

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C07H 21/04



[12] 发 明 专 利 说 明 书

C12N 15/00 C12P 21/00

C07K 14/00 A61K 39/395

A61K 38/00 A61K 31/00

[21] ZL 专利号 96180284.7

[43] 授权公告日 2003 年 4 月 30 日

[11] 授权公告号 CN 1107072C

[22] 申请日 1996. 10. 31 [21] 申请号 96180284.7

[30] 优先权

[32] 1996. 3. 22 [33] US [31] 60/013,923

[86] 国际申请 PCT/US96/16966 1996. 10. 31

[87] 国际公布 WO97/34911 英 1997. 9. 25

[85] 进入国家阶段日期 1996. 11. 4

[71] 专利权人 人类基因组科学公司

地址 美国马里兰州

[72] 发明人 R·埃伯纳 S·M·鲁宾 余国良

[56] 参考文献

INTERNATIONAL IMMUNOLOGY 1993. VOL5, NO6. PAGES
691 - 694 1993 - 05 - 01 takeda et al. rapid acceleration
of neutrophil apoptosis by tumor necrosis factor - alpha

审查员 陈 真

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 黄革生

权利要求书 3 页 说明书 55 页 附图 15 页

[54] 发明名称 编程性细胞死亡诱导分子 II

[57] 摘要

本发明涉及 TNF—配基超家庭的一个新成员，编程性细胞死亡诱导分子 II (AIM II)。具体地说，本发明提供了编码人 AIM II 蛋白的分离的核酸分子。本发明也提供了 AIM II 多肽，以及生产该多肽的载体、宿主细胞和重组方法。本发明还涉及筛选用于鉴定 AIM II 活性激动剂和拮抗剂的方法。本发明还提供了治疗淋巴结病，自身免疫疾病、移植植物抗宿主疾病及抑制癌症如肿瘤细胞生长的治疗方法。

I S S N 1 0 0 8 - 4 2 7 4

1.一种分离的核酸分子，其含有与选自下组的序列具有至少 95 % 同一性的核苷酸序列的多核苷酸，所述组含有：

(a)编码具有图 1A - C 中全部氨基酸序列 (SEQ ID NO: 2) 的 AIM II 多肽的核苷酸序列；

(b)编码 AIM II 多肽的核苷酸序列，所述多肽具有 ATCC 保藏号 NO. 97689 所含 cDNA 克隆编码的全部氨基酸序列；

(c)编码 AIM II 多肽胞外结构域的核苷酸序列；

(d)编码 AIM II 多肽跨膜结构域的核苷酸序列；

(e)编码 AIM II 多肽胞内结构域的核苷酸序列；

(f)编码可溶性 AIM II 多肽的核苷酸序列，所述可溶性 AIM II 多肽含有胞外和胞内结构域，但缺乏跨膜结构域；和

(g)与上述(a)、(b)、(c)、(d)、(e)或(f)中任何一种核苷酸序列互补的核苷酸序列。

2.权利要求 1 的核酸分子，其中所说的多核苷酸具有图 1A - C 中全部的核苷酸序列 (SEQ ID NO: 1) 。

3.权利要求 1 的核酸分子，其中所说的多核苷酸具有图 1A - C (SEQ ID NO: 1) 中编码 AIM II 多肽的核苷酸序列，所述 AIM II 多肽具有图 1A - C 中全部氨基酸序列 (SEQ ID NO: 2) 。

4.权利要求 1 的核酸分子，其中所述多核苷酸具有 ATCC 保藏号 97689 中所含 cDNA 克隆的全部核苷酸序列。

5.权利要求 1 的核酸分子，其中所述多核苷酸具有编码 AIM II 多肽的核苷酸序列，该 AIM II 多肽具有 ATCC 保藏号 97689 中所含 cDNA 克隆编码的全部氨基酸序列。

6.一种分离的核酸分子，其含有的一段多核苷酸在严格杂交条件下可跟含有与权利要求 1 的(a)、(b)、(c)、(d)、(e)、(f)或(g)中一个核苷酸序列相同的核苷酸序列的多核苷酸杂交，其中所说的多核苷酸在严格杂交条件下与仅含 A 残基或仅含 T 残基的核苷酸序列不能杂交上。

7.一种分离的核酸分子，其含有编码 AIM II 多肽中具有表位的区段

的氨基酸序列的多核苷酸, 所述 AIM II 多肽具有权利要求 1 的(a)、(b)、(c)、(d、)(e)或(f)中的氨基酸序列。

8. 权利要求 7 的分离的核酸分子, 其编码 AIM II 多肽具有表位的区段, 所述区段选自下述的组中: 含有图 1A - C (SEQ ID NO: 2) 中大约第 13 位至大约第 20 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 23 位至大约第 36 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 69 位至大约第 79 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 85 位至大约第 94 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 167 位至大约第 178 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 184 位至大约第 196 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 221 至大约第 233 位氨基酸残基的多肽。

9. 一种制备重组载体的方法, 所述方法包括将权利要求 1 的分离的核酸分子插入至载体中。

10. 按权利要求 9 的方法制备的重组载体。

11. 一种制备重组宿主细胞的方法, 所述方法包括将权利要求 10 的重组载体导入宿主细胞中。

12. 按权利要求 11 的方法制备的重组宿主细胞。

13. 一种重组生产 AIM II 多肽的方法, 所述方法包括在所述多肽能表达的条件下培养权利要求 12 的重组宿主细胞; 以及回收所述多肽。

14. 一种分离的 AIM II 多肽, 其含有与选自下述组的一个序列具有至少 95 % 同一性的氨基酸序列, 该组含有:

(a) 具有图 1A - C (SEQ ID NO: 2) 中全部氨基酸序列的 AIM II 多肽的氨基酸序列;

(b) 具有 ATCC 保藏号 NO. 97689 所含 cDNA 克隆所编码的全部氨基酸序列的 AIM II 多肽的氨基酸序列;

(c) AIM II 多肽的胞外结构域的氨基酸序列;

(d) AIM II 肽的跨膜结构域的氨基酸序列;

(e) AIM II 多肽的胞内结构域的氨基酸序列;

(f) 可溶性 AIM II 多肽的氨基酸序列, 所述可溶性 AIM II 多肽具有

全部或部分胞外和胞内结构域，但缺乏跨膜结构域；和

(g) (a)、(b)、(c)、(d)、(e)或(f)中任一种多肽的含有表位的区段的氨基酸序列。

15.一种含有 AIM II 蛋白中具有表位的区段的分离的多肽，其中所述区段选自由下述组成之组：含有图 1A - C (SEQ ID NO: 2) 中大约第 13 位至大约第 20 位氨基酸残基的多肽；含有图 1A - C (SEQ ID NO: 2) 中大约第 23 位至大约第 36 位氨基酸残基的多肽；含有图 1A - C (SEQ ID NO: 2) 中大约第 69 位至大约第 79 位氨基酸残基的多肽；含有图 1A - C (SEQ ID NO: 2) 中大约第 85 位至大约第 94 位氨基酸残基的多肽；含有图 1A - C (SEQ ID NO: 2) 中大约第 167 位至大约第 178 位氨基酸残基的多肽；含有图 1A - C (SEQ ID NO: 2) 中大约第 184 位至大约第 196 位氨基酸残基的多肽；以及含有图 1A - C (SEQ ID NO: 2) 中大约第 221 位至大约第 233 位氨基酸残基的多肽。

16.一种分离的抗体，其能特异性结合权利要求 14 的 AIM II 多肽。

17.权利要求 14 的(a)、(b)、(c)、(d)、(e)或(f)中的多肽在制备可用于治疗选自由下述组成之组的状态或疾病的药物组合物中的用途：淋巴结病、自身免疫疾病和移植物抗宿主疾病。

18.权利要求 14 的(a)、(b)、(c)、(d)、(e)或(f)中的多肽在制备可用于抑制瘤形成的药物组合物中的用途。

19. 权利要求 14 的多肽的拮抗物在制备可用于预防选自由下述组成之组的状态或疾病的药物组合物中的用途：败血病性休克、感染、脑型疟、HIV - 病毒活化、骨吸收、类风湿性关节炎和恶病质。

编程性细胞死亡诱导分子 II

发明领域

本发明涉及 TNF - 配基超家族的一个新成员。具体地说, 本发明提供了编码人编程性细胞死亡诱导分子 II (Apoptosis Inducing Molecular II, AIM II) 的分离的核酸分子。本发明还提供了 AIM II 多肽, 以及生产该多肽的载体、宿主细胞和重组方法。本发明还涉及鉴定 AIM II 活性的激动剂和拮抗剂的筛选方法。本发明也提供了治疗淋巴结病、自身免疫疾病、移植物抗宿主疾病以及抑制肿瘤如肿瘤细胞生长的治疗方法。

背景技术

人肿瘤坏死因子 α (TNF - α) 和 β (TNF - β , 或淋巴细胞毒素) 是一个多肽介导分子大类中的相关成员, 该类多肽介导分子包括干扰素、白细胞介素和生长因子, 总称为细胞因子 (Beutler B. 和 Cerami A., Annu. Ret. Immunol., 7:625-655 (1989))。

肿瘤坏死因子 (TNF - α 和 TNF - β) 最初是由于其抗肿瘤活性效果而被发现的, 然而目前它被看作是一种多效细胞因子, 具有多种生物学活性, 包括参与一些转化细胞系的编程性细胞死亡、介导细胞的活化和增殖, 并在免疫调控和炎症中也有重要作用。

迄今为止, TNF - 配基超家族的已知成员有: TNF - α 、TNF - β (淋巴细胞毒素 - α)、LT - β 、OX40L、Fax 配基、CD30L、CD27L、CD40L 以及 4 - 1BBL。TNF 配基超家族的配基是酸性的、类似 TNF 的分子, 在胞外结构域有大约 20 % (范围, 12 % - 36 %) 的序列同源性, 并主要以膜结合形式存在, 其具生物活性的类型是三聚体/多聚体复合物。目前仅从 TNF、LT α 和 Fas 配基鉴定出 TNF 配基超家族的可溶性形式(一般综述见 Gruss H. 和 Dower S. K., 血液, 85 (12): 3378-3404 (1995))。此处全文引入此文献作为参考。

这些蛋白质参与细胞增殖、活化和分化过程的调控, 包括通过编程

性细胞死亡或细胞毒性控制细胞的存活或死亡 (Armitage R. J., Curr. Opin. Immunol., 6: 407 (1994)及 Smith C. A., 细胞, 75: 959 (1994))。

哺乳动物的发育依赖于细胞的增殖和分化, 以及在编程性细胞死亡过程中发生的细胞程序性死亡 (Walker 等, Methods Achiev. Exp. Pathol., 13:18 (1988)。编程性细胞死亡在识别自身抗原的免疫胸腺细胞的破坏过程中起着非常重要的作用。正常排除过程的失效可能在自身免疫疾病中具有一定作用 (Gammon 等, 今日免疫学, 12: 193 (1991))。

Itoh 等 (细胞, 66:233 (1991)) 描述了一种细胞表面抗原 - Fas/ CD 95。它介导编程性细胞死亡, 参与 T - 细胞的克隆缺失。Fas 在已活化的 T - 细胞、B - 细胞和嗜中性白细胞以及胸腺、肝、心脏、肺中都有表达。Fas 除了在已活化的 T - 细胞、B - 细胞和嗜中性白细胞中表达外, 它还在成年小鼠卵巢中表达 (Watanabe-Fukunaga 等, 免疫学杂志, 148: 1274 (1992))。在将 Fas 的单克隆抗体与 Fas 交联的实验中, 会引发编程性细胞死亡 (Yonehara 等, 实验医学杂志, 169: 1747 (1989); Trauth 等, 科学, 245: 301 (1989))。另外, 有例子表明, 单克隆抗体与 Fas 的结合有可能在一定条件下刺激 T - 细胞 (Alderson 等, 实验医学杂志, 178: 2231 (1993))。

Fas 抗原是一种细胞表面蛋白, 共相对分子量为 45kd。人和鼠的 Fas 基因已由 Watanabe-Fukunaga 等 (免疫学杂志, 148: 1274 (1992) 和 Itoh 等 (细胞, 66: 233 (1991)) 克隆。两种基因所编码的蛋白质均为跨膜蛋白, 与神经生长因子/肿瘤坏死因子受体超家族具有结构上的同源性, 所述超家族包括两种 TNF 受体、低亲和性神经生长因子受体及 LT β 受体 CD40、CD27、CD30 和 OX40。

近来 Fas 配基已有描述 (Suda 等, 细胞 75: 1169 (1993))。其氨基酸序列表明 Fas 配基是属于 TNF 家族的 II 型跨膜蛋白。Fas 配基在脾细胞和胸腺细胞中表达, 经纯化的 Fas 配基分子量为 40kd。

近来, 已经证明 Fas/Fas 配基间的相互作用是 T 细胞激活后的编程性细胞死亡过程所必需的 (Ju 等, 自然, 373: 444 (1995); Brunner 等, 自然, 373: 441 (1995))。T 细胞的活化可诱导细胞表面上的这两种蛋白质, 随后配基与受体间的相互作用导致了细胞的编程性细胞死

亡。由此证明,在正常免疫反应中, Fas/Fas 配基相互作用所诱导的编程性细胞死亡过程可能具有一定的调控作用。

根据结构和生物学上的相似性,本发明的多肽已被鉴定为 TNF 配基超家族的一个新成员。

显然,应该存在一些因子,它们可调控正常和异常细胞的活化和分化。因此,有必要鉴定和表征在正常和疾病状态下调控正常和异常细胞的活化和分化的这类因子。具体地说,有必要分离和鉴定控制编程性细胞死亡以治疗自身免疫疾病、移植物抗宿主疾病及淋巴结病的其它的 Fas 配基。

发明概述

本发明提供了含有编码 AIM II 多肽的多核苷酸的分离的核酸分子,所述 AIM II 多肽具有图 1A - C (SEQ ID NO: 2) 所示氨基酸序列或由保藏于细菌宿主的 cDNA 克隆所编码的氨基酸序列,该细菌宿主于 1996 年 8 月 22 日以 ATCC 保藏号 97689 进行了保藏。

本发明也涉及含有本发明的分离核酸分子的重组载体、含有所述重组载体的宿主细胞,以及通过重组技术制备所述载体和宿主细胞并利用它们生产 AIM II 多肽或肽的方法。

本发明还提供了具有由本文所述多核苷酸编码的氨基酸序列的分离的 AIM II 多肽。

本文所用术语“AIM II”多肽包括保留了 AIM II 多肽活性的膜结合蛋白质(含有一个细胞质结构域、一个跨膜结构域,以及一个胞外结构域)及其截短的蛋白质。在一个实施方案中,可溶性 AIM II 多肽含有 AIM II 蛋白全部或部分胞外结构域,但缺乏能将所述多肽保留于细胞膜上的跨膜区。只要可溶性 AIM II 蛋白能分泌出来,它也可含有部分跨膜区或部分胞质结构域或其它序列。可将一段异源信号肽融合于可溶性 AIM II 多肽的 N - 末端,以使该可溶性 AIM II 多肽在表达时能分泌出来。

本发明也提供了可用于治疗淋巴结病、自身免疫疾病及移植物抗宿主疾病的 AIM II 多肽,特别是人 AIM 多肽,所述 AIM II 多肽可用于刺激周边耐受性、破坏一些转化细胞系、介导细胞活化和增殖,在功能上

也可作为免疫调节和炎症反应的初级介导分子。

本发明还提供了含有 AIM II 多核苷酸或 AIM II 多肽的组合物，所述组合物可施用于体外细胞、离体（*ex vivo*）细胞和体内细胞，或可施用于多细胞生物体。在本发明此方面某些特别优选的实施方案中，该组合物中含有 AIM II 多核苷酸，以便在宿主生物中表达用于治疗疾病的 AIM II 多肽。就此而言，特别优选在病人体内进行表达，以便治疗与 AIM II 内源活性异常相关的疾病。

本发明也提供了鉴定可增强或抑制由 AIM II 所诱导的细胞反应的化合物的筛选方法，该筛选方法包括将表达 AIM II 的细胞与候选化合物接触、测试细胞反应以及将该细胞反应与标准细胞反应相比较，所述标准是在无该候选化合物的接触下测定的；由此，细胞反应比标准增强则表明该化合物是激动剂，细胞反应比标准降低则表明该化合物是拮抗剂。

另一方面，本发明提供了 AIM II 激动剂和拮抗剂的筛选试验。拮抗剂可用于预防败血症性休克、炎症、脑型疟、HIV 病毒的活化、移植物-宿主排斥反应、骨吸收、类风湿性关节炎及恶病质（消瘦或营养不良）。

本发明另一方面涉及需要更高水平的体内 AIM II 活性的个体的治疗方法，所述方法包括对这类个体施用含有治疗有效剂量的本发明的分离 AIM II 多肽或其激动剂的组合物。

本发明另一方面还涉及需要更低水平的体内 AIM II 活性的个体的治疗方法，所述方法包括对这类个体施用含有治疗有效剂量的 AIM II 拮抗剂的组合物。

附图简述

图 1A - C 示 AIM II 的核苷酸序列（SEQ ID NO：1）及推断的氨基酸序列（SEQ ID NO：2）。该蛋白质推断的分子量大约为 26.4kDa。AIM II 蛋白推测的跨膜结构域用下划线表示。

图 2A - F 示 AIM II 蛋白和人 TNF- α （SEQ ID NO：3）、人 TNF- β （SEQ ID NO：4）、人淋巴细胞毒素（SEQ ID NO：5）及人 Fas 配基（SEQ ID NO：6）的氨基酸序列之间的相似区。

图 3A - F 示 AIM II 氨基酸序列分析。图中示意了 α 、 β 、转角和卷曲区；亲水性和疏水性；两亲区、柔性区；抗原性指数和表面可能性。在“抗原性指数 - Jameson-Wolf”图表中，图 1A - C 中大约第 13 - 20、23 - 36、69 - 79、85 - 94、167 - 178、184 - 196、221 - 233 位氨基酸对应于 AIM II 蛋白中显示出高抗原性的区域。

发明详述

本发明提供了含有编码 AIM II 多肽的多核苷酸的分离核酸分子，该核酸分子通过对克隆的 cDNA 进行测序而测定，所述 AIM II 多肽具有图 1A - C 所示氨基酸序列 (SEQ ID NO: 2)。本发明的 AIM II 蛋白与人 TNF - α (SEQ ID NO: 3)、人 TNF - β (SEQ ID NO: 4)、人淋巴细胞毒素 (SEQ ID NO: 5) 和人 Fas 配基 (SEQ ID NO: 6) 有序列同源性 (图 2A - F)。图 1A - C 中所示核苷酸序列得自于 cDNA 克隆的测序结果，所述 cDNA 克隆于 1996 年 8 月 22 日保藏于美国典型培养物保藏中心，12301 Park Lawn Drive, Rockville, Maryland 20852, 保藏号为 97689。该保藏克隆包含于 pBluescript SK (-) 质粒 (Stratagene, La Jolla, CA) 中。

除非特别指出，本文由 DNA 分子测序确定的所有核苷酸序列都是利用自动化 DNA 测序仪 (如 Applied Biosystems, Inc. 的 373 型) 测定的，由本文所测定的 DNA 分子所编码的多肽的所有氨基酸序列都是对上述所测定的 DNA 序列进行翻译而推测的。因此，如同本领域对由此自动化途径测定的任何 DNA 序列所知的那样，本文所测定的任何核苷酸序列都可能存在一些差错。通过自动化所测定的核苷酸序列与测序 DNA 分子的真实核苷酸序列之间通常至少有约 90 % 同一性、更通常是至少有约 95 % 至至少约 99.9 % 同一性。通过其它途径可更准确地测定真实序列，这些途径包括本领域所熟知的手工 DNA 测序。本领域亦知，与真实序列相比，在测定的核苷酸序列内单一插入或缺失可造成核苷酸序列翻译时的移框，因此导致由测定的核苷酸序列所编码的推测氨基酸序列与由测序 DNA 分子真实编码的氨基酸序列从这类插入或缺失点处开始完全不同。

利用本文所提供的信息，如图 1A - C 中的核苷酸序列，通过标准的克隆和筛选方法，如使用 mRNA 作为起始材料克隆 cDNA 时所使用的

方法,有可能获得编码 AIM II 多肽的本发明的核酸分子。如本发明所述,图 1A - C 中所述核酸分子 (SEQ ID NO: 1) 发现于来自人巨噬细胞 ox LDL (HMCCB64) 的 cDNA 文库中。在来自活化 T 细胞 (HT4CC72) 的 cDNA 文库中也鉴定到此基因。图 1A - C 中 AIM II cDNA 测出的核苷酸序列 (SEQ ID NO: 1) 含有编码 240 个氨基酸残基的蛋白质的开放阅读框,其具有的起始密码子位于图 1A - C (SEQ ID NO: 1) 中核苷酸序列第 49 - 51 位,胞外结构域含有图 1A - C (SEQ ID NO: 2) 中约第 60 至约第 240 位的氨基酸残基,跨膜结构域含有图 1A - C (SEQ ID NO: 2) 中约第 37 至约第 59 位的氨基酸残基,胞内结构域含有图 1A - C (SEQ ID NO: 2) 中由约第 1 至约第 36 位的氨基酸残基,推断其分子量大约 26.4kDa。图 1A - C (SEQ ID NO: 2) 所示 AIM II 蛋白与人 Fas 配基 (图 2A - F) 的氨基酸序列具有大约 27 % 同一性和大约 51 % 相似性,与人 TNF - α (图 2A - F) 的氨基酸序列具有大约 26 % 同一性和大约 47 % 相似性。

作为一名普通技术人员,应能理解由于存在上面所讨论的测序错误的可能性,由保藏的 cDNA 所编码的推测 AIM II 多肽含有大约 240 个氨基酸,但氨基酸数可能是 230 - 250 个。进一步应理解,根据所使用的标准,AIM II 多肽的胞外、胞内和跨膜结构域的确切位点可能有微小差异。比如,图 1A - C (SEQ ID NO: 2) 中 AIM II 胞外结构域的确切位点可能会由于用来限定该区域的标准而会有微小不同(如具体位点可能在约第 1 至 5 位残基间“漂移”)。

如上所述,本发明的核酸分子可以是 RNA 形式,如 mRNA,或 DNA 形式,包括如通过克隆或合成产生的 cDNA 和基因组 DNA。DNA 可以是双链或单链。单链 DNA 或 RNA 可以是编码链(也称被为有义链),也可以是非编码链(也被称为反义链)。

“分离”的核酸分子意指从其天然环境中分离出的核酸分子—DNA 或 RNA。例如,为了本发明目的,将包含于载体中的重组 DNA 分子也看作是“分离的”。分离的 DNA 分子的其它例子包括存在于异源宿主细胞内的重组 DNA 分子或以溶液状态存在的已纯化(部分或基本上) DNA 分子。分离的 RNA 分子包括本发明的 DNA 分子的体内或体外 RNA 转录

本。根据本发明，分离的核酸分子还包括合成生产的此类分子。

本发明的分离核酸分子包括含有图 1A - C (SEQ ID NO: 1) 所示开放阅读框 (ORF) 的 DNA 分子; 含有图 1A - C (SEQ ID NO: 2) 所示 AIM II 蛋白编码序列的 DNA 分子; 以及含有与上述序列基本上不同、但由于遗传密码简并性而仍编码 AIM II 蛋白的序列的 DNA 分子。当然, 遗传密码在本领域是众所周知的。因此, 对本领域技术人员而言, 产生这类简并性变体的方法是很常规的。

另一方面, 本发明提供了编码 AIM II 多肽的分离核酸分子, 所述 AIM II 多肽具有 cDNA 克隆所编码的氨基酸序列, 该 cDNA 克隆位于 1996 年 8 月 22 日以 ATCC 保藏号 97689 保藏的质粒中。优选地, 该核酸分子编码由上述保藏 cDNA 克隆所编码的多肽。本发明还提供了具有图 1A - C (SEQ ID NO: 1) 所示核苷酸序列或具有上述保藏克隆所含 AIM II cDNA 的核苷酸序列的分离核酸分子, 或具有与上述序列之一互补的序列的核酸分子。这类分离的分子, 尤其是 DNA 分子, 可有效用作探针, 通过染色体原位杂交进行基因作图, 以及通过如 Northern 印迹分析检测人组织中 AIM II 基因的表达。

本发明也涉及本文所述的分离核酸分子的片段。具有保藏 cDNA 的核苷酸序列或具有图 1A - C (SEQ ID NO: 1) 所示核苷酸序列的分离核酸的片段意指如本文所述可有效用作诊断探针和引物的长度至少大约 15nt、更优选至少约 20nt、更更优选至少约 30nt、甚至更优选至少约 40nt 的片段。当然, 按照本发明, 50 - 1500nt 长度的更大片段也是有用的, 与保藏 cDNA 或图 1A - C (SEQ ID NO: 1) 所示的核苷酸序列绝大部分 (如果不是全部的话) 一致的片段同样是有用的。例如, 长度至少约 20nt 的片段意指含有保藏 cDNA 的核苷酸序列或图 1A - C (SEQ ID NO: 1) 所示的核苷酸序列来源的 20 或更多个连续碱基的片段。

优选本发明的核酸片段包括编码 AIM II 蛋白中含表位的区域的核酸分子。具体地说, 本发明的这类核酸片段包括编码下述多肽的核酸分子: 含有图 1A - C (SEQ ID NO: 2) 中约第 13 位至约第 20 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中约第 23 位至约

第 36 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中约第 69 位至约第 79 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中约第 85 位至约第 94 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中约第 167 位至约第 178 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中约第 184 位至约第 196 位氨基酸残基的多肽; 以及含有图 1A - C (SEQ ID NO: 2) 中的第 221 位至约第 233 位氨基酸残基的多肽。本发明人已测出上述多肽片段都是 AIM II 蛋白的抗原性区域。下面详述测定 AIM II 蛋白其它的含表位的区域的方法。

根据本发明, AIM II 多核苷酸可用于多种应用中, 尤其是利用了 AIM II 的化学和生物学特性的应用中。这些应用涉及自身免疫疾病和异常细胞增殖。其它一些应用涉及细胞、组织和生物体异常的诊断和治疗。

本发明也涉及将 AIM II 多核苷酸用于检测互补性多核苷酸链的用途, 例如用作诊断试剂。对与机能障碍相关的 AIM II 突变形式的检测将提供一种诊断工具, 这种诊断工具可增强或限定起因于 AIM II 过低表达、过高表达或异常表达的疾病或疾病敏感性(如自身免疫疾病)的诊断中。由于 AIM II 基因存在于多种肿瘤细胞系中, 包括胰腺肿瘤、睾丸肿瘤、子宫内膜肿瘤和 T 细胞淋巴瘤, 因此, 编码 AIM II 的多核苷酸也可用作诊断标记以检测本发明多肽的表达。

另一方面, 本发明提供了含有某种多核苷酸的分离核酸分子, 该多核苷酸在严格杂交条件下与本发明的上述核酸分子(如 ATCC 保藏号 97689 所含 cDNA 克隆)中的多核苷酸区段能杂交。“严格杂交条件”意指在含有下述成分的溶液中于 42 °C 保温过夜: 50 % 甲酰胺, 5 × SSC(150mM NaCl, 15mM 柠檬酸三钠), 50mM 磷酸钠(pH7.6), 5 × Denhardt's 溶液, 10 % 硫酸葡聚糖及 20g/ml 变性的经剪切的鲑精 DNA, 随后在 0.1 × SSC 中于大约 65 °C 下洗涤滤膜。

能与一个多核苷酸的“区段”杂交的多核苷酸意指一个与标准多核苷酸中至少大约 15 个核苷酸(nt)、更优选至少约 20nt、更更优选至少约 30nt、甚至更优选大约 30 - 70nt 的区段能杂交的多核苷酸(DNA 或 RNA)。如上文和下文详细论述, 它们可有效用作诊断探针和引物。

例如, “长度至少为 20nt”的多核苷酸的区段意指来源于标准多核

苷酸（如保藏的 cDNA 或图 1A - C(SEQ ID NO: 1)所示核苷酸序列）的核苷酸序列中 20 个或更多个连续核苷酸。

当然，仅与 poly A 序列〔如图 1A - C(SEQ ID NO: 1)所示 AIM II cDNA 的 3'末端 poly(A)区〕或 T（或 U）残基的互补延伸链能杂交的多核苷酸并不包括在用于与本发明核酸区段的杂交中的本发明多核苷酸内，这是由于这种多核苷酸分子与含有 poly(A)延伸链或其互补区段的任何核酸分子（如实际中的任何双链 cDNA 克隆）都能杂交上。

如本文所述，编码 AIM II 多肽的本发明的核酸分子可包括但并不限于下述核酸分子：编码所述多肽的氨基酸序列的那些核酸分子；多肽编码序列和其它序列，如编码前导序列或分泌序列（如前蛋白或蛋白原或前蛋白原序列）的序列；含有或不含有前述的其它编码序列的多肽编码序列，所述多肽编码序列还含有其它的非编码序列，包括但不限于如内含子、5'和 3'非编码序列（如在转录和 mRNA 剪切中具有如结合核糖体、稳定 mRNA 等功能的转录、但不翻译的序列，包括剪切和聚腺苷酸化信号）；编码其它氨基酸如提供其它功能的氨基酸的其它编码序列。因此，编码该多肽的序列可与标记序列融合，所述标记序列例如是编码能简化融合多肽的纯化的多肽的序列。在本发明此方面的某些优选实施方案中，标记氨基酸序列是六-组氨酸肽，如 pQE 载体（Qiagen, Inc.）上提供的标记，这类肽中多数可商购。例如，如同 Gentz 等，美国国家科学院院报，86: 821-824 (1989)所述，六-组氨酸的存在便于对融合蛋白进行纯化。“HA”标记是另一个可有效用于纯化的肽，此“HA”标记对应于流感血凝素蛋白来源的表位，该表位已被记载于 Wilson 等，细胞，37: 767 (1984)。如下文所述，其它这类融合蛋白包括在 N 末端或 C 末端与 Fc 融合的 AIM II。

本发明的核酸分子还包括编码缺乏 N - 末端甲硫氨酸的全长 AIM - II 多肽的核酸分子。

本发明还涉及本发明的核酸分子的变体，这些变体编码 AIM II 蛋白的区段、类似物或衍生物。变体可以是天然存在的，如天然的等位基因变体。一个“等位基因变体”意指位于某种生物体中某条染色体上特定位点的某基因的若干种变体形式之一（基因 II，Lewin B. 编，John

Wiley & Sons, New York (1985))。非天然存在的变体可利用本领域已知的致突技术产生。

这类变体包括可涉及一个或多个核苷酸的核苷酸取代、缺失或添加而产生的核酸分子。这些变体可以在编码区、非编码区或两种区域中都发生了改变。编码区内的改变可产生保守的或非保守的氨基酸取代、缺失或添加。这些改变中特别优选沉默取代、添加和缺失，它们并不改变 AIM II 蛋白或其区段的性质和活性。就此而言，也特别优选保守取代。

本发明的另一实施方案中包括含有某种多核苷酸和分离核酸分子，该多核苷酸与下述核苷酸序列之一的同一性至少为 90 %、更优选至少为 95 %、96 %、97 %、98 % 或 99 %：(a) 编码具有图 1A - C (SEQ ID NO: 2) 中全部氨基酸序列的 AIM II 多肽的核苷酸序列；(b) 编码具有 cDNA 克隆所编码的全部氨基酸序列的 AIM II 多肽的核苷酸序列，该 cDNA 克隆包含于 ATCC 保藏号 97689 中；(c) 编码 AIM II 多肽胞外结构域的核苷酸序列；(d) 编码 AIM II 多肽跨膜结构域的核苷酸序列；(e) 编码 AIM II 多肽胞内结构域的核苷酸序列；(f) 编码可溶性 AIM II 多肽的核苷酸序列，该可溶性 AIM II 多肽具有胞外和胞内结构域，但缺乏跨膜结构域；以及 (g) 与上述 (a)、(b)、(c)、(d)、(e) 或 (f) 中任一种核苷酸序列互补的核苷酸序列。

例如，含有与编码 AIM II 多肽的标准核苷酸序列具有至少 95 % “同一性”的核苷酸序列的多核苷酸意指所述多核苷酸的核苷酸序列与标准序列大致上是相同的，只是该多核苷酸序列在编码 AIM II 多肽的标准核苷酸序列中每 100 个核苷酸内可含有不超过 5 个点突变。换句话说，为了获得含有与标准核苷酸序列具有至少 95 % 同一性的核苷酸序列的多核苷酸，标准序列内不超过 5 % 的核苷酸可缺失或由其它核苷酸所取代；或者是，在标准序列内可插入占标准序列总核酸数不超过 5 % 的数目的核苷酸。标准序列的这些突变可发生在标准核苷酸序列的 5' 或 3' 末端位置，也可发生在这些末端位置之间的任何地方，它们或者是在标准序列内核苷酸序列之间分散存在，或者是在标准序列内以一个或数个连续组的形式存在。

作为一项实际工作，任一特定核酸分子与如图 1A - C 所示的核苷

酸序列或保藏 cDNA 克隆的核苷酸序列是否具有至少 90 %、95 %、96 %、97 %、98 % 或 99 % 同一性可常规地利用如 Bestfit 程序 (Wisconsin Sequence Analysis Package, 575 Science Drive, Madison, WI 53711) 等已知计算机程序进行测定。Bestfit 程序使用了 Smith 和 Waterman, 应用数学进展, 2: 482- 489 (1981) 的局部同源性算法以寻找两序列间具同源性的最佳区段。利用 Bestfit 或其它任何一种序列比较程序确定某一特定序列与本发明的标准序列是否具有至少如 95 % 同一性时, 参数的设置当然是为了确保同一性百分比是相对标准核苷酸序列的全长计算的, 并且也允许存在占标准序列的核苷酸总数不超过 5 % 的同源性缺口。

本专利申请涉及与图 1A - C (SEQ ID NO: 1) 所示核酸序列或与保藏 cDNA 的核酸序列具有至少 90 %、95 %、96 %、97 %、98 % 或 99 % 同一性的核酸分子, 并且毋需要求这些核酸分子是否编码具有 AIM II 活性的多肽。这是因为即使某一特定核酸分子不编码具有 AIM II 活性的多肽, 本领域技术人员也应知道如何利用这些核酸分子, 如用作杂交探针或聚合酶链反应 (PCR) 引物。本发明中不编码具有 AIM II 活性的多肽的核酸分子包括如下应用: (1) 从 cDNA 文库中分离 AIM II 基因或其等位基因变体; (2) 如 Verma 等, 《人类染色体: 基础技术手册》, Pergamon 出版社, 纽约 (1988) 所述, 与中期染色体进行原位杂交 (如 “ FISH ”) 以进行 AIM II 基因的准确染色体定位; 以及 Northern 印迹分析以检测特异组织中 AIM mRNA 表达情况。

然而, 优选的核酸分子是具有与图 1A - C (SEQ ID NO: 1) 所示的核酸序列或保藏 cDNA 的核酸序列之间存在至少 90 %、95 %、96 %、97 %、98 % 或 99 % 同一性的序列并且确实编码具有 AIM II 蛋白活性的多肽的核酸分子。“具有 AIM II 活性的多肽”意指通过特定生物学测试, 具有与本发明 AIM II 蛋白活性相似的、但毋需完全相同的活性的多肽。例如, 如同 Zarres 等, 细胞, 70:31-46 (1992) 所述, AIM II 蛋白细胞毒活性可通过 propidium iodide 染色分析编程性细胞死亡情况而进行测定。或者, 也可通过 Gavriel 等, 细胞生物学杂志, 119:493-501 (1992) 所述的 TUNEL 染色测定 AIM II 诱导的编程性细胞死亡。

简要地说, propidium iodide 染色的操作如下: 将组织或培养物来源的细胞在甲醛中固定, 制成冰冻切片, 再用 propidium iodide 进行染色。利用共焦点荧光显微镜通过 propidium iodide 可观察到细胞核。通过固缩核(染色体凝集、皱缩和/或核的片段化)可示出细胞死亡情况。

当然, 由于遗传密码的简并性, 本领域普通技术人员很快会明白, 含有与保藏 cDNA 的核酸序列或图 1A - C (SEQ ID NO: 1) 所示的核酸序列具有至少 90 %、95 %、96 %、97 %、98 % 或 99 % 同一性的序列的大量核酸分子也编码“具有 AIM II 蛋白活性”的多肽。事实上, 由于这些核苷酸序列的简并性变体都编码同样的蛋白质, 即使不进行上述对比测试, 本领域技术人员也清楚这一点。本领域技术人员还明白, 对于那些不是简并性变体的核酸分子, 相当一部分也能编码具有 AIM II 蛋白活性的多肽。这是由于本领域技术人员完全了解那些极少可能或不大可能显著影响蛋白质活性的氨基酸取代(例如, 一种脂肪族氨基酸被另一种脂肪族氨基酸所取代)。

例如, Bowie J. U.等, “解译蛋白质序列内的信息: 氨基酸取代的耐受性”, 科学, 247: 1306-1310 (1990)中记载了有关如何进行无表型变化的氨基酸取代的规则, 其中作者指出蛋白质出人意料地能耐受氨基酸取代。

载体和宿主细胞

本发明也涉及含有本发明的分离 DNA 分子的载体、利用该重组载体进行了遗传学操作的宿主细胞, 以及通过重组技术进行的 AIM II 多肽或其片段的生产。

该多核苷酸可与含有选择性标记以在宿主中进行增殖的载体连接起来。通常质粒载体可通过沉淀物如磷酸钙沉淀物或具有带电脂类的复合物导入宿主。如果载体是一种病毒, 利用适当的包装细胞系可在体外将其包装起来, 再转导至宿主细胞中。

DNA 插入片段应与适当的启动子有效连接, 如噬菌体 λ PL 启动子、大肠杆菌 lac、trp 和 tac 启动子、SV40 早期和晚期启动子、逆转录病毒 LTR 启动子等。本领域技术人员也熟知适当的其它启动子。表达构建体还含有转录起始点、终止点, 以及转录区内用于翻译的核糖体结合位

点。由此构建体表达的成熟转录本的编码区段优选地在起始处含有翻译起始序列，在待翻译的多肽末端适当位置处含有终止密码子（UAA、UGA 或 UAG）。

如上所述，表达载体优选地含有至少一个选择性标记。这类标记包括用于真核细胞培养的新霉素抗性基因、二氢叶酸还原酶基因，用于大肠杆菌和其它细菌培养的四环素或氨苄青霉素抗性基因。适当宿主的代表性例子有细菌细胞如大肠杆菌、链霉菌和鼠伤寒沙门氏菌细胞；真菌细胞，如酵母细胞；昆虫细胞，如果蝇 S2 和 *Spodoptera Sf9* 细胞；动物细胞，如 CHO、COS 和 Bowes 黑素瘤细胞；以及植物细胞，但并不局限于上述细胞。本领域熟知上述宿主细胞的适当培养介质和培养条件。

用于细菌的优选载体包括可以 Qiagen 获得的 pQE70、pQE60 和 pEQ-9；可从 Stratagene 获得的 pBS 载体、Phagescript 载体、Bluescript 载体、pNH8A、pNH16a、pNH18A 和 pNH46A；以及可从 Pharmacia 获得的 ptrc99a、pKK223-3、pKK233-3、pDR540 和 pRIT5；等等。优选的真核生物载体有可从 Stratagene 获得的 pWLNEO、pSV2CAT、pOG44、pXT1 和 pSG；以及可从 Pharmacia 获得的 pSVK3、pBPV、pMSG 和 pSVL；等。对本领域技术人员而言，适宜的其它载体也是显而易见的。

可通过下列方法将构建体导入宿主细胞：磷酸钙沉淀转染、DEAE-葡聚糖介导的转染、阳离子类脂介导的转染、电击法、转导、感染或其它方法。这些方法已记载于多种标准实验手册中，如 Davis 等，《分子生物学基本方法》（1986）。

多肽可以修饰形式如融合蛋白形式表达，并且不仅可含有分泌信号，还可含有其它的异源性功能区。例如，可在多肽的 N-末端添加含其它氨基酸特别是带电氨基酸的区域，以便在宿主细胞内、在纯化过程中或在随后的使用 and 贮存过程中提供稳定性和持续性。另外，也可将小肽添加到多肽中以便于纯化，并可在多肽的终制备前去除这类区域。将小肽添加到多肽中以分泌去除这类区域。将小肽添加到多肽中以分泌和排出多肽、提高多肽稳定性、方便纯化多肽的技术是本领域的熟练和常规技术。优选的融合蛋白含有有助于蛋白质溶解的来源于免疫球蛋白的

异源区域。例如，EP - A - O 464 533（加拿大对应号 2045869）公开了含有免疫球蛋白分子恒定区的各种区段以及其它人类蛋白质或其部分的融合蛋白。在多种情形下，融合蛋白中的 Fc 部分在治疗和诊断中是非常有应用价值的，并能提高如药物动力学特性（EP - A 0 232 262）。另一方面，在有些应用中，最好是在通过已述的较好方法表达、检测并纯化该融合蛋白后去除其中的 Fc 部分。当 Fc 区段被证实对于融合蛋白在治疗和诊断中的应用具有不良作用如融合蛋白用作免疫抗原时，就是如此。例如，在药物的发掘中，人类蛋白质如 hIL-5 就与 Fc 区段融合以达到鉴定 hIL-5 拮抗物的筛选试验的高容量目的。参见 D. Bennett 等，分子识别杂志，第 8 卷：52 - 58(1995)和 K. Johanson 等，生物化学杂志，第 270 卷，16 期：9459- 9471 (1995)。

通过众所周知的方法，包括硫酸铵或乙醇沉淀、酸提取、阴离子或阳离子交换层析、磷酸纤维素层析、疏水性相互作用层析、亲和层析、羟磷灰石层析及外源凝集素层析等方法，可从重组细胞培养物中回收和纯化 AIM II 多肽。在纯化过程中，最优选使用高效液相层析方法（“HPLC”）。本发明的多肽包括天然纯化产物、化学合成途径产生的产物以及从原核生物或真核生物宿主包括如细菌、酵母、更高等的植物、昆虫和哺乳动物细胞等通过重组技术产生的产物。根据在重组生产方法中应用的宿主类型，本发明的多肽可能被糖基化，也可能未被糖基化。另外，本发明的多肽也可能含有已修饰的起始甲硫氨酸残基，在有些情形下这是由于宿主介导的过程产生的。

AIM II 多肽和片段

本发明还提供了具有保藏 cDNA 所编码的氨基酸序列或图 1A - C（SEQ ID NO：2）中的氨基酸序列的分离 AIM II 多肽，或含有上述多肽的一个区段的肽或多肽。

本领域深知可改变 AIM II 多肽的一些氨基酸序列而不显著影响此蛋白的结构或功能。如能仔细分析这类序列差异，应能清楚蛋白质中存在决定其活性的重要区域。

因此，本发明还包括具有显著的 AIM II 多肽活性或含有 AIM II 蛋白区域（如下述的蛋白质区段）的 AIM II 多肽的变体。这类突变体包括

缺失、插入、转换、重复和类型取代。如上所述，有关何种氨基酸变化有可能无表型差异的规则可见于 Bowie J. U.等，“解译蛋白质序列中的信息：对氨基酸替换的耐受性”，科学，247: 1306- 1310 (1990)。

因此，图 1A - C (SEQ ID NO: 2) 的多肽或由保藏 cDNA 所编码的多肽的片段、衍生物或类似物可以是：(i) 其中一个或多个氨基酸残基被保守的或非保守的氨基酸残基（优选保守的氨基酸残基）所取代，这种取代的氨基酸残基可以是、也可以不是由该遗传密码所编码的氨基酸；或(ii) 其中一个或多个氨基酸残基含有取代基；或(iii) 其中成熟多肽与另一种化合物如能提高此多肽半衰期的化合物（如聚乙二醇）相融合；或(iv) 其中成熟多肽中融合了其它的氨基酸，如 IgG Fc 融合区肽、前导序列、分泌性序列或可用于成熟多肽纯化的序列，或蛋白原序列。这类片段、衍生物和类似物也在本文所教导的本领域技术人员的范围之内。

尤其引人注目的是带电氨基酸被另一种带电氨基酸、中性或带负电的氨基酸所取代，后者能产生带更少正电荷的蛋白质，因而能改进 AIM II 蛋白的特性。极有必要能避免凝集现象。蛋白质的凝集不仅会破坏活性，而且在制备药物制剂时也会由于具有免疫原性而带来大问题（Pinckard 等，临床实验免疫学，2: 331-340 (1967)；Robbins 等，糖尿病，36: 838-845 (1987)）；Cleland 等，Crit. Rev. Therapeutic Drug Carrier Systems, 10: 307- 377 (1993)）。

氨基酸的置换也可改变多肽与细胞表面受体的选择性结合。Ostade 等，自然，361: 266- 268 (1993) 中描述了某些突变能导致 TNF - α 仅与 TNF 受体的两种已知类型中的一种选择性结合。因此，本发明的 AIM II 受体可含有一个或多个氨基酸取代、缺失或添加，这类改变可以是天然突变或人工操作的结果。

如所述，优选影响较小的改变，如不显著影响蛋白质折叠或活性的保守氨基酸取代（见表 1）。

表 1 保守氨基酸取代

芳族氨基酸	苯丙氨酸 色氨酸 酪氨酸
疏水性氨基酸	亮氨酸 异亮氨酸 缬氨酸
极性氨基酸	谷氨酰胺 天冬酰胺
碱性氨基酸	精氨酸 赖氨酸 组氨酸
酸性氨基酸	天冬氨酸 谷氨酸
小氨基酸	丙氨酸 丝氨酸 苏氨酸 甲硫氨酸 甘氨酸

通过本领域已知方法,如位点特异性诱变方法或丙氨酸扫描诱变方法 (Cunningham 和 Wells, 科学, 244: 1081-1085 (1989)), 可鉴定出本发明的 AIM II 蛋白内对其功能是必不可少的氨基酸。后一方法是在分子内任何一个残基处引入单一丙氨酸突变,然后测定由此获得的突变体分子的生物学活性,如受体结合力,或体内或体外增殖活性。也可通过结构分析,如结晶过程、核磁共振或光亲和标记 (Smith 等, 分子生物学杂志, 224: 899- 904 (1992)和 de Vos 等, 科学, 255: 306- 312 (1992)) 测定出对配基 - 受体的结合至关重要的位点。

优选以分离形式提供本发明的多肽。“分离的多肽”意指从其

天然环境中提取的多肽。因此，为了本发明目的，重组宿主细胞内生产和包含的多肽也被认为是“分离的”。“分离的多肽”还指已从重组宿主内部分或基本上纯化的多肽。例如，可通过 Smith 和 Hohnson，基因，67: 31-40 (1988)所述的一步法基本纯化重组生产型 AIM II 多肽。

本发明的多肽包括由保藏 cDNA 编码的多肽、图 1A - C (SEQ ID NO: 2) 的多肽、缺乏 N - 末端甲硫氨酸的图 1A - C (SEQ ID NO: 2) 的多肽、胞外结构域、跨膜结构域、胞内结构域、含有全部或部分胞外和胞内结构域但缺乏跨膜结构域的可溶性多肽，以及与保藏 cDNA 编码的多肽或图 1A - C (SEQ ID NO: 2) 的多肽间的同一性至少为 80 %、更优选至少为 90 % 或 95 %、更更优选至少为 96 %、97 %、98 % 或 99 % 的多肽，还包括含有至少 30 个氨基酸、更优选至少 50 个氨基酸的此类多肽的区段。

含有与 AIM II 多肽的标准氨基酸序列具有至少如 95 % “同一性”的氨基酸序列的多肽意指此多肽的氨基酸序列与标准序列基本相同，只是此多肽序列在 AIM II 多肽的标准氨基酸中每 100 个氨基酸可含有不多于 5 个氨基酸的改变。换句话说，为了获得含有与标准氨基酸序列具有至少 95 % 同一性的氨基酸序列的多肽，可对标准序列中不超过 5 % 的氨基酸残基进行缺失或由其它的氨基酸所取代；或者是，可将占标准序列内总氨基酸残基不超过 5 % 数量的氨基酸插入至标准序列中。标准序列的这些改变可位于标准氨基酸序列的氨基或羧基末端位置，可这些末端位置之间的任何部位，它们或者单独分散于标准序列的残基中，或以一个或多个连续群组的形式存在于标准序列内。

作为一项实际工作，可利用如 Bestfit 程度 (Wisconsin Sequence analysis Package, 575 Science Drive, Madison, WI 53711) 等已知计算机程序常规地测定任一特定多肽与如图 1A - C (SEQ ID NO: 2) 所示的氨基酸序列或保藏 cDNA 克隆所编码的氨基酸序列是否具有至少 90 %、95 %、96 %、97 %、98

% 或 99 % 同一性。利用 Bestfit 或其它任何一种序列比较程序分析某一特定序列与本发明的标准序列间是否具有至少如 95 % 同一性时, 参数的设置当然是为了确保同一性百分比是相对于标准氨基酸序列的全长计算的, 并且也允许存在占标准序列的氨基酸总数达 5 % 的同源性缺口。

本文所用术语“AIM II”多肽包括膜结合性蛋白(含有胞质结构域、跨膜结构域及胞外结构域), 以及保留有 AIM II 多肽活性的截短蛋白质。在一个实施方案中, 可溶性 AIM II 多肽含有 AIM II 蛋白胞外结构域的全部或部分, 但缺乏能将多肽滞留于细胞膜上的跨膜区段。只要可溶性 AIM II 能被分泌出来, 它也可含有部分跨膜结构域或部分胞质结构域或其它序列。可在可溶性 AIM II 多肽的 N - 末端融合一段异源信号肽, 以便在表达可溶性 AIM II 多肽后将其分泌出来。

利用本领域技术人员熟知的方法, 可将本发明的多肽用作 SDS - PAGE 凝胶或分子筛凝胶过滤柱上的分子量标记。

另一方面, 本发明提供了含有本发明多肽的含表位的区段的肽或多肽。此多肽区段的表位是本文所述多肽的免疫原性或抗原性表位。“免疫原性表位”是指整个蛋白质作为免疫时能引发抗体反应的蛋白质部分, 而能与抗体结合的蛋白质分子中的区域则被称为“抗原性表位”。蛋白质内免疫原性表位的数目一般比抗原性表位少。参见如 Geysen 等, 美国国家科学院院报, 81: 3998 - 4002 (1983)。

至于含有抗原性表位的肽或多肽(即含有蛋白质分子中能与抗体结合的区域肽或多肽)的选择, 本领域熟知与部分蛋白南序列极为相似的相对较短的合成肽一般能刺激产生可与该部分相似的蛋白质发生反应的抗血清。参见如 Sutcliffe H. G., Shinnick T. M., Green N. 和 Learner R. A. (1983), “可与蛋白质中预定位点发生反应的抗体”, 科学, 219: 660 - 666。能引发产生蛋白质反应性血清的肽多以蛋白质的一级序列表示, 它们可通过一套简单的化学规则确定, 并且不局限于完整蛋白质的免疫显性区域(即免疫原性

表位) 或者其氨基或羧基末端。

因此, 本发明中含有抗原性表位的肽和多肽可有效用于制备能与本发明多肽特异性结合的抗体, 包括单克隆抗体。参见如 Wilson 等, 细胞, 37: 767 - 778 (1984), 于 777 页。

本发明携有抗原性表位的肽和多肽所含的一段序列是本发明的多肽的氨基酸序列内所含的优选至少 7 个、更优选至少 9 个、最优选至少 15 个左右至 30 个左右的氨基酸。

能用于制备 AIM II 特异性抗体的抗原性多肽或肽的例子包括但并不局限于如下: 含有图 1A - C (SEQ ID NO: 2) 中大约第 13 位至大约第 20 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 23 位至大约第 36 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 69 位至大约第 79 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 85 位至大约第 94 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 167 位至大约第 178 位氨基酸残基的多肽; 含有图 1A - C (SEQ ID NO: 2) 中大约第 184 位至大约第 196 位氨基酸残基的多肽; 以及含有图 1A - C (SEQ ID NO: 2) 中大约第 221 位至大约第 233 位氨基酸残基的多肽。如上所述, 本发明人已确定上述多肽片段是 AIM II 蛋白的抗原性区域。

可通过任何传统途径生产本发明携有表位的肽和多肽。Houghten R. A. (1985), “快速固相合成大量多肽的一般方法: 抗原 - 抗体反应在单个氨基酸水平上的特异性”, 美国国家科学院院报, 82: 5131-5135。这种“多个多肽同时合成 (SMPS)”途径被进一步描述于 Houghten 等的美国专利号 4, 631, 211 中。

可将本发明的 AIM II 多肽用于治疗诱发淋巴结病的淋巴组织增生性疾病。由于 AIM II 可通过刺激 T - 细胞的克隆缺失进而介导编程性细胞死亡, 因而, 可将 AIM II 应用于治疗自身免疫疾病、刺激周边耐受性和 T - 细胞介导的具细胞毒性的编程性细胞死亡过程。也可将 AIM II 用作研究工具, 以阐明自身免疫紊乱和过敏反应的生物学机制, 并治疗移植物抗宿主疾病, 所述自身免疫紊乱

包括系统性红斑狼疮（SLE）、免疫增生性疾病、淋巴结病（IPL）、血管免疫增生性淋巴结病（AIL）、淋巴瘤细胞性淋巴结病（IBL）、类风湿性关节炎、糖尿病和多发性硬化症。

也可将本发明的 AIM II 多肽用来抑制肿瘤，如肿瘤细胞生长。AIM II 多肽可能通过编程性细胞死亡和对某些细胞的细胞毒作用而对肿瘤的破坏起着一定的作用。由于 AIM II 能刺激内皮来源细胞的增殖，因而也可将 AIM II 用于治疗一些需生长促进活性的疾病，如再狭窄性疾病。并且，AIM II 也可用于调节内皮细胞发育过程中的造血功能。

本发明还提供了鉴定能与 AIM II 结合的分子如受体分子的方法。通过本领域技术人员已知的众多方法，如配基选择和 FACS 分类，可鉴定出编码能与 AIM II 结合的蛋白质如受体蛋白的基因。在许多实验室手册如 Coligan 等，免疫学现代方法，1（2）：第 5 章（1991）中，都记载有这类方法。

例如，为了此目的，可使用表达克隆法。为达此目的，首先从对 AIM II 具有反应的细胞中制备聚腺苷酸化的 RNA，从此 RNA 构建 cDNA 文库，将文库分成几组，再分别转化对 AIM II 无反应的细胞。然后将转化细胞与标记 AIM II 混合处理。（AIM II 的标记可通过多种熟知技术，包括放射性碘处理或包含位点特异性蛋白激酶识别位点等标准方法。）处理后，固定细胞，并测定 AIM II 的结合率。此方法可在载玻片上简便进行。

鉴定出含有能产生 AIM II 结合性细胞的 cDNA 的各组。将这些阳性组分分成亚组，再按上述方法转化宿主细胞并进行筛选。通过这种重复性亚分组 - 重新筛选（sub-pooling and re-screening）过程，可分离出编码推定的结合分子如受体分子的一个或多个单克隆。

或者是将经标记的配基与细胞提取物如膜或膜提取物进行光亲和性连接，这种细胞提取物是从能表达可与该配基结合的分子如受体分子的细胞中制备的。通过聚丙烯酰胺凝胶电泳（“PAGE”）分离交联物质，再对 X - 光胶片进行曝光。切下含有配基 - 受体的

标记复合物，将其裂解为肽片段，再进行蛋白质微量测序。利用从微量测序结果中获得的氨基酸序列，可设计出特异性或简并性寡核苷酸探针，以对 cDNA 文库进行筛选，鉴定出编码推定的受体分子的基因。

也可将本发明的受肽用于验证细胞或无细胞制备物中的 AIM II 结合性分子如受体分子结合 AIM II 的能力。

本领域技术人员深知，可将本发明的 AIM II 多肽或上述的其含表位的片段与免疫球蛋白 (IgG) 恒定结构域的一部分相组合产生嵌合多肽。这类融合蛋白可便于纯化，并在体内具有更长的半衰期。如含有人 CD4 多肽前两个结构域和哺乳动物免疫球蛋白重链或轻链的恒定区内各个结构域的嵌合蛋白就体现了上述特点 (EPA394, 827; Traunecker 等, 自然, 331: 84-86 (1988))。由于 IgG 部分而具有二硫键连接的双体结构的融合蛋白与单体 AIM II 蛋白或单独的蛋白片段相比, 前者也能更有效结合和中和其它分子 (Fountoulakis 等, 生物化学杂志, 270: 3958-3964 (1995))。

本发明人已经发现 AIM II 在脾脏、胸腺和骨髓组织中都有表达。对一系列紊乱如败血症性休克、炎症、脑型疟、HIV 病毒的活化、移植物-宿主排斥性、骨吸收、类风湿性关节炎和恶病质而言, 在具有这类紊乱的个体的某些组织 (如脾脏、胸腺和骨髓组织) 或体液 (如血清、细胞液、尿液、滑液或脊柱液) 中, 能检测到明显高于或低于“标准” AIM II 基因表达水平的 AIM II 基因的表达, 其中“标准” AIM II 基因表达水平即为不具有这类紊乱的个体的组织或体液中 AIM II 的表达水平。因此, 本发明提供了一种在紊乱诊断过程中非常有用的诊断方法, 这种方法包括: (a) 测定个体的细胞或体液中 AIM II 基因表达水平; (b) 将测得的 AIM II 基因表达水平与标准 AIM II 基因表达水平进行比较, 其中测得的 AIM II 基因表达水平高于或低于标准表达水平都表明存在某种紊乱。

AIM II 激动剂和拮抗剂

本发明也提供了筛选化合物以鉴定出能增强或阻抑 AIM II 对

细胞的作用如与 AIM II 结合性分子（如受体分子）的相互作用的化合物的方法。激动剂是能增强 AIM II 天然生物学功能或以类似于 AIM II 的方式发挥作用的化合物；而拮抗剂则能降低或消除这类功能。

例如，从表达可结合 AIM II 的分子如由 AIM II 通过信号或调节途径调控的分子的细胞中可制备细胞组分，如膜制备物。在缺乏或存在待测分子（该待测分子可能是 AIM II 的激动剂或拮抗剂）的情况下将制备物与已标记的 AIM II 一起保温。标记配基结合力的降低水平反映了待测分子与结合性分子或 AIM II 本身相结合的能力。能“无偿”结合即与 AIM II 结合性分子结合时不会诱导 AIM II 反应的分子最有可能是好的拮抗剂。能很好结合并且作用效果与 AIM II 相同或密切相关的分子是好的激动剂。

例如，通过潜在激动剂或拮抗剂与细胞或适当的细胞制备物间的相互作用而测定第二信使系统的活性，并与 AIM II 或具有与 AIM II 相同功能的分子的活性相比较，可测出待测分子的 AIM II 相似性活性。在这方面，有效的第二信使系统包括但不限于 AMP 鸟苷酸环化酶、离子通道或磷酸肌醇水解第二信使系统。

AIM II 拮抗剂测试法的另一个例子是竞争性测试法，这种测试方法是在适于竞争抑制测试的适当条件下，将 AIM II 和潜在拮抗剂与膜结合的 AIM II 受体分子或重组 AIM II 受体分子相组合。可通过如放射性活性等对 AIM II 进行标记，以便能准确测定与受体分子结合的 AIM II 分子的数量，从而估计该潜在拮抗剂的效果。

潜在的拮抗剂包括能与本发明的多肽结合从而能抑制或消除其活性的一些小的有机分子、肽、多肽和抗体。潜在的拮抗剂也可以是能与结合性分子如受体分子的相同位点结合而不诱导出 AIM II 诱导性活性、因而通过阻止 AIM II 的结合而抑制 AIM II 的作用的一些小的有机分子、肽或多肽，如密切相关的蛋白质或抗体。

其它潜在的拮抗剂包括反义分子。可通过反义 DNA 或反义 RNA 或通过三联体螺旋形成而利用反义技术控制基因表达。反义

技术已被记载于如 Okano, 神经化学杂志, 56: 560 (1991); 《用作基因表达反义抑制剂的寡脱氧核苷酸》, CRC Press, Boca Raton, RL(1988). 三联体螺旋形成记载于如 Lee等, 核酸研究, 6: 3073 (1979); Cooney等, 科学, 241: 456 (1988); 以及 Dervan等, 科学, 251: 1360 (1991)。这些方法都是基于多核苷酸与互补性 DNA 或 RNA 的结合。例如, 可利用编码本发明成熟多肽的多核苷酸的 5' 编码区, 设计长度约 10 ~ 40 个碱基对的反义 RN 寡核苷酸。DNA 寡核苷酸设计为与参与转录的基因区互补, 因而可抑制 AIM II 的转录和生成。反义 RNA 寡核苷酸与体内 mRNA 杂交, 从而阻断 mRNA 分子翻译为 AIM II 多肽的过程。也可将上述寡核苷酸导入至细胞内以便在体内表达该反义 RNA 或 DNA, 从而抑制 AIM II 的生成。

可将拮抗剂与药用载体(如后述)一起制成组合物进行应用。

例如, 可将拮抗剂用于治疗恶病质, 这种病是由于脂蛋白脂肪酶的系统性缺乏导致的类脂清除缺陷, 据称这种脂蛋白脂肪酶可被 AIM II 抑制。也可将 AIM II 拮抗剂用于治疗脑型症, AIM II 可能在此病中具有病理学作用。也可应用此拮抗剂治疗是类风湿性关节炎, 这是通过抑制 AIM II 诱导的炎症因子如滑膜细胞内的 IL1 的产生而进行的。在治疗关节炎时, 优选将 AIM II 拮抗剂注入关节内。

也可利用 AIM II 拮抗剂, 通过阻止免疫系统的刺激而预防移植物存在时发生的移植物-宿主排斥反应。

也可利用 AIM II 拮抗剂抑制骨吸收, 因而治疗和/或预防骨质疏松症。

也可将拮抗剂用作抗炎症试剂, 并治疗内毒素性休克。这种危险状态起因于对细菌或其它类型感染作出的超常反应。

癌症预测

与相应的“正常”哺乳动物相比, 患有癌症的哺乳动物的某些组织表达 AIM II 蛋白和编码 AIM II 蛋白的 mRNA 的水平显著降低。正常哺乳动物即是指与患癌哺乳动物属同一种类的、不患癌症的

个体。另外，与属同一种类但不患癌症的哺乳动物的血清相比，在患有癌症的哺乳动物的某些体液（如血清、胞质、尿液和脊柱液）中测到的 AIM II 水平明显降低。因此，本发明提供了可在肿瘤诊断中有效应用的诊断方法，这种方法包括测定哺乳动物细胞或体液中编码 AIM II 蛋白的基因的表达水平，并将测得的基因表达水平与标准 AIM II 基因表达水平相比较，如果测得的基因表达水平低于标准即表明患有某种肿瘤。

按照常规方法已能进行肿瘤诊断时，本发明可有效用作预测指示剂，其中具有强化 AIM II 基因表达的病人比基因表达水平更低的病人的临床效果可能要更好一些。

“测定编码 AIM II 蛋白的基因的表达水平”意指直接地（如通过测定或估计蛋白质水平或 mRNA 水平的绝对量）或相对地（如通过与第二种生物样品的 AIM II 蛋白水平或 mRNA 水平相比较）定性或定量地测定或估计第一种样品中 AIM II 蛋白水平或编码 AIM II 蛋白的 mRNA 的水平。

优选地，测定或估计第一种生物样品内 AIM II 蛋白水平或 mRNA 水平并与标准 AIM II 蛋白水平或 mRNA 水平相比较，此标准值是从得自未长癌的个体的第二种生物样品中测得的。本领域深知，一旦已测得标准 AIM II 蛋白水平或 mRNA 水平，即可将该值重复用作比较标准。

“生物样品”意指得自含有 AIM II 蛋白或 mRNA 的个体、细胞系、组织培养物或其它来源的任何生物样品。生物样品包括含有分泌的成熟 AIM II 蛋白的哺乳动物体液（如血清、胞质、尿液、滑液和脊柱液），以及卵巢、前列腺、心脏、胎盘、胰脏、肝、脾、肺、胸和脐组织。

本发明可有效用于哺乳动物的癌症诊断中。本发明尤其可有效用于哺乳动物内下列各类器官的癌症的诊断中：胸、卵巢、前列腺、骨、肝、肺、胰和脾。优选的哺乳动物包括猴、猿、猫、狗、牛、猪、马、兔和人类。尤其优选人类。

利用 Chomczynski 和 Sacchi，生物化学年鉴（Anal.

Biochem.), 162: 156 - 159 (1987) 记载的硫氰酸胍 - 苯酚 - 氯仿一步法可从生物样品中分离总细胞 RNA。然后利用适当的方法测定编码 AIM II 蛋白的 mRNA 的水平。这些方法包括 Northern 印迹分析 (Harada 等, 细胞, 63: 303 - 312(1990))、S1 核酸酶作图 (Fujita 等, 细胞, 49: 357 - 367(1987))、聚合酶链反应 (PCR)、逆转录与聚合酶链反应的组合 (RT - PCR) (Makino 等, 技术, 2: 295 - 301(1990)), 以及逆转录与连接酶链反应的组合 (RT - LCR)。

利用基于抗体的技术可测定生物样品中 AIM II 蛋白的水平。例如, 通过传统的免疫组织学方法 (Jalkanen M.等, 细胞生物学杂志, 101: 976 - 985(1985); Jalkanen M.等, 细胞生物学杂志, 105: 3087 - 3096(1987)), 可研究组织中 AIM II 蛋白的表达情况。

其它可有效用于 AIM II 蛋白基因表达测定的基于抗体的方法包括免疫测定, 如酶连免疫吸附测定 (ELISA) 和放射免疫测定 (RIA)。

本领域熟知各种适用标记, 它们包括酶标记如葡萄糖氧化酶, 放射性同位素如碘 (^{125}I 、 ^{121}I)、碳 (^{14}C)、硫 (^{35}S)、氚 (^3H)、铟 (^{112}In) 和锝 ($^{99\text{m}}\text{Tc}$), 以及荧光标记如荧光素、若丹明和生物素。

疗法

可将 AIM II 多肽尤其是人 AIM II 多肽用于治疗瘤形成、淋巴结病、自身免疫疾病、移植物抗宿主疾病。另外, 本发明的 AIM II 多肽也可用于抑制瘤生成, 如肿瘤细胞的生长。AIM II 多肽可能通过编程性细胞死亡和对某些细胞的细胞毒作用而参与破坏肿瘤。由于 AIM II 对内皮来源的细胞具有增殖效果, 故也可将 AIM II 用于需生长促进活性的疾病治疗中, 如再狭窄。因此, 也可将 AIM II 用于调节内皮细胞发育过程中的造血功能。

给药方式

应该明白, 上述病症可通过服用 AIM II 蛋白而得到治疗。因

此，本发明还提供了治疗需提高 AIM II 活性水平的个体的方法，这种方法包括给这类个体服用含有有效量的本发明的分离 AIM II 多肽、特别是 AIM II 的成熟形式的药用组合物，以有效提高这种个体体内的 AIM II 活性水平。

尽管如上所述可根据治疗需要灵活掌握用量，但作为一般建议，经胃肠外服用的 AIM II 多肽的总治疗有效量每剂为大约 $1\mu\text{g/kg}$ 病人体重/天 $\sim 10\text{mg/kg}$ 病人体重/天。此剂量更加优选为至少 0.01mg/kg/天 ，对人类最优选此激素服用剂量为大约 $0.01\sim 1\text{mg/kg/天}$ 。如果持续给药，通常可通过每天 1 $\sim$ 4 次注射或连续性皮下输液（如利用微型泵）完成，其中 AIM II 多肽服用的剂量速率为大约 $1\mu\text{g/kg/hour}\sim 50\mu\text{g/kg/hour}$ 。也可使用静脉内包溶液。

含有本发明的 AIM II 的药用组合物可通过口服给药、直肠内给药、胃肠外给药、脑池内给药、阴道内给药、腹膜内给药、局部给药（通过粉末、软膏、滴剂或表皮贴片）、颊部（*bucally*）给药或通过口腔或鼻腔喷雾剂给药。“药用载体”意指无毒的固体、半固体或液体配方、稀释剂、胶囊类物质或任何类型的制剂辅料。本文所用术语“胃肠外的”意指包括静脉内、肌肉内、腹膜内、胸骨内、皮下和关节内注射和输液等的给药方式。

除可使用可溶性 AIM II 多肽外，也可使用通过加入去污剂如曲拉通 X - 100 和缓冲液而适当溶解了的含有跨膜区的 AIM II 多肽。

染色体检测

本发明的核酸分子在染色体鉴定工作中也是很有价值的。此序列可特异性导向至人个体染色体的特定位点并与其杂交。按照本发明，将 DNA 定位于染色体上是在联系序列与疾病相关基因时非常重要的首要步骤。

在这方面的某些优选实施方案中，可利用本文所公开的 cDNA 克隆 AIM II 蛋白的基因组 DNA。利用本领域熟知的多种技术和文库即可完成此项工作。这些文库一般可商购。然后，通过本领域熟

知的可进行原位染色体作图的技术，利用此基因组 DNA 进行原位染色体作图。

另外，在某些情况下，通过从 cDNA 制备 PCR 引物（优选 15 - 25bp）可将序列定位在染色体上。通过计算机分析该基因的 3' 非翻译区，可快速挑选出并不跨越基因组 DNA 中一个以上外显子的引物，因而可使扩增过程复杂化。然后，利用这些引物，对含有人个体染色体的体细胞杂合体进行 PCR 筛选。

利用 cDNA 克隆对中期染色体进行荧光原位杂交（“FISH”）可一步提供精确的染色体定位。该项技术中可使用来自 cDNA 的短至 50 或 60bp 的探针。有关该技术的综述参见 Verma 等，人类染色体：基础技术手册，Pergamon Press，纽约（1988）。

一旦一个序列已定位于染色体上某一精确位点时，即可将此序列在染色体上的物理位置与遗传图谱数据联系起来。这类数据可见于如 V.McKusick，《人类孟德尔遗传》，通过与 Johns Hopkins University, Welch Medical Library 联机获得。再通过连锁分析（物理上邻接基因的共遗传）对已定位于同一染色体区的基因和疾病之间的关系进行鉴定。

接着，有必要对感染个体与未感染个体的 cDNA 或基因组序列之间的差异进行分析测定。若在一些或全部感染个体内观察到突变的存在，而在任何正常个体中均未发现突变，那么，这种突变很可能是该病的病因。

上面已对本发明进行了一般性描述，下述实施例将有助于更好地理解本发明，这些实施例是以举例方式提供的，但并不用来限制本发明。

实施例

实施例 1：在大肠杆菌内表达和纯化 AIM II

A. 带有 N - 末端 HA 标记的 AIM II 的表达

利用特异于 AIM II 蛋白氨基末端序列及此基因的 3' 侧的载体序列的 PCR 寡核苷酸引物，扩增出保藏 cDNA 克隆内编码 AIM II 蛋白的 DNA 序列。分别在 5' 和 3' 序列中添加了含有便于克隆的限

制性位点的额外的核苷酸。

利用下述引物可表达一个 22kDa 的 AIM II 蛋白片段（缺乏 N - 末端和跨膜区）：

5'寡核苷酸引物的序列为 5'GCGGGATCCGGAGAGATGGT CACC 3'（SEQ ID NO： 3），它对应于图 1A - C（SEQ ID NO： 1）中 AIM II 蛋白编码序列内第 244 - 258 位核苷酸，并在其中含有一个由下划线标出的 BamHI 限制性位点；

3'引物的序列为 5'CGCAAAGCTTCCTTCACACCATGAAAG C 3'（SEQ ID NO： 8），它含有一个由下划线标出的 Hind III 限制性位点及紧随其后的图 1A - C 中所示第 756 - 783 位核苷酸。

通过下述引物可表达完整 AIM II 蛋白：

5'寡核苷酸引物的序列为 5'GACCGGATCCATG GAG GAG AGT GTC GTA CGG C3'（SEQ ID NO： 9），它对应于图 1A - C（SEQ ID NO： 1）中 AIM II 蛋白编码序列内 22 个核苷酸，并在其中含有一个由下划线标出的 BamHI 限制性位点；

3'引物的序列为 5'CGCAAAGCTTCCT TCA CAC CAT GAA AGC 3'（SEQ ID NO： 10），它含有一个由下划线标出的 Hind III 限制性位点及紧随其后的图 1A - C 中所示第 756 - 783 位核苷酸。

这些限制性位点与细菌表达载体 pQE9 内的限制性酶切位点是相匹配的，可在这些实施例中用于细菌的表达。（Qiagen, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311）。pQE9 编码氨苄青霉素抗生素抗性（“Amp^r”），并含有一个细菌复制起点（“ori”），一个 IPTG 诱导型启动子、一个核糖体结合位点（“RBS”）、一个 6 - His 标记以及一些限制性酶切位点。

利用 BamHI 和 Hind III 消化经扩增的 AIM II DNA 和载体 pQE9，然后将这些已消化的 DNA 连接在一起。在已经限制性切割的 pQE9 载体内插入 AIM II 蛋白 DNA 后，AIM II 蛋白编码区被置于载体内 IPTG 诱导型启动子的下游，并与其有效相连，而且，

AIM II 蛋白编码区的插入位置使其与适当放置以便于 AIM II 蛋白翻译的起始 AUG 密码子读框一致。

B. 具有 C - 末端 HA 标记的 AIM II 的表达

本实施例中使用细菌表达载体 pQE60 进行细菌表达。(QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311)。pQE60 编码氨苄青霉素抗生素抗性 (“ Amp^r ”), 并含有一个细菌复制起点 (“ ori ”)、一个 IPTG 诱导型启动子、一个核糖体结合位点 (“ RBS ”)、编码组氨酸残基的六个密码子, 以及适当的单一限制性酶切位点, 其中所述的组氨酸残基可允许通过利用 QIAGEN 公司 (公司地址同前) 所售镍 - 次氨基三乙酸 (“ Ni-NTA ”) 亲和树脂进行亲和纯化。这些元件排列的方式能保证编码多肽的 DNA 片段插入后所表达的多肽在其羧基末端可共价连接 6 个组氨酸残基 (即 “ 6x His 标记 ”)。

利用分别可与所需的 AIM II 蛋白区段的氨基末端序列及保藏构建体中 cDNA 编码序列的 3' 侧的序列可退火的 PCR 寡核苷酸引物, 从保藏 cDNA 克隆中扩增编码所需的 AIM II 蛋白区段的 DNA 序列。在 5' 和 3' 序列中分别添加了含有可便于克隆至 pQE60 载体的限制性位点的额外核苷酸。

为了克隆蛋白质, 所用 5' 引物的序列为 5' GACGC CCATGG AG GAG GAG AGT GTC GTA CGC C 3' (SEQ ID NO: 17), 此序列中含有由下划线标出的 NcoI 限制性位点, 其后为可与图 1A - C 中 AIM II 序列的氨基末端编码序列互补的核苷酸。当然, 本领域普通技术人员深知, 可改变蛋白质编码序列内 5' 引物起始位点, 以扩增编码完整蛋白质中任何所需区段的 DNA 片段 (更短或更长)。3' 引物的序列为 5' GACC GGATCC CAC CAT GAA AGC CCC GAA GTA AG 3' (SEQ ID NO: 18), 此序列中含有由下划线标出的 BamHI 限制性位点, 其后为可与图 1A - C 中 AIM II DNA 序列内终止密码子前的编码序列 3' - 末端互补的核苷酸, 编码序列与限制性位点排列的方式能确保其读框与 pQE60 载体内 6 个 His 密码子的读框一致。

利用 **BamHI** 和 **NcoI** 消化扩增的 **AIM II** DNA 片段及载体 **pQE60**，然后将已消化的 DNA 连接在一起。将 **AIM II** DNA 插入于已限制性消化的 **pQE60** 载体中使 **AIM II** 蛋白编码区置于 **IPTG** 诱导型启动子的下游，并与起始密码子 **AUG** 和其中的 6 个组氨酸密码子读框一致。

利用标准方法，将上述 A 或 B 中制成的含有 **HA** 标记的表达构建体的连接混合物转化至感受态大肠杆菌细胞中。**Sambrook** 等，《分子克隆：实验室手册》，第二版，冷泉港实验室出版社，冷泉港，纽约（1989）中记载了这类方法。将含有多拷贝质粒 **pREP4** 的大肠杆菌 **M15/rep4** 菌株用于进行本文所述的实施例，其中，所述质粒 **pREP4** 可表达 **lac** 阻遏子，并赋予卡那霉素抗性（“**Kan^r**”）。该菌株仅是可适用于表达 **AIM II** 蛋白的多种菌株中的一种，它可购自 **Qiagen** 公司。

通过在含有氨苄青霉素和卡那霉素的 **LB** 平板上的生长能力鉴定出转化体。从抗性菌落中提取质粒 DNA，再通过限制性酶切分析确证克隆 DNA 的同一性。

在添加有氨苄青霉素（100 μ g/ml）和卡那霉素（25 μ g/ml）的液体 **LB** 培养基中过夜（“**O/N**”）培养含有所需构建体的克隆。

按约 1：100 至 1：250 的稀释度，将 **O/N** 培养物接种至大量培养物中。培养细胞至 600nm 处的光密度值（“**OD600**”）为 0.4~0.6，然后加入异丙基- β -D-硫代半乳糖苷（“**IPTG**”）至终浓度为 1mM，以通过灭活 **lacI** 阻遏子而诱导从 **lac** 阻遏子敏感性启动子转录。随后进一步培养细胞 3 至 4 小时。然后，通过离心收集细胞，并通过标准方法裂解细胞。利用常规收集技术从裂解细胞内纯化包涵体，再将包涵体内的蛋白质溶解于 8M 尿素中。在 2 $\times$ 磷酸缓冲盐溶液（“**PBS**”）中将含有已溶解的蛋白质的 8M 尿素溶液过一个 **PD-10** 柱，由此去除尿素，更换缓冲液，并使蛋白质重新折叠。再通过另一个步骤-层析，对蛋白质进一步纯化以除去内毒素。然后，进行无菌过滤。将无菌过滤的蛋白质制备液以 95 μ /ml 的浓度保存于 2 $\times$ **PBS** 中。

C. 不含有 HA 标记的 AIM II 蛋白的表达和纯化

本实施例中使用细菌表达载体 pQE60 进行细菌表达。(QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311)。pQE60 编码氨苄青霉素抗生素抗性 (“ Amp^r ”), 并含有一个细菌复制起点 (“ ori ”)、一个 IPTG 诱导型启动子、一个核糖体结合位点 (“ RBS ”)、编码组氨酸残基的六个密码子, 以及适当的单一限制性酶切位点, 其中所述的组氨酸残基可允许通过利用 QIAGEN 公司 (公司地址同前) 所售镍 - 次氨基三乙酸 (“ Ni-NTA ”) 亲和树脂进行亲和纯化。这些元件排列的方式能保证编码多肽的 DNA 片段插入后所表达的多肽在其羧基末端可共价连接 6 个组氨酸残基 (即 “ 6x His 标记 ”)。然而, 此实施例中, 该多肽编码序列的插入方式使这六个组氨酸密码子不能被翻译, 因而产生的多肽并不含有 6 × His 标记。

利用分别可与所需的 AIM II 蛋白区段的氨基末端序列及保藏构建体中 cDNA 编码序列的 3' 侧的序列退火的 PCR 寡核苷酸引物, 从保藏 cDNA 克隆中扩增编码所需的缺乏疏水性前导序列的 AIM II 蛋白区段的 DNA 序列。在 5' 和 3' 序列中分别添加了含有可便于克隆至 pQE60 载体的限制性位点的额外核苷酸。

为了克隆蛋白质, 所用 5' 引物的序列为 5' GACGC CCATGG AG GAG GAG AGT GTC GTA CGC C 3' (SEQ ID NO: 17), 此序列中含有由下划线标出的 NcoI 限制性位点, 其后为可与图 1A - C 中 AIM II 序列的氨基末端编码序列互补的核苷酸。当然, 本领域普通技术人员深知, 可改变蛋白质编码序列内 5' 引物起始位点, 以扩增编码完整蛋白质中任何所需区段的 DNA 片段 (更短或更长)。3' 引物的序列为 5' CGC AAGCTT CCTT CAC ACC ATG AAA GC 3' (SEQ ID NO: 19), 此序列中含有由下划线标出的 Hind III 限制性位点, 其后为可与图 1A - C 中 AIM II DNA 序列内 3' 末端非编码序列互补的核苷酸。

利用 Hind III 和 NcoI 消化扩增获得的 AIM II DNA 片段及载体 pQE60, 然后将已消化的 DNA 连接在一起。将 AIM II DNA 插

入于已限制性消化的 pQE60 载体中使 AIM II 蛋白编码区及其终止密码子一起置于 IPTG 诱导型启动子的下游并与起始密码子 AUG 读框一致。该终止密码子能阻止插入位点下游的 6 个组氨酸密码子的翻译。

利用标准方法，如 Sambrook 等，《分子克隆：实验室手册》，第二版，冷泉港实验室出版社，冷泉港，纽约（1989）中所记载的那些方法，将连接混合物转化至感受态大肠杆菌细胞中。将含有多拷贝质粒 pREP4 的大肠杆菌 M15/rep4 菌株用于进行本文所述的实施例，其中，所述质粒 pREP4 可表达 lac 阻遏子，并赋予卡那霉素抗性（“Kan^r”）。此菌株仅是可适用于表达 AIM II 蛋白的多种菌株中的一种，它可购自 Qiagen 公司（地址同前）。通过在含有氨苄青霉素和卡那霉素的 LB 平板上的生长能力鉴定出转化体。从抗性菌落中提取质粒 DNA，再通过限制性酶切分析、PCR 和 DNA 测序确证克隆 DNA 的同一性。

在补加有氨苄青霉素（100 μ g/ml）和卡那霉素（25 μ g/ml）的液体 LB 培养基中过夜（“O/N”）培养含有所需构建体的克隆。再按大约 1：25 至 1：250 的稀释度，将 O/N 培养物接种至大量培养物中。培养细胞至 600nm 处的光密度值（“OD₆₀₀”）为 0.4~0.6，然后加入异丙基- β -D-硫代半乳糖苷（“IPTG”）至终浓度为 1mM，以通过灭活 lacI 阻遏子而诱导从 lac 阻遏子敏感性启动子转录。随后进一步培养细胞 3 至 4 小时，再通过离心收集细胞。

然后，在 6M 盐酸胍、pH8 中，于 4℃ 将细胞匀浆 3 - 4 小时。通过离心去除细胞碎片，在添加有 200mM NaCl 的 50mM 乙酸钠缓冲液 pH6 中透析含有 AIM II 的上清液。或者，通过在含有蛋白酶抑制剂的 500mM NaCl、20 % 甘油、25mM Tris · HCl pH7.4 中透析上清液，可成功地对蛋白质进行重新折叠。复性后，蛋白质可通过离子交换、疏水性相互作用和尺寸排阻层析进行纯化。或者，可通过亲和层析步骤，如利用抗体柱以获得纯的 AIM II 蛋白。将已纯化的蛋白质贮于 4℃，或冻于 - 80℃。

实施例 2：在杆状病毒表达系统内克隆和表达 AIM II 蛋白

利用对应于 AIM II 基因 5'和 3'序列的 PCR 寡核苷酸引物扩增保藏克隆中编码全长 AIM II 蛋白的 cDNA 序列：

5'引物的序列为 5' GCT CCA GGA TCC GCC ATC ATG GAG GAG AGT GTC GTA CGG C 3' (SEQ ID NO: 1)，它含有由下划线标出的 BamHI 限制性酶切位点及随后的图 1A - C 中 AIM II 蛋白序列内的 22 个碱基。如下所述，插入至表达载体后，编码 AIM II 的扩增片段的 5'末端提供了一种有效的信号肽。如 Kozak M.,分子生物学杂志, 196: 947 - 950 (1987) 所记载的在真核细胞内用于翻译起始的有效信号被置于构建体中载体区段内的适当位置；

3'引物的序列为 5' GA CGC GGT ACC GTC CAA TGC ACC ACG CTC CTT CCT TC 3' (SEQ ID NO: 12)，它含有由下划线标出的 Asp 718 限制性位点及随后的与图 1A - C 中 AIM II 内 769 - 795 位核苷酸互补的核苷酸。

利用可商购的试剂盒(“ Geneclean”, BIO 101 Inc., La Jolla, Ca.), 从 1 % 琼脂糖凝胶中分离扩增片段。然后利用 BamHI 和 Asp 718 消化该片段，再在 1 % 琼脂糖凝胶上对它进行纯化。此处将该片段命名为 F2。

如 Summers 等，杆状病毒载体方法和昆虫细胞培养技术手册，Texas Agricultural Experimental Station Bulletin No. 1555 (1987)所述，通过标准方法，利用载体 pA2-GP 在杆状病毒表达载体中表达 AIM II 蛋白。此表达载体中含有苜蓿银纹夜蛾 (Autographa californica) 核型多角体病毒 (AcMNPV) 的强多角体蛋白基因启动子，其后还含有便利的限制性位点。包括 N - 末端甲硫氨酸在内的 AcMNPVgp67 的信号肽恰好位于 BamHI 位点上游。此载体中还使用了猴病毒 40 (“ SV40 ”) 的聚腺苷酸化位点以有效完成聚腺苷酸化。为能简单进行重组病毒的筛选，此载体中还与多角体蛋白基因启动子方向一致地插入了大肠杆菌来源的 β - 半乳糖苷酶基因，随后是多角体蛋白基因的多聚腺苷酸化信

号。多角体蛋白基因序列的两侧都含有病毒序列，以便与野生型病毒 DNA 进行细胞介导的同源重组，从而产生能表达此克隆多核苷酸的活性病毒。

如本领域技术人员所知，只要构建过程能提供所需的恰当排列的用于转录翻译、运输等的信号，如读框一致的 AUG 和信号肽，则可利用其它多种杆状病毒载体代替 pA2-GP。这类载体与其它载体一起记载于 Luckow 等，病毒学，170：31 - 39。

通过本领域常规技术，利用限制性酶 BamHI 和 Asp 718 消化该质粒，再利用小牛肠道磷酸酯酶对其进行去磷酸化。接着利用可商购的试剂盒（“Geneclean” BIO 101 Inc., La Jolla, Ca），从 1% 琼脂糖凝胶分离此 DNA。此处将该载体 DNA 命名为“V”。

通过 T4 DNA 连接酶将片段 F2 和已去磷酸化的质粒 V2 连接在一起。利用连接混合物转化大肠杆菌 HB101 细胞，再平铺于培养平板上。经 XbaI 消化各个菌落来源的 DNA，再经凝胶电泳分析消化产物，鉴定出含有携带人 AIM II 基因的质粒的细菌。通过 DNA 测序确证克隆片段的序列。此处将该质粒命名为 pBac AIM II。

利用 Felgner 等，美国国家科学院院报，84：7413 - 7417（1987）所述的脂质转染法，将 5 μ g 质粒 pBac AIM II 与 1.0 μ g 可商购的线性化杆状病毒 DNA（“Baculo GoldTM baculovirus DNA”，Pharmlngen, San Diego, CA.）一起进行共转染。在微滴定板上含有 50 μ l 无血清的 Grace's 培养基（Life Technologies Inc., Gaithersburg, MD）的无菌孔中，混合 1 μ g BaculoGoldTM 病毒 DNA 和 5 μ g 质粒 pBac AIM II，然后加入 10 μ l Lipofectin 和 90 μ l Grace's 培养基，混合后于室温放置 15 分钟。然后将转染混合物逐滴加入至接种于含有 1ml 不含血清的 Grace's 培养基的 35mm 组织培养板上的 Sf9 昆虫细胞（ATCC CRL 1711）。前后摇动此板以混合新加入的溶液，然后将板在 27℃ 培养 5 小时。5 小时后，从板中除去转染溶液，再加入 1ml 补充有 10% 胎牛血清的 Grace's 昆虫培养基。将板放回至培养箱中，于 27℃ 继续培养 4 天。

4 天后，收集上清，并按照上文列举的 Summers 和 Smith 所

述的方法进行噬斑测定。利用含有“Blue Gal”(Life technologies Inc., Gaithersburg)的琼脂糖凝胶以便对表达半乳糖苷酶的克隆进行简便的鉴定和分离,其中该表达半乳糖苷酶的克隆能产生染成蓝色的噬斑(在Life Technologies Inc., Gaithersburg所提供的昆虫细胞培养和杆状病毒学使用说明第9-10页中也可见到此类“噬斑测定”的详细说明)。

连续稀释4天后,将病毒加入至细胞中。进行适当培养后,利用Eppendorf取液器的吸头挑出染成蓝色的噬斑,然后将含有重组病毒的琼脂重新悬浮于含有200 μ l Grace's培养基的Eppendorf管中。通过轻微离心去除其中的琼脂,再利用含有重组杆状病毒的上清液感染接种于35mm板的Sf9细胞。4天后,收集这些培养板的上清液,将它们贮于4 $^{\circ}$ C。通过限制性酶谱测定和测序等DNA分析方法鉴定出含有正确插入的hESSBI、II和III的克隆。本文将此克隆命名为V-AIM II。

在补充有10%加热灭活FBS的Grace's培养基中培养Sf9细胞。利用重组杆状病毒V-AIM II,以2左右(1左右~3左右)的感染复数(“MOI”)感染该细胞。6小时后,除去培养基,再换入无甲硫氨酸和半胱氨酸的SF900 II培养基(可从Life Technologies Inc., Gaithersburg获得)。42小时后,加入5 μ Ci 35 S-甲硫氨酸和5 μ Ci 35 S-半胱氨酸(可从Amersham获得)。继续培养细胞16小时,然后经离心收集细胞,进行细胞裂解,并通过SDS-PAGE和放射自显影观察已被标记的蛋白质。

实施例3:哺乳动物细胞中的克隆和表达

用于在哺乳动物细胞内瞬时表达AIM II蛋白的绝大部分载体都应携有SV 40复制起始点。这有助于在能表达病毒DNA合成起始所需的T抗原的细胞(如COS细胞)内确保载体复制至高拷贝数。为达此目的,也可使用其它任何的哺乳动物细胞系。

一个典型的哺乳动物表达载体含有可介导mRNA转录起始的启动子元件、蛋白质编码序列,以及转录终止、转录本聚腺苷酸化所需的信号。其它的元件包括增强子、Kozak序列及侧翼含有RNA

剪接所需的供体和受体位点的间插序列。利用 SV40 的早期、晚期启动子、逆转录病毒如 RSV、HTLVI、HIVI 的长末端重复区 (LTRs) 及巨细胞病毒 (CMV) 的早期启动子, 可获得高效转录。然而, 也可使用细胞信号 (如人肌动蛋白启动子)。可用于完成本发明的适当表达载体包括 pSVL、pMSG (Pharmacia, Uppsala, Sweden)、pRSVcat(ATCC 37152)、pSV2dhfr(ATCC 37146)和 pBC12MI(ATCC 67109)等载体。可用的哺乳动物细胞包括人 Hela 细胞、283 细胞、H9 细胞和 Jurkat 细胞、小鼠 NIH3T3 细胞和 C127 细胞, Cos1 细胞, Cos7 细胞和 CV1 细胞, 非洲绿猴细胞、鹤鹑 QC1 - 3 细胞, 小鼠 L 细胞和中国仓鼠卵巢细胞。

或者, 也可在染色体整合了该基因的稳定细胞系中表达该基因。通过与选择性标记如 dhfr、gpt、新霉素、潮霉素基因一起共转染可鉴定和分离出已转染的细胞。

也可通过扩增转染基因以大量表达它所编码的蛋白质。在开发携有数百或甚至数千拷贝的感兴趣基因的细胞系时, DHFR (二氢叶酸还原酶) 是一个很有用的标记。另一种有效的选择性标记是谷氨酰胺合成酶 (GS) (Murphy 等, 生物化学杂志, 227: 277 - 279 (1991); Bebbington 等, Bio/Technology 10: 169-175 (1992))。利用这些标记, 在选择性培养基中培养这些哺乳动物细胞, 筛选出具有最高抗性的细胞。这些细胞系中含有已整合至染色体内的扩增基因。中国仓鼠卵巢 (CHO) 细胞常被用来生产蛋白质。

表达载体 pC1 和 pC4 含有添加了 CMV 增强子片段 (Boshart 等, 细胞, 41:521-530 (1985)) 的劳氏肉瘤病毒强启动子 (LTR) (Cullen 等, 分子和细胞生物学, 438 - 447 (1985 年 3 月))。多克隆位点, 如含有 BamHI、XbaI 和 Asp 718 限制性酶切位点等, 可便于克隆感兴趣基因。载体中还含有 3' 内含子, 大鼠前胰岛素原基因的多聚腺苷酸化和终止信号。

实施例 3(a): COS 细胞内的克隆和表达

通过将编码 AIM II 的 cDNA 克隆至表达载体 pcDNA I/Amp

(可从 Invitrogen Inc. 获得) 中可制成表达质粒 pAIM II HA。

此表达载体 pcDNA I/amp 含有: (1) 可有效用于大肠杆菌和其它原核细胞内扩增的大肠杆菌复制起点; (2) 可用于含有质粒的原核细胞筛选的氨苄青霉素抗性基因; (3) 可用于真核细胞内扩增的 SV40 复制起点; (4) CMV 启动子、多位点接头、SV40 内含子和聚腺苷酸化信号, 这些序列的排列方式可保证通过多位点接头内的限制性位点便利地将 cDNA 置于 CMV 启动子的表达调控下, 并与 SV40 内含子和聚腺苷酸化信号有效相连。

将编码 AIM II 蛋白及在其 3' 末端读框一致地融合了 HA 标记的 DNA 片段克隆于载体的多位点接头区, 以便在 CMV 启动子的指导下进行重组蛋白质表达。此 HA 标记对应于 Wilson 等, 细胞, 37: 767 (1984) 所述的来源于流感血凝素蛋白的一个表位。靶蛋白内 HA 标记的融合可便于通过识别 HA 表位的抗体简便地检测重组蛋白质。

质粒构建的策略如下。与上述构建在大肠杆菌内表达 AIM II 的表达载体的过程相似, 利用含有便利的限制性位点的引物, 扩增出保藏克隆内的 AIM II cDNA。为便于对表达的 AIM II 进行检测、纯化和鉴定, 引物之一含有上述的血凝素标记(“HA 标记”)。

适当的引物有此实施例中所使用的下述引物: 含有下划线标出的 BamHI 位点和 AUG 起始密码子的 5' 引物, 其序列如下:

5' G CTC GGA TCC GCC ATC ATG 3' (SEQ ID NO: 13);

含有下划线标出的 XbaI 位点、终止密码子、可形成血凝素 HA 标记的 9 个密码子以及 31bp 的 3' 编码序列(在其 3' 末端)的 3' 引物, 其序列如下:

5' GAT GTT CTA GAA AGC GTA GTC TGG GAC GTC GTA TGG GTA CAC CAT GAA AGC CCC GAA GTA AGA CCG GGT AC 3' (SEQ ID NO: 14)。

通过 HindIII 和 XhoI 消化 PCR 扩增 DNA 片段和载体 pcDNA I/Amp, 再将它们连接起来。将连接混合物转化至大肠杆菌 SURE

菌株(可从 Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA92037 获得)中, 将已转化的培养物铺板于含氨苄青霉素培养基的平板上, 再培养这些板以确保氨苄青霉素抗性菌落的生长。从抗性菌落中分离质粒 DNA, 并通过限制性分析和凝胶电泳检测质粒 DNA 中 AIM II 编码片段的存在。

为了表达重组 AIM II, 可利用如 Sambrook 等, 《分子克隆: 实验室手册》, 冷泉港实验室出版社, 冷泉港, 纽约 (1989) 所述的 DEAE - DEXTRAN, 通过上述的表达载体转染 COS 细胞。在适于载体表达 AIM II 的条件下培养细胞。

利用如 Harlow 等, 《抗体: 实验室手册》, 第二版, 冷泉港实验室出版社, 冷泉港, 纽约 (1988) 中所记载的方法, 通过放射性标记和免疫沉淀测定 AIM II HA 融合蛋白的表达。为达此目的, 转染 2 天后, 通过在含有 ^{35}S - 半胱氨酸的培养基中培养细胞而对其进行标记。收集细胞和培养基, 洗涤细胞, 再按照 Wilson 等 (出处同前) 所述的方法, 利用含有去污剂的 RIPA 缓冲液: 150mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50mM Tris, pH7.5 裂解细胞。利用 HA 特异性单克隆抗体, 从细胞裂解物和组织培养基中沉淀蛋白质。再通过 SDS - PAGE 凝胶和放射显影分析已沉淀的蛋白质。在细胞裂解物中可观察到预期大小的表达产物, 这在负对照中是不存在的。

实施例 3 (b): CHO 细胞内的克隆和表达

利用载体 pC4 表达 AIM II 蛋白。质粒 pC1 是质粒 pSV2-dhfr [ATCC 保藏号 37146] 的衍生物。两种质粒都含有在 SV40 早期启动子控制下的小鼠 DHFR 基因。通过在补充有化疗剂氨甲喋呤的选择性培养基 (alpha minus MEM, Life Technologies) 中培养细胞可筛选出转染有这些质粒的无二氢叶酸活性的中国仓鼠卵巢细胞或其它细胞。抗氨甲喋呤 (MTX) 的细胞内 DHFR 基因的扩增已有很好的记载 (参见如 Alt F. W., Kellems R. M., Bertino J. R. 和 Schimke R. T., 1978, 生物化学杂志, 253: 1357 - 1370; Hamlin J. L. 和 Ma C., 1990, Biochem et Biophys. Acta, 1097:

107-143; Page M. J.和 Sydenham M. A., 1991, 生物技术, 第 9 卷: 64 - 68)。生长于更高浓度的 MTX 的细胞能通过 DHFR 基因的扩增、过量表达靶酶 DHFR 而增强其对此药物的抗性。如果 DHFR 基因连接了另一个基因, 通常它也被共扩增和过量表达。培育出含有 1000 多个拷贝的基因的细胞系是本领域的重要工作。因此, 撤去氨甲喋呤后, 细胞系中仍含有已整合至染色体中的扩增基因。

质粒 pC4 中含有用于表达目的基因的劳氏肉瘤病毒长末端重复区 (LTR) 的强启动子 (Cullen 等, 分子和细胞生物学, 1985 年 3 月: 438 - 447) 以及从人巨细胞病毒 (CMV) 立即早期基因的增强子中分离出的一个片段 (Boshart 等, 细胞, 41: 521 - 530 (1985))。该启动子下游含有可允许基因插入的 BamHI、XbaI 和 Asp 718 限制性酶切位点。此质粒在这些克隆位点后还含有大鼠前胰岛素原基因的 3' 内含子和聚腺苷酸化位点。也可使用其它高效启动子进行表达, 如人 β -肌动蛋白基因启动子、SV40 的早期或晚期启动子, 或其它逆转录病毒如 HIV 和 HTLV I 的长末端重复区。可利用 Clontech 公司的 Tet-Off 和 Tet-On 基因表达系统在哺乳动物细胞内以调控方式表达 AIM II (Gossen M.和 Bujard H., 1992, 美国国家科学院院报, 89: 5547 - 5551)。至于该 mRNA 的聚腺苷酸化, 也可使用其它的信号, 如来源于人生长激素或珠蛋白基因的信号。通过与选择性标记如 gpt、G418 或潮霉素共转化, 也可筛选出携有已整合至染色体内的目的基因的稳定细胞系。最好是开始时使用多个选择性标记, 如同时使用 G418 和氨甲喋呤。

利用限制性酶 BamHI 和 Asp 718 消化质粒 pC4, 然后通过本领域已知方法, 利用小牛肠道磷酸酶对其进行去磷酸化处理。再通过 1% 琼脂糖凝胶分离该载体。

利用分别对应于该基因的 5' 和 3' 序列的 PCR 寡核苷酸引物扩增编码完整 AIM II 蛋白及其前导序列的 DNA 序列。该 5' 引物的序列为: 5' GCT CCA GGA TCC GCC ATC ATG GAG GAG AGT GTC GTA CGG C 3' (SEQ ID NO: 15), 其中含有下划线表示

的 BamHI 限制性酶切位点和其后的如 Kozak M., 分子生物学杂志, 196: 947 - 950 (1987) 所述的真核生物内用于翻译起始的有效信号, 以及图 1A - C (SEQ ID NO: 1) 所示的 AIM II 编码序列内的 22 个碱基。3'引物的序列为: 5' GA CGC GGT ACC GTC CAA TGC ACC ACG CTC CTT CCT TC 3' (SEQ ID NO: 16), 其中含有下划线表示的 Asp 718 限制性位点及其后的与图 1A - C (SEQ ID NO: 1) 所示 AIM II 基因第 769 - 795 位核苷酸互补的核苷酸。

该扩增片段经内切核酸酶 BamHI 和 Asp 718 消化后, 再在 1% 琼脂糖凝胶上重新纯化。然后通过 T4 DNA 连接酶将该分离片段和去磷酸化载体连接起来, 再转化大肠杆菌 HB101 或 XL1 - Blue 细胞, 并通过如限制性酶切分析鉴定出含有已插入至质粒 pC4 的该片段的细菌。

使用缺乏活性 DHFR 基因的中国仓鼠卵巢细胞进行转化。利用 Lipofectin (Felgner 等, 出处同前), 将 5 μ g 表达质粒 pC4 与 0.5 μ g 质粒 pSV2-neo 一起进行共转化。质粒 pSV2 neo 含有一个显性选择性标记, 即 Tn5 来源的 neo 基因, 该 neo 基因编码可赋予对包括 G418 在内的一组抗生素的抗性的酶类。将该细胞接种于补加了 1mg/ml G418 的 alpha minus MEM 培养基中。2 天后, 通过胰蛋白酶消化该细胞, 并将其接种于含有同时补加了 10.25 或 50ng/ml 氨甲喋呤和 1mg/ml G418 的 alpha minus MEM 培养基的杂交瘤克隆板 (Greiner, Germany) 中。大约 10 - 14 天后, 经胰蛋白酶消化单个克隆, 然后将它们接种于 6 - 孔培养皿或 10ml 锥形瓶中, 并使用不同浓度的氨甲喋呤 (50nM、100nM、200nM、400nM、800nM)。将在最高浓度氨甲喋呤中生长的克隆转移至含有更高浓度氨甲喋呤 (1 μ M、2 μ M、5 μ M、10mM、20mM) 的新 6 - 孔板上。重复此步骤直至获得能在 100 ~ 200 μ M 浓度下生长的克隆。通过如 SDS - PAGE、Western 杂交或反相 HPLC 分析等方法, 对目的基因产物的表达进行分析。

实施例 4: AIM II 蛋白表达的组织分布

利用 Sambrook 等人（出处同前）及其它文献所记载的方法进行 Northern 印迹分析，以研究人组织中 AIM II 基因的表达。按照生产商说明，利用 rediprime™ DNA 标记系统（Amersham Life Science）对含有 AIM II 蛋白全部核苷酸序列（SEQ ID NO: 1）的 cDNA 探针进行 ³²P 标记。标记后，按照生产商 PT 1200 - 1 号技术，利用 CHROMA SPIN - 100™ 柱（Clontech Laboratories 公司）对该探针进行纯化。然后利用该经纯化的标记探针分析各种人组织中 AIM II mRNA 的水平。

从 Clontech 获得含有多个人组织（H）或人免疫系统组织（IM）的多组织 Northern (Multiple Tissue Northern, MTN) 印迹，并按照生产商的 PT 1190 - 1 号技术方案，利用 ExpressHyb™ 杂交溶液（Clontech），通过标记探针对这些印迹进行分析。杂交和洗涤后，固定印迹并在 - 70 °C 对胶片曝光过夜，再按照标准方法对胶片进行显影。

显然，可通过除本文详细叙述的说明书和实施例外的其它方式操作本发明。

在上文的指导下，可对本发明进行多种改进和改变，因此，它们也在本文附加权利要求的范围内。

本文所列举的所有出版物（包括专利、专利申请、杂志文章、实验室手册、书籍或其它文件）的全部公开内容均引入本文作为参考。

序列表

(1) 一般信息:

- (i) 申请人: 人类基因组科学公司
9410Key West Avenue
Rockville, MD 20850
美国
申请人 / 发明人: Ebner, Reinhard
Yu, Guo- Liang
Ruben, Steven M.
- (ii) 发明名称: 编程性细胞死亡诱导分子 II
- (iii) 序列数: 19
- (iv) 通讯地址:
 - (A)收件人: Sterne, Kessler, Goldstein & Fox, P. L. L.C
 - (B)街道: 1100 New York Ave., Suite 600
 - (C)城市: Washington
 - (D)州: DC
 - (E)国家: 美国
 - (F)邮编: 20005 - 3934
- (v) 计算机可读形式:
 - (A)介质类型: Floppy disk
 - (B)计算机: IBM PC 兼容机
 - (C)操作系统: PC- DOS/MS- DOS
 - (D)软件: Patent In Release #1.0, Version#1.30
- (vi) 目前的申请信息
 - (A)申请号: TBA
 - (B)申请日: 本文所附
 - (C)分类:
- (vii) 在先申请信息:
 - (A)申请号: 60/ 013, 923
 - (B)申请日: 1996 年 3 月 22 日
 - (C)分类:
- (viii) 代理人/代理机构信息:

(A)姓名: Goldstein, Jorge A.

(B)登记号: 29, 021

(C)资料/文档号: 1488.065 PC 01/JAG/EKS/KMT

(ix) 通讯信息:

(A)电话: 202 - 371 - 2600

(B)电传: 202 - 371 - 2540

(2) SEQ ID NO: 1 信息:

(i)序列特征:

(A)长度: 1169 个碱基对

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(ix)特性:

(A)名称/关键词: CDS

(B)位置: 49..768

(xi)序列描述: SEQ ID NO: 1:

```

GAGGTTGAAG GACCCAGGCG TGTCAGCCCT GCTCCAGAGA CCTTGGGC ATG GAG GAG      57
                                   Met Glu Glu
                                   1

AGT GTC GTA CGG CCC TCA GTG TTT GTG GTG GAT GGA CAG ACC GAC ATC      105
Ser Val Val Arg Pro Ser Val Phe Val Val Asp Gly Gln Thr Asp Ile
    5                10                15

CCA TTC ACG AGG CTG GGA CGA AGC CAC CGG AGA CAG TCG TGC AGT GTG      153
Pro Phe Thr Arg Leu Gly Arg Ser His Arg Arg Gln Ser Cys Ser Val
    20                25                30                35

GCC CGG GTG GGT CTG GGT CTC TTG CTG TTG CTG ATG GGG GCT GGG CTG      201
Ala Arg Val Gly Leu Gly Leu Leu Leu Leu Leu Met Gly Ala Gly Leu
                40                45                50

GCC GTC CAA GGC TGG TTC CTC CTG CAG CTG CAC TGG CGT CTA GGA GAG      249
Ala Val Gln Gly Trp Phe Leu Leu Gln Leu His Trp Arg Leu Gly Glu
                55                60                65

ATG GTC ACC CGC CTG CCT GAC GGA CCT GCA GGC TCC TGG GAG CAG CTG      297
Met Val Thr Arg Leu Pro Asp Gly Pro Ala Gly Ser Trp Glu Gln Leu
                70                75                80

```

ATA CAA GAG CGA AGG TCT CAC GAG GTC AAC CCA GCA GCG CAT CTC ACA	345
Ile Gln Glu Arg Arg Ser His Glu Val Asn Pro Ala Ala His Leu Thr	
85 90 95	
GGG GCC AAC TCC AGC TTG ACC GGC AGC GGG GGG CCG CTG TTA TGG GAG	393
Gly Ala Asn Ser Ser Leu Thr Gly Ser Gly Gly Pro Leu Leu Trp Glu	
100 105 110 115	
ACT CAG CTG GGC CTG GCC TTC CTG AGG GGC CTC AGC TAC CAC GAT GGG	441
Thr Gln Leu Gly Leu Ala Phe Leu Arg Gly Leu Ser Tyr His Asp Gly	
120 125 130	
GCC CTT GTG GTC ACC AAA GCT GGC TAC TAC TAC ATC TAC TCC AAG GTG	489
Ala Leu Val Val Thr Lys Ala Gly Tyr Tyr Tyr Ile Tyr Ser Lys Val	
135 140 145	
CAG CTG GGC GGT GTG GGC TGC CCG CTG GGC CTG GCC AGC ACC ATC ACC	537
Gln Leu Gly Gly Val Gly Cys Pro Leu Gly Leu Ala Ser Thr Ile Thr	
150 155 160	
CAC GGC CTC TAC AAG CGC ACA CCC CGC TAC CCC GAG GAG CTG GAG CTG	585
His Gly Leu Tyr Lys Arg Thr Pro Arg Tyr Pro Glu Glu Leu Glu Leu	
165 170 175	
TTG GTC AGC CAG CAG TCA CCC TGC GGA CGG GCC ACC AGC AGC TCC CGG	633
Leu Val Ser Gln Gln Ser Pro Cys Gly Arg Ala Thr Ser Ser Ser Arg	
180 185 190 195	
GTC TGG TGG GAC AGC AGC TTC CTG GGT GGT GTG GTA CAC CTG GAG GCT	681
Val Trp Trp Asp Ser Ser Phe Leu Gly Gly Val Val His Leu Glu Ala	
200 205 210	
GGG GAG GAG GTG GTC GTC CGT GTG CTG GAT GAA CGC CTG GTT CGA CTG	729
Gly Glu Glu Val Val Val Arg Val Leu Asp Glu Arg Leu Val Arg Leu	
215 220 225	
CGT GAT GGT ACC CGG TCT TAC TTC GGG GCT TTC ATG GTG TGAAGGAAGG	778
Arg Asp Gly Thr Arg Ser Tyr Phe Gly Ala Phe Met Val	
230 235 240	
AGCGTGGTGC ATTGGACATG GGTCTGACAC GTGGAGAACT CAGAGGGTGC CTCAGGGGAA	838
AGAAAACTCA CGAAGCAGAG GCTGGGCGTG GTGGCTCTCG CCTGTAATCC CAGCACTTTG	898
GGAGGCCAAG GCAGGCGGAT CACCTGAGGT CAGGAGTTCTG AGACCAGCCT GGCTAACATG	958
GCAAAAACCCC ATCTCTACTA AAAATACAAA AATTAGCCGG ACGTGGTGGT GCCTGCCTGT	1018
AATCCAGCTA CTCAGGAGGC TGAGGCAGGA TAATTTTGCT TAAACCCGGG AGGCGGAGGT	1078
TGCAGTGAGC CGAGATCACA CCACTGCACT CCAACCTGGG AAACGCAGTG AGACTGTGCC	1138
TCAAAAAAAAA AAAAAAAAA AAAAAAAAA A	1169

(2) SEQ ID NO: 2 信息:

(i)序列特征:

(A)长度: 240 个氨基酸

(B)类型: 氨基酸

(D)拓扑结构: 线性

(ii)分子类型: 蛋白质

(xi)序列描述: SEQ ID NO: 2:

```

Met Glu Glu Ser Val Val Arg Pro Ser Val Phe Val Val Asp Gly Gln
 1             5             10             15

Thr Asp Ile Pro Phe Thr Arg Leu Gly Arg Ser His Arg Arg Gln Ser
      20             25             30

Cys Ser Val Ala Arg Val Gly Leu Gly Leu Leu Leu Leu Leu Met Gly
      35             40             45

Ala Gly Leu Ala Val Gln Gly Trp Phe Leu Leu Gln Leu His Trp Arg
      50             55             60

Leu Gly Glu Met Val Thr Arg Leu Pro Asp Gly Pro Ala Gly Ser Trp
      65             70             75             80

Glu Gln Leu Ile Gln Glu Arg Arg Ser His Glu Val Asn Pro Ala Ala
      85             90             95

His Leu Thr Gly Ala Asn Ser Ser Leu Thr Gly Ser Gly Gly Pro Leu
      100            105            110

Leu Trp Glu Thr Gln Leu Gly Leu Ala Phe Leu Arg Gly Leu Ser Tyr
      115            120            125

His Asp Gly Ala Leu Val Val Thr Lys Ala Gly Tyr Tyr Tyr Ile Tyr
      130            135            140

Ser Lys Val Gln Leu Gly Gly Val Gly Cys Pro Leu Gly Leu Ala Ser
      145            150            155            160

Thr Ile Thr His Gly Leu Tyr Lys Arg Thr Pro Arg Tyr Pro Glu Glu
      165            170            175

Leu Glu Leu Leu Val Ser Gln Gln Ser Pro Cys Gly Arg Ala Thr Ser
      180            185            190

Ser Ser Arg Val Trp Trp Asp Ser Ser Phe Leu Gly Gly Val Val His
      195            200            205

Leu Glu Ala Gly Glu Glu Val Val Val Arg Val Leu Asp Glu Arg Leu
      210            215            220

Val Arg Leu Arg Asp Gly Thr Arg Ser Tyr Phe Gly Ala Phe Met Val
      225            230            235            240

```

(2) SEQ ID NO: 3 信息:**(i)序列特征:****(A)长度: 455 个氨基酸****(B)类型: 氨基酸****(C)链型: 不相关****(D)拓扑结构: 不相关****(ii)分子类型: 蛋白质****(xi)序列描述: SEQ ID NO: 3:**

```

Met Gly Leu Ser Thr Val Pro Asp Leu Leu Pro Leu Val Leu Leu
1           5           10           15

Glu Leu Leu Val Gly Ile Tyr Pro Ser Gly Val Ile Gly Leu Val Pro
20           25           30

His Leu Gly Asp Arg Glu Lys Arg Asp Ser Val Cys Pro Gln Gly Lys
35           40           45

Tyr Ile His Pro Gln Asn Asn Ser Ile Cys Cys Thr Lys Cys His Lys
50           55           60

Gly Thr Tyr Leu Tyr Asn Asp Cys Pro Gly Pro Gly Gln Asp Thr Asp
65           70           75           80

Cys Arg Glu Cys Glu Ser Gly Ser Phe Thr Ala Ser Glu Asn His Leu
85           90           95

Arg His Cys Leu Ser Cys Ser Lys Cys Arg Lys Glu Met Gly Gln Val
100          105          110

Glu Ile Ser Ser Cys Thr Val Asp Arg Asp Thr Val Cys Gly Cys Arg
115          120          125

Lys Asn Gln Tyr Arg His Tyr Trp Ser Glu Asn Leu Phe Gln Cys Phe
130          135          140

Asn Cys Ser Leu Cys Leu Asn Gly Thr Val His Leu Ser Cys Gln Glu
145          150          155          160

Lys Gln Asn Thr Val Cys Thr Cys His Ala Gly Phe Phe Leu Arg Glu
165          170          175

Asn Glu Cys Val Ser Cys Ser Asn Cys Lys Lys Ser Leu Glu Cys Thr
180          185          190

Lys Leu Cys Leu Pro Gln Ile Glu Asn Val Lys Gly Thr Glu Asp Ser
195          200          205

Gly Thr Thr Val Leu Leu Pro Leu Val Ile Phe Phe Gly Leu Cys Leu
210          215          220

```

Leu Ser Leu Leu Phe	Ile Gly Leu Met Tyr Arg Tyr Gln Arg Trp Lys	
225	230	235 240
Ser Lys Leu Tyr Ser	Ile Val Cys Gly Lys Ser Thr Pro Glu Lys Glu	
245	250	255
Gly Glu Leu Glu Gly Thr Thr Thr	Lys Pro Leu Ala Pro Asn Pro Ser	
260	265	270
Phe Ser Pro Thr Pro Gly Phe Thr Pro Thr Leu Gly Phe Ser Pro Val		
275	280	285
Pro Ser Ser Thr Phe Thr Ser Ser Ser Thr Tyr Thr Pro Gly Asp Cys		
290	295	300
Pro Asn Phe Ala Ala Pro Arg Arg Glu Val Ala Pro Pro Tyr Gln Gly		
305	310	315 320
Ala Asp Pro Ile Leu Ala Thr Ala Leu Ala Ser Asp Pro Ile Pro Asn		
325	330	335
Pro Leu Gln Lys Trp Glu Asp Ser Ala His Lys Pro Gln Ser Leu Asp		
340	345	350
Thr Asp Asp Pro Ala Thr Leu Tyr Ala Val Val Glu Asn Val Pro Pro		
355	360	365
Leu Arg Trp Lys Glu Phe Val Arg Arg Leu Gly Leu Ser Asp His Glu		
370	375	380
Ile Asp Arg Leu Glu Leu Gln Asn Gly Arg Cys Leu Arg Glu Ala Gln		
385	390	395 400
Tyr Ser Met Leu Ala Thr Trp Arg Arg Arg Thr Pro Arg Arg Glu Ala		
405	410	415
Thr Leu Glu Leu Leu Gly Arg Val Leu Arg Asp Met Asp Leu Leu Gly		
420	425	430
Cys Leu Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro Ala Ala Leu Pro		
435	440	445
Pro Ala Pro Ser Leu Leu Arg		
450	455	

(2) SEQ ID NO : 4 信息:**(i)序列特征:****(A)长度: 205 个氨基酸****(B)类型: 氨基酸****(C)链型: 不相关****(D)拓扑结构: 不相关****(ii)分子类型: 蛋白质****(xi)序列描述: SEQ ID NO : 4:**

```

Met Thr Pro Pro Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr
1          5          10          15
Leu His Leu Leu Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala
20          25          30
Gln Gly Leu Pro Gly Val Gly Leu Thr Pro Ser Ala Ala Gln Thr Ala
35          40          45
Arg Gln His Pro Lys Met His Leu Ala His Ser Thr Leu Lys Pro Ala
50          55          60
Ala His Leu Ile Gly Asp Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg
65          70          75          80
Ala Asn Thr Asp Arg Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn
85          90          95
Asn Ser Leu Leu Val Pro Thr Ser Gly Ile Tyr Phe Val Tyr Ser Gln
100         105         110
Val Val Phe Ser Gly Lys Ala Tyr Ser Pro Lys Ala Thr Ser Ser Pro
115         120         125
Leu Tyr Leu Ala His Glu Val Gln Leu Phe Ser Ser Gln Tyr Pro Phe
130         135         140
His Val Pro Leu Leu Ser Ser Gln Lys Met Val Tyr Pro Gly Leu Gln
145         150         155         160
Glu Pro Trp Leu His Ser Met Tyr His Gly Ala Ala Phe Gln Leu Thr
165         170         175
Gln Gly Asp Gln Leu Ser Thr His Thr Asp Gly Ile Pro His Leu Val
180         185         190
Leu Ser Pro Ser Thr Val Phe Phe Gly Ala Phe Ala Leu
195         200         205

```

(2) SEQ ID NO : 5 信息:**(i)序列特征:****(A)长度: 205 个氨基酸****(B)类型: 氨基酸****(C)链型: 不相关****(D)拓扑结构: 不相关****(ii)分子类型: 蛋白质****(xi)序列描述: SEQ ID NO : 5:**

```

Met Thr Pro Pro Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr
1           5           10           15

Leu His Leu Leu Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala
          20           25           30

Gln Gly Leu Pro Gly Val Gly Leu Thr Pro Ser Ala Ala Gln Thr Ala
          35           40           45

Arg Gln His Pro Lys Met His Leu Ala His Ser Thr Leu Lys Pro Ala
          50           55           60

Ala His Leu Ile Gly Asp Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg
65           70           75           80

Ala Asn Thr Asp Arg Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn
          85           90           95

Asn Ser Leu Leu Val Pro Thr Ser Gly Ile Tyr Phe Val Tyr Ser Gln
          100          105          110

Val Val Phe Ser Gly Lys Ala Tyr Ser Pro Lys Ala Thr Ser Ser Pro
          115          120          125

Leu Tyr Leu Ala His Glu Val Gln Leu Phe Ser Ser Gln Tyr Pro Phe
          130          135          140

His Val Pro Leu Leu Ser Ser Gln Lys Met Val Tyr Pro Gly Leu Gln
          145          150          155          160

Glu Pro Trp Leu His Ser Met Tyr His Gly Ala Ala Phe Gln Leu Thr
          165          170          175

Gln Gly Asp Gln Leu Ser Thr His Thr Asp Gly Ile Pro His Leu Val
          180          185          190

Leu Ser Pro Ser Thr Val Phe Phe Gly Ala Phe Ala Leu
          195          200          205

```

(2) SEQ ID NO: 6 信息:**(i)序列特征:****(A)长度: 281 个氨基酸****(B)类型: 氨基酸****(C)链型: 不相关****(D)拓扑结构: 不相关****(ii)分子类型: 蛋白质****(xi)序列描述: SEQ ID NO: 6:**

```

Met Gln Gln Pro Phe Asn Tyr Pro Tyr Pro Gln Ile Tyr Trp Val Asp
1           5           10           15

Ser Ser Ala Ser Ser Pro Trp Ala Pro Pro Gly Thr Val Leu Pro Cys
          20           25           30

Pro Thr Ser Val Pro Arg Arg Pro Gly Gln Arg Arg Pro Pro Pro Pro
          35           40           45

Pro Pro Pro Pro Pro Leu Pro Pro Pro Pro Pro Pro Pro Pro Leu Pro
          50           55           60

Pro Leu Pro Leu Pro Pro Leu Lys Lys Arg Gly Asn His Ser Thr Gly
65           70           75           80

Leu Cys Leu Leu Val Met Phe Phe Met Val Leu Val Ala Leu Val Gly
          85           90           95

Leu Gly Leu Gly Met Phe Gln Leu Phe His Leu Gln Lys Glu Leu Ala
          100           105           110

Glu Leu Arg Glu Ser Thr Ser Gln Met His Thr Ala Ser Ser Leu Glu
          115           120           125

Lys Gln Ile Gly His Pro Ser Pro Pro Pro Glu Lys Lys Glu Leu Arg
          130           135           140

Lys Val Ala His Leu Thr Gly Lys Ser Asn Ser Arg Ser Met Pro Leu
          145           150           155           160

Glu Trp Glu Asp Thr Tyr Gly Ile Val Leu Leu Ser Gly Val Lys Tyr
          165           170           175

Lys Lys Gly Gly Leu Val Ile Asn Glu Thr Gly Leu Tyr Phe Val Tyr
          180           185           190

Ser Lys Val Tyr Phe Arg Gly Gln Ser Cys Asn Asn Leu Pro Leu Ser
          195           200           205

His Lys Val Tyr Met Arg Asn Ser Lys Tyr Pro Gln Asp Leu Val Met
          210           215           220

```

Met	Glu	Gly	Lys	Met	Met	Ser	Tyr	Cys	Thr	Thr	Gly	Gln	Met	Trp	Ala
225					230					235					240
Arg	Ser	Ser	Tyr	Leu	Gly	Ala	Val	Phe	Asn	Leu	Thr	Ser	Ala	Asp	His
				245					250					255	
Leu	Tyr	Val	Asn	Val	Ser	Glu	Leu	Ser	Leu	Val	Asn	Phe	Glu	Glu	Ser
			260					265					270		
Gln	Thr	Phe	Phe	Gly	Leu	Tyr	Lys	Leu							
		275					280								

(2) SEQ ID NO: 7 信息:

(i)序列特征:

(A)长度: 24 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 7:

GCGGGATCCG GAGAGATGGT CACC 24

(2) SEQ ID NO: 8 信息:

(i)序列特征:

(A)长度: 27 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 8:

CGCAAGCTTC CTTACACCA TGAAAGC 27

(2) SEQ ID NO: 9 信息:

(i)序列特征:

(A)长度: 32 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 9:

GACCGGATCC ATGGAGGAGA GTGTCGTACG GC 32

(2) SEQ ID NO: 10 信息:

(i)序列特征:

(A)长度: 27 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 10:

CGCAAGCTTC CTTACACCA TGAAAGC 27

(2) SEQ ID NO: 11 信息:

(i)序列特征:

(A)长度: 40 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 11:

GCTCCAGGAT CCGCCATCAT GGAGGAGAGT
GTCGTACGGC 40

(2) SEQ ID NO: 12 信息:

(i)序列特征:**(A)长度: 37 个碱基对****(B)类型: 核酸****(C)链型: 单链****(D)拓扑结构: 线性****(ii)分子类型: cDNA****(xi)序列描述: SEQ ID NO: 12:****GACGCGGTAC CGTCCAATGC ACCACGCTCC TTCCTTC 37****(2) SEQ ID NO: 13 信息:****(i)序列特征:****(A)长度: 19 个碱基对****(B)类型: 核酸****(C)链型: 单链****(D)拓扑结构: 线性****(ii)分子类型: cDNA****(xi)序列描述: SEQ ID NO: 13:****GCTCGGATCC GCCATCATG 19****(2) SEQ ID NO: 14 信息:****(i)序列特征:****(A)长度: 71 个碱基对****(B)类型: 核酸****(C)链型: 单链****(D)拓扑结构: 线性****(ii)分子类型: cDNA****(xi)序列描述: SEQ ID NO: 14:****GATGTTCTAG AAAGCGTAGT CTGGGACGTC GTATGGGTAC ACCATGAAAG CCCCGAAGTA****60****AGACCGGGTA C****71**

(2) SEQ ID NO : 15 信息:**(i)序列特征:****(A)长度: 40 个碱基对****(B)类型: 核酸****(C)链型: 单链****(D)拓扑结构: 线性****(ii)分子类型: cDNA****(xi)序列描述: SEQ ID NO : 15:**

GCTCCAGGAT CCGCCATCAT GGAGGAGAGT
GTCGTACGGC 40

(2) SEQ ID NO : 16 信息:**(i)序列特征:****(A)长度: 37 个碱基对****(B)类型: 核酸****(C)链型: 单链****(D)拓扑结构: 线性****(ii)分子类型: cDNA****(xi)序列描述: SEQ ID NO : 16:**

GACGCGGTAC CGTCCAATGC ACCACGCTCC TTCCTTC 37

(2) SEQ ID NO : 17 信息:**(i)序列特征:****(A)长度: 32 个碱基对****(B)类型: 核酸****(C)链型: 单链****(D)拓扑结构: 线性****(ii)分子类型: cDNA**

(xi)序列描述: SEQ ID NO: 17:

GACGCCCATG GAGGAGGAGA GTGTCGTACG GC 32

(2) SEQ ID NO: 18 信息:

(i)序列特征:

(A)长度: 33 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 18:

GACCGGATCC CACCATGAAA GCCCCGAAGT AAG 33

(2) SEQ ID NO: 19 信息:

(i)序列特征:

(A)长度: 27 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: cDNA

(xi)序列描述: SEQ ID NO: 19:

CGCAAGCTTC CTTACACCA TGAAAGC 27

```

      10              30              50
GAGGTTGAAGGACCCAGGCGTGTGAGCCCTGCTCCAGAGACCTTGGGCATGGAGGAGAGT
-----+-----+-----+-----+-----+-----+-----+
                                     M E E S
      70              90             110
GTCGTACGGCCCTCAGTGTTTGTGGTGGATGGACAGACCGACATCCCATTCACGAGGCTG
-----+-----+-----+-----+-----+-----+-----+
V V R P S V F V V D G Q T D I P F T R L
      130             150             170
GGACGAAGCCACCGGAGACAGTCGTGCAGTGTGGCCCGGGTGGGTCTGGGTCTCTTGCTG
-----+-----+-----+-----+-----+-----+-----+
G R S H R R Q S C S V A R V G L G L L L
      190             210             230
TTGCTGATGGGGGCTGGGCTGGCCGTCCAAGGCTGGTTCCTCCTGCAGCTGCACTGGCGT
-----+-----+-----+-----+-----+-----+-----+
L L M G A G L A V Q G W F L L Q L H W R
      250             270             290
CTAGGAGAGATGGTCACCCGCTGCCTGACGGACCTGCAGGCTCCTGGGAGCAGCTGATA
-----+-----+-----+-----+-----+-----+-----+
L G E M V T R L P D G P A G S W E Q L I
      310             330             350
CAAGAGCGAAGGTCTCACGAGGTCAACCCAGCAGCGCATCTCACAGGGGCCAACTCCAGC
-----+-----+-----+-----+-----+-----+-----+
Q E R R S H E V N P A A H L T G A N S S
      370             390             410
TTGACCGGCAGCGGGGGGCGCTGTTATGGGAGACTCAGCTGGGCCTGGCCTTCCTGAGG
-----+-----+-----+-----+-----+-----+-----+
L T G S G G P L L W E T Q L G L A F L R
      430             450             470
GGCCTCAGCTACCACGATGGGGCCCTTGTGGTCACCAAAGCTGGCTACTACTACATCTAC
-----+-----+-----+-----+-----+-----+-----+
G L S Y H D G A L V V T K A G Y Y Y I Y

```

图 1A

```

      490              510              530
TCCAAGGTGCAGCTGGGCGGTGTGGGCTGCCCCTGGGCTGGCCAGCACCATCACCCAC
-----+-----+-----+-----+-----+-----+
S  K  V  Q  L  G  G  V  G  C  P  L  G  L  A  S  T  I  T  H
      550              570              590

GGCCTCTACAAGCGCACACCCCGCTACCCCGAGGAGCTGGAGCTGTTGGTCAGCCAGCAG
-----+-----+-----+-----+-----+-----+
G  L  Y  K  R  T  P  R  Y  P  E  E  L  E  L  L  V  S  Q  Q
      610              630              650

TCACCTGCGGACGGGCCACCCAGCAGCTCCCGGGTCTGGTGGGACAGCAGCTTCCTGGGT
-----+-----+-----+-----+-----+-----+
S  P  C  G  R  A  T  S  S  S  R  V  W  W  D  S  S  F  L  G
      670              690              710

GGTGTGGTACACCTGGAGGCTGGGGAGGAGGTGGTCGTCCGTGTGCTGGATGAACGCCTG
-----+-----+-----+-----+-----+-----+
G  V  V  H  L  E  A  G  E  E  V  V  V  R  V  L  D  E  R  L
      730              750              770

GTTCGACTGCGTGATGGTACCCGGTCTTACTTCGGGGCTTTCATGGTGTGAAGGAAGGAG
-----+-----+-----+-----+-----+-----+
V  R  L  R  D  G  T  R  S  Y  F  G  A  F  M  V  *
      790              810              830

CGTGGTGCATTGGACATGGGTCTGACACGTGGAGAACTCAGAGGGTGCCTCAGGGGAAAAG
-----+-----+-----+-----+-----+-----+
      850              870              890

AAAACCTACGAAGCAGAGGCTGGGCGTGGTGGCTCTCGCCTGTAATCCAGCACTTTGGG
-----+-----+-----+-----+-----+-----+
      910              930              950

AGGCCAAGGCAGGCGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCTAACATGGC
-----+-----+-----+-----+-----+-----+
      970              990              1010

AAAACCCCATCTCTAC  TAAAAATACAAAAATTAGCCGGACGTGGTGGTGCCTGCCTGTAA
-----+-----+-----+-----+-----+-----+

```

图 1B

1030 1050 1070
TCCAGCTACTCAGGAGGCTGAGGCAGGATAATTTTGCTTAAACCCGGGAGGCGGAGGTTG
-----+-----+-----+-----+-----+-----+
1090 1110 1130
CAGTGAGCCGAGATCACACCACTGCACTCCAACCTGGGAAACGCAGTGAGACTGTGCCTC
-----+-----+-----+-----+-----+-----+
1150
AAAAAAAAAAAAAAAAAAAAAAAAAAAA
-----+-----+-----

图 1C

	10										20										30										大多数
	M	T	P	P	E	--	R	L	F	L	P	R	V	--	V	D	--	P	--	--	--	G	--	AIM II氨基酸							
1	M	E	--	E	--	S	V	V	R	P	S	V	F	V	D	G	Q	T	D	I	P	F	I	R	L	G	--	人TNF-α蛋白			
1	M	G	L	S	T	V	P	D	L	L	L	P	L	V	L	E	L	L	V	G	I	Y	P	S	G	V	I	G	--	人TNF-β蛋白	
1	M	T	P	P	E	--	R	L	F	L	P	R	V	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	人淋巴瘤细胞毒素蛋白		
1	M	T	P	P	E	--	R	L	F	L	P	R	V	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	人Fas配基蛋白		
1	M	Q	Q	P	F	--	N	Y	P	Y	P	Q	I	Y	W	V	D	S	S	A	S	S	P	W	A	P	P	G	T		
	--	--	--	--	--	--	--	--	--	--	--	--	--	--	C	G	T	I	L	H	--	P	--	R	R	--	C	--	大多数		
	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	AIM II氨基酸		
26	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	人TNF-α蛋白		
31	V	P	H	L	G	D	R	E	K	R	D	S	V	C	P	Q	G	K	Y	I	H	P	Q	N	N	S	I	C	T	人TNF-β蛋白	
13	--	--	--	--	--	--	--	--	--	--	--	--	--	--	C	G	T	T	L	H	--	--	--	--	--	--	--	--	人淋巴瘤细胞毒素蛋白		
13	--	--	--	--	--	--	--	--	--	--	--	--	--	--	C	G	T	T	L	H	--	--	--	--	--	--	--	--	人Fas配基蛋白		
29	V	L	P	--	--	--	--	--	--	--	--	--	--	--	C	P	T	S	V	P	R	R	P	G	Q	R	R	P	P	P	
	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	大多数		
	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	AIM II氨基酸		
35	V	A	R	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	人TNF-α蛋白		
61	K	C	H	K	G	T	Y	L	N	D	C	P	G	P	G	Q	D	T	D	C	R	E	C	E	S	G	S	F	T	人TNF-β蛋白	
19	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	人淋巴瘤细胞毒素蛋白		
19	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	人Fas配基蛋白		
49	P	P	P	P	P	L	--	--	--	--	--	--	--	--	P	P	P	P	P	P	P	P	L	P	P	L	P	P	L	K	
	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	大多数		

38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	AIM II氨基酸 人TNF-α蛋白 人TNF-β蛋白 人淋巴毒素蛋白 人Fas配基蛋白	
91	A	S	E	N	H	L	R	H	C	L	S	C	S	K	C	R	K	E	M	G	V	G	L	G	L	L	L	L	L	M	G	A	G	L		
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
73	K	R	G	N	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	T	G	L	C	L	L	V	M	F	F	M	V	L	V		
	-	-	G	L	P	G	-	-	-	-	-	-	-	-	-	-	-	-	-	V	G	L	-	-	-	F	-	L	F	-	-	-	-	-	大多数	
52	-	-	A	V	Q	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
121	R	D	T	V	C	G	C	R	K	N	Q	Y	R	H	Y	W	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
34	-	-	G	L	P	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
34	-	-	G	L	P	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
93	-	-	A	L	V	G	-	-	-	-	-	-	-	-	-	-	-	-	-	L	G	L	G	M	F	Q	L	F	H	L	Q	K	E	L		
	-	E	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	大多数	
66	G	E	M	V	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
151	N	G	T	V	H	L	S	C	Q	E	K	Q	N	T	V	C	T	C	H	A	G	F	F	L	R	E	N	E	C	V	-	-	-	-		
41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
112	A	E	L	R	E	S	T	S	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	Q	T	A	R	Q	H	P	S	M	E	L	A	K	S	T	L	K	P	A	A	H	L	I	G	D	P	S	S	-	-	-	-	-	-	-	大多数

图 2B

		190										200										210																											
77		A	G	S	W	E	Q	L	I	Q	E	R	R	S	H	E	V	N	P	A	A	H	L	T	G	A	N	S	S	L	T																		
181		S	C	S	N	C	K	S	L	E	C	T	K	L	C	L	P	Q	I	E	N	V	K	G	T	E	D	S	G	I																			
46		Q	T	A	R	Q	H	P	K	M	H	L	A	H	S	T	L	K	P	A	A	H	L	I	G	D	P	S	K	-	-																		
46		Q	T	A	R	Q	H	P	K	M	H	L	A	H	S	T	L	K	P	A	A	H	L	I	G	D	P	S	K	-	-																		
128		E	K	Q	I	G	H	P	S	P	P	E	K	K	E	L	R	K	V	A	H	L	T	G	K	S	N	S	-	-																			
		AIM II氨基酸												人TNF-α蛋白												人TNF-β蛋白										人淋巴细胞毒素蛋白													
		人Fas配基蛋白																																															
		大多数																																															
		- Q N - P L L -												- - - - -												- - - - -														W E									
		220												230												240																							
107		G	S	G	G	P	L	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
211		T	V	L	L	P	L	V	I	F	F	G	L	C	L	L	S	L	L	F	I	G	L	M	Y	R	Y	Q	R	-	-																		
74		-	Q	N	-	S	L	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
74		-	Q	N	-	S	L	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
156		-	R	S	M	P	L	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
		AIM II氨基酸												人TNF-α蛋白												人TNF-β蛋白										人淋巴细胞毒素蛋白													
		人Fas配基蛋白																																															
		大多数																																															
		A N L G R A F -												- - - - -												- - - - -														L Q D G									
		250												260												270																							
116		T	Q	L	G	L	A	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
241		S	K	L	Y	S	I	V	C	G	K	S	T	P	E	K	E	G	E	L	E	G	T	T	K	P	L	A	P	N																			
81		A	N	T	D	R	A	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
81		A	N	T	D	R	A	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
164		D	T	Y	G	I	V	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-																			
		AIM II氨基酸												人TNF-α蛋白												人TNF-β蛋白										人淋巴细胞毒素蛋白													
		人Fas配基蛋白																																															
		大多数																																															
		F S L S N G S L V V P T S G I Y F V Y S Q V V F S G K A Y S																																															

图 2C

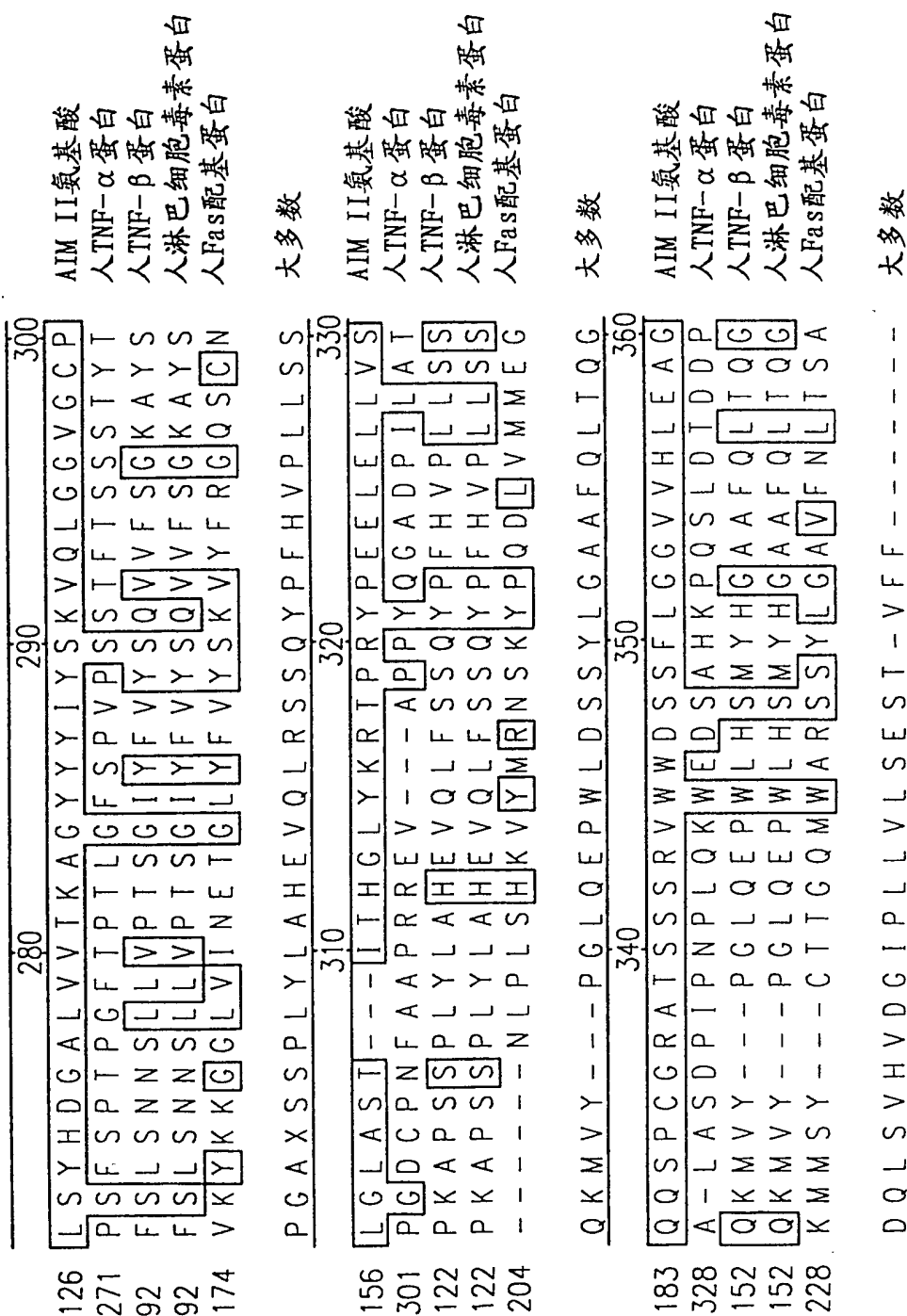


图 2D

	370	380	390	
213	E E V V V R V L D E R L V R L R D G T R S Y F			AIM II氣基酸
357	A T L Y A V V E N V P P L R W K E F V R L G L S D H E I D			人TNF-α 蛋白
179	D Q L S T H T D G I P H L V L S P S T - V F F			人TNF-β 蛋白
179	D Q L S T H T D G I P H L V L S P S T - V F F			人淋巴细胞毒蛋白
255	D H L Y V N V S E L S L V N F E E S Q - T F F			人Fas配基蛋白
	- - - - - G A F A - L - - - - -			大多数
	400	410	420	
236	- - - - - G A F M V			AIM II氣基酸
387	R L E L Q N G R C L R E A Q Y S M L A T W R R R T P R R E A			人TNF-α 蛋白
201	- - - - - G A F A - L			人TNF-β 蛋白
201	- - - - - G A F A - L			人淋巴细胞毒蛋白
277	- - - - - G L Y K - L			人Fas配基蛋白
	- - - - -			大多数
	430	440	450	
241	T L E L L G R V L R D M D L L G C L E D I E E A L C G P A A			AIM II氣基酸
417				人TNF-α 蛋白
205				人TNF-β 蛋白
205				人淋巴细胞毒蛋白
281				人Fas配基蛋白
	- - - - -			大多数

2E

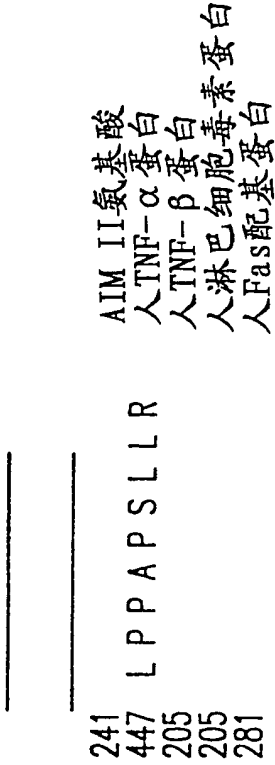


图 2F

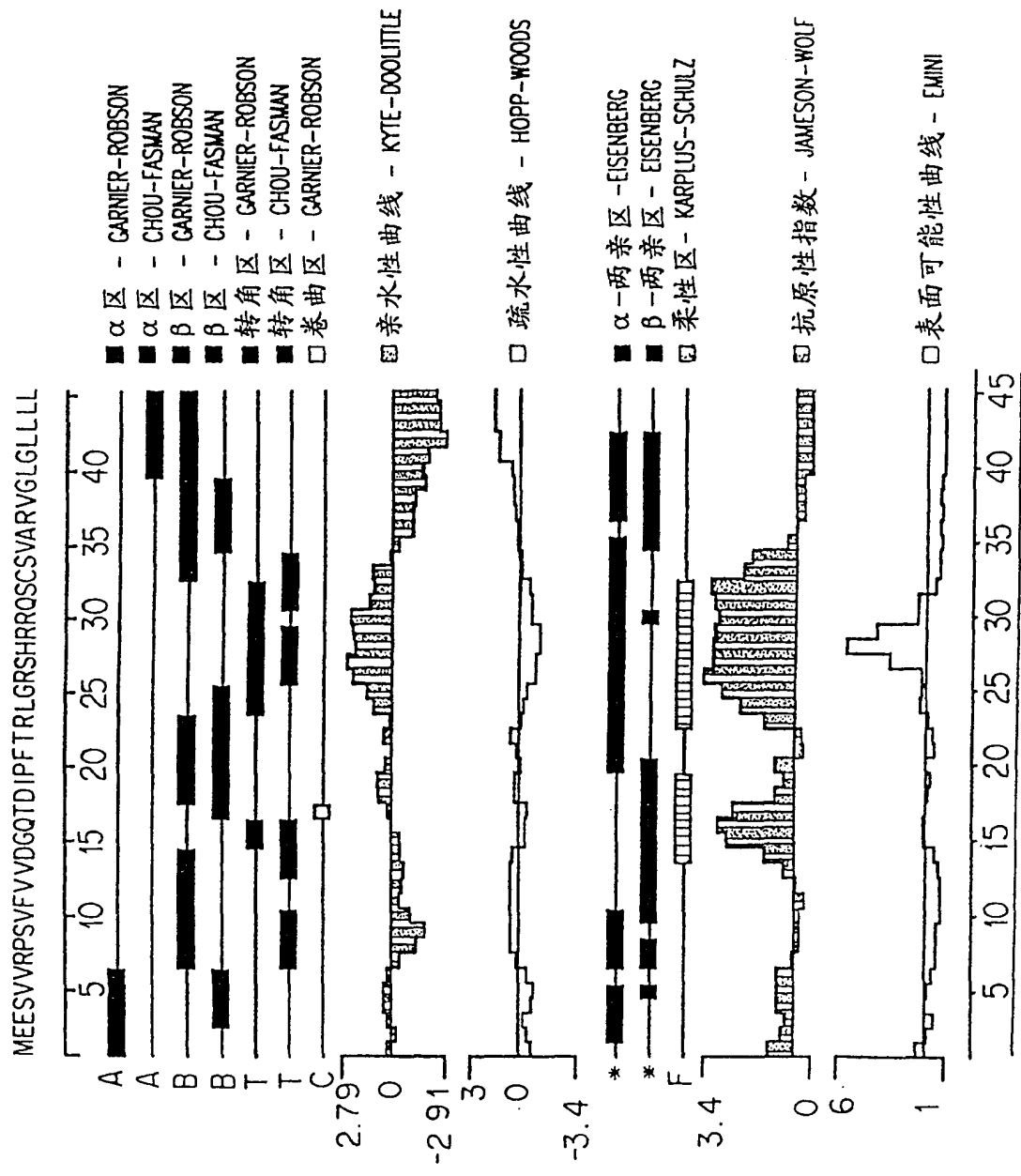


图 3A

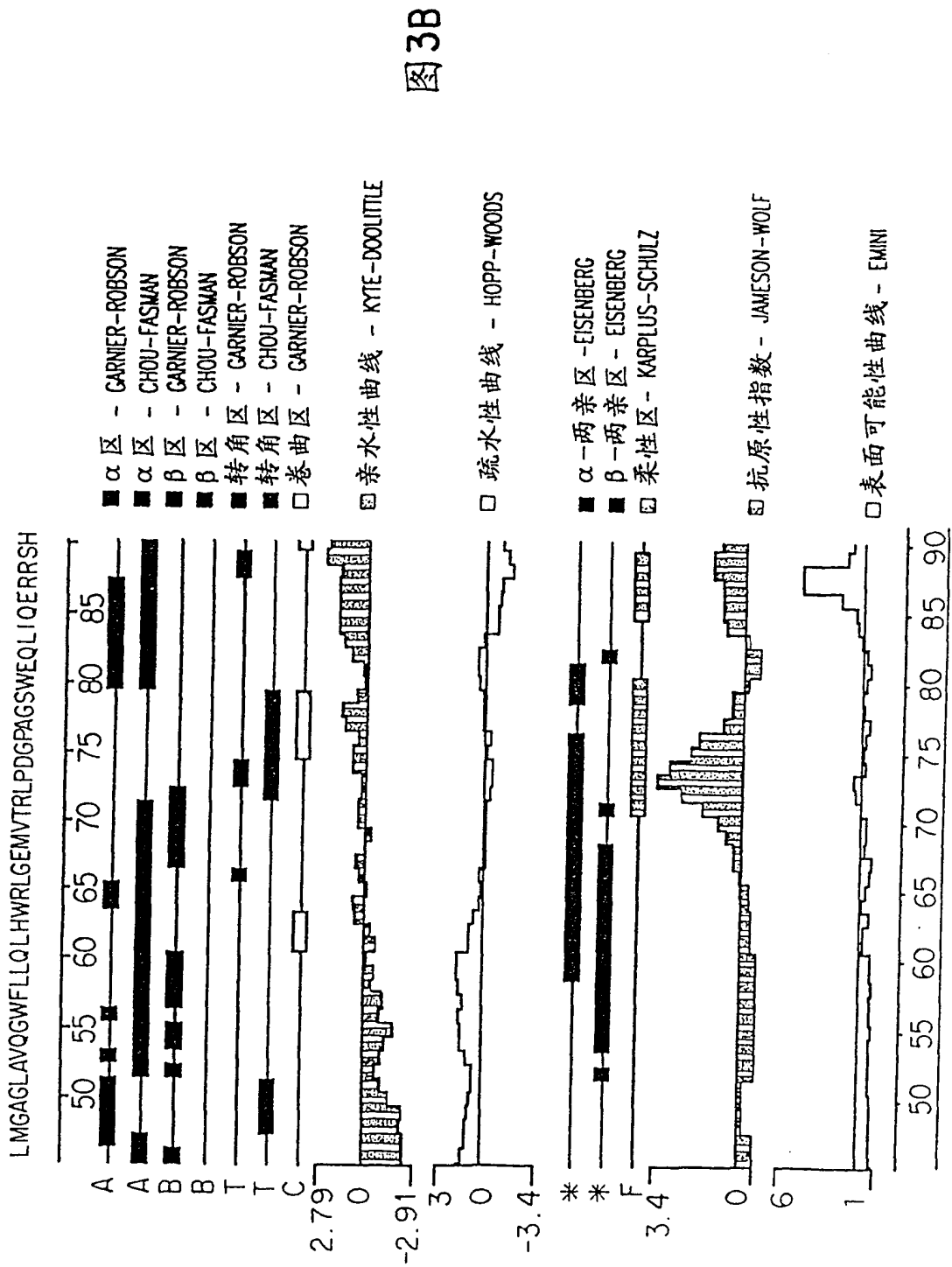


图 3B

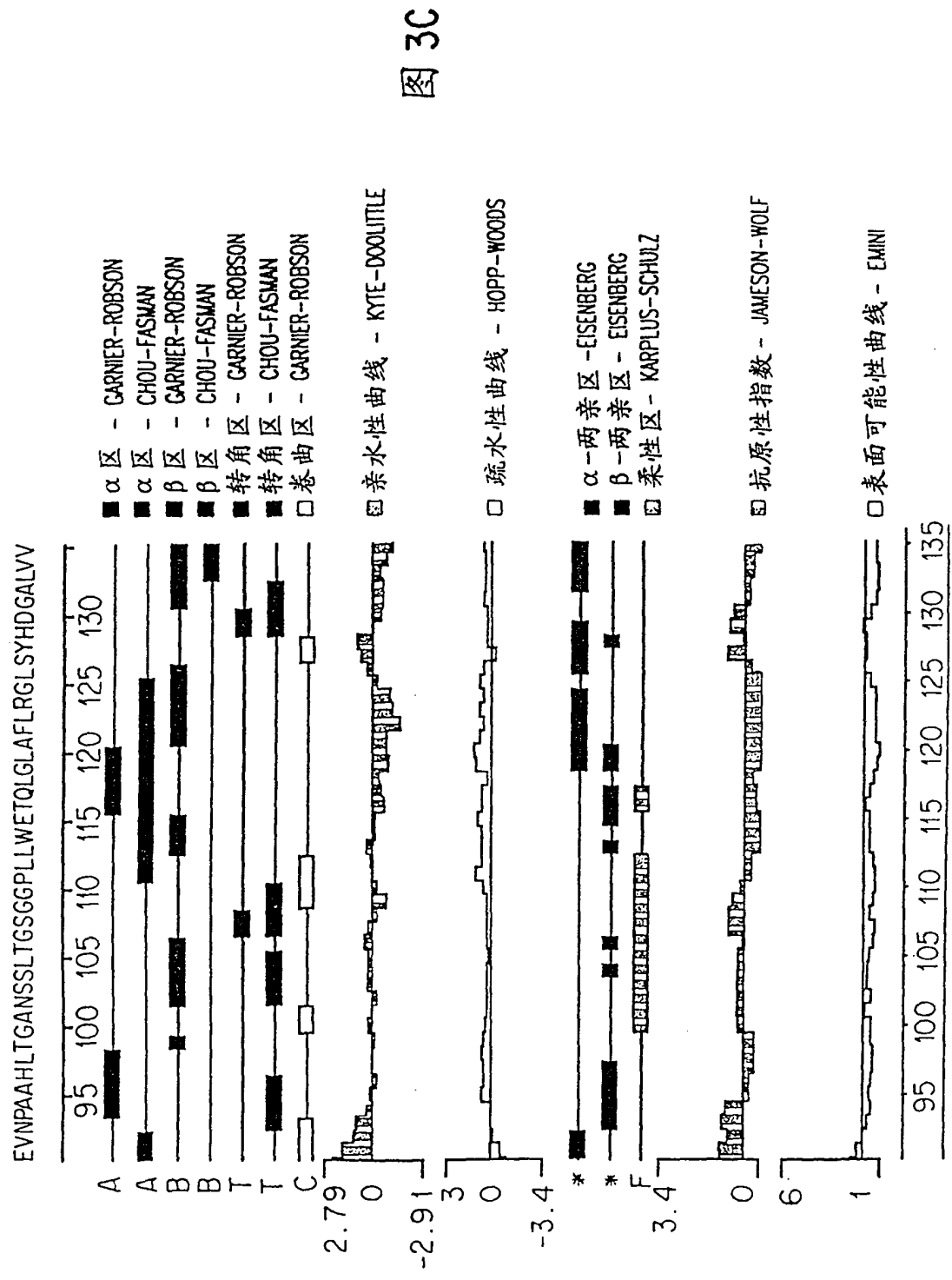


图 3C

