

人巨细胞病毒 AD169 株 UL115 基因编码 gL 蛋白 序列分析及 B 细胞表位预测

魏中南¹ 胡洪波¹ 彭巧英²

[摘要] 目的:分析人巨细胞病毒(HCMV)AD169 株 UL115 基因序列,预测 UL115 基因编码的 gL 蛋白 B 细胞优势表位。方法:基于 UL115 基因编码的 gL 蛋白的氨基酸序列,结合亲水性参数、可及性参数、抗原性参数、柔韧性参数及二级结构方案对 HCMV gL 蛋白的 B 细胞表位进行预测,参照已建立的预测方法综合评价 B 细胞优势表位。结果:①UL115 核酸变异集中在序列的 N 端,大部分是同义突变,氨基酸序列高度保守;②HCMV gL 蛋白 B 细胞表位可能位于编码蛋白 N 段 197~205、253~261 位。结论:多参数预测 gL 蛋白的 B 细胞优势表位,为进一步研究蛋白特征、制备单克隆抗体及表位疫苗提供依据。

[关键词] 人巨细胞病毒;AD169 株;UL115 基因;gL 蛋白;B 细胞表位

[中图分类号] R552 **[文献标志码]** A **[文章编号]** 1004-2806(2012)08-0505-03

Study on B-cell epitopes of glycoprotein L expressed by human cytomegalovirus UL115 gene of AD169 strain

WEI Zhongnan¹ HU Hongbo¹ PENG Qiaoying²

(¹Department of Clinical Laboratory, Hubei Maternal and Child Health Hospital, Wuhan, 430070, China; ²Department of Neonatology, Hubei Maternal and Child Health Hospital)

Corresponding author: HU Hongbo, E-mail: dingguang666@hotmail.com

Abstract Objective: To analyze the sequence of genome of cytomegalovirus UL115 and predict the B cell epitopes of the gL protein expressed by cytomegalovirus UL115. **Method:** The hydrophilicity, accessibility, antigenicity and flexibility index were used to predict the potential B cell epitopes of gL protein based on gL genome sequence. **Result:** ① Most amino acid sequence of gL was highly conserved although several strains had variation. Those mutations focused on the N end of UL115, but most of them were sense mutation. ② The B cell epitopes of gL protein generated by combined application were predicted at gL protein N-terminal 197—205, 253—261. **Conclusion:** The B-epitopes of gL protein was predicted successfully, which established the basis of the characterization of the protein, development of epitopes based vaccine, and preparation of monoclonal antibody against fusion protein.

Key words human cytomegalovirus; AD169 strain; UL115 gene; envelope glycoprotein L; B cell epitopes

HCMV 属于疱疹病毒科 β 亚科,直径大约 200 nm,呈球形,其主要结构蛋白分为衣壳蛋白、被膜蛋白和包膜糖蛋白。包膜糖蛋白不仅能够诱导宿主的免疫应答,而且还帮助病毒识别宿主细胞

膜受体,黏附、进入细胞,导致病毒最终致病。HCMV 膜融合机制至少涉及 4 个包膜糖蛋白 gB、gH、gL、gO。gL 蛋白位于 HCMV 氨基酸序列的 G 区,由 HCMV UL115 编码,与识别宿主细胞膜受体、吸附和穿入细胞、病毒的致病作用等密切相关。本文就 HCMV AD169 株 gL 蛋白基因序列加以分析,并应用生物学软件对临床分离株的 B 细胞优势

¹ 湖北省妇幼保健院检验科(武汉,430070)

² 湖北省妇幼保健院新生儿科

通信作者:胡洪波, E-mail: dingguang666@hotmail.com

分析[J]. 中国实用医学, 2008, 3(4): 54—55.

[4] 钟娜, 郑文爱, 朱海花, 等. 6 235 例孕妇梅毒血清学检测结果分析[J]. 中国艾滋病性病, 2007, 13(1): 52—53.

[5] 刘丽君, 魏米. 丙型肝炎病毒的流行病学[J]. 传染病信息, 2007, 20(5): 1007—1013.

[6] 叶庆华. 乙型肝炎病毒宫内感染的研究[J]. 国外医学妇产科学分册, 2002, 29: 18—20.

[7] 乐杰. 妇产科学[M]. 北京: 人民卫生出版社, 2001: 171—171.

[8] 龚向东, 叶顺章, 张君炎, 等. 1991~2001 年我国性病

流行病学分析[J]. 中华皮肤科杂志, 2002, 35(3): 178—182.

[9] 毛小恩, 周惠耕. 13 272 例婚前医学检查人群感染梅毒情况分析[J]. 中国初级卫生保健, 2003, 17(2): 78—78.

[10] 曹韵珍, 李关汉, 王世一, 等. 中国人类免疫缺陷病毒 I 母婴传播的现状、危机及对策[J]. 中华传染病杂志, 2002, 20(3): 185—187.

[11] 李甲芬. HBV 标志物阳性者产乳乳汁 HBV-DNA 测定[J]. 临床检验杂志, 2005, 23(2): 11—11.

(收稿日期: 2012-02-22)

表位进行预测,为在分子水平上进一步研究蛋白特征、制备单克隆抗体及表位疫苗提供依据。

1 材料与方法

1.1 材料

HCMV gL 蛋白核苷酸及氨基酸序列分析 AD169 (序列号: AAM96170.1)、Merlin (AAR31659.1)、Toledo (ADD39150.1) 和 Towne (ACM48085.1) 株 gL 蛋白核苷酸及氨基酸序列均检索自 Gene Bank, 共 837 个核苷酸, 编码 278 个氨基酸残基。

1.2 方法

AD169 株 gL 蛋白 B 细胞表位分析 gL 蛋白的 B 细胞表位根据以下 6 项参数预测: ① Garnier-Robson 方法预测 β 转角; ② Garnier-Robson 方法预测无规卷曲; ③ Karplus-Schulz 方法预测蛋白柔韧性; ④ Kyte-Doolittle 方法分析蛋白亲水性; ⑤ Emini 方法预测蛋白可及性; ⑥ Jameson-Wolf 方法预测蛋白抗原指数。根据分子生物学特性和相关分析结果, 筛选程序如下。

1.2.1 跨膜区预测 Goldman-EngLeman-Steitz 方法预测跨膜区, 以非跨膜区域作为抗原表位的待选区。

1.2.2 入选参数选择 入选参数包括 3 项: ① Kyte-Doolittle 方法分析氨基酸亲水性; ② Jameson-Wolf 方法预测蛋白抗原指数; ③ Emini 方法预测蛋白可及性, 以 4~6 个以上残基区域作为抗原性待选区域。此外应满足 β 转角、无规卷曲、柔韧性区域 3 项参数中有 2 项参数满足要求。

1.2.3 B 细胞表位选择 将待研究肽段中的氨基酸按文献报道已确立的统计分值计算各类氨基酸的抗原性数值总和, 最后求待选各肽段的加权抗原性分值, 通过比较评估, 取分值大的肽段, 作为该蛋白的 B 细胞表位区域。

2 结果

HCMV AD169 株 gL 蛋白核苷酸及氨基酸序列分析 gL 蛋白核苷酸变异比较普遍, 主要集中在序列的 N 端, 大部分是同义突变, 氨基酸序列高度保守。AD169 与参比株 gL 蛋白氨基酸序列同源性比较参见表 1。

表 1 AD169 与参比株 gL 蛋白氨基酸序列同源性比较

序列	AD169	Merlin	Toledo	Towne
AD169	—	5	2	4
Merlin	0.982	—	5	6
Toledo	0.992	0.982	—	6
Towne	0.985	0.978	0.978	—

注: 上方三角数据显示氨基酸变异数; 下方三角数据显示氨基酸序列同源性。

HCMV AD169 株 gL 蛋白跨膜区预测显示 gL 蛋白编码蛋白跨膜区在编码蛋白 N 段 7~29 位, 胞质区在编码蛋白 N 段 1~6 位, 因此, 编码蛋白 N 段 30~278 位的肽段为 B 细胞表位预测的候选区, 结果见图 1。

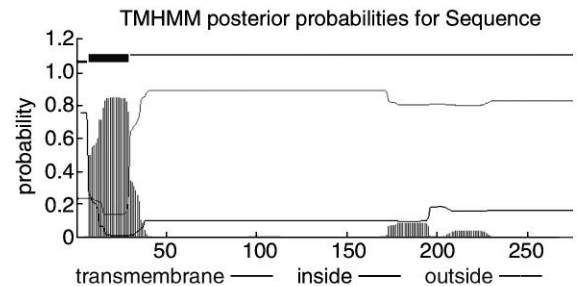


图 1 HCMV AD169 株 gL 蛋白跨膜区预测

HCMV AD169 株 gL 蛋白亲水性、可及性、抗原性、柔韧性及二级结构参数 (β 转角、无规卷曲) 预测多种预测方法均显示 gL 蛋白 N 段 62~68, 197~205 和 255~261 具有较好的亲水性、可及性和较高的抗原性指数, 并且在二级结构上含有易形成抗原表位的转角和不规则卷曲结构 (表 2)。

HCMV AD169 株 gL 蛋白抗原性指数预测 gL 蛋白 N 段 197~205、253~261 位具有较高的抗原指数 (表 3)。

表 2 HCMV AD169 株 gL 蛋白亲水性、可及性、抗原性、柔韧性及二级结构参数

Predicted Method	Predicted amino acid sequence position
Hydrophilicity	41~54 60~68 75~95 115~120 130~149 160~175 197~206 235~278
Accessibility	62~68 114~117 197~205 253~261
Antigenicity	41~54 60~68 75~82 91~97 127~133 141~149 156~164 167~176 195~207 237~251 253~261 268~278
β sheet	30~36 56~59 68~77 82~93 118~127 134~140 149~161 163~169 172~177 190~194 203~214 219~224 248~253 260~266 273~278
Coil	79~81 114~116 131~132 225~226 238~240 253~257 272~273
Flexibility	37~43 49~52 62~67 75~83 90~94 114~118 126~132 141~150 198~204 254~261 269~273

表 3 临床分离株 HCMV gL 蛋白的 B 细胞
表位平均抗原性指数

B cell epitopes	Amino acid sequence	Antigenic index, AI
62~68	GDKYESW	-0.026
197~205	NEATRTNRA	0.047
253~261	LPPELKQTR	0.056

3 讨论

gL 蛋白的功能与 gH 蛋白相关, gL、gH、和 gO 以糖蛋白复合物的形式存在, gL 能够主动结合 gH, 将其运输到细胞表面。当细胞同时表达 gH/gL 时, 就形成了明显的合胞体结构, 在允许细胞表面结合细胞受体, 导致蛋白质磷酸化和钙离子流动, 引发细胞融合^[1-2]。尽管 gL 蛋白核苷酸变异比较普遍, 但大部分是同义突变, 氨基酸序列高度保守, 不同细胞株 gL 蛋白氨基酸序列比对显示, 序列同源性高达 97.8%~99.2%, 说明该结构区对于 HCMV 生物学功能有着重要意义, 选择压力保留了这段基因的稳定性。

本研究按 Kyte-Doolittle 亲水性方案、Emini 表面可及性方案和 Jameson-Wolf 抗原性指数方案分别预测 UL115 编码蛋白可能的 B 细胞抗原表位。从预测结果发现, 不同的预测方法预测的抗原表位个数和抗原表位可能出现的肽段有所差异, 故筛选采用了多项参数一致的原则。由于 β 转角和无规卷曲多显于抗原表面, 利于与相应抗体嵌合, 所以将蛋白分子柔韧性, 具备 β 转角、无规卷曲等二级结构也纳入 B 细胞表位的筛选条件^[3-5]。

吴玉章等^[6]综合考虑蛋白质的许多性质如片段的活性、结构、构象和氨基酸侧链的排列等, 建立了抗原性指数预测方法, 应用该方法计算可能是 B 细胞表位的平均抗原指数。结果提示 gL 蛋白 N 段 197~205 和 255~261 平均抗原指数较高, 可能是 B 细胞表位的优势区段。保守性结构分析显示, 两细胞表位均位于 gL 蛋白保守区段, 发生变异的可能性

相对较小, 推测如果用该段表位诱生的单克隆抗体, 可以对在病毒发生变异的情况下对 HCMV 的诊断。

根据氨基酸序列及其二级结构来进行表位预测, 其结果多为线性表位, 但绝大多数的 B 细胞表位为构象表位, 通过三维结构表现其抗原性, 因此预测结果只能作为 B 细胞潜在表位的参考。本研究通过对 UL115 基因的序列分析, 并结合 UL115 编码 gL 蛋白的二级结构、亲水性、表面可及性与抗原性指数等多种参数, 预测了 gL 蛋白的 B 细胞抗原表位, 为研究蛋白特征、制备单克隆抗体及表位疫苗提供依据。

参考文献

- [1] VANARSDALL A L, CHASE M C, JOHNSON D C. Human cytomegalovirus glycoprotein go complexes with gH/gL, promoting interference with viral entry into human fibroblasts but not entry into epithelial cells [J]. J Virol, 2011, 85: 11638-11645.
- [2] BOWMAN J J, LACAYO J C, BURBELO P, et al. Rhesus and human cytomegalovirus glycoprotein L are required for infection and cell-to-cell spread of virus but cannot complement each other [J]. J Virol, 2011, 85: 2089-2099.
- [3] KOLASKAR A S, TONGAONKAR P C. A semi-empirical method for prediction of antigenic determinants on protein antigens [J]. FEBS Lett, 1990, 276: 172-174.
- [4] KARPLUS P A, SCHULZ G E. Prediction of chain flexibility in proteins [J]. Naturwissenschaften, 1985, 72: 212-213.
- [5] EMINI E A, HUGHES J V, PERLOW D S, et al. Induction of hepatitis a virus-neutralizing antibody by a virus-specific synthetic peptide [J]. J Virol, 1985, 55: 836-839.
- [6] 吴玉章, 朱锡华. 一种病毒蛋白 B 细胞表位预测方法的建立 [J]. 科学通报, 1994, 39(24): 2275-2279.

(收稿日期: 2012-01-19)

本刊对“关键词”书写的要求

关键词是科技论文的文献检索标志, 是表达文献主题概念的词或词组。每篇论文应选取 1~5 个关键词。关键词应尽量从美国国立医学图书馆编印的“Medical Subject Headings (MeSH)”中选取, 其中文译名可参照中国医学科学院信息研究所编译《医学主题词注释字顺表》。未被词表收录的词(自由词)必要时也可以作为关键词使用。凡有英文摘要的文章, 应标注与中文对应的英文关键词。中、英文关键词分别用通栏排在中、英文摘要下方。无摘要文章的关键词排在正文前。