

## 【论著摘要】

文章编号:1000-7423(2003)-01-0057-01

## 旋毛虫新生幼虫期特异性基因丝氨酸蛋白酶的结构预测与功能分析

任瑞文<sup>1</sup> 徐晓立<sup>2</sup> 卢强<sup>3</sup> 刘明远<sup>3</sup>

中图分类号: R383.15 R392.11

文献标识码: B

我国旋毛虫病患者数及死亡率居三大食源性人兽共患寄生虫病之首<sup>[1]</sup>。旋毛虫新生幼虫肠型期(NBL)为非细胞内寄生并在血液中移行,对宿主的免疫作用较敏感。利用基因工程技术制备 NBL 抗原用于旋毛虫病的诊断和预防具有重要意义。本文对旋毛虫新生幼虫 cDNA 文库进行筛选,获得其特异性全长基因,采用生物信息学原理对其进行分析,结果报告如下。

## 1 材料与方法

1.1 材料 NBL 特异性全长 cDNA (编号: SSC2): 由 NBL 特异性探针从其 cDNA 文库调取。克隆入 λZAP 载体,由解放军军需大学动科系刘明远博士提供<sup>[2]</sup>。

1.2 基因片段同源性分析 登录 NCBI BLAST 搜索基因和蛋白数据库,进行核酸、蛋白同源性分析。登录 PROSITE 数据库进行蛋白位点和序列模式分析。

1.3 编码蛋白理化特性分析 用蛋白序列分析软件包 ANTHEPROT 4.3,分析其编码蛋白氨基酸组成、潜在信号肽断裂位点、疏水性及抗原性等理化特性。

1.4 编码蛋白结构预测 通过蛋白序列分析软件包 ANTHEPROT 4.3,分别采用 GOR 1、Gibrat、DPM、Predator 算法,对其二级结构进行预测。登录 SWISS-MODEL 自动蛋白同源建模服务器进行三维结构预测。

## 2 结果

2.1 NCBI BLASTn 基因数据库同源性检索 未发现该序列与其他生物已知 DNA 序列有明显的相似性。

2.2 NCBI BLASTp 蛋白数据库同源性检索 结果表明该编码蛋白与旋毛虫丝氨酸蛋白酶相似性高达 38%。登录 PROSITE 数据库进行蛋白位点和序列模式分析,发现两个丝氨酸蛋白酶活性中心保守序列 (histidine active site 129~134 LTAAHC 和 serine active site 269~280 DSCQGDSGGPLI),根据丝氨酸蛋白酶的归类准则将其归类为胰蛋白酶家族丝氨酸蛋白酶。

2.3 蛋白理化特性分析 结果显示,其编码蛋白分子量为 50.1 kDa,等电点为 9.735。氨基酸组成带负电荷氨基酸 50 个,占 10.8%,带正电荷氨基酸 26 个,占 5.6%,整个蛋白带正电荷 24。疏水性氨基酸 99 个,占 21.3%。N 末端第 31~64 个氨基酸以疏水性为主,其亲水性区域位于 C 末端。前 75 个氨基酸区域抗原性较弱,其抗原决定簇主要集中在第 191~460 氨基酸区域。

2.4 二级结构预测 结果显示,氨基酸区域 1~64 以 α 螺旋

及 β 折叠为主,第 64~128 以 β 转角及松散结构为主,第 126~348 α 螺旋、β 转角和 β 折叠分布较均匀,结构复杂;第 348~464 表现为高度无规则卷曲。

2.5 三维结构预测 登录 SWISS-MODEL 自动蛋白质同源建模服务器,搜索、优化后获得其三维结构模型。理论上,柔性区的存在有利于蛋白质两侧功能区的正确折叠,使接头两侧的多肽片断和氨基酸侧链能够自由运动,形成具有生物活性的空间结构。本文其疏水性凹陷区周围及两结构域的绞合区原子的可移动性相对要大于其他部分。

## 3 讨论

分析此编码蛋白的 N 末端前 31~64 个氨基酸,以疏水性氨基酸为主形成一疏水性核心,在此区域的 N 末端有一小段带正电核的序列 (Lys<sup>+</sup>-Met-Ile-Arg<sup>+</sup>-Arg<sup>+</sup>),C 末端有一信号肽酶识别位点 (A+任意氨基酸)。提示可能为信号肽断裂位点。64~126 间 62 个氨基酸中 Pro/Ser/Gly 有 22 个,均匀分布,结构上表现为无规则卷曲,推测为结构域的绞合区。126~348 间 α 螺旋、β 转角和 β 折叠分布较均匀,其 lys、Arg 含量较高,Lys 常与结合功能有关,Arg 常与催化功能有关,推测可能为重要的功能区。不同的酶,其活性中心虽有某些共同的氨基酸残基(如: His、Ser、Cys、Asp 等),但邻近氨基酸序列却很少相同。

本文对 PROSITE 数据库检索发现,SSC2 编码蛋白含有两处保守序列: 129~134 处有一组氨酸活性位点信号区 LTAAHC, 269~280 处有一丝氨酸活性位点信号区 DSCQGDSGGPLI,同时具有丝氨酸蛋白酶活性中心高度保守的氨基酸序列 G-D-S-G-G-P (273~278),因此初步将其归为丝氨酸蛋白酶。

丝氨酸蛋白酶是广泛存在于蠕虫类的蛋白水解酶,旋毛虫新生幼虫无坚硬的口器,入侵宿主很可能是通过分泌特殊蛋白酶类。旋毛虫第 I 期幼虫是经小肠侵入肠系膜淋巴结或血液,通过组织,最终侵入骨骼肌。旋毛虫幼虫期特异性基因的表达产物丝氨酸蛋白酶可能与此过程有关。Todorova<sup>[3]</sup>等用宿主 IgG 使此酶失活之后,发现虫体在敏感宿主体内的存活率大大降低。表明该酶对虫体在宿主体内存活起重要作用,在旋毛虫的免疫中可能起重要作用。

## 参 考 文 献

[1] 刘明远. 我国旋毛虫流行病学概况与研究前瞻[J]. 中国农业大学学报, 1998, 3: 102-104.

[2] 刘明远, 徐克诚, 黎诚耀, 等. 中国旋毛虫分离株成虫和新生幼虫基因文库的构建及筛选[J]. 中国兽医学报, 1998, 18: 147-150.

[3] Todorova VK, Stoyanov DI. Partial characterization of serine proteinases secreted by adult *Trichinella spiralis*[J]. Parasitol Res, 2000, 86: 684-687.

基金项目: 国家自然科学基金(NSFC-39670561, NSFC-39770565)及中法先进研究计划(BT96-12)资助项目

作者单位: 1 广州军区联勤部军事医学研究所, 广州 510507;

2 新疆塔里木农垦大学动物科技学院, 阿拉尔 843300;

3 解放军军需大学动物医学科学系, 长春 130062

(收稿日期: 2001-09-24 编辑: 富秀兰)