

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200710087457.0

[51] Int. Cl.

G01N 33/566 (2006.01)

G01N 23/20 (2006.01)

G01N 1/28 (2006.01)

[43] 公开日 2007 年 10 月 31 日

[11] 公开号 CN 101063682A

[22] 申请日 2007. 3. 19

[21] 申请号 200710087457.0

[71] 申请人 南开大学

地址 300071 天津市南开区卫津路 94 号

共同申请人 清华大学

中国科学院生物物理研究所

[72] 发明人 饶子和 姜智勇 孙玉娜 马 明

[74] 专利代理机构 北京三聚阳光知识产权代理有限公司

代理人 李红团

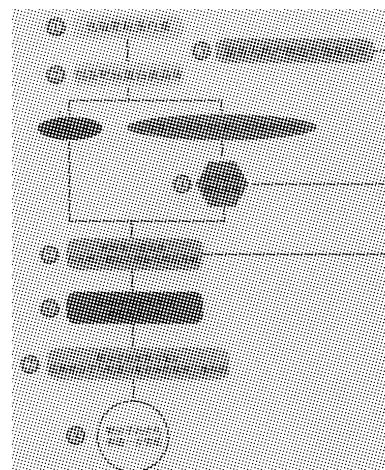
权利要求书 4 页 说明书 33 页 附图 10 页

[54] 发明名称

从中药等天然产物混合体系中高效、快速获取靶分子与靶蛋白复合体精细三维结构的新方法

[57] 摘要

本发明提供了一种从中药等天然产物混合体系中高效、快速获取靶分子与靶蛋白复合体精细三维结构的新方法，本发明以靶标蛋白质晶体结构为基础，使用 X 射线晶体学的手段，对我国传统中药等天然产物进行活性成分的筛选，是从一种新的角度去研究中药等天然产物学和生物学。本发明的方法与目前的药物筛选方法相比，具有筛选效率高、快捷、准确、且相对简单、成本较低的优点。



1、一种从天然产物混合物备选库中高效、快速获取靶分子与靶蛋白复合体精细三维结构的方法，其特征在于包括如下步骤：基于蛋白质晶体结构从天然产物混合物备选库中获取靶分子与靶标蛋白复合体的精细三维结构。

2、根据权利要求1所述的方法，其特征在于包括如下步骤：将靶标蛋白质的晶体与天然产物混合物备选样品接触，分析靶标蛋白质的晶体与天然产物混合物备选样品接触后的晶体的 X 射线衍射数据，获得与靶标蛋白结合的靶分子的电子密度，进而获得靶分子与靶蛋白复合体的精细三维结构。

3、根据权利要求2所述的方法，其特征在于包括如下步骤：

(1) 选取靶标蛋白质；

(2) 制备靶标蛋白质的晶体；

(3) 制备天然产物混合物备选库；

(5) 将靶标蛋白质的晶体与天然产物混合物备选样品接触；

(6) 收集并分析步骤(5)中将靶标蛋白质的晶体与天然产物混合物备选样品接触后的晶体的 X 射线衍射数据，获得与靶标蛋白结合的靶分子的电子密度；；

(7) 鉴定靶分子，获得靶分子与靶蛋白复合体的精细三维结构。

4、根据权利要求3所述的方法，其特征在于包括如下步骤：

(8) 靶分子的优化和进一步的生物学实验。

5、根据权利要求3所述的方法，其特征在于步骤(1)中选取的靶标蛋白质若已报道其具有体外活性测定方法，则在步骤(3)与步骤(5)之间要进行如下步骤(4)：天然产物混合物备选样品对靶标蛋白质的体外活性测定实验，选取对靶标蛋白质具有抑制活性的样品，这些样品称为“有抑制活性的”样品，将“有抑制活性的”样品与靶标蛋白质进行步骤(5)。

6、根据权利要求 3 所述的方法，其特征在于步骤（1）中，选取的靶标蛋白质为通过 X 射线衍射晶体学的方法解析出其三维结构的蛋白质。

7、根据权利要求 3 所述的方法，其特征在于所述天然产物混合物备选库的制备包括如下步骤：

(a)取一定量的初级原料，用乙醇溶液回流提取，合并提取液，浓缩得浸膏；对于微生物代谢产物的提取物制备，则是将含有适量微生物的培养液过滤，得到固体微生物菌丝体和培养液滤液，菌丝体再用丙酮溶液超声破碎，过滤，滤液浓缩得浸膏，此浸膏与培养液滤液都分别作为微生物代谢产物的提取物进行下一步的处理；

(b)称取一定量步骤(a)中所得的浸膏，用双蒸水悬浮，用有机溶剂萃取，分离有机溶剂相和水相；合并水相，备用；合并每次有机溶剂萃取物，蒸除有机溶剂，得到干膏状有机溶剂相成分，备用；

对于微生物提取物的培养液滤液部分则直接用有机溶剂萃取，合并每次有机溶剂萃取物，蒸除有机溶剂，得到干膏状有机溶剂相成分，备用；

(c)将步骤(b)中萃取后的水相部分蒸除水中溶解的少量有机溶剂，用大孔树脂柱层析分离，用水为流动相洗脱，弃去水洗部分；用乙醇溶液洗脱，将乙醇溶液洗脱部分蒸除溶剂，得干膏状样品备用；

(d)上述步骤(b)中制得的干膏状有机溶剂相部分、步骤(c)中制得的干膏状样品为天然产物混合物备选样品，构成一份初级原料所提供的天然产物混合物备选样品，所有初级原料提供的天然产物混合物备选样品即组成了本天然产物混合物备选库。

8、根据权利要求 7 所述的方法，其特征在于步骤（a）中，当初级原料为中药处方时，除用乙醇溶液回流提取，得浸膏外，还另取初级原料，用水煎煮，煎煮液合并，浓缩得浸膏；用乙醇溶液和水提取得到的两份浸膏，都分别作为中药处方的提取物浸膏进行下一步的处理。

9、根据权利要求 8 所述的方法，其特征在于步骤 (a) 为：取一定量的初级原料，用 70%-95%乙醇回流提取，合并提取液，浓缩得浸膏；对于中药处方的提取物浸膏的制备，则按照处方中规定的药材种类和 1-6 倍用量组成初级原料，分别用 70%-95%乙醇按上述同样的方法制得浸膏一；另外，同样则按照处方中规定的药材种类和 1-6 倍用量组成初级原料，用双蒸水煎煮，煎煮液合并，浓缩得浸膏二；浸膏一和浸膏二都分别作为中药处方的提取物浸膏进行下一步的处理；对于微生物代谢产物的提取物制备，则是将含有适量微生物的培养液过滤，得到固体微生物菌丝体和培养液滤液，菌丝体再用 60%-90%丙酮超声破碎，过滤，滤液浓缩得浸膏，此浸膏与培养液滤液都分别作为微生物代谢产物的提取物进行下一步的处理。

10、根据权利要求 9 所述的方法，其特征在于步骤(a)中所用乙醇量与初级原料质量比为 6:1-12:1；丙酮与菌丝体质量比为 8:1-30:1；步骤(b)中浸膏与悬浮用水的质量比为 1:5-1:15，萃取用有机溶剂为乙酸乙酯或氯仿，其体积与悬浮用水的体积比为 6:1-10:1，萃取用有机溶剂的体积与培养液滤液的体积比为 1:8-1:1。

11、根据权利要求 10 所述的方法，其特征在于步骤(c)为：

将步骤(b)中萃取后的水相部分蒸除水中溶解的少量有机溶剂，用大孔树脂柱层析分离，依次用水、50%乙醇、95%乙醇为流动相洗脱，弃去水洗部分；将 50%乙醇和 95%乙醇洗脱部分分别蒸除溶剂，得干膏状样品备用。

12、根据权利要求 1-11 任一项权利要求所述的方法，其中所述的天然产物混合物备选库的初级原料选自以下中的任一种：传统药材或潜在的药用动植物、按中药处方规定的药材种类和用量组成的混合药材、海藻或海绵等海洋生物、微生物。

13、根据权利要求 12 所述的方法，其中所述的传统药材为《中药药典》、《中华本草》中收录的品种；其中所述的潜在的药用动植物为《中药药典》、《中华本草》

中未收录而在《中国植物志》中收录的品种。

14、根据权利要求 12 所述的方法，其特征在于步骤（5）中的靶分子是天然产物混合物备选样品中存在的能够与靶标蛋白质结合的包括抑制剂在内的配体分子。

15、根据权利要求 12 所述的方法，其特征在于使靶标蛋白晶体与天然产物混合物备选样品接触，使用的方法为晶体浸泡或共结晶。

16、根据权利要求 15 所述的方法，其特征在于晶体浸泡可以采取“单独浸泡”、“分组浸泡”、或“全部混合浸泡”的方式。

17、根据权利要求 13 所述的方法，其特征在于所述靶标蛋白是 SARS 冠状病毒主蛋白酶。

18、根据权利要求 13 所述的方法，其特征在于所述靶标蛋白是 HCV NS5B 蛋白。

从中药等天然产物混合体系中高效、快速获取靶分子与靶蛋白复合体精细三维结构的新方法

技术领域

本发明涉及一种从中药等天然产物混合体系（天然产物混合物备选库）中高效、快速获取靶分子与靶蛋白复合体精细三维结构的新方法，属于药物筛选方法领域。

背景技术

天然产物又称次级代谢产物(Secondary Metabolites)，具有结构的多样性和生物活性的多样性。天然产物及其衍生物在以往的疾病治疗中发挥了无可限量的作用，也是当今药物研发过程中最具潜力的资源之一（Newman DJ, 等人 *Nat Prod Rep*, 2000, 17(3):215-234; Lee K-H. *J Nat Prod*. 2004, 67(2):273-283)。但是，在过去 10 年中，制药工业对天然产物的研究兴趣有所降低，这主要是由于传统的天然产物粗提物库与常规的高通量的筛选方法缺乏兼容性（Frank E. and Guy T. *Nature Rev. Drug Discov*. 2005, 4:206-220）。

上个世纪 90 年代组合化学（Combinchem）的出现，使得从天然产物中发现新药的研究被冷落了多年。组合化学能够合成数目巨大的化合物（理论上据估计有 10^{18} 到 10^{200} 个化合物，通常认为是 10^{60} 个），这就为常规的高通量的筛选方法提供了理想的筛选化合物库，因此，组合化学这一技术的出现，使得大部分制药公司都转向了以组合化学库为基础的高通量筛选路线上，他们期待采用这种技术手段能够大大加快活性化合物的发现速度。然而，近年来制药工业的新药产出率并没有像预期的那样提高，而是持续下跌。造成这种结果的原因很多，一个最主要的原因可能就是由于组合化学技术出现之后，制药工业开始采用组合化学合成的含有数目巨大的化合物库来进行高通量的筛选，因此对天然产物研发的兴趣有所降低（David J. 等人. *J. Nat. Prod*. 2003, 66, 1022-1037）。组合化学方法是从共同的结构模板出发，选择具有相同功能的多种组建模块，通过同种键反应实现的分子多样性。这样合成的化合物的结构骨架缺乏多样性。因此，组合化学作为药物发现的工具目前仍有其局限性，有待进一步发展和完善。而天然产物在药物研发中的独特作用，

近年来重新引起人们的关注。

几十年来，天然产物一直是药物和药物先导化合物的重要源泉，天然产物在新药和先导化合物的发现中起着重要作用。美国国立癌症研究院最近的调查显示，1981—2002年上市的877种小分子新化学实体药物中，来源于天然产物的占61%。其中，6%直接来源于天然产物，27%为天然产物的衍生物，5%来源于天然产物药效团的合成物，23%是根据天然产物结构进行模拟设计的合成物 (David J. 等人. *J. Nat. Prod.* 2003, 66, 1022-1037)。

目前，用于药物筛选的天然产物库共有三种形式，一是天然产物粗提物库，为混合物库，包含有10—100种组分；二是天然产物提取组分，半纯品库，包含有5—10种组分；三是纯品库，单一组分 (Frank E. and Guy T. *Nature Reviews Drug discovery.* 2005, 4:206-220)。

对于天然产物纯品库，命中物 (hit) 的检测方法与人工化学合成的纯品库相同，可以用于常规的高通量的筛选方法中，因此，当前的许多化学家们不辞辛劳的进行天然产物的分离纯化工作。然而，对于每一种天然产物粗提物来说，分离得到其每种成分的纯品化合物，需要较长的周期和大量的工作，并带有明显的盲目性，获得有效化合物的准确性非常低。

对于天然产物粗提物库，简单易得，但用常规的高通量的筛选方法进行筛选的可靠性差，这主要体现在如下方面：混合物库成分较复杂且其各个成分的浓度并不确定，有可能相差10倍、100倍甚至更多。因此，对于那些浓度低的组分，可能在常规的高通量的活性筛选中被忽略掉，而含量较丰富的组分有可能表现出非特异性的抑制，比如破坏检测体系的pH值或者其它的物理特性，造成假阳性结果 (Frank E. and Guy T. *Nature Reviews Drug discovery.* 2005, 4:206-220)。而本发明的通过使用X射线晶体学等生物物理的方法可以克服这一缺点，使用这种技术可以检测到与靶标蛋白有较低亲和力的化合物，因此使得天然产物粗提物库能够应用于X射线晶体学的筛选方法中，大大提高筛选效率，并提高命中率，可谓是当今药物研发的革命性进展。

对于天然产物尤其是来源于中药等的天然产物的研究是很有必要的，因为中药在我国已有数千年应用的历史，是一个伟大的宝库。但是传统中药在使用时，多以复方药物、采用煎煮的方法，由于有效成分相对含量较低，往往需要长期服用，才能取得一定的疗效，而对于很多急性病症，中药的疗效较差；而且，由于传统中药成分极其复杂，以例如汤剂经煎煮等方式使用后，其中化学成分又可能产生化合或分解等新的变化，难于确定具体何种成分在起作用，从而无法控制药材种植、药材炮制，以及拟定成药后的质量标准，也给药理、临床实验带来困难。

中药所含的化学成分复杂，有效成分却往往含量低微。中药化学成分的

分离、纯化和鉴定则需要反复运用各种色谱方法和多种波谱方法去完成，一味中药的化学成分研究已经是一项周期长、工作量大的工作，这还是仅限那些含量相对较大的成分。在结构清楚的各种成分里面寻找有效成分又需要对纯品进行各种体外、体内的生物活性测试，这其实是一种盲目的筛选模式，往往大多数分离得到的纯品化合物是无效的，而且这种简单的生物活性测试模型能说明的问题也是有限的。即使我们能够确定一些有效成分，那么这些成分是如何对某些疾病起到治疗作用的，它们是如何通过与机体的相互作用而对机体进行调节的，以及关于这些小分子成分与体内生物大分子相互作用细节，目前几乎很少有详细明确的阐述。

随着基因组学、生物信息学等的快速发展，产生了基于分子生物学和结构生物学的药物筛选方法。人类以及微生物的基因组计划提供了数量空前的潜在药物靶标，仅仅在人类基因组的工作草图完成后不到 1 年的时间内就发现了 1778 个疾病基因。这些工作推动了蛋白质表达(蛋白质组学)以及蛋白质结构(结构生物学)的研究，与人类疾病相关基因的识别、鉴定及其所表达的蛋白质的结构与功能的研究为药物筛选与发现开拓了不可限量的前景，给新药发现带来了契机。随着科学技术的进一步发展，研究者开始以蛋白质的晶体结构为基础进行先导化合物的药物筛选，这类筛选方法主要包括：

(1) 基于蛋白质三维结构的虚拟药物筛选和设计(计算机辅助的药物设计)

分子生物学和结构生物学的发展，提供了许多靶标蛋白质的三维结构，因此，按照互补性原则，根据蛋白质的三维结构的药物设计方法已被广泛使用。它与传统的药物化学方法不同，是根据蛋白质的三维结构，在结合位点找到与之相适配的小分子结构。一般可分为两种途径：一是基于数据库搜寻的也称为虚拟筛选，即按照受体结合部位的形状和结构特征，将有机分子库的化合物以其三维结构逐一对接到受体结合部位，在小分子数据库中找出与结合部位相适配的分子。另一种方法称为全新药物设计，即按照受体结合部位的形状和结构特征，由原子、化学基团或片断拼接成与之互补的配体分子。事实上，前者相当于在许多已有的钥匙中寻找可开启锁的钥匙，而后者是按照锁的结构制造新的钥匙。

可是，目前用于虚拟筛选的小分子数据库主要是剑桥小分子库，此库中新化合物种类较少、更新慢，而且许多化合物的成药性较差，即使能够有活性，进一步开发的潜力不大。全新药物设计的方法能够解决上述问题，计算机全新药物设计可以在靶标蛋白质的活性口袋内放置片断，然后通过优化静电、范德华和氢键作用“生长”出适合活性口袋的分子。但这种方法又往往产生难于合成的分子。

(2) 基于片断(fragment-based)的筛选

基于片断筛选中的片断指的是分子量小于 200Da 的小分子基团，通常是组成化合物的基本基团，如萘醌、苯环等等。通常，这种小分子片断对靶标蛋白具有较低的结合力，因此往往不被常规的高通量筛选方法所发现。

但是这种方法也存在着本身的局限性：首先，分子量较小的基团在低分辨率(低于 2.5Å 的分辨率)的电子密度中不易与结晶溶液中本身存在的小分子(如 DMSO 等)、离子(如酸根离子等)等区别开，因此对靶标蛋白的晶体质量要求较高，极大的限制了靶蛋白的范围；其次，能够稳定存在的、分子量在 200Da 水平上的小分子基团种类有限，目前使用较多的用于基于片断筛选的备选库中也只有数百至数千种成分，这就极大的限制了命中的可能性；再次，这种筛选方法将很多不能够合成、无法分解为基本基团的天然产物排除在药物筛选的范围之外，限制了筛选范围。

截至到目前为止，现有技术中没有描述以蛋白质的晶体结构为基础，从中药等天然产物混合体系中高效、快速获取靶分子和靶蛋白精细三维结构的方法。对于具有结构多样性和生物活性多样性的天然产物的研究，仍然以传统方式进行，即用药物靶蛋白筛选粗提物，如果发现活性，就将该提取物分割成若干组分，再分离和鉴定活性化合物。这个过程不仅缓慢、效率低、劳动强度大，而且也不能保证筛选出的先导化合物在化学上是可行的。因此，有待发展一种全新的天然产物药物研发模式，并迫切需要建立一种新的、能满足科学技术发展需要的来源于中药等的天然产物混合物药物筛选备选库，其能够用于从天然产物混合物中进行快速、高效、全面的药物筛选方法。

发明内容

本发明的目的在于提供一种从中药等天然产物混合体系中高效、快速获取靶分子与靶蛋白复合体精细三维结构的新方法，该方法能够高效进行药物筛选，该方法以蛋白质晶体结构为基础，直接以经过优化的能够用于 X 射线晶体学方法的来源于中药等的天然产物混合物备选库进行筛选，能够高效、准确的寻找特异性的靶分子（能够与靶标蛋白结合的包括抑制剂在内的配体小分子），提高了筛选的效率和准确性。本发明的方法较之目前已存在的药物筛选的方法相对简单、成本较低，但筛选效率高、快捷、准确，并可以从天然产物混合体系中，得到结合力最好的靶分子。

具体而言，本发明以靶标蛋白质晶体结构为基础，使用 X 射线晶体学的手段，对我国传统中药等天然产物进行活性成分的筛选，是从一种新的角度去研究中药等天然产物学和生物学。其实质内容可包括两方面：第一，靶蛋白的表达、分离纯化、结晶乃至晶体衍射数据的收集和分析，这属于生物大分子的研究范畴；第二，打破本身为一个复杂的有机生命体的中药等天然产物中所含的千万种非孤立存在的小分子间原有的“生命联系状态”，使这些小分

子化合物自由地与靶蛋白进行竞争性的结合，这又属于化学尤其是天然产物化学研究的范畴。把天然产物粗提物中的每一种小分子分离纯化出来，单独进行靶蛋白的活性筛选，能最大限度地排除掉实验中的偶然因素，能保证对靶蛋白有作用（包括抑制作用）的靶分子确实来源于结构确定的某种小分子，但是这样做不但工作量巨大，研究周期长，而且对筛选本身来说是盲目的。为此，本发明采用了如下方案：将靶蛋白的晶体放在含有很多种小分子的天然产物粗提物混合物样品中，像钓鱼似的“钓出”对其最具结合能力的靶分子。这样做的优点是显而易见的，因为它可针对特定的靶标蛋白高效率地在天然产物混合物库中进行药物筛选。为了将天然产物粗提物样品中会对此过程产生影响的因素去除（如粗提物本身所含的生物大分子、重金属、无机盐等成分，以及由于样品本身太杂造成的分子间的相互作用），又能保留所有可能与靶蛋白结合的靶分子以保证此天然产物粗提物样品提供最大的可筛选成分，避免出现遗憾的假阴性，需要对用于 X 射线晶体学筛选的天然产物混合物样品进行一定的优化。

本发明提供的一种从天然产物混合物备选库中高效、快速获取靶分子与靶蛋白复合体精细三维结构的方法，其包括如下步骤：基于蛋白质晶体结构从天然产物混合物备选库中获取靶分子与靶标蛋白复合体的精细三维结构。

本发明提供的方法可包括如下步骤：将靶标蛋白质的晶体与天然产物混合物备选样品接触，分析靶标蛋白质的晶体与天然产物混合物备选样品接触后的晶体的 X 射线衍射数据，获得与靶标蛋白结合的靶分子的电子密度，进而获得靶分子与靶蛋白复合体的精细三维结构。

本发明提供的方法可包括如下步骤：

- (1) 选取靶标蛋白质；
- (2) 制备靶标蛋白质的晶体；
- (3) 制备天然产物混合物备选库；
- (5) 将靶标蛋白质的晶体与天然产物混合物备选样品接触；
- (6) 收集并分析步骤（5）中将靶标蛋白质的晶体与天然产物混合物备选样品接触后的晶体的 X 射线衍射数据，获得与靶标蛋白结合的靶分子的电子密度；；
- (7) 鉴定靶分子，获得靶分子与靶蛋白复合体的精细三维结构。

本发明的方法还可包括如下步骤：

- (8) 靶分子的优化和进一步的生物学实验。

本发明提供的方法中，当步骤（1）中选取的靶标蛋白质若已报道其具有体外活性测定方法，则在步骤（3）与步骤（5）之间要进行如下步骤（4）：天然产物混合物备选样品对靶标蛋白质的体外活性测定实验，选取对靶标蛋白质具有抑制活性的样品，这些样品称为“有抑制活性的”样品，将“有抑

制活性的”样品与靶标蛋白质进行步骤(5)。

本发明的优选技术方案如下：

1. 靶标蛋白质的选取

所有用 X 射线晶体学的方法解析出三维结构的蛋白质都可以作为本发明方法的靶标蛋白质。针对这些蛋白质的重要性、蛋白表达量大小以及晶体生长难易程度等，优先选取与人类重要疾病或重要病毒相关的蛋白质（例如人类 HIV 相关蛋白质包括（HIV 蛋白酶、HIV 整合酶、HIV 反转录酶）、HCV NS5B 蛋白、SARS 主蛋白酶等）；优先选取能够在原核表达系统（主要指 *E. coil* 表达系统）中获得大量表达（包括可溶性表达或者不可溶的以包涵体形式存在的表达），晶体能够大量获得的蛋白质。

在确定了靶标蛋白之后，可以采用从公司购买、由合作者提供或从 cDNA 文库中克隆的方法，参照已经发表的关于该靶标蛋白的克隆、表达、纯化文献中的方法（对于 HCV NS5B 蛋白可参考：Ferrari, Wright-Minogue 等人. 1999; Lesburg, Cable 等人. 1999; 对于 SARS 主蛋白酶可参考：Yang H 等人. 2003. The Crystal Structures of SARS Virus Main Protease Mpro and Its Complex with an Inhibitor. PNAS, 100(23): 13190-13195），得到正确的构建（construct），并参照已发表文献中报道的方法（对于 HCV NS5B 蛋白可参考 Ferrari, Wright-Minogue 等人. 1999; 对于 SARS 主蛋白酶可参考：Yang H 等人. 2003. The Crystal Structures of SARS Virus Main Protease Mpro and Its Complex with an Inhibitor. PNAS, 100(23): 13190-13195）进行该靶标蛋白质的表达和纯化，并可对该靶标蛋白的表达条件和该靶标蛋白的纯化条件进行适当的改进（如改换蛋白表达载体、改变蛋白表达温度，以及改变蛋白纯化溶液等）。

2 制备靶标蛋白的晶体

使用上述经过纯化的靶标蛋白，参照已报道的该靶标蛋白的结晶条件或经过自己优化的该靶标蛋白的结晶条件（优化靶标蛋白的结晶条件包括改变靶标蛋白的结晶浓度，改变靶标蛋白的结晶池液的组成等），生长大量、衍射能力强的晶体，备用。

3 天然产物混合物备选库的制备

天然产物混合物备选库中的样品的初级原料的来源很广泛，传统中药在我国已有几千年的应用，至今还用于各种疾病的治疗，首当其冲地是我们筛选备选库的初级原料的来源。传统经方、验方等中药处方提取物也可以作为初级原料的来源。另外，我们也不拘泥于传统常用的中草药，一些少数民族药物、民间应用的天然药物中也往往含有新颖的、令人感兴趣的小分子化合物。海洋生物是提供新颖结构分子巨大的生物资源，由于生态环境与陆地截然不同，导致体内生物合成过程不同，使得海洋生物与陆地生物相比，具

有更丰富的生物分子多样性，因此海洋生物也是我们备选库原料的重要来源。其它的如苔藓、真菌等代谢产物以及微生物转化产物由于其特有的生理过程和生长环境，也含有及其丰富的特有天然产物小分子，也是我们初级原料的来源。总之，我们要制备的是最大可能的包括最广泛的天然产物的药物筛选备选样品，这些样品的集合组成本发明的天然产物混合物备选库。

由于天然产物粗提物所含的成分是相当复杂的，因此必须对天然产物的粗提物进行优化，目的是去除干扰因素如蛋白、多糖等，保留成药性较好的小分子并使之尽量集中，提高相对浓度。本发明为实现这一目的，采用本文如下的技术方案制备天然产物混合物备选库，其基本原理为：对初级原料提取后，提取物（粗提物）先使用不同极性的溶剂如乙酸乙酯进行萃取、萃取后的水相再用大孔树脂分离的方法进行处理，这样既能保证去除干扰因素，又能利用两种不同分离机制的差别，使小分子集中到极性不同的段位，由此得到能够用于X射线晶体学等生物物理方法的天然产物混合物备选库。

天然产物混合物备选库的制备方法包括如下步骤：

(a) 取一定量的初级原料，用乙醇溶液回流提取，合并提取液，浓缩得浸膏；对于微生物代谢产物的提取物制备，则是将含有适量微生物的培养液过滤，得到固体微生物菌丝体和培养液滤液，菌丝体再用丙酮溶液超声破碎，过滤，滤液浓缩得浸膏，此浸膏与培养液滤液都分别作为微生物代谢产物的提取物进行下一步的处理；

(b) 称取一定量步骤(a)中所得的浸膏，用双蒸水悬浮，用有机溶剂萃取，分离有机溶剂相和水相；合并水相，备用；合并每次有机溶剂萃取物，蒸除有机溶剂，得到干膏状有机溶剂相成分，备用；

对于微生物提取物的培养液滤液部分则直接用有机溶剂萃取，合并每次有机溶剂萃取物，蒸除有机溶剂，得到干膏状有机溶剂相成分，备用；

(c) 将步骤(b)中萃取后的水相部分蒸除水中溶解的少量有机溶剂，用大孔树脂柱层析分离，用水为流动相洗脱，弃去水洗部分；用乙醇溶液洗脱，将乙醇溶液洗脱部分蒸除溶剂，得干膏状样品备用；

(d) 上述步骤(b)中制得的干膏状有机溶剂相部分、步骤(c)中制得的干膏状样品为天然产物混合物备选样品，构成一份初级原料所提供的天然产物混合物备选样品，所有初级原料提供的天然产物混合物备选样品即组成了本天然产物混合物备选库。

本发明天然产物混合物备选库的制备方法优选方案为：

(a) 制备提取物（粗提物）浸膏

取一定量的初级原料，用实验用乙醇回流提取 2-4 次，每次 1-3 小时，合并提取液，浓缩得浸膏。

所选用的初级原料，可选自上述作为本发明的天然产物混合物备选库的

原料除微生物外的任一种物质；其中所述的实验用乙醇浓度为 70%-95%乙醇，优选 95%乙醇；其中回流提取次数优选为 3 次，每次优选 2 小时；其中所用乙醇量与药材量质量比为 6:1-12:1。对于中药处方的提取物浸膏的制备，则按照处方中规定的药材种类和 1-6 倍用量组成初级原料，分别用 70%-95%乙醇按上述同样的方法制得浸膏一；另外，同样则按照处方中规定的药材种类和 1-6 倍用量组成初级原料，用双蒸水煎煮 3 次，每次 1-3 小时，煎煮液合并，浓缩得浸膏二。浸膏一和浸膏二都作为中药处方的提取物浸膏进行下一步的处理。对于微生物代谢产物的提取物制备，则是将含有适量微生物的培养液过滤，得到固体微生物菌丝体和培养液滤液，菌丝体再用 80%丙酮超声破碎 20-60 分钟，过滤，滤液浓缩得浸膏，此浸膏与培养液滤液都作为微生物代谢产物的提取物进行下一步的处理。

(b) 制得有机溶剂萃取部分

称取一定量步骤(a)中的浸膏，用双蒸水悬浮，倾入分液漏斗中，用适宜量的乙酸乙酯萃取 4-6 次，萃取时需充分振摇，静置 5-8 小时分层后，分离有机溶剂相和水相；合并每次有机溶剂相萃取物，用旋转蒸发器蒸除有机溶剂，得到干膏状有机溶剂相成分，备用。

对于微生物提取物的培养液滤液部分则直接用适宜量有机溶剂萃取 4-6 次，合并每次有机溶剂萃取物，用旋转蒸发器蒸除有机溶剂，得到干膏状有机溶剂相成分，备用；

其中，萃取用有机溶剂用乙酸乙酯或氯仿，萃取前浸膏与悬浮用水的质量比为 1:5-1:15，优选 1:10；分液漏斗选用适宜的分液漏斗，例如，如果加入 100ml 水，则选用 1000ml 的分液漏斗；萃取用有机溶剂的体积与悬浮用水的体积比为 6:1-10:1，优选 8:1；选用的旋转蒸发器优选 EYELA N1001 型旋转蒸发器。

(c) 制得水相大孔树脂 50%乙醇和 95%乙醇洗脱部分：

将步骤(b)中萃取后的水相部分蒸除水中溶解的少量有机溶剂，用大孔树脂柱层析分离，依次用水、50%乙醇、95%乙醇为流动相洗脱，水洗脱 6-10 次，优选 7 次，每次洗脱用水为 2-4 倍体积量的步骤 2 中悬浮用水量，弃去水洗部分；50%乙醇和 95%乙醇各洗脱 4-6 次，优选 5 次，每次洗脱用乙醇量为 2-4 倍体积量的步骤 2 中悬浮用水量。将 50%乙醇和 95%乙醇洗脱部分分别使用旋转蒸发器蒸除溶剂，得干膏状样品备用。

其中，所述的旋转蒸发器优选 EYELA N1001 型旋转蒸发器；所述的洗脱用悬浮用水量和洗脱用乙醇量优选为 3 倍体积量的步骤 2 中悬浮用双蒸水；所述的大孔树脂优选 HP-20 型，优选地，大孔树脂用 180 毫升，玻璃柱规格 30×300 mm。

(4) 制得天然产物混合物备选库：

上述步骤(b)中制得的有机溶剂相部分、步骤(c)中制得的水相大孔树脂 50%乙醇洗脱部分和水相大孔树脂 95%乙醇洗脱部分等 3 部分天然产物混合物备选样品, 构成一份初级原料所提供的 3 种天然产物混合物备选样品(初级原料为中药处方或微生物时, 每一份初级原料提供 6 种天然产物混合物备选样品), 所有初级原料提供的天然产物混合物备选样品(设为 N*种)即组成了本天然产物混合物备选库。

上述天然产物混合物备选库的组成示于图 2

4 天然产物备选样品对靶标蛋白体外抑制活性的测定

步骤(1)中所选取的靶标蛋白若已报道其具有成熟的体外活性测定方法, 在进行靶标蛋白晶体与天然产物混合物备选库中的样品(称为天然产物混合物备选样品)(按照步骤 3 中制备)的浸泡实验之前, 最好将天然产物混合物备选样品对靶标蛋白进行体外活性的测定(靶标蛋白活性测定方法参照相关文献(如 HCV NS5B 蛋白, 参见: Robert A. 等人. *Journal of Virology*, July 2003:7575-7581; SARS 冠状病毒主蛋白酶, 参见: Yang H, Xie W 等人. 2005 *PloS Biol* 3(10):e324)), 从不同的天然产物混合物备选样品中选取对靶标蛋白具有抑制活性的样品(命名为“有抑制活性”的样品), 进行下一步的晶体浸泡或共结晶工作, 这样可以针对性的选取对特定靶标蛋白有抑制活性的天然产物混合物备选样品, 减少下面步骤 5 中晶体浸泡或共结晶的次数, 从而有效减少工作量, 并提高命中率。

不同靶标蛋白的活性测定条件, 可以参考已报道的本领域已知方法进行, 如 HCV NS5B 蛋白的活性测定条件, 参见: Robert A. 等人. *Journal of Virology*, July 2003:7575-7581; SARS 冠状病毒主蛋白酶的活性测定条件, 参见: Yang H, Xie W 等人. 2005 *PloS Biol* 3(10):e324。

对于其他的、试图通过本发明的方法进行功能探索的靶标蛋白(如 SARS 非结构蛋白 nsp10), 可不考虑此一步骤, 而直接实施步骤(5), 以确定是否在靶标蛋白上结合有靶分子, 用来研究该靶标蛋白的功能。

5. 靶标蛋白质的晶体与天然产物混合物备选样品接触

对于已报道具有体外活性测定方法的靶蛋白, 在步骤 4 中得到“有抑制活性”的样品(可能是一种, 也可能是多种, 设为 N1 种)之后, 使用步骤 2 中制备的靶标蛋白的晶体与“有抑制活性”的样品的浸泡实验或者使用步骤 1 得到的经过纯化后的靶标蛋白与“有抑制活性”的样品进行共结晶实验。主要目的是获得靶标蛋白与靶分子(该靶分子为“有抑制活性”的样品中可能存在的靶标蛋白的配体分子)复合体的能够稳定存在, 并具有可供 X 射线衍射实验的衍射能力的晶体。

使用的方法主要包括晶体浸泡和共结晶, 优选晶体浸泡的方法可采用如下三种方式进行:

方式（一）：单独浸泡

(1) 将步骤 4 中得到的 N1 种“有抑制活性”的样品溶解于靶标蛋白的结晶池液（结晶池液是靶蛋白晶体生长的溶液，通常由沉淀剂和缓冲溶液组成）（例如 SARS 主蛋白酶的结晶缓冲液组成为：1.8~12% 的 PEG 6000，3% 二甲基亚砷（DMSO），1mM DTT，100 mM MES 缓冲液（pH 5.0~PH6.0）中，制得晶体浸泡液。

(2) 利用尼龙晶体环等工具，将晶体从结晶池液中取出，分别加入至上述（1）中所述的晶体浸泡液中浸泡 2 小时至 48 小时不等。采用这种方式需要浸泡 N1 次。

方式（二）分组浸泡

特别需要强调的是，步骤 4 中得到的 N1 种“有抑制活性”的样品，其对靶标蛋白的抑制作用不一定是特异性的抑制作用，可能存在假阳性结果，为了减少浸泡次数，可以将这 N1 种“有抑制活性”的样品按照其对靶标蛋白抑制活性的高低进行分组（假设 10 种为一组），然后按照方式（一）进行，这样只需要浸泡 N1/10 次，减少了浸泡次数，提高了筛选效率和命中率。

方式（三）全部混合浸泡

将步骤 4 中得到的 N1 种“有抑制活性”的样品全部混合起来，按照步骤 5 方式（一）中的（1）制备得到 1 种晶体浸泡液，然后再进行晶体浸泡实验，采用这种方式只需要浸泡一次，最大程度的减少了浸泡次数，最大程度的提高了筛选效率和命中率。

对于其他的、试图通过本发明的方法进行功能探索的靶标蛋白（如 SARS 非结构蛋白 nsp10），将天然产物混合物备选样品与靶标蛋白晶体直接采用以上所述的方式进行接触。

6. 晶体 X 射线衍射数据的收集和分析

在步骤 4、5 的基础之上，使用具备较好衍射能力（通常指的是最高分辨率不低于 3Å）的靶标蛋白质的晶体与天然产物混合物备选样品接触后的晶体，进行 X 射线衍射数据的收集和分析。

对于实验室进行的实验，可选用使用 Rigaku MM007 铜靶旋转阳极发生器产生 X 射线的 X 光机；通过 OSMIC 公司的聚焦镜对 X 射线进行聚焦，使其能够达到最高的强度；使用 Rigaku RAXIS IV++ 成像板探测器或 Marreseach 公司的 Mar345dtb 成像系统收集衍射数据。

对于高通量 X 射线晶体学的筛选工作，可使用 Rigaku 公司的 ACTOR 自动晶体安装机械手系统和 X 射线强度更高的 FR-E X 射线衍射仪，以加快实验速度。

进行数据收集时，首先使用 Hampton Research 公司的尼龙晶体环，从晶体浸泡液中获取浸泡一定时间（为了减小浸泡液对靶标蛋白质晶体的损害，

浸泡时间一般选择在 2 小时至 48 小时之间；为了加快速度，浸泡时间控制在 2—6 小时之间）的晶体；并迅速使用 Rigaku 公司的冷却系统或 Oxford Cyrosystem 公司的冷却系统，将晶体在以上所述的两种冷冻系统产生的低温氮气流中冷冻至 -150°C — -180°C ；使 X 射线通过晶体，使用旋进法收集 X 射线衍射数据。

在收集到将靶标蛋白质的晶体与天然产物混合物备选样品接触后的晶体的 X 射线衍射数据后，需要按照下述步骤进行相应的数据处理：

首先，使用 HKL2000 等（包括 Mosflm、D*trek 等）衍射数据处理程序包等本领域已知的常用的数据处理的方法，对上一步骤中收集得到的衍射数据进行处理，获得完整的数据文件；

其次，使用 CNS 或者 CCP4 程序包中的 Phaser、Molrep 等各种分子置换（Molecular Replacement）程序等本领域已知的常用的数据处理方法，利用分子置换（Molecular Replacement）的方法，以靶标蛋白的母体晶体结构为初始模型，得到候选的靶标蛋白与靶分子的复合体的精细三维结构；

再次，使用程序 O、COOT、XtalView 等可视化程序包等本领域已知的常用的数据处理的方法，观测可能的靶标蛋白与抑制剂或者配体小分子复合体的电子密度。如果此时发现，在靶标蛋白相应的活性位点，有不属于蛋白质分子本身的、不属于溶剂分子的、不属于结晶溶液中存在的分子的未知电子密度存在，并且该电子密度与靶标蛋白的活性位点的氨基酸有包括氢键作用、共价作用等在内合理的化学环境，则可以确认此天然产物混合物备选样品中，有抑制剂小分子等配体分子与靶标蛋白结合。

对于其他的希望使用本发明方法进行功能研究的靶标蛋白（例如 SARS 非结构蛋白 nsp10），则需要在整个蛋白质分子周围进行仔细观察，看是否有不属于蛋白质分子本身的、不属于溶剂分子的、不属于结晶溶液中存在的分子的未知电子密度存在，如果有未知电子密度存在的话，则此处可能是一处具有研究价值的活性位点。

7. 鉴定靶分子，获得靶分子与靶蛋白复合体的精细三维结构

一旦在步骤 6 中发现结合在靶标蛋白质分子上的未知电子密度，一般来说可按照下面两种方法进行进一步工作，以确定结合小分子的具体成分等：

（1）对于分辨率较高的复合体的晶体结构，通常指分辨率超过 2.5\AA ，可以通过直接观察分析电子密度，分析结合小分子与靶标蛋白质分子的化学作用，并结合该天然产物混合物备选样品研究，对结合小分子的成分和结构进行合理的推测（可能是一种、也可能是多种），然后通过购买、合成等途径获取可能的小分子，并进行对照实验，以确定小分子的最终成分。

（2）对于分辨率较低的复合体的晶体结构，通常指分辨率低于 2.5\AA 。首先，利用质谱的方法，对蛋白质和天然产物混合物备选样品进行分析，确

定结合小分子的分子量。其次，通过查阅相关文献等手段，对该备选样品的已有的成分进行具体分析，统计在上述用质谱方法确定的分子量左右（通常指分子量 ± 10 以内）可能存在的化合物（可能是一种，也可能是多种）。然后，通过合成、购买等手段获取该种或者这些化合物，并进行对照实验，以确定小分子的最终成分。

8. 靶分子的优化和进一步的生物学实验

一旦按照步骤1—步骤7的方法，从某种天然产物混合物备选样品中获得了靶标蛋白质分子与靶分子结合的复合体的结构，并确定靶分子的具体成分和结构式之后，可以根据获得的复合体结构对靶分子进行结构上的设计和优化。以期得到能够设计并合成出与靶标蛋白质分子活性位点更加匹配、结合能力更高的小分子化合物，称为优化的靶分子。在得到优化的靶分子之后，可以对靶分子、优化的靶分子进行进一步的生物学实验，包括体外活性测定、细胞学实验、动物实验等生物学实验，进一步测定它们的生物学参数（如结合常数、酶动力学参数、细胞毒性等等），以及对活体动物的疗效，以便为进一步的临床实验做好准备。

本发明提供的“从中药等天然产物混合体系中高效、快速获取靶分子与靶蛋白复合体精细三维结构的新方法”，将X射线衍射晶体学的方法应用到天然产物药物研究中，解决了当前天然产物研究中存在的单一成分分离纯化的瓶颈，大大提高了研发效率，使天然产物的研究进入了一个全新的阶段，为当前的制药工业带来了新的契机。

本发明提供的方法与现有方法相比，利用本发明方法进行药物筛选的优点在于：

(1) 首先，该发明方法使用具有结构多样性和生物活性多样性的天然产物混合物作为药物筛选的备选库，解决了目前以蛋白质晶体结构为基础进行药物筛选的方法中的诸多缺点（具体地，对于计算机辅助药物设计中的对接问题较困难、虚拟库中的分子新颖性较差、化合物种类少、化合物成药性差、理论计算对实验验证结果的预示程度低，对于基于片断的筛选中的分子量较小的基团在的分辨率的电子密度中不易与结晶溶液中本身存在的小分子区别、能够稳定存在的分子量在200Da水平的小分子基团的种类有限等）。

(2) 其次，该发明方法使用X射线晶体学等生物物理的方法，可以检测到与靶标蛋白亲和力较弱的配体小分子，因此使得不能用于常规的高通量筛选方法中的简单易得的天然产物粗提物等混合物样品，可以用于高通量的X射线晶体学筛选的方法中，这样就省去了繁琐的天然产物的分离纯化步骤，解决了目前制药工业对于天然产物研究中以分离纯化为特点的效率低、劳动强度大等缺陷。

(3) 该发明方法使用天然产物粗提物的混合物样品作为备选库，可以有

效的减少筛选中体外活性测定次数，减少晶体浸泡或共结晶的次数，大大提高靶标蛋白晶体与天然产物小分子相结合的命中率，从而有效的提高筛选的效率。本发明方法和传统的基于蛋白质晶体结构为基础进行筛选的方法的筛选效率的对比示于下表1（如步骤3所示，设每一份制得的天然产物混合物备选样品中含有 n 种天然产物小分子）

筛选方法		天然产物小分子数目	体外活性测定次数	有抑制活性的样品数目	晶体浸泡或共结晶次数
传统的以蛋白质晶体结构为基础进行筛选的方法		$N \times n$	$N \times n$	$N * (N \gg N1)$	$N * (>> N1)$
本发明的方法	步骤5方式（一）	$N \times n$	$N *$	$N1$	$N1$
	步骤5方式（二）	$N \times n$	$N *$	$N1$	$N1/10$
	步骤5方式（三）	$N \times n$	$N *$	$N1$	1

(4) 使用靶标蛋白的晶体作为“诱饵”，浸泡有活性的天然产物混合物库（形象的比喻为富含药物小分子的池塘），可以把对靶标蛋白有特异性结合包括特异性抑制结合的天然产物小分子从混合物中分离出来（“钓鱼”的过程），获得靶标蛋白质与天然产物小分子结合的精确三维模型，根据其原子、分子的空间配置和大小进行进一步的药物设计与优化。

(5) 使用本发明方法，通过使用靶标蛋白的晶体浸泡天然产物混合物备选样品，可以从天然产物混合物中得到与靶标蛋白结合最好的包括抑制剂在内的配体小分子（竞争结合的原理），优化选择了包括抑制剂在内的小分子配体。众所周知，对于每一种天然产物粗提物来说，里面含有很多结构类似的成分，以紫草为例，紫草的有效成分主要是紫草素及其衍生物，当我们用靶标蛋白的晶体浸泡经过紫草提取物样品时，紫草中含有的紫草素及其各种衍生物竞争结合靶标蛋白的活性位点，最后能够结合到靶标蛋白活性位点的，一定是与靶标蛋白结合力最好的配体小分子。

附图说明

为了使本发明的内容更容易被清楚的理解，下面根据本发明的具体实施例并结合附图，对本发明作进一步详细的说明，其中

图 1 本发明方法的实验流程图

图 2 本发明的天然产物混合物备选库的组成示意图

图 3 曲线 1—曲线 11 分别为金银花、连翘、野菊花、芦根、栀子、四季青、玄参、马勃、射干、鸭跖草、地骨皮的水相 50%乙醇洗脱样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：100 μ g/mL。曲线 12 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。

图 4 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2—曲线 7 分别为山豆根、知母、穿心莲、鱼腥草、大青叶、黄芩的水相 50%乙醇洗脱样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：100 μ g/mL。

图 5 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2—曲线 11 分别为玄参、金银花、穿心莲、四季青、鸭跖草、地骨皮、芦根、大青叶、马勃、黄芩的乙酸乙酯萃取样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：100 μ g/mL。

图 6 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2—曲线 8 分别为连翘、射干、野菊花、鱼腥草、山豆根、知母、栀子的乙酸乙酯萃取样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：100 μ g/mL。

图 7 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2—曲线 6 分别为鱼腥草、山豆根、黄芩、知母、栀子的乙酸乙酯萃取样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：10 μ g/mL

图 8 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2 和曲线 3 分别为川藏香茶菜水相 50%乙醇洗脱样品和川藏香茶菜乙酸乙酯萃取样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：100 μ g/mL

图 9 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2 和曲线 3 分别为川藏香茶菜水相 50%乙醇洗脱样品和川藏香茶菜乙酸乙酯萃取样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度：10 μ g/mL

图 10 本发明的 SARS 主蛋白酶活性位点附近电子密度对比图。a 为 SARS 主蛋白酶母体活性位点的电子密度显示；b 为 SARS 主蛋白酶与川藏香茶菜乙酸乙

酯萃取样品浸泡后晶体活性位点的电子密度显示；c 为 SARS 主蛋白酶与川藏香茶菜丙素浸泡后晶体活性位点的电子密度显示。图中所示的电子密度，暗色（白色圆圈外）的为 $2fofc$ 电子密度，为 1.0σ ；亮色（白色圆圈内）的为 $1fofc$ 差值电子密度，为 2.5σ 。

图 11 本发明的川藏香茶菜乙酸乙酯萃取样品中的存在的小分子的电子密度与 SARS 主蛋白酶活性位点 Cys145 残基电子密度形成共价键的显示图。蛋白质分子用飘带图表示，活性位点的 Cys145 残基用球棍模型表示；圆圈 2 中电子密度表示 Cys145 的电子密度，圆圈 1 中电子密度表示未知小分子的电子密度。

图 12 本发明的 SARS 主蛋白酶与川藏香茶菜提取物乙酸乙酯萃取样品作用前后的质谱图。左边的峰表示 SARS 主蛋白酶本身的质谱峰；右边的峰表示 SARS 主蛋白酶与川藏香茶菜乙酸乙酯萃取样品作用后的质谱峰。

图 13 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2 为川藏香茶菜丙素样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度 $20\ \mu\text{M}$ 。

图 14 曲线 1 为对照（不加入其它样品的 SARS 主蛋白酶的活性曲线图）。曲线 2 为川藏香茶菜丙素样品对 SARS 主蛋白酶的活性抑制曲线，样品浓度 $2\ \mu\text{M}$ 。

图 15 是本发明中所述的对于 HCV NS5B 晶体进行改进后的结果。其中 a 为按照文献(Ago, Adachi 等人, 1999)中描述的结晶溶液进行晶体生长后 HCV NS5B 的晶体，b 为该文献中的结晶溶液进行改进后进行晶体生长后 HCV NS5B 的晶体

图 16 A1-A12 分别为紫草、板蓝根、香附、荔枝核、山药、地骨皮、石韦、天花粉、百部、辛夷、桑枝、虎帐的乙酸乙酯样品对 HCV NS5B 蛋白的活性抑制图、A13 为对照（不加入其它样品的 HCV NS5B 活性图）。T5：时间为 5 分钟时取样、T10：时间为 10 分钟时取样、T15：时间为 15 分钟时取样、T20：时间为 20 分钟时取样。a： $20\ \mu\text{g/mL}$ ：12 个中药乙酸乙酯样品浓度为 $20\ \mu\text{g/mL}$ 。b： $10\ \mu\text{g/mL}$ ：12 个中药乙酸乙酯样品浓度为 $10\ \mu\text{g/mL}$ 。

图 17 HCV NS5B 蛋白 His502 活性位点附近电子密度对比图。a 为 HCV NS5B 蛋白母体活性位点的电子密度；b 为 HCV NS5B 蛋白的晶体浸泡紫草乙酸乙酯萃取样品后的 HCV NS5B 蛋白活性位点处的电子密度显示；c 为 b 旋转 90° 后的显示。图中所示的电子密度，暗色（白色圆圈外）的为 $2fofc$ 电子密度，

为 1.0 σ ; 亮色 (白色圆圈内) 的为 1fofc 差值电子密度, 为 2.5 σ 。

图 18. 乙酰紫草素小分子与未知电子密度的吻合。1 球棍为蛋白质的氨基酸残基, 2 表示未知电子密度, 3 球棍为乙酰紫草素的结构式。

具体实施方式

实施例 1

天然产物混合物备选库中的样品的制备

川藏香茶菜 (*I. pharicus* (Prain) Murata) 备选样品的制备

1 制备提取物浸膏:

取 500 克干燥粉碎的川藏香茶菜药材, 用 4500 毫升 95% 乙醇回流提取, 提取 3 次, 每次 2 小时, 合并提取液, 浓缩得浸膏;

2 制得乙酸乙酯部分:

称取此浸膏 10.5 克, 用 85 毫升双蒸水悬浮, 倾入 1000 毫升分液漏斗中, 用乙酸乙酯萃取。每次萃取用乙酸乙酯 850 毫升, 充分振摇, 静置 6 小时分层。共萃取 6 次, 合并各次乙酸乙酯萃取物。萃取后得到乙酸乙酯相和水相。乙酸乙酯萃取物用 EYELA N1001 型旋转蒸发器蒸除溶剂, 得干膏状样品备用;

3 制得大孔树脂 50% 乙醇和 95% 乙醇洗脱部分:

萃取后的水相部分用 EYELA N1001 型旋转蒸发器蒸除水中溶解的少量乙酸乙酯, 用 HP-20 型大孔树脂柱层析分离 (大孔树脂用 170 毫升, 玻璃柱规格 30×300 mm), 依次用水、50% 乙醇、95% 乙醇为流动相洗脱。水洗脱 8 次, 每次洗脱 250 毫升, 弃去水洗部分。50% 乙醇和 95% 乙醇各洗脱 5 次, 每次洗脱 300 毫升。每部分均使用 EYELA N1001 型旋转蒸发器蒸除溶剂, 得干膏状样品备用;

4 得到川藏香茶菜备选样品:

上述步骤 (2) 制得的乙酸乙酯相部分、步骤 (3) 制得的 50% 乙醇洗脱部分和 95% 乙醇洗脱部分等 3 部分即组成了初级原料川藏香茶菜提供的 3 份混合物备选样品。

其它 17 种 (鱼腥草, 野菊花, 大青叶, 芦根, 金银花, 连翘, 黄芩, 知

母，四季青，穿心莲，鸭跖草，栀子，射干，山豆根，马勃，玄参，地骨皮）初级原料提供的备选样品的制备与川藏香茶菜备选样品的制备基本相同。

实施例 2

以 SARS 冠状病毒主蛋白酶的晶体结构为基础，从实施例 1 的天然产物混合物备选样品中快速获取川藏香茶菜丙素与 SARS 主蛋白酶复合体精细三维结构的方法。包括如下步骤：

1. 靶标蛋白质的选取：SARS 冠状病毒主蛋白酶

2003 年，全球大范围爆发了 SARS 疫情，而引起 SARS 的病原体最终被确认为一种未知的冠状病毒—SARS 冠状病毒（Drosten, C., Gunther, S., Preiser, W. 等人. *N Engl J Med.* 2003, 348:1967-76）。进一步的研究发现，SARS 冠状病毒的基因组编码了两个大的复制酶多蛋白 (replicase polyproteins) pp1a(486kDa) 和 pp1ab(790kDa)，这两个蛋白是由占冠状病毒 2/3 到 3/4 的基因组编码的。这两个蛋白被水解后产生病毒复制复合体的很多功能亚基。在这一水解过程中，SARS 冠状病毒的主要蛋白水解酶 (main protease, 简称为主蛋白酶, 缩写为 M^{pro}, 分子量 33.8 kDa, 有时也称为 3C-Like Protein) 起到了非常关键的作用。主蛋白酶也是 pp1a 和 pp1ab 中的一段，它的释放是通过自催化水解完成的，自催化水解发生在该蛋白酶旁侧的 Gln(Ser, Ala) 位点，通过反式剪接(双分子反应)完成。在主蛋白酶的作用下，复制酶多蛋白 pp1a 和 pp1ab 被水解成十多个功能肽段，从而进一步发挥作用。如果能够抑制 SARS 冠状病毒主蛋白酶的水解作用，那么将会有效地抵御 SARS 冠状病毒对人体的侵染。因此，SARS 冠状病毒的主蛋白酶是抗 SARS 药物筛选的一个主要靶标蛋白。SARS 冠状病毒主蛋白酶（氨基酸序列见序列表部分 1）在大肠杆菌菌株 BL21 (DE3) 中表达后进行进一步分离纯化（参见：Yang H et al. 2003. The Crystal Structures of SARS Virus Main Protease M^{pro} and Its Complex with an Inhibitor. *PNAS*, 100(23): 13190-13195）。

2 SARS 冠状病毒主蛋白酶的晶体生长

SARS 病毒的主蛋白酶的结晶仍然参见 Yang H et al. 2003. The Crystal Structures of SARS Virus Main Protease M^{pro} and Its Complex with an Inhibitor. *PNAS*, 100(23): 13190-13195。具体地：采用悬滴气相扩散法，

在 289K—297K 下，用组织培养板进行的。用 Hampton Research 公司的晶体生长试剂盒 (Crystal Screening Kit I & II, PEG/ION Kit, PEG 6000 Grid Kit 和 (NH₄)₂SO₄ Kit) 对初始结晶条件进行筛选。在 PEG 6000 Grid Kit 的 C3 (20% PEG 6000, 0.1 M MES (pH 6.0)) 条件下的液滴中有红色球状沉淀，并在沉淀的边缘出现微晶。在此结晶条件下，我们尝试了不同 pH 值的缓冲液(例如 pH 值从 5.5 到 6.5)、不同聚合度的 PEG(如从 PEG 3350 到 PEG 10000 之间进行)、不同蛋白浓度 (例如 2mg/ml 到 30mg/ml) 和各种添加剂 (使用 Hampton Research 的 Additive Screen Kit (添加剂筛选试剂盒) 进行) 等来优化晶体。最后，我们在 1.8—12% 的 PEG 6000, 3% 二甲基亚砷 (DMSO), 1mM DTT, 100 mM MES 缓冲液 (pH 6.0) 以及蛋白浓度为 9—12mg/ml 的情况下获得了适合衍射的晶体。典型的晶体尺寸为 0.03 × 0.3 × 0.5 mm。

3 天然产物混合物备选样品的制备

参照实施例 1

4. 天然产物混合物备选样品对 SARS 主蛋白酶抑制活性的测定

SARS 主蛋白酶的活性测定是使用荧光底物 MCA-AVLQSGFR-Lys (Dnp)-Lys-NH₂ (纯度大于 95%，上海吉尔生化有限公司) 来完成的。该荧光底物的氨基酸序列来源于 SARS 主蛋白酶的 N 端自剪切序列。用于荧光强度测定的仪器为 Fluoraskan Ascent 荧光仪 (ThermoLabsystems, Helsinki, Finland)，激发光和发射光的波长分别为 320nm 和 405nm。

将实施例 1 中得到的干膏状天然产物混合物备选样品溶于二甲基亚砷 (DMSO) 中使其终浓度为 50mg/mL

在缓冲溶液 (50mM Tris-HCl (pH 7.3), 1mM EDTA (含或不含 DTT)) 中加入 SARS 主蛋白酶 (终浓度 0.5 μM)，天然产物混合物备选样品的 DMSO 溶解物 (如川藏香茶菜乙酸乙酯萃取物样品的 DMSO 溶解物) 使其终浓度为：100 μg/mL，底物浓度为 20 μM，298 K 放置 10 分钟后，迅速加入荧光标记底物 (MCA-AVLQ↓SGFRL(DNP)L-NH₂，终浓度 20 μM)。激发波长和发射波长分别为 330 nm 和 395 nm，温度保持 298 K，每 2 秒钟记录一次荧光读数。

对照：不加入备选样品，其余条件相同。结果示于图 3—图 9 中。

图 3—图 6 以及图 8 为 18 种中药制得的天然产物混合物备选样品对 SARS 主蛋白酶抑制活性的测定结果，其中每种备选样品的浓度为 $100\ \mu\text{g/mL}$ 。

由以上图 3—图 9 可以看出，其中大部分备选样品均对 SARS 主蛋白酶有抑制活性，其中抑制活性较好的有 6 种，分别是黄芩、鱼腥草、山豆根、栀子、知母、川藏香茶菜的乙酸乙酯萃取物样品。

图 7 和图 9 是以上 6 种抑制活性较好的：黄芩、鱼腥草、山豆根、栀子、知母、川藏香茶菜的乙酸乙酯萃取物样品的浓度稀释至 $10\ \mu\text{g/mL}$ 时的对 SARS 主蛋白酶抑制活性测定结果，可以发现，将黄芩、鱼腥草、山豆根、栀子、知母、川藏香茶菜的乙酸乙酯萃取物浓度为 $10\ \mu\text{g/mL}$ 时，对 SARS 主蛋白酶仍有较强的抑制活性。

这就说明在这 5 种中药的乙酸乙酯萃取样品中可能含有某种成分能够有效的抑制 SARS 主蛋白酶的活性，也说明这 5 种中药有可能成为治疗 SARS 冠状病毒的有效药物。为了了解它们对 SARS 主蛋白酶抑制作用确切的分子机理，我们可以选择这 5 种中药的乙酸乙酯萃取样品进行 SARS 主蛋白酶的晶体浸泡或共结晶实验。

5. SARS 主蛋白酶晶体浸泡川藏香茶菜乙酸乙酯萃取样品

以川藏香茶菜乙酸乙酯萃取样品为例，由于川藏香茶菜乙酸乙酯萃取样品中含有不只一种的化合物，若采用共结晶的方法，靶蛋白结晶的影响较大，无法获得衍射质量好、数量充足的靶标蛋白质晶体，甚至可能无法生长晶体，因此主要采用晶体浸泡的方法用来获得 SARS 主蛋白酶与川藏香茶菜乙酸乙酯萃取物样品接触之后的晶体。

为了尽可能的在保证晶体质量的基础上，提高浸泡时备选样品中小分子化合物的浓度，采用两种方法制备晶体浸泡液。

方法一，首先将川藏香茶菜提取物乙酸乙酯萃取物 DMSO 的溶解物（按照步骤 4 制备）10 倍稀释至晶体生长池液（1.8~12% 的 PEG 6000，3% 二甲基亚砷（DMSO），1mM DTT，100 mM MES 缓冲液（pH 5.0~PH6.0））中，13000—15000rpm 高速离心，之后取上清液，得浸泡液 1，备用。

方法二，将干膏状川藏香茶菜提取物乙酸乙酯萃取物（按照实施例 1 中

制备)直接溶解于晶体生长的池液中(1.8~12%的PEG 6000, 3%二甲基亚砷(DMSO), 1mM DTT, 100 mM MES 缓冲液(pH 5.0~PH6.0)),待充分溶解后,13000-15000rpm高速离心,之后取上清液,得浸泡液2,备用。

为了保证晶体浸泡时,不至于使过高的小分子浓度影响晶体的衍射质量,因此将浸泡液1和2各自稀释至原液浓度的10倍,备用。

浸泡时,利用尼龙晶体环等工具,将已生长好的SARS主蛋白酶的蛋白质晶体从结晶池液中取出,分别加入浸泡液1和2以及相应的10倍稀释的浸泡液中,浸泡4-6小时后,制得SARS主蛋白酶与川藏香茶菜乙酸乙酯萃取物接触后的晶体。

6. 晶体衍射数据的收集和分析

在步骤5的基础之上,使用具备较好衍射能力(分辨率优选大于2.6Å)的晶体,进行X射线衍射数据的收集和分析。

在实验室进行的筛选实验,可选用使用Rigaku MM007铜靶旋转阳极发生器产生X射线的X光机;通过OSMIC公司的聚焦镜对X射线进行聚焦,使其能够达到最高的强度;使用Rigaku RAXIS IV++成像板探测器或Marresearch公司的Mar345dtb成像系统收集衍射数据。

对于高通量的药物筛选工作,可使用Rigaku公司的ACTOR自动晶体安装机械手系统和FR-E X射线衍射仪。

进行数据收集时,首先使用Hampton Research公司的尼龙晶体环,从浸泡液中获取浸泡一定时间(通常为2-48小时)的晶体;并迅速使用Rigaku公司的冷却系统或Oxford Cyrosystem公司的冷却系统,将晶体在以上所述的两种冷冻系统产生的低温氮气流中冷冻至-150°C--180°C;使X射线通过晶体,使用旋进法收集X射线衍射数据。

选择衍射质量好的晶体(分辨率最好大于2.6Å)进行数据收集、处理,检测是否有小分子的结合。数据收集的统计如下表2:

表2: SARS主蛋白酶母体、SARS主蛋白酶与川藏香茶菜乙酸乙酯萃取物样品浸泡后的复合体数据收集统计

统计参数		
	SARS 主蛋白酶母体晶体	SARS 主蛋白酶与川藏香茶 菜乙酸乙酯萃取物样品浸 泡后的晶体
空间群	C2	C2
晶胞参数 (Å, °)	a = 108.4 b = 81.8 c = 53.6 β = 104.7	a = 107.8 b = 82.5 c = 52.9 β = 104.9
分辨率 (Å)	50-1.50 (1.55-1.50)	50.0-3.0 (3.1-3.0)
所有衍射点	219, 294	63, 908
独立衍射点	63, 241	9, 039
冗余度	3.5 (2.5)	7.1 (6.4)
I/ σ (I)	30.4 (2.0)	13.9 (3.1)
Rmerge (%)	3.8 (28.6)	8.2 (67.5)
完整度 (%)	87.7 (40.6)	99.9 (100.0)

通过观察相同角度的 SARS 主蛋白酶母体和复合体活性位点电子密度图 (图 10) (暗色是蛋白质本身的电子密度), 可以发现, 在复合体的活性位点能够观察到清晰的非蛋白质本身的电子密度 (白色圆圈内亮色表示), 提示可能有小分子结合在蛋白质分子上。图 10a 中显示的是 SARS 主蛋白酶母体结构的活性位点附近的电子密度, 图 10b 为 SARS 主蛋白酶与川藏香茶菜乙酸乙酯萃取物样品浸泡后晶体活性位点附近的电子密度。

仔细观察该处的电子密度 (如图 10 所示), 我们可以发现存在于川藏香茶菜乙酸乙酯萃取物样品中的未知小分子的电子密度与 SARS 主蛋白酶活性位点处 Cys145 的电子密度紧密结合在一起, 提示我们这应该是小分子与 Cys145 的 γ 硫原子形成了共价键, 从而抑制了 SARS 主蛋白酶的活性 (如图 11 所示)。

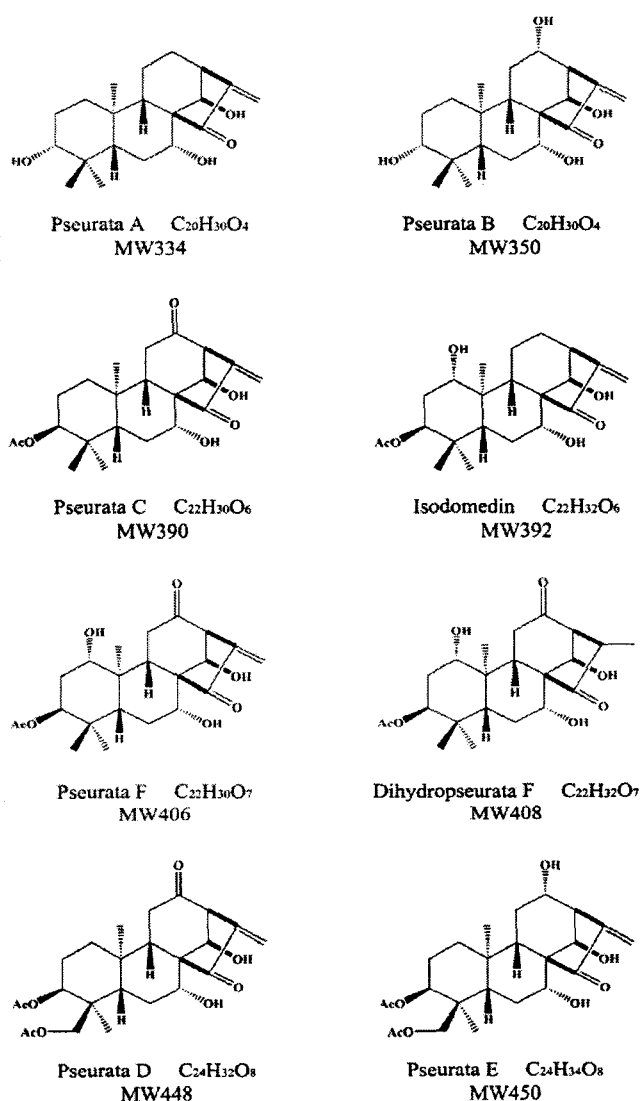
7. 结合小分子的分析与鉴定

我们将 SARS 主蛋白酶 (浓度为 12mg/ml 的 SARS 主蛋白酶水溶液) 与川藏香茶菜乙酸乙酯萃取物样品按照质量浓度比 1:1-1:4 的比例混合, 并在 4° C -16° C 下作用 6-18 小时。将作用后的样品, 13,000-15,000rpm 高速离

心，弃去沉淀后备用于质谱分析。质谱采用 MALDI (matrix-assisted laser-desorption ionization) 谱对样品进行分析，结果参见图 12。

从质谱的结果我们可以发现，在加入川藏香茶菜乙酸乙酯萃取物样品后，分子量出现了明显的偏移，这主要是因为蛋白质分子与川藏香茶菜乙酸乙酯萃取物样品中的小分子结合的原因。通过质谱结果可以计算（用较高的质谱峰值减去较低的质谱峰）发现，结合的小分子分子量约为 390Da。

通过分析已有的川藏香茶菜成分，我们发现，在该分子量水平上，主要存在如下所示的化合物



我们可以发现在川藏香茶菜的主要成分中，在该分子量范围内，主要成分以二萜类化合物为主，包括川藏香茶甲、乙、丙、丁、戊、己素（分别对应上面 PseurataA、PseurataB、PseurataC、PseurataD、PseurataE、PseurataF），Isodomedin 和 Dihydropseurata）等。又由于结合在 SARS 主蛋白酶活性口袋处的小分子，和 Cys145 的 γ 硫原子形成了共价键，推知该小分子为 Michael 受体类小分子，具有 α ， β 共轭不饱和双键。

基于分子量和化学环境的分析，得出结论：结合的小分子为川藏香茶菜

丙素或者 Isodomedin。

为了进一步验证，我们从中国科学院云南昆明植物所获得了川藏香茶菜丙素及 Isodomedin 的纯化合物，按照以上所述的方法制备川藏香茶菜丙素以及 Isodomedin 与 SARS 主蛋白酶复合体的晶体，并进行了数据收集（数据统计结果示于表 3）。通过观察电子密度，我们只在川藏香茶菜丙素浸泡过的 SARS 主蛋白酶的晶体发现了相似的电子密度，如图 10. c 所示。因此可以确定，从川藏香茶菜乙酸乙酯萃取物样品中发现的与 SARS 主蛋白酶相结合的抑制剂小分子应为川藏香茶菜丙素。

在川藏香茶菜乙酸乙酯萃取物样品和川藏香茶菜丙素浸泡过的 SARS 主蛋白酶的晶体活性位点，未知电子密度还有所区别，主要是因为不同晶体的数据质量不同，而且在川藏香茶菜乙酸乙酯萃取物样品中川藏香茶菜丙素的相对浓度比纯品肯定要低，因此可能占有率也有所下降。

因此，可以确定结合的小分子即为川藏香茶菜丙素。

表 3: SARS 主蛋白酶母体、SARS 主蛋白酶与川藏香茶菜丙素的复合体数据收集统计

统计参数	SARS 病毒主蛋白酶及相关复合体	
	SARS 病毒主蛋白酶母体	SARS 病毒主蛋白酶与川藏香茶菜丙素复合体
空间群	<i>C</i> 2	<i>C</i> 2
晶胞参数 (Å, °)	a = 108.4	a = 107.3
	b = 81.8	b = 83.3
	c = 53.6	c = 53.3
	β = 104.7	β = 105.2
分辨率 (Å)	50-1.50 (1.55-1.50)	50.0-2.4 (2.5-2.4)
所有衍射点	219, 294	176, 898
独立衍射点	63, 241	19, 456
冗余度	3.5 (2.5)	9.1 (5.3)
I/ σ (I)	30.4 (2.0)	15.2 (4.1)
Rmerge (%)	3.8 (28.6)	5.4 (56.6)
完整度 (%)	87.7 (40.6)	99.9 (100.0)

由此，我们获得了 SARS 冠状病毒主蛋白酶与川藏香茶菜丙素复合体的三维结构。

这里，我们可以看到本发明的新方法与传统药物筛选方法相比的巨大优势。按照我们的方法，(1) 仅需要筛选一次就可以从川藏香茶菜乙酸乙酯萃取物中快速发现能够与 SARS 主蛋白酶结合的小分子，(2) 该小分子是川藏香茶菜乙酸乙酯萃取物中与 SARS 的结合能力最好的小分子，(3) 无需在药物筛选时制备众多的纯品化合物；而如果利用传统方法，不仅需要在进行筛选前分离提取制备出至少 8 种化合物，而且需要至少进行 8 次的活性测定、并进行至少 8 次晶体浸泡实验等，才有可能发现结合能力最好的小分子化合物。

8 进一步的生物学实验：川藏香茶菜丙素对 SARS 主蛋白酶抑制活性的测定及其抑制动力学参数测定

我们使用纯品化合物小分子（川藏香茶菜丙素），测定了其对于 SARS 主蛋白酶的抑制活性。川藏香茶菜丙素对 SARS 主蛋白酶抑制活性的测定按照如下步骤进行：在缓冲溶液（50mM Tris-HCl (pH 7.3), 1mM EDTA (含或不含 DTT)) 中加入 SARS 主蛋白酶（终浓度 0.5 μ M），川藏香茶菜丙素（终浓度为：20 μ M），底物浓度为 20 μ M，298 K 放置 10 分钟后，迅速加入荧光标记底物（MCA-AVLQ \downarrow SGFRL(DNP)L-NH₂, 终浓度 20 μ M）。激发波长和发射波长分别为 330 nm 和 395 nm，温度保持 298 K，每 2 秒钟记录一次荧光读数（结果示于图 13）。然后改变川藏香茶菜丙素的浓度，在 2 μ M 测定其对 SARS 主蛋白酶的抑制活性（结果示于图 14）。对照：不加入备选样品，其余条件相同。活性测定结果见图 13—图 14。从图 13—图 14 中可以发现，20 μ M 的川藏香茶菜丙素对 SARS 主蛋白酶具有较强的抑制活性，当其浓度为 2 μ M 时对 SARS 主蛋白酶仍然有抑制活性。这就说明中药川藏香茶菜里面含有的川藏香茶菜丙素能够有效的抑制 SARS 主蛋白酶的活性，也说明川藏香茶菜丙素有可能成为治疗 SARS 冠状病毒的有效药物。

同时，我们测定了川藏香茶菜丙素对 SARS 主蛋白酶的精确的抑制动力学参数。测定原理如下：通过将时间依赖型抑制剂的反应过程曲线按照一阶指数形式拟合，即可以得到表观一级抑制速率常数 (k_{obs})。其中 P 是产物的荧光强度， v_0 是初始反应速率， t 为时间， D 是为了解决零时刻荧光强度不为零而引入的位移参量。通过公式 (1) 可求得 k_{obs} 。在公式 (2) 中，利用 $1/k_{obs}$ 对 $1/[I]$ 进行线性拟合，即可求得 K_i 和 k_3 。其中 $[I]$ 是抑制剂的浓度， $[S]$ 是底物的浓度， K_m 是米氏常数， k_3 是抑制剂使酶失活的反应速率常数， K_i 是抑制剂与酶之间非共价结合的平衡常数。

$$P = (v_0 / k_{obs})(1 - \exp(-k_{obs}t)) + D \quad (1)$$

$$\frac{1}{k_{obs}} = \frac{1}{k_3} + \frac{K_i}{k_3} (1 + [S]/Km) \cdot \frac{1}{[I]} \quad (2)$$

在川藏香茶菜丙素动力学参数测定实验中，酶活测定方法与实施例 2 步骤 4 基本相同，SARS 主蛋白酶的浓度（[E]表示）：[E]=0.2 μ M，川藏香茶菜丙素的浓度（用[I]表示）在[I]=25~100[E]范围内，选取 8 个不同值测定 8 组荧光强度，将荧光强度结果对时间的反应曲线用 Origin 程序进行非线性拟合，由此来计算川藏香茶菜丙素的 K_i 和 k_3 的值。通过如上所述的方法，我们测得川藏香茶菜丙素的动力学参数为：

$$k_3: 7.5 \pm 0.7 \cdot 10^{-3} \text{ s}^{-1}$$

$$K_i: 10.3 \pm 1.7 \mu\text{M}$$

实施例 3

紫草、板蓝根、香附、荔枝核、山药、地骨皮、石韦、天花粉、百部、辛夷、荔枝、虎帐 12 种药材的乙酸乙酯萃取样品的制备与实施例 1 中川藏香茶菜乙酸乙酯萃取样品的制备基本相同。

实施例 4 以 HCV NS5B 蛋白的晶体结构为基础，从实施例 3 的备选样品中快速获取乙酰紫草素与 HCV NS5B 蛋白复合体的精细三维结构的新方法，包括如下步骤：

1. 靶标蛋白质的选取：HCV NS5B 蛋白

HCV (Hepatitis C Virus) 属于黄病毒科 (flaviviridae)，是丙型肝炎的病原体。由于缺乏疫苗和有效的治疗药物，HCV 感染已经成为一种世界性的流行病。由于缺乏合适的动物模型和组织培养系统，目前还不能有效的进行病毒的扩增和相应的病毒学研究，因此开发针对 HCV 的特效药物成为一个世界性的难题。随着一些 HCV 复制过程重要蛋白以及这些蛋白与相关配体或抑制剂的精确三维结构的解析，基于这些蛋白的三维结构设计、筛选治疗 HCV 感染的药物，成为目前开发治疗 HCV 药物的重要手段。NS5B 是 HCV 的 RNA 依赖的 RNA 聚合酶 (RNA-dependent RNA polymerase, RdRp)，在 HCV 复制过程中起着关键的作用 (Behrens, S. E 等人 *EMBO J* 1996: 15(1)12-22)。但全长的 HCV NS5B 蛋白溶解性较差，Ferrari (Ferrari, 等人 1999) 等为提高 NS5B 的溶解性，在构建 NS5B 表达载体时切除了 C 端的 21 个氨基酸 (称为 HCV NS5B Δ 21)，在 *E. Coil* 中获得了保持酶活性的高度可溶的蛋白。虽然在 HCV 不同的基因型和亚型中，NS5B 的蛋白质序列不尽相同，但是相关活性位点却非常保守，这就为我们从已解析结果的 NS5B 出发，筛选针对不同基因型的药物提供了重要的依据。

到目前为止，在 NCBI PubMed 上可以检索到与 HCV NS5B 抑制剂相关的文章已有 53 篇，但是在蛋白质结构数据库 (Protein Database Bank, PDB) 中

能够检索到的 NS5B 与抑制剂结合的复合物的三维结构只有 15 个，也就是说相当一部分抑制剂的开发只进行到体外酶活测定，而没有得到相应的抑制剂与蛋白质结合的复合体三维结构，因此没有给出该抑制剂抑制 HCV NS5B 蛋白确切的分子机制。

本发明以 HCV NS5B 蛋白为靶标蛋白，介绍从中药等天然产物混合物备选库中快速获取靶分子与靶蛋白精细三维结构，高效地进行抗 HCV 药物筛选的新方法。

构建在 pET-21b 载体上的 HCV NS5B Δ 21: genotypelb, isolate BK (丙型肝炎病毒基因型 1b 分离株 BK 缺失 C 端 21 个残基的 NS5B: 氨基酸序列见序列表部分 2) 质粒获赠于中科院生物物理所高光侠教授实验室，该质粒在大肠杆菌菌株 BL21 (DE3) 中表达后进行进一步分离纯化参见 (Ago, 等人 1999)。

2. HCV NS5B 蛋白的晶体生长

HCV NS5B 的晶体生长参考文献 (Ago, 等人 1999)，并进行了初步的实验，结果发现在文献中报道的蛋白结晶条件下 (21 - 28% (w/v) polyethylene (PEG) 4000, 0.2 - 0.35 M ammonium acetate, 0.1 M sodium acetate, and 0.02 M TES (pH 6.0 - 7.5).) 虽然能够生长出文献报道的晶体，但是该晶体不但尺寸较小，孪晶现象严重，而且很难得到衍射质量好的晶体 (图 15. a)，不利于我们进行进一步的晶体浸泡和共结晶试验，从而进行相应的药物筛选工作。因此我们对结晶条件进行了相应的改进，结晶条件如下：20-25% (w/v) PEG4000, 10% (v/v) 甘油, 5mM DTT, 50mM MES pH5.0~6.0。在该结晶条件和 25~30mg/ml 的蛋白浓度下，能够得到尺寸较大 (通常能够达到 $3 \times 1.5 \times 1$ mm, 图 15. b)、衍射能力很强的晶体，为进一步的药物筛选工作提供了良好的晶体学基础。

3. 天然产物混合物备选样品的制备

参照实施例 3

4. 天然产物混合物备选样品对 HCV NS5B 抑制活性的测定

HCV NS5B 的体外活性测定方法参见 (Robert A. 等人. Journal of Virology, July 2003:7575-7581) 天然产物混合物备选样品对 HCV NS5B 的抑制活性如下测定：在缓冲溶液 20mM Tris-HCl (pH 7.5), 5mM MgCl₂, 0.5mM MnCl₂, 1mM DTT, 1mg/ml polyA, 1mg/ml OligodT (18T), 1mg/ml UTP, 1mg/ml P³²-UTP 中加入 HCV NS5B 蛋白 (终浓度 0.8 μ g/ml)，天然产物混合物备选样品的 DMSO 溶解物 (终浓度 20 μ g/ml、10 μ g/ml)，每 5 分钟取样一次，将到达取样时间的 2-4 μ l 反应后的溶液滴于处理好的 DEAE 滤纸上以终止反应。当反应完成后，用 2 \times SSC 溶液和 95% 乙醇先后洗涤 DEAE 滤纸 4 次，以除去残余 P³²-UTP，晾干后，使用 PhosphorImage 或者底片曝光，以测定中药混合物备选样品对 HCV NS5B 的抑制活性，结果示于图 16。由

图 16 可以看出紫草、山药、虎帐的乙酸乙酯萃取样品对 HCV NS5B 蛋白有较好的抑制活性。选定这些样品进行下一步的晶体浸泡工作。

5. HCV NS5B 蛋白的晶体浸泡紫草乙酸乙酯萃取样品

采用晶体浸泡的方法使 HCV NS5B 蛋白的晶体与紫草乙酸乙酯萃取样品接触。

HCV NS5B 蛋白的晶体浸泡液的制备与实施例 2 中 5 SARS 主蛋白酶的晶体浸泡液的制备方法基本相同。HCV NS5B 蛋白的晶体浸泡紫草乙酸乙酯萃取样品的方法与 SARS 主蛋白酶的晶体浸泡川藏香茶菜乙酸乙酯萃取样品的方法基本相同。

6. 晶体衍射数据的收集和分析

HCV NS5B 蛋白的晶体与紫草乙酸乙酯萃取样品接触后的晶体（简称复合体）的数据分析与实施例 2 步骤 6 中基本相同。

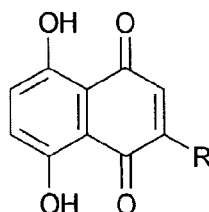
通过观察相同角度的 HCV NS5B 蛋白母体的晶体和复合体活性位点电子密度图（图 17），可以发现，在复合体的活性位点能够观察到清晰的非蛋白质本身的电子密度（白色圆圈内亮色部分），提示可能有小分子结合在蛋白质分子上。

7 结合小分子的分析与鉴定

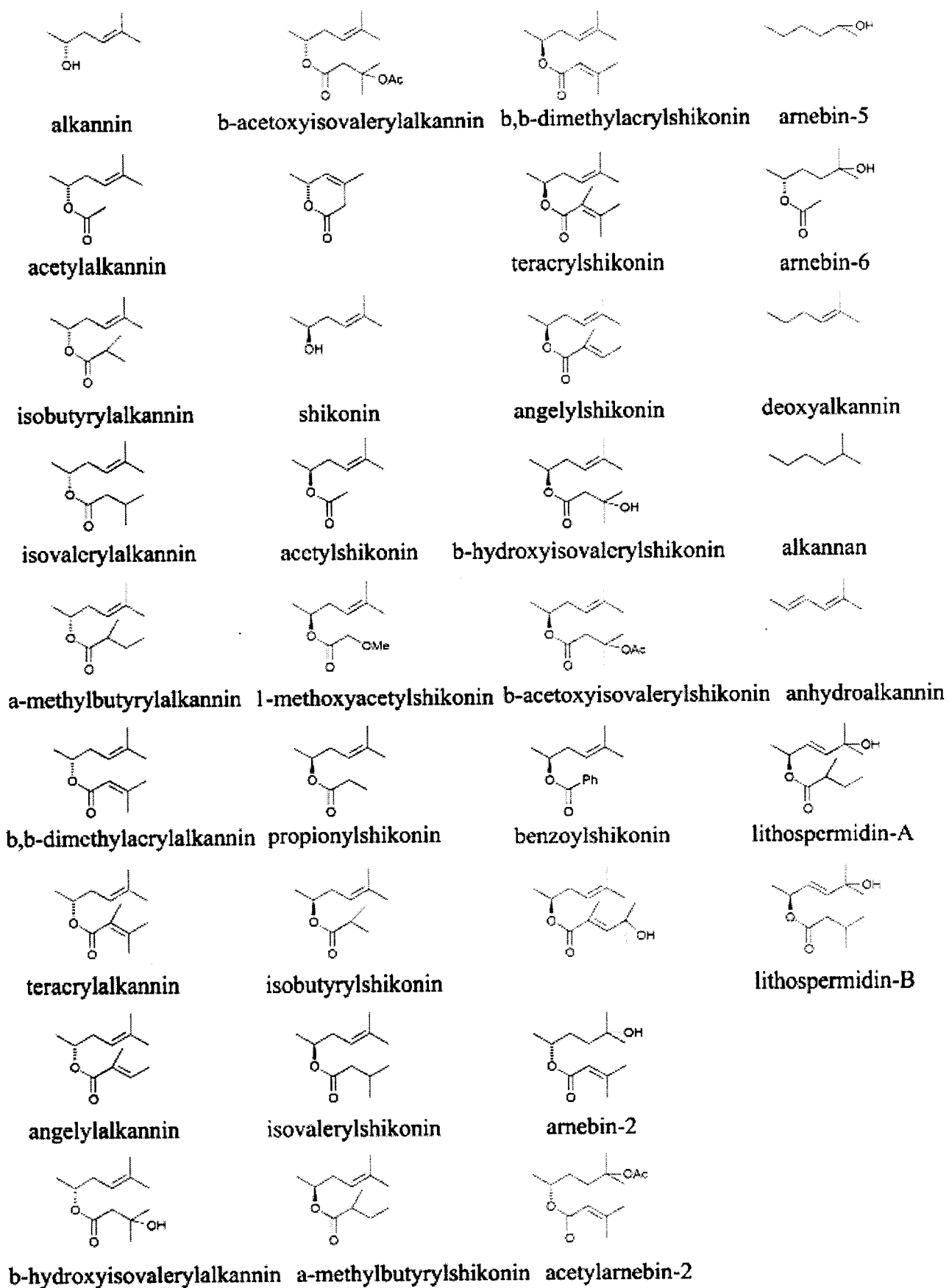
采用本发明方法实施方案的步骤 7 提出的另外一种分析结合小分子的方法对结合的小分子进行分析和鉴定。

紫草（Sinkiang）属紫草科药用多年生草本植物，味甘、性寒，归心肺经。因其凉血活血、解毒透疹的显著功效，在各中医药典籍中均有详细记载。在《本草纲目》中以解表凉血、清热解毒、活血化瘀等广谱疗效被列为上品（以根入药）。在现代医学研究中，因其具有抗感染、促进伤口愈合、抗肿瘤、抗微生物和抗血栓等众多功效，被广泛地用于医学研究和疾病治疗（Papageorgiou 等人，1999. "The Chemistry and Biology of Alkannin, Shikonin, and Related Naphthazarin Natural Products." *Angew. Chem. Int. Ed.* 39: 270-300）。

到目前为止，已经鉴定出紫草中的多种主要成分，这些主要成分对紫草在治疗疾病方面的疗效起着关键性的作用。而相关研究（Papageorgiou 等人，1999. "The Chemistry and Biology of Alkannin, Shikonin, and Related Naphthazarin Natural Products." *Angew. Chem. Int. Ed.* 39: 270-300）表明这些成分主要都是紫草素（skikonin/alkannin）母核的衍生物，该母核的结构式为：



紫草中主要成分的母核的结构式。其中 R 基团不同代表不同的化合物。目前已知的 R 基团包括如下面所示的 36 种。我们首先根据未知电子密度的显示,去除一些基团过大而不可能放入该处电子密度的化合物。然后准备候选的小分子的 PDB 数据,读入 COOT 后,利用 COOT 的 Realspace Refinement 的功能,将不同的小分子在电子密度中进行修正。结果显示,乙酰紫草素与未知电子密度的吻合较好,结果如图 18 所示。由此获得了乙酰紫草素与 HCV NS5B 蛋白复合体的精细三维结构。



序列表

<110> 清华大学、南开大学、中国科学院生物物理研究所

<120> 从中药等天然产物混合体系中高效、快速获取靶分子与靶蛋白复合物精细三维结构的新方法

<160> 2

<210> 1

<211> 306

<212> prt

<213> SARS 冠状病毒 (SARS Coronavirus)

<400> 1

```

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

```

Cys Gly Ser Val Gly Phe Asn Ile Asp Tyr Asp Cys Val Ser Phe Cys
 145 150 155
 160
 Tyr Met His His Met Glu Leu Pro Thr Gly Val His Ala Gly Thr Asp
 165 170 175
 Leu Glu Gly Lys Phe Tyr Gly Pro Phe Val Asp Arg Gln Thr Ala Gln
 180 185 190
 Ala Ala Gly Thr Asp Thr Thr Ile Thr Leu Asn Val Leu Ala Trp Leu
 195 200 205
 Tyr Ala Ala Val Ile Asn Gly Asp Arg Trp Phe Leu Asn Arg Phe Thr
 210 215 220
 Thr Thr Leu Asn Asp Phe Asn Leu Val Ala Met Lys Tyr Asn Tyr Glu
 225 230 235 240
 Pro Leu Thr Gln Asp His Val Asp Ile Leu Gly Pro Leu Ser Ala Gln
 245 250 255
 Thr Gly Ile Ala Val Leu Asp Met Cys Ala Ala Leu Lys Glu Leu Leu
 260 265 270
 Gln Asn Gly Met Asn Gly Arg Thr Ile Leu Gly Ser Thr Ile Leu Glu
 275 280 285
 Asp Glu Phe Thr Pro Phe Asp Val Val Arg Gln Cys Ser Gly Val Thr
 290 295 300
 Phe Gln
 305

<210> 2

<211> 566

<212> PRT

<213> 丙型肝炎病毒基因型 1b 分离株 BK (Hepatitis C Virus HCV, genotype 1b, isolate BK)

<400> 2

Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro Cys Ala Ala
 1 5 10 15
 Glu Glu Thr Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg
 20 25 30

305 310 315 320
 Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Glu Ala Ser Leu
 325 330 335
 Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp
 340 345 350
 Pro Pro Lys Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser
 355 360 365
 Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr Leu
 370 375 380
 Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala
 385 390 395 400
 Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Tyr Ala
 405 410 415
 Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe Phe Ser Ile
 420 425 430
 Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys Gln Ile Tyr
 435 440 445
 Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln Ile Ile Gln
 450 455 460
 Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr Ser Pro Gly
 465 470 475 480
 Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly Val Pro Pro
 485 490 495
 Leu Arg Val Trp Arg His Arg Ala Arg Ser Val Arg Ala Arg Leu Leu
 500 505 510
 Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu Phe Asn Trp
 515 520 525
 Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala Ala Ser Gln
 530 535 540
 Leu Asp Leu Ser Ser Trp Phe Val Ala Gly Tyr Ser Gly Gly Asp Ile
 545 550 555 560
 Tyr His Ser Leu Ser Arg
 565

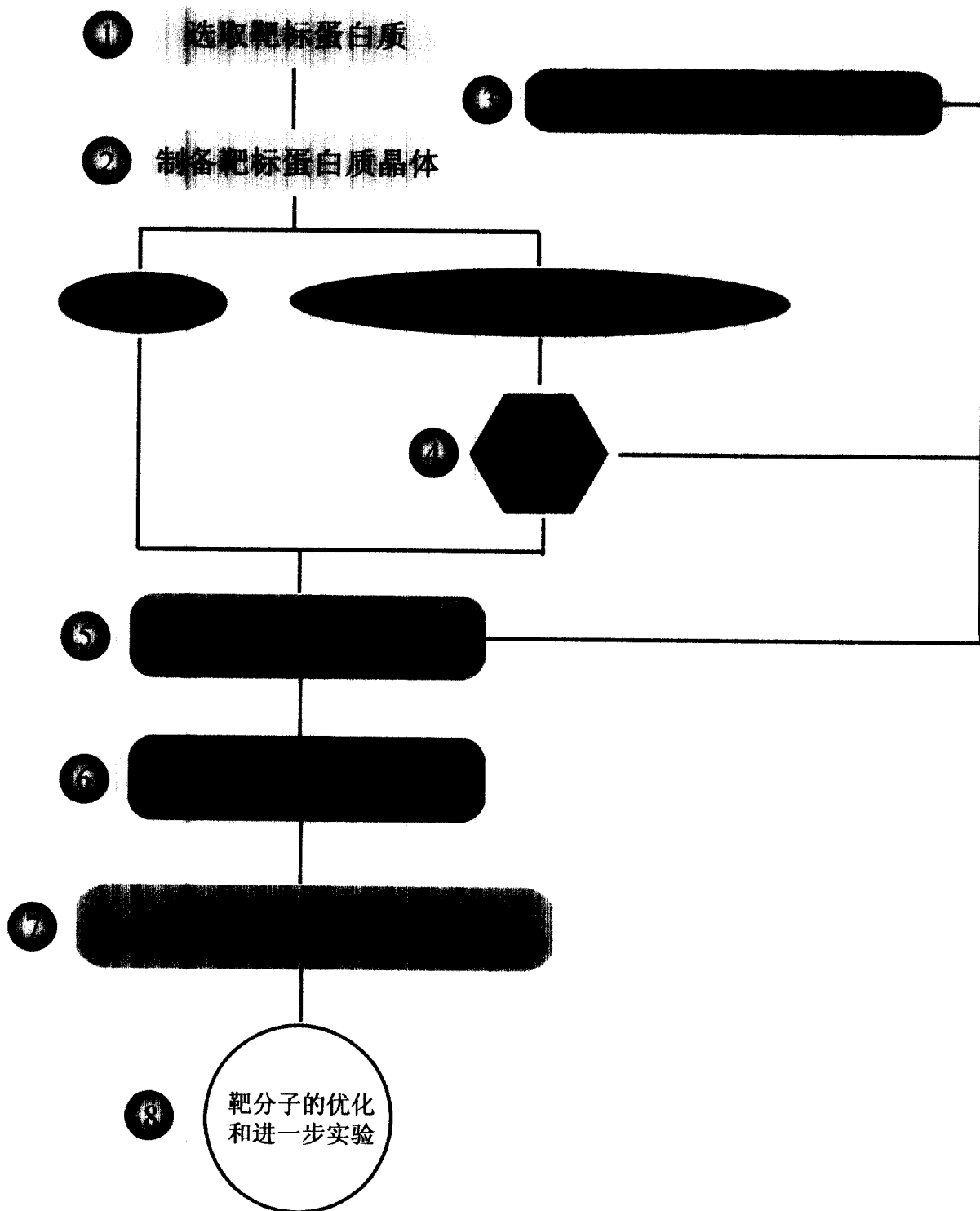


图 1

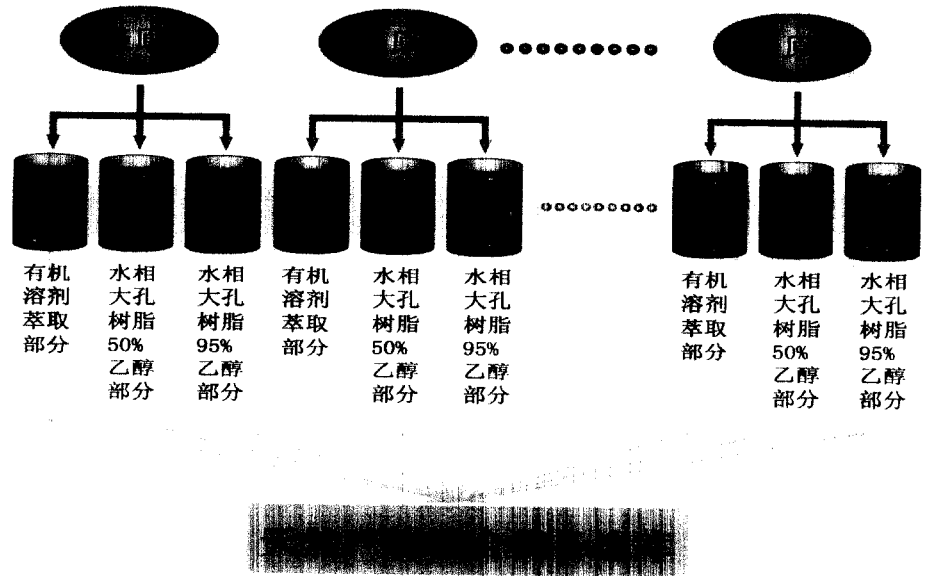
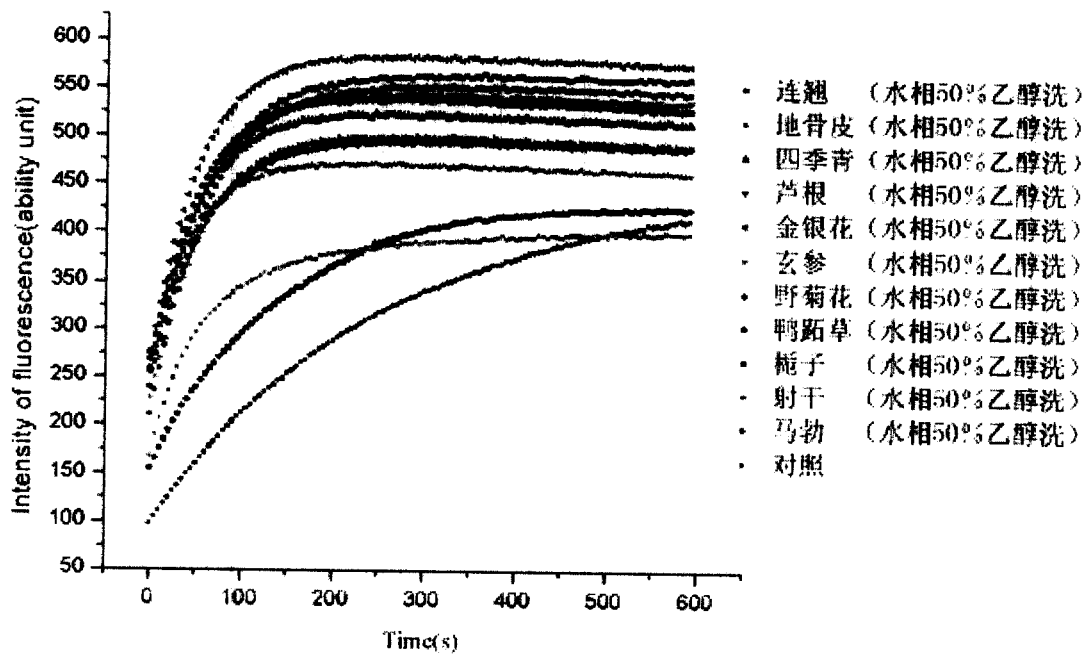


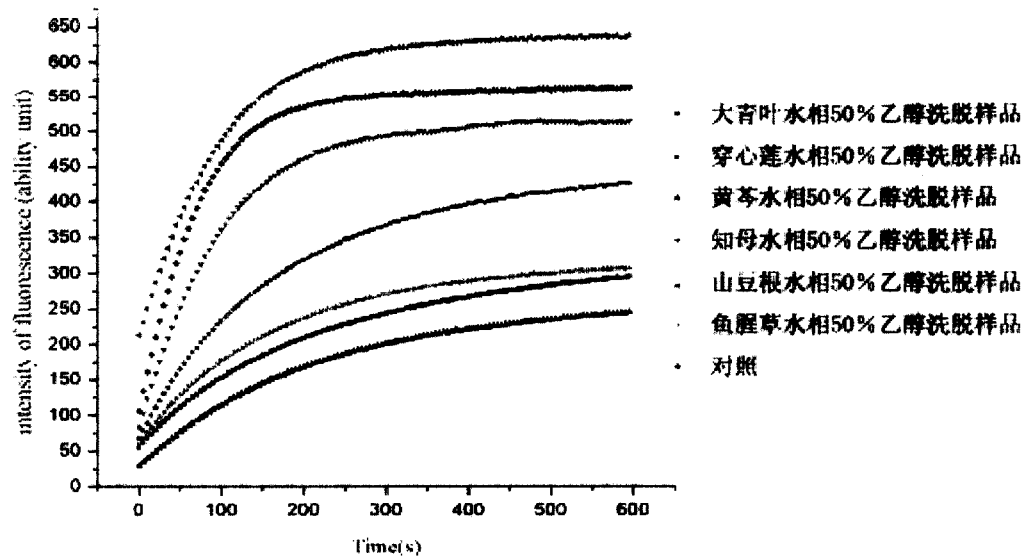
图 2



天然产物水相50%乙醇洗脱样品对SARS主蛋白酶活性的抑制曲线

天然产物样品浓度: 100ug/ml

图 3



天然产物水相50%乙醇洗脱样品对SARS主蛋白酶活性的抑制曲线
天然产物样品的浓度为100ug/ml

图 4

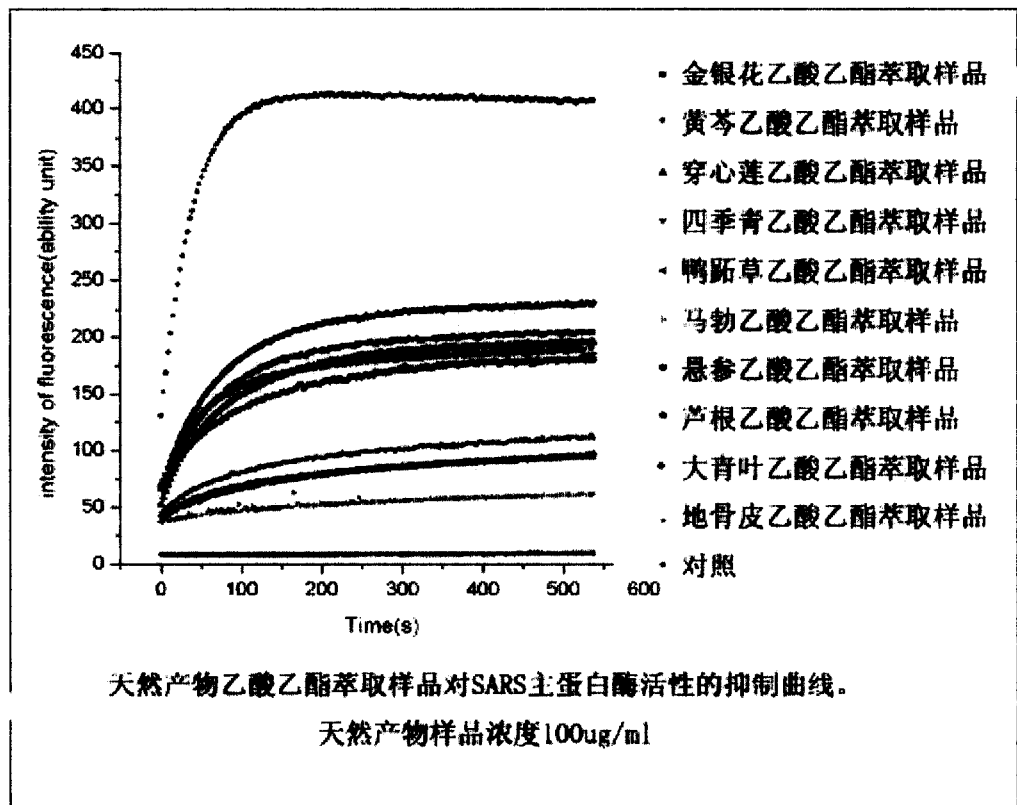
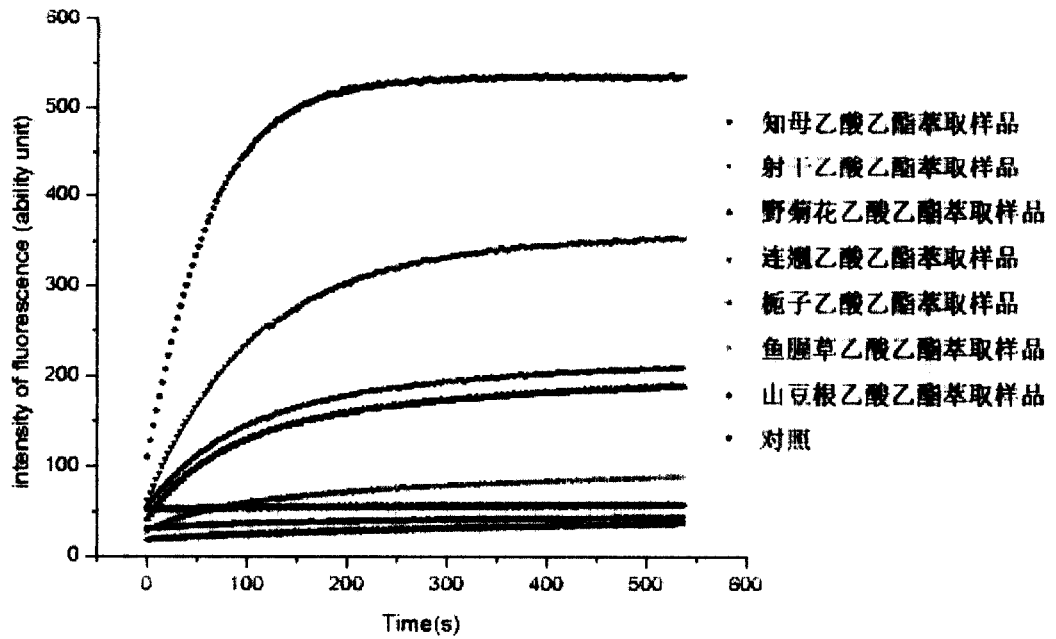


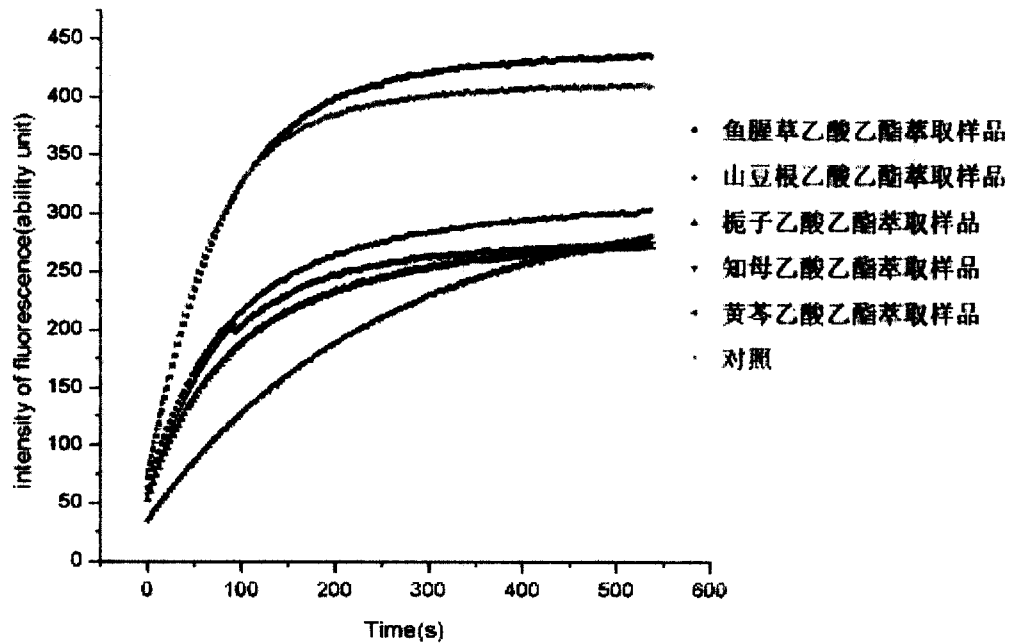
图 5



天然产物乙酸乙酯萃取样品对SARS主蛋白酶活性的抑制曲线

天然产物样品浓度: 100ug/ml

图 6



天然产物乙酸乙酯萃取样品对SARS主蛋白酶活性的抑制曲线

天然产物样品浓度: 10ug/ml

图 7

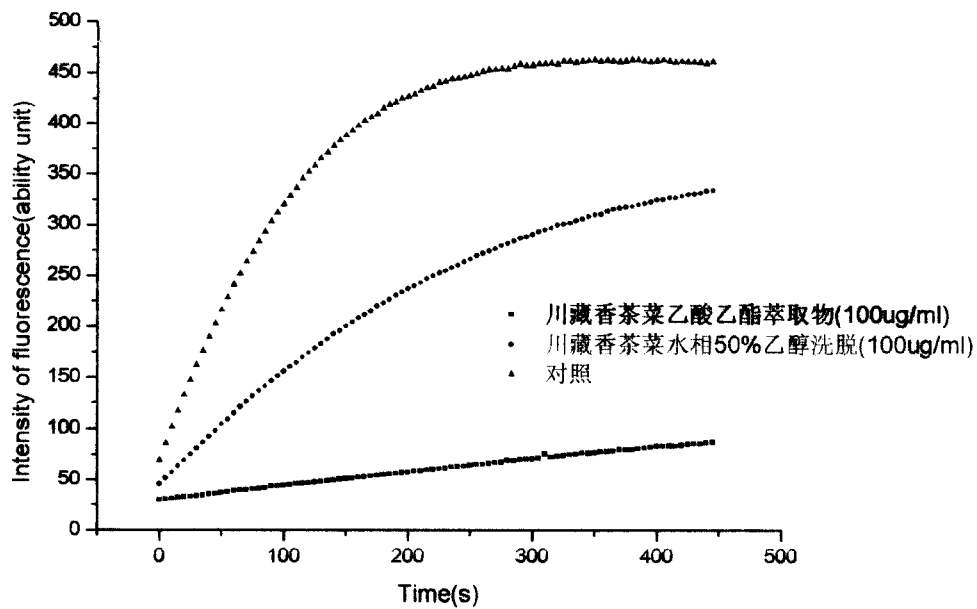


图 8

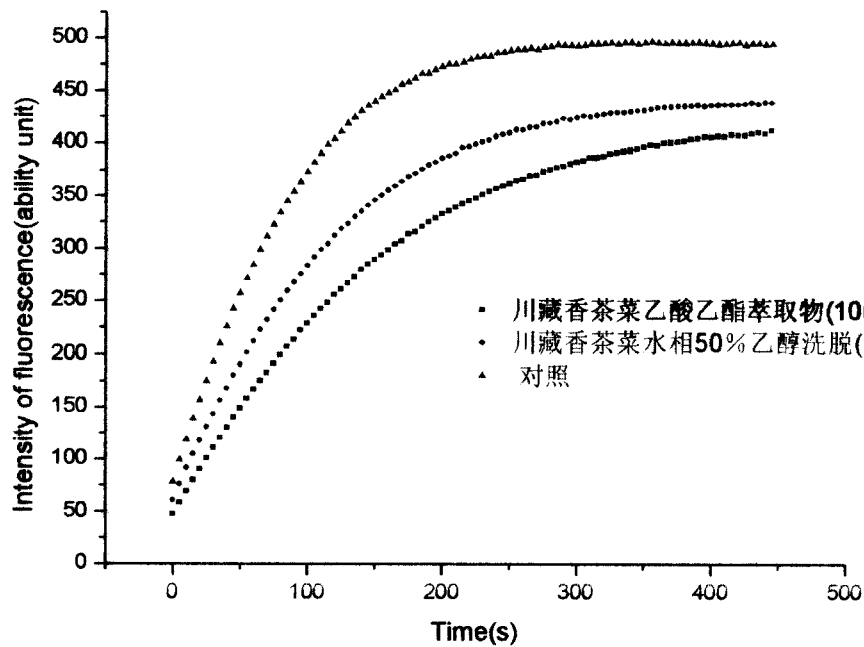
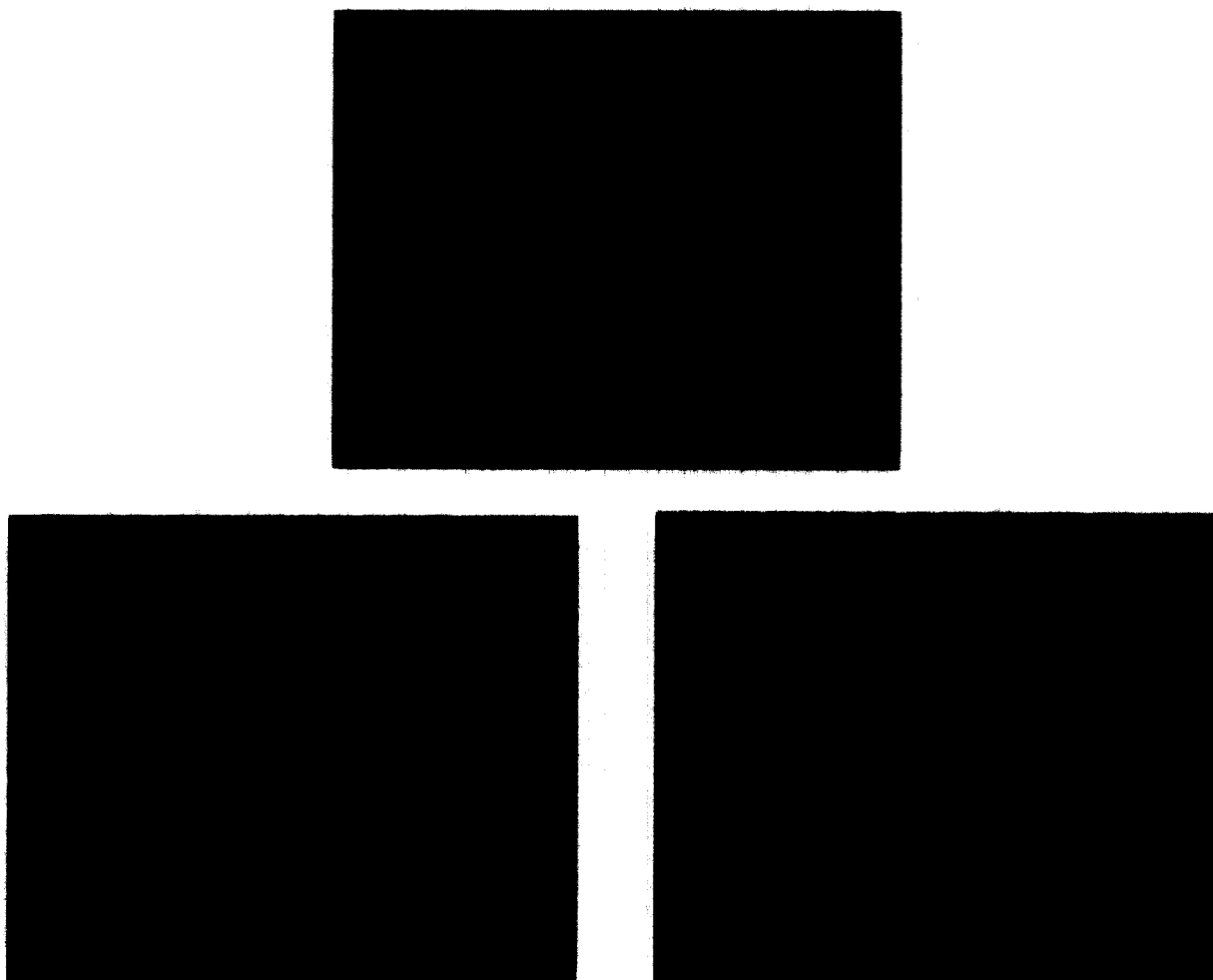


图 9



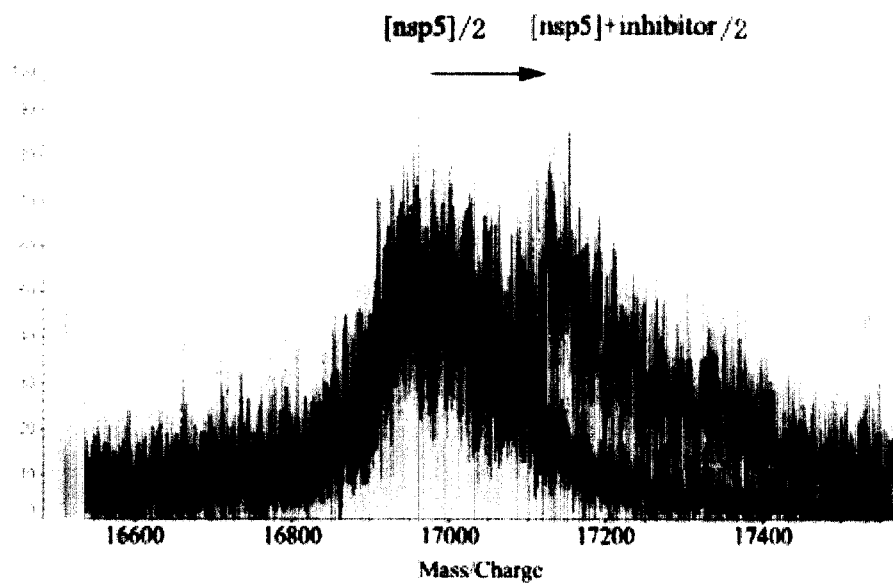
SARS主蛋白酶活性位点电子密度图对比。a为SARS主蛋白酶母体晶体活性位点的电子密度显示；b为SARS主蛋白酶与川藏香茶菜乙酸乙酯萃取物浸泡后晶体活性位点的电子密度显示；c为SARS主蛋白酶与川藏香茶菜丙素浸泡后晶体活性位点的电子密度显示。图中所示的电子密度，蓝色为 2σ 电子密度，为 1.0σ ；青色和绿色为 1σ 差值电子密度，为 2.5σ 。

图 10



未知小分子的电子密度与 SARS 主蛋白酶活性位点 Cys145 残基电子密度形成共价键。蛋白质分子用白色飘带图表示，活性位点的 Cys145 残基用球棍模型表示；蓝色电子密度表示 Cys145 的电子密度，红色电子密度表示未知小分子的电子密度。

图 11



蓝色表示 SARS 主蛋白酶本身的质谱峰；红色表示 SARS 主蛋白酶与川藏香茶菜乙酸乙酯萃取物作用后的质谱峰

图 12

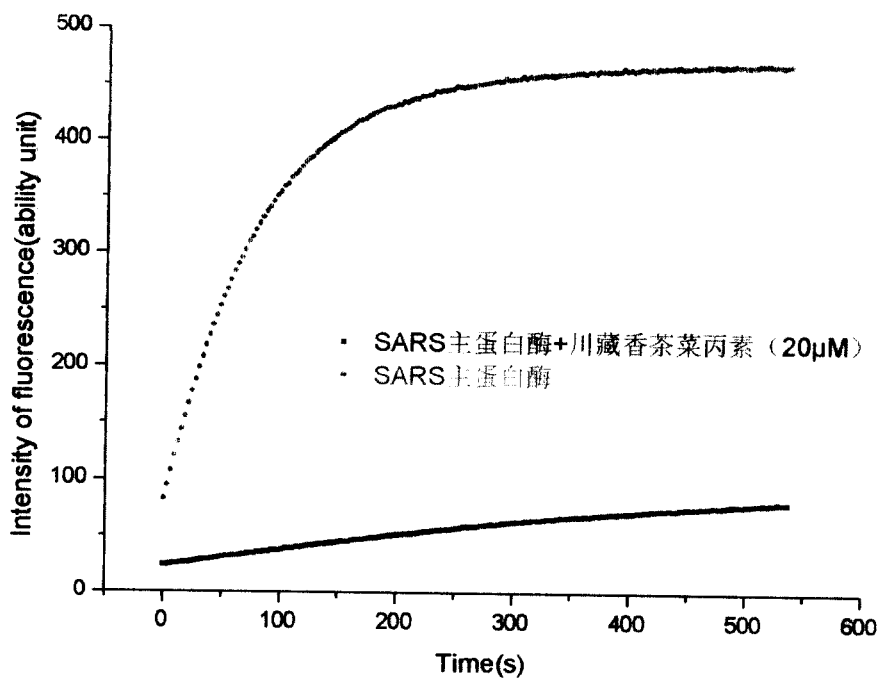


图 13

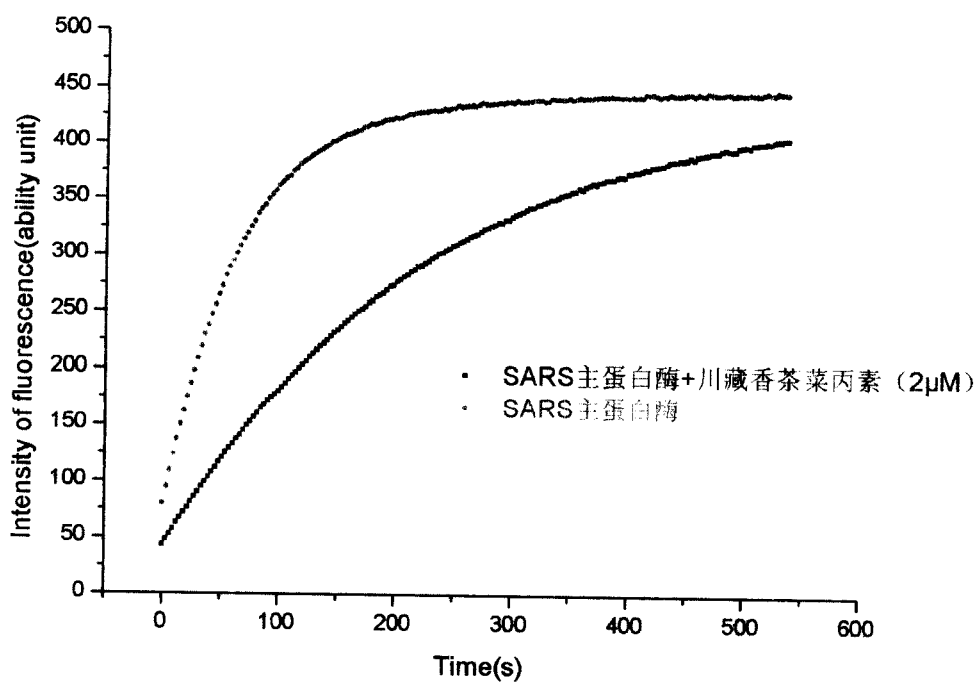


图 14

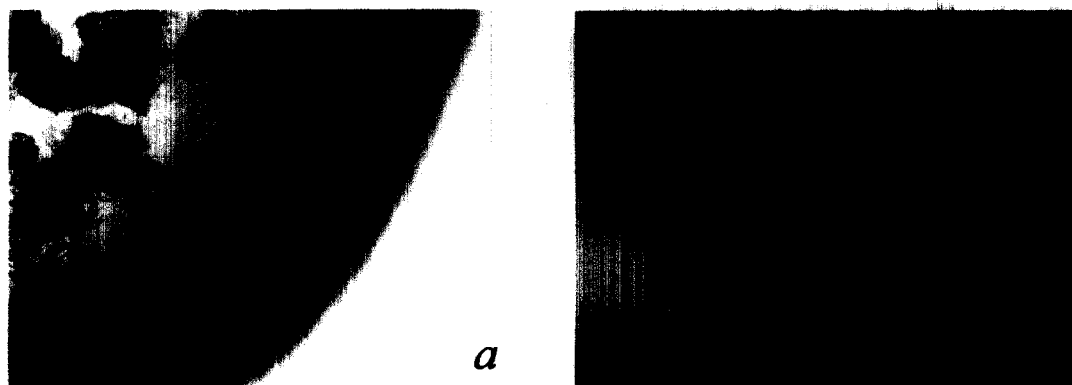


图 15

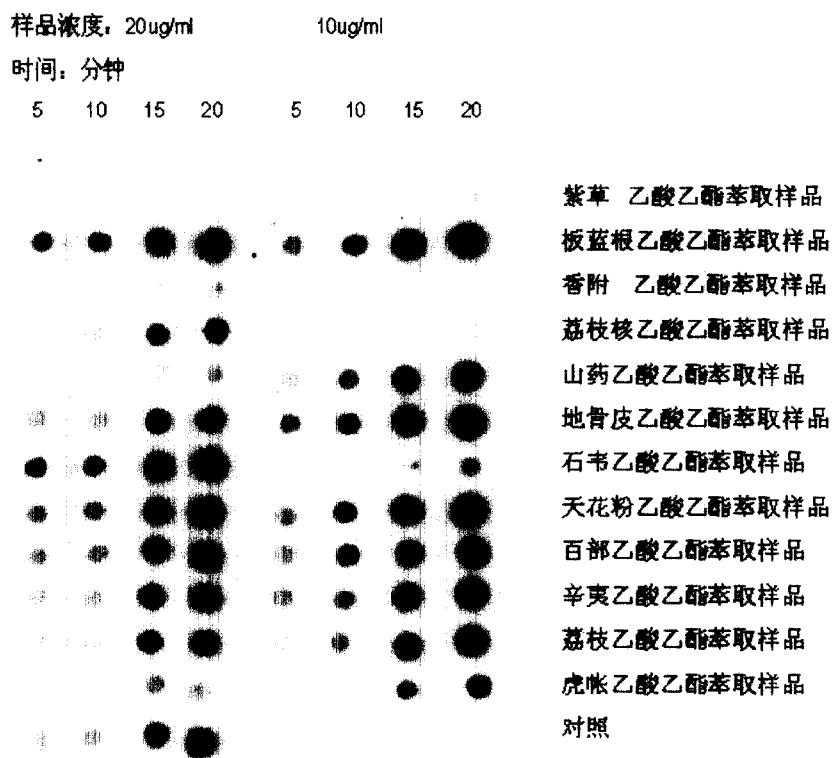
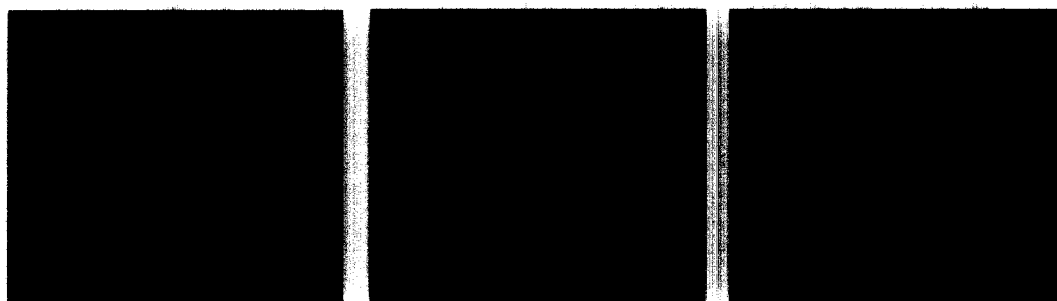


图 16



HCV NS5B 蛋白 His502 活性位点附近电子密度对比。a 为母体活性位点的电子密度；b 为浸泡过紫草乙酸乙酯萃取物之后该处的电子密度，c 为 b 旋转 90°后的显示。图中 $2fofc$ 电子密度为 1.0σ ， $1fofc$ 差值电子密度为 2.3σ 。

图 17



图 18