

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200510114233.5

[51] Int. Cl.

C07K 19/00 (2006.01)

C12N 15/62 (2006.01)

C12N 15/63 (2006.01)

A61K 38/16 (2006.01)

A61P 31/18 (2006.01)

[43] 公开日 2006 年 4 月 26 日

[11] 公开号 CN 1763106A

[22] 申请日 2005.10.21

[21] 申请号 200510114233.5

[71] 申请人 中国科学院生物物理研究所

地址 100101 北京市朝阳区大屯路中国科学院
生物物理研究所

[72] 发明人 李志强 高光侠

[74] 专利代理机构 北京纪凯知识产权代理有限公司
代理人 关 畅

权利要求书 1 页 说明书 28 页 附图 2 页

[54] 发明名称

一种抗病毒的融合蛋白及其编码基因与应用

[57] 摘要

本发明公开了一种抗病毒的融合蛋白及其编码基因与应用。本发明的抗病毒融合蛋白，是具有下述氨基酸残基序列之一的蛋白质：1) 序列表中的 SEQ ID No: 1; 2) 将序列表中 SEQ ID No: 1 的氨基酸残基序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加且可降解 HIV 病毒 RNA 的蛋白质。该融合蛋白对 HIV 病毒感染具有明显的抑制作用，可用于制备预防和/或治疗艾滋病的药物。

- 1、一种抗病毒的融合蛋白，是具有下述氨基酸残基序列之一的蛋白质：
 - 1) 序列表中的 SEQ ID No: 1;
 - 2) 将序列表中 SEQ ID No: 1 的氨基酸残基序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加且可特异性降解 HIV 病毒 RNA 的蛋白质。
- 2、根据权利要求 1 所述的抗病毒的融合蛋白，其特征在于：所述病毒为 HIV 病毒。
- 3、权利要求 1 或 2 所述的抗病毒的融合蛋白的编码基因。
- 4、根据权利要求 3 所述基因，其特征在于：所述抗病毒的融合蛋白的编码基因具有下述核苷酸序列之一：
 - 1) 序列表中 SEQ ID No: 2 的核苷酸序列;
 - 2) 编码序列表中 SEQ ID No: 1 蛋白质序列的 DNA;
 - 3) 在高严谨条件下可与序列表中的 SEQ ID No: 2 限定的 DNA 序列杂交的核苷酸序列。
- 5、含有权利要求 3 或 4 所述抗病毒的融合蛋白的编码基因的表达载体。
- 6、含有权利要求 3 或 4 所述抗病毒的融合蛋白的编码基因的细胞系。
- 7、含有权利要求 3 或 4 所述抗病毒的融合蛋白的编码基因的宿主菌。
- 8、权利要求 1、2 所述的抗病毒的融合蛋白及其编码基因在制备抗病毒药物中的应用。
- 9、根据权利要求 8 所述的应用，其特征在于：所述病毒为 HIV 病毒。
- 10、根据权利要求 9 所述的应用，其特征在于：所述抗病毒药物为预防和/或治疗艾滋病的药物。

一种抗病毒的融合蛋白及其编码基因与应用

技术领域

本发明涉及一种抗病毒的融合蛋白及其编码基因与应用。

背景技术

核糖核酸（RNA）是生命活动中不可缺少的重要分子，具有传递遗传信息、编码蛋白质等许多功能。有害 RNA 的产生导致多种疾病的发生。

HIV 病毒是一种 RNA 病毒，其基因组 RNA 是病毒遗传信息的携带者，病毒 RNA 也是翻译病毒蛋白的模板。HIV 所引起的艾滋病是当今世界上对人类健康威胁最大的流行性传染病。在过去的十几年中，各国投入了巨大的努力和巨额资金致力于研制开发抗艾滋病的药物，目前世界上批准用于临床的抗艾滋病药物已有近二十种。但是，已有的任何一种药物都不能完全清除人体内的 HIV 病毒，几种药物组合使用的效果会优于单一药物治疗，但也不能完全清除 HIV 病毒。HIV 在复制过程中有很高的突变率，病人体内的病毒呈异源性，其中一些病毒具有抗药性，当治疗进行到一定阶段时，这些抗药病毒的数量会逐渐占据优势，使得现有的治疗无效。而且这些抗病毒药物的长期毒性已在临床上显示出来。开发一种无副作用而且不会产生抗药性的抗艾滋病药物是本领域每一个科学家的梦想。抗艾滋病药物的副作用来源于药物作用的不专一性；抗药性的产生则是因为 HIV 病毒具有非常高的突变率，使作为药物靶点的病毒蛋白可以通过少至一个氨基酸的改变而逃避药物的抑制作用，但不影响病毒蛋白自身的功能。因此，理想的抗艾滋病药物应对 HIV 病毒具有极高的特异性，对作用靶点的突变又有一定的包容性。

HIV 是一种逆转录病毒，在感染过程中首先通过病毒表面包膜蛋白（envelope）与其受体的相互作用吸附到细胞表面，然后病毒质膜与细胞质膜融合，裸露的病毒颗粒进入到细胞中。在细胞质中病毒的 RNA 被逆转录成双链 DNA，这段 DNA 存在于由病毒蛋白及宿主蛋白组成的整合前复合体（Pre-Integration Complex, 简称 PIC）中。PIC 由细胞质膜附近迁移到核膜附近并进入细胞核。在细胞核中病毒 DNA 由病毒的整

合酶整合到宿主细胞的染色体中。上述过程被称为逆转录病毒生命周期的前期。整合的病毒 DNA 被转录成 RNA，其中大部分经过剪切后作为模板翻译产生病毒蛋白，小部分全长 RNA 作为病毒遗传物质被病毒蛋白包装成病毒颗粒出芽到细胞外成为未成熟病毒。在未成熟病毒中一些病毒蛋白以融合蛋白形式存在，经过病毒蛋白酶剪切后病毒颗粒结构重组成为有感染性的成熟病毒，此过程称为逆转录病毒生命周期的后期。

在 HIV 复制过程中，病毒自身编码的 Tat 蛋白发挥着重要作用。Tat 是一种 RNA 结合蛋白，能够特异性结合到 HIV 病毒 RNA 上的 TAR 序列：gggtctctctggtagaccag atctgagcctgggagctctctggctaactaggaacccactgcttaagcctcaataaagct (ROBERT A. MARCINIAK, MARIANO A. GARCIA-BLANCO, AND PHILLIP A. SHARP. (1990) Identification and characterization of a HeLa nuclear protein that specifically binds to the trans-activation-response (TAR) element of human immunodeficiency virus. PNAS, Vol. 87, pp. 3624-3628.)。Tat 通过结合 TAR 序列实现 HIV 基因组的持续转录，是 HIV 生命周期中不可缺少的一环。

ZAP 是一个抗病毒因子，通过在胞质中特异性降解病毒 RNA 从而抑制某些病毒的复制，如鼠白血病病毒 (MLV)、辛德比斯病毒 (SIN)。但是 ZAP 对 HIV 病毒作用并不明显。ZAP 作用机理的研究结果表明，ZAP 本身不具有 RNA 酶的活性，但是能特异性结合到病毒 RNA 的特定序列上，并招募细胞内具有降解 RNA 功能的蛋白复合物 Exosome 而特异性地降解所结合的 RNA。

发明内容

本发明的目的是提供一种抗病毒的融合蛋白及其编码基因与应用。

本发明所提供的抗病毒的融合蛋白，名称为 ZAP-Tat，是具有下述氨基酸残基序列之一的蛋白质：

1) 序列表中的 SEQ ID No: 1;

2) 将序列表中 SEQ ID No: 1 的氨基酸残基序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加且可特异性降解 HIV 病毒 RNA 的蛋白质。

所述一个或几个氨基酸残基的取代和/或缺失和/或添加是指不多于十个氨基酸残基的取代和/或缺失和/或添加。

序列表中的 SEQ ID No: 1 由 904 个氨基酸残基组成。自氨基端的第 1 位至 776 位氨基酸残基为 ZAP, 自氨基端的第 777 位至 805 位氨基酸残基为 linker, 自氨基端的第 806 位至 881 位氨基酸残基为 Tat, 自氨基端的第 884 位至 893 位氨基酸残基为 myc, 自氨基端的第 899 位至 904 位氨基酸残基为 6×HIS。

上述抗病毒的融合蛋白 (ZAP-Tat) 的编码基因也属于本发明的保护范围。

抗病毒的融合蛋白 (ZAP-Tat) 的编码基因, 可具有下述核苷酸序列之一:

- 1) 序列表中 SEQ ID No: 2 的核苷酸序列;
- 2) 编码序列表中 SEQ ID No: 1 蛋白质序列的 DNA;
- 3) 在高严谨条件下可与序列表中的 SEQ ID No: 2 限定的 DNA 序列杂交的核苷酸序列。

序列表中的 SEQ ID No: 2 由 2712 个核苷酸组成。自 5' 端的第 1 位至 2328 位核苷酸为 ZAP 编码序列, 自 5' 端的第 2329 位至 2415 位核苷酸为 linker 的编码序列, 自 5' 端的第 2416 位至 2643 位核苷酸为 Tat 的编码序列, 自 5' 端的第 2650 位至 2679 位核苷酸为 myc 的编码序列, 自 5' 端的第 2695 位至 2712 位核苷酸为 myc 的编码序列。

所述高严谨条件可为用 0.1×SSPE (或 0.1×SSC), 0.1% SDS 的溶液, 在 65℃ 下杂交并洗膜。

含有本发明基因的表达载体、细胞系及宿主菌均属于本发明的保护范围。

本发明的融合蛋白包括可与病毒 TAR 序列特异结合的 RNA 结合蛋白亚基 (Tat) 和具有降解 RNA 功能的蛋白亚基 (ZAP), 该融合蛋白通过 RNA 结合蛋白亚基结合到特定的靶 RNA 上, RNA 降解亚基招募细胞内的 RNA 酶, 降解融合蛋白所结合的特定 RNA。本发明的融合蛋白对 HIV 病毒具有明显的抑制效果可用于制备抗病毒的药物, 如制备预防和/或治疗艾滋病的药物。

附图说明

图 1 为 ZAP-Tat 等质粒酶切电泳图。

图 2 为利用 Western 免疫杂交方法检测 ZAP-Tat 等在细胞内的表达。

图 3 为分析 ZAP-Tat 对 HIV 载体表达的抑制效果。

具体实施方式

下述实施例中的实验方法，如无特别说明，均为常规方法。

实施例 1、融合蛋白 ZAP-Tat 的表达

本实施例中，首先在 ZAP 的 C-端接上联接肽序列（命名为 ZAP-linker），并证明了该联接肽序列不影响 ZAP 的功能。然后，将 Tat 通过联接肽序列（linker）融合到 ZAP 的 C-端，成为 ZAP-Tat。使用联接肽序列的目的是将 ZAP 和 Tat 两个结构域分开，使之不互相干扰彼此的活性。作为对照，还构建了 N-端含有 69 个来自 ZAP 的氨基酸残基的 Tat（命名为 dTat）。

1、linker 的制备

利用化学合成方法合成 5' → 3' 链具有以下序列的一段双链 DNA：5' -GGT GGA GGC TCT GGG GGA GGT ACA GGA GGT GGC TCA GGT GGG GGC ACT GGT GGA GGC AGT GGA GGT GGC-3'。该序列编码一段联接肽序列。然后对其用如下的引物进行 PCR 扩增，所用引物引入克隆所需的酶切位点：

上游引物：linker sp:

5' - ATATAGGCGGCCGCTCGAGTGGTGGAGGCTCTGG-3'（下划线部分核苷酸为 *Not* I 酶切识别位点，方框内的核苷酸序列为 *Xho* I 酶切识别位点）

下游引物：linker ap:

5' - ATATAGGGGCCCTTGGCGCCGATATCGCCACCTCCACTGCCTC-3'（下划线部分核苷酸为 *Apa* I 酶切识别位点，方框内的核苷酸序列为 *Eco*R V 酶切识别位点，*GGCGCC* 为 *Kas* I 酶切识别位点）

PCR 的反应体系为：

1μl 模板 DNA (5pM/μl)，1μl 上游引物(50μM)，1μl 下游引物(50μM)，1μl Tag 酶，4μl dNTP 混合物(2.5mM)，5μl PCR 反应缓冲液 (10×)，37μl dH₂O。

PCR 的反应程序为：94℃退火 2min，然后 94℃，30sec；55℃，30sec；72℃，30sec 共 25 个循环。

得到的 PCR 产物进行凝胶电泳，结果表明得到 124bp 的条带，结果如图 1 中的泳道 2 所示，回收该条带的 DNA，对其测序表明所获得的 PCR 片段即包含 *Not*I—*Xho*I 一联接序列—*Eco*RV—*Kas*I—*Apa* I 的 DNA 片段。把该 DNA 片段命名为 linker。

图 1 中，泳道 1 和 9 是 DL2000 plus 的 marker；泳道 2 是 linker 的 PCR 产物；泳道 3 是 Tat 的 PCR 产物；泳道 4 是 pNZAP-Tat 用 *Xba* I 和 *Nhe* I 双酶切的片段；泳

道 5 是 pZAP-Tat 用 Nhe I 和 Not I 双酶切的片段；泳道 6 是 pNZAPlinker 用 EcoR I 和 Not I 双酶切的片段；泳道 7 是 pZAPlinker 用 Nhe I 和 Not I 双酶切的片段；泳道 8 是 pdTat 用 Xho I 和 EcoR I 双酶切的片段。

2、在质粒 pcDNA4/TO/myc-His B 中插入 linker:

pcDNA4/TO/myc-His B(从 Invitrogen 购买), 用 Not I 和 Apa I 双酶切 pcDNA4/TO/myc-His B, 切胶回收得到 5133bp 的大片段, 将步骤 1 所获得的 PCR 片段也用 Not I 和 Apa I 双酶切, 切胶回收得到 109bp 片段, 二者连接, 转化大肠杆菌 TOP10, 在含有 50 μ g/ml 氨苄青霉素的 LB 固体培养基上培养 12 小时, 挑取单克隆进行培养。提取质粒, 用引物 linker sp 和 linker ap 进行 PCR 验证, 得到 124bp 片段, 结果如图 1 所示。对质粒进行测序, 结果表明该质粒含有该段 linker 序列, 将该质粒命名为 pcDNA4/TO/linker/myc-His B。

3、在质粒 pcDNA4/TO/linker/myc-His B 中插入 NZAP 序列:

将 pcDNA4/TO/linker/myc-His B 用 EcoRI 和 NotI 双酶切, 切胶回收得到 5215bp 的大片段。质粒 pNZAP-myc (pNZAP-myc 的构建方法见: Guangxia Gao, Xuemin Guo, Stephen P. Goffl. (2002) Inhibition of Retroviral RNA Production by ZAP, a CCCH-Type Zinc Finger Protein. SCIENCE VOL 297, 1703-1706), 含有序列表中序列 14 (编码序列列表中序列 13 的 NZAPmyc) 的 NZAPmyc 编码基因, 也用 EcoRI 和 NotI 双酶切, 切胶回收得到 1088 的片段, 二者连接, 转化大肠杆菌 TOP10, 在含有 50 μ g/ml 氨苄青霉素的 LB 固体培养基上培养 12 小时, 挑取单克隆进行培养。提取质粒, 用 EcoRI 和 NotI 双酶切得到了 5215bp 的 pcDNA4/TO/linker/myc-His B 片段和 1088bp 的 NZAP 片段, 结果如图 1 中泳道 6 所示。对质粒进行测序, 结果表明该质粒含有序列表中序列 6 的核苷酸序列, 编码序列列表中序列 5 的 NZAP-linker-myc, 将该质粒命名为 pNZAPlinker。

4、在载体 pNZAPlinker 上插入 Tat 蛋白编码序列

以质粒 pCMVTat (pCMVTat 的构建方法见文献: NARASIMHACHAR SRINIVASAKUMAR, NATHALIE CHAZAL, C. HELGA-MARIA, SUSAN PRASAD, MARIE-LOUISE HAMMARSKJOLD, AND DAVID REKOSH. (1997) The Effect of Viral Regulatory Protein Expression on Gene Delivery by Human Immunodeficiency Virus Type 1 Vectors Produced in Stable Packaging Cell Lines. JVI, Vol. 71, No. 8, 5841 - 5848) 作为模板, 然后对其用如下的引物进行 PCR 扩增, 所用引物引入克隆所需的酶切位点:

上游引物: TAT SP:

5' - GGCCGCGCTGCTGAGCC - 3'

下游引物: TAT AP:

5' - ATATAGTTCGAACCGCGGCTGCTTTGATAGA - 3' (下划线部分核苷酸为 BstB I 酶切识别位点)

PCR 的反应体系为:

1 μ l 模板 (0.1 μ g/ μ l), 1 μ l 上游引物(50 μ M), 1 μ l 下游引物(50 μ M), 0.25 μ l Taq 酶, 4 μ l dNTP 混合物(2.5mM), 5 μ l PCR buffer (10 \times), 37.75 μ l dH₂O。

PCR 的反应程序为:

94 $^{\circ}$ C 退火 2min, 然后 94 $^{\circ}$ C, 30sec; 55 $^{\circ}$ C, 30sec; 72 $^{\circ}$ C, 30sec 共 25 个循环。

得到的 PCR 产物进行凝胶电泳, 结果表明得到 243bp 的条带, 结果如图 1 中的泳道 3 所示, 回收该条带的 DNA, 用 BstB I 酶切, 切胶回收含有 Tat 基因的小片段 233bp; 载体 pNZAPlinker 用 *EcoR* V 和 *BstB* I 酶切, 切胶回收含有 NZAP 和 linker 编码基因的大片段 6271bp, 二者连接, 转化大肠杆菌 TOP10, 在含有 50 μ g/ml 氨苄青霉素的 LB 固体培养基上培养 12 小时, 挑取单克隆进行培养, 提取质粒, 用 *Xba* I 和 *Nhe* I 双酶切, 得到 6027bp 和 477bp 的片段, 结果如图 1 中泳道 4 所示。该质粒命名为 pNZAP-Tat, 含有序列表中 SEQ ID No: 10 的核苷酸序列, 编码序列列表中序列 9 的 NZAP-Tat/myc。

5、ZAP-Tat 的表达载体 pZAP-Tat 的构建

将 pNZAP-Tat 用 *Afl* II 和 *Not* I 双酶切, 凝胶电泳回收 5372bp 大片段作为载体, pZAP-myc (pZAP-myc 的构建方法见: Guangxia Gao, Xuemin Guo, Stephen P. Goffl. (2002) Inhibition of Retroviral RNA Production by ZAP, a CCCH-Type Zinc Finger Protein. SCIENCE VOL 297, 1703-1706) 用 *Afl* II 和 *Not* I 双酶切, 回收 2701bp 的 ZAP 片段, 与上面制备的载体连接, 转化大肠杆菌 TOP10, 在含有 50 μ g/ml 氨苄青霉素的 LB 固体培养基上培养 12 小时, 挑取单克隆进行培养, 提取质粒用 *Nhe* I 和 *Not* I 双酶切, 得到 6141bp 的载体片段和 1932bp 片段, 结果如图 1 中泳道 5 所示, 该质粒命名为 pZAP-Tat, 含有序列表中序列 2 的融合蛋白 ZAP-Tat 的编码基因。

6、ZAP、ZAP-linker、dTat 的表达质粒的构建

含有序列表中序列 4 (编码序列列表中序列 3 的 ZAP-myc) 的 ZAP-myc 编码基因的质粒 pZAP-myc 的构建方法见文献: Guangxia Gao, Xuemin Guo, Stephen P. Goffl. (2002) Inhibition of Retroviral RNA Production by ZAP, a CCCH-Type Zinc Finger Protein. SCIENCE VOL 297, 1703-1706。

将质粒 pNZAPlinker 用 *Afl* II 和 *Not* I 双酶切, 凝胶电泳回收 5162bp 大片段作为载体, 质粒 pZAP-myc 也用 *Afl* II 和 *Not* I 双酶切, 回收 2701bp 的 ZAP 片段,

二者连接,转化大肠杆菌 TOP10,在含有 50 μ g/ml 氨苄青霉素的 LB 固体培养基上培养 12 小时,挑取单克隆进行培养,提取质粒用 Nhe I 和 Not I 双酶切,得到 5931bp 的载体片段和 1932bp 的片段,结果如图 1 中泳道 7 所示。该质粒命名为 pZAPlinker,含有序列表中序列 8 的核苷酸序列,编码序列列表中序列 7 的 ZAPlinker-myc。

将质粒 pZAP-Tat 用 Xho I 单酶切,回收 5651bp 的片段,用 Klenow (New England Biolab) 补平,自身连接,转化大肠杆菌 TOP10,在含有 50 μ g/ml 氨苄青霉素的 LB 固体培养基上培养 12 小时,挑取单克隆进行培养,提取质粒,用 Xho I 和 Not I 双酶切验证,结果如图 1 中泳道 8 所示。对用 Xho I 和 EcoR I 双酶切为单一片段的质粒测序,结果表明该质粒含有该段 dTat-myc 序列,将该质粒命名为 pdTat,含有序列表中序列 12 的核苷酸序列,编码序列列表中序列 11 的 dTat-myc。

7、ZAP、ZAP-linker、dTat、ZAP-Tat 的表达

将 pZAP-myc、pZAPlinker、pdTat、pZAP-Tat 分别用磷酸钙法转染入 293T 细胞(购自 ATCC),质粒用量为 5 μ g/3.5cm 培养皿,8 小时后换液,用 PBS 缓冲液洗一次,加入新鲜培养基 (DMEM: FBS (胎牛血清): PS (青霉素和链霉素) = 100: 10: 1 (体积比)),48 小时后收集细胞,裂解后进行 SDS-PAGE 电泳,然后将蛋白转到 PVDF 膜上,用 myc 抗体 (9E10) (购自 Sant Cruz) 进行 Western 免疫检测,检测到 ZAP、ZAP-linker、dTat、ZAP-Tat 的蛋白表达,结果如图 2 所示。

实施例 2、抗 HIV 功能检测

实施例 1 中,利用 Western 免疫杂交方法,首先证明了 ZAP、ZAP-linker、dTat、ZAP-Tat 在细胞内的表达,本实施例分析了上述蛋白对 HIV RNA 的抑制情况,方法如下:

将表达 ZAP、ZAPlinker、dTat、ZAP-Tat 的质粒与携带荧光素酶报告基因的 HIV 载体 pHR' -CMV-luc 共表达 (pHR' -CMV-luc 含有 HIV 的大部分骨架序列,包括 5' 和 3' LTR, 5' 剪切位点,包装信号,部分的 gag 基因编码序列, 3' 剪切位点等, pHR' -CMV-luc 的构建方法见 David Derse, Shawn A. Hill, Patricia A. Lloyd, Hye-kyung Chung, and Barry A. Morse. Examining Human T-Lymphotropic Virus Type 1 Infection and Replication by Cell-Free Infection with Recombinant Virus Vectors. J Virol. 2001 September; 75(18): 8461 - 8468), 分析上述蛋白对 HIV 载体表达的抑制情况。为了检测转染效率,使用了表达海肾 (Renilla) 荧光素酶的质粒 pRL-TK (购自 Promega) 作为转染效率的内参。pHR' -CMV-luc 和 pRL-TK 分别与下述四种质粒中的一种用磷酸钙法共转染 293A 细胞 (购自 ATCC): pcDNA4/TO/myc-His B, pZAP-myc, pZAPlinker 和 pZAP-Tat。所用质粒的质量比为:

pcDNA4/T0/myc-His B 或 pZAP-myc 或 pZAP-linker 或 pZAP-Tat: pHR' -CMV-luc: pRL-TK = 100: 10: 1。先用胰蛋白酶消化以便收获生长到对数期的 293A 细胞, 以 5×10^5 个细胞 / 3.5cm 培养皿的数量重新种入 3.5cm 培养皿, 加入新鲜 DMEM 培养基于 37°C、5% CO₂ 的培养箱中培养 16-24 小时。然后用质粒 pcDNA4/T0/myc-His B 或 pZAP 或 pZAP-linker 或 pZAP-Tat: 5μg, pHR' -CMV-luc: 0.5μg, pRL-TK: 0.05μg 的量, 磷酸钙法分别转染细胞。转染后 8 小时, 吸出培养皿中的培养基, 用 PBS 缓冲液洗一次, 重新加入新鲜培养基, 培养 48 小时后, 用 Promega 公司的 Dual-luciferase 试剂盒检测荧光素酶活性, 分别得到萤火虫荧光素酶和海肾荧光素酶的荧光值。海肾荧光素酶的数值用来作为内参, 表明转染效率的高低, 萤火虫荧光素酶的数值作为报告参数, 表明在某一转染效率下报告基因表达量的高低。萤火虫荧光素酶和海肾 (Renilla) 荧光素酶的荧光值比值是将每组数值做归一化处理, 便于与某一标准值比较。以空质粒 pcDNA4/T0/myc-His B 共转染组的萤火虫荧光素酶和海肾荧光素酶的荧光值的比值作为对照, 其它组的比值去除该值, 所得比值即为抑制倍数, 即: 抑制倍数=空质粒共转染组比值/功能蛋白共转染组比值。结果如图 3 所示, 表明 pZAP-Tat 表达的 ZAP-Tat 对 HIV 的抑制作用达到 16 倍以上, 其它质粒表达的蛋白对 HIV 的抑制作用不足 4 倍。图 3 中, pcDNA4 表示 pcDNA4/T0/myc-His B 和 pHR' -CMV-luc 和 pRL-TK 共转染组; ZAP 表示 pZAP-myc 和 pHR' -CMV-luc 和 pRL-TK 共转染组; ZAPlinker 表示 pZAPlinker 和 pHR' -CMV-luc 和 pRL-TK 共转染组; ZAP-Tat 表示 pZAP-Tat 和 pHR' -CMV-luc 和 pRL-TK 共转染组。

序列表

<160> 14

<210> 1

<211> 904

<212> PRT

<213> 人工序列

<220>

<223>

<400> 1

Met Ala Asp Pro Gly Val Cys Cys Phe Ile Thr Lys Ile Leu Cys Ala

5 10 15

His Gly Gly Arg Met Thr Leu Glu Glu Leu Leu Gly Glu Ile Arg Leu

20 25 30

Pro Glu Ala Gln Leu Tyr Glu Leu Leu Glu Thr Ala Gly Pro Asp Arg

35 40 45

Phe Val Leu Leu Glu Thr Gly Gly Gln Ala Gly Ile Thr Arg Ser Val

50 55 60

Val Ala Thr Thr Arg Ala Arg Val Cys Arg Arg Lys Tyr Cys Gln Arg

65 70 75 80

Pro Cys Asp Ser Leu His Leu Cys Lys Leu Asn Leu Leu Gly Arg Cys

85 90 95

His Tyr Ala Gln Ser Gln Arg Asn Leu Cys Lys Tyr Ser His Asp Val

100 105 110

Leu Ser Glu Gln Asn Phe Gln Ile Leu Lys Asn His Glu Leu Ser Gly

115 120 125

Leu Asn Gln Glu Glu Leu Ala Cys Leu Leu Val Gln Ser Asp Pro Phe

130 135 140

Phe Leu Pro Glu Ile Cys Lys Ser Tyr Lys Gly Glu Gly Arg Lys Gln

145 150 155 160

Thr Cys Gly Gln Pro Gln Pro Cys Glu Arg Leu His Ile Cys Glu His

165 170 175

Phe Thr Arg Gly Asn Cys Ser Tyr Leu Asn Cys Leu Arg Ser His Asn

180 185 190

Leu Met Asp Arg Lys Val Leu Thr Ile Met Arg Glu His Gly Leu Ser

195 200 205

Pro Asp Val Val Gln Asn Ile Gln Asp Ile Cys Asn Asn Lys His Ala

210 215 220

Arg Ala His Gln Asn Ser Gln Thr His Gln Ala Ser Leu Ser Lys Gln
 865 870 875 880
 Pro Arg Phe Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met His
 885 890 895
 Thr Gly His His His His His His
 900

<210> 2

<211> 2712

<212> DNA

<213> 人工序列

<220>

<223>

<400> 2

```

atggcagatc cgggggatg ctgtttcatc accaagatcc tgtgcgccca cgggggcccgt    60
atgacctgga aggaactgct gggtagatc aggctccccg aggcgcagct ctacgagctg    120
ctggagacgg cggggcccga tcgcttcgtg ctattggaga ctggaggcca gcccgggatc    180
actcggctctg tagtggctiac tactcgagcc cgcgtctgcc gtcggaagta ctgccagaga    240
ccctgcgaca gcctgcacct ctgcaagctt aatctgctcg gccggtgcca ctatgcacag    300
tctcagcgga acctctgcaa atattctcac gatgttctct cggaacagaa cttccagatc    360
ctgaagaatc atgagctctc tgggttaac caagaggagc tagcttgctt cctgggtccaa    420
agcgaccttt ttttctgccc cgagatatgc aagagttaca aaggagaggg ccgaaaacag    480
acctgtgggc agccacagcc atgcgagaga ctccacatct gtgagcactt caccgggggc    540
aactgcagtt acctcaactg tctcaggtct cacaacctga tggacagaaa ggtgttgacc    600
atcatgaggg agcacgggct gactcctgat gtggtccaga acatccagga catctgcaac    660
aacaacacag ccaggaggaa cccgcctggc acgagagctg cccatccaca ccgcagagge    720
ggcgcacaca gagacagaag caaaagcaga gaccgcttcc ttcacaacag tctagaatth    780
ctctcacctg ttgtctcacc tctgggatct ggtccgccta gcccagatgt caccagctgt    840
aaagattccc tggaggatgt gtctgtggat gtcaccaga agttcaagta cttggggacg    900
catgaccgtg cgcagctctc cccagctctc tctaaggctg ctggtgttca aggaccagat    960
caaatgagag caagccaaga gttttcagag gatgggaatc tagatgacat attttctagg   1020
aatcgttctg attcatcatc aagtcgagcc tccgctgcca aggtggcaca aagaaatgaa   1080
getgtggcca tgaanaatgg catggaggtc aagggaaga aggaggtctc agacatcgat   1140
cgggtcccat ttttaaatag ttatatgat ggggtgacca tggaaaaagc atcggctctc   1200
ggaattccag gcaaaaagt caccagcaat gatctgaaa atttgctatt acttaacgac   1260
acttggaga atgtggctaa gccccagat ctgcagacca caggcagaat cactgacagt   1320
ggccaagaca aggcattcct gcagaataaa tatggaggaa acccagtggt ggcaagtgca   1380
tccaccata atgccccaaa tggctctagt caaattatgg atgaaactcc taatgtctct   1440
aaaagtagta ccagtggttt tgccataaaa ccagcaattg ctggaggaaa agaagcagtc   1500
tattctggag ttcagagctc gagaagccag gtcttagctg tgcctgggga ggctactacc   1560

```

cctgtacaga gcaacaggct gcctcagtcg cctctgtctt cctcaagcca cagagctgca 1620
 gcctctggga gccctggcaa gaactccacc catacctctg tgagcccagc catcgagtet 1680
 tcaaggatga catcagaccc cgatgagtat ctctacgct acatcctaaa tcctttattt 1740
 aggatggata atcatggccc gaaggaaatc tgcaggacc atctgtacaa gggctgtcaa 1800
 cagagccact ggcacaggag tcaacttccat ctgccctacc ggtggcagat gttegtatat 1860
 accacttga gggacttcca ggacatggag tctategaac aggctattg tgatccccac 1920
 gttgaactca tttgataga aaaccatcag atcaatttc agaaaatgac ctgtgactcc 1980
 taccceatcc gacgcctctc cactcctca tatgaggaaa agccacttag tgctgtcttc 2040
 gccaccaagt ggatttggta ttggaagaat gaatttaatg aatatatcca gtatgggaat 2100
 gagagcccag gccacaccag ctctgacatc aactctgcgt acctggagtc tttcttccag 2160
 tcttgtccca ggggagtttt gccattccag gctggttcac agaagtaga gttaagcttc 2220
 caagggatga ttcagacaaa tatagcttcc aagactcaa ggcatgttgt cagaaggcca 2280
 gtatttgttt ctctgaacga tgtggagcag aagagaagag gtccagaggg cggccgctcg 2340
 agtggiggag gctctggggg aggtacagga ggtggctcag gtgggggcac tgggtggaggc 2400
 agtggagggtg gcgatggccg cgctgctgag ccagtagatc ctagtctaga gccctggaag 2460
 catccaggaa gtcagcctag gactgcttgt acaattgct attgtaaaaa gtgttgcttt 2520
 catgtctacg cgtgtttcac aagaaaaggc ttaggcatct cctatggcag gaagaagcgg 2580
 agacagcgac gaagagctca tcagaacagt cagactcacc aagcttctct atcaaagcag 2640
 ccgcggttcg aacaaaaact catctcagaa gaggatctga atatgcatac cggtcacat 2700
 caccatcacc at 2712

<210> 3

<211> 808

<212> PRT

<213> 人工序列

<220>

<223>

<400> 3

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

His Tyr Ala Gln Ser Gln Arg Asn Leu Cys Lys Tyr Ser His Asp Val
 100 105 110
 Leu Ser Glu Gln Asn Phe Gln Ile Leu Lys Asn His Glu Leu Ser Gly
 115 120 125
 Leu Asn Gln Glu Glu Leu Ala Cys Leu Leu Val Gln Ser Asp Pro Phe
 130 135 140
 Phe Leu Pro Glu Ile Cys Lys Ser Tyr Lys Gly Glu Gly Arg Lys Gln
 145 150 155 160
 Thr Cys Gly Gln Pro Gln Pro Cys Glu Arg Leu His Ile Cys Glu His
 165 170 175
 Phe Thr Arg Gly Asn Cys Ser Tyr Leu Asn Cys Leu Arg Ser His Asn
 180 185 190
 Leu Met Asp Arg Lys Val Leu Thr Ile Met Arg Glu His Gly Leu Ser
 195 200 205
 Pro Asp Val Val Gln Asn Ile Gln Asp Ile Cys Asn Asn Lys His Ala
 210 215 220
 Arg Arg Asn Pro Pro Gly Thr Arg Ala Ala His Pro His Arg Arg Gly
 225 230 235 240
 Gly Ala His Arg Asp Arg Ser Lys Ser Arg Asp Arg Phe Leu His Asn
 245 250 255
 Ser Leu Glu Phe Leu Ser Pro Val Val Ser Pro Leu Gly Ser Gly Pro
 260 265 270

 Pro Ser Pro Asp Val Thr Ser Cys Lys Asp Ser Leu Glu Asp Val Ser
 275 280 285
 Val Asp Val Thr Gln Lys Phe Lys Tyr Leu Gly Thr His Asp Arg Ala
 290 295 300
 Gln Leu Ser Pro Val Ser Ser Lys Ala Ala Gly Val Gln Gly Pro Ser
 305 310 315 320
 Gln Met Arg Ala Ser Gln Glu Phe Ser Glu Asp Gly Asn Leu Asp Asp
 325 330 335
 Ile Phe Ser Arg Asn Arg Ser Asp Ser Ser Ser Ser Arg Ala Ser Ala
 340 345 350
 Ala Lys Val Ala Gln Arg Asn Glu Ala Val Ala Met Lys Met Gly Met
 355 360 365
 Glu Val Lys Gly Lys Lys Glu Ala Pro Asp Ile Asp Arg Val Pro Phe
 370 375 380
 Leu Asn Ser Tyr Ile Asp Gly Val Thr Met Glu Lys Ala Ser Val Ser
 385 390 395 400
 Gly Ile Pro Gly Lys Lys Phe Thr Ala Asn Asp Leu Glu Asn Leu Leu

725 730 735
 Glu Leu Ser Phe Gln Gly Met Ile Gln Thr Asn Ile Ala Ser Lys Thr
 740 745 750
 Gln Arg His Val Val Arg Arg Pro Val Phe Val Ser Ser Asn Asp Val
 755 760 765
 Glu Gln Lys Arg Arg Gly Pro Glu Gly Gly Arg Ser Ser Leu Glu Gly
 770 775 780
 Pro Arg Phe Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met His
 785 790 795 800
 Thr Gly His His His His His His
 805

<210> 4

<211> 2424

<212> DNA

<213> 人工序列

<220>

<223>

<400> 4

```

atggcagatc cgggggatg ctgtttcatc accaagatcc tgtgcgcca cgggggccgt    60
atgaccttgg aggaactgct gggtagatc aggcctcccc aggcgcagct ctacgagctg    120
ctggagacgg cggggcccga tcgttctgtg ctattggaga ctggaggcca ggccggggtc    180
actcggctcg tagtggttac tactcgagcc cgcgtctgcc gtcggaagta ctgccagaga    240
ccctcgcaca gectgcacct ctgcaagett aatctgctcg gccgggtcca ctatgcacag    300
tctcagcgga acctctgcaa atattctcac gatgttctct cggaacagaa cttccagatc    360
ctgaagaatc atgagetctc tgggcttaac caagaggagc tagcttgctt cctggtccaa    420
agcgaccttt ttttctgccc cgagatatgc aagagttaca aaggagaggg ccgaaaacag    480
acctgtgggc agccacagcc atgcgagaga ctccacatct gtgagcactt caccgggggc    540
aactgcagtt acctcaactg tctcaggtct cacaacctga tggacagaaa ggtgttgacc    600
atcatgaggg agcacgggct gaggctgat gtggtccaga acatccagga catctgcaac    660
aacaaacacg ccaggaggaa cccgcctggc acgagagctg cccatccaca ccgcagaggc    720
ggcgcacaca gagacagaag caaaagcaga gaccgcttcc ttcacaacag tctagaattt    780
ctctcacctg ttgtctcacc tctgggatct ggtccgccta gcccagatgt caccagctgt    840
aaagattccc tggaggatgt gtctgtggat gtcacccaga agttcaagta cttggggagc    900
catgaccctg cgcagctctc cccagctcca tctaaggetg ctggtgttca aggaccagct    960
caaatgagag caagccaaga gttttcagag gatgggaatc tagatgacat attttctagg    1020
aatcgttctg attcatcacc aagtcgagcc tccgctgcca agtggtgcaca aagaaatgaa    1080
gctgtggcca tgaanaatgg catggaggtc aagggaaga aggaggctcc agacatcgat    1140
cgggtcccat ttttaaatag ttatattgat ggggtgacca tggaaaaagc atcgggtctca    1200

```

ggaattccag gcaaaaagtt cacagccaat gatctgaaa atttgctatt acttaacgac 1260
 acttgggaaga atgtggctaa gccccaggat ctgcagacca caggcagaat cactgacagt 1320
 ggccaagaca aggcatcct gcagaataaa tatggaggaa acccagtgtg ggcaagtgca 1380
 tccaccata atgccccaaa tgctctagt caaattatgg atgaaactcc taatgtctct 1440
 aaaagtagta ccagtggttt tgccataaaa ccagcaattg ctggaggaaa agaagcagtc 1500
 tattctggag ttcagagtcc gagaagccag gtcctagctg tgccctggga ggctactacc 1560
 cctgtacaga gcaacaggct gctcagctg cctctgtctt cctcaagcca cagagctgca 1620
 gcctctggga gcctggcaa gaactccacc catacctctg tgagcccagc catcgagtct 1680
 tcaaggatga catcagacc cgatgagtat ctcctacgct acatcctaaa tcctttattt 1740
 aggatggata atcatggccc gaaggaaatc tgtcaggacc atctgtacaa gggctgtcaa 1800
 cagagccact gcgacaggag tcacttccat ctgcctacc ggtggcagat gttcgtatat 1860
 accacttga gggacttcca ggacatggag tctatcgaac aggcctattg tgatccccac 1920
 gttgaactca ttttgataga aaaccatcag atcaatttc agaaaatgac ctgtgactcc 1980
 tacccatcc gacgcctctc cactccctca tatgaggaaa agccacttag tgctgtcttc 2040
 gccaccaagt ggatttggta ttggaagaat gaatttaatg aatataatcca gtatgggaat 2100
 gagagcccag gccacaccag ctctgacatc aactctgct acctggagtc tttcttcag 2160
 tcttgtccca ggggagtttt gccattccag gctggttcac agaagtacga gttaagcttc 2220
 caagggatga ttcagacaaa tatagcttcc aagactcaa ggcatgtgt cagaaggcca 2280
 gtatttgttt ctctgaacga tgtggagcag aagagaagag gtccagagg cggccgctcg 2340
 agtctagagg gcccgcggtt cgaacaaaa ctcactcag aagaggatct gaatatgcat 2400
 accggtcatc atcaccatca ccat 2424

<210> 5

<211> 311

<212> PRT

<213> 人工序列

<220>

<223>

<400> 5

Met Ala Asp Pro Gly Val Cys Cys Phe Ile Thr Lys Ile Leu Cys Ala

1 5 10 15

His Gly Gly Arg Met Thr Leu Glu Glu Leu Leu Gly Glu Ile Arg Leu

20 25 30

Pro Glu Ala Gln Leu Tyr Glu Leu Leu Glu Thr Ala Gly Pro Asp Arg

35 40 45

Phe Val Leu Leu Glu Thr Gly Gly Gln Ala Gly Ile Thr Arg Ser Val

50 55 60

Val Ala Thr Thr Arg Ala Arg Val Cys Arg Arg Lys Tyr Cys Gln Arg

<400> 6

```

atggcagatc cgggggtatg ctgtttcatic accaagatcc tgtgcgcccc cggggggccgt      60
atgaccctgg aggaactgct gggtagatc aggctcccc aggcgcagct ctacgagctg      120
ctggagacgg cggggcccga tcgcttcgtg ctattggaga ctggaggcca ggcccgggatc      180
actcggctctg tagtggetac tactcgagcc cgcgtctgcc gtcggaagta ctgccagaga      240
ccctgegaca gectgcacct ctgcaagctt aatctgctcg gccggtgcca ctatgcacag      300
tctcagcggga acctctgcaa atattctcac gatgttctct cggaacagaa cticcagatc      360
ctgaagaatc atgagctctc tgggcttaac caagaggagc tagcttgctt cctgggtccaa      420
agegaccctt ttttctgcc cgagatatgc aagagttaca aaggagaggg ccgaaaacag      480
acctgtgggc agccacagcc atgcgagaga ctccacatct gtgagcactt cacccggggc      540
aactgcagtt acctcaactg tctcaggtct cacaacctga tggacagaaa ggtgttgacc      600
atcatgaggg agcacgggct gagtcctgat gtggtccaga acatccagga catctgcaac      660
aacaaacacg ccaggaggaa cccgcctggc acgagagctg cccatccaca ccgcagaggc      720
ggcgcacaca gagacagaag caaaagcaga gaccgttctt gcggccgctc gagtggtgga      780
ggctctgggg gaggtacagg aggtggctca ggtgggggca ctggtggagg cagtggaggt      840
ggcgatatcg gcgccaaggg cccgcgggtc gaacaaaaac tcattctcaga agaggatctg      900
aatatgcata ccggtcatca tcaccatcac cat      933

```

<210> 7

<211> 834

<212> PRT

<213> 人工序列

<220>

<223>

<400> 7

```

Met Ala Asp Pro Gly Val Cys Cys Phe Ile Thr Lys Ile Leu Cys Ala
1           5           10           15
His Gly Gly Arg Met Thr Leu Glu Glu Leu Leu Gly Glu Ile Arg Leu
           20           25           30
Pro Glu Ala Gln Leu Tyr Glu Leu Leu Glu Thr Ala Gly Pro Asp Arg
           35           40           45
Phe Val Leu Leu Glu Thr Gly Gly Gln Ala Gly Ile Thr Arg Ser Val
           50           55           60
Val Ala Thr Thr Arg Ala Arg Val Cys Arg Arg Lys Tyr Cys Gln Arg
65           70           75           80
Pro Cys Asp Ser Leu His Leu Cys Lys Leu Asn Leu Leu Gly Arg Cys
           85           90           95
His Tyr Ala Gln Ser Gln Arg Asn Leu Cys Lys Tyr Ser His Asp Val
           100          105          110

```

Leu Ser Glu Gln Asn Phe Gln Ile Leu Lys Asn His Glu Leu Ser Gly
 115 120 125
 Leu Asn Gln Glu Glu Leu Ala Cys Leu Leu Val Gln Ser Asp Pro Phe
 130 135 140
 Phe Leu Pro Glu Ile Cys Lys Ser Tyr Lys Gly Glu Gly Arg Lys Gln
 145 150 155 160
 Thr Cys Gly Gln Pro Gln Pro Cys Glu Arg Leu His Ile Cys Glu His
 165 170 175
 Phe Thr Arg Gly Asn Cys Ser Tyr Leu Asn Cys Leu Arg Ser His Asn
 180 185 190

 Leu Met Asp Arg Lys Val Leu Thr Ile Met Arg Glu His Gly Leu Ser
 195 200 205
 Pro Asp Val Val Gln Asn Ile Gln Asp Ile Cys Asn Asn Lys His Ala
 210 215 220
 Arg Arg Asn Pro Pro Gly Thr Arg Ala Ala His Pro His Arg Arg Gly
 225 230 235 240
 Gly Ala His Arg Asp Arg Ser Lys Ser Arg Asp Arg Phe Leu His Asn
 245 250 255
 Ser Leu Glu Phe Leu Ser Pro Val Val Ser Pro Leu Gly Ser Gly Pro
 260 265 270
 Pro Ser Pro Asp Val Thr Ser Cys Lys Asp Ser Leu Glu Asp Val Ser
 275 280 285
 Val Asp Val Thr Gln Lys Phe Lys Tyr Leu Gly Thr His Asp Arg Ala
 290 295 300
 Gln Leu Ser Pro Val Ser Ser Lys Ala Ala Gly Val Gln Gly Pro Ser
 305 310 315 320
 Gln Met Arg Ala Ser Gln Glu Phe Ser Glu Asp Gly Asn Leu Asp Asp
 325 330 335
 Ile Phe Ser Arg Asn Arg Ser Asp Ser Ser Ser Ser Arg Ala Ser Ala
 340 345 350
 Ala Lys Val Ala Gln Arg Asn Glu Ala Val Ala Met Lys Met Gly Met
 355 360 365
 Glu Val Lys Gly Lys Lys Glu Ala Pro Asp Ile Asp Arg Val Pro Phe
 370 375 380
 Leu Asn Ser Tyr Ile Asp Gly Val Thr Met Glu Lys Ala Ser Val Ser
 385 390 395 400
 Gly Ile Pro Gly Lys Lys Phe Thr Ala Asn Asp Leu Glu Asn Leu Leu
 405 410 415
 Leu Leu Asn Asp Thr Trp Lys Asn Val Ala Lys Pro Gln Asp Leu Gln

420 425 430
 Thr Thr Gly Arg Ile Thr Asp Ser Gly Gln Asp Lys Ala Phe Leu Gln
 435 440 445
 Asn Lys Tyr Gly Gly Asn Pro Val Trp Ala Ser Ala Ser Thr His Asn
 450 455 460
 Ala Pro Asn Gly Ser Ser Gln Ile Met Asp Glu Thr Pro Asn Val Ser
 465 470 475 480
 Lys Ser Ser Thr Ser Gly Phe Ala Ile Lys Pro Ala Ile Ala Gly Gly
 485 490 495
 Lys Glu Ala Val Tyr Ser Gly Val Gln Ser Pro Arg Ser Gln Val Leu
 500 505 510
 Ala Val Pro Gly Glu Ala Thr Thr Pro Val Gln Ser Asn Arg Leu Pro
 515 520 525
 Gln Ser Pro Leu Ser Ser Ser Ser His Arg Ala Ala Ala Ser Gly Ser
 530 535 540
 Pro Gly Lys Asn Ser Thr His Thr Ser Val Ser Pro Ala Ile Glu Ser
 545 550 555 560
 Ser Arg Met Thr Ser Asp Pro Asp Glu Tyr Leu Leu Arg Tyr Ile Leu
 565 570 575
 Asn Pro Leu Phe Arg Met Asp Asn His Gly Pro Lys Glu Ile Cys Gln
 580 585 590
 Asp His Leu Tyr Lys Gly Cys Gln Gln Ser His Cys Asp Arg Ser His
 595 600 605
 Phe His Leu Pro Tyr Arg Trp Gln Met Phe Val Tyr Thr Thr Trp Arg
 610 615 620
 Asp Phe Gln Asp Met Glu Ser Ile Glu Gln Ala Tyr Cys Asp Pro His
 625 630 635 640
 Val Glu Leu Ile Leu Ile Glu Asn His Gln Ile Asn Phe Gln Lys Met
 645 650 655
 Thr Cys Asp Ser Tyr Pro Ile Arg Arg Leu Ser Thr Pro Ser Tyr Glu
 660 665 670
 Glu Lys Pro Leu Ser Ala Val Phe Ala Thr Lys Trp Ile Trp Tyr Trp
 675 680 685
 Lys Asn Glu Phe Asn Glu Tyr Ile Gln Tyr Gly Asn Glu Ser Pro Gly
 690 695 700
 His Thr Ser Ser Asp Ile Asn Ser Ala Tyr Leu Glu Ser Phe Phe Gln
 705 710 715 720
 Ser Cys Pro Arg Gly Val Leu Pro Phe Gln Ala Gly Ser Gln Lys Tyr
 725 730 735
 Glu Leu Ser Phe Gln Gly Met Ile Gln Thr Asn Ile Ala Ser Lys Thr

cggggccat ttttaaatag ttatattgat ggggtgacca tggaaaaagc atcggtctca 1200
 ggaattccag gcaaaaagtt cacagccaat gatctggaaa atttgctatt acttaacgac 1260
 acttgaaga atgtggctaa gccccaggat ctgcagacca caggcagaat cactgacagt 1320
 ggccaagaca aggcattcct gcagaataaa tatggaggaa acccagtggtg ggcaagtgca 1380
 tccaccata atgccccaaa tggctctagt caaattatgg atgaaactcc taatgtctct 1440
 aaaagtagta ccagtggttt tgccataaaa ccagcaattg ctggaggaaa agaagcagtc 1500
 tattctggag ttcagagtcc gagaagccag gtcctagctg tgccctggga ggctactacc 1560
 cctgtacaga gcaacaggct gcctcagctg cctctgtctt cctcaagcca cagagctgca 1620
 gcctctggga gccctggcaa gaactccacc catacctctg tgagcccagc catcgagtct 1680
 tcaaggatga catcagacce cgatgagtat ctctacgct acatcctaaa tccittattt 1740
 aggatggata atcatgccc gaaggaaatc tgtcaggacc atctgtacaa gggctgtcaa 1800
 cagagccact ggcagaggag tcacttccat ctgccctacc ggtggcagat gttcgtatat 1860
 accacttggg gggacttcca ggacatggag tctatcgaac aggcctattg tgatccccac 1920
 gtgaaactea tttgataga aaaccatcag atcaatttcc agaaaatgac ctgtgactcc 1980
 taccceatcc gacgcctctc cactccctca tatgaggaaa agccacttag tgctgtcttc 2040
 gccaccaagt ggatttggtg ttggaagaat gaatttaag aatatatcca gtatgggaat 2100
 gagagcccag gccacaccag ctctgacatc aactctgcgt acctggagtc tttcttcag 2160
 tcttgtccca ggggagtttt gccattccag gctgggtcac agaagtacga gttaagcttc 2220
 caagggatga ttcagacaaa tatagcttcc aagactcaaa ggcatgttgt cagaaggcca 2280
 gtatttgttt ctctgaacga tgtggagcag aagagaagag gtccagaggg cggccgctcg 2340
 agtggaggag gctctggggg aggtacagga ggtggctcag gtgggggcac tggaggagge 2400
 agtggagggt gcgatatcgg cgccaagggc ccgcggttcg aacaaaaact catctcagaa 2460
 gaggatctga atatgatac cggatcatc caccatcacc at 2502

<210> 9

<211> 381

<212> PRT

<213> 人工序列

<220>

<223>

<400> 9

Met Ala Asp Pro Gly Val Cys Cys Phe Ile Thr Lys Ile Leu Cys Ala

1 5 10 15

His Gly Gly Arg Met Thr Leu Glu Glu Leu Leu Gly Glu Ile Arg Leu

20 25 30

Pro Glu Ala Gln Leu Tyr Glu Leu Leu Glu Thr Ala Gly Pro Asp Arg

	355		360		365								
	Asp	Leu	Asn	Met	His	Thr	Gly	His	His	His	His	His	His
	370		375		380								

<210> 10

<211> 1143

<212> DNA

<213> 人工序列

<220>

<223>

<400> 10

atggcagatc	cegggggatg	ctgtttcatc	accaagatcc	tgtgcgccca	cgggggccgt	60
atgaccctgg	aggaactgct	gggtgagatc	aggctccccg	aggcgcagct	ctacgagctg	120
ctggagacgg	cggggcccca	tcgcttcgtg	ctattggaga	ctggaggcca	ggccgggatc	180
actcggcttg	tagtggctac	tactcgagcc	cgcgtctgcc	gtcggaaagta	ctgccagaga	240
ccctgcgaca	gectgcacct	ctgcaagctt	aatctgctcg	gccgggtcca	ctatgcacag	300
tctcagcggg	acctctgcaa	atattctcac	gatgttctct	cggaacagaa	cttccagatc	360
ctgaagaatc	atgagctctc	tggccttaac	caagaggagc	tagcttgcc	cctgggtccaa	420
agcgaccctt	ttttctgcc	cgagatatgc	aagagttaca	aaggagaggg	ccgaaaacag	480
acctgtgggc	agccacagcc	atgcgagaga	ctccacatct	gtgagcactt	caccgggggc	540
aactgcagtt	acctcaactg	tctcaggctc	cacaacctga	tggacagaaa	ggtgttgacc	600
atcatgaggg	agcacgggct	gagtcctgat	gtggtccaga	acatccagga	catctgcaac	660
aacaaacacg	ccaggaggaa	cccgcctggc	acgagagctg	cccatccaca	ccgcagaggc	720
ggcgcacaca	gagacagaag	caaaagcaga	gaccgttctc	gcggccgctc	gagtgggtga	780
ggctctgggg	gaggtacagg	agggtgctca	ggtgggggca	ctggtggagg	cagtggaggt	840
ggcgatggcc	gcgctgctga	gccagtagat	cctagtctag	agccctggaa	gcatccagga	900
agtcagccta	ggactgcttg	taacaattgc	tattgtaaaa	agtgttgctt	tcattgctac	960
gcgtgtttca	caagaaaagg	cttaggcata	tcctatggca	ggaagaagcg	gagacagcga	1020
cgaagagctc	ctcaggacag	tcagactcat	caagcttctc	tatcaaagca	gccgcggttc	1080
gaacaaaaac	tcattctcaga	agaggatctg	aatatgcata	ccggtcatca	tcaccatcac	1140
cat						1143

<210> 11

<211> 194

<212> PRT

<213> 人工序列

<220>

<223>

<400> 11

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

<210> 12

<211> 582

<212> DNA

<213> 人工序列

<220>

<223>

<400> 12

atggcagatc ccggggtatg ctgtttcadc accaagatcc tgtgcgccca cgggggccgt

60

```

atgacctgg aggaactgct gggtagatc aggtcctcccg aggcgcagct ctacgagctg    120
ctggagacgg cggggcccga tcgcttcgtg ctattggaga ctggaggcca ggccgggatc    180
actcggctcg tagtggctac tactcgatcg agtggtagag gctctggggg aggtacagga    240
ggtaggctcag gtggggggcac tggtaggagc agtggagggtg gcgatggccg cgctgctgag    300
ccagtagatc ctagtctaga gccctggaag catccaggaa gtcagcctag gactgcttgt    360
aacaattgct attgtaaaaa gtgttgcttt cattgctacg cgtgtttcac aagaaaaggc    420
ttaggcatct cctatggcag gaagaagcgg agacagcgac gaagagctcc tcaggacagt    480
cagactcadc aagcttctct atcaaagcag ccgcggttcg aacaaaaact catctcagaa    540
gaggatctga atatgcatac cggatcatcat caccatcacc at                    582

```

<210> 13

<211> 285

<212> PRT

<213> 人工序列

<220>

<223>

<400> 13

```

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

```

	165	170	175
Phe Thr Arg Gly Asn Cys Ser Tyr Leu Asn Cys Leu Arg Ser His Asn			
	180	185	190
Leu Met Asp Arg Lys Val Leu Thr Ile Met Arg Glu His Gly Leu Ser			
	195	200	205
Pro Asp Val Val Gln Asn Ile Gln Asp Ile Cys Asn Asn Lys His Ala			
	210	215	220
Arg Arg Asn Pro Pro Gly Thr Arg Ala Ala His Pro His Arg Arg Gly			
225	230	235	240
Gly Ala His Arg Asp Arg Ser Lys Ser Arg Asp Arg Ser Cys Gly Arg			
	245	250	255
Ser Ser Leu Glu Gly Pro Arg Phe Glu Gln Lys Leu Ile Ser Glu Glu			
	260	265	270
Asp Leu Asn Met His Thr Gly His His His His His His His			
	275	280	285

<210> 14

<211> 855

<212> DNA

<213> 人工序列

<220>

<223>

<400> 14

```

atggcagatc cgggggatg ctgtttcatc accaagatcc tgtgcgcccc cgggggcccgt    60
atgaccctgg aggaactgct gggtagatc aggcctcccc aggcgcagct ctacgagctg    120
ctggagacgg cggggcccga tcgcttcgtg ctattggaga ctggaggcca ggcccgggatc    180
actcggctct tagtggctac tactcgagcc cgcgtctgcc gtcggaagta ctgccagaga    240
ccctgcgaca gcctgcacct ctgcaagctt aatctgctcg gccggtgcca ctatgcacag    300
tctcagcgga acctctgcaa atattctcac gatgttctct cggaacagaa cttccagatc    360
ctgaagaatc atgagctctc tgggcttaac caagaggagc tagcttgctt cctggtccaa    420
agcgacctt ttttctgcc cgagatatgc aagagtaca aaggagaggg ccgaaaacag    480
acctgtgggc agccacagcc atgcgagaga ctccacatct gtgagcactt cacccggggc    540
aactgcagtt acctcaactg tctcaggctt cacaacctga tggacagaaa ggtgttgacc    600
atcatgaggg agcacgggct gaticctgat gtggctcaga acatccagga catctgcaac    660
aacaaacacg ccaggaggaa cccgcctggc acgagagctg cccatccaca ccgagaggc    720
ggcgcacaca gagacagaag caaaagcaga gaccgttctt gggcggctc gactctagag    780
ggcccgggt tcgaacaaaa actcatctca gaagaggatc tgaatatgca taccggtcat    840
catcaccatc accat

```

855

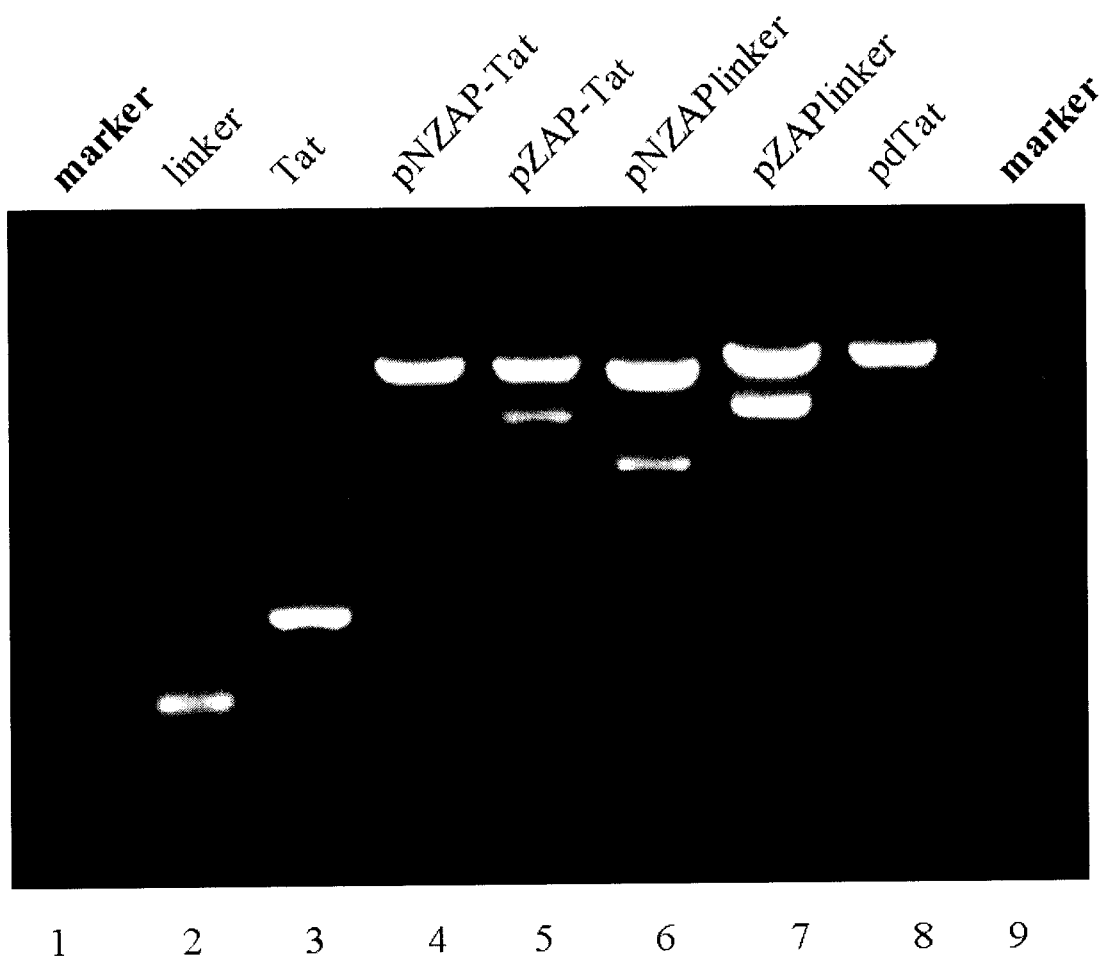


图 1

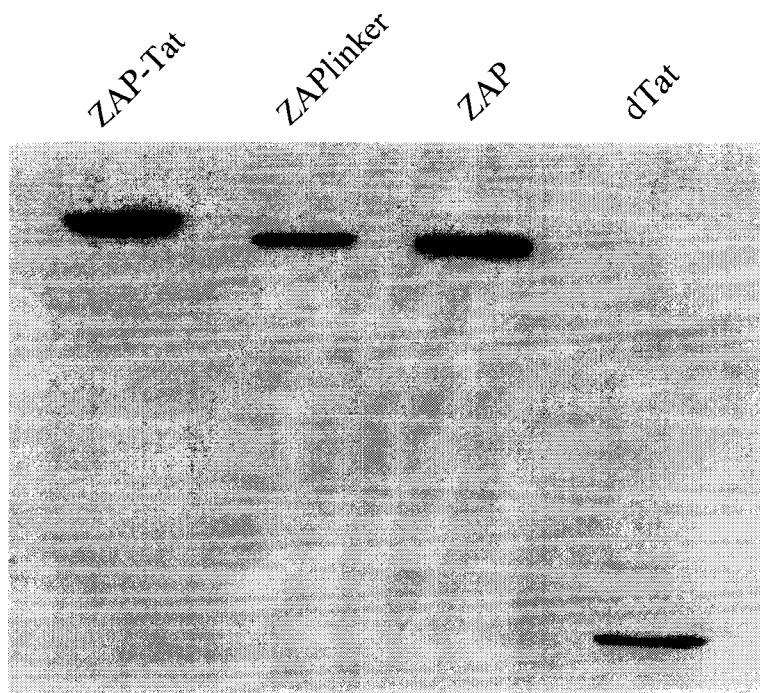


图 2

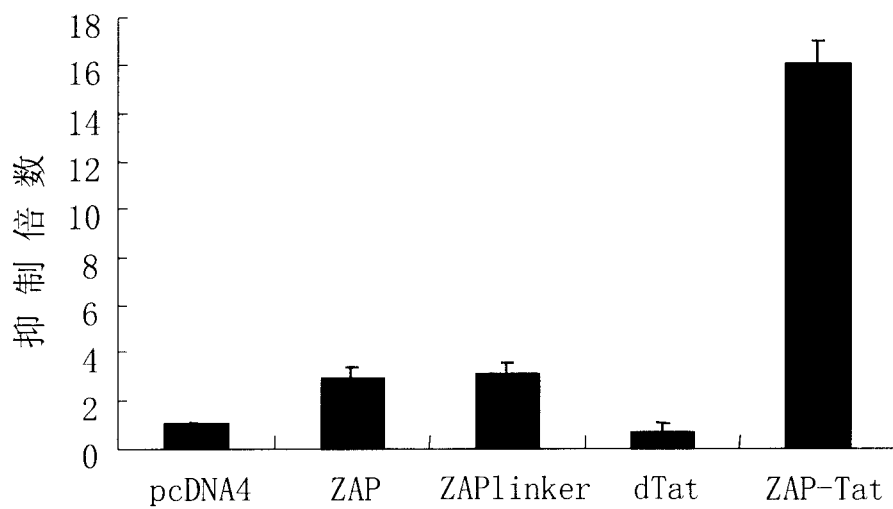


图 3