



(12)发明专利申请

(10)申请公布号 CN 109715188 A

(43)申请公布日 2019.05.03

(21)申请号 201780045021.7

(22)申请日 2017.07.20

(30)优先权数据

16180430.7 2016.07.20 EP

(85)PCT国际申请进入国家阶段日

2019.01.18

(86)PCT国际申请的申请数据

PCT/EP2017/068297 2017.07.20

(87)PCT国际申请的公布数据

WO2018/015468 EN 2018.01.25

(71)申请人 慕尼黑工业大学

地址 德国慕尼黑

申请人 马普协会

(72)发明人 阿娜希塔·贾瓦希里

托比亚斯·克鲁泽

马库斯·格哈德

伯恩哈德·B·辛格尔

丹尼尔·霍恩伯格 汗·雷莫特

马蒂亚斯·曼 费利克斯·迈斯纳

斯特芬·巴克特

(74)专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 郑斌 刘振佳

(51)Int.Cl.

A61K 38/02(2006.01)

A61K 48/00(2006.01)

A61K 31/00(2006.01)

C07K 16/12(2006.01)

C07K 16/28(2006.01)

G01N 33/68(2006.01)

A61P 31/12(2006.01)

A61P 35/00(2006.01)

权利要求书3页 说明书44页

序列表25页 附图27页

(54)发明名称

用于预防或治疗幽门螺杆菌感染的药剂和方法

(57)摘要

本发明涉及幽门螺杆菌(H.pylori)HopQ与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂,并且涉及基于幽门螺杆菌HopQ的免疫原性组合物。本发明还涉及所述抑制剂和免疫原性组合物用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的用途。

1. 幽门螺杆菌 (*Helicobacter pylori*) HopQ 与癌胚抗原相关细胞黏附分子 (CEACAM) 家族成员之间相互作用的抑制剂, 其用于预防或治疗由幽门螺杆菌 (*H. pylori*) 引起或与其相关的疾病或病症的方法。

2. 根据权利要求1所述应用的抑制剂, 其中所述抑制剂抑制幽门螺杆菌 HopQ 与所述 CEACAM 家族成员的结合和/或 HopQ-CEACAM 介导的信号传导, 其中优选地所述抑制剂抑制幽门螺杆菌 HopQ 与所述 CEACAM 家族成员的胞外结构域的结合, 优选与所述 CEACAM 家族成员的 N-结构域的结合。

3. 根据权利要求1或2所述应用的抑制剂, 其中所述 CEACAM 家族成员选自人 CEACAM 家族成员、非人灵长类 CEACAM 家族成员和大鼠 CEACAM 家族成员, 其中优选地所述 CEACAM 家族成员选自 CEACAM1、CEACAM3、CEACAM5 和 CEACAM6。

4. 根据权利要求1至3中任一项所述应用的抑制剂, 其中所述抑制剂选自:

(a) 与幽门螺杆菌 HopQ 结合、优选与幽门螺杆菌 HopQ 的胞外结构域结合的(多)肽配体或肽模拟物配体;

(b) 与所述 CEACAM 家族成员结合、优选与所述 CEACAM 家族成员的胞外结构域结合、更优选与所述 CEACAM 家族成员的 N-结构域结合的(多)肽配体或肽模拟物配体;

(c) 编码(a)和(b)的所述(多)肽配体的核酸分子;

(d) 与幽门螺杆菌 HopQ 结合、优选与幽门螺杆菌 HopQ 的胞外结构域结合的核酸配体;

(e) 与所述 CEACAM 家族成员结合、优选与所述 CEACAM 家族成员的胞外结构域结合、更优选与所述 CEACAM 家族成员的 N-结构域结合的核酸配体;

(f) 抑制所述 CEACAM 家族成员或幽门螺杆菌 HopQ 的表达的抑制性核酸分子;

(g) 与幽门螺杆菌 HopQ 结合、优选与幽门螺杆菌 HopQ 的胞外结构域结合的小分子; 以及

(h) 与所述 CEACAM 家族成员结合、优选与所述 CEACAM 家族成员的胞外结构域结合、更优选与所述 CEACAM 家族成员的 N-结构域结合的小分子。

5. 根据权利要求4所述应用的抑制剂,

其中所述(多)肽配体选自所述 CEACAM 家族成员或幽门螺杆菌 HopQ 的抗体、抗体衍生物、抗体模拟物、肽适配体和可溶性片段;

其中所述肽模拟物配体选自类肽、 β -肽和 D-肽;

其中所述核酸配体选自 DNA 适配体、RNA 适配体和 XNA 适配体; 和/或

其中所述抑制性核酸分子选自 siRNA、shRNA、miRNA 和反义 DNA 或 RNA 分子。

6. 根据权利要求4或5所述应用的抑制剂, 其中幽门螺杆菌 HopQ 的所述胞外结构域是幽门螺杆菌 HopQ 的插入结构域、环 A、环 B、环 C 或环 D, 其中:

环 A 位于幽门螺杆菌 HopQ 的螺旋 H3 与链 S1 之间;

环 B 位于幽门螺杆菌 HopQ 的链 S2 与螺旋 H4 之间;

环 C 位于幽门螺杆菌 HopQ 的螺旋 H5 与螺旋 H6 之间; 并且

环 D 位于幽门螺杆菌 HopQ 的螺旋 H7 与螺旋 H8 之间。

7. 根据权利要求1至6中任一项所述应用的抑制剂, 其中所述由幽门螺杆菌引起或与其相关的疾病或病症选自幽门螺杆菌感染和由幽门螺杆菌引起的胃十二指肠病症, 其中优选地所述由幽门螺杆菌引起的胃十二指肠病症选自胃炎、慢性胃炎、胃萎缩、胃或十二指肠溃疡、胃癌和 MALT 淋巴瘤。

8. 用于鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的体外方法,所述方法包括:

(a) 使 (i) CEACAM蛋白或其功能性片段与 (ii) 幽门螺杆菌HopQ蛋白或其功能性片段和 (iii) 受试化合物接触,以及

(b) 确定所述受试化合物是否抑制所述CEACAM蛋白或其功能性片段与所述幽门螺杆菌HopQ蛋白或其功能性片段之间的相互作用,

其中抑制所述CEACAM蛋白或其功能性片段与所述幽门螺杆菌HopQ蛋白或其功能性片段之间相互作用的受试化合物被鉴定为用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物,

其中优选地步骤 (b) 包括确定所述受试化合物是否抑制所述幽门螺杆菌HopQ蛋白或其功能性片段与所述CEACAM蛋白或其功能性片段的结合,其中优选地所述幽门螺杆菌HopQ蛋白的功能性片段包含胞外结构域或其片段,和/或所述CEACAM蛋白的功能性片段包含胞外结构域或其片段,优选N-结构域;和/或确定所述受试化合物是否抑制HopQ-CEACAM介导的信号传导。

9. 幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子 (CEACAM) 家族成员之间相互作用的抑制剂,其中所述抑制剂选自:

- (a) 与幽门螺杆菌HopQ的胞外结构域结合的(多)肽配体或肽模拟物配体;
- (b) 与所述CEACAM家族成员的N-结构域结合的(多)肽配体或肽模拟物配体;
- (c) 编码 (a) 和 (b) 的所述(多)肽配体的核酸分子;
- (d) 与幽门螺杆菌HopQ的胞外结构域结合的核酸配体;
- (e) 与所述CEACAM家族成员的N-结构域结合的核酸配体;
- (f) 抑制所述CEACAM家族成员或幽门螺杆菌HopQ的表达的抑制性核酸分子;
- (g) 与幽门螺杆菌HopQ的胞外结构域结合的小分子;以及
- (h) 与所述CEACAM家族成员的N-结构域结合的小分子。

10. 根据权利要求9所述的抑制剂,其中幽门螺杆菌HopQ的所述胞外结构域是幽门螺杆菌HopQ的插入结构域、环A、环B、环C或环D,其中:

- 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间;
- 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间;
- 环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间;并且
- 环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。

11. 免疫原性组合物,其包含:

(a) 至少一种分离的(多)肽,其包含 (i) 幽门螺杆菌HopQ的氨基酸序列;或 (ii) 其免疫原性变体;或 (iii) (i) 或 (ii) 的免疫原性片段;或者

(b) 编码根据项目 (a) 的分离的(多)肽的至少一种核酸分子。

12. 根据权利要求11所述的免疫原性组合物,其中所述免疫原性片段包含幽门螺杆菌HopQ的胞外结构域,其中优选地所述胞外结构域是幽门螺杆菌HopQ的插入结构域、环A、环B、环C或环D,或者前述任一种的功能性片段,其中:

- 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间;
- 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间;

环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间;并且
环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。

13. 根据权利要求11或12所述的免疫原性组合物,其中所述免疫原性组合物引发包含抗体分泌的免疫应答,其中优选地所述抗体抑制幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间的相互作用。

14. 根据权利要求11至13中任一项所述的免疫原性组合物,其用作药物。

15. 根据权利要求11至14中任一项所述的免疫原性组合物,其用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,其中优选地所述疾病或病症选自幽门螺杆菌感染和由幽门螺杆菌引起的胃十二指肠病症,其中优选地所述胃十二指肠病症选自胃炎、慢性胃炎、胃或十二指肠溃疡、胃癌和MALT淋巴瘤。

用于预防或治疗幽门螺杆菌感染的药剂和方法

技术领域

[0001] 本发明涉及幽门螺杆菌 (*H. pylori*) HopQ 与癌胚抗原相关细胞黏附分子 (carcinoembryonic antigen-related cell adhesion molecule, CEACAM) 家族成员之间相互作用的抑制剂, 并且涉及基于幽门螺杆菌 HopQ 的免疫原性组合物。本发明还涉及所述抑制剂和所述免疫原性组合物用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的用途。

背景技术

[0002] 幽门螺杆菌 (*Helicobacter pylori*, *H. pylori*) 是微需氧革兰氏阴性细菌, 能够在人的胃中终身持续存在。幽门螺杆菌感染是人中最常见的细菌感染性疾病: 根据地区的社会经济状况, 全球约一半人口感染有幽门螺杆菌 (Perez-Perez 等, 2004)。这种感染与许多胃疾病 (例如慢性萎缩性胃炎、消化性溃疡、胃癌和黏膜相关性淋巴组织 (mucosa associated lymphoid tissue, MALT) 淋巴瘤) 相关 (Nomura 等, 1994; Forman, 1996; Parsonnet 等, 1991; Blaser 等, 1995)。幽门螺杆菌是胃癌 (第三常见的癌症类型, 2011 年全球有 983,000 例) 的主要原因 (Jemal 等, 2011)。

[0003] 胃癌伴随着相当大的社会经济成本。目前治疗单个胃癌患者的费用为约 50,000 欧元。预防胃癌包括早期治疗由幽门螺杆菌引起的感染。据估计, 患有由幽门螺杆菌引起的感染的个体中至少有三分之一需要治疗。目前, 难以预测哪些患者会发生与幽门螺杆菌感染相关的后续疾病。根据大量研究的结果, 用于预防胃癌的对幽门螺杆菌感染的一般治疗是具有成本效益的, 因为其将会预防超过 95% 的病例 (Graham & Shiotani, 2005)。患有胃溃疡、癌前或确定性胃癌的患者, 胃癌患者的亲属, 以及需要长期用非甾体抗炎药 (包括用于心血管疾病的阿司匹林) 治疗的患者被明确指出进行治疗。由于在日本胃癌的比例较高, 在那里建议所有感染幽门螺杆菌的个体都进行治疗, 但是抗生素抗性比例稳定地提高 (Shiota 等, 2010)。

[0004] 迄今为止, 对于由幽门螺杆菌引起的感染的标准治疗由与质子泵抑制剂 (例如奥美拉唑) 组合的两种抗生素组成。一周治疗的费用为每位患者约 200 欧元。这种治疗在一些患者中具有显著的副作用, 并且导致抗性病原体急剧增多。由于二线和三线治疗经常失败, 现今所有患者中约 10% 不能再接受治疗 (Gao 等, 2010), 截止 2020 年这可上升至估计 60%。

[0005] 此外, 近年来已经鉴定出越来越多定植动物和人的肠肝道 (enterohepatic tract) 的螺杆菌属 (*Helicobacter*) 物种 (幽门螺杆菌除外), 并表明其参与多种疾病 (Fox, 2002)。例如, 胆汁螺杆菌 (*H. bills*) 与例如胆囊炎、胆囊癌和胆道恶性肿瘤的疾病相关 (Fox 等, 1998; Matsukura 等, 2002; Pisani 等, 2008)。

[0006] 因此, 需要用于预防或治疗由螺杆菌属 (例如幽门螺杆菌或胆汁螺杆菌) 引起或与其相关的疾病或病症的新治疗方法。例如, 如果有可用的针对幽门螺杆菌的疫苗, 则可使数百万患者受益, 并显著降低医疗成本。疫苗在对抗流行性感染性疾病方面高度有效。事实上, 美国疾病控制中心 (U.S. Center of Disease Control) 称疫苗接种是用于预防感染性

疾病的最有效方法 (U.S.CDC, 2011)。然而, 迄今为止, 还没有可用的针对幽门螺杆菌的用于人的有效疫苗。在设计疫苗时, 靶标筛选和选择不利于成功实现泛保护 (pan protection) (Gómez-Gascón等, 2012)。用于疫苗接种的最佳抗原不仅应该是保守的, 而且应该对于定植、维持感染或致病性是必需的。因此, 使得细菌能够与其宿主直接相互作用的抗原通常可为疫苗接种和治疗提供优选的靶标。

[0007] 发明概述

[0008] 一方面, 本发明涉及幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子 (CEACAM) 家族成员之间相互作用的抑制剂, 其用于预防或治疗由或幽门螺杆菌引起或与其相关的疾病或病症的方法。

[0009] 在一个实施方案中, 抑制剂抑制幽门螺杆菌HopQ与CEACAM家族成员结合和/或HopQ-CEACAM介导的信号传导。

[0010] 在一个实施方案中, 抑制剂抑制幽门螺杆菌HopQ与CEACAM家族成员结合, 优选抑制与CEACAM家族成员的胞外结构域结合, 更优选抑制与CEACAM家族成员的N-结构域 (N-domain) 结合。

[0011] 在一个实施方案中, CEACAM家族成员在上皮细胞、内皮细胞和/或免疫细胞的表面上表达。

[0012] 在一个实施方案中, CEACAM家族成员选自人CEACAM家族成员、非人灵长类CEACAM家族成员和大鼠CEACAM家族成员。

[0013] 在一个实施方案中, CEACAM家族成员选自CEACAM1、CEACAM3、CEACAM5和CEACAM6。

[0014] 在一个实施方案中, 幽门螺杆菌HopQ是I型HopQ蛋白或II型HopQ蛋白。

[0015] 在一个实施方案中, 抑制剂选自:

[0016] (a) 与幽门螺杆菌HopQ结合、优选与幽门螺杆菌HopQ的胞外结构域结合的(多)肽配体或肽模拟物配体 (peptidomimetic ligand);

[0017] (b) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的(多)肽配体或肽模拟物配体;

[0018] (c) 编码(a)和(b)的(多)肽配体的核酸分子;

[0019] (d) 与幽门螺杆菌HopQ结合、优选与幽门螺杆菌HopQ的胞外结构域结合的核酸配体;

[0020] (e) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的核酸配体;

[0021] (f) 抑制CEACAM家族成员或幽门螺杆菌HopQ的表达的抑制性核酸分子;

[0022] (g) 与幽门螺杆菌HopQ结合、优选与幽门螺杆菌HopQ的胞外结构域结合的小分子; 以及

[0023] (h) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的小分子。

[0024] 在一个实施方案中, (多)肽配体选自CEACAM家族成员或幽门螺杆菌HopQ的抗体、抗体衍生物、抗体模拟物、肽适配体和可溶性片段。

[0025] 在一个实施方案中, 肽模拟物配体选自类肽、 β -肽和D-肽。

[0026] 在一个实施方案中, 核酸配体选自DNA适配体、RNA适配体和XNA适配体。

- [0027] 在一个实施方案中,抑制性核酸分子选自siRNA、shRNA、miRNA和反义DNA或RNA分子。
- [0028] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的插入结构域。
- [0029] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B、环C或环D,其中:
- [0030] 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间;
- [0031] 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间;
- [0032] 环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间;并且
- [0033] 环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。
- [0034] 在一个实施方案中,抑制剂包含在药物组合物中。
- [0035] 在一个实施方案中,由幽门螺杆菌引起或与其相关的疾病或病症选自幽门螺杆菌感染和幽门螺杆菌引起的胃十二指肠病症。
- [0036] 在一个实施方案中,由幽门螺杆菌引起的胃十二指肠病症选自胃炎、慢性胃炎、胃萎缩、胃或十二指肠溃疡、胃癌和MALT淋巴瘤。
- [0037] 在另一方面,本发明涉及用于鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的体外方法,该方法包括:
- [0038] (a)使(i)CEACAM蛋白或其功能性片段与(ii)幽门螺杆菌HopQ蛋白或其功能性片段和(iii)受试化合物接触,以及
- [0039] (b)确定受试化合物是否抑制CEACAM蛋白或其功能性片段与幽门螺杆菌HopQ蛋白或其功能性片段之间的相互作用,其中抑制CEACAM蛋白或其功能性片段与幽门螺杆菌HopQ蛋白或其功能性片段之间相互作用的受试化合物被鉴定为用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物。
- [0040] 在一个实施方案中,步骤(b)包括确定受试化合物是否抑制幽门螺杆菌HopQ蛋白或其功能性片段与CEACAM蛋白或其功能性片段的结合,其中优选地幽门螺杆菌HopQ蛋白的功能性片段包含胞外结构域或其片段,和/或CEACAM蛋白的功能性片段包含胞外结构域或其片段,优选N-结构域;和/或确定受试化合物是否抑制HopQ-CEACAM介导的信号传导。
- [0041] 在一个实施方案中,CEACAM蛋白选自人CEACAM蛋白、非人灵长类CEACAM蛋白和老鼠CEACAM蛋白。
- [0042] 在一个实施方案中,CEACAM蛋白选自CEACAM1、CEACAM3、CEACAM5和CEACAM6。
- [0043] 在一个实施方案中,幽门螺杆菌HopQ蛋白是I型HopQ蛋白或II型HopQ蛋白。
- [0044] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的插入结构域或其功能性片段。
- [0045] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B、环C或环D,或者前述任一种的功能性片段,其中:
- [0046] 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间;
- [0047] 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间;
- [0048] 环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间;并且
- [0049] 环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。

[0050] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B或环C,或者前述任一种的功能性片段。

[0051] 在一个实施方案中,受试化合物选自(多)肽、肽模拟物、核酸分子和小分子。

[0052] 在另一方面,本发明涉及能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段用于研究幽门螺杆菌感染或者鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0053] 在另一方面,本发明涉及异源表达能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段的细胞用于研究幽门螺杆菌感染或者鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0054] 在另一方面,本发明涉及异源表达能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段的非人转基因动物用于研究幽门螺杆菌感染或者鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0055] 在上述用途的一个实施方案中,CEACAM蛋白选自人CEACAM蛋白、非人灵长类CEACAM蛋白和大鼠CEACAM蛋白。

[0056] 在一个实施方案中,CEACAM蛋白选自CEACAM1、CEACAM3、CEACAM5和CEACAM6。

[0057] 在上述方法或用途的一个实施方案中,由幽门螺杆菌引起或与其相关的疾病或病症选自幽门螺杆菌感染和由幽门螺杆菌引起的胃十二指肠病症。

[0058] 在一个实施方案中,由幽门螺杆菌引起的胃十二指肠病症选自胃炎、慢性胃炎、胃萎缩、胃或十二指肠溃疡、胃癌和MALT淋巴瘤。

[0059] 在另一方面,本发明涉及幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂,其中抑制剂选自:

[0060] (a) 与幽门螺杆菌HopQ的胞外结构域结合的(多)肽配体或肽模拟物配体;

[0061] (b) 与CEACAM家族成员的N-结构域结合的(多)肽配体或肽模拟物配体;

[0062] (c) 编码(a)和(b)的(多)肽配体的核酸分子;

[0063] (d) 与幽门螺杆菌HopQ的胞外结构域结合的核酸配体;

[0064] (e) 与CEACAM家族成员的N-结构域结合的核酸配体;

[0065] (f) 抑制CEACAM家族成员或幽门螺杆菌HopQ的表达的抑制性核酸分子;

[0066] (g) 与幽门螺杆菌HopQ的胞外结构域结合的小分子;以及

[0067] (h) 与CEACAM家族成员的N-结构域结合的小分子。

[0068] 在一个实施方案中,(多)肽配体选自CEACAM家族成员或幽门螺杆菌HopQ的抗体、抗体衍生物、抗体模拟物、肽适配体和可溶性片段。

[0069] 在一个实施方案中,肽模拟物配体选自类肽、 β -肽和D-肽。

[0070] 在一个实施方案中,核酸配体选自DNA适配体、RNA适配体和XNA适配体。

[0071] 在一个实施方案中,抑制性核酸分子选自siRNA、shRNA、miRNA和反义DNA或RNA分子。

[0072] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的插入结构域。

[0073] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B、环C或环D,其中:

- [0074] 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间；
- [0075] 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间；
- [0076] 环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间；并且
- [0077] 环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。
- [0078] 在一个实施方案中，(多)肽配体或肽模拟物配体分别选自CEACAM家族成员或幽门螺杆菌HopQ的可溶性片段及其肽模拟物变体。
- [0079] 在一个实施方案中，幽门螺杆菌HopQ的可溶性片段包含幽门螺杆菌HopQ的插入结构域或其功能性片段。
- [0080] 在一个实施方案中，幽门螺杆菌HopQ的可溶性片段包含幽门螺杆菌HopQ的环A、环B、环C或环D，或者前述任一种的功能性片段，其中：
- [0081] 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间；
- [0082] 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间；
- [0083] 环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间；并且
- [0084] 环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。
- [0085] 在一个实施方案中，CEACAM家族成员选自人CEACAM家族成员、非人灵长类CEACAM家族成员和大鼠CEACAM家族成员。
- [0086] 在一个实施方案中，CEACAM家族成员选自CEACAM1、CEACAM3、CEACAM5和CEACAM6。
- [0087] 在另一方面，本发明涉及免疫原性组合物，其包含
- [0088] (a) 至少一种分离的(多)肽，其包含(i) 幽门螺杆菌HopQ的氨基酸序列；或(ii) 其免疫原性变体；或(iii) (i) 或(ii) 的免疫原性片段；或者
- [0089] (b) 编码根据项目(a)的分离的(多)肽的至少一种核酸分子。
- [0090] 在一个实施方案中，分离的(多)肽是重组(多)肽。
- [0091] 在一个实施方案中，免疫原性片段包含幽门螺杆菌HopQ的胞外结构域。
- [0092] 在一个实施方案中，幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的插入结构域或其功能性片段。
- [0093] 在一个实施方案中，幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B、环C或环D，或者前述任一种的功能性片段，其中：
- [0094] 环A位于幽门螺杆菌HopQ的螺旋H3与链S1之间；
- [0095] 环B位于幽门螺杆菌HopQ的链S2与螺旋H4之间；
- [0096] 环C位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间；并且
- [0097] 环D位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间。
- [0098] 在一个实施方案中，分离的(多)肽是融合蛋白。
- [0099] 在一个实施方案中，核酸分子是DNA或RNA，其中优选地核酸分子包含在载体中。
- [0100] 在一个实施方案中，免疫原性组合物还包含至少一种佐剂。
- [0101] 在一个实施方案中，免疫原性组合物是疫苗。
- [0102] 在一个实施方案中，免疫原性组合物引发包含抗体分泌的免疫应答，其中优选地抗体抑制幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间的相互作用。
- [0103] 在另一方面，本发明涉及如上定义的免疫原性组合物，其用作药物。

[0104] 在另一方面,本发明涉及如上定义的免疫原性组合物,其用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,其中优选地疾病或病症选自幽门螺杆菌感染和由幽门螺杆菌引起的胃十二指肠病症。

[0105] 在一个实施方案中,胃十二指肠病症选自胃炎、慢性胃炎、胃或十二指肠溃疡、胃癌和MALT淋巴瘤。

[0106] 在另一方面,本发明涉及能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段,其用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,其中CEACAM蛋白或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0107] 在一个实施方案中,疾病或病症选自幽门螺杆菌感染和由幽门螺杆菌引起的胃十二指肠病症,其中优选地胃十二指肠病症选自胃炎、慢性胃炎、胃或十二指肠溃疡、胃癌和MALT淋巴瘤。

[0108] 在一个实施方案中,CEACAM蛋白选自CEACAM1、CEACAM3、CEACAM5和CEACAM6。

[0109] 在一个实施方案中,固体支持物是微球体。

附图说明

[0110] 图1示出了幽门螺杆菌使用hu-CEACAM1的N端结构域并结合CEACAM5和CEACAM6而不结合CEACAM8。通过Western印迹和流式细胞术分别分析活的幽门螺杆菌和(a) hu-CEACAM1-Fc和(b) hu-CEACAM5-Fc、hu-CEACAM6-Fc或hu-CEACAM8-Fc的牵出实验(pull down experiment) (n=3)。(c) 人正常胃、胃炎和胃癌针对CEACAM1、CEACAM5和CEACAM6的IHC染色。比例尺条,50 μ m。(d) 通过Western印迹检测hu-CEACAM1 Δ N-Fc或用 α -hu-IgG-FITC对细胞进行染色,并通过流式细胞术分析细菌的荧光强度。(e) 通过流式细胞术分析用GFP标记的CEACAM1变体孵育的幽门螺杆菌,并测量与细菌结合的CEACAM变体的比例。单因素ANOVA, P值=0.009, n.s.:不显著。误差棒表示s.e.m。(f) 与去糖基化的hu-CEACAM1-Fc一起孵育的幽门螺杆菌菌株的牵出实验。

[0111] 图2示出了人病原体对Hu-CEACAM1的应用具有高度选择性。不同的幽门螺杆菌菌株和其他细菌与CEACAM家族成员的结合定量。(a,c和f) 将幽门螺杆菌菌株、卡他莫拉菌(*Moraxella catarrhalis*)、腔隙莫拉菌(*Moraxella lacunata*)和空肠弯曲杆菌(*Campylobacter jejuni*)与hu-CEACAM1、3、5、6和8-Fc一起孵育。洗涤后,裂解细菌,对蛋白质进行SDS-凝胶/Western印迹,并用相应的抗体进行检测,或(b,d和g)用抗hu-IgG-FITC进行染色,并通过流式细胞术分析细菌的荧光强度(3个技术重复)。单因素ANOVA, n.s.:不显著。误差棒表示s.e.m。(e) 在首次用于实验的(**naïve**)健康个体和幽门螺杆菌阳性胃炎的胃活检物中对CEACAM表达的评分。(h) 比较了hu-CEACAM1 (P13688;SEQ ID NO:28)、CEACAM5 (P06731;SEQ ID NO:29)、CEACAM6 (P40199;SEQ ID NO:30)和CEACAM8 (P31997;SEQ ID NO:31)的N端结构域的氨基酸序列。(i) 示出了经纯化的hu-CEACAM1、1 Δ N、6和8-FC表达的考马斯染色的SDS-PAGE。

[0112] 图3示出了与CEACAM1直系同源物结合的幽门螺杆菌。(a) 将来自不同物种的可溶性GFP标记的CEACAM与幽门螺杆菌一起孵育,并通过流式细胞术确定荧光(3个技术重复)。通过Western印迹检测使用(b) 大鼠-CEACAM1-Fc和(c) 大鼠-CEACAM1 Δ N-Fc的细菌牵出(pull-down),或(d)用 α -Hu-IgG-FITC进行染色并通过流式细胞术进行分析。(e) 与表达人、

大鼠和小鼠CEACAM1的CHO细胞结合的幽门螺杆菌的代表性共聚焦图像。未经转染的CHO作为对照。比例尺条：左侧图，25 μ m；右侧图，10 μ m。(f)与幽门螺杆菌一起孵育的未经转染的人-、小鼠-和大鼠-CEACAM1转染的CHO细胞的全细胞裂解物的牵出。洗涤后，裂解细胞，并对蛋白质进行SDS-凝胶/Western印迹，并用相应的 α -CEACAM1抗体进行检测。

[0113] 图4涉及CEACAM1直系同源物和非幽门螺杆菌。(a)比较了人CEACAM1 (P13688; SEQ ID NO:28)、鼠CEACAM1 (P31809; SEQ ID NO:33)和大鼠CEACAM1 (P16573; SEQ ID NO:32)的N端结构域的氨基酸序列。在人和大鼠中相同但与小鼠-CEACAM1序列不同的氨基酸用箭头指示。(b)将活的非幽门菌株(non-pylori strain)与hu-CEACAM1-Fc、CEACAM5-Fc、CEACAM6-Fc、CEACAM8-Fc和CEACAM1 Δ N-Fc一起孵育。在严格洗涤后，裂解细菌并在SDS-凝胶上分离蛋白质并用相应的抗体进行检测。(c)细菌牵出和抗hu-IgG-FITC染色后，通过流式细胞术进行分析细菌的荧光强度比例(3个技术重复)。单因素ANOVA, *:P=0.03, ***:P=0.0001, n.s.:不显著。(d)用于分析幽门螺杆菌-CEACAM相互作用的细菌牵出方案。

[0114] 图5示出了幽门螺杆菌通过HopQ与CEACAM1结合。(a)与hu-CEACAM1-Fc一起孵育的多种幽门螺杆菌敲除菌株的牵出。(b)将幽门螺杆菌菌株G27的全裂解物与hu-CEACAM1-Fc一起孵育，并用蛋白G琼脂糖凝胶(sepharose)进行沉淀。对于质谱分析，蛋白质变性，胰蛋白酶消化，通过MS/MS对肽进行分析，并随后针对幽门螺杆菌G27蛋白质组进行检索。示出了鉴定的选择的外膜蛋白。还分析了显示出高Sequest得分的HopQ和HopZ。(c)与hu-CEACAM1-、CEACAM5-和CEACAM6-Fc结合的幽门螺杆菌菌株P12、P12 Δ hopQ和P12 Δ hopZ的牵出和Western印迹以及FACS分析。

[0115] 图6涉及HopQ作为幽门螺杆菌的hu-CEACAM相互作用配偶体的鉴定。(a)将幽门螺杆菌菌株G27的全裂解物与hu-CEACAM1-Fc一起孵育，并用蛋白G琼脂糖凝胶进行沉淀。对于质谱分析，将蛋白质变性，胰蛋白酶消化，通过MS/MS对肽进行分析，并随后针对幽门螺杆菌G27蛋白质组进行检索。(b)将CHO-CEACAM1、AGS、MKN45和MKN28与myc-His-标记的HopQ一起孵育，并随后与抗c-myc mAb一起孵育，随后与FITC缀合的山羊抗小鼠F(ab')₂一起孵育。并行地，通过用兔抗CEApAb (Dianova) 进行染色来控制CEACAM的存在。通过将细胞与对照IgG抗体而不是HopQ蛋白或一抗一起孵育来测量背景荧光(细线)。通过流式细胞术对样品进行分析。(c)如上所述，将指定的CHO转染子与HopQ和抗CEA pAb一起孵育。随后，通过流式细胞术对样品进行分析(n=3)。(d)从所有大陆(continent)的幽门螺杆菌分离株中收集hopQ基因(NCBI数据库<http://www.ncbi.nlm.nih.gov/>)。应用MEGA6程序以使用邻接法(Neighbor-Joining method)推断DNA相关性。最大复合似然法(Maximum Composite Likelihood method)用于计算进化距离。hopQ基因分组为两种主要的等位基因变体(I型和II型)。如所示，I型等位基因更加多样化并进一步分为Ia和Ib两个亚组。

[0116] 图7示出了HopQ黏附素(adhesin)结构域的X射线结构和结合特性。(a)示出了3+4-螺旋束拓扑(topology)的HopQ^{AD}的带状示意图，该拓扑还见于BabA和SabA黏附素中(图8d)。在大多数Hop家族成员中保守的三个Cys对(Cys102-Cys131、Cys237-Cys269和Cys361-Cys384)夹断位于HopQ黏附素结构域远端的延伸环，以形成具有提高的序列多样性的共同蛋白质表面区域。类似于BabA的4链插入结构域(BabA-ID;图8d)，其包含(house)黏附素的碳水化合物结合位点，HopQ^{AD}具有在螺旋H4与H5之间的 β -发夹插入结构域(HopQ-ID)。(b) HopQ^{AD}或突变体HopQ^{AD} (HopQ^{AD} Δ ID)的平均ELISA滴度(n=4; \pm s.d.)，该突变体HopQ^{AD}缺乏

与提高浓度的C1ND结合的HopQ-ID。HopQ-ID的损失导致结合亲和力降低约10倍。(c) 在远Western印迹实验中用考马斯蓝(“C”)或用HopQ^{AD}(“HopQ”)染色的C1ND的SDS和非变性PAGE。SDS和非变性PAGE显示出除了非糖基化蛋白质(较低条带)之外的C1ND的三种糖基化形式。HopQ^{AD}在非变性条件下选择性地结合C1ND,表明在HopQ-CEACAM相互作用中隐含强蛋白质-蛋白质组分。

[0117] 图8示出了与人CEACAM1N-结构域结合的HopQ^{AD}的等温滴定量热法(ITC)。用250μM HopQ^{AD}滴定的(a) 25μM C1ND或(b) 大肠杆菌表达的C1ND(Ec-C1ND)的ITC注射热(上部)和标准化结合等温线(下部)分别显示出在存在或不存在N-糖基化时等效的等摩尔相互作用。结合亲和力和热力学曲线如插图所示。(c) 在HopQ^{AD}的H5螺旋周围以1.0σ轮廓化的2mFo-DFc电子密度图。(d) BabA₁₋₅₂₇(PDB登记代码4ZH0)、HopQ^{AD}和SabA₁₋₄₆₀(PDB登记代码405J)结构的叠加。BabA和SabA结构都在它们的N末端末端显示了扭结,使它们垂直于核心结构域定位,但是HopQ结构中缺少这种方向变化。α-螺旋核心结构域在所有结构中是保守的,而HopQ中的2链插入结构域(insertion domain, ID)在BabA中被两个另外的β-链延长。先前已经示出了BabA的ID负责黏附至血型受体。根据HopQ拓扑对链和螺旋命名。(e) 本研究中使用的HopQ^{AD}和HopQ^{AD} Δ ID片段的SDS-PAGE和示意图。(f) HopQ^{AD}或突变体HopQ^{AD}(HopQ^{AD} Δ ID)的平均ELISA滴度(n=4; ±s.d.),该突变体HopQ^{AD}缺乏与提高浓度的CEACAM5或8结合的HopQ-ID。

[0118] 图9示出了幽门螺杆菌中hopQ的缺失导致细菌细胞黏附降低并消除CagA递送、IL-8释放和细胞伸长。(a) 与MOI 10的CFSE-DA标记的细菌菌株P12、G27、P12 Δ hopQ和G27 Δ babA Δ sabA一起孵育的CHO-hu-CEACAM1-L、MKN45和AGS细胞的流式细胞术分析(3个技术重复)。示出了平均值±s.e.m。双尾t检验,*P≤0.03。(b) 将CHO-CEACAM1-L细胞与幽门螺杆菌一起培养和不与幽门螺杆菌一起培养。随后通过IP和Western印迹分析CEACAM1的Tyr-磷酸化。过钒酸盐(pervanadate)处理作为阳性对照,CEACAM1的检测作为上样对照(底部)。(c) 用P12、NCTC11637和相应的同基因hopQ-突变体感染AGS细胞。用α-磷酸酪氨酸和α-CagA对印迹进行探测。(d) 通过ELISA测定的AGS的IL-8产生。(e) 用指定的幽门螺杆菌wt和敲除菌株或重新表达wt hopQ基因的NCTC11637 Δ hopQ感染HA标记的HEK293-hu-CEACAM1转染子。(f) 不同感染的AGS的代表性相差显微照片。(g) 用P12、P12 Δ hopQ或重新表达wt hopQ基因的P12 Δ hopQ/hopQ感染AGS细胞6小时(3个技术重复)。(h) 每4×10⁵个AGS细胞用2、5、10或20μg α-CEACAM Ab预处理(泳道3-6)。孵育30分钟后,向细胞添加MOI 20的野生型幽门螺杆菌。(i) 野生型幽门螺杆菌以每8×10⁶个细菌2、5、10或20μg α-HopQ预处理(泳道3-6),然后添加至AGS细胞。感染6小时后,对细胞进行照相并收获细胞,随后用α-PY99和α-CagA进行免疫印迹。底部图示出了在五个不同的0.25-mm²视野中每个实验中伸长的AGS细胞的定量(3个技术重复)。误差棒示出了平均值±s.d.。(j) 在低微摩尔浓度下用对应于HopQ-ID的HopQ来源肽(aa 190-218)预孵育细胞抑制了HopQ依赖性CagA磷酸化以及细胞伸长。

[0119] 图10示出了胃细胞系中CEACAM表达模式和CagA磷酸化的表征。用mAb对不同细胞生长阶段的胃细胞系的hu-CEACAM1、CEACAM5和CEACAM6进行染色,并且(a) 通过FITC-缀合的二抗染色,并随后通过流式细胞术监测CEACAM细胞表面表达,或(b) 对细胞裂解物进行SDS-凝胶/Western印迹,并用相应的抗体进行检测。(c) 用wt幽门螺杆菌菌株P12和重要黏附素的多种同基因突变体(BabA、SabA和OipA或双突变体BabA/SabA)感染AGS细胞。使用MOI 50感染细胞6小时。用α-PY-99对印迹进行探测。(d) 用载体对照转染HEK293细胞,随后用指

定的cagPAI阳性幽门螺杆菌菌株P12、NCTC11637和cagPAI阴性菌株Ka89进行MOI 50感染持续6小时。(e)用所示的CEACAM表达载体转染HEK293细胞48小时,随后用wt幽门螺杆菌菌株NCTC11637进行MOI 50感染持续6小时。抗CagA检测用作相等细菌加载的对照。抗GAPDH检测用作细胞裂解物加载对照。(f)用针对转录因子Myc、STAT3、CreATF2/CREB、GRE的指定的萤光素酶报道构建体和作为阴性对照的pTAL-萤光素酶转染CHO-hu-CEACAM1-4L。然后用幽门螺杆菌wt、同基因hopQ缺失突变体感染经转染的细胞或不处理,随后如所示出的那样将萤光素酶活性作为相对光单位(Relative Light Unit,RLU)进行测量(n=3)。(g)在亚微摩尔浓度下用HopQ^{AD}预孵育AGS细胞抑制P12诱导的CagA磷酸化,以及细胞伸长,而HopQ-ID缺失突变体(HopQ^{AD} Δ ID)则没有。类似地,用HopQ-ID肽(aa 190-218;参见图9j)预孵育阻断了P12诱导的CagA磷酸化,尽管与完整HopQ^{AD}相比浓度高约10至20倍。

[0120] 图11示出了幽门螺杆菌在大鼠胃的定植依赖于HopQ。(a)针对大鼠-HCACAM1的大鼠胃的IHC染色。(b)雄性Sprague dawley大鼠(数据来自一个实验,其中每组8只大鼠)用SS1和SS1 Δ hopQ菌株经口感染两次。示出了平均值±s.e.m。双尾t检验,*P=0.02。(c)感染的大鼠胃的苏木精/伊红染色。

[0121] 图12示出了仅幽门螺杆菌的SS1菌株可定植大鼠胃。(a)用通过流式细胞术和Western印迹分析的幽门螺杆菌wt菌株SS1、SS1 Δ hopQ、重新表达wt hopQ基因的SS1 Δ hopQ以及hu-CEACAM1-Fc和大鼠-CEACAM1-Fc进行的牵出实验。(b)大鼠胃活检物分离的RNA中大鼠-CEACAM1的表达。NTC:无模板对照,NEC:无酶对照。

[0122] 图13(a)示出了与人CEACAM1的N端结构域(huCEACAM1-ND)结合的HopQ黏附素结构域(HopQ^{AD})的X-射线结构。形成与CEACAM1-ND的接触界面的HopQ环包含SEQ ID NO:15(菌株P12的HopQ)的第123-136位残基(环A)、第152-180位残基(环B)和第258-290位残基(环C)。HopQ插入结构域(参见例如SEQ ID NO:1的第210至238位残基)和SEQ ID NO:15的环371-407(环D)与直接结合界面相邻。针对位于CEACAM结合界面内或其附近的肽产生的抗体将具有中和作用,通过空间位阻抑制HopQ-CEACAM缔合。(b)见于幽门螺杆菌菌株P12中的代表性序列以及四个环的共有序列。共有序列基于来自不同临床幽门螺杆菌分离株的87个代表性HopQ等位基因的多序列比对,其中单个氨基酸上方的柱的高度表示HopQ等位基因之间的同一性程度。序列保守标志表示各环中可能的氨基酸序列变异,其中氨基酸单字母符号的高度代表其概率。

[0123] 发明详述

[0124] 虽然在上下文中详细描述了本发明,但应当理解的是,本发明不限于本文中所述的具体方法、方案和试剂,因为这些可变化。还应当理解的是,本文中使用的术语仅用于描述一些特定实施方案的目的,并不旨在限制本发明的范围,本发明的范围将仅由所附权利要求书限定。除非另外限定,否则本文中使用的所有技术和科学术语均具有与本领域普通技术人员通常理解的相同的含义。

[0125] 在下文中,将描述本发明的某些要素。这些要素可随一些具体实施方案而列出,然而,应当理解的是,其可以以任何方式和以任意数量组合以产生另外的实施方案。不应将不同描述的实施例和优选实施方案解释为将本发明限制于仅明确描述的一些实施方案。该描述应当被理解为支持并且涵盖将明确描述的实施方案与任意数量的公开和/或优选要素组合的实施方案。此外,除非上下文另有指明,否则本申请中所有描述的要素的任意排列和组

合都应被认为被本申请的说明书公开。

[0126] 优选地,本文中使用的术语如以下中所述进行限定:“*A multilingual glossary of biotechnological terms (IUPAC Recommendations)*”,H.G.W.Leuenerger,B.Nagel,和H. Kölbl,编辑,Helvetica Chimica Acta,CH-4010 Basel,Switzerland,(1995)。

[0127] 除非另有说明,否则本发明的实践将采用在本领域文献中解释的化学、生物化学、细胞生物学、免疫学和重组DNA技术的常规方法(参见例如*Molecular Cloning:A Laboratory Manual*,第3版,J.Sambrook等编辑,Cold Spring Harbor Laboratory Press,Cold Spring Harbor 2000)。

[0128] 除非上下文另有要求,否则在该说明书和所附权利要求书通篇,词语“包含/包括”及其变化形式将被理解为意在包括所述成员、整数或步骤或者成员、整数或步骤的组而不排除任何其他成员、整数或步骤或者成员、整数或步骤的组,但是在一些实施方案中,可排除这样的其他成员、整数或步骤或者成员、整数或步骤的组,即该主题在于包括所述成员、整数或步骤或者成员、整数或步骤的组。除非本文另有说明或明显与上下文矛盾,否则在描述本发明的情况下(特别是在权利要求书的情况下)使用的没有数量词修饰的名词以及类似提及应解释为涵盖一个/种或更多个/种。本文中对数值范围的记载仅是为了作为单独指代落入该范围内的每个单独值的速记方法。除非本文另有说明,否则每个单独的值都并入说明书中,如同其在本文中单独记载一样。除非本文中另有说明或在以其他方式明显与上下文矛盾,否则本文中描述的所有方法都可以以任意合适的顺序进行。本文中提供的任何和所有实例或示例性语言(例如,“例如/如”)的使用仅旨在更好地举例说明本发明,而不是对本发明另外要求保护的范围进行限制。本说明书中的任何语言都不应被解释为表示对本发明的实践必不可少的任何未要求保护的要素。

[0129] 在本说明书的正文通篇引用了数个文件。本文中无论是在上文还是在下文中引用的每个文件(包括所有专利、专利申请、科学出版物、制造商的说明书、用法说明等)均在此通过引用整体并入。本文中的任何内容均不应被解释为承认本发明无权凭借在先发明而早于这样的公开内容。

[0130] 幽门螺杆菌特异性地定植人胃上皮细胞,并且是溃疡病和胃癌发生的主要致病因子。发明人已经鉴定了作为所有人幽门螺杆菌分离株的重要受体的癌胚抗原相关细胞黏附分子(CEACAM)家族成员,并且表明HopQ是特异性结合人CEACAM1、CEACAM3、CEACAM5和CEACAM6的新的表面暴露的黏附素。与CEACAM1 N-结构域结合的幽门螺杆菌诱导CEACAM1介导的信号传导,并且HopQ-CEACAM1相互作用使得毒力因子CagA能够易位到宿主细胞中,允许定植于大鼠感染模型中并增强促炎介质(例如白介素-8)的释放。基于HopQ和HopQ-CEACAM复合体的晶体结构,发明人已经发现HopQ的细胞外3+4螺旋束结构域中的 β -发夹插入结构域和四个特定的环结合区参与CEACAM结合。来源于插入结构域的肽竞争性抑制HopQ介导的CagA毒力途径的激活,遗传或抗体介导的HopQ功能的消除也是如此。总之,本发明将HopQ-CEACAM相互作用鉴定为对抗幽门螺杆菌相关疾病的新治疗靶标。

[0131] 本发明提供了幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂,其用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法。

[0132] 本发明还提供了幽门螺杆菌HopQ与CEACAM家族成员之间相互作用的抑制剂在制

备用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物中的用途。

[0133] 本发明还提供了在对象中预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,所述方法包括向所述对象施用幽门螺杆菌HopQ与CEACAM家族成员之间相互作用的抑制剂。

[0134] 根据本发明,由幽门螺杆菌引起或与其相关的疾病或病症优选选自幽门螺杆菌感染和由幽门螺杆菌引起的胃十二指肠疾病。

[0135] 本文中使用的术语“感染”是指致病因子(例如,幽门螺杆菌)侵入对象的身体组织、其增殖以及组织对这些因子及其产生的毒素的反应。

[0136] 本文中使用的术语“胃十二指肠病症”(或简称“胃病”)是指影响胃和邻近的十二指肠的病症。“由幽门螺杆菌引起的胃十二指肠病症”是本领域技术人员已知的,并且包括例如胃炎、慢性胃炎、胃萎缩、胃或十二指肠溃疡、胃癌(也称为胃癌症)和MALT淋巴瘤。

[0137] 本文中使用的术语“对象”涉及任何生物体,例如脊椎动物,特别是任何哺乳动物,包括人和另外的哺乳动物,例如啮齿动物、兔、牛、绵羊、马、狗、猫、美洲驼(lama)、猪或非人灵长类(例,如猴)的动物。啮齿动物可以是小鼠、大鼠、仓鼠、豚鼠或灰鼠(chinchilla)。在一个实施方案中,对象是人、大鼠或非人灵长类。优选地,对象是人。在一个实施方案中,对象是患有或怀疑患有疾病或病症(特别是本文中所公开的疾病或病症)的对象,在本文中也称为“患者”。

[0138] 本文中使用的术语“预防”可以指停止/抑制疾病或病症的发作(例如,通过预防性治疗)。其还可以指发作的延迟、症状的频率降低,或与疾病或病症相关的症状的严重程度降低(例如,通过预防性治疗)。

[0139] 本文中使用的术语“治疗”涉及改善患者的健康状况和/或延长(提高)生命的任何治疗。

[0140] 本文中使用的术语“药物”是指用于进行治疗(即用于预防或治疗疾病或病症)的物质/组合物。根据本发明,术语“疾病”或“病症”是指任何病理状态,特别是本文中所限定的疾病或病症。

[0141] 癌胚抗原相关细胞黏附分子(CEACAM)家族是免疫球蛋白相关脊椎动物糖蛋白的公知的家族(参见例如Tchoupa等,2014)。CEACAM家族成员通常包含N末端细胞外Igv样结构域,其后可以是多至6个细胞外Igc₂样结构域,并通过C端跨膜结构域(TM螺旋)或C端GPI-锚锚定在细胞膜中。Igv样结构域也称为N端结构域或N-结构域。例如,人CEACAM1包含N-结构域,随后是三个(A1,B,A2) Igc₂样结构域。在一个实施方案中,人CEACAM1的N-结构域包含人CEACAM1的第35至142位氨基酸残基、基本上由其组成或由其组成。

[0142] 根据本发明,CEACAM家族成员优选在上皮细胞、内皮细胞和/或免疫细胞(特别是白细胞,例如T细胞、B细胞和中性粒细胞)的表面上表达。在一个实施方案中,CEACAM家族成员在上皮细胞(例如,胃上皮细胞)的表面上表达,优选在上皮细胞的顶侧表达。

[0143] 根据本发明,CEACAM家族成员优选选自人CEACAM家族成员、非人灵长类CEACAM家族成员和大鼠CEACAM家族成员。在一个实施方案中,CEACAM家族成员是人CEACAM家族成员。在一个实施方案中,CEACAM家族成员不是CEACAM 8。在一个实施方案中,CEACAM家族成员不是CEACAM4、CEACAM7和CEACAM8。在一个实施方案中,CEACAM家族成员选自CEACAM1、CEACAM3,CEACAM5和CEACAM6。在一个实施方案中,CEACAM家族成员选自CEACAM1、CEACAM5和

CEACAM6。在一个实施方案中,CEACAM家族成员是CEACAM1。人CEACAM1的UniProt ID是P13688。人CEACAM3的UniProt ID是P40198。人CEACAM5的UniProt ID是P06731。人CEACAM6的UniProt ID是P40199。

[0144] 术语“幽门螺杆菌HopQ”和“HopQ”在本文中可互换使用。HopQ是幽门螺杆菌特异性旁系蛋白外膜蛋白家族成员。幽门螺杆菌hopQ(omp27;幽门螺杆菌参考菌株26695中的HP1177)表现出代表两个等位基因家族(I型和II型)的遗传多样性。根据本发明,术语“幽门螺杆菌HopQ”涵盖I型和II型HopQ蛋白。在一个实施方案中,幽门螺杆菌HopQ是I型HopQ蛋白或II型HopQ蛋白。在一个实施方案中,I型HopQ蛋白具有SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:15的氨基酸序列,或者与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:15的氨基酸序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%或99%相似性,优选相同的氨基酸序列。在一个实施方案中,II型HopQ蛋白具有SEQ ID NO:5的氨基酸序列或者与SEQ ID NO:5的氨基酸序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%或99%相似性,优选相同的氨基酸序列。

[0145] “序列相似性”表示相同或代表保守氨基酸替换的氨基酸的百分比。两个氨基酸序列之间的“序列同一性”表示序列之间相同的氨基酸的百分比。用于确定序列相似性(优选序列同一性)的比对可用本领域已知的工具完成,优选使用最佳序列比对,例如,使用CLC主工作台(CLC bio)或Align,使用标准设置,优选EMBOSS::针,矩阵:Blosum62,缺口开放10.0,缺口延伸0.5。

[0146] 在一个实施方案中,抑制剂抑制幽门螺杆菌HopQ与CEACAM家族成员的结合和/或HopQ-CEACAM介导的信号传导。

[0147] 本文中使用的术语“HopQ-CEACAM介导的信号传导”是指CagA毒力途径的激活和/或CagA的磷酸化和/或CagA易位到细胞(例如,上皮细胞)中和/或IL-8诱导和/或细胞伸长。在一个实施方案中,HopQ-CEACAM介导的信号传导是指CagA易位到细胞(例如,上皮细胞)中、IL-8诱导和细胞伸长。在一个实施方案中,HopQ-CEACAM介导的信号传导是指CagA易位到细胞(例如,上皮细胞)中。

[0148] 在一个实施方案中,抑制剂抑制(例如竞争性抑制)幽门螺杆菌HopQ与CEACAM家族成员的结合、优选抑制与CEACAM家族成员的胞外结构域的结合。

[0149] 本文中使用的术语“胞外结构域”意指非胞质/胞浆蛋白质或包埋在膜中的蛋白质的那些部分,并且包括位于/暴露于细胞表面和/或在周质空间的部分。可通过使用本领域技术人员已知的标准生物信息学工具和/或公共数据库来鉴定此类序列/结构域。在一个实施方案中,胞外结构域还缺乏N端分泌序列。

[0150] 与CEACAM家族成员结合,术语“胞外结构域”可以指所述成员整个细胞外部分,其优选地包含N-结构域,取决于具体的CEACAM家族成员,可随后是一个或更多个Igc2样结构域。在一个实施方案中,CEACAM家族成员的胞外结构域包含N-结构域和1、2、3、4、5或6个Igc2样结构域、基本上由其组成或由其组成。在一个实施方案中,CEACAM家族成员的胞外结构域包含N-结构域、基本上由其组成或由其组成。在一个实施方案中,胞外结构域是N-结构域。当与CEACAM家族成员的胞外结构域结合使用时,术语“片段”可以指N-结构域和/或一个或更多个Igc2样结构域。术语“片段”还可以指N-结构域和/或一个或更多个Igc2样结构域的片段,前提是这些片段能够与幽门螺杆菌HopQ相互作用和/或与其结合(也称为HopQ结合片

段)。

[0151] 与幽门螺杆菌HopQ结合,术语“胞外结构域”可以指幽门螺杆菌HopQ的整个细胞外部分,即缺乏C端跨膜结构域的全长蛋白质。在一个实施方案中,胞外结构域还缺少N端β-链和/或分泌序列。在一个实施方案中,胞外结构域对应于包含SEQ ID NO:1的第37至463位残基、基本上由其组成或由其组成的氨基酸序列。在一个实施方案中,胞外结构域包含幽门螺杆菌HopQ的插入结构域、基本上由其组成或由其组成。在一个实施方案中,胞外结构域是幽门螺杆菌HopQ的插入结构域。在一个实施方案中,幽门螺杆菌HopQ的胞外结构域包含幽门螺杆菌HopQ的环A、环B、环C和/或环D(环A、环B和/或环C)、基本上由其组成或由其组成。当与幽门螺杆菌HopQ的胞外结构域结合使用时,术语“片段”优选是指能够与CEACAM家族成员相互作用和/或与其结合的片段(也称为CEACAM结合片段)。

[0152] 本文中使用的术语“插入结构域”是指幽门螺杆菌HopQ的细胞外3+4螺旋束结构域中的β-发夹插入结构域,更特别是螺旋H4与H5之间的β-发夹插入结构域,其参与CEACAM结合。插入结构域在本文中也称为HopQ-ID。在一个实施方案中,插入结构域对应于包含SEQ ID NO:1的第210至238位残基、基本上由其组成或由其组成的氨基酸序列。

[0153] 本文中使用的术语“环A”是指位于幽门螺杆菌HopQ的螺旋H3与链S1之间的环。

[0154] 在一个实施方案中,环A包含氨基酸序列CGGYX_{a5}X_{a6}X_{a7}PX_{a9}EX_{a11}X_{a12}QK (SEQ ID NO: 17)、基本上由其组成或由其组成,

[0155] 其中

[0156] X_{a5}是选自T和Y的氨基酸或缺失;

[0157] X_{a6}是选自K和N的氨基酸或缺失;

[0158] X_{a7}是选自S、K、N和T的氨基酸或缺失;

[0159] X_{a9}是选自G、S、Q、R、T、I和V的氨基酸或缺失;

[0160] X_{a11}是选自N和G的氨基酸或缺失;并且

[0161] X_{a12}是选自N和H的氨基酸或缺失。

[0162] 在一个实施方案中,环A包含SEQ ID NO:21的氨基酸序列、基本上由其组成或由其组成。在一个实施方案中,环A对应于包含SEQ ID NO:15的第123至136位残基、基本上由其组成或由其组成的氨基酸序列。

[0163] 本文中使用的术语“环B”是指位于幽门螺杆菌HopQ的链S2与螺旋H4之间的环。

[0164] 在一个实施方案中,环B包含氨基酸序列CGGX_{b4}X_{b5}X_{b6}X_{b7}X_{b8}GX_{b10}X_{b11}X_{b12}X_{b13}X_{b14}X_{b15}GX_{b17}X_{b18}X_{b19}LX_{b21}AX_{b23}KX_{b25}X_{b26}SLSI (SEQ ID NO:18)、基本上由其组成或由其组成,

[0165] 其中

[0166] X_{t4}是选自S、G、N、T和F的氨基酸或缺失;

[0167] X_{b5}是选自T和I的氨基酸或被删除;

[0168] X_{b6}是选自N、G和K的氨基酸或缺失;

[0169] X_{b7}是选自S和A的氨基酸或缺失;

[0170] X_{b8}是选自N和D的氨基酸或缺失;

[0171] X_{b10}是选自Q、K和R的氨基酸或缺失;

[0172] X_{b11}是选自T、V和S的氨基酸或缺失;

[0173] X_{b12}是选自H、Q和Y的氨基酸或缺失;

- [0174] X_{b13}是选自S和N的氨基酸或缺失；
- [0175] X_{b14}是选自S、P和N的氨基酸或缺失；
- [0176] X_{b15}是选自N和S的氨基酸或缺失；
- [0177] X_{b17}是选自T和V的氨基酸；
- [0178] X_{b18}是选自N和S的氨基酸；
- [0179] X_{b19}是选自T、L和M的氨基酸或缺失；
- [0180] X_{b21}是选自K和P的氨基酸或缺失；
- [0181] X_{b23}是选自D、G和A的氨基酸或缺失；
- [0182] X_{b25}是选自N和G的氨基酸或缺失；并且
- [0183] X_{b26}是选自V和S的氨基酸或缺失。
- [0184] 在一个实施方案中，环B包含SEQ ID NO:22或SEQ ID NO:23的氨基酸序列、基本上由其组成或由其组成。在一个实施方案中，环B对应于包含SEQ ID NO:15的第152-180位残基的氨基酸序列、基本上由其组成或由其组成。
- [0185] 本文中使用的术语“环C”是指位于幽门螺杆菌HopQ的螺旋H5与螺旋H6之间的环。
- [0186] 在一个实施方案中，环C包含氨基酸序列CPX_{c3}LIX_{c6}X_{c7}X_{c8}X_{c9}X_{c10}X_{c11}X_{c12}X_{c13}X_{c14}X_{c15}X_{c16}X_{c17}X_{c18}NX_{c20}PSWQX_{c25}X_{c26}X_{c27}X_{c28}X_{c29}KNX_{c32}C (SEQ ID NO:19)、基本上由其组成或由其组成，
- [0187] 其中，
- [0188] X_{c3}是选自M、I和V的氨基酸或缺失；
- [0189] X_{c6}是选自A和G的氨基酸或缺失；
- [0190] X_{c7}是选自K和R的氨基酸或缺失；
- [0191] X_{c8}是选自S和T的氨基酸或缺失；
- [0192] X_{c9}是选自S和T的氨基酸或缺失；
- [0193] X_{c10}是选自S、N和G的氨基酸或缺失；
- [0194] X_{c11}是选自G、N、E、S和D的氨基酸或缺失；
- [0195] X_{c12}是选自S、G和N的氨基酸或缺失；
- [0196] X_{c13}是选自S、M、G、N和T的氨基酸或缺失；
- [0197] X_{c14}是选自G、A、T、S、N和M的氨基酸或缺失；
- [0198] X_{c15}是选自G、N、T、A和V的氨基酸或缺失；
- [0199] X_{c16}是选自A、N、G和S的氨基酸或缺失；
- [0200] X_{c17}是选自T、N、A、G和S的氨基酸或缺失；
- [0201] X_{c18}是选自T和A的氨基酸或缺失；
- [0202] X_{c20}是选自T和A的氨基酸或缺失；
- [0203] X_{c25}是选自T和I的氨基酸或缺失；
- [0204] X_{c26}是选自A、S、T和N的氨基酸或缺失；
- [0205] X_{c27}是选自G和S的氨基酸或缺失；
- [0206] X_{c28}是选自G和N的氨基酸或缺失；
- [0207] X_{c29}是选自G、L和S的氨基酸或缺失；并且
- [0208] X_{c32}是选自S和A的氨基酸或缺失。

[0209] 在一个实施方案中,环C包含SEQ ID NO:24或SEQ ID NO:25的氨基酸序列、基本上由其组成或由其组成。在一个实施方案中,环C对应于包含SEQ ID NO:15的第258至290位残基的氨基酸序列、基本上由其组成或由其组成。

[0210] 当与环A、B和/或C结合使用时,术语“功能性片段”优选是指能够与CEACAM家族成员相互作用和/或与其结合的片段(也称为CEACAM结合片段)。

[0211] 本文中使用的术语“环D”是指位于幽门螺杆菌HopQ的螺旋H7与螺旋H8之间的环。

[0212] 在一个实施方案中,环D包含氨基酸序列SSX_{d3}X_{d4}LKX_{d7}Y1X_{d10}KCDX_{d14}SX_{d16}X_{d17}SX_{d19}X_{d20}X_{d21}X_{d22}X_{d23}NMX_{d26}X_{d27}X_{d28}X_{d29}X_{d30}KX_{d32}X_{d33}X_{d34}WGX_{d37}GCAG (SEQ ID NO:20)、基本上由其组成或由其组成,

[0213] 其中

[0214] X_{d3}是选自G和D的氨基酸或缺失;

[0215] X_{d4}是选自H和Y的氨基酸或缺失;

[0216] X_{d7}是选自D和N的氨基酸或缺失;

[0217] X_{d10}是选自G和R的氨基酸或缺失;

[0218] X_{d14}是选自M、A和V的氨基酸或缺失;

[0219] X_{d16}是选自A和G的氨基酸或缺失;

[0220] X_{d17}是选自I和V的氨基酸或缺失;

[0221] X_{d19}是选自S和G的氨基酸或缺失;

[0222] X_{d20}是任意氨基酸或缺失;

[0223] X_{d21}是任意氨基酸或缺失;

[0224] X_{d22}是任意氨基酸或缺失;

[0225] X_{d23}是选自T、A和S的氨基酸或缺失;

[0226] X_{d26}是选自T和A的氨基酸或缺失;

[0227] X_{d27}是选自M、P、A和Q的氨基酸或缺失;

[0228] X_{d28}是选自Q、R、K和H的氨基酸或缺失;

[0229] X_{d29}是选自S和N的氨基酸或缺失;

[0230] X_{d30}是选自Q和M的氨基酸或缺失;

[0231] X_{d32}是选自N和S的氨基酸或缺失;

[0232] X_{d33}是选自N和T的氨基酸或缺失;

[0233] X_{d34}是选自T、N和I的氨基酸或缺失;并且

[0234] X_{d37}是选自N和K的氨基酸或缺失。

[0235] 在一个实施方案中,环D包含SEQ ID NO:26或SEQ ID NO:27的氨基酸序列、基本上由其组成或由其组成。在一个实施方案中,环D对应于包含SEQ ID NO:15的第371至407位残基的氨基酸序列、基本上由其组成或由其组成。

[0236] 在一个实施方案中,抑制剂选自:

[0237] (a) 与幽门螺杆菌HopQ结合、优选与幽门螺杆菌HopQ的胞外结构域结合的(多)肽配体或肽模拟物配体;

[0238] (b) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的(多)肽配体或肽模拟物配体;

- [0239] (c) 编码(a)和(b)的(多)肽配体的核酸分子;
- [0240] (d) 与幽门螺杆菌HopQ结合、优选与幽门螺杆菌HopQ的胞外结构域结合的核酸配体;
- [0241] (e) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的核酸配体;
- [0242] (f) 抑制CEACAM家族成员或幽门螺杆菌HopQ的表达的抑制性核酸分子;
- [0243] (g) 与幽门螺杆菌HopQ结合、优选与幽门螺杆菌HopQ的胞外结构域结合的小分子; 以及
- [0244] (h) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的小分子。
- [0245] 本文中使用的术语“(多)肽配体”意指CEACAM家族成员的配体或幽门螺杆菌HopQ的配体,其为(多)肽,其中术语“(多)肽”是指作为肽或多肽的分子。
- [0246] 术语“肽”一般涉及包含至少2个、至少3个、至少4个、至少6个、至少8个、至少10个、至少12个或至少14个,并且优选多至8个、10个、12个、14个、16个、18个、20个、25个、30个、50个或100个通过肽键连接在一起的连续氨基酸的物质。术语“多肽”和“蛋白质”涉及大肽,优选具有多于100个氨基酸的肽,但术语“肽”、“多肽”和“蛋白质”在本文中一般可互换使用。
- [0247] 根据本发明的(多)肽优选是分离的。术语“分离的(多)肽”意指该(多)肽与其天然环境分离。分离的(多)肽可处于基本上纯化和/或纯的状态。术语“基本上纯化的”或“基本上纯的”意指该(多)肽基本上不含其他物质,例如天然或在体内与其一起存在和/或与其相关的物质,例如其他蛋白质、核酸、脂质和碳水化合物。在一些实施方案中,根据本发明的(多)肽是(化学)合成的。
- [0248] 根据本发明,(多)肽配体优选选自CEACAM家族成员或幽门螺杆菌HopQ的抗体、抗体衍生物、抗体模拟物、肽适配体和可溶性片段。
- [0249] 术语“抗体”(也称为免疫球蛋白,Ig)是指至少包含通过二硫键相互连接的两条重(H)链和两条轻(L)链的糖蛋白。每条重链包含重链可变区(在本文中缩写为VH)和重链恒定区。每条轻链包含轻链可变区(在本文缩写为VL)和轻链恒定区。VH和VL区可被进一步细分为称为互补决定区(CDR)的高变区,其散布有称为框架区(FR)的更保守的区域。每个VH和VL由按照以下顺序从氨基端至羧基端排列的三个CDR和四个FR构成:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4。重链和轻链的可变区包含与抗原相互作用的结合结构域。抗体的恒定区可介导免疫球蛋白与宿主组织或因子的结合,所述宿主组织或因子包括免疫系统的多种细胞(例如,效应细胞)和经典补体系统的第一组分(C1q)。
- [0250] 本文中使用的术语“抗体衍生物”是指包含至少一个抗体可变结构域而不具有抗体(例如IgA、IgD、IgE、IgG、IgM、IgY或IgW)的总体结构,但是仍然能够结合靶分子的分子。所述衍生物可以是但不限于功能性(即靶标结合,特别是特异性靶标结合)抗体片段,例如Fab、Fab2、scFv、Fv、或其部分,或免疫球蛋白的其他衍生物或组合,例如纳米抗体、双抗体、小抗体(minibody)、骆驼(camelid)单结构域抗体、单结构域或Fab片段、可变区的重链和轻链结构域(例如Fd、VL(包括V λ 和V κ)、VH、VHH)以及由通过至少两个结构环连接的免疫球蛋白结构域的两条 β 链组成的小结构域。优选地,抗体衍生物是单价的。更优选地,衍生物是单链抗体,最优选地具有这样的结构:VL-肽接头-VH或VH-肽接头-VL。

[0251] 本文中使用的术语“抗体模拟物”是指与抗体类似可特异性结合抗原但与抗体在结构上不相关的人工(多)肽。其通常显著小于抗体,其中摩尔质量为约3至20kDa。抗体模拟物的一些非限制性实例是affibody、affilin、affimer、alphabody、affitin、anticalin、avimer、DARPin、fynomer、Kunits结构域肽、monobody、蛋白A的Z结构域、 γ B结晶、泛素、胱抑素(cystatin)、来自嗜酸热硫化叶菌(*Sulfolobus acidocaldarius*)的Sac7D、脂质运载蛋白(lipocalin)、膜受体的A结构域、锚蛋白重复基序(ankyrin repeat motive)、Fyn的SH3结构域、蛋白酶抑制剂的Kunits结构域、纤连蛋白的第十个III型结构域、3-或4-螺旋束蛋白、狢狢重复结构域、富含亮氨酸的重复结构域、PDZ结构域、SUMO或SUMO样结构域、免疫球蛋白样结构域、磷酸酪氨酸结合结构域、pleckstrin同源结构域、src同源2结构域或合成肽配体,其例如来自(随机)肽文库。合成肽配体具有用于结合特定靶分子的非天然存在氨基酸序列。

[0252] 肽适配体是设计成用于干扰其他蛋白质相互作用的蛋白质。它们通常由在两端附接于蛋白质支架的可变肽环组成。可变环长度通常由10至20个氨基酸构成,并且支架可以是具有良好溶解性和致密性(compactness)的任何蛋白质,例如硫氧还蛋白-A。本文中使用的术语“肽适配体”还涵盖肽适配体的衍生物,例如affimer蛋白。

[0253] 术语“部分”或“片段”在本文中可互换使用并且是指连续元件。例如,例如氨基酸序列或蛋白质的结构的一部分是指所述结构的连续元件。蛋白质序列的部分或片段优选包含该蛋白质序列的至少6个(特别地至少8个、至少12个、至少15个、至少20个、至少30个、至少50个、至少100个、至少150个、至少160个、至少170个、至少180个、至少190个或至少200个)连续氨基酸的序列。根据本发明,蛋白质序列的一部分或片段优选不包含与蛋白质序列的其他N端和/或C端氨基酸序列的部分或片段的连续。

[0254] 与CEACAM家族成员或幽门螺杆菌HopQ的片段结合使用时,术语“可溶性”是指主要溶于水溶液(例如水、PBS或胞质)的(多)肽,(例如,在pH 6-8下)。术语“主要可溶的”意指在所述水溶液中,大多数(例如>50%或>60%或>70%或>80%或>90%的)(多)肽分子处于可溶状态。在一个实施方案中,这种可溶性片段缺少跨膜结构域或GPI锚。

[0255] 在一个实施方案中,CEACAM家族成员的可溶性片段包含CEACAM家族成员的胞外结构域或其HopQ结合片段、基本上由其组成或由其组成。在一个实施方案中,可溶性片段包含N-结构域或其HopQ结合片段、基本上由其组成或由其组成。

[0256] 在一个实施方案中,幽门螺杆菌HopQ的可溶性片段包含幽门螺杆菌HopQ的胞外结构域或其CEACAM结合片段、基本上由其组成或由其组成。在一个实施方案中,可溶性片段包含以下、基本上由其组成或由其组成:幽门螺杆菌HopQ的插入结构域、环A、环B、环C和/或环D(优选环A、环B和/或环C),或者前述任一种的功能性片段。

[0257] 本发明还涵盖CEACAM家族成员或幽门螺杆菌HopQ的可溶性片段的肽模拟物变体。如下文进一步描述的,还涵盖氨基酸插入变体、氨基酸添加变体、氨基酸缺失变体和/或氨基酸替换变体。根据本发明,这些变体是抑制幽门螺杆菌HopQ与CEACAM家族成员之间相互作用的功能性变体。

[0258] 在一个实施方案中,如下文进一步描述的,可溶性片段还包含可检测标记或标签。在一个实施方案中,可溶性片段还包含一种或多种提高稳定性和/或防止可溶性片段聚集的修饰,如下文结合免疫原性片段进一步描述的。

[0259] 本文中使用的术语“肽模拟物配体”意指CEACAM家族成员的配体或幽门螺杆菌HopQ的配体,其是肽模拟物。

[0260] 本文中使用的术语“肽模拟物”是指具有与相应的(多)肽基本相同的一般结构的化合物,其具有提高其稳定性和/或生物学功能的修饰。肽模拟物包括例如包含相应(多)肽的相同氨基酸序列的那些化合物,其在两个或更多个氨基酸之间具有改变的主链。作为替代或补充,肽模拟物可包含合成或非天然存在的氨基酸代替天然存在的氨基酸。示例性肽模拟物包括拟肽、 β -肽和D-肽。

[0261] 本文中使用的术语“肽模拟物变体”意指给定天然亲本(多)肽的肽模拟物衍生物,例如CEACAM家族成员或幽门螺杆菌HopQ的可溶性片段的肽模拟物衍生物。

[0262] 本文中使用的术语“拟肽”是指肽模拟物,其中每个氨基酸的侧链附加到与 α 碳相对的氨基酸的氮原子。例如,拟肽可以被认为是N-取代的甘氨酸,其具有NRCH₂CO的一般结构的重复单元,并且具有与相应多肽相同或基本相同的氨基酸序列。

[0263] β -肽由 β 氨基酸组成,其氨基与 β 碳键合而不是与 α 碳键合,如在20种标准生物学氨基酸中那样。 β -肽在体外和体内针对蛋白水解降解是稳定的。

[0264] D-肽是D-氨基酸的序列。正如 β -肽一样,D-肽不易于通过蛋白水解在胃中或细胞内降解。

[0265] 根据本发明,核酸分子可以是这样形式的分子,其为单链或双链的,并且是线性的或共价闭合以形成环。在一个实施方案中,核酸分子是DNA或RNA或XNA。

[0266] 在本发明的上下文中,术语“DNA”涉及包含脱氧核糖核苷酸残基并且优选地完全或基本上由脱氧核糖核苷酸残基构成的分子。“脱氧核糖核苷酸”涉及在 β -D-呋喃核糖基的2'-位缺乏羟基的核苷酸。术语“DNA”包括分离的DNA,例如部分或完全纯化的DNA、基本上纯的DNA、合成DNA和重组产生的DNA,并且包括经修饰的DNA,其通过添加、缺失、替换和/或改变一个或更多个核苷酸而与天然存在的DNA不同。这样的改变可包括例如向DNA的末端或在内部(例如在DNA的一个或更多个核苷酸处)添加非核苷酸物质。DNA分子中的核苷酸还可包含非标准核苷酸,例如非天然存在的核苷酸或化学合成的核苷酸。这些改变的DNA可被称为类似物或天然存在DNA的类似物。

[0267] 在本发明的情况中,术语“RNA”涉及包含核糖核苷酸残基并且优选地完全或基本上由核糖核苷酸残基构成的分子。“核糖核苷酸”涉及在 β -D-呋喃核糖基的2'-位具有羟基的核苷酸。术语“RNA”包括分离的RNA,例如部分或完全纯化的RNA、基本上纯的RNA、合成RNA和重组产生的RNA,并且包括经修饰的RNA,其通过添加、缺失、替换和/或改变一个或更多个核苷酸而与天然存在的RNA不同。这样的改变可包括例如向RNA的末端或在内部(例如在RNA的一个或更多个核苷酸处)添加非核苷酸物质。RNA分子中的核苷酸还可包含非标准核苷酸,例如非天然存在的核苷酸或化学合成的核苷酸或脱氧核苷酸。这些改变的RNA可被称为类似物或天然存在RNA的类似物。根据本发明,“RNA”是指单链RNA或双链RNA。在一个实施方案中,RNA是mRNA。在一个实施方案中,RNA是体外转录RNA(in vitro transcribed RNA,IVT RNA)或合成RNA。

[0268] 异种核酸(XNA)是合成的DNA/RNA类似物,其含有非天然组分,例如替代的核碱基、糖或具有不同化学结构的连接骨架。

[0269] 本文中使用的术语“核酸配体”意指CEACAM家族成员的配体或幽门螺杆菌HopQ的

配体,其是核酸分子,例如核酸适配体。

[0270] 核酸适配体(即RNA适配体、DNA适配体和XNA适配体)是一类小RNA核酸配体,其由RNA或者单链DNA或XNA寡核苷酸构成,并且对其靶标具有高特异性和亲和力。与抗体相似,适配体通过识别特定的三维结构与其靶标相互作用。

[0271] 本文中使用的术语“抑制性核酸分子”是指抑制靶分子(例如CEACAM家族成员或幽门螺杆菌HopQ)表达的核酸分子。示例性抑制性核酸分子包括小干扰RNA(siRNA)、小/短发夹RNA(shRNA)、微RNA(miRNA)和反义DNA或RNA分子,所有这些都是本领域技术人员公知的。

[0272] 本文中使用的术语“小分子”是指低分子量(例如, $<900\text{Da}$ 或 $<500\text{Da}$)有机化合物。

[0273] 术语“结合”在本发明的上下文中(例如,与本文中定义的多肽配体、核酸配体或小分子结合)可以指特异性结合。本文中使用的术语“特异性的结合”或“特异性地结合”意指与另一靶标的结合相比,与结合剂(例如多肽配体(例如,抗体))对其具有特异性的靶标(例如表位)的结合较强。“较强的结合”可例如通过较低的解离常数(KD)来表征。在一个实施方案中,如果结合剂能够与预定靶标结合而其不能与其他靶标结合,则该结合剂对所述预定靶标是特异性的。在一个实施方案中,“特异性结合”靶标的结合剂对于该靶标具有小于 10^{-5}M (例如 10^{-6} 、 10^{-7} 、 10^{-8} 、 10^{-9} 、 10^{-10} 、 10^{-11} 和 10^{-12} 或更小)的 K_D 值。给定结合剂的 K_D 值受结合剂的结合速率和解离速率二者影响,并且随温度而变化。在本发明的情况下,优选地在室温下 K_D 值低于上述值。结合条件优选为生理条件。技术人员知晓确定 K_D 值的不同测定。优选的测定系统是竞争测定。

[0274] 在一个实施方案中,与幽门螺杆菌HopQ结合的(多)肽配体或肽模拟物配体或核酸配体或小分子与幽门螺杆菌HopQ的表位结合,该表位包含包含在幽门螺杆菌HopQ的插入结构域、环A、环B、环C和/或环D(优选环A、环B和/或环C)中的至少1、2、3、4、5、6、7或8个氨基酸残基。

[0275] 在一个实施方案中,抑制剂包含在药物组合物中。因此,本发明还提供了药物组合物,其包含如本文中所限定的幽门螺杆菌HopQ与CEACAM家族成员之间相互作用的抑制剂。

[0276] 本发明还提供了鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的体外方法,该方法包括:

[0277] (a) 使(i) CEACAM蛋白或其功能性片段与(ii) 幽门螺杆菌HopQ蛋白或其功能性片段和(iii) 受试化合物接触,以及

[0278] (b) 确定受试化合物是否抑制CEACAM蛋白或其功能性片段与幽门螺杆菌HopQ蛋白或其功能性片段之间的相互作用,

[0279] 其中抑制CEACAM蛋白或其功能性片段与幽门螺杆菌HopQ蛋白或其功能性片段之间相互作用的受试化合物被鉴定为用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物。

[0280] 在一个实施方案中,步骤(b)包括确定受试化合物是否抑制幽门螺杆菌HopQ蛋白或其功能性片段与CEACAM蛋白或其功能性片段结合,其中优选地幽门螺杆菌HopQ蛋白的功能性片段包含胞外结构域或其片段,和/或CEACAM蛋白的功能性片段包含胞外结构域或其片段,优选N-结构域;和或确定受试化合物是否抑制HopQ-CEACAM介导的信号传导。

[0281] 在一个实施方案中,受试化合物选自(多)肽、肽模拟物、核酸分子和小分子。

[0282] 本发明还提供了能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段用于研究幽门螺杆菌感染或者鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0283] 本文中与CEACAM蛋白结合使用的术语“功能性片段”可例如是指CEACAM蛋白的胞外结构域或其片段。

[0284] 本发明还提供了异源表达能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段的细胞用于研究幽门螺杆菌感染或者鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0285] 这种细胞(也称为宿主细胞)可以是原核细胞(例如,细菌细胞)或真核细胞(例如,真菌、植物或动物细胞)。在一个实施方案中,细胞是哺乳动物细胞,例如CHO细胞或HEK293细胞。优选地,细胞是分离的细胞。

[0286] 本发明还提供了异源表达能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段的非人转基因动物用于研究幽门螺杆菌感染或者鉴定用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0287] 本文中使用的术语“非人转基因动物”特别涉及非人哺乳动物,例如啮齿动物、兔、牛、绵羊、马、狗、猫、美洲驼、猪或非人灵长类(例如,猴)。啮齿动物可以是小鼠、大鼠、仓鼠、豚鼠或灰鼠。在一个实施方案中,非人转基因动物是大鼠。

[0288] 本发明还提供了幽门螺杆菌HopQ与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂,其中抑制剂选自:

[0289] (a) 与幽门螺杆菌HopQ的胞外结构域结合的(多)肽配体或肽模拟物配体;

[0290] (b) 与CEACAM家族成员的N-结构域结合的(多)肽配体或肽模拟物配体;

[0291] (c) 编码(a)和(b)的(多)肽配体的核酸分子;

[0292] (d) 与幽门螺杆菌HopQ的胞外结构域结合的核酸配体;

[0293] (e) 与CEACAM家族成员的N-结构域结合的核酸配体;

[0294] (f) 抑制CEACAM家族成员或幽门螺杆菌HopQ的表达的抑制性核酸分子;

[0295] (g) 与幽门螺杆菌HopQ的胞外结构域结合的小分子;以及

[0296] (h) 与CEACAM家族成员的N-结构域结合的小分子。

[0297] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的插入结构域。

[0298] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B、环C或环D。

[0299] 在一个实施方案中,(多)肽配体或肽模拟物配体分别选自CEACAM家族成员或幽门螺杆菌HopQ的可溶性片段及其肽模拟物变体。

[0300] 在一个实施方案中,幽门螺杆菌HopQ的可溶性片段包含幽门螺杆菌HopQ的插入结构域或其功能性片段。

[0301] 在一个实施方案中,幽门螺杆菌HopQ的可溶性片段包含幽门螺杆菌HopQ的环A、环B、环C或环D,或者前述任一种的功能性片段。

[0302] 本发明还提供了免疫原性组合物,其包含

[0303] (a) 至少一种(例如一种、两种、三种、四种或五种或更多种)分离的(多)肽,其包含

(i) 幽门螺杆菌HopQ的氨基酸序列;或(ii) 其免疫原性变体;或(iii) (i) 或(ii) 的免疫原性片段;或者

[0304] (b) 编码根据项目(a) 的分离的(多)肽的至少一种,例如一种、两种、三种、四种或五种或更多种核酸分子。

[0305] 本文中使用的术语“免疫原性的”意指在对象中引发免疫应答的能力,即诱导体液免疫应答和/或细胞介导的免疫应答的能力。“体液免疫应答”由见于细胞外体液中的大分子(例如分泌形抗体、补体蛋白和某些抗微生物肽)介导。“细胞介导的免疫应答”涉及响应于抗原的吞噬细胞、抗原特异性T淋巴细胞的激活,以及多种细胞因子的释放。在一个实施方案中,免疫应答由抗体介导(=抗体应答)。本文中使用的术语“免疫原性片段”和“免疫原性变体”优选地是指这样的片段和变体,其能够引发对该片段和变体所来源的(多)肽具有特异性的免疫应答。

[0306] 在一个实施方案中,幽门螺杆菌HopQ的氨基酸序列是SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:15的氨基酸序列,或者与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:15的氨基酸序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%或99%相似性、优选相同的氨基酸序列。

[0307] 在一个实施方案中,免疫原性片段包含幽门螺杆菌HopQ的胞外结构域、基本上由其组成或由其组成。在一个实施方案中,免疫原性片段包含幽门螺杆菌HopQ的插入结构域、环A、环B、环C和/或环D。在一个实施方案中,免疫原性片段缺少N端β-链和/或N端分泌序列(=信号肽)和/或C端跨膜(TM)结构域。在一个实施方案中,免疫原性片段缺少N端β-链和N端分泌序列和C端TM结构域。在一个实施方案中,免疫原性片段包含SEQ ID NO:1的第37至463位残基、基本上由其组成或由其组成。

[0308] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的插入结构域或其功能性片段。

[0309] 在一个实施方案中,幽门螺杆菌HopQ的胞外结构域是幽门螺杆菌HopQ的环A、环B、环C或环D,或者前述任一种的功能性片段。

[0310] 根据本发明的术语“免疫原性变体”特别地是指免疫原性突变体、剪接变体、构象变体、同种型、等位基因变体、物种变体和同源物,特别是天然存在的那些。等位基因变体涉及基因的正常序列的改变,其意义常常不清楚。完全基因测序通常鉴定给定基因的多种等位基因变体。同源物是与给定核酸或氨基酸序列具有不同物种(或品系)来源的核酸或氨基酸序列。术语“变体”应涵盖任何翻译后修饰变体和构象变体。

[0311] 出于本发明的目的,氨基酸序列的“免疫原性变体”包含免疫原性氨基酸插入变体、氨基酸添加变体、氨基酸缺失变体和/或氨基酸替换变体。在蛋白质的N端和/或C端包含缺失的氨基酸缺失变体也称为N端和/或C端截短变体。

[0312] 氨基酸插入变体包含在特定氨基酸序列中插入单个或两个或更多个氨基酸。在具有插入的氨基酸序列变体的情况下,一个或更多个氨基酸残基被插入到氨基酸序列中的特定位点,但是随机插入并适当地筛选所得产物也是可能的。

[0313] 氨基酸添加变体包含一个或更多个氨基酸(例如1个、2个、3个、5个、10个、20个、30个、50个或更多个氨基酸)的N端和/或C端融合体。

[0314] 氨基酸缺失变体的特征在于从序列中去除一个或更多个氨基酸,例如去除1个、2

个、3个、5个、10个、20个、30个、50个或更多个氨基酸。缺失可发生在蛋白质的任何位置，例如在N端和/或C端。

[0315] 氨基酸替换变体的特征在于序列中的至少一个残基被去除，并且在其位置插入另一个残基。在一个实施方案中，氨基酸替换变体包含多至10个、9个、8个、7个、6个、5个、4个、3个或2个氨基酸的替换。优选修饰发生在氨基酸序列中在同源蛋白或肽之间不保守的位置和/或用具有类似特性的其他氨基酸替换氨基酸。优选地，蛋白质变体中的氨基酸替换是保守氨基酸替换。保守氨基酸替换涉及用同一氨基酸家族（即，其侧链（例如，在电荷和/或大小方面）相关的氨基酸）中的一个氨基酸替换另一个氨基酸。天然存在的氨基酸一般分为四个家族：酸性氨基酸（天冬氨酸、谷氨酸）、碱性氨基酸（赖氨酸、精氨酸、组氨酸）、非极性氨基酸（丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸）和不带电荷的极性氨基酸（甘氨酸、天冬酰胺、谷氨酰胺、半胱氨酸、丝氨酸、苏氨酸、酪氨酸）。苯丙氨酸、色氨酸和酪氨酸有时被共同归类为芳香族氨基酸。然而，也可用具有不同特性的其他氨基酸替换氨基酸，例如用一个或多个亲水性氨基酸替换一个或多个（表面暴露的）疏水性氨基酸以降低或抑制分离的（多）肽的聚集，其中优选地，这些（多）肽的其他特性（例如，其免疫原性或结合特性）不被这样的氨基酸替换损害。

[0316] 根据本发明，给定的参考氨基酸序列和作为所述给定氨基酸序列的变体的氨基酸序列之间的相似性（优选同一性）优选为至少约60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%。相似性或同一性程度优选地针对参考氨基酸序列的整个长度的至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、至少约90%或约100%的氨基酸区域来给出。例如，如果参考氨基酸序列由200个氨基酸组成，则相似性或同一性程度优选地针对至少约20个、至少约40个、至少约60个、至少约80个、至少约100个、至少约120个、至少约140个、至少约160个、至少约180个或约200个氨基酸（优选连续氨基酸）来给出。在一些优选实施方案中，相似性或同一性的程度/百分比针对参考氨基酸序列的整个长度来给出。

[0317] 在一个实施方案中，免疫原性变体是来自另一幽门螺杆菌菌株的等同蛋白。在一个实施方案中，等同蛋白是同源物，优选直系同源物。“直系同源物”是通过从单一祖先基因/蛋白质物种形成（speciation）而不是通过基因重复而相关的同源基因/蛋白质。

[0318] 在一个实施方案中，免疫原性变体包含与选自SEQ ID NO:1、SEQ ID NO:1的第37至463位残基、SEQ ID NO:3、SEQ ID NO:15和SEQ ID NO:5的氨基酸序列具有至少60%、至少70%、至少80%、至少90%、至少95%或至少97%（例如，97%或98%或99%）同一性的氨基酸序列。

[0319] 在一个实施方案中，分离的（多）肽是重组（多）肽。

[0320] 本文中使用的术语“重组（多）肽”意指由重组核酸分子（例如，DNA）在活细胞内表达（例如借助于特定的表达载体）产生的（多）肽。重组核酸分子是通过遗传重组的实验室方法（例如，分子克隆）形成的核酸分子。

[0321] 在一个实施方案中，分离的（多）肽在宿主细胞（优选原核宿主细胞（例如大肠杆菌））中产生。

[0322] 在一个实施方案中，本文中所述的分离的（多）肽还包含可检测标记或标签。本文

中使用的术语“可检测标记或标签”是指允许检测和/或分离和/或固定本文中所述的分离的(多)肽的可检测标记或标签,并且意在包括用于这些目的的本领域中已知的任何标记/标签。特别优选的是亲和标签,例如甲壳质结合蛋白(CBP)、麦芽糖结合蛋白(MBP)、谷胱甘肽-S-转移酶(GST)、聚(His)(例如,6×His或His₆)、**Strep-tag®**、**Strep-tag II®**和**Twin-Strep-tag®**;增溶标签,例如硫氧还蛋白(TRX)、聚(NANP)和SUMO;色谱标签,例如FLAG-标签;表位标签,例如V5-标签、myc-标签和HA-标签;荧光标记或标签(即,荧光色素/荧光团),例如荧光蛋白质(例如GFP、YFP、RFP等)和荧光染料(例如FITC、TRITC、香豆素和菁(cyanine));发光标记或标签,例如萤光素酶;以及其他)酶标记(例如,过氧化物酶、碱性磷酸酶、β-半乳糖苷酶、脲酶或葡萄糖氧化酶)。还包括任何前述标记或标签的组合。

[0323] (多)肽标记或标签的氨基酸序列可以引入在本文中所述的分离的(多)肽的氨基酸序列内的任何位置。例如,其可添加至所述分离的(多)肽的N端和/或C端和/或添加至氨基酸侧链,例如通过与赖氨酸进行EDC-NHS偶联。其适用于非肽标记或标签。

[0324] 在一个实施方案中,分离的(多)肽是融合蛋白。

[0325] 术语“融合蛋白”是指通过将两个或更多个不同的(多)肽或蛋白质连接(优选头尾(即N端至C端或反之亦然)连接)得到具有来源于每种原始蛋白质的功能特性的单个蛋白质而产生的蛋白质。

[0326] 本发明还提供了如本文中所定义的融合蛋白。

[0327] 根据本发明的分离的(多)肽还可包含一种或更多种提高分离的(多)肽的稳定性和/或防止其聚集的修饰。术语分离的(多)肽的“稳定性”特别地涉及其“半衰期”,例如体内半衰期。“半衰期”涉及消除分子的一半活性、量或数目所需的时间。防止聚集还将提高分离的(多)肽的储存稳定性。

[0328] 例如,分离的(多)肽可与半衰期延长模块融合或缀合。这样的模块是本领域技术人员已知的并且包括例如白蛋白、白蛋白结合结构域、免疫球蛋白的Fc区/结构域、免疫球蛋白结合结构域、FcRn结合基序和聚合物。特别优选的聚合物包括聚乙二醇(PEG)、羟乙基淀粉(HES)、透明质酸、聚唾液酸和PEG模拟肽序列。防止分离的(多)肽聚集的修饰对技术人员来说也是已知的,并且包括例如用一个或更多个亲水性氨基酸替换一个或更多个疏水性氨基酸,优选表面暴露的疏水性氨基酸。在一个实施方案中,分离的(多)肽或其免疫原性变体或前述任一种的免疫原性片段包含用亲水性氨基酸替换多至10个、9个、8个、7个、6个、5个、4个、3个或2个(优选5个、4个、3个或2个)疏水性氨基酸,优选表面暴露的疏水性氨基酸。优选地,分离的(多)肽的其他特性(例如其免疫原性)不被这样的替换损害。

[0329] 根据本发明的分离的(多)肽还可与载体材料例如钥孔血蓝蛋白(Keyhole Limpet Hemocyanin, KLH)、BSA、卵清蛋白等融合或缀合,以允许或促进引发免疫应答并且特别是高滴度抗体的方式将相应的抗原呈递给对象的免疫系统。

[0330] 本文中使用的术语“与……融合”特别地是指基因融合,例如通过重组DNA技术。

[0331] 本文中使用的术语“与……缀合”特别地是指得到稳定共价连接的化学和/或酶促缀合。

[0332] 根据本发明的分离的(多)肽还可包含允许将分离的(多)肽靶向递送至给定细胞、组织或器官的氨基酸序列,优选将分离的(多)肽靶向至特定细胞类型(例如树突细胞)的氨

氨基酸序列。合适的氨基酸序列在例如Sioud等,2013和Apostolopoulos等,2013中描述,并且包括例如具有氨基酸序列NWYLPWLGTDNDW (SEQ ID NO:7)的肽。

[0333] 在一个实施方案中,核酸分子是DNA或RNA。

[0334] 本发明还涵盖在严格杂交条件下与根据以上项目(b)的核酸分子杂交的核酸分子。

[0335] 本文中限定的“严格杂交条件”包括在68°C下在5×SSC/5×Denhardt溶液/1.0% SDS中杂交,并在室温下在0.2×SSC/0.1% SDS中洗涤,或包括其本领域公认的等同方案(例如,其中在60°C下在2.5×SSC缓冲液中进行杂交然后在37°C下在低缓冲液浓度下进行数个洗涤步骤并且保持稳定的条件)。可改变盐浓度和温度的参数以实现寡核苷酸与靶核酸之间的最佳同一性水平。关于这样的条件的指导可在本领域中例如通过以下获得:Molecular Cloning:A Laboratory Manual,第3版,J.Sambrook等编辑,Cold Spring Harbor Laboratory Press,Cold Spring Harbor 2000和Ausubel等(编辑),1995,Current Protocols in Molecular Biology,(John Wiley和Sons,N.Y),Unit 2.10。

[0336] 在一个实施方案中,核酸分子是经密码子优化的,例如被密码子优化成用于在除幽门螺杆菌之外的细菌(例如大肠杆菌)中表达或用于在真核细胞例如哺乳动物细胞(例如,CHO细胞、BHK细胞、COS细胞和HEK293细胞)或昆虫细胞(例如,SF9细胞、SF21细胞和High Five™细胞)中表达。

[0337] 在一个实施方案中,核酸分子被容纳/包含在载体中。

[0338] 本文中使用的术语“载体”包括技术人员已知的任何载体,包括质粒载体、黏粒载体、噬菌体载体(例如λ噬菌体)、病毒载体(例如腺病毒、AAV或杆状病毒载体)、或人工染色体载体(例如细菌人工染色体(bacterial artificial chromosome,BAC)、酵母人工染色体(yeast artificial chromosome,YAC)或P1人工染色体(P1artificial chromosome,PAC))。所述载体包括表达载体以及克隆载体。表达载体包括质粒以及病毒载体,并且通常包含期望的编码序列和在特定宿主生物体(例如,细菌、酵母、植物、昆虫或哺乳动物)或在体外表达系统中表达可操作地连接的编码序列所必需的合适DNA序列。克隆载体通常用于改造和扩增某种期望的DNA片段,并且可缺乏表达期望的DNA片段所需的功能序列。

[0339] 在一个实施方案中,免疫原性组合物还包含来自幽门螺杆菌的至少一种另外的抗原。

[0340] 本文中使用的术语“来自幽门螺杆菌的另外的抗原”优选地是指与根据以上项目(a)和(b)的物质(即分离的(多)肽和核酸分子)不同的抗原。

[0341] 在一个优选实施方案中,另外的抗原选自幽门螺杆菌的外膜蛋白和毒力因子蛋白、其免疫原性片段,以及编码这些蛋白质或片段的核酸分子。

[0342] 术语“外膜蛋白”是指与幽门螺杆菌的外膜缔合的蛋白质,其包括整合膜蛋白以及通过N端附接的脂质锚定于膜的脂蛋白。其结构和功能进一步描述在例如Koebnik等,2000中。根据本发明使用的特别优选的幽门螺杆菌外膜蛋白选自BabA、HpaA、Omp18、Omp22和SabA。

[0343] 本文中使用的术语“毒力因子蛋白”是指参与幽门螺杆菌的致病性的蛋白质,例如功能蛋白,例如酶(参见例如Kalali等,2014)。根据本发明的特别优选的毒力因子蛋白是幽门螺杆菌的γ-谷氨酰转肽酶(gGT)(也称为HPGGT或HPG)。合适的HPG蛋白是例如WO 2008/

046650 A1中描述的那些,并且包括HPG的酶促失活形式(S451/452A),其任选地缺乏N端分泌序列。

[0344] 可以是根据本发明的免疫原性组合物的一部分的另外的抗原也是US 2007/0042448 A1或WO 2004/094467 A2中描述的那些。

[0345] 在一个实施方案中,免疫原性组合物还包含至少一种佐剂。

[0346] 术语“佐剂”是指增强针对抗原(例如针对根据以上项目(a)和(b)的物质或本文中限定的来自幽门螺杆菌的另外的抗原)的免疫应答的物质,例如通过提供免疫系统的一般刺激来增强。合适的佐剂是本领域技术人员已知的并且包括基于毒素的佐剂、基于TLR配体的佐剂、基于核酸/载体的佐剂、基于蛋白质的佐剂、基于聚合物的佐剂、黏膜佐剂、ISCOM基质和任何前述物质的组合。特定佐剂包括但不限于聚阳离子聚合物/肽、免疫刺激性脱氧核苷酸(ODN)、合成KLK肽、神经活性化合物(例如人生长激素)、明矾(alumn)、弗氏完全或不完整佐剂、霍乱毒素(CT)、CTA1-DD、热不稳定性肠毒素(LT)、CT或LT的突变体、聚IC、树突细胞(DC)结合肽和C3d融合蛋白。在一个实施方案中,基于TLR配体的佐剂是TLR5配体,其例如来自细菌鞭毛蛋白的组群,例如在WO 2010/050903 A1、Mori等,2012和Song等,2015中描述的那些。在一个实施方案中,佐剂选自霍乱毒素(CT)、CTA1-DD和热不稳定性肠毒素(LT)。

[0347] 在一个实施方案中,免疫原性组合物是疫苗或包含在疫苗中。

[0348] 术语“疫苗”是指赋予或提高针对特定疾病的免疫力的制剂。根据本发明的疫苗赋予或提高针对由幽门螺杆菌引起或与其相关的疾病或病症(特别是本文中提及的特定疾病)的免疫力。

[0349] 在一个实施方案中,本发明的免疫原性组合物引发包含抗体分泌的免疫应答,其中优选地所述抗体抑制幽门螺杆菌HopQ与如本文中所定义的癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间的相互作用。在一个实施方案中,抗体抑制幽门螺杆菌HopQ与CEACAM家族成员的结合和/或HopQ-CEACAM介导的信号传导。在一个实施方案中,抗体与幽门螺杆菌HopQ的胞外结构域或其片段(例如幽门螺杆菌HopQ的插入结构域、环A、环B、环C和/或环D)结合。在一个实施方案中,抗体与幽门螺杆菌HopQ的表位结合,该表位包含包含在幽门螺杆菌HopQ的插入结构域、环A、环B、环C和/或环D(优选环A、环B和/或环C)中的至少1、2、3、4、5、6、7或8个氨基酸残基。

[0350] 根据本发明,免疫原性/药物组合物包含有效量的活性剂,例如本文中所述的(多)肽或肽模拟物或核酸分子或小分子,以产生期望的反应或期望的作用。

[0351] 根据本发明的免疫原性/药物组合物优选地是无菌的。免疫原性/药物组合物可以以均一剂型提供并且可以以本身已知的方式制备。根据本发明的免疫原性/药物组合物可例如是溶液剂或混悬剂的形式。

[0352] 免疫原性/药物组合物还可包含一种或更多种载体和/或赋形剂,其全部优选地是可药用的。本文中使用的术语“可药用的”是指材料的无毒性,其优选地不与免疫原性/药物组合物的活性剂的作用相互作用。特别地,“可药用”意指由联邦或州政府的管理机构批准或在美国药典、欧洲药典或其他公认的药典中列出用于动物并且更特别是用于人。

[0353] 术语“载体”是指天然或合成性质的有机或无机组分,其中组合了活性组分以促进、增强或实现应用。根据本发明,术语“载体”还包括一种或更多种适合施用于对象的相容性固体或液体填充剂、稀释剂或包封物质。可能的载体物质(例如,稀释剂)是例如无菌水、

林格溶液(Ringer's solution)、乳酸盐林格溶液(Lactated Ringer's solution)、生理盐水、抑菌盐水(例如,含有0.9%苜醇的盐水)、磷酸缓冲盐水(PBS)、Hank溶液、固定油、聚亚烷基二醇、氢化萘和生物相容性丙交酯聚合物、丙交酯/乙交酯共聚物或聚氧乙烯/聚氧丙烯共聚物。在一个实施方案中,载体是PBS。所得溶液或混悬剂优选地与接受者的血液等张。合适的载体及其制剂在Remington's Pharmaceutical Sciences,第17版,1985,Mack Publishing Co.中更详细地描述。

[0354] 本文中使用的术语“赋形剂”旨在包括可存在于药物组合物中并且不是活性成分的例如以下所有物质:盐、黏合剂(例如,乳糖、右旋糖、蔗糖、海藻糖、山梨醇、甘露醇)、润滑剂、增稠剂、表面活性剂、防腐剂、乳化剂、缓冲物质、稳定剂、矫味剂或着色剂。

[0355] 不可药用的盐可用于制备可药用盐并且包括在本发明中。这类可药用盐以非限制性方式包括由以下酸制备的那些:盐酸、氢溴酸、硫酸、硝酸、磷酸、马来酸、乙酸、水杨酸、柠檬酸、甲酸、丙二酸、琥珀酸等。可药用盐也可制备成碱金属盐或碱土金属盐,例如钠盐、钾盐或钙盐。可添加盐来调节离子强度或张度(tonicity)。

[0356] 用于药物组合物的合适防腐剂包括抗氧化剂、柠檬酸、柠檬酸钠、苯扎氯铵、氯丁醇、半胱氨酸、甲硫氨酸、对羟基苯甲酸酯和硫柳汞。

[0357] 用于药物组合物的合适缓冲物质包括盐中的乙酸、盐中的柠檬酸、盐中的硼酸和盐中的磷酸。其他合适的缓冲物质包括精氨酸盐酸盐和精氨酸磷酸盐。

[0358] 合适的稳定剂包括甘油、抗坏血酸盐/酯和组氨酸。

[0359] 根据本发明的免疫原性组合物也可以如US 6,838,089 B1和US 6,372,260 B1中所述配制。

[0360] 根据本发明的免疫原性/药物组合物也可配制成稳定的冻干产品,其用合适的稀释剂重构,所述稀释剂任选地包含一种或更多种如上所述的赋形剂。

[0361] 本发明还提供了如本文中定义的免疫原性组合物,其用作药物。

[0362] 本发明还提供了如本文中所定义的免疫原性组合物,其用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法。

[0363] 本发明还提供了如本文中所定义的免疫原性组合物在制备用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物中的用途。

[0364] 本发明还提供了在对象中预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,所述方法包括向对象施用如本文中所定义的免疫原性组合物。

[0365] 本发明还提供了CEACAM蛋白或其功能性片段,其能够与幽门螺杆菌HopQ相互作用,用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,其中CEACAM蛋白或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0366] 本发明还提供了能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段在制备用于预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的药物的用途,其中CEACAM蛋白或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0367] 本发明还提供了在对象中预防或治疗由幽门螺杆菌引起或与其相关的疾病或病症的方法,所述方法包括向对象施用能够与幽门螺杆菌HopQ相互作用的CEACAM蛋白或其功能性片段。其中CEACAM蛋白或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0368] 本文中使用的术语“固体支持物”是指能够与如本文中所定义的CEACAM蛋白或其

功能性片段结合的任何固体支持物。在一个实施方案中,固体支持物是非细胞固体支持物。这样的非细胞固体支持物可包括支持材料,例如聚合物,特别是生物黏附阳离子聚合物(例如,壳聚糖、聚半乳糖胺、聚赖氨酸、二乙氨基乙基葡聚糖(DEAE)、DEAE-亚胺)。支持物可具有任何可能的结构构型,只要与其结合的分子能够与其相应的结合配偶体(例如,螺杆菌属细菌)结合即可。合适的构型包括球形构型,例如微球体(参见例如WO 2013/164652 A2)。在一个实施方案中,固体支持物是微球体。

[0369] 本发明还提供了抑制胆汁螺杆菌(*Helicobacter bilis*, *H. bilis*)与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂,其用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的方法。

[0370] 本发明还提供了胆汁螺杆菌与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂在制备用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物中的用途。

[0371] 本发明还提供了在对象中预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的方法,所述方法包括向对象施用胆汁螺杆菌与癌胚抗原相关细胞黏附分子(CEACAM)家族成员之间相互作用的抑制剂。

[0372] 根据本发明,由胆汁螺杆菌引起或与其相关的疾病或病症优选选自胆汁螺杆菌感染、胆囊炎、胆结石、胆囊癌和胆管癌。

[0373] 在一个实施方案中,抑制剂抑制(例如,竞争性抑制)胆汁螺杆菌与CEACAM家族成员的结合、优选抑制与CEACAM家族成员的胞外结构域结合、更优选抑制与CEACAM家族成员的N-结构域结合。

[0374] 在一个实施方案中,CEACAM家族成员在上皮细胞、内皮细胞和/或免疫细胞(特别是白细胞,例如T细胞、B细胞和中性粒细胞)的表面上表达。在一个实施方案中,CEACAM家族成员在上皮细胞(例如,胆管上皮细胞)的表面上表达,优选在上皮细胞的顶侧表达。

[0375] 在一个实施方案中,CEACAM家族成员选自人CEACAM家族成员、非人灵长类CEACAM家族成员和大鼠CEACAM家族成员。在一个实施方案中,CEACAM家族成员选自CEACAM1、CEACAM5和CEACAM6。

[0376] 在一个实施方案中,抑制剂选自:

[0377] (a) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的(多)肽配体或肽模拟物配体;

[0378] (b) 编码(a)的(多)肽配体的核酸分子;

[0379] (c) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的核酸配体;

[0380] (d) 抑制CEACAM家族成员表达的抑制性核酸分子;以及

[0381] (e) 与CEACAM家族成员结合、优选与CEACAM家族成员的胞外结构域结合、更优选与CEACAM家族成员的N-结构域结合的小分子。

[0382] 在一个实施方案中,抑制剂包含在药物组合物中。

[0383] 本发明还提供了用于鉴定预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物候选物的体外方法,所述方法包括:

[0384] (a) 使(i) CEACAM蛋白质或其功能性片段与(ii)胆汁螺杆菌和(iii)受试化合物接

触,以及

[0385] (b) 确定受试化合物是否抑制CEACAM蛋白或其功能性片段与胆汁螺杆菌之间的相互作用,

[0386] 其中,抑制CEACAM蛋白或其功能性片段与胆汁螺杆菌之间相互作用的受试化合物被鉴定为用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物候选物。

[0387] 在一个实施方案中,步骤(b)包括确定受试化合物是否抑制胆汁螺杆菌与CEACAM蛋白或其功能性片段的结合,其中优选地CEACAM蛋白的功能性片段包含胞外结构域或其片段,优选N-结构域。

[0388] 在一个实施方案中,CEACAM蛋白选自人CEACAM蛋白、非人灵长类CEACAM蛋白和大鼠CEACAM蛋白。在一个实施方案中,CEACAM蛋白选自CEACAM1、CEACAM5和CEACAM6。

[0389] 在一个实施方案中,受试化合物选自(多)肽、肽模拟物、核酸分子和小分子。

[0390] 在另一方面,本发明涉及能够与胆汁螺杆菌相互作用的CEACAM蛋白或其功能性片段用于研究胆汁螺杆菌感染或者鉴定用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0391] 在另一方面,本发明涉及异源表达能够与胆汁螺杆菌相互作用的CEACAM蛋白或其功能性片段的细胞用于研究胆汁螺杆菌感染或者鉴定用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0392] 另一方面,本发明涉及异源表达能够与胆汁螺杆菌相互作用的CEACAM蛋白或其功能性片段的非人转基因动物用于研究胆汁螺杆菌感染或者鉴定用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物候选物的用途。

[0393] 在上述用途的一个实施方案中,CEACAM蛋白选自人CEACAM蛋白、非人灵长类CEACAM蛋白和大鼠CEACAM蛋白。在一个实施方案中,CEACAM蛋白选自CEACAM1、CEACAM5和CEACAM6。

[0394] 本发明还提供了能够与胆汁螺杆菌相互作用的CEACAM蛋白或其功能性片段,其用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的方法,其中CEACAM蛋白或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0395] 本发明还提供了能够与胆汁螺杆菌相互作用的CEACAM蛋白或其功能性片段在制备用于预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的药物的用途,其中CEACAM蛋白质或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0396] 本发明还提供在对象中预防或治疗由胆汁螺杆菌引起或与其相关的疾病或病症的方法,所述方法包括向对象施用能够与胆汁螺杆菌相互作用的CEACAM蛋白或其功能性片段,其中CEACAM蛋白或其功能性片段与固体支持物(优选非细胞固体支持物)衔接。

[0397] 在上述用途的一个实施方案中,CEACAM蛋白选自人CEACAM蛋白、非人灵长类CEACAM蛋白和大鼠CEACAM蛋白。在一个实施方案中,CEACAM蛋白选自CEACAM1、CEACAM5和CEACAM6。

[0398] 本文中所述的药剂和组合物可通过任何常规途径施用,例如通过肠内施用或通过肠胃外施用(包括通过注射或输注)。在一个实施方案中,施用肠胃外进行,例如皮内、皮下或肌内进行。在一个实施方案中,使用经黏膜施用,例如经口或舌下施用。

[0399] 本文中所述的药剂和组合物以有效量施用。“有效量”是指单独或与另外的剂量一

起实现期望反应或期望作用的量。在治疗特定疾病或治疗特定病症的情况下,期望的反应优选涉及抑制疾病进程。这包括减慢疾病的进展,并且特别是中断或逆转疾病的进展。治疗疾病或治疗病症中的期望的反应也可以是延迟所述疾病或所述病症的发生或预防其发生。本文中所述的药剂或组合物的有效量将取决于待治疗的病症;疾病的严重程度;对象的个体参数,包括年龄、生理状况、身材大小和体重;治疗的持续时间、伴随治疗(如果存在的话)的类型、具体施用途径和类似因素。因此,本文中所述药剂的施用剂量可取决于多个这样的参数。在对象中的反应在初始剂量下不足的情况下,可使用更高的剂量(或通过不同的、更局部化的施用途径实现的有效更高剂量)。

[0400] 本发明还提供了药盒(kit),其包含:如本文中限定的(i)抑制剂或(ii)免疫原性组合物或(iii)CEACAM蛋白或其功能性片段。

[0401] 本文中使用的术语“药盒套件(kit of parts)(简称:药盒)”是指包含一个或更多个容器和任选地数据载体的制品。所述一个或更多个容器可填充有本文中公开的一种或更多种装置或试剂。药盒中可包含另外的容器,其含有例如稀释剂、缓冲剂和另外的试剂。所述数据载体可以是非电子数据载体,例如图形数据载体(例如信息页(information leaflet)、信息表(informatin sheet)、条形码(bar code)或存取码(access code));或电子数据载体,例如软盘、光盘(compact disk,CD)、数字多功能盘(digital versatile disk,DVD)、微芯片或其他基于半导体的电子数据载体。存取码可允许访问数据库,例如互联网数据库、集中式或分散式数据库。所述数据载体可包含关于使用根据本发明的药盒的说明书。

[0402] 通过以下实施例进一步举例说明本发明,所述实施例不应解释为限制本发明的范围。

实施例

[0403] 实施例1:幽门螺杆菌与人胃中表达的CEACAM结合

[0404] 使用牵出和流式细胞术方法,发现了幽门螺杆菌菌株G27和卡他莫拉菌相当的(图2a和b)与重组人CEACAM1-Fc的稳健相互作用(图1a)。作为阴性对照,腔隙莫拉菌不与人CEACAM1结合,作为与幽门螺杆菌密切相关的病原体的空肠弯曲杆菌也是如此(图2a和b)。当测试CEACAM特异性时,观察到幽门螺杆菌与CEACAM5和6的明显相互作用,但未观察到与CEACAM8的相互作用(图1b),并且各N-结构域的比较表明数个残基在CEACAM1、5和6中保守,而非在CEACAM8中(图2h)。幽门螺杆菌也与CEACAM3相互作用(图2c和d)。重要的是,到目前为止测试的所有幽门螺杆菌菌株都与这些CEACAM结合(图2f和g),包括充分表征的参考菌株(26695,J99)和小鼠适应菌株SS1。然而,菌株之间的结合强度不同,其中一些优先与CEACAM1结合,而另一些与CEACAM5和/或CEACAM6结合(图2f和g)。引人注目的是,CEACAM1结合物主要来自高度毒力菌株,具有编码IV型分泌系统(T4SS)用于递送CagA的cag致病岛(cag pathogenicity island,cagPAI),而作为经典cagPAI阴性菌株的TX30显示优选与CEACAM5和6结合(图2f和g)。然后,本发明人分析了CEACAM1、CEACAM5和CEACAM6在正常和发炎的人胃组织以及胃癌中的表达谱。CEACAM1和CEACAM5在上皮细胞的顶侧表达,并且它们的表达以及CEACAM6的表达在胃炎时和胃肿瘤中上调(图1c和图2e)。在感染期间,幽门螺杆菌诱导的应答可因此导致其CEACAM-受体的表达提高。

[0405] 本发明人发现幽门螺杆菌与CEACAM1的N-结构域结合(图1d),因为重组CEACAM1 Δ N不与细菌相互作用,并还观察到幽门螺杆菌与含有N-结构域的所有CEACAM结合并且与单独的N-结构域结合(图1e)。然而,与单独的N结构域结合弱于与N-A1-B CEACAM1变体的结合,N-A1-B CEACAM1变体结合小于N-A1-B-A2变体(图1e),表明这些结构域使CEACAM1-幽门螺杆菌相互作用稳定化,然而结合仅部分依赖于糖基化(图1f)。这种CEACAM结合特性为幽门螺杆菌提供了稳健的上皮黏附,而不依赖于BabA和SabA黏附素使用的Lewis血型抗原。尽管充分描述了CEACAM在胃肠道肿瘤中的过表达,但迄今为止尚未报道它们在胃中幽门螺杆菌诱导的炎症期间的上调,表明病原体具有塑造其自身黏合生态位(adhesive niche)的能力。CEACAM调节的似乎合理的途径是通过转录因子NF- κ B和AP1,两者都在幽门螺杆菌感染期间被诱导并且已知调节CEACAM表达。这些CEACAM受体的上调可补偿定植期间所描述的BabA表达的丧失,从而实现持久的定植。

[0406] 实施例2:螺杆菌属的物种特异性-CEACAM相互作用

[0407] 迄今为止,幽门螺杆菌仅被描述为感染人和非人灵长类。尽管CEACAM在大多数哺乳动物物种中发现并且具有高度保守性,但发明人发现幽门螺杆菌选择性地与人而非小鼠、牛或犬CEACAM1直系同源物结合(图3a)。然而,出乎意料的是,发现了幽门螺杆菌菌株与大鼠-CEACAM1的强相互作用(图3b和d)。同样在这里,这种相互作用是通过大鼠-CEACAM1的N-结构域介导的(图3c和d)。为了证实这些发现,本发明人将人-、小鼠-或大鼠-CEACAM1转染到通常不允许幽门螺杆菌黏附的CHO细胞中。使用共聚焦激光扫描显微术进行观察,观察到幽门螺杆菌重新黏附至表达人和大鼠但非小鼠CEACAM1的CHO细胞(图3e),这可通过来自经转染细胞的裂解物的牵出和Western印迹来证实(图3f和图4d)。这一发现使幽门螺杆菌成为其CEACAM结合不限于一个物种的第一病原体。比较CEACAM1-N结构域的蛋白质序列,在人和大鼠中保守的数个氨基酸在小鼠中不同(即ash10、glu26、asn42、tyr48、pro59、thr66、asn77、val79、val89、ile90、glu103、tyr108)(图4a)。此外,它们的缺乏与小鼠CEACAM1结合的发现可解释感染小鼠与人之间病理学上所见的差异。

[0408] 螺杆菌属包括数个其他物种,即猫螺杆菌(*H. felis*)、猪螺杆菌(*H. suis*)和毕氏螺杆菌(*H. bizzozeronii*),以及人致病性的胆汁螺杆菌和海尔曼螺杆菌(*H. heilmannii*)。当评估这些螺杆菌与人CEACAM的相互作用时,只有胆汁螺杆菌与hu-CEACAM1、5和6结合(图4b和c)。如同幽门螺杆菌,胆汁螺杆菌与hu-CEACAM1的N-结构域相互作用(图4b和c)。这种相互作用可解释胆汁螺杆菌如何设法定植人胆管,其中存在高水平的组成型表达的CEACAM1。

[0409] 实施例3:HopQ是与CEACAM相互作用的螺杆菌黏附素

[0410] 为了鉴定螺杆菌中的CEACAM结合配偶体,本发明人最初筛选了许多螺杆菌突变体,其缺乏已被示出参与宿主细胞相互作用的多种模式的限定的毒力因子(BabA、SabA、AlpA/B、VacA、gGT、脲酶(urease)和CagPAI)。所有这些突变体仍然与hu-CEACAM1结合(图5a)。因此,它们建立了使用幽门螺杆菌裂解物和与蛋白G偶联的重组hu-CEACAM1-Fc的免疫沉淀方法(图6a)。共沉淀物的质谱分析鉴定出两种高度保守的幽门螺杆菌外膜蛋白作为候选CEACAM1黏附素:HopQ和HopZ(图5b)。与hopZ突变体不同,hopQ缺失突变体缺乏CEACAM1结合(图5c)。重要的是,hopQ突变体也不能够与CEACAM5和6结合(图5c)。

[0411] 接下来,发明人测试了重组HopQ与不同胃癌细胞系的结合,并发现HopQ与内源性表达CEACAM的AGS和MKN45二者相互作用(图6b)。HopQ不与CEACAM阴性细胞系MKN28结合。利

用它们的CHO转染子,本发明人发现重组HopQ优先与CEACAM1和5相互作用,并且在较低程度上与CEACAM3和6相互作用。未观察到表达CEACAM4、7或8的CHO细胞的结合(图6c)。

[0412] HopQ是幽门螺杆菌特异性外膜蛋白家族成员,并且与来自其他革兰氏阴性细菌的其他CEACAM结合黏附素(即分别来自脑膜炎奈瑟菌(*Neisseria meningitidis*)和淋病奈瑟菌(*Neisseria gonorrhoeae*)或卡他莫拉菌的Opa蛋白或UspA1)没有显著的同源性,并且因此,它是劫持CEACAM的新细菌因子。与Opa和UspA1相似,HopQ靶向CEACAM中的N端结构域(发明人发现需要折叠蛋白质并依赖于CEACAM序列的相互作用),导致对人CEACAM1、3、5和6的特异性。幽门螺杆菌hopQ(omp27;幽门螺杆菌参考菌株26695中的HP1177)表现出代表两个等位基因家族(Cao&Cover,2002)(I型和II型)的遗传多样性(图6d),其中发现I型等位基因更常见于cag(+)/类型s1-vacA菌株。这两个等位基因共享75至80%的核苷酸序列,并且在氨基酸水平上显示出70%的同源性(Cao&Cover,2002)。发明人观察到HopQ对于CEACAM的结合强度的等位基因差异,其中hopQ I型等位基因看来与CEACAM1结合得更强,而在TX30菌株中发现的II型等位基因更偏好CEACAM5和6。重要的是,hopQ基因型显示出地理变异,与西方菌株相比,hopQ I型等位基因在亚洲更为普遍;并且还发现其与菌株毒力相关,其中I型等位基因与更高的炎症和胃萎缩相关。

[0413] 实施例4:HopQ黏附素结构域的结构和结合特性

[0414] HopQ属于幽门螺杆菌外膜蛋白(Hop's)的旁系同源家族,其中血型抗原结合黏附素BabA和SabA也属于该家族。为了深入了解其结构-功能关系,发明人确定了对应于其预计的胞外结构域的HopQ片段的X射线结构(成熟蛋白质的第17-443位残基,即去除信号肽后;HopQ^{AD};图7a和表1)。在ELISA中,HopQ^{AD}显示出与人CEACAM1的N端结构域(C1ND;第35-142位残基)的强烈的剂量依赖性结合(图7b)。通过用C1ND滴定HopQ^{AD}的等温滴定量热法(isothermal titration calorimetry,ITC)测量的结合曲线显示1:1化学计量,其中解离常数K_d为296±40nM(图8a)。

[0415] HopQ^{AD} X射线结构显示,与BabA和SabA相似,HopQ胞外结构域采用3+4螺旋束拓扑结构,但缺乏将胞外结构域与跨膜区域连接的扩展卷曲螺旋“茎”结构域(图8d)。在BabA中,碳水化合物结合位点完全位于插入螺旋H4与H5之间的4链β-结构域中(图8d)。在HopQ中,在该位置(第180-218位残基)发现了2链β-发夹。去除β-发夹导致稳定的蛋白质,其显示CEACAM1结合亲和力降低约10倍,表明尽管HopQ插入结构域参与结合,但它不包含如BabA中所发现的完整结合位点(图7b)。BabA和SabA黏附素是分别结合Lewis b和唾液酸化的Lewis x和聚糖的凝集素。为了验证HopQ-CEACAM相互作用是否类似地由聚糖驱动,本发明人评价了在天然或变性条件下HopQ与C1ND的结合。远western分析显示HopQ特异性结合折叠但未变性的CEACAM1-N(图7c)。相反,细菌牵出实验显示CEACAM1-Fc去糖基化后结合仅略微降低(图1f),证实蛋白质-蛋白质相互作用是HopQ-CEACAM结合的主要贡献者。

[0416] 表1.HopQAD结构的数据收集和细化统计。

		HopQ ^{AD}
数据收集		
[0417]	空间群	P 1 2 ₁ 1
	晶胞尺寸	
	<i>a, b, c</i> (Å)	57.7, 57.7, 285.7
	α, β, γ (°)	90.0, 90.1, 90.0
	分辨率 (Å)	49.38-2.6 (2.74-2.6)*
	<i>R</i> _{合并}	14.7 (121.0)*
	<i>I/σI</i>	7.3 (0.9)*
CC1/2		
	完全性 (%)	99.3 (57.6)*†
	冗余	99.7 (98.2)*
		4.7 (4.5)*
细化		
[0418]	分辨率 (Å)	285.6-2.6
	反射号	55541
	<i>R</i> _{起作用} / <i>R</i> _{游离}	20.9 / 23.6
	原子数	
	蛋白质	10946
	水	35
	B-因子	
	蛋白质	42.8
	水	60.7
	R.m.s偏差	
	键长 (Å)	0.014
	键角 (°)	1.71

[0419] *最高分辨率显示在括号中。

[0420] †通过应用基于约为0.5的半数据集的平均强度相关系数 (CC1/2) 的截止值来确定分辨率限值。

[0421] 实施例5:HopQ-CEACAM1相互作用引发细胞应答

[0422] 为了进一步研究HopQ可如何影响黏附和细胞应答, 本发明人试图建立可分析HopQ-CEACAM介导的黏附的细胞发病机理模型。因此, 本发明人关于它们的CEACAM表达而表征了通常用于幽门螺杆菌体外实验的多种胃细胞系, 并观察到MKN45、KatoIII和AGS确实表达CEACAM1、CEACAM5和CEACAM6, 而MKN28显示不存在CEACAM (图10a和b)。用CEACAM1-L (含有ITIM基序) 稳定转染CHO细胞。在用幽门螺杆菌野生型菌株P12及其同基因hopQ缺失突变体感染后, 发明人观察到对CHO-CEACAM1-L、MKN45和AGS细胞的黏附显著降低, 而缺乏黏附素BabA和SabA的菌株仅显示略微降低黏附 (图9a)。在CHO-CEACAM1-L细胞中, 发明人观察到在暴露于幽门螺杆菌后CEACAM1 ITIM结构域的酪氨酸磷酸化, 其在5分钟内是明显的, 并且保持长达1小时 (图9b)。CEACAM1 ITIM结构域的磷酸化是引发诱导下游信号级联的SHP1/2募集的公知的初始事件。通过CEACAM的接触依赖性信号传导是调节与感染、炎症和癌症相关的免疫应答的常用手段, 并且这些免疫抑制级联 (immune-dampening cascade) 可能反映了在多种黏膜革兰氏阴性病原体中多种独立的非同源CEACAM相互作用蛋白的出现, 这些病原体包括奈瑟菌属 (*Neisseria*)、嗜血杆菌属 (*Haemophilus*)、埃希菌属 (*Escherichia*)、沙门菌属 (*Salmonella*)、莫拉菌属 (*Moraxella*)。对于幽门螺杆菌, 通过HopQ与人CEACAM1的相互作用可代表在人的定植和慢性感染期间免疫调节信号传导的关键参数。

[0423] 另外, hopQ突变体幽门螺杆菌菌株显示出T4SS依赖性CagA易位的几乎完全丧失(图9c)和IL-8诱导的强烈降低(图9d), 然而其他已知黏附素的丧失对CagA递送没有影响(图10c和d)。

[0424] 为了在独立模型中证实这些数据并补偿稳定转染的细胞中的潜在克隆效应, 本发明人用人CEACAM(1-L、3、4、5、6、7、8)表达质粒瞬时转染HEK293细胞。这些细胞的感染证实了在CHO-CEACAM1-L细胞中观察到的CagA易位缺陷, 其在补充hopQ突变株(P12 Δ hopQ/hopQ)后恢复(图9e和图10e)。此外, 在hopQ缺失后, 细胞伸长(所谓的“蜂鸟表型”)显著降低(图9f和g)。此外, 本发明人观察到幽门螺杆菌以hopQ依赖性方式调节重要的宿主转录因子(例如Myc、STAT3、CreATF2/CREB、GRE和NF- κ B)(图10f)。这些结果表明, HopQ-CEACAM结合导致宿主细胞信号级联的直接和间接改变, 并开始阐明这些HopQ相关的毒力情况。鉴于这些信号传导事件对胃癌发生的重要性, 本发明人探索了CEACAM-HopQ相互作用是否可被靶向以阻止CagA易位和下游效应。实际上, 使用 α -CEACAM1抗体、 α -HopQ抗血清或对应于Hop-ID的HopQ来源肽(aa 190-218)以剂量依赖性方式降低CagA易位(图9h-j), 但相应的对照则没有(图10g)。这些数据表明HopQ-CEACAM1相互作用对于癌蛋白CagA成功转运到上皮细胞中以及调节炎症信号传导是必需的, 并且干扰这种相互作用可阻止CagA易位, 从而表明HopQ靶向幽门螺杆菌疫苗接种或免疫治疗的转化潜力。

[0425] 实施例6: hopQ的缺失消除了幽门螺杆菌感染在大鼠模型中的定植

[0426] 当发明人发现HopQ与人和大鼠CEACAM结合但不与小鼠CEACAM结合时, 他们使用幽门螺杆菌感染的大鼠模型确定了HopQ在体内的作用。已经观察到CEACAM1在正常大鼠胃中表达(图11a和图12b), 发明人用已知感染啮齿动物的不同幽门螺杆菌菌株感染大鼠。虽然所有菌株在体外都与大鼠CEACAM1结合, 但只有SS1能够高效地定植大鼠(图12a)。与野生型SS1菌株相比, hopQ缺陷型SS1菌株在可检测水平下不能够定植大鼠, 并且不可诱导炎症应答(图11b和c)。因此, 在该模型中, HopQ看来也是介导幽门螺杆菌定植的重要因素。

[0427] 实施例7: HopQ^{AD}和C1ND复合体的结构

[0428] 确定了从大肠杆菌重组产生和纯化的人CEACAM1的HopQ黏附素结构域与非糖基化N端结构域之间的复合体的结构(图13a和表2)。该结构显示结合CEACAM N端结构域的HopQ的接触表面由三个扩展环形成: HopQ_123-136(环A)、HopQ_152-180(环B)和HopQ_258-290(环C)。通过与由HopQ-ID形成的 β -发夹键合的主链氢键使环HopQ_123-136的结合构象稳定化。因此, 发现HopQ-ID的缺失导致HopQ-CEACAM结合的急剧降低。HopQ黏附素结构域上的第四个扩展环HopQ_371-407(环D)位于HopQ-CEACAM结合界面附近。尽管HopQ_371-407与CEACAM N结构域之间没有直接接触, 但结合HopQ_371-407的抗体或抗体衍生物将通过空间位阻破坏HopQ-CEACAM相互作用。

[0429] 表2. HopQ^{AD}-hC 1ND结构的数据收集和细化统计。

HopQ ^{AD} -hC1ND	
数据收集	
空间群	C2
晶胞尺寸	
<i>a, b, c</i> (Å)	118.0, 174.0, 118.1
α, β, γ (°)	90.0, 118.4, 90.0
分辨率 (Å)	50.00-3.55 (3.64-3.55)*
<i>R</i> _{合并}	11.8 (88.0)*
<i>I</i> / σ <i>I</i>	8.0 (1.6)*
CC1/2	99.5 (59.7)*
完全性 (%)	99.3 (99.5)*
冗余	3.8 (3.8)*
[0430]	
细化	
分辨率 (Å)	48.81-3.55
反射数	25252
<i>R</i> _{起作用} / <i>R</i> _{游离}	28.1 / 33.8
原子数	
蛋白质	9621
水	0
B-因子	
蛋白质	81.5
水	NA
R.m.s偏差	
键长 (Å)	0.008
键角 (°)	1.12

[0431] *最高分辨率显示在括号中。

[0432] †通过应用基于约为0.5的半数据集的平均强度相关系数 (CC1/2) 的截止值来确定分辨率限值。

[0433] 材料和方法

[0434] 细菌和细菌培养条件

[0435] 幽门螺杆菌菌株G27、PMSS1、SS1、J99 (ATCC, 700824)、2808、26695 (ATCC, 70039)、TX30、60190、P12、NCTC11637 (ATCC, 43504)、Ka89和胆汁螺杆菌 (ATCC43879) 在微需氧条件下 (10%CO₂, 5%O₂, 8.5%N₂和37°C) 在Wilkins-Chalgren血琼脂平板上培养。猪螺杆菌和海尔曼螺杆菌在布鲁氏菌琼脂上生长, 并且猫螺杆菌 (ATCC 49179) 和毕氏螺杆菌在补充有10%马血的脑心浸出液 (brain-heart infusion, BHI) 琼脂上培养。将卡他莫拉菌 (ATCC, 25238)、腔隙莫拉菌 (ATCC 17967) 和空肠弯曲杆菌 (ATCC, 33560) 在补充有5%加热马血的脑心浸出液 (BHI) 琼脂上在37°C下在CO₂培养箱中培养过夜。如所述 (Belogolova等, 2013), 通过用氯霉素抗性盒替换整个基因来完成同基因 Δ hopQ突变体的产生。

[0436] CEACAM蛋白的产生

[0437] 如所述 (Singer等, 2014), 使分别编码以下的胞外结构域的cDNA与人重链Fc结构域融合并克隆到pcDNA3.1 (+) 表达载体 (Invitrogen, San Diego, CA) 中, 测序并稳定转染到HEK293 (ATCC CRL-1573) 细胞中: 人CEACAM1-Fc (由N-A1-B1-A2结构域组成)、人CEACAM1dN-Fc (由A1-B1-A2组成)、大鼠CEACAM1-Fc (由N-A1-B1-A2组成)、大鼠CEACAM1dN-Fc (由A1-B1-A2组成)、人CEACAM3-Fc (由N组成)、人CEACAM6-Fc (由N-A-B组成)、人CEACAM8-Fc (由N-A-B组成)。Fc嵌合CEACAM-Fc蛋白在无血清Pro293s-CDM培养基 (Lonza) 中积累, 并通过蛋白A/

G-琼脂糖凝胶亲和色谱法 (Pierce) 回收。通过 SDS-PAGE 分析蛋白质并用考马斯蓝染色, 表明产生的融合蛋白的等量和完整性 (图2i)。从 Sino Biological Inc. 订购重组人 CEACAM5-Fc。为了产生重组人 CEACAM1 N-结构域 (C1ND), 使用 GeneOptimizer® (Life Technologies) 对注释结构域 (UniProt ID:P13688) 进行第一反向翻译, 并添加 Igk 链的前导序列以及 C 端 Strep-标签 II。合成该基因并将其无缝克隆到 pCDNA3.4-TOPO (Life Technologies) 中。根据 Expi293 表达系统说明书 (Life Technologies), 在 Expi293 细胞的 2L 培养物中产生蛋白质。将所得上清液浓缩并通过交叉流过滤使用 Hydrosart 5kDa 分子量截留膜 (Sartorius) 对 10 体积的 1×SAC 缓冲液 (100mM Tris, 140mM NaCl, 1mM EDTA, pH 8.0) 渗滤。将渗余物上样至 StrepTrap HP 柱 (GE Healthcare) 上, 并用补充有 2.5mM D-脱硫生物素 (IBA) 的 1×SAC 洗脱。将蛋白质储存在 +4°C。

[0438] 对于 C1ND (Ec-C1ND) 的细菌表达, 氨基酸序列经密码子优化以在大肠杆菌中表达, 通过 GeneArt 从头基因合成 (Life Technologies) 进行合成, 并使用 Gateway 技术 (Invitrogen) 与 pDEST™14 载体中的 C 端 His6 标签一起克隆。用得到的构建体转化大肠杆菌 C43 (DE3) 细胞, 并使其在 37°C 下在摇动下在补充有 100µg/mL 氨苄青霉素的 LB 中培养。在 OD₆₀₀=1 时, 在 30°C 下用 1mM IPTG 诱导 Ec-C1ND 表达过夜。通过在 4°C 下以 6.238g 离心 15 分钟收集细胞, 并将其重悬于 50mM Tris-HCl pH 7.4, 500mM NaCl (4mL/g 湿细胞) 中, 其补充有 5µM 亮抑蛋白酶肽和 1mM AEBSF, 100µg/mL 溶菌酶和 20µg/mL DNase I。然后, 在 4°C 下通过 Constant System Cell Cracker 中的单通道以 20kPsi 裂解细胞, 并通过在 48.400g 下离心 40 分钟除去碎片。将细胞质提取物通过 0.45µm 孔过滤器过滤, 并上样至用缓冲液 A (50mM Tris-HCl pH 7.4, 500mM NaCl 和 20mM 咪唑) 平衡的 5mL 预填充 Ni-NTA 柱 (GE Healthcare) 上。然后用 40 个床体积的缓冲液 A 洗涤柱, 并用线性梯度的 0-75% 缓冲液 B (50mM Tris-HCl pH 7.4, 500mM NaCl 和 500mM 咪唑) 洗脱结合的蛋白质。合并如通过 SDS-PAGE 确定的含有 Ec-C1ND 的级分并在 10kDa MW 截止旋转浓缩器中浓缩至最终体积为 5mL。为了除去少量蛋白质污染物, 将浓缩的样品注射到用含有 50mM Tris-HCl pH 8.0, 150mM NaCl 的缓冲液预平衡的 Hi-Prep™ 26/60 Sephacryl S-100HR 柱 (GE Healthcare) 上。合并含有 Ec-C1ND 复合体的级分, 并使用 10kDa MW 截止旋转浓缩器进行浓缩。

[0439] HopQ^{AD}和HopQ^{AD} Δ ID 克隆、产生和纯化

[0440] 为了获得可溶性 HopQ 片段, 使用来自幽门螺杆菌 G27 菌株 (登记号 CP001173; 区域: 1228696..1230621; SEQ ID NO:1) 的 HopQ 基因, 并产生范围为第 37-463 位残基的 HopQ 片段 (成熟蛋白质的第 17-443 位残基), 从而除去 N 端 β-链和信号肽, 以及预期代表 TM 结构域的 C 端 β-结构域。在 HopQ^{AD} Δ ID 中, 成熟蛋白质的第 190-218 位氨基酸被两个甘氨酸替代 (图 8e)。在 T7 启动子下, 使用包含与 pPRkana-1 (pPR-IBA 1 (IBA GmbH) 的衍生物, 其中氨苄青霉素抗性盒替换为卡那霉素抗性盒) 的侧翼靶载体序列具有 30bp 重叠的引物 (正向:

[0441]

GTTTAACTTTAAGAAGGAGATATACAAATGGCGGTTCAAAAAGTGAAAAACGC (SEQ ID NO:

8)

;

[0442] 反向:

[0443]

TCAAGCITATTAATGATGATGATGATGGTGGGCGCCGTTATTCGTGGTGTG (SEQ ID

NO: 9)

),

[0444] 从幽门螺杆菌G27基因组DNA中PCR扩增对应于HopQ I型片段的DNA编码序列。平行地,使用引物(正向:

[0445] CACCATCATCATCATCATTAATAAGCTTGATCCGGCTGCTAAC (SEQ ID NO:10);反向: GTTTAACTTTAAGAAGGAGATATACAAATG (SEQ ID NO:11)使用相同的重叠序列以相反方向对载体进行PCR扩增。正向引物另外携带6×His标签的序列。使用Gibson Assembly (New England Biolabs GmbH)无缝地克隆扩增子。基于密码子优化的HopQ^{AD}质粒克隆HopQ^{AD} Δ ID构建体。通过用两个甘氨酸替换ID区域的5'磷酸化引物(正向:

[0446] GGTGACGCTCAGAACCTGCTGAC (SEQ ID NO:12);反向:

[0447] ACCACCTTTAGAGTTCAGCGGAG (SEQ ID NO:13))扩增质粒,用DpnI (NEB)消化并通过T4连接酶(NEB)平末端连接。

[0448] 用pPRkana-1构建体转化大肠杆菌BL21 (DE3)细胞(NEB GmbH),使细胞在根据{Studier:2005ku}的补充有2mM MgSO₄、100mg/L卡那霉素-硫酸盐(Carl Roth GmbH+Co.KG)、0.2g/L PPG2000 (Sigma Aldrich)和0.2%w/v乳糖-一水合物(Sigma Aldrich)自身诱导的TB培养基(terrific broth,TRB)上在37°C下以275rpm培养,直至OD达到1-2。然后,将温度降至25°C并自身诱导过夜,第二天早晨达到10-15的最终OD。通过在Sorvall RC-6 Plus离心机(Thermo Fischer)中使用SLA-3000转子在4°C下以6000g离心15分钟收获细胞。在细胞破碎之前,将细胞重悬于10ml补充有0.1mM AEBSF-HCl、150U/g BWW DNase I和5mM MgCl₂/克生物湿重(BWW)的冷NiNTA缓冲液A(500mM NaCl,100mM Tris,25mM咪唑,pH 7.4)中,并用Ultra-Turrax T25 digital (IKA GmbH+Co.KG)分散。在4°C下通过高压匀浆用PANDA2000 (GEA Niro Soavi)以800-1200巴(bar)在3个通道中进行细胞破碎。通过在4°C下在SLA-1500转子中以25000g离心30分钟使细胞裂解物澄清,并通过0.2μm滤器过滤除去剩余颗粒。

[0449] 通过连续镍亲和力和尺寸排阻色谱法纯化HopQ片段。简言之,将澄清的细胞裂解物上样至用缓冲液A预平衡的5ml预填充的Ni-NTA HisTrap FF粗柱(GE Healthcare)上,用10倍柱体积(CV)的缓冲液A洗涤,并用15CV线性梯度至75%的NiNTA缓冲液B(500mM NaCl,100mM Tris,500mM咪唑,pH 7.4)洗脱结合的蛋白质。收集洗脱的峰级分,合并并使用10kDa分子量截止旋转浓缩器浓缩至终浓度8-10mg/ml。随后,将5ml浓缩蛋白质上样到用缓冲液C(5mM Tris,140mM NaCl,pH 7.3)预平衡的HiLoad 16/600Superdex 75pg柱(GE Healthcare)上,并以1ml min⁻¹洗脱。最后,仅合并对应于单体峰的蛋白质并在结晶前在+4°C下储存。为了分析HopQ^{AD}的多聚化状态,在具有24ml床体积的Superdex 200 10/300 GL (GE Healthcare)上进行SEC。用缓冲液C预使柱平衡,并随后注射25μg蛋白质并以0.5ml/min的流量分离。

[0450] 对表示HopQ插入结构域(HopQ-ID)的肽进行HA标记,合成(EKLEAHVTTSKYQQDNQTKTTTSVIDTTNYPYDVPDYA (SEQ ID NO:14,下划线为HA标签))并进行HPLC纯化(Peptide Specialty Laboratories,Heidelberg,Germany)。对于细胞测定,将冻干肽溶解在无菌PBS中至浓度为1mM,并用0.1-0.5kDa分子量截止膜对PBS透析以除去剩余的TFA。将肽溶液储存

在-20℃直至进一步使用。

[0451] 表3.HopQ序列。

蛋白质	幽门螺杆菌菌株	UniProt ID	UniProt参考群集 90%序列同一性 (07/2016)	SEQ ID NO (蛋白质/DNA)	注释
[0452] HopQ I型	G27	-	-	1 / 2	登记号CP001173, 区域: 1228696..1230621
HopQ I型	26695	O25791	173	3 / 4	还称为Omp27和HP1177
HopQ I型	P12	H6A3H4	173	15/16	-
HopQ II型	Tx30a	Q8GDI6	77	5 / 6	-
trxA (参考)	J99	P66929	231	-	-

[0453] 通过ELISA检测HopQ-CEACAM相互作用

[0454] 为了检测CEACAM与HopQ^{AD}之间的相互作用,在4℃下将PBS中的重组C1ND (1μg/ml) 在96孔免疫板 (Nunc MaxiSorb) 上包被过夜。在RT下用SmartBlock (Candor) 将孔封闭2小时。随后,在室温下将HopQ片段以10μg/ml至0.05ng/ml的5倍系列稀释度添加持续2小时。接下来,将α-6xHis-HRP缀合物 (克隆3D5, LifeTechnologies) 以1:5000稀释,并在室温下孵育1小时。为了检测,使用1-StepTMUltra TMB-ELISA底物溶液 (LifeTechnologies) 并用2N H₂SO₄终止酶促反应。用PBS/0.05% Tween20进行孵育步骤之间的洗涤 (3-5×)。

[0455] 等温滴定量热法

[0456] ITC测量用MicroCal iTC200量热计 (Malvern) 进行。将HopQ^{AD} I型 (50μM) 或C1ND (25μM) 加载到量热计的小室中,并分别将CEACAM (50μM或500μM) 或HopQ^{AD}I型 (250μM) 加载到注射器中。所有测量均在25℃下进行,其中搅拌速度为600rpm,并在20mM HEPES缓冲液 (pH 7.4)、150mM NaCl、5% (v/v) 甘油和0.05% (v/v) Tween-20中进行。使用MicroCal LLC ITC200软件分析结合数据。

[0457] 用于免疫印迹的SDS-PAGE和非变性PAGE

[0458] 在冰冷的25mM Tris、250mM甘氨酸缓冲液中用SDS-PAGE和非变性PAGE (分别在15%和7.5%聚丙烯酰胺凝胶上) 分离CEACAM。随后,通过在冰冷的转移缓冲液 (25mM Tris、250mM甘氨酸和20%甲醇) 中在25V下湿印迹60分钟将样品转移至PVDF膜。在10%乳粉 (MP)、1×PBS和0.005%吐温-20中将膜封闭1小时。将这两种膜洗涤并在2μM HopQ^{AD} I型的存在下在5%MP、1×PBS、0.005% Tween-20中一起孵育1小时,以允许HopQ^{AD} I与CEACAM之间形成复合体。洗涤步骤后,通过分别在1小时和30分钟期间在5%MP、1×PBS、0.005% Tween-20中连续添加小鼠α-His (AbD Serotec) 和山羊α-小鼠抗体 (Sigma-Aldrich) 检测HopQ的C端His标签 (CEACAM是strep标记的)。洗涤步骤后,通过在显影缓冲液 (10mM Tris-HCl pH 9.5, 100mM NaCl, 50mM MgCl₂) 中添加BCIP/NBT底物 (5-溴-4-氯-3-吲哚基-磷酸/硝基蓝四唑) (Roche) 进行显影。

[0459] 细菌牵出

[0460] 细菌在WCdent琼脂平板上培养过夜。从平板上刮下细菌，悬浮在PBS中，并根据标准曲线通过光密度600读数评估菌落形成单位(colony forming unit, cfu)。用PBS洗涤细菌两次，并在37℃下将 2×10^8 个细胞/ml与可溶性CEACAM-Fc或CEACAM-GFP蛋白或CHO细胞裂解物在头对头旋转(head-over-head rotation)下一起孵育1小时。孵育后，用PBS洗涤细菌5次，并在SDS样品缓冲液(62.5mM Tris-HCl [pH 6.8], 2% w/v SDS, 10% 甘油, 50mM DTT和0.01% w/v溴酚蓝)中煮沸，随后进行SDS-PAGE和Western印迹或溶于FACS缓冲液(PBS/0.5% BSA)中用于流式细胞术分析。

[0461] 免疫沉淀和质谱

[0462] 通过在冰上超声处理($10 \times$, 20s)裂解含有蛋白酶和磷酸酶抑制剂(Roche)的冷PBS中的细菌(2×10^8 个)。通过在4℃下以15,000rpm离心30分钟，然后用预洗涤的蛋白G-琼脂糖(Roche Diagnostics)预清除，以从裂解物中除去细胞碎片。将CEACAM1-Fc添加至裂解物(10 μ g)并在4℃下孵育1小时。将预洗涤的蛋白质G-琼脂糖(60 μ L)添加至抗体和裂解物混合物，并在4℃下孵育2小时。用PBS洗涤珠五次以除去非特异性结合的蛋白质。分离出三分之二的珠并用于质谱分析样品制备。除去上清液，并将珠在50 μ l溶解于20mM Hepes (pH 7.5)中的7M尿素/2M硫脲中重悬两次用于蛋白质的变性。通过离心将珠沉淀，合并上清液并转移至新的Eppendorf管。随后，将蛋白质在1mM DTT中还原45分钟，并在5.5mM碘乙酰胺的终浓度下在黑暗中烷基化30分钟。通过将DTT浓度升高至5mM使烷基化步骤淬灭持续30分钟。所有孵育步骤均在RT下在剧烈摇动(Eppendorf振荡器, 450rpm)下进行。为了消化蛋白质，添加1 μ l LysC (0.5 μ g/ μ l)并将样品在RT下孵育4小时。为了降低尿素浓度，将样品用50mM三乙基碳酸氢铵1:4稀释，并随后与1.5 μ l胰蛋白酶(0.5 μ g/ μ l)一起在37℃下孵育过夜。最后通过用甲酸酸化使胰蛋白酶失活。将上清液转移至新的Eppendorf管，并与含有0.1%甲酸的珠的以下洗涤级分合并。用甲酸(100% v/v)将样品调节至pH 3，并用SepPak C18柱(50mg, Waters)进行肽脱盐。简言之，随后用1ml 100%乙腈和500 μ l 80%乙腈、0.5%甲酸洗涤柱。用1ml 0.1% TFA使柱平衡，将样品上样并用1ml 0.1% TFA再次洗涤柱。在用500 μ l甲酸进行另外的洗涤步骤后，用250 μ l 80%乙腈、0.5%甲酸将肽洗脱两次。然后通过真空离心除去有机相，并将肽储存在-80℃。临测量之前，将肽溶于20 μ l 0.1%甲酸中，超声处理5分钟(水浴)，并随后用预洗涤和平衡的滤器(0.45 μ m低蛋白结合滤器, VWR International, LLC)过滤样品。测量用LC-MS系统进行，该系统由直接连接至Orbitrap XL仪器(Thermo Scientific)的Ultimate 3000纳米HPLC组成。将样品上样到捕获柱(2 μ m, 100A, 2cm长)上并在5至30%乙腈、0.1%甲酸的150分钟梯度期间在15cm C18柱(2 μ m, 100A, Thermo Scientific)上分离。在具有分辨率为60,000的轨道阱(orbitrap)中在m/z 400下获得测量谱(survey spectra)。对于蛋白质鉴定，通过碰撞诱导的解离依次分离多至五个最强烈的全扫描离子并将其片段化。用Proteome Discoverer Software version 1.4 (Thermo Scientific)分析接收的数据，并在SEQUEST算法中针对幽门螺杆菌(菌株G27)数据库(1501种蛋白质)进行检索。添加蛋白质N端乙酰化和甲硫氨酸氧化作为可变修饰，半胱氨酸上的脲基甲基化作为静态修饰。将酶特异性设定为胰蛋白酶，并且前体和碎片离子的质量耐受性分别设定为10ppm和0.8Da。对于电荷状态+1、+2、+3和+4，仅考虑分别满足1.5、2.0、2.25和2.5的 X_{corr} 值的肽用于数据分析。

[0463] 细胞、细胞-细菌共培养和延伸表型定量测定

[0464] 胃癌细胞系MKN45、KatoIII (ATCC, HTB-103)、MKN28和AGS (ATCC, CRL-1739) 从ATCC和DSMZ获得, 通过利用短串联重复 (Short Tandem Repeat, STR) 分析鉴定, 在含有2mM L-谷氨酰胺 (GIBCO, Invitrogen, CA, USA) 的补充有10%FBS (GIBCO, Invitrogen, CA, USA) 和1%青霉素/链霉素 (GIBCO, Invitrogen, CA, USA) 的DMEM (GIBCO, Invitrogen, Carlsbad CA, USA) 中培养至稀疏或紧密汇合。将所有细胞系保持在37°C, 5%CO₂和100%湿度的培养箱中, 并且常规通过DAPI染色和PCR每年两次进行支原体测试。将平板培养的细菌悬浮于DMEM中, 并在微量离心机中以150g离心5分钟进行洗涤。在重悬于DMEM中后, 确定600nm处的光密度, 并将细菌以不同的细菌/细胞比例 (MOI) 在37°C下添加至过夜血清剥夺的细胞以开始感染。在指定时间后, 用PBS洗涤细胞两次, 并随后用蛋白酶和磷酸酶抑制剂PBS中的1%NP-40裂解。选择HEK293细胞用于CEACAM转染研究, 因为发现该细胞对huCEACAM表达是阴性的, 并且易于转染。HEK细胞在含有补充有25mM HEPES缓冲液和10%热灭活FBS (Biochrom, Berlin, Germany) 的RPMI 1640培养基 (Invitrogen) 的6孔板中培养2天至约70%汇合。将细胞血清剥夺过夜, 并对于每个图中指定的时间点以MOI 50用幽门螺杆菌感染。感染后, 在含有1mmol/L Na₃VO₄ (Sigma-Aldrich) 的冰冷PBS中收获细胞。使用Olympus IX50相差显微镜在5个不同的0.25-mm²视野中对每个实验中伸长的AGS细胞进行定量。

[0465] 转染

[0466] 根据制造商的方案 (Invitrogen) 使用lipofectamine 2000程序, 通过用4μg pcDNA3.1-huCEACAM1-4L、pcDNA3.1-huCEACAM1-4S、pcDNA3.1-msCEACAM1-L、pcDNA3.1-ratCEACAM1-L质粒 (Singer) 分别稳定转染细胞, 产生永久表达hu-CEACAM1-4L、小鼠-CEACAM1-L和大鼠-CEACAM1-L的CHO细胞系 (ATCC)。在含有1mg/ml Geneticinsulfat (G418, Biochrom, Berlin, Germany) 的培养基中选择稳定转染的细胞。通过流式细胞术 (FACScalibur, BD) 确定在log期生长的单个克隆中CEACAM1的表面表达。根据制造商的说明书, 用TurboFect试剂 (Fermentas, Germany), 用4μg HA-标记的CEACAM构建体或萤光素酶报道构建体 (Clontech, Germany) 转染HEK293细胞持续48小时。

[0467] Western印迹

[0468] 将等体积的细胞裂解物上样至8%SDS-PAGE凝胶上, 电泳后, 将分离的蛋白质转移至硝酸纤维素膜 (Whatman/GE Healthcare, Freiburg, Germany)。将膜在5%脱脂乳中在室温下封闭1小时, 并用与以下结合的一抗mAb 18/20孵育过夜: CEACAM1、3、5、B3-17和C5-1X (对于hu-CEACAM1具有单特异性, Singer), 4/3/17 (与CEACAM1、5结合, Genovac), 以及5C8C4 (对于hu-CEACAM5具有单特异性, Singer), 1H7-4B (对于hu-CEACAM6具有单特异性, Singer), 6/40c (对于hu-CEACAM8具有单特异性, Singer), Be9.2 (α-大鼠-CEACAM1), mAb 11-1H (α-大鼠-CEACAM1 ΔN, Singer), 磷酸酪氨酸抗体PY-99 (Santa Cruz, LaJolla, CA, USA), α-CagA磷酸酪氨酸抗体PY-972, 小鼠单克隆α-CagA抗体 (Austral Biologicals, San Ramon, CA, USA), 小鼠单克隆α-CEACAM1 (克隆D14HD11 Genovac/Aldevron, Freiburg, Germany) 或山羊α-GAPDH (Santa Cruz)。洗涤后, 将膜与二抗 [HRP缀合的α-小鼠IgG (Promega)] 一起孵育, 并通过ECL Western印迹检测试剂检测蛋白质。通过LabImage 1D软件 (INTAS) 进行定量。

[0469] 流式细胞术

[0470] 将Fc标记的CEACAM (2.5 μ g/ml) 与幽门螺杆菌 (OD:1) 一起孵育,并随后与FITC缀合的山羊 α -人IgG (Sigma) 一起孵育。用FACS缓冲液洗涤后,通过对细菌进行设门(基于前向和侧向散射)并测量细菌相关荧光对样品进行分析。在每种情况下,每个样品获得10,000个事件。用FACS CyAn (Beckman Coulter) 进行分析,并用FlowJo软件 (Treestar) 评价数据。为了分析CEACAM介导的HopQ结合,将指定的细胞类型 (50 μ l中的 5×10^5 个) 与20 μ g/ml的幽门螺杆菌菌株P12来源的myc和6 \times His标记的在3%FCS/PBS中稀释的重组HopQ一起在冰上孵育1小时。用3%FCS/PBS洗涤三次后,用20 μ g/ml小鼠 α -c-myc mAb (克隆9E10, AbD Serotec) 标记样品,并随后用FITC缀合的山羊 α -小鼠F(ab')₂标记 (Dianova, Germany)。平行地,通过使用兔抗CEA pAb (A0115, Dianova) 对细胞进行染色,然后用FITC缀合的山羊 α -兔F(ab')₂ (Dianova, Germany) 对细胞进行染色来控制CEACAM的存在。使用同种型匹配的Ig mAb测定背景荧光。在FACS Calibur流式细胞仪 (BD Biosciences, San Diego, CA) 中检查染色的细胞样品,并使用CellQuest软件分析数据。通过PI染色鉴定的死细胞被排除在测量之外。

[0471] 免疫组化

[0472] 在当地伦理委员会批准后,从Institut für Pathologie, Klinikum Bayreuth, Germany的组织库中随机选择石蜡包埋的人正常胃、胃炎和癌症样品。如果组织质量差,则排除组织学样品。在压力锅 (pressure cooker) 中用10mmol/L柠檬酸钠缓冲液 (pH 6) 进行抗原修复后,将切片与 α -hu-CEACAM1、5、6和 α -大鼠-CEACAM1抗体 (分别为克隆B3-17、5C8C4、1H7-4B和Be9.2) 一起孵育。按照制造商的说明使用SignalStain DAB (Cell Signaling) 对切片进行显影。切片用苏木精 (Morphisto) 复染。使用Olympus Virtual Slide System VS120 (Olympus, Hamburg, Germany) 进行自动图像采集。

[0473] 黏附测定

[0474] 根据Hytonen等,2006进行黏附测定。简言之,使人胃上皮细胞 (MKN45和AGS) 和CEACAM1转染的CHO细胞在5%CO₂气氛中在组织培养96孔板 (Bioscience) 上的补充有5% FCS和1-谷氨酰胺 (2mmol, Sigma, St. Louis, USA) 的无抗生素DMEM (Gibco, Gaithersburg, MD) 中培养2天。为了在黏附测定中使幽门螺杆菌细胞可视化,用CFDA-SE (Molecular Probes) 对OD:1的细菌进行荧光标记并用PBS洗涤。在37 $^{\circ}$ C下在黑暗中恒定旋转下将CFDA-SE以10 μ mol/L的浓度添加30分钟。通过3次PBS洗涤除去过量染料。将细菌重悬于PBS中直至进一步使用。将标记的细菌与细胞在37 $^{\circ}$ C下共孵育 (MOI 10), 同时温和搅拌1小时。用PBS (1ml, \times 3) 洗涤以除去非黏附细菌后,将细胞在多聚甲醛中固定 (2%, 10分钟)。通过使用流式细胞术分析测量结合荧光标记细菌的细胞的百分比来确定细菌结合。

[0475] IL-8细胞因子ELISA

[0476] 如上所述用幽门螺杆菌感染AGS细胞系,并将PBS孵育的对照细胞用作阴性对照。收集培养物上清液并储存在-20 $^{\circ}$ C下直至测定。根据所述程序,通过标准ELISA用市售的测定试剂盒 (Becton Dickinson, Germany) 确定上清液中的IL-8浓度。

[0477] CagA毒力途径的HopQ依赖性

[0478] 如果没有另外说明,则将AGS细胞系 (ATCC CRL-1730) 用多种幽门螺杆菌菌株以感染复数 (multiplicity of infection, MOI) 50感染6小时。然后在1mmol/L Na₃VO₄ (Sigma-Aldrich) 的存在下在冰冷的PBS中收获细胞。在每个实验中,使用相差显微镜 (Olympus IX50) 在10个不同的0.25-mm²视野中对伸长的AGS细胞的数量进行定量。使用检测Tyr-磷酸

化CagA的指定抗体确定CagA易位。所有实验一式三份进行。对于抑制实验,在感染前将细胞与指定的抗体或肽一起孵育。

[0479] 共聚焦显微术

[0480] 使CHO细胞在室玻片(chamber slide)(Thermo Scientific)上培养,在多聚甲醛中固定(4%,10分钟),并用PBS/5%牛血清白蛋白封闭。CFDA-SE标记的细菌(10 μ mol/L,在37 $^{\circ}$ C下在黑暗中恒定旋转30分钟)以MOI 5与细胞在37 $^{\circ}$ C恒定旋转下孵育1小时。5X PBS洗涤后,细胞膜用Deep Red(Life Technology)染色,并且细胞核用DAPI(Life Technology)染色。使用Leica SP5共聚焦显微镜拍摄细胞的共聚焦图像。

[0481] HopQ^{AD}和HopQ^{AD}与C1ND的复合体的结晶和结构确定

[0482] 将HopQ^{AD}浓缩至40mg/mL并通过在20 $^{\circ}$ C下使用0.12M醇(0.02M 1,6-己二醇;0.02M 1-丁醇;0.02M 1,2-丙二醇;0.02M 2-丙醇;0.02M 1,4-丁二醇;0.02M 1,3-丙二醇)、0.1M Tris(碱)/BICINE pH 8.5,20%v/v PEG500*MME进行滴蒸气扩散(drop vapor diffusion)进行结晶;10%w/v PEG20000作为结晶缓冲液。将晶体进行环安装并在液氮中快速冷却。在束线Proximal(SOLEIL,Gif-sur-Yvette,France)下以100K收集数据,并使用XDS包对其进行索引、处理和缩放。所有晶体都在P2₁空间群中,其中近似的晶胞尺寸为**a =57.7 \AA , b =57.7 \AA , c =285.7 \AA** 和 $\beta =90.1$,并且每个不对称单位有4个HopQ^{AD}拷贝。通过使用BabA结构的分子置换获得相,并且通过使用Refmac5的手动重建和最大似然细化的迭代循环来对模型进行细化。表1总结了晶体参数、数据处理和结构细化统计。为了在HopQ^{AD}与人CEACAM1的N-结构域(C1ND)之间形成复合体,相对于经纯化HopQ^{AD},以1.2倍摩尔过量添加纯化的重组C1ND,并将混合物注射到在20mM Tris-HCl pH 8.0,500mM NaCl缓冲液中预平衡的Hi-PrepTM26/60 Sephacryl S-100HR柱(GE Healthcare)上。将含有HopQ^{AD}-C1ND复合体的级分合并在一起,并使用30kDa MW截止旋转浓缩器浓缩至终浓度为30mg/mL。在0.03M氟化钠、0.03M溴化钠、0.03M碘化钠、0.1M MES pH 6.5、20%v/v乙二醇和10%w/v PEG 8000中获得晶体。晶体进行环安装并在液氮中闪蒸冷却,在束线Proxima 1(Soleil,Gif-sur-Yvette,France)下以100K收集数据。晶体位于C2空间群中,其中近似的晶胞尺寸为**a =118.0 \AA , b =174.0 \AA , c =118.1 \AA** , $\beta =118.4$,并且每个不对称单位有三个HopQ^{AD}-C1ND拷贝。通过使用HopQ^{AD}和C1ND(PDB代码4WHD)结构的分子置换获得相,并且通过使用Refmac5的手动重建和最大似然细化的迭代循环来细化模型。表2总结了晶体参数、数据处理和结构细化统计。

[0483] 氨基酸序列比对

[0484] 使用CLC主工作台(CLC bio)进行人-、小鼠-和大鼠-CEACAM1和人CEACAM(1、5、6和8)的N端结构域的氨基酸序列比对。

[0485] 荧光素酶报道测定

[0486] 根据制造商说明书(Promega,USA),用幽门螺杆菌感染用多种荧光素酶报道子和对照构建体(Clontech)转染的CHO-CEACAM1-L细胞5小时,并通过荧光素酶测定使用双荧光素酶报道测定系统进行分析。简言之,通过被动裂解收获细胞,用Precision Red(Cytoskeleton,USA)测量蛋白质浓度,并通过添加被动裂解缓冲液使裂解物均衡。通过使用Plate Luminometer(MITHRAS LB940,来自Berthold,Germany)测量荧光素酶活性。

[0487] 动物实验

[0488] 从Charles River Laboratories,Sulzfeld,Germany获得4周龄的无特定病原体的120-150克雄性Sprague dawley大鼠。由不参与实验的动物护理者将动物随机分配到不同的实验组中,并且预先建立排除动物的标准。对所有结果和数据的评估进行研究者盲法,由独立研究者以盲法方式进行组织学。在8个组中用约 1×10^9 个活幽门螺杆菌在2个间隔天内对动物进行两次胃内攻击。根据伦理委员会和州兽医办公室的准许和指导方针(Regierung von Oberbayern,55.2-1.54-2532-160-12),实验在Zentrum für **Präklinische** Forschung Klinikum r.d.Isar der TU München的无特定病原体单元中进行。

[0489] 统计分析

[0490] 对于体外实验,通过Shapiro-Wilk测试确定正态分布。使用Graph Pad Prism软件,使用双尾t检验(Student t-test)和单因素ANOVA连同事后Bonferroni检验(比较多于两组)对所有数据进行分析。数据显示为平均值 \pm s.e.m或s.d.,至少进行三个独立实验。P值 <0.05 被认为是显著的。对于动物研究,在使用尽可能低的数字以符合实验动物的伦理准则的同时,基于先前的动物实验进行检验力计算(power calculation)以实现0.05的双侧显著性。

[0491] 参考文献

[0492] 1.Apostolopoulos,V.et al.,2013.Targeting antigens to dendritic cell receptors for vaccine development.Journal of Drug Delivery,2013:869718.

[0493] 2.Belogolova.E.et al.,2013.Helicobacter pylori outer membrane protein HopQ identified as a novel T4SS-associated virulence factor.Cell Microbiol.15.pp.1896-1912.

[0494] 3.Blaser,M.J.et al.,1995.Infection with Helicobacter pylori strains possessing cagA is associated with an increased risk of developing adenocarcinoma of the stomach.Cancer Research.55(10),pp.2111-2115.

[0495] 4.Cao,P.&Cover,T.L.,2002.Two different families of hopQ alleles in Helicobacter pylori.Journal of Clinical Microbiology.40,pp.4504-4511.

[0496] 5.Forman,D.,1996.Helicobacter pylori and gastric cancer.Scandinavian Journal of Gastroenterology.Supplement,214,pp.31-3-discussion 40-3.

[0497] 6.Fox,J.G.,2002.The non-Helicobacter pylori helicobacters:their expanding role in gastrointestinal and systemic diseases.Gut,50.pp.273-283.

[0498] 7.Fox,J.G.et al.,1998.Hepatic Helicobacter species identified in bile and gallbladder tissue from Chileans with chronic cholecystitis.Gastroenterology,114,pp.755-763.

[0499] 8.Gao,W.et al.,2010.The evolution of Helicobacter pylori antibiotics resistance over 10 years in Beijing,China.Helicobacter,15(5),pp.460-466.

[0500] 9.Gómez-Gaseón,L.et al.,2012.Exploring the pan-spectrum of Streptococcus suis:looking for common protein antigens.Journal of Proteomics,75(18),pp.5654-5666.

- [0501] 10.Graham,D.Y.&Shiotani,A.,2005.The time to eradicate gastric cancer is now.*Gut*,54(6),pp.735-738.
- [0502] 11.Ilylonen.J.et al.,2006.Use of flow cytometry for the adhesion analysis of *Streptococcus pyogenes* mutant strains to epithelial cells: investigation of the possible role of surface pullulanase and cysteine protease,and the transcriptional regulator Rgg.*BMC Microbiol.*,6,18,doi: 10.1186/1471-2180-6-18.
- [0503] 12.Jemal,A.et al.,2011.Global cancer statistics.*CA:a cancer journal for clinicians*.61(2),pp.69-90.
- [0504] 13.Kalali,B.et al.,2014.*H. pylori* virulence factors:influence on immune system and pathology.*Mediators of Inflammation*,2014:426309.
- [0505] 14.Koebnik.R.et al.,2000.Structure and function of bacterial outer membrane proteins:barrels in a nutshell.*Molecular Microbiology*.37(2),pp.239-253.
- [0506] 15.Matsukura.N.et al.,2002.Association between *Helicobacter bilis* in bile and biliary tract malignancies:*H. bilis* in bile from Japanese and Thai patients with benign and malignant diseases in the biliary tract.*Jpn J Cancer Res.*,93(7),pp.842-7.
- [0507] 16.Mori,J.et al.,2012.Chimeric flagellin as the self-adjucanting antigen for the activation of immune response against *Helicobacter pylori*,*Vaccine*.30(40),pp.5856-5863.
- [0508] 17.Nomura.A.et al.,1994.*Helicobacter pylori* infection and the risk for duodenal and gastric ulceration.*Annals of Internal Medicine*,120(12), pp.977-981.
- [0509] 18.Parsonnet,J.et al.,1991.*Helicobacter pylori* infection and the risk of gastric carcinoma.*New England Journal of Medicine*,325(16),pp.1127-1131.
- [0510] 19.Perez-Perez,G.I.,et al.,2004.Epidemiology of *Helicobacter pylori* infection.*Helicobacter*,9 Suppl 1.pp.1-6.
- [0511] 20.Pisani,P.et al.,2008.Cross-Reactivity between Immune Responses to *Helicobacter bilis* and *Helicobacter pylori* in a Population in Thailand at High Risk of Developing Cholangiocarcinoma.*Clin Vaccine Immunol.*,15(9) .pp.1363-1368.
- [0512] 21.Shiola,S.et al.,2010,Population-based strategies for *Helicobacter pylori*-associated disease management:a Japanese perspective.*Expert Review of Gastroenterology&Hepatology*,4(2),pp.149-156.
- [0513] 22.Singer.B.B.et al.,2014.Soluble CEACAM8 interacts with CEACAM1 inhibiting TLR2-triggered immune responses.*PLoS One*.9,e94106.
- [0514] 23.Sioud,M.et al.,2013.A novel peptide carrier for efficient targeting of antigens and nucleic acids to dendritic cells.*FASLB J.*,27(8),

pp.3272-3283.

[0515] 24.Song,II.et al.,2015.A novel chimeric flagellum fused w ith the multi-epitope vaccine CTB-UE prevents Helicobacler pyloti-inlueed gastrie cancer in a BAI.B/c monse mod • 1.Appl Microbiol Bioteehnol.,99 (22) .pp.9495-9502.

[0516] 25.Tchoupa,A.K.et al.,2014.Signaling by epithelial members of the CEACAM family-mucosal docking sites for pathogenic bacteria.Cc11 Commun Signal.12:27.

[0517] 26.United States Centers for Discase Control and Prevention(2011) .“A CDC framework for preventing infectious diseascs”,accessed 20.12.2012.

序列表

<110> 慕尼黑工业大学等
 <120> 用于预防或治疗幽门螺杆菌感染的药剂和方法
 <130> 914-8 PCT
 <150> EP 16 180 430.7
 <151> 2016年7月20日
 <160> 33
 <170> PatentIn version 3.5
 <210> 1
 <211> 641
 <212> PRT
 <213> 幽门螺杆菌
 <400> 1
 Met Lys Lys Thr Lys Lys Thr Ile Leu Leu Ser Leu Thr Leu Ala Ala
 1 5 10 15
 Ser Leu Leu His Ala Glu Asp Asn Gly Val Phe Leu Ser Val Gly Tyr
 20 25 30
 Gln Ile Gly Glu Ala Val Gln Lys Val Lys Asn Ala Asp Lys Val Gln
 35 40 45
 Lys Leu Ser Asp Thr Tyr Glu Gln Leu Ser Arg Leu Leu Thr Asn Asp
 50 55 60
 [0001] Asn Gly Thr Asn Ser Lys Thr Ser Ala Gln Ala Ile Asn Gln Ala Val
 65 70 75 80
 Asn Asn Leu Asn Glu Arg Ala Lys Thr Leu Ala Gly Gly Thr Thr Asn
 85 90 95
 Ser Pro Ala Tyr Gln Ala Thr Leu Leu Ala Leu Arg Ser Val Leu Gly
 100 105 110
 Leu Trp Asn Ser Met Gly Tyr Ala Val Ile Cys Gly Gly Tyr Thr Lys
 115 120 125
 Ser Pro Gly Glu Asn Asn Gln Lys Asp Phe His Tyr Thr Asp Glu Asn
 130 135 140
 Gly Asn Gly Thr Thr Ile Asn Cys Gly Gly Ser Thr Asn Ser Asn Gly
 145 150 155 160
 Thr His Ser Tyr Asn Gly Thr Asn Thr Leu Lys Ala Asp Lys Asn Val
 165 170 175
 Ser Leu Ser Ile Glu Gln Tyr Glu Lys Ile His Glu Ala Tyr Gln Ile
 180 185 190
 Leu Ser Lys Ala Leu Lys Gln Ala Gly Leu Ala Pro Leu Asn Ser Lys
 195 200 205
 Gly Glu Lys Leu Glu Ala His Val Thr Thr Ser Lys Tyr Gln Gln Asp
 210 215 220

Asn Gln Thr Lys Thr Thr Thr Ser Val Ile Asp Thr Thr Asn Asp Ala
 225 230 235 240

Gln Asn Leu Leu Thr Gln Ala Gln Thr Ile Val Asn Thr Leu Lys Asp
 245 250 255

Tyr Cys Pro Ile Leu Ile Ala Lys Ser Ser Ser Ser Asn Gly Gly Thr
 260 265 270

Asn Asn Ala Asn Thr Pro Ser Trp Gln Thr Ala Gly Gly Gly Lys Asn
 275 280 285

Ser Cys Ala Thr Phe Gly Ala Glu Phe Ser Ala Ala Ser Asp Met Ile
 290 295 300

Asn Asn Ala Gln Lys Ile Val Gln Glu Thr Gln Gln Leu Ser Ala Asn
 305 310 315 320

Gln Pro Lys Asn Ile Thr Gln Pro His Asn Leu Asn Leu Asn Ser Pro
 325 330 335

Ser Ser Leu Thr Ala Leu Ala Gln Lys Met Leu Lys Asn Ala Gln Ser
 340 345 350

Gln Ala Glu Ile Leu Lys Leu Ala Asn Gln Val Glu Ser Asp Phe Asn
 355 360 365

[0002] Lys Leu Ser Ser Gly His Leu Lys Asp Tyr Ile Gly Lys Cys Asp Ala
 370 375 380

Ser Ala Ile Ser Ser Ala Asn Met Thr Met Gln Asn Gln Lys Asn Asn
 385 390 395 400

Trp Gly Asn Gly Cys Ala Gly Val Glu Glu Thr Gln Ser Leu Leu Lys
 405 410 415

Thr Ser Ala Ala Asp Phe Asn Asn Gln Thr Pro Gln Ile Asn Gln Ala
 420 425 430

Gln Asn Leu Ala Asn Thr Leu Ile Gln Glu Leu Gly Asn Asn Pro Phe
 435 440 445

Arg Asn Met Gly Met Ile Ala Ser Ser Thr Thr Asn Asn Gly Ala Leu
 450 455 460

Asn Gly Leu Gly Val Gln Val Gly Tyr Lys Gln Phe Phe Gly Glu Lys
 465 470 475 480

Lys Arg Trp Gly Leu Arg Tyr Tyr Gly Phe Phe Asp Tyr Asn His Ala
 485 490 495

Tyr Ile Lys Ser Asn Phe Phe Asn Ser Ala Ser Asp Val Trp Thr Tyr
 500 505 510

Gly Val Gly Ser Asp Leu Leu Phe Asn Phe Ile Asn Asp Lys Asn Thr
 515 520 525

Asn Phe Leu Gly Lys Asn Asn Lys Ile Ser Phe Gly Leu Phe Gly Gly
 530 535 540

Ile Ala Leu Ala Gly Thr Ser Trp Leu Asn Ser Gln Phe Val Asn Leu
 545 550 555 560

Lys Thr Ile Ser Asn Val Tyr Ser Ala Lys Val Asn Thr Ala Asn Phe
 565 570 575

Gln Phe Leu Phe Asn Leu Gly Leu Arg Thr Asn Leu Ala Arg Pro Lys
 580 585 590

Lys Lys Asp Ser His His Ala Ala Gln His Gly Met Glu Leu Gly Val
 595 600 605

Lys Ile Pro Thr Ile Asn Thr Asn Tyr Tyr Ser Phe Leu Asp Thr Lys
 610 615 620

Leu Glu Tyr Arg Arg Leu Tyr Ser Val Tyr Leu Asn Tyr Val Phe Ala
 625 630 635 640

Tyr

<210> 2
 <211> 1926
 <212> DNA
 <213> 幽门螺杆菌

[0003]

<400> 2
 atgaaaaaaaa cgaaaaaaac aattttgctt tctctaactc ttgcgccate actcttgcac 60
 gctgaagaca acggcgcttt ttaagcgtg ggttatcaaa tcggtgaage ggttcaaaaa 120
 gtgaaaaacg ccgacaaggt gcaaaaactt tcagacactt atgaacaatt aagccggctt 180
 ttaaccaacg ataatggcac aaactcaaag acaagcgcgc aagcgatcaa ccaagcggtt 240
 aataatttga acgaacgcgc aaaaacttta gccggtgga caaccaattc cctgcctat 300
 caagccacgc ttttagcgtt gagatcgggtg ttaggctat ggaatagcat gggttatgca 360
 gtcatatgcg gaggtatata caaaagtcca ggcgaaaaca atcaaaaaga tttccactac 420
 accgatgaga atggcaacgg cactacaatc aattgcggtg ggagcacaaa tagtaatggc 480
 actcatagtgt ataatggcac aaatacatta aaagcagaca aaaatgttct tctatctatt 540
 gagcaatattg aaaaaatcca tgaagcttat cagattcttt caaaagcttt aaaacaagct 600
 ggcttgcctc cttaaatag caaaggggaa aaattagaag cgcattgtaac cacatcaaag 660
 tatcaacaag ataatcaaac taaaacgaca acttctgtta ttgatacgac taatgatgca 720
 caaaatcttt tgactcaagc gaaacgatt gtcaatacc ttaaagatta ttgccccata 780
 ttgatagcga aatcttctag tagtaatggc ggaactaata acgcaaacac cccttcatgg 840
 caaacagccg gtggcgcaaa aaattcatgc gcgacttttg gcgaggagtt tagtgccgct 900
 tcagacatga ttaataatgc gcaaaaaate gttcaagaaa cccaacaact cagcgccaac 960
 caacaaaaaa atatcacaca accccataat ctcaacctta actctcctag cagtcttacg 1020
 gcttttagctc aaaaaatgct caaaaacgcg caatctcaag cagaaatctt aaaactagcc 1080

aatcaagtgg agagcgattt taacaaactt tcttcaggcc atcttaaaga ctacataggg 1140
 aaatgcgatg cgagcgctat aagcagtgcg aatatgacaa tgcaaaatca aaagaacaat 1200
 tgggggaacg ggtgtgctgg cgtggaagaa actcagtcct tgtaaanaac aagcgccgct 1260
 gattttaaca accaaaagcc tcaaatcaat caagcacaaa acctagccaa cacccttatt 1320
 caagaacttg gcaacaacce ttttaggaat atgggcatga tcgcttcttc aaccacgaat 1380
 aacggcgcc tgaatggtct tggggtgcaa gtgggttata agcaattttt tggagaaaag 1440
 aaaagatggg ggtaagga ttatggttc tttgattaca accacgccta catcaaatcc 1500
 aatttcttta actcggcttc tgatggtgg acttatgggg tgggtagcga tttgtgttt 1560
 aatttcatca acgataaaaa caccaacttt ttaggcaaga ataacaagat ttcttttggg 1620
 ctttttgag gcatcgcttc agcagggact tcatggctta atttcaatt cgtgaattta 1680
 aaaaccatca gcaatgtcta tagcgctaaa gtgaatacgg ctaatttcca atttttatc 1740
 aatttgggct tgagaaccaa tctcgetagg cctaagaaaa aagatagcca tcatgcagct 1800
 caacatggca tggattggg cgtgaaaate cctaccatta acacgaatta ctattctttt 1860
 ctagacacca aactagaata tagaaggctt tatagcgtgt atctcaatta tgtgttcgct 1920
 tattaa 1926

<210> 3
 <211> 641
 <212> PRT
 <213> 幽门螺杆菌

<400> 3

[0004]

Met Lys Lys Thr Lys Lys Thr Ile Leu Leu Ser Leu Thr Leu Ala Ala
 1 5 10 15
 Ser Leu Leu His Ala Glu Asp Asn Gly Val Phe Leu Ser Val Gly Tyr
 20 25 30
 Gln Ile Gly Glu Ala Val Gln Lys Val Lys Asn Ala Asp Lys Val Gln
 35 40 45
 Lys Leu Ser Asp Thr Tyr Glu Gln Leu Ser Arg Leu Leu Thr Asn Asp
 50 55 60
 Asn Gly Thr Asn Ser Lys Thr Ser Ala Gln Ala Ile Asn Gln Ala Val
 65 70 75 80
 Asn Asn Leu Asn Glu Arg Ala Lys Thr Leu Ala Gly Gly Thr Thr Asn
 85 90 95
 Ser Pro Ala Tyr Gln Ala Thr Leu Leu Ala Leu Arg Ser Val Leu Gly
 100 105 110
 Leu Trp Asn Ser Met Gly Tyr Ala Val Ile Cys Gly Gly Tyr Thr Lys
 115 120 125
 Ser Pro Gly Glu Asn Asn Gln Lys Asp Phe His Tyr Thr Asp Glu Asn
 130 135 140
 Gly Asn Gly Thr Thr Ile Asn Cys Gly Gly Ser Thr Asn Ser Asn Gly
 145 150 155 160

Thr His Ser Ser Ser Gly Thr Asn Thr Leu Lys Ala Asp Lys Asn Val
 165 170 175

Ser Leu Ser Ile Glu Gln Tyr Glu Lys Ile His Glu Ala Tyr Gln Ile
 180 185 190

Leu Ser Lys Ala Leu Lys Gln Ala Gly Leu Ala Pro Leu Asn Ser Lys
 195 200 205

Gly Glu Lys Leu Glu Ala His Val Thr Thr Ser Lys Pro Glu Asn Asn
 210 215 220

Ser Gln Thr Lys Thr Thr Thr Ser Val Ile Asp Thr Thr Asn Asp Ala
 225 230 235 240

Gln Asn Leu Leu Thr Gln Ala Gln Thr Ile Val Asn Thr Leu Lys Asp
 245 250 255

Tyr Cys Pro Met Leu Ile Ala Lys Ser Ser Ser Glu Ser Ser Gly Ala
 260 265 270

Ala Thr Thr Asn Ala Pro Ser Trp Gln Thr Ala Gly Gly Gly Lys Asn
 275 280 285

Ser Cys Ala Thr Phe Gly Ala Glu Phe Ser Ala Ala Ser Asp Met Ile
 290 295 300

[0005] Asn Asn Ala Gln Lys Ile Val Gln Glu Thr Gln Gln Leu Ser Ala Asn
 305 310 315 320

Gln Pro Lys Asn Ile Thr Gln Pro His Asn Leu Asn Leu Asn Thr Pro
 325 330 335

Ser Ser Leu Thr Ala Leu Ala Gln Lys Met Leu Lys Asn Ala Gln Ser
 340 345 350

Gln Ala Glu Ile Leu Lys Leu Ala Asn Gln Val Glu Ser Asp Phe Asn
 355 360 365

Lys Leu Ser Ser Gly His Leu Lys Asp Tyr Ile Gly Lys Cys Asp Ala
 370 375 380

Ser Ala Ile Ser Ser Ala Asn Met Thr Met Gln Asn Gln Lys Asn Asn
 385 390 395 400

Trp Gly Asn Gly Cys Ala Gly Val Glu Glu Thr Leu Ser Ser Leu Lys
 405 410 415

Thr Ser Ala Ala Asp Phe Asn Asn Gln Thr Pro Gln Ile Asn Gln Ala
 420 425 430

Gln Asn Leu Ala Asn Thr Leu Ile Gln Glu Leu Gly Asn Asn Pro Phe
 435 440 445

Arg Asn Met Gly Met Ile Ala Ser Ser Thr Thr Asn Asn Gly Ala Leu
 450 455 460

Asn Gly Leu Gly Val Gln Val Gly Tyr Lys Gln Phe Phe Gly Glu Lys
465 470 475 480

Lys Arg Trp Gly Leu Arg Tyr Tyr Gly Phe Phe Asp Tyr Asn His Ala
485 490 495

Tyr Ile Lys Ser Asn Phe Phe Asn Ser Ala Ser Asp Val Trp Thr Tyr
500 505 510

Gly Val Gly Ser Asp Leu Leu Phe Asn Phe Ile Asn Asp Lys Asn Thr
515 520 525

Asn Phe Leu Gly Lys Asn Asn Lys Ile Ser Val Gly Phe Phe Gly Gly
530 535 540

Ile Ala Leu Ala Gly Thr Ser Trp Leu Asn Ser Gln Phe Val Asn Leu
545 550 555 560

Lys Thr Ile Ser Asn Val Tyr Ser Ala Lys Val Asn Thr Ala Asn Phe
565 570 575

Gln Phe Leu Phe Asn Leu Gly Leu Arg Thr Asn Leu Ala Arg Pro Lys
580 585 590

Lys Lys Asp Ser His His Ala Ala Gln His Gly Met Glu Leu Gly Val
595 600 605

[0006] Lys Ile Pro Thr Ile Asn Thr Asn Tyr Tyr Ser Phe Leu Asp Thr Lys
610 615 620

Leu Glu Tyr Arg Arg Leu Tyr Ser Val Tyr Leu Asn Tyr Val Phe Ala
625 630 635 640

Tyr

<210> 4
<211> 1926
<212> DNA
<213> 幽门螺杆菌

<400> 4
atgaaaaaa cgaaaaaac gattctgctt tccttaactc tcgcggcgtc attgctccat 60
gctgaagaca acggcgtttt ttaagcgtg ggttatcaaa tcggtgaagc ggttcaaaaa 120
gtgaaaaacg cgcacaaggt gcaaaaactt tcagacactt atgaacaatt aagccggtt 180
ttaaccaacg ataatggcac aaactcaaag acaagcgcgc aagcgatcaa ccaagcggtt 240
aataatttga acgaacgcgc aaaaacttta gccggtggga caaccaattc cctgcctat 300
caagccacgc ttttagcgtt gagatcggtg ttagggctat ggaatagcat gggttatgcg 360
gtcatatgcg gaggttatac caaaagtcca ggcgaaaaca atcaaaaaga tticcactac 420
accgatgaga atggcaatgg cactacaatc aattgcggtg ggagcacaaa tagtaatggc 480
actcatagtt ctagtggcac aaatacatta aaagcagaca aaaatgtttc tctatctatt 540
gagcaatgat aaaaaatcca tgaagcttat cagattcttt caaaagcttt aaaacaagcc 600

gggcttgctc ctttaaatag caaaggggaa aagttagaag cgcatgtaac cacatcaaaa 660
 ccagaaaata atagtcaaac taaaacgaca acttctgta ttgatacgac taatgatgcg 720
 caaaatcttt tgactcaagc gcaaacgatt gtcaataccc ttaaagatta ttgcccattg 780
 ttgatagcga aatctagtag taaaagtagt ggcgcagcta ctacaaaacgc cccttcatgg 840
 caaacagccg gtggcggcaa aaattcatgt ggcacttttg gtgcggagtt tagtgccgct 900
 tcagacatga ttaataatgc gcaaaaaatc gttcaagaaa cccaacaact cagcgccaac 960
 caacaaaaaa atatcacaca accccataat ctcaacctta acacccttag cagtcttacg 1020
 gcttttagctc aaaaaatgct caaaaatgcg caatctcaag cagaaatttt aaaactagcc 1080
 aatcaagtgg agagcgattt taacaaactt tcttcaggcc atcttaaaga ctacataggg 1140
 aaatgcgatg cgagcgctat aagcagtgcg aatatgacaa tgcaaaatca aaagaacaat 1200
 tgggggaacg ggtgtgctgg cgtggaagaa actctgtctt cattaaaaac aagtgccgct 1260
 gattttaaca accaaacgcc acaaatcaat caagcgcaaa acctagccaa cacccttatt 1320
 caagaacttg gcaacaaccc ttttaggaat atgggcatga tcgcttcttc aaccacgaat 1380
 aacggcgctt tgaatggcct tggggtgcaa gtgggttata agcaattttt tggggaaaag 1440
 aaaagatggg ggtaagga ttatggttc tttgattaca accacgccta tatcaaatcc 1500
 aatttcttta actcggcttc tgatgtgtgg acttatgggg tgggcagcga tttattgttt 1560
 aatttcatca atgataaaaa caccaacttt ttaggcaaga ataacaagat ttcagtggga 1620
 ttttttgtag gtatgcctt agcagggact tcatggccta attctcaatt cgtgaattta 1680
 aaaacatca gcaatgttta tagcgctaaa gtgaatacgg ctaacttcca atttttattc 1740
 aatttgggct tgagaaccaa tctcgtaga cctaagaaaa aagatagtca tcatgaggct 1800
 caacatggca tggaaatggg cgtgaaaatc cctaccatta acacgaatta ttattctttt 1860
 ctagacacta aactagaata tcgaaggctt tatagcgtgt atctcaatta tgtgtttgcc 1920
 tattaa 1926

[0007]

<210> 5
 <211> 632
 <212> PRT
 <213> 幽门螺杆菌

<400> 5

Met Lys Lys Thr Lys Lys Thr Ile Leu Leu Ser Leu Thr Leu Ala Ser
 1 5 10 15
 Ser Leu Leu His Ala Glu Asp Asn Gly Val Phe Leu Ser Val Gly Tyr
 20 25 30
 Gln Ile Gly Glu Ala Val Gln Lys Val Lys Asn Ala Asp Lys Val Gln
 35 40 45
 Lys Leu Ser Asp Ala Tyr Glu Asn Leu Asn Lys Leu Leu Ala Asn His
 50 55 60
 Ser His Ser Asn Pro Glu Ala Ile Asn Ala Asn Ser Ala Thr Ala Ile
 65 70 75 80
 Asn Gln Ala Ile Gly Asn Leu Asn Ala Asn Thr Gln Asn Leu Ile Asp
 85 90 95

Lys Thr Asp Asn Ser Pro Ala Tyr Gln Ala Thr Leu Leu Ala Leu Lys
 100 105 110
 Ser Thr Val Gly Leu Trp Asn Ser Ile Ala Tyr Ala Val Ile Cys Gly
 115 120 125
 Gly Tyr Thr Asp Lys Pro Asn His Asn Thr Thr Glu Thr Phe Tyr Asn
 130 135 140
 Gln Pro Gly Gln Gly Ser Asp Ser Ile Thr Cys Gly Gly His Val Gly
 145 150 155 160
 Leu Leu Gln Ala Gly Lys Asn Asn Ser Leu Ser Ile Glu Gln Phe Ala
 165 170 175
 Thr Leu Asn Lys Ala Tyr Gln Ile Ile Gln Ala Ala Leu Lys Gln Gly
 180 185 190
 Leu Pro Ala Leu Ser Asp Thr Lys Lys Thr Val Glu Val Thr Ile Lys
 195 200 205
 Thr Ala Thr Asn Ala Asn Asn Ile Asn Val Asn Asn Asn Asn Asn
 210 215 220
 Ala Ala Asp Thr Thr Val Ser Ile Thr Asp Thr Phe Ile Asn Asp Ala
 225 230 235 240
 [0008] Gln Asn Leu Leu Thr Gln Ala Gln Thr Ile Ile Asn Thr Leu Gln Asp
 245 250 255
 Asn Cys Pro Gln Leu Lys Gly Lys Ser Ser Ser Asn Gly Gly Thr Asn
 260 265 270
 Gly Ala Asn Thr Pro Ser Trp Gln Thr Gly Ala Asn Gln Asn Ser Cys
 275 280 285
 Ser Val Phe Gly Thr Glu Phe Ser Ala Ile Ser Asp Met Ile Ser Asn
 290 295 300
 Ala Gln Asn Ile Val Gln Glu Thr Gln Gln Leu Asn Thr Thr Pro Leu
 305 310 315 320
 Lys Ser Ile Ala Gln Pro Asn Asn Phe Asn Leu Asn Ser Pro Asn Ser
 325 330 335
 Ile Ala Leu Ala Gln Ser Met Leu Lys Asn Ala Gln Ser Gln Ala Ala
 340 345 350
 Val Leu Lys Leu Ala Asn Gln Val Gly Ser Asp Phe Asn Arg Ile Ser
 355 360 365
 Thr Gly Val Leu Lys Asn Tyr Ile Glu Glu Cys Asn Ala Asn Ala Ser
 370 375 380
 Ser Glu Ser Val Ser Ser Asn Thr Trp Gly Lys Gly Cys Ala Gly Val
 385 390 395 400

Lys Gln Thr Leu Thr Ser Leu Glu Asn Ser Asn Ala Ser Phe Ser Ser
 405 410 415

Gln Thr Pro Gln Ile Asn Gln Ala Gln Asn Leu Ala Asn Thr Ile Val
 420 425 430

Gln Glu Leu Gly His Asn Pro Phe Lys Arg Val Gly Ile Ile Ser Ser
 435 440 445

Gln Thr Asn Asn Gly Ala Met Asn Gly Leu Gly Val Gln Val Gly Tyr
 450 455 460

Lys Gln Phe Phe Gly Glu Lys Lys Arg Trp Gly Leu Arg Tyr Tyr Gly
 465 470 475 480

Phe Phe Asp Tyr Asn His Thr Tyr Ile Lys Ser Asn Phe Phe Asn Ser
 485 490 495

Ala Ser Asp Val Trp Thr Tyr Gly Val Gly Ser Asp Leu Leu Phe Asn
 500 505 510

Phe Ile Asn Asp Lys Asn Thr Asn Phe Leu Gly Lys Asn Asn Gln Ile
 515 520 525

Ser Val Gly Leu Phe Gly Gly Ile Ala Leu Ala Gly Thr Ser Trp Leu
 530 535 540

[0009]

Asn Ser Gln Phe Val Asn Leu Lys Thr Ile Ser Asn Val Tyr Ser Ala
 545 550 555 560

Lys Val Asn Thr Ala Asn Phe Gln Phe Leu Phe Asn Leu Gly Leu Arg
 565 570 575

Thr Asn Leu Ala Arg Pro Lys Lys Lys Asp Ser His His Ala Gly Gln
 580 585 590

His Gly Met Glu Leu Gly Val Lys Ile Pro Thr Ile Asn Thr Asn Tyr
 595 600 605

Tyr Ser Phe Leu Asp Thr Lys Leu Glu Tyr Arg Arg Leu Tyr Ser Val
 610 615 620

Tyr Leu Asn Tyr Val Phe Ala Tyr
 625 630

<210> 6

<211> 1899

<212> DNA

<213> 幽门螺杆菌

<400> 6

atgaaaaaaaa cgaaaaaaac gattctgctt tctctaactc ttgcgtcatc attgctccat 60

gctgaagaca acggcgctttt ttaagcgtg ggctatcaaa tcggtgaagc ggttcaaaaa 120

gtgaaaaacg cgcacaaggt gcaaaaactt tcagacgctt atgaaaactt aaacaagctt 180

ttagctaate acagccacte caatccagaa gcgattaacg caaacagcgc cacagcgate 240

[0010]

```

aatcaagcga ttgtaattt aaacgcaaac acgcaaaatt taattgataa aacagacaat 300
tcccctgcct atcaagccac gcttttagcg ctaaaatcca cgggtggggt atggaatagc 360
atagcctacg cgcgcatatg cggaggctat acggataaac ccaatcacia caccacagaa 420
actttttaca accagccagg acaaggttca gattcaatca cttgcggtgg gcatgtgggg 480
ttacttcaag caggcaaaaa taattctcta tccattgaac aatttgcaac gctcaataaa 540
gcgatcaaaa tcatccaagc cgctttgaaa caaggtctcc ctgctttaag cgatacaaaa 600
aaaaagctgg aagtaacat taaaacagca accaacgcta acaacattaa tgtaataat 660
aacaataaca atgctgctga cactacagtt agcataactg atacttttat taacgatgca 720
caaaaccttt taaccaaacg gcaaacatc atcaaacacc ttcaagacia ttgcccgcga 780
ttgaaaggga agtctagtag caatgggtga actaatggcg caaacacccc ttcattggcaa 840
acaggcgcta accaaaaatc gtgcagcgtt tttggcacgg aatttagcgc tatttcagac 900
atgattagta acgctcaaaa catcgttcaa gaaacccaac agcttaatac caccocacta 960
aaaagcatcg cgcaacccaa caatttcaac cttaactccc ctaatagtat cgctttggct 1020
caaagcatcg tcaaaaacgc tcaatctcaa gcagcgggtt taaaactggc caatcaagt 1080
gggagcgatt ttaatagaat ttctacagga gttcttaaaa actatataga agaatgcaat 1140
gcgaatgctt caagtgaag cgtttctagt aacacttggg ggaaaggctg cgcgggcgtg 1200
aaacaaactc taacttcgct agaaaatagc aacgcttctt ttctagcca aacgcctcaa 1260
atcaatcaag cgcaaaacct cgtaacacc attgttcaag aactcggta taaccctttc 1320
aaacgggtgg gcatcattag ctctcaaac aataacgggg cgatgaatgg ccttgggggtg 1380
caagtgggtt ataagcaatt ttttgagaa aagaaaagat gggggtaag gtattatggt 1440
ttctttgatt acaaccacac ctatatcaaa tccaatttct ttaactggc ttctgatgtg 1500
tggacttatg ggttgggcag cgatttattg ttaacttca tcaacgataa aaacaccaac 1560
tttttaggta agaataacca gatttcagtg gggctttttg ggggaatcgc cttagcaggg 1620
acttcatgac ttaattctca attcgtgaat ttaaaaacca tcagcaatgt ctatagcgt 1680
aaagtgaata cggctaattt ccaattttta ttcaatttgg gcttgagaac caatctcgt 1740
agacctaaag aaaaagatag ccatcatcgc ggtcaacatg gcatggaatt gggcgtgaaa 1800
atccctacca ttaacacgaa ttactattct ttctagaca ctaaactaga atataggagg 1860
ctttatagcg tgtatctcaa ttatgtgttt gcctattaa 1899

```

```

<210> 7
<211> 12
<212> PRT
<213> 人工序列

```

```

<220>
<223> 靶向肽

```

<400> 7

```

Asn Trp Tyr Leu Pro Trp Leu Gly Thr Asn Asp Trp
1           5           10

```

```

<210> 8
<211> 53
<212> DNA
<213> 人工序列

```

	<220>		
	<223>	正向引物	
	<400>	8	
		gtttaacttt aagaaggaga tatacaaatg gcggttcaaa aagtgaaaaa cgc	53
	<210>	9	
	<211>	50	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	反向引物	
	<400>	9	
		tcaagcttat taatgatgat gatgatggtg ggcgccgta ttcgtggtg	50
	<210>	10	
	<211>	43	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	正向引物	
	<400>	10	
		caccatcatc atcatcatta ataagcttga tccggctgct aac	43
	<210>	11	
	<211>	30	
	<212>	DNA	
	<213>	人工序列	
	<220>		
[0011]	<223>	反向引物	
	<400>	11	
		gtttaacttt aagaaggaga tatacaaatg	30
	<210>	12	
	<211>	23	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	正向引物	
	<400>	12	
		ggtgacgctc agaacctgct gac	23
	<210>	13	
	<211>	23	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	反向引物	
	<400>	13	
		accaccttta gagttcagcg gag	23
	<210>	14	
	<211>	38	
	<212>	PRT	
	<213>	人工序列	
	<220>		
	<223>	HopQ-ID+HA	
	<400>	14	

Glu Lys Leu Glu Ala His Val Thr Thr Ser Lys Tyr Gln Gln Asp Asn
 1 5 10 15

 Gln Thr Lys Thr Thr Thr Ser Val Ile Asp Thr Thr Asn Tyr Pro Tyr
 20 25 30

 Asp Val Pro Asp Tyr Ala
 35

 <210> 15
 <211> 641
 <212> PRT
 <213> 幽门螺杆菌

 <400> 15

 Met Lys Lys Thr Lys Lys Thr Ile Leu Leu Ser Leu Thr Leu Ala Ser
 1 5 10 15

 Ser Leu Leu His Ala Glu Asp Asn Gly Val Phe Leu Ser Val Gly Tyr
 20 25 30

 Gln Ile Gly Glu Ala Val Gln Lys Val Lys Asn Ala Asp Lys Val Gln
 35 40 45

 Lys Leu Ser Asp Thr Tyr Glu Gln Leu Ser Arg Leu Leu Thr Asn Asp
 50 55 60

 [0012] Asn Gly Thr Asn Ser Lys Thr Ser Ala Gln Ala Ile Asn Gln Ala Val
 65 70 75 80

 Asn Asn Leu Asn Glu Arg Ala Lys Thr Leu Ala Gly Gly Thr Thr Asn
 85 90 95

 Ser Pro Ala Tyr Gln Ala Thr Leu Leu Ala Leu Arg Ser Val Leu Gly
 100 105 110

 Leu Trp Asn Ser Met Gly Tyr Ala Val Ile Cys Gly Gly Tyr Thr Lys
 115 120 125

 Ser Pro Gly Glu Asn Asn Gln Lys Asn Phe His Tyr Thr Asp Glu Asn
 130 135 140

 Gly Asn Gly Thr Thr Ile Asn Cys Gly Gly Ser Thr Asn Ser Asn Gly
 145 150 155 160

 Thr His Ser Ser Asn Gly Thr Asn Thr Leu Lys Ala Asp Lys Asn Val
 165 170 175

 Ser Leu Ser Ile Glu Gln Tyr Glu Lys Ile His Glu Ser Tyr Gln Ile
 180 185 190

 Leu Ser Lys Ala Leu Lys Gln Ala Gly Leu Ala Pro Leu Asn Ser Lys
 195 200 205

 Gly Glu Lys Leu Glu Ala His Val Thr Thr Ser Lys Tyr Gln Gln Asp
 210 215 220

Ser Gln Thr Lys Thr Thr Thr Ser Val Ile Asp Thr Thr Asn Asp Ala
 225 230 235 240
 Gln Asn Leu Leu Thr Gln Ala Gln Thr Ile Val Asn Thr Leu Lys Asp
 245 250 255
 Tyr Cys Pro Met Leu Ile Ala Lys Ser Ser Ser Gly Ser Gly Gly Gly
 260 265 270
 Ala Ala Thr Asn Thr Pro Ser Trp Gln Thr Ala Gly Gly Gly Lys Asn
 275 280 285
 Ser Cys Glu Thr Phe Gly Ala Glu Phe Ser Ala Ala Ser Asp Met Ile
 290 295 300
 Asn Asn Ala Gln Lys Ile Val Gln Glu Thr Gln Gln Leu Ser Ala Asn
 305 310 315 320
 Gln Pro Lys Asn Ile Thr Gln Pro His Asn Leu Asn Leu Asn Thr Pro
 325 330 335
 Ser Ser Leu Thr Ala Leu Ala Gln Lys Met Leu Lys Asn Ala Gln Ser
 340 345 350
 Gln Ala Glu Ile Leu Lys Leu Ala Asn Gln Val Glu Ser Asp Phe Asn
 355 360 365
 [0013] Lys Leu Ser Ser Gly His Leu Lys Asp Tyr Ile Gly Lys Cys Asp Met
 370 375 380
 Ser Ala Ile Ser Ser Thr Asn Met Thr Met Gln Ser Gln Lys Asn Asn
 385 390 395 400
 Trp Gly Asn Gly Cys Ala Gly Val Glu Glu Thr Leu Thr Ser Leu Lys
 405 410 415
 Thr Ser Ala Ala Asp Phe Asn Asn Gln Thr Pro Gln Ile Asn Gln Ala
 420 425 430
 Gln Asn Leu Ala Asn Thr Leu Ile Gln Glu Leu Gly Asn Asn Pro Phe
 435 440 445
 Arg Asn Met Gly Met Ile Ala Ser Ser Thr Thr Asn Asn Gly Ala Leu
 450 455 460
 Asn Gly Leu Gly Val Gln Val Gly Tyr Lys Gln Phe Phe Gly Glu Lys
 465 470 475 480
 Lys Arg Trp Gly Leu Arg Tyr Tyr Gly Phe Phe Asp Tyr Asn His Ala
 485 490 495
 Tyr Ile Lys Ser Asn Phe Phe Asn Ser Ala Ser Asp Val Trp Thr Tyr
 500 505 510
 Gly Val Gly Ser Asp Leu Leu Phe Asn Phe Ile Asn Asp Lys Asn Thr
 515 520 525

Asn Phe Leu Gly Lys Asn Asn Gln Ile Ser Phe Gly Leu Phe Gly Gly
 530 535 540

Ile Ala Leu Ala Gly Thr Ser Trp Leu Asn Ser Gln Phe Val Asn Leu
 545 550 555 560

Lys Thr Ile Ser Asn Val Tyr Ser Ala Lys Val Asn Thr Ala Asn Phe
 565 570 575

Gln Phe Leu Phe Asn Leu Gly Leu Arg Thr Asn Leu Ala Arg Pro Lys
 580 585 590

Lys Lys Asp Ser His His Ala Ala Gln His Gly Ile Glu Leu Gly Val
 595 600 605

Lys Ile Pro Thr Ile Asn Thr Asn Tyr Tyr Ser Phe Leu Asp Thr Lys
 610 615 620

Leu Glu Tyr Arg Arg Leu Tyr Ser Val Tyr Leu Asn Tyr Val Phe Ala
 625 630 635 640

Tyr

<210> 16
 <211> 1926
 <212> DNA
 <213> 幽门螺杆菌

[0014]

<400> 16
 atgaaaaaaaa cgaaaaaaaaa gattctgctt tctctaacte ttgcgtcadc attgctccat 60
 gctgaagaca acggcgcttt ttaagcgtg ggttatcaaa ttggtgaagc ggttcaaaaa 120
 gtgaaaaacg ccgacaaggt gcaaaaactt tcagacactt atgaacaatt aagccggctt 180
 ttaaccaacg ataatggcac aaactcaag acaagcgcgc aagcgatcaa ccaagcggtt 240
 aataatttga acgaacgcgc aaaaacttta gccggtgga caaccaattc cctgcctat 300
 caagccacgc ttttagcgtt aagatcgggt ttaggctat ggaatagcat gggttatgca 360
 gtcatatgca gaggttatac caaaagtcca ggcgaaaaca atcaaaaaaa tttccactac 420
 accgatgaga atggcaacgg cactacaatc aattgcgggt ggagcacaac tagtaatggc 480
 actcatagtt ctaatggcac aaatacatta aaagcagaca aaaatgtttc tctatctatt 540
 gagcaatatg aaaaaatcca tgaatcctat cagattcttt caaaagcttt aaaacaagcc 600
 ggccttgctc ctttaaatag caaaggggaa aagttagaag cgcattgtaac cacatcaaac 660
 tatcaacaag atagtcaaac taaaacgaca acttctgtta ttgatacgac taatgatgca 720
 caaaatcttt tgactcaagc gcaaacgatt gtcaatacc ttaaagatta ttgcccatg 780
 ttgatagcga aatctagtag tggaagtgtt ggcggagctg ctacaacac cccttcatgg 840
 caaacagccg gtggcgcaaa aaattcatgc gagacttttg gtgcggagtt tagtgccgct 900
 tcagacatga ttaataatgc gcaaaaaatc gttcaagaaa cccaacaact cagcgcacaac 960
 caaccaaaaa atatacaca accccataat ctcaacctta acaccctag cagtcttacg 1020
 gcttttagctc aaaaaatgct caaaaacgcg caatctcaag cagaaatctt aaaactagcc 1080
 aatcaagtgg agagcgattt taacaaactt tcttcaggcc atcttaaga ctacataggg 1140

```

aatgcgata tgagtgtat aagcagtacg aatatgacaa tgcaaagtca aaagaacaat 1200
tgggggaacg ggtgcgctgg cgtggaagaa actctaactt cattaaaaac aagcgccgct 1260
gattttaaca accaaacgcc tcaaatcaat caagcgcaaa acctagctaa cacccttatt 1320
caagaacttg gcaacaaccc ttttaggaat atgggcatga tcgcttcttc aaccacgaat 1380
aacggcgctc tgaatggcct tggggtgcaa gtgggttata agcaattttt tggagaaaag 1440
aaaagatggg ggttagata ttatggtttc ttgattaca accacgcta tatcaaatcc 1500
aatttcttta attcggcttc tgatgtgtgg acttatgggg tgggtagcga tttattgttt 1560
aatttcatca acgataaaaa caccaacttt ttaggcaaga ataaccagat ttcttttggg 1620
ctttttggag gaatcgctt agcagggact tcatggctta attctcaatt cgtgaattta 1680
aaaaccatca gcaatgtcta tagcgctaaa gtgaatacgg ctaacttcca atttttattc 1740
aatctcgctc tgagaaccaa tctcgctagg cctaagaaaa aagatagcca tcatggcgct 1800
caacatggca tagaattagg cgtgaaaate cctaccatta acacgaatta ctattctttt 1860
ctagacacta aactagaata tagaaggctt tatagcgtgt atctcaatta tgtgttcgcg 1920
tattaa 1926

```

```

<210> 17
<211> 14
<212> PRT
<213> 人工序列

```

```

<220>
<223> 一般共有序列

```

[0015]

```

<220>
<221> misc_feature
<222> (5)..(5)
<223> Xaa是Thr或Tyr或者缺失

```

```

<220>
<221> misc_feature
<222> (6)..(6)
<223> Xaa是Lys或Asn或者缺失

```

```

<220>
<221> misc_feature
<222> (7)..(7)
<223> Xaa是Ser、Lys、Asn或Thr或者缺失

```

```

<220>
<221> misc_feature
<222> (9)..(9)
<223> Xaa是Gly、Ser、Gln、Arg、Thr、Ile或Val或者缺失

```

```

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa是Asn或Gly或者缺失

```

```

<220>
<221> misc_feature
<222> (12)..(12)
<223> Xaa是Asn或His或者缺失

```

```

<400> 17

```

```

Cys Gly Gly Tyr Xaa Xaa Xaa Pro Xaa Glu Xaa Xaa Gln Lys
1          5          10

```

```

<210> 18
<211> 30

```

- <212> PRT
 <213> 人工序列
- <220>
 <223> 一般共有序列
- <220>
 <221> misc_feature
 <222> (4)..(4)
 <223> Xaa是Ser、Gly、Asn、Thr或Phe或者缺失
- <220>
 <221> misc_feature
 <222> (5)..(5)
 <223> Xaa是Thr或Ile或者缺失
- <220>
 <221> misc_feature
 <222> (6)..(6)
 <223> Xaa是Asn、Gly或Lys或者缺失
- <220>
 <221> misc_feature
 <222> (7)..(7)
 <223> Xaa是Ser或Ala或者缺失
- <220>
 <221> misc_feature
 <222> (8)..(8)
 <223> Xaa是Asn或Asp或者缺失
- <220>
 <221> misc_feature
 <222> (10)..(10)
 <223> Xaa是Gln、Lys或Arg或者缺失
- [0016] <220>
 <221> misc_feature
 <222> (11)..(11)
 <223> Xaa是Thr、Val或Ser或者缺失
- <220>
 <221> misc_feature
 <222> (12)..(12)
 <223> Xaa是His、Gln或Tyr或者缺失
- <220>
 <221> misc_feature
 <222> (13)..(13)
 <223> Xaa是Ser或Asn或者缺失
- <220>
 <221> misc_feature
 <222> (14)..(14)
 <223> Xaa是Ser、Pro或Asn或者缺失
- <220>
 <221> misc_feature
 <222> (15)..(15)
 <223> Xaa是Asn或Ser或者缺失
- <220>
 <221> misc_feature
 <222> (17)..(17)
 <223> Xaa是Thr或Val
- <220>
 <221> misc_feature
 <222> (18)..(18)
 <223> Xaa是Asn或Ser
- <220>
 <221> misc_feature
 <222> (19)..(19)
 <223> Xaa是Thr、Leu或Met或者缺失

<220>
 <221> misc_feature
 <222> (21)..(21)
 <223> Xaa是Lys或Pro或者缺失

<220>
 <221> misc_feature
 <222> (23)..(23)
 <223> Xaa是Asp、Gly或Ala或者缺失

<220>
 <221> misc_feature
 <222> (25)..(25)
 <223> Xaa是Asn或Gly或者缺失

<220>
 <221> misc_feature
 <222> (26)..(26)
 <223> Xaa是Val或Ser或者缺失

<400> 18
 Cys Gly Gly Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Gly
 1 5 10 15

Xaa Xaa Xaa Leu Xaa Ala Xaa Lys Xaa Xaa Ser Leu Ser Ile
 20 25 30

<210> 19
 <211> 33
 <212> PRT
 <213> 人工序列

<220>
 <223> 一般共有序列

[0017]

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> Xaa是Met、Ile或Val或者缺失

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> Xaa是Ala或Gly或者缺失

<220>
 <221> misc_feature
 <222> (7)..(7)
 <223> Xaa是Lys或Arg或者缺失

<220>
 <221> misc_feature
 <222> (8)..(8)
 <223> Xaa是Ser或Thr或者缺失

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> Xaa是Ser或Thr或者缺失

<220>
 <221> misc_feature
 <222> (10)..(10)
 <223> Xaa是Ser、Asn或Gly或者缺失

<220>
 <221> misc_feature
 <222> (11)..(11)
 <223> Xaa是Gly、Asn、Glu、Ser或Asp或者缺失

<220>

```

<221> misc_feature
<222> (12)..(12)
<223> Xaa是Ser、Gly或Asn或者缺失

<220>
<221> misc_feature
<222> (13)..(13)
<223> Xaa是Ser、Met、Gly、Asn或Thr或者缺失

<220>
<221> misc_feature
<222> (14)..(14)
<223> Xaa是Gly、Ala、Thr、Ser、Asn或Met或者缺失

<220>
<221> misc_feature
<222> (15)..(15)
<223> Xaa是Gly、Asn、Thr、Ala或Val或者缺失

<220>
<221> misc_feature
<222> (16)..(16)
<223> Xaa是Ala、Asn、Gly或Ser或者缺失

<220>
<221> misc_feature
<222> (17)..(17)
<223> Xaa是Thr、Asn、Ala、Gly或Ser或者缺失

<220>
<221> misc_feature
<222> (18)..(18)
<223> Xaa是Thr或Ala或者缺失

<220>
<221> misc_feature
<222> (20)..(20)
<223> Xaa是Thr或Ala或者缺失

<220>
<221> misc_feature
<222> (25)..(25)
<223> Xaa是Thr或Ile或者缺失

<220>
<221> misc_feature
<222> (26)..(26)
<223> Xaa是Ala、Ser、Thr或Asn或者缺失

<220>
<221> misc_feature
<222> (27)..(27)
<223> Xaa是Gly或Ser或者缺失

<220>
<221> misc_feature
<222> (28)..(28)
<223> Xaa是Gly或Asn或者缺失

<220>
<221> misc_feature
<222> (29)..(29)
<223> Xaa是Gly、Leu或Ser或者缺失

<220>
<221> misc_feature
<222> (32)..(32)
<223> Xaa是Ser或Ala或者缺失

<400> 19

Cys Pro Xaa Leu Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1          5          10          15

Xaa Xaa Asn Xaa Pro Ser Trp Gln Xaa Xaa Xaa Xaa Xaa Lys Asn Xaa

```

[0018]

	20	25	30
	Cys		
	<210> 20		
	<211> 41		
	<212> PRT		
	<213> 人工序列		
	<220>		
	<223> 一般共有序列		
	<220>		
	<221> misc_feature		
	<222> (3)..(3)		
	<223> Xaa是Gly或Asp或者缺失		
	<220>		
	<221> misc_feature		
	<222> (4)..(4)		
	<223> Xaa是His或Tyr或者缺失		
	<220>		
	<221> misc_feature		
	<222> (7)..(7)		
	<223> Xaa是Asp或Asn或者缺失		
	<220>		
	<221> misc_feature		
	<222> (10)..(10)		
	<223> Xaa是Gly或Arg或者缺失		
[0019]	<220>		
	<221> misc_feature		
	<222> (14)..(14)		
	<223> Xaa是Met、Ala或Val或者缺失		
	<220>		
	<221> misc_feature		
	<222> (16)..(16)		
	<223> Xaa是Ala或Gly或者缺失		
	<220>		
	<221> misc_feature		
	<222> (17)..(17)		
	<223> Xaa是Ile或Val或者缺失		
	<220>		
	<221> misc_feature		
	<222> (19)..(19)		
	<223> Xaa是Ser或Gly或者缺失		
	<220>		
	<221> misc_feature		
	<222> (20)..(20)		
	<223> Xaa是任意氨基酸或者缺失		
	<220>		
	<221> misc_feature		
	<222> (21)..(21)		
	<223> Xaa是任意氨基酸或者缺失		
	<220>		
	<221> misc_feature		
	<222> (22)..(22)		
	<223> Xaa是任意氨基酸或者缺失		
	<220>		
	<221> misc_feature		
	<222> (23)..(23)		
	<223> Xaa是Thr、Ala或Ser或者缺失		

<220>
 <221> misc_feature
 <222> (26)..(26)
 <223> Xaa是Thr或Ala或者缺失

 <220>
 <221> misc_feature
 <222> (27)..(27)
 <223> Xaa是Met、Pro、Ala或Gln或者缺失

 <220>
 <221> misc_feature
 <222> (28)..(28)
 <223> Xaa是Gln、Arg、Lys或His或者缺失

 <220>
 <221> misc_feature
 <222> (29)..(29)
 <223> Xaa是Ser或Asn或者缺失

 <220>
 <221> misc_feature
 <222> (30)..(30)
 <223> Xaa是Gln或Met或者缺失

 <220>
 <221> misc_feature
 <222> (32)..(32)
 <223> Xaa是Asn或Ser或者缺失

 <220>
 <221> misc_feature
 <222> (33)..(33)
 <223> Xaa是Asn或Thr或者缺失

 [0020]
 <220>
 <221> misc_feature
 <222> (34)..(34)
 <223> Xaa是Thr、Asn或Ile或者缺失

 <220>
 <221> misc_feature
 <222> (37)..(37)
 <223> Xaa是Asn或Lys或者缺失

 <400> 20

 Ser Ser Xaa Xaa Leu Lys Xaa Tyr Ile Xaa Lys Cys Asp Xaa Ser Xaa
 1 5 10 15

 Xaa Ser Xaa Xaa Xaa Xaa Xaa Asn Met Xaa Xaa Xaa Xaa Xaa Lys Xaa
 20 25 30

 Xaa Xaa Trp Gly Xaa Gly Cys Ala Gly
 35 40

 <210> 21
 <211> 14
 <212> PRT
 <213> 人工序列

 <220>
 <223> 肽

 <400> 21

 Cys Gly Gly Tyr Thr Lys Ser Pro Gly Glu Asn Asn Gln Lys
 1 5 10

 <210> 22
 <211> 29
 <212> PRT

<213> 人工序列

<220>

<223> 肽

<400> 22

Cys Gly Gly Ser Thr Asn Ser Asn Gly Thr His Ser Ser Asn Gly Thr
1 5 10 15

Asn Thr Leu Lys Ala Asp Lys Asn Val Ser Leu Ser Ile
20 25

<210> 23

<211> 30

<212> PRT

<213> 人工序列

<220>

<223> 共有序列

<400> 23

Cys Gly Gly Ser Thr Asn Ser Asn Gly Gln Thr His Ser Ser Asn Gly
1 5 10 15

Thr Asn Thr Leu Lys Ala Asp Lys Asn Val Ser Leu Ser Ile
20 25 30

<210> 24

<211> 33

<212> PRT

<213> 人工序列

[0021]

<220>

<223> 肽

<400> 24

Cys Pro Met Leu Ile Ala Lys Ser Ser Ser Gly Ser Ser Gly Gly Ala
1 5 10 15

Ala Thr Asn Thr Pro Ser Trp Gln Thr Ala Gly Gly Gly Lys Asn Ser
20 25 30

Cys

<210> 25

<211> 33

<212> PRT

<213> 人工序列

<220>

<223> 共有序列

<400> 25

Cys Pro Met Leu Ile Ala Lys Ser Ser Ser Gly Ser Ser Gly Gly Ala
1 5 10 15

Thr Thr Asn Thr Pro Ser Trp Gln Thr Ala Gly Gly Gly Lys Asn Ser
20 25 30

Cys

<210> 26
 <211> 37
 <212> PRT
 <213> 人工序列

<220>
 <223> 肽

<400> 26

Ser Ser Gly His Leu Lys Asp Tyr Ile Gly Lys Cys Asp Met Ser Ala
 1 5 10 15

Ile Ser Ser Thr Asn Met Thr Met Gln Ser Gln Lys Asn Asn Trp Gly
 20 25 30

Asn Gly Cys Ala Gly
 35

<210> 27
 <211> 38
 <212> PRT
 <213> 人工序列

<220>
 <223> 共有序列

<400> 27

Ser Ser Gly His Leu Lys Asp Tyr Ile Gly Lys Cys Asp Ala Ser Ala
 1 5 10 15

[0022] Ile Ser Ser Ala Asn Met Thr Met Gln Asn Gln Lys Asn Asn Thr Trp
 20 25 30

Gly Asn Gly Cys Ala Gly
 35

<210> 28
 <211> 108
 <212> PRT
 <213> 人

<400> CEAM1_HUMAN N-结构域

Gln Leu Thr Thr Glu Ser Met Pro Phe Asn Val Ala Glu Gly Lys Glu
 1 5 10 15

Val Leu Leu Leu Val His Asn Leu Pro Gln Gln Leu Phe Gly Tyr Ser
 20 25 30

Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Arg Gln Ile Val Gly Tyr
 35 40 45

Ala Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Asn Ser Gly Arg
 50 55 60

Glu Thr Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Val Thr Gln
 65 70 75 80

Asn Asp Thr Gly Phe Tyr Thr Leu Gln Val Ile Lys Ser Asp Leu Val
 85 90 95

Asn Glu Glu Ala Thr Gly Gln Phe His Val Tyr Pro
 100 105

<210> 29
 <211> 110
 <212> PRT
 <213> 人

<400> CEAM5_HUMAN N-结构域

Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val Ala Glu Gly Lys Glu
 1 5 10 15

Val Leu Leu Leu Val His Asn Leu Pro Gln His Leu Phe Gly Tyr Ser
 20 25 30

Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Arg Gln Ile Ile Gly Tyr
 35 40 45

Val Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Tyr Ser Gly Arg
 50 55 60

Glu Ile Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Ile Ile Gln
 65 70 75 80

Asn Asp Thr Gly Phe Tyr Thr Leu His Val Ile Lys Ser Asp Leu Val
 85 90 95

Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu
 100 105 110

[0023]

<210> 30
 <211> 108
 <212> PRT
 <213> 人

<400> CEAM6_HUMAN N-结构域

Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val Ala Glu Gly Lys Glu
 1 5 10 15

Val Leu Leu Leu Ala His Asn Leu Pro Gln Asn Arg Ile Gly Tyr Ser
 20 25 30

Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Ser Leu Ile Val Gly Tyr
 35 40 45

Val Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Tyr Ser Gly Arg
 50 55 60

Glu Thr Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Val Thr Gln
 65 70 75 80

Asn Asp Thr Gly Phe Tyr Thr Leu Gln Val Ile Lys Ser Asp Leu Val
 85 90 95

Asn Glu Glu Ala Thr Gly Gln Phe His Val Tyr Pro
 100 105

<210> 31
 <211> 108

<212> PRT
 <213> 人
 <400> CEAM8_HUMAN N-结构域
 Gln Leu Thr Ile Glu Ala Val Pro Ser Asn Ala Ala Glu Gly Lys Glu
 1 5 10 15
 Val Leu Leu Leu Val His Asn Leu Pro Gln Asp Pro Arg Gly Tyr Asn
 20 25 30
 Trp Tyr Lys Gly Glu Thr Val Asp Ala Asn Arg Arg Ile Ile Gly Tyr
 35 40 45
 Val Ile Ser Asn Gln Gln Ile Thr Pro Gly Pro Ala Tyr Ser Asn Arg
 50 55 60
 Glu Thr Ile Tyr Pro Asn Ala Ser Leu Leu Met Arg Asn Val Thr Arg
 65 70 75 80
 Asn Asp Thr Gly Ser Tyr Thr Leu Gln Val Ile Lys Leu Asn Leu Met
 85 90 95
 Ser Glu Glu Val Thr Gly Gln Phe Ser Val His Pro
 100 105

<210> 32
 <211> 99
 <212> PRT
 <213> 大鼠

[0024]

<400> CEAM1_RAT N-结构域
 Pro Pro Asn Val Val Glu Glu Ser Ser Val Leu Leu Leu Thr His Asn
 1 5 10 15
 Leu Pro Gln Glu Phe Gln Val Phe Tyr Trp Tyr Lys Val Thr Thr Thr
 20 25 30
 Gly Leu Asn Ser Glu Ile Ala Arg Tyr Ile Arg Ser Ser Asn Thr Ser
 35 40 45
 Gln Thr Glu Pro Ala Tyr Ser Gly Arg Val Thr Ile Tyr Ser Asn Gly
 50 55 60
 Ser Leu Phe Phe Gln Asn Val Asn Lys Thr Asp Glu Gly Pro Tyr Thr
 65 70 75 80
 Leu Ser Val Ile Asp Lys Gln Phe Asn Pro Ile Gln Thr Ser Val Gln
 85 90 95
 Phe Arg Val

<210> 33
 <211> 108
 <212> PRT
 <213> 小鼠

<400> CEAM1_MOUSE N-结构域
 Glu Val Thr Ile Glu Ala Val Pro Pro Gln Val Ala Glu Asp Asn Asn

	1	5	10	15
	Val	Leu Leu Leu	Val His Asn Leu	Pro Leu Ala Leu Gly Ala Phe Ala
		20	25	30
	Trp Tyr Lys	Gly Asn Thr Thr	Ala Ile Asp Lys Glu	Ile Ala Arg Phe
		35	40	45
[0025]	Val Pro Asn Ser	Asn Met Asn Phe Thr	Gly Gln Ala Tyr Ser	Gly Arg
	50	55	60	
	Glu Ile Ile Tyr Ser	Asn Gly Ser Leu Leu	Phe Gln Met Ile Thr	Met
	65	70	75	80
	Lys Asp Met Gly	Val Tyr Thr Leu Asp	Met Thr Asp Glu Asn Tyr	Arg
		85	90	95
	Arg Thr Gln Ala Thr	Val Arg Phe His	Val His Pro	
		100	105	

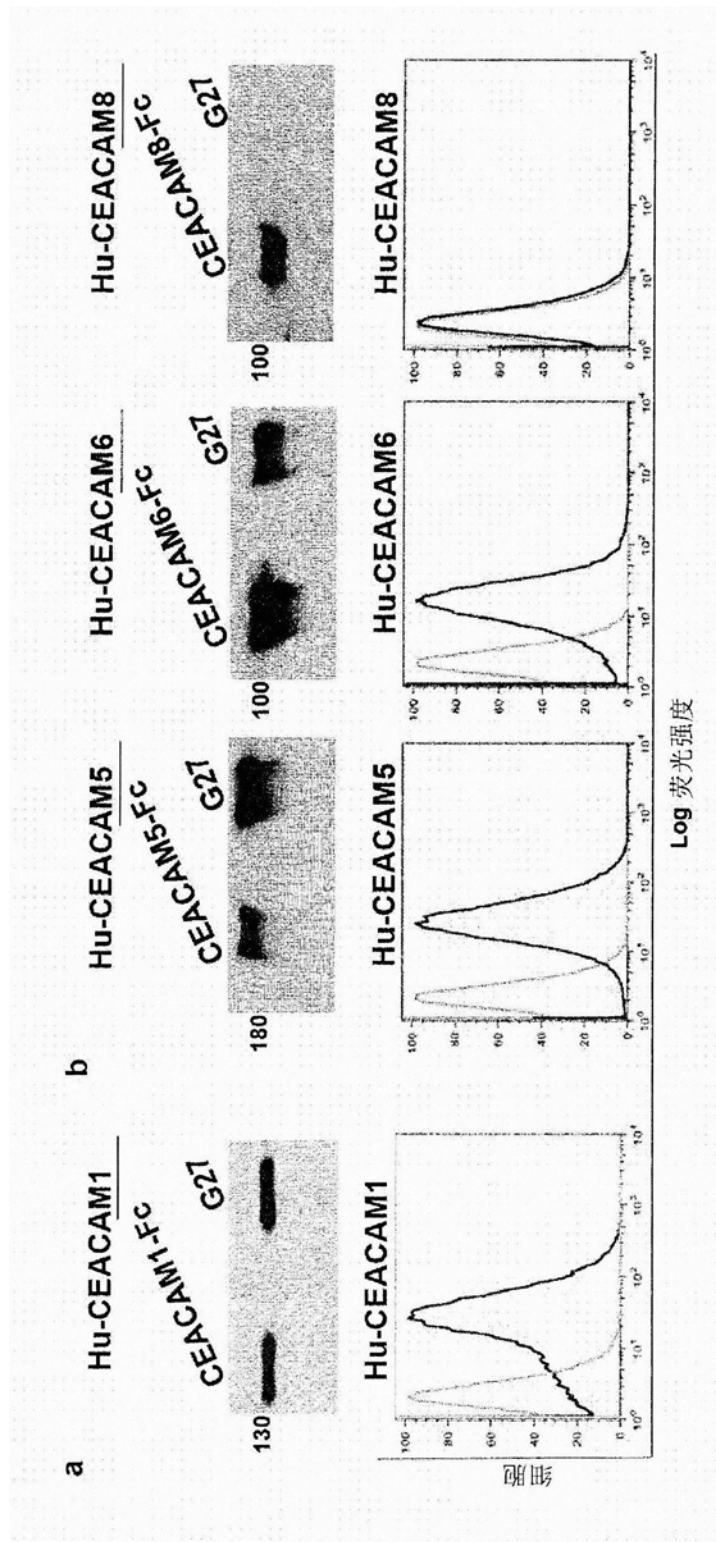


图1

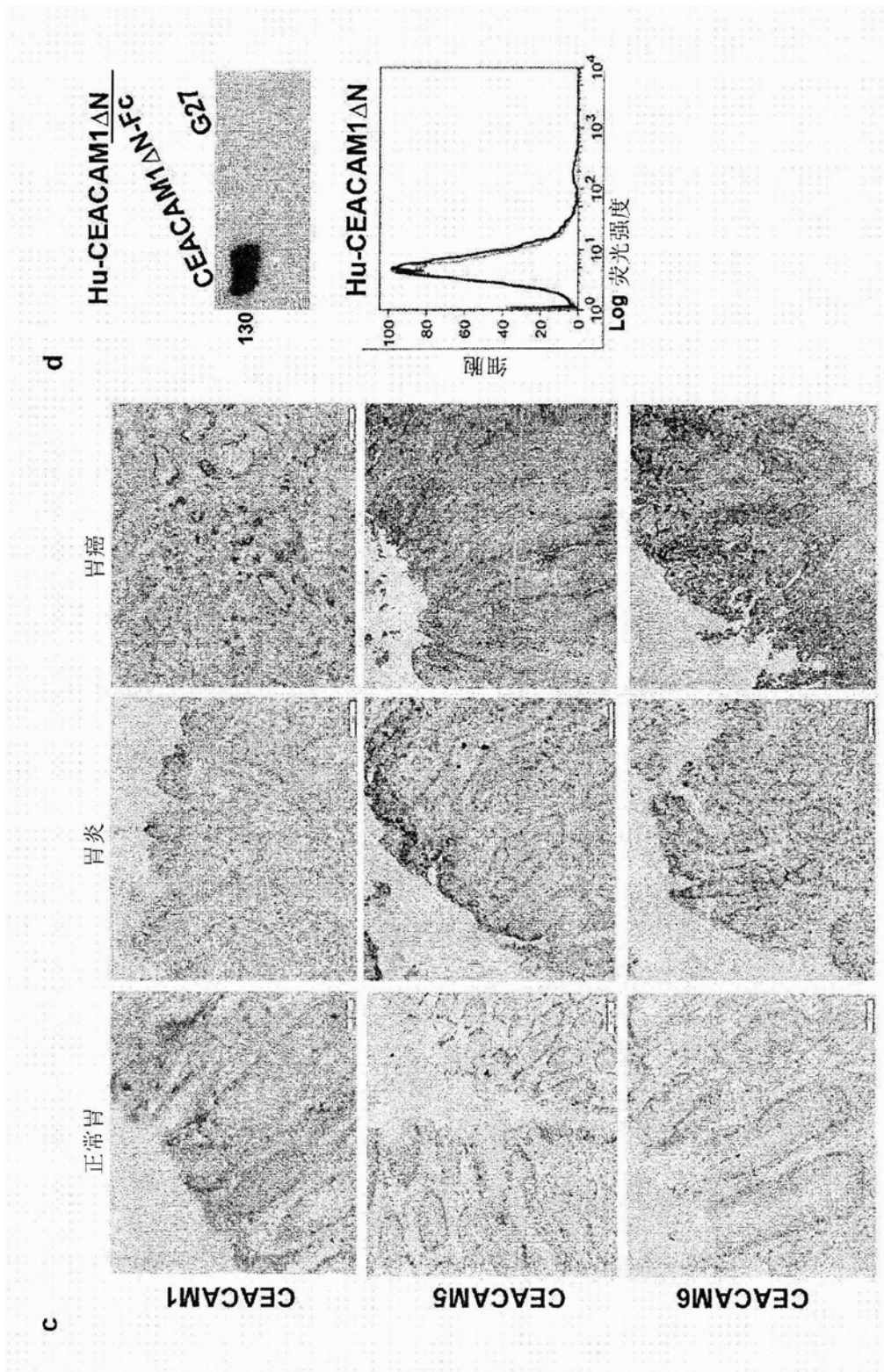


图1(续)

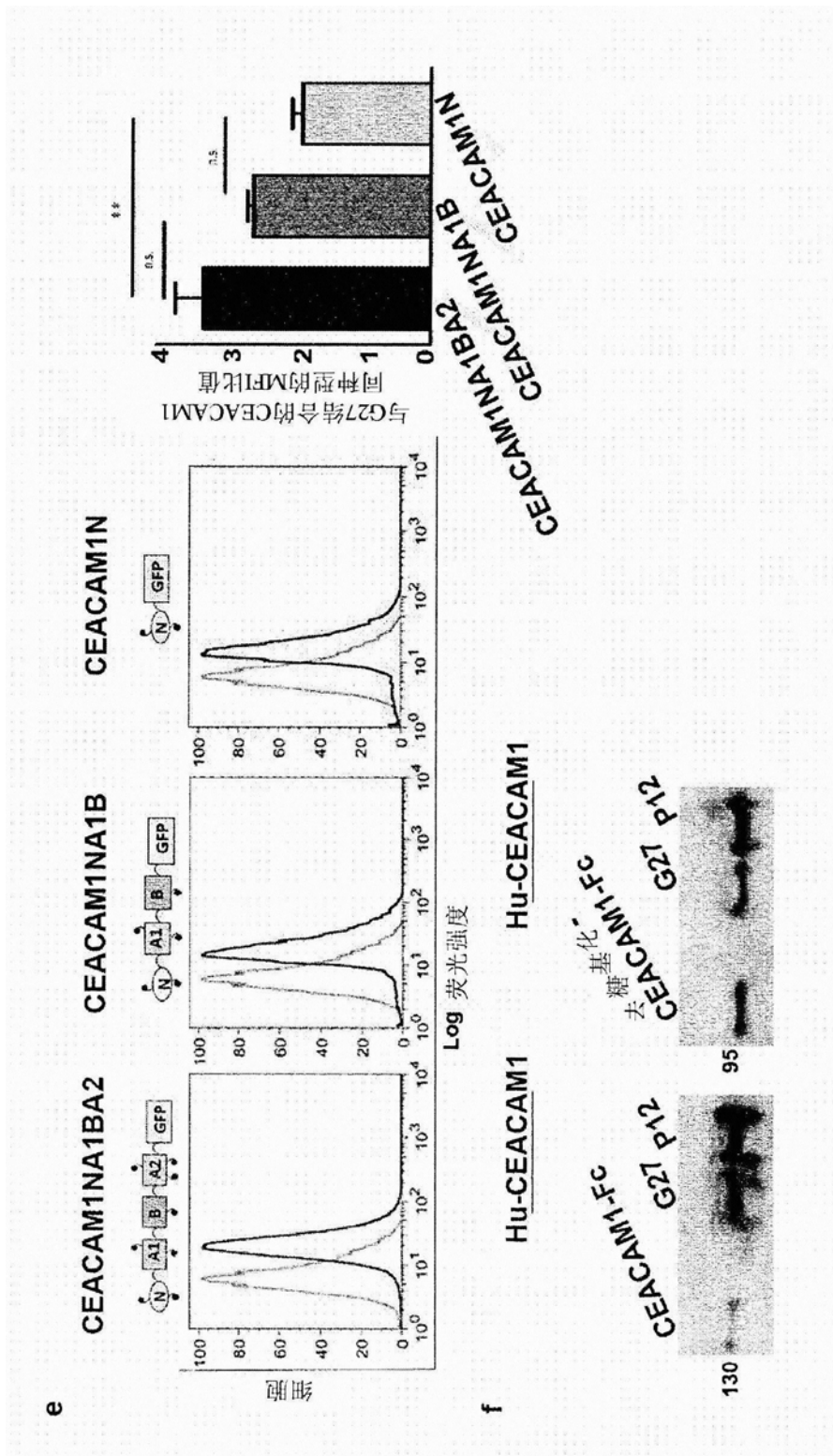


图1(续)

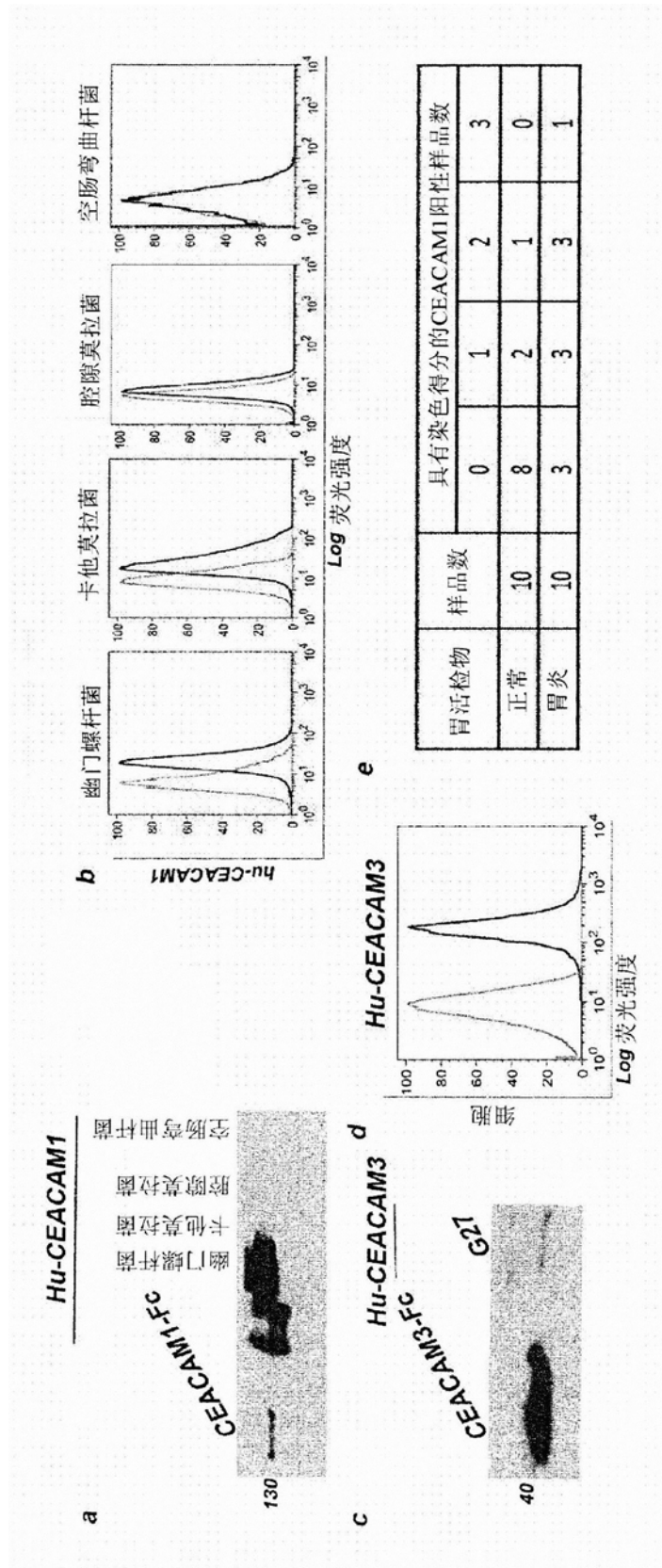


图2

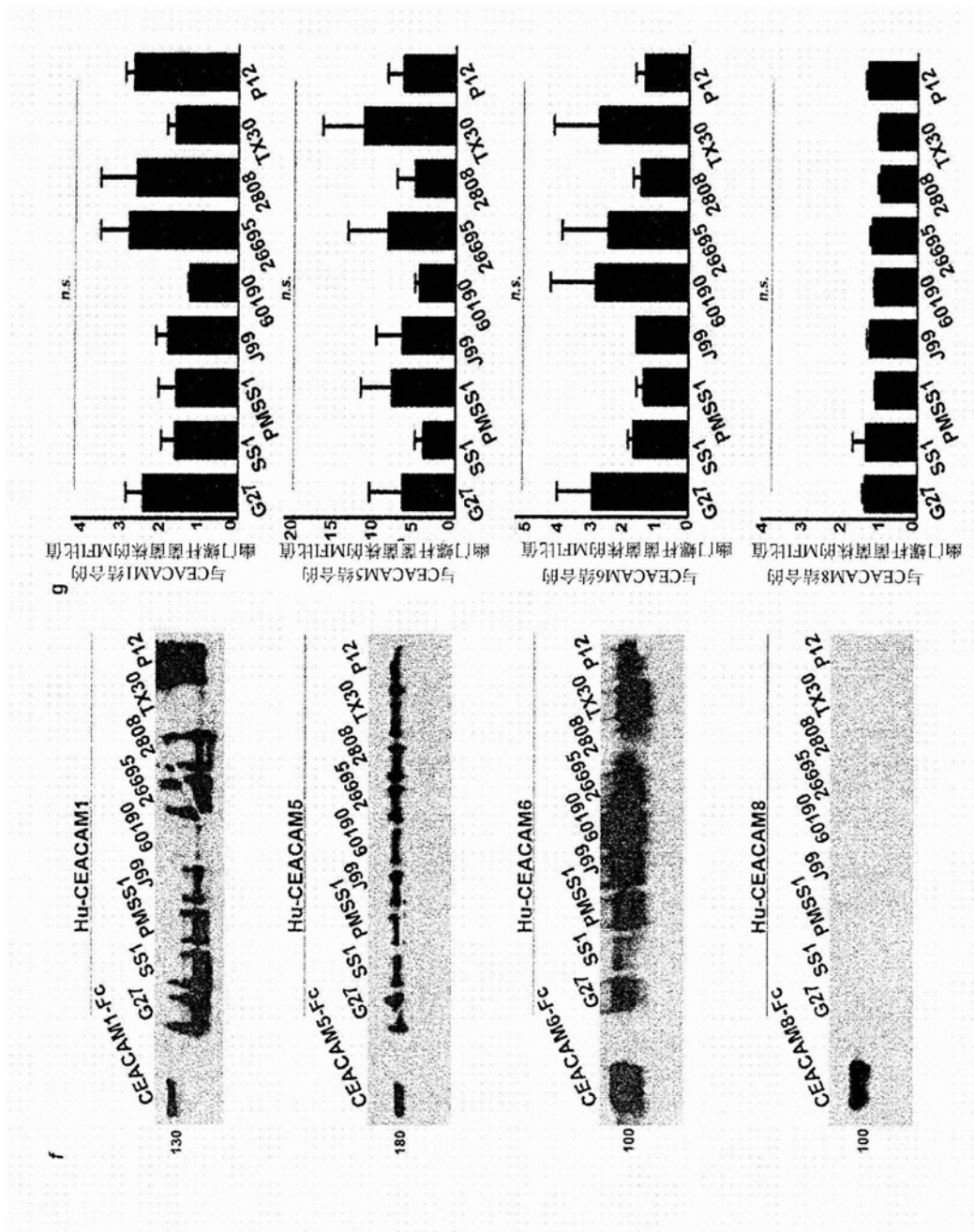


图2(续)

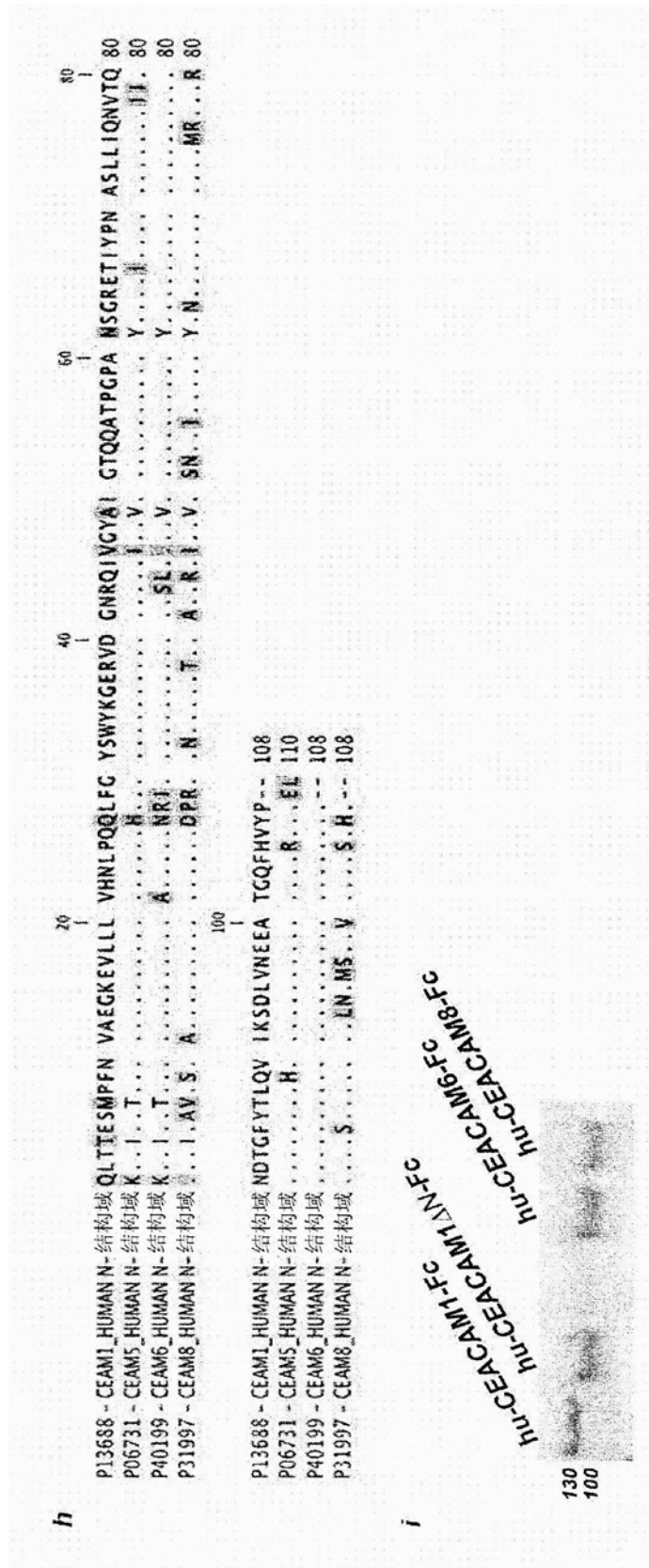


图2(续)

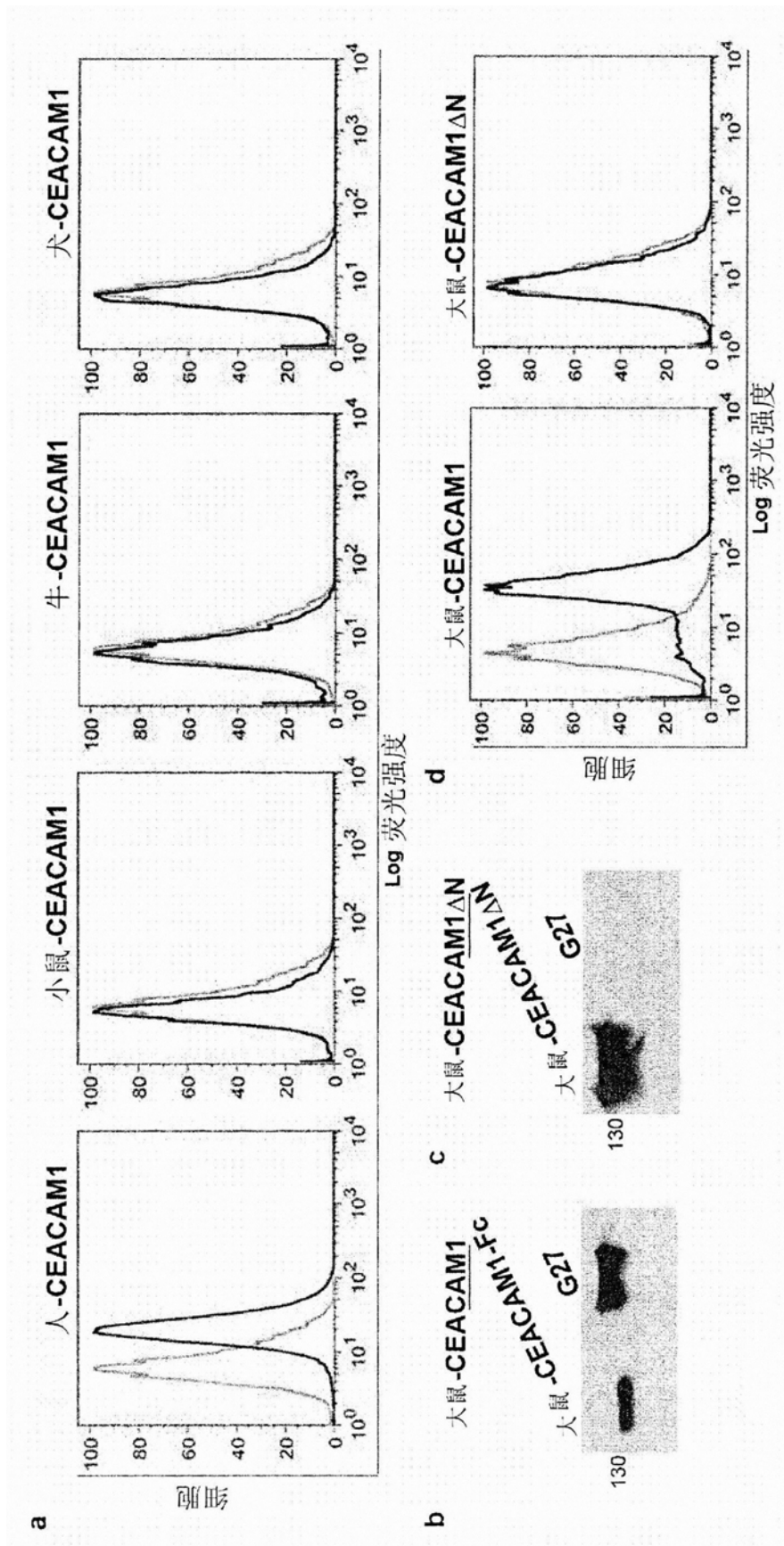


图3

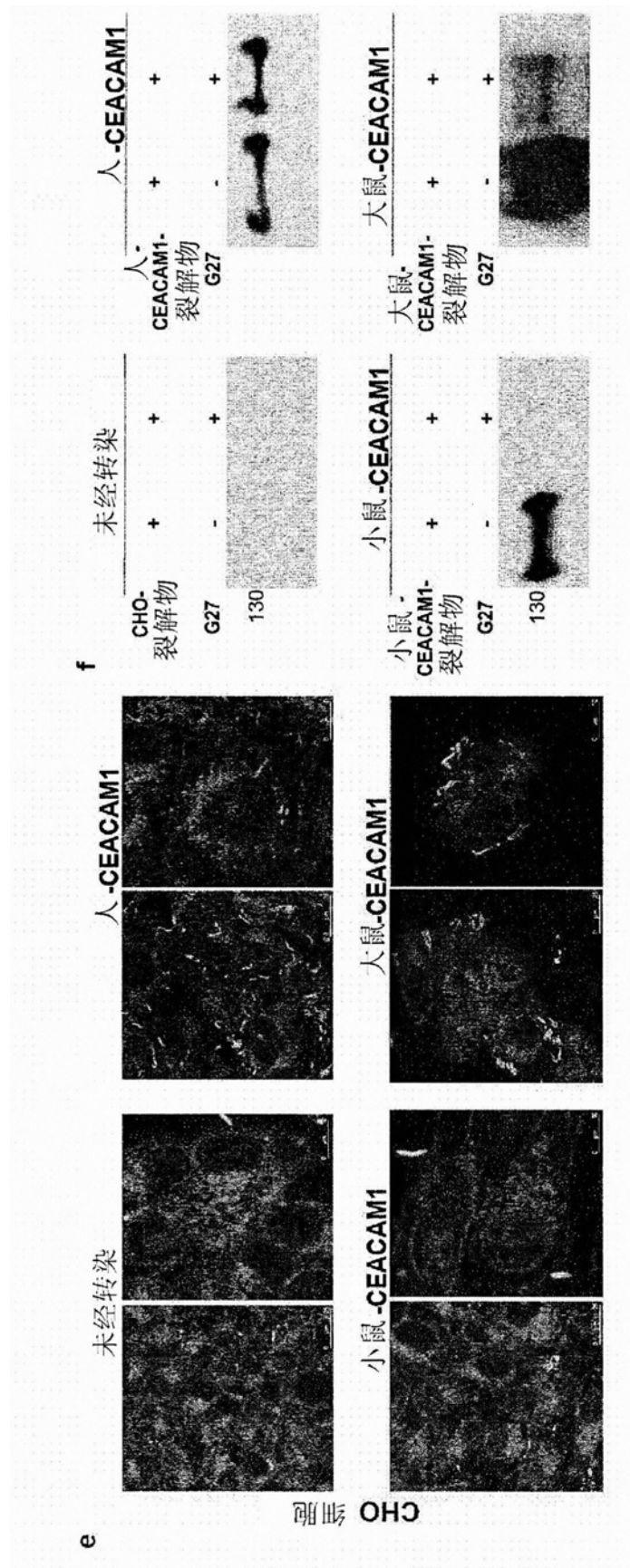


图3(续)

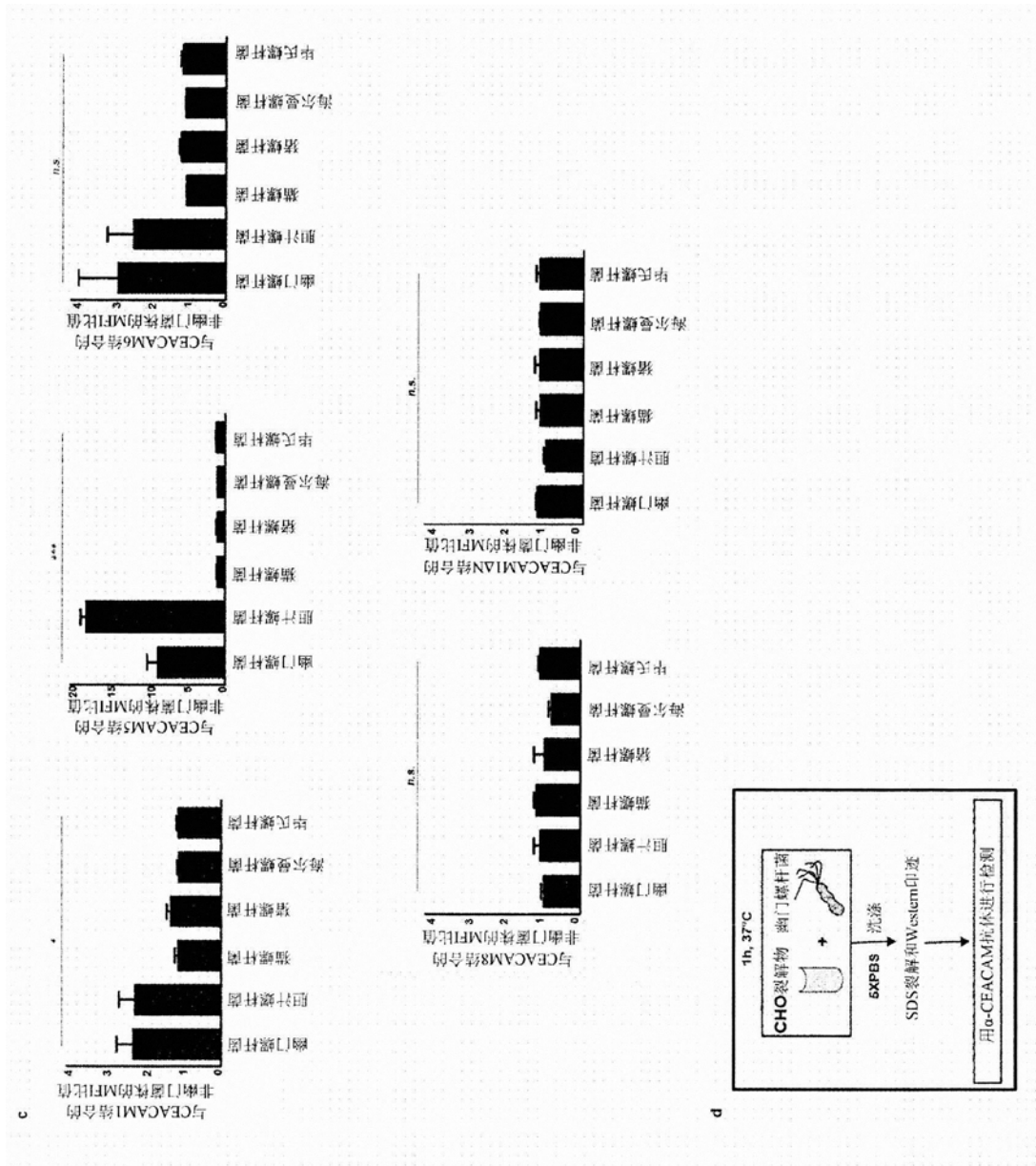


图4(续)

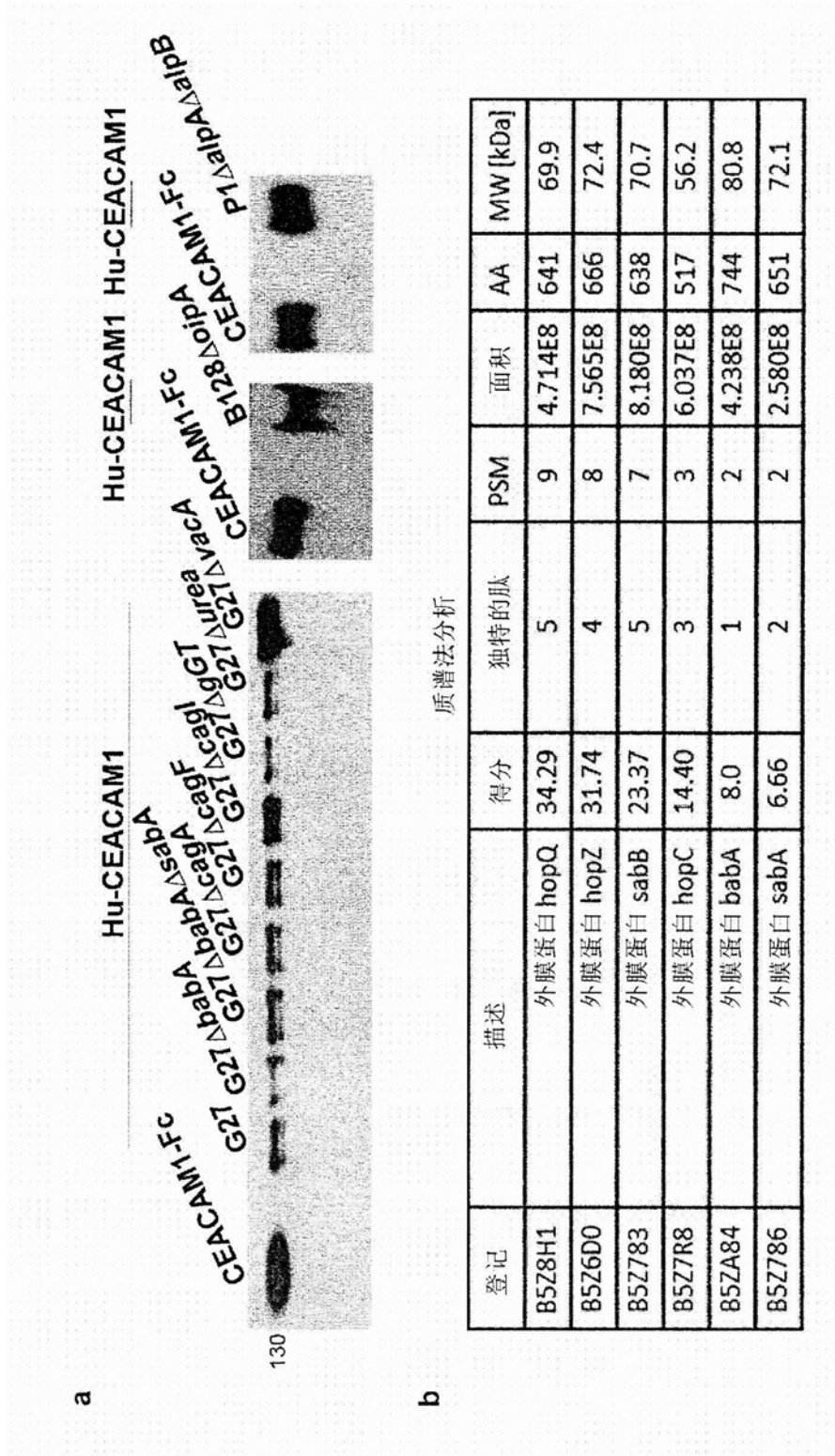


图5

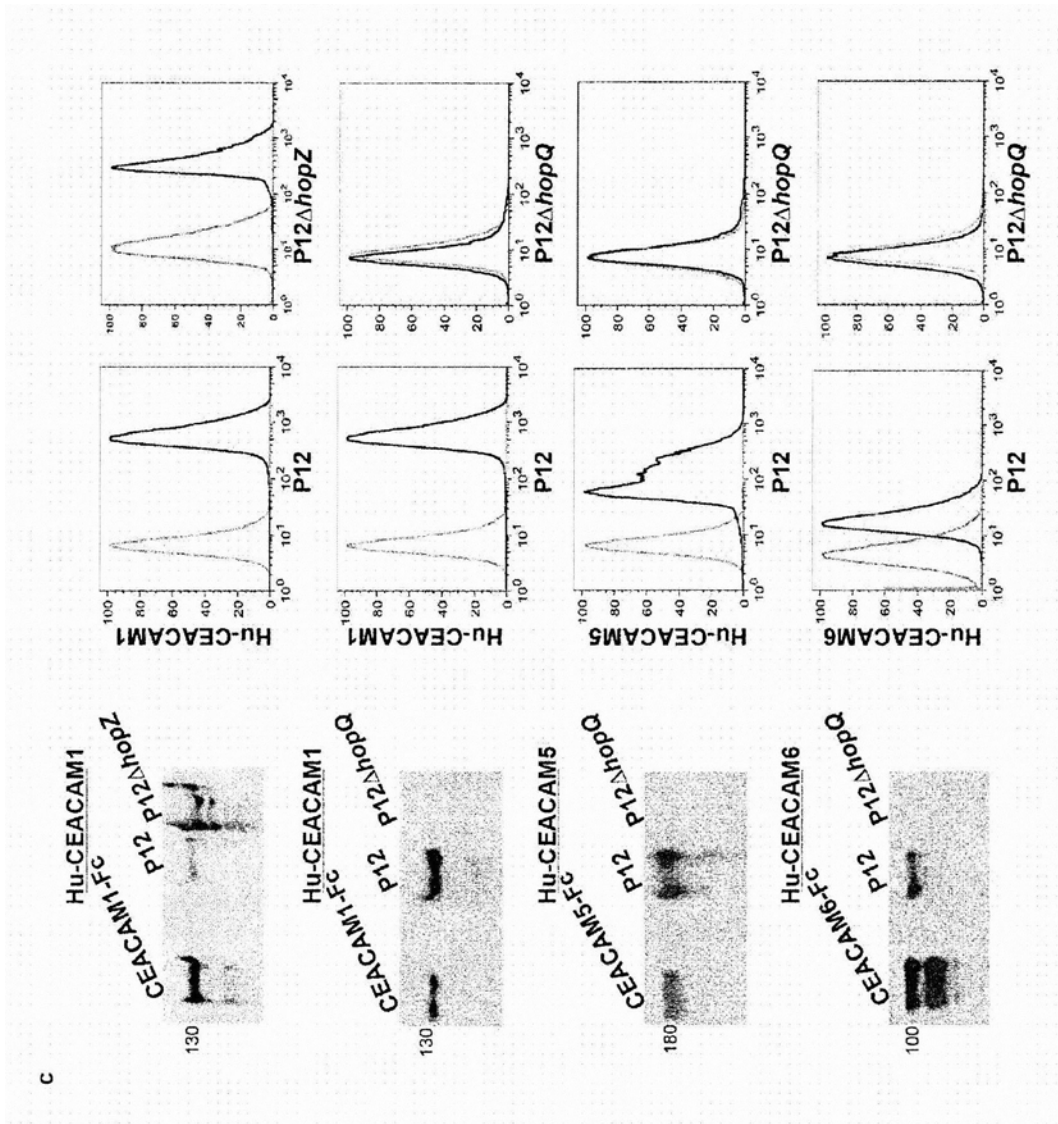


图5(续)

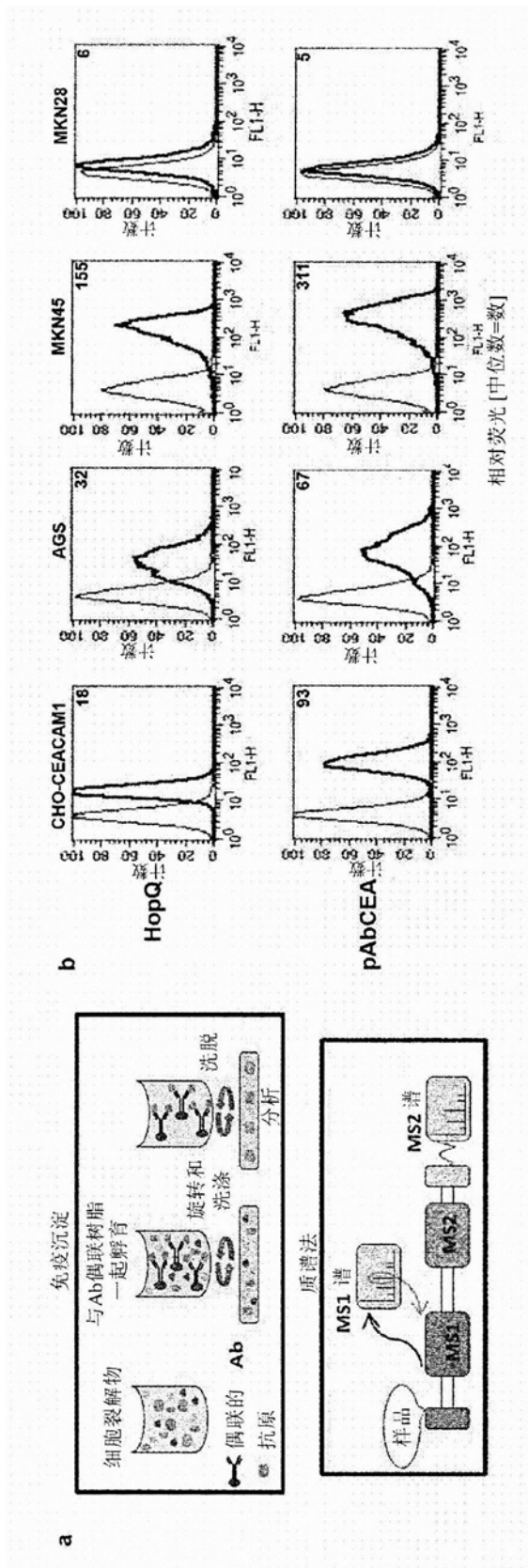


图6

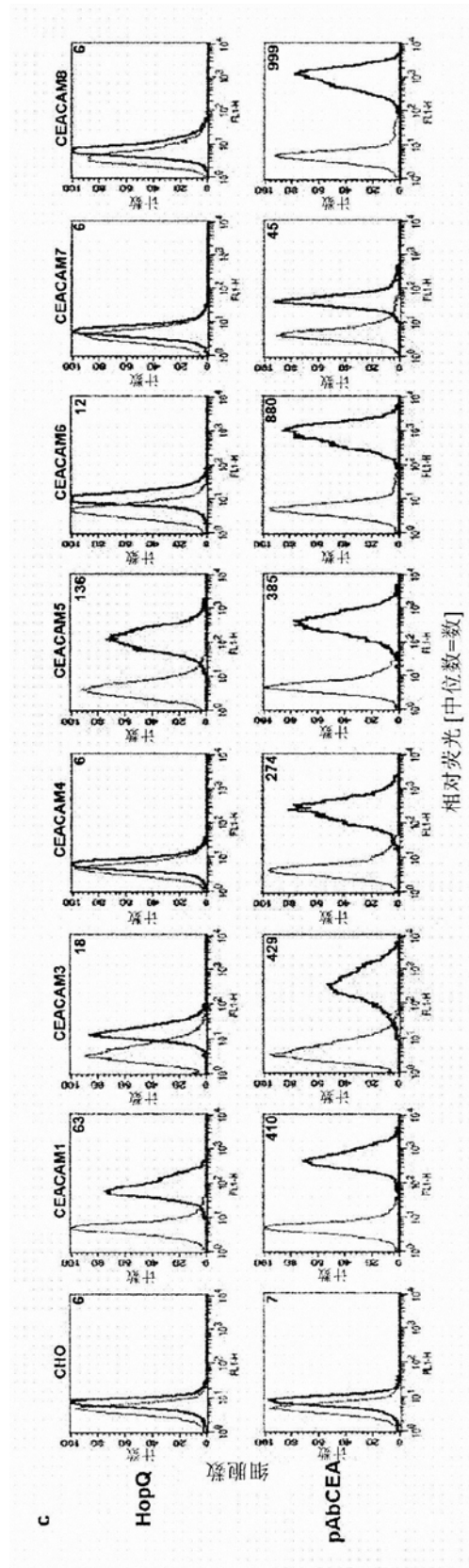


图6(续)

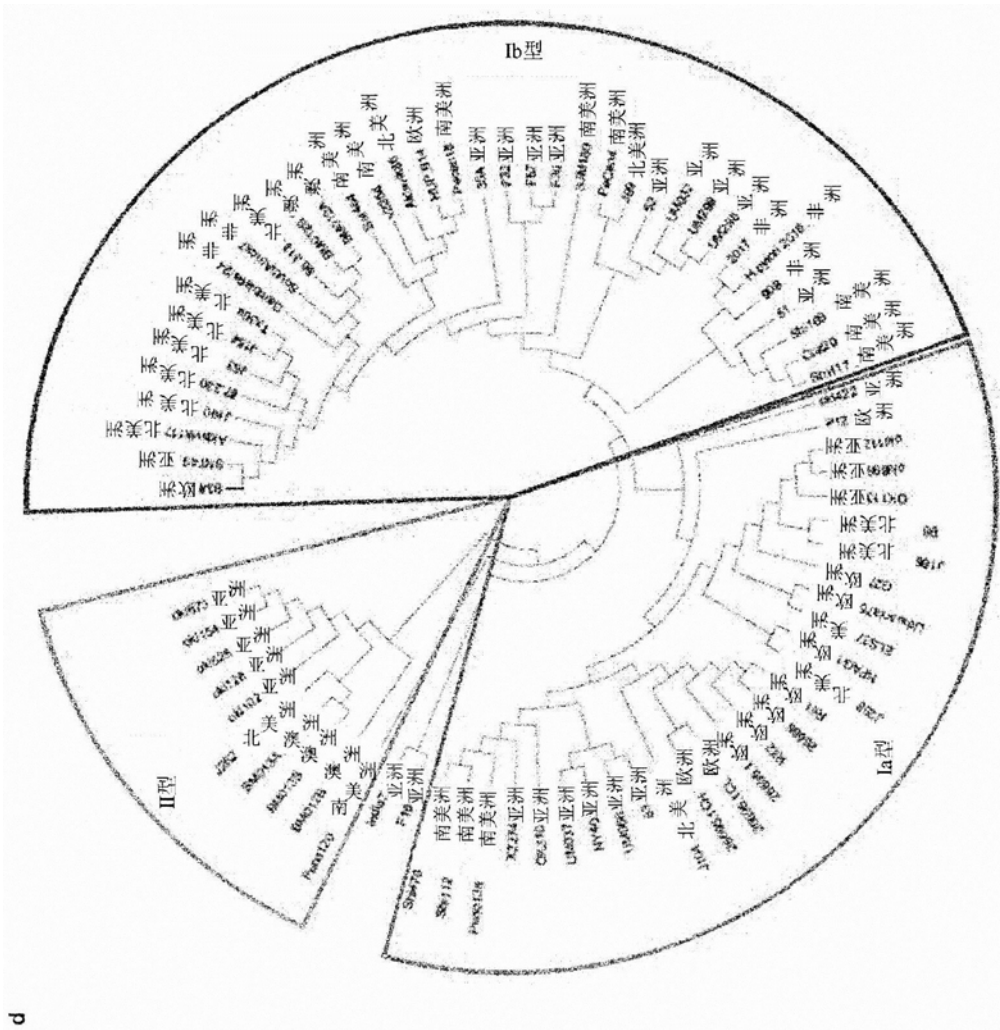


图6(续)

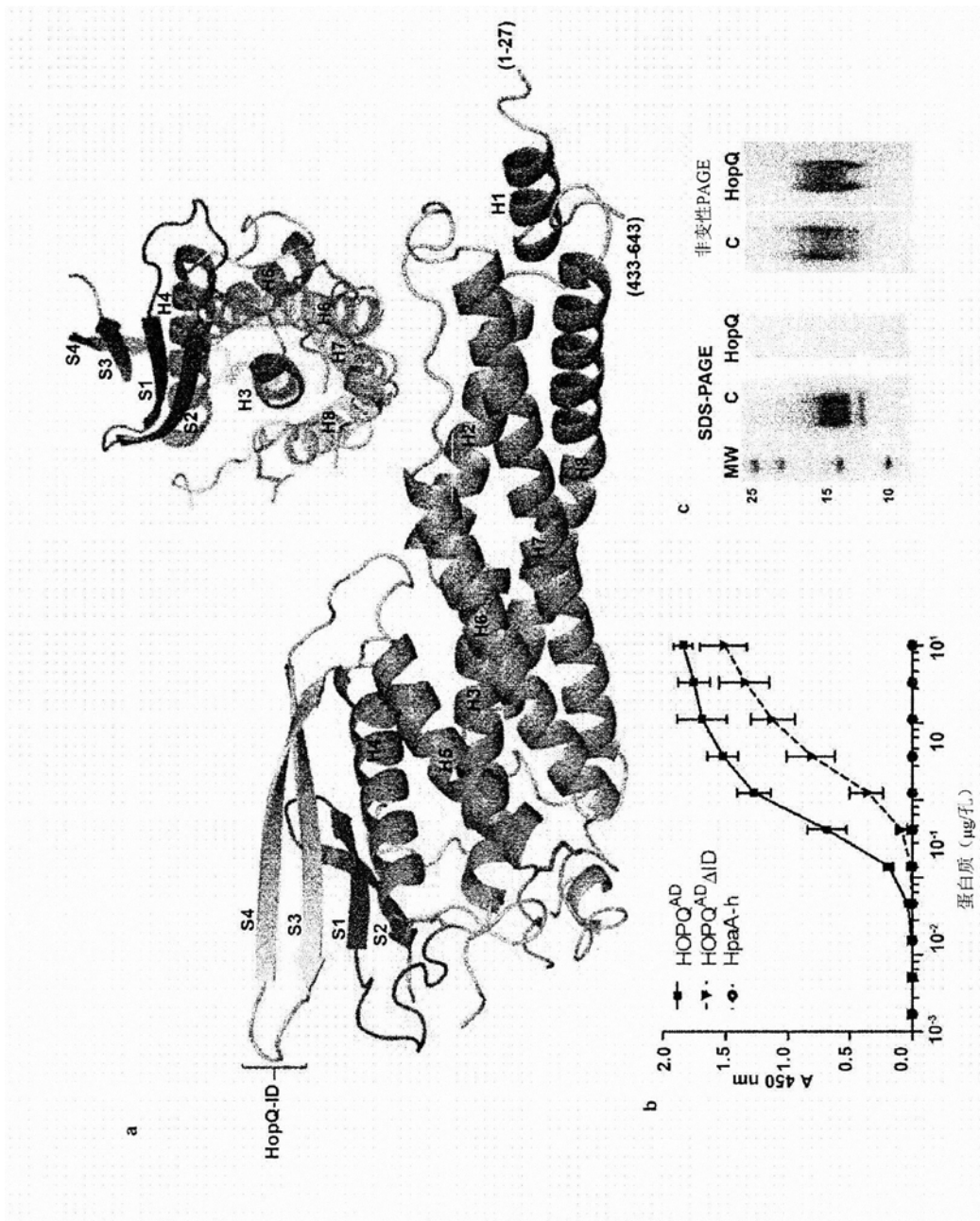


图7

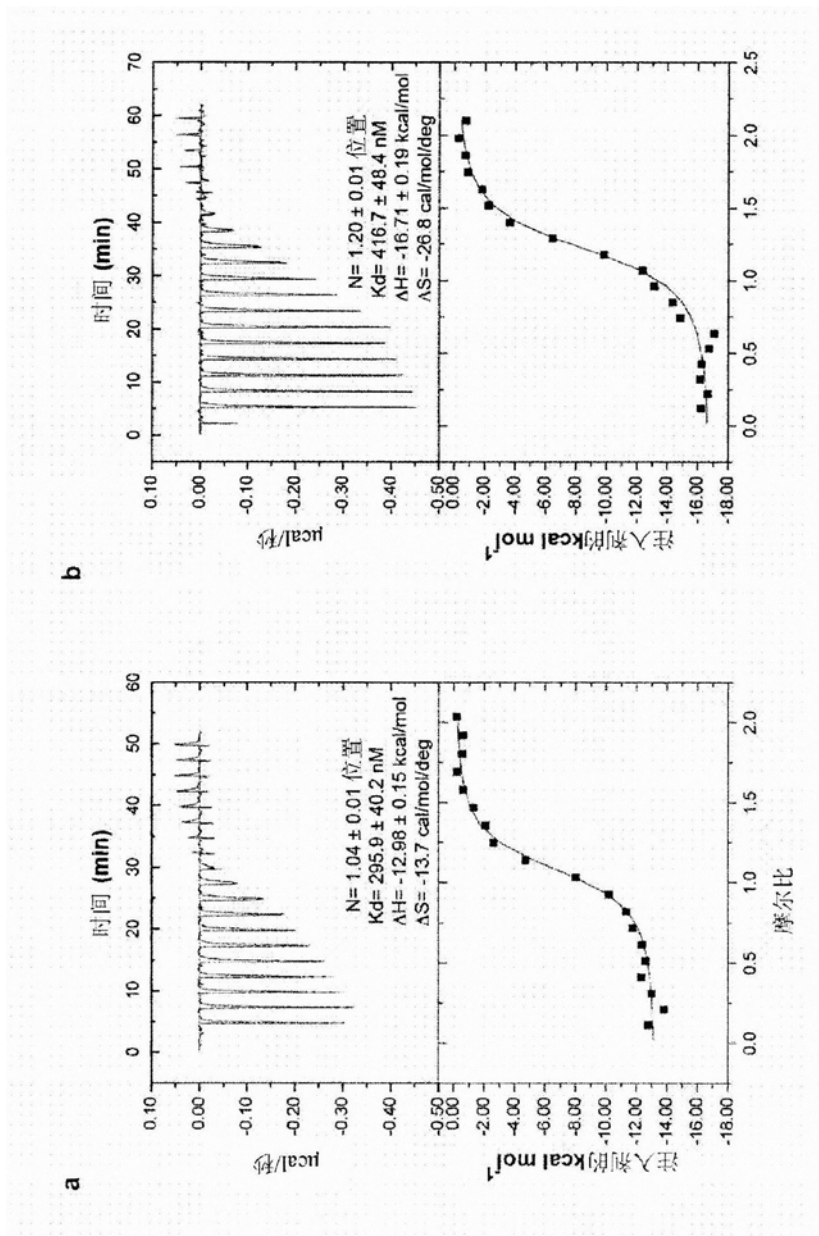


图8

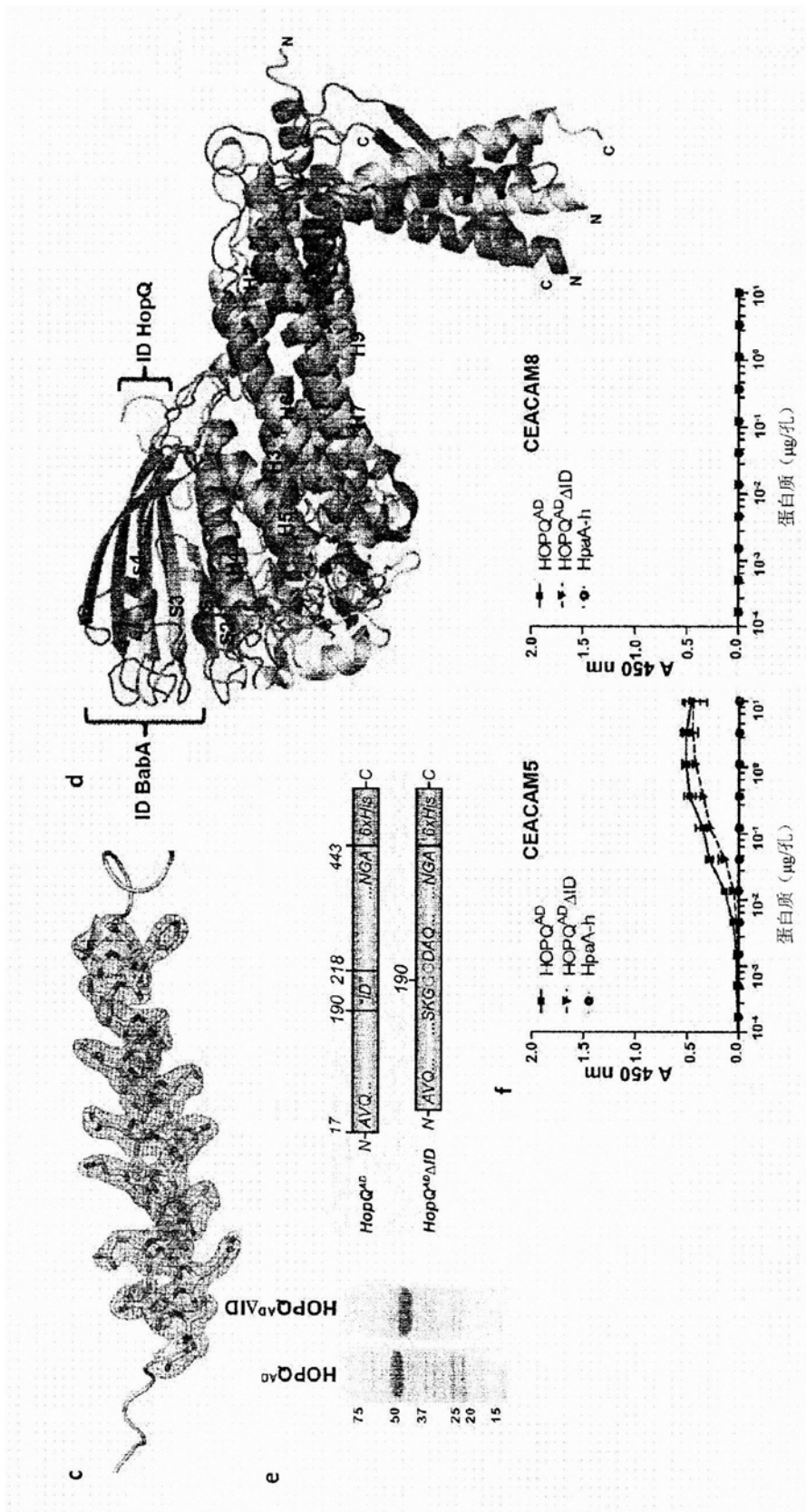


图8(续)

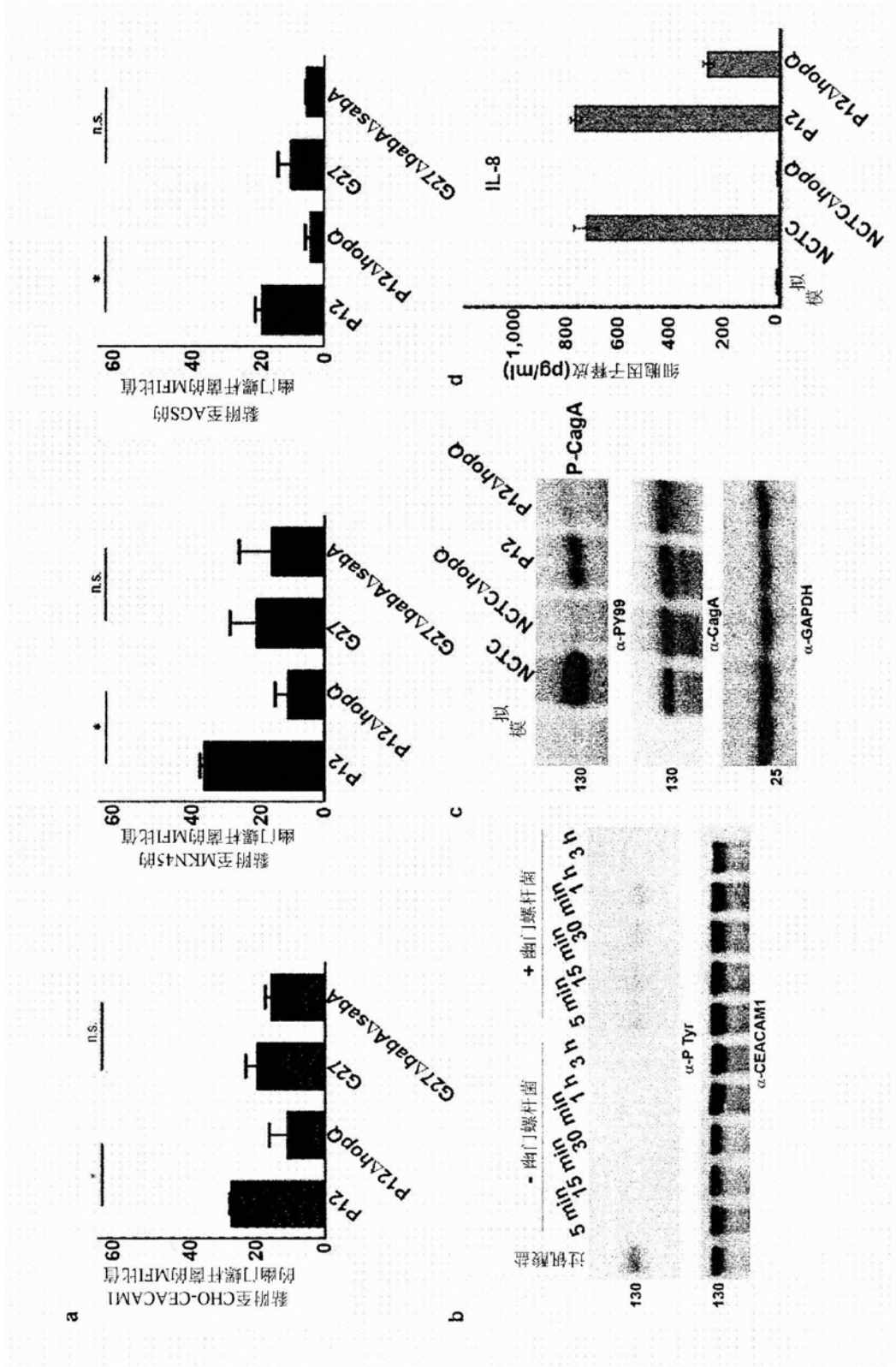


图9

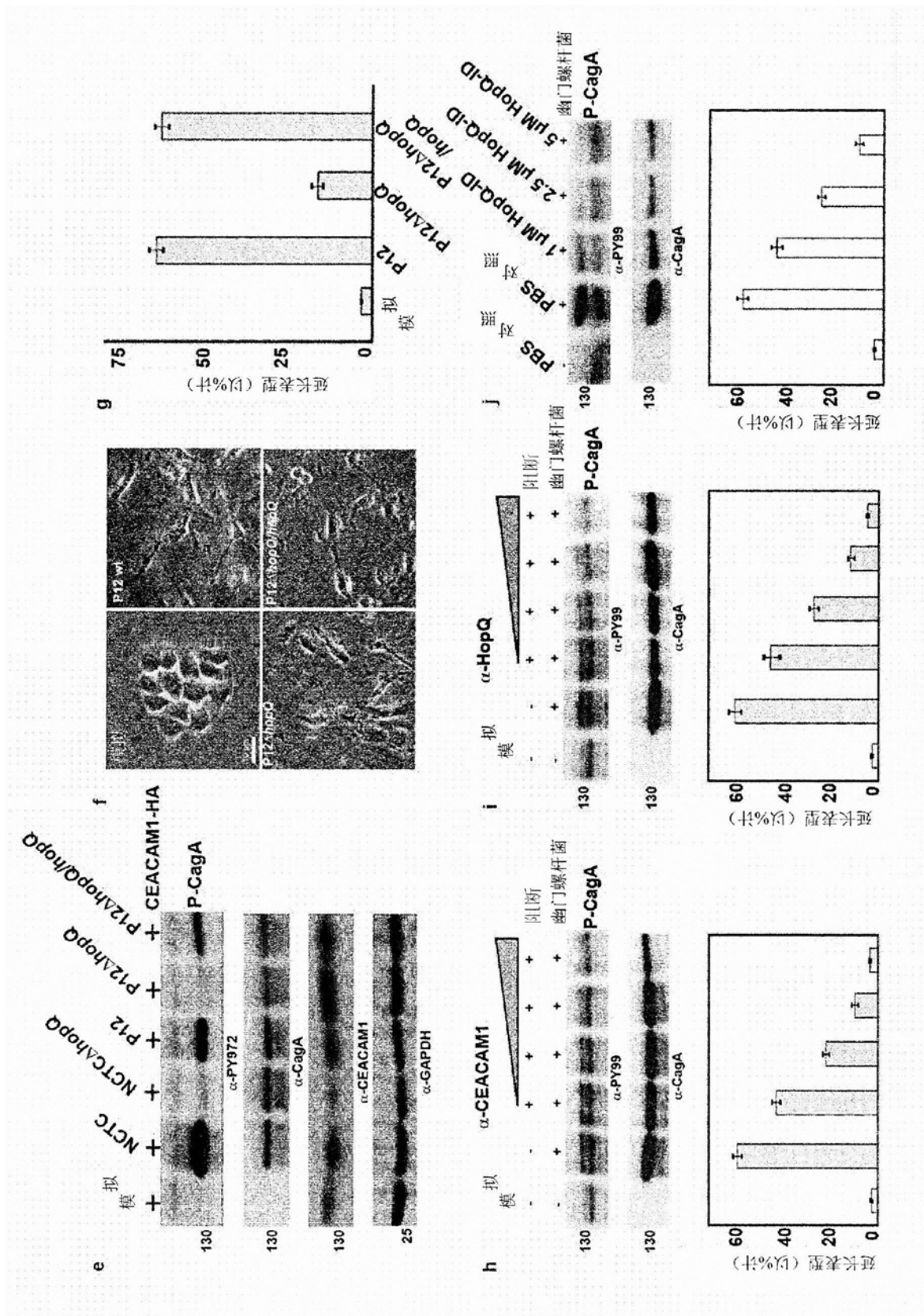


图9(续)

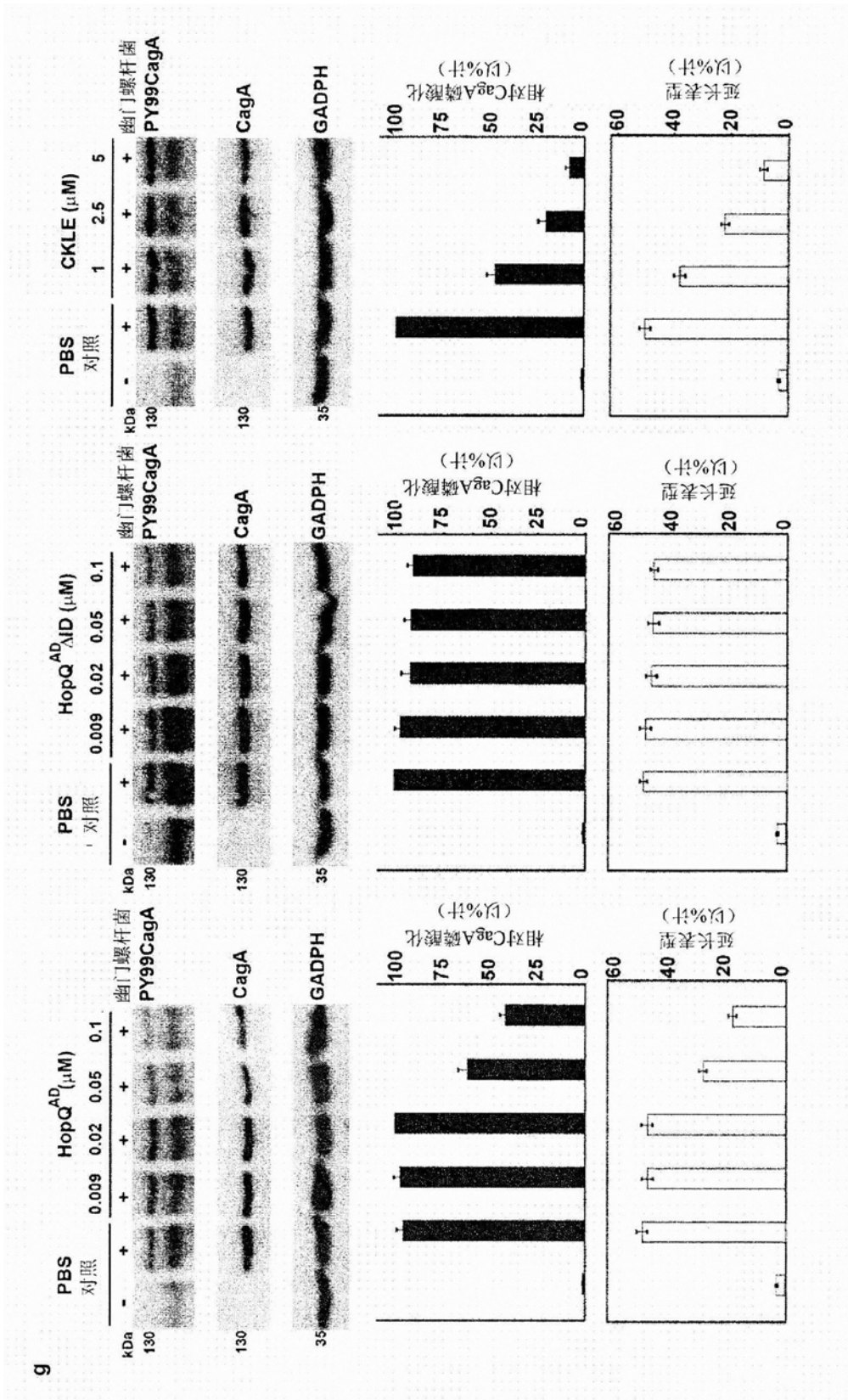


图10 (续)

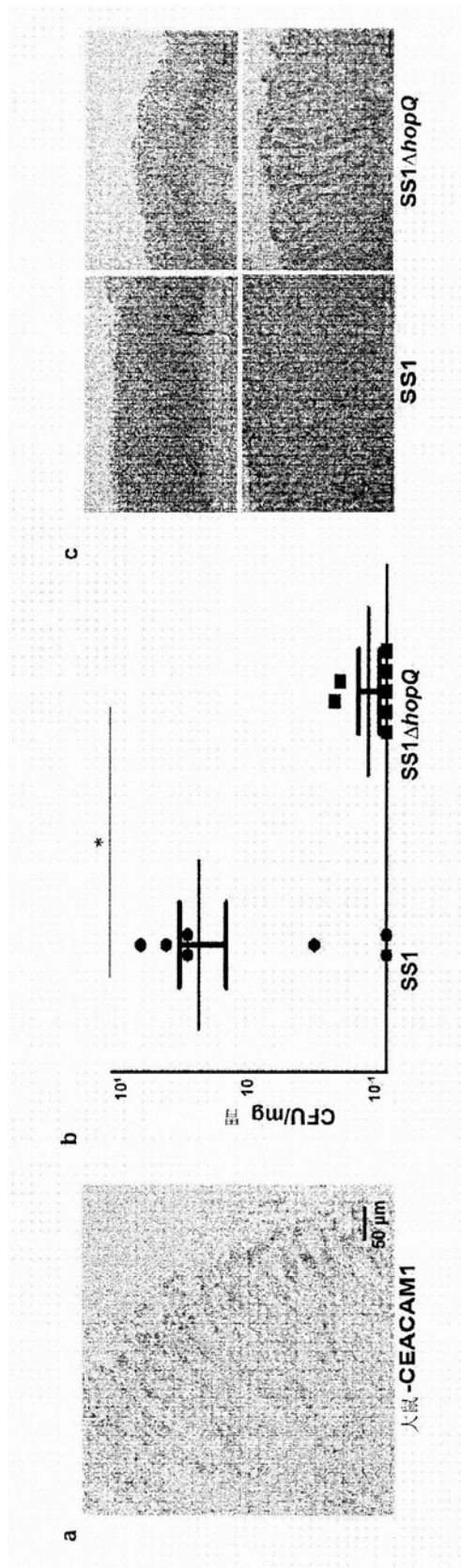


图11

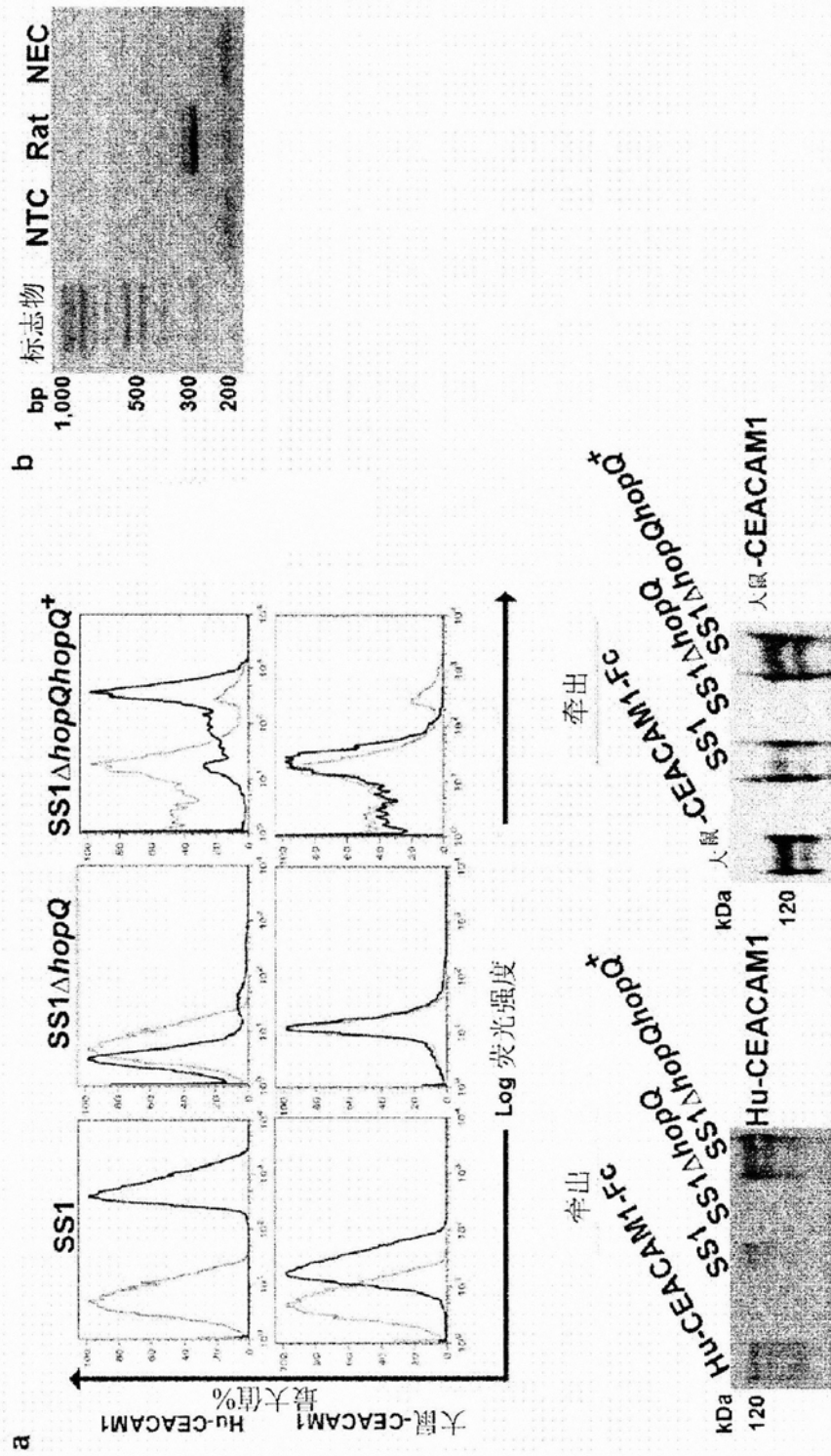


图12

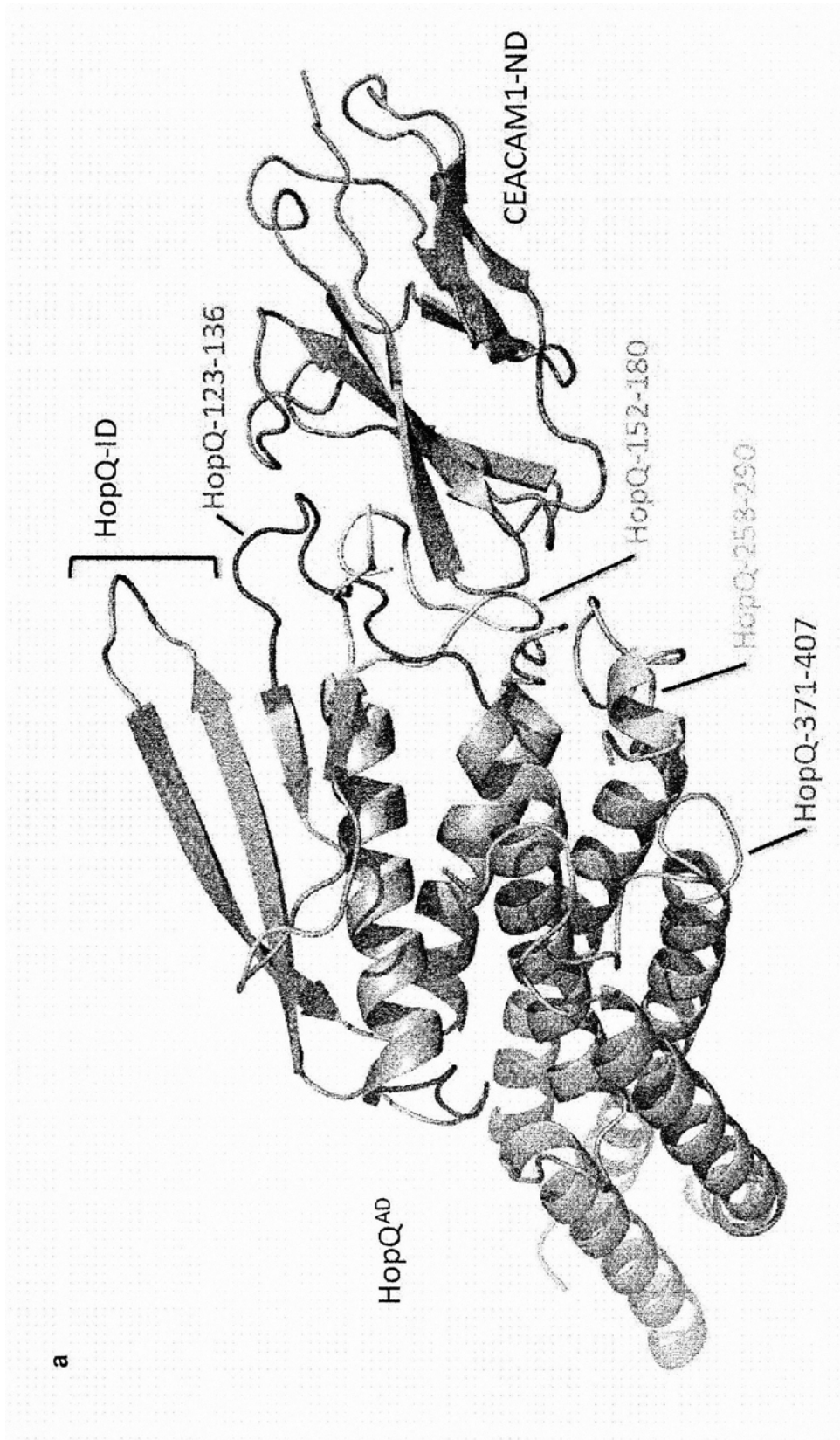


图13

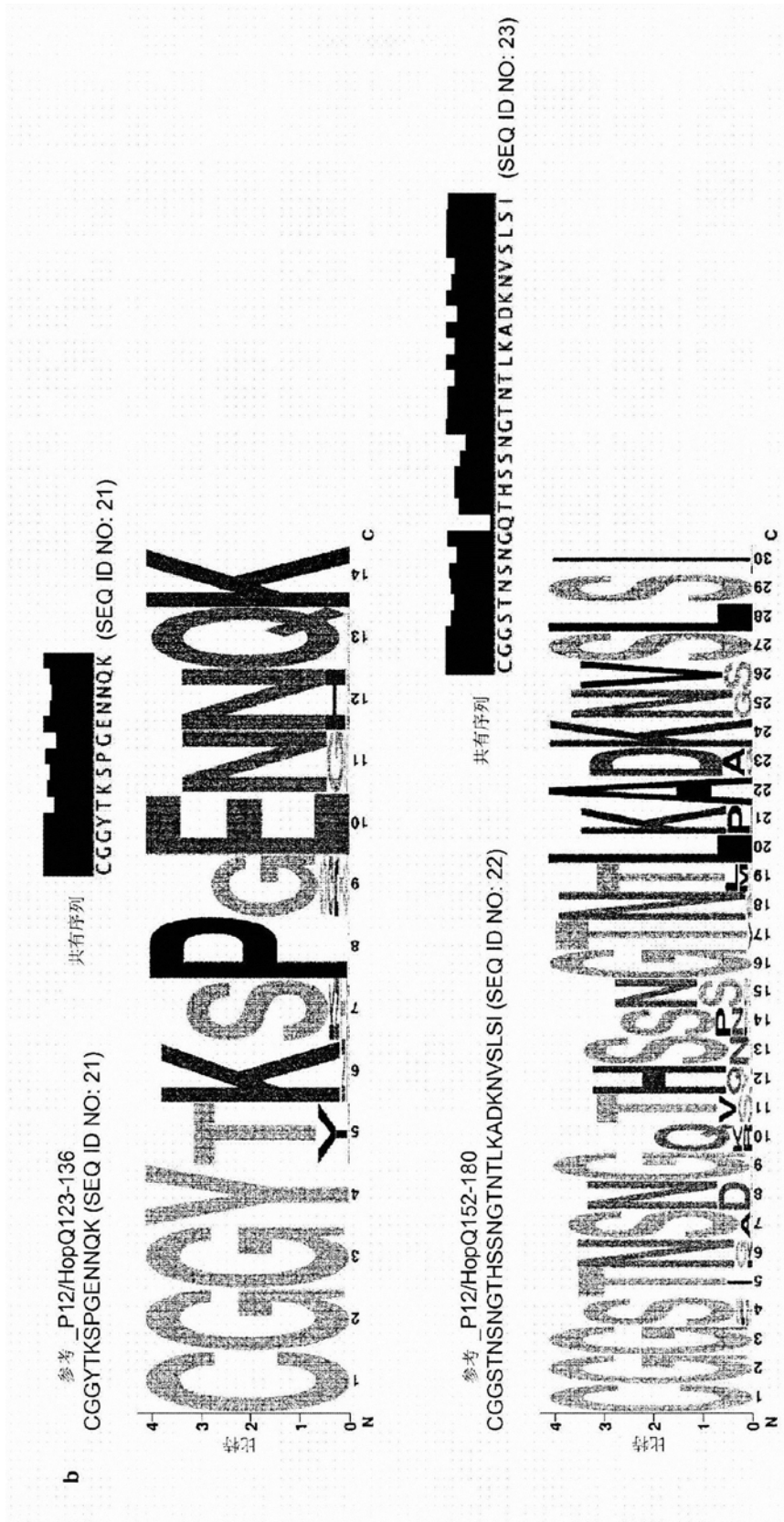


图13(续)

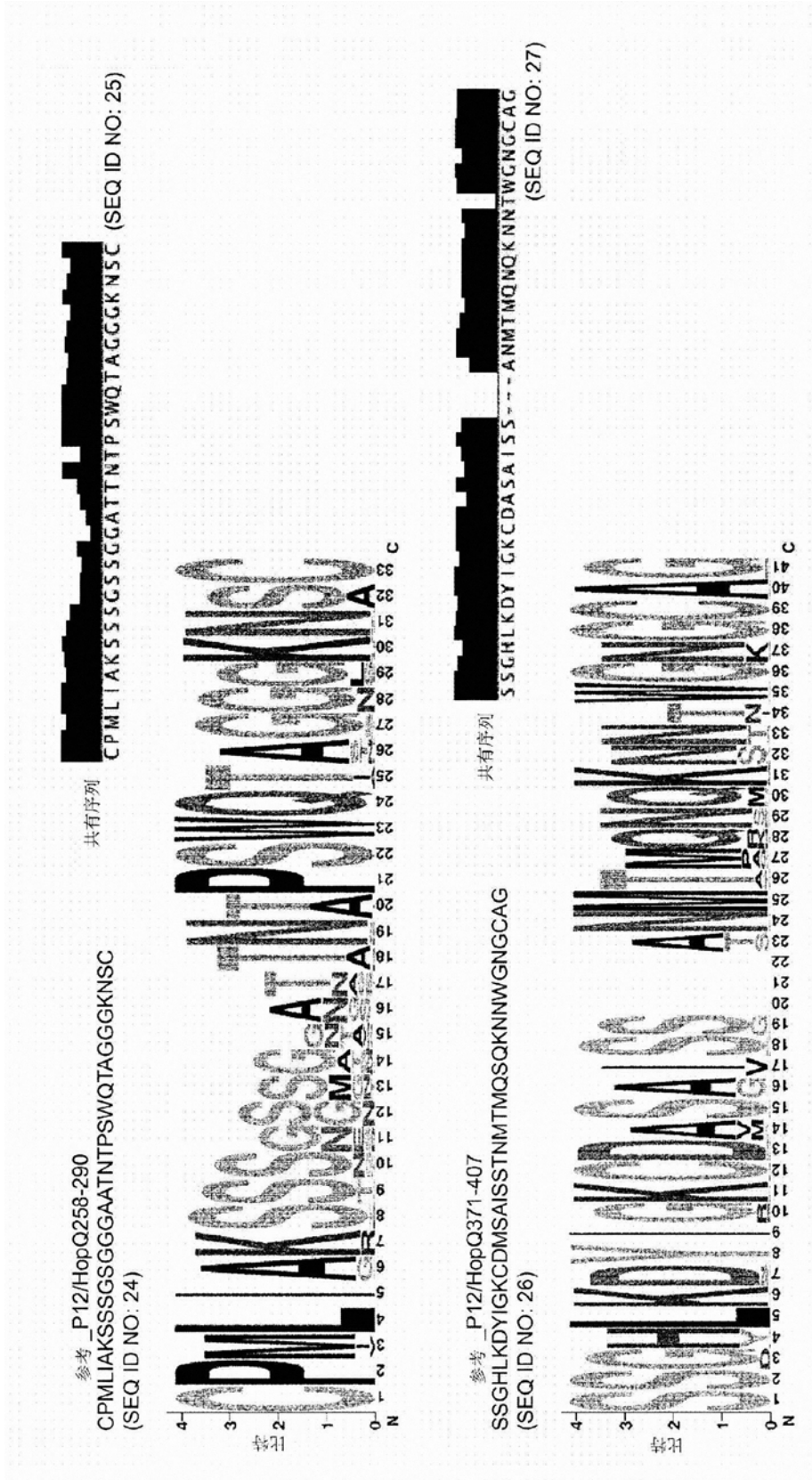


图13(续)