



(12) 发明专利

(10) 授权公告号 CN 112739781 B

(45) 授权公告日 2024. 12. 13

(21) 申请号 201980061143.4

大卫·多德

(22) 申请日 2019.07.12

(74) 专利代理机构 北京品源专利代理有限公司

(65) 同一申请的已公布的文献号

11332

申请公布号 CN 112739781 A

专利代理师 刘明海 胡彬

(43) 申请公布日 2021.04.30

(51) Int.Cl.

(30) 优先权数据

C09B 57/00 (2006.01)

62/698,006 2018.07.13 US

C09B 11/24 (2006.01)

(85) PCT国际申请进入国家阶段日

C07H 21/00 (2006.01)

2021.03.15

C07D 413/14 (2006.01)

G01N 33/58 (2006.01)

(86) PCT国际申请的申请数据

(56) 对比文件

PCT/US2019/041717 2019.07.12

US 6133445 A, 2000.10.17

(87) PCT国际申请的公布数据

W02020/014681 EN 2020.01.16

Nicole Kretschy等.Comparison of the Sequence-Dependent Fluorescence of the Cyanine Dyes Cy3, Cy5, DyLight DY547 and DyLight DY647 on Single-Stranded DNA. 《PLOS ONE》.2014,第9卷(第1期),

(73) 专利权人 宽腾矽公司

地址 美国康涅狄格州

审查员 吴宏霞

(72) 发明人 乔纳森·M·罗斯伯格

杰瑞米·拉基 罗杰·奈尼

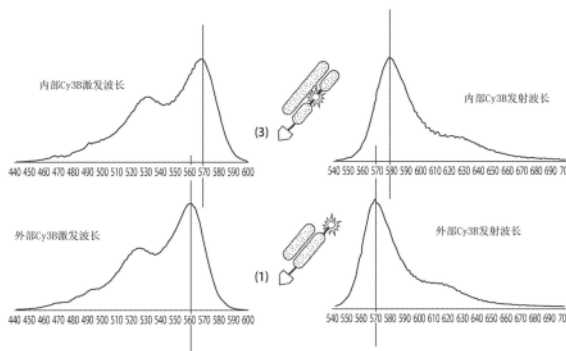
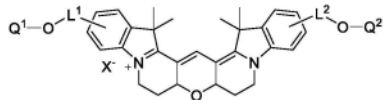
权利要求书6页 说明书68页 附图7页

(54) 发明名称

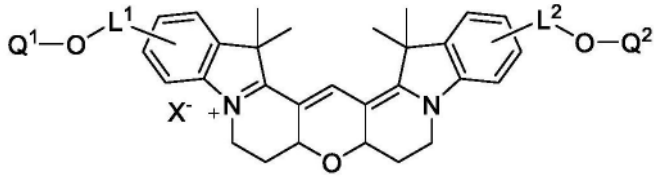
双缀合标记物和使用方法

(57) 摘要

本文提供了一种生物分子,其包括荧光标记物。该标记物是具有式(I)的Cy3B,其中Q¹和Q²独立地是单体或低聚生物分子。



1. 下式的带标记的生物分子：



其中：

Q^1 和 Q^2 独立地是单体或低聚生物分子；

X^- 是抗衡离子或不存在；并且

L^1 和 L^2 独立地是选自以下的连接基：任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合，

其中 L^1 和 L^2 直接连接至所述多环结构的一个或多个环。

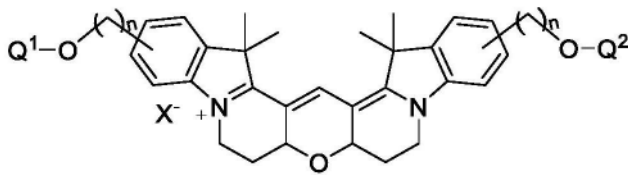
2. 根据权利要求1所述的带标记的生物分子，其中 L^1 是任选取代的亚烷基。

3. 根据权利要求2所述的带标记的生物分子，其中 L^1 是未取代的 C_{1-20} 亚烷基。

4. 根据权利要求1或2所述的带标记的生物分子，其中 L^2 是任选取代的亚烷基。

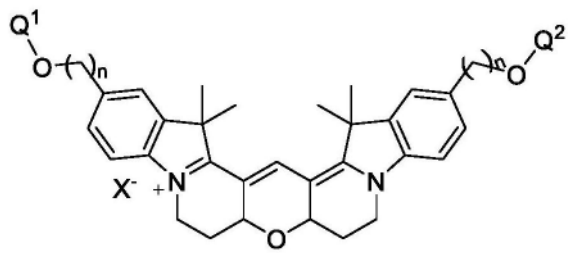
5. 根据权利要求4所述的带标记的生物分子，其中 L^2 是未取代的 C_{1-20} 亚烷基。

6. 根据权利要求1所述的带标记的生物分子，其中所述带标记的生物分子具有下式：



其中 n 独立地是1至20的整数，包括1和20。

7. 根据权利要求1或2所述的带标记的生物分子，其中所述带标记的生物分子具有下式：



8. 根据权利要求1或2所述的带标记的生物分子，其中 Q^1 和 Q^2 独立地是核苷、核苷酸、寡核苷酸、核酸或其衍生物或片段。

9. 根据权利要求8所述的带标记的生物分子，其中 Q^1 和 Q^2 独立地是脱氧核糖核酸、核糖核酸、肽核酸、锁核酸或其衍生物或片段。

10. 根据权利要求8所述的带标记的生物分子，其中所述带标记的生物分子是单链核酸，所述单链核酸包括第一寡核苷酸链。

11. 根据权利要求10所述的带标记的生物分子，其还包括与所述第一寡核苷酸链杂交的第二寡核苷酸链。

12. 根据权利要求11所述的带标记的生物分子，其中所述第二寡核苷酸链与 Q^1 和 Q^2 杂

交。

13. 根据权利要求11所述的带标记的生物分子,其中所述第二寡核苷酸链与 Q^1 或 Q^2 杂交。

14. 根据权利要求1或2所述的带标记的生物分子,其中 Q^1 和 Q^2 独立地是氨基酸、寡肽、多肽、蛋白质或其片段。

15. 根据权利要求14所述的带标记的生物分子,其中 Q^1 和 Q^2 连接在一起形成环状肽或环状蛋白质。

16. 根据权利要求1或2所述的带标记的生物分子,其中 Q^1 和 Q^2 独立地是单糖、寡糖、多糖或其片段。

17. 根据权利要求1或2所述的带标记的生物分子,其中相对于包括多环荧光团的未缀合分子,所述带标记的生物分子的一种或多种发射特性增加。

18. 根据权利要求17所述的带标记的生物分子,其中所述一种或多种发射特性选自由以下组成的组:发光寿命、发光强度、发光量子产率、亮度和最大发射波长。

19. 根据权利要求18所述的带标记的生物分子,其中相对于未缀合的分子,所述带标记的生物分子的发光寿命增加了至少10%。

20. 根据权利要求19所述的带标记的生物分子,其中相对于未缀合的分子,所述带标记的生物分子的发光寿命增加了10%至25%。

21. 根据权利要求19所述的带标记的生物分子,其中相对于未缀合的分子,所述带标记的生物分子的发光寿命增加了25%至50%。

22. 根据权利要求18所述的带标记的生物分子,其中相对于未缀合的分子,所述带标记的生物分子的最大发射波长增加了至少1%。

23. 根据权利要求22所述的带标记的生物分子,其中相对于未缀合的分子,所述带标记的生物分子的最大发射波长增加了1%至10%。

24. 根据权利要求17所述的带标记的生物分子,其中所述未缀合的分子不包括 Q^1 和 Q^2 之一或任一个。

25. 根据权利要求1或2所述的带标记的生物分子,其中 Q^1 和 Q^2 独立地任选地与一种或多种除多环荧光团之外的荧光团缔合。

26. 根据权利要求25所述的带标记的生物分子,其中每个荧光团与任意其他荧光团至少相距5埃。

27. 根据权利要求1或2所述的带标记的生物分子,其中 Q^1 和 Q^2 独立地任选地与配置为用作反应底物的反应剂缔合。

28. 根据权利要求11所述的带标记的生物分子,其中第一寡核苷酸链和第二寡核苷酸链独立地任选地与配置为用作反应底物的反应剂缔合。

29. 根据权利要求27所述的带标记的生物分子,其中所述反应是聚合反应。

30. 根据权利要求27所述的带标记的生物分子,其中当处于聚合反应条件下时,所述反应剂被聚合酶从所述带标记的生物分子上切割下来。

31. 一种带标记的核苷酸,其包括与权利要求1-30中任一项所述的带标记的生物分子缔合的一个或多个核苷酸。

32. 根据权利要求31所述的带标记的核苷酸,其中所述一个或多个核苷酸包括选自鸟

嘌呤、胞嘧啶、腺嘌呤和胸腺嘧啶或尿嘧啶的一种类型的核苷酸。

33. 根据权利要求31所述的带标记的核苷酸,其中当处于聚合反应条件下时,所述一个或多个核苷酸被聚合酶从所述带标记的生物分子上切割下来。

34. 根据权利要求31所述的带标记的核苷酸,其中所述一个或多个核苷酸包括三磷酸核苷。

35. 根据权利要求31所述的带标记的核苷酸,其中所述一个或多个核苷酸包括六磷酸核苷。

36. 根据权利要求34所述的带标记的核苷酸,其中所述一个或多个核苷酸通过末端磷酸酯连接至所述带标记的生物分子。

37. 一种组合物,其包括权利要求31-36中任一项所述的带标记的核苷酸。

38. 一种核酸测序反应组合物,其在反应混合物中包括两种或更多种不同类型的带标记的核苷酸,其中至少一种类型的带标记的核苷酸是根据权利要求31-36中任一项所述的带标记的核苷酸。

39. 根据权利要求38所述的核酸测序反应组合物,其包括四种不同类型的带标记的核苷酸。

40. 根据权利要求39所述的核酸测序反应组合物,其中所述四种不同类型的带标记的核苷酸包括第一带标记的核苷酸、第二带标记的核苷酸、第三带标记的核苷酸和第四带标记的核苷酸,其中所述第一带标记的核苷酸包括鸟嘌呤,所述第二带标记的核苷酸包括胞嘧啶,所述第三带标记的核苷酸包括腺嘌呤,所述第四带标记的核苷酸包括胸腺嘧啶或尿嘧啶。

41. 一种确定模板核酸序列的方法,所述方法包括:

(i) 将目标体积的复合物暴露于根据权利要求38-40中任一项所述的核酸测序反应组合物中,其中所述复合物包括模板核酸、引物和聚合酶;

(ii) 将一系列具有一个或多个激发能的脉冲引向目标体积附近;

(iii) 在顺序掺入到包含引物的核酸中期间,检测从发光标记的核苷酸发射的多个光子;以及

(iv) 通过确定发射的光子的时间和任选的发光强度,来鉴定掺入的核苷酸的序列。

42. 一种用于对模板核酸进行测序的试剂盒,所述试剂盒包括:

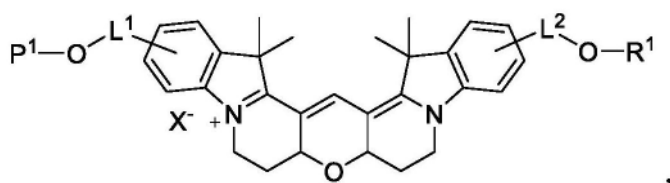
两种或更多种不同类型的带标记的核苷酸,其中至少一种类型的带标记的核苷酸包括根据权利要求31-36中任一项所述的带标记的核苷酸。

43. 根据权利要求42所述的试剂盒,其包括四种不同类型的带标记的核苷酸。

44. 根据权利要求42所述的试剂盒,其还包括聚合酶。

45. 根据权利要求42所述的试剂盒,其还包括与所述模板核酸互补的引物。

46. 下式的化合物或其盐:



其中:

X⁻是抗衡离子或不存在；

L¹和L²独立地是选自以下的连接基：任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合；

P¹是氧保护基团；并且

R¹是反应性部分，

其中L¹和L²直接连接至所述多环结构的一个或多个环。

47. 根据权利要求46所述的化合物，其中R¹是亚磷酰胺基团。

48. 根据权利要求46或47所述的化合物，其中L¹是任选取代的亚烷基。

49. 根据权利要求48所述的化合物，其中L¹是未取代的C₁₋₂₀亚烷基。

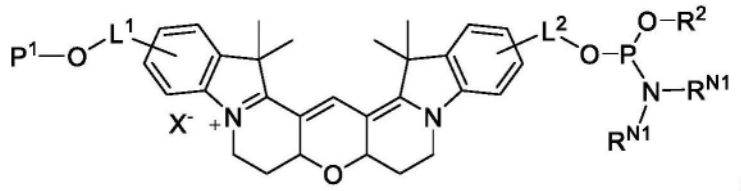
50. 根据权利要求46或47所述的化合物，其中L²是任选取代的亚烷基。

51. 根据权利要求50所述的化合物，其中L²是未取代的C₁₋₂₀亚烷基。

52. 根据权利要求46或47所述的化合物，其中P¹是任选取代的三苯甲基保护基团。

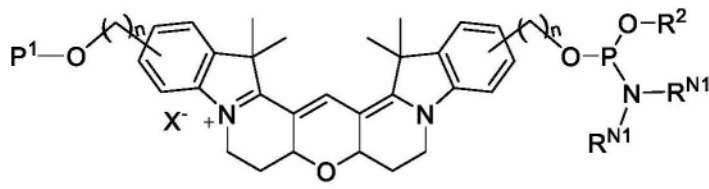
53. 根据权利要求52所述的化合物，其中P¹是4-甲氧基三苯甲基 (MMT)。

54. 根据权利要求46或47所述的化合物，其中所述化合物具有下式：



或其盐。

55. 根据权利要求46或47所述的化合物，其中所述化合物具有下式：



或其盐。

56. 根据权利要求55所述的化合物，其中R²是任选取代的C₁-C₆烷基。

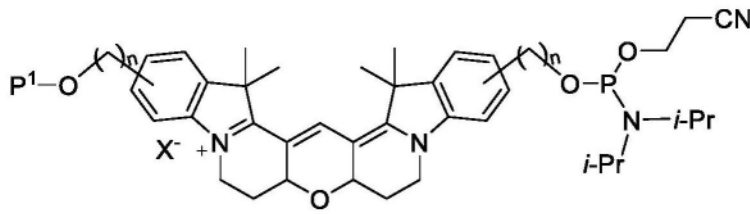
57. 根据权利要求56所述的化合物，其中R²是：

58. 根据权利要求55所述的化合物，其中每个R^N都是任选取代的C₁₋₆烷基。

59. 根据权利要求58所述的化合物，其中每个R^N都是异丙基。

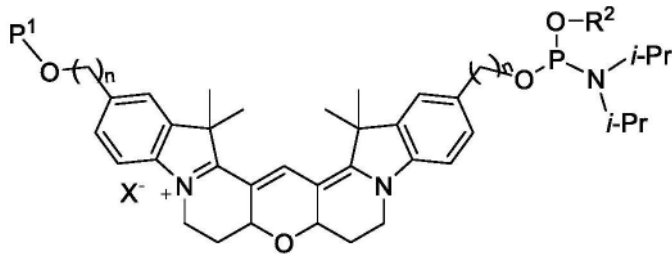
60. 根据权利要求55所述的化合物，其中每个n都为4。

61. 根据权利要求46或47所述的化合物，其中所述化合物具有下式：



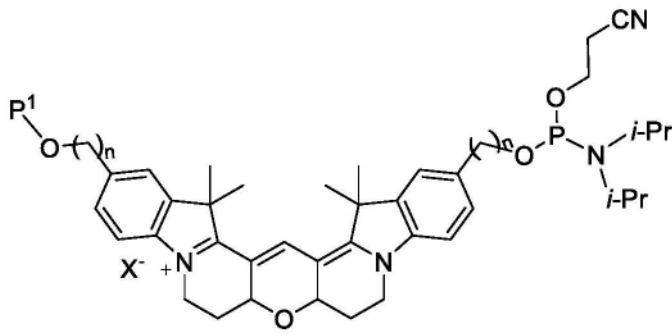
或其盐。

62. 根据权利要求46或47所述的化合物,其中所述化合物具有下式:



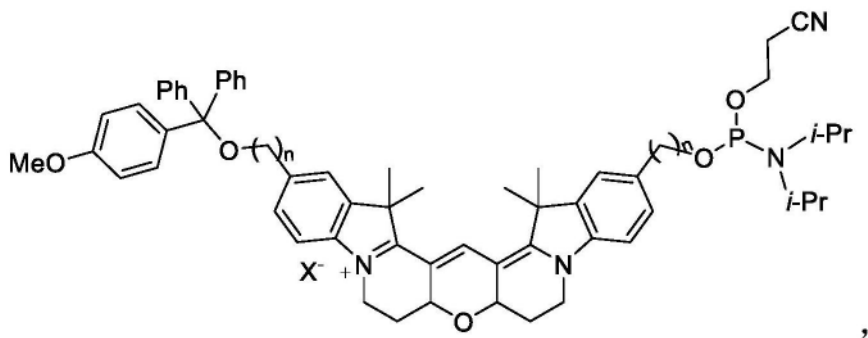
或其盐。

63. 根据权利要求46或47所述的化合物,其中所述化合物具有下式:



或其盐。

64. 根据权利要求46或47所述的化合物,其中所述化合物具有下式:

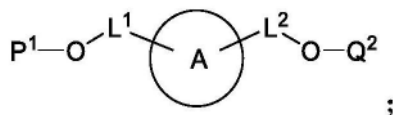


或其盐。

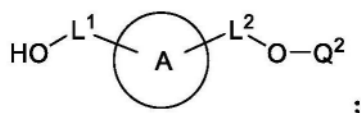
65. 一种组合物,其包括权利要求46或47所述的化合物。

66. 一种制备根据权利要求1所述的带标记的生物分子的方法,所述方法包括:

(i) 在足以促进缀合的条件下,使式Q²-OH的单体或低聚生物分子或其盐与权利要求46或47的化合物或其盐接触,得到下式缀合物:



(ii) 在足以裂解P¹保护基团并产生下式缀合物的条件下,对步骤(i)中形成的缀合物进行脱保护:



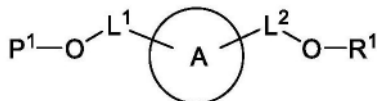
(iii) 在足以促进缀合的条件下,使步骤(ii)中形成的缀合物与式 Q^1-O-R^1 的单体或低聚生物分子或其盐接触,得到根据权利要求1所述的带标记的生物分子。

[0009] 在一些方面,本文提供了一种核酸测序反应组合物,其在反应混合物中包括两种或更多种不同类型的带标记的核苷酸。在一些实施方案中,核酸测序反应组合物中至少一种类型的标记核苷酸是根据本申请的标记核苷酸。在一些实施方案中,核酸测序反应组合物包括四种不同类型的带标记的核苷酸。在一些实施方案中,核酸测序反应组合物包括第一带标记的核苷酸、第二带标记的核苷酸、第三带标记的核苷酸和第四带标记的核苷酸,其中所述第一带标记的核苷酸包括鸟嘌呤,所述第二带标记的核苷酸包括胞嘧啶,所述第三带标记的核苷酸包括腺嘌呤,所述第四带标记的核苷酸包括胸腺嘧啶或尿嘧啶。

[0010] 在一些方面,本文提供了确定模板核酸的序列的方法。在一些实施方案中,所述方法包括将目标体积的复合物暴露于本申请的核酸测序反应组合物中,其中所述复合物包括模板核酸、引物和聚合酶。在一些实施方案中,该方法还包括将一系列具有一个或多个激发能的脉冲引向目标体积附近。在一些实施方案中,该方法还包括在顺序掺入到包含引物的核酸中期间,检测从发光标记的核苷酸发射的多个光子。在一些实施方案中,该方法还包括通过确定发射的光子的时间和任选的发光强度,来鉴定掺入的核苷酸的序列。

[0011] 在一些方面,本文提供了用于对模板核酸进行测序的试剂盒。在一些实施方案中,该试剂盒包括两种或更多种不同类型的带标记的核苷酸。在一些实施方案中,这两种或更多种不同类型的带标记的核苷酸中的至少一种包括本申请的带标记的核苷酸。

[0012] 在一些方面,本申请提供了式 (II) 化合物或其盐:

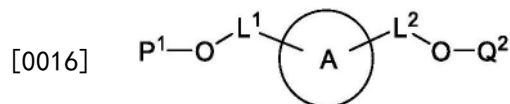


[0013]

(II),

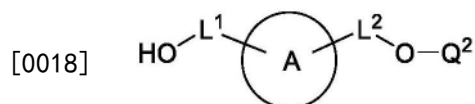
[0014] 其中:A是多环荧光团; L^1 和 L^2 独立地是选自以下的连接基:任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合; P^1 是氧保护基团;并且 R^1 是反应性部分。在一些方面,本文提供了组合物,所述组合物包括式 (II) 化合物。

[0015] 在一些方面,本文提供了用于制备本申请的带标记的生物分子的方法。在一些实施方案中,所述方法包括:(i) 在足以促进缀合的条件下,使式 $Q^2\text{-OH}$ 的单体或低聚生物分子或其盐与式 (II) 化合物或其盐接触,得到下式缀合物:



[0016]

[0017] 在一些实施方案中,该方法还包括:(ii) 在足以裂解 P^1 保护基团并产生下式缀合物的条件下,对步骤 (i) 中形成的缀合物进行脱保护:



[0018]

[0019] 在一些实施方案中,该方法还包括:(iii) 在足以促进缀合的条件下,使步骤 (ii) 中形成的缀合物与式 $Q^1\text{-O-R}^1$ 的单体或低聚生物分子或其盐接触,得到式 (I) 的带标记的生

物分子。

[0020] 如下所述,在某些实施例的详细描述中阐述了本发明的某些实施例的细节。通过定义、实施例、附图和权利要求书,本发明的其他特征、目的和优点将显而易见。

附图说明

[0021] 构成本说明书一部分的附图说明了本发明的几种实施方案,并与本描述一起,起到解释本发明原理的作用。

[0022] 图1A-1G示出了本申请的带标记的生物分子的各种实例。图1A示出了具有一个内部标记物的带标记的生物分子。图1B示出了具有两个内部标记物的带标记的生物分子。图1C示出了环化的带标记的生物分子。图1D示出了带标记的寡核苷酸,其包括一条内部标记的链,该内部标记的链与未标记的链杂交。图1E示出了带标记的寡核苷酸,其包括一条内部标记的链,该内部标记的链与另一条内部标记的链杂交。图1F示出了带标记的寡核苷酸,其包括一条未标记的链,该未标记的链与具有两个内部标记物的带标记的链杂交。图1G示出了带标记的寡核苷酸链,其自杂交,形成了茎环基序。

[0023] 图2A-2C描绘了本申请的带标记的生物分子的用途的实例。图2A描述了与靶核酸序列杂交的内部标记的寡核苷酸。图2B描述了与靶蛋白结合的内部标记的生物分子。图2C描述了与靶蛋白结合的内部标记的抗体。

[0024] 图3A-3B说明了本申请的带标记的生物分子的实例,该带标记的生物分子被功能性部分修饰。图3A描绘了用猝灭部分修饰的内部标记的寡核苷酸,该猝灭部分阻止了内部标记物的检测,除非从生物分子上切割下来。图3B描绘了用配体修饰的内部标记的生物分子,该配体与靶蛋白结合,其中该配体与靶蛋白的结合允许检测内部标记物。

[0025] 图4A-4C描述了根据本申请与聚合酶结合的带标记的核苷酸的实例。图4A描述了与核苷酸结合的聚合酶,该核苷酸包括外部标记物。图4B描述了与核苷酸结合的聚合酶,该核苷酸包括外部标记的生物分子。图4C描述了与核苷酸结合的聚合酶,该核苷酸包括内部标记的寡核苷酸。

[0026] 图5A-5B示出了根据本申请的外部标记的生物分子和内部标记的生物分子的测序分析比较。图5A概括地描绘了一组外部标记的核苷酸和内部标记的核苷酸,它们被制备并进行进一步分析。图5B示出了使用外部标记的核苷酸和内部标记的核苷酸进行的单分子测序反应的结果。

[0027] 图6示出了外部标记的核苷酸和内部标记的核苷酸的激发和发射光谱的比对。

具体实施方式

[0028] 在其他方面中,本公开提供了带标记的生物分子,所述带标记的生物分子包括内部缀合的发光标记物(例如,内部标记物)。在一些实施方案中,内部标记物配置有增强的构象约束,从而限制旋转并阻断缩短发光寿命的失活途径。在一些实施方案中,带标记的生物分子被配置成为内部标记物提供刚性分子支架,从而避免了标记物-标记物的相互作用和其他可能降低发光强度或其他发射特性的猝灭作用。

[0029] 不希望受到任何特定理论的束缚,本文提供的带标记的生物分子具有许多明显的优点,例如量子产率提高和发光寿命延长、发光强度和/或亮度增加,并且对大量溶剂分子

的暴露降低,从而限制了反应性物质的形成。因此,在一些实施方案中,本文提供了带标记的生物分子及其使用方法。在一些实施方案中,本公开提供了与制备带标记的生物分子有关的组合物和方法。在一些实施方案中,本公开提供了与制备双缀合标记物有关的组合物和方法。

[0030] 内部标记的生物分子

[0031] 本发明的一方面涉及内部标记的生物分子(“带标记的生物分子”)。如本文所述,带标记的生物分子(例如,寡核苷酸、核酸、多肽、蛋白质、多糖)包括内部缀合的发光化合物(“发光标记物”或“标记物”,例如多环荧光团)。“内部标记的”或“内部缀合的”在本文中可互换使用,是指生物分子的一部分与发光化合物的第一个位点缀合,生物分子的另一部分与发光化合物的第二个位点缀合。发光化合物的内部缀合具有多个优点。例如,在生物分子的内部缀合一种标记物可以改变该标记物的寿命和其他光物理性质(例如,通过限制标记物的旋转)。作为另一个实例,在内部缀合标记物可以使该标记物固定,将其与其他标记物分开,从而减轻自猝灭。此外,在内部缀合标记物可能会限制标记物接触溶液中的大量溶剂,从而减轻自由基的形成,自由基的形成可能会损害溶液中的其他成分。在内部掺入标记物具有本文所述的其他优点。带标记的生物分子的一组非限制性实例示于图1A-1G。

[0032] 图1A-1G概括地描述了本申请的带标记的生物分子的各种构型。每个实例都显示为具有内部标记物和生物分子(显示为点状形状)。图1A是一种带标记的生物分子,其包括一个内部标记物,该内部标记物将生物分子的一部分与生物分子的另一部分缀合。在一些实施方案中,带标记的生物分子包括两个或更多个内部标记物。

[0033] 图1B是一种带标记的生物分子,包括两个内部标记物。如图所示,一个标记物将生物分子的第一部分与生物分子的第二部分缀合,另一个标记物将生物分子的第二部分与生物分子的第三部分缀合。在一些实施方案中,相对于具有一个内部标记物拷贝的生物分子,具有两个或更多个相同内部标记物拷贝的带标记的生物分子的发光强度和/或亮度增加。在一些实施方案中,具有两种或更多种不同类型的内部标记物的带标记的生物分子可以提供两种或更多种独特的可检测信号。多重标记的生物分子的各种构型和用途的实例在本文其他地方描述。

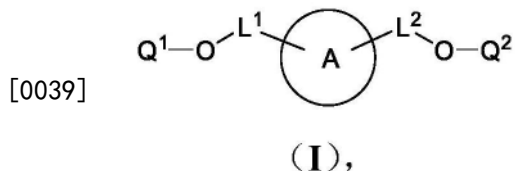
[0034] 根据本申请,已证明,内部缀合的标记物的刚性增加可以增强带标记的生物分子的一种或多种发光特性,例如通过内部标记物的多环荧光团提供的构象约束。有利的是,可以对在其上缀合有内部标记物的生物分子支架进行工程设计,以进一步促进刚性,并进一步增强这些系统中的一种或多种发光特性。

[0035] 图1C是一种带标记的生物分子,其包括与环化的生物分子缀合的内部标记物。如该实例所描述的,在一些实施方案中,内部标记物将生物分子的一端与生物分子的另一端缀合,以使带标记的生物分子环化。不希望受任何特定理论的束缚,认为环化作用可以促进整个生物分子主链的结构刚性,这增强了本文提供的内部标记物的有利的发光性质。环状生物分子的实例包括但不限于环状肽、环状蛋白质和环状核酸(例如,环状RNA、DNA质粒)。

[0036] 发明人还认识并意识到,寡核苷酸(例如,多核苷酸、核酸)为本申请的内部标记物提供了刚性的、高度可调的生物分子支架。内部标记的寡核苷酸的各种实例示于图1D-1G。这些与内部标记的寡核苷酸有关的示例构建体和其他实施方案在本文其他地方进行了详细描述。

[0037] 如在图1A-1G所示的示例结构中一般说明的,在一些实施方案中,如本文所述,内部标记物100对应于式(I)的A。在一些实施方案中,如本文所述,示例结构的生物分子(显示为点状形状)对应于式(I)的 Q^1 和/或 Q^2 。在一些实施方案中,如本文所述,内部标记物100对应于A,并且示例结构的生物分子(显示为点状形状)对应于式(I)的 Q^1 和 Q^2 。

[0038] 一方面,本文提供了式(I)的带标记的生物分子:




[0040] 其中:

[0041] Q^1 和 Q^2 独立地是单体或低聚生物分子;

[0042] A是多环荧光团;并且

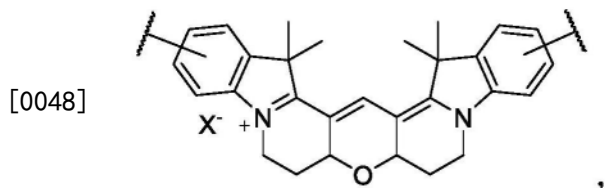
[0043] L^1 和 L^2 独立地是选自以下的连接基:任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合。

[0044] 如本文所述,  (在本文中也描述为“A”)是发光化合物或染料。如式(I)所示,A通过由 L^1 表示的连接基与生物分子的一部分(Q^1 基团)缀合,并且通过连接基 L^2 与生物分子的另一部分(Q^2 基团)缀合。 Q^1 和 Q^2 一起形成生物分子,该分子被 $-O-L^1-A-L^2-O-$ 结构打断,从而在内部被标记。

[0045] 在某些实施方案中,A是多环荧光团。在某些实施方案中,掺入多环荧光团是有利的,因为与线性或非多环荧光团相比,多环结构可以给该体系带来更大的刚性。在某些实施方案中, L^1 和 L^2 直接连接(例如,通过共价键)到多环结构的一个或多个环(例如,苯环或杂芳环)上。这种直接连接也可以赋予该体系更大的刚性(例如,通过标记物的固定化/限制旋转)。在某些实施方案中, L^1 和 L^2 直接连接到A上的不同环上,这种设计特征也可以给该体系带来更大的刚性和/或帮助固定染料。

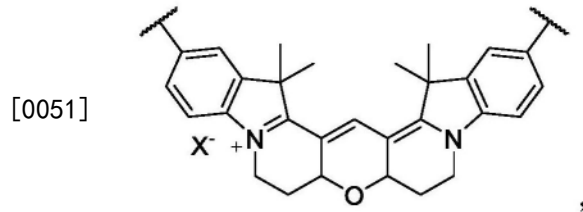
[0046] 在某些实施方案中,A是多环花菁、荧光酮(fluorone)、吡啶、吩噻嗪、香豆素或硼二吡咯亚甲基(BODIPY)荧光团。在某些实施方案中,A是卟啉、酞菁或萘二甲酰亚胺。这些是非限制性实例。在某些实施方案中,A是多环花菁荧光团。在某些实施方案中,A是多环荧光酮荧光团。在某些实施方案中,A是多环吡啶荧光团。在某些实施方案中,A是多环吩噻嗪荧光团。在某些实施方案中,A是多环香豆素荧光团。在某些实施方案中,A是多环BODIPY荧光团。在某些实施方案中,A是卟啉。在某些实施方案中,A是酞菁。在某些实施方案中,A是萘二甲酰亚胺。在下文和本文中描述了环A的其他实施方案。

[0047] 在某些实施方案中,A是多环花菁荧光团。在某些实施方案中,A是任选取代的Cy3B染料。在某些实施方案中,A是Cy3B。在某些实施方案中,A具有下式:



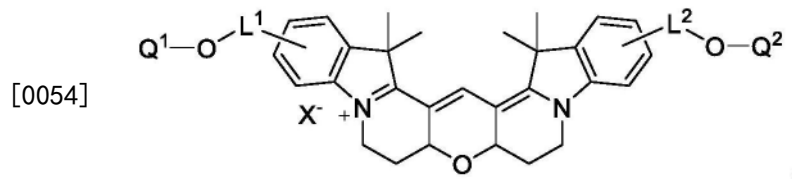
[0049] 其中 X^- 是抗衡离子或不存在;并且其中该结构任选地在任意位置被取代。在某些实施方案中,该结构是未取代的。

[0050] 在某些实施方案中,A具有下式:



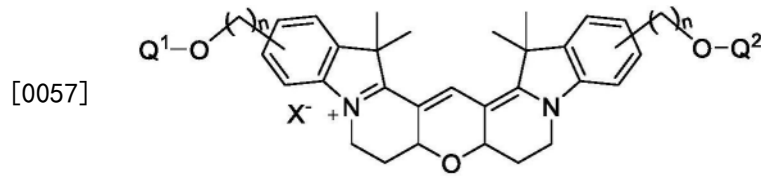
[0052] 其中 X^- 是抗衡离子或不存在;并且其中该结构任选地在任意位置被取代。在某些实施方案中,该结构是未取代的。

[0053] 因此,在某些实施方案中,带标记的生物分子具有下式:



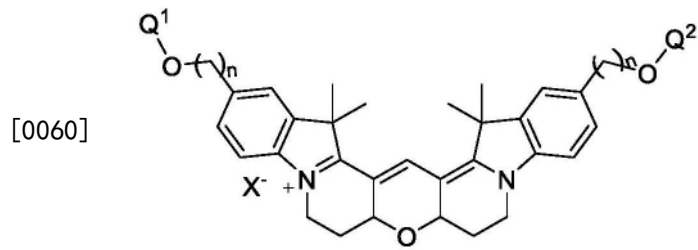
[0055] 其中 X^- 是抗衡离子或不存在。

[0056] 在某些实施方案中,带标记的生物分子具有下式:

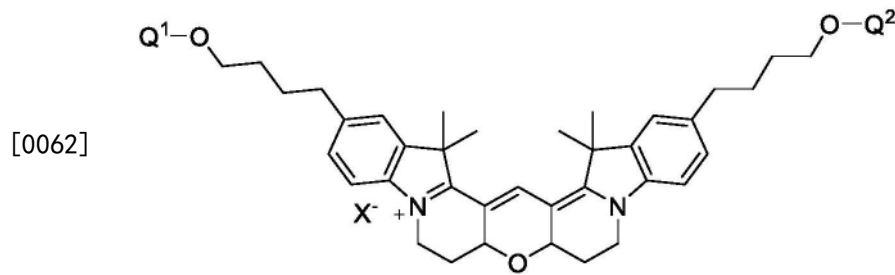


[0058] 其中,n独立地是1至20的整数,包括1和20。

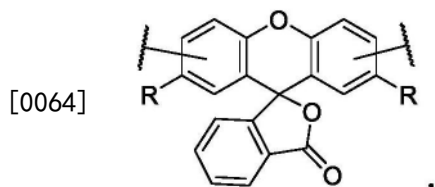
[0059] 在某些实施方案中,带标记的生物分子具有下式:



[0061] 在某些实施方案中,带标记的生物分子具有下式:



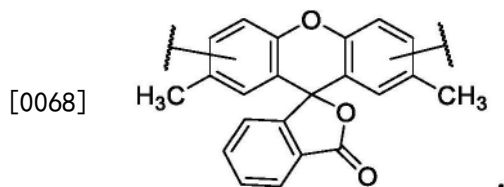
[0063] 如本文所述,在某些实施方案中,环A是多环荧光酮荧光团(例如,荧光素或若丹明)。在某些实施方案中,A是荧光酮染料。在某些实施方案中,A是罗丹明染料。在某些实施方案中,A具有下式:



[0065] 其中每个R独立地是氢、卤素、 $-N_3$ 、 $-CN$ 、 $-NO_2$ 、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、 $-OR^0$ 、 $-SR^S$ 或 $-N(R^N)_2$ 。在某些实施方案中,该结构任选地在任意位置被取代。在某些实施方案中,该结构是未取代的。

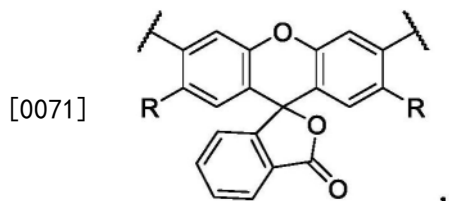
[0066] 如本文所定义的,每个R独立地是氢、卤素、 $-N_3$ 、 $-CN$ 、 $-NO_2$ 、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、 $-OR^0$ 、 $-SR^S$ 或 $-N(R^N)_2$ 。在某些实施方案中,R是氢。在某些实施方案中,R是卤素。在某些实施方案中,R是 $-N_3$ 。在某些实施方案中,R是 $-CN$ 。在某些实施方案中,R是 $-NO_2$ 。在某些实施方案中,R是任选取代的烷基。在某些实施方案中,R是任选取代的烯基。在某些实施方案中,R是任选取代的炔基。在某些实施方案中,R是任选取代的碳环基。在某些实施方案中,R是任选取代的芳基。在某些实施方案中,R是任选取代的杂环基。在某些实施方案中,R是任选取代的杂芳基。在某些实施方案中,R是 $-OR^0$ 。在某些实施方案中,R是 $-SR^S$ 。在某些实施方案中,R是 $-N(R^N)_2$ 。在某些实施方案中,R是任选取代的 C_{1-6} 烷基。在某些实施方案中,R是未取代的 C_{1-6} 烷基。在某些实施方案中,R是任选取代的 C_{1-3} 烷基。在某些实施方案中,R是未取代的 C_{1-3} 烷基。在某些实施方案中,R选自自由以下组成的组:甲基、乙基、正丙基、异丙基、正丁基、异丁基、仲丁基和叔丁基。在某些实施方案中,R是甲基。在某些实施方案中,每个R都是甲基。

[0067] 在某些实施方案中,A具有下式:



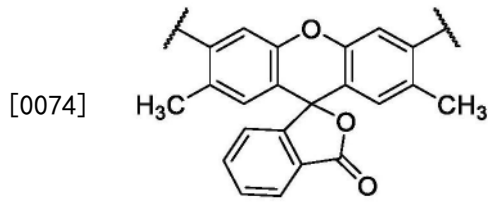
[0069] 其中该结构任选地在任意位置被取代。在某些实施方案中,该结构是未取代的。

[0070] 在某些实施方案中,A具有下式:



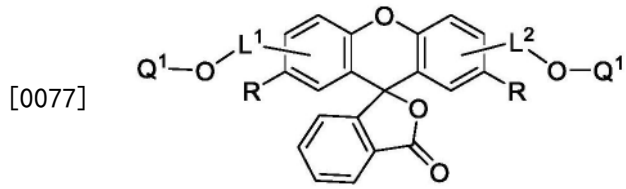
[0072] 其中该结构任选地在任意位置被取代。在某些实施方案中,该结构是未取代的。

[0073] 在某些实施方案中,A具有下式:

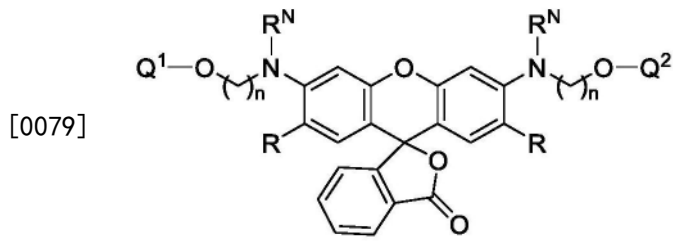


[0075] 其中该结构任选地在任意位置被取代。在某些实施方案中,该结构是未取代的。

[0076] 因此,在某些实施方案中,带标记的生物分子具有下式:



[0078] 在某些实施方案中,带标记的生物分子具有下式:

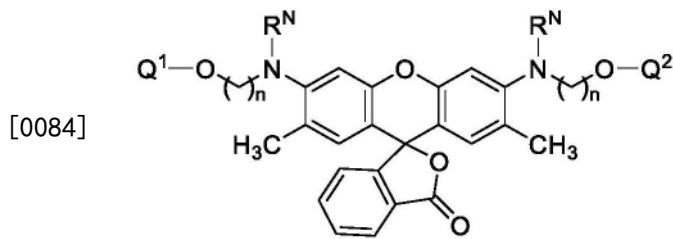


[0080] 其中:

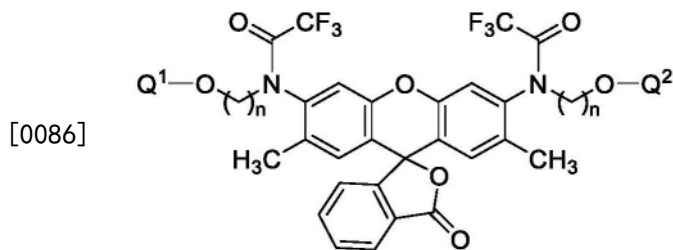
[0081] n独立地是1至20的整数,包括1和20;并且

[0082] R^N的每种情况都独立地是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氮保护基团。

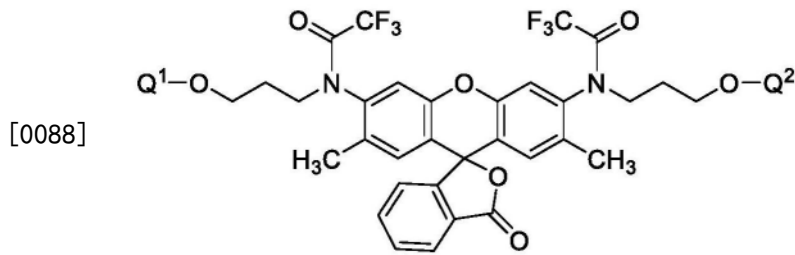
[0083] 在某些实施方案中,带标记的生物分子具有下式:



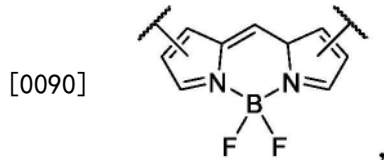
[0085] 在某些实施方案中,带标记的生物分子具有下式:



[0087] 在某些实施方案中,带标记的生物分子具有下式:

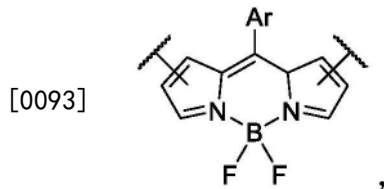


[0089] 如本文所述,在某些实施方案中,A是硼二吡咯亚甲基 (BODIPY) 荧光团。在某些实施方案中,A具有下式:



[0091] 其中该结构任选地在任意位置被取代。

[0092] 在某些实施方案中,A具有下式:



[0094] 其中Ar是任选取代的芳基或任选取代的杂芳基;并且其中该结构任选地在任意位置被取代。

[0095] 如本文所定义,Ar是任选取代的芳基或任选取代的杂芳基。在某些实施方案中,Ar是任选取代的芳基。在某些实施方案中,Ar是任选取代的杂芳基。在某些实施方案中,Ar是

任选取代的苯基。在某些实施方案中,A是聚氟苯基。在某些实施方案中,Ar具有式:

在某些实施方案中,Ar具有式:

在某些实施方案中,Ar具有式:

某些实施方案中,Ar具有式:

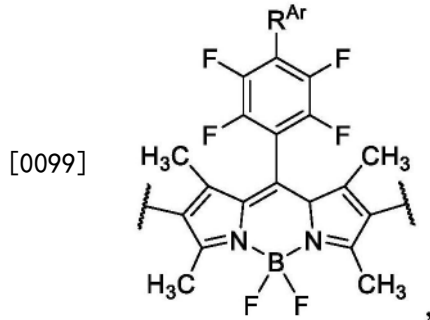
其中m如本文所定义。

[0096] 如本文所定义的, R^{Ar} 是氢、卤素、 $-N_3$ 、 $-CN$ 、 $-NO_2$ 、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、 $-OR^0$ 、 $-SR^S$ 或 $-N(R^N)_2$ 。在某些实施方案中, R^{Ar} 是氢。在某些实施方案中, R^{Ar} 是卤素($-Cl$ 、 $-I$ 、 $-Br$ 、 $-F$)。在某些实施方案中, R^{Ar} 是 $-N_3$ 。在某些实施方案中, R^{Ar} 是 $-CN$ 。在某些实施方案中, R^{Ar} 是 $-NO_2$ 。在某些实施方案中, R^{Ar} 是任选取代的烷基。在某些实施方案中, R^{Ar} 是任选

取代的烯基。在某些实施方案中, R^{Ar} 是任选取代的炔基。在某些实施方案中, R^{Ar} 是任选取代的碳环基。在某些实施方案中, R^{Ar} 是任选取代的芳基。在某些实施方案中, R^{Ar} 是任选取代的杂环基。在某些实施方案中, R^{Ar} 是任选取代的杂芳基。在某些实施方案中, R^{Ar} 是 $-OR^0$ 。在某些实施方案中, R^{Ar} 是 $-SR^S$ 。在某些实施方案中, R^{Ar} 是或 $-N(R^N)_2$ 。在某些实施方案中, R^{Ar} 是 $-F$ 。在某些实施方案中, R^{Ar} 是 $-S(CH_2CH_2O)_mOCH_3$, 其中 m 如本文所定义。

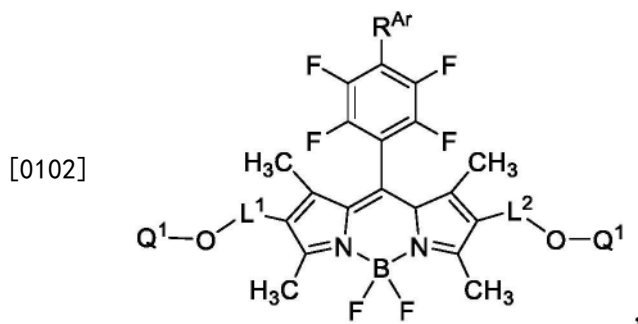
[0097] 如本文所定义, m 是 1 至 6 的整数, 包括 1 和 6。

[0098] 在某些实施方案中, A 具有下式:



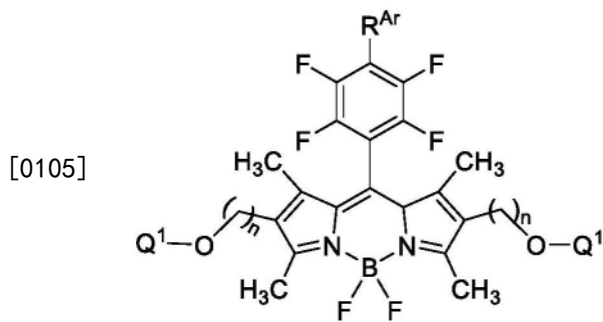
[0100] 其中 R^{Ar} 是氢、卤素、 $-N_3$ 、 $-CN$ 、 $-NO_2$ 、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、 $-OR^0$ 、 $-SR^S$ 或 $-N(R^N)_2$ 。在某些实施方案中, 该结构在任意位置被取代。在某些实施方案中, 该结构是未取代的。

[0101] 因此, 在某些实施方案中, 带标记的生物分子具有下式:

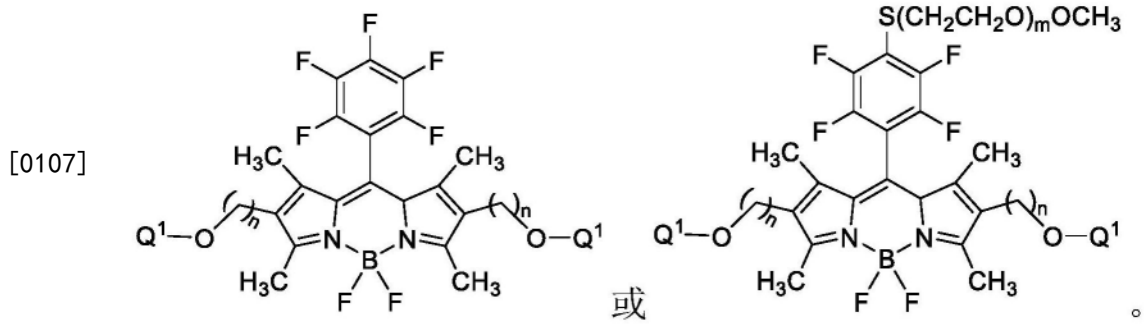


[0103] 其中 R^{Ar} 是氢、卤素、 $-N_3$ 、 $-CN$ 、 $-NO_2$ 、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、 $-OR^0$ 、 $-SR^S$ 或 $-N(R^N)_2$ 。

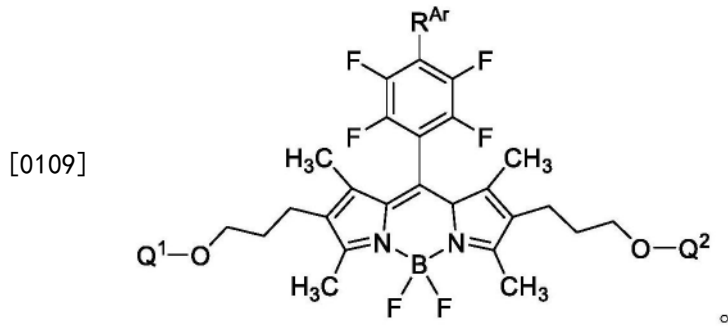
[0104] 在某些实施方案中, 带标记的生物分子具有下式:



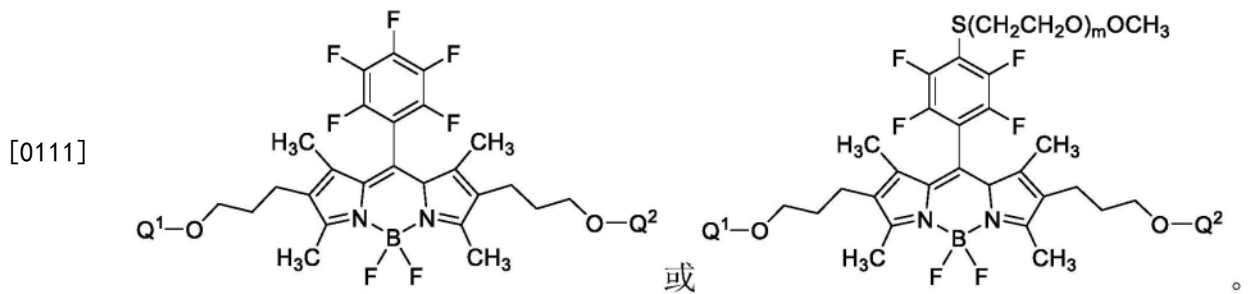
[0106] 在某些实施方案中, 带标记的生物分子具有下式之一:



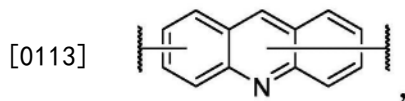
[0108] 在某些实施方案中,带标记的生物分子具有下式:



[0110] 在某些实施方案中,带标记的生物分子具有下式之一:

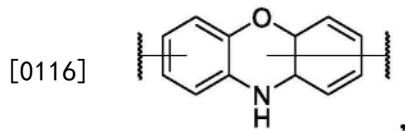


[0112] 如本文所述,在某些实施方案中,A是吲哚荧光团。在某些实施方案中,A是下式的吲哚荧光团:



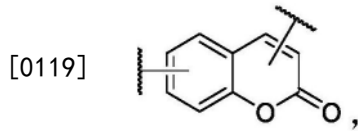
[0114] 其中该结构任选地被取代。

[0115] 如本文所述,在某些实施方案中,A是吩噁嗪荧光团。在某些实施方案中,A是下式的吩噁嗪荧光团:

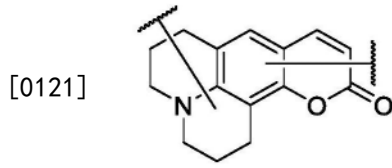


[0117] 其中该结构是任选取代的。


[0118] 如本文所述,在某些实施方案中,A是香豆素荧光团。在某些实施方案中,A是下式的香豆素荧光团:

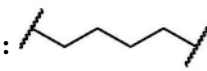
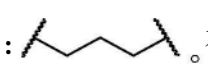


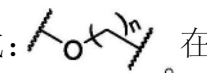
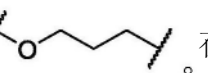
[0120] 其中该结构任选地被取代。在某些实施方案中,A是下式的香豆素荧光团:

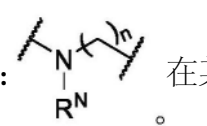
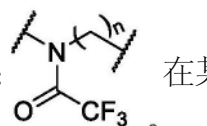


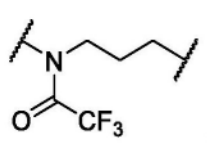
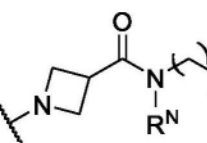
[0122] 其中该结构任选地被取代。

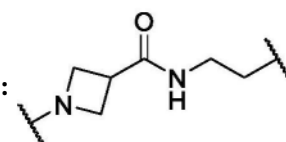
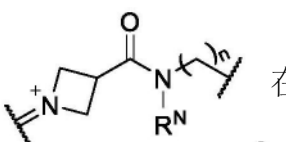
[0123] 如本文所定义的, L^1 和 L^2 独立地是选自以下的连接基:任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合。在某些实施方案中, L^1 包括任选取代的亚烷基。在某些实施方案中, L^1 包括任选取代的亚烯基。在某些实施方案中, L^1 包括任选取代的亚炔基。在某些实施方案中, L^1 包括任选取代的亚杂烷基。在某些实施方案中, L^1 包括任选取代的亚杂烯基。在某些实施方案中, L^1 包括任选取代的亚杂炔基。在某些实施方案中, L^1 包括任选取代的亚碳环基。在某些实施方案中, L^1 包括任选取代的亚杂环基。在某些实施方案中, L^1 包括任选取代的亚芳基。在某些实施方案中, L^1 包括任选取代的亚杂芳基。在某些实施方案中, L^1 是任选取代的 C_{1-20} 亚烷基。在某些实施方案中, L^1 是任选取代的 C_{1-10} 亚烷基。在某些实施方案中, L^1 是任选取代的 C_{1-6} 亚烷基。在某些实施方案中, L^1 具有式:  , 其中n如本文所定义。在某些实施方案

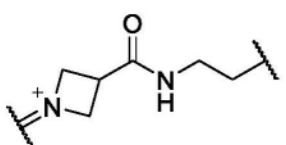
中, L^1 具有式: 。在某些实施方案中, L^1 具有式: 。在某些实施方案

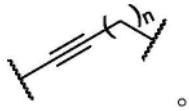
中, L^1 具有式: 。在某些实施方案中, L^1 具有式: 。在某些实施方案


中, L^1 具有式: 。在某些实施方案中, L^1 具有式: 。在某些实施方案中, L^1 具

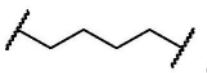
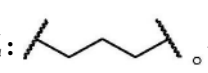
有式: 。在某些实施方案中, L^1 具有式: 。在某些实施方案

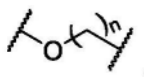
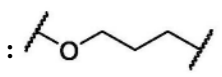
中, L^1 具有式: 。在某些实施方案中, L^1 具有式: 。在某

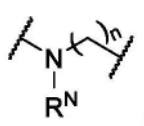
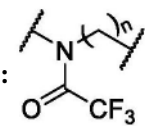
些实施方案中, L^1 具有式: 。在某些实施方案中, L^1 具有式:

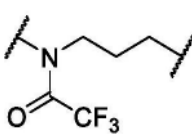
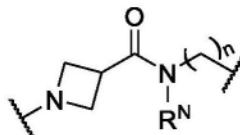


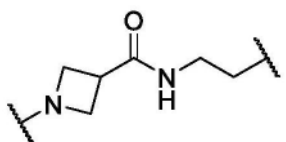
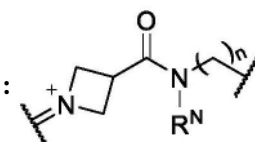
[0124] 在某些实施方案中, L^2 包括任选取代的亚烷基。在某些实施方案中, L^2 包括任选取代的亚烯基。在某些实施方案中, L^2 包括任选取代的亚炔基。在某些实施方案中, L^2 包括任选取代的亚杂烷基。在某些实施方案中, L^2 包括任选取代的亚杂烯基。在某些实施方案中, L^2 包括任选取代的亚杂炔基。在某些实施方案中, L^2 包括任选取代的亚碳环基。在某些实施方案中, L^2 包括任选取代的亚杂环基。在某些实施方案中, L^2 包括任选取代的亚芳基。在某些实施方案中, L^2 包括任选取代的亚杂芳基。在某些实施方案中, L^2 是任选取代的 C_{1-20} 亚烷基。在某些实施方案中, L^2 是任选取代的 C_{1-10} 亚烷基。在某些实施方案中, L^2 是任选取代的 C_{1-6} 亚烷基。在某些实施方案中, L^2 具有式: , 其中 n 如本文所定义。在某些实施方案中, L^2 具有

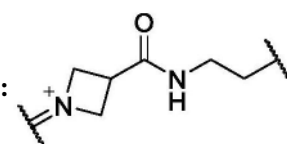
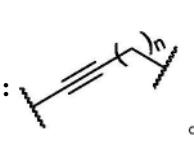
式: 。在某些实施方案中, L^2 具有式: 。在某些实施方案中, L^2 具有

式: 。在某些实施方案中, L^2 具有式: 。在某些实施方案中, L^2 具有

式: 。在某些实施方案中, L^2 具有式: 。在某些实施方案中, L^2 具有式:

。在某些实施方案中, L^2 具有式: 。在某些实施方案中, L^2 具

有式: 。在某些实施方案中, L^2 具有式: 。在某些实施

方案中, L^2 具有式: 。在某些实施方案中, L^2 具有式: 。

[0125] 如本文所定义, n 独立地是 1 至 20 的整数, 包括 1 和 20。在某些实施方案中, n 是 1。在某些实施方案中, n 是 2。在某些实施方案中, n 是 3。在某些实施方案中, n 是 4。在某些实施方案中, n 是 5。在某些实施方案中, n 是 6。在某些实施方案中, n 是 7。在某些实施方案中, n 是 8。在某些实施方案中, n 是 9。在某些实施方案中, n 是 10。在某些实施方案中, n 是 11。在某些实施方案中, n 是 12。在某些实施方案中, n 是 13。在某些实施方案中, n 是 14。在某些实施方案中, n 是 15。在某些实施方案中, n 是 16。在某些实施方案中, n 是 17。在某些实施方案中, n 是 18。在某些实施方案中, n 是 19。在某些实施方案中, n 是 20。

[0126] 如本文所述, Q^1 和 Q^2 独立地是单体或低聚生物分子。如本文所述, Q^1 和 Q^2 一起形成寡聚或聚合生物分子, 该生物分子被 $-O-L^1-A-L^2-O-$ 打断, 从而在内部被标记 (例如, 形成带标记的生物分子)。

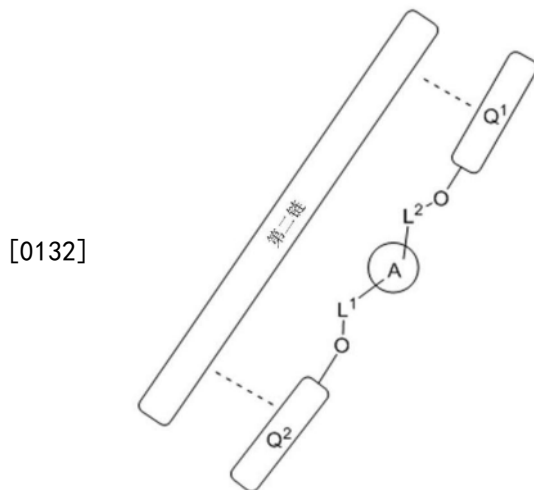
[0127] 在一些实施方案中,带标记的生物分子是包括至少5个单体生物分子(例如,至少5个核苷酸、至少5个氨基酸、至少5个单糖)的寡聚或聚合生物分子。在一些实施方案中,寡聚或聚合生物分子包括至少10个单体生物分子。在一些实施方案中,寡聚或聚合生物分子包括至少10个且少于200个单体生物分子。例如,在一些实施方案中,寡聚或聚合生物分子包括至少10个且少于150个单体生物分子,至少10个且少于100个单体生物分子,至少10个且少于50个单体生物分子,至少10个且少于40个单体生物分子,至少10个且少于30个单体生物分子,或至少10个且少于20个单体生物分子。

[0128] 在某些实施方案中,带标记的生物分子是寡核苷酸或核酸。

[0129] 在某些实施方案中, Q^1 和 Q^2 独立地是核苷、核苷酸、寡核苷酸、核酸或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 独立地是核苷或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 独立地是核苷酸衍生物或其片段。在某些实施方案中, Q^1 和 Q^2 独立地是寡核苷酸或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 独立地是核酸或其衍生物或片段。

[0130] 在某些实施方案中, Q^1 和 Q^2 独立地是脱氧核糖核酸、核糖核酸、肽核酸、锁核酸或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 独立地是脱氧核糖核酸或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 独立地是核糖核酸或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 独立地是肽核酸或其衍生物或片段。在某些实施方案中, Q^1 和 Q^2 是锁核酸或其衍生物或片段。

[0131] 如本文所述, Q^1 和 Q^2 一起形成寡聚或聚合生物分子,该生物分子被 $-O-L^1-A-L^2-O-$ 打断,从而在内部被标记(例如,形成带标记的生物分子)。在某些实施方案中,带标记的生物分子是单链核酸。在这种情况下,该单链核酸包括第一寡核苷酸链(Q^1 和/或 Q^2)。在某些实施方案中,带标记的生物分子包括与第一寡核苷酸链杂交的第二寡核苷酸链。例如,在某些实施方案中,所述第二寡核苷酸链与 Q^1 和/或 Q^2 杂交。这种内部标记体系的视觉表示如下(其中----表示杂交相互作用(例如,一个或多个Watson-Crick碱基相互作用)):



[0133] 该体系的视觉表示也可见于图1D。

[0134] 内部标记的寡核苷酸的其他实例示于图1E-1G。如本文描述的这些和其他例子所示,提供了各种寡核苷酸链杂交策略,以促进刚性和/或屏蔽内部标记物免受大量溶剂的影响。相应地,寡核苷酸链杂交可以作为制备本申请的带标记的寡核苷酸的一般性设计策略。在一些实施方案中,寡核苷酸链杂涉及自链杂交(例如,在单链内自杂交)。在一些实施方案中,寡核苷酸链杂涉及不同寡核苷酸链的杂交。

[0135] 图1D是一种带标记的寡核苷酸,其包括内部标记的寡核苷酸链,该内部标记的寡核苷酸链与未标记的寡核苷酸链杂交。在一些实施方案中,使用未标记的寡核苷酸链来增加带标记的寡核苷酸链的特定区域(例如,包含内部标记物的区域)的刚性。在一些实施方案中,将未标记的寡核苷酸链与本文提供的内部标记的寡核苷酸链杂交,该内部标记的寡核苷酸链包含两个或更多个内部标记物。

[0136] 图1E是一种带标记的寡核苷酸,其包括一条内部标记的寡核苷酸链,该内部标记的寡核苷酸链与另一条内部标记的寡核苷酸链杂交。在一些实施方案中,一条寡核苷酸链的内部标记物包括与另一条寡核苷酸链的内部标记物相同的荧光团。在一些实施方案中,一条寡核苷酸链的内部标记物包括与另一条寡核苷酸链的内部标记物不同的荧光团。在一些实施方案中,内部标记的寡核苷酸链之一包括两个或更多个本申请的内部标记物。在一些实施方案中,两条内部标记的寡核苷酸链都包括两个或更多个本申请的内部标记物。

[0137] 在一些实施方案中,寡核苷酸链杂交能够促进一种或多种结构基序的形成,例如茎环、结点、假节点和双螺旋。根据本申请,结构基序可用于增强带标记的寡核苷酸的刚性和/或限制内部标记物暴露于大量溶剂的程度。图1F-1G描述了带标记的寡核苷酸的实例,这些带标记的寡核苷酸具有通过链杂交形成的更高阶结构基序。

[0138] 图1F是一种带标记的寡核苷酸,其包含未标记的寡核苷酸链,该未标记的寡核苷酸链与包含两个内部标记物的寡核苷酸链杂交。如本实例所概括显示的,双螺旋的形成可以促进同一寡核苷酸链的两个内部标记物的分离。在一些实施方案中,杂交的寡核苷酸链形成双螺旋,每圈具有约10至12个碱基对。

[0139] 因此,在一些实施方案中,当同一寡核苷酸链的两个内部标记物在链内占据与一个核苷酸大约相同的空间时,内部标记物沿链最少分离5至6个核苷酸,使得标记物在双螺旋的大约相反的两侧。在一些实施方案中,内部标记物沿寡核苷酸链分隔4至8个(例如4、5、6、7或8个)核苷酸。不希望受理论约束,由于中间螺旋结构吸收了全部辐射和/或非辐射衰变,因此可以利用这样的设计策略来限制标记物-标记物相互作用(例如,淬灭效应)的程度。

[0140] 在一些实施方案中,将本公开的内部标记物整合到寡核苷酸骨架中,以使标记物-标记物相互作用的程度最小化。因此,在一些实施方案中,内部标记物沿着相同寡核苷酸链被1至3个(例如1、2或3个)或9至13个(例如9、10、11、12或13个)核苷酸隔开。如该实例所示,可以通过考虑预测的或已知的螺旋结构来设计带标记的寡核苷酸,从而可以操纵一个内部标记物相对于另一个内部标记物穿过空间的相对位置,以达到所需的应用。在设计带标记的寡核苷酸时可以使用的结构基序的其它实例是本技术中已知的,并在本文描述。

[0141] 图1G是一种内部标记的寡核苷酸链,其自杂交,形成了茎环基序。茎环,或发夹环,是寡核苷酸链上的未配对核苷酸环,当寡核苷酸链折叠并与同一条链的另一段形成碱基对时形成。在一些实施方案中,茎环的未配对环包括3至10个核苷酸。因此,茎环可以由具有反向互补序列的寡核苷酸链的两个区域形成,其中所述反向互补序列杂交形成茎,其中这两个区域被形成未配对环的3至10个核苷酸分开。在一些实施方案中,茎可以设计成具有一个或多个G/C核苷酸,与A/T/U核苷酸相比,G/C核苷酸可以通过形成额外的氢键相互作用提供额外的稳定性。在一些实施方案中,茎包括紧邻未配对环序列的G/C核苷酸。在一些实施方案中,茎包括与未配对环序列相邻的前2、3、4或5个核苷酸内的G/C核苷酸。

[0142] 如本文所述,在一些实施方案中,内部标记物将寡核苷酸链的一部分与寡核苷酸链的另一部分缀合。如图1D-1G的示例结构中一般描绘的那样,在一些实施方案中,由内部标记物共轭的寡核苷酸链的两部分都与同一寡核苷酸链杂交。因此,内部标记物可以与杂交的寡核苷酸链的一个或多个(例如1、2、3、4、5或更多)未配对碱基相邻。

[0143] 在一些实施方案中,内部标记物可以通过辐射性和/或非辐射性衰变与鸟嘌呤核碱基相互作用,以影响减弱的发光寿命。在一些实施方案中,与内部标记物相邻的一个或多个未配对碱基被设计成排除鸟嘌呤或使鸟嘌呤降至最低。在一些实施方案中,围绕内部缀合标记物的区域被设计成排除G/C含量或使G/C降至最低。在一些实施方案中,内部缀合标记物与寡核苷酸链上的G或C核苷酸至少相隔2个核苷酸(例如,与G或C核苷酸相隔2、3、4、5、6、7、8、9或超过10个核苷酸)。因此,在一些实施例中,每个内部标记物的两侧至少有2个选自A或T/U的连续核苷酸。

[0144] 本文提供的标记物在寡核苷酸和核酸以外的系统中也具有应用。在某些实施方案中,带标记的生物分子是多肽或蛋白质。例如,在某些实施方案中, Q^1 和 Q^2 独立地是氨基酸、寡肽、多肽、蛋白质或其片段。在某些实施方案中, Q^1 和 Q^2 独立地是氨基酸。在某些实施方案中, Q^1 和 Q^2 独立地是寡肽或其片段。在某些实施方案中, Q^1 和 Q^2 独立地是多肽或蛋白质或其片段。在某些实施方案中, Q^1 和 Q^2 连接在一起形成环状肽或环状蛋白质。

[0145] 适用于本申请的带标记的生物分子的寡肽和多肽的非限制性实例包括但不限于寡肽、环肽和小蛋白(例如,基于禽胰肽的微型蛋白,如Hodges、A.M. and Schepartz、A. (2007) *J. Am. Chem. Soc.* 129:11024-11025中所描述的)。在多肽中工程化设计结构约束(structural constraint)的方法是本领域众所周知的,并且被认为是特别有用的,例如,赋予刚性并增强本文所述的一种或多种发光性质。例如,可以修饰肽氨基酸序列的脯氨酸含量,从而控制肽的形状并赋予刚性(参见,例如,Kritzer, J.A., et al. (2006) *ChemBioChem* 7:29-31)。可使用的肽工程技术的其他非限制性实例包括肽环化(参见,例如Maltsev、O.V., et al. (2016) *Angewandte Chemie* 55(4):1535-1539)、通过钉合和/或H键替代物的 α -螺旋肽约束(参见,例如Douse、C.H., et al. (2014) *ACS Chem. Biol.* 9:2204-2209)、通过环状 β -折叠和 β -发夹模拟物的肽约束(参见,例如Gibbs、A.C., et al. (1998) *Nat. Struct. Biol.* 5:284-288)。

[0146] 在某些实施方案中,带标记的生物分子是寡糖或多糖。例如,在某些实施方案中, Q^1 和 Q^2 独立地是单糖、寡糖、多糖或其片段。在某些实施方案中, Q^1 和 Q^2 独立地是单糖。在某些实施方案中, Q^1 和 Q^2 独立地是寡糖或其片段。在某些实施方案中, Q^1 和 Q^2 独立地是多糖或其片段。适用于本申请的带标记的生物分子的寡糖和多糖的实例是本领域已知的(例如,如Solid Support Oligosaccharide Synthesis and Combinatorial Carbohydrate Libraries, Wiley 2001中所述)。

[0147] 根据本申请,本文提供的带标记的生物分子包括充当刚性支架的生物分子,在其上掺入了本申请的内部标记物。在一些实施方案中,生物分子包含一种或多种特征,这一种或多种特征赋予可以在各种检测、定量分析和成像方法中使用的附加功能。例如,图2A至图2C示出了内部标记的生物分子的非限制性实例,所述内部标记的生物分子包括与目标靶分子相互作用的生物分子。

[0148] 在一些实施方案中,包含寡核苷酸的带标记的生物分子可以用作杂交探针。杂交

探针是DNA或RNA的带标记片段(例如,寡核苷酸),可以将其添加到已知或未知含量的样品中,以检测与杂交探针互补的目标靶核酸的存在。例如,图2A描绘了与靶核酸210杂交的内部标记的杂交探针200。如该实例概括说明的,内部标记物的杂交探针200与靶核酸210形成碱基配对相互作用,导致来自内部标记物的可检测的发光增加。然而,在一些实施方案中,内部标记的杂交探针200与靶核酸210的杂交导致可检测的发光降低。

[0149] 在一些实施方案中,本申请的内部标记的杂交探针包含与靶核酸序列基本互补的序列,使得该探针和靶标在杂交条件下形成碱基配对相互作用。如本文所述,内部标记物使生物分子的一部分与生物分子的另一部分缀合。因此,被内部标记物缀合的生物分子(例如,寡核苷酸)的一个或两个部分可以被设计成与靶核酸序列杂交。在一些实施方案中,靶核酸包含RNA(例如,mRNA)。在一些实施方案中,靶核酸包含DNA(例如,cDNA、基因组DNA或其片段)。

[0150] 本申请的内部标记的杂交探针可以在本领域已知的利用杂交探针的任何方法中使用。此类技术的示例包括但不限于荧光原位杂交(FISH)、Northern印迹、Southern印迹以及涉及SNP检测、实时核酸检测、实时PCR定量、等位基因判别和鉴定、多重PCR的通用技术分析和临床诊断分析。

[0151] 本申请的带标记的生物分子,在一些实施方案中,包括充当蛋白质配体的生物分子。例如,图2B描绘了被靶蛋白211结合的带标记的生物分子201。如该实例概括说明的,靶蛋白211与带标记的生物分子201的生物分子(如虚线所示)至少一部分缔合(例如,结合),导致来自内部标记物的可检测的发光增加。然而,在一些实施方案中,靶蛋白211与带标记的生物分子201的结合会导致可检测的发光降低。

[0152] 在一些实施方案中,带标记的生物分子201和靶蛋白211包括已知的结合对。在一些实施方案中,可将带标记的生物分子201加入到已知或未知含量的样品中,以检测靶蛋白211的存在。在一些实施方案中,靶蛋白211是受体,带标记的生物分子201包括受体配体。在一些实施方案中,靶蛋白211是对带标记的生物分子201的至少一部分具有特异性的抗体。在一些实施方案中,靶蛋白211是抗体,带标记的生物分子201包括抗原。在一些实施方案中,靶蛋白211是核酸结合蛋白(例如,DNA结合蛋白),带标记的生物分子201包括核酸。据预测,这种带标记的生物分子在采用带标记的蛋白质配体来检测靶蛋白的存在或评估蛋白质-配体结合相互作用的方法学(例如,荧光偏振和本技术中已知或本文描述的其他技术)中是有用的。

[0153] 在一些实施方案中,包含多肽的带标记的生物分子可以用作抗体。图2C描绘了与靶蛋白212结合的内部标记的抗体202。如本实例所概括描述的,内部标记的抗体202包括一个Fab区域,该Fab区域被配置为特异性结合靶蛋白212,导致来自内部标记物的可检测的发光增加。然而,在一些实施方案中,内部标记的抗体202与靶蛋白212的结合会导致可检测的发光降低。

[0154] 本申请的内部标记的抗体可以用于本领域中利用发光标记抗体的已知方法。此类技术的实例包括但不限于荧光原位杂交(FISH)、Western印迹、涉及免疫标记的一般方法如免疫细胞化学和免疫组织化学技术以及本技术中已知的或本文描述的其它技术。

[0155] 如上所述,本申请的内部标记物可以与生物分子缀合,所述生物分子与所需靶分子相互作用。在一些实施方案中,内部标记物与蛋白质缀合,所述蛋白质与靶配体缔合(例

如结合)。在一些实施方案中,蛋白质是抗体或抗体的抗原结合部分。在一些实施方案中,蛋白质是酶,例如肽酶(例如,外肽酶或内肽酶)、核糖酶、适体酶(apozyme)、连接酶、转移酶或tRNA合成酶。在一些实施方案中,内部标记物与核酸缀合,所述核酸与靶配体缔合(例如结合)。在一些实施方案中,核酸是核酸适体(例如,DNA适体、RNA适体或其衍生物或类似物)。

[0156] 在一些实施方案中,带标记的生物分子包含被一个或多个功能性部分修饰的生物分子。例如,图3A-3B示出了带标记的生物分子的非限制性实例,所述标记的生物分子包含与靶分子和/或内部标记物相互作用的部分。

[0157] 根据本申请,生物分子支架提供了刚性标记支架,这对于需要内部标记物牢固限定位置的技术特别有益。例如,Förster共振能量转移(FRET)和荧光相关谱(FCS)已成为用于体外和体内研究生物分子构象动力学的重要工具。这些方法依赖于通过荧光受体荧光团(FRET)或非荧光猝灭剂,例如在光致电子转移(PET)FCS中使用的那样,对供体荧光团的荧光信号的距离依赖性猝灭。

[0158] 在一些实施方案中,带标记的生物分子包含与该带标记的生物分子的内部标记物相互作用的一个或多个猝灭部分(例如,荧光和/或非荧光猝灭部分)。在一些实施方案中,这样的部分可用于实时PCR,其中相对于被核酸外切酶活性切割的猝灭剂,内部标记物的位置被很好地定义。该处理的示例示于图3A。

[0159] 如图I所示,内部标记的杂交探针与靶核酸杂交,所述内部标记的杂交探针包含猝灭部分300。在一些实施方案中,猝灭部分300是非荧光猝灭剂,它吸收来自内部标记物的发射光。在一些实施方案中,猝灭部分300是荧光猝灭剂,它吸收来自内部标记物的一个波长的发射光并以另一波长发射。

[0160] 如图II所示,猝灭部分300已从内部标记的杂交探针上切割下来(例如,通过核酸外切酶)。猝灭部分300与内部标记物的这种分离可以消除距离依赖性猝灭效应,从而允许检测到来自内部标记物的发光。应当理解,在一些实施方案中,内部标记的杂交探针包括内部标记物,该内部标记物用作杂交探针的另一标记物的猝灭剂。

[0161] 被猝灭部分修饰的杂交探针是本领域已知的,并且预期可以与本申请的内部标记物一起使用。此类杂交探针的实例包括但不限于分子信标、TaqMan探针、激子控制的杂交敏感荧光寡核苷酸(ECHO)探针和循环探针技术(CPT)探针。

[0162] 因此,在一些实施方案中,可以根据本文提供的带标记的生物分子的所需发光性质,将多个内部标记物(例如,两个、三个、四个、五个或更多个内部标记物)掺入到生物分子中。例如,在一些实施方案中,相对于具有一个内部标记物的生物分子,具有两个或更多个内部标记物的带标记的生物分子的发光强度和/或亮度增加。在一些实施例中,这两个或更多个内部标记物被配置为提供独立的报告信号。在一些实施方案中,这两个或更多个内部标记物被配置为提供相关的报道子信号(例如,FRET的供体标记物和受体标记物对)。

[0163] 在一些方面,本申请提供了被配置用于常规固相合成技术的内部标记物,例如可用于寡核苷酸合成的亚磷酰胺类似物。因此,在一些实施方案中,本文提供的标记物可以容易地掺入到生物分子中,以产生具有一定数量的内部标记物的带标记的生物分子,内部标记物的数量仅受所需的生物分子大小限制。例如,在一些实施方案中,带标记的生物分子包含两个或更多个(例如2、3、4、5、6、7、8、9、10或更多个)内部标记物。在一些实施方案中,带

标记的生物分子包括2至5、2至10、5至10、5至15、10至15、15至20或更多个内部标记物。

[0164] 在一些实施方案中,本文提供的带标记的生物分子还包括除式(I)的多环荧光团以外的一个或多个发光标记物。例如,在一些实施方案中,带标记的生物分子包括至少一个根据式(I)的内部标记物,和一个或多个内部标记物,该一个或多个内部标记物包括线性或非多环荧光团。在一些实施方案中,带标记的生物分子包括至少一个根据式(I)的内部标记物,和一个或多个外部标记物。

[0165] 在一些实施方案中,外部标记物是指与本文提供的带标记的生物分子的单个位点缀合的标记物(例如,荧光团)。在一些实施方案中,外部标记物在带标记的生物分子的末端缀合。例如,在一些实施方案中,外部标记物与带标记的寡核苷酸的5'或3'端缀合。在一些实施方案中,外部标记物与带标记的多肽的N-或C-末端缀合。在一些实施方案中,外部标记物在生物分子的末端单体上与带标记的生物分子缀合(例如,与寡核苷酸链中末端核苷酸的碱基缀合,与多肽链中末端氨基酸的侧链缀合)。

[0166] 在一些实施方案中,外部标记物在生物分子的非末端位点与带标记的生物分子缀合。在一些实施方案中,外部标记物在生物分子的单体之间的位点上与带标记的生物分子缀合。在一些实施方案中,外部标记物与带标记的寡核苷酸的无碱基位点缀合。在一些实施方案中,外部标记物在生物分子的非末端单体上与带标记的生物分子缀合(例如,与寡核苷酸链中非末端核苷酸的碱基缀合,与多肽链中非末端氨基酸的侧链缀合)。

[0167] 在一些实施方案中,本申请的带标记的生物分子包括被一个或多个部分修饰的生物分子,其中该一个或多个部分充当蛋白质配体。例如,图3B阐明了一种方法,其中包含配体部分的带标记的生物分子被靶蛋白以可检测地方式结合。如图I所示,靶蛋白暴露于带标记的生物分子,该带标记的生物分子包含被配置为与靶蛋白结合的配体部分301。在靶蛋白和配体部分301之间不存在结合的情况下,带标记的生物分子的内部标记物不发出可检测的信号。

[0168] 如图II所示,靶蛋白与配体部分301缔合(例如结合),导致来自内部标记物的可检测的发光增加。然而,在一些实施方案中,靶蛋白与配体部分301的结合会导致可检测的发光降低。在一些实施方案中,结合时可检测的发光的变化是由于内部标记生物分子被约束于观察区域(例如,在表面上固定足以允许检测的时间段)的结果。在一些实施方案中,结合时发光的变化是由于FRET相互作用(例如,与靶蛋白或表面缀合的受体/供体)的结果。

[0169] 在一些实施方案中,包含一个或多个配体部分的带标记的生物分子,例如,可以用于将带标记的生物分子固定到表面或材料上的目的,用于检测已知或未知样品中的靶蛋白,用于检测(例如,定量)靶蛋白与配体部分之间的结合相互作用,其中该配体部分已知或未知与靶蛋白结合。

[0170] 在某些实施方案中,带标记的生物分子与反应剂缔合,该反应剂被配置为用作反应底物。例如,在某些实施方案中,式(I)的 Q^1 和 Q^2 独立地任选地与被配置为用作反应底物的反应剂缔合。例如,在寡核苷酸或核酸系统的情况下,第一和第二寡核苷酸链独立地任选地与被配置为用作反应底物的反应剂缔合。在一些实施方案中,反应剂被配置为用作聚合反应的底物。在一些实施方案中,当处于聚合反应条件下时,该反应剂被聚合酶从带标记的生物分子上切割下来。例如,在一些实施方案中,反应物是核苷酸(例如,用于核酸测序方法中)。

[0171] 带标记的核苷酸

[0172] 本文还提供了带标记的核苷酸,所述带标记的核苷酸包括与本文所述的带标记的生物分子缔合的一个或多个核苷酸。在一些实施方案中,所述一个或多个核苷酸包括选自鸟嘌呤、胞嘧啶、腺嘌呤和胸腺嘧啶或尿嘧啶的一种类型的核苷酸。在一些实施方案中,当处于聚合反应条件下时,该一个或多个核苷酸被聚合酶从带标记的生物分子上切割下来。

[0173] 不希望受任何特定理论的束缚,除了本文其他地方描述的优点之外,与测序反应中当前使用的那些核苷酸相比,本文提供的带标记的核苷酸具有许多明显的优点,例如阅读长度增加和准确性增加。图4A-4C突出了本公开的带标记的核苷酸的几个特征。

[0174] 图4A-4C中的每一个都描绘了与聚合酶410结合的核苷酸400。图4A的核苷酸400在外部与发光标记物缀合,而图4B和4C的核苷酸与本文描述的不同的带标记的生物分子缔合。如图4A所示,外部缀合的标记物相对靠近聚合酶,并且接近大量溶剂分子的程度相对较高。在一些实施方案中,这样的特性不利于聚合反应。例如,在一些实施方案中,标记物与聚合酶之间的较短距离可以通过辐射性和/或非辐射性衰变而导致标记物引起的对聚合酶损伤(如路径(i)所示)。在某些实施方案中,标记物接触大量溶剂分子的程度越高,可导致活性氧(ROS)形成的发生率越高。一旦形成,ROS就可以破坏聚合酶,从而不利地影响酶活性(显示为路径(ii))。

[0175] 图4B描绘了与带标记的生物分子,例如带标记的寡核苷酸链缔合的核苷酸。如相对于外部缀合的标记物所示,生物分子增加了标记物和聚合酶之间的分离。另外,核苷酸和内部标记物之间的生物分子部分在标记物和聚合酶之间提供了保护性屏障。因此,由于标记物-聚合酶的分离和/或由于生物分子吸收了从标记物发出的全部衰变(示为路径(iii)),因此可以减少标记物诱导的对聚合酶损伤的发生。

[0176] 此外,相对于图4A所示,将内部染料整合到图4B中所示的生物分子中降低了标记物暴露于大量溶剂分子的程度。因此,由于标记物接近大量溶剂的程度降低导致ROS形成的发生率降低,和/或由于生物分子吸收了全部由ROS引起的损伤,和/或由于游离自由基随着标记物-聚合酶分隔距离而衰变(如路径(iv)所示),因此由ROS引起的损伤降低了。

[0177] 图4C描绘了与带标记的寡核苷酸缔合的核苷酸。如图所示,带标记的寡核苷酸与未标记的寡核苷酸链杂交。根据本申请,这样的构建体赋予了高度的刚性,增强了针对图4B的带标记的核苷酸所描述的上述每个优点。

[0178] 例如,杂交链提供了增加的刚度,意味着总体柔韧性更少,这进一步促进了标记物-聚合酶的分离。杂交链还提供了标记物和聚合酶之间的另一道屏障,可以吸收任何标记物诱导的衰变(显示为路径(v))。另外,由于杂交链进一步限制了标记物与大量溶剂的接触,和/或由于杂交链吸收了任何由ROS引起的损害,和/或由于游离自由基随着标记物-聚合酶分隔距离的增加而衰变,因此由ROS引起的损伤进一步降低。

[0179] 因此,在上述每个实例中,通过限制光诱导的聚合酶损伤程度,本申请的带标记的生物分子所提供的优点在测序反应中增加了阅读长度。另外,例如,由于一种或多种发射特性的增强,本申请的带标记的核苷酸提供了提高的准确性。

[0180] 应该理解的是,在带标记的生物分子的上下文中,与其连接的“核苷酸”或“多聚核苷”是指被配置为掺入到增长的核酸链中(例如,在测序反应期间)的一个或多个核苷酸(例如,核苷多磷酸)。在一些实施方案中,该一个或多个核苷酸包括一个或多个核苷单磷酸或

核苷多磷酸。在一些实施方案中,多磷酸核苷的实例包括:二或三磷酸核苷,或具有多于三个的5'磷酸的核苷,例如六磷酸核苷。在本申请中描述的任何组合物或方法的一些实施方案中,核苷酸的磷酸酯部分(例如,多磷酸酯部分)包括一个或多个磷酸酯(例如1、2、3、4、5、6、7、8、9、10或更多个磷酸基团)或其变体。例如,在一些实施方案中,核苷酸的磷酸酯部分(例如,多磷酸酯部分)可包括磷酸酯、硫酸酯、氨基磷酸酯、烷基磷酸酯键、其他合适的键或多个这样的修饰,或其中两个或多个的组合。

[0181] 带标记的核苷酸可以是末端磷酸酯标记的核苷酸,使得本申请的带标记的生物分子与核苷酸的末端磷酸酯连接。例如,在一些实施方案中,可以通过末端磷酸酯将一个或多个核苷酸与形成如本申请中所述的带标记的生物分子的一部分的生物分子(例如,式(I)的 Q^1 和/或 Q^2)连接。因此,在一些实施方案中,本申请的“带标记的核苷酸”是指与式(I)的带标记的生物分子连接的核苷酸。在一些实施方案中,可以通过末端磷酸酯将一个或多个核苷酸与形成如本申请中所述的带标记的生物分子的一部分的寡核苷酸(例如,未标记的寡核苷酸链)连接。

[0182] 带标记的生物分子可以通过连接基与核苷酸的末端磷酸酯连接。连接基可包括例如至少一个或多个羟基基团、巯基基团、氨基基团或卤代烷基基团,其可能适合于在天然或修饰核苷酸的末端磷酸酯上形成例如磷酸酯、硫酸酯、氨基磷酸酯或烷基磷酸酯键。连接基可以是可裂解的,以便例如借助聚合酶将标记物与末端磷酸酯分离。核苷酸和连接基的实例在美国专利号7,041,812中提供,其通过引用整体并入本文。在一些实施方案中,连接基包括任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合。可用于将标记物连接至核苷酸的连接基的其他实例可见于共同待决的美国专利申请号15/600,979,其相关部分通过引用整体并入本文。

[0183] 核苷酸(例如,多磷酸核苷)可包括腺嘌呤(A)、胞嘧啶(C)、鸟嘌呤(G)、胸腺嘧啶(T)和尿嘧啶(U)或其变体中的任何一种。核苷酸(例如,多磷酸核苷)可包括甲基化的核碱基。例如,甲基化核苷酸可以是包含一个或多个与核碱基相连(例如,直接与核碱基的环相连,与核碱基的环的取代基相连)的甲基基团的核苷酸。示例性的甲基化核碱基包括1-甲基胸腺嘧啶、1-甲基尿嘧啶、3-甲基尿嘧啶、3-甲基胞嘧啶、5-甲基胞嘧啶、1-甲基腺嘌呤、2-甲基腺嘌呤、7-甲基腺嘌呤、N6-甲基腺嘌呤、N6,N6-二甲基腺嘌呤、1-甲基鸟嘌呤、7-甲基鸟嘌呤、N2-甲基鸟嘌呤和N2,N2-二甲基鸟嘌呤。

[0184] 如本文所用,术语“核酸”通常是指包含一个或多个核酸亚单位的分子。核酸可包含一个或多个选自腺嘌呤(A)、胞嘧啶(C)、鸟嘌呤(G)、胸腺嘧啶(T)和尿嘧啶(U)或其变体的亚单元。在一些实例中,核酸是脱氧核糖核酸(DNA)或核糖核酸(RNA)或其衍生物。在一些实施方案中,核酸是修饰的核酸,包括但不限于锁核酸(LNA)、肽核酸(PNA)、三唑连接的核酸、2'-F修饰的核酸,及其衍生物和类似物。核酸可以是单链或双链的。在一些实施方案中,核酸通常是指核苷酸的任意聚合物。

[0185] 发射特性

[0186] 如本文所述,在生物分子的内部缀合标记物可以改变标记物的光物理性质(例如,通过限制标记物在生物分子内部的旋转和固定)。因此,在某些实施方案中,相对于包括标

记物的未缀合分子,所述带标记的生物分子的一种或多种发射特性有所改变(例如增加)。如本文所述,包含标记物的“非缀合分子”不包含 Q^1 和 Q^2 之一或两者。内部缀合所改变的一个或多个发射特性可以包括但不限于发光寿命、发光强度、亮度、最大发射波长、发光量子产率和耐光性。

[0187] 在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光寿命有所增加。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光寿命增加了至少10%。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光寿命增加了约10%至50%(例如,约10%至25%,约10%至15%,约25%至50%,约40%至50%)。发光寿命增加的实例示于图5A-5B。

[0188] 根据图5A中概括性描述的构建体制备了一组包含核苷酸的带标记的生物分子。带标记的核苷酸(1)和(2)代表缀合物,其中Cy3B染料分别在寡核苷酸的末端或分支点与DNA接头外部缀合。相比之下,带标记的核苷酸(3)将寡核苷酸链中的一个碱基交换为Cy3B染料本身,从而内部整合到DNA骨架中。这些带标记的核苷酸用于了测序实验,并获得了每种缀合物的寿命测量值。外部缀合的Cy3B构建体(1)和(2)产生的寿命约为2.2纳秒,而内部缀合的Cy3B构建体(3)的实测寿命为2.6纳秒——内部缀合的染料的寿命增加了约15-20%。

[0189] 在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光强度有所增加。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光强度增加了约5%和25%(例如,约5%至20%,约5%至15%,约5%至10%,约10%至25%,约15%至25%,约20%至25%)。在一些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光强度增加了约5%,约10%,约15%,约20%,约25%或更多。

[0190] 在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的亮度有所增加。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的亮度增加了约5%至10%。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的亮度增加了约5%和25%(例如,约5%至20%,约5%至15%,约5%至10%,约10%至25%,约15%至25%,约20%至25%)。在一些实施方案中,相对于未缀合的分子,所述带标记的生物分子的亮度增加了约5%,约10%,约15%,约20%,约25%或更多。

[0191] 在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的最大发射波长增加了至少1%。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的最大发射波长增加了约1%和10%(例如,约1%至5%,约5%至10%)。最大发射波长增加的一个实例示于图6。在一些实施方案中,相对于未缀合的分子,所述带标记的生物分子的最大发射波长增加了约1%、约2%、约5%、约10%或更多。

[0192] 内部缀合的Cy3B和外部缀合的Cy3B所获得的大量荧光数据如图6所示。如图所示,相对于外部染料,内部染料的激发光谱红移了8nm,这使振动肩峰(轨迹左侧的驼峰)更靠近532nm。此外,如图所示,内部染料的发射光谱是红移的,这被发现有利地增加了可检测的信号,因为更多的光被允许通过过滤器。

[0193] 在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光量子产率有所增加。在某些实施方案中,相对于未缀合的分子,所述带标记的生物分子的发光量子产率增加了约5%和25%(例如,约5%至20%,约5%至15%,约5%至10%,约10%至25%,约15%至25%,约20%至25%)。在一些实施方案中,相对于未缀合的分子,所述带标记的生

物分子的发光量子产率增加了约5%，约10%，约15%，约20%，约25%或更多。

[0194] 在某些实施方案中，相对于未缀合的分子，带标记的生物分子的耐光性有所增加。如本文所用，在一些实施方案中，耐光性是指发光分子随时间持续发荧光的能力。在一些实施方案中，可以通过测量光漂白速率来评估耐光性。例如，在一些实施方案中，可以测量带标记的生物分子（例如，具有内部缀合的标记物的生物分子）的光漂白速率，并将其与未缀合的分子所测量的光漂白速率进行比较。测得的光漂白速率的降低将表明耐光性增加。测量光漂白速率的方法是本领域已知的，例如，如Wüstner, D., et al. (2014) *Molecules*, 9: 11096-11130; Brakenhoff, G. J., et al. (1994) *Journal of Microscopy*, 175(2): 154-161; 和Song, L., et al. (1995) *Biophys J.* 68(6): 2588-2600所述。在一些实施方案中，相对于未缀合的分子，带标记的生物分子的光漂白降低，如通过荧光漂白后恢复测定 (FRAP) 所测量的，例如，如Meyvis, T., et al. (1999) *Pharmaceutical Research*, 16(8): 1153-1162所述。在一些实施方案中，相对于未缀合的分子，带标记的生物分子的光漂白降低，如通过荧光漂白损失 (FLIP) 所测量的，例如，如Wüstner, D., et al. (2012) *BMC Bioinformatics*, 13: 296所述。

[0195] 在一些实施方案中，本公开提供了用于基于单分子的一种或多种发光性质来鉴定这些分子的新组合物。在一些实施方案中，基于分子的亮度、发光寿命、吸收光谱、发射光谱、发光量子产率、发光强度或其两种或更多种的组合来鉴定分子（例如，发光标记的核苷酸）。鉴定可能是指指定一个分子的确切分子身份，或者可能是指从一组可能的分子中辨别或区分特定的分子。在一些实施方案中，可以基于不同的亮度、发光寿命、吸收光谱、发射光谱、发光量子产率、发光强度或其两种或更多种的组合来将多个单分子彼此区分。在一些实施方案中，通过将单分子暴露于一系列独立的光脉冲，并评估从该分子发射的每个光子的时间或其他性质，来鉴定该分子（例如，与其他分子区别）。在一些实施例中，将从单分子顺序发射的多个光子的信息聚集并评估，从而鉴定该分子。在一些实施方案中，由从分子顺序发射的多个光子来确定分子的发光寿命，并且使用发光寿命来鉴定该分子。在一些实施方案中，由从分子顺序发射的多个光子来确定分子的发光强度，并且使用发光强度来鉴定该分子。在一些实施方案中，由从分子顺序发射的多个光子来确定分子的发光寿命和发光强度，并且使用发光寿命和发光强度来鉴定该分子。

[0196] 因此，在本申请的一些方面，使反应样品暴露于多个独立的光脉冲，并检测和分析一系列发射的光子。在一些实施方案中，这一系列发射的光子可以提供有关反应样品中存在且在实验时间内不会改变的单分子的信息。然而，在一些实施方案中，这一系列发射的光子提供有关反应样品中不同时间（例如，随着反应或过程的进行）存在的一系列不同分子的信息。

[0197] 分子的发光寿命的确定可以使用任何合适的方法来进行（例如，通过使用合适的技术来测量寿命，或通过确定发射的时间相关性）。在一些实施方案中，确定分子的发光寿命包括确定相对于一个或多个分子（例如，测序反应中不同的发光标记的核苷酸）的寿命。在一些实施方案中，确定分子的发光寿命包括确定相对于参照品的寿命。在一些实施方案中，确定分子的发光寿命包括测量寿命（例如，荧光寿命）。在一些实施方案中，确定分子的发光寿命包括确定指示寿命的一个或多个时间特性。在一些实施方案中，分子的发光寿命可以基于相对于激发脉冲在一个或多个时间窗内发生的多个发射事件（例如，1、2、3、4、

5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100或更多个发射事件)的分布来确定。例如,基于相对于激发脉冲所测量的光子到达时间的分布,可以将单分子的发光寿命与具有不同发光寿命的多个分子区分开。

[0198] 应当理解,单分子的发光寿命指示在单分子达到激发态之后发射的光子的时间,并且可以通过指示光子时间的信息来区分单分子。一些实施方案可以包括通过测量与分子发射的光子相关的时间,基于分子的发光寿命将分子与多个分子区分开。时间的分布可以提供发光寿命的指示,发光寿命可以从该分布确定。在一些实施方案中,基于时间的分布,例如通过将时间的分布与对应于已知分子的参考分布进行比较,可以将单分子与多个分子区分开。在一些实施例中,发光寿命的值由时间分布确定。

[0199] 如本文中用于单分子的那样,发光强度是指单位时间内由分子发射的发射光子的数量,该分子通过递送脉冲激发能而被激发。在一些实施方案中,发光强度是指单位时间内由分子发射的、并由一个或一组特定传感器检测到的发射光子的检测数量,该分子通过递送脉冲激发能而被激发。

[0200] 在一些方面,本公开提供了与具有增强的发射亮度的带标记的生物分子有关的方法和组合物。如本文所用,在一些实施方案中,“亮度”(及其变化词,例如,“明度”、“明亮度”等)是指报告每个带标记的反应物分子的平均发射强度的参数。因此,在一些实施方案中,“发射强度”可用于泛指包含亮标反应物的组合物的亮度。在一些实施方案中,带标记的反应物的亮度等于其量子产率与消光系数的乘积。在一些实施方案中,将本公开的带标记的生物分子设计为使量子产率最大化,以促进亮度增加。

[0201] 发光量子产率是指在给定的波长或在给定的光谱范围内导致发射事件的激发事件的分数的,并且通常小于1。在一些实施方案中,本文所述的分子的发光量子产率为0至约0.001,约0.001至约0.01,约0.01至约0.1,约0.1至约0.5,约0.5至0.9或约0.9至1。在一些实施方案中,通过确定或估计发光量子产率来鉴定分子。

[0202] 在一些实施方案中,本文所述的内部标记物允许将连续的发光标记物添加到带标记的生物分子中,以增加亮度和/或发光强度。在一些实施方案中,包含两个或多个发光标记物的内部标记的生物分子根据公式 $L_n(x)$ 表现出亮度和/或发光强度,其中 L_n 等于带标记的反应剂上的发光标记物的总数, x 等于相应的单一标记的反应物的测量亮度或荧光强度。因此,在一些实施方案中,双染料标记的反应组分的亮度和/或发光强度是单染料标记的类似物的两倍。在一些实施方案中,三或四染料标记的反应组分的亮度和/或发光强度分别是单染料标记的类似物的三倍或四倍。在一些实施方案中,本文所述的亮标反应剂表现出的亮度和/或发光强度是 $L_n(x)$ 预测值的至少70%、至少80%、至少90%、至少95%、至少98%或至少99%。

[0203] 核酸测序反应组合物

[0204] 本文还提供了核酸测序反应组合物,其在反应混合物中包括两种或更多种不同类型的带标记的核苷酸,其中至少一种类型的带标记的核苷酸是包括本文所述的带标记的生物分子的带标记的核苷酸。

[0205] 在一些实施方案中,核酸测序反应组合物包括两种或更多种(例如,两种、三种、四种、五种或更多种)不同类型的带标记的核苷酸。在一些实施方案中,核酸测序反应组合物包括四种不同类型的带标记的核苷酸。在一些实施方案中,所述四种不同类型的带标记的

核苷酸包括第一带标记的核苷酸、第二带标记的核苷酸、第三带标记的核苷酸和第四带标记的核苷酸,其中所述第一带标记的核苷酸包括鸟嘌呤,所述第二带标记的核苷酸包括胞嘧啶,所述第三带标记的核苷酸包括腺嘌呤,所述第四带标记的核苷酸包括胸腺嘧啶或尿嘧啶。

[0206] 在一些实施方案中,核酸测序反应组合物中的每种类型的带标记的核苷酸存在的浓度为约100至1000nM(例如约100至800nM、约150至700nM、约200至600nM或约250至500nM)。

[0207] 在一些实施方案中,核酸测序反应组合物包含测序模板,所述测序模板包含与靶核酸复合的聚合酶。在一些实施方案中,复合物还包含引物寡核苷酸,所述引物寡核苷酸具有与靶核酸的一部分互补的序列。在一些实施方案中,复合物存在的浓度为约10pM至10nM(例如约25pM至5nM、约50pM至2nM、约50pM至1nM、约50pM至500pM、约50pM至100pM、约250pM至5nM、约250pM至2nM、约250pM至1nM或约250pM至500pM)。

[0208] 在一些实施方案中,核酸测序反应组合物包含一种或多种缓冲剂(例如, MES、MOPS、MOPSO、HEPES、Tris、TAPS和本领域已知的其他此类合适的缓冲剂)。在一些实施方案中,该一种或多种缓冲剂包括MOPS。在一些实施方案中,缓冲剂存在的浓度为约25至100mM(例如约25至75mM、约50至75mM或约65mM)。

[0209] 在一些实施方案中,核酸测序反应组合物包括二价阳离子(例如,镁离子、钙离子)。在一些实施方案中,二价阳离子包括镁或钙盐(例如,包括镁或钙和乙酸盐、氯化物、磷酸盐、硫酸盐的盐)。在一些实施方案中,盐是乙酸镁。在一些实施方案中,二价阳离子存在的浓度为约5至50mM(例如、约10至40mM、约15至35mM、约20至30mM或约25mM)。

[0210] 在一些实施方案中,核酸测序反应组合物包含一种或多种一价盐(例如,钠盐或钾盐,如氯化钠、乙酸钠、氯化钾或乙酸钾)。在一些实施方案中,一价盐存在的浓度为约10至200mM(例如约25至150mM、约25至40mM、约50至150mM、约100至150mM)。在一些实施方案中,核酸测序反应组合物包括约40mM单价盐,例如40mM氯化钠。在一些实施方案中,核酸测序反应组合物包括约120mM单价盐,例如120mM乙酸钾。

[0211] 在一些实施方案中,核酸测序反应组合物包含一种或多种光稳定剂(例如,一种或多种光保护添加剂,如抗氧化剂、氧清除剂、三态淬灭剂和本领域中已知的类似的吸能添加剂)。在一些实施方案中,光稳定剂包括原儿茶酸(PCA)。在一些实施方案中,光稳定剂包括4-硝基苄醇(NBA)。在一些实施方案中,光稳定剂包括trolox或其衍生物。在一些实施方案中,光稳定剂存在的浓度为约0.1mM至约20mM。在一些实施方案中,Trolox的浓度为约5mM。在一些实施方案中,PCA的浓度为约3mM。在一些实施方案中,PCA的浓度为约8mM。在一些实施方案中,NBA的浓度为约3mM。含光稳定剂(例如,PCA)的混合物还可以包含用于使光稳定剂再生的酶(例如,原儿茶酸二氧酶(PCD))。在一些实施方案中,PCD的浓度为约0.3mM。在一些实施方案中,PCD的浓度为约0.5mg/mL。

[0212] 在一些实施方案中,核酸测序反应组合物包含一种或多种还原剂。例如,在一些实施方案中,核酸测序反应组合物包含约10至100mM DTT(例如,约40mM DTT)。

[0213] 测序

[0214] 本申请的一些方面可用于对于生物聚合物,例如核酸和蛋白质进行测序。在一些方面,本申请中描述的组合物和技术可用于鉴定一系列掺入到核酸或蛋白质中的核苷酸或

氨基酸单体(例如,通过检测掺入一系列带标记的核苷酸或氨基酸单体的时间过程)。在一些实施方案中,本申请中描述的组合物和技术可用于鉴定一系列掺入到由聚合酶合成的模板依赖性核酸测序反应产物中的核苷酸。

[0215] 因此,本文还提供了使用本申请的核酸测序反应组合物来确定模板核酸序列的方法。在一些实施方案中,测序方法包括以下步骤:(i)将目标体积的复合物暴露于本公开的核酸测序反应组合物(例如,至少一种包含本文所述的带标记的生物分子的带标记的核苷酸)中,其中所述复合物包括模板核酸、引物和聚合酶,(ii)将一系列具有一个或多个激发能的脉冲引向目标体积附近;(iii)在顺序掺入到包含引物的核酸中期间,检测从发光标记的核苷酸发射的多个光子;以及(iv)通过确定发射的光子的时间和任选的发光强度,来鉴定掺入的核苷酸的序列。

[0216] 在一些实施方案中,如本文所用,激发能是来自光源的光脉冲。在一些实施方案中,激发能在可见光谱中。在一些实施方案中,激发能在紫外线光谱中。在一些实施方案中,激发能在红外光谱中。在一些实施方案中,激发能处于或接近发光标记的分子(从该分子中要检测多个发射光子)的吸收最大值。在一些实施方案中,激发能为约500nm至约700nm(例如,约500nm至约600nm、约600nm至约700nm、约500nm至约550nm、约550nm至约600nm、约600nm至约650nm或约650nm至约700nm)。在一些实施方案中,激发能可以是单色的或限定在一个光谱范围内。在一些实施方案中,光谱范围为约0.1nm至约1nm、约1nm至约2nm或约2nm至约5nm。在一些实施方案中,光谱范围为约5nm至约10nm、约10nm至约50nm或约50nm至约100nm。

[0217] 当靶核酸的核碱基与互补多磷酸核苷(例如,dNTP)之间碱基配对时,聚合酶通过在新合成的核酸链的3'羟基端和dNTP的 α -磷酸酯之间形成磷酸二酯键,将dNTP掺入新合成的核酸链中。在与dNTP缀合的发光分子(例如,本文所述的带标记的生物分子)包含荧光团的实例中,它的存在通过激发发出信号,并且在掺入步骤期间和/或之后检测发射脉冲。对于与dNTP的末端(γ)磷酸酯共轭的发光分子(例如,带标记的生物分子),将dNTP掺入到新合成的链中会导致 β 和 γ 磷酸酯和发光分子的释放,在样品孔中自由扩散,导致从荧光团检测到的发射减少。

[0218] 在某些实施方案中,模板依赖性核酸测序产物通过天然存在的核酸聚合酶进行。在一些实施方案中,聚合酶是天然存在的聚合酶的突变体或修饰变体。在一些实施方案中,模板依赖性核酸测序产物将包含一个或多个与模板核酸链互补的核苷酸区段。在一个方面,本申请提供了一种通过确定其互补核酸链的序列来确定模板(或靶标)核酸链的序列的方法。

[0219] 如本文所用,术语“聚合酶”通常是指能够催化聚合反应的任何酶(或聚合酶)。聚合酶的实例包括但不限于核酸聚合酶、转录酶或连接酶。聚合酶可以是聚合作用酶。针对单分子核酸延伸的实施方案(例如,用于核酸测序)可以使用能够合成与靶核酸分子互补的核酸的任何聚合酶。在一些实施方案中,聚合酶可以是DNA聚合酶、RNA聚合酶、逆转录酶和/或其一种或多种的突变或改变形式。

[0220] 聚合酶的实例包括但不限于DNA聚合酶、RNA聚合酶、热稳定聚合酶、野生型聚合酶、修饰型聚合酶、大肠杆菌DNA聚合酶I、T7 DNA聚合酶、噬菌体T4 DNA聚合酶 ϕ 29(psi29) DNA聚合酶、Taq聚合酶、Tth聚合酶、Tli聚合酶、Pfu聚合酶、Pwo聚合酶、VENT聚合酶、

DEEPVENT聚合酶、EX-Taq聚合酶、LA-Taq聚合酶、Sso聚合酶、Poc聚合酶、Pab聚合酶、Mth聚合酶、ES4聚合酶、Tru聚合酶、Tac聚合酶、Tne聚合酶、Tma聚合酶、Tca聚合酶、Tih聚合酶、Tfi聚合酶、Platinum Taq聚合酶、Tbr聚合酶、Tfl聚合酶、Tth聚合酶、Pfutubo聚合酶、Pyrobest聚合酶、Pwo聚合酶、KOD聚合酶、Bst聚合酶、Sac聚合酶、Klenow片段、具有3'至5'核酸外切酶活性的聚合酶及其变体、修饰产物和衍生物。在一些实施方案中,聚合酶是单亚基聚合酶。DNA聚合酶的非限制性实例及其性质在DNA Replication 2nd edition, Kornberg and Baker, W.H. Freeman, New York, N.Y. (1991) 等中有详细描述。

[0221] 在另一方面,本申请提供了通过对多个核酸片段进行测序来对靶核酸进行测序的方法,其中靶核酸包括这些片段。在某些实施方案中,该方法包括将多个片段序列合并,以提供亲本靶核酸的序列或部分序列。在一些实施方案中,合并步骤由计算机硬件和软件执行。本文描述的方法可以允许对一组相关靶核酸,例如整个染色体或基因组进行测序。

[0222] 在测序期间,可以将聚合酶偶连(例如,连接)至靶核酸分子的引物位置。引物位置可以是与靶核酸分子的一部分互补的引物。作为另一种选择,引物位置是在靶核酸分子的双链段内提供的空隙或缺口。空隙或缺口的长度可以是0到至少1、2、3、4、5、6、7、8、9、10、20、30或40个核苷酸。缺口可以在双链序列的一条链中提供断裂,可以为聚合酶例如链置换聚合酶提供引物位置。

[0223] 在一些情况下,可以将测序引物退火到目标核酸分子上,该目标核酸分子可以或不固定在固体支持物上。固体支持物可以包括例如用于核酸测序的芯片上的样品孔(例如,纳米孔、反应室)。在一些实施方案中,测序引物可以固定到固体支持物上,靶核酸分子的杂交也将靶核酸分子固定到固体支持物上。在一些实施方案中,将聚合酶固定到固体支持物上,将可溶性引物和靶核酸与聚合酶接触。然而,在一些实施方案中,在溶液中形成包含聚合酶、靶核酸和引物的复合物,并将该复合物固定到固体支持物上(例如,通过聚合酶、引物和/或靶核酸的固定)。在一些实施方案中,样品孔(例如,纳米孔、反应室)中的任何组分都不固定到固体支持物上。例如,在一些实施方案中,在溶液中形成包含聚合酶、靶核酸和引物的复合物,并且该复合物不固定到固体支持物上。

[0224] 在适当的条件下,与退火的引物/靶核酸接触的聚合酶可以将一个或多个核苷酸加入或并入到引物上,核苷酸可以以5'至3'、模板依赖的方式加入到引物上。将核苷酸并入到引物上(例如,通过聚合酶的作用)的这一过程一般可称为引物延伸反应。每个核苷酸都可以与可检测的标记物相缔合,可以在核酸延伸反应期间检测并鉴定(例如,基于其发光寿命和/或其它特征)该标记物,用于确定并入到延伸引物中的每个核苷酸,从而确定新合成的核酸分子的序列。通过新合成的核酸分子的序列互补性,也可以确定目标核酸分子的序列。在某些情况下,测序引物与靶核酸分子的退火和核苷酸与测序引物的结合可以在相似的反应条件下(例如,相同或相似的反应温度)或在不同的反应条件下(例如,不同的反应温度)发生。在一些实施方案中,通过合成方法进行测序可以包括:存在靶核酸分子的群体(例如,靶核酸的拷贝)和/或扩增靶核酸以取得目标核酸群体的步骤。然而,在一些实施方案中,通过合成测序来确定正在评估的每个反应中的单分子的序列(并且不需要核酸扩增来制备用于测序的靶模板)。在一些实施方案中,根据本申请的方面,并行地(例如,在单个芯片上)进行多个单分子测序反应。例如,在一些实施方案中,分别在单个芯片上的独立反应室(例如,纳米孔、样品孔)中进行多个单分子测序反应。

[0225] 实施方案能够以高准确度和长读数长度对单核酸分子进行测序,例如准确度为至少约50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、99.9%、99.99%、99.999%或99.9999%,和/或读取长度大于或等于约10

[0226] BY20IM3317FGPC-CN.电话0A个碱基对(bp)、50bp、100bp、200bp、300bp、400bp、500bp、1000bp、10,000bp、20,000bp、30,000bp、40,000bp、50,000bp或100,000bp。在一些实施方案中,单分子测序中使用的靶核酸分子是单链靶核酸(例如,脱氧核糖核酸(DNA)、DNA衍生物、核糖核酸(RNA)、RNA衍生物)模板,将该模板添加或固定到含有至少一种另外的测序反应组分(例如,聚合酶如DNA聚合酶、测序引物)的样品孔(例如纳米孔)中,该另外的组分被固定或连接到固体支撑物如样品孔的底部或侧壁。可以将靶核酸分子或聚合酶直接或通过连接基连接到样品壁上,例如在样品孔的底部或侧壁。样品孔(例如,纳米孔)还可以包含通过引物延伸反应合成核酸所需的任何其它试剂,例如,合适的缓冲液、辅助因子、酶(例如,聚合酶)和多磷酸脱氧核苷如脱氧核苷三磷酸,包括脱氧腺苷三磷酸(dATP)、脱氧胞苷三磷酸(dCTP)、脱氧鸟苷三磷酸(dGTP)、脱氧尿苷三磷酸(dUTP)和脱氧胸苷三磷酸(dTTP) dNTPs,其包括发光标记物,如本文提供的带标记的生物分子的发光标记物。

[0227] 在一些实施方案中,每一类dNTPs(例如,含腺嘌呤的dNTPs(例如,dATP)、含胞嘧啶的dNTPs(例如,dCTP)、含鸟嘌呤的dNTPs(例如,dGTP)、含尿嘧啶的dNTPs(例如,dUTPs)和含胸腺素的dNTPs(例如,dTTP))都与发光分子缀合,该发光分子包含不同的发光特性,从而检测从发光分子发射的光可以指示被整合到新合成的核酸中的dNTP的身份。可以通过任何合适的装置和/或方法检测来自发光分子的发射光(例如,来自包含至少一个发光标记物的带标记的生物分子的发射光),并将其归属于其合适的发光分子(因此,归属于相关的dNTP)。发光分子可以在任何位置与dNTP缀合,使得发光分子(例如,本申请的带标记的生物分子)的存在不会抑制dNTP并入到新合成的核酸链中或者不会抑制聚合酶的活性。在一些实施方案中,发光分子与dNTP的末端磷酸(例如, γ 磷酸)缀合。

[0228] 在一些实施方案中,单链靶核酸模板可以与测序引物、dNTPs、聚合酶和核酸合成所需的其他试剂接触。在一些实施方案中,所有合适的dNTPs可以同时与单链靶核酸模板接触(例如,所有dNTPs同时存在),这样可以连续地进行dNTPs的并入。在其它实施方案中,dNTPs可以依次与单链靶核酸模板接触,其中单链靶核酸模板分别与每个合适的dNTP接触,在单链靶核酸模板与不同dNTPs接触之间有洗涤步骤。对于待鉴定的单链靶核酸模板的每一个连续的碱基位置,可以重复这样的循环,即,使单链靶核酸模板与每种dNTP分别接触,然后进行洗涤。

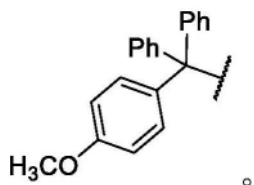
[0229] 在一些实施方案中,测序引物与单链靶核酸模板退火,聚合酶根据单链靶核酸模板连续将dNTPs(或其他多磷酸核苷)并入到引物中。与每个并入的dNTP相缔合的独特的发光分子,例如本文所述的带标记的生物分子,可以在dNTP并入到引物中期间或之后用适当的激发光激发,并且可以随后使用任意合适的装置和/或方法检测其发射。检测到特定的光发射(例如,具有特定的发射寿命、强度、光谱和/或其组合)可归因于并入的特定dNTP。然后,从检测到的发光分子的集合获得序列,可以使用该序列通过序列互补性确定单链靶核酸模板的序列。

[0230] 在一些实施方案中,本公开提供了可以在共同待决的美国专利申请号:14/543,865、14/543,867、14/543,888、14/821,656、14/821,686、14/821,688、15/161,067、15/161,

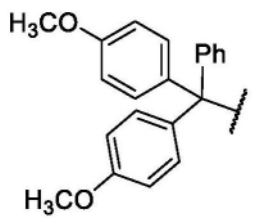
供的定义和实施方案,均适用于本文提供的化合物。

[0244] 如本文所定义, P^1 是氧保护基团。本文提供了氧保护基团的几个实例。在某些实施方案中, P^1 是任选取代的三苯基保护基(例如,三苯甲基)。在某些实施方案中, P^1 是三苯甲

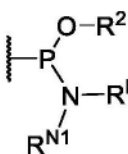
基, 具有式: 。在某些实施方案中, P^1 是 4-单甲氧基三苯甲基 (MMT), 具有式:

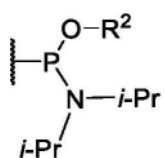


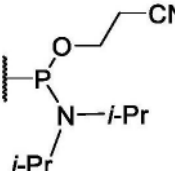
在某些实施方案中, P^1 是 4,4-二甲氧基三苯甲基 (DMT), 具有式:



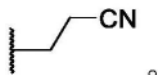
[0245] 在某些实施方案中, R^1 是反应性部分。例如, R^1 是可用于多核苷酸合成、多肽合成、多糖合成等的反应性手柄。反应性部分可以是能够与第二反应性部分(例如, $-OH$ 或 $-NH_2$ 基团)反应以形成共价键的任何基团。本领域技术人员将会知道哪些反应性部分可用于在多核苷酸合成、多肽合成、多糖合成等中形成键。在某些实施方案中, R^1 是可用于多核苷酸合成的部分(例如, 亚磷酰胺)。在某些实施方案中, R^1 是亚磷酰胺。在某些实施方案中, R^1 是具

有式:  的亚磷酰胺, 其中 R^{N1} 和 R^2 如本文所定义。在某些实施方案中, R^1 是具有式:



的亚磷酰胺。在某些实施方案中, R^1 是具有式:  的亚磷酰胺。

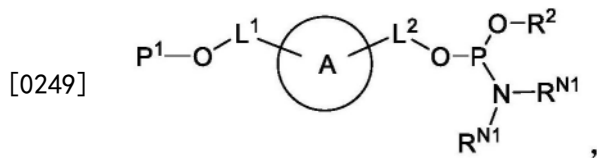
[0246] 如本文所定义的, R^2 是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氧保护基团。在某些实施方案中, R^2 是任选取代的烷基。在某些实施方案中, R^2 是任选取代的 C_{1-6} 烷基。在某些实施方案中, R^2 是任选取代的 C_{1-3} 烷基。在某些实施方案中, R^2 具有式:



[0247] 如本文所定义的, R^{N1} 的每种情况都独立地是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氮保护基团; 任选地, 其中, 与相同氮键合的两个 R^{N1} 与中间的原子连接在一起, 形成任选取代的杂环基或任选取代的杂芳基。在某些实施方案中, R^{N1} 是任选取代的烷基。在某些实施方案中, R^{N1} 是任选取代的 C_{1-6} 烷基。在某些实施方案中, R^{N1} 是未取代的 C_{1-6} 烷基。在某些实施方案中, R^{N1} 是任选取代的 C_{1-3} 烷基。在某些实施方案中, R^{N1} 是未取

代的C₁₋₃烷基。在某些实施方案中,R^{N1}选自由以下组成的组:甲基、乙基、正丙基、异丙基、正丁基、异丁基、仲丁基和叔丁基。在某些实施方案中,R^{N1}是异丙基。在某些实施方案中,R^{N1}的两种情况都是异丙基。

[0248] 在某些实施方案中,R¹是在核苷偶联反应中具有反应性的部分,例如亚磷酰胺。在某些实施方案中,式(II)的化合物具有下式:

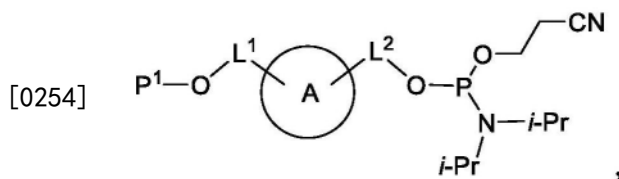


[0250] 或其盐,其中:

[0251] R²是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氧保护基团;

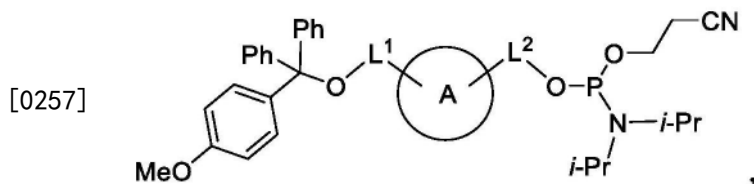
[0252] R^{N1}的每种情况都独立地是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氮保护基团;任选地,其中,与相同氮键合的两个R^{N1}与中间的原子连接在一起,形成任选取代的杂环基或任选取代的杂芳基。

[0253] 在某些实施方案中,式(II)的化合物具有下式:



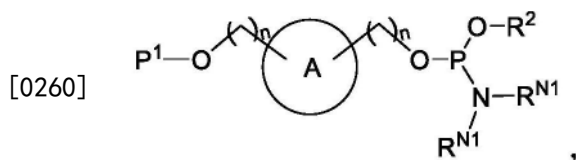
[0255] 或其盐。

[0256] 在某些实施方案中,式(II)的化合物具有下式:



[0258] 或其盐。

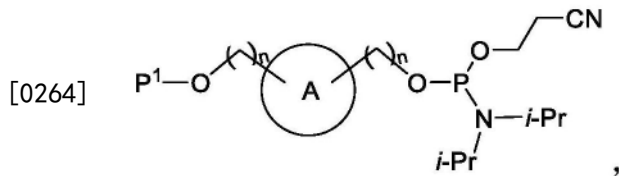
[0259] 在某些实施方案中,式(II)的化合物具有下式:



[0261] 或其盐,其中:

[0262] n的每种情况都是1至20的整数,包括1和20。

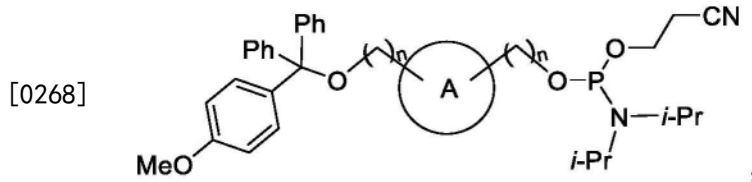
[0263] 在某些实施方案中,式(II)的化合物具有下式:



[0265] 或其盐,其中:

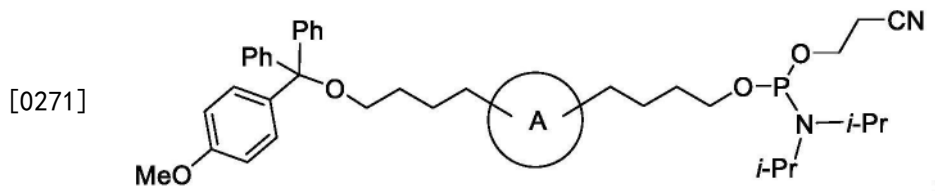
[0266] n的每种情况都是1至20的整数,包括1和20。

[0267] 在某些实施方案中,式(II)的化合物具有下式:



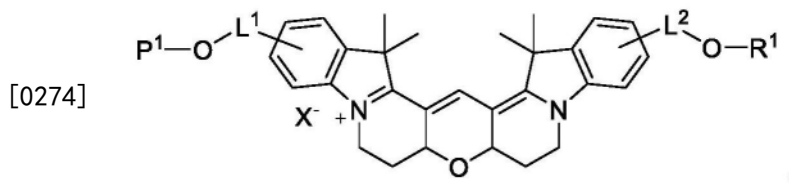
[0269] 或其盐。

[0270] 在某些实施方案中,式(II)的化合物具有下式:



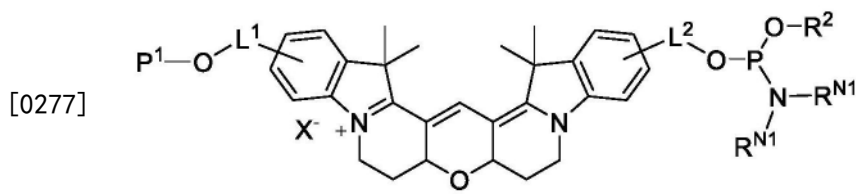
[0272] 或其盐。

[0273] 如本文所述,环A可以是多环花菁,例如Cy3B。在某些实施方案中,式(II)的化合物具有下式:



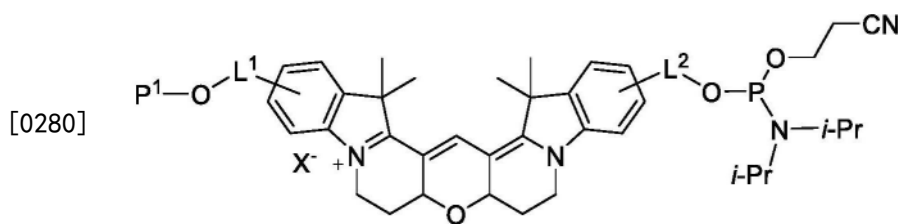
[0275] 或其盐,其中X⁻是抗衡离子或不存在。

[0276] 在某些实施方案中,式(II)的化合物具有下式:



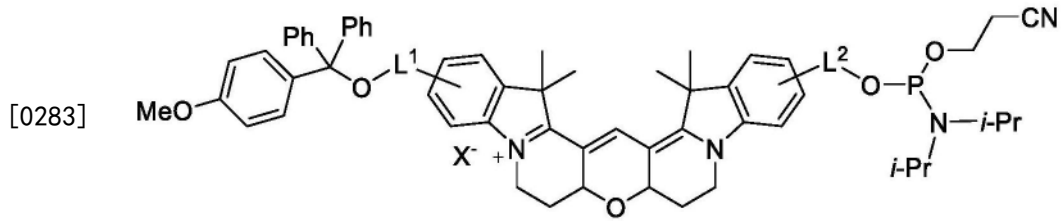
[0278] 或其盐。

[0279] 在某些实施方案中,式(II)的化合物具有下式:



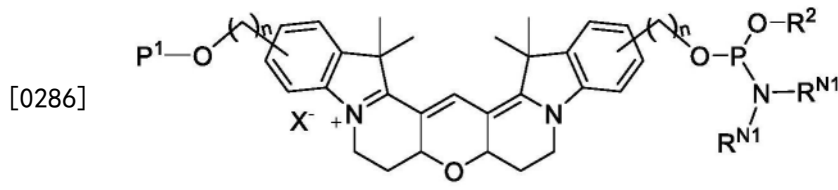
[0281] 或其盐。

[0282] 在某些实施方案中,式(II)的化合物具有下式:



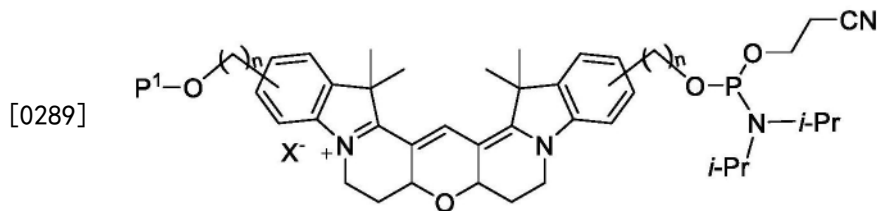
[0284] 或其盐。

[0285] 在某些实施方案中,式(II)的化合物具有下式:



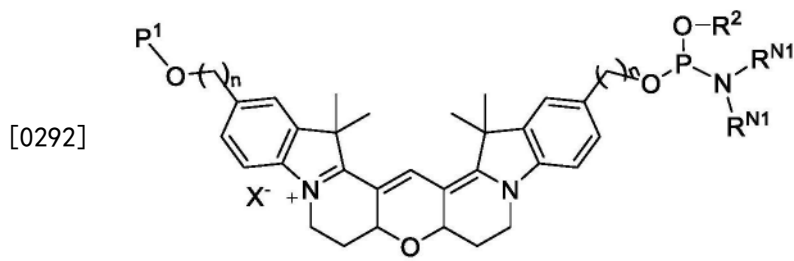
[0287] 或其盐。

[0288] 在某些实施方案中,式(II)的化合物具有下式:



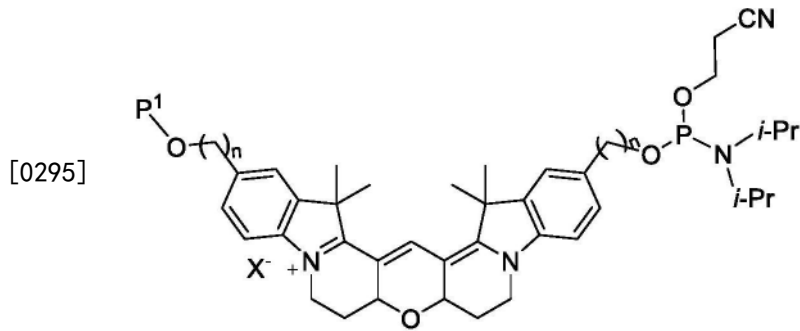
[0290] 或其盐。

[0291] 在某些实施方案中,式(II)的化合物具有下式:



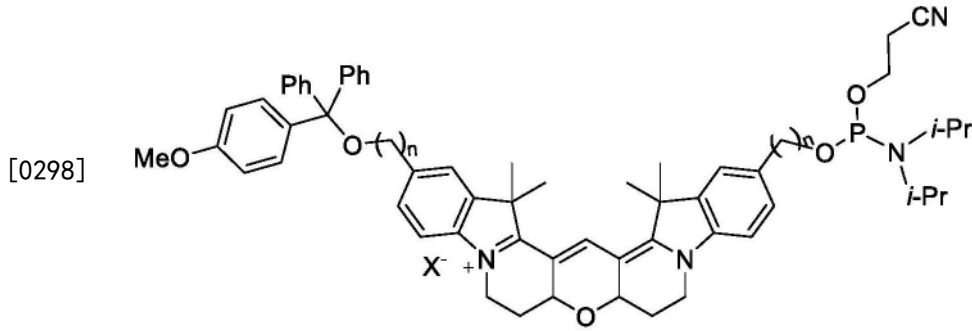
[0293] 或其盐。

[0294] 在某些实施方案中,式(II)的化合物具有下式:



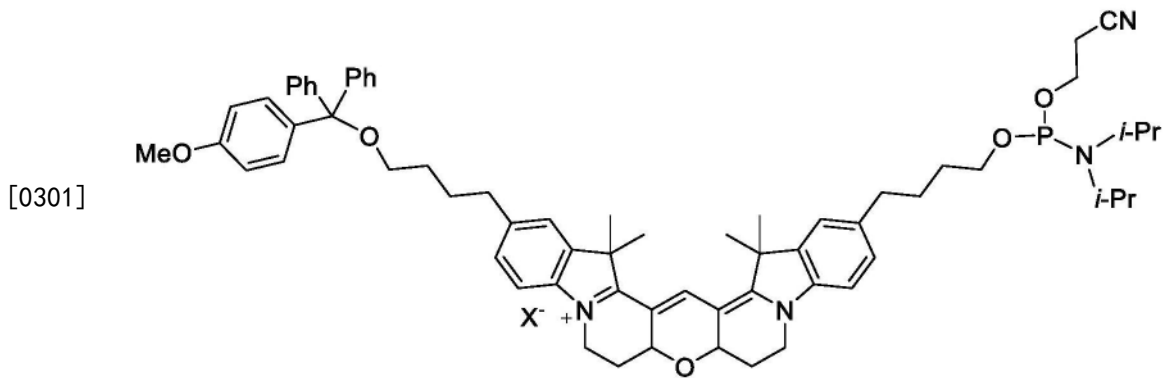
[0296] 或其盐。

[0297] 在某些实施方案中,式(II)的化合物具有下式:



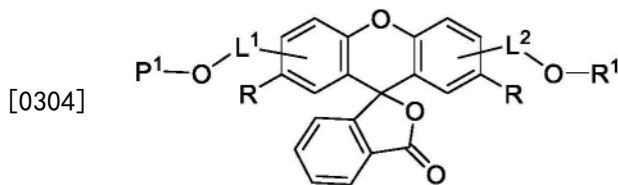
[0299] 或其盐。

[0300] 在某些实施方案中,式(II)的化合物具有下式:



[0302] 或其盐。

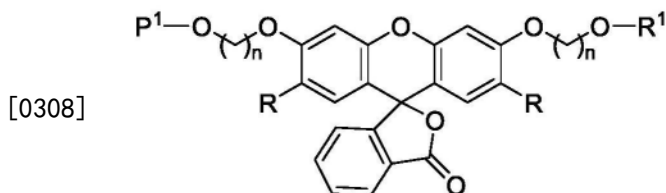
[0303] 在某些实施方案中,A可以是荧光酮荧光团,例如荧光素或若丹明。在某些实施方案中,式(II)的化合物具有下式:



[0305] 或其盐,其中:

[0306] 每个R独立地是氢、卤素、-N₃、-CN、-NO₂、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、-OR⁰、-SR^S或N(R^N)₂。

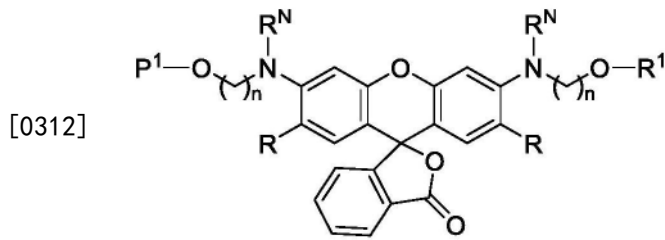
[0307] 在某些实施方案中,式(II)的化合物具有下式:



[0309] 或其盐,其中:

[0310] n独立地是1至20的整数,包括1和20。

[0311] 在某些实施方案中,式(II)的化合物具有下式:

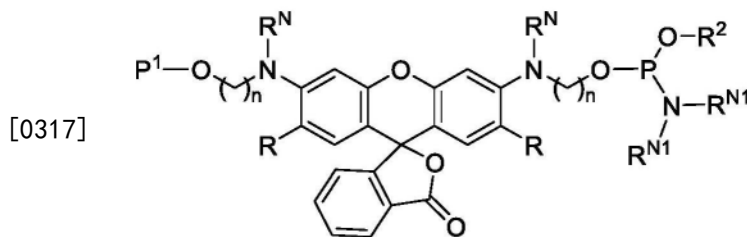


[0313] 或其盐,其中:

[0314] n独立地是1至20的整数,包括1和20;并且

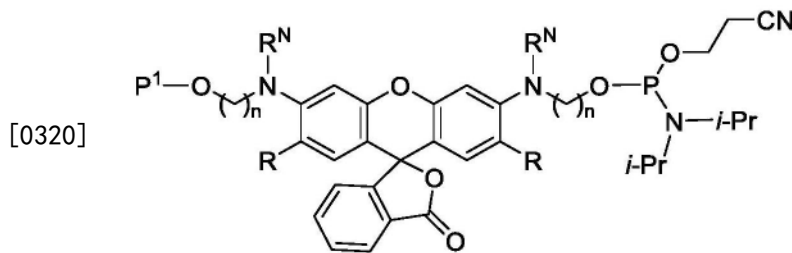
[0315] R^N的每种情况都独立地是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氮保护基团。

[0316] 在某些实施方案中,式(II)的化合物具有下式:



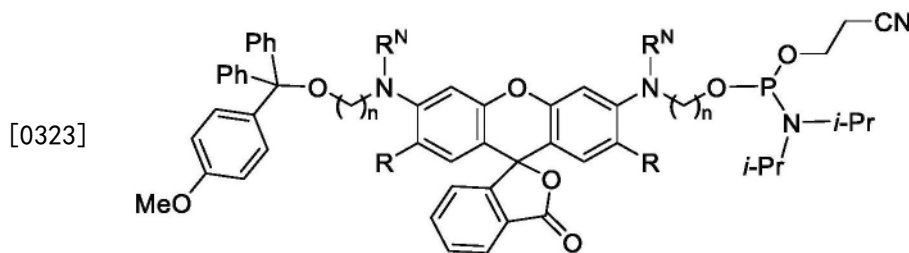
[0318] 或其盐。

[0319] 在某些实施方案中,式(II)的化合物具有下式:



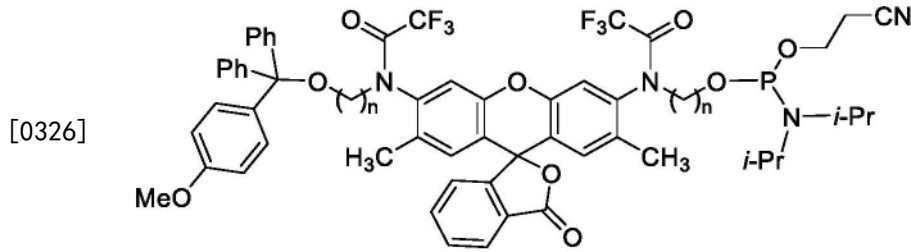
[0321] 或其盐。

[0322] 在某些实施方案中,式(II)的化合物具有下式:



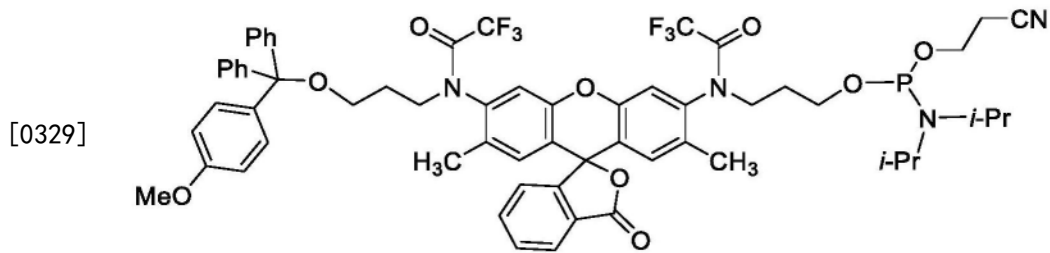
[0324] 或其盐。

[0325] 在某些实施方案中,式(II)的化合物具有下式:



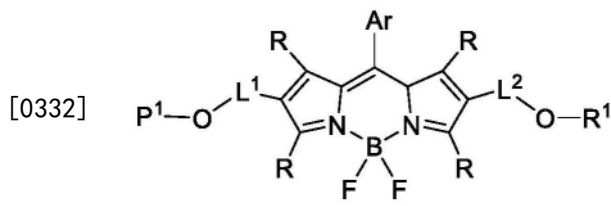
[0327] 或其盐。

[0328] 在某些实施方案中,式(II)的化合物具有下式:



[0330] 或其盐。

[0331] 如本文所述,在某些实施方案中,A是BODIPY荧光团。因此,在某些实施方案中,式(II)的化合物具有下式:

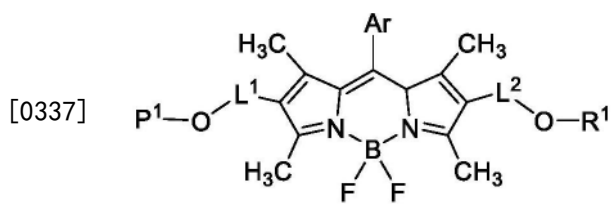


[0333] 或其盐,其中:

[0334] 每个R独立地是氢、卤素、-N₃、-CN、-NO₂、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、-OR⁰、-SR^S或N(R^N)₂;并且

[0335] Ar是任选取代的芳基或任选取代的杂芳基。

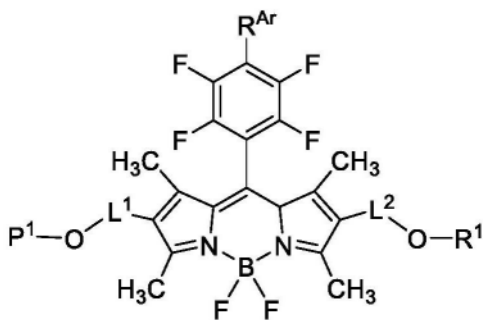
[0336] 在某些实施方案中,式(II)的化合物具有下式:



[0338] 或其盐。

[0339] 在某些实施方案中,式(II)的化合物具有下式:

[0340]



[0341] 或其盐,其中:

[0342] R^{Ar} 是氢、卤素、 $-N_3$ 、 $-CN$ 、 $-NO_2$ 、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的芳基、任选取代的杂环基、任选取代的杂芳基、 $-OR^0$ 、 $-SR^S$ 或 $N(R^N)_2$ 。

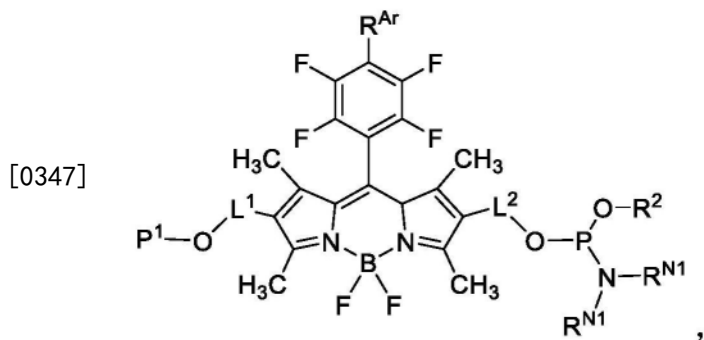
[0343] 如本文一般定义, R^0 是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氧保护基团。在某些实施方案中, R^0 是氢。在某些实施方案中, R^0 是任选取代的烷基。在某些实施方案中, R^0 是任选取代的 C_{1-6} 烷基。在某些实施方案中, R^0 是未取代的 C_{1-6} 烷基。在某些实施方案中, R^0 选自: 甲基、乙基、正丙基、异丙基、正丁基、异丁基、仲丁基和叔丁基。在某些实施方案中, R^0 是任选取代的烯基。在某些实施方案中, R^0 是任选取代的炔基。在某些实施方案中, R^0 是任选取代的碳环基。在某些实施方案中, R^0 是任选取代的杂环基。在某些实施方案中, R^0 是任选取代的芳基。在某些实施方案中, R^0 是任选取代的杂芳基。在某些实施方案中, R^0 是任选取代的酰基。在某些实施方案中, R^0 是氧保护基团。在某些实施方案中, R^0 是 $-C(=O)Ph$ 。

[0344] 如本文一般定义, R^N 的每个情况都独立地是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或氮保护基团; 或者任选地, 两个 R^N 与中间的原子连接在一起, 形成任选取代的杂环基或任选取代的杂芳基。在某些实施方案中, R^N 是氢。在某些实施方案中, R^N 是任选取代的烷基。在某些实施方案中, R^N 是任选取代的 C_{1-6} 烷基。在某些实施方案中, R^N 是未取代的 C_{1-6} 烷基。在某些实施方案中, R^N 选自: 甲基、乙基、正丙基、异丙基、正丁基、异丁基、仲丁基和叔丁基。在某些实施方案中, R^N 是任选取代的烯基。在某些实施方案中, R^N 是任选取代的炔基。在某些实施方案中, R^N 是任选取代的碳环基。在某些实施方案中, R^N 是任选取代的杂环基。在某些实施方案中, R^N 是任选取代的芳基。在某些实施方案中, R^N 是任选取代的杂芳基。在某些实施方案中, R^N 是任选取代的酰基。在某些实施方案中, R^N 是氮保护基团。在某些实施方案中, 两个 R^N 与中间的原子连接在一起, 形成任选取代的杂环基或任选取代的杂芳基。

[0345] 如本文一般定义, R^S 是氢、任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、任选取代的酰基或硫保护基团。在某些实施方案中, R^S 是氢。在某些实施方案中, R^S 是任选取代的烷基。在某些实施方案中, R^S 是任选取代的 C_{1-6} 烷基。在某些实施方案中, R^S 是未取代的 C_{1-6} 烷基。在某些实施方案中, R^S 选自: 甲基、乙基、正丙基、异丙基、正丁基、异丁基、仲丁基和叔丁基。在某些实施方案中, R^S 是任选取代的烯基。在某些实施方案中, R^S 是任选取代的炔基。在某些实施

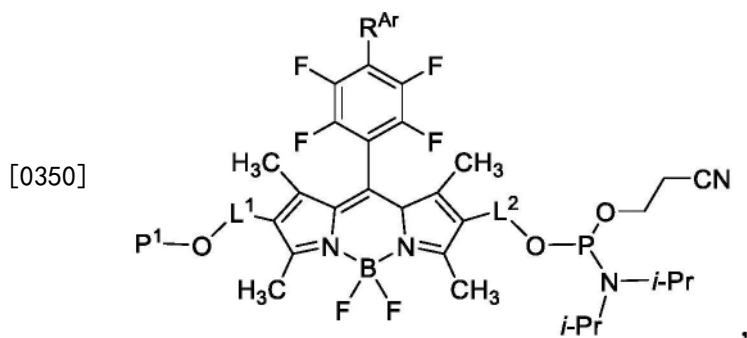
方案中, R^S 是任选取代的碳环基。在某些实施方案中, R^S 是任选取代的杂环基。在某些实施方案中, R^S 是任选取代的芳基。在某些实施方案中, R^S 是任选取代的杂芳基。在某些实施方案中, R^S 是任选取代的酰基。在某些实施方案中, R^S 是硫保护基团。在某些实施方案中, R^S 是 PEG (聚乙二醇)。在某些实施方案中, R^S 是 $-(CH_2CH_2O)_mOCH_3$ 。在某些实施方案中, $-SR^S$ 是 $-S(CH_2CH_2O)_mOCH_3$ 。

[0346] 在某些实施方案中, 式 (II) 的化合物具有下式:



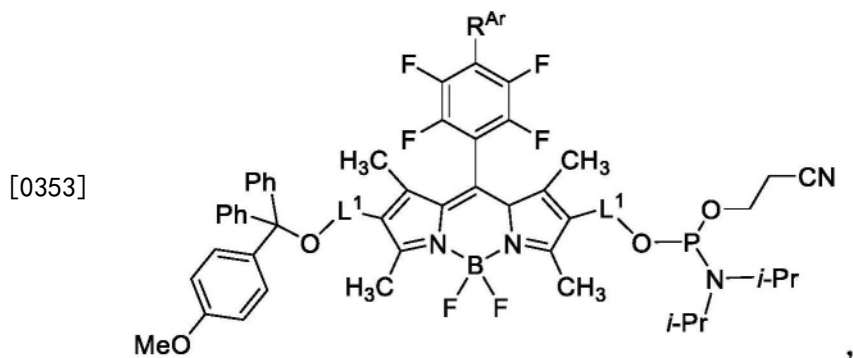
[0348] 或其盐。

[0349] 在某些实施方案中, 式 (II) 的化合物具有下式:



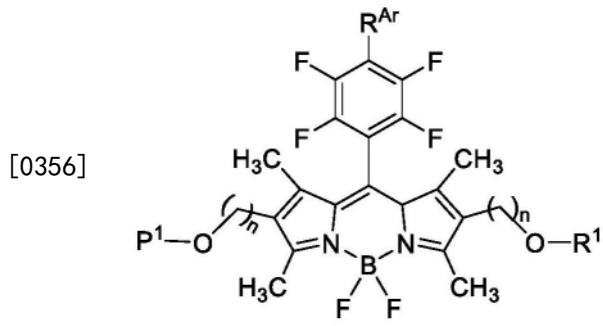
[0351] 或其盐。

[0352] 在某些实施方案中, 式 (II) 的化合物具有下式:



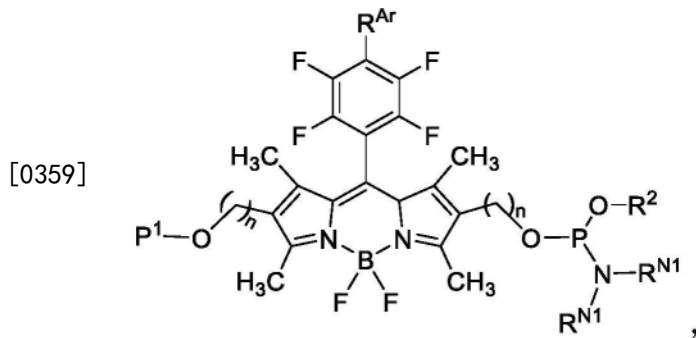
[0354] 或其盐。

[0355] 在某些实施方案中, 式 (II) 的化合物具有下式:



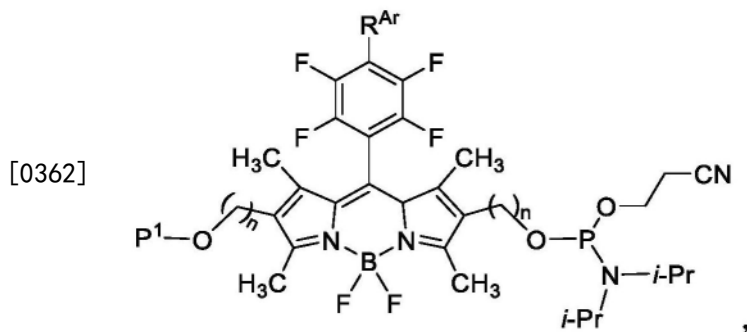
[0357] 或其盐,其中,n独立地是1至20的整数,包括1和20。

[0358] 在某些实施方案中,式(II)的化合物具有下式:



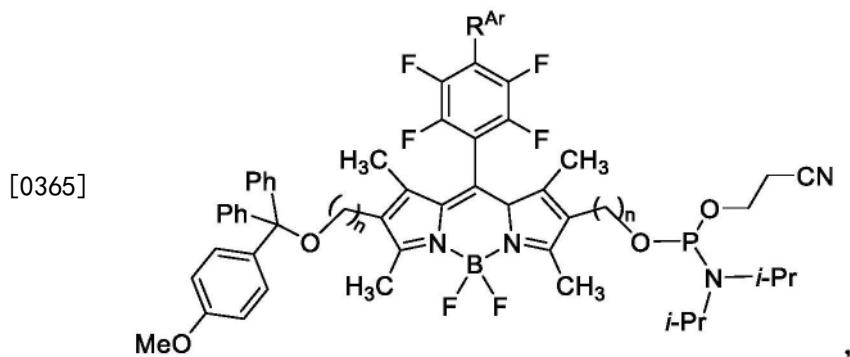
[0360] 或其盐。

[0361] 在某些实施方案中,式(II)的化合物具有下式:



[0363] 或其盐。

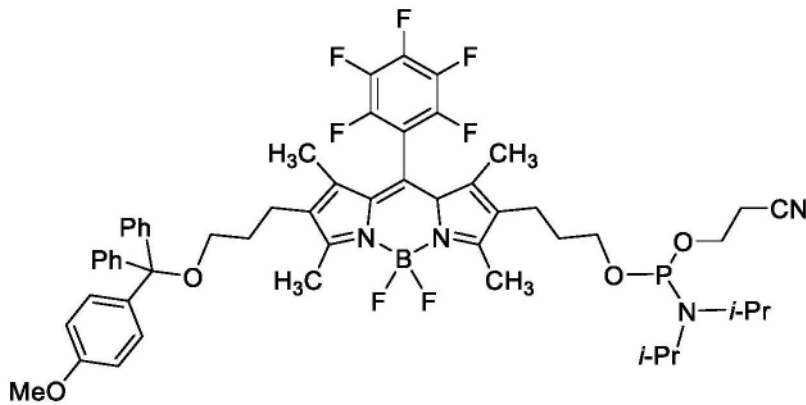
[0364] 在某些实施方案中,式(II)的化合物具有下式:



[0366] 或其盐。

[0367] 在某些实施方案中,式(II)的化合物具有下式:

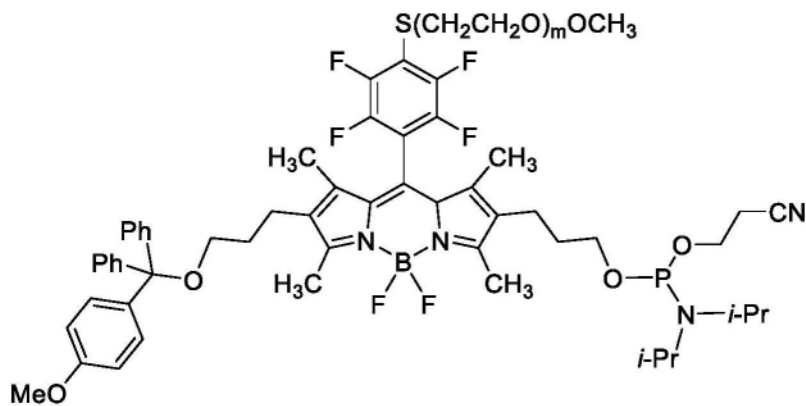
[0368]



[0369] 或其盐。

[0370] 在某些实施方案中,式(II)的化合物具有下式:

[0371]



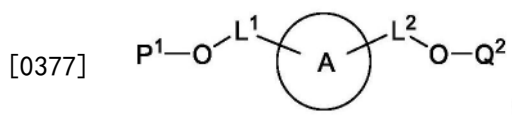
[0372] 或其盐。

[0373] 制备内部标记的生物分子的方法

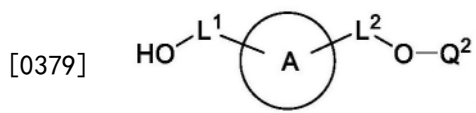
[0374] 本文还提供了制备本文所述的内部标记的生物分子(例如,式(I)的生物分子)的方法。通常,该方法包括两个随后的缀合步骤,涉及本文提供的双官能化合物(例如,式(II)的化合物)。

[0375] 因此,本文提供了制备带标记的生物分子的方法,所述方法包括:

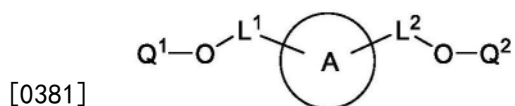
[0376] (i) 在足以促进缀合的条件下,使式 Q^2-OH 的单体或低聚生物分子或其盐与式(II)化合物或其盐接触,得到下式缀合物或其盐:



[0378] (ii) 在足以裂解 P^1 保护基团并产生下式缀合物或其盐的条件下,对步骤(i)中形成的缀合物进行脱保护:



[0380] (iii) 在足以促进缀合的条件下,使步骤(ii)中形成的缀合物与式 Q^1-O-R^1 的单体或低聚生物分子或其盐接触,得到式(I)的带标记的生物分子:



(I)。

[0382] 例如,在某些实施方案中,当制备内部标记的寡核苷酸时,可以使用标准亚磷酰胺化学方法将式(II)的化合物与 Q^2-OH 偶联。此外,在某些实施方案中,当制备内部标记的寡核苷酸时,将在步骤(ii)中形成的带有羟基部分的未保护的缀合物与烷基-(2-氰基乙基)-N,N-二异丙基)-亚磷酰胺偶联,例如在标准寡核苷酸合成中所使用的。该反应可以在活化剂如1H-四唑、乙硫基四唑、苄基硫基四唑、二氰基咪唑或其他合适的弱酸的存在下进行。然后,在某些实施方案中,步骤(iii)可以使用标准亚磷酰胺化学方法进行。

[0383] 制备双缀合标记物的方法

[0384] 在另一方面,本发明提供了用于制备本文所述的双缀合标记物(例如,式(II)的化合物)的合成方法。

[0385] 一般反应参数

[0386] 以下实施方案适用于本文所述的所有合成方法。

[0387] 本文提供和描述的反应可能涉及一种或多种试剂。在某些实施方案中,试剂可以以催化量存在。在某些实施方案中,催化量为0-1mol%、0-5mol%、0-10mol%、1-5mol%、1-10mol%、5-10mol%、10-20mol%、20-30mol%、30-40mol%、40-50mol%、50-60mol%、60-70mol%、70-80mol%、80-90mol%或90-99mol%。在某些实施方案中,试剂可以以化学计量(例如约1当量)存在。在某些实施方案中,试剂可以过量(例如,大于1当量)存在。在某些实施方案中,过量是约1.1、1.2、1.3、1.4、1.5、2.0、2.5、3.0、3.5、4.0、4.5、5.0、5.5、6.0、6.5、7.0、7.5、8.0、8.5、9.0、9.5、10、15或20当量。在某些实施方案中,过量是约1.1-2、2-3、3-4、4-5、1.1-5、5-10、10-15、15-20或10-20当量。在某些实施方案中,过量是大于20当量。

[0388] 本文所述的反应可以在任何温度下进行。在某些实施方案中,反应在室温(rt)(21℃或70°F)或其大约值进行。在某些实施方案中,反应是在低于室温(例如-100℃至21℃)下进行的。在某些实施方案中,反应在-78℃或大约-78℃进行。在某些实施方案中,反应在-10℃或大约-10℃进行。在某些实施方案中,反应在0℃或大约0℃进行。在某些实施方案中,反应在高于室温的温度进行。在某些实施方案中,反应在30、40、50、60、70、80、110、120、130、140或150℃进行。在某些实施方案中,反应在高于150℃的温度进行。

[0389] 本文所述的反应可以在溶剂或溶剂混合物(例如助溶剂)中进行。溶剂可以是极性或非极性、质子或非质子的。在本文所述的反应中可以使用任何溶剂,并且反应不限于特定的溶剂或溶剂的组合。可用于本文所述方法的常见有机溶剂包括但不限于丙酮、乙腈、苯、苄腈、1-丁醇、2-丁酮、乙酸丁酯、叔丁基甲基醚、二硫化碳四氯化碳、氯苯、1-氯丁烷、氯仿、环己烷、环戊烷、1,2-二氯苯、1,2-二氯乙烷、二氯甲烷(DCM)、N,N-二甲基乙酰胺N,N-二甲基甲酰胺(DMF)、1,3-二甲基-3,4,5,6-四氢-2-嘧啶酮(DMPU)、1,4-二噁烷、1,3-二噁烷、乙醚、2-乙氧基乙醚、乙酸乙酯、乙醇、乙二醇、二甲醚、庚烷、正己烷、己烷、六甲基磷酰胺(HMPA)、2-甲氧基乙醇、2-甲氧基乙酸乙酯、甲醇、2-甲基丁烷、4-甲基-2-戊酮、2-甲基-1-丙醇、2-甲基-2-丙醇、1-甲基-2-吡咯烷酮、二甲基亚砜(DMSO)、硝基甲烷、1-辛醇、戊烷、3-戊酮、1-丙醇、2-丙醇、吡啶、四氯乙烯、四氢呋喃(THF)、2-甲基四氢呋喃、甲苯、三氯苯、1,

1,2-三氯三氟乙烷、2,2,4-三甲基戊烷、三甲胺、三乙胺、N,N-二异丙基乙胺、二异丙胺、水、邻二甲苯、对二甲苯。

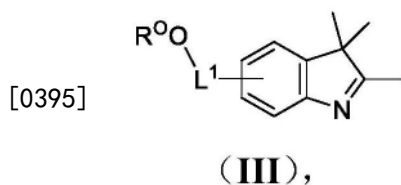
[0390] 本文所述的反应可以在任何时间量内进行。在某些实施方案中,反应可以进行几秒钟、几分钟、几小时或几天。

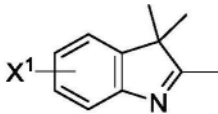
[0391] 本文描述的方法可以用于以任何化学收率制备化合物。在某些实施方案中,以1-10%、10-20%、20-30%、30-40%、40-50%、50-60%、60-70%、70-80%、80-90%或90-100%的收率生产化合物。在某些实施方案中,收率是一个合成步骤后的百分比收率。在某些实施方案中,收率是多于一个合成步骤(例如2、3、4或5个合成步骤)后的百分比收率。

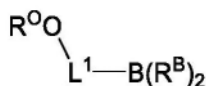
[0392] 本文所述的方法可以进一步包括一个或多个纯化步骤。例如,在某些实施方案中,通过本文描述的方法产生的化合物可以通过色谱、萃取、过滤、沉淀、结晶或本领域已知的任何其他方法纯化。在某些实施方案中,化合物或混合物无需纯化即可直接用于下一个合成步骤(例如粗品)。

[0393] 基于Cy3B的标记物的制备方法

[0394] 本文提供了一种制备式(III)化合物或其盐的方法:



[0396] 所述方法包括在钯的存在下,使式  化合物或其盐,与式

 化合物或其盐偶联,以得到式(III)化合物或其盐,其中:

[0397] X¹是卤素或离去基团;

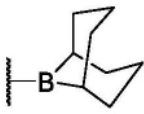
[0398] B(R^B)₂是硼烷、硼酸或硼酸酯;

[0399] L¹是连接基,选自由以下组成的组:任选取代的亚烷基、任选取代的亚烯基、任选取代的亚炔基、任选取代的亚杂烷基、任选取代的亚杂烯基、任选取代的亚杂炔基、任选取代的亚碳环基、任选取代的亚杂环基、任选取代的亚芳基、任选取代的亚杂芳基及其组合;并且

[0400] R⁰是氧保护基团。

[0401] 在某些实施方案中,生产式(III)化合物的反应在除钯以外的金属的存在下进行。例如,该反应可以被钯催化,或被不同的金属催化。在某些实施方案中,该金属是过渡金属。

[0402] 如本文所定义,基团-B(R^B)₂是硼烷、硼酸或硼酸酯。在某些实施方案中,-B(R^B)₂是硼烷。在某些实施方案中,-B(R^B)₂是硼酸。在某些实施方案中,-B(R^B)₂是硼酸酯。在某些实

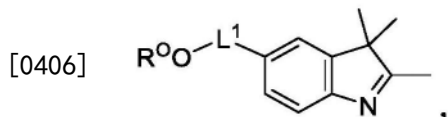
施方案中,-B(R^B)₂是式:  的硼烷。如本文所定义,R^B的每个实例独立地是任选取代的烷基、任选取代的碳环基、任选取代的杂环基、任选取代的芳基、任选取代的杂芳基、-OH或-OR⁰。任选地,两个R^B与中间的原子连接在一起,形成任选取代的碳环基或任选取代的杂

环基。通常, $-B(R^B)_2$ 是可用于金属促进的或金属催化的交叉偶联反应的任何合适的硼烷、硼酸或硼酸酯。

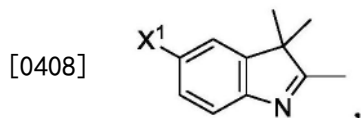
[0403] 连接基 L^1 如本文所定义。

[0404] 如本文所定义, X^1 是卤素或离去基团。在某些实施方案中, X^1 是卤素。在某些实施方案中, X^1 是离去基团。在某些实施方案中, X^1 是 $-Cl$ 、 $-Br$ 或 $-I$ 。在某些实施方案中, X^1 是 $-I$ 。

[0405] 在偶联反应的某些实施方案中, 式 (III) 的化合物具有下式或其盐:

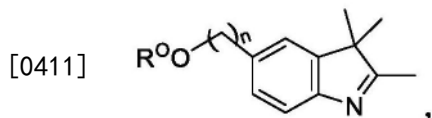


[0407] 因此起始原料的化学式为:

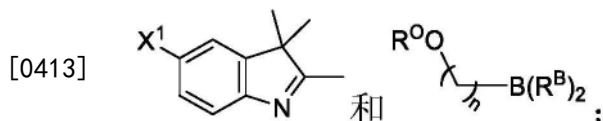


[0409] 或其盐。

[0410] 在某些实施方案中, 式 (III) 的化合物具有下式或其盐:



[0412] 其中 n 是 1-20 的整数, 包括 1 和 20; 因此, 起始原料具有下式:



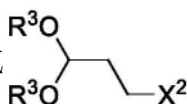
[0414] 或其盐。

[0415] 如上所述, 偶联反应在钯的存在下进行。在某些实施方案中, 钯是钯配合物。在某些实施方案中, 钯配合物是钯 (II) 配合物。在某些实施方案中, 钯络合物是 $PdCl_2(dppf)$ 。在某些实施方案中, 钯以催化量存在。在其他实施方案中, 钯以化学计量或过量存在。

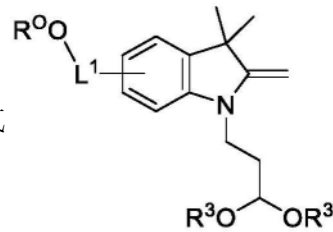
[0416] 在某些实施方案中, 偶联反应在碱的存在下进行。在某些实施方案中, 碱是碳酸盐碱。在某些实施方案中, 碱是 Cs_2CO_3 。

[0417] 在某些实施方案中, 偶联反应在溶剂中进行。在某些实施方案中, 溶剂是 THF、DMF 或其混合物。

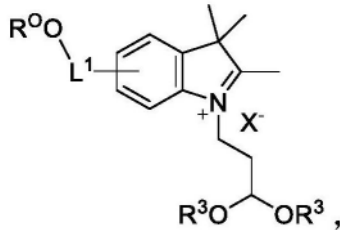
[0418] 偶联反应可以在任何温度下进行。在某些实施方案中, 反应在室温下进行。在某些实施方案中, 反应在高于室温 (即高温) 下进行。在某些实施方案中, 反应在室温至 $100^\circ C$ 之间进行。在某些实施方案中, 反应在 $50^\circ C$ 至 $100^\circ C$ 之间进行。在某些实施方案中, 反应在约 $70^\circ C$ 下进行。

[0419] 在某些实施方案中, 该方法还包括以下步骤: 用式  化合物或其盐, 将

式(III)化合物烷基化,得到式



化合物,或其互变异构体:

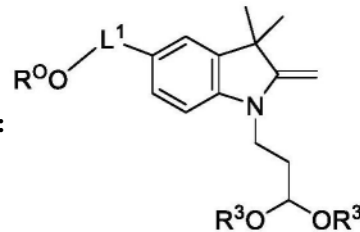


或其盐,其中:

[0420] X^2 是卤素或离去基团;并且

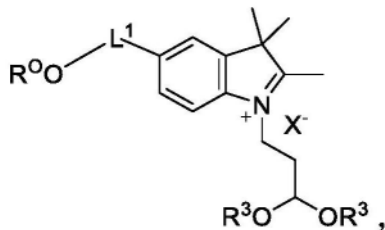
[0421] 每个 R^3 独立地是任选取代的烷基、任选取代的酰基或氧保护基团;或者任选地,两个 R^3 与中间的原子连接在一起,形成任选取代的杂环基。

[0422] 在某些实施方案中,产物是式:



的化合物,或其互变异构

体:

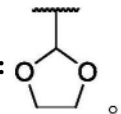


或其盐。

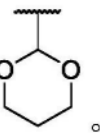
[0423] 在某些实施方案中,烷基化步骤在碱的存在下进行。在某些实施方案中,碱是卤化物盐。在某些实施方案中,碱是碘化物盐。在某些实施方案中,碱是KI。在某些实施方案中,烷基化步骤在溶剂中进行。在某些实施方案中,溶剂是乙腈(MeCN)。在某些实施方案中,反应在室温下进行。在某些实施方案中,该反应在高温下进行。在某些实施方案中,反应在约100°C下进行。

[0424] 如本文所定义, X^2 是卤素或离去基团。在某些实施方案中, X^2 是卤素。在某些实施方案中, X^2 是-Cl、-Br或-I。在某些实施方案中, X^2 是-Br。在某些实施方案中, X^2 是离去基团。

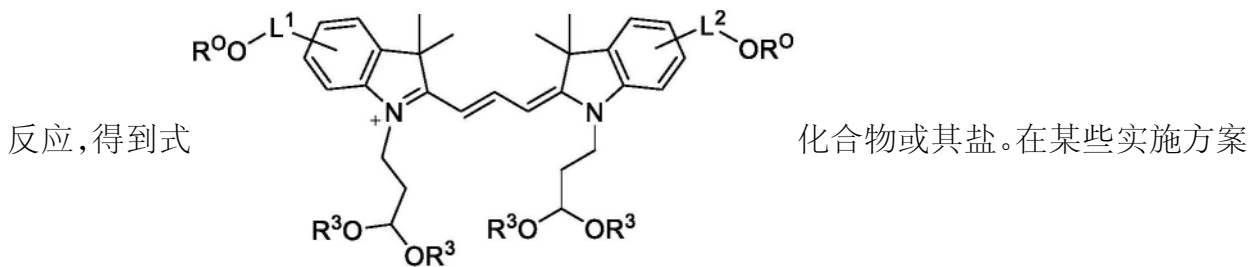
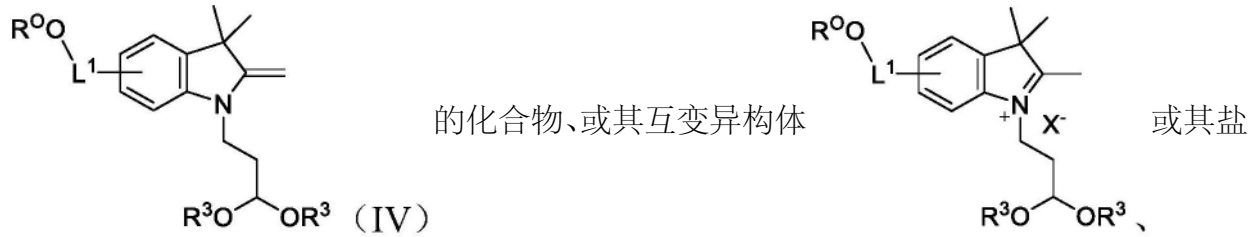
[0425] 如本文所定义的,每个 R^3 独立地是任选取代的烷基、任选取代的酰基或氧保护基团;或者任选地,两个 R^3 与中间的原子连接在一起,形成任选取代的杂环基。在某些实施方案中, R^3 是任选取代的烷基。在某些实施方案中, R^3 是任选取代的酰基。在某些实施方案中, R^3 是氧保护基团。在某些实施方案中,两个 R^3 与中间的原子连接在一起,形成任选取代的杂

环基。在某些实施方案中,两个 R^3 与中间的原子连接在一起形成: 在某些实施方案

中,两个R³与中间的原子连接在一起形成:



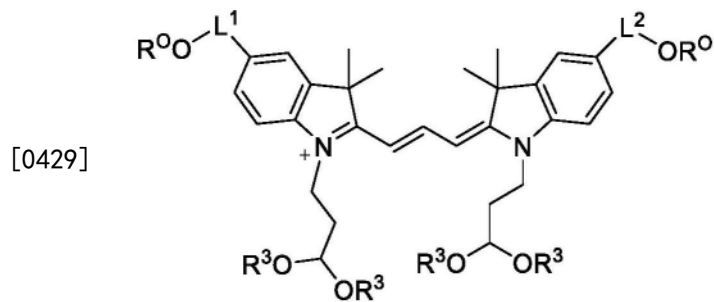
[0426] 在某些实施方案中,所述方法还包括:在甲脒的存在下,使式



中,该反应需要 (i) 使式 (IV) 的化合物或其盐在甲脒存在下反应,形成中间体;以及 (ii) 使步骤 (i) 中形成的中间体与另一种式 (IV) 化合物或其盐反应,得到上述产物。

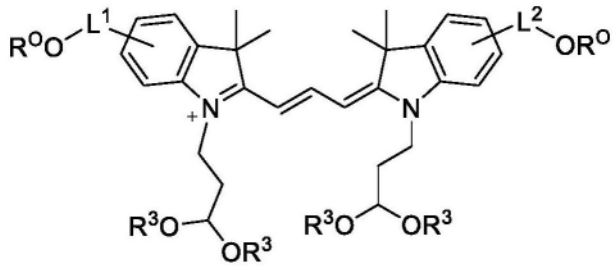
[0427] 在某些实施方案中,步骤 (i) 中的甲脒为二苯基甲脒。在某些实施方案中,步骤 (i) 中的反应在碱的存在下进行。在某些实施方案中,碱是吡啶碱。在某些实施方案中,碱是 DMAP。在某些实施方案中,步骤 (i) 中的反应在酸酐的存在下进行。在某些实施方案中,酸酐是乙酸酐 (Ac₂O)。在某些实施方案中,步骤 (i) 中的反应在高温 (例如约 125°C) 下进行。在某些实施方案中,该反应在溶剂中进行。在某些实施方案中,步骤 (ii) 中的反应在碱的存在下进行。在某些实施方案中,碱是胺碱 (例如三烷基胺碱)。在某些实施方案中,碱是 Et₃N。其中反应在碱的存在下进行。在某些实施方案中,该反应在溶剂中进行。在某些实施方案中,溶剂是 EtOH。在某些实施方案中,反应在高温 (例如约 80°C) 下进行。

[0428] 在某些实施方案中,产物具有下式:

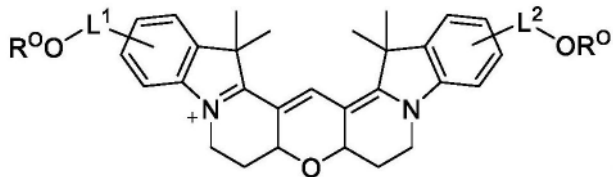


[0430] 或其盐。

[0431] 在某些实施方案中,所述方法还包括:在酸的存在下,将式



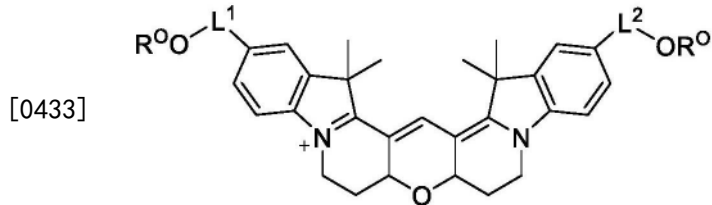
化合物或其盐环化,以得到式



化合物或其盐。在某些实施方案中,酸是磺酸。

在某些实施方案中,酸是硫酸。在某些实施方案中,该反应在溶剂中进行。在某些实施方案中,溶剂是 CH_2Cl_2 。在某些实施方案中,反应在高温(例如约 60°C)下进行。

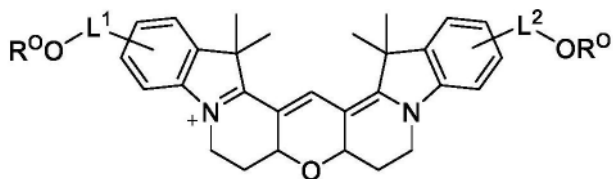
[0432] 在某些实施方案中,产物具有下式:



[0433]

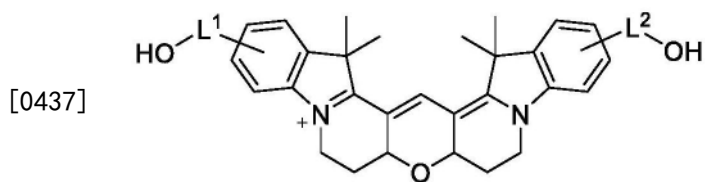
[0434] 或其盐。

[0435] 在某些实施方案中,该方法还包括以下步骤:将式



化合物或其盐脱保护,

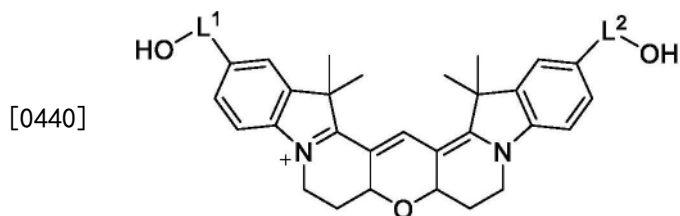
[0436] 得到下式的化合物:



[0437]

[0438] 或其盐。

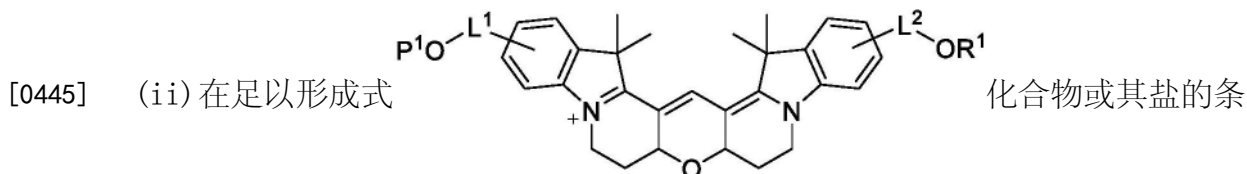
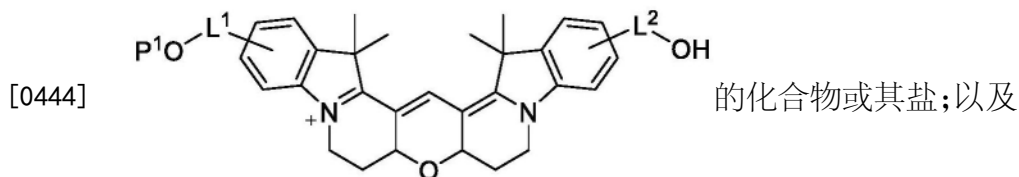
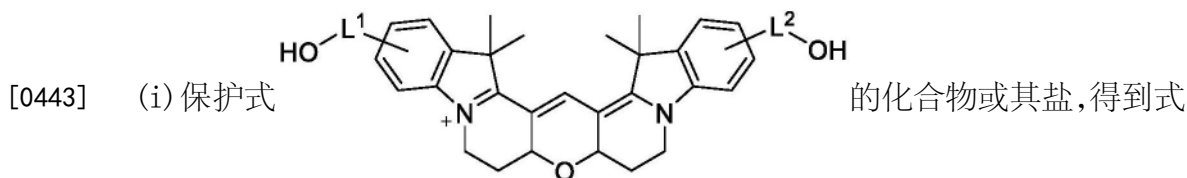
[0439] 在某些实施方案中,化合物具有下式:



[0440]

[0441] 或其盐。

[0442] 在某些实施方案中,该方法还包括以下步骤:



件下,使步骤(i)制得的化合物反应。

[0446] 定义

[0447] 化学定义

[0448] 特定官能团和化学术语的定义在下面更详细地描述。根据元素周期表(CAS版本,化学和物理手册,第75版,内页)来鉴定化学元素,并且通常如本文所述来定义特定的官能团。另外,Organic Chemistry,Thomas Sorrell,University Science Books,Sausalito,1999;Smith and March,March's Advanced Organic Chemistry,5th Edition,John Wiley&Sons,Inc.,New York,2001;Larock,Comprehensive Organic Transformations,VCH Publishers,Inc.,New York,1989;和Carruthers,Some Modern Methods of Organic Synthesis,3rd Edition,Cambridge University Press,Cambridge,1987中描述了有机化学的一般原理以及特定的功能部分和反应性。

[0449] 本文所述的化合物可包含一个或多个不对称中心,因此可以以各种立体异构形式存在,例如对映异构体和/或非对映异构体。例如,本文所述的化合物可以是单个对映异构体、非对映异构体或几何异构体的形式,或者可以是立体异构体的混合物的形式,包括外消旋混合物和富含一种或多种立体异构体的混合物。异构体可以通过本领域技术人员已知的方法从混合物中分离,包括手性高压液相色谱法(HPLC)以及手性盐的形成和结晶;或优选的异构体可以通过不对称合成制备。参见,例如,Jacques et al.,Enantiomers,Racemates and Resolutions(Wiley Interscience,New York,1981);Wilén et al.,Tetrahedron 33:2725(1977);Eliel,E.L.Stereochemistry of Carbon Compounds(McGraw-Hill,NY,1962);和Wilén,S.H.,Tables of Resolving Agents and Optical Resolutions p.268(E.L.Eliel,Ed.,Univ.of Notre Dame Press,Notre Dame,IN 1972)。本发明另外包括基本上没有其他异构体的单个异构体形式的化合物,或者各种异构体混合物形式的化合物。

[0450] 除非另有说明,否则本文所述的结构还旨在包括不同之处仅在于存在一个或多个同位素富集原子的化合物。例如,除了用氘或氚代替氢、用¹⁸F代替¹⁹F或用¹³C或¹⁴C代替¹²C之外,具有当前结构的化合物,都在本发明的范围内。这样的化合物例如可以用作生物学测定中的分析工具或探针。

[0451] 当列出值的范围时,其旨在涵盖该范围内的每个值和子范围。例如,“C₁₋₆烷基”旨

在涵盖 C_1 、 C_2 、 C_3 、 C_4 、 C_5 、 C_6 、 C_{1-6} 、 C_{1-5} 、 C_{1-4} 、 C_{1-3} 、 C_{1-2} 、 C_{2-6} 、 C_{2-5} 、 C_{2-4} 、 C_{2-3} 、 C_{3-6} 、 C_{3-5} 、 C_{3-4} 、 C_{4-6} 、 C_{4-5} 和 C_{5-6} 烷基。

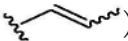
[0452] 术语“脂族基”是指烷基、烯基、炔基和碳环基团。同样，术语“杂脂族基”是指杂烷基、杂烯基、杂炔基和杂环基。

[0453] 术语“烷基”是指具有1至10个碳原子的直链或支链饱和烃基的基团（“ C_{1-10} 烷基”）。在一些实施方案中，烷基基团具有1至9个碳原子（“ C_{1-9} 烷基”）。在一些实施方案中，烷基基团具有1至8个碳原子（“ C_{1-8} 烷基”）。在一些实施方案中，烷基基团具有1至7个碳原子（“ C_{1-7} 烷基”）。在一些实施方案中，烷基基团具有1至6个碳原子（“ C_{1-6} 烷基”）。在一些实施方案中，烷基基团具有1至5个碳原子（“ C_{1-5} 烷基”）。在一些实施方案中，烷基基团具有1至4个碳原子（“ C_{1-4} 烷基”）。在一些实施方案中，烷基基团具有1至3个碳原子（“ C_{1-3} 烷基”）。在一些实施方案中，烷基基团具有1至2个碳原子（“ C_{1-2} 烷基”）。在一些实施方案中，烷基基团具有1个碳原子（“ C_1 烷基”）。在一些实施方案中，烷基基团具有2至6个碳原子（“ C_{2-6} 烷基”）。 C_{1-6} 烷基基团的实例包括甲基（ C_1 ）、乙基（ C_2 ）、丙基（ C_3 ）（例如，正丙基、异丙基）、丁基（ C_4 ）（例如，正丁基、叔丁基、仲丁基、异丁基）、戊基（ C_5 ）（例如，正戊基、3-戊基、戊基、新戊基、3-甲基-2-丁基、叔戊基）和己基（ C_6 ）（例如，正己基）。烷基基团的其他实例包括正庚基（ C_7 ）、正辛基（ C_8 ）等。除非另有说明，否则烷基基团的每种情况都独立地是未取代的（“未取代的烷基”）或被一个或多个取代基（例如，卤素，如F）取代的（“取代的烷基”）。在某些实施方案中，烷基基团是未取代的 C_{1-10} 烷基（例如，未取代的 C_{1-6} 烷基，例如- CH_3 （Me）、未取代的乙基（Et）、未取代的丙基（Pr，例如未取代的正丙基（n-Pr）、未取代的异丙基（i-Pr））、未取代的丁基（Bu，例如未取代的正丁基（n-Bu）、未取代的叔丁基（tert-Bu或t-Bu）、未取代的仲丁基（sec-Bu）、未取代的异丁基（i-Bu））。在某些实施方案中，烷基基团是取代的 C_{1-10} 烷基（例如取代的 C_{1-6} 烷基，例如- CF_3 、Bn）。

[0454] 术语“卤代烷基”是取代的烷基基团，其中一个或多个氢原子独立地被卤素，例如氟、溴、氯或碘取代。在一些实施方案中，卤代烷基部分具有1至8个碳原子（“ C_{1-8} 卤代烷基”）。在一些实施方案中，卤代烷基部分具有1至6个碳原子（“ C_{1-6} 卤代烷基”）。在一些实施方案中，卤代烷基部分具有1至4个碳原子（“ C_{1-4} 卤代烷基”）。在一些实施方案中，卤代烷基部分具有1至3个碳原子（“ C_{1-3} 卤代烷基”）。在一些实施方案中，卤代烷基部分具有1至2个碳原子（“ C_{1-2} 卤代烷基”）。卤代烷基的实例包括- CHF_2 、- CH_2F 、- CF_3 、- CH_2CF_3 、- CF_2CF_3 、- $CF_2CF_2CF_3$ 、- CCl_3 、- $CFCl_2$ 、- CF_2Cl 等。

[0455] 术语“杂烷基”是指一种烷基基团，其还包括至少一个（例如1、2、3或4个杂原子）选自氧、氮或硫的杂原子，所述杂原子在母链内（例如，插入到母链的相邻碳原子之间）和/或置于母链的一个或多个末端位置上。在某些实施方案中，杂烷基基团是指在母链内具有1至10个碳原子和1个或多个杂原子的饱和基团（“杂 C_{1-10} 烷基”）。在一些实施方案中，杂烷基基团是在母链内具有1至9个碳原子和1个或多个杂原子的饱和基团（“杂 C_{1-9} 烷基”）。在一些实施方案中，杂烷基基团是在母链内具有1至8个碳原子和1个或多个杂原子的饱和基团（“杂 C_{1-8} 烷基”）。在一些实施方案中，杂烷基基团是在母链内具有1至7个碳原子和1个或多个杂原子的饱和基团（“杂 C_{1-7} 烷基”）。在一些实施方案中，杂烷基基团是在母链内具有1至6个碳原子和1个或多个杂原子的饱和基团（“杂 C_{1-6} 烷基”）。在一些实施方案中，杂烷基基团是在母链内具有1至5个碳原子和1个或2个杂原子的饱和基团（“杂 C_{1-5} 烷基”）。在一些实施方案

中,杂烷基基团是在母链内具有1至4个碳原子和1个或2个杂原子的饱和基团(“杂C₁₋₄烷基”)。在一些实施方案中,杂烷基基团是在母链内具有1至3个碳原子和1个杂原子的饱和基团(“杂C₁₋₃烷基”)。在一些实施方案中,杂烷基基团是在母链内具有1至2个碳原子和1个杂原子的饱和基团(“杂C₁₋₂烷基”)。在一些实施方案中,杂烷基基团是具有1个碳原子和1个杂原子的饱和基团(“杂C₁烷基”)。在一些实施方案中,杂烷基基团是在母链内具有2至6个碳原子和1个或2个杂原子的饱和基团(“杂C₂₋₆烷基”)。除非另有说明,否则杂烷基基团的每种情况都独立地是未取代的(“未取代的杂烷基”)或被一个或多个取代基取代的(“取代的杂烷基”)。在某些实施方案中,杂烷基基团是未取代的杂C₁₋₁₀烷基。在某些实施方案中,杂烷基基团是取代的杂C₁₋₁₀烷基。

[0456] 术语“烯基”是指具有2至10个碳原子和一个或多个碳-碳双键(例如1、2、3或4个双键)的直链或支链烃基的基团。在一些实施方案中,烯基基团具有2至9个碳原子(“C₂₋₉烯基”)。在一些实施方案中,烯基基团具有2至8个碳原子(“C₂₋₈烯基”)。在一些实施方案中,烯基基团具有2至7个碳原子(“C₂₋₇烯基”)。在一些实施方案中,烯基基团具有2至6个碳原子(“C₂₋₆烯基”)。在一些实施方案中,烯基基团具有2至5个碳原子(“C₂₋₅烯基”)。在一些实施方案中,烯基基团具有2至4个碳原子(“C₂₋₄烯基”)。在一些实施方案中,烯基基团具有2至3个碳原子(“C₂₋₃烯基”)。在一些实施方案中,烯基基团具有2个碳原子(“C₂烯基”)。所述一个或多个碳-碳双键可以是内部的(例如在2-丁烯基中)或末端的(例如在1-丁烯基中)。C₂₋₄烯基基团的实例包括乙烯基(C₂)、1-丙烯基(C₃)、2-丙烯基(C₃)、1-丁烯基(C₄)、2-丁烯基(C₄)、丁二烯基(C₄)等。C₂₋₆烯基基团的实例包括上述的C₂₋₄烯基基团以及戊烯基(C₅)、戊二烯基(C₅)、己烯基(C₆)等。烯基的其他实例包括庚烯基(C₇)、辛烯基(C₈)、八三烯基(C₈)等。除非另有说明,否则烯基基团的每种情况都独立地是未取代的(“未取代的烯基”)或被一个或多个取代基取代的(“取代的烯基”)。在某些实施方案中,烯基基团是未取代的C₂₋₁₀烯基。在某些实施方案中,烯基基团是取代的C₂₋₁₀烯基。在烯基基团中,未指定其立体化学的C=C双键(例如, -CH=CHCH₃或)可以是(E)-或(Z)-双键。

[0457] 术语“杂烯基”是指一种烯基基团,其还包括至少一个(例如1、2、3或4个杂原子)选自氧、氮或硫的杂原子,所述杂原子在母链内(例如,插入到母链的相邻碳原子之间)和/或置于母链的一个或多个末端位置上。在某些实施方案中,杂烯基基团是在母链中具有2至10个碳原子、至少一个双键和1个或多个杂原子的基团(“杂C₂₋₁₀烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至9个碳原子、至少一个双键和1个或多个杂原子的基团(“杂C₂₋₉烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至8个碳原子、至少一个双键和1个或多个杂原子的基团(“杂C₂₋₈烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至7个碳原子、至少一个双键和1个或多个杂原子的基团(“杂C₂₋₇烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至6个碳原子、至少一个双键和1个或多个杂原子的基团(“杂C₂₋₆烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至5个碳原子、至少一个双键和1个或2个杂原子的基团(“杂C₂₋₅烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至4个碳原子、至少一个双键和1个或2个杂原子的基团(“杂C₂₋₄烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至3个碳原子、至少一个双键和1个杂原子的基团(“杂C₂₋₃烯基”)。在一些实施方案中,杂烯基基团是在母链中具有2至6个碳原子、至少一个双键和1个或2个杂原子的基团(“杂C₂₋₆烯基”)。除非另有说明,否则杂烯基基团的每种情况都独

立地是未取代的(“未取代的杂烯基”)或被一个或多个取代基取代的(“取代的杂烯基”)。在某些实施方案中,杂烯基基团是未取代的杂 C_{2-10} 烯基。在某些实施方案中,杂烯基基团是取代的杂 C_{2-10} 烯基。

[0458] 术语“炔基”是指具有2至10个碳原子和一个或多个碳-碳三键(例如1、2、3或4个三键)的直链或支链烃基的基团(“ C_{2-10} 炔基”)。在一些实施方案中,炔基基团具有2至9个碳原子(“ C_{2-9} 炔基”)。在一些实施方案中,炔基基团具有2至8个碳原子(“ C_{2-8} 炔基”)。在一些实施方案中,炔基基团具有2至7个碳原子(“ C_{2-7} 炔基”)。在一些实施方案中,炔基基团具有2至6个碳原子(“ C_{2-6} 炔基”)。在一些实施方案中,炔基基团具有2至5个碳原子(“ C_{2-5} 炔基”)。在一些实施方案中,炔基基团具有2至4个碳原子(“ C_{2-4} 炔基”)。在一些实施方案中,炔基基团具有2至3个碳原子(“ C_{2-3} 炔基”)。在一些实施方案中,炔基基团具有2个碳原子(“ C_2 炔基”)。所述一个或多个碳-碳三键可以是内部的(例如在2-丁炔基中)或末端的(例如在1-丁炔基中)。 C_{2-4} 炔基基团的实例包括但不限于乙炔基(C_2)、1-丙炔基(C_3)、2-丙炔基(C_3)、1-丁炔基(C_4)、2-丁炔基(C_4)等。 C_{2-6} 炔基基团的实例包括上述的 C_{2-4} 炔基以及戊炔基(C_5)、己炔基(C_6)等。炔基的其他实例包括庚炔基(C_7)、辛炔基(C_8)等。除非另有说明,否则炔基基团的每种情况都独立地是未取代的(“未取代的炔基”)或被一个或多个取代基取代的(“取代的炔基”)。在某些实施方案中,炔基基团是未取代的 C_{2-10} 炔基。在某些实施方案中,炔基基团是取代的 C_{2-10} 炔基。

[0459] 术语“杂炔基”是指一种炔基基团,其还包括至少一个(例如1、2、3或4个杂原子)选自氧、氮或硫的杂原子,所述杂原子在母链内(例如,插入到母链的相邻碳原子之间)和/或置于母链的一个或多个末端位置上。在某些实施方案中,杂炔基基团是指在母链中具有2至10个碳原子、至少一个三键和1个或多个杂原子的基团(“杂 C_{2-10} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至9个碳原子、至少一个三键和1个或多个杂原子(“杂 C_{2-9} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至8个碳原子、至少一个三键和1个或多个杂原子(“杂 C_{2-8} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至7个碳原子、至少一个三键和1个或多个杂原子(“杂 C_{2-7} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至6个碳原子、至少一个三键和1个或多个杂原子(“杂 C_{2-6} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至5个碳原子、至少一个三键和1个或2个杂原子(“杂 C_{2-5} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至4个碳原子、至少一个三键和1个或2个杂原子(“杂 C_{2-4} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至3个碳原子、至少一个三键和1个杂原子(“杂 C_{2-3} 炔基”)。在一些实施方案中,杂炔基基团在母链中具有2至6个碳原子、至少一个三键和1个或2个杂原子(“杂 C_{2-6} 炔基”)。除非另有说明,否则杂炔基基团的每种情况都独立地是未取代的(“未取代的杂炔基”)或被一个或多个取代基取代的(“取代的杂炔基”)。在某些实施方案中,杂炔基基团是未取代的杂基 C_{2-10} 炔基。在某些实施方案中,杂炔基基团是取代的杂 C_{2-10} 炔基。

[0460] 术语“碳环基”或“碳环型”是指在非芳环体系中具有3至14个环碳原子(C_{3-14} 碳环基)和零个杂原子的非芳环烃基基团的基团。在一些实施方案中,碳环基基团具有3至10个环碳原子(“ C_{3-10} 碳环基”)。在一些实施方案中,碳环基基团具有3至8个环碳原子(“ C_{3-8} 碳环基”)。在一些实施方案中,碳环基基团具有3至7个环碳原子(“ C_{3-7} 碳环基”)。在一些实施方案中,碳环基基团具有3至6个环碳原子(“ C_{3-6} 碳环基”)。在一些实施方案中,碳环基基团具

有4至6个环碳原子(“C₄₋₆碳环基”)。在一些实施方案中,碳环基基团具有5至6个环碳原子(“C₅₋₆碳环基”)。在一些实施方案中,碳环基基团具有5至10个环碳原子(“C₅₋₁₀碳环基”)。示例性的C₃₋₆碳环基基团包括但不限于环丙基(C₃)、环丙烯基(C₃)、环丁基(C₄)、环丁烯基(C₄)、环戊基(C₅)、环戊烯基(C₅)、环己基(C₆)、环己烯基(C₆)、环己二烯基(C₆)等。示例性的C₃₋₈碳环基基团包括但不限于上述C₃₋₆碳环基基团以及环庚基(C₇)、环庚烯基(C₇)、环庚二烯基(C₇)、环庚三烯基(C₇)、环辛基(C₈)、环辛烯基(C₈)、双环[2.2.1]庚基(C₇)、双环[2.2.2]辛基(C₈)等。示例性的C₃₋₁₀碳环基基团包括但不限于上述C₃₋₈碳环基基团以及环壬基(C₉)、环壬烯基(C₉)、环癸基(C₁₀)、环癸烯基(C₁₀)、八氢-1H-茛基(C₉)、十氢萘基(C₁₀)、螺[4.5]癸基(C₁₀)等。如上述示例所示,在某些实施方案中,碳环基基团是单环的(“单环碳环基”)或多环的(例如,含有稠和环、桥接环或螺环系统,如双环系统(“双环碳环基”)或三环系统(“三环碳环基”)),可以是饱和的,也可以含有一个或多个碳-碳双键或三键。“碳环基”还包括环系统,其中如本文所定义的碳环基环与一个或多个芳基或杂芳基基团稠合,其中连接的点在碳环基环上,并且在这种情况下,碳的数目继续表示碳环型环系统中的碳的数目。除非另有说明,否则碳环基基团的每种情况都独立地是未取代的(“未取代的碳环基”)或被一个或多个取代基取代的(“取代的碳环基”)。在某些实施方案中,碳环基基团是未取代的C₃₋₁₄碳环基。在某些实施方案中,碳环基基团是取代的C₃₋₁₄碳环基。

[0461] 在一些实施方案中,“碳环基”是具有3至14个环碳原子的单环饱和碳环基(“C₃₋₁₄环烷基”)。在一些实施方案中,环烷基基团具有3至10个环碳原子(“C₃₋₁₀环烷基”)。在一些实施方案中,环烷基基团具有3至8个环碳原子(“C₃₋₈环烷基”)。在一些实施方案中,环烷基基团具有3至6个环碳原子(“C₃₋₆环烷基”)。在一些实施方案中,环烷基基团具有4至6个环碳原子(“C₄₋₆环烷基”)。在一些实施方案中,环烷基基团具有5至6个环碳原子(“C₅₋₆环烷基”)。在一些实施方案中,环烷基基团具有5至10个环碳原子(“C₅₋₁₀环烷基”)。C₅₋₆环烷基基团的实例包括环戊基(C₅)和环己基(C₆)。C₃₋₆环烷基基团的实例包括上述的C₅₋₆环烷基基团以及环丙基(C₃)和环丁基(C₄)。C₃₋₈环烷基基团的实例包括上述的C₃₋₆环烷基基团以及环庚基(C₇)和环辛基(C₈)。除非另有说明,否则环烷基基团的每种情况都独立地是未取代的(“未取代的环烷基”)或被一个或多个取代基取代的(“取代的环烷基”)。在某些实施方案中,环烷基基团是未取代的C₃₋₁₄环烷基。在某些实施方案中,环烷基基团是取代的C₃₋₁₄环烷基。

[0462] 术语“杂环基”或“杂环型”是指具有环碳原子和1至4个环杂原子的3至14元非芳族环系统的基团,其中每个杂原子独立地选自氮、氧和硫(“3-14元杂环基”)。在化合价允许的情况下,在包括一个或多个氮原子的杂环基中,连接点可以是碳或氮原子。杂环基基团可以是单环的(“单环杂环基”)或多环的(例如,稠和环、桥接环或螺环系统,如双环系统(“双环杂环基”)或三环系统(“三环杂环基”)),可以是饱和的,也可以含有一个或多个碳-碳双键或三键。杂环基多环体系可以在一个或两个环中包括一个或多个杂原子。“杂环基”还包括环系统,其中如本文所定义的杂环基环与一个或多个碳环基基团稠合,其中连接点在碳环基或杂环基环上,或者包括环系统,其中如本文所定义的杂环基环与一个或多个芳基或杂芳基基团稠合,其中连接点在杂环基环上,并且在这种情况下,环成员的数目继续表示杂环基环系统中环成员的数目。除非另有说明,否则杂环基的每种情况都独立地是未取代的(“未取代的杂环基”)或被一个或多个取代基取代的(“取代的杂环基”)。在某些实施方案中,杂环基基团是未取代的3-14元杂环基。在某些实施方案中,杂环基基团是取代的3-14元

杂环基。

[0463] 在一些实施方案中,杂环基基团是具有环碳原子和1-4个环杂原子的5-10元非芳族环系统,其中每个杂原子独立地选自氮、氧和硫(“5-10元杂环基”)。在一些实施方案中,杂环基基团是具有环碳原子和1-4个环杂原子的5-8元非芳族环系统,其中每个杂原子独立地选自氮、氧和硫(“5-8元杂环基”)。在一些实施方案中,杂环基基团是具有环碳原子和1-4个环杂原子的5-6元非芳族环系统,其中每个杂原子独立地选自氮、氧和硫(“5-6元杂环基”)。在一些实施方案中,5-6元杂环基具有1-3个选自氮、氧和硫的环杂原子。在一些实施方案中,5-6元杂环基具有1-2个选自氮、氧和硫的环杂原子。在一些实施方案中,5-6元杂环基具有1个选自氮、氧和硫的环杂原子。

[0464] 含有1个杂原子的示例性3元杂环基基团包括但不限于氮杂环丙烷基(azirdinyl)、氧杂环丙烷基(oxiranyl)、硫杂环丙烷基(thiiranyl)。包含1个杂原子的示例性4元杂环基基团包括但不限于氮杂环丁烷基(azetidiny)、氧杂环丁烷基(oxetanyl)和硫杂环丁烷基(thietanyl)。含有1个杂原子的示例性5元杂环基基团包括但不限于四氢呋喃基、二氢呋喃基、四氢噻吩基、二氢噻吩基、吡咯烷基、二氢吡咯基和吡咯基-2,5-二酮。含有2个杂原子的示例性5元杂环基基团包括但不限于二氧杂环戊烷基(dioxolanyl)、氧杂硫杂环戊烷基(oxathiolanyl)和二硫杂环戊烷基(dithiolanyl)。含有3个杂原子的示例性5元杂环基基团包括但不限于三唑啉基、噁二唑啉基(oxadiazoliny)和噻二唑啉基(thiadiazoliny)。含有1个杂原子的示例性6元杂环基基团包括但不限于哌啶基、四氢吡喃基、二氢吡啶基和硫杂环己烷基(thianyl)。包含2个杂原子的示例性6元杂环基基团包括但不限于哌嗪基、吗啉基、二硫杂环己烷基和二氧杂环己烷基(dioxanyl)。含有3个杂原子的示例性6元杂环基基团包括但不限于三嗪基。含有1个杂原子的示例性7元杂环基基团包括但不限于氮杂环庚烷基(azepanyl)、氧杂环庚烷基(oxepanyl)和硫杂环庚烷基(thiepanyl)。含有1个杂原子的示例性8元杂环基基团包括但不限于氮杂环辛烷基(azocanyl)、氧杂环辛烷基(oxecanyl)和硫杂环辛烷基(thiocanyl)。示例性的双环杂环基基团包括但不限于吲哚基、异吲哚基、二氢苯并呋喃基、二氢苯并噻吩基、四氢苯并噻吩基、四氢苯并呋喃基、四氢吲哚基、四氢喹啉基、四氢异喹啉基、十氢喹啉基、十氢异喹啉基、八氢色烯基(octahydrochromenyl)、八氢异色烯基、十氢萘啶基、十氢-1,8-萘啶基、八氢吡咯并[3,2-b]吡咯、吲哚基、邻苯二甲酰亚胺基、萘二甲酰亚胺基、色满基(chromanyl)、色烯基、1H-苯并[e][1,4]二氮杂卓基、1,4,5,7-四氢吡喃并[3,4-b]吡咯基、5,6-二氢-4H-呋喃并[3,2-b]吡咯基、6,7-二氢-5H-呋喃并[3,2-b]吡喃基、5,7-二氢-4H-噻吩并[2,3-c]吡喃基、2,3-二氢-1H-吡咯并[2,3-b]吡啶基、2,3-二氢呋喃[2,3-b]吡啶基、4,5,6,7-四氢-1H-吡咯并[2,3-b]吡啶基、4,5,6,7-四氢呋喃[3,2-c]吡啶基、4,5,6,7-四氢噻吩并[3,2-b]吡啶基、1,2,3,4-四氢-1,6-萘啶基等。

[0465] 术语“芳基”是指在芳香环系统中具有6-14个环碳原子和零个杂原子的单环或多环(如:双环或三环) $4n+2$ 芳环体系(例如,在环状阵列中具有6个、10个或14个共享电子)的基团(“ C_{6-14} 芳基”)。在一些实施方案中,芳基基团具有6个环碳原子(“ C_6 芳基”;例如苯基)。在一些实施方案中,芳基基团具有10个环碳原子(“ C_{10} 芳基”;例如,萘基,例如1-萘基和2-萘基)。在一些实施方案中,芳基基团具有14个环碳原子(“ C_{14} 芳基”;例如,蒽基)。“芳基”还包括环系统,其中如本文所定义的芳基环与一个或多个碳环基或杂环基稠合,其中连接的基

团或点在芳基环上,并且在这种情况下,碳原子数继续表示芳基环系统中的碳原子数。除非另有说明,否则芳基基团的每种情况都独立地是未取代的(“未取代的芳基”)或被一个或多个取代基取代的(“取代的芳基”)。在某些实施方案中,芳基基团是未取代的C₆₋₁₄芳基。在某些实施方案中,芳基基团是取代的C₆₋₁₄芳基。

[0466] 术语“杂芳基”是指在芳香环系统中具有环碳原子和1-4个环杂原子的5-14元单环或多环(如:双环、三环)4n+2芳环体系(例如,在环状阵列中具有6个、10个或14个共享电子)的基团。其中每个杂原子独立地选自氮、氧和硫(“5-14元杂芳基”)。在化合价允许的情况下,在包括一个或多个氮原子的杂芳基中,连接点可以是碳或氮原子。杂芳基多环体系可以在一个或两个环中包括一个或多个杂原子。“杂芳基”包括环系统,其中如本文所定义的杂芳基环与一个或多个碳环基或杂环基稠合,其中连接的点在杂芳基环上,并且在这种情况下,环成员的数目继续表示杂芳基环系统中环成员的数目。“杂芳基”还包括环系统,其中如上所定义的杂芳基环与一个或多个芳基基团稠合,其中连接点在芳基或杂芳基环上,并且在这种情况下,环成员的数目表示稠合多环(芳基/杂芳基)环系统中环成员的数目。其中一个环不包含杂原子的多环杂芳基基团(例如,吡啶基、喹啉基、咪唑基等),连接点可以在任一环上,例如,带有杂原子的环(例如2-吡啶基)或不包含杂原子的环(例如5-吡啶基)。

[0467] 在一些实施方案中,杂芳基基团是在芳香族环体系中具有环碳原子和1-4个环杂原子的5-10元芳香族环体系,其中每个杂原子独立地选自氮、氧和硫(“5-10元杂芳基”)。在一些实施方案中,杂芳基基团是在芳香族环体系中具有环碳原子和1-4个环杂原子的5-8元芳香族环体系,其中每个杂原子独立地选自氮、氧和硫(“5-8元杂芳基”)。在一些实施方案中,杂芳基基团是在芳香族环体系中具有环碳原子和1-4个环杂原子的5-6元芳香族环体系,其中每个杂原子独立地选自氮、氧和硫(“5-6元杂芳基”)。在一些实施方案中,5-6元杂芳基具有1-3个选自氮、氧和硫的环杂原子。在一些实施方案中,5-6元杂芳基具有1-2个选自氮、氧和硫的环杂原子。在一些实施方案中,5-6元杂芳基具有1个选自氮、氧和硫的环杂原子。除非另有说明,否则杂芳基基团的每种情况都独立地是未取代的(“未取代的杂芳基”)或被一个或多个取代基取代的(“取代的杂芳基”)。在某些实施方案中,杂芳基基团是未取代的5-14元杂芳基。在某些实施方案中,杂芳基基团是取代的5-14元杂芳基。

[0468] 含有1个杂原子的示例性5元杂芳基基团包括但不限于吡咯基、呋喃基和噻吩基。含有2个杂原子的示例性5元杂芳基基团包括但不限于咪唑基、吡唑基、噁唑基、异噁唑基、噻唑基和异噻唑基。含有3个杂原子的示例性5元杂芳基包括但不限于三唑基、噁二唑基和噻二唑基。含有4个杂原子的示例性5元杂芳基包括但不限于四唑基。含有1个杂原子的示例性6元杂芳基包括但不限于吡啶基。包括2个杂原子的示例性6元杂芳基包括但不限于哒嗪基、嘧啶基和吡嗪基。含有3或4个杂原子的示例性6元杂芳基分别包括但不限于三嗪基和四嗪基。含有1个杂原子的示例性7元杂芳基包括但不限于氮杂卓基(azepinyl)、氧杂卓基(oxepinyl)和硫杂卓基(thiepinyl)。示例性的5,6-双环杂芳基基团包括但不限于吡啶基、异吡啶基、吡唑基、苯并三唑基、苯并噻吩基、异苯并噻吩基、苯并呋喃基、苯并异呋喃基、苯并咪唑基、苯并噁唑基、苯并异噁唑基、苯并噁二唑基、苯并噻唑基、苯并异噻唑基、苯并噻二唑基、吡啶嗪基(indoliziny)和嘌呤基。示例性的6,6-双环杂芳基基团包括但不限于萘啶基、蝶啶基、喹啉基、异喹啉基、噌啉基(cinnolinyl)、喹喔啉基、酞嗪基(phthalazinyl)和喹唑啉基。示例性的三环杂芳基基团包括但不限于菲基吡啶基、二苯并呋喃基、咪唑基、

吡啶基、吩噻嗪基、吩噁嗪基和吩嗪基。

[0469] 术语“不饱和键”是指双键或三键。

[0470] 术语“不饱和的”或“部分不饱和的”是指包括至少一个双键或三键的部分。

[0471] 术语“饱和的”是指不包括双键或三键的部分,即,该部分仅包括单键。

[0472] 在基团上附加前缀“-亚”表示该基团是二价部分,例如,亚烷基是烷基的二价部分,亚烯基是烯基的二价部分,亚炔基是炔基的二价部分,亚杂烷基是杂烷基的二价部分,亚杂烯基是杂烯基的二价部分,亚杂炔基是杂炔基的二价部分,亚碳环基是碳环基的二价部分,亚杂环基是杂环基的二价部分,亚芳基是芳基的二价部分,亚杂芳基是杂芳基的二价部分。

[0473] 除非另有明确说明,否则基团是任选取代的。术语“任选取代的”是指被取代或未被取代。在某些实施方案中,烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基基团是任选取代的。“任选取代的”是指可以被取代或未被取代的基团(例如,“取代的”或“未取代的”烷基、“取代的”或“未取代的”烯基、“取代的”或“未取代的”炔基、“取代的”或“未取代的”杂烷基、“取代的”或“未取代的”杂烯基、“取代的”或“未取代的”杂炔基、“取代的”或“未取代的”碳环基、“取代的”或“未取代的”杂环基、“取代的”或“未取代的”芳基或“取代的”或“未取代的”杂芳基基团)。通常,术语“取代的”是指基团上存在的至少一个氢被允许的取代基取代,例如,取代基在取代时会产生稳定的化合物,例如不会通过例如重排、环化、消除或其他反应自发进行转化的化合物。除非另有说明,否则“取代的”基团在该基团的一个或多个可取代的位置上具有取代基,并且当任何给定结构中的一个以上的位置被取代时,该取代基在每个位置上可以相同或不同。预期术语“取代的”包括被有机化合物的所有允许的取代基取代,并且包括导致形成稳定化合物的本文所述的任何取代基。本发明考虑了任何和所有这样的组合以便获得稳定的化合物。为了本发明的目的,杂原子如氮可以具有氢取代基和/或满足杂原子化合价并导致形成稳定部分的本文所述的任何合适的取代基。本发明无意以任何方式受限于本文所述的示例性取代基。

[0474] 示例性的碳原子取代基包括但不限于卤素、-CN、-NO₂、-N₃、-SO₂H、-SO₃H、-OH、-OR^{aa}、-ON(R^{bb})₂、-N(R^{bb})₂、-N(R^{bb})₃⁺X⁻、-N(OR^{cc})R^{bb}、-SH、-SR^{aa}、-SSR^{cc}、-C(=O)R^{aa}、-CO₂H、-CHO、-C(OR^{cc})₃、-CO₂R^{aa}、-OC(=O)R^{aa}、-OCO₂R^{aa}、-C(=O)N(R^{bb})₂、-OC(=O)N(R^{bb})₂、-NR^{bb}C(=O)R^{aa}、-NR^{bb}CO₂R^{aa}、-NR^{bb}C(=O)N(R^{bb})₂、-C(=NR^{bb})R^{aa}、-C(=NR^{bb})OR^{aa}、-OC(=NR^{bb})R^{aa}、-OC(=NR^{bb})OR^{aa}、-C(=NR^{bb})N(R^{bb})₂、-OC(=NR^{bb})N(R^{bb})₂、-NR^{bb}C(=NR^{bb})N(R^{bb})₂、-C(=O)NR^{bb}SO₂R^{aa}、-NR^{bb}SO₂R^{aa}、-SO₂N(R^{bb})₂、-SO₂R^{aa}、-SO₂OR^{aa}、-OSO₂R^{aa}、-S(=O)R^{aa}、-OS(=O)R^{aa}、-Si(R^{aa})₃、-OSi(R^{aa})₃-C(=S)N(R^{bb})₂、-C(=O)SR^{aa}、-C(=S)SR^{aa}、-SC(=S)SR^{aa}、-SC(=O)SR^{aa}、-OC(=O)SR^{aa}、-SC(=O)OR^{aa}、-SC(=O)R^{aa}、-P(=O)(R^{aa})₂、-P(=O)(OR^{cc})₂、-OP(=O)(R^{aa})₂、-OP(=O)(OR^{cc})₂、-P(=O)(N(R^{bb})₂)₂、-OP(=O)(N(R^{bb})₂)₂、-NR^{bb}P(=O)(R^{aa})₂、-NR^{bb}P(=O)(OR^{cc})₂、-NR^{bb}P(=O)(N(R^{bb})₂)₂、-P(R^{cc})₂、-P(OR^{cc})₂、-P(R^{cc})₃⁺X⁻、-P(OR^{cc})₃⁺X⁻、-P(R^{cc})₄、-P(OR^{cc})₄、-OP(R^{cc})₂、-OP(R^{cc})₃⁺X⁻、-OP(OR^{cc})₂、-OP(OR^{cc})₃⁺X⁻、-OP(R^{cc})₄、

[0475] BY20IM3317FGPC-CN. 电话OA-OP(OR^{cc})₄、-B(R^{aa})₂、-B(OR^{cc})₂、-BR^{aa}(OR^{cc})、C₁₋₁₀烷基、C₁₋₁₀全卤代烷基、C₂₋₁₀烯基、C₂₋₁₀炔基、杂C₁₋₁₀烷基、杂C₂₋₁₀烯基、杂C₂₋₁₀炔基、C₃₋₁₀碳环基、3-14元杂环基、C₆₋₁₄芳基和5-14元杂芳基,其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{dd}基团取代;其中X⁻是抗

衡离子；

[0476] 或者碳原子上的两个孪位氢原子被基团=O、=S、=NN(R^{bb})₂、=NNR^{bb}C(=O)R^{aa}、=NNR^{bb}C(=O)OR^{aa}、=NNR^{bb}S(=O)₂R^{aa}、=NR^{bb}或=NOR^{cc}取代；

[0477] R^{aa}的每个情况都独立地选自C₁₋₁₀烷基、C₁₋₁₀全卤代烷基、C₂₋₁₀烯基、C₂₋₁₀炔基、杂C₁₋₁₀烷基、杂C₂₋₁₀烯基、杂C₂₋₁₀炔基、C₃₋₁₀碳环基、3-14元杂环基、C₆₋₁₄芳基和5-14元杂芳基，或者两个R^{aa}基团连接形成3-14元杂环基或5-14元杂芳基环，其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{dd}基团取代；

[0478] R^{bb}的每个情况都独立地选自氢、-OH、-OR^{aa}、-N(R^{cc})₂、-CN、-C(=O)R^{aa}、-C(=O)N(R^{cc})₂、-CO₂R^{aa}、-SO₂R^{aa}、-C(=NR^{cc})OR^{aa}、-C(=NR^{cc})N(R^{cc})₂、-SO₂N(R^{cc})₂、-SO₂R^{cc}、-SO₂OR^{cc}、-SOR^{aa}、-C(=S)N(R^{cc})₂、-C(=O)SR^{cc}、-C(=S)SR^{cc}、-P(=O)(R^{aa})₂、-P(=O)(OR^{cc})₂、-P(=O)(N(R^{cc})₂)₂、C₁₋₁₀烷基、C₁₋₁₀全卤代烷基、C₂₋₁₀烯基、C₂₋₁₀炔基、杂C₁₋₁₀烷基、杂C₂₋₁₀烯基、杂C₂₋₁₀炔基、C₃₋₁₀碳环基、3-14元杂环基、C₆₋₁₄芳基和5-14元杂芳基，或者两个R^{bb}基团连接形成3-14元杂环基或5-14元杂芳基环，其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{dd}基团取代；其中X⁻是抗衡离子；

[0479] R^{cc}的每个情况都独立地选自氢、C₁₋₁₀烷基、C₁₋₁₀全卤代烷基、C₂₋₁₀烯基、C₂₋₁₀炔基、杂C₁₋₁₀烷基、杂C₂₋₁₀烯基、杂C₂₋₁₀炔基、C₃₋₁₀碳环基、3-14元杂环基、C₆₋₁₄芳基和5-14元杂芳基，或者两个R^{cc}基团连接形成3-14元杂环基或5-14元杂芳基环，其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{dd}基团取代；

[0480] R^{dd}的每个情况都独立地选自卤素、-CN、-NO₂、-N₃、-SO₂H、-SO₃H、-OH、-OR^{ee}、-ON(R^{ff})₂、-N(R^{ff})₂、-N(R^{ff})₃⁺X⁻、-N(OR^{ee})R^{ff}、-SH、-SR^{ee}、-SSR^{ee}、-C(=O)R^{ee}、-CO₂H、-CO₂R^{ee}、-OC(=O)R^{ee}、-OCO₂R^{ee}、-C(=O)N(R^{ff})₂、-OC(=O)N(R^{ff})₂、-NR^{ff}C(=O)R^{ee}、-NR^{ff}CO₂R^{ee}、-NR^{ff}C(=O)N(R^{ff})₂、-C(=NR^{ff})OR^{ee}、-OC(=NR^{ff})R^{ee}、-OC(=NR^{ff})OR^{ee}、-C(=NR^{ff})N(R^{ff})₂、-OC(=NR^{ff})N(R^{ff})₂、-NR^{ff}C(=NR^{ff})N(R^{ff})₂、-NR^{ff}SO₂R^{ee}、-SO₂N(R^{ff})₂、-SO₂R^{ee}、-SO₂OR^{ee}、-OSO₂R^{ee}、-S(=O)R^{ee}、-Si(R^{ee})₃、-OSi(R^{ee})₃、-C(=S)N(R^{ff})₂、-C(=O)SR^{ee}、-C(=S)SR^{ee}、-SC(=S)SR^{ee}、-P(=O)(OR^{ee})₂、-P(=O)(R^{ee})₂、-OP(=O)(R^{ee})₂、-OP(=O)(OR^{ee})₂、C₁₋₆烷基、C₁₋₆全卤代烷基、C₂₋₆烯基、C₂₋₆炔基、杂C₁₋₆烷基、杂C₂₋₆烯基、杂C₂₋₆炔基、C₃₋₁₀碳环基、3-10元杂环基、C₆₋₁₀芳基、5-10元杂芳基，其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{gg}基团取代，或者两个孪位R^{dd}取代基可以连接形成=O或=S；其中X⁻是抗衡离子；

[0481] R^{ee}的每个情况都独立地选自C₁₋₆烷基、C₁₋₆全卤代烷基、C₂₋₆烯基、C₂₋₆炔基、杂C₁₋₆烷基、杂C₂₋₆烯基、杂C₂₋₆炔基、C₃₋₁₀碳环基、C₆₋₁₀芳基、3-10元杂环基和3-10元杂芳基，其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{gg}基团取代；

[0482] R^{ff}的每个情况都独立地选自氢、C₁₋₆烷基、C₁₋₆全卤代烷基、C₂₋₆烯基、C₂₋₆炔基、杂C₁₋₆烷基、杂C₂₋₆烯基、杂C₂₋₆炔基、C₃₋₁₀碳环基、3-10元杂环基、C₆₋₁₀芳基和5-10元杂芳基，或者两个R^{ff}基团连接形成3-10元杂环基或5-10元杂芳基环，其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{gg}基团取代；

并且

[0483] R^{sg} 的每个情况都独立地选自卤素、-CN、-NO₂、-N₃、-SO₂H、-SO₃H、-OH、-OC₁₋₆烷基、-ON(C₁₋₆烷基)₂、-N(C₁₋₆烷基)₂、-N(C₁₋₆烷基)₃⁺X⁻、-NH(C₁₋₆烷基)₂⁺X⁻、-NH₂(C₁₋₆烷基)⁺X⁻、-NH₃⁺X⁻、-N(OC₁₋₆烷基)(C₁₋₆烷基)、-N(OH)(C₁₋₆烷基)、-NH(OH)、-SH、-SC₁₋₆烷基、-SS(C₁₋₆烷基)、-C(=O)(C₁₋₆烷基)、-CO₂H、-CO₂(C₁₋₆烷基)、-OC(=O)(C₁₋₆烷基)、-OCO₂(C₁₋₆烷基)、-C(=O)NH₂、-C(=O)N(C₁₋₆烷基)₂、-OC(=O)NH(C₁₋₆烷基)、-NHC(=O)(C₁₋₆烷基)、-N(C₁₋₆烷基)C(=O)(C₁₋₆烷基)、-NHCO₂(C₁₋₆烷基)、-NHC(=O)N(C₁₋₆烷基)₂、-NHC(=O)NH(C₁₋₆烷基)、-NHC(=O)NH₂、-C(=NH)O(C₁₋₆烷基)、-OC(=NH)(C₁₋₆烷基)、-OC(=NH)OC₁₋₆烷基、-C(=NH)N(C₁₋₆烷基)₂、-C(=NH)NH(C₁₋₆烷基)、-C(=NH)NH₂、-OC(=NH)N(C₁₋₆烷基)₂、-OC(=NH)NH(C₁₋₆烷基)、-OC(=NH)NH₂、-NHC(=NH)N(C₁₋₆烷基)₂、-NHC(=NH)NH₂、-NHSO₂(C₁₋₆烷基)、-SO₂N(C₁₋₆烷基)₂、-SO₂NH(C₁₋₆烷基)、-SO₂NH₂、-SO₂(C₁₋₆烷基)、-SO₂O(C₁₋₆烷基)、-OSO₂(C₁₋₆烷基)、-SO(C₁₋₆烷基)、-Si(C₁₋₆烷基)₃、-OSi(C₁₋₆烷基)₃-C(=S)N(C₁₋₆烷基)₂、C(=S)NH(C₁₋₆烷基)、C(=S)NH₂、-C(=O)S(C₁₋₆烷基)、-C(=S)SC₁₋₆烷基、-SC(=S)SC₁₋₆烷基、-P(=O)(OC₁₋₆烷基)₂、-P(=O)(C₁₋₆烷基)₂、-OP(=O)(C₁₋₆烷基)₂、-OP(=O)(OC₁₋₆烷基)₂、C₁₋₆烷基、C₁₋₆全卤代烷基、C₂₋₆烯基、C₂₋₆炔基、杂C₁₋₆烷基、杂C₂₋₆烯基、杂C₂₋₆炔基、C₃₋₁₀碳环基、C₆₋₁₀芳基、3-10元杂环基、5-10元杂芳基；或者两个孪位R^{sg}取代基可以连接形成=O或=S；其中X⁻是抗衡离子。

[0484] 在某些实施方案中，碳原子取代基包括：卤素、-CN、-NO₂、-N₃、-SO₂H、-SO₃H、-OH、-OC₁₋₆烷基、-ON(C₁₋₆烷基)₂、-N(C₁₋₆烷基)₂、-N(C₁₋₆烷基)₃⁺X⁻、-NH(C₁₋₆烷基)₂⁺X⁻、-NH₂(C₁₋₆烷基)⁺X⁻、-NH₃⁺X⁻、-N(OC₁₋₆烷基)(C₁₋₆烷基)、-N(OH)(C₁₋₆烷基)、-NH(OH)、-SH、-SC₁₋₆烷基、-SS(C₁₋₆烷基)、-C(=O)(C₁₋₆烷基)、-CO₂H、-CO₂(C₁₋₆烷基)、-OC(=O)(C₁₋₆烷基)、-OCO₂(C₁₋₆烷基)、-C(=O)NH₂、-C(=O)N(C₁₋₆烷基)₂、-OC(=O)NH(C₁₋₆烷基)、-NHC(=O)(C₁₋₆烷基)、-N(C₁₋₆烷基)C(=O)(C₁₋₆烷基)、-NHCO₂(C₁₋₆烷基)、-NHC(=O)N(C₁₋₆烷基)₂、-NHC(=O)NH(C₁₋₆烷基)、-NHC(=O)NH₂、-C(=NH)O(C₁₋₆烷基)、-OC(=NH)(C₁₋₆烷基)、-OC(=NH)OC₁₋₆烷基、-C(=NH)N(C₁₋₆烷基)₂、-C(=NH)NH(C₁₋₆烷基)、-C(=NH)NH₂、-OC(=NH)N(C₁₋₆烷基)₂、-OC(=NH)NH(C₁₋₆烷基)、-OC(=NH)NH₂、-NHC(=NH)N(C₁₋₆烷基)₂、-NHC(=NH)NH₂、-NHSO₂(C₁₋₆烷基)、-SO₂N(C₁₋₆烷基)₂、-SO₂NH(C₁₋₆烷基)、-SO₂NH₂、-SO₂(C₁₋₆烷基)、-SO₂O(C₁₋₆烷基)、-OSO₂(C₁₋₆烷基)、-SO(C₁₋₆烷基)、-Si(C₁₋₆烷基)₃、-OSi(C₁₋₆烷基)₃-C(=S)N(C₁₋₆烷基)₂、C(=S)NH(C₁₋₆烷基)、C(=S)NH₂、-C(=O)S(C₁₋₆烷基)、-C(=S)SC₁₋₆烷基、-SC(=S)SC₁₋₆烷基、-P(=O)(OC₁₋₆烷基)₂、-P(=O)(C₁₋₆烷基)₂、-OP(=O)(C₁₋₆烷基)₂、-OP(=O)(OC₁₋₆烷基)₂、C₁₋₆烷基、C₁₋₆全卤代烷基、C₂₋₆烯基、C₂₋₆炔基、杂C₁₋₆烷基、杂C₂₋₆烯基、杂C₂₋₆炔基、C₃₋₁₀碳环基、C₆₋₁₀芳基、3-10元杂环基、5-10元杂芳基；或者两个孪位R^{sg}取代基可以连接形成=O或=S；其中X⁻是抗衡离子。

[0485] 术语“卤代”或“卤素”是指氟(氟，-F)、氯(氯，-Cl)、溴(溴，-Br)或碘(碘，-I)。

[0486] 术语“羟基”是指基团-OH。延伸地，术语“取代的羟基”是指一种羟基基团，其中与母体分子直接连接的氧原子被氢以外的基团取代，并且包括选自以下的基团：-OR^{aa}、-ON(R^{bb})₂、-OC(=O)SR^{aa}、-OC(=O)R^{aa}、-OCO₂R^{aa}、-OC(=O)N(R^{bb})₂、-OC(=NR^{bb})R^{aa}、-OC(=NR^{bb})OR^{aa}、-OC(=NR^{bb})N(R^{bb})₂、-OS(=O)R^{aa}、-OSO₂R^{aa}、-OSi(R^{aa})₃、-OP(R^{cc})₂、-OP(R^{cc})₃⁺X⁻、-OP(OR^{cc})₂、-OP(OR^{cc})₃⁺X⁻、-OP(=O)(R^{aa})₂、-OP(=O)(OR^{cc})₂和-OP(=O)(N(R^{bb})₂)₂，其中X⁻、R^{aa}、R^{bb}和R^{cc}如本文所定义。

[0487] 术语“氨基”是指基团 $-\text{NH}_2$ 。扩展地,术语“取代的氨基”是指单取代的氨基、双取代的氨基或三取代的氨基。在某些实施方案中,“取代的氨基”是单取代的氨基或双取代的氨基。

[0488] 术语“单取代的氨基”是指一种氨基基团,其中与母体分子直接连接的氮原子被一个氢和一个除氢以外的基团取代,并且包括选自以下的基团: $-\text{NH}(\text{R}^{\text{bb}})$ 、 $-\text{NHC}(=\text{O})\text{R}^{\text{aa}}$ 、 $-\text{NHCO}_2\text{R}^{\text{aa}}$ 、 $-\text{NHC}(=\text{O})\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{NHC}(=\text{NR}^{\text{bb}})\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{NHSO}_2\text{R}^{\text{aa}}$ 、 $-\text{NHP}(=\text{O})(\text{OR}^{\text{cc}})_2$ 和 $-\text{NHP}(=\text{O})\text{N}(\text{R}^{\text{bb}})_2$,其中 R^{aa} 、 R^{bb} 和 R^{cc} 如本文所定义,并且其中基团 $-\text{NH}(\text{R}^{\text{bb}})$ 的 R^{bb} 不是氢。

[0489] 术语“双取代的氨基”是指一种氨基基团,其中与母体分子直接连接的氮原子被两个除氢以外的基团取代,并且包括选自以下的基团: $-\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{NR}^{\text{bb}}\text{C}(=\text{O})\text{R}^{\text{aa}}$ 、 $-\text{NR}^{\text{bb}}\text{CO}_2\text{R}^{\text{aa}}$ 、 $-\text{NR}^{\text{bb}}\text{C}(=\text{O})\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{NR}^{\text{bb}}\text{C}(=\text{NR}^{\text{bb}})\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{NR}^{\text{bb}}\text{SO}_2\text{R}^{\text{aa}}$ 、 $-\text{NR}^{\text{bb}}\text{P}(=\text{O})(\text{OR}^{\text{cc}})_2$ 和 $-\text{NR}^{\text{bb}}\text{P}(=\text{O})\text{N}(\text{R}^{\text{bb}})_2$,其中 R^{aa} 、 R^{bb} 和 R^{cc} 如本文所定义,条件是与母体分子直接连接的氮原子不被氢取代。

[0490] 术语“三取代的氨基”是指一种氨基基团,其中与母体分子直接连接的氮原子被三个基团取代,并且包括选自以下的基团: $-\text{N}(\text{R}^{\text{bb}})_3$ 和 $-\text{N}(\text{R}^{\text{bb}})_3^+\text{X}^-$,其中 R^{bb} 和 X^- 如本文所定义。

[0491] 术语“磺酰基”是指选自以下的基团: $-\text{SO}_2\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{SO}_2\text{R}^{\text{aa}}$ 和 $-\text{SO}_2\text{OR}^{\text{aa}}$,其中 R^{aa} 和 R^{bb} 如本文所定义。

[0492] 术语“亚磺酰基”是指基团 $-\text{S}(=\text{O})\text{R}^{\text{aa}}$,其中 R^{aa} 如本文所定义。

[0493] 术语“酰基”是指具有通式 $-\text{C}(=\text{O})\text{R}^{\text{X1}}$ 、 $-\text{C}(=\text{O})\text{OR}^{\text{X1}}$ 、 $-\text{C}(=\text{O})-\text{O}-\text{C}(=\text{O})\text{R}^{\text{X1}}$ 、 $-\text{C}(=\text{O})\text{SR}^{\text{X1}}$ 、 $-\text{C}(=\text{O})\text{N}(\text{R}^{\text{X1}})_2$ 、 $-\text{C}(=\text{S})\text{R}^{\text{X1}}$ 、 $-\text{C}(=\text{S})\text{N}(\text{R}^{\text{X1}})_2$ 、 $-\text{C}(=\text{S})\text{O}(\text{R}^{\text{X1}})$ 、 $-\text{C}(=\text{S})\text{S}(\text{R}^{\text{X1}})$ 、 $-\text{C}(=\text{NR}^{\text{X1}})\text{R}^{\text{X1}}$ 、 $-\text{C}(=\text{NR}^{\text{X1}})\text{OR}^{\text{X1}}$ 、 $-\text{C}(=\text{NR}^{\text{X1}})\text{SR}^{\text{X1}}$ 和 $-\text{C}(=\text{NR}^{\text{X1}})\text{N}(\text{R}^{\text{X1}})_2$,其中 R^{X1} 是氢;卤素;取代或未取代的羟基;取代或未取代的硫醇;取代或未取代的氨基;取代或未取代的酰基、环状或非环状的、取代或未取代的、支链或直链的脂族基;环状或非环状的、取代或未取代的、支链或直链的杂脂族基;环状或非环状的、取代或未取代的、支链或直链的烷基;环状或非环状的、取代或未取代的、支链或直链的烯基;取代或未取代的炔基;取代或未取代的芳基、取代或未取代的杂芳基、脂族氧基、杂脂族氧基、烷基氧基、杂烷基氧基、芳氧基、杂芳氧基、脂族硫基、杂脂族硫基、烷基硫基、杂烷基硫基、芳基硫基、杂芳基硫基、单或二脂族氨基、单或二杂脂族氨基、单-或二烷基氨基、单或二杂烷基氨基、单或二芳基氨基或单或二杂芳基氨基;或者两个 R^{X1} 基团一起形成5至6元杂环。示例性的酰基基团包括醛($-\text{CHO}$)、羧酸($-\text{CO}_2\text{H}$)、酮、酰基卤、酯、酰胺、亚胺、碳酸酯、氨基甲酸酯和脲。酰基取代基包括但不限于本文所述的导致形成稳定部分的任意取代基(例如脂族、烷基、烯基、炔基、杂脂族、杂环、芳基、杂芳基、酰基、氧代、亚氨基、硫代氧代(thiooxo)、氰基、异氰基、氨基、叠氮基、硝基、羟基、硫醇基、卤素、脂族氨基、杂脂族氨基、烷基氨基、杂烷基氨基、芳基氨基、杂芳基氨基、烷基芳基、芳基烷基、脂族氧基、杂脂族氧基、烷基氧基、杂烷基氧基、芳氧基、杂芳氧基、脂族硫氧基、杂脂族硫基、烷基硫基、杂烷基硫基、芳基硫基、杂芳基硫基、酰氧基等,它们各自可以被或可以不被进一步取代)。

[0494] 术语“羰基”是指一种基团,其中与母体分子直接连接的碳被 sp^2 杂化,并且被氧、氮或硫原子取代,例如,选自以下的基团:酮(例如 $-\text{C}(=\text{O})\text{R}^{\text{aa}}$)、羧酸(例如 $-\text{CO}_2\text{H}$)、醛($-\text{CHO}$)、酯(例如 $\text{CO}_2\text{R}^{\text{aa}}$ 、 $-\text{C}(=\text{O})\text{SR}^{\text{aa}}$ 、 $-\text{C}(=\text{S})\text{SR}^{\text{aa}}$)、酰胺(例如 $-\text{C}(=\text{O})\text{N}(\text{R}^{\text{bb}})_2$ 、 $-\text{C}(=\text{O})\text{NR}^{\text{bb}}\text{SO}_2\text{R}^{\text{aa}}$ 、 $-\text{C}(=\text{S})\text{N}(\text{R}^{\text{bb}})_2$)和亚胺(例如 $-\text{C}(=\text{NR}^{\text{bb}})\text{R}^{\text{aa}}$ 、 $-\text{C}(=\text{NR}^{\text{bb}})\text{OR}^{\text{aa}}$ 、 $-\text{C}(=\text{NR}^{\text{bb}})\text{N}$

(R^{bb})₂), 其中R^{aa}和R^{bb}如本文所定义。

[0495] 术语“甲硅烷基”是指基团-Si(R^{aa})₃, 其中R^{aa}如本文所定义。

[0496] 术语“氧代”是指基团=O, 术语“硫代氧”是指基团=S。

[0497] 在化合价允许的情况下, 氮原子可以被取代或未被取代, 并且包括伯、仲、叔和季氮原子。示例性的氮原子取代基包括但不限于氢、-OH、-OR^{aa}、-N(R^{cc})₂、-CN、-C(=O)R^{aa}、-C(=O)N(R^{cc})₂、-CO₂R^{aa}、-SO₂R^{aa}、-C(=NR^{bb})R^{aa}、-C(=NR^{cc})OR^{aa}、-C(=NR^{cc})N(R^{cc})₂、-SO₂N(R^{cc})₂、-SO₂R^{cc}、-SO₂OR^{cc}、-SOR^{aa}、-C(=S)N(R^{cc})₂、-C(=O)SR^{cc}、-C(=S)SR^{cc}、-P(=O)(OR^{cc})₂、-P(=O)(R^{aa})₂、-P(=O)(N(R^{cc})₂)₂、C₁₋₁₀烷基、C₁₋₁₀全卤代烷基、C₂₋₁₀烯基、C₂₋₁₀炔基、杂C₁₋₁₀烷基、杂C₂₋₁₀烯基、杂C₂₋₁₀炔基、C₃₋₁₀碳环基、3-14元杂环基、C₆₋₁₄芳基和5-14元杂芳基, 或者连接到N原子的两个R^{cc}基团连接形成3-14元杂环基或5-14元杂芳基环, 其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{dd}基团取代, 并且其中R^{aa}、R^{bb}、R^{cc}和R^{dd}如上文所定义。

[0498] 在某些实施方案中, 氮原子上存在的取代基是氮保护基团(在本文中也称为“氨基保护基团”)。氮保护基团包括但不限于-OH、-OR^{aa}、-N(R^{cc})₂、-C(=O)R^{aa}、-C(=O)N(R^{cc})₂、-CO₂R^{aa}、-SO₂R^{aa}、-C(=NR^{cc})R^{aa}、-C(=NR^{cc})OR^{aa}、-C(=NR^{cc})N(R^{cc})₂、-SO₂N(R^{cc})₂、-SO₂R^{cc}、-SO₂OR^{cc}、-SOR^{aa}、-C(=S)N(R^{cc})₂、-C(=O)SR^{cc}、-C(=S)SR^{cc}、C₁₋₁₀烷基(例如, 芳烷基、杂芳烷基)、C₂₋₁₀烯基、C₂₋₁₀炔基、杂C₁₋₁₀烷基、杂C₂₋₁₀烯基、杂C₂₋₁₀炔基、C₃₋₁₀碳环基、3-14元杂环基、C₆₋₁₄芳基和5-14元杂芳基基团, 其中每个烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、碳环基、杂环基、芳烷基、芳基和杂芳基独立地被0、1、2、3、4或5个R^{dd}基团取代, 并且其中R^{aa}、R^{bb}、R^{cc}和R^{dd}如本文所定义。氮保护基团是本领域公知的, 包括在Protecting Groups in Organic Synthesis, T.W.Greene and P.G.M.Wuts, 3rd edition, John Wiley&Sons, 1999(其通过引用并入本文)中详细描述的那些。

[0499] 例如, 氮保护基团例如酰胺基团(例如, -C(=O)R^{aa})包括但不限于甲酰胺、乙酰胺、氯乙酰胺、三氯乙酰胺、三氟乙酰胺、苯基乙酰胺、3-苯基丙酰胺、吡啶甲酰胺、3-吡啶基羧酰胺、N-苯甲酰基苯丙氨酰基衍生物、苯甲酰胺、对苯基苯甲酰胺、邻硝基苯乙酰胺、邻硝基苯氧基乙酰胺、乙酰乙酰胺、(N'-二硫代苄氧基酰氨基)乙酰胺、3-(对羟基苯基)丙酰胺、3-(邻硝基苯基)丙酰胺、2-甲基-2-(邻硝基苯氧基)丙酰胺、2-甲基-2-(邻苯基偶氮苯氧基)丙酰胺、4-氯丁酰胺、3-甲基-3-硝基丁酰胺、邻硝基肉桂酰胺、N-乙酰甲硫氨酸衍生物、邻硝基苯甲酰胺和邻-(苯甲酰氧基甲基)苯甲酰胺。

[0500] 氮保护基团例如氨基甲酸酯基团(例如, -C(=O)OR^{aa})包括但不限于氨基甲酸甲酯、氨基甲酸乙酯、9-苄基甲基氨基甲酸酯(Fmoc)、9-(2-磺基)苄基甲基氨基甲酸酯、9-(2,7-二溴)苄基甲基氨基甲酸酯、2,7-二叔丁基-[9-(10,10-二氧代-10,10,10,10-四氢噻吨基)]甲基氨基甲酸酯(DBD-Tmoc)、4-甲氧基苯甲酰基氨基甲酸酯(Phenoc)、2,2,2-三氯乙基氨基甲酸酯(Troc)、2-三甲基甲硅烷基乙基氨基甲酸酯(Teoc)、2-苯基乙基氨基甲酸酯(hZ)、1-(1-金刚烷基)-1-甲基乙基氨基甲酸酯(Adpoc)、1,1-二甲基-2-卤代乙基氨基甲酸酯、1,1-二甲基-2,2-二溴乙基氨基甲酸酯(DB-t-BOC)、1,1-二甲基-2,2,2-三氯乙基氨基甲酸酯(TCBOC)、1-甲基-1-(4-联苯基)乙基氨基甲酸酯(Bpoc)、1-(3,5-二叔丁基苯基)-1-甲基乙基氨基甲酸酯(t-Bumeoc)、2-(2'-和4'-吡啶基)乙基氨基甲酸酯(Pyoc)、2-(N,N-二环己基羧酰胺基)乙基氨基甲酸酯、氨基甲酸叔丁酯(BOC或Boc)、1-金刚烷基氨基甲酸酯

(Adoc)、乙烯基氨基甲酸酯(Voc)、烯丙基氨基甲酸酯(Alloc)、1-异丙烯基丙基氨基甲酸酯(Ipaoc)、肉桂基氨基甲酸酯(Coc)、4-硝基肉桂基氨基甲酸酯(Noc)、8-喹啉基氨基甲酸酯、N-羟基哌啶基氨基甲酸酯、烷基二硫代氨基甲酸酯、苄基氨基甲酸酯(Cbz)、对甲氧基苄基氨基甲酸酯(Moz)、对硝基苄基氨基甲酸酯、对溴苄基氨基甲酸酯、对氯苄基氨基甲酸酯、2,4-二氯苄基氨基甲酸酯、4-甲基亚磺酰基苄基氨基甲酸酯(Msz)、9-蒎基甲基氨基甲酸酯、二苯基甲基氨基甲酸酯、2-甲基硫乙基氨基甲酸酯、2-甲基磺酰基乙基氨基甲酸酯、2-(对甲苯磺酰基)乙基氨基甲酸酯、[2-(1,3-二硫杂环己烷基)]甲基氨基甲酸酯(Dmoc)、4-甲基硫代苯基氨基甲酸酯(Mtpc)、2,4-二甲基硫代苯基氨基甲酸酯(Bmpc)、2-膦酰乙基氨基甲酸酯(2-phosphonioethyl carbamate, Peoc)、2-三苯基膦酰异丙基氨基甲酸酯(2-triphenylphosphonioisopropyl carbamate, Ppoc)、1,1-二甲基-2-氰基乙基氨基甲酸酯、间氯-对-酰氧基苄基氨基甲酸酯、对-(二羟基硼基)苄基氨基甲酸酯、5-苯并异噁唑基甲基氨基甲酸酯、2-(三氟甲基)-6-色酮基甲基氨基甲酸酯(2-(trifluoromethyl)-6-chromonylmethyl carbamate, Tcroc)、间硝基苯基氨基甲酸酯、3,5-二甲氧基苄基氨基甲酸酯、邻硝基苄基氨基甲酸酯、3,4-二甲氧基-6-硝基苄基氨基甲酸酯、苯基(邻硝基苯基)甲基氨基甲酸酯、叔戊基氨基甲酸酯、S-苄基硫代氨基甲酸酯、对氰基苄基氨基甲酸酯、氨基甲酸环丁酯、氨基甲酸环己酯、氨基甲酸环戊酯、环丙基甲基氨基甲酸酯、对癸氧基苄基氨基甲酸酯、2,2-二甲氧基酰基乙烯基氨基甲酸酯、邻-(N,N-二甲基氨基甲酰氨基)苄基氨基甲酸酯、1,1-二甲基-3-(N,N-二甲基氨基甲酰氨基)丙基氨基甲酸酯、1,1-二甲基丙炔基氨基甲酸酯、二(2-吡啶基)甲基氨基甲酸酯、2-咪喃基甲基氨基甲酸酯、2-碘乙基氨基甲酸酯、异硼炔基氨基甲酸酯、异丁基氨基甲酸酯、异烟碱基氨基甲酸酯、对-(p'-甲氧基苯基偶氮)苄基氨基甲酸酯、1-甲基环丁基氨基甲酸酯、1-甲基环己基氨基甲酸酯、1-甲基-1-环丙基甲基氨基甲酸酯、1-甲基-1-(3,5-二甲氧基苯基)乙基氨基甲酸酯、1-甲基-1-(对苯基偶氮苯基)乙基氨基甲酸酯、1-甲基-1-苯基乙基氨基甲酸酯、1-甲基-1-(4-吡啶基)乙基氨基甲酸酯、氨基甲酸苯酯、对(苯基偶氮)苄基氨基甲酸酯、2,4,6-三叔丁基苯基氨基甲酸酯、4-(三甲基铵)苄基氨基甲酸酯和2,4,6-三甲基苄基氨基甲酸酯。

[0501] 氮保护基团例如磺酰胺基团(例如, $-S(=O)_2R^{aa}$)包括但不限于对甲苯磺酰胺(Ts)、苯磺酰胺、2,3,6-三甲基-4-甲氧基苯磺酰胺(Mtr)、2,4,6-三甲氧基苯磺酰胺(Mtb)、2,6-二甲基-4-甲氧基苯磺酰胺(Pme)、2,3,5,6-四甲基-4-甲氧基苯磺酰胺(Mte)、4-甲氧基苯磺酰胺(Mbs)、2,4,6-三甲基苯磺酰胺(Mts)、2,6-二甲氧基-4-甲基苯磺酰胺(iMds)、2,2,5,7,8-五甲基色满-6-磺酰胺(Pmc)、甲磺酰胺(Ms)、 β -三甲基甲硅烷基乙磺酰胺(SES)、9-蒎磺酰胺、4-(4',8'-二甲氧基萘甲基)苯磺酰胺(DNMBS)、苄基磺酰胺、三氟甲基磺酰胺和苯甲酰磺酰胺。

[0502] 其他氮保护基团包括但不限于吩噻嗪基-(10)-酰基衍生物、N'-对甲苯磺酰基氨基酰基衍生物、N'-苯基氨基硫酰基衍生物、N-苯甲酰基苯基丙氨酰基衍生物、N-乙酰甲硫氨酸衍生物、4,5-二苯基-3-噁唑啉-2-酮、N-邻苯二甲酰亚胺、N-二硫代琥珀酰亚胺(Dts)、N-2,3-二苯基马来酰亚胺、N-2,5-二甲基吡咯、N-1,1,4,4-四甲基二甲硅烷基氮杂环戊烷加合物(STABASE)、5-取代的1,3-二甲基-1,3,5-三氮杂环己-2-酮、5-取代的1,3-二苄基-1,3,5-三氮杂环己-2-酮、1-取代的3,5-二硝基-4-吡啶酮、N-甲胺、N-烯丙胺、N-[2-(三甲基甲硅烷基)乙氧基]甲胺(SEM)、N-3-乙酰氧基丙胺、N-(1-异丙基-4-硝基-2-氧代-3-吡咯

啉-3-基)胺、季铵盐、N-苄基胺、N-二(4-甲氧基苯基)甲胺、N-5-二苯并环庚胺(N-5-dibenzosuberylamine)、N-三苯基甲胺(Tr)、N-[(4-甲氧基苯基)二苯基甲基]胺(MMTr)、N-9-苯基苄基胺(PhF)、N-2,7-二氯-9-苄基亚甲基胺、N-二茂铁基甲基氨基(Fcm)、N-2-吡啶甲基氨基N'-氧化物(N-2-picolylamino N'-oxide)、N-1,1-二甲基硫代亚甲基胺、N-亚苄基胺、N-对甲氧基亚苄基胺、N-二苯基亚甲基胺、N-[(2-吡啶基)间甲苯基]亚甲基胺、N-(N',N'-二甲基氨基亚甲基)胺、N,N'-异丙二烯二胺、N-对硝基亚苄基胺、N-亚水杨基胺(N-salicylideneamine)、N-5-氯亚水杨基胺、N-(5-氯-2-羟苯基)苯基亚甲基胺、N-亚环己基胺、N-(5,5-二甲基-3-氧代-1-环己烯基)胺、N-硼烷衍生物、N-二苯基硼酸衍生物、N-[苯基(五酰基铬或钨)酰基]胺、N-铜螯合物、N-锌螯合物、N-硝基胺、N-亚硝胺、N-胺氧化物、二苯基膦酰胺(Dpp)、二甲基硫代膦酰胺(Mpt)、二苯基硫代膦酰胺(Ppt)、二烷基膦酰胺、二苄基膦酰胺、二苯基膦酰胺、苯亚磺酰胺(benzenesulfenamide)、邻硝基苯亚磺酰胺(Nps)、2,4-二硝基苯亚磺酰胺、五氯苯亚磺酰胺、2-硝基-4-甲氧基苯亚磺酰胺、三苯基甲基亚磺酰胺和3-硝基吡啶亚磺酰胺(Npys)。在某些实施方案中,氮保护基团为苄基(Bn)、叔丁氧基羰基(BOC)、碳苄基氧基(Cbz)、9-苄基甲氧基羰基(Fmoc)、三氟乙酰基、三苯基甲基、乙酰基(Ac)、苯甲酰基(Bz)、对甲氧基苄基(PMB)、3,4-二甲氧基苄基(DMPM)、对甲氧基苯基(PMP)、2,2,2-三氯乙氧基(Troc)、三苯甲基(Tr)、甲苯磺酰基(Ts)、溴苯磺酰基(Bs)、硝苯磺酰基(nosyl,Ns)、甲磺酰基(Ms)、三氟甲磺酰基(triflyl,Tf)或丹磺酰基(dansyl,Ds)。

[0503] 在某些实施方案中,氧原子上存在的取代基是氧保护基团(在本文中也称为“羟基保护基团”)。氧保护基团包括但不限于 $-R^{aa}$ 、 $-N(R^{bb})_2$ 、 $-C(=O)SR^{aa}$ 、 $-C(=O)R^{aa}$ 、 $-CO_2R^{aa}$ 、 $-C(=O)N(R^{bb})_2$ 、 $-C(=NR^{bb})R^{aa}$ 、 $-C(=NR^{bb})OR^{aa}$ 、 $-C(=NR^{bb})N(R^{bb})_2$ 、 $-S(=O)R^{aa}$ 、 $-SO_2R^{aa}$ 、 $-Si(R^{aa})_3$ 、 $-P(R^{cc})_2$ 、 $-P(R^{cc})_3^+X^-$ 、 $-P(OR^{cc})_2$ 、 $-P(OR^{cc})_3^+X^-$ 、 $-P(=O)(R^{aa})_2$ 、 $-P(=O)(OR^{cc})_2$ 和 $-P(=O)(N(R^{bb})_2)_2$,其中 X^- 、 R^{aa} 、 R^{bb} 和 R^{cc} 如本文所定义。氧保护基团是本领域公知的,包括在Protecting Groups in Organic Synthesis,T.W.Greene and P.G.M.Wuts,3rd edition, John Wiley&Sons,1999(通过引用并入本文)中详细描述的那些。

[0504] 示例性氧保护基团包括但不限于甲基、甲氧基甲基(MOM)、甲基硫代甲基(MTM)、叔丁基硫代甲基、(苯基二甲基甲硅烷基)甲氧基甲基(SMOM)、苄氧基甲基(BOM)、对甲氧基苄氧基甲基(PMBM)、(4-甲氧基苯氧基)甲基(p-AOM)、邻甲氧基苯酚甲基(guaiacolmethyl,GUM)、叔丁氧基甲基、4-戊烯氧基甲基(POM)、甲硅烷氧基甲基、2-甲氧基乙氧基甲基(MEM)、2,2,2-三氯乙氧基甲基、双(2-氯乙氧基)甲基、2-(三甲基甲硅烷基)乙氧基甲基(SEMOR)、四氢吡喃基(THP)、3-溴四氢吡喃基、四氢硫吡喃基、1-甲氧基环己基、4-甲氧基四氢吡喃基(MTHP)、4-甲氧基四氢硫吡喃基、4-甲氧基四氢硫吡喃基S,S-二氧化物、1-[(2-氯-4-甲基)苯基]-4-甲氧基哌啶-4-基(CTMP)、1,4-二噁烷-2-基、四氢呋喃基、四氢硫代呋喃基、2,3,3a,4,5,6,7,7a-八氢-7,8,8-三甲基-4,7-甲基苯并呋喃-2-基、1-乙氧基乙基、1-(2-氯乙氧基)乙基、1-甲基-1-甲氧基乙基、1-甲基-1-苄氧基乙基、1-甲基-1-苄氧基-2-氟乙基、2,2,2-三氯乙基、2-三甲基甲硅烷基乙基、2-(苯基硒基)乙基、叔丁基、烯丙基、对氯苯基、对甲氧基苯基、2,4-二硝基苯基、苄基(Bn)、对甲氧基苄基、3,4-二甲氧基苄基、邻硝基苄基、对硝基苄基、对卤苄基、2,6-二氯苄基、对氰基苄基、对苯苄基、2-吡啶甲基、4-吡啶甲基、3-甲基-2-吡啶甲基N-氧化物、二苯基甲基、p,p'-二硝基二苯甲基(p,p'-dinitrobenzhydryl)、5-二苯并环庚基(5-dibenzosuberyl)、三苯基甲基、 α -萘基二苯基甲

基、对甲氧基苄基二苄基甲基、二(对甲氧基苄基)苄基甲基、三(对甲氧基苄基)甲基、4-(4'-溴苄甲酰氧基苄基)二苄基甲基、4,4',4''-三(4,5-二氯邻苯二甲酰亚胺基苄基)甲基、4,4',4''-三(乙酰丙酰氧基苄基)甲基(4,4',4''-tris(levulinoyloxyphenyl)methyl)、4,4',4''-三(苄甲酰氧基苄基)甲基、3-(咪唑-1-基)双(4',4''-二甲氧基苄基)甲基、1,1-双(4-甲氧基苄基)-1'-苄基甲基、9-蒎基、9-(9-苄基)氧杂蒎基、9-(9-苄基-10-氧代)蒎基、1,3-苯并二硫杂-2-基(1,3-benzodithiolan-2-yl)、苯并异噻唑基S,S-二氧化物、三甲基甲硅烷基(TMS)、三乙基甲硅烷基(TEs)、三异丙基甲硅烷基(TIPS)、二甲基异丙基甲硅烷基(IPDMS)、二乙基异丙基甲硅烷基(DEIPS)、二甲基叔己基甲硅烷基(dimethylhexylsilyl)、叔丁基二甲基甲硅烷基(TBDMS)、叔丁基二苄基甲硅烷基(TBDPS)、三苄基甲硅烷基、三对二甲苯基甲硅烷基、三苄基甲硅烷基、二苄基甲基甲硅烷基(DPMS)、叔丁基甲氧基苄基甲硅烷基(TBMPS)、甲酸酯、苄甲酰甲酸酯、乙酸酯、氯乙酸酯、二氯乙酸酯、三氯乙酸酯、三氟乙酸酯、甲氧基乙酸酯、三苄基甲氧基乙酸酯、苄氧乙酸酯、对氯苄氧基乙酸酯、3-苄基丙酸酯、4-氧戊酸(乙酰丙酸酯)、4,4-(亚乙基二硫代)戊酸(乙酰丙酰基二硫代乙醛)、新戊酸酯、金刚烷酸酯、巴豆酸酯(crotonate)、4-甲氧基巴豆酸酯、苄甲酸酯、对苄基苄甲酸酯、2,4,6-三甲基苄甲酸酯、碳酸甲酯、9-苄基甲基碳酸酯(Fmoc)、碳酸乙酯、2,2,2-三氯乙基碳酸酯(Troc)、2-(三甲基甲硅烷基)乙酯碳酸酯(TMSEC)、2-(苄磺酰基)乙基碳酸酯(Psec)、2-(三苄基磷基)乙基碳酸酯(Peoc)、碳酸异丁酯、碳酸乙烯基酯、碳酸烯丙酯、碳酸叔丁酯(BOC或Boc)、对硝基苄基碳酸酯、碳酸苄基酯、对甲氧基苄基碳酸酯、3,4-二甲氧基苄基碳酸酯、邻硝基苄基碳酸酯、对硝基苄基碳酸酯、S-苄基硫代碳酸酯、4-乙氧基-1-萘基碳酸酯、甲基二硫代碳酸酯、2-碘代苄甲酸酯、4-叠氮基丁酸酯、4-硝基-4-甲基戊酸酯、邻(二溴甲基)苄甲酸酯、2-甲酰基苄磺酸酯、2-(甲基硫代甲氧基)乙基、4-(甲基硫代甲氧基)丁酸酯、2-(甲基硫代甲氧基甲基)苄甲酸酯、2,6-二氯-4-甲基苄氧基乙酸酯、2,6-二氯-4-(1,1,3,3-四甲基丁基)苄氧基乙酸酯、2,4-双(1,1-二甲基丙基)苄氧基乙酸酯、氯二苄基乙酸酯、异丁酸酯、单琥珀酸酯、(E)-2-甲基-2-丁烯酸酯、邻(甲氧基酰基)苄甲酸酯、 α -萘甲酸酯、硝酸酯、烷基N,N,N',N'-四甲基磷酸二氨基酯、烷基N-苄基氨基甲酸酯、硼酸酯、二甲基磷硫酰基(dimethylphosphinothioyl)、烷基2,4-二硝基苄基亚硫酸酯(alkyl2,4-dinitrophenylsulfenate)、硫酸酯、甲磺酸酯、苄基磺酸酯和甲苯磺酸酯(Ts)。在某些实施方案中,氧保护基团是甲硅烷基。在某些实施方案中,氧保护基团是叔丁基二苄基甲硅烷基(TBDPS)、叔丁基二甲基甲硅烷基(TBDMS)、三异丙基甲硅烷基(TIPS)、三苄基甲硅烷基(TPS)、三乙基甲硅烷基(TEs)、三甲基甲硅烷基(TMS)、三异丙基甲硅烷氧基甲基(TOM)、乙酰基(Ac)、苄甲酰基(Bz)、碳酸烯丙酯、2,2,2-三氯乙基碳酸酯(Troc)、2-三甲基甲硅烷基乙基碳酸酯、甲氧基甲基(MOM)、1-乙氧基乙基(EE)、2-甲氧基-2-丙基(MOP)、2,2,2-三氯乙氧基乙基、2-甲氧基乙氧基甲基(MEM)、2-三甲基甲硅烷基乙氧基甲基(SEM)、甲基硫代甲基(MTM)、四氢吡喃基(THP)、四氢呋喃基(THF)、对甲氧基苄基(PMP)、三苄甲基(Tr)、甲氧基三苄甲基(MMT)、二甲氧基三苄甲基(DMT)、烯丙基、对甲氧基苄基(PMB)、叔丁基、苄基(Bn)、烯丙基或新戊酰基(Piv)。

[0505] 在某些实施方案中,硫原子上存在的取代基为硫保护基团(也称为“硫醇保护基团”)。硫保护基团包括但不限于: $-R^{aa}$ 、 $-N(R^{bb})_2$ 、 $-C(=O)SR^{aa}$ 、 $-C(=O)R^{aa}$ 、 $-CO_2R^{aa}$ 、 $-C(=O)N(R^{bb})_2$ 、 $-C(=NR^{bb})R^{aa}$ 、 $-C(=NR^{bb})OR^{aa}$ 、 $-C(=NR^{bb})N(R^{bb})_2$ 、 $-S(=O)R^{aa}$ 、 $-SO_2R^{aa}$ 、 $-Si(R^{aa})_3$ 、 $-P$

$(R^{cc})_2$ 、 $-P(R^{cc})_3^+X^-$ 、 $-P(OR^{cc})_2$ 、 $-P(OR^{cc})_3^+X^-$ 、 $-P(=O)(R^{aa})_2$ 、 $-P(=O)(OR^{cc})_2$ 和 $-P(=O)(N(R^{bb})_2)_2$,其中 R^{aa} 、 R^{bb} 和 R^{cc} 如本文所定义。硫保护基团是本领域公知的,包括在Protecting Groups in Organic Synthesis,T.W.Greene and P.G.M.Wuts,3rd edition,John Wiley& Sons,1999(通过引用并入本文)中详细描述的那些。在某些实施方案中,硫保护基团是乙酰氨基甲基、叔丁基、3-硝基-2-吡啶亚磺酰基、2-吡啶-亚磺酰基或三苯基甲基。

[0506] “抗衡离子”或“阴离子抗衡离子”是与带正电的基团相缔合以维持电子中性的带负电的基团。阴离子抗衡离子可以是单价的(即,包括一个形式上的负电荷)。阴离子抗衡离子也可以是多价的(即,包括一个以上的形式上的负电荷),例如二价或三价。示例性的抗衡离子包括卤离子(例如, F^- 、 Cl^- 、 Br^- 、 I^-)、 NO_3^- 、 ClO_4^- 、 OH^- 、 $H_2PO_4^-$ 、 HCO_3^- 、 HSO_4^- 、磺酸根离子(例如,甲磺酸根、三氟甲磺酸根、对甲苯磺酸根、苯磺酸根、10-樟脑磺酸根、萘-2-磺酸根、萘-1-磺酸-5-磺酸根、乙烷-1-磺酸-2-磺酸根等)、羧酸根离子(例如乙酸根、丙酸根、苯甲酸根、甘油酸根、乳酸根、酒石酸根、乙醇酸根、葡萄糖酸根等)、 BF_4^- 、 PF_4^- 、 PF_6^- 、 AsF_6^- 、 SbF_6^- 、 $B[3,5-(CF_3)_2C_6H_3]_4^-$ 、 $B(C_6F_5)_4^-$ 、 BPh_4^- 、 $Al(OC(CF_3)_3)_4^-$ 和碳硼烷阴离子(例如, $CB_{11}H_{12}^-$ 或 $(HCB_{11}Me_5Br_6)^-$)。可以是多价的示例性抗衡离子包括 CO_3^{2-} 、 HPO_4^{2-} 、 PO_4^{3-} 、 $B_4O_7^{2-}$ 、 SO_4^{2-} 、 $S_2O_3^{2-}$ 、羧酸根阴离子(例如,酒石酸根、柠檬酸根、富马酸根、马来酸根、苹果酸根、丙二酸根、葡萄糖酸根、琥珀酸根、戊二酸根、己二酸根、庚二酸酯、辛二酸酯、壬二酸酯、癸二酸酯、水杨酸酯、邻苯二甲酸酯、天冬氨酸、谷氨酸等)和碳硼烷。

[0507] 术语“离去基团”在合成有机化学领域中具有其通常的含义,是指能够被亲核试剂取代的原子或基团。参见,例如,Smith, March Advanced Organic Chemistry 6th ed. (501-502)。合适的离去基团的实例包括但不限于卤素(例如,F、Cl、Br或I(碘))、烷氧基羰氧基、芳氧基羰氧基、烷磺酰基氧基、芳磺酰基氧基、烷基-羰氧基(例如,乙酰氧基)、芳基羰氧基、芳氧基、甲氧基、N,0-二甲基羟氨基、pixyl和卤甲酸酯。在某些情况下,离去基团是磺酸酯例如甲苯磺酸酯(-OTs)、甲磺酸酯(-OMs)、对溴苯磺酰氧基(溴甲酸酯,-OBs)、 $-OS(=O)_2(CF_2)_3CF_3$ (壬酸酯,-ONf)或三氟甲磺酸酯(-OTf)。在某些情况下,离去基团是溴甲酸酯,例如对溴苯磺酰氧基。在某些情况下,离去基团是一种壬酸酯,例如2-硝基苯磺酰氧基。离去基团还可以是磷氧化物(例如,在Mitsunobu反应过程中形成的)或一个内部离去基团,例如环氧化物或环状硫酸盐。离去基团的其他非限制性实例是水、氨、醇、醚部分、硫醚部分、卤化锌、镁部分、重氮盐和铜部分。其他示例性的离去基团包括但不限于卤素(例如,氯、溴、碘)和活化的取代的羟基基团(例如, $-OC(=O)SR^{aa}$ 、 $-OC(=O)R^{aa}$ 、 $-OC(O)_2R^{aa}$ 、 $-OC(=O)N(R^{bb})_2$ 、 $-OC(=O)NR^{bb}R^{aa}$ 、 $-OC(=O)NR^{bb}OR^{aa}$ 、 $-OC(=O)NR^{bb}N(R^{bb})_2$ 、 $-OS(=O)R^{aa}$ 、 $-OSO_2R^{aa}$ 、 $-OP(R^{cc})_2$ 、 $-OP(R^{cc})_3$ 、 $-OP(=O)_2R^{aa}$ 、 $-OP(=O)(R^{aa})_2$ 、 $-OP(=O)(OR^{cc})_2$ 、 $-OP(=O)_2N(R^{bb})_2$ 和 $-OP(=O)(NR^{bb})_2$,其中 R^{aa} 、 R^{bb} 和 R^{cc} 如本文所定义。

[0508] 如本文所使用的,短语“至少一个实例”的使用是指1、2、3、4或更多个实例,但是还涵盖例如1至4、1至3、1至2、2至4、2至3或3至4个实例的范围,包括端值。

[0509] “非氢基团”是指为特定变量定义的任何非氢基团。

[0510] 以下定义是在整个本申请中使用的更通用的术语。

[0511] 如本文所用,术语“盐”是指任何和所有盐,包括药学上可接受的盐。术语“药学上可接受的盐”是指在合理的医学判断范围内适合与人和低等动物的组织接触而没有过度的毒性、刺激性、过敏反应等的盐,并且它们与合理的效益/风险比相称。药学上可接受的盐是

本领域众所周知的。例如, Berge et al. describe pharmaceutically acceptable salts in detail in J. Pharmaceutical Sciences, 1977, 66, 1-19 中详细描述了药学上可接受的盐, 该文献通过引用并入本文。本发明化合物的药学上可接受的盐包括衍生自合适的无机和有机酸和碱的那些。药学上可接受的无毒酸加成盐的实例是氨基基团与无机酸, 例如盐酸、氢溴酸、磷酸、硫酸和高氯酸, 或者与有机酸, 例如乙酸、草酸、马来酸、酒石酸、柠檬酸、琥珀酸或丙二酸, 或者使用本技术中已知的其它方法, 例如离子交换而形成的盐。其他药学上可接受的盐包括己二酸盐、藻酸盐、抗坏血酸盐、天冬氨酸盐、苯磺酸盐、苯甲酸盐、硫酸氢盐、硼酸盐、丁酸盐、樟脑酸盐、樟脑磺酸盐、柠檬酸盐、环戊丙酸盐、二葡萄糖酸盐、十二烷基硫酸盐、乙磺酸盐、甲酸盐、富马酸盐、葡庚糖酸盐、甘油磷酸盐、葡萄糖酸盐、半硫酸盐、庚酸盐、己酸盐、氢碘酸盐、2-羟基-乙磺酸盐、乳糖酸盐、乳酸盐、月桂酸盐、月桂基硫酸盐、苹果酸盐、马来酸盐、丙二酸盐、甲磺酸盐、2-萘磺酸盐、烟酸盐、硝酸盐、油酸盐、草酸盐、棕榈酸盐、双羟萘酸盐 (pamoate)、果胶酸盐、过硫酸盐、3-苯基丙酸盐、磷酸盐、苦味酸盐、新戊酸盐、丙酸盐、硬脂酸盐、琥珀酸盐、硫酸盐、酒石酸盐、硫氰酸盐、对甲苯磺酸盐、十一烷酸盐、戊酸盐等。衍生自适当碱的盐包括碱金属、碱土金属、铵和 $N^+(C_{1-4}\text{烷基})_4$ 盐。代表性的碱金属或碱土金属盐包括钠、锂、钾、钙、镁等。另外的药学上可接受的盐在适当时包括无毒的铵、季铵和使用抗衡离子例如卤离子、氢氧根、羧酸根、硫酸根、磷酸根、硝酸根、低级烷基磺酸根和芳基磺酸根形成的胺阳离子。

[0512] 还应理解, 具有相同分子式但其原子键合的性质或顺序或其原子在空间中的排列不同的化合物称为“异构体”。其原子在空间中的排列不同的异构体被称为“立体异构体”。

[0513] 彼此不是镜像的立体异构体被称为“非对映异构体”, 而彼此是不可重叠的镜像的立体异构体被称为“对映异构体”。当化合物具有不对称中心时, 例如, 它与四个不同的基团键合, 可能有一对对映异构体。对映异构体的特征在于其不对称中心的绝对构型, 并由 Cahn 和 Prelog 的 R- 和 S- 排序规则描述, 或由分子旋转偏振光平面的方式描述, 并称为右旋或左旋 (即, 分别为 (+) 或 (-) - 异构体)。手性化合物可以以单独的对映异构体或以其混合物存在。含有等比例对映异构体的混合物被称为“外消旋混合物”。

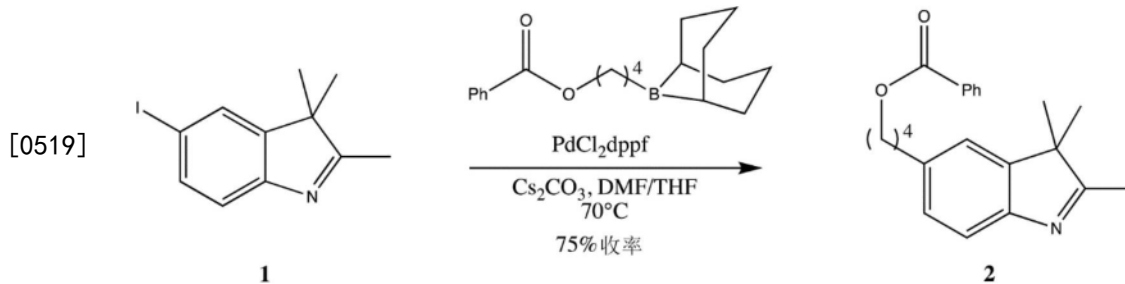
[0514] 术语“催化”是指由于称为“催化剂”的物质的参与而使化学反应的速率增加。在某些实施方案中, 催化剂的量和性质在反应过程中基本保持不变。在某些实施方案中, 催化剂在反应后再生, 或使催化剂的性质基本恢复。一种催化剂可以参与多种化学转化。催化剂的效果可能会由于称为抑制剂或毒药 (降低催化活性) 或促进剂 (增加活性) 的其它物质的存在而变化。催化反应比相应的未催化反应具有较低的活化能 (活化的限速自由能), 从而在相同温度下具有较高的反应速率。催化剂可能有利地影响反应环境, 与试剂结合以极化键, 形成通常不由未催化反应产生的特定中间体, 或导致试剂解离成反应性形式。

[0515] 术语“溶剂”是指能溶解一种或多种溶质, 从而产生溶液的物质。溶剂可作为本文所述的任何反应或转化的介质。溶剂可将一种或多种反应物或试剂溶解在反应混合物中。溶剂可促进反应混合物中一种或多种试剂或反应物的混合。溶剂还可以起到相对于在不同溶剂中的反应增加或降低反应速率的作用。溶剂可以是极性的或非极性的、亲性的或无性的。本文描述了在方法中有用的常见有机溶剂。

[0516] 实施例

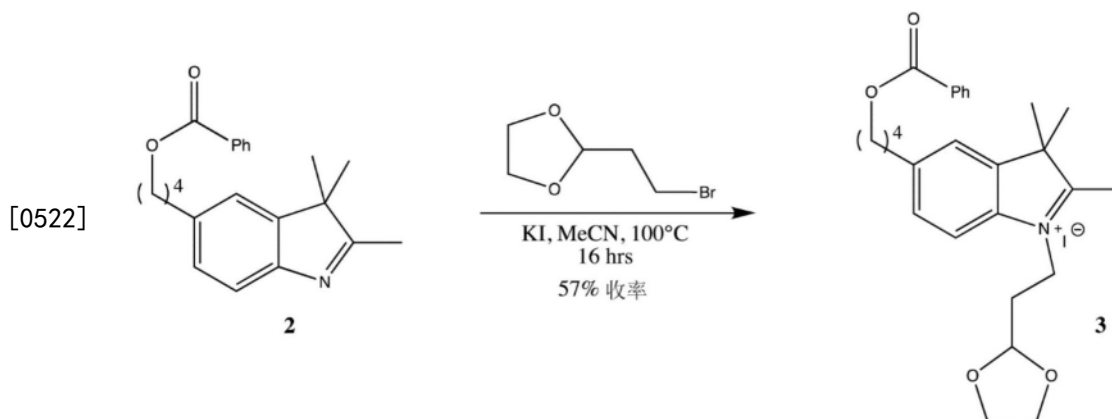
[0517] 双缀合标记物的合成

[0518] 假吲哚苯甲酸酯2 (Indolenine benzoate 2)



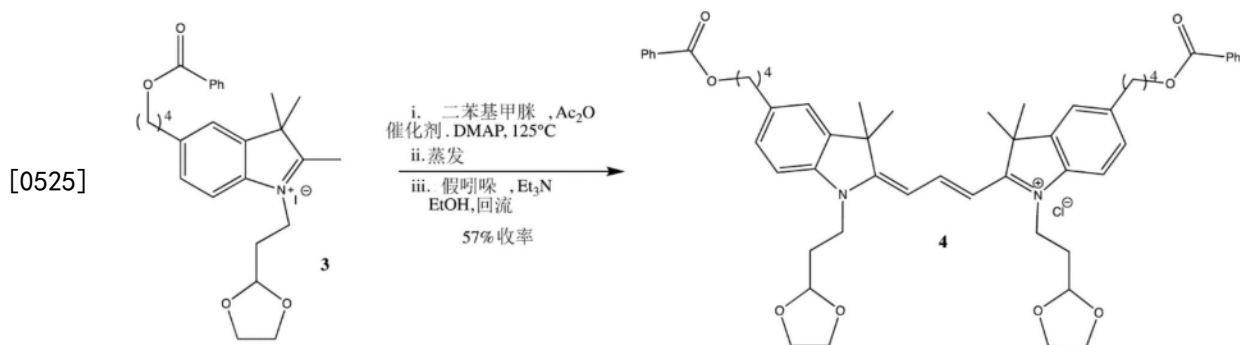
[0520] 向含有3-丁-1-醇苯甲酸酯 (3.0g, 17.1mmol) 的烧瓶中加入0.5M 9-硼双环[3.3.1]壬烷 (34.1mL, 17.1mmol) 的THF溶液。将澄清的无色溶液在室温搅拌3小时。在装有冷凝器的另一个烧瓶中, 装入碘假吲哚 (1, 3.4g, 11.9mmol)、[1,1'-双(二苯基膦基)二茂铁]二氯化钯(II) (0.7g, 0.8mmol) 和氯化铯 (6.0g, 18.5mmol)。向烧瓶中加入DMF (20mL), 并将深色悬浮液用氩气鼓泡10分钟。加入苯甲酸酯硼烷/THF溶液后, 将反应加热至70°C持续12小时。此时间之后, HPLC表明起始的假吲哚1已完全转化。将反应冷却至室温, 用EtOAc (50mL) 和己烷 (50mL) 稀释, 并通过硅藻土过滤。将有机层用水洗涤 (3x), 用硫酸镁干燥。过滤并蒸发后, 将粗残余物通过正相色谱法纯化 (0→50% EtOAc/己烷, SiO₂), 得到假吲哚2 (3.0g, 产率为75%), 为粘性黄色油。C₂₂H₂₆NO₂ (M+H)⁺的HRMS (ESI) 计算值为336.1964, 实测值为336.1962。

[0521] 烷基化的假吲哚3



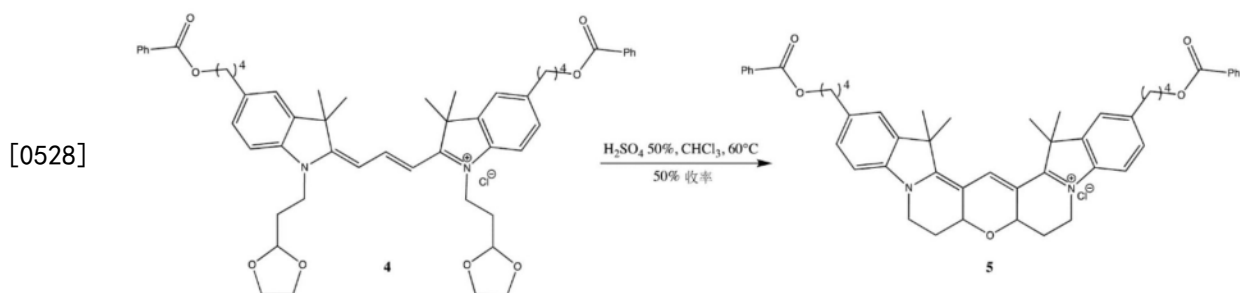
[0523] 向schlenk烧瓶中加入碘化钾 (2.0g, 12.0mmol), 并用氩气冲洗。加入假吲哚 (2, 2.0g, 6.0mmol), 为乙腈 (8.0mL) 溶液。加入2-(2-溴乙基)-1,3-二氧戊环 (1.4mL, 12.0mmol), 并将密封的容器加热至100°C保持14小时。将反应物用二氯甲烷 (20mL) 稀释, 并将悬浮液通过烧结的玻璃漏斗过滤。通过正相色谱法 (0→7% MeOH/DCM, SiO₂) 直接纯化滤液, 得到假吲哚3 (1.9g, 57% 收率), 为米色固体。C₂₇H₃₄NO₄ (M)⁺的HRMS (ESI) 计算值为436.2488, 实测值为436.2487。

[0524] 三甲胺花菁4



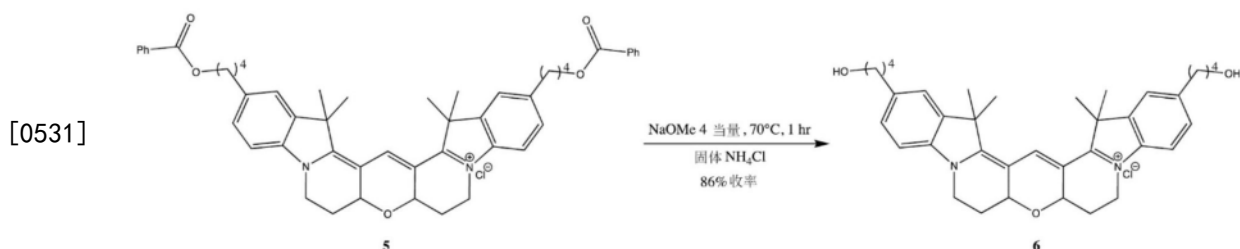
[0526] 向假吡啶 (3, 0.96g, 1.7mmol)、二苯基甲脞 (0.40g, 2.1mmol) 和 DMAP (21mg, 0.17mmol) 的混合物中加入乙酸酐 (5mL)。将棕色混合物加热至 120°C 持续 1 小时。冷却至室温后, 将挥发物真空浓缩。向粗制中间体中添加另外一部分的假吡啶 (3, 1.6g, 2.8mmol), 然后添加乙醇 (5mL) 和三甲胺 (1.2mL, 8.6mmol)。将反应在氩气下加热至回流 1 小时。将反应物用氯化钠水溶液稀释, 并用 DCM 萃取。有机层经硫酸镁干燥、过滤并真空浓缩。通过正相色谱法 (100% EtOAc, 然后 0→10% MeOH/DCM, SiO₂) 纯化粗残余物, 得到 Cy3 4 (0.9g, 收率 57%), 为深紫色固体。C₅₅H₆₅N₂O₈ (M)⁺ 的 HRMS (ESI) 计算值为 881.4735, 实测值为 881.4717。

[0527] Cy3B 类似物 5



[0529] 向含有 Cy3 4 (400mg, 0.437mmol) 的烧瓶中加入氯仿 (6mL) 和硫酸 (4mL, 50% v/v 水)。将两相混合物在 60°C 下剧烈搅拌 30 分钟, 从深红色转变为紫色。冷却至室温后, 将反应用水 (20mL) 稀释, 并用 EtOAc (50mL) 萃取。有机层用饱和氯化钠水溶液洗涤, 用硫酸镁干燥, 过滤, 并真空浓缩。通过正相色谱法 (100% EtOAc, 然后 0→15% MeOH/DCM, SiO₂) 纯化粗残余物, 得到 Cy3B 5 (177mg, 收率 50%), 为深紫色固体。C₅₁H₅₅N₂O₅ (M)⁺ 的 HRMS (ESI) 计算值为 775.4105, 实测值为 775.4091。

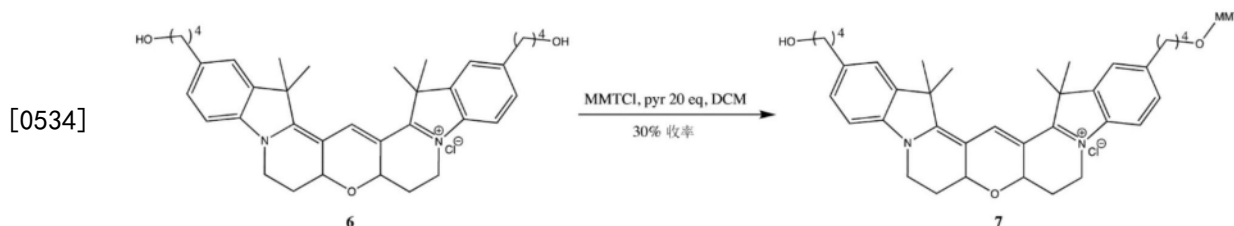
[0530] Cy3B 二醇 6



[0532] 向 Cy3B 二苯甲酸酯 5 (150mg, 0.185mmol) 在甲醇 (4mL) 中的溶液中加入甲醇钠 (0.37mL, 0.74mmol, 2.0M 的 MeOH 溶液)。将反应加热至 70°C 保持 1 小时。通过添加固体氯化铵 (66mg) 将反应淬灭, 并在室温下搅拌 30 分钟。将挥发物真空浓缩, 并将粗产物重新溶解于 DCM (10mL)。将悬浮液过滤并真空浓缩。通过正相色谱法 (100% EtOAc, 然后 0→25% MeOH/

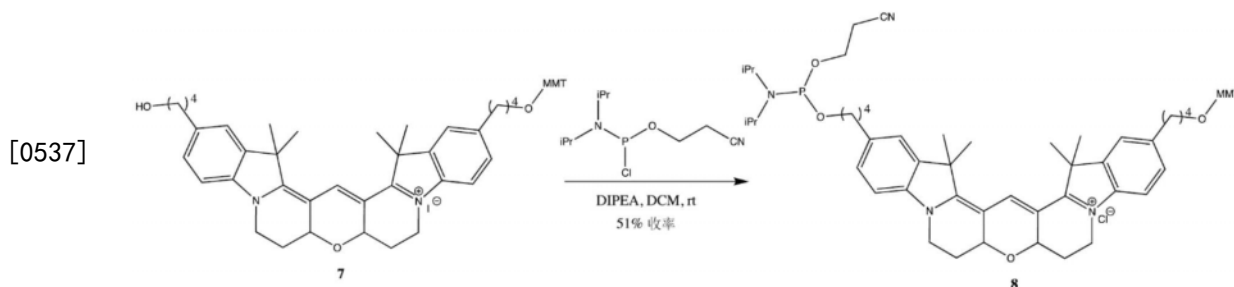
DCM, SiO₂) 纯化粗残余物, 得到Cy3B 6 (99mg, 收率86%), 为深紫色固体。C₃₇H₄₇N₂O₃ (M)⁺的HRMS (ESI) 实测值为567.3581, 观察值为567.3568

[0533] Cy3B单-MMT 7



[0535] 向Cy3B 二醇6 (100mg, 0.166mmol) 和单甲氧基三苯甲基氯 (62mg, 0.20mmol) 的二氯甲烷溶液中, 加入吡啶 (0.26mL, 3.3mmol)。将暗紫色反应在室温下搅拌30分钟。将反应物用DCM稀释, 并依次用水和饱和氯化钠水溶液洗涤。有机层经硫酸镁干燥, 过滤, 并真空浓缩。通过正相色谱法 (100% EtOAc, 然后0→25% 含1% Et₃N的MeOH/DCM, SiO₂) 纯化粗残余物, 得到Cy3B 7 (43mg, 30% 收率), 为深紫色固体。C₅₇H₆₃N₂O₄ (M)⁺的HRMS (ESI) 计算值为839.4782, 实测值为839.4751。

[0536] 亚磷酰胺8



[0538] 在室温下, 向Cy3B 7 (35mg, 0.036mmol) 和N,N-二异丙基乙胺 (14μL, 0.079mmol) 的无水二氯甲烷溶液中加入2-氰基乙基N,N-二异丙基氯代亚磷酰胺 (9.4mg, 0.040mmol)。将反应物用脱氧的DCM稀释, 用氯化钾水溶液洗涤, 经硫酸镁干燥, 过滤, 并真空浓缩。粗残余物通过正相色谱法纯化 (0→2% 含1% Et₃N的MeOH/DCM, 碱性氧化铝)。将产物重新溶解在DCM (1mL) 中, 并沉淀入己烷 (25mL) 中。在高真空下干燥, 得到Cy3B 8 (20mg, 51% 产率), 为深紫色固体。C₆₆H₈₀N₄O₅P (M)⁺的HRMS (ESI) 计算值为1039.5861, 实测值为1039.5835。

[0539] 等价物和范围

[0540] 在权利要求书中, 诸如“一”和“所述”的冠词可以表示一个或多个, 除非有相反的说明或从上下文中明显看出。如果组中的一个、多于一个或所有组成员在给定产品或过程中存在、在其中使用或以其他方式相关, 则认为包括组中一个或多个成员之间的“或”的声明或描述是满足的, 除非有相反的说明或从上下文中明显看出。本发明包括其中恰好有一个成员存在于、在其中使用或以其他方式与给定产品或工艺相关的实施方案。本发明包括其中多个或全部成员存在于、在其中使用或以其他方式与给定产品或工艺相关的实施方案。

[0541] 此外, 本发明涵盖所有变型、组合和置换, 其中来自所列权利要求中的一个或多个的一个或多个限制、要素、从句和描述性术语被引入到另一个权利要求中。例如, 可以修改依赖于另一个权利要求的任何权利要求以包括在依赖于相同基本权利要求的任何其他权利要求中找到的一个或多个限制。在元素被呈现为列表的情况下, 例如, 以马库什组的形

式,这些元素的每个子组也被公开,并且可以从组中移除任何元素。应该理解的是,通常,在本发明或本发明的方面被称为包括特定元素和/或特征的情况下,本发明的某些实施方案或本发明的方面包括、或基本上由这些元素和/或特征组成。出于简化的目的,这些实施方案未在本文中具体阐述。

[0542] 本说明书和权利要求书中使用的短语“和/或”应理解为指如此连接的元素中的“一个或两个”,即在某些情况下联合存在,在其他情况下分开存在的元素。用“和/或”列出的多个元素应以同样的方式解释,即,如此连接的元素中的“一个或多个”。除了由“和/或”子句具体确定的元素之外,其他元素可以任选地存在,无论其与具体确定的那些元素相关还是无关。因此,作为非限制性的例子,当与诸如“包括”这样的开放式语言结合使用时,对“A和/或B”的引用在一个实施方案中可以仅指A(任选地包括B以外的元素);在另一个实施方案中,仅指B(任选地包括A以外的元素);在另一个实施方案中,同时指A和B(任选地包括其他元素);等等。

[0543] 如本文在说明书和权利要求书中所使用的,“或”应理解为具有与上述定义的“和/或”相同的含义。例如,当分隔列表中的项目时,“或”或“和/或”应被解释为是包含性的,即包含若干元素或一系列元素中的至少一个,但也包括多个,以及,任选地,额外的未列出的项目。只有明确指出相反的术语,例如“只有一个”或“恰好一个”,或者当在权利要求中使用“由……组成”时,才表示包括若干元素或一系列元素中的恰好一个元素。一般来说,本文使用的术语“或”只有当前面有排他性术语,如“二者之一”、“其中之一”、“仅其中之一”或“恰好其中之一”时,才应解释为表示排他性选择(即“一个或另一个,但不是两个”)。在权利要求中使用的“基本上由……组成”,应具有其在专利法领域中使用的普通含义。

[0544] 如本文在说明书和权利要求中所使用的,“至少一个”这一短语在提及一个或多个元素的列表时,应理解为从元素列表中的任何一个或多个元素中选取的至少一个元素,但不一定包括元素列表中具体列出的每个元素中的至少一个,且不排除元素列表中的任何元素组合。该定义还允许,除了短语“至少一个”所指的元素列表内具体确定的元素之外,还可以任选地存在其他元素,无论这些元素与具体确定的元素相关还是不相关。因此,作为一个非限制性的实例,“A和B中的至少一个”(或者,等效地,“A或B中的至少一个”,或者,等效地,“A和/或B中的至少一个”)在一个实施例中,可以指至少一个,任选地包括一个以上的A,没有B存在(并且任选地包括B以外的元素);在另一个实施方案中,指至少一个,任选地包括一个以上的B,没有A存在(并任选地包括A以外的元素);在另一个实施方案中,指至少一个,任选地包括一个以上的A,以及至少一个,任选地包括一个以上的B(并任选地包括其他元素);等等。

[0545] 还应理解,除非有相反的明确指示,在本权利要求的任何方法中,包括一个以上的步骤或行为,该方法的步骤或行为的顺序不一定限于该方法的步骤或行为的叙述顺序。

[0546] 在权利要求书以及上述说明书中,所有的过渡性短语如“包括”、“包含”、“带有”、“具有”、“含”、“涉及”、“持有”、“构成组成”等都应理解为是开放式的,即指包括但不限于此。只有“由……组成”和“基本上由……组成”的过渡短语应分别为美国专利局《专利审查程序手册》第2111.03节规定的封闭或半封闭的过渡短语。应当理解的是,在本文中使用开放式过渡短语(例如,“包括”)描述的实施方案,在其他实施方案中,也被认为是“由”和“基本上由”开放式过渡短语所描述的特征组成的。例如,如果公开内容描述了“包含A和B的组合

物”，则公开还考虑了其他实施方案“由A和B组成的组合物”和“基本上由A和B组成的组合物”。

[0547] 在给出范围的情况下，包括端点。此外，除非另外指出或从本领域普通技术人员的上下文和理解中显而易见，否则表示为范围的值可以假设在本发明的不同实施方案中的所述范围内的任何特定值或子范围，除非上下文另有明确规定，否则该范围下限单位的十分之一。

[0548] 本申请涉及各种已发布的专利，已公布的专利申请，期刊文章和其他出版物，所有的这些通过引用并入本文。如果任何并入的参考文献与本说明书之间存在冲突，则应以本说明书为准。另外，落入现有技术内的本发明的任何特定实施方案可以明确地从任何一个或多个权利要求中排除。因为这些实施方案被认为是本领域普通技术人员已知的，所以即使在此未明确阐述排除，也可以排除它们。无论是否与现有技术存在相关，本发明的任何特定实施方案可以出于任何原因被排除在任何权利要求之外。

[0549] 本领域技术人员将认识到或能够使用不超过常规的实验确定本文所述的具体实施方案的许多等价物。本文描述的本发明实施方案的范围不旨在限于以上说明书，而是在所附权利要求书中阐述。本领域普通技术人员将理解，在不脱离如所附权利要求书所限定的本发明的精神或范围的情况下，可以对本说明书进行各种改变和修改。

[0550] 在本文中对变量的任何定义中对化学基团的列举包括该变量作为任何单一基团或所列基团的组合的定义。本文对一个变量的实施方案的叙述包括该实施方案作为任何单一的实施方案或与任何其他实施方案或其部分的组合。本文对一个实施方案的叙述包括该实施方案作为任何单一的实施方案或与任何其他实施方案或其部分的组合。

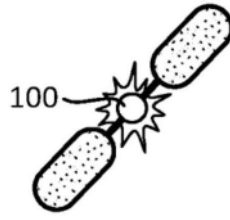


图1A

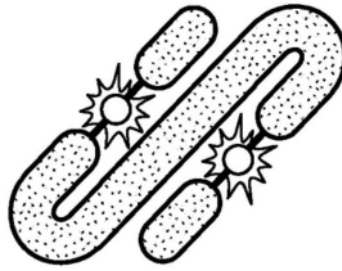


图1B



图1C

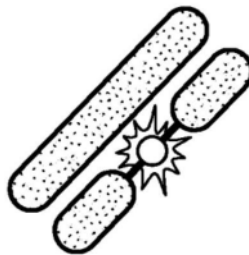


图1D

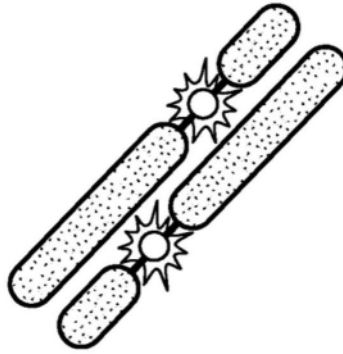


图1E

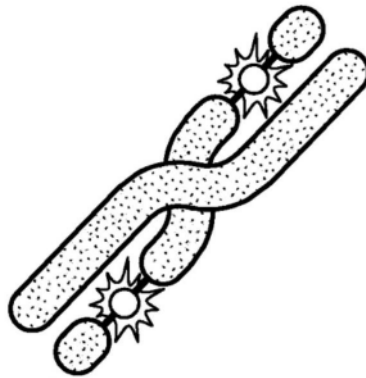


图1F



图1G

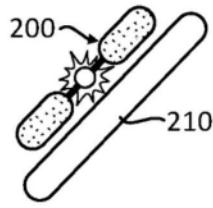


图2A

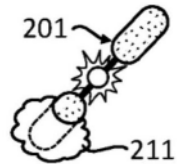


图2B

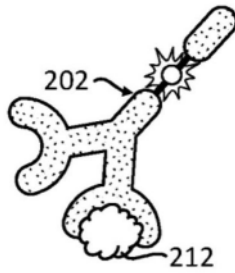


图2C

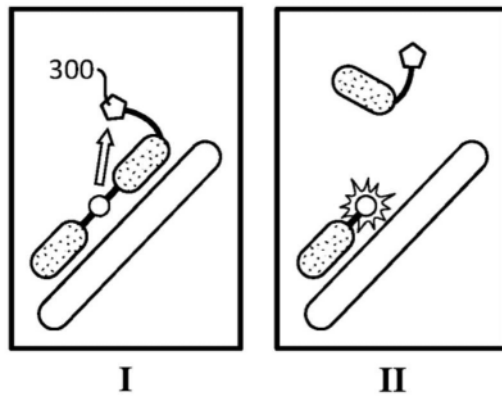


图3A

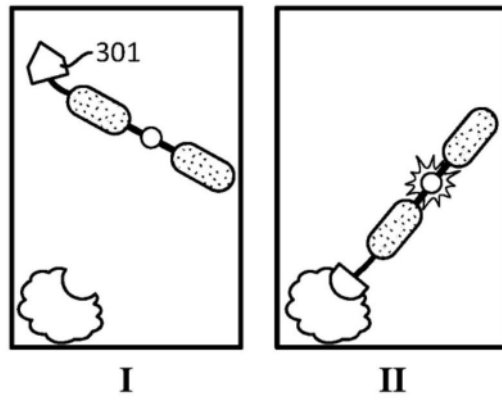


图3B

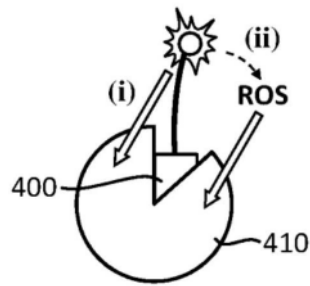


图4A

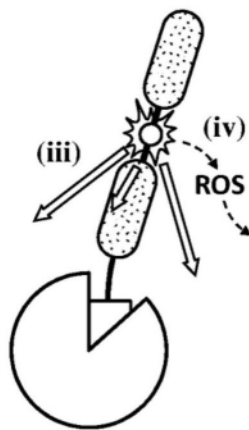


图4B

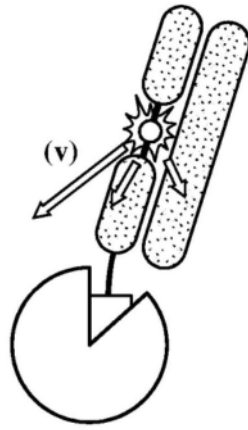


图4C

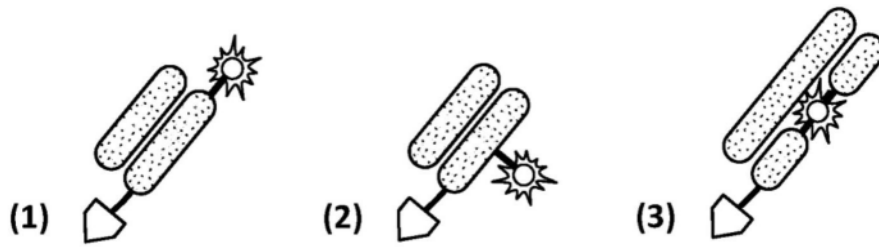


图5A

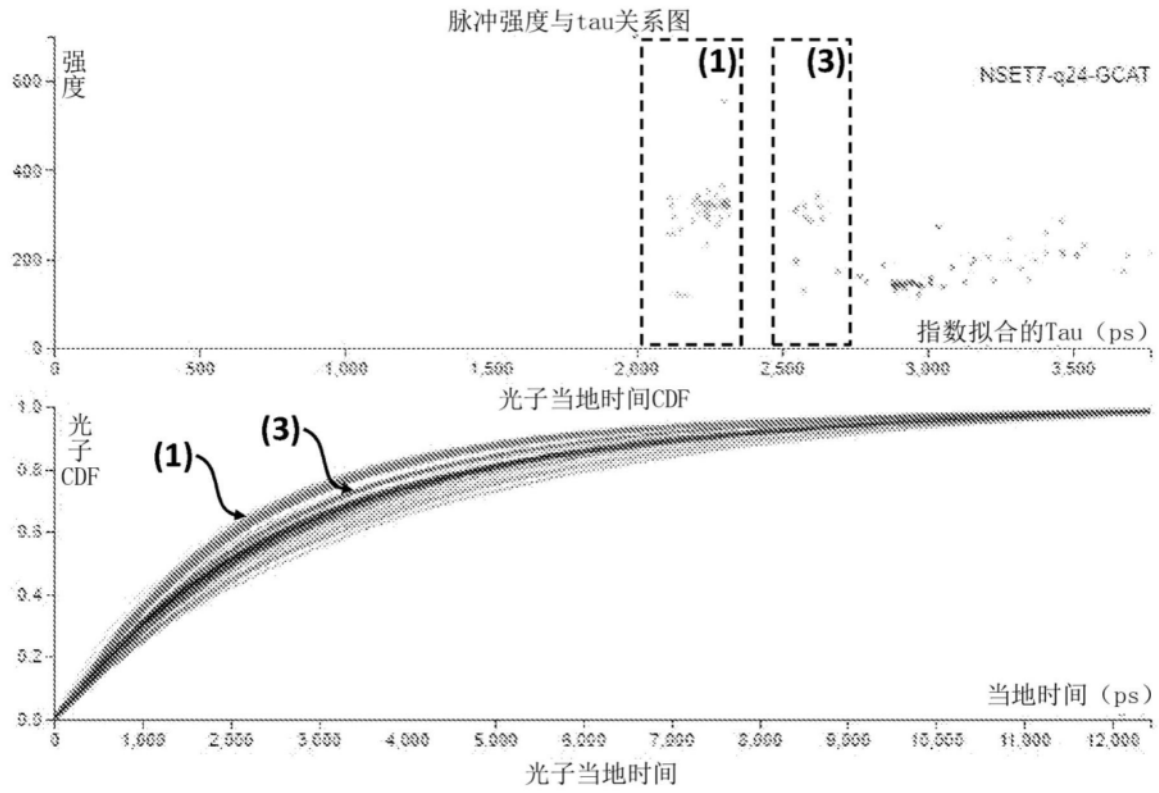


图5B

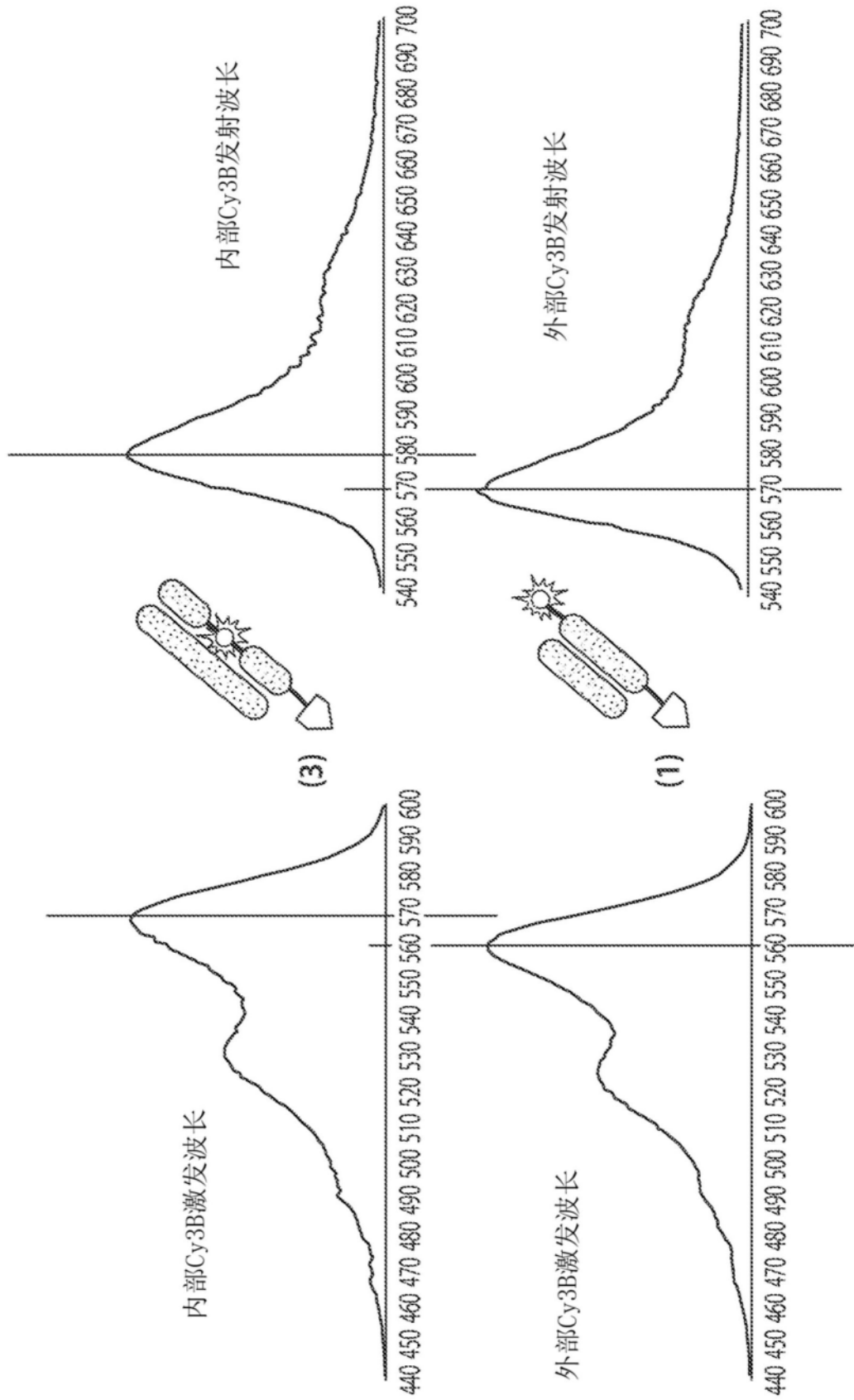


图6