

学校编码: 10384

分类号 _____ 密级 _____

学号: 32320141153424

UDC _____

厦 门 大 学

硕 士 学 位 论 文

P-TEFb 磷酸化 RNA 聚合酶 II CTD 的质谱分析

Mass Spectrometry-based Analysis of Phosphorylation
Pattern of RNA Polymerase II CTD by P-TEFb

史中文

指导教师姓名: 高祥副教授

专 业 名 称: 药物分析学

论文提交日期: 2017 年 4 月

论文答辩时间: 2017 年 5 月

学位授予日期: 2017 年 6 月

答辩委员会主席: _____

评 阅 人: _____

2017 年 5 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月

摘要

在真核细胞中，RNA 聚合酶 II (RNA Pol II) 负责大多数编码蛋白质的 mRNA 和核内小 RNA (snRNA) 基因的转录，其碳端结构域 (CTD) 为 RNA 聚合酶 II 所特有，在结合转录循环和基因表达中扮演着重要的角色，转录循环过程中，TFIIH 和 P-TEFb 是参与其磷酸化修饰调控的重要通用转录因子。其中，P-TEFb 不仅是一种通用转录因子，同时也是艾滋病毒 (HIV-1) 转录复制过程中不可或缺的宿主转录因子之一。P-TEFb 对 CTD 磷酸化模式的分析一直是研究热点。然而，CTD 高度磷酸化状态及其七肽重复结构 (YSPTSPS) 等特性使其质谱分析面临巨大挑战。

本课题首先建立 P-TEFb 磷酸化 RNA 聚合酶 II CTD 的体外实验模型，基于生物质谱技术对磷酸化模式进行分析。主要内容包括以下几点：(1) 建立高活性 P-TEFb 的亲分和分离制备方法，运用合成肽段底物对激酶活性进行验证；(2) 通过胰酶酶切成功鉴定 CTD 片段第 1 位至第 18 位及第 228 位至第 388 位氨基酸片段，共 9 个肽段，对没有酶切位点的第 19 位至第 227 位氨基酸片段进行微波辅助酸水解方法对 CTD 进行化学方法剪切，共鉴定 53 个肽段，发现 CTD 片段酸水解的断裂位点具有 Y、P、K 及 R 选择性，对两部分肽段样品分别进行质谱分析，CTD 蛋白覆盖率达到 92%；(3) 在体外激酶反应开始后，选取 10、30、120 及 240 分钟等时间点，运用上述方法对 CTD 的磷酸化状态进行跟踪分析，发现 P-TEFb 偏好于磷酸化 CTD 的七肽重复结构的第五位丝氨酸残基，当磷酸化反应进行足够长时间时，第七位丝氨酸残基也被磷酸化，第二位丝氨酸残基可以发生低程度的磷酸化，第 50 个七肽重复序列以后的 31 个氨基酸残基几乎没有被 P-TEFb 磷酸化。

总之，本课题通过构建基于液质联用的检测方法研究 P-TEFb 高度磷酸化 RNA 聚合酶 II CTD 过程，为揭示 CTD 发生磷酸化的位置、频率及 P-TEFb 对 CTD 不同氨基酸的特异性等问题提供了直观且精确的方法。

关键词： RNA 聚合酶 II CTD；P-TEFb；磷酸化

Abstract

The RNA polymerase II is responsible not only for transcription of thousands of protein-coding genes constituting the largest group of distinct individual genes in the eukaryotic genome, but also for genes encoding small noncoding RNAs (snRNA). Furthermore, it is found that the carboxyl-terminal domain (CTD) of RNA polymerase II is unique to this eukaryotic RNA polymerase. TFIIF and P-TEFb are important general transcription factors involved in the progress of transcription cycle. P-TEFb is not only a kind of general transcription factor, but also one of necessary host transcription factors in HIV-1 gene transcription elongation. Thus, the phosphorylation status of CTD by P-TEFb has attracted great attention for a long time. However, its intrinsic structure with hyperphosphorylation and heptapeptide repeat makes it difficult to be analyzed by using mass spectrometry.

In this study, *in vitro* phosphorylation patterns of full-length CTD by P-TEFb has been studied using liquid chromatography-mass spectrometry (LC-MS). Main contents are as summarized as below: (1) P-TEFb was purified through affinity purification and its kinase activity was verified using synthetic peptide as substrate; (2) 9 peptides were detected after tryptic digest in MS analysis, covering 1 to 18 and 228 to 388 amino acids of CTD, meanwhile, remained 19 to 227 amino acids were cleaved by microwave assisted acid hydrolysis method and 53 peptides were then detected, finally the coverage rate after combination of two parts was up to 92%; (3) phosphorylation status of CTD at different time points was revealed by results of ESI-MS analysis. Firstly, Ser-5 phosphorylation was preferred by P-TEFb followed by Ser-7, further more, if reaction time was long enough, low level phosphorylation of Ser-2 could happen, finally, the sequence after fiftieth heptapeptide repeat have barely been phosphorylated.

Generally, in this research, a new method in combination of enzymatic digestion and microwave assisted acid hydrolysis was successfully developed for supplying a new way to study position, frequency and preference of phosphorylation of CTD by P-TEFb directly and accurately.

Keywords: RNA polymerase II CTD; P-TEFb; Phosphorylation

目录

摘要	I
Abstract	II
目录	III
Catalogue	VI
英文缩略对照表	IX
第 1 章 前言	1
1.1 RNA 聚合酶 II 及其负责的真核基因转录	2
1.1.1 真核生物 RNA 聚合酶	2
1.1.2 RNA 聚合酶 II (RNA Pol II) 的结构	2
1.1.3 RNA 聚合酶 II 参与的真核基因转录过程	4
1.1.4 P-TEFb 复合体及其功能	6
1.2 RNA 聚合酶 II CTD 及其磷酸化循环	8
1.2.1 CTD 七肽重复序列的结构	8
1.2.2 CTD 的磷酸化循环	10
1.3 磷酸化蛋白质组学定量技术	12
1.3.1 放射性标记法	13
1.3.2 荧光检测法	13
1.3.3 稳定同位素标记技术	13
1.4 质谱技术分析 CTD 磷酸化模式的研究进展	16
1.5 研究内容和意义	17
第 2 章 实验材料和方法	19
2.1 实验药品、试剂与仪器	19

2.1.1 细胞株、菌株和质粒资源.....	19
2.1.2 主要试剂和耗材.....	19
2.1.3 主要仪器.....	20
2.2 常用溶液配方	22
2.3 实验方法	26
2.3.1 大肠杆菌感受态细胞的制备.....	26
2.3.2 质粒转化.....	27
2.3.3 质粒小量提取.....	27
2.3.4 质粒大量提取.....	28
2.3.5 贴壁细胞的培养.....	28
2.3.6 真核细胞瞬时转染.....	29
2.3.7 全细胞裂解.....	29
2.3.8 免疫亲和纯化.....	30
2.3.9 原核蛋白表达.....	30
2.3.10 GST pulldown	31
2.3.11 AKTA pure 系统纯化蛋白	31
2.3.12 SDS-PAGE 蛋白电泳	32
2.3.13 蛋白免疫印迹实验 (Western blot)	32
2.3.14 考马斯亮蓝 R-250 染色	33
2.3.15 体外激酶反应	33
2.3.16 截留柱辅助的质谱检测蛋白样品制备 (FASP)	33
2.3.17 微波辅助酸水解蛋白 (MAAH)	34
2.3.18 肽段样品除盐	35
2.3.18.1 Ziptip 除盐	35
2.3.18.2 Stop and go extraction tips 除盐.....	35

第 3 章 实验结果与讨论	37
3.1 实验结果	37
3.1.1 P-TEFb 纯化条件的优化	37
3.1.2 GST-CTD 的原核表达及纯化	39
3.1.3 胰酶酶切 GST-CTD	42
3.1.4 微波辅助酸水解法 (MAAH) 消化未酶解氨基酸片段	45
3.1.5 P-TEFb 体外磷酸化 GST-CTD	49
3.2 讨论	63
第 4 章 小结与展望	65
4.1 小结	65
4.2 展望	66
参考文献	67
附录	78
致谢	79

Catalogue

Abstract in Chinese	I
Abstract in English	II
Catalogue in Chinese	III
Catalogue in English	VI
Abbreviation	IX
Chapter 1 Forewords	1
1.1 RNA Polymerase II and transcription mediated by it	2
1.1.1 RNA Polymerase in eukaryotes	2
1.1.2 Structure of RNA Polymerase II (RNA Pol II).....	2
1.1.3 Transcription mediated by RNA Polymerase II	4
1.1.4 P-TEFb complex and its function	6
1.2 RNA Polymerase II CTD and its phosphorylation cycle	7
1.2.1 Structure of CTD heptapeptide repeats.....	7
1.2.2 Phosphorylation cycle of CTD.....	10
1.3 Quantitative phosphoproteomics	12
1.3.1 Radioactive labeling.....	12
1.3.2 Fluorescence assay	13
1.3.3 Stable isotope labeling	13
1.4 Progress of CTD analysis based on mass spectrometry	16
1.5 Contents and significance	17
Chapter 2 Materials and methods	19
2.1 Reagents and instruments	19
2.1.1 Cell lines, E. coli and plasmids	19
2.1.2 Reagents and materials	19

2.1.3 Instruments and expendable supplies.....	20
2.2 Commonly used solution	22
2.3 Methods.....	26
2.3.1 Preparation of competent cells.....	26
2.3.2 Plasmid transformation	27
2.3.3 Mini plasmid extraction	27
2.3.4 Max plasmid extraction.....	28
2.3.5 Adherent cell culture.....	29
2.3.6 Transient transfection.....	29
2.3.7 Whole cell lysis.....	30
2.3.8 Immune affinity purificaion.....	30
2.3.9 Protein expression in eukaryotic cells.....	30
2.3.10 GST pulldown.....	31
2.3.11 Purification of proteins through AKTA pure system.....	32
2.3.12 SDS-PAGE electrophoresis	32
2.3.13 Western blot.....	32
2.3.14 Commassie staining with R-250	33
2.3.15 In vitro kinas assay.....	33
2.3.16 Filter-assisted sample preparation (FASP)	34
2.3.17 Microwave-assisted acid hydrolysis (MAAH)	35
2.3.18 Desalting before mass spectrometry analysis	35
2.3.18.1 Desalting using Ziptip.....	35
2.3.18.2 Desalting using stop and go extraction tips	36
Chapter 3 Results and discussion.....	37
3.1 Results	37
3.1.1 Affinity purification of P-TEFb and its activity	37
3.1.2 Expression and purification of GST-CTD	39
3.1.3 Digestion of GST-CTD with trypsin.....	42

3.1.4 Microwave assisted acid hydrolysis (MAAH).....	45
3.1.5 P-TEFb phosphorylating GST-CTD in vitro	49
3.2 Discussion	63
Chapter 4 Summary and outlook.....	65
4.1 Summary.....	65
4.2 Outlook.....	66
Reference	67
Appendix.....	78
Acknowledgement.....	79

英文缩略对照表

Abbreviation	Full Name
7SK snRNP	7SK small nuclear ribonucleoprotein, HEXIM1/7SK/LARP7/MePCE/P-TEFb
TF	Transcription factor
ActD	Actinomycin D
AIDS	Acquired immunodeficiency syndrome
TBP	TATA-binding protein
TAF	TBP-associated factor
CTD	C-terminal domain
DMEM	Dulbecco's modified of Eagle's medium
DSIF	DRB-sensitivity inducing factor
ESI	Electrospray ionization
FASP	Filter aided sample preparation
HIV-1	Human immunodeficiency virus type 1
HMBA	Hexamethylene bisacetamide
IT	Iontrap
LTR	Long terminal repeats
MALDI	Matrix assisted laser desorption ionization
NELF	Negative elongation factor

英文缩略对照表

P-TEFb	Positive transcriptional elongation factor b
PIC	Preinitiation complex
RNA Pol I	RNA Polymerase I
RNA Pol II	RNA polymerase II
RNA Pol III	RNA Pol III
snRNP	Small nuclear ribonucleoprotein particle
SEC	Super elongation complex, AFF1/4、 ELL2、 ENL/AF9、 P-TEFb
SILAC	Stable isotope labeling with amino acids in cell culture
TAR	Transacting-response
TEC	Transcript elongation complex
TOF	Time of flight

第 1 章 前言

在真核细胞中，RNA 聚合酶 II (RNA Pol II) 负责大多数编码蛋白质的 mRNA 和核内小 RNA (snRNA) 基因的转录，这个过程由一系列复杂而有序的调控机制完成。从真菌到人类，基因转录都由四个基本的步骤组成：转录前起始复合物的组装、转录起始、转录延伸和转录结束。细胞应答信号、基因转录的调节一直以来被认为发生在转录起始前复合物的组装和调控阶段，但之后人们发现转录延伸也存在着高度有序且精细的调控方式，甚至这一阶段的调控机理比转录起始阶段更加复杂，这与正性转录延伸因子 b (Positive transcription elongation factor b, P-TEFb) 的功能密切相关。

作为影响细胞生长与分化的基本转录因子之一，P-TEFb 不仅与心肌肥大和癌症的发生有直接的关系，也是艾滋病毒 (HIV-1) 转录复制过程中不可或缺的宿主转录因子之一。RNA 聚合酶 II 在结合转录循环和基因表达中扮演着重要的角色，它的最大亚基 Rpb1 的 C 端由一段七肽重复序列 Tyr-Ser-Pro-Thr-Ser-Pro-Ser 组成，通过在这段序列的不同位点上发生磷酸化而产生的结合表面的多样性介导了其一系列细胞核因子的相互作用，在转录过程中，RNA 聚合酶 II CTD 的 5 位丝氨酸残基磷酸化后启动子区彻底暴露、转录暂停、加帽开始进行，此时通过 P-TEFb 中的 CDK9 磷酸化 RNA 聚合酶 II CTD 的 2 位丝氨酸，从而使中止的转录过程得以继续并偶联 RNA 成熟和加工机制。因此，本文将建立一种依赖于质谱的研究 RNA 聚合酶 II CTD 的磷酸化修饰的分析方法，进一步了解 P-TEFb 磷酸化 RNA 聚合酶 II CTD 的模式，包括 CTD 的 52 个七肽重复序列发生磷酸化的数量，发生单磷酸化或双磷酸化的频率，以及 P-TEFb 对 CTD 磷酸化的特异性。

1.1 RNA 聚合酶 II 及其负责的真核基因转录

1.1.1 真核生物 RNA 聚合酶

转录反应，即在 DNA 指导下 RNA 的合成，是通过 RNA 聚合酶来实现的。在原核生物中，一种 RNA 聚合酶足以转录原核细胞所需要的所有 RNA 分子。然而在真核生物中，RNA 聚合酶有 I、II、III 三类来完成细胞中基因的转录工作，它们的相对分子质量均在 500 000 左右，通常有 8~14 个亚基，含有 Zn^{2+} 。

根据对 α -鹅膏蕈碱敏感程度的不同，可以对三类 RNA 聚合酶进行区分：RNA 聚合酶 I 对 α -鹅膏蕈碱不敏感，它负责转录前体 45 S rRNA，经转录后加工产生 5.8 S rRNA、18 S rRNA、28 S rRNA；RNA 聚合酶 II 可以被低浓度的 α -鹅膏蕈碱 (10^{-9} ~ 10^{-8} M) 抑制，它负责转录所有前体 mRNA 和大多数核内小 RNA (snRNA)；RNA 聚合酶 III 可以被高浓度的 α -鹅膏蕈碱 (10^{-5} ~ 10^{-4} M) 抑制，它负责转录 tRNA、5 S rRNA、U6 snRNA 和不同的胞质小 RNA (scRNA) 等小分子转录物。

1.1.2 RNA 聚合酶 II (RNA Pol II) 的结构

真核生物 RNA 聚合酶 II (RNA Pol II) 由 10 个亚基组成的核心酶及 Rpb4/Rpb7 组成的异源二聚体构成，分子量为 514 kDa^[1-3] (表 1.1)。核心酶的组成亚基包括 Rpb1, Rpb2, Rpb3, Rpb5, Rpb6, Rpb8, Rpb9, Rpb10 及 Rpb12。

2000 年，Patrick Cramer 等通过 X 射线衍射得到了酵母 RNA 聚合酶 II 核心酶的骨架模型^[4]。这个骨架模型显示，RNA 聚合酶 II 最大的两个亚基 Rpb1 和 Rpb2 处于中心位置，形成了 RNA 聚合酶 II 的活性中心，其它较小的亚基分布排列在其周围。此外有趣的是，RNA 聚合酶 II 有一条“尾巴”，这条“尾巴”是 Rpb1 亚基羧基端 (C 端) 的一段灵活连接在核心酶上的特殊氨基酸序列，称为羧基端结构域 (C-terminal domain, CTD)。在老鼠、果蝇或者酵母中，敲除 CTD 具有致死性，说明这一结构对生命活动有着重要作用。然而，在体外实验

或瞬时转染的模板中，转录均可在没有 CTD 的情况下进行^[5-7]，表明 CTD 发挥着辅助性作用。

表 1.1 RNA 聚合酶 II 亚基^[8,9]

RNA 聚合酶 II 亚基	结构特点	功能作用
Rpb1	RNA 聚合酶 II 的最大亚基；包含 CTD 结构	参与结合 DNA；与其它 Rpb9 结合，形成转录反应进行的凹槽结构
Rpb2	RNA 聚合酶 II 的第二大亚基；	与其它亚基一起负责维系 DNA 模板与新合成 RNA 的结合
Rpb3	RNA 聚合酶 II 的第三大亚基	与 Rpb11 组成的异源二聚体构成聚合酶的核心组件
由 10 个亚基组成的核心酶 Rpb5	每个 RNA 聚合酶 II 中有两个 Rpb5 亚基	与 Rpb1、Rpb3、Rpb6 相互作用
Rpb6	—	与其它亚基一起稳定转录过程中的 RNA 聚合酶 II
Rpb8	—	与 Rpb1、Rpb2、Rpb3、Rpb5、Rpb7 相互作用
Rpb9	—	与 Rpb1 一起组成转录反应进行的凹槽结

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库