

SARS 冠状病毒辅助蛋白的研究进展

黄灵芝

(军事医学科学院野战输血研究所, 北京 100850)

摘要: 严重急性呼吸道综合征冠状病毒(SARS-CoV)2002~2003年在全球范围内引起了 SARS 的大规模暴发。其基因组约 30kb, 含 14 个潜在的开放阅读框(ORF), 其中 6 个 ORF 编码病毒复制的必需蛋白如复制酶(ORF 1a 和 1b)和 4 个结构蛋白:核衣壳、刺突、膜和包膜。其余 8 个 ORF 编码 SARS-CoV 特有的辅助蛋白。本文主要在 3 个方面概述了 SARS-CoV 辅助蛋白的研究进展:检测感染病人辅助蛋白的抗体;辅助蛋白的表达、加工和定位;辅助蛋白对细胞功能的影响,从而揭示 SARS-CoV 辅助蛋白的分子和生化特征,为研究 SARS-CoV 为何有如此强大的致病力提供线索。

关键词: 严重急性呼吸综合征; 冠状病毒; 病毒辅助蛋白

中图分类号: R373; R978.7 **文献标识码:** A **文章编号:** 1001-0971(2007)03-0208-03

1 引言

2003 年严重急性呼吸道综合征(SARS)在世界范围内暴发,全球共报道了 8 000 多例死亡病例,促使全世界的科学家快速分离了这种冠状病毒,命名为 SARS-CoV,属于包膜单正链 RNA 病毒,基因组达 30 kb,可以感染多种物种包括人类。从人体分离的 SARS-CoV 具有高度同源性。流行病学调查显示 SARS-CoV 很可能由动物传播到人体,而且可能再次出现,因此必须深入研究。

本文概述了目前对 SARS-CoV 的研究进展。SARS-CoV 属于典型的冠状病毒科,2/3 的 SARS-CoV 基因编码复制酶,可以翻译成两个大的蛋白:pp1a(486 ku)和 pp1ab(790 ku)。这些蛋白的水解由半胱氨酸蛋白酶介导,产生非结构蛋白,其中一些与病毒复制相关,或产生 mRNA 表达其他开放阅读框(ORF),进而翻译成结构蛋白如刺突(S)、包膜(E)、膜(M)和核衣壳(N)。每种冠状病毒编码不同的辅助蛋白,SARS-CoV 基因组编码 8 个辅助蛋白,长度从 39~274 个氨基酸不等,SARS-CoV 复制酶和结构蛋白基因具有一定的序列同源性,但其辅助蛋白与其他已知冠状病毒蛋白在氨基酸序列上同源性较低。然而研究表明,SARS-CoV 的辅助蛋白如 3a,3b,9b 具有其他冠状病毒蛋白的部分结构特点和生物学特征。

2 病人血清中抗体的检测

在 SARS 病人血清中检测到 SARS-CoV 辅助蛋白的抗体,提示这些蛋白在病毒感染过程中表达,但抗体只在少量病人中发现,提示其表达水平很低或无高度免疫原性。研究表明,3a 抗体在多数 SARS 病人中发现,提示 3a 可能是丰富的具有免疫原性的辅助蛋白,表达于感染细胞表面,其拓扑学结构证实,3a 第一个跨膜区的 34 个氨基酸面对细胞外基质,第三个跨膜区后的 C 端面对胞浆,其 N 端外功能区伸出病毒体,而且在 2 例病人中发现 B 细胞识别 3a 的 N 端。研究表明 48.8% 的 SARS 康复病人具有 3a 的 N 端抗体,死亡病例中仅 7.4%。但 3a 抗体的存在目前并未证明能够抑制病毒感染。然而,Zhong 等发现 3a 的 N 端外功能区可以诱发强烈的抗原抗体反应,裂解感染细胞,提示 3a 可能激发了病人的体液免疫保护机制。

3 表达、包装和细胞定位

SARS-CoV 通过转录减弱机制合成负链 RNA,作为模板合成长基因组 mRNA 和亚基因组 mRNA (sgRNA),基因组 mRNA 被翻译成第一个 ORF,其余 ORF 由 8 个 sgRNA 翻译而来。9 个不同长度的 mRNA 均可从被 SARS-CoV 感染的细胞中检测到,SARS-CoV 辅助蛋白由 sgRNA 3,6,7,8,9 编码.sgRNA6 为单顺反子,sgRNA 3,7,8,9 为双顺反子。

3.1 ORFs 3a 和 3b

SARS-CoV 的 sgRNA 3 含 2 个 ORF:3a 和 3b。体内外实验表明,3a 蛋白表达于感染细胞,定位于

收稿日期:2006-11-20

作者简介:黄灵芝,女,在读硕士研究生,研究方向:免疫血液学,
Tel:010-66931904, E-mail:huanglz1982@163.com

高尔基体,可以转运到细胞表面,其胞浆区由2个并列的分选基序 $Yxx\Phi$ 和 ExD 组成。3a 胞浆区缺失后,3a 被滞留在细胞内。3a 属于冠状病毒结构蛋白,被包装进入病毒样颗粒,在杆状病毒中可以与 M, E 蛋白共表达。经转录后修饰存在两种形式的 3a 蛋白,分子量分别为 37 ku (3a-1) 和 31 ku (3a-2),均分泌到细胞外。在感染细胞中,3a-1 的密度接近于 SARS 病毒颗粒,提示其很可能被并入病毒体中。而 3a-2 密度较低,可能作为细胞外膜结构被释放。3a 胞浆分选基序的缺失使其不能到达细胞表面,进而不能分泌出胞。SARS-CoV 的 N 端只在哺乳动物细胞中表达,可以出胞。N 端蛋白可在病人血清中检测到,证明 N 蛋白可以分泌到细胞外。

未糖基化的 3a 蛋白分子量为 31 ku,与 3a-2 蛋白相似。31 ku 的蛋白被 O 连接糖基化后分子量增加。3a 与冠状病毒 M 蛋白极为相似,均为 3 次跨膜糖蛋白,具有短 N 端外功能区和长 C 端内功能区,主要定位在高尔基体,尽管序列同源性较低,但 3a 和 M 蛋白的疏水区高度相似,提示两者可能具有相似的结构和生物学特性。

3a 蛋白可与 S 蛋白相互作用,在 SARS-CoV 感染的细胞外基质中可以检测到,提示 3a 和 S 可能通过病毒颗粒的形成一起分泌到细胞外。通过比较从病人和狸猫中分离的 SARS-CoV 序列发现,S 和 3a 基因非同义变化速率比同义变化速率要高,提示在病毒进化过程中 S 和 3a 蛋白有阳性选择性,使得病毒改善在宿主中的适应性,因此这些蛋白在病毒生活周期和疾病的发展中发挥重要作用。Schwegmann-Wessel 等报道了细胞外定位的新的分选信号在多数冠状病毒的 S 蛋白如 TGEV 表达,但对 SARS-CoV S 缺失、定点突变研究证实 $Yxx\Phi$ 基序单独表达时使 TGEV S 蛋白滞留于细胞内,而 SARS-CoV S 蛋白可以有效地转运到细胞表面。如果 3a 蛋白可以经过细胞内摄作用表达在细胞表面,推测 3a 可能通过与 S 蛋白相互作用调节细胞表面 S 蛋白的表达,但还需进一步实验证实。研究证明 3a 蛋白可以与其他结构蛋白如 M, E 相互作用,但是否与 3a 的功能或与 3a 并入病毒体有关仍需继续研究。

在感染病毒的 Vero E6 细胞中发现了 3a 基因的突变,而且在临床病人样本中也发现有相似的突变,提示 3a 的突变可能并不是由于对体外培养环境的适应。最新的研究表明,如果 3a 基因包含移码突变,全长的 3a 蛋白仍可以表达,但表达水平很低,只

有部分临床样本包含 3a 基因移码突变的病毒,暗示很少或没有全长 3a 蛋白表达,从而也可以解释为什么一些病人无 3a 抗体。

3b 蛋白也可以在 SARS-CoV 感染的细胞中检测到,而且在 SARS 病人血清中也能检测到 3b 抗体,3b 蛋白包含 2 个可预测的核定位信号,使其定位在核中。其他冠状病毒感染型支气管炎病毒 (IBV) 的 3b 蛋白也定位在核中,然而 SARS-CoV 的 3a 基因与 IBV 氨基酸序列相似性较低,而且不知这两种蛋白是否通过相同的机制转运到核中。也有报道表明 SARS-CoV 3b 对细胞周期和对细胞凋亡的调节具有一定作用。

3.2 ORF6

ORF6 存在于胞浆,有时存在于线粒体和高尔基体,当将其构建到载体上并通过脂质体转染时,其蛋白稳定性较低,但通过重组病毒的方法表达量高,提示在病毒感染时防止 ORF 细胞内降解的细胞因子上调,但在病毒复制和病理过程中的作用需要深入研究。

3.3 ORF 7a 和 7b

SARS-CoV 7a 编码 sgRNA 7 的第一个 ORF,在感染 SARS-CoV 的细胞表达,定位于线粒体和高尔基体的中间小室,同时也存在于 SARS 病人的肺组织中,特别是支气管上皮细胞、外周血红细胞和白细胞。使用晶体衍射和 NMR 对 7a 的结构分析发现,7a 的 16~85 位氨基酸形成了一个压缩结构,存在 Ig 样折叠,可将其定义为 Ig 亚家族。Hanel 等发现 7a 的结构与表达于细胞表面的细胞间粘附分子 ICAM-1 或 ICAM-2 的 D1 区相似,85~97 位氨基酸在晶体状态和溶液状态均具有伸缩性,可能包含了与 7a 细胞定位相关的信号。7a 蛋白是迄今为止唯一一个已经解析其结构的 SARS-CoV 辅助蛋白,由于 7a 与 ICAM-1 具有结构相似性,从而预测 7a 可以与 LFA-1 的整合素 I 区亚基 I 区结合,但还未经实验证明。

尽管在 SARS 病人中检测到 7b 抗体,但 7b 在感染细胞中是否表达还未经证明。有研究报道,Frankfurt 分离株在体外培养 3 代后 7b ORF 的 45 个氨基酸缺失,提示 7b 可能在病毒复制中不是十分重要。而且 Yount 等也证明 7a, 7b 的缺失对病毒的复制无显著影响,但仍需研究 7a, 7b 在 SARS-CoV 宿主中的作用。

3.4 ORFs 8a 和 8b

8a 和 8b 在感染 SARS-CoV 的 Vero E6 细胞表达,分别编码 39 和 84 个氨基酸,但从动物体内分离的 SARS-CoV 多 29 个氨基酸,使得 8a,8b 融合成一个蛋白 8ab,共 122 个氨基酸,其 N 端与 8a 相同,C 端与 8b 相同。对 63 例 SARS 病人的病毒分离株研究发现病毒基因组的这个区域都存在缺失,但由于病情程度不同,缺失的类型也不同。尽管 ORF8 的缺失对病毒的生存无显著影响,但 8a,8b,8ab 具有不同的稳定性和功能,在病毒复制和其病理过程中发挥不同作用。由于 SARS-CoV 可以从动物传播到人,提示这些辅助蛋白可能使得病毒更适应环境。从 Vero E6 和 FRhk4 细胞分离的病毒 8a 基因存在移码突变,但未在临床病人样本中发现,提示这种突变可能是在体外培养过程中产生的。

3.5 ORF 9b

9b 在感染的细胞中表达,而且在 SARS 病人血清中也发现 9b 抗体,由含 N 基因的内部 ORF 编码,与小鼠冠状病毒和牛冠状病毒类似。

4 对细胞功能的影响

研究发现,辅助蛋白不是病毒复制所必需,但对于病毒-宿主相互作用,以及病毒的稳定性和病理过程非常重要。如猫冠状病毒的 7b 基因在体外培养中缺失,其致病力降低。对辅助蛋白在病毒复制中作用的研究表明,5 个辅助蛋白 3a,3b,6,7a 和 7b 对病毒复制不重要,当 ORF 缺失后,病毒仍可复制。

4.1 ORF 3a 和 3b

A549 肺上皮细胞稳定表达 3a,而且人纤维蛋白原 3 个亚基 $\text{A}\alpha$, $\text{B}\beta$ 和 γ 的 mRNA 表达均上调,细胞内人纤维蛋白原的分泌也增加。在感染 Vero E6 细胞中人纤维蛋白原水平升高,与人纤维蛋白原 mRNA 在 SARS-CoV 感染的人外周血单核细胞中表达升高一致,提示人纤维蛋白原的过量产生对 SARS-CoV 感染的病理进程非常重要。

在感染急性期,损伤或肿瘤形成时,肝纤维蛋白原增加恢复内环境稳定性,但纤维蛋白原的过量产生和纤维蛋白的形成也会加强细胞因子的生成,导致促凝血和纤维蛋白溶解的不平衡。对 SARS 病人尸检发现存在严重肺损伤、血液学指标异常且活化部分凝血酶时间延长,提示在 SARS-CoV 感染中凝血功能和纤维蛋白聚合紊乱。3a 的单独表达可以上调肺细胞中纤维蛋白原的表达,提示 3a 对于

SARS 病理进程也有作用。在 Vero E6 细胞中 3a 的过表达可以激活半胱天冬酶(caspase)-8 依赖的凋亡信号通路。3b 是 sgRNA 的第二个 ORF,对其研究较少,但 3b 的过表达可使细胞周期阻滞在 G0/G1 期并诱导凋亡,但是否依赖 caspase 通路还不清楚。3b 蛋白很可能通过内部核糖体进入机制被表达,其起始密码子不是 AUG。Hussain 等报道 3b 蛋白可以从独立的 sgRNA 中表达,但很少发生。病毒复制周期中是否存在 3b 的短暂表达尚待研究。

4.2 ORF6

SARS-CoV ORF6 蛋白可以增强小鼠冠状病毒(MHV)致病性。研究表明将 ORF6 插入 MHV 基因组中可促进病毒繁殖,并增加对小鼠的致死性。尽管 ORF6 在 SARS-CoV 感染中的重要性仍需探讨,但研究表明 ORF6 对于加强病毒复制和组装有直接作用,而且 ORF6 的过表达能加强细胞内 DNA 合成。

4.3 ORF7

7a 的过表达能够通过 caspase 依赖的通路诱导凋亡。但 7a 可能不是感染 SARS-CoV Vero E6 细胞中唯一诱导凋亡的因子。因此应继续研究不同的病毒蛋白在 SARS-CoV 的感染中对细胞的作用。

7a 可以通过细胞周期蛋白(cyclin) D3 的表达和视网膜母细胞瘤蛋白的磷酸化将细胞周期阻滞在 G0/G1 期。过表达 7a 可以抑制细胞蛋白合成,使得 p38 MAPK 磷酸化并激活。然而,对 p38 磷酸化的抑制并不能阻止对翻译的抑制或 7a 诱导的凋亡。除此之外,还需研究 7a 与 3a 的相互作用,而 3a 又可以与 M, E 相互作用,所以提示这些病毒蛋白可能会在 SARS-CoV 感染的细胞中形成复合体。

最新研究表明 7a 可与含 TPR 基序的细胞蛋白 SGT 结合。而含 TPR 基序的蛋白与细胞周期的控制、转录和剪接方式、蛋白转运和蛋白折叠等相关,因此 7a 和 SGT 的生物学特性仍需继续研究。

5 展望

研究表明,辅助蛋白 3a,3b,6,7a 和 7b 对于 SARS-CoV 的复制不是必需的,但在病毒稳定性和宿主病理进程中发挥作用,而且可以精细地调节病毒复制。以后还需研究辅助蛋白和其他病毒蛋白或细胞因子的关系,从而更好研究 SARS-CoV 和人体宿主之间的相互作用,为抗病毒治疗和疾病控制提供有效的方案。