



[12] 发明专利申请公开说明书

[21]申请号 93117790.1

[51]Int.Cl⁵

C07K 15/12

[43]公开日 1994年8月24日

[22]申请日 93.8.5

[30]优先权

[32]92.8.5 [33]US[31]923,740

[71]申请人 马克斯普朗克科学促进协会

地址 联邦德国格丁根

[72]发明人 A·乌尔里希 N·P·H·莫勒
K·B·莫勒

[74]专利代理机构 中国专利代理(香港)有限公司
代理人 汪 洋

C07K 15/14 C07K 15/04
C12N 9/50 C12N 15/09
C12P 21/02 C12Q 1/68
C12Q 1/37 G01N 33/50
A61K 37/02 A61K 37/54

说明书页数:

附图页数:

[54]发明名称 蛋白质酪氨酸磷酸酶的PTP-D亚家族

[57]摘要

鉴别一种蛋白质酪氨酸磷酸酶的新亚家族(PTP—D)。该家族中包括的是,已知蛋白质酪氨酸磷酸酶的催化磷酸酶区,以前定义的一致序列中,有一个,两个,或3个所鉴定的氨基酸变化的PTP—D蛋白质或糖蛋白。可以用重组方法生产PTP—D蛋白质或糖蛋白。提供抗PTP—D蛋白质或糖蛋白的抗体以及为其编码的核酸构建体,筛选能与PTP—D蛋白质或糖蛋白结合并抑制或刺激其酶活性的分子的方法。

权利要求书

1. 一种 P T P - D 蛋白质或糖蛋白，其中若所述的 P T P - D 蛋白质或糖蛋白是天然产物，则所述的 P T P - D 蛋白质或糖蛋白基本没有与其天然相关的其它蛋白质或糖蛋白。

2. 根据权利要求 1 的 P T P - D 蛋白质或糖蛋白，含有包括一个，2 个或 3 个选自下列氨基酸序列的催化磷酸酶区：

1. GYINAS/N [SEQ ID NO. 5]
2. SXXYWP [SEQ ID NO. 6]
3. IAMVXXXXE [SEQ ID NO. 7]。

3. 根据权利要求 1 的 P T P - D 蛋白质或糖蛋白，它不是天然存在的。

4. 根据权利要求 3 的 P T P - D 蛋白质或糖蛋白，它有缺失的氨基酸。

5. 根据权利要求 3 的 P T P - D 蛋白质或糖蛋白，有额外的氨基酸。

6. 根据权利要求 3 的 P T P - D 蛋白质或糖蛋白，它有取代的氨基酸。

7. 根据权利要求 1 的 P T P - D 蛋白质或糖蛋白，它是天然产生的，且基本没有与其相关的其它蛋白质或糖蛋白。

8. 根据权利要求 1 的 P T P - D 蛋白质或糖蛋白，含有与图 1 中所示的 P T P - D 1 的氨基酸序列有 7 0 % 或更多同源性的氨基酸序列。

9. 根据权利要求 8 的 P T P - D 蛋白质或糖蛋白，是 P T P - D 1 。

1 0. 根据权利要求 1 的 P T P - D 蛋白质或糖蛋白，含有与图 2 所示的 P T P - D 2 的氨基酸序列有 7 0 % 或更多相同性的氨基酸序列。

1 1. 根据权利要求 1 0 的 P T P - D 蛋白质或糖蛋白，它是 P T P - D 2 。

1 2. 一种含编码 P T P - D 蛋白质或糖蛋白的核苷酸序列的重组核酸构建体。

1 3. 根据权利要求 1 2 的重组核酸构建体，是 c D N A 序列。

1 4. 根据权利要求 1 2 的重组核酸构建体，它是基因组 D N A 序列。

1 5 . 根据权利要求 1 2 的重组核酸构建体, 有SEQ ID NO. 2 3 的核苷酸序列 (图 1) 。

1 6 . 根据权利要求 1 2 的重组核酸构建体, 有SEQ ID NO. 2 5 的核苷酸序列 (图 2) 。

1 7 . 根据权利要求 1 2 的重组核酸构建体, 它是表达载体。

1 8 . 权利要求 1 7 的重组核酸构建体, 是载体。

1 9 . 权利要求 1 7 的重组核酸构建体, 是质粒。

2 0 . 一种分离和纯化的核酸构建体, 含有一个编码 P T P - D 蛋白质或糖蛋白的核苷酸序列。

2 1 . 根据权利要求 2 0 的分离的和纯化的核酸构建体, 含有 SEQ ID NO. 2 3 的核苷酸序列 (图 1) 。

2 2 . 根据权利要求 2 0 的分离的和纯化的核酸构建体, 含有 SEQ ID NO. 2 5 的核苷酸序列 (图 2) 。

2 3 . 用权利要求 1 7 的核酸构建体转化的原核宿主细胞。

2 4 . 根据权利要求 2 3 的原核宿主细胞, 是细菌。

2 5 . 用权利要求 1 7 的核酸构建体转化的真核宿主细胞。

2 6 . 根据权利要求 2 5 的真核细胞，是酵母细胞或哺乳动物细胞。

2 7 . 一种制备 P T P - D 蛋白质或糖蛋白的方法，包括：

- (a) 在培养条件下，培养能表达所述 P T P - D 蛋白质或糖蛋白的宿主细胞；
- (b) 表达所述的 P T P - D 蛋白质或糖蛋白；和
- (c) 从所述培养物中回收所述的 P T P - D 蛋白质或糖蛋白。

2 8 . 根据权利要求 2 7 的方法，其中所述宿主细胞是原核细胞。

2 9 . 根据权利要求 2 7 的方法，其中所述宿主细胞是真核细胞。

3 0 . 一种 P T P - D 分子特异性的抗体。

3 1 . 根据权利要求 3 0 的抗体，是单克隆的。

3 2 . 根据权利要求 3 0 的抗体，是多克隆的。

3 3 . 一种检测细胞中 P T P - D 蛋白质或糖蛋白的存在或测量其数量的方法，所述方法包括：

- (a) 将所述细胞或其提取物与 P T P - D 蛋白质或糖蛋白的抗原决定基特异性的抗体接触；和
- (b) 检测所述抗体与所述细胞或其提取物的结合，或测量结合的抗体数量。

由此检测 P T P - D 蛋白质或糖蛋白的数量。

3 4 . 一种检测细胞或主体中编码 P T P - D 蛋白的核酸序列，或编码突变 P T P - D 蛋白质或糖蛋白的核酸构建体的存在的方法，包括：

- (a) 在杂交条件下，将所述主体的细胞或其提取物与编码至少部分所述 P T P - D 蛋白质或糖蛋白或至少部分所述突变 P T P - D 蛋白质或糖蛋白的寡核苷酸探针接触；和
- (b) 检测所述探针与所述细胞核酸的杂交，由此检测所述核酸构建体的存在。

3 5 . 权利要求 3 4 的方法，在步骤 (a) 之前还包括：

- (c) 选择性地扩增编码所述 P T P - D 蛋白质或糖蛋白的所述细胞的所述核酸的数量。

3 6 . 一种鉴别化学或生物制备中，能与 P T P - D 蛋白质或糖蛋白结合的化合物的方法，所述方法包括：

- (a) 将所述 P T P - D 蛋白质或糖蛋白，或结合其部分的化合物附着到固相基质上；

- (b) 将所述的化学或生物制品与所述固相基质接触以便使所述化合物结合，洗涤掉未结合的物质；和
- (c) 检测与所述固相结合所述化合物的存在。

37. 一种从复合混合物中分离能与 P T P - D 蛋白质或糖蛋白结合的化合物的方法，所述方法包括：

- (a) 将所述的 P T P - D 蛋白质或糖蛋白，或结合其部分的化合物附着到固相基质上；
 - (b) 将所述复合混合物与所述固相基质接触以使所述化合物结合，并洗涤掉未结合的物质；和
 - (c) 洗脱所述结合的化合物，
- 由此分离所述化合物。

38. 一种鉴别能刺激或抑制 P T P - D 蛋白质或糖蛋白酶活性的方法，所述方法包括：

- (a) 将所述化合物与纯化形式的，膜制剂中的，或完整活的细胞或固定细胞中的所述 P T P - D 蛋白质或糖蛋白接触；
- (b) 将步骤 (a) 的所述混合物培养足够的时间；
- (c) 测定所述 P T P - D 蛋白质或糖蛋白的酶活性；
- (d) 与没有所述化合物培养的所述 P T P - D 蛋白质或糖蛋白的酶活性进行比较，

由此，检测所述 P T P - D 蛋白质或糖蛋白是否刺激或抑制所述活性。

39. 根据权利要求 38 的方法，其中步骤 (a) 中的所述 P T P

- D 蛋白质或糖蛋白是 P T P - D 蛋白质或糖蛋白的片段。

4 0 . 一种治疗或防止带有不正常 P T P - D 蛋白质或糖蛋白的疾病的药物组合物，所述组合物含有 P T P - D 蛋白质或糖蛋白和药理学可接受的载体。

4 1 . 一种治疗或防止带有不正常 P T P - D 蛋白质或糖蛋白的疾病的药物组合物，所述组合物含有一种以刺激或抑制 P T P - D 蛋白质或糖蛋白的酶活性有效量的分子。

4 2 . 一种治疗或防止主体中，带有不正常 P T P - D 蛋白质或糖蛋白的疾病的方法，包括给所述主体施用权利要求 4 0 的所述组合物。

4 3 . 一种治疗或防止主体中，带有不正常 P T P - D 蛋白质或糖蛋白的疾病的方法，包括给所述主体施用权利要求 4 1 的所述组合物。

4 4 . 一种药物组合物，含有 P T P - D 蛋白质或糖蛋白和药理学可接受的载体。

4 5 . 一种药物组合物，含有一种以刺激或抑制 P T P - D 蛋白质或糖蛋白有效量的分子。

4 6 . 一种治疗或防止主体中疾病的方法，包括给所述主体施用权利要求 4 4 的所述组合物。

4 7 . 一种治疗或防止主体中疾病的方法，包括给所述主体施用权利要求 4 5 的所述组合物。

说明书

蛋白质酪氨酸磷酸酶的 P T P - D 亚家族

本发明涉及生物化学和细胞以及分子生物学领域。本发明涉及作为蛋白质酪氨酸磷酸酶 (PTPases) 新亚家族成员的 P T P 蛋白质或糖蛋白, 编码 P T P - D 蛋白质或糖蛋白的核酸构建体, 携带核酸构建体的重组表达载体, 含重组表达载体的细胞, 生产并鉴定 P T P - D 蛋白质和糖蛋白以及为此编码的 D N A 构建体的方法, P T P - D 蛋白质和糖蛋白特异性的抗体, 筛选能结合并抑制或刺激 P T P - D 蛋白质或糖蛋白的蛋白质酪氨酸磷酸酶活性的化合物的方法。

蛋白质的磷酸化是调节各种细胞过程的基本机制。大部分的蛋白质磷酸化作用发生在丝氨酸和苏氨酸残基, 由于发现许多致肿瘤基因产物和生长因子受体具有内蛋白质酪氨酸激酶活性, 所以在酪氨酸残基上的磷酸化作用很令人感兴趣。已经证实了蛋白质酪氨酸磷酸化在生长因子信号转导, 细胞周期演化以及肿瘤变态方面的重要性

(Hunter et al., Ann. Rev. Biochem. 54 : 987 - 930 (1985), Ullrich et al., Cell 61 : 203 - 212 (1990), Nurse, Nature 344 : 503 - 508 (1990), Cantley et al., Cell 64 : 281 - 302 (1991))。

生化研究已经表明, 在各种细胞蛋白质酪氨酸残基上的磷酸化作用是涉及竞争磷酸化和去磷酸化反应的动态过程。由蛋白质酪氨酸激

酶 (PTKases) 和蛋白质酪氨酸磷酸酶 (PTPases) 的相应作用介导调节蛋白质酪氨酸的磷酸化作用。由PTKases 催化酪氨酸磷酸化反应。通过PTPases 的作用可以特异性地将酪氨酸磷酸化的蛋白质去磷酸化。通过平衡PTKases 和PTPases 的活性, 确定细胞内物质的蛋白质酪氨酸磷酸化水平。(Hunter, T., Cell 58 : 1013 - 1016 (1989))。

蛋白质酪氨酸激酶 (PTKases) 是一个大的蛋白质家族, 包括许多生长因子受体和潜在的致肿瘤基因。(Hanks et al., Science 241 : 42 - 52 (1988))。已经将许多PTKases 与诱发细胞周期所需的起始信号 (Weaver et al., Mol. and cell. Biol. 11 (9) : 4415 - 4422 (1991)) 连接在一起。PTKases 包括不同家族的, 与丝氨酸/苏氨酸-特异性蛋白质激酶有共同祖先, 但与其有很大差别的酶 (Hanks, et al., 同上文)。在具有跨膜拓扑结构的受体-型PTKases 的情况下, 更能清楚地理解导致PTKases 活性改变的机制 (Ullrich et al. (1990) 同上文)。据认为特异性配体与受体-型PTKases 成员细胞外区的结合是为了诱发其导致酪氨酸激酶活性增加和信号转导途径激活的低聚化作用 (Ullrich et al., (1990) 同上文)。通过突变或过度表达而去调节激酶活性可以很好地证实细胞变形的机制 (Hunter et al., (1985) 同上文; Ullrich et al., (1990) 同上文)。

蛋白质磷酸酶至少由两种独立的且性质截然不同的家族 (Hunter,

T. (1989) 同上文), 蛋白质丝氨酸/苏氨酸磷酸酶和蛋白质酪氨酸磷酸酶 (PTPases) 组成。

蛋白质酪氨酸磷酸酶 (PTPases) 是一个已经被分为两个亚组的蛋白质家族。第一亚组由低分子量的细胞内酶组成, 所述酶含有单个保守的催化性磷酸酶区域。所有已知的胞内型PTPases 含有单个保守的催化性磷酸酶区。第一组PTPases 的实例包括 (1) 胎盘PTPases 1 B (Charbonneau et al., Proc. Natl. Acad. Sci. USA 86 : 5252 - 5256 (1989); Chernoff et al., Proc. Natl. Acad. Sci. USA 87 : 2735 - 2789 (1989)), (2) T细胞PTPase (Cool et al., Proc. Natl. Acad. Sci. USA 86 : 5257 - 5261 (1989)); (3) 大鼠脑PTPase (Guan et al., Proc. Natl. Acad. Sci. USA 87: 1501-1502 (1990)), (4) 神经元磷酸酶 (STEP) (Lombroso et al., Proc. Natl. Acad. Sci. USA 88 : 7242 - 7246 (1991)), 和 (5) 含有一个与细胞骨架蛋白质同源的区域 of 的细胞质磷酸酶 (Gu et al., Proc. Natl. Acad. Sci. USA 88 : 5867 - 57871 (1991); Yang et al., Proc. Natl. Acad. Sci. USA 88 : 5949 - 5953 (1991))。

第二亚组由高分子量, 称为 R-PTPases 的连接至PTPase 上的受体组成。R-PTPase 由 (a) 细胞内催化区, (b) 单个跨膜片段, 和 (c) 一个推测的配体-结合细胞外区组成 (Gebbinck et al., 同上文)。

R-PTPases 的 (c) 推测的配体-结合的细胞外“受体”区的结构和大小有很大不同。相反, (a) R-PTPases 的细胞内催化区是高度同源的。所有 R-PTPases 有两个串联复制的催化磷酸酶同源

区，除称为HPTP β 的 R - PTPase 外，它只有一个催化磷酸酶区。
(Tsai et al., Journal of Biol. Chem. 266 (16) : 10534 - 10543 (1991))。

R - PTPases 的一个实例是白细胞共同抗原 (LCA) (Ralph, S. J., EMBO J. 6 : 1251 - 1257 (1987))。LCA 是在所有白细胞和其造血细胞前身表面上表达的一个高分子量糖蛋白家族 (Thomas, Ann. Rev. Immunol. 7 : 339 - 369 (1989))。用几个种的 LCA 序列检测显著的相似程度 (Charbonneau et al., Proc. Natl. Acad. Sci. USA 85 : 7182 - 7186 (1988))。在文献中，将 LCA 以不同的名字命名，包括 T 2 0 0 (Trowbridge et al., Eur. J. Immunol. 6 : 557 - 562 (1962))，B 细胞形式的 B 2 2 0 (Coffman et al., Nature 289 : 681 - 683 (1981))，小鼠的同种异型标记 Ly - 5 (Komuro et al., Immunogenetics 1 : 452 - 456 (1975))，及最近的 CD 4 5 (Cobbold et al., Leucocyte Typing III, ed. A. J. McMichael et al., pp. 788 - 803 (1987))。

几种研究表明 CD 4 5 在 T 细胞激活中起重要作用。在 Weiss A., Ann. Rev. Genet. 25 : 487 - 510 (1991) 中总结了这些研究。在一个研究中，经 NSG 诱变处理并经选择不能表达 CD 4 5 的 T 细胞克隆对 T 细胞刺激产物应答，(Weaver et al., (1991) 同上文)。这些 T 细胞克隆在其对通过 T 细胞抗原受体转移的信号应答，包括适当靶的细胞溶解，增殖，产生淋巴激活素方面是功能缺陷的 (Weaver et al., (1991) 同上文)。

其它研究表明，CD 4 5 的 PTPase 活性在 pp 5 6^{lek}，一种淋巴细胞 - 特异的 PTKase 的激活中起作用 (Mustelin et al., Proc.

Natl. Acad. Sci. USA 86 : 6302 - 6306 (1989) ; Ostergaard et al., Proc. Natl. Acad. Sci. USA 86 : 8959 - 8963 (1989)) 。
这些作者假设 CD 4 5 的磷酸酶活性通过使 C 末端的酪氨酸残基去磷酸化而激活 pp 5 6 ^{1ex} ，这可能依次与 T 细胞激活有关。

R - PTPases 的另一实例是白细胞共同抗原相关分子 (L A R) (Streuli et al., J. Exp. Med. 168 : 1523 - 1530 (1988)) 。
开始，将 L A R 鉴别成 L C A 的同系物 (Streuli et al., 同上文) 。
虽然 (a) L A R 分子的细胞内催化区含两个催化磷酸酶同源区域 (区域 I 和区域 II) ，但突变分析表明，仅区域 I 有催化磷酸酶活性，而区域 II 是酶失活的 (Streuli et al., EMBO J. 9 (8) : 2399 - 2407 (1990)) 。在氨基酸 1 3 7 9 位的酪氨酸变成苯丙氨酸的化学诱导的 L A R 突变体是温度敏感的 (Tsai et al., J. Biol. Chem. 266 (16) : 10534 - 10543 (1991)) 。

已经克隆了一种新的小鼠 R - P T P ，叫做 mRPTP μ ，它具有与 L A R 共享某些结构 motifs 的 (a) 细胞外区域 (Gebbink et al., (1 9 9 1) 同上文) 。另外， 这些作者已经克隆了 RPTP μ 的人类同系物，并将其定位于人染色体 1 8 上的基因。

以 c D N A 克隆的序列为基础，已经预测了两个果蝇 PTPases，称为 DLAR 和 DPTP (Streuli et al., Proc. Natl. Acad. Sci. USA 86 : 8698 - 8702 (1989)) 。已经克隆了编码另一果蝇 R - PTPase 称为 DPTP 9 9 A 的 c D N A 并鉴定其特定 (Hariharan et al., Proc. Natl. Acad. Sci. USA 88 : 11266 - 11270 (1991)) 。

R - PTPases 的其它实例包括 R - PTPase - α ， β ， γ ， 和 ζ (Krueger et al., EMBO J. 9 : 3241 - 3252 (1990), Sap et al.,

Proc. Natl. Acad. Sci. USA 87 : 6112 - 6116 (1990), Kaplan et al., Proc. Natl. Acad. Sci. USA 87 : 7000 - 7004 (1990), Jirik et al., FEBS Lett. 273 : 239 - 242 (1990), Mathews et al., Proc. Natl. Acad. Sci. USA 87 : 4444 - 4448 (1990), Ohagi et al., Nucl. Acids Res. 18 : 7159 (1990))。公开的申请 WO 92 / 01050 公开了人 R - PTPase - α , β , 和 γ , 并报告了有关这三种 R - PTPases 及该蛋白质家族其它成员的保守区之间存在的结构同源性特点。鼠 R - PTPase - α 有 794 个氨基酸, 而人 R - PTPase - α 有 802 个氨基酸。R - PTPase - α 有一个与其它酪氨酸磷酸酶的催化区同源的细胞内区域。142 个氨基酸的细胞外区域 (包括 R - PTPase - α 的信号肽) 有很高的丝氨酸和苏氨酸含量 (32%) 和 8 个潜在的 N - 糖基化位点。已经生产出编码 R - PTPase - α 的 cDNA 克隆, 并从真核宿主中表达了 R - PTPase - α 。已经用 Northern 分析鉴别各种细胞和组织中天然表达的 R - PTPase - α 。通过用合成的 R - PTPase - α 肽免疫接种, 已经生产了 R - PTPase - α 的多克隆抗体, 用编码部分 R - PTPase - α 的 cDNA 克隆转染的细胞中, 该抗体可鉴别一个 130 KDa 的蛋白质。

R - PTPases 的另一实例是 HePTP (Jirik et al., FASEB J. 4 : 82082 (1990) Abstract 2253)。Jirik 等人用编码 LCA 的两个磷酸酶区的探针从肝胚细胞瘤细胞系 Hep G 2 中筛选 cDNA 文库, 并发现了一个编码新 R - PTPase, 命名为 HePTP 的 cDNA 克隆。看来, HePTP 基因在各种人和鼠细胞系及组织中都表达。

由于初步进行了 PTPase 的纯化, 定序、以及克隆, 因此, 可以迅速地鉴别其它的潜在的 PTPases。已经鉴别的大量不同的 PTPases

正在稳步增加，因此推测该家族可能与PTKase 家族一样大（Hunter（1989）同上文）。

已经鉴别并定义了已知PTPases 催化区中的保守氨基酸序列（Krueger et al., EMBO J. 9 : 3241 - 3252 (1990) 和 Yi et al., Mol. Cell. Biol. 12 : 836 - 846 (1992), 该文献插入本文作为参考）。本文将这些氨基酸序列称为“一致序列”。

Yi 等人列出了下列PTPases: LCA, PTPIB, TCPTP, LAR, DLAR 和HPTP α , HPTP β 和HPTP γ 的催化磷酸酶区序列。该组合包括下列“一致序列”（Yi et al., 同上文, 图 2 (A), 行 1 和 2）:

1. DYINAS/N [SEQ ID NO: 1]
2. CXXYWP [SEQ ID NO: 2]
3. I/VVMXXXXE [SEQ ID NO: 3]

Krueger 等人列出了PTPIB, TCPTP, LAR, LCA, HPTP α , β , δ , ϵ 和 ζ , 及DLAR 和DPTP 的催化磷酸酶区序列。该排列包括下列“一致序列”（Krueger et al., 同上文, 图 7, 行 1 和 2）:

1. D/NYINAS/N [SEQ ID NO: 4]
2. CXXYWP [SEQ ID NO: 2]
3. I/VVMXXXXE [SEQ ID NO: 3]

越来越清楚，酪氨酸残基的去磷酸化可能作为一种重要的调节机制单独起作用。已经表明，就酪氨酸激酶的src 家族来说，c 末端酪氨酸残基的去磷酸化可激活酪氨酸激酶活性（Hunter, T. Cell 49: 1 - 4 (1987)）。认为酪氨酸的去磷酸化是成熟 - 促进因子（MPF）激酶的有丝分裂激活中的专性步骤（Morla et al., Cell

58:193-203 (1989))。这些观察结果满足了本领域关于了解调节酪氨酸磷酸酶活性机制的需要。

很清楚需要进一步分析PTPases 之间结构-功能的关系，以增进对信号转变，细胞周期演化和细胞生长，以及肿瘤变态的了解的重要性。

表 I 给出了蛋白质化学家普遍使用且本文所用的氨基酸的一个字母缩写。

氨基酸名称	符号
甘氨酸	G
丙氨酸	A
缬氨酸	V
亮氨酸	L
异亮氨酸	I
丝氨酸	S
苏氨酸	T
半胱氨酸	C
甲硫氨酸	M
天冬氨酸	D
天冬酰胺	N
谷氨酸	E
谷氨酰胺	Q
精氨酸	R
赖氨酸	K

组氨酸	H
苯丙氨酸	F
酪氨酸	Y
色氨酸	W
脯氨酸	P
丝氨酸或天冬酰胺	S/N
天冬氨酸或天冬酰胺	D/N
异亮氨酸或缬氨酸	I/V
(非特定的氨基酸)	X

本发明人在本文中描述了蛋白质酪氨酸磷酸酶的新亚家族 (PTP-D) 的鉴别。该新亚家族 (下文称“PTP-D 亚家族”) 在结构上可与以前报告的PTPases 有显著的不同。本发明因此提供来自PTP-D 亚家族PTPases 的 P T P - D 蛋白质或糖蛋白。

最好, P T P - D 蛋白质或糖蛋白含有来自具有催化磷酸酶区的PTPases 的 P T P - D 亚家族的PTPase, 其中氨基酸序列选自:

1 . GYINAS/N [SEQ ID NO . 5]

2 . SXXYWP [SEQ ID NO . 6]

3 . IAMVXXXXE [SEQ ID NO . 7]

选自

1 . GYINAS/N [SEQ ID NO . 5]

2 . SXXYWP [SEQ ID NO . 6]

3 . IAMVXXXXE [SEQ ID NO . 7]

的氨基酸序列与前文在PTPases 催化磷酸酶区域中定义的氨基酸一致序列有下列氨基酸不同 (划线处表示不同):

(a) 在培养条件下，培养能表达 P T P - D 蛋白质或糖蛋白的宿主，

(b) 表达 P T P - D 蛋白质或糖蛋白；和

(c) 从培养物中回收 P T P - D 蛋白质或糖蛋白。

本发明也涉及特异于 P T P - D 蛋白质或糖蛋白或特异于 P T P - D 蛋白质或糖蛋白的抗原决定簇的抗体，多克隆抗体，单克隆抗体，或嵌合抗体。

本发明还涉及检测细胞或受试体中 P T P - D 蛋白质或糖蛋白存在，或检测其数量的方法，包括：

(a) 将所述的细胞或其提取物与 P T P - D 蛋白质或糖蛋白的抗原决定簇特异性抗体接触；和

(b) 检测与细胞或其提取物结合的抗体，或测量结合的抗体数量。

由此检测 P T P - D 蛋白质或糖蛋白的存在或测量其数量。

本发明也涉及检测细胞或受试体中核酸构建体存在的方法，该构建体编码正常或突变 P T P - D 或糖蛋白，该方法包括：

(a) 在杂交条件下，将来自受试体的细胞或提取物与编码至少部分正常或突变 P T P - D 蛋白质或糖蛋白的寡核苷酸探针接触；和

(b) 测量与细胞核酸杂交的探针，

由此检测核酸构建体的存在。

检测前，可以用多聚酶链反应选择性地扩增细胞的核酸。

本发明也涉及鉴别并分离化学或生物制品中，能与 P T P - D 蛋白质或糖蛋白结合的化合物的方法，所述方法包括：

- (a) 将 P T P - D 蛋白质或糖蛋白或其化合物 - 结合的部分吸附到固相基质上；
- (b) 将化学或生物制品与固相基质接触以便使化合物进行结合，并洗涤掉任何未结合的物质；
- (c) 检测与固相结合化合物的存在，且用于分离目的，
- (d) 洗脱结合的化合物，

由此分离该化合物。

最后，本发明包括一种鉴别分子的方法，该分子能刺激或抑制 P T P - D 蛋白质或糖蛋白的酶活性，该方法包括：

- (a) 将化合物与纯化形式，膜制品中或整个活的或固定细胞中的 P T P - D 蛋白质或糖蛋白接触；
- (b) 将步骤 (a) 中的混合物培养足够的时间间隔；
- (c) 测量 P T P - D 蛋白质或糖蛋白的酶活性；
- (d) 与没有该化合物存在时培养的 P T P - D 蛋白质或糖蛋白的酶活性进行比较，

由此确定该分子是刺激还是抑制活性。在该方法中，也可以用 P T P - D 蛋白质或糖蛋白片段鉴别能刺激或抑制该活性的分子。

图 1 表示 P T P - D 1 的部分 c D N A 序列和推测的氨基酸序列，它是 P C R 片段。

图 2 表示 P T P - D 2 的部分 c D N A 序列和推测的氨基酸序列，它是 P C R 片段。

图 3 A 表示 P T P - D 1 和 P T P - D 2 的推测氨基酸序列与

PTPase 1 B 氨基酸序列的比较，使用了CLUSTAL 程序 (Higgins, C. Fuch, R. and Bleasby, D., Multiple Sequence Alignment; CABIOS (1 9 9 1) (在印刷))。

图 3 B 表示 P T P - D 1 和 P T P - D 2 之间核苷酸的比较。

图 3 C 表示 P T P - D 1 和 P T P - D 2 之间氨基酸的比较。

图 4 表示 P T P - D 1 的部分 cD N A 序列和推测的氨基酸序列。该部分 cD N A 序列包括图 1 所示的 P C R 片段的 cD N A 序列。

本发明定义了结构与以前报告的PTPases 明显不同的蛋白质酪氨酸磷酸酶 (PTPases) 的新亚家族 (P T P - D 亚家族)。与以前定义的PTPases 催化磷酸酶区的氨基酸一致序列相比，该 P T P - D 亚家族成员的特征在于有一个、两个或三个下列氨基酸不同 (划线的不同)：

- 1 . P T P - D : GYINAS/N [SEQ ID NO. 5]
一致序列 : DYINAS/N [SEQ ID NO. 1]
NYINAS/N [SEQ ID NO. 8]
- 2 . P T P - D 1 / D 2 : SXXYWP [SEQ ID NO. 6]
一致序列 : CXXYWP [SEQ ID NO. 2]
- 3 . P T P - D 1 / D 2 : IAMVXXXXE [SEQ ID NO. 7]
一致序列 : (I/V/L)V(M/I/L)(V/L/I/M)XXXXE
[SEQ ID NO. 9]

术语“亚家族”用于说明一组结构上与上文限定的特定氨基酸残基相关的PTPases。

以前定义的氨基酸一致序列是指Krueger 等人, (EMBO J. 9 : 3241 - 3252 (1990) 和 Yi 等人 (Mol. Cell. Biol. 12 : 836 - 846 (1992) 中所述的已知PTPases 催化磷酸酶区中的保守氨基酸序列, 该文献引入本文作为参考。

相应地, 本发明涉及 P T P - D 蛋白质或糖蛋白, 含有来自 PTPases 的 P T P - D 亚家族的PTPase, 具有含 1、2、或 3 个选自下列氨基酸序列的催化磷酸酶区:

- | | |
|---------------|-------------------|
| 1 . GYINAS/N | [SEQ ID NO . 5] |
| 2 . SXXYWP | [SEQ ID NO . 6] |
| 3 . IAMVXXXXE | [SEQ ID NO . 7] |

目前, 还不知道PTPases 的新 P T P - D 亚家族的 P T P - D 蛋白质和糖蛋白是否受体结合的PTPases 还是细胞内PTPases。

一方面, 本发明涉及一种天然产生的哺乳动物 P T P - D 蛋白质或糖蛋白。另一方面, 本发明涉及一种重组哺乳动物 P T P - D 蛋白质或糖蛋白。另一方面, 本发明涉及一种化学合成的哺乳动物 P T P - D 蛋白质或糖蛋白。本发明优选的 P T P - D 蛋白质或糖蛋白是来自人的。

本发明提供一种基本上没有与其天然相关的其它蛋白质或糖蛋白的天然产生的 P T P - D 蛋白质或糖蛋白。“基本上没有其它蛋白质或糖蛋白”是指将 P T P - D 蛋白质或糖蛋白纯化, 去掉至少 9 0 % (以重量计), 如果需要甚至至少 9 9 % 的与 P T P - D 蛋白质或糖蛋白天然相连的其它蛋白质或糖蛋白, 且因此基本上没有它们。这可以经过将含 P T P - D 蛋白质或糖蛋白的细胞、组织或液体经标准蛋

白质纯化技术如带与该蛋白质特定反应的单克隆抗体的免疫吸附柱完成。其它形式的亲和纯化可利用能结合催化磷酸酶区的固相基质或与 P T P - D 蛋白质或糖蛋白中存在的受体区结合的配体来完成。另外，可以用组合的标准方法，如硫酸铵沉淀，分子筛层析，和离子交换层析完成纯化。

可以理解，可以从各种细胞和组织中，用生化方法纯化本发明的哺乳动物 P T P - D 蛋白质或糖蛋白。为了制备天然产生的 P T P - D 蛋白质或糖蛋白，组织如骨骼肌，特别是人的，是优选的。可以使用细胞系，如横纹肌肉瘤细胞系（R D）。

另外，由于可以分离或合成编码 P T P - D 蛋白质或糖蛋白的基因，所以如果需要，可以在原核生物或非哺乳动物真核生物中合成基本上没有其它哺乳动物蛋白质或糖蛋白的 P T P - D 蛋白质。按本发明的设想，在哺乳动物细胞，如转染的 C O S，N I H - 3 7 3 或 C H O 细胞中生产的重组 P T P - D 蛋白质或糖蛋白，例如要么是天然产生的蛋白质序列，要么是修饰的蛋白质序列，其中有氨基酸缺失和 / 或插入和 / 或取代。用重组方法生产天然产生的 P T P - D 蛋白质或糖蛋白时，基本上没有与其天然相连的其它蛋白质或糖蛋白。

另外，在固相支持物上合成所需序列的多肽并随后将其从支持物上分离出来的方法是已知的。

另一方面，本发明提供 P T P - D 蛋白质或糖蛋白的片段。术语“片段”用于说明通过适当修饰该编码 P T P - D 蛋白质或糖蛋白的 D N A 序列，导致 C - 末端，N - 末端一个或多个位点上和天然序列内缺失一个或多个氨基酸，而从具有天然产生的蛋白质序列的 P T P - D 或糖蛋白中得到的多肽。用 P T P - D 蛋白质或糖蛋白的片段筛

选作为拮抗剂或兴奋剂（如下文所定义的）的化合物。应明白，上述 P T P - D 蛋白质或糖蛋白片段可保留天然 P T P - D 蛋白质或糖蛋白的特征部分。特别是，上述 P T P - D 蛋白质或糖蛋白片段应保留作为完整 P T P - D 蛋白质或糖蛋白特征的一种或多种生物活性或功能。（不以任何方式限制本发明要求的范围的）P T P - D 片段的实例是：a) 催化区；b) 在完整细胞内，与其它分子相互作用的 P T P - D 蛋白质或糖蛋白区域；c) P T P - D 的调节区。

另一方面，本发明通过适当修饰编码 P T P - D 蛋白质或糖蛋白的 D N A 序列，导致 C 末端，N 末端的一个或多个位点上和天然序列内，增加一个或多个氨基酸，提供从天然产生的 P T P - D 蛋白质或糖蛋白中得到具有附加氨基酸的 P T P - D 蛋白质或糖蛋白。应理解，上述含有附加氨基酸的 P T P - D 蛋白质或糖蛋白可以保留天然 P T P - D 或糖蛋白的特征部分。尤其是，含附加氨基酸的上述 P T P - D 蛋白质或糖蛋白应保留作为完整 P T P - D 蛋白质或糖蛋白特征的一种或多种生物活性或功能。至少应被保留的上述特征的实例是：a) 催化活性；b) 底物特异性；c) 与完整细胞内其它分子的相互作用；d) P T P - D 的调节功能。这些实例不以任何方式限制本发明要求的范围。

另一方面，本发明通过适当修饰或突变编码 P T P - D 蛋白质或糖蛋白的 D N A 序列，导致 C 末端，N 末端，一个或多个位点上，和天然氨基酸序列内一个或多个氨基酸的取代，而提供了从天然产生的 P T P - D 蛋白质或糖蛋白中得到的含取代氨基酸的 P T P - D 蛋白质或糖蛋白。应明白，含取代氨基酸的上述 P T P - D 蛋白质或糖蛋白可保留 P T P - D 蛋白质或糖蛋白的特征部分。尤其是，含取代氨

基酸的上述 P T P - D 蛋白质或糖蛋白应保留作为完整 P T P - D 蛋白质或糖蛋白特征的一种或多种生物活性或功能。至少应保留的上述特征的实例是：a) 催化活性；b) 底物特异性；c) 在完整细胞内与其它分子的相互作用；d) P T P - D 的调节功能。这些实例不以任何方式限制本发明所要求的范围。

也可以将缺失、插入和取代的方法结合进行以得到 P T P - D 蛋白质或糖蛋白的最终构建体，提供的最终构建体具有完整 P T P - D 蛋白质或糖蛋白中所存在的所需活性或功能。上述活性和功能的实例是：a) 催化活性；b) 底物特异性；c) 在体外和体内与其它分子的相互作用；d) 调节功能。在缺失、插入和取代任意组合后，必需仅保留上述活性或功能之一。这些实例不以任何方式限制本发明所要求的范围。显然，在编码 P T P - D 蛋白质或糖蛋白的 D N A 中进行的修饰或突变必须不改变阅读框架并且最好不产生改变二级 m R N A 结构的互补区（参见欧洲专利申请 No. E P 7 5 , 4 4 4 ）。

在遗传水平上，通过位点特异诱变（如 Adelman et al., D N A 2 : 1 8 3 (1 9 8 3) 举证的）编码肽分子的 D N A 中的核苷酸，由此生产编码 P T P - D 蛋白质或糖蛋白，此后在重组细胞培养中表达 D N A （见下文），从而制备含氨基酸缺失，和 / 或插入，和 / 或取代的这些 P T P - D 蛋白质或糖蛋白。含氨基酸的典型缺失和 / 或插入和 / 或叠加的 P T P - D 蛋白质或糖蛋白在通常情况表现出与天然 P T P - D 蛋白质或糖蛋白性质一样的生物活性。

另外，用本领域已知的方法，通过直接的化学合成常规制备具有氨基酸缺失和 / 或插入和 / 或取代的 P T P - D 蛋白质或糖蛋白。

另一方面，本发明提供由其它 P T Pases 组成的所谓嵌合分子，其

中用来自 P T P - D 蛋白质或糖蛋白的同源序列代替一个或多个特定的氨基酸序列。嵌合分子包括，如，具有另一个 PTPase 的配体 - 结合的细胞外区的 P T P - D 蛋白质或糖蛋白，所述的细胞外区被移到 P T P - D 蛋白质或糖蛋白的一部分上。其它嵌合分子包括：a) 已经用 P T P - D 蛋白质的磷酸酶区取代了其催化磷酸酶区的其它 PTPases。这种情况下，优选的氨基酸数是 220 到 260 之间。b) 用其它 PTPases 的同源部分取代了其一部分或几部分催化区的 P T P - D 蛋白质或糖蛋白。c) 由一个 P T P - D 亚家族成员取代的嵌合分子，其中，其一部分或几部分已被 P T P - D 亚家族一个（或几个）其它成员的同源部分取代。

将“同源序列”定义为在两种或多种 PTPases 中，位于一级序列中的相似位置且可表现序列同源性的序列。应强调，“同源序列”不应限定为具有高度同源性的情况。嵌合分子是阐明结构 - 功能关系和鉴别特定化合物（药物）的重要工具。因此，最有用的嵌合体经常是（但不总是）其中分子的一部分已被另一分子中位于相似位置（但相异）的序列（换句话说，同源的）取代的分子。因此，交换的部位经常代表分子中区别最大的部分。

P T P - D 蛋白质或糖蛋白可含有不是正常 P T P - D 蛋白质或糖蛋白部分的其他化学部分。共价修饰物也包括在本发明范围内。通过使 P T P - D 蛋白质或糖蛋白的靶氨基酸残基与能与所选择的侧链或末端残基反应的有机衍生剂反应，将上述修饰物导入 P T P - D 蛋白质或糖蛋白中。

半胱氨酰残基通常都与 α - 卤代乙酸基（和相应的胺），如氯乙酸或氯乙酰胺反应以得到羧甲基或羧酰胺甲基衍生物。也可以通过与

溴三氟丙酮， α -溴- β -(5-imidazolyl)丙酸，氯乙酰磷酸酯（或盐），N-烷基马来酰亚胺，3-硝基-2-吡啶二硫化物，甲基-2-吡啶二硫化物，对氯高汞苯甲酸，2-氯汞基-4-硝基苯酚，或氯-7-2-噁-1,3-二唑反应衍生半胱氨酰残基。

通过与焦碳酸二乙酯在PH 5.5-7.0反应，衍生组氨酰残基，这是由于这种试剂对于组氨酰基侧链是特异性相关的。对-溴苯甲酰甲基溴也是有用的，最好是在卡可酸钠，PH 6.0完成反应。

赖氨酰和氨末端残基与琥珀或其它羧酸酐反应。用这些试剂衍生之后，影响或逆转赖氨酰残基的电荷。衍生含 α -氨基残基的其它适宜的试剂包括亚氨酸酯如methyl picolinimide，磷酸吡哆醛；吡哆醛；氯氢硼化物；三硝基苯磺酸；邻甲基异脲；2,4-戊二酮；和转氨酶-催化的与乙醛酸的反应。

通过与一种或几种常规试剂，包括苯甲酰甲醛，2,3-丁二酮，1,2-环己二酮，和茚三酮反应，修饰精氨酰残基。由于胍功能基团的高pKa，所以精氨酸残基的衍生要求在碱性条件下完成。此外，这些试剂可以与赖氨酸以及精氨酸 ϵ -氨基的基团反应。

特别有兴趣地在于将光谱标记导入酪氨酰残基，通过与芳香重氮化合物或四硝基甲烷反应，已经广泛地研究了酪氨酰残基本身的特定改变。通常，用N-乙酰基咪唑和四硝基甲烷分别形成邻-乙酰基酪氨酰和3-硝基衍生物。

通过反应碳化二亚胺($R^1-N-C-N-R^1$)如环己基-3-(2-咪唑基-(4-乙基)碳化二亚胺或1-乙基-3-(4-氨基-4,4-二甲基苯基)碳化二亚胺，可选择性地修饰羧基侧基团(天冬氨酰基或谷氨酰基)。此外，通过与铵离子反应，可将天冬

氨酰和谷氨酰残基变成天冬酰胺酰和谷氨酰胺酰残基。

经常将谷氨酰胺酰和天冬酰胺酰残基再酰胺化变成相应的谷氨酰和天冬氨酰残基。另外，在弱酸条件下再酰胺化这些残基。这些残基的两种形式都包括在本发明范围内。

将双功能试剂的衍生作用用于将肽交联到水不溶的支持物基质或其它大分子载体上。一般使用的交联剂包括：如，1，1-双(重氨基乙酰基)-2-苯乙烷，戊二醛，N-羧琥珀酰亚胺酯，例如，含叠氨基水杨酸的酯，同双官能亚氨酸酯，包括二琥珀酰亚胺酯如，3'3-二硫双(琥珀酰亚胺-丙酸酯)和双官能马来酰亚胺如双-N-马来酰亚胺基-1，8-辛烷。象甲基-3-[对-叠氨基苯基]二硫]丙酰亚胺酯这样的衍生剂产生在光存在下能形成交联的可光敏化中间产物。将美国专利Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; 和4,330,440中描述的易反应的水不溶基质用于蛋白质固定。

其它修饰作用包括脯氨酸和赖氨酸的羟基化，丝氨酸或苏氨酸残基羟基的磷酸化，赖氨酸，精氨酸和组氨酸侧链 α -氨基的甲基化(Creighton, T. E., Proteins: Structure and Molecule Properties, W. H. Freeman & Co., San Francisco, pp. 79 - 86 (1983))，N末端胺的乙酰化，以及某些情况下，C末端羧基的酰胺化。

上述衍生的部分可提高溶解性，吸附作用，生物半周期，等。这些部分可以可变地消除或减弱蛋白质等的任何不需的副作用。已公开了能介导上述作用的部分，如，在Remington's Pharmaceutical

Sciences, 16th. ed., Mack Publishing Co., Easton, PA (1980)。

另一方面，本发明涉及与（如图 1 和图 2 分别提供的）PTP-D 1 或 PTP-D 2 的氨基酸序列有 70% 或更多相同性的如上文定义的 PTP-D 蛋白质或糖蛋白。PTP-D 1 和 PTP-D 2 是在人骨骼肌中表达的 PTP-D 亚家族成员的 PTP-D 蛋白质。

另一方面，本发明涉及含有 PTP-D 1 或 PTP-D 2 的 PTP-D 蛋白质或糖蛋白。

已经证明，PTP-D 亚家族的成员，PTP-D 1 和 PTP-D 2 在人的骨骼肌中表达。因此，本发明涉及（但不以任何方式限制）在该组织中表达的 PTP-D 1 和 PTP-D 2 PTP-D 亚家族的成员。

另一方面，本发明涉及核酸构建体，它含有编码 PTP-D 蛋白质或糖蛋白，或编码具有氨基酸缺失和/或插入和/或取代的 PTP-D 蛋白质或糖蛋白的核苷酸序列。本发明还涉及表达载体如重组表达载体形式的核酸序列，以及含表达载体的原核和真核宿主细胞。

在本发明的另一方面，提供表达核酸构建体的方法，该构建体编码 PTP-D 蛋白质或糖蛋白。通过在促进上述 PTP-D 蛋白质或糖蛋白表达的条件下，在适当的营养培养基中培养细胞从而生产 PTP-D 蛋白质或糖蛋白。本领域任一普通专业人员可以知道用未充分实验的本发明核酸构建体和寡核苷酸如何鉴别并克隆与本文所述的 PTP-D 蛋白质或糖蛋白有同源性的或其他哺乳动物种的其他 PTP-D 蛋白质或糖蛋白。此外，本发明的核酸操作可以将来自特定 PTPase 的配体-结合的细胞外区移到 PTP-D 蛋白质或糖蛋白部分上，得到嵌合 PTP-D 蛋白质或糖蛋白。上述嵌合分子的非限制

实例包括含是一种表皮生长因子受体，成纤维细胞生长因子受体等配体—结合细胞外区的 P T P - D 蛋白质或糖蛋白。遗传工程方法加工的嵌合受体是现有技术已知的（参见，如 Riedel et al., Nature 324 : 628 - 670 (1986)）。在基因治疗中，可以使用如上述的，编码 P T P - D 蛋白质或糖蛋白，以及编码含氨基酸缺失和 / 或插入和 / 或取代的 P T P - D 蛋白质或糖蛋白，和编码嵌合 P T P - D 蛋白质或糖蛋白的核酸构建体。通过融合用正常 P T P - D 蛋白质或糖蛋白转染的所需谱系的细胞（如造血细胞，）代替导致疾病的不正常或机能障碍的 P T P - D 蛋白质或糖蛋白。另外，可将携带所选配体（如 E G F）受体的嵌合 P T P - D 蛋白质或糖蛋白的细胞用于上述基因治疗。

可通过各种方法，如 D N A 或 R N A 合成，或更好的，通过重组 D N A 技术生产作为本发明重组分子的核酸构建体。由，例如，Wu 等人（Prog. Nucl. Acid. Res, Molec. Biol. 21 : 101 - 141 (1978)）公开了有关合成上述分子的技术。由 Sambrook 等人（Molecular Cloning : A Laboratory Manual, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989)）公开了按上述方法生产构建重组分子的过程。

最好处理本发明重组 D N A 分子的 3' 末端，使其不适合于聚合作用。通过用化学方法封阻末端，或修饰末端碱基使得它们从空间上干扰多聚酶反应，从而完成上述处理。在优选的实施方案中，通过固定 3' 末端，如将其与固体支持物（如玻璃，塑料，乳汁，等）结合完成上述处理。支持物可以是任何形式的，如片，棒条，球形，卵形体等。上述固定的方法是本领域普通专业人员已知的。在最优选的实

施方案中，将重组DNA分子的3'末端共价地结合到固体支持物上。可以用间隔区域将探针延伸到固体支持物外，只有（1）它不会从空间上妨碍重组分子的功能或特征，和（2）间隔区的序列不参予检测中的杂交或聚合反应。通常需将数个，且最好大量的上述重组DNA分子固定到支持物上。

将代表PTP-D蛋白质或糖蛋白一部分的寡核苷酸用于筛选编码上述PTP-D蛋白质和糖蛋白的存在以及筛选PTP-D基因克隆。由，例如，Wu等人（同上文）公开了合成上述寡核苷酸的技术。

用溴化氰，或蛋白酶如木瓜蛋白酶，胰凝乳蛋白酶，胰蛋白酶等（Oike et al., *J. Biol. Chem.* 257 : 9751 - 9758 (1982); Liu et al., *Int. J. Pept. Protein Res.* 21 : 209 - 215 (1983)）将蛋白质分子切成片段。由于遗传密码是简并的，可以用一种以上密码子编码一种特定的氨基酸（Watson, J. D., In: *Molecular Biology of Gene*, 4th. ed. Benjamin / Cummings Publishing Co., Inc., Menlo Park, CA (1987)）。用遗传密码，可以鉴别一种或多种不同的寡核苷酸，其每一种均能编码氨基酸。在真核细胞中考虑不正常的碱基配对关系，和其中一种特定密码子实际使用（以编码一种特定氨基酸）的频率。可以估计事实上一种特定寡核苷酸组成实际XXX-编码序列的可能性。由Lathe等人（*J. Molec. Biol.* 183 : 1 - 12 (1985)）公开了上述“密码使用原则”。用Lathe的“密码使用原则”，可鉴别含理论上“最大可能的”能编码PTP-D序列的核苷酸序列的一种寡核苷酸，或一组寡核苷酸。

虽然偶然地，一种氨基酸序列可以仅由一种寡核苷酸编码，但经常地，氨基酸序列可以由任何一组相似的寡核苷酸编码。重要的是，

该组的所有成员都含有能编码肽片段的寡核苷酸，因此，潜在地含有与编码肽片段的基团相同的寡核苷酸序列，而这组中仅有一个成员含有与上述基因的核苷酸序列相同的核苷酸序列。由于该成员存在于该组内，并且，甚至该组其它成员存在的情况下能与DNA杂交，因此，可以在相同的方法中，使用未分离的一组寡核苷酸，其中只有1种作为一种寡核苷酸，以克隆编码该肽的基因。

将含理论上能编码PTP-D片段的“最可能的”序列的寡核苷酸，或一组寡核苷酸用于鉴别能与“最可能的”序列，或一组序列杂交的互补寡核苷酸或一组寡核苷酸序列。可以将含上述互补序列的寡核苷酸用作探针以鉴别并分离PTP-D基因（Sambrook et al., 同上文）。

鉴别（用上述方法），合成并用现有技术中已知的方法与DNA或最好是从能表达PTP-D基因的细胞中得到的cDNA制剂杂交能编码PTP-D基因片段的（或与所述寡核苷酸、或一组寡核苷酸互补的）适宜的寡核苷酸，或一组寡核苷酸。可以用本领域专业人员已知的方法（Belagaje et al., *J. Biol. Chem.* 254 : 5765-5780 (1979); Maniatis et al., In : *Molecular Mechanisms in the Control of Gene Expression*, Nierlich et al., Ed., Acad. Press, NY (1976); Wu et al., *Prog. Nucl. Acid Res. Molec. Biol.* 21 : 101 - 141 (1978); Khorana, R. G. *Science* 203 : 614 - 625 (1979)) 合成互补至“最可能”PTP-D肽编码序列的单链寡核苷酸分子。另外，用自动的合成仪器可以完成DNA合成。由Sambrook 等人，（同上文），和由Haymes 等人（In: *Nucleic Acid Hybridization, A Practical Approach*, IRL Press,

Washington, DC (1985)) (所述文献引入本文作为参考) 公开了核酸杂交技术。如, 或与上述那些相似的技术已经足够确保人醛脱氢酶 (Hsu et al., Proc. Natl. Acad. Sci. USA 82 : 3771 - 3775 (1985)), 粘连蛋白 (Suzuki et al., EMBO J. 4 : 2519 - 2524 (1985)), 人雌激素受体基因 (Walter et al., Proc. Natl. Acad. Sci. USA 82 : 7889 - 7893 (1985)), 组织型血纤维蛋白溶酶原激活剂 (Pennica et al., Nature 301 : 214 - 221 (1983)) 和人末期胎盘碱性磷酸酶互补 DNA (Kam et al., Proc. Natl. Acad. Sci. USA 82 : (715 - 8719 (1985)) 的克隆。

在其它克隆 P T P - D 基因途径中, 通过将 DNA 或最好 cDNA (来自能表达 P T P - D 的细胞) 克隆到表达载体中, 制备表达载体文库。然后从该文库中筛选能表达蛋白质的成员, 所述蛋白能与抗 - P T P - D 抗体结合, 并含有能编码多肽的核苷酸序列。所述多肽与 P T P - D 或其片段有相同的核苷酸序列。在该实施方案中, 从能表达 P T P - D 蛋白质的细胞中提取并纯化 DNA, 或最好是 cDNA。将纯化的 cDNA 制成片段 (经切割, 内切酶消化, 等) 以得到 DNA 或 cDNA 片段库。然后将来自该库的 DNA 或 cDNA 片段克隆到表达载体中, 以便得到其每个成员含唯一一个克隆 DNA 或 DNA 片段的表达载体基因文库。

“表达载体”是一种 (由于存在适宜的转录和 / 或转译控制序列) 能表达克隆到载体中的 DNA (或 cDNA) 分子并由此能生产多肽或蛋白质的载体。当将表达载体导入适当的宿主细胞, 则会发生克隆序列的表达。如果使用原核表达载体, 则适当的宿主细胞可以是任何能表达克隆序列的原核细胞。相似地, 如果使用真核表达载体,

则适当的宿主细胞可以是任何能表达克隆序列的真核细胞。重要的是，由于真核DNA可含有插入序列，且由于在原核细胞中不能正确地加工上述序列，所以，最好使用能表达PTP-D的细胞的cDNA以便生产原核基因表达载体文库。由Sambrook等人（同上文）公开了制备cDNA和生产基因文库的方法。

用常规技术，包括用于连接的平整末端或交错末端，限制酶消化以提供适当的末端，按适当的，碱性磷酸酶处理填平粘性末端，从而避免不必要的连接，并用适当的连接酶连接，可以将编码本发明PTP-D蛋白质或糖蛋白，或编码本发明有氨基酸缺失和/或插入和/或取代的PTP-D蛋白质或糖蛋白，或编码本发明嵌合分子的DNA序列与载体DNA重组。上述操作技术是由Sambrook等人（同上文）公开的并且是本领域已知的。

如果核酸构建体包含了含转录和转译调节信息的核苷酸序列并将上述序列“可操作地连接”到编码多肽的核苷酸序列中，则该核酸构建体，如DNA是能表达多肽的。一种可操作连接是其中以上述方法将调节DNA序列与需表达的DNA序列连接以便进行基因表达。基因表达所需的调节区的精确性质随生物不同而不同，但通过应包括一个启动子区，在原核细胞中，它含有启动子（引起RNA转录开始）以及，转录成RNA后，将发生蛋白质合成开始信号的DNA序列。上述区域通常包括涉及转录和转译开始的5'-非编码序列，如TATA盒，帽序列，CAAT序列，等。

如果需要，用上述方法也可以得到3'到蛋白质编码基因序列的非编码区。保留该区域作为转录终止调节序列，如终止和聚腺苷作用。因此，通过保留编码蛋白质DNA序列的天然邻接的3'区，可以提

供转录终止信号。若在表达宿主细胞中，转录终止信号的功能不能令人满意，则可取代宿主细胞中的3'功能区。

如果两个DNA序列之间连接的性质不(1)导致导入框移突变，(2)干扰启动子区序列指导PTP-D基因序列转录的能力，或(3)干扰由启动子区序列转录PTP-D基因序列的能力，则就说这两个DNA序列(如一种启动子区序列和一种PTPase编码区)是可操作连接的。如果启动子可影响该DNA序列的转录，则可将启动子区操作连接到DNA上。因此，为了表达蛋白质，由适当宿主识别的转录和转译信号是必需的。

启动子是一种能结合RNA聚合酶并启动“可操作连接的”核酸序列转录的双链DNA或RNA。按本文所用的，“启动子序列”是在由RNA聚合酶转录的DNA或RNA链上存在的启动子序列。

“启动子序列补体”是其序列为“启动子序列”补充物的核酸分子。因此，如果朝“启动子序列”或“启动子序列补体”延伸，则将毗连单链“启动子序列补码”的引物DNA或RNA或“启动子序列”延伸时，产生含功能启动子的双链分子。该功能启动子将指导与含“启动子序列”的双链分子的那条链(并不是含“启动子序列补体”的那条分子链)可操作相连的核酸分子的转录。

特定的RNA聚合酶对上述启动子表现出高度的特异性。已非常详细地定性了噬菌体T7，T3和SP-6的RNA聚合酶的特征，且它们表现出高度的启动子特异性。对每个RNA聚合酶特异的启动子序列也直接控制聚合酶只利用双螺旋DNA模板两个链中的一条链。由启动子序列的定向决定选择哪条链转录。由于通过将核苷酸5'磷酸加到3'羟基末端，仅仅是酶聚合RNA，所以这种选择确定了转

录的方向。

若以要么使两个序列都被转录到相同的RNA转录本上，要么允许转录一个RNA转录本，在一个序列中开始，延伸到第二个序列中的方式将核酸分子的两个序列互相连接起来，则称这两个序列是“可操作连接的”。如果在启动子序列中开始的转录会产生一个可操作连接的第二序列的RNA转录本，则，这两个序列，如一个启动子序列和任何其它的DNA或RNA“第二”序列是可操作连接的。为了进行“可操作连接”，这两个序列与另一个不必是紧密邻近的。

因此，如上文所述，为了起启动子的作用，启动子序列必须作为双链分子存在。为了本发明的目的，功能启动子序列的两个链一个叫做“转录本”链，一个叫“互补链”。“转录本”链是由RNA聚合酶转录的双螺旋的那条链（即作为转录模板的）。“互补”链是具有与“转录本”互补的序列且必须存在的，与“转录本”链杂交的链以便发生转录。因此，当启动子序列的“转录本”链与第二序列是可操作连接时，在聚合酶存在的情况下，“转录本”链与“互补”链的杂交可引起“转录本”链的转录，并且用“转录本”链作为模板可产生一个RNA转录本。

本发明的启动子序列可以是原核的，真核的或病毒的。适宜的启动子是可抑制的，或最好是组成型的。适宜的原核启动子的实例包括能识别T4(Malik et al., J. Biol. Chem. 263 : 1174-1181 (1984); Rosenberg et al., Gene 59 : 191 - 200 (1987); Shinedling et al., J. Molec. Biol. 195 : 471 - 480 (1987); Hu et al., Gene 42 : 21 - 30 (1986))。T3, Sp6, 和T7 (Chamberlin et al., Nature 228 : 227 - 231 (1970); Bailey et al., Proc.

Natl. Acad. Sci (U. S. A) 80 : 2814 - 2818 (1983); Davanlook et al., Proc. Natl. Acad. Sci (U. S. A) 81 : 2035 - 2039 (1984)) 聚合酶的启动子; 噬菌体 λ 的 P_R 和 P_L 启动子 (The Bacteriophage Lambda, Hershey, A. D. Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1973); Lambda II, Hendrix, R. W. Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1980)); 大肠杆菌的 *trp*, *recA*, 热激活, 和 *lacZ* 启动子, 枯草杆菌的 α 淀粉酶 (Ulmanen et al., J. Bacteriol. 162 : 176 - 182 (1985)) 和 $\zeta - 28$ - 特异性启动子 (Gilman et al., Gene 32 : 11 - 20 (1984)); 芽孢杆菌的噬菌体启动子 (Gryczan, T. J., In : The Molecular Biology of the Bacilli Academic Press, Inc. NY (1982)); 链霉菌启动子 (Ward et al., Mol. Gen. Genet. 203 : 468 - 478 (1986)); λ 噬菌体的 *int* 启动子; pBR 322 β - 乳糖酶的 *bla* 启动子, (pRB 325 的乙酰氨基霉素转移酶基因的 CAT 启动子); 等。由 Glick, B. R. (J. Ind. Microbiol. 1 : 277 - 282 (1987)); Cenatiempo, Y. (Biochimie 68 : 505 - 516 (1986)); Watson 等人 (In : Molecular Biology of Gene, Fourth. Edition. Benjamin Cummins, Menlo Park, CA (1987)); 和 Gottesman, S. (Ann. Rev. Genet. 18 : 415 - 442 (1984)) 综述了原核启动子。优选的真核启动子包括小鼠金属硫蛋白 I 基因的启动子 (Hamer et al., J. Mol. Appl. Gen. 1 : 273-288 (1982)); SV 40 早期启动子 (Benoist et al., Nature (London) 290 : 304 - 310 (1981)); 和酵母 *gal 4* 基因启动子 (Johnston et al., Proc. Natl. Acad. Sci (USA) 79 : 6971 - 6975 (1982); Silver

et al., Proc. Natl. Acad. Sci (USA) 81 : 5951 - 5955 (1984))。上文所列的所有文献均插入本文作为参考。

强启动子是优选的。上述优选的启动子的实例是识别 T 3 , S P 6 和 T 7 聚合酶的启动子, 小鼠金属硫蛋白 I 基因的 P_L 启动子。用于 P T P - D 真核表达最优选的启动子如在 pLSV 载体中启动转录的 SV 4 0 启动子 (Livneh et al., (1986) J. Biol. Chem. 261 : 12490 - 12497)。由 Watson 等人 (In : Molecular Biology of the Gene, Fourth Edition, Benjamin / Cummings Publishing Co., Inc., Menlo Park, (1987)) 公开了上述聚合酶识别位点的序列。

另一方面, 本发明涉及能特异性识别 P T P - D 蛋白质或糖蛋白或者能特异性识别 P T P - D 蛋白质或糖蛋白的抗原决定簇的抗体。

可以在 P T P - D 蛋白质或糖蛋白不正常表达或激活引起的疾病诊断方法中使用重组表达或天然产生的 P T P - D 蛋白质或糖蛋白, 和 / 或识别 P T P - D 蛋白质或糖蛋白的抗体。本发明提供判断主体中正常或突变 P T P - D 蛋白质或糖蛋白的存在或水平的方法。个体中, P T P - D 蛋白质或糖蛋白的缺乏, 或更典型地, 低水平表达, 突变 P T P - D 蛋白质或糖蛋白的存在可以作为对致肿瘤基因变态和癌症发生敏感度的重要预测指标。另外, 可能归因于对负调节不敏感的突变受体 / 酶系统, 或归因于体内刺激配体过多而产生的 P T P - D 蛋白质或糖蛋白的过度表达, 可作为对糖尿病敏感度的重要预测指标。

本发明也涉及用上述抗体检测细胞, 细胞或组织提取物, 或生物体液中 P T P - D 蛋白质或糖蛋白的存在, 或测量其数量或其浓度。

一方面, 本发明涉及检测细胞中 P T P - D 蛋白质或糖蛋白的存

在或测量其数量的方法，包括：

(a) 将所述细胞或其提取物与对 P T P - D 蛋白质或糖蛋白的抗原决定簇特异性的抗体接触；和

(b) 检测所述抗体与所述细胞或其提取物的结合，或测量结合的抗体量。

由此，检测所述 P T P - D 蛋白质或糖蛋白的存在或测量其数量。

术语“抗体”包括多克隆抗体，单克隆抗体 (mAbs)，嵌合抗体，抗个体基因型 (抗 - Id) 抗体。

多克隆抗体是从抗原免疫接种的动物血清中得到的抗体分子异源种群。

单克隆抗体 (mAbs) 是针对特异性抗原的抗体的基本同源的种群。可以用本领域专业人员已知的方法得到 MAb 参见，如，Kohler 等人，(Nature 256 : 495 - 497 (1975)) 和美国专利 No. 4, 376, 110。上述抗体可以是包括 IgG, IgM, IgE, IgA, GILD 的任何免疫球蛋白型和其任何亚型的。可在体外或体内培养生产本发明 mAbs 的杂交瘤。高滴定度 mAbs 产品的体内生产使其成为目前优选的生产方法。总之，将来自个体杂交瘤的细胞经腹膜内注射到姥鲛烷灌注的 BALB/C 小鼠中，以得到含高浓度所需 mAbs 的腹水。用本领域专业人员已知的柱层析法，可以从腹水，或培养上清液中纯化同型 IgM 或 IgG 的 MAb。

嵌合抗体是其不同部分来自不同动物种的分子，如含有鼠 MAb 的可变区和人免疫球蛋白不变区的抗体。嵌合抗体和其生产方法是现有技术已知的 (Cabilly et al., Proc. Natl. Acad. Sci USA 71 : 3273 - 3277 (1984); Morrison et al., Proc. Natl. Acad. Sci

USA 81 : 6851 - 6855 (1984); Boulianne et al., Nature 312 : 643 - 646 (1984); Cabilly et al., 欧洲专利申请 1 2 5 0 2 3 (1 9 8 4 年 1 1 月 1 4 日 公 开) ; Neuberger et al., Nature 314 : 268 - 270 (1985); Taniguchi et al., 欧洲专利申请 171496 (1 9 8 5 年 2 月 1 9 日 公 开) ; Morrison et al., 欧洲专利申请 1 7 3 4 9 4 (1 9 8 6 年 3 月 5 日 公 开) ; Neuberger et al., PCT 申请 WO 8 6 / 0 1 5 3 3 (1 9 8 6 年 3 月 1 3 日 公 开) ; Kudo et al., 欧洲专利申请 1 8 4 1 8 7 (1 9 8 6 年 6 月 1 1 日 公 开) ; Sahagan et al., J. Immunol. 137 : 1066 - 1074 (1986); Robinson et al., 国际专利申请 # P C T / U S 8 6 / 0 2 2 6 9 (1 9 8 7 年 5 月 7 日 公 开) ; Liu et al., Proc. Natl. Acad. Sci USA 84 : 3439 - 3443 (1987); Sun et al., Proc. Natl. Acad. Sci USA 84 : 214 - 218 (1987); Better et al., Science 140 : 1041 - 1043 (1988))。这些文献由此插入本文作为参考。

抗个体基因型 (抗 - Id) 抗体是一种抗体，它识别通常与抗体的抗原 - 结合位点相关的独特的决定簇。针对所要用制备的抗 - Id 的 mAb，通过免疫接种与 mAb 源相同种和基因型的动物 (如小鼠种) ，可以制备抗 - Id 抗体。通过产生针对这些个体基因型的决定簇 (抗 - Id 抗体) ，被免疫的动物将识别并应答免疫抗体的个体基因型决定簇。

也可以用抗 - Id 抗体作为“免疫原”以诱导另一动物中的免疫应答，产生所谓的抗 - 抗 - Id 抗体。抗 - 抗 - Id 可以与诱导抗 - Id 的最初 mAb 是抗原决定基相同的。因此，使用针对 mAb 个体基因型决定簇的抗体，可以鉴别表达相同特异性抗体的其它克隆。

相应地，可以在适当的动物，如BALB/c小鼠中，使用产生的抗本发明PTP-D蛋白质或糖蛋白的mAbs，以诱导抗-Id抗体。可以用上述免疫小鼠的脾细胞以生产分泌抗-Id mAbs的抗-Id杂交瘤。此外，可以将抗-Id mAbs与载体如匙孔蛾形血蓝蛋白(KLH)结合，并可将其用于免疫接种其它的BALB/c小鼠。这些小鼠的血清将含有抗-抗-Id抗体，所述抗体具有对PTP-D抗原决定基特异性的最初mAb的结合特性。

因此抗-Id mAbs有其自己的个体基因型抗原决定基，或具有与所判断的，如PTP-D蛋白质或糖蛋白抗原决定基结构相似的“idiotopes”。

术语“抗体”意思是包括能结合抗原的完整分子及其片段，如Fab和F(ab')₂。Fab和F(ab')₂片段没有完整抗体的Fc片段，能更快速地从循环中清除，且比完整抗体有更低的非特异性组织结合(Wahl et al., J. Nucl. Med. 24 : 316 - 325 (1983))。

应知道，按本文公开的用于完整抗体分子的方法，可以将本发明中使用的Fab和F(ab')₂以及抗体的其它片段用于检测并定量PTP-D蛋白质或糖蛋白。用酶，如木瓜酶(以生产Fab片段)或胃蛋白酶(以生产F(ab')₂片段)，经蛋白水解裂解，通常产生上述片段。

可以将本发明中使用的抗体，或抗体片段用于定量或定性地检测是否存在表达PTP-D蛋白质或糖蛋白的细胞。通过使用与光学显微镜、流式细胞计数或荧光检测相结合的荧光标记抗体(见下文)的免疫荧光技术，达到上述目的。

可以将本发明中使用的抗体(或其片段)用于组织学上，如在免疫荧光或免疫电子显微镜检查中，原位检测PTP-D蛋白质或糖蛋

白。从患者中取出组织学样品，并将本发明的标记抗体提供给上述样品，从而可完成原位检测。最好通过将标记抗体（或片段）用于或覆盖在生物学样品上，从而提供抗体（或片段）。通过使用上述方法，不仅可确定 P T P - D 蛋白质或糖蛋白的存在，而且可确定其在检测组织上的分布。使用本发明，普通技术人员可以很容易理解，可改变任何各种组织学方法（如染色方法）以便完成上述原位检测。用于 P T P - D 蛋白质或糖蛋白的上述检测方法，一般包括：存在能鉴别 P T P - D 蛋白质或糖蛋白的可检测标记抗体的情况下，培养生物样品，如生物体液，组织提取物，新鲜收集的细胞，如淋巴细胞或白细胞，或已在组织培养中培养的细胞，并用本领域许多已知技术中的任何一种方法检测抗体。

可以用固相支持物如硝酸纤维素，或其它能固定细胞，细胞颗粒或可溶蛋白质的固体支持物处理生物样品。然后，可以用适当的缓冲液洗涤支持物，接着用可检测标记的 P T P - D 特异性抗体处理。然后可以用缓冲液第二次洗涤固相支持物以除去未结合的抗体。用常规方法可以检测在所述固体支持物上结合的标记量。

“固相支持物”是指能结合抗原或抗体的任何支持物。已知的支持物，或载体包括玻璃，聚苯乙烯，聚丙烯，聚乙烯，葡聚糖，尼龙，淀粉酶，天然和改性的纤维素，聚丙烯酰胺，辉长岩，及磁铁矿。从本发明目的来说，载体的性质可以是溶解到一定程度或不可溶的。支持物可以有实质上任何可能的结构形状，只要被结合的分子能够与抗原或抗体结合。因此，支持物的形状可以是球形的，如小珠，或圆柱体，如检测试管的内表面或棒的外表面。另外，表面可以是平的，如片，检测条，等。优选的支持物包括聚苯乙烯珠。本领域专业人员还

知道许多其它的适用于结合抗体或抗原的载体，或者通过常规实验的使用，可以确定它们。

按已知的方法，可以确定所得的许多抗 - P T P - D 抗体的结合活性。使用常规实验，本领域专业人员可以确定每个检测的可操作和最佳检测条件。

按具体情况是常规的还是必需，可将其它步骤如洗涤、搅拌、摇动、过滤等加到检测分析中。

一个方法是将 P T P - D 特异性抗体结合到酶上并在酶免疫检测 (E I A) 中使用，其中 P T P - D 特异性抗体是可检测标记的。若此后与适当的底物接触，酶依次以上述方法与底物反应以产生如，用分光光度，荧光或肉眼法可检测的化学部分。可用于检测标记的酶包括，但不限于，苹果酸脱氢酶，葡萄球菌核酸酶， δ - 5 - 甾类异构酶，酵母醇脱氢酶， α - 磷酸甘油脱氢酶，磷酸丙糖异构酶，辣根过氧化物酶，碱性磷酸酶，天冬酰胺酶，葡糖氧化酶， β - 半乳糖苷酶，核糖核酸酶，脲酶，过氧化氢酶，葡糖 - 6 - 磷酸脱氢酶，葡糖淀粉酶和乙酰胆碱酯酶。通过使用有关酶生色底物的比色方法完成检测。与相似制备的标准物比较，通过肉眼比较底物的酶反应程度，完成检测。

可以用任何各种其它的免疫检测法完成检测。例如，用放射性标记的抗体或抗体片段，通过使用放射免疫检测法 (R I A) (参见，如 Work et al., Laboratory Techniques and Biochemistry in Molecular Biology, North Holland Publishing Company New York, (1 9 7 8)，该文献引入本文作为参考) 可检测 P T P - D 蛋白质或糖蛋白。借助于如 γ 计数器或闪烁计数器的方法或放射自显影检测

放射性同位素。

也可以用荧光化合物标记抗体。当荧光标记的抗体与适当波长的光接触量，由于荧光可检测其存在。最常用的荧光标记化合物是荧光素异硫氰酸盐，若丹明，藻红素，藻青素，别藻兰蛋白，邻-苯二醛和荧光胺。

用荧光发射金属如 ^{152}Eu 或镧系的其它金属，也可以可检测性标记抗体。用上述金属整合基团如二亚乙基三胺五乙酸（DTPA）或乙二胺四乙酸（EDTA），可将这些金属吸附到抗体上。

通过与化学发光化合物结合，也可以标记抗体使之可检测。检测化学反应期间产生的发光，可以确定化学发光-标记抗体的存在。特别有用的化学发光标记化合物的实例是鲁米诺，异鲁米诺，theromatic acridinium ester，咪唑，acridinium 盐和草酸酯。

同样地，也可以用生物发光化合物标记本发明的抗体。生物发光是生物系统中发现的一种化学发光类型，其中催化蛋白质提高了化学发光反应的效率。通过检测是否存在发光可以确定生物发光蛋白质的存在。用于标记目的的重要发光化合物是荧光素，荧光素酶和 aeluorin。

本发明也涉及检测主体中是否存在编码 P T P - D 蛋白质或糖蛋白的核酸构建体，或编码突变 P T P - D 蛋白质或糖蛋白的核酸构建体的方法，包括：

(a) 在杂交条件下，将所述主体的细胞或其提取物与编码至少部分所述正常或突变 P T P - D 蛋白质或糖蛋白的寡核苷酸探针接触，和

(b) 测量所述探针与所述细胞核酸进行杂交，由此检测所述核

酸构建体的存在。

该方法在步骤(a)前可包括步骤(c)。步骤(c)用于经多聚酶链反应，选择性地扩增所述细胞编码所述P T P - D蛋白质或糖蛋白的核酸数量。

用编码各种P T P - D蛋白质或糖蛋白(见上文)部分的寡核苷酸探针检测来自主体的细胞是否存在编码P T P - D蛋白质或糖蛋白的D N A或R N A序列。由，例如Wu等人(Prog. Nucl. Acid Res. Molec. Biol. 21 : 101 - 141 (1978))公开了合成上述探针的技术。优选的探针是指向核酸序列的探针，该核酸序列编码本发明P T P - D蛋白质或糖蛋白的至少4个氨基酸残基，且最好是5个氨基酸残基(参见下文实施例4)。用上述探针可进行定性或定量的检测。例如，用Northern分析(见下文实施例3)测量细胞或组织制品中P T P - D m R N A，如P T P - D 1 m R N A和P T P - D 2 m R N A的表达。

甚至用从个体中得到的很少量的D N A，接着使用选择性扩增技术，完成上述方法。能扩增纯化核酸片段的重组D N A方法已被认识很久了。一般，上述方法涉及将核酸片段导入D N A或R N A载体中，载体的克隆扩增，扩增核酸片段的回收。由Cohen等人(美国专利No. 4, 237, 224)；和Sambrook等人(同上文)(上述文献引入本文作为参考)提供上述方法的实施例。

最近，已经描述了能增加上述所需核酸分子浓度的体外酶法。该方法叫做“聚合酶链反应”或“P C R”(Mullis et al., Cold Spring Harbor Symp. Quant. Biol. 51 : 263 - 273 (1986)；Erlich, E P 50, 424；E P 84, 796；E P 258, 017；

EP 2 3 7, 3 6 2; Mullis, K., EP 2 0 1, 1 8 4; Mullis et al., US. 4, 6 8 3, 2 0 2; Erlich, H. U.S. 4, 5 8 2, 7 8 8; 和Saiki et al., U.S. 4, 6 8 3, 1 9 4)。

本发明也涉及鉴别化学或生物制品中能与 P T P - D 蛋白质或糖蛋白结合的化合物的方法，该方法包括：

- (a) 将所述 P T P - D 蛋白质或糖蛋白或其化合物结合的部分附着到固相基质上；
- (b) 将所述化学或生物制品与所述固相基质接触以使所述化合物结合，并洗涤掉未结合的物质；和
- (c) 检测与所述固相结合所述化合物的存在。

本发明也涉及从复合混合物中分离能与 P T P - D 蛋白质或糖蛋白结合的化合物的方法，该方法包括：

- (a) 将所述 P T P - D 蛋白质或糖蛋白，或化合物结合的部分附着到固相基质上；
- (b) 将所述的复合混合物与所述固相基质接触以使所述化合物结合，并洗涤掉任何未结合的物质；和
- (c) 洗脱所结合的化合物。

由此分离所述的化合物。

“能与 P T P - D 蛋白质或糖蛋白结合的化合物”是指与 P T P - D 外的磷酸酶区催化位点相互作用的天然产生或合成生产的分子。

“催化位点”是指含磷酸酶活性的 P T P - D 的最小邻近的部分。所述化合物可直接或间接地调整 P T P - D 蛋白质或糖蛋白的酶活性。上述化合物的实施是 (i) 与 P T P - D 蛋白质或糖蛋白相互作用且由 P T P - D 蛋白质或糖蛋白去磷酸化的细胞内蛋白质，(ii) 由其它

类型细胞生产的天然产生的分子。

P T P - D 蛋白质或糖蛋白的“化合物结合部分”是指催化位点外的分子部分。不是催化位点部分的 P T P - D 蛋白质或糖蛋白的任何部分都可以是化合物结合部分。通过蛋白水解裂解，接着用本领域专业人员已知的常规纯化方法，可以从天然产生或重组表达的 P T P - D 蛋白质或糖蛋白中制备“化合物结合的部分”。另外，可以用本领域专业人员已知的重组技术，通过在适当细胞中只表达 P T P - D 的这些部分来制备 P T P - D 蛋白质或糖蛋白的化合物结合的部分。

另一方面，本发明涉及筛选定义为直接或间接抑制 P T P - D 蛋白质或糖蛋白酶活性或激活的分子的拮抗剂的方法。另一方面，本发明涉及筛选定义为直接或间接提高 P T P - D 蛋白质或糖蛋白的酶活性或激活作用的兴奋剂的方法。

在筛选能激活或抑制磷酸酶活性，并由此影响细胞代谢主要途径的药物和其它试剂的方法中，使用本发明的 P T P - D 蛋白质或糖蛋白。通过将完整的 P T P - D 蛋白质或糖蛋白或 P T P - D 蛋白质或糖蛋白片段附着到固相基质上，产生一种亲和探针，可以用该探针，基于它的自结合活性，筛选与 P T P - D 蛋白质或糖蛋白有相互作用能力的生物产品或化学试剂。然后从亲和探针上洗脱纯化形式的结合物质。

可以用 P T P - D 蛋白质或糖蛋白，或者有氨基酸缺失和/或插入和/或取代以及有酶活性的 P T P - D 蛋白质或糖蛋白检测能增强或抑制磷酸酶活性的化合物。可以在体外系统检测所检测化合物改变磷酸酶活性的能力，其中将检测化合物加到纯化 P T P - D 蛋白质或糖蛋白，或者有氨基酸缺失和/或插入和/或取代并有酶活性的 PTP

- D 蛋白质或糖蛋白中，用本领域专业人员已知的标准酶法测定对酶活性的影响。

通过有限的蛋白水解处理天然产生的或重组表达的 P T P - D 蛋白质或糖蛋白，可以制备筛选中所用的 P T P - D 蛋白质或糖蛋白的适当片段。另外，通过重组技术可以生产 P T P - D 的适当片段。作为一个实例（该实例不以任何方式限制本发明所要求的范围），可以最好只使用用于筛选目的的催化区。可以用本领域专业人员已知的重组技术，在适宜的宿主（如大肠杆菌）中，生产单独或作为融合蛋白的一部分的上述催化区，它们只由对于酶活性所必需的最小数目的氨基酸组成。

另外，用活的或固定细胞，或从活的或固定细胞中得到的膜部分测定在整个细胞制剂中，一种化合物对 PTPase 活性的影响。该方法用于筛选直接对 P T P - D 蛋白质或糖蛋白的酶活性部分起作用的化合物。如果 P T P - D 分子或糖蛋白有细胞外受体部分，则该方法可用于筛选经细胞外受体部分起作用的化合物。将检测化合物与（表达大量本发明 P T P - D 蛋白质或糖蛋白的）细胞，如转染的 C O S 或 N I H - 3 T 3 细胞或从中得到的膜制剂一起培养。然后用现有技术已知的方法（Honegger, et al., Cell 51 : 199 - 209 (1987); Margolis et al., Cell 57 : 1101-1107 (1989)）测量细胞磷酸酪氨酸的数量。将结果与没有要检测化合物，或没有或有已知的 P T P - D 蛋白质或糖蛋白激活剂时得到的结果比较。在上述研究中，也可以测定在酪氨酸激酶激活剂存在的条件下，检测化合物的作用。

刺激 PTPase 活性的化合物会引起磷酸酪氨酸量的净减少，而抑制 PTPase 活性的化合物会引起磷酸酪氨酸量的增加。

生长因子受体是酪氨酸激酶的情况下，如表皮生长因子（EGF）和血小板-衍生的生长因子（PDGF）的受体，酪氨酸磷酸化都与细胞生长和肿瘤基因变态有关。导致去磷酸化的PTPase 激活可作为防止或抑制生长的反调节机制，并可能作为抵抗癌症的内源调节机制。因此，这种受体酶系统的突变或不良调节可能促进对癌症的敏感性。

胰岛素受体也是一种酪氨酸激酶，携带胰岛素受体的细胞中，酪氨酸的磷酸化可能与正常的生理功能有关。与细胞生长和癌症的情况相反PTPase 的激活会抵消胰岛素的作用。次正常的PTPase 水平或酶活性会起作用以除去正常的反调节机制。可能更重要的是，希望通过PTPase 的超-活性，或不适宜的激活，可抑制或完全阻止胰岛素对细胞的作用而导致糖尿病（胰岛素-抗性变种）。因此，对糖尿病敏感可能与PTPase 失调有关。

因此，本发明中鉴别正常或突变 P T P - D 基因，或测量与细胞或组织有关的 P T P - D 蛋白质或糖蛋白的数量或活性的方法，可作为鉴别易感癌症，糖尿病，或其它与细胞磷酸酪氨酸代谢改变有关的疾病的方法。

本发明也涉及使用药物组合物中的上述所鉴别的拮抗剂或兴奋剂，所述药物组合物用于治疗带有正常或不正常 P T P - D 蛋白质或糖蛋白表达的疾病。该组合物一般可以是用于系统化或局部的注射或输液的形式，并且例如，也可与适当的载体配制在一起用于注射或输液。

本发明也涉及预防或治疗与 P T P - D 蛋白质或糖蛋白的激活有关的方法，该方法包括，给患者施用其需要的，有效剂量的本发明 P T P - D 蛋白质或糖蛋白或本发明的抗体或者刺激或抑制本发明 P T P - D 蛋白质或糖蛋白酶活性的分子。

在下文所示的实施例中将进一步说明本发明。这些实施例不以任何方式限制本发明的范围。

实施例 1

用多聚酶链反应 (P C R) 鉴别新的PTPase 亚家族

用硫氰酸胍 / CsCl 方法 (Chirgwin et al., Biochem. 18 . 5293 - 5299 (1979)) 从人骨骼肌中分离总的 R N A 。在寡 (dT) 纤维素柱上分离 Poly (A⁺) R N A (Aviv et al., Proc. Natl. Acad. Sci. USA 58 : 1408 - 1412 (1972)) 。用寡 (dT) 引发和来自 Gibco BRL 的 Moloney Murine Leukemia Virus RNase H 逆转录酶 (Gaithersburg MD 20877 U.S.A) 按制造商的说明, 从 2 μg poly (A)⁺ 合成第一条 c D N A 链。

用多聚酶链反应 (Saiki et al., Science 239 : 487 - 491 (1988)) 分离相当于骨骼肌中所表达的PTPases 的 c D N A 。总之, 用 Gene Amp Kit (Perkin Elmer Cetus, Norwalk, CT, U.S.A) 用下列一组简并的寡核苷酸引物扩增从上文得到的人肌肉第一条 c D N A 链 (相当于约 5 0 mg) 。

有意义引物 (寡核苷酸 no. 5 8)

5' A (CT) TT (CT) TGG (ACG) (AG) AGATG (AG) T (TCGA) TGG 3'

[SEQ. ID NO. 1 0]

有意义引物 (寡核苷酸 no. 5 7)

5' CC (TCGA) A (CT) (AGT) CC (ATC) GC (AG) CT (GA) CAGTG 3'

[SEQ. ID NO. 1 1]

相当于下列氨基酸一致序列的引物 :

有意义引物 (寡核苷酸 # 5 8) F W X M X W

[SEQ ID NO. 1 2]

反意义引物 (寡核苷酸 # 5 7) H C S A G (S / I / V /) G

[SEQ ID NO. 1 3]

每个 P C R 循环包括在 9 4 °C 经 1 分钟的变性步骤, 3 7 °C 经 2 分钟的退火步骤, 7 2 °C, 2 分钟的延伸步骤。完成 3 0 到 4 0 个循环。将反应产物进行琼脂糖凝胶电泳。分离所期望大小的片段 (以已经描述的 PTPases 的结构为基础), 用 T 4 克隆系统 (Invitrogen, San Diego, California) 亚克隆并用标准技术 (如在 Current Protocols in Molecular Biology, eds. F. M. Ausubel et al., John Wiley & Sons, New York, (1988) 中描述的), 用 Sanger 等人描述 (Proc. Natl. Acad. Sci. USA 74 : 5463 - 5467 (1977)) 的酶链终止法定序 (Sequenase, U. S. Biochemicals)。在图 1 和图 2 中分别列出了两个 P C R 片段的部分 D N A 序列和推测的氨基酸序列, 叫做 P T P - D 1 和 P T P - D 2。用 CLUSTAL 程度 (Higgins, C., Fuch, R., and Bleasby, D., Multiple Sequence Alignment; CABIOS (1991) (在印刷)) 将推测的氨基酸序列与图 3 A 中的 PTPase 1 B 比较。

看起来, 两个片段都与其它已知的 PTPases 有明显的同源性, 但令人惊奇的是, 还有有关该酶类的未描述的特征 (用 University of Wisconsin, Genetics Computer Group Program 分析)。

下面列出了与以前描述的已知 PTPases 的一致序列相比, P T P - D 1 和 P T P - D 2 的特有特征 (划线处表示不同) :

1. P T P - D 1 / D 2 : SXXYWP [SEQ ID NO. 6]

一致序列: CXXYWP [SEQ ID NO. 2]

3. P T P - D 1 / D 2 : IAMVXXXXE [SEQ ID NO. 7]

一致序列: (I/V/L)V(M/I/L)(V/L/I/M)XXXXE
[SEQ ID NO. 9]

实施例 2

P T P - D 亚家族成员的cD N A 克隆

按实施例 1 中所述, 从人骨骼肌中制备mR N A。用Okayama 和 Berg (Mol. Cell. Biol. 2 : 161 - 170 (1982); Okayama and Berg, Mol. Cell. Biol. 3 : 280 - 289 (1983)) 描述的方法构建cD N A 文库。用pCDVI-PL制备引物片段 (Noma et al., Nature 319 : 640 - 646 (1986))。按最近的描述 (Boel et al., Bio Techniques 11 : xx (1991)) 用一个短的合成连接物作为第二条链的引物。按 H. Inuoue 等人 (Gene 96 : 23 - 28 (1990)) 的方案, 将大肠杆菌 D H 5 2 (Gibco BRL, Gaithersburg, MD 20877, U. S. A) 用于转化。转化后, 将细胞以每板 1 5, 0 0 0 - 2 0, 0 0 0 个菌落的密度在 L B 平板 (含 5 0 μg 氨苄青霉素 / ml) 上培养。

用标准菌落杂交技术 (Maniatis et al., Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, second edition (1988)) 筛选硝酸纤维素复制物滤膜 (Schleicher & Schuell, BA 8 5)。合成下列的寡核苷酸, 用 T 4 多核苷酸激酶和 [γ - ³²P] A T P (Amersham) 在 5' 末端标记, 并用于筛选

cDNA 文库：

5' ATA GCA ATG GTG ACA GCA GAA 3' [SEQ. ID NO. 14]

该寡核苷酸相当于实施例 1 中第 1 个 PCR 片段的氨基酸序列 Ile-Ala-Met-Val-Thr-Ala-Glu。将 50 ml 杂交溶液 (6 × SSC, 5 × Denhardt's 溶液, 0.05% SDS (Current Protocols in Molecular Biology, eds. F. M. Ausubel et al., John Wiley & Sons, New York (1988))) 中的 10 皮摩尔标记的寡核苷酸加到复制硝酸纤维滤膜中, 并在 42 °C 杂交 3 小时。在室温用 6 × SSC, 0.05% SDS 将滤膜洗涤 3 次, 在 42 °C 洗一次, 最后在 48 °C 洗一次。用标准技术 (Maniatis et al., Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, second edition (1988)) 分离经放射自显影鉴别的阳性菌落。在图 4 中显示了称为 PTP-D1 的一个阳性克隆的部分序列以及推测的氨基酸序列。该部分序列包括上述第 1 号 PCR 片段的序列, 从而证明所分离的 cDNA 克隆的等同性。再者, 与前面描述的 PTPases 比较结果表明 PTP-D1 至少有一个额外的独有特征 (不同处划线表示) :

3. PTP-D1 GYINAS [SEQ ID NO. 5]

一致序列 N/DYINAS/N [SEQ ID NO. 4]

实施例 3

PTP-D1 和 PTP-D2 的 Northern 印迹分析

按 Puissant 等人 (Bio Techniques 8 : 148 - 149 (1990)) 描述的酸性硫氰酸胍-酚-氯仿提取法, 从人骨骼肌中分离总 RNA。

在寡(dT)柱(Avir et al., Proc. Natl. Acad. Sci. USA 69 : 1408 - 1412 (1972))上分离poly(A)⁺RNA。将15 μg poly(A)⁺RNA放在检测PTP-D表达的泳道中,将7.5 μg放在分析PTP-D 2表达的泳道中,在琼脂糖-甲醛凝胶中分离RNA并用标准技术(Current Protocols in Molecular Biology, eds. F. M. Ausubel et al., John Wiley & Sons, New York (1988))吸印到硝酸纤维滤膜上。将滤膜与相当于PTP-D 1和PTP-D 2所示序列的³²P-标记的cDNA片段杂交。用Random Primers DNA Labeling System (Cat. no. 8187SA, Bethesda Research Laboratories, Gaithersburg, MD 20877, U. S. A)按制造商的说明,完成³²P-标记。随后,将滤膜暴露于X射线胶片。Northern印迹分析表明,骨骼肌中PTP-D 1和PTP-D的主要转录物分别为6.5 Kb和1.1 kb。

实施例 4

鉴别PTP-D亚家族的新成员

分别从下列组织和细胞系:骨骼肌、肝、胎盘, Hep G 2 (美国典型培养物收集中心(ATCC)HB 8 0 6 5) RD (ATCC CCL 136) (Puissant et al., Biotechniques 8 : 148 - 149 (1990))中分离总RNA。在寡(dT)柱(Avir et al., Proc. Natl. Acad. Sci. USA 69 : 1408 - 1412 (1972))上分离poly(A)⁺RNA。按实施例 1 中的描述完成第一条cDNA链的合成。用标准条件(PCR Technology-Principles and Applications for DNA Amplification,

Erlich, H. E. ed., Stockton Press, New York (1989)), 将来自上述组织和细胞系中的cDNA制剂单独进行多聚酶链反应。用下列引物进行扩增:

有意义引物:

(寡核苷酸no. 58, 参见本发明实施例1)与另两个反意义引物结合:

寡核苷酸no. 250 (相当于PTP-D1和PTP-D2的氨基酸序列CYATTG [SEQ ID NO. 15]):

C 5' AG (TCGA)CC (TCGA)GT (TCGA)GT (TCGA)GC (AG)TA (AG)CA
FR [SEQ. ID NO. 16]

寡核苷酸no. 251 (相当于PTP-D1的氨基酸序列QERTVW [SEQ ID NO. 17]):

5' GGT (TCGA)AC (TCGA)GT (TCGA)C (TG) (TC)TC (TC)T
[SEQ. ID NO. 18]

完成30到40个PCR循环。反应产物经琼脂糖凝胶电泳。分离所期望大小的片段(寡核苷酸nos. 58和250结合物约为190bp左右;寡核苷酸的Nos. 58和251结合物约为235bp左右),切成平整末端,并用标准技术(Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, (1988))亚克隆到pGEM3载体(Promega)中。用测序酶(United States Biochemical, Cleveland, Ohio 44122, U. S. A)经酶链终止方法(Sanger et al., Proc. Natl. Acad. Sci. USA 74: 5463-5467 (1977))测定亚克隆PCR片段的序列。将核苷酸序列和相应的氨基

酸序列与 P T P - D 1 和 P T P - D 2 的序列比较。将与 P T P - D 1 / P T P - D 2 和 / 或 P T P - D 1 / D 2 序列相比表现出 7 0 % 或更多相同性的克隆鉴定为本发明的 P T P - D 亚家族成员。

实施例 5

检测编码 P T P - D 蛋白质的核酸的存在

从细胞系 Hep G 2 (美国典型培养物收集中心 (ATCC) H B 8 0 6 5) 和 R D (ATCC CCL 1 3 6) (Puissant et al., Biotechniques 8 : 148 - 149 (1990)) 中分离总 R N A 。在寡 (dT) 柱 (Aviv & Leder, Proc. Natl. Acad. Sci. USA 69 : 1408 - 1412 (1972)) 上分离 poly (A) ⁺ R N A 。按实施例 1 中所述, 完成第一条 c D N A 链的合成。用标准条件 (PCR Technology - Principles and Application for DNA Amplification, Erlich, H. E. ed., Stockton Press, New York (1988)), 使 c D N A 进行多聚酶链反应。用下列引物进行扩增:

有意义引物 5' ATAGCAATGGTGACAGCAGAA 3'

[SEQ. ID NO. 1 9]

反意义引物:

5' CGCCC (A / G) A (C / T) (T / C / G / A) CC (T / C / G / A) GC (T / C / G / A) CT
(G / A) CAGTG [SEQ. ID NO. 2 0] 3'

完成 3 5 个循环。反应产物进行琼脂糖凝胶电泳。分离所需大小的片段 (3 6 0 bp), 切成平整末端, 并用标准技术 (Ausubel et al., Current Protocols in Molecular Biology, John Wiley &

Sons, New York, (1988)) 亚克隆到pGEM 3 载体 (Promega) 中。用测序酶 (United States Biochemical, Cleveland, Ohio 44122, U.S.A) 吸收使用酶链终止法 (Sanger et al., Proc. Natl. Acad. Sci. USA 74 : 5463 - 5467 (1977)) 进行定序。确定亚克隆 P C R 片段的 P T P - D 的相同性。

实施例 6

检测细胞中 P T P - D 蛋白质或糖蛋白的存在并对其定量

原核表达载体pGEX 的改变

为了提供来自 P T P D 1 (见下文) 的cD N A 片段, 用标准技术 (Current Protocols in Molecular Biology, eds. F. M. Ausubel et al., John Wiley & Sons, New York, 1988) 改变pGEX 2 T 载体 (Pharmacia, Uppsala, Sweden) 克隆位点。用限制酶 BamH I 和EcoR I 消化pGEX 2 T 载体, 并分离。将下列寡核苷酸连接到消化的pGEX 2 T 载体中。

5' GATCTCCGAATTCCATGGATCCAGGCCTCTAGAAGCTTAC 3'

[SEQ. ID NO. 2 1]

3' AGGCTTAAGGTACCTAGGTCCGGAGATCTTCGAATGTAA 5'

[SEQ. ID NO. 2 2]

由此得到有下列克隆位点的载体pGEX- A K 2 :

5' EcoR I , Nco I , BamH I , Stu I , Xba I , HindIII 3'

在大肠杆菌中表达 GST - P T P D 1 融合蛋白

用限制酶EcoR I和BglII消化编码P T P D 1的cDNA（实施例2）。消化后，用标准技术（Current Protocols in Molecular Biology, eds. F. M. Ausubel et al., John Wiley & Sons, New York, 1988）分离一个约1600bp的片段并将其连接到pGEX - A K 2（用EcoR I和BamH I消化）。插入片段相当于本发明图4中所示P T P D 1的编码区（即：它编码272个氨基酸）和3'非转译区的约800bp。将编码谷胱甘肽S - 转移酶和P T P D 1（Smith et al., Proc. Natl. Acad. Sci. USA 83 : 8703 - 8707 (1988)）的pGEX - A K 2 / P T P D 1载体构建体导入大肠杆菌的DH5 α 菌株（Cat. No. 8263SA, Bethesda Research Laboratories, Gaithersburg, MD）和SURE™（Cat. No. 200294, Stratagene, La Jolla, CA 92037）中。

使培养过夜的已转化大肠杆菌在LB培养基中生长并用新鲜的培养基以1 : 10稀释，并生长1小时。加入异丙基 - 1 - 硫代 - β - D - 吡喃半乳糖苷（IPTG）以达到0.5mM（DH5 α ）和5mM（SURE）的最终浓度，并进一步培养4小时。对照：1）含或不含IPTG的pGEX - A K 2；2）不加入IPTG的pGEX - A K 2 / P T P D 1。从包含体中将GST - P T P D 1融合蛋白分离为一种不可溶产物（在含1.0%（体积/体积）Triton \times 100的50mM N - 2 - 羟乙基哌嗪 - N' - 2 - 乙烷磺酸（HEPES）缓冲液，PH, 7.5中洗涤3次）或者按制造商说明用谷胱甘肽 - Sepharose 4B亲和层析（Cat. No. 17 - 0756 - 01 Pharmacia, Uppsala, Sweden）分离为一种可溶蛋白。

生产对 P T P D 1 特异性的抗体

用标准技术 (Practical Immunology 3rd. Edition, L. Hudson & F. C. Hay, Blackwell, Oxford (1989)) 生产对 P T P D 1 特异性的抗血清。简而言之, 将在 $200\ \mu\text{l}$ 磷酸缓冲盐中的 $200\ \mu\text{g}$ GST - P T P D 1 融合蛋白与等体积弗氏完全佐剂 (Sigma, Cat. No. F 5 8 8 1) 结合, 并肌肉内注射到两只新西兰兔的大腿肌肉中。每个兔接受 $100\ \mu\text{g}$ 的融合蛋白。第一次注射两个星期后, 进行辅助注射 (与开始免疫的方法一样, 但没有弗氏佐剂)。再过 2 个星期后, 从每个兔中取 $20\ \text{ml}$ 血液。在玻璃试管中, 使血液在室温凝集 1 小时, 从试管孔, 将凝块弄松散后, 离心。将血清转移到新试管中并在 $-20\ ^\circ\text{C}$ 贮存至使用。

为了除去与谷胱甘肽 S - 转移酶 (GST) 反应的抗体, 用上述方法 (在大肠杆菌中表达 GST - P T P D 1 融合蛋白), 使血清通过已用谷胱甘肽 S - 转移酶饱和的谷胱甘肽 - Sepharose 4 B 柱。利用 pGEX - A K 2 构建体生产 GST 蛋白质。使血清通过柱 3 次以确保完全除去抗 - GST 抗体。按下文所述 (“检测细胞系中 P T P - D 1 的存在并测量其数量”) 通过 Western 印迹评估除去的效率。

检测细胞系中 P T P - D 1 的存在并测量其数量

用抗 - P T P D 1 抗体检测哺乳动物细胞中 P T P D 1 的表达。按标准方法的荧光免疫将提供有关在特定细胞系和组织中表达的资料。甚至更重要的是, 可以用抗体制剂确定细胞系和组织中的数量。作为抗 - P T P D 1 抗体的后面申请的一个实施例, 下文将描述 R D 细胞系 (美国典型培养物收集中心 C C L 1 3 6) 中 P T P D 1 的检

测。应强调，该实施例不以任何方式限制使用可用于检测其它细胞和组织中存在的 P T P D 1 的抗体。同样地，在纯化 P T P D 1 和确定其它类型的检测方法时，该抗体制剂也是有用的。

用标准技术，在含有带 Hank's 平衡盐溶液的 2 倍正常浓度的氨基酸和维生素以及 10% (V/V) 胎牛血清 (Gibco-BRL) 的基本培养基 (Eagle, Cat No. 041-022570, Gibco, Life Technologies Ltd Paisley, Scotland) 中培养 R D (胚胎横纹肌肉瘤; 人) 细胞系。

用磷酸缓冲盐将细胞洗涤 2 次并除去上清液。将一个 10 cm 组织培养平板上的细胞溶解于 800 μ l Triton \times 100 溶菌缓冲液 (20 mM HEPES PH 7.5, 50 mM NaCl, 10% 甘油, 1.0% Triton \times 100, 1.5 mM MgCl₂, 4 mM 乙二醇双 (β -氨基乙醚) N, N, N', N' - 四乙酸 (EGTA, Sigma E D 2 S S), 10 μ g/ml 抑肽酶, 1 mM 氟化苯基乙基磺酰 (PMSF)), 离心并将上清液以等分贮存在 -80 $^{\circ}$ C, 直到使用。将 1 到 50 μ l 的这种溶菌产物与 25 μ l S D S 样品缓冲液 (62.5 mM Tris-Cl PH 7.0, 3.0% (W/V) S D S, 10% (V/V) 甘油, 10% 2-巯基乙醇, 和 0.05% (W/V) 溴酚蓝) 混合, 煮沸 5 分钟, 经 S D S - 聚丙烯酰胺凝胶电泳 (7.5%) 分离, 并用标准技术 (Burnette, W. N. (1981) Anlyt. Biochem. 112 : 195 - 201) 吸印在硝酸纤维素上。通过使用来自与 R D 细胞溶解产物平行的大肠杆菌生产的 G S T - P T P D 1 融合蛋白的限量。制定定量确定 P T P D 1 的标准曲线。将硝酸纤维滤膜与每升磷酸缓冲盐 (P B S) 中 2 克奶粉 (Carnation, 脱脂干牛奶, Carnation, Los Angeles) 一起培养 30

分钟以阻断非特异性结合，用含 0.02% (V/V) Tween 20 (Sigma P1379) (PBS-Tween) 和 0.2% (W/V) 明胶 (BioRad Cat. No. 170 - 6537, Richmond, California) 的 PBS 洗涤 2 次，用 PBS-Tween 洗涤 3 次，最后与 1:200 稀释 (用 PBS-Tween) 的上述抗-PTP D1 抗体制剂一起培养 4 小时。用 PBS-Tween 洗涤 3 次后，将滤膜与辣根过氧化物酶结合的山羊抗-兔 IgG (Cat. No. 170 - 6525, BioRad) 一起培养。用 PBS-Tween 将滤膜洗涤 3 次，按制造商的说明 (Cat. No. RPN 2106, Amersham, UK) 用增强化学发光的技术 (ECL) 确定兔抗的数量，由此确定 PTP D1 的数量。通过将 RD 细胞中得到的信号与用大肠杆菌产生的 GST-PTP D1 融合蛋白得到的标准曲线比较可以确定 RD 细胞系生产的 PTP D1 量。

实施例 7

鉴别刺激或抑制 PTP-D 蛋白质或糖蛋白酶活性的分子

用标准技术 (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, (1988)), 将含 PTP D1 完整编码区或其功能部分的 cDNA 插入到哺乳动物表达载体 pcDNA 1 (Cat. No. V490 - 20, Invitrogen, San Diego) 中。用由 Gorman 等人 (Virology 171: 377 - 385 (1989)) 描述的 293 细胞瞬时表达系统生产有酶活性的 PTP D1。用标准技术，在 5% CO₂，37 °C 时在用 10% (V/V) 胎牛血清 (Gibco) 补充的 Dulbecco's Modified Eagle Medium (Cat. No. 0041 - 02430,

Gibco, Life Technologies Ltd., Paisley, Scotland) 中培养293细胞。

将10 μ g 质粒载体 P T P - D 1 / pcDNA 1 与 0.5 ml 0.25 M CaCl₂ 和 0.5 ml 2 \times B B S (50 mM N, N - 双 (羧乙基) - 2 - 氨基乙烷 - 磺酸 (B E S) , 280 mM NaCl, 1.5 mM Na₂ HPO₄) 混合, 并如 Chen & Okayama (Mol. Cell. Biol. 7 : 2745 - 2752 (1987)) 描述, 将其用于转染在 10 cm 培养皿中的 1.5 \times 10⁶ - 203 细胞。加入 Ca - 磷酸盐 DNA 沉淀后, 细胞在 3% CO₂ 下, 于 37 $^{\circ}$ C 培养 24 小时, 并用由 10% 胎牛血清补充的 DMEM 洗涤一次, 并在 5% CO₂ 下, 在新鲜培养基中于 37 $^{\circ}$ C 再培养 24 小时。除去培养基, 并将细胞溶解在 1.0 ml 的溶解缓冲液 (20 mM HEPES PH 7.5, 150 mM NaCl, 10% 甘油, 1.0% Triton \times 100, 1.5 mM MgCl₂, 4 mM 乙二醇双 (β - 乙氨醚) N, N, N', N' - 四乙酸 (EGTA, Sigma E D 2 S S), 10 μ g/ml 抑肽酶, 1 mM PMSF) 中。在 4 $^{\circ}$ C 将细胞溶解物以 2500 \times g 离心 2 分钟。除去上清液, 将 100 μ l 等分样品迅速冷冻在液氮中并在 -70 $^{\circ}$ C 贮存至使用。

用不同的底物估算 P T P - D 1 磷酸酶活性的潜在抑制剂或刺激剂: 1) 磷酸对硝基苯酯 (p N P - P ; Sigma 104 - 0); 2) ³²P - 标记的 Raytide (Oncogene Science Inc. Manhasset, NY); 3) ³²P - 标记的牛髓磷脂碱性蛋白 (M B P)。进一步分析可降低或增加 P T P - D 1 对一种或多种这些底物的活性的底物。

基本按 N. K Tonks 等人 (J. Biol. Chem. 236 : 6731 - 6737 (1988)) 所述测定 P T P - D 1 对 p N P - P 的活性。用微滴定板,

在室温，将 $10\ \mu\text{l}$ 上述 293 溶解产物与 $100\ \mu\text{l}$ pNP-P (分别 30, 和 $100\ \text{mM}$) 一起培养。用 Dynatech MR 500 读数器每隔一分钟读一次吸收值。在加入 pNP-P 前 5 分钟，将需分析具刺激或抑制活性的底物加到 PTP-D1 / 293 细胞溶解产物中。

用 ^{32}P 标记 Raytide 和髓磷脂碱性蛋白

基本按 Krueger 等人 (EMBO J. 9 : 3241 - 3252 (1990)) 所述测量 PTP-D1 对 ^{32}P - 标记的 Raytide 的活性。按制造商的说明 (Oncogene Science) 作微小改变，使用酪氨酸激酶 p60^{c-arc} 将合成的肽 Raytide 进行 ^{32}P 标记。简而言之，将 $2\ \mu\text{l}$ 的 p60^{c-arc} 与 $20\ \mu\text{l}$ Raytide ($1\ \text{mg}/\text{ml}$) 和 $108\ \mu\text{l}$ 的激酶缓冲液 (含 $10\ \text{mM MgCl}_2$, $0.2\% (V/V)$ β -巯基乙醇, $30\ \mu\text{M ATP}$ 和 $50\ \mu\text{Ci} [\gamma - ^{32}\text{P}] \text{ATP}$ 的 $50\ \text{mM HEPES PH} 7.5$) 混合。将混合物在 $37\ ^\circ\text{C}$ 培养 16 小时，并通过加入 $500\ \mu\text{l}$ 溶于 $20\ \text{mM Na}_2\text{HPO}_4$ 和 $100\ \mu\text{l}$ $5\ \text{mg}/\text{ml}$ 的乙酰化牛血清白蛋白中的 $20\% (W/V)$ 三氯乙酸 (TCA)，来停止反应。将混合物离心，用 $20\% \text{TCA} / 20\ \text{mM NaH}_2\text{PO}_4$ 将沉淀洗涤 3 次，最后，重新溶解于 $0.2\ \text{M Tris-Cl pH} 8.0$ 中。

按 Guan 等人 (Nature 350 : 359 - 362 (1991)) 所述，用与标记 Raytide 相似的方法标记髓磷脂碱性蛋白 (Sigma)。在含下列组成的 $60\ \mu\text{l}$ 反应物中标记 $30\ \mu\text{g MBP}$: $50\ \text{mM HEPES}$ 缓冲液, $\text{PH} 7.5$, $10\ \text{mM MgCl}_2$, $0.067\% \beta$ -巯基乙醇; 包括 $150\ \mu\text{Ci} [\gamma - ^{32}\text{P}] \text{ATP}$ 的 $0.05\ \text{mM ATP}$, 和 $4\ \text{U p}43^{\text{v-abl}}$ 激酶 (Oncogene Science)。在 $30\ ^\circ\text{C}$ ，将混合物培养 60 分钟，通

过加入冰-冷三氯乙酸以达到 20% 的终浓度而停止反应。在冰上 30 分钟后，用 20% TCA 洗涤 3 次并重新溶于 100 μ l H₂O 中。

使用 Raytide 或 MBP 检测 PTPase 活性

将 5 μ l 10x PTPase 缓冲液 (25 mM HEPES pH 7.3, 5 mM EDTA, 10 mM 二硫苏糖醇) 与 a) 5 μ l ³²P-标记的 Raytide 或 MBP (相当于每分钟 10 - 20 \times 10⁴ 计算), b) 分别为 5, 10 和 25 μ l 的 PTP-D1/293 细胞溶解产物, 和 c) H₂O 混合以达到 50 μ l 的终体积。在 37 $^{\circ}$ C 培养 30 分钟后停止反应。如果是 Raytide, 通过加入 0.75 ml 的酸性活性炭混合物 (Krueger et al., EMBO J. 9: 3241 - 3252 (1990): 0.9 M HCl, 90 mM 焦磷酸钠, 2 mM NaH₂PO₄, 4% (V/V) Norit A (Sigma)) 终止反应。混合并离心后, 除去 400 μ l 的上清液, 并测量放射性活性量。若用 MBP 作为底物, 则加入 20% TCA (最终体积) 停止反应。测量上清液中的 ³²P 量。

在开始检测前 5 分钟, 将需分析其刺激或抑制活性的底物加到 PTP-D1/293 细胞中。

序列目录

1. 一般资料:

(i) 申请人: Ullrich, Axel

Moller, Niels P. H.

Moller, Karin B.

(ii) 发明的题目：蛋白质酪氨酸磷酸酶的 P T P - D 亚家族

(iii) 序列数：3 8

(iv) 有关的地址：

(A) 收信人：Sterne, Kessler, Goldstein & Fox

(B) 街道：1 2 2 5 Connecticut Avenun, N. W.

(C) 城市：华盛顿

(D) 州：D. C

(E) 国家：U. S. A

(F) 邮政编码：2 0 0 3 6

(v) 计算机可读形式：

(A) 介质类型：Floppy disk

(B) 计算机：IBM PC 兼容性

(C) 操作系统：PC - DOS / MS - DOS

(D) 软件：Patent In Release #1.0, Version #1.25

(vi) 目前的申请资料：

(A) 申请号：待授予

(B) 申请日：

(C) 分类号：

(viii) 代理人 / 代理资料：

(A) 姓名：Shaffer, Robin E.

(B) 登记号：34, 249

(C) 证明 / 公文号 : 1406. 0150000

(ix) 通讯资料 :

(A) 电话 : (202) 833-7533

(B) 传真 : (202) 833-8716

2. SEQ ID NO: 1 的资料

(i) 序列特征 :

(A) 长度 : 6 个氨基酸

(B) 类型 : 氨基酸

(D) 拓扑物型 : 两种都有

(ii) 分子类型 : 肽

(ix) 特征 :

(A) 名称 / 关键词 : 肽

(B) 位置 : 6

(D) 其它资料 : / 注 = “该氨基酸可以由天冬酰氨取代”。

(xi) 序列描述 : SEQ ID NO: 1 :

Asp Tyr Ile Asn Ala Ser

1

5

(2) SEQ ID NO: 2 的资料

(i) 序列特征 :

(A) 长度 : 6 个氨基酸

(B) 类型 : 氨基酸

(D) 拓扑结构 : 二种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO： 2

Cys Xaa Xaa Tyr Trp Pro
1 5

(2) SEQ ID NO: 3 的资料

(i) 序列特征：

(A) 长度：8 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：二种

(ii) 分子类型：肽

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：1

(D) 其它资料： / 注 = “该氨基酸可以由缬氨酸取代”。

(xi) 序列描述：SEQ ID NO： 3

Ile Val Met Xaa Xaa Xaa Xaa Glu
1 5

(2) SEQ ID NO: 4 的资料

(i) 序列特征：

(A) 长度：6 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：1

(D) 其它资料： / 注 = “该氨基酸可以由天冬酰胺取代”。

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：6

(D) 其它资料： / 注 = “该氨基酸可以由天冬酰胺取代”。

(xi) 序列描述：SEQ ID NO: 4

Asp Tyr Ile Asn Ala Ser

(2) SEQ ID NO: 5 的资料

(i) 序列特征：

(A) 长度：6 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：6

(D) 其它资料： / 注 = “该氨基酸可以由天冬酰胺取代”。

(xi) 序列描述：SEQ ID NO: 5

Gly Tyr Ile Asn Ala Ser

(2) SEQ ID NO: 6 的资料

(i) 序列特征：

(A) 长度：6 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 6：

Ser Xaa Xaa Tyr Trp Pro

(2) SEQ ID NO: 7 的资料

(i) 序列特征：

(A) 长度：9 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 7：

Ile Ala Met Val Xaa Xaa Xaa Xaa Glu

(2) SEQ ID NO: 8 的资料

(i) 序列特征：

(A) 长度：6 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：6

(D) 其它资料： / 注 = “该氨基酸可以由天冬酰胺取代”。

(xi) 序列描述：SEQ ID NO：8：

Asn Tyr Ile Asn Ala Ser

(2) SEQ ID NO: 9 的资料

(i) 序列特征：

(A) 长度：9 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：1

(D) 其它资料： / 注 = “该氨基酸可以由缬氨酸或亮氨酸取代”。

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：3

(D) 其它资料： / 注 = “该氨基酸可由亮氨酸或异亮氨酸取代”。

(ix) 特征：

(A) 名称 / 关键词：肽

(B) 位置：4

(D) 其它资料： / 注 = “该氨基酸可由亮氨酸，异亮氨酸
或甲硫氨酸取代”。

(xi) 序列描述：SEQ ID NO：9：

Ile Val Met Val Xaa Xaa Xaa Xaa Glu
1 5

(2) SEQ ID NO: 10 的资料

(i) 序列特征：

(A) 长度：30 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO：10：

ACTTTCTTGG ACGAGAGATG AGTTCGATGG 30

(2) SEQ ID NO: 11 的资料

(i) 序列特征：

(A) 长度：30 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 1 1 :

CCTCGAACTA GTCCATCGCA GCTGACAGTG

30

(2) SEQ ID NO: 1 2 的资料

(i) 序列特征：

(A) 长度：6 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 1 2 :

Phe Trp Xaa Met Xaa Trp

1

5

(2) SEQ ID NO: 1 3 的资料

(i) 序列特征：

(A) 长度：7 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(ix) 特征

(A) 名称 / 关键词：肽

(B) 位置：6

(2) SEQ ID NO: 1 6 的资料

(i) 序列特征：

(A) 长度：3 4 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 1 6：

AGTCGACCTC GAGTTCGAGT TCGAGCAGTA AGCA

34

(2) SEQ ID NO: 1 7 的资料

(i) 序列特征：

(A) 长度：6 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 1 7：

Gln Glu Arg Thr Val Trp

(2) SEQ ID NO: 1 8 的资料

(i) 序列特征：

(A) 长度：2 9 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 18:

GGTTCGAACT CGAGTTCGAC TGTCTCTCT

29

(2) SEQ ID NO: 19 的资料

(i) 序列特征：

(A) 长度：21 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 19:

ATAGCAATGG TGACAGCAGA A

21

(2) SEQ ID NO: 20 的资料

(i) 序列特征：

(A) 长度：23 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(ix) 特征：

(A) 名称 / 关键词：混合特征

(B) 位置：6

(D) 其它资料：/注 = “该核苷酸可以由鸟嘌呤取代”

(ix) 特征：

(A) 名称 / 关键词：混合特征

(B) 位置：8

(D) 其它资料：/注 = “该核苷酸可以由胸腺嘧啶取代”

(ix) 特征：

(A) 名称 / 关键词：混合特征

(B) 位置：9

(D) 其它资料：/注 = “该核苷酸可以由胞嘧啶，鸟嘌呤或腺嘌呤取代”

(ix) 特征：

(A) 名称 / 关键词：混合特征

(B) 位置：1 2

(D) 其它资料：/注 = “该核苷酸可以由胞嘧啶，鸟嘌呤或腺嘌呤取代”

(ix) 特征：

(A) 名称 / 关键词：混合特征

(B) 位置：1 5

(D) 其它资料：/注 = “该核苷酸可以由胞嘧啶，鸟嘌呤或腺嘌呤取代”

(ix) 特征：

(A) 名称 / 关键词：混合特征

(B) 位置：1 8

(D) 其它资料： / 注 = “该核苷酸可以由腺嘌呤取代”

(xi) 序列描述：SEQ ID NO： 2 0：

CGCCCAACTC CTGCTCTGCA CTG 23

(2) SEQ ID NO: 2 1 的资料

(i) 序列特征：

(A) 长度： 4 0 个碱基对

(B) 类型： 核酸

(C) 链型： 两种

(D) 拓扑结构： 两种

(ii) 分子类型： D N A

(xi) 序列描述：SEQ ID NO： 2 1：

GATCTCCGAA TTCCATGGAT CCAGGCCTCT AGAAGCTTAC 40

(2) SEQ ID NO: 2 2 的资料

(i) 序列特征：

(A) 长度： 4 0 个碱基对

(B) 类型： 核酸

(C) 链型： 两种

(D) 拓扑结构： 两种

(ii) 分子类型： D N A

(xi) 序列描述：SEQ ID NO： 2 2：

AATTGTAAGC TTCTAGAGGC CTGGATCCAT GGAATTCGGA 40

(2) SEQ ID NO: 2 3 的资料

(i) 序列特征：

(A) 长度：261个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 23：

```
TTTTGGCAGA TGGTATGGCA ACAGGGAATT GCAATTATAG CAATGCTCAC AGCAGAAGAG      60
GAGGGTNGAN NGGAGAAGAG CTTTAGGTAC TGGCCACGAC TTGGTTCCAG GCACAACACT      120
GTCACCTATG GAAGGTTTAA GATCACGACC CCGTTCGCA CAGACTCTGG CTGCTATGCC      180
ACACACAGGCC TGAAGATGAA GCACCTCCTT ACCGGGCAAG AGAGGACCGT CTGGCANNCTC      240
CAATACACAG ACTGGCCTGA A                                                    261
```

(2) SEQ ID NO: 24 的资料

(i) 序列特征：

(A) 长度：87个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 24：

```
Phe Trp Gln Met Val Trp Glu Gln Gly Ile Ala Ile Ile Ala Met Val
1                               5                               10
Thr Ala Glu Glu Glu Gly Xaa Xaa Glu Lys Ser Phe Arg Tyr Trp Pro
20                               25                               30
Arg Leu Gly Ser Arg His Asn Thr Val Thr Tyr Gly Arg Phe Lys Ile
35                               40                               45
Thr Thr Arg Phe Arg Thr Asp Ser Gly Cys Tyr Ala Thr Thr Gly Leu
```

50	55	60
Lys Met Lys His Leu	Leu Thr Gly Gln Glu Arg Thr Val Trp Xaa Leu	
65	70	75
Gln Tyr Thr Asp Trp	Pro Glu	
	85	

(2) SEQ ID NO: 2 5 的资料

(i) 序列特征：

(A) 长度：2 7 0 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：D N A

(xi) 序列描述：SEQ ID NO：2 5：

TTCTGGCGGA TGATCTGGGA GCAGGGAGTG AATGTGATTG CCATGGTCAC TGCAGAGGAG	60
GAGGGTGGAC GAACCAAAAG CCACCGATAC TGGCCCAAAC TAGGTTCAA GCACAGCTCA	120
GCCACCTATG GCAAGTTCAA GGTCAACCACG AAGTTTCGAA CGGATTCTGT TTGCTATGCA	180
ACCACGGGCT TGAAGGTCAA GCACCTTTTG TCTGGGNAAG AAAGCACGGT GTGGCATTTA	240
CAATATACTG ACTGGCCTGA CTTCGGCGCC	270

(2) SEQ ID NO: 2 6 的资料

(i) 序列特征：

(A) 长度：9 0 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO：2 6：

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

(2) SEQ ID NO: 27 的资料

(i) 序列特征:

(A) 长度: 60 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构:

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 27:

Met Val Trp Glu Gln Lys Ser Arg Gly Val Val Met Leu Asn Arg Val
 1 5 10 15
 Met Glu Lys Gly Ser Leu Lys Cys Ala Gln Tyr Trp Pro Gln Lys Glu
 20 25 30
 Glu Lys Glu Met Ile Phe Glu Asp Thr Asn Leu Lys Leu Thr Leu Ile
 35 40 45
 Ser Glu Asp Ile Lys Ser Tyr Tyr Thr Val Arg Gln
 50 55 60

(2) SEQ ID NO: 28 的资料

(i) 序列特征:

(A) 长度: 60 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 28 :

```
Met Val Trp Glu Gln Gly Ile Ala Ile Ile Ala Met Val Thr Ala Glu
1                               5                               10                               15
Glu Glu Gly Xaa Xaa Glu Lys Ser Phe Arg Tyr Trp Pro Arg Leu Gly
                20                               25                               30
Thr Arg His Asn Thr Val Thr Tyr Gly Arg Phe Lys Ile Thr Thr Arg
                35                               40                               45
Phe Arg Thr Asp Ser Gly Cys Tyr Ala Thr Thr Gly
                50                               55                               60
```

(2) SEQ ID NO: 29 的资料

(i) 序列特征：

(A) 长度：60 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 29 :

```
Met Ile Trp Glu Gln Gly Val Asn Val Ile Ala Met Val Thr Ala Glu
1                               5                               10                               15
Glu Glu Gly Gly Arg Ile Lys Ser His Arg Tyr Trp Pro Lys Leu Gly
                20                               25                               30
Ser Lys His Ser Ser Ala Thr Tyr Gly Lys Phe Lys Val Thr Thr Lys
                35                               40                               45
Phe Arg Thr Asp Ser Val Cys Tyr Ala Thr Thr Gly
                50                               55                               60
```

(2) SEQ ID NO: 30 的资料

(i) 序列特征：

(A) 长度：60 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 30：

```
Leu Glu Leu Glu Asn Leu Thr Thr Gln Glu Thr Arg Glu Ile Leu His
1                               5                               10           15
Phe His Tyr Thr Thr Trp Pro Asp Phe Gly Val Pro Glu Ser Pro Ala
                20                               25           30
Ser Phe Leu Asn Phe Leu Phe Lys Val Arg Glu Ser Gly Ser Leu Ser
                35                               40           45
Pro Glu His Gly Pro Val Val Val His Cys Ser Ala
50                               55           60
```

(2) SEQ ID NO: 31 的资料

(i) 序列特征：

(A) 长度：29 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 31：

```
Leu Lys Met Lys His Leu Leu Thr Gly Gln Glu Arg Thr Val Trp Xaa
1                               5                               10           15
Leu Gln Tyr Thr Asp Trp Pro Glu His Gly Cys Pro Glu
                20                               25
```

(2) SEQ ID NO: 32 的资料

(i) 序列特征：

(A) 长度：27 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 3 2 :

```
Leu Lys Val Lys His Leu Leu Ser Gly Xaa Glu Arg Thr Val Trp His
1           5           10           15
Leu Gln Tyr Thr Asp Trp Pro Asp Phe Gly Ala
                20           25
```

(2) SEQ ID NO: 3 3 的资料

(i) 序列特征：

(A) 长度：300 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 3 3 :

```
CATTTTTGGC AGATGGTATG GGAACAGGGA ATTGCAATTA TAGCAATGGT GACAGCAGAA      60
GAGGAGGGTN GANNGGAGAA GAGCTTTAGG TACTGGCCAC GACTTGGTTC CAGGCACAAC      120
ACTGTCACCT ATGGAAGGTT TAAGATCACG ACCCGGTTCC GCACAGACTC TGGCTGCTAT      180
GCCACCACAG GCCTGAAGAT GAAGCACCTC CTTACCGGGC AAGAGAGGAC CGTCTGGCNN      240
CTCCAATACA CAGACTGGCC TGAACATGGC TGTCCAGAAG ACCTCAAGGG AATTTTATCA      300
```

(2) SEQ ID NO: 3 4 的资料

(i) 序列特征：

(A) 长度：270 个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO: 34:

```
TTCTGGCGGA TGATCTGGGA GCAGGGAGTG AATGTGATTG CCATGGTCAC TGCAGAGGAG      60
GACGGTGGAC GAACCAAAG CCACCGATAC TGGCCCAAAC TAGGTTCAA GCACAGCTCA      120
GCCACCTATG GCAAGTTCAA GGTCACCACG AAGTTTCGAA CGGATTCTGT TTGCTATGCA      180
ACCACGGGCT TGAAGGTCAA GCACCTTTTG TCTGGGNAAG AAAGGACGGT GTGGCATTTA      240
CAATATACTG ACTGGCCTGA CTTGCGCGCC      270
```

(2) SEQ ID NO: 35 的资料

(i) 序列特征：

(A) 长度：150 个氨基酸

(B) 类型：氨基酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO: 35:

```
Glu Trp Asp Tyr Ile Ala Thr Gln Gly Pro Leu Gln Asn Thr Cys Gln
1           5           10
Asp Phe Trp Gln Met Val Trp Glu Gln Gly Ile Ala Ile Ile Ala Met
20          25          30
Val Thr Ala Glu Glu Glu Gly Xaa Xaa Glu Lys Ser Phe Arg Tyr Trp
35          40          45
Pro Arg Leu Gly Ser Arg His Asn Thr Tyr Thr Tyr Gly Arg Phe Lys
50          55          60
Ile Thr Thr Arg Phe Arg Thr Asp Ser Gly Cys Tyr Ala Thr Thr Gly
65          70          75          80
```


(A) 长度：877个碱基对

(B) 类型：核酸

(C) 链型：两种

(D) 拓扑结构：两种

(ii) 分子类型：DNA

(xi) 序列描述：SEQ ID NO：37：

```
GAATTCTTAA GAAACGGCTA GTTGATGGGG AGTGCTCAAC AGCACGACTC CCTGAAAATG      60
CAGAAAGAAA TCGATTCCRA GATGTTCTTC CTTATGATGA TCGGAGAGTG GAGTTGGTCC      120
CAACTAAAGA AAACAACACT GGTACATCA ACGCATCACA TATTAAGGTC TCTGTCAGTG      180
GAATCGAATG GGATTATATT GCCACACAGG GACCATTACA GAATACCTGT CAAGATTTTT      240
GCCACATGGT ATGGGAACAG GGAATTGCAA TTATAGCAAT GGTGACAGCA GAAGAGGAGG      300
GTNGANNGGA GAAGAGCTTT AGGTACTGGC CACGACTTGG TTCCAGGCAC AACACTGTCA      360
CCTATGGAAG GTTTAAGATC ACGACCCGGT TCCGCACAGA CTCTGGCTGC TATGCCACCA      420
CAGGCCTGAA GATGAAGCAC CTCCTTACCG GGCAAGAGAG GACCGTCTGG CNNCTCCAAT      480
ACACAGACTG GCCTGAACAT GGCTGTCCAG AAGACCTCAA GGGATTTTTA TCATATCTTG      540
AAGAGATCCA GTCTGTTCGA CGCCATACAA ATAGCACAAG TGATCCCCAA AGCCCCAACC      600
CTCCGTTGTT GGTCCACTGC AGTGCTGGGG TAGGAAGGAC TGGCGTGGTG ATTTTGTCGG      660
AGATCATGAT CGCCTGCCTG GAACACAATG AGGTGCTGGA CATCCCGAGA GTGCTGGACA      720
TGCTGAGGCA ACAGAGAATG ATGCTGGTGC ARACTCTCTG CCAGTACACA TTTGTGTACA      780
GAGTCCTCAT CCAGTTCCTG AAAAGCTCCA GGCTCATCTA AGCTCCCACA ATTTCTTACG      840
GGCCCACTCA TGTGAAGCGT TTACAGCTTA AAAAAAAA      878
```

(2) SEQ ID NO: 38 的资料

(i) 序列特征：

(A) 长度：272个氨基酸

(B) 类型：核酸

(D) 拓扑结构：两种

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO：38：

Ile Leu Lys Lys Arg Leu Val Asp Gly Glu Cys Ser Thr Ala Arg Leu
1 5 10 15
Pro Glu Asn Ala Clu Arg Asn Arg Phe Gln Asp Val Leu Pro Tyr Asp
20 25 30
Asp Ala Arg Val Glu Leu Val Pro Thr Lys Glu Asn Asn Thr Gly Tyr
35 40 45
Ile Asn Ala Ser His Ile Lys Val Ser Val Ser Gly Ile Glu Trp Asp
50 55 60
Tyr Ile Ala Thr Gln Gly Pro Leu Gln Asn Thr Cys Gln Asp Phe Trp
65 70 75 80
Gln Met Val Trp Glu Gln Gly Ile Ala Ile Ile Ala Met Val Thr Ala
85 90 95
Glu Glu Glu Gly Xaa Xaa Glu Lys Ser Phe Arg Tyr Trp Pro Arg Leu
100 105 110
Gly Ser Arg His Asn Thr Val Thr Tyr Gly Arg Phe Lys Ile Thr Thr
115 120 125
Arg Phe Arg Thr Asp Ser Gly Cys Tyr Ala Thr Thr Gly Leu Lys Met
130 135 140
Lys His Leu Leu Thr Gly Gln Glu Arg Thr Val Trp Xaa Leu Gln Tyr
145 150 155 160
Thr Asp Trp Pro Glu His Gly Cys Pro Glu Asp Leu Lys Gly Phe Leu
165 170 175
Ser Tyr Leu Glu Glu Ile Gln Ser Val Arg Arg His Thr Asn Ser Thr
180 185 190
Ser Asp Pro Gln Ser Pro Asn Pro Pro Leu Leu Val His Cys Ser Ala
195 200 205
Gly Val Gly Arg Thr Gly Val Val Ile Leu Ser Glu Ile Met Ile Ala
210 215 220
Cys Leu Glu His Asn Glu Val Leu Asp Ile Pro Arg Val Leu Asp Met
225 230 235 240
Leu Arg Gln Gln Arg Met Met Leu Val Gln Thr Leu Cys Gln Tyr Thr
245 250 255
Phe Val Tyr Arg Val Leu Ile Gln Phe Leu Lys Ser Ser Arg Leu Ile
260 265 270

说明书附图

10 30 50
llllggcagolgglatgggocogggoollgcoollolagcoollgglgacogcogooag
F W Q M V W E Q G I A I I A M V T A E E

70 90 110
gaggglxgoxxggogooagcilllogglactggccocgocllggllccaggcacocact
E G X X E K S F R Y W P R L G S R H N T

130 150 170
glcactolggooaglllaogalcocgocccggllccgcocogoclcggclgclalgcc
V T Y G R F K I T T R F R T D S G C Y A

190 210 230
occacoggcclgaogalgaogcacclcllaccgggcoogogaggocclggcxccl
T T G L K M K H L L T G Q E R T V W X L

250
coolocacogoclgccclgaa
Q Y T D W P E

图 1

10 30 50
llclggcggolgalclgggocogggoglgoolgtgallgccolgglcoclgcagaggag
F W R M I W E Q G V N V I A M V T A E E

70 90 110
gaggglggocgaoocooogccoccgolactggccoooclogglcsoogcocogclca
E G G R T K S H R Y W P K L G S K H S S

130 150 170
gccacclolggcoagllcsooglcocccogooaglllcgaoocggolclglllgclalgcc
A T Y G K F K V T T K F R T D S V C Y A

190 210 230
occacoggccllgaogglcsoogcaccllllgclggxooogooogocggclggcollto
T T G L K V K H L L S G X E R T V W H L

250 270
coololactgoclgccclgocllcggcgcc
Q Y T D W P D F G A

图 2

PIB MVWEQKSRGVVMLNRYMEKGS LKCAQYWPQKEEKEMIFEDTNLKLTLISEDIKSYT VRQ
 PD1- MVWEQGI A I I AMVTAEEEGXXEKSFRYWPRLGSRHNTV TYGRFKITTRFR TDSGCYATTG
 PD2- MIWEQGVNVIAMVTAEEEGGRTKSHRYWPKLGSKHSSATYGKFKVTTKFR TDSV CYATTG
 *.*** . *. * * .***. .. .*. * .

PIB LELENLTTQETREILHFHYTTWPDFGVPE SPASFLNFLFKVRESGSL SPEHGPV VVHCSA
 PD1- LKMKHLLTGQERTVWXLQYTDWPEHGCPE-----
 PD2- LKVKHLLSGXERTVMHLQYTDWPDFGA-----
 * . . * . * . . . ** ** . *

图 3A

质量: 195.0 长度: 845
 平均配对率: 1.000 比率: 0.722 缺口: 0
 相似性百分比: 72.222 相同性百分比: 70.000
 缺口重: 6.000
 长度重: 0.300 平均错配率: 0.000

P-D1 201 gatlllltggcagatgglatgggaacagggaoollgcaallatagcoolggl 250
 || |||| |||| | ||||| ||||| | | || || |||||
 || |||| |||| | ||||| ||||| | | || || |||||
 P-D2 1 ...llctggcggatgatctgggagcagggaglgaoalgatlgccatggl 47

251 gacagcagooagggaggglxgaxxggagooagaclllagglactggccac 300
 || ||||| ||||| ||||| · ||· || ||| | ||||| |||||
 || ||||| ||||| ||||| · ||· || ||| | ||||| |||||
 48 cacctgcagagggagggaggglggacgaoaccsooagccaccgatctggccca 97

301 gacttggllccoggcacaacactgacclatggaggltaagatcacg 350
 ||| ||||| | ||||| | | | ||||| ||||| | ||| ||| |||||
 ||| ||||| | ||||| | | | ||||| ||||| | ||| ||| |||||
 98 oactagglccaagcacagctcagccacctatggcoagllcaaggtcacc 147

351 occcggllccgcacagactctggctgclatgccaccocagggcctgaagat 400
 || ||| || || || || ||||| ||||| ||||| ||| ||||| |
 || ||| || || || || ||||| ||||| ||||| ||| ||||| |
 148 acgaoagllcgaocggallctglgllgclatgcoaccocgggcllgaaggl 197

401 gaogcaccclccltaccgggcaagagaggaccgctcggcxcclccaalaca 450
 || ||||| | | ||| · ||||| ||||| || ||||| · | ||||| |
 || ||||| | | |||| · ||||| ||||| || ||||| · | ||||| |
 198 caogcaccllllgcltgggxaagooaggocggltgagcaltlccaalata 247

451 cogactggcclgaoacolggctgclcagaagocctcaagggatllllatca 500
 | ||||| ||||| |||
 | ||||| ||||| |||
 248 ctgactggcclgactlcggcggc..... 270

图 3B

