

基于富含硼酸金属有机框架的痕量样本蛋白组学前处理技术

张爽, 丁显廷*

上海交通大学生物医学工程学院, 个性化医学研究院, 上海 200030

* 联系人, E-mail: dingxianting@sjtu.edu.cn

2024-10-29 收稿, 2025-01-17 修回, 2025-02-23 接受, 2025-02-28 网络版发表

国家重点研发计划(2022YFF0710200, 2022YFC2601700)、国家自然科学基金(T2122002, 22077079)、上海市“科技创新行动计划”国际科技合作项目(22Z510202478)、上海市教育委员会/上海市教育发展基金会曙光计划(21SG10)、上海交通大学项目(YG2021ZD19, Agri-X20200101, 2020 SJTU-HUJI)和上海市卫生健康委员会项目(2019CXJQ03)资助

摘要 在单细胞和微量组织样本的蛋白质组学研究中, 样本前处理技术的改进对于提高分析的灵敏度和可靠性至关重要. 本文开发了一种富含硼酸的镧系金属有机框架(MOF)纳米材料, 该材料通过氮-硼酸络合和阳离子- π 相互作用高效捕获样品中的蛋白质. 我们进一步建立了基于MOF材料的样品制备标准化工作流程(MASP), 该流程缩小了样本体积, 简化了样本前处理过程, 并集成了细胞裂解、蛋白质捕获、消化和肽段收集, 尽可能减少样本损耗. 通过MASP技术, 我们在10个HEK-293T细胞中成功定量1885个蛋白质. 这一成果不仅展示了MASP在提高低丰度蛋白质样品分析深度上的潜力, 也为单细胞和空间蛋白质组学研究提供一种高效、可靠的样本前处理方案, 具有重要的临床应用前景.

关键词 金属有机框架, 蛋白质组学, 痕量样本, 蛋白吸附

传统的自下而上的蛋白质组学工作流程的技术路线是将样本中的蛋白质酶解为短肽. 该流程包括细胞和组织中蛋白质的提取、半胱氨酸的还原与烷基化、蛋白质消化、肽段净化和浓缩, 最后用于液相色谱串联质谱的数据采集^[1-4]. 在空间蛋白质组学的研究中, 常常通过激光显微切割组织切片以获得空间信息. 切割后的组织片与单细胞样本都是超痕量的生物样本. 在对这类样本进行前处理的过程中, 存在潜在的样品污染和吸附损失, 导致质谱结果的重现性差, 同时增加了时间成本和投资成本, 这限制了生物医学科学家进行基于质谱的痕量样本蛋白质组学实验^[2,5-11].

为了应对上述问题, 目前已开发出多种纳米颗粒, 通过纳米-生物相互作用捕获样品中的蛋白质, 以进行全蛋白质组学研究^[12-20]. 设计用于富集单细胞和组织

切片中微量蛋白质的功能材料十分重要, 也是一个持续的挑战^[21].

蛋白质表面通常存在阳离子基团和阴离子羧酸盐基团, 因此, 开发能够结合蛋白表面阴离子和阳离子基团的材料可有效结合蛋白质. 苯硼酸能够通过氮-硼酸配位与富电子氨基酸残基结合, 其中的芳香环可以通过阳离子- π 相互作用与蛋白质上的胍基发生相互作用^[22-24]. 蛋白质上的阴离子羧酸可以通过配位与金属阳离子相结合^[25]. 金属有机框架(metal-organic framework, MOF)由配体和金属自组装形成, 通常带正电, 能够与蛋白质的游离羧基发生配位. 因此, 开发新型MOF以捕获全蛋白质, 并分析痕量样品中的全蛋白质组学具有重要意义.

本研究采用3,5-二羧基苯基硼酸作为配体, 以镧系

引用格式: 张爽, 丁显廷. 基于富含硼酸金属有机框架的痕量样本蛋白组学前处理技术. 科学通报, 2026, 71: 69–77

Zhang S, Ding X. Trace sample proteomics pre-treatment technology based on boric acid metal-organic frameworks (in Chinese). Chin Sci Bull, 2026, 71: 69–77, doi: [10.1360/TB-2024-1132](https://doi.org/10.1360/TB-2024-1132)

金属盐作为原料,合成了一类富含硼酸的铜系MOFs.在此基础上,进一步开发了MOFs辅助样品制备(MOFs aided sample preparation, MASP)流程.如图1所示,该流程首先将MOF添加至生物样本中以捕获蛋白质,随后通过离心去除上清液,并洗涤MOF表面的蛋白质.接着,将蛋白质酶解为肽段,脱盐后进行质谱数据采集.

MASP简化了蛋白样本前处理步骤,将裂解/冲洗/消化集成在一个PCR管中,减少了由于表面吸附导致的样本损失,能提升超痕量样本中的蛋白组学鉴定深度.本文在10~500个HEK 293T细胞中测试了MASP,并验证了MASP在处理痕量细胞方面的可行性和效率.MASP提供了一种通用的新方法,能在痕量细胞样本中进行深入的蛋白质组分析.

1 实验

1.1 材料制备

本文通过溶剂热法制备球形的富含硼酸的金属有机框架.将0.072 mmol $\text{DyCl}_3 \cdot 6\text{H}_2\text{O}$ 、0.036 mmol间苯二甲酸和0.036 mmol的3,5-二羧基苯基硼酸加入36 mL N,N -二甲基甲酰胺(DMF)/ H_2O (体积比7:3)的混合溶液中,超声10 min,室温搅拌2 h,后在120℃下反应12 h.20000 g离心10 min后收集产物,然后用DMF、无水乙醇和去离子水分别洗涤3次.将洗涤后的MOF在去离子水中透析1天. Eu-MOF、Gd-MOF和Tb-MOF的合成分别以 $\text{EuCl}_3 \cdot 6\text{H}_2\text{O}$ 、 $\text{GdCl}_3 \cdot 6\text{H}_2\text{O}$ 和 $\text{TbCl}_3 \cdot 6\text{H}_2\text{O}$ 作为金属源,其他步骤与Dy-MOF相同.4种MOF经冷冻干燥后,保存在干燥锅中备用.

1.2 材料表征

使用扫描电子显微镜(SEM)观察4种MOF的形态特征.在室温下对4种MOF进行动态光散射(DLS)测试,评估流体动力学尺寸和表面电位,结果为同一样品十

次测量的平均值和三次重复测量的标准差.使用X射线光电子能谱(XPS)分析4种MOF表面元素组成.使用傅里叶变换红外光谱(FT-IR)检测4000~400 cm^{-1} 范围内功能团的吸收峰.

1.3 蛋白组学样本制备

将5 μL 细胞裂解缓冲液(0.1% N -十二烷基 β -D-麦芽糖苷、1 mmol/L TECP和2 mmol/L CAA溶于50 mmol/L NH_4HCO_3)转移至PCR管中.将HEK 293T细胞在PBS中稀释至500、100、50和10个细胞/ μL 的浓度.取1 μL 细胞悬浮液并加入PCR管中的5 μL 细胞裂解缓冲液中.将混合物在95℃下加热30 min,超声处理1 h,再将样品在60℃下加热1 h进行还原烷基化.将4 μL MOF悬浮液(2.5 mg/mL)加入细胞裂解物.将混合物在24℃下以1000 r/min的速度振荡10 min,使蛋白质充分吸附在MOF上.然后,将混合物在20000 g下离心10 min,弃去上清液.随后用20 μL 的50 mmol/L NH_4HCO_3 洗涤MOF表面的蛋白质3次.随后,重悬MOF/蛋白质混合物,以酶/蛋白质质量比为1:20的比例加入胰蛋白酶,在37℃下酶解过夜,将MOF表面的蛋白质消化成肽.最后加入终浓度为1%的甲酸以终止酶解反应并将洗脱肽段.采用Ziptip(Millipore)进行脱盐处理,纯化的肽段在真空浓缩仪中干燥,储存在-80℃冰箱.

1.4 原始数据采集

使用0.1%的甲酸将肽段重悬,在timsTOF Flex (Bruker)上进行分析.重悬的肽段在内径75 μm 的Aurora Elite CSI第三代15 cm HPLC柱(AUR3-15075C18-CSIonOpticks, Australia)上以以下梯度程序进行洗脱:在75 min内,B相溶液从2%升高到22%,随后在5 min内从22%升高到37%,然后在5 min内增加到80%,最后在80%的B相溶液中平衡5 min.其中A相缓冲液为体积分

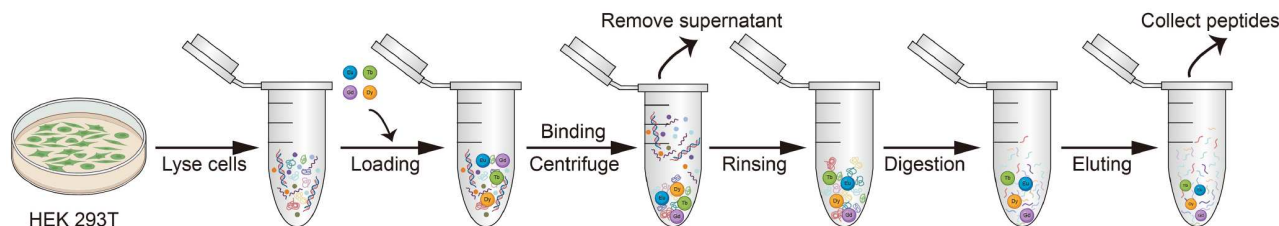


图1 MASP处理生物样本的流程图

Figure 1 MASP workflow for handling biological samples

数0.1%的甲酸水溶液, B相缓冲液为体积分数0.1%的甲酸乙腈溶液.

1.5 原始数据分析

对于采用DDA采集方法获取的原始文件, 我们使用MaxQuant(版本号1.6.17.0), 在human protein database(UP000005640, 9606)中进行检索, 对蛋白质进行鉴定和定量, 具体检索参数设置如下: 搜索6~144个氨基酸长度的肽, 酶特异性设定为胰蛋白酶将C端切割为精氨酸和赖氨酸, 最多允许两次胰蛋白酶漏切, 蛋氨酸氧化和蛋白N端乙酰化被设定为可变修饰, 半胱氨酸氨基甲基化被设定为固定修饰. 在肽和蛋白质水平上, 错误发现率(FDR)设置为1%.

对于采用DIA采集方法获取的原始文件, 我们使用了Data-Independent Acquisition by Neural Networks (DIANN, 版本号1.8.1)的开源软件用于dDIA的搜库. 按照文献中对DIANN检索策略的说明, 采用默认参数, 在Human Protein Database(UP000005640, 9606)中检索原始文件.

2 结果与讨论

2.1 材料表征

本文以Eu、Gd、Tb、Dy为金属节点, 以间苯二甲酸和3,5-二羧基苯基硼酸配体为原料, 通过溶剂热法制备了四种富含硼酸的镧系金属有机框架. 采用扫描电子显微镜对四种MOF进行了形态学表征, 如图2(a)~(d)所示, 四种MOF均呈球形. 进一步使用TEM对四种MOF进行表征, 如图2(e)~(g)所示, 四种球形MOF的粒径分布较均匀.

采用XPS技术对四种MOF表面的元素组成进行了表征. 如图2(i)所示, 在四种MOF中均观察到了B 1s峰, 表明目标B-O基团成功引入至MOF中. 此外, 在每种MOF中都检测到了相应的镧系元素, 进一步证明了富含硼酸的镧系金属有机框架的成功制备.

本文进一步使用DLS表征了四种MOF的动力学尺寸. 图3(a)结果显示, Eu-MOF、Gd-MOF、Tb-MOF和Dy-MOF的水合粒径分别为845、817、769和803 nm, 四种MOF的粒径接近. 此外, 本文还对间苯二甲酸、

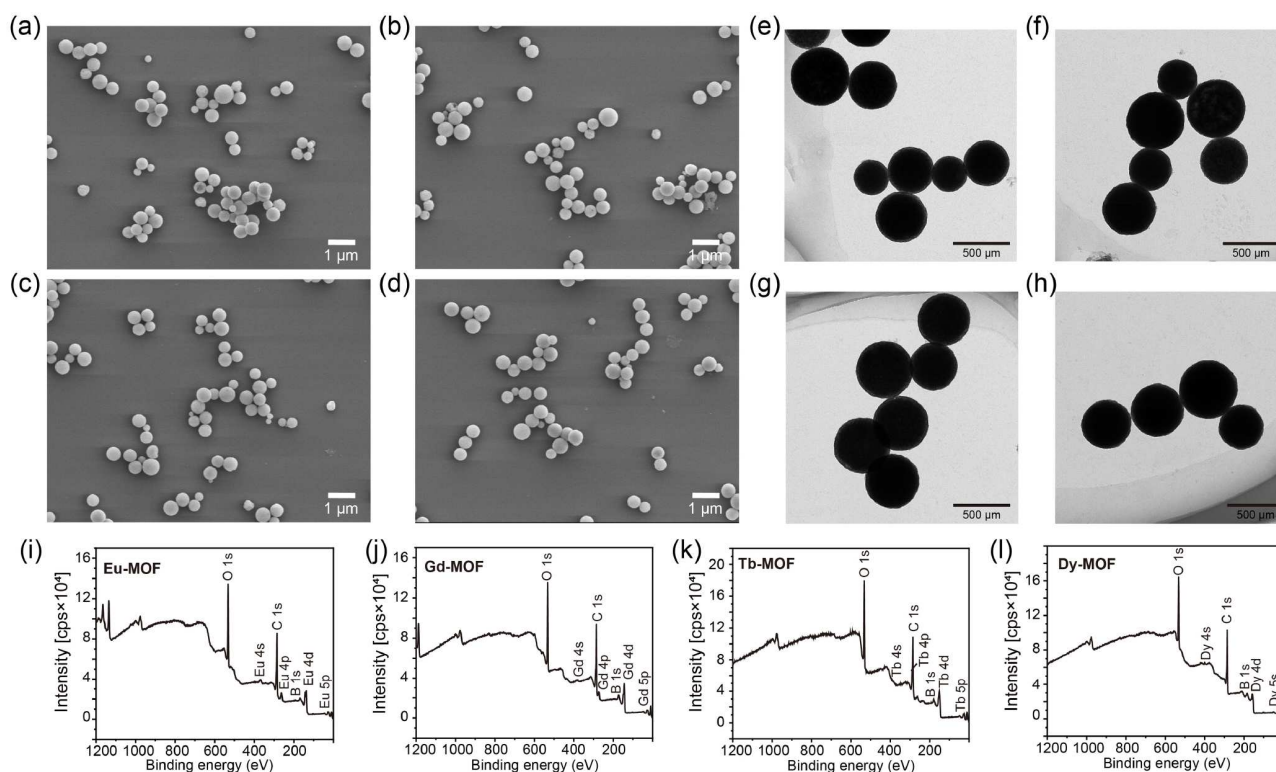


图 2 四种MOF的材料表征. Eu-MOF(a)、Gd-MOF(b)、Tb-MOF(c)和Dy-MOF(d)的SEM图; Eu-MOF(e)、Gd-MOF(f)、Tb-MOF(g)和Dy-MOF(h)的TEM图; Eu-MOF(i)、Gd-MOF(j)、Tb-MOF(k)和Dy-MOF(l)的XPS谱图

Figure 2 Material characterization of four MOFs. SEM images of (a) Eu-MOF, (b) Gd-MOF, (c) Tb-MOF, and (d) Dy-MOF. TEM images of (e) Eu-MOF, (f) Gd-MOF, (g) Tb-MOF, and (h) Dy-MOF. XPS spectra of (i) Eu-MOF, (j) Gd-MOF, (k) Tb-MOF, and (l) Dy-MOF

3,5-二羧基苯基硼酸这两种配体和4种MOF进行了傅里叶变换红外光谱分析. 如图3(b)所示, 在配体中观察到 1679 cm^{-1} 处的C=O伸缩振动峰, 但该峰并未在四种MOF中出现, 这证明了羧酸根与镧系金属离子的成功配位^[26]. 此外, 在3,5-二羧基苯基硼酸配体和MOF中均观察到 1310 cm^{-1} 处的吸收峰, 该特征峰源自B-O, 说明MOF中含有硼酸基团, 并且硼酸基团未与金属离子配位, 进一步证明了MOF的成功制备^[27].

本文对四种MOF的Zeta电位进行了表征, 如图3(c)所示, Eu-MOF、Gd-MOF、Tb-MOF和Dy-MOF的Zeta电位分别为26.8、19.7、13.0和7.5 mV, 表明它们具有不同的表面物理化学性质.

本文通过简单有效的一步溶剂热法设计并合成了四种富含硼酸的镧系金属有机框架, 该方法能够同时合成这四种MOF, 从而大大缩短了合成时间. 此外, 本研究通过SEM、TEM、DLS和FT-IR等表征手段确认了本合成路线的可行性.

2.2 MOF的蛋白质吸附表征

我们将这四种MOF分别与异硫氰酸荧光素-牛血清白蛋白(FTIC-BSA)共孵育, 使用共聚焦显微镜对荧光蛋白和MOF材料进行定位, 以探究这四种MOF对蛋白的吸附能力. 如图4(a)~(d)所示, 上部分为MOF材料的明场图像, 中间为FTIC-BSA的荧光分布, 下部分为FTIC-BSA的荧光分布与MOF明场图像的共定位. 可以观察到, 在MOF/蛋白混合物中, MOF的位置与FTIC-BSA的荧光信号分布一致, 这说明MOF表面成功捕获了FTIC-BSA.

将这四种富含硼酸的MOF等质量混合, 称为mix-MOF, 并将mix-MOF与FTIC-BSA共孵育. 如图4(e)所

示, 在激光共聚焦显微镜下观察到MOF材料与荧光的共定位, 进一步验证了MOF对模型蛋白的富集能力.

2.3 mix-MOF的蛋白组学鉴定表现

在质谱分析之前, 分别使用四种单一MOF和mix-MOF处理100个HEK 293T细胞, 以评估MOF辅助样品制备(MASP)工作流程在蛋白质组学分析中的实用性. 本实验优化和调整了单次质谱采集的参数, 包括分辨率和LC-MS/MS梯度, 以提高蛋白质组学分析的覆盖范围.

将四种单一MOF和mix-MOF加入细胞裂解液中进行孵育, 以捕获细胞裂解液中的蛋白质. 在酶解MOF上的蛋白质后, 加入1%甲酸(FA)来破坏MOF并洗脱肽段, 以用于后续的质谱采集. 如图5(a)所示, mix-MOF在100个HEK 293T细胞中平均定量了5788种蛋白质, Eu-MOF、Gd-MOF、Tb-MOF和Dy-MOF分别平均定量了5581、5606、5523和5544种蛋白质. mix-MOF显示出比单一MOF更深的蛋白质组学覆盖深度. 虽然这四种MOF拥有相同的有机配体, 但不同的金属节点赋予了它们不同的物理化学性质, 因此, 四种MOF表面形成的蛋白冠组成也存在差异. mix-MOF结合了四种MOF上形成的所有蛋白冠, 从而有助于增强其蛋白质组覆盖率. 如图5(b)所示, 在韦恩图中可以观察到mix-MOF与单一MOF鉴定出的蛋白质种类存在大量重叠, 这表明mix-MOF中鉴定的蛋白质主要来源于四个单一MOF鉴定的蛋白质的集合.

2.4 MASP蛋白鉴定性能表征

本文设计的MOF通过物理化学作用, 能够在液体环境中自发富集蛋白质. 本研究在此基础上提出了一

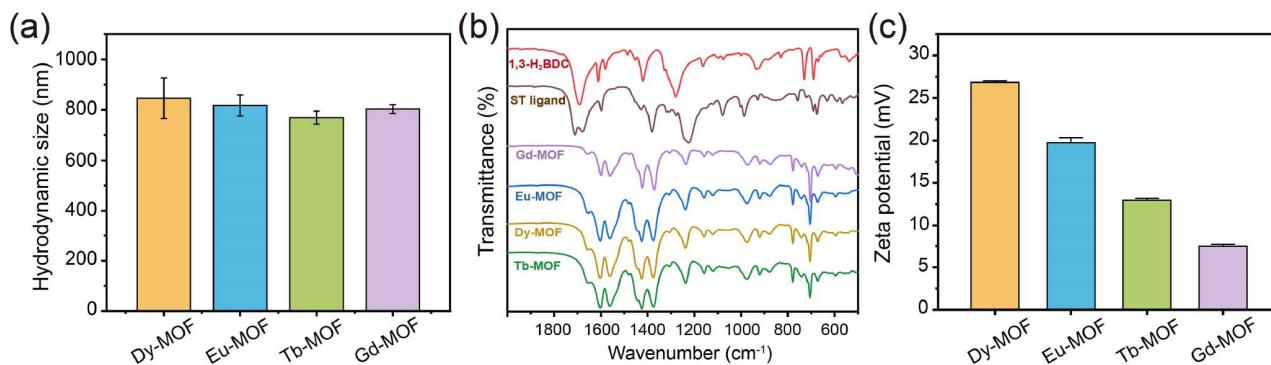


图3 四种MOF的物理化学性质表征. (a) 四种MOF的水合粒径; (b) 四种MOF和配体的红外光谱; (c) 四种MOF的zeta电位

Figure 3 Physicochemical characterization of four MOFs. (a) Hydrodynamic size of four MOFs; (b) FT-IR spectrum of ligand and the four MOFs; (c) zeta potentials of four MOFs

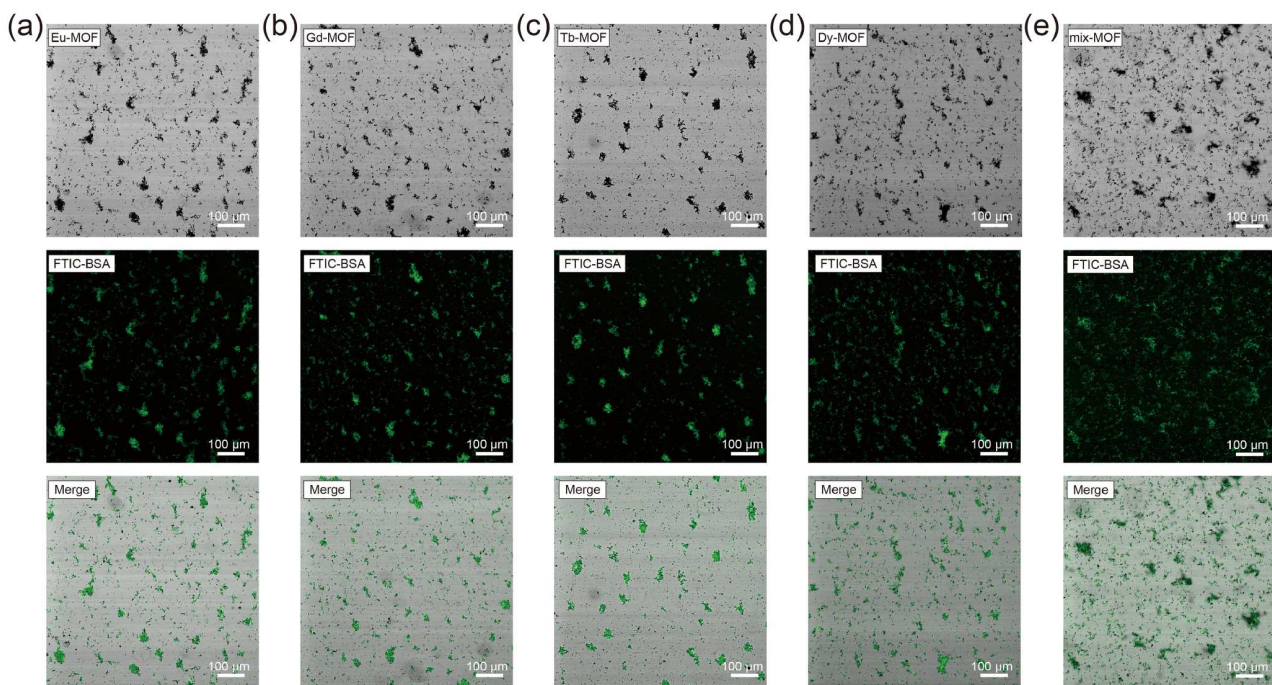


图 4 Eu-MOF(a)、Gd-MOF(b)、Tb-MOF(c)、Dy-MOF(d)和mix-MOF(e)吸附FITC-BSA后的荧光位置

Figure 4 The fluorescence after adsorption of FITC-BSA on Eu-MOF (a), Gd-MOF (b), Tb-MOF (c), Dy-MOF (d), and mix-MOF (e)

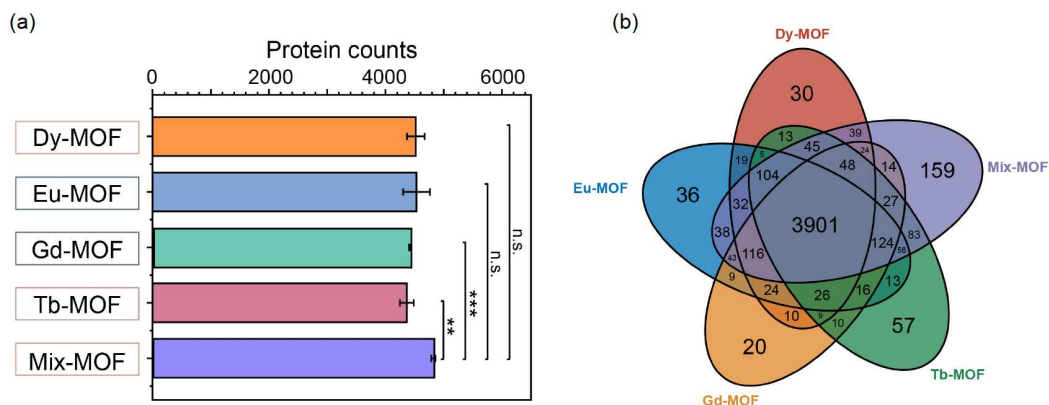


图 5 单一MOF和mix-MOF的蛋白组学鉴定表现. (a) 用DDA方法测定100个HEK 293T细胞中定量蛋白质的数量, 数据以平均值 $\pm$ SD表示($n=3$ 个独立技术重复), ***, $P<0.001$, **, $0.01<P<0.001$, *, $0.01<P<0.05$, n.s. $P>0.05$; (b) 单一MOF和mix-MOF量定的蛋白质重叠的维恩图

Figure 5 Proteomic identification performance of single MOF and mix-MOFs. (a) The number of protein groups quantified in 100 HEK 293T cells was determined by DDA mode. Mean $\pm$ SD ($n=3$ independent technical replicates), ***, $P<0.001$, **, $0.01<P<0.001$, *, $0.01<P<0.05$, n.s. $P>0.05$. (b) Venn diagram depicting the overlap of the proteins quantified by one single MOF and mix-MOFs

种MOF辅助的蛋白质样本制备新方法MASP, 如图1所示. 该方法在细胞裂解液中添加MOF, 以实现无偏蛋白质捕获, 接着通过离心去除杂质. 随后, 使用胰蛋白酶消化MOF表面结合的蛋白质. 消化完成后, 加入终浓度为1%的甲酸以破坏MOF的外部结构, 并洗脱肽段. 脱

盐后, 将肽段重悬于0.1%的甲酸中, 进行质谱分析.

进一步使用数据非依赖型采集方式(data-independent acquisition, DIA)来测试MASP定量的有效性. 我们使用dDIA模式系统地研究了MASP的蛋白质组学鉴定性能, 包括灵敏度、蛋白质组覆盖率和可重复性.

使用MASP在10、25、50、100和500个HEK 293T细胞中进行样本前处理流程,以评估MASP在不同蛋白丰度的生物样本中的蛋白质和肽段鉴定深度.如图4(a)和(b)所示,随着起始细胞量的增加,鉴定到的蛋白质和肽段的数量上升. MASP在10个HEK 293T细胞中可重复定量1885种蛋白质,在500个细胞中可重复定量7573种蛋白质.随着细胞数量的增加,变异系数呈下降趋势(图6(b)).而相关系数呈上升趋势(图6(c)).与10个细胞相比,500个细胞的蛋白质动态范围扩大了1.5个数量级(图6(d))

为了表征独立实验组之间的检测可重复性,本研究使用MASP对100个HEK 293T细胞进行了10次独立的技术重复实验.如图6(e)所示,10次重复实验中蛋白质丰度的分布高度一致. MASP在100个HEK 293T细胞中平均鉴定出6149种蛋白质,对应47508个独特肽段.

3 结论

本文以Eu、Gd、Tb、Dy为金属节点,以间苯二甲酸和3,5-二羧基苯基硼酸配体为原料,过溶剂热法制备

了一类富硼酸的铜系金属有机框架,能够通过氮-硼酸盐络合、阳离子- π 相互作用以及离子相互作用等去结合蛋白质.本文进一步提出了一种蛋白质组学样本前处理工作流程,即MOF辅助样品制备流程MASP,用于超痕量细胞样本的深度蛋白质组学解析. MASP在10和500个HEK-293T细胞中分别定量了1885和7573个蛋白,并显示出良好的可重复性和低变异性.在痕量细胞尺度的蛋白组学分析中,与已开发的auto-POTS和SP3等技术相比, MASP展现了相当的鉴定深度,显示出了在单细胞和空间蛋白质组学样本前处理方面的应用潜力.

随着自动化仪器的发展和质谱仪的更新,目前已开发出各种工作流程和设备来用于处理单细胞和微量组织切片样品^[28-33].其中,一锅法样本制备将样品裂解、蛋白质的还原烷基化和酶解等步骤整合,减少了样品转移环节,以最大限度地降低表面吸附造成的损失,来提高微量样品的蛋白质鉴定深度^[34-38].然而,一锅法制备的样品不可避免地含有核酸、盐、脂质和细胞碎片,如果将这些样品直接注入质谱仪会影响仪器

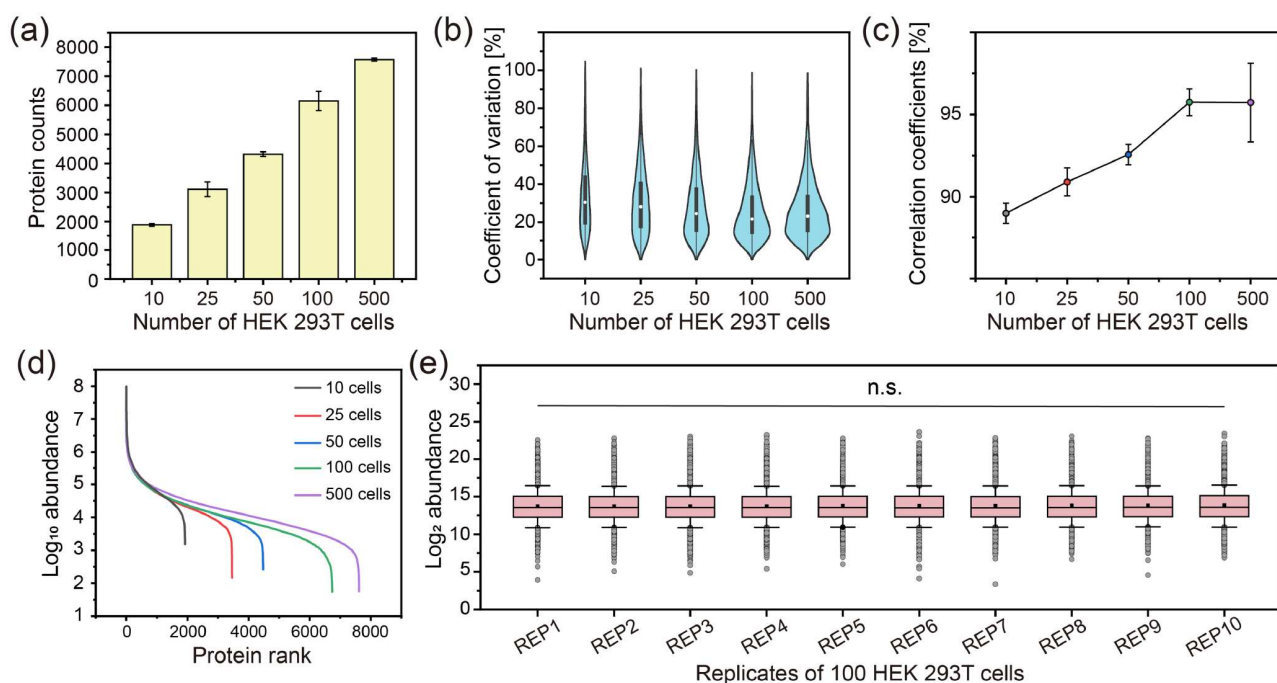


图 6 MASP的蛋白组学性能评估. (a) MASP在10~500个HEK 293T细胞中定量蛋白的个数($n=5$ 个技术重复); (b) MASP在10~500个HEK 293T细胞中定量蛋白的变异系数; (c) MASP在10~500个HEK 293T细胞中定量蛋白的相关系数; (d) MASP在10~500个HEK 293T细胞中定量蛋白的动态范围评估; (e) MASP在100个HEK 293T细胞中进行10次独立实验定量的总蛋白丰度分布

Figure 6 Proteomic performance of MASP. (a) The number of quantified proteins across 10–500 HEK 293T cells using MASP. (b) The coefficient of variation of proteins obtained across 10–500 HEK 293T cells by MASP. (c) The correlation coefficients of proteins obtained across 10–500 HEK 293T cells by MASP. (d) Assessment of dynamic range based on protein abundance rank of the whole-cell protein across 10–500 HEK 293T cells using MASP. (e) Distributions of total protein abundances in 10 independent MASP tests of 100 HEK 293T cells

的检测限和色谱柱的寿命^[39]。本研究开发的MASP能够有效富集微量样本中的蛋白质，同时去除样本中的其他杂质，最终生成适合质谱分析的纯净肽段。这一特点令MASP方法在质谱分析中表现良好，适用于多种研究场景。此外，该方法的实施成本较低，不需要购置商业化的机器人移液平台或自动进样器，显著降低了实验费用，使其在许多常规实验室中更适合进行超痕量

样本的无标记蛋白质组学分析。

MASP是一种通用的、可扩展的和高度可重复的蛋白质组学工作流程，适用于低输入样品的应用场景。MASP能够让研究人员以快速、稳健和精简的方式分析超痕量细胞，有望推动单细胞蛋白质组学和空间蛋白质组学的应用。这一方法有助于在各种疾病的生物标志物临床发现中，提供更好的分类、诊断和治疗方案。

致谢 感谢上海交通大学先进电子材料与器件平台和实验动物中心的支持。

参考文献

- 1 Yates III J R. A century of mass spectrometry: from atoms to proteomes. *Nat Methods*, 2011, 8: 633–637
- 2 Mund A, Coscia F, Kriston A, et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat Biotechnol*, 2022, 40: 1231–1240
- 3 Bennett H M, Stephenson W, Rose C M, et al. Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nat Methods*, 2023, 20: 363–374
- 4 Alfaro J A, Bohländer P, Dai M, et al. The emerging landscape of single-molecule protein sequencing technologies. *Nat Methods*, 2021, 18: 604–617
- 5 Perkel J M. Single-cell proteomics takes centre stage. *Nature*, 2021, 597: 580–582
- 6 Kulak N A, Pichler G, Paron I, et al. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods*, 2014, 11: 319–324
- 7 Rosenberger F A, Thielert M, Mann M. Making single-cell proteomics biologically relevant. *Nat Methods*, 2023, 20: 320–323
- 8 Zhu Y, Aebersold R, Mann M, et al. SnapShot: clinical proteomics. *Cell*, 2021, 184: 4840–4840.e1
- 9 Rosenberger F A, Thielert M, Strauss M T, et al. Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. *Nat Methods*, 2023, 20: 1530–1536
- 10 Nordmann T M, Anderton H, Hasegawa A, et al. Spatial proteomics identifies JAKi as treatment for a lethal skin disease. *Nature*, 2024, 635: 1001–1009
- 11 Coscia F, Doll S, Bech J M, et al. A streamlined mass spectrometry-based proteomics workflow for large-scale FFPE tissue analysis. *J Pathol*, 2020, 251: 100–112
- 12 Hughes C S, Moggridge S, Müller T, et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc*, 2019, 14: 68–85
- 13 Virant-Klun I, Leicht S, Hughes C, et al. Identification of maturation-specific proteins by single-cell proteomics of human oocytes. *Mol Cell Proteomics*, 2016, 15: 2616–2627
- 14 Blume J E, Manning W C, Troiano G, et al. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat Commun*, 2020, 11: 3662
- 15 Ferdosi S, Stukalov A, Hasan M, et al. Enhanced competition at the nano-bio interface enables comprehensive characterization of protein corona dynamics and deep coverage of proteomes. *Adv Mater*, 2022, 34: e2206008
- 16 Corbo C, Li A A, Poustchi H, et al. Analysis of the human plasma proteome using multi-nanoparticle protein corona for detection of Alzheimer's disease. *Adv Healthcare Mater*, 2021, 10: 2000948
- 17 Sadeghi S A, Ashkarran A A, Wang Q, et al. Mass spectrometry-based top-down proteomics in nanomedicine: proteoform-specific measurement of protein corona. *ACS Nano*, 2024, 18: 26024–26036
- 18 Wang M, Dai X, Yang X, et al. Serum protein fishing for machine learning-boosted diagnostic classification of small nodules of lung. *ACS Nano*, 2024, 18: 4038–4055
- 19 Hadjidemetriou M, Papaflippou L, Unwin R D, et al. Nano-scavengers for blood biomarker discovery in ovarian carcinoma. *Nano Today*, 2020, 34: 100901
- 20 Overbey E G, Kim J K, Tierney B T, et al. The Space Omics and Medical Atlas (SOMA) and international astronaut biobank. *Nature*, 2024, 632: 1145–1154

- 21 Sun N, Wu H, Shen X, et al. Proteomics: nanomaterials in proteomics (Adv. Funct. Mater. 26/2019). *Adv Funct Mater*, 2019, 29: 1970182
- 22 Liu C, Wan T, Wang H, et al. A boronic acid-rich dendrimer with robust and unprecedented efficiency for cytosolic protein delivery and CRISPR-Cas9 gene editing. *Sci Adv*, 2019, 5: eaaw8922
- 23 Tan E, Wan T, Yu C, et al. ROS-responsive polypeptides for intracellular protein delivery and CRISPR/Cas9 gene editing. *Nano Today*, 2022, 46: 101617
- 24 Kennedy C R, Lin S, Jacobsen E N. Die Kation- π -Wechselwirkung in der Katalyse mit niedermolekularen Verbindungen. *Angew Chem*, 2016, 128: 12784–12814
- 25 Ren L, Lv J, Wang H, et al. A coordinative dendrimer achieves excellent efficiency in cytosolic protein and peptide delivery. *Angew Chem Int Ed*, 2020, 59: 4711–4719
- 26 Wang H, Wang X, Liang M, et al. A boric acid-functionalized lanthanide metal-organic framework as a fluorescence “turn-on” probe for selective monitoring of Hg^{2+} and CH_3Hg^+ . *Anal Chem*, 2020, 92: 3366–3372
- 27 Li S, Qin Y, Zhong G, et al. Highly efficient separation of glycoprotein by dual-functional magnetic metal-organic framework with hydrophilicity and boronic acid affinity. *ACS Appl Mater Interfaces*, 2018, 10: 27612–27620
- 28 Croteck C, Clark N M, Boyle B W, et al. Automated single-cell proteomics providing sufficient proteome depth to study complex biology beyond cell type classifications. *Nat Commun*, 2024, 15: 5707
- 29 Ye Z, Sabatier P, Martin-Gonzalez J, et al. One-Tip enables comprehensive proteome coverage in minimal cells and single zygotes. *Nat Commun*, 2024, 15: 2474
- 30 Wang Y, Guan Z Y, Shi S W, et al. Pick-up single-cell proteomic analysis for quantifying up to 3000 proteins in a Mammalian cell. *Nat Commun*, 2024, 15: 1279
- 31 Chen W, Wang S, Adhikari S, et al. Simple and integrated spintip-based technology applied for deep proteome profiling. *Anal Chem*, 2016, 88: 4864–4871
- 32 Li Z Y, Huang M, Wang X K, et al. Nanoliter-scale oil-air-droplet chip-based single cell proteomic analysis. *Anal Chem*, 2018, 90: 5430–5438
- 33 Shao X, Wang X, Guan S, et al. Integrated proteome analysis device for fast single-cell protein profiling. *Anal Chem*, 2018, 90: 14003–14010
- 34 Leduc A, Khoury L, Cantlon J, et al. Massively parallel sample preparation for multiplexed single-cell proteomics using nPOP. *Nat Protoc*, 2024, 19: 3750–3776
- 35 Gebreyesus S T, Siyal A A, Kitata R B, et al. Streamlined single-cell proteomics by an integrated microfluidic chip and data-independent acquisition mass spectrometry. *Nat Commun*, 2022, 13: 37
- 36 Zhu Y, Piehowski P D, Zhao R, et al. Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. *Nat Commun*, 2018, 9: 882
- 37 Brunner A, Thielert M, Vasilopoulou C, et al. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol Syst Biol*, 2022, 18: e10798
- 38 Ye Z, Sabatier P, van der Hoeven L, et al. Enhanced sensitivity and scalability with a Chip-Tip workflow enables deep single-cell proteomics. *Nat Methods*, 2025, 22: 499–509
- 39 Matzinger M, Müller E, Dürnberger G, et al. Robust and easy-to-use one-pot workflow for label-free single-cell proteomics. *Anal Chem*, 2023, 95: 4435–4445

Summary for “基于富含硼酸金属有机框架的痕量样本蛋白组学前处理技术”

Trace sample proteomics pre-treatment technology based on boric acid metal-organic frameworks

Shuang Zhang & Xianting Ding^{*}

Institute for Personalized Medicine, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China

^{*} Corresponding author, E-mail: dingxianting@sjtu.edu.cn

Proteins are the most information-rich biomolecules which are essential for normal biochemical functions and implicated in various disease processes. Mass spectrometry (MS)-based label-free proteomics has emerged as a crucial technique for protein identification. In bottom-up proteomic analysis, the effective acquisition of data necessitates the digestion of proteins extracted from complex biological samples into pure peptides. The Filter Aided Sample Preparation (FASP) workflow is widely regarded as the gold standard for this process. However, conventional FASP workflow involves multiple redundant steps, such as repeated pipetting, rinsing, and ultrafiltration, which can result in substantial protein loss due to non-specific adsorption and contamination. This issue is particularly critical when dealing with limited and often non-renewable clinical samples, including rare cell subsets and small volumes of blood or cerebrospinal fluid.

To address these challenges, various nanoparticles have been developed to capture proteins through nano-bio interactions in clinical samples before mass spectrometry (MS) acquisition for comprehensive proteomic analysis. The utilization of surface-functionalized magnetic nanoparticles, in conjunction with robotic nanoliter liquid handling systems, enables rapid and high-throughput proteomic analysis. Despite these advancements, the manufacturing of compatible automated instruments remains a significant barrier to accessibility. The design of functional materials for the enrichment of trace proteins in single cells and tissue sections is of great importance and remains a persistent challenge.

Protein surfaces are usually occupied by cationic groups such as amine, imidazole, guanidinium, and anionic carboxylate groups. Hence, developing functional materials that are capable of embracing both the anionic and cationic groups on the protein surface could effectively entrap proteins. Phenylboronic acid (PBA) is a highly electron-deficient group that could bind with electron-rich amino acid residues such as amine on lysine and imidazole on histidine via nitrogen-boronate coordination. The aromatic ring in PBA allows interaction with the guanidinium or ammonium groups on proteins via cation- π interactions. Metal-organic frameworks (MOFs) are formed through the self-assembly of ligands and metal ions, enabling coordination with the free carboxyl groups of proteins. Developing novel MOFs that are designed to capture entire proteins to profile the whole-proteomics in trace samples is highly significant but has been overlooked.

Herein, we synthesized a class of boronic acid-rich lanthanide metal-organic frameworks (MOFs) using 3,5-dicarboxyphenylboronic acid and lanthanide metal salts. The thus manufactured MOFs with high boronic acid grafting ratios enable high binding affinity with different types of proteins via a combination of nitrogen-boronate complexation, cation- π interaction, and coordination between the abundant open metal sites and the bare amino acids residues on the protein surface. We further developed a MOFs Aided Sample Preparation (MASP) pipeline for deep proteome of trace biological samples. Specifically, the MOFs were loaded and proteins in biological samples were captured, followed by discarding the supernatant containing impurities through centrifugation. Subsequently, the proteins on the surface of MOFs were rinsed and enzymatically digested into peptides, which were eluted and desalted before MS acquisition. The MASP process streamlines the protein sample preparation steps by integrating lysis, washing, and digestion within a single PCR tube, thereby reducing sample loss due to surface adsorption and enhancing the depth of proteomic identification in ultra-trace samples. We evaluated the MASP method with 10 to 5000 HEK 293T cells and demonstrated the feasibility and efficiency of MASP in processing trace cells. Our results indicate that MASP enables label-free proteomic analysis in as few as 10 HEK-293T cells, making in-depth proteome analysis more feasible and robust in practical clinical scenarios.

metal-organic framework (MOF), proteomics, trace sample, protein adsorption

doi: [10.1360/TB-2024-1132](https://doi.org/10.1360/TB-2024-1132)