论 文

健康人的尿液中为什么会有蛋白质?

刘永涛¹, 赵敏迪², 潘宣圳¹, 高友鹤^{1*}

1. 北京师范大学生命科学学院, 生物化学与分子生物学系, 基因工程药物及生物技术北京市重点实验室, 北京 100875;

2. 北京医院国家老年医学中心检验医学科, 北京 100730

* 联系人: E-mail: gaoyouhe@bnu.edu.cn

收稿日期: 2019-10-31; 接受日期: 2020-02-18; 网络版发表日期: 2020-03-04

国家重点研发计划(批准号: 2018YFC0910202, 2016YFC1306300)、北京自然科学基金(批准号: 7172076)、北京合作建设项目(批准号: 110651103)、北京师范大学项目(批准号: 11100704)和北京协和医学院(批准号: 2016-2.27)资助

摘要 蛋白质是人必需的营养物质, 但为什么会有一小部分蛋白质通过尿液被排掉? 健康人尿液中的蛋白质含量较低, 而其种类却十分丰富, 其中存在的蛋白质可能是机体不需要的废弃物, 也可能是会对机体产生毒害作用而必须被排出。相比小分子代谢物而言, 蛋白质的结构复杂、功能多样, 即使分子水平上的一个小的改变都会影响其后续的生物学功能, 通过全面地比较血浆和尿液蛋白质组分子水平修饰上的差异, 或许可为解释尿液中存在蛋白质的现象提供线索。在这项研究中, 共收集了9例健康人的尿液和9例健康人的血浆样品, 通过高分辨串联质谱及非标记定量的蛋白质组学技术对样本进行分析, 并使用pFind非限制性修饰鉴定算法整体比较了两种类型样品之间的分子修饰差异情况。结果表明, 尿液蛋白质组中的半胱氨酸(cysteine, Cys)修饰变为脱氢丙氨酸(dehydroalanine, Dha)的数量显著高于血浆。[Cys→Dha]这种分子修饰破坏了原有蛋白质的二硫键, 从而影响甚至改变蛋白质的结构与生物功能。综上, 本研究揭示了血浆和尿液在整体蛋白质组修饰上的差异, 并指出可能由于蛋白质分子修饰的差异造成了其结构不可逆的改变, 进而使得失去功能或者有毒的蛋白质经血液被排进尿液中。

关键词 尿液, 血浆, 蛋白质组, 分子修饰, 非限制性修饰鉴定算法

蛋白质是人所必需的营养物质, 也是行使生命活动的最基本单位, 其重要地位不言而喻, 但为什么会有一小部分蛋白质通过尿液被排掉? 健康人尿液中虽然蛋白质总量较低, 但种类却异常丰富。已知正常人尿液中可以鉴定到6000余种不同的蛋白质^[1]。关于尿液中为何会含有种类丰富的蛋白质, 至今没有确切的定论。正常情况下, 肾小球可筛过一部分蛋白质, 其中98%在肾小管被重吸收回体内, 剩余2%的蛋白质与肾小管及其他尿路上皮细胞分泌的少量黏蛋白一起排

出。尿液产自肾脏, 并在膀胱中储存了几个小时后, 因此尿液中蛋白质信息可直观地反映泌尿系统的健康状况。但尿液中大多数蛋白质是通过肾小球过滤血浆后汇集在一起的, 理论上它们都来源于血液, 在经过肾脏的过滤与重吸收依旧能够保留下来, 足够说明剩余的2%蛋白质也不可能是被健康的肾脏随机、非选择性地“漏出”, 这确实是一个耐人寻味的现象。在进行这项研究之前, 本课题组^[2]曾尝试进行了一个简单的动物实验: 收集若干只经一段时间禁食大鼠的尿液并对

引用格式: 刘永涛, 赵敏迪, 潘宣圳, 等. 健康人的尿液中为什么会有蛋白质? 中国科学: 生命科学, 2020, 50: 338–348

Liu Y T, Zhao M D, Pan X Z, et al. Why are there proteins in the urine of healthy people? (in Chinese). Sci Sin Vitae, 2020, 50: 338–348, doi: 10.1360/SSV-2019-0235

其样本进行蛋白质组学分析,出乎意料地发现,该样本中仍存在种类与含量丰富的蛋白质。蛋白质是行使生命活动最重要的物质之一,蛋白质的降解和合成一直是体内最高效的生化反应。既然蛋白质如此宝贵,机体或是肾脏不会随便将其浪费,尤其是在严重饥饿的情况下。那么为什么可以在健康人的尿液中找到种类如此丰富的蛋白质呢?既然出现在尿液中,可以推测,这些蛋白质应为机体代谢的废弃物,或许它们无法再被重新利用,甚至是对机体会产生有害作用,必须立刻被清除,以免导致机体正常生理代谢功能的紊乱。

蛋白质是生命活动的主要承担者,以其氨基酸组成种类多样,高级结构复杂而有别于其他类型的化合物,凭此向世界展示出绚烂多彩的生命形态。蛋白质中不同的氨基酸分子也会参与种类众多的酶促或非酶促反应,这些作用于氨基酸残基上的化学反应可被认为是蛋白质分子水平上的修饰,而不同的修饰作用可对蛋白质的高级结构产生影响,甚至决定了蛋白质功能的完整性。同样,蛋白质也需要被修饰。通常在基因表达后,大多数的蛋白质需要经过不同程度的分子修饰才能执行功能。蛋白质的分子修饰可通过修饰蛋白质链中氨基酸残基上的现有基团或引入新的基团来实现^[3],像蛋白质链中半胱氨酸残基之间形成的二硫键也是一种分子修饰类型^[4]。经过分子修饰的蛋白质还可以更客观地反映生命的各种生理和病理过程^[5,6],如蛋白质的羰基化修饰会使蛋白质降解并导致蛋白质聚集体的形成^[7]。同时,特定的氨基酸修饰可用作生物标志物对健康进行诊断,如指示氧化损伤的生物标记等^[8]。蛋白质分子修饰水平上的一点改变就会对整个蛋白质的结构与功能产生重要影响,而全面地比较血浆与尿液蛋白质组分子修饰间的差异或许能为发现尿液中存在蛋白质的现象提供一定的线索。

如今生物信息学与蛋白质组学融合越来越紧密,同时也为蛋白质组学在分子修饰水平等领域提供了强大的帮助,高分辨的质谱仪(mass spectrometry, MS)已成为分析样品蛋白质组及其分子修饰最重要的工具。据人类蛋白质参考数据库(human protein reference database, HPRD, <http://www.hprd.org/>)信息显示,在其收录的30047种人类蛋白质中,鉴定出93710个不同的蛋白质修饰位点。通过高分辨的质谱分析已发现近1500种不同类型的蛋白质分子修饰种类,UniMod数据库(<http://www.unimod.org/>)将所有已知类型的蛋白质修

饰和基于质谱的检测信息进行了汇总,极大方便了人们对蛋白质分子修饰信息的检索工作^[9]。

目前,关于血浆和尿液蛋白质组整体修饰水平之间的研究较少。根据本课题组^[10,11]之前的工作基础,本次借助于open-pFind这种生物信息学算法,继续在血浆和尿液蛋白质组分子修饰之间进行更深入、全面且非限制性的搜索鉴定。pFind Studio是一款用于蛋白质组数据库检索的专业软件(<http://pfind.ict.ac.cn/>),可以自动识别二级质谱(second-order mass spectrometry, MS2)数据中的肽和蛋白质的信息。它具有模块化的定制平台,供用户测试和比较其算法,并可以根据开放应用程序编程接口(application programming interface, API)标准创建自己的模块,然后将其添加到工作流中,而不必传统地使用默认模块。其开放式搜索功能(open search)是把UniMod数据库与二级质谱肽段匹配的结果集成在一起,并通过算法提供了有关肽段和蛋白质修饰的信息,以实现非限制性鉴定^[12-14]。

在这项研究中,本课题组试图探索有关尿液中含有蛋白质的问题,同时提出假设,即健康人尿液中的蛋白质或许是因为血液中的蛋白质在某些方面发生改变,失去了利用价值而被排出。基于以上推论,期待从蛋白质分子修饰层面寻找线索。通过液相色谱串联高分辨质谱与bottom-up的非标记定量蛋白质组学技术相结合,对人尿液和血浆蛋白质组之间修饰信息进行采集。使用open-pFind算法分析实验数据,以便获取样本整体的修饰情况信息。

1 材料与方法

1.1 样品的收集

9例健康人的血浆样品从北京医院的检验科收集获得。样品在健康人群中随机选择,并且对其饮食、药物和其他因素没有任何限制或要求。9例健康人的尿液样本从实验室的健康志愿者中收集。该研究得到北京医院和北京师范大学伦理委员会的批准,本实验已向志愿者提供研究的详细情况,内容包括研究目的、方法及过程,同时对志愿者的个人资料严格保密。样本信息详见表1。

1.2 蛋白质样品制备和胰蛋白酶酶解

血浆样品($n=9$)用Milli-Q水稀释40倍,尿液样品

表 1 获得的血浆与尿液样品信息与统计结果

Table 1 Statistical data of plasma and urine samples

		血浆样本(n=9)	尿液样本(n=9)
性别	男性	2	5
	女性	7	4
年龄	中位数	31±4	26±2
	平均值	26~37	22~29

($n=9$)则不需要稀释. 血浆样品取用100 μL , 尿液样品取用20 mL. 使用20 mmol/L的二硫苏糖醇(dithiothreitol, DTT)在37℃下与样品反应1 h, 使蛋白质结构中的二硫键变性, 随后添加55 mmol/L的碘乙酰胺(iodoacetamide, IAM)并避光反应30 min以使二硫键结合位点烷基化. 在-20℃下, 用三倍体积的预冷丙酮沉淀上清液2 h, 然后在4℃和12000 $\times g$ 下离心30 min, 获得蛋白质沉淀. 随后将沉淀物重悬于适量蛋白溶解液(8 mol/L尿素, 2 mol/L硫脲, 25 mmol/L DTT和50 mmol/L Tris)中. 使用Bradford蛋白定量法测量蛋白质提取液浓度. 结合膜上酶解(filter-aided sample preparation, FASP)的方法^[15], 每个样品按质量比50:1的比例使用胰蛋白酶(Trypsin Gold, Mass Spec Grade, Promega, 美国)酶解100 μg 的蛋白质样品. 37℃酶解14 h后, 向溶液中加入10%甲酸溶液使甲酸终浓度控制在0.1%, 终止酶解, 经10 kD超滤管离心后获得多肽溶液, 使用BCA(bicinchoninic acid)蛋白定量法测定多肽的浓度, 计算多肽量. 并通过真空离心浓缩仪(Thermo Fisher, 美国)进行干燥, 干燥后的肽段密封至于-80℃保存.

1.3 液相色谱-串联质谱分析和数据库搜索

分析前, 需将干燥的多肽样品溶于0.1%甲酸, 控制多肽样本终浓度在0.1 $\mu\text{g}/\mu\text{L}$, 每个样品按1 μg 多肽质量进行分析, 使用Hela细胞蛋白裂解液(Thermo Fisher, 美国)作为整个分析过程的质量控制(quality control, QC): Thermo EASY-nLC1200色谱系统加载至预柱与分析柱上. 通过Thermo Orbitrap Fusion Lumos质谱系统(Thermo Fisher, 德国)采集蛋白质组数据. 液相色谱分析方法: 预柱: 75 $\mu\text{m}\times 2\text{ cm}$, nanoViper C18, 2 μm , 100 $\text{\AA}$; 分析柱: 50 $\mu\text{m}\times 15\text{ cm}$, nanoViper C18, 2 μm , 100 $\text{\AA}$; 进样体积: 10 μL , 流速: 250 nL/min. 流动相: A

相: 100%质谱级水(Thermo Fisher, 西班牙)/1%甲酸(Thermo Fisher, 美国), B相: 80%乙腈(Fisher Scientific, 美国)/20%水/1%甲酸, 120 min梯度洗脱: 0 min, 3% B相; 0~3 min, 8% B相; 3~93 min, 22% B相; 93~113 min, 35% B相; 113~120 min, 90% B相. 质谱分析方法, 离子源: nanoESI, 喷雾电压: 2.0 kV, 毛细管温度: 320℃, S-lens RF Level: 30; 分辨率设置: 一级(Orbitrap)120000@m/z 200, 二级(Orbitrap)30000@m/z 200; 母离子扫描范围: m/z 350~1350; 子离子扫描范围: 从m/z 110开始扫描, MS1 AGC: 4e5, 电荷范围: 2~7, 离子注入时间: 50 ms, MS2 AGC: 1e5, 离子注入时间: 50 ms, 离子筛选窗口: m/z 2.0, 碎裂模式: 高能碰撞诱导裂解(higher-energy collisional dissociation, HCD), 能量NCE 32, Data-dependent MS/MS: Top 20, 动态排除时间: 15 s, 内部校准质量数: 445.12003.

使用pFind Studio软件(3.1.3版, 中国科学院计算技术研究所)对质谱采集的数据进行无标记定量分析. 目标检索数据库来自Uniprot下载的Homo Sapiens数据库(更新至2018年10月). 检索时, 仪器类型为HCD-FTMS, 酶全特异性为胰蛋白酶, 最多有2个漏切位点. 选择开放式搜索. 筛选条件: 在肽水平上的假阳性率(false discovery rate, FDR)小于1%, 在蛋白水平上的Q值小于1%. 数据同时使用正向和反向数据库检索策略进行分析.

2 结果

2.1 通过使用bottom-up的蛋白质组学技术对总蛋白进行鉴定

在非标记分析中, 通过高分辨串联质谱仪分析了18个(9/9)蛋白质组样品. 在基于open-pFind软件检索数据(.raw)之后, 可以在pBuild中浏览并导出分析结果. 将样品结果($n=18$)进行整理和统计, 样品中蛋白质和多肽的鉴定结果如表2所示.

通过120 min液相梯度分析, 尿液样品中鉴定出1247~1921种蛋白质(平均值为1576), 血浆样品中鉴定出605~777种蛋白质(平均值为718).

2.2 两种类型样品中蛋白质修饰位点的数量统计

通过pFind Studio的开放式搜索, 在所有样品中总共鉴定出1247个不同的修饰. 含量最高的修饰类型是由人工烷基化试剂(碘乙酰胺, IAM)作用于蛋白质链

表 2 样品中蛋白质和多肽的鉴定结果

Table 2 The results of proteins and peptides identification in samples

序号	谱图数	扫描数目	多肽数目	序列数目	蛋白质数目	平均鉴定覆盖率
U1	54513	41426	16672	11673	1528	26.35%
U2	58458	44158	16938	11581	1569	27.31%
U3	52400	40187	16757	11803	1544	26.71%
U4	51519	39645	15800	11151	1490	25.84%
U5	57983	43913	18569	13041	1818	27.16%
U6	59448	45099	18181	12000	1554	29.77%
U7	61770	45901	19328	14183	1921	28.70%
U8	47555	38449	15701	10354	1247	28.43%
U9	54102	42610	18109	11967	1573	28.65%
P1	31779	27506	7672	4694	759	37.71%
P2	32401	27976	7526	4621	605	38.99%
P3	32620	28144	8131	5028	758	40.31%
P4	40627	34992	10758	6128	685	48.56%
P5	42032	35817	10649	6105	758	45.79%
P6	40953	34904	10483	5944	707	44.63%
P7	40543	34584	10154	5781	721	45.26%
P8	41380	35352	10710	6109	777	45.95%
P9	41311	35318	10253	5830	691	47.02%

U: 尿液; P: 血浆

中半胱氨酸残基(Cys, C)巯基上所发生的脲甲基化(carbamidomethyl)修饰. 除了IAA引入的几种修饰之外, 甲硫氨酸(Met, M)的氧化、天冬酰胺(Asp, D)和谷氨酰胺(Gln, Q)的脱酰胺化、谷氨酸(Glu, E)或谷氨酰胺的焦谷氨酸化(pyrp-Glu)、蛋白质中N末端的乙酰化也是常见的修饰. 图1展示了尿液和血浆样品中数量排在前十名的修饰类型统计结果与比较.

2.3 血浆蛋白质组和尿液蛋白质组修饰的比较

通过分析发现, 尿蛋白中修饰位点的平均数为10000, 而血浆中的修饰位点的平均数为6000. 在血浆中, 赖氨酸(Lys, K)的氨甲酰化量显著高于尿液, 而尿液中蛋白质中Cys的脱巯基化(-SH)变成脱氢丙氨酸(Dha)的量要比血浆高得多. 表3以样品中主要的高丰度蛋白质人血清白蛋白(human serum albumin, HSA)为例, 统计了质谱鉴定的结果和修饰信息. 数据经过搜库比对与分析后, 会给出每种蛋白质的鉴定覆盖率, 其

值越高, 提示蛋白质碎片的“复原”程度越高. 经对样品统计, 在血浆中HSA的鉴定覆盖率为93.5%±1.1%, 尿液中为85.6%±2.3%, 依靠对样品中白蛋白的质谱鉴定覆盖率进行归一化, 并对每种发现修饰所占总肽段数的进行计算统计, 算法如下: 占总肽段数的比例=修饰数量÷样品中多肽鉴定总量×鉴定覆盖率, 图2展示了血浆和尿液中白蛋白的几种修饰情况比较.

如表3所示, 通过质谱对HSA鉴定覆盖率进行归一化校正后, 两种类型的样品中脲甲基化修饰所占比例无统计学差异(显著性大于0.1), 这说明使用该方法的归一化效果较好. 尿液中HSA的[Cys→Dha]修饰是血浆中的两倍, 而血浆HSA中的赖氨酸氨甲酰化明显比尿液中该修饰的数量高出数倍, 但是血浆样品组内该修饰的差异也较大. 图3为几种类型肽段修饰的MS2鉴定结果.

通过观察HSA半胱氨酸位点修饰的情况发现, 蛋白质链中的半胱氨酸残基是一种容易参与化学修饰的

图 1 样品中蛋白质修饰数量的统计数据. A: 尿液样品中修饰数量排名前10的修饰类型, 红色标注的为半胱氨酸变为脱氢丙氨酸修饰, 其在尿液中排名显著高于血浆; B: 血浆样品中修饰数量排名前10的修饰类型, 绿色标注的为赖氨酸氨甲酰化修饰, 其在血浆中排名显著高于尿液; C: 血浆和尿液中的主要修饰类型之间的比较

Figure 1 Statistics of the number of protein modifications in the sample. A: Top 10 types of modifications in the urine sample, the red mark is cysteine to dehydroalanine modification, which rank significantly higher in urine than plasma; B: top 10 modification types in the number of modifications in plasma samples, the green labelled lysine carbamyl modifications, which rank significantly higher in plasma than urine; C: comparison of major modification types in plasma and urine

图 2 几种类型修饰在人血浆和尿液中占总肽段鉴定水平的比值

Figure 2 The ratio of some modifications to the total peptide identification level in human plasma and urine

氨基酸种类。将两种类型样本中鉴定到与半胱氨酸有关的修饰信息进行了整理与统计,发现在血浆中的200种非低丰度修饰(指该种修饰的数量至少鉴定到10次)中,与半胱氨酸有关的修饰种类有37种,占修饰类型总数的18.5%,而在尿液的171种非低丰度修饰中,与半胱氨酸有关的修饰种类有29种,占修饰类型总数的17.0%。数量上看,与半胱氨酸修饰相关的修饰数量分别占血浆和尿液修饰总量的27.0%及28.7%。由此可见,半胱氨酸是一种在蛋白质分子中极为重要且活泼的氨基酸残基,不论是种类还是数量都占据了蛋白质分子修饰中重要的一部分。表4详细统计了血浆与尿液样本中与半胱氨酸有关的修饰信息。

3 讨论

生命体中的蛋白质不论在细胞内还是细胞外,均

可在其翻译后进行某些种类的分子修饰。不同类型的修饰会影响蛋白质中的电荷状态、疏水性、构象或稳定性,并最终影响其功能。影响蛋白质结构的因素很多,其中之一就是其分子水平上的修饰。基于这项研究的结果,本课题组提出了一个假设,血浆中某些蛋白质在分子水平的修饰发生了变化,引起了结构和功能的改变,从而导致它们无法被重复利用,甚至对机体有毒害作用,因此必须通过尿液排出体外。这将导致在尿液中观察到的蛋白质分子修饰与血液中的蛋白质分子修饰不同。通过对两类样本的整体修饰检索比较后,本课题组观察到的现象印证了这个假设。

通过分析样品中的高丰度蛋白质,发现半胱氨酸脱巯基(-SH)成为脱氢丙氨酸[Cys→Dha]这种修饰在尿液蛋白质中比在血浆中的修饰程度更高。半胱氨酸是一种重要的氨基酸,也很容易参与种类繁多的修饰,UniMod中收录的信息显示,有数百种蛋白质修饰与半胱氨酸残基有关。人血清白蛋白(HSA)是由609个氨基酸组成的蛋白质,其中包含35个半胱氨酸残基,共有17对二硫键。Paris等人^[16]对HSA蛋白质的高级分子结构及其中的17对二硫键的相关位置已有详细报道。

尽管在蛋白质分子中,绝大部分的半胱氨酸残基处的巯基应处于二硫键(S-S)状态,但实际上在蛋白质中也可以检测到游离巯基,这使得它们有机会可以参与多种分子修饰。半胱氨酸残基之间形成的共价二硫键在稳定许多蛋白质的天然结构中起特殊作用,二硫键可以更好地保持蛋白质的疏水性。但是,在蛋白质的“生命周期”中,有时可能会在还原性环境下打开二硫键,使其还原。还原态的二硫键是不稳定且活泼的,易发生多种化学修饰,其中半胱氨酸变成脱氢丙氨酸(Dha)是最常见的修饰类型。半胱氨酸的脲甲基化修饰与[Cys→Dha]修饰的比例可以反映样品蛋白质中游离

表 3 血浆和尿液中人血清白蛋白修饰的类型和数量的比较

Table 3 Comparison of the types and quantities of modification in HSA between plasma and urine samples

	尿液样本(n=9)中的白蛋白		血浆样本(n=9)中的白蛋白		显著性
	PSM数	占总肽段数比例	PSM数	占总肽段数比例	
多肽鉴定总数	4067±613	—	14041±1843	—	
Cys→Dha修饰	58±23	1.21%±0.4%	85±17	0.56%±0.06%	0.0004
K氨甲酰修饰	12±4	0.26%±0.08%	458±376	3.06%±2.40%	0.004
脲甲基化修饰	1473±97	31.5%±3.6%	4408±691	29.2%±1.3%	0.11

注: PSM即peptide-spectrum matching, 指通过搜库得到质谱图谱与肽段的比较结果

图 3 二级质谱图(MS2)鉴定出样品中部分肽的相关修饰及b/y离子信息. A: 半胱氨酸的[Cys→Dha]修饰; B: 赖氨酸的氨甲酰化修饰

Figure 3 MS2 identifies relevant modification and b/y ions information in the sample partial peptides. A: [Cys→Dha] of cysteine; B: lysine carbamoyl

巯基及完整二硫键的数目, 而游离巯基数目可以反映蛋白质结构的稳定性, 即[Cys→Dha]越高的样品, 其蛋白质结构越不稳定, 也表明其结构被破坏的可能性就越大. 在开发基于重组单克隆抗体的生物治疗药物时, 应密切注意这些新的二硫键相关结构, 抗体结构和稳定性变化与二硫键之间的重要性已得到证实. 这表明, 二硫键的用处与IgG(immunoglobulin G)分子的

稳定性、结构和生物学功能有关^[17].

以上的例子充分说明了二硫键对于蛋白质结构的重要性, 甚至可以推测, 哪怕是蛋白质分子中一个关键性的二硫键被破坏也极有可能会整个蛋白质结构崩溃, 进而失去功能而被排出体外. 与此同时, 最近的一项研究表明, 在慢性肝病及慢性肾脏病或糖尿病患者血液样本中, 人血清白蛋白中的半胱氨酸Cys34

表 4 血浆与尿液样本中半胱氨酸残基位点所参与分子修饰的信息

Table 4 Information on modification of cysteine residue sites in plasma and urine samples

序号	修饰名称	血浆中比例(%)	尿液中比例(%)	修饰类型
1	脲甲基化修饰	18.37	22.45	人工引入修饰
2	双脲甲基化修饰	1.32	1.99	人工引入修饰
3	Cys→Dha修饰	0.42	0.61	化学衍生修饰
4	脲甲基DTT修饰	0.34	0.15	人工引入修饰
5	羧甲基化修饰	2.42	1.34	化学衍生修饰
6	马来酰亚胺修饰	0.16	0.23	化学衍生修饰
7	N-甲基马来酰亚胺+水修饰	—	0.17	化学衍生修饰
8	消除_H(4)C(3)O(1)	0.25	0.15	其他
9	HMVK修饰	0.10	0.11	化学衍生修饰
10	谷胱甘肽修饰	0.17	0.14	翻译后修饰
11	半胱氨酰修饰	—	0.10	聚合
12	丙酰胺修饰	0.17	0.08	人工引入修饰
13	二氧化硫结合	—	0.09	翻译后修饰
14	氧化O2	0.16	0.08	翻译后修饰
15	乙基马来酰亚胺+水	—	0.07	化学衍生修饰
16	消除_H(6)C(3)O(1)	0.11	0.09	其他
17	丙酰胺_2H(3)修饰	0.09	0.07	同位素标记
18	维生素K-HQ修饰	—	0.07	化学衍生修饰
19	脱氢	0.10	0.08	聚合
20	羧甲基_13C(2)修饰	0.13	0.06	同位素标记
21	氧化	—	0.06	翻译后修饰
22	丁烯醛修饰	0.11	0.08	其他
23	羟肉桂基修饰	—	0.08	翻译后修饰
24	硫化物修饰	—	0.07	翻译后修饰
25	BMOE修饰	0.12	0.06	化学衍生修饰
26	Cys→Ser	—	0.06	氨基酸替换
27	N-甲基马来酰亚胺修饰	—	0.06	化学衍生修饰
28	亚硝基	0.09	0.06	翻译后修饰
29	二吡咯基甲烷甲基化修饰	0.10	0.06	翻译后修饰
30	甲基化修饰	0.16	—	翻译后修饰
31	NEIAA修饰	0.16	—	同位素标记
32	氧化O3	0.13	—	化学衍生修饰
33	N-乙马来酰亚胺硫化	0.21	—	化学衍生修饰
34	泛素-Br2修饰	0.21	—	化学衍生修饰
35	丙二酰修饰	0.15	—	化学衍生修饰
36	NO_SMX_SIMD修饰	0.12	—	化学衍生修饰
37	羧甲基DTT修饰	0.11	—	人工引入修饰
38	磷酸二乙酯修饰	0.11	—	化学衍生修饰
39	DNCB_半抗原修饰	0.10	—	化学衍生修饰
40	NO_SMX_SMCT修饰	0.10	—	化学衍生修饰
41	2-二甲基琥珀酰修饰	0.09	—	化学衍生修饰
42	2-琥珀酰修饰	0.10	—	化学衍生修饰
43	NIPCAM修饰	0.11	—	化学衍生修饰
44	氧化+N-乙马来酰亚胺修饰	0.10	—	化学衍生修饰
45	氨甲酰化	0.12	—	人工引入修饰
46	乙酰化	0.09	—	翻译后修饰
47	DNPS修饰	0.09	—	化学衍生修饰
总计	47	26.99	28.72	

修饰可用作氧化应激相关疾病的标志物^[18]。真核细胞中的氧化还原酶包括蛋白质二硫键异构酶^[19]和硫氧还蛋白^[20], 它们会加速二硫键的还原。同时, 分子内或域间的二硫键对于分泌途径中许多蛋白质的结构稳定性至关重要。Bechtel和Weerapana^[21]全面概述了二硫键的功能作用及其与人类疾病的相关性。因此, 如果二硫键被破坏, 蛋白质的功能将不可避免地受到影响。

理论上, 机体应将某些功能丧失的蛋白质(例如[Cys→Dha]修饰的蛋白质)降解并重新利用、合成新的蛋白质。但实际上, 大多数此类蛋白质会被排入尿液。为什么它们不能被循环利用? 也许这是由于半胱氨酸修饰变成脱氢丙氨酸(Dha)的这个反应是不可逆的。Dha是一种亲电试剂, 在含有Dha的肽段中存在一个不饱和的碳-碳双键, 而它在体外蛋白质链中起烷化剂的作用^[22]。许多含有脱氢丙氨酸的肽具有一定的毒性^[23], 因此, 发生这种修饰的蛋白质或肽也都具有一定的毒性, 它们在体内停留久了便会对机体产生毒害作用, 最好立即将其排出体外。因此, 尿液中会发现大量含有修饰成为脱氢丙氨酸的蛋白质。

除HSA以外, 其他蛋白质和蛋白质家族在尿液中具有大量[Cys→Dha]修饰, 例如衰变促进因子, 表皮生长因子, 免疫球蛋白(IgG), α 球蛋白和许多免疫球蛋白(Ig κ 链), 补体(补体因子B/D等)等。某些蛋白质(例如淀粉酶和纤连蛋白)中[Cys→Dha]修饰的比例不高, 可能是这些蛋白质中的半胱氨酸残基所占比例较低导致的。在主要尿蛋白家族(major urinary protein)的分子中, [Cys→Dha]修饰率较正常值更高。据推测, 在体内被[Cys→Dha]修饰的肽或蛋白质可以通过组合剪接形成。推测这种蛋白质可能没有实质性的功能与作用, 只是为了帮助将这些含有此类修饰的多肽排泄到尿液中。

上述发现的[Cys→Dha]修饰或许并不是唯一会导致尿液中含有蛋白质的原因, 因为研究同时发现, 在尿液中的某些蛋白质(如B型骨桥蛋白、分泌磷蛋白等)不含有半胱氨酸, 所以它们不能发生与半胱氨酸相关

的各种修饰。本研究仅从一个方面深入探讨了尿液中存在蛋白质的现象, 而所发现的[Cys→Dha]修饰的具体发生位置和影响因素仍然未知。可大胆猜想一下, 倘若是尿液的特殊环境使得其中的蛋白质发生[Cys→Dha]修饰从而破坏了二硫键, 更可能的是破坏该蛋白中所有的二硫键而不是选择性的去除一部分。完成上述工作后, 本研究还通过使用相同的非限制性鉴定搜索分析了部分小鼠与大鼠尿液蛋白质组的相关实验数据。结果发现, [Cys→Dha]修饰不受物种的限制, 可以在小型啮齿动物的尿液中发现, 其数量和百分比均高于人尿液。另一种值得关注的修饰是赖氨酸的氨甲酰化修饰, 氨甲酰化是一种不可逆的非酶修饰过程, 其过程是尿素的分解产物与蛋白质的N末端或赖氨酸残基的侧链反应, 并与蛋白质老化有关^[24]。赖氨酸氨基甲酰化可以促进金属离子对特定酶活性的配位作用^[25,26]。据报道, 尿素水平升高的患者(例如肾病患者)的血浆中氨甲酰化的量显著增加^[27]。具体血浆中该修饰数量高于尿液的原因, 目前还不得而知。

不仅存在[Cys→Dha]修饰的影响, 有报道表明许多疾病是由修饰水平变化引起的^[28]。在帕金森综合征中, 有些蛋白质通过各种修饰以及与其他细胞成分的相互作用来影响神经元的正常功能^[29]。Regan等人^[30]发现, 糖基化是另一种复杂的修饰类型, 是蛋白质的一种常见且重要的修饰, 与多种疾病状态有关, 可导致复杂的聚糖被整合到生物分子中, 如阿尔兹海默症的异常糖基。因此, 蛋白质修饰的类型和频率也许可以作为诊断中的潜在生物标志物。

“尿液中为什么会有蛋白质”这一问题一直是人们关心的问题。蛋白质分子修饰的改变, 可能影响了蛋白质的电荷、形状, 进而影响了肾脏对其滤过和(或)重吸收, 从而达到使其从血液排到尿液的结果。而实际机体的生理情况或许更加复杂。本研究所提出的猜想或许只是众多可能性中的一种, 但从蛋白质组整体分子修饰角度出发, 在扩大蛋白质组学应用研究的同时, 也为后续继续深入研究这一科学问题提供了参考。

参考文献

- 1 Zhao M, Li M, Yang Y, et al. A comprehensive analysis and annotation of human normal urinary proteome. *Sci Rep*, 2017, 7: 3024
- 2 Yuan Y, Gao Y. Urinary protein changes in a rat starvation model. *MOJ Proteom Bioinform*, 2017, 6
- 3 Khoury G A, Baliban R C, Floudas C A. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-

- prot database. *Sci Rep*, 2011, 1: 90
- 4 Sambataro F, Pennuto M. Post-translational modifications and protein quality control in motor neuron and polyglutamine diseases. *Front Mol Neurosci*, 2017, 10: 82
 - 5 Wilhelm M, Schlegl J, Hahne H, et al. Mass-spectrometry-based draft of the human proteome. *Nature*, 2014, 509: 582–587
 - 6 Witze E S, Old W M, Resing K A, et al. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods*, 2007, 4: 798–806
 - 7 Dalle-Donne I, Aldini G, Carini M, et al. Protein carbonylation, cellular dysfunction, and disease progression. *J Cell Mol Med*, 2006, 10: 389–406
 - 8 Gianazza E, Crawford J, Miller I. Detecting oxidative post-translational modifications in proteins. *Amino Acids*, 2007, 33: 51–56
 - 9 Creasy D M, Cottrell J S. Unimod: Protein modifications for mass spectrometry. *Proteomics*, 2004, 4: 1534–1536
 - 10 Liu L, Liu X, Sun W, et al. Unrestrictive identification of post-translational modifications in the urine proteome without enrichment. *Proteome Sci*, 2013, 11: 1
 - 11 Qin W, Wang T, Huang H, et al. Profiling of lysine-acetylated proteins in human urine. *Sci China Life Sci*, 2019, 62: 1514–1520
 - 12 Chi H, He K, Yang B, et al. pFind-Alioth: A novel unrestricted database search algorithm to improve the interpretation of high-resolution MS/MS data. *J Proteomics*, 2015, 125: 89–97
 - 13 Wang L H, Li D Q, Fu Y, et al. Pfind 2.0: A software package for peptide and protein identification via tandem mass spectrometry. *Rapid Commun Mass Spectrom*, 2007, 21: 2985–2991
 - 14 Li D, Fu Y, Sun R, et al. Pfind: A novel database-searching software system for automated peptide and protein identification via tandem mass spectrometry. *Bioinformatics*, 2005, 21: 3049–3050
 - 15 Wiśniewski J R, Zougman A, Nagaraj N, et al. Universal sample preparation method for proteome analysis. *Nat Methods*, 2009, 6: 359–362
 - 16 Paris G, Kraszewski S, Ramseyer C, et al. About the structural role of disulfide bridges in serum albumins: Evidence from protein simulated unfolding. *Biopolymers*, 2012, 97: 889–898
 - 17 Liu H, May K. Disulfide bond structures of IgG molecules: Structural variations, chemical modifications and possible impacts to stability and biological function. *mAbs*, 2012, 4: 17–23
 - 18 Nagumo K, Tanaka M, Chuang V T G, et al. Cys34-cysteinylated human serum albumin is a sensitive plasma marker in oxidative stress-related chronic diseases. *PLoS ONE*, 2014, 9: e85216
 - 19 Jensen K S, Hansen R E, Winther J R. Kinetic and thermodynamic aspects of cellular thio-disulfide redox regulation. *Antioxid Redox Signal*, 2009, 11: 1047–1058
 - 20 Mahmood D F D, Abderrazak A, El Hadri K, et al. The thioredoxin system as a therapeutic target in human health and disease. *Antioxid Redox Signal*, 2013, 19: 1266–1303
 - 21 Bechtel T J, Weerapana E. From structure to redox: The diverse functional roles of disulfides and implications in disease. *Proteomics*, 2017, 17: 1600391
 - 22 Friedman M. Chemistry, biochemistry, nutrition, and microbiology of lysinoalanine, lanthionine, and histidinoalanine in food and other proteins. *J Agric Food Chem*, 1999, 47: 1295–1319
 - 23 Siodlak D. α,β -Dehydroamino acids in naturally occurring peptides. *Amino Acids*, 2015, 47: 1–17
 - 24 Kollipara L, Zahedi R P. Protein carbamylation: *In vivo* modification or in vitro artefact? *Proteomics*, 2013, 13: 941–944
 - 25 Badar A, Arif Z, Alam K. Role of carbamylated biomolecules in human diseases. *IUBMB Life*, 2018, 70: 267–275
 - 26 Hsieh Y C, Chen M C, Hsu C C, et al. Crystal structures of vertebrate dihydropyrimidinase and complexes from *Tetraodon nigroviridis* with lysine carbamylation: Metal and structural requirements for post-translational modification and function. *J Biol Chem*, 2013, 288: 30645–30658
 - 27 Shi J, van Veelen P A, Mahler M, et al. Carbamylation and antibodies against carbamylated proteins in autoimmunity and other pathologies. *Autoimmun Rev*, 2014, 13: 225–230
 - 28 Mnatsakanyan R, Shema G, Basik M, et al. Detecting post-translational modification signatures as potential biomarkers in clinical mass spectrometry. *Expert Rev Proteomics*, 2018, 15: 515–535
 - 29 Dixit A, Mehta R, Singh A K. Proteomics in human Parkinson's disease: Present scenario and future directions. *Cell Mol Neurobiol*, 2019, 39: 901–915
 - 30 Regan P, McClean P L, Smyth T, et al. Early stage glycosylation biomarkers in Alzheimer's disease. *Medicines (Basel)*, 2019, 6: 92

Why are there proteins in the urine of healthy people?

LIU YongTao¹, ZHAO MinDi², PAN XuanZhen¹ & GAO YouHe¹

1 Beijing Key Laboratory of Gene Engineering Drug and Biotechnology, Department of Biochemistry and Molecular Biology, College of Life Science, Beijing Normal University, Beijing 100875, China;

2 Department of Laboratory Medicine, National Geriatrics Center, Beijing Hospital, Beijing 100730, China

Proteins are essential nutrients for humans, but why is there a small amount of proteins excreted through urine? Although healthy people have low levels of protein in the urine, they are rich in variety. The proteins in them may be unnecessary wastes of the body or may even have a toxic effect on the body and must be excreted. Compared with small-molecule metabolites, proteins have a complex structure and various functions. Even a small change at the molecular level will affect their subsequent biological functions. A comprehensive comparison of the molecular level modifications of plasma and urine proteome, the difference may provide clues to the presence of protein in the urine. In this study, a total of 9 healthy people's urine and 9 healthy people's plasma samples were collected. The samples were analyzed by high-resolution tandem mass spectrometry and unlabeled quantitative proteomics techniques, and the non-limiting modification identification algorithm was used to compare molecular modifications between two types of samples as a whole. The results showed that the amount of cysteine (Cys) modification to dehydroalanine (Dha) in the urine proteome was higher than that in plasma. The molecular modification [Cys→Dha] destroys the disulfide bond of the original protein, thereby affecting or even changing the structure and biological function of the protein. Therefore, this study revealed the differences in the overall proteome modification of plasma and urine and pointed out that the structure of protein may be irreversibly changed due to the difference in modification of protein molecules, which in turn caused proteins which lost their functions and even toxic proteins from the blood to be excreted into the urine.

urine, plasma, proteome, molecular modification, non-limiting modification identification

doi: [10.1360/SSV-2019-0235](https://doi.org/10.1360/SSV-2019-0235)