



(12) 发明专利申请

(10) 申请公布号 CN 116615557 A

(43) 申请公布日 2023. 08. 18

(21) 申请号 202180079320.9

(22) 申请日 2021.11.22

(30) 优先权数据

63/118,544 2020.11.25 US

(85) PCT国际申请进入国家阶段日

2023.05.25

(86) PCT国际申请的申请数据

PCT/US2021/060378 2021.11.22

(87) PCT国际申请的公布数据

W02022/115388 EN 2022.06.02

(71) 申请人 索尼集团公司

地址 日本东京

(72) 发明人 T·马特雷 S·辛格

(74) 专利代理机构 中国贸促会专利商标事务所

有限公司 11038

专利代理师 于巧玲

(51) Int.Cl.

C12Q 1/6818 (2006.01)

权利要求书22页 说明书47页

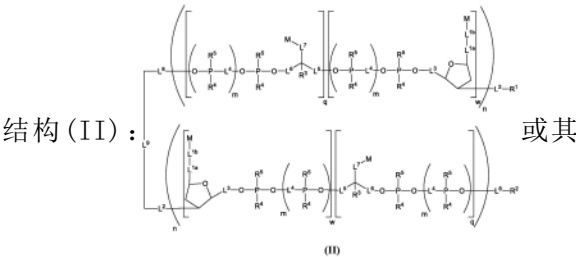
序列表6页 附图12页

(54) 发明名称

聚合物染料

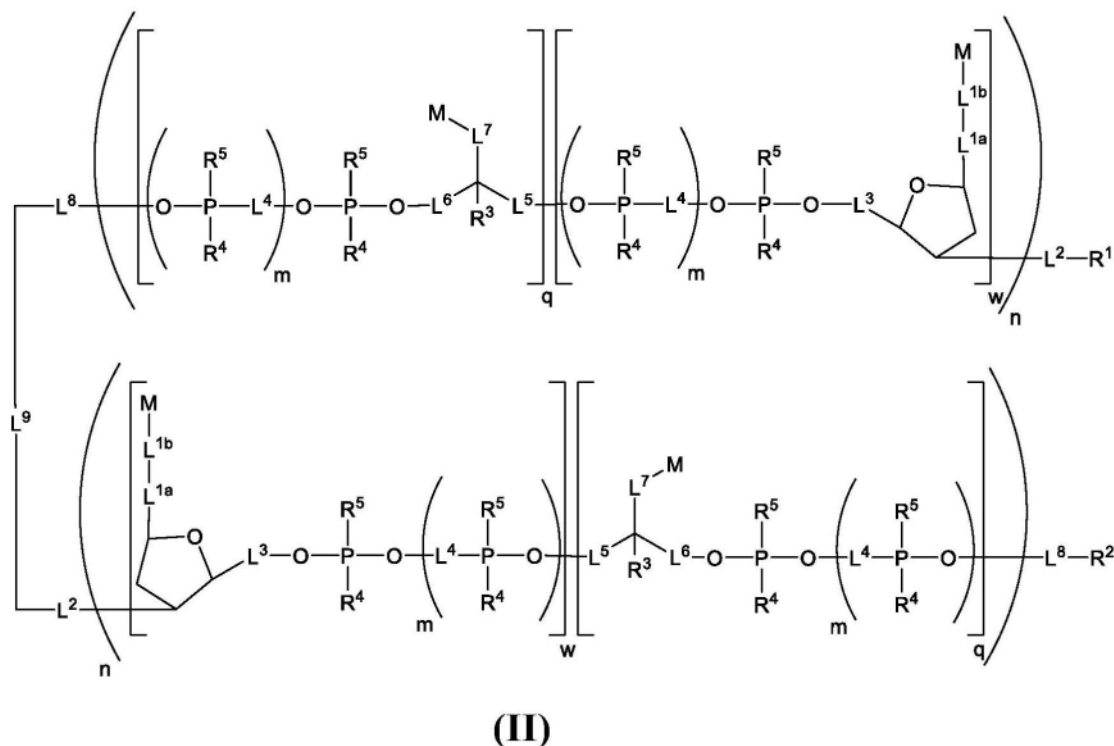
(57) 摘要

公开了用作探针的化合物。化合物具有如下



立体异构体、互变异构体或盐,其中M、L<sup>1a</sup>、L<sup>1b</sup>、L<sup>2</sup>、L<sup>3</sup>、L<sup>4</sup>、L<sup>5</sup>、L<sup>6</sup>、L<sup>7</sup>、L<sup>8</sup>、L<sup>9</sup>-R<sup>1</sup>、R<sup>2</sup>、R<sup>3</sup>、R<sup>4</sup>、R<sup>5</sup>、m、n、q和w如本文所定义。此外,还描述了可用于检测靶标分析物的组合物、试剂盒和方法。

1. 具有以下结构(II)的化合物:



或其立体异构体、盐或互变异构体,其中:

M在每次出现时独立地是生色团;

L<sup>1a</sup>在每次出现时独立地是杂亚芳基连接基;

L<sup>2</sup>和L<sup>8</sup>是独立任意的连接基;

L<sup>1b</sup>、L<sup>3</sup>、L<sup>5</sup>、L<sup>6</sup>和L<sup>7</sup>在每次出现时是独立任意的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

L<sup>4</sup>在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

L<sup>9</sup>是包含多核苷酸的连接基;

R<sup>1</sup>和R<sup>2</sup>各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基或-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>;

R<sup>3</sup>在每次出现时独立地为H、烷基或烷氧基;

R<sup>4</sup>在每次出现时独立地为OH、SH、O<sup>-</sup>、S<sup>-</sup>、OR<sub>d</sub>或SR<sub>d</sub>;

R<sup>5</sup>在每次出现时独立地为氧代、硫代或不存在;

R<sub>a</sub>为O或S;

R<sub>b</sub>为OH、SH、O<sup>-</sup>、S<sup>-</sup>、OR<sub>d</sub>或SR<sub>d</sub>;

R<sub>c</sub>为OH、SH、O<sup>-</sup>、S<sup>-</sup>、OR<sub>d</sub>、OL'、SR<sub>d</sub>、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;

R<sub>d</sub>是抗衡离子;

L'在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;

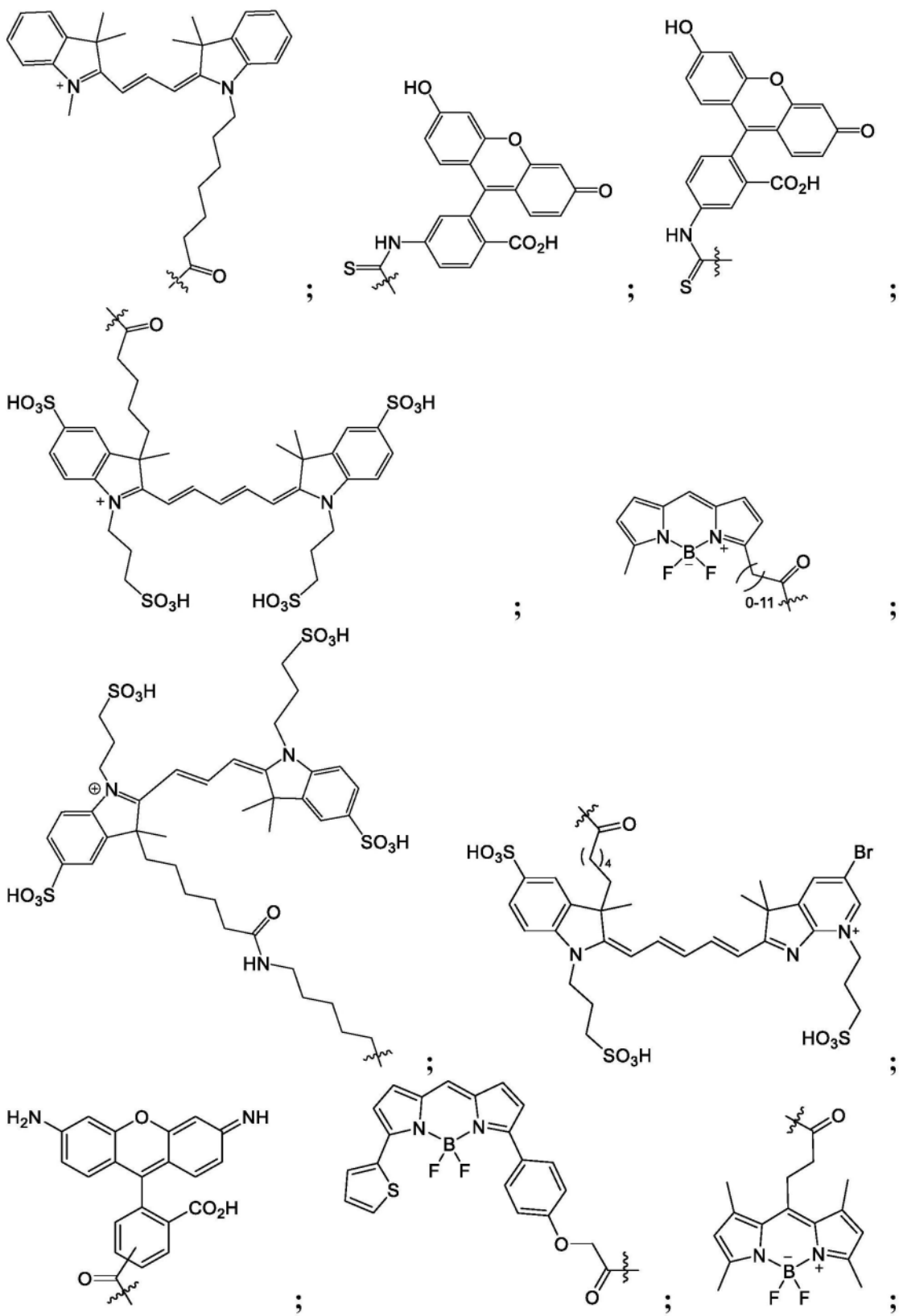
m在每次出现时独立地为0或更大的整数;

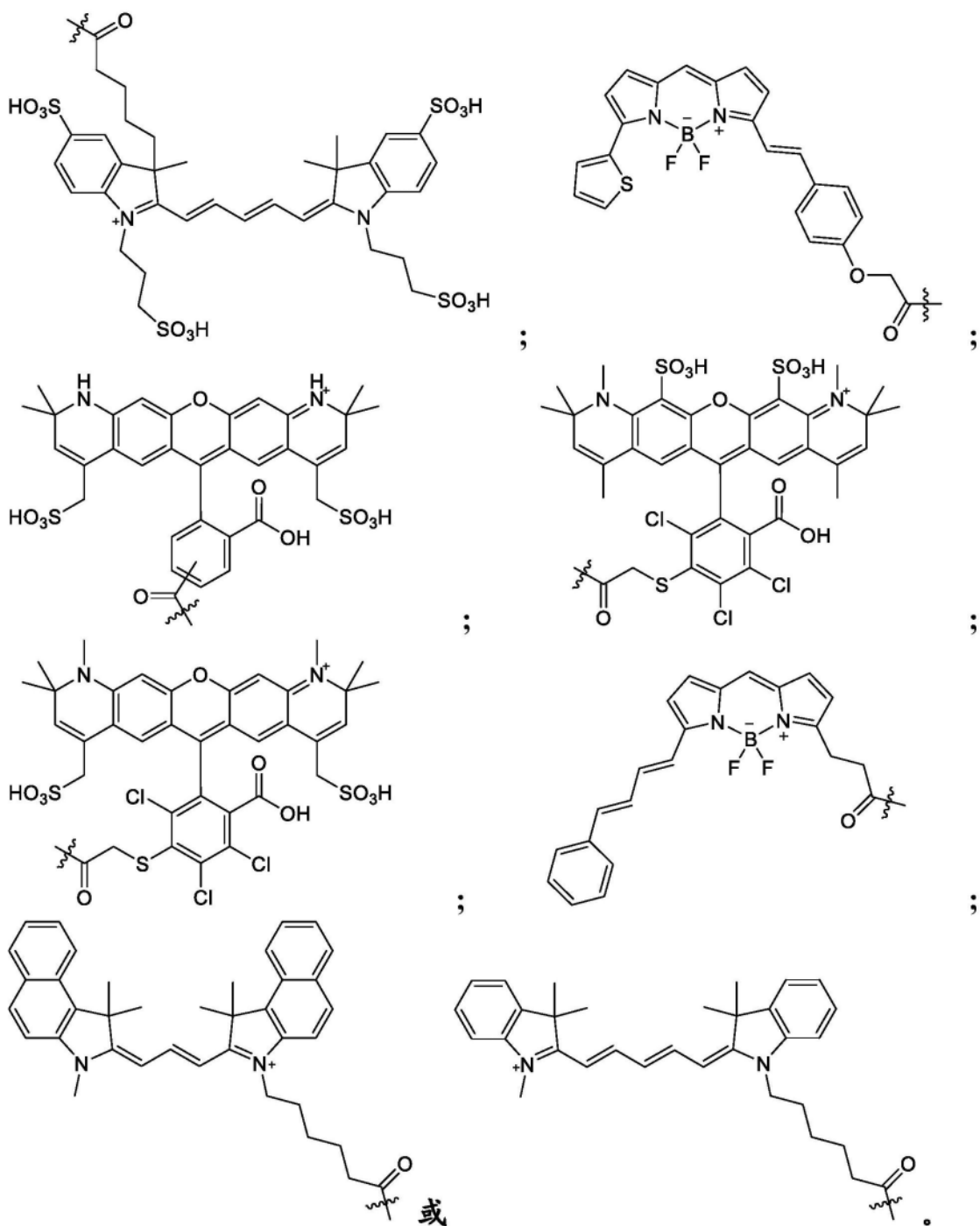
n是一或更大的整数;

- q和w在每次出现时独立地为0或1,条件是q或w至少有一次出现时是1。
2. 根据权利要求1所述的化合物,其中M在每次出现时独立地是相同或不同的供体生色团。
3. 根据权利要求1或2所述的化合物,其中M在每次出现时是相同的供体生色团。
4. 根据权利要求1所述的化合物,其中M在每次出现时独立地是相同或不同的受体生色团。
5. 根据权利要求1或4所述的化合物,其中M在每次出现时是相同的受体生色团。
6. 根据权利要求1-5中任一项所述的化合物,其中M在每次出现时独立地是相同或不同的荧光团。
7. 根据权利要求6所述的化合物,其中M在每次出现时独立地是相同或不同的供体荧光团。
8. 根据权利要求7所述的化合物或组合物,其中M在每次出现时是相同的供体荧光团。
9. 根据权利要求6所述的化合物,其中M在每次出现时独立地是相同或不同的受体荧光团。
10. 根据权利要求9所述的化合物,其中M在每次出现时是相同的受体荧光团。
11. 根据权利要求1的化合物,其中M在每次出现时独立地为二甲基氨基芘,喹吖啶酮,氟苯基-二甲基-BODIPY, his-氟苯基-BODIPY, 吡啶, 骈三萘, 联六苯, 卟啉, 苯并芘, (氟苯基-二甲基-二氟硼杂-二氮杂-引达省)苯基, (二-氟苯基-二氟硼杂-二氮杂-引达省)苯基, 四联苯, 联-苯并噻唑, 三联-苯并噻唑, 联-萘, 联-蒽, 方酸, 方酸鎗, 9,10-乙炔基蒽, 或三联-萘部分。
12. 根据权利要求1的化合物,其中M在每次出现时独立地为芘、茈、茈一酰亚胺、5-FAM或6-FAM或其衍生物。
13. 根据权利要求1的化合物,其中M在每次出现时独立地具有以下结构之一:

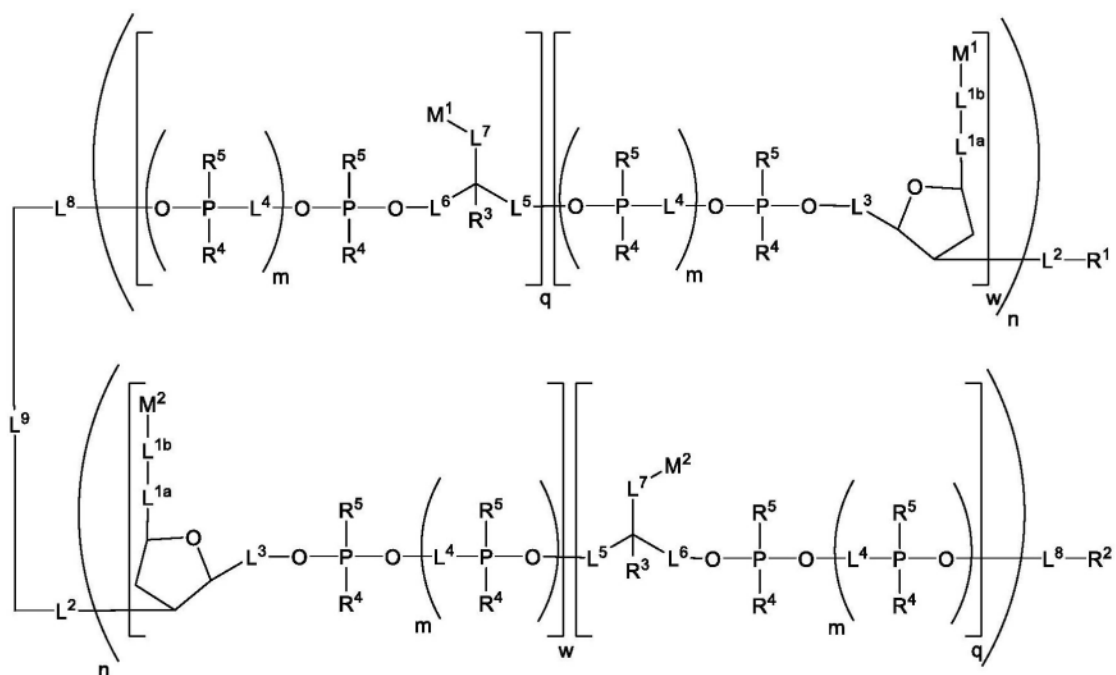






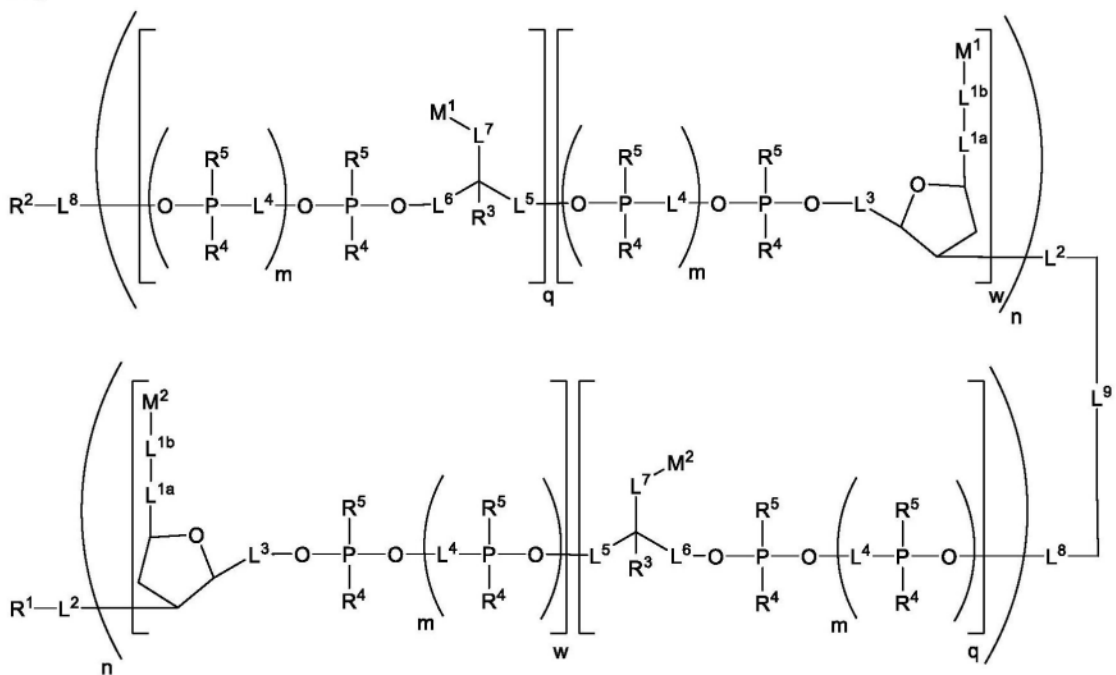


15. 根据权利要求1的化合物, 具有以下结构(IIa)或(IIb):



(IIa)

或



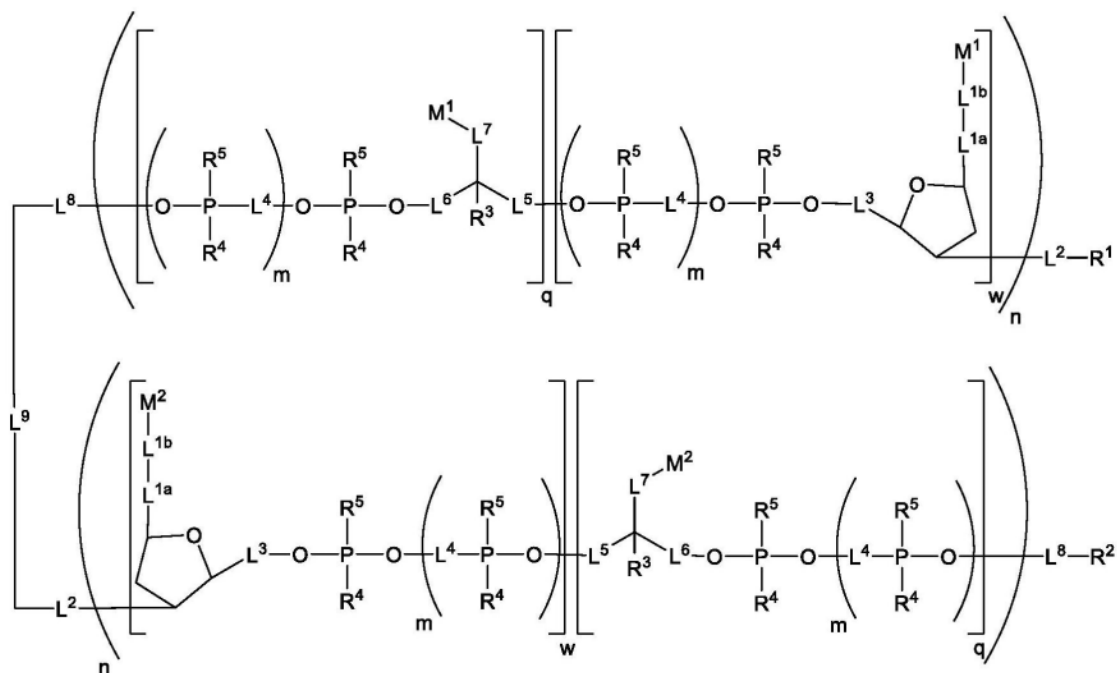
(IIb)

或其立体异构体、盐或互变异构体,其中:

$M^1$ 在每次出现时独立地是相同或不同的供体生色团;

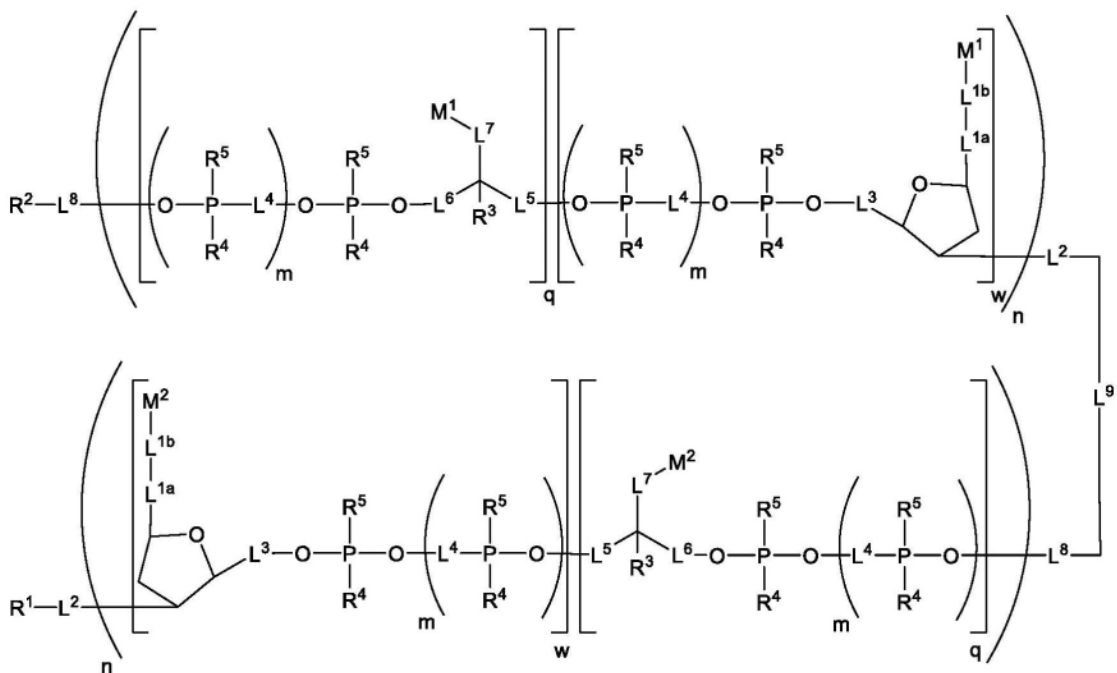
$M^2$ 在每次出现时独立地是相同或不同的受体生色团。

16. 根据权利要求15所述的化合物,其具有以下结构(IIa):



(IIa) .

17. 根据权利要求15的化合物, 具有以下结构(IIb):



(IIb) .

18. 根据权利要求15-17中任一项所述的化合物, 其中 $M^1$ 在每次出现时独立地是相同或不同的供体荧光团。

19. 根据权利要求15-18中任一项所述的化合物, 其中 $M^2$ 在每次出现时独立地是相同或不同的受体荧光团。

20. 根据权利要求15-19中任一项所述的化合物, 其中 $M^1$ 在每次出现时是相同的供体荧光团。

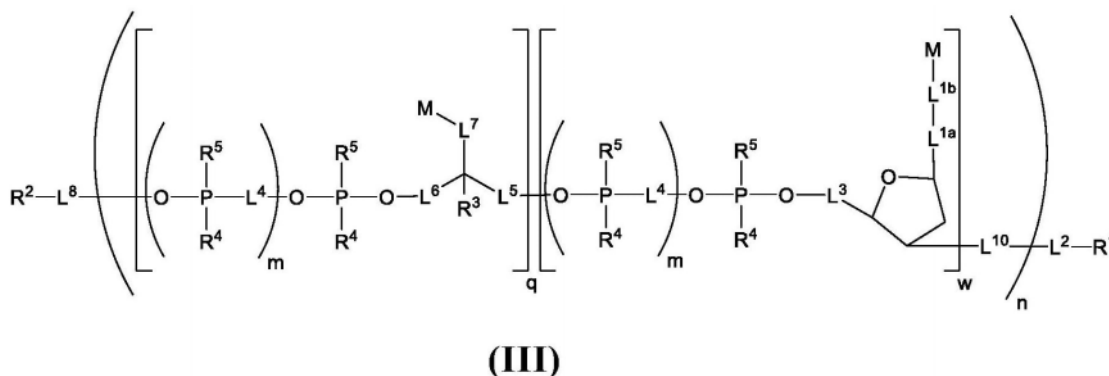
21. 根据权利要求15-20中任一项所述的化合物,其中 $M^2$ 在每次出现时是相同的受体荧光团。

22. 根据权利要求1-21中任一项所述的化合物,其中 $L^9$ 多核苷酸的长度范围为10个核苷酸至40个核苷酸。

23. 根据权利要求1-22中任一项所述的化合物,其中 $L^9$ 多核苷酸的长度范围为10个核苷酸至30个核苷酸。

24. 根据权利要求1-23中任一项所述的化合物,其中 $L^9$ 多核苷酸的长度为15个核苷酸至25个核苷酸。

25. 具有以下结构(III)的化合物:



或其立体异构体、盐或互变异构体,其中:

$M$ 在每次出现时独立地为相同或不同的生色团;

$L^{1a}$ 在每次出现时独立地是杂亚芳基连接基;

$L^2$ 和 $L^8$ 是独立任选的连接基;

$L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立地任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

$L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

$L^{10}$ 是包含多核苷酸的连接基;

$R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基、 $-OP(=R_a)(R_b)R_c$ 或为包含多核苷酸的部分,条件是 $R^1$ 和 $R^2$ 中的至少一个为包含多核苷酸的部分;

$R^3$ 在每次出现时独立地为H、烷基或烷氧基;

$R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R^5$ 在每次出现时独立地为氧代、硫代或不存在;

$R_a$ 为O或S;

$R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

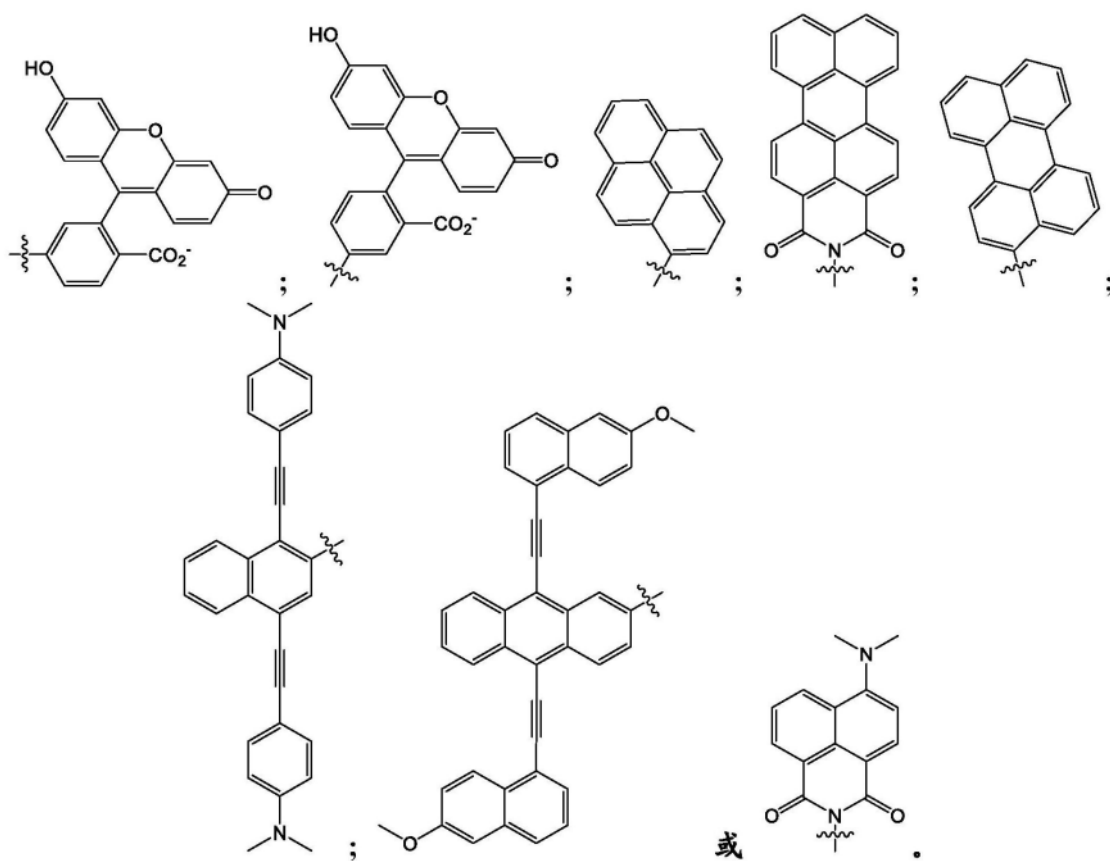
$R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;

$R_d$ 是抗衡离子;

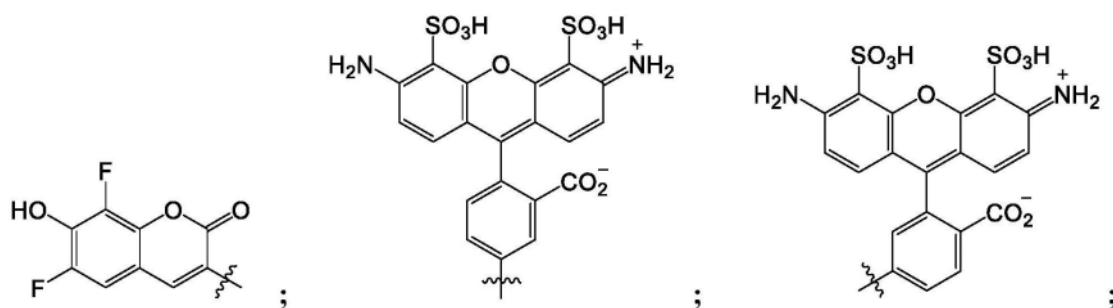
$L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;

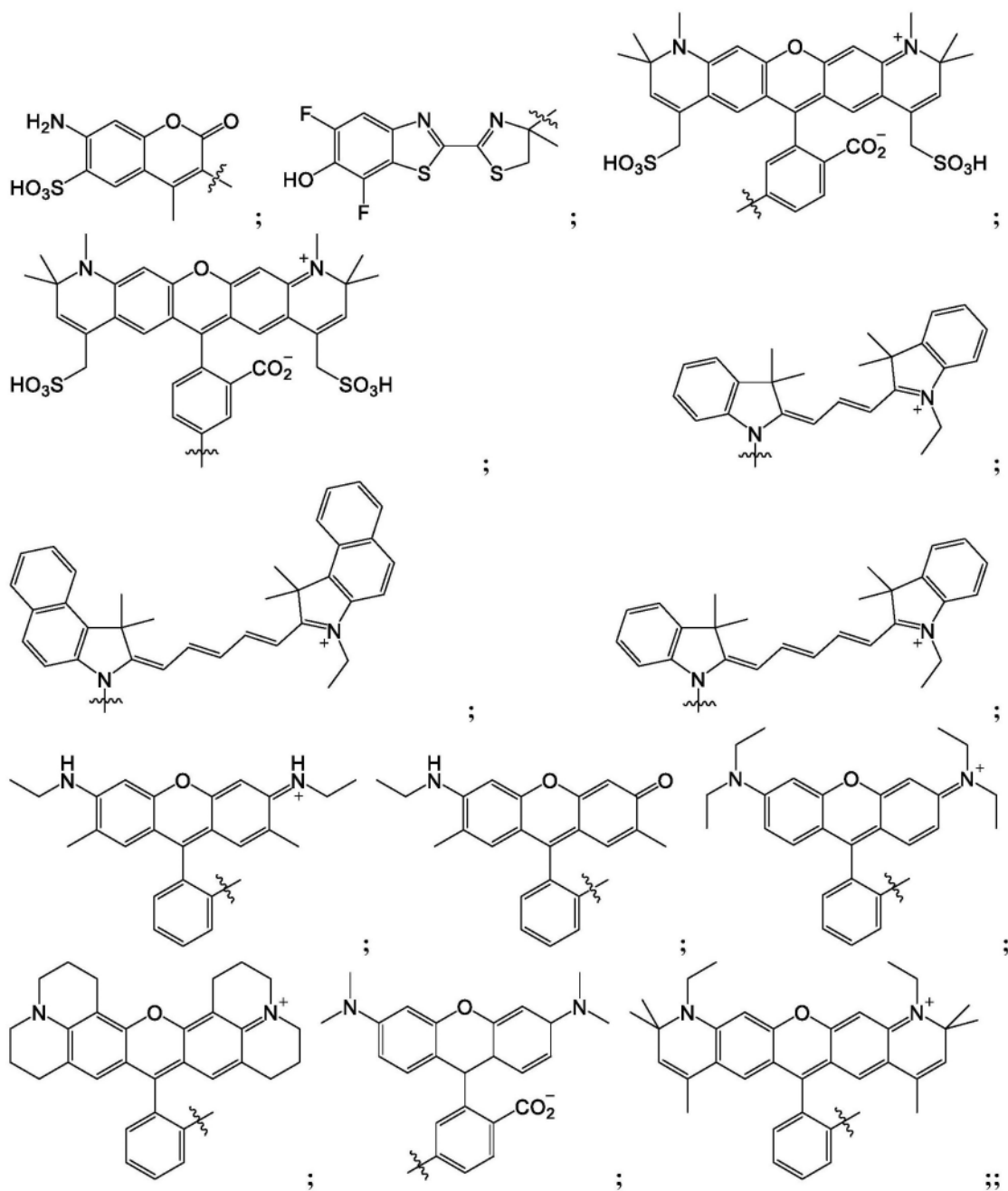
$m$ 在每次出现时独立地为0或更大的整数;

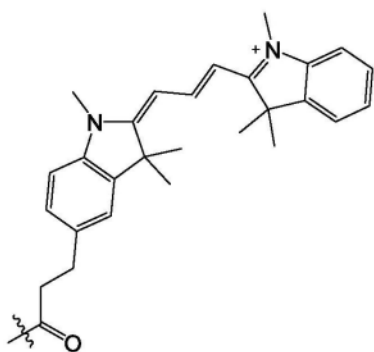
- n是2或更大的整数；
- q和w在每次出现时独立地为0或1，条件是q或w至少有一次出现时是1。
26. 根据权利要求25所述的化合物，其中R<sup>1</sup>包含多核苷酸。
27. 根据权利要求25或26所述的化合物，其中R<sup>2</sup>包含多核苷酸。
28. 根据权利要求26或27所述的化合物，其中R<sup>1</sup>包含结合至所述多核苷酸的靶向部分。
29. 根据权利要求27或28所述的化合物，其中R<sup>2</sup>包含结合至所述多核苷酸的靶向部分。
30. 根据权利要求25-29中任一项所述的化合物，其中M在每次出现时独立地是相同或不同的供体生色团。
31. 根据权利要求30所述的化合物，其中M在每次出现时是相同的供体生色团。
32. 根据权利要求25-29中任一项所述的化合物，其中M在每次出现时独立地是相同或不同的受体生色团。
33. 根据权利要求32所述的化合物，其中M在每次出现时是相同的受体生色团。
34. 根据权利要求25-29中任一项所述的化合物，其中M在每次出现时独立地是相同或不同的荧光团。
35. 根据权利要求34的化合物，其中M在每次出现时独立地是相同或不同的供体荧光团。
36. 根据权利要求34所述的化合物，其中M在每次出现时是相同的供体荧光团。
37. 根据权利要求34所述的化合物，其中M在每次出现时独立地是相同或不同的受体荧光团。
38. 根据权利要求37所述的化合物，其中M在每次出现时是相同的受体荧光团。
39. 根据权利要求25-29中任一项的化合物，其中M在每次出现时独立地为二甲基氨基苣，喹吖啶酮，氟苯基-二甲基-BODIPY，his-氟苯基-BODIPY，吖啶，骈三苯，联六苯，卟啉，苯并苣，(氟苯基-二甲基-二氟硼杂-二氮杂-引达省)苯基，(二-氟苯基-二氟硼杂-二氮杂-引达省)苯基，四联苯，联-苯并噻唑，三联-苯并噻唑，联-苯，联-萘，方酸，方酸鎗，9,10-乙炔基萘，或三联-萘部分。
40. 根据权利要求25-29中任一项的化合物，其中M在每次出现时独立地为苣、花、花-酰亚胺、5-FAM或6-FAM或其衍生物。
41. 根据权利要求25-29中任一项的化合物，其中M在每次出现时独立地具有以下结构之一：



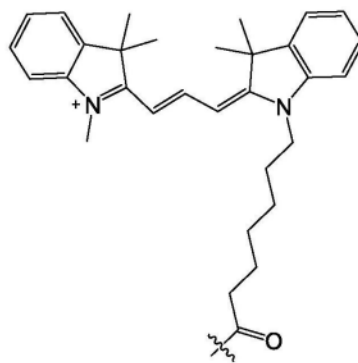
42. 根据权利要求25-29中任一项的化合物,其中M在每次出现时独立地具有以下结构之一:



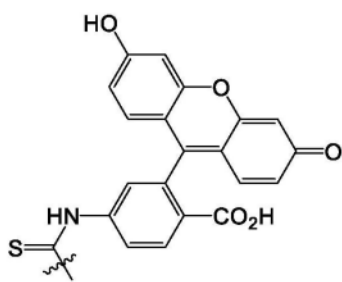




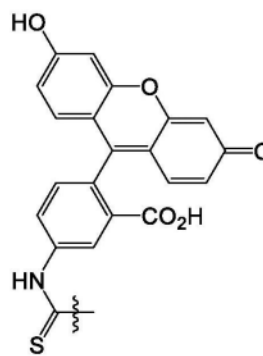
;



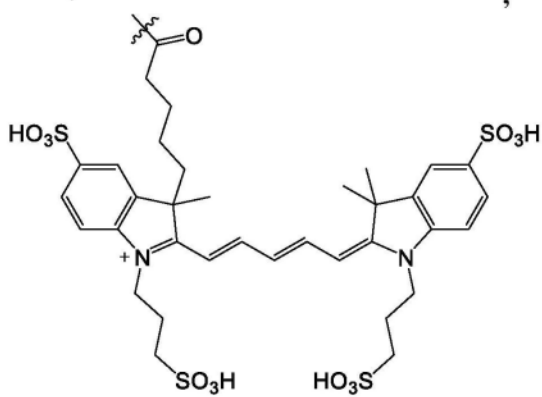
;



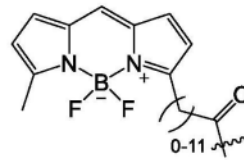
;



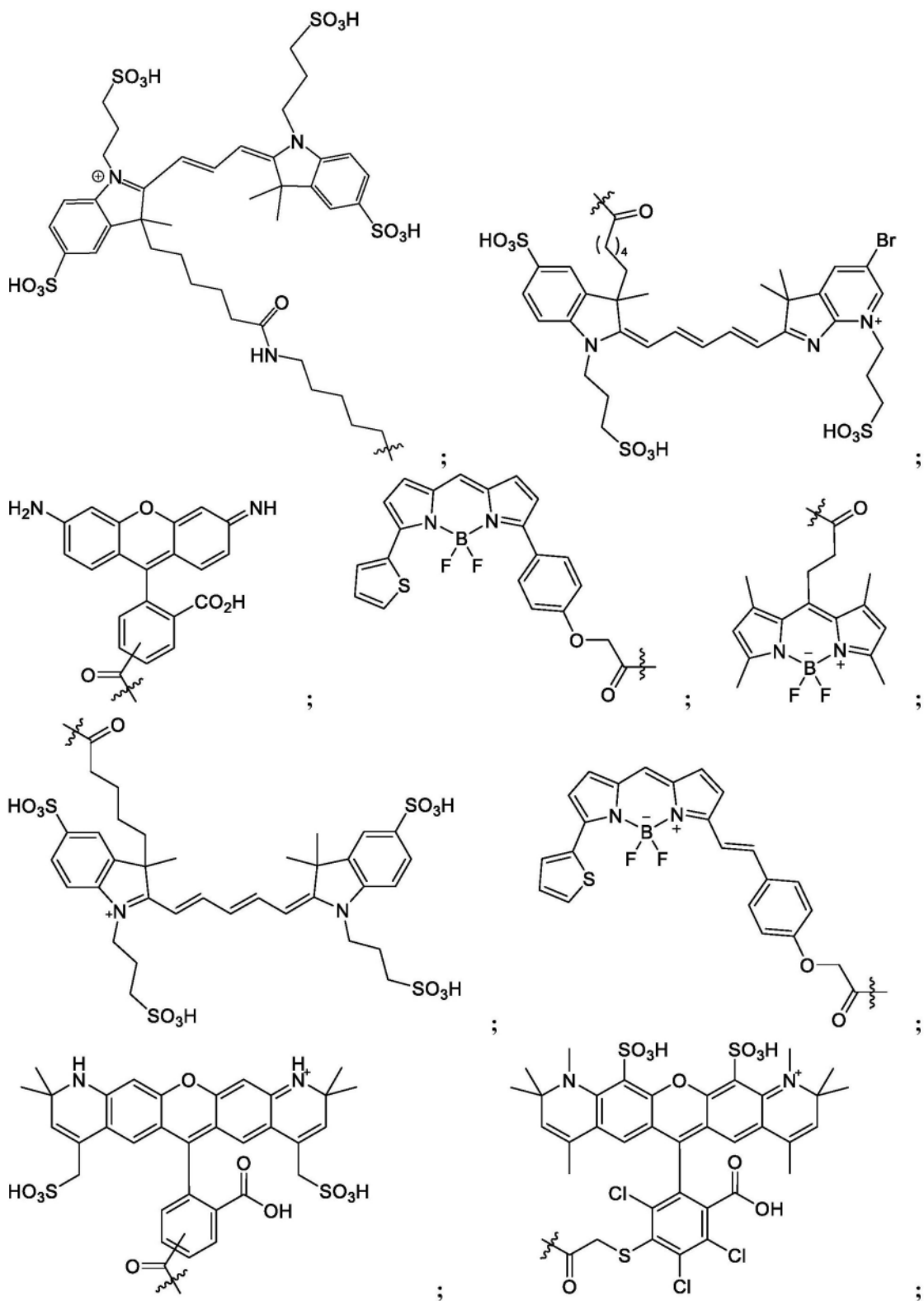
;

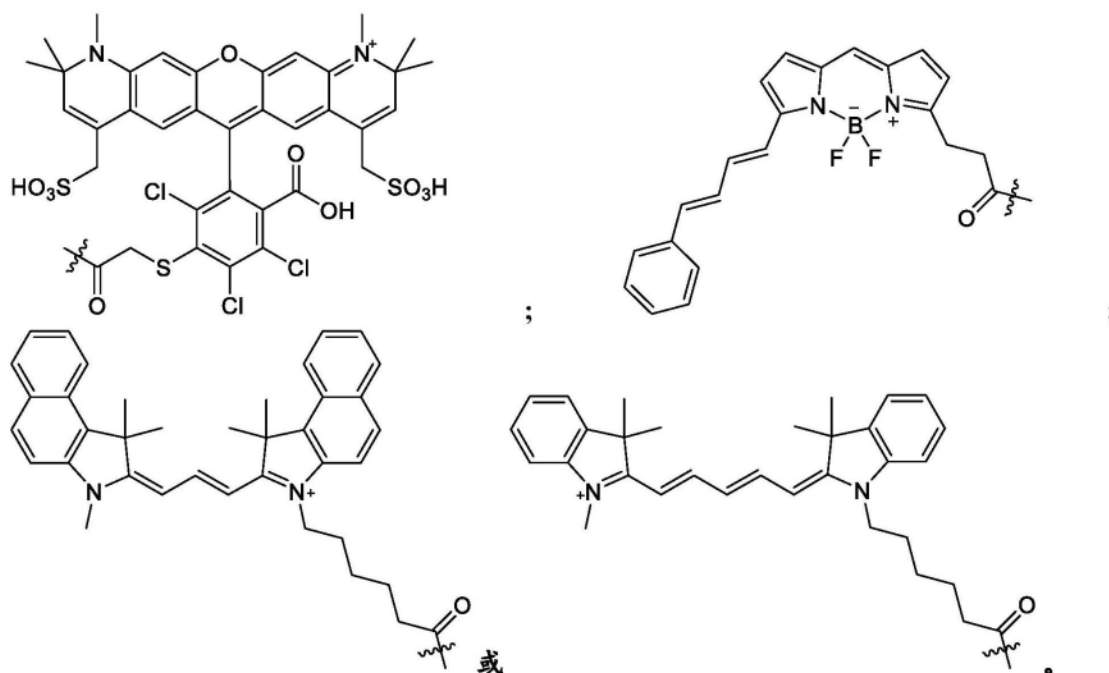


;



;



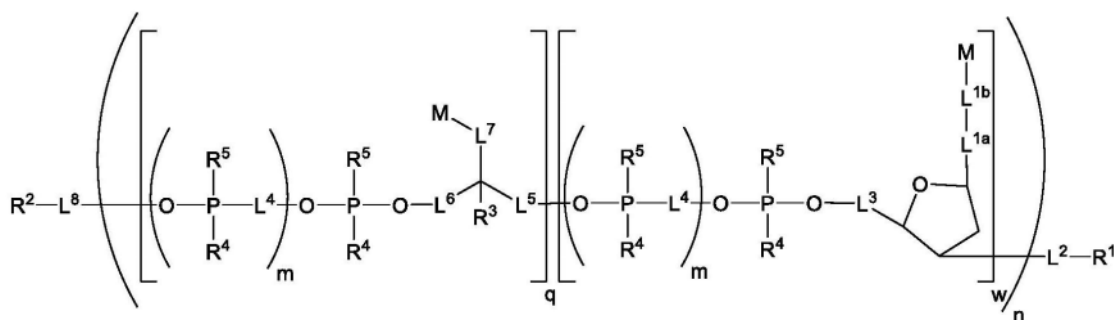


43. 组合物, 包含: 权利要求1-24中任一项的化合物和权利要求25-42中任一项的化合物。

44. 根据权利要求43的组合物, 其中 $L^9$ 的多核苷酸具有第一序列,  $L^{10}$ 的多核苷酸具有第二序列, 第一序列与第二序列的至少一部分具有至少90%的互补性。

45. 组合物, 其包含权利要求1-24中任一项的化合物和捕获探针。

46. 根据权利要求43-45中任一项所述的组合物, 其进一步包含: 具有以下结构(I)的化合物:



多核苷酸的部分,条件是 $R^1$ 和 $R^2$ 中的至少一个为包含多核苷酸的部分;

$R^3$ 在每次出现时独立地为H、烷基或烷氧基;

$R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R^5$ 在每次出现时独立地为氧代、硫代或不存在;

$R_a$ 为O或S;

$R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;

$R_d$ 是抗衡离子;

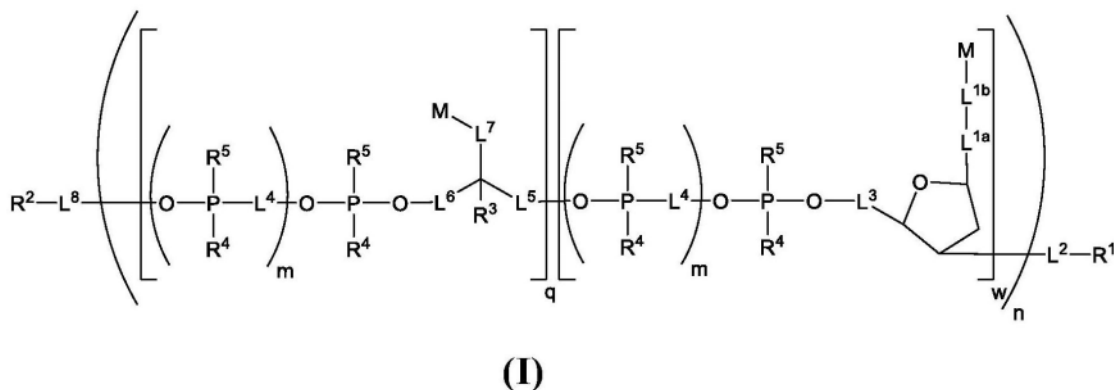
$L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;

$m$ 在每次出现时独立地为0或更大的整数;

$n$ 是1或更大的整数;

$q$ 和 $w$ 在每次出现时独立地为0或1,条件是 $q$ 或 $w$ 至少有一次出现时是1。

47. 组合物,其包含捕获探针和具有以下结构(I)的化合物:



或其立体异构体、盐或互变异构体,其中:

$M$ 在每次出现时独立地相同或不同的生色团;

$L^{1a}$ 在每次出现时独立地是杂亚芳基连接基;

$L^2$ 和 $L^8$ 是独立任选的连接基;

$L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立地任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

$L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基

$R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基、 $-OP(=R_a)(R_b)R_c$ 或为包含多核苷酸的部分,条件是 $R^1$ 和 $R^2$ 中的至少一个为包含多核苷酸的部分;

$R^3$ 在每次出现时独立地为H、烷基或烷氧基;

$R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R^5$ 在每次出现时独立地为氧代、硫代或不存在;

$R_a$ 为O或S;

$R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚；

$R_d$ 是抗衡离子；

$L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基；

$m$ 在每次出现时独立地为0或更大的整数；

$n$ 是1或更大的整数；

$q$ 和 $w$ 在每次出现时独立地为0或1，条件是 $q$ 或 $w$ 至少有一次出现时是1。

48. 根据权利要求45-47中任一项的组合物，其中 $R^1$ 或 $R^2$ 的多核苷酸具有第一序列，捕获探针具有第二序列，第一序列与第二序列的至少第一部分具有至少90%的互补性。

49. 根据权利要求45-47中任一项所述的组合物，其进一步包含分支连接基。

50. 根据权利要求49所述的组合物，其中所述分支连接基包含第一多核苷酸和多个第二多核苷酸。

51. 根据权利要求50所述的组合物，其中所述分支连接基的第一多核苷酸具有第一序列并且捕获探针具有第二序列，第一序列与第二序列的至少第一部分具有至少90%的互补性。

52. 根据权利要求50或51所述的组合物，其中所述分支连接基的每个第二多核苷酸具有第三序列并且 $R^1$ 或 $R^2$ 的多核苷酸具有第四序列，第三序列具有与第四序列至少90%的互补性。

53. 根据权利要求52所述的组合物，其中第三序列具有与第四序列至少92%的互补性。

54. 根据权利要求52或53的组合物，第三序列具有与第四序列至少95%的互补性。

55. 根据权利要求52-54中任一项的组合物，第三序列具有与第四序列至少97%的互补性。

56. 根据权利要求52-55中任一项的组合物，第三序列具有与第四序列至少98%的互补性。

57. 根据权利要求52-56中任一项的组合物，第三序列具有与第四序列至少99%的互补性。

58. 根据权利要求52-57中任一项的组合物，其中所述第三序列、第四序列或两者的长度范围为10个核苷酸至40个核苷酸。

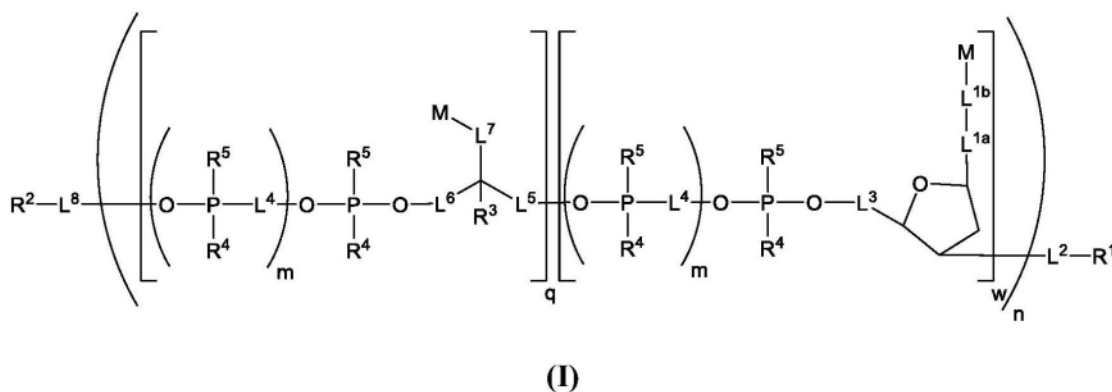
59. 根据权利要求52-58中任一项所述的组合物，其中所述第三序列、第四序列或两者的长度范围为10个核苷酸至30个核苷酸。

60. 根据权利要求52-59中任一项所述的组合物，其中所述第三序列、第四序列或两者的长度范围为15个核苷酸至25个核苷酸。

61. 根据权利要求45-60中任一项的组合物，其中所述捕获探针包含靶向部分。

62. 根据权利要求61所述的组合物，其中所述靶向部分是抗体。

63. 组合物，其包含权利要求25-42中任一项的化合物；和具有以下结构(I)的化合物：



或其立体异构体、盐或互变异构体,其中:

M在每次出现时独立地是相同或不同的生色团;

$L^{1a}$ 在每次出现时独立地是杂亚芳基连接基;

$L^2$ 和 $L^8$ 是独立任意的连接基;

$L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立地任意的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

$L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

$R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基、 $-OP(=R_a)(R_b)R_c$ 或为包含多核苷酸的部分,条件是 $R^1$ 和 $R^2$ 中的至少一个为包含多核苷酸的部分;

$R^3$ 在每次出现时独立地为H、烷基或烷氧基;

$R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R^5$ 在每次出现时独立地为氧代、硫代或不存在;

$R_a$ 为O或S;

$R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

$R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;

$R_d$ 是抗衡离子;

$L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;

m在每次出现时独立地为0或更大的整数;

n是1或更大的整数;

q和w在每次出现时独立地为0或1,条件是q或w至少有一次出现时是1。

64. 根据权利要求63所述的组合物,其中 $R^1$ 或 $R^2$ 的多核苷酸具有第一序列, $L^{10}$ 具有第二序列,第一序列与第二序列的至少第一部分具有至少90%的互补性。

65. 根据权利要求44、48、51或64中任一项的组合物,其中第一序列与第一序列的部分具有至少92%的互补性。

66. 根据权利要求44、48、51、64或65中任一项的组合物,第一序列与第一序列的部分具有至少95%的互补性。

67. 根据权利要求44、48、51或64-66中任一项的组合物,第一序列与第一序列的部分具

有至少97%的互补性。

68. 根据权利要求44、48、51或64-67中任一项的组合物,第一序列与第一序列的部分具有至少98%的互补性。

69. 根据权利要求44、48、51或64-68中任一项的组合物,第一序列与第一序列的部分具有至少99%的互补性。

70. 根据权利要求44、48、51或64-69中任一项的组合物,其中第一序列、第二序列的部分或两者的长度范围为10个核苷酸至40个核苷酸。

71. 根据权利要求44、48、51或64-70中任一项的组合物,其中第一序列、第二序列的部分或两者的长度范围为10个核苷酸至30个核苷酸。

72. 根据权利要求44、48、51或64-71中任一项所述的组合物,其中第一序列、第二序列的部分或两者的长度范围为15个核苷酸至25个核苷酸。

73. 根据权利要求44、48、51或64-72中任一项所述的组合物,其中第二序列包含与第一部分相同的第二部分。

74. 根据权利要求44、48、51或64-73中任一项所述的组合物,还包含结合至具有第三序列的连接多核苷酸的靶向部分,该第三序列与第二序列的至少第三部分具有至少90%的互补性。

75. 根据权利要求1-74中任一项的化合物或组合物,其中q在每次出现是0。

76. 根据权利要求1-74中任一项所述的化合物或组合物,其中q在至少一次出现时是1。

77. 根据权利要求1-74或76中任一项的化合物或组合物,其中w每次出现是0。

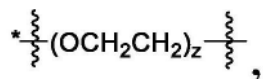
78. 根据权利要求1-77中任一项所述的化合物或组合物,其中w在至少一次出现时是1。

79. 根据权利要求1-78中任一项所述的化合物或组合物,其中至少一次出现的 $L^4$ 是亚杂烷基,或者其中每次出现的 $L^4$ 是亚杂烷基。

80. 根据权利要求79所述的化合物或组合物,其中所述杂亚烷基包括氧化烯烃。

81. 根据权利要求80所述的化合物或组合物,其中所述杂亚烷基包括氧化乙烯。

82. 根据权利要求79-81中任一项的化合物或组合物,其中 $L^4$ 具有以下结构:



其中:

z是1至100的整数;

\*表示与相邻磷原子的键。

83. 根据权利要求82所述的化合物或组合物,其中z是3至6的整数或22至26的整数。

84. 根据权利要求1-78中任一项所述的化合物或组合物,其中至少一次出现的 $L^4$ 是亚烷基,或者其中每次出现的 $L^4$ 是亚烷基。

85. 根据权利要求84所述的化合物或组合物,其中至少一个亚烷基是亚乙基,或其中每个亚烷基是亚乙基。

86. 根据权利要求1-85中任一项的化合物或组合物,其中至少一次出现的 $R^3$ 是H,或其中每次出现的 $R^3$ 是H。

87. 根据权利要求1-86中任一项所述的化合物或组合物,其中 $L^{1a}$ 在每次出现时独立地是任选取代的5-7元杂亚芳基连接基。

88. 根据权利要求1-86中任一项所述的化合物或组合物, 其中 $L^{1a}$ 具有以下结构之一:



89. 根据权利要求1-88中任一项所述的化合物或组合物, 其中至少一次出现的 $L^3$ 是亚烷基连接基, 或者其中每次出现的 $L^3$ 是亚烷基连接基。

90. 根据权利要求1-89中任一项所述的化合物或组合物, 其中 $L^2$ 和/或 $L^8$ 在至少一次出现时不存在, 或者其中 $L^2$ 和/或 $L^8$ 在每次出现是不存在。

91. 根据权利要求1-90中任一项所述的化合物或组合物, 其中至少一次出现的 $L^5$ 和/或 $L^6$ 是亚烷基, 或者其中每次出现的 $L^5$ 和/或 $L^6$ 是亚烷基。

92. 根据权利要求1-91中任一项所述的化合物或组合物, 其中 $L^{1b}$ 在每次出现时独立地包含酰胺官能团或三唑基官能团。

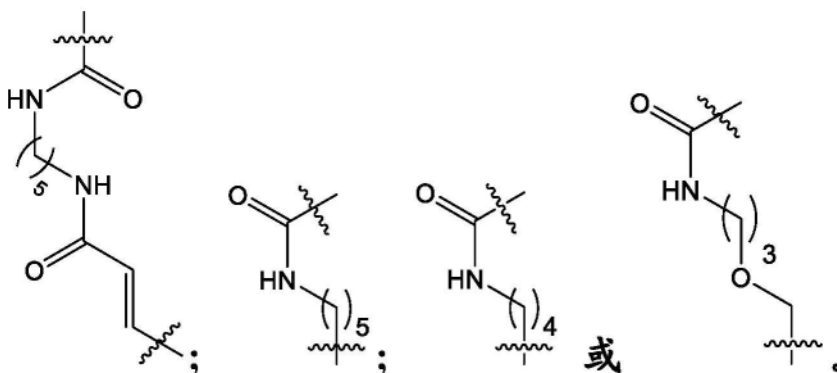
93. 根据权利要求1-92中任一项所述的化合物或组合物, 其中 $R^5$ 在每次出现时独立地是 $OH$ 、 $O^-$ 或 $OR_d$ 。

94. 根据权利要求1-93中任一项所述的化合物或组合物, 其中 $R^4$ 在每次出现时是氧代。

95. 根据权利要求1-94中任一项所述的化合物或组合物, 其中至少一次出现的 $L^7$ 是任取代的杂亚烷基连接基, 或者其中每次出现的 $L^7$ 独立地是任取代的杂亚烷基连接基。

96. 根据权利要求1-95中任一项所述的化合物或组合物, 其中 $L^7$ 包含酰胺或三唑基官能团。

97. 根据权利要求1-96中任一项所述的化合物或组合物, 其中每次出现的 $L^7$ 具有以下结构之一:



98. 根据权利要求1-96中任一项所述的化合物或组合物, 其中 $n$ 是1-100的整数, 或者其中 $n$ 是1-10的整数。

99. 根据权利要求1-98中任一项所述的化合物或组合物, 其中 $m$ 是3至6的整数, 或者其中 $m$ 是3。

100. 组合物, 包含:

捕获探针, 包含:

第一片段, 包含具有第一序列的第一多核苷酸; 和

与第一片段缀合的第二片段, 第二片段包含多个交替的第二多核苷酸和第一聚合物, 每个第一聚合物包含第一生色团, 每个第二多核苷酸包含第二序列; 和

多个可检测探针,所述多个可检测探针中的每个包含第三多核苷酸,第三多核苷酸共价结合至至少一个包含第二生色团的第二聚合物,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。

101.根据权利要求100所述的组合物,其中第一聚合物中的每个包含多个第一生色团。

102.根据权利要求100或101所述的组合物,其中每个可检测探针包含多个第二生色团。

103.根据权利要求100-102中任一项所述的组合物,其中,第一生色团是受体生色团。

104.根据权利要求100-102中任一项所述的组合物,其中第一生色团是供体生色团。

105.根据权利要求100-104中任一项所述的组合物,其中,第二生色团是受体生色团。

106.根据权利要求100-104中任一项所述的组合物,其中,第二生色团是供体生色团。

107.根据权利要求100-106中任一项所述的组合物,其中,第一生色团是荧光团。

108.根据权利要求100-107中任一项所述的组合物,其中,第二生色团是荧光团。

109.根据权利要求100-108中任一项所述的组合物,其中各自的第二聚合物共价结合至多个可检测探针中的至少一个可检测探针的第三多核苷酸的每一末端。

110.组合物,包含:

捕获探针,包含:

第一片段,包含具有第一序列的第一多核苷酸;和

与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸,每个第二多核苷酸包含第二序列;和

多个可检测探针,多个可检测探针中的每个包含具有第一末端和第二末端的第三多核苷酸,第一末端共价结合至包含受体生色团的第一聚合物,第二末端共价结合至包含供体生色团的第二聚合物,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。

111.根据权利要求100-110中任一项所述的组合物,其中所述捕获探针具有权利要求25-42中任一项所述的结构。

112.根据权利要求100-111中任一项所述的组合物,其中至少一个可检测探针独立地是权利要求1-24中任一项所述的化合物。

113.根据权利要求100-112中任一项的组合物,进一步包含共价结合至靶向部分的第四多核苷酸,第四多核苷酸包含具有与第一序列至少90%互补性的第四序列。

114.根据权利要求100-113中任一项的组合物,其中第四序列具有与第一序列至少92%的互补性。

115.根据权利要求100-114中任一项的组合物,其中第四序列具有与第一序列至少95%的互补性。

116.根据权利要求100-115中任一项的组合物,其中第四序列具有与第一序列至少96%的互补性。

117.根据权利要求100-116中任一项的组合物,其中第四序列具有与第一序列至少97%的互补性。

118.根据权利要求100-117中任一项的组合物,其中第四序列具有与第一序列至少98%的互补性。

119.根据权利要求100-118中任一项的组合物,其中第四序列具有与第一序列至少

99%的互补性。

120. 根据权利要求100-119中任一项的组合物,其中第二序列具有与第三序列至少92%的互补性。

121. 根据权利要求100-120中任一项的组合物,其中第二序列具有与第三序列至少95%的互补性。

122. 根据权利要求100-121中任一项的组合物,其中第二序列具有与第三序列至少96%的互补性。

123. 根据权利要求100-122中任一项的组合物,其中第二序列具有与第三序列至少97%的互补性。

124. 根据权利要求100-123中任一项的组合物,其中第二序列具有与第三序列至少98%的互补性。

125. 根据权利要求100-124中任一项的组合物,其中第二序列具有与第三序列至少99%的互补性。

126. 根据权利要求110-125中任一项所述的组合物,其中所述受体生色团是荧光团。

127. 根据权利要求110-126中任一项所述的组合物,其中所述供体生色团是荧光团。

128. 根据权利要求100-127中任一项的组合物,其中所述捕获探针包含靶向部分。

129. 根据权利要求128所述的组合物,其中所述靶向部分是抗体。

130. 试剂盒,其包含权利要求1-129中任一项的化合物或组合物以及使用说明书。

131. 用于鉴定靶标分析物的存在的方法,包括:

通过在测定条件下使样品与权利要求43-130中任一项的组合物接触来产生混合物;和在检测条件下对所述混合物进行成像。

132. 根据权利要求131所述的方法,其中使所述混合物成像包括:

在第一波长激发供体荧光团;和

在第二波长检测受体荧光团的发射。

## 聚合物染料

[0001] 关于序列表的声明

[0002] 与本申请相关的序列表代替纸质副本以文本格式提供,并特此通过引用并入说明书中。包含序列表的文本文件的名称是870268\_430WO\_SEQUENCE\_LISTING.txt。文本文件为4.1KB,创建于2021年11月19日,正在通过EFS-Web以电子方式提交。

### 背景技术

#### 技术领域

[0003] 本公开一般涉及包含与至少一个多核苷酸共价结合的聚合物生色团的化合物(例如,包含与核苷酸探针结合的聚合物荧光团部分的化合物),以及包含其的组合物和试剂盒,以及它们的方法制备和在各种分析方法中的用途。

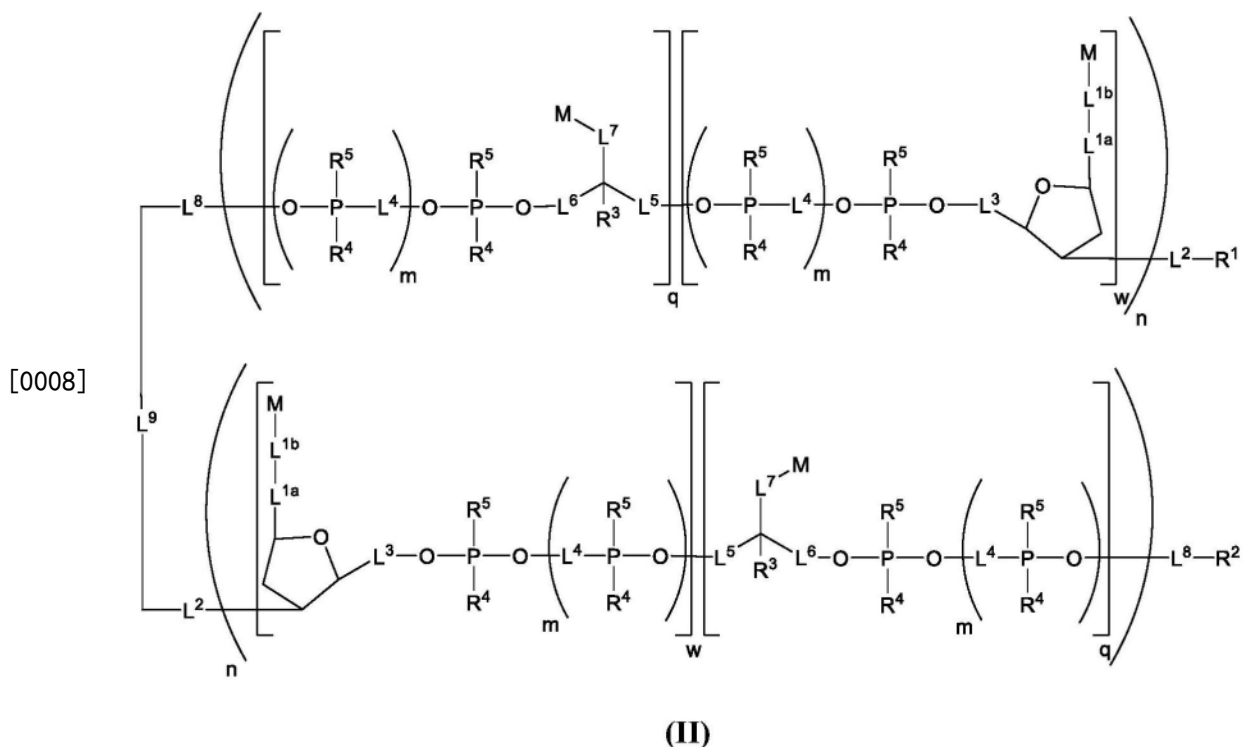
#### [0004] 相关技术的描述

[0005] 探针用于在各种诊断和分析环境中检测特定靶标序列或其他分析环境。例如,常规的异质杂交测定通常包括以下步骤:靶标核酸的固定化(例如,在纸、珠或塑料表面上);添加与靶标序列互补的标记探针;杂交;去除未杂交的探针;和检测仍然与固定靶标结合的探针。然而,使用固体表面来固定靶标核酸会通过限制探针的移动性或接近靶标而延长杂交所需的时间。此外,固体表面可能会干扰来自探头的信号或导致信号中出现噪声。分离探针-靶杂交物的要求也妨碍体内检测和合成反应期间核酸的同时检测(实时检测)。

[0006] 因此本领域需要改进的探针,其具有增加的亮度并产生较低的信噪比。本公开满足了这种需要并提供了进一步的相关优势。

### 发明内容

[0007] 在一个方面中,本文描述的是具有以下结构(II)的化合物:



[0009] 或其立体异构体、盐或互变异构体,其中:

[0010] M在每次出现时独立地是生色团;

[0011]  $L^{1a}$ 在每次出现时独立地是杂亚芳基连接基;

[0012]  $L^2$ 和 $L^8$ 是独立任选的连接基;

[0013]  $L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

[0014]  $L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

[0015]  $L^9$ 是包含多核苷酸的连接基;

[0016]  $R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基或 $-OP(=R_a)(R_b)R_c$ ;

[0017]  $R^3$ 在每次出现时独立地为H、烷基或烷氧基;

[0018]  $R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

[0019]  $R^5$ 在每次出现时独立地为氧代、硫代或不存在;

[0020]  $R_a$ 为O或S;

[0021]  $R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

[0022]  $R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;

[0023]  $R_d$ 是抗衡离子;

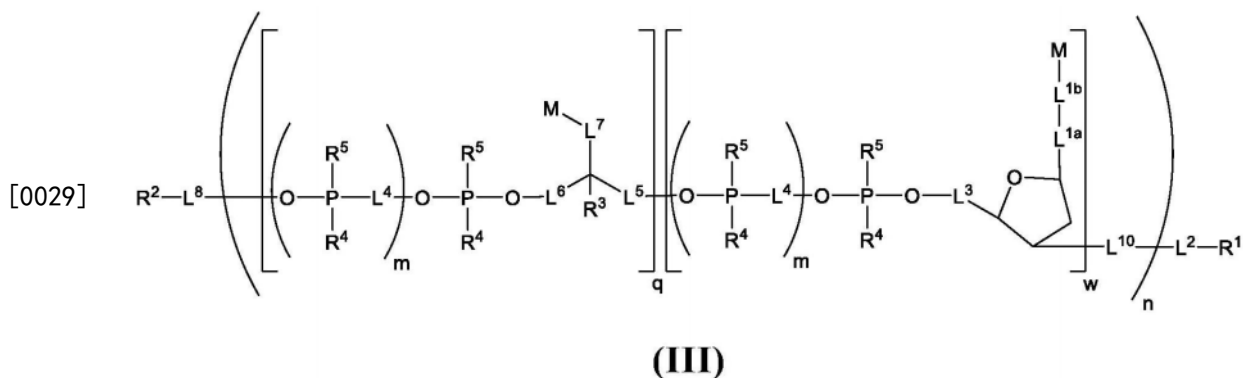
[0024]  $L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;

[0025] m在每次出现时独立地为0或更大的整数;

[0026] n是1或更大的整数;

[0027] q和w在每次出现时独立地为0或1,条件是q或w至少有一次出现时是1。

[0028] 另一方面,本公开提供具有以下结构(III)的化合物:



[0030] 或其立体异构体、盐或互变异构体,其中:

[0031] M在每次出现时独立地为相同或不同的生色团;

[0032] L<sup>1a</sup>在每次出现时独立地是杂亚芳基连接基;

[0033] L<sup>2</sup>和L<sup>8</sup>是独立任选的连接基;

[0034] L<sup>1b</sup>、L<sup>3</sup>、L<sup>5</sup>、L<sup>6</sup>和L<sup>7</sup>在每次出现时是独立地任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

[0035] L<sup>4</sup>在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;

[0036] L<sup>10</sup>是包含多核苷酸的连接基;

[0037] R<sup>1</sup>和R<sup>2</sup>各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基、-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>或为包含多核苷酸的部分,条件是R<sup>1</sup>和R<sup>2</sup>中的至少一个为包含多核苷酸的部分;

[0038] R<sup>3</sup>在每次出现时独立地为H、烷基或烷氧基;

[0039] R<sup>4</sup>在每次出现时独立地为OH、SH、O<sup>-</sup>、S<sup>-</sup>、OR<sub>d</sub>或SR<sub>d</sub>;

[0040] R<sup>5</sup>在每次出现时独立地为氧代、硫代或不存在;

[0041] R<sub>a</sub>为O或S;

[0042] R<sub>b</sub>为OH、SH、O<sup>-</sup>、S<sup>-</sup>、OR<sub>d</sub>或SR<sub>d</sub>;

[0043] R<sub>c</sub>为OH、SH、O<sup>-</sup>、S<sup>-</sup>、OR<sub>d</sub>、OL'、SR<sub>d</sub>、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;

[0044] R<sub>d</sub>是抗衡离子;

[0045] L'在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;

[0046] m在每次出现时独立地为0或更大的整数;

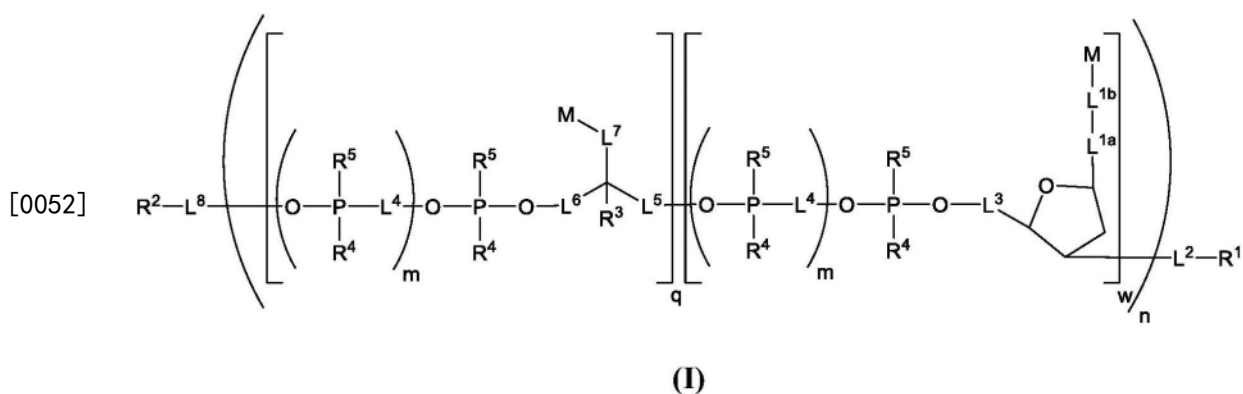
[0047] n是2或更大的整数;

[0048] q和w在每次出现时独立地为0或1,条件是q或w至少有一次出现时是1。

[0049] 另一方面,本文描述了包含结构(II)的化合物和捕获探针的组合物。

[0050] 本文还描述了包含结构(II)的化合物和结构(III)的化合物的组合物。

[0051] 在另一方面,本文描述的是包含结构(III)的化合物和具有以下结构(I)的化合物的组合物:



[0053] 或其立体异构体、盐或互变异构体,其中:

[0054] M在每次出现时独立地为相同或不同的生色团:

[0055] L<sup>1a</sup>,在每次出现时,独立地是杂亚芳基连接基;

[0056]  $L^2$ 和 $L^8$ 是独立任选的连接基:

[0057]  $L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立地任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基；

[0058]  $L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基；

[0059]  $R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基、 $-OP(=R_a)(R_b)R_c$ 或为包含多核苷酸的部分,条件是 $R^1$ 和 $R^2$ 中的至少一个为包含多核苷酸的部分;

[0060]  $R^3$ 在每次出现时独立地为H、烷基或烷氧基;

[0061]  $R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;

[0062] R<sup>5</sup>在每次出现时独立地为氧代、硫代或不存在;

[0063]  $R_a$  为0或S;

[0064]  $R_b$  为  $OH$ 、 $SH$ 、 $O^-$ 、 $S^-$ 、 $OR_d$  或  $SR_d$ ;

[0065]  $R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚；

[0066]  $R_d$ 是抗衡离子;

[0067] L'在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;m在每次出现时独立地为0或更大的整数;

[0068] n是1或更大的整数:

[0069]  $q$ 和 $w$ 在每次出现时独立地为0或1,条件是 $q$ 或 $w$ 至少有一次出现时是1。

[0070] 本文还描述了一种组合物,其包含:(1)捕获探针,其包含:(i)第一片段,其包含具有第一序列的第一多核苷酸;(ii)与第一片段缀合的第二片段,第二片段包含多个交替的第二多核苷酸和第一聚合物,每个第一聚合物包含第一生色团,每个第二多核苷酸包含第二序列;(2)多个可检测探针,多个可检测探针中的每个都包含第三多核苷酸,所述第三多核苷酸共价结合至至少一个包含第二生色团的第二聚合物,第三多核苷酸具有第三序列,该第三序列具有与第二序列至少90%的互补性。

[0071] 在另一方面,本文描述的是一种组合物,其包含:(1)捕获探针,其包含:(i) 第一片

段,其包含具有第一序列的第一多核苷酸;(ii)与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸,每个第二多核苷酸包含第二序列;(2)多个可检测探针,多个可检测探针中的每个包含具有第一末端和第二末端的第三多核苷酸,第一末端共价结合至包含受体生色团的第一聚合物,第二末端共价结合结合到包含供体生色团的第二聚合物,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。

[0072] 另外,本文描述的是包含本文描述的化合物或组合物的试剂盒。

[0073] 在又一个实施方案中,本公开提供了一种用于鉴定靶标分析物的存在的方法,包括:通过在测定条件下使样品与本文所述的组合物接触来产生混合物;并在检测条件下对混合物进行成像

[0074] 附图的简要说明

[0075] 本专利或申请文件包含至少一幅彩色绘图。带有彩色附图的本专利或专利申请公开的副本将在请求和支付必要费用后由专利局提供。在附图中,相同的参考数字鉴定相似的要素。要素在图中的尺寸和相对位置不一定按比例画出和这些要素中的某些被扩大和定位以改善图的可辨认性。此外,画出的要素具体形状并不期望传递有关该特定要素的真实形状的任何信息,而是仅仅为了便于在图中识别而选用。

[0076] 图1显示了探针复合物的一个实施方案的示意图。

[0077] 图2A显示了探针复合物的另一个实施方案的示意图。

[0078] 图2B显示了探针复合物的另一个实施方案的图示。

[0079] 图2C显示了探针复合物的另一个实施方案的图示。

[0080] 图3A显示了探针复合物的另一个实施方案的图示。

[0081] 图3B显示了探针复合物的另一个实施方案的图示。

[0082] 图4显示了实施例3中描述的测试结果。

[0083] 图5显示了实施例5中描述的测试结果。

[0084] 图6显示了实施例6中描述的测试结果。

[0085] 图7说明了如实施例7中所述的探针串联的一个实施方案。

[0086] 图8A显示了实施例8中描述的测试结果。

[0087] 图8B显示了实施例8中描述的测试结果。

[0088] 发明详述

[0089] 在下文描述中,描述某些特定细节以提供对本发明各种实施方式的透彻理解。然而,本领域技术人员将理解本发明可以不用这些细节来实施。

[0090] 除非上下文另有需要,在整个说明书和权利要求中,措辞“包含”及其变型比如“含有”和“包括”解释为开放内含意义,也即“包括,但不限于”。

[0091] 在本说明书通篇提及“一种实施方式”或“一个实施方式”意指的是,在实施方式中描述的特定性质、结构或特征包括在至少一个本公开实施方式中。从而,短语“在一种实施方式中”或“在一个实施方式中”在本说明书通篇各种位置出现不一定全部是指同一个的实施方式。另外,特定的性质、结构或特征可以在一个或多个实施方式中以任何适宜方式组合。

[0092] “氨基”是指-NH<sub>2</sub>基团。

[0093] “羧基”是指-CO<sub>2</sub>H基团。

[0094] “氰基”是指-CN基团。

[0095] “甲酰基”是指-C(=O)H基团。

[0096] “羟基”或“羟基”是指-OH基团。

[0097] “亚氨基”是指=NH基团。

[0098] “硝基”是指-NO<sub>2</sub>基团。

[0099] “氧代”是指=O取代基基团。

[0100] “氢硫基”是指-SH基团。

[0101] “硫代”是指=S基团。

[0102] “烷基”是指仅由碳和氢原子组成的直链或支化的烃链基团,不含不饱和度,具有一至十二个碳原子(C<sub>1</sub>-C<sub>12</sub>烷基)、一至八个碳原子(C<sub>1</sub>-C<sub>8</sub>烷基)或一至六个碳原子(C<sub>1</sub>-C<sub>6</sub>烷基),并且其通过单键连接至分子其余部分,例如甲基、乙基、正丙基、1-甲基乙基(异丙基)、正丁基、正戊基、1,1-二甲基乙基(叔丁基)、3-甲基己基、2-甲基己基等。除非特别在说明书中另有说明,烷基是任选经取代的。

[0103] “亚烷基”或“亚烷基链”是指将分子其余部分连接至残基的直链或支化的二价烃链,仅由碳和氢组成,不含不饱和度,并且具有一至十二个碳原子,例如亚甲基、亚乙基、亚丙基、正-亚丁基、亚乙烯基、亚丙烯基、正-亚丁烯基、亚丙炔基、正-亚丁炔基等。亚烷基链通过单键连接至分子其余部分并且通过单键连接至残基。亚烷基链与分子其余部分的连接点和与残基的连接点能够通过链中的一个碳或任何两个碳。除非说明书中特别另有说明,亚烷基是任选经取代的。

[0104] “亚烯基”或“亚烯基链”是指将分子其余部分连接至残基的直链或支化的二价烃链,仅由碳和氢组成,含有至少一个碳-碳双键和具有二至十二个碳原子,例如亚乙烯基、亚丙烯基、正-亚丁烯基等。亚烯基链通过单键连接至分子其余部分和通过双键或单键连接至残基。亚烯基链与分子其余部分的连接点和与残基的连接点能够通过链中的一个碳或任何两个碳。除非说明书中特别另有说明,亚烯基是任选经取代的。

[0105] “亚炔基”或“亚炔基链”是指将分子其余部分连接至残基的直链或支化的二价烃链,仅由碳和氢组成,含有至少一个碳-碳三键和具有二至十二个碳原子,例如亚乙烯基,亚丙烯基,正-亚丁烯基等。亚炔基链通过单键连接至分子其余部分和通过双键或单键连接至残基。亚炔基链与分子其余部分的连接点和与残基的连接点能够通过链中的一个碳或任何两个碳。除非说明书中特别另有说明,亚炔基是任选经取代的。

[0106] “烷基醚”是指如前文所定义的任何烷基,其中至少一个碳-碳键用碳-氧键替换。碳-氧键可以在末端(如烷氧基)或碳氧键可以在内部(也即C-O-C)。烷基醚包括至少一个碳氧键,但是可以包括多于一个。例如,聚乙二醇(PEG)包括烷基醚的含义当中。除非说明书中特别另有说明,烷基醚基团是任选经取代的。例如,在某些实施方式中烷基醚用醇或-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>取代,其中R<sub>a</sub>,R<sub>b</sub>和R<sub>c</sub>各自如对结构(I)化合物的定义。

[0107] “烷氧基”是指式-OR<sub>a</sub>基团,其中R<sub>a</sub>是如前文所定义的含有一至十二个碳原子的烷基。除非说明书中特别另有说明,烷氧基是任选经取代的。

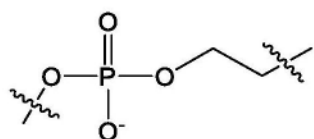
[0108] “烷氧基烷基醚”是指通式为-OR<sub>a</sub>R<sub>b</sub>的基团,其中R<sub>a</sub>是如上所定义的含有1至12个碳原子的亚烷基,R<sub>b</sub>是如本文所定义的烷基醚基团。除非说明书中另有具体说明,否则烷氧基烷基醚基团任选被取代,例如被醇或-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>取代,其中R<sub>a</sub>、R<sub>b</sub>和R<sub>c</sub>中的每个如结构

(I)的化合物所定义。

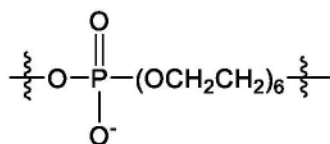
[0109] “杂烷基”是指如上所定义的烷基,其在烷基内或烷基末端包含至少一个杂原子(例如,N、O、P或S)。在一些实施方案中,杂原子在烷基内(即,杂烷基包含至少一个碳-[杂原子]<sub>x</sub>-碳键,其中x是1、2或3)。在其他实施方案中,杂原子位于烷基的末端,因此用于将烷基连接到分子的其余部分(例如,M<sub>1</sub>-H-A),其中M<sub>1</sub>是分子的一部分,H是杂原子,并且A是烷基)。除非说明书中另有具体说明,否则杂烷基任选地被取代。示例性的杂烷基包括氧化乙烯(例如,聚氧化乙烯),任选地包括磷-氧键,例如磷酸二酯键。

[0110] “杂烷氧基”是指式-OR<sub>a</sub>的基团,其中R<sub>a</sub>是如上所定义的包含一到十二个碳原子的杂烷基。除非说明书中另有具体说明,否则杂烷氧基任选地被取代。

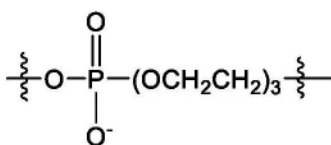
[0111] “杂亚烷基”是指如上所定义的亚烷基,其在亚烷基链内或亚烷基链的末端包含至少一个杂原子(例如,N、O、P或S)。在一些实施方案中,杂原子在亚烷基链内(即,杂亚烷基包含至少一个碳-[杂原子]-碳键,其中x是1、2或3)。在其他实施方案中,杂原子位于亚烷基的末端,因此用于将亚烷基连接到分子的其余部分(例如,M<sub>1</sub>-H-A-M<sub>2</sub>,其中M<sub>1</sub>和M<sub>2</sub>是分子的一部分,H是杂原子并且A是亚烷基)。除非在说明书中另有具体说明,杂亚烷基任选地被取代。示例性杂亚烷基包括氧化乙烯(例如,聚氧化乙烯)和“C”、“HEG”、“TEG”、“PEG 1K”及其变体,连接基如下所示:



“C 连接基”

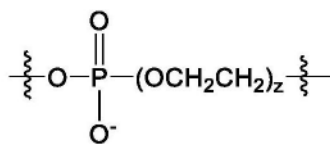


“HEG 连接基”



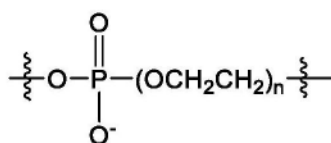
[0112]

“TEG 连接基”



$z > 0$ ; e.g.,  $z = 3-10$

“EG 连接基”

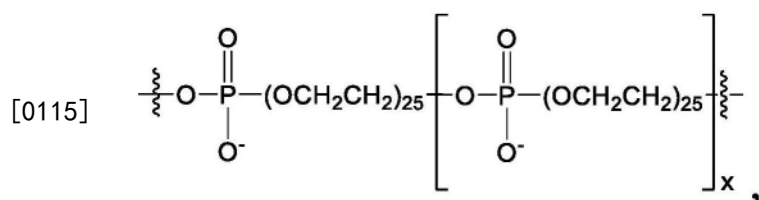


$n = 22-26$

“PEG 1K 连接基”

[0113] 上述C-连接基、HEG连接基的多聚体和/或PEG 1K连接基包括在杂亚烷基连接基的各种实施方案中。

[0114] 在PEG 1K连接基的一些实施方案中,n是25。多聚体可包含例如以下结构:



[0116] 其中x是0或大于0的整数,例如,x在0-100范围内(例如,1、2、3、4、5、6、7、8、9或10)。

[0117] “杂亚烯基”是如前文所定义的杂亚烷基,包含至少一个碳-碳双键。除非说明书中特别另有说明,杂亚烯基是任选经取代的。

[0118] “杂亚炔基”是包含至少一个碳-碳三键的杂亚烷基。除非说明书中特别另有说明,杂亚炔基是任选经取代的。

[0119] 在提及“杂原子连接体”时“杂原子”是指由一个或多个杂原子组成的连接体基团。示范性杂原子连接体包括单个原子选自O、N、P和S,和多个杂原子例如具有式-P(O<sup>-</sup>)(=O)O<sup>-</sup>或-OP(O<sup>-</sup>)(=O)O<sup>-</sup>的连接体及其多聚体和组合。

[0120] “磷酸盐/酯”是指-OP(=O)(R<sub>a</sub>)R<sub>b</sub>基团,其中R<sub>a</sub>是OH, O<sup>-</sup>或OR<sub>c</sub>;而R<sub>b</sub>是OH, O<sup>-</sup>, OR<sub>c</sub>, 硫代磷酸盐/酯基团或其它磷酸盐/酯基团,其中R<sub>c</sub>是平衡离子(例如Na<sup>+</sup>等)。

[0121] “磷酸基烷基”是指-OP(=O)(R<sub>a</sub>)R<sub>b</sub>基团,其中R<sub>a</sub>是OH, O<sup>-</sup>或OR<sub>c</sub>;而R<sub>b</sub>是-O烷基,其中R<sub>c</sub>是平衡离子(例如Na<sup>+</sup>等)。除非说明书中特别另有说明,磷酸基烷基是任选经取代的。例如,在某些实施方式中,磷酸基烷基中的-O烷基部分任选用一个或多个羟基,氨基,氢硫基,磷酸盐/酯,硫代磷酸盐/酯,磷酸基烷基,硫代磷酸基烷基,磷酸基烷基醚、硫代磷酸基烷基醚或-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>取代,其中每个R<sub>a</sub>, R<sub>b</sub>和R<sub>c</sub>如对于结构(I)的化合物所定义。

[0122] “磷酸基烷基醚”是指-OP(=O)(R<sub>a</sub>)R<sub>b</sub>基团,其中R<sub>a</sub>是OH, O<sup>-</sup>或OR<sub>c</sub>;而R<sub>b</sub>是-O烷基醚,其中R<sub>c</sub>是平衡离子(例如Na<sup>+</sup>等)。除非说明书中特别另有说明,磷酸基烷基醚基团是任选经取代的。例如,在某些实施方式中,磷酸基烷基醚基团中的-O烷基醚部分任选用一个或多个羟基,氨基,氢硫基,磷酸盐/酯,硫代磷酸盐/酯,磷酸基烷基,硫代磷酸基烷基,磷酸基烷基醚或硫代磷酸基烷基醚或-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>取代,其中每个R<sub>a</sub>, R<sub>b</sub>和R<sub>c</sub>如对于结构(I)的化合物所定义。

[0123] “硫代磷酸盐/酯”是指-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>基团,其中R<sub>a</sub>是O或S, R<sub>b</sub>是OH, O<sup>-</sup>, S<sup>-</sup>, OR<sub>d</sub>或SR<sub>d</sub>;而R<sub>c</sub>是OH, SH, O<sup>-</sup>, S<sup>-</sup>, OR<sub>d</sub>, SR<sub>d</sub>, 磷酸盐/酯基团或其它硫代磷酸盐/酯基团,其中R<sub>d</sub>是平衡离子(例如Na<sup>+</sup>等)且条件是:i) R<sub>a</sub>是S; ii) R<sub>b</sub>是S<sup>-</sup>或SR<sub>d</sub>; iii) R<sub>c</sub>是SH, S<sup>-</sup>或SR<sub>d</sub>;或iv) i), ii)和/或iii)的组合。

[0124] “硫代磷酸基烷基”是指-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>基团,其中R<sub>a</sub>是O或S, R<sub>b</sub>是OH, O<sup>-</sup>, S<sup>-</sup>, OR<sub>d</sub>或SR<sub>d</sub>;而R<sub>c</sub>是-O烷基,其中R<sub>d</sub>是平衡离子(例如Na<sup>+</sup>等)且条件是:i) R<sub>a</sub>是S; ii) R<sub>b</sub>是S<sup>-</sup>或SR<sub>d</sub>;或iii) R<sub>a</sub>是S而R<sub>b</sub>是S<sup>-</sup>或SR<sub>d</sub>。除非说明书中特别另有说明,硫代磷酸基烷基是任选经取代的。例如,在某些实施方式中,硫代磷酸基烷基中的-O烷基部分任选用一个或多个羟基,氨基,氢硫基,磷酸盐/酯,硫代磷酸盐/酯,磷酸基烷基,硫代磷酸基烷基,磷酸基烷基醚或硫代磷酸基烷基醚或-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>取代,其中每个R<sub>a</sub>, R<sub>b</sub>和R<sub>c</sub>如对于结构(I)的化合物所定义。

[0125] “硫代磷酸基烷基醚”是指-OP(=R<sub>a</sub>)(R<sub>b</sub>)R<sub>c</sub>基团,其中R<sub>a</sub>是O或S, R<sub>b</sub>是OH, O<sup>-</sup>, S<sup>-</sup>, OR<sub>d</sub>或SR<sub>d</sub>;而R<sub>c</sub>是-O烷基醚,其中R<sub>d</sub>是平衡离子(例如Na<sup>+</sup>等)且条件是:i) R<sub>a</sub>是S; ii) R<sub>b</sub>是S<sup>-</sup>或SR<sub>d</sub>;

或iii)  $R_a$  是S而 $R_b$ 是 $S^-$ 或 $SR_d$ 。除非说明书中特别另有说明,硫代磷酸基烷基醚基团是任选经取代的。例如,在某些实施方式中,硫代磷酸基烷基中的-O烷基醚部分任选用一个或多个羟基,氨基,氢硫基,磷酸盐/酯,硫代磷酸盐/酯,磷酸基烷基,硫代磷酸基烷基,磷酸基烷基醚或硫代磷酸基烷基醚或 $-OP(=R_a)(R_b)R_c$ ,取代,其中每个 $R_a$ ,  $R_b$ 和 $R_c$ 如对于结构(I)的化合物所定义。

[0126] “碳环”是指稳定的3-至18-元芳族或非芳族环,其包含3至18个碳原子。除非说明书中特别另有说明,碳环可以是单环、双环、三环或四环的环系,其可以包括稠合的或桥连的环系,和可以是部分或完全饱和的。非芳族碳环基残基包括环烷基,而芳族碳环基残基包括芳基。除非说明书中特别另有说明,碳环基团是任选经取代的。

[0127] “环烷基”是指稳定的非芳族单环或多环碳环,其可以包括稠合的或桥连的环系,具有三至十五个碳原子,优选具有三至十个碳原子,并且其是饱和的或不饱和的和通过单键连接至分子其余部分。单环环烷基包括例如环丙基,环丁基,环戊基,环己基,环庚基和环辛基。多环环烷基包括例如金刚烷基,降莰烷基,十氢萘基,7,7-二甲基-二环-[2.2.1]庚烷基等。除非说明书中特别另有说明,环烷基是任选经取代的。

[0128] “芳基”是指环系包含至少一个碳环芳族环。在某些实施方式中,芳基包含6至18个碳原子。芳基环可以是单环、双环、三环或四环的环系,其可以包括稠合的或桥连的环系。芳基包括例如衍生自下述的芳基:醋葱,茈,醋菲烯,葱,萸,苯, **蒽**, 荧葱, 茈, 不对称引达省, 对称引达省, 茈满, 茈, 萘, 非那烯, 菲, 七曜烯, 茈和三亚苯。除非说明书中特别另有说明,芳基是任选经取代的。

[0129] “杂环”是指稳定的3-至18-元芳族或非芳族环,包含一至十二个碳原子和一至六个选自氮、氧和硫的杂原子。除非说明书中特别另有说明,杂环可以是单环、双环、三环或四环的环系,其可以包括稠合的或桥连的环系;和杂环中的氮、碳或硫原子可以任选被氧化;氮原子可以任选被季铵化;和杂环可以是部分或完全饱和的。芳族杂环的实例是在下文中列于杂芳基的定义中(也即杂芳基是杂环的子集)。非芳族杂环的实例包括例如二氧杂环戊烷基,噻吩基[1,3]二噻烷基,十氢异喹啉基,咪唑啉基,咪唑烷基,异噻唑烷基,异噻唑烷基,吗啉基,八氢吲哚基,八氢异吲哚基,2-氧代哌嗪基,2-氧代哌啶基,2-氧代吡咯烷基,噁唑烷基,哌啶基,哌嗪基,4-哌啶酮基,吡咯烷基,吡唑烷基,吡唑并嘧啶基,奎宁环基,噻唑烷基,四氢呋喃基,三噻烷基,三硫杂环己烷基,三噻烷基,四氢吡喃基,硫吗啉基,硫杂吗啉基,1-氧代-硫吗啉基和1,1-二氧代-硫吗啉基。除非说明书中特别另有说明,杂环基团是任选经取代的。

[0130] “杂芳基”是指5-至14-元环系统包含一至十三个碳原子,一至六个选自氮、氧和硫的杂原子,和至少一个芳族环。出于本发明的某些实施方式的意图,杂芳基残基可以是单环、双环、三环或四环的环系,其可以包括稠合的或桥连的环系;和杂芳基残基中的氮、碳或硫原子可以任选被氧化;氮原子可以任选被季铵化。实例包括例如氮杂**萘**基,吡啶基,苯并咪唑基,苯并噻唑基,苯并吲哚基,苯并二氧杂环戊二烯基,苯并呋喃基,苯并噁唑基,苯并噻唑基,苯并噻二唑基,苯并[b][1,4]二氧杂**萘**基,1,4-苯并二噻烷基,苯并萘并呋喃基,苯并噁唑基,苯并二氧杂环戊二烯基,苯并二氧杂环己二烯基,苯并吡喃基,苯并吡喃酮基,苯并呋喃基,苯并呋喃酮基,苯并噻吩基(苯并噻吩基),苯并三唑基,苯并[4,6]咪唑并[1,



[0135] “着色”是指分子吸收着色光谱范围内的光(也即红、黄、蓝色等)。

[0136] “FRET”是指**Förster**共振能量转移,其是指能量从一部分(例如第一生色团或“供体”)的激发转移至相邻部分(例如第二生色团或“受体”)的物理相互作用。“FRET”有时也与荧光共振能量转移(即在各生色团是荧光部分的情况下)可互换地使用。一般地,FRET需要(1)受体生色团的激发或吸收光谱与供体生色团的发射光谱重叠;(2)受体和供体生色团的跃迁偶极矩实质上平行(也即在约 $0^\circ$ 或 $180^\circ$ );和(3)受体和供体生色团具有空间邻近性(也即相互接近)。能量从供体向受体转移通过无辐射的偶极-偶极耦合发生,并且供体生色团与受体生色团之间的距离一般远小于光的波长。

[0137] “供体”或“供体生色团”是指生色团(例如荧光团),其被或能够被诱导为激发电子态并且可以通过长程偶极-偶极相互作用以无辐射方式将其激发能量转移至附近的受体生色团。不受理论限制,据信能量转移发生的原因是各自生色团的振荡偶极具有相似的共振频率。具有这些相似共振频率的供体和受体称为“供体-受体对”,其与“FRET部分”或“FRET染料”等可互换地使用。

[0138] “受体”或“受体生色团”是指来自供体生色团的激发能量经由无辐射转移通过长程偶极-偶极相互作用向其转移的生色团(例如荧光团)。

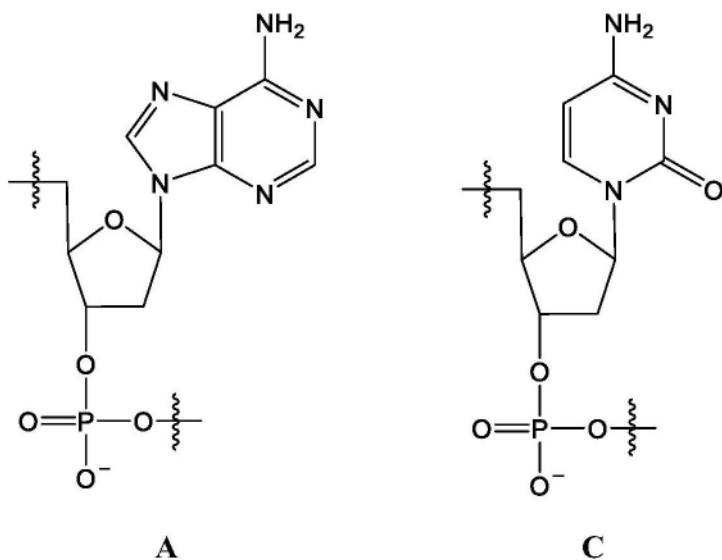
[0139] “斯托克斯位移”是指电子跃迁(例如从激发态至非激发态或反之)吸收谱图与发射谱图之间的带最大值的位置(例如波长)差异。在某些实施方式中,化合物具有大于25nm,大于30nm,大于35nm,大于40nm,大于45nm,大于50nm,大于55nm,大于60nm,大于65nm,大于70nm,大于75nm,大于80nm,大于85nm,大于90nm,大于95nm,大于100nm,大于110nm,大于120nm,大于130nm,大于140nm,大于150nm,大于160nm,大于170nm,大于180nm,大于190nm,或大于200nm的斯托克斯位移。

[0140] “连接基”是指至少一个原子比如碳、氧、氮、硫、磷及其组合的连续链,其将分子的一部分与同一分子的又一部分连接或者与不同分子、部分或固体载体(例如微粒)连接。连接基可以经由共价键或其它手段比如离子键或氢键相互作用连接分子。

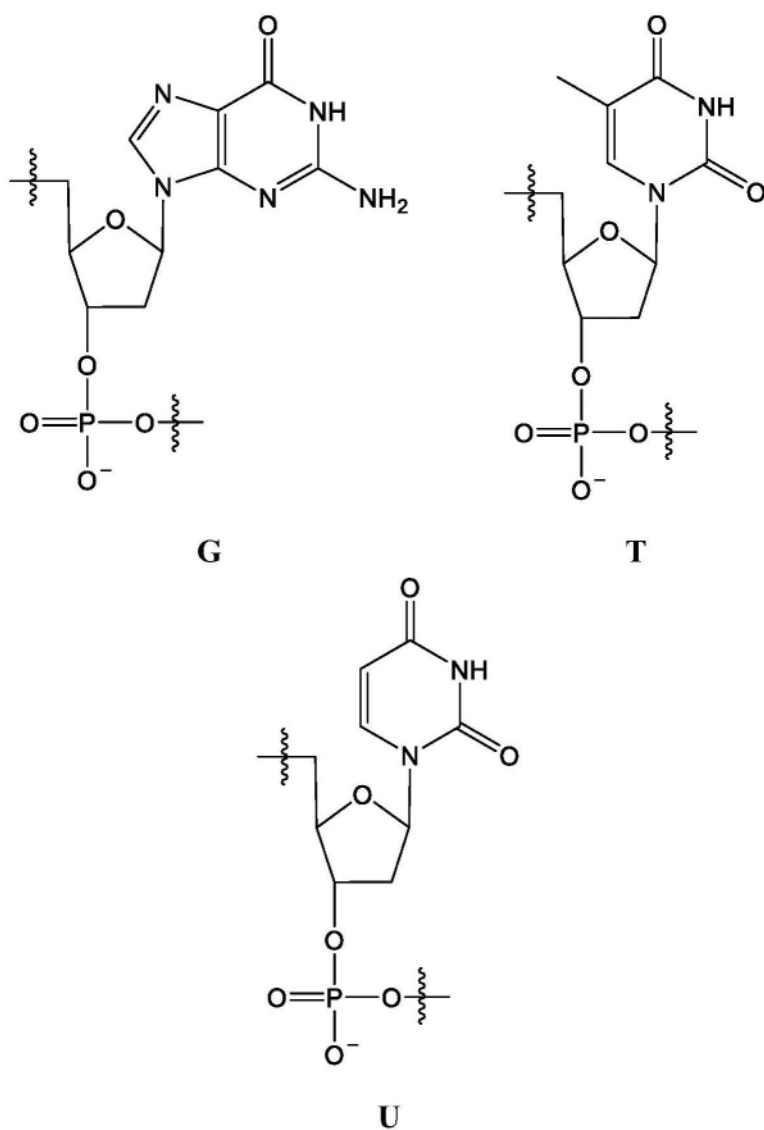
[0141] 如本文所用,“核酸”或“核酸分子”或“多核苷酸”是指任何脱氧核糖核酸(DNA)或核糖核酸(RNA),包括寡核苷酸。

[0142] 核酸可以代表编码链或其互补链。核酸可以由天然存在的核苷酸的单体(例如脱氧核糖核苷酸和核糖核苷酸)、天然存在的核苷酸的类似物(例如天然存在的核苷酸的 $\alpha$ -对映体形式)或两者的组合组成。核酸序列中的核苷酸根据标准IUPAC惯例命名。具体来说,“A”为腺嘌呤,“C”为胞嘧啶,“G”为鸟嘌呤,“T”为胸腺嘧啶,“U”为尿嘧啶,分别指以下结构:

[0143]



[0144]



[0145] 多核苷酸的序列是指核苷酸在多核苷酸中排列的顺序。

[0146] 天然存在的核苷酸的类似物,也称为修饰的核苷酸,可以对糖部分或嘧啶或嘌呤碱基部分进行修饰或替换。核酸单体可以通过磷酸二酯键或此类连接的类似物连接。磷酸

二酯键的类似物包括硫代磷酸酯、二硫代磷酸酯、硒代磷酸酯、二硒代磷酸酯、苯胺基硫代磷酸酯(phosphoroanilothioate)、苯胺基磷酸酯(phosphoranilidate)、氨基磷酸酯等。在各种实施方案中,使用修饰的核苷酸间连接。修饰的核苷酸间连接是本领域众所周知的并且包括甲基磷酸酯、硫代磷酸酯、二硫代磷酸酯、亚磷酰胺和磷酸酯连接。核酸分子可以是单链或双链的。

[0147] 天然存在的核苷酸的类似物还包括锁核酸(LNA),也称为桥接核酸(BNA)。LNA包含一个2'-氧、4'-碳亚甲基桥,其“锁定”3'-内构象,从而限制呋喃核糖环的柔韧性。

[0148] 本文所用的术语“杂交”是指第一条核酸链通过碱基配对与第二条核酸链结合的任何过程。杂交可以发生在完全互补的核酸链之间或包含次要错配区域的“基本上互补的”核酸链之间。尽管为了便于理解本文一般参照双链体结构(即,两个核酸的结构)讨论杂交,但是本领域的普通技术人员将理解下面的实施方案还涵盖其中所描述的核酸形成三链体结构的构型(即三种核酸的结构)。例如,两个探针可以与单个靶序列杂交,或者两个可检测探针可以与捕获探针的相同区域杂交。

[0149] 如本文所用,如果在足够严格的条件下使用时探针主要仅与靶核酸杂交,则探针是对靶序列“特异性”的多核苷酸。通常,如果探针-靶双链体(或三链体)稳定性大于探针与样品中发现的任何其他序列之间形成的双链体(或三链体)的稳定性,则探针对靶序列是特异性的。本领域技术人员将认识到各种因素,例如盐条件以及探针的碱基组成和错配的位置,将影响探针特异性,并且在大多数情况下需要对探针特异性进行常规实验确认。可以选择杂交条件,在该条件下探针可以仅与靶序列形成稳定的双链体(或三链体)。因此,在适当严格的条件下使用靶标特异性探针能够特异性扩增那些包含靶标探针结合位点的靶标序列。序列特异性条件的使用使得探针能够特异性结合到包含完全互补的探针结合位点的靶序列。

[0150] 只有完全互补的核酸链才能杂交的条件被称为“严格杂交条件”或“序列特异性杂交条件”。如本文所用的术语“严格”是指本领域通常理解的杂交条件以定义杂交程序的条件。严格条件可以是低的、高的或中等的,因为这些术语是本领域公知的并且为普通技术人员所熟知。在各种实施方案中,严格条件可以包括例如高度严格条件和/或中等严格(即,中等严格)条件。

[0151] 可以在不太严格的杂交条件下获得基本互补序列的稳定双链体(或三链体);可以通过适当调整杂交条件来控制可容忍的错配程度。核酸技术领域的技术人员可以根据经验确定双链体(或三链体)稳定性,考虑许多变量,包括例如寡核苷酸的长度和碱基对组成、离子强度和错配碱基对的发生率,遵循本领域提供的指导。

[0152] 如本文所用,“互补”是指核酸分子可以通过传统的Watson-Crick碱基配对或其他非传统类型的配对(例如Hoogsteen或反向Hoogsteen氢结合)与另一个核酸分子在互补核苷或核苷酸之间形成氢键。

[0153] 本领域理解核酸分子不需要与靶核苷酸序列100%互补才能特异性杂交。也就是说,两个或更多个核酸分子可能不完全互补,并由核酸分子中可与第二个核酸分子形成氢键的连续残基的百分比表示。例如,如果第一个核酸分子有10个核苷酸,第二个核酸分子有10个核苷酸,则第一个和第二个核酸分子之间的5、6、7、8、9或10个核苷酸的碱基配对5代表分别50%、60%、70%、80%、90%和100%的互补性。“完全(“Perfect ly”)”或“完全(“ful

ly”)”互补的核酸分子是指其中第一个核酸分子的所有连续残基将与第二个核酸分子中相同数量的连续残基形成氢键的核酸分子,其中核酸分子要么都具有相同数量的核苷酸(即具有相同的长度)或两个分子具有不同的长度。

[0154] 本文所用的术语“杂交复合物”是指两个核苷酸序列之间由于互补的G和C碱基之间以及互补的A和T碱基之间形成氢键而形成的复合物;这些氢键可以通过碱基堆积相互作用进一步稳定。两个互补的核苷酸序列以反平行构型形成氢键。杂交复合物可以在溶液形成(例如,Cot或Rot分析)或在溶液中存在的一个核苷酸序列与固定在固体载体(例如,纸、膜、滤器、芯片、针或载玻片,或已固定细胞和/或核酸的任何其他合适的底物)的另一个核苷酸序列之间形成。

[0155] “反应性基团”是能够与第二反应性基团(例如,“互补反应性基团”)反应以形成一个或多个共价键的部分,例如通过取代、氧化、还原、加成或环加成反应。示范性反应性基团包括例如亲核物质,亲电试剂,

[0156] 双烯,亲双烯体,醛,肟,肟,炔,胺,叠氮化物,酰基叠氮化物,

[0157] 酰卤,腈,硝酮,氢硫基,二硫化物,磺酰卤,异硫氰酸酯,亚氨酸酯,活化的酯,酮, $\alpha,\beta$ -不饱和羰基,烯烃,马来酰亚胺, $\alpha$ -卤代酰亚胺,环氧化物,氮丙啶,四嗪,四唑,磷,生物素,硫杂环丙烷等。

[0158] 术语“可视的”和“视觉可检测的”在本文中用于指通过视觉检查可以观察的物质,而不用预先光照或者化学或酶活化。所述视觉可检测的物质吸收和发出约300至约900nm光谱范围的光。优选,所述物质是强着色的,优选具有至少约40,000,更优选至少约50,000,还更优选至少约60,000,又还更优选至少约70,000,和最优选至少约80,000M<sup>-1</sup>cm<sup>-1</sup>的摩尔消光系数。本发明化合物可以检测如下:肉眼观察,或者借助基于光学的检测装置包括但不限于吸收分光光度计、透射光学显微镜、数字照相机和扫描仪。视觉可检测的物质并不局限于发出和/或吸收可见光谱光的那些。发出和/或吸收紫外(UV)区域的光(约10nm至约400nm)、红外(IR)区域的光(约700nm至约1mm)的物质,和发出和/或吸收其它区域的电磁光谱的物质也包括在“视觉可检测的”物质的范围内。

[0159] 出于本发明实施方式的意图,术语“光稳定的可视染料”是指化学部分,其如上文所定义是视觉可检测的,并且在暴露于光的情况下不显著变化或分解。优选,光稳定的可视染料在暴露于光至少1小时之后不展示显著漂白或分解。更优选,可视染料在暴露于光至少12小时、还更优选至少24小时、又还更优选至少1周和最优选至少1个月之后是稳定的。适用于本发明化合物和方法的光稳定的可视染料的非限制性实例包括偶氮染料,硫靛蓝染料,喹吖啶酮色素,二噁嗪,酞菁,紫环酮(perinone),二酮基吡咯并吡咯,喹酞酮,和truarycarbonium。

[0160] 如本文所用,术语“花衍生物”期望包括视觉可检测的任何取代的花。然而,该术语并不期望包括花本身。术语“蒽衍生物”,“蔡衍生物”和“茈衍生物”类似地使用。在某些优选实施方式中,衍生物(例如花、茈、蒽或蔡衍生物)是花、蒽、蔡或茈的酰亚胺、双酰亚胺或肟酰亚胺衍生物。

[0161] 本发明各种实施方式的视觉可检测分子用于各式各样的分析应用比如生物化学和生物医药应用,其中需要确定特定分析物(例如生物分子)的存在、位置或量。因此,在又一方面本发明提供视觉可检测生物分子的方法,包括:(a)提供具有视觉可检测的生物分子

的生物学系统,所述分子包含与生物分子连接的结构(I)的化合物;和(b)通过其可视特性检测生物分子。出于本发明意图,短语“通过其可视特性检测所述生物分子”意指,在不用光照或化学或酶活化的情况下,用肉眼或者借助基于光学的检测装置包括但不限于吸收分光光度计、透射光学显微镜、数字照相机和扫描仪来观察生物分子。密度计可以用来定量存在的视觉可检测的生物分子的量。例如,生物分子在两个样品中的相对量能够通过测量相对光密度来测量。如果已知染料分子/生物分子的化学计量,且已知染料分子的消光系数,那么还能够从光密度的测量确定生物分子的绝对浓度。如本文所用,术语“生物学系统”用于指除了视觉可检测的生物分子之外还包含一种或多种生物分子的任何溶液或混合物。所述生物学系统的实例包括细胞,细胞提取物,组织样品,电泳凝胶,测试混合物,和杂交反应混合物。

[0162] “固体载体”是指本领域对于分子固相载体已知的任何固体基底,例如“微粒”是指用于连接至本发明化合物的任何各种小颗粒,包括例如玻璃珠、磁珠、聚合物珠、非聚合物珠等。在某些实施方式中,微粒包含聚苯乙烯珠。

[0163] “固相载体残基”指的是当分子从固相载体上裂解时仍然附着在分子上的官能团。固体载体残基是本领域已知的并且可以基于固体载体的结构和将分子与其连接的基团容易地推导出来。

[0164] 如本文所用,词“靶标”是指待检测的靶标分析物。“靶向部分”是选择性结合或缔合特定靶标例如分析物分子的部分。“选择性”结合或缔合是指靶向部分相对于其他靶标优先与所需靶标缔合或结合。在一些实施方案中,本文公开的化合物包括与靶向部分的连接,用于选择性结合或缔合化合物与感兴趣的分析物(即,靶向部分的靶标)的目的,从而允许检测分析物。示例性的靶向部分包括例如抗体、抗原、核酸序列、酶、蛋白质、细胞表面受体拮抗剂等。在一些实施方案中,靶向部分是诸如抗体的部分,其选择性地与细胞上或细胞内的靶标特征结合或缔合,例如细胞膜或其他细胞结构上的靶标特征,从而允许检测感兴趣的细胞。在某些实施方案中,选择性结合或缔合所需分析物的小分子也被视为靶向部分。本领域技术人员将理解将在各种实施方案中有用的其他分析物和相应的靶向部分。

[0165] “碱基配对部分”是指能够与互补杂环部分经由氢键(例如Watson-Crick碱基配对)杂交的杂环部分。碱基配对部分包括天然和非天然碱基。碱基配对部分的非限制性实例是RNA和DNA碱基比如腺苷,鸟苷,胸苷,胞嘧啶和尿苷及其类似物。

[0166] 如本文所用,“捕获探针”是结合(例如,共价结合或杂交)到固体载体或靶向部分的单链DNA分子。通常,单链DNA分子包含与可检测探针或分支连接基的核苷酸序列互补的重复核苷酸序列,使得多个可检测探针或分支连接基能够与单个捕获探针杂交。

[0167] “可检测探针”包含与至少一个生色团结合的核苷酸序列。

[0168] 如本文所用,“分支连接基”是指聚合物主链,多个多核苷酸结合至该聚合物主链,使得多核苷酸延伸远离聚合物主链。例如,支链连接基可以是与多个DNA分子缀合的聚炔烃。多个多核苷酸包含至少一个具有第一序列的多核苷酸。第一序列与捕获探针的重复核苷酸序列互补。此外,多个多核苷酸包含多个具有第二序列的多核苷酸,该第二序列与可检测探针的核苷酸序列互补。

[0169] 本公开的本发明的实施方式也意在涵盖同位素标记的全部结构(I)、(II)或(III)化合物,其中一个或多个原子用具有不同的原子质量或质量数的原子替换。能够掺入公开

化合物的同位素的实例包括氢,碳,氮,氧,磷,氟,氯和碘的同位素,分别比如 $2\text{H}$ , $3\text{H}$ , $11\text{C}$ , $13\text{C}$ , $14\text{C}$ , $13\text{N}$ , $15\text{N}$ , $15\text{O}$ , $17\text{O}$ , $18\text{O}$ , $31\text{P}$ , $32\text{P}$ , $35\text{S}$ , $18\text{F}$ , $36\text{Cl}$ , $123\text{I}$ 和 $125\text{I}$ 。

[0170] 同位素标记的结构(I)或(II)化合物能够一般地通过本领域技术人员已知的常规技术或通过与下文和下述实例描述的那些类似的过程用适当同位素标记的试剂代替原先使用的非标记试剂来制备。

[0171] “稳定化合物”和“稳定结构”意指化合物,其是足够稳健的从而经历从反应混合物分离为有用程度的纯度,和配制为有效的治疗剂还存在。

[0172] “任选的”或“任选地”意指随后描述的事件或环境可以或可以不发生,并且该描述包括其中所述事件或环境发生的情况和其并不发生的情况。例如,“任选经取代的烷基”意指烷基可以或可以不被取代并且该描述包括取代的烷基和不具有取代的烷基。

[0173] “盐”包括酸和碱加成盐。

[0174] “酸加成盐”是指与无机酸形成的那些盐,所述无机酸是比如盐酸,氢溴酸,硫酸,硝酸,磷酸等,和与有机酸形成的那些盐,所述有机酸是比如乙酸,2,2-二氯乙酸,己二酸,藻酸,抗坏血酸,天冬氨酸,苯磺酸,苯甲酸,4-乙酰氨基苯甲酸,樟脑酸,樟脑-10-磺酸,癸酸,己酸,辛酸,碳酸,肉桂酸,柠檬酸,环拉酸,十二烷基硫酸,乙烷-1,2-二磺酸,乙磺酸,2-羟基乙磺酸,甲酸,富马酸,粘酸,龙胆酸,葡庚糖酸,葡萄糖酸,葡萄糖醛酸,谷氨酸,戊二酸,2-氧代-戊二酸,甘油磷酸,羟基乙酸,马尿酸,异丁酸,乳酸,乳糖酸,月桂酸,马来酸,苹果酸,丙二酸,扁桃酸,甲磺酸,粘酸,萘-1,5-二磺酸,萘-2-磺酸,1-羟基-2-萘甲酸,烟酸,油酸,乳清酸,草酸,棕榈酸,双羟萘酸,丙酸,焦谷氨酸,丙酮酸,水杨酸,4-对氨水杨酸,癸二酸,硬脂酸,丁二酸,酒石酸,硫氰酸,对-甲苯磺酸,三氟乙酸,十一烯酸等。

[0175] “碱加成盐”是指将无机碱或有机碱加至游离酸制备的那些盐。衍生自无机碱的盐包括例如钠,钾,锂,铵,钙,镁,铁,锌,铜,锰,铝盐等。衍生自有机碱的盐包括例如下述的盐:伯、仲和叔胺,取代的胺包括天然取代的胺,环状胺和碱性离子交换树脂,比如氨,异丙胺,三甲胺,二乙胺,三乙胺,三丙胺,二乙醇胺,乙醇胺,地阿诺,2-二甲基氨基乙醇,2-二乙基氨基乙醇,二环己基胺,赖氨酸,精氨酸,组氨酸,咖啡因,普鲁卡因,海巴明(hydrabamine),胆碱,内铵盐,苄乙苄胺, $\text{N,N}'$ -二苄基乙二胺,乙二胺,氨基葡萄糖,甲基葡萄糖胺,可可碱,三乙醇胺,氨丁三醇,嘌呤类,哌嗪,哌啶, $\text{N}$ -乙基哌啶,多元胺树脂等。特别优选的有机碱是异丙胺,二乙胺,乙醇胺,三甲胺,二环己基胺,胆碱和咖啡因。

[0176] 结晶可以产生本文描述的化合物的溶剂化物。本发明的实施方式包括所描述化合物的全部溶剂化物。如本文所用,术语“溶剂化物”是指聚集体,其包含一个或多个本发明化合物分子与一个或多个溶剂分子。溶剂可以是水,在该情况下溶剂化物可以是水合物。另选地,溶剂可以是有机溶剂。从而,本发明化合物可以作为水合物存在,包括一水合物,二水合物,半水合物,倍半水合物,三水合物,四水合物等,以及相应的溶剂化形式。本发明化合物可以是真正的溶剂化物,而在另外的情况下本发明化合物可以仅保留外来水或其它溶剂或者是水和一些外来溶剂的混合物。

[0177] 本发明化合物(例如结构I、II或III化合物)或它们的盐、互变异构体或溶剂化物的实施方式可以含有一个或多个不对称中心和可以从而产生对映体、非对映体和其它立体异构形式,其在绝对立体化学方面可以定义为(R)-或(S)-或者对于氨基酸的(D)-或(L)-。本公开的实施方式意在包括全部所述可能异构体,以及它们的外消旋和光学纯形式。光学

活性的 (+) 和 (-)、(R) - 和 (S) - 或 (D) - 和 (L) - 异构体可以用手性合成子或手性试剂制备, 或用常规技术例如色谱法和分步结晶拆分。用来制备/分离单独对映体的常规技术包括从适宜的光学纯前体手性合成或者用例如手性高压液体色谱法 (HPLC) 拆分外消旋体 (或盐或衍生物的外消旋体)。在本文描述的化合物含有烯式双键或其它几何不对称中心的情况下, 并且除非另有指定, 期望的是化合物包括 E 和 Z 几何异构体。类似地, 也期望包括全部互变异构体形式。

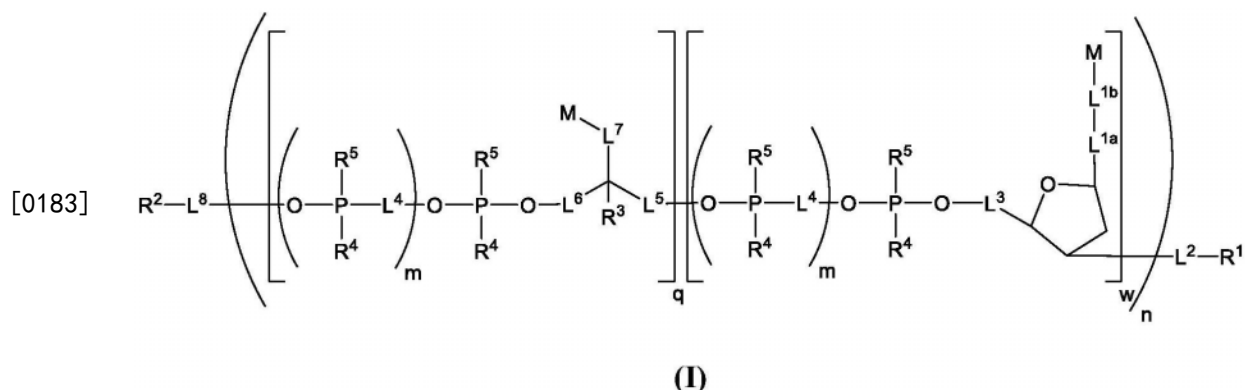
[0178] “立体异构体”是指由通过相同键键合的相同原子构成的化合物, 但是具有不可互换的不同三维结构。本发明预期各种立体异构体及其混合物和包括“对映体”, 其是指两个立体异构体, 其分子是彼此不可重叠的镜像。

[0179] “互变异构体”是指质子从分子的一个原子转移到相同分子的又一个原子。本发明包括任何所述化合物的互变异构体。化合物的各种互变异构体形式可通过本领域普通技术人员容易地衍生。

[0180] 本文使用的化学命名方案和结构图解是 I.U.P.A.C. 命名系统的修饰形式, 采用 ACD/Name 版本 9.07 软件程序和/或 ChemDraw Ultra 版本 11.0 软件命名程序 (CambridgeSoft)。也使用本领域普通技术人员熟悉的通用名。

[0181] 如上所述, 在本发明的一个实施方案中, 提供了在各种分析方法中用作探针的化合物。一般而言, 本申请的实施方案涉及可用作捕获或可检测探针的化合物, 以及包含其的组合物。

[0182] 在一些实施方案中, 本公开的化合物具有以下结构 (I) :

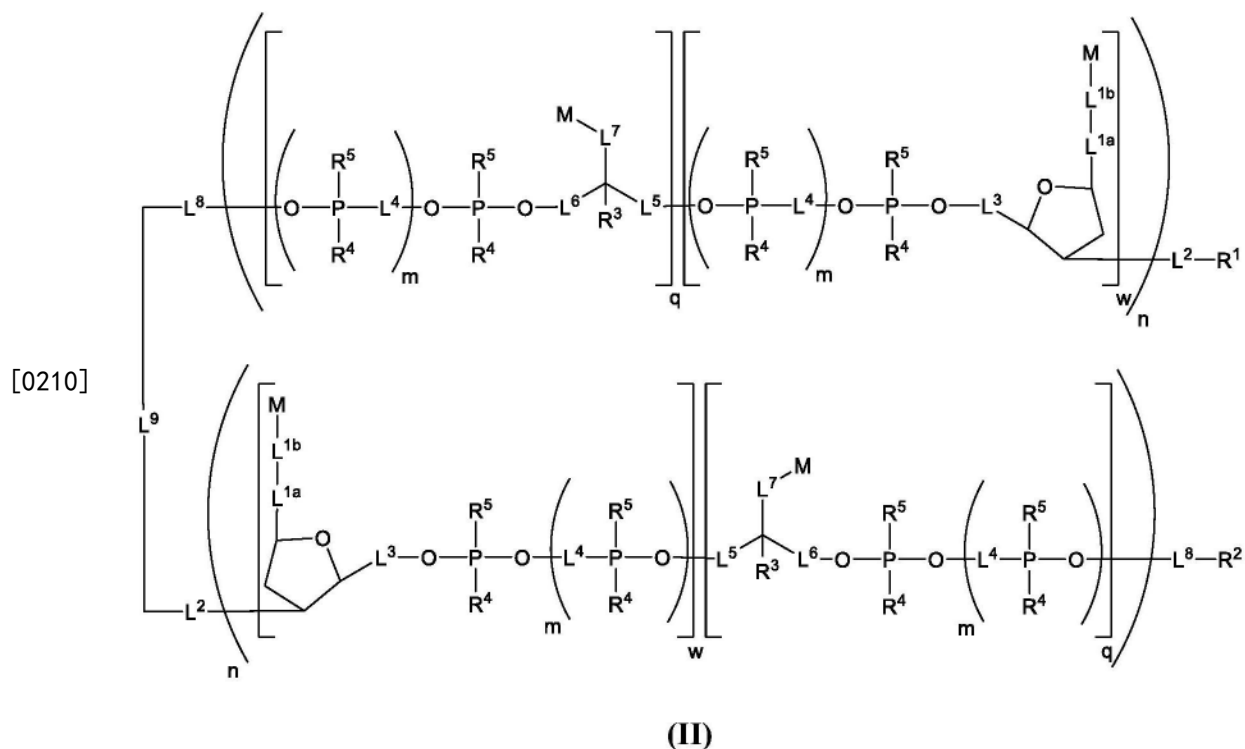


- [0192]  $R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ；
- [0193]  $R^5$ 在每次出现时独立地为氧代、硫代或不存在；
- [0194]  $R_a$ 为O或S；
- [0195]  $R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ；
- [0196]  $R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚；
- [0197]  $R_d$ 是抗衡离子；
- [0198]  $L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基；
- [0199]  $m$ 在每次出现时独立地为0或更大的整数；
- [0200]  $n$ 是1或更大的整数；
- [0201]  $q$ 和 $w$ 在每次出现时独立地为0或1，条件是 $q$ 或 $w$ 至少有一次出现时是1。
- [0202] 在一些实施方案中， $M$ 在每次出现时独立地是相同或不同的供体生色团。在其他实施方案中， $M$ 在每次出现时独立地是相同或不同的受体生色团。在某些实施方案中， $M$ 在每次出现时是相同的供体生色团。在其他实施方案中， $M$ 在每次出现时是相同的受体生色团。
- [0203] 在实施方案中， $M$ 在每次出现时独立地是相同或不同的荧光团。在一些实施方案中， $M$ 在每次出现时独立地为相同或不同的供体荧光团。在其他实施方案中， $M$ 在每次出现时独立地是相同或不同的受体荧光团。在某些实施方案中， $M$ 在每次出现时是相同的供体荧光团。在其他实施方案中， $M$ 在每次出现时是相同的受体荧光团。
- [0204] 在一些实施方案中， $R^1$ 和 $R^2$ 各自是包含多核苷酸的部分。在一些这样的实施方案中， $R^1$ 的多核苷酸与 $R^2$ 的多核苷酸互补，使得第一化合物的 $R^1$ 的多核苷酸能够与第二化合物的 $R^2$ 的多核苷酸杂交，等等。这方面的一个例子在图3A中示出。靶向部分301(在这种情况下，包覆的珠子)结合到第一多核苷酸320。探针305a、305b、305c、305d独立地是结构(I)的化合物。每个探针独立地包含布置在第二多核苷酸316a、316b、316c、316d和第三多核苷酸317a、317b、317c、317d之间的聚合物314a、314b、314c、314d。如图所示3A，探针305b的第二多核苷酸316b与探针305a的第三多核苷酸317a杂交。在一些实施方案中，三链体结构由三个探针的多核苷酸形成。
- [0205] 在实施方案中，第二序列具有与第三序列至少92%的互补性。在一些实施方案中，第二序列具有与第三序列至少95%的互补性。在一些实施方案中，第二序列具有与第三序列至少96%的互补性。在一些实施方案中，第二序列具有与第三序列至少97%的互补性。在一些实施方案中，第二序列具有与第三序列至少98%的互补性。在一些实施方案中，第二序列具有与第三序列至少99%的互补性。
- [0206] 另外，在一些实施方案中，多个探针包括以某种模式交替或重复的两个或更多个探针305、307(如图3B所示)。在一些这样的实施方案中，每个第一探针305独立地包含布置在第二多核苷酸316和第三多核苷酸317之间的聚合物314，并且每个第二探针307包含布置在第四多核苷酸319和第五多核苷酸321之间的聚合物314。第三多核苷酸317与第四多核苷酸319杂交，第五多核苷酸321与第二多核苷酸杂交。在一些实施方案中，三链体结构由三个探针的多核苷酸形成。
- [0207] 在一些实施方案中，每个第一聚合物包含多个生色团。在实施方案中，第二多核苷

酸、第三多核苷酸或两者的长度范围为10个核苷酸至40个核苷酸。在一些实施方案中，第二多核苷酸、第三多核苷酸或两者的长度范围为10个核苷酸至30个核苷酸。在一些实施方案中，第二多核苷酸、第三多核苷酸或两者的长度范围为15个核苷酸至25个核苷酸。

[0208] 第二多核苷酸之一与第一多核苷酸杂交。在各种实施方案中，第一多核苷酸包含具有与第二序列至少90%互补性的第一序列。在一些实施方案中，第一序列具有与第二序列至少92%的互补性。在一些实施方案中，第二序列具有与第一序列至少95%的互补性。在一些实施方案中，第二序列具有与第一序列至少96%的互补性。在一些实施方案中，第二序列具有与第一序列至少97%的互补性。在一些实施方案中，第二序列具有与第一序列至少98%的互补性。在一些实施方案中，第二序列具有与第一序列至少99%的互补性。在其他实施方案中，靶向部分共价结合至捕获探针。

[0209] 在其他实施方案中，本公开的化合物具有以下结构(II)：



[0211] 或其立体异构体、盐或互变异构体，其中：

[0212] M在每次出现时独立地是生色团；

[0213]  $L^{1a}$ ，在每次出现时，独立地是杂亚芳基连接基；

[0214]  $L^2$ 和 $L^8$ 是独立任选的连接基；

[0215]  $L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基；

[0216]  $L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基；

[0217]  $L^9$ 是包含多核苷酸的连接基；

[0218]  $R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基或 $-OP(=R_a)(R_b)R_c$ ；

[0219]  $R^3$ 在每次出现时独立地为H、烷基或烷氧基；

[0220]  $R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、ORd或SRd； $R^5$ 在每次出现时独立地为氧代、

硫代或不在于；

[0221]  $R_a$  为 O 或 S；

[0222]  $R_b$  为 OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$  或  $SR_d$ ；

[0223]  $R_c$  为 OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚；

[0224]  $R_d$  是抗衡离子；

[0225]  $L'$  在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基；

[0226]  $m$  在每次出现时独立地为 0 或更大的整数；

[0227]  $n$  是 1 或更大的整数；

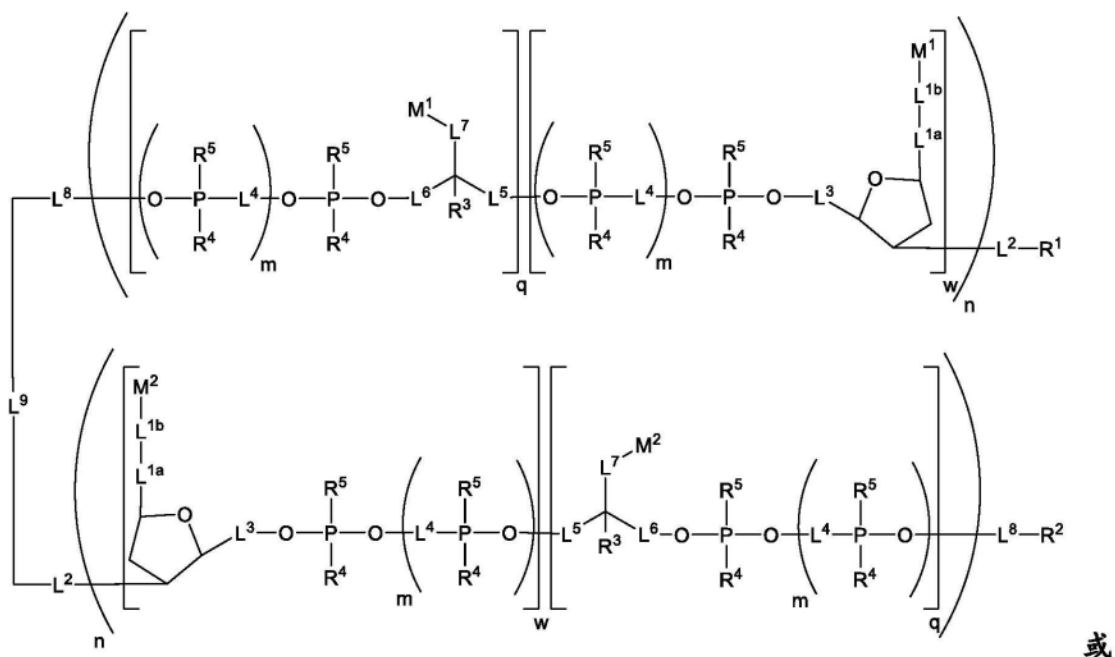
[0228]  $q$  和  $w$  在每次出现时独立地为 0 或 1，条件是  $q$  或  $w$  至少有一次出现时是 1。

[0229] 在一些实施方案中， $M$  在每次出现时独立地是相同或不同的供体生色团。

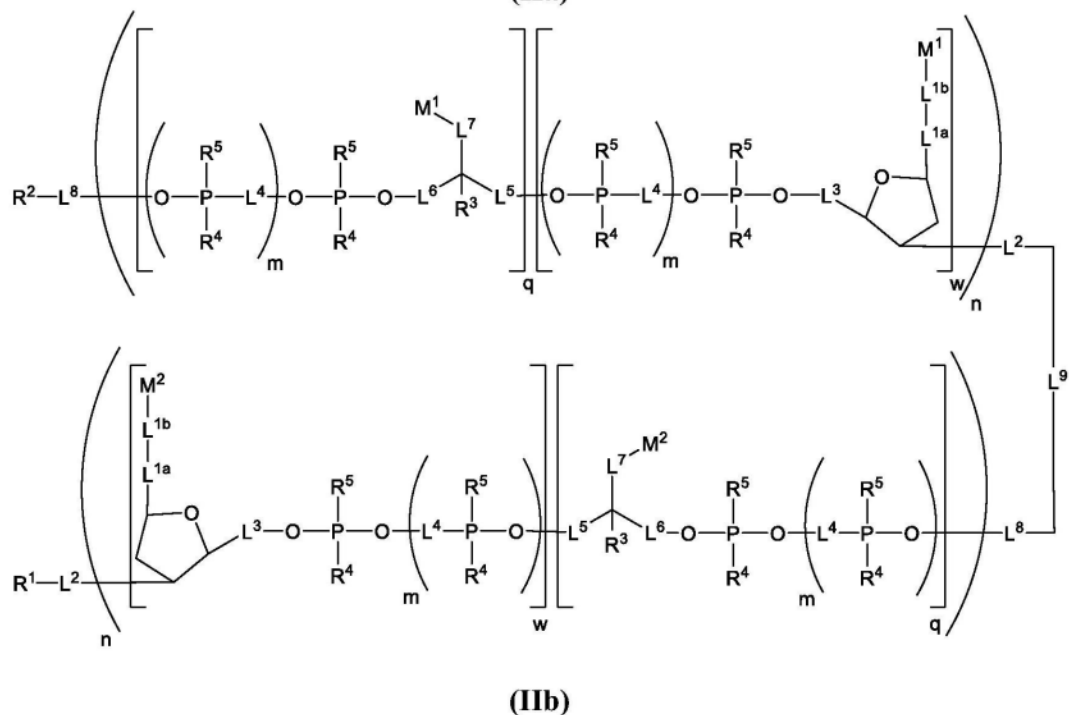
[0230] 在其他实施方案中， $M$  在每次出现时独立地是相同或不同的受体生色团。在某些实施方案中， $M$  在每次出现时是相同的供体生色团。在其他实施方案中， $M$  在每次出现时是相同的受体生色团。

[0231] 在实施方案中， $M$  在每次出现时独立地是相同或不同的荧光团。在一些实施方案中， $M$  在每次出现时独立地为相同或不同的供体荧光团。在其他实施方案中， $M$  在每次出现时独立地是相同或不同的受体荧光团。在某些实施方案中， $M$  在每次出现时是相同的供体荧光团。在其他实施方案中， $M$  在每次出现时是相同的受体荧光团。

[0232] 在一些实施方案中，本公开的化合物具有以下结构 (IIa) 或 (IIb)：



[0233]



[0234] 或其立体异构体、盐或互变异构体, 其中:

[0235]  $M^1$ 是在每次出现时独立地相同或不同的供体生色团;[0236]  $M^2$ 在每次出现时都是独立的相同或不同的受体生色团;[0237]  $L^{1a}$ , 在每次出现时, 独立地是杂亚芳基连接基;[0238]  $L^2$ 和 $L^8$ 是独立任选的连接基;[0239]  $L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;[0240]  $L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔

基连接基；

[0241]  $L^9$  是包含多核苷酸的连接基；

[0242]  $R^1$  和  $R^2$  各自独立地为 H、OH、SH、烷基、烷氧基、烷基醚、杂烷基或  $-OP(=R_a)(R_b)R_c$ ；

[0243]  $R^3$  在每次出现时独立地为 H、烷基或烷氧基；

[0244]  $R^4$  在每次出现时独立地为 OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$  或  $SR_d$ ；

[0245]  $R^5$  在每次出现时独立地为氧代、硫代或不存在；

[0246]  $R_a$  为 O 或 S；

[0247]  $R_b$  为 OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$  或  $SR_d$ ；

[0248]  $R_c$  为 OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚；

[0249]  $R_d$  是抗衡离子；

[0250]  $L'$  在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基；

[0251]  $m$  在每次出现时独立地为 0 或更大的整数；

[0252]  $n$  是 1 或更大的整数；

[0253]  $q$  和  $w$  在每次出现时独立地为 0 或 1，条件是  $q$  或  $w$  至少有一次出现时是 1。

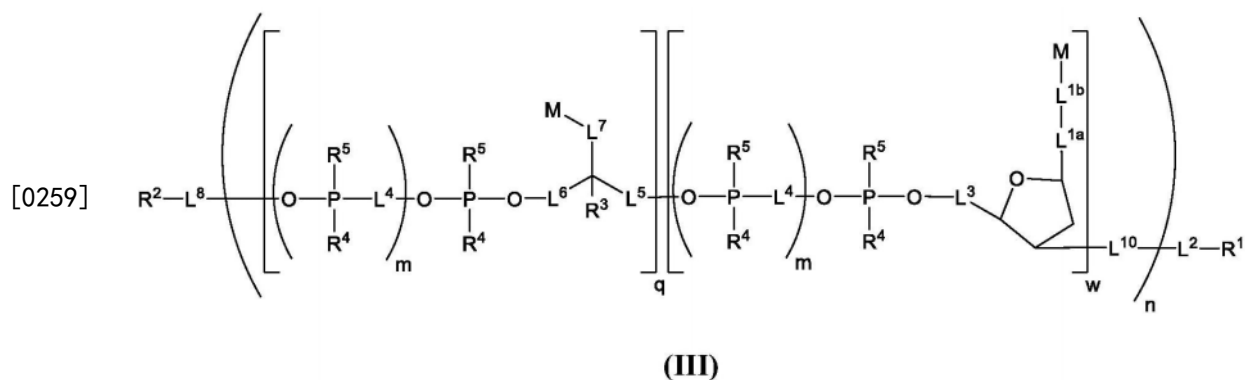
[0254] 在一些实施方案中，本公开的化合物具有结构 (IIa)。在一些实施方案中，本公开的化合物具有结构 (IIb)。

[0255] 在一些实施方案中， $M^1$  在每次出现时独立地是相同或不同的供体荧光团。在其他实施方案中， $M^2$  在每次出现时独立地是相同或不同的受体荧光团。在某些实施方案中， $M^1$  在每次出现时都是相同的供体荧光团。在其他实施方案中， $M^2$  在每次出现时是相同的受体荧光团。

[0256] 在实施方案中， $L^9$  的多核苷酸长度范围为 10 个核苷酸至 40 个核苷酸。在一些实施方案中， $L^9$  的多核苷酸的长度范围为 10 个核苷酸至 30 个核苷酸。在一些实施方案中， $L^9$  的多核苷酸的长度范围为从 15 个核苷酸到 25 个核苷酸。

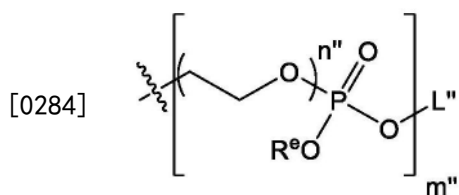
[0257] 在实施方案中，本文所述的具有结构 (I)、(II)、(IIa)、(IIb) 或其组合的化合物是可检测探针。在一些实施方案中，本文所述的具有结构 (I) 的化合物是可检测探针。在一些实施方案中，本文所述的具有结构 (II) 的化合物是可检测探针。在一些实施方案中，本文所述的具有结构 (IIa) 的化合物是可检测探针。在一些实施方案中，本文所述的具有结构 (IIb) 的化合物是可检测探针。

[0258] 因此，在一些实施方案中，本公开的化合物具有以下结构 (III)：



- [0260] 或其立体异构体、盐或互变异构体,其中:
- [0261] M在每次出现时独立地相同或不同的生色团;
- [0262]  $L^{1a}$ 在每次出现时独立地是杂亚芳基连接基;
- [0263]  $L^2$ 和 $L^8$ 是独立任选的连接基;
- [0264]  $L^{1b}$ 、 $L^3$ 、 $L^5$ 、 $L^6$ 和 $L^7$ 在每次出现时是独立地任选的亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;
- [0265]  $L^4$ 在每次出现时独立地为亚烷基、亚烯基、亚炔基、杂亚烷基、杂亚烯基或杂亚炔基连接基;
- [0266]  $L^{10}$ 是包含多核苷酸的连接基;
- [0267]  $R^1$ 和 $R^2$ 各自独立地为H、OH、SH、烷基、烷氧基、烷基醚、杂烷基、 $-OP(=R_a)(R_b)R_c$ 或为包含多核苷酸的部分,条件是 $R^1$ 和 $R^2$ 中的至少一个为包含多核苷酸的部分;
- [0268]  $R^3$ 在每次出现时独立地为H、烷基或烷氧基;
- [0269]  $R^4$ 在每次出现时独立地为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;
- [0270]  $R^5$ 在每次出现时独立地为氧代、硫代或不存在;
- [0271]  $R_a$ 为O或S;
- [0272]  $R_b$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 或 $SR_d$ ;
- [0273]  $R_c$ 为OH、SH、 $O^-$ 、 $S^-$ 、 $OR_d$ 、 $OL'$ 、 $SR_d$ 、烷基、烷氧基、杂烷基、杂烷氧基、烷基醚、烷氧基烷基醚、磷酸酯、硫代磷酸酯、磷烷基、硫代磷烷基、磷烷基醚或硫代磷烷基醚;
- [0274]  $R_d$ 是抗衡离子;
- [0275]  $L'$ 在每次出现时独立地是包含与固体载体的共价键的连接基、包含与固体载体残基的共价键的连接基或包含与核苷的共价键的连接基;
- [0276] m在每次出现时独立地为0或更大的整数;
- [0277] n是2或更大的整数;
- [0278] q和w在每次出现时独立地为0或1,条件是q或w至少有一次出现时是1。
- [0279] 在一些实施方案中,M在每次出现时独立地是相同或不同的供体生色团。在其他实施方案中,M在每次出现时独立地是相同或不同的受体生色团。在某些实施方案中,M在每次出现时是相同的供体生色团。在其他实施方案中,M在每次出现时是相同的受体生色团。
- [0280] 在实施方案中,M在每次出现时独立地是相同或不同的荧光团。在一些实施方案中,M在每次出现时独立地为相同或不同的供体荧光团。在其他实施方案中,M在每次出现时独立地是相同或不同的受体荧光团。在某些实施方案中,M是每次出现相同的供体荧光团。在其他实施方案中,M在每次出现时是相同的受体荧光团。
- [0281] 在实施方案中, $R^1$ 包含多核苷酸。在一些实施方案中, $R^2$ 包含多核苷酸。在具体实施方案中, $R^1$ 包含与多核苷酸结合的靶向部分。在具体实施方案中, $R^2$ 包含与多核苷酸结合的靶向部分。
- [0282] 在实施方案中,结构(III)的化合物是捕获探针。
- [0283] 结构(I)、(II)或(III)中任一结构中的多核苷酸,当位于化合物的末端部分时,可以终止于任何可接受的基团。例如,某些多核苷酸将终止于羟基或磷酸基团。在其他各种实施方案中,多核苷酸将终止于 $-OP(=R_a)(R_b)R_c$ ,其中 $R_c$ 为 $OL'$ 且 $R_a$ 和 $R_b$ 如上所定义。在那些实施例的一些中, $L'$ 是连接到固体载体、固体载体残基或核苷的杂亚烷基连接基。在一些实

施方案中, L' 包含氧化烯烃 (alkylene oxide) 或磷酸二酯部分或其组合。在某些实施方案中, L' 具有以下结构:



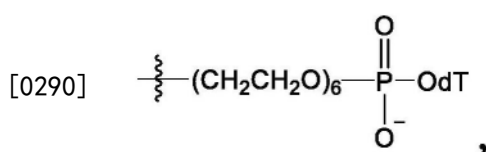
[0285] 其中:

[0286] m'' 和 n'' 独立地为 1 至 10 的整数;

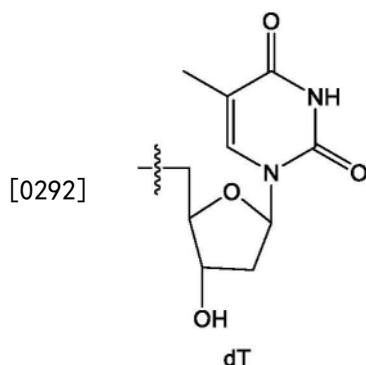
[0287] R<sup>e</sup> 为 H, 电子对或抗衡离子;

[0288] L'' 是 R<sup>e</sup> 或与固体载体、固体载体残基或核苷 (例如, 脱氧胸苷) 的直接键合或连接。

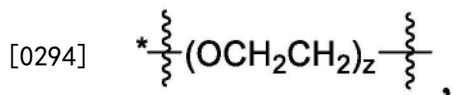
[0289] 在一些实施方案中, 多核苷酸终止于以下结构:



[0291] 其中 dT 是脱氧胸苷:



[0293] 在实施方案中, q 的每次出现都是 0。在实施方案中, q 在至少一次出现时是 1。在实施方案中, w 的每次出现都是 0。在实施方案中, w 在至少一次出现时是 1。在实施方案中, q 的每次出现都是 0。在实施方案中, 至少一次出现的 L<sup>4</sup> 是亚杂烷基, 或其中每次出现的 L<sup>4</sup> 是亚杂烷基。在实施方案中, 杂亚烷基包括氧化烯烃。在其他实施方案中, 杂亚烷基包括氧化乙烯。在一些实施方案中, L<sup>4</sup> 具有以下结构:

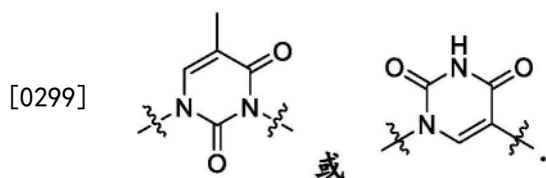


[0295] 其中:

[0296] z 为 1 至 100 的整数;

[0297] \* 表示与相邻磷原子的键。

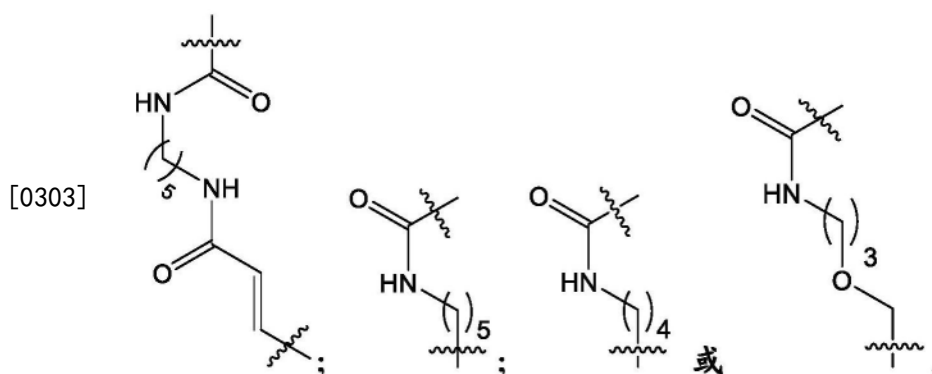
[0298] 在实施方案中, z 是 3 至 6 的整数或 22 至 26 的整数。在实施方案中, 至少一次出现的 L<sup>4</sup> 是亚烷基, 或其中每次出现的 L<sup>4</sup> 是亚烷基。在各种实施方案中, 至少一个亚烷基是亚乙基, 或其中每个亚烷基是亚乙基。在进一步的实施方案中, R<sup>3</sup> 在至少一次出现时是 H, 或其中 R<sup>3</sup> 的每次出现是 H。在一些实施方案中, L<sup>1a</sup> 在每次出现时独立地是任选取代的 5-7 元杂亚芳基连接基。在实施方案中, L<sup>1a</sup> 具有以下结构之一:



[0300] 在实施方案中,至少一次出现的 $L^3$ 是亚烷基连接基,或其中每次出现的 $L^3$ 是亚烷基连接基。在一些实施方案中, $L^2$ 和/或 $L^8$ 的至少一次出现不存在,或者其中 $L^2$ 和/或 $L^8$ 的每一次出现都不存在。在实施方案中, $L^5$ 和/或 $L^6$ 至少一次出现是亚烷基,或其中 $L^5$ 和/或 $L^6$ 每次出现是亚烷基。在各种实施方案中, $L^{1b}$ 在每次出现时独立地包含酰胺官能团或三唑基官能团。

[0301] 在实施方案中, $R^5$ 在每次出现时独立地为OH、O-或 $OR_d$ 。在实施方案中,其中 $R^4$ 在每次出现时是氧代基。

[0302] 在实施方案中,至少一次出现的 $L^7$ 是任选取代的杂亚烷基连接基,或其中每次出现的 $L^7$ 独立地是任选取代的杂亚烷基连接基。在实施方案中, $L^7$ 包含酰胺或三唑基官能团。在一些实施方案中,每次出现的 $L^7$ 具有以下结构之一:



[0304] 在实施方案中, $n$ 是1至100的整数,或其中 $n$ 是1至10、或2至10的整数。例如,1、2、3、4、5、6、7、8、9或10。

[0305] 在实施方案中, $m$ 为3至6的整数,或其中 $m$ 为3。在实施方案中,荧光团在每次出现时独立地是为二甲基氨基芘,喹吖啶酮,氟苯基-二甲基-BODIPY,His-氟苯基-BODIPY,吡啶,骈三苯,联六苯,卟啉,苯并芘,(氟苯基-二甲基-二氟硼杂-二氮杂-引达省)苯基,(二-氟苯基-二氟硼杂-二氮杂-引达省)苯基,四联苯,联-苯并噻唑,三联-苯并噻唑,联-萘,联-蒽,方酸,方酸鎗,9,10-乙炔基蒽,或三联-萘部分。

[0306] 在实施方案中,荧光团在每次出现时独立地为芘、芘-芘一酰亚胺、5-FAM或6-FAM或其衍生物。

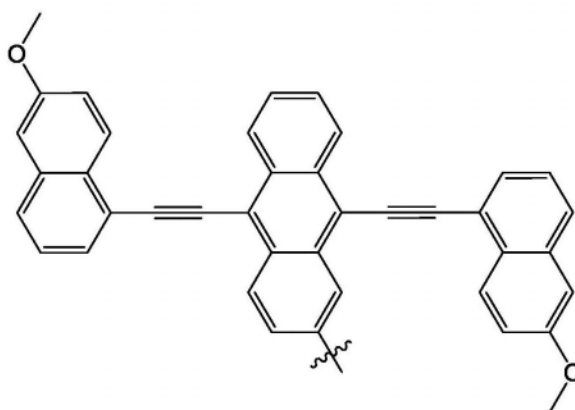
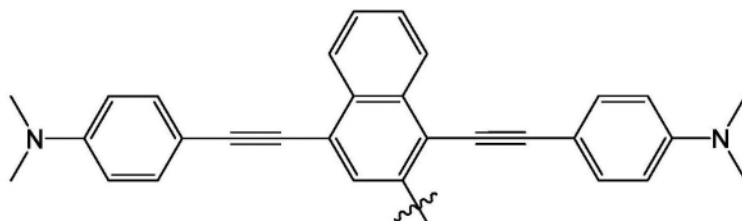
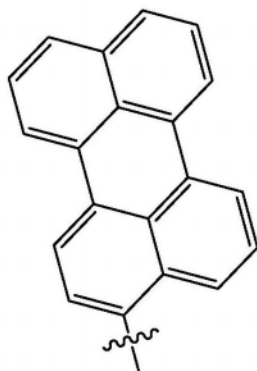
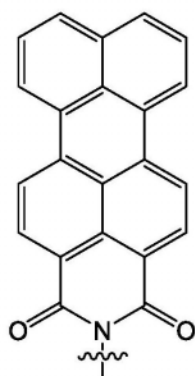
[0307] 在进一步的实施方案中,荧光团在每次出现时独立地选自表1。

[0308] 表1. 示例性荧光团

|        |                                  |
|--------|----------------------------------|
|        | <p><b>结构</b></p> <p><b>F</b></p> |
|        |                                  |
| [0309] | <p>;</p>                         |
|        | <p>;</p>                         |

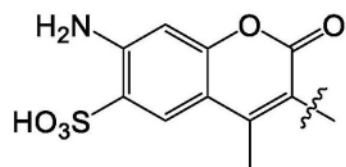
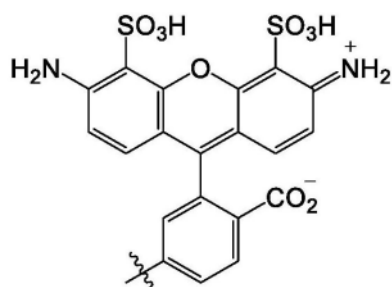
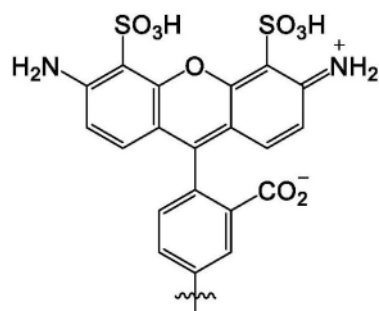
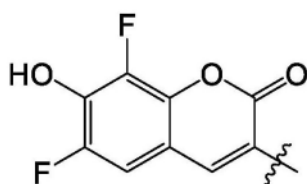
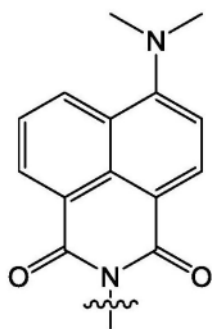
[0310]

结构

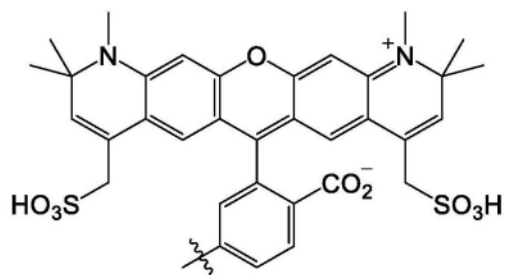
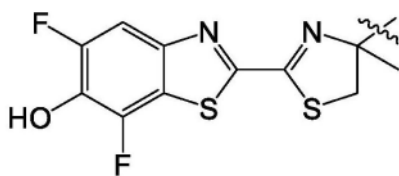


[0311]

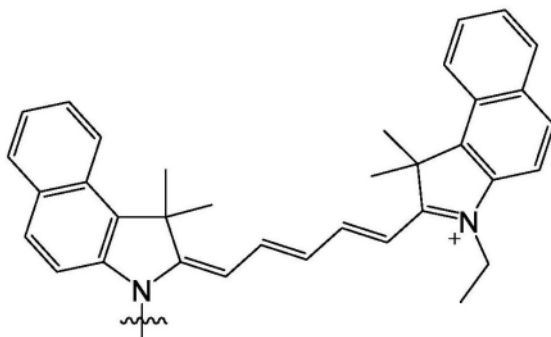
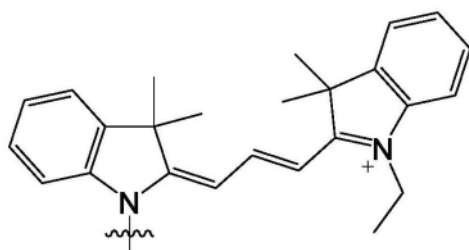
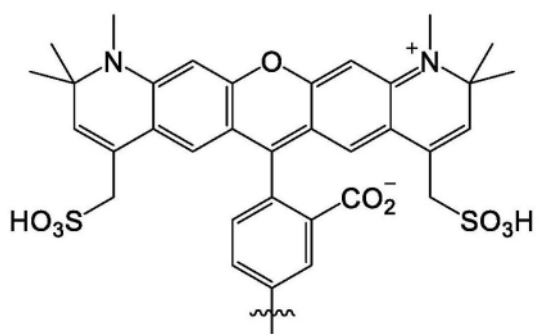
结构



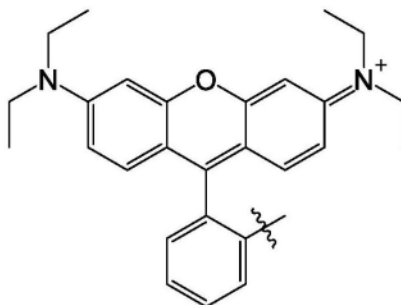
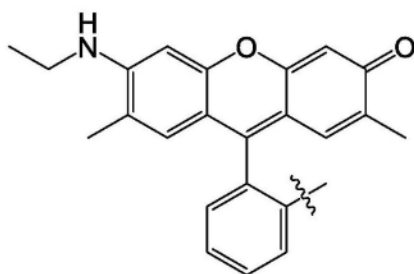
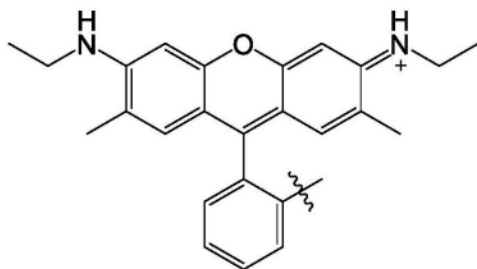
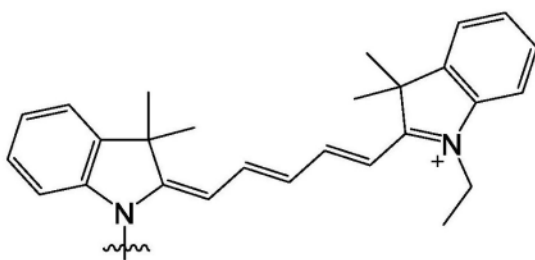
结构



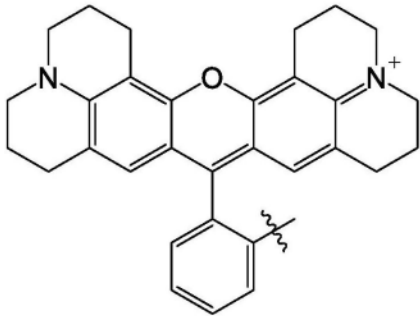
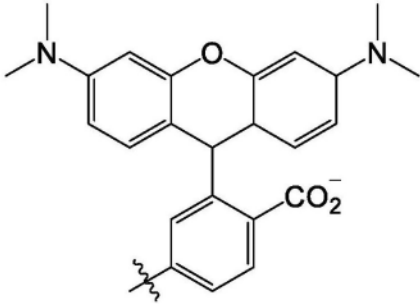
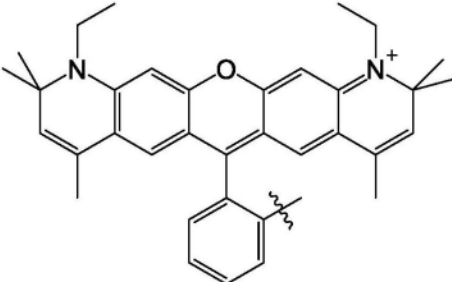
[0312]



结构



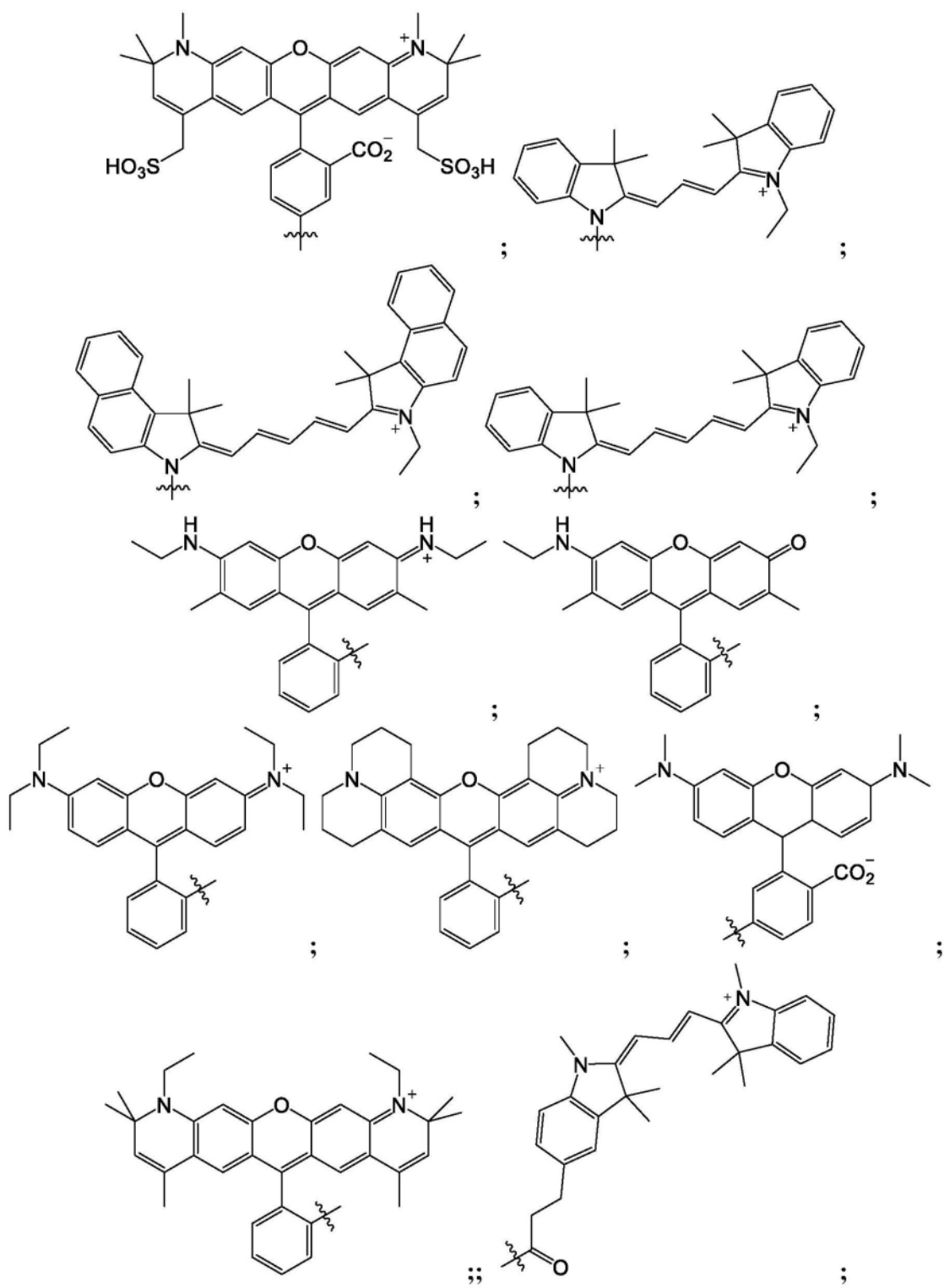
[0313]

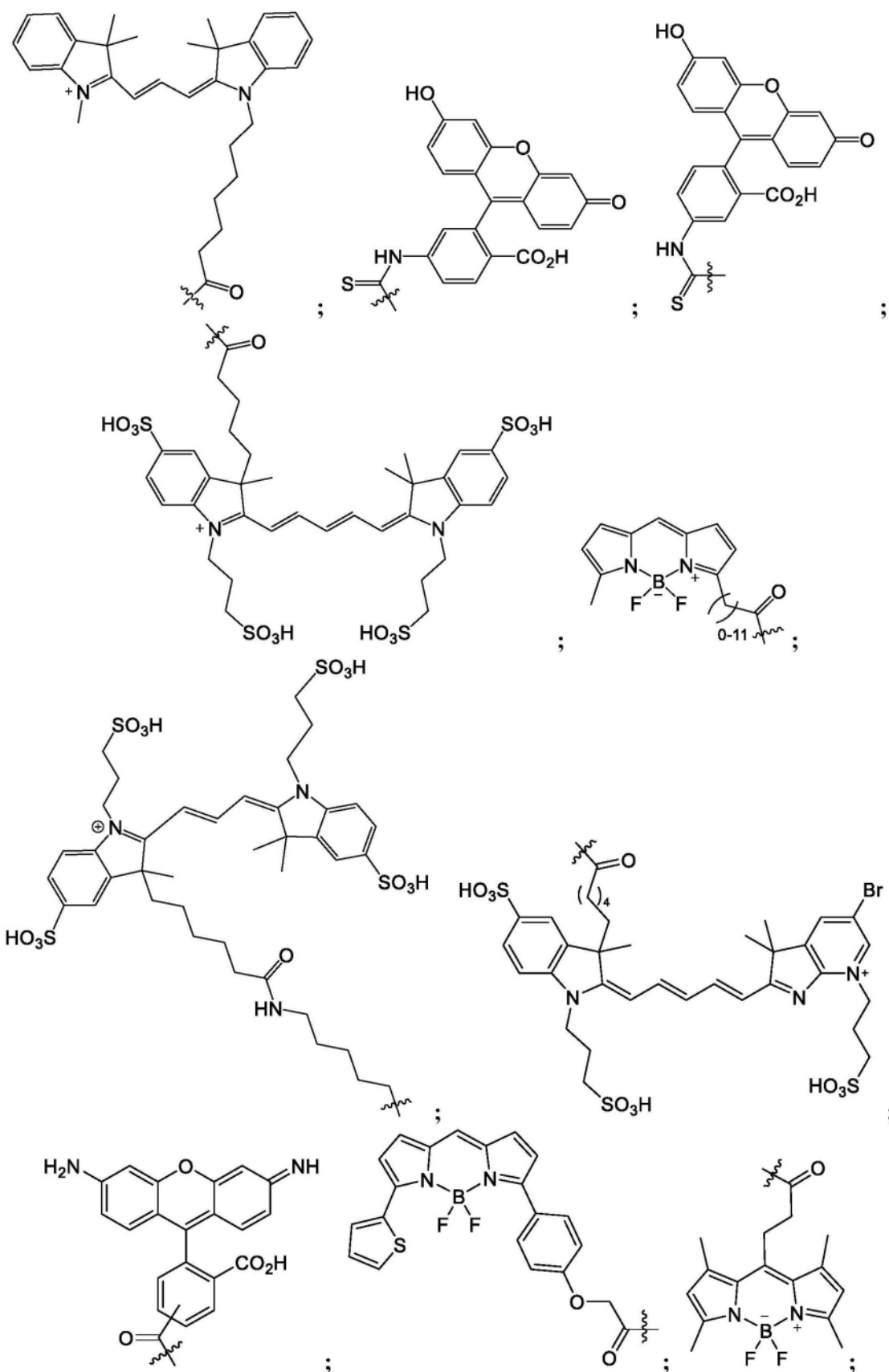
|        |                                                                                                                                |
|--------|--------------------------------------------------------------------------------------------------------------------------------|
|        | <p style="text-align: center;"><b>结构</b></p>  |
| [0314] |                                               |
|        |                                             |

[0315] 在特定实施方案中, 荧光团在每次出现时独立地具有以下结构之一:

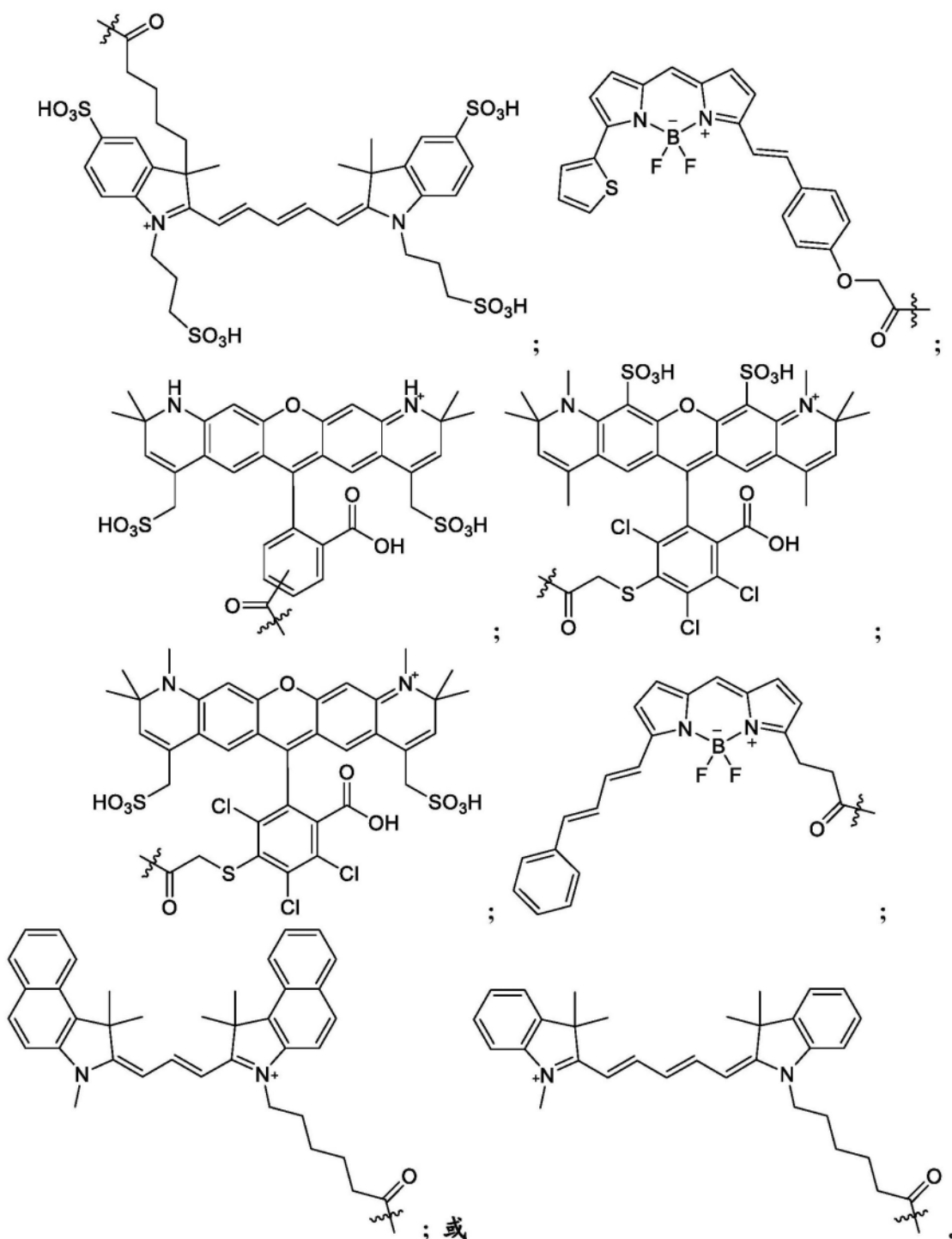


[0319]





[0321]



[0322] FRET过程的效率部分取决于生色团的特性。具体而言,高效FRET需要供体生色团的吸收光谱与受体生色团的发射光谱之间有很大的重叠。此外,生色团的距离和方向起着重要作用。FRET效率与生色团之间距离的6次方成反比,跃迁偶极矩的角度应基本使其平行成平行的(即接近 $0^\circ$ 或 $180^\circ$ )。因此,在某些实施方案中,选择第一和第二生色团与聚合物主链的共价连接,使得第一和第二生色团之间的距离最小化并且跃迁偶极矩基本平行。FRET的效率可以根据以下等式表示:

[0323] 
$$E_{FRET} = \frac{Ro^6}{Ro^6 + R^6}$$

[0324] 其中 $E_{\text{FRET}}$ 是FRET效率,R是生色团之间的距离,Ro根据以下等式表示:

$$[0325] \quad Ro = (8.8 \times 10^{23} JK^2 Q_0 n^{-4})^{1/6} <$$

[0326] 其中J是受体的吸收光谱和供体的发射光谱的光谱重叠, $Q_0$ 是供体的量子效率, $n^{-4}$ 是供体和受体之间的介质的指数(常数), $K^2$ 是偶极子方向匹配。

[0327] 因此,一个实施例提供了一种聚合物化合物,其包含具有受体跃迁偶极矩并共价连接至聚合物主链的受体生色团,以及具有供体跃迁偶极矩并共价连接至聚合物主链的供体生色团,其中该聚合物化合物在生理条件下采用溶液中的确认,其中受体生色团和供体生色团之间的有效距离小于约50.0nm并且受体跃迁偶极子和供体跃迁偶极子基本平行。

[0328] 在一些实施方案中,受体生色团与供体生色团之间的有效距离小于约25.0nm。在一些实施方案中,受体生色团与供体生色团之间的有效距离小于约10.0nm。在一些实施方案中,受体生色团与供体生色团之间的有效距离小于约30.0nm、小于约27.0nm、小于约22.0nm、小于约20.0nm、小于约17.0nm、小于约15.0nm、小于约12.0nm、小于约11.0nm、小于约9.0nm、小于约8.0nm、小于约7.0nm、小于约6.0nm、小于约5.0nm、小于约4.0nm、小于约3.0nm、小于约2.0nm或小于约1.0nm。

[0329] 在一些实施方案中,受体生色团是荧光染料部分。在某些实施方案中,供体生色团是荧光染料部分。在某些相关实施方案中,受体生色团和供体生色团都是荧光染料部分。

[0330] 在一些实施方案中,受体跃迁偶极矩和供体跃迁偶极矩之间的角度在120°至180°的范围内。例如,在一些实施方案中,受体跃迁偶极矩和供体跃迁偶极矩之间的角度范围为125°至180°、130°至180°、140°至180°、150°至180°,从160°到180°,从170°到180°,从172°到180°,从175°到180°,或从177°到180°。

[0331] 在某些实施方案中,受体跃迁偶极矩与供体跃迁偶极矩之间的角度范围为0°至60°。例如,例如,在一些实施方案中,受体跃迁偶极矩和供体跃迁偶极矩之间的角度范围为0°至50°、0°至40°、0°至30°、0°至20°、从0°到10°、从0°到8°、从0°到5°、从0°到3°、或从0°到2°。

[0332] 在实施方案中, $M^1$ 是供体生色团且 $M^2$ 是受体生色团。在这样的实施方案中, $M^1$ 和 $M^2$ 是基于期望的光学特性来选择的,例如基于期望的斯托克斯位移、吸收/发射重叠、特定颜色和/或荧光发射波长。在一些实施方案中, $M^1$ (或备选地 $M^2$ ) 在每次出现时相同;然而,重要的是要注意,每次出现的 $M^1$ 或 $M^2$ 不必分别是相同的 $M^1$ 或 $M^2$ 。某些实施方案包括其中 $M^1$ 在每次出现时不相同的化合物。一些实施方案包括其中 $M^2$ 在每次出现时不相同的化合物。

[0333] 在一些实施方案中,选择 $M^1$ 和 $M^2$ 以具有用于FRET方法的吸收和/或发射特性。例如,在这样的实施方案中,选择不同的 $M^1$ 和 $M^2$ 部分,使得 $M^1$ 在一个波长具有辐射吸收,其通过FRET机制诱导 $M^2$ 在不同波长辐射发射。示例性的 $M^1$ 和 $M^2$ 部分可以由本领域普通技术人员基于期望的最终用途适当地选择。

[0334] 每一个各自的 $M^1$ 和 $M^2$ 可以从 $M^1$ 或 $M^2$ 上的任何位置(即原子)连接到分子的其余部分。本领域技术人员将认识到用于将 $M^1$ 或 $M^2$ 连接到分子的其余部分的手段。示例性方法包括本文所述的“点击”反应。

[0335] 在一些实施方案中, $M^1$ 或 $M^2$ 是FRET、荧光或有色部分。可以使用任何荧光和/或有色部分来形成FRET供体-受体对,例如可以使用本领域已知的并且通常用于比色、UV和/或荧光测定的那些。在一些实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地发荧光或着色。在本

发明的各种实施方案中有用的 $M^1$ 或 $M^2$ 部分的例子包括,例如:咕吨衍生物(例如荧光素,罗丹明,Oregon绿,曙红或Texas红);菁蓝衍生物(例如菁蓝,靛青,氧杂羰花青,硫杂羰花青或部花青);方酸(squaraine)衍生物和环取代的方酸,包括Seta、SeTau和Square染料;萘衍生物(例如丹磺酰和prodan衍生物);香豆素衍生物;噁二唑衍生物(例如吡啶基噁唑,硝基苯并噁二唑或苯并噁二唑);蒽衍生物(例如蒽醌,包括DRAQ5、DRAQ7和CyTRAK橙);芘衍生物比如级联蓝;噁嗪衍生物(例如Ni 1e红,Ni 1e蓝,甲酚基紫,噁嗪170);吡啶衍生物(例如原黄素,吡啶橙,吡啶黄);芳基次甲基衍生物:金胺,晶紫,孔雀绿;和四吡咯衍生物(例如卟吩,酞菁或胆红素)。其它示范性 $M$ 部分包括:菁类染料,黄原酸盐/酯染料(例如Hex,Vic,Nedd,Joe或Tet);Yakima黄;Redmond红;tamra;texas红和alexa**fluor**®染料。

[0336] 在前述任一的其他实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地包含三个或更多个芳基或杂芳基环或其组合,例如四个或更多个芳基或杂芳基环或其组合,或甚至五个或更多个芳基或杂芳基环,或其组合。在一些实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地包含六个芳环或杂芳环或其组合。在进一步的实施方案中,环是稠合的。例如在一些实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地包含三个或更多个稠合环、四个或更多个稠合环、五个或更多个稠合环、或甚至六个或更多个稠合环。

[0337] 在一些实施方案中, $M^1$ 或 $M^2$ 或两者是环状的。例如,在一些实施方案中, $M^1$ 或 $M^2$ 或两者是碳环。在其他实施方案中, $M^1$ 或 $M^2$ 或两者为杂环。在前述任一的其他实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地包含芳基部分。在这些实施方案的一些中,芳基部分是多环的。在其他更具体的实例中,芳基部分是稠合多环芳基部分,例如其可包含至少3个、至少4个或甚至超过4个芳基环。

[0338] 在任何前述结构(I)、(II)、(IA)、(IB)、(IC)、(ID)或(IE)的化合物的其他实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时,独立地包含至少一个杂原子。例如,在一些实施方案中,杂原子是氮、氧或硫。

[0339] 在前述任一的更多实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地包含至少一个取代基。例如,在一些实施方案中,取代基是氟,氯,溴,碘,氨基,烷基氨基,芳基氨基,羟基,巯基,烷氧基,芳氧基,苯基,芳基,甲基,乙基,丙基,丁基,异丙基,叔丁基,羧基,磺酸盐/酯基团,酰胺,或甲酰基。

[0340] 在前述的一些甚至更具体的实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地是为二甲基氨基芘,喹吡啶酮,氟苯基-二甲基-BODIPY,his-氟苯基-BODIPY,吡啶,骈三萘,联六苯,卟啉,苯并芘,(氟苯基-二甲基-二氟硼杂-二氮杂-引达省)苯基,(二-氟苯基-二氟硼杂-二氮杂-引达省)苯基,四联苯,联-苯并噻唑,三联-苯并噻唑,联-萘,联-蒽,方酸,方酸鎗,9,10-乙炔基蒽,或三联-萘部分。

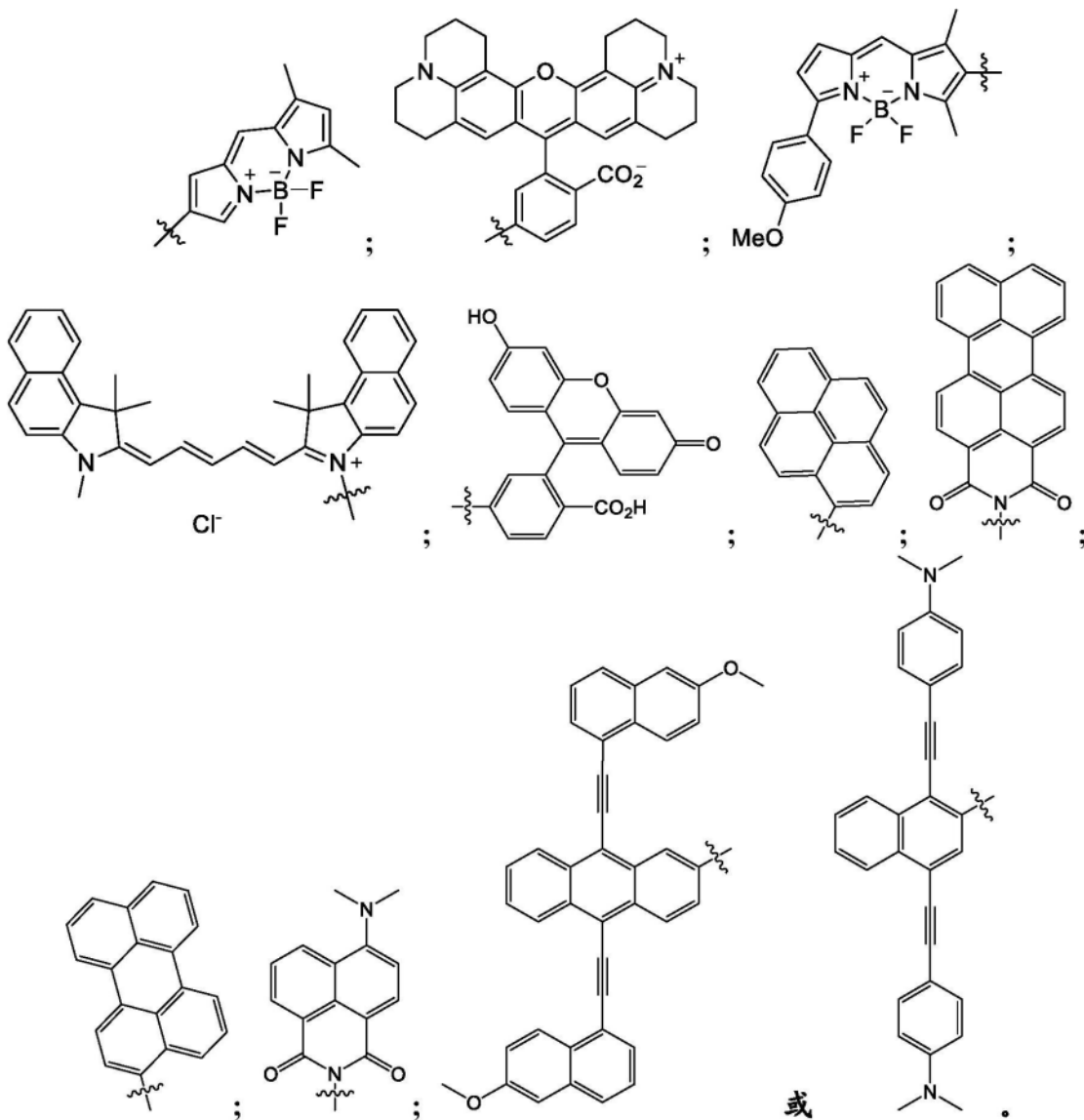
[0341] 在其他实施方案中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地为p-三联苯,花,偶氮苯,吩嗪,菲咯啉,吡啶,噻吨, **蒽**,红荧烯,蒽,青蓝,花酰亚胺或花酰胺或其衍生物。在更多的实施方式中, $M^1$ 或 $M^2$ 或两者在每次出现时独立地是香豆素染料,试卤灵染料,二吡咯亚甲基二氟化物染料,钆联吡啶染料,能量转移染料,噻唑橙染料,聚甲炔,和N-芳基-1,8-萘二甲酰亚胺染料。在某些实施方案中, $M^1$ 或 $M^2$ 在每次出现时独立地为硼-二吡咯亚甲基、罗丹明、青蓝、芘、花、花-酰亚胺或6-FAM或其衍生物。

[0342] 任何前述的更多实施方案中, $M^1$ 在每次出现时是相同的。在其他实施方案中,每个

$M^1$ 是不同的。在更多实施方案中,一个或多个 $M^1$ 相同且一个或多个 $M^1$ 不同。

[0343] 在任何前述的更多实施方案中,  $M^2$  在每次出现时是相同的。在其他实施方案中, 每个  $M^2$  是不同的。在更多实施方案中, 一个或多个  $M^2$  相同且一个或多个  $M^2$  不同。

[0344] 在一些实施方案中, M<sup>1</sup>或M<sup>2</sup>或两者在每次出现时独立地为硼-二吡咯亚甲基、罗丹明、青蓝、芘、茈、茈一酰亚胺,或6-FAM或其衍生物。在一些其他实施方案中, M<sup>1</sup>和M<sup>2</sup>在每次出现时独立地具有以下结构之一:



[0346] 如所理解的,本文描述的化合物可以使用任何合适的方法形成,例如US 2017/0292957,US 2016/0208100,US 2016/0341736,US 2018/0065998,US 2018/0079909,和US 2019/0016898的此类教导通过引用整体并入。

[0347] 说明性捕获探针102和说明性可检测探针104、106的示意性版本如图1所示。捕获探针102包含(1)包含具有第一序列的第一多核苷酸108的第一片段,和(2)与第一片段缀合的第二片段,第二片段包含多个交替的第二多核苷酸110和第一聚合物112a、112b,每个第一聚合物包含第一生色团114,每个第二多核苷酸包含第二序列。在一些实施方案中,捕获探针是如上所述的结构(III)的化合物。

[0348] 还存在多个可检测探针104、106,多个可检测探针中的每个包含第三多核苷酸116,该第三多核苷酸116共价结合至至少一种包含第二生色团118a、118b的第二聚合物。在实施方案中,第三多核苷酸具有第三序列,其具有与第二序列至少90%互补性。

[0349] 因此,本文描述的是组合物,其包含:(1)捕获探针,其包含:(i)第一片段,其包含具有第一序列的第一多核苷酸;(ii)与第一片段缀合的第二片段,第二片段包含多个交替的第二多核苷酸和第一聚合物,每个第一聚合物包含第一生色团,每个第二多核苷酸包含第二序列;(2)多个可检测探针,多个可检测探针中的每个都包含第三多核苷酸,该第三多核苷酸共价结合至至少一个包含第二生色团的第二聚合物,该第三多核苷酸具有第三序列,该第三序列具有与第二序列至少90%的互补性。

[0350] 在实施方案中,第二序列具有与第三序列至少92%的互补性。在一些实施方案中,第二序列具有与第三序列至少95%的互补性。在一些实施方案中,第二序列具有与第三序列至少96%的互补性。在一些实施方案中,第二序列具有与第三序列至少97%的互补性。在一些实施方案中,第二序列具有与第三序列至少98%的互补性。在一些实施方案中,第二序列具有与第三序列至少99%的互补性。在一些实施方案中,三链体(triplex)结构由两个可检测探针的多核苷酸和捕获探针的多核苷酸形成。

[0351] 如图所示参照图1,在一些实施方案中,每个第一聚合物包含多个第一生色团。类似地,在一些实施方案中,每个可检测探针包含多个第二生色团。在一些实施方案中,第二聚合物118a、118b共价结合至至少一个可检测探针106的第三多核苷酸116的每一末端。

[0352] 在一些实施方案中,例如图1所示的实施方案。第一个生色团是受体生色团。在其他实施方案中,第一生色团是供体生色团。如图1所示,在一些实施方案中,第二生色团是供体生色团。在其他实施方案中,第二生色团是供体生色团。

[0353] 在实施方案中,第三多核苷酸的长度范围为10个核苷酸至40个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为10个核苷酸至30个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为15个核苷酸至25个核苷酸。

[0354] 在实施方案中,第三多核苷酸包含第一部分、与第一部分相邻的第二部分和通过第二部分与第一部分分开的第三部分。在一些此类实施方案中,第三多核苷酸与第二序列的一部分的互补程度高于第一部分与第三部分的互补程度。因此,第三个多核苷酸将优先与第二个序列杂交而不是形成发夹结构。在实施方案中,至少一部分可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,至少一个可检测探针具有如上所述的结构(I)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(II)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIb)。

[0355] 在实施方案中,每个可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,每个可检测探针具有如上所述的结构(I)。在一些实施方案中,每个可检测探针具有如上所述的结构(II)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIb)。

[0356] 也如图1所示,第四多核苷酸120共价结合至靶向部分122。在其他实施方案中,靶向部分共价结合至捕获探针。

[0357] 在实施方案中,组合物包含如本文所述的捕获探针和可检测探针。因此,本文描述

的是组合物,其包含:(1)捕获探针,其包含:(i)第一片段,其包含具有第一序列的第一多核苷酸;(ii)与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸,每个第二多核苷酸包含第二序列;(2)多个可检测探针,多个可检测探针中的每个包含具有第一末端和第二末端的第三多核苷酸,第一末端和第二末端中的至少一个共价结合至第一聚合物,所述第一聚合物包含生色团,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。

[0358] 在实施方案中,第三多核苷酸的长度范围为10个核苷酸至40个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为10个核苷酸至30个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为15个核苷酸至25个核苷酸。

[0359] 在实施方案中,第二序列具有与第三序列至少92%的互补性。在一些实施方案中,第二序列具有与第三序列至少95%的互补性。在一些实施方案中,第二序列具有与第三序列至少96%的互补性。在一些实施方案中,第二序列具有与第三序列至少97%的互补性。在一些实施方案中,第二序列具有与第三序列至少98%的互补性。在一些实施方案中,第二序列具有与第三序列至少99%的互补性。

[0360] 在一些实施方案中,第一聚合物中的每个包含多个生色团。类似地,在一些实施方案中,每个第二聚合物包含多个生色团。在一些实施方案中,还存在第二种类型的可检测探针。在一些这样的实施方案中,可检测探针包含与第一聚合物或第二聚合物结合的第三多核苷酸。

[0361] 在实施方案中,至少一部分可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,至少一个可检测探针具有如上所述的结构(I)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(II)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIb)。

[0362] 在实施方案中,每个可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,每个可检测探针具有如上所述的结构(I)(如图2A所示)。在一些实施方案中,每个可检测探针具有如上所述的结构(II)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIb)。

[0363] 在实施方案中,第四多核苷酸共价结合至靶向部分。在各种实施方案中,第四多核苷酸包含具有与第一序列至少90%互补性的第四序列。在一些实施方案中,第四序列具有与第一序列至少92%的互补性。在一些实施方案中,第四序列具有与第一序列至少95%的互补性。在一些实施方案中,第四序列具有与第一序列至少96%的互补性。在一些实施方案中,第四序列具有与第一序列至少97%的互补性。在一些实施方案中,第四序列具有与第一序列至少98%的互补性。在一些实施方案中,第四序列具有与第一序列至少99%的互补性。在其他实施方案中,靶向部分共价结合至捕获探针。

[0364] 说明性捕获探针202和说明性可检测探针204a、204b、204c的另一个实施方案的示意性版本示于图2A中。捕获探针102包含(1)第一片段,其包含具有第一序列的第一多核苷酸,和(2)与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸210a、210b、210c,每个包含第二序列。

[0365] 还存在多个可检测探针204a、204b、204c,多个可检测探针中的每个包含第三多核苷酸216a、216b、216c。第三多核苷酸在一末端共价结合到包含生色团214的第一聚合物。在

实施方案中,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。

[0366] 因此,本文描述的是组合物,其包含:(1)捕获探针,其包含:(i)第一片段,其包含具有第一序列的第一多核苷酸;(ii)与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸,每个第二多核苷酸包含第二序列;(2)多个可检测探针,多个可检测探针中的每个包含具有第一末端和第二末端的第三多核苷酸,第一末端共价结合至包含生色团的第一聚合物,第三多核苷酸具有第三序列,其具有与第二序列至少90%互补性。

[0367] 在实施方案中,第三多核苷酸的长度范围为10个核苷酸至40个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为10个核苷酸至30个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为15个核苷酸至25个核苷酸。

[0368] 在实施方案中,第二序列具有与第三序列至少92%的互补性。在一些实施方案中,第二序列具有与第三序列至少95%的互补性。在一些实施方案中,第二序列具有与第三序列至少96%的互补性。在一些实施方案中,第二序列具有与第三序列至少97%的互补性。在一些实施方案中,第二序列具有与第三序列至少98%的互补性。在一些实施方案中,第二序列具有与第三序列至少99%的互补性。在一些实施方案中,三链体结构由两个可检测探针的多核苷酸和捕获探针的多核苷酸形成。

[0369] 如图2A所示,在一些实施方案中,第一聚合物中的每个包含多个生色团。在一些实施方案中,还存在第二种类型的可检测探针。在一些这样的实施方案中,可检测探针包含与第一聚合物和/或第二聚合物结合的第三多核苷酸。

[0370] 在实施方案中,至少一部分可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,至少一个可检测探针具有如上所述的结构(I)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(II)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIb)。

[0371] 在实施方案中,每个可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,每个可检测探针具有如上所述的结构(I)(如图2A所示)。在其他实施方案中,每个可检测探针具有如上所述的结构(II)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIb)。

[0372] 也如图2A所示,第四多核苷酸220共价结合至靶向部分222。在各种实施方案中,第四多核苷酸包含具有与第一序列至少90%互补性的第四序列。在一些实施方案中,第四序列具有与第一序列至少92%的互补性。在一些实施方案中,第四序列具有与第一序列至少95%的互补性。在一些实施方案中,第四序列具有与第一序列至少96%的互补性。在一些实施方案中,第四序列具有与第一序列至少97%的互补性。在一些实施方案中,第四序列具有与第一序列至少98%的互补性。在一些实施方案中,第四序列具有与第一序列至少99%的互补性。在其他实施方案中,靶向部分共价结合至捕获探针。

[0373] 说明性捕获探针202和说明性可检测探针204、206a、206b的另一个实施方式的示意版本示于图2B中。捕获探针102包含(1)第一片段,其包含具有第一序列的第一多核苷酸,和(2)与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸210a、210b、210c,每个包含第二序列。

[0374] 还存在多个可检测探针204、206a、206b,多个可检测探针中的每个包含第三多核

苷酸216a、216b、216c。第三多核苷酸在一末端共价结合到包含受体生色团214的第一聚合物,并且在另一末端共价结合到包含供体生色团218的第二聚合物。换言之,多个可检测探针中的每个包含具有第一末端和第二末端的第三多核苷酸,第一末端共价结合至包含受体生色团的第一聚合物,第二末端共价结合至包含供体生色团的第二聚合物。在实施方案中,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。

[0375] 因此,本文描述的是组合物,其包含:(1)捕获探针,其包含:(i)第一片段,其包含具有第一序列的第一多核苷酸;(ii)与第一片段缀合的第二片段,第二片段包含多个第二多核苷酸,每个第二多核苷酸包含第二序列;(2)多个可检测探针,多个可检测探针中的每个包含具有第一末端和第二末端的第三多核苷酸,第一末端共价结合至包含受体生色团的第一聚合物,第二末端共价结合到包含供体生色团的第二聚合物,第三多核苷酸包含具有与第二序列至少90%互补性的第三序列。在实施方案中,第二序列具有与第三序列至少92%的互补性。在一些实施方案中,第二序列具有与第三序列至少95%的互补性。

[0376] 在一些实施方案中,第二序列具有与第三序列至少96%的互补性。在一些实施方案中,第二序列具有与第三序列至少97%的互补性。在一些实施方案中,第二序列具有与第三序列至少98%的互补性。在一些实施方案中,第二序列具有与第三序列至少99%的互补性。在一些实施方案中,三链体结构由两个可检测探针的多核苷酸和捕获探针的多核苷酸形成。

[0377] 如图2B所示,在一些实施方案中,每个第一聚合物包含多个生色团。类似地,在一些实施方案中,每个第二聚合物包含多个生色团。在一些实施方案中,还存在第二种类型的可检测探针204。在一些此类实施方案中,可检测探针204包含结合到第一聚合物或第二聚合物的第三多核苷酸。

[0378] 在实施方案中,第三多核苷酸的长度范围为10个核苷酸至40个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为10个核苷酸至30个核苷酸。在一些实施方案中,第三多核苷酸的长度范围为15个核苷酸至25个核苷酸。

[0379] 在实施方案中,至少一部分可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,至少一个可检测探针具有如上所述的结构(I)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(II)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,至少一个可检测探针具有如上所述的结构(IIb)。

[0380] 在实施方案中,每个可检测探针具有上述结构(I)、(II)、(IIa)、(IIb)或其组合。在一些实施方案中,每个可检测探针具有如上所述的结构(I)。在一些实施方案中,每个可检测探针具有如上所述的结构(II)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIa)。在一些实施方案中,每个可检测探针具有如上所述的结构(IIb)。

[0381] 也如图2B所示,第四多核苷酸220共价结合至靶向部分222。在各种实施方案中,第四多核苷酸包含具有与第一序列至少90%互补性的第四序列。在一些实施方案中,第四序列具有与第一序列至少92%的互补性。在一些实施方案中,第四序列具有与第一序列至少95%的互补性。在一些实施方案中,第四序列具有与第一序列至少96%的互补性。在一些实施方案中,第四序列具有与第一序列至少97%的互补性。在一些实施方案中,第四序列具有与第一序列至少98%的互补性。在一些实施方案中,第四序列具有与第一序列至少99%的

互补性。在其他实施方案中，靶向部分共价结合至捕获探针。

[0382] 另一个实施方案在图2C中所示。捕获探针102包含(1)第一片段，其包含具有第一序列的第一多核苷酸，和(2)与第一片段缀合的第二片段，第二片段包含多个第二多核苷酸210，每个第二多核苷酸210包含第二序列。

[0383] 还存在多个分支连接基226和多个可检测探针204。多个可检测探针中的每个包含第三多核苷酸216(具有第三序列)并且多个分支连接基中的每个包含第四多核苷酸228(具有第四序列)和第五多核苷酸224(具有第五序列)。第三多核苷酸在一末端共价结合到包含生色团214的第一聚合物。

[0384] 在实施方案中，第二序列具有与第四序列至少90%的互补性。在实施方案中，第二序列具有与第四序列至少92%的互补性。在一些实施方案中，第二序列具有与第四序列至少95%的互补性。在一些实施方案中，第二序列具有与第四序列至少96%的互补性。在一些实施方案中，第二序列具有与第四序列至少97%的互补性。在一些实施方案中，第二序列具有与第四序列至少98%的互补性。在一些实施方案中，第二序列具有与第四序列至少99%的互补性。

[0385] 在实施方案中，第三序列具有与第五序列至少90%的互补性。在实施方案中，第三序列具有与第五序列至少92%的互补性。在一些实施方案中，第三序列具有与第五序列至少95%的互补性。在一些实施方案中，第三序列具有与第五序列至少96%的互补性。在一些实施方案中，第三序列具有与第五序列至少97%的互补性。在一些实施方案中，第三序列具有与第五序列至少98%的互补性。在一些实施方案中，第三序列具有与第五序列至少99%的互补性。

[0386] 在实施方案中，第二、第三、第四和/或第五多核苷酸的长度范围为10个核苷酸至40个核苷酸。在一些实施方案中，第二、第三、第四和/或第五多核苷酸的长度范围为10个核苷酸至30个核苷酸。在一些实施方案中，第二、第三、第四和/或第五多核苷酸长度范围为15个核苷酸至25个核苷酸。在一些实施方案中，三链体结构由两个可检测探针的多核苷酸和捕获探针的多核苷酸形成。

[0387] 在一些实施方案中，第一聚合物中的每个包含多个生色团。

[0388] 在实施方案中，至少一部分可检测探针具有结构(I)。在一些实施方案中，至少一个可检测探针具有如上所述的结构(I)。在一些实施方案中，每个可检测探针具有如上所述的结构(I)。

[0389] 在一些实施方案中，第六多核苷酸220共价结合至靶向部分222。在各种实施方案中，第六多核苷酸包含具有与第一序列至少90%互补性的第六序列。在一些实施方案中，第六序列具有与第一序列至少92%的互补性。在一些实施方案中，第六序列具有与第一序列至少95%的互补性。在一些实施方案中，第六序列具有与第一序列至少96%的互补性。在一些实施方案中，第六序列具有与第一序列至少97%的互补性。在一些实施方案中，第六序列具有与第一序列至少98%的互补性。在一些实施方案中，第六序列具有与第一序列至少99%的互补性。在其他实施方案中，靶向部分共价结合至捕获探针。

[0390] 因此，本文描述的是组合物，其包含：(1)捕获探针，其包含：(i)第一片段，其包含具有第一序列的第一多核苷酸；(ii)与第一片段缀合的第二片段，第二片段包含多个第二多核苷酸，每个第二多核苷酸包含第二序列；(2)多个可检测探针，多个可检测探针中的每

个包含具有第一末端和第二末端的第三多核苷酸,第一末端共价结合至包含生色团的第一聚合物,第三多核苷酸具有第三序列;(3)多个分支连接基,其包含与第四多核苷酸和多个第五多核苷酸结合的聚合物主链,第四多核苷酸包含具有与第二序列至少90%互补性的第四序列,第五多核苷酸中的每个包含具有与第三序列至少90%互补性的第五序列。

[0391] 本文还描述了包含本文所述的捕获探针、可检测探针或两者的试剂盒。在一些这样的实施方案中,本公开的试剂盒还包括用于使用化合物、组合物或可检测探针来鉴定靶核苷酸序列的说明书。

[0392] 进一步描述了用于鉴定靶标分析物存在的方法,包括:通过在测定条件下使样品与本文描述的组合物接触来产生混合物;并在检测条件下对所述混合物进行成像。在一些实施方案中,使混合物成像包括:在第一波长激发供体荧光团;在第二波长检测受体荧光团的发射。

## 实施例

[0393] 一般方法使用MassLynx 4.1采集软件在Waters/Micromass Quattro micro MS/MS系统(仅MS模式)上进行质谱分析。用于染料LC/MS的流动相为100mM 1,1,1,3,3,3-六氟-2-丙醇(HFIP)、8.6mM三乙胺(TEA),pH 8。还使用配备2.1mm x 50mm Acquity BEH-C18色谱柱(保持在45℃)的Waters Acquity UHPLC系统分析磷酸胺和前体分子,采用乙腈/水流动相梯度。单体中间体的分子量是在Waters/Micromass Quattro微型MS/MS系统(仅MS模式)上使用卓鎬(tropylium)阳离子输注增强电离获得的。在Cary Eclipse光谱光度计上记录激发和发射曲线实验。

[0394] 除非另有说明,否则所有反应均在氮气气氛下在烘箱干燥的玻璃器皿中进行。市售DNA合成试剂购自Glen Research(Sterling,VA)。无水吡啶、甲苯、二氯甲烷、二异丙基乙胺、三乙胺、醋酸、吡啶和THF购自Aldrich。所有其他化学品均购自Aldrich或TCI,并按原样使用,无需另外纯化。

[0395] 实施例1

[0396] 具有乙二醇间隔基的染料的合成

[0397] 具有氧化乙烯连接基的化合物的制备如下:

[0398] 寡氟苷(oligofluoroside)构建体(例如,结构(I)、(II)或(III)的化合物)在Applied Biosystems 394DNA/RNA合成器以1μmol规模合成,具有3'-磷酸基团或3'-S<sub>2</sub>-(CH<sub>2</sub>)<sub>6</sub>-OH基团或本文所述的任何其他基团。合成是直接在CPG珠子上或在聚苯乙烯固体载体上使用标准亚磷酸胺化学进行的。使用标准固相DNA方法在3'到5'方向合成低聚氟苷,并使用标准β-氰乙基亚磷酸胺化学偶联。将氟苷和核苷亚磷酸胺和间隔基(例如,六乙氧基-乙二醇亚磷酸胺、三乙氧基-乙二醇亚磷酸胺、聚乙二醇亚磷酸胺)和连接基(例如,5'-氨基-修饰亚磷酸胺和硫醇-修饰S2亚磷酸胺)溶解在乙腈中,制成0.1M溶液,并使用以下合成循环按顺序添加:1)在二氯甲烷中用二氯乙酸去除5'-二甲氧基三苯甲基保护基,2)在乙腈中将下一个亚磷酸胺与活化剂试剂偶联,3)用碘/吡啶/水氧化P(III)形成稳定的P(V),和4)用乙酸酐/1-甲基咪唑/乙腈封端任何未反应的5'-羟基。重复合成循环,直到全长寡氟苷构建体组装完成。在链组装结束时,单甲氧基三苯甲基(MMT)基团或二甲氧基三苯甲基(DMT)基团在二氯甲烷中用二氯乙酸去除。

[0399] 在标记的Eppendorf管中以0.2 $\mu$ mol标度在可控孔径玻璃(CPG)载体上提供化合物。加入400 $\mu$ L的20-30% $\text{NH}_4\text{OH}$ 并轻轻混合。将打开的试管在55 $^{\circ}\text{C}$ 下放置约5分钟或直到释放出过量气体,然后盖紧关闭并孵育2小时(+/-15分钟)。将试管从加热区块中取出并使其达到室温,然后以13,400RPM离心30秒以巩固上清液和固体。小心移去上清液并置于标记管中,然后加入150 $\mu$ L乙腈洗涤载体。将洗涤液加入试管后,将试管置于40 $^{\circ}\text{C}$ 的CentriVap装置中直至干燥。

[0400] 通过ESI-MS、紫外吸光度和荧光光谱对产品进行表征。

[0401] 实施例2

[0402] 化合物的光谱测试

[0403] 将干燥的化合物在150 $\mu$ L的0.1M  $\text{Na}_2\text{CO}_3$ 缓冲液中重构以制备 $\sim$ 1mM储备液。浓缩储备液在0.1x PBS中稀释50倍,并在NanoDrop紫外光谱仪上进行分析以获得吸光度读数。吸光度读数与消光系数(每个FAM单元为75,000 $\text{M}^{-1}\text{cm}^{-1}$ )和比尔定律一起用于确定储备液的实际浓度。

[0404] 根据计算出的储备液浓度,在0.1M  $\text{Na}_2\text{CO}_3$  (pH9)中制备约4mL5 $\mu$ M溶液,并在Cary 60紫外光谱仪上的1x 1cm石英比色皿中使用300nm至700nm的光谱范围进行分析,以测量相对于组的整体吸光度。从这些5 $\mu$ M溶液中,以50nM或25nM(也在0.1M  $\text{Na}_2\text{CO}_3$ , pH9中)进行第二次稀释,用于Cary Eclipse荧光计上光谱分析。激发设置在494nm,从499到700nm收集发射光谱。

[0405] 实施例3

[0406] 通过串联体化(CONCATEMERIZATION)的信号放大

[0407] 生物素化靶标探针(5'-B-heg-ATGCACAGTCGG-dT-3') (SEQ ID NO:1)单独结合中性抗生物素蛋白珠以及与(1)第一个探针(5'-CGA CGC TTA CAG-heg-F-(heghegheg-F)3-heg-CCG ACT GTG CA-dT-3') (SEQ ID NO:2)和多个第二探针(5'-CTGTAAGCGTCG-heg-F-(heghegheg-F)9-heg-GACATTCGCAGC-dT-3') (SEQ ID NO:3), (2)第一个探针(5'-CGA CGC TTA CAG-heg-F-(heghegheg-F)3-heg-CCG ACT GTG CA-dT-3') (SEQ ID NO:2)混合,使用流式细胞术进行分析。

[0408] 使用热循环仪将储备液加热至94 $^{\circ}\text{C}$  10分钟,然后冷却并保持在37 $^{\circ}\text{C}$ 直至使用。每个样品结合五纳克(ng)捕获探针通过在黑暗室温下孵育30分钟,添加至30 $\mu$ L的中性抗生物素蛋白包被珠。用杂交缓冲液(10mM Tris-HCL、4mM  $\text{MgCl}_2$ 、15mM KCl、pH 8.2)洗涤珠子并以400xg离心5分钟。去除上清液并将样品重悬浮于细胞染色缓冲液中。对于条件1和2,添加1 $\mu$ L的100 $\mu$ M第一探针,并将样品在室温下孵育10分钟,然后用杂交缓冲液洗涤。样品在400xg下离心5分钟,去除上清液并将样品重悬浮于细胞染色缓冲液中或用第二个探针处理。对于条件2,添加1 $\mu$ L的100 $\mu$ M第二探针,样品在室温下孵育10分钟,然后用杂交缓冲液洗涤。样品以400xg离心5分钟,去除上清液。将样品重悬浮于细胞染色缓冲液中,并在SA3800流式细胞仪上进行分析。结果示于图4。

[0409] 实施例4

[0410] 分支扩增

[0411] 将结合到中性抗生物素蛋白珠的5ng生物素化靶标探针(5'-生物素-heg-TTT CTT TGA GGT TTA GGA TTC-dT-3') (SEQ ID NO:4)与靶标放大器(3'-dT-AAA GAA ACT CCA AAT

CCT AAG-GGT TGG CCT TAG GGT TC AGA-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5) 200pmol和多个检测探针混合,所述多个检测探针选自:(1)5'-CCA ACC GGA ATC CCA AG X-F-dT-3' (SEQ ID NO:6),5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)3-heg-dT-3' (SEQ ID NO:7);(2)5'-CCA ACC GGA ATC CCA AG X-F-dT-3' (SEQ ID NO:6),5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)4-heg-dT-3' (SEQ ID NO:8);和(3)5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)9-heg-dT-3' (SEQ ID NO:9)。

[0412] 使用热循环仪将探针储备液加热至94℃10分钟,然后冷却并保持在37℃直至使用。通过在黑暗中在室温下孵育30分钟,将每个样品5ng捕获探针与中性抗生物素蛋白珠结合。用杂交缓冲液洗涤珠子,以400xg离心5分钟并去除上清液。添加200pmol靶标放大器并将样品在37℃下孵育10分钟,然后如上所述用杂交缓冲液洗涤。添加400pmole检测探针并将样品在37℃下孵育10分钟。样品如上所述用杂交缓冲液洗涤,并重悬浮于细胞染色缓冲液中用于在SA3800流式细胞仪上进行分析。

[0413] 实施例5

[0414] 分支扩增

[0415] 捕获探针单独或与下述混合:(1)靶标放大器(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5) 200pmol,和(2)多个检测探针和靶标放大器(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5) 200pmol,以及多个检测探针(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)3-heg-dT-3') (SEQ ID NO:7) 400pmol。结果示于图5。在4X放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)3-heg-dT-3';A03) (SEQ ID NO:7) 杂交后,荧光增加约9倍。

[0416] 实施例6

[0417] 分支扩增

[0418] 样品与下述在37℃下孵育10分钟:捕获探针(A01);捕获探针+靶标(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5) - (A02);捕获探针+靶标(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5)+5x放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)4-heg-dT-3') (SEQ ID NO:8) - (A03);捕获探针+靶标(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5)+10x放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)9-heg-dT-3') (SEQ ID NO:9) - (A04);靶标(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5)+5x放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)4-heg-dT-3') (SEQ ID NO:8) - (A05);以及靶标(3'-dT-AAA GAA ACT CCA AAT CCT AAG-GGT TGG CCT TAG GGT TC AGA-heg-F-5') (SEQ ID NO:5)+10x放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)9-heg-dT-3') (SEQ ID NO:9) - (A06)。结果示于图6。在5X放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)4-heg-dT-3') (SEQ ID NO:8) 杂交后荧光增加约11倍,和在10X放大器(5'-CCA ACC GGA ATC CCA AG-heg-F(heghegheg-F)9-heg-dT-3') (SEQ ID NO:9) 杂交后荧光

增加约16倍。

[0419] 实施例7

[0420] 串联体扩增

[0421] 生物素化靶探针 (5'-B-heg-ATGCACAGTCGG-dT-3') (SEQ ID NO:1) 结合到 (3'-dTACGTGTCAGCC-heg-F (heghegheg-F) 3-heg-GACATTTCGAGC-5') (SEQ ID NO:10), 与第二探针 (5'-CTGTAAGCGTCG-heg-F (heghegheg-F) 4-heg-CGACGCTTACAG-dT-3') (SEQ ID NO:11), 它之以字线方式串联 (如图7所示)。然后测试所得结构。

[0422] 实施例8

[0423] 在溶液中的逐步线性探针杂交

[0424] 第一个探针 (3'-dT-AAA GAA ACT CCA ATC CTA AG-heg-F (heghegheg-F) -heg-GGT TGG CCT TAG GGT TC-5') (SEQ ID NO:12) 与第二个探针 (5'-CCA ACC GGA ATC CCA AG-heg-F (heghegheg-F) -heg-CTT TGA GGT TTA GGA T-dT-3') (SEQ ID NO:13) 在各种杂交条件下混合 (将每个探针的200uM储备液加热至94℃, 持续5分钟, 以1:1的比例组合, 每个探针的最终浓度为100uM)。结果示于图8。在~7.6m处发现单个探针, 在~6.8m处开始杂交探针 (可能是单杂交), 从4.2-5.5m处可以看到更长的多聚体, 这些都没有完全分辨。在37℃下孵育会产生更长的链。

[0425] 此外, 如图8B所示, 第一个探针与第二个探针的杂交导致荧光增加~1.6倍。然而, 用第一个探针再用第二个探针进行的另外的孵育并没有增加荧光。用含有多聚体的预杂交材料孵育从溶液分析中没有显示荧光的另外增加。

[0426] 可以组合上述各种实施例以提供进一步的实施方案。本说明书中引用和/或申请数据表中列出的所有美国专利、美国专利申请公开、美国专利申请、外国专利、外国专利申请和非专利出版物, 包括2020年11月25日提交的美国临时专利申请号63/118,544号, 其全部内容通过引用并入本文。如果需要采用各种专利、申请和出版物的概念来提供又进一步的实施方案, 则可以修改实施方案的各方面。

[0427] 可以根据以上详细描述对实施方案进行这些和其他改变。一般而言, 在所附权利要求中, 所使用的术语不应被解释为将权利要求限制为说明书和权利要求中公开的具体实施方案, 而应被解释为包括所有可能的实施方案以及这些权利要求享有的等同物的全部范围。因此, 权利要求不受本公开的限制。

## 序列表

<110> 索尼集团公司  
<120> 聚合物染料  
<130> 870268.430W0  
<140> PCT  
<141>  
<150> US 63/118,544  
<151> 2020-11-25  
<160> 13  
<170> 用于Windows版本 4.0的 FastSEQ  
<210> 1  
<211> 12  
<212> DNA  
<213> 人工序列  
<220>  
<223> 生物素化靶标探针  
<220>  
<221> misc\_特征  
<222> 1  
<223> B-heg  
<220>  
<221> misc\_特征  
<222> 12  
<223> dT  
<400> 1  
atgcacagtc gg  
<210> 2  
<211> 23  
<212> DNA  
<213> 人工序列  
<220>  
<223> 生物素化靶标探针  
<220>  
<221> misc\_特征  
<222> (12) ... (13)  
<223> heg-F- (heghegheg-F) 3-heg  
<220>  
<221> misc\_特征

12

|                               |    |
|-------------------------------|----|
| <222> 23                      |    |
| <223> dT                      |    |
| <400> 2                       |    |
| cgacgcttac agccgactgt gca     | 23 |
| <210> 3                       |    |
| <211> 24                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 生物素化靶标探针                |    |
| <220>                         |    |
| <221> misc_特征                 |    |
| <222> (12) ... (13)           |    |
| <223> heg-F(heghegheg-F)9-heg |    |
| <220>                         |    |
| <221> misc_特征                 |    |
| <222> 24                      |    |
| <223> dT                      |    |
| <400> 3                       |    |
| ctgtaagcgt cggacattcg cagc    | 24 |
| <210> 4                       |    |
| <211> 21                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 生物素化靶标探针                |    |
| <220>                         |    |
| <221> misc_特征                 |    |
| <222> 1                       |    |
| <223> 生物素-heg                 |    |
| <220>                         |    |
| <221> misc_特征                 |    |
| <222> 21                      |    |
| <223> dT                      |    |
| <400> 4                       |    |
| tttctttgag gtttaggatt c       | 21 |
| <210> 5                       |    |
| <211> 61                      |    |
| <212> DNA                     |    |

|                                                                   |    |
|-------------------------------------------------------------------|----|
| <213> 人工序列                                                        |    |
| <220>                                                             |    |
| <223> 生物素化靶标探针                                                    |    |
| <220>                                                             |    |
| <221> misc_特征                                                     |    |
| <222> 1                                                           |    |
| <223> dT                                                          |    |
| <220>                                                             |    |
| <221> misc_特征                                                     |    |
| <222> 61                                                          |    |
| <223> heg-F                                                       |    |
| <400> 5                                                           |    |
| aaagaaactc caaatcctaa gggttggcct tagggttcag aggttggcct tagggttcag | 60 |
| a                                                                 | 61 |
| <210> 6                                                           |    |
| <211> 17                                                          |    |
| <212> DNA                                                         |    |
| <213> 人工序列                                                        |    |
| <220>                                                             |    |
| <223> 生物素化靶标探针                                                    |    |
| <220>                                                             |    |
| <221> misc_特征                                                     |    |
| <222> 17                                                          |    |
| <223> X-F-dT                                                      |    |
| <400> 6                                                           |    |
| ccaaccggaa tcccaag                                                | 17 |
| <210> 7                                                           |    |
| <211> 17                                                          |    |
| <212> DNA                                                         |    |
| <213> 人工序列                                                        |    |
| <220>                                                             |    |
| <223> 生物素化靶标探针                                                    |    |
| <220>                                                             |    |
| <221> misc_特征                                                     |    |
| <222> 17                                                          |    |
| <223> heg-F(heghegheg-F)3-heg-dT                                  |    |
| <400> 7                                                           |    |
| ccaaccggaa tcccaag                                                | 17 |
| <210> 8                                                           |    |

|                                  |    |
|----------------------------------|----|
| <211> 17                         |    |
| <212> DNA                        |    |
| <213> 人工序列                       |    |
| <220>                            |    |
| <223> 生物素化靶标探针                   |    |
| <220>                            |    |
| <221> misc_特征                    |    |
| <222> 17                         |    |
| <223> heg-F(heghegheg-F)4-heg-dT |    |
| <400> 8                          |    |
| ccaaccggaa tcccaag               | 17 |
| <210> 9                          |    |
| <211> 17                         |    |
| <212> DNA                        |    |
| <213> 人工序列                       |    |
| <220>                            |    |
| <223> 生物素化靶标探针                   |    |
| <220>                            |    |
| <221> misc_特征                    |    |
| <222> 17                         |    |
| <223> heg-F(heghegheg-F)9-heg-dT |    |
| <400> 9                          |    |
| ccaaccggaa tcccaag               | 17 |
| <210> 10                         |    |
| <211> 24                         |    |
| <212> DNA                        |    |
| <213> 人工序列                       |    |
| <220>                            |    |
| <223> 生物素化靶标探针                   |    |
| <220>                            |    |
| <221> misc_特征                    |    |
| <222> 1                          |    |
| <223> d                          |    |
| <220>                            |    |
| <221> misc_特征                    |    |
| <222> (12) ... (13)              |    |
| <223> heg-F(heghegheg-F)3-heg    |    |
| <400> 10                         |    |
| tacgtgtcag ccgacattcg cagc       | 24 |

|                                          |    |
|------------------------------------------|----|
| <210> 11                                 |    |
| <211> 24                                 |    |
| <212> DNA                                |    |
| <213> 人工序列                               |    |
| <220>                                    |    |
| <223> 生物素化靶标探针                           |    |
| <220>                                    |    |
| <221> misc_特征                            |    |
| <222> (12) ... (13)                      |    |
| <223> heg-F(heghegheg-F) 4-heg           |    |
| <220>                                    |    |
| <221> misc_特征                            |    |
| <222> 23 24                              |    |
| <223> dT                                 |    |
| <400> 11                                 |    |
| ctgtaagcgt cgcgacgctt acag               | 24 |
| <210> 12                                 |    |
| <211> 37                                 |    |
| <212> DNA                                |    |
| <213> 人工序列                               |    |
| <220>                                    |    |
| <223> 生物素化靶标探针                           |    |
| <220>                                    |    |
| <221> misc_特征                            |    |
| <222> 1                                  |    |
| <223> dT                                 |    |
| <220>                                    |    |
| <221> misc_特征                            |    |
| <222> (20) ... (21)                      |    |
| <223> heg-F(heghegheg-F) -heg            |    |
| <400> 12                                 |    |
| aaagaaactc caatcctaag ggttggcctt agggttc | 37 |
| <210> 13                                 |    |
| <211> 33                                 |    |
| <212> DNA                                |    |
| <213> 人工序列                               |    |
| <220>                                    |    |
| <223> 生物素化靶标探针                           |    |
| <220>                                    |    |

<221> misc\_特征

<222> (17) ... (18)

<223> heg-F(heghegheg-F) -heg

<220>

<221> misc\_特征

<222> 33

<223> dT

<400> 13

ccaaccggaa tcccaagctt tgaggtttag gat

33

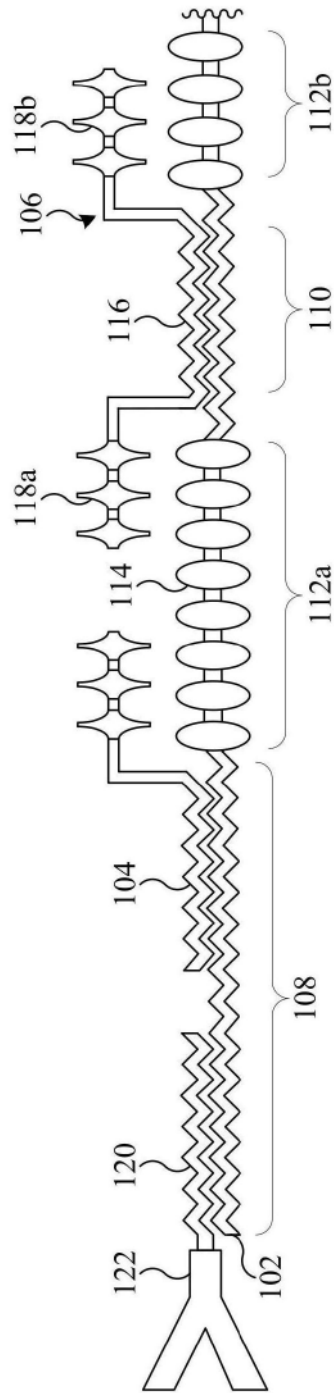


图1

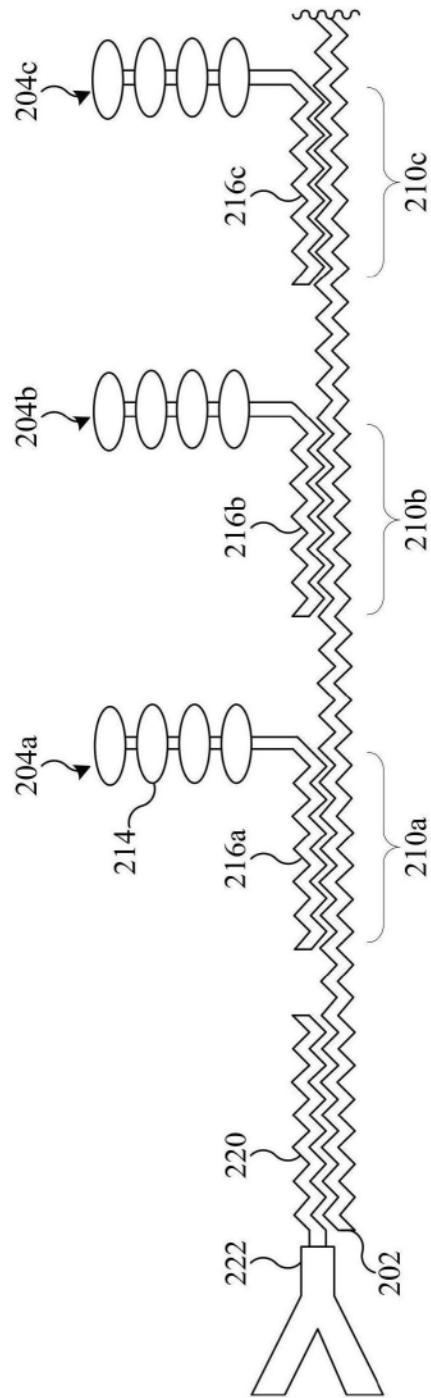


图2A

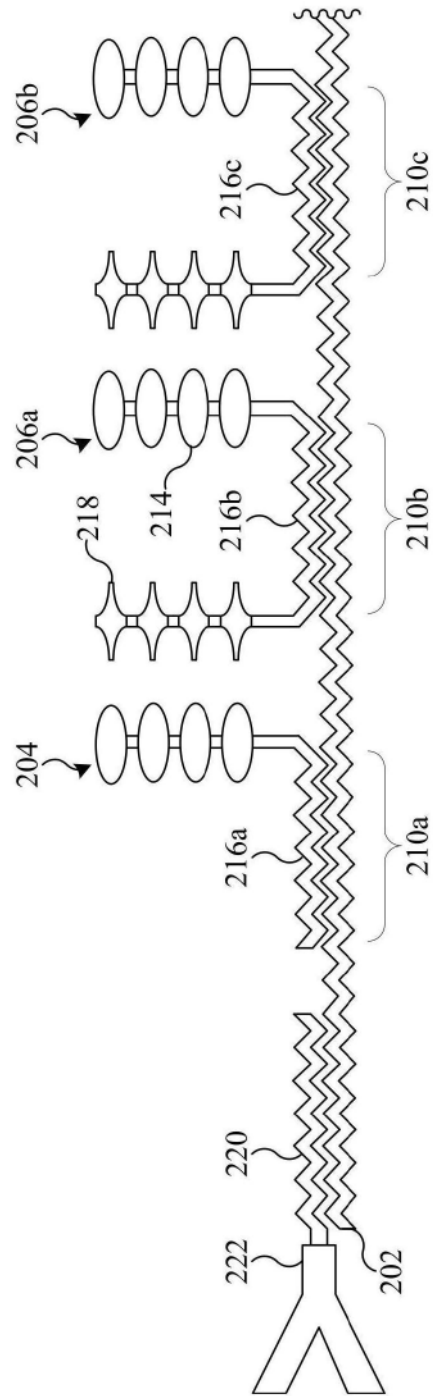


图2B

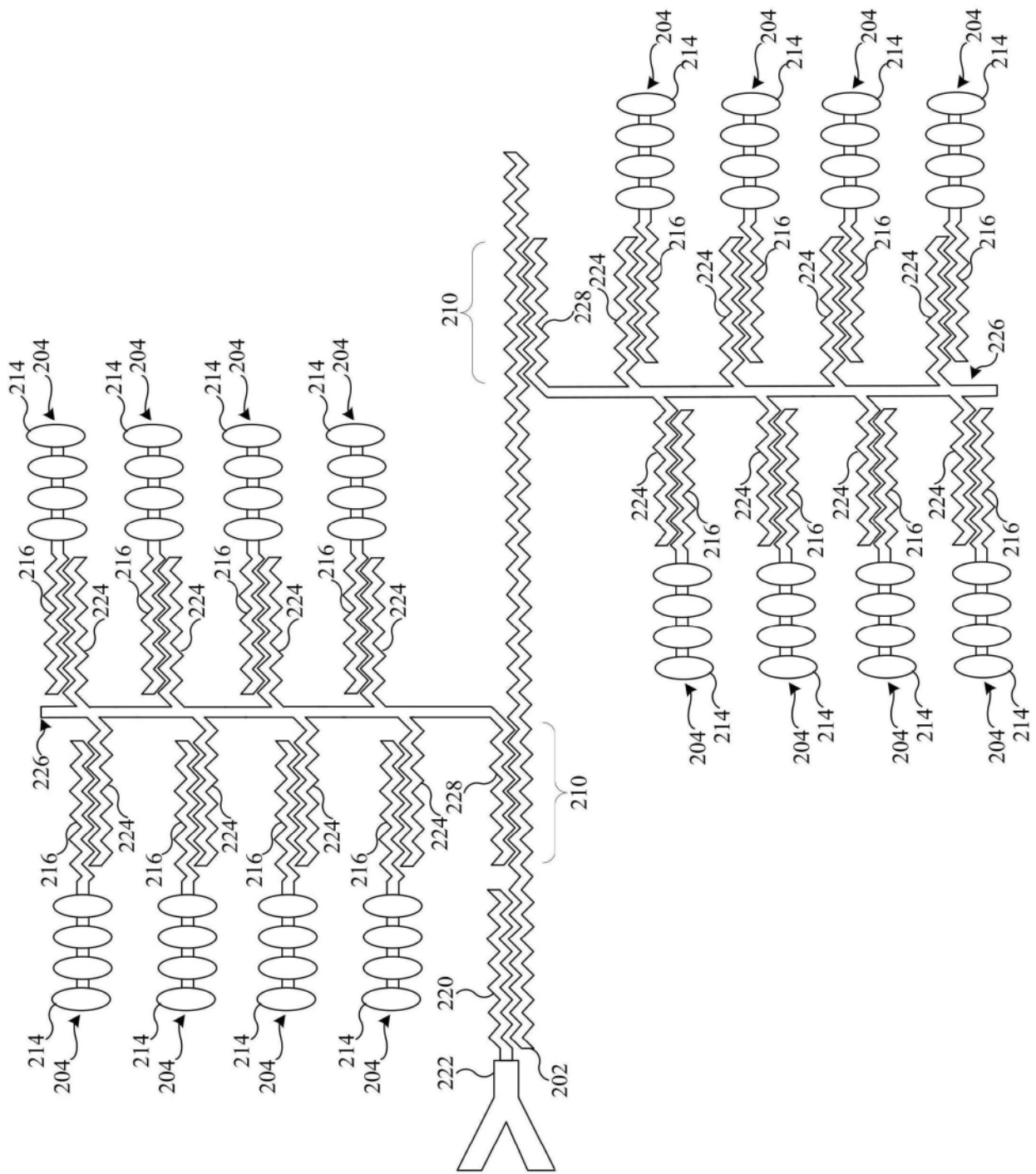


图2C

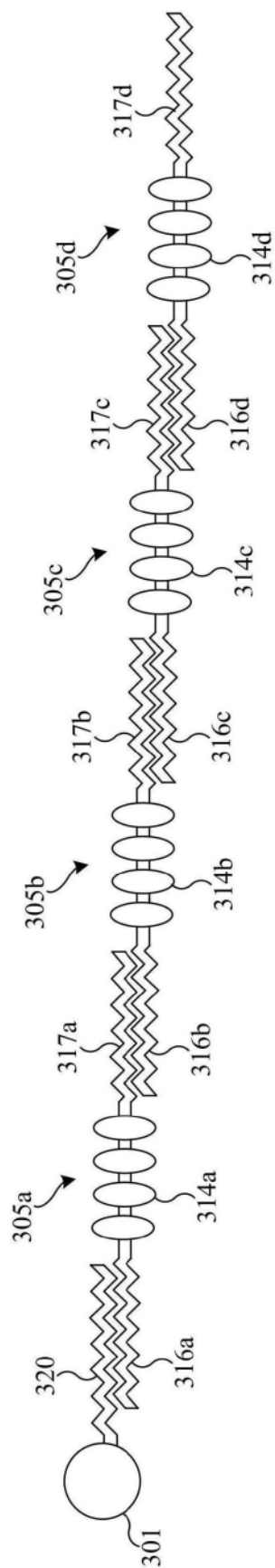


图3A

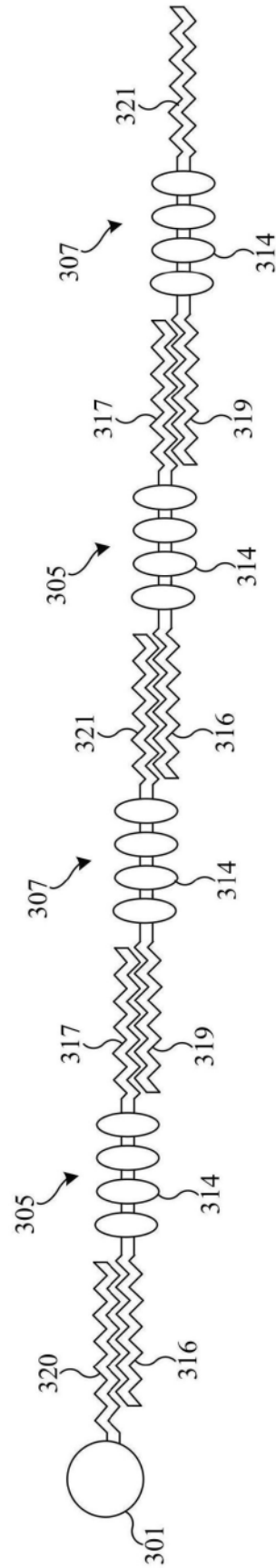


图3B

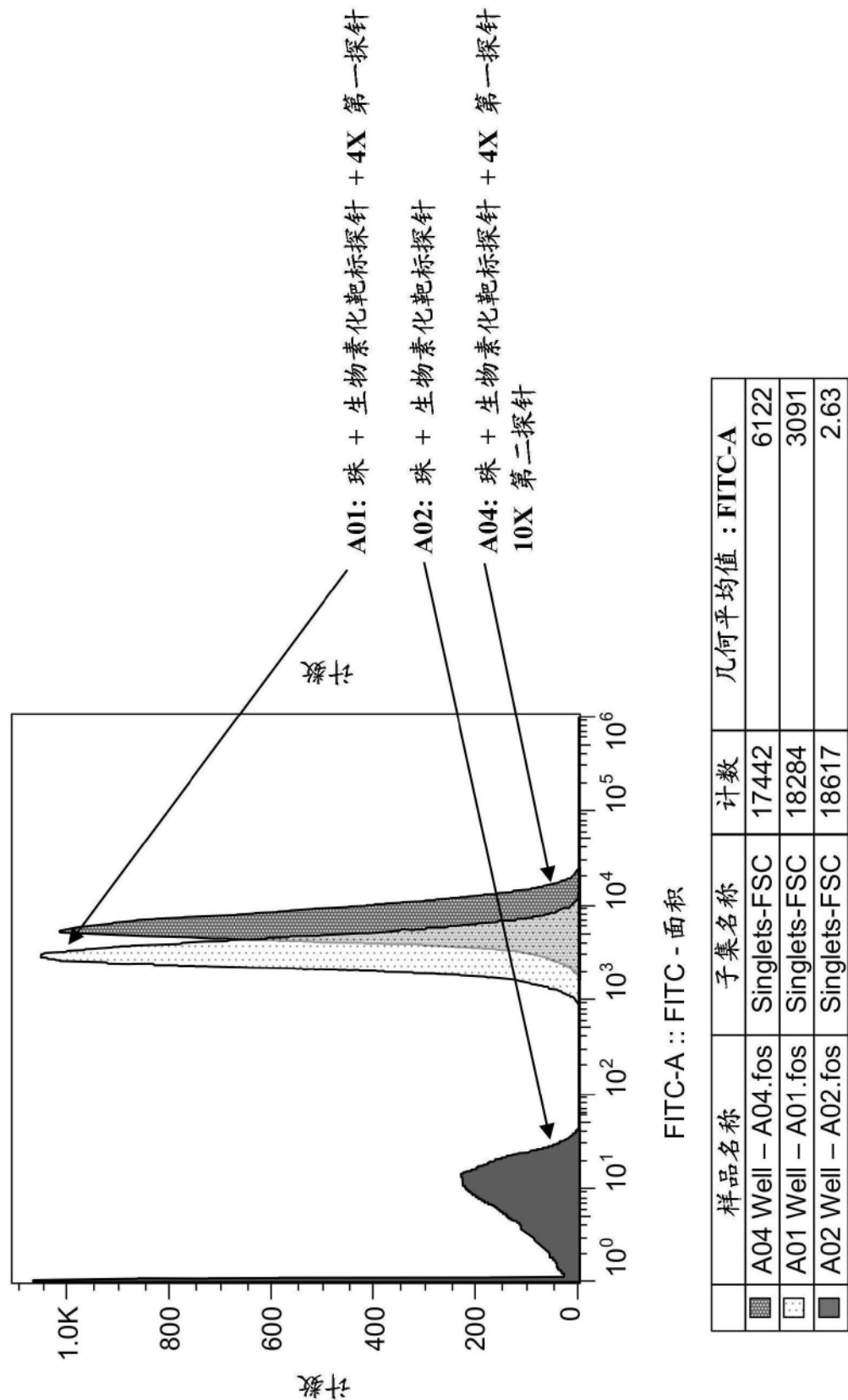


图4

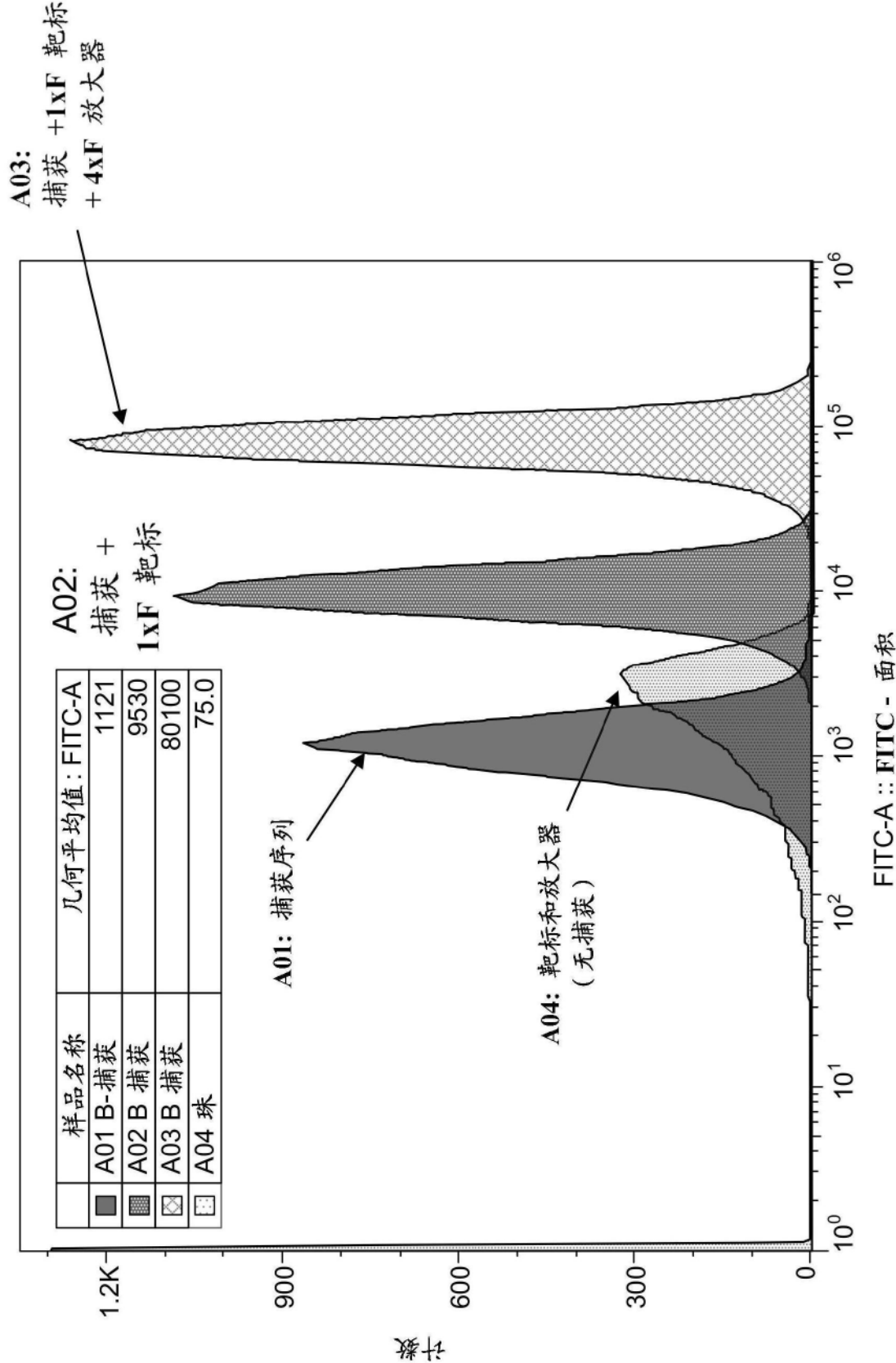
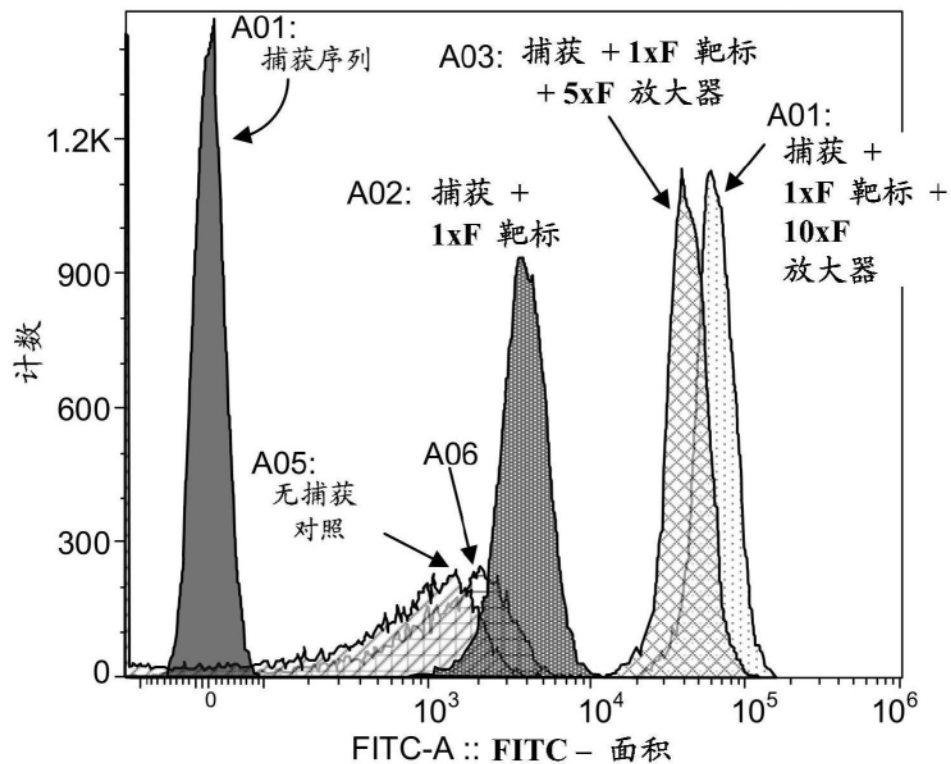


图5



|   | 样品名称     | 子集名称        | 计数    | 几何平均值 : FITC-A |
|---|----------|-------------|-------|----------------|
| ■ | A01 B-捕获 | SSC singlet | 17721 | 13.3           |
| ▤ | A02 B-捕获 | SSC singlet | 17790 | 3822           |
| ▨ | A03 B-捕获 | SSC singlet | 17850 | 42474          |
| ▧ | A04 B-捕获 | SSC singlet | 17556 | 64001          |
| ▩ | A05 珠    | SSC singlet | 17380 | 152            |
| ▦ | A06 珠    | SSC singlet | 17401 | 187            |

图6

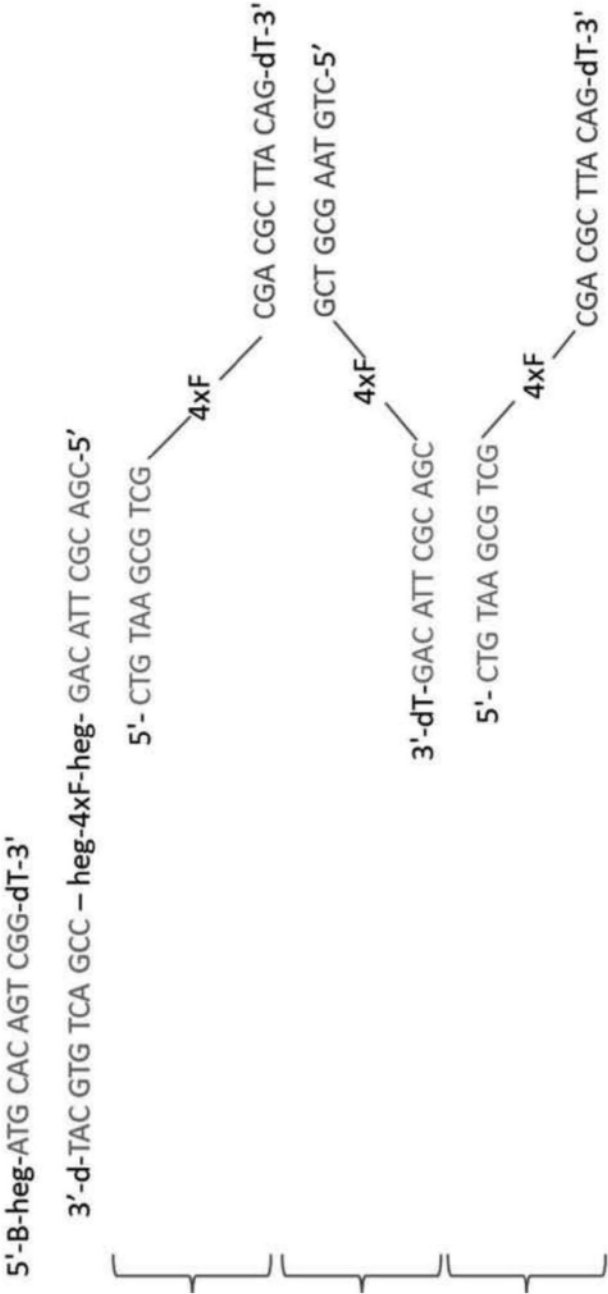


图7

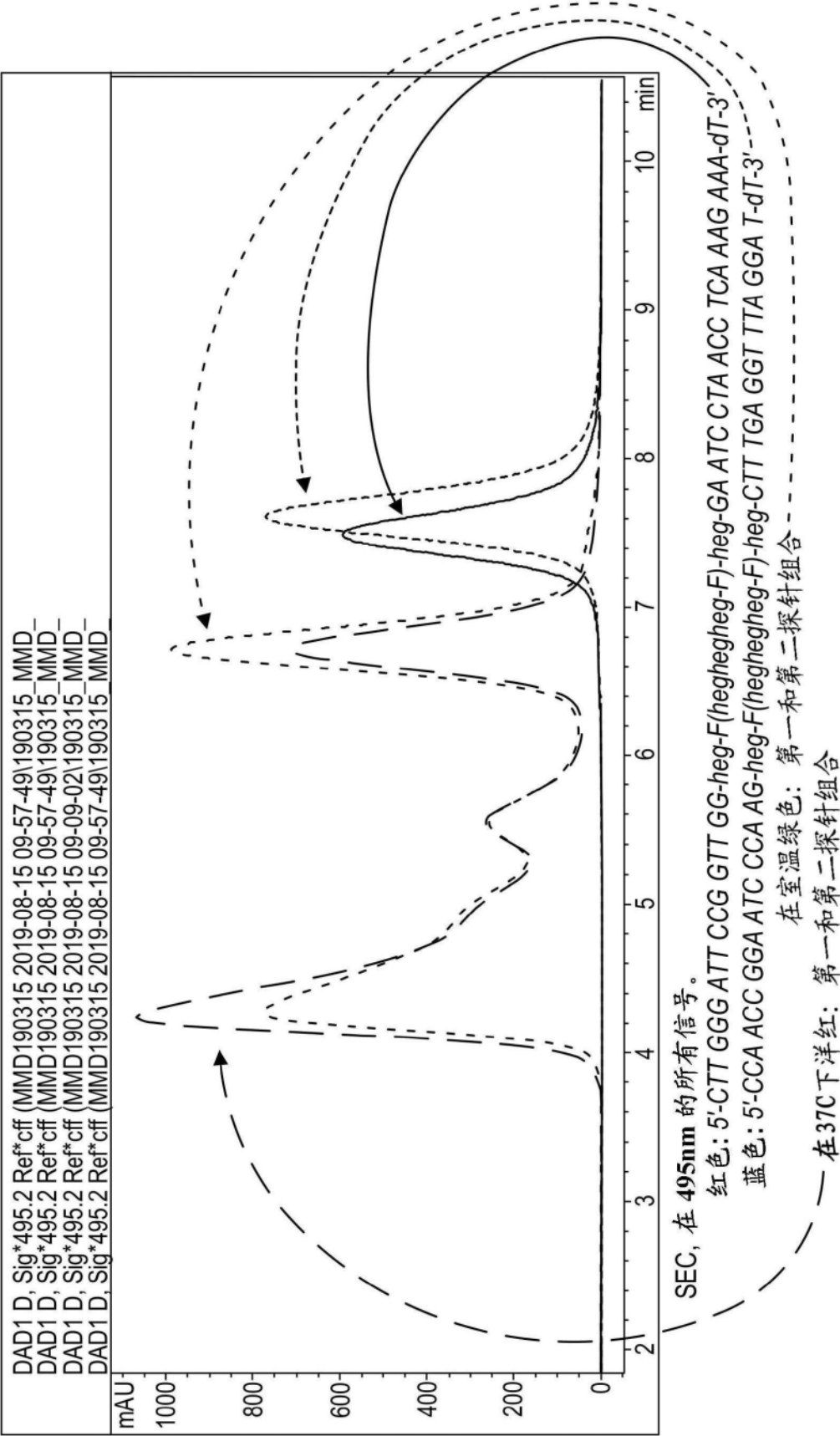
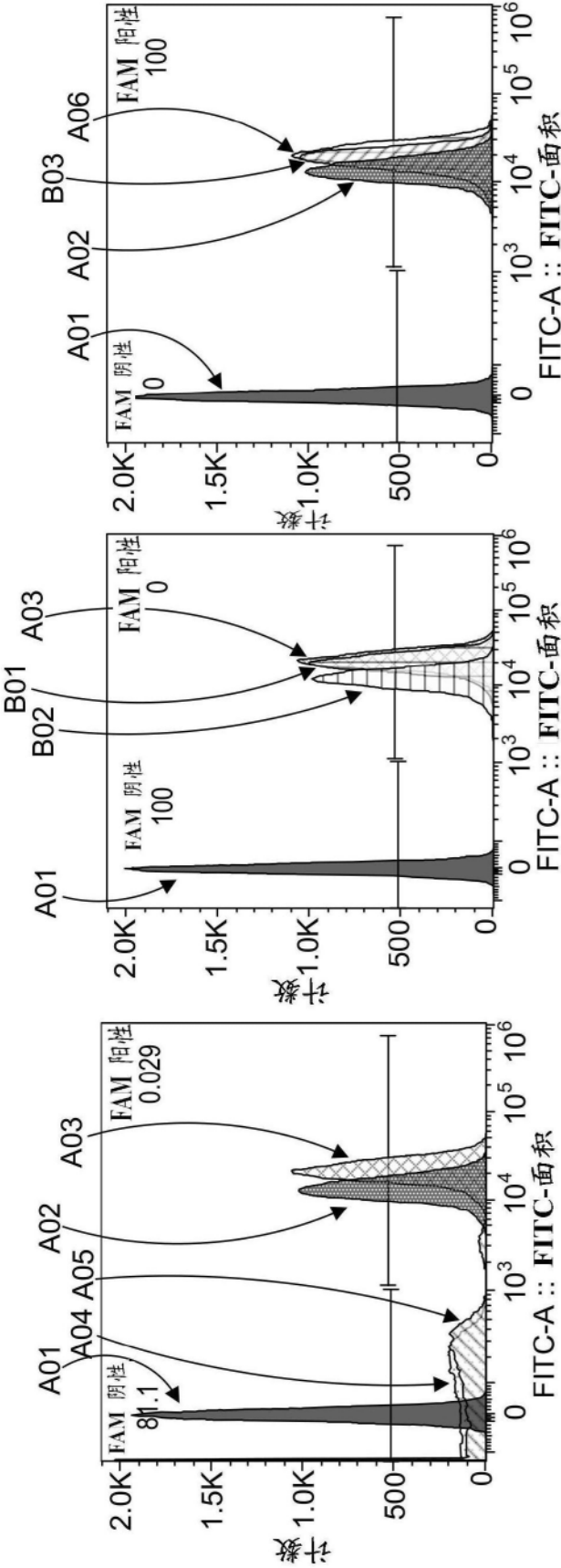


图8A



| FAM 阴性   几何平均值 (FITC-A) | FAM 阳性   几何平均值 (FITC-A) | 减去背景  | 对 2X 的比率 |
|-------------------------|-------------------------|-------|----------|
| 10.1                    | 0                       |       |          |
| 847                     | 13251                   | 12404 | 1.0      |
| 647                     | 21096                   | 20449 | 1.6      |
| 87.4                    | 2311                    | 2224  | 0.2      |
| 117                     | 3108                    | 2991  | 0.2      |
| 0                       | 20602                   | 20602 | 1.7      |
| -20.2                   | 19422                   | 19422 | 1.6      |
| -118                    | 11836                   | 11954 | 1.0      |
| 764                     | 17319                   | 16555 | 1.3      |

样品

|   |                 |
|---|-----------------|
| ■ | A01 B-捕获        |
| ▨ | A02 B-捕获        |
| ▩ | A03 B-捕获        |
| ▧ | A04 B-捕获        |
| ▦ | A05 珠           |
| ▤ | A06 B 捕获预杂交的    |
| ▣ | B01 B 捕获        |
| ▢ | B02 B 捕获        |
| □ | B03 B 捕获预杂交混合物上 |

图8B