



# [12] 发明专利说明书

[21] ZL 专利号 94105555.8

[45] 授权公告日 2004 年 6 月 23 日

[11] 授权公告号 CN 1154640C

[22] 申请日 1994.4.13 [21] 申请号 94105555.8

[30] 优先权

[32] 1993.10.1 [33] US [31] 130271

[32] 1993.11.30 [33] US [31] 159861

[71] 专利权人 纽约市哥伦比亚大学理事

地址 美国纽约州

共同专利权人 冷泉港实验室

[72] 发明人 W·C·施蒂尔 M·H·韦格勒

M·H·J·奥尔迈耶

L·W·迪拉德 J·C·里德

审查员 刘桂明

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 王景朝

权利要求书 2 页 说明书 103 页 附图 4 页

[54] 发明名称 用标示物编码的多元组合化学库

[57] 摘要

本文提供了编码组合化学, 其中用有机分子记录连续合成方法, 这些有机分子作为相同或不同的信息量子确定了反应物和阶段的选择。各种产物可在多步合成方法中产生, 例如低聚物和合成的非重复性有机分子。便利地是, 有多族化合物可用作识别剂, 其中取代基的数量和/或位置可确定该选择。另外, 可使用可测定的官能团, 例如放射性同位素、荧光、卤素等, 其中两个不同基团的存在和比例可用于确定阶段或选择。特别是, 可用多种识别剂以提供二进制或高级代码, 以便仅用少数可分离标示物就可确定多种选择。可筛选颗粒的所需特性, 特别是结合亲合力, 其中产物可与颗粒分离或留在颗粒上。可通过释放标示物和分析确定颗粒的反应历程来测量定具有所述特性的颗粒的反应历程。

1. 一种对在多种独特固体载体的每种固体载体上的反应系列的反应历程进行记录的方法，其中所述反应系列包括至少两个阶段，  
5 这些阶段需要不同试剂或反应条件，对多种所述独特固体载体产生不同的改变，在不同的独特固体载体上产生多种不同的最终产物，使用组合识别剂记录所述反应历程，所述识别剂的特征在于确定试剂或反应条件的选择和反应系列中的阶段，并且对该选择和阶段来说识别剂可被分析，所述的方法包括：

10 在反应系列的第一或中间阶段，不同试剂或在使用不同反应条件下与一组独特固体载体中的每个固体载体和组合识别剂反应，其中该独特固体载体组包括至少一种所述的独特固体载体，对于每组独特固体载体，组合识别剂确定试剂的选择和反应系列中的阶段，每个识别剂各自直接连接到独特固体载体上或通过不同于在前识别  
15 剂的识别剂连接到独特固体载体上；

将所述的独特固体载体组混合在一起，然后将这些混合在一起的独特固体载体分成许多组以用于第二中间或最后阶段；

重复该反应至少一次以提供多种最终产物，在不同的各个独特固体载体上具有不同的产物。

20 2. 根据权利要求 1 的方法，包括对独特固体载体上的最终产物筛选所需特性；和识别具有所需特性的至少一种最终产物的反应历程的附加阶段。

3. 根据权利要求 1 的方法，还包括从固体载体解离最终产物和筛选最终产物。

25 4. 根据权利要求 1 的方法，还包括处理识别剂以便从固体载体上分离标示物组分和所述标示物组分与可用荧光或电子俘获检测的部分反应。

5. 根据权利要求 4 的方法，其中分离是通过光化学或氧化进行，

可分离的部分是由丹磺酰氯或多卤苯甲酰卤得到。

6. 根据权利要求 1 的方法，其中识别剂包括具有两种特征的标示物组分，一种是可分离特征，另一种是可检测特征。

5 7. 根据权利要求 6 的方法，其中可检测的特征是可用质谱分析检测的能力。

8. 根据权利要求 7 的方法，其中质谱分析检测氘原子的存在。

### 用标示物编码的多元组合化学库

本申请是1993年11月30日申请的序列号为08/159,861的美国专利申请的部分连续申请,美国专利申请08/159,861是1993年10月1日申请的序列号为08/130,271的美国专利申请的部分连续申请。因此,这两个专利申请的内容作为参考引入到本申请中。

人们对大量性质不同的化合物的合成方法感兴趣,这些化合物可根据各种可能的生理活性或其他活性进行筛选。已经研究出一些方法,在这些方法中,人们顺序地加入各个单元作为化学合成的一部分,以制备所有的或大部分可能的化合物,这些化合物可在合成的每个连续步骤中可能存在的不同选择得到。对于这些将要成功的技术,必须要有适合测定这些化合物的方法,通过这些方法人们可以测定如此制备的化合物的结构。例如, Brenner和Lerner Proc. Natl. Acad. Sci. USA (1992) 81, 5381-5383和WO 93/20242描述了一种合成方法,其中同时产生了低聚核苷酸,并且该低聚核苷酸作为遗传标示物化学连接到低聚肽上成为所需的化合物。WO 93/06121介绍了一些无规低聚物的粒基合成方法,其中颗粒上的识别标示物用来帮助识别合成的低聚物序列。可分开的标记体系描述于Ohlmeyer等人, Proc. Natl. Acad. Sci. USA (1993) 90, 10922-10926中。

这些方法和组合物提供了对组合化学的编码,从而在合成的每

个阶段，唯一地标记一种载体，例如在其上合成化合物的一种颗粒，以确定与在该载体上合成化合物有关的特殊事件，通常是指化学行为。标记是用识别剂分子进行，这些识别剂分子标出了连续事件，在合成过程中载体颗粒暴露于这些连续事件中，这样就提供了在该载体上生成化合物的反应历程。

每种识别剂分子具有如下特征：在所用的合成条件下是稳定的、在合成阶段中保持与载体相联系、在合成过程中唯一地确定了一种特殊事件，该特殊事件反映了在给定的合成阶段下特殊的反应选择、可与在测定过程中存在的其他组分区别开、可分开用方便分析技术可辨别出的标示物组分。

本发明的识别剂可与另一种识别剂结合使用以形成二进制或更高级的编码体系，这样使用较小数目的识别剂就可编码相当大数目的反应产物。例如，当使用二进制编码时， $N$ 个识别剂，就可独特地编码出高达 $2^N$ 个不同的化合物。

此外，本发明的识别剂不需通过在先的识别剂连续地连接到基质上，而是直接地或通过合成的产物单独地连接到基质上。这些识别剂可以是不连续的。而且，这些识别剂含有可分开的成分或部分，这些成分或部分可使容易分析出的标示组分分离。

便利的是，该组合合成法使用可确定的固体载体，在该固体载体上进行反应，并且识别剂结合在该固体载体上。对于感兴趣的特征可以筛选出带有最终产物化合物的单独的固体载体或基质，并且通过分析相关的识别剂标示物来测定反应历程。

图1说明了通过质谱分析所得到的标示物4的分析结果。可看到有两个相应于标示物4的信号。

图2说明了用质谱分析所得到的标示物11的分析结果。可看到有两个相应于标示物11的信号。

图3说明了用质谱分析所得到的标示物13的分析结果。可看到有两个相应于标示物13的信号。

图4说明了当大约相同量的每种标示物混合在一起时,用正化学离子质谱(PCIMS)分析所得到的标示物4、11和13的分析结果。可容易地区别出相应于每种分开标示物的两个信号。

本申请中所用的术语“标示物”或“T”是指具有两种性质的化学部分。第一,它可区别于所有的其他化学部分;第二,当它以 $10^{-18}$ 至 $10^{-9}$ 摩尔量存在时可检测出。这两种性质在单一的化学结构中可体现出来。另外,这些性质在连接在一起的分开的化学结构中可体现出来。在后一种情况下,这些化学结构中的一个化学结构(可表示为C(或在多于一种这种结构C, C'等情况下))提供了使该标示物可区别于其他标示物的性质,而另一个化学结构E提供了使该标示物可检测的性质,并且可选择性地提供使该标示物可与其他标示物分离的性质。

本申请中所用的术语“连接体”或“L”是指具有三种性质的化学部分。第一,它可连接到固体载体上,第二,它可连接到标示物上;第三,当它连接到固体载体和标示物上时,它可分裂开以便标示物可与固体载体分离。这三个性质可在单一的化学结构中体现出来。另外,这些性质可在三个连接在一起的化学结构中体现出来。在后一种情况下,这些化学结构中的一个化学结构(可表示为F')提供了使连接体可连接到固体载体上的性质;第二个化学结构(可表示为V)提供了使连接体可分裂的性质;第三个化学结构(可表示为

A') 提供了使连接体可连接到标示物上的性质。所希望的是, 化学结构 V 和 A' 是一个和相同的化学结构, 其中 V-A' 化学结构可表示为 F<sup>2</sup>。

本申请中所用的术语“识别剂”是指包括标示物和连接体的化学个体。因此, 在广义上, 识别剂可由式 L-T 表示, 而识别剂的具体含义由式 F'-V-A'-T; F'-V-A'-C-E (或 F'-V-A'-E-C); L-C-E (或 L-E-C) 和 L-C-E-C' 表示。

本文中所用的术语“选择”是指组合合成中给定阶段的交替变量, 例如, 反应物、试剂、反应条件及其组合。术语“阶段”是相应于化合物或配位体的连续合成中的步骤; 该化合物或配位体是组合合成的最终产物。

术语“烷基”包括直链、支链和环状结构及其组合。因此, 该术语包括甲基、乙基、丙基、异丙基、丁基、仲和叔丁基、环丙基、环丁基、环戊基、2-甲基环丙基等。低级烷基是 C<sub>1</sub>-C<sub>6</sub> 烷基。低级链烯基是直链、支链或环状构型及其组合的 C<sub>2</sub>-C<sub>6</sub> 链烯基。

除非另有说明, 在特定分子中任何取代基 (例如 R<sup>1</sup>, R<sup>2</sup>, Z 等) 的定义与其在该分子中其他地方的定义是无关的。因此, NR<sup>4</sup>R<sup>4</sup> 代表 NHH、NHCH<sub>3</sub>、NHCH<sub>2</sub>CH<sub>3</sub>、N(CH<sub>3</sub>)<sub>2</sub> 等。

本文所述化合物中的某些化合物含有一个或多个不对称中心, 因此, 可产生对映体、非对映异构体和其他立体异构形式。本发明包括所有这些可能的立体异构体以及它们的外消旋和旋光纯形式。旋光活性 (R) 和 (S) 异构体可用手性合成纤维、手性试剂来制备, 或用常规技术拆分。当本文所述的化合物含有烯烃双键时, 则这些化合物包括 E 和 Z 几何异构体。

在其上进行本发明组合合成的物质在本文中可互换地称作小球、固体表面、(固体)基质、颗粒、载体等。这些术语将包括:

a) 固体载体, 例如小球、丸、园环、毛细管、空心纤维、针状物、固体纤维、纤维素小球、多孔玻璃小球、硅胶、与二乙烯基苯任意交联的聚苯乙烯小球、接枝共聚小球、聚丙烯酰胺小球、胶乳小球、与N, N'-二-丙烯酰乙二胺任意交联的二甲基丙烯酰胺小球、涂有疏水聚合物的玻璃颗粒等, 即, 具有刚性或半刚性表面的物质;

b) 可溶载体, 例如低分子量非交联聚苯乙烯。

这些物质必须含有官能团或必须能被官能化以便识别剂或产物中间物能够连接到这些物质上。

另外, 下列缩写词具有指定的含义:

AcOH= 乙酸

BSA = 二(三甲基甲硅烷基)乙酰胺

CAN = 硝酸铈(iv)铵

DEAD= 偶氮二甲酸二乙酯

DCM = 二氯甲烷

DIC = 二异丙基碳化二亚胺

DMF = N, N-二甲基甲酰胺

Fmoc= 9-芴基甲氧基羰基

HOBt= 1-羟基苯并三唑

PhMe= 甲苯

r. t. = 室温

TFA = 三氟乙酸



## THF = 四氢呋喃

本发明涉及产物，即化合物库的制备，其中存在于这些库中的各个产物或化合物可以相互物理分开，并且可以对或者在固体载体上或者与固体载体分开的感兴趣的特性进行筛选。通过连续合成，其中在合成的每个阶段每个单独的中间体用各种方法处理，产生了很多个产物，每种产物的少量存在，常少于100微微摩尔，更常少于10毫微微摩尔。因为如此产生的最终产物或化合物的量很少，所以通过分离和结构上阐明产物来识别这些产物通常是不可行的。而且，在涉及到加入非类似单元的连续合成中，如果使用一般可得到的产物的量进行分析不是不可能的话，那么这样的分析也将是很困难的。然而，通过将连续合成的每个选择或组合选择（例如，“加试剂A”或“加试剂A，然后加试剂B，并加热至100℃2小时”）与可确定变量（例如反应物、试剂、反应条件、或这些变量的组合）的组合识别剂相联系，人们就可使用识别剂来确定每个可确定的和可分离的基质的反应历程。通过对与识别剂分开的标示物的分析可在微微摩尔或更低浓度，例如毫微微或更低浓度下容易的识别反应历程。根据与产物相联系的标示物，人们可用各种筛选技术来测定合成产物的特性，一般是化学特性或生物特性，然后识别反应历程，从而确定产物的结构，该产物具有所需要的性质。

重要的是，本发明方法使用可与合成的配位体或化合物分离的标示物，这也是为了消除编码。这种可分离性也使标示物在多于一个依据的情况下可辨别出。特别是，这些标示物可分离（例如，根据色谱保留时间），然后可以分析（例如，第二个依据是光谱性质如质谱分析 $m/e$ ，或电泳性 (electrophoricity)）。由于具有多个区别

依据,所以可用少数标示物来编码大量信息。

因此,本发明提供了用于编码组合合成的方法和组合物,从而在合成的每个阶段提供一种或多种识别剂,这些识别剂可对与在载体或颗粒上进行化合物合成中的颗粒阶段相关的事件编码。该事件包括在反应的那个阶段选择反应物和 / 或反应条件,其中每个这样的阶段可涉及一种或多种反应物,这些反应物在相同的或不同的条件,例如部分反应、多步加入、加入速度、试剂的不同组合等条件下是相同的或不同的。另外,颗粒的基团可与其他颗粒的基团整合,并在连续合成过程中的任何时间经受一系列不同事件。

通过提供N个识别剂,并且每个识别剂具有M个可辨别状态,就可独特地确定 $M^N$ 个不同的合成。在 $M=2$ 的情况下,这时这两个状态可以是有识别剂存在或没有识别剂存在,这样该合成将用基数2或二进制代码确定。在 $M=3$ 的情况下,这时这三个状态可以是在两个可辨别浓度下存在识别剂或不存在识别剂,则该合成将用基数3代码确定。在本文中,这种 $M>2$ 的基数M代码称作高级代码。与二进制代码相比高级代码的优点是可用较少的识别剂来编码相同量的关于合成的信息。所产生的产物将被定义为是由一系列合成所产生的。在合成的每个阶段,存在许多反应物和 / 或试剂和 / 或条件,它们产生了与可识别的和通常可分离的本体如标示物有关的产物的特性。关于反应物和试剂,意图是指反应物能多数引入到产物中,例如氨基酸、核苷酸、亲核试剂、亲电子试剂、三烯、烷基化剂或酰化剂、二胺或任何其它合成纤维等,而试剂则可以或不可以引入到产物中,例如碱、酸、热量、氧化剂或还原剂,但反应物和试剂均包括在术语“试剂”中。该合成方法中可涉及到可引入到产

物中的单个反应物。另外，阶段可包括一个或多个使反应中间体改变的反应。在许多情况下，将包括这些可能性的组合。

使用基数2或二进制代码 ( $M=2$ ) 和三个识别剂 ( $N=3$ )，则在合成中的一个给定阶段中编码  $8 (2^3)$  个试剂。如果三个识别剂表示为 T1、T2、和 T3，每个识别剂在存在时或不存在时分别表示为“0”或“1”，那么8个不同试剂可用二进制代码表示如下：

	试剂 1	试剂 2	试剂 3	试剂 4
T1, T2, T3	0, 0, 0	1, 0, 0	0, 1, 0	1, 1, 0
	试剂 5	试剂 6	试剂 7	试剂 8
T1, T2, T3	0, 0, 1	1, 0, 1	0, 1, 1	1, 1, 1

同样，用更多的识别剂可以编码更多的关于合成的信息。例如，9个识别剂 ( $N=9$ ) 和基数2代码 ( $M=2$ ) 可编码高达  $2^9$  或512个不同的试剂选择。使用基数3代码 ( $M=3$ ) 和三个识别剂 ( $N=3$ )，可编码27个 ( $3^3$ ) 试剂选择。如果这三个识别剂表示为 T1, T2 和 T3，不存在识别剂时表示为“0”，识别剂以  $\sim 0.5$  微微摩尔 / 小球量存在时表示为“1”，识别剂以  $\sim 1.0$  微微摩尔 / 小球量存在时表示为“2”，那么，这27个不同试剂可用三个识别剂的基数3代码表示如下：

	试剂 1	试剂 2	试剂 3	试剂 4
T1, T2, T3	0, 0, 0	1, 0, 0	2, 0, 0	0, 1, 0
	试剂 5	试剂 6	.....	试剂 27
T1, T2, T3	1, 1, 0	2, 1, 0	.....	2, 2, 2

为了使这种高级编码图切实可行, 应将一种给定量的 (例如-1.0 微微摩尔 / 小球) 附加识别剂加入到该库的所有成员中以便提供一个标准, 与这个标准对照可测定出所有识别剂的量。用各种检测方法经气相色谱或 HPLC 可测定出这些识别剂的量。在 HPLC 情况下, 如果识别剂用不同量的放射性核素如氚 ( $^3\text{H}$ ) 进行放射性示踪, 那么通过闪烁计数可方便地测定出这些识别剂的量。通过测定  $^3\text{H}$  与  $^{14}\text{C}$  之比可特别容易地进行定量, 因此, 使用  $^{14}\text{C}$  作为标准。在这种方法中, 可辨别出 10 个  $^3\text{H}$  的量, 从而得到了基数 10 或十进制代码 ( $M=10$ ), 用为数很少的识别剂该代码可编码大量信息。

### 产物和合成方法

本发明方法的产物多半是有机化合物, 在此该连续合成将包括加入或除去化学单元, 反应包括改变或引入一个或多个官能团, 开环作用, 闭环作用等。化学单元可呈许多形式, 可以是天然形成的

和合成的，例如亲核试剂、亲电子试剂、二烯、烷基化剂或酰化剂、二胺、核苷酸、氨基酸、糖、类脂类、或它们的衍生物、有机单体、合成纤维，和它们的混合物。另外，反应可包括产生烷基化、酰化、硝化、卤化、氧化、还原、水解、取代、消去、加成等作用的反应。该方法可产生极少量的非低聚物、低聚物或它们的混合物，其中反应历程和在适当情况下的成分可由存在的标示物确定。非低聚物包括各种各样的有机分子，例如杂环、芳族、脂环族、脂族分子及其组合分子，包括类固醇、抗菌素、酶抑制剂、配位体、激素、药、生物碱、类鸦片、萜烯、卟啉、毒素、催化剂，以及它们的混合物。低聚物包括低聚肽、低聚核苷酸、低聚糖、聚类脂类、聚酯、聚酰胺、聚尿烷、聚脲、聚醚、聚（磷衍生物）例如磷酸酯、磷酯、磷酰胺、磷酰胺、亚磷酸酯、亚磷酰胺等、聚（硫衍生物）例如砷、磺酸酯、亚硫酸酯、氨磺酰、亚磺酰胺等，其中对磷和硫衍生物来说，指出的杂原子多半将结合到C、H、N、O或S上，和它们的混合物。

反应可包括在中心核分子结构的各种任意位置上的变型或在特定位置上的变型。例如，人们可以溴化一种多环化合物，其中溴化作用可在许多位置上发生，或者使用溴化剂，该溴化剂对特殊位置具有专属性，例如N-溴琥珀酰亚胺。这些反应多半将涉及单一位置或等价位置，例如，乙二醇的两个羟基中的一个羟基。

本发明的合成反应大部分具有至少两个阶段，其中除了双官能化合物外，化合物是用相同的连接官能团连接，例如氨基酸和酰胺键、核苷酸和磷酸酯键、或它们的类似化合物，例如氨基异氰酸酯和脲键。

本发明的方法允许在每个阶段的反应有所变化，这取决于所选择的试剂和所用的条件。合成方法将随人们希望制备的产物的性质而变化，因此，该方法必须考虑分段改变产物性质的能力，同时能够保持前面阶段的结果，并预先考虑到以后阶段的需要。

在研究合成方法中，人们可以进行少数化合物的分批合成，这些化合物将在组合合成过程中制备。通过一些极限实例，人们可优选条件以提高化合物的产率，这些化合物用其他方法不能形成或只以低产率形成。人们可使用分批合成法，这些方法提供了比组合完成法更高浓度的特产物，以开发一些说明化合物活性的检测法。

载体：连接和分离

---

合成方案要求人们提供许多不同的反应，这些反应涉及在合成的每个阶段产生许多不同中间体的不同反应物。尽管其他技术也是适用的，这可通过使用可确定的小固体基质而很容易地实现，例如市售的小球，这些小球可容易混合，分离，并可用作连续合成的固体基质。这些固体基质可以是固体的、多孔的、可变形的或坚硬的，并且具有任何适合的结构和形状。在某些情况下，磁性小球或荧光小球也是适用的。这些小球的直径通常为至少10-2000  $\mu\text{m}$ ，优选至少20-500  $\mu\text{m}$ ，更优选至少50-250  $\mu\text{m}$ 。

任何适合的成分均可用于颗粒或小球中，该小球成分应具有下述特性：在各种加工阶段保持其机械牢固性，可被官能化，具有官能基或可与活性物反应，可以连续合成以及连接识别剂，可容易地混合和分离，并且能够容易地与标示物和产物分离。可使用的小球包

括纤维素小球、微孔玻璃小球、硅胶、聚苯乙烯小球、接枝共聚物小球如聚乙二醇 / 聚苯乙烯、聚丙烯酰胺小球、胶乳小球、二甲基丙烯酰胺小球；特别是与N, N'-二-丙烯酰乙二胺交联的物质，包括N-叔丁氧基羰基- $\beta$ -丙氨酰-N'-丙烯酰六亚甲基二胺、复合物，例如涂有疏水聚合物如交联的聚苯乙烯或接枝线性聚苯乙烯的氟化乙烯聚合物的玻璃颗粒等。

根据合成过程或最终产物的检测方法的特性，各种各样的小球大体上是合乎要求的。尽管小球是特别适合的，其他的固体载体也可使用，例如毛细管、空心丝、针状体、固体纤维等，这时固体载体的大小在反应过程中可根据需要而变化。

根据合成方法的特性，这些小球可用各种方法官能化以便能够连接初始反应物。这些反应物可通过非易变键连接，例如酯键、酰胺键、胺键、醚键，或通过硫、硅或碳原子连接，这取决于人们是否希望从小球上能够除去产物。适合的是，连到小球上的键可以是永久性的，但在小球和产物之间可提供一个连接体，该连接体是可裂开的，如表1所列举的连接体。可使用两个或多个不同的键以便可有区别地释放标示物和 / 或产物。

根据连接到颗粒上的连接基的性质，小球上的活性官能团可以不是必须的，在此连接的方式应便于引入单键或双键，例如可用碳烯和氮烯或其他高活性物来实现。在这种情况下，将可裂开的键提供于连接基，该连接基将产物或标示物连接到小球上。

希望的是，当产物是永久性地连接时，与小球连接的键将是持久的，以便在筛选过程中，小球在空间上不与产物的连接相干扰。可使用各种键，特别是亲水键，例如聚乙烯氧基、糖类、多元醇、

酯、酰胺、它们的组合等。

存在于小球上的官能团可包括羧基、羟基、偕卤代亚胺、氨基、硫代、活性卤素 (Cl 或 Br) 或假卤素 (例如  $-\text{CF}_3$ 、 $-\text{CN}$  等)、羰基、甲硅烷基、甲苯磺酰基、甲磺酰基、对溴苯磺酰基、三氟甲磺酰基等。在选择官能团中, 应考虑到识别剂通常也将连接到小球上。这些考虑将包括相同的或不同的官能团是否应与产物和识别剂相联系, 以及这两个官能团是否与适合的产物或识别剂的连接和标示物的分离阶段相适应。不同的连接基可用于产物, 以便可选择释放出特定量的产物。在某些情况下, 颗粒可具有保护的官能团, 这些官能团可在每个阶段之前部分或完全脱保护, 在完全脱保护的情况下可以再保护。例如, 在多肽合成中氨基可用苄酯基保护, 羟基用苄醚保护等。

在需要产物分离时, 可使用许多官能团和反应物。适合的是, 可使用醚, 其中取代的苄醚或其衍生物, 例如二苯甲基醚、2, 3-二氢化茛基醚等, 可用酸性或温和的还原条件分裂。另外, 人们可使用  $\beta$  消去反应, 其中可用弱碱来释放出产物。可使用乙缩醛, 包括其硫代类似物, 其中弱酸是适合的, 特别是在俘获羰基化合物存在下。通过将甲醛、HCl 和醇部分混合, 形成了  $\alpha$ -氯代醚, 然后在小球上与羟基官能团偶联形成了乙缩醛。可使用对光敏感的键合, 例如 7-硝基 2, 3-二氢化茛基、2-硝基二苯甲基醚或酯等。酯或酰胺可用作连接体, 其中特别是与环酐反应, 接着在使用偶合剂如碳化二亚胺的情况下通过与小球上的羟基或氨基官能团反应, 形成了半酸性酯或酰胺。可使用肽作为连接体, 其中该序列经过了酶催化水解, 特别是其中酶识别了一个特殊的序列。可用碳酸衍生物, 如光气、



羰基二咪唑等和弱碱制备碳酸酯和氨基甲酸酯。特别是对碳酸酯来说,该键合可用酸、碱或强还原剂(如 $\text{LiAlH}_4$ )分裂。可分裂键的一览表参见,例如, Greene和Wuts, <<有机合成中的保护基>>, 第二版, Wiley, 1991。已经开发出的各种体系的多方面的适应性可使产物和识别剂的连接和产物和标示物的不同分离的条件在较宽范围内随意变化。

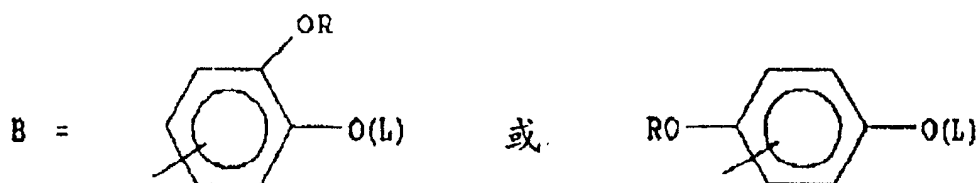
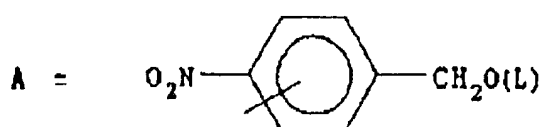
下表给出了各种示例性的连接单元(即式 I 中的 $\text{F}^2$ )和它们可分裂的方式:

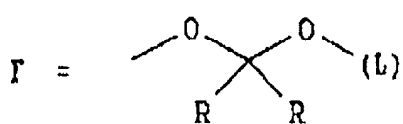
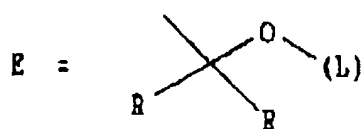
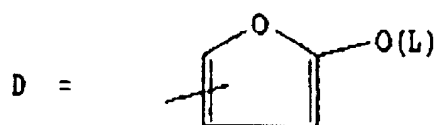
表 1. 各种示例性的连接单元和它们可分裂的方式

连接基	分裂试剂
甲硅烷基	氟化物或酸
A	$h\nu$
B	$\text{Ce}(\text{NH}_4)_2(\text{NO}_3)_6$
$-\text{NCO}_2(\text{L})^*$	$\text{OH}^-$ , $\text{H}^+$ , 或 $\text{LiAlH}_4$
C	$\text{O}_3$ , $\text{OsO}_4/\text{IO}_4^-$ , 或 $\text{KMnO}_4$
D	1) $\text{O}_2$ 或 $\text{Br}_2$ , $\text{MeOH}$ 2) $\text{H}_3\text{O}^+$
$-\text{Si}-(\text{L})$	氧化作用等 $\text{H}^+$ , $\text{Br}_2$ , $\text{Cl}_2$ ,
E	$\text{H}_3\text{O}^+$
F	$\text{H}_3\text{O}^+$
G	$\text{F}^-$ 或 $\text{H}^+$

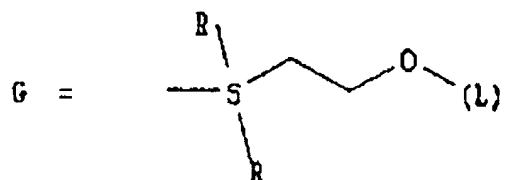
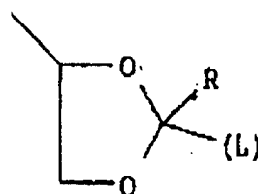
H	X= 酮、酯、酰胺、 NO <sub>2</sub> 、硫化物、 亚砷、砷、和相 关的吸电子基	碱 OH <sup>-</sup>
I		H <sub>3</sub> O <sup>+</sup> 或还原作用 (例如 Li/NH <sub>3</sub> )
J		( $\phi_3P$ ) <sub>3</sub> RhCl(H)
K		Li, Mg, 或 BuLi
M		Hg <sup>+2</sup>
N		Zn 或 Mg
	x = 卤素或假卤素	
O		氧化作用 (例如 Pb(OAc) <sub>4</sub> 或 H <sub>5</sub> IO <sub>6</sub> )
P		碱
	x = 吸电子基	

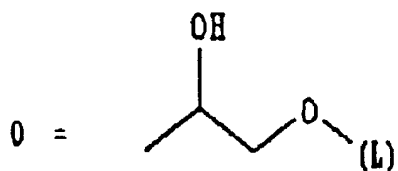
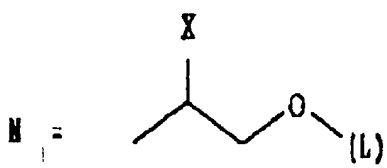
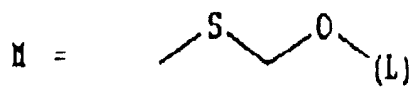
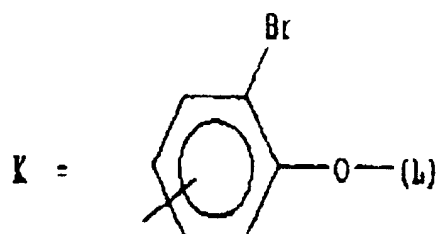
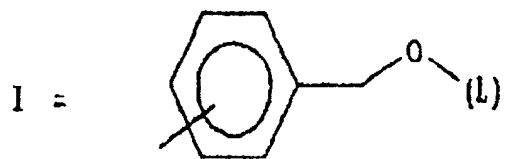
(L) 表示标示物或产物的连接点。

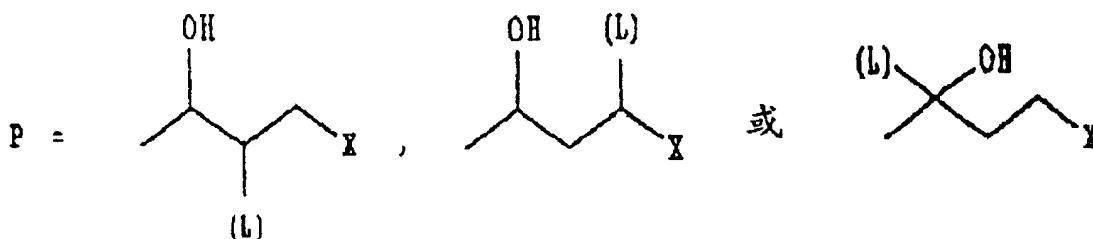




或







L是或者直接连接到指示原子或通过连接基如C(O)O间接连接到指示原子上的标示物或产物，该连接基可提供适合的官能团。

R是H或低级烷基。

### 连接体

对配位体来说连接体的选择将是合成方法的一部分，因为该连接基在产物上可产生剩余官能团。产物与小球分离后对其进一步改性通常是困难的，但也是有可能的。在设计合成方法中，可用保留在产物中的一个官能团作为连接基的连接点。另外，当产物的性质允许时，人们可采用分裂或分离方法除去连接官能团，例如，用金属氢化物或酸来除去芳磺基醚或甲硅烷基。因为在许多情况下，该合成方法可包括一个用于连接的官能化部位，在选择连接基时，该官能团优先使用。在某些情况下，要求在产物与载体的连接部位具有不同的官能团，这需要使用不同的连接方式，这些连接方式必须适应相同的分离方法或不同的分离方法，这些方法可同时进行或连

续进行, 例如用光照射和酸性水解。

用于将识别剂与颗粒连接的特别有意义的连接基是碳烯和氮烯, 该碳烯和氮烯可插入到碳和氢原子之间形成共价键, 或插入到烯烃键中形成环丙烷 (在碳烯情况下) 或氮丙啶 (在氮烯情况下)。

各种取代的苯可与碳烯或氮烯连接基一起使用, 其中苯被一个可提供碳烯的基团:  $\text{CHN}_2$ 、 $\text{COCHN}_2$ 、 $\text{SO}_2\text{CHN}_2$ ; 或可提供氮烯的基团:  $\text{N}_3$ 、 $\text{NO}_2$ 、 $\text{NO}$ 、 $\text{SO}_2\text{N}_3$  所取代。碳烯可由重氮基烷烃衍生物通过光解、热解, 或用低价过渡金属化合物如  $\text{Rh}(\text{OAc})_2$  处理而产生。氮烯可由叠氮化物通过光解或热解而产生, 和由硝基、亚硝基和叠氮化物通过使用三价磷化合物或低价过渡金属而产生。

特别有意义的一类连接体部分 ( $\text{F}^1\text{-F}^2$ -) 包括 2-硝基-4-羧基苄氧基、2-硝基-4-重氮乙酰基苄氧基、4 或 5 叠氮基甲基羧基-2-甲氧基苄氧基、和 2-甲氧基-4, 或 5-羧基苄氧基部分。

其中 T 代表标识物, Z 代表碳烯或氮烯前体或羧基, R 是 H 或低级烷基的示例性化合物如下: 对于光化学标示物分离 (例如用约 350 nm 紫外光): T3-Z-2-硝基苄醚、T4-Z-2-硝基苄醚、T5-Z-2-硝基苄醚、T6-Z-2-硝基苄醚、T2-Z-4-硝基苄醚、T3-Z-4-硝基苄醚、T3-Z-2-硝基苄基碳酸酯、T5-Z-2-硝基苄基碳酸酯、T6-Z-2-硝基苄基碳酸酯、T2-Z-4-硝基苄基碳酸酯和 T3-Z-4-硝基苄基碳酸酯。对于氧化分离 (例如用硝酸高铷铵): 1-OT-2-OR-3-Z-苯、1-OT-2-OR-4-Z-苯、1-OT-2-OR-5-Z-苯、1-OT-2-OR-6-Z-苯、1-OT-4-OR-2-Z-苯和 1-OT-4-OR-3-Z-苯。对于还原或烷基化分离 (例如用锂 / 氨或甲基碘): T (2-Z-苯基) 硫醚、T (3-Z-苯基) 硫醚、和 T (4-Z-苯基) 硫醚。对于脱甲硅基分离 (例如用氟化四丁基铵或酸): T 二烷基

- (2-Z- 苯基) 甲硅烷基醚、T-二烷基 - (3-Z- 苯基) 甲硅烷基醚、T-二烷基 - (4-Z- 苯基) 甲硅烷基醚、T-二烷基 - (2-Z- 苯基) 硅烷、T-二烷基 - (3-Z- 苯基) 硅烷和 T-二烷基 - (4-Z- 苯基) 硅烷。

## 组合合成

---

该合成方法通常包括具有至少 2 个选择的阶段，可包括 10 个选择或更多选择。通常，每个阶段的选择数量将不超过约 100，更优选不超过约 50。阶段的数量通常至少为约 3 个，并不多于约 30 个。

这些选择和阶段的数量通常使得至少一些化合物可具有足够的变形，以提供至少一种化合物具有所需特性的适合的可能性。通常这意味着有至少 20 种化合物，但可以有  $10^6$  或更多化合物。

## 反应过程

---

在进行合成反应时，人们可先从许多小球开始，通常有至少  $10^3$  个小球。根据第一阶段中选择数量，将这些颗粒分开相应地装入许多容器中。人们可使用微量滴定井板、分隔容器、柱、凝胶、Terasaki 板、烧瓶、Merrifield 合成容器等。这些颗粒通常被分成几个组，每个组有至少一个颗粒，一般有许多颗粒，通常有 1000 或更多个颗粒。

然后将适合的试剂加入到分隔容器的每个容器中，在阶段中对试剂进行处理，并加入编码试剂和阶段的识别剂。每个阶段将提供所需的反应。一旦反应完成，可洗涤小球，使其不带任何试剂，接着

将所有的小球合并在一起,然后根据下个阶段选择的数量将这些小球分开。重复进行分开小球、接着标记和合成阶段(或反之亦然),然后再合并小球这个过程,直到组合合成完成。

在某些情况下,可在2个或多个容器中进行相同的反应,以便提高与其他选择相比在特殊阶段下进行特殊反应的产物的比例。在其他情况下,一个或多个阶段可包括一部分被搁在边上而未经过反应的小球,以便提高与最终产物相关的变化性。在其他情况下,批料可沿不同的合成路线进行。

为了标出或编码在每个阶段小球上的合成历程,用其特有的识别剂组合来标记与每个选择和阶段有关的小球。另外,人们可使用单个标示物来标出或编码该合成历程。根据所涉及的化学性质,可在包括每个选择的反应之前、之后或同时进行标记。此外,作为对比,可在任何阶段选出样品小球,分解出一部分小球的标示物并解码以验证连接在样品小球上的标示物的正确性。

如上所述,在某些情况下,将部分颗粒分成小部分,然后将每部分颗粒进行不同的反应系列。任何时候都可将这些部分的颗粒再合并成单一的混合物以用于以后的反应。

产物的合成完成之后,或者在从小球分离配位体之后或者在仍然连接的时候,对产物筛选所需的特性。

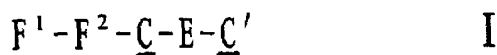
在合成过程不包括阶段地加入相同单元或反应副产物形成的情况下,有可能出现下列情况,即在单一小球上有许多化合物或活性化合物的结构从其反应历程不可得知。根据本发明,由于知道合成历程,人们可在批量地重复合成过程,得到足够量的产物,以便分离出产物和在结构上识别活性化合物。



可制备各种各样的药物类似物,例如下列药剂的类似物:抗高血压剂,例如埃那拉普利尔; $\beta$ -阻滞剂,例如茶心安;抗溃疡药( $H_2$ -受体拮抗药)例如甲腈咪胍和呋喃硝胺;抗真菌剂(胆固醇脱甲基酶(demethylase)抑制剂)例如异康唑;抗焦虑药,例如安定;止痛药,例如阿司匹林、苯乙酰胺和芬太尼;抗生素,例如万古霉素、毒雷素和先锋霉素;消炎药,例如可的松;避孕药,例如孕激素;堕胎药,例如RU-456;抗组胺药,例如扑尔敏;镇咳药,例如可待因;镇静药,例如巴比妥等。

## 识别剂

本发明的识别剂可用式 I 表示:



其中  $F^1-F^2$  是能够连接到载体上和从载体上分离标示物的连接体;

$\underline{C}-E-\underline{C}'$  是能够测定和识别的标示物;

E 是标示物组分,该标示物组分 (a) 可测定,例如用气相色谱法或质谱法可分析出的电泳基团或 (b) 可测定和分离;

$\underline{C}$  和  $\underline{C}'$  是标示组分,该组分可从所有其他标示物中个别地识别出一个标示物,通常可通过改变取代基长度或取代部位而分离,例如,改变色谱保留时间或质谱比例  $m/e$ ;

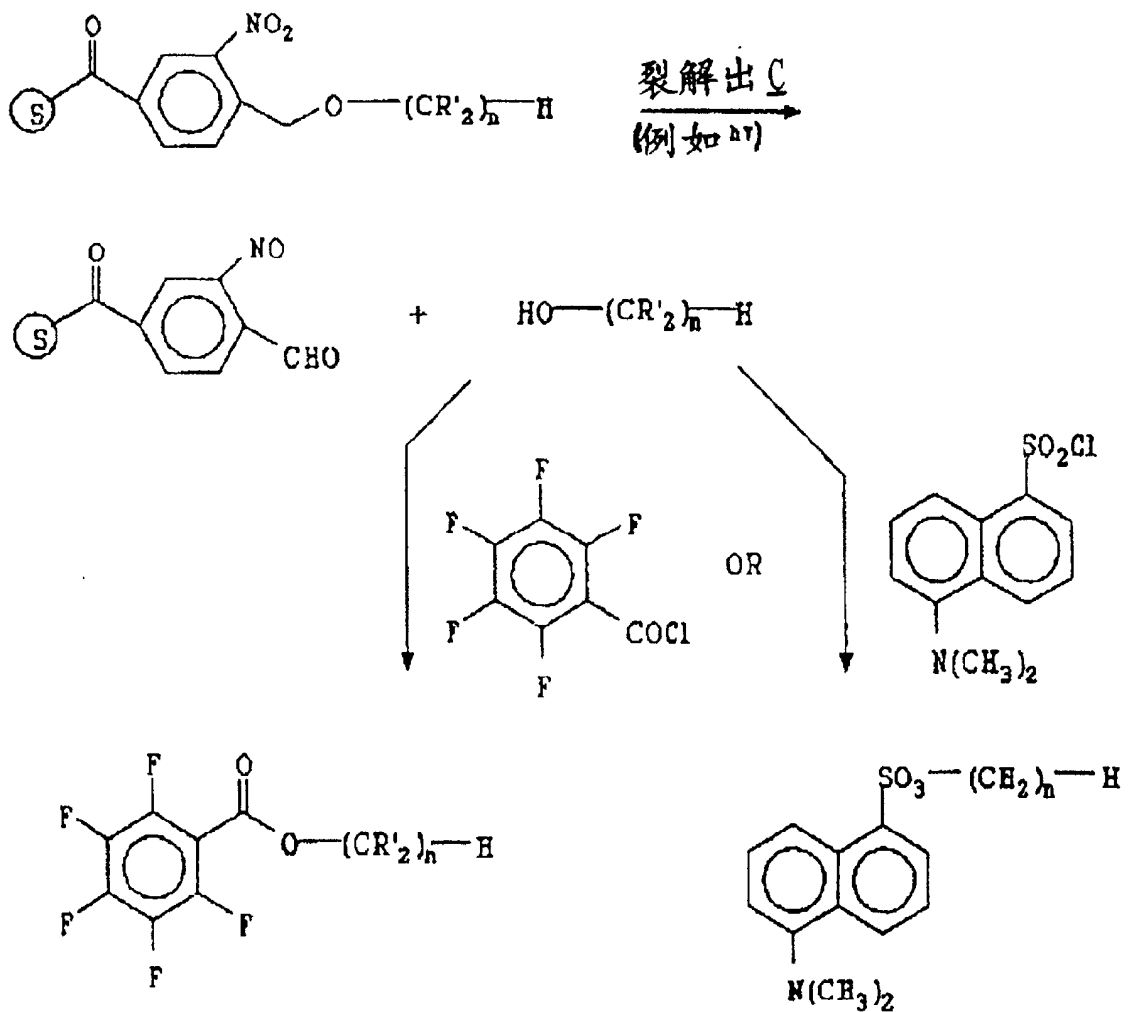
$F^2$  是能够选择分裂以释放标示组分的连接组分;

$F^1$  是提供连接到载体上的官能基;或

当  $F^1$  是可分裂基团,例如 OH 或羧基时,  $F^2$  是单键。

虽然在组合合成过程中一般是在每个适当的阶段和选择加入

式 I 的识别剂, 但是, 部分 E 可在从基质分裂 (优选光化学分裂或氧化分裂) 之前或之后在合成结束时加入。具体地说在 C 含有 OH、 $\text{NHR}^4$  或 SH 时, E 可在分裂之前连接到 C 上。另外, 如 E 是在分裂之后连接, 那么在 C 上的连接点可以是  $\text{F}^2$  的连接点。这在下页的合成路线中举例说明。



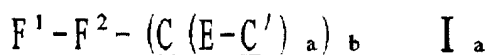
其中  $S$  = 基质, 并且

$$n = 1-40。$$

可用下式表示识别剂与基质的连接:

$F^1-F^2-\underline{C}-E-\underline{C}'+S \rightarrow S-F^{1'}-F^2-\underline{C}-E-\underline{C}'$  其中  $F^{1'}-F^2-\underline{C}-E-\underline{C}'$  表示与基质相连的识别剂残基。例如, 当小球以氨基进行功能化并且  $F^1$  是  $CO_2H$  时, 那么  $F^{1'}$  是  $-C(O)-$ ; 当小球含有不饱和键并且  $F^1$  是  $N_2CH-C(O)-$  时, 那么  $F^{1'}$  是  $=CH-C(O)-$  或  $-CH_2-C(O)-$ 。

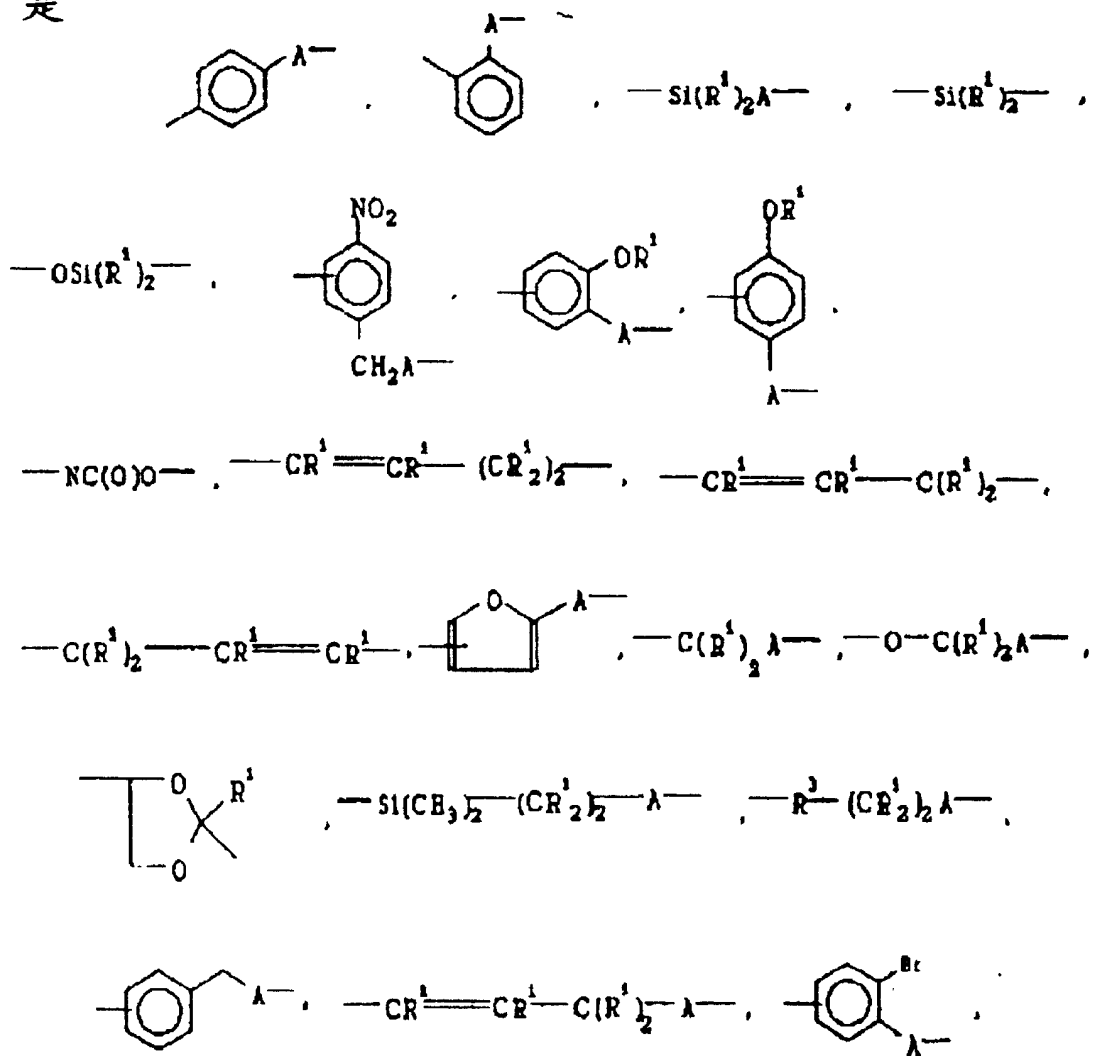
用作识别剂的特别重要的化合物是结构式 I<sub>a</sub> 的式 I 化合物:

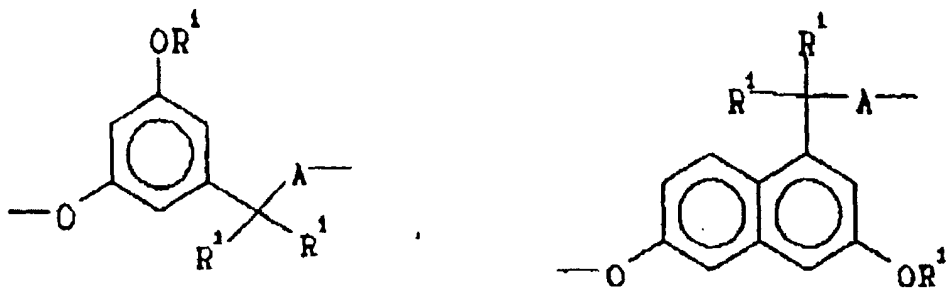
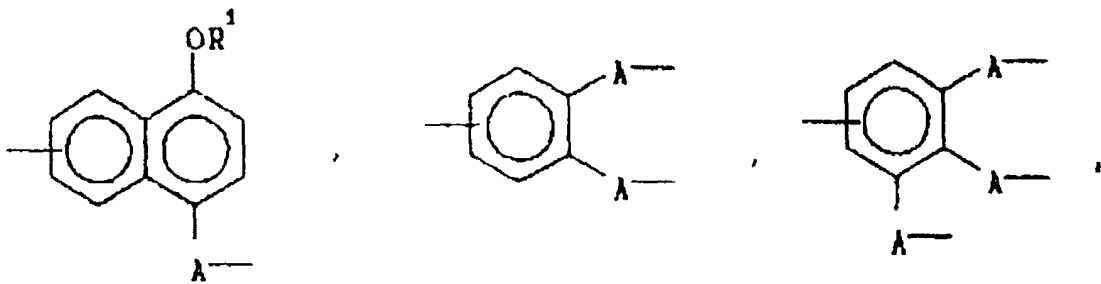
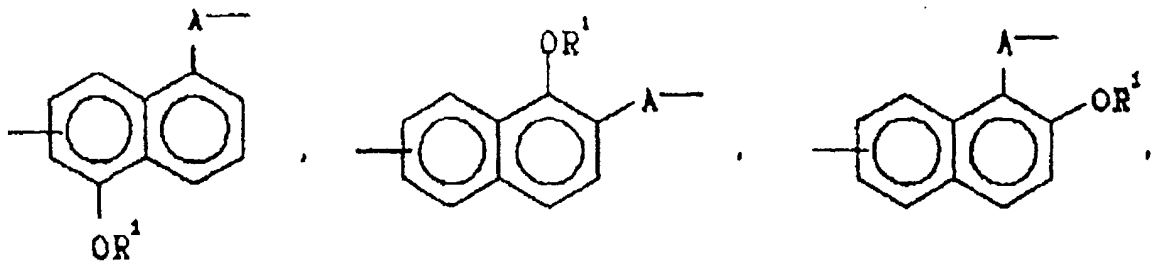
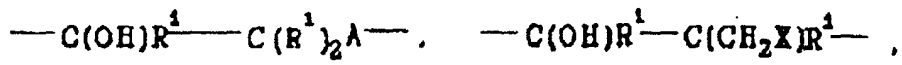
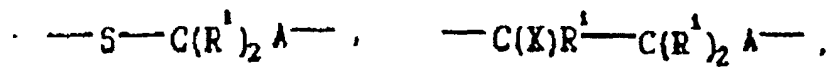


其中:

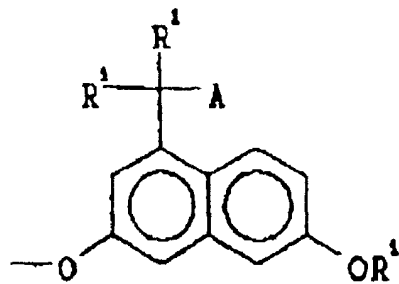
$F^1$  是  $CO_2H$ ,  $CH_2X$ ,  $NR^1R^1$ ,  $C(O)R^1$ ,  $OH$ ,  $CHN_2$ ,  $SH$ ,  $C(O)CHN_2$ ,  $S(O_2)Cl$ ,  $S(O_2)CHN_2$ ,  $N_3$ ,  $NO_2$ ,  $NO$ ,  $S(O_2)N_3$ ,  $OC(O)X$ ,  $C(O)X$ ,  $NCO$ , 或  $NCS$ ;

F<sup>2</sup> 是





或



A 是  $-O-$ 、 $-OC(O)O-$ 、 $-OC(O)-$  或  $-NHC(O)-$ ;

C 是单键、被 1-40 个 F、Cl、Br、 $C_1-C_6$  烷基、 $NR^4R^4$ 、 $OR^4$  或  $NR^4$  作意取代的  $C_1-C_{20}$  亚烷基、或者  $-(C(R^4)_2)_m-Y-Z-Y-(C(R^4)_2)_nY-Z-Y|_p-$ ; 但须  $C+C'$  中的最大碳原子数优选为 20;

C' 是 H; F; Cl; 被 1-40 个 F、Cl、Br、 $C_1-C_6$  烷基、 $NR^4R^4$ 、 $OR^4$  或  $NR^4$  任意取代的  $C_1-C_{20}$  亚烷基、或者  $-(C(R^4)_2)_m-Y-Z-Y-(C(R^4)_2)_nY-Z-Y|_p-$ ; E 是被 1-20 个 F、Cl 或 Br 取代的  $C_1-C_{10}$  烷基; 或者 Q-芳基, 其中芳基被 1-7 个 F、Cl、 $NO_2$ 、 $SO_2R^5$  或取代苯基取代, 其中苯基的取代基是 1-5 个 F、Cl、 $NO_2$  或  $SO_2R^5$ ;

E-C' 可以是  $-H-$ 、 $-OH$  或氨基;

R<sup>1</sup> 是 H 或  $C_1-C_6$  烷基;

R<sup>3</sup> 是  $C=O$ 、 $C(O)O$ 、 $C(O)NR^1$ 、S、SO 或  $SO_2$ ;

R<sup>4</sup> 是 H 或  $C_1-C_6$  烷基;

R<sup>5</sup> 是  $C_1-C_6$  烷基;

a 是 1-5;

b 是 1-3;

m 和 n 各自是 0-20;

p 是 1-7;

Q 是单键、O、S、 $NR^4$ 、 $C=O$ 、 $-C(O)NR^5$ 、 $-NR^5C(O)-$ 、 $-C(O)O-$  或  $-OC(O)-$ ;

X 是诸如 Br、Cl、三氟甲磺酰基、甲磺酰基、甲苯磺酰基或  $OC(O)OR^5$  的离去基团;

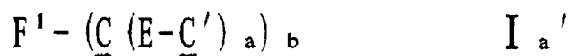
Y 是单键、O、S、或  $NR^4$ ;

Z 是单键; 被 1-4 个 F、Cl、Br、 $C_1-C_6$  烷基、 $C_1-C_6$  烷基、 $C_1-C_6$  取

代烷基[取代基是1-13个F、Cl或C<sub>1</sub>-C<sub>6</sub>取代烷氧基(取代基是1-13F、Cl或Br)]任意取代的亚苯基; (C(R<sup>4</sup>)<sub>2</sub>)<sub>1-20</sub>; 或(CF<sub>2</sub>)<sub>1-20</sub>; 但须, 当Z是单键时, 与Z相邻的一个Y也是单键; 以及芳基是含有不多于10个碳原子和不多于2个选自O、S和N的杂原子的单环或双环芳环。

在式 I<sub>a</sub> 中的 F<sup>2</sup> 的定义中, 所示出的左边的键与 F<sup>1</sup> 相连。

式 I<sub>a'</sub> 的化合物也可用作识别剂:



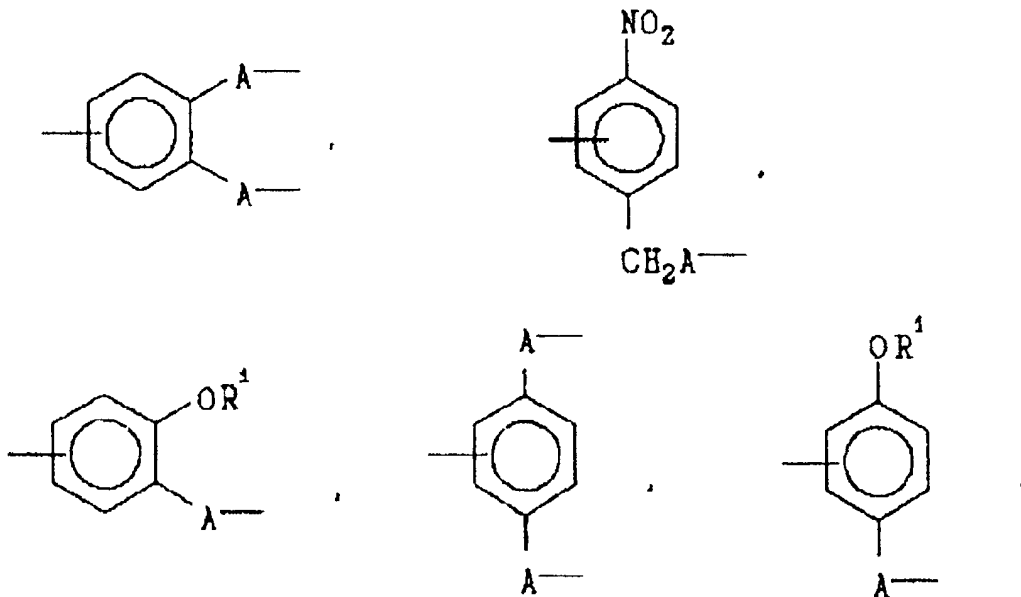
其中

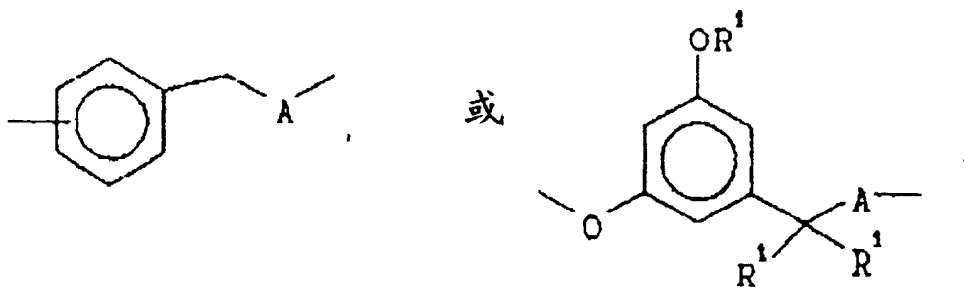
F<sup>1</sup> 是 OH 或 COOH; 并且其余基团的定义与式 I<sub>a</sub> 相同。

优选的式 I<sub>a</sub> 化合物是下面这些化合物, 其中:

F<sup>1</sup> 是

F<sup>2</sup> 是 CO<sub>2</sub>H, OH, CHN<sub>2</sub>, C(O)CHN<sub>2</sub>, C(O)X, NCS, 或 CH<sub>2</sub>X:



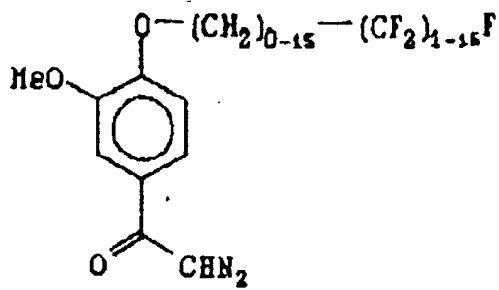
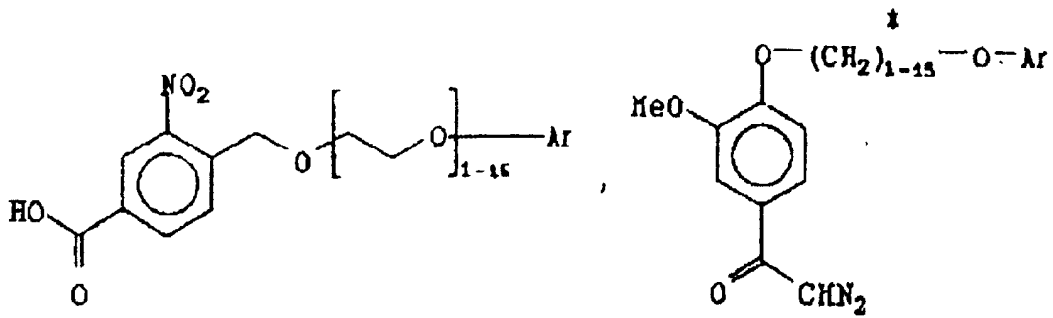
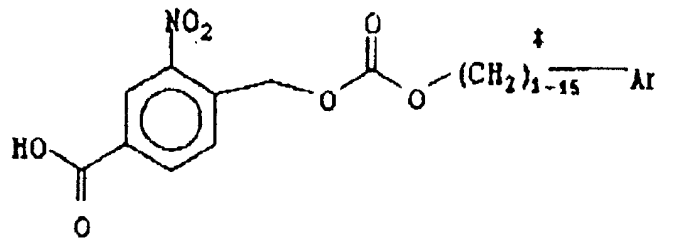
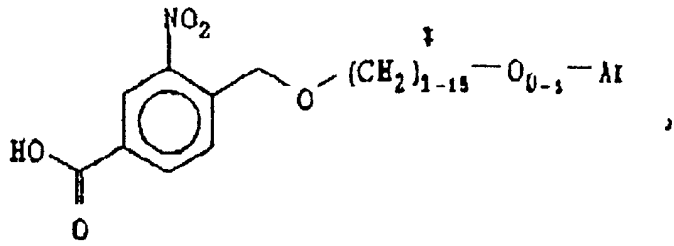
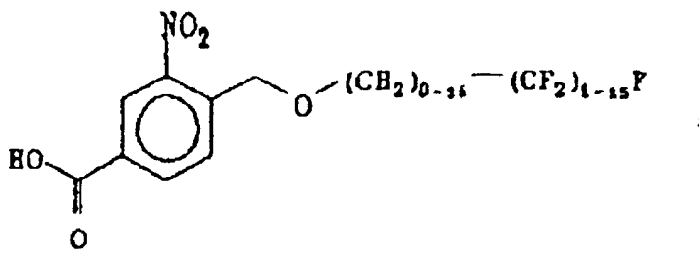


C和C'各自独立地是未取代的或被1-40个F或Cl取代的C<sub>1</sub>-C<sub>20</sub>亚烷基, 或者 [O-(CH<sub>2</sub>)<sub>2-3</sub>]<sub>p</sub>; E是被1-20个F或Cl取代的C<sub>1</sub>-C<sub>10</sub>烷基; Q芳基, 其中芳基是被1-7个F或Cl取代的双环芳环; 或被1-5个F、Cl、NO<sub>2</sub>或SO<sub>2</sub>R<sup>s</sup>取代的Q-苯基; 以及Q是单键、O、-NR<sup>s</sup>C(O)-或-OC(O)-。

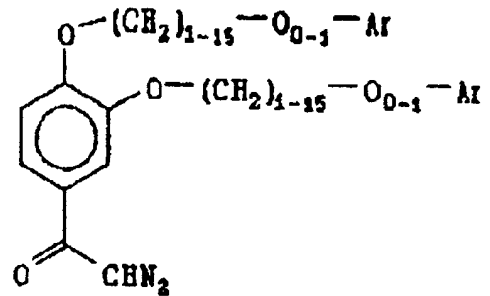
优选的式 I<sub>a</sub> 化合物是其中 -C(E-C')<sub>a</sub> 代表 -(CH<sub>2</sub>)<sub>3-15</sub>-(CF<sub>2</sub>)<sub>1-15</sub>F、-(CH<sub>2</sub>)<sub>3-15</sub>-(CCl<sub>2</sub>)<sub>1-15</sub>Cl、-(CH<sub>2</sub>CH<sub>2</sub>-O)<sub>1-5</sub>-Ar、-(CH<sub>2</sub>CH<sub>2</sub>O)<sub>1-5</sub>-Ar 或 -(CH<sub>2</sub>)<sub>1-12</sub>-O-Ar 的那些化合物; 其中 Ar 是五氟、五氯或五溴苯基; 2, 3, 5, 6-四氟-4 (2, 3, 4, 5, 6-五氟苯基) 苯基, 2, 4, 6-三氯苯基, 2, 4, 5-三氯苯基, 2, 6-二氯-4-氟苯基或 2, 3, 5, 6-四氟苯基。

其他优选的式 I<sub>a</sub> 化合物由下列结构式代表:





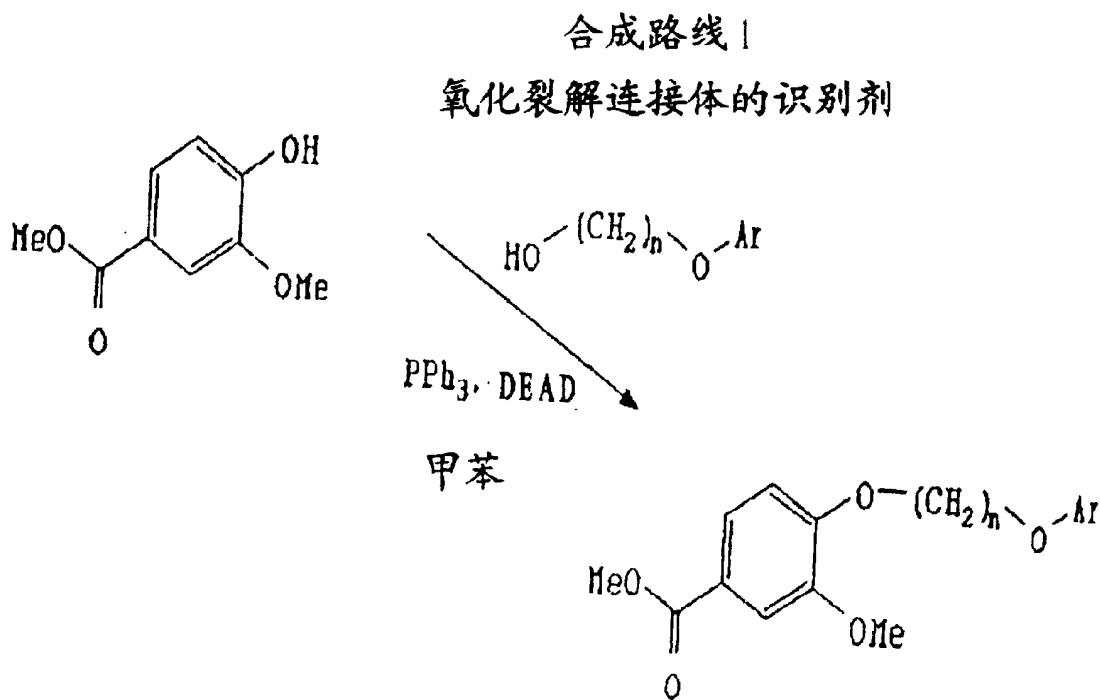
或

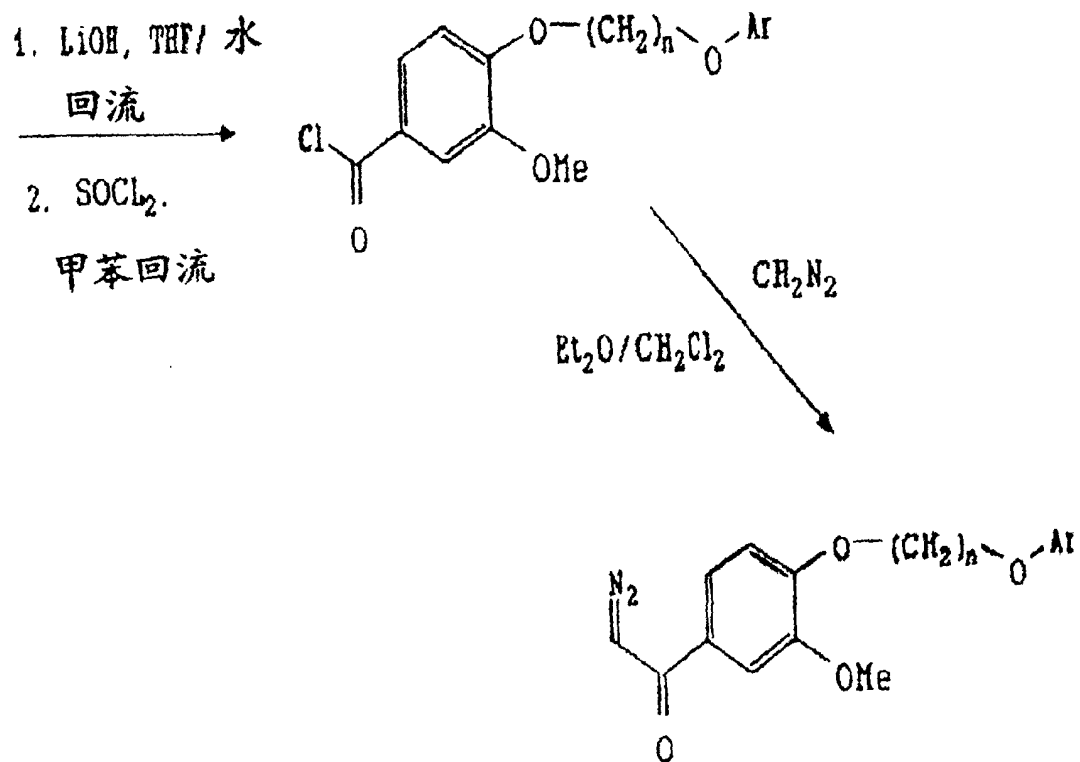


其中 Ar 是五氟、五氯或五溴苯基；2, 3, 5, 6-四氟-4 (2, 3, 4, 5, 6-五氟苯基) 苯基, 2, 4, 6-三氯苯基, 2, 4, 5-三氯苯基, 2, 6-二氯-4-氟苯基, 或 2, 3, 5, 6-四氟苯基。

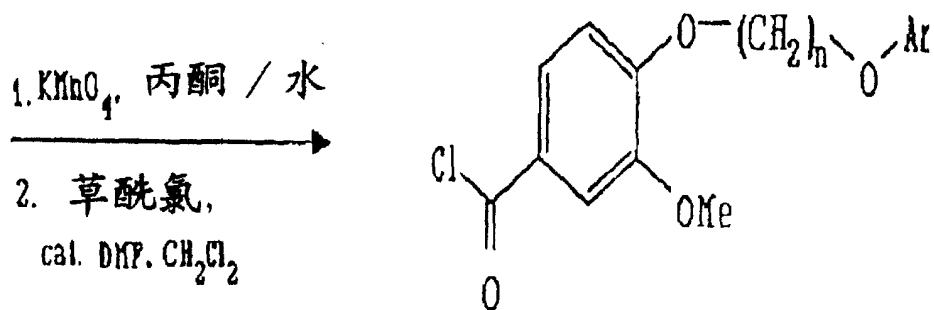
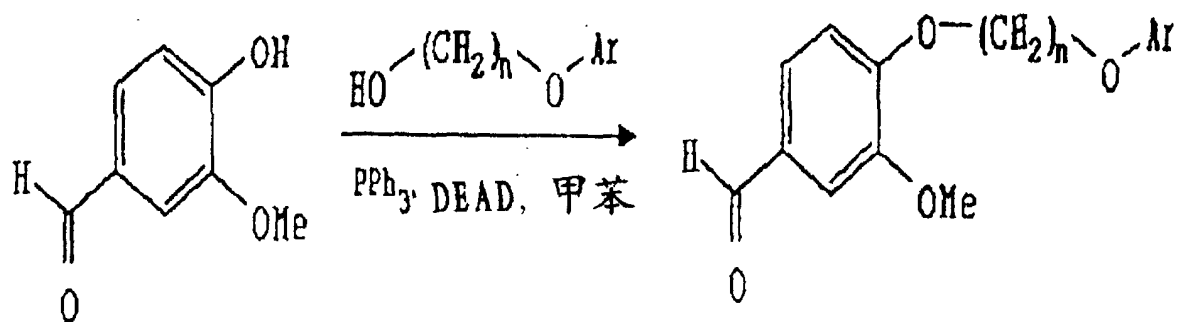
其他优选的式 I<sub>a</sub> 化合物是那些其中 E-C' 是 H, OH 或 NH<sub>2</sub> 的化合物。这些化合物用来在组合合成的最后与 E 反应特别有用, 用来与通过荧光或电子捕获识别的 E 反应更是如此, 如丹磺酰氯或多卤代苯甲酰卤化物。

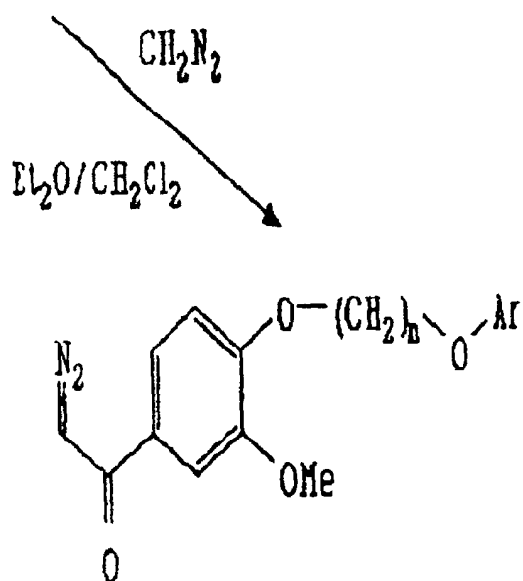
式 I 化合物可根据下列示例性的合成路线或其他本领域技术人员已知的方法制备。





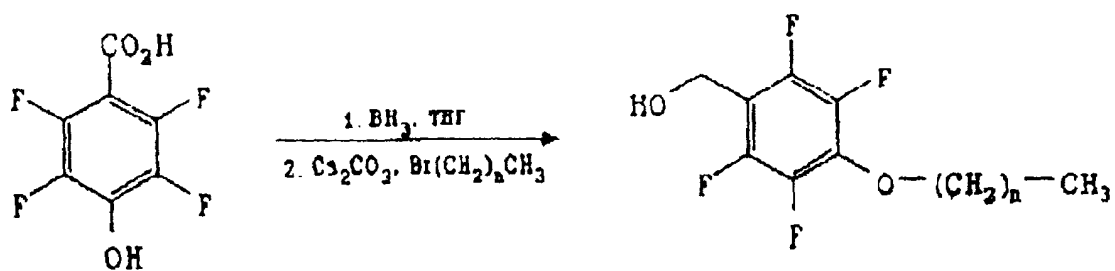
合成路线2  
另一氧化分离连接体的识别剂





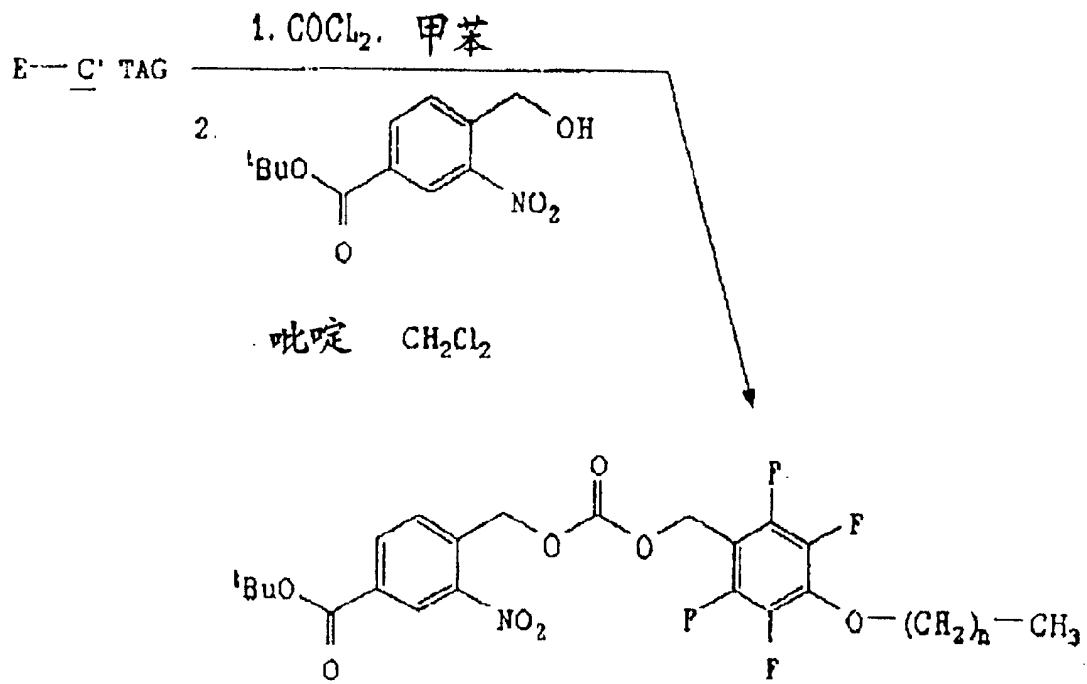
## 合成路线3

E-C' 标示物 (Tag)



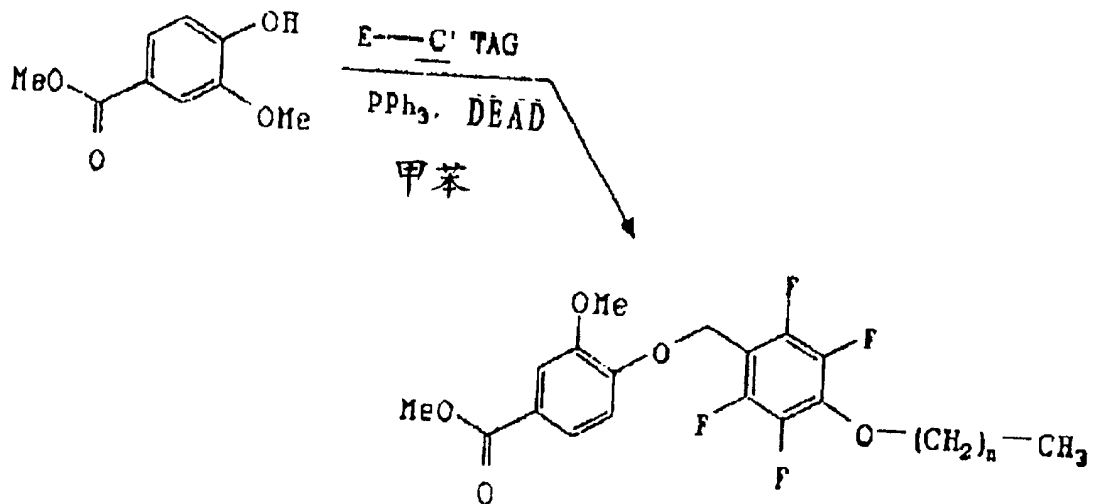
## 合成路线4

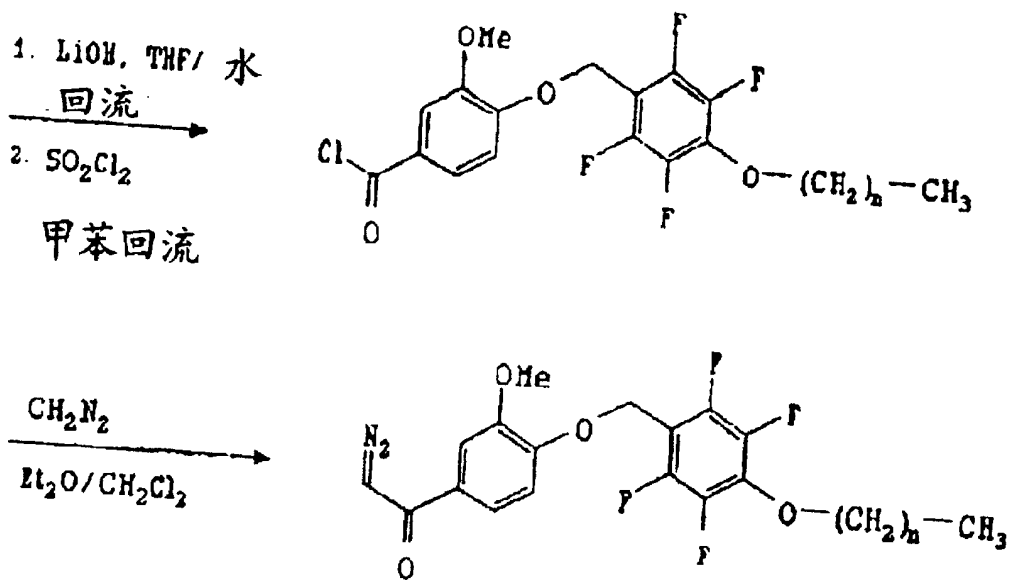
光解裂解连接体的识别剂



## 合成路线 5

## 氧化裂解连接体的识别剂





识别剂可以含有一个或多个相同的标示物。识别剂将是单个的化合物，这些化合物可以从其他化合物中区分出来并将独一无二地识别出不同的选择和阶段。照这样，就可以用相对少量的识别剂，通常少于50个标示物，制备很大的组合库。在每阶段都将加入一组识别剂，该组识别剂就决定了所述的阶段和选择。每个识别剂将以共价键的形式或非共价键的形式与小球或产物连结在一起，但通常是与小球连结在一起。识别剂的组合在每阶段用于提供二进制代码或其他代码，借此代码就决定了选择和阶段。该组识别剂可包括零个或仅仅一个识别剂。

## 标示物

---

至于标示物 ( $\underline{C}$ -E- $\underline{C}'$ )，被使用的标示物具有下列特征：通过取决于 $F^2$ 的方法，优选光解或氧化，可从小球上除去；可被单独稍稍分开，通常是分开着的；在合成条件下是稳定的；可编码阶段或选择，这样可在合成的每一阶段独一无二地确定使用试剂的选择；最好具有识别各种标示物的简便方法，该方法使用不需复杂技术水平即可操作的易得设备；这些设备应相对便宜并能基于少量的分子提供较强的信号；并且标示物应提供足够的敏感性，从而使得标示物能从其他在标示物测定时存在的成分中区分开来。

标示物在结构上可以是有联系的或无联系的，如可以是同系物系列、重复的官能团、周期表中的有关元素、不同的同位素、上述情形的组合等。标示物可用作二进制代码的单元，这样一个标示物可决定两种选择，两个标示物可决定四种选择，三个标示物可决定八种选择，五个标示物可决定三十二种选择，等等。因此，在合成的每个阶段，相对少量的标示物可指明大量数目的选择。对每一阶段来说，包括识别剂的标示物与其他阶段可以有关也可以无关。对于任何组合合成，每个标示物都必须使其能够从其他标识物中区别开来。

每个小球通常具有至少0.01毫微微摩尔、更经常地是具有0.001-50微微摩尔的每种标示物。产物的数量也可至少在相同的范围以及多达至少 $10^4$ 或更多。取决于小球的数目、阶段的数目以及每阶段选择的数目，产生的产物数目通常超过 $10^2$ ，更经常超过 $10^3$ ，并可超过 $10^{10}$ 。



标示物大部分是有机分子。每个标示物除氘外含有的原子数通常少于约100、更经常是少于约80、一般少于约60；这不包括连接部分的原子数，连接部分在从小球上分离标示物时将不被保留。连接部分可以是任意大小，通常除氘外包括的原子数少于约30，更经常是少于约20。连接部分的大小不是关键性的，但应以方便为宜。标示物可形成一类化合物，此时所有的化合物都具有类似的性质；标示物也可以是不同类化合物的组合，此时化合物可以是脂族、脂环族、芳香族、杂环化合物或这些化合物的组合。区别特征可以是若干个重复单元，如烷基部分中的亚甲基，聚亚烷基氧部分中的亚烷基氧基，多卤化合物中的卤素， $\alpha$ -和/或 $\beta$ -取代的乙烯类化合物，其中取代基可包括烷基、氧基、羧基、氨基、卤素等；同位素等。

## 标示物分析

---

依据基团 $F^2$ 的性质，可采用还原、氧化、热解、水解或光解条件从小球上除去标示物。

可借助于物理性质的差别区分标示物，如借助于标示物的分子量，或者气相或液相色谱的色谱保留时间。位置异构体可具有不同的保留时间。如果位置异构体或立体异构体不适于物理分离，那么为了得到所需要的分离，人们可以使用与不同数目的单元如亚甲基或亚乙基氧基相连的可变数目的取代基，例如卤素（如氟）、甲基、氧基或其他侧链。也可使用放射性同位素比例，例如 $^{14}C$ 和 $^3H$ ，这里放射性同位素可提供有差别的辐射。物理性质的差别，如质量数，

可提供有关选择和阶段的信息。

除了使用  $^{14}\text{C} / ^3\text{H}$  比例外,人们还可使用一定组合的非放射性同位素,例如  $-\text{CH}_m\text{D}_n$ , 其中  $m$  是  $0 \sim 3$  并且  $n$  是  $3-m$ 。例如,应用质谱检测至多四个不同甲基的可变数量,人们可确定很大数目的选择。

当  $E$  是单键并且  $C'$  是  $H$  时,从载体上释放得到的标示物具有与标记试剂反应的活性官能团,该标记试剂引入了可检测的标示物成分  $E$ 。为方便起见,官能团可以是双键,尤其是活化的双键,羟基,硫代基、氨基、羧基等。然后,标示物再与过量的标记试剂反应,从而提供用于分析的产物 ( $E-C$ )。这样,作为识别系统的一部分,可以使用各种各样的标记试剂,而这些标记试剂可以不与所需产物的合成策略相匹配。可用于检测的标记试剂包括卤代芳烃(例如全氟苄基溴)、荧光物质(例如丹磺酰氯)、放射性同位素、化学发光物质等。

尽管已经给出了示例性的标示物和反应,但应当理解,许多其他的组合也可以使用。

依据标示物的化学和物理性质,应当选择适宜的分离方法,理想的分离方法选自各种色谱法,包括气相色谱(GC),液相色谱(LC),特别是高效液相色谱(HPLC),薄层色谱(TLC),电泳等。除了色谱法,还可通过质量数用质谱来分离。标示物包括:

对于GC: 具有不同分子量的化学惰性有机分子,包括烷烃、烯烃、芳烃、卤化碳、醚、醇、硅烷、硫醚等,尤其是卤代化合物,这些化合物带或不带其他官能团,可用电子捕获检测或质谱检测(MS)与毛细管GC分离相结合进行分离,以及用原子发射检测与毛细管GC分离相结合进行分离的带有有机化学中非常见元素(例如  $\text{Sn}$ ,  $\text{Ge}$ ) 的化合物;

对于LC、HPLC或TLC: 见上述对于GC的化合物, 可方便分离的化合物是直链醚或烃, 这些化合物带有用于分离后经放射性检测的放射性同位素取代基或放射性同位素组合取代基, 或者这些化合物带有用于分离后经荧光检测的适宜取代基;

对于电泳: 见上述, 特别是官能团带电分子, 例如阳离子或阴离子, 尤其是有机或无机酸基, 这里的分子可通过带有在电泳中检测的可检测放射性同位素或荧光性物质而进一步改性;

对于质谱: 见上述, 特别是由于下列因素引起的不同质数: 不同的同位素, 相同官能团或不同官能团的不同数目, 同系列的不同数目, 或这些因素的组合。

标示物彼此之间的分离可涉及到单一的技术或组合的技术, 例如色谱和电泳; 气相色谱和质谱等。

本发明的标示物具有使其在很低含量的情况下可被检测的性质, 检测到的含量通常不大于毫微摩尔, 优选不大于微微摩尔或更少, 更优选不大于毫微微摩尔或更少, 检测是在很高含量的其他化合物存在下进行的。为此, 可能用到特殊的原子取代, 用以提供易被检测的标记。这些取代包括:

(a) 被电负性元素如氟或氯取代, 检测时用电子捕获检测与毛细管GC或负离子质谱检测相结合;

(b) 被非常用元素(不包括C、H和O)取代, 检测时用原子发射检测与毛细管GC相结合;

(c) 被几个非常用元素取代, 用原子发射检测以确定元素间的比例;

(d) 被放射性元素如 $^3\text{H}$ 取代, 用放射自显影法或闪烁计数法与

LC、TLC或电泳相结合检测；

(e) 被具有不同发射频率的若干个放射性元素如 $^3\text{H}$ 和 $^{14}\text{C}$ 取代，用放射自显影法或闪烁计数法确定不同放射性元素间的比例。

对于单元素取代(上述a, b, d), A个标示物(可检测出其存在与否)的可分离的混合物能够编码出多达 $2^A$ 个不同的合成方式。对于多元素取代(见上述c和e), A个标示物(各自具有B个可区分的状态, 例如不同的 $^3\text{H}/^{14}\text{C}$ 比例, 不同的 $\text{Si}/\text{Sn}$ 比例)的可分离的混合物能够编码出多达 $B^A$ 个不同的合成方式。

存在有许许多多的同位素, 这里同位素的存在或比例可提供关于阶段和选择的信息。同位素可以是放射性的或非放射性的。特别有意义的同位素包括氘、氚、 $^{14}\text{C}$ 、 $^{32}\text{P}$ 、 $^{131}\text{I}$ 等。

通过使用同位素改性化合物的混合物, 人们能够极大地拓宽从单个标示化合物获得的信息, 而区别该标示化合物的唯一依据是同位素的存在。例如, 人们可制备一定比例的氢与氘的混合物, 其中各种比例的区别可小到10%。通过用另一种原子, 例如氟, 取代氢, 人们可得到氢、氘和氟的不同变化的混合物, 从而提供很大数目的不同的可区别标示物。

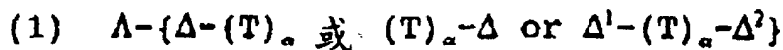
其他可能涉及到的基因可以是芳环, 该芳环可在不同的位置用不同的官能团取代。因此, 通过取代的苯环, 这里取代的位置和取代的性质可被确定, 人们可获得可被区分的许多分子以及阶段和选择方面的信息。例如, 如果C保持不变, 当E是多卤代芳环时, 人们可通过E上的取代类型进行检测和区分。

也存在使用荧光标示物的可能性。尽管单独的荧光标示物用很大数目的选择不能足以确定很大数目阶段, 但参考上面所述, 通

过基于C或C'中的变化提供分离荧光标示分子的方法,人们通过标示物的荧光可分别检测到标示物。

与具体小球相连接的标示物混合物可被分割开,并可进行初步分离,此处需要分别检测每个标示物。一旦该组标示物被分离开,就可依据标示物的具体官能团和与众不同的性质分析每个标示物。可以使用的检测具体标示物的各种技术包括放射自显影法或闪烁计数法、电子捕获检测、负离子或正离子质谱、红外光谱、紫外光谱、顺磁共振光谱、荧光等。

另一个成分可具有至少6个不同的标记,这些标记在一个化合物子库(kil)或普通介质中相互关联,每个标记带有一个本质上化学惰性的可区分部分,彼此之间分子量不同。所述标记的结构式如下:



其中 $\Lambda$ 是连接基团,该连接基团带有与固体载体相连接的官能团和用于与固体载体相分离的官能团,用于分离的官能团可以包括在连接固体载体的官能团之中;

$\Delta$ 是区分基团,该基团使得每一个标记通过其物理性质以及通过除荧光以外的方法与其他标记相区分,从而提供能够编码多步骤合成方法的一套标记;该区分基团包括分割后的剩余官能团,该官能团已事先与连接基团相连接;

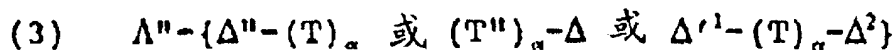
$\Delta^1$  和  $\Delta^2$  是区分基团部分, 它们一起限定了区分基团; 它们连接在一起时就落入  $\Delta$  的定义之中;

T 是可检测基团, 当其与区分基团连接时使得能够检测到低含量的标记, 这里可检测基团可以存在在子库中的标记上或后来被加到区分基团上, 并且如果连到连接基团上, 可检测基团包括分离后剩下的任何官能团, 这些官能团事先与连接基团相连接;

以及  $\alpha$  是 0 或 1, 表示可检测基团存在或不存在;



其中所有的符号同前述定义, 但下面的定义除外: SS 是固体载体;  $\Delta^1$  是以共价键形式与 SS 相连的连接基团; 以及  $\beta$  对每个固体载体来说是整数, 该整数至少为 6 并通常不大于约 30;



其中所有的符号同前述定义, 但下面的定义除外:  $\Delta''$  是氢或经过导致与固体载体分开的光解裂解、消去或其他化学反应后的连接基团的残基;  $\Delta'^1$  或  $\Delta''$ , 作为标记与固体载体分开的结果, 分别是  $\Delta$  或  $\Delta^1$  或改性的  $\Delta$  或  $\Delta^1$ ; 作为标记与固体载体分开的结果,  $T''$  是

T或改性的T;



其中所有的标记同前述定义, 并且  $\Delta''$  是单键或与T连接后连接基团的剩余部分; 但有一个附加条件是, 只能有一个  $\alpha$  是1。

### 鉴定

为了确定产物的所需性质, 使用了许多鉴定手段和技术。

在筛选小球时, 人们经常地或者使用单一的小球或者使用小球混合物, 确定小球或混合物是否是显示活性。于是混合物可包括10、100、1000个或更多的小球。这样可迅速将一大类化合物筛分成较小类的化合物。

有一个技术是人们在连接具体的双原子分子如受体时所感兴趣的。在这种情况下, 人们可进行两步筛选, 从而人们首先使用连接作为初次筛选, 接着再通过具有生命力细胞的生物活性进行第二次筛选。

鉴定可以使用单一颗粒或几类颗粒或其组合分步进行。例如, 进行组合合成后, 约50至10,000个颗粒的组可在分立的容器中细分。在每个容器中, 对于每个颗粒, 与颗粒相连的产物的一部分被释放出来。分级释放可能是产物与颗粒不同连接的结果或是使用有限

量的试剂、条件等的结果，这样每个颗粒释放出的产物分子的平均数目要少于每个颗粒的产物分子的总数目。那么人们在一个较小体积中可得到产物混合物。该混合物可用于连接鉴定，这里，连接可以抑制已知的连接配位体与受体的连接、活化或抑制细胞的新陈代谢过程等。检测连接活性可以使用各种各样的鉴定条件，这在下文将给以描述。一旦一个实验组显示出活性，就可以通过相同或不同的鉴定筛选单个颗粒。当然，人们也可采用三步或四步法，这里将大组分成较小的组等等，最后筛选出单一颗粒。在每种情况下，颗粒上的部分产物将会释放出来，而得到的混合物将用于适宜的鉴定中。鉴定可以是相同的也可以是不同的，在以后的步骤或最后步骤中会用到更复杂和更耗时的鉴定。

人们也可提供空间排列，其中将颗粒分布在多孔盘中，多孔盘的每个孔中具有0或1个颗粒。

可以使用科目分类法找到具有催化性质的化学物质，催化性质的实例有水解活性，例如酯酶活性。为此目的，人们可将小球置入被扩散性实验基质包围的半固体基体中。如果催化活性可以通过不干扰基体的方法在局部测定，例如通过改变光的吸收或检测裂解基质的荧光测定，那么在催化活性区的小球可被分离出来，它们的标记也可以消除掉。

除了催化活性外，还可开发出带有抑制活性或活化活性的化合物。可以找到这样的化合物，它们抑制或活化酶或者阻止连接反应。为检测到抑制酶的小球，该小球具有带此所需性质的连接产物，有利的是可从小球中释放产物，使该产物扩散到半固体基体中或滤料上，而在基体中或滤料上可现察到抑制、活化或阻止作用。构成可



见的或换句话说可检测到的抑制、活化或阻止区域的小球，可被挑选出来，并且标示物可以去掉编码。在这种情况下，合成产物的一部分必须通过可裂解的连接、优选对光不稳定的连接与小球相连，而标示物的一部分仍与小球相连，挑选出来后可通过与以前不同的方法释放。

可以使用渗析膜，其中小球层与放射性标记配位体 / 受体对层分开。小球层可用紫外光照射，从小球中释放出的产物将扩散到配位体 / 受体对层，这里，放射性标记配位体将与化合物对受体的亲合性成比例地释放出来。放射性标记配位体会反过来扩散到小球层中。由于放射性标记与小球最接近，就可以分析与放射性发射有关的小球。

寻找到具有生物活性的产物具有特别重要的意义。在某些应用中，需要寻找到对存活细胞产生作用的产物，例如，能够抑制细菌生长，抑制病毒生长，抑制基因表达或活化基因表达。在小球上筛选化合物可以很容易地实现，例如，通过将小球嵌入到半固体介质中，从小球释放出的产物分子库（而小球被保留）使得化合物向周围的介质扩散。可以观察到某些效果，例如，细菌带产生了菌斑。还可观察到生长抑制区域、生长活化区域或对基因表达产生作用的区域，并且在区域中心的小球可以挑选出并加以分析。

一个鉴定方案涉及到凝胶，其中被作用的分子或体系例如细胞被基本上均一地置入凝胶中。可以使用各种各样的凝胶试剂，例如聚丙烯酰胺、琼脂糖、明胶等。然后将颗粒分布在凝胶中，以便颗粒之间获得足够的分离，从而进行单独检测。如果所需产物具有水解活性，凝胶中存在有能够提供荧光产物的基质。于是，人们为荧

光筛选凝胶，并机械地选择与荧光信号有关的颗粒。

人们可在凝胶中置入细胞，其作用是产生细胞带。如上所述，可将颗粒分布开。当然，人们也可在凝胶中放置网格，限定有一个颗粒或没有颗粒的区域。如果细胞毒性是鉴定标准，人们可释放出产物，培育足够长的时间，接着将活体染剂分布在凝胶中。然后就可区分出吸收染剂的细胞或没有吸收染剂的细胞。

如上所述，对细胞可进行遗传工程，以便在转换信号时能有所显示。存在许多其基因已知的受体，这些基因的表达可被活化。通过将外生基因嵌入某一部位，而在该部位的基因受到对该受体敏感的促进剂的转录控制，可产生提供检测信号如荧光信号的酶。接着就可分析与荧光细胞有关颗粒的反应历程。

### 化合物库和子库

---

为方便起见，提供了化合物库和 / 或子库。化合物库包括向其中加入产物和标示物库的颗粒，从而可筛选连接小球的产物；或者包括从小球中除去并且为了筛选被单独分组或以 $10-100-1000$ 个化合物为一组的产物。子库提供在进行化合物库合成时用作标示物的各种试剂。子库通常具有至少4个、经常是至少5个、更经常是至少10个在分立容器中的不同化合物，并可包括至少 $10^2$ 个不同的彼此分开的有机化合物，通常不多于约 $10^2$ 个、更经常是不多于约36个不同的化合物。对于二进制测定，检测方式通常为化合物分析所共有，因此要有一个共有的生色团、一个共有的被检测原子等。当予先制备每个识别剂时，每个识别剂的特征在于具有一个可区分

成分, 该成分为能够被物理测量确定的选择和阶段编码, 并包括共有至少一个普通官能团的基团或全部化合物。

另一方面, 子库可以提供这样的试剂, 它们相互组合可提供各种各样的识别剂。在这种情况下, 子库包括若干个分开的第一类官能团、经常是双官能团、通常是四个或更多个官能团的有机化合物, 一般是一个官能团为合成中的每个步骤服务, 其中带官能团的有机化合物共有一个相同的官能团, 并能为至少一个测定性质所区分。而且, 人们还要有至少一个、通常是至少两个第二类有机化合物, 该类有机化合物能够与官能有机化合物的官能团反应, 并能依据每个所述第二类有机化合物的量形成可被区分的混合物。例如, 人们可使用二元醇、氨基酸或乙醇酸, 其中各个双官能团化合物可根据存在的氟或氯原子数目区分, 从而确定阶段; 人们也可以使用碘甲烷, 其中一类碘甲烷没有放射性同位素, 另一类具有 $^{14}\text{C}$ , 还有一类具有一个或多个 $^3\text{H}$ 。通过使用两类或更多类的碘甲烷, 人们可提供能够被其放射同位素发射所测定的各种混合物。另外, 人们可以使用一些能够用于二进制代码的第二类有机化合物。

如前所述, 人们可使释放后的标示物与能够被检测的分子反应。在这种情况下, 标示物可以相当简单, 具有与连接颗粒和连接检测部分相同的官能团。例如, 通过与羟基羧基连接, 释放出了羟基, 该羟基可用能够被检测的分子酯化或醚化。再如, 通过在二进制代码中使用氟代烷基和氯代烷基的组合, 氟和 / 或氯基团的数目可以确定选择, 而碳原子的数目表示阶段。

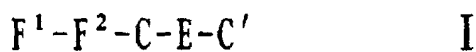
具有重要意义的化合物的基团包括与下列基团连接在一起的连接体: 取代的邻硝基苄氧基、2, 3-二氢化茛苄氧基或苄氧基, 或其

他能够进行光解或其他选择性解离的基团。连接基团可以是2-20个碳原子的亚烷基,聚亚烷基氧基,特别是2-3个碳原子的亚烷基氧基,4-8个碳原子的环烷基,卤代烷基,特别是2-20个碳原子的氟代烷基,一个或多个芳环等等,其中连接体通过带有不同数目的单元和/或取代基提供各种基团之间的差异。

可以商业产品形式提供单一的颗粒或若干种颗粒,尤其是当颗粒显示出令人感兴趣的性质时更是这样。基于有关的标示物,可以去除反应历程的编码。这样就可以大规模合成产物。当反应历程能够明确地决定结构时,就可运用相同或类似的反应系列大批量地生产产物。当反应历程不能明确地决定结构时,人们就要大批量地重复反应历程;并用得到的产物进行结构分析。在一些情况下,可能发现组合化学的反应系列不是大量生产产物的优选途径。

因此,本发明的实施方案是包括若干个相互分离的有机化合物的子库,每个化合物的特征在于其具有可区分的成分,该成分能够编码出可被物理测量测定的至少一个不同信息量子,并共有至少一个普通官能团。一个优选方案是包括至少4个不同的官能有机化合物的子库。

更优选的是其中所述官能有机化合物具有结构式为



的子库,其中 $F^1-F^2$ 是使其与固体颗粒连接并从固体颗粒上除去的连接体; $\underline{C}-E-\underline{C}'$ 是能为物理测量测定的标示物,尤其是,其中所述官能有机化合物的不同之处在于含有的亚甲基和/或卤素、氮或硫的数目不同。

下面的子库也是优选的:其中 $\underline{C}-E-\underline{C}'$ 部分是经光化学除去的子

库，或者其中 C-E-C' 部分是由氧化、水解、热解或还原除去的子库。

在一个实施方案中，本发明是包括至少 6 个不同组分的成分，每个组分带有可区分的部分。组分的特征在于，每个部分在本质上都是化学稳定的或惰性的，并带有与每个其他部分不同的可识别性质。每个部分都与连接基团相连，连接基团具有能够在连接基团和单个分立的固体表面之间形成共价键的活性官能团，或者每个部分与在低于 1 毫微摩尔能够被检测的基团相连，但附加条件是，当各个部分与连接基团相连时，组分在物理上是彼此分立的。优选地，固体载体是小球。在一个实施方案中，每个组分包括与单个分立固体表面相连的不同化合物分子，其中分子在固体表面上。优选地，本发明的这些部分确定分子核上的同系列和 / 或取代系列。

本发明还是关于包括至少 100 个独特固体载体的化合物库。在该化合物库中，每种固体载体都具有 (1) 与固体载体连接的单个化合物，其作为与载体相连的主要化合物；和 (2) 许多标示物，例如不能程序化的标示物，其中标示物是单个标示物分子，该种分子由于物理上可分开因此是物理上可区分的，并且是取代的，从而在低于约 1 毫微摩尔时可被检测出来，或者带有与取代基相连的官能基，该取代基可在低于 1 毫微摩尔时被检测出来。优选地，在该化合物库中，每种固体载体带有至少 6 个标示物。在另一个实施方案中，标示物在化合物库中确定二进制代码，而二进制代码为用于合成固体载体上化合物的合成计划编码。

本发明还提供决定合成计划和确定二进制代码的方法，其中合成计划是由彼此分离的一系列物理上不同的标示物编码的。在该方法中，为确定合成计划的每个阶段至少使用两个标示物，这样至

少有六个标示物。该方法的步骤包括借助于标示物的物理差别分离标示物,并检测标示物。合成计划由不同标示物的二进制代码确定。

本发明的化合物可用作镇痛药和 / 或用于治疗炎症,特别是在氮杂三环的情况下,氮杂三环起到神经激肽1 / 舒缓激肽受体拮抗剂的作用。苯并二吡啶庚因库的化合物可用作肌肉松弛剂和 / 或镇定剂和 / 或镇静剂。23.5 百万混合酰胺库 (实施例3) 的化合物能够起到内皮激肽拮抗剂的作用,可用于治疗高血压或Raynaud 综合症。

## 实施例1

### 放射性标记的标示物

---

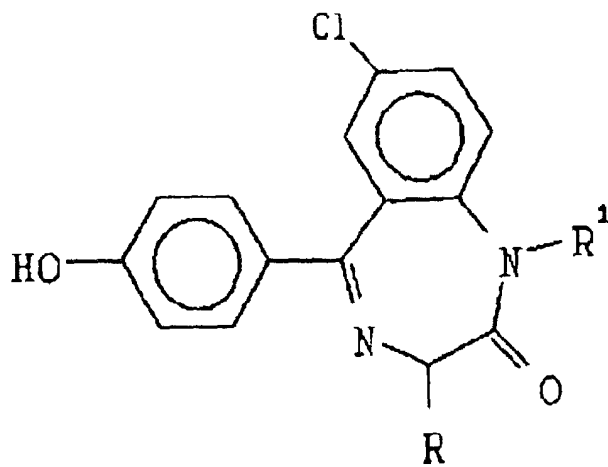
使用的标示物是直链烷基- $\alpha$ ,  $\omega$ -二醇的单甲基醚。该二醇带有 $N+2$ 个碳原子,而 $N$ 表示阶段。甲基可以是放射性标记试剂,该试剂各种比例为 $1/1-m/1$ 的 $^3\text{H}/^{14}\text{C}$ ,其中 $m$ 是选择数。双重放射性标记能够精确测定标示物中存在的氘。由10个不同的亚烷基和10个不同的放射性标记比例,就能得到 $10^{10}$ 个独特的十元组标示物。为了连接上标示物,先使标示物与活化剂反应,如与光气反应形成氯甲酸酯,接着再与 $F^1-F^2$ 组分反应。在这种情况下, $F^1-F^2$ 是以叔丁基酯保护的邻硝基对羧基苄醇。在合成阶段进行时,随时直接向小球加入去酯的识别剂,其中小球具有共价键结合的氨或羟基,与酸形成酰胺或酯,其中酸用常规的化学法活化,如用碳化二亚胺偶合法活化。在序列合成结束时,用各种受体或酶筛选小球以确定具体

特性。接着分离出显示某种特性的小球，用HPLC分割和分离标示物得到一系列二醇单甲基醚，对该醚用常规的放射性同位素识别法分析放射性。例如，如果从HPLC柱流出的第一个和第二个标示物分别具有5:1和7:1的 $^3\text{H}/^{14}\text{C}$ 比例，那么显示活性的产物是用阶段1的5号试剂和阶段2的7号试剂合成的。

## 实施2

### 苯并二吡啶因库

组合的苯并二吡啶因库包括30个结构式VIII的化合物：



VIII

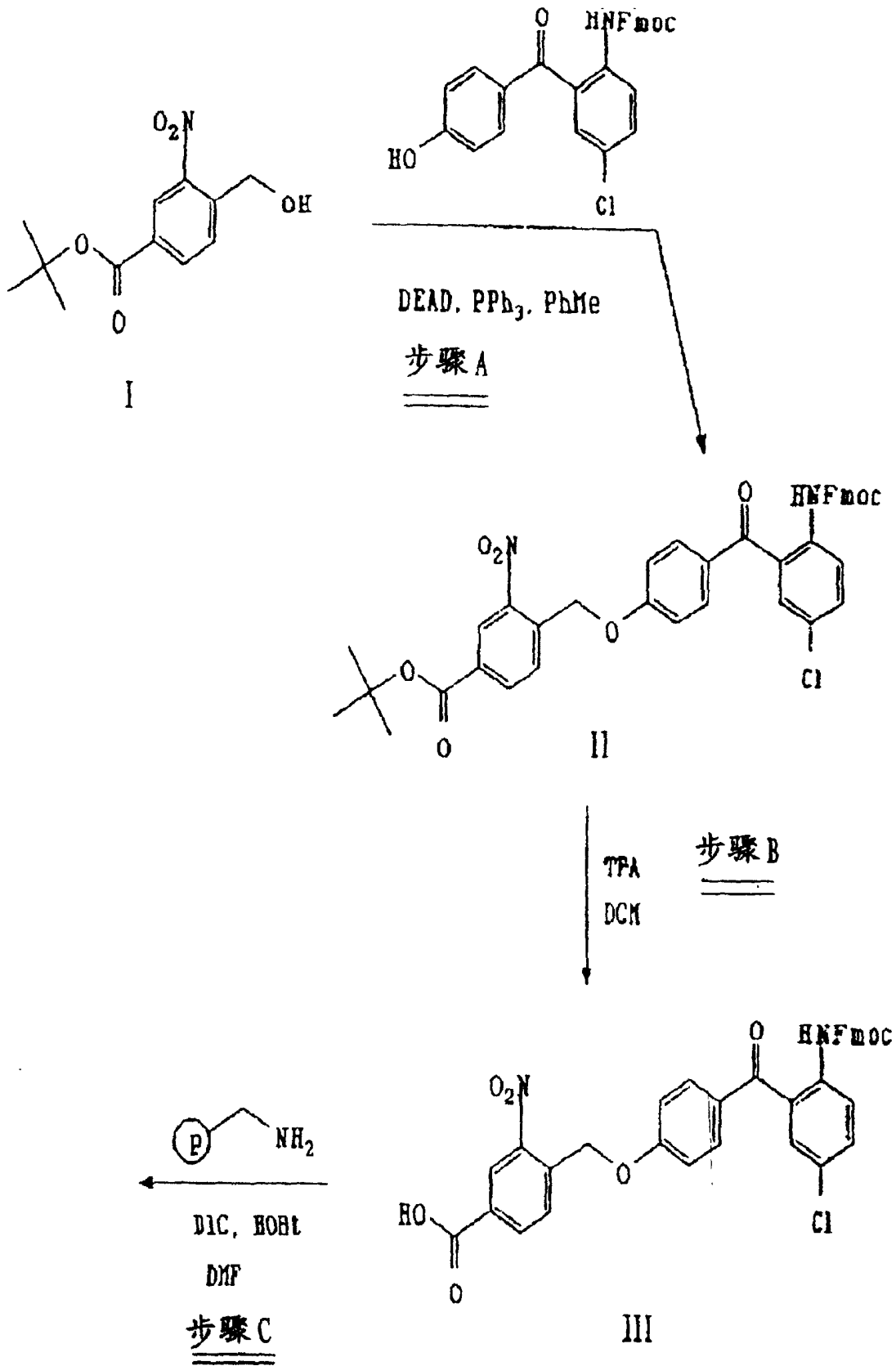
其中：

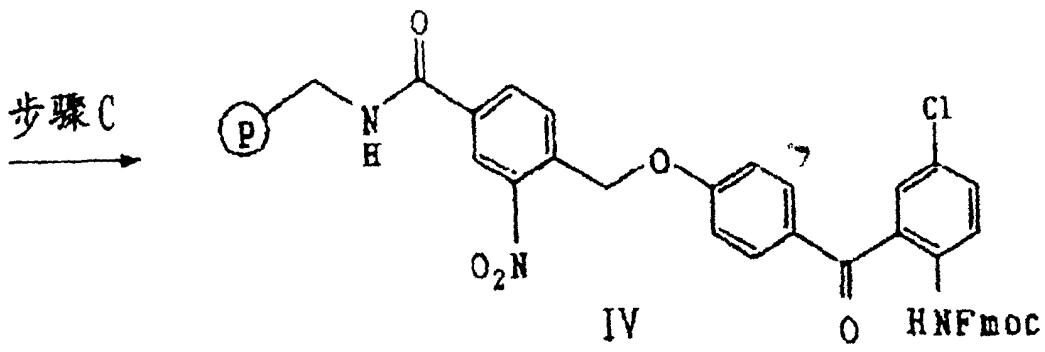
R 是  $\text{CH}_3$ ,  $\text{CH}(\text{CH}_3)_2$ ,  $\text{CH}_2\text{CO}_2\text{H}$ ,  $(\text{CH}_2)_4\text{NH}_2$ ,  $\text{CH}_2\text{C}_6\text{H}_4\text{OH}$ , 或  $\text{CH}_2\text{C}_6\text{H}_5$  和  
R' 是 H,  $\text{CH}_3$ ,  $\text{C}_2\text{H}_5$ ,  $\text{CH}_2\text{CH}=\text{CH}_2$ , 或  $\text{CH}_2\text{C}_6\text{H}_5$

---

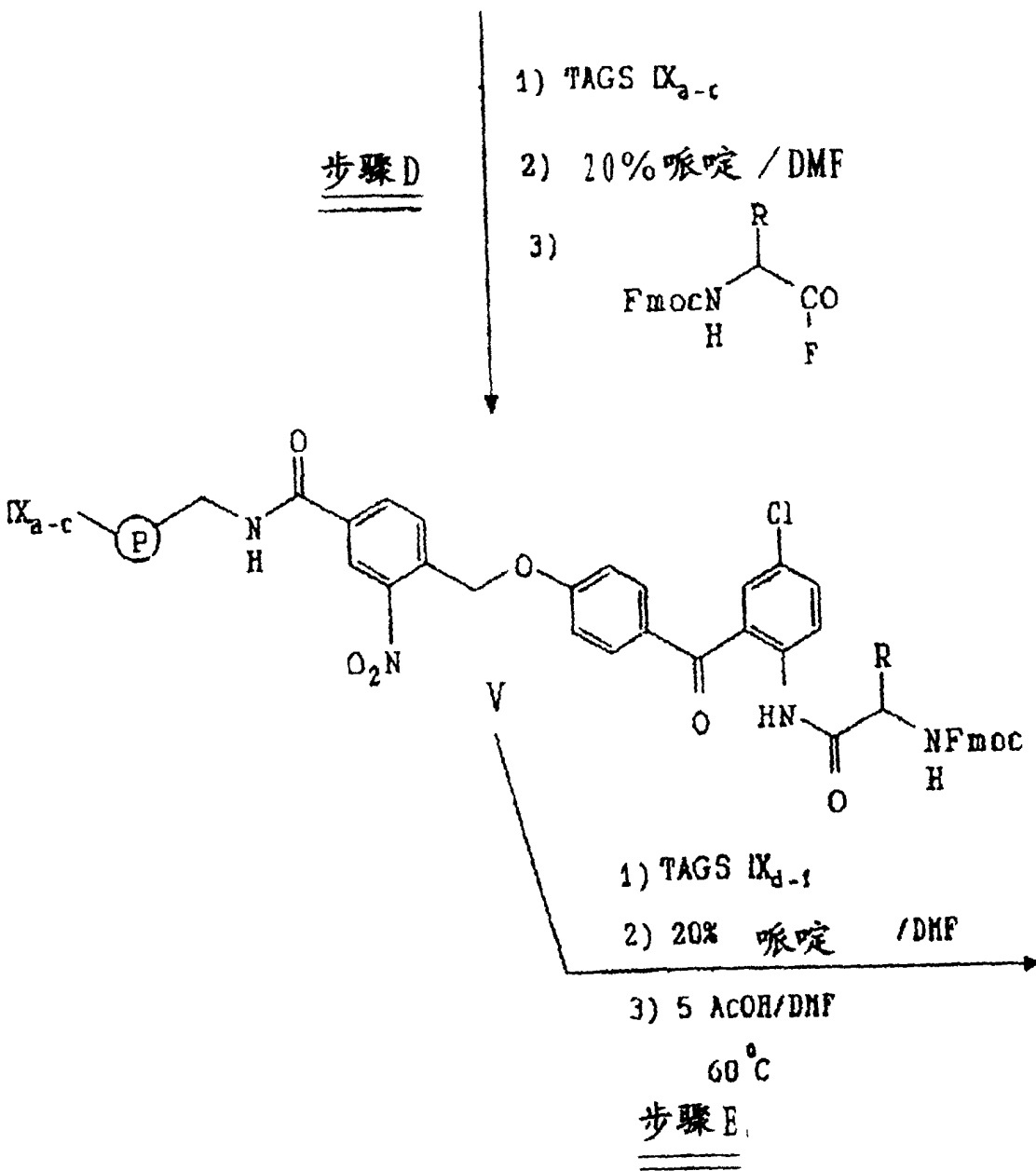
该苯并二吡啶因库是通过下列合成路线制备的。

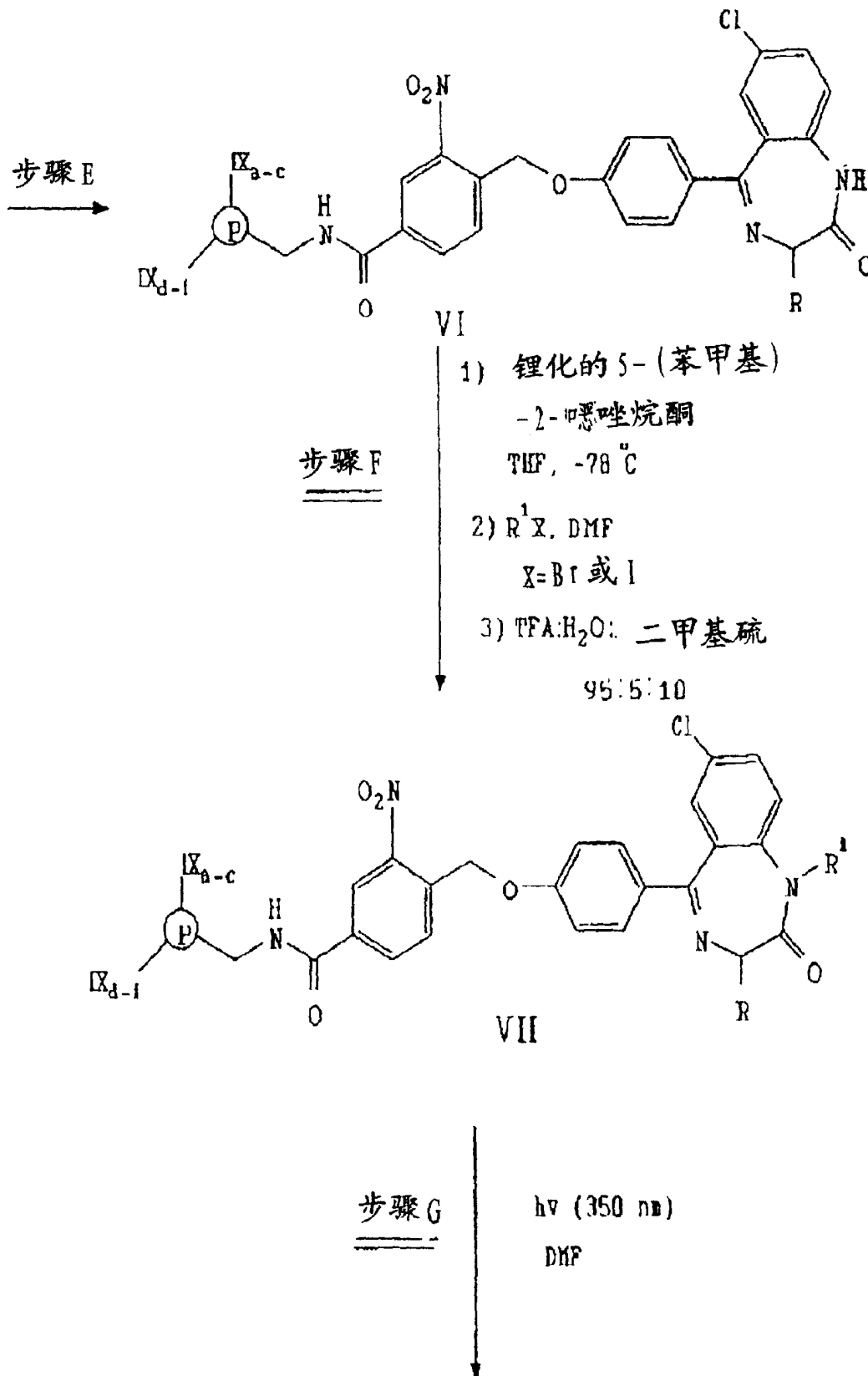


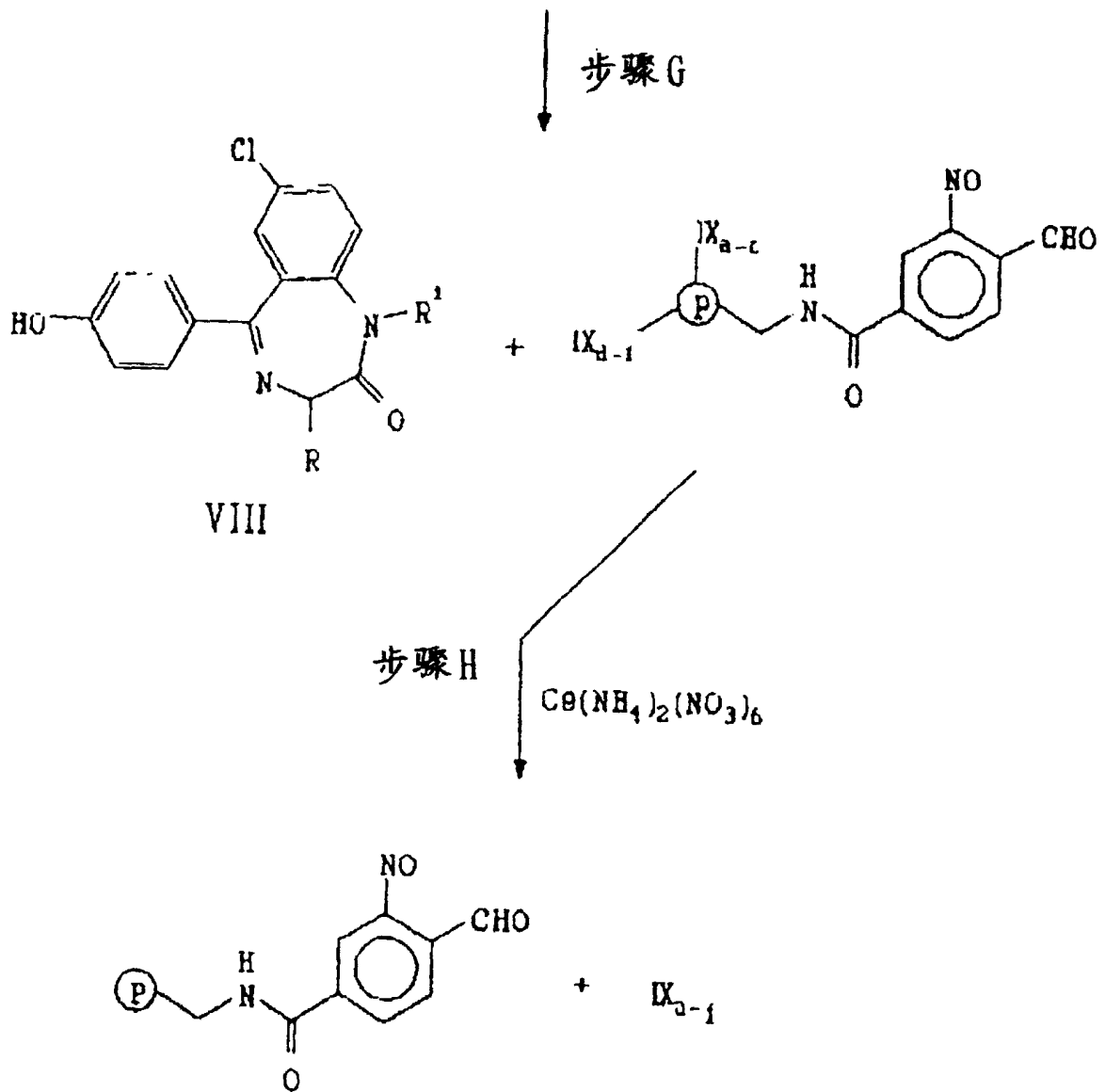




Ⓟ = 聚苯乙烯树脂

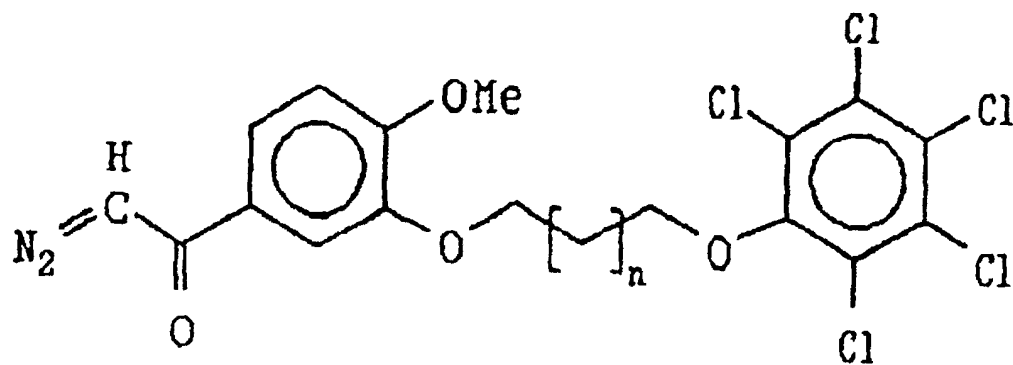






按照类似于 Bunin 和 Ellman (JACS, 114, 10997-10998 [1992]) 的方法, 在聚苯乙烯小球上制备苯并二吡啶因 VIII, 但是要在小球和苯并二吡啶因之间引入易光解的连接体 (见步骤 A、B 和 C), 这样就可以在步骤 G 不通过水解而是用紫外光照射 (在 DMF 中, 350nm, 10 分钟-12 小时) 除去苯并二吡啶因。而且, 在步骤 D 和 E 引入二进制代码, 从而精确限定用于引入每个 6R'S 和 5R''S 的反应顺序。在根据步骤 H 除去标示物并接着 GC 分离用电子捕获检测分析后, 就可确定单个 R 和 R' 基团的性质。

步骤 D、E 和 F 基本上按照 Bunin 和 Ellman 的方法进行, 但还包括在步骤 D 引入识别剂 IX<sub>a</sub>-C 和在步骤 E 引入 IX<sub>d</sub>-f。识别剂者用结构式 IX 代表:



IX

其中:

IX <sub>a</sub>	表示	n=6;
IX <sub>b</sub>	表示	n=5;
IX <sub>c</sub>	表示	n=4;
IX <sub>d</sub>	表示	n=3;
IX <sub>e</sub>	表示	n=2; 和
IX <sub>f</sub>	表示	n=1.

每个R和R'的代码如下:

表 2-1

IX	R
a	CH <sub>3</sub>
b	CH(CH <sub>3</sub> ) <sub>2</sub>
a, b	CH <sub>2</sub> CO <sub>2</sub> H
c	(CH <sub>2</sub> ) <sub>4</sub> NH <sub>2</sub>
a, c	CH <sub>2</sub> -C <sub>6</sub> H <sub>4</sub> -4-OH
b, c	CH <sub>2</sub> C <sub>6</sub> H <sub>5</sub>
IX	R'
d	H
e	CH <sub>3</sub>
d, e	C <sub>2</sub> H <sub>5</sub>
f	CH <sub>2</sub> CH=CH <sub>2</sub>
d, f	CH <sub>2</sub> C <sub>6</sub> H <sub>5</sub>

步骤 A

向I (1当量) 的甲苯溶液 (浓度=0.5M) 中加入Fmoc保护的2-氨基-5-氯-4'-羟基二苯酮 (1.3当量) 和氮杂二羧酸二乙酯 (1.3当量) 以及三苯基膦 (1.3当量)。混合物在室温下搅拌24小时。减压除

去溶剂, 剩余物用乙醚研制并过滤, 再减压除去溶剂。得到的产物 II 用硅胶色谱纯化。

### 步骤 B

---

向在室温下搅拌的 II 在 DCM 的溶液 (0.2M) 中加入 TFA (3 当量), 并搅拌溶液 12 小时。减压将溶液蒸发至干, 剩余物溶于 DCM, 用盐水洗涤一次, 并干燥 ( $\text{Na}_2\text{SO}_4$ )。过滤和蒸发溶剂得到 III。

### 步骤 C

---

在肽反应容器 (Merrifield 容器) 中, 将带有氨基官能团 (1.1mEq/g) 的 1%DVB (二乙烯基苯) 交联的聚苯乙烯小球 (50 $\mu$ ) 悬浮在 DMF 中。加入在 DMF 中的 III (2 当量) 和 HOBt (3 当量), 摇动容器 10 分钟。加入 DIC (3 当量), 摇动容器, 直至阴性水合茚三酮实验表明 12 小时后反应进行完全。

除去 DMF, 树脂在真空干燥前另用 DMF (X5) 和 DCM (X5) 洗涤。

### 步骤 D

---

将干燥树脂分装在 6 个反应容器中, 并用 DCM 悬浮。将适宜组合的识别剂 IX<sub>a-</sub>。(见表 2-1) 加到烧瓶中, 搅动 1 小时。Rh (TFA)<sub>2</sub> 催化剂 (1mol%) 加到每个烧瓶中, 再摇动 2 小时。将烧瓶中液体排干, 用 DCM (X5) 洗涤树脂。接着用 TFA 的 DCM 溶液 (0.01M) 处理树脂, 摇动 30



分钟,再用DCM(X3)接着用DMF(X2)洗涤。用20%的吡啶DMF溶液处理树脂,摇动30分钟,然后用DMF(X3)和DCM(X3)洗涤。

向每只烧瓶中加入适宜的Fmoc保护的氨基酰基氟(3当量)(当需要时,侧链官能基团可以叔丁基酯(Asp)、叔丁基醚(Tyr)或叔丁氧基羰基(Lys)形式保护)和2,6-二叔丁基-4-甲基吡啶(10当量),烧瓶搅动夜或直至得到阴性的水合茚三酮实验结果。树脂用DCM洗涤一次,然后将六份混合,真空干燥前再次洗涤(DCM×5)。

### 步骤E

---

干燥树脂分装在5个反应容器中,用DCM悬浮。将适宜组合的识别剂IX<sub>d-r</sub>(见表2-1)加到烧瓶中,摇动1小时。向每个烧瓶中加入Rh(TFA)<sub>2</sub>催化剂(1mol%),再摇动2小时。将烧瓶中的液体沥干,树脂用DCM(X5)洗涤。接着树脂用TFA的DCM溶液(0.01M)处理,摇动30分钟,然后用DMF(X3)和DCM(X3)洗涤。

向每只烧瓶中加入5%的乙酸DMF溶液,混合物加热至60℃并摇动过夜。沥干溶剂,然后树脂用DMF(X5)洗涤。

### 步骤F

---

将每批树脂都悬浮在THF中。烧瓶冷却至-78℃。向每只烧瓶加入锂化5-(苯甲基)-2-噁唑烷酮(2当量)的THF溶液,混合物在-78℃下摇动1小时。再向每只反应瓶加入适宜的烷基化试剂(表2-2)(4当量),接着加入催化量的DMF。将容器温热至室温,并在此温度下

摇动5小时。过滤除去溶剂，树脂用THF (X1) 洗涤，再真空干燥。将每份树脂合并，用THF (X2) 和DCM (X2) 洗涤，合并后树脂再用TFA: 水: 二甲基硫的95: 5: 10混合物处理2小时，从而除去侧链保护基。

表 2-2

识别剂	烷基化试剂
e	$H_3Cl$
d, e	$C_2H_5Br$
f	$BrCH_2-CH=CH_2$
d, f	$BrCH_2C_6H_3$

### 步骤G

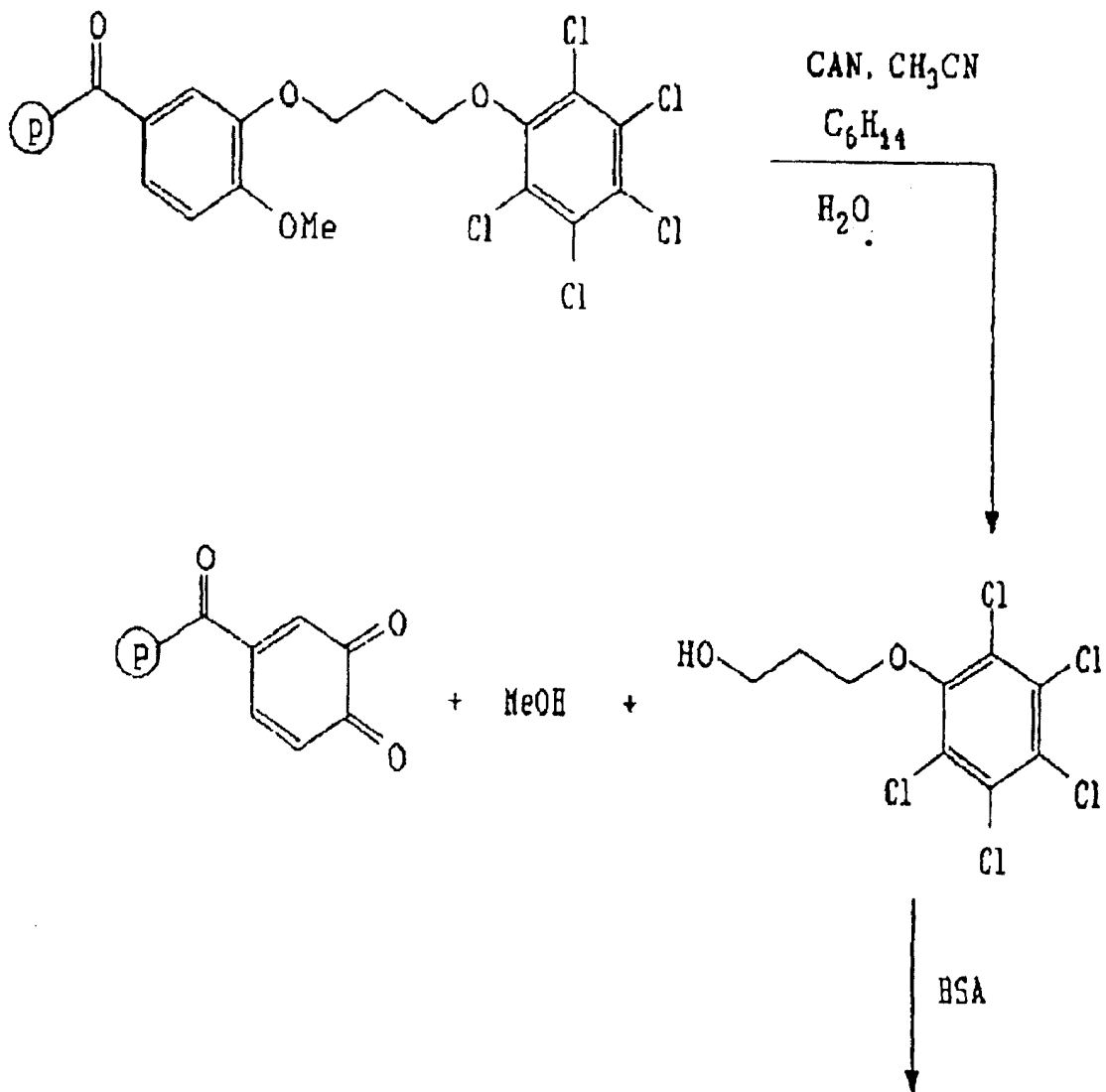
通过将小球悬浮在DMF中并用U. V. (350nm) 照射12小时，将得到的苯并二吡啶因从聚苯乙烯小球上剥离下来。

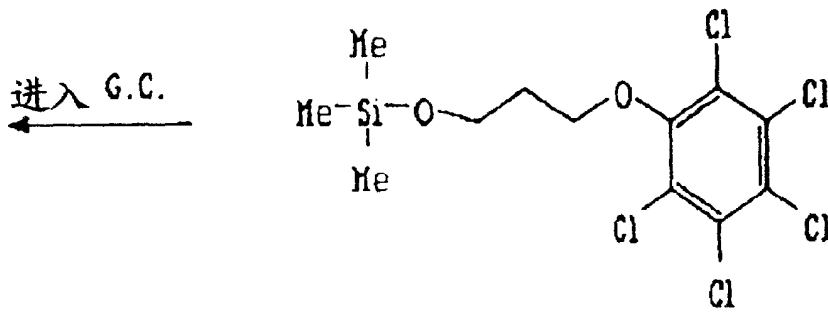
### 步骤H

将感兴趣的小球装入玻璃毛细管中，用注射器向毛细管注入1  $\mu$ l 1M硝酸铵铈(IV) (CAN) 水溶液、1  $\mu$ l 乙腈和2  $\mu$ l 己烷。用火焰

密封毛细管，然后离心以保证小球被试剂浸没。将毛细管放在超声器中，用声波处理1-10小时，优选处理2-6小时。

使毛细管破损开口，取出 $\sim 1\mu\text{l}$ 上部己烷层与 $\sim 0.2\mu\text{l}$ 双(三甲基甲硅烷基)乙酰胺 (BSA) 混合，而后注入GC，并象下列路线示例的那样，用电子捕获检测确定每个标示物。





### 实施例 3

#### 23, 540, 625 混合酰胺库

通过制备包括肽和其他酰胺化合物的由 23, 540, 625 个成员组成的组合库来进一步实验编码技术。

在 5 个步骤中使用 15 种不同的试剂并在第六步中使用 31 种不同的试剂进行上述合成。使用 4 个识别剂为共有 15 种试剂的五个步骤的每个步骤编码，使用 5 个识别剂为共有 31 种试剂的最后步编码。因此制备了包括 25 个识别剂的识别组。使用了 2-硝基-4-羧基苄基、0-芳基取代的  $\omega$ -羟基烷基碳酸酯识别剂，其中标示物组分由 3-12 个碳原子的烷基部分组成，芳基部分是 (A) 五氯苯基，(B) 2, 4, 5-三

氯苯基。(C) 2, 4, 6-三氯苯基, 或 (D) 2, 6-二氯-4-氯苯基。使用适宜长度的烷基链和 A、B、C 或 D 制备一组为 25 个的标示物, 用 0.2mm × 25M 聚甲基硅氧烷 GC 柱分离。标示物 T1~T25 (其中 T1 代表保留时间最长的标示物, T25 代表保留时间最短的标示物) 的化学组成总结如下:

T1 10A	T6 10C	T11 7B	T16 5C	T21 2B
T2 9A	T7 9B	T12 7C	T17 4B	T22 2C
T3 8A	T8 9C	T13 6B	T18 4C	T23 1B
T4 7A	T9 8B	T14 6C	T19 3B	T24 1C
T5 10B	T10 8C	T15 5B	T20 3C	T25 2D

标示物以 NAr 表示, 其中 N 是亚甲基的数目减去 2, Ar 是芳基。因此, 标示物 2A 具有通过氧与五氯苯基相连的亚丁基。

在开始五个阶段中使用的 15 种试剂以及识别它们的代码表示如下, 其中 1 代表有标示物存在, 0 代表没有标示物存在。

试 剂	代 码
L-丝氨酸	(0001)
D-丝氨酸	(0010)
L-谷氨酸	(0011)
D-谷氨酸	(0100)
L-谷氨酰胺	(0101)
D-谷氨酰胺	(0110)
L-赖氨酸	(0111)
D-赖氨酸	(1000)
L-脯氨酸	(1001)
D-脯氨酸	(1010)
L-苯丙氨酸	(1011)
D-苯丙氨酸	(1100)
3-氨基苯甲酸	(1101)
4-氨基苯乙酸	(1110)
3,5-二氨基苯甲酸	(1111)

第六阶段中的31种试剂以及代表它们的代码表示如下:

试剂	代码
L-丝氨酸	(00001)
D-丝氨酸	(00010)
L-谷氨酸	(00011)
D-谷氨酸	(00100)
L-谷氨酰胺	(00101)
D-谷氨酰胺	(00110)
L-赖氨酸	(00111)
D-赖氨酸	(01000)
L-脯氨酸	(01001)
D-脯氨酸	(01010)
L-苯丙氨酸	(01011)
D-苯丙氨酸	(01100)
3-氨基苯甲酸	(01101)
4-氨基苯乙酸	(01110)
3,5-二氨基苯甲酸	(01111)
琥珀酸酐	(10000)
惕各酸	(10001)

2-吡嗪羧酸	(10010)
(±) 硫辛酸	(10011)
1-哌啶丙酸	(10100)
胡椒基酸	(10101)
6-甲基烟酸	(10110)
3-(2-噻吩基)丙烯酸	(10111)
甲基碘	(11000)
甲苯碘酰氯	(11001)
异氰酸对甲苯磺酰酯	(11010)
3-氟基苯甲酸	(11011)
邻苯二甲酸酐	(11100)
乙酸酐	(11101)
氯甲酸乙酯	(11110)
甲磺酰氯	(11111)

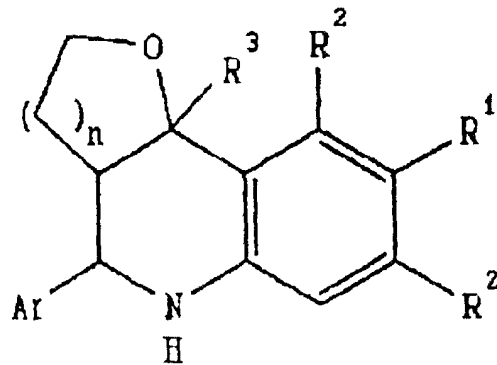


用普通方法在小球上制备含有六个甘氨酸单元的间隔基团。采用丁基侧链保护制备可变区域, 并且以 Fmoc 衍生物的形式保护氨基。用 DIC 和 HOBT 活化羧酸来形成酰胺键。

#### 实施例 4

#### 杂-Diels-Alder 库

组合的杂 Diels-Alder 库包括如下结构式的 42 个化合物:

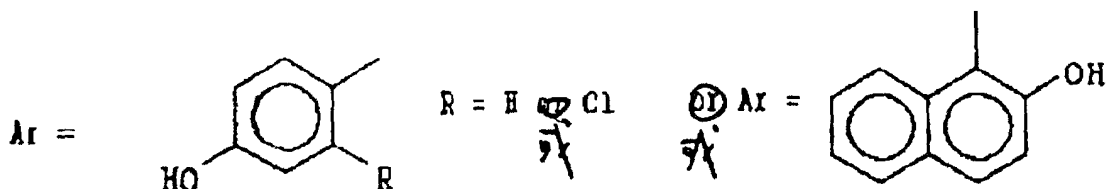


其中:

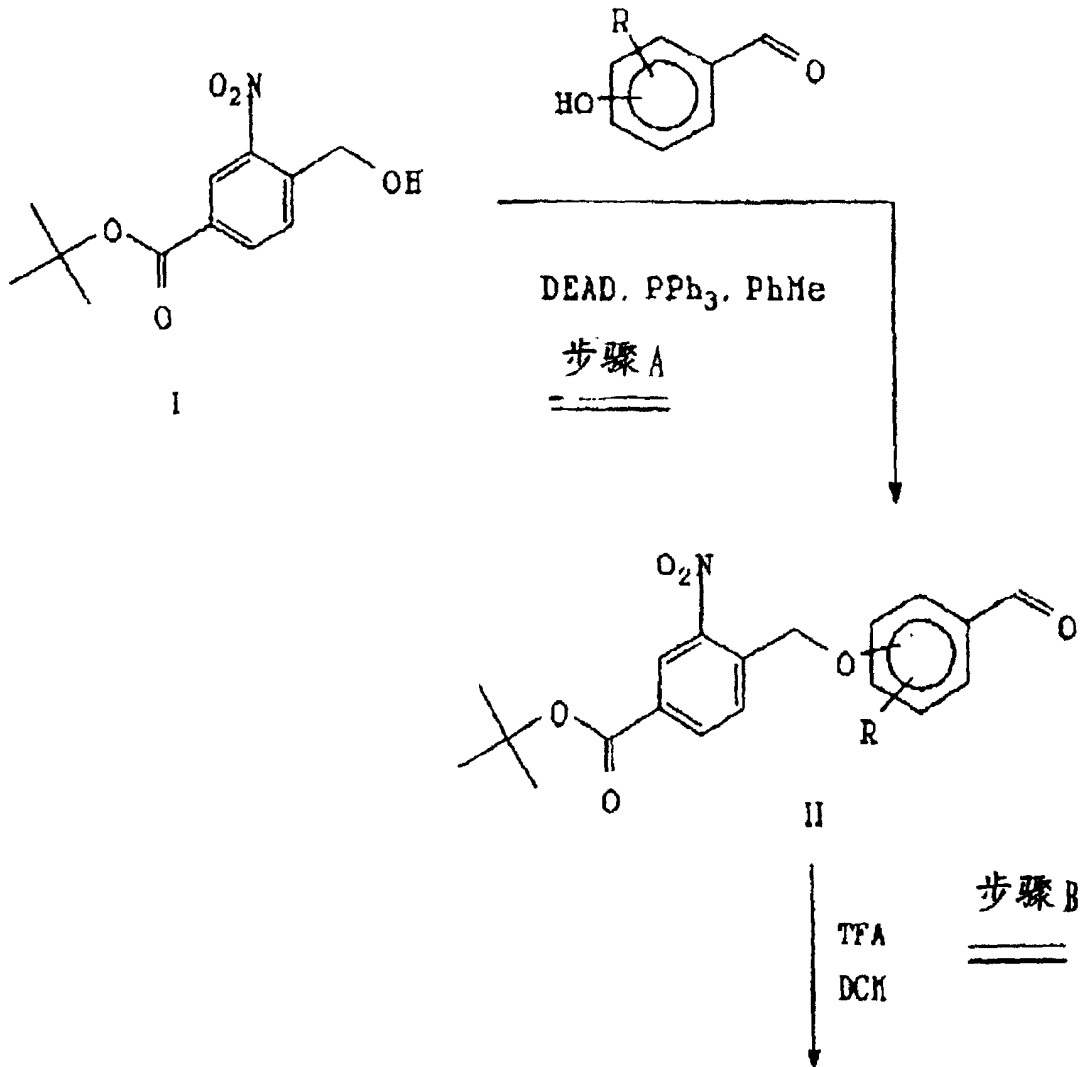
$R^1$  是 H,  $CH_3O$ ,  $F_3C$ ,  $F_3CO$ ,  $H_3C_6O$ , 或  $C_6H_{11}$ ;

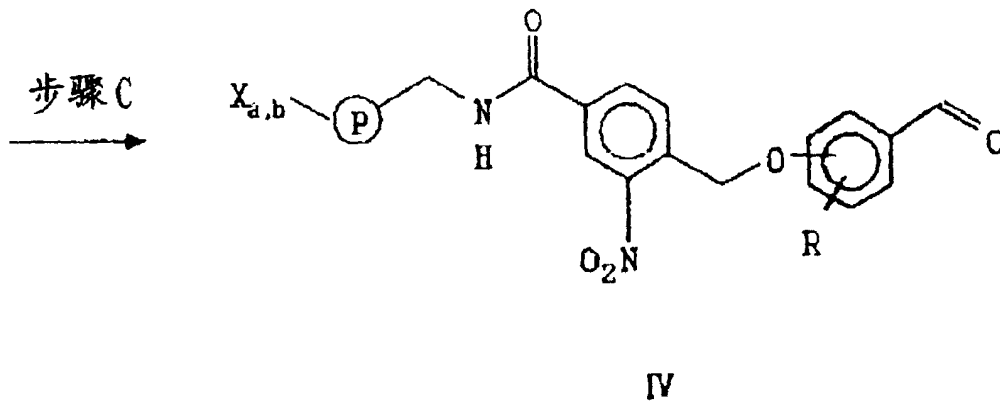
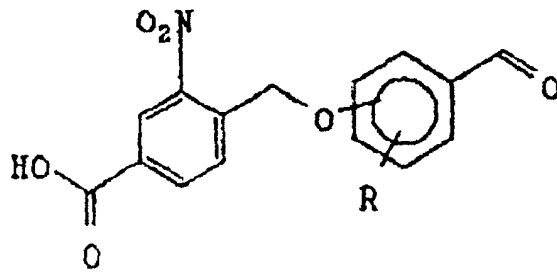
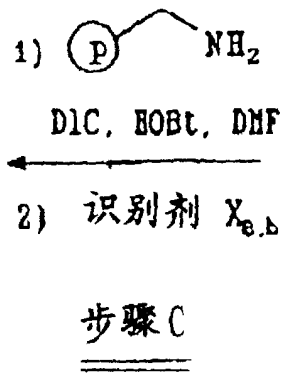
$R^2$  是 H,  $CH_3$ , 或  $CH_3O$ ;

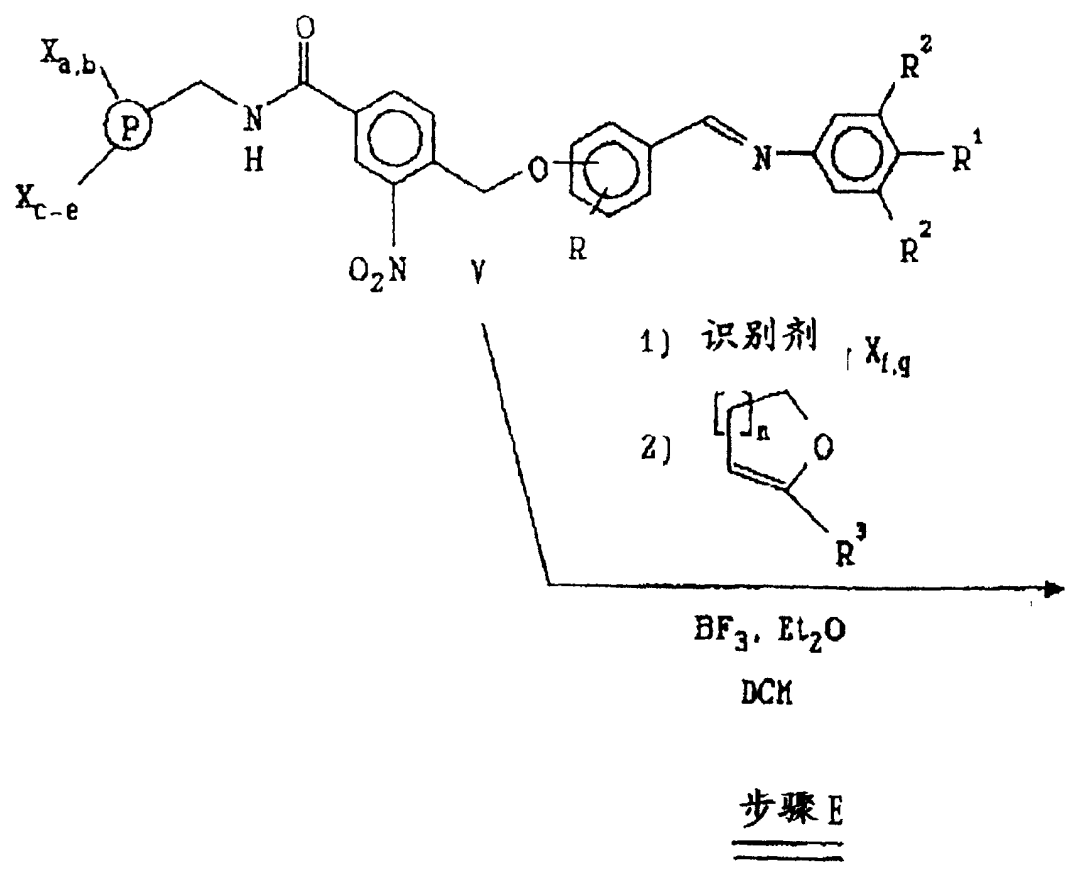
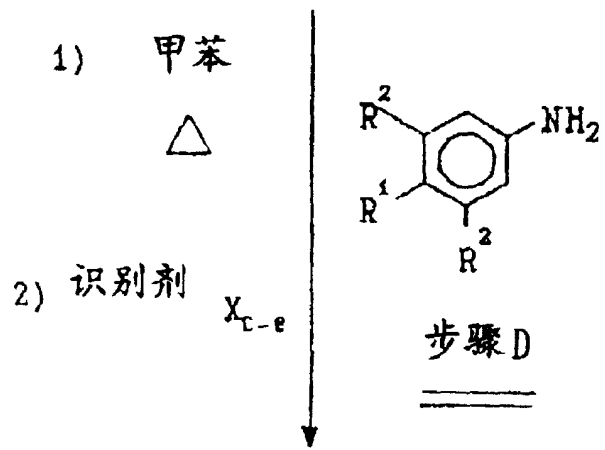
$R^3$  是 H (当  $n=2$  时), 或  $CH_3$  (当  $n=1$  时); 和

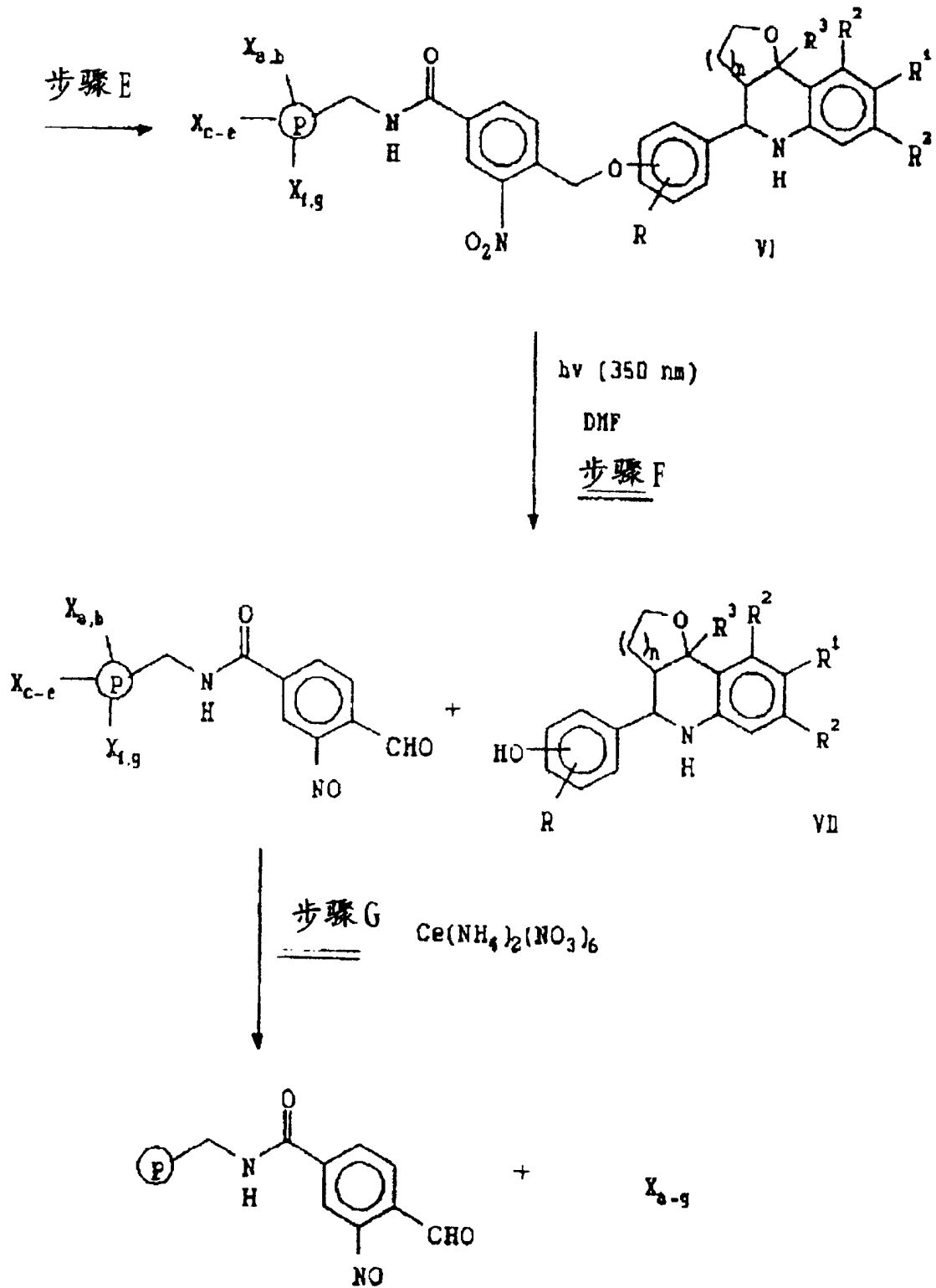


该化合物库按照下列合成路线制备：



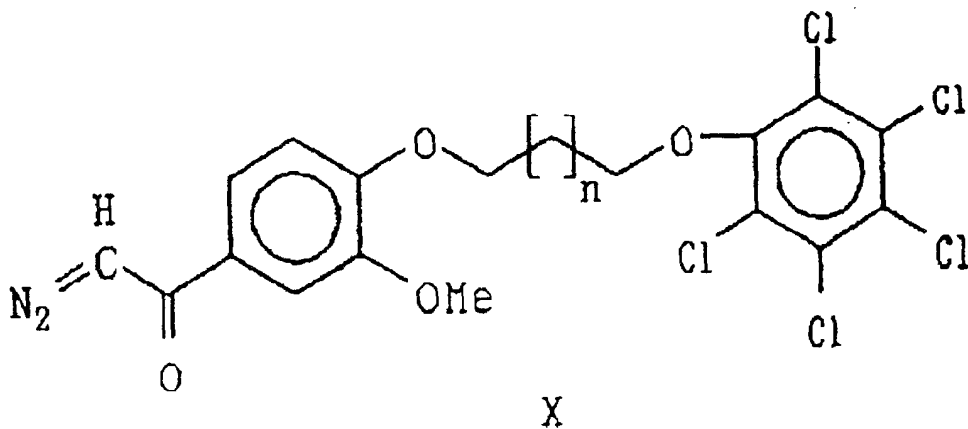






在聚苯乙烯小球上制备氮杂三环产物 (VI), 产物与小球之间通过光离解的连接体连接. 这样, 用紫外光 (350nm, 在DMF中) 照射即可从小球上除去氮杂三环 (VII). 在步骤C、D和E中引入的二进制代码独一无二地确定了用于引入ArR、R<sup>1</sup>、R<sup>2</sup>和R<sup>3</sup>的反应顺序. 按照步骤G除去编码的标示物, 并在GC分离后用电子捕获得检测对其进行分析.

在该合成路线中使用的识别剂用结构式X表示:



其中:

X <sub>a</sub>	表示	n=10
X <sub>b</sub>	表示	n=9
X <sub>c</sub>	表示	n=8
X <sub>d</sub>	表示	n=7
X <sub>e</sub>	表示	n=6
X <sub>f</sub>	表示	n=5
X <sub>g</sub>	表示	n=4

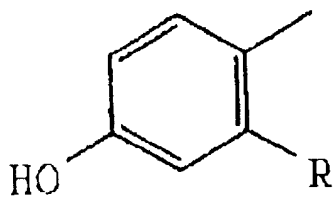
每个R、R<sup>1</sup>、R<sup>2</sup>、R<sup>3</sup>的代码如下:

表 4-1

X

a

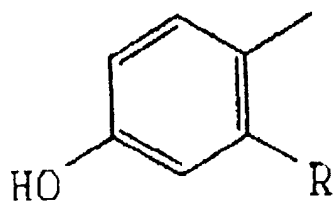
Ar =



R = H

b

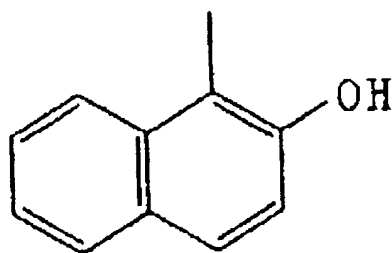
Ar =



R = Cl

a, b

Ar =



c	$R^1=H$	$R^2=H$
d	$R^1=H$	$R^2=CH_3$
d, c	$R^1=OCH_3$	$R^2=OCH_3$
e	$R^1=CF_3$	$R^2=H$
e, c	$R^1=C_6H_5O$	$R^2=H$
e, d	$R^1=F_3CO$	$R^2=H$
e, d, c	$R^1=C_6H_{11}$	$R^2=H$
f	$R^3=CH_3$	$n=1$
g	$R^3=H$	$n=2$

### 步骤A

在30分钟内, 向在 $0^\circ$ 下搅拌的 I (2.03g, 8mmol)、4-羟基苯甲醛 (1.17g, 9.6mmol) 和三苯基膦 (2.73g, 10.4mmol) 在甲苯 (20ml) 的溶液中加入偶氮二羧酸二乙酯。使溶液温热, 在达到室温后搅拌1小时。通过在减压下除去大约一半的溶剂浓缩溶液, 然后用乙醚研磨。接着过滤混合物, 剩余物用乙醚彻底洗涤。减压除去溶剂, 剩余物用硅胶色谱 (乙酸乙酯占己烷的15%) 纯化, 得到1.3g 醚 II<sub>a</sub> (产率47%)。



按类似的方式使2-氯-4-羟基苯甲醛和2-羟基-1-萘甲醛与 I 偶合, 分别以91%和67%的产率得到醚 II<sub>b</sub>和C。

### 步骤 B

---

向在室温下搅拌的醚 II<sub>a</sub> (0.407g, 1.14mmol) 的DCM (20ml) 溶液中加入TFA (8ml)。搅拌该溶液6小时。减压将溶液蒸发至干得到0.343g 酸 III<sub>a</sub> (产率100%)。类似地, 使醚 II<sub>b</sub>和 II<sub>c</sub>脱去保护基, 分别以产率92%和100%得到酸 III<sub>b</sub>和C。

### 步骤 C

---

向肽反应容器 (Merrifield 容器) 中称量加入带有氨甲基官能团 (1.1meq/g) 的1%DVB (二乙烯基苯) 交联的聚苯乙烯小球 (50-80 $\mu$ ) (200mg 树脂)。用DMF (2ml) 悬浮树脂, 并搅动20分钟。加入酸 III<sub>a</sub> (38mg, 2当量)、1-羟基苯并三唑 (40mg, 2当量) 和二异丙基碳化二亚胺 (38mg, 2当量), 搅动混合物直至获得阴性的水合茚三酮实验结果 (22小时)。过滤除去溶液, 并用DCM (8 $\times$ 10ml) 洗涤树脂。

将树脂再次悬浮到DCM (5ml) 中, 加入识别剂 X<sub>a</sub> (15mg), 摇动烧瓶1小时。加入Rh (TFA)<sub>2</sub> 催化剂 (1mol%), 摇动烧瓶2小时。过滤除去溶剂, 将树脂再悬浮到DCM (5ml) 中, 加入三氟乙酸 (1滴), 摇动容器20分钟。过滤除去溶剂, 并用DCM (8 $\times$ 10ml) 洗涤树脂。

以类似的方式使酸 III<sub>b</sub>和 III<sub>c</sub>连接到树脂上, 并用适宜的识别剂编码, 例如, 用 X<sub>b</sub>对酸 III<sub>b</sub>编码, 用 X<sub>a</sub>和 X<sub>b</sub>对酸 III<sub>c</sub>编码。将三批树脂

合并、混合、洗涤并干燥。

#### 步骤D

---

将干燥树脂分成相等的7份(每份87mg),分别投入7个肽反应容器(Merrifield容器)中,容器用加热带包封。在每个容器中,用甲苯(10ml)悬浮树脂,摇动20分钟。然后向每只烧瓶加入适量的苯胺(见表4-2)。

表 4-2

烧瓶	苯 胺	加入量
1	苯 胺	3 ml
2	3, 5-二甲基苯胺	3 ml
3	3, 4, 5-三甲氧基苯胺	2 g
4	4-三氟甲基苯胺	3 ml
5	4-苯氧基苯胺	2 g
6	4-三氟甲氧基苯胺	3 ml
7	4-环己基苯胺	2 g

接上加热带电源, 反应混合物在 70°C 摇动 18 小时。断开加热带电源, 过滤除去溶剂, 每份树脂都用干燥 DCM (4 × 10 ml)、乙醚 (10 ml)、甲苯 (10 ml) 和 DCM (2 × 10 ml) 洗涤。再用 DCM (5 ml) 分别悬浮每份树脂, 向每只烧瓶中加入适宜的识别剂或识别剂的组合 (X<sub>c-e</sub>) (15 mg) (见表 4-1)。摇动烧瓶 1 小时, 然后向每只烧瓶中加入 Rb (TFA)<sub>2</sub>

(1mol%)，并继续搅拌2小时。

除去溶剂，每份树脂都悬浮在DCM(5ml)中，并加入TFA(1滴)。混合物摇动20分钟；然后过滤除去溶剂。洗涤每份树脂(DCM, 1×10ml)并将其合并，再用DCM(3×10ml)洗涤，然后在真空下彻底干燥。

### 步骤E

---

干燥树脂分成相等的两份(每份0.3g)，每份都放入肽反应器中。每份树脂都用DCM(2×10ml)洗涤，然后悬浮在DCM(5ml)中。向一只烧瓶中加入识别剂 $X_f$ (15mg)，向另一只烧瓶中加入 $X_g$ (15mg)。摇动烧瓶1小时，然后加入Rh(TFA)<sub>2</sub>催化剂(1mol%)。摇动烧瓶2小时，然后过滤除去溶剂。每份树脂都用DCM(3×10ml)洗涤，然后每份树脂再悬浮在DCM(5ml)中。

向烧瓶中加入适宜的烯醇醚(1ml)(见表4-1)，搅动容器30分钟。向每只烧瓶中加入BF<sub>3</sub>·OEt<sub>2</sub>溶液(0.5ml在DCM中的5%溶液)，摇动烧瓶24小时。过滤除去溶剂，然后树脂用DCM(10ml)洗涤，再将树脂合并。小球再用DCM(5×10ml)、DMF(2×10ml)、甲醇(2×10ml)和DCM(2×10ml)进一步洗涤。接着树脂在真空下彻底干燥。

### 步骤F

---

为了确认在杂-Diels-Alder库中制备的产物，在投料量很大的基础上完成了一个实施例，从而可通过光谱手段确认产物的结构。采用的方法基本上与制备组合库时所述的方法相同。在步骤A, 4-

羟基苯甲醛与易光解基团相偶合。在步骤D, 苯胺与醛缩合。在步骤E, 用4, 5-二氢-2-甲基咪喃生成了烯醇醚。

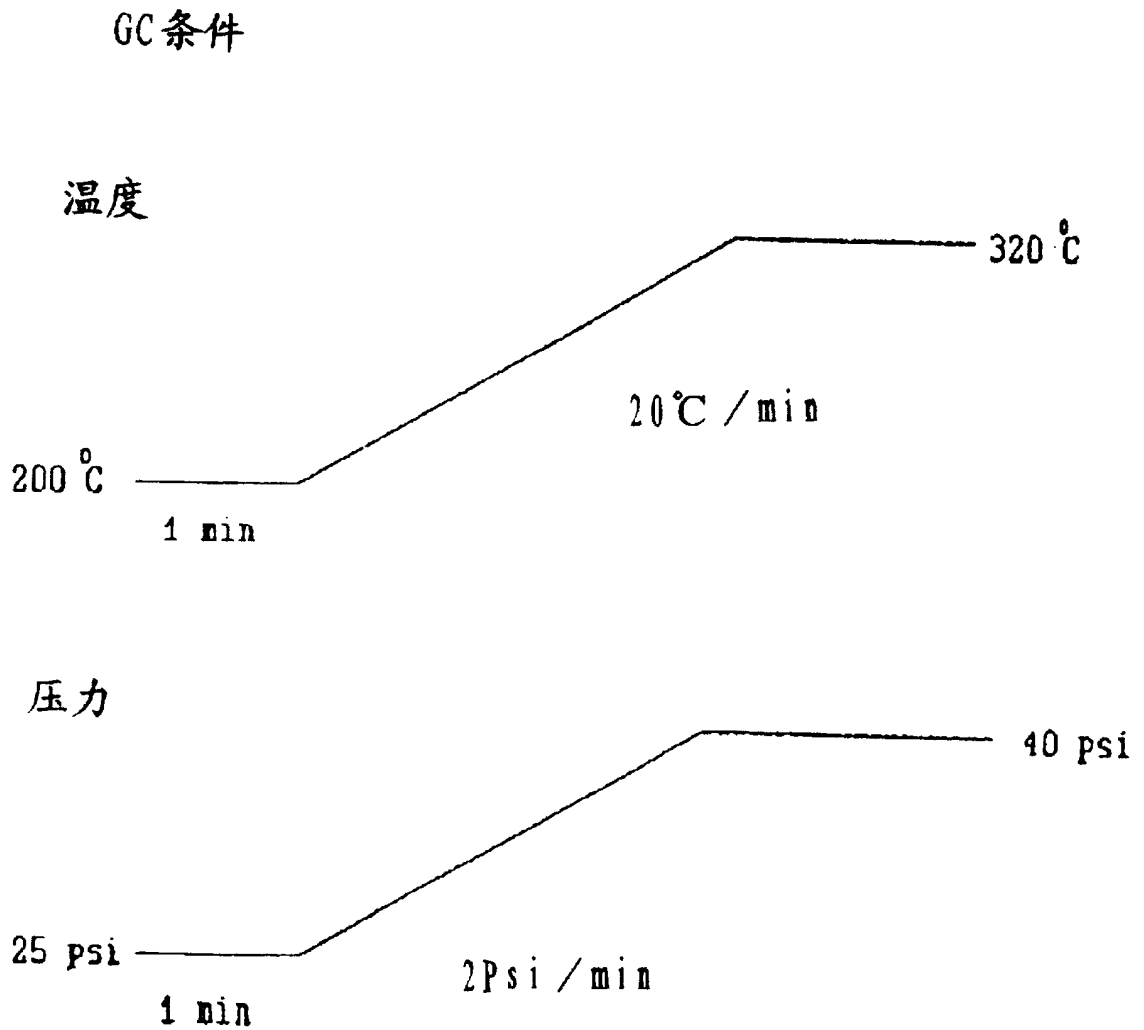
化合物的光解(步骤F)按照下述的方法进行: 将100mg小球悬浮在DMF(0.3ml)中, 用UVP“Black Ray”UVL56型手动366nm灯照射小球16小时。用吸液管将DMF移至一侧, 再用一些DMF(2×3ml)漂洗小球。合并原始溶液和洗涤溶液, 减压除去溶剂。通过与真实样品比较, NMR对反应混合物的分析表明, 反应混合物中含有需要的氮杂三环。

### 步骤G

将待实验小球装入一端封闭的派热克斯玻璃毛细管中, 用注射器注入1M硝酸铵铈(IV)水溶液和乙腈(1:1)(1 $\mu$ l), 离心毛细管, 这样小球停留在毛细管底部并被试剂溶液完全浸没。用注射器注入己烷(2 $\mu$ l), 毛细管再次离心。用火焰密封毛细管的开口端, 并将毛细管放在超声器中达4小时。然后, 将毛细管倒置在离心机中旋转, 这样迫使水层穿过己烷层至毛细管底部。该萃取方法重复3或4次, 然后使毛细管开封, 用注射器移去己烷层(1.5 $\mu$ l), 放入另一个含有BSA(0.2 $\mu$ l)的毛细管中。密封该毛细管并离心, 直至里面的试剂彻底混合。为了分离和解释标示物分子, 取出一部分溶液(约1 $\mu$ l)注入带有电子捕获检测器的气相色谱仪, 气相色谱仪的色谱柱是25M×0.2mm聚甲基硅氧烷熔成的硅胶柱。

在200℃和25psi的载气(H<sub>2</sub>)压力下, 将样品注入GC柱。1分钟后, 以20℃/分的速率升温至320℃, 并以2psi/分的速率升压至

40 psi。这些条件示于下图中：



由4个随机选择的小球得到了下列结果：

## 小球1

	检出的标示物		
	Xf	Xe Xd Xc	Xb Xa
Ar			2-羟基萘基
R <sup>1</sup>		C <sub>6</sub> H <sub>11</sub>	
R <sup>2</sup>		H	
R <sup>3</sup>	CH <sub>3</sub> (n=1)		

## 小球2

	检出的标示物		
	Xg	Xe Xd Xc	Xb
Ar			2-氯-4-羟基苯基
R <sup>1</sup>		C <sub>6</sub> H <sub>11</sub>	
R <sup>2</sup>		H	
R <sup>3</sup>	H (n=2)		



## 小球3

	检出的标示物		
	Xg	Xe Xd	Xb Xa
Ar			2-羟基萘基
R <sup>1</sup>		F <sub>3</sub> CO	
R <sup>2</sup>		H	
R <sup>3</sup>	H (n=2)		

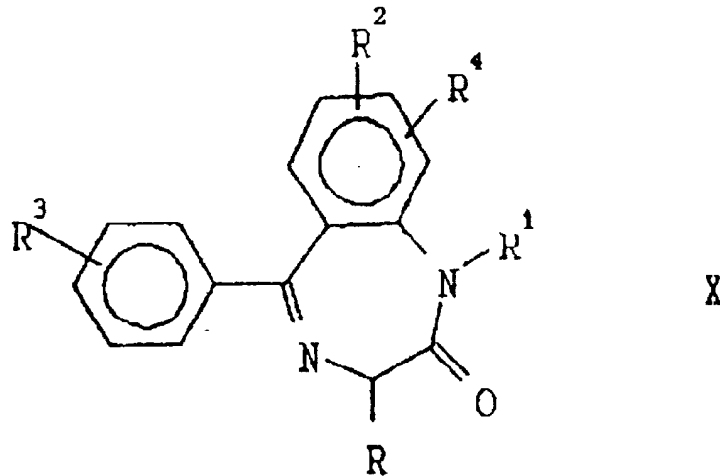
## 小球4

检出的标示物			
	Xf	Xc Xd	Xb
Ar			2-氯-4-羟基苯基
R <sup>1</sup>		F <sub>3</sub> CO	
R <sup>2</sup>		H	
R <sup>3</sup>	CH <sub>3</sub> (n=1)		

## 实施例5

## 苯并二吡啶因库

按照实施例2的方法，制备了结构式X的组合库，



其中

R是天然存在的D或L氨基酸基团；

R¹是H、C<sub>1</sub>-C<sub>6</sub>烷基、低级链烯基、C<sub>1</sub>-C<sub>6</sub>烷基氨基、羧基C<sub>1</sub>-C<sub>6</sub>烷基或苯基C<sub>1</sub>-C<sub>6</sub>烷基，其中苯基可被低级烷基、F、Cl、Br、OH、NH<sub>2</sub>、CO<sub>2</sub>H或O-低级烷基任意取代；

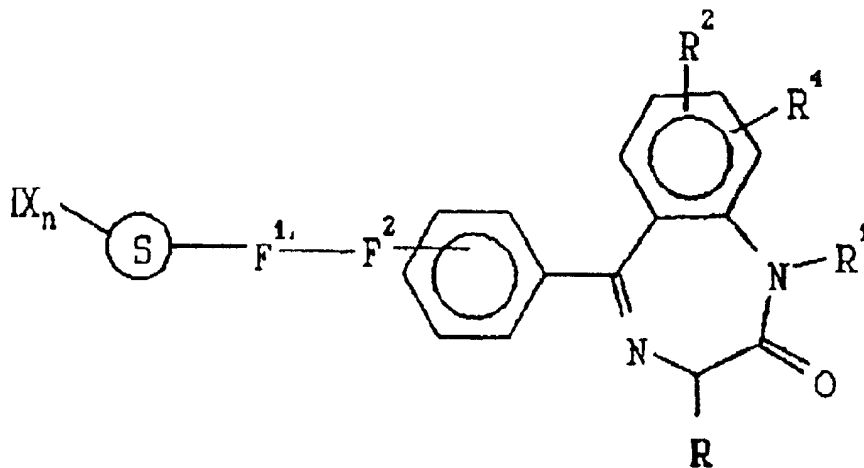
R²是H或CO<sub>2</sub>H；

R³是H或OH；

R⁴是H或Cl；

但须，当R³是OH时，R²是H，并且当R²是羧基时，R³是H。

该化合物库由许多结构式如下的编码小球释放得到，



其中

$IX_n$  是许多结构式  $I_a$  的识别剂，其中所述的“许多”代表编码的合成路线；

$S$  是基质；

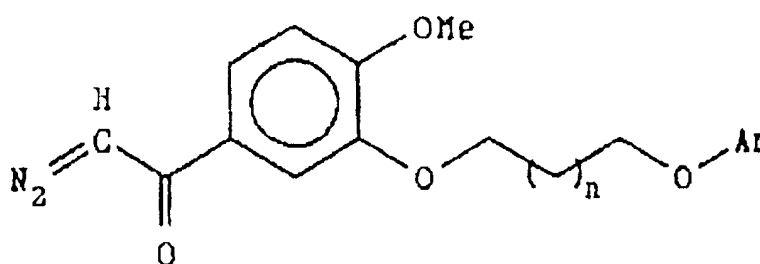
$F^{1'}-F^2$  是结构式  $I_a$  的连接体残基；以及  $R$ 、 $R^1$ 、 $R^2$  和  $R^4$  同结构式  $X$  中的定义。

## 实施例 6

### 典型识别剂的制备

通过形成卡宾连到树脂上的重氮化合物识别剂的制备示例如下。

下列通式的化合物制备如下。



其中

n 是 0-10, 以及

Ar 是五氯苯酚、2, 4, 6-三氯苯酚、2, 4, 5-三氯苯酚或 2, 6-二氯-4-氟苯酚。

向 1-羟基-4-(2, 6-二氯-4-氟苯氧基) 丁烷 (0.38g, 1.5mmol)、异香草酸甲酯 (0.228g, 1.5mmol) 和三苯基膦 (0.393g, 1.5mmol) 在 THF (8ml) 的溶液中加入偶氮二羧酸二乙酯 (0.287g, 1.7mmol)。在室温下搅拌上述溶液 36 小时。减压除去溶剂, 剩余物用硅胶色谱 (用 20% 乙酸乙酯和 80% 石油醚的混合物洗脱) 纯化, 得到 0.45g 醛 (产率为 77%)。

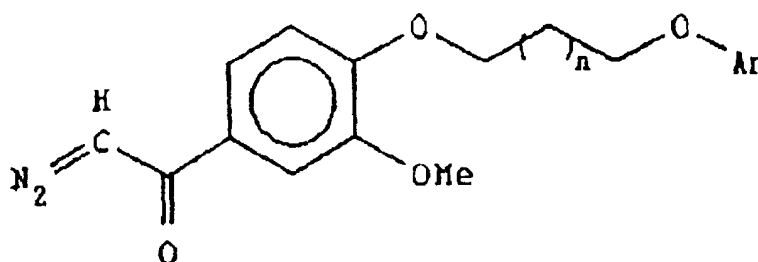
将上述醛 (100mg, 0.26mmol) 溶于丙酮 (8ml) 中, 并用  $\text{KMnO}_4$  (61mg, 0.39mmol) 在丙酮 (4ml) 和水 (4ml) 中的溶液处理。反应液在室温下搅拌 13 小时。反应混合物用乙酸乙酯 (100ml) 和水 (50ml) 稀释, 使有机层和水层分开。再用乙酸乙酯 (2×100ml) 萃取水层。合并有机层, 用水 (50ml) 洗涤并干燥 ( $\text{MgSO}_4$ )。除去溶剂得到 109mg 苯甲酸 (产率 93%)。

用草酰氯 (36mg, 0.28mmol) 和催化量的 DMF 处理上述酸 (76mg, 0.188mmol) 的二氯甲烷 (2ml) 溶液。在室温下搅拌 10 分钟后, 观察到缓慢而稳定的气体放出。继续搅拌 2 小时, 然后反应液用 DCM (15ml) 稀释, 并用饱和碳酸氢钠水溶液 (5ml) 洗涤。分层, 有机层用  $\text{Na}_2\text{SO}_4$  干燥, 挥发除去溶剂得到浅黄色晶体状的苯甲酰氯。

将苯甲酰氯溶于二氯甲烷 (5ml), 并在  $-78^\circ\text{C}$  下加到搅拌的重氮甲烷的乙醚溶液中。使冷浴温热, 反应混合物在室温下搅拌 5 小时。减压除去溶剂和过量的重氮甲烷, 剩余物用硅胶色谱纯化, 采用梯

度洗脱法洗脱, 其中乙酸乙酯在己烷中的浓度范围是10%-40%, 得到48mg重氮化合物(产率60%)。

下列通式的化合物按下述方法制备:



其中

$n$  是 0-10, 以及

Ar 是五氯苯酚、2, 4, 6-三氯苯酚、2, 4, 5-三氯苯酚或 2, 6-二氯-4-氯苯酚。

在氩气氛下, 将香草酸甲酯 (0.729g, 4.0mmol)、1-羟基-9-(2, 3, 4, 5, 6-五氯苯氧基) 壬烷 (1.634g, 4.0mmol) 和三苯基膦 (1.259g, 4.8mmol) 溶于 20ml 干燥甲苯中, 滴加 DEAD (0.76ml, 0.836g, 4.8mmol), 混合物在 25℃ 下搅拌 1 小时。反应液浓缩至一半, 用快速色谱纯化, 以 DCM 洗脱得到 1.0g (1.7mmol, 43%) 产物, 为白色结晶固体。

上述甲基酯 (1.0g, 1.7mmol) 溶于 50ml THF, 加入 2ml 水, 接着加入氢氧化锂 (1.2g, 50mmol)。混合物在 25℃ 下搅拌 1 小时, 再回流 5 小时。冷却至 25℃ 后, 将混合物倾入乙酸乙酯 (200ml) 中, 所得溶液先用 1M HCl (50ml × 3) 洗涤, 再用饱和 NaCl 水溶液 (1 × 50ml) 洗涤, 然后用硫酸钠干燥。除去溶剂, 粗产物酸用甲苯共沸一次。

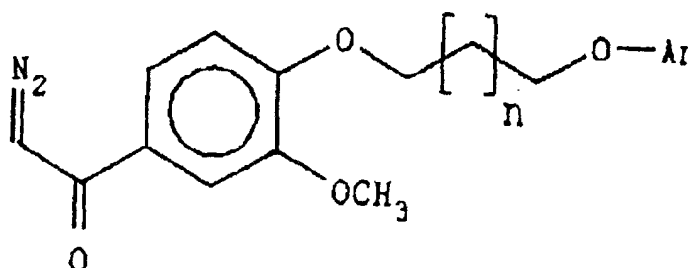
上述粗产物溶于 100ml 甲苯, 加入 10ml (1.63g, 14mmol) 亚硫酰氯, 混合物回流 90 分钟。通过蒸馏将溶液体积降至约 30ml, 然后蒸发除去剩余甲苯。在氩气下, 将粗酰基氯溶于 20ml 干燥 DCM, 并冷却至 -78℃, 加入约 10mmol 重氮甲烷在 50ml 无水乙醚中的溶液。混合物温热至室温, 搅拌 90 分钟。用氩气通过溶液鼓泡 10 分钟, 然后挥发除去溶剂, 粗产物用快速色谱纯化, 用 10-20% 乙酸乙酯 / 己烷洗脱, 得到浅黄色固体状的重氮酮 (0.85g, 1.4mmol, 三步的收率为 82%)。

按照上述方法制备了下列识别剂:

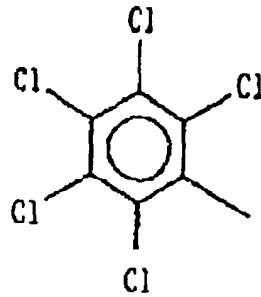
### 氧化裂解类型 I

制备了如下结构式的 7 个

如下



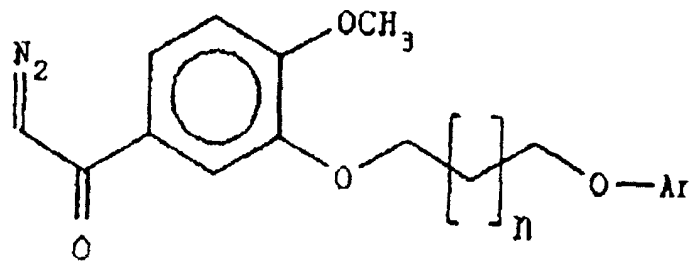
其中：  
Ar 是



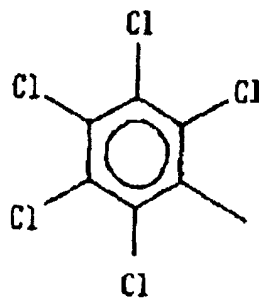
以及  $n$  is 4, 5, 6, 7, 8, 9, 和 10.

### 氧化裂解类型 II

制备了如下列结构式的 13 个识别剂：



其中：  
Ar 是

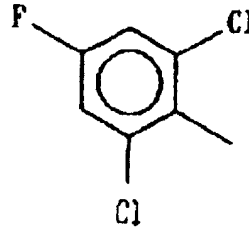




以及  $n$  is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10;

并且其中:

Ar 是



以及  $n$  是 0, 3, 和 9.

### 实施例 7

#### 用可被质谱识别的标示物

##### 编码组合库

标示物 4、11 和 13 (合成路线 7) 结构相同, 但由于不同数目的氘取代其分子量不同, 它们可分别被合成出来 (合成路线 8 和 9) 并可通过质谱 (MS) 分别加以分析。在 MS 技术中, 正化学电离质谱 (PCIMS) 可给出最小的标示物片段, 这样只有分子离子 ( $[M+NH_4]^+$ ) 和另一个片段, ( $[MH-H_2O]^+$ ) 可给出信号 (图 1、2 和 3)。这实际上就可通过观察两个信号来确定标示物的存在与否, 从而消除了分析较复杂结构时任何可能出现的混淆。随后混合了大约相等数量的三个标示物,

并用PCIMS分析(图5)。同样,可以容易地辨认出相应于每个单独标示物的两个信号。

现在将标示物4转变成重氮酮前体8(合成路线8),然后与Tentagel树脂相连得到9(合成路线11)。接着除去9的一个小球,并用硝酸高铈铵氧化释放出4。PCIMS分析再次清楚地显示出标示物4的存在。

总之,合成出了一组结构相同但分子量不同的标示物4、11和13。所有标示物可同时在混合物中被PCIMS容易地检测到。PCIMS可以检测到从单个Tentagel树脂小球(在组合库合成中使用)释放出的少量4。MS是有效而又灵敏的标示物检测方法,并可用作编码合成组合库的基本手段。

用PCIMS分析4、11和13的结果是通过使用1% $\text{NH}_3$  /  $\text{CH}_4$ 的试剂气体混合物获得的。

(2). 在 $0^\circ\text{C}$ 和Ar气氛下,向11.1ml (125mmol, 5.00当量) 1,4-丁二醇(1)、6.97ml (50.0mmol, 2.00当量)  $\text{Et}_3\text{N}$ 和0.153g (1.25mmol, 0.05当量) 4-二甲氨基吡啶在干燥 $\text{CH}_2\text{Cl}_2$  (100ml) 中的溶液加入3.88g (25.0mmol, 1.00当量) 97%叔丁基二甲基甲硅烷基氯。得到的溶液在 $0^\circ\text{C}$ 下搅拌15分钟,然后在 $25^\circ\text{C}$ 下搅拌16小时。接着反应液用 $\text{CH}_2\text{Cl}_2$  (250ml) 稀释,用1M  $\text{HCl}$  (100ml)、饱和 $\text{NaHCO}_3$ 水溶液 (100ml) 和水 (100ml) 洗涤,然后干燥( $\text{MgSO}_4$ )。减压除去挥发性物质得到油状的粗产物2。

(3). 在 $0^\circ\text{C}$ 和Ar气氛下,向~10.0mmol粗醇 2、1.93g (10.5mmol, 1.05当量) 五氟苯酚和2.89g (11.0mmol, 1.10当量) 三苯基膦在干燥 $\text{CH}_2\text{Cl}_2$  (40ml) 中的溶液加入1.73ml (11.0mmol, 1.10当量) 偶

氮二羧酸二乙酯。得到的桔黄色溶液在0℃下搅拌5分钟, 再在25℃下搅拌15小时。接着反应液用CH<sub>2</sub>Cl<sub>2</sub> (250ml) 稀释, 并用饱和Na<sub>2</sub>CO<sub>3</sub>水溶液 (100ml)、饱和NH<sub>4</sub>Cl水溶液 (100ml) 和水 (100ml) 洗涤, 然后干燥 (MgSO<sub>4</sub>)。减压除去挥发性物质并用快速色谱法纯化 (0-20% EtOAc / 己烷) 得到油状产物3。

(4). 在25℃下, 向1.85g (5.00mmol, 1.00当量) 甲硅烷基保护的醇3在THF (20ml) 中的溶液加入10.0ml (10.0mmol, 2.00当量) 1.0M氯化四丁基铵的THF溶液。得到的桔黄色溶液在25℃下搅拌4小时。减压除去挥发性物质并用快速色谱法纯化 (20-40% EtOAc / 己烷) 得到1.10g (86%) 油状产物4。

(5). 在0℃和Ar气氛下。向0.800g (3.125mmol, 1.00当量) 醇4、0.569g (3.125mmol, 1.00当量) 香草酸甲酯和0.984g (3.75mmol, 1.20当量) 三苯基膦在干燥CH<sub>2</sub>Cl<sub>2</sub> (20ml) 中的溶液加入0.591ml (3.75mmol, 1.20当量) 偶氮二羧酸二乙酯。得到的浅黄色溶液在0℃下搅拌5分钟, 再在25℃下搅拌19小时。接着反应液用CH<sub>2</sub>Cl<sub>2</sub> (100ml) 稀释, 并用1MNaOH (50ml)、饱和NH<sub>4</sub>Cl水溶液 (50ml) 和水 (50ml) 洗涤, 然后干燥 (MgSO<sub>4</sub>)。减压除去挥发性物质, 并用快速色谱法纯化 (20% EtOAc / 己烷) 得到油状产物5。

(6). 向3.125mmol酯5的THF (12ml) 溶液中加入1.31g (31.3mmol, 10.0当量) 氢氧化锂一水合物。向得到的悬浮液中加入MeOH (24ml), 使其形成溶液, 该溶液在25℃搅拌1小时, 再回流1天。减压除去挥发性物质, 加入1MHCl直至溶液的PH值达到~1。收集形成的白色沉淀产物, 干燥得到0.968g (76%-2步) 6。

(7). 在Ar气氛下向0.968g (2.38mmol, 1.00当量) 羧酸6中加入

2.43ml 亚硫酸氯。得到的悬浮液回流1.5小时,然后形成了黄色溶液。减压除去挥发性物质,得到的剩余物用甲苯共沸三次,得到无色晶体状的产物7。

(8). 在0℃和Ar气氛下,向2.38mmol 酰基氯7的1:1THF:MeCN (20ml) 溶液中加入1.16ml (8.33mmol, 3.50当量) Et<sub>3</sub>N, 然后加入3.57ml (7.14mmol, 3.00当量) 2.0M (三甲基甲硅烷基) 重氮甲烷的己烷溶液。得到的黄色溶液在0℃搅拌1小时,再在25℃搅拌1天。反应液用EtOAc (150ml) 稀释,并用饱和NaHCO<sub>3</sub>水溶液 (2×75ml) 和饱和NaCl水溶液 (2×75ml) 洗涤,然后干燥 (MgSO<sub>4</sub>)。减压除去挥发性物质得到油状粗产物8。

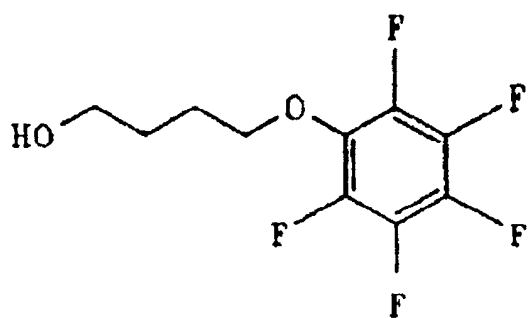
(11). 用市场上可购到的1,4-丁二醇-2,2,3,3-d<sub>4</sub> (10) 代替1,按照类似于将1转变为4时所述的方法,分三步以41%收率得到了11。

用市上可购到的1,4-丁二醇-2,2,3,3,4,4-d<sub>8</sub> (10) 代替1,按照类似于将1转变为4时所述方法,分三步以42%收率得到了13。

按照基本上与实施例4 (杂Diels-Alder库) 所给方法的相同方法,用5-50% (w/w vs. 树脂) 前体重氮酮8将标示物4引至固体载体上得到9。随后按照基本上与实施例4步骤G相同的方法从9中除去标示物4。

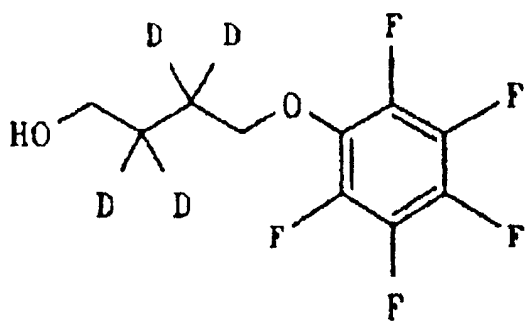
使用相应于标示物11和13的重氮酮将这些标示物引至固体载体上,这样与8一起,得到了二进制编码组的成员。

## 合成路线7



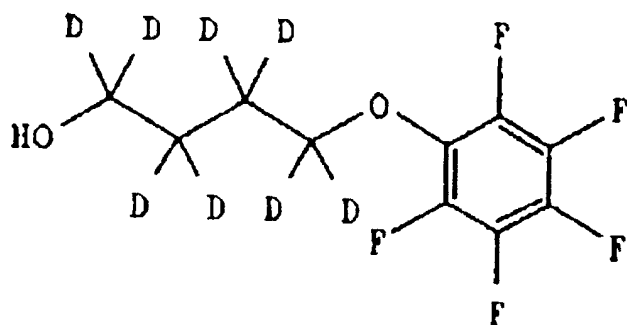
4

M.W. = 256



11

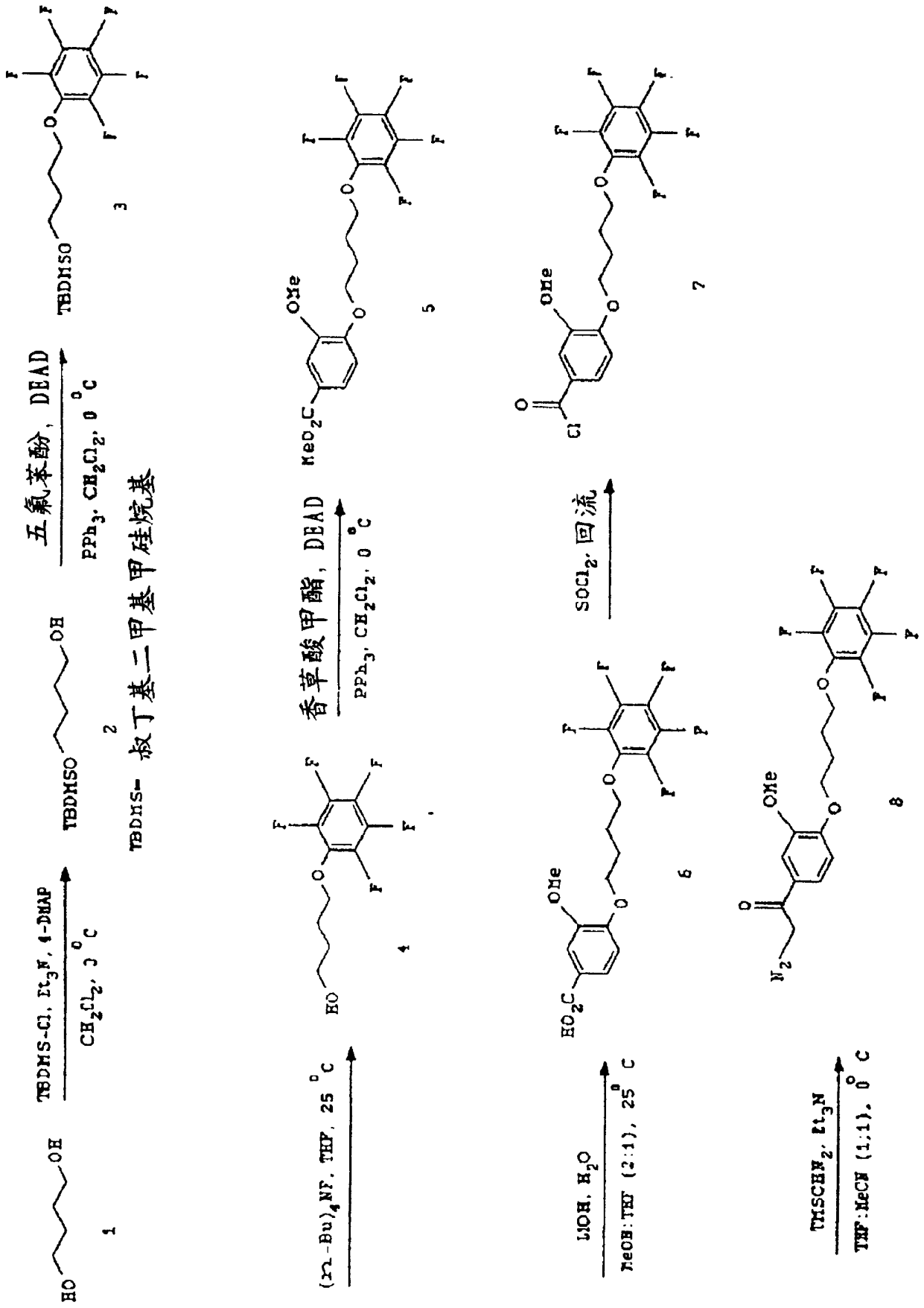
M.W. = 260



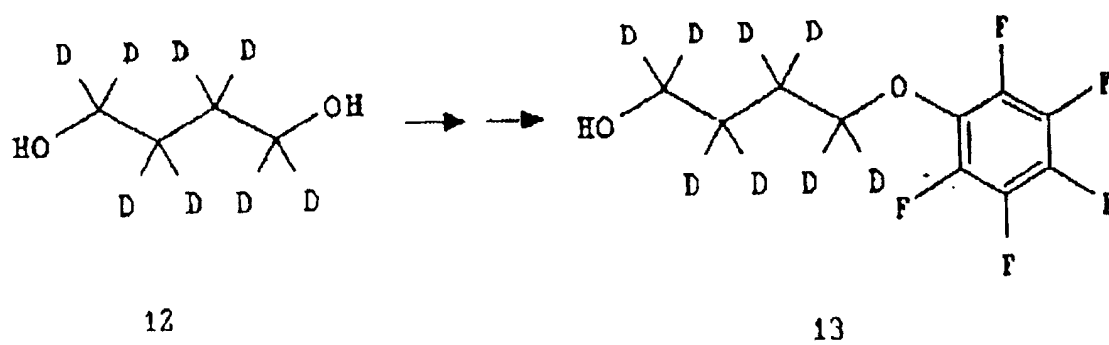
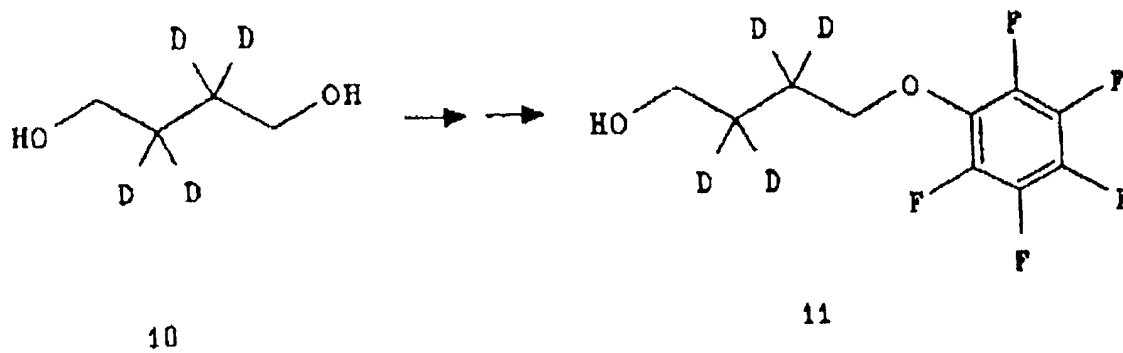
13

M. W. = 264

合成路线 8

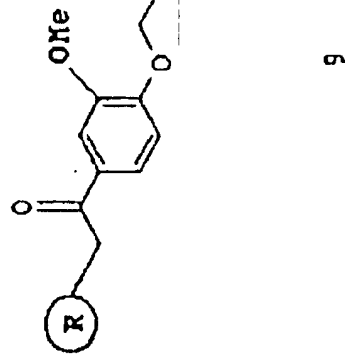
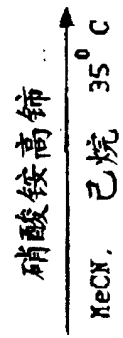
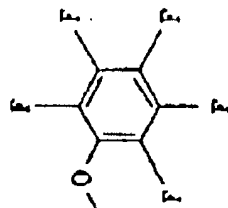
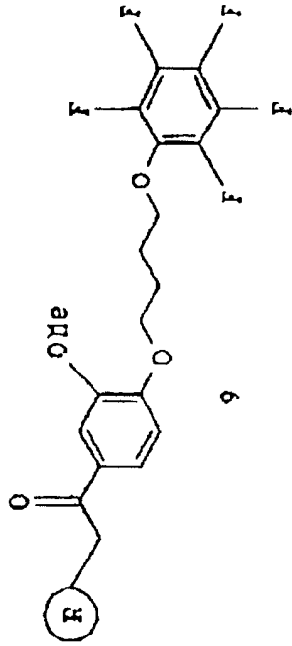
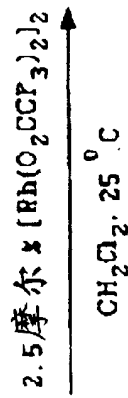
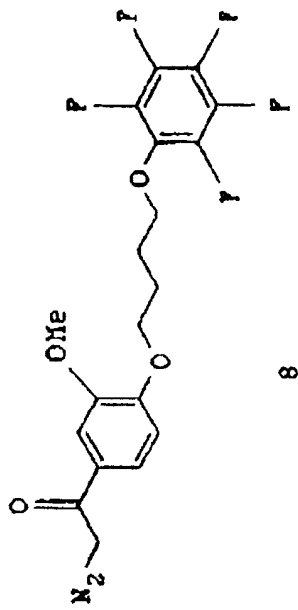


## 合成路线 9





合成路线 10



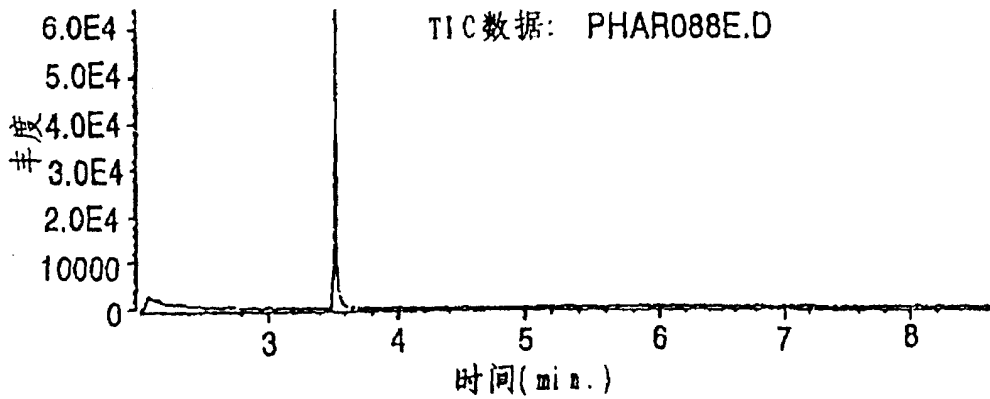
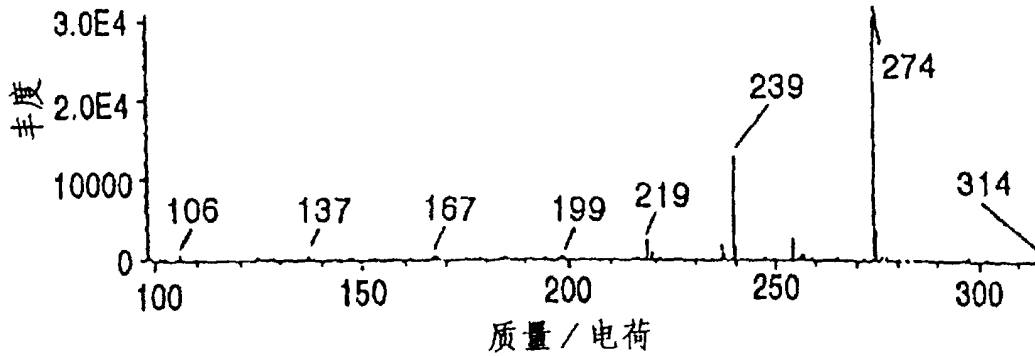
从上面的叙述可以清楚地看出,本发明提供了多用而又简单的识别化合物的方法,其中化合物的存在量不能保证其反应历程的精确测定。该方法可以生产很大数目的不同化合物,这些化合物可用于各种筛选技术中,从而确定生物活性或其他有意义的活性。在操作条件下呈化学惰性的标示物的使用,使得由各种合成技术(这些技术用于生产所需产物)产生的各种各样的环境呈现出很大的多样性。标示物可以很容易地合成并可进行精确的分析,从而可以精确地确定成分的性质。

本说明书中引用的所有出版物和专利申请都引作本文的参考文献,就好象是每份单独的出版物或专利申请都分别而又明确地指明作为本文的参考文献一样。

尽管为了清楚理解的目的,上述发明已通过图表和实施例给予了详细说明,但本领域的普通技术人员在受到本发明的启示后可以很容易地做出某些变化和改进,但这些变化和改进都不会背离所附权利要求的精神或范围。

图 1

扫描 131(3.513 min) 数据: PHAR088E.D

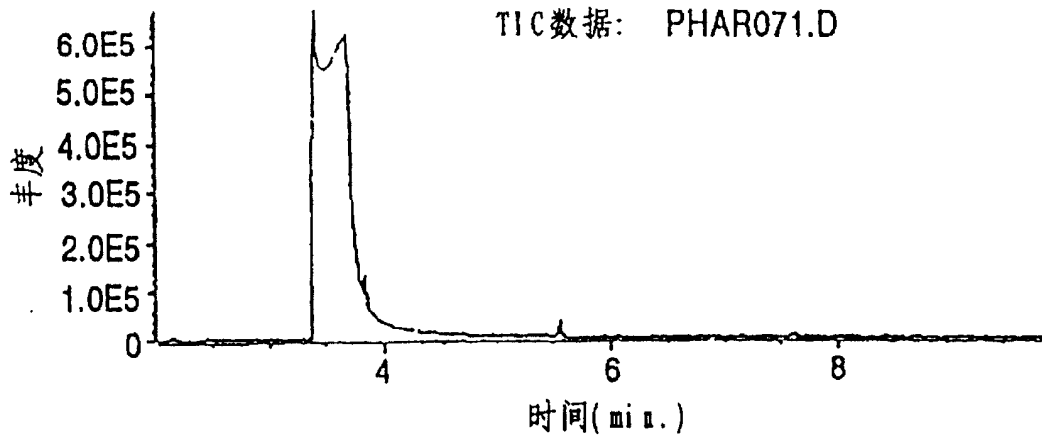
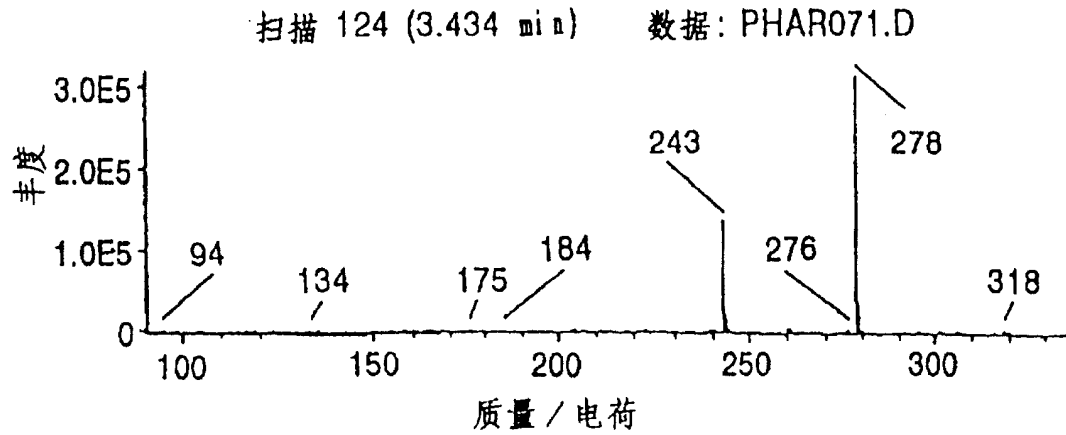


扫描 131(3.513 min) 数据 : PHAR088E.D

药典 GC/CIMS CH4/NH3(g) 90/350 A/D=1

m/z	丰度	m/z	丰度	m/z	丰度	m/z	丰度
203.30	2	237.05	17	255.20	12	272.05	11
216.20	4	239.05	411	255.45	12	274.05	1000
218.05	70	240.05	48	256.20	23	275.05	111
219.20	25	241.20	4	257.05	15	276.20	11
221.20	5	247.30	5	259.05	6	277.20	3
233.95	2	254.20	84	265.05	8	297.15	4
236.20	52						

图 2

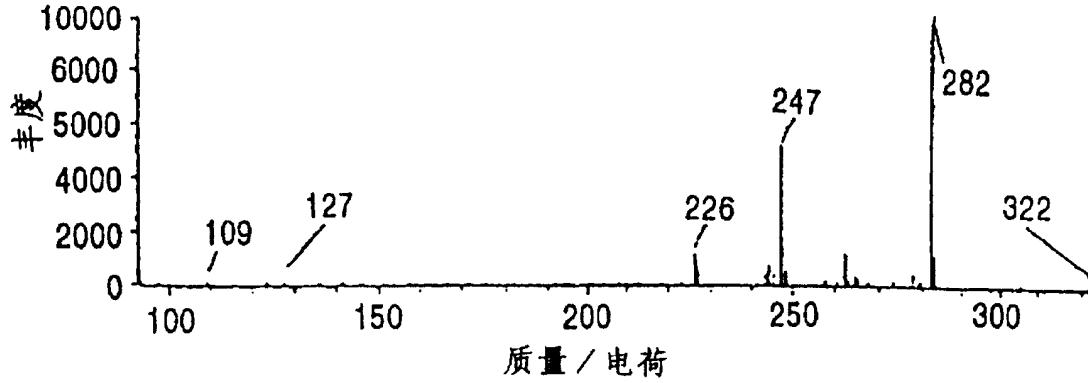


扫描 124 (3.434 min) 数据: PHAR071.D  
药典 GC/CIMS CH4/NH3(g) 90/350 A/D=1

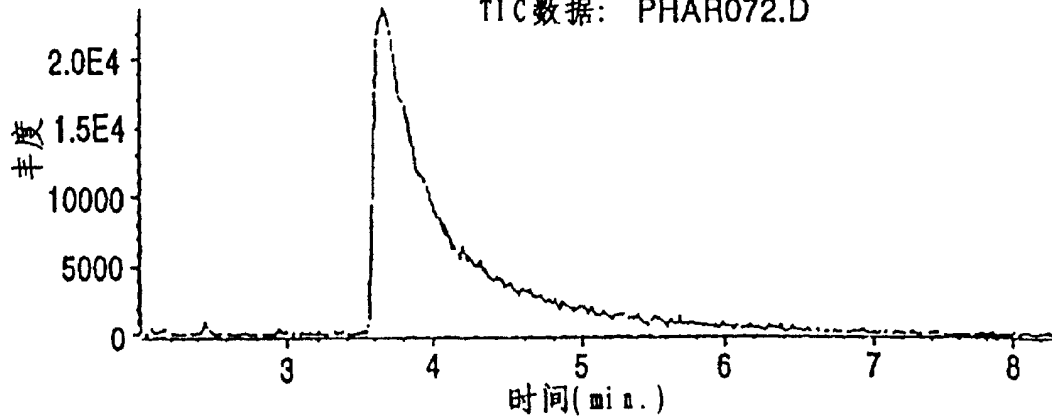
m/z	丰度	m/z	丰度	m/z	丰度	m/z	丰度
203.05	1	238.05	1	260.20	11	279.05	113
210.20	1	239.20	2	261.05	13	280.05	10
213.20	1	240.20	8	262.20	2	281.05	1
220.20	1	241.20	7	263.20	1	289.20	1
222.05	8	243.05	429	269.05	1	290.05	2
223.05	3	244.05	47	274.05	2	291.05	1
224.20	1	245.05	3	275.30	4	292.05	1
224.55	1	251.20	1	276.20	13	295.15	3
225.30	1	258.20	9	278.05	1000		

图 3

扫描 143 (3.648 min) 数据: PHAR072.D



TIC数据: PHAR072.D

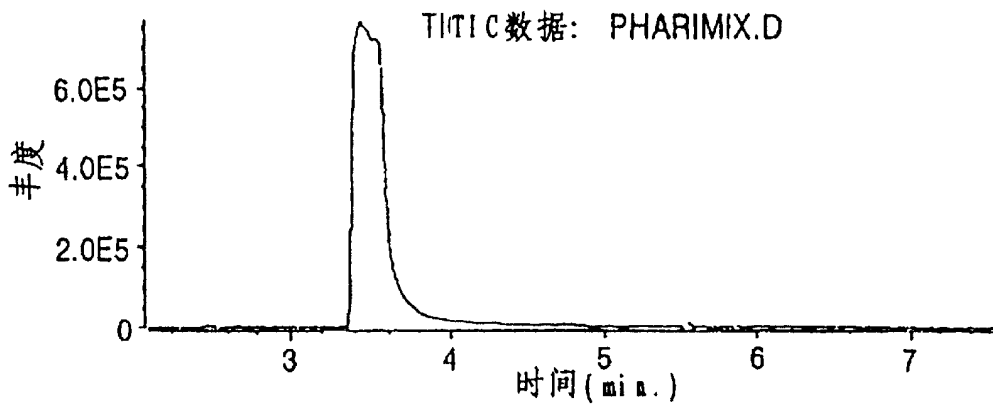
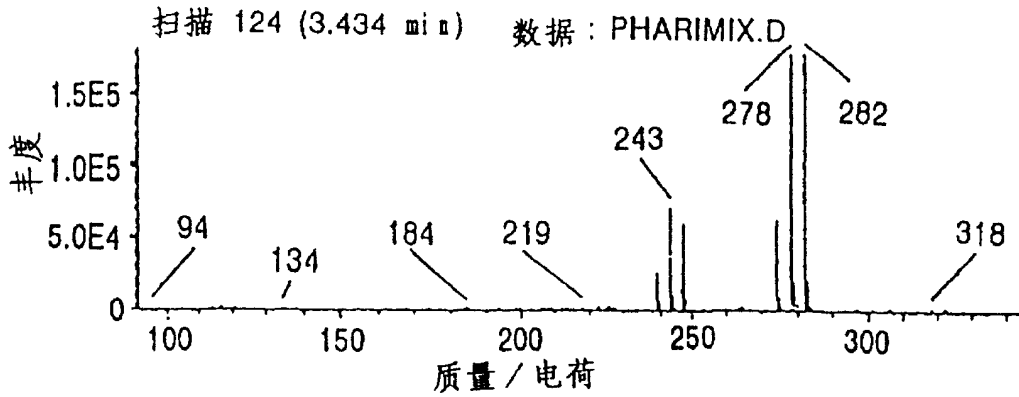


扫描 143 (3.648 min) 数据: PHAR072.D

药典 GC/CIMS CH4/NH3(g) 90/350 A/D=1

m/z	丰度.	m/z	丰度.	m/z	丰度.	m/z	丰度.
208.20	7	244.05	66	262.20	124	278.20	39
223.20	7	245.20	35	263.05	18	279.20	13
226.20	102	247.20	493	264.05	25	280.20	11
227.30	42	248.20	49	265.20	21	282.20	1000
241.20	9	257.95	12	267.20	7	283.20	115
242.95	34	260.55	13	273.20	12	284.20	9

图 4



扫描 124 (3.434 min) 数据: PHARIMIX.D  
GC/CIMS CH4/NH3(g) 90/350 A/D=1

m/z	丰度	m/z	丰度	m/z	丰度	m/z	丰度
203.05	1	237.20	2	261.05	13	283.20	112
206.95	1	239.05	178	262.05	9	284.05	11
210.30	1	240.05	27	263.20	3	285.05	1
213.05	1	241.05	8	264.20	12	285.95	1
213.95	1	243.05	409	265.20	13	288.30	1
218.05	4	244.05	49	266.05	2	290.05	1
219.20	2	245.05	7	267.20	1	291.20	2
220.05	1	247.05	402	268.30	1	292.05	2
222.05	9	248.05	44	269.05	1	293.40	1
223.05	3	249.20	4	274.05	432	293.75	1
224.20	2	251.20	1	275.05	51	294.40	2
225.20	2	254.05	5	276.05	18	295.25	3
226.20	9	256.20	6	278.05	996	296.25	2
227.05	3	257.05	7	279.05	112	297.25	2
228.20	1	258.20	12	280.20	18	298.15	1
229.20	1	259.20	3	282.20	1000	299.25	4
236.05	2	260.20	10				