

## 论著·基础研究

## 核小体重塑及组蛋白去乙酰化酶复合物的负染电镜结构分析

段昱娟, 黄 晶

上海交通大学医学院附属第九人民医院上海精准医学研究院, 上海 200125

**[摘要]** **目的**·利用负染电镜技术分析人源核小体重塑及组蛋白去乙酰化酶复合物(nucleosome remodeling and deacetylase complex, NuRD复合物)结构, 获得人源NuRD复合物的轮廓信息。**方法**·将C端带有3×Flag标签的MBD3(methyl-CpG binding domain protein 3)和N端带有10×His标签的GATAD2A(GATA zinc finger domain containing 2A)克隆至pMLink表达载体中, 采用聚乙烯亚胺瞬时转染过表达的方式在人源Expi293F悬浮细胞里表达NuRD复合物中的蛋白质组分; 依次通过Ni-NTA亲和层析、Flag(DYKDDDDK)标签亲和层析和Superose 6 Increase 5/150凝胶过滤层析分离纯化NuRD复合物; 利用蛋白质印迹法(Western blotting)和液相色谱-串联质谱法(liquid chromatography-tandem mass spectrometry, LC-MS/MS)对复合物进行组分鉴定; 利用负染电镜技术结合单颗粒重构技术研究NuRD复合物的空间结构; 通过UCSF Chimera软件将蛋白质数据库(Protein Data Bank, PDB)中已有亚复合物的原子结构模型(7AO9, 5FXV)与生成的结构模型进行自动匹配及比对, 预测多个蛋白组分在负染结构模型中的定位。**结果**·利用两步亲和层析, 成功富集了带有纯化标签的MBD3、GATAD2A蛋白及其他内源蛋白组分, 通过进一步的凝胶过滤层析分离得到了均一性良好的复合物; 通过Western blotting和LC-MS/MS鉴定, 确认纯化得到的复合物为组分完整的人源NuRD复合物。利用负染电镜技术及单颗粒重构技术初步解析了NuRD复合物的空间结构, 其整体轮廓特征明显, 呈现为不对称的长条形; 通过进一步的三维优化处理, 最终获得了人源NuRD复合物分辨率约为17 Å (1 Å=0.1 nm)的初步三维结构模型; 已有的亚复合物原子结构模型(PDB: 7AO9, 5FXV)与NuRD复合物的初步三维结构模型自动匹配后, 初步确定了MTA1/2/3(metastasis-associated protein 1/2/3)、HDAC1/2(histone deacetylase 1/2)、RBBP4/7(retinoblastoma-binding protein 4/7)及MBD2/3(methyl-CpG-binding domain protein 2/3)蛋白亚基在NuRD复合物负染结构模型中的定位。**结论**·利用单颗粒重构技术搭建了人源NuRD复合物的低分辨率负染结构模型。

**[关键词]** 核小体重塑及组蛋白去乙酰化酶复合物; 表观遗传调控; 去乙酰化; 负染电镜技术

**[DOI]** 10.3969/j.issn.1674-8115.2022.04.007 **[中图分类号]** Q518.2; R34 **[文献标志码]** A

## Negative-stain electron microscopic study of the nucleosome remodeling and deacetylase complex

DUAN Yujuan, HUANG Jing

Shanghai Institute of Precision Medicine, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200125, China

**[Abstract]** **Objective**·To study the structure of human nucleosome remodeling and deacetylase complex (NuRD complex) with negative-stain electron microscopy to obtain the profile information of human NuRD complex. **Methods**·Full length MBD3 (methyl-CpG binding domain protein 3) and GATAD2A (GATA zinc finger domain containing 2A) constructs were cloned into the pMLink vectors with an amino terminal 10×His tag and a carboxyl terminal 3×Flag affinity tag, respectively. Proteins were expressed in human Expi293F cells grown in suspension cultures by using polyethylenimine as the transfection reagent, and were isolated via affinity purifications with Ni-NTA and anti-DYKDDDDK (Flag) G1 affinity resin sequentially. The complex was further purified by Superose 6 Increase 5/150 gel filtration chromatography to improve the homogeneity, identified by Western blotting and liquid chromatography-tandem mass spectrometry (LC-MS/MS), and was then studied by negative-stain electron microscopy and

**[基金项目]** 国家自然科学基金(32022036)。

**[作者简介]** 段昱娟(1997—), 女, 硕士生; 电子信箱: yujuan\_duan@sjtu.edu.cn。

**[通信作者]** 黄 晶, 电子信箱: huangjing@shsmu.edu.cn。

**[Funding Information]** National Natural Science Foundation of China (32022036)。

**[Corresponding Author]** HUANG Jing, E-mail: huangjing@shsmu.edu.cn

**[网络首发]** <https://kns.cnki.net/kcms/detail/31.2045.R.20220425.1121.004.html> (2022-04-26 12:02:40)。

single particle analysis. The existing atomic structural model of the subcomplex (7AO9, 5FXY) in Protein Data Bank (PDB) was automatically docked into the generated structural model by the UCSF Chimera software to predict the localization of multiple protein components in the structural model. **Results** The Flag-tagged MBD3 and His-tagged GATAD2A proteins could effectively pull down the other endogenous components of the NuRD complex through two-step affinity purifications. The high-purity complex was obtained by gel filtration chromatography and confirmed as the NuRD complex by LC-MS/MS and Western blotting identification. Preliminary study on the three-dimensional (3D) structure of the NuRD complex was carried out with negative-stain electron microscopy and single particle analysis, which revealed that the NuRD complex was in the obvious shape of a long asymmetrical rod. The 3D structural model of the human NuRD complex was finally obtained by further 3D refinement at an overall resolution of 17 Å (1 Å=0.1 nm). The existing atomic structural model (PDB: 7AO9, 5FXY) was automatically docked into the negative staining structural model, and the localizations of MTA1/2/3 (metastasis-associated protein 1/2/3), HDAC1/2 (histone deacetylase 1/2), RBBP4/7 (retinoblastoma-binding protein 4/7) and MBD2/3 (methyl-CpG-binding domain protein 2/3) subunits in the NuRD complex were preliminarily confirmed. **Conclusion** A low-resolution negative-staining structural model of human NuRD complex is obtained by single particle analysis.

**[Key words]** nucleosome remodeling and deacetylase (NuRD) complex; epigenetic regulation; deacetylation; negative-stain electron microscopy

组蛋白修饰是生物体表观遗传调控的一种主要方式,多种不同类型组蛋白修饰酶通过在染色质核小体的核心组蛋白上添加或去除甲基、乙酰基、磷酸基或泛素等共价修饰,动态地调节染色质的结构与功能,从而广泛参与基因表达调控、干细胞多能性维持、DNA损伤修复及个体发育等重要的生命活动过程<sup>[1]</sup>。组蛋白上的乙酰化动态修饰由组蛋白乙酰转移酶(histone acetyltransferase, HAT)和组蛋白去乙酰化酶(histone deacetylase, HDAC)共同调控<sup>[2]</sup>。HDAC通过去除HAT向组蛋白尾部加入的乙酰基修饰,使染色质形成更高级、浓缩的结构,从而有效抑制基因转录。组蛋白上乙酰化修饰平衡的紊乱往往会使基因表达失控,导致肿瘤的发生<sup>[3]</sup>。有研究发现,HDAC抑制剂能通过增加细胞内组蛋白的乙酰化程度,提高*p21*等基因的表达水平,抑制肿瘤细胞的增殖,诱导细胞分化或凋亡<sup>[4]</sup>,现已成为肿瘤靶向治疗的一个新热点。目前已经研发出的多种能够有效抑制HDAC酶活性的抑制剂大多为针对第I、II、IV类HDAC的广谱抑制剂,在临床治疗中具有较强的不良反应。另外,在通常情况下,HDAC会被招募到Sin3复合物、核小体重塑和组蛋白去乙酰化酶复合物(nucleosome remodeling and deacetylase complex, NuRD复合物)、CoREST (corepressor of RE1-silencing transcription factor)复合物等不同的大分子复合物中<sup>[5-8]</sup>,与其他亚基协同作用,发挥功能活性。HDAC抑制剂对不同的HDAC复合物也表现出明显的选择偏向性<sup>[9]</sup>。单纯靶向于HDAC蛋白自身的癌症治疗药物具有较大的局限性,设计更为有效的、靶向于特定复合物的药物是目前药物开发新的突破点,也是

研究者面临的巨大挑战。

NuRD复合物是一种同时具备染色质重塑酶和去乙酰化酶活性的大分子复合物,由ATP依赖的染色质重塑因子CHD3/4(chromodomain helicase DNA binding protein 3/4)、转录抑制因子GATAD2A/B(GATA zinc finger domain containing 2A/B)、甲基化CpG结合蛋白MBD2/3(methyl-CpG-binding domain protein 2/3)、组蛋白去乙酰化酶HDAC1/2(histone deacetylase 1/2)、肿瘤转移相关蛋白MTA1/2/3(metastasis-associated protein 1/2/3)和组蛋白结合蛋白RBBP4/7(retinoblastoma-binding protein 4/7)等多种蛋白质亚基组成<sup>[10]</sup>。基于NuRD复合物在基因转录调控、DNA损伤修复及干细胞多能性维持上的重要作用,其对于机体的正常生长发育至关重要,其水平也与人类癌症的发生发展及转移潜能密切相关<sup>[11-14]</sup>。受限于超大复合物表达纯化的难度,目前已有的NuRD复合物结构研究多聚焦于去乙酰化核心模块MTA1/2/3-HDAC1/2-RBBP4/7,完整复合物的研究是较为缺失的。MILLARD等<sup>[15-17]</sup>自2013年以来,通过结构解析初步揭示了NuRD复合物去乙酰化核心模块中MTA1/2/3作为支架蛋白分别招募HDAC1/2和RBBP4/7的分子机制。LOW等<sup>[18]</sup>通过金标准绝对定量的质谱分析,初步推断NuRD复合物中各亚基RBBP4/7、HDAC1/2、MTA1/2/3、MBD2/3、GATAD2A/B、CHD3/4的化学计量比为4:2:2:1:1:1,并通过交联质谱、静态光散射等多种方法证实了NuRD复合物是由MTA1/2/3-HDAC1/2-RBBP4/7去乙酰化核心模块和MBD2/3-GATAD2A/B-CHD3/4染色质重塑模块两部分组成,

且整体的NuRD复合物可能呈现为非对称的、模块化的状态。

解析获得完整NuRD复合物的高分辨率结构信息,了解各亚基间的相互作用关系,获得更多的位点信息,对于探索NuRD复合物参与基因表达调控的分子机制以及设计相应的靶向药物具有重要意义。本研究拟通过哺乳细胞瞬时转染过表达和蛋白纯化系统,富集得到完整的人源NuRD复合物,借助120 kV透射电镜和单颗粒重构技术获得NuRD复合物的负染三维结构模型,以期为后续解析高分辨率NuRD复合物冷冻电镜结构奠定基础。

## 1 材料与方法

### 1.1 实验材料

**1.1.1 主要试剂与仪器** Fast Pfu DNA聚合酶(全式金),限制性核酸内切酶(Thermo Fisher Scientific),DNA连接酶Ligation high (Toyobo),质粒大量抽提试剂盒(康为世纪),聚乙烯亚胺(polyethylenimine, PEI) 40 000 (Polysciences),Balance CD 293培养基(Cell wise),Ni-NTA琼脂糖(Qiagen),抗DYKDDDDK G1亲和树脂(金斯瑞),兔源HDAC1单克隆抗体(Proteintech),兔源MTA1单克隆抗体、兔源RBBP4多克隆抗体、兔源CHD4单克隆抗体(ABclonal),鼠源抗His单克隆抗体、鼠源抗Flag单克隆抗体、山羊抗兔IgG、山羊抗兔IgG(AbHO),ECL<sup>TM</sup> Prime Western blotting检测试剂(Cytiva),Superose 6 Increase 5/150、快速蛋白液相色谱仪(GE healthcare),eBlot L1快速湿转仪(金斯瑞),高分辨轨道阱质谱仪(型号Q-Exactive HF, Thermo Fisher Scientific),碳膜铜网(北京中镜科仪技术有限公司),透射电镜(型号Talos L120C, FEI),3%乙酸双氧铀(uranyl acetate, UA)(河南锐欣实验用品有限公司)。

**1.1.2 细胞、菌株与载体** 人胚肾来源的Expi293F细胞,大肠埃希菌DH5 $\alpha$ 菌种、Stbl3菌种,哺乳动物细胞表达载体pMLink均由本实验室保存。

### 1.2 实验方法

**1.2.1 引物合成和DNA序列测定** 实验所用引物均由上海生工生物工程股份有限公司合成。DNA序列测定由上海铂尚生物技术有限公司完成。

**1.2.2 质粒的构建** 利用PCR从HEK293T细胞cDNA文库中获得人源MBD3全长(氨基酸残基:1~291)、GATAD2A全长(氨基酸残基:1~633)的基因编码序列,分别插入到带有Flag或His纯化标签的pMLink载体中,获得相应的pMLink-MBD3-3 $\times$ Flag、pMLink-10 $\times$ His-GATAD2A重组蛋白表达质粒。

**1.2.3 瞬时转染过表达蛋白** 将重组质粒pMLink-MBD3-3 $\times$ Flag和pMLink-10 $\times$ His-GATAD2A以1:1的比例加入到20 mL哺乳动物细胞悬浮培养基中,室温孵育10 min。另外将转染试剂PEI加入到20 mL哺乳动物细胞悬浮培养基中室温孵育10 min。将转染试剂稀释液与质粒稀释液混合,继续室温孵育20 min,最后将转染混合物加入到密度为 $2.5 \times 10^6$ 个/mL的Expi293F细胞中(每毫升细胞转染1 mg目的基因质粒,使用2 mL 1 mg/mL pH 7.0的PEI转染试剂)。加入转染体系后将细胞在37 $^{\circ}$ C、5% CO<sub>2</sub>的摇床中125 r/min继续培养,48~52 h后收集细胞用于蛋白纯化。

**1.2.4 NuRD复合物纯化** 将转染表达有目的蛋白复合物的Expi293F细胞用裂解缓冲液(含50 mmol/L pH 7.5 Tris-HCl、300 mmol/L NaCl、10%甘油、20 mmol/L咪唑、1 mmol/L  $\beta$ -巯基乙醇、1 mmol/L PMSF、5 mmol/L benzamidine、1  $\mu$ g/mL亮抑蛋白酶肽、1  $\mu$ g/mL胃蛋白酶抑制剂)重悬,冰浴条件下超声破碎。将匀浆于4 $^{\circ}$ C、46 000 $\times$ g高速离心1 h,离心后的上清液与平衡好的Ni-NTA填料于旋转混合仪上4 $^{\circ}$ C缓慢混合1 h。混合完成后将混合液135 $\times$ g离心5 min,弃去上清液,将填料转移至层析管中进行重力流动的亲和层析实验,用裂解缓冲液清洗填料,直至考马斯亮蓝检测液检测不到蛋白组分流后,再用相应的洗脱缓冲液I(含25 mmol/L pH 7.5 Tris-HCl、150 mmol/L NaCl、250 mmol/L咪唑、1 mmol/L  $\beta$ -巯基乙醇、1 mmol/L PMSF、5 mmol/L benzamidine、1  $\mu$ g/mL亮抑蛋白酶肽、1  $\mu$ g/mL胃蛋白酶抑制剂)洗脱蛋白。将洗脱下来蛋白与抗DYKDDDDK G1亲和树脂低温旋转混合约5 h,之后用洗涤缓冲液(含25 mmol/L pH 7.5 Tris-HCl、150 mmol/L NaCl、20 mmol/L咪唑、1 mmol/L  $\beta$ -巯基乙醇、1 mmol/L PMSF、5 mmol/L benzamidine、1  $\mu$ g/mL亮抑蛋白酶肽、1  $\mu$ g/mL胃蛋白酶抑制剂)清洗填料直至无杂蛋白流出,再用洗脱缓冲液II(含25 mmol/L pH 7.5 Tris-HCl、150 mmol/L NaCl、500  $\mu$ g/mL Flag多肽)洗脱目的蛋白。将蛋白样品



用截留分子量为 30 000 的浓缩管浓缩至约 100  $\mu\text{L}$ , 利用分子筛 Superose 6 Increase 5/150 进行进一步分离纯化。最后通过十二烷基磺酸钠-聚丙烯酰胺凝胶电泳 (sodium dodecylsulfate-polyacrylamide gel electrophoresis, SDS-PAGE) 分析鉴定, 合并组分完整的 NuRD 复合物蛋白溶液。

### 1.2.5 NuRD 复合物组分的鉴定

(1) 利用蛋白质印迹法 (Western blotting) 鉴定 NuRD 复合物中的各蛋白组分。通过 SDS-PAGE 分离蛋白样品后, 用湿转膜仪将蛋白转移至甲醇活化后的聚偏二氟乙烯 (PVDF) 膜上。之后将 PVDF 膜浸润在含 5% 脱脂奶粉的封闭液中, 室温封闭 2 h。倒弃封闭液, 分别加入稀释至合适浓度的 HDAC1、MTA1、RBBP4、CHD4 蛋白特异性抗体及抗 His、Flag 标签特异性抗体, 4  $^{\circ}\text{C}$  孵育过夜。用膜清洗液 TBST (含 20 mmol/L pH 7.5 Tris-HCl、150 mmol/L NaCl、0.05% tween-20) 清洗 PVDF 膜, 重复 3 次, 每次 10 min。之后加入稀释后的偶联了辣根过氧化物酶 (HRP) 的山羊抗鼠抗体或山羊抗兔抗体, 室温孵育 1 h。用 TBST 清洗 3 次, 每次 10 min。使用化学发光成像法显色鉴定。

(2) 通过质谱鉴定检测纯化产物中的蛋白组分。取约 20  $\mu\text{g}$  蛋白, 经烷基化处理后, 加入胰蛋白酶, 37  $^{\circ}\text{C}$  酶解过夜; 酶解肽段经 C18 柱除盐后, 采用液相色谱-串联质谱法 (liquid chromatography-tandem mass spectrometry, LC-MS/MS) 进行分析; 质谱原始数据提交到 Proteome Discoverer 2.3 软件中, 使用 UniProt 数据库中的人蛋白质数据库 (Protein Data Bank, PDB) 进行检索。

**1.2.6 负染样品的制备、数据收集及模型构建** 因 UA 溶液渗透力较强、颗粒细腻, 故选择其作为样品的负染色液。负染电镜实验步骤如下: 使用辉光放电仪对碳膜铜网进行亲水化处理后, 于碳膜铜网中央滴加 3  $\mu\text{L}$  浓度约为 0.05 mg/mL 的蛋白质溶液, 静置 1 min; 用滤纸从碳膜铜网边缘轻轻吸去多余的溶液, 滴加 3  $\mu\text{L}$  1.5% UA 溶液, 染色 1 min; 用滤纸吸去多余液体, 置于数显式红外烘烤灯下烘烤; 之后将样品置于 120 kV 透射电镜中观察, 放大倍数设置为 57 000 倍, 欠焦值设置为 -1.0~-2.5  $\mu\text{m}$ ; 选择样品颗粒分散性和均一性较好的区域进行数据收集。随后利用数据处理软件 EMAN2 和 RELION3.1.1 对所收集的图像进行处理并生成相应

的三维模型<sup>[19-20]</sup>。用傅里叶空间中不同分辨率壳层中 2 个独立的计算结构的相关性估计三维结构的分辨率, 通常将获得的结构的分辨率定义为傅里叶壳层关联函数 (Fourier shell correlation, FSC) = 0.143 时的空间频率的倒数。

**1.2.7 NuRD 复合物蛋白质结构分析** 将 PDB 中已提交的亚复合物 MTA1-HDAC1-MBD2 冷冻电镜结构 (PDB: 7AO9) 及亚基 RBBP4 晶体结构 (PDB: 5FXV) 通过 UCSF Chimera 软件自动匹配至本研究获得的三维结构模型中<sup>[21]</sup>, 定位二聚体 MTA1-HDAC1、RBBP4 及 MBD2 在 NuRD 复合物负染结构模型中的位置, 并预测其他亚基的定位。

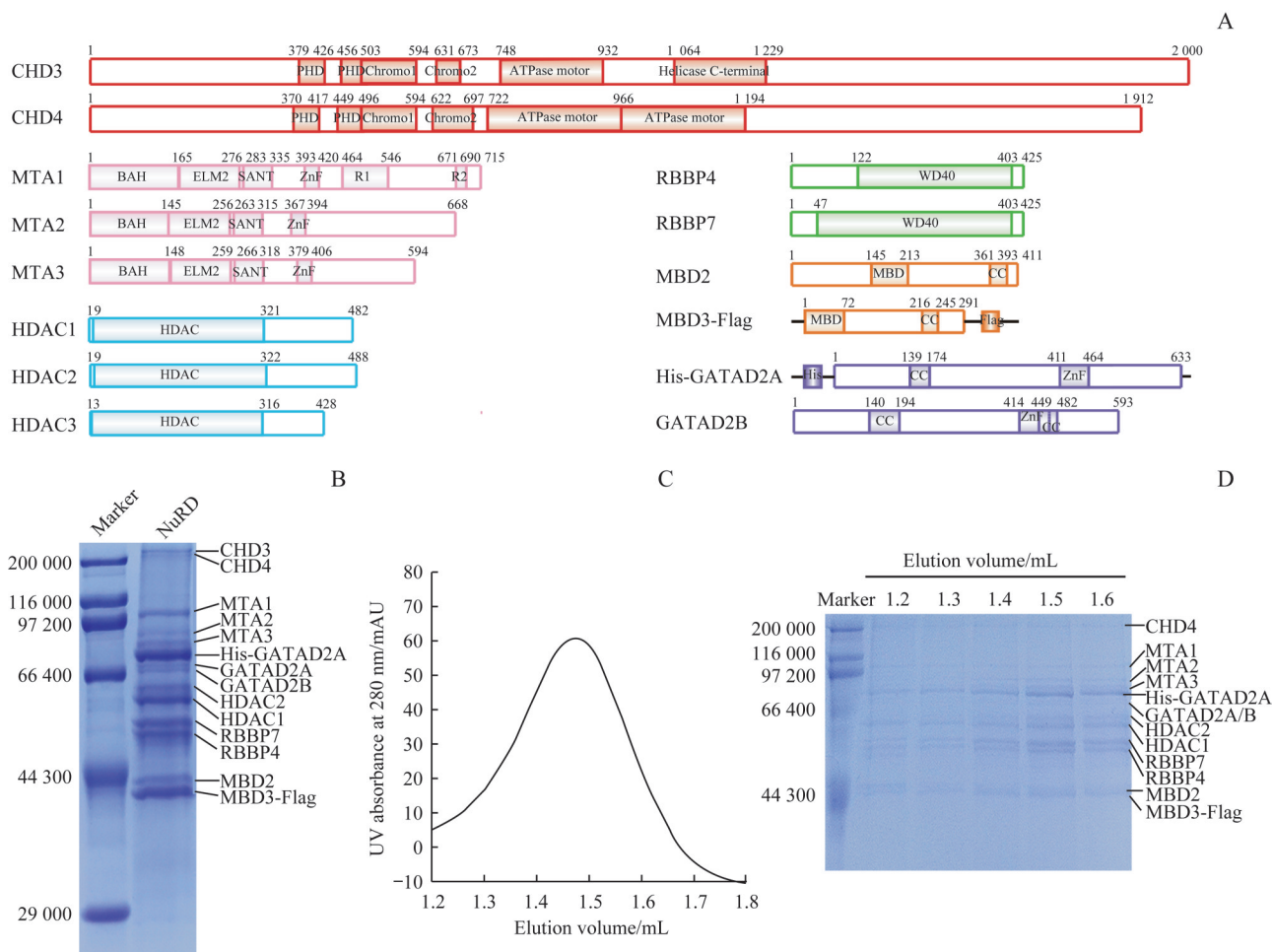
## 2 结果

### 2.1 NuRD 复合物的表达与纯化

经过多重筛选最终决定构建 C 端带有 3 $\times$ Flag 标签的 MBD3 重组蛋白表达质粒和 N 端带有 10 $\times$ His 标签的 GATAD2A 重组蛋白表达质粒, 用于后续的表达纯化。将构建好的质粒在哺乳动物细胞 Expi293F 中瞬时转染共表达, 依次通过 Ni-NTA 纯化和 Flag 标签纯化, 将纯化得到的蛋白通过 SDS-PAGE 分离及考马斯亮蓝染色鉴定。结果显示, 哺乳动物细胞内过表达的 MBD3 和 GATAD2A 能成功共纯化出细胞内的多个内源蛋白, 在与 NuRD 复合物各组分蛋白 (图 1A) 的分子量进行比对分析后, 初步认为纯化产物为组分齐全、纯度较高的 NuRD 复合物样品 (图 1B)。为了确认复合物的状态以及提高样品的均一性, 将亲和层析得到的蛋白样品通过分子筛 Superose 6 Increase 5/150 进一步分离纯化, 得到的复合物在 1.3~1.6 mL 的体积处出峰 (图 1C); 该出峰位置对应分子量大于 669 000 的球状蛋白物, 这与 NuRD 复合物的预测大小 (分子量 720 000) 接近。分管收集流出组分, 通过 SDS-PAGE 分离及考马斯亮蓝染色鉴定, 结果显示复合物中多组分蛋白比例适当, 且于 1.3~1.6 mL 处共同出峰 (图 1D), 合并收集可能为 NuRD 复合物的 1.3~1.6 mL 流出组分, 浓缩用于后续实验。经过整个分离纯化步骤, 从 2 L 的 Expi293F 细胞中可纯化出约 0.1 mg NuRD 复合物。

### 2.2 NuRD 复合物的组分鉴定

针对通过两步亲和层析及分子筛纯化得到的复合物, 利用蛋白特异性抗体成功检测到 MTA1、



**Note:** A. Composition and domain organization of human NuRD complex. The NuRD complex comprised at least six subunits, and there were two or three paralogs for each subunit. The plasmids of MBD3 were constructed with Flag tag at the C-terminus and GATAD2A with His tag at the N-terminus for subsequent protein expression and purification. B. SDS-PAGE analysis of the purified NuRD complex derived sequentially from the Ni-NTA and anti-DYKDDDDK G1 affinity resin. The identities of the major polypeptides and protein size markers were indicated. C. Gel filtration chromatography profile of the NuRD complex on Superose 6 Increase 5/150 GL column. D. SDS-PAGE analysis of fractions across the major peak shown in C. PHD—plant homeodomain finger; BAH—bromo adjacent homology; ELM2—Egl-27 and MTA1 homology 2; SANT—switching defective protein 3 (Swi3), adaptor 2 (Ada2), nuclear receptor co-repressor (N-CoR), transcription factor (TF); ZnF—Zinc finger; R1/R2—RBBP4 recruitment domains; WD40—conserved tryptophan (W) and aspartic acid (D) residues and a repeat length of approximately 40 amino acids; CC—coiled coil.

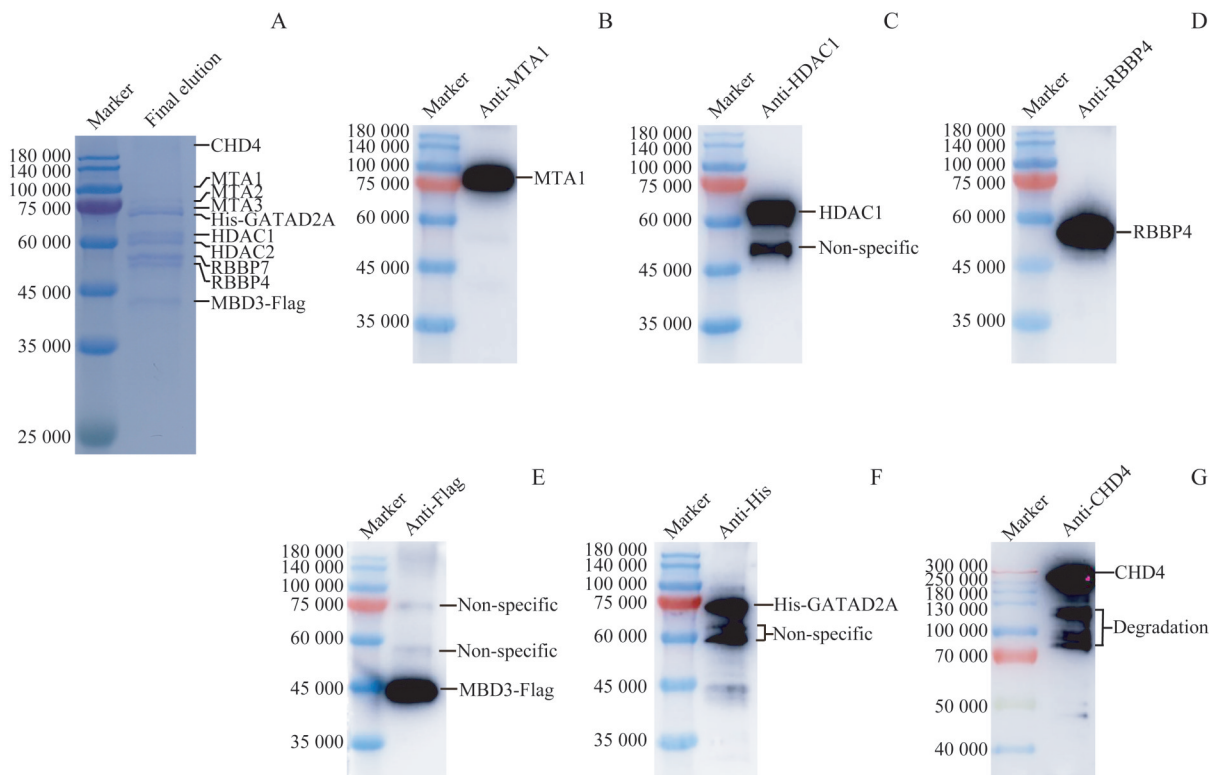
**图1** NuRD复合物的表达与纯化

**Fig 1** Expression and purification of the NuRD complex

HDAC1、RBBP4 和 CHD4 蛋白（图 2A）。采用 Western blotting 对检测到的各组蛋白进行进一步分析，其中，MTA1 和 RBBP4 条带清晰，特异性强（图 2B、2D）；HDAC1 的主条带清晰，但利用该兔源 HDAC1 蛋白单克隆抗体能在约 50 000 处同时检测到一条非特异性结合条带（图 2C），与文献报道结果一致<sup>[22]</sup>；CHD4 对应条带信号强，检测到的多条分子量较小的条带可能为 CHD4 降解产物或其他非特异性结合条带（图 2G）。通过 His、Flag 标签特异性抗体成功鉴定到了相应蛋白 His-GATAD2A 和 MBD3-Flag，虽有非特异性结合条带，但主带信号强且条带清晰（图 2E、2F）。根据分子量大小，

Western blotting 检测到的蛋白与考马斯亮蓝染色后 SDS-PAGE 胶上的主要蛋白条带基本吻合。以上结果均证实 NuRD 复合物中的各主要组分蛋白被有效地共纯化。

LC-MS/MS 鉴定结果显示，纯化得到的复合物终产物中含有 NuRD 复合物的染色质重塑模块蛋白 GATAD2A/B、MBD2/3、CHD3/4 及组蛋白去乙酰化模块蛋白 MTA1/2/3、RBBP4/7、HDAC1/2。分别以蛋白群组中鉴定到的二级肽段谱图数（peptide spectrum matches, PSMs）排序，可粗略估计各蛋白的相对含量，如表 1 所示。上述结果表明纯化得到的产物为 NuRD 复合物。



**Note:** A. The Coomassie-blue staining of the SDS-PAGE gel clearly showed that the final elution contained all major protein components of the NuRD complex. B–G. Through Western blotting analysis, each component of the NuRD complex was detected by using MTA1 rabbit polyclonal antibody (B), HDAC1 rabbit polyclonal antibody (C), RBBP4 rabbit monoclonal antibody (D), anti-Flag mouse polyclonal antibody (E), anti-His mouse polyclonal antibody (F) and CHD4 rabbit polyclonal antibody (G).

**图2** 纯化的NuRD复合物的Western blotting鉴定

**Fig 2** Western blotting identification of the purified NuRD complex

**表1** 纯化的NuRD复合物样品蛋白组分LC-MS/MS分析结果

**Tab 1** Analysis of protein components in the purified NuRD complex by LC-MS/MS

Module	Description		Coverage	Peptide	PSM	Unique peptide	AA
	Full name	Protein					
Nucleosome remodeling module	Transcriptional repressor p66-α	GATAD2A	71	44	412	41	633
	Methyl-CpG-binding domain protein 3	MBD3	75	27	240	25	291
	Chromodomain-helicase-DNA-binding protein 4	CHD4	58	91	189	74	1 912
	Chromodomain-helicase-DNA-binding protein 3	CHD3	21	27	42	13	2 000
	Chromodomain-helicase-DNA-binding protein 5	CHD5	14	20	41	2	1 954
	Transcriptional repressor p66-β	GATAD2B	6	4	23	1	593
	Methyl-CpG-binding domain protein 2	MBD2	18	5	9	3	411
Histone deacetylase module	Metastasis-associated protein MTA2	MTA2	73	41	142	37	668
	Histone-binding protein RBBP7	RBBP7	78	22	121	16	425
	Metastasis-associated protein MTA1	MTA1	72	47	119	40	715
	Histone deacetylase 2	HDAC2	62	21	107	13	488
	Histone deacetylase 1	HDAC1	56	21	93	13	482
	Histone-binding protein RBBP4	RBBP4	69	17	72	11	425
	Metastasis-associated protein MTA3	MTA3	57	28	64	20	594

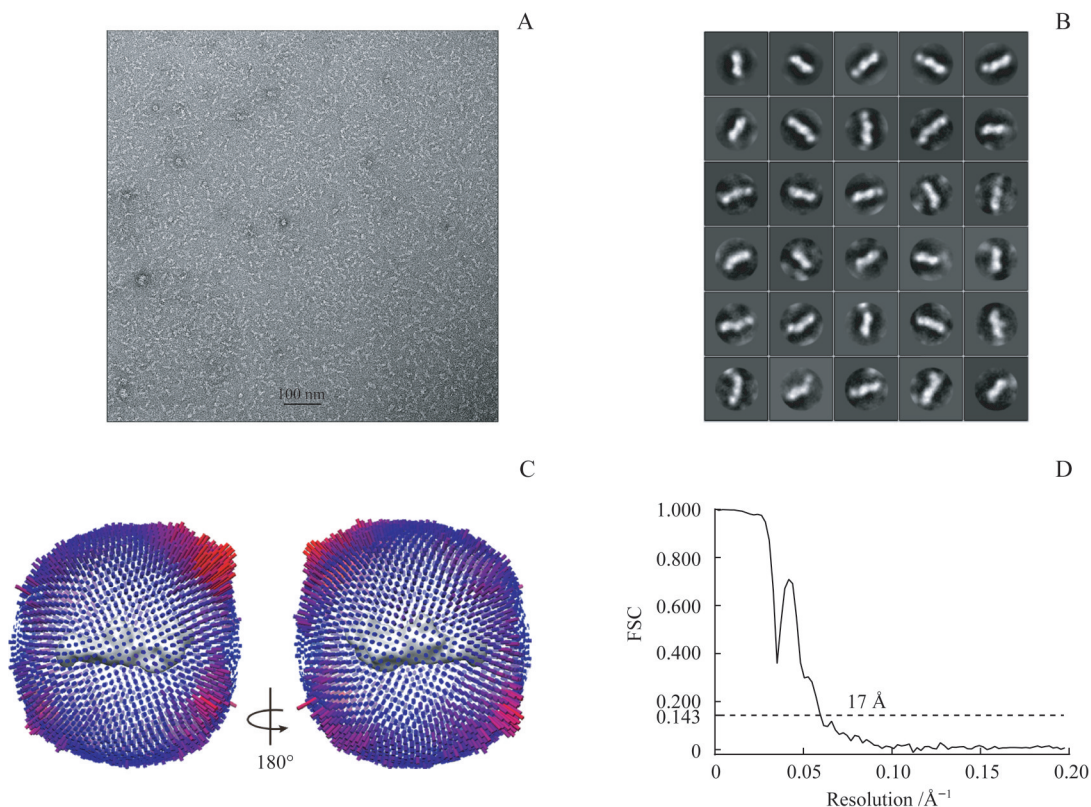
**Note:** Coverage—the percentage of the protein sequence covered by identified peptides; Peptide—the number of distinct peptide sequences in the the protein; PSM—the total number of identified peptide sequences for the protein; Unique peptide—the number of peptide sequences unique to a protein group; AA—the number of the amino acid of the protein.



### 2.3 NuRD复合物的负染电镜数据收集和三维模型构建

在获得高纯度且均一性较好的NuRD复合物蛋白样品后,通过负染电镜技术和单颗粒重构技术分析复合物的三维模型。样品吸附到经亲水化处理的碳膜铜网后,经重金属UA染色,通过120 kV透射电镜观察,样品在负染条件下蛋白质浓度合适,颗粒分散性、均一性良好,虽存在少量细小的小蛋白颗粒,但不影响后续的颗粒挑选及结构解析。电镜照片中大部分颗粒为长条形,长度为25~32 nm,与预期大小基

本相符,其长度差异可能体现复合物不同角度下的截面状态(图3A)。在57 000放大倍数下共收集负染电镜照片233张,运用RELION3.1.1软件挑选颗粒并进行二维分类和三维分类分析。二维分类下,NuRD复合物结构特征明显,大部分呈现出多节球状样的形态(图3B),用于三维重构的颗粒角度分布均匀(图3C)。经过进一步的三维优化处理,最终获得分辨率约为17 Å(1 Å=0.1 nm)的人源NuRD复合物负染三维结构模型(图3D)。



**Note:** A. A representative negative-staining electron micrograph image of the NuRD complex. B. Representative two-dimensional class averages obtained from electron microscope particles of the NuRD complex. C. Angular distributions of particles for the final reconstruction of the NuRD complex. D. The gold-standard FSC curve calculated between two halves of the NuRD complex dataset.

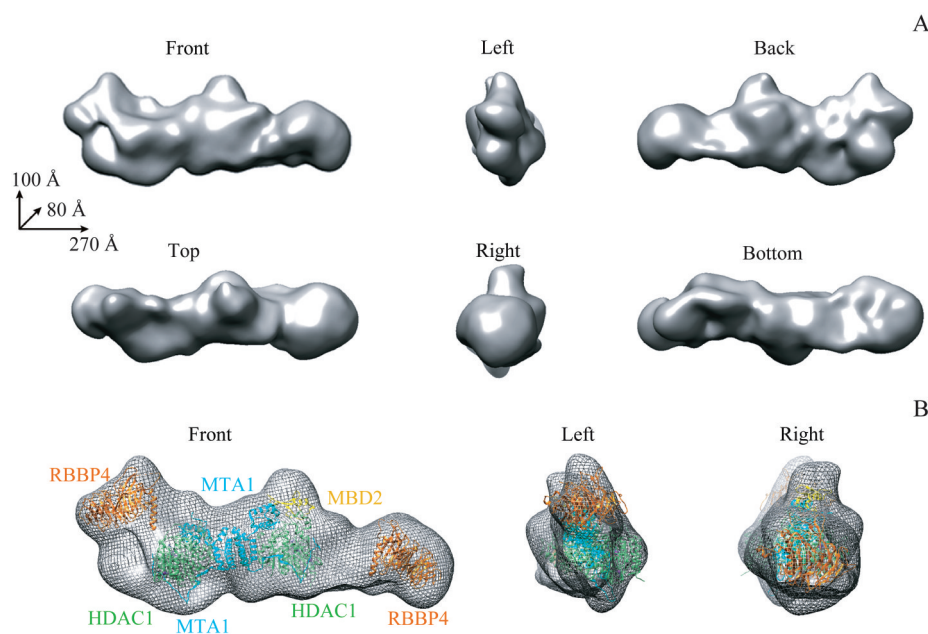
图3 NuRD复合物的负染电镜分析

Fig 3 Negative-staining electron microscopy analysis of the NuRD complex

### 2.4 NuRD复合物结构的初步分析

针对生成的三维结构进行分析,NuRD复合物的尺寸约为270 Å×100 Å×80 Å,呈现为多节球状的长条形,具有与对称二聚体较为吻合的模块,但整体的NuRD复合物呈现出一定的不对称性(图4A),其轮廓和已报道的亚复合物MTA1-HDAC1-MBD2、MTA1-HDAC1-MBD2-RBBP4结构轮廓有所类似<sup>[18]</sup>。将PBD中已有的亚复合物结构(7AO9, 5FXF)通

过UCSF Chimera自动匹配<sup>[21]</sup>,大致确定对称的二聚体MTA1-HDAC1、2个RBBP4亚基及MBD2亚基在NuRD复合物负染结构模型中的定位(图4B)。另外,通过三维结构模型信息可以推测其他亚基的定位,其中,MBD2附近的多余电镜密度可能为GATAD2A/B亚基,HDAC1旁的多余电镜密度可能为RBBP4/7亚基。



**Note:** A. Surface view of the electron micrograph density map of the NuRD complex shown in six orthogonal views. The three-dimensional model size of the NuRD complex is denoted on the left. B. The crystal structures of RBBP4 (5FXY) and the electron microscopy structure of MTA1-HDAC1-RBBP4-MBD2 (7AO9) are docked into corresponding mass of NuRD complex shown in three orthogonal views. MTA1 is colored in cyan, HDAC1 in green, RBBP4 in orange, and MBD2 in yellow.

图4 NuRD复合物三维模型结构和多亚基的定位

Fig 4 Three-dimensional reconstruction of NuRD complex and localization of multiple domains

### 3 讨论

由于在基因转录和DNA损伤修复及肿瘤发生发展过程中发挥着重要的调控功能, NuRD复合物一直是表观遗传学及肿瘤学领域的研究热点<sup>[12-13, 23-24]</sup>。相较于NuRD复合物的功能研究, 关于完整NuRD复合物的结构研究较为缺失。解析NuRD复合物的结构信息, 对于揭示复合物中各亚基组成及相互作用关系、研究NuRD复合物影响基因表达的调控机制、进一步揭示疾病的发生发展过程及研发后续的治疗药物具有重大意义。

有研究<sup>[25]</sup>表明, 染色质结合蛋白PWWP2A/B (PWWP domain-containing protein 2A/B) 能与MBD2/3等组分竞争性地结合染色质重塑模块 (MTA1/2/3-HDAC1/2-RBBP4/7), 形成稳定复合物, 使去乙酰化核心模块 (MTA1/2/3-HDAC1/2-RBBP4/7) 招募染色质重塑模块的能力减弱, 这可能造成在选择去乙酰化核心模块添加纯化标签进行亲和纯化时难以纯化得到完整NuRD复合物。与之前的研究不同, 我们将蛋白纯化重点从较为稳定的组蛋白去乙酰化核心模块转向染色质重塑模块。经过Ni-NTA纯化和Flag标签纯化, 首次分离纯化得到了完整NuRD复

合物。通过凝胶过滤层析Superose 6 Increase 5/150按分子量进一步分离, 为后续的电镜结构研究提供了更为均一完整的复合物状态的蛋白样品。之后, 运用负染电镜和单颗粒重构技术, 获得了NuRD复合物的低分辨率结构模型。从整体轮廓上看, NuRD复合物呈现出条形、多节球状样的结构特征。已有的亚复合物MTA1-HDAC1-MBD2及RBBP4结构可以通过UCSF Chimera软件较好地自动匹配进去, 其他多余密度推测可能为GATAD2A/B、RBBP4/7等亚基组分, 但染色质重塑模块中的CHD4亚基由于分子量较大, 难以在我们解析得到的结构模型中找到合适的定位空间。我们推测由于在纯化产物中CHD4的含量较低, 最终得到的含有CHD4的完整NuRD复合物可能较少, 以至于后续使用RELION3.1.1软件进行结构解析数据处理时二维、三维分类挑选到的颗粒大多为缺失CHD4组分的亚复合物, 最终得到的结构模型可能为缺失CHD4亚基的亚复合物结构模型。在后续的研究中, 可通过额外添加CHD4蛋白进行体外组装的方式提高复合物中CHD4的含量, 以期得到更多更为完整的NuRD复合物用于结构生物学研究。

受限于负染电镜结构解析的局限性, 染色剂中



的重金属可能会遮蔽部分结构信息,且在吸附过程中,所涉及的分子力可引起构象变化,使负染电镜解析的结构信息与真实蛋白质结构间存在些许差异。另外,负染电镜结构解析只能提供粗略的轮廓信息,精确的亚基定位及相互作用信息还需要进一步通过冷冻电镜高分辨率结构解析提供依据。为了得到完整人源 NuRD 复合物的高分辨率结构,未来可通过加入交联剂戊二醛的方式提高复合物的完整稳定性,获得更为均一的样品,进行冷冻电镜结构分析,从而获得 NuRD 复合物的高分辨率冷冻电镜结构;另外,也可尝试寻找合适的核小体底物进一步稳定复合物的结构,解析核小体-NuRD 复合物的高分辨率冷冻电镜结构,进而阐明 NuRD 复合物在基因表达调控上的结构基础。

#### 利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

#### 作者贡献/Authors' Contributions

黄晶设计并指导了整个课题的研究;段昱娟参与了实验设计并完成了所有的实验;段昱娟和黄晶参与了整篇论文的写作与修改。所有作者均阅读并同意了最终稿件的提交。

The study was designed and supervised by HUANG Jing. DUAN Yujuan participated in research design and performed all experiments. The manuscript was drafted and revised by DUAN Yujuan and HUANG Jing. All the authors have read the last version of paper and consented for submission.

• Received: 2022-01-13

• Accepted: 2022-03-07

• Published online: 2022-04-26

#### 参 · 考 · 文 · 献

- [1] CHEN T P, DENT S Y R. Chromatin modifiers and remodellers: regulators of cellular differentiation[J]. *Nat Rev Genet*, 2014, 15(2): 93-106.
- [2] ITO K, J BARNES P, M ADCOCK I. Histone acetylation and deacetylation[J]. *Methods Mol Med*, 2000, 44: 309-319.
- [3] WANG Z B, ZANG C Z, CUI K R, et al. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes[J]. *Cell*, 2009, 138(5): 1019-1031.
- [4] MINUCCI S, PELICCI P G. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer[J]. *Nat Rev Cancer*, 2006, 6(1): 38-51.
- [5] ZHANG Y, IRATNI R, ERDJUMENT-BROMAGE H, et al. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex[J]. *Cell*, 1997, 89(3): 357-364.
- [6] ZHANG Y, NG H H, ERDJUMENT-BROMAGE H, et al. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation[J]. *Genes Dev*, 1999, 13(15): 1924-1935.
- [7] ZHANG Y, LEROY G, SEELIG H P, et al. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities[J]. *Cell*, 1998, 95(2): 279-289.
- [8] YOU A, TONG J K, GROZINGER C M, et al. CoREST is an integral component of the CoREST- human histone deacetylase complex[J]. *Proc Natl Acad Sci U S A*, 2001, 98(4): 1454-1458.
- [9] BANTSCHKEFF M, HOPF C, SAVITSKI M M, et al. Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes[J]. *Nat Biotechnol*, 2011, 29(3): 255-265.
- [10] TORCHY M P, HAMICHE A, KLAHOLZ B P. Structure and function insights into the NuRD chromatin remodeling complex[J]. *Cell Mol Life Sci*, 2015, 72(13): 2491-2507.
- [11] FARNUNG L, OCHMANN M, CRAMER P. Nucleosome-CHD4 chromatin remodeler structure maps human disease mutations[J]. *eLife*, 2020, 9: e56178.
- [12] LI D Q, KUMAR R. Mi-2/NuRD complex making inroads into DNA-damage response pathway[J]. *Cell Cycle*, 2010, 9(11): 2071-2079.
- [13] RAIS Y, ZVIRAN A, GEULA S, et al. Deterministic direct reprogramming of somatic cells to pluripotency[J]. *Nature*, 2013, 502(7469): 65-70.
- [14] LUO M, LING T, XIE W B, et al. NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells[J]. *Stem Cells*, 2013, 31(7): 1278-1286.
- [15] MILLARD C J, WATSON P J, CELARDO I, et al. Class I HDACs share a common mechanism of regulation by inositol phosphates[J]. *Mol Cell*, 2013, 51(1): 57-67.
- [16] MILLARD C J, VARMA N, SALEH A, et al. The structure of the core NuRD repression complex provides insights into its interaction with chromatin[J]. *eLife*, 2016, 5: e13941.
- [17] MILLARD C J, FAIRALL L, RAGAN T J, et al. The topology of chromatin-binding domains in the NuRD deacetylase complex[J]. *Nucleic Acids Res*, 2020, 48(22): 12972-12982.
- [18] LOW J K K, SILVA A P G, TABAR M S, et al. The nucleosome remodeling and deacetylase complex has an asymmetric, dynamic, and modular architecture[J]. *Cell Rep*, 2020, 33(9): 108450.
- [19] TANG G, PENG L W, BALDWIN P R, et al. EMAN2: an extensible image processing suite for electron microscopy[J]. *J Struct Biol*, 2007, 157(1): 38-46.
- [20] SCHERES S H W. RELION: implementation of a Bayesian approach to cryo-EM structure determination[J]. *J Struct Biol*, 2012, 180(3): 519-530.
- [21] PETERSEN E F, GODDARD T D, HUANG C C, et al. UCSF Chimera: a visualization system for exploratory research and analysis[J]. *J Comput Chem*, 2004, 25(13): 1605-1612.
- [22] SANNA S, ESPOSITO S, MASALA A, et al. HDAC1 inhibition ameliorates TDP-43-induced cell death *in vitro* and *in vivo*[J]. *Cell Death Dis*, 2020, 11(5): 369.
- [23] DOS SANTOS R L, TOSTI L, RADZISHEUSKAYA A, et al. MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner[J]. *Cell Stem Cell*, 2014, 15(3): 392.
- [24] LUO J, SU F, CHEN D, et al. Deacetylation of p53 modulates its effect on cell growth and apoptosis[J]. *Nature*, 2000, 408(6810): 377-381.
- [25] ZHANG T Y, WEI G F, MILLARD C J, et al. A variant NuRD complex containing PWWP2A/B excludes MBD2/3 to regulate transcription at active genes[J]. *Nat Commun*, 2018, 9(1): 3798.

[本文编辑] 崔黎明

