

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 00815837.1

[51] Int. Cl.

C12N 9/64 (2006.01)

C12N 15/57 (2006.01)

C07K 16/40 (2006.01)

A61K 38/48 (2006.01)

A61K 48/00 (2006.01)

C12Q 1/68 (2006.01)

[45] 授权公告日 2006年4月12日

[11] 授权公告号 CN 1250717C

[51] Int. Cl. (续)

C12Q 1/37 (2006.01)

G01N 33/50 (2006.01)

A61P 9/00 (2006.01)

[22] 申请日 2000.11.17 [21] 申请号 00815837.1

[30] 优先权

[32] 1999.11.19 [33] EP [31] 9920386.0

[32] 1999.11.19 [33] NL [31] 1013616

[32] 2000.5.31 [33] EP [31] 00201937.0

[32] 2000.5.31 [33] NL [31] 1015356

[86] 国际申请 PCT/EP2000/011532 2000.11.17

[87] 国际公布 WO2001/036610 英 2001.5.25

[85] 进入国家阶段日期 2002.5.17

[71] 专利权人 索尔瓦药物有限公司

地址 荷兰韦斯普

[72] 发明人 W·德利尔斯奈德 R·韦格斯

M·韦斯克

审查员 吴希哲

[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 唐伟杰

权利要求书 2 页 说明书 83 页 附图 11 页

[54] 发明名称

金属蛋白酶家族中的人酶

[57] 摘要

本发明涉及最近鉴定的下文称为 IGS5 的多肽和编码这些多肽的多核苷酸，它们在治疗中和在鉴定可能是潜在可用于治疗的刺激物和/或抑制物的化合物中的用途，以及这些多肽和多核苷酸的生成。所述多肽和多核苷酸涉及金属蛋白酶家族，优选人型。本发明还涉及抑制或刺激/激活这些多肽和多核苷酸的作用，包含所述多核苷酸的载体，以及包含所述载体的宿主细胞。本发明还涉及用于筛选能够担当所述 IGS5 酶的刺激物或抑制物的化合物的方法。本发明的多肽在几种机能障碍、紊乱和疾病方面受到关注。本发明的多肽在心血管疾病中特别受到关注。

1. 一种分离的多肽，其具有选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6之一的氨基酸序列并且具有IGS 5多肽活性。
2. 权利要求1的分离的多肽，其是人类多肽。
3. 一种分离多核苷酸，其具有编码与选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6之一的氨基酸序列相同并具有IGS 5多肽活性的多肽的核苷酸序列，或具有与所述分离多核苷酸互补的核苷酸序列。
4. 一种分离多核苷酸，其具有与选自SEQ ID NO: 1、SEQ ID NO: 3、和SEQ ID NO: 5之一的核苷酸序列相同的核苷酸序列，或具有与所述分离多核苷酸互补的核苷酸序列。
5. 选自下组的分离多核苷酸：
 - (a) 具有编码选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6之一的多肽的核苷酸序列的多核苷酸；
 - (b) 选自SEQ ID NO: 1、SEQ ID NO: 3、和SEQ ID NO: 5的多核苷酸；或
 - (c) 与所述分离多核苷酸互补的核苷酸序列。
6. 权利要求3-5中任一项的多核苷酸，其是DNA或RNA。
7. 权利要求3-5中任一项的多核苷酸，其是人类多核苷酸。
8. 一种包含多核苷酸的表达系统，当所述表达系统存在于相容宿主细胞中时，所述多核苷酸能够生成权利要求1-2中任一项的多肽。
9. 包含权利要求8的表达系统的宿主细胞或其表达权利要求1-2中任一项的多肽的膜。
10. 用于生产权利要求1-2中任一项的多肽的方法，包括在足以生成所述多肽的条件下培养权利要求9的宿主细胞并由培养物培养基中回收多肽。
11. 对权利要求1-2中任一项的多肽具有免疫特异性的抗体。
12. 用于鉴定可影响权利要求1-2中任一项所述多肽之活性的化

化合物的筛选方法，包括：

(a) 使携带权利要求1-2中任一项所述多肽的权利要求9所述细胞或膜接触候选化合物；并

(b) 评估所述候选化合物是否导致对所述多肽活性的刺激或抑制。

13. 用于鉴定可刺激或抑制权利要求1-2中任一项所述多肽之功能的化合物的筛选方法，包括选自下组的方法：

(a) 在存在所述多肽的合适底物的情况下，测量候选化合物对该多肽、携带该多肽的细胞或膜、或者其融合蛋白的活性的影响；

(b) 在存在所述多肽的竞争物的情况下，测量候选化合物对该多肽、携带该多肽的细胞或膜、或者其融合蛋白的活性的影响；

(c) 使用适用于该多肽活性或者携带该多肽的细胞或膜的检测系统，测试候选化合物是否引起由所述多肽的激活或抑制生成的信号；

(d) 混合候选化合物与含权利要求1所述多肽的溶液和合适底物以形成混合物，测量混合物中该多肽的活性，并比较混合物与不含候选化合物的标准之间的活性；

(e) 检测候选化合物对细胞中编码所述多肽的mRNA和所述多肽的生成的影响；或

(f) 用ELISA检测法检测候选化合物对细胞中编码所述多肽的mRNA和所述多肽的生成的影响。

14. 权利要求12或13的方法，其中筛选用于鉴定适用于治疗和/或预防心血管疾病的化合物。

金属蛋白酶家族中的人酶

发明领域

本发明涉及最近鉴定的多肽和编码这些多肽的多核苷酸，它们在治疗中和在鉴定可能是潜在可用于治疗的刺激物和/或抑制物的化合物中的用途，以及这些多肽和多核苷酸的生成。更具体的说，本发明的多肽和多核苷酸所涉及的酶是金属蛋白酶多肽家族或结构和功能相关多肽的特定家族的成员。这些酶在下文中称为IGS5。本发明还涉及抑制或刺激/激活这些多肽和多核苷酸的作用，包含所述多核苷酸的载体，以及包含所述载体的宿主细胞。本发明还涉及用于筛选能够担当所述IGS5酶的刺激物或抑制物的化合物的方法。

发明背景

药物开发过程目前正经历基本变化，因为它采用了“功能基因组学”，即基于高通量基因组或基因的生物学。这种方法作为鉴定作为治疗靶的基因和基因产物的手段，正迅速取代基于“位置克隆”的早期方法。首先鉴定表型，即生物学功能或遗传学疾病，然后根据遗传图定位追踪对此负有责任的基因。

功能基因组学非常依赖高通量DNA测序技术和用于由目前可利用的许多分子生物学数据库鉴定潜在关注的基因序列的多种生物信息学工具。仍旧需要鉴定和表征作为药物开发的靶的其它基因及其相关多肽/蛋白质。

在药物开发所关注的多肽中，存在金属蛋白酶及结构和功能相关酶的特定家族。已经鉴定了几种疾病，其中金属蛋白酶在该疾病的病理学中担当关键角色。例如，现有技术中已经鉴定和表征了许多锌金属蛋白酶或结构和功能相关酶的特定家族，而且已经变得显然的是，这些酶（如锌金属蛋白酶）的参与在涵盖正常和疾病两种状况的生物

学功能多样性中担当角色。锌金属蛋白酶是这些酶中催化功能严重依赖于活性位点的锌离子的子集。这组酶包括根据序列和结构信息分类的多个家族，而且据描述密切涉及诸如胚胎的发育、软骨和骨的形成、肽激素的加工、生殖、心血管疾病、关节炎、癌症等过程。针对有些锌金属蛋白酶的活性位点定向抑制物早已用于治疗，如用作抗高血压药物。

根据锌金属蛋白酶中锌结合位点周围的序列和结构信息，可以将这些酶分成几个家族，进而可以分成超家族，诸如“metzincin”（astacin、serratin、reprolysin、基质素）、“gluzincin”（嗜热芽孢菌蛋白酶、neprilysin、血管紧张素转化酶、氨肽酶）、或“zincin”（包含metzincin和gluzincin超家族）。这种分组不仅有助于阐明常见的催化和生物合成加工机制，而且在阐明最近鉴定的具有相似锌结合基元的蛋白质的功能中是无价的。本领域早已鉴定的金属蛋白酶（如锌酶）的一些个别范例包括neprilysin、内皮素转化酶、血管紧张素转化酶、嗜热芽孢菌蛋白酶、氨肽酶、astacin、serratin、reprolysin、基质素、胰岛素酶、羧肽酶、和DD-羧肽酶。

根据本领域的上述证据，显然，金属蛋白酶和特定结构和功能相关酶在健康和疾病中担当关键角色。因而，仍然需要进一步揭开这组酶的重要功能和潜在治疗应用以及对提供新型金属蛋白酶和随后开发的新型合成刺激物（激活物）或抑制物，它们有助于为多种社会经济学重要疾病提供新型疗法。

发明概述

一方面，本发明涉及IGS5，特别是IGS5多肽和IGS5多核苷酸，优选涉及人类物种的那些，以及用于生成它们的重组材料和方法。另一方面，本发明涉及使用这些多肽、多核苷酸、和重组材料的方法，包括治疗疾病，其中多肽像金属蛋白酶或结构和功能相关酶的特定家族在待治疗的机能障碍、紊乱、或疾病（下文总称为“疾病”）的病理学中担当重要角色。认为本发明的多肽和多核苷酸在其中有用的疾病

的范例包括：CNS紊乱，包括精神分裂症，发作性阵发性焦虑症（EPA）紊乱诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经退化疾病，严重精神发育迟缓，和运动障碍，诸如亨廷顿病或图雷特综合症，厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意力缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心率失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食道反流病（GERD），活动能力障碍，和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒病；糖尿病并发症等。还有一个方面，本发明涉及使用本发明提供的材料来鉴定激动剂和拮抗剂或抑制剂的方法，以及使用经鉴定的化合物来治疗与IGS5不平衡有关的状况。还有另一个方面，本发明涉及用于检测与不适当的IGS5活性或水平有关的疾病的诊断测定法。

本发明的多肽在心血管疾病领域特别受到关注。

表1: SEQ ID NO: 1的IGS5-DNA ("IGS5DNA")

5' -
TGCACCACCCCTGGCTGCGTGATAGCAGCTGCCAGGATCCTCCAGAACATGGACCCGACC
ACGGAACCGTGTGACGACTTCTACCAGTTTGCATGCGGAGGCTGGCTGCGGCGCCACGTG
ATCCCTGAGACCAACTCAAGATACAGCATCTTTGACGTCCTCCGCGACGAGCTGGAGGTC
ATCCTCAAAGCGGTGCTGGAGAATTGACTGCCAAGGACCGGCCGGCTGTGGAGAAGGCC
AGGACGCTGTACCGCTCCTGCATGAACCAGAGTGTGATAGAGAAGCGAGGCTCTCAGCCC
CTGCTGGACATCTTGGAGGTGGTGGGAGGCTGGCCGGTGGCGATGGACAGGTGGAACGAG
ACCGTAGGACTCGAGTGGGAGCTGGAGCGGCAGCTGGCGCTGATGAACTCACAGTTCAAC
AGGCGCGTCCTCATCGACCTCTTCATCTGGAACGACGACCAGAACTCCAGCCGGCACATC
ATCTACATAGACCAGCCCACCTTGGGCATGCCCTCCCAGAGTACTACTTCAACGGCGGC
AGCAACCGGAAGGTGCGGGAAGCCTACCTGCAGTTCATGGTGTGAGTGGCCACGTTGCTG
CGGGAGGATGCAAACCTGCCAGGGACAGCTGCCTGGTGCAGGAGGACATGATGCAGGTG
CTGGAGCTGGAGACACAGCTGGCCAAGGCCACGGTACCCAGGAGGAGAGACACGACGTC
ATCGCCTTGTACCACCGGATGGGACTGGAGGAGCTGCAAAGCCAGTTTGGCCTGAAGGGA
TTTAACTGGACTCTGTTTACATAAACTGTGCTATCCTCTGTCAAATCAAGCTGCTGCCA
GATGAGGAAGTGGTGGTCTATGGCATCCCCTACCTGCAGAACCTTGAAAACATCATCGAC
ACCTACTCAGCCAGGACCATAACAGAACTACCTGGTCTGGCGCCTGGTGTGCTGGACCGCATT
GGTAGCCTAAGCCAGAGATTCAGGACACACGAGTGAAGTACCGCAAGGCGCTGTTTGGC
ACAATGGTGGAGGAGGTGCGCTGGCGTGAATGTGTGGGCTACGTCAACAGCAACATGGAG
AACGCCGTGGGCTCCCTCTACGTGAGGAGCGTTCCTGGAGACAGCAAGAGCATGGTC
AGAGAACTCATTGACAAGGTGCGGACAGTGTGTTGTGGAGACGCTGGACGAGCTGGGCTGG
ATGGACGAGGAGTCCAAGAAGAAGGCGCAGGAGAAGGCCATGAGCATCCGGGAGCAGATC
GGGCACCCTGACTACATCCTGGAGGAGATGAACAGGCGCCTGGACGAGGAGTACTCCAAT
CTGAACTTCTCAGAGGACCTGTACTTTGAGAACAGTCTGCAGAACCTCAAGGTGGGCGCC
CAGCGGAGCCTCAGGAAGCTTCGGGAAAAGGTGGACCCAAATCTCTGGATCATCGGGGCG
GCGGTGGTCAATGCGTTCTACTCCCCAAACCGAAACCAGATTGTATTCCCTGCCGGGATC
CTCCAGCCCCCTTCTTTCAGCAAGGAGCAGCCACAGGCCTTGAACCTTGGAGGCATTGGG
ATGGTGATCGGGCACGAGATCACGCACGGCTTTGACGACAATGGCCGGAACCTTCGACAAG
AATGGCAACATGATGGATTGGTGGAGTAACTTCTCCACCAGCACTTCGGGAGCAGTCA
GAGTGCATGATCTACCAGTACGGCAACTACTCCTGGGACCTGGCAGACGAACAGAACGTG
AACGGATTCAACACCCTTGGGGAAAACATTGCTGACAACGGAGGGGTGCGGCAAGCCTAT

```

AAGGCCTACCTCAAGTGGATGGCAGAGGGTGGCAAGGACCAGCAGCTGCCCGGCCTGGAT
CTCACCCATGAGCAGCTCTTCTTCATCAACTACGCCAGGTGTGGTGCGGGTCTACCGG
CCCAGTTCGCCATCCAATCCATCAAGACAGACGTCCACAGTCCCCTGAAGTACAGGGTA
CTGGGGTTCGCTGCAGAACCTGGCCGCCTTCGCAGACACGTTCCACTGTGCCCGGGGCACC
CCCATGCACCCCAAGGAGCGATGCCGCGTGTGGTAG - 3'

```

表2: SEQ ID NO: 2的IGS5-蛋白质 ("IGS5PROT")

```

CTTPGCVIAAARILQNMDPTTEPCDDFYQFACGGWLRRHVI PETNSRYSIFDVL RDELEV
ILKAVLENSTAKDRPAVEKARTLYRSCMNQSVIEKRSQPLLDILEVVGWPVAMDRWNE
TVGLEWELERQLALMNSQFNRRVLIDLF IWNDQNSSRHIIYIDQPTLGMP SREYYFNGG
SNRKVREAYLQFMVSVATLLREDANLPRDSCLVQEDMMQVLELETQLAKATVPQEERHDV
IALYHRMGLEELQSQFGLKGFNWTLFIQTVLSSVKIKLLPDEEVVYGI PYLQNL ENIID
TYSARTIQNYLVWRLVLD RIGSLSQRFKDTRVNYRKALFGTMVEEVRWRECVGYVNSNME
NAVGSLYVREAFPGDSKSMVRELIDKVRTV FVETLDELGWMDEESKKKAQEKAMSIREQI
GHPDYILEEMNRRLDEEYSNLNFS EDLYFENSLQNLKVG AQRSRLRKLREKVDPNLWIIGA
AVVNAFYSPNRNQIVFPAGILQPPFFSKEQPQALNFGGIGM VIGHEITHGFDDNGRNF DK
NGNMMDWWSNFSTQHFREQSECMIYQYGNYSWDLADEQNVNGFNTLGENIADNGGVRQAY
KAYLKWMAEGGKDQQLPGLDLTHEQLFFINYAQVWCGSYRPEFAIQSIKTDVHSP LKYRV
LGS LQNLAAFADTFHCARGT PMHPKERCRVW

```

表3: SEQ ID NO: 3的IGS5-DNA-1 ("IGS5DNA1")

5' -

```

ATGGGGAAGTCCGAAGGCCCGTGGGGATGGTGGAGAGCGCTGGCCGTGCAGGGCAGAAG
CGCCCGGGTTCCTGGAGGGGGGCTGCTGCTGCTGCTGCTGCTGGTGACCGCTGCCCTG
GTGCCCTTGGGTGTCCTCTACGCCGACCGCAGAGGGAAGCAGCTGCCACGCCTTGCTAGC
CGGCTGTGCTTCTTACAGGAGGAGAGGACCTTTGTAAAACGAAAACCCCGAGGGATCCCA
GAGGCCCAAGAGGTGAGCGAGGTCTGCACCACCCTGGCTGCGTGATAGCAGCTGCCAGG
ATCCTCCAGAACATGGACCCGACCACGGAACCGTGTGACGACTTCTACCAGTTTGCATGC
GGAGGCTGGCTGCGGCGCCACGTGATCCCTGAGACCAACTCAAGATACAGCATCTTTGAC
GTCCTCCGCGACGAGCTGGAGGTCATCCTCAAAGCGGTGCTGGAGAATTCGACTGCCAAG
GACCGGCCGGCTGTGGAGAAGGCCAGGACGCTGTACCGCTCCTGCATGAACCAGAGTGTG
ATAGAGAAGCGAGGCTCTCAGCCCCTGCTGGACATCTTGGAGGTGGTGGGAGGCTGGCCG

```


GTGGCGATGGACAGGTGGAACGAGACCGTAGGACTCGAGTGGGAGCTGGAGCGGCAGCTG
GCGCTGATGAACTCACAGTTCAACAGGCGCGTCCTCATCGACCTCTTCATCTGGAACGAC
GACCAGAACTCCAGCCGGCACATCATCTACATAGACCAGCCACCTTGGGCATGCCCTCC
CGAGAGTACTACTTCAACGGCGGCAGCAACCGGAAGGTGCGGGAAGCCTACCTGCAGTTC
ATGGTGTCTAGTGGCCACGTTGCTGCGGGAGGATGCAAACCTGCCAGGGACAGCTGCCTG
GTGCAGGAGGACATGATGCAGGTGCTGGAGCTGGAGACACAGCTGGCCAAGGCCACGGTA
CCCCAGGAGGAGAGACACGACGTATCGCCTTGTACCACCGGATGGGACTGGAGGAGCTG
CAAAGCCAGTTTGGCCTGAAGGGATTAACTGGACTCTGTTCATACAACTGTGCTATCC
TCTGTCAAATCAAGCTGCTGCCAGATGAGGAAGTGGTGGTCTATGGCATCCCCTACCTG
CAGAACCTTGAAAACATCATCGACACCTACTCAGCCAGGACCATACAGAACTACCTGGTC
TGGCGCCTGGTGTCTGGACCGCATTGGTAGCCTAAGCCAGAGATTCAAGGACACACGAGTG
AACTACCGCAAGGCGCTGTTTGGCACAATGGTGGAGGAGGTGCGCTGGCGTGAATGTGTG
GGCTACGTCAACAGCAACATGGAGAACGCCGTGGGCTCCCTCTACGTCAGGGAGGCGTTC
CCTGGAGACAGCAAGAGCATGGTCAGAGAACTCATTGACAAGGTGCGGACAGTGTTTGTG
GAGACGCTGGACGAGCTGGGCTGGATGGACGAGGAGTCCAAGAAGAAGGCGCAGGAGAAG
GCCATGAGCATCCGGGAGCAGATCGGGCACCTGACTACATCCTGGAGGAGATGAACAGG
CGCCTGGACGAGGAGTACTCCAATCTGAACTTCTCAGAGGACCTGTACTTTGAGAACAGT
CTGCAGAACCTCAAGGTGGGCGCCAGCGGAGCCTCAGGAAGCTTCGGGAAAAGGTGGAC
CCAAATCTCTGGATCATCGGGGCGGCGGTGGTCAATGCGTTCCTACTCCCCAAACCGAAAC
CAGATTGTATTCCCTGCCGGGATCCTCCAGCCCCCTTCTTCAGCAAGGAGCAGCCACAG
GCCTTGAACCTTGGAGGCATTGGGATGGTGTATCGGGCACGAGATCACGCACGGCTTTGAC
GACAATGGCCGGAACCTTCGACAAGAATGGCAACATGATGGATTGGTGGAGTAACTTCTCC
ACCCAGCACTTCCGGGAGCAGTCAGAGTGCATGATCTACCAGTACGGCAACTACTCCTGG
GACCTGGCAGACGAACAGAACGTGAACGATTCAACACCCTTGGGGAAAACATTGCTGAC
AACGGAGGGGTGCGGCAAGCCTATAAGGCCTACCTCAAGTGGATGGCAGAGGGTGGCAAG
GACCAGCAGCTGCCCGCCTGGATCTACCCATGAGCAGCTCTTCTTCATCAACTACGCC
CAGGTGTGGTGCGGGTCTACCGGCCCGAGTTCGCCATCCAATCCATCAAGACAGACGTC
CACAGTCCCCTGAAGTACAGGGTACTGGGGTCTGCTGCAGAACCTGGCCGCCTTCGCAGAC
ACGTTCCACTGTGCCCGGGCACCCCATGCACCCCAAGGAGCGATGCCGCGTGTGGTAG

- 3'

表4: SEQ ID NO: 4的IGS5-蛋白质-1 ("IGS5PROT1")

```

MGKSEGPVGMVESAGRAGQKRPGFLEGGLLLLLLLVTAALVALGVLYADRRGKQLPRLAS
RLCFLQEERTFVKRKPRGIPEAQEVSEVCTTPGCVIAAARILQNMDPTTEPCDDFYQFAC
GGWLRRHVPIPETNSRYSIFDVLRDELEVILKAVLENSTAKDRPAVEKARTLYRSCMNQSV
IEKRGSQPLLDILEVVGGWPVAMDRWNETVGLEWELERQLALMNSQFNRRVLIDLFIWND
DQNSSRHI IYIDQPTLGMPSREYYFNNGSNRKVREAYLQFMVSVATLLREDANLPRDSCL
VQEDMMQVLELETQLAKATVPQEERHDVIALYHRMGLEELQSQFGLKGFNWTLFIQTVLS
SVKIKLLPDEEVVYGI PYLQNLNI IDTYSARTIQNYLVWRLVLD RIGSLSQRFKDTRV
NYRKALFGTMVEEVRWRECVGYVNSNMENAVGSLYVREAFPGDSKSMVRELIDKVRTVFV
ETLDELGWMDEESKKAQEKAMSIREQIGHPDYILEEMNRRLDEEYSNLNFSEDLYFENS
LQNLKVG AQRSRLRKLREKVDPNLWI IGA AVVNAFYSPNRNQIVFPAGILQPPFFSKEQPQ
ALNFGGIGMVI GHEITHGFDDNGRNF DKNMMDWWSNFS TQHFREQSECMIYQYGNYSW
DLADEQNVNGFNTLGENIADNGGVRQAYKAYLKWMAEGGKDQQLPGLDLTHEQLFFINYA
QVWCGSYRPEFAIQSIKTDVHSP LKYRVLGSLQNLAAFADTFHCARGT PMPHKERC RVW

```

表5: SEQ ID NO: 5的IGS5-DNA-2 ("IGS5DNA2")

5' -

```

ATGGGGAAGTCCGAAGGCCAGTGGGGATGGTGGAGAGCGCCGGCCGTGCAGGGCAGAAG
CGCCCGGGGTTCTCTGGAGGGGGGCTGCTGCTGCTGCTGCTGCTGGTGACCGCTGCCCTG
GTGGCCTTGGGTGTCCTCTACGCCGACCGCAGAGGGATCCCAGAGGCCCAAGAGGTGAGC
GAGGTCTGCACCACCCCTGGCTGCGTGATAGCAGCTGCCAGGATCCTCCAGAACATGGAC
CCGACCACGGAACCGTGTGACGACTTCTACCAGTTTGCATGCGGAGGCTGGCTGCGGCGC
CACGTGATCCCTGAGACCAACTCAAGATACAGCATCTTTGACGTCCTCCGCGACGAGCTG
GAGGTCATCCTCAAAGCGGTGCTGGAGAATTCGACTGCCAAGGACCGCCGGCTGTGGAG
AAGGCCAGGACGCTGTACCGCTCCTGCATGAACCAGAGTGTGATAGAGAAGCGAGGCTCT
CAGCCCCTGCTGGACATCTTGGAGGTGGTGGGAGGCTGGCCGGTGGCGATGGACAGGTGG
AACGAGACCGTAGGACTCGAGTGGGAGCTGGAGCGGCAGCTGGCGCTGATGAACTCACAG
TTCAACAGGCGCGTCCTCATCGACCTCTTCATCTGGAACGACGACCAGAACTCCAGCCGG
CACATCATCTACATAGACCAGCCCACCTTGGGCATGCCCTCCCGAGAGTACTACTTCAAC
GGCGGCAGCAACCGGAAGGTGCGGGAAGCCTACCTGCAGTTCATGGTGTGAGTGGCCAG
TTGCTGCGGGAGGATGCAAACCTGCCAGGGACAGCTGCCTGGTGCAGGAGGACATGATG
CAGGTGCTGGAGCTGGAGACACAGCTGGCCAAGGCCACGGTACCCAGGAGGAGAGACAC

```

GACGTCATCGCCTTGTACCACCGGATGGGACTGGAGGAGCTGCAAAGCCAGTTTGGCCTG
AAGGGATTTAACTGGACTCTGTTTCATACAAACTGTGCTATCCTCTGTCAAATCAAGCTG
CTGCCAGATGAGGAAGTGGTGGTCTATGGCATCCCCTACCTGCAGAACCTTGAAAACATC
ATCGACACCTACTCAGCCAGGACCATACAGAACTACCTGGTCTGGCGCCTGGTGCTGGAC
CGCATTGGTAGCCTAAGCCAGAGATTCAAGGACACACGAGTGAACTACCGCAAGGCGCTG
TTTGGCACAAATGGTGGAGGAGGTGCGCTGGCGTGAATGTGTGGGCTACGTCAACAGCAAC
ATGGAGAACGCCGTGGGCTCCCTCTACGT CAGGGAGGCGTTCCCTGGAGACAGCAAGAGC
ATGGTCAGAGAACTCATTGACAAGGTGCGGACAGTGT TGTGGAGACGCTGGACGAGCTG
GGCTGGATGGACGAGGAGTCCAAGAAGAAGGCGCAGGAGAAGGCCATGAGCATCCGGGAG
CAGATCGGGCACCCCTGACTACATCCTGGAGGAGATGAACAGGCGCCTGGACGAGGAGTAC
TCCAATCTGAACTTCTCAGAGGACCTGTACTTTGAGAACAGTCTGCAGAACCTCAAGGTG
GGCGCCACGCGGAGCCTCAGGAAGCTTCGGGAAAAGGTGGACCCAAATCTCTGGATCATC
GGGGCGGCGGTGGTCAATGCGTTCTACTCCCCAAACCGAAACCAGATTGTATTCCTGCC
GGGATCCTCCAGCCCCCTTCTTCAGCAAGGAGCAGCCACAGGCCTTGAACTTTGGAGGC
ATTGGGATGGTGATCGGGCACGAGATCACGCACGGCTTTGACGACAATGGCCGGA ACTTC
GACAAGAATGGCAACATGATGGATTGGTGGAGTAACTTCTCCACCCAGCACTTCCGGGAG
CAGTCAGAGTGCATGATCTACCAGTACGGCAACTACTCCTGGGACCTGGCAGACGAACAG
AACGTGAACGGATTCAACACCCTTGGGGAAAACATTGCTGACAACGGAGGGGTGCGGCAA
GCCTATAAGGCCTACCTCAAGTGGATGGCAGAGGGTGGCAAGGACCAGCAGCTGCCCGGC
CTGGATCTCACCCATGAGCAGCTCTTCTTCATCAACTACGCCAGGTGTGGTGCGGGTCC
TACCGGCCCGAGTTCGCCATCCAATCCATCAAGACAGACGTCCACAGTCCCCTGAAGTAC
AGGGTACTGGGGTCGCTGCAGAACCTGGCCGCCTTCGCAGACACGTTCCACTGTGCCCGG
GGCACCCCATGCACCCCAAGGAGCGATGCCGCGTGTGGTAG - 3'

表6: SEQ ID NO: 6的IGS5-蛋白质-2 ("IGS5PROT2")

```
MGKSEGPVGMVESAGRAGQKRPGFLEGLLLLLLLLVTAALVALGVLYADRRGIPEAQEVS
EVCTTPGCVIAAARILQNMPTTEPCDDFYQFACGGWLRRHVIPETNSRYSIFDVLRDEL
EVILKAVLENSTAKDRPAVEKARTLYRSCMNQSVIEKRGSQPLLDILEVGGWFPVAMDRW
NETVGLEWELERQLALMNSQFNRRVLIDLFIWNDDQNSSRHIIYIDQPTLGMPREYYFN
GGSNRKVREAYLQFMVSVATLLREDANLPRDSCLVQEDMMQVLELETQLAKATVPQEERH
DVIALYHRMGLEELQSQFGLKGFNWTLFIQTVLSSVKIKLLPDEEVVYGIPLYQNLNI
IDTYSARTIQNYLVWRLVLDRIGSLSQRFKDTRVNYRKALFGTMVEEVRWRECVGYVNSN
MENA VGSLYVREAFPGDSKSMVRELIDKVRTVFVETLDELGWMDEESKKKAQEKAMSIRE
```

```

QIGHDPDYILEEMNRRLDEEYSNLNFS2EDLYFEN3SLQNLK4VGAQRS5LRKLREK6VDPNLWII
GAAVVNAFYSPNRNQIVFPAGILQPPFFSKEQPQALNFGGIGM2VIGHEITHGFDDNGRNF
DKNGNMMDWWSNFS2TQHFREQSECM1IYQYGNYSWDLADEQNVNGFNTLGENIADNGGVRO
AYKAYLKWMAEGGKDQQLPGLDLTHEQLFFINYAQVWCGSYRPEFAIQSIKTDVHSPLKY
RVLGSLQNLAAAFADTFHCARGTPMHPKERCRVW

```

发明详述

定义

提供下列定义以便于理解本文中频繁使用的某些术语。

“IGS5”指包含SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6之一列出的氨基酸序列的多肽或其变体。因而，“IGS5”具体包括IGS5PROT、IGS5PROT1、和IGS5PROT2。

“酶活性”或“生物学活性”指所述IGS5的代谢或生理学功能，包括相似的活性、改进的活性、或非期望副作用降低的活性。还包括所述IGS5的抗原性和免疫原性。

“IGS5基因”指包含SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一列出的核苷酸序列的多核苷酸或其变体（如等位基因变体）和/或它们的互补物。

在用于本文时，“抗体”包括多克隆和单克隆抗体；嵌合的、单链的、和人源化的抗体；以及Fab片段，包括Fab或其它免疫球蛋白表达文库的产物。

“分离的”指“人为的”由天然状态改变和/或与天然环境分开。因而，若自然界中存在的组合物或物质成为“分离的”，则它发生了变化或与其最初环境分开，或者二者兼之。例如，存活动物中天然存在的多核苷酸或多肽不是“分离的”，但是与天然状态的共存物质分开的相同多核苷酸或多肽是“分离的”，正如本文采用的术语。

“多核苷酸”一般指任何多核糖核苷酸或多脱氧核糖核苷酸，可以是未修饰的RNA或DNA，或者是经修饰的RNA或DNA。“多核苷酸”包

括但不限于单链和双链DNA、单链与双链区混合的DNA、单链和双链RNA、单链与双链区混合的RNA、包含DNA与RNA的杂种分子（其中DNA和RNA可以是单链，通常是双链，或者是单链与双链区的混合）。另外，“多核苷酸”包括包含RNA或DNA或二者的三链区。术语“多核苷酸”还包括包含一种或多种经修饰碱基的DNA或RNA，以及为了稳定或其它原因而修饰了主链的DNA或RNA。“经修饰”碱基包括例如三苯甲基化碱基和非常见碱基（诸如肌苷）。可以对DNA和RNA进行多种修饰；因而，“多核苷酸”涵盖自然界中常常发现的多核苷酸的化学、酶促、或代谢修饰形式，以及病毒和细胞特征性的DNA和RNA化学形式。“多核苷酸”还涵盖相对较短的多核苷酸，通常称为寡核苷酸。

“多肽”指包含通过肽键或经修饰肽键（即peptide isosteres）彼此相连的两个或多个氨基酸的任何肽或蛋白质。“多肽”既包括短链，常常称为肽、寡肽、或寡聚物；又包括长链，一般称为蛋白质。多肽可包含20种基因编码氨基酸以外的氨基酸。“多肽”包括通过天然加工（诸如翻译后加工）或通过本领域众所周知的化学修饰技术而进行了修饰的氨基酸序列。这些修饰详细描述于基础教科书、更详细的专论、以及庞大的研究文献中。修饰可发生于多肽中的任何位置，包括肽主链、氨基酸侧链、和氨基或羧基末端。应当领会，相同类型的修饰可以相同或不同程度存在于指定多肽中的几个位点。同样，指定多肽可包含许多类型的修饰。作为遍在蛋白化作用的结果，多肽可以是分支的，也可以是具有或不具有分支的环状。环状的、分支的、和分支环状的多肽可来自翻译后天然加工，或者由合成方法产生。修饰包括乙酰化、酰化、ADP-核糖基化、酰胺化、生物素化、黄素的共价吸附、血红素部分的共价吸附、核苷酸或核苷酸衍生物的共价吸附、脂类或脂类衍生物的共价吸附、磷脂酰肌醇的共价吸附、交联、环化、二硫键的形成、脱甲基、共价交联的形成、胱氨酸的形成、焦谷氨酸的形成、甲酰化、 γ -羧化、糖基化、GPI锚的形成、羟化、碘化、甲基化、肉豆蔻酰化、氧化、蛋白水解加工、磷酸化、异戊二烯化、外消旋、硒化、硫化、由tRNA介导向蛋白质添加氨基酸（诸如精氨酰化）、

和遍在蛋白化作用。参阅例如《Proteins – Structure and Molecular Properties》（蛋白质 – 结构和分子特性），第2版，T. E. Creighton, W. H. Freeman and Company, 纽约, 1993; F. Wold, “翻译后蛋白质修饰: 观点和展望” (Posttranslational Protein Modifications: Perspectives and Prospects), 在《Post-translational Covalent Modification of Proteins》（蛋白质的翻译后共价修饰）一书中, 第1-12页, B. C. Johnson编, Academic出版社, 纽约, 1983; Seifter等人, 蛋白质修饰和非蛋白质辅助因子的分析 (Analysis for protein modifications and nonprotein cofactors), Meth. Enzymol., 182: 626 - 646, 1990; 和Rattan等人, “蛋白质合成: 翻译后修饰和衰老” (Protein Synthesis: Posttranslational Modifications and Aging), Ann. NY Acad. Sci., 663: 48 - 62, 1992.

“变体”指分别与参考多核苷酸或多肽不同但保留了本质特性的多核苷酸或多肽。典型的多核苷酸变体因核苷酸序列而不同于另一种参考多核苷酸。变体的核苷酸序列中的变化有可能改变或未改变参考多核苷酸所编码的多肽的氨基酸序列。正如下文所述, 核苷酸变化可在参考多核苷酸所编码的多肽中导致氨基酸替代、添加、删除、融合、和截短。典型的多肽变体因氨基酸序列而不同于另一种参考多肽。一般而言, 差异是有限的, 因此参考多肽与变体的序列在整体上密切相似, 在许多区域相同。变体与参考多肽的氨基酸序列可能因一处或多处替代、添加、和删除及其任意组合而不同。替代或插入的氨基酸残基可以是或不是由遗传密码子编码的。多核苷酸或多肽的变体可以是天然存在的 (诸如等位基因变体), 或者是非天然存在的。可通过诱变技术或直接合成来生成多核苷酸和多肽的非天然存在变体。

作为本领域一种对同一性的度量, “同一性”指通过比较序列而测定的两种或多种多肽序列或者两种或多种多核苷酸序列之间的相关性。在本领域, “同一性”还指可以通过这些序列之间的匹配, 例如总体上通过比对序列使得获得最大程度匹配, 由此而测定的多肽或多核苷酸之间的序列相关程度。因而, “同一性”和/或另一用词“相似

性”具有本领域认可的含义，而且可以通过已知方法容易的测定，包括但不限于《Computational Molecular Biology》（计算分子生物学），A. M. Lesk编，牛津大学出版社，纽约，1988；《Biocomputing: Informatics of Genome Projects》（生物计算：信息学和基因组计划），D. W. Smith编，Academic出版社，纽约，1993；《Computer Analysis of Sequence Data》（序列数据的计算机分析），第1部分，A. M. Griffin和H. G. Griffin编，Humana出版社，新泽西，1994；《Sequence Analysis in Molecular Biology》（分子生物学中的序列分析），G. von Heinje, Academic出版社，1987；《Sequence Analysis Primer》（序列分析引物），M. Gribskov和J. Devereux编，M Stockton出版社，纽约，1991；和H. Carillo和D. Lipton, SIAM J. Applied Math., 48: 1073, 1988中所描述的方法。用于测定同一性的优选方法设计用于给出待测序列之间的最大匹配。用于测定同一性和相似性的方法已编译成公众可获得的计算机程序。用于测定两种序列之间的同一性和相似性的优选计算机程序方法包括但不限于GCG程序软件包（J. Devereux等人, Nucleic Acids Research, 12 (1): 387, 1984）、BLASTP、BLASTN、FASTA（S. F. Altschul等人, J. Molec. Biol., 215: 403, 1990）。公众可以由NCBI和其它来源获得BLAST X程序（BLAST手册, S. Altschul等人, NCBI NLM NIH Bethesda, MD 20894; S. Altschul等人, J. Mol. Biol., 215: 403 - 410, 1990）。众所周知的Smith Waterman算法也可用于测定同一性。公众可获得的分别可用于测定多肽序列或多核苷酸序列的同一性或相似性的程序是来自Genetics Computer Group（Madison, WI）的“gap”（缺口）程序，常常以用于比较的缺省参数（而且末端缺口没有罚分）运行该程序。用于多肽序列比较的优选（即缺省）参数包括：Needleman和Wunsch, J. Mol. Biol., 48: 443 - 453, 1970描述的算法；Hentikoff和Hentikoff, Proc. Natl. Acad. Sci. USA, 89: 10915 - 10919, 1992的比较矩阵BLOSSUM62；缺口罚分 = 12；缺口长度罚分 = 14。用于多核苷酸序列比较的优选（即缺省）参数包括：Needleman和Wunsch, J. Mol. Biol., 48: 443 - 453, 1970

描述的算法；比较矩阵，匹配 = +12，错配 = 0；缺口罚分 = 50；缺口长度罚分 = 3。用词“同源性”可替代用词“同一性”。

例如，具有与参考核苷酸序列（例如选自SEQ ID NO: 1、SEQ ID NO: 3、和SEQ ID NO: 5的参考核苷酸序列）具有至少95%“同一性”的核苷酸序列的多核苷酸意指，多核苷酸的核苷酸序列与参考序列大致相同，只是在参考核苷酸序列的每100个核苷酸中多核苷酸序列可以包含多达5个点突变。换言之，为了获得具有与参考核苷酸序列至少95%同一的核苷酸序列的多核苷酸，可以在参考序列中删除或用其它核苷酸替代多达5%的核苷酸，或者可以在参考序列中插入多达参考序列核苷酸总数5%的任何数目核苷酸，或者参考序列中多达核苷酸总数5%的任何数目核苷酸可以是删除、插入、或替代的组合。参考序列的这些突变可以发生于参考核苷酸序列的5'或3'末端位置或者这些末端位置之间的任何位置，它们或是以单个核苷酸散布于参考序列中，或是以一个或多个连续组散布于参考序列中。

相似的，例如，具有与参考氨基酸序列（例如选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的参考氨基酸序列）具有至少95%“同一性”的氨基酸序列的多肽意指，多肽的氨基酸序列与参考序列大致相同，只是在参考氨基酸序列的每100个氨基酸中多肽氨基酸序列可以包含多达5个氨基酸改变。换言之，为了获得具有与参考氨基酸序列至少95%同一的氨基酸序列的多肽，可以在参考序列中删除或用其它氨基酸替代多达5%的氨基酸残基，或者可以在参考序列中插入多达参考序列氨基酸残基总数5%的任何数目氨基酸。参考序列的这些改变可以发生于参考氨基酸序列的氨基或羧基末端位置或者这些末端位置之间的任何位置，它们或是以单个残基散布于参考序列中，或是以一个或多个连续组散布于参考序列中。

“同系物”是本领域用于指示与目的序列具有高度序列相关性的多核苷酸或多肽序列的一般术语。如本文所述，这种相关性可以通过测定待比较序列间的同一性和/或相似性程度来量化。这个一般术语包括术语“直向进化同源物（ortholog）”，指其它物种中功能与该多

核苷酸或多肽等同的多核苷酸或多肽，以及“平行进化同源物（paralog）”，指相同物种中的功能相似序列。因此，例如在人类中，在内皮素转化酶家族中，ECE-1是其它成员（如ECE-2）的平行进化同源物。

“融合蛋白”指由两个（通常是无关的）基因或其片段融合在一起而编码的蛋白质。该术语可以例示为包含免疫球蛋白分子恒定区各部分与另一种人类蛋白质或其部分的融合蛋白。在许多情况中，采用免疫球蛋白Fc区作为融合蛋白的一部分对于治疗和诊断是有利的，例如可导致药代动力学特性得到改进（参阅例如EP-A 0232 262）。另一方面，对于有些用途，将会希望能够在表达、检测、和纯化融合蛋白之后删除Fc部分。

本发明的多肽

本发明涉及IGS5多肽（或IGS5酶，如IGS5PROT、IGS5PROT1、或IGS5PROT2），特别是人类IGS5多肽（或人类IGS5酶），还涉及包含所述完整IGS5多肽实质性部分的IGS5多肽片段。因而，在第一个方面，本发明的IGS5多肽包括分离多肽，特别是分离的人类物种多肽，所述多肽所含氨基酸序列与选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的氨基酸序列在其全长上具有至少70%同一性、优选至少80%特别是至少85%同一性、更优选至少90%同一性、仍更优选至少95%同一性、最优选至少97-99%同一性。这些多肽包括包含选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的氨基酸序列的多肽。

在第二个方面，本发明的IGS5多肽包括分离多肽，特别是分离人类IGS5多肽，所述多肽所具有的氨基酸序列与选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的氨基酸序列在其全长上具有至少70%同一性、优选至少80%特别是至少85%同一性、更优选至少90%同一性、仍更优选至少95%同一性、最优选至少97-99%同一性。这些多肽包括SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的IGS5多肽。

本发明的其它多肽包括分离IGS5多肽，特别是人类物种多肽，所

述多肽包含SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6之一所含序列。

本发明的多肽是金属蛋白酶多肽家族的成员。它们受到关注是因为已经鉴定了几种机能障碍、紊乱、或疾病，其中金属蛋白酶在疾病的病理学中担当重要角色。认为本发明多肽和多核苷酸有用的疾病的范例包括：CNS紊乱，包括精神分裂症，发作性阵发性焦虑症（EPA）紊乱诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经退化疾病，严重精神发育迟缓，和运动障碍，诸如亨廷顿病或图雷特综合症，厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意力缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心率失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食道反流病（GERD），活动能力障碍，和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒症；糖尿病并发症等。本发明的多肽在心血管疾病领域特别受到关注。此外，本发明的IGS5多肽在鉴定这些多肽的刺激物或移植物、提供用于检测与不适当IGS5活性或水平有关的疾病的诊断测定法、和使用经鉴定为刺激物或抑制物的化合物来治疗与IGS5不平衡有关的状况方面受到关注。因此，本发明的IGS5多肽可用于设计或筛选选择性刺激物或抑制物，并由此可导致新型药物的开发。本发明的IGS5多肽（特别是人类物种IGS5多肽）的特性在下文中称为“IGS5活

性”、“IGS5多肽活性”、或“IGS5的生物学活性”。这些活性还包括所述IGS5多肽的抗原性和免疫原性，特别是选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的多肽的抗原性和免疫原性。优选的是，本发明的多肽展示IGS5（优选人类IGS5）的至少一种生物学活性。

本发明的IGS5多肽可以是“成熟”蛋白质的形式，或者是较大蛋白质（诸如前体或融合蛋白）的一部分。包含额外的氨基酸序列常常是有利的，这些额外氨基酸序列包括分泌序列或前导序列、原序列、有助于纯化的序列（诸如多个组氨酸残基）、或在重组生产过程中有助于稳定的额外序列。

本发明还包括上述多肽的变体，即因保守氨基酸替代（用具有相似特征的残基进行替代）而与参照物不同的多肽。典型的这类替代是Ala、Val、Leu、与Ile之间；Ser与Thr之间；酸性残基Asp与Glu之间；Asn与Gln之间；和碱性残基Lys与Arg之间；或芳香族残基Phe与Tyr之间的替代。特别优选的是以任意组合替代、删除、或添加几个、5-10、1-5、1-3、1-2或1个氨基酸的变体。

本发明还包括IGS5多肽的片段，特别是包含完整IGS5多肽的实质性部分的IGS5多肽片段。片段指具有与上述IGS5多肽氨基酸序列的一部分而非全部相同的氨基酸序列的多肽。与IGS5多肽一样，片段可以是“独立的”，或者包含于较大多肽中而形成一部分或区域，最优选作为单个连续区域。

优选片段包括例如具有IGS5多肽的氨基酸序列但删除了包含氨基末端的一系列连续残基或包含羧基末端的一系列连续残基或删除了两段连续残基（一段包含氨基末端，另一段包含羧基末端）的截短型多肽。同样优选由结构或功能属性表征的片段，诸如包含 α -螺旋和 α -螺旋形成区、 β -折叠和 β -折叠形成区、转角和转角形成区、卷曲和卷曲形成区、亲水区、疏水区、 α 两亲区、 β 两亲区、柔性区、表面形成区、底物结合区、和高抗原指数区的片段。其它优选片段是生物学活性片段。生物学活性片段是那些介导酶活性的片段，包括那些具有相似活性、改进活性、或非期望活性降低的片段。还包括那些在动

物（尤其是人）中具有抗原性或免疫原性的片段。

关于本发明的变体，即包含SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6之一所示完整IGS5多肽的实质性部分的多肽片段，术语“实质性”指该IGS5多肽片段具体具有至少大约50个氨基酸的大小，优选至少大约100个氨基酸的大小，更优选至少大约200个氨基酸的大小，最优选至少大约300个氨基酸的大小。在本文中，“大约”包括比具体叙述的数值多或少几个、5、4、3、2、或1个氨基酸。本发明的IGS5多肽片段优选至少在一定程度上展示IGS5多肽本身特征性特性的至少一种。

关于本发明的IGS5多肽，发现它们可能涉及生物学活性肽的代谢。具体而言，发现这些IGS5多肽是可能作用于多种血管活性肽的金属蛋白酶类酶。现有技术中已知的血管活性肽包括例如心房钠尿肽（ANP）、缓激肽、大内皮素（大ET-1）、内皮素（ET-1）、物质P、和血管紧张素-1。在本发明的内容中，发现IGS5胞外结构域是一种新型人类金属蛋白酶，例如它在体外水解多种上述血管活性肽，包括大ET-1、ET-1、ANP、和缓激肽。

此外，本发明的IGS5金属蛋白酶类酶可能受到参考化合物的抑制，所述参考化合物用于测定对具有ECE/NEP特征的酶的抑制特性，例如受到诸如磷酸二肽等化合物的抑制。没有观察到特异性抑制NEP的参考化合物对IGS5具有抑制作用，例如IGS5不受诸如thiorphan等化合物的抑制。也没有观察到特异性抑制ECE的参考化合物对IGS5具有抑制作用，例如IGS5不受诸如选择性ECE抑制剂CGS-35066等化合物的抑制（De Lombart等人，J. Med. Chem., 43(3): 488-504, 2000年2月10日）。这些参考化合物对本发明IGS5金属蛋白酶类酶的抑制数据描述于下文实验部分，特别是实施例7。

可以任何合适方式制备本发明的多肽。这些多肽包括分离的天然存在的多肽、重组生产的多肽、合成生产的多肽、或由这些方法联合产生的多肽。用于制备这些多肽的方法在本领域是众所周知的。

本发明的多核苷酸

另一方面，本发明涉及IGS5多核苷酸（如IGS5DNA、IGS5DNA1、或IGS5DNA2），特别是人类IGS5多核苷酸。这些多核苷酸包括分离多核苷酸，优选分离人类物种多核苷酸，所述多核苷酸所含核苷酸序列所编码的多肽与选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的氨基酸序列在其全长上具有至少70%同一性、优选至少80%特别是至少85%同一性、更优选至少90%同一性、仍更优选至少95%同一性。在这点上，高度优选编码的多肽具有至少97%同一性的多核苷酸，更高度优选具有至少98-99%和至少99%同一性的多核苷酸，特别是最高度优选具有99.9%同一性的多核苷酸。这些多核苷酸包括包含分别编码多肽SEQ ID NO: 2、SEQ ID NO: 4、或SEQ ID NO: 6的SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一所含核苷酸序列的多核苷酸。

在这方面的变体中，本发明的多核苷酸包括分离多核苷酸，特别是分离人类多核苷酸，所述多核苷酸所含核苷酸序列与编码选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的多肽的核苷酸序列在完整编码区上具有至少70%同一性、优选至少80%特别是至少85%同一性、更优选至少90%同一性、仍更优选至少95%同一性。在这点上，高度优选具有至少97%同一性的多核苷酸，更高度优选具有至少98-99%和至少99%同一性的多核苷酸，特别是最高度优选具有99.9%同一性的多核苷酸。

本发明的其它多核苷酸包括分离多核苷酸，特别是分离人类多核苷酸，所述多核苷酸所含核苷酸序列与选自SEQ ID NO: 1、SEQ ID NO: 3、和SEQ ID NO: 5的核苷酸序列之一在其全长上具有至少70%同一性、优选至少80%特别是至少85%同一性、更优选至少90%同一性、仍更优选至少95%同一性。具体而言，本发明的多核苷酸包括具有与参考多核苷酸序列在其全长上具有至少70%同一性、优选至少80%特别是至少85%同一性、更优选至少90%同一性、仍更优选至少95%同一性的核苷酸序列的分离多核苷酸。在这点上，高度优选包含或具有与选自SEQ ID NO: 1、SEQ ID NO: 3、和SEQ ID NO: 5的核苷酸序列之一具有至少97%同一性、更高度优选具有至少98-99%和至少99%同一

性、特别是最高度优选具有99.9%同一性的核苷酸序列的多核苷酸。这些多核苷酸包括包含SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5的多核苷酸的多核苷酸，以及SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5的多核苷酸自身，特别是人类物种多核苷酸。

本发明还提供了与所有上述多核苷酸互补的多核苷酸。

SEQ ID NO: 1的核苷酸序列（称为“IGS5DNA”）是人类起源（Homo sapiens）的cDNA序列，长度为2076个核苷酸，所含多肽编码序列（第1-2073位核苷酸）编码691个氨基酸的多肽，即SEQ ID NO: 2的多肽（称为“IGS5PROT”）。编码SEQ ID NO: 2多肽的核苷酸序列可以与SEQ ID NO: 1所含多肽编码序列相同，或者可以是与SEQ ID NO: 1所含序列不同的序列，但是因为遗传密码的冗余（简并性）也编码SEQ ID NO: 2的多肽。

SEQ ID NO: 3的核苷酸序列（称为“IGS5DNA1”）是人类起源（Homo sapiens）的cDNA序列，长度为2340个核苷酸（包含终止密码子标签），所含多肽编码序列（第1-2337位核苷酸）编码779个氨基酸的多肽，即SEQ ID NO: 4的多肽（称为“IGS5PROT1”）。编码SEQ ID NO: 4多肽的核苷酸序列可以与SEQ ID NO: 3所含多肽编码序列相同，或者可以是与SEQ ID NO: 3所含序列不同的序列，但是因为遗传密码的冗余（简并性）也编码SEQ ID NO: 4的多肽。

SEQ ID NO: 5的核苷酸序列（称为“IGS5DNA2”）是人类起源（Homo sapiens）的cDNA序列，长度为2262个核苷酸（包含终止密码子标签），所含多肽编码序列（第1-2259位核苷酸）编码753个氨基酸的多肽，即SEQ ID NO: 6的多肽（称为“IGS5PROT2”）。编码SEQ ID NO: 6多肽的核苷酸序列可以与SEQ ID NO: 5所含多肽编码序列相同，或者可以是与SEQ ID NO: 5所含序列不同的序列，但是因为遗传密码的冗余（简并性）也编码SEQ ID NO: 6的多肽。

下文更详细的描述了由本发明多核苷酸编码的多肽种类的特征。

金属蛋白酶的生物学和药物学特性

本发明的多肽（特别是人类物种多肽）与金属蛋白酶家族的其它蛋白质在结构和功能上相关，例如显示与金属蛋白酶或相关酶（诸如MMP、ACE、ECE、或NEP）具有同源性和/或结构相似性。因而，例如SEQ ID NO: 2的多肽与金属蛋白酶家族的其它蛋白质在结构和功能上相关，具有与酶诸如NEP或ECE（如ECE-1）特别是NEP的同源性和/或结构相似性。因而，预计本发明的优选多肽和多核苷酸特别具有与其同源多肽和多核苷酸相似的生物学功能/特性。此外，本发明的优选多肽和多核苷酸具有至少一种IGS5活性。

特别是对本发明而言，上文早已描述了金属蛋白酶的一般特性及活性。为了进一步理解本发明多肽和多核苷酸的本质和特征，特别是这些多肽和多核苷酸的功能，下文分别概述了每一种酶像MMP、ACE、ECE、或NEP的一些更具体特性。

基质金属蛋白酶（MMP），也称为基质素（matrixin），是在胞外基质成份的更新中发挥功能的锌金属蛋白酶家族。至今，在人中已经鉴定了基质素家族的几个成员。许多细胞类型，诸如成纤维细胞、上皮细胞、巨噬细胞、淋巴细胞、和癌细胞，合成并分泌MMP。在所有生成细胞中（除了嗜中性粒细胞以外），MMP以前酶原的形式合成，并注定以酶原的形式分泌。在生理学条件下，这些酶在形态发生、组织重建、和吸收中担当重要角色。在过量情况中，它们参与破坏与许多结缔组织疾病诸如关节炎、牙周炎、肾小球肾炎、癌细胞侵入和转移有关的胞外基质。因而，MMP在例如正常胚胎形成和组织重建中和在许多疾病诸如关节炎、癌症、牙周炎、肾小球肾炎、脑脊髓炎、动脉粥样硬化、和组织溃疡中担当重要角色。基质素在结缔组织基质的生理学和病理学分解代谢中的重要性已经受到重视，因为在正常稳态组织中很少能够检测到MMP活性，但是许多MMP的合成受到炎症细胞因子、激素、生长因子、和细胞转化的转录调控。MMP的生物学活性还在无活性前体（MMP原）的激活步骤中以及通过与胞外基底和内源抑制物的相互作用受到胞外控制。MMP是涉及胞外基质降解和重建的一类重要锌依赖性金属蛋白酶。这些酶的抑制物在例如癌症、关节炎、骨质疏松、和

阿尔茨海默症中具有治疗潜能，而且这些抑制物中的几种正在进行临床评估。

血管紧张素-1转化酶（ACE；氨基二肽酶A；EC 3.4.15.1）是锌金属蛋白酶中血管紧张素转化酶家族的成员。ACE主要表达于内皮细胞、上皮细胞、和神经上皮细胞的表面（躯体ACE）作为胞外酶，即包括它的催化位点在内的大部分质量锚定在质膜上并面向胞外环境。在血管内皮细胞的质膜中发现了ACE，在肺部血管内皮细胞表面发现了高水平ACE，推测ACE的活性位点可代谢循环物质。除了内皮细胞定位以外，ACE还表达于小肠吸收性上皮细胞和肾近曲小管的刷状缘中。ACE还发现于单核的细胞（诸如巨噬细胞分化后的单核细胞和T淋巴细胞）和成纤维细胞中。采用放射性标记的特异性ACE抑制物的体外放射自显影和免疫组织化学研究已经将ACE的重要位置定位于脑。ACE主要发现于脉络丛，它可能是脑脊液、室管膜、穹窿下器、基底神经节（尾状核—豆状核和苍白球）、黑质、和垂体中ACE的来源。已经在许多生物学流体中检测到可溶性形式的ACE，诸如血清、精液、羊水、和脑脊液。可溶性形式的ACE似乎衍生自内皮细胞中的膜结合形式的酶。ACE的主要生物学活性是由血管紧张素-1切除C端二肽而生成有效的血管加压肽—血管紧张素-2，并通过随后除去两个C端二肽而灭活血管舒张肽—缓激肽。由于ACE涉及这两种血管活性肽（血管紧张素-2和缓激肽）的新陈代谢，ACE已成为了高血压和充血性心力衰竭治疗的决定性分子靶。这导致了对高度有效且特异性的ACE抑制物的开发，这些抑制物已经成为临床上重要的且广泛使用的口服活性药物，用于控制这些高血压和充血性心力衰竭状况。代谢血管活性肽是对ACE了解得最透彻的生理学功能，然而，由于ACE的定位和/或对一系列生物学活性肽的体外切割，推测这种酶还涉及与血压调控无关的一系列其它生理学过程，诸如免疫、生殖、和神经肽代谢。

中性内肽酶（NEP, neprilysin, EC 3.4.24.11）是一种锌金属蛋白酶，归类为neprilysin家族的成员。NEP最初由兔肾刷状缘膜分离。稍后，在大鼠脑中鉴定了涉及阿片类肽脑啡肽降解的NEP样酶。胞外酶

NEP的克隆及随后的定点诱变实验显示,除了具有与嗜热芽孢菌蛋白酶相似的特异性以外,它还具有相似的活性位点组织。对于切割肽中疏水性残基的N端一侧,NEP还显示嗜热芽孢菌蛋白酶样特异性。关于NEP的一般分布,已经测定了脑和脊髓,而且损伤和电子显微镜研究总体上支持NEP的主要神经元定位现象,然而该酶能够存在于条纹灰质和条纹黑质途径的纤维周围的少突胶质细胞上和周围神经系统的施旺细胞上。NEP并未显示集中在特定膜界面(诸如突触)上,而是均一的分布于神经元核周质和树突的表面。在周边,NEP在肾和肠的刷状缘膜、淋巴结、和胎盘中特别丰富,而且以更低浓度发现于许多其它组织中,包括主动脉血管壁。已发现常见急性成淋巴细胞白血病抗原就是NEP,现有技术还显示酶瞬时存在于淋巴造血细胞表面,并在某些疾病状态下在成熟淋巴细胞上以升高的水平存在。对NEP的临床关注,特别是对NEP抑制物作为潜在临床试剂的关注,起源于NEP与另一种锌金属蛋白酶氨肽酶N(APN,膜丙氨酰氨肽酶,EC 3.4.11.2)在脑啡肽降解中的作用和它在心房钠尿肽(ANP)降解中的作用。例如,已知NEP和血管紧张素转化酶(ACE)的双重抑制物是有效的抗高血压药物,因为它同时升高心房钠尿肽的循环水平(由于NEP抑制)和降低血管紧张素-2的信号水平(由于ACE抑制)。由于周边酶显示降解循环钠尿肽和利尿肽(心房钠尿肽),这引起了对NEP抑制物的临床潜能的进一步关注。因此对NEP抑制物研究抗高血压特性。由另一个范例知道,合成NEP抑制物thiorphan对脑啡肽代谢的抑制在小鼠中引起纳洛酮(naloxone)可逆的抗疼痛应答。这说明有可能通过升高靶受体区域的内源阿片样物质的水平来镇痛,而相对没有吗啡或其它经典鸦片剂药物的副作用。现已认识到,为了实现任何显著效果,还应当抑制其它脑啡肽代谢酶,特别是氨肽酶N(APN)。这些双重NEP/APN抑制物完全阻断脑啡肽代谢,而且具有强烈的抗疼痛特性。

内皮素转化酶(ECE)催化有效血管收缩肽内皮素(ET)生物合成的最终步骤。这包括切割无活性中间物大内皮素的Trp-Val键。ECE-1是与中性内肽酶(NEP; neprilysin; EC 3.4.24.11,见上文)同源

的一种锌金属蛋白酶。与NEP一样，ECE-1受到化合物膦酰二肽的抑制，而且是II型整合膜蛋白。然而，与NEP不同的是，ECE-1以二硫键连接的二聚体形式存在，而且不受其它NEP抑制物诸如thiorphan的抑制。免疫细胞化学研究指示ECE-1主要定位于细胞表面，其中它作为胞外酶存在。ECE-1定位于内皮细胞和一些分泌细胞（如胰 β 细胞和平滑肌细胞）。有效且选择性的ECE抑制物或ECE/NEP双重抑制物可能在心血管和肾医疗中具有治疗应用。内皮素（ET）是由21个氨基酸通过两个分子内二硫键构成的双环肽，它是至今鉴定的最有效的血管收缩肽之一，而且对动物的施用导致血压持续升高，说明了它在心血管调控中的潜在角色。ET-1在人体中的内源生成有助于维持基础血管张力。因此，内皮素系统和相关酶（像ECE）代表了新型制药学试剂开发的可能候选者。因而，对ECE的临床关注，特别是对ECE抑制物作为潜在临床试剂的关注，起源于ECE的作用，特别是在ET生物合成方面。因此，显示显著内皮素转化酶抑制活性的化合物可用于治疗和预防由ET诱导或怀疑由它诱导的多种疾病，诸如心血管疾病，包括心力衰竭、心绞痛、心率失常、心肌梗死、心脏肥大、低血压、高血压（如原发性高血压、肾性高血压、或肺动脉高血压）、血栓形成、动脉硬化、脑血管痉挛、蛛网膜下出血、脑缺血、脑梗死、周围血管疾病、雷诺病、肾病（如肾衰竭）、哮喘、中风、阿尔茨海默症；糖尿病并发症；溃疡，诸如胃溃疡；癌症，诸如肺癌；内毒素休克；脓毒病等。

本发明的多肽在心血管疾病领域特别受到关注。

用于获得本发明多核苷酸的流程

可以通过标准克隆和筛选技术，使用表达序列标签（EST）分析，由衍生自人睾丸组织细胞mRNA的cDNA文库获得本发明的多核苷酸（M. D. Adams等人，*Science*, 252: 1651 - 1656, 1991; M. D. Adams等人，*Nature*, 355: 632 - 634, 1992; M. D. Adams等人，*Nature*, 增刊 377: 3 - 174, 1995）。还可以由天然来源诸如基因组DNA文库获得本发明的多核苷酸，或者可以使用众所周知的商品化技术进行合成（如

F.M. Ausubel等人, 《Current Protocols in Molecular Biology》(分子生物学现行方案), 2000)。

在将本发明多核苷酸用于本发明多肽的重组生产时, 多核苷酸自身可包含成熟多肽的编码区; 或者在相同读码框中包含成熟多肽的编码区以及其它编码序列, 诸如编码前导序列或分泌序列、蛋白原序列、前蛋白序列、前蛋白原序列、或其它融合肽部分的序列。例如, 可编码有助于纯化融合多肽的标记序列。在本发明这方面的某些优选实施方案中, 标记序列是六组氨酸肽, 如pQE载体(Qiagen公司)中提供的, 且描述于Gentz等人, Proc. Natl. Acad. Sci. USA, 86: 821-824, 1989, 或HA标签。多核苷酸还可包含5'和3'非编码序列, 诸如转录但不翻译的序列、剪接和聚腺苷酸化信号、核糖体结合位点、和稳定mRNA的序列。

本发明的其它实施方案包括编码包含选自SEQ ID NO: 2、SEQ ID NO: 4、和SEQ ID NO: 6的氨基酸序列之一但其中几个例如5-10、1-5、1-3、1-2、或1个氨基酸残基以任意组合被替代、删除、或添加的多肽变体的多核苷酸。

与SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一所含核苷酸序列同一或足够同一的多核苷酸, 可作为cDNA和基因组DNA的杂交探针或者作为核酸扩增反应(PCR)的引物, 用于分离编码本发明多肽的全长cDNA和基因组克隆, 和用于分离与SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一具有高度序列相似性的其它基因(包括编码来自人类来源的平行进化同源物的基因和来自非人物种的直向进化同源物和平行进化物的基因)的cDNA和基因组克隆。通常, 这些核苷酸序列与参照物至少70%同一, 优选至少80%特别是至少85%同一, 更优选至少90%同一, 最优选至少95%同一。探针或引物通常将包含至少15个核苷酸, 优选至少30个核苷酸, 而且可具有至少50个核苷酸。特别优选的探针将具有30-50个核苷酸。特别优选的引物将具有20-25个核苷酸。

可以通过包括下列步骤的方法来获得编码本发明多肽(包括同系

物和来自非人物种的直向进化同源物)的多核苷酸:在严谨杂交条件下用具有SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一的序列或其片段的经标记探针筛选合适文库;并分离包含所述多核苷酸序列的全长cDNA和基因组克隆。这些杂交技术对熟练技术人员是众所周知的。优选的严谨杂交条件包括在含50% 甲酰胺、5x SSC(150mM NaCl、15mM 柠檬酸三钠)、50mM 磷酸钠(pH7.6)、5x Denhardt氏液、10% 硫酸葡聚糖(w/v)、和20 μg/ml 变性剪碎鲑鱼精DNA的溶液中于42℃保温过夜,随后用0.1x SSC于大约65℃清洗滤膜。因而,本发明还包括可通过在严谨杂交条件下用具有SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一的序列或其片段的经标记探针筛选合适文库而获得的多核苷酸。

熟练技术人员将领会,在许多情况中,分离cDNA序列将是不完整的,即多肽编码区在cDNA的5'端是切短的。这是逆转录酶具有固有低“持续合成能力”(对酶在聚合反应过程中保持附着在模板上的能力的量度),不能在第一条cDNA链合成过程中完成mRNA模板的DNA拷贝导致的。

本领域熟练技术人员众所周知几种商品化方法可用于获得全长cDNA或延伸短cDNA,例如基于cDNA末端快速扩增(RACE)的方法(参阅例如Frohman等人,PNAS USA, 85: 8998-9002, 1988)。这种技术的最新修改,例如Marathon™技术(Clontech Laboratories公司),显著简化了对较长cDNA的搜索。在Marathon™技术中,由提取自选定组织的mRNA制备cDNA,并将“衔接头”序列连接到每个末端。然后使用基因特异的和衔接头特异的寡核苷酸引物对进行核酸扩增反应(PCR)以扩增cDNA中“缺失的”5'末端。然后使用“嵌套”引物对重复PCR反应,其中“嵌套”引物即为设计成在扩增产物内退火的引物,通常是在衔接头序列3'端退火的衔接头特异引物,和在已知基因序列5'端退火的基因特异引物。然后通过DNA测序分析此反应的产物,并通过将产物直接连接现有cDNA以产生完整序列,或通过将新的序列信息用于设计5'端引物而进行另一次全长PCR,由此构建全长cDNA。

载体、宿主细胞、表达

可以通过本领域众所周知的过程由经遗传工程改造的包含表达系统的细胞制备本发明的重组多肽。因此，在另一个方面，本发明涉及包含本发明多核苷酸的表达系统，用这些表达系统遗传改造的宿主细胞，和通过重组技术生产本发明的多肽。使用由本发明DNA构建物衍生的RNA，无细胞翻译系统也可用于生产这些蛋白质。

为了重组生产，可遗传改造宿主细胞以掺入本发明多核苷酸的表达系统或其部分。可通过许多标准实验室指南(诸如Davis等人,《Basic Methods in Molecular Biology》(分子生物学基本方法), 1986; 和Sambrook等人,《Molecular Cloning: A Laboratory Manual》(分子克隆: 实验室指南), 第2版, 冷泉港实验室出版社, 冷泉港, 纽约, 1989)中描述的方法, 将多核苷酸导入宿主细胞。优选方法包括例如磷酸钙转染、由DEAE-葡聚糖介导的转染、转位、显微注射、由阳离子脂类介导的转染、电穿孔、转导、刮擦装载(scrape loading)、轰击导入、或感染。

合适宿主的代表性范例包括细菌细胞, 诸如链球菌、葡萄球菌、大肠杆菌、链霉菌、和枯草芽孢杆菌细胞; 真菌细胞, 诸如酵母细胞和曲霉细胞; 昆虫细胞, 诸如果蝇S2和Spodoptera Sf9细胞; 动物细胞, 诸如CHO、COS、HeLa、C127、3T3、BHK、HEK293、和Bowes黑素瘤细胞; 和植物细胞。

可使用很多种表达系统, 例如由染色体、游离体、和病毒衍生的系统, 如由细菌质粒, 细菌噬菌体, 转座子, 酵母游离体, 插入元件, 酵母染色体元件, 病毒诸如杆状病毒、乳多空病毒(诸如SV40)、痘苗病毒、腺病毒、禽痘病毒、假狂犬病病毒、和逆转录病毒衍生的载体, 以及由上述联合衍生的载体, 诸如由质粒和噬菌体遗传元件衍生的载体, 诸如粘粒和噬菌粒。表达系统可包含调控以及引起表达的控制区。一般而言, 可使用能够在宿主中维持、增殖、或表达多核苷酸产生多肽的任何系统或载体。可通过多种众所周知的常规技术, 诸如

Sambrook等人《Molecular Cloning: A Laboratory Manual》(见上文)中所述,将合适的核苷酸序列插入表达系统。为了使翻译产生的蛋白质分泌进入内质网内腔、周质空间、或胞外环境,可将合适的分泌信号掺入目的多肽。这些信号对于多肽而言可以是内源的,或者是外源的(即衍生自不同物种)。

若要表达本发明多肽用于筛选实验,通常有可能在细胞表面产生多肽或者是可溶性蛋白质的形式表达。若将多肽分泌进入培养基,则可回收培养基以回收并纯化多肽。若生成于胞内,则在回收多肽前首先必须裂解细胞。若多肽结合于细胞表面(膜结合型多肽),则通常制备膜部分,以累积膜结合型多肽。

可通过众所周知的方法由重组细胞培养物回收并纯化本发明的多肽,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、疏水相互作用层析、亲和层析、羟基磷灰石层析、和凝集素层析。最优选的是,采用高效液相层析来进行纯化。当多肽在胞内合成、分离、和/或纯化过程中发生变性时,可采用众所周知的用于蛋白质重新折叠的技术来重新产生有活性的构象。

诊断实验

本发明还涉及本发明多核苷酸作为诊断剂的用途。与机能障碍有关且表征为选自SEQ ID NO: 1、SEQ ID NO: 3、和SEQ ID NO: 5的多核苷酸的基因突变型的检测将提供诊断工具,从而增加或明确由基因表达的不足、过度、或时空改变引起的疾病或其易感性的诊断。通过多种技术,可以在DNA水平上检测基因内携带突变的个体。

可由受试者的细胞,诸如由血液、尿液、唾液、组织切片、或尸解材料,获得用于诊断的核酸。基因组DNA可直接用于检测,或者可在分析前通过PCR或其它扩增技术进行酶促扩增。RNA或cDNA也可以相似方式使用。可通过扩增产物与正常基因型相比的大小变化来检测删除和插入。可通过将扩增后的DNA与经标记IGS5核苷酸序列进行杂交来鉴定点突变。可通过RNA酶消化或溶解温度的差异来区分完美匹配的序列

与错配的双链体。还可通过DNA片段在含或不含变性剂的凝胶中的电泳迁移率的变化，或者通过直接的DNA测序来检测DNA序列差异（参阅如Myers等人，*Science*, 230: 1242, 1985）。还可通过核酸酶保护实验（诸如RNA酶和S1保护）或化学切割法来揭示特定位置的序列变化（参阅Cotton等人，*Proc. Natl. Acad. Sci. USA*, 85: 4397 - 4401, 1985）。在另一个实施方案中，可构建包含IGS5核苷酸序列或其片段的寡核苷酸探针阵列来进行如遗传突变的高效筛选。阵列技术的方法是众所周知的，而且具有普遍适用性，可用于在分子遗传学中解决多种问题，包括基因表达、遗传连锁、和遗传可变性（参阅例如M.Chee等人，*Science*, 274: 610 - 613, 1996）。

诊断实验提供了通过上述方法检测IGS5基因中的突变来诊断或确定对疾病易感性的方法。另外，可通过包括由衍生自受试者的样品测定多肽或mRNA水平的异常降低或升高的方法来诊断这些疾病。可使用任何本领域众所周知的用于多核苷酸定量的技术在RNA水平测量表达的降低或升高，诸如PCR、RT-PCR、RNA酶保护、Northern印迹、和其它杂交方法。可用于测定衍生自宿主的样品中蛋白质（诸如本发明多肽）水平的实验技术对本领域技术人员是众所周知的。这些实验方法包括放射性免疫实验、竞争性结合实验、Western印迹分析、和ELISA实验。

因而，另一方面，本发明涉及诊断试剂盒，其中包含：

(a) 本发明的多核苷酸，优选SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5之一的多核苷酸或其片段；

(b) 与(a)互补的核苷酸序列；

(c) 本发明的多肽，优选SEQ ID NO: 2、SEQ ID NO: 4、或SEQ ID NO: 6之一的多肽或其片段；或

(d) 针对本发明多肽，优选SEQ ID NO: 2、SEQ ID NO: 4、或SEQ ID NO: 6之一的多肽的抗体。

应当领会，在任何这些试剂盒中，(a)、(b)、(c)、或(d)可构成所述试剂盒的实质性成份。这样的试剂盒将用于诊断疾病或对

疾病的易感性，特别是上文（即本发明多肽的内容中）所述疾病等。

染色体实验

本发明的核苷酸序列对于染色体定位也是有价值的。将序列特异靶向单个人染色体上的特定位置，并能够与其发生杂交。依照本发明将相关序列对染色体进行作图是将那些序列与基因相关疾病联系起来的重要的第一步。一旦将序列定位于精确的染色体位置，可使序列在染色体上的物理位置与遗传图谱资料相关联。这些资料可在例如 V. McKusick 的《Mendelian Inheritance in Man》（人类孟德尔遗传）（可通过网络由 Johns Hopkins 大学 Welch 医学图书馆获得）中找到。然后通过连锁分析（物理上相邻基因的共遗传）鉴定基因与已定位于相同染色体区域的疾病之间的关系。

还可测定受影响与不受影响的个体之间的 cDNA 或基因组序列差异。若在有些或所有受影响的个体中观察到突变，而在任何正常个体中没有观察到该突变，则它有可能是疾病的病因。

组织定位

本发明的核苷酸序列对于组织定位也是有价值的。这些技术能够通过检测编码 IGS5 多肽的 mRNA 来测定它们在组织中的表达模式。这些技术包括原位杂交技术和核苷酸扩增技术（例如 PCR）。这些技术在本领域是众所周知的。来自这些研究的结果可指示多肽在生物体中的正常功能。另外，IGS5 mRNA 正常表达模式与由 IGS5 基因编码的 mRNA 的表达模式之间的比较性研究，可以对疾病中突变体 IGS5 多肽的作用或正常 IGS5 多肽不适当表达的作用提供有价值洞察。这种不适当表达可能是时间、空间、或仅仅是量的差异。

本发明的多肽或其片段或其类似物或者表达它们的细胞还可用作免疫原以生成针对本发明多肽的免疫特异性抗体。术语“免疫特异性”指抗体对本发明多肽的亲和力显著高于它们对现有技术其它相关多肽的亲和力。

抗体

使用常规方案，通过对动物（优选非人动物）施用多肽或者携带表位的片段、类似物、或细胞，可获得针对本发明多肽而产生的抗体。对于单克隆抗体的制备，可使用能提供由连续细胞系培养物产生的抗体的任何技术。范例包括杂交瘤技术（G. Kohler和C. Milstein, Nature, 256: 495 - 597, 1975）、trioma技术、人B细胞杂交瘤技术（Kozbor等人, Immunology Today, 4: 72, 1983）、和EBV-杂交瘤技术（Cole等人, 《Monoclonal Antibodies and Cancer Therapy》（单克隆抗体和癌症治疗），第77 - 96页, Alan R. Liss公司, 1985）。也可采用用于生成单链抗体的技术（诸如美国专利号4,946,778中所述）来生成针对本发明多肽的单链抗体。同样，转基因小鼠或其它生物体（包括其它哺乳动物）可用于表达人源化抗体。

上述抗体可用于分离或鉴定表达多肽的克隆，或者用于通过亲和层析纯化多肽。

针对本发明多肽的抗体还可用于治疗上文所述疾病等。

融合蛋白

另一方面，本发明涉及经遗传工程改造的包含本发明多肽或其片段和各亚类免疫球蛋白（IgG、IgM、IgA、IgE）重链或轻链恒定区各种部分的可溶性融合蛋白。优选的免疫球蛋白是人类IgG，特别是IgG1重链的恒定部分，其中融合发生于铰链区。在具体的实施方案中，可以简单的通过掺入可以被血液凝结因子Xa切割的序列来除去Fc部分。另外，本发明涉及通过遗传工程来制备这些融合蛋白的方法，和这些融合蛋白在药物筛选、诊断、和治疗中的用途。本发明的另一个方面还涉及编码这些融合蛋白的多核苷酸。融合蛋白技术的范例可以在国际专利申请号WO 94/29458和WO 94/22914中找到。

疫苗

本发明的另一个方面涉及用于在哺乳动物中诱导免疫学应答的方法，包括对哺乳动物施用（例如通过接种）本发明的多肽，施用量足以生成抗体和/或T细胞免疫应答，从而保护所述动物免于上文提到的疾病。本发明的还有一个方面涉及用于在哺乳动物中诱导免疫学应答的方法，包括借助在体内指导多核苷酸表达并编码多肽的载体来投递本发明的多肽，从而诱导这种免疫学应答以生成抗体，保护所述动物免于疾病。

本发明的另一个方面涉及免疫学/疫苗制剂（组合物），在导入哺乳动物宿主后，将在该哺乳动物中诱导针对本发明多肽的免疫学应答，其中所述组合物包含本发明的多肽或多核苷酸。这些免疫学/疫苗制剂（组合物）可以是治疗性的免疫学/疫苗制剂，或者是预防性的免疫学/疫苗制剂。疫苗制剂还可以包含合适载体。由于多肽可能在胃中降解，因此优选胃肠外施用（例如皮下、肌肉内、静脉内、或皮内注射）。适用于胃肠外施用的制剂包括水性和非水性无菌注射液，可包含抗氧化剂、缓冲液、抑菌剂、和使得制剂与接受者血液等渗的溶质；以及水性和非水性无菌悬浮液，可包含悬浮剂或增稠剂。制剂可存在于单位剂量或多份剂量的容器中，例如，密封的安瓿瓶和药水瓶，而且可以冻干状态储存，只需要在临使用前加入无菌液体载体。疫苗制剂还可包含用于增强制剂免疫原性的佐剂系统，诸如水包油系统和本领域知道的其它系统。剂量将取决于疫苗的比活性，而且易于通过常规实验测定。

筛选实验

本发明的多肽负责一种或多种生物学功能，包括一种或多种疾病状况，特别是上文所述疾病。因此，渴望找到用于鉴定可刺激或抑制所述多肽功能的化合物的筛选方法。因此，另一方面，本发明提供了用于鉴定那些可刺激或抑制所述多肽功能的化合物的筛选方法。可以由多种来源鉴定化合物，例如细胞、无细胞制剂、化学文库、和天然产物混合物。如此鉴定的这些刺激物或抑制物可以是天然的或经修饰

的底物、配体、受体、酶、等（在多肽的情况下）；或者是结构或功能模拟物（参阅Coligan等人，Current Protocols in Immunology, 1(2)：第5章，1991）。

筛选方法可以是只测量候选化合物对多肽活性或者对携带多肽的细胞或膜的影响。或者，筛选方法可包括竞争剂的竞争。另外，这些筛选方法可使用适用于多肽或者携带多肽的细胞或膜的活性的检测系统来测试候选化合物是否导致由多肽的激活或抑制生成的信号。通常在存在已知底物的情况中测定对多肽活性的抑制，并通过活性改变来观察候选化合物的效果，例如通过测试候选化合物是否导致对多肽的抑制或刺激。例如，筛选方法可只包括下列步骤：混合候选化合物与含本发明多肽的溶液和合适底物以形成混合物，测量混合物的IGS5活性，并比较混合物的IGS5活性与不含候选化合物的标准。

本发明的多核苷酸、多肽、和针对多肽的抗体还可用于构成用于检测加入的化合物对细胞中mRNA和多肽生成的影响的筛选方法。例如，可以建立ELISA测定法，通过本领域知道的标准方法，使用单克隆和多克隆抗体来测量多肽的分泌水平或细胞结合水平。这可用于发现可抑制或增强适当操作的细胞或组织中的多肽生成的试剂。

潜在多肽抑制物的范例包括抗体或（在有些情况中）与多肽的配体、底物、受体、酶、等密切相关的寡核苷酸或蛋白质，可以是例如如配体、底物、受体、酶、等的片段；或结合本发明多肽但不引发应答从而可预防多肽活性的小分子。

因而，另一方面，本发明涉及用于鉴定特别是抑制物、刺激物、配体、受体、底物、酶、等；本发明的多肽；或者降低或增强这些多肽生成的筛选试剂盒，其包含：

- (a) 本发明的多肽；
- (b) 表达本发明多肽的重组细胞；
- (c) 表达本发明多肽的细胞膜；或

(d) 针对本发明多肽的抗体，所述多肽优选是SEQ ID NO: 2、SEQ ID NO: 4、或SEQ ID NO: 6之一的多肽。

应当领会，在任何这些试剂盒中，成份（a）、（b）、（c）、或（d）可构成所述试剂盒的实质性部分。

熟练技术人员易于领会的是，本发明的多肽也可用于基于结构来设计多肽的刺激物或抑制物的方法中，包括：

- （a）在第一种情况中测定多肽的三维结构；
- （b）由三维结构推测刺激物或抑制物可能的反应或结合位点；
- （c）合成预测可与推测的结合或反应位点结合或反应的候选化合物；并
- （d）测试候选化合物是不是真正的刺激物或抑制物。

还应当领会，这通常是重复的过程。

预防和治疗方法

另一方面，本发明提供了用于治疗涉及IGS5多肽活性表达过度或不足的异常状况的方法，诸如上文（在本发明的多肽内容中）一般称为“疾病”的那些待治疗机能障碍、紊乱、或疾病。

若该多肽的活性过度，则可使用几种方法。一种方法是任选与制药学可接受载体一起，以有效抑制该多肽功能（诸如通过阻断结合底物、酶、等）的量，对有此需要的对象施用上文所述的抑制物化合物，由此减轻异常状况。在另一种方法中，可以施用该多肽仍能与内源性多肽竞争结合底物、酶等的可溶性形式。

还有一种方法是使用表达阻断技术来抑制编码内源IGS5多肽的基因的表达。已知这种技术包括反义序列的使用，其或是内部生成或是分开施用的（参阅例如O'Connor, J Neurochem, 56: 560, 1991; 《Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression》（寡脱氧核苷酸作为基因表达的反义抑制剂），CRC出版社，Boca Raton, 佛罗里达，美国，1988）。或者，可使用与基因形成三股螺旋（三链体）的寡核苷酸（参阅例如Lee等人, Nucleic Acids Res, 6: 3073, 1979; Cooney等人, Science, 241: 456, 1988; Dervan等人, Science, 251: 1360, 1991）。可施用这些寡聚物本身，或者

可在体内表达相关寡聚物。合成的反义或三链体寡核苷酸可包含经修饰碱基或经修饰主链。后者的范例包括甲基膦酸酯、硫代磷酸酯、或肽核酸主链。为了免受核酸酶的降解作用而提供保护，可以将这些主链掺入反义或三链体寡核苷酸，这在本领域是众所周知的。用这些或其它经修饰主链合成的反义和三链体分子也构成本发明的一个部分。

另外，可使用对IGS5 mRNA序列特异的核酶来预防IGS5多肽的表达。核酶是具有催化活性的天然或合成的RNA（参阅例如N. Usman等人，*Curr. Opin. Struct. Biol.*, 6(4): 527-533, 1996）。合成核酶可设计成在选定位置特异切割IGS5 mRNA，从而预防IGS5 mRNA翻译成多肽。可用天然磷酸核糖主链和天然碱基合成核酶，正如RNA分子中常见的。或者，为了免受核糖核酸酶的降解作用而提供保护，可用非天然主链（例如2'-O-甲基RNA）合成核酶，而且可包含经修饰碱基。

对于涉及IGS5表达和活性不足的异常状况的治疗，也可使用几种方法。一种方法是与制药学可接受载体一起，对对象施用治疗有效量的可刺激本发明多肽的化合物，由此减轻异常状况。或者，可采用基因疗法来实现对象体内相关细胞内源生成IGS5。例如，如上所述，可改造本发明的多核苷酸，从而在复制缺陷型逆转录病毒载体中进行表达。然后可分离逆转录病毒表达构建物，并导入用包含编码本发明多肽的RNA的逆转录质粒载体转导的包装细胞，使得包装细胞产生包含目的基因的感染性病毒颗粒。可对对象施用这些生产者细胞，从而在体内改造细胞和在体内表达多肽。基因疗法的概述参阅《*Human Molecular Genetics*》（人类分子遗传学），第20章“基因疗法和其它基于分子遗传学的治疗方法”（及其引用的参考文献），T. Strachan和A. P. Read, BIOS Scientific Publishers有限公司，1996。另一种方法是与合适制药学载体联合施用治疗量的本发明多肽。

配方和施用

另一方面，本发明提供了包含治疗有效量的多肽（诸如本发明多

肽的可溶性形式)、刺激或抑制肽、或小分子化合物与制药学可接受载体或赋形剂联合而成的制药学组合物。合适载体包括但不限于盐水、缓冲盐水、右旋糖、水、甘油、乙醇、及其组合。本发明还涉及制药学包装和试剂盒,其中包含装有一种或多种本发明上述组合物成份的一个或多个容器。可以单独给与本发明的多肽和其它化合物,或者联合其它化合物(诸如治疗性化合物)。

组合物应适合于施用路径,例如全身或口服路径。全身施用的优选路径包括注射,通常是静脉内注射。也可使用其它注射路径,诸如皮下、肌肉内、或腹膜内。全身施用的其它方法包括使用渗透剂(诸如胆盐或梭链孢酸或其它去污剂)的穿粘膜和穿皮施用。另外,若本发明的多肽或其它化合物能够配制成肠配方或胶囊配方,则口服施用也是可能的。这些化合物的施用也可以是局部的和/或定位的,以药膏、糊剂、凝胶等形式进行。

需要的剂量范围取决于选择的本发明肽或其它化合物、施用的路径、配方的性质、对象状况的本质、和主治医生的判断。然而,合适的剂量范围是 $0.1 - 100 \mu\text{g}/\text{kg}$ 对象体重。然而,考虑到可利用化合物的多样性和不同施用路径的效率不同,预计所需剂量的变化很大。例如,预计口服施用要求的剂量比静脉内注射要大。正如本领域完全领会的,可使用标准经验性路径来调整这些剂量水平的变化以进行优化。

在治疗中使用的多肽也可在对象体内内源生成,这就是上文所述的常常称为“基因疗法”的治疗形式。因而,例如可使用多核苷酸(诸如DNA或RNA),例如通过使用逆转录病毒质粒载体,改造来自对象的细胞,从而离体编码多肽。然后,将细胞导入对象体内。

多核苷酸和多肽序列构成了有价值的信息来源,借此有可能鉴定具有相似同源性的其它序列。最容易的便于该目的的方法是:将序列保存于计算机可读介质中,然后使用众所周知的搜索工具(诸如GCC和Lasergene软件包中的搜索工具)用保存的数据来搜索序列数据库。因此,另一方面,本发明提供了计算机可读介质,其中贮存了包含SEQ ID NO: 1、SEQ ID NO: 3、或SEQ ID NO: 5的序列的多核苷酸和/或

由其编码的多肽序列。

将本说明书中引用的所有发表物(包括但不限于专利和专利申请)收入本文作为参考,就像特异且个别的指定将个别发表物完整收入本文作为参考。

实施例

下列实施例只是为了进一步更详细的例示本发明,因此不应当认为这些实施例以任何方式限制本发明的范围。

实施例1: 编码NEP/ECE金属蛋白酶家族新成员的cDNA的克隆

实施例1a: NEP/ECE金属蛋白酶家族新成员的cDNA片段的同源性PCR克隆

在表达序列标签(EST)的DNA数据库中,检测到4个交叠的EST序列(编号AA524283、AI088893、AI217369、和AI380811),它们所包含的小型开放读码框所编码的蛋白质片段显示与中性内肽酶24.11/内皮素转化酶(NEP/ECE)金属蛋白酶蛋白质家族成员C端部分具有相似性(A. J. Turner等人, *Faseb J*, 11: 355 - 364, 1997)。这些EST中的NEP/ECE样小型开放读码框由终止密码子终止(在AA524283的情况下),而且在所有4个EST中前面都有包含所有3种读码框形式的终止密码子的序列。这种居先序列似乎与NEP/ECE金属蛋白酶家族成员完全无关。尽管小型开放读码框的极性与衍生这些EST的mRNA的5'→3'取向相反,仍然决定将这些序列用作RT-PCR同源性克隆方法的基础。并行的是,在公开的结构域数据库中观察到了显示与上述4种EST相同结构的其它EST序列,例如编号AI825876、AI888306、AI422224、AI422225、AI469281、AA975272、AA494534、AW006103、AI827701、AI650385、AI827898、AI934499、和AA422157。使用在EST簇(显示与NEP/ECE家族具有相似性的区域内)上设计的反向引物(IP11689; SEQ ID NO: 7)

和集中在NEP/ECE家族保守肽基元[VNA(F, Y)Y]上的简并正向引物 (IP11685; SEQ ID NO: 8) 进行RT-PCR反应。

为了合成cDNA, 混合2 μ g 人肺总RNA (Clontech #64023-1)、1 μ l oligo(dT)₁₂₋₁₈ (500 μ g/ml)、和9 μ l H₂O (终体积12 μ l), 加热至70 $^{\circ}$ C达10分钟, 然后在冰上冷却。加入4 μ l 5x 第一条链缓冲液 (250mM Tris-HCl pH8.3、375mM KCl、15mM MgCl₂)、2 μ l 0.1M DTT、1 μ l 10mM dNTP混合液、和1 μ l 200U SuperscriptTMII (Life Technologies) 逆转录酶。将混合物于42 $^{\circ}$ C保温50分钟, 并通过于70 $^{\circ}$ C加热15分钟来灭活反应。

在50 μ l终体积中进行PCR反应, 其中包含1 μ l cDNA合成反应物、5 μ l GeneAmpTM 10x PCR缓冲液 (500mM KCl、100mM Tris pH8.3、15mM MgCl₂、0.01% (w/v) 明胶; Perkin Elmer)、2 μ l 10mM dNTP混合液、每种各10pmol的正向和反向引物、和5U AmpliTaqTM聚合酶 (Perkin Elmer)。在初次变性 (95 $^{\circ}$ C 5min) 后, 将PCR反应进行40个循环: 变性 (94 $^{\circ}$ C 1min)、退火 (60 $^{\circ}$ C 1min)、和延伸 (72 $^{\circ}$ C, 1min)。通过琼脂糖凝胶电泳分析PCR反应产物。IP11685/IP11689 RT-PCR反应生成了大约600碱基对 (bp) 的扩增子。使用Qiaex-IITM纯化试剂盒 (Qiagen) 由凝胶纯化片段, 并使用pGEM-T Easy系统 (Promega) 依照供应商推荐的流程将片段连接到pGEM-T Easy质粒中。然后将重组质粒用于转化感受态大肠杆菌SURETM2细菌 (Stratagene)。将经转化细胞涂布在含100 μ g/l 氨苄青霉素、0.5mM IPTG、和50 μ g/ml X-gal的LB琼脂平板上。使用BioRobotTM9600核酸纯化系统 (Qiagen) 由单个菌落的微量培养物纯化质粒DNA。使用ABI PrismTM BigDyeTM终止物循环测序简易反应试剂盒 (PE-ABI), 使用插入片段的侧翼或内部 (IGS5特异) 引物, 对经纯化质粒DNA进行DNA测序反应。通过EtOH/NaOAc沉淀纯化循环测序反应产物, 并在ABI 373型自动化测序仪上进行分析。重组克隆YCE14、YCE15、和YCE16 (衍生自IP11685/IP11689扩增子) 所含插入片段的DNA序列在N端方向上延伸了最初EST簇的开放读码框, 并进一步确认了这个开放读码框来源于NEP/ECE金属蛋白酶蛋白质家

族新成员（见图1）。因而，这种上游序列完全偏离EST序列中存在的上游序列。这种新序列在本发明的内容中一般称为“IGS5”。

实施例1b：包含IGS5推定胞外结构域（ectodomain）的cDNA的克隆

为了获得其它IGS5 cDNA序列，使用对IGS5特异的反向引物IP12190（SEQ ID NO: 9）和集中在NEP/ECE家族保守肽基元[LXXLXWMD]上的简并正向引物IP12433（SEQ ID NO: 10）在上文所述条件下对人肺RNA进行另一轮RT-PCR反应。IP12190/IP12433 RT-PCR反应生成了大约600bp的扩增子，将其克隆到pGEM-T Easy载体中而生成克隆YCE19、YCE22、和YCE23。将所有克隆完整测序，从而能够进一步向上游延伸IGS5开放读码框（见图1）。

为了获得覆盖IGS5转录本5'末端的cDNA克隆，使用与Marathon™ cDNA扩增试剂盒（Clontech K1802-1）一起提供的衔接头引物1（AP1: SEQ ID NO: 11）与IGS5特异引物IP12189（SEQ ID NO: 12）和IP12585（SEQ ID NO: 13）的组合，对人心脏Marathon-Ready™ cDNA进行半嵌套5'-RACE PCR。依照由Clontech提供的Marathon-Ready™ cDNA用户手册的指示进行PCR RACE反应。将RACE产物在琼脂糖凝胶上分开，用溴化乙锭染色，并印迹到Hybond-N⁺膜上。将印迹在修改后的Church缓冲液（0.5M 磷酸盐、7% SDS、10mM EDTA）中于65℃预杂交2小时，然后在含 2×10^6 cpm/ml ³²P标记的克隆YCE23插入片段的相同缓冲液中于65℃杂交过夜。使用Prime-It™ MII试剂盒（Stratagene），依照供应商提供的指示，通过随机引发掺入[α -³²P]dCTP而将YCE23插入片段放射性标记至比活 $>10^9$ cpm/ μ g。在高严谨度（2x SSC/0.1% SDS，室温，2x30min；随后0.1x SSC/0.1% SDS，65℃，2x40min）清洗杂交印迹，并放射自显影过夜。由凝胶纯化杂交片段，克隆到pGEM™-T Easy载体中（生成克隆YCE59、YCE64、和YCE65），并如上所述测序。

可以将所有分离克隆的DNA序列装配成单个毗连序列群（IGS5CONS；见图1），它将IGS5的开放读码框进一步向上游延伸，但是仍未遭遇翻译起始密码子ATG。根据EST AI380811设计了引物

IP11689, 它不包含经比对EST序列中存在的终止密码子之前的最后4个核苷酸。为了生成在终止密码子处终止的开放读码框, 在IGS5CONS的整体装配中包括经比对EST序列的最后(共有的)22个核苷酸。

同源性搜索显示, (部分)编码的蛋白质与中性内肽酶(NEP; 见实施例2)最相似。然而, IGS5CONS开放读码框的最初20个氨基酸显示与NEP不具有任何相似性。这有可能归因于它们衍生自内含子的事实。事实上, 人NEP外显子4的起始位置大致对应于这20个氨基酸的下游位置(L. D'Adamio等人, Proc. Natl. Acad. Sci. USA, 86: 7103-7107, 1989)。亲水性分析(J. Kyte等人, J. Mol. Biol., 157: 105-132, 1982; P. Klein等人, Biochim. Biophys. Acta, 815: 468-476, 1985)未指示在推测的IGS5CONS氨基酸序列中存在跨膜结构域, 而预计在最初20个氨基酸中应存在这种跨膜结构域。由于这些原因, 优选排除IGS5毗连序列群的最初序列部分(图1)。产生的DNA序列(IGS5DNA; SEQ ID NO: 1)长度为2076个核苷酸(包括终止密码子), 编码691个残基的蛋白质(IGS5PROT; SEQ ID NO: 2)。IGS5PROT与人NEP蛋白质序列的比对显示, IGS5PROT序列对应于NEP的完整胞外结构域序列。因而, 能够预计IGS5PROT具有推定IGS5酶的全部酶活性, 正如NEP胞外结构域所证明的(F. Fossiez等人, Biochem. J., 284: 53-59, 1992)。

表7: 实施例1中所用寡聚引物的概述

SEQ ID NO:7	IP11689:	5'-ACACGGCATCGTCCTTG-3'
SEQ ID NO:8	IP11685:	5'-CCCCCTGGACGGTGAA(C或T)GC(A,C,G或T)T(A或T)(C或T)TA-3'
SEQ ID NO:9	IP12190:	5'-AATCCGTTACGTTCTGTTCGTCTGCC-3'
SEQ ID NO:10	IP12433:	5'-CCTGGAGGAGCTG(A,C或G)(A,C或T)(A,C,G或T)TGGATG(A或G)A-3'
SEQ ID NO:11	AP1:	5'-CCATCCTAATACGACTCACTATAGGGC-3'
SEQ ID NO:12	IP12189:	5'-GTCCTTGCCACCCTCTGCCATCC-3'
SEQ ID NO:13	IP12585:	5'-ACCACCCCCGCCCGATGATCCAGAG-3'

实施例2: IGS5与NEP/ECE金属蛋白酶家族成员的蛋白质序列的比对
对于实施例1a中克隆的IGS5序列, 使用BLSAT算法(S. F. Altschul

等人, *Nucleic Acids Res.*, 25: 3389 - 3402, 1997) 执行对现有蛋白质数据库和翻译后DNA数据库的同源性搜索。这些搜索显示, 该IGS5蛋白质与小鼠、大鼠、和人中性内肽酶 (SW: NEP_MOUSE, 编号Q61391; SW: NEP_RAT, 编号P07861; 和SW: NEP_HUMAN, 编号P08473) 最相似 (在大约700个比对残基上具有54-55% 同一性)。

因而, 几乎完整的IGS5蛋白质序列与其它NEP/ECE家族成员的这一比对显示了IGS5与金属蛋白酶 (一般而言) 和与NEP和/或ECE金属蛋白酶家族 (具体而言) 的关系。由这种结构比对得出结论, IGS5具有金属蛋白酶的功能性, 因而它在动物和人的几种机能障碍、紊乱、或疾病方面受到关注。

实施例3: 编码NEP/ECE金属蛋白酶家族新成员的cDNA的克隆

实施例3a: NEP/ECE金属蛋白酶家族新成员的cDNA片段的同源性PCR克隆

在表达序列标签 (EST) 的DNA数据库中, 检测到4个交叠的EST序列 (编号AA524283、AI088893、AI217369、和AI380811), 它们所包含的小型开放读码框所编码的蛋白质片段显示与中性内肽酶24.11/内皮素转化酶 (NEP/ECE) 金属蛋白酶蛋白质家族成员C端部分有相似性 (A. J. Turner等人, *Faseb J*, 11: 355 - 364, 1997)。这些EST中的NEP/ECE样小型开放读码框由终止密码子终止 (在AA524283的情况中), 而且在所有4个EST中前面都有包含所有3种读码框形式的终止密码子的序列。这种居先序列似乎与NEP/ECE金属蛋白酶家族成员完全无关。尽管该小型开放读码框的极性与衍生这些EST的mRNA的5'→3'取向相反, 仍然决定将这些序列用作RT-PCR同源性克隆方法的基础。并行的是, 在公开的结构域数据库中观察到了显示与上述4种EST相同结构的其它EST序列, 例如编号AI825876、AI888306、AI422224、AI422225、AI469281、AA975272、AA494534、AW006103、AI827701、AI650385、AI827898、AI934499、和AA422157。使用在EST簇 (显示与NEP/ECE家

族有相似性的区域内)上设计的反向引物(IP11689; SEQ ID NO: 7)和集中在NEP/ECE家族保守肽基元[VNA(F,Y)Y]上的简并正向引物(IP11685; SEQ ID NO: 8)进行RT-PCR反应。

为了合成cDNA,混合2 μ g 人肺总RNA(Clontech #64023-1)、1 μ l oligo(dT)₁₂₋₁₈(500 μ g/ml)、和9 μ l H₂O(终体积12 μ l),加热至70 $^{\circ}$ C达10分钟,然后在冰上冷却。加入4 μ l 5x 第一条链缓冲液(250mM Tris-HCl pH8.3、375mM KCl、15mM MgCl₂)、2 μ l 0.1M DTT、1 μ l 10mM dNTP混合液、和1 μ l 200U SuperscriptTMII(Life Technologies)逆转录酶。将混合物于42 $^{\circ}$ C保温50分钟,并通过于70 $^{\circ}$ C加热15分钟来灭活反应。

在50 μ l终体积中进行PCR反应,其中包含1 μ l cDNA合成反应物、5 μ l GeneAmpTM 10x PCR缓冲液(500mM KCl、100mM Tris pH8.3、15mM MgCl₂、0.01% (w/v)明胶; PE Biosystems)、2 μ l 10mM dNTP混合液、每种各10pmol的正向和反向引物、和5U AmpliTaqTM聚合酶(PE Biosystems)。在初次变性(95 $^{\circ}$ C 5min)后,将PCR反应进行40个循环:变性(94 $^{\circ}$ C 1min)、退火(60 $^{\circ}$ C 1min)、和延伸(72 $^{\circ}$ C, 1min)。通过琼脂糖凝胶电泳分析PCR反应产物。IP11685/IP11689 RT-PCR反应生成了大约600碱基对(bp)的扩增子。使用Qiaex-IITM纯化试剂盒(Qiagen)由凝胶纯化片段,并使用pGEMTM-T Easy系统(Promega)依照供应商推荐的流程将片段连接到pGEMTM-T Easy质粒中。然后将重组质粒用于转化感受态大肠杆菌SURETM2细菌(Stratagene)。将经转化细胞涂布在含100 μ g/ml 氨苄青霉素、0.5mM IPTG、和50 μ g/ml X-gal的LB琼脂平板上。使用BioRobotTM9600核酸纯化系统(Qiagen)由单个菌落的微量培养物纯化质粒DNA。使用ABI PrismTM BigDyeTM终止物循环测序简易反应试剂盒(PE Biosystems),使用插入片段的侧翼或内部引物,对经纯化质粒DNA进行DNA测序反应。对质粒的两条链完整测序。通过EtOH/NaOAc沉淀纯化循环测序反应产物,并在ABI 373型自动化测序仪上进行分析。重组克隆YCE14、YCE15、和YCE16(衍生自IP11685/IP11689扩增子)所含插入片段的DNA序列在N端方向上延伸

了最初EST簇的开放读码框，并进一步支持这个开放读码框衍生自NEP/ECE金属蛋白酶蛋白质家族新成员的假设（见图2）。因而，这种上游序列完全偏离EST序列中存在的上游序列。这种新序列在本发明的内容中一般称为“IGS5”。

实施例3b：包含IGS5全长编码序列的cDNA片段的克隆

为了获得其它IGS5 cDNA序列，使用对IGS5特异的反向引物IP12190（SEQ ID NO: 9）和集中在NEP/ECE家族保守肽基元[LXXLXWMD]上的简并正向引物IP12433（SEQ ID NO: 10）在上文所述条件下对人肺RNA进行另一轮RT-PCR反应。IP12190/IP12433 RT-PCR反应生成了大约600bp的扩增子，将其克隆到pGEM™-T Easy载体中而生成克隆YCE19、YCE22、和YCE23。将所有克隆完整测序，从而能够进一步向上游延伸IGS5开放读码框（见图2）。

为了获得覆盖IGS5转录本5'末端的cDNA克隆，使用与Marathon™cDNA扩增试剂盒（Clontech K1802-1）一起提供的衔接头引物1（AP1: SEQ ID NO: 11）与IGS5特异引物IP12189（SEQ ID NO: 12）和IP12585（SEQ ID NO: 13）的组合对人心脏Marathon-Ready™cDNA进行半嵌套5'-RACE PCR。依照由Clontech提供的Marathon-Ready™cDNA用户手册的指示进行PCR RACE反应。将RACE产物在琼脂糖凝胶上分开，用溴化乙锭染色，并转移到Hybond™-N⁺膜（Amersham）上。将印迹在修改后的Church缓冲液（0.5M 磷酸盐、7% SDS、10mM EDTA）中于65℃预杂交2小时，然后在含 2×10^6 cpm/ml ³²P标记的克隆YCE23插入片段的相同缓冲液中于65℃杂交过夜。使用Prime-It™II试剂盒（Stratagene），依照供应商提供的指示，通过随机引发掺入[α -³²P]dCTP将YCE23插入片段放射性标记至比活 $>10^9$ cpm/ μ g。在高严谨度（2x SSC/0.1% SDS，室温，2x30min；随后0.1x SSC/0.1% SDS，65℃，2x40min）清洗杂交印迹，并放射自显影过夜。由凝胶纯化杂交片段，克隆到pGEM™-T Easy载体中（生成克隆YCE59、YCE64、和YCE65），并如上所述测序。

可以将所有分离克隆的DNA序列装配成单个毗连序列群，它将IGS5的开放读码框进一步向上游延伸，然而仍未遭遇翻译起始密码子。根据EST AI380811设计了引物IP11689，其中未掺入经比对EST序列中存在的终止密码子之前的最后4个核苷酸。为了生成在终止密码子处终止的开放读码框，在毗连序列群中包括经比对EST序列的最后（共有的）22个核苷酸。

通过5'RACE PCR延伸或通过筛选cDNA文库来克隆IGS5编码序列中仍然缺失的氨基末端部分的几次尝试都失败了。因此，现试图获得初步IGS5毗连序列群5'末端附近和上游的基因组序列信息。将在 λ EMBL3噬菌体载体中构建的人基因组DNA文库（Clontech HL1067j）的大约550,000个噬斑复制到HybondTM-N⁺膜上。将膜复制物在修改后的Church缓冲液中于65℃预杂交2小时，然后在含 2×10^6 cpm/ml ³²P标记的位于克隆YCE59 5'末端的大约150bp *EcoRI/EcoRII*片段的相同缓冲液中于65℃杂交过夜。使用Prime-ItTMII试剂盒（Stratagene），依照供应商提供的指示，通过随机引发掺入 $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ 将cDNA探针放射性标记至比活 $>10^9$ cpm/ μg 。在高严谨度（ $2 \times \text{SSC}/0.1\% \text{ SDS}$ ，室温， $2 \times 30\text{min}$ ；随后 $0.1 \times \text{SSC}/0.1\% \text{ SDS}$ ，65℃， $1 \times 40\text{min}$ ）清洗杂交印迹，并放射自显影。将杂交噬斑进行第二轮筛选，并获得纯的单个噬斑。使用QiagenTM λ Midi试剂盒（Qiagen）由受感染液体培养物纯化重组噬菌体DNA，并使用侧翼EMBL3载体引物和IGS5内部引物如上所述进行测序。由克隆IGS5/S1的插入片段测得了初步IGS5毗连序列群上游大约5000个核苷酸的序列。对翻译后DNA数据库的同源性搜索显示，这5000bp片段所含一个78bp片段编码的肽与小鼠SEP序列（GenBank编号AF157105）中其它方式剪接的69bp片段最相似（在25个经比对残基上15个相同），而后者是最近描述的NEP/ECE家族新成员（Ikeda等人，JBC, 274: 32469-32477, 1999）。这一78bp人片段的前后分别是推定的共有剪接受体和供体位点，但是不含翻译起始密码子ATG。

为了获得包含IGS5编码序列氨基末端部分的cDNA克隆，使用与MarathonTM cDNA扩增试剂盒（Clontech K1802-1）一起提供的衔接头

引物1 (AP1: SEQ ID NO: 11) 与在上文所述78bp基因组片段内设计的IGS5特异性反义引物IP14241 (SEQ ID NO: 14) 和IP14242 (SEQ ID NO: 15) 的组合对人睾丸Marathon-Ready™ cDNA (Clontech 7414-1) 进行半嵌套5'-RACE PCR。依照由Clontech提供的Marathon-Ready™ cDNA用户手册的指示进行PCR RACE反应 (反应体积 = 25 μl)。将RACE产物在琼脂糖凝胶上分开, 用溴化乙锭染色, 并通过Southern印迹分析。

为了生成用于印迹RACE产物的特异性杂交探针, 使用反向寡核苷酸引物IP14241 (SEQ ID NO: 14) 和集中在小鼠SEP蛋白质跨膜结构域中的肽基元[GLMVLLLL]上的简并正向引物IP13798 (SEQ ID NO: 16) 对上文获得的嵌套RACE产物进行半同源PCR反应。在25 μl终体积中进行PCR反应, 其中包含1 μl 半嵌套5'RACE PCR反应产物、2.5 μl GeneAmp™ 10x PCR缓冲液 (500mM KCl、100mM Tris pH8.3、15mM MgCl₂、0.01% (w/v) 明胶; PE Systems)、1 μl 10mM dNTP混合液、每种各10pmol的正向和反向引物、和2.5U AmpliTaq-Gold™ 聚合酶 (PE Biosystems)。在初次变性 (95°C 10min) 后, 将PCR反应进行35个循环: 变性 (95°C 1min)、退火 (50°C 30sec)、和延伸 (72°C, 30sec)。通过琼脂糖凝胶电泳分析PCR反应产物。半同源性PCR反应生成了大约110个碱基对的扩增子。使用Qiaex-II™ 纯化试剂盒 (Qiagen) 由凝胶纯化片段, 并使用pGEM™-T系统 (Promega) 依照供应商推荐的流程将片段连接到pGEM™-T质粒中。然后将重组质粒用于转化感受态大肠杆菌SURE™2细菌 (Stratagene)。将经转化细胞涂布在含100 μg/ml 氨苄青霉素、0.5mM IPTG、和50 μg/ml X-gal的LB琼脂平板上。使用BioRobot™9600核酸纯化系统 (Qiagen) 由单个菌落的微量培养物纯化质粒DNA, 并如上所述测序。重组克隆YCE207、YCE212、YCE216、YCE217、YCE218、和YCE219所含插入片段的DNA序列能够与上文所述78bp基因组片段一起装配成单个毗邻序列群 (见图2)。

将半嵌套5'RACE PCR反应产物的Southern印迹在修改后的Church缓冲液中于65°C预杂交1小时, 然后在含2x10⁶cpm/ml ³²P标记的克隆YCE207插入片段的相同缓冲液中于65°C杂交过夜。在高严谨度清洗杂

交印迹，并放射自显影。由凝胶纯化杂交片段，克隆到pGEMTM-T载体中（生成克隆YCE223、YCE224、和YCE226），并如上所述测序。这些克隆的DNA序列能够与78bp基因组片段以及克隆YCE207、YCE212、YCE216、YCE217、YCE218、和YCE219装配成单个毗邻序列群（图2）。产生的毗邻序列群包含一个起始于ATG起始密码子的开放读码框，编码的蛋白质显示与小鼠SEP蛋白质N端序列有高度相似性。

为了获得覆盖IGS5编码序列的氨基末端部分且与克隆YCE59交叠的cDNA克隆，使用基于IGS5 5'UTR序列的特异正向引物（IP14535；SEQ ID NO: 17）与位于YCE59内的特异反向引物（IP14537；SEQ ID NO: 18），对人睾丸Marathon-ReadyTMcDNA（Clontech 7414-1）进行PCR反应。在25 μ l体积中进行PCR反应，其中包含2.5 μ l人睾丸Marathon-ReadyTMcDNA、2.5 μ l GeneAmpTM 10x PCR缓冲液（500mM KCl、100mM Tris pH8.3、15mM MgCl₂、0.01% (w/v) 明胶；PE Biosystems）、1 μ l 10mM dNTP混合液、每种各10pmol的正向和反向引物、和2.5U AmpliTaq-GoldTM聚合酶（PE Biosystems）。在初次变性（95 $^{\circ}$ C 10min）后，将PCR反应进行41个循环：变性（95 $^{\circ}$ C 1min）、退火（53 $^{\circ}$ C 1min）、和延伸（72 $^{\circ}$ C，1min）。通过琼脂糖凝胶电泳分析PCR反应产物。PCR反应生成了分别为大约300和380碱基对的两扩增子。由凝胶纯化该300bp和380bp片段，克隆到pGEMTM-T载体中，并如上所述测序。这生成了克隆YCE231、YCE233、和YCE235（300bp片段）以及YCE229（380bp片段）。

将所有分离克隆的DNA序列装配起来显示存在因最初在基因组克隆IGS5/S1中鉴定的78bp片段的有无而不同的两种cDNA序列。这两种序列有可能起源于不同剪接的RNA分子。最长的转录本包含2337个核苷酸（编码779个残基的蛋白质）的开放读码框，而较短的转录本包含2259个核苷酸（编码753个残基的蛋白质）的开放读码框。我们将长型的编码序列和蛋白质序列分别称为IGS5DNA1（SEQ ID NO: 3, 2340bp, 包括终止密码子标签）和IGS5PROT1（SEQ ID NO: 4），而将短型的编码序列和蛋白质序列分别称为IGS5DNA2（SEQ ID NO: 5, 2262bp, 包

括终止密码子标签)和IGS5PROT2 (SEQ ID NO: 6)。在IGS5DNA1和IGS5DNA2中假定甲硫氨酸起始密码子的下游,在第10位密码子处存在另一个符合读码框的甲硫氨酸密码子。尽管我们已经采用了第一个甲硫氨酸密码子作为起始密码子,由于两个甲硫氨酸似乎处于同等有利的翻译“Kozak”起始范围内,因而不能排除翻译有时(或甚至完全的)由第10位密码子起始(M. Kozak, Gene, 234: 187-208, 1999)。IGS5PROT1和IGS5PROT2序列的亲水性分析(J. Kyte等人, J. Mol. Biol., 157: 105-132, 1982; P. Klein等人, Biochim. Biophys. Acta, 815: 468-476, 1985)显示,在第22-50位残基间存在单个跨膜结构域。这指示IGS5PROT1和IGS5PROT2是II型整合膜蛋白,因而具有与NEP/ECE蛋白质家族其它成员相似的膜拓扑学。

表8: 实施例3中所用寡核苷酸引物的概述

SEQ ID NO:7	IP11689:	5'-ACACGGCATCGCTCCTTG-3'
SEQ ID NO:8	IP11685:	5'-CCCCCTGGACGGTGAA(C或T)GC(A,C,G或T)T(A或T)(C或T)TA-3'
SEQ ID NO:9	IP12190:	5'-AATCCGTTACGTTCTGTTCGTCTGCC-3'
SEQ ID NO:10	IP12433:	5'-CCTGGAGGAGCTG(A,C或G)(A,C或T)(A,C,G或T)TGGATG(A或G)A-3'
SEQ ID NO:11	AP1:	5'-CCATCCTAATACGACTCACTATAGGGC-3'
SEQ ID NO:12	IP12189:	5'-GTCCTTGCCACCCTCTGCCATCC-3'
SEQ ID NO:13	IP12585:	5'-ACCACCCCCGCCCCGATGATCCAGAG-3'
SEQ ID NO:14	IP14241:	5'-ACAGCCGGCTAGCAAGGCGTGGCAGCTG-3'
SEQ ID NO:15	IP14242:	5'-ACGACAGCCGGCTAGCAAGGCGTGGCAG-3'
SEQ ID NO:16	IP13798:	5'-GG(A,C,G或T)CT(C或G)ATGGT(A,C,G或T)CT(C或G)CT(C或G)CT(C或G)CT(C或G)-3'
SEQ ID NO:17	IP14535:	5'-CTCCTGAGTGAGCAAAGGTTCC-3'
SEQ ID NO:18	IP14537:	5'-GCAAAGTGGTAGAAGTCGTACAC-3'

实施例4: IGS5与NEP/ECE金属蛋白酶家族成员的蛋白质序列的比对

对于实施例3中克隆的IGS5序列,使用BLSAT算法(S. F. Altschul等人, Nucleic Acids Res., 25: 3389-3402, 1997)执行对现有蛋

白质数据库和翻译后DNA数据库的同源性搜索。这些搜索显示，IGS5PROT1与小鼠SEP（GenBank编号AF157105）最相似（在778个比对残基上具有76%同一性），还在696个比对残基上显示与小鼠、大鼠、和人中性内肽酶（SW: NEP_MOUSE, 编号Q61391; SW: NEP_RAT, 编号P07861; 和SW: NEP_HUMAN, 编号P08473）有54-55%同一性。IGS5PROT2的同源性搜索显示，这种序列与小鼠SEP（GenBank编号AF157106）最相似（在752个比对残基上具有78%同一性）。在与小鼠SEP和SEP蛋白质的类比中，预计IGS5PROT1和IGS5PROT2分别代表IGS5蛋白质的膜结合形式和可溶性形式。这通过可变剪接的78bp外显子3'端编码的二碱性残基（KRK）的存在得到了确证。

因而，完整IGS5蛋白质序列与其它NEP/ECE家族成员的比对显示了IGS5与NEP/ECE金属蛋白酶（一般而言）和与SEP和NEP家族成员（具体而言）的关系。由这种结构比对得出结论，IGS5蛋白具有金属蛋白酶的功能性，因而在动物和人的几种机能障碍、紊乱、或疾病方面受到关注。

实施例5: IGS5的RNA表达分析

对人RNA Master印迹™的IGS5表达分析

如下制备Express-Hyb™（Clontech #8015-1）和剪碎鲑鱼睾丸DNA的溶液：于50-60℃预热15ml Express-Hyb™。将1.5mg剪碎鲑鱼睾丸DNA于95℃加热5分钟，然后在冰上迅速冷却。将热变性的剪碎鲑鱼睾丸DNA与预热的Express-Hyb™混合。将人RNA Master印迹™（Clontech #7770-1）在上文制备的10ml溶液于65℃预杂交30分钟，连续搅动。将使用Prime-It™II试剂盒（Stratagene）³²P标记的YCE15探针加热变性，并加到剩余的5ml Express-Hyb™溶液中。于65℃杂交过夜。在2x SSC/1% SDS中于65℃清洗100分钟（5x20min）。在200ml 0.1x SSC/0.5% SDS中于55℃再清洗2x20分钟。最后使用X射线胶片使Master印迹放射自显影。

IGS5探针在Master印迹™上的杂交显示在广泛组织中表达,特别是睾丸、小肠、前列腺、和胃中表达(图3)。

对人脑多种组织 Northern 印迹 II 和 IV (分别为#7755-1 和#7769-1) 的 IGS5 表达分析

于68℃预热Express-Hyb™溶液(Clontech #8015-1)。将印迹于68℃预杂交1小时。将100 μg剪碎鲑鱼睾丸DNA加到使用Prime-It™II试剂盒(Stratagene) ³²P标记的YCE15探针中,并于95℃加热变性10分钟。将探针加到剩余的5ml Express-Hyb™溶液中,并于68℃杂交2小时。在2x SSC/0.05% SDS中于室温清洗40分钟(2x20min)。在200ml 0.1x SSC/0.1% SDS中于55℃再清洗2x20分钟。使用X射线胶片使印迹放射自显影。

此次Northern印迹分析显示,在所研究的所有组织中存在大约3kb的主要杂交条带和大约5.5-6kb的次要条带。

实施例6: His标记的人IGS5胞外结构域的表达和纯化

本实验目的是使用杆状病毒表达系统生成可溶性IGS5蛋白质。构建在感染Sf9细胞系后表达His6标记的IGS5胞外结构域的重组杆状病毒。然后在包括扁豆凝集素和Zn-IMAC层析的两步流程中由培养物上清液纯化可溶性IGS5蛋白质。

我们将阿黑皮素原(POMC)前体的信号肽融合His标记的IGS5胞外结构域部分的编码序列。由于该蛋白质的酶活性位点(金属蛋白酶)位于C末端,我们优选将His纯化标签加到蛋白质N端。此外,将Gly-Ser接头插在POMC信号肽与IGS5胞外结构域之间。表达的IGS5蛋白质由IGS5PROT2的第60位残基起始(SEVC...),因而包含几乎整个IGS5胞外结构域。克隆策略包括合成寡核苷酸装配、交叠PCR、和3点连接。这导致由POMC信号(分泌时切除)、Gly-Ser接头、His6肽、和IGS5胞外结构域构成的蛋白质的表达。

实施例6a: pAcSG2SOLhuIGS5His6杆状病毒转移载体的构建

为了构建pAcSG2SOLhuIGS5杆状病毒转移载体,生成下列DNA片段:

1. 用StuI/NotI消化pAcSG2载体(BD PharMingen). 使用Qiaex-II™提取试剂盒(Westburg)由琼脂糖凝胶提取5527bp片段,并溶于30 μl 10mM Tris-HCl pH8.5.

2. 通过PCR与限制性消化/连接的联合由克隆YCE15、YCE22、YCE64、和YCE65装配pGEM™-T克隆YCE174. 将与IP11689相反而确实包含IGS5编码序列的最后4个核苷酸和终止密码子的引物IP13541用于该流程(表9). 因此YCE174包含huIGS5胞外结构域直至(且包括)终止密码子的几乎整个编码区(图2). 用XhoI/NotI消化YCE174,如琼脂糖凝胶电泳所示,产生3025bp、1723bp、和448bp片段. 由凝胶提取(Qiaex-II™, Qiagen)包含huIGS5胞外结构域编码区的1723bp片段,并溶于20 μl 10mM Tris-HCl pH8.5.

3. 通过联合寡核苷酸IP14165、IP14114、IP14115、IP14116、IP14117、IP14118、IP14119、和IP14120,及随后使用引物IP14166和IP14110的交叠PCR,装配5'端含StuI识别位点、随后是POMC信号序列、Gly-Ser接头、His6标签、和IGS5胞外结构域编码序列5'末端65bp的合成核酸片段(180bp)(表9;也可见图4). 通过在第57位核苷酸处导入沉默突变来除去天然POMC信号肽编码序列中存在的StuI位点(IP14115,第30位核苷酸G→A).

表9: 实施例6中所用寡核苷酸引物的概述

SEQ ID NO:19	IP14165:	5'-GACAAGGCCTATTATGCCGAGATCGTGCTGCAGCCGCTCG-3'
SEQ ID NO:20	IP14114:	5'-AAGGCCAGCAACAGGGCCCCGAGCGGCTGCAGCACGATC-3'
SEQ ID NO:21	IP14115:	5'-GGGGCCCTGTTGCTGGCCTTGCTGCTTCAAGCCTCCATGG-3'
SEQ ID NO:22	IP14116:	5'-GTGAGAACCGCCACGCACTTCCATGGAGGCTGAAGCAGC-3'
SEQ ID NO:23	IP14117:	5'-AAGTGCCTGGCGGTTCTCACCATCACCACCATCACAGCGA-3'
SEQ ID NO:24	IP14118:	5'-AGCCAGGGGTGGTGCAGACCTCGCTGTGATGGTGGTGATG-3'
SEQ ID NO:25	IP14119:	5'-GGTCTGCACCACCCCTGGCTGCGTGATAGCAGCTGCCAGG-3'
SEQ ID NO:26	IP14120:	5'-GGGTCCATGTTCTGGAGGATCCTGGCAGCTGCTATCACGC-3'
SEQ ID NO:27	IP14166:	5'-GACAAGGCCTATTATG-3'
SEQ ID NO:28	IP14110:	5'-GGGTCCATGTTCTG-3'
SEQ ID NO:29	IP14111:	5'-AGCGAGGTCTGCAC-3'
SEQ ID NO:30	IP14112:	5'-GTAGATGATGTGCCG-3'
SEQ ID NO: 31	IP13541:	5'-GCACTAGTCTTGGCTACCACACGCGCATCGTCTCTTG-3'

使用引物IP14111和IP14112由克隆YCE174模板扩增第二种PCR片段(495bp)。该第一种和第二种PCR产物共享长度为65bp的交叠区。使用两种PCR产物作为模板,使用引物IP14166和IP14112,进行交叠PCR,生成最终的610bp PCR片段。将该片段在QiaQuick旋转柱(Qiagen)上纯化,StuI/XhoI消化,并由琼脂糖凝胶提取(Qiaex-II™)产生的496bp片段。

在连接反应中将如上所述生成的3种DNA片段(5527、1723、和496bp)联合起来。将连接混合物于16℃保温过夜,并用于转化感受态DH5α F'细胞。将经转化细菌涂布在含100 μg/ml氨苄青霉素的LB琼脂平板上。将平板于37℃保温过夜。在添加100 μg/ml氨苄青霉素的5ml LB培养基中培养30个随机菌落。使用BioRobot™9600核酸纯化系统(Qiagen)制备质粒DNA,并通过BamHI消化进行筛选。对展示正确限制性模式的7个克隆,通过XhoI、StuI、AlwNI、HindIII、和HincII消化和插入片段序列分析进行进一步分析。最后选择一个具有正确限

制性模式和预期插入片段序列的克隆，并作为ICCG4502保藏于培养物收藏中心(菌株表)(该克隆在转移载体的第878位包含沉默突变G→A)。

由保藏的克隆进行无菌Qiagen Midi DNA制备(Westburg)，生成110 μg DNA。XhoI、AlwNI、StuI、HIndIII、和HincII消化的限制性分析揭示了正确的限制性模式，如琼脂糖凝胶电泳所示。序列分析确认了预期序列。图5显示了pAcSG2SOLhuIGS5His6杆状病毒转移载体的图谱。

实施例6b: 用于表达可溶性huIGS5His6的重组杆状病毒的生成

通过pAcSG2SOLhuIGS5His6转移载体DNA(ICCG4502)与野生型杆状病毒基因组修改版的线性化基因组DNA(BaculoGold; Pharmingen 目录编号21100D)在宿主昆虫细胞(*Spodoptera frugiperda* Sf9细胞)中共转染(磷酸钙转染法)，生成表达N端His标记的人IGS5胞外结构域(减去一些氨基酸)的重组杆状病毒。BaculoGold DNA包含致死删除，而且不编码可存活病毒。BaculoGold DNA与互补杆状病毒转移载体的共转染通过同源重组挽救了该致死删除。使用这种方法，噬斑纯化了3个候选重组体。扩增所有候选重组体。

实施例6c: 真核表达

动力学表达分析

通过低速离心收集在旋转烧瓶中在添加10%灭活胎牛血清(Gibco BRL 目录编号10084168)的TC100培养基(JRH Biosciences 目录编号56941)中于27℃指数悬浮生长的Sf9细胞(IGCL 83.0)，并以 5×10^5 细胞/瓶(25cm²)接种在无血清TC100培养基中。以感染复数(MOI) 3pfu/细胞加入候选重组病毒克隆，随后将细胞/病毒培养物于27℃保温。在感染后24、48、和72小时，通过两次连续低速离心收集条件培养基(CM)。通过SDS-PAGE凝胶电泳和Western印迹分析样品。

Western印迹揭示了所有候选克隆的CM中都有大约81kDa的清晰条

带，符合该成熟蛋白质的理论Mr (81.2kDa) (未显示)。所有3个克隆的表达水平在感染后48-72小时达到峰值。选择2号克隆进一步扩增，并作为IGBV73保藏。最佳收获时间设定为感染后72小时。

脱糖基研究

可溶性人IGS5蛋白质序列包含8个潜在N-糖基化位点(图6)。既然纯化方案包括在扁豆凝集素柱上结合糖基，那么将感染后72小时收集的所有候选重组病毒克隆的CM样品用于N-糖苷酶F的脱糖基化研究，以检查重组体可溶性His6IGS5蛋白质是否确实作为糖基化蛋白质表达。

向样品中加入SDS至终浓度1%，并于95℃保温5分钟。加入1倍体积的2x保温缓冲液(250mM 磷酸盐缓冲液、50mM EDTA、5% N-辛基糖苷、1% 2-巯基乙醇)并于95℃再次保温5分钟后，将样品冷却至37℃。加入1U N-糖苷酶F(Boehringer Mannheim 目录编号1365177)，并在于37℃保温过夜后用100mM DTT(终浓度)还原样品。

经N-糖苷酶F处理的CM样品和未处理对照的Western印迹分析显示，样品脱糖基化后Mr发生变化(图7)，证明可溶性人His标签IGS5是作为糖基化蛋白质表达的。

实施例6d: 纯化

制备性生产和样品处理

通过低速离心收集在旋转烧瓶中在添加10% 灭活胎牛血清(Gibco BRL 目录编号10084168)的TC100培养基(JRH Biosciences 目录编号56941)中于27℃指数悬浮生长的Sf9细胞(IGCL 83-2)，并以 2×10^6 细胞/ml的密度重悬于添加0.013TIU/ml 抑酶肽(Pentex)的TC100培养基中。以感染复数(MOI) 2.25pfu/细胞，将重组病毒IGBV73加到细胞中。随后将细胞/病毒悬浮物在玻璃滚瓶($3 \times 500\text{ml}/1260\text{cm}^2$)中于27℃保温72小时。然后通过两次连续低速离心将CM(1.5L)与细胞和细胞碎片分开。向300ml 澄清杆状病毒CM中加入1片无EDTA的完全EFC

(Roche Biochemicals 目录编号1873580)。分别加入HEPES、甘油、和吐温20至终浓度20mM、5% (v/v)、和0.005% (w/v)。将CM的pH调至7.4, 并将样品过滤 (Durapore膜滤器 0.2 μ GV)。所有纯化步骤都于4 $^{\circ}$ C进行。

扁豆凝集素层析

将杆状病毒样品以0.3ml/min上样至装在C10/10柱 (Pharmacia) 中用添加1片EFC /500ml的缓冲液A (20mM Hepes pH7.4、150mM NaCl、5% 甘油、0.005% 吐温20) 平衡的5ml扁豆凝集素Sephrose树脂上过夜。用平衡缓冲液清洗层析柱直至280nm吸光值达到基线水平, 并用含0.5M α -甲基吡喃糖苷的缓冲液A以1ml/min洗脱结合的蛋白质。通过应用100mM醋酸、500mM NaCl (pH5.0) 再生柱。手工收集洗脱液和再生液, 并通过12.5% Phast凝胶 (Pharmacia) 的SDS-PAGE和银染分析合并液。包括有预染标准 (Gibco) 作为相对分子量 (Mr) 标准。在流出物中回收蛋白质的主要量, 而且在1-3号洗脱合并液中观察到Mr大约85,000的IGS5候选条带 (图8)。抗His标签mAb对扁豆层析的Western印迹分析显示, 可溶性hIGS5蛋白质 (Mr大约85,000) 定量结合于扁豆凝集素树脂, 而且His标记的蛋白质整个洗脱峰中回收, 但是主要在1号和2号合并液中 (图9)。在Zn-IMAC柱上进一步加工1号和2号扁豆凝集素洗脱合并液 (运行A和B)。

固定化金属亲和层析 (IMAC) 和透析

如制造商所述将锌离子上样至1ml螯合HiTrap (Pharmacia) 上, 并用缓冲液B (20mM Hepes、100mM NaCl、5% 甘油、0.005% (w/v) 吐温20, pH7.2) 平衡。将1号和2号扁豆洗脱合并液以0.5ml/min分别上样至HiTrap柱中 (IMAC运行A和IMAC运行B)。包括空白运行以比较层析吸光值图谱。用缓冲液B清洗层析柱直至基线水平, 并用溶于缓冲液B的咪唑分级梯度 (20、50、100、和200mM) 洗脱结合的蛋白质。手工收集各级分。用20mM Hepes、50mM EDTA、500mM NaCl pH7.2再生IMAC柱。通过SDS-PAGE (12.5% Phast凝胶, Pharmacia) 和银染分析洗脱和再生合并液。将200mM咪唑合并液转移至slide alyzer盒 (MWC0

10,000, Pierce) 并在缓冲液B (130倍过量, 不更换缓冲液) 中透析过夜。使用micro-BCA法 (Pierce) 测定透析后合并液中可溶性IGS5的量。包括BSA作为参照物。通过(1)还原和非还原条件下的SDS-PAGE和(2)使用抗His标签的mAb (21E1B4, Innogenetics) 的Western印迹, 随后与碱性磷酸酶标记的兔抗小鼠Ig (Dako) 一起保温, 并通过NBT/BCIP染色进行检测, 由此在生物化学上表征透析后杆状病毒IGS5。通过PGN酶F处理 (Biorad) 确认了可溶性IGS5的糖基化状态。

SDS-PAGE分析和银染显示, 20mM和50mM咪唑梯度洗脱了大量杂质蛋白质(图10)。在100mM和200mM咪唑洗脱梯度中回收到hIGS5蛋白质。200mM咪唑合并液中的85kDa条带在SDS-PAGE上是单一条带, 它与抗His标签mAb发生反应(图9)。银染未揭示由IMAC运行A和运行B获得的hIGS5材料之间有任何纯度差异。由300ml杆状病毒CM开始, 通过两步纯化流程获得了340 μ g超过95%纯的含His标签hIGS5胞外结构域(即产量为大约1mg/L)。

SDS-PAGE后, 将200mM咪唑Zn-IMAC合并液转印迹至PVDF膜, 并通过酰胺黑染色使蛋白质显影。先后用20%乙腈和20%甲醇清洗PVDF条带, 并干燥。使用Procise™492A (Applied Biosystems) 依照制造商的说明书通过Edman降解进行氨基末端测序。氨基末端测序确认了在杆状病毒IGS5回收在85kDa蛋白质条带中。

实施例7: 酶抑制实验

针对生物学活性肽的代谢测试本发明IGS5多肽的酶活性。具体而言, 测试这些IGS5多肽是否作用于本领域现阶段知道的多种血管活性肽, 诸如心房钠尿肽 (ANP)、缓激肽、大内皮素 (大ET-1)、内皮素 (ET-1)、物质P、和血管紧张素-1。在本发明的内容中, 具体而言, 测试IGS5胞外结构域 (新型人金属蛋白酶) 是否水解上述血管活性肽。为了比较, 还对早先描述成可溶性分泌型内肽酶 (SEP) (Emoto等人, J. Biol. Chem., 274: 32469 - 32477, 1999) 的金属蛋白酶家族已知成员进行测定。此外, 测试了IGS5转化大ET-1类似物 (所谓的17个

氨基酸的大ET)的活性是否受到参考化合物的抑制,所述参考化合物用于测定对具有ECE和/或NEP特征的酶的抑制特性。用于测试IGS5对大ET-1类似物的活性的抑制作用的化合物是抑制内肽酶像NEP和ECE的化合物磷酸二肽、特异抑制NEP的化合物thiorphan、和作为选择性ECE抑制剂的化合物CGS-35066。

实施例7a: 材料

酶: IGS5(sol hu)(his6); 或 His6 标记的 IGS5 胞外结构域;
储存液: 53mg/ml, 溶于 20mM HEPES pH7.2、5% 甘油、
0.005% 吐温 20、100mM NaCl, 纯度>99%; 保存于 4℃。
工作液: 用测定缓冲液将储存液稀释至 10mg/ml。

底物: Mca-Asp-Ile-Ala-Trp-Phe-Dpa-Thr-Pro-Glu-His-
Val-Val-Pro-Tyr-Gly-Leu-Gly-COOH;
荧光淬灭的大 ET-1 类似物;
Mca = (7-甲氧基香豆素-4-基);
Dpa = (3-[2,4-二硝基苯基]-L-2,3-二氨基丙酰基);
储存液: 100 μM, 溶于测定缓冲液; 保存于-20℃。

可以由供应商 Polypeptide Laboratories (Wolfenbüttel, 德国) 处购买。

测定缓冲液: 100mM Tris pH7.0, 250mM NaCl。

将所有待测化合物以10mM溶于DMSO,并用测定缓冲液进一步稀释。

实施例7b: 实验流程

将70 μl 测定缓冲液、10 μl 酶工作液、和10 μl 待测化合物溶液在Eppendorf管中混合,并于37℃预先保温15分钟。然后,加入10 μl底物储存液,并将反应化合物于37℃保温60分钟,从而能够进行酶促水解。随后,通过于95℃加热5分钟来终止酶促反应。离心(Heraeus Biofuge B, 3min)后,对上清液进行HPLC分析。

实施例7c: HPLC流程

为了将剩余底物与切割产物分开,使用CC 125/4 Nucleosil 300/5 C18RP柱和CC 8/4 Nucleosil 100/5 C18预柱(可以由Macherey-Nagel (Düren, 德国)处购买)进行反相HPLC技术。因而,将60 μ l在实施例7b中获得的反应样品注射到HPLC中,并以1ml/min的流速用下列梯度和溶液洗脱层析柱:

溶液 A: 100% H₂O+0.5M H₃PO₄, pH2.0

溶液 B: 100% 乙腈+0.5M H₃PO₄

0-2min 20% B

2-6min 20-60% B

6-8min 60% B

8-10min 60-90% B

10-13min 90% B

13-15min 90-100% B

通过214nm的吸光值和通过激发波长328nm和发射波长393nm的荧光来检测肽。

实施例7d: 计算

将水解后含未淬灭Mca荧光团肽的HPLC峰的荧光信号增值作为任何计算的基础。

对含和不含抑制物的样品比较该信号,并根据各自的峰面积计算%抑制。

$$\% \text{抑制} = 100 \times (1 - A_{\text{抑制}} / A_{\text{对照}})$$

所有样品运行双份,并使用平均值。

在每次测定运行中加入标准抑制物(10nM和100nM磷酸二肽)和溶剂对照(0.1%)。

实施例7e: 结果

关于本发明的IGS5多肽,实施例7的结果显示这些IGS5金属蛋白酶

多肽在体外水解本领域现阶段已知的多种血管活性肽，诸如大ET-1、ET-1、ANP、和缓激肽。表10显示了与SEP活性相比的水解测定的结果。由这些结果得出结论，IGS5可能特别涉及所述生物学活性肽的代谢。

表10: 与SEP (可溶性分泌型内肽酶) 相比IGS5多肽对血管活性肽的水解		
血管活性肽	% 水解作用 IGS5多肽	% 水解作用 SEP (Emoto等人)
	条件: 100 μ g IGS5多肽; 0.5 μ M 底物; 2小时; 37 $^{\circ}$ C	条件: 10 μ g SEP; 0.5 μ M 底物; 12 小时; 37 $^{\circ}$ C
ANP	5 (80*)	>95
缓激肽	100 (62**)	>95
大ET-1	(?) ***	42
ET-1	30	92
物质P	n. d.	>95
血管紧张素-1	n. d.	>95
17个氨基酸的大ET	41	n. d.

* 500 μ g IGS5 多肽

** 10 μ g IGS5 多肽

*** 检测到活性，但是因 HPLC 检测的问题而不能量化

此外，参考化合物抑制ECE和/或NEP活性的实验结果显示，本发明IGS5金属蛋白酶多肽转化大ET-1类似物即17个氨基酸的大ET的活性受到磷酸二肽(抑制ECE/NEP的参考化合物)的抑制，但是IGS5未受到NEP抑制物thiorphan的有效抑制。表11显示了这些结果。IGS5多肽也不受选择性ECE抑制物CGS-35066的抑制，后者是内皮素转化酶-1的有效且选择性的非肽抑制物，而且具有持久的作用时间(De Lombart等人, J. Med. Chem., 43(3): 488-504, 2000年2月10日)。

表11: 对IGS5多肽转化大ET-1类似物17个氨基酸的大ET的活性的抑制	
抑制物化合物	IC ₅₀ nM
磷酸二肽	18
thiorphan	>1000
CGS-35066	1300

图的简述

图1: 分离并完整测序以生成部分IGS5共有cDNA序列的不同cDNA克隆相对位置的示意图。图中指示了用于5'RACE和半同源性PCR克隆的PCR引物, 而且在本文件中也有所描述(分别以IP#表示)。IGS5CONS指合并所有获得的序列后获得的共有毗邻序列群。假定包含IGS5酶(IGS5DNA, IGS5PROT)胞外结构域的IGS5毗邻序列群中存在的长度为691个氨基酸的长型开放读码框以空心方框(“□□”)表示。经比对EST序列(编号AA524283、AI088893、AI217369、和AI380811)中与NEP/ECE家族成员具有同源性的部分以“+==+”表示。“bp”= 碱基对。

图2: 分离并完整测序以生成IGS5DNA1和IGS5DNA2 cDNA序列的不同cDNA克隆相对位置的示意图。图中指示了用于PCR、5'RACE、和半同源性PCR克隆的PCR引物, 而且在本文件中也有所描述(分别以IP#表示)。IGS5CONS1和IGS5CONS2指合并所有获得的序列后获得的两种不同的共有毗邻序列群。IGS5DNA1和IGS5DNA2分别指在IGS5CONS1和IGS5CONS2中存在的开放读码框(“**”)。经比对EST序列(编号AA524283、AI088893、AI217369、和AI380811)中与NEP/ECE家族成员具有同源性的部分以“+==+”表示(IGS5EST)。“bp”= 碱基对。在基因组克隆IGS5/S1中鉴定的78bp片段以“IGS5/S1/78bp”表述。78bp可变外显子序列在克隆YCE231、YCE233、和YCE235中和在IGS5CONS2和IGS5DNA2中的缺乏以缺口表示。

图3: IGS5基因的RNA Master™印迹分析。

图4: 使用不同寡核苷酸通过交叠PCR装配而成的180bp片段的序列, 其编码POMC信号序列、Gly-Ser接头、His6标签、和IGS5胞外结构

域序列的起始。（*沉默突变，POMC信号序列的第57bp）。

图5: 载体pAcSG2SOLhuIGS5His6的质粒图。

图6: 在用重组杆状病毒IGBV73感染Sf9细胞后表达的成熟重组体可溶性His标记的人IGS5的预测蛋白质序列(切除长度为26个氨基酸的POMC信号序列后)。潜在的N-糖基化位点标有下划线。

图7: 脱糖基研究 - Western印迹分析。用3种重组体可溶性His6IGS5克隆(克隆1: 泳道1-3, 克隆2: 泳道4-6, 克隆3: 泳道7-9)感染后72小时收获CM, 并如文中所述加入和不加N-糖苷酶F进行处理。将10 μ l CM等同物上样至凝胶, 20 μ l未处理CM作为对照。使用抗His抗体(21E1B4EPR300, Innogenetics, 1 μ g/ml终浓度)进行检测。二抗是缀合碱性磷酸酶的兔抗小鼠抗体(Sigma A-1902)。用NBT-BCIP使条带显影。Mr标准是Biolabs宽范围MW标准(目录编号7707S)。

图8: 在还原条件下(+DTT)在12.5% PHAST凝胶(Pharmacia; 4 μ l/加样孔)上进行的对扁豆层析步骤的SDS-PAGE分析。通过银染使蛋白质显影。

图9: 纯化流程中不同阶段的IGS5的Western印迹分析。将样品在7.5% Minigel(Biorad MINI-Protean II)上分开, 并通过Western印迹进行分析, 其中使用抗His6一抗21E1B4, 缀合碱性磷酸酶的兔抗小鼠Ig作为二抗, 并通过NBT/BCIP检测。+/- DTT分别表示用或未用DTT还原蛋白质。

图10: 在还原条件下(+DTT)在12.5% PHAST凝胶(Pharmacia; 4 μ l/加样孔)上进行的对扁豆层析洗脱液1号合并液进行Zn-IMAC层析的不同咪唑洗脱合并液的SDS-PAGE分析。通过银染使蛋白质显影。

序列表

<110> SOLVAY PHARMACEUTICALS B.V.

<120> 金属蛋白酶家族中的新型人酶

<130> SPW 99.09 /H 99.26-WO

<140> .

<141>

<160> 31

<170> PatentIn Ver. 2.1

<210> 1

<211> 2076

<212> DNA

<213> 人类 (Homo sapiens)

<220>

<221> CDS

<222> (1)..(2073)

<400> 1

```

tgc acc acc cct ggc tgc gtg ata gca gct gcc agg atc ctc cag aac      48
Cys Thr Thr Pro Gly Cys Val Ile Ala Ala Ala Arg Ile Leu Gln Asn
   1             5             10             15

atg gac ccg acc acg gaa ccg tgt gac gac ttc tac cag ttt gca tgc      96
Met Asp Pro Thr Thr Glu Pro Cys Asp Asp Phe Tyr Gln Phe Ala Cys
           20             25             30

gga ggc tgg ctg cgg cgc cac gtg atc cct gag acc aac tca aga tac      144
Gly Gly Trp Leu Arg Arg His Val Ile Pro Glu Thr Asn Ser Arg Tyr
           35             40             45

agc atc ttt gac gtc ctc cgc gac gag ctg gag gtc atc ctc aaa gcg      192
Ser Ile Phe Asp Val Leu Arg Asp Glu Leu Glu Val Ile Leu Lys Ala
           50             55             60

gtg ctg gag aat tcg act gcc aag gac cgg ccg gct gtg gag aag gcc      240
Val Leu Glu Asn Ser Thr Ala Lys Asp Arg Pro Ala Val Glu Lys Ala
           65             70             75             80

agg acg ctg tac cgc tcc tgc atg aac cag agt gtg ata gag aag cga      288
Arg Thr Leu Tyr Arg Ser Cys Met Asn Gln Ser Val Ile Glu Lys Arg
           85             90             95

ggc tct cag ccc ctg ctg gac atc ttg gag gtg gtg gga ggc tgg ccg      336
Gly Ser Gln Pro Leu Leu Asp Ile Leu Glu Val Val Gly Gly Trp Pro
           100            105            110

gtg gcg atg gac agg tgg aac gag acc gta gga ctc gag tgg gag ctg      384
Val Ala Met Asp Arg Trp Asn Glu Thr Val Gly Leu Glu Trp Glu Leu
           115            120            125

gag cgg cag ctg gcg ctg atg aac tca cag ttc aac agg cgc gtc ctc      432
Glu Arg Gln Leu Ala Leu Met Asn Ser Gln Phe Asn Arg Arg Val Leu
           130            135            140

```


atc gac ctc ttc atc tgg aac gac gac cag aac tcc agc cgg cac atc	480
Ile Asp Leu Phe Ile Trp Asn Asp Asp Gln Asn Ser Ser Arg His Ile	
145 150 155 160	
atc tac ata gac cag ccc acc ttg ggc atg ccc tcc cga gag tac tac	528
Ile Tyr Ile Asp Gln Pro Thr Leu Gly Met Pro Ser Arg Glu Tyr Tyr	
165 170 175	
ttc aac ggc ggc agc aac cgg aag gtg cgg gaa gcc tac ctg cag ttc	576
Phe Asn Gly Gly Ser Asn Arg Lys Val Arg Glu Ala Tyr Leu Gln Phe	
180 185 190	
atg gtg tca gtg gcc acg ttg ctg cgg gag gat gca aac ctg ccc agg	624
Met Val Ser Val Ala Thr Leu Leu Arg Glu Asp Ala Asn Leu Pro Arg	
195 200 205	
gac agc tgc ctg gtg cag gag gac atg atg cag gtg ctg gag ctg gag	672
Asp Ser Cys Leu Val Gln Glu Asp Met Met Gln Val Leu Glu Leu Glu	
210 215 220	
aca cag ctg gcc aag gcc acg gta ccc cag gag gag aga cac gac gtc	720
Thr Gln Leu Ala Lys Ala Thr Val Pro Gln Glu Glu Arg His Asp Val	
225 230 235 240	
atc gcc ttg tac cac cgg atg gga ctg gag gag ctg caa agc cag ttt	768
Ile Ala Leu Tyr His Arg Met Gly Leu Glu Glu Leu Gln Ser Gln Phe	
245 250 255	
ggc ctg aag gga ttt aac tgg act ctg ttc ata caa act gtg cta tcc	816
Gly Leu Lys Gly Phe Asn Trp Thr Leu Phe Ile Gln Thr Val Leu Ser	
260 265 270	
tct gtc aaa atc aag ctg ctg cca gat gag gaa gtg gtg gtc tat ggc	864
Ser Val Lys Ile Lys Leu Leu Pro Asp Glu Glu Val Val Tyr Gly	
275 280 285	
atc ccc tac ctg cag aac ctt gaa aac atc atc gac acc tac tca gcc	912
Ile Pro Tyr Leu Gln Asn Leu Glu Asn Ile Ile Asp Thr Tyr Ser Ala	
290 295 300	
agg acc ata cag aac tac ctg gtc tgg cgc ctg gtg ctg gac cgc att	960
Arg Thr Ile Gln Asn Tyr Leu Val Trp Arg Leu Val Leu Asp Arg Ile	
305 310 315 320	
ggt agc cta agc cag aga ttc aag gac aca cga gtg aac tac cgc aag	1008
Gly Ser Leu Ser Gln Arg Phe Lys Asp Thr Arg Val Asn Tyr Arg Lys	
325 330 335	
gcg ctg ttt ggc aca atg gtg gag gag gtg cgc tgg cgt gaa tgt gtg	1056
Ala Leu Phe Gly Thr Met Val Glu Glu Val Arg Trp Arg Glu Cys Val	
340 345 350	
ggc tac gtc aac agc aac atg gag aac gcc gtg ggc tcc ctc tac gtc	1104
Gly Tyr Val Asn Ser Asn Met Glu Asn Ala Val Gly Ser Leu Tyr Val	
355 360 365	
agg gag gcg ttc cct gga gac agc aag agc atg gtc aga gaa ctc att	1152
Arg Glu Ala Phe Pro Gly Asp Ser Lys Ser Met Val Arg Glu Leu Ile	
370 375 380	
gac aag gtg cgg aca gtg ttt gtg gag acg ctg gac gag ctg ggc tgg	1200

Asp Lys Val Arg Thr Val Phe Val Glu Thr Leu Asp Glu Leu Gly Trp 385	390	395	400
atg gac gag gag tcc aag aag aag gcg cag gag aag gcc atg agc atc Met Asp Glu Glu Ser Lys Lys Lys Ala Gln Glu Lys Ala Met Ser Ile	405	410	415
1248			
cgg gag cag atc ggg cac cct gac tac atc ctg gag gag atg aac agg Arg Glu Gln Ile Gly His Pro Asp Tyr Ile Leu Glu Glu Met Asn Arg	420	425	430
1296			
cgc ctg gac gag gag tac tcc aat ctg aac ttc tca gag gac ctg tac Arg Leu Asp Glu Glu Tyr Ser Asn Leu Asn Phe Ser Glu Asp Leu Tyr	435	440	445
1344			
ttt gag aac agt ctg cag aac ctc aag gtg ggc gcc cag cgg agc ctc Phe Glu Asn Ser Leu Gln Asn Leu Lys Val Gly Ala Gln Arg Ser Leu	450	455	460
1392			
agg aag ctt cgg gaa aag gtg gac cca aat ctc tgg atc atc ggg gcg Arg Lys Leu Arg Glu Lys Val Asp Pro Asn Leu Trp Ile Ile Gly Ala	465	470	480
1440			
gcg gtg gtc aat gcg ttc tac tcc cca aac cga aac cag att gta ttc Ala Val Val Asn Ala Phe Tyr Ser Pro Asn Arg Asn Gln Ile Val Phe	485	490	495
1488			
cct gcc ggg atc ctc cag ccc ccc ttc ttc agc aag gag cag cca cag Pro Ala Gly Ile Leu Gln Pro Pro Phe Phe Ser Lys Glu Gln Pro Gln	500	505	510
1536			
gcc ttg aac ttt gga ggc att ggg atg gtg atc ggg cac gag atc acg Ala Leu Asn Phe Gly Gly Ile Gly Met Val Ile Gly His Glu Ile Thr	515	520	525
1584			
cac ggc ttt gac gac aat ggc cgg aac ttc gac aag aat ggc aac atg His Gly Phe Asp Asp Asn Gly Arg Asn Phe Asp Lys Asn Gly Asn Met	530	535	540
1632			
atg gat tgg tgg agt aac ttc tcc acc cag cac ttc cgg gag cag tca Met Asp Trp Trp Ser Asn Phe Ser Thr Gln His Phe Arg Glu Gln Ser	545	550	560
1680			
gag tgc atg atc tac cag tac ggc aac tac tcc tgg gac ctg gca gac Glu Cys Met Ile Tyr Gln Tyr Gly Asn Tyr Ser Trp Asp Leu Ala Asp	565	570	575
1728			
gaa cag aac gtg aac gga ttc aac acc ctt ggg gaa aac att gct gac Glu Gln Asn Val Asn Gly Phe Asn Thr Leu Gly Glu Asn Ile Ala Asp	580	585	590
1776			
aac gga ggg gtg cgg caa gcc tat aag gcc tac ctc aag tgg atg gca Asn Gly Gly Val Arg Gln Ala Tyr Lys Ala Tyr Leu Lys Trp Met Ala	595	600	605
1824			
gag ggt ggc aag gac cag cag ctg ccc ggc ctg gat ctc acc cat gag Glu Gly Gly Lys Asp Gln Gln Leu Pro Gly Leu Asp Leu Thr His Glu	610	615	620
1872			
cag ctc ttc ttc atc aac tac gcc cag gtg tgg tgc ggg tcc tac cgg Gln Leu Phe Phe Ile Asn Tyr Ala Gln Val Trp Cys Gly Ser Tyr Arg			1920

625	630	635	640	
ccc gag ttc gcc atc caa tcc atc aag aca gac gtc cac agt ccc ctg				1968
Pro Glu Phe Ala Ile Gln Ser Ile Lys Thr Asp Val His Ser Pro Leu	645	650	655	
aag tac agg gta ctg ggg tcg ctg cag aac ctg gcc gcc ttc gca gac				2016
Lys Tyr Arg Val Leu Gly Ser Leu Gln Asn Leu Ala Ala Phe Ala Asp	660	665	670	
acg ttc cac tgt gcc cgg ggc acc ccc atg cac ccc aag gag cga tgc				2064
Thr Phe His Cys Ala Arg Gly Thr Pro Met His Pro Lys Glu Arg Cys	675	680	685	
cgc gtg tgg tag				2076
Arg Val Trp	690			

<210> 2

<211> 691

<212> PRT

<213> 人类 (Homo sapiens)

<400> 2

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

Met Val Ser Val Ala Thr Leu Leu Arg Glu Asp Ala Asn Leu Pro Arg
 195 200 205
 Asp Ser Cys Leu Val Gln Glu Asp Met Met Gln Val Leu Glu Leu Glu
 210 215 220
 Thr Gln Leu Ala Lys Ala Thr Val Pro Gln Glu Glu Arg His Asp Val
 225 230 235 240
 Ile Ala Leu Tyr His Arg Met Gly Leu Glu Glu Leu Gln Ser Gln Phe
 245 250 255
 Gly Leu Lys Gly Phe Asn Trp Thr Leu Phe Ile Gln Thr Val Leu Ser
 260 265 270
 Ser Val Lys Ile Lys Leu Leu Pro Asp Glu Glu Val Val Val Tyr Gly
 275 280 285
 Ile Pro Tyr Leu Gln Asn Leu Glu Asn Ile Ile Asp Thr Tyr Ser Ala
 290 295 300
 Arg Thr Ile Gln Asn Tyr Leu Val Trp Arg Leu Val Leu Asp Arg Ile
 305 310 315 320
 Gly Ser Leu Ser Gln Arg Phe Lys Asp Thr Arg Val Asn Tyr Arg Lys
 325 330 335
 Ala Leu Phe Gly Thr Met Val Glu Glu Val Arg Trp Arg Glu Cys Val
 340 345 350
 Gly Tyr Val Asn Ser Asn Met Glu Asn Ala Val Gly Ser Leu Tyr Val
 355 360 365
 Arg Glu Ala Phe Pro Gly Asp Ser Lys Ser Met Val Arg Glu Leu Ile
 370 375 380
 Asp Lys Val Arg Thr Val Phe Val Glu Thr Leu Asp Glu Leu Gly Trp
 385 390 395 400
 Met Asp Glu Glu Ser Lys Lys Lys Ala Gln Glu Lys Ala Met Ser Ile
 405 410 415
 Arg Glu Gln Ile Gly His Pro Asp Tyr Ile Leu Glu Glu Met Asn Arg
 420 425 430
 Arg Leu Asp Glu Glu Tyr Ser Asn Leu Asn Phe Ser Glu Asp Leu Tyr
 435 440 445
 Phe Glu Asn Ser Leu Gln Asn Leu Lys Val Gly Ala Gln Arg Ser Leu
 450 455 460
 Arg Lys Leu Arg Glu Lys Val Asp Pro Asn Leu Trp Ile Ile Gly Ala
 465 470 475 480
 Ala Val Val Asn Ala Phe Tyr Ser Pro Asn Arg Asn Gln Ile Val Phe
 485 490 495
 Pro Ala Gly Ile Leu Gln Pro Pro Phe Phe Ser Lys Glu Gln Pro Gln
 500 505 510
 Ala Leu Asn Phe Gly Gly Ile Gly Met Val Ile Gly His Glu Ile Thr

50	55	60	
tta cag gag gag agg acc ttt gta aaa cga aaa ccc cga ggg atc cca			240
Leu Gln Glu Glu Arg Thr Phe Val Lys Arg Lys Pro Arg Gly Ile Pro			
65	70	75	80
gag gcc caa gag gtg agc gag gtc tgc acc acc cct ggc tgc gtg ata			288
Glu Ala Gln Glu Val Ser Glu Val Cys Thr Thr Pro Gly Cys Val Ile			
	85	90	95
gca gct gcc agg atc ctc cag aac atg gac ccg acc acg gaa ccg tgt			336
Ala Ala Ala Arg Ile Leu Gln Asn Met Asp Pro Thr Thr Glu Pro Cys			
	100	105	110
gac gac ttc tac cag ttt gca tgc gga ggc tgg ctg cgg cgc cac gtg			384
Asp Asp Phe Tyr Gln Phe Ala Cys Gly Gly Trp Leu Arg Arg His Val			
	115	120	125
atc cct gag acc aac tca aga tac agc atc ttt gac gtc ctc cgc gac			432
Ile Pro Glu Thr Asn Ser Arg Tyr Ser Ile Phe Asp Val Leu Arg Asp			
	130	135	140
gag ctg gag gtc atc ctc aaa gcg gtg ctg gag aat tcg act gcc aag			480
Glu Leu Glu Val Ile Leu Lys Ala Val Leu Glu Asn Ser Thr Ala Lys			
	145	150	155
gac cgg ccg gct gtg gag aag gcc agg acg ctg tac cgc tcc tgc atg			528
Asp Arg Pro Ala Val Glu Lys Ala Arg Thr Leu Tyr Arg Ser Cys Met			
	165	170	175
aac cag agt gtg ata gag aag cga ggc tct cag ccc ctg ctg gac atc			576
Asn Gln Ser Val Ile Glu Lys Arg Gly Ser Gln Pro Leu Leu Asp Ile			
	180	185	190
ttg gag gtg gtg gga ggc tgg ccg gtg gcg atg gac agg tgg aac gag			624
Leu Glu Val Val Gly Gly Trp Pro Val Ala Met Asp Arg Trp Asn Glu			
	195	200	205
acc gta gga ctc gag tgg gag ctg gag cgg cag ctg gcg ctg atg aac			672
Thr Val Gly Leu Glu Trp Glu Leu Glu Arg Gln Leu Ala Leu Met Asn			
	210	215	220
tca cag ttc aac agg cgc gtc ctc atc gac ctc ttc atc tgg aac gac			720
Ser Gln Phe Asn Arg Arg Val Leu Ile Asp Leu Phe Ile Trp Asn Asp			
	225	230	235
gac cag aac tcc agc cgg cac atc atc tac ata gac cag ccc acc ttg			768
Asp Gln Asn Ser Ser Arg His Ile Ile Tyr Ile Asp Gln Pro Thr Leu			
	245	250	255
ggc atg ccc tcc cga gag tac tac ttc aac ggc ggc agc aac cgg aag			816
Gly Met Pro Ser Arg Glu Tyr Tyr Phe Asn Gly Gly Ser Asn Arg Lys			
	260	265	270
gtg cgg gaa gcc tac ctg cag ttc atg gtg tca gtg gcc acg ttg ctg			864
Val Arg Glu Ala Tyr Leu Gln Phe Met Val Ser Val Ala Thr Leu Leu			
	275	280	285
cgg gag gat gca aac ctg ccc agg gac agc tgc ctg gtg cag gag gac			912
Arg Glu Asp Ala Asn Leu Pro Arg Asp Ser Cys Leu Val Gln Glu Asp			
	290	295	300

atg atg cag gtg ctg gag ctg gag aca cag ctg gcc aag gcc acg gta	960
Met Met Gln Val Leu Glu Leu Glu Thr Gln Leu Ala Lys Ala Thr Val	
305 310 315 320	
ccc cag gag gag aga cac gac gtc atc gcc ttg tac cac cgg atg gga	1008
Pro Gln Glu Glu Arg His Asp Val Ile Ala Leu Tyr His Arg Met Gly	
325 330 335	
ctg gag gag ctg caa agc cag ttt ggc ctg aag gga ttt aac tgg act	1056
Leu Glu Glu Leu Gln Ser Gln Phe Gly Leu Lys Gly Phe Asn Trp Thr	
340 345 350	
ctg ttc ata caa act gtg cta tcc tct gtc aaa atc aag ctg ctg cca	1104
Leu Phe Ile Gln Thr Val Leu Ser Ser Val Lys Ile Lys Leu Leu Pro	
355 360 365	
gat gag gaa gtg gtg gtc tat ggc atc ccc tac ctg cag aac ctt gaa	1152
Asp Glu Glu Val Val Val Tyr Gly Ile Pro Tyr Leu Gln Asn Leu Glu	
370 375 380	
aac atc atc gac acc tac tca gcc agg acc ata cag aac tac ctg gtc	1200
Asn Ile Ile Asp Thr Tyr Ser Ala Arg Thr Ile Gln Asn Tyr Leu Val	
385 390 395 400	
tgg cgc ctg gtg ctg gac cgc att ggt agc cta agc cag aga ttc aag	1248
Trp Arg Leu Val Leu Asp Arg Ile Gly Ser Leu Ser Gln Arg Phe Lys	
405 410 415	
gac aca cga gtg aac tac cgc aag gcg ctg ttt ggc aca atg gtg gag	1296
Asp Thr Arg Val Asn Tyr Arg Lys Ala Leu Phe Gly Thr Met Val Glu	
420 425 430	
gag gtg cgc tgg cgt gaa tgt gtg ggc tac gtc aac agc aac atg gag	1344
Glu Val Arg Trp Arg Glu Cys Val Gly Tyr Val Asn Ser Asn Met Glu	
435 440 445	
aac gcc gtg ggc tcc ctc tac gtc agg gag gcg ttc cct gga gac agc	1392
Asn Ala Val Gly Ser Leu Tyr Val Arg Glu Ala Phe Pro Gly Asp Ser	
450 455 460	
aag agc atg gtc aga gaa ctc att gac aag gtg cgg aca gtg ttt gtg	1440
Lys Ser Met Val Arg Glu Leu Ile Asp Lys Val Arg Thr Val Phe Val	
465 470 475 480	
gag acg ctg gac gag ctg ggc tgg atg gac gag gag tcc aag aag aag	1488
Glu Thr Leu Asp Glu Leu Gly Trp Met Asp Glu Glu Ser Lys Lys Lys	
485 490 495	
gcg cag gag aag gcc atg agc atc cgg gag cag atc ggg cac cct gac	1536
Ala Gln Glu Lys Ala Met Ser Ile Arg Glu Gln Ile Gly His Pro Asp	
500 505 510	
tac atc ctg gag gag atg aac agg cgc ctg gac gag gag tac tcc aat	1584
Tyr Ile Leu Glu Glu Met Asn Arg Arg Leu Asp Glu Glu Tyr Ser Asn	
515 520 525	
ctg aac ttc tca gag gac ctg tac ttt gag aac agt ctg cag aac ctc	1632
Leu Asn Phe Ser Glu Asp Leu Tyr Phe Glu Asn Ser Leu Gln Asn Leu	
530 535 540	

<210> 4
 <211> 779
 <212> PRT
 <213> 人类 (Homo sapiens)
 <400> 4
 Met Gly Lys Ser Glu Gly Pro Val Gly Met Val Glu Ser Ala Gly Arg
 1 5 10 15
 Ala Gly Gln Lys Arg Pro Gly Phe Leu Glu Gly Gly Leu Leu Leu
 20 25 30
 Leu Leu Leu Val Thr Ala Ala Leu Val Ala Leu Gly Val Leu Tyr Ala
 35 40 45
 Asp Arg Arg Gly Lys Gln Leu Pro Arg Leu Ala Ser Arg Leu Cys Phe
 50 55 60
 Leu Gln Glu Glu Arg Thr Phe Val Lys Arg Lys Pro Arg Gly Ile Pro
 65 70 75 80
 Glu Ala Gln Glu Val Ser Glu Val Cys Thr Thr Pro Gly Cys Val Ile
 85 90 95
 Ala Ala Ala Arg Ile Leu Gln Asn Met Asp Pro Thr Thr Glu Pro Cys
 100 105 110
 Asp Asp Phe Tyr Gln Phe Ala Cys Gly Gly Trp Leu Arg Arg His Val
 115 120 125
 Ile Pro Glu Thr Asn Ser Arg Tyr Ser Ile Phe Asp Val Leu Arg Asp
 130 135 140
 Glu Leu Glu Val Ile Leu Lys Ala Val Leu Glu Asn Ser Thr Ala Lys
 145 150 155 160
 Asp Arg Pro Ala Val Glu Lys Ala Arg Thr Leu Tyr Arg Ser Cys Met
 165 170 175
 Asn Gln Ser Val Ile Glu Lys Arg Gly Ser Gln Pro Leu Leu Asp Ile
 180 185 190
 Leu Glu Val Val Gly Gly Trp Pro Val Ala Met Asp Arg Trp Asn Glu
 195 200 205
 Thr Val Gly Leu Glu Trp Glu Leu Glu Arg Gln Leu Ala Leu Met Asn
 210 215 220
 Ser Gln Phe Asn Arg Arg Val Leu Ile Asp Leu Phe Ile Trp Asn Asp
 225 230 235 240
 Asp Gln Asn Ser Ser Arg His Ile Ile Tyr Ile Asp Gln Pro Thr Leu
 245 250 255
 Gly Met Pro Ser Arg Glu Tyr Tyr Phe Asn Gly Gly Ser Asn Arg Lys
 260 265 270
 Val Arg Glu Ala Tyr Leu Gln Phe Met Val Ser Val Ala Thr Leu Leu
 275 280 285
 Arg Glu Asp Ala Asn Leu Pro Arg Asp Ser Cys Leu Val Gln Glu Asp

290	295	300
Met Met Gln Val Leu Glu Leu Glu Thr Gln Leu Ala Lys Ala Thr Val 305	310	315
Pro Gln Glu Glu Arg His Asp Val Ile Ala Leu Tyr His Arg Met Gly 325	330	335
Leu Glu Glu Leu Gln Ser Gln Phe Gly Leu Lys Gly Phe Asn Trp Thr 340	345	350
Leu Phe Ile Gln Thr Val Leu Ser Ser Val Lys Ile Lys Leu Leu Pro 355	360	365
Asp Glu Glu Val Val Val Tyr Gly Ile Pro Tyr Leu Gln Asn Leu Glu 370	375	380
Asn Ile Ile Asp Thr Tyr Ser Ala Arg Thr Ile Gln Asn Tyr Leu Val 385	390	395
Trp Arg Leu Val Leu Asp Arg Ile Gly Ser Leu Ser Gln Arg Phe Lys 405	410	415
Asp Thr Arg Val Asn Tyr Arg Lys Ala Leu Phe Gly Thr Met Val Glu 420	425	430
Glu Val Arg Trp Arg Glu Cys Val Gly Tyr Val Asn Ser Asn Met Glu 435	440	445
Asn Ala Val Gly Ser Leu Tyr Val Arg Glu Ala Phe Pro Gly Asp Ser 450	455	460
Lys Ser Met Val Arg Glu Leu Ile Asp Lys Val Arg Thr Val Phe Val 465	470	475
Glu Thr Leu Asp Glu Leu Gly Trp Met Asp Glu Glu Ser Lys Lys Lys 485	490	495
Ala Gln Glu Lys Ala Met Ser Ile Arg Glu Gln Ile Gly His Pro Asp 500	505	510
Tyr Ile Leu Glu Glu Met Asn Arg Arg Leu Asp Glu Glu Tyr Ser Asn 515	520	525
Leu Asn Phe Ser Glu Asp Leu Tyr Phe Glu Asn Ser Leu Gln Asn Leu 530	535	540
Lys Val Gly Ala Gln Arg Ser Leu Arg Lys Leu Arg Glu Lys Val Asp 545	550	555
Pro Asn Leu Trp Ile Ile Gly Ala Ala Val Val Asn Ala Phe Tyr Ser 565	570	575
Pro Asn Arg Asn Gln Ile Val Phe Pro Ala Gly Ile Leu Gln Pro Pro 580	585	590
Phe Phe Ser Lys Glu Gln Pro Gln Ala Leu Asn Phe Gly Gly Ile Gly 595	600	605
Met Val Ile Gly His Glu Ile Thr His Gly Phe Asp Asp Asn Gly Arg 610	615	620

Asn Phe Asp Lys Asn Gly Asn Met Met Asp Trp Trp Ser Asn Phe Ser
 625 630 635 640
 Thr Gln His Phe Arg Glu Gln Ser Glu Cys Met Ile Tyr Gln Tyr Gly
 645 650 655
 Asn Tyr Ser Trp Asp Leu Ala Asp Glu Gln Asn Val Asn Gly Phe Asn
 660 665 670
 Thr Leu Gly Glu Asn Ile Ala Asp Asn Gly Gly Val Arg Gln Ala Tyr
 675 680 685
 Lys Ala Tyr Leu Lys Trp Met Ala Glu Gly Gly Lys Asp Gln Gln Leu
 690 695 700
 Pro Gly Leu Asp Leu Thr His Glu Gln Leu Phe Phe Ile Asn Tyr Ala
 705 710 715 720
 Gln Val Trp Cys Gly Ser Tyr Arg Pro Glu Phe Ala Ile Gln Ser Ile
 725 730 735
 Lys Thr Asp Val His Ser Pro Leu Lys Tyr Arg Val Leu Gly Ser Leu
 740 745 750
 Gln Asn Leu Ala Ala Phe Ala Asp Thr Phe His Cys Ala Arg Gly Thr
 755 760 765
 Pro Met His Pro Lys Glu Arg Cys Arg Val Trp
 770 775

<210> 5

<211> 2262

<212> DNA

<213> 人类 (Homo sapiens)

<220>

<221> CDS

<222> (1)..(2259)

<400> 5

atg ggg aag tcc gaa ggc cca gtg ggg atg gtg gag agc gcc ggc cgt 48
 Met Gly Lys Ser Glu Gly Pro Val Gly Met Val Glu Ser Ala Gly Arg
 1 5 10 15
 gca ggg cag aag cgc ccg ggg ttc ctg gag ggg ggg ctg ctg ctg ctg 96
 Ala Gly Gln Lys Arg Pro Gly Phe Leu Glu Gly Gly Leu Leu Leu Leu
 20 25 30
 ctg ctg ctg gtg acc gct gcc ctg gtg gcc ttg ggt gtc ctc tac gcc 144
 Leu Leu Leu Val Thr Ala Ala Leu Val Ala Leu Gly Val Leu Tyr Ala
 35 40 45
 gac cgc aga ggg atc cca gag gcc caa gag gtg agc gag gtc tgc acc 192
 Asp Arg Arg Gly Ile Pro Glu Ala Gln Glu Val Ser Glu Val Cys Thr
 50 55 60
 acc cct ggc tgc gtg ata gca gct gcc agg atc ctc cag aac atg gac 240
 Thr Pro Gly Cys Val Ile Ala Ala Ala Arg Ile Leu Gln Asn Met Asp

65	70	75	80	
ccg acc acg gaa ccg tgt gac gac ttc tac cag ttt gca tgc gga ggc				288
Pro Thr Thr Glu Pro Cys Asp Asp Phe Tyr Gln Phe Ala Cys Gly Gly	85	90	95	
tgg ctg cgg cgc cac gtg atc cct gag acc aac tca aga tac agc atc				336
Trp Leu Arg Arg His Val Ile Pro Glu Thr Asn Ser Arg Tyr Ser Ile	100	105	110	
ttt gac gtc ctc cgc gac gag ctg gag gtc atc ctc aaa gcg gtg ctg				384
Phe Asp Val Leu Arg Asp Glu Leu Glu Val Ile Leu Lys Ala Val Leu	115	120	125	
gag aat tcg act gcc aag gac cgg ccg gct gtg gag aag gcc agg acg				432
Glu Asn Ser Thr Ala Lys Asp Arg Pro Ala Val Glu Lys Ala Arg Thr	130	135	140	
ctg tac cgc tcc tgc atg aac cag agt gtg ata gag aag cga ggc tct				480
Leu Tyr Arg Ser Cys Met Asn Gln Ser Val Ile Glu Lys Arg Gly Ser	145	150	155	160
cag ccc ctg ctg gac atc ttg gag gtg gtg gga ggc tgg ccg gtg gcg				528
Gln Pro Leu Leu Asp Ile Leu Glu Val Val Gly Gly Trp Pro Val Ala	165	170	175	
atg gac agg tgg aac gag acc gta gga ctc gag tgg gag ctg gag cgg				576
Met Asp Arg Trp Asn Glu Thr Val Gly Leu Glu Trp Glu Leu Glu Arg	180	185	190	
cag ctg gcg ctg atg aac tca cag ttc aac agg cgc gtc ctc atc gac				624
Gln Leu Ala Leu Met Asn Ser Gln Phe Asn Arg Arg Val Leu Ile Asp	195	200	205	
ctc ttc atc tgg aac gac gac cag aac tcc agc cgg cac atc atc tac				672
Leu Phe Ile Trp Asn Asp Asp Gln Asn Ser Ser Arg His Ile Ile Tyr	210	215	220	
ata gac cag ccc acc ttg ggc atg ccc tcc cga gag tac tac ttc aac				720
Ile Asp Gln Pro Thr Leu Gly Met Pro Ser Arg Glu Tyr Tyr Phe Asn	225	230	235	240
ggc ggc agc aac cgg aag gtg cgg gaa gcc tac ctg cag ttc atg gtg				768
Gly Gly Ser Asn Arg Lys Val Arg Glu Ala Tyr Leu Gln Phe Met Val	245	250	255	
tca gtg gcc acg ttg ctg cgg gag gat gca aac ctg ccc agg gac agc				816
Ser Val Ala Thr Leu Leu Arg Glu Asp Ala Asn Leu Pro Arg Asp Ser	260	265	270	
tgc ctg gtg cag gag gac atg atg cag gtg ctg gag ctg gag aca cag				864
Cys Leu Val Gln Glu Asp Met Met Gln Val Leu Glu Leu Glu Thr Gln	275	280	285	
ctg gcc aag gcc acg gta ccc cag gag gag aga cac gac gtc atc gcc				912
Leu Ala Lys Ala Thr Val Pro Gln Glu Glu Arg His Asp Val Ile Ala	290	295	300	
ttg tac cac cgg atg gga ctg gag gag ctg caa agc cag ttt ggc ctg				960
Leu Tyr His Arg Met Gly Leu Glu Glu Leu Gln Ser Gln Phe Gly Leu	305	310	315	320

aag gga ttt aac tgg act ctg ttc ata caa act gtg cta tcc tct gtc	1008
Lys Gly Phe Asn Trp Thr Leu Phe Ile Gln Thr Val Leu Ser Ser Val	
325 330 335	
aaa atc aag ctg ctg cca gat gag gaa gtg gtg gtc tat ggc atc ccc	1056
Lys Ile Lys Leu Leu Pro Asp Glu Glu Val Val Val Tyr Gly Ile Pro	
340 345 350	
tac ctg cag aac ctt gaa aac atc atc gac acc tac tca gcc agg acc	1104
Tyr Leu Gln Asn Leu Glu Asn Ile Ile Asp Thr Tyr Ser Ala Arg Thr	
355 360 365	
ata cag aac tac ctg gtc tgg cgc ctg gtg ctg gac cgc att ggt agc	1152
Ile Gln Asn Tyr Leu Val Trp Arg Leu Val Leu Asp Arg Ile Gly Ser	
370 375 380	
cta agc cag aga ttc aag gac aca cga gtg aac tac cgc aag gcg ctg	1200
Leu Ser Gln Arg Phe Lys Asp Thr Arg Val Asn Tyr Arg Lys Ala Leu	
385 390 395 400	
ttt ggc aca atg gtg gag gag gtg cgc tgg cgt gaa tgt gtg ggc tac	1248
Phe Gly Thr Met Val Glu Glu Val Arg Trp Arg Glu Cys Val Gly Tyr	
405 410 415	
gtc aac agc aac atg gag aac gcc gtg ggc tcc ctc tac gtc agg gag	1296
Val Asn Ser Asn Met Glu Asn Ala Val Gly Ser Leu Tyr Val Arg Glu	
420 425 430	
gcg ttc cct gga gac agc aag agc atg gtc aga gaa ctc att gac aag	1344
Ala Phe Pro Gly Asp Ser Lys Ser Met Val Arg Glu Leu Ile Asp Lys	
435 440 445	
gtg cgg aca gtg ttt gtg gag acg ctg gac gag ctg ggc tgg atg gac	1392
Val Arg Thr Val Phe Val Glu Thr Leu Asp Glu Leu Gly Trp Met Asp	
450 455 460	
gag gag tcc aag aag aag gcg cag gag aag gcc atg agc atc cgg gag	1440
Glu Glu Ser Lys Lys Lys Ala Gln Glu Lys Ala Met Ser Ile Arg Glu	
465 470 475 480	
cag atc ggg cac cct gac tac atc ctg gag gag atg aac agg cgc ctg	1488
Gln Ile Gly His Pro Asp Tyr Ile Leu Glu Glu Met Asn Arg Arg Leu	
485 490 495	
gac gag gag tac tcc aat ctg aac ttc tca gag gac ctg tac ttt gag	1536
Asp Glu Glu Tyr Ser Asn Leu Asn Phe Ser Glu Asp Leu Tyr Phe Glu	
500 505 510	
aac agt ctg cag aac ctc aag gtg ggc gcc cag cgg agc ctc agg aag	1584
Asn Ser Leu Gln Asn Leu Lys Val Gly Ala Gln Arg Ser Leu Arg Lys	
515 520 525	
ctt cgg gaa aag gtg gac cca aat ctc tgg atc atc ggg gcg gcg gtg	1632
Leu Arg Glu Lys Val Asp Pro Asn Leu Trp Ile Ile Gly Ala Ala Val	
530 535 540	
gtc aat gcg ttc tac tcc cca aac cga aac cag att gta ttc cct gcc	1680
Val Asn Ala Phe Tyr Ser Pro Asn Arg Asn Gln Ile Val Phe Pro Ala	
545 550 555 560	

```

ggg atc ctc cag ccc ccc ttc ttc agc aag gag cag cca cag gcc ttg 1728
Gly Ile Leu Gln Pro Pro Phe Phe Ser Lys Glu Gln Pro Gln Ala Leu
                    565                                570                                575

aac ttt gga ggc att ggg atg gtg atc ggg cac gag atc acg cac ggc 1776
Asn Phe Gly Gly Ile Gly Met Val Ile Gly His Glu Ile Thr His Gly
                    580                                585                                590

ttt gac gac aat ggc cgg aac ttc gac aag aat ggc aac atg atg gat 1824
Phe Asp Asp Asn Gly Arg Asn Phe Asp Lys Asn Gly Asn Met Met Asp
                    595                                600                                605

tgg tgg agt aac ttc tcc acc cag cac ttc cgg gag cag tca gag tgc 1872
Trp Trp Ser Asn Phe Ser Thr Gln His Phe Arg Glu Gln Ser Glu Cys
                    610                                615                                620

atg atc tac cag tac ggc aac tac tcc tgg gac ctg gca gac gaa cag 1920
Met Ile Tyr Gln Tyr Gly Asn Tyr Ser Trp Asp Leu Ala Asp Glu Gln
                    625                                630                                635                                640

aac gtg aac gga ttc aac acc ctt ggg gaa aac att gct gac aac gga 1968
Asn Val Asn Gly Phe Asn Thr Leu Gly Glu Asn Ile Ala Asp Asn Gly
                    645                                650                                655

ggg gtg cgg caa gcc tat aag gcc tac ctc aag tgg atg gca gag ggt 2016
Gly Val Arg Gln Ala Tyr Lys Ala Tyr Leu Lys Trp Met Ala Glu Gly
                    660                                665                                670

ggc aag gac cag cag ctg ccc ggc ctg gat ctc acc cat gag cag ctc 2064
Gly Lys Asp Gln Gln Leu Pro Gly Leu Asp Leu Thr His Glu Gln Leu
                    675                                680                                685

ttc ttc atc aac tac gcc cag gtg tgg tgc ggg tcc tac cgg ccc gag 2112
Phe Phe Ile Asn Tyr Ala Gln Val Trp Cys Gly Ser Tyr Arg Pro Glu
                    690                                695                                700

ttc gcc atc caa tcc atc aag aca gac gtc cac agt ccc ctg aag tac 2160
Phe Ala Ile Gln Ser Ile Lys Thr Asp Val His Ser Pro Leu Lys Tyr
                    705                                710                                715                                720

agg gta ctg ggg tcg ctg cag aac ctg gcc gcc ttc gca gac acg ttc 2208
Arg Val Leu Gly Ser Leu Gln Asn Leu Ala Ala Phe Ala Asp Thr Phe
                    725                                730                                735

cac tgt gcc cgg ggc acc ccc atg cac ccc aag gag cga tgc cgc gtg 2256
His Cys Ala Arg Gly Thr Pro Met His Pro Lys Glu Arg Cys Arg Val
                    740                                745                                750

tgg tag 2262
Trp

<210> 6
<211> 753
<212> PRT
<213> 人类 (Homo sapiens)

<400> 6
Met Gly Lys Ser Glu Gly Pro Val Gly Met Val Glu Ser Ala Gly Arg
1 5 10 15

```

Ala Gly Gln Lys Arg Pro Gly Phe Leu Glu Gly Gly Leu Leu Leu Leu
20 25 30

Leu Leu Leu Val Thr Ala Ala Leu Val Ala Leu Gly Val Leu Tyr Ala
35 40 45

Asp Arg Arg Gly Ile Pro Glu Ala Gln Glu Val Ser Glu Val Cys Thr
50 55 60

Thr Pro Gly Cys Val Ile Ala Ala Ala Arg Ile Leu Gln Asn Met Asp
65 70 75 80

Pro Thr Thr Glu Pro Cys Asp Asp Phe Tyr Gln Phe Ala Cys Gly Gly
85 90 95

Trp Leu Arg Arg His Val Ile Pro Glu Thr Asn Ser Arg Tyr Ser Ile
100 105 110

Phe Asp Val Leu Arg Asp Glu Leu Glu Val Ile Leu Lys Ala Val Leu
115 120 125

Glu Asn Ser Thr Ala Lys Asp Arg Pro Ala Val Glu Lys Ala Arg Thr
130 135 140

Leu Tyr Arg Ser Cys Met Asn Gln Ser Val Ile Glu Lys Arg Gly Ser
145 150 155 160

Gln Pro Leu Leu Asp Ile Leu Glu Val Val Gly Gly Trp Pro Val Ala
165 170 175

Met Asp Arg Trp Asn Glu Thr Val Gly Leu Glu Trp Glu Leu Glu Arg
180 185 190

Gln Leu Ala Leu Met Asn Ser Gln Phe Asn Arg Arg Val Leu Ile Asp
195 200 205

Leu Phe Ile Trp Asn Asp Asp Gln Asn Ser Ser Arg His Ile Ile Tyr
210 215 220

Ile Asp Gln Pro Thr Leu Gly Met Pro Ser Arg Glu Tyr Tyr Phe Asn
225 230 235 240

Gly Gly Ser Asn Arg Lys Val Arg Glu Ala Tyr Leu Gln Phe Met Val
245 250 255

Ser Val Ala Thr Leu Leu Arg Glu Asp Ala Asn Leu Pro Arg Asp Ser
260 265 270

Cys Leu Val Gln Glu Asp Met Met Gln Val Leu Glu Leu Glu Thr Gln
275 280 285

Leu Ala Lys Ala Thr Val Pro Gln Glu Glu Arg His Asp Val Ile Ala
290 295 300

Leu Tyr His Arg Met Gly Leu Glu Glu Leu Gln Ser Gln Phe Gly Leu
305 310 315 320

Lys Gly Phe Asn Trp Thr Leu Phe Ile Gln Thr Val Leu Ser Ser Val
325 330 335

Lys Ile Lys Leu Leu Pro Asp Glu Glu Val Val Val Tyr Gly Ile Pro

340	345	350
Tyr Leu Gln Asn Leu Glu Asn Ile Ile Asp Thr Tyr Ser Ala Arg Thr 355	360	365
Ile Gln Asn Tyr Leu Val Trp Arg Leu Val Leu Asp Arg Ile Gly Ser 370	375	380
Leu Ser Gln Arg Phe Lys Asp Thr Arg Val Asn Tyr Arg Lys Ala Leu 385	390	395
Phe Gly Thr Met Val Glu Glu Val Arg Trp Arg Glu Cys Val Gly Tyr 405	410	415
Val Asn Ser Asn Met Glu Asn Ala Val Gly Ser Leu Tyr Val Arg Glu 420	425	430
Ala Phe Pro Gly Asp Ser Lys Ser Met Val Arg Glu Leu Ile Asp Lys 435	440	445
Val Arg Thr Val Phe Val Glu Thr Leu Asp Glu Leu Gly Trp Met Asp 450	455	460
Glu Glu Ser Lys Lys Lys Ala Gln Glu Lys Ala Met Ser Ile Arg Glu 465	470	475
Gln Ile Gly His Pro Asp Tyr Ile Leu Glu Glu Met Asn Arg Arg Leu 485	490	495
Asp Glu Glu Tyr Ser Asn Leu Asn Phe Ser Glu Asp Leu Tyr Phe Glu 500	505	510
Asn Ser Leu Gln Asn Leu Lys Val Gly Ala Gln Arg Ser Leu Arg Lys 515	520	525
Leu Arg Glu Lys Val Asp Pro Asn Leu Trp Ile Ile Gly Ala Ala Val 530	535	540
Val Asn Ala Phe Tyr Ser Pro Asn Arg Asn Gln Ile Val Phe Pro Ala 545	550	555
Gly Ile Leu Gln Pro Pro Phe Phe Ser Lys Glu Gln Pro Gln Ala Leu 565	570	575
Asn Phe Gly Gly Ile Gly Met Val Ile Gly His Glu Ile Thr His Gly 580	585	590
Phe Asp Asp Asn Gly Arg Asn Phe Asp Lys Asn Gly Asn Met Met Asp 595	600	605
Trp Trp Ser Asn Phe Ser Thr Gln His Phe Arg Glu Gln Ser Glu Cys 610	615	620
Met Ile Tyr Gln Tyr Gly Asn Tyr Ser Trp Asp Leu Ala Asp Glu Gln 625	630	635
Asn Val Asn Gly Phe Asn Thr Leu Gly Glu Asn Ile Ala Asp Asn Gly 645	650	655
Gly Val Arg Gln Ala Tyr Lys Ala Tyr Leu Lys Trp Met Ala Glu Gly 660	665	670

Gly Lys Asp Gln Gln Leu Pro Gly Leu Asp Leu Thr His Glu Gln Leu
675 680 685

Phe Phe Ile Asn Tyr Ala Gln Val Trp Cys Gly Ser Tyr Arg Pro Glu
690 695 700

Phe Ala Ile Gln Ser Ile Lys Thr Asp Val His Ser Pro Leu Lys Tyr
705 710 715 720

Arg Val Leu Gly Ser Leu Gln Asn Leu Ala Ala Phe Ala Asp Thr Phe
725 730 735

His Cys Ala Arg Gly Thr Pro Met His Pro Lys Glu Arg Cys Arg Val
740 745 750

Trp

<210> 7

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 反向引物

<400> 7

acacggcatc gtccttg

18

<210> 8

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 简并正向引物

<400> 8

ccccctggac ggtgaaygcn twyta

25

<210> 9

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 反向引物

<400> 9

aatccgttca cgttctgttc gtctgcc

27

<210> 10

<211> 24

<212> DNA

<213>人工序列

<220>

<223>人工序列的描述: 简并正向引物

<400> 10

cctggaggag ctgvhntgga tgra

24

<210> 11

<211> 27

<212> DNA

<213>人工序列

<220>

<223>人工序列的描述: 衔接头引物

<400> 11

ccatcctaatac gactcact atagggc

27

<210> 12

<211> 23

<212> DNA

<213>人工序列

<220>

<223>人工序列的描述: 引物

<400> 12

gtccttgcca ccctctgcca tcc

23

<210> 13

<211> 26

<212> DNA

<213>人工序列

<220>

<223>人工序列的描述: 引物

<400> 13

accacccccg ccccgatgat ccagag

26

<210> 14

<211> 28

<212> DNA

<213>人工序列

<220>

<223>人工序列的描述: 特异反义引物/反向寡核苷酸引物

<400> 14

acagccggct agcaaggcgt ggcagctg

28

<210> 15

<211> 28

<212> DNA	
<213>人工序列	
<220>	
<223>人工序列的描述: 特异反义引物	
<400> 15	
acgacagccg gctagcaagg cgtggcag	28
<210> 16	
<211> 24	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列的描述: 简并正向引物	
<400> 16	
ggnctsatgg tncstctstct sctc	24
<210> 17	
<211> 22	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列的描述: 特异正向引物	
<400> 17	
ctcctgagtg agcaaagggt cc	22
<210> 18	
<211> 24	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列的描述: 特异反向引物	
<400> 18	
gcaaactggt agaagtcgtc acac	24
<210> 19	
<211> 40	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列的描述: 引物	
<400> 19	
gacaaggcct attatgccga gatcgtgctg cagccgctcg	40

<210> 20
 <211> 40
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 引物
 <400> 20
 aaggccagca acagggcccc cgagcggctg cagcacgatc 40

<210> 21
 <211> 40
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 引物
 <400> 21
 .ggggccctgt tgctggcctt gctgcttcaa gcctccatgg 40

<210> 22
 <211> 40
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 简并引物
 <400> 22
 gtgagaaccg ccacgcactt ccatggaggc ttgaagcagc 40

<210> 23
 <211> 40
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 引物
 <400> 23
 aagtgcgtgg cggttctcac catcaccacc atcacagcga 40

<210> 24
 <211> 40
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 引物
 <400> 24
 agccaggggt ggtgcagacc tcgctgtgat ggtggtgatg 40

<210> 25
 <211> 40
 <212> DNA
 <213> 人工序列

 <220>
 <223> 人工序列的描述: 引物

 <400> 25
 ggtctgcacc acccctggct gcgtgatagc agctgccagg 40

 <210> 26
 <211> 40
 <212> DNA
 <213> 人工序列

 <220>
 <223> 人工序列的描述: 引物

 <400> 26
 ggggccatgt tctggaggat cctggcagct gctatcacgc 40

 <210> 27
 <211> 16
 <212> DNA
 <213> 人工序列

 <220>
 <223> 人工序列的描述: 引物

 <400> 27
 gacaaggcct attatg 16

 <210> 28
 <211> 14
 <212> DNA
 <213> 人工序列

 <220>
 <223> 人工序列的描述: 引物

 <400> 28
 ggggccatgt tctg 14

 <210> 29
 <211> 14
 <212> DNA
 <213> 人工序列

 <220>
 <223> 人工序列的描述: 引物

 <400> 29
 agcgaggctc gcac 14

 <210> 30

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 引物

<400> 30

gtagatgatg tgccg

15

<210> 31

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 引物

<400> 31

gcactagtct tggctaccac acgcggcatc gctccttg

38

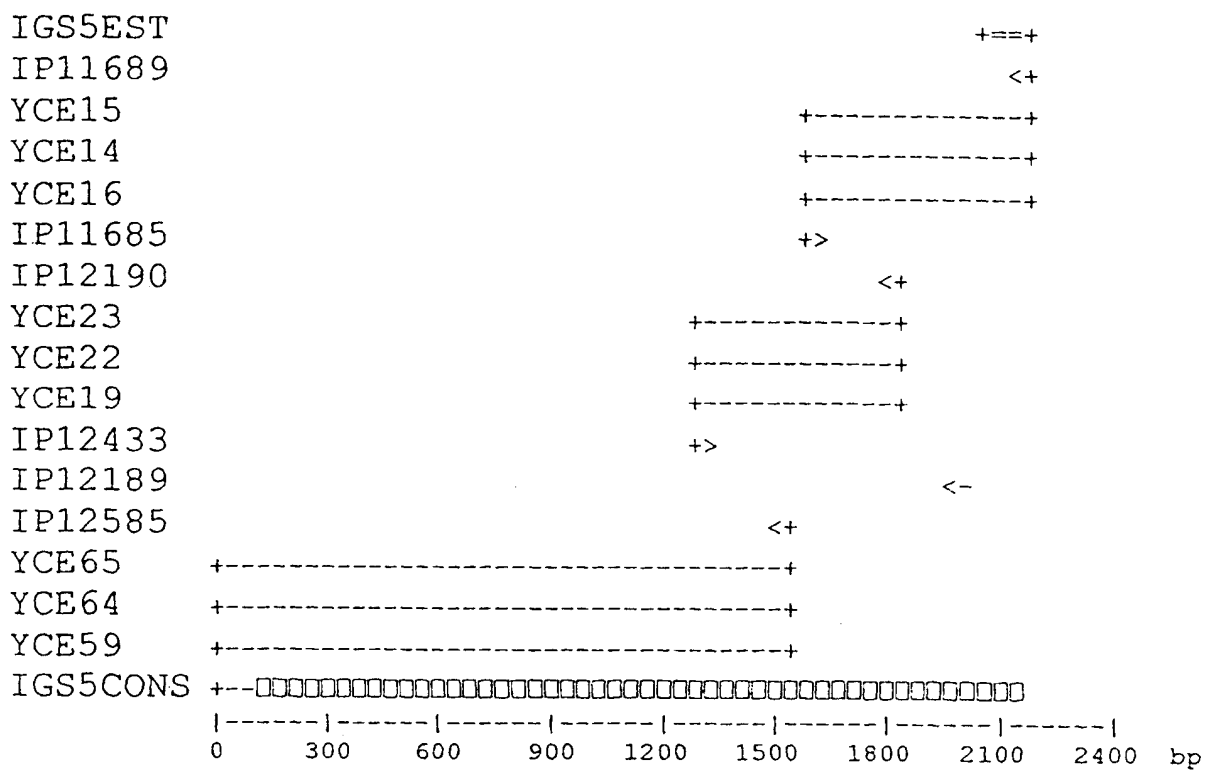


图 1

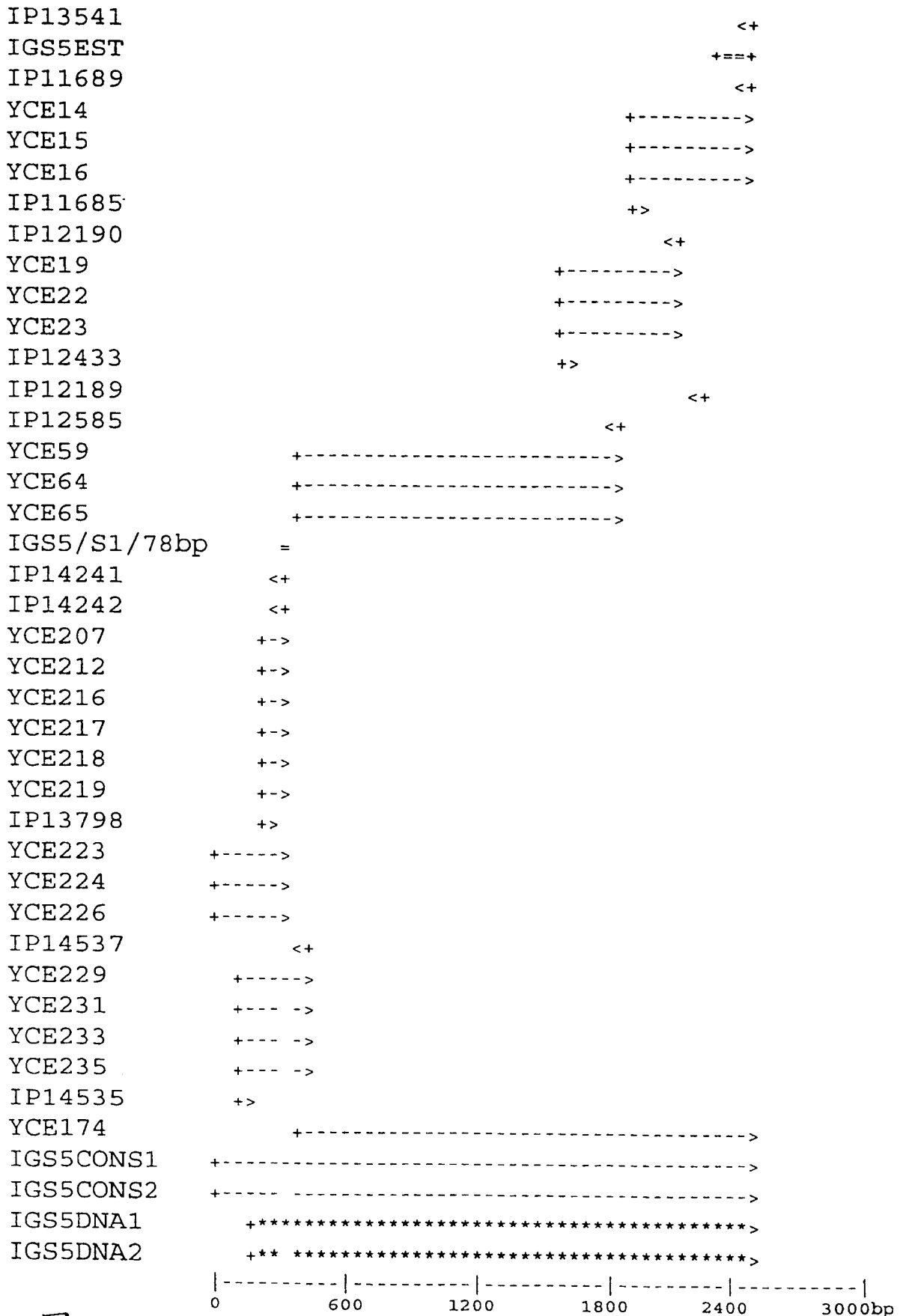


图 2

图 3A

	1	2	3	4	5	6	7	8
A	全脑	扁桃体	尾状核	小脑	大脑皮层	额叶	海马	延髓
B	枕叶	豆状核	黑质	颞叶	丘脑	伏核	脊髓	
C	心脏	主动脉	骨骼肌	结肠	膀胱	子宫	前列腺	胃
D	睾丸	卵巢	胰	垂体腺	肾上腺	甲状腺	唾液腺	乳腺
E	肾	肝	小肠	脾	胸腺	外周 淋巴细胞	淋巴结	骨髓
F	阑尾	肺	气管	胎盘				
G	胎脑	胎心	胎肾	胎肝	胎脾	胎胸腺	胎肺	
H	酵母 总RNA 100ng	酵母 tRNA 100ng	大肠杆菌 rRNA 100ng	大肠杆菌 DNA 100ng	Poly r[A] 100ng	人 Got-1DNA 100ng	人 DNA 100ng	人 DNA 500ng

图 3B

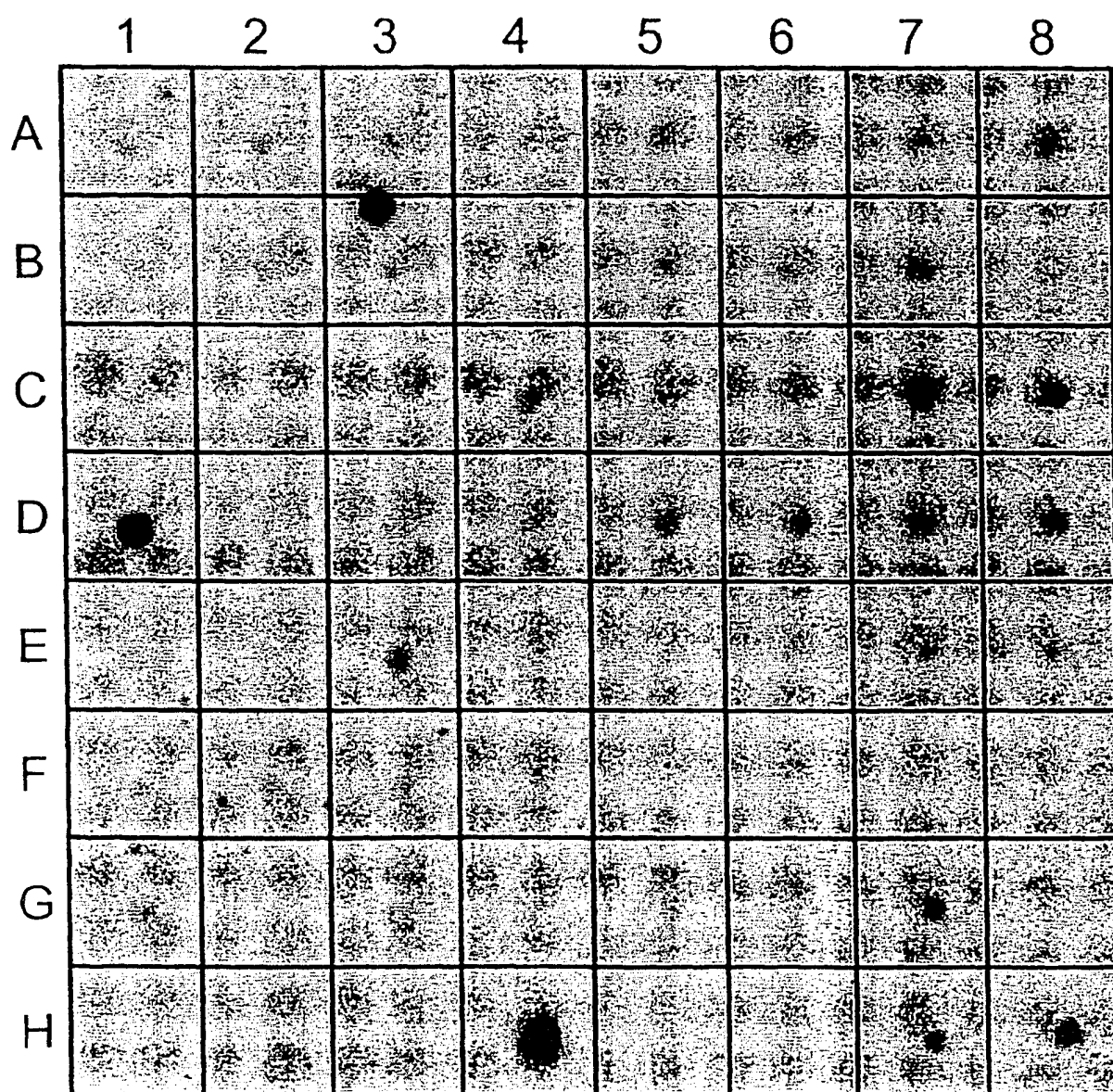
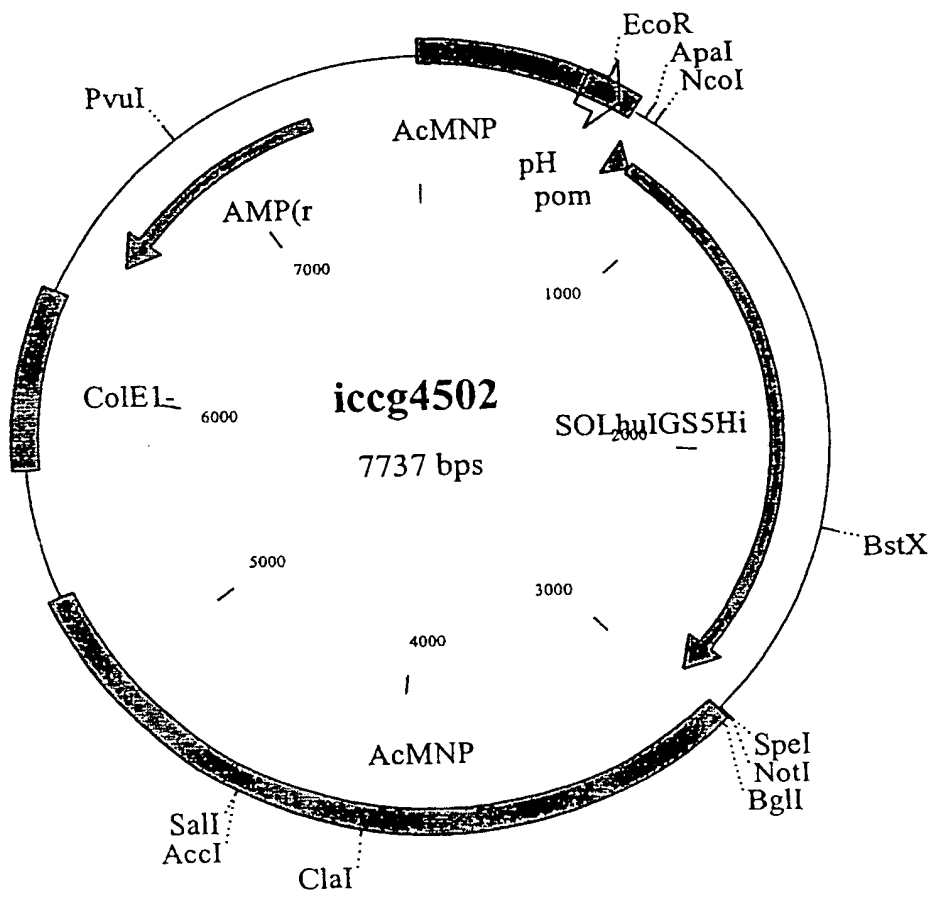


图 5



GSHHHHHHSEVCTTPGCVIAAARILQNMDPTTEPCDDFYQFACGGWLRRHVIPETNS
RYSIFDVL RDELEVILKAVLE**N**STAKDRPAVEKARTLYRSCM**N**OSVIEKRGSQPLLD
ILEVVGGWPV**A**MDRW**N**ETVGLEWELERQLALMNSQFNRRVLIDLFIWNDDO**N**SSRHI
IYIDQPTLGMP SREYYFNGG SNRKVREAYLQFMVSVATLLREDANLPRDSCLVQEDM
MQVLELETQLAKATVPQEERHDVIALYHRMGLEELQSQFGLKGF**N**WTLFIQTVLSSV
KIKLLPDEEVVYGI PYLQNL ENI IDTYSARTIQNYLVWRLVLD RIGSLSQRFKDTR
VNYRKALFGTMVEEVRWRECVGYVNSNMENAVGSLYVREAFPGDSKSMVRELIDKVR
TVFVETLDELGWMDEESKKAQEKAMSIREQIGHPDYILEEMNRRLD E EY**S**NL**N**FSE
DLYFENSLQNLKVG AQRSRKLREKVDPNLWII GA AVVNAFYSPNRNQIVFPAGILQ
PPFFSKEQPQALNFGGIGMVI GHEITHGFDDNGRNF DKNGNMMDWWS**N**FSTOHFREQ
SECMIYOY**G**NYSWDLADEQNVNGFNTLGENIADNGGVRQAYKAYLKWMAEGGKDQQL
PGLDLTHEQLFFINYAQVWCGSYRPEFAIQSIKTDVHSPLKYRVLGSLQNLAAFADT
FHCARGTPMHPKERCRVW

图 6

图 7

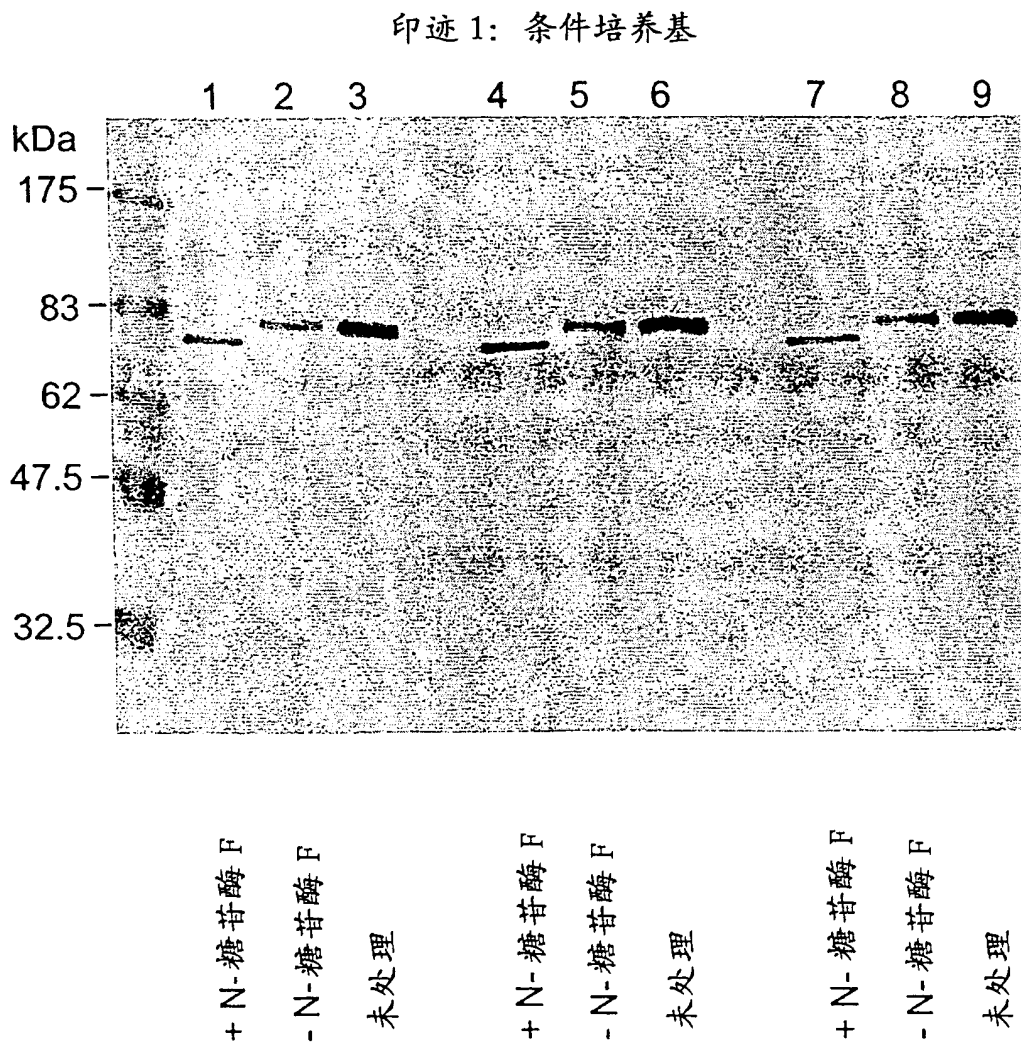
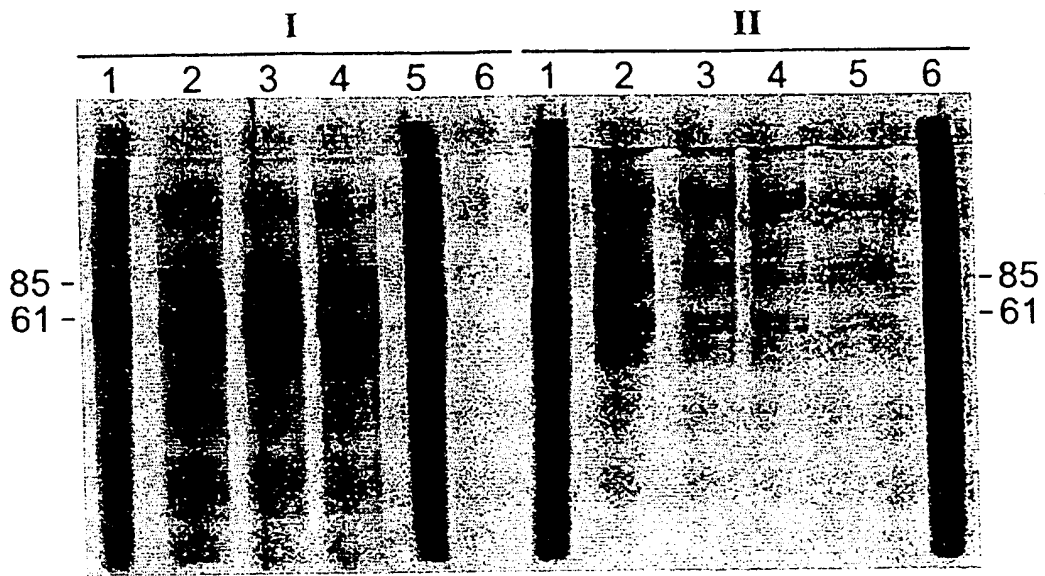
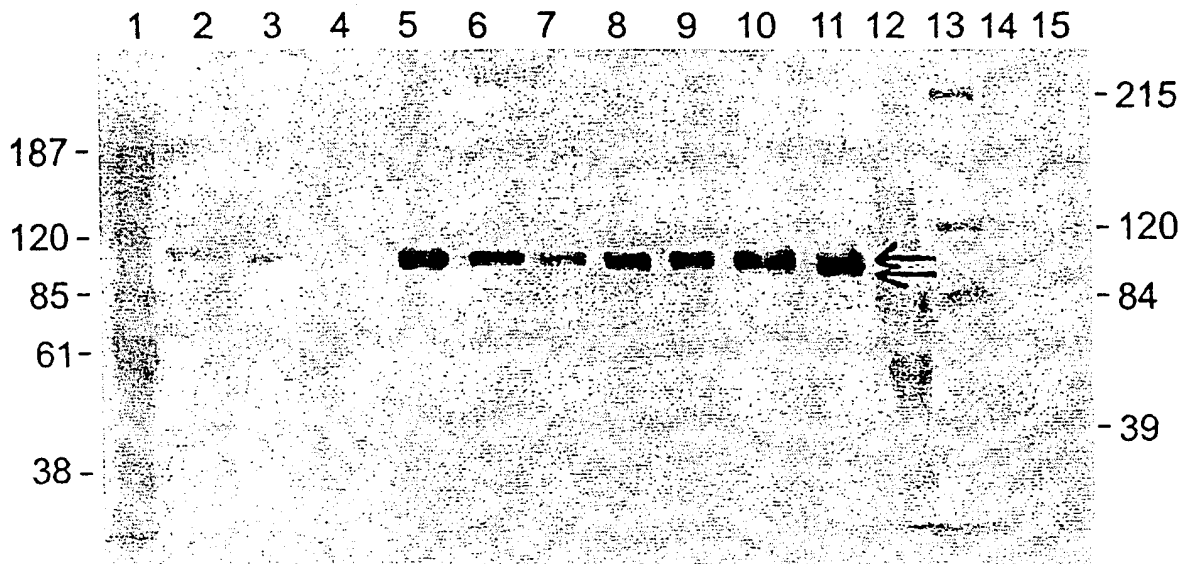


图 8



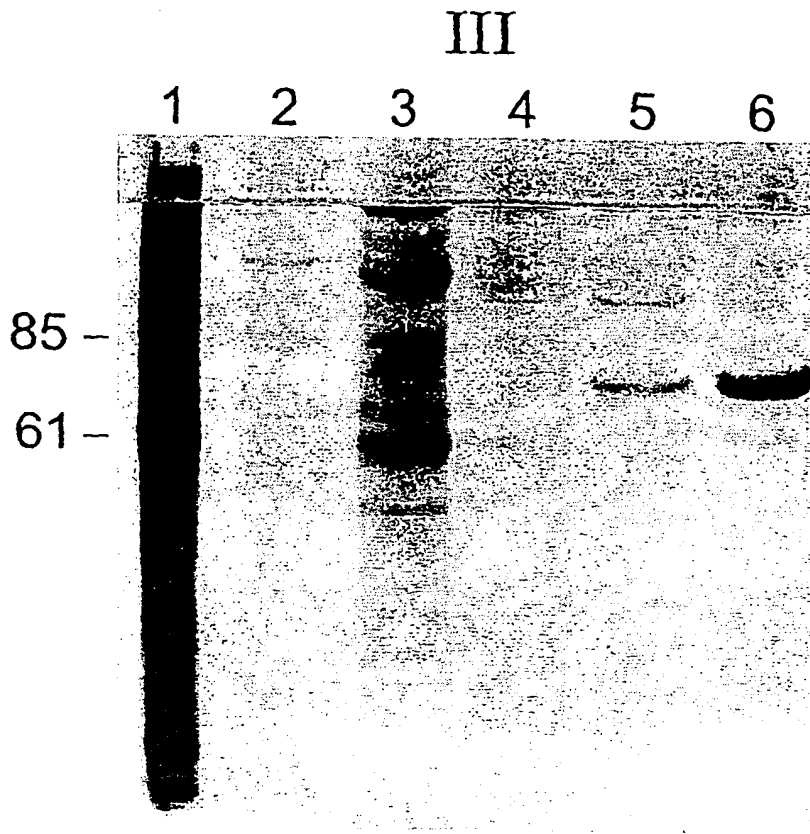
泳道	样品	DTT
I.1	预染标准	+
2	预处理前起始	+
3	预处理后起始	+
4	流过液	+
5	预染标准	+
6	/	
II.1	预染标准	+
2	1号合并液 500mM MMP 洗脱液	+
3	2号合并液"	+
4	3号合并液"	+
5	4号合并液"	+
6	预染标准	+

图 9



泳道	样品	DTT
1	预染标准 (Gibco)	+
2	起点初制	+
3	起始 Lentil	+
4	流通 Lentil	+
5	500mM MMP 洗脱物 (合并液 1)	+
6	500mM MMP 洗脱物 (合并液 2)	+
7	500mM MMP 洗脱物 (合并液 3)	+
8	200mM 咪唑 IMAC 洗脱物 (<合并液 1)	+
9	200mM 咪唑 IMAC 洗脱物 (<合并液 2)	+
10	透析后 200mM 合并液	+
11	透析后 200mM 合并液	-
12	预染标准 (Gibco)	+
13	预染标准 (Pierce)	+

图 10



泳道	样品	DTT
III.1	预染标准	+
2	流过液	+
3	20mM 合并液	+
4	50mM 合并液	+
5	100mM 合并液	+
6	200mM 合并液	+