

復旦大學

本科毕业论文（设计）



论文题目：N4-甲基胞嘧啶 DNA 甲基转移酶（N4CMT）的结构与
机制研究

姓 名：张雪婷 学 号：21300700017

院 系：生命科学学院

专 业：生物科学（强基计划）

指导教师：甘建华 职 称：研究员

单 位：复旦大学生命科学学院

完成日期： 2025 年 5 月 10 日

N4-甲基胞嘧啶 DNA 甲基转移酶 (N4CMT) 的结构与机制研究

完成人

张雪婷

指导小组成员

甘建华 研究员

目 录

摘要	I
Abstract	II
一、前 言	1
1.1 DNA 甲基化修饰的研究概述	2
1.2 N4CMT: 真核生物中首个催化 4mC 生成的 DNA 甲基转移酶	3
二、材料与amp;方法	5
2.1 实验材料	5
2.1.1 实验用 DNA	5
2.1.2 实验用酶	5
2.1.3 负染实验材料	5
2.1.4 实验常用材料	5
2.1.5 实验仪器	5
2.2 试剂	*
2.2.1 质粒构建实验使用的试剂	5
2.2.2 蛋白表达实验使用的试剂	5
2.2.3 蛋白纯化实验使用的试剂	5
2.3 实验方法	*
2.3.1 目的基因片段的质粒构建	5
2.3.2 目的蛋白的表达	5
2.3.3 目的蛋白的纯化	5

2.3.3 负染色技术.....	5
三、研究结果	*
3.1 克隆载体和表达菌株优化.....	*
3.2 全长蛋白表达.....	*
3.3 蛋白结构预测.....	*
3.4 截断蛋白表达.....	*
四、讨 论	*
参考文献	*
致谢	*

摘要

DNA 甲基化作为重要的表观遗传修饰，在基因调控、基因组稳定性和疾病发生中发挥关键作用。其中，N4-甲基胞嘧啶（4mC）在真核生物中的存在及功能长期存疑，直至蛭形轮虫中首次鉴定出催化 4mC 生成的 DNA 甲基转移酶 N4CMT。本研究以 N4CMT 为对象，旨在解析其结构特性与催化机制，为揭示 4mC 在真核生物中的功能提供理论基础。

通过分子克隆技术构建了 RSiN4CMT 基因的四种表达载体（pET28a、pET-28a(+)-SUMO 等），并优化表达菌株（Rosetta、Coldon Plus 等）。利用镍柱亲和层析、肝素亲和层析及凝胶过滤层析等多步纯化策略，成功获得全长 RSiN4CMT 蛋白。但凝胶过滤层析与透射电镜负染结果表明，全长蛋白呈现异质性聚集，推测其 N 端 54 个氨基酸的无序结构可能导致内在不稳定性。进一步通过 AlphaFold3 结构预测设计截断体（RSiN4CMT 55-401），但其纯化后仍存在酶切后降解或聚集现象。

本研究系统优化了 N4CMT 的表达与纯化流程，揭示了融合标签对蛋白稳定性的影响，为后续研究提供了技术参考。尽管全长蛋白的聚集问题尚未完全解决，但结果表明 N4CMT 的催化结构域可能受无序区域干扰，需进一步通过截断设计或缓冲条件优化提升蛋白均一性。未来工作可结合结构生物学手段解析 N4CMT 的三维构象，并探索其底物特异性与表观遗传调控机制，为阐明 4mC 在真核生物中的生物学意义奠定基础。

关键词：N4-甲基胞嘧啶，DNA 甲基转移酶，蛋白纯化，结构解析

Abstract

DNA methylation, as a crucial epigenetic modification, plays a pivotal role in gene regulation, genome stability, and disease development. The existence and function of N4-methylcytosine (4mC) in eukaryotes remained uncertain until the discovery of the DNA methyltransferase N4CMT, which catalyzes 4mC formation in bdelloid rotifers. This study focuses on N4CMT to elucidate its structural features and catalytic mechanisms, aiming to provide a theoretical foundation for understanding the role of 4mC in eukaryotes.

Four prokaryotic expression vectors (e.g., pET28a and pET-28a(+)-SUMO) for the RSiN4CMT gene were constructed via molecular cloning, with optimized expression strains (e.g., Rosetta and Codon Plus). A multi-step purification strategy, including nickel affinity chromatography, heparin affinity chromatography, and gel filtration chromatography, successfully yielded full-length RSiN4CMT protein. However, gel filtration and transmission electron microscopy revealed heterogeneous aggregation of the full-length protein, likely due to the intrinsically disordered structure of its N-terminal 54 amino acids. A truncated variant (RSiN4CMT 55-401) designed based on AlphaFold3 structural prediction still exhibited degradation or aggregation after purification.

This study systematically optimized the expression and purification workflow for N4CMT, highlighting the impact of fusion tags on protein stability and providing technical references for future research. Although aggregation of the full-length protein remains unresolved, the results suggest that its catalytic domain may be disrupted by disordered regions, necessitating further optimization of truncation designs or buffer conditions. Future efforts should integrate structural biology approaches to resolve the three-dimensional conformation of N4CMT and explore its substrate specificity and epigenetic regulatory mechanisms, thereby advancing the understanding of 4mC's biological significance in eukaryotes.

Key words: N4-methylcytosine, DNA methyltransferase, protein purification, structural characterization

一、前言

1.1 DNA 甲基化修饰的研究概述

在近几十年来，DNA 化学修饰得到了系统性研究。在这些修饰中，由 S-腺苷甲硫氨酸（S-Adenosylmethionine, SAM）依赖的 DNA 甲基转移酶（MTases）特异性催化腺嘌呤或胞嘧啶生成的三种甲基化修饰在生物体内最为普遍，分别为：5-甲基胞嘧啶（5-methylcytosine, 5mC）、N4-甲基胞嘧啶（5-methylcytosine, 4mC）和 N6-甲基腺嘌呤（N6-Methyladenine, 6mA）^{[1][2]}。DNA 甲基化修饰的功能涵盖基因表达调控、X 染色体失活、基因组印记的建立与维持等，并在肿瘤发生等生物学过程中发挥关键作用^{[3][4]}。

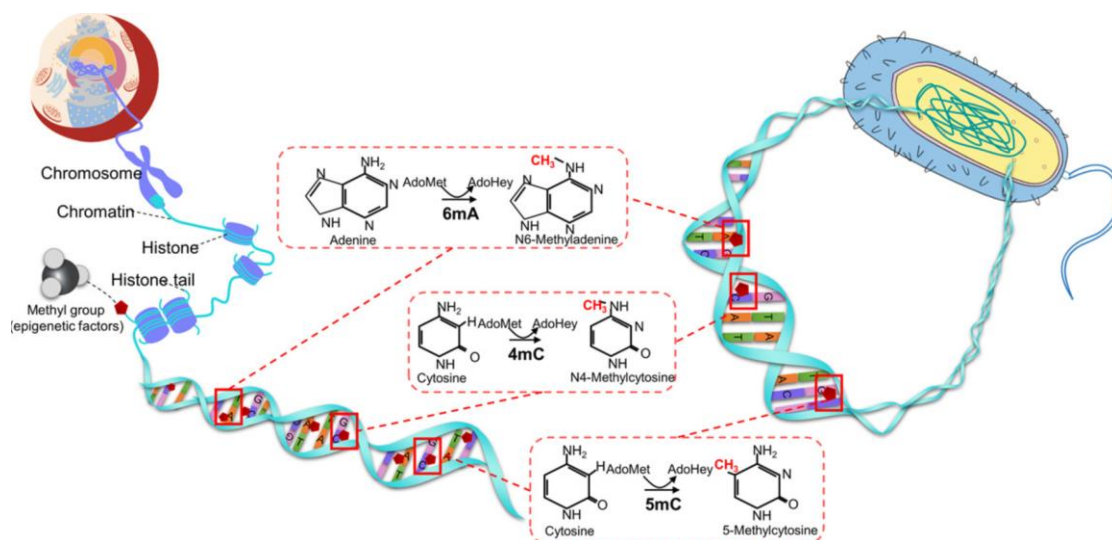


图 1.1 真核生物和原核生物中三种 DNA 甲基化修饰^[5]

1925 年，Johnson 和 Coghill 从结核分支杆菌中分离核酸并从核酸水解产物中首次发现天然存在的 5mC，通过晶体结构、光学性质等结晶学数据确认与其合成样品完全一致^[6]。在这项早期研究发表的 22 年后，Häggqvist 采用纸色谱法和紫外吸收光谱分析证实哺乳动物、昆虫和植物等真核生物的 DNA 中存在大量的 5mC^[7]。该研究表明 5mC 是一种广泛存在的 DNA 修饰，从而推动对 DNA 的甲基化水平进一步广泛探索和量化比较。在过去的二十年间，随着测序技术灵敏度和通量的发展，来自不同生物体的数百种细胞类型和发育阶段的甲基化组图谱被成功绘制^[8]，5mC 的表观遗传调控机制及生物学功能也逐渐被系统性解析和完善。各种研究表明，在真核生物的多个发育和生理过程中，5mC

在各种基因的调控中起着重要作用^[9]，包括基因表达、基因组维护和亲本印记^{[10][11]}。此外，异常的 CpG 岛启动子 5mC 被认为通过沉默肿瘤抑制基因的表达来促进肿瘤进展^{[12][13]}。有研究证明，存在 TET 双加氧酶能够依次将 5mC 氧化为 5-羟甲基胞嘧啶（5-Hydroxymethylcytosine, 5hmC）、5-甲酰胞嘧啶（5-Formylcytosine, 5fC）和 5-羧基胞嘧啶（5-Carboxylcytosine, 5caC）^{[14][15][16][17]}，而不同的氧化形式会招募有特定倾向性的转录因子^{[18][19][20][21][22][23]}。5mC 的形成依赖于甲基转移酶的催化作用，该酶以 SAM 为甲基供体，将甲基基团共价连接到胞嘧啶嘧啶环的第五位碳原子上^[10]。在哺乳动物基因组中生成 5mC 的 DNA 甲基转移酶（DNMTs）的基因 Dnmt1^[26]、Dnmt3a/3b^{[27][28][29]}被克隆和测序。在啮齿动物中，Dnmt3c 是 Dnmt3b 的一个基因复制产物，它是一种在雄性生殖细胞中特异性表达的催化活性 DNA 甲基转移酶，对小鼠的生育能力至关重要^{[27][28][30]}。同时发现，这些哺乳动物 DNMTs 的催化结构域与细菌 II 型 DNA 胞嘧啶-C5 甲基转移酶具有显著的相似性^{[29][33][34]}。

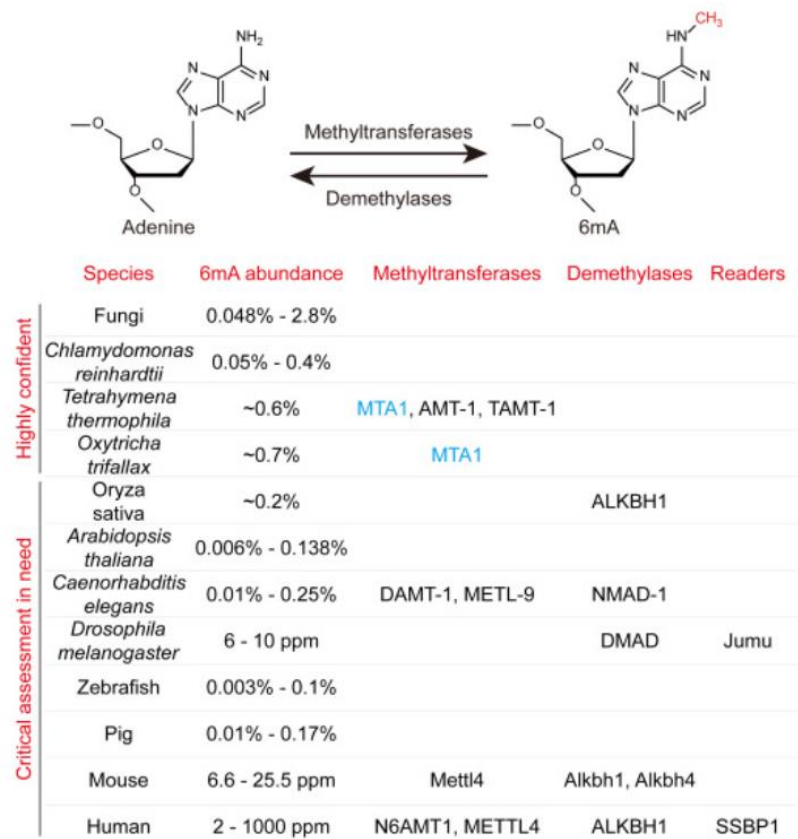


图 1.2 6mA 与假设的 6mA 甲基转移酶和去甲基化酶在各种物种中的丰度^[35]

作为表观遗传标记，6mA 最初在细菌中报道，主要在限制-修饰（restriction-modification, R-M）系统以及在复制、修复、转录和转座中发挥作用[4][36][37][38][39][40]。在单细胞真核生物中，6mA 的丰度普遍处于较低水平。莱茵衣藻细胞内约 0.4% 的腺嘌呤被修饰为 6mA^[41]，而嗜热四膜虫虽在同类生物中修饰程度较高，6mA 占腺嘌呤的比例也仅达到 0.66%^[42]。在多细胞真核生物中，6mA 的存在仍存在争议。目前检测到的 6mA 水平极低（约 0.0001%-0.1%），难以完全排除细菌污染或技术假阳性的可能性^{[43][44]}。近期基于 SMRT 测序的开发的具有 6mA 高分辨率的 6mASCOPE 技术证实，昆虫和植物样本中检测到的大部分 6mA 信号实际来源于其共生细菌基因组。值得注意的是，即使在先前报道 6mA 富集的胶质母细胞瘤干细胞等人类细胞样本中，该技术也未检测到高峰度的 6mA 修饰水平^[45]。鉴定 6mA 甲基转移酶和去甲基化酶是解决关于 6mA 作为哺乳动物表观遗传标记作用的持续争论的最佳方法。目前，METTL4 和 N6AMT1 被推测为哺乳动物中潜在的 6mA 甲基转移酶^{[46][47]}，但其催化活性仍需通过体外重组蛋白实验进行进一步验证。值得注意的是，研究发现 RNA 的 m6A 甲基转移酶复合体 METTL3/METTL14 在体外能以双链 DNA 为底物催化 6mA 生成，并且 METTL3 与细菌甲基转移酶 MTA1 共享保守的 DPPW 基序^[48]。然而，该复合物是否在体内具有相同活性尚未阐明。在鉴定出真正的 6mA 甲基转移酶之前，关于 6mA 是否作为功能性表观遗传标记存在于哺乳动物细胞中仍存在根本性争议。

与关于真核生物 DNA 中 5mC 的大量信息和关于 6mA 的迅速增长的研究相比，关于真核生物 DNA 中 4mC 的研究非常少。在细菌中，4mC 是一种常见的 DNA 修饰方式。然而，4mC 在真核生物基因组中是否具有表观遗传标记的功能一直未得到明确证实。

1.2 N4CMT：真核生物中首个催化 4mC 生成的 DNA 甲基转移酶

2022 年，一项突破性研究首次在蛭形轮虫（bdelloid rotifer）中发现了一种 DNA 甲基转移酶 N4CMT，这是目前已知的唯一能够在动物体内催化 DNA 生成 N4-甲基胞嘧啶（4mC）修饰的酶。^[49]

蛭形轮虫纲的轮虫是一种体长不足毫米的微型淡水无脊椎动物，其特点是采用克隆繁殖方式、体细胞染色体数目恒定、直接发育模式、具有合胞体组织

以及具有古老的基因组四倍体化特征^[50]。特别注意的是，蛭形轮虫展现出惊人的水平基因转移能力，能够高效整合并保留外源基因的功能^[51]。基因组测序分析显示，在其编码序列中，8%-12%来自非后生动物，其中大部分来源于细菌^{[52][53][54]}。

在该研究中，作者使用 4mC 的抗体检测到轮虫 *Adineta vaga* 的基因组 DNA 中存在 4mC。甲基化 DNA 免疫沉淀测序（DIP-seq）的结果显示转座原件（TE）插入位置附近 4mC 有着显著富集。为了在单碱基分辨率下验证这些结果，进行了单分子实时（SMRT）测序，并在超过 20,000 个胞嘧啶中检测到 4mC，约占基因组中总胞嘧啶的 0.06%。^[49]

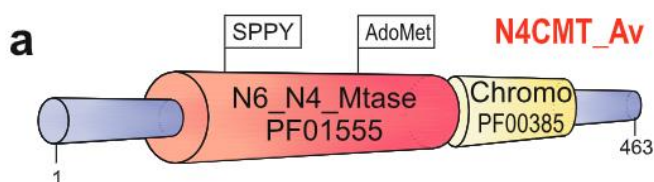


图 1.3 蛭形轮虫中 N4CMT 的结构域^[49]

SPPY：保守的催化基序；AdoMet：S-腺苷甲硫氨酸；数字表示氨基酸位置

作者还鉴定了一种 N4-胞嘧啶甲基转移酶，称为 N4CMT。编码该酶的基因具有甲基转移酶等位基因所代表的 N6_N4_MTase 结构域（PF01555），属于重排型，其催化结构域位于 S-腺苷甲硫氨酸（AdoMet）结合结构域的 N 端。其催化基序为 SPPY，与大多数细菌 N4C-MTases 的催化基序相同，且不同于细菌 N6A-MTases（DPPY）、真核生物 N6AMT1（NPPY）、N6AMT2

（DPPY/F）或 METTL4 样酶（DPPW）的催化基序。该结构域的 C 端还有一个属于 HP1/染色质盒亚家族（该亚家族是结合甲基化赖氨酸的 Royal 家族结构折叠成员）的真核染色质结构域。其直系同源物仅存在于蛭形轮虫纲各个主要科的已测序代表物种中，但在蛭形轮虫纲的姐妹类群单巢纲（Monogononta）或其他已测序的真核生物中却不存在。实验表明，N4CMT 在体外和体内均具有 4mC 甲基转移酶活性，4mC 修饰与转座子沉默和组蛋白修饰有关。^[49] 本研究以 N4CMT 为对象，旨在解析其结构特性与催化机制，为揭示 4mC 在真核生物中的功能提供理论基础。

通过分子克隆技术构建了 RSiN4CMT 基因的克隆载体并优化表达菌株。利用镍柱亲和层析、肝素亲和层析及凝胶过滤层析等多步纯化策略，成功获得全

长 RSiN4CMT 蛋白。但由于全长蛋白呈现异质性聚集，进一步通过 AlphaFold3 结构预测，推测其 N 端的无序结构可能导致内在不稳定性，设计其截断体进行纯化。

二、材料与方法

2.1 实验材料

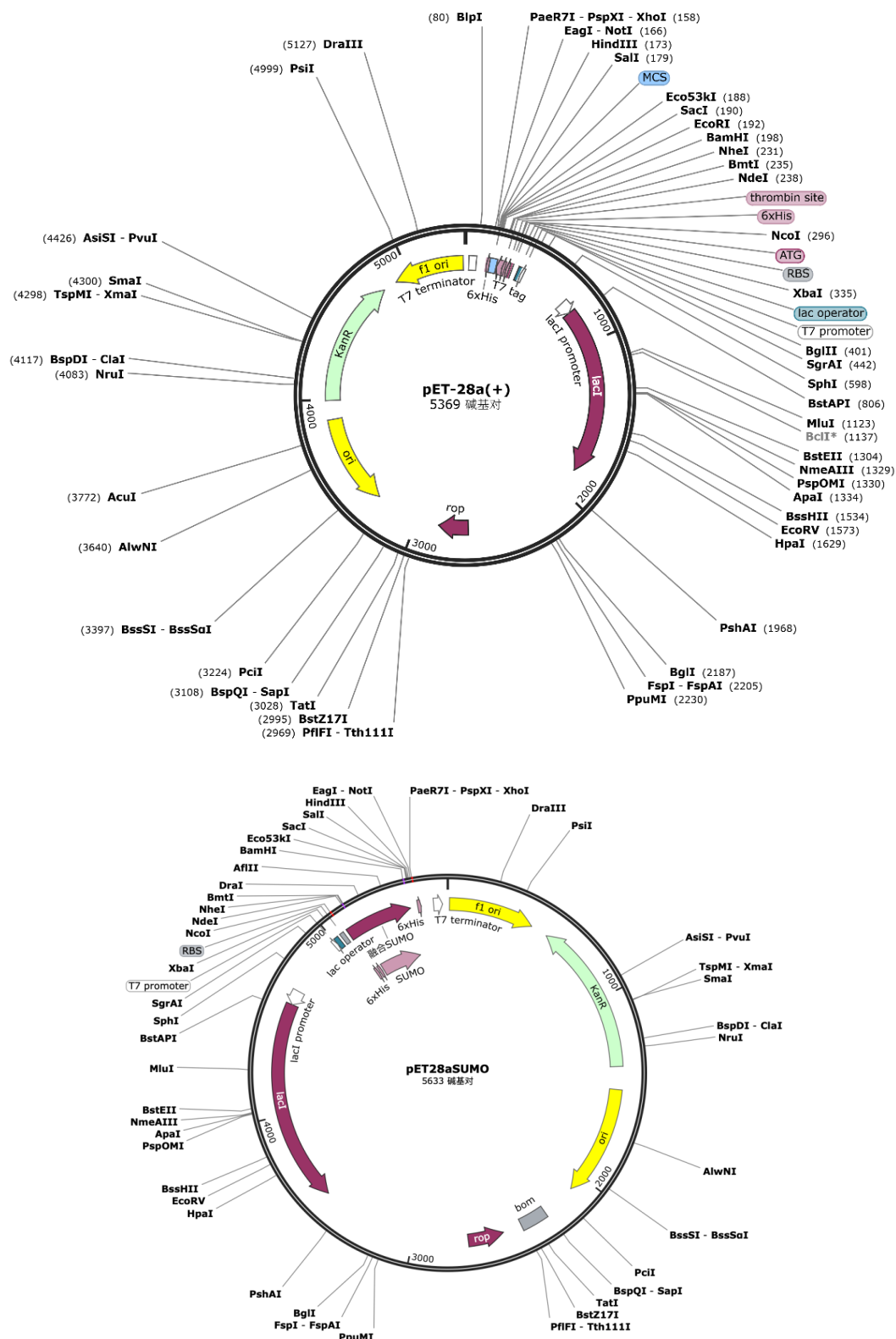
2.1.1 实验用 DNA

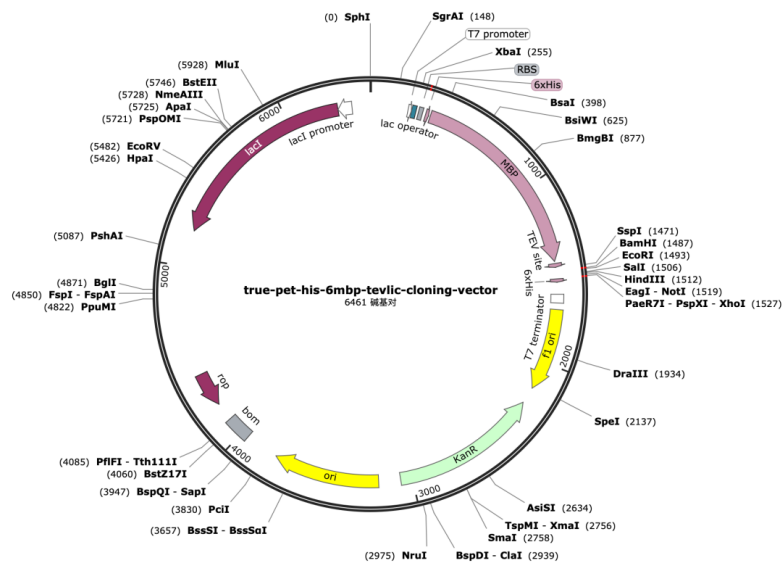
编码 RSiN4CMT 的全长基因是从北京擎科生物科技股份有限公司订购的。编码 RSiN4CMT 的全长基因经过了密码子优化，使其更适合在大肠杆菌中表达。在目的片段的质粒构建过程中，所使用的引物序列均在北京擎科生物科技股份有限公司合成，具体的引物序列如表 2.1 所示。

表 2.1 质粒构建所使用的引物序列

引物名称	引物序列（从 5 ‘到 3’）
RSiN4CMT-28a-F	CAGCAAATGGGTCGCGGATCCATGCATCGTATG ACCACCAGC
RSiN4CMT-28a-R	GTGGTGGTGGTGGTGGTCTCGAGCTAATCATCATC ATCGCTCAGATCGCT
RSiN4CMT-28a-SUMO-F	AGAGAACAGATTGGTGGATCCATGCATCGTATG ACCACCAGC
RSiN4CMT-28a-SUMO-R	GTGGTGGTGGTGGTGGTCTCGAGCTAATCATCATC ATCGCTCAGATCGCT
RSiN4CMT-His6-MBP-F	ATTGGAAGTGGAACGGATCCATGCATCGTATG ACCACCAGC
RSiN4CMT-His6-MBP-R	GTGGTGGTGGTGGTGGTCTCGAGCTAATCATCATC ATCGCTCAGATCGCT
RSiN4CMT-28b-His6-GST-F	AACCTGTATTTTCAGGGATCC ATGCATCGTATGACCACCAGC
RSiN4CMT-28b-His6-GST-R	GTGGTGGTGGTGGTGGTCTCGAGCTAATCATCATC ATCGCTCAGATCGCT
RSiN4CMT_55-401-His6-MBP-F	ATTGGAAGTGGAACGGATCCGGCGGCGGTACC AGCATTAATCATCAGCAGAG
RSiN4CMT_55-401-His6-MBP-R	GTGGTGGTGGTGGTGGTCTCGAGTTAATCATCATCA TCGCTCAGATCG
RSiN4CMT_55-401-28a-SUMO-F	AGAGAACAGATTGGTGGATCCGGCGGCGGTACC AGCATTAATCATCAGC
RSiN4CMT_55-401-28a-SUMO-R	GTGGTGGTGGTGGTGGTCTCGAGTTAATCATCATCA TCGCTCAGATCG

本研究利用同源重组将经过密码子优化的目的基因插入到 pET28a、pET-28a(+)-SUMO、pET His6 MBP TEV LIC、Pet28b-His-GST 四种克隆载体的 BamH I\XhoI 两个酶切位点之间。实验中使用的载体如下图 2.1 所示。





a. pET28a 质粒图谱; b. pET-28a(+)-SUMO 质粒图谱; c. pET His6 MBP TEV LIC 质粒图谱; d. pET28b-His-GST 质粒图谱

本研究中所使用的 ULP-1 蛋白酶和 TEV 蛋白酶来源于本实验室成员表达。限制性内切酶 BamH 1\Xho1 购于 New England Biolabs 公司。DNA 聚合酶 TransStart FastPfu Fly DNA polymerase 采购于北京全式金生物技术股份有限公司。同源重组所用 2×Taq Mix 采购于近岸蛋白质科技有限公司。

2.1.3 负染实验材料

2.1.4 实验常用材料

表 2.2 实验常用材料

2	离心管	50 mL	NEST 生物科技有限公司
3	EP 管	1.5 mL	Axygen 公司
4	EP 管	2 mL	Axygen 公司
5	超滤离心管	15 mL	Millipore 公司
6	超滤离心管	50 mL	Millipore 公司
7	圆底拉盖管	14 mL	Corning 公司

2.1.5 实验仪器

蛋白表达实验过程中所使用到的仪器，具体如表 2.3 所示：

表 2.3 蛋白表达实验使用仪器

编号	仪器名称	仪器型号	仪器生产公司
1	组合式摇床	HYL-C	强乐实验设备有限公司
2	生化培养箱	LRH-150	上海一恒科学仪器有限公司
3	恒温培养箱	DHP-9082	上海天呈实验仪器制造有限公司
4	离心机	Centrifuge 5418	Eppendorf
5	离心机	Centrifuge 5424	Eppendorf
6	离心机	Centrifuge 5424R	Eppendorf
7	离心机	Centrifuge 5804R	Eppendorf
8	离心机	Sorvall ST 40R Centrifuge	Thermo Fisher Scientific
9	离心机	Sorvall ST 1R PLUS Centrifuge	Thermo Fisher Scientific
10	离心机	Sorvall LYNX 6000 Superspeed Centrifuge	Thermo Fisher Scientific
11	离心机	HSC-2015L	宁波新芝生物科技股份有限公司
12	手掌型离心机	LX-200	其林贝尔仪器制造有限公司

蛋白纯化实验过程中所使用到的仪器，具体如表 2.4 所示：

表 2.4 蛋白纯化实验使用仪器

编号	仪器名称	仪器型号	仪器生产公司
1	Ni 亲和层析柱	Fast Flow 5 mL	Cytiva
2	阴离子交换柱	High Performance 5 mL	Cytiva
3	凝胶过滤层析柱	HiLoad Superdex 200 16/60 GL	Cytiva
4	凝胶过滤层析柱	HiLoad Superdex 200 10/300 GL	Cytiva
5	液相色谱系统	ÄKTApurifier UPC100	Cytiva
6	液相色谱系统	ÄKTApurifier 10	Cytiva
7	液相色谱系统	ÄKTApure	Cytiva
8	液相色谱系统	ÄKTA Avant	Cytiva

9	蠕动泵	Peristaltic Pump P-1	Cytiva
10	水浴锅	LabServ LS-B4	Thermo Fisher Scientific
11	高压破碎仪		
12	二级纯水系统	B LabTower EDI 30	Thermo Fisher Scientific
13	超纯水系统	Milli-Q Advantage A10	MERCK
14	实验室雪花制冰机	ST-70	Science-Tool

配制试剂过程中所使用到的仪器，具体如表 2.5 所示：

表 2.5 配制试剂使用仪器

编号	仪器名称	仪器型号	仪器生产公司
1	涡旋混合器	VORTEX-6	其林贝尔仪器制造有限公司
2	多用途旋转摇床	QB-208	其林贝尔仪器制造有限公司
3	加热磁力搅拌器	Cimarec TM	Thermo Fisher Scientific
4	磁力搅拌器	GL-3250C	其林贝尔仪器制造有限公司
5	无油真空泵	R400	Science-Tool
6	pH 仪	FiveEasy TM	METTLER TOLEDO
7	分光光度计	BioMate TM 3S	Thermo Fisher Scientific

2.2 试剂

2.2.1 质粒构建实验使用的试剂

1. 50×TAE 缓冲液的配制

称取 57 mL 冰醋酸并将其加入 100 mL 500 mM EDTA 母液中，搅拌将两者充分混匀。称取 24.2 g Tris base 并将其加入到上述混合液中，搅拌使 Tris base 充分溶解（在通风橱操作）。加入超纯水至 800 mL，磁力搅拌溶解后定容至 1 L。室温保存备用。

实验时，使用超纯水将其稀释 50 倍至 1×TAE 使用。

2. 6×DNA 上样缓冲液的配制

混合 30 mL 甘油（60%）、3 mL 500 mM EDTA 母液、500 μ L 1%溴酚蓝及 500 μ L 1%二甲苯青并使用超纯水将上述溶液定容至 50 mL。放入 4℃冰箱保存备用。

2.2.2 蛋白表达实验使用的试剂

1. LB 固体培养基的配制

称取预混好的 LB 培养基粉末 12.5g 和琼脂粉 7.5 g，将其放入 500 mL 玻璃瓶中并加入 500mL 去离子水，拧上瓶盖（但不要拧紧，防止灭菌时瓶内气压过高发生危险），用铝箔纸包裹瓶口并标记姓名日期等信息。将其放入高压灭菌锅中进行高温高压湿热灭菌（121℃条件下灭菌 20 分钟），灭菌后拧紧瓶盖，放入 60℃烘箱中保存。每次使用时转移至生物安全柜中，取出 50mL 放入离心管中，无菌条件下加入抗生素溶液。根据实验需求，可选择添加 50 μ L 50mg/mL 的卡那霉素或 50 μ L 100mg/mL 的氨苄青霉素，使培养基中抗生素终浓度达到 50 μ g/mL 或 100 μ g/mL（即 1‰）。充分混匀后，将适量的培养基均匀倒入无菌培养皿中，确保完全覆盖皿底。待培养基完全凝固后，倒置放于 4℃冰箱中保存备用。

2. LB 液体培养基的配制

称取 LB 培养基粉末 25 g 加入三角锥形瓶中，并加入 1L 的超纯水，用铝箔纸将瓶口封住并标记姓名日期等信息。将其放入高压灭菌锅中进行高温高压湿热灭菌（121℃条件下灭菌 20 分钟），灭菌后冷却取出，室温保存。

3. 50 mg/ml 卡那霉素母液的配置

称取 2.5 g 硫酸卡那霉素并将其加入到 40 mL 超纯水中，搅拌使卡那霉素充分溶解。使用超纯水将上述卡那霉素溶液定溶至 50mL，并使用无菌 0.22 μ m 滤膜将其进行过滤。放入 4℃冰箱中保存备用。

4. 100 mg/ml 氨苄青霉素母液的配制

称取 5 g 氨苄青霉素粉末并将其加入到 40 mL 超纯水中，搅拌使氨苄青霉素充分溶解。使用超纯水将上述氨苄青霉素溶液定溶至 50mL，并使用无菌 0.22 μ m 滤膜将其进行过滤。放入 4℃冰箱中保存备用。

5. 1M IPTG 的配制

称取 11.92 g IPTG 粉末并将其加入到 45 mL 超纯水中，搅拌使 IPTG 充分溶解。使用超纯水将上述 IPTG 溶液定容至 50mL，并使用无菌 0.22 μ m 滤膜将其进行过滤。每管 1mL 分装后放入-20℃冰箱保存备用。

2.2.3 蛋白纯化实验使用的试剂

1. 5M NaCl 母液的配制

称取 292.2 g 的 NaCl 粉末并将其加入到 900 mL 超纯水中，搅拌使 NaCl 充分溶解。使用超纯水将上述 NaCl 溶液定容至 1L，并使用无菌 0.45 μm 滤膜将其进行抽滤。室温保存备用。

2. 1M Tris-HCl pH 6.8 母液的配制

称取 121.1 g Tris Base 粉末并将其加入到 900 mL 超纯水中，搅拌使 Tris Base 充分溶解。使用 NaOH 或 HCl 将上述 Tris Base 溶液的 pH 值调节至 6.8。使用超纯水将上述 Tris Base 溶液定容至 1 L，并使用无菌 0.45 μm 滤膜将其进行抽滤。室温保存备用。

3. 1M Tris-HCl pH 8.0 母液的配制

称取 121.1 g Tris Base 粉末并将其加入到 900 mL 超纯水中，搅拌使 Tris Base 充分溶解。使用 NaOH 或 HCl 将上述 Tris Base 溶液的 pH 值调节至 8.0。使用超纯水将上述 Tris Base 溶液定容至 1 L，并使用无菌 0.45 μm 滤膜将其进行抽滤。室温保存备用。

4. 1M DTT 的配制

称取 7.71 g DTT 粉末并将其加入到 40mL 超纯水，搅拌使 DTT 充分溶解。使用超纯水将上述 DTT 溶液定容至 50mL，并使用无菌 0.22 μm 滤膜将其进行过滤。每管 1mL 分装后放入-20℃冰箱保存备用。

5. 4M 咪唑 pH 8.0 母液的配制

称取 272.3 g 咪唑粉末并将其加入到 800 mL 超纯水中，搅拌使咪唑充分溶解。使用 NaOH 或 HCl 将上述咪唑溶液的 pH 值调节为 8.0。使用超纯水将上述咪唑溶液定容至 1 L，并使用无菌 0.45 μm 滤膜将其进行抽滤。室温保存备用。

6. 0.5M EDTA pH 8.0 母液的配制

称取 186.1 g 乙二胺四乙酸二钠盐 ($\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$) 并将其加入到 800 mL 超纯水中，磁力搅拌使乙二胺四乙酸二钠盐充分溶解。分次加入 NaOH 颗粒 (约 20 g) 将上述溶液 pH 值调节至 8.0，此过程需缓慢操作以避免过量。使用超纯水将上述 EDTA 溶液定容至 1 L，并使用无菌 0.45 μm 滤膜将其进行抽滤。室温避光保存备用。

7. 10% (w/v) SDS 母液的配制

称取 100 g SDS 粉末并将其加入到 800 mL 超纯水中，60℃加热搅拌使 SDS 粉末充分溶解。使用超纯水将上述溶液定容至 1 L，并使用无菌 0.45 μm 滤膜将其过滤。室温保存备用。

8. 15% SDS-PAGE 分离胶母液的配制

混合 187.5 mL 40% 丙烯酰胺/甲叉双丙烯酰胺 Arc-Bis (29:1)、126 mL 1.5 M Tris-HCl (pH 8.8)、5 mL 10% (w/v) SDS。使用超纯水将上述溶液定容至 500 mL。放入 4℃冰箱保存备用。

9. 5% SDS-PAGE 浓缩胶母液的配制

混合 50mL 40% 丙烯酰胺/甲叉双丙烯酰胺 Arc-Bis (29:1)、50mL 1.5 M Tris-HCl (pH 8.8)、5mL 10% (w/v) SDS。使用超纯水将上述溶液定容至 500mL。放入 4℃冰箱保存备用。

10. 10×TGS 缓冲液的配制

称取 30.2g Tris base 并将其加入到 900 mL 超纯水中。称取 10g SDS 粉末并将其加入到上述 Tris base 溶液中。称取 144g 甘氨酸并将其加入到上述混合溶液中，搅拌使上述物质充分溶解。使用超纯水将上述溶液定容至 1L。室温保存备用。

实验时，使用超纯水将其稀释 10 倍至 1×TGS 使用。

11. 考马斯亮蓝染色液的配制

称取 1.25 g 考马斯亮蓝 R250 并将其加入到 400 mL 无水乙醇中，加入 500 mL 超纯水及 80 mL 冰醋酸，搅拌使考马斯亮蓝 R250 充分溶解。室温避光保存备用，使用前需过滤去除沉淀。

12. 5×SDS-PAGE 上样缓冲液的配制

量取 12.5 mL 1 M Tris-HCl pH 6.8 并将其加入到 25 mL 100% 甘油中，磁力搅拌使两者充分混合。称取 5 g SDS 粉末缓慢加入溶液中，60℃水浴加热并搅拌使 SDS 充分溶解（避免泡沫溢出）。将 2.5 mL 1% 溴酚蓝溶液加入到上述混合溶液中，搅拌使其混匀。使用超纯水将上述混合溶液定容至 50 mL。每管 1 mL 分装后放入-20℃冰箱保存备用。

实验时，需补加还原剂。可每 1 mL 分装溶液中加入 50 μ L β -巯基乙醇（终浓度为 5%），混匀备用；或每 1 mL 分装溶液中加入 20 μ L 1 M DTT（终浓度为 20 mM），涡旋混匀后立即使用。

13. 10×PBS 缓冲液的配制

分别称取 80g NaCl、32.3g $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 和 4.5g $\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$ 、2g KCl 并将其加入到 900 mL 超纯水中，搅拌使上述各物质充分溶解。使用超纯水将上述溶液定容至 1L。室温保存备用。

2.3 实验方法

2.3.1 目的基因片段的质粒构建

为了构建 RSiN4CMT 蛋白的各种克隆表达载体，需要对目的基因片段进行扩增。

2.3.1.1 PCR 扩增目的基因片段

PCR(polymerase Chain Reaction)即多聚酶链式反应，是扩增目的基因片段的常用方法。PCR 经过变性、退火、延伸这三个连续步骤对目的基因片段进行扩增。变性过程将含有目的基因片段的 DNA 双链打开；退火过程使引物与特定的 DNA 位置进行结合；延伸过程以各类 dNTP 为底物，以合成目的基因片段。PCR 的反应体系如表 2.6 所示，包括引物、模板、酶和原料。实验中所用的 PCR 试剂盒来自于北京全式金生物技术公司生产的 TransStart® FastPfu Fly DNA Polymerase 试剂盒。模板来自于合成的 DNA 或测序成功的质粒。将反应体系各项加入 PCR 管中混匀，加样顺序按照体积从大到小排序，最后加入 DNA 聚合酶。

表 2.6 PCR 扩增目的基因片段的反应体系

反应体系组分	使用体积
5× FastPfu Buffer	10 μ L
dNTPs	4 μ L
模版	1 μ L
正向引物（10 μ M）	1 μ L
反向引物（10 μ M）	1 μ L
FastPfu DNA 聚合酶	1 μ L
超纯水	32 μ L
总计	50 μ L

PCR 的退火温度为引物理论 T_m 值-5℃，一般正向引物和反向引物退火温度相同。延伸的时间根据目的基因片段长度来计算，一般每 1kb 延伸 1 分钟。加样完成后将 PCR 管经短暂离心后放入 PCR 仪器中，设置程序开始反应，具体反应程序设置如表 2.7 所示。程序结束后的产物放入 4℃冰箱保存。

表 2.7 PCR 反应程序

反应过程	温度	时间
预变性	95℃	3 min
变性	95℃	30s
退火	T_m -5℃	30s
延伸	72℃	1 kb/min
充分延伸	72℃	5 min

PCR 的反应程序如表 2.7 所示，其中第二步变性、第三步退火、第四步延伸要循环 28 次以上再进入下一个步骤，保证产生足够的 DNA 扩增产物。

2.3.1.2 目的基因片段的琼脂糖凝胶鉴定

称取 1.0 g 电泳级琼脂糖（Agarose I 型）并将其加入到 100 mL 1×TAE 缓冲液中，微波炉加热使琼脂糖充分溶解至透明。待上述溶液冷却至手可以触碰的温度时（约 55℃），吸取万分之一体积的 GelRed 核酸染料，充分混匀后倒入制胶板，室温静置 45 分钟固化，即配制成 1%的琼脂糖凝胶。

吸取 10 μL 的 6×DNA 上样缓冲液并将其加入到 50 μL 的 PCR 产物中，并将样品加入到 1%的琼脂糖凝胶上。使用 150V 的电压，将样品在 1×TAE 中电泳 20 分钟左右，直至目的条带被分离出来。电泳结束后，将琼脂糖凝胶在紫外灯下进行观察。

2.3.1.3 目的基因片段的回收

上述琼脂糖凝胶如果能在紫外灯下看到清晰的荧光条带且分子量大小无误，可以切下含有目的条带的凝胶放入 1.5ml EP 管中，然后使用 DNA 凝胶回收试剂盒对目的基因片段进行回收。使用的 DNA 凝胶回收试剂盒购于广州美基生物公司。具体实验操作如下：

1. 精确称量切取的凝胶质量（精度为 0.1 mg），按质量体积比加入 1-3 倍体积的 Buffer GDP，55℃加热 10-15 分钟，期间颠倒混匀 3 次使胶块充分溶解。

2. 将胶溶液置于在室温下冷却或插在冰上快速冷却，冷却后转移至 DNA 纯化柱，12,000g 离心 1 分钟，重复上柱 1 次以提高结合效率，弃去废液。
3. 吸取 600 μ L Buffer GDP 并加入至 DNA 纯化柱，静置 2 分钟，12,000g 离心 1 分钟，弃去废液。
4. 吸取 700 μ L Buffer DW2（含乙醇）并加入至 DNA 纯化柱，12,000g 离心 1 分钟，弃去废液。
5. 12,000g 空离 2 分钟，以去除残留乙醇。
6. 吸取 65℃ 预热的 35 μ L Elution Buffer 并加入至 DNA 纯化柱，静置 2 分钟，12,000g 离心 1 分钟，将 DNA 洗脱至 Elution Buffer 中。经 Nanodrop 检测回收产物 A260/A280 比值 1.82 ± 0.05 ，标记后放入-20℃冰箱保存。

2.3.1.4 质粒线性化酶切

载体线性化可以通过以下两种方法实现，第一种是限制性酶切法，第二种是体外扩增线性化技术。

限制性酶切法基于 II 型限制性内切酶对特定识别位点的精准切割能力，通过选择载体上预置的酶切位点实现定向线性化。

体外扩增线性化技术利用高保真 DNA 聚合酶的链置换能力，通过反向引物设计实现载体任意位点的体外扩增线性化。相较传统酶切法，其优势体现在可以扩增位点选择自由度高，可突破多克隆位点限制并且产物得率显著提升，通常达 2-5 μ g/反应。但体外扩增过程存在理论错配率（高保真酶错配率约 $1/10^5$ - $1/10^6$ 碱基），需通过测序验证；且扩增产物缺乏细菌体内复制所需的调控元件，可能影响载体稳定性。

在本研究中，采用限制性酶切法，先把目的质粒转化至克隆感受态细胞 DH5 α 中，然后挑取单克隆菌落进行扩增，最后利用碱裂解法质粒抽提试剂盒得到一定量的高浓度的环状原始质粒。抽提质粒的具体实验操作如下：

1. 接种单克隆菌落于 5mL 含质粒相应抗性抗生素的 LB 液体培养基中，置于恒温摇床中 37℃ 220rpm 培养 12-16h（OD₆₀₀≈0.6-0.8）。
2. 吸取 1mL 菌液至 1.5mL EP 管中，12,000g 离心 1min，弃上清。重复收集至菌体总量≥5mg（湿重）。

3. 吸取 250 μ L Solution I（含 RNase A）并加入至 EP 管中，重悬菌体，涡旋震荡至无团块。
4. 吸取 250 μ L Solution II 并加入至 EP 管中，轻柔颠倒 6-8 次（ ≤ 5 min），至溶液透明，。
5. 吸取 350 μ L Solution III 并加入至 EP 管中，立即颠倒 6-8 次，12,000g 离心 10min。
6. 吸取 500 μ L BL 平衡液并加入至吸附柱中，12,000g 离心 10min。
7. 将步骤 5 中的上清转移至平衡后的吸附柱中，12,000g 离心 1min。
8. 吸取 500 μ L PW 漂洗液（含无水乙醇）并加入至吸附柱中，静置 2 分钟，12,000g 离心 1min。重复一次上述操作后，12,000g 空离 2min 彻底去除乙醇。
9. 将吸附柱转移至新 EP 管中，向膜中央滴加 50-100 μ L 预热至 65 $^{\circ}$ C 的超纯水或 TE 缓冲液，室温静置 2 分钟，12,000g 离心 2min，将质粒 DNA 洗脱至超纯水或 TE 缓冲液中。
10. 经 Nanodrop 检测回收质粒 DNA 的浓度，标记后进行后续操作，或放入-20 $^{\circ}$ C 冰箱保存备用。

得到足量的原始环状载体后，使用选取的限制性内切酶将质粒进行双酶切反应，具体反应体系如表 2.8 所示。

表 2.8 载体双酶切反应体系

反应体系组分	使用体积/质量
载体	1-5 μ g
10 $\times$ rCutSmart Buffer	5 μ L
内切酶 1	1 μ l 对应 5 μ g 质粒
内切酶 2	1 μ l 对应 5 μ g 质粒
总计	50 μ L

37 $^{\circ}$ C 酶切 3h 以上，同时留取少量不酶切的原始质粒来进行对照。酶切结束后，吸取 10 μ L 的 6 $\times$ DNA 上样缓冲液并将其加入到 50 μ L 的未进行/进行双酶切的质粒产物中，并将样品加入到 1% 的琼脂糖凝胶上。使用 150V 的电压，将样品在 1 $\times$ TAE 中电泳 20 分钟左右，对照 marker 比对条带大小，线性化载

体的电泳速度比环状质粒慢。回收被完全切开的载体片段，使用 Nanodrop 测量浓度。标记后放入-20℃冰箱保存备用。

2.3.1.5 目的片段的重组与转化

使用同源重组的方法将体外扩增的目的基因与线性化载体相连接，构建含目的基因片的表达载体。本实验使用的是南京诺唯赞的试剂盒，原理如图 2.2 所示。在重组酶的催化下，50℃ 反应 5-15 分钟即可进行转化，完成定向克隆。

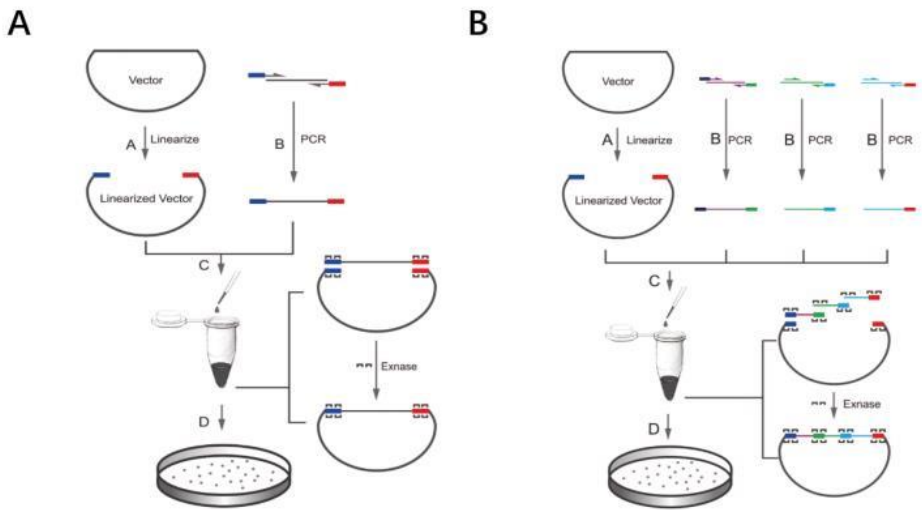


图 2.2 同源重组原理图

本研究使用 ClonExpress Ultra One Step Cloning Kit 进行 DNA 片段重组。该重组克隆试剂盒基于同源重组原理，可高效组装 1-5 个 DNA 片段。其核心组分 2×ClonExpress Mix 含有经优化的 Exnase 重组酶，不仅显著提升重组效率，且对反应体系中的杂质具有较强耐受性。操作步骤如下：首先计算线性化载体与插入片段的使用量。然后计算重组反应所需 DNA 量。配制表 2.9 所示反应体系：反应 50℃，5 min。

表 2.9 同源重组反应体系

反应组成成分	使用体积
线性化载体	X μ L
DNA 插入片段	Y μ L
2 × ClonExpress Mix	5 μ L
超纯水	10-X-Y μ L
总计	10 μ L

重组产物的转化：

1. 将感受态细胞置于冰上解冻。
2. 吸取同源重组的产物并将其加入到感受态细胞中，冰上孵育 30 分钟。
3. 将感受态细胞 42° C 热激 45 s 后，冰上孵育 2-3 分钟。加入 600 μ L 无抗 LB 液体培养基，置于摇床中 37° C 220rpm 培养 1 小时。
4. 5,000rpm 离心 5 分钟，弃上清至管内总体积剩余 100 μ L 左右。用剩余培养基将菌体重悬，涂布在含有质粒相应抗性抗生素的 LB 固体平板上。
5. 倒置于 37°C 培养箱中培养过夜。

2.3.1.6 目的基因片段重组质粒的筛选与测序

经过同源重组转化和过夜培养后，转化平板上可观察到大量单克隆菌落形成。为验证重组效率，需随机挑选若干克隆进行菌落 PCR 验证。建议 PCR 引物设计中 至少包含一条载体通用测序引物，阳性克隆应能扩增出比插入片段略大的特异性条带。对于 PCR 验证为阳性的菌落，可将剩余菌液接种至含相应抗生素的 LB 液体培养基中进行扩增培养。后续可通过质粒提取后的限制性酶切分析，或直接送样进行 Sanger 测序以进一步确认重组质粒的正确性。菌落 PCR 反应体系如表 2.10 所示。

PCR 之后进行电泳并成像，进行鉴定，将成功的送到公司测序比对。

表 2.10 菌落 PCR 反应体系

反应组成成分	使用体积
正向引物 (10 μ M)	1 μ L
反向引物 (10 μ M)	1 μ L
菌液	1 μ L
2 $\times$ Taq Master Mix	5 μ L
超纯水	1 μ L
总计	10 μ L

将测序正确的重组质粒放入-20°C冰箱保存备用。

2.3.2 目的蛋白的表达

2.3.2.1 小量培养

将序列正确的重组质粒转入感受态细胞中。挑取 LB 固体培养基平板上的单克隆菌落，并将其置于 5mL 的含有质粒相应抗性抗生素的 LB 液体培养基中。置于摇床中 37° C 220rpm 培养 3-5 小时，培养至 LB 液体培养基浑浊。

2.3.2.2 大量培养与目的蛋白的诱导表达

将感受态细胞进一步进行大量培养，并诱导目的蛋白表达，具体实验操作如下：

1. 吸取适量的小量培养的菌液并将其加入到 1L 的 LB 液体培养基中，LB 培养基含有质粒抗性相应的抗生素。
2. 置于摇床中 37℃ 220rpm 培养 5-6 小时，培养至对数培养期，OD 值在 0.6-0.8 之间。
3. 置于摇床中 14-18℃ 220rpm 继续降温培养 1 小时。
4. 等到菌液温度变低后，吸取 200μL 1M IPTG 并将其加入到 LB 液体培养基中，置于摇床中 14-18℃ 220rpm 诱导表达 18-20 小时。

2.3.2.3 收集

将上述菌液收集到收菌桶中，5000rpm 离心 10 分钟，去上清。加入适量 1×PBS 溶液重悬，4000rpm 离心 25-30 分钟，去上清，放入-20℃或-80℃冰箱保存备用。

2.3.3 目的蛋白的纯化

2.3.3.1 目的蛋白的预处理

从-20℃冰箱中取出保存的大肠杆菌菌体样品（50mL 离心管），加入在 4℃冰箱预冷的 His Buffer A。待样品在水浴中完全解冻后，使用涡旋振荡器充分重悬菌体，随后将离心管置于冰上保持低温。根据实验需求，可在重悬过程中加入蛋白酶抑制剂（如 PMSF）或甘油等保护剂，以维持目标蛋白的稳定性。

细胞破碎前需进行设备准备工作：首先开启 JN-3000 mini 高压细胞破碎仪及其配套水冷系统，待冷却温度稳定在 4℃后，首先将进样杯中留存的 20%乙醇溶液走干，用超纯水冲洗 2 次破碎管路，然后用 His Buffer A 平衡 1 次破碎管路。将系统压力调节至 1200-1400 MPa 范围内，将菌悬液倒入进样杯中进行破碎处理（重复 3-4 次）。观察发现，初次破碎后溶液通常呈现粘稠状态，经多次破碎后逐渐变得澄清、流动性增强。

破碎完成后，需进行设备维护：先用 His Buffer A 冲洗残留样品，再用超纯水清洗 2 次破碎管路，最后用 20%乙醇溶液灌注管路以防腐蚀和微生物污染。将破碎后的菌液分装至超速离心管，平衡后 14000rpm 离心 1 小时。离心后小心收集上清液，若上清仍显浑浊，可进行二次离心处理以确保澄清度。

2.3.3.2 镍柱亲和层析对目的蛋白进行纯化

亲和层析是一种基于生物分子间特异性相互作用的分离技术，其核心原理是利用金属离子与特定氨基酸残基之间的配位作用实现选择性吸附。在镍柱亲和层析中， Ni^{2+} 通过其六个配位位点与固定化配体（如 IDA 或 NTA）形成稳定的螯合物，剩余的空配位点则能够特异性识别并结合蛋白质中的组氨酸标签

（His-Tag），使带有 $6\times\text{His}$ 标签的目的蛋白结合在镍柱上，这种选择性结合的特性使得镍柱亲和层析成为纯化重组蛋白的常用方法。洗脱过程通常采用竞争性洗脱策略，即通过梯度增加咪唑浓度（10-500mM）来依次洗脱不同结合强度的蛋白质组分。具体实验操作如下：

首先需要进行系统平衡，这是确保纯化效果的重要前提。具体操作包括用超纯水冲洗管路 50mL 以去除杂质，随后以 3mL/min 的流速用 His Buffer A 平衡 5-6 个柱体积，约 35mL。随后将收集的上清液转移至 Ni 柱进行上样，上样阶段需要特别注意流速控制，通常设置在 1-3mL/min 范围内，并根据样品粘度进行适当调整，以防止柱压过高。对于核酸结合蛋白这类特殊样品，还需进行高盐洗涤步骤，使用含 2M NaCl Tris-HCl 8.0 缓冲液冲洗 50-100mL，这一步骤能有效去除非特异性结合的核酸和其他杂质。

最终的梯度洗脱过程使用液相色谱系统 ÄKTA 来完成。系统配置包括 A 泵（His Buffer A）和 B 泵（His Buffer B），在正式洗脱前需执行 Autozero UV 校准并确保基线稳定。洗脱程序设置为 3mL/min 流速下的 20%-100% B 泵线性梯度，每个洗脱峰都应取样进行 SDS-PAGE 分析以评估纯化效果，收集合适的纯化蛋白。值得注意的是，对于某些特殊情况的处理，如 His 标签脱落或弱结合现象，可考虑优化螯合介质或引入辅助纯化步骤来改善纯化效果。整个操作过程中，保持系统的稳定性和避免气体进入层析柱是确保实验成功的关键因素。

2.3.3.2 目的蛋白的酶切与透析

蛋白进过 Ni 初次纯化后，通过 SDS-PAGE 分析筛选目标峰收集组分，合并后加入 Ulp-1 蛋白酶进行位点特异性酶切反应。由于 $6\times\text{His}$ 标签位于 SUMO 标签上游，所以在切除 SUMO 标签的同时， $6\times\text{His}$ 标签也被切掉。

为优化后续纯化效果，需同步进行缓冲液置换处理。此步骤采用透析法降低体系中的咪唑浓度，主要基于以下考量：① 残留咪唑会竞争性抑制 His 标签

与镍柱的结合；② 高浓度咪唑将导致切割后的标签片段无法有效吸附，从而影响二次纯化的分离效率。实际操作中，将样品置于透析装置内，于 4℃ 层析柜中持续搅拌透析约 3 小时或过夜，确保充分去除小分子杂质。

2.3.3.3 使用 Heparin 亲和层析对目的蛋白进行纯化

Heparin 柱固定的肝素配体是一种高度硫酸化的带负电荷的线性多糖（糖胺聚糖），富含磺酸基（ $-\text{SO}_3^-$ ）和羧酸基（ $-\text{COO}^-$ ），可以与核酸结合蛋白发生相互作用，以达到纯化此类蛋白的目的。使用 Heparin 亲和层析对目的蛋白进行纯化，具体实验步骤如下：首先需要对收集的蛋白样品降盐，将 4 倍体积的 20mM Tris-HCl pH 8.0 加入到蛋白样品中，使得目的蛋白样品中的 NaCl 浓度降至 100mM。随后进行平衡系统，具体操作包括用超纯水冲洗管路 50mL 以去除杂质，随后以 3mL/min 的流速用 H Buffer A 平衡 5-6 个柱体积，约 35mL。将稀释后的目的蛋白转移至 Heparin 柱进行上样，流速控制在 2mL/min。使用 Buffer A 将结合的目的蛋白的 Heparin 柱进行充分清洗，流速控制在 2mL/min。

最终的梯度洗脱过程使用液相色谱系统 ÄKTA 来完成。系统配置包括 A 泵（H Buffer A）和 B 泵（H Buffer B），在正式洗脱前需执行 Autozero UV 校准并确保基线稳定。洗脱程序设置为 2mL/min 流速下的 20%-100% B 泵线性梯度，每个洗脱峰都应取样进行 SDS-PAGE 分析以评估纯化效果，收集合适的纯化蛋白。

2.3.3.4 第二次镍柱亲和层析对目的蛋白进行纯化

二次镍柱亲和层析主要用于消除酶切产物中的标签片段、限制性内切酶（带有 His 标签）及非特异性吸附的杂蛋白，从而实现更高纯度的蛋白制备。具体操作流程如下：将上述处理的蛋白样品转移至经 His Buffer A 平衡过的镍柱，在上样过程中采用 Bradford 法实时监测流穿液成分。具体监测方法为：取 5μL 流穿液与 50μL Bradford 显色液混合，通过溶液蓝色深浅变化判断蛋白残留情况。当流穿液开始显色时启动样品收集，上样结束后继续用 His Buffer A 冲洗层析柱直至 Bradford 检测呈阴性。随后直接采用 His Buffer B 对镍柱进行洗脱。所有收集组分均需通过 SDS-PAGE 进行质量评估。需特别注意的是，部分目的蛋白可能因非特异性吸附作用而保留在镍柱上，导致标签切除后仍无法通过二次镍柱有效去除。针对此特殊情况，可选用两种替代方案：① 通过

ÄKTA 纯化系统进行梯度洗脱优化；② 转换其他层析策略，如基于电荷特性的离子交换层析或基于疏水作用的疏水层析等。

2.3.3.5 使用凝胶过滤层析对目的蛋白进行纯化

将蛋白分子通过具有不同大小孔径的凝胶颗粒，较大的蛋白分子无法进入凝胶颗粒的多孔结构而只经过凝胶颗粒外部，较早洗脱下来；较小的蛋白分子进入凝胶颗粒内部，较晚洗脱下来。因此，凝胶过滤层析依据蛋白分子的大小与形状，以达到分离目的蛋白的效果。使用凝胶过滤层析对目的蛋白进行纯化，具体实验操作如下：

1. 色谱柱准备

首先根据目标蛋白的分子量范围选择合适的凝胶过滤介质（如 Superdex 200 用于 50-600 kDa 蛋白）。将干燥的凝胶颗粒用缓冲液（如 20mM Tris-HCl, 150mM NaCl, pH 8.0）充分溶胀，按照厂家推荐时间进行水化。将溶胀后的凝胶浆液缓慢倒入色谱柱中，避免产生气泡。完成装柱后，用 3-5 个柱体积（CV）的缓冲液进行平衡柱子，直至紫外吸收基线稳定。

2. 目的蛋白样品处理

样品体积不应超过柱体积的 2-5%（对于精细分离）或 10-30%（对于脱盐），因此使用浓缩管将收集的目的蛋白进行浓缩，使目的蛋白体积浓缩至 5mL 左右。将浓缩后的蛋白样品通过 0.22 μ m 滤膜过滤，去除可能堵塞柱子的颗粒物。

3. 上样与洗脱

使用进样环或样品泵将样品加载到平衡好的柱子上。上样后立即连接缓冲液进行等度洗脱，保持流速恒定。建议在 280 nm 波长下监测洗脱曲线，收集出现吸收峰对应的馏分。对于复杂样品，可配合 SDS-PAGE 分析各馏分的纯度。典型洗脱体积为 1.5-2 个柱体积，可根据标准蛋白的保留时间估算目标蛋白的洗脱窗口。

4. 数据分析

通过比较标准蛋白的保留体积绘制校准曲线，计算目标蛋白的表观分子量。使用峰面积积分法评估纯度，注意区分主峰与二聚体/多聚体峰。对于脱盐应用，可通过电导率检测确认盐离子是否完全去除。

2.3.4 负染色技术

负染色技术是一种利用透射电子显微镜（TEM）观察微小颗粒的样品制备方法。其原理是通过高电子密度的重金属盐染色剂包裹样品，在背景形成高对比度，使未被染色的样品在电子显微镜下呈现明亮结构。具体的实验步骤包括吸取 5 μ L 乙酸铀与蛋白样品孵育 1 分钟后用滤纸光面吸干，倾斜放置，形成乙酸铀梯度，风干 10 分钟后置于透射显微镜下观察蛋白样品颗粒。

三、研究结果

3.1 克隆载体和表达菌株优化

在本研究中选择来自轮虫动物门的 *Rotaria* sp. ‘Silwood-1’中存在的 N4CMT 直系同源基因，GenBank 序号为 CAF4877480.1，下文简称为 RSiN4CMT。全长野生型的 RSiN4CMT 的分子量约为 47.0kDa，理论等电点为 9.24。本研究首先利用同源重组的方法，将经过密码子优化的目的基因插入到 pET28a、pET-28a(+)-SUMO、pET His6 MBP TEV LIC、Pet28b-His-GST 这四种载体的 BamH I\XhoI 1 两个酶切位点之间，目的基因为 1206 bp，然后将重组质粒转化至克隆感受态 DH5 α 中，并对带有目的基因片段的重组质粒进行筛选与测序。

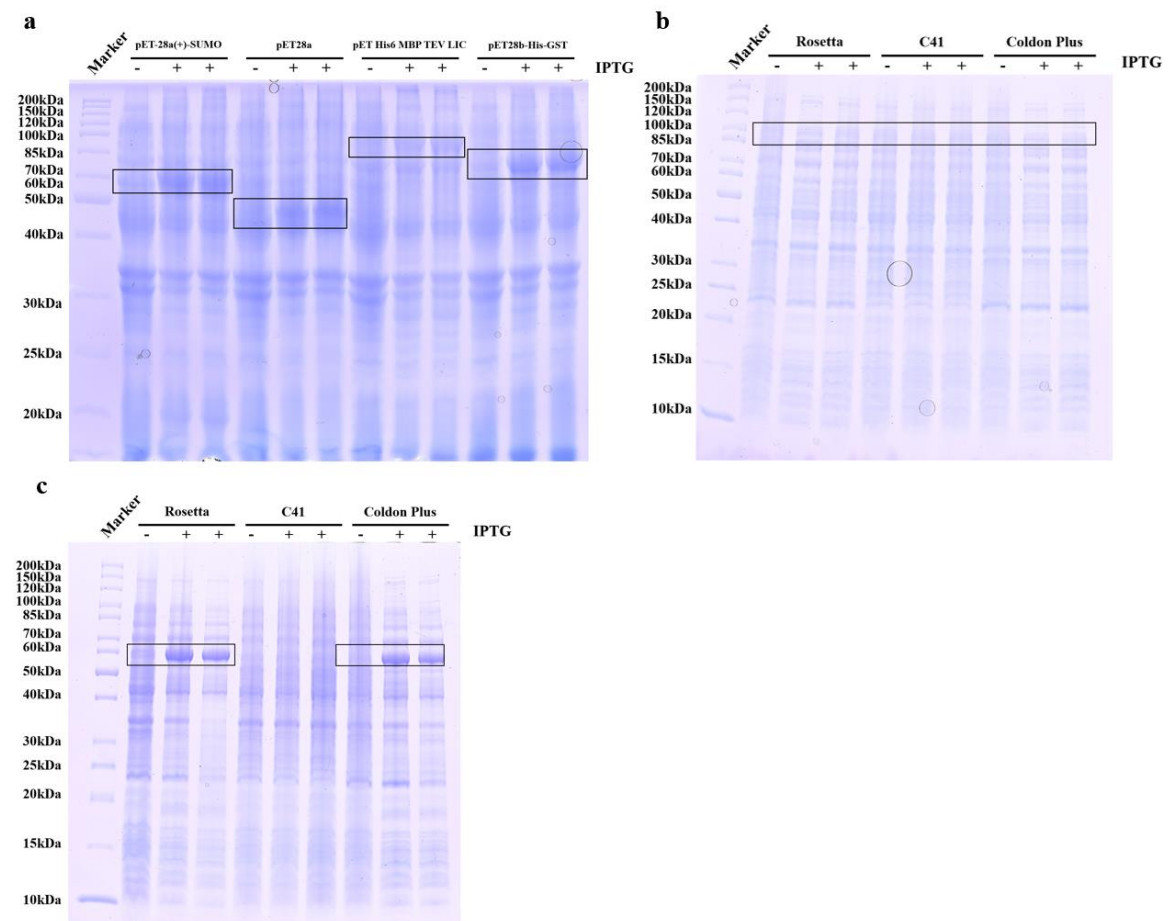


图 3.1 RSiN4CMT 小量诱导表达的 SDS-PAGE 鉴定

a. RSiN4CMT 使用不同载体表达的 SDS-PAGE 图；b. pET28b-His-GST-RSiN4CMT 融合蛋白使用不同表达感受态的 SDS-PAGE 图；c. pET-28a(+)-SUMO-RSiN4CMT 融合蛋白使用不同表达感受态的 SDS-PAGE 图

测序成功后，将四种带有目的基因片段的重组质粒转化到表达感受态 BL21 (DE3) 菌株中，取少量菌液在 18℃ 的温度下加入 0.2% IPTG 诱导 16 小时，将未诱导的菌液作为阴性对照，高温破菌后通过 SDS-PAGE 对目的蛋白的表达情况进行鉴定，如图 3.1 所示。其中，pET28a-RSiN4CMT 融合蛋白的分子量约为 50.1kDa，理论等电点为 9.33；pET-28a(+)-SUMO-RSiN4CMT 融合蛋白的分子量约为 60.4kDa，理论等电点为 8.91；pET His6 MBP TEV LIC-RSiN4CMT 融合蛋白的分子量约为 90.4kDa，理论等电点为 8.10；pET28b-His-GST-RSiN4CMT 融合蛋白的分子量约为 75.1kDa，理论等电点为 8.83。实验发现，pET-28a(+)-SUMO 和 Pet28b-His-GST 这两个载体的目的蛋白表达量相对较高。随后对表达菌株进行优化，将带有目的基因片段的 pET-28a(+)-SUMO 和 pET28b-His-GST 转化至表达感受态 Rosetta、C41 和 Coldon Plus 中，并取少量菌液在 18℃ 的温度下加入 0.2% IPTG 诱导 16 小时，将未诱导的菌液作为阴性对照，高温破菌后通过 SDS-PAGE 对目的蛋白的表达情况进行鉴定，如图 3.1 所示。经实验发现，在表达感受态 Rosetta 和 Coldon Plus 中，pET-28a(+)-SUMO 上目的蛋白表达量相对较高。pET28b-His-GST-RSiN4CMT 融合蛋白在三种表达感受态中的表达量与阴性对照并未有明显差异。

3.2 全长蛋白表达

通过上述表达载体和菌株的优化，后续蛋白表达条件选择使用 pET-28a(+)-SUMO 载体和表达感受态 Rosetta 和 Coldon Plus，并在在 18℃ 的温度下诱导 16 小时。

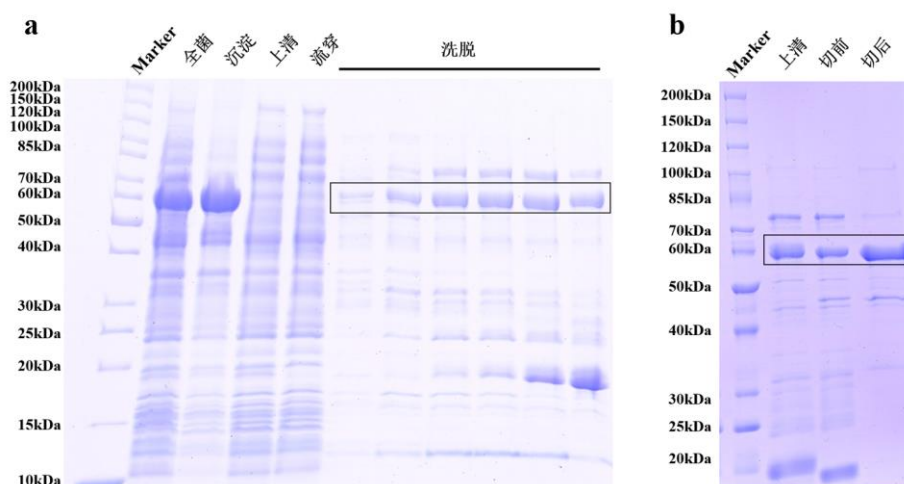


图 3.2 Rosetta 中诱导表达 RSiN4CMT 的 SDS-PAGE 鉴定

a. 镍柱亲和层析后的 SDS-PAGE 图；b. ULP-1 过夜酶切后的 SDS-PAGE 图

本研究首先使用表达感受态 Rosetta 对 RSiN4CMT 进行纯化。将重组质粒转化到 Rosetta 中，培养至对数培养期时，在 18℃ 的温度下加入 0.2% IPTG 诱导 16 小时。通过细胞破碎、镍柱亲和层析后使用 ULP-1 过夜切除 sumo-tag，SDS-PAGE 胶图如图 3.2 所示。由于 pET-28a(+)-SUMO-RSiN4CMT 融合蛋白的理论等电点为 8.91，进行镍柱亲和层析的 His Buffer A/B 的 pH 值均使用 NaOH 或 HCl 将 pH 值调节至 6.5 左右，以避免缓冲液 pH 值与蛋白等电点相近而导致蛋白沉淀。His Buffer A/B 的具体成分为：20mM MES 6.0、500mM NaCl、25/500mM 咪唑。实验发现，将镍柱亲和层析分离收集到的目的条带使用 ULP-1 过夜酶切后无法有效将 SUMO-tag 切除。推测原因为 MBP 标签的酶切位点可能隐藏在蛋白三维结构内部或形成聚集体，阻碍蛋白酶接近酶切位点从而无法去除 MBP 标签。随后尝试使用表达感受态 Coldon Plus 对 RSiN4CMT 进行纯化。将重组质粒转化到 Coldon Plus 中，培养至对数生长期时，在 18℃ 的温度下加入 0.2% IPTG 诱导 16 小时。通过细胞破碎、镍柱亲和层析后使用 ULP-1 切除 sumo-tag，再使用 Heparin 亲和层析后进行二次镍柱亲和层析，最后通过凝胶过滤层析色谱纯化得到 RSiN4CMT，各步骤的 SDS-PAGE 胶图和凝胶过滤层析色谱的洗脱峰图如图 3.3 所示。根据使用表达感受态 Rosetta 对 RSiN4CMT 的纯化实验推测，缓冲液的 pH 值对目的蛋白产量不会产生较大影响，因此在该目的蛋白后续纯化过程中，缓冲液的 pH 值均为 8.0。His Buffer A 的成分为：20mM Tris-HCl 8.0、500mM NaCl、25mM 咪唑，His Buffer A/B 的成分为：20mM Tris-HCl 8.0、150mM NaCl、500mM 咪唑。实验发现，在二次镍柱亲和层析后仅能收集少量目的蛋白条带，并且凝胶过滤层析目的蛋白条带在高聚位置出峰，推测目的蛋白的聚集状态不均一。

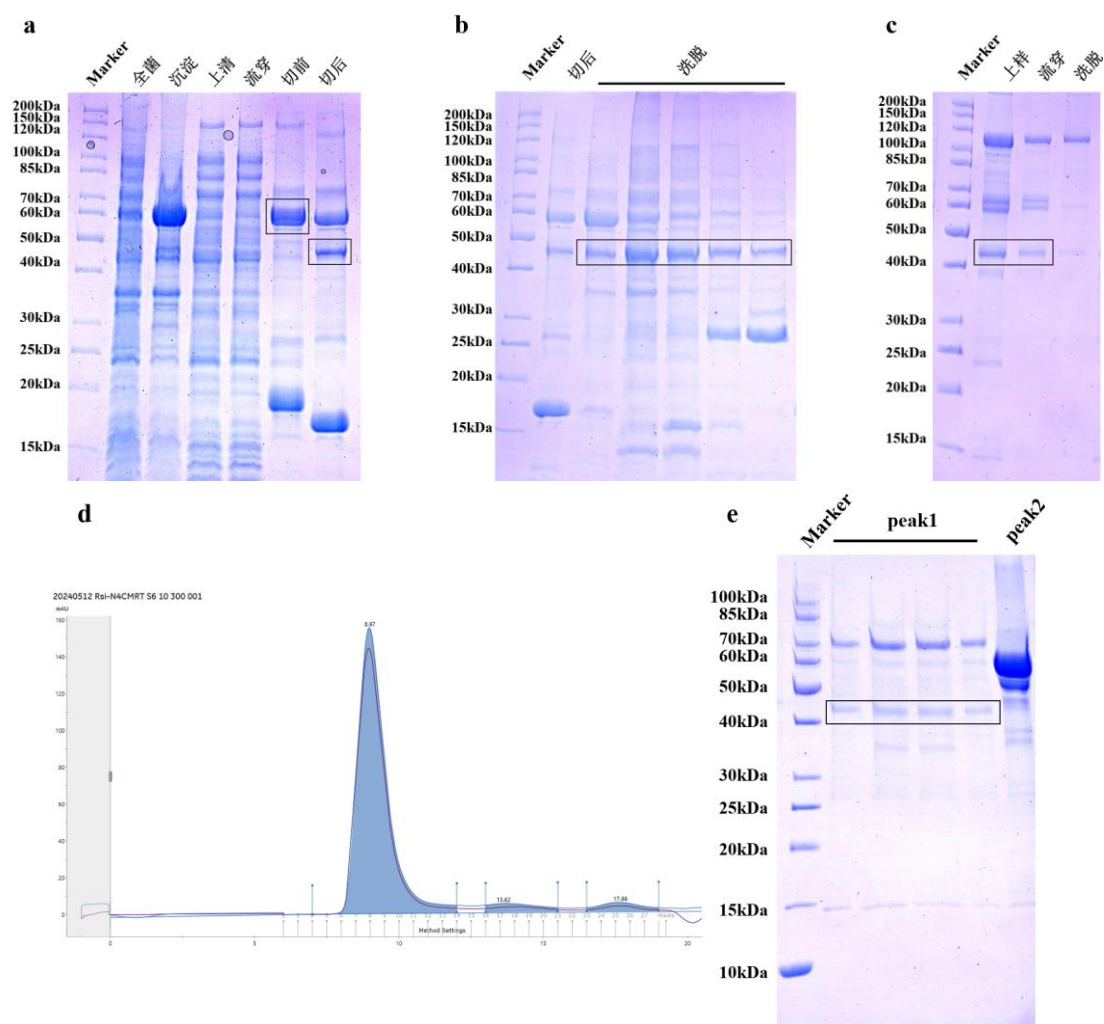


图 3.3 Coldon Plus 中诱导表达 RSIN4CMT 的 SDS-PAGE 鉴定

a. 镍柱亲和层析及 ULP-1 过夜酶切后的 SDS-PAGE 图；b. Heparin 亲和层析后的 SDS-PAGE 图；c. 二次镍柱亲和层析后的 SDS-PAGE 图；d. 凝胶过滤层析洗脱图；e. 凝胶过滤层析后的 SDS-PAGE 图

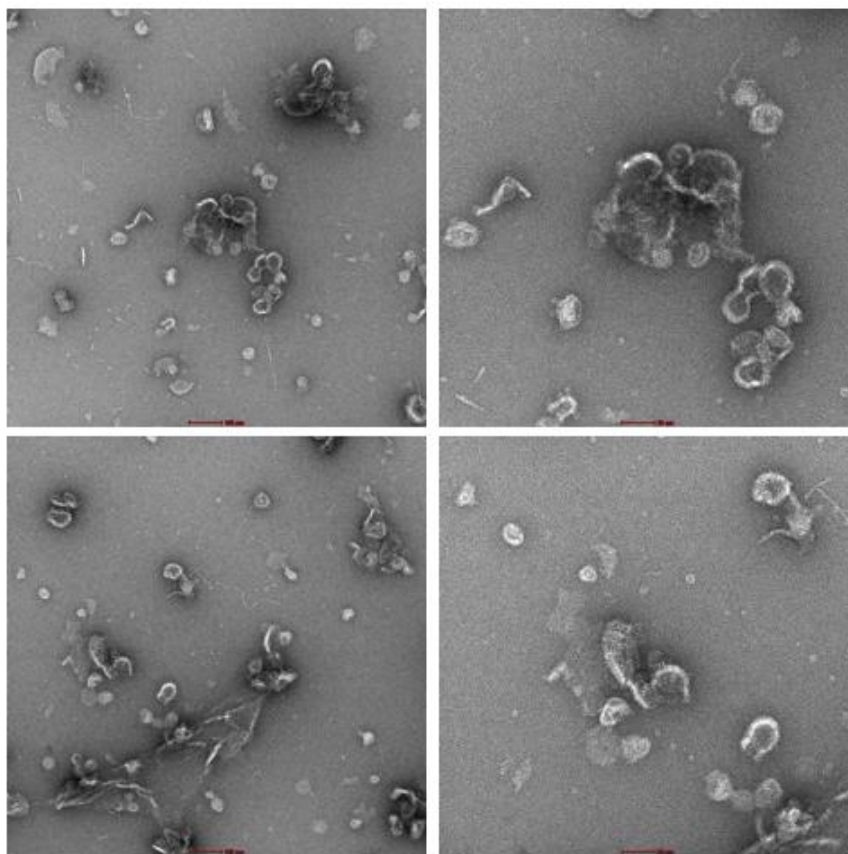


图 3.4 RSiN4CMT 的透射电子显微镜负染结果

将收集到的 RSiN4CMT 样品进行负染样品的制备并在透射电镜下进行观察。如图 3.4 所示，RSiN4CMT 蛋白质样品呈现异质性的聚集状态，提示除了蛋白自身性质存在内在不稳定性的原因之外，在蛋白样品纯化过程中的一些不当操作或条件可能导致蛋白存在部分变性。为了改善目的蛋白的聚集状态，后尝试在镍柱亲和层析和 ULP-1 酶切过程中向缓冲液添加 0.5%的去垢剂 Tween20，但并未有明显改善效果。

3.3 蛋白结构预测

由于全长 RSiN4CMT 的表达纯化存在较多问题，主要包括目的蛋白多以包涵体形式存在，可溶蛋白较少；可溶蛋白带有的蛋白标签难以切除；可溶蛋白混杂较多核酸且很难改善；可溶蛋白形成不规则的聚集状态等，因此尝试通过表达纯化 RSiN4CMT 截断体蛋白来获得状态均一的目的蛋白。图 3.5 中 Alphafold3 对 RSiN4CMT 的预测结构显示，在该蛋白 N 端存在约 54 个氨基酸的无序结构，可能导致蛋白存在内在不稳定性。因此尝试删除 N 端的这段序列，表达纯化 RSiN4CMT 55-401 位氨基酸的截短体蛋白。在截断位点设计正向

引物，保留起始密码子，使用全长目的蛋白的基因质粒作为模版扩增目标序列，通过同源重组的方法将目标序列插入到 pET-28a(+)-SUMO、pET His6 MBP TEV LIC 这两种载体的 BamH I\XhoI 两个酶切位点之间，然后将重组质粒转化至克隆感受态 DH5 α 中，并对带有目的基因片段的重组质粒进行筛选与测序。

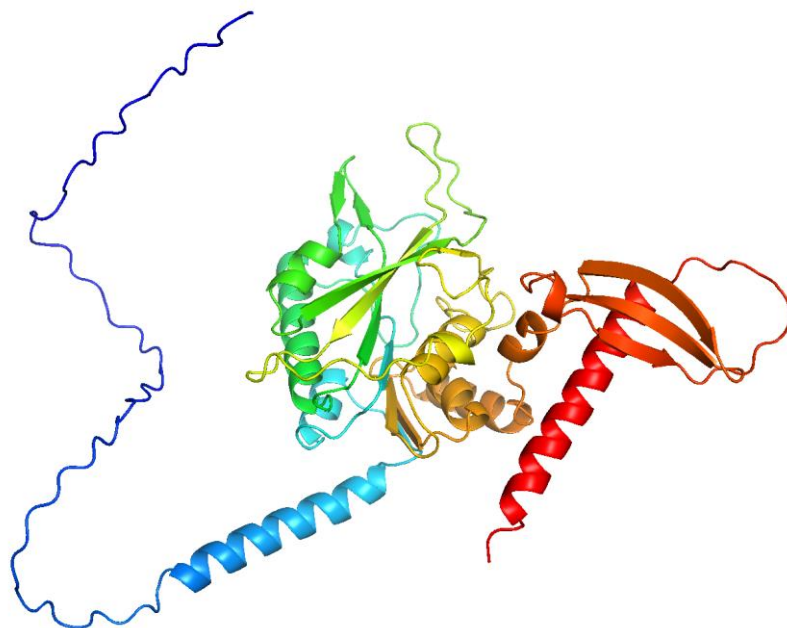


图 3.5 全长 RSiN4CMT 的 AlphaFold3 预测模型

3.4 截断蛋白表达

测序成功后，将两种带有目的基因片段的重组质粒转化至表达感受态 BL21 (DE3) 菌株中，在 18℃ 的温度下加入 0.2% IPTG 诱导 16h。首先纯化 pET His6 MBP TEV LIC 克隆载体上表达的 RSiN4CMT 55-401，通过细胞破碎、镍柱亲和层析后使用 TEV 过夜切除 MBP-tag，透析 4 小时以降低咪唑浓度，最后进行二次镍柱亲和层析，各步骤的 SDS-PAGE 的胶图如下图 3.6 所示。实验发现，在将 MBP-tag 酶切后进行二次镍柱亲和层析，流穿液中没有发现目的蛋白条带。推测原因可能是蛋白在酶切促溶标签后其疏水区域或敏感位点暴露，导致蛋白水解或聚集。

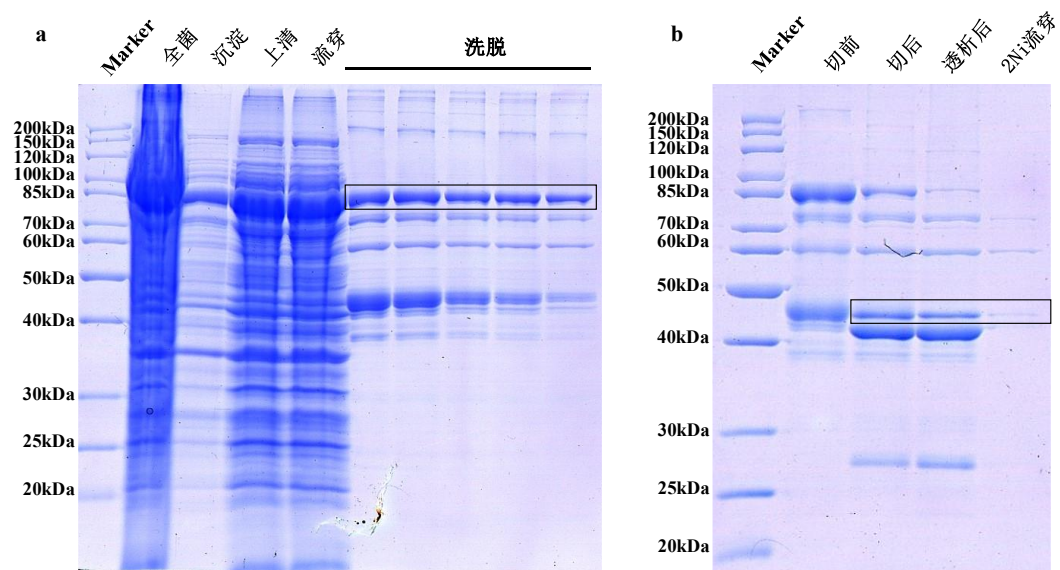


图 3.6 BL21(DE3)中诱导表达 RSIN4CMT55-401 的 SDS-PAGE 鉴定

a. 镍柱亲和层析后的 SDS-PAGE 图；b. TEV 过夜酶切、二次镍柱亲和层析后的 SDS-PAGE 图

四、讨 论

本研究首先通过分子克隆技术构建了 RSiN4CMT 基因的 4 种重组表达载体，并在大肠杆菌表达系统中进行小量诱导鉴定从而选择克隆载体以及表达感受态，随后进行优化表达纯化，包括优化缓冲溶液成分、缓冲液 pH 值和纯化流程等。经过镍柱亲和层析、离子交换层析和凝胶过滤层析等多步色谱纯化成功获得 RSiN4CMT 蛋白。但凝胶过滤层析和透射电镜负染结果均显示，RSiN4CMT 蛋白质样品呈现异质性的聚集状态，无法进行后续的结构和功能研究，并且无法通过在缓冲溶液中添加去垢剂的方式改善，提示全长 RSiN4CMT 蛋白可能存在内在不稳定性。全长 RSiN4CMT 的表达纯化还存在较多问题，主要包括目的蛋白多以包涵体形式存在，可溶蛋白较少；可溶蛋白带有的蛋白标签难以切除；可溶蛋白混杂较多核酸且很难改善；可溶蛋白形成不规则的聚集状态等。因此随后根据 Alphafold3 对 RSiN4CMT 的预测结构选择基因截断位点并设计相关引物，通过分子克隆技术构建了 RSiN4CMT 55-401 截断体基因的 2 种表达载体。初步在大肠杆菌表达系统中进行了表达纯化，但 MBP-tag 酶切后进行二次镍柱亲和层析没有发现目的蛋白条带。通过构建 4 种重组载体（如 pET28a、pET-28a(+)-SUMO 等）并筛选最佳表达菌株（Rosetta 和 Coldon Plus），建立了针对 N4CMT 的高效表达与纯化流程。尽管全长蛋白存在聚集问题，但提示了融合标签（如 SUMO、MBP）对蛋白活性和稳定性的影响，为后续研究其他难溶性真核甲基转移酶提供了参考。

全长 RSiN4CMT 蛋白在大肠杆菌表达系统中主要以包涵体形式存在，可溶蛋白呈现异质性聚集（见图 3.4）。这种聚集现象可能源于蛋白本身的内在无序性（N 端 54 个氨基酸预测为无序结构）或实验条件（如缓冲液成分、缓冲液 pH 值）与目的蛋白的理化性质（如等电点、疏水性或电荷分布）不匹配，导致蛋白稳定性下降。针对当前蛋白表达与纯化的局限性，后续实验条件的优化可围绕以下策略展开：（1）更换缓冲溶液的 pH 值，通过偏离目的蛋白的等电点以降低静电相互作用导致的非特异性聚集；（2）开展去垢剂与稳定剂的筛选，在缓冲液中加入更多类型的非离子型去垢剂（如 Triton X-100、DDM）或小分

子稳定剂（如甘油、海藻糖），通过破坏疏水相互作用或增强蛋白构象稳定性抑制聚集倾向；优化融合标签与载体设计，以提升可溶性组分的占比及层析纯化效率。同时，设计多种的目的蛋白截断体以改善蛋白纯化的问题，具体可基于以下多维分析框架构建截断方案：（1）运用 Phyre2 或 SWISS-MODEL 等工具进行催化结构域预测，精确界定包含保守性甲基转移酶基序；（2）结合 AlphaFold 3 预测的三维结构模型，评估 N/C 端延伸结构域对催化口袋构象的影响；（3）采用系统发育树分析，识别可能参与物种特异性功能适应的可变区段。此方法学不仅可通过去除潜在的结构干扰元件优化蛋白可溶性，更能通过模块化截断策略揭示各结构域在底物识别、SAM 结合亲和力调控及二聚化界面形成中的特异性贡献。在成功获得聚集状态均一的 RSiN4CMT 蛋白样品后，设计一系列对称和非对称的 DNA 底物，使用荧光素酶法（MTase-Glo）等方法检测甲基化反应产物，测定蛋白样品在体外的甲基转移酶活性和底物特异性。通过体外退火、凝胶分离和回收实验获得活性状态的 N4CMT，通过高通量晶体筛选系统进行筛选。对于从筛选试验中获得的初始晶体生长条件进行优化，以获得高质量的晶体用于 X 射线衍射数据的采集。通过分子置换、单波长（SAD）或多波长反常散射（MAD）方法进行结构解析。通过结构分析阐明 N4CMT 识别和修饰底物的分子基础，鉴定出其中参与结合和催化的关键结构。在结构的基础上，设计和合成突变的 N4CMT 或底物，对从晶体结构分析中得出的信息进行进一步的确认。

参考文献

- [1] Roberts, R. J. (2005). How restriction enzymes became the workhorses of molecular biology. *Proceedings of the National Academy of Sciences of the United States of America*, 102(17), 5905–5908.
- [2] Anton, B. P., & Roberts, R. J. (2021). Beyond restriction modification: Epigenomic roles of DNA methylation in prokaryotes. *Annual Review of Microbiology*, 75, 129–149.
- [3] Jones, P. A., & Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science*, 293(5532), 1068–1070.
- [4] Wion, D., & Casadesús, J. (2006). N6-methyl-adenine: An epigenetic signal for DNA-protein interactions. *Nature Reviews Microbiology*, 4(3), 183–192.
- [5] Lv, H., Dao, F.-Y., Zhang, D., Yang, H., & Lin, H. (2021). Advances in mapping the epigenetic modifications of 5-methylcytosine (5mC), N6-methyladenine (6mA), and N4-methylcytosine (4mC). *Biotechnology and Bioengineering*, 118(12), 4204–4216.
- [6] Johnson, T. B., & Coghill, R. D. (1925). The identification of 5-methyl cytosine in tuberculinic acid. *Journal of the American Chemical Society*, 47(11), 2838–2844.
- [7] Vischer, E., & Chargaff, E. (1947). The separation and characterization of purines in minute amounts of nucleic acid hydrolysates. *Journal of Biological Chemistry*, 168(2), 781–782.
- [8] Mattei, A. L., Bailly, N., & Meissner, A. (2022). DNA methylation: A historical perspective. *Trends in Genetics*, 38(7), 676–707.
- [9] Ehrlich, M., & Wang, R. Y. (1981). 5-Methylcytosine in eukaryotic DNA. *Science*, 212(4501), 1350–1357.
- [10] Arand, J., Spieler, D., Karius, T., Branco, M. R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V., & Walter, J. (2012). In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLOS Genetics*, 8(6), e1002750.
- [11] Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13(7), 484–492.
- [12] Pfeifer, G. P. (2018). Defining driver DNA methylation changes in human cancer. *International Journal of Molecular Sciences*, 19(4), 1166.
- [13] Sproul, D., Kitchen, R. R., Nestor, C. E., Dixon, J. M., Sims, A. H., Harrison, D. J., Ramsahoye, B. H., & Meehan, R. R. (2012). Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biology*, 13(10), R84.
- [14] Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., & Rao, A. (2009).

- Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324(5929), 930–935.
- [15] Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C., & Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, 466(7310), 1129–1133.
- [16] He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C. X., Zhang, K., He, C., & Xu, G. L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 333(6047), 1303–1307.
- [17] Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C., & Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 333(6047), 1300–1303.
- [18] Hashimoto, H., Olanrewaju, Y. O., Zheng, Y., Wilson, G. G., Zhang, X., & Cheng, X. (2014). Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes & Development*, 28(20), 2304–2313.
- [19] Wang, L., Zhou, Y., Xu, L., Xiao, R., Lu, X., Chen, L., & Patel, D. J. (2015). Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature*, 523(7562), 621–625.
- [20] Jin, S. G., Zhang, Z. M., Dunwell, T. L., Harter, M. R., Wu, X., Johnson, J., Li, Z., Liu, J., Szabó, P. E., Lu, Q., Xu, G. L., Song, J., & He, C. (2016). Tet3 reads 5-carboxylcytosine through its CXXC domain and is a potential guardian against neurodegeneration. *Cell Reports*, 14(3), 493–505.
- [21] Wang, D., Hashimoto, H., Zhang, X., Barwick, B. G., Lonial, S., Boise, L. H., Vertino, P. M., & Cheng, X. (2017). MAX is an epigenetic sensor of 5-carboxylcytosine and is altered in multiple myeloma. *Nucleic Acids Research*, 45(5), 2396–2407.
- [22] Ren, R., Horton, J. R., Zhang, X., Blumenthal, R. M., & Cheng, X. (2018). Detecting and interpreting DNA methylation marks. *Current Opinion in Structural Biology*, 53, 88–99.
- [23] Yang, J., Horton, J. R., Li, J., Huang, Y., Zhang, X., Blumenthal, R. M., & Cheng, X. (2019). Structural basis for preferential binding of human TCF4 to DNA containing 5-carboxylcytosine. *Nucleic Acids Research*, 47(15), 8375–8387.
- [24] Howard, M. J., Foley, K. G., Shock, D. D., Batra, V. K., & Wilson, S. H. (2019). Molecular basis for the faithful replication of 5-methylcytosine and its oxidized forms by DNA polymerase beta. *Journal of Biological Chemistry*, 294(18), 7194–7201.
- [25] Yang, J., Zhang, X., Blumenthal, R. M., & Cheng, X. (2020). Detection of DNA modifications by sequence-specific transcription factors. *Journal of Molecular Biology*, 432(6), 1661–1673.
- [26] Bestor, T., Laudano, A., Mattaliano, R., & Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology*, 203(4), 971–983.

- [27] Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R., & Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*, 122(10), 3195–3205.
- [28] Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, 19(3), 219–220.
- [29] Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99(3), 247–257.
- [30] Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F., & Bourc’his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science*, 354(6314), 909–912.
- [31] Jain, D., Meydan, C., Lange, J., Claeys Bouuaert, C., Lailier, N., Mason, C. E., Anderson, K. V., & Keeney, S. (2017). Rahu is a mutant allele of Dnmt3c, encoding a DNA methyltransferase homolog required for meiosis and transposon repression in the mouse male germline. *PLOS Genetics*, 13(7), e1006964.
- [32] Dura, M., Teissandier, A., Armand, M., Barau, J., Lapoujade, C., Fouchet, P., Bonneville, L., Schulz, M., Weber, M., & Bourc’his, D. (2022). DNMT3A-dependent DNA methylation is required for spermatogonial stem cells to commit to spermatogenesis. *Nature Genetics*, 54(4), 469–480.
- [33] Posfai, J., Bhagwat, A. S., & Roberts, R. J. (1988). Sequence motifs specific for cytosine methyltransferases. *Gene*, 74(1), 261–265.
- [34] Posfai, J., Bhagwat, A. S., Posfai, G., & Roberts, R. J. (1989). Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Research*, 17(7), 2421–2435.
- [35] Meng, W. Y., Wang, Z. X., Zhang, Y., Hou, Y., & Xue, J. H. (2024). Epigenetic marks or not? The discovery of novel DNA modifications in eukaryotes. *Journal of Biological Chemistry*, 300(4), 106791.
- [36] Braaten, B. A., Nou, X., Kaltenbach, L. S., & Low, D. A. (1994). Methylation patterns in pap regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli*. *Cell*, 76(3), 577–588.
- [37] Laengle-Rouault, F., Maenhaut-Michel, G., & Radman, M. (1986). GATC sequence and mismatch repair in *Escherichia coli*. *EMBO Journal*, 5(8), 2009–2013.
- [38] Lu, M., Campbell, J. L., Boye, E., & Kleckner, N. (1994). SeqA: A negative modulator of replication initiation in *E. coli*. *Cell*, 77(3), 413–426.
- [39] Luria, S. E., & Human, M. L. (1952). A nonhereditary, host-induced variation of bacterial viruses. *Journal of Bacteriology*, 64(4), 557–569.
- [40] Roberts, D., Hoopes, B. C., McClure, W. R., & Kleckner, N. (1985). IS10 transposition is regulated by DNA adenine methylation. *Cell*, 43(1), 117–130.
- [41] Fu, Y., Luo, G. Z., Chen, K., Deng, X., Yu, M., Han, D., Hao, Z., Liu, J., Lu, X., Dore, L. C., Weng, X., Ji, Q., Mets, L., & He, C. (2015). N6-

- methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. *Cell*, 161(4), 879–892.
- [42] Wang, Y., Chen, X., Sheng, Y., Liu, Y., & Gao, S. (2017). N6-adenine DNA methylation is associated with the linker DNA of H2A.Z-containing well-positioned nucleosomes in Pol II-transcribed genes in *Tetrahymena*. *Nucleic Acids Research*, 45(20), 11594–11606.
- [43] O’Brown, Z. K., Boulias, K., Wang, J., Wang, S. Y., O’Brown, N. M., Hao, Z., Shibuya, H., Fady, P. E., Shi, H., He, C., & Megason, S. G. (2019). Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA. *BMC Genomics*, 20(1), 445.
- [44] Ye, P., Luan, Y., Chen, K., Liu, Y., Xiao, C., & Xie, Z. (2017). MethSMRT: An integrative database for DNA N6-methyladenine and N4-methylcytosine generated by single-molecular real-time sequencing. *Nucleic Acids Research*, 45(D1), D85–D89.
- [45] Kong, Y., Cao, L., Deikus, G., Fan, Y., Mead, E. A., Lai, W., Zhang, Y., Yong, R., Sebra, R., Wang, H., & Fang, G. (2022). Critical assessment of DNA adenine methylation in eukaryotes using quantitative deconvolution. *Science*, 375(6581), 515–522.
- [46] Xiao, C.-L., Zhu, S., He, M., Chen, D., Zhang, Q., Chen, Y., Yu, G., Liu, J., Xie, S.-Q., Luo, F., Liang, Z., Wang, D.-P., Bo, X.-C., Gu, X.-F., Wang, K., & Yan, G.-R. (2018). N6-methyladenine DNA modification in the human genome. *Molecular Cell*, 71(2), 306–318.e7.
- [47] Hao, Z., Wu, T., Cui, X., Zhu, P., Tan, C., Dou, X., Hsu, K.-W., Lin, Y.-T., Peng, P.-H., Zhang, L.-S., Gao, Y., Hu, L., Sun, H.-L., Zhu, A., Liu, J., Wu, K.-J., & He, C. (2020). N6-deoxyadenosine methylation in mammalian mitochondrial DNA. *Molecular Cell*, 78(3), 382–395.e8.
- [48] Yu, D., Horton, J. R., Yang, J., Hajian, T., Vedadi, M., Sagum, C. A., Bedford, M. T., Blumenthal, R. M., Zhang, X., & Cheng, X. (2021). Human METTL3-METTL14 RNA adenine methyltransferase complex is active on double-stranded DNA containing lesions. *Nucleic Acids Research*, 49(20), 11629–11642.
- [49] Rodriguez, F., Yushenova, I. A., DiCorpo, D., & Arkhipova, I. R. (2022). Bacterial N4-methylcytosine as an epigenetic mark in eukaryotic DNA. *Nature Communications*, 13(1), 1072.
- [50] Mark Welch, D. B., Mark Welch, J. L., & Meselson, M. (2008). Evidence for degenerate tetraploidy in bdelloid rotifers. *Proceedings of the National Academy of Sciences of the United States of America*, 105(13), 5145–5149.
- [51] Gladyshev, E. A., Meselson, M., & Arkhipova, I. R. (2008). Massive horizontal gene transfer in bdelloid rotifers. *Science*, 320(5880), 1210–1213.
- [52] Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E. G., Hejnal, A., Henrissat, B., Koszul, R., Aury, J.-M., Barbe, V., Barthélémy, R.-M., Bast, J., Bazykin, G. A., Chabrol, O., Couloux, A., Da Rocha, M., Da Silva, C., Gladyshev, E., ... Van Doninck, K. (2013). Genomic evidence for

- ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature*, 500(7463), 453–457.
- [53] Boschetti, C., Carr, A., Crisp, A., Eyres, I., Wang-Koh, Y., Lubzens, E., Barraclough, T. G., Micklem, G., & Tunnacliffe, A. (2012). Biochemical diversification through foreign gene expression in bdelloid rotifers. *PLOS Genetics*, 8(11), e1003035.
- [54] Nowell, R. W., Almeida, P., Wilson, C. G., Smith, T. P., Fontaneto, D., Crisp, A., Micklem, G., Tunnacliffe, A., Boschetti, C., & Barraclough, T. G. (2018). Comparative genomics of bdelloid rotifers: Insights from desiccating and nondesiccating species. *PLOS Biology*, 16(4), e2004830.