



(12) 发明专利

(10) 授权公告号 CN 101107357 B

(45) 授权公告日 2013. 01. 02

(21) 申请号 200680002864. 0

(22) 申请日 2006. 01. 20

(30) 优先权数据

60/646, 584 2005. 01. 21 US

(85) PCT申请进入国家阶段日

2007. 07. 23

(86) PCT申请的申请数据

PCT/US2006/002420 2006. 01. 20

(87) PCT申请的公布数据

W02006/079061 EN 2006. 07. 27

(73) 专利权人 哈佛大学校长及研究员协会

地址 美国马萨诸塞州

(72) 发明人 戴维德·刘 樱井馨

(74) 专利代理机构 北京东方亿思知识产权代理

有限责任公司 11258

代理人 李剑

(51) Int. Cl.

C12N 15/10 (2006. 01)

C12Q 1/68 (2006. 01)

(56) 对比文件

W0 02103008 A2, 2002. 12. 27, 说明书第
53-54 页, 附图 43.

W0 2004016767 A2, 2004. 02. 26, 附图 59.

W0 2004013070 A2, 2004. 02. 12, 实施例 1,
附图 25-26.Xiaoyu Li, David R. Liu. DNA-Templated
Organic Synthesis: Nature's Strategy for
Controlling Chemical Reactivity Applied
to Synthetic Molecules. Angew. Chem. Int.
Ed. 43. 2004, 434849-4867.

审查员 李娟娟

权利要求书 3 页 说明书 25 页

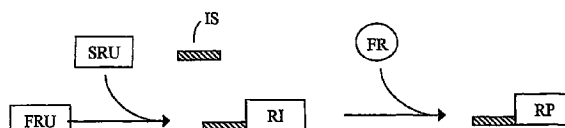
序列表 5 页 附图 11 页

(54) 发明名称

用于以核酸为模板的合成中的游离反应物

(57) 摘要

本发明提供了方法和组合物, 用于扩大可在以核酸为模板的有机合成期间进行的化学反应的范围。特别地, 以核酸为模板的化学方法用于产生反应中间产物, 其与可用于鉴定反应中间产物和/或得到的反应产物的寡核苷酸附着。然后反应中间产物与游离反应物(例如, 与寡核苷酸的偶联是困难的或不切实际的反应物)反应, 产生反应产物。该手段扩大了可用于以核酸为模板的合成中的试剂的范围, 将其扩大为: 不需要或者不能与寡核苷酸联结的试剂。但是, 这些试剂仍允许合成与可用于鉴定反应产物的寡核苷酸附着反应产物。



1. 一种通过以核酸为模板的合成来合成反应产物的方法,该方法包括如下步骤:

(a) 提供一种混合物,其包含 (i) 与包含密码子序列的第一寡核苷酸附着的第一反应单元和 (ii) 与包含与所述密码子序列互补的反密码子序列的第二寡核苷酸附着的第二反应单元;

(b) 将所述第一寡核苷酸的所述密码子序列与所述第二寡核苷酸的所述反密码子序列退火,诱发所述第一和第二反应单元之间的反应,以形成至少与所述第一寡核苷酸共价附着的反应中间产物;以及

(c) 将与所述第一反应单元共存的所述反应中间产物与能选择性地与所述反应中间产物反应的游离反应物混合,由此合成与所述第一寡核苷酸序列附着的反应产物,其中,所述游离反应物与所述反应中间产物的反应活性较之所述游离反应物对来自步骤 (a) 中提供的所述混合物的反应单元中的至少一种的反应活性更高。

2. 如权利要求 1 所述的方法,其中,所述反应产物以高于或等于 50% 的产率合成。

3. 如权利要求 1 所述的方法,其中,所述反应产物以高于或等于 75% 的产率合成。

4. 如权利要求 1 所述的方法,其中,所述反应产物以高于或等于 85% 的产率合成。

5. 一种通过以核酸为模板的合成来合成反应产物的方法,该方法包括如下步骤:

(a) 提供一种混合物,其包含多种不同的第一反应单元,所述第一反应单元与包含密码子序列的其各自的第一寡核苷酸相连,其中,所述寡核苷酸序列是与其附着的第一反应单元的指示物;

(b) 提供与第二寡核苷酸附着的第二反应单元,所述第二寡核苷酸包含与至少一种第一反应单元的所述密码子序列互补的反密码子序列;

(c) 将至少一种所述第一寡核苷酸的所述密码子序列与所述第二寡核苷酸的所述反密码子序列退火,以诱发所述第一和第二反应单元之间的反应,以形成至少与第一寡核苷酸共价附着的第一反应中间产物;以及

(d) 将与所述第一反应单元共存的所述第一反应中间产物与能选择性地与所述第一反应中间产物反应的游离反应物混合,由此合成与所述第一寡核苷酸相连的第一反应产物,其中,所述游离反应物与所述第一反应中间产物的反应活性较之所述游离反应物与来自所述混合物的所述反应单元的至少一种的反应活性更高。

6. 如权利要求 5 所述的方法,其还包括如下步骤:

(e) 提供第三反应单元,其与包含下述反密码子序列的第三寡核苷酸相连,所述反密码子序列与至少一种第一反应单元的所述密码子序列互补;

(f) 将所述第一寡核苷酸中至少一种的所述密码子序列与所述第三寡核苷酸的所述反密码子序列退火,以诱发所述第一和第三反应单元之间的反应,以形成至少与至少一种所述第一寡核苷酸附着的第二反应中间产物;以及

(g) 将所述第二反应中间产物与能选择性地与所述第二反应中间产物反应的游离反应物混合,由此合成与至少一种所述第一寡核苷酸附着的第二反应产物,其中,所述游离反应物与所述第二反应中间产物的反应活性高于所述游离反应物与所述混合物中的所述反应单元中的至少一种的反应活性。

7. 一种进行以核酸为模板的合成的体外方法,所述方法包括如下步骤:

(a) 提供一种混合物,其包含 (i) 每种都包含第一反应单元的多种不同模板,所述第一

反应单元与限定了密码子序列的第一寡核苷酸共价附着,以及(ii)包含第二反应单元的转移单元,所述第二反应单元与限定了反密码子序列的第二寡核苷酸共价附着,所述反密码子序列与所述模板的所述密码子序列互补;

(b) 将一种模板的所述密码子序列和所述转移单元的所述反密码子序列退火,以使所述第一反应单元与所述第二反应单元进入反应邻近状态,使得所述第一和第二反应单元相互反应,产生反应中间产物;以及

(c) 在所述中间反应产物与未反应的模板共同存在的情况下,将所述反应中间产物与游离反应物接触,所述游离反应物与所述反应中间产物发生化学反应,产生反应产物,其中,所述游离反应物与所述反应中间产物的反应活性比所述游离反应物与起始混合物中至少一种所述反应单元的反应活性高,并且其中,所述第一寡核苷酸保持与所述反应产物附着。

8. 如权利要求7所述的方法,其中,所述第一反应单元是小分子骨架。

9. 如权利要求7所述的方法,其中,步骤(b)中,所述第一反应单元的官能团被转化为所述反应中间产物中不同的化学基元。

10. 如权利要求8所述的方法,其中,所述小分子骨架含有受保护的官能团。

11. 如权利要求10所述的方法,其中,在步骤(b)中,所述官能团被去保护,以产生反应中间产物,其中,所述小分子骨架含有去保护的官能团。

12. 如权利要求7所述的方法,其中,所述反应产物

(a) 不是核酸;或

(b) 不是核苷酸或核苷酸类似物。

13. 如权利要求7所述的方法,其中,在步骤(b)中,所述第一和第二反应单元互相反应,以在没有核糖体的协助下产生所述反应中间产物。

14. 如权利要求7所述的方法,还包括选择与所述第一寡核苷酸附着的所述反应产物的步骤。

15. 如权利要求14所述的方法,还包括扩增所述第一寡核苷酸的步骤。

16. 如权利要求15所述的方法,包括如下额外步骤:测定与所述反应产物附着的所述第一寡核苷酸的序列,以确定所述反应产物的身份或合成历史。

17. 如权利要求1-16中任一项所述的方法,其中,所述游离反应物与所述反应中间产物的反应活性是所述游离反应物与起始混合物中至少一种所述反应单元的反应活性的至少5倍。

18. 如权利要求1-16中任一项所述的方法,其中,所述游离反应物与所述反应中间产物的反应活性是所述游离反应物与起始混合物中至少一种所述反应单元的反应活性的至少50倍。

19. 如权利要求1-16中任一项所述的方法,其中,所述游离反应物与所述反应中间产物的反应活性是所述游离反应物与起始混合物中至少一种所述反应单元的反应活性的至少1000倍。

20. 如权利要求17所述的方法,其中,所述游离反应物与所述反应中间产物的反应活性是所述游离反应物与起始混合物中任何所述反应单元的反应活性的至少5倍。

21. 如权利要求18中任一项所述的方法,其中,所述游离反应物与所述反应中间产

物的反应活性是所述游离反应物与起始混合物中任何所述反应单元的反应活性的至少 50 倍。

22. 如权利要求 19 中任一项所述的方法,其中,所述游离反应物与所述反应中间产物的反应活性是所述游离反应物与起始混合物中任何所述反应单元的反应活性的至少 1000 倍。

用于以核酸为模板的合成中的游离反应物

[0001] 相关申请

[0002] 本申请要求提交于 2005 年 1 月 21 日的美国专利申请序列号 No. 60/646, 584 的优先权和权益, 该公开的全文通过引用并入本文。

[0003] 政府基金

[0004] 本申请中描述的研究部分受 the Office of Naval Research(合同号 No. N00014-03-1-0749) 和 NIH/NIGMS(补助 R01GM065865) 的资助。美国政府可能具有本发明的某些权利。

[0005] 发明领域

[0006] 一般而言, 本发明涉及用于进行以核酸为模板的合成的方法和组合物。更具体地, 本发明涉及下述方法和组合物, 其用于开展以核酸为模板的合成, 以生产反应中间产物, 反应中间产物然后可使用游离反应物被化学转化为反应产物, 所述游离反应物与反应中间产物反应, 产生反应产物。

[0007] 发明背景

[0008] 以核酸为模板的有机合成使得可采用在传统合成方式中不可能的反应活性控制模式, 并且允许使用以前仅可被生物大分子使用的翻译、选择和扩增方法对合成分子加以操作 (Gartner et al. (2001) J. AM. CHEM. SOC. 123 :6961-3 ;Gartner et al. (2002) ANGEW. CHEM. , INT. ED. ENGL. 123 :61796-1800 ;Gartner et al. (2002) J. AM. CHEM. Soc. 124 : 10304-6 ;Calderone et al. (2002) ANGEW. CHEM. , INT. ED. ENGL. 41 :4104-8 ;Gartner et al. (2003) ANGEW. CHEM. , INT. ED. ENGL. 42 :1370-5 ;Li et al. (2004) J. AM. CHEM. SOC. 124 : 5090-2 ;Kanan et al (2004) NATURE 431 :545-9 ;(Gartner et al. (2004) SCIENCE 305 :1601-5 ; Li et al. (2004) ANGEW. CHEM. INT. ED. 43 :4848-70 ;Brenner et al. (1992) PROC. NATL. ACAD. SCI. USA 89 :5181 ;Doyon et al. (2003) J. AM. CHEM. SOC. 125 :12372-3 ;Halpin et al (2004) PLoS BIOL. 2 :e174)。可通过以核酸为模板合成 (特别地, 以 DNA 为模板的有机合成, 或 DTS) 得到的结构已基本被限制为由两个与核酸连接的反应物之间的偶联反应的产物。但是, 在一些情况下, 难于或不可能将反应物与寡核苷酸联结。因此, 对能使不与寡核苷酸连接的小分子试剂以序列程序化或序列记录的方式反应的策略的开发, 将显著扩大以核酸为模板的合成的合成能力。

发明内容

[0009] 本发明提供了方法和组合物, 用于扩大以核酸为模板的有机合成的范围, 这通过解决对与寡核苷酸联结 (tether) 的试剂的需要来实现。在试剂与核酸 (例如 DNA) 的连接是不可能或不方便的情况下, 这些转化仍允许此类试剂作用于小分子合成, 并同时保持核酸序列与产物结构之间的对应性。此外, 通过从偶联反应去偶联核酸模板的步骤, 该手段允许在不一定支持核酸杂交的条件下发生键的形成。

[0010] 在一个方面, 本发明提供了一种合成反应产物的方法。该方法包含如下步骤: (a) 提供一种混合物, 其包含第一反应单元和第二反应单元, 所述混合物在允许第一和第二反

应单元之间发生反应的条件下提供,以形成反应中间产物;(b) 提供寡核苷酸,其包含与反应中间产物附着(attach)的鉴定序列(identifying sequence);以及(c) 将反应中间产物与能选择性地与反应中间产物反应的游离反应物混合,由此产生与鉴定序列相连的反应产物。在此方法中,游离反应物与反应中间产物的反应活性较之游离反应物与起始混合物中的任何反应单元的反应活性更高。

[0011] 在另一个方面,本发明提供了一种合成反应产物的方法,这通过例如美国专利申请序列号 10/643,752(其以 US2004/0180412 的美国专利申请公开号公开)所述的以核酸为模板的合成来进行。该方法包括如下步骤:(a) 提供一种混合物,其包含(i) 与包含密码子序列的第一寡核苷酸附着的第一反应单元和(ii) 与包含与所述密码子序列互补的反密码子序列的第二寡核苷酸附着的第二反应单元;(b) 将第一寡核苷酸的密码子序列与第二寡核苷酸的反密码子序列退火,诱发第一和第二反应单元之间的反应,以形成至少与第一寡核苷酸附着的反应中间产物;以及(c) 将反应中间产物与能选择性地与反应中间产物反应的游离反应物混合,由此合成仍与第一寡核苷酸序列附着的反应产物。优选地,游离反应物与反应中间产物的反应活性比游离反应物与起始混合物中反应单元中的至少一种的反应活性更高。

[0012] 类似地,在多种不同的第一反应单元(以及可选地,第二反应单元)存在于初始反应混合物中时,在某些反应条件下,可能可以产生多种不同的反应中间产物。因此,对于游离反应物来说,选择性地仅与混合物中一种特定类型的反应中间产物反应可能是有利的。或者,对于游离反应物来说,选择性地与一组或一亚组反应中间产物反应可能是有利的,其中,所述反应中间产物具有有特定化学功能性的官能团。例如,游离反应物可能选择性地与含有游离胺的反应中间产物发生反应,而不与缺少游离胺的其它反应中间产物反应。

[0013] 此外,通过知晓密码子和/或反密码子序列,可以确定何种第二反应单元与第一反应单元反应,以产生反应中间产物和/或反应产物。此外,如果第一寡核苷酸提供了针对与第一寡核苷酸附着的第一反应单元的序列鉴定物(identifier),那么可以确定何种第一反应单元与第二反应单元反应,以产生反应中间产物和/或反应产物。基于反应条件,含有反应中间产物的反应混合物中存在的反应物,以及关于某些游离反应物何时添加进含有反应中间产物的混合物的信息,可以确定何种游离反应物与反应中间产物反应,以产生特定的反应产物。该信息可用于鉴定反应产物以及制造反应产物的反应途径。

[0014] 将参照下述附图、详述、实施例和权利要求来进一步理解本发明的前述方面和特征。

[0015] 定义

[0016] 本文中使用的术语“密码子”和“反密码子”指分别在模板和转移单元中的互补寡核苷酸序列,其允许转移单元在以核酸为模板的合成期间退火到模板上。

[0017] 本文中使用的术语“游离反应物”指不与能参与以核酸为模板的合成的寡核苷酸相连的化学试剂或化学基元(moiety)。相比之下,第一和第二反应单元和转移单元与能参与以核酸为模板的合成的寡核苷酸相附着。

[0018] 本文中使用的术语“寡核苷酸”或“核酸”指核苷酸的聚合物。聚合物可包括但不限于:天然核苷(例如,腺嘌呤核苷、胸腺嘧啶核苷、鸟嘌呤核苷、胞嘧啶核苷、尿嘧啶核苷、脱氧腺嘌呤核苷、脱氧胸腺嘧啶核苷、脱氧鸟嘌呤核苷以及脱氧胞嘧啶核苷)、核苷类

似物（例如，2- 氨基腺嘌呤核苷、2- 硫代胸腺嘧啶核苷、次黄苷、吡咯并嘧啶、2- 甲基腺嘌呤核苷、5- 甲基胞嘧啶核苷、C5- 溴尿嘧啶核苷、C5- 氟尿嘧啶核苷、C5- 碘尿嘧啶核苷、C5- 丙炔基 - 尿嘧啶核苷、C5- 丙炔基 - 胞嘧啶核苷、C5- 甲基胞嘧啶核苷、7- 脱氮腺嘌呤核苷、7- 脱氮鸟嘌呤核苷、8- 羟腺嘌呤核苷、8- 羟鸟嘌呤核苷、O(6)- 甲基鸟嘌呤和 2- 硫代胞嘧啶核苷）、经化学修饰的碱基、经生物修饰的碱基（例如，甲基化碱基）、嵌入碱基（intercalated base）、经修饰的糖（例如，2' - 氟核糖、核糖、2' - 脱氧核糖、阿拉伯糖和己糖）或者经修饰的磷酸基团（例如，硫代磷酸酯和 5' -N- 亚磷酰胺连接）。核酸和寡核苷酸还可包括具有经修饰的主链的碱基的其它聚合物，例如锁核酸（LNA）、肽核酸（PNA）、苏糖核酸（TNA）以及能用作模板用于使用扩增技术（例如聚合酶链式反应、连接酶链式反应或非酶促模板定点复制）的扩增反应的任何其它聚合物。

[0019] 本文中使用的术语“反应单元”指下述化学试剂或化学基元（包括：例如但不限于，构造块（building block）、单体、单体单元、小分子骨架（scaffold）或者可用于以核酸为模板的化学合成的其它反应物），其能与另一种化学试剂或化学基元一起参与化学反应以产生反应中间产物和 / 或反应产物。

[0020] 本文中使用的术语“反应中间产物”指下述化学试剂或化学基元，其能被游离反应物化学转化为不同的试剂或化学基元。

[0021] 本文中使用的术语“小分子”指在实验室合成或在自然界发现的、具有小于每摩尔 10000 克，可选地，小于每摩尔 5000 克，以及可选地，小于每摩尔 2000 克的有机化合物。

[0022] 本文中使用的术语“小分子骨架”指具有适合用于官能化（functionalization）的至少一个位点或化学基元的化学化合物。小分子骨架或分子骨架可能具有适合用于官能化的两个、三个、四个、五个或更多个位点或化学基元。这些官能化位点可被保护或被掩蔽，这是本领域技术人员已知的。位点还可被发现于基础环状结构或主链上。小分子骨架并非核酸、核苷酸或核苷酸类似物。

[0023] 本文中使用的术语“转移单元”指包含下述寡核苷酸的分子，所述寡核苷酸具有与反应单元附着的反密码子序列，所述反应单元包括，例如但不限于，构造块、单体、单体单元、小分子骨架或者可用于以核酸为模板的化学合成的其它反应物。

[0024] 本文中使用的术语“模板”指包含下述寡核苷酸的分子，所述寡核苷酸具有适合用于以核酸为模板的化学合成的至少一条密码子序列。可选地，模板可以包含 (i) 多条密码子序列，(ii) 扩增设施，例如，PCR 引物结合位点或与其互补的序列，(iii) 与其相联的反应单元，(iv) (i) 和 (ii) 的组合，(v) (i) 和 (iii) 的组合，(vi) (ii) 和 (iii) 的组合，或者 (i)、(ii) 和 (iii) 的组合。

[0025] 在说明书全文中，在组合物被描述为具有、包括或包含特定的多个组分的情况下，或者方法被描述为具有、包括或包含特定的多个工艺步骤的情况下，则认为，本发明的组合物基本上由，或由提到的组分构成，以及，本发明的方法也基本上由，或由提到的工艺步骤构成。此外，应当理解，用于进行某动作（action）的顺序或步骤顺序是不重要的，只要本发明保持可操作即可。此外，除非有相反指明，可同时进行两个或多个步骤或动作。

附图说明

[0026] 图 1 是本发明的一方面的示意性图示，其中，第一反应单元（FRU）与第二反应单元

(SRU) 反应,形成反应中间产物 (RI)。反应中间产物 (RI) 与鉴定序列 (IS) 附着。RI-IS 复合体与游离反应物 (FR) 混合,该游离反应物 (RU) 选择性地与 RI 反应,产生与 IS 相连的反应产物 (RP)。

[0027] 图 2 是本发明一个方面的示意性图示,其中,在反应中间产物 (RI) 形成之前,第一反应单元 (FRU) 和第二反应单元 (SRU) 分别与密码子序列 (CS) 和互补的反密码子序列 (ACS) 连接。CS 和 ACS 互相退火,以允许 FRU 和 SRU 互相反应,产生 RI。CS 保持与 RI 相连。当 RI 已与游离反应物 (FR) 反应时,CS 保持与反应产物 (RP) 的附着。

[0028] 图 3 是对图 1 所示方案的一种实施方式的示意性图示,其中,多种第一反应单元 (FRU₁-FRU₄) 的混合物中,第一反应单元中的至少一种 (FRU₁) 与第二反应单元 (SRU₁) 反应,形成反应中间产物 (RI₁),其与多种第一反应单元 (FRU₂-FRU₄) 共同存在。反应中间产物 (RI₁) 与鉴定序列 (IS) 相连。然后,令与 IS 相连并与多种第一反应单元 (FRU₂-FRU₄) 共同存在的反应中间产物 (RI₁) 与游离反应物 (FR) 反应,游离反应物选择性地与 RI₁ 反应,产生与 IS 相连的反应产物 (RP₁)。

[0029] 图 4 是图 2 所示的方案的一种实施方式的示意性图示,其中,在反应中间产物 (RI₁) 形成之前,第一反应单元 (FRU₁-FRU₄) 和第二反应单元 (SRU) 分别与密码子序列 (CS) 和互补的反密码子序列 (ACS) 连接。CS 保持与 RI₁ 相连。当 RI₁ 已与游离反应物 (FR) 反应产生 RP₁ 时,CS 保持与反应产物 (RP₁) 的附着。

[0030] 图 5A 描述了通过以 DNA 为模板的方式将叠氮化物转化为一级胺(最上面的方案)、羧酸(中间的方案)和硫醇(底下的方案)的过程。图 5B 描述了使用底物叠氮化物 1-12 进行的示例性的以 DNA 为模板的反应,其中,列出的产率代表下限。

[0031] 图 6A 是在来自图 5B 的叠氮化物 3 经过以 DNA 为模板的从叠氮化物向胺的转化之后,变性聚丙烯酰胺凝胶电泳 (PAGE) 凝胶的示意图。道 1 含有连接叠氮化物的 30 体模板。道 2 含有连接叠氮化物的 30 体模板 + 连接羧酸的 20 体捕获试剂 (capture reagent) + 30mM EDC + 15mM 硫 -NHS (显示没有产物形成)。道 3 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 (10 体不可见)。道 4 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 + 连接羧酸的 20 体捕获试剂 + 30mM EDC + 15mM 硫 -NHS,从叠氮化物到胺的转化之后进行以 DNA 为模板的胺酰化反应产生的 50 体次级产物是可见的。道 5 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 (含有错配序列) + 连接羧酸的 20 体捕获试剂 + 30mM EDC + 15mM 硫 -NHS。道 6 含有连接叠氮化物的 30 体模板 + 5mM TCEP-HCl + 连接羧酸的 20 体捕获试剂 + 30mM EDC + 15mM 硫 -NHS (阳性对照,其中叠氮化物被 TCEP 原位还原)。看起来,在电泳开始时 30 体模板和 20 体捕获反应物之间的双体的不完全变性导致了模糊的条带 (道 2 和 4-6)。

[0032] 图 6B 是在来自图 5B 的叠氮化物 7 经过以 DNA 为模板的从叠氮化物向胺的转化之后,变性 PAGE 凝胶的示意图。道 1 含有连接叠氮化物的 30 体模板 + 连接醛的 20 体捕获试剂 + 3mM NaBH₃CN。道 2 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 + 连接醛的 20 体捕获试剂 + 3mM NaBH₃CN。道 3 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 (含有错配序列) + 连接醛的 20 体捕获试剂 + 3mM NaBH₃CN。在道 1 和道 3 中观察到了少量的 50 体捕获产物的形成,这是在连接底物的模板制备期间和以 DNA 为模板的反应期间苯基叠氮化物缓慢自发还原产生的。减去道 1 和道 3 中观察到的背景反应活性 (<13%),以测定道

2 的报道产率。

[0033] 图 6C 是在来自图 5B 的叠氮化物 8 经过以 DNA 为模板的从叠氮化物向羧酸的转化之后,变性 PAGE 凝胶的示意图。道 1 含有连接叠氮化物的 30 体模板 + 连接胺的 20 体捕获试剂 + 30mM EDC + 15mM 硫 - NHS。道 2 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 + 连接胺的 20 体捕获试剂 + 30mM EDC + 15mM 硫 - NHS。道 3 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 (含有错配序列) + 连接胺的 20 体捕获试剂 + 30mM EDC + 15mM 硫 - NHS。

[0034] 图 6D 是在来自图 5B 的叠氮化物 11 经过以 DNA 为模板的从叠氮化物向硫醇的转化之后,变性 PAGE 凝胶的示意图,其中使用 10% 的聚丙烯酰胺凝胶。道 1 含有连接叠氮化物的 30 体模板 + 连接烷基溴的 20 体捕获试剂。道 2 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 + 连接烷基溴的 20 体捕获试剂。道 3 含有连接叠氮化物的 30 体模板 + 连接膦的 10 体试剂 (含有错配序列) + 连接烷基溴的 20 体捕获试剂。

[0035] 图 7 是来自以 DNA 为模板的官能团转化的代表性 MALDI-TOF 谱 (在该情况下,是来自叠氮化物 1 的胺产物)。

[0036] 图 8 描述了含有四种叠氮化物的单种溶液与四种未连接 DNA 的小分子亲电试剂的反应,所述反应生成四种序列程序化的磺胺、氨基甲酸酯、尿素和硫脲产物。模板 13 与具有密码子序列 a 的寡核苷酸附着。含有三苯基膦的转移单元 17 与具有反密码子序列 a' 的寡核苷酸附着。在模板化合成期间,密码子序列 a 退火到反密码子序列 a' 上。类似地,模板 14、15 和 16 分别含有密码子序列 b、c 和 d,它们分别通过反密码子序列 b'、c' 和 d' 退火到转移单元 18、19 和 20 上。

[0037] 图 9 是示出了图 8 所用的起始试剂和生成的反应产物的示意图。

[0038] 图 10A 到图 10D 是在对连接胺的模板和小分子试剂的反应进行 HPLC 分析之后的 HPLC 痕迹图 (在 260nm 监测的)。SM 代表未反应的起始材料,连接胺的模板,的峰。PRD 代表经衍生的产物。如无特别指明,除标记为 SM 或 PRD 之外的峰不对应于连接 DNA 的物质 (通过在 230nm 的 UV 吸收和 MALDI-TOF 分析判断的)。图 10A 示出了连接胺的模板 13 与丹磺酰氯 21 (图 8 中的试剂 13 和 21) 的反应结果。图 10B 示出了连接胺的模板 14 与氯甲酸乙酯 22 (图 8 中的试剂 14 和 22) 的反应结果。图 10C 示出了连接胺的模板 15 与 4- 甲氧基苯基异氰酸酯 23 (图 8 中的试剂 15 和 23) 的反应结果。图 10D 示出了连接胺的模板 16 与 6- 吗啉代吡啶基 3- 甲氧基苯基异氰酸酯 24 (图 8 中的试剂 16 和 24) 的反应结果。

[0039] 图 11A 示出了一种溶液中四种叠氮化物起始材料 (图 8 中的试剂 13-16) 的 MALDI-TOF 谱。图 11B 示出了叠氮化物起始材料 13-16 的四种序列特异性转化产物 (图 8 中的产物 25-28) 的 MALDI-TOF 谱。

[0040] 发明详述

[0041] 本发明可用于合成分子 (例如小分子) 的文库 (library)。本文所述的官能团转化特别有助于扩大以核酸为模板的有机合成的范围,这通过解决对与寡核苷酸联结的试剂的需要来实现。在试剂与寡核苷酸的连接是不可能或不方便的情况下,这些转化仍允许此类试剂作用于小分子合成,并同时保持核酸序列与产物结构之间的对应性。

[0042] 在一个方面,本发明提供了一种合成反应产物的方法。该方法包含如下步骤:(a) 提供一种混合物,其包含第一反应单元和第二反应单元,所述混合物在允许第一和第二反应单元之间发生反应的条件下提供,以形成反应中间产物;(b) 提供寡核苷酸,其包含与反

应中间产物附着的鉴定序列；以及 (c) 将反应中间产物与能选择性地与反应中间产物反应的游离反应物混合，由此产生与鉴定序列相连的反应产物。在此方法中，游离反应物与反应中间产物的反应活性较之游离反应物与起始混合物中的任何反应单元的反应活性更高。

[0043] 该手段示意性示于图 1 中。简言之，第一反应单元 (FRU) 与相异的第二反应单元 (SRU) 反应，产生反应中间产物 (RI)。RI 与鉴定序列 (IS) 附着，优选地，共价附着。IS 可以是寡核苷酸（例如，DNA 或其衍生物，RNA 或其衍生物）。然后 RI-IS 复合体与游离反应物 (FR) 在允许 FR 将 RI 化学转化为反应产物 (RP) 的条件下混合。RP 仍与 IS 相连，这可用于对 RP 以及 RP 的合成历史加以鉴定。

[0044] 在一种手段中，在产生 RI 的反应之前，IS，例如，限定了特定密码子序列或反密码子序列的核酸序列与 FRU 相连。在反应之后 IS 保持与 RI 相连，从而提供与 RI 相连的 IS。在产生 RP 之后，IS 保持与 RP 相连，使得可以对 RP 及其合成历史加以鉴定。应当考虑，SRU 还可能与互补于 IS 的序列相连。结果，在步骤 (a) 中，IS 与互补于 IS 的序列杂交，从而使 FRU 和 SRU 进入反应邻近状态 (reactive proximity)。

[0045] 在另一种手段中，在通过 FRU 和 SRU 之间的反应形成 RI 之后，IS，例如，限定了特定密码子序列或反密码子序列的核酸序列与 RI 相连。然后可通过 FR 将 RI 化学转化为 RP。IS 保持与 RP 相连。RI 形成后，可通过酶促方式，例如通过聚合酶或连接酶，将 IS 与 RI 相连。

[0046] 在一种实施方式中，FR 与 RI 的反应活性是 FR 与起始混合物中的其它反应中间产物或反应单元中的至少一种以及可选地全部的反应活性的至少五倍。此外，取决于反应物和反应条件，在其它实施方式中，RF 与 RI 的反应活性是 RF 与起始混合物中其它反应中间产物或反应单元中的至少一种以及可选地全部的反应活性的至少 10 倍，至少 50 倍，至少 100 倍，至少 250 倍，至少 500 倍，或者至少 1000 倍。此外，取决于反应物和反应条件，以高于或等于 50%，高于或等于 75%，高于或等于 85%，或高于或等于 98% 的产率合成 RP。

[0047] 可以实验测定 FR 对 RI 的反应活性（相对于起始材料 FRU 和 SRU 而言）。可以测定在标准反应条件下 FR 和 RI 混合产生的产物的量。可以测定在同样的反应条件下，FR 和 FRU 或 SRU 以等摩尔混合所产生的产物的量。可以通过化学领域的标准技术来测定产物的产率。基于每种反应中产生的产物的相对量，可以测定 FR 是否比 FRU 或 SRU 更具反应活性，以及如果是的话，反应活性大多少。类似的手段可用于测定游离反应物与一种反应中间产物的反应活性是否比其与其它不同的反应中间产物的反应活性更高。

[0048] 在另一个方面，本发明提供了一种合成反应产物的方法，该方法通过以核酸为模板的合成来实现，例如，美国专利申请公开号 US2004/0180412 所述的合成。该方法包括如下步骤：(a) 提供一种混合物，其包含 (i) 与包含密码子序列的第一寡核苷酸附着的第一反应单元和 (ii) 与包含与所述密码子序列互补的反密码子序列的第二寡核苷酸附着的第二反应单元，其中，所述反密码子序列是第二反应单元的指示物；(b) 将第一寡核苷酸的密码子序列与第二寡核苷酸的反密码子序列退火，诱发第一和第二反应单元之间的反应，以形成至少与第一寡核苷酸附着的反应中间产物；以及 (c) 将反应中间产物与能选择性地与反应中间产物反应的游离反应物混合，由此合成与第一寡核苷酸序列附着的反应产物。优选地，游离反应物与反应中间产物的反应活性比游离反应物与起始混合物中其它中间产物或反应单元中的至少一种以及可选地全部的反应活性更高。

[0049] 该手段示意性示于图 2 中,其中,第一反应单元 (FRU) 与包含密码子序列 (CS) 的第一寡核苷酸附着,例如,共价附着。第一反应单元和寡核苷酸的组合可称为模板。密码子序列可鉴定 FRU,例如,如 IS。或者,CS 可还包含能鉴定 FRU 的单独的鉴定序列 (IS)。在后一情况下,CS 可鉴定与 FRU 反应产生反应中间产物的第二反应单元 (SRU),IS 可鉴定 FRU。

[0050] 在此方法中,SRU 与第二寡核苷酸附着,例如,共价附着,第二寡核苷酸含有与 CS 互补的反密码子序列 (ACS)。第二反应单元与反密码子序列的组合被称为转移单元。当在合适的反应条件下混合模板和转移单元时,CS 和 ACS 互相退火,使得 FRU 和 SRU 进入反应邻近状态。然后 FRU 和 SRU 互相反应,例如通过邻近状态催化来进行,产生仍与 CS 相连的 RI。当与游离反应物 (FR) 混合时,RI 被 FR 化学转化为仍与 CS 相连的反应产物 (RP)。假设附着到 RP 上的寡核苷酸含有 CS,那么可以确定何种 SRU 参与对 RI 和 / 或 RP 的合成。类似地,如果附着到 RP 上的寡核苷酸含有 IS,那么可以确定何种 FRU 参与对 RI 和 / 或 RP 的合成。因此,该信息可用于确定 RI 和 / 或 RP 的身份以及合成历史。

[0051] 此外,可考虑,FRU 可以是小分子骨架,其可在以核酸为模板的合成期间使用,以产生小分子。特别地,小分子骨架可用作为核心,在其上装配小分子的取代基。

[0052] 在一种实施方式中,FR 与 RI 的反应活性是 FR 与起始混合物中其它反应中间产物或反应单元中至少一种以及可选地全部的反应活性的至少五倍。此外,取决于反应物和反应条件,在其它实施方式中,FR 与 RI 的反应活性是 FR 与起始混合物中其它反应中间产物或反应单元中至少一种以及可选地全部的反应活性的至少 10 倍,至少 50 倍,至少 100 倍,至少 250 倍,至少 500 倍,或者至少 1000 倍。此外,取决于反应物和反应条件,以高于或等于 50%,高于或等于 75%,高于或等于 85%,或高于或等于 98% 的产率合成 RP。

[0053] 在另一个方面,本发明提供了一种合成反应产物的方法。所述方法包括如下步骤:(a) 提供第二反应单元和多种不同的第一反应单元的混合物,该混合物在能诱发第一反应单元中的至少一种与第二反应单元之间发生反应的条件提供,由此形成与多种第一反应单元共同存在的反应中间产物;(b) 提供与反应中间产物附着的鉴定序列,其中,该序列能将反应中间产物与第一反应单元区别开;以及将与第一反应单元共同存在的反应中间产物与能选择性地与反应中间产物反应的游离反应物混合,由此合成与鉴定序列相连的反应产物,反应产物与多种第一反应单元共同存在。游离反应物与反应中间产物的反应活性比游离反应物与起始混合物中其它反应中间产物或反应单元中的至少一种的反应活性更高。

[0054] 该手段示意性地展示于图 3 中,其与图 1 示意性展示的手段相似,不同之处在于,第一反应单元作为多种第一反应单元的混合物存在。简言之,含有被称为 FRU_1 、 FRU_2 、 FRU_3 和 FRU_4 的四种第一反应单元的起始混合物与被称为 SRU_1 第二反应单元混合。在合适的条件下, FRU_1 和 SRU_1 相互反应,产生被称为 RI_1 的反应中间产物。鉴定序列 (IS) 可与 RI_1 附着,其能鉴定 RI_1 。之后,在允许将 RI 化学转化为被称为 RP_1 的反应产物的条件下,将游离反应物 (FR) 混合进前述的混合物。 RP_1 仍与可用于鉴定 RP 及 RP 的合成历史的 IS 相连。此外,可以考虑, RP_1 可经过多轮其它官能团转化,尤其是 FRU 是小分子骨架时,以产生经过进一步修饰的产物。

[0055] 在另一个方面,本发明提供了一种合成反应产物的方法,该方法通过以核酸为模板的合成来进行。该方法包括如下步骤:(a) 提供一种混合物,其中包含 (i) 多种不同的第一反应单元,它们每种都与包含密码子序列的第一寡核苷酸相连,其中,每种寡核苷酸序列

还是与其附着的第一反应单元的指示物；(b) 提供与第二寡核苷酸附着的第二反应单元，第二寡核苷酸包含与至少一种第一反应单元的密码子序列互补的反密码子序列，其中，所述反密码子序列是第二反应单元的指示物；(c) 将至少一种第一寡核苷酸的密码子序列与第二寡核苷酸的反密码子序列退火，以诱发第一和第二反应单元之间的反应，以形成至少与第一寡核苷酸相连的第一反应中间产物；以及 (d) 将第一反应中间产物与能选择性地与第一反应中间产物反应的游离反应物混合，由此合成与鉴定序列相连的第一反应产物，其中，所述游离反应物与第一反应中间产物的反应活性高于游离反应物与混合物中存在的其它反应中间产物或反应单元的至少一种以及可选地全部的反应活性。

[0056] 该手段示意性展示于图 4 中，其与图 2 示意性展示的手段类似，不同之处仅在于，第一反应单元作为多种第一反应单元的混合物存在。简言之，最初的反应含有多种模板，其中每种模板含有第一反应单元（被称为 FRU_1 、 FRU_2 、 FRU_3 和 FRU_4 ），它们分别与其各自的寡核苷酸（含有其自己的密码子独特序列（CS））附着，优选地，共价附着。寡核苷酸优选还含鉴定序列，该序列能鉴定出何种第一反应单元与模板的何种密码子序列附着。含有与含有反密码子序列的寡核苷酸附着（优选地，共价附着）的第二反应单元（SRU）的转移单元与模板混合，所述反密码子序列与密码序列互补。转移单元的 ACS 退火到模板的 CS 上，以使 FRU_1 和 SRU 进入反应邻近状态，由此 FRU_1 和 SRU 相互反应，产生仍保持与含 CS 的寡核苷酸附着的反应中间产物（ RI_1 ）。当与游离反应物（FR）混合时， RI 被 FR 化学转化以产生仍与 CS 相连的反应产物（RP）。假设附着到 RP 上的寡核苷酸含有 CS，那么可以确定何种 SRU 参与对 RI 和 / 或 RP 的合成。类似地，如果附着到 RP 上的寡核苷酸含有 IS，那么可以确定何种 FRU 参与对 RI 和 / 或 RP 的合成。因此，该信息可用于确定 RI 和 / 或 RP 的身份以及合成历史。如所讨论的，可以考虑，第一反应单元可以是用于设计和合成小分子文库的小分子骨架。

[0057] 此外，多种不同的官能团转化在同一反应容器中同时发生是可能的。因此，所述方法还可包括如下额外步骤：提供相异的第三反应单元，其与包含反密码子序列的第三寡核苷酸相连，所述反密码子序列与至少一种不同的第一反应单元的密码子序列互补，其中，该反密码子序列是第三反应单元的指示物；将一种不同的第一寡核苷酸的密码子序列与第三寡核苷酸的反密码子序列退火，以诱发第一和第三反应单元之间的反应，以形成至少与不同的第一寡核苷酸附着的第二反应中间产物；以及将第二反应中间产物与能选择性地与第二反应中间产物反应的游离反应物混合，由此合成与鉴定序列附着的第二反应产物，其中，所述游离反应物与第二反应中间产物的反应活性高于游离反应物与混合物中存在的其它反应中间产物或反应单元的至少一种以及可选地全部的反应活性。

[0058] 此外，可以考虑到，反应产物可经过其它轮次的官能团转化（例如，当 FRU_1 是小分子骨架时），用于产生经过进一步修饰的产物。

[0059] 本领域技术人员将理解，本发明的方法可用于扩大在以核酸为模板的化学合成期间可以使用的化学剂的范围。关于对模板、转移单元、反应条件、反应化学、选择方案的选择和使用方面的一般性考虑是本领域已知的。关于这些考虑的一般性讨论见下文。

[0060] I. 模板考虑

[0061] 核酸模板可指导多种化学反应，没有明显的结构要求，这通过序列特异性地引来与互补寡核苷酸相连的反应物实现。在合成期间，模板与一种或多种转移单元杂交或退火，

以指导反应中间产物的合成,反应中间产物随后可被游离反应物转化为反应产物。然后可基于某些标准,例如与预先选择的目标分子结合的能力,对反应产物加以选择或筛选。一旦已鉴定出了反应产物,那么可对其相联的模板加以测序,以破译反应中间产物和 / 或反应产物的合成历史。

[0062] (i) 模板形式

[0063] 模板的长度可在很大范围变化,这取决于所关注的以核酸为模板的合成的类型。例如,在某些实施方式中,模板可以为 10 至 10000 个核苷酸长,20 至 1000 个核苷酸长,20 至 400 个核苷酸长,40 至 1000 个核苷酸长,或者 40 至 400 个核苷酸长。模板的长度当然将取决于,例如,密码子的长度、文库的复杂程度、反应产物的复杂程度和 / 或大小、间隔序列的使用等。

[0064] 模板可在一个末端包括以反应单元结尾的发夹环 (hairpin),其可与一种或多种与转移单元相联的反应单元反应。例如,DNA 模板可包含以 5'-氨基结尾的发夹环,其可以或可以不是受保护的。氨基可充当形成非天然聚合物或小分子的起始点。

[0065] (ii) 密码子使用

[0066] 可以考虑到,可以以多种方式来设计模板序列。例如,必须确定密码子长度,必须设定密码子序列。如果使用的密码子长度为 2,那么使用四种天然存在的碱基,将仅有 16 种可能的组合用于对文库加以编码。如果密码子的长度增加至 3(编码蛋白质天然使用的数字),可能的组合的数量就增加至 64。如果密码子的长度增加至 4,那么可能的组合的数量就增加至 256。在确定密码子的长度中要考虑的其它因素是错配、读码框位移 (frame-shifting)、文库复杂程度等。随着密码子长度增加至某个点,错配的数量会降低;但是,过长的密码子在有多个错配碱基对时可能也将杂交。

[0067] 虽然密码子长度可变,但密码子可能在 2 至 50 个核苷酸,2 至 40 个核苷酸,2 至 30 个核苷酸,2 至 20 个核苷酸,2 至 15 个核苷酸,2 至 10 个核苷酸,3 至 50 个核苷酸,3 至 40 个核苷酸,3 至 30 个核苷酸,3 至 20 个核苷酸,3 至 15 个核苷酸,3 至 10 个核苷酸,4 至 50 个核苷酸,4 至 40 个核苷酸,4 至 30 个核苷酸,4 至 20 个核苷酸,4 至 15 个核苷酸,4 至 10 个核苷酸,5 至 50 个核苷酸,5 至 40 个核苷酸,5 至 30 个核苷酸,5 至 20 个核苷酸,5 至 15 个核苷酸,5 至 10 个核苷酸,6 至 50 个核苷酸,6 至 40 个核苷酸,6 至 30 个核苷酸,6 至 20 个核苷酸,6 至 15 个核苷酸,6 至 10 个核苷酸,7 至 50 个核苷酸,7 至 40 个核苷酸,7 至 30 个核苷酸,7 至 20 个核苷酸,7 至 15 个核苷酸,7 至 10 个核苷酸,8 至 50 个核苷酸,8 至 40 个核苷酸,8 至 30 个核苷酸,8 至 20 个核苷酸,8 至 15 个核苷酸,8 至 10 个核苷酸,9 至 50 个核苷酸,9 至 40 个核苷酸,9 至 30 个核苷酸,9 至 20 个核苷酸,9 至 15 个核苷酸,9 至 10 个核苷酸的范围内。但是,密码子优选为 3、4、5、6、7、8、9 或 10 个核苷酸长。

[0068] 优选地,用于模板中的密码子的组使得密码子组中任何两个密码子之间错配的数量最大化,以确保仅正确的转移单元反密码子退火到模板的密码子位点上。此外,下述是重要的:模板在一个密码子组的所有成员间以及不同密码子组的所有密码子间都具有错配,以确保反密码子不会与错的密码子组错误结合。对示例性密码子组的选择以及产生功能性密码子组的方法在例如美国专利公开号 US2004/0180412 中有描述。采用这种手段和其它手段,可以产生不同的密码子组,使得没有密码子重复。

[0069] 当以核酸为模板的合成被用于产生聚合物或小分子时,还可在密码子之间放置间

隔序列,以防止读码框位移。例如,可对编码聚合物亚基的模板(即,用于聚合物的“遗传密码”)的碱基加以选择,以使得读码框外退火的可能性最小化。这些遗传密码减少了不想要的对经过读码框位移的以核酸为模板的聚合物的翻译,并且预期熔点温度范围不同,在读码框外的退火期间导致的错配最小数量不同。

[0070] (iii) 模板合成

[0071] 可以使用本领域公知的方法来合成模板。这些方法包括体内和体外方法,包括PCR、质粒制备、内切核酸酶消化、固相合成(例如,使用自动合成仪)、体外转录、链分离等。合成之后,如果需要的话,可使用本领域已知的标准偶联化学,将模板附着(例如,共价或非共价附着)到感兴趣的反应单元上。

[0072] 用于合成大量模板的一种有效方法是使用“分裂库(split-pool)”技术。使用标准的3'至5'化学方法来合成寡核苷酸。首先,合成恒定的3'末端。然后将其分进n个不同的容器,n是在该模板中的该位置出现的不同密码子的数量。对于每个容器,在恒定的3'末端的(伸长中的)5'末端上合成n个不同密码子中的一个。由此,每个容器从5'到3'含有与恒定的3'末端附着的不同的密码子。然后合并n个容器,使得单个容器含有与恒定的3'末端附着的n个不同密码子。现在合成了临近密码子的5'末端的任何恒定碱基。然后将该库分进m个不同的容器,其中m是在模板的下一个(更5')位置出现的不同密码子的数量。在m个容器的每个中合成不同的密码子(在伸长中的寡核苷酸的5'末端)。在单个容器中合并得到的寡核苷酸。按照需要重复分开、合成及合并,以合成寡核苷酸中的所有密码子和恒定区域。

[0073] II. 转移单元

[0074] 转移单元包含含有反密码子序列的寡核苷酸和反应单元。反密码子被设计为与模板中存在的密码子互补。因此,当设计反密码子时应当考虑模板中使用的序列和密码子长度。与模板中使用的密码子互补的任何分子都可使用,包括天然的或非天然的核苷酸。在某些实施方式中,密码子包括一个或多个自然界中发现的碱基(即,胸腺嘧啶、尿嘧啶、鸟嘌呤、胞嘧啶和腺嘌呤)。因此,反密码子可包括一个或多个具有碱基、糖以及可选的磷酸基团的、在自然界中正常发现的核苷酸。

[0075] 如上文所讨论的,反密码子与特定种类的反应单元相联,以形成转移单元。反应单元可代表不同的实体,或者可以是反密码子单元的官能部分。在某些其它实施方式中,当要生成小分子文库而非聚合物文库时,通常将反密码子与用于修饰小分子骨架的反应物或反应单元相联。在某些实施方式中,通过长到足以允许反应物与小分子骨架进入反应邻近状态的接头(linker),将反应物与反密码子连接起来。接头优选具有允许分子内反应但最小化分子间反应的长度和组成。反应物包括大量试剂,如可用于以核酸为模板的合成的多种反应所证实的,并且其可以是化学领域已知的任何化学基团、催化剂(例如,有机金属化合物)或反应基元(例如,亲电试剂、亲核试剂)。

[0076] 在某些实施方式中,每种反密码子序列与一种单体类型相联。例如,反密码子序列ATTAG可与具有异丁基侧链的氨基甲酸残基相联,反密码子序列CATAG可与具有苯基侧链的氨基甲酸残基相联。这种将反密码子与单体单元一一对应的方式允许通过对合成中所用的核酸模板进行测序来对文库中任何聚合物加以破译,以及允许通过知道原来的聚合物的序列而对同样的聚合物或相关的聚合物加以合成。通过改变(突变)模板的序列,可以引

入不同的单体单元,由此允许合成相关聚合物,其随后可被选择及发展 (evolve)。在某些优选的实施方式中,若干反密码子可针对一种单体编码,与自然界中情况一样。

[0077] 反密码子可通过接头基元与反应物相联。连接可被光照、氧化、水解、暴露给酸、暴露给碱、还原等切开。Fruchtel et al. (1996) ANGEW. CHEM. INT. ED. ENGL. 35 :17 描述了多种可用于实践本发明的连接。接头协助反应物与小分子骨架的接触,以及在某些实施方式中,取决于想要的反应,将 DNA 定位为离去基团 (“可自动切割”策略),或者可以通过“无痕”接头策略 (其产生不留下具有化学功能性的额外一个或多个原子的产物) 或“有用的疤痕”策略 (其中,接头的一部分留下,以在接头切割后的后续步骤中被官能化) 将反应基团与模板连接起来。有用的接头、对它们的设计和用途在美国专利申请公开号 No. US2004/0180412 中有所描述。

[0078] 转移单元对模板的特异性退火允许以比很多传统有机合成中所用的浓度更低的浓度来使用转移单元。因此,转移单元可以以小于 1 毫摩尔的浓度使用 (例如,小于 100 μ M, 小于 10 μ M, 小于 1 μ M, 小于 100nM, 或者小于 10nM)。

[0079] III. 化学反应

[0080] 可以使用本文所述的方法来制备多种化合物和 / 或文库。在某些实施方式中,根据本发明的方法合成不是或者不类似核酸或其类似物的化合物。在某些其它实施方式中,根据本发明的方法合成不是或不类似蛋白质、肽或其类似物的化合物。

[0081] (i) 用于小分子合成的偶联反应

[0082] 在一些实施方式中,可使用本文所述的方法制造化合物,例如小分子。这些小分子可以例如是类天然产物,非多聚的,和 / 或非寡聚的。小分子中的实质性兴趣部分在于它们在很多药物制剂中作为活性成分的用途,但是它们还可用作,例如催化剂、材料或添加剂。

[0083] 在使用本发明的方法合成小分子的过程中,可以使用可发展的 (evolvable) 模板。模板可包括小分子骨架,在其上将构建小分子,或者可向模板中加入小分子骨架。小分子骨架可以是具有用于官能化的两个或多个位点的任何化学化合物。例如,小分子骨架可包括具有可官能化的基团的环系统 (例如,在胆固醇中发现的 ABCD 类固醇环系统),所述基团与组成环的原子偶联。在另一个例子中,小分子可以是药物试剂的基础结构,例如,吗啡、埃博霉素 (epothilone) 或头孢菌素抗生素。可用本领域已知的方法和保护基团对小分子骨架上将被官能化的位点或基团进行保护。用于小分子骨架的保护基团可以是互相垂直的,使得保护基团可以一次被去除一个。

[0084] 在此方法中,转移单元包含下述反密码子,其与用于修饰小分子骨架或加到小分子骨架上或从小分子骨架消去的反应物或构造块相连。试剂或反应块可以是,例如,亲电试剂 (例如,乙酰基、酰胺、酰氯 (acidchloride)、酯、腈、亚胺)、亲核试剂 (例如,胺、羟基、硫醇)、催化剂 (例如,有机金属催化剂) 或侧链。在杂交条件下令转移单元与模板接触。作为寡核苷酸退火的结果,附着的反应物或构造块被允许与小分子骨架上的位点反应,产生一种或多种反应中间产物。然后反应中间产物可与游离反应物反应,产生反应产物。

[0085] 对反应条件、接头、反应物和将被官能化的位点进行选择,以避免分子间反应,以及促进分子内反应。可将模板与转移单元顺序或同时接触,这取决于将被合成的特定化合物。

[0086] 在骨架上的位点被修饰之后,新合成的小分子保持与编码其合成的模板相接触。对模板序列的破译允许揭示 (deconvolution) 合成历史,以及由此揭示小分子的结构。还可对模板加以扩增,以制造更多的目标小分子和 / 或模板可被发展 (诱变),以制造相关的小分子。还可从模板切下小分子,用于纯化或筛选。

[0087] (ii) 用于聚合物合成的偶联反应

[0088] 在某些实施方式中,可使用本文所述的技术来制备聚合物,特别是非天然聚合物。示例性非天然聚合物包括但不限于,肽核酸 (PNA) 聚合物、聚氨基甲酸酯、聚脲、聚酯、聚丙烯酸酯、聚烯烃 (聚乙烯、聚丙烯)、聚碳酸酯、具有非天然立体化学的多肽、具有非天然氨基酸的多肽及其组合。在某些实施方式中,聚合物包含 10、25、75、100、125、150 个单体单元或更多。使用本文所述的方法合成的聚合物可用作为,例如催化剂、药物、金属螯合剂或催化剂。

[0089] 在制备某些非天然聚合物的过程中,附着到反密码子的单体单元可以是能连接起来形成聚合物的单体或寡聚体。单体单元可以是,例如,氨基甲酸酯、D-氨基酸、非天然氨基酸、PNAs、尿素、羧酸、酯、碳酸酯、丙烯酸酯或醚。

[0090] (iii) 反应条件

[0091] 应当理解,以核酸为模板的反应,例如用于生产反应中间产物的以核酸为模板的反应,可在水性或非水性 (即有机) 溶液或一种或多种水性及非水性溶液的混合物中发生。在水性溶液中,反应可以在大约 2 至大约 12 的 pH 范围进行,或者优选地,大约 2 至大约 10,或者更优选地,大约 4 至大约 10。用于以 DNA 为模板的化学方法中的反应优选不需要非常碱性的条件 (例如, $\text{pH} > 12$ 、 $\text{pH} > 10$) 或者非常酸性的条件 (例如 $\text{pH} < 1$ 、 $\text{pH} < 2$ 、 $\text{pH} < 4$), 因为极端条件可能导致核酸模板和 / 或被合成的分子 (例如,聚合物或小分子) 的降解或改性。水性溶液可含有多种浓度下的一种或多种无机盐,其包括但不限于 NaCl 、 Na_2SO_4 、 KCl 、 Mg^{+2} 、 Mn^{+2} 等。

[0092] 适合用于以核酸为模板的反应的有机溶剂包括但不限于:二氯甲烷、氯仿、二甲基甲酰胺和有机醇 (包括甲醇和乙醇)。为允许将反应组分定量溶解于有机溶剂中,可以加入季铵盐,例如,长链四烷基铵盐 (Jost et al (1989) NUCLEIC ACIDS RES. 17 :2143 ; Mel'nikov et al. (1999) LANGMUIR 15 :1923-1928)。

[0093] 以核酸为模板的反应可能需要催化剂,例如,均相的、非均相的、相转移和不对称催化。在其它一些实施方式中,并不需要催化剂。不与核酸相连的额外辅助试剂的存在在某些实施方式中是优选的。有用的辅助试剂可包括,例如,氧化剂 (例如 NaIO_4); 还原剂 (例如 NaCNBH_3); 活化剂 (例如 EDC、NHS 和硫 -NHS); 过渡金属,例如镍 (例如 $\text{Ni}(\text{NO}_3)_2$)、铑 (例如 RhCl_3)、钌 (例如 RuCl_3)、铜 (例如 $\text{Cu}(\text{NO}_3)_2$)、钴 (例如 CoCl_2)、铁 (例如 $\text{Fe}(\text{NO}_3)_3$)、钼 (例如 OsO_4)、钛 (例如 TiCl_4 或四异丙醇钛)、钯 (例如 NaPdCl_4) 或 Ln; 过渡金属配体 (例如,膦、胺和卤化物); Lewis 酸; 和 Lewis 碱。

[0094] 优选地,对反应条件加以优化,使其适合于所用的反应单元和寡核苷酸的性质。应当理解,对用于制造反应中间产物以及将反应物中间产物转化为终产物的反应条件和试剂 (例如游离反应物) 的选择取决于将要生产的特定化合物和文库。但是应当考虑到,对试剂和反应条件的选择在本领域技术人员水平范围内。

[0095] (iv) 化学反应的类型

[0096] 应当理解,大量不同的化学反应可用于制造反应中间产物和 / 或从反应中间产物制造反应产物。用于合成聚合物、小分子或其它分子的已知化学反应可用于以核酸为模板的反应中。因此,可以使用例如 March' sAdvanced Organic Chemistry, Organic Reactions, Organic Syntheses、有机教科书、杂志,例如 Journal of the American Chemical Society、Journal of Organic Chemistry、Tetrahedron 等,和 Carruther' s Some Modern Methods of Organic Chemistry 中列出的反应。选择的反应优选与核酸(例如 DNA 或 RNA) 兼容,或者与用作模板的经修饰的核酸兼容。

[0097] 尽管如上所述,但是应当考虑到,本发明特别有利于进行某些官能团转化,其包括但不限于:从叠氮化物到胺的转化、从叠氮化物到硫醇的转化、从叠氮化物到羧酸的转化、从羟基到胺的转化、从羟基到硫醇的转化、从缩醛到醛的转化、从缩酮到酮的转化、从碳酸酯到羟基的转化、从氨基甲酸酯到胺的转化、从硫代碳酸酯到硫醇的转化、从硝基到胺的转化、从磺胺到胺的转化、从烯烃到环氧化物的转化、从 α , β - 不饱和酮到环氧化物的转化、从环氧化物到 1,2- 二醇的转化、从环氧化物到 1,2- 羟胺的转化、从环氧化物到 1,2- 羟基硫化物的转化、从烯烃到氮丙啶的转化、从氮丙啶到 1,2- 二胺的转化、从氮丙啶到 1,2- 氨基硫化物的转化、从磷酸酯到磷酸的转化、从酰亚胺到胺的转化以及从腈到羧酸的转化。其它示例性转化可在例如 Greene et al. (ed.) (1999) PROTECTIVE GROUPS IN ORGANIC SYNTHESIS 3RD ED., Wiley-Interscience 和 Kocienski (1994) PROTECTING GROUPS, Thieme 中找到。

[0098] IV. 对产物的选择、筛选和鉴定

[0099] (i) 选择和筛选手段

[0100] 针对具有想要的活性(例如,催化活性、结合亲和性或者活性试验中的特定效果)的反应产物的选择和 / 或筛选可使用本领域使用的及本领域已知的方法来进行。例如,亲和性选择可按照基于文库的选择方法(例如,噬菌体展示、多核糖体展示以及 mRNA 融合蛋白展示的肽)中所用的原则来进行。对催化活性的选择可通过在过渡态类似物亲和柱上的亲和性选择 (Baca et al. (1997) PROC. NATL. ACAD. SCI. USA 94 (19) :10063-8) 来进行或者通过基于功能的选择方案 (Pedersen et al. (1998) PROC. NATL. ACAD. SCI. USA 95 (18) :10523-8) 来进行。因为可以通过 PCR 扩增微量 DNA (大约 10^{-20} 摩尔) (Kramer et al. (1999) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (ed. Ausubel, F. M.) 15.1-15.3, Wiley), 因此可以用比通过目前的方法进行反应分析所需要的数量少十个或更多数量级的规模来进行这些选择,进行真正宽广的经济高效搜索。

[0101] 可针对对靶分子的结合,对模板和反应产物加以选择(或筛选)。在该上下文中,选择或分开 (partitioning) 指下述任何过程,通过所述过程,与靶分子结合的文库成员与未和靶分子结合的文库成员分离开。选择可通过本领域已知的多种方法来进行。

[0102] 本发明的模板含有用于直接选择和扩增的内建功能。在多种应用中,与靶分子结合优选是选择性的,使得模板和得到的反应产物优先与特定靶分子结合,这有可能阻止或诱发特定的生物学效应。最终,使用本发明鉴定的结合分子可作为治疗和 / 或诊断试剂。一旦完成了选择,可选地,选出的模板可被扩增和测序。选出的反应产物如果以足够量存在,那么其可与模板分离,被纯化(例如,通过 HPLC、柱色谱或其它色谱方法),以及被进一步分析。

[0103] 可开展选择策略,以允许针对几乎任何靶来选择。重要地,选择策略不需要关于靶分子或文库中分子的任何详细的结构信息。整个过程由文库中分子对给定的靶的特异性识别和结合所涉及的结合亲和性驱动。

[0104] 模板分子和反应产物之间的连接允许使用多种选择策略对结合分子进行快速鉴定。以核酸为模板的合成广泛地允许鉴定出针对任何已知靶分子的结合分子。此外,可通过下述方法发现新颖的未知靶:分离出针对未知抗原(表位)的结合分子,使用这些结合分子进行鉴定和确认。

[0105] 从文库中选择结合分子可以以任何形式进行,以鉴定最优结合分子。典型地,结合选择包括:固定想要的靶分子,加入可能的结合子的文库,以及通过洗涤除去未结合物。当展示出针对被固定靶的低亲和性的分子被洗去时,通常,具有更强的亲和性的分子保持与靶的附着。优选地,对严谨洗涤后保持与靶结合的经富集群体进行洗脱,这用例如酸、离液 chaotropic 盐、加热、用已知配体的竞争性洗脱或者通过靶和 / 或模板分子的蛋白水解性释放来进行。经洗脱的模板适合用于 PCR,导致很多数量级的扩增,由此,几乎每种选出的模板都可以剧烈增加的拷贝数获得,用于克隆、测序和 / 或进一步富集和多样化 (diversification)。

[0106] 靶分子(肽、蛋白、DNA 或其它抗原)可被固定到固体支持物上,例如,容器壁、微滴定板孔壁。优选地,将文库溶解于一个罐中的水性结合缓冲液中,在存在固定的靶分子的情况下对其进行平衡。用缓冲液洗去未结合物。可通过用缺少 PCR 引物结合位点的非官能化模板对经结合的文库加以洗涤,除去可能通过其附着的 DNA 模板而非通过其合成基元与靶分子结合的那些分子。然后例如通过变形可洗脱保持结合的文库成员。

[0107] 或者,可将靶分子固定到珠粒 (bead) 上,特别是当怀疑靶分子将充分吸附到容器壁上时,如在从 SDS-PAGE 凝胶洗脱的未折叠靶的情况下。然后可将经过衍生化的珠粒用于分离高亲和性文库成员和未结合物,这通过在台式离心机上简单地沉淀珠粒来进行。或者,可将珠粒用于制造亲和柱。在此类情况下,文库经过该柱一次或多次,以允许结合。然后对该柱进行洗涤,以除去未结合的文库成员。磁性珠粒本质上是上述珠粒的变体;靶与磁性珠粒附着,然后珠粒用于选择。

[0108] 可通过变性、酸或离液盐来释放结合靶分子的文库成员。或者,洗脱条件可以更特异,以减少背景,或者针对想要的特异性来选择。可使用蛋白水解,以切割靶分子和固定表面之间的接头,或者切割反应产物和模板之间的接头来完成洗脱。或者,洗脱可以通过与针对靶分子的竞争性配体的竞争来完成。或者,可在选择程序末尾,存在洗过的靶分子时,直接进行 PCR 反应。由此,结合分子无须可从靶上洗脱即具有选择性,因为进一步扩增或克隆仅需要模板,而非反应产物本身。事实上,一些靶分子与最渴望的配体非常紧密地结合,使得难于进行洗脱。

[0109] (ii) 鉴定产物

[0110] 一旦完成了所有轮次的选择,优选地,使用任何合适的技术对与或曾经与选出的反应产物附着的模板加以扩增,以便于测序或者对模板的后续操作。可通过本领域的任何已知方法来扩增天然的寡核苷酸。这些方法包括,例如,聚合酶链式反应 (PCR);基于核酸序列的扩增(见,例如,Compton (1991) NATURE 350:91-92),扩增的反义 RNA(见,例如, vanGelder et al. (1988) PROC. NATL. ACAD. SCI. USA 85:77652-77656);自我维持的序列复

制系统 (Gnatelli et al. (1990) PROC. NATL. ACAD. Sci. USA 87 :1874-1878) ;不依赖聚合酶的扩增 (见,例如 Schmidt et al. (1997) NUCLEIC ACIDS RES. 25 :4797-4802) 以及对携带克隆 DNA 片段的质粒的体内扩增。对 PCR 方法的描述例如可以在 Saiki et al. (1985) Science 230 :1350-1354 ;Scharf et al. (1986) Science 233 :1076-1078 ;以及美国专利 No. 4683202 中找到。还可使用连接酶介导的扩增方法,例如连接酶链式反应 (LCR)。通常,允许对选出的核酸序列进行可靠、有效扩增的任何手段可用于本发明的方法中。优选地,但是非必须地,扩增后序列的成比例情况反映了扩增前混合物中序列的相对比例。

[0111] 对于非天然核苷酸来说,对高效扩增程序的选择较少。因为非天然核苷酸可被某些酶 (包括聚合酶) 包括进去,因此,将可能通过在每个延伸循环期间加入聚合酶来进行人工的聚合酶链式反应。

[0112] 一旦被扩增,编码感兴趣的产物的模板序列即可被确定。测序例如可通过标准的双脱氧链终止法来进行,或者通过化学测序,例如使用 Maxam-Gilbert 测序程序来进行。或者,可通过杂交实验来测定模板 (或者,如果使用较长模板的话,其可变的部分) 的序列。例如,与可探测基元 (例如荧光基元) 相联的单链模板分子被暴露给芯片,该芯片携带有已知序列的单链核酸或核酸类似物的大量克隆群体,每种克隆群体存在于芯片上可设定地址的特定位置。模板序列被允许退火到芯片序列上。然后测定芯片上可探测基元的位置。基于可探测基元的位置以及处于该位置的被固定序列,可以确定模板的序列。可以考虑在芯片或其它固体支持物上的阵列中固定大量的此类核酸。

[0113] 可通过在 DNA 水平引入突变来发展文库,例如,使用易错 PCR (Cadwell et al. (1992) PCR METHODS APPL. 2 :28) 或通过对 DNA 进行体外同源重组 (Stemmer (1994) PROC. NATL. ACAD. SCI. USA 91 :10747 ;Stemmer (1994) NATURE 370 :389) 或通过盒诱变来进行。例如在美国专利申请公开号 No. US2004/0180412 中描述了模板发展和发展性合成。

[0114] 下述实施例含有重要的额外信息、示例和指导,其可用于在本发明的多种实施方式或其等同方式中实践本发明。将从下面这些实施例更好地理解本发明的实践,但实施例仅作阐述目的,而不应被理解为以任何方式进行限制。

[0115] 实施例

[0116] 下述实施例证实了序列程序化官能团转化的可行性。实施例 1 和 2 描述了三种序列程序化官能团转化,即,从叠氮化物到胺、从叠氮化物到硫醇以及从叠氮化物到羧酸的转化,其中,已通过凝胶电泳 (实施例 1) 或通过质谱 (实施例 2) 对转化的终产物进行了分析。实施例 3 示出,可使用小分子试剂,例如不连接到 DNA 的磺酰氯、氯甲酸酯、异氰酸酯以及异硫氰酸酯反应物,将连接胺的模板转化为磺胺、氨基甲酸酯、尿素或硫脲。实施例 4 示出,可在单种溶液中,将有机叠氮化物的混合物以序列特异性方式转化为胺中间产物,然后使用不连接到 DNA 的游离反应物 (例如,磺酰氯、氯甲酸酯、异氰酸酯和异硫氰酸酯) 将胺中间产物以序列特异性方式转化为磺胺、氨基甲酸酯、尿素和硫脲产物。

[0117] 实施例 1 以 DNA 为模板将叠氮化物转化为一级胺、羧酸和硫醇 (通过 PAGE 分析)

[0118] 本实施例描述序列程序化官能团转化,其中,可将叠氮化物特异性转化为胺、硫醇或羧酸。各反应方案和得到的反应产率示于图 5A 和 5B 中。

[0119] I. 材料和方法

[0120] (i) 合成叠氮酸

[0121] 按下文所述,从相应的羧酸前体来制备用于合成图 5B 所示的化合物 1-12 的叠氮底物:

[0122] 叠氮乙酸(用于生产图 5B 的模板 1)。按照 Lundquist et al. (2001)ORG LETT. 3: 781 所述来生产该试剂。该产物被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 3.96(2H, s)。

[0123] 叠氮-3-甲基戊酸(用于生产图 5B 的模板 2)。按照上述 Lundquist et al. (2001) 所述来生产该试剂。该产物被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 3.92(1H, br), 3.88(2H, dd, J = 5.7Hz, J = 9Hz), 1.77(2H, m), 0.99(6H, t, J = 6.6Hz)。

[0124] 4-叠氮甲基苯甲酸(用于制备图 5B 的模板 3)。将叠氮化钠(1.3g, 20mmol)和 18-冠-6 醚(0.2mL, 1mmol)溶解于 DMSO(4mL)中。向得到的溶液中加入 4-氯甲基苯甲酸(1.71g, 10mmol),在 25℃对反应混合物进行 12 小时的搅拌。在 EtOAc 中稀释反应体系,用 0.1N HCl(2x)洗涤,然后用盐水洗。用 Na₂SO₄ 干燥有机层,浓缩以提供白色固体(1.75g, 定量的)。得到的产物被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 8.15(2H, d, J = 8.4Hz) 7.45(2H, d, J = 8.4Hz)(1H, s);针对 C₈H₆N₃O₂ 计算的 ESMS:176.0460;观察到的:176.0461。

[0125] 按照上述 Lundquist et al. (2001) 中描述来通过重氮转移合成叠氮酸的方法,合成 1-叠氮环己基羧酸(用于制备图 5B 的模板 5)和 叠氮异谷氨酸(用于制备图 5B 的模板 4)。1-叠氮环己基羧酸被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 1.86(4H, m) 1.63(4H, m) 1.36(2H, m);针对 C₇H₇N₃O₂(M+NH₄⁺) 计算的 CIMS:1187.1195;观察到的:1187.1188。叠氮异谷氨酸被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 6.43(2H, d, J = 17.4Hz) 3.14(1Hdd, J = 14.1Hz, J = 6.9Hz) 2.54(2H, tJ = 7.5Hz) 2.23(2H, dd, J = 13.6HzJ = 6.6Hz);针对 C₇H₇N₃O₂(M+NH₄⁺) 计算的 CIMS:190.0931;观察到的:190.040。

[0126] 1-叠氮甲基苯甲酸(用于生产图 5B 的模板 6)。按照 Wada et al. (2001) TETRAHEDRON LETT. 42:1069-72 以及还按照 Love et al. (2001). J. ORG CHEM. 66:68165-76 所述来生产该试剂。该产物被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 8.18(1H, dd, J = 7.6Hz, J = 1.2Hz), 7.63(1H, td, J = 7.5Hz, J = 1.2Hz), 7.55(1H, d, J = 6.6Hz), 7.46(1H, td, J = 7.5Hz, J = 1.5Hz), 4.894(2H, s)。

[0127] 4-叠氮苯甲酸(用于生产图 5B 的模板 7)。该试剂购自 Sigma-Aldrich(St. Louis, MO)。

[0128] 4-叠氮苯甲基-环己基二羧酸单酯(用于制备图 5B 的模板 8)。将反-环己基二羧酸(200mg, 1.16mmol)、EDC(223mg, 2.32mmol)和 N,N'-二异丙基乙胺(0.4mL, 2.32mmol)溶解于 CH₂Cl₂(4mL)中,在 25℃搅拌 30 分钟。向该混合物中加入 4-叠氮苯甲醇(86.6mg, 0.58mmol)。在 25℃对反应体系进行 12 小时的搅拌。浓缩反应混合物,通过快速色谱纯化(30% EtOAc/己烷)。想要的酯作为黄色固体获得(18.2mg, 5%)。得到的产物被发现具有如下特征:¹H NMR(300MHz, CDCl₃) δ 7.34(2H, d, J = 8.4Hz) 7.03(2H, d, J = 8.4Hz) 5.08(1H, s) 2.33(2H, m) 2.09(4H, d, J = 9.3Hz) 1.47(4H, t, J = 9.9Hz);针对 C₁₅H₁₇N₃O₄(M+HCO₂⁻) 计算的 ESMS:348.1196;观察到的:348.1195。

[0129] 4-叠氮苯甲基-琥珀酸单酯(用于制备图 5B 的模板 9)。将 4-叠氮苯甲醇(100mg, 0.67mmol)、琥珀酸酐(134mg, 1.37mmol)和 N,N'-甲基氨基吡啶(3.7mg, 30 μmol)溶解于 DMF(1mL)中,在 25℃搅拌 12 小时。浓缩反应混合物,通过快速色谱纯化(30% EtOAc/己

烷)。想要的酯作为黄色固体获得 (75.9mg, 45%)。得到的产物被发现具有如下特征：¹H NMR(300MHz, CDCl₃) δ 7.28(2H, d, J = 7.8Hz) 6.96(2H, d, J = 7.8Hz), 5.05(2H, s), 2.63(4H, m); 针对 C₇H₆N₃O₂ 计算的 ESMS :248.0672; 观察到的 :248.0660。

[0130] 4-叠氮苯甲基-联苯甲酸单酯 (用于制备图 5B 的模板 10)。将 4-叠氮苯甲醇 (112mg, 0.5mmol) 和联苯甲酸酐 (74.5mg, 0.5mmol) 溶解于吡啶 (1mL) 中, 在 25℃ 搅拌 12 小时。将反应混合物稀释于 EtOAc 中, 用磷酸缓冲液 (pH6.0, 2x) 洗, 然后用盐水洗。在 Na₂SO₄ 上干燥有机层, 真空下浓缩。通过快速色谱 (25% EtOAc/ 己烷) 纯化粗产物。想要的酯作为黄色固体获得 (193mg, 99%)。得到的产物被发现具有如下特征：¹H NMR(300MHz, CDCl₃) δ 9.87(1H, s) 7.92(1H, dd, J = 7.2Hz, J = 1.2Hz) 7.86(1H, dd, J = 7.8Hz, J = 1.2Hz) 7.43(2H, dd, J = 5.7Hz, J = 1.5Hz) 7.38(2H, dd, J = 5.7Hz, J = 1.2Hz) 7.32(2H, dd, J = 7.5Hz, J = 1.2Hz) 7.28(2H, dd, J = 7.3Hz, J = 1.2Hz) 6.98(2H, J = 8.4Hz) 6.85(2H, J = 8.4Hz) 4.91(2H, J = 2.7Hz); 针对 C₂₁H₁₆N₃O₄ (M+H⁺) 计算的 ESMS :374.1141; 观察到的 :374.1149。

[0131] 1-叠氮甲基苯甲酰硫代乙酸硫酯 (用于制备图 5B 的模板 11)。将 2-叠氮甲基苯甲酸 (40mg, 0.23mmol) 与 EDC (64.9mg, 0.34mmol) 和 N-羟基琥珀酰亚胺 (NHS) (39.1mg, 0.34mmol) 在 CH₂Cl₂ 中于 25℃ 混合 2 小时。用 NaHCO₃ (2x) 洗反应混合物, 然后用盐水洗。浓缩有机层, 粗产物不经进一步纯化即直接用于下一步。令 2-叠氮甲基苯甲酸 N-羟基琥珀酰亚胺酯 (16.4mg, 47 μmol) 和 DMF (250 μL) 中的硫代乙酸 (3.2 μL) 在 25℃ 反应 24 小时。在 EtOAc 中稀释反应混合物, 用 NaHCO₃ (2x) 洗, 然后用盐水洗。在 Na₂SO₄ 上干燥有机层, 真空下浓缩。通过快速色谱 (30% EtOAc/ 己烷) 纯化粗制混合物, 以提供硫酯 (9.8mg, 83%)。得到的产物被发现具有如下特征：¹H NMR(300MHz, CDCl₃) δ 7.95(d, 1H, J = 7.8Hz) 7.56(t, 1H, J = 7.8Hz) 7.50(d, 1H, J = 6.6Hz) 7.41(t, 1H, J = 7.8Hz) 4.64(s, 2H) 3.88(s, 2H); 针对 C₇H₆N₃O₂ 计算的 ESMS :250.0287; 观察到的 :250.0284。

[0132] 2-叠氮甲基苯甲酰硫丙酸硫酯 (用于制备图 5B 的模板 12)。通过将等体积的 1-叠氮甲基苯甲酸 (用于生产图 5B 中的模板 6) (900mM, DMF 中)、EDC (900mM, DMF 中) 和 NHS (900mM, DMF 中) 在 25℃ 混合 1 小时, 来制备 2-叠氮甲基苯甲酰 N-羟基琥珀酰亚胺 (NHS) 酯。在平行制备中, 将硫醇基团附着到 DNA 寡核苷酸上, 在模板形成时将其并入模板 (见下文, 制备 5' 连接 2-叠氮甲基苯甲酰硫丙酸硫酯的 DNA)。

[0133] (ii) 制备官能化寡核苷酸

[0134] 在本实施例以及下述实施例中, 使用标准的亚磷酰胺 (phosphoramidite) 方案, 在 Perseptive Biosystems Expedite8090DNA 合成仪上合成寡核苷酸, 并使用制备规模的反相 HPLC 进行纯化。用于自动固相寡核苷酸合成的试剂购自 Glen Research。针对下述以胺结尾的、生物素化的 DNA 寡核苷酸, 用 5' 氨基-修饰物 (modifier) 5 (Glen Research) 制备 5'-氨基经修饰的寡核苷酸; 用 3' 氨基-修饰物 C7CPG (Glen Research) 制备 3'-氨基经修饰的寡核苷酸; 以及, 用生物素 TEG CPG (Glen Research) 制备 3'-经生物素标记的寡核苷酸。通过分析规模的反相 HPLC 来纯化官能化的 DNA 寡核苷酸。

[0135] 基于在 Hewlett-Packard8453UV- 可见光分光光度计 (Agilent Technologies) 上测量到的 260nm 吸收, 测定溶液中经纯化的寡核苷酸的浓度。通过 UV 透视法, 以及使用 Eagle Eye II densitometer (Stratagene) 进行密度测定, 来观察和定量用溴化乙啶染色

的寡核苷酸。

[0136] (a) 模板寡核苷酸

[0137] 5' 连接叠氮化物的 DNA 寡核苷酸模板 (用于生产图 5B 的模板 1-11)。将等体积的各叠氮酸 (900mM, DMF 中)、EDC (900mM, DMF 中) 和 NHS (900mM, DMF 中) 在 25℃ 混合 1 小时, 来制备想要的叠氮酸的 N- 羟基琥珀酰亚胺 (NHS) 酯。将粗制的 NHS 酯以两部分 (每份 50 μ L) 加入到含有处于 100mM 磷酸钠缓冲液 (pH7.2, 350 μ L) 中的 5' 氨基经修饰的 DNA 寡核苷酸 (50 μ L, 典型地, 300 μ M) 的溶液中。用于这些制备中所用的 30 体模板是 5' NH₂(C₂H₄O)₂-PO₃H-GGT ACG AAT TGC ACT CGG GAA ATC CAC CTT (SEQ ID NO: 1)。偶联反应在 25℃ 进行 1 小时。将得到的反应混合物直接上样到 NAP-5 尺寸排除柱 (Amersham Biosciences) 上, 去除有机溶剂、盐和过量的小分子, 通过分析规模的反相 HPLC (8-30% MeCN/0.1M TEAA 梯度) 进一步纯化连接叠氮化物的 DNA 寡核苷酸。通过 MALDI-TOF 质谱分析想要的寡核苷酸产物。

[0138] 5' 连接 2-叠氮甲基苯甲酰硫丙酸硫酯的 DNA (用于生产图 5B 的模板 12)。将 DMF 中的 2,2'-二硫二丙酸 (900mM) 溶液与等体积的 EDC (900mM, DMF 中) 和 NHS (900mM, DMF 中) 在 25℃ 混合 1 小时。将粗制的 NHS 酯加入到处于 100mM 磷酸钠缓冲液 (pH7.2, 350 μ L) 中的含有 5' 氨基经修饰的 DNA 寡核苷酸 (50 μ L, 典型地, 300 μ M) 的溶液中。偶联反应在 25℃ 进行 1 小时。将得到的反应混合物直接上样到 NAP-5 尺寸排除柱 (Amersham Biosciences) 上, 通过分析规模的反相 HPLC (8-30% MeCN/0.1M TEAA 梯度) 对其进行进一步纯化。通过 MALDI-TOF 质谱分析连接二硫化物的寡核苷酸产物。通过在 25℃ 用 20mM DTT 对 100mM CAPS 缓冲液 (pH8) 中的上述连接二硫化物的寡核苷酸 (典型地, 10 μ M) 处理 0.5 小时, 来制备连接 2-硫丙酸的寡核苷酸。通过将反应混合物通过凝胶过滤柱来除去过量的 DTT。平行地, 按上文所述来制备 2-叠氮甲基苯甲酰 N-羟基琥珀酰亚胺 (NHS) 酯 (见上文对 2-叠氮甲基苯甲酰硫丙酸硫酯的制备)。将粗制的 NHS 酯 (100 μ L) 加入到处于 100mM 磷酸钠缓冲液 (pH7.2, 300 μ L) 中的 5' 连接硫醇的寡核苷酸 (100 μ L) 的溶液中。偶联反应在 25℃ 进行 1 小时。将得到的反应混合物直接上样到 NAP-5 尺寸排除柱上, 通过分析规模的反相 HPLC (8-30% MeCN/0.1M TEAA 梯度) 进一步纯化。通过 MALDI-TOF 质谱分析想要的寡核苷酸产物。

[0139] (b) 转移单元

[0140] 3' 连接三苯基膦的 DNA。在连接到 CPG 树脂上的 3' 氨基经修饰的寡核苷酸上进行三苯基膦基团的附着。与模板完全互补的 10 体试剂具有下述结构: 5' AAT TCG TACC-OP₃H-CH₂CH(CH₂OH)(CH₂)₄NHCOC₆H₄PPh₂ (SEQ ID NO: 2)。含有相对模板而言三碱基错配的 10 体试剂具有下述结构: 5' AAT ACA TCC C-OP₃H-CH₂CH(CH₂OH)(CH₂)₄NHCOC₆H₄PPh₂ (SEQ ID NO: 3)。后者在这些实验中用作为对照。

[0141] 通过三个循环的下述过程: (i) 用 DMF 中的 20% 哌啶处理 10 分钟; (2) 用 DMF 洗; 以及 (iii) 用 MeCN 洗, 去除 3' Fmoc-NH- 寡核苷酸上的 Fmoc 基团。在氮气流下干燥树脂。将 4-二苯基膦基苯甲酸 (30.6mg, 100 μ mol)、EDC (19.1mg, 100 μ mol)、N,N-二异丙基乙胺 (36.8 μ L, 211 μ mol) 的处于 DMF (0.6mL) 中的溶液加入树脂, 混合物在 37℃ 温育 2 小时。用 DMF (2x) 以及用 MeCN (2x) 洗树脂, 然后在氮气下干燥。从 CPG 树脂上切割下经衍生化的寡核苷酸, 这通过在 1:1 的羟基铵: 甲基胺 (AMA) (含有盐酸三 (2-羧乙基) 膦 (TCEP-HCl,

1mg)) 中于 55℃温育 45 分钟来实现。过滤切割溶液,通过分析规模的反相 HPLC(8-30% MeCN/0.1M TEAA 梯度)对其进行纯化。通过 MALDI-TOF 质谱分析想要的寡核苷酸产物。

[0142] (c) 捕获试剂

[0143] 为了进行聚丙烯酰胺凝胶电泳,使用退火到模板上的 20 体第二试剂(捕获试剂)来捕获反应产物。

[0144] 氨基经修饰的 DNA(该试剂用于捕获模板 8-10 的产物)。20 体第二试剂含有序列: 5' TCC CGA GTG CAA TTC GTA CC- $\text{OPO}_3\text{H}-\text{CH}_2\text{CH}(\text{CH}_2\text{OH})(\text{CH}_2)_4\text{NH}_2$ (SEQ ID NO:4)。该寡核苷酸用作用于下述捕获试剂的起始材料。

[0145] 3' 连接溴乙酸酯的 DNA(该试剂用于捕获模板 11 和 12 的产物)。通过将等体积的溴乙酸(900mM, DMF 中)、EDC(900mM, DMF 中)和 NHS(900mM, DMF 中)在 25℃混合 1 小时,来制备溴乙酸的 NHS 酯。将粗制的 NHS 酯(100 μL)加入到 5' 氨基经修饰的 DNA 寡核苷酸(50 μL , 典型地, 300 μM) 在 100mM 磷酸钠缓冲液(pH7.2, 350 μL) 中的溶液中。偶联反应在 25℃进行 1 小时。将得到的反应混合物直接上样到 NAP-5 尺寸排除柱上,去除有机溶剂、盐和过量的小分子,通过分析规模的反相 HPLC(8-30% MeCN/0.1M TEAA 梯度)对其进行进一步纯化。通过 MALDI-TOF 质谱分析想要的寡核苷酸产物。

[0146] 3' 连接 4- 甲酰苯甲酸酯的 DNA(该试剂用于捕获模板 7 的产物)。按照用于连接溴乙酸酯的 DNA 的方案,制备连接 4- 甲酰苯甲酸酯的 20 体的 DNA,其中使用 4- 甲酰苯甲酸代替溴乙酸。

[0147] 3' 连接琥珀酸单酯的 DNA(该试剂用于捕获模板 1-6 的产物)。在 25℃用 DMF(200 μL) 中的 NHS(10mg, 0.1mmol) 对琥珀酸酐(22mg, 0.1mmol) 进行 15 分钟活化。将 100 μL 该混合物加入到 100mMHEPES 缓冲液(pH8.5, 850 μL) 中的 3' 氨基经修饰的模板(50 μL , 典型地, 300 μM) 中,在 37℃温育 16 小时。通过 NAP-5 尺寸排除柱对反应混合物进行脱盐,通过分析规模 HPLC(8-30% MeCN/0.1M TEAA 梯度)对其进行进一步纯化。通过 MALDI-TOF 质谱分析想要的寡核苷酸产物。

[0148] II. 结果和结论

[0149] (i) 以 DNA 为模板的从叠氮化物到胺的转化

[0150] 用多种与 30 体 DNA 寡核苷酸模板 5' 末端相连的有机叠氮化物与互补的 DNA10 体 3' 末端接合的三苯基膦反应(见图 5A)。通过将 5' 连接叠氮化物的 30 体模板(12pmol) 和 3' 连接三苯基膦的 10 体试剂(24pmol)(在总体积为 200 μL 的 100mM CAPS 缓冲液(pH10) 中,其中含有 500mM NaCl) 在 25℃混合 16 小时,来进行以 DNA 为模板的从叠氮化物到胺的官能团转化。对于底物 4 和 5, 1M NaCl, 加入 0.5mM DTT 以抑制膦氧化被发现能增加产率。代表性反应条件包括:对 1-7 而言, 60nM 叠氮化物, 120nM 膦, 0.1M CAPS(pH10), 0.5M NaCl; 对 8-11 而言, 同上, 差别在于 0.1M MES(pH6.0), 1M NaCl; 对 12 而言, 同上, 差别在于 0.1M MOPS(pH7.5), 1M NaCl。

[0151] 与以 DNA 为模板的偶联反应不同,从叠氮化物到胺的转化可能不能被变性聚丙烯酰胺凝胶电泳(PAGE) 直接监测,因为起始材料和产物具有相近的分子量。为检测这些反应的进程,在存在碳二亚胺的情况下用连接 20 体的羧酸,在存在 NaBH_3CN 的情况下用连接 20 体的醛来捕获可能的胺产物。这些第二试剂或捕获试剂置换连接 10 体的氧化膦,高效偶联一级胺,而不会偶联叠氮化物。得到的酰胺或二级胺产物获得了 20 体的分子量,易于通过

PAGE 与起始的叠氮化物区分开。

[0152] 为了捕获从底物 1-6 获得的胺产物,向反应混合物(含有处于 MES 缓冲液(pH6.5)中的 EDC(30mM)和硫-NHS(15mM))中加入 3'连接羧酸的 20 体试剂(24pmol)。为了捕获从底物 7 获得的胺产物,在存在处于 MES 缓冲液(pH6.5)中的 NaBH_3CN (3mM)时,用 3'连接醛的 20 体试剂来捕获产物。产物捕获之后,用 NaOAc(pH5)、乙醇和糖原(glycogen)来沉淀连接 DNA 的物质。

[0153] 将得到的沉淀溶解于变性凝胶上样缓冲液中,进行变性 PAGE 分析。如无特别指明,用 15%聚丙烯酰胺凝胶(TBE-尿素)来进行 PAGE 分析。

[0154] 通过凝胶的溴化乙啶染色、产物和模板条带的 UV 显示以及基于 CCD 的密度测定,来定量反应产率。产率计算假设变性凝胶中的产物和模板按照每碱基等密度的方式被染色。产物在定量期间是部分双链的情况下,染色密度的改变可能导致更高的表观产率。

[0155] 通过变性 PAGE 分析获得典型结果示于图 6。图 6A 示出了对图 5B 中叠氮化物 3 的以 DNA 为模板的、从叠氮化物到胺的转化的变性 PAGE 分析。图 6B 示出了对图 5B 中叠氮化物 7 的以 DNA 为模板的、从叠氮化物到胺的转化的变性 PAGE 分析。

[0156] 对测试的七种叠氮化物(图 5B 中的底物 1-7)而言,以 DNA 为模板的叠氮化物还原在 pH10 时高效进行。反应产物的真实产率概括于图 5B 中。在每种情况下,其中磷与非互补的错配寡核苷酸相连的对照反应不产生显著的酰胺或二级胺产物,这表明这些以 DNA 为模板的从叠氮化物到胺的转化以序列特异性的方式进行。

[0157] (ii) 以 DNA 为模板的从叠氮化物到羧酸或硫醇的转化

[0158] 反应范围被进一步扩展到实现从叠氮化物到羧酸以及从叠氮化物到硫醇的官能团转化(见图 5A)。在两种情况下,叠氮化物还原诱发了自发片段化,以暴露羧酸或硫醇基团。为评估这些反应的效率,在存在碳二亚胺的情况下,用连接 DNA 的胺来捕获羧酸(从底物 8-10 获得的产物),而连接 DNA 的烷基溴则被用于捕获硫醇产物(从底物 11 和 12 获得的产物)。

[0159] 以 DNA 为模板的从叠氮化物到羧酸的转化的进行类似从叠氮化物到胺的转化,差别在于缓冲液含有 0.1M MES(pH6.0)和 1M NaCl。为捕获羧酸产物,向反应混合物(含有处于 MES 缓冲液(pH6.5)中的 EDC(30mM)和硫-NHS(15mM))中加入 3'连接胺的 20 体试剂。来自变性 PAGE 的典型结果示于图 6C(代表为使用来自图 5B 的试剂 8 的情况)。

[0160] 以 DNA 为模板的从叠氮化物到硫醇的转化按上文进行,差别在于缓冲液含有 0.1M MES(pH6.0)(对底物 11 而言)或 MOPS(pH7.5)(对底物 12 而言)和 1M NaCl。为捕获硫醇产物,向反应混合物中加入 3'连接烷基溴的 20 体试剂,在 37°C 温育 6 小时。来自变性 PAGE 的典型结果示于图 6D(代表为使用来自图 5B 的试剂 11 的情况)。

[0161] 对于从羧酸到胺以及从硫醇到胺的转化而言,变性 PAGE 分析表明,暴露羧酸和硫醇基团(图 5B 中的底物 8-12)的以 DNA 为模板的官能团转化也以高效且序列特异性的方式进行。

[0162] 实施例 2 以 DNA 为模板将叠氮化物转化为一级胺、羧酸和硫醇(通过质谱分析)

[0163] 本实施例类似实施例 1,除了通过质谱而非 PAGE 来分析反应产物。为方便起见,在同样的或相似条件下,采用更小的模板以及不同的捕获系统。

[0164] I. 材料和方法

[0165] (i) 合成叠氮酸

[0166] 按照实施例 1 所述,来制备用于合成图 5B 所示的化合物 1-12 的叠氮底物。

[0167] (ii) 制备经官能化的寡核苷酸

[0168] 以与实施例 1 类似的方式来制备用于本实施例的寡核苷酸,其中具有如下改变。

[0169] (a) 模板寡核苷酸

[0170] 按照实施例 1 所述来制备模板寡核苷酸,区别在于不使用 30 体模板,而使用下述 10 体模板:5' -NH₂(C₂H₄O)₂-PO₃H-GGT ACG AAT T-OP₃H-CH(CH₂OH)CH₂(OC₂H₄)₄CH₂NHCO- 生物素 (SEQ ID NO :5)。

[0171] (b) 转移单元

[0172] 按照实施例 1 所述制备连接三苯基磷的试剂。

[0173] (iii) 质谱分析

[0174] 在 Applied Biosystems Voyager-DE Pro Biospectrometry Workstation 上进行 MALDI-TOF 质谱,用 Voyager Data Explorer 软件进行处理。在所有实验中,九份羟基吡啶甲酸 (HPA, 50mg/mL, 处于 50% MeCN/H₂O 中) 和一份柠檬酸铵 (50mg/mL, 处于 H₂O 中) 的混合物被用作作为基质。

[0175] II. 结果和结论

[0176] 向含有 500mM NaCl 的 100mM CAPS 缓冲液 (pH10) 中的 5' 连接叠氮化物、3' 生物素化的 10 体模板 (12pmol) 溶液中加入连接互补性 DNA 的磷试剂 (24pmol)。在 25℃ 对混合物进行 0.5 小时的搅动,然后在 37℃ 进行 12 小时搅动。通过用连接链亲和素磁性颗粒 (Roche) 处理反应混合物来纯化经生物素化的产物和未反应的模板,按照厂商方案进行洗脱。用乙醇和糖原沉淀洗脱液中的 DNA。对底物 11-12 直接进行后续质谱分析。通过使用 ZipTip C18 柱 (Millipore) 对溶解于基质溶液中的沉淀脱盐,来制备用于 MALDI-TOF 分析的样品。

[0177] 通过 MALDI-TOF 质谱鉴定得到的亚胺基正磷 (iminophosphorane),发现其出人意料地对水解稳定,尤其是在酸性条件下,推测这是由于形成了稳定的 HCl 盐 (Shalev, et al. (1996) J. ORG. CHEM. 61 :1689-1701)。但是,用连接 DNA 的磷在 pH10 的缓冲液中于 25℃ 对连接模板的叠氮化物的处理却产生了定量的亚胺基正磷水解,产生了相应的一级胺。

[0178] 来自 MALDI-TOF 分析的结果概括于表 1 中,其中,试剂 1-12 如图 5B 所指示的。由于在用于 MALDI-TOF 实验的条件下不稳定,通过在以 DNA 为模板的 Staudinger 反应之后,用碘乙酰胺 (5mM) 处理,捕获作为烷基硫醚加合物 (adduct) 的连接硫醇的产物 (11-12) (表 1 中针对 11-12 的 MALDI-TOF 数据是被捕获的硫醚加合物的)。

[0179] 表 1

[0180]

试剂 (见图 5B)	预期质量	观察到的质量
1	5866.05	5868.02 ± 9
2	5922.16	5926.50 ± 9
3	5942.15	5945.18 ± 9
4	5937.13	5940.98 ± 9
5	5934.17	5934.61 ± 9
6	5942.15	5944.46 ± 9

[0181]

7	5928.12	5930.91±9
8	5963.16	5968.42±9
9	5909.07	5915.97±9
10	6032.23	6038.24±9
11	5941.16	5944.61±9
12	5955.19	5957.01±9

[0182] 每种情况下,对叠氮化物还原反应的质谱分析都与预期的一级胺产物形成一致。代表性谱示于图 7 中。

[0183] 基于表 1 所示的质谱分析,序列特异性的从叠氮化物到胺、从叠氮化物到羧酸以及从叠氮化物到硫醇的转化都产生了正确的产物。

[0184] 实施例 3 使用小分子试剂来转化连接胺的模板

[0185] 为进一步探索以 DNA 为模板的官能团转化使不连接 DNA 的试剂参与序列程序化合成的能力,制备了四种 DNA 模板(模板 13-16,图 8),每种都含有:处于 5' 末端的不同的叠氮化物,四种独特的六碱基密码子之一以及在 3' 末端的生物素基团,以便于模板操作和纯化。然后将连接叠氮化物的模板化学转化为连接胺的模板,这通过暴露给 TCEP-HCl 来实现。然后用游离反应物与得到的胺反应,以确定胺中间产物向终产物的转化是否可行。特别地,丹磺酰氯 (21)、氯甲酸乙酯 (22)、4- 甲氧基苯基异氰酸酯 (23) 和 6- 吗啉代吡啶基 3- 异硫氰酸酯 (24) 全部选用为胺反应试剂,因为它们不能容易地附着到 DNA 上(由于其结构或其与水的反应活性)。图 9 示出了简单化的反应方案,其中示出了起始试剂和理论终产物。

[0186] I. 材料和方法

[0187] 通过计算机筛选,设计用于模板的编码序列,以 (i) 确保任何两个不同密码子之间至少存在六个非互补的碱基对, (ii) 保持每密码子的恒定 % GC,以最小化试剂间熔点的区别,以及 (iii) 质量上有所变化,使得 16 个理论小分子偶联产物的分子量不同,并可被 MALDI-TOF 质谱鉴定出来。

[0188] 用于图 8 所示的方案中的每种模板(模板 13-16)都含有 5' NH₂(C₂H₄O)₂-PO₃H-TT-(密码子)-GTA_n-OP₃H-CH(CH₂OH)CH₂(OC₂H₄)₄CH₂NHCO-生物素。用于每种模板的密码子如下所示:用于模板 13 的密码子是 GTG CAA CGT CAT, n = 0 (SEQ ID NO :6);用于模板 14 的密码子是 CCT AGT CGT CAT, n = 3 (SEQ ID NO :7);用于模板 15 的密码子是 TAA GCC CGT CAT, n = 2 (SEQ ID NO :8);用于模板 16 的密码子是 AGC TTG CGT CAT, n = 1 (SEQ ID NO :9)。

[0189] 按照实施例 1 所述,制备含有叠氮化物的模板 13-16(分别见模板 1、3、4 和 2)。

[0190] 通过暴露给 TCEP-HCl,将模板 13-16 中的叠氮化物基团化学转化为氨基。简言之,通过用 100mM MOPS 缓冲液 (pH7.5) 中的 5mM TCEP-HCl 在 25℃ 对连接叠氮化物的模板(模板 13-16) 进行 3 小时处理,由此来制备连接胺的模板。通过 HPLC 来纯化得到的模板。之后,用可溶试剂与得到的模板(胺中间产物)反应,以确定官能团转化是否可行。

[0191] II. 结果和结论

[0192] 一旦制造出了连接胺的模板 13-16,然后将它们暴露给可溶试剂,以观察官能团转化是否可行。每种转化都在下文中详细讨论。

[0193] 将 100 μL 100mM 水性 NaHCO₃ (pH9.0) 中的连接胺的模板 13 (400pmol) 与 100 μL DMF 中的 20mM 丹磺酰氯 21 混合,在 37℃ 搅动 1 小时。在 200 μL 0.1M TEAA 中稀释反应混

合物,使其经过 NAP-5 尺寸排除柱。通过分析规模的反相 HPLC(8-30% MeCN/0.1M TEAA 梯度)来分析 1mL 0.1M TEAA 中的洗脱液。基于起始材料、产物和任何副产物的积分峰面积(基于 260nm 的 UV 吸收)来计算产物产率。代表性色谱示于图 10A 中。

[0194] 将 100 μ L 200mM 水性 NaHCO_3 (pH9.0) 中的连接胺的模板 14(400pmol) 与 100 μ L DMF 中的 40mM 氯甲酸乙酯 22 混合,在 37°C 搅动 1 小时。通过加入 NaOAc 缓冲液 (pH5.0) 中的糖原来淬灭反应,接着进行乙醇沉淀。将沉淀物溶解于 0.1M TEAA 中,通过分析规模的反相 HPLC(8-30% MeCN/0.1M TEAA 梯度)来分析。代表性色谱示于图 10B 中。

[0195] 将 100 μ L 500mM 水性三乙胺 (pH10) 中的连接胺的模板 15(400pmol) 与 100 μ L DMF 中的 20mM 4-甲氧基苯基异氰酸酯 23 混合,在 37°C 搅动 1 小时。按上文所述淬灭反应混合物,并对其进行分析。代表性色谱示于图 10C 中。

[0196] 将 100 μ L 500mM 水性三乙胺 (pH10) 中的连接胺的模板 16(400pmol) 与 100 μ L DMF 中的 20mM 6-吗啉代吡啶基 3-异硫氰酸酯 24 混合,并允许它们反应,在 37°C 搅动 1 小时。按上文所述淬灭反应混合物,并对其进行分析。代表性色谱示于图 10D 中。

[0197] 当在碱性条件下 (pH9-10),将可溶试剂(游离反应物)21、22、23 或 24 在 DMF 中过量(10 或 20mM 的最终浓度)加入到连接模板的一级胺中时,高效产生相应的磺胺、氨基甲酸酯、尿素或硫脲(对 21 而言,70%的产率,对 22、23 和 24 而言,产率 >86%)。这些结果证实:可使用游离反应物,将与酰胺偶联的模板转化为磺胺、氨基甲酸酯、尿素或硫脲。

[0198] 实施例 4 使用小分子试剂对四种连接叠氮化物的模板进行序列特异性转化

[0199] 本实施例证实,可对结合模板的反应物进行序列特异性转化,产生反应中间产物,反应中间产物接着可与游离反应物反应,产生反应产物。特别地,多种连接叠氮化物的模板的单种溶液混合物被序列特异性地转化为胺中间产物。然后使用不与 DNA 相连的磺酰氯、氯甲酸酯、异氰酸酯和异硫氰酸酯,将胺中间产物序列特异性地修饰为磺胺、氨基甲酸酯、尿素或硫脲产物。

[0200] I. 材料和方法

[0201] (i) 制备官能化的寡核苷酸

[0202] (a) 模板寡核苷酸

[0203] 按照实施例 3 所述,制备模板 13-16。

[0204] (b) 转移单元

[0205] 按照实施例 1 所述,制备下述连接三苯基膦的寡核苷酸中的每种。

[0206] 寡核苷酸 17(图 9)具有结构 5'-CGT TGC ACA A- $\text{OPO}_3\text{H}-\text{CH}_2\text{CH}(\text{CH}_2\text{OH})(\text{CH}_2)_4\text{NHCOC}_6\text{H}_4\text{PPh}_2$ (SEQ ID NO:10)。寡核苷酸 18(图 9)具有结构 5'-CGACTAGGAA- $\text{OPO}_3\text{H}-\text{CH}_2\text{CH}(\text{CH}_2\text{OH})(\text{CH}_2)_4\text{NHCOC}_6\text{H}_4\text{PPh}_2$ (SEQ ID NO:11)。寡核苷酸 19(图 9)具有结构 5'-CGGGCT TAAA- $\text{OPO}_3\text{H}-\text{CH}_2\text{CH}(\text{CH}_2\text{OH})(\text{CH}_2)_4\text{NHCOC}_6\text{H}_4\text{PPh}_2$ (SEQ ID NO:12)。寡核苷酸 20(图 9)具有结构 5'-CGCAAGCTAA- $\text{OPO}_3\text{H}-\text{CH}_2\text{CH}(\text{CH}_2\text{OH})(\text{CH}_2)_4\text{NHCOC}_6\text{H}_4\text{PPh}_2$ (SEQ ID NO:13)。

[0207] II. 结果和结论

[0208] 将模板 13-16 的混合物与序列特异性反应物 17 混合,然后与游离反应物 21 混合。类似地,得到的溶液与序列特异性反应物 18 混合,然后与游离反应物 22 混合;与序列特异性反应物 19 混合,然后与游离反应物 23 混合;与序列特异性反应物 20 混合,然后与游离反

应物 24 混合。

[0209] 更特别地,将 3' 连接三苯基膦的寡核苷酸 17(8 当量的)加入到处于 100mM CAPS 缓冲液 (pH10) 和 500mM NaCl 中的四种 5' 连接叠氮化物的模板 (模板 13-16, 每种模板 100nM) 的单一溶液混合物中,进行从叠氮化物到胺的转化。在 25℃对混合物温育 0.5 小时,然后在 37℃温育 12 小时。通过加入 NaOAc 缓冲液 (pH5.0) 中的糖原和乙醇,来沉淀寡核苷酸。将沉淀溶解于 100 μ L 100mM 水性 NaHCO_3 中,令其与 100 μ L DMF 中的丹磺酰氯 21 (20mM) 在 37℃反应 1 小时。通过乙醇沉淀对反应混合物脱盐。如果以 DNA 为模板的、从叠氮化物到胺的转化以序列特异性方式进行,那么应当仅有来自模板 13 的胺与序列特异性反应物 21 反应,产生磺胺 25,同时,模板 14-16 应当保持不变 (见图 8)。乙醇沉淀时除去过量的磺酰氯,使用连接 N- 羟基琥珀酰亚胺 (NHS) 酯的树脂除去任何未反应的胺。

[0210] 然后使用连接膦的寡核苷酸 18,重复上文所述的以 DNA 为模板的从叠氮化物到胺的转化。将沉淀溶解于 100 μ L 200mM 水性 NaHCO_3 中,令其与 100 μ L DMF 中的氯甲酸乙酯 22 (40mM) 在 37℃反应 1 小时。通过乙醇沉淀对反应混合物脱盐,并进行干燥。

[0211] 然后使用连接膦的寡核苷酸 19,重复上文所述的以 DNA 为模板的从叠氮化物到胺的转化。将沉淀溶解于 100 μ L 500mM 三乙胺水溶液中,令其与 100 μ L DMF 中的 4- 甲氧基苯基异氰酸酯 23 (20mM) 在 37℃反应 1 小时。通过乙醇沉淀对反应混合物脱盐,并进行干燥。

[0212] 然后使用连接膦的寡核苷酸 20,重复上文所述的以 DNA 为模板的从叠氮化物到胺的转化。将沉淀溶解于 100 μ L 500mM 三乙胺水溶液中,令其与 100 μ L DMF 中的 6- 吗啉代吡啶基 3- 异硫氰酸酯 24 (20mM) 在 37℃反应 1 小时。通过乙醇沉淀对反应混合物脱盐,并进行干燥。将沉淀溶解于 100mM MES 缓冲液 (pH6) 中,首先用 TCEP-HCl (5mM) 在 25℃处理 2 小时,然后用经 NHS 活化的树脂 (Amersham Biosciences ;对 100pmol 模板而言,5 μ L 树脂溶液) 再处理 2 小时。通过过滤除去树脂,用 0.1M TEAA 洗三次。

[0213] 通过用与磁性颗粒相连的链亲和素捕获连接模板的生物素基团,来纯化产物的最终混合物。按照厂商的方案,从颗粒洗脱下捕获的寡核苷酸。用 NaOAc (pH5.0)、糖原和乙醇沉淀洗脱液中的 DNA。通过在 260nm 处,监测起始材料库和终产物库的 UV 吸收,以分光光度方式测定 DNA 回收。含等量产物 25-28 (见图 8) 的、具有 1.0 的 260nm UV 吸收的混合物的浓度被估计为 5.5 μ M。按照实施例 2 所述来制备用于 MALDI-TOF 分析的样品。图 11A 和图 11B 分别示出了起始材料 (模板 13、14、15 和 16) 和产物 (产物 25、26、27 和 28) 的代表性谱。

[0214] MALDI-TOF 质谱揭示:终产物混合物主要含有四种序列程序化产物 (磺胺 25、氨基甲酸酯 26、尿素 27 和硫脲 28)。没有观察到 12 种可能的不想要的交叉产物 (cross-product) 中的任何一种。

[0215] UV 分光光度分析表明,对以四种连续 DNA 和小分子偶联序列为模板的还原而言,以 51% 的总产率产生终产物混合物。这些结果表明:以 DNA 为模板的官能团转化能使不连接 DNA 的小分子参与序列程序化反应。该方法的效率也强调了以这种小 (小于 1nmol) 的规模进行有机合成时成为可能的基于分子生物学的纯化和洗涤策略的价值。

[0216] 综合考虑,该实施例中描述的以 DNA 为模板的官能团转化扩展了以核酸为模板的合成的合成能力,这通过解决对与寡核苷酸联结 (tether) 的试剂的需要来实现。在试剂

与寡核苷酸的连接是不可能或不方便的情况下,这些转化仍允许此类试剂作用于小分子合成,并同时保持核酸序列与产物结构之间的对应性。

[0217] 通过引用并入本文

[0218] 为了一切目的,本文提到的每份专利文献和科学文章的全部公开内容都通过引用并入本文。

[0219] 等同物

[0220] 本发明可以在不背离其原则和重要特征的情况下以其它特定形式实施。因此,在所有方面,前述实施方式将被认为是阐述性的,而非对本文所述的本发明的限制。由此,本发明的范围由所附的权利要求而非前述说明书指出,在权利要求的等同性的范围和含义内进行的所有改变都意欲包括进本发明。

序列表

<110> 哈佛大学校长及研究员协会

戴维德·刘

樱井馨

<120> 用于以核酸为模板的合成中的游离反应物

<130>LS1-003PC

<150>US 60/646, 584

<151>2005-01-21

<160>13

<170>PatentIn version 3.3

<210>1

<211>30

<212>DNA

<213> 人工

<220>

<223>30 体寡核苷酸模板（用于通过 PAGE 的分析）

<400>1

ggtacgaatt gcactcggga aatccacctt

30

<210>2

<211>10

<212>DNA

<213> 人工

<220>

<223>10 体寡核苷酸,与磷相连

<400>2

aattcgtagc

10

<210>3

<211>10

<212>DNA

<213> 人工

<220>

<223>10 体寡核苷酸,具有三碱基错配

<400>3

aatacatccc

10

<210>4

<211>20

<212>DNA

<213> 人工

<220>

<223>20 体寡核苷酸捕获试剂

<400>4

tcccgagtgc aattcgtacc

20

<210>5

<211>10

<212>DNA

<213> 人工

<220>

<223>10 体寡核苷酸模板（用于通过质谱的分析）

<400>5

ggtacgaatt

10

<210>6

<211>14

<212>DNA

<213> 人工

<220>

<223> 模板 13 寡核苷酸

<400>6	
ttgtgcaacg tcat	14
<210>7	
<211>23	
<212>DNA	
<213> 人工	
<220>	
<223> 模板 14 寡核苷酸	
<400>7	
ttcctagtcg tcatgtagta gta	23
<210>8	
<211>20	
<212>DNA	
<213> 人工	
<220>	
<223> 模板 15 寡核苷酸	
<400>8	
tttaagcccg tcatgtagta	20
<210>9	
<211>17	
<212>DNA	
<213> 人工	
<220>	
<223> 模板 16 寡核苷酸	
<400>9	
ttagcttgcg tcatgta	17
<210>10	
<211>10	
<212>DNA	

<213> 人工

<220>

<223> 反应物 17 寡核苷酸

<400>10

cgttgcacaa

10

<210>11

<211>10

<212>DNA

<213> 人工

<220>

<223> 反应物 18 寡核苷酸

<400>11

cgactaggaa

10

<210>12

<211>10

<212>DNA

<213> 人工

<220>

<223> 反应物 19 寡核苷酸

<400>12

cgggcttaaa

10

<210>13

<211>10

<212>DNA

<213> 人工

<220>

<223> 反应物 20 寡核苷酸

<400>13

cgcaagctaa	10
------------	----

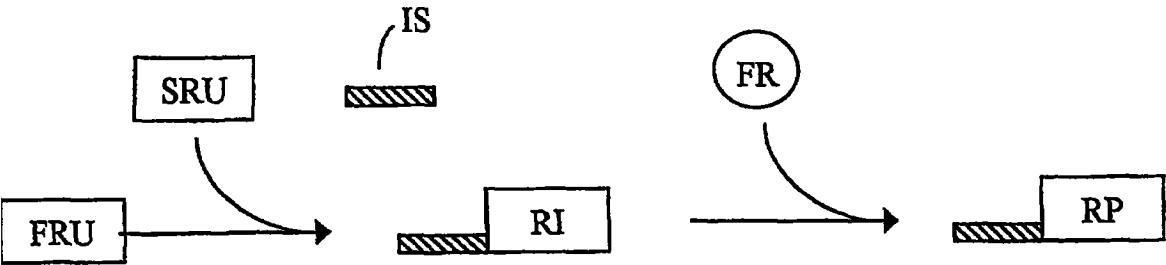


图1

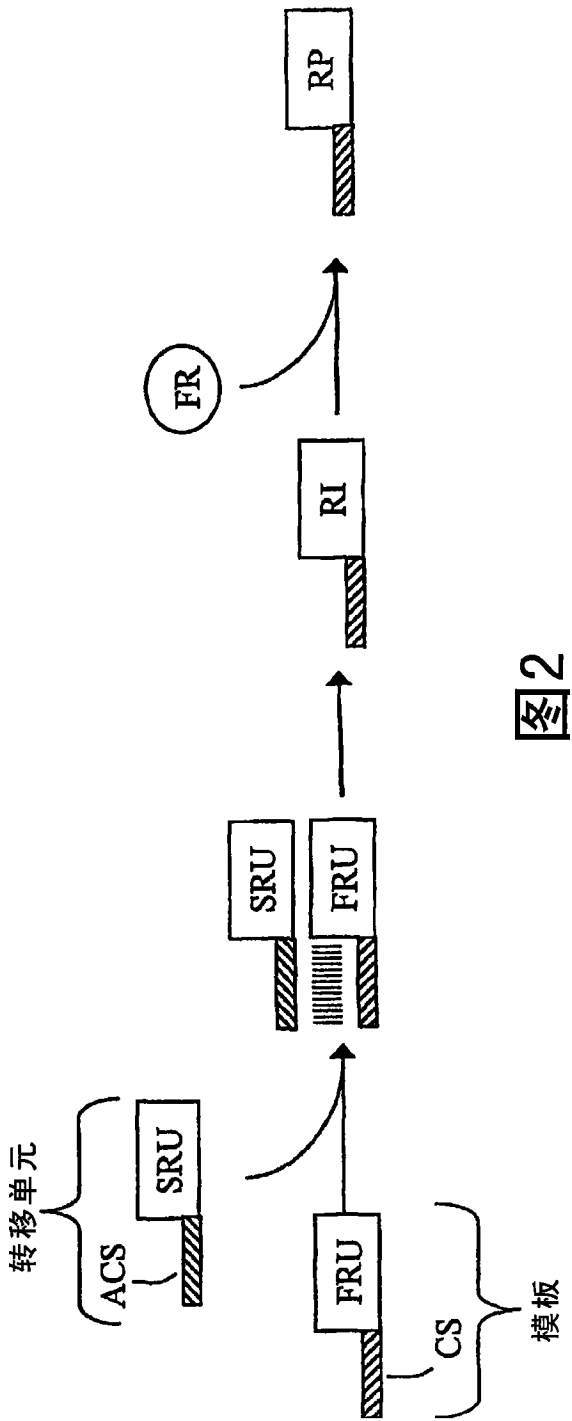


图2

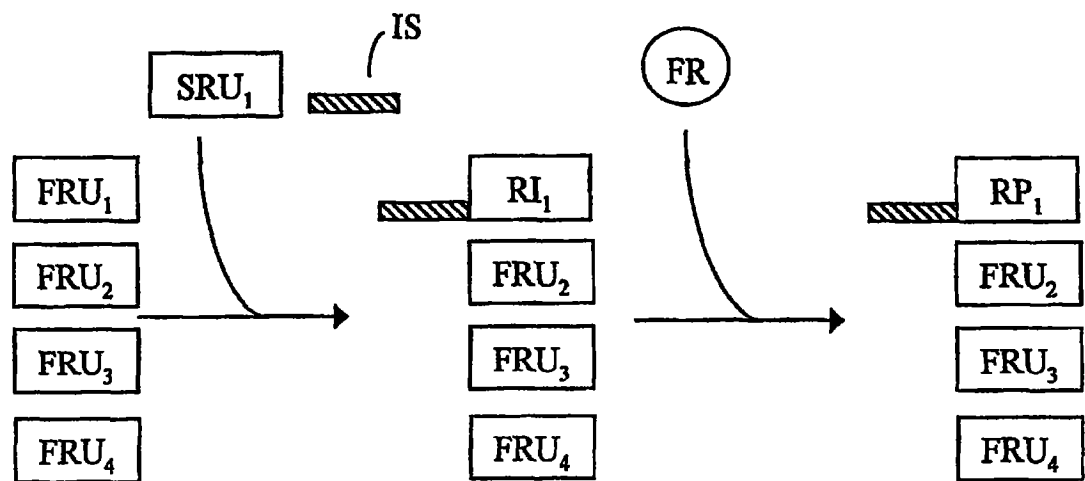


图3

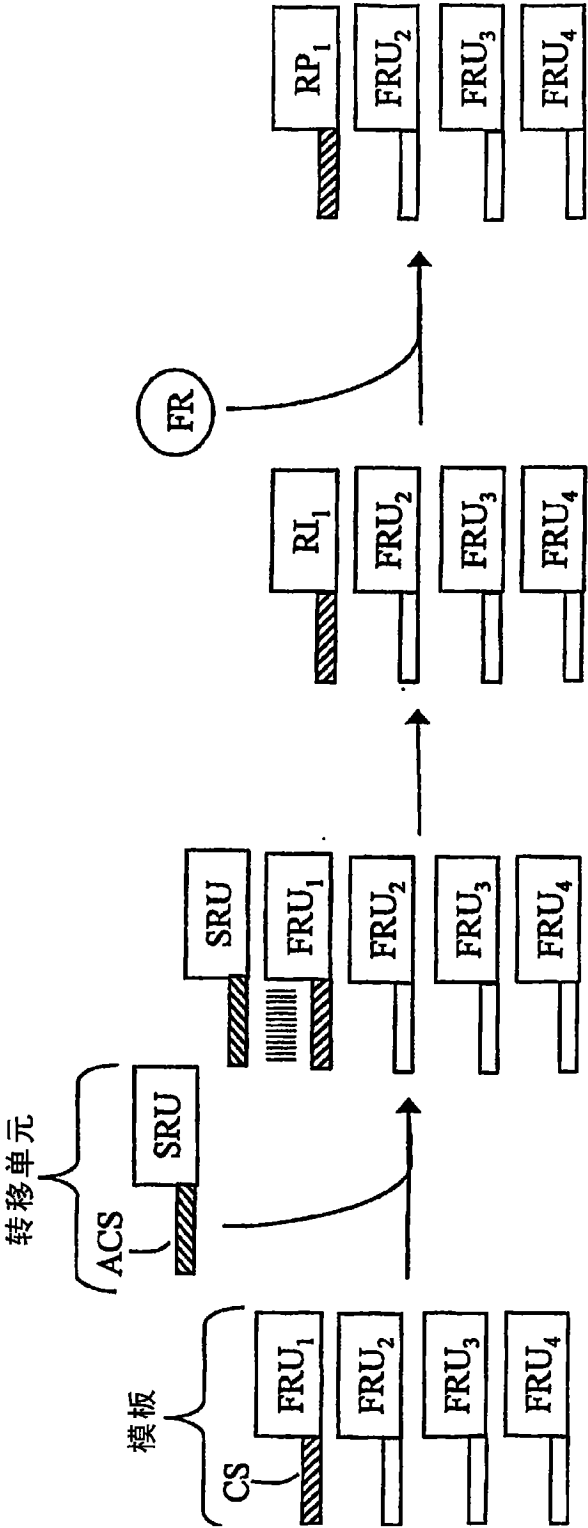


图4

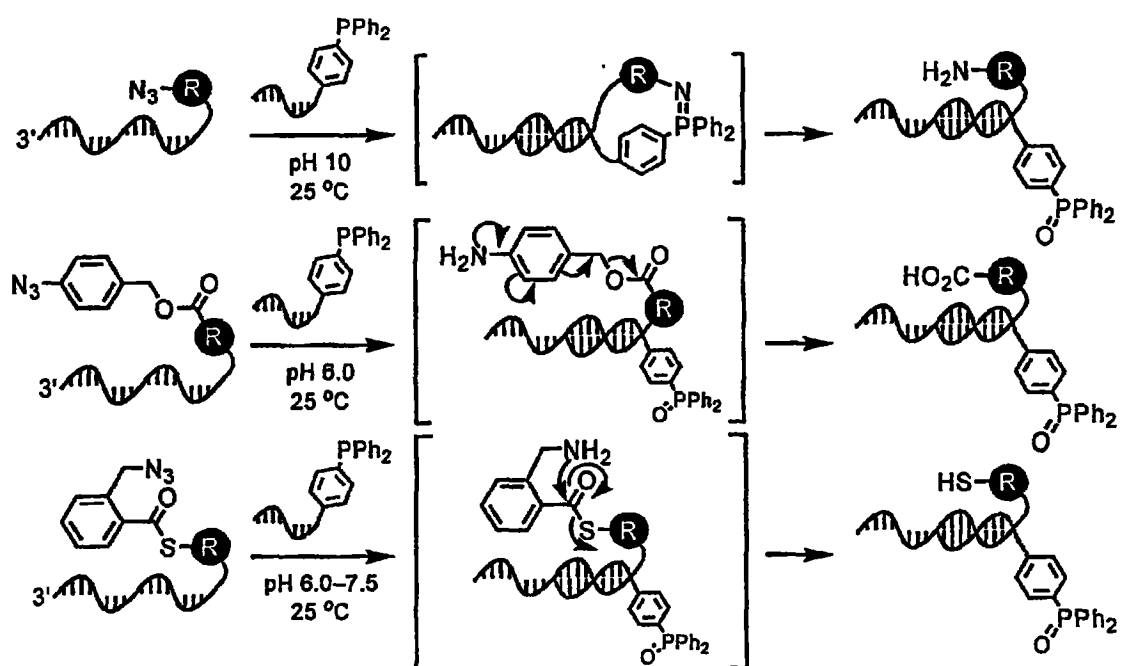


图5A

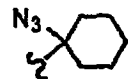
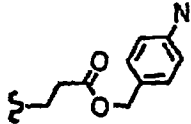
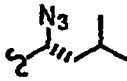
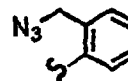
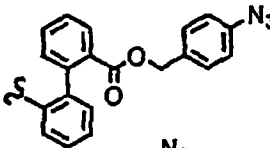
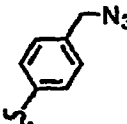
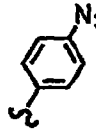
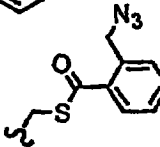
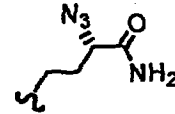
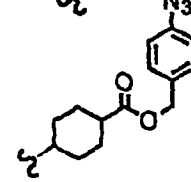
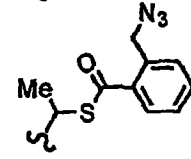
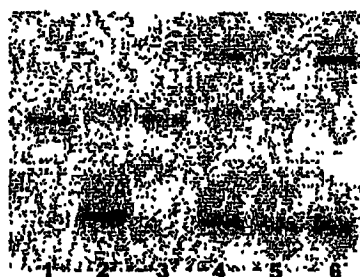
叠氮化物	% 产率	叠氮化物	% 产率	叠氮化物	% 产率
1 	77	5 	58	9 	84
2 	87	6 	81	10 	67
3 	78	7 	67	11 	71
4 	72	8 	71	12 	55

图5B

30体模板



被捕获的50体产物

20体捕获试剂

图 6A

被捕获的50体产物

30体模板

20体捕获试剂

1 2 3

被捕获的50体产物

30体模板

20体捕获试剂

1 2 3

图 6B

图 6C

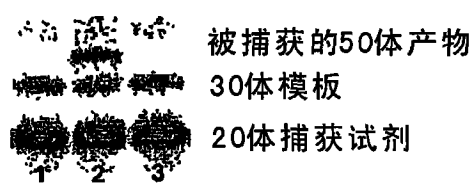


图 6D

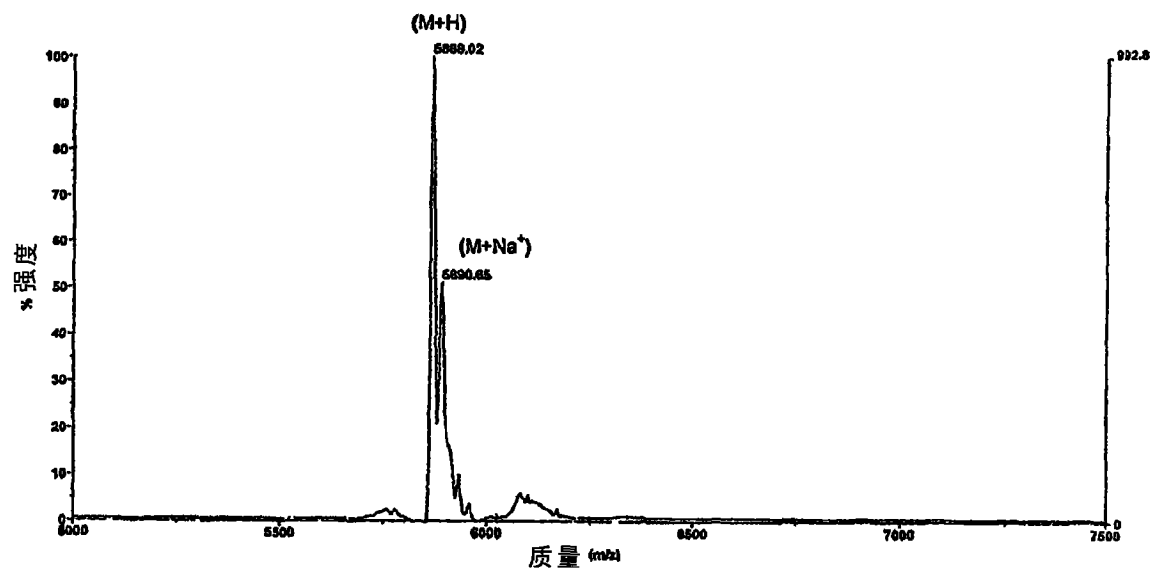


图 7

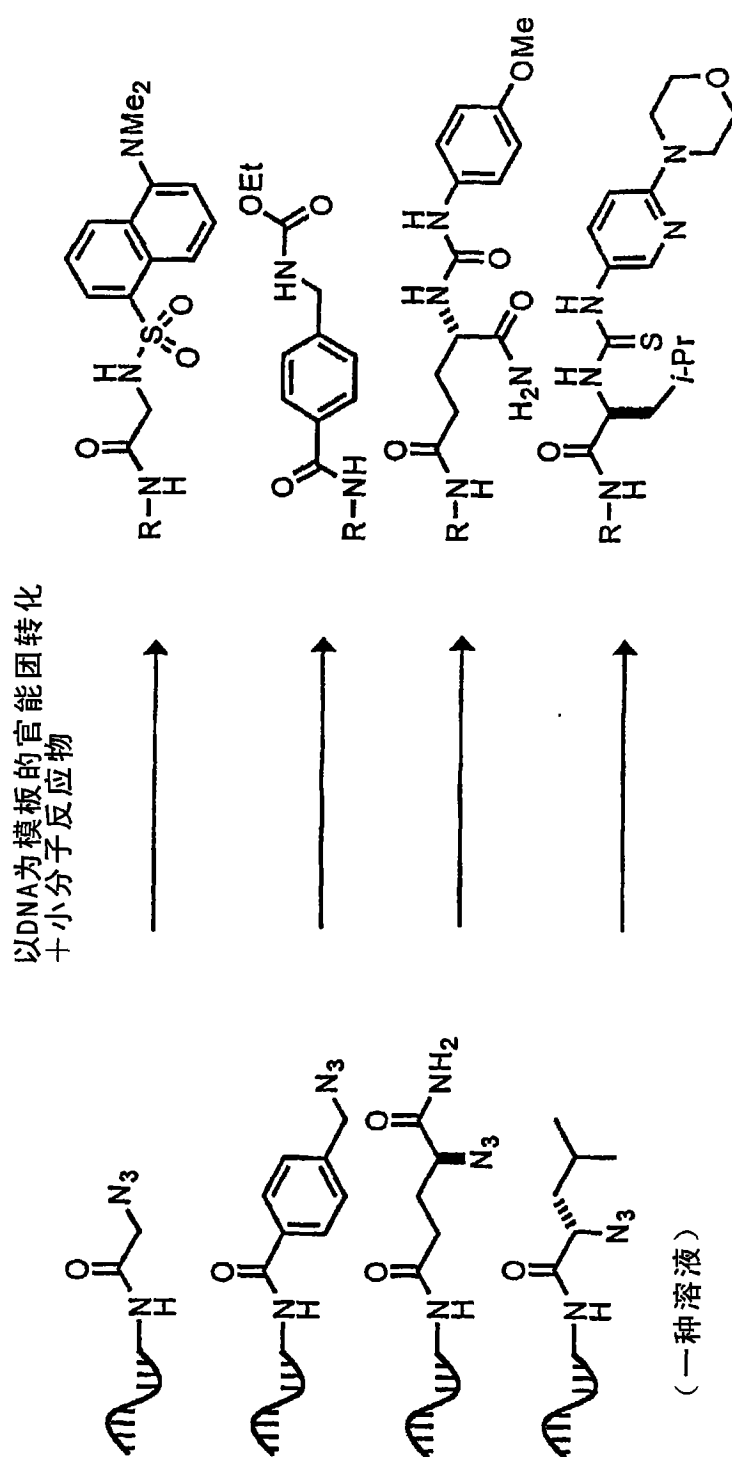


图9

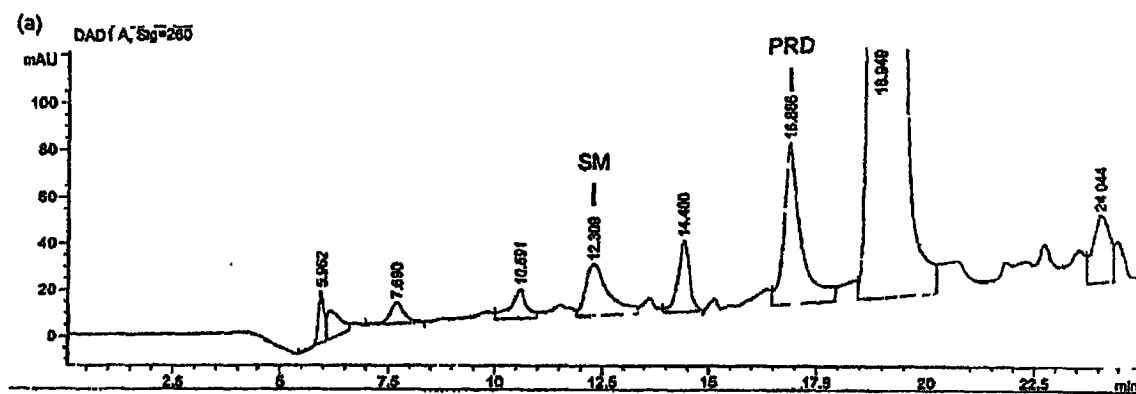


图 10A

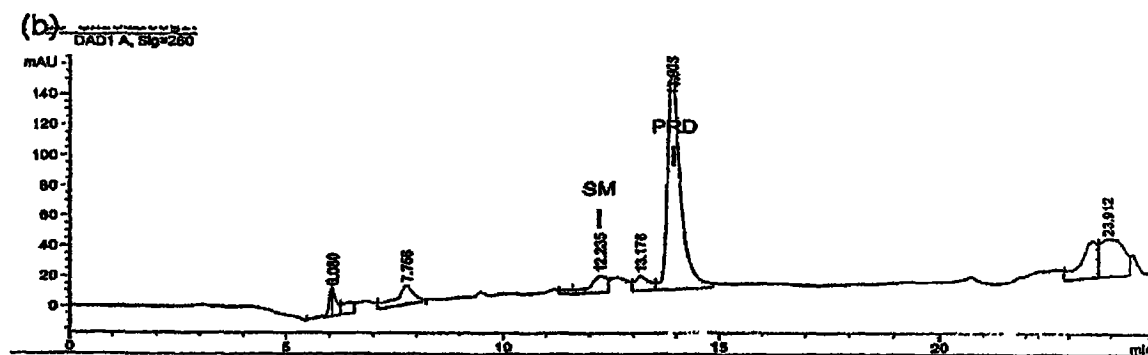


图 10B

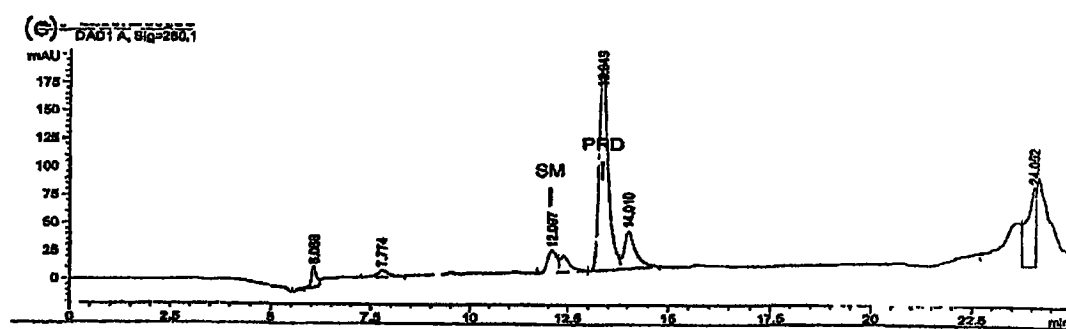


图 10C

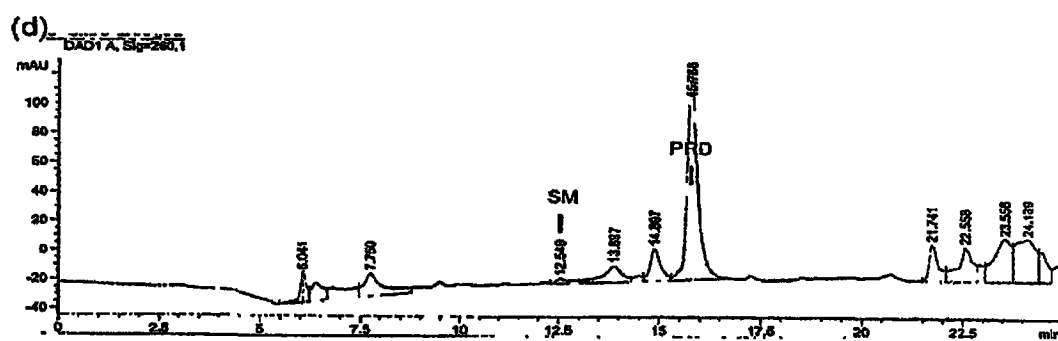
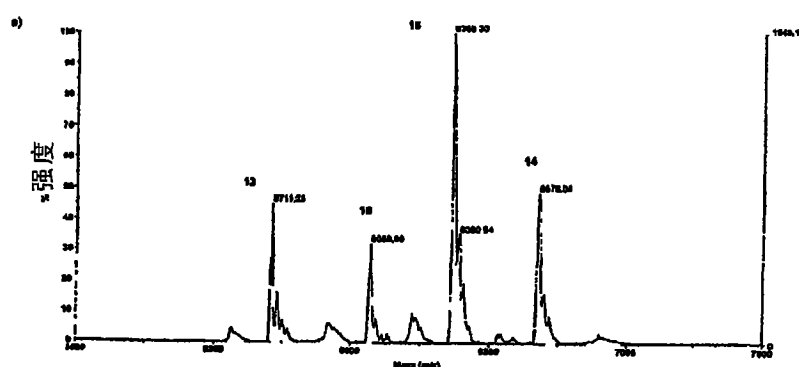


图 10D



预期质量 观察剂的质量

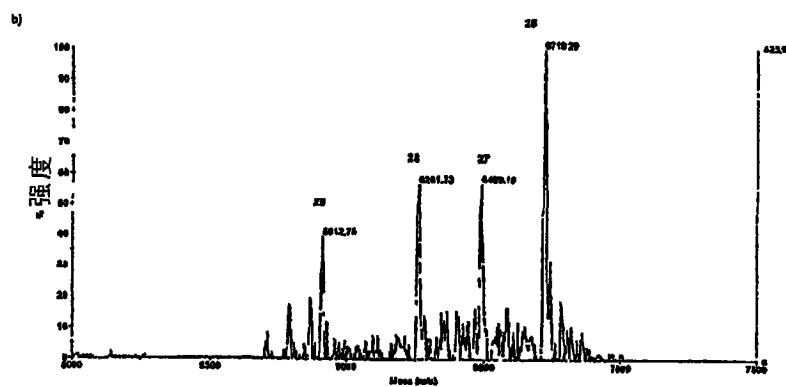
25: 5912.61 5912.75 ± 9

26: 6718.21 6719.29 ± 10

27: 6262.54 6261.73 ± 9

28: 6486.12 6489.18 ± 9

图 11A



预期质量 观察剂的质量

13: 5705.55 5711.25 ± 9

14: 6672.18 6678.04 ± 10

15: 6363.06 6368.30 ± 9

16: 6065.81 6068.96 ± 9

图 11B