



(12) 发明专利申请

(10) 申请公布号 CN 102274208 A

(43) 申请公布日 2011.12.14

(21) 申请号 201110112717.1

(22) 申请日 2011.05.03

(71) 申请人 中国科学院化学研究所

地址 100080 北京市海淀区中关村北一街 2 号

申请人 中国科学院生物物理研究所

(72) 发明人 李林 常圣海 向俊锋 梁欢欢

李骞 唐亚林 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.

A61K 31/216(2006.01)

A61P 31/16(2006.01)

权利要求书 1 页 说明书 11 页

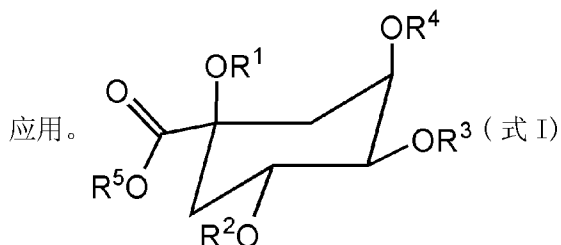
序列表 6 页 附图 9 页

(54) 发明名称

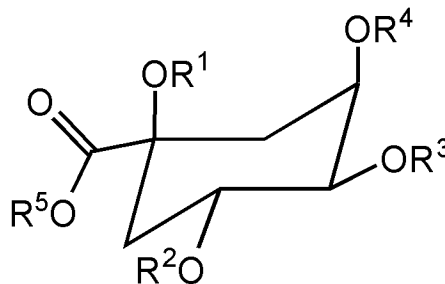
咖啡酰奎宁酸类化合物的新用途

(57) 摘要

本发明公开了咖啡酰奎宁酸类化合物的新用途。本发明提供的式 I 所示咖啡酰奎宁酸类化合物的新用途是一方面是其作为流感病毒 RNA 聚合酶抑制剂的应用；另一方面是其在制备预防和/或治疗与流感病毒 RNA 聚合酶相关的疾病的药物中的应用，特别是在制备抗流感药物中的应用，更特别是在制备抗禽流感病毒 (H5N1) 的药物中的



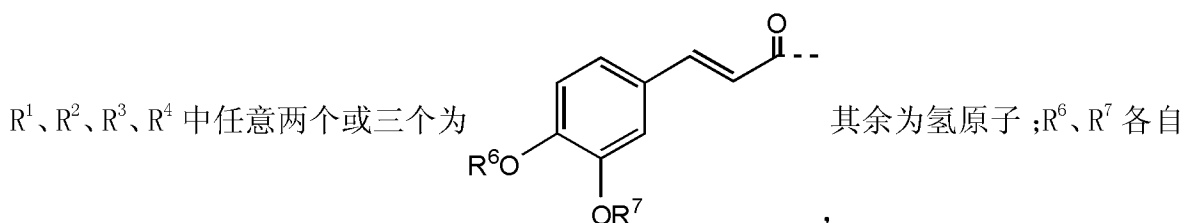
1. 式 I 所示化合物或其药学上可接受的盐作为流感病毒 RNA 聚合酶抑制剂的应用；



(式 I)

其中：

R^5 为氢原子或 C_1-C_4 烷基；



独立地为氢原子或 C_1-C_4 烷基。

2. 根据权利要求 1 所述的应用，其特征在于：所述式 I 中的 R^5 为氢原子、甲基或乙基。

3. 根据权利要求 1 或 2 所述的应用，其特征在于：所述式 I 中的 R^6, R^7 各自独立地为氢原子、甲基或乙基。

4. 根据权利要求 1-3 中任一项所述的应用，其特征在于：所述式 I 所示化合物选自下述任意一种：1,3,4-三咖啡酰奎宁酸、1,3,5-三咖啡酰奎宁酸、1,4,5-三咖啡酰奎宁酸、3,4,5-三咖啡酰奎宁酸、1,5-二咖啡酰奎宁酸、1,3-二咖啡酰奎宁酸、1,4-二咖啡酰奎宁酸、3,5-二咖啡酰奎宁酸、3,4-二咖啡酰奎宁酸和 4,5-二咖啡酰奎宁酸；优选为 3,5-二咖啡酰奎宁酸、3,4-二咖啡酰奎宁酸或 4,5-二咖啡酰奎宁酸。

5. 式 I 所示化合物或其药学上可接受的盐在制备预防和 / 或治疗与流感病毒 RNA 聚合酶相关疾病药物中的应用。

6. 式 I 所示化合物或其药学上可接受的盐在制备抗流感病毒药物中的应用。

7. 根据权利要求 3 所述的应用，其特征在于：所述流感病毒为禽流感病毒，优选 H5N1 型禽流感病毒。

8. 根据权利要求 5-7 中任一项所述的应用，其特征在于：所述式 I 中的 R^5 为氢原子、甲基或乙基。

9. 根据权利要求 5-8 中任一项所述的应用，其特征在于：所述式 I 中的 R^6, R^7 各自独立地为氢原子、甲基或乙基。

10. 根据权利要求 5-9 中任一项所述的应用，其特征在于：所述式 I 式 I 所示化合物选自下述任意一种：1,3,4-三咖啡酰奎宁酸、1,3,5-三咖啡酰奎宁酸、1,4,5-三咖啡酰奎宁酸、3,4,5-三咖啡酰奎宁酸、1,5-二咖啡酰奎宁酸、1,3-二咖啡酰奎宁酸、1,4-二咖啡酰奎宁酸、3,5-二咖啡酰奎宁酸、3,4-二咖啡酰奎宁酸和 4,5-二咖啡酰奎宁酸；优选为 3,5-二咖啡酰奎宁酸、3,4-二咖啡酰奎宁酸或 4,5-二咖啡酰奎宁酸。

咖啡酰奎宁酸类化合物的新用途

技术领域

[0001] 本发明涉及咖啡酰奎宁酸类化合物的新用途。

背景技术

[0002] 来源于鸟类的高致病性禽流感病毒 (H5N1) 严重威胁着人类健康。自从 1997 年香港报告第一例鸡传播给人的禽流感病例以来, 该病毒在东南亚地区的规模爆发, 并迅速扩散到许多国家和地区。禽流感病毒复制迅速, 经常变异, 能从鸟类传播给人类, 使感染者患上严重的呼吸系统疾病并且具有很高的致病率和死亡率。

[0003] 抗禽流感药物依据其作用靶点主要分为三类:

[0004] (1) 神经氨酸苷酶抑制剂

[0005] 在流感病毒与携有 N-乙酰神经氨酸 (NANA, 唾液酸) 的受体结合的过程中, 病毒红血球凝聚素 (H) 是必需的。然后, 病毒神经氨酸苷酶 (N) 从细胞表面糖蛋白上切下 N-乙酰神经氨酸, 使得新生成的病毒微粒得以离开感染细胞并扩散到其它宿主细胞。所以, 神经氨酸苷酶抑制剂可以通过阻碍病毒微粒的释放来阻止病毒的扩散。

[0006] 在神经氨酸苷酶抑制剂中, 扎那米韦 (zanamivir) 和奥司他韦 (oseltamivir, 达菲®) 是最著名的预防和治疗季节性流感的药物。两种药物对于流感病毒神经氨酸苷酶都具有很强的抑制效果 ($IC_{50} \leq 1\text{ng/ml}$), 在体外和体内实验中能抑制流感病毒的复制, 在临床上对于人感染 A 型和 B 型流感具有很好的预防和治疗效果。

[0007] 然而, 流感病毒对于扎那米韦和奥司他韦的抗药性是不可忽视的问题。在接受奥司他韦给药的病人中, 出现了抗奥司他韦的流感病毒, 还出现了抗奥司他韦的 B 型流感病毒在人和人之间传播。此外, 抗奥司他韦的 H1N1 病毒出现于 2007-2008 年, 而现今大部分的 H1N1 病毒都对奥司他韦具有抗药性。奥司他韦在用于治疗禽流感 H5N1 时, 同样有报道出现抗药性。研究表明, 病毒神经氨酸苷酶在与抑制剂作用位点的变异导致了神经氨酸苷酶抑制剂的抗药性。

[0008] (2) M2 离子通道拮抗剂

[0009] 金刚烷衍生物金刚胺和金刚乙胺对于 A 型流感病毒是有效的。它们通过直接与 M2 蛋白作用来干扰病毒去皮 (uncoating) 过程。但是, 对于人流感 (H1N1 和 H3N2) 和禽流感 (H5N1) 病毒, 该类化合物会产生抗药性, 因此其应用受到了很大的限制。

[0010] (3) 流感病毒 RNA 聚合酶抑制剂 (RNA polymerase inhibitors)

[0011] 流感病毒 RNA 聚合酶是由三个含病毒编码的蛋白 (PA, PB1 和 PB2) 组成的复合物。除了参与 RNA 复制以外, 还与核酸内切酶活性有关, 以保证促使转录和随后的翻译步骤的“戴帽机制”(cap snatching)。流感病毒 RNA 聚合酶复合物与禽流感病毒 H5N1 的高危害性有很大关联。基于此, 靶向流感病毒 RNA 聚合酶的新型抗病毒化合物对于开发预防和治疗人或禽流感病毒感染的药物尤为重要。

[0012] T-705 (Favipiravir, 6-fluoro-3-hydroxy-2-pyrazinecarboxamide) 是迄今为止最有代表意义的通过抑制流感病毒 RNA 聚合酶活性而达到抗流感病毒效果的化合

物。研究表明, T-705 对于 A、B 和 C 型流感病毒在体外和体内实验中都有很强的抑制作用。以 A 型流感病毒为例, 对于 H1N1, $IC_{50} \leq 0.2 \mu\text{g/ml}$; 对于 H2N2, $IC_{50} \leq 0.3 \mu\text{g/ml}$; 对于 H3N2, $IC_{50} \leq 0.48 \mu\text{g/ml}$ 。此外, 对于神经氨酸酶抑制剂 (奥司他韦) 和金刚胺类化合物具有抗药性的流感病毒, T-705 也有很好的活性 (Furuta, Y. et al. In vitro and in vivo activities of antiinfluenza virus compound T-705. Antimicrob. Agents Chemother. 46, 977-981 (2002))。对于 T-705 抗流感病毒的作用机理研究表明, 虽然 T-705 本身在小于 $100 \mu\text{M}$ 时对流感病毒 RNA 聚合酶活性不具有抑制作用, 但是其在感染细胞内的代谢产物 T-705RTP (T-705-4-呋喃核糖-5'-三磷酸盐, T-705-4-ribofuranosyl-5'-triphosphate) 却被证实是有很强的抑制流感病毒 RNA 聚合酶活性 ($IC_{50} = 0.14 \mu\text{M}$) 的作用。而且, 最近的研究表明, T-705 对于高致病性禽流感病毒 H5N1, 包括抗奥司他韦的 H5N1 病毒, 具有很好的 (大鼠) 体内活性 (Maki Kiso et al., T-705 (favipiravir) activity against lethal H5N1 influenza A viruses, PNAS 2010, 107, 882-887)。

[0013] 基于以上研究, 寻找新一代的抗流感化合物对于预防和治疗潜在的流感大规模流行尤为关键。以 T-705 为代表的流感病毒 RNA 聚合酶抑制剂已经被证实对于多种流感病毒有抑制作用, 其对于抗金刚胺和抗奥司他韦的流感病毒也具有很好的活性。我们有理由相信, 流感病毒 RNA 聚合酶抑制剂很有希望成为新一代的抗流感药物。

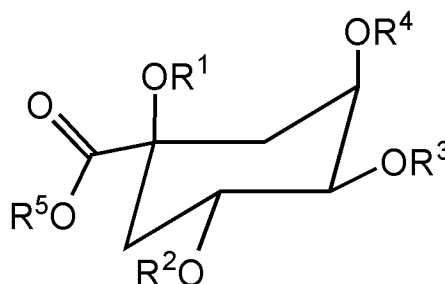
发明内容

[0014] 本发明的目的是提供咖啡酰奎宁酸类化合物的新用途。

[0015] 本发明所提供的咖啡酰奎宁酸类化合物的新用途之一是其作为流感病毒 RNA 聚合酶抑制剂的应用。

[0016] 本发明中所述咖啡酰奎宁酸类化合物其结构式如式 I 所示:

[0017]

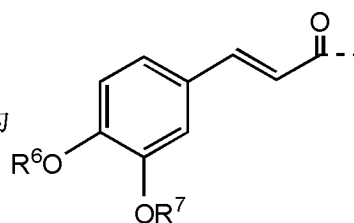


[0018] (式 I)

[0019] 其中:

[0020] R^5 为氢原子或 C_1-C_4 烷基;

[0021] R^1, R^2, R^3, R^4 中任意两个或三个为



其余为氢原子;

[0022] R^6, R^7 各自独立地为氢原子或 C_1-C_4 烷基。

[0023] 所述式 I 中的 R^5 可进一步优选为氢原子、甲基或乙基。

[0024] 所述式 I 中的 R^6 、 R^7 可各自独立地优选为氢原子、甲基或乙基；如 R^6 和 R^7 为氢原子，或 R^6 为氢原子、 R^7 为甲基，或 R^6 为甲基、 R^7 为氢原子。

[0025] 所述式 I 所示化合物更具体地选自下述化合物中的任意一种：1,3,4-三咖啡酰奎宁酸、1,3,5-三咖啡酰奎宁酸、1,4,5-三咖啡酰奎宁酸、3,4,5-三咖啡酰奎宁酸、1,5-二咖啡酰奎宁酸、1,3-二咖啡酰奎宁酸、1,4-二咖啡酰奎宁酸、3,5-二咖啡酰奎宁酸、3,4-二咖啡酰奎宁酸和 4,5-二咖啡酰奎宁酸；优选 3,5-二咖啡酰奎宁酸、3,4-二咖啡酰奎宁酸或 4,5-二咖啡酰奎宁酸。

[0026] 上述咖啡酰奎宁酸类化合物药学上可接受的盐作为流感病毒 RNA 聚合酶抑制剂的应用也属于本发明的保护范围。

[0027] 本发明还提供了式 I 所示咖啡酰奎宁酸类化合物或其药学上可接受的盐在制备预防和 / 或治疗与流感病毒 RNA 聚合酶相关疾病药物中的应用。

[0028] 本发明进一步提供了式 I 所示咖啡酰奎宁酸类化合物或其药学上可接受的盐在制备抗流感病毒药物中的应用，尤其是在制备抗禽流感病毒 (H5N1) 药物中的应用。

[0029] 此外，以式 I 所示咖啡酰奎宁酸类化合物或其药学上可接受的盐为活性成分与药学上可接受的赋形剂组成的组合物，也属于本发明的保护范围。

[0030] 上述组合物可用于制备预防和 / 或治疗与流感病毒 RNA 聚合酶相关疾病药物；更具体的可用于制备抗流感病毒药物，尤其用于制备抗禽流感病毒 (H5N1) 药物。

[0031] 基于本领域技术人员公知的常规技术，本发明的药物组合物可以通过本领域常用的药学上可接受的赋形剂制剂，最终提供多种单剂量和多剂量形式。剂型的非限定性的例子包括但不限于，口服制剂例如粉末、颗粒、片剂、胶囊、悬液、乳液、糖浆和气雾剂 (aerosol)，注射用制剂例如溶液、悬液、乳液和冻干粉针等。

[0032] 对本领域技术人员来说，本发明中的咖啡酰奎宁酸类化合物还包括通过本领域技术人员惯用的化学方法取代得到的上述具体化合物的衍生物。以上述具体化合物为母核，可以使用本领域公知的取代基通过化学方法获得衍生物。例如，可以预期那些通过连接羟基、卤素、硝基或碳 1-3 的烷基到母核来构造出的衍生物很可能具有与母核相同或相似的功能。这些取代的衍生物也落入本发明的范围。

[0033] 本发明以流感病毒 RNA 聚合酶为靶点，通过核磁共振技术寻找流感病毒 RNA 聚合酶抑制剂。结果发现咖啡酰奎宁酸类化合物具有流感病毒 RNA 聚合酶活性，并进一步通过分子学实验得以证实。这些抑制剂可以用于制备与流感病毒 RNA 聚合酶相关的疾病的药物，特别是抗流感的药物，更特别是抗禽流感病毒 (H5N1) 的药物。

附图说明

[0034] 图 1 为流感病毒 RNA 聚合酶蛋白 PA_C 亚基的结构。

[0035] 图 2 为流感病毒 RNA 聚合酶蛋白 PA_C 亚基的纯化鉴定结果。

[0036] 图 3 为 1,5-二咖啡酰奎宁酸在加入 PA_C 前 (a, b) 后 (c, d) 的弛豫编辑谱，其中，1,5-二咖啡酰奎宁酸和 PA_C 蛋白的浓度分别是 $1.0 \times 10^{-3} \text{ mol/L}$ 和 $4.4 \times 10^{-6} \text{ mol/L}$ 。

[0037] 图 4 为以小分子横向弛豫拟合法对 1,5-二咖啡酰奎宁酸同 PA_C 相互作用亲和力的拟合，其中，a) $T_{2\text{obs}}$ 对 C_p/C_L 作图，b) $R_{2\text{obs}}$ 对 C_p/C_L 作图， $R_{2\text{obs}} = 1/T_{2\text{obs}}$ 。

[0038] 图 5 为以小分子横向弛豫拟合法对 1,3-二咖啡酰奎宁酸同 PA_C 相互作用亲和力

的拟合,其中,a) $T_{2\text{obs}}$ 对 C_p/C_L 作图,b) $R_{2\text{obs}}$ 对 C_p/C_L 作图, $R_{2\text{obs}} = 1/T_{2\text{obs}}$ 。

[0039] 图 6 为以小分子横向弛豫拟合法对 3,5-二咖啡酰奎宁酸同 PA_C 相互作用亲和力的拟合,其中,a) $T_{2\text{obs}}$ 对 C_p/C_L 作图,b) $R_{2\text{obs}}$ 对 C_p/C_L 作图, $R_{2\text{obs}} = 1/T_{2\text{obs}}$ 。

[0040] 图 7 为以小分子横向弛豫拟合法对 3,4-二咖啡酰奎宁酸同 PA_C 相互作用亲和力的拟合,其中,a) $T_{2\text{obs}}$ 对 C_p/C_L 作图,b) $R_{2\text{obs}}$ 对 C_p/C_L 作图, $R_{2\text{obs}} = 1/T_{2\text{obs}}$ 。

[0041] 图 8 为以小分子横向弛豫拟合法对 4,5-二咖啡酰奎宁酸同 PA_C 相互作用亲和力的拟合,其中,a) $T_{2\text{obs}}$ 对 C_p/C_L 作图,b) $R_{2\text{obs}}$ 对 C_p/C_L 作图, $R_{2\text{obs}} = 1/T_{2\text{obs}}$ 。

[0042] 图 9 为 3,4-二咖啡酰奎宁酸 (A)、1,5-二咖啡酰奎宁酸 (B)、4,5-二咖啡酰奎宁酸 (C) 以及 3,5-二咖啡酰奎宁酸 (D) 在 5mM 浓度下的 ApG 引物延长实验结果。阴性对照为无聚合酶、无小分子时获得的结果,阳性对照为有聚合酶、无小分子时获得的结果。图 (b) 为通过磷光成像 (phosphorimaging) 获得的图 (a) 的定量结果。图 (b) 中数据为两次独立实验的平均值,标准偏差标于图中。

具体实施方式

[0043] I. 流感病毒 RNA 聚合酶:

[0044] A 型流感病毒含有 8 个负链 RNA 基因片段,每个片段都与含有病毒编码的基于 RNA 的流感病毒 RNA 聚合酶复合物 (大约 250kDa) 相连。该复合物在病毒的复制和致病机理中负责很多最重要的任务,例如“戴帽机制” (cap-snatching),即切掉宿主细胞前信使 RNA (pre-messenger RNA) 以将其“帽子”用于病毒转录。因为流感病毒 RNA 聚合酶复合物在病毒复制周期中负责很多重要的步骤,使得其序列在 A, B, C 型流感中存在高度的保守性而不易发生突变。此外,它与人体蛋白没有关联,使得其能够成为抗流感化合物的新靶点 (E. Obayashi et. Al., The structural basis for an essential subunit interaction in influenza virus RNA polymerase, Nature 2008, 454, 1127-1132)。

[0045] 流感病毒 RNA 聚合酶复合物是一个异三聚体,包含 PA, PB1 和 PB2 蛋白。PB1 拥有流感病毒 RNA 聚合酶活性位点,而且与核酸内切酶活性有关,而 PB2 负责“帽结合” (cap binding)。PA 蛋白可以分为两个域:较小的氨基端 (N-terminal, PA_N , 1-256 残基,约 25kDa) 和较大的羧基端 (C-terminal, PA_C , 257-716 残基,约 55kDa)。 PA_N 是 PA 蛋白的重要功能区,与病毒复制过程中的许多步骤都有关,包括核酸内切酶和蛋白酶的活性,以及病毒 RNA 与互补 RNA 启动子的相互作用。而对 PA_C 的结构分析发现,在流感病毒 RNA 聚合酶复合物中, PB1 的氨基端与 PA_C 相连。很显然,破坏该连接就会破坏流感病毒 RNA 聚合酶复合物的结构因而抑制其功能,从而干扰病毒的复制 (Kawaguchi, A., Naito, T. & Nagata, K. Involvement of influenza virus PA subunit in assembly of functional RNA polymerase complexes. J. Virol. 79, 732-744 (2005))。基于以上研究, PA 蛋白,特别是 PA_C 与 PB1 作用的位点,能够成为抗流感化合物靶点。

[0046] 有文献报道,利用 X-射线晶体法解析了 PA_C 亚基的结构并且提出了多个潜在的抗流感化合物作用位点。如附图 1 中所示, PA 包括 13 个 α -螺旋,一个短的 3_{10} 螺旋 ($\eta 1$), 9 个 β -折叠和数个环链 (loop) 或者转角 (turn)。从外观上看, PA_C 亚基酷似一个“龙头”,而且可以分为两个亚区:一区“头部”,和二区“嘴部”。位于“头部”的 $\beta 1$ - $\beta 7$ 折叠构成了一个扭曲的平面,并且被 5 个 α -螺旋 ($\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 6$ 和 $\alpha 7$) 和 $\eta 1$ 螺旋包围。“嘴部”

包括 $\beta 8$ 、 $\beta 9$ 折叠,和 $\alpha 4$ 、 $\alpha 5$ 、 $\alpha 8$ - $\alpha 12$ 螺旋。其中, $\alpha 11$ 和 $\alpha 13$ 构成下颚,而 $\alpha 8$ 和 $\alpha 10$ 构成上颚(X. J.He et. Al., Crystal structure of the polymerase PAC-PB1N complex from an avian influenza H5N1 virus, Nature 454(2008)1123-1127)。

[0047] 通过对 PA_C 亚基的进一步分析可以发现,多个位置的氨基酸变异会造成流感病毒 RNA 聚合酶活性下降从而抑制流感病毒的复制,提示这些位置可能成为抗流感化合物的作用靶点。例如:

[0048] (1) 一个半径约 25Å 的半环形深沟槽被认为是很好的 RNA 结合位点,分布在该沟槽表面的 K539 氨基酸变异 (K539A) 会明显破坏互补 RNA (cRNA) 和病毒 RNA (vRNA) 的合成 (Fodor, E. et al., A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. J. Virol. 76, 8989-9001 (2002))。

[0049] (2) E410, K461, E524 和 K536 残基分布于半径约 14Å 的椭圆形通道内表面。其中 E524A, E410A, K536A/W537A 双变异都会削弱流感病毒 RNA 聚合酶活性 (Regan, J. F., Liang, Y. & Parslow, T. G. Defective assembly of influenza A virus due to a mutation in the polymerase subunit PA. J. Virol. 80, 252-261 (2006))。

[0050] (3) PB1 结合在 PA_C 的疏水核心,包括 F411, M595, L666, W706, F710, V636 和 L640 等。W706A/Q670A、L666G/F710E、L666G/F710G 和 W706A/F710Q 双变异破坏 PB1 同 PA_C 的结合从而影响流感病毒 RNA 聚合酶活性 (X. J.He et. al., Crystal structure of the polymerase PAC-PB1N complex from an avian influenza H5N1 virus, Nature 454(2008)1123-1127)。

[0051] 此外,几乎所有这些氨基酸对于 A、B、C 三种类型的流感病毒都高度保守,提示基于这些靶点设计的抗流感化合物很有可能对 A、B、C 型流感都有效,而且具有较低的抗药性风险。

[0052] II. 核磁共振 (NMR) 技术

[0053] 核磁共振 (NMR) 技术是研究生物大分子与小分子之前相互作用的有效手段。很多 NMR 参数都可以用于判断大、小分子间的相互作用,其中灵敏度高且易于检测的参数更具有可用性。发明人采用以弛豫编辑 NMR 谱 (relaxation-edited NMR) 监测小分子横向弛豫 (transverse relaxation) 变化的方法来判断小分子同生物大分子的结合。

[0054] 在受体蛋白存在与否时对比配体质子的横向弛豫速率 (transverse relaxation rate, $R_2 = 1/T_2$) 的方法,是研究分子间相互作用最成熟、最经典的方法。 R_2 的大小正比于分子转动相关时间 τ_c ,而 τ_c 与分子尺寸直接相关。一般情况下,靶蛋白受体的 τ_c 要远大于自由态小分子的 ($\tau_{c, target} \gg \tau_{c, free-ligand}$),所以有 $R_{2, target} \gg R_{2, free-ligand}$ 。当小分子与靶蛋白结合时,结合态小分子的 R_2 表现为结合态复合物的 R_2 ,因而远大于自由态小分子的 R_2 ($R_{2, bound} \gg R_{2, free}$)。而在 NMR 方法中,所检测到的小分子 R_2 ($R_{2, obs}$) 是自由态与结合态 R_2 的加权平均 ($R_{2, obs} = X_{bound} R_{2, bound} + X_{free} R_{2, free}$),其中 X_{bound} 和 X_{free} 分别是结合态和自由态小分子所占比例。据此,在小分子配体与靶蛋白受体结合时,其 $R_{2, obs}$ 自然要远大于受体蛋白不存在时的 $R_{2, free}$, R_2 因而成为判断大、小分子间结合的有效“探针”。

[0055] 观察 R_2 变化的办法是弛豫编辑 NMR (relaxation-edited NMR),或者称为 T_2 滤波 NMR (T_2 -filtered NMR)。该实验可以采用 Carr-Purcell-Meiboom-Gill (CPMG) 自旋

锁定序列 (H. Y. Carr et. al., Rev. Sci. Instrum. 1958, 29, 688 ;Deverell, C. et. al., Mol. Phys. 1970, 18, 553) 来完成, 其实验脉冲可以简单地描述为 $[D-90_x-(\Delta-180_y-\Delta)_n-acquire]$ 。其中自旋锁定时间为 $2 \times n \times \Delta$, 当该时间大于或者等于某组分的 R_2 时, 该组分的磁化矢量就会在自旋锁定时间内完全弛豫, 因而在最后的采样过程中不会出现其共振峰。显然, CPMG 脉冲可以削弱甚至消除横向弛豫较快 (R_2 较大) 组分的谱峰, 而几乎不影响横向弛豫慢的组分。在判定小分子配体同靶蛋白受体的相互作用时, 通常分别检测小分子在加入蛋白前后的弛豫编辑 NMR 谱, 如果小分子和蛋白发生相互作用, 那么其 R_2 就会受到蛋白的影响而变大, 因而在弛豫编辑 NMR 谱中, 在相同的自旋锁定时间下, 小分子谱峰在蛋白存在时会发生明显的衰减, 甚至被完全消除。该方法还可以用于确定靶蛋白的配体小分子, 在蛋白存在的情况下, 配体小分子的谱峰会被选择性地压制, 而自由态小分子谱峰几乎不受影响。

[0056] 当然, 也可以通过直接测定小分子 R_2 (或者 T_2) 的方法来判定其和靶蛋白的相互作用。在不同浓度比的蛋白 / 小分子混合物中, 通过测定小分子的 R_2 (或者 T_2), 还可以计算相互作用的亲和力。

[0057] 如前所述, 把 PA_C 蛋白与小分子混合后, 如果小分子能与蛋白相互作用, 该小分子的横向弛豫速率 (R_2) 就会受到蛋白的影响而变快。本发明以流感病毒 RNA 聚合酶为靶点, 根据上述原理通过核磁共振技术寻找流感病毒 RNA 聚合酶抑制剂。

[0058] 结果发现了一系列流感病毒 RNA 聚合酶抑制剂, 并通过分子学实验证实了这些抑制剂对流感病毒 RNA 聚合酶的抑制作用。这些抑制剂可以用于制备与流感病毒 RNA 聚合酶相关的疾病的药物, 特别是抗流感的药物, 更特别是抗禽流感病毒 (H5N 1) 的药物。

[0059] 下面通过具体实施例对本发明做进一步详细地描述, 本领域技术人员可知, 这些实施例是为了更具体地阐述, 本发明的保护范围不受实施例的限制。

[0060] 下述实施例中所述实验方法, 如无特殊说明, 均为常规方法; 所述试剂和生物材料, 如无特殊说明, 均可从商业途径获得。

[0061] 实施例中所用的药物: 1, 5- 二咖啡酰奎宁酸、1, 3- 二咖啡酰奎宁酸、3, 5- 二咖啡酰奎宁酸、3, 4- 二咖啡酰奎宁酸、4, 5- 二咖啡酰奎宁酸均购于成都曼斯特生物科技有限公司。重水 (D_2O)、四氘代 3-(三甲基硅基) 丙酸钠 ($TSP-d_4$) 购于美国 Sigma 公司。

[0062] 实施例中所用的 PA_C 蛋白的表达与纯化方法如下: 根据 X. J. He 等在 Nature, 2008, 454, 1123-1127 中的记载制备 PA_C 蛋白。简要描述为: 禽流感病毒 H5N1 (A/Goose/Guangdong/1/96) RNA 流感病毒 RNA 聚合酶 PA 亚基的 257-716 残基的编码基因 (其核苷酸序列如序列 1 所示) 克隆于 pGEX-6p 载体 (GE Healthcare) (酶切位点 SmaI 和 Not I) 并单独转移到 Escherichia 大肠杆菌 BL21 株。重组细胞于 37°C 下在含有 100mg/L 氨苄西林的 LB 介质中培养。当 OD600 达到 0.6-0.8 时, 培养基在 16°C 下以 0.5mM 异丙基-β-D-半乳糖苷酶 (IPTG) 诱导。经过 20 小时的孵化, 表达 PA_C 的细胞被集中并通过离心 (10 分钟, 5000rpm) 收集。PBS 缓冲液重悬得到的细胞, 用高压破碎的方法裂解细胞, 37000g 离心去除沉淀, 取上清液; 得到的上清液以谷胱甘肽亲和柱 (GE Healthcare) 纯化, 10 倍柱体积的 PBS 缓冲液洗去杂蛋白, 1 倍柱体积的 PBS 缓冲液中加入 PPase (1 单位每 100 μg 融合蛋白) 酶切纯化的融合蛋白 16h, 2 倍柱体积 PBS 缓冲液洗下 PA_C , 20mM Tris 8.0 50mM NaCl 低盐透析, 浓缩到 10mg/ml, 并进一步用 Q 凝胶 FF 离子交换色谱 (HiTrap™ 5ml Q HP) (低

盐 20mM tris8.0 50mM NaCl,高盐 20mM tris8.0 1MNaCl) 和 Superdex-200 凝胶过滤色谱 (20mM tris8.0 100mM NaCl) 纯化。纯化后的 PAc 进行 SDS-PAGE 分析,如图 2 所示,大小在 45KD 到 66KD 之间,PA_c 的大小为 57KD。用得到的 PA_c 进行结晶,结构解析。

[0063] H5N1 病毒 RNA 流感病毒 RNA 聚合酶的制备参考文献 Journal of Virology,2002, 13,8989-9001 报道的方法,简要描述为:流感病毒 RNA 聚合酶 (PA、PB1 和 PB2) 复合物以哺乳动物 293T 细胞表达。首先通过基因克隆技术,将含有流感病毒毒株中 PA(其核苷酸序列如序列 2 所示),PB1(其核苷酸序列如序列 3 所示),PB2(其核苷酸序列如序列 4 所示)的负链基因组序列连接至 pPOLI 载体中,分别产生 pPOLI-PA-RT, pPOLI-PB1-RT, pPOLI-PB2-RT。经 NgoMIV 核酸内切酶作用后,获得含有 PA-, PB1-, PB2- 正义链信息的 cDNA,将这些 cDNA 通过 AgeI 位点连接至 pCDNA3 载体 (Invitrogen) 上,形成含有编码链序列信息(而不是病毒负义 RNA 的基因组序列信息)pCDNA3-PA, pCDNA3-PB1, pCDNA3-PB2 载体。其次,在完成上述质粒的构建工作后,我们将 pCDNA3-PA, pCDNA3-PB1 以及 pCDNA3-PB2 以等比例混合共转染进稳定培养的 293T 细胞中。经过 48 小时的培养后,收集细胞,用移液器将细胞培养基弃去,并加入 1xPBS 溶液,轻轻摇晃后吸去上清,以彻底去除残留的培养基。然后,按照 1ml/3.5cm 培养皿的量加入 1xPBS 溶液,用移液器将细胞从培养皿上吹打下来,置于 1.5mlEP 管内低速离心 (3000rpmx3min),弃去上清。重复加入 1ml 1xPBS 溶液重悬细胞,并离心后弃去上清。接着,向细胞沉积物内加入 90 μ l lysis buffer (50mM HEPES pH8.0,200mMNaCl,1mM DTT,0.5% Nonidet P-40(乙基苯基聚乙二醇),1mM phenylmethylsulfonyl fluoride(PMSF,苯甲基磺酰氟),25%甘油,0.1mM β -mercaptoethanol(β -巯基乙醇),1tablet/10ml protease inhibitor),剧烈重悬细胞,并置冰上裂解 15min 以上,在裂解过程中可反复振荡细胞重悬液。裂解结束后,将 EP 管置于冷冻离心机内高速 (13000rpm) 离心 15min,吸取上清转移至一干净的 EP 管内,并置冰上 4 $^{\circ}$ C 保存。在完成上述步骤后,取出 IgG sepharose 6 Fast Flow(GE)beads,用水重悬清洗一次,2200rpm 离心 1min,弃去上清,用 binding buffer(10mM HEPES pH8.0,150mM NaCl, 1mM DTT,0.1% Nonidet P-40,0.2mM phenylmethylsulfonyl fluoride(PMSF),10%甘油)重悬 IgGsepharose beads,清洗两次,2200rpm 离心 1min 收集 beads。处理结束后,将 4 $^{\circ}$ C 保存的裂解液上清加入 IgG sepharose 6 Fast Flow beads 中,置 4 $^{\circ}$ C 振荡器上结合 2 小时。之后,取出结合完毕的 EP 管,置冷冻离心机内 2000rpm 离心 1min。弃去上清,加入 binding buffer,重悬洗两次后加入 TEV cleavage buffer(10mM HEPES pH8.0,150mM NaCl,,1mM DTT,0.1% Nonidet P-40,0.2mM phenylmethylsulfonyl fluoride(PMSF),10%甘油,1mM DL-Dithiothreitol) 重悬洗一次,弃净上清。加入 80 μ l TEV cleavage buffer,同时加如 1 μ l TEV protease(Invitrogen 公司),置于 16 $^{\circ}$ C 酶切 2 小时以上。最后,将酶切完成的 EP 管用 3000rpm 离心 2min,将上清转移至一干净的 EP 管内,完成聚合酶复合物的提纯工作。

[0064] 实施例 1、通过 NMR 实验测定咖啡酰奎宁酸类化合物同 PA_c 相互作用的解离常数 K_d

[0065] NMR 实验采用 Bruker AVANCE 600 谱仪(装配 5mm BBI 探头,提供 z 方向梯度)完成,实验操作及数据处理利用 TOPSPIN(Bruker, version 2.1) 工作站完成,所有 NMR 实验都在 298.2K 下进行(Bruker, B3000 温控单元)。弛豫编辑 NMR 谱采用 [D/presaturation-90_x-(Δ -180_y- Δ)_n-acquire] 脉冲序列,其中运用 Carr-Purcell-Meiboom-Gill(CPMG) 序列进行自旋锁定 (spin-lock),并采用了预饱和和脉

冲压制水峰。对于所有弛豫编辑实验, $D = 3s$, $P_{90}(90^\circ \text{ 脉冲})$ 根据不同样品分别设定, $\Delta = 1.5ms$, $2 \times n \times \Delta =$ 总自旋锁定时间, 自旋锁定循环数 n 根据不同实验分别设定, 采样点数 64k, 扫描次数 32。横向弛豫时间 (transverse relaxation time, T_2) 通过 CPMG 法测得, 采用 $[D-90_x-(\Delta-180_y-\Delta)_n-acquire]_m$ 脉冲序列。其中 $D = 3s$, P_{90} 根据不同样品分别设定, $\Delta = 1.5ms$, $2 \times n \times \Delta =$ 总自旋锁定时间, 自旋锁定循环数 $n = 0, 10, 20, 30, 40, 50, 80, 100, 150, 200, 250, 300, 350, 400, 500, 600$, F2 维采样点数 32k (td2, 真实采样维), F1 维采样点数 16 (td1 = m, 虚拟采样维), 扫描次数 16。 T_2 利用 TOPSPIN 工作站拟合谱峰衰减曲线得到。

[0066] 通过 R. S. Luo 等在 Spectrochim. ActaA 1999, 55, 1897-1901 中报道的方法, 通过小分子横向弛豫变化拟合蛋白大分子和小分子之间相互作用解离常数。概括如下:

[0067] 当配体 (L) 和蛋白 (P) 形成复合物 (LP) 时, 溶液中存在一个解离平衡 $LP = L+P$, 该平衡的解离常数 K_d 是描述配体和蛋白相互作用亲和力的关键参数, 且可以通过配体在蛋白存在时的横向弛豫时间 (T_2) 计算。横向弛豫速率 ($R_2 = 1/T_2$) 的测量值 (R_{2obs}) 是小分子在结合态 (R_{2B}) 和自由态 (R_{2F}) 的横向弛豫速率的加权平均值:

$$[0068] \quad R_{2obs} = (1-X_B)R_{2F} + X_B R_{2B}$$

[0069] 其中 X_B 是结合态配体占配体总量的摩尔分数。解离平衡的 K_d 即表达为:

$$[0070] \quad K_d = [L][P]/[LP]$$

[0071] 其中 [L], [P] 和 [LP] 分别是自由态配体, 自由态蛋白和结合态复合物的平衡浓度。则有:

$$[0072] \quad nC_p = [LP] + [P]$$

$$[0073] \quad C_L = [L] + [LP]$$

$$[0074] \quad X_B = [LP]/C_L$$

[0075] 其中 n 代表蛋白的结合位点数。基于这些关系式, R_{2obs} 和 C_p/C_L 的关系式就可以表达为:

$$[0076] \quad 2(R_{2obs} - R_{2F}) / (R_{2B} - R_{2F}) = 1 + (K_d/C_p) (C_p/C_L) + n(C_p/C_L) - \{ [1 + (K_d/C_p) (C_p/C_L) + n(C_p/C_L)]^2 - 4n(C_p/C_L) \}^{1/2}$$

[0077] 由于 C_p 和 C_L 已知, 就可以通过拟合 T_{2obs} 对 C_p/C_L 的曲线得到 T_{2F} 和 T_{2B} (外推曲线到 $C_p/C_L = 0$ 得到 T_{2F} , C_p/C_L 到无穷得到 T_{2B})。那么, 通过上述方程拟合 R_{2obs} 对 C_p/C_L 的曲线就可以得到解离常数 K_d 。小分子同 PA_C 的亲合力以 pK_d 值, 即结合态复合物解离常数的负对数, 表示。 pK_d 值越大表示配体同靶分子结合能力越强。

[0078] 为了证明小分子同 PA_C 之间的相互作用, 采用了弛豫编辑 NMR 实验。

[0079] 以 1,5-二咖啡酰奎宁酸为例, 结果如图 3 所示, 在 600ms 的自选锁定时间下, 1,5-二咖啡酰奎宁酸在未加入 PA_C 时谱峰有些许衰减 (图 3a, 3b), 而在加入 PA_C 后, 1,5-二咖啡酰奎宁酸的谱峰在相同自选锁定时间下完全衰减 (图 3c, 3d), 说明该化合物能与 PA_C 结合。

[0080] (1) 1,5-二咖啡酰奎宁酸同 PA_C 相互作用的解离常数 K_d 的测定

[0081] 利用横向弛豫拟合的方法得到 1,5-二咖啡酰奎宁酸同 PA_C 相互作用的解离常数 K_d 。

[0082] 制备一系列含有不同浓度比的 PA_C 和 1,5-二咖啡酰奎宁酸的混合物并利用翻转

恢复 NMR 法测定其横向弛豫时间 T_2 , 结果如表 1 所示:

[0083] 表 1

	$C_P(10^{-6}\text{mol/L})$	$C_L(10^{-3}\text{mol/L})$	C_P/C_L	$T_{2\text{obs}}(\text{S})$
	0	1.00	0	1.12
	0.545	1.00	0.545	0.561
[0084]	1.64	1.00	1.64	0.354
	2.18	1.00	2.18	0.274
	2.73	1.00	2.73	0.227
	3.27	1.00	3.27	0.177

[0085] C_P 为流感病毒 RNA 聚合酶蛋白 PA_C 的浓度

[0086] C_L 为 1,5-二咖啡酰奎宁酸的浓度

[0087] $T_{2\text{obs}}$ 为 1,5-二咖啡酰奎宁酸的 T_2 观测值。

[0088] 通过表 1 可见, 随着蛋白和配体浓度比 (C_P/C_L) 的增大, 配体的 T_2 观测值 ($T_{2\text{obs}}$) 变小并且符合指数衰减规律 (图 4a)。通过拟合 $R_{2\text{obs}}$ ($R_{2\text{obs}} = 1/T_{2\text{obs}}$) 对 C_P/C_L 的曲线 (图 4b), 就可以得到 $K_d = 1.91 \times 10^{-6}\text{M}$, $\text{p}K_d = 5.72$ 。

[0089] (2)、1,3-二咖啡酰奎宁酸同 PA_C 相互作用的解离常数 K_d 的测定

[0090] 制备一系列含有不同浓度比的 PA_C 和 1,3-二咖啡酰奎宁酸的混合物并利用翻转恢复 NMR 法测定其横向弛豫时间 T_2 , 结果如表 2 所示:

[0091] 表 2

	$C_P(10^{-6}\text{mol/L})$	$C_L(10^{-3}\text{mol/L})$	C_P/C_L	$T_{2\text{obs}}(\text{S})$
	0	1.00	0	1.28
	1.91	1.00	1.91	0.774
	2.73	1.00	2.73	0.618
[0092]	3.82	1.00	3.82	0.474
	4.36	1.00	4.36	0.441
	4.91	1.00	4.91	0.399
	5.45	1.00	5.45	0.375

[0093] C_P 为流感病毒 RNA 聚合酶蛋白 PA_C 的浓度

[0094] C_L 为 1,3-二咖啡酰奎宁酸的浓度

[0095] $T_{2\text{obs}}$ 为 1,3-二咖啡酰奎宁酸的 T_2 观测值。

[0096] 通过表 2 可见, 随着蛋白和配体浓度比 (C_P/C_L) 的增大, 配体的 T_2 观测值 ($T_{2\text{obs}}$) 变小并且符合指数衰减规律 (图 5a)。通过拟合 $R_{2\text{obs}}$ ($R_{2\text{obs}} = 1/T_{2\text{obs}}$) 对 C_P/C_L 的曲线 (图 5b), 就可以得到 $K_d = 1.05 \times 10^{-4}\text{M}$, $\text{p}K_d = 3.98$ 。

[0097] (3) 3,5-二咖啡酰奎宁酸同 PA_C 相互作用的解离常数 K_d 的测定

[0098] 制备一系列含有不同浓度比的 PA_C 和 3,5-二咖啡酰奎宁酸的混合物并利用翻转

恢复 NMR 法测定其横向弛豫时间 T_2 , 结果如表 3 所示:

[0099] 表 3

	$C_P(10^{-6}\text{mol/L})$	$C_L(10^{-3}\text{mol/L})$	C_P/C_L	$T_{2\text{obs}}(\text{S})$
	0	0.998	0	1.17
	1.74	0.998	1.75	0.720
[0100]	2.62	0.998	2.62	0.588
	3.49	0.998	3.50	0.507
	4.36	0.998	4.37	0.444
	5.24	0.998	5.24	0.381
	6.98	0.998	6.99	0.330

[0101] C_P 为流感病毒 RNA 聚合酶蛋白 PA_C 的浓度

[0102] C_L 为 3,5-二咖啡酰奎宁酸的浓度

[0103] $T_{2\text{obs}}$ 为 3,5-二咖啡酰奎宁酸的 T_2 观测值

[0104] 通过表 3 可见, 随着蛋白和配体浓度比 (C_P/C_L) 的增大, 配体的 T_2 观测值 ($T_{2\text{obs}}$) 变小并且符合指数衰减规律 (图 6a)。通过拟合 $R_{2\text{obs}}$ ($R_{2\text{obs}} = 1/T_{2\text{obs}}$) 对 C_P/C_L 的曲线 (图 6b), 就可以得到 $K_d = 1.29 \times 10^{-5}\text{M}$, $pK_d = 4.89$ 。

[0105] (4) 3,4-二咖啡酰奎宁酸同 PA_C 相互作用的解离常数 K_d 的测定

[0106] 制备一系列含有不同浓度比的 PA_C 和 3,4-二咖啡酰奎宁酸的混合物并利用翻转恢复 NMR 法测定其横向弛豫时间 T_2 , 结果如表 4 所示:

[0107] 表 4

	$C_P(10^{-6}\text{mol/L})$	$C_L(10^{-3}\text{mol/L})$	C_P/C_L	$T_{2\text{obs}}(\text{S})$
	0	1.00	0	0.870
	1.75	1.00	1.75	0.483
[0108]	2.62	1.00	2.62	0.417
	3.49	1.00	3.49	0.330
	4.36	1.00	4.36	0.287
	5.24	1.00	5.24	0.251
	6.98	1.00	6.98	0.208

[0109] C_P 为流感病毒 RNA 聚合酶蛋白 PA_C 的浓度

[0110] C_L 为 3,4-二咖啡酰奎宁酸的浓度

[0111] $T_{2\text{obs}}$ 为 3,4-二咖啡酰奎宁酸的 T_2 观测值

[0112] 通过表 4 可见, 随着蛋白和配体浓度比 (C_P/C_L) 的增大, 配体的 T_2 观测值 ($T_{2\text{obs}}$) 变小并且符合指数衰减规律 (图 7a)。通过拟合 $R_{2\text{obs}}$ ($R_{2\text{obs}} = 1/T_{2\text{obs}}$) 对 C_P/C_L 的曲线 (图 7b), 就可以得到 $K_d = 1.58 \times 10^{-6}\text{M}$, $pK_d = 5.80$ 。

[0113] (5) 4,5-二咖啡酰奎宁酸同 PA_C 相互作用的解离常数 K_d 的测定

[0114] 制备一系列含有不同浓度比的 PA_C 和 4,5-二咖啡酰奎宁酸的混合物并利用翻转恢复 NMR 法测定其横向弛豫时间 T₂, 结果如表 5 所示:

[0115] 表 5

	C _P (10 ⁻⁶ mol/L)	C _L (10 ⁻³ mol/L)	C _P /C _L	T _{2obs} (S)
	0	1.07	0	0.867
[0116]	1.75	1.07	1.63	0.390
	2.18	1.07	2.03	0.324
	2.62	1.07	2.44	0.297
	3.49	1.07	3.26	0.238
[0117]	4.36	1.07	4.07	0.205
	6.98	1.07	6.51	0.166

[0118] C_P 为流感病毒 RNA 聚合酶蛋白 PA_C 的浓度

[0119] C_L 为 4,5-二咖啡酰奎宁酸的浓度

[0120] T_{2obs} 为 4,5-二咖啡酰奎宁酸的 T₂ 观测值

[0121] 通过表 5 可见,随着蛋白和配体浓度比 (C_P/C_L) 的增大,配体的 T₂ 观测值 (T_{2obs}) 变小并且符合指数衰减规律 (图 8a)。通过拟合 R_{2obs} (R_{2obs} = 1/T_{2obs}) 对 C_P/C_L 的曲线 (图 8b), 就可以得到 K_d = 8.91 × 10⁻⁶M, pK_d = 5.05。

[0122] 实施例 2、RNA 流感病毒聚合酶活性测试

[0123] 采用 T. Deng 等在 J. Virol. 80 (2006) 2337-2348 中报道的 ApG 引物延长法测试 RNA 流感病毒聚合酶活性。具体方法如下:在 5mM MgCl₂, 5mM 小分子 (咖啡酰奎宁酸类化合物), 5mM 二硫苏糖醇, 1mM 三磷酸腺苷 (ATP), 0.5mM 三磷酸尿苷 (UTP), 三磷酸胞苷 0.5mM (CTP), 0.1 μM [α-³²P] 三磷酸鸟苷 (GTP 3000Ci/mmol), 和 2U/μl RNasin (Promega 公司) 存在时, 同 2.5 μL H5N1 禽病毒 RNA 聚合酶 (3P) 和 0.7 μM 模版病毒 RNA 启动子 (等摩尔 5' -end vRNA 5' -AGUAGAAACAAGGCC-3' 和 3' -end vRNA 5' -GGCCUGCUUUUGCU-3' 的混合物) 进行了 5 μL 的反应。0.5mM ApG 引物序列 (Sigma 公司) 被加入到反应体系中, 并在 30°C 下孵化 1 小时。加入 5 μL 2× 甲酰胺 / 溴苯酚蓝 / EDTA 负载缓冲液后, 混合物在 95°C 下加热 2 分钟。结果分析利用放射自显影法进行于 18% PAGE, 1× Tris- 硼酸盐 -EDTA 和 8M 尿素中。

[0124] 图 9 显示了 3,4-二咖啡酰奎宁酸 (A)、1,5-二咖啡酰奎宁酸 (B)、4,5-二咖啡酰奎宁酸 (C) 以及 3,5-二咖啡酰奎宁酸 (D) 在 5mM 浓度下的 ApG 引物延长实验结果。阴性对照为无聚合酶、无小分子时获得的结果, 阳性对照为有聚合酶、无小分子时获得的结果。图 (b) 为通过磷光成像 (phosphorimaging) 获得的图 (a) 的定量结果。图 (b) 中数据为两次独立实验的平均值, 标准偏差标于图中。结果显示这些化合物在 5mM 浓度下显示出流感病毒 RNA 聚合酶抑制活性。

[0125] 本发明描述了优选的实施例, 应当理解的是, 其中的变化和改进也在本发明精神的范围内, 对本领域技术人员是显而易见的, 本发明的范围由所附权利要求和其同等替换范围确定。

[0001]

序列表

<110> 中国科学院化学研究所 中国科学院生物物理研究所

<120> 咖啡酰奎宁酸类化合物的新用途

<130> CGGNA01120412

<160> 4

<210> 1

<211> 1383

<212> DNA

<213> 人工序列

<400> 1

attgagccat ttctgaagac aacaccacgc cctcttagat tacctgatgg gcctccctgc	60
tctcagecggc egaagttctt gctgatggat gcccttaaatt tgagcatcga agacccgagt	120
catgaggggg aggggatacc gctgtatgat gcaatcaaatt gcatgaaaac atttttcggc	180
tggaaagagc ccaacattgt aaaaccacat gaaaaaggca taaaccccaa ttacctctg	240
gcttggaagc aggtgctggc agagctccaa gatattgaaa acgaggagaa aattccaaag	300
acaaagaaca tgaggaaaac aagccaattg aagtgggcac ttggtgagaa tatggcacca	360
gagaaagtag actttgagga ttgcaaagat gttagegate taaggcagta tgacagtgat	420
gaaccaaagc ctagatcact agcaagctgg atccagagtg aattcaacaa ggcatgcaaa	480
ttgacagatt caagttgat tgaacttgat gaaatagggg aagacgttgc tccaattgag	540
cacattgcaa gtatgagaag gaactatttc acagctgaag tateccattg cagggtact	600
gaatacataa tgaaggagat gtacataaac acagctttgt tgaatgcatc ctgtgcegce	660
atggatgact tccaactgat tccaatgata agcaaatgca gaaccaagga agggagacgg	720
aaaactaacc tgtatggatt cattataaaa ggaagatccc atttgagaaa tgacaccgat	780
gtggtaaact ttgtgagtat ggaattctct cttactgacc cgaggctgga gccacacaag	840
tgggaaaagt actgcgttct cgagatagga gacatgctcc tacggactgc aataggccaa	900
gtatcaagge ccatgtttct ttatgtgaga accaatggaa cctccaagat caagatgaaa	960
tggggcatgg aaatgaggcg atgccttctt caatcccttc aacagattga gagcatgatt	1020
gaggcagagt cttctgtcaa agagaaagac atgactaaag aattctttga aaacaaatca	1080

[0002]

gaaacalggc caallggaga atcacceaaa gggatggagg aaggctccal cgggaagggtg	1140
tgcagaacct tactggctaa atctgttttc aacagtctat atgcatctcc acaactcgag	1200
gggttttctgg ctgaatcaag aaaattgctt ctcatgttgc aggcacttag ggacaacctg	1260
gaacctggga ccttcgatct tggggggcta tatgaagcaa ttgaggagtg cctgattaat	1320
gatccctggg ttttgcttaa tgcactttgg ttcaactcct tcctcacaca tgcactaaaa	1380
taa	1383

<210> 2

<211> 2151

<212> DNA

<213> 人工序列

<400> 2

atggaagact ttgtgcgaca atgettcaat ccaatgattg tcgagcttgc ggaaaaggca	60
atgaaagaat atggggaaga tccgaaaatc gaaacgaaca aatttgccgc aatatgcacg	120
cacttagaag tetgettcat gtattcagat ttccacttta ttgatgaacg gggcgaatca	180
acaattgtag aatctggcga tccaatgca ttattgaaac accggttga aataatcgaa	240
gggagggacc gaacaatggc ctggacagtg gtgaatagta tctgcaacac cacaggagtt	300
gagaagccta aatttctccc agatttgtat gactacaaag agtaccgatt tattgaaatt	360
ggagtgaecg ggagggaagt teacacatac tatctagaaa aagccaaca gataaaatct	420
gagaagacac acallcacal allctcaltc actggagagg aaalggccac caaagcggac	480
tacacccttg atgaagagag cagggcccga atcaaaaacca ggctgtacac tataaggcag	540
gaaatggcca gtagggggtt atgggattcc tttcgtcaat ccgagagagg cgaagagaca	600
attgaagaaa gatttgaat cacaggaact atgcgcaggc ttgccgacca aagtctccca	660
cctaatttct ccagccttga aaactttaga gcctatgttg atggattcga accgaacggc	720
tgcattgagg gcaagcttct tcaaatgtca aaagaagtga acgccagaat tgagccattt	780
ctgaagacaa caccacgccc tcttagatta cctgatgggc ctccctgctc tcagcggctc	840
aagttcttgc tgatggatgc ccttaaattg agcatcgaag acccgagtca tgagggggag	900
gggataccgc tgtatgatgc aatcaaatgc atgaaaacat ttttcggctg gaaagagccc	960
aacattgtaa aaccacatga aaaaggcata aaccccaatt acctcctggc ttggaagcag	1020
gtgctggcag agctccaaga tattgaaaac gaggagaaaa ttccaaagac aaagaacatg	1080
aggaaaacaa gccaatgaa gtgggcactt ggtgagaata tggcaccaga gaaagtagac	1140

[0003]

tttgaggatt gcaaagatgt tagcgatcta aggcagtatg acagtgatga accaaagect	1200
agatcactag caagctggat ccagagtga ttcaacaagg catgcgaatt gacagattca	1260
agtlggallg aactlgaalga aataggggaa gacgtlgtct caatlgaagca callgcaagi	1320
atgagaagga actattttcac agctgaagta tcccattgca gggctactga atacataatg	1380
aaggagtgat acataaacac agctttgttg aatgcatcct gtgcggccat ggatgacttc	1440
caactgattc caatgataag caaatgcaga accaaggaag ggagacggaa aactaacctg	1500
tatggattca ttataaaagg aagatcccat ttgagaaatg acaccgatgt ggtaaaacttt	1560
gtgagtatgg aattctctct tactgaccgg aggctggagc cacacaagtg ggaaaagtac	1620
tgcgttctcg agataggaga catgctccta cggactgcaa taggccaagt atcaaggecc	1680
atgtttcttt atgtgagaac caatggaacc tccaagatca agatgaaatg gggcatggaa	1740
atgaggcgat gcctttctca atcccttcaa cagattgaga gcatgattga ggcagagtct	1800
tctgtcaaag agaaagacat gactaaagaa ttctttgaaa acaaatcaga aacatggcca	1860
attggagaat cacccaaagg gatggaggaa ggctccatcg ggaaggtgtg cagaacctta	1920
ctggctaaat ctgttttcaa cagtctatat geatctccac aactcgaggg gttttcggt	1980
gaatcaagaa aattgcttct cattgttcag gcacttaggg acaacctgga acctgggacc	2040
ttcgatcttg gggggctata tgaagcaatl gaggagtgcc tgattaalga tccctgggtt	2100
ttgettaatg catcttggtt caactcttct ctcacacatg cactaaaata a	2151

<210> 3

<211> 2274

<212> DNA

<213> 人工序列

<400> 3

atggatgtca atccgacttt acttttetta aaagtgccag cgcaaaatgc tataagtacc	60
acattccett aacttgaga tctccatac agccatggaa caggaacagg atacaccatg	120
gacacagtca acagaacaca tcaatattca gaaaagggga aatggacaac gaacacagag	180
actggagcac cccaactcaa tccgattgat ggaccactgc ctgaggataa tgagccgagt	240
gggtalgcac aaacagallg lgtallggaa gcaatggctt tcttgaaaa glcccacca	300
gggatctttg aaaactcgtg tcttgaaacg atggaaattg ttcagcaaac aagagtggat	360
aaactgacct aaggtegeca gacctatgac tggacattga atagaaacca gccggctgca	420
accgctttgg ccaacactat agaggtcttc agatcgaatg gtctaacagc caatgaatcg	480
ggaaggctaa tagatttctt caaggacgtg atgggatcaa tggataaggg agagatggaa	540

[0004]

ataacaacac atttccagag aaagagaaga gtgagggaca acatgaccaa gaaaatggtc	600
acacaaagaa caatagggaa gaaaaaacia aggctgaaca aaaggagcta cetaataaga	660
gactgacac tgaacacaat gacaaaggac gcagaaagag gcaaattgaa gaggagggca	720
attgcaacac ceggaatgca aatcagagga ttcgtgtact ttgtcgaaac actagegagg	780
aglatctgtg agaaacttga gcaatctgga ctcccagctg gagggaaatga aaagaaggct	840
aaattggcaa atgtctgtgag gaagatgatg actaactcac aagatacaga gctctccttt	900
acaattactg gagacaacac caaatggaat gagaatcaga accctcggat gtttctagca	960
atgataacat acatcacaag gaaacaacct gaatggttta gaaatgtctt aagcattgct	1020
cctataatgt tctcaaacia gatggcaaga ttagggaaag gatacatggt cgaaagtaag	1080
agcatgaagc tacggacaca aataccagca gaaatgcttg caagcattga cctgaaatac	1140
ttcaacgaat caacgagaaa gaaaatcgag aaaataagac ctctactaat agatggcaca	1200
gcctcattga gtcttggaat gatgatgggc atgttcaata tgctgagtac agtcttagga	1260
gtttcaatcc tgaatcttgg gcagaagagg tacaccaaaa ccacatactg gtgggacgga	1320
ctccaatct ctgatgattt cgtctcata gtgaatgcac cgaatcatga gggaaatacaa	1380
gcaggggtgg ataggttcta taggacttgc aaactagttg gaatcaatat gagcaagaag	1440
aagtcttaca taaatcggac aggaacattt gaattcacia gcttcttcta ccgctatggg	1500
ttcgtagcca acttccagat ggagctgccc agctttggag tgtctgggat taatgaatcg	1560
gctgacatga gcattggtgt tacagtgata aagaacaata tgataaacia cgaccttga	1620
ccagcaacag ctccagatggc tcttccagta ttcatacagg actacagata cacataccga	1680
tgccacaggg gggatacaca aatccaaacg aggagatcat tcgagctgaa gaggctgtgg	1740
gagcagacce gttcaaaggc aggactgttg gtttcagatg gaggaccaaa cctatacaat	1800
atccggaatc tccacattcc ggaggtctgc ttgaagtggg aattgatgga tgaagactac	1860
cagggcagac tgtgtaatcc tctgaaccgc tttgttagtc ataaggaaat tgagtctgtc	1920
aacaatgctg lgglaatgcc agctcatggc ccagccaaga gcatggaata cgatgcccgtt	1980
gagactacac attcatggat tctaagagg aatcgttcca ttctcaacac cagccaaagg	2040
ggaattcttg aggatgaaca gatgtatcag aagtctgca atctattcga gaaattcttc	2100
cctagcagct catatcggag gccagttgga atttccagca tggtggaggc catggtgtct	2160
agggccccgaa ttgatgcacg aattgacttc gagtctggaa ggattaagaa agaagagttt	2220
gctgagatca tgaagatctg ttccaccatt gaagagctca gacggcaaaa atag	2274

<210> 4

<211> 2268

<212> DNA

[0005]

<213> 人工序列

<400> 4

atggagagaa	taaaagaatt	aagagatcta	atgtcgcagt	cccgcactcg	cgagatacta	60
acaaaaacca	ctgtggacca	tatggccata	atcaagaaat	acacatcagg	aagacaagag	120
aagaacctg	ctctcagaat	gaaatggatg	atggcaatga	aatatccaat	cacagcagac	180
aagagaataa	tggagatgat	tctgaaagg	aatgaacaag	gacaaacgct	ttggagcaag	240
acaaatgatg	ctgggtcgga	cagagtgatg	gtgtctcccc	tagctgtaac	ttggtggaac	300
aggaatgggc	cgacaacaag	tacagtccat	tatccaaagg	tttacaaaac	atactttgag	360
aaggttgaaa	ggttaaaaca	tggagccttc	ggtcccgtcc	atttccgaaa	ccaagttaaa	420
atacgcgcc	gagtggatat	aaacccgggc	catgcagatc	ttagtctaa	agaagcacia	480
gatgttatca	tggaggtegt	tttccaaaat	gaagtgggag	ctagaatatt	gacatcagag	540
tcgcaattga	caataacaaa	agagaagaaa	gaagagctcc	aggattgtaa	aattgctcct	600
ttaatggtgg	catacatggt	ggaaagagaa	ctggtcgca	aaaccagatt	tctaccgta	660
gcaggcggaa	caagcagtgt	gtacattgag	gtattgcatt	tgactcaagg	gacctgttgg	720
gaacagatgt	acactcccgg	cggagaagtg	agaaatgatg	atgttgacca	gagtttgatc	780
atcgtcgcca	gaaacattgl	taggagagca	acagtalcag	cggatccact	ggcatcacta	840
ctggagatgt	gtcacagcac	acaaattggg	ggaataagga	tggtggacat	ccttaggcaa	900
aacceactg	aggagcaagc	tgtggatata	tgcaaagcag	caatgggttt	gaggatcagt	960
tcatecttta	gctttggggg	cttcactttc	aaaagaacaa	gtggatcacc	cgtaagaag	1020
gaagaggaag	tgcttacagg	caacctccaa	acattgaaaa	taagagtaca	tgaggggtat	1080
gaagaattca	caatggttgg	gcggagagca	acagctatcc	tgaggaaagc	aactagaaag	1140
ctgattcagt	tgatagtaag	tggaagagat	gaacaatcaa	tcgctgaagc	gatcattgta	1200
gcaatggltg	ctcacagga	ggallgcattg	alaaaggcag	tccgaggcga	ctgaatttc	1260
gtgaacagag	caaaccaaag	attgaacccc	atgcatcaac	tctgaggcca	cttccaaaaa	1320
gatgcaaaag	tgctgtttca	gaactgggga	attgaacctc	ttgacaatgt	catgggggatg	1380
atcggagtat	tacctgacat	gactccaagc	gcagagatgt	cactgagagg	agtgagagtt	1440
agtaagatgg	gagtggatga	atattccagc	acggagagag	tggtggtgag	tattgaccgt	1500
ttcttgaggg	tccgagatca	gcaggggaac	gtactcttat	ctctgaaga	ggttagtgaa	1560
acacagggaa	cagagaagtt	gacaataaca	tattcatccc	caatgatgtg	ggaaatcaac	1620
ggctcctgagt	cagtccttgg	taacaattat	caatggatca	tcaggaattg	ggagactgta	1680
aagattcaat	ggtctcaaga	tcccacaatg	ctgtacaata	agatggagtt	tgaatcgttc	1740
caatccttgg	tgccaaagc	tgccagaggt	caatacagtg	ggtttgtgag	aacactatcc	1800

[0006]

caacaaatgc	gtgatgtact	ggggacattt	gatactgtcc	aaataataaa	gctgctacca	1860
tttgcagcag	ccccaccgga	gcagagcaga	atgcagtttt	cttctctaac	tgtgaatgtg	1920
agaggctcag	gaatgagaat	actcgtgagg	ggtaactccc	ccgtgttcaa	ctacaacaag	1980
gcaacaaaa	ggcttacagt	cctcggaaag	gacgcaggcg	cattaacaga	agatccagac	2040
gagggaacag	ccggggtgga	atctgcggta	ctgaggggat	tcctaattct	aggcaaagag	2100
gacaaaagat	atggaccagc	attaagcatc	aatgaactga	gcaatcttgc	aaaaggggag	2160
aaggctaatg	tgctgatagg	gcaaggagac	gtggtgttgg	taatgaaacg	gaaacgggac	2220
tctagcatac	ttactgacag	ccagacagcg	acaaaaagaa	ttcggatg		2268

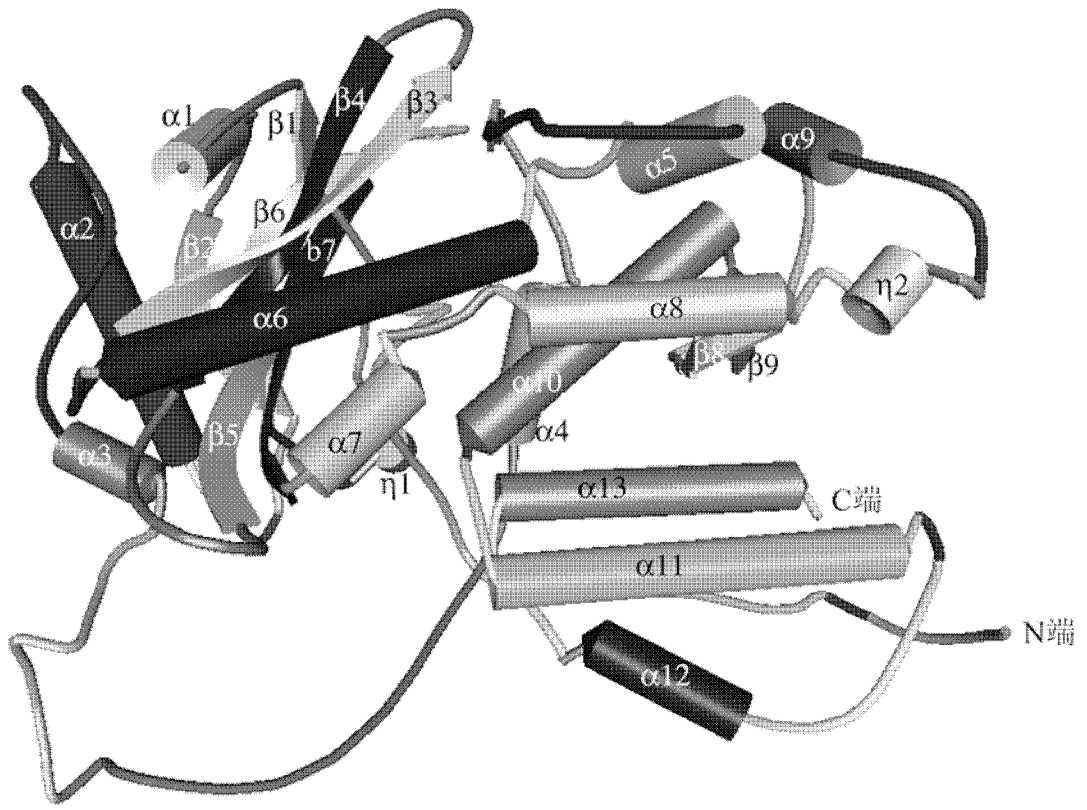


图 1

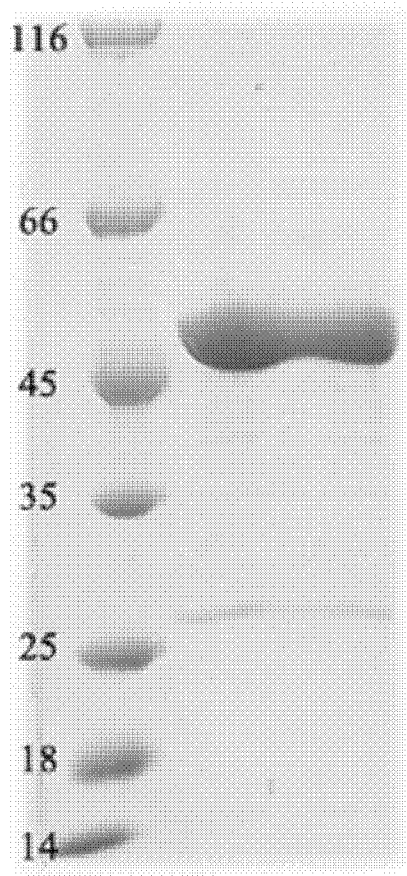


图 2

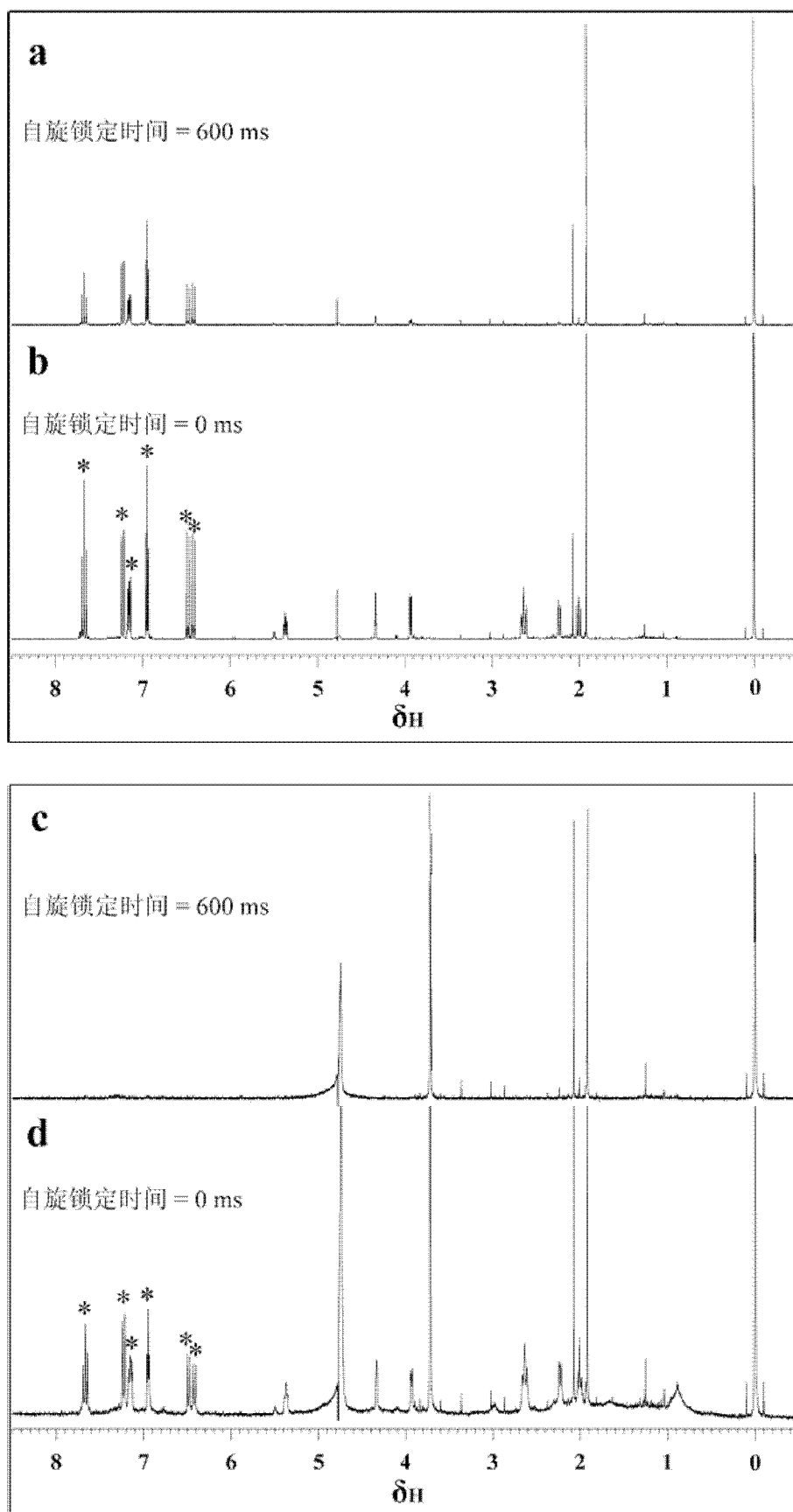


图 3

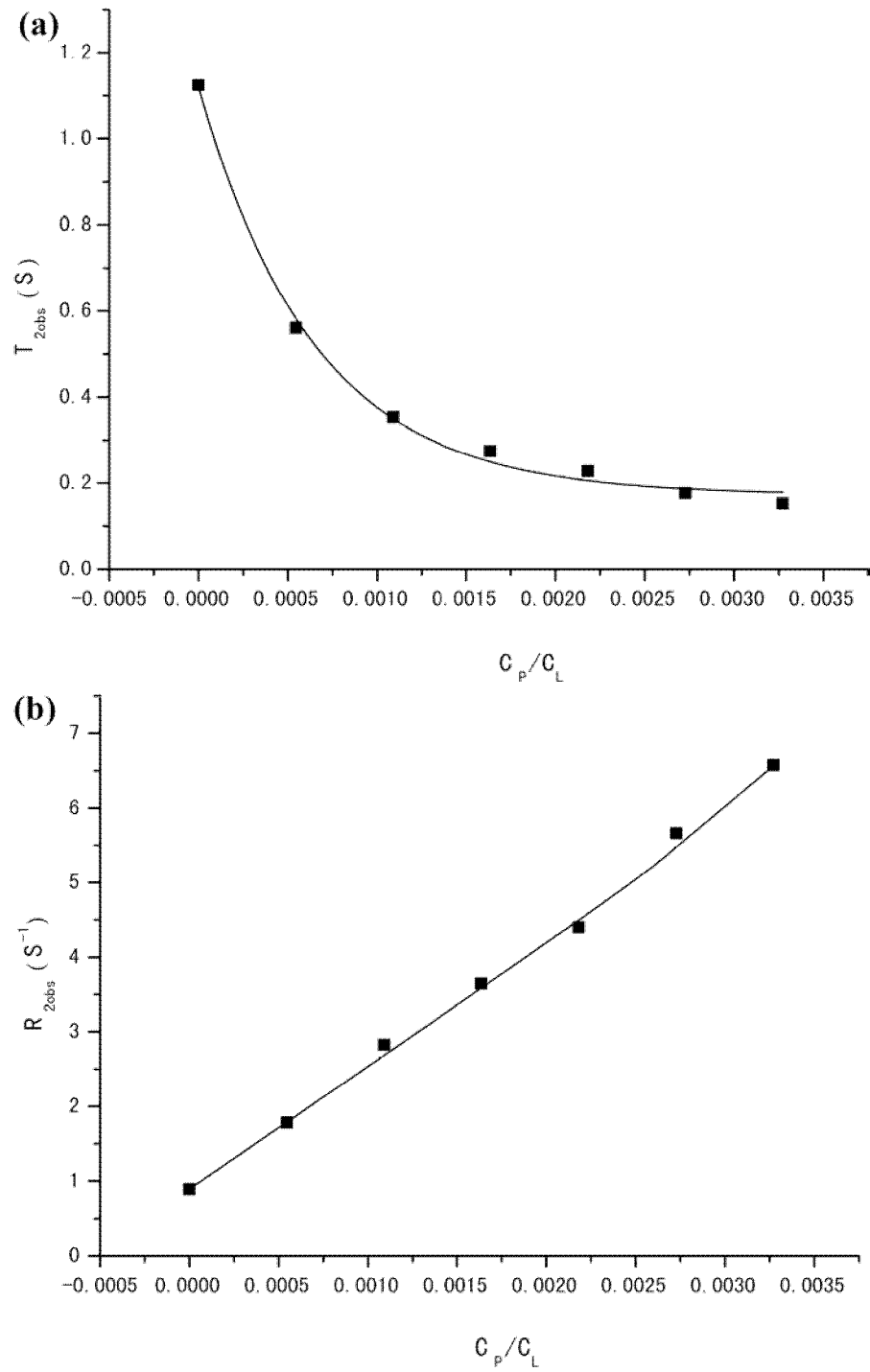


图 4

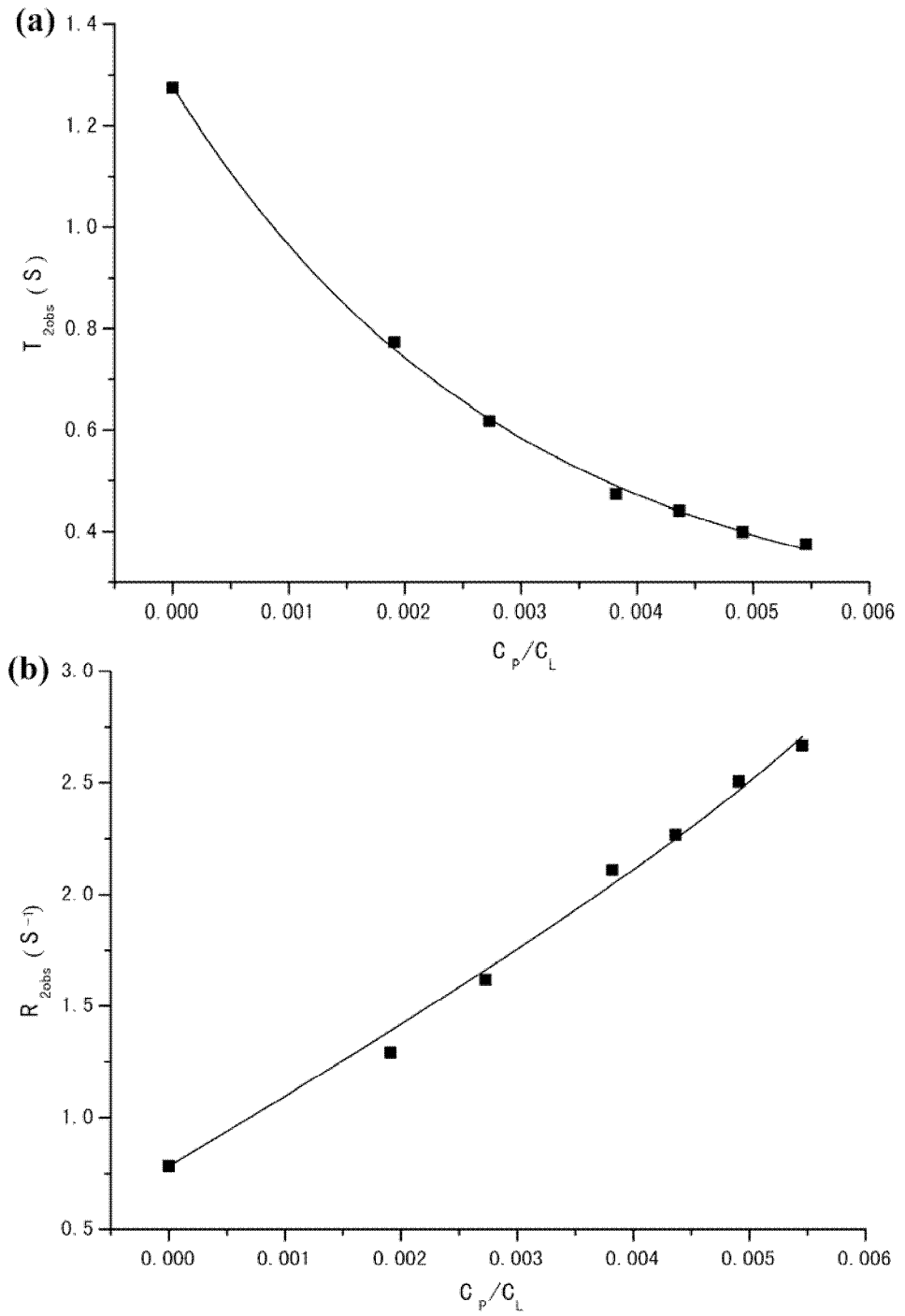


图 5

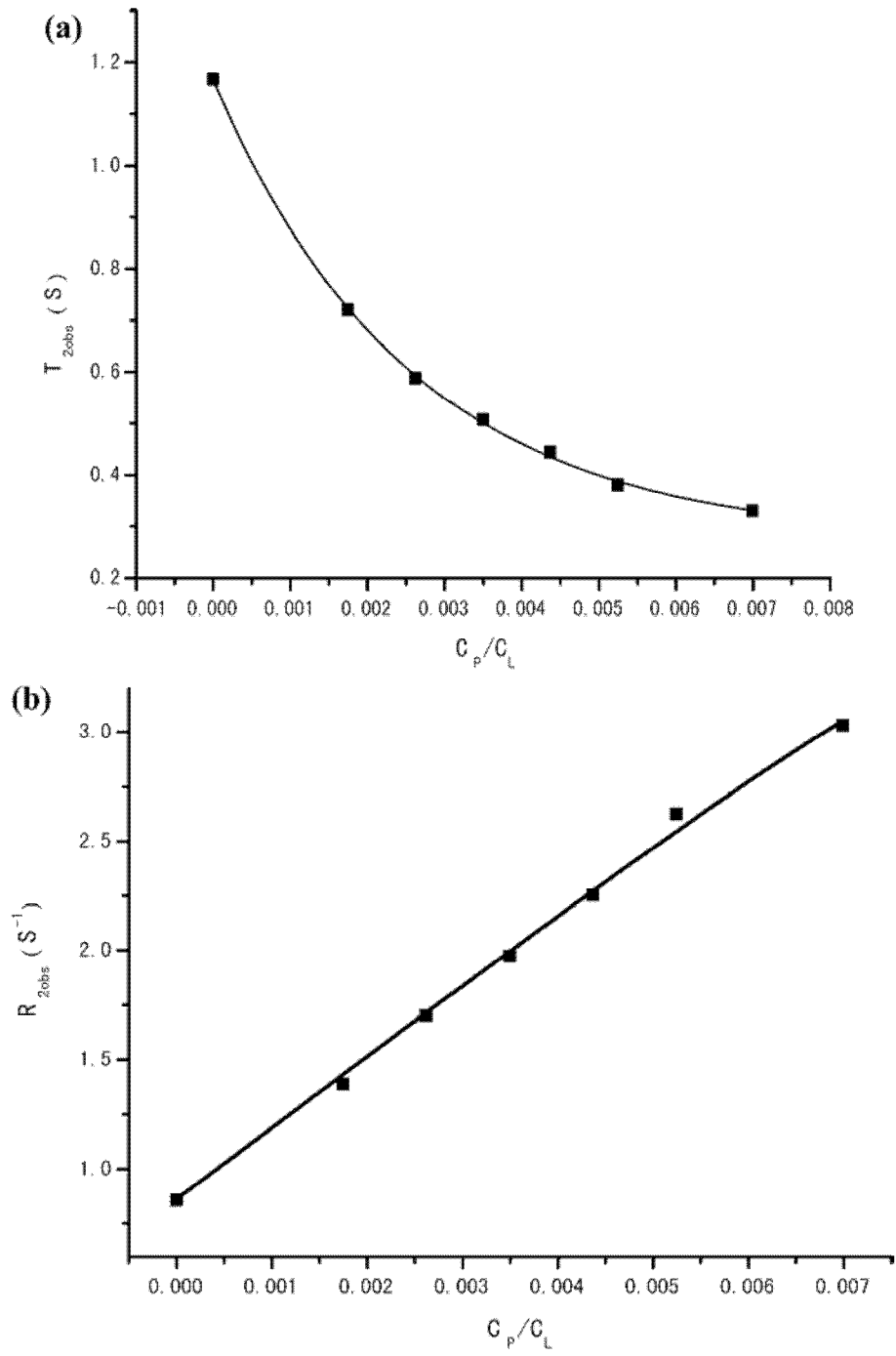


图 6

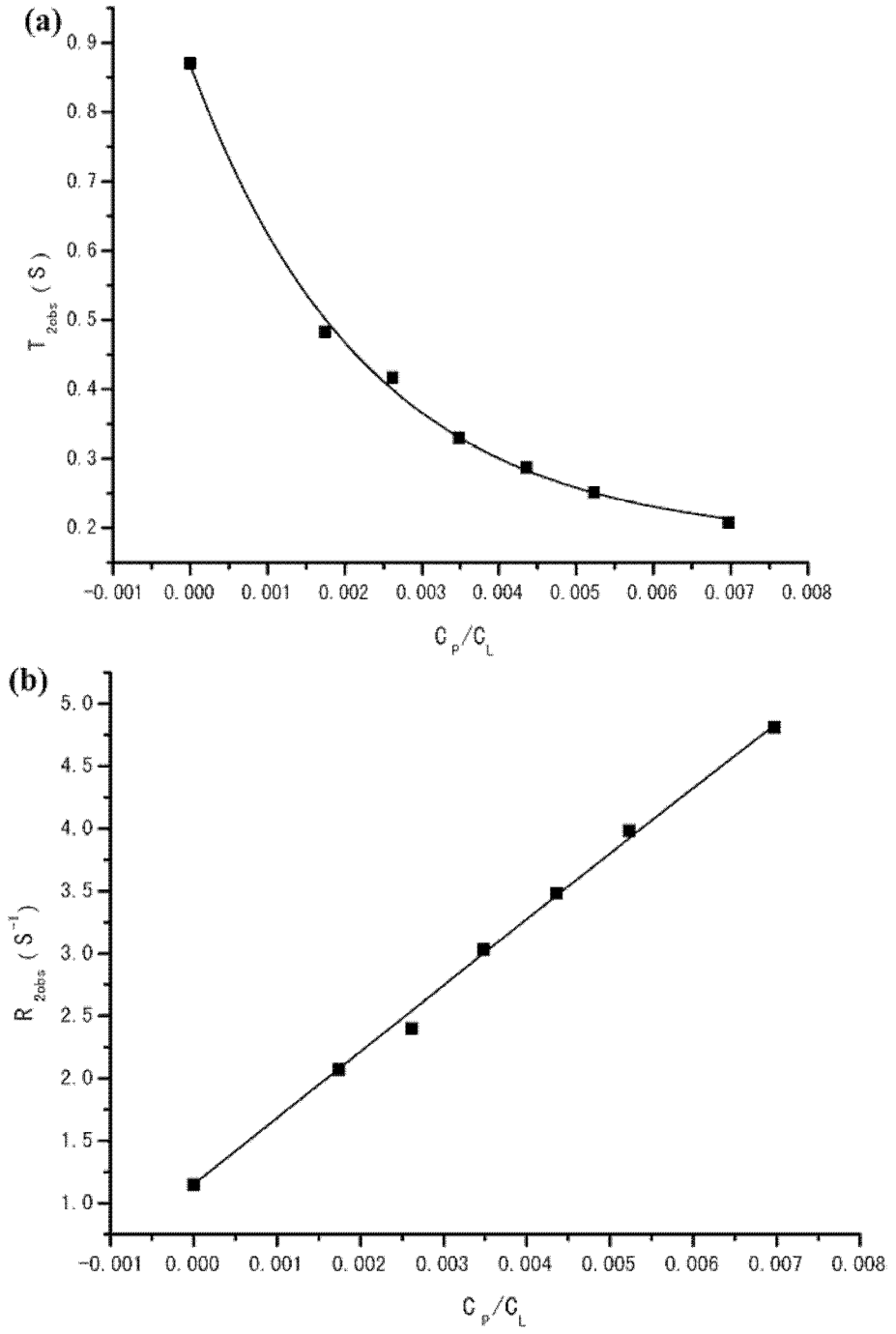


图 7

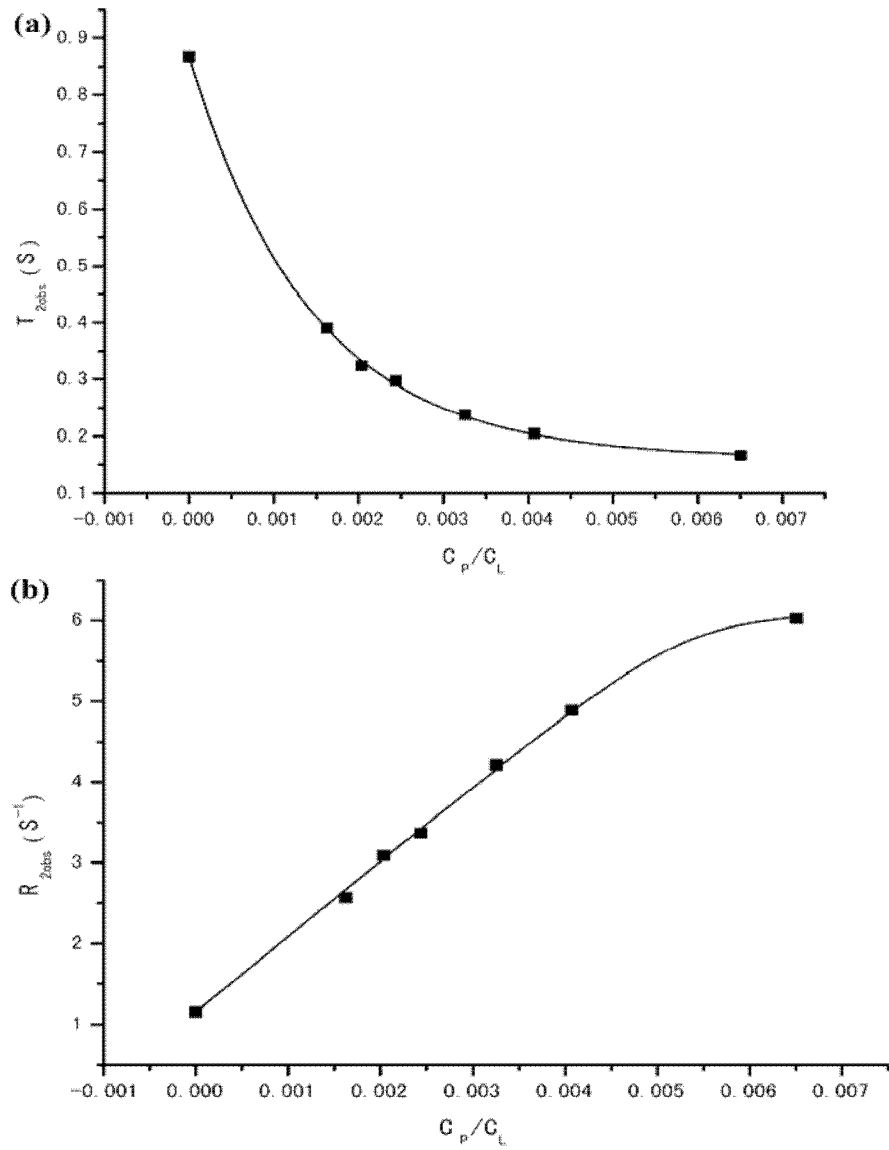


图 8

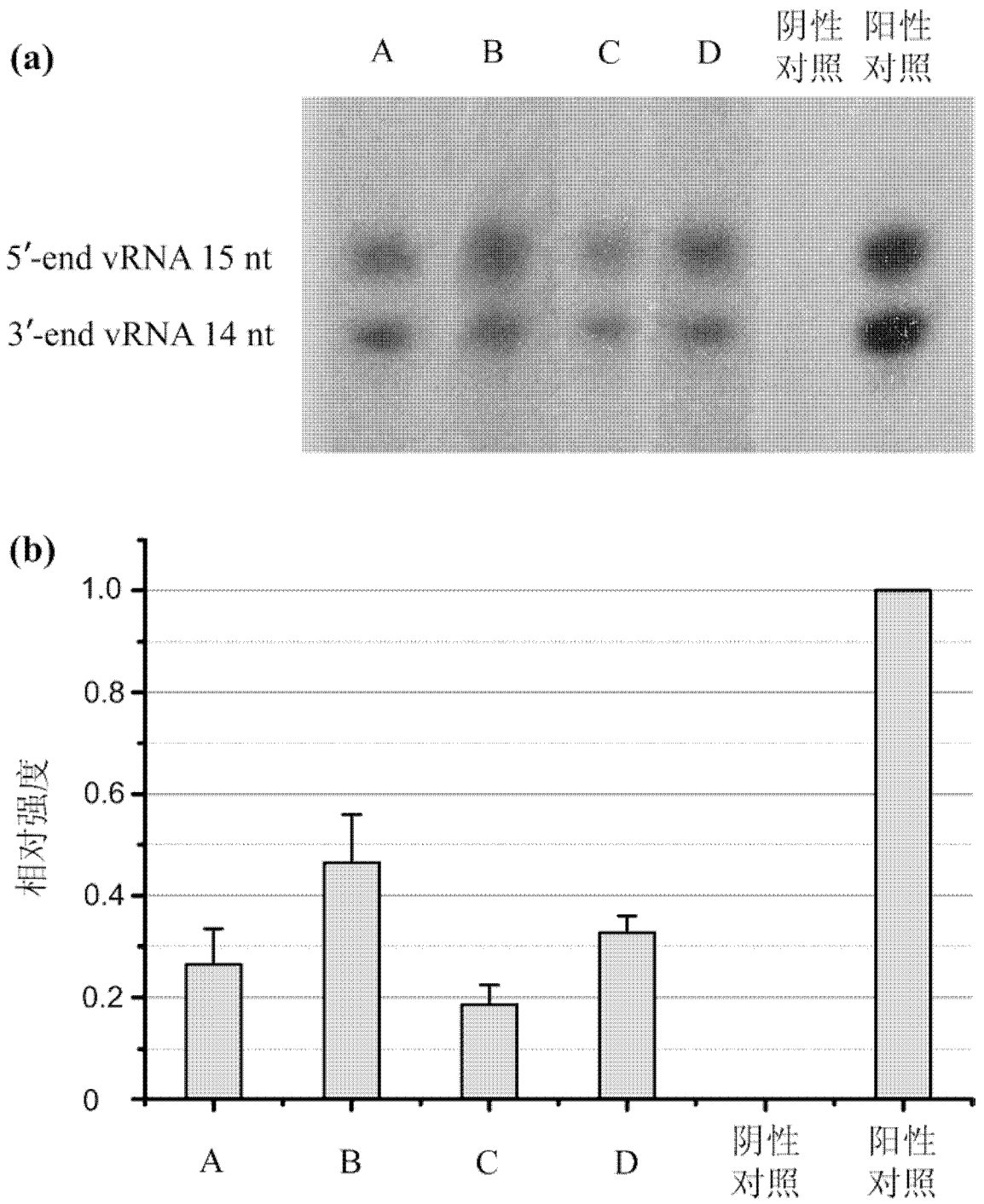


图 9