[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
C12N 9/50
C12N 9/82 C12N 15/63
C12P 21/00

[12] 发明专利申请公开说明书

[21] 申请号 200310119663.7

[43] 公开日 2004 年 6 月 9 日

[11] 公开号 CN 1502692A

[22] 申请日 1999.9.23
[21] 申请号 200310119663.7
分案原申请号 99811202.X
[30] 优先权
[32] 1998. 9.24 [33] US [31] 60/101,594
[71] 申请人 法玛西雅厄普约翰美国公司
地址 美国密执安州
[72] 发明人 M·E·格尼
M·J·比安科沃斯基
R·L·希里克森 L·A·帕罗迪
R·严
[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 李 瑛
权利要求书 1 页 说明书 115 页 附图 18 页

[54] 发明名称 阿尔茨海默氏疾病分泌酶

[57] 摘要

本发明提供了阿尔茨海默氏疾病分泌酶以及切割 APP 蛋白的 β 分泌酶切割位点的酶促过程，及相关的核酸，肽，载体，细胞，细胞分离物及测定方法。

ISSN1008-4274

1. 一种分离或纯化的多肽，包含选自下列组的氨基酸序列:

（a）图 2（SEQ ID NO: 6）中所列举的氨基酸序列;

（b）图 3 （SEQ ID NO: 4）中所列举的氨基酸序列;

（c）显示出天冬氨酰蛋白酶活性的（a）或（b）的片段; 和

（d）除了保守取代之外具有与（a）或（b）或（c）相同的氨基酸序列的（a）－（c）的保守取代变体，并且其中的保守取代变体显示出天冬氨酰蛋白酶活性。

2. 根据权利要求 1 的分离或纯化的多肽，包含图 3（SEQ ID NO: 4）中所列举的氨基酸序列，或包含保留了天冬氨酰蛋白酶活性的其片段。

3. 根据权利要求 1 的分离或纯化的多肽，包含图 2（SEQ ID NO: 6）中所列举的氨基酸序列，或包含保留了天冬氨酰蛋白酶活性的其片段。

4. 根据权利要求 2 或 3 的分离或纯化的多肽，其中所说的片段包括 Asp2 天冬氨酰蛋白酶活性位点三肽 DTG 和 DSG，并且其中多肽显示出将 APP 加工成淀粉样蛋白β过程中所涉及的天冬氨酰蛋白酶活性。

5. 根据权利要求 1 的分离或纯化的多肽，包含图 2（SEQ ID NO: 6）或图 3（SEQ ID NO: 4）中所列举的氨基酸序列。

6. 根据权利要求 1 的分离或纯化的多肽，其中保守取代变体包含与（a）或（b）或（c）至少 95％相同的氨基酸序列。

7. 根据权利要求 1 的分离或纯化的多肽，由包括下列步骤的方法生产:

在其中的细胞可表达由核酸所编码的多肽的条件下培养用编码多肽的核酸转化成转染的宿主细胞，其中的核酸包含编码多肽的核苷酸序列，所述多肽包含图 2（SEQ ID NO: 6）或图 3（SEQ ID NO: 4）中所列举的氨基酸序列，和

分离所表达的多肽。

阿尔茨海默氏疾病分泌酶

本申请是申请号为99811202.X，申请日为1999年9月23日，发明名称为“ 阿尔茨默氏疾病分泌酶”的专利申请的分案申请。

发明所属技术领域

本发明涉及到阿尔茨海默氏疾病，APP，淀粉样蛋白β肽及人天冬酰氨蛋白酶等领域，同时也涉及到调控这些多肽活性的生物因子的鉴别方法。

背景技术

由于淀粉样蛋白蚀斑的形成，神经原纤维缠结，神经胶质过多及二乙基溴乙酰胺降低等情况的发生，阿尔茨海默氏疾病（AD）可致进行性早老性痴呆。该疾病既可遗传发生，也可随机形成，其临床特征及病理学特点都非常相似。已经发现有三个基因的突变将致常染色体上阿尔茨海默氏疾病显性结构的形成。其编码淀粉样蛋白蛋白前体(APP)及两个相关的蛋白，早老蛋白-1(presenilin-1)(PS1)及早老蛋白-2 (PS2)，其正如其名字所示那样在结构及功能上都非常相关。三个基因任一发生突变都可通过细胞内途径加速APP蛋白酶水解进程，其产生淀粉样蛋白β肽或者Aβ肽（有时侯也写为 A beta），即一种为AD中淀粉样蛋白蚀斑最基本组分的40-42氨基酸长的多肽。蛋白水解酶处理的胞内途径的调节异常可能是AD病理生理学的主要特征。在蚀斑形成的情形下，APP PS1或者PS2的突变持续加速APP的蛋白水解进程从而加快Aβ1-42的形成，Aβ1-42是Aβ肽的一种形式，具有独特的淀粉样遗传特性，因而在AD中非常重要。APP的不同形态大小不一，从695到770个氨基酸残基，位于细胞表面上，并具有单一的C端横跨膜区域。Aβ肽源于APP的一个区域，其临近并含有横跨膜区域的一部分。正常情况下，APP在α-分泌酶位点的处理过程将靠近膜的Aβ序列的中间区域分开，并释放出来自于细胞表面的APP的可溶的

细胞外区域。这些α-分泌酶的APP处理过程，产生了可溶性的APP-α，其为正常现象，并未被认为导致AD的发生。

APP在β-及γ-分泌酶位点上的病理学处理过程产生了与其在α-分泌酶位点上处理的不同的结果。在β-及γ-分泌酶位点上的连续处理释放出一种可能在AD发病机理上非常重要的Aβ肽，细胞表面上的APP再内在化（在所有的细胞中）后，β-及γ-分泌酶位点上的处理可能同时发生在内质网（在神经元中）及红细胞浆质/溶酶体途径中。尽管经过艰苦的努力，10年或者更长的时间，来鉴别负责在β-及γ-分泌酶位点上的处理APP以产生Aβ肽的酶，直到本发明前依然未知。这里，我们首次报道了β-分泌酶的鉴别及并对其特征进行了描述。我们也揭示了一些已知的及一些新的人类天冬氨酸蛋白酶，其可以如β-分泌酶一样产生作用，且首次解释了这些蛋白酶在AD中所发挥的作用。我们也描述了这些蛋白酶中有着独特功能的区域，并且首次对其作用底物的特征进行了描述。这是该种类型活性蛋白表达分离纯化的首次描述，也是该蛋白的使用，及有用细胞系及抑制剂的鉴别及生产的首次描述。

发明概述

本发明公开了asp2基因及多肽的大量突变体。

编码可以在APP的β-分泌酶切割位点进行切割的蛋白酶的任一分离或纯化的多核苷酸中含有两个或多个特殊核苷酸序列，这些特殊的核苷酸为编码100-300个氨基酸位置的核苷酸所分开。在这些位点上的氨基酸可以是任何一种氨基酸，第一组特殊的核酸含有可以编码多肽DTG的核酸，这儿第一组特殊的核酸的第一种核酸是第一种特殊的核酸，第二组特殊的核酸编码多肽DSG或者DTG，第二组核酸的最后一种核酸是最后一种特殊的核酸，在附件中并未包括揭示于SEQ ID NO.1 及 SEQ. ID NO.5 的核苷酸。权利要求1所述的核酸多聚核苷酸其中两组核酸为编码大约125-222氨基酸位点的核酸所分开，这些氨基酸可以为任何一种氨基酸。权利要求2所述的核酸多聚核苷酸其编码大约150-172氨基酸位点，这些

氨基酸可以为任何一种氨基酸。权利要求3中所述的编码172氨基酸位点的核酸多聚核苷酸，这些氨基酸可以为任何一种氨基酸。权利要求4所述的核酸多聚核苷酸其中的核苷如SEQ. ID. NO.3所示。权利要求5所述的核酸多聚核苷酸其中两组核酸为编码大约150-196氨基酸位点的核酸所分开。权利要求6所述的核酸多聚核苷酸其中两组核苷为编码大约196氨基酸（位点）的核酸所分开。权利要求7所述的核酸多聚核苷酸其中两组核酸被与分开如SEQ.ID.NO.5所示的同一套特殊核酸的同一核酸序列所分开。权利要求4所述的核酸多聚核苷酸其中两组核酸为编码大约150-190氨基酸（位点）的核酸所分开。权利要求9所述的核酸多聚核苷酸其中两组核酸为编码大约190氨基酸（位点）的核酸所分开。权利要求10所述的核酸多聚核苷酸其中两组核酸为分开如SEQ.ID.NO.1所示特殊核酸的相同的核酸所分开。权利要求1-11中第一组特殊氨基酸的第一种核酸，是第一种特殊的核酸，其被可操作性地连接到任一密码子上，其中该密码子的核酸编码任一含有1-10，000氨基酸（位点）的多肽。权利要求1-12中所述的核酸多聚核苷酸其中第一组特殊的核酸可操作性地连接到核酸聚合体上，该聚合体编码选自下述的任一多肽：任一报道蛋白或利于纯化的蛋白。权利要求1-13中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到核酸聚合体上其编码任一多肽选自于含有下列组分的组： 免疫球蛋白的重链，麦芽糖结合蛋白，谷胱甘肽 S 转染，绿色荧光蛋白，及泛素。权利要求1-14 中第二组特殊氨基酸的最后一种核酸，即最后一个特异性核酸可操作性地连接到核酸聚合体上，其编码任一含有从1到10，000个氨基酸的多肽。权利要求1-15 其中最后一种特殊的核酸可操作性地连接到任一密码子上，该密码子被连到核酸聚合体上，其编码选自下述的任一多肽报道蛋白：任一报道蛋白或利于纯化的蛋白。权利要求1-16中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到核酸聚合体上其编码任一多肽选自于含有下列组分的组： 免疫球蛋白的重链，麦芽糖结合蛋白，谷胱甘肽 S 转染，绿色荧光蛋白，及泛素。

任一分离或纯化的核酸多聚核苷酸其编码一种能够裂解APPβ-分泌

酶切割位点的蛋白酶，其含有两个或更多的特殊核酸，其中所述的特殊核酸可以为编码大约100～300个氨基酸（位点）的核酸所分开，其中这些位点上的氨基酸可以为任何一种氨基酸，这儿第一组特殊的核酸含有可以编码多肽DTG的核酸，这儿第一组特殊的核酸的第一种核酸是第一种特殊的核酸，第二组特殊的核酸编码多肽DSG或者DTG，第二组特殊核酸的最后一种核酸是最后一种特殊的核酸，这里第一种特殊的核酸可操作性地连接到编码从0到81个任一数量的氨基酸的核酸上，这里每一种密码子可以编码任一氨基酸。在权利要求18中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到编码从64到77个任一数量的氨基酸的核酸上，这里每一种密码子可以编码任一氨基酸。在权利要求19中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到编码71个氨基酸的核酸上。在权利要求20中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到71个氨基酸上，这71个氨基酸的第一种是氨基酸T。在权利要求21中所述的核酸多聚核苷酸中，多聚核苷酸含有一种序列，其至少有95%与序列SEQ. ID. (实施例 11)相似。在权利要求22中所述的核酸多聚核苷酸中，其有一种完整的含有序列SEQ. ID. (实施例 11)的多聚核苷酸。在权利要求18中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到编码从40到54个任一数量的氨基酸的核酸上，这里每一种密码子可以编码任一氨基酸。在权利要求24中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到编码47个氨基酸的核酸上。在权利要求20中所述的核酸多聚核苷酸其中第一种特殊的核酸可操作性地连接到47个氨基酸上，这47个氨基酸的第一种是氨基酸E。在权利要求21中所述的核酸多聚核苷酸中，多聚核苷酸含有一种序列，其至少有95%与序列SEQ. ID. (实施例 10)相似。在权利要求22中所述的核酸多聚核苷酸中，其有一种完整的含有序列SEQ. ID. (实施例 10)的多聚核苷酸。

任一分离或纯化的核酸多聚核苷酸其编码一种能够裂解APPβ-分泌酶切割位点的蛋白酶，其含有两个或更多的特殊核酸，其中特殊核酸可以为编码大约100～300个氨基酸（位点）的核酸所分开，其中在这些位

点上的氨基酸可以为任何一种氨基酸，这儿第一组特殊的核酸含有可以编码多肽DTG的核酸，这儿第一组特殊的核酸的第一种核酸是第一种特殊的核酸，第二组特殊的核酸编码多肽DSG或者DTG，第二组特殊核酸的最后一种核酸，即最后一种特殊的核酸，可操作性地连接到编码从50-170任一数量的密码子的核酸上。在权利要求29中所述的核酸多聚核苷酸其中最后一种特殊的核酸可操作性地连接到含有从100到170个密码子的核酸上。在权利要求30中所述的核酸多聚核苷酸其中最后一种特殊的核酸可操作性地连接到含有从142到163个密码子的核酸上。在权利要求31中所述的核酸多聚核苷酸其中最后一种特殊的核酸可操作性地连接到含有大约142个密码子的核酸上。在权利要求32中所述的核酸多聚核苷酸其中多聚核苷酸含有与SEQ. ID （实施例9或10）具有至少95%相似性的一段序列。在权利要求33中所述的核酸多聚核苷酸其中完整的多聚核苷酸含有序列SEQ. ID （实施例9或10）。在权利要求31中所述的核酸多聚核苷酸其中最后一种特殊的核酸可操作性地连接到含有大约163个密码子的核酸上。在权利要求35中所述的核酸多聚核苷酸其中多聚核苷酸含有一种序列其与SEQ. ID （实施例9或10）有至少95%的相似性。在权利要求36中所述的核酸多聚核苷酸其中完整的多聚核苷酸含有序列SEQ. ID （实施例9或10）。在权利要求31中所述的核酸多聚核苷酸其中最后一种特殊的核酸可操作性地连接到含有大约170个密码子的核酸上。权利要求1-38中编码多肽DSG的第二组特殊的核酸，并且任选地第一组核酸多聚核苷酸可操作性地连接到一种多肽纯化标记上。权利要求1-39中核酸多聚核苷酸可操作性地连接到一种六组氨酸的多肽纯化标记上。权利要求1-40中第一组特殊的核酸在一种多聚核苷酸上，第二组特殊的核酸在另一种多聚核苷酸上，其中第一种及第二个多聚核苷酸都至少有50个密码子。权利要求1-40中第一组特殊的核酸在一种多聚核苷酸上，第二组特殊的核酸在另一种多聚核苷酸上，其中第一种及第二个多聚核苷酸都至少有50个密码子，而在这里所述的多聚核苷酸都在同一种溶液中。一种载体其含有一种如权利要求1-42所述的多聚核苷酸。一种细胞或一种细胞系其含有一种如权利要求1-42所述的多聚核苷酸。

任一分离或纯化的多肽或蛋白质，其中含有一氨基酸多聚体，该多聚体是一种能够切割APP中β-分泌酶切割位点的蛋白酶，其中含有两组或多组特殊氨基酸，其中所述的特殊核氨基酸可以为大约100~300个氨基酸（位点）所分开，所述位点上的氨基酸可以为任何一种氨基酸，这儿第一组特殊的氨基酸含有多肽DTG，这儿第一组特殊的氨基酸的第一种氨基酸是第一种特殊的氨基酸，第二组特殊的氨基酸选自含有DSG或者DTG的多肽，第二组特殊氨基酸的最后一种氨基酸，即最后一种特殊的氨基酸，但SEQ ID NO.2 及 SEQ. ID NO.6所示的蛋白酶不包括在其中。权利要求45所述的氨基酸多肽其中两组氨基酸为大约125到222个氨基酸位点所分开，这里的每一个位点可以是任何一种氨基酸。权利要求46所述的氨基酸多肽其中两组氨基酸为大约150到172个氨基酸位点所分开。权利要求47所述的氨基酸多肽其中两组氨基酸为大约172个氨基酸位点所分开。权利要求48所述的氨基酸多肽其中的蛋白酶如SEQ. ID. NO.4所示的那样。权利要求46所述的氨基酸多肽其中两组氨基酸为大约150到196个氨基酸位点所分开。权利要求50所述的氨基酸多肽其中两组氨基酸为大约196个氨基酸位点所分开。权利要求51所述的氨基酸多肽其中两组氨基酸为分开SEQ. ID. NO.6所示的同组特殊氨基酸的相同的氨基酸序列所分开。权利要求46所述的氨基酸多肽其中两组氨基酸为大约150到190个氨基酸位点所分开。权利要求53所述的氨基酸多肽其中两组氨基酸为大约190个氨基酸位点所分开。权利要求54所述的氨基酸多肽其中两组核苷为分开SEQ. ID. NO.2所示的同组特殊氨基酸的相同的氨基酸序列所分开。权利要求45-55其中第一组特殊的氨基酸的第一种氨基酸是第一种特殊的氨基酸，其可操作性地连接到从1到10,000个氨基酸的任一多肽上。权利要求45-56所述的氨基酸多肽其中第一种特殊的氨基酸可操作性地连接到任一选自含有任一一种报道蛋白或利于纯化的蛋白的组的任一多肽上。权利要求45-57所述的氨基酸多肽其中第一种特殊的氨基酸可操作性地连接到任一选自含有免疫球蛋白的重链，麦芽糖结合蛋白，谷胱甘肽 S 转染，绿色荧光蛋白，及泛素的组的任一多肽上。在权利要求45-58中，其中第二组特殊氨基酸的最后一种氨基酸，也即最后一种特殊的氨基酸，

可操作性地连接到从1到10,000个氨基酸的任一多肽上。权利要求45-59所述的氨基酸多肽其中最后一种特殊的氨基酸可操作性地连接到任一选自含有任何一种报道蛋白或利于纯化的蛋白的组的任一多肽上。权利要求45-60所述的氨基酸多肽其中第一种特殊的氨基酸可操作性地连接到任一选自含有免疫球蛋白的重链，麦芽糖结合蛋白，谷胱甘肽 S 转染，绿色荧光蛋白，及泛素的组的任一多肽上。

任一分离或纯化的多肽或一种含有编码能够裂解APPβ-分泌酶切割位点的蛋白酶的氨基酸多肽的蛋白，其含有两组或更多的特殊氨基酸，其中特殊氨基酸可以为大约100~300个氨基酸（位点）所分开，其中在这些位点上的氨基酸可以为任何一种氨基酸，这儿第一组特殊的氨基酸含有氨基酸DTG，这儿第一组特殊的氨基酸的第一种氨基酸是第一种特殊的氨基酸D，第二组氨基酸为DSG或者DTG，第二组特殊氨基酸的最后一种氨基酸，即最后一种特殊的氨基酸G，这里第一种特殊的氨基酸可操作性地连接到编码从0到81个任一数量的氨基酸位点的氨基酸上，这里的每一个位点可以是任一氨基酸。权利要求62所述的氨基酸多肽，其中第一种特殊的氨基酸可操作性地连接到来自大约64到77个氨基酸位点的一种多肽上，这里的每一个位点可以是任一氨基酸。权利要求63所述的氨基酸多肽，其中第一种特殊的氨基酸可操作性地连接到来自大约77个氨基酸位点的一种多肽上。权利要求64所述的氨基酸多肽，其中第一种特殊的氨基酸可操作性地连接到71个氨基酸位点上，而且这71个氨基酸的第一种为氨基酸T。权利要求65所述的氨基酸多肽，其中的多肽含有一种序列，其与SEQ. ID. (实施例 11)具有至少95%的相似性。权利要求66所述的氨基酸多肽，其中的完整多肽含有序列SEQ. ID. (实施例 11)。权利要求62所述的氨基酸多肽，其中第一种特殊的氨基酸可操作性地连接到来自大约40到54个氨基酸位点的任何一种多肽上，这里的每一个位点可以是任一氨基酸。权利要求68所述的氨基酸多肽，其中第一种特殊的的氨基酸可操作性地连接到编码47个氨基酸的多肽的氨基酸上。权利要求69所述的氨基酸多肽，其中第一种特殊的的氨基酸可操作性地连接到一种47个氨基酸的多肽上，其中这47个氨基酸的第一种是氨基酸E。权利要求70所

述的氨基酸多肽，其中的多肽含有一种序列其与SEQ. ID.（实施例10)有至少95%的相似性。该氨基酸多肽其中的多肽含有实施例10）。

任一分离或纯化的多肽其是一种能够裂解APPβ-分泌酶切割位点的蛋白酶，其含有两组或更多的特殊氨基酸，其中特殊氨基酸可以为大约100~300个氨基酸（位点）所分开，其中在这些位点上的氨基酸可以为任何一种氨基酸，这儿第一组特殊的氨基酸含有编码DTG的氨基酸，这儿第一组特殊的氨基酸的第一种氨基酸是第一种特殊的氨基酸D，第二组氨基酸为DSG或者DTG，第二组特殊氨基酸的最后一种氨基酸，即最后一种特殊的氨基酸G，其可操作性地连接到从50到170的任一数目的氨基酸上，该氨基酸可以是任意一种氨基酸。权利要求73所述的氨基酸多肽其中最后一种特殊的氨基酸可操作性地连接到一种大约100~170个氨基酸的多肽上。权利要求74所述的氨基酸多肽其中最后一种特殊的氨基酸可操作性地连接到一种大约142~163个氨基酸的多肽上。权利要求75所述的氨基酸多肽其中最后一种特殊的氨基酸可操作性地连接到一种大约142个氨基酸的多肽上。权利要求76所述的氨基酸多肽，其中的多肽含有一种序列其与SEQ. ID.（实施例9或10）有至少95%的相似性。权利要求75所述的氨基酸多肽其中最后一种特殊的氨基酸可操作性地连接到一种大约163个氨基酸的多肽上。权利要求79所述的氨基酸多肽，其中的多肽含有一种序列其与SEQ. ID.（实施例9或10）有至少95%的相似性。权利要求79所述的氨基酸多肽，其中的完整的多肽含有一种序列SEQ. ID.（实施例9或10）。权利要求74所述的氨基酸多肽其中最后一种特殊的氨基酸可操作性地连接到一种大约170个氨基酸的多肽上。权利要求46-81其中第二组特殊的氨基酸由具有氨基酸序列DSG的多肽组成。权利要求45-82 其中的氨基酸多肽可操作性地连接到一种多肽纯化标记上。权利要求45-83其中的氨基酸多肽可操作性地连接到一种为六个组氨酸的多肽纯化标记上。权利要求45-84其中第一组特殊的氨基酸在一种多肽上，而第二组特殊的氨基酸在另一种多肽上，而第一种及第二个多肽都至少有50个氨基酸，所述氨基酸可以是任意一种氨基酸。权利要求45-84其中第一组特殊的氨基酸在一种多肽上，而第二组特殊的氨基酸在另一种多肽上，而第

一种及第二个多肽都至少有50个氨基酸，且所述多肽在同一溶液中。一种载体其含有一种如权利要求45-86所述的多聚核苷酸。一种细胞或一种细胞系其含有一种如权利要求45-87所述的多聚核苷酸。制备权利要求1-44所述的任一多聚核苷酸，载体，或细胞系的方法。制造权利要求45-88所述的任一多聚核苷酸，载体，或细胞系的方法。利用权利要求89和90中所述方法制备而成的如权利要求1-88中所述的任一种多聚核苷酸、多肽、载体、细胞或细胞系。

任一分离或纯化的多肽或蛋白质，其中含有一氨基酸多肽，该氨基酸多肽编码一种能够裂解APPβ-分泌酶切割位点的蛋白酶，所述蛋白酶中含有两组或更多的特殊氨基酸，其中特殊氨基酸可以为大约100～300个氨基酸（位点）所分开，其中在这些位点上的氨基酸可以为任何一种氨基酸，这儿第一组特殊的氨基酸含有氨基酸DTG，这儿第一组特殊的氨基酸的第一种氨基酸是第一种特殊的氨基酸D，第二组氨基酸为DSG或者DTG，第二组特殊氨基酸的最后一种氨基酸，即最后一种特殊的氨基酸G，这里第一种特殊的氨基酸可操作性地连接到0-81个氨基酸位置上任一数目的氨基酸上，这里的每一个位点可以是任一氨基酸。

权利要求62所述的氨基酸多肽，其中第一种特殊的氨基酸可操作地连接到一长约30～77个氨基酸位点的多肽上，这里的每一个位点可以是任一氨基酸。权利要求63所述的氨基酸多肽，其中第一种特殊的氨基酸可操作性地连接到一长为35，47，71或77个氨基酸位点的多肽上。

权利要求63所述的氨基酸多肽，其中第一种特殊的氨基酸可操作性连接到来自SEQ. ID. NO.3长度为35、47、 71、 或77的一组相应多肽上，从第一种特殊序列DTG的氨基酸上开始，向SEQ. ID. NO.3的N端方向计数。

权利要求65所述的氨基酸多肽，其中的多肽含有一种序列，其与相同的SEQ. ID. NO.4所示相应的氨基酸具有至少95%的相似，也就是说，与SEQ. ID. NO.4所示的部分序列相同。包括所有从第一和/或第二个特殊核苷酸开始朝向N末端，经过并包括第一段特殊氨基酸之前的71，47，35个氨基酸（实施例10及11）的所有序列。

权利要求65所述的氨基酸多肽，其中的完整多肽含有一种71个氨基

酸的多肽，其中的第一种氨基酸为T，第二个为Q。权利要求21所述的核酸多聚核苷酸，其中的多聚核苷酸含有一段与SEQ. ID. NQ. 3所示氨基酸有着至少95%同源性的序列。也就是说，与SEQ. ID. NO.3所示的序列相同，其中包括从第一和/或者第二个特殊核苷酸朝向N末端，经过并包括71个氨基酸的所有序列，参见实施例10，开始于DTG位点且包括编码71个氨基酸的核苷。

在权利要求22中所述的核酸多聚核苷酸中，其有所述的完整的多聚核苷酸含有与序列SEQ. ID. . NO.3 所示氨基酸相同的序列。也就是说，与SEQ. ID. NO.3中的序列相同，其中包括从第一和/或者第二个特殊核苷酸开始朝向N末端，经过并包括71个氨基酸的所有序列，参见实施例10，开始于DTG位点且包括编码71个氨基酸的核苷。

在权利要求18中所述的核酸多聚核苷酸中，其中第一组特殊的核酸可操作性地连接到编码任一从30到54数目的氨基酸的核酸上，其中的每一种密码子可以编码任一氨基酸。

在权利要求20中所述的核酸多聚核苷酸中，其中的第一组特殊的核酸可操作性地连接到47个密码子上，其中这些35或47个氨基酸的第一种为氨基酸E或G。

权利要求21所述的核酸多聚核苷酸，其中的多聚核苷酸含有一种序列其与SEQ. ID. NQ. 3所示氨基酸有着至少95%的同源性。也就是说，与SEQ. ID. NO.3所示的部分序列相似，包括所有从或者第一或者第二个特殊核苷酸直到N末端，经过及包括35或47个氨基酸，参见实施例11的47个例子，开始于DTG位点且包括在DTG位点前编码35或47个氨基酸的核苷。权利要求22所述的核酸多聚核苷酸，其中的多聚核苷酸具有一定的与SEQ. ID. NO.3所示相同或相应的氨基酸的同源性。也就是说，与SEQ. ID. NO.3所示的序列相似，包括所有从或者第一或者第二个特殊核苷酸直到N末端，经过及包括35或47个氨基酸，参见实施例11的47个例子，开始于DTG位点且包括在DTG位点前编码35或47个氨基酸的核苷。

一种分离的核酸分子，其中包括一多聚核苷酸，该多聚核苷酸编码一Hu-Asp 多肽，其核苷酸序列与选自下列的序列有至少95%的同源性:

(a)编码一种Hu-Asp 多肽的一段核苷酸序列，该多肽选自：Hu-Asp 1，Hu-Asp2(a)，及 Hu-Asp2(b)，其中所述Hu-Asp 1，Hu-Asp2(a)，及 Hu-Asp2(b) 多肽分别具有如SEQ ID No.2，SEQ ID No.4,及 SEQ ID No.6所示的完全的氨基酸序列，及

(b)与(a)中核苷酸序列互补的核苷酸序列。

权利要求92所述的核酸分子，其中所述Hu-Asp 多肽为Hu-Asp 1，所述的1(a)多聚核苷酸包括有SEQ ID No.1所示的核苷酸序列。权利要求92所述的核酸分子，其中所述Hu-Asp 多肽为Hu-Asp 2(a)，所述的1(a)多聚核苷酸包括有SEQ ID No.4所示的核苷酸序列。权利要求92所述的核酸分子，其中所述Hu-Asp 多肽为Hu-Asp 2(b)，所述的1(a)多聚核苷酸包括有SEQ ID No.5所示的核苷酸序列。一种分离的核酸分子包括一种多聚核苷酸，该多聚核苷酸在严格条件下与具有如权利要求92的(a)或(b)的核苷酸序列的多聚核苷酸进行杂交。一种含有权利要求96所示核酸分子的载体。在权利要求97所述的载体中，其中所述核酸分子可操作性地连接到表达Hu-Asp 多肽的启动子上。在权利要求98所述的载体中，其中所述Hu-Asp 多肽为Hu-Asp1。在权利要求98所述的载体中，其中所述Hu-Asp多肽为Hu-Asp2(a)。在权利要求98所述的载体中，其中所述Hu-Asp 多肽为Hu-Asp2(b)。一种含有权利要求98所述的载体的宿主细胞。一种获取Hu-Asp 多肽的方法，其包括培养权利要求102所述的宿主细胞，且分离出所述的Hu-Asp 多肽。一种分离的Hu-Asp1 多肽，其包括一种含有与如SEQ ID No.2所示的氨基酸序列有至少95%相似性的氨基酸序列。一种分离的Hu-Asp2(a) 多肽，其包括一种含有与如SEQ ID No.4所示的氨基酸序列有至少95%相似性的氨基酸序列。一种分离的Hu-Asp12(b) 多肽，其包括一种含有与如SEQ ID No.8所示的氨基酸序列有至少95%相似性的氨基酸序列。一种分离的抗体，其可与权利要求104-107任一所述的Hu-Asp多肽特异性地结合。

本发明中公开了多种分析该种酶的方法。

一种鉴别可用于筛选β分泌酶活性抑制剂的细胞的方法，该方法包括下述步骤:

(a) 鉴别一种细胞，该细胞能表达一种能够在APPβ分泌酶位点进行切割的蛋白酶：包括以下步骤：

i) 收集细胞或待鉴定细胞的上清液；

ii)测量核心多肽的产量，其中该核心多肽选自APPC-末端肽或者可溶性的APP；

iii)选择产生核心多肽的细胞。

权利要求108所述的方法其中细胞被收集且核心多肽为用β分泌酶进行裂解生成的APPC末端肽。权利要求108所述的方法其中上清液被收集且核心多肽为可溶性的APP，其中所述可溶的APP 具有一种因β分泌酶裂解而生成的C末端。权利要求108所述的方法，其中所述细胞含有任一种如权利要求1-86所述的核酸或多肽且这里细胞被显示可以裂解含有下列多肽结构的任一多肽中的β分泌酶位点：P2、P1、P1'、 P2'其中P2为K或N，P1为M或L，P1'为D，P2'为A的β分泌酶位点。权利要求111所述的方法，其中P2为K，P1为M。权利要求112所述的方法，其中P2为N，P1为L。

任一细菌细胞，其中含有如权利要求1-86及92-107所述的任一种核酸或多肽。权利要求114所述的细菌细胞，其中所述细菌为大肠杆菌。含有如权利要求1-86及92-107所述的任一种核酸或多肽的任一真核细胞。

含有如权利要求1-86及92-107所述的任一种核酸或多肽的任一昆虫细胞。一种如权利要求117所述的昆虫细胞，其中所述昆虫为sf9,或者High5。一种如权利要求100所述的昆虫细胞，其中的昆虫细胞为High5。任一哺乳动物细胞含有如权利要求1-86及92-107所述的任何的核酸或多肽。一种权利要求120所述的哺乳动物细胞，其中所述的哺乳动物细胞选自下列类型的细胞：人，啮齿类，兔类及灵长类。一种如权利要求121所述的哺乳动物细胞其中的哺乳动物细胞选自人类细胞。一种如权利要求122所述的哺乳动物细胞其中的人类细胞选自HEK293及IMR-32。一种如权利要求121所述的哺乳动物细胞其中的细胞为灵长类动物细胞。权利要求124所述的灵长类动物细胞其中的灵长类动物细胞为COS-7细胞。一种如权利要求121所述的哺乳动物细胞其中的细胞选自一种啮齿类动物细胞。权利要求126所述的一种啮齿类动物细胞选自CHO-K1，Neuro-2A，3T3 细

胞。一种如权利要求115所述的酵母细胞。一种如权利要求115所述的鸟类细胞。

APP的任一同工型，其中最后两个羧基端氨基酸都为赖氨酸残基。在说明书中，所定义的同工型为任一APP多肽，包括APP变体（包括突变体），及存在于人体中的APP片段，如US 5,766,846，第7栏，第 45-67行（其作为参考资料引入本文件）所述的APP片段，在此引入作为参考。 权利要求114所述的APP同工型，包括被称为APP695的同工型，其经修饰从而使最后两个具有赖氨酸残基成为其最后两个羧基端氨基酸。权利要求130所述的APP同工型，包括SEQ. ID. 16。权利要求130所述的APP同工型变体，包括SEQ. ID. 18,及20。任一真核细胞系，包括如权利要求130-132所述的核酸及多肽。任一如权利要求133所述的细胞系，其为哺乳动物细胞系（HEK293，Neuro2a，及其它）。一种用于鉴别针对能够裂解APP β-分泌酶切割位点的酶的抑制剂的方法，包括下列步骤:

a）在一种培养基中培养细胞，采用能够实现下述目的的条件：酶可以对APP进行处理，及释放淀粉样蛋白β肽到培养基中，并可使得APP的CTF99片段聚集于细胞溶胞产物中;

b）将已培养的细胞暴露于一种待测组合物；并且特异性检测该待测组合物是否抑制了酶的功能，方法是测试释放进培养基中的淀粉样蛋白β肽或者细胞溶胞产物中的APP的CTF99片段的数量;

c）测定待测组合物减少了存在于培养基中的可溶性淀粉样蛋白β肽的数量，以及细胞溶胞产物中的作为Asp2抑制剂的APP的CTF99片段的减少。

如权利要求135所述的方法，其中所培养的细胞是人类，啮齿类或昆虫细胞系。如权利要求136所述的方法，其中的人类或啮齿类细胞系呈现出了β分泌酶活性，因而发生了APP的处理，并伴有淀粉样蛋白β肽释放入培养基以及CTF99片段在细胞溶胞产物中的聚集。如权利要求137所述的方法，用直接针对具有β分泌酶活性的酶的抑制剂处理人或啮齿类细胞，减少了可溶性淀粉样蛋白β肽释放进培养基中的数量及CTF99在细胞溶胞产物中的聚集。鉴别一种试剂的方法，该试剂可以减少选自Hu-Asp 1，Hu-

Asp2(a)，及 Hu-Asp2(b)的Hu-Asp多肽的活性，该方法包括:

a) 在试剂存在及不存在的情况下测定所述Hu-Asp多肽的活性，及

b) 比较所述试剂存在及不存在的情况下所测定的Hu-Asp多肽的活性;

试剂存在的情况下所述Hu-Asp多肽的活性比试剂不存在情况下要低，这表明该试剂减少了所述Hu-Asp多肽的活性。本发明所述的核酸、多肽、蛋白、载体、细胞及细胞系以及测试方法。

本发明所提供了分离的核酸分子，其包括一编码一种多肽多聚核苷酸，所述多肽选自人类天冬氨酰蛋白酶。具体地为人类天冬氨酰蛋白酶1(Hu-Asp1) 及人类天冬氨酰蛋白酶2 (Hu-Asp2)的两个可选择的剪接变体(Hu-Asp2(a)及Hu-Asp2 (b))。正如这儿所使用的，所有涉及的"Hu-Asp"应该被理解为指的是所有的Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)。此外，正如这儿所使用的,所有涉及的"Hu-Asp2"应该被理解为指的是Hu-Asp2(a)及Hu-Asp2 (b)。Hu-Asp1在胰脏及前列腺组织表达量最高，而Hu-Asp2(a)及Hu-Asp2 (b) 在胰脏及脑组织中的表达量最高。本发明同时也提供了分离的Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)多肽及显示出天冬氨酰蛋白酶活性的片段。

在一种优选的实施例当中,核酸分子含有一个多聚核苷酸，其核苷酸序列选自SEQ ID NO:1所示的1-1554残基，其编码Hu-Asp1，SEQ ID NQ:3的1-1503残基，其编码Hu-Asp2(a)，SEQ ID NQ:5的1-1428残基，其编码Hu-Asp2 (b)。在另一方面，本发明也提供了一种分离的核酸分子其包括一种多聚核苷酸，其在严格条件下能够与编码Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)或者其中片段的多聚核苷酸杂交。欧洲专利申请EP 0 848 062公开了一种多肽Asp 1，其与Hu-Asp1基本上同源，而国际申请WO 98/22597揭示了一种多肽Asp 2，其与Hu-Asp2(a)具有很大的同源性。

本发明还提供了含有本发明所述分离的核酸分子的载体，引入该载体的宿主细胞，以及通过培养上述宿主细胞及分离相关多肽以获取Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)多肽的重组方法。

另一方面，本发明同时也提供了分离的Hu-Asp1，Hu-Asp2(a)，及

Hu-Asp2(b) 多肽及其片段。在优选的实施例当中，Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)多肽分别具有如SEQ ID NO:2，SEQ ID NO:4，或 SEQ ID NO:6所示的氨基酸序列。本发明同时也描述了Hu-Asp2(a)的活性结构，制备该活性结构的方法，制备可溶性结构的方法，测量Hu-Asp2活性的方法，及Hu-Asp2作用的底物。本发明同时也描述了靶向Hu-Asp1，Hu-Asp2(a)，及Hu-Asp2(b) mRNA 转录物的反义低聚体及该反义低聚体用于减少该mRNA及其相应多肽的产量的用途。分离的多克隆或单克隆抗体，都可特异性连接到本发明所提供的Hu-Asp1，Hu-Asp2(a)，及Hu-Asp2(b) 中的等任一多肽上。

本发明同时也提供了一种方法，其可以鉴别用于调控Hu-Asp1，Hu-Asp2(a)，及Hu-Asp2(b)任一之活性的试剂。本发明描述了一种方法，其可以通过加入Hu-Asp2多肽，无细胞测试该种试剂，以及在Hu-Asp2存在的条件下在人体或其它哺乳类动物细胞中测试该试剂的方法。

所列序列的简要描述

Sequence ID No. 1-人类Asp-1， 核苷酸序列

Sequence ID No. 2-人类Asp-1，推测的氨基酸序列

Sequence ID No. 3-人类Asp-2(a),核苷酸序列

Sequence ID No.4-人类Asp-2(a)，推测的氨基酸序列

Sequence ID No.5-人类 Asp-2(b)，核苷酸序列

Sequence ID No.6-人类 Asp-2(b)，推测的氨基酸序列

Sequence ID No. 7-鼠Asp-2(a)，核苷酸序列

Sequence ID No. 8-鼠Asp-2(a)，推测的氨基酸序列

Sequence ID No.9-人类APP695，核苷酸序列

Sequence ID No.10-人类APP695，推测的氨基酸序列

Sequence ID No.1 1-人类 APP695-Sw，核苷酸序列

Sequence ID No.12-人类 APP695-Sw. 推测的氨基酸序列

Sequence ID No.13-人类 APP695-VF, 核苷酸序列

Sequence ID No. 14-人类APP695-VF, 推测的氨基酸序列

Sequence ID No.15-人类 APP695-KK, 核苷酸序列

Sequence ID No. 16-人类APP695-KK, 推测的氨基酸序列

Sequence ID No.17-人类 APP695-Sw-KK, 核苷酸序列

Sequence ID No.18-人类ApP695-Sw-KK, 推测的氨基酸序列

Sequence ID No.19-人类APP695-VF-KK, 核苷酸序列

Sequence ID No.20-人类APP695-VF-KK, 推测的氨基酸序列

Sequence ID No.21-T7-人类-pro-Asp-2(a)ΔTM, 核苷酸序列

Sequence ID No.22-T7-人类-pro-Asp-2(a) ΔTM, 氨基酸序列

Sequence ID No.23-Caspase-人类-pro-Asp-2(a) ΔTM, 核苷酸序列

Sequence ID No.24-T7-Caspase-人类-pro-Asp-2(a) ΔTM, 氨基酸序列

Sequence ID No.25-人类-pro-Asp-2(a) ΔTM (低GC含量), 核苷酸序列

Sequence ID No.26-人类-pro-Asp-2(a) ΔTM, (低GC含量), 氨基酸序列

Sequence ID No.27-T7-Caspase-Caspase 8 cleavage-人类-pro-Asp-2(a) ΔTM,核苷酸序列

Sequence ID No.28-T7-Caspase-Caspase 8 cleavage-人类-pro-Asp-2(a) ΔTM, 氨基酸序列

Sequence ID No.29-人类 Asp-2(a) ΔTM, 核苷酸序列

Sequence ID No.30-人类Asp-2(a) ΔTM, 氨基酸序列

Sequence ID No.3 1-人类Asp-2(a) ΔTM$(His)_6$, 核苷酸序列

Sequence ID No.32-人类 Asp-2(a) ΔTM$(His)_6$, 氨基酸序列

Sequence ID No.s 33-46在下述本发明的详细描述中说明。

附图的简要说明

图1: 图1显示出人Asp1的核苷酸序列 (SEQ ID NO: 1) 及推测的氨基酸序列(SEQ ID NO:2);

图2: 图2显示出人Asp2(a)的核苷酸序列 (SEQ ID NO: 3) 及推测的氨基酸序列(SEQ ID NO:4);

图3：图3 显示出人Asp2(b)的核苷酸序列 (SEQ ID NO: 5) 及推测的氨基酸序列(SEQ ID NO:6)；推测的Hu-Asp2(b)生物横跨膜区域附于括弧中；

图4：图4 显示出鼠Asp2(a)的核苷酸序列 (SEQ ID NO: 7) 及推测的氨基酸序列(SEQ ID NO:8)；

图5：图5显示出Hu-Asp2(a)及鼠Asp2(a)的推测的氨基酸序列的最匹配序列；

图6：图6显示出T7-人类-pro-Asp-2(a)ΔTM的核苷酸序列 (SEQ ID NO: 21) 及推测的氨基酸序列(SEQ ID NO:22)；

图7：图7显示出T7- caspase-人类-pro-Asp-2(a)ΔTM的核苷酸序列 (SEQ ID NO: 23) 及推测的氨基酸序列(SEQ ID NO:24)；

图8：图8显示出人类-pro-Asp-2(a)ΔTM (低GC含量) 的核苷酸序列 (SEQ ID NO: 25) 及推测的氨基酸序列(SEQ ID NO:26)；

图9：蛋白质印迹显示通过用反义低聚体靶定Hu-Asp2 Mrna 转染HEK125.3细胞可以减少CTF99的产量；

图10：蛋白质印迹实验显示：在用或不用Hu-Asp2与APP-KK共转染的鼠Neuro-2a细胞中，只有那些用HuAsp2共转染的细胞中才出现CTF99产量的增加。用HuAsp2或不用HuAsp2与APP-Sw-KK共转染的细胞中，只有那些共转染了Hu-Asp2的细胞内才出现CTF99产量的进一步增加。

图11：图11显示出人类-Asp2(a) ΔTM的推测的氨基酸序列(SEQ ID NO:30)；

图12：图11显示出人类-Asp2(a) ΔTM$(His)_6$的推测的氨基酸序列(SEQ ID NO:30)。

本发明的详细描述

本发明所使用的一些定义见下文，大多数定义为本领域普通技术人员所使用的定义。

当β淀粉样蛋白多肽，任一来自APPβ分泌酶切割位点的多肽，其包

括39，40，41，42 及 43个氨基酸的多肽，从β分泌酶切割位点向外延伸39，40，41，42及43个氨基酸。β淀粉样蛋白多肽还可为序列1-6，该序列参见美国专利US 5,750,349的SEQ. ID. NO.1-6，于1998年5月12日颁布（在此引入作为参考）。本发明公开的β分泌酶切割片段被称为CTF-99，其由β分泌酶切割位点延伸至APP的羧基端。

当APP的同工型被讨论当其被认为是任一APP多肽时，包括APP变体（包括突变体），以及存在于人体中的APP片段如美国专利US 5,766,846，第7栏，45-67行所描述的APP 片段，在此引入作为参考。

这儿所使用的词组"β-淀粉样蛋白前体蛋白"(APP)被定义为这样一种多肽，其为同名的位于人类第21于染色体长臂上的基因所编码，且其包括"βAP-这里指β-淀粉样蛋白蛋白"参见上文，在其第三羧基范围内。APP是一种蛋白糖基化的，单膜形成的蛋白，其在许多哺乳动物组织的很广泛的细胞内进行表达。现在已知存在于人类的APP特异同工型的实例为Kang 等 (1987) 自然325:733-736一文中所描述的695—氨基酸多肽，称为"正常"APP; Ponte等. (1988) 自然 331:525-527 (1988) 及Tanzi 等. (1988) 自然 331:528-530 所描述的751—氨基酸多肽，及Kitaguchi 等. (1988) 自然 331:530-532所描述的770—氨基酸多肽。APP变体的具体实例包括点突变其可以因位置和显型而有不同的变化，(参照已知的变种突变，参见flardy (1992) Nature Genet. 1:233-234). 所有文献在此引入作为参考。这里所使用的词组"APP片段"意指除了那些仅包括βAP 或βAP 片段之外的APP片段，也就是说，APP片段包括那些除完整βAP 或βAP 片段之外的APP片段的APP氨基酸序列。

当词组"任一氨基酸"被使用时，所指氨基酸选自下述的三个字母或单个字母的简写，提供如下:

丙胺酸 Ala，A；精氨酸，Arg，R；天冬酰胺酸，Asn，N；天(门)冬氨酸，Asp，D；半胱氨酸，Cys，C;谷氨酸，Gln，Q；lu；E-谷氨酸，Glu，E；氨基乙酸，Gly，G；组氨酸，His，H；异亮氨酸，Ile，I；亮氨酸，leu，L；赖氨酸，Lys，K；蛋氨酸，Met，M；苯丙氨酸，Phe，F；脯氨酸，Pro，P；丝氨酸，Ser，S；苏氨酸，Thr，T；色氨酸，Trp，W；酪

氨酸，Tyr，Y；缬氨酸，Val，V；天冬氨酸或者天冬酰胺酸，Asx，B；谷氨酸或谷氨酸盐，Glx,Z;其它的氨基酸，Xaa，X.

本发明提供了一种可从基因数据库中筛选出具有天冬氨酰蛋白酶活性的单活性位点基元的方法。真核天冬氨酰蛋白酶例如胃蛋白酶及血管紧张肽原酶其具有一种两功能域的结构其折叠可形成两个天冬氨酰残基在活性位点之间更为靠近。它们包埋在短的三肽基序DTG中，或更不常见的包埋于DSG。 许多天冬氨酰蛋白酶作为酶原出现，其N端被切割活化。DTG 或 DSG活性位点基序在酶原中出现在大约第65-70个残基上(凝乳酶原,胃蛋白酶原)，但在N端功能域切割后出现在活性酶的大约第25-30个残基上。活性位点基序的有限长度使得搜索收集短的表达序列标记(EST)以发现新的天冬氨酰蛋白酶的工作变得困难。EST 序列一般平均为250 个核苷或者更短，并由此编码80-90个氨基酸残基或者更短。这样的序列太短而无法跨越所述的两个活性位点基序。优选的方法就是筛选假设的或装配的蛋白编码序列的数据库。本发明描述了一种计算机方法在该蛋白序列数据库中鉴别候选的天冬氨酰蛋白酶。该方法被用于在Caenorhabditis elegans基因组中鉴别7个候选的天冬氨酰蛋白酶序列。这些序列经过同源性检索找到了Hu-Asp1及Hu-Asp2的两个可选择的剪接变体，这里将其命名为Hu-Asp2(a)及Hu-Asp2(b)。

本发明的一种重要的方面我们这里所披露的是有关APP处理的新的信息。淀粉样蛋白前体蛋白（APP）经过Aβ途经的病理学处理需要有β-分泌酶及γ-分泌酶两种蛋白酶的连续作用。经过用β-分泌酶及γ-分泌酶两种蛋白酶切割APP分别产生Aβ肽的N端及C端。因为Aβ肽的过量生产，具体为$A\beta_{1-42}$,可以引发阿尔茨海默氏疾病，因而β-分泌酶和/或γ-分泌酶的抑制剂对治疗阿尔茨海默氏疾病有潜在的作用。尽管β-分泌酶和γ-分泌酶在APP的病理学过程中的重要性，但是这些酶的分子定义至今尚未完成。那就是说，尚未知在β-分泌酶或γ-分泌酶切割位点需要那种酶。这些位点其本身已知是因为APP及$A\beta_{1-42}$已知，参见US 5,766,846及 US 5,837,672，(作为参考引入，除了"可溶性"多肽外)。但是那种酶涉及生产$A\beta_{1-42}$至今尚未知。

本发明涉及到几种新的人类天冬氨酰蛋白酶的分子学定义，其中之一指的是Hu-Asp-2(a)及Hu-Asp2(b)，其特征已作了详细描述。已经公开的asp1及asp2的现有形式，参见EP0848062A2及EP 0855444 A2,发明者为David Powel等，申请人为Smith Kline Beecham公司(已作为参考资料引入)。本发明公开了Hu-Asp 2的新型及旧型。其第一次以一活性形式表达，它们的底物及其特异性都是首次公开。在该发明前，细胞或细胞提取物需要在β-分泌酶位点进行裂解，现在纯化蛋白可以被用于如此分析。基于下述实验结果：(1) 反义敲除实验，(2)瞬时转染敲入实验，及(3)使用纯化重组的Hu-Asp-2进行的生化实验，我们发现Hu-Asp-2是参与APP处理过程的β-分泌酶。尽管已经报道了Hu-Asp-2(a)的核酸及推测的氨基酸序列，如上看到的那样，参见EP 0848062 A2 及 EP 085544 A2，但并未披露这些酶的功能性特征， 这里作者描述了Hu-Asp-2 酶的特点，并能够解释为什么其在$A\beta_{1-42}$多肽的形成过程中为一种重要且必要的酶及在AD的形成过程中为一种重要且必要的步骤。

在另外一种实施例中，本发明同样描述了一种Hu-Asp2的新型剪切变体，其为Hu-Asp2(b)，其在本发明前从未被揭示过。

在另外一种实施例中，本发明提供了一种分离的核酸分子，其包括这样一种多聚核苷酸其编码一种多肽，该多肽选自人天冬氨酰蛋白酶1(Hu-Asp1)及人天冬氨酰蛋白酶2(Hu-Asp2)的两个可选择性变体，即本发明中的Hu-Asp-2(a)及Hu-Asp2(b)。正如这儿所使用的，所有涉及的"Hu-Asp2"应该被理解为指的是Hu-Asp2(a)及Hu-Asp2 (b)。Hu-Asp1在胰脏及前列腺组织中有着丰富的表达量，而Hu-Asp2(a)及Hu-Asp2 (b) 在胰脏及脑组织中有着丰富的表达量。本发明同时也提供了分离的Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)多肽及其显示天冬氨酰蛋白酶活性的片段。

Hu-Asp1，Hu-Asp2(a)及Hu-Asp2 (b)的推测的氨基酸序列与先前所鉴别的哺乳动物天冬氨酰蛋白酶如胃蛋白酶原A，胃蛋白酶原B，组织蛋白酶B，组织蛋白酶E，及血管紧张肽原酶具有显著的同源性。P.B.Szecs, Scand. J. Clin. Lab. Invest. 52:(Suppl. 2105-22(1992)). 这些酶

的共同特征是具有一种重复的DTG/DSG 序列基元。本发明公开的Hu-Asp1及HuAsp2 多肽与来自蛋白质序列资料库的编码天冬氨酰蛋白酶的ProSite共有基序有着极高的同源性。

给出的如SEQ ID NO:1所示的1-1554残基的核苷酸序列对应于编码Hu-Asp1的核苷酸序列，给出的如SEQ ID NO:3所示的1-1503残基的核苷酸序列对应于编码Hu-Asp2(a)的核苷酸序列，而且给出的如SEQ ID NO:5所示的1-1428残基的核苷酸序列对应于编码Hu-Asp2(b)的核苷酸序列。在实施例1及2中将描述编码Hu-Asp1及HuAsp2(a)及Hu-Asp2(b)的DNA的分离及测序。

正如实施例1及2所描述的，自动的测序方法将被用于获得Hu-Asp1，Hu-Asp2(a)及Hu-Asp-2(b)的核苷酸序列。本发明所获得的Hu-Asp核苷酸序列为DNA双链，并且相信为100%的准确。然而，正如本领域公知的那样，用自动测序法所获得的核苷酸序列可能发生一些错误。自动测序法所测定的核苷酸序列与给定的核酸分子的实际核苷酸序列一般有至少90%，更一般的为至少95%至大约99.9%的同一性。实际上核苷酸序列通过用本领域公知的手工测序的方法进行测定将更为准确。序列中的一点错误如加入或删除一个或多个核苷将导致翻译时发生移码其结果是从点突变位点开始，推测的氨基酸序列将完全不同于核酸分子的实际核苷酸序列所能够推测的氨基酸序列。本发明的Hu-Asp DNA 包括 cDNA、化学合成的DNA、通过PCR分离的DNA、 基因组DNA、及其结合体。基因组Hu-Asp DNA可以通过用这里所述的Hu-Asp2 cDNA使用本领域公知的技术常识或者使用寡聚核苷酸（其选自Hu-Asp2序列，其将引导聚合酶链式反应（PCR））筛选基因组文库获得。本发明也同样包括由Hu-Asp DNA转录获得的RNA。

因为遗传密码的简并性，两个DNA序列虽不同，但也可能编码相同的氨基酸序列。本发明因此提供了分离的核酸分子，其多聚核苷酸序列编码本发明所述的任一Hu-Asp多肽，这里所述的多聚核苷酸序列可编码具有如SEQ ID NO:2，SEQ ID NO:4，SEQ ID NO:6或其片段所示的完整的氨基酸序列或其片段的Hu-Asp多肽。

本发明还提供了纯化的Hu-Asp多肽，包括重组和非重组形式，更为

重要的是，提供了一种生产活性Hu-Asp2多肽的方法。其包括在细菌细胞，昆虫细胞及哺乳动物细胞中生产Hu-Asp2 多肽及其衍生物，其中其允许Hu-Asp2多肽从细菌，昆虫及哺乳动物细胞中分泌进入培养基中，以及用于制备掺入了氨基酸标记以利于后续纯化的Hu-Asp2多肽变体的方法。在本发明的一种优选的实施方案中，Hu-Asp2多肽在转化细胞中或者经过纯化及在无细胞系统中用另一种蛋白酶进行裂解被转化为一种蛋白水解活性形态，Hu-Asp2多肽这样的活性形式其N端以序列TQHGIR 或 ETDEEP起始。Hu-Asp蛋白的变体及衍生体，包括片段，只要它们具有如SEQ ID Nos: 2, 4, 和 6 所示的天然氨基酸序列且保留了Hu-Asp 的任一种生物活性，那么就也都在本发明的保护范围内。当然本技术领域的任一普通技术如用标准的天冬氨酰蛋白酶分析化验Hu-Asp蛋白的一种变种，衍生体或者片段也很容易测定其是否可以展示Hu-Asp的活性。本发明范围内的Hu-Asp片段包括那些含有具有氨基酸序列DTG的活性位点区域的片段，及那些含有具有氨基酸序列DSG的活性位点区域的片段，及那些含有同时具有氨基酸序列DSG的活性位点区域的片段，DTG及DSG活性位点序列间隔已被加长的片段，及其间隔被缩短的片段。同样也包括在本发明保护范围内的Hu-Asp片段其跨膜区域被删除以允许Hu-Asp2以一可溶性形式生产。在本发明的另一个实施例中，Hu-Asp2的两个组件，可以重组多肽的形式单独进行生产，然后合并在溶液中，并在溶液中重新组配成活性蛋白酶，上述的每个组件中各含一单个的活性位点DTG或DSG序列。

自然编码Hu-Asp的核苷酸序列可以通过突变获取Hu-Asp变种，例如，一种Hu-Asp变种，这里所指的是与自然的Hu-Asp多肽基本同源的一种多肽，但其中的一段氨基酸序列因为发生一处或多处缺失，插入，或替代而与自然的Hu-Asp的氨基酸序列不同。变体的氨基酸序列或核苷酸序列与自然的Hu-Asp序列同源性为至少80%，优选至少为90%，最优选至少95%。这样，一种变体的核苷酸序列其与自然的Hu-Asp基因序列相比不同，例如，每100个核苷发生5点突变，其将与自然的蛋白有95%的同源性。自然及变种的Hu-Asp序列的相似性，或同源性的百分比可以通过下列方法测定，例如，通过使用一般用于此目的计算机程序，例如，Gap程序(威斯

康星序列分析软件包，Unix第8版，Genetics Computer Group，University Research Park，Madison Wisconsin)其使用Smith 及Waterman的运算法则(Adv. AppL Math. 2.’482489(1981))可以进行测定。

自然的Hu-Asp的氨基酸序列的改变，可以通过大量现有技术中的任一种来实现。例如，用本领域已知方法在特定座位引入突变，如寡聚核苷酸介导的诱变，正如下述几位描述的那样: Walder 等 (*基因* 42:133 (1986))；Bauer 等 (*基因* 37:73 (1985))；Craik (*生物技术*，1985,1月，12-19页)；Smith 等.(*基因工程原理和方法*，Plenum Press (1981))；及美国专利文献Nos. 4,518,584 及 4,737,462。

本发明保护范围内的Hu-Asp变种可以包括保守的替代序列，意指Hu-Asp多肽的一种或几个氨基酸残基为其它不同的残基所取代但其并不会改变Hu-Asp多肽的二级或三级结构。这种替代可以包括氨基酸残基的取代为另一种理化性质相似的多肽的残基所取代，例如替代一种脂肪簇的残基 (Ile，Val，leu or Ala) 为其它，或者替代发生在两个基本的残基之间，如Lys及 Arg，酸性残基 Glu及 Asp，氨基组合物残基Gln 及 Asn，羟基氢氧基残基 Ser 及 Tyr，或者芳香残基Phe及Tyr。有关表型阴性的氨基酸残基的交换的进一步的信息可以在Bowie等，科学 247:1306-1310 (1990)发现。另一种其能充分地表现出的物学活性的Hu-Asp变种为那些氨基酸残基的替代发生在蛋白结构域以外Hu-Asp变种。

另一方面，本发明提供了一种分离的核酸分子，其包括一种多聚核苷酸其能在严格条件下与上述核酸分子的一部分发生杂交，例如，先前所描述的核酸分子的至少大约15个核苷，更好为大约20个核苷，再好为至少大约30个核苷，最好为至少大约30个至100个核苷。这些核酸分子的一部分具有参考资料所述的长度，如，至少有大约15个临近核苷为参考资料所述。在严格的杂交条件下过夜培养，温度大约为42℃，时间为大约2.5个小时，在6 × SSC/0.1% SDS中，接着在65℃下，在1.0 × SSC，0.1% SDS中过滤洗涤。

这里所描述的编码Hu-Asp的核酸分子的片段，以及能与该核酸分子杂交的多聚核苷酸，都可以被用做探针，或可在聚合酶链式反应(PCR)的

过程中作为引物。该探针可以被使用，例如，用于体外测试方法去检测Hu-Asp核酸的存在，以及用于Southern和Northern印迹实验。表达Hu-Asp的细胞类型可以通过使用该探针进行鉴别。这种处理过程是公知的，本领域技术人员可以根据不同的应用选择不同长度的探针。对于PCR来说，相应于所期望的Hu-Asp核酸分子5’及 3’引物可以通过使用传统的技术来分离及扩增该序列。

另外的有用的Hu-Asp 核酸分子为反义或有义的寡聚核苷酸包括一单链核酸序列，其能够与一目标的Hu-Asp mRNA（使用一条有义链），或者Hu-Asp DNA（使用一条反义链）序列进行结合。在本发明的一种优选的实施方案中，这些Hu-Asp 反义寡聚核苷酸减少了Hu-Asp mRNA，并随之减少了Hu-Asp多肽的生产量。

在另一方面，本发明包括有或者没有自然模式糖基化的Hu-Asp 多肽。Hu-Asp1 及Hu-Asp2具有良好的天冬酰氨联结糖基受体位点，其中Hu-Asp 1具有两个这样的位点,Hu-Asp2具有四个这样的位点。酵母或者哺乳类动物细胞表达系统中（下面接着讨论）表达的Hu-Asp与自然状态下的Hu-Asp多肽在分子量或糖基化模式上或者相似或者有显著的不同。Hu-Asp 在细菌系统中的表达将提供一种非糖基化的Hu-Asp。

本发明的多肽更适宜以一分离且充分纯化的形式提供，Hu-Asp多肽可以通过众所周知的方法包括硫化铵或乙醇沉淀，阴阳离子交换色谱，磷酸纤维素色谱， 疏水作用色谱，亲合色谱，羟磷灰石色谱，植物血凝素色谱，及高压液相色谱（HPLC）等从组织，培养细胞，或重组的细胞培养物中恢复和纯化。在一种优选的实施例中，使用基因工程技术将一种氨基酸标记加入到一种Hu-Asp 多肽中为本领域技术人员所共知，其包括六个组氨酸氨基酸残基的加入以允许结合在一适宜支持物上的镍而纯化，单克隆或多克隆的表位 包括但并不限于T7表位，myc抗原决定基及V5a表位,Hu-Asp2与一适宜的蛋白伴侣融合，其包括但并不限于谷胱甘肽-S-转移酶 或者麦芽糖结合蛋白。在一具体的实施例中，这些另外的氨基酸序列加入到Hu-Asp的C端，但其也可以加入到Hu-Asp多肽的N端或其中间的插入位置。

本发明同样涉及到一种载体，其含有本发明的多聚核苷酸分子，以及用该载体转化后的宿主细胞。本发明的任一的多聚核苷酸分子都可被加入到一种载体中，其一般包括一种选择性标记及一种复制的起始位点以在宿主细胞中进行繁殖。因为本发明同样也提供了一种由上述多聚核苷酸分子所表达的Hu-Asp多肽，也选择了一些表达Hu-Asp的载体。这些载体包括编码上述或下述任一Hu-Asp多肽的DNA，其被可操作地连接到一种适宜的转录或翻译调控序列上，例如那些获自哺乳动物细胞，细菌，病毒或昆虫基因的调控序列。调控序列包括转录启动子，操纵子，或增强子，信使RNA核糖体结合位点,及可以调控转录或翻译的适宜序列。当调控序列与编码Hu-Asp的DNA相关时将这些序列可操作地连接在一起。这样，如果一种启动子核苷酸序列可介导Hu-Asp序列的转录的话，那么就将其可操作地连接一段Hu-Asp DNA序列上。

为了克隆编码Hu-Asp的多聚核苷酸分子，或者为了表达Hu-Asp多肽，选择适宜载体的，这当然取决于载体所转化宿主细胞。而且如果可以的话，宿主细胞应选自Hu-Asp多肽可以被表达的细胞。表达Hu-Asp多肽的适宜的宿主细胞包括原核生物细胞，酵母细胞及高等的真核生物细胞，每一种下面都要进行论述。

在这些宿主细胞中表达的Hu-Asp多肽也可以为融合性蛋白，其包括来自异种蛋白的区域。其可以带来这样的优点，允许分泌，稳定性的提高，或利于蛋白质的纯化。例如，一段编码一定信号肽的序列，可以被引入表达载体中。一种编码该信号肽的DNA序列(导向分泌) 可以被融入Hu-Asp序列的阅读框内，这样Hu-Asp被作为一含有该信号肽的融合性蛋白被翻译。在目标宿主细胞中该信号肽具有一定的功能，其可以促进Hu-Asp多肽的胞外分泌。优选地，一旦Hu-Asp多肽从细胞中分泌完毕，信号肽将与Hu-Asp多肽分离。信号序列无任何限制，在实践该发明时其可以包括酵母细胞的I因子，及在sf9昆虫细胞中的蜜蜂褪黑激素导向因子。

在一种优选的实施方案中，Hu-Asp多肽作为一融合蛋白其包括一种异源结构域其可被用于促进该多肽的纯化。许多可用于此功能的蛋白允许该分泌蛋白选择性结合在一结合伴侣的融合性蛋白上。例如，Hu-Asp

多肽可被修饰以包括一种多肽以形成一种融合蛋白，其特异性结合在一结合伴侣上，或一多肽标记上。该多肽标记的非限制性实例包括6-His 标记，硫氧还蛋白标记，血球凝集素标记，GST标记，及 OmpA 信号序列标记。正如本技术领域人员所能理解的那样，其能够被认识及结合在一多肽上的结合伴侣可以是任一分子或组合物包括金属离子（例如 金属亲合柱），抗体，或其片段。任一结合该多肽的蛋白或多肽，例如FLAG标记。

表达Hu-Asp多肽的适宜的宿主细胞包括原核生物细胞,酵母细胞，及高等的真核生物细胞。用于表达Hu-Asp的适宜的原核生物宿主细胞包括埃希氏杆菌属细菌，杆菌属细菌，及沙门氏菌属细菌，以及假单胞菌属，链酶菌属，及葡萄球菌属细菌。对于在大肠杆菌中进行的表达，一种Hu-Asp 多肽可能包括一种N末端甲硫氨酸残基，以利于在一原核宿主细胞中表达该重组多肽。N末端甲硫氨酸残基可以随意地从表达的Hu-Asp多肽中进行分离。另一种可以被加入到Hu-Asp 多肽中以利于在大肠杆菌中进行表达的N末端氨基酸残基包括但并不限于T7前导序列，T7caspase 8前导序列,及其它的前导序列包括用于纯化的标记如6-His 标记(见实施例9)。在大肠杆菌中表达的Hu-Asp多肽可以通过移去细胞质尾端，横跨膜区域，或膜最接近区域而被缩短。在大肠杆菌中表达的Hu-Asp多肽可以溶解形式或者以包含体在或非包含体的不溶形式而获得。非溶解多肽可以通过加入HCl胍，尿素或其它蛋白质变性剂然后通过稀释或透析进入合适的含水缓冲液之前或之后纯化前或后重新折叠形可溶形式。如果使用重组方法生产失活的Hu-Asp前体，可以用另一种适宜的蛋白酶如人类免疫性缺陷病毒蛋白酶切除一前体节段（prosegment）使其具有活性。

用于原核宿主细胞中的表达载体一般包括一种或多个可选择性标记的表型基因，该基因一般编码，例如，具有抗生素抗性或满足一营养缺陷型要求的蛋白。该载体具有一种较宽的选择范围其很容易从商业渠道获得。例如有 pSPORT 载体，pGEM 载体 (Promega)，pPRQEX 载体 (LTI Bethesda, MD)，Bluescript 载体 (Stratagene)，pET载体 (Novagen) 及 pQE 载体 (Qiagen)。

Hu-Asp 同样也可在酵母宿主细胞中表达，其可选自下列的种属，

包括酵母属，毕赤酵母属，及克鲁维酵母属。最适宜的酵母宿主细胞为酿酒酵母及巴斯德毕赤酵母。酵母载体通常包括一种来自2T酵母质粒的复制序列的起始位点，一段自动复制序列(ARS)，一个启动子区，一段聚腺甘化的序列，一段转录终止的序列，及一种选择性标记基因。同时用到在酵母及大肠杆菌中都可进行复制的载体（命名为穿梭质粒），除了上述的酵母载体的特征外，穿梭质粒同样也包括一种在大肠杆菌中进行选择和复制的序列。在酵母宿主细胞中表达的Hu-Asp多肽的直接分泌可以通过在编码Hu-Asp的核苷酸序列的5'末端融入编码酵母I因子前导序列的核苷酸序列来完成。

昆虫宿主细胞培养系统可同样被用于表达Hu-Asp多肽，在一优选的实施方案中，本发明的Hu-Asp多肽使用一昆虫细胞表达系统进行表达（参见实施例10）。此外,一杆状病毒表达系统也可用于在昆虫表达系统中表达Hu-Asp多肽，如luckow及Summers，生化技术 6:47(1988)所述那样。

在另一种优选实施方案中,Hu-Asp多肽在哺乳动物宿主细胞中进行表达。适宜的哺乳动物细胞系的非限制性实施例包括猴肾细胞COS-7细胞系(Gluzman等,细胞 23:175 (1981))，人胚胎肾细胞系293，及中国鼠卵巢(CHO)细胞。最好是，中国仓鼠卵巢(CHO)细胞被用于表达Hu-Asp蛋白（见实施例11）。

本发明的Hu-Asp多肽的适宜性表达载体的选择将当然依赖于所使用的哺乳动物宿主细胞，这是本领域技术人员所共知的常识。适宜性表达载体例如包括pcDNA3 (Invitrogen)及pSVL (Pharmacia Biotech)。用于表达Hu-Asp多肽的载体为pcDNA3.1-Hygro (Invitrogen)。在哺乳动物宿主细胞中进行表达的载体可以包括获自病毒基因组的转录及转译控制序列。本发明一般所使用的启动序列及增强序列包括但并不限于，那些获自人类细胞巨化病毒(CMV)，腺病毒2， 多形瘤病毒,及猿病毒40 (SV40)。这里也描述了构建哺乳动物细胞表达载体的方法，例如，在Okayama 及 Berg (MoL CelL BioL 3:280 (1983))；Cosman 等(MoL，免疫学 23:935 (1986))；Cosman 等 （自然 312:768(1984))；EP-A-0367566；及 WO 91/18982。

本发明的多肽同样可被用于制备单克隆或多克隆抗体，这些抗体可以用于检测Hu-Asp 多肽表达的诊断实验中。这些抗体可以通过传统的技术手段来进行制备。参见，例如， 抗体：实验室指南，Harlow 及 Land (eds.)，Cold Spring Harbor Laboratory Press，Cold Spring Harbor，N.Y.，(1988)；单克隆抗体,杂交瘤：生物分析的一种新的方法，Kennet 等 (eds.)，Plenum Press，New York (1980)。包括Hu-Asp第5至20位氨基酸残基的合成多肽也可用于生产单克隆或多克隆抗体，但在这之前须先连接一适宜的载体蛋白，该蛋白包括但并不限于匙孔血蓝蛋白(KLH)，鸡卵白蛋白，或者牛血清白蛋白，使用不同的交联试剂包括碳二亚胺，戊二醛，或者如果该多肽含有一种半胱氨酸，N-甲基顺丁烯二酰亚胺。用于免疫目的的一适宜性多肽当与KLH结合时含有分别含有QRRPRDPEVVNDESSLVRHRWK或者LRQQHDDFADDISLLK的Hu-Asp1或者Hu-Asp2的C末端。

本发明的Hu-Asp核酸分子对于染色体鉴别来说同样是有用的，因为其能够与人染色体的特定位点进行杂交。Hu-Asp1已被定位于第21号染色体上，而Hu-Asp2被定位于染色体11q23.3-24.1上。 在染色体上进行特定位点的鉴别是一现实需要，因为基于实际序列数据库（重复多态性）的少数染色体标记试剂对于在染色体位点上标记是现实可用的。一旦一种序列被绘制至一种精确的染色体位点，该序列在染色体上的自然位点可与基因图谱库进行关联。基因与疾病（其已被绘制在同一染色体区域）的关系可以通过连接分析进行测定，这里自然上接近的基因的共遗传可以被测定。看起来与一特定疾病相关的一种基因事实上是否该疾病的导致原因可以通过比较受疾病影响及未受其影响的个体的核酸序列来进行测定。

在另一种实施例中，本发明也涉及到一种分析Hu-Asp特别是Hu-Asp2的功能的方法，其中所述， 在溶液中与一适宜底物温育Hu-Asp多肽，该底物包括但并不限于含有APPβ分泌酶切割位点的合成多肽，最好是一种含有在瑞典人家族中的遗传性AD发现的突变如KM变为NL的多肽，这样的多肽包括有序列SEVNLDAEFR，在酸性缓冲液中，最好是pH为5.5(见实施

例12)的酸性缓冲液中使用高压液相色谱来监控该多肽的分解以来作为对Hu-Asp活性的测量。蛋白酶水解活性的最佳测试方法使用内部熄灭多肽化验底物的方法。该适宜性底物包括附有一对荧光团及一种含有但并不限于香豆素及二硝基酚的淬火剂的多肽。这样一来Hu-Asp切割多肽的结果将因为荧光团与淬火剂的的分离而导致荧光的增加。Hu-Asp蛋白酶水解活性的比色分析将使用另一合适的底物，其包括P2及具有通过一氨基组合物连接至o-硝基酚的识别位点的P1氨基酸，这样一来通过Hu-Asp切割改变分析缓冲液至一碱性pH后将导致光学密度的增加。

在另一实施例中，本发明涉及到一种方法，其可以鉴别这样一种试剂，其增强了Hu-Asp多肽的活性，该多肽选自含有Hu-Asp1，Hu-Asp2(a)，及Hu-Asp2（b）的组，该方法包括:

(a)在待测试剂的存在下或不存在下测定所述Hu-Asp多肽的活性，且，

(b)比较在待测试剂的存在下与不存在下所测定的所述Hu-Asp多肽的活性;

在试剂存在的情况下所述Hu-Asp多肽的活性比较在试剂不存在的情况下所述Hu-Asp多肽的活性要高显示出该试剂增加了所述Hu-Asp多肽的活性。该种与Hu-Asp多肽一起进行的测试可以在无细胞系统中也可以在表达Hu-Asp的培养细胞中及其衍生体或变体进行。

在另一实施例中，本发明涉及到一种方法，其可以鉴别这样一种试剂，其减弱了Hu-Asp多肽的活性，该多肽选自含有Hu-Asp1，Hu-Asp2(a)，及Hu-Asp2（b）的组，该方法包括:

(a)在待测试剂的存在下或不存在下测定所述Hu-Asp多肽的活性，且，

(b)比较在待测试剂的存在下与不存在下所测定的所述Hu-Asp多肽的活性;

在试剂存在的情况下所述Hu-Asp多肽的活性比较在试剂不存在的情况下所述Hu-Asp多肽的活性要低显示出该试剂减弱了所述Hu-Asp多肽的活性。该种与Hu-Asp多肽一起进行的测试可以在无细胞系统中也可以在

表达Hu-Asp的培养细胞中及其衍生体或变体进行。

在另一实施例中，本发明涉及到一种新的细胞系(HEK125.3 细胞)以测量淀粉样蛋白β肽(Aβ)从淀粉样蛋白蛋白前体(APP)发生的进程。这些细胞都是来自人类胚胎肾293细胞(HEK293)的稳定的转化子，其具有一种双顺反子载体其获自含有一种修饰的人类APP cDNA，一种内部核糖体进入位点及在另一种顺反子上的一种增强的绿色荧光蛋白(EGFP) cDNA 的pIRES-EGFP (Clontech)， APP cDNA通过在一APP编码序列的羧基末端上加入两个赖氨酸密码子来进行修饰，这样一来其提高了来自人类APP的Aβ肽2-4倍。Aβ肽的处理水平相比较未转化的HEK293细胞提高了60倍。HEK125.3细胞将对分析抑制Aβ肽的处理过程的组合物来讲是有用的。本发明同样也包括了在APP的另一同工型包括751及770氨基酸同工型的C末端增加两个赖氨酸残基，或在具有人类AD发现的突变（包括瑞典KM→NL及V717→F的突变）的APP的同工型的C末端增加两个赖氨酸残基，在APP片段的C末端增加两个赖氨酸残基，例如那些以β分泌酶切割位点开始APP片段的C末端，及在含有β分泌酶切割位点（其可以被连接到一种N端信号肽上以进行膜插入或分泌）的APP片段的C末端增加两个赖氨酸残基，及一种APP片段的C末端（其可以被连接到一种N端信号肽上以进行膜插入或分泌及一种信号序列（包括但并不限于绿色荧光蛋白或者碱性磷酸脂酶这样β分泌酶切割就会从表达该多肽的细胞表面释放该信号蛋白）上）增加两个赖氨酸残基。

上文已对本发明作了概括性描述，通过下文的实施例将对本发明有更进一步的理解，提供这些实施例是为了说明本发明而非进行任何限制。

实施例

实施例 1： 可用于在Wormpep 12中进行C. elegans天冬氨酰蛋白酶基因及天冬氨酰蛋白酶的鉴别的搜寻规则系统的发展

材料和方法：

一般的天冬氨酰蛋白酶例如胃蛋白酶及血管紧张肽原酶具有两个结构域结构，该结构的折叠可使活性位点中的两个天冬氨酰残基成为邻近残基。它们包埋在短的三肽基序DTG中，或更为罕见地包埋于DSG。DTG或者DSG活性位点基序出现在酶的大约25-30个氨基酸残基之间， 而酶原中（凝乳酶原，胃蛋白酶原）为约65-70个氨基酸残基之间。该基序再一次出现在酶的下游大约150-200个氨基酸残基之间。通过切除N端前结构域激活该酶原。下面的横式例示了通过基因复制及趋异明显引发的天冬氨酰蛋白酶的双重结构域结构:

NH_2------X-----$D^{25}TG$----------Y----$D^{Y+25}TG$----------C

X表示酶的开始，位于N末端前结构域之后， Y表示分子中心在这里基因再一次重复开始。

就逆转录病毒酶，例如HIV蛋白酶而言，它们仅具备众所周知的蛋白酶如胃蛋白酶，组织蛋白酶D，血管紧张肽原酶等的二结构域结构的一半。它们无前体节段（prosegment），而是从含有病毒gag 及 pol蛋白的多蛋白前体中切割出来，其可以被表示为:

NH_2---------$D^{25}TG$----------C 100

这种"单体"仅具有约100个氨基酸，因此与其它的前结构域不计在内的N端氨基酸具有330个左右氨基酸的天冬氨酰蛋白酶二聚物相比是极为节省的。

真核天冬氨酰蛋白酶活性位点基序的有限长度对于为搜寻收集新序列的EST来说是很困难的。典型地EST序列平均长为250个核苷，因此在这种情况下不可能横跨两个天冬氨酰蛋白酶活性位点的基序。作为替代，如果我们转向C. elegans基因组。C. elegans基因组估计含有大约13,000个基因。当然，大约有12,000个被测序，而其相应的假定阅读框 (ORF) 被放置于数据库Wormpep12中。我们使用这个数据库作为对高等真核生物新型天冬氨酰蛋白酶的整个基因组浏览的基础。所使用的方法为我们开发

专门用于此目的的运算法则。下面的这个AWK程序（其专门用于定位含有两个DTG或DSG基序的蛋白）被用来作为搜寻，其可以被重复四次以恢复天冬氨酰基序的成双联合。

```
BEGIN{RS=">"}            /* defines ">"as record separator for FASTA format */
{
pos = index($0,"DTG")    /*finds "DTG" in record*/
 if (pos>0) {
      rest = substr($0,pos+3)  /*get rest of record after first DTG*/
      pos2 = index(rest,"DTG") /*find second DTG*/
      if (pos2>0) printf ("%s%s\n",">",$0)}     /* report hits*/
}
}
```

上述的 AWK 程序被用于搜寻 Wormpep12,其可以从下列网址进行下载： fttp.sanger.ac.uk/pub/databases/wormpep，该序列至少含有两个 DTG 或 DSG 基序。使用 AWK 限制每一记录为 3000 个特征或更少。这样，35 个左右大的记录从 Wormpep12 中手工进行删除因为不管在什么情况下其都不可能编码天冬氨酰蛋白酶。

结果及讨论:

Wormpep 12数据库具有12,178个记录，尽管其中有些(<10%) 代表来自相同基因的可选择性剪接转录子。估计在C. elegans 基因组中的编码基因数量为13,000个，因此，Wormpep12可被估计为覆盖了C. elegans 基因组的大于90%的基因组。

真核天冬氨酰蛋白酶含有一种双结构域结构，其大概源自祖先基因的复制。每一种结构域都含有活性位点基序D(S/T)G 其位于每一结构域的20-25个氨基酸残基之间。逆转录病毒(例如., HIV蛋白酶）或者反转座子蛋白酶都为亚基的同聚二聚体，其与单一的真核天冬氨酰蛋白酶结

构域有一定的同源性。一种AWK 程序被用于搜寻Wormpep12数据库以找出这样的蛋白其中D(S/T)G基序至少出现两次。这样鉴别出超过60个蛋白质具有两个DTG或者DSG基序。 病毒检测被用于选择蛋白其中天冬氨酰结构域的位置被提示为一种双结构域的结构满足上述的机理。

此外，PROSITE 真核及病毒天冬氨酰蛋白酶活性位点模式PS00141被用于搜寻Wormpep12数据库以找出候选的天冬氨酰蛋白酶(Bairoch A.，Bucher P.，Hofmann K.，PROSITE 数据库：其1997年状况，核酸研究. 24:217-221(1997))．其产生了一种与Wormpep12一样的序列，当然，这7个序列含有两个DTG 或者 DSG 基序及PROSITE天冬氨酰蛋白酶活性位点模式。其中三个发现在同一装配性质粒克隆中(F21F8.3，F21F8.4，及F21F8.7) 启示其代表了一种因祖先基因复制而激活的蛋白质家族。其它两个具有与F21F8.3，F21F8.4 及 F21F8.7有很强同源性的ORFs位于相同的基因簇(F21F8.2 及F21F8.6)中，然而，其仅仅含有一种DTG基序。 在Wormpep12数据库中进行更为详尽的BLAST搜索这7个序列并未发现在C. elegans 基因组（含有两个重复序列DTG 或DSG基序）中其它的候选天冬氨酰蛋白酶。

用BLASTX 在SWISS-PROT，GenPep 及TREMBL中搜寻每一种C. elegans 序列揭示出R12H7.2为与哺乳动物天冬氨酰蛋白酶同源的最为接近的蠕虫， T18H9.2为有稍微远一点的相关性，而CEASP1，F21F8.3，F21F8.4，及F21F8.7 形成了一种亚簇其序列与哺乳动物序列同源性最低。

讨论:

APP，早老蛋白（presenilins），及 p35，cdk5的活化剂,所有都要经历胞内蛋白酶水解，在该位点上遵循HIV蛋白酶底物特异性的特点。因此，具有相同的底物特异性的胞内天冬氨酰蛋白酶的调控紊乱，可能为提供了一种AD中的和混乱病理的统一机制。因此，我们着力去辨别新的人类天冬氨酰蛋白酶。在C. elegans 中进行的全基因组搜寻辨别出7个开放阅读框，其附于我们已经鉴别出的天冬氨酰蛋白酶谱上。这7个天冬

氨酰蛋白酶可能包含有在一种简单的，多细胞真核生物中的该蛋白酶完整的互补序列。其包括4个唯一于C. elegans （其可能因祖先基因的复制而激活）的4个最为接近相关的天冬氨酰蛋白酶。其它三个候选的天冬氨酰蛋白酶(T18H9.2，R12H7.2及C11D2.2）被发现具有与哺乳动物基因序列的同源性。

实施例2:通过基因组桥连用数据库Mining来鉴别新型的人类天冬氨酰蛋白酶

材料和方法:

EST数据库，cDNA，及推测的多肽序列的计算机辅助分析:

用CEASP1，F21F8.3，F21F8.4,及F21F8.7序列对EST数据库进行详尽的同源性检索并未发现任何新的哺乳动物的同源物。TBLASTN用R12H7.2搜索显示其与组织蛋白酶D,组织蛋白酶E,胃蛋白酶原A，胃蛋白酶原C及血管紧张肽原酶有很强的同源性，尤其是在活性位点的DTG基序左右，但并未识别出其它额外的新的哺乳动物细胞的天冬氨酰蛋白酶。这显示了C. elegans 基因组可能仅含有一种单链的溶酶体天冬氨酰蛋白酶其在哺乳动物细胞中为一种基因族所表现并通过一种祖先基因的复制及随之产生的修饰而激活。

TBLASTN用T18H9.2进行搜索，剩余的C. elegans序列，辨别出几个ESTs其被装配进入一种编码一新型人类天冬氨酰蛋白酶 (Hu-Asp 1)的邻近序列群。正如上述实施例1所述，用Hu-Asp1的邻近序列群进行的BLASTX搜索SWISS-PROT 发现序列的活性位点基序与其它的天冬氨酰蛋白酶的活性位点排列成一行。详尽的，反复地进行BLASTN搜索LifeSeq，LifeSeqFL,及公开的EST收集物，从已被装配进入一单链邻近序列群的多倍CDNA文库中鉴别出102EST。发现于公开EST收集物中的该邻近序列群的51序列同样亦被基因组研究学院（TIGR）装配进入一单链邻近序列群(THC213329)。该TIGR的注解显示其并未发现在邻近序列群的数据库中有所要寻找的序列。注意TIGR邻近序列群是我们所装配的LifeSeq邻近序列群的反义互

补。用Hu-Asp1在ZooSeq中进行BLASTN搜索鼠类EST序列发现在每一数据库中有一种同源性的EST (分别是Incyte clone 700311523及 IMAGE clone 313341, GenBank 序列号 W 10530,)。

用装配的编码Hu-Asp1的DNA序列进行TBLASTN搜索LifeSeqFL及公开的EST数据库显示出另一种相关的人类序列(Hu-Asp2) 其用一单链EST (2696295)来体现。该部分CDNA序列的翻译揭示出一种单链DTG基序其与一牛天冬氨酰蛋白酶（NM 1）的活性位点基序具有同源性。.

BLAST搜索邻近序列群集合及多倍序列集合可以通过使用来自Incyte的LifeSeq, LifeSeqFL及LifeSeq的装配数据库所提供的生物信息学工具来完成。推测蛋白基序可以通过使用ProSite 文(Motifs in GCG 9)库或者Pfam 数据库来鉴别。

Hu-Asp1的全长cDNA克隆:

C. elegans的基因T18H9.2CE的开放阅读框被用于搜索Incyte LifeSeq 及LifeSeq-FL 数据库，检测到一单电子集合1863920CE1。5'端更多的cDNA克隆位于这个邻近序列群1863920, 其获自Incyte且两条链已被完全测序。含有克隆1863920的开放阅读框的翻译揭示出复制的天冬氨酰蛋白酶的活性位点基序(DTG/DSG)的存在但5'末端却是不完整的。Hu-Asp1编码序列的残余物可通过使用一种人类胎盘制备的cDNA模板(Clonetech)由5' Marathon RACE 分析进行测定. 一种为5' 末端克隆1863920特异性的3'-反义寡聚核苷酸引物与在PCR中为Marathon制备cDNA调节子特异性的5'-有义引物进行配对。特异的PCR产品通过循环测序被直接序列化且结果产生装配有克隆1863920序列以产生完整的Hu-Asp-1 (SEQ ID No.1)的编码序列的序列。

几个相关特征都已体现在Hu-Asp1的原始的氨基酸序列中(图1, SEQ ID No.2)。该序列包括一种信号肽(SEQ ID No.2中的残基1-20), 一前体节段,及一种含有天冬氨酰蛋白酶的活性位点基序(DTG/DSG)的两个拷贝的

酶促结构域。第一种及第二个活性位点基序之间的间隔大约为200个残基其符合所期望的单个真核天冬氨酰蛋白酶结构域的大小。更为相关的是，该序列包含一种靠近C端的所期望的横跨膜的结构域(SEQ ID No.2中的残基469-492)其提示该蛋白酶是镶嵌在膜上的。该特征在其它的天冬氨酰蛋白酶上并未发现。

Hu-Asp2的全长cDNAs克隆:

正如实施例1所述的那样，Caenorhabditis elegans的为推定的天冬氨酰蛋白酶数据库WormPep12基因组大范围扫描并降低了人类EST数据库揭示了一种人类与C. elegans基因T18H9.2 的定向进化同源基因其这里所指为Hu-Asp1。装配的Hu-Asp1的邻近序列群通过使用BLAST搜索工具在人类EST数据库中被用于调查人类平行进化同源基因结果与Lifeseq FL数据库相比较有大约60%的同源性其具有一定的匹配性(2696295CE1)。相同的调查对或者gb105PubEST或者可能来自TIGR人类数据库家族并未发现相似的EST克隆。CDNA克隆2696295,辨别于一来自人类子宫cDNA文库的单链继代序列分析，其获自Incyte双链的完整序列。该克隆包括一种不完整的1266个bp的开放阅读框其编码一种422个氨基酸多肽但在5'末端缺乏一种起始密码子ATG。检查推测序列显示复制的天冬氨酰蛋白酶的活性位点基序(DTG/DSG)的存在，其被194个氨基酸残基分开。随后检索LifeSeq EST数据库的后期版本发现一种额外的ESTs序列，其来自人类天冬氨酰cDNA文库(4386993)，其表现出含有额外的与克隆2696295相关的5'末端序列。克隆4386993获自Incyte及双链具有完整序列。对克隆4386993及克隆2696295进行比较分析证实克隆4386993用31个氨基酸残基（包括两个基本点框内翻译起始密码子）扩展了开放阅读框。尽管两个框内密码子ATGs的存在，在ATG 的下游并未发现框内终止密码子，这说明4386993 并非全长。此外，克隆2696295 及4386993的序列比较揭示了于发现相对于克隆4386993克隆2696295的插入75个碱基对中因而25个额外氨基酸残基插入进克隆2696295中。Hu-Asp2编码序列的余值可以通过使用一人类海马

Marathon 制备cDNA模板用5’ Marathon RACE 分析进行测定。一种特异于所共享的克隆2696295及4386993的3’-反义寡聚核苷酸引物与一种特异于在PCR中Marathon制备cDNA合成的调节子的5’-有义引物进行配对。具体的PCR产品可通过循环序列化进行直接测序，结果用克隆2696295及4386993的序列装配的最终序列分别产生了一种完整的编码Hu-Asp2(a) (SEQ ID No.3) 及 Hu-Asp2（b） (SEQ ID No.5)的序列。

几个相关特征列于Hu-Asp2(a) (图2及SEQ ID No.4)及Hu-Asp-2(b) (图3, SEQ ID No.6)的原始氨基酸序列中。这两个序列都含有一种信号多肽(在SEQ ID No.4及SEQ ID No.6中的残基1-21),一种前体节段，及一种含有两个天冬氨酰蛋白酶活性位点基序(DTG/DSG)的两个拷贝的酶促结构域。第一种及第二个活性位点基序之间的间隔因为在Hu-Asp-2（b)中25个氨基酸残基的删除而有所变化，Hu-Asp2（b)及Hu-Asp-2(a)中分别含有168及194个氨基酸残基。更为相关的是，两个序列都含有一接近C末端的推测的横跨膜功能域(在SEQ ID No.4 中的残基455-477及在SEQ ID No.6中的残基430-452)其启示该蛋白酶镶嵌于膜上。该特征在除了Hu-Asp 1外的其它种的天冬氨酰蛋白酶中并未发现。

实施例3: 鼠Asp2 cDNA及基因组DNA的分子克隆，鼠类Asp2 cDNA的克隆及特征分析

鼠类动物的Hu-Asp2的定向进化同源基因使用重组cDNA文库筛选,PCR,及基因组克隆进行克隆。来自鼠大脑cDNA文库的大约500,000个独立的克隆通过使用32P标记的编码序列探针（用Hu-Asp2制备）进行筛选。按照DNA序列分析进行正链复制，最长的cDNA含有3’端未翻译的全部区域及编码区的47个氨基酸。先经过DNA序列分析，用特异于上面测定的5’端最多的cDNA序列的反义寡聚核苷酸引物及特异于人类5’端Asp2序列的有义引物对相同的鼠脑cDNA文库进行PCR扩增以提供一种980个bp的编码序列。鼠类Asp-2的5’端序列的残余起源于基因组序列(参见下文)。

鼠类Asp-2基因的分离及序列分析---------编码鼠类Asp2 cDNA一部

分的鼠类EST序列通过使用BLAST搜索工具及Hu-Asp2编码序列作为调查以在GenBank EST数据库中进行鉴别。克隆g3160898与超过352个bp的人类序列显示出88%的同源性。与鼠类Asp2的该区域特异性配对的寡聚核苷酸引物被合成且用于扩增鼠类基因的所述区域。鼠类基因组DNA，来自种属129/SvJ，其用PCR进行扩增（25个循环），其使用特异于鼠类Asp2的不同引物对及并用琼脂糖凝胶电泳对产品进行分析。引物对Zoo-1及Zoo4扩增了一种750个bp的片段其基于与已知的cDNA序列进行的比较含有大约600个bp的内含子序列。该引物对然后被用于通过PCR筛选一种鼠类BAC文库，单个基因组克隆被分离及该克隆通过DNA序列分析被证实含有鼠科动物Asp2基因。该Asp2基因组克隆的ShotgunDNA序列测定及与Hu_Asp2的cDNA序列及被全长鼠类Asp2序列(SEQ ID No.7)所测定的部分鼠类cDNA序列进行比较。该推测的鼠类Asp2的氨基酸序列(SEQ ID No.8)显示出相比较人类序列(图4)与18/501氨基酸残基取代有96.4%的同源性(GCG最佳配对运算法则)。

实施例4: Hu-Asp2转录表达的组织分布

材料和方法:

Hu-Asp2表达的组织分配可以通过使用获自Clonetech (Palo Alto, CA)的多倍体组织RNA印迹法来进行测定。在载体pINCY中的Incyte克隆2696295用EcoRI/NotI进行消化，插入1.8 kb cDNA，用制备型琼脂糖凝胶电泳进行纯化。该片段被荧光标记至一特定的活性其大于1×10^9 dpm/μg，在[a-^{32}P-dATP] (>3000 Ci/mmol, Amersham, Arlington Heights, IL)及DNA多聚酶I的Klenow片段的存在下经过任意引导，从不同人类组织分离的尼龙膜过滤的含有变性的，大小分级分离的多聚A+ RNAs与2×10^6dpm/ml 的探针在ExpressHyb缓冲液 (Clonetech, Palo Alto, CA)中，在68 ^{0}C下进行杂交1个小时，然后通过推荐的手工方法进行洗涤。通过自动射线照射术（使用BioMax XR 胶卷 (Kodak, Rochester, NY)）显示杂

交信号，并在-80 ℃下筛选鉴别。

结果和讨论:

因为相对较少量的ESTs 通过使用上述方法进行检测，来自于数据库分析的关于Hu-Asp2转录表达的组织分配的信息是很有限的(<5)。在获得Hu-Asp2基因表达的更多的信息的努力中，应用Northern分析进行测定Hu-Asp2转录子的大小和数量。经过在变性环境下分离，分离自一系列外周组织及脑组织的PolyA$^+$ RNAs被显示于一固体支持物上，Hu-Asp2转录子通过高度严格杂交条件下杂交一来自于克隆2696295的荧光标记插入物而被显现。该 2696295 cDNA探针显现出一转录子群其中表面尺寸为3.0kb，4.4 kb 及8.0 kb 的转录子共移动，其中后两者为最丰富的。

经过组织学调查，Hu-Asp2转录子在胰腺及脑组织中是最为丰富的，在其它所检测的组织中除了胸腺及PBLs其含量虽低但是还可以达到可测量的水平。假定在脑组织中的Hu-Asp2转录子的相对含量，脑组织的区域表达也是可以建立的。一相似尺寸的转录子群在所检测的脑区的所有区域都是可以检测到的。[小脑，大脑皮层，枕骨柱，脑前叶，颞叶，及大脑皮层]，在骨髓及脊柱中有最高的含量。

实施例5: 在人类细胞系中对Hu-Asp-1及Hu-Asp-2转录子进行RNA印迹检测

对多种人类细胞系进行测试以测定其产生Hu-Asp 1及Hu-Asp2 mRNA的能力。人类胚胎肾(HEK-293)细胞，非洲绿猴(Cos-7)细胞，中国仓鼠卵巢(CHO)细胞，HELA细胞，及成神经瘤细胞系IMR-32都获自ATCC。除了CHO细胞保持在α-MEM/10% FCS外，其它细胞被培养在含有10% FCS的DME中，在37℃下，5%的CO_2下直到几近铺满。洗过的细胞单层，(3×10^7)被溶解在盘中，使用Qiagen Oligotex介导的信使RNA试剂盒抽提poly A$^+$ RNA。来自每一细胞系的含有2 μg poly A+ RNA的样品在变性条件下分馏

(乙二醛处理)，通过毛细管作用转移到一种固体尼龙膜支持物上，通过与任意引导标记(32P)（获自Hu-Asp1或Hu-Asp2的编码序列探针）杂交使转录子发生显色反应。通过暴露于X射线胶卷检测到放射性信号，然后使用磷光计进行图象分析。

Hu-Asp1 cDNA探针显示出一种相似的转录子群(2.6 kb及3.5 kb)其先前在人类组织器官中被检测到过。放射性信号定量分析测定其相对含量为Cos-7 > HEK 292= HELA> IMR32。

Hu-Asp2 cDNA探针与组织相比较同样也显示出一种相似的转录子群(3.0 kb，4.4 kb，及 8.0 kb)，其具有下列相对含量：HEK 293>Cos7>IMR32>HELA。

实施例6：修饰APP以增加Aβ处理用于体外筛选

从APP中加工Aβ肽的人类细胞系提供了一种方法其可以以胞内实验筛选β-分泌酶及γ-分泌酶的抑制物。Aβ肽的生产及释放进入培养物上清液中可以通过酶联免疫试验(EIA)进行调控。尽管APP的表达有很宽的范围，不管是神经细胞系或非神经细胞系都处理及释放Aβ肽，内生APP的处理水平却是很低的且不容易为EIA所检测到。Aβ肽的处理过程可以通过在APP突变的转化细胞系中表达以进行提高，我们突然也发现到在APP695的羧基末端增加两个赖氨酸残基可以提高Aβ肽的处理过程，这就允许我们制造了一种转化细胞系其释放Aβ肽进入培养物上清液中达到一种显著的水平，20,000 pg/ml。

材料和方法:

材料:

人类胚胎肾细胞系293 (HEK293细胞)由内部获得，载体pIRES-EGFP购自Clontech。使用聚合酶链反应（PCR）的进行突变的寡聚核苷酸O购

自Genosys。含有人类APP695的质粒(SEQ ID No.9 [核苷] 及 SEQ ID No.10 [氨基酸])获自西北大学医学院。其在Not1位点被亚克隆进p5K (Stratagene)以产生一种质粒APP695。

诱变方法:

通过使用Stratagene快速诱变试剂盒将瑞典突变 (K670N, M671L)引入pAPP695制备出所述质粒pAPP695NL (SEQ ID No. 11 [核苷]及 SEQ ID No.12 [氨基酸])。 为了在APP695的C末端引入一种双赖氨酸基序。向前的引物#276 5' GACTGACCACTCGACCAGGTTC (SEQ ID No.47) 被用于使用一种"补丁"引物#2745'CGAATTAAATTCCAGCACACTGGCTACTTCTTGTTCTGCA TCTCAAAGAAC(SEQ ID No.48) 及一种侧面引物 #275 CGAATTAAATTCCAGCA CACTGGCTA(SEQ ID No.49)以修饰APP695 cDNA的3'末端(SEQ ID No.15 [核苷]及 SEQ ID No.16 [氨基酸])。其同样加入一种BstX1 限制酶切位点其与在pIRES-EGFP的多克隆位点的BstX1位点是相匹配的。用一种Clontech HF Advantage cDNA PCR试剂盒进行PCR扩增, 使用厂商提供的多聚酶混合物及缓冲液。对于"补丁" PCR,所使用的修补引物的浓度为侧面引物的摩尔浓度的1/20。PCR扩增产品通过使用QIAquick PCR纯化试剂盒(Qiagen)来进行纯化。经过用限制性酶进行消化, 该产品在0.8%的琼脂糖凝胶中进行分离。然后用QIAquick 凝胶抽提试剂盒(Qiagen)进行纯化。

为了重新形成一种修饰的APP695-Sw cDNA, APP695-Sw cDNA的5'端Not 1-Bgl2片段及获自PCR的APP695 cDNA的3'端Bgl2-BstX1 片段被共同接入在Not1及BstX1位点打开的pIRES-EGFP质粒DNA中。使用快速DNA结合试剂盒(Boehringer Mannheim)在室温下结合5分钟,然后转化进入一种Library Efficiency DHSa Competent细胞中去(GibcoBRL Life Technologies)。使用引物#276 及#275通过PCR扩增对细菌菌落进行筛选插入。使用一种QIAprep Spin Miniprep试剂盒(Qiagen) 进行哺乳动物细胞转染使得质粒DNA纯化。所获得的构建体被命名为pMG125.3 (APPSW-KK, SEQ ID

No.17[核苷] 及SEQ ID No. 18 [氨基酸])。

哺乳动物细胞转染:

用于转染的HEK293细胞在Dulbecco's改良的Eagle's培养基 (DMEM)中经过使用10%牛胎血清被培养至80%汇合点处, 使用LipofectAmine (Gibco-BRL)进行共转染。用3μg的pMG125.3 DNA 及9μg的pcDNA3.1 DNA每10×10^6个细胞。转染后三天,细胞在400μg/ml的浓度下经过一含有G418的培养基。在选择性培养基中经过三天生长, 细胞因其荧光被分类筛选。

通过FACS克隆选择125.3细胞:

在一种为空调氩激光所提供的装备有488 nm激发线的EPICS Elite ESP流体细胞记数器(Coulter, Hialeah, FL)中对细胞样品进行分析, EGFP的释放可以通过一种525 nm的袋状过滤膜进行测量, 经过向前或向右角度的光散射来调控该成活细胞后, 荧光强度被展示于一种4的十进制的对数的范围内。单一的绿色细胞被分离至一种96孔平板(其含有生长培养基并没有G418)的每一孔中。经过一种4天的恢复期, G418被加入到该培养基中至一最终浓度为400μg/ml。经过选择, 32%的孔含有扩展的菌落。具有菌落的孔从96孔平板上转移到一种24孔平板上, 然后至一具有快速生长菌落(其被选于在每一通道处进行扩展)的6孔平板上。最终所选择的细胞系为以最快速度生长的最后6个生长的细胞)。该菌落, 其被命名为125.3,已被保存在G418以400 μg/ml的浓度, 每4天换入一新鲜的培养基。经过23次传代后, 并未发现EGFP 荧光的Aβ产品有任何损失。

Aβ EIA 分析 (双重抗体夹心ELISA分析hAβ 1-40/42):

经过转染后48小时, 收集细胞培养物上清液, 并以下述标准的Aβ EIA方法进行分析。在一种双抗体夹心ELISA中, 人类A1-40或者1-42使用单

克隆抗体(mAb) 6E10 (Senetek, St. Louis, MO)及生物素化的兔抗血清162或者164 (纽约州立基础研究所, Staten Island, NY)进行测量。捕获抗体6E10特异性针对呈现其在hAβ N端氨基酸残基1-16上的抗原表位。缀合的检测抗体162及164, 其分别特异于hAβ140及142。简而言之, 一种Nunc Maxisorp 96孔免疫平板上涂覆100μg/孔的mAb 6E10(5μg/ml), 稀释于0.1M碳酸盐-碳酸氢盐缓冲液, pH 9.6 及在40℃下过夜培养。通过用含有 0.05% of Tween-20 (DPBST)的3倍的0.01M DPBS (来自Pierce, Rockford, II Modified Dulbecco's磷酸盐缓冲盐水(0.008M磷酸钠, 0.002M磷酸钾, 0.14M氯化钠, 0.01 M 氯化钾, pH 7.4))洗涤平板,该平板用200μl的10%的在0.01M DPBS中的正常绵羊血清(Sigma)封闭60分钟, 以避免非特异性结合。洗涤平板后加入人类Aβ1-40或者1-42的标准品100μl/孔(Bachem, Torrance, CA)(由在DMSO中的1mg/ml的储备溶液用培养基稀释)及100μl/孔的样品, 例如, 转染细胞的条件培养基。该平板在室温下培养2小时及4℃下过夜保存。 第二天, 在洗涤平板后, 加入100μg/孔的生物素化的兔抗血清162 1:400或者164 1:50稀释于DPBST + 0.5% BSA, 在室温下培养1小时15分钟。洗涤后, 加100μg/孔的中性抗生物素蛋白-辣根过氧化物酶(Pierce, Rockford, II)(在DPBST中稀释10,000倍), 在室温下培养1小时。经过最后一次洗涤100μg/孔的邻苯二胺二盐酸化物(Sigma Chemicals, St. Louis, MO), 在50mM的柠檬酸/100mM磷酸钠缓冲液(Sigma Chemicals, St. Louis, MO), pH 5.0中, 作为底物加入, 在一种运动的微量平板阅读仪上使用Soft max Pro软件, 监控颜色变化20分钟。所有的标准品和样品都以一式三份操作。吸收值落入标准曲线内的样品通过使用Soft max Pro软件由标准曲线进行外推且以pg/ml培养基来表示。

结果:

在APP695的C末端增加两个赖氨酸残基可以极大地提高在HEK293细胞中的Aβ处理过程, 正如瞬时表达(表1)所显示的那样。在APP695中增加二赖氨酸基序可使Aβ加工提高到在含有瑞典突变的APP695中所看到的那

样。将二赖氨酸基序与瑞典突变结合可以将处理过程再提高2.8倍。

HEK293细胞与pMG125.3及pcDNA3.1的共转染可以允许对于G4 18抗性的转化细胞的双重选择及EGFP的高水平表达。通过用FACS进行克隆选择，经过在24孔平板上培养36小时所获得的细胞系产生了一种极高的在每毫升的培养基中的20,000 pg的Aβ肽。在不同生长条件下Aβ肽的产量总结于表2中。

表1. HEK293细胞经过与含有野生型或者修饰的APP的指定载体进行瞬时转染后48小时Aβ肽向培养基中的释放量。表列值为使用学生t-测试来估计不等方差的逐对比较的平均值+ SD 及P-值。

APP结构	Aβ1-40 多肽 (pg/ml)	提高的倍数	P-值
pIRES-EGFP载体	147+28	1.0	
wtAPP695(142.3)	194+15	1.3	0.051
wtAPP695-KK(124.1)	424+34	2.8	$3x\ 10^{-5}$
APP695-Sw (143.3)	457 + 65	3.1	$2\ x\ 10^{-3}$
APP695-SwKK (125.3)	1308+98	8.9	$3\ x\ 10^{-4}$

表2. 在不同生长条件下Aβ肽从HEK293细胞中的释放量

培养平板类型	培养基体积	培养时间	Ab1-40 (pg/ml)	Ab1-42 (pg/ml)
24孔平板	400μl	36小时	28,036	1,439

实施例 7:在HEK125.3 细胞中加工Aβ肽的反义低聚物抑制

为了选择目标序列及使用专利技术(Sequitur Ver. D Pat pending #3002)设计第二代嵌合反义低聚体，向Sequitur, Inc (Natick, MA)提

供Hu-Asp1及Hu-Asp2序列。批号为S644，S645，S646 及S647的反义低聚体被作为Asp1的目标靶，批号为S648，S649，S650 及 S651的反义低聚体被作为Asp2的目标靶。批号为S652，S653，S655,及S674的对照反义低聚体被作为不相关基因的目标靶，批号为S656，S657,S658,及S65的反义低聚体被作为另一种不相关基因的目标靶。

为了用反义低聚体进行转染，HEK125.3细胞被培养至以补充有10%牛胎血清的最少必需量的培养基（MEM）的6孔平板的50%覆盖度。2mg/ml oligofectin G (Sequitur Inc.，Natick，MA)的储备液在无血清MEM中被稀释至50 μg/ml。单独地，在Opti-MEM (GIBCO-BRL，GrandIsland，NY)中，100 μM的反义低聚体储备液被稀释至800 nM。oligofectin G及反义低聚体被稀释的储备液以1:1的比例进行混合，然后在室温下进行温育。经过15分钟后，该试剂被稀释10倍进入含有10%牛胎血清的MEM中， 经过第一次转移掉旧的培养基后，在六孔平板上每一孔加入上述2 ml稀释的试剂。转染后，细胞依然在oligofectin Glantisense低聚体的存在下进行继续培养。为了调控Ab多肽的释放，从转染结束后24小时开始，400μl条件培养基定期从培养孔中移去，取而代之以新鲜的培养基。所报道的数据来自培养结束后48小时收获的培养物上清液。

结果:

获自Sequitur Inc的16个不同的反义低聚体分别被转染进入HEK125.3细胞以测定其对Aβ肽处理过程所产生的效果。只有标定Asp1及Asp2的反义低聚体减少了Aβ肽的处理过程，其中含有标定Asp2的反义低聚体的HEK125.3细胞具有更强的抑制效果。Aβ(1-40)及Aβ(1-42)都以相同的程度被抑制。在表3中,百分比抑制度可以通过与未转染细胞相比较来计算。提供大于50%抑制度的反义低聚体试剂用一种星号来进行标记。在所测试的试剂中，标定Asp1的反义低聚体的4个中的3个给出了一种对于Aβ(1-40)处理的平均52%的抑制度，及对于Aβ(1-42) 处理的平均47%的抑制度。对于ASP2，4个反义低聚体中的4个给出大于50%的抑制度，而对于Aβ(1-40)

处理的平均62%的抑制度，及对于Aβ(1-42)处理的平均60%的抑制度。

表3. 用反义低聚体处理的HEK125.3细胞中释放的Aβ肽的抑制

靶向的基因	反义低聚体	Aβ (1-40)	Aβ(1-42)
Asp1-1	S644	62%*	56%*
Asp1-2	S645	41%*	38%*
Asp1-3	S646	52%*	46%*
Asp1-4	S647	6%	25%
Asp2-1	S648	71%*	67%*
Asp2-2	S649	83%*	76%*
Asp2-3	S650	46%*	50%*
Asp2-4	S651	47%*	46%*
Con1-1	S652	13%	18%
Con1-2	S653	35%	30%
Con1-3	S655	9%	18%
Con1-4	S674	29%	18%
Con2-1	S656	12%	18%
Con2-2	S657	16%	19%
Con2-3	S658	8%	35%
Con2-4	S659	3%	18%

实施例8: 在所培养细胞中Hu-Asp2β分泌酶活性的证明

与早期发作阿尔茨海默氏疾病相关的APP的几个突变已经表现出可改变了Aβ肽的处理过程。其使得N-端及C-端切割位点侧平以从APP释放出Aβ肽。这些切割位点分别标为β-分泌酶及γ-分泌酶切割位点。APP在β-分泌酶切割位点的切割释放出一种具有11,145道尔顿分子量的99个氨基酸的APPC末端片段。紧接β-分泌酶切割位点上游的瑞典突变KM→NL导致

了Aβ肽的1-40及1-42的氨基酸结构的产量的提高。伦敦VF突变(在APP770同工型中V717→F)对于总的Aβ肽的产量却鲜有影响，但是在APP处理过程中通过影响所使用的γ-分泌酶切割位点的选择优先提高了Aβ肽较长的1-42氨基酸结构的百分比含量，这样，我们应该测定是否这些突变改变了由一与介导Hu-Asp2表达的结构共转染的培养细胞所产生的Aβ肽的数量和类型。

做了两个实验其证明了在培养细胞中Hu-Asp2β-分泌酶的活性。在第一种实验中，用介导如实施例7所述的Hu-Asp2转录子的反义低聚体处理HEK125.3细胞，结果发现其减少了由β-分泌酶(CTF99) (图9)切割所形成的APPC端片段的数量。这也显示了Hu-Asp2直接或间接地有利于β-分泌酶的切割。在另一实验中，在已转染的鼠Neuro2A细胞中Hu-Asp2的扩增表达表现出增加了CTF99β-分泌酶切割片段(图10)的积累。当一含有一种C端两个赖氨酸基序的突变的APP-KK菌落被用于转染时该种增加可以很容易被发现。当Hu-Asp2与含有瑞典突变KM→NL的APP-Sw-KK共转染可以看到其进一步的提高。瑞典突变已知可以提高β-分泌酶对APP的切割。

另一种实验证明了Hu-Asp2有利于γ-分泌酶与人类胚胎肾HEK293细胞共转染的实验中的活性。Hu-Asp2与一种APP-KK菌落的共转染极大地提高了来自HEK293细胞中的可溶性Aβ 1-40及Aβ 1-42多肽的产量及释放量。Aβ 1-42多肽更以成倍增加的量释放。Aβ 1-42多肽产量的进一步提高可以通过Hu-Asp2与APP-VF (SEQ D No.13 [核苷] 及 SEQ D No.14 [氨基酸]) 或者含有伦顿突变V717→F的APP-VF-KK SEQ D No.19 [核苷] 及SEQ D No.20 [氨基酸])菌落的共转染看到。V717→F突变已知可以改变APPγ-分泌酶切割的特异性以至于在Aβ 42位点切割的可能性增加。这样，Asp2直接或间接地有利于γ-分泌酶在β 42位点对APP的切割。

材料:

抗体6E10及4G8购自Senetek (St. Louis, MO)。抗体369获自

Rockefeller大学的Paul Greengard实验室。抗体C8获自哈佛医学院及伯明翰妇女医院的Dennis Selkoe实验室。

所使用的APP构建体:

用于转染实验的APP构建体包括下列:

APP	野生型APP695 (SEQ D No.9及 No.10)
APP-Sw	含有瑞典KM→NL突变的APP695 (SEQ D No.11及 No.12),
APP-VF	含有伦顿突变V→F的APP695 (SEQ ID No.13 及No.14)
APP-KK	含有C末端KK基序的APP695 (SEQ D No.15及No. 16),
APP-Sw-KK	含有C末端KK基序的APP695-Sw (SEQ D No.17及 No.18),
APP-VF-KK	含有C末端KK基序的APP695-VF (SEQ ID No.19及 No.20).

通过PCR用合适的连接序列将上述结构在Not1及BstX1位点处插入到载体pIRES-EGFP (Clontech, Palo Alto CA)中。

在HEK293细胞, HEK125.3细胞及Neuro-2A细胞中的反义低聚体或质粒DNA结构的转染:

人类胚胎肾细胞HEK293细胞及鼠Neuro-2a细胞使用来自Gibco/BRL的Lipofectamine Plus试剂用表达构建体进行转染, 细胞被培养在24孔组织培养平板上至70-80%汇合处的密度。其中每一平板上有4孔用2 μg DNA (3:1, APP:共转染子)进行转染, 8μl Plus试剂, 及4μl在OptiMEM下的ipofectamine 加入到一种总体积为1 ml, 平均每孔分配有200μl的体积, 然后温育3小时。小心保持两种用于共转染的质粒的比例及在转染过程中的总的DNA的数量。转染培养基用DMEM, 10%FBS, NaPyruvate, 及抗生素/抗真菌剂等来替代。细胞在正常条件下进行培养(37℃, 5% CO_2) 48小时。该条件培养基被移至一聚丙烯试管中, 在-80°C下储存直到如前实施例所

述的EIA分析Aβ1-40 及Aβ1-42 的内容物。如实施例7所述那样反义低聚体转染进入HEK125.3细胞。

细胞提取物的制备，蛋白质印迹方法:

经过与质粒DNA转染60小时后收获细胞。首先，细胞被转移进入一种平板上的15-ml的圆锥试管，在1,500 rpm下离心5分钟以除去培养基。细胞沉淀用PBS洗涤一次。我们然后用细胞溶解缓冲液(10 mM HEPES，pH 7.9，150 mM NaCl，10%甘油，1 mM EGTA，1 mM EDTA，0.1 mM钒酸钠及 1% NP-40)对细胞进行溶解。然后将裂解后的细胞混合物在5000 rpm下离心，上清液作为提取物被储存在-20°C温度下。来自用Asp2反义低聚体转染的HEK125.3细胞的同等量的提取物，对照实验与抗体369（其可以识别APP的C端）一起离心。然后与抗体6E10一起进行免疫沉淀反应检测出CTF99。使用C8（同样可以识别APP的C末端的另一沉淀抗体）进行重复实验。对于用Hu-Asp2 及APP-KK，APP-Sw-KK，APP-VF-KK 或者APP-VF进行共转染的鼠类Neuro-2a 细胞提取物的蛋白质印迹实验，同等量的细胞提取物经过4-10%或10-20% Tricine梯度凝胶(NOVEX，San Diego，CA)下进行电泳。全长APP及CTF99 β分泌酶产品使用抗体6E10可以检测到。

结果:

与用具有反向序列(Asp2- 1反向及Asp2-2 反向)的对照寡聚体类似转移的细胞相比，用Asp2-1或者Asp2-2反义低聚体转染HEK125.3细胞可减少CTFβ分泌酶产物的产量。在共转染的实验中，Hu-Asp2与APP-KK构建体共转染进入鼠Neuro-2a细胞增加了CTF99的形成量。如果Hu-Asp2与APP-S w-KK（含有可以提高β分泌酶处理的瑞典突变KM→NL的APP的一种变异体）进行共表达，CTF99的形成量可以进一步增加。

Hu-Asp2与APP的共转染对于Aβ40产量鲜有影响但在上述条件下却增加了Aβ42产量(表4)。在APP的C末端增加两个赖氨酸基序将Aβ肽的处理过程提高了2倍,尽管Aβ40及Aβ42的产量保持在较低的水平(分别为352 pg/ml及21pg/ml,)。Asp2与APP-KK的共转染进一步提高了Aβ40及Aβ42的产量。Hu-Asp2刺激Aβ40使其产量提高了至少3倍,而使得Aβ42使其产量提高了至少10倍。这样, Hu-Asp2与APP-KK构建体的共转染优先提高了Aβ42的产量。

APP的V717→F突变已被显示增加了γ-分泌酶在Aβ42切割位点的处理进程。Hu-Asp2与APP-VF 或APP-VF-KK结构的共转染增加了Aβ42的产量(用APP-VF提高了2倍, 用APP-VF-KK提高了4倍, 表4),但是对于Aβ40的产量有混合效果(使用APP-VF有轻微的减少,使用APP-VF-KK将其产量提高了2倍, 相比较pcDNA共转染对照组)。这样, Asp2对于Aβ42产量的效果按比例增加结果导致Aβ42/总的Aβ的比例的增加。事实上, Aβ42/总的Aβ的比例在HEK293细胞(该细胞用Hu-Asp2及APP-VF-KK进行共转染)中达到了一种非常高的值(42%)。

蛋白质印迹显示HEK125.3细胞(用靶定Hu-Asp2 mRNA的反义低聚体进行转染的细胞).所产生的CTF99产量的减少。(右)蛋白质印迹显示出鼠Neuro-2a细胞(其与Hu-Asp2及APP-KK进行共转染)所产生的CTF99产量的增加。CTF99产量的进一步提高在Hu-Asp2与APP-Sw-KK共转染的细胞中也可以发现)。

表-4. Hu-Asp2或者pcDNA质粒DNA与具有V717→F突变(其修饰了γ-分泌酶的处理过程)的不同的APP结构的共转染的结果, 与Asp2的共转染持续增加了Aβ42/总的Aβ的比例, 表列值为Aβ肽pg/ml.

	pcDNA 共转染			Asp2 共转染		
	Aβ40	Aβ42	Aβ42/总量	Aβ40	Aβ42	Aβ42/总量
App	192+18	<4	<2%	188+40	8+10	3.9%
APP-VF	118+15	15+19	11.5%	85+7	24+12	22A%

APP-KK	352+24	21+6	5.5%	1062+101	226+49	17.5%
APP-VF-KK	230+31	88+24	27.7%	491+35	355+36	42%

实施例9: 人类Asp2L的细菌表达

重组Hu_Asp2L在大肠杆菌中的表达

加入N末端序列后Hu-Asp2L可以在大肠杆菌中进行表达，例如加入一种T7标记(SEQ ID No.21及No.22)或者后接一aspase 8前导序列(SEQ ID No.23及No.24)的T7标记。可以选择地，通过定点诱变使得5'端序列GC含量的减少可以用于增加Hu-Asp2 (SEQ ID No.25 及 No. 26)的产量。此外，可以用一蛋白酶水解位点构建Asp2 (SEQ ID No.27 及 No.28)。为了在表达及重折叠后产生一可溶性蛋白，跨膜区域及细胞质尾的缺失，或者临近膜区域，跨膜区域及细胞质尾的缺失是优选的。

方法:

使用含有合适连接序列的引物进行PCR，将Asp2编码序列与包括T7标记(SEQ ID Nos. 21 及 22) 或者一种T7-caspase 8前导序列(SEQ ID Nos. 23及24)的N端序列修饰物组装融合。这些构建体被克隆进一种表达载体pet23a(+) [Novagen]中其中一种T7启动子介导一种在一种多倍克隆位点之前的T7标记的表达。为了在pet23a+的T7前导序列之后克隆一种Hu-Asp2序列，下述寡聚核苷酸被用于对所选择的Hu-Asp2序列进行扩增:
#553=GTGGATCCACCCAGCACGGCATCCGGCTG (SEQ ID No.35),
#554=GAAAGCTTTCATGACTCATCTGTCTGTG(3AATGTTG (SEQ ID No.36),
这些引物分别将BamHI及HindⅢ位点置于插入片段的5'及3'端。Asp2序列从全长的Asp2（b）cDNA（其按照厂商提供的方案使用在68°C下退火及延伸的两步PCR循环进行25个循环的Advantage-GC cDNA PCR [Clontech]被克隆进入pcDNA3.1）中进行扩增。插入子及载体都用BamHI及HindⅢ进行

切割，通过琼脂糖凝胶电泳进行纯化，然后用快速的DNA连接试剂盒进行连接[Boerhinger Mannheim]。连接反应被用于转化大肠杆菌菌株JM109(Promega)，挑选菌落用于质粒纯化(Qiagen,Qiaprep minispin)及DNA序列分析。对于使用异丙基-β-D-硫代吡喃半乳糖苷 (IPTG)进行诱导的可诱导性表达,表达载体被转化进入大肠杆菌菌株BL21 (Statagene)中。细菌培养物在有100 μg/ml氨苄青霉素存在的条件下在LB肉汤中进行培养。然后在对数生长期进行诱导，其使用OD600为0.6-1.0的1mM的IPTG在37°C下诱导4个小时。细胞沉淀物通过离心进行收获。

为了在T7标记及caspase前导序列(SEQ ID Nos. 23 及24)后克隆Hu-Asp2序列，将上面所述含有T7-Hu-Asp2 序列 (SEQ ID Nos. 21及22)的在BamHI位点打开。然后将磷酸化后的caspase 8 前导寡聚核苷酸。#559=GATCGATGACTATCTCTGACTCTCCGCGTGAACAGGACG (SEQ ID No.37)，#560=GATCCGTCCTGTTCACGCGGAGAGTCAGAGATAGTCATC (SEQ ID No.38) 退并连接至载体DNA。设计每一列寡聚核苷酸的5'端的突出端使得其允许连接进入BamHI位点但使得BamHI 不能再对其进行消化。如上分析连接反应转化进入大肠杆菌菌株系BL21后的蛋白质表达，连接反应也可转化进入JM109细胞。

为了减少Asp2 5'端的GC含量，设计了一对反平行低聚体以改变退化密码子碱基，在从G/C至A/T (SEQ ID Nos. 25 及 26).的大约15个氨基酸的位置。在Asp2 5'末端的新的核苷酸序列并未改变所编码的氨基酸，选用其来优化为大肠杆菌的表达。有义连接子序列为5' CGGCATCCGGCTGCC CCTGCGTAGCGGTCTGGGTGGTGCTCCACTGGGTCTGCG TCTGCCCCGGGAGACCGACGAA G 3' (SEQ ID No.39). 反义连接子序列为5'CTTCGTCGGTCTCCCGGGGCAGACGC AGACCCAGTGGAGCACCACCCAGACCGCTACGCAGGGGCAGCCGGATGCCG 3' (SEQ ID No.40)。在0.1 M NaCl-10 mM Tris, pH7.4下经过退火磷酸化的连接子其被连接进入一种位于在载体pTAC 上的Hu-Asp2的唯一的Cla I及Sma I位点。对于使用异丙基-β-D-硫代吡喃乳糖苷(IPTG)进行诱导的可诱导性表达。细菌培养物在有100μg /ml氨苄青霉素存在的条件下在LB肉汤中进行培养。然后在对数生长期进行诱导，其使用OD600为0.6-1.0的1mM

的IPTG在37°C下诱导4个小时。通过离心收获细胞沉淀。

为了产生一种载体，其中前导序列可以通过使用caspase 8 进行限制性蛋白水解而移去，这样就构建成以N末端序列GSFV (SEQ ID Nos. 27 及28)开始Hu-Asp2多肽文库，随后进行下面的程序，两个含有caspase 8切割位点IETG的磷酸化寡聚核苷，#571=5'GATCGATGACTATCTCTGACTCTC CGCTGGACTCTGGTATCGAAACCGACG (SEQ ID No.41)和#572=GATCCGTCGGT' TTCGATACCAGAGTCCAGCGGAGAGTCAGAGATAGTCATC (SEQ ID No.42)被退火且连接进入pET23a+（其已经被BamHI切开）。经过转化进入JM109,纯化载体DNA被恢复，通过DNA序列分析证实插入子的方向。+，下面的寡聚核苷被用于扩增所选择的Hu-Asp2序列:

#573=5'AAGGATCCTI ~ GTGGAGATGGTGGACAACCTG , (SEQ ID No.43) #554=GAAAGCTTI'CATGACTCATCTGTCTGTGGAATGTTG (SEQ ID No.44)，其分别将BamHI及HindⅢ位点侧面接入插入子的5'及3'端。Asp2序列从全长的Asp2cDNA（其按照厂商提供的方案使用在68℃下退火及扩展的两步PCR循环进行25个循环的Advantage-GC cDNA PCR [Clontech]被克隆进入pcDNA3.1）中进行扩增。插入子及载体都用BamHI及HindⅢ进行切割，通过琼脂糖凝胶电泳进行纯化，然后用快速的DNA连接试剂盒进行连接[Boerhinger Mannheim]。连接反应被用于转化大肠杆菌菌株JM109 (Promega)，精选菌落以纯化质粒(Qiagen,Qiaprep minispin)及进行DNA序列分析。对于使用异丙基-β-D-硫代吡喃半乳糖苷(IPTG)进行诱导的可诱导性表达,表达载体被转化进入大肠杆菌菌株BL21 (Statagene)中。细菌培养物在有100 μg/ml氨卞青霉素存在的条件下在LB肉汤中进行培养。然后在对数生长期进行诱导，其使用OD600为0.6-1.0的1mM的IPTG在37℃下诱导4个小时。细胞沉淀物通过离心进行收获。

为了帮助纯化，一种6-His标记通过在BamHI位点对构建体进行切割而被引入任一上述构建体的在T7前导序列之后的位点，然后在退火，及对含有六个组氨酸序列#565=GATCGCATCATCACCATCACCATG(SEQ ID No.45), #566=GATCCATGGTGATGGTGATGATGC (SEQ ID No.46)的寡聚核苷酸进行磷酸化等条件下进行连接。设计每一列寡聚核苷酸的5'端的突出端使得其

允许连接进入BamHI位点但结果并不使得BamHI 对其进行消化。

制备细菌小球:

10.8L的36.34g细菌小球被扩散进入一种总体积为200ml,使用一种20mm组织均化器探子在3000 至5000 rpm下在2M KCl, 0.1M Tris, 0.05M EDTA, 1mM DTT进行均化。用水调节其电导率为193mMhos。

细胞小球被扩散后,加入额外的KCl溶液,使其总体积变为500 ml。悬浮液然后在5000 rpm下使用相同的探针进一步均化3小时,混合液然后通过一种在10,000psi 下的Rannie高压均化器。

在所有情况下,小球材料被收集,而溶解的片段被丢弃。结果的溶液在一种GSA转轴下在12,500 rpm下进行离心1小时。小球在同样的溶液中使用相同的组织均化器探子在2000 rpm下进行重悬浮(没有DTT)。然后在3000 rpm下均化5分钟后,在相同的溶液条件下其体积被调整至500ml。然后在12,500 rpm下离心1小时。如前一样重悬浮小球,但这时候最终体积被调整至1 .5L,然后均化5分钟。经过在相同的速度下离心30分钟后,重复该程序。小球然后在大约150ml的冷水中重悬浮,集中来自六个离心试管(其使用GSA转轴)的小球。然后在3,000 rpm下均化5分钟,用冷水调整其体积为250ml ,然后离心30分钟,生成的小球重量为17.75g。

概述: 细菌小球在KCl溶液中进行溶解,在此之前以一种GSA转轴进行离心以开始制备小球。然后使用相同的溶液3次进行重悬浮/均化。最后使用水洗/均化以除去多余的KCl及EDTA。

rHuAsp2L的可溶解性:

一种比率为9- 10ml/克的细胞小球被用于溶解来自先前所述小球的rfluAsp2L。解冻17.75g小球,然后加入150ml 8M胍 HCl, 5mM βME, 0.1% DEA。3M Tris被用于调整pH至8.6。使用一20mm的组织均化器探子在1000rpm下开始重悬浮进入胍溶液中,混合液然后在4°C下搅动1小时,在12,500 rpm

下以GSA转子离心1小时。得到的上清液再在40,000 x g以一种SS-34转子离心30分钟。除了50ml外，最终的上清液储存于-20°C下。

溶解的rHuAsp2L的固定化镍亲合色谱：

使用下列溶液：

A) 6M 脉 HCI, 0. 1M NaP, pH 8.0, 0.01M Tris, 5mM βME, 0.5mM咪唑

A') 6M 尿素, 20mM NaP, pH 6.80, 50mM NaCl

B') 6M 尿素, 20mM NaP, pH 6.20, 50m M NaCl, 12mM咪唑

C') 6M 尿素, 20mM NaP, pH 6.80, 50m M NaCl, 300mM咪唑

注意：缓冲液 A' 及 C' 以适当的比例混合以提供一种咪唑的中间浓度。

50ml的溶解材料与50ml的缓冲液A进行混合，然后加入一种100-125ml的 Qiagen Ni-NTA Superflow (用缓冲液A进行预平衡)至一种5 x 10cm的 Bio-Rad 的econo柱。然后在4°C的冷室中过夜平稳振荡。

色谱步骤：

1) 排去生成的流通液，
2) 用50ml缓冲液A进行洗涤(收集到流通液级分中)
3) 用250ml缓冲液A进行洗涤(洗涤1)
4) 用250ml缓冲液A进行洗涤(洗涤2)
5) 用250ml缓冲液A'进行洗涤
6) 用250ml缓冲液B'进行洗涤
7) 用250ml缓冲液A'进行洗涤
8) 用250ml 75mM 咪唑进行洗脱
9) 用250ml 150mM 咪唑进行洗脱(150-1)
10) 用250ml 150mM 咪唑进行洗脱 (150-2)
11) 用250ml 300mM 咪唑进行洗脱 (300-1)
12) 用250ml 300mM 咪唑进行洗脱(300-2)

13) 用250ml 300mM 咪唑进行洗脱 (300-3)

色谱结果:

以75mM至300mM咪唑进行洗脱rHuAsp，75mM级分以及第一种150mM咪唑(150-1)级分含有污染蛋白如考马斯蓝染色凝胶上所显示的那样。因此，级分150-2及300-1将被用于进行重折叠实验，因为其含有最大量的蛋白(参见考马斯蓝染色凝胶)。

rHuAsp2L的重折叠实验:

实验1:

40ml的150-2 用1M DTT，3M Tris，pH 7.4 及DEA使其值升高至一最终浓度分别表示为6mM，50mM，及0.1%。其一面搅动一面用200 ml的(4°C)冷的20mM NaP，pH 6.8,150mM NaCl进行突然稀释。该稀释使得尿素的最终浓度为1M。该溶液保持清澈，即使在RT或4°C下开口于空气。

经过在4°C下开口于空气4-5小时，然后用20mM NaP，pH 7.4，150mM NaCl，20%甘油透析该溶液一夜。该方法有效地取除了溶液中的尿素而并不使蛋白沉降。

实验2:

一些150-2洗脱液在一种Amicon Centriprep 10,000 MWCO,中浓缩2倍。然后以实验1那样进行处理。该实验材料依然是溶液，且无明显的沉降。

实验3:

89ml的150-2洗出液用1M DTT, 3M Tris, pH 7.4及DEA使其值升高至一最终浓度分别为6mM, 50mM,及0.1%。其一面搅动一面用445ml的(4°C)冷的20mM NaP, pH 6.8,150mM NaCl进行突然稀释。该溶液保持清澈,且无明显沉淀。该溶液被移至RT且搅拌10分钟,然后加入MEA使其最终浓度变为0.1mM。在RT下缓慢搅拌一小时。加入胱胺及$CuSO_4$使最终浓度分别为1mM及10μM。该溶液在RT下缓慢搅拌10分钟。然后被移至4°C的冷室中过夜缓慢振荡且开口于空气。

第二天,溶液(依然清澈,且无明显沉淀)在100,000 x g下离心1小时。收集多次离心的上清液,用20mM NaP, pH 7.4, 150mM NaCl, 20%甘油对稳定的蛋白进行透析。透析后,该物质被保存在-20°C下。

保存一些(大约10ml)的蛋白溶液(依然在1M的尿素中)以用于生化分析,在-20°C下冷冻储存。

实施例10: Hu-Asp2的表达及其在昆虫细胞中的衍生细胞

杆状病毒感染表达---- Hu-Asp2的编码序列及几个衍生物通过在昆虫细胞中使用PCR工程化。对于全长序列,修饰翻译起始位点以适合Kozak共有序列的5'-有义寡聚核苷引物与含有在Hu-Asp2 序列中的自然翻译终止密码子的3'-反义寡聚核苷引物进行配对。PCR扩增pcDNA3.1(hygro)/Hu-Asp2模板 (参见实施例12)。删去了C端跨膜区域(SEQ D No.29 及No.30)或者删去了跨膜区域而在C端(SEQ D No. 31及No.32)引入了一种六个组氨酸的标记的Hu-Asp2的两个衍生物同样可以使用PCR进行工程化。与上述相同的5'-有义寡聚核苷引物与下列任一进行配对: 或者一种3'-反义引物其(1) 在密码子453 (SEQ ID No.3)之后引入了一种翻译终止密码子或者一种3'-反义引物其(2)在使用pcDNA3.1(hygro)/Hu-Asp-2L作为模板进行PCR中翻译终止密码子之后有一种六个组氨酸的标记。在所有的情况下,按照厂商指导,使用PwoI DNA聚合酶(Boehringer-Mannheim)进行PCR反应扩增15个循环。反应产物用BamHI 及NotI 进行消化,然后连接至BamHI 及NotI消化的杆状病毒转移

载体pVL1393 (Invitrogen)。连接产品的一部分转化进入合适的E. coli DH5α细胞，然后在LB-Amp上使用抗生素进行选择。质粒DNA通过使用标准的碱性溶液进行溶解及在CsCl下连接以产生杆状病毒转移载体pVL1393/Asp2ΔTM及pVL I 393/ASp2ΔTM$(His)_6$。重组杆状病毒的产生及sf9昆虫细胞的感染使用标准方法进行操作。

感染表达----在High 5昆虫细胞中使用昆虫表达载体pIZ/V5-His进行Hu-Asp2ΔTM及Hu-Asp2ΔTM$(His)_6$ 的瞬时及稳定表达。来自于表达质粒载体pVL1393IAsp2, pVL1393/Asp2AΔTM 及pVL1393/ASp2ΔTM$(His)_6$的DNA插入子同时使用BamHI 及NotI 进行消化切割，然后使用标准方法将其亚克隆进入BamHI及NotI 消化的pIZ/V5-His。结果的表达质粒，指的是pIZ/Hu-Asp2ΔTM 及pIZ/Hu-Asp2ΔTM$(His)_6$,其以上述方法进行制备。

对于转染，High 5昆虫细胞在27 ^{0}C下在一种封闭的瓶中被培养在用10 μg/ml 庆大霉素补充的高五无血清培养基中，使用高五细胞进行转染，使用标准方法在高五无血清培养基中补充10 μg/ml 庆大霉素及Insectin Plus 脂质体 (Invitrogen, Carlsbad, CA)。

对于大规模瞬时转染， 1.2 x 10^7 高五细胞被放置于一种150 mm的组织培养盘中且允许在室温下固着15-30分钟。在固着期内，DNA/脂质体混合物可以通过混合6 ml无血清培养基，60 μg Asp2ΔTM/pIZ (+/- His) DNA及120 μl insectin Plus 进行制备，且在室温下温育15分钟。平板接种培养基从细胞碟上移去，且与DNA/脂质体混合物在室温下以2 rpm均匀振荡混合4小时。另外加入6 ml培养基至细胞碟上，然后在27 ^{0}C下在一潮湿的孵卵器中培养四天。转染后4天收获培养基，在500 x g下离心纯化，用蛋白质印迹法去测试Asp2的表达。对于稳定的表达，细胞用50 μg/ml Zeocin进行处理，通过限制性稀释，活细胞被用于制备克隆细胞，然后如上分析其表达程度。

Hu-Asp2ΔTM 及 Hu-Asp2ΔTM$(His)_6$ 的纯化-----从Hu-Asp2中除去跨膜片段结果使得多肽分泌进入培养基中。随着通过杆状病毒感染或转染生产蛋白，收获条件培养基，通过离心进行纯化，以Tris-HCl (pH 8.0)进行透析。该物质然后通过离子交换(Tris-HCl, pH 8.0)连续色谱进行

纯化，然后使用NaCl梯度进行阳离子交换色谱(pH 4.5下的醋酸缓冲液)。洗脱曲线可以通过下列进行绘制：(1) 蛋白质印迹及(2)使用如实施例12所述的多肽作用物进行活性分析。对于Hu-Asp2ΔTM$(His)_6$，条件培养基在Tris缓冲液下进行培养(pH 8.0)，然后通过在IMAC树脂下连续色谱进行纯化，然后进行阴离子交换色谱。

序列分析纯化的Hu-Asp2ΔTM$(His)_6$蛋白揭示信号肽已经被切除。[TQHGIRLPLR].

实施例11: 在CHO细胞中Hu-Asp2的表达

Hu-Asp-2L在CHO-K1细胞中的异源表达----- Hu-Asp2的全部编码序列被克隆进入哺乳动物细胞表达载体PcDNA3.1(+)Hygro (Invitrogen, Carlsbad, CA)(其含有一种紧在以驱动表达的启动子及bGH多腺苷酸化信号的前面的CMV)中。表达质粒，pcDNA3. 1(+)Hygro/Hu-Asp2,通过碱性溶解及在CsCl中进行连接来制备，两条链的完全序列化证实了编码序列的完整性。

野生型的中国仓鼠卵巢细胞(CHO-K1) 可以从ATCC获得。细胞被保存在α-MEM(其含有在37°C下，5% CO_2的10% FCS)单层细胞培养物上。两个100mm的CHO-K1细胞平皿(60% 融合) 仅用pcDNA3.1(+)/Hygro (仿制品)或pcDNA3.1(+)Hygro/Hu-Asp2使用阳离子性的脂质体DOTAP按照厂商所推荐那样进行转染。细胞用质粒DNA/脂质体混合物处理15小时，然后培养基使用含有500单位/ml潮霉素B的生长培养基进行替换。在pcDNA3.1(+)Hygro/Hu-Asp2转染CHO-K1细胞的情况下，每个潮霉素B抗性细胞通过限制性稀释进行克隆。随着每个细胞系的克隆表达，Hu-Asp2蛋白的表达通过使用一种多克隆兔抗血清提高重组的在大肠杆菌中进行表达的Hu-Asp2的蛋白质印迹分析获得。每一种细胞系接近融合的皿通过在PBS中经过离心恢复细胞获得。细胞沉淀物在冷的含有蛋白酶抑制剂的溶解缓冲液中(25 mM Tris-HCl(8.0)/5 mM EDTA)重悬浮。细胞通过超声波进行溶解。溶解的及膜级分通过离心(105,000 x g, 60 分钟)进行分离。

来自每一部分一定量的蛋白质通过SDS-PAGE进行分离。随着所分离多肽电转移至PVDF膜上，Hu-Asp-2L蛋白使用兔抗Hu-Asp2血清(1/1000稀释)进行检测，抗原抗体复合体使用碱性磷酸酶缀合羊抗兔抗体(1/2500)显现。一种具有表观平均分子量为65 kDa 的特异性免疫反应蛋白在pcDNA3.1(+)Hygro/Hu-Asp2转染细胞中进行检测(其无仿制细胞)。同样，Hu-Asp2多肽仅在膜片段中检测到，其与在推测序列中信号肽及信号跨膜区域的存在是一致的。基于上述分析， #5克隆具有Hu-Asp2蛋白最高的表达水平。这样产生的细胞系经过放大试验以提供所要纯化的物质。

来自CHO-K1/Hu-Asp2 #5克隆的重组Hu-Asp-2L的纯化-------- 在一种典型的纯化实验中，来自20个150 mm融合细胞平皿的#5克隆细胞沉淀物，被用作起始物。细胞沉淀物在如上所述那样在50 ml 冷的溶解缓冲液中重悬浮。细胞通过使用polytron均化器(2 x 20 sec)进行溶解，溶胞产物再在338,000 x g下离心20分钟。膜沉淀物然后在20ml的含有50 mM β-辛基葡糖苷的冷的溶解缓冲液中重悬浮。然后在4°C下振荡1小时。去污剂提取物然后通过在338,000 x g下离心20分钟而澄清，取上清液做进一步的分析。

β-辛基葡糖苷的提取物被应用一种Mono Q阴离子交换柱其先前已用25 mM Tris-HCl(pH 8.0)/50 mM β-辛基葡糖苷进行平衡。上样后，柱子用一种递增的NaCl浓度(30分钟内0-1.0 M)的线性梯度进行洗脱并由蛋白质印迹分析测定各个级分的β分泌酶活性。(参见下文)。 含有 Hu_Asp-2L 免疫活性及β分泌酶活性的级分被收集且在25 mM NaOAc (pH 4.5)/50 mM 的β-辛基葡糖苷下进行透析。透析后，沉淀物通过离心除去，可溶物在一种单S阳离子交换柱上进行色谱分析，其预先使用25 mM NaOAc (pH 4.5)/50 mMβ-辛基葡糖苷进行平衡。柱子用一种递增的NaCl浓度(30分钟内0-1.0 M)的线性梯度进行洗脱并由蛋白质印迹分析测定各个级分的β分泌酶活性。含有Hu-Asp-2 免疫活性及β分泌酶活性的级分被合并且使用SDS-PAGE/考马斯蓝染色测得其有大于90%的纯度。

实施例12: 使用多肽底物来分析Hu-Asp2 β分泌酶活性

β分泌酶的分析———β分泌酶活性可以通过量化含有APP瑞典突变（使用RP-HPLC及UV检测）的合成肽的水解来测量。每一种反应含有50 mM Na-MES (pH 5.5)，1%β-辛基葡糖苷，多肽作用底物 (SEVNLDAEFR，70 μM)及酶(1-5μg 蛋白)。反应在37 °C下温育几次，反应产物通过RP-HPLC使用一种线性梯度从0-70 B经过30分钟(A=0. 1% TFA在水中，B=0. 1 %TFA/10%水/90%AcCN)进行溶解。洗脱曲线通过在214 nm处的吸收来测绘。在预备实验中，有两个峰值其在多肽作用底物未反应前进行洗脱，同时使用Edman测序及MADLI-TOF质谱分析法证实其具有序列DAEFR及SEVNL。 多肽作用底物的水解百分比可以通过比较两个产品肽完整的峰面积及在214 nm处吸收所取得的起始物质的峰面积而计算出来。蛋白酶裂解反应的特异性可以通过在蛋白酶抑制剂混合物(8μM 抑肽素A，10μM亮肽素，10μM E64，及5 mM EDTA)的存在下进行β分泌酶试验而测定。

一种可供选择的β分泌酶试验使用内部淬火荧光底物在一单一样品或多孔形式下使用荧光色谱来检测酶的活性，每一反应含有50 mM Na-MES (pH 5.5)，多肽作用底物MCA-EVKMDAEF[K-DNP] (BioSource International) (50 μM)及纯化的Hu-Asp-2 酶。该组分被平衡至37 °C几次，反应通过加入作用底物而激活。在330 nm处激活，反应动力学通过测量在390 nm处的荧光释放而计算。为了检测调节Hu-Asp-2活性的组合物，该待测组合物在反应的预培养阶段加入，反应动力学如上进行计算。催化剂活性以提高表观荧光的速率的组合物来计算而抑制剂则以减少表观荧光的速率的组合物来计算。

已经非常清楚本发明如前面说明书及实施例所述是具有实用性的。本发明的大量的修饰和变化都可能是以本发明前文所述为基础或提示的，因此，都在本发明的保护范围内。

这儿所引用的全部的出版物的全文都作为参考资料而纳入本说明书。

序列表

<110> Gurney, Mark E.
Bienkowski, Michael J.
Heinrikson, Robert L.
Parodi, Luis A.
Yan, Riqiang
Pharmacia & Upjohn Company

<120> 阿尔茨海默氏疾病分泌酶

<130> 6177.P CP

<140>
<141>

<150> 60/101,594
<151> 1998-09-24

<160> 49

<170> PatentIn Ver. 2.0

<210> 1
<211> 1804
<212> DNA
<213> Homo sapiens

<400> 1
atgggcgcac tggcccgggc gctgctgctg cctctgctgg cccagtggct cctgcgcgcc 60
gccccggagc tggcccccgc gcccttcacg ctgcccctcc gggtggccgc ggccacgaac 120
cgcgtagttg cgcccacccc gggacccggg acccctgccg agcgccacgc cgacggcttg 180
gcgctcgccc tggagcctgc cctggcgtcc cccgcgggcg ccgccaactt cttggccatg 240
gtagacaacc tgcaggggga ctctggccgc ggctactacc tggagatgct gatcgggacc 300
cccccgcaga agctacagat tctcgttgac actggaagca gtaactttgc cgtggcagga 360
accccgcact cctacataga cacgtacttt gacacagaga ggtctagcac ataccgctcc 420
aagggctttg acgtcacagt gaagtacaca caaggaagct ggacgggctt cgttggggaa 480
gacctcgtca ccatccccaa aggcttcaat acttcttttc ttgtcaacat tgccactatt 540
tttgaatcag agaatttctt tttgcctggg attaaatgga atggaatact tggcctagct 600
tatgccacac ttgccaagcc atcaagttct ctggagacct tcttcgactc cctggtgaca 660
caagcaaaca tccccaacgt tttctccatg cagatgtgtg gagccggctt gcccgttgct 720
ggatctggga ccaacggagg tagtcttgtc ttgggtggaa ttgaaccaag tttgtataaa 780
ggagacatct ggtatacccc tattaaggaa gagtggtact accagataga aattctgaaa 840
ttggaaattg gaggccaaag ccttaatctg gactgcagag agtataacgc agacaaggcc 900
atcgtggaca gtggcaccac gctgctgcgc ctgccccaga aggtgtttga tgcggtggtg 960
gaagctgtgg cccgcgcatc tctgattcca gaattctctg atggtttctg gactgggtcc 1020
cagctggcgt gctggacgaa ttcggaaaca ccttggtctt acttccctaa aatctccatc 1080
tacctgagag atgagaactc cagcaggtca ttccgtatca caatcctgcc tcagctttac 1140

attcagccca tgatgggggc cggcctgaat tatgaatgtt accgattcgg catttcccca 1200
tccacaaatg cgctggtgat cggtgccacg gtgatggagg gcttctacgt catcttcgac 1260
agagcccaga agagggtggg cttcgcagcg agcccctgtg cagaaattgc aggtgctgca 1320
gtgtctgaaa tttccgggcc tttctcaaca gaggatgtag ccagcaactg tgtccccgct 1380
cagtctttga gcgagcccat tttgtggatt gtgtcctatg cgctcatgag cgtctgtgga 1440
gccatcctcc ttgtcttaat cgtcctgctg ctgctgccgt tccggtgtca gcgtcgcccc 1500
cgtgaccctg aggtcgtcaa tgatgagtcc tctctggtca gacatcgctg gaaatgaata 1560
gccaggcctg acctcaagca accatgaact cagctattaa gaaaatcaca tttccagggc 1620
agcagccggg atcgatggtg gcgctttctc ctgtgcccac ccgtcttcaa tctctgttct 1680
gctcccagat gccttctaga ttcactgtct tttgattctt gattttcaag ctttcaaatc 1740
ctccctactt ccaagaaaaa taattaaaaa aaaaacttca ttctaaacca aaaaaaaaaa 1800
aaaa 1804

<210> 2
<211> 518
<212> PRT
<213> Homo sapiens

<400> 2
Met Gly Ala Leu Ala Arg Ala Leu Leu Leu Pro Leu Leu Ala Gln Trp
1 5 10 15

Leu Leu Arg Ala Ala Pro Glu Leu Ala Pro Ala Pro Phe Thr Leu Pro
20 25 30

Leu Arg Val Ala Ala Ala Thr Asn Arg Val Val Ala Pro Thr Pro Gly
35 40 45

Pro Gly Thr Pro Ala Glu Arg His Ala Asp Gly Leu Ala Leu Ala Leu
50 55 60

Glu Pro Ala Leu Ala Ser Pro Ala Gly Ala Ala Asn Phe Leu Ala Met
65 70 75 80

Val Asp Asn Leu Gln Gly Asp Ser Gly Arg Gly Tyr Tyr Leu Glu Met
85 90 95

Leu Ile Gly Thr Pro Pro Gln Lys Leu Gln Ile Leu Val Asp Thr Gly
100 105 110

Ser Ser Asn Phe Ala Val Ala Gly Thr Pro His Ser Tyr Ile Asp Thr
115 120 125

Tyr Phe Asp Thr Glu Arg Ser Ser Thr Tyr Arg Ser Lys Gly Phe Asp
130 135 140

Val Thr Val Lys Tyr Thr Gln Gly Ser Trp Thr Gly Phe Val Gly Glu
145 150 155 160

```
Asp Leu Val Thr Ile Pro Lys Gly Phe Asn Thr Ser Phe Leu Val Asn
                165                 170                 175

Ile Ala Thr Ile Phe Glu Ser Glu Asn Phe Phe Leu Pro Gly Ile Lys
            180                 185                 190

Trp Asn Gly Ile Leu Gly Leu Ala Tyr Ala Thr Leu Ala Lys Pro Ser
        195                 200                 205

Ser Ser Leu Glu Thr Phe Phe Asp Ser Leu Val Thr Gln Ala Asn Ile
    210                 215                 220

Pro Asn Val Phe Ser Met Gln Met Cys Gly Ala Gly Leu Pro Val Ala
225                 230                 235                 240

Gly Ser Gly Thr Asn Gly Gly Ser Leu Val Leu Gly Gly Ile Glu Pro
                245                 250                 255

Ser Leu Tyr Lys Gly Asp Ile Trp Tyr Thr Pro Ile Lys Glu Glu Trp
            260                 265                 270

Tyr Tyr Gln Ile Glu Ile Leu Lys Leu Glu Ile Gly Gly Gln Ser Leu
        275                 280                 285

Asn Leu Asp Cys Arg Glu Tyr Asn Ala Asp Lys Ala Ile Val Asp Ser
    290                 295                 300

Gly Thr Thr Leu Leu Arg Leu Pro Gln Lys Val Phe Asp Ala Val Val
305                 310                 315                 320

Glu Ala Val Ala Arg Ala Ser Leu Ile Pro Glu Phe Ser Asp Gly Phe
                325                 330                 335

Trp Thr Gly Ser Gln Leu Ala Cys Trp Thr Asn Ser Glu Thr Pro Trp
            340                 345                 350

Ser Tyr Phe Pro Lys Ile Ser Ile Tyr Leu Arg Asp Glu Asn Ser Ser
        355                 360                 365

Arg Ser Phe Arg Ile Thr Ile Leu Pro Gln Leu Tyr Ile Gln Pro Met
    370                 375                 380

Met Gly Ala Gly Leu Asn Tyr Glu Cys Tyr Arg Phe Gly Ile Ser Pro
385                 390                 395                 400

Ser Thr Asn Ala Leu Val Ile Gly Ala Thr Val Met Glu Gly Phe Tyr
                405                 410                 415
```

```
Val Ile Phe Asp Arg Ala Gln Lys Arg Val Gly Phe Ala Ala Ser Pro
            420                 425                 430

Cys Ala Glu Ile Ala Gly Ala Ala Val Ser Glu Ile Ser Gly Pro Phe
        435                 440                 445

Ser Thr Glu Asp Val Ala Ser Asn Cys Val Pro Ala Gln Ser Leu Ser
    450                 455                 460

Glu Pro Ile Leu Trp Ile Val Ser Tyr Ala Leu Met Ser Val Cys Gly
465                 470                 475                 480

Ala Ile Leu Leu Val Leu Ile Val Leu Leu Leu Leu Pro Phe Arg Cys
                485                 490                 495

Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu Ser Ser Leu
            500                 505                 510

Val Arg His Arg Trp Lys
        515
```

<210> 3
<211> 2070
<212> DNA
<213> Homo sapiens

<400> 3
atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60
ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggggcgc ccccctgggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540
gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600
ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660
cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720
gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780
gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960
ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020
ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1080
atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140

tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200
ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320
gaagactgtg gctacaacat tccacagaca gatgagtcaa ccctcatgac catagcctat 1380
gtcatggctg ccatctgcgc cctcttcatg ctgccactct gcctcatggt gtgtcagtgg 1440
cgctgcctcc gctgcctgcg ccagcagcat gatgactttg ctgatgacat ctccctgctg 1500
aagtgaggag gcccatgggc agaagataga gattcccctg gaccacacct ccgtggttca 1560
ctttggtcac aagtaggaga cacagatggc acctgtggcc agagcacctc aggaccctcc 1620
ccacccacca aatgcctctg ccttgatgga gaaggaaaag gctggcaagg tgggttccag 1680
ggactgtacc tgtaggaaac agaaaagaga agaaagaagc actctgctgg cgggaatact 1740
cttggtcacc tcaaatttaa gtcgggaaat tctgctgctt gaaacttcag ccctgaacct 1800
ttgtccacca ttcctttaaa ttctccaacc caaagtattc ttcttttctt agtttcagaa 1860
gtactggcat cacacgcagg ttaccttggc gtgtgtccct gtggtaccct ggcagagaag 1920
agaccaagct tgtttccctg ctggccaaag tcagtaggag aggatgcaca gtttgctatt 1980
tgctttagag acagggactg tataaacaag cctaacattg gtgcaaagat tgcctcttga 2040
attaaaaaaa aaaaaaaaaa aaaaaaaaaa 2070

<210> 4
<211> 501
<212> PRT
<213> Homo sapiens

<400> 4
Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
1 5 10 15

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

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp

130 135 140

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

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp
180 185 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro
195 200 205

Asn Leu Phe Ser Leu His Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln
210 215 220

Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile
225 230 235 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg
245 250 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln
260 265 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val
275 280 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala
290 295 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp
305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr
325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val
340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg
355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala
370 375 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu

385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala
405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu
420 425 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro
435 440 445

Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala
450 455 460

Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp
465 470 475 480

Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp
485 490 495

Ile Ser Leu Leu Lys
500

<210> 5
<211> 1977
<212> DNA
<213> Homo sapiens

<400> 5
atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60
ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggcgc ccccctgggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540
gggctggcct atgctgagat tgccaggctt tgtggtgctg gcttcccccc caaccagtct 600
gaagtgctgg cctctgtcgg agggagcatg atcattggag gtatcgacca ctcgctgtac 660
acaggcagtc tctggtatac acccatccgg cgggagtggt attatgaggt gatcattgtg 720
cgggtggaga tcaatggaca ggatctgaaa atggactgca aggagtacaa ctatgacaag 780
agcattgtgg acagtggcac caccaacctt cgtttgccca agaaagtgtt tgaagctgca 840
gtcaaatcca tcaaggcagc ctcctccacg gagaagttcc ctgatggttt ctggctagga 900
gagcagctgg tgtgctggca agcaggcacc acccctgga acattttccc agtcatctca 960
ctctacctaa tgggtgaggt taccaaccag tccttccgca tcaccatcct tccgcagcaa 1020
tacctgcggc cagtggaaga tgtggccacg tcccaagacg actgttacaa gtttgccatc 1080

tcacagtcat ccacgggcac tgttatggga gctgttatca tggagggctt ctacgttgtc 1140
tttgatcggg cccgaaaacg aattggcttt gctgtcagcg cttgccatgt gcacgatgag 1200
ttcaggacgg cagcggtgga aggccctttt gtcaccttgg acatggaaga ctgtggctac 1260
aacattccac agacagatga gtcaaccctc atgaccatag cctatgtcat ggctgccatc 1320
tgcgccctct tcatgctgcc actctgcctc atggtgtgtc agtggcgctg cctccgctgc 1380
ctgcgccagc agcatgatga ctttgctgat gacatctccc tgctgaagtg aggaggccca 1440
tgggcagaag atagagattc ccctggacca cacctccgtg gttcactttg gtcacaagta 1500
ggagacacag atggcacctg tggccagagc acctcaggac cctccccacc caccaaatgc 1560
ctctgccttg atggagaagg aaaaggctgg caaggtgggt tccagggact gtacctgtag 1620
gaaacagaaa agagaagaaa gaagcactct gctggcggga atactcttgg tcacctcaaa 1680
tttaagtcgg gaaattctgc tgcttgaaac ttcagccctg aacctttgtc caccattcct 1740
ttaaattctc caacccaaag tattcttctt ttcttagttt cagaagtact ggcatcacac 1800
gcaggttacc ttggcgtgtg tccctgtggt accctggcag agaagagacc aagcttgttt 1860
ccctgctggc caaagtcagt aggagaggat gcacagtttg ctatttgctt tagagacagg 1920
gactgtataa acaagcctaa cattggtgca aagattgcct cttgaaaaaa aaaaaaa 1977

<210> 6
<211> 476
<212> PRT
<213> Homo sapiens

<400> 6
Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
1 5 10 15

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

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
130 135 140

```
Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
145                 150                 155                 160

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
                165                 170                 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Leu Cys Gly
            180                 185                 190

Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly
        195                 200                 205

Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu
    210                 215                 220

Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val
225                 230                 235                 240

Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr
                245                 250                 255

Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu
            260                 265                 270

Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser
        275                 280                 285

Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val
    290                 295                 300

Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser
305                 310                 315                 320

Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile
                325                 330                 335

Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln
            340                 345                 350

Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val
        355                 360                 365

Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala
    370                 375                 380

Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu
385                 390                 395                 400
```

```
Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu
                405                 410                 415

Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser Thr Leu Met Thr
            420                 425                 430

Ile Ala Tyr Val Met Ala Ala Ile Cys Ala Leu Phe Met Leu Pro Leu
        435                 440                 445

Cys Leu Met Val Cys Gln Trp Arg Cys Leu Arg Cys Leu Arg Gln Gln
    450                 455                 460

His Asp Asp Phe Ala Asp Asp Ile Ser Leu Leu Lys
465                 470                 475
```

<210> 7
<211> 2043
<212> DNA
<213> Mus musculus

<400> 7

```
atggccccag cgctgcactg gctcctgcta tgggtgggct cgggaatgct gcctgcccag 60
ggaacccatc tcggcatccg gctgcccctt cgcagcggcc tggcagggcc acccctgggc 120
ctgaggctgc cccgggagac tgacgaggaa tcggaggagc ctggccggag aggcagcttt 180
gtggagatgg tggacaacct gaggggaaag tccggccagg gctactatgt ggagatgacc 240
gtaggcagcc ccccacagac gctcaacatc ctggtggaca cgggcagtag taactttgca 300
gtgggggctg ccccacaccc tttcctgcat cgctactacc agaggcagct gtccagcaca 360
tatcgagacc tccgaaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaggggaa 420
ctgggcaccg acctggtgag catccctcat ggccccaacg tcactgtgcg tgccaacatt 480
gctgccatca ctgaatcgga caagttcttc atcaatggtt ccaactggga gggcatccta 540
gggctggcct atgctgagat tgccaggccc gacgactctt tggagccctt ctttgactcc 600
ctggtgaagc agacccacat tcccaacatc ttttccctgc agctctgtgg cgctggcttc 660
cccctcaacc agaccgaggc actggcctcg gtgggaggga gcatgatcat tggtggtatc 720
gaccactcgc tatacacggg cagtctctgg tacacaccca tccggcggga gtggtattat 780
gaagtgatca ttgtacgtgt ggaaatcaat ggtcaagatc tcaagatgga ctgcaaggag 840
tacaactacg acaagagcat tgtggacagt gggaccacca accttcgctt gcccaagaaa 900
gtatttgaag ctgccgtcaa gtccatcaag gcagcctcct cgacggagaa gttcccggat 960
ggcttttggc taggggagca gctggtgtgc tggcaagcag gcacgacccc ttggaacatt 1020
ttcccagtca tttcacttta cctcatgggt gaagtcacca atcagtcctt ccgcatcacc 1080
atccttcctc agcaatacct acggccggtg gaggacgtgg ccacgtccca agacgactgt 1140
tacaagttcg ctgtctcaca gtcatccacg ggcactgtta tgggagccgt catcatggaa 1200
ggtttctatg tcgtcttcga tcgagcccga aagcgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcggca gtggaaggtc cgtttgttac ggcagacatg 1320
gaagactgtg gctacaacat tccccagaca gatgagtcaa cacttatgac catagcctat 1380
gtcatggcgg ccatctgcgc cctcttcatg ttgccactct gcctcatggt atgtcagtgg 1440
cgctgcctgc gttgcctgcg ccaccagcac gatgactttg ctgatgacat ctccctgctc 1500
```

aagtaaggag gctcgtgggc agatgatgga gacgcccctg gaccacatct gggtggttcc 1560
ctttggtcac atgagttgga gctatggatg gtacctgtgg ccagagcacc tcaggaccct 1620
caccaacctg ccaatgcttc tggcgtgaca gaacagagaa atcaggcaag ctggattaca 1680
gggcttgcac ctgtaggaca caggagaggg aaggaagcag cgttctggtg gcaggaatat 1740
ccttaggcac cacaaacttg agttggaaat tttgctgctt gaagcttcag ccctgaccct 1800
ctgcccagca tcctttagag tctccaacct aaagtattct ttatgtcctt ccagaagtac 1860
tggcgtcata ctcaggctac ccggcatgtg tccctgtggt accctggcag agaaagggcc 1920
aatctcattc cctgctggcc aaagtcagca gaagaaggtg aagtttgcca gttgctttag 1980
tgatagggac tgcagactca agcctacact ggtacaaaga ctgcgtcttg agataaacaa 2040
gaa 2043

<210> 8
<211> 501
<212> PRT
<213> Mus musculus

<400> 8
Met Ala Pro Ala Leu His Trp Leu Leu Leu Trp Val Gly Ser Gly Met
1 5 10 15

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

Gly Leu Ala Gly Pro Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Ser Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
130 135 140

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

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp

165 170 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp
180 185 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Ile Pro
195 200 205

Asn Ile Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln
210 215 220

Thr Glu Ala Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile
225 230 235 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg
245 250 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln
260 265 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val
275 280 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala
290 295 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp
305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr
325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val
340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg
355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala
370 375 380

Val Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu
385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala
405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu

420 425 430

Gly Pro Phe Val Thr Ala Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro
435 440 445

Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala
450 455 460

Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp
465 470 475 480

Arg Cys Leu Arg Cys Leu Arg His Gln His Asp Asp Phe Ala Asp Asp
485 490 495

Ile Ser Leu Leu Lys
500

<210> 9
<211> 2088
<212> DNA
<213> Homo sapiens

<400> 9
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440

gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

<210> 10
<211> 695
<212> PRT
<213> Homo sapiens

<400> 10

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160

```
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350

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

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415
```

```
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670
```

```
His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685

Phe Phe Glu Gln Met Gln Asn
    690                 695


<210> 11
<211> 2088
<212> DNA
<213> Homo sapiens

<400> 11
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag              2088
```

<210> 12
<211> 695
<212> PRT
<213> Homo sapiens

<400> 12
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
195 200 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
210 215 220

```
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350

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

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480
```

```
Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590

Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685

Phe Phe Glu Gln Met Gln Asn
    690                 695
```

<210> 13
<211> 2088
<212> DNA
<213> Homo sapiens

<400> 13
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcttcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

<210> 14
<211> 695
<212> PRT
<213> Homo sapiens

<400> 14
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro

```
            20                  25                  30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
         35                  40                  45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
     50                  55                  60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
 65                  70                  75                  80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                 85                  90                  95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
        115                 120                 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
```

```
        275                 280                 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350

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

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
```

```
    530                 535                 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640

Ile Phe Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685

Phe Phe Glu Gln Met Gln Asn
    690                 695
```

<210> 15
<211> 2094
<212> DNA
<213> Homo sapiens

<400> 15
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540

```
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag       2094

<210> 16
<211> 697
<212> PRT
<213> Homo sapiens

<400> 16
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
  1               5                  10                  15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
             20                  25                  30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
         35                  40                  45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
     50                  55                  60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
 65                  70                  75                  80
```

```
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
                  85                  90                  95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
            100                 105                 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
        115                 120                 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
    130                 135                 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335
```

```
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350

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

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590
```

```
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685

Phe Phe Glu Gln Met Gln Asn Lys Lys
    690                 695
```

<210> 17
<211> 2094
<212> DNA
<213> Homo sapiens

<400> 17
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200

tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

<210> 18
<211> 697
<212> PRT
<213> Homo sapiens

<400> 18
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

```
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145                 150                 155                 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
                165                 170                 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
            180                 185                 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
        195                 200                 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350

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

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400
```

```
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
    450                 455                 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590

Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655
```

```
His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685

Phe Phe Glu Gln Met Gln Asn Lys Lys
    690                 695
```

<210> 19
<211> 2094
<212> DNA
<213> Homo sapiens

<400> 19

```
atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcatcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggtgg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
```

ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcttcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccccagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

<210> 20
<211> 697
<212> PRT
<213> Homo sapiens

<400> 20
Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
180 185 190

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val

```
        195                 200                 205

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
    210                 215                 220

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
225                 230                 235                 240

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu
                245                 250                 255

Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile
            260                 265                 270

Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg
        275                 280                 285

Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu
    290                 295                 300

Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys
305                 310                 315                 320

Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg
                325                 330                 335

Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp
            340                 345                 350

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

Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala
    370                 375                 380

Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn
385                 390                 395                 400

Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe
                405                 410                 415

Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His
            420                 425                 430

Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala
        435                 440                 445

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu
```

```
    450                 455                 460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala
465                 470                 475                 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn
                485                 490                 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser
            500                 505                 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr
        515                 520                 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln
    530                 535                 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn
545                 550                 555                 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr
                565                 570                 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser
            580                 585                 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
        595                 600                 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
    610                 615                 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
625                 630                 635                 640

Ile Phe Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                645                 650                 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
            660                 665                 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
        675                 680                 685

Phe Phe Glu Gln Met Gln Asn Lys Lys
    690                 695
```

<210> 21
<211> 1341
<212> DNA
<213> Homo sapiens

<400> 21
atggctagca tgactggtgg acagcaaatg ggtcgcggat ccacccagca cggcatccgg 60
ctgcccctgc gcagcggcct ggggggcgcc cccctggggc tgcggctgcc ccgggagacc 120
gacgaagagc ccgaggagcc cggccggagg ggcagctttg tggagatggt ggacaacctg 180
aggggcaagt cggggcaggg ctactacgtg gagatgaccg tgggcagccc cccgcagacg 240
ctcaacatcc tggtggatac aggcagcagt aactttgcag tgggtgctgc cccccacccc 300
ttcctgcatc gctactacca gaggcagctg tccagcacat accgggacct ccggaagggt 360
gtgtatgtgc cctacaccca gggcaagtgg gaaggggagc tgggcaccga cctggtaagc 420
atcccccatg gccccaacgt cactgtgcgt gccaacattg ctgccatcac tgaatcagac 480
aagttcttca tcaacggctc caactgggaa ggcatcctgg ggctggccta tgctgagatt 540
gccaggcctg acgactccct ggagcctttc tttgactctc tggtaaagca gacccacgtt 600
cccaacctct tctccctgca cctttgtggt gctggcttcc ccctcaacca gtctgaagtg 660
ctggcctctg tcggagggag catgatcatt ggaggtatcg accactcgct gtacacaggc 720
agtctctggt atacacccat ccggcgggag tggtattatg aggtcatcat tgtgcgggtg 780
gagatcaatg gacaggatct gaaaatggac tgcaaggagt acaactatga caagagcatt 840
gtggacagtg gcaccaccaa ccttcgtttg cccaagaaag tgtttgaagc tgcagtcaaa 900
tccatcaagg cagcctcctc cacggagaag ttccctgatg gtttctggct aggagagcag 960
ctggtgtgct ggcaagcagg caccacccct tggaacattt tcccagtcat ctcactctac 1020
ctaatgggtg aggttaccaa ccagtccttc cgcatcacca tccttccgca gcaatacctg 1080
cggccagtgg aagatgtggc cacgtcccaa gacgactgtt acaagtttgc catctcacag 1140
tcatccacgg gcactgttat gggagctgtt atcatggagg gcttctacgt tgtctttgat 1200
cgggcccgaa aacgaattgg ctttgctgtc agcgcttgcc atgtgcacga tgagttcagg 1260
acggcagcgg tggaaggccc ttttgtcacc ttggacatgg aagactgtgg ctacaacatt 1320
ccacagacag atgagtcatg a 1341

<210> 22
<211> 446
<212> PRT
<213> Homo sapiens

<400> 22
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Thr Gln
1 5 10 15

His Gly Ile Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu
20 25 30

Gly Leu Arg Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly
35 40 45

Arg Arg Gly Ser Phe Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser
50 55 60

```
Gly Gln Gly Tyr Tyr Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr
 65                  70                  75                  80

Leu Asn Ile Leu Val Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala
                85                  90                  95

Ala Pro His Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser
            100                 105                 110

Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly
        115                 120                 125

Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly
    130                 135                 140

Pro Asn Val Thr Val Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp
145                 150                 155                 160

Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala
                165                 170                 175

Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp
            180                 185                 190

Ser Leu Val Lys Gln Thr His Val Pro Asn Leu Phe Ser Leu His Leu
        195                 200                 205

Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val
    210                 215                 220

Gly Gly Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly
225                 230                 235                 240

Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile
                245                 250                 255

Ile Val Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys
            260                 265                 270

Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu
        275                 280                 285

Arg Leu Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala
    290                 295                 300

Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln
305                 310                 315                 320
```

```
Leu Val Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val
                325                 330                 335

Ile Ser Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile
            340                 345                 350

Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr
        355                 360                 365

Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly
    370                 375                 380

Thr Val Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp
385                 390                 395                 400

Arg Ala Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His
                405                 410                 415

Asp Glu Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp
            420                 425                 430

Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser
        435                 440                 445
```

<210> 23
<211> 1380
<212> DNA
<213> Homo sapiens

<400> 23
atggctagca tgactggtgg acagcaaatg ggtcgcggat cgatgactat ctctgactct 60
ccgcgtgaac aggacggatc cacccagcac ggcatccggc tgcccctgcg cagcggcctg 120
gggggcgccc ccctggggct gcggctgccc cgggagaccg acgaagagcc cgaggagccc 180
ggccggaggg gcagctttgt ggagatggtg gacaacctga ggggcaagtc ggggcagggc 240
tactacgtgg agatgaccgt gggcagcccc ccgcagacgc tcaacatcct ggtggataca 300
ggcagcagta actttgcagt gggtgctgcc ccccacccct tcctgcatcg ctactaccag 360
aggcagctgt ccagcacata ccgggacctc cggaagggtg tgtatgtgcc ctacacccag 420
ggcaagtggg aaggggagct gggcaccgac ctggtaagca tcccccatgg ccccaacgtc 480
actgtgcgtg ccaacattgc tgccatcact gaatcagaca agttcttcat caacggctcc 540
aactgggaag gcatcctggg gctggcctat gctgagattg ccaggcctga cgactccctg 600
gagcctttct ttgactctct ggtaaagcag acccacgttc ccaacctctt ctccctgcac 660
ctttgtggtg ctggcttccc cctcaaccag tctgaagtgc tggcctctgt cggagggagc 720
atgatcattg gaggtatcga ccactcgctg tacacaggca gtctctggta tacacccatc 780
cggcgggagt ggtattatga ggtcatcatt gtgcgggtgg agatcaatgg acaggatctg 840
aaaatggact gcaaggagta caactatgac aagagcattg tggacagtgg caccaccaac 900
cttcgtttgc ccaagaaagt gtttgaagct gcagtcaaat ccatcaaggc agcctcctcc 960
acggagaagt tccctgatgg tttctggcta ggagagcagc tggtgtgctg gcaagcaggc 1020

accacccctt ggaacatttt cccagtcatc tcactctacc taatgggtga ggttaccaac 1080
cagtccttcc gcatcaccat ccttccgcag caatacctgc ggccagtgga agatgtggcc 1140
acgtcccaag acgactgtta caagtttgcc atctcacagt catccacggg cactgttatg 1200
ggagctgtta tcatggaggg cttctacgtt gtctttgatc gggcccgaaa acgaattggc 1260
tttgctgtca gcgcttgcca tgtgcacgat gagttcagga cggcagcggt ggaaggccct 1320
tttgtcacct tggacatgga agactgtggc tacaacattc cacagacaga tgagtcatga 1380

<210> 24
<211> 459
<212> PRT
<213> Homo sapiens

<400> 24
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Thr
1 5 10 15

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

Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg
35 40 45

Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly
50 55 60

Ser Phe Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly
65 70 75 80

Tyr Tyr Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile
85 90 95

Leu Val Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His
100 105 110

Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg
115 120 125

Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu
130 135 140

Gly Glu Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val
145 150 155 160

Thr Val Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe
165 170 175

Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu
180 185 190

```
Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val
        195                 200                 205

Lys Gln Thr His Val Pro Asn Leu Phe Ser Leu His Leu Cys Gly Ala
    210                 215                 220

Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser
225                 230                 235                 240

Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp
                245                 250                 255

Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg
            260                 265                 270

Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn
        275                 280                 285

Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro
    290                 295                 300

Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser
305                 310                 315                 320

Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys
                325                 330                 335

Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu
            340                 345                 350

Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu
        355                 360                 365

Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp
    370                 375                 380

Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met
385                 390                 395                 400

Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg
                405                 410                 415

Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe
            420                 425                 430

Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp
        435                 440                 445
```

```
Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser
    450                 455


<210> 25
<211> 1302
<212> DNA
<213> Homo sapiens

<400> 25
atgactcagc atggtattcg tctgccactg cgtagcggtc tgggtggtgc tccactgggt 60
ctgcgtctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 120
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 180
gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 240
gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 300
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 360
ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 420
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 480
gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 540
ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 600
cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 660
gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 720
gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 780
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 840
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 900
ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 960
ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1020
atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1080
tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1140
ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1200
catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1260
gaagactgtg gctacaacat tccacagaca gatgagtcat ga                    1302

<210> 26
<211> 433
<212> PRT
<213> Homo sapiens

<400> 26
Met Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly
  1               5                  10                  15

Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu
             20                  25                  30

Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val Asp Asn Leu Arg
         35                  40                  45
```

```
Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr Val Gly Ser Pro
     50                  55                  60

Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser Ser Asn Phe Ala
 65                  70                  75                  80

Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln
                 85                  90                  95

Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr
            100                 105                 110

Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu Val Ser Ile
        115                 120                 125

Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile Ala Ala Ile Thr
    130                 135                 140

Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu
145                 150                 155                 160

Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro
                165                 170                 175

Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro Asn Leu Phe Ser
            180                 185                 190

Leu His Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu
        195                 200                 205

Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu
    210                 215                 220

Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr
225                 230                 235                 240

Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met
                245                 250                 255

Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr
            260                 265                 270

Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser
        275                 280                 285

Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu
    290                 295                 300
```

```
Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile
305                 310                 315                 320

Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser
                325                 330                 335

Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp
            340                 345                 350

Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser
        355                 360                 365

Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val
    370                 375                 380

Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys
385                 390                 395                 400

His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val
                405                 410                 415

Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu
            420                 425                 430

Ser
```

<210> 27
<211> 1278
<212> DNA
<213> Homo sapiens

<400> 27
atggctagca tgactggtgg acagcaaatg ggtcgcggat cgatgactat ctctgactct 60
ccgctggact ctggtatcga aaccgacgga tcctttgtgg agatggtgga caacctgagg 120
ggcaagtcgg ggcagggcta ctacgtggag atgaccgtgg gcagcccccc gcagacgctc 180
aacatcctgg tggatacagg cagcagtaac tttgcagtgg gtgctgcccc ccaccccttc 240
ctgcatcgct actaccagag gcagctgtcc agcacatacc gggacctccg gaagggtgtg 300
tatgtgccct acacccaggg caagtgggaa ggggagctgg gcaccgacct ggtaagcatc 360
ccccatggcc ccaacgtcac tgtgcgtgcc aacattgctg ccatcactga atcagacaag 420
ttcttcatca acggctccaa ctgggaaggc atcctggggc tggcctatgc tgagattgcc 480
aggcctgacg actccctgga gcctttcttt gactctctgg taaagcagac ccacgttccc 540
aacctcttct ccctgcacct ttgtggtgct ggcttccccc tcaaccagtc tgaagtgctg 600
gcctctgtcg gagggagcat gatcattgga ggtatcgacc actcgctgta cacaggcagt 660
ctctggtata cacccatccg gcgggagtgg tattatgagg tcatcattgt gcgggtggag 720
atcaatggac aggatctgaa aatggactgc aaggagtaca actatgacaa gagcattgtg 780

gacagtggca ccaccaacct tcgtttgccc aagaaagtgt ttgaagctgc agtcaaatcc 840
atcaaggcag cctcctccac ggagaagttc cctgatggtt tctggctagg agagcagctg 900
gtgtgctggc aagcaggcac caccccttgg aacattttcc cagtcatctc actctaccta 960
atgggtgagg ttaccaacca gtccttccgc atcaccatcc ttccgcagca atacctgcgg 1020
ccagtggaag atgtggccac gtcccaagac gactgttaca agtttgccat ctcacagtca 1080
tccacgggca ctgttatggg agctgttatc atggagggct tctacgttgt ctttgatcgg 1140
gcccgaaaac gaattggctt tgctgtcagc gcttgccatg tgcacgatga gttcaggacg 1200
gcagcggtgg aaggcccttt tgtcaccttg gacatggaag actgtggcta caacattcca 1260
cagacagatg agtcatga 1278

<210> 28
<211> 425
<212> PRT
<213> Homo sapiens

<400> 28
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Thr
1 5 10 15

Ile Ser Asp Ser Pro Leu Asp Ser Gly Ile Glu Thr Asp Gly Ser Phe
20 25 30

Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr
35 40 45

Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val
50 55 60

Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe
65 70 75 80

Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu
85 90 95

Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu
100 105 110

Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val
115 120 125

Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn
130 135 140

Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala
145 150 155 160

Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln
165 170 175

```
Thr His Val Pro Asn Leu Phe Ser Leu His Leu Cys Gly Ala Gly Phe
            180                 185                 190

Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile
        195                 200                 205

Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr
    210                 215                 220

Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu
225                 230                 235                 240

Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp
                245                 250                 255

Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys
            260                 265                 270

Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu
        275                 280                 285

Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln
    290                 295                 300

Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu
305                 310                 315                 320

Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln
                325                 330                 335

Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys
            340                 345                 350

Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala
        355                 360                 365

Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg
    370                 375                 380

Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr
385                 390                 395                 400

Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly
                405                 410                 415

Tyr Asn Ile Pro Gln Thr Asp Glu Ser
            420                 425
```

<210> 29
<211> 1362
<212> DNA
<213> Homo sapiens

<400> 29
atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60
ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggcgc ccccctgggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540
gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600
ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660
cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720
gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780
gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960
ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020
ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1080
atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140
tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200
ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320
gaagactgtg gctacaacat tccacagaca gatgagtcat ga 1362

<210> 30
<211> 453
<212> PRT
<213> Homo sapiens

<400> 30
Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
1 5 10 15

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

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val

```
     50                  55                  60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
 65                  70                  75                  80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
                 85                  90                  95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
            100                 105                 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
        115                 120                 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
    130                 135                 140

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

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
                165                 170                 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp
            180                 185                 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro
        195                 200                 205

Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln
    210                 215                 220

Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile
225                 230                 235                 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg
                245                 250                 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln
            260                 265                 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val
        275                 280                 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala
    290                 295                 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp
```

```
305                 310                 315                 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr
                325                 330                 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val
            340                 345                 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg
        355                 360                 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala
    370                 375                 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu
385                 390                 395                 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala
                405                 410                 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu
            420                 425                 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro
        435                 440                 445

Gln Thr Asp Glu Ser
    450
```

<210> 31
<211> 1380
<212> DNA
<213> Homo sapiens

<400> 31
atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgcccac 60
ggcacccagc acggcatccg gctgcccctg cgcagcggcc tgggggggcgc ccccctgggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtaag catcccccat ggccccaacg tcactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcatcctg 540
gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600
ctggtaaagc agacccacgt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660
cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720

gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780
gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960
ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020
ttcccagtca tctcactcta cctaatgggt gaggttacca accagtcctt ccgcatcacc 1080
atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140
tacaagtttg ccatctcaca gtcatccacg ggcactgtta tgggagctgt tatcatggag 1200
ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320
gaagactgtg gctacaacat tccacagaca gatgagtcac agcagcagca gcagcagtga 1380

<210> 32
<211> 459
<212> PRT
<213> Homo sapiens

<400> 32
Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
1 5 10 15

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

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
35 40 45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
50 55 60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
65 70 75 80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
85 90 95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
100 105 110

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
115 120 125

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
130 135 140

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

```
Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
                165                 170                 175

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp
            180                 185                 190

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro
        195                 200                 205

Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln
    210                 215                 220

Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile
225                 230                 235                 240

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg
                245                 250                 255

Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln
            260                 265                 270

Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val
        275                 280                 285

Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala
    290                 295                 300

Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp
305                 310                 315                 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr
                325                 330                 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val
            340                 345                 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg
        355                 360                 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala
    370                 375                 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu
385                 390                 395                 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala
                405                 410                 415
```

```
Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu
            420                 425                 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro
        435                 440                 445

Gln Thr Asp Glu Ser His His His His His His
    450                 455


<210> 33
<211> 25
<212> PRT
<213> Homo sapiens

<400> 33
Ser Glu Gln Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu
  1               5                  10                  15

Ser Ser Leu Val Arg His Arg Trp Lys
             20                  25


<210> 34
<211> 19
<212> PRT
<213> Homo sapiens

<400> 34
Ser Glu Gln Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp Ile Ser
  1               5                  10                  15

Leu Leu Lys


<210> 35
<211> 29
<212> DNA
<213> Homo sapiens

<400> 35
gtggatccac ccagcacggc atccggctg                                    29

<210> 36
<211> 36
<212> DNA
<213> Homo sapiens
```

<400> 36
gaaagctttc atgactcatc tgtctgtgga atgttg 36

<210> 37
<211> 39
<212> DNA
<213> Homo sapiens

<400> 37
gatcgatgac tatctctgac tctccgcgtg aacaggacg 39

<210> 38
<211> 39
<212> DNA
<213> Homo sapiens

<400> 38
gatccgtcct gttcacgcgg agagtcagag atagtcatc 39

<210> 39
<211> 77
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述 : Hu-Asp2

<400> 39
cggcatccgg ctgcccctgc gtagcggtct gggtggtgct ccactgggtc tgcgtctgcc 60
ccgggagacc gacgaag 77

<210> 40
<211> 77
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述 : Hu-Asp2

<400> 40
cttcgtcggt ctcccggggc agacgcagac ccagtggagc accacccaga ccgctacgca 60
ggggcagccg gatgccg 77

<210> 41
<211> 51
<212> DNA
<213> 人工序列

<220>
<223>人工序列的描述 : Caspase 8
切割位点

<400> 41
gatcgatgac tatctctgac tctccgctgg actctggtat cgaaaccgac g 51

<210> 42
<211> 51
<212> DNA
<213>人工序列 l Sequence

<220>
<223>人工序列的描述 : Caspase 8
切割位点

<400> 42
gatccgtcgg tttcgatacc agagtccagc ggagagtcag agatagtcat c 51

<210> 43
<211> 32
<212> DNA
<213> Homo sapiens

<400> 43
aaggatcctt tgtggagatg gtggacaacc tg 32

<210> 44
<211> 36
<212> DNA
<213> Homo sapiens

<400> 44
gaaagctttc atgactcatc tgtctgtgga atgttg 36

<210> 45
<211> 24
<212> DNA
<213>人工序列

<220>
<223>人工序列的描述 : 6-His 标记

<400> 45
gatcgcatca tcaccatcac catg 24

<210> 46

<211> 24
<212> DNA
<213>人工序列

<220>
<223>人工序列的描述 : 6-His 标记

<400> 46
gatccatggt gatggtgatg atgc 24

<210> 47
<211> 22
<212> DNA
<213>人工序列

<220>
<223>人工序列的描述 : 引入KK基序

<400> 47
gactgaccac tcgaccaggt tc 22

<210> 48
<211> 51
<212> DNA
<213>人工序列

<220>
<223>人工序列的描述 : 引入KK基序

<400> 48
cgaattaaat tccagcacac tggctacttc ttgttctgca tctcaaagaa c 51

<210> 49
<211> 26
<212> DNA
<213>人工序列

<220>
<223>人工序列的描述 : 引入KK基序

<400> 49
cgaattaaat tccagcacac tggcta 26

图1（1）

```
ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC
M  G  A  L  A  R  A  L  L  L  P  L  L  A  Q  W  L  L  R  A

CCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC
A  P  E  L  A  P  A  P  F  T  L  P  L  R  V  A  A  A  T  N

CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG
R  V  V  A  P  T  P  G  P  G  T  P  A  E  R  H  A  D  G  L

GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG
A  L  A  L  E  P  A  L  A  S  P  A  G  A  A  N  F  L  A  M

GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC
V  D  N  L  Q  G  D  S  G  R  G  Y  Y  L  E  M  L  I  G  T

CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA
P  P  Q  K  L  Q  I  L  V  D  T  G  S  S  N  F  A  V  A  G

ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC
T  P  H  S  Y  I  D  T  Y  F  D  T  E  R  S  S  T  Y  R  S

AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA
K  G  F  D  V  T  V  K  Y  T  Q  G  S  W  T  G  F  V  G  E

GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT
D  L  V  T  I  P  K  G  F  N  T  S  F  L  V  N  I  A  T  I

TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT
F  E  S  E  N  F  F  L  P  G  I  K  W  N  G  I  L  G  L  A

TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA
Y  A  T  L  A  K  P  S  S  S  L  E  T  F  F  D  S  L  V  T

CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT
Q  A  N  I  P  N  V  F  S  M  Q  M  C  G  A  G  L  P  V  A

GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA
G  S  G  T  N  G  G  S  L  V  L  G  G  I  E  P  S  L  Y  K

GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA
G  D  I  W  Y  T  P  I  K  E  E  W  Y  Y  Q  I  E  I  L  K

TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC
L  E  I  G  G  Q  S  L  N  L  D  C  R  E  Y  N  A  D  K  A

ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG
I  V  D  S  G  T  T  L  L  R  L  P  Q  K  V  F  D  A  V  V

GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC
E  A  V  A  R  A  S  L  I  P  E  F  S  D  G  F  W  T  G  S

CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC
Q  L  A  C  W  T  N  S  E  T  P  W  S  Y  F  P  K  I  S  I

TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC
Y  L  R  D  E  N  S  S  R  S  F  R  I  T  I  L  P  Q  L  Y

ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA
I  Q  P  M  M  G  A  G  L  N  Y  E  C  Y  R  F  G  I  S  P

TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC
S  T  N  A  L  V  I  G  A  T  V  M  E  G  F  Y  V  I  F  D

AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA
```

图1（2）

```
R  A  Q  K  R  V  G  F  A  A  S  P  C  A  E  I  A  G  A  A

GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT
V  S  E  I  S  G  P  F  S  T  E  D  V  A  S  N  C  V  P  A

CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA
Q  S  L  S  E  P  I  L  W  I  V  S  Y  A  L  M  S  V  C  G

GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC
A  I  L  L  V  L  I  V  L  L  L  L  P  F  R  C  Q  R  R  P

CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATA
R  D  P  E  V  V  N  D  E  S  S  L  V  R  H  R  W  K

GCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGC
AGCAGCCGGGATCGATGGTGGCGCTTTCTCCTGTGCCCACCCGTCTTCAATCTCTGTTCT
GCTCCCAGATGCCTTCTAGATTCACTGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATC
CTCCCTACTTCCAAGAAAAATAATTAAAAAAAAAACTTCATTCTAAACCAAAAAAAAAAA
AAAA
```

图 2（1）

```
ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAGTGCTGCCTGCCCAC
M  A  Q  A  L  P  W  L  L  L  W  M  G  A  G  V  L  P  A  H

GGCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCCTGGGG
G  T  Q  H  G  I  R  L  P  L  R  S  G  L  G  G  A  P  L  G

CTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  P  E  E  P  G  R  R  G  S  F

GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T

GTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A

GTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T

TACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E

CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I

GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L

GGGCTGGCCTATGCTGAGATTGCCAGGCTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCT
G  L  A  Y  A  E  I  A  R  L  C  G  A  G  F  P  L  N  Q  S

GAAGTGCTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTAC
E  V  L  A  S  V  G  G  S  M  I  I  G  G  I  D  H  S  L  Y

ACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTGATCATTGTG
T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y  E  V  I  I  V

CGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAGTACAACTATGACAAG
R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E  Y  N  Y  D  K

AGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCA
S  I  V  D  S  G  T  T  N  L  R  L  P  K  K  V  F  E  A  A

GTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGA
V  K  S  I  K  A  A  S  S  T  E  K  F  P  D  G  F  W  L  G

GAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCCCAGTCATCTCA
E  Q  L  V  C  W  Q  A  G  T  T  P  W  N  I  F  P  V  I  S

CTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACCATCCTTCCGCAGCAA
L  Y  L  M  G  E  V  T  N  Q  S  F  R  I  T  I  L  P  Q  Q

TACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGTTACAAGTTTGCCATC
```

图 2（2）

```
Y  L  R  P  V  E  D  V  A  T  S  Q  D  D  C  Y  K  F  A  I

TCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTC
S  Q  S  S  T  G  T  V  M  G  A  V  I  M  E  G  F  Y  V  V

TTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAG
F  D  R  A  R  K  R  I  G  F  A  V  S  A  C  H  V  H  D  E

TTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTCACCTTGGACATGGAAGACTGTGGCTAC
F  R  T  A  A  V  E  G  P  F  V  T  L  D  M  E  D  C  G  Y

AACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTATGTCATGGCTGCCATC
N  I  P  Q  T  D  E  S  T  L  M  T  I  A  Y  V  M  A  A  I

TGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGGCGCTGCCTCCGCTGC
C  A  L  F  M  L  P  L  C  L  M  V  C  Q  W  R  C  L  R  C

CTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAAGTGAGGAGGCCCA
L  R  Q  Q  H  D  D  F  A  D  D  I  S  L  L  K

TGGGCAGAAGATAGAGATTCCCCTGGACCACACCTCCGTGGTTCACTTTGGTCACAAGTA
GGAGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCCTCCCCACCCACCAAATGC
CTCTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCCAGGGACTGTACCTGTAG
GAAACAGAAAAGAGAAGAAAGAAGCACTCTGCTGGCGGGAATACTCTTGGTCACCTCAAA
TTTAAGTCGGGAAATTCTGCTGCTTGAAACTTCAGCCCTGAACCTTTGTCCACCATTCCT
TTAAATTCTCCAACCCAAAGTATTCTTCTTTTCTTAGTTTCAGAAGTACTGGCATCACAC
GCAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAAGAGACCAAGCTTGTTT
CCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCTATTTGCTTTAGAGACAGG
GACTGTATAAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGAAAAAAAAAAAAA
```

图3（1）

```
ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAGTGCTGCCTGCCCAC
M  A  Q  A  L  P  W  L  L  L  W  M  G  A  G  V  L  P  A  H

GGCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCCTGGGG
G  T  Q  H  G  I  R  L  P  L  R  S  G  L  G  G  A  P  L  G

CTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  P  E  E  P  G  R  R  G  S  F

GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T

GTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A

GTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T

TACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E

CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I

GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L

GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT
G  L  A  Y  A  E  I  A  R  P  D  D  S  L  E  P  F  F  D  S

CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTC
L  V  K  Q  T  H  V  P  N  L  F  S  L  Q  L  C  G  A  G  F

CCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATC
P  L  N  Q  S  E  V  L  A  S  V  G  G  S  M  I  I  G  G  I

GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT
D  H  S  L  Y  T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y

GAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG
E  V  I  I  V  R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E

TACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAA
Y  N  Y  D  K  S  I  V  D  S  G  T  T  N  L  R  L  P  K  K

GTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGAT
V  F  E  A  A  V  K  S  I  K  A  A  S  S  T  E  K  F  P  D
```

图 3 (2)

GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATT
G F W L G E Q L V C W Q A G T T P W N I

TTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACC
F P V I S L Y L M G E V T N Q S F R I T

ATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGT
I L P Q Q Y L R P V E D V A T S Q D D C

TACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG
Y K F A I S Q S S T G T V M G A V I M E

GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC
G F Y V V F D R A R K R I G F A V S A C

CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTCACCTTGGACATG
H V H D E F R T A A V E G P F V T L D M

GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTAT
E D C G Y N I P Q T D E S T L M T I A Y

GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGG
V M A A I C A L F M L P L C L M V C Q W

CGCTGCCTCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG
R C L R C L R Q Q H D D F A D D I S L L

AAGTGAGGAGGCCCATGGGCAGAAGATAGAGATTCCCCTGGACCACACCTCCGTGGTTCA
K

CTTTGGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCCTCC
CCACCCACCAAATGCCTCTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCCAG
GGACTGTACCTGTAGGAAACAGAAAAGAGAAGAAAGAAGCACTCTGCTGGCGGGAATACT
CTTGGTCACCTCAAATTTAAGTCGGGAAATTCTGCTGCTTGAAACTTCAGCCCTGAACCT
TTGTCCACCATTCCTTTAAATTCTCCAACCCAAAGTATTCTTCTTTTCTTAGTTTCAGAA
GTACTGGCATCACACGCAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAAG
AGACCAAGCTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCTATT
TGCTTTAGAGACAGGGACTGTATAAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGA
ATTAAAAAAAAAAAAAAAAAAAAAAAAAA

图 4

```
ATGGCCCCAGCGCTGCACTGGCTCCTGCTATGGGTGGGCTCGGGAATGCTGCCTGCCCAG
M  A  P  A  L  H  W  L  L  L  W  V  G  S  G  M  L  P  A  Q
GGAACCCATCTCGGCATCCGGCTGCCCCTTCGCAGCGGCCTGGCAGGGCCACCCCTGGGC
G  T  H  L  G  I  R  L  P  L  R  S  G  L  A  G  P  P  L  G
CTGAGGCTGCCCCCGGGAGACTGACGAGGAATCGGAGGAGCCTGGCCGGAGAGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  S  E  E  P  G  R  R  G  S  F
GTGGAGATGGTGGACAACCTGAGGGGAAAGTCCGGCCAGGGCTACTATGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T
GTAGGCAGCCCCCCACAGACGCTCAACATCCTGGTGGACACGGGCAGTAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A
GTGGGGGCTGCCCCACACCCTTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T
TATCGAGACCTCCGAAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAGGGGGAA
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E
CTGGGCACCGACCTGGTGAGCATCCCTCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I
GCTGCCATCACTGAATCGGACAAGTTCTTCATCAATGGTTCCAACTGGGAGGGCATCCTA
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L
GGGCTGGCCTATGCTGAGATTGCCAGGCCCGACGACTCTTTGGAGCCCTTCTTTGACTCC
G  L  A  Y  A  E  I  A  R  P  D  D  S  L  E  P  F  F  D  S
CTGGTGAAGCAGACCCACATTCCCAACATCTTTTCCCTGCAGCTCTGTGGCGCTGGCTTC
L  V  K  Q  T  H  I  P  N  I  F  S  L  Q  L  C  G  A  G  F
CCCCTCAACCAGACCGAGGCACTGGCCTCGGTGGGAGGGAGCATGATCATTGGTGGTATC
P  L  N  Q  T  E  A  L  A  S  V  G  G  S  M  I  I  G  G  I
GACCACTCGCTATACACGGGCAGTCTCTGGTACACACCCATCCGGCGGGAGTGGTATTAT
D  H  S  L  Y  T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y
GAAGTGATCATTGTACGTGTGGAAATCAATGGTCAAGATCTCAAGATGGACTGCAAGGAG
E  V  I  I  V  R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E
TACAACTACGACAAGAGCATTGTGGACAGTGGGACCACCAACCTTCGCTTGCCCAAGAAA
Y  N  Y  D  K  S  I  V  D  S  G  T  T  N  L  R  L  P  K  K
GTATTTGAAGCTGCCGTCAAGTCCATCAAGGCAGCCTCCTCGACGGAGAAGTTCCCGGAT
V  F  E  A  A  V  K  S  I  K  A  A  S  S  T  E  K  F  P  D
GGCTTTTGGCTAGGGGAGCAGCTGGTGTGCTGGCAAGCAGGCACGACCCCTTGGAACATT
G  F  W  L  G  E  Q  L  V  C  W  Q  A  G  T  T  P  W  N  I
TTCCCAGTCATTTCACTTTACCTCATGGGTGAAGTCACCAATCAGTCCTTCCGCATCACC
F  P  V  I  S  L  Y  L  M  G  E  V  T  N  Q  S  F  R  I  T
ATCCTTCCTCAGCAATACCTACGGCCGGTGGAGGACGTGGCCACGTCCCAAGACGACTGT
I  L  P  Q  Q  Y  L  R  P  V  E  D  V  A  T  S  Q  D  D  C
TACAAGTTCGCTGTCTCACAGTCATCCACGGGCACTGTTATGGGAGCCGTCATCATGGAA
Y  K  F  A  V  S  Q  S  S  T  G  T  V  M  G  A  V  I  M  E
GGTTTCTATGTCGTCTTCGATCGAGCCCGAAAGCGAATTGGCTTTGCTGTCAGCGCTTGC
G  F  Y  V  V  F  D  R  A  R  K  R  I  G  F  A  V  S  -  A  C
CATGTGCACGATGAGTTCAGGACGGCGGCAGTGGAAGGTCCGTTTGTTACGGCAGACATG
H  V  H  D  E  F  R  T  A  A  V  E  G  P  F  V  T  A  D  M
GAAGACTGTGGCTACAACATTCCCCAGACAGATGAGTCAACACTTATGACCATAGCCTAT
E  D  C  G  Y  N  I  P  Q  T  D  E  S  T  L  M  T  I  A  Y
GTCATGGCGGCCATCTGCGCCCTCTTCATGTTGCCACTCTGCCTCATGGTATGTCAGTGG
V  M  A  A  I  C  A  L  F  M  L  P  L  C  L  M  V  C  Q  W
CGCTGCCTGCGTTGCCTGCGCCACCAGCACGATGACTTTGCTGATGACATCTCCCTGCTC
R  C  L  R  C  L  R  H  Q  H  D  D  F  A  D  D  I  S  L  L
AAGTAAGGAGGCTCGTGGGCAGATGATGGAGACGCCCCTGGACCACATCTGGGTGGTTCC
K
CTTTGGTCACATGAGTTGGAGCTATGGATGGTACCTGTGGCCAGAGCACCTCAGGACCCT
CACCAACCTGCCAATGCTTCTGGCGTGACAGAACAGAGAAATCAGGCAAGCTGGATTACA
GGGCTTGCACCTGTAGGACACAGGAGAGGGAAGGAAGCAGCGTTCTGGTGGCAGGAATAT
CCTTAGGCACCACAAACTTGAGTTGGAAATTTTGCTGCTTGAAGCTTCAGCCCTGACCCT
CTGCCCAGCATCCTTTAGAGTCTCCAACCTAAAGTATTCTTTATGTCCTTCCAGAAGTAC
TGGCGTCATACTCAGGCTACCCGGCATGTGTCCCTGTGGTACCCTGGCAGAGAAAGGGCC
AATCTCATTCCCTGCTGGCCAAAGTCAGCAGAAGAAGGTGAAGTTTGCCAGTTGCTTTAG
TGATAGGGACTGCAGACTCAAGCCTACACTGGTACAAAGACTGCGTCTTGAGATAAACAA
GAA
```

图 5

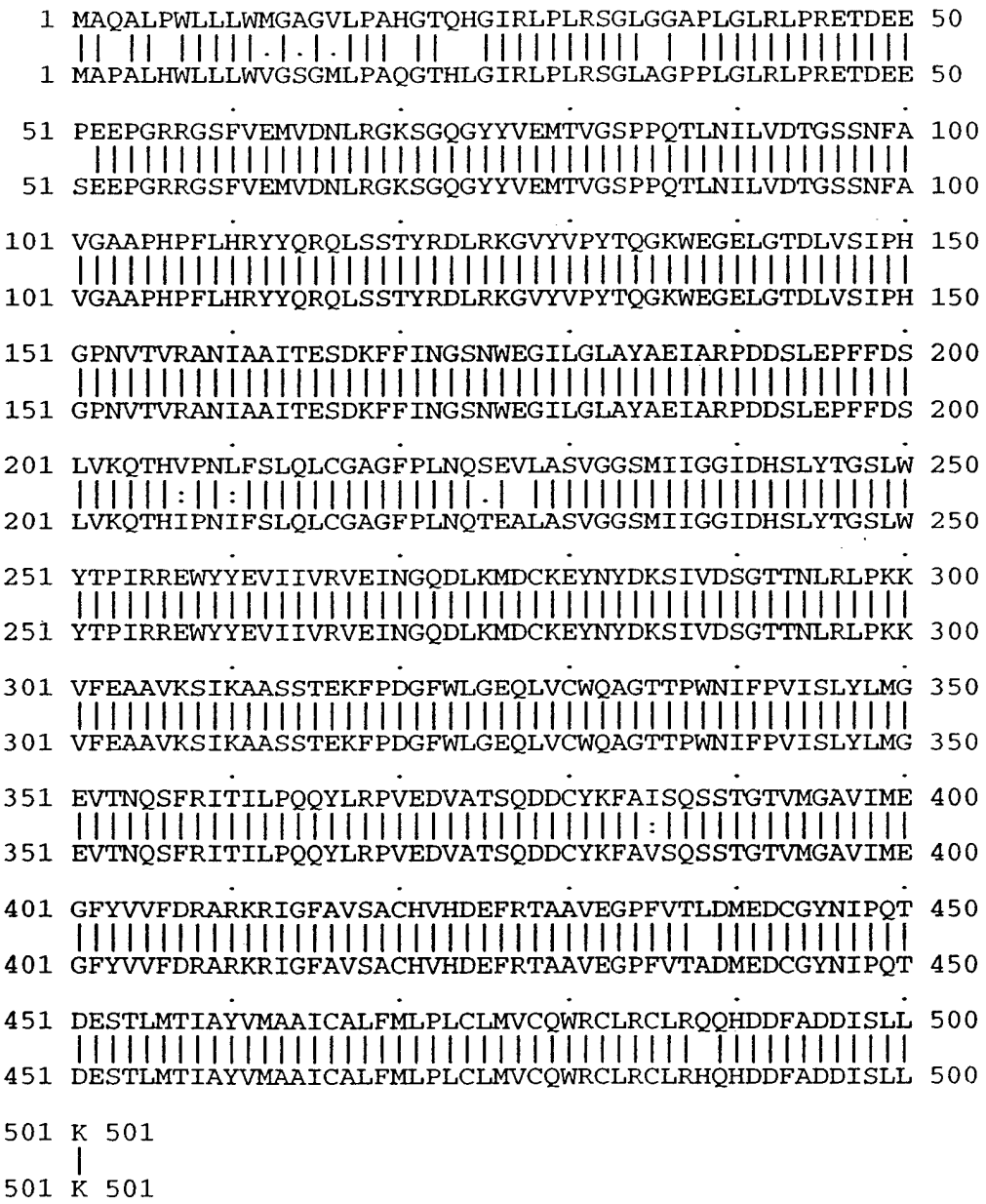

1 MAQALPWLLLWMGAGVLPAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEE 50
1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLAGPPLGLRLPRETDEE 50
51 PEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA 100
51 SEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA 100
101 VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH 150
101 VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH 150
151 GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS 200
151 GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS 200
201 LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW 250
201 LVKQTHIPNIFSLQLCGAGFPLNQTEALASVGGSMIIGGIDHSLYTGSLW 250
251 YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK 300
251 YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK 300
301 VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG 350
301 VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG 350
351 EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME 400
351 EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIME 400
401 GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT 450
401 GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQT 450
451 DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRQQHDDFADDISLL 500
451 DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRHQHDDFADDISLL 500
501 K 501
501 K 501

图6（1）

```
ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCACCCAGCACGGCATCCGG
M  A  S  M  T  G  G  Q  Q  M  G  R  G  S  T  Q  H  G  I  R

CTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCCTGGGGCTGCGGCTGCCCCGGGAGACC
L  P  L  R  S  G  L  G  G  A  P  L  G  L  R  L  P  R  E  T

GACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTTGTGGAGATGGTGGACAACCTG
D  E  E  P  E  E  P  G  R  R  G  S  F  V  E  M  V  D  N  L

AGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACCGTGGGCAGCCCCCCGCAGACG
R  G  K  S  G  Q  G  Y  Y  V  E  M  T  V  G  S  P  P  Q  T

CTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCCACCCC
L  N  I  L  V  D  T  G  S  S  N  F  A  V  G  A  A  P  H  P

TTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGGC
F  L  H  R  Y  Y  Q  R  Q  L  S  S  T  Y  R  D  L  R  K  G

GTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGC
V  Y  V  P  Y  T  Q  G  K  W  E  G  E  L  G  T  D  L  V  S

ATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGAC
I  P  H  G  P  N  V  T  V  R  A  N  I  A  A  I  T  E  S  D

AAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTGGGGCTGGCCTATGCTGAGATT
K  F  F  I  N  G  S  N  W  E  G  I  L  G  L  A  Y  A  E  I

GCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGTT
A  R  P  D  D  S  L  E  P  F  F  D  S  L  V  K  Q  T  H  V

CCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTG
P  N  L  F  S  L  Q  L  C  G  A  G  F  P  L  N  Q  S  E  V

CTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGC
L  A  S  V  G  G  S  M  I  I  G  G  I  D  H  S  L  Y  T  G

AGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTCATCATTGTGCGGGTG
S  L  W  Y  T  P  I  R  R  E  W  Y  Y  E  V  I  I  V  R  V

GAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAGTACAACTATGACAAGAGCATT
E  I  N  G  Q  D  L  K  M  D  C  K  E  Y  N  Y  D  K  S  I

GTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCAGTCAAA
V  D  S  G  T  T  N  L  R  L  P  K  K  V  F  E  A  A  V  K

TCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGAGCAG
S  I  K  A  A  S  S  T  E  K  F  P  D  G  F  W  L  G  E  Q

CTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCCCAGTCATCTCACTCTAC
L  V  C  W  Q  A  G  T  T  P  W  N  I  F  P  V  I  S  L  Y

CTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACCATCCTTCCGCAGCAATACCTG
L  M  G  E  V  T  N  Q  S  F  R  I  T  I  L  P  Q  Q  Y  L

CGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAG
```

图6（2）

```
R  P  V  E  D  V  A  T  S  Q  D  D  C  Y  K  F  A  I  S  Q

TCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGAT
S  S  T  G  T  V  M  G  A  V  I  M  E  G  F  Y  V  V  F  D

CGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTCAGG
R  A  R  K  R  I  G  F  A  V  S  A  C  H  V  H  D  E  F  R

ACGGCAGCGGTGGAAGGCCCTTTTGTCACCTTGGACATGGAAGACTGTGGCTACAACATT
T  A  A  V  E  G  P  F  V  T  L  D  M  E  D  C  G  Y  N  I

CCACAGACAGATGAGTCATGA
P  Q  T  D  E  S  *
```

图 7 (1)

```
ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCGATGACTATCTCTGACTCT
M  A  S  M  T  G  G  Q  Q  M  G  R  G  S  M  T  I  S  D  S

CCGCGTGAACAGGACGGATCCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTG
P  R  E  Q  D  G  S  T  Q  H  G  I  R  L  P  L  R  S  G  L

GGGGGCGCCCCCCTGGGGCTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCC
G  G  A  P  L  G  L  R  L  P  R  E  T  D  E  E  P  E  E  P

GGCCGGAGGGGCAGCTTTGTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGC
G  R  R  G  S  F  V  E  M  V  D  N  L  R  G  K  S  G  Q  G

TACTACGTGGAGATGACCGTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACA
Y  Y  V  E  M  T  V  G  S  P  P  Q  T  L  N  I  L  V  D  T

GGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAG
G  S  S  N  F  A  V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q

AGGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGGCGTGTATGTGCCCTACACCCAG
R  Q  L  S  S  T  Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q

GGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTC
G  K  W  E  G  E  L  G  T  D  L  V  S  I  P  H  G  P  N  V

ACTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCC
T  V  R  A  N  I  A  A  I  T  E  S  D  K  F  F  I  N  G  S

AACTGGGAAGGCATCCTGGGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTG
N  W  E  G  I  L  G  L  A  Y  A  E  I  A  R  P  D  D  S  L

GAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAG
E  P  F  F  D  S  L  V  K  Q  T  H  V  P  N  L  F  S  L  Q

CTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCGGAGGGAGC
L  C  G  A  G  F  P  L  N  Q  S  E  V  L  A  S  V  G  G  S

ATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATC
M  I  I  G  G  I  D  H  S  L  Y  T  G  S  L  W  Y  T  P  I

CGGCGGGAGTGGTATTATGAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTG
R  R  E  W  Y  Y  E  V  I  I  V  R  V  E  I  N  G  Q  D  L

AAAATGGACTGCAAGGAGTACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAAC
K  M  D  C  K  E  Y  N  Y  D  K  S  I  V  D  S  G  T  T  N

CTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCC
L  R  L  P  K  K  V  F  E  A  A  V  K  S  I  K  A  A  S  S

ACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGC
T  E  K  F  P  D  G  F  W  L  G  E  Q  L  V  C  W  Q  A  G

ACCACCCCTTGGAACATTTTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAAC
T  T  P  W  N  I  F  P  V  I  S  L  Y  L  M  G  E  V  T  N
```

图 7（2）

```
CAGTCCTTCCGCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCC
Q  S  F  R  I  T  I  L  P  Q  Q  Y  L  R  P  V  E  D  V  A

ACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATG
T  S  Q  D  D  C  Y  K  F  A  I  S  Q  S  S  T  G  T  V  M

GGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGC
G  A  V  I  M  E  G  F  Y  V  V  F  D  R  A  R  K  R  I  G

TTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCT
F  A  V  S  A  C  H  V  H  D  E  F  R  T  A  A  V  E  G  P

TTTGTCACCTTGGACATGGAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCATGA
F  V  T  L  D  M  E  D  C  G  Y  N  I  P  Q  T  D  E  S  *
```

图8（1）

```
ATGACTCAGCATGGTATTCGTCTGCCACTGCGTAGCGGTCTGGGTGGTGCTCCACTGGGT
M  T  Q  H  G  I  R  L  P  L  R  S  G  L  G  G  A  P  L  G  -

CTGCGTCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT
L  R  L  P  R  E  T  D  E  E  P  E  E  P  G  R  R  G  S  F  -

GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC
V  E  M  V  D  N  L  R  G  K  S  G  Q  G  Y  Y  V  E  M  T  -

GTGGGCAGCCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA
V  G  S  P  P  Q  T  L  N  I  L  V  D  T  G  S  S  N  F  A  -

GTGGGTGCTGCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA
V  G  A  A  P  H  P  F  L  H  R  Y  Y  Q  R  Q  L  S  S  T  -

TACCGGGACCTCCGGAAGGGCGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG
Y  R  D  L  R  K  G  V  Y  V  P  Y  T  Q  G  K  W  E  G  E  -

CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT
L  G  T  D  L  V  S  I  P  H  G  P  N  V  T  V  R  A  N  I  -

GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCCTG
A  A  I  T  E  S  D  K  F  F  I  N  G  S  N  W  E  G  I  L  -

GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT
G  L  A  Y  A  E  I  A  R  P  D  D  S  L  E  P  F  F  D  S

CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTC
L  V  K  Q  T  H  V  P  N  L  F  S  L  Q  L  C  G  A  G  F  -

CCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCGGAGGGAGCATGATCATTGGAGGTATC
P  L  N  Q  S  E  V  L  A  S  V  G  G  S  M  I  I  G  G  I  -

GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT
D  H  S  L  Y  T  G  S  L  W  Y  T  P  I  R  R  E  W  Y  Y  -

GAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG
E  V  I  I  V  R  V  E  I  N  G  Q  D  L  K  M  D  C  K  E

TACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAA
Y  N  Y  D  K  S  I  V  D  S  G  T  T  N  L  R  L  P  K  K  -

GTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGAT
V  F  E  A  A  V  K  S  I  K  A  A  S  S  T  E  K  F  P  D  -

GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATT
G  F  W  L  G  E  Q  L  V  C  W  Q  A  G  T  T  P  W  N  I  -

TTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTTCGCATCACC
F  P  V  I  S  L  Y  L  M  G  E  V  T  N  Q  S  F  R  I  T  -

ATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGT
I  L  P  Q  Q  Y  L  R  P  V  E  D  V  A  T  S  Q  D  D  C  -
```

图 8（2）

```
TACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG
Y  K  F  A  I  S  Q  S  S  T  G  T  V  M  G  A  V  I  M  E  -

GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC
G  F  Y  V  V  F  D  R  A  R  K  R  I  G  F  A  V  S  A  C  -

CATTAG
H  *
```

图 9

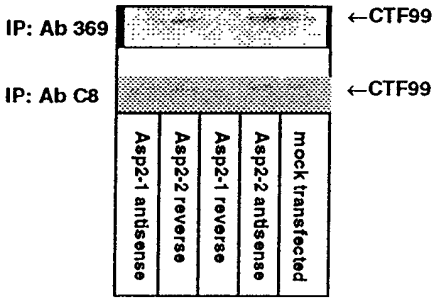

IP: Ab 369
←CTF99
IP: Ab C8
←CTF99
Asp2-1 antisense
Asp2-2 reverse
Asp2-1 reverse
Asp2-2 antisense
mock transfected

图 10

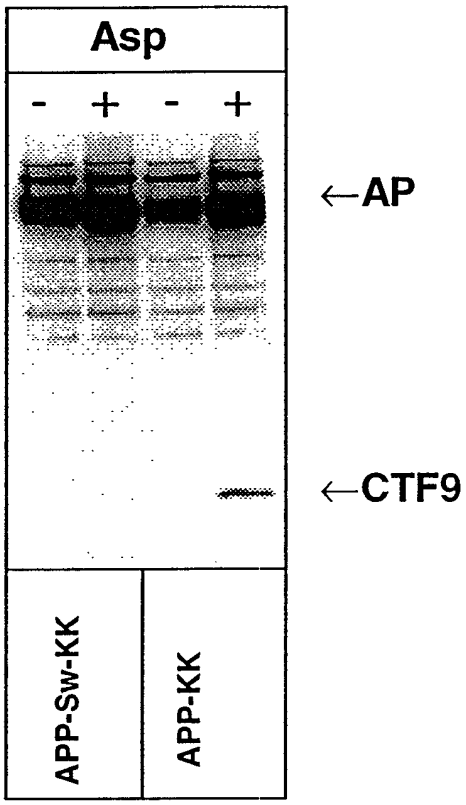


图 11

MAQALPWLLLWMGAGVLPAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEE
PEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA
VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH
GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS
LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW
YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK
VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG
EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT
DES

图 12

MAQALPWLLLWMGAGVLPAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEE
PEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFA
VGAAPHPFLHRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSIPH
GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS
LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW
YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK
VFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG
EVTNQSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT
DESHHHHHH