

DOI:10.13350/j.cjpb.221103

• 论著 •

## 结核分枝杆菌 PPE2 的生物信息学分析\*

冯楠<sup>1</sup>, 柳小玲<sup>1</sup>, 张杰<sup>2</sup>, 董江涛<sup>2</sup>, 梁粟<sup>2</sup>, 王菊<sup>1</sup>, 张辉<sup>1</sup>, 丰启盈<sup>1</sup>, 邹炎臻<sup>1</sup>, 吴江东<sup>1</sup>, 章乐<sup>1</sup>, 吴芳<sup>1\*</sup>, 张万江<sup>1\*\*</sup>

(1. 石河子大学医学院病理生理学教研室, “新疆地方与民族高发病”教育部重点实验室, 新疆石河子 832002;

2. 石河子大学医学院第一附属医院)

**【摘要】** 目的 应用生物信息学软件预测分析结核分枝杆菌 PPE2 蛋白的结构与功能。方法 从 NCBI 中获得 PPE2 序列及其编码基因序列, 通过 ORF Finder、ProtParam、ProtScaleon Expsy、TMHMM Sever v. 2. 0、SignalP 4. 1 Sever、PSORT、NetPhos 3. 1 Sever、BLAST、SOPMA、SWISS-MODEL、ABCpred、SYFPEITHI 等工具预测分析 PPE2 编码蛋白的生物学结构与特性。结果 PPE2 蛋白由 556 个氨基酸组成, 相对分子质量为  $56\ 951.97 \times 10^3$ , 等电点 4. 78, 为亲水性蛋白, 无跨膜区域及信号肽; 编码基因 Rv0256c 总长度为 1 671 bp, 有 10 个开放阅读框架; PPE2 蛋白有 53 个磷酸化位点, 1 个保守结构域; 蛋白二级结构中  $\alpha$ -螺旋占 41. 91%,  $\beta$ -折叠占 11. 33%,  $\beta$ -转角占 3. 96%, 无规则卷曲占 42. 81%; 有 57 个(得分 > 0. 5) B 细胞抗原表位和 113 个(得分 > 15) T 细胞抗原表位。结论 PPE2 蛋白为亲水蛋白, 具有潜在的 T、B 细胞抗原表位, 可作为研发结核病疫苗的候选蛋白。

**【关键词】** PPE2 蛋白; 结核分枝杆菌; Rv0256c; 生物信息学

**【中图分类号】** R378. 911

**【文献标识码】** A

**【文章编号】** 1673-5234(2022)11-1252-04

[Journal of Pathogen Biology. 2022 Nov. ;17(11):1252-1255,1260.]

Bioinformatics analysis of *Mycobacterium tuberculosis* PPE2

FENG Nan<sup>1</sup>, LIU Xiao-ling<sup>1</sup>, ZHANG Jie<sup>2</sup>, DONG Jiang-tao<sup>2</sup>, LIANG Su<sup>2</sup>, WANG Ju<sup>1</sup>, ZHANG Hui<sup>1</sup>, FENG Qi-ying<sup>1</sup>, ZOU Zhi-zhen<sup>1</sup>, WU Jiang-dong<sup>1</sup>, ZHANG Le<sup>1</sup>, WU Fang<sup>1</sup>, ZHANG Wan-jiang<sup>1</sup> (1. Department of Pathophysiology, Shihezi University School of Medicine/Xinjiang High Incidence of Local and Ethnic Diseases, Key Laboratory of the Ministry of Education, Shihezi, Xingjiang 832002, China; 2. The First Affiliated Hospital of Medical College of Shihezi University)

**【Abstract】** **Objective** To predict and analyze the structure and function of the *Mycobacterium tuberculosis* PPE2 protein by using various bioinformatics software. **Methods** PPE2 gene(Rv0256c) and its coding sequence were obtained from NCBI. The ORF Finder was used to find the open reading frames of the Rv0256c gene. ProtParam was used to analyze the physicochemical properties, and ProtScaleon Expsy was used to analyze the hydrophilicity of the PPE2 protein. The transmembrane regions, the signal peptides, the subcellular localizations and the phosphorylation sites of the protein were predicted by TMHMM Sever V. 2. 0, SignalP 4. 1 Sever, PSORT and NetPhos 3. 1 Sever. The BLAST tool was used to predict its conserved domain and homology. SOPMA and Swiss-Model were used to predict the secondary structure and 3D model of the protein. ABCpred and SYFPEITHI were used to predict its T-cell and B-cell epitopes.

**Results** PPE2 protein was composed of 556 amino acids with an isoelectric point of 4. 78. It was a hydrophilic protein without transmembrane region and signal peptide. The coding gene Rv0256c was 1 671 bp in length and had 10 open reading frames. PPE2 protein had 53 phosphorylation sites and 1 conserved domain. In protein secondary structure,  $\alpha$ -helix accounted for 41. 91%,  $\beta$ -folding accounted for 11. 33%,  $\beta$ -turning accounted for 3. 96%, random crimping accounted for 42. 81%. In the 3D model, the QmeaDisCo global score was  $0.64 \pm 0.07$ . There were 57 B cell epitopes (score > 0. 5) and 113 T cell epitopes (score > 15) in the protein. **Conclusion** PPE2 protein is a water protein with potential T and B cell epitopes, which can be used as a candidate protein for developing tuberculosis vaccine.

**【Key words】** PPE2 protein; *Mycobacterium tuberculosis*; Rv0256c; bioinformatics

\*\*\* 结核病是结核分枝杆菌(MTB)引起的慢性传染性疾病, 是全球传染病致死的主要原因之一<sup>[1]</sup>。与其他病原体不同的是, MTB 不产生毒素, 主要通过其荚膜、脂类及蛋白发挥致病作用, MTB 基因组中包括许多表达蛋白质的基因, 其中许多表达蛋白是与其致病性相关的毒力因子<sup>[2]</sup>。

\* **【基金项目】** 国家自然科学基金项目(No. 82060021); NSFC-新疆联合基金项目(No. U1803127); 兵团重点领域科技攻关计划项目(No. 2018AB019); 石河子大学 2020 年度自主资助支持校级科研立项项目(No. ZZZC202013A)。

\*\* **【通讯作者】** 吴芳, E-mail: 1743142853@qq.com;  
张万江, E-mail: E-mail: zwj1117@126.com

**【作者简介】** 冯楠(1989-), 男, 新疆石河子人, 硕士研究生, 主要从事感染性疾病的诊断与防治研究。E-mail: 244920124@qq.com

近年研究发现,PE/PPE 基因家族所编码的蛋白对于 MTB 的毒力与免疫机制起关键作用,已经证明有几种 PE/PPE 蛋白能够通过直接与 Toll 样受体结合以调节宿主免疫系统。同时,该蛋白家族也是宿主适应性 T 细胞反应的主要靶点<sup>[3]</sup>。PE/PPE 蛋白家族也是潜在的候选疫苗抗原<sup>[4]</sup>。该蛋白家族成员 PPE2 能通过影响宿主正常生物学进程,利于 MTB 在宿主体内存活,但其结构与功能尚不清楚。本研究通过对 PPE2 基因编码蛋白进行生物信息学分析,预测其结构与生物学功能,以期作为结核病疫苗的候选蛋白。

## 材料与方法

### 1 基因信息

从美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI)获取 MTB 的基因组相关信息,其中 Rv0256c 基因是 PPE2 的编码基因。

### 2 生物信息学分析

通过 NCBI 的 ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>)工具分析 Rv0256c 基因的开放阅读框架;通过 ProtParam 工具 (<https://web.expasy.org/protparam/>)预测分析 Rv0256 基因编码 PPE2 蛋白的理化性质;通过 ProtScale (<https://web.expasy.org/protscale/>)预测分析 PPE2 蛋白的亲水性;通过 TMHMM Server v. 2.0 (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>)预测分析 PPE2 蛋白跨膜区;通过 SignalP 4.1 Server (<https://services.healthtech.dtu.dk/service.php?SignalP-6.0>)预测分析 PPE2 蛋白信号肽,通过 PSORT (<https://www.genscript.com/psort.html>)预测分析 PPE2 蛋白亚细胞定位;通过 NetPhos 3.1 Server (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>)预测分析 PPE2 蛋白磷酸化位点,通过 NCBI 中 BLAST 工具 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)预测分析 PPE2 蛋白保守域及同源性;通过 SOPMA ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html))预测分析 PPE2 蛋白二级结构,通过 SWISS-MODEL (<https://www.expasy.org/resources/swiss-model>)预测分析 PPE2 蛋白三级结构;通过 ABCpred ([https://webs.iitd.edu.in/raghava/abcpred/ABC\\_submission.html](https://webs.iitd.edu.in/raghava/abcpred/ABC_submission.html))预测分析 PPE2 蛋白 B 细胞抗原表位,通过 SYFPEITHI (<https://cn.string-db.org/>)预测分析 PPE2 蛋白 T 细胞抗原表位。

## 结果

### 1 Rv0256c 基因及其编码蛋白的基本信息和开放阅读框架

NCBI 数据显示,Rv0256c 基因的 ID 为 886684,编码蛋白编号为 YP\_177704.1,所在基因组登录号为 NC\_000962.3。该基因全长为 1 671 bp,位于基因组的第 307 877-309 547 位。起始密码子为 ATG,终止密码子为 TGA。Rv0256c 基因共 10 个开放阅读框架(图 1)。

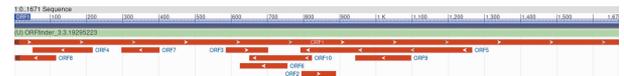


图 1 Rv0256c 基因开放阅读框架  
Fig. 1 Open reading frame of Rv0256c gene

### 2 PPE2 蛋白的理化性质

利用 ProtParam 分析 PPE2 蛋白的氨基酸组成,其中丙氨酸(Ala)占 18.5%,甘氨酸(Gly)占 7.9%,亮氨酸(Leu)占 9.4%,氨基酸总数为 556。预测其相对分子质量为  $56\ 951.97 \times 10^3$ ,分子式为  $C_{2589}H_{4005}N_{661}O_{761}S_{12}$ ,等电点为 4.78,消光系数(280 nm)为 74 370。蛋白质不稳定系数为 43.73,为不稳定蛋白。

### 3 PPE2 蛋白亲水性

ProtParam 预测 PPE2 蛋白的脂肪族指数为 86.62,亲水性平均值为 0.191,预测为亲水蛋白。ProtScale 分析 PPE2 蛋白为亲水蛋白(图 2),其中第 374 位氨基酸亲水性得分最高,为 2.178;第 474 位氨基酸亲水性得分最低,为 -3.567。

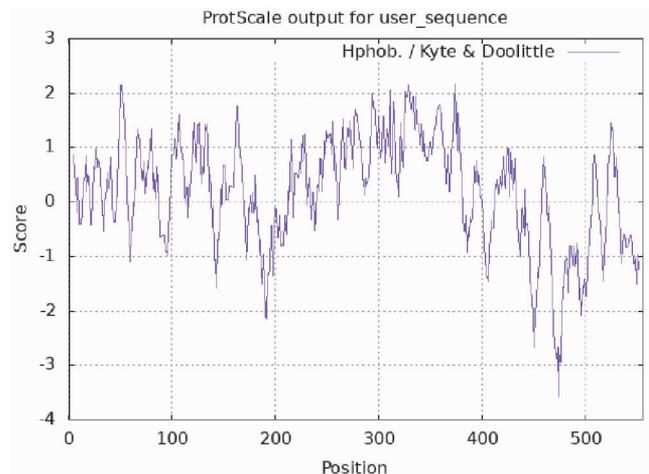


图 2 PPE2 蛋白的亲水性分析  
Fig. 2 Hydrophilicity of PPE2 protein

### 4 PPE2 蛋白跨膜区

应用 TMHMM Server v. 2.0 预测 PPE2 蛋白跨膜螺旋数为 0(图 3),跨膜氨基酸数为 83.956 29(>18



表 1 PPE2 蛋白 B 细胞抗原表位预测  
Table 1 Epitope prediction of B cell of PPE2 protein

等级 Rank	序列 Sequence	开始位置 Start position	得分 Score
1	LQKIGYTDYFNNVIQP	200	0.95
2	GGSPRTPMMPGTWDTD	535	0.94
2	HGVLMATNFFGINTIP	120	0.94
3	TGVRKKTPEPDSAEAP	444	0.92
3	PTPPGPPPPPVTAAPP	400	0.92
4	APTAVTAPTPPPGPPP	393	0.9
4	PVAAIAPSIPTPTTP	373	0.9
4	RMWQAATTMASYQAV	145	0.9
5	AGLITLPNDAFGGSPR	524	0.89
5	TAPIWMASPPEVHSAL	2	0.89
6	TGSPQGAGTLGFAGTT	500	0.88
6	MFSGFDPWLPGLNPL	229	0.88
7	RGYEYLDLDPETGHP	484	0.87
7	AQARKAVGTGVRKKTTP	436	0.87
7	APQIVKANAPTAASDE	174	0.87
8	PMMPGTWDTDSATRVE	541	0.86
8	LGYPMDIGSYVAFLSQ	256	0.86
9	AWLVQASANSAAMATR	77	0.85
9	TPEPDSAEAPAAAAAP	450	0.85
9	VSAEGWHLSIAAYET	29	0.85

表 2 PPE2 蛋白 HLA-A \* 02:01 限制性 CTL 抗原表位预测  
Table 2 Prediction of HLA-A \* 02:01 restricted CTL epitope of PPE2 protein

位置 Pos	序列 (123456789)	得分 Score
332	LLAAPLAPL	31
317	TLLEQTLAL	28
47	ELAALLAAV	27
105	AAMPTLAEL	27
212	VIQPFINWL	27
239	SLGNPLTFL	27
318	LLEQTLALL	27
324	ALLPAALPL	27
325	LLPAALPLL	27
307	HTDTIALV	26
322	TLALLPAAL	26
329	ALPLLAAPL	25
346	SAAGGFAGL	25
427	YLVGDLNSA	25

结核分枝杆菌通过调节多种代谢相关的基因表达,进而在宿主环境中创造利于自身生存的条件。Ppe2-cobq1-cobu 操纵子是基因调控的罕见例子,其中 PE/PPE 家族基因序列位于该操纵子中<sup>[7]</sup>,有研究证明该操纵子的启动子区域位于 PPE2 基因的上游(300 bp 以内)<sup>[8]</sup>。但由于 PPE2 蛋白结构与功能未知,尚无法阐明其与 Ppe2-cobq1-cobu 操纵子的功能相关性。已有研究报道 MTB 的 PE/PPE 蛋白家族成员中有超过 35 种定位于分枝杆菌细胞膜或细胞壁上,与细胞壁相关的蛋白可能介导病原体与宿主的相互作用,所以细胞定位可能参与调控 MTB 不同的生物学功能<sup>[9]</sup>。本研究预测 PPE2 蛋白可能为非跨膜蛋白,且

是一个膜外蛋白,是由 Rv0256c 编码的细胞膜相关蛋白,暴露于细胞表面,可能通过抑制活性氮氧化物的产生发挥生物学效应<sup>[5]</sup>。PPE2 蛋白有多个磷酸化位点,说明其可能参与细胞间的信号传导。 $\alpha$ -螺旋在 DNA 模体中起重要作用,由于模体序列表现较为保守, $\alpha$ -螺旋占比高的蛋白多为保守蛋白。本研究预测 PPE2 蛋白二级结构中  $\alpha$ -螺旋占 41.91%,表明该蛋白可能具有功能性。无规则卷曲是蛋白质重要的功能区域<sup>[10]</sup>,SOPMA 软件预测 PPE2 蛋白无规则卷曲占 42.81%,同样说明该蛋白可能为重要的功能蛋白。已有研究表明,PPE 蛋白在结核分枝杆菌感染进程中参与影响宿主体液与细胞免疫的生理过程<sup>[11]</sup>。Pal 等<sup>[12]</sup>报道了一种 PPE 家族蛋白,即 PPE2 蛋白,在感染条件下具有影响成人造血的作用。这些细胞数量的减少可能会在宿主内部创造一个免疫剥夺的环境,因此可能会增加结核病的易感性。PPE2 表达的耻垢分枝杆菌在体外和体内的细胞内存活时间得到延长<sup>[13]</sup>,PPE2 可通过抑制 iNOS 基因转录来限制一氧化氮(NO)的产生。PPE2 蛋白具有一个亮氨酸拉链 DNA 结合基序和一个功能性的核定位信号,通过经典的 Importin $\alpha/\beta$  途径进入巨噬细胞核,与 iNOS 启动子的 TATA 盒重叠的 GATA 结合部位相互作用,抑制 NO 的产生。这些信息提示结核分枝杆菌感染过程中,PPE2 在宿主与病原体的相互作用中十分关键。

Delogu 等<sup>[14]</sup>报道 MTB 的毒力因子,特别是表面成分已成为疫苗制备的潜在靶标。含有 B 细胞抗原表位的蛋白可以合成多肽,替代完成抗原诱导机体产生 B 细胞抗体,用以研制疫苗<sup>[15]</sup>。在结核病保护性免疫应答中 T 细胞具有重要作用,所以含 T 细胞表位的 MTB 蛋白有可能成为疫苗制备的候选蛋白<sup>[16]</sup>。PPE 家族中 PPE42 为抗原的亚单位疫苗 ID93/GLA-SE 已经进入临床实验<sup>[17]</sup>,证实了 PPE 家族蛋白具有成为疫苗抗原的可能性。本研究预测 PPE2 含有得分 > 0.85 的 B 细胞表位 20 个,>25 分的 T 细胞表位有 14 个,该蛋白含有较多高分预测表位,可作为结核病疫苗候选抗原。

Rv0256c 基因编码的 PPE2 蛋白具有多种生物学功能,但相关机制和通路尚不清楚。本研究对 PPE2 蛋白进行生物信息学分析,预测了该蛋白结构与性质,可为该蛋白的相关研究提供基础资料。

【参考文献】

[1] Liu J, Chen XF, Wang J, et al. Prediction and identification of CD4<sup>+</sup> T cell epitope for the protective antigens of *Mycobacterium tuberculosis*[J]. Medicine, 2021, 100(6): e24619.

(下转 1260 页)

- infection[J]. Infection,2018,46(3):303-315.
- [8] Contera S, Bernardino de la Serna J, Tetley TD. Biotechnology, nanotechnology and medicine[J]. Emerg Top Life Sci, 2020, 4(6):551-554.
- [9] Wu J, Wang X, Wang Q, et al. Nanomaterials with enzyme-like characteristics (nanozymes): next-generation artificial enzymes (II)[J]. Chem Soc Rev, 2019, 48(4):1004-1076.
- [10] Sun FY, Liang YY, Jin LH, et al. Weak interaction-tailored catalytic interface of ultrasmall gold nanoclusters as enzyme mimics for enhanced colorimetric biosensing[J]. ACS Appl Mater Interfaces, 2021, 13(48):58209-58219.
- [11] Long YJ, Li YF, Liu Y, et al. Visual observation of the mercury-stimulated peroxidase mimetic activity of gold nanoparticles[J]. Chem Commun (Camb), 2011, 47(43):11939-11941.
- [12] Ni P, Chen C, Jiang Y, et al. Gold nanoclusters-based dual-channel assay for colorimetric and turn-on fluorescent sensing of alkaline phosphatase [J]. Sensor Actuat B-Chem, 2019, 301:127080.
- [13] Hu DH, Sheng ZH, Fang ST, et al. Folate Receptor-targeting gold nanoclusters as fluorescence enzyme mimetic nanoprobe for tumor molecular colocalization diagnosis[J]. Theranostics, 2014, 4(2):142-153.
- [14] Weeratunga P, Herath TUB, Kim TH, et al. Dense granule protein-7 (GRA-7) of *Toxoplasma gondii* inhibits viral replication *in vitro* and *in vivo*[J]. J Microbiol, 2017, 55(11):909-917.
- [15] Jacobs D, Vercammen M, Saman E. Evaluation of recombinant dense granule antigen 7 (GRA7) of *Toxoplasma gondii* for detection of immunoglobulin G antibodies and analysis of a major antigenic domain[J]. Clin Diagn Lab Immunol, 1999, 6(1):24-29.
- [16] Rhee SG. Cell signaling. H<sub>2</sub>O<sub>2</sub>, a necessary evil for cell signaling[J]. Science, 2006, 312(5782):1882-1883.
- 【收稿日期】 2022-06-18 【修回日期】 2022-09-10

(上接 1255 页)

- [2] 王璞, 蔡玉荣, 张刚, 等. 结核分枝杆菌 PE/PPE 蛋白家族生物学研究进展[J]. 中国人兽共患病学报, 2021, 37(10):937-944.
- [3] Delogu G, Brennan MJ, Manganelli R. PE and PPE Genes: A Tale of Conservation and Diversity [J]. Adv Exp Med Biol, 2017(1019):191-207.
- [4] Brennan MJ. The enigmatic PE/PPE multigene family of mycobacteria and tuberculosis vaccination[J]. Infect Immun, 2017, 85(6):e00969-16.
- [5] Srivastava S, Battu MB, Khan MZ, et al. *Mycobacterium tuberculosis* PPE2 protein interacts with p67<sup>bhox</sup> and inhibits reactive oxygen species production[J]. J Immunol, 2019, 203(5):1218-1229.
- [6] Oliver GR, Hart SN, Klee EW. Bioinformatics for clinical next generation sequencing[J]. Clin Chem, 2015, 61(1):124-135.
- [7] Young DB, Comas I, de Carvalho LP. Phylogenetic analysis of vitamin B12-related metabolism in *Mycobacterium tuberculosis* [J]. Front Mol Biosci, 2015(2):6.
- [8] Srivastava S, Mukhopadhyay S. *Mycobacterium tuberculosis* protein PPE2 binds to DNA region containing promoter activity [J]. Biochem Biophys Res Commun, 2021(567):166-170.
- [9] Wang Q, Boshoff HIM, Harrison JR, et al. PE/PPE proteins mediate nutrient transport across the outer membrane of *Mycobacterium tuberculosis* [J]. Science, 2020, 367(6482):1147-1151.
- [10] 张西燕, 付玉荣, 伊正君. 结核分枝杆菌 *eis* 基因及其编码蛋白的生物信息学分析[J]. 中国病原生物学杂志, 2018, 13(2):140-145, 151.
- [11] Mukhopadhyay S, Balaji KN. The PE and PPE proteins of *Mycobacterium tuberculosis* [J]. Tuberculosis, 2011, 91(5):441-447.
- [12] Pal R, Mukhopadhyay S. PPE2 protein of *Mycobacterium tuberculosis* affects myeloid hematopoiesis in mice [J]. Immunobiology, 2021, 226(1):152051.
- [13] Bhat KH, Srivastava S, Kotturu SK, et al. The PPE2 protein of *Mycobacterium tuberculosis* translocates to host nucleus and inhibits nitric oxide production[J]. Sci Rep, 2017(7):39706.
- [14] Delogu G, Provvedi R, Sali M, et al. *Mycobacterium tuberculosis* virulence: insights and impact on vaccine development[J]. Future Microbiol, 2015, 10(7):1177-1194.
- [15] Cornaby C, Gibbons L, Mayhew V, et al. B cell epitope spreading: mechanisms and contribution to autoimmune diseases [J]. Immunol Lett, 2015, 163(1):56-68.
- [16] Coscolla M, Copin R, Sutherland J, et al. *Mycobacterium tuberculosis* T cell epitope analysis reveals paucity of antigenic variation and identifies rare variable TB antigens[J]. Cell Host Microbe, 2015, 18(5):538-548.
- [17] Penn-Nicholson A, Tameris M, Smit E, et al. Safety and immunogenicity of the novel tuberculosis vaccine ID93+GLA-SE in BCG-vaccinated healthy adults in South Africa: a randomised, double-blind, placebo-controlled phase 1 trial[J]. Lancet Respir Med, 2018, 6(4):287-298.
- 【收稿日期】 2022-05-25 【修回日期】 2022-08-16