SUMO2位点，对应6747个蛋白质，实现了对人源细胞中外源性SUMO2位点的深度覆盖和全面分析^[30]。但是，基于K0的富集方法也存在弊端，将SUMO2上的所有K突变为R后，突变前K位点上的各种翻译后修饰信息将被丢失，包括重要的多聚SUMO化链的形成。

2 内源性SUMO化蛋白质组的富集方法

尽管针对外源性SUMO化蛋白质组目前已经报道了上万个SUMO化位点，但是将含有富集标签且氨基酸突变的外源性SUMO引入生物体内，不仅无法准确反映SUMO化修饰在生物体内的真实表达水平，而且不能用于实际的临床和组织样品的分析。因此，发展可以特异性富集内源性SUMO化修饰底物并对内源性SUMO化肽段进行有效鉴定的蛋白质组样品预处理方法是当前该领域的研究热点和难点。

2.1 蛋白质水平的富集

受限于抗体技术的发展，研究人员最初仅针对SUMO蛋白研发相应的抗体，并在蛋白质水平对SUMO修饰底物进行富集。Matafora等^[31]将经过蛋白酶抑制剂MG132处理的HeLa细胞进行细胞核蛋白提取，再利用SUMO1抗体对其进行免疫共沉淀(IP)，酶解后进行质谱分析，从中鉴定到193个SUMO1化蛋白质。此后，Becker等^[32]利用两个已知的SUMO1和SUMO2/3的单克隆抗体(SUMO1 21C7和SUMO2 8A2)，在对IP的洗脱条件进行优化后，将富集的蛋白进行胰蛋白酶酶解和质谱鉴定，从HeLa细胞中共富集到1000多个蛋白质，通过严格的数据筛选去除其中的非特异吸附蛋白，最终得到232个高可信的SUMO候选蛋白。然而以上研究均未对富集到的SUMO化蛋白的具体位点信息进行报道。

2.2 肽段水平的富集

仅在蛋白质水平富集内源性SUMO化修饰底物无法对内源性SUMO位点进行深度覆盖，一方面，抗体富集后的蛋白样品中仍然存在很多非特异性吸附蛋白的干扰，另一方面，酶解后产生的大量非SUMO化肽段会显著抑制SUMO化肽段的质谱响应。基于此，针对内源性SUMO化肽段的抗体富集策略在最近几年受到广泛关注。

2017年，Lumpkin等^[33]利用 α -溶菌酶(WaLP)能够特异性酶切苏氨酸(T)残基^[34]，使经过WaLP酶切后留在底物肽段上的SUMO1/2/3修饰只有GG-，并利用泛素化抗体(anti K- ϵ -GG)实现了对内源性SUMO1/2/3化肽段的同时富集，从人源细胞中鉴定到1209个内源性SUMO化位点。然而，该方法也存在一些不足，如WaLP的酶切特异性不高，泛素化位点上的精氨酸被WaLP酶切后产生的GG-残基将对鉴定结果产生假阳性干扰；且该方法无法区分鉴定到的SUMO化修饰的类别。同年，Cai等^[35]开发了一种特异性识别SUMO1被胰蛋白酶酶切后留在底物肽段上的氨基酸残基的泛-SUMO1抗体，并在肽段水平对SUMO1修饰底物进行富集，再利用Glu-C对SUMO1化肽段进行二次酶切以降低肽段长度，最后从小鼠睾丸中鉴定到了53个高可信的SUMO1位点。2018年，Hendriks等^[36]发展了利用抗体在肽段水平富集内源性SUMO2/3修饰位点的策略。他们利用可以特异性识别SUMO2/3被Lys-C酶切后留在底物肽段上的氨基酸残基的SUMO2/3抗体(SUMO2/3 8A2)富集SUMO2/3化肽段，并进一步采用Asp-N二次酶切和高pH C₁₈反相分级策略提高内源性SUMO2/3位点的鉴定覆盖度，从正常生长、热休克和MG132刺激的人源HEK293细胞中鉴定到14869个SUMO2/3位点和3870个内源性SUMO化蛋白，是目前报道的内源性SUMO2/3化位点的最大数据集。

目前采用免疫沉淀的抗体富集策略已覆盖到哺乳动物中一万多个SUMO化位点，3000多个底物蛋

白。然而, 抗体方法由于自身特性往往只局限于一类SUMO化修饰的富集, 难以实现样品中所有SUMO化修饰的同时富集和分析。针对上述问题, 本课题组^[37]利用胰蛋白酶酶切后的SUMO化肽段与非SUMO化肽段在碱性条件下带电性质的差异, 发展了一种基于强阴离子交换(SAX)色谱分离的内源性SUMO化肽段富集方法, 实现了对人源细胞中多种类型SUMO化修饰的同时富集。此外, 还通过采用Asp-N/Glu-C二次酶切降低SUMO化肽段长度, 提高了质谱对SUMO化肽段的鉴定效率; 进一步结合二维高低pH RPLC-ESI-MS/MS分析技术, 从HeLa细胞的酶解产物中富集到了177个内源性SUMO1位点和74个内源性SUMO2/3位点。

3 展 望

随着基因工程和抗体等富集技术的不断发展, 针对外源性SUMO化修饰位点的规模化鉴定已经日趋完善。然而, 只有内源性的SUMO化修饰才能真正反映实际细胞/组织样品中SUMO的表达水平和动态变化。相比于其他翻译后修饰, 内源性SUMO化修饰的覆盖深度远未达到研究的需要。因此, 发展针对内源性SUMO化位点的高通量富集方法仍是当前SUMO化蛋白质组学的首要任务。此外, 目前亟需开发专门针对SUMO化位点鉴定的数据库检索软件。随着富集方法的不断创新、高灵敏质谱和生物信息学技术的不断发展, 人们对SUMO化修饰的认识会更加深入, SUMO化蛋白质组学也将会为发现更多关键的疾病标志物和新的治疗靶标提供新的方向。

参考文献:

- [1] Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. *Cell*, **1997**, 88(1): 97-107.
- [2] Celen A B, Sahin U. *FEBS J.*, **2020**, 287(15): 3110-3140.
- [3] Eifler K, Vertegaal A C. *FEBS J.*, **2015**, 282(19): 3669-3680.
- [4] Johnson E S. *Annu. Rev. Biochem.*, **2004**, 73: 355-382.
- [5] Guo D H, Li M Y, Zhang Y, Yang P, Eckenrode S, Hopkins D, Zheng W P, Purohit S, Podolsky R H, Muir A, Wang J Z, Dong Z, Brusko T, Atkinson M, Pozzilli P, Zeidler A, Raffel L J, Jacob C O, Park Y, Serrano-Rios M, Larrad M T M, Zhang Z X, Garchon H J, Bach J F, Rotter J I, She J X, Wang C Y. *Nat. Genet.*, **2004**, 36(8): 837-841.
- [6] Liang Y C, Lee C C, Yao Y L, Lai C C, Schmitz M L, Yang W M. *Sci. Rep.*, **2016**, doi:10.1038/srep26509.
- [7] Chang H M, Yeh E T H. *Physiol. Rev.*, **2020**, 100(4): 1599-1619.
- [8] Boulanger M, Chakraborty M, Tempe D, Piechaczyk M, Bossis G. *Molecules*, **2021**, 26(4): 828-866.
- [9] Kunz K, Piller T, Muller S. *J. Cell Sci.*, **2018**, doi:10.1242/jcs.211904.
- [10] Hickey C M, Wilson N R, Hochstrasser M. *Nat. Rev. Mol. Cell Biol.*, **2012**, 13(12): 755-766.
- [11] Rodriguez M S, Dargemont C, Hay R T. *J. Biol. Chem.*, **2001**, 276(16): 12654-12659.
- [12] Matic I, Schimmel J, Hendriks I A, van Santen M A, van de Rijke F, van Dam H, Gnad F, Mann M, Vertegaal A C. *Mol. Cell*, **2010**, 39(4): 641-652.
- [13] Morozko E L, Smith-Geater C, Monteys A M, Pradhan S, Lim R G, Langfelder P, Kachemov M, Hill A, Stocksdale J T, Cullis P R, Wu J, Ochaba J, Miramontes R, Chakraborty A, Hazra T K, Lau A, St-Cyr S, Orellana I, Kopan L, Wang K Q, Yeung S, Leavitt B R, Reidling J C, Yang X W, Steffan J S, Davidson B L, Sarkar P S, Thompson L M. *Proc. Natl. Acad. Sci. U. S. A.*, **2021**, doi: 10.1073/pnas.2021836118.
- [14] Kessler J D, Kahle K T, Sun T T, Meerbrey K L, Schlabach M R, Schmitt E M, Skinner S O, Xu Q K, Li M Z, Hartman Z C, Rao M, Yu P, Dominguez-Vidana R, Liang A C, Solimini N L, Bernardi R J, Yu B, Hsu T, Golding I, Luo J, Osborne C K, Creighton C J, Hilsenbeck S G, Schiff R, Shaw C A, Elledge S J, Westbrook T F. *Science*, **2012**, 335(6066): 348-353.
- [15] Hendriks I A, Treffers L W, Verlaan-de Vries M, Olsen J V, Vertegaal A C O. *Cell Rep.*, **2015**, 10(10): 1778-1791.
- [16] Hannoun Z, Maarifi G, Chelbi-Alix M K. *Cytokine Growth Factor Rev.*, **2016**, 29: 3-16.
- [17] Li T, Huang S S, Dong M H, Gui Y P, Wu D L. *Urol. Oncol. - Semin. Orig. Invest.*, **2013**, 31(8): 1539-1545.
- [18] Bawa-Khalfe T, Yang F-M, Ritho J, Lin H-K, Cheng J, Yeh E T H. *Oncotarget*, **2017**, 8(11): 17651-17664.

- [19] Marcelli S, Ficulle E, Piccolo L, Corbo M, Feligioni M. *Pharmacol. Res.*, **2018**, 130: 420–437.
- [20] Hendriks I A, Vertegaal A C. *Nat. Rev. Mol. Cell Biol.*, **2016**, 17(9): 581–595.
- [21] Galisson F, Mahrouche L, Courcelles M, Bonneil E, Meloche S, Chelbi–Alix M K, Thibault P. *Mol. Cell. Proteomics*, **2011**, 10(2): M110004796.
- [22] Lamoliatte F, Caron D, Durette C, Mahrouche L, Maroui M A, Caron–Lizotte O, Bonneil E, Chelbi–Alix M K, Thibault P. *Nat. Commun.*, **2014**, 5: 5409.
- [23] Hendriks I A, D’Souza R C, Yang B, Verlaan–de Vries M, Mann M, Vertegaal A C. *Nat. Struct. Mol. Biol.*, **2014**, 21(10): 927–936.
- [24] Tammsalu T, Matic I, Jaffray E G, Ibrahim A F, Tatham M H, Hay R T. *Nat. Protoc.*, **2015**, 10(9): 1374–1388.
- [25] Tirard M, Hsiao H–H, Nikolov M, Urlaub H, Melchior F, Brose N. *Proc. Natl. Acad. Sci. U. S. A.*, **2012**, 109(51): 21122–21127.
- [26] Schimmel J, Eifler K, Sigurdsson J O, Cuijpers S A G, Hendriks I A, Verlaan–de Vries M, Kelstrup C D, Francavilla C, Medema R H, Olsen J V, Vertegaal A C O. *Mol. Cell*, **2014**, 53(6): 1053–1066.
- [27] Hendriks I A, D’Souza R C, Chang J G, Mann M, Vertegaal A C. *Nat. Commun.*, **2015**, 6: 7289.
- [28] Impensa F, Radoshevich L, Cossart P, Ribet D. *Proc. Natl. Acad. Sci. U. S. A.*, **2014**, 111(34): 12432–12437.
- [29] Lamoliatte F, McManus F P, Maarifi G, Chelbi–Alix M K, Thibault P. *Nat. Commun.*, **2017**, 8: 14109.
- [30] Hendriks I A, Lyon D, Young C, Jensen L J, Vertegaal A C, Nielsen M L. *Nat. Struct. Mol. Biol.*, **2017**, 24(3): 325–336.
- [31] Matafora V, D’Amato A, Mori S, Blasi F, Bachi A. *Mol. Cell. Proteomics*, **2009**, 8(10): 2243–2255.
- [32] Becker J, Barysch S V, Karaca S, Dittner C, Hsiao H H, Berriel Diaz M, Herzig S, Urlaub H, Melchior F. *Nat. Struct. Mol. Biol.*, **2013**, 20(4): 525–531.
- [33] Lumpkin R J, Gu H, Zhu Y, Leonard M, Ahmad A S, Clauser K R, Meyer J G, Bennett E J, Komives E A. *Nat. Commun.*, **2017**, 8(1): 1171.
- [34] Meyer J G, Kim S, Maltby D A, Ghassemian M, Bandeira N, Komives E A. *Mol. Cell. Proteomics*, **2014**, 13(3): 823–835.
- [35] Cai L L, Tu J, Song L, Qin J, Ding C, He F C. *Mol. Cell. Proteomics*, **2017**, 16(5): 717–727.
- [36] Hendriks I A, Lyon D, Su D, Skotte N H, Daniel J A, Jensen L J, Nielsen M L. *Nat. Commun.*, **2018**, 9(1): 2456.
- [37] Li Y, Sun M W, Hu Y C, Shan Y C, Liang Z, Zhang L H, Zhang Y K. *Anal. Chim. Acta*, **2021**, doi:10.1016/j.aca.2021.338324.

(责任编辑: 盛文彦)