

(19) 中华人民共和国国家知识产权局



(12) 发明专利申请

(10) 申请公布号 CN 105018441 A

(43) 申请公布日 2015. 11. 04

(21) 申请号 201410162712. 3

(22) 申请日 2014. 04. 22

(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 15 号

(72) 发明人 刘迎芳

(74) 专利代理机构 北京纪凯知识产权代理有限公司

11245

代理人 关畅

(51) Int. Cl.

C12N 9/12(2006. 01)

C12N 15/54(2006. 01)

C12N 15/866(2006. 01)

权利要求书2页 说明书14页

序列表12页 附图4页

(54) 发明名称

流感病毒 RNA 聚合酶结晶的方法

(57) 摘要

本发明公开了一种流感病毒 RNA 聚合酶结晶的方法。本发明提供的流感病毒 RNA 聚合酶结晶的方法,为先将流感病毒 RNA 聚合酶的 PB1 亚基编码基因、PA 亚基编码基因和 PB2 亚基截短体的编码基因在昆虫细胞中共表达,得到 RNA 聚合酶,再进行纯化、结晶,得到流感病毒 RNA 聚合酶晶体。本发明的实验证明,本发明利用了不同长度的截短保留氨基端的 PB2 蛋白片段与其它两个亚基全长蛋白进行共表达纯化及复合体组装,这一截短后的小复合体可以在多种截短情况下、使用不同毒株获得晶体。本方法不仅为流感病毒聚合酶的结构解析奠定基础,同时,对抗流感病毒药物筛选、广谱性疫苗和药物的设计研发进程也将具有重要意义。

1. 一种流感病毒 RNA 聚合酶复合体表达、纯化及结晶的方法,为先将流感病毒 RNA 聚合酶的 PB1 亚基编码基因、PA 亚基编码基因和 PB2 亚基截短体的编码基因在离体的昆虫细胞中共表达、纯化,得到 RNA 聚合酶,再将所述 RNA 聚合酶与 RNA 结合形成复合体,再将所述复合体结晶,得到流感病毒 RNA 聚合酶晶体;

所述 PB2 亚基截短体由 PB2 亚基自 N 末端起第 1 至 N 位氨基酸残基组成;N = 37-280 中的任一自然数。

2. 根据权利要求 1 所述的方法,其特征在于:

所述 PB2 亚基截短体由 PB2 亚基自 N 末端起第 1 至 N 位氨基酸残基组成;N = 37-130 中的任一自然数。

3. 根据权利要求 2 所述的方法,其特征在于:

所述 PB2 亚基截短体为如下 1) 或 2) 或 3):

1) 为 PB2 亚基氨基酸序列自 N 末端起前 37 位的氨基酸残基形成的蛋白;

2) 为 PB2 亚基氨基酸序列自 N 末端起前 130 位的氨基酸残基形成的蛋白;

3) 为 PB2 亚基氨基酸序列自 N 末端起前 103 位的氨基酸残基形成的蛋白。

4. 根据权利要求 1-3 中任一所述的方法,其特征在于:

所述昆虫细胞为 High Five 细胞;

所述流感病毒为 A 型流感病毒,所述 A 型流感病毒的病毒株具体为 H5N1、H1N1 或 H3N2。

5. 根据权利要求 1-4 中任一所述的方法,其特征在于:

所述将流感病毒 RNA 聚合酶的 PB1 亚基编码基因、PA 亚基编码基因和 PB2 亚基截短体的编码基因在昆虫细胞中共表达采用 Bac-to-Bac 杆状病毒表达系统,所述 Bac-to-Bac 杆状病毒表达系统中的供体质粒为 pFastBac,大肠杆菌为 DH10Bac,昆虫细胞为 Sf9。

6. 根据权利要求 1-5 中任一所述的方法,其特征在于:

所述 PB1 亚基的氨基酸序列为序列表中序列 1 所示;

所述 PA 亚基的氨基酸序列为序列表中序列 3 所示;

所述 PB2 亚基的氨基酸序列为序列表中序列 2 所示;

1) 所示的所述 PB2 亚基截短体的氨基酸序列为序列表中序列 2 自 N 末端第 1-37 位;

2) 所示的所述 PB2 亚基截短体的氨基酸序列为序列表中序列 2 自 N 末端第 1-130 位;

3) 所示的所述 PB2 亚基截短体的氨基酸序列为序列表中序列 2 自 N 末端第 1-103 位。

7. 根据权利要求 1-6 中任一所述的方法,其特征在于:

所述 PB1 亚基的编码基因的核苷酸序列为序列表中序列 4 所示;

所述 PA 亚基的编码基因的核苷酸序列为序列表中序列 6 所示;

1) 所示的所述 PB2 亚基截短体的编码基因的核苷酸序列为序列表中序列 5 自 5' 末端第 1-111 位;

2) 所示的所述 PB2 亚基截短体的编码基因的核苷酸序列为序列表中序列 5 自 5' 末端第 1-390 位;

3) 所示的所述 PA 亚基截短体的编码基因的核苷酸序列为序列表中序列 5 自 5' 末端第 1-309 位。

8. 根据权利要求 1-7 中任一所述的方法,其特征在于:

所述纯化为将共表达得到的所述 RNA 聚合酶依次经亲和层析、离子交换层析和凝胶层

析,得到纯化后的 RNA 聚合酶;

所述结晶采用的缓冲液由终浓度为 2% -30% 的 PEG1000-20000、终浓度为 50mM-500mM NaCl 和浓度为 100mM、pH 值为 8.5 的 TrisHCl 缓冲液组成;

所述结晶采用的温度为 4℃ -20℃。

9. 由权利要求 1-8 任一所述的方法制备的 RNA 聚合酶晶体。

10. 权利要求 9 所述的 RNA 聚合酶晶体在抗流感病毒药物筛选或疫苗设计和制备中的应用。

流感病毒 RNA 聚合酶结晶的方法

技术领域

[0001] 本发明涉及生物技术领域,涉及流感病毒 RNA 聚合酶结晶的方法,尤其涉及一种流感病毒 RNA 聚合酶复合体的表达、纯化及结晶方法。

背景技术

[0002] 流感病毒,尤其是甲型流感病毒一直是威胁人类社会的一个大问题,它具有广泛的宿主范围包括人、猪、马和禽类等,可以造成人类及动物上呼吸道感染,即流行性感,由于其可以借空气迅速的传播,在上个世纪造成数次世界性的大流行。流感病毒属于正黏病毒科,为负链单链 RNA 分节段基因组病毒,其基因组分为 8 段,目前发现至少可编码 16 个蛋白。虽然目前存在不少抗流感病毒药物和疫苗,但是这些药物和疫苗主要针对流感病毒表面蛋白,如 HA, NA 和 M2。由于流感病毒在药物压力下可以通过突变或者基因重组改变其表面抗原,这些表面蛋白经过突变的毒株将具有抗药性,从而可能引起无法控制的流感流行。为了解决药物和疫苗的抗药性问题,人们在考虑改变思路来研发广谱性抗流感药物。

[0003] 随着对流感病毒复制机制的研究,人们逐渐将目光集中在病毒的核糖核蛋白 RNP 上。流感病毒 RNP 是流感病毒复制的核心组分。与病毒表面蛋白不同,流感病毒的 RNP 更为保守,它由病毒基因组、RNA 聚合酶和缠绕在病毒基因组 RNA 上的多个核蛋白组成,负责病毒的转录和复制。流感病毒聚合酶是流感病毒复制的核心机器,它是由 PA、PB1 和 PB2 三个亚基组成的大小约为 250 千道尔顿 (KDa) 的三元复合物,可以产生三种类型的 RNA,分别是 cRNA、vRNA 和 mRNA,该聚合酶在病毒的生命活动中既参与病毒基因组的转录复制过程,同时有文章报道该复合物也参与病毒的组装。

[0004] 尽管人们通过生物化学,遗传学,生物信息学和结构生物学对流感病毒聚合酶进行了各方面的深入研究,提出了越来越精细的模型,关于聚合酶的工作机制方面仍然有很多关键问题尚无法解决,例如:聚合酶如何调控转录和复制过程;在复制过程中,两个聚合酶分子如何协调以启动和终止复制等;此外,聚合酶还参与病毒的组装和抵御宿主免疫系统。为了解决这些问题,十分需要一个原子分辨率的聚合酶整体结构,然而由于流感病毒聚合酶是一个 250KD 的复合物,为了以行使其多种功能,可能具有多种构象,这给结构生物学研究、尤其是晶体结构的获得造成了极大的障碍,因此即使经过了数十年的努力,仍没有流感病毒聚合酶整体复合物的结构得到报道。

[0005] 目前,人们在重组蛋白表达纯化,晶体优化和结构解析方面积累了大量的经验,通常情况下,纯化得到足够量的稳定均一的蛋白质、并由此获得结晶仍然是大分子晶体结构分析的主要瓶颈。

发明内容

[0006] 本发明的目的是提供一种流感病毒 RNA 聚合酶结晶的方法。

[0007] 本发明提供的方法,包括如下步骤:为先将流感病毒 RNA 聚合酶的 PB1 亚基编码基因、PA 亚基编码基因和 PB2 亚基截短体的编码基因在昆虫细胞中共表达、纯化,得到 RNA 聚

合酶,再将所述 RNA 聚合酶与 RNA 结合形成复合体,再将所述复合体结晶,得到流感病毒 RNA 聚合酶晶体;

[0008] 所述 PB2 亚基截短体由 PB2 亚基自 N 末端起第 1 至 N 位氨基酸残基组成;N = 37-280 中的任一自然数。

[0009] 上述方法中,所述 PB2 亚基截短体由 PB2 亚基自 N 末端起第 1 至 N 位氨基酸残基组成;N = 37-130 中的任一自然数。

[0010] 上述方法中,所述 PB2 亚基截短体为如下 1) 或 2) 或 3) :

[0011] 1) 为 PB2 亚基氨基酸序列自 N 末端起前 37 位的氨基酸残基形成的蛋白;

[0012] 2) 为 PB2 亚基氨基酸序列自 N 末端起前 130 位的氨基酸残基形成的蛋白;

[0013] 3) 为 PB2 亚基氨基酸序列自 N 末端起前 103 位的氨基酸残基形成的蛋白。

[0014] 上述方法中,所述昆虫细胞为 High Five 细胞;

[0015] 所述流感病毒为 A 型流感病毒,所述 A 型流感病毒的病毒株具体为 H5N1、H1N1 或 H3N2。

[0016] 上述方法中,所述将流感病毒 RNA 聚合酶的 PB1 亚基编码基因、PA 亚基编码基因和 PB2 亚基截短体的编码基因在昆虫细胞中共表达采用 Bac-to-BacBac-to-Bac 杆状病毒表达系统,所述 Bac-to-BacBac-to-Bac 杆状病毒表达系统中的供体质粒为 pFastBac,大肠杆菌为 DH10Bac,昆虫细胞为 Sf9。

[0017] 上述方法中,所述 PB1 亚基的氨基酸序列为序列表中序列 1 所示;

[0018] 所述 PA 亚基的氨基酸序列为序列表中序列 3 所示;

[0019] 所述 PB2 亚基的氨基酸序列为序列表中序列 2 所示;

[0020] 1) 所示的所述 PB2 亚基截短体的氨基酸序列为序列表中序列 2 自 N 末端第 1-37 位;

[0021] 2) 所示的所述 PB2 亚基截短体的氨基酸序列为序列表中序列 2 自 N 末端第 1-130 位;

[0022] 3) 所示的所述 PB2 亚基截短体的氨基酸序列为序列表中序列 2 自 N 末端第 1-103 位。

[0023] 上述方法中,所述 PB1 亚基的编码基因的核苷酸序列为序列表中序列 4 所示;

[0024] 所述 PA 亚基的编码基因的核苷酸序列为序列表中序列 6 所示;

[0025] 1) 所示的所述 PB2 亚基截短体的编码基因的核苷酸序列为序列表中序列 5 自 5' 末端第 1-111 位;

[0026] 2) 所示的所述 PB2 亚基截短体的编码基因的核苷酸序列为序列表中序列 5 自 5' 末端第 1-390 位;

[0027] 3) 所示的所述 PA 亚基截短体的编码基因的核苷酸序列为序列表中序列 5 自 5' 末端第 1-309 位。

[0028] 上述方法中,所述纯化为将共表达得到的所述 RNA 聚合酶依次经亲和层析、离子交换层析和凝胶层析,得到纯化后的 RNA 聚合酶;

[0029] 所述结晶采用的缓冲液由终浓度为 2% -30% 的 PEG1000-20000、终浓度为 50mM-500mM NaCl 和浓度为 100mM、pH 值为 8.5 的 TrisHCl 缓冲液组成;

[0030] 所述结晶采用的温度为 4℃ -20℃。

[0031] 由上述的方法制备的 RNA 聚合酶晶体也是本发明保护的范围。

[0032] 上述的 RNA 聚合酶晶体在抗流感病毒药物筛选或疫苗设计和制备中的应用也是本发明保护的范围。

[0033] 本发明的实验证明,本发明利用了不同长度保留氨基端的 PB2 蛋白截短体与流感病毒 RNA 聚合酶复合体其它两个亚基 PA 和 PB1 全长蛋白进行共表达、纯化及复合体组装获得的不同的小复合体,这些小复合体均可在加入 vRNA 后结晶形成流感病毒 RNA 聚合酶晶体。因此推断,不论使用何种流感病毒毒株,PB2 亚基氨基端的这一区间内任何截短体都可能帮助获得流感病毒 RNA 聚合酶复合体结晶。

[0034] 本发明的方法解决了长久以来的一个重要问题,即如何有效表达纯化一个含有大部分流感病毒 RNA 聚合酶复合体成分的复合体,特别是含有完整 PB1 结构的该蛋白复合物,并使之结晶,这一方法不仅为流感病毒聚合酶的结构解析奠定基础,同时,对抗流感病毒药物筛选、广谱性疫苗和药物的设计研发进程也将具有重要意义和应用价值。

附图说明

[0035] 图 1 为含有 PB2-37 截短体的流感病毒 RNA 聚合酶复合体纯化过程中图谱

[0036] 图 2 为含有 PB2-37 截短体的流感病毒 RNA 聚合酶复合体纯化 SDS-PAGE 电泳图谱

[0037] 图 3 为含有 PB2-130 截短体的流感病毒 RNA 聚合酶复合体分子筛纯化图谱

[0038] 图 4 为含有 PB2-130 截短体的流感病毒 RNA 聚合酶复合体的纯化结果

[0039] 其中,图 4A 为含有 PB2-130 截短体的流感病毒 RNA 聚合酶复合体电泳图,1-2 为镍柱层析后电泳、3 为离子交换后电泳、4-5 为分子筛层析后电泳

[0040] 图 4B 为离子交换图谱,蓝色曲线(上面的曲线)为样品在 UV280 的吸收情况;红色曲线(下面的曲线)代表 UV260 吸收;黑色曲线代表离子交换盐浓度的增加过程。

[0041] 图 4C 为分子筛纯化图谱,其中黑色曲线代表不加 RNA 时 UV280 和 UV260 吸收情况,蓝色曲线代表加入 RNA 后复合体洗脱 UV280 和 UV260 吸收情况。

[0042] 图 5 为含有不同 PB2 截短体的流感病毒 RNA 聚合酶复合体纯化后的电泳图谱

[0043] 图 6 为含有 PB2-130 截短体的流感病毒 RNA 聚合酶复合体电镜结果及结晶图片

[0044] 其中 A 和 B 为二体及四体的电镜结构观察结果、C 为四体加入 vRNA 后的结晶照片、D 为结晶电泳检测图

[0045] 图 7 为含有不同 PB2 截短体的流感病毒 RNA 聚合酶复合体的结晶图片

具体实施方式

[0046] 下述实施例中所使用的实验方法如无特殊说明,均为常规方法。

[0047] 下述实施例中所用的材料、试剂等,如无特殊说明,均可从商业途径得到。

[0048] 实验所用昆虫表达载体 pFASTBAC 和 pFastBac-Dual 及细胞培养基 SF900II SFM 均购自 Invitrogen 公司。所需 RNA 均由 Takara 公司合成,镍亲和柱购自 Qiagen 公司(Ni-NTA Agarose,产品目录号:30230),SP 阳离子交换柱(HITRAP SP HP,产品目录号:17-1152-01)和 Superdex20016/60 凝胶层析柱(HiLoad16/600Superdex200pg,产品目录号:28-9893-35)购自 GE 公司。检测 PA、PB1 和 PB2 所用抗体通过表达相应蛋白多肽,免疫兔获得多克隆抗体,经过实验验证特异性良好。

细胞系,72 小时分别收集上清液,得到表达 PB2-37 的重组昆虫杆状病毒、表达 PB1 的重组昆虫杆状病毒、表达 PA 的重组昆虫杆状病毒,由此获得第一代 (P1) 各亚基表达的病毒。

[0090] 将上述产生的表达 PB2-37 的重组昆虫杆状病毒、表达 PB1 的重组昆虫杆状病毒、表达 PA 的重组昆虫杆状病毒分别通过两次转染昆虫细胞系 Sf9 细胞以进行病毒扩增,得到扩增的 P3 代表达 PB2-37 的重组昆虫杆状病毒、P3 代表达 PB1 的重组昆虫杆状病毒、P3 代表达 PA 的重组昆虫杆状病毒。

[0091] 检测滴度,P3 代病毒滴度可以达到 10^7 pfu/mL 以上,可以用于蛋白质表达纯化。若病毒滴度较低,可用 P3 代病毒再次感染 Sf9 细胞系,得到滴度较高的 P4 代杆状病毒用于蛋白质表达。

[0092] 3)、流感病毒 RNA 聚合酶的表达

[0093] 将上述 2) 得到 P3 代表达 PB2-37 的重组昆虫杆状病毒、P3 代表达 PB1 的重组昆虫杆状病毒、P3 代表达 PA 的重组昆虫杆状病毒等滴度量 (滴度均为 10^8 pfu/mL) 混合后按 10^8 virus particles/Cell (MOI = 10 ; MOI: multiplicity of infection) 转染昆虫细胞 High Five, 48 小时后收集培养液, $500 \times g$ 4°C 离心 10 分钟收集细胞。将上述收获的细胞按照 1:10 (质量: 体积) 溶解于裂解缓冲液中 (25mM Tris, 250mM NaCl, pH8.0), 使用超声破碎细胞, 具体在 0°C 冰液中保温放置装有细胞悬液的容器, 在 30w 功率 /50mL 细胞溶液比率下, 按照超声 3 秒停 6 秒方式总计超声破碎 30 分钟的方式充分裂解, 之后, $40,000 \times g$ 4°C 离心 30 分钟, 收集细胞裂解液上清液, 即为含有 PB2-37 截短体流感病毒 RNA 聚合酶的上清液。可以反复离心 2 次以便尽可能除去杂质。

[0094] 上述裂解缓冲液按照如下方法制备: 将如下溶质按照如下终浓度溶解在水中得到的溶液: 250mM NaCl、25mM Tris, 调节裂解缓冲液的 pH 值为 8.0。

[0095] 2、流感病毒 RNA 聚合酶的镍柱亲和层析纯化

[0096] 将上述 1 得到的含有 PB2-37 截短体流感病毒 RNA 聚合酶的上清液吸取到预冷的烧杯中, 加入适量用裂解缓冲液清洗过的镍柱磁珠 (Nickel Beads), 4°C 结合 1 小时。 $500 \times g$ 4°C 离心 3 分钟收集 Beads, 用裂解缓冲液清洗 Beads 5 次。加入洗脱缓冲液洗脱 Beads, 收集洗脱液; $30,000 \times g$ 4°C 离心 30 分钟, 取上清液, 即为纯化的 PB2-37 截短体流感病毒 RNA 聚合酶。

[0097] 用截留分子量为 30,000 道尔顿的超滤管浓缩纯化的 PB2-37 截短体流感病毒 RNA 聚合酶至 1mL。

[0098] 上述洗脱缓冲液按照如下方法制备: 将如下溶质按照如下终浓度溶解在水中得到的溶液: 25mM 4-羟乙基哌嗪乙磺酸、200mM NaCl、250mM 咪唑, 10% 丙三醇; 调节洗脱缓冲液的 pH 值为 7.8。

[0099] 将上述纯化的 PB2-37 截短体流感病毒 RNA 聚合酶用 SDS-PAGE 检测。

[0100] 结果如图 2 所示, Lane1、2、3 分别代表 5 微克纯化后的 PB2-37 截短体流感病毒 RNA 聚合酶、2.5 微克纯化后的 PB2-37 截短体流感病毒 RNA 聚合酶、3 微克纯化后的 PB2-37 截短体流感病毒 RNA 聚合酶; 可以看出, 得到分子量约为 80KD 的 PB1 亚基条带、稍小于 80KD 的 PA 条带以及隐约可见的含有 37 个氨基酸残基的 PB2 条带 (4KD), 表明, 得到 PB2-37 截短体流感病毒 RNA 聚合酶。

[0101] 分别回收每个亚基的条带, 送去进行蛋白 N 端测序, 验证得确得到 H5N1 流感病毒

RNA 聚合酶的三个亚基,表明电泳检测的目的大小的条带得确为 H5N1 流感病毒 RNA 聚合酶的三个亚基。

[0102] 计算含量,每升培养液可获得 1mg 的纯化 RNA 聚合酶 (PA-PB2-PB2-37)。

[0103] 二、流感病毒 RNA 聚合酶的大规模表达、纯化

[0104] 1、流感病毒 RNA 聚合酶大规模表达

[0105] 使用波式反应器或锥形瓶或其它类型的细胞培养瓶,使用 SF900II SFM 培养基悬浮培养 20L High Five 细胞,待细胞密度生长至 2×10^6 cells/mL 时,将上述一得到的 P3 代表达 PB2-37 的重组昆虫杆状病毒、P3 代表达 PB1 的重组昆虫杆状病毒、P3 代表达 PA 的重组昆虫杆状病毒等滴度量 (滴度均为 108pfu/mL) 混合后按 1% (体积 / 体积) 加入,27°C 培养 60 小时,使用离心机在 4°C、离心力 4000g 下离心 10 分钟收集细胞。

[0106] 在培养过程中,开启波式反应器的给氧系统给培养袋不断补充氧气。如使用锥形瓶或其它类型细胞培养瓶则只需保持通气即可,不需要特别补氧方式。

[0107] 将上述收集细胞按照 1:10 (质量 : 体积) 溶解于裂解缓冲液中,超声破碎方法破碎细胞,使用超声破碎细胞是在 0°C 冰液中保温放置装有细胞悬液的容器,在 30w 功率 /50mL 细胞溶液比率下,按照超声 3 秒停 6 秒方式总计超声破碎 10-30 分钟的方式充分裂解,之后, $40,000 \times g$ 4°C 离心 60 分钟,收集细胞裂解液上清液,即为含有 PB2-37 截短体流感病毒 RNA 聚合酶 的上清液。

[0108] 上述裂解缓冲液按照如下方法制备 :将如下溶质按照如下终浓度溶解在水中得到的溶液 :250mM NaCl、25mM Tris,调节裂解缓冲液的 pH 值为 8.0。

[0109] 2、流感病毒 RNA 聚合酶的纯化

[0110] A :亲和层析

[0111] 将上述 1 得到的含有 PB2-37 截短体流感病毒 RNA 聚合酶的上清液吸取到预冷的烧杯中,加入适量用裂解缓冲液清洗过的镍柱磁珠 (Nickel Beads),4°C 结合 3 小时。 $500 \times g$ 4°C 离心 3 分钟收集 Beads,用裂解缓冲液清洗 Beads 5 次。此时可将 Beads 加入一个相应大小的空的层析柱中,进一步用清洗液清洗。(该步骤也可将离心获得的上清液直接加入含有镍柱磁珠的柱子中进行结合,然后在柱上进行冲洗及洗脱过程以便亲和纯化目的复合物)清洗充分后,加入洗脱缓冲液冲洗 Beads,此时蛋白复合物从 beads 上被洗脱下来,通常需要用 2 倍 -5 倍 Beads 体积洗脱目的蛋白,收集洗脱液,得亲和层析纯化 PB2-37 截短体流感病毒 RNA 聚合酶 (电泳检测,结果与图 2 一致,说明得到目的蛋白)。

[0112] 用截留分子量为 30,000 道尔顿的超滤管浓缩至 5mL,得到亲和层析后浓缩样品。

[0113] 将浓缩样品 $16,000 \times g$ 4°C 离心 30 分钟,收集上清液 (浓缩后蛋白聚合沉淀,收集在上清液中存在的蛋白进行下述实验)。

[0114] 上述洗脱缓冲液按照如下方法制备 :将如下溶质按照如下终浓度溶解在水中得到的溶液 :25mM 4-羟乙基哌嗪乙磺酸、200mM NaCl、250mM 咪唑,10% 丙三醇 ;调节洗脱缓冲液的 pH 值为 7.8。

[0115] B、离子交换层析

[0116] 通过亲和层析初步纯化的目的蛋白需要进一步纯化,首选阳离子交换层析的方法。

[0117] 过程如下 :在 AKTA FPLC 或 AKTA Purifier 等层析设备上用低盐缓冲液平衡阳离

子交换柱 HiTrap S HP (含有介质 beads 5 mL, 来自于 GE Healthcare, 层析柱直径约为 1 厘米, 长约为 3 厘米), 用购自 GE 公司的 FPLC 系统 Purifier 10 的上样环将上述 A 得到的浓缩样品离心后的上清液) 注入层析系统, 使上述 A 获得的亲和层析后浓缩样品上清液缓缓流过离子交换柱, 目的蛋白将结合在离子交换柱上。为了保险起见, 收集未结合到离子交换柱的样品, 可以再次将此流穿样品上样, 以便获得进一步的结合, 之后继续用低盐缓冲液清洗离子交换柱, 直至紫外吸收值没有明显变化为止。调整缓冲液洗脱方式, 用 200 mM 至 400 mM 盐浓度线性梯度洗脱样品。

[0118] 结果如图 1A 所示, 收集 200 mM 至 400 mM 洗脱得到的洗脱液 (对应的 Peak 1 峰) 为离子交换层析纯化后 PB2-37 截短体流感病毒 RNA 聚合酶 (电泳检测, 结果与图 2 一致, 说明得到目的蛋白)。

[0119] 将离子交换层析纯化后流感病毒 RNA 聚合酶用截留分子量为 30,000 道尔顿的超滤管浓缩至 1 mL, 得到离子交换层析后浓缩样品。

[0120] 将浓缩样品 $16,000 \times g$ 4°C 离心 30 分钟, 收集上清液。

[0121] 上述低盐缓冲液按照如下方法制备: 将如下溶质按照如下终浓度溶解在水中得到的溶液: 25 mM 4-羟乙基哌嗪乙磺酸、200 mM NaCl; pH 值为 7.8;

[0122] 上述高盐缓冲液按照如下方法制备: 将如下溶质按照如下终浓度溶解在水中得到的溶液: 25 mM 4-羟乙基哌嗪乙磺酸、1 M NaCl; pH 7.8。

[0123] C、凝胶层析

[0124] 用凝胶层析缓冲液平衡凝胶层析柱 Superdex 200 16/60 分子筛 (GE healthcare, 柱径 1.6 厘米, 柱长 60 厘米) 将上述 B 获得的离子交换层析后浓缩样品上清液注入凝胶层析系统, 按每分钟 1 mL 的流速用凝胶层析洗脱缓冲液洗脱样品。

[0125] 观察紫外吸收值, 根据出现紫外吸收峰对应的位置收集样品。

[0126] 结果如图 1B, 收集 280 nm 或 260 nm 下 66 ± 2 mL 处出峰对应的洗脱液 (电泳检测, 结果与图 2 一致, 说明得到目的蛋白), 为纯化 PB2-37 截短体流感病毒 RNA 聚合酶。

[0127] 上述凝胶层析洗脱缓冲液按照如下方法制备: 将如下溶质按照如下终浓度溶解在水中得到的溶液: 25 mM HEPES、200 mM NaCl、1 mM EDTA, 调节凝胶层析缓冲液的 pH 值为 7.8。

[0128] 用截留分子量为 30,000- 道尔顿的超滤管浓缩纯化 PB2-37 截短体流感病毒 RNA 聚合酶至 1 mL, 得到浓缩 PB2-37 截短体流感病毒 RNA 聚合酶。

[0129] 将浓缩样品 $16,000 \times g$ 4°C 离心 30 分钟, 收集上清液。

[0130] 三、流感病毒 RNA 聚合酶与流感病毒的核酸 promoter 双链 RNA 结合

[0131] 为了结晶, 必须将流感病毒 RNA 聚合酶结合核酸, 通常在病毒中通常流感病毒 RNA 聚合酶与其 promoter 核酸 (RNA) 结合, 具体如下:

[0132] 1、promoter 核酸 (RNA) 的获得

[0133] 首先, 合成流感病毒的 cRNA 或 vRNA promoter 的两条链, 共四条 RNA, 序列分别为: 第一条: 5'-AGUAGAAACAAGGGUG-3' 含有 16 个核苷酸; 第二条: 5'-CACCCUGCUUUUGCU-3', 含有 15 个核苷酸; 第三条: 5'-AGCAAAAGCAGGGUG-3', 含有 15 个核苷酸; 第四条: 5'-CACCCUUGUUUCUACU-3', 含有 16 个核苷酸。其中, 第一与第二条 RNA 为 vRNA promoter 的两条配对链, 它们按 1:1 摩尔比混合, 构成 vRNA promoter。

[0134] 第三与第四条 RNA 为 cRNA promoter 的两条配对链, 它们按 1:1 摩尔比混合, 构成

cRNApromoter。

[0135] 将 vRNA promoter 的两条链按 1:1 混合,进行退火 (Annealing) 操作 (95°C 5 分钟后,迅速转移到 0°C 保温),得到 vRNA promoter。

[0136] 同样,将 cRNA promoter 的两条链按 1:1 混合,进行退火 (Annealing) 操作 (95°C 5 分钟后,迅速转移到 0°C 保温),得到 cRNA promoter。

[0137] 2、流感病毒 RNA 聚合酶与 promoter 双链 RNA 结合

[0138] 在上述二的 C 方法凝胶层析所获得的浓缩样品上清液中加入 vRNA promoter,在 0°C 中保温 0.5 小时,得到结合 vRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶。

[0139] 将上述结合 vRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶按照上述二的 C 方法凝胶层析纯化,结果图 1C 所示,收集 280nm 或 260nm 下 66±2mL 处出峰对应的洗脱液,为纯化结合 vRNA promoter 的流感病毒 RNA 聚合酶,说明加入 RNA promoter 之后蛋白质复合物出峰位置仍然为 66mL,说明结合 vRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶为单体 (约 165KD)。

[0140] 采用同样的方法将 vRNA promoter 结合 PB2-37 截短体流感病毒 RNA 聚合酶,结果与上述一致。

[0141] 四、流感病毒 RNA 聚合酶的结晶

[0142] 由三得到的结合 vRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶和结合 cRNApromoter 的 PB2-37 截短体流感病毒 RNA 聚合酶,经浓缩后,达到 20mg/mL 浓度,可以用于进行结晶实验。

[0143] 采用商购的结晶试剂盒 (hampton reaserch company),将上述浓缩后结合 vRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶和结合 cRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶,均在结晶缓冲液条件为 5% PEG20000、100mMTrisHCL (pH8.5)、250mMNaCL 以及 16°C 温度环境下结晶。

[0144] 结合 vRNA promoter 的 PB2-37 截短体流感病毒 RNA 聚合酶结晶图如图 7A 所示,结合 cRNApromoter 的 PB2-37 截短体流感病毒 RNA 聚合酶结晶图如图 7B 所示,可以看出,均获得类似于长方形外观的流感病毒 RNA 聚合酶晶体。

[0145] 为了确定此为目的蛋白结晶是否为流感病毒 RNA 聚合酶晶体,取出晶体后,使用沉淀剂反复洗涤,除去粘在晶体上的蛋白溶液,洗净的晶体进行 SDS-PAGE 电泳分析,结果与图 2 所示的结果一致,说明该晶体为 PB2-37 截短体流感病毒 RNA 聚合酶结晶晶体。

[0146] 实施例 2、含有 PB2-130 截短体的流感病毒 RNA 聚合酶的结晶

[0147] 本实施例预结晶的流感病毒 RNA 聚合酶聚合体的三个亚基包括 PB1 蛋白、PA 蛋白和 PB2-130 截短体;

[0148] PB2-130 截短体的氨基酸序列为序列 2 自 5' 末端第 1-130 位氨基酸,对应的核苷酸序列为序列 5 自 5' 末端第 1-390 位核苷酸。

[0149] 一、流感病毒 RNA 聚合酶的获得

[0150] 1、流感病毒 RNA 聚合酶各亚基的表达

[0151] 1)、表达流感病毒 RNA 聚合酶各亚基的重组载体的构建

[0152] 由于后期要进行镍柱纯化,因此需要在亚基 PB2-130 或 PB1 的羧基端加入 HIS 标签,两种都进行实验,发现在复合物表达过程中,含有组氨酸标签的蛋白只需在亚基

PB2-130 和 PB1 任一个亚基上即可,没有必要 PB1 及 PB2 亚基同时都含有组蛋白标签。

[0153] 在亚基 PB2-130 的羧基端加入 HIS 标签,扩增亚基 PB2-130 编码基因的羧基端引物为 5' ATCTCGAGCTAGTGGTGGTGGTGGTGGTGAAGGTTCCATGTTTTAACCTTTCAACC3' 含终止子,含 6 个串联的组氨酸标签;

[0154] 在亚基 PB2-130 的羧基端不加入 HIS 标签,扩增亚基 PB2-130 编码基因的羧基端引物为:5' ATCTCGAGCTAGAAGGTTCCATGTTTTAACCTTTCAACC3';

[0155] 下面的实验以均为在 PB1 的羧基端加入 HIS 标签为例。

[0156] (1) 重组载体 pFastBac-PB2-130 的获得

[0157] 氨基端引物序列为:5' ATGGATCCATGGAGAGAATAAAAAGAATTAAGAG3';羧基端引物为:5' ATCTCGAGCTAGAAGGTTCCATGTTTTAACCTTTCAACC3' (含终止子,不含组氨酸标签)。

[0158] 以 PB2 蛋白的编码基因为模板,用上述氨基端引物和羧基端引物进行 PCR 扩增,得到 PCR 产物。

[0159] 将该 PCR 产物分别酶切后连接到载体 pFastBac 质粒上的克隆位点上,从而构建出表达 PB2 氨基端含 1-130 位氨基酸残基片段的重组载体。

[0160] 经过测序,该重组载体为将序列表中序列 5 自 5' 末端第 1-390 位核苷酸所示的 PB2-130 截短体蛋白的编码基因插入表达载体 pFASTBAC 载体的 BamHI 和 XhoI 酶切位点间,表达 PB2-130 截短体蛋白,得到的重组载体,命名为 pFastBac-PB2-130。

[0161] (2) 重组载体 pFastBac-PB1 的获得:与实施例 1 方法相同。

[0162] (3) 重组载体 pFastBac-PA 的获得:与实施例 1 方法相同。

[0163] 2) 重组昆虫杆状病毒的制备:与实施例 1 方法基本相同,不同的是将 pFastBac-PB2-37 替换为 pFastBac-PB2-130,得到 P3 代表 PB2-130 的重组昆虫杆状病毒、P3 代表 PB1 的重组昆虫杆状病毒、P3 代表 PA 的重组昆虫杆状病毒。

[0164] 3)、流感病毒 RNA 聚合酶的表达

[0165] 按照实施例 1 的方法,将上述 2) 得到 P3 代表 PB2-130 的重组昆虫杆状病毒、P3 代表 PB1 的重组昆虫杆状病毒、P3 代表 PA 的重组昆虫杆状病毒等滴度量(滴度均为 108pfu/mL)混合后转入细胞 High Five,裂解,得到含有 PB2-130 截短体流感病毒 RNA 聚合酶的上清液。

[0166] 2、流感病毒 RNA 聚合酶的镍亲和层析纯化

[0167] 将上述 1 得到的含有 PB2-130 截短体流感病毒 RNA 聚合酶的上清液按照上述实施例 1 的方法进行镍柱亲和层析纯化,得到纯化的 PB2-130 截短体流感病毒 RNA 聚合酶。

[0168] 将上述纯化的 PB2-130 截短体流感病毒 RNA 聚合酶用 SDS-PAGE 检测。

[0169] 结果如图 5 所示,1-130 为纯化的 PB2-130 截短体流感病毒 RNA 聚合酶,可以看出,得到分子量约为 80KD 的 PB1 亚基条带、稍小于 80KD 的 PA 条带以及 14KD PB2 的 130 个氨基酸残基条带,表明,得到 PB2-130 截短体流感病毒 RNA 聚合酶。

[0170] 分别回收每个亚基的条带,送去进行蛋白 N 端测序,验证得确得到 H5N1 流感病毒 RNA 聚合酶的三个亚基,表明电泳检测的目的大小的条带得确为 H5N1 流感病毒 RNA 聚合酶的三个亚基。

[0171] 计算含量,每升培养液可获得 0.8mg 的纯化 RNA 聚合酶 (PA-PB1-PB2-130)。

[0172] 二、流感病毒 RNA 聚合酶的大规模表达、纯化

[0173] 1、流感病毒 RNA 聚合酶大规模表达

[0174] 按照实施例 1 的方法将 P3 代表 PB2-130 的重组昆虫杆状病毒、P3 代表 PB1 的重组昆虫杆状病毒、P3 代表 PA 的重组昆虫杆状病毒等滴度量（滴度均为 108pfu/mL）混合后按 1%（体积/体积）加入 High Five 细胞、培养、裂解，收集上清液，得到含有 PB2-130 截短体流感病毒 RNA 聚合酶的上清液。

[0175] 2、流感病毒 RNA 聚合酶的纯化

[0176] A：亲和层析

[0177] 按照实施例 1 的方法将上述 1 得到的含有 PB2-130 截短体流感病毒 RNA 聚合酶的上清液进行镍柱亲和层析，用 2 倍-5 倍 Beads 体积洗脱目的蛋白，收集洗脱液，得亲和层析纯化 PB2-130 截短体流感病毒 RNA 聚合酶。

[0178] 电泳检测，结果如图 4A 的泳道 1 和 2 所示，可以看出，得到分子量约为 80KD 的 PB1 亚基条带、稍小于 80KD 的 PA 条带以及 14KD PB2 的 130 个氨基酸残基条带，表明，得到 PB2-130 截短体流感病毒 RNA 聚合酶。

[0179] B、离子交换层析

[0180] 按照实施例 1 的方法将上述 A 纯化后的 PB2-130 截短体流感病毒 RNA 聚合酶进行离子交换层析，用 200mM 至 400mM 盐浓度线性梯度洗脱样品。

[0181] 结果如图 4B 所示，收集 200mM 至 400mM 洗脱得到的洗脱液（对应的主峰）为离子交换层析纯化后 PB2-130 截短体流感病毒 RNA 聚合酶。

[0182] 电泳检测，结果如图 4A 的泳道 3 所示，说明得到目的蛋白。

[0183] C、凝胶层析

[0184] 按照实施例 1 的方法将上述 B 纯化后的 PB2-130 截短体流感病毒 RNA 聚合酶的上清液进行凝胶层析，结果如图 3 所示，收集 280nM 或 260nM 下大约为 58 毫升处出峰对应的洗脱液，对应分子量为复合体二体大小，这与纯化表达的含有 PB2-37 的流感病毒 RNA 聚合酶复合体样品出峰位置不同，为纯化 PB2-130 截短体流感病毒 RNA 聚合酶。

[0185] 电泳检测，结果如图 4A 的泳道 5 和 6 所示，说明得到目的蛋白。

[0186] 三、流感病毒 RNA 聚合酶与流感病毒的核酸 promoter 双链 RNA 结合

[0187] 1、promoter 核酸 (RNA) 的获得：与实施例 1 相同。

[0188] 2、流感病毒 RNA 聚合酶与 promoter 双链 RNA 结合

[0189] 与实施例 1 基本相同，不同的是替换为纯化 PB2-130 截短体流感病毒 RNA 聚合酶，得到结合 vRNA promoter 的 PB2-130 截短体流感病毒 RNA 聚合酶和结合 cRNA promoter 的 PB2-130 截短体流感病毒 RNA 聚合酶。

[0190] 将结合 vRNA promoter 的 PB2-130 截短体流感病毒 RNA 聚合酶和未结合 RNA 的 PB2-130 截短体流感病毒 RNA 聚合酶凝胶层析纯化。

[0191] 结果图 4C 所示，收集 280nM 下 53 毫升处出峰对应的洗脱液，为结合 vRNA promoter 的 PB2-130 截短体流感病毒 RNA 聚合酶（四体形式，分子量约为 720KD），收集 280nM 下 68 毫升出峰对应的洗脱液，为 PB2-130 截短体流感病毒 RNA 聚合酶（单体形式，分子量约为 180KD），58 毫升处出峰对应的洗脱液，为结合 vRNA promoter 的 PB2-130 截短体流感病毒 RNA 聚合酶（二体形式，分子量约为 360KD），表明 PB2-130 截短体流感病毒 RNA 聚合酶主要以二体（未加 RNA 启动子时）和四体（加入 RNA 启动子后）形式存在，而不是

PB-103 截短体蛋白的编码基因插入表达载体 pFASTBAC 载体的 BamHI 和 XhoI 酶切位点间, 表达 PB-103 截短体蛋白, 得到的重组载体, 命名为 pFastBac-PB2-103。

[0212] (2) 重组载体 pFastBac-PB1 的获得: 与实施例 1 方法相同。

[0213] (3) 重组载体 pFastBac-PA 的获得: 与实施例 1 方法相同。

[0214] 2) 重组昆虫杆状病毒的制备: 与实施例 1 方法基本相同, 不同的是将 pFastBac-PB2-37 替换为 pFastBac-PB2-103, 得到 P3 代表 PB2-103 的重组昆虫杆状病毒、P3 代表 PB1 的重组昆虫杆状病毒、P3 代表 PA 的重组昆虫杆状病毒。

[0215] 3)、流感病毒 RNA 聚合酶的表达

[0216] 按照实施例 1 的方法, 将上述 2) 得到 P3 代表 PB2-103 的重组昆虫杆状病毒、P3 代表 PB1 的重组昆虫杆状病毒、P3 代表 PA 的重组昆虫杆状病毒等滴度量 (滴度均为 108pfu/mL) 混合后转入细胞 High Five, 裂解, 得到含有 PB2-103 截短体流感病毒 RNA 聚合酶的上清液。

[0217] 2、流感病毒 RNA 聚合酶的镍亲和层析纯化

[0218] 将上述 1 得到的含有 PB2-103 截短体流感病毒 RNA 聚合酶的上清液按照上述实施例 1 的方法进行镍柱亲和层析纯化, 得到纯化的 PB2-103 截短体流感病毒 RNA 聚合酶。

[0219] 将上述纯化的 PB2-103 截短体流感病毒 RNA 聚合酶用 SDS-PAGE 检测。

[0220] 结果如图 5 所示, 1-103 为纯化的 PB2-103 截短体流感病毒 RNA 聚合酶, 可以看出, 得到分子量约为 80KD 的 PB1 亚基条带、稍小于 80KD 的 PA 条带以及 11KD PB2 的 103 个氨基酸残基条带, 表明, 得到 PB2-103 截短体流感病毒 RNA 聚合酶。

[0221] 分别回收每个亚基的条带, 送去进行蛋白 N 端测序, 验证得确得到 H5N1 流感病毒 RNA 聚合酶的三个亚基, 表明电泳检测的目的大小的条带得确为 H5N1 流感病毒 RNA 聚合酶的三个亚基。

[0222] 计算含量, 每升培养液可获得 5mg 的纯化 RNA 聚合酶 (PA-PB1-PB2-103)。

[0223] 二、流感病毒 RNA 聚合酶的大规模表达、纯化

[0224] 1、流感病毒 RNA 聚合酶大规模表达

[0225] 按照实施例 1 的方法将 P3 代表 PB2-103 的重组昆虫杆状病毒、P3 代表 PB1 的重组昆虫杆状病毒、P3 代表 PA 的重组昆虫杆状病毒等滴度量 (滴度均为 108pfu/mL) 混合后按 1% (体积/体积) 加入 High Five 细胞、培养、裂解, 收集上清液, 得到含有 PB2-103 截短体流感病毒 RNA 聚合酶的上清液。

[0226] 2、流感病毒 RNA 聚合酶的纯化

[0227] A: 亲和层析

[0228] 按照实施例 1 的方法将上述 1 得到的含有 PB2-103 截短体流感病毒 RNA 聚合酶的上清液进行镍柱亲和层析, 用 2 倍-5 倍 Beads 体积洗脱目的蛋白, 收集洗脱液, 得亲和层析纯化 PB2-103 截短体流感病毒 RNA 聚合酶。

[0229] 电泳检测, 与图 5 的 1-103 大小一致。表明, 得到 PB2-103 截短体流感病毒 RNA 聚合酶。

[0230] B、离子交换层析

[0231] 按照实施例 1 的方法将上述 A 纯化后的 PB2-103 截短体流感病毒 RNA 聚合酶进行离子交换层析, 用 200mM 至 400mM 盐浓度线性梯度洗脱样品, 结果与图 4B 一致, 收集 200mM

至 400mM 洗脱得到的洗脱液（对应的主峰）为离子交换层析纯化后 PB2-103 截短体流感病毒 RNA 聚合酶。

[0232] 电泳检测,结果图 5 的 1-103 大小一致,说明得到目的蛋白。

[0233] C、凝胶层析

[0234] 按照实施例 1 的方法将实施例 1 的方法将上述 B 纯化后的 PB2-103 截短体流感病毒 RNA 聚合酶的上清液进行凝胶层析,结果与图 3 一致。

[0235] 电泳检测,结果图 5 的 1-103 大小一致,说明得到目的蛋白。

[0236] 三、流感病毒 RNA 聚合酶与流感病毒的核酸 promoter 双链 RNA 结合

[0237] 1、promoter 核酸 (RNA) 的获得:与实施例 1 相同。

[0238] 2、流感病毒 RNA 聚合酶与 promoter 双链 RNA 结合

[0239] 与实施例 1 基本相同,不同的是替换为纯化 PB2-103 截短体流感病毒 RNA 聚合酶,得到结合 vRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶和结合 cRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶。

[0240] 将结合 vRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶和未结合 RNA 的 PB2-103 截短体流感病毒 RNA 聚合酶凝胶层析纯化。

[0241] 结果与图 4C 一致所示,收集 280nM 下 53 毫升处出峰对应的洗脱液,为结合 vRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶(四体形式,分子量为 720KD);收集 280nM 下 68 毫升出峰对应的洗脱液,为 PB2-103 截短体流感病毒 RNA 聚合酶(二体形式,分子量为 360KD),收集 280nM 下 58 毫升出峰对应的洗脱液,为 PB2-103 截短体流感病毒 RNA 聚合酶(单体形式,分子量为 180KD);表明 PB2-103 截短体流感病毒 RNA 聚合酶主要以二体(未加 RNA 启动子时)和四体(加入 RNA 启动子后)形式存在,而不是 PB2-37 截短体流感病毒 RNA 聚合酶的单体存在形式。

[0242] 结合 vRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶和结合 cRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶结果无显著差异。

[0243] 四、流感病毒 RNA 聚合酶的结晶

[0244] 采用商购的结晶试剂盒(hampton reaserch company),将上述浓缩后结合 vRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶和结合 cRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶,均在结晶缓冲液条件为 30% PEG20000、100mM TrisHCL(pH8.5)、500mM NaCl 以及 20℃ 温度环境下结晶。

[0245] 结合 vRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶结晶图如图 7E 所示,结合 cRNA promoter 的 PB2-103 截短体流感病毒 RNA 聚合酶结晶图如图 7F 所示,可以看出,均获得类似于长方形外观的流感病毒 RNA 聚合酶晶体。

[0246] 为了确定此为目的蛋白结晶是否为流感病毒 RNA 聚合酶晶体,取出晶体后,使用沉淀剂反复洗涤,除去粘在晶体上的蛋白溶液,洗净的晶体进行 SDS-PAGE 电泳分析,结果与图 5 的 1-103 一致,说明该晶体为 PB2-103 截短体流感病毒 RNA 聚合酶晶体。

[0001]

序列表

<110>中国科学院生物物理研究所
 <120>流感病毒 RNA 聚合酶结晶的方法
 <160> 6

<170> PatentIn version 3.5

<210> 1

<211> 757

<212> PRT

<213>人工序列

<220>

<223>

<400> 1

```

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

```

[0002]

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

[0003]

515	520	525	
Val Ile Lys Asn Asn Met Ile Asn Asn Asp Leu Gly Pro Ala Thr Ala			
530	535	540	
Gln Met Ala Leu Gln Leu Phe Ile Lys Asp Tyr Arg Tyr Thr Tyr Arg			
545	550	555	560
Cys His Arg Gly Asp Thr Gln Ile Gln Thr Arg Arg Ser Phe Glu Leu			
	565	570	575
Lys Lys Leu Trp Glu Gln Thr Arg Ser Lys Ala Gly Leu Leu Val Ser			
	580	585	590
Asp Gly Gly Pro Asn Leu Tyr Asn Ile Arg Asn Leu His Ile Pro Glu			
	595	600	605
Val Cys Leu Lys Trp Glu Leu Met Asp Glu Asp Tyr Gln Gly Arg Leu			
	610	615	620
Cys Asn Pro Leu Asn Pro Phe Val Ser His Lys Glu Ile Glu Ser Val			
625	630	635	640
Asn Asn Ala Val Val Met Pro Ala His Gly Pro Ala Lys Ser Met Glu			
	645	650	655
Tyr Asp Ala Val Ala Thr Thr His Ser Trp Ile Pro Lys Arg Asn Arg			
	660	665	670
Ser Ile Leu Asn Thr Ser Gln Arg Gly Ile Leu Glu Asp Glu Gln Met			
	675	680	685
Tyr Gln Lys Cys Cys Asn Leu Phe Glu Lys Phe Phe Pro Ser Ser Ser			
	690	695	700
Tyr Arg Arg Pro Val Gly Ile Ser Ser Met Val Glu Ala Met Val Ser			
705	710	715	720
Arg Ala Arg Ile Asp Ala Arg Ile Asp Phe Glu Ser Gly Arg Ile Lys			
	725	730	735
Lys Glu Glu Phe Ala Glu Ile Met Lys Ile Cys Ser Thr Ile Glu Glu			
	740	745	750
Leu Arg Arg Gln Lys			
	755		
<210> 2			
<211> 759			
<212> PRT			
<213>人工序列			
<220>			
<223>			
<400> 2			
Met Glu Arg Ile Lys Glu Leu Arg Asp Leu Met Ser Gln Ser Arg Thr			

[0004]

1	5	10	15												
Arg	Glu	Ile	Leu	Thr	Lys	Thr	Thr	Val	Asp	His	Met	Ala	Ile	Ile	Lys
	20	25	30												
Lys	Tyr	Thr	Ser	Gly	Arg	Gln	Glu	Lys	Asn	Pro	Ala	Leu	Arg	Met	Lys
	35	40	45												
Trp	Met	Met	Ala	Met	Lys	Tyr	Pro	Ile	Thr	Ala	Asp	Lys	Arg	Ile	Met
	50	55	60												
Glu	Met	Ile	Pro	Glu	Arg	Asn	Glu	Gln	Gly	Gln	Thr	Leu	Trp	Ser	Lys
65		70		75		80									
Thr	Asn	Asp	Ala	Gly	Ser	Asp	Arg	Val	Met	Val	Ser	Pro	Leu	Ala	Val
	85	90	95												
Thr	Trp	Trp	Asn	Arg	Asn	Gly	Pro	Thr	Thr	Ser	Thr	Val	His	Tyr	Pro
	100	105	110												
Lys	Val	Tyr	Lys	Thr	Tyr	Phe	Glu	Lys	Val	Glu	Arg	Leu	Lys	His	Gly
	115	120	125												
Thr	Phe	Gly	Pro	Val	His	Phe	Arg	Asn	Gln	Val	Lys	Ile	Arg	Arg	Arg
	130	135	140												
Val	Asp	Ile	Asn	Pro	Gly	His	Ala	Asp	Leu	Ser	Ala	Lys	Glu	Ala	Gln
145		150		155		160									
Asp	Val	Ile	Met	Glu	Val	Val	Phe	Pro	Asn	Glu	Val	Gly	Ala	Arg	Ile
	165	170	175												
Leu	Thr	Ser	Glu	Ser	Gln	Leu	Thr	Ile	Thr	Lys	Glu	Lys	Lys	Glu	Glu
	180	185	190												
Leu	Gln	Asp	Cys	Lys	Ile	Ala	Pro	Leu	Met	Val	Ala	Tyr	Met	Leu	Glu
	195	200	205												
Arg	Glu	Leu	Val	Arg	Lys	Thr	Arg	Phe	Leu	Pro	Val	Ala	Gly	Gly	Thr
	210	215	220												
Ser	Ser	Val	Tyr	Ile	Glu	Val	Leu	His	Leu	Thr	Gln	Gly	Thr	Cys	Trp
225		230		235		240									
Glu	Gln	Met	Tyr	Thr	Pro	Gly	Gly	Glu	Val	Arg	Asn	Asp	Asp	Val	Asp
	245	250	255												
Gln	Ser	Leu	Ile	Ile	Ala	Ala	Arg	Asn	Ile	Val	Arg	Arg	Ala	Thr	Val
	260	265	270												
Ser	Ala	Asp	Pro	Leu	Ala	Ser	Leu	Leu	Glu	Met	Cys	His	Ser	Thr	Gln
	275	280	285												
Ile	Gly	Gly	Ile	Arg	Met	Val	Asp	Ile	Leu	Arg	Gln	Asn	Pro	Thr	Glu
	290	295	300												
Glu	Gln	Ala	Val	Asp	Ile	Cys	Lys	Ala	Ala	Met	Gly	Leu	Arg	Ile	Ser
305		310		315		320									
Ser	Ser	Phe	Ser	Phe	Gly	Gly	Phe	Thr	Phe	Lys	Arg	Thr	Asn	Gly	Ser

[0005]

		325						330						335	
Ser	Val	Lys	Lys	Glu	Glu	Glu	Val	Leu	Thr	Gly	Asn	Leu	Gln	Thr	Leu
		340						345						350	
Lys	Ile	Lys	Val	His	Glu	Gly	Tyr	Glu	Glu	Phe	Thr	Met	Val	Gly	Arg
		355						360						365	
Arg	Ala	Thr	Ala	Ile	Leu	Arg	Lys	Ala	Thr	Arg	Arg	Leu	Ile	Gln	Leu
		370						375						380	
Ile	Val	Scr	Gly	Arg	Asp	Glu	Gln	Ser	Ile	Ala	Glu	Ala	Ile	Ile	Val
385						390					395				400
Ala	Met	Val	Phe	Ser	Gln	Glu	Asp	Cys	Met	Ile	Lys	Ala	Val	Arg	Gly
						405					410				415
Asp	Leu	Asn	Phe	Val	Asn	Arg	Ala	Asn	Gln	Arg	Leu	Asn	Pro	Met	His
						420					425				430
Gln	Leu	Leu	Arg	His	Phe	Gln	Lys	Asp	Ala	Lys	Val	Leu	Phe	Gln	Asn
						435									445
Trp	Gly	Ile	Glu	Pro	Ile	Asp	Asn	Val	Met	Gly	Met	Ile	Gly	Ile	Leu
						455									460
Pro	Asp	Met	Thr	Pro	Ser	Ala	Glu	Met	Ser	Leu	Arg	Gly	Val	Arg	Val
465						470									480
Ser	Lys	Met	Gly	Val	Asp	Glu	Tyr	Ser	Ser	Thr	Glu	Arg	Val	Val	Val
															485
															490
Ser	Ile	Asp	Arg	Phe	Leu	Arg	Val	Arg	Asp	Gln	Gln	Gly	Asn	Val	Leu
															500
															505
Leu	Ser	Pro	Glu	Glu	Val	Ser	Glu	Thr	Gln	Gly	Thr	Glu	Lys	Leu	Thr
															515
															520
Ile	Thr	Tyr	Ser	Ser	Ser	Met	Met	Trp	Glu	Ile	Asn	Gly	Pro	Glu	Ser
															530
															535
Val	Leu	Val	Asn	Thr	Tyr	Gln	Trp	Ile	Ile	Arg	Asn	Trp	Glu	Thr	Val
545															550
															555
Lys	Ile	Gln	Trp	Ser	Gln	Asp	Pro	Thr	Met	Leu	Tyr	Asn	Lys	Met	Glu
															565
															570
Phe	Glu	Ser	Phe	Gln	Ser	Leu	Val	Pro	Lys	Ala	Ala	Arg	Ser	Gln	Tyr
															580
															585
Ser	Gly	Phe	Val	Arg	Thr	Leu	Phe	Gln	Gln	Met	Arg	Asp	Val	Leu	Gly
															595
															600
															605
Thr	Phe	Asp	Thr	Val	Gln	Ile	Ile	Lys	Leu	Leu	Pro	Phe	Ala	Ala	Ala
															610
															615
															620
Pro	Pro	Glu	Pro	Ser	Arg	Met	Gln	Phe	Ser	Ser	Leu	Thr	Val	Asn	Val
625															630
															635
Arg	Gly	Ser	Gly	Met	Arg	Ile	Leu	Val	Arg	Gly	Asn	Ser	Pro	Val	Phe

[0006]

	645		650		655
Asn Tyr Asn Lys	Ala Thr Lys Arg	Leu Thr Val	Leu Gly Lys Asp	Ala	
	660		665		670
Gly Ala Leu Thr	Glu Asp Pro Asp	Glu Gly Thr	Ala Gly Val	Glu Ser	
	675		680		685
Ala Val Leu Arg	Gly Phe Leu Ile	Leu Gly Arg	Glu Asp Lys	Arg Tyr	
	690		695		700
Gly Pro Ala Leu	Ser Ile Asn Glu	Leu Ser Asn	Leu Ala Lys	Gly Glu	
705		710		715	
Lys Ala Asn Val	Leu Ile Met Gln	Gly Asp Val	Val Leu Val	Met Lys	
	725		730		735
Arg Lys Arg Asp	Phe Ser Ile Leu	Thr Asp Ser	Gln Thr Ala	Thr Lys	
	740		745		750
Arg Ile Arg Met	Ala Ile Asn				
	755				
<210>	3				
<211>	716				
<212>	PRT				
<213>	人工序列				
<220>					
<223>					
<400>	3				
Met Glu Asp Phe	Val Arg Gln Cys	Phe Asn Pro	Met Ile Val	Glu Leu	
1	5	10		15	
Ala Glu Lys Ala	Met Lys Glu Tyr	Gly Glu Asp	Pro Lys Ile	Glu Thr	
	20	25		30	
Asn Lys Phe Ala	Ala Ile Cys Thr	His Leu Glu	Val Cys Phe	Met Tyr	
	35	40		45	
Ser Asp Phe His	Phe Ile Asp	Glu Arg Gly	Glu Ser Thr	Ile Ile Glu	
	50	55		60	
Ser Gly Asp Pro	Asn Ala Leu	Leu Lys His	Arg Phe Glu	Ile Ile Glu	
65	70	75		80	
Gly Arg Asp Arg	Thr Met Ala	Trp Thr Val	Val Asn Ser	Ile Cys Asn	
	85	90		95	
Thr Thr Gly Val	Glu Lys Pro	Lys Phe Leu	Pro Asp Leu	Tyr Asp Tyr	
	100	105		110	
Lys Glu Asn Arg	Phe Ile Glu	Ile Gly Val	Thr Arg Arg	Glu Val His	
	115	120		125	
Thr Tyr Tyr Leu	Glu Lys Ala	Asn Lys Ile	Lys Ser Glu	Lys Thr His	

[0007]

130		135		140															
Ile	His	Ile	Phe	Ser	Phe	Thr	Gly	Glu	Glu	Met	Ala	Thr	Lys	Ala	Asp				
145					150					155					160				
Tyr	Thr	Leu	Asp	Glu	Glu	Ser	Arg	Ala	Arg	Ile	Lys	Thr	Arg	Leu	Phe				
				165						170					175				
Thr	Ile	Arg	Gln	Glu	Met	Ala	Ser	Arg	Gly	Leu	Trp	Asp	Ser	Phe	Arg				
				180						185					190				
Gln	Ser	Glu	Arg	Gly	Glu	Glu	Thr	Ile	Glu	Glu	Arg	Phe	Glu	Ile	Thr				
				195						200					205				
Gly	Thr	Met	Cys	Arg	Leu	Ala	Asp	Gln	Ser	Leu	Pro	Pro	Asn	Phe	Ser				
				210						215					220				
Ser	Leu	Glu	Lys	Phe	Arg	Ala	Tyr	Val	Asp	Gly	Phe	Glu	Pro	Asn	Gly				
				225						230					235				240
Cys	Ile	Glu	Gly	Lys	Leu	Ser	Gln	Met	Ser	Lys	Glu	Val	Asn	Ala	Arg				
				245						250					255				
Ile	Glu	Pro	Phe	Leu	Lys	Thr	Thr	Pro	Arg	Pro	Leu	Arg	Leu	Pro	Asp				
				260						265					270				
Gly	Pro	Pro	Cys	Ser	Gln	Arg	Ser	Lys	Phe	Leu	Leu	Met	Asp	Ala	Leu				
				275						280					285				
Lys	Leu	Ser	Ile	Glu	Asp	Pro	Ser	His	Glu	Gly	Glu	Gly	Ile	Pro	Leu				
				290						295					300				
Tyr	Asp	Ala	Ile	Lys	Cys	Met	Lys	Thr	Phe	Phe	Gly	Trp	Lys	Glu	Pro				
				305						310					315				320
Asn	Ile	Val	Lys	Pro	His	Glu	Lys	Gly	Ile	Asn	Pro	Asn	Tyr	Leu	Leu				
				325						330					335				
Ala	Trp	Lys	Gln	Val	Leu	Ala	Glu	Leu	Gln	Asp	Ile	Glu	Asn	Glu	Glu				
				340						345					350				
Lys	Ile	Pro	Lys	Thr	Lys	Asn	Met	Arg	Lys	Thr	Ser	Gln	Leu	Lys	Trp				
				355						360					365				
Ala	Leu	Gly	Glu	Asn	Met	Ala	Pro	Glu	Lys	Val	Asp	Phe	Glu	Asp	Cys				
				370						375					380				
Lys	Asp	Val	Ser	Asp	Leu	Arg	Gln	Tyr	Asp	Ser	Asp	Glu	Pro	Lys	Pro				
				385						390					395				400
Arg	Ser	Leu	Ala	Ser	Trp	Ile	Gln	Ser	Glu	Phe	Asn	Lys	Ala	Cys	Glu				
				405						410					415				
Leu	Thr	Asp	Ser	Ser	Trp	Ile	Glu	Leu	Asp	Glu	Ile	Gly	Glu	Asp	Val				
				420						425					430				
Ala	Pro	Ile	Glu	His	Ile	Ala	Ser	Met	Arg	Arg	Asn	Tyr	Phe	Thr	Ala				
				435						440					445				
Glu	Val	Ser	His	Cys	Arg	Ala	Thr	Glu	Tyr	Ile	Met	Lys	Gly	Val	Tyr				

[0008]

450	455	460
Ile Asn Thr Ala Leu Leu Asn Ala Ser Cys Ala Ala Met Asp Asp Phe		
465	470	475
Gln Leu Ile Pro Met Ile Ser Lys Cys Arg Thr Lys Glu Gly Arg Arg		
	485	490
Lys Thr Asn Leu Tyr Gly Phe Ile Ile Lys Gly Arg Ser His Leu Arg		
	500	505
Asn Asp Thr Asp Val Val Asn Phe Val Ser Met Glu Phe Ser Leu Thr		
	515	520
Asp Pro Arg Leu Glu Pro His Lys Trp Glu Lys Tyr Cys Val Leu Glu		
	530	535
Ile Gly Asp Met Leu Leu Arg Thr Ala Ile Gly Gln Val Ser Arg Pro		
545	550	555
Met Phe Leu Tyr Val Arg Thr Asn Gly Thr Ser Lys Ile Lys Met Lys		
	565	570
Trp Gly Met Glu Met Arg Arg Cys Leu Leu Gln Ser Leu Gln Gln Ile		
	580	585
Glu Ser Met Ile Glu Ala Glu Ser Ser Val Lys Glu Lys Asp Met Thr		
	595	600
Lys Glu Phe Phe Glu Asn Lys Ser Glu Thr Trp Pro Ile Gly Glu Ser		
	610	615
Pro Lys Gly Met Glu Glu Gly Ser Ile Gly Lys Val Cys Arg Thr Leu		
625	630	635
Leu Ala Lys Ser Val Phe Asn Ser Leu Tyr Ala Ser Pro Gln Leu Glu		
	645	650
Gly Phe Ser Ala Glu Ser Arg Lys Leu Leu Leu Ile Val Gln Ala Leu		
	660	665
Arg Asp Asn Leu Glu Pro Gly Thr Phe Asp Leu Gly Gly Leu Tyr Glu		
	675	680
Ala Ile Glu Glu Cys Leu Ile Asn Asp Pro Trp Val Leu Leu Asn Ala		
	690	695
Ser Trp Phe Asn Ser Phe Leu Thr His Ala Leu Lys		
705	710	715
<210> 4		
<211> 2270		
<212> DNA		
<213>人工序列		
<220>		
<223>		
<400> 4		

[0009]

atgtcaatcc gactttactt ttcttaaaag tgccagcgca aaatgctata agtaccacat	60
tccttatac tggagatcct ccatacagcc atggaacagg aacaggatac accatggaca	120
cagteaacag aacacatcaa tattcagaaa aggggaaatg gacaacgaac acagagactg	180
gagcacccea actcaatceg attgatggac cactacctga ggataatgag ccgagtgggt	240
atgcacaaac agattgtgta ttggaagcaa tggetttcct tgaaaaatcc cacccaggga	300
tctttgaaaa ctctgtctt gaaacgatgg aaattgttca gcaacaaga gtggataaac	360
tgaccaagg tegccaaacc tatgactgga cattgaatag aaaccagccg getgcaaccg	420
ctttggccaa cactatagag gtcttcagat cgaatggtct aacagccaat gaatcgggaa	480
ggetaataga tttctcaaa gacgtgatgg aatcaatgga taagggagaa atggaaataa	540
taacacattt ccagagaaa agaaagatga gggacaacat gaccaagaaa atggtcacac	600
aaagaacaat agggaagaag aaacaaagc tgaacaaaag gagctaccta ataagacac	660
tgacactgaa cacaatgaca aaagacgcag aaagaggcaa attgaagagg cgggcaattg	720
caacaccceg gatgcaaate agaggattcg tgtactttgt cgaaacacta gcgaggagta	780
tctgtgagaa acttgagcaa tctggactcc cagtcggagg gaatgaaaag aaggctaaat	840
tggcaaatgt cgtgaggaag atgatgacta acteacaaga tacagagctc tectttacaa	900
ttactggaga caacacaaa tggaatgaga atcagaacce teggatgttt cttagcaatga	960
taacatacat cacaaggaac caacctgaat ggtttagaaa tgtcttaagc attgctccta	1020
taatgttctc aaacaagatg gcaagattag ggaaaggata catgttcgaa agtaagagca	1080
tgaagetacg gacacaaata ccagcagaaa tgettgaag cattgacctg aaatacttca	1140
acgaatcaac gagaaagaaa atcgagaaaa taagacctct actaatagat ggcacagct	1200
cattgagtec tggaatgatg atgggatgt teaatatget gagtacagtc ttaggagttt	1260
caatcctgaa tcttggcag aagaggtaca ccaaaaccac atactgggtg gacggactcc	1320
aatcctctga tgattctct ctcatagtga atgeaccgaa tcatgaggga atacaagcag	1380
gggtggatag gttctatagg acttgcaaac tagttggaat caatatgagc aagaagaagt	1440
cttacataaa tcggacagga acatttgaat tcacaagctt cttctaccgc tatgggttcg	1500
tagccaactt cagtatggag ctgcccagct ttggagtgtc tgggattaat gaatcggctg	1560
acatgagcat tgggtttaca gtgataaaga acaatatgat aaacaacgac cttggaccag	1620
caacagctca gatgctctt cagctattca tcaaggacta cagatacaca taccgatgcc	1680
acagggggga tacacaaate caaacgagga gatcattcga gctgaagaag ctgtgggagc	1740
agacccttc aaaggcagga ctgttggtt cagatggagg accaaaccta tacaatatcc	1800
ggaatctcca cattccggag gtctgcttga agtgggaatt gatggatgaa gactaccagg	1860
gcagactgtg taatcctctg aacctgtt ttagtcataa ggaaattgag tctgtcaaca	1920
atctgttgtt aatgccagct catggcccag ccaagagcat ggaatacgat gcagttgca	1980
ctacacatte atggattcct aagaggaatc gttccattct caacaccagc caaaggggga	2040
ttcttgagga tgaacagatg tatcagaagt gctgcaatct attcagagaaa ttcttcccta	2100
gcagttcata tcggaggcca gttggaattt ccagcatggt ggaggccatg gtgtctaggg	2160
cccgaattga tgcacgaatt gacttegagt ctggaaggat taagaaagaa gatttctctg	2220
agatcatgaa gatctgttcc accattgaag agctcagacg gcaaaaaatg	2270

<210> 5

<211> 2280

[0010]

<212> DNA

<213>人工序列

<220>

<223>

<400> 5

atggagagaa	taaaagaatt	aagagatcta	atgtcgcagt	cccgcactcg	cgagatacta	60
acaaaaacca	ctgtggacca	tatggccata	atcaagaaat	acacatcagg	aagacaagag	120
aagaacctcg	ctctcagaat	gaaatggatg	atggcaatga	aatatccaat	cacagcagac	180
aagagaataa	tggagatgat	tcctgaaagg	aatgaacaag	gacaaacgct	ttggagcaag	240
acaaatgatg	ctgggtcggg	cagagtgatg	gtgtctcccc	tagctgtaac	ttgggtggaac	300
aggaatgggc	cgacaacaag	tacagtccat	tatccaaagg	tttacaaaac	atactttgag	360
aaggttgaaa	ggttanaaca	tggaaacctc	ggtcccgttc	atttccgaaa	ccaagttaaa	420
atacgtcgcc	gggtggatat	aaaccegggc	catgcagatc	tcagtgctaa	agaagcacia	480
gatgttatca	tggaggtcgt	tttcccaaat	gaagtgggag	ctagaatatt	gacatcagag	540
tcgcaatiga	caataacaaa	agagaagaaa	gaagagctcc	aggattgtaa	aattgctcct	600
ttaatggtgg	catacatggt	ggaaagagaa	ctggtccgca	aaaccagatt	tctaccggta	660
gcaggcggaa	caagcagtgt	gtacattgag	gtattgcatt	tgactcaagg	gacctgttgg	720
gaacagatgt	acactcccgg	eggagaagtg	agaaatgatg	atggtgacca	gagtttgatc	780
atcgtgcca	gaaacattgt	taggagagca	acagtatcag	cggatccact	ggcatcacta	840
ctggagatgt	gtcacagcac	acaaattggg	ggaataagga	tggtagcat	ccttaggcaa	900
aacccaactg	aggagcaage	tgtggatata	tgcaaagcag	caatgggttt	gaggatcagt	960
tcattcctta	gctttggagg	cttcaacttc	aaaagaacaa	gtggatcate	cgtcaagaag	1020
gaagaggaag	tgettacagg	caacctccaa	acattgaaaa	taagagtaca	tgaggggtat	1080
gaagaattca	caatggttgg	gcgagagca	acagctatcc	tgaggaaagc	aactagaagg	1140
ctgattcagt	tgatagtaag	tggaaagatg	gaacaatcaa	tcgctgaagc	gatcattgta	1200
gcaatggtgt	tctcacagga	ggattgcatg	ataaaggcag	tccgaggcga	tctgaatttc	1260
gtgaacagag	caaaccaaac	attgaacccc	atgcatcaac	tctgaggcca	cttccaaaaa	1320
gatgcaaaag	tgctgtttca	gaactgggga	attgaaccta	ttgacaatgt	catggggatg	1380
atcggaatat	tacctgacat	gactccaagc	gcagagatgt	cactgagagg	agtgagagtt	1440
agtaagatgg	gagtagatga	atattccagc	acggagagag	tggtagtgag	tattgaccgt	1500
ttcttgaggg	tccgagatca	gcaggggaac	gtactcttat	ctctgaaga	ggttagtgaa	1560
acacagggaa	cagagaagtt	gacaataaca	tattcactct	caatgatgtg	ggaaatcaac	1620
ggctctgagt	cagtgcttgt	taaaccttat	caatggatca	tcaggaattg	ggagactgta	1680
aagattcaat	ggtctcaaga	tcccacaatg	ctgtacaata	agatggagtt	tgaatcgttc	1740
caatccttgg	tgccataaagc	tgccagaggt	caatacagtg	ggtttgtgag	aacactattc	1800
caacaaatgc	gtgatgtaet	ggggacattt	gatactgtcc	aaataataaa	gctgctacca	1860
tttgcagcag	ccccaccgga	gcagagcaga	atgcagtttt	cttctctaac	tgtgaatgtg	1920
agaggctcag	gaatgagaat	actcgtgagg	ggtaactccc	ccgtgttcaa	ctacaacaag	1980
gcaacccaaa	ggcttacagt	cctcggaaaag	gacgcaggtg	cattaacaga	agatccagac	2040
gaggaacag	ccggggtgga	atctgcagta	ttgaggggat	tcctaattct	aggcagagag	2100

[0011]

gacaaaagat atggaccage attgagcatc aatgaactga gcaatcttgc aaaaggggag	2160
aaggetaatg tgctgatagg gcaaggagac gtggtgttgg taatgaaacg gaaacgggac	2220
tctageatac ttaactgacag tcagacagcg accaaaagaa ttcggatggc catcaattag	2280

<210> 6

<211> 2151

<212> DNA

<213>人工序列

<220>

<223>

<400> 6

atggaagact ttgtgcgaca atgcttcaat ccaatgattg tcgagcttgc ggaaaaggca	60
atgaaagaat atggggaaga tccgaagatc gaaacgaaca aatttgccgc aatatgcacg	120
caettagaag tctgtttcat gtatteagat ttccacttta ttgatgaacg gggcgaatca	180
acaattatag aatctggcga tcccaatgea ttattgaaac accggtttga aataatcgaa	240
gggagggacc gaacaatggc ctggacagtg gtgaatagta tctgcaacac cacaggagtt	300
gagaagccta aatttctccc agatttgiat gactacaaag agaaccgatt tattgaaatt	360
ggagtgcacac ggagggaaat tcacacatac tatctagaaa aagccaacaa gataaaatct	420
gagaagacac acattcaeat attctcattc actggagagg aaatggccac caaagcggac	480
tacacccttg atgaagaaag cagggcccga atcaaaacca ggctgttcac tataaggcag	540
gaaatggcca gtaggggttt atgggattcc tttcgtcagt ccgagagagg cgaagagaca	600
attgaagaaa gatttgaat cacaggaact atgtgcagge ttgccgacca aagtctccca	660
cctaatttct ccagccttga aaaatttaga gcctatgtgg atggattcga accgaacggc	720
tgcattgagg gcaagcttcc tcaaatgtca aaagaagtaa acgccagaat tgagccattt	780
ctgaagacaa caccacgccc tcttagatta cctgatgggc ctccctgctc tcagcggctg	840
aagttcttgc tgatggatgc ccttaaatta agcctcgaag acccgagtca tgagggggag	900
gggataaccgc tatatgatgc aatcaaatgc atgaaaacat ttttcggctg gaaagagccc	960
aacattgtaa aaccacatga aaaaggcata aacccecaatt acctcctggc ttggaagcag	1020
gtctgtggcag agctccaaga tattgaaaac gaggagaaaa ttccaaagac aaagaacatg	1080
aggaaaacaa gcccaattgaa gtgggcaett ggtgagaata tggcaccaga gaaagtagac	1140
tttgaggatt gcaaagatgt tagcgatcta aggcagtatg acagtgatga accaaagcct	1200
agatcactag caagctggat ccagagtgaa ttcaacaagg catgcgaatt gacagattca	1260
agttggattg aacttgatga aataggggaa gacgttgcct caattgagca cattgcaagt	1320
atgagaagga actatttcaac agcggaaagta tcccattgca gggetactga atacataatg	1380
aaggagtggt acataaacac agctttgttg aatgcatcct gtgcggccat ggatgacttc	1440
caactgattc caatgataag caaatgcaga accaaagaag ggagacggaa aactaacctg	1500
tatggattca ttataaaagg aagateccat ttgagaaatg acaccgatgt ggtaaacctt	1560
gtgagtatgg aattctctct tactgatecg aggetggagc cacacaagtg ggaaaagtac	1620
tgcgttctcg agataggaga catgctccta egaactgcaa taggccaagt gteaaggccc	1680
atgtttcttt atgtgagaac caatggaacc tecaagatca agatgaaatg gggcatggaa	1740

[0012]

atgaggcgat	gccttcttca	atcccttcaa	cagattgaga	gcatgattga	ggcagagtct	1800
tctgtcaaag	aaaaagacat	gactaaagaa	ttctttgaaa	acaaatcaga	aacatggcca	1860
attggagaat	cacccaaagg	gatggaggaa	ggctccatcg	ggaaggtgtg	cagaacctta	1920
ctggctaaat	ctgttttcaa	cagtctatat	gcctctccac	aactcgaggg	gtttteagct	1980
gaatcaagaa	aattgcttct	cattgttcag	gcacttaggg	acaacctgga	acctgggacc	2040
ttegatcttg	ggggctata	tgaagcaatt	gaggagtgcc	tgattaatga	tccttgggtt	2100
ttgettaatg	catcttggtt	caactccttc	ctcacacatg	cactaaaata	g	2151

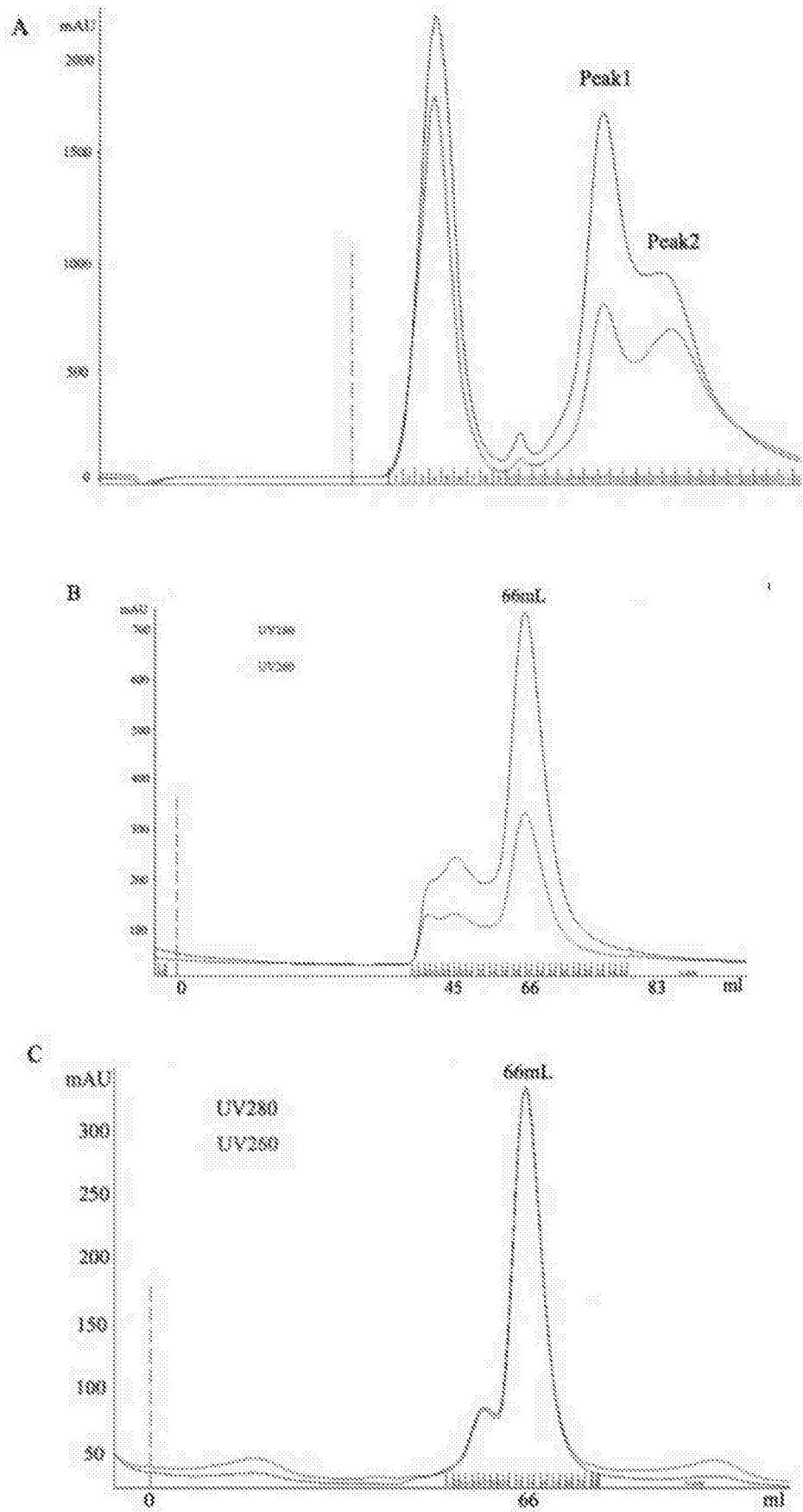


图 1

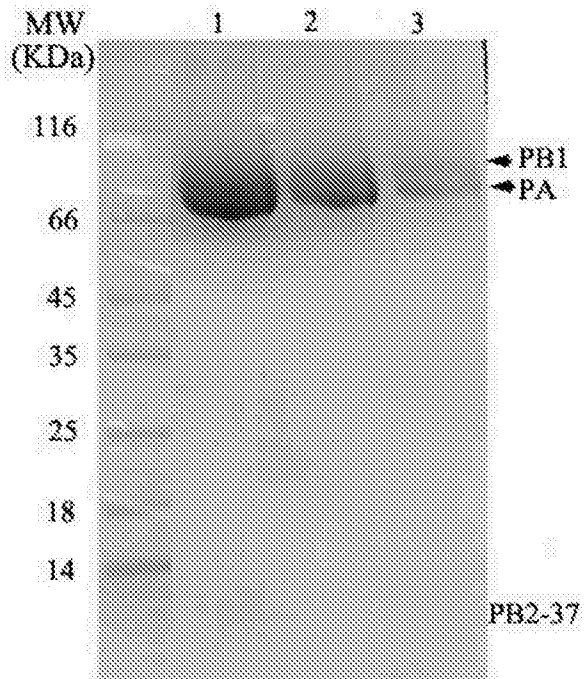


图 2

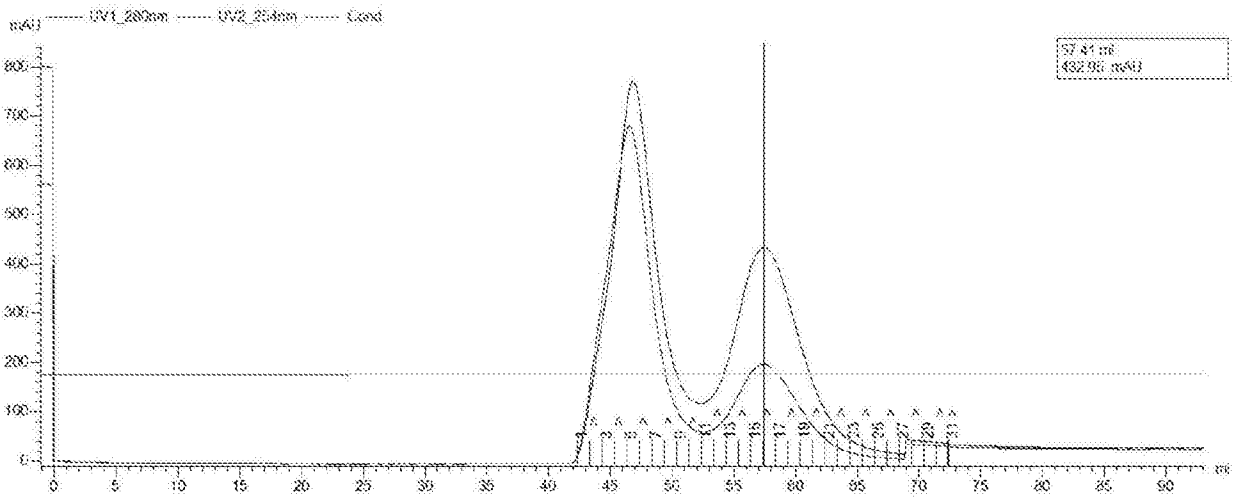


图 3

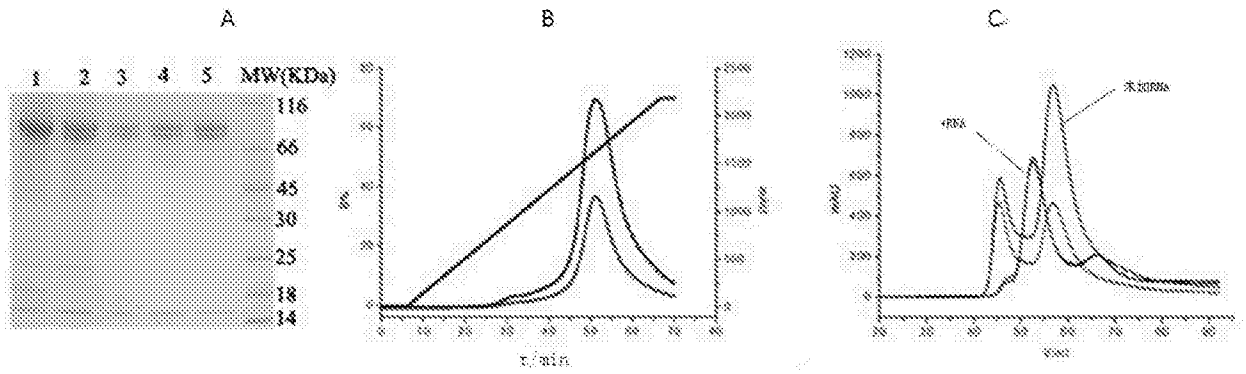


图 4

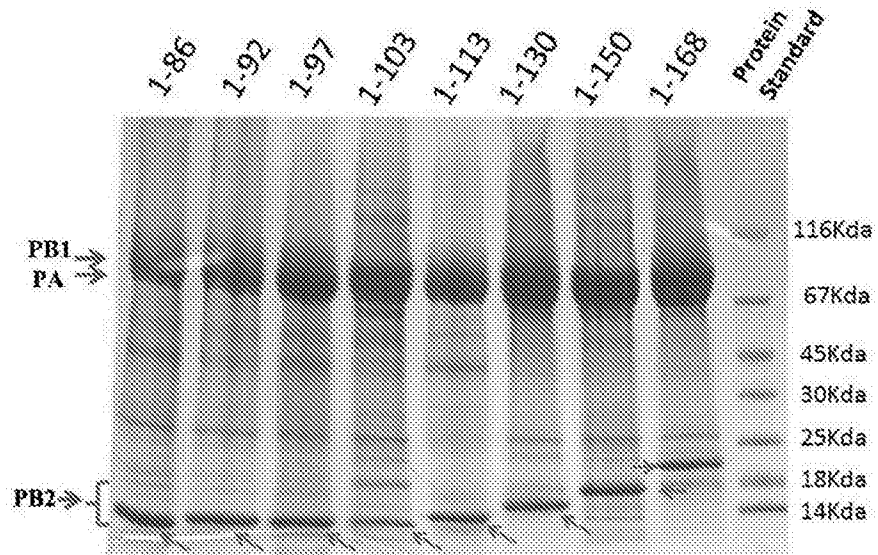


图 5

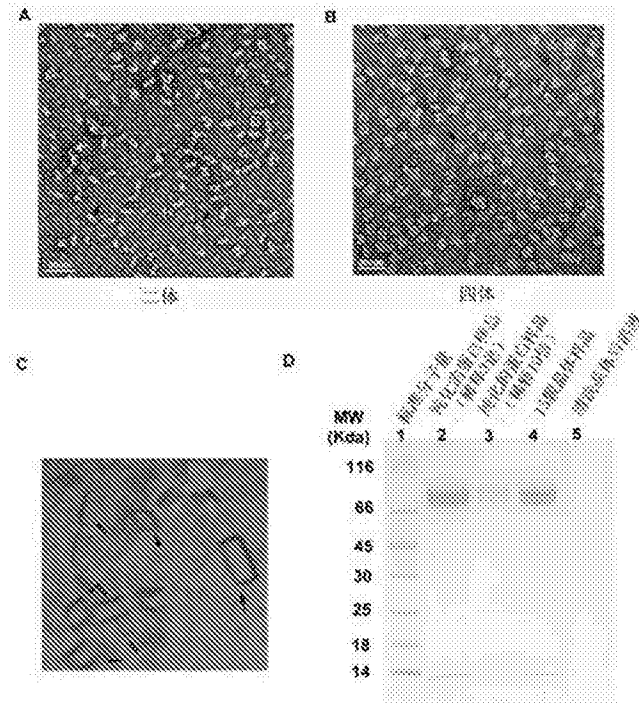


图 6

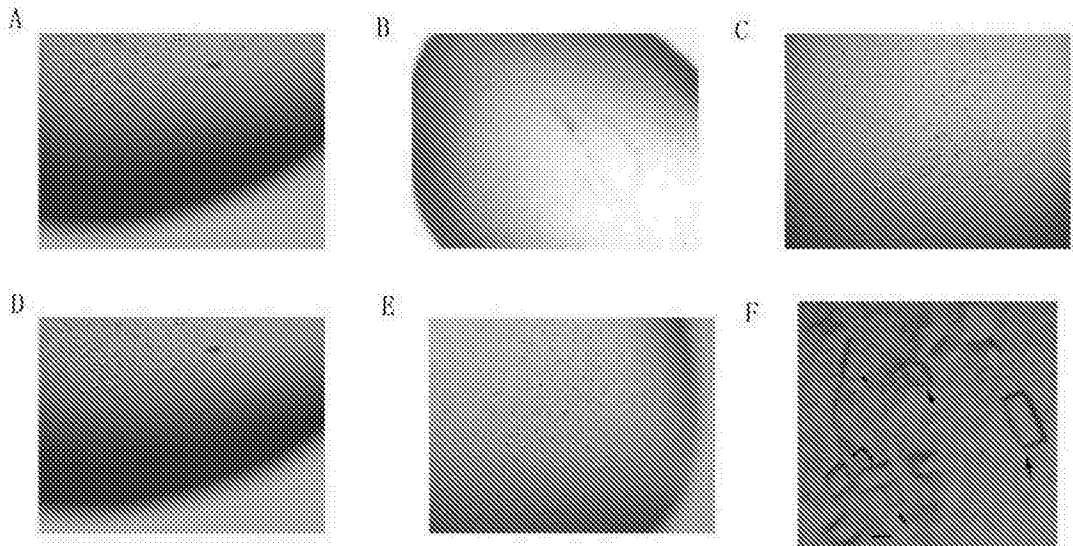


图 7