

用于测序的核苷酸类似物

技术领域

本发明涉及核酸测序领域。具体地，本发明涉及用于测序的核苷酸类似物。

背景技术

NGS 测序的出现克服了 Sanger 测序成本高、测序时间长等缺点，极大地推动了基因测序技术的应用。目前，NGS 测序已经在产前筛查、肿瘤诊断、肿瘤治疗、动植物育种等领域深度应用，带动了科技与医学的进步。

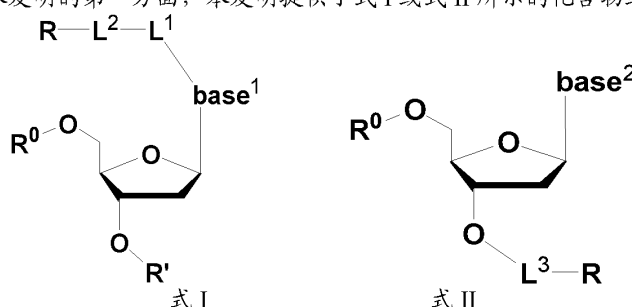
带有可逆阻断基团的核苷三磷酸 (dNTP) 类似物是 NGS 测序中关键原材料。由于可逆阻断基团的引入，使得 dNTP 中 3'-OH 基团可以保留，克服了 Sanger 测序中的缺点，同时保证了碱基识别的准确性。可以说带有可逆阻断基团的核苷三磷酸类似物 (dNTP) 是 NGS 测序中最为关键的技术。而 NGS 测序中带有可切断 linker 与可检测标记的核苷酸更是实现碱基检测的关键所在。

发明内容

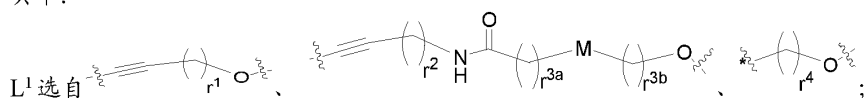
本发明旨在开发一类在碱基上或者核苷酸糖环 3'-OH 带有可切断可检测标记的酯基、碳酸酯等 dNTP 类似物用于 NGS 测序。

第一类 dNTP 类似物的结构通式如图 1 所示，碱基带有可切断可检测标记的核苷酸，包括 2'-脱氧尿苷三磷酸、2'-脱氧胞苷三磷酸、7-脱氮-2'-脱氧腺苷三磷酸与 7-脱氮-2'-脱氧鸟苷三磷酸的主体结构。第一类 dNTP 类似物中，糖环 3'-OH 上连接有本领域常规使用的连接在 3'-OH 上效果优良的任意可逆阻断基团。第二类 dNTP 类似物的结构通式如图 2 所示，糖环 3'-OH 可逆阻断基团上带有可检测标记的核苷酸。

为此，在本发明的第一方面，本发明提供了式 I 或式 II 所示的化合物或其盐，



其中：



r^1 、 r^2 、 r^{3a} 、 r^{3b} 、 r^4 各自独立地选自 1-6 之间的任意整数；

M 选自直接键、CH₂、NH、O、S；

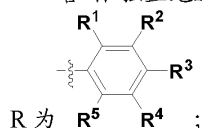
L^2 为 $\begin{matrix} \text{X} \\ \text{Y} \end{matrix}$ ，优选地，W 与 R 相连；

L^3 为 $\begin{matrix} \text{X} \\ \text{W} \end{matrix}$ ，优选地，W 与 R 相连；

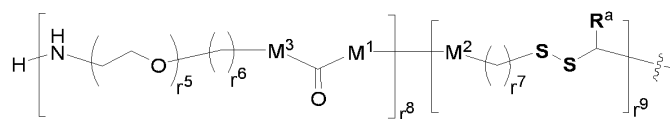
各 X 独立地选自 O、S、NH；

Y 选自直接键、O、S、NH；

各 W 独立地选自直接键、O、S、NH；



R 为 R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个为



、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）、芳香族烷基（如苯基-C1-C6 烷基）、F、I、Br、Cl、烷氧基（如 C1-C6 烷氧基）；

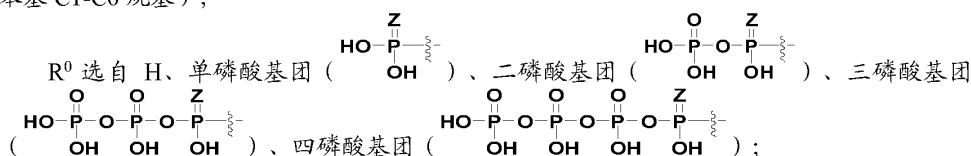
r^5 、 r^6 、 r^7 各自独立地选自 1-6 之间的任意整数;

r^8, r^9 各自独立地选自 0、1, 且 r^8, r^9 不同时为 0;

M¹、M²、M³各自独立地选自直接键、NH、O、S;

R^a 选自 H、脂肪族烷基（如 C1-C6 烷基，例如 Me、Et、iPr、tBu）、芳香族烷基（如苯基-C1-C6 烷基）、环烷基（如 C3-C6 环烷基）；

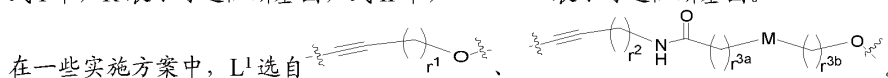
R^b、R^c中,任意一个选自-N₃、-SS-C1-C6烷基(如-SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu)、-ONH₂、-OCOR_x、-OCONHR_x,另一个选自H、脂肪族烷基(如C1-C6烷基,例如Me、Et、iPr、tBu)、芳香族烷基(如苯基-C1-C6烷基)、环烷基(如C3-C6环烷基),其中,各R_x独立地选自脂肪族烷基(如C1-C6烷基),环烷基(如C3-C6环烷基)或芳香族烷基(如苯基-C1-C6烷基);



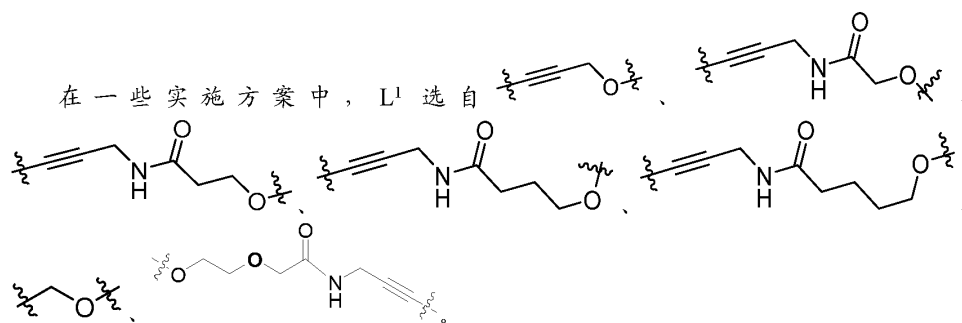
各 Z 独立地选自 O, S, BH:

base¹、base² 各自独立地选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；

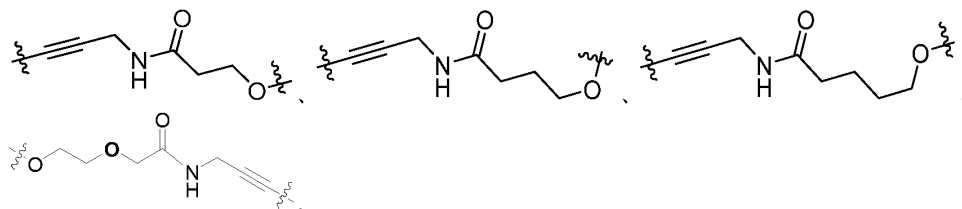
式 I 中, R' 表示可逆阻断基团, 式 II 中, $\text{L}^3\text{---R--}$ 表示可逆阻断基团。

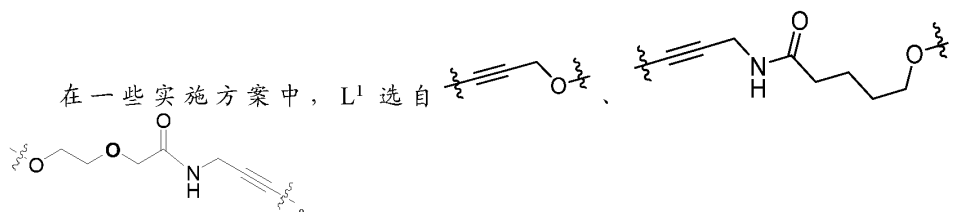


在一些实施方案中， L^1 选自



在一些实施方案中，L¹选自 、





在一些实施方案中， r^1 选自 1、2、3。

在一些实施方案中， r^1 为 1。

在一些实施方案中， r^2 选自 1、2、3。

在一些实施方案中， r^2 为 1。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3。

在一些实施方案中， r^{3a} 为 1。

在一些实施方案中， r^{3b} 为 2。

在一些实施方案中， r^4 选自 1、2、3。

在一些实施方案中， r^4 为 1。

在一些实施方案中， M 选自 CH_2 、 O 。

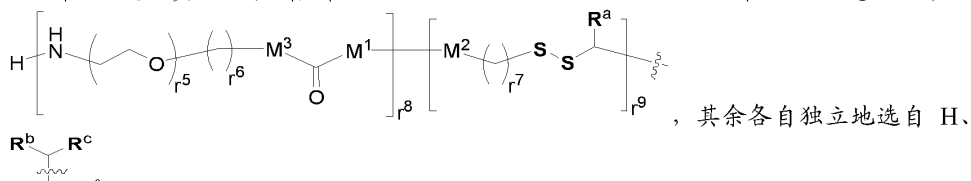
在一些实施方案中，各 X 独立地选自 O 、 S 。

在一些实施方案中， X 为 O 。

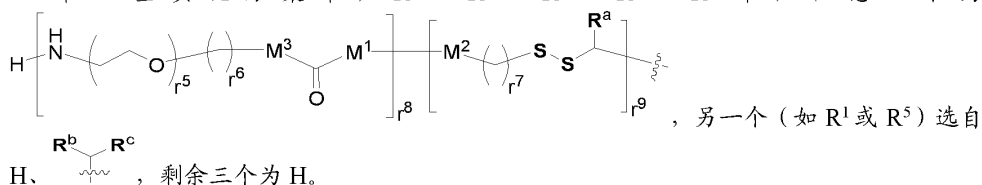
在一些实施方案中， Y 为直接键。

在一些实施方案中， W 为直接键。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个为



在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个为



在一些实施方案中， r^5 选自 1、2、3。

在一些实施方案中， r^5 为 2。

在一些实施方案中， r^6 选自 1、2、3。

在一些实施方案中， r^6 为 1 或 2。

在一些实施方案中， r^7 选自 1、2、3。

在一些实施方案中， r^7 为 2。

在一些实施方案中， M^1 选自直接键、 NH 、 O 。

在一些实施方案中， M^2 为 NH 。

在一些实施方案中， M^3 选自直接键、 NH 。

在一些实施方案中， R^a 选自 H 、 $C1-C6$ 烷基。

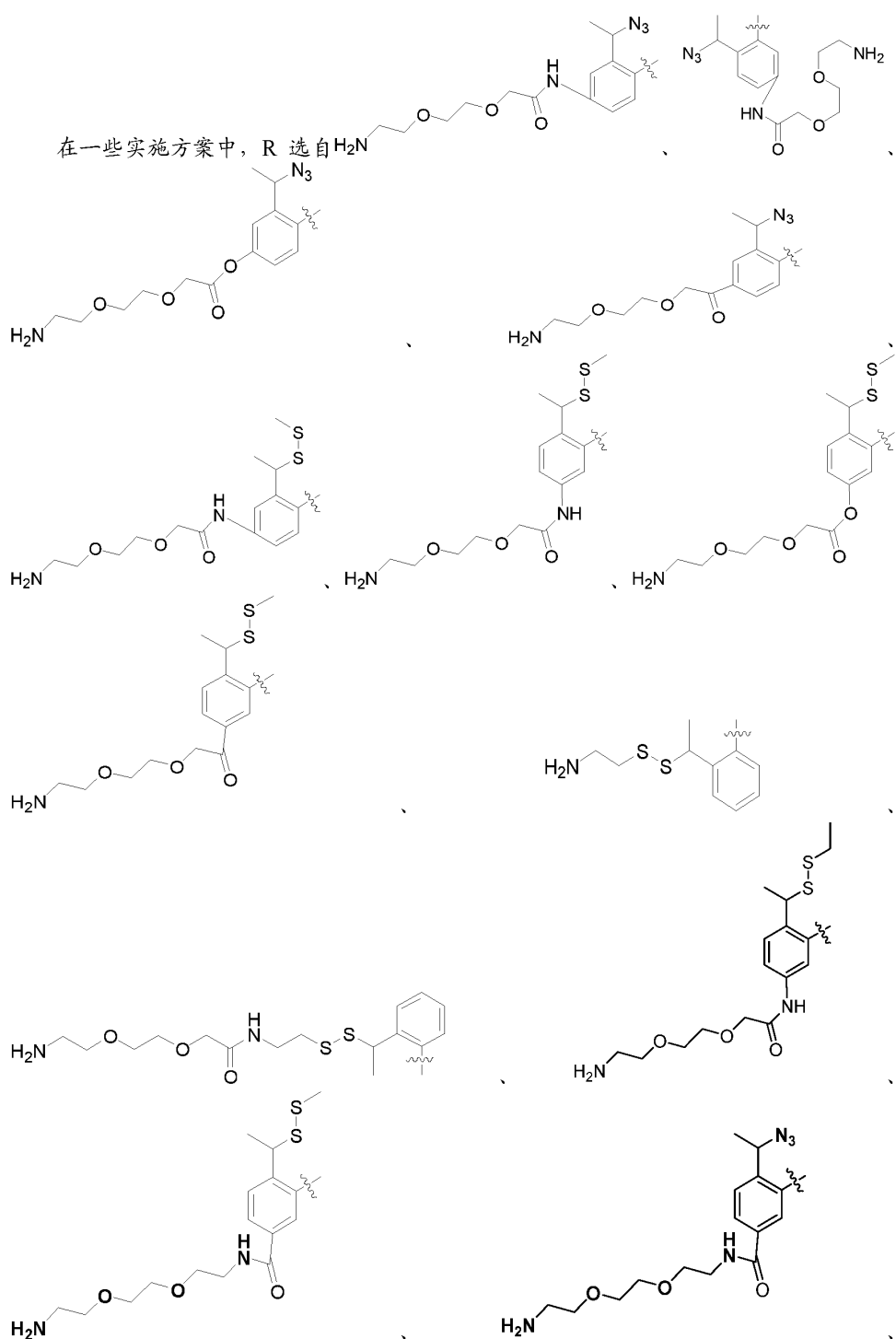
在一些实施方案中， R^a 为甲基。

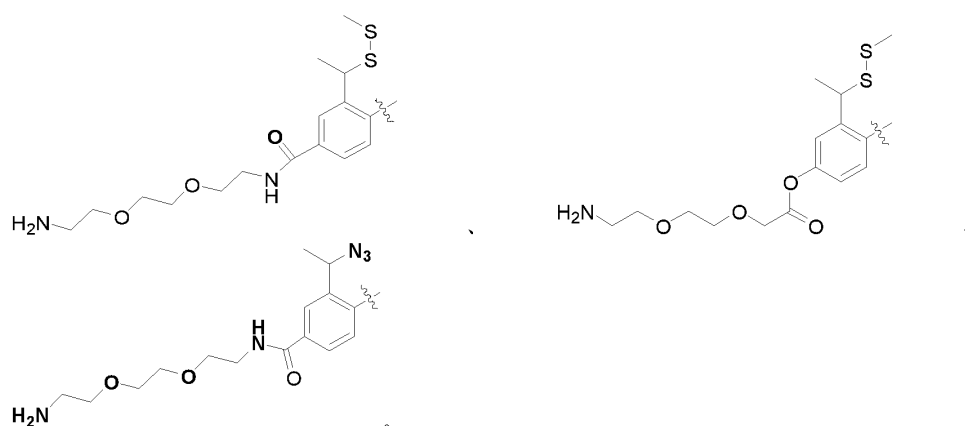
在一些实施方案中， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基（如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$ ），另一个选自 $C1-C6$ 烷基。

在一些实施方案中， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-Me$ 、 $-SS-Et$ ，另一个为甲基。

在一些实施方案中， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-Me$ ，另一个为甲基。

在一些实施方案中，R 选自





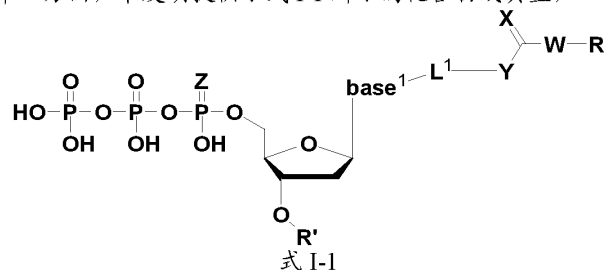
在一些实施方案中, R^0 为三磷酸基团 ($\text{HO}-\text{P}(=\text{O})(\text{OH})-\text{O}-\text{P}(=\text{O})(\text{OH})-\text{O}-\text{P}(=\text{O})(\text{OH})-\text{Z}$)。

在一些实施方案中, Z 为 O。

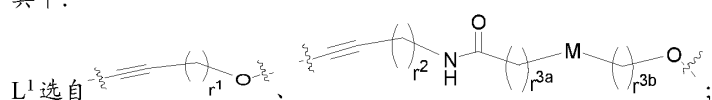
在一些实施方案中, base^1 选自

在一些实施方案中, base^2 选自

在本发明的第二方面, 本发明提供了式 I-1 所示的化合物或其盐,



其中:



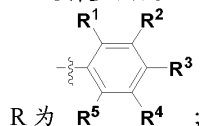
r^1 、 r^2 、 r^{3a} 、 r^{3b} 各自独立地选自 1-6 之间的任意整数;

M 选自直接键、 CH_2 、NH、O、S;

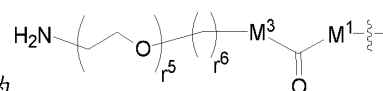
X 选自 O、S、NH;

Y 选自直接键、O、S、NH;

W 选自直接键、O、S、NH;



R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个(如 R^2 、 R^3 或 R^4)为



其余各自独立地选自 H、、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）、芳香族烷基（如苯基-C1-C6 烷基）、F、I、Br、Cl、烷氧基（如 C1-C6 烷氧基）；

r^5 、 r^6 各自独立地选自 1-6 之间的任意整数；

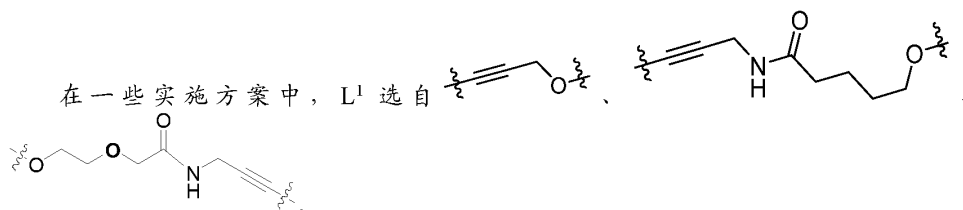
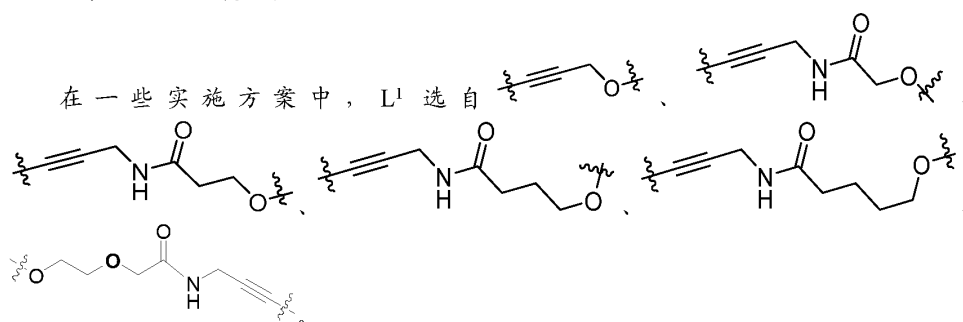
M^1 、 M^3 各自独立地选自直接键、NH、O、S；

R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基（如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$ ）、 $-ONH_2$ 、 $-OCOR_x$ 、 $-OCONHR_x$ ，另一个选自 H、脂肪族烷基（如 C1-C6 烷基，例如 Me、Et、iPr、tBu）、芳香族烷基（如苯基-C1-C6 烷基）、环烷基（如 C3-C6 环烷基），其中，各 R_x 独立地选自脂肪族烷基（如 C1-C6 烷基），环烷基（如 C3-C6 环烷基）或芳香族烷基（如苯基 C1-C6 烷基）；

Z 选自 O、S、BH；

base¹ 选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；

R' 表示可逆阻断基团。



在一些实施方案中， r^1 选自 1、2、3。

在一些实施方案中， r^1 为 1。

在一些实施方案中， r^2 选自 1、2、3。

在一些实施方案中， r^2 为 1。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3。

在一些实施方案中， r^{3a} 为 1。

在一些实施方案中， r^{3b} 为 2。

在一些实施方案中，M 选自 CH_2 、O。

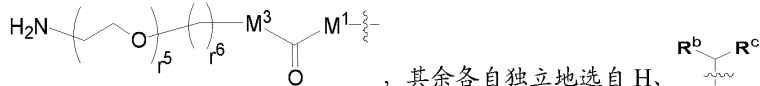
在一些实施方案中，X 选自 O、S。

在一些实施方案中，X 为 O。

在一些实施方案中，Y 为直接键。

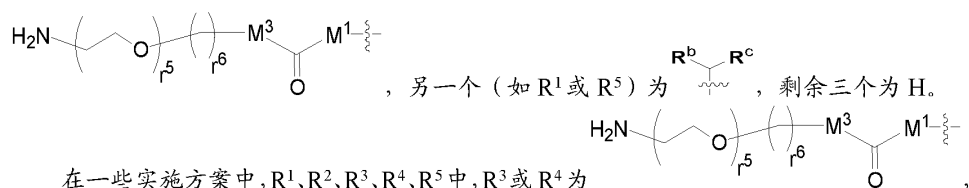
在一些实施方案中，W 为直接键。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个（如 R^2 、 R^3 或 R^4 ）为

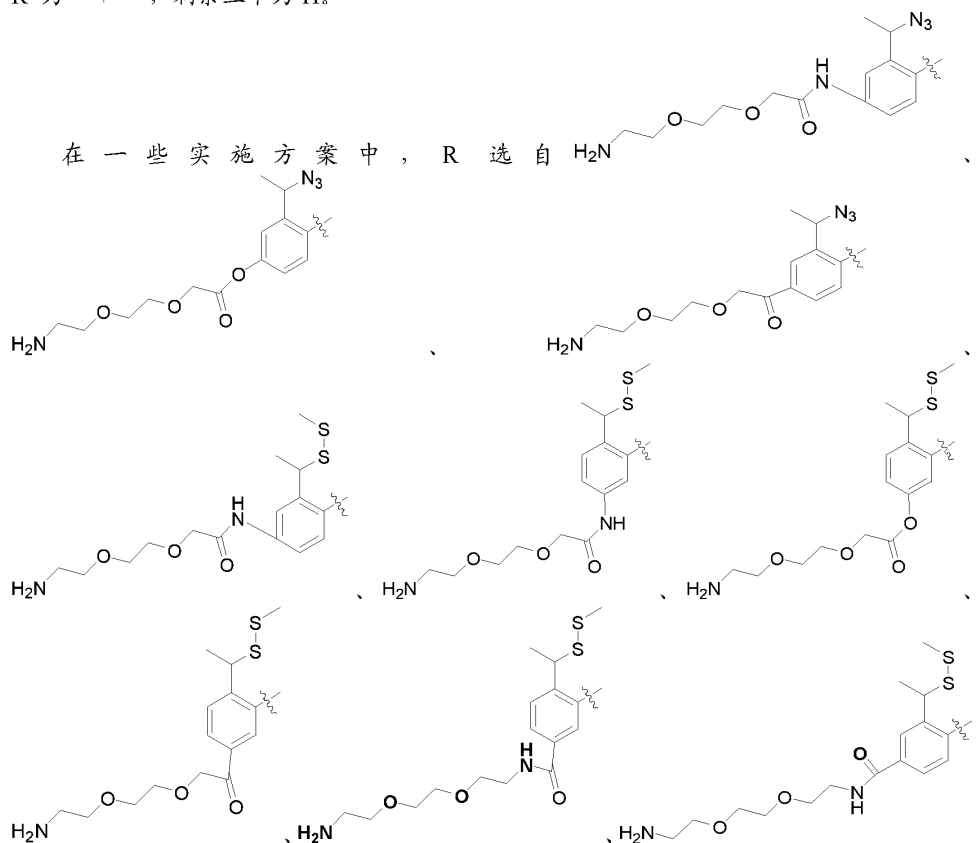


其余各自独立地选自 H、。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个（如 R^2 、 R^3 或 R^4 ）为



在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中， R^3 或 R^4 为 R^b-R^c ， R^1 为 R^b-R^c ，剩余三个为 H。



在一些实施方案中， r^5 选自 1、2、3。

在一些实施方案中， r^5 为 2。

在一些实施方案中， r^6 选自 1、2、3。

在一些实施方案中， r^6 为 1 或 2。

在一些实施方案中， M^1 选自直接键、NH、O。

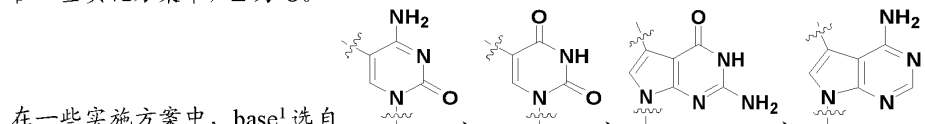
在一些实施方案中， M^3 选自直接键、NH。

在一些实施方案中， M^1 为 NH、O 或 S 时， M^3 为直接键，且 M^3 为 NH、O 或 S 时， M^1 为直接键。

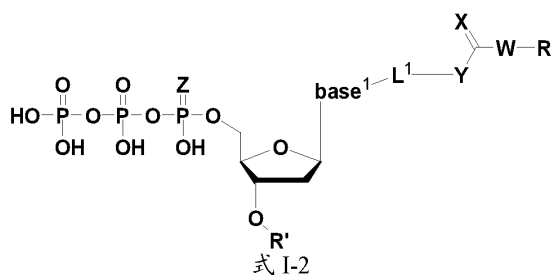
在一些实施方案中， R^b 、 R^c 中，任意一个选自 $-\text{N}_3$ 、 $-\text{SS}-\text{C}_1-\text{C}_6$ 烷基（如 $-\text{SS}-\text{Me}$ 、 $-\text{SS}-\text{Et}$ 、 $-\text{SS}-\text{iPr}$ 、 $-\text{SS}-\text{t-Bu}$ ），另一个选自 C_1-C_6 烷基。

在一些实施方案中， R^b 、 R^c 中，任意一个选自 $-\text{N}_3$ 、 $-\text{SS}-\text{Me}$ ，另一个为甲基。

在一些实施方案中，Z 为 O。



在本发明的第三方面，本发明提供了式 I-2 所示的化合物或其盐，



其中:



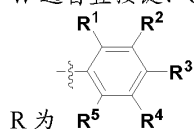
r^1 、 r^2 、 r^{3a} 、 r^{3b} 各自独立地选自 1-6 之间的任意整数;

M 选自直接键、CH₂、NH、O、S;

X 选自 O、S、NH;

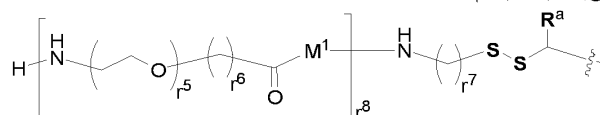
Y 选自直接键、O、S、NH;

W 选自直接键、O、S、NH;



R 为 R⁵、R⁴、R³、R²、R¹;

R¹、R²、R³、R⁴、R⁵中,任意一个(如 R¹或 R⁵)为



其余各自独立地选自 H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基(如 C1-C6 烷基)、环烷基(如 C3-C6 环烷基)、芳香族烷基(如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基(如 C1-C6 烷氧基);

r^5 、 r^6 、 r^7 各自独立地选自 1-6 之间的任意整数;

r^8 选自 0、1,

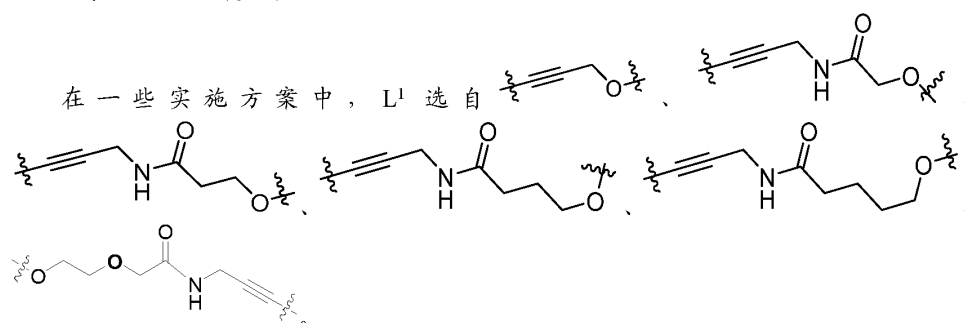
M¹选自直接键、NH、O、S;

R^a选自 H、脂肪族烷基(如 C1-C6 烷基,例如 Me、Et、iPr、tBu)、芳香族烷基(如苯基-C1-C6 烷基)、环烷基(如 C3-C6 环烷基);

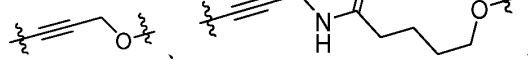
Z 选自 O、S、BH;

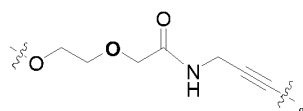
base¹选自碱基、脱氮碱基或其互变异构体,例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;

R'表示可逆阻断基团。



在一些实施方案中, L¹ 选自





在一些实施方案中， r^1 选自 1、2、3。

在一些实施方案中， r^1 为 1。

在一些实施方案中， r^2 选自 1、2、3。

在一些实施方案中， r^2 为 1。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3。

在一些实施方案中， r^{3a} 为 1。

在一些实施方案中， r^{3b} 为 2。

在一些实施方案中，M 选自 CH_2 、O。

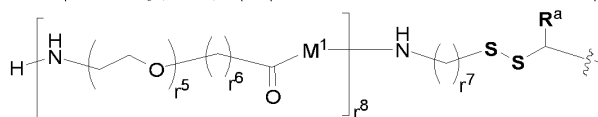
在一些实施方案中，X 选自 O、S。

在一些实施方案中，X 为 O。

在一些实施方案中，Y 为直接键。

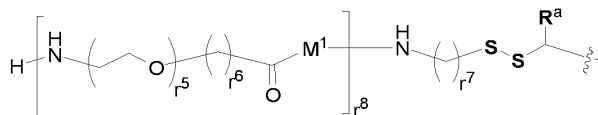
在一些实施方案中，W 为直接键。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个（如 R^1 或 R^5 ）为



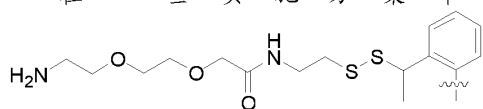
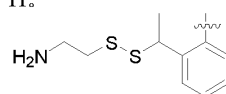
，其余为 H。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中， R^1 为



，其余为 H。

在一些实施方案中，R 选自



在一些实施方案中， r^5 选自 1、2、3。

在一些实施方案中， r^5 为 2。

在一些实施方案中， r^6 选自 1、2、3。

在一些实施方案中， r^6 为 1。

在一些实施方案中， r^7 选自 1、2、3。

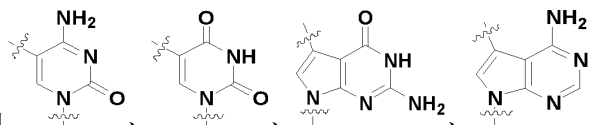
在一些实施方案中， r^7 为 2。

在一些实施方案中， M^1 为直接键。

在一些实施方案中， R^a 选自 C1-C6 烷基。

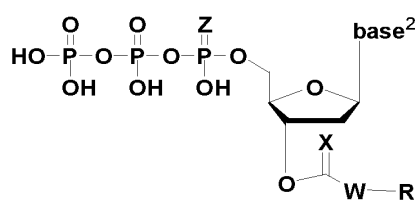
在一些实施方案中， R^a 为甲基。

在一些实施方案中，Z 为 O。



在一些实施方案中， base^1 选自

在本发明的第四方面，本发明提供了式 II-1 所示的化合物或其盐，

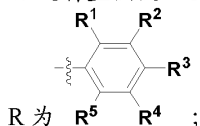


式 II-1

其中:

X 选自 O、S、NH;

W 选自直接键、O、S、NH;



R 为

R¹、R²、R³、R⁴、R⁵ 中,任意一个(如 R³ 或 R⁴)为



其余各自独立地选自 H、、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基(如 C1-C6 烷基)、环烷基(如 C3-C6 环烷基)、芳香族烷基(如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基(如 C1-C6 烷氧基);

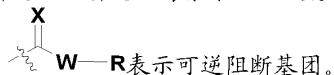
r⁵、r⁶ 各自独立地选自 1-6 之间的任意整数;

M¹、M³ 各自独立地选自直接键、NH、O、S;

Rᵇ、Rᶜ 中,任意一个选自 -N₃、-SS-C1-C6 烷基(如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu)、-ONH₂、-OCORₓ、-OCONHRₓ, 另一个选自 H、脂肪族烷基(如 C1-C6 烷基,例如 Me、Et、iPr、tBu)、芳香族烷基(如苯基-C1-C6 烷基)、环烷基(如 C3-C6 环烷基), 其中,各 Rₓ 独立地选自脂肪族烷基(如 C1-C6 烷基),环烷基(如 C3-C6 环烷基)或芳香族烷基(如苯基 C1-C6 烷基);

Z 选自 O, S, BH;

base² 选自碱基、脱氮碱基或其互变异构体,例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;

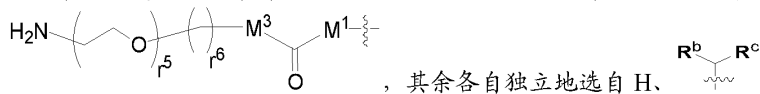


在一些实施方案中, X 选自 O、S。

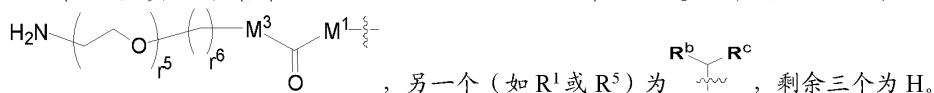
在一些实施方案中, X 为 O。

在一些实施方案中, W 为直接键。

在一些实施方案中, R¹、R²、R³、R⁴、R⁵ 中,任意一个(如 R³ 或 R⁴)为

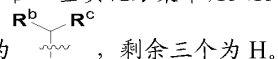


在一些实施方案中, R¹、R²、R³、R⁴、R⁵ 中,任意一个(如 R³ 或 R⁴)为

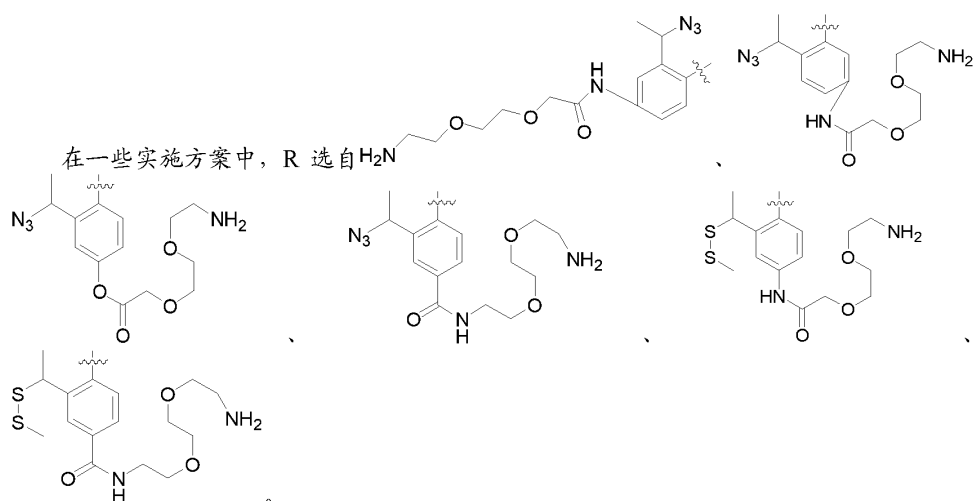


另一个(如 R¹ 或 R⁵)为 , 剩余三个为 H。

在一些实施方案中, R¹、R²、R³、R⁴、R⁵ 中, R³ 或 R⁴ 为



R¹ 为 , 剩余三个为 H。



在一些实施方案中, r^5 选自 1、2、3。

在一些实施方案中, r^5 为 2。

在一些实施方案中, r^6 选自 1、2、3。

在一些实施方案中, r^6 选自 1、2。

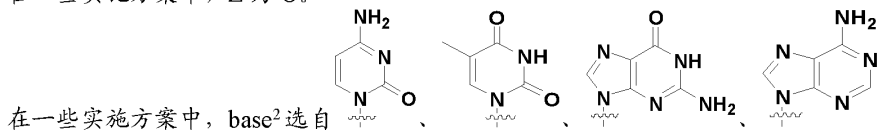
在一些实施方案中, M^1 选自直接键、NH、O。

在一些实施方案中, M^3 选自直接键、NH。

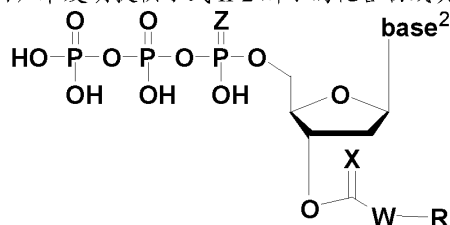
在一些实施方案中, R^b 、 R^c 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基 (如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$) , 另一个选自 C1-C6 烷基。

在一些实施方案中, R^b 、 R^c 中, 任意一个为 $-N_3$ 或 $-SS-Me$, 另一个为甲基。

在一些实施方案中, Z 为 O。



在本发明的第五方面, 本发明提供了式 II-2 所示的化合物或其盐,

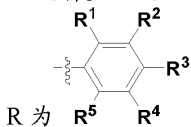


式 II-2

其中:

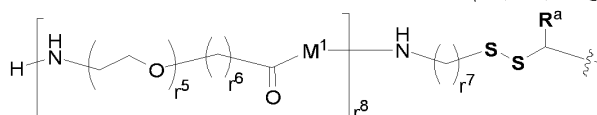
X 选自 O、S、NH;

W 选自直接键、O、S、NH;



R 为 R^5 ;

R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^1 或 R^5) 为



, 其余各自独立地选自 H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基)、

芳香族烷基（如苯基-C1-C6 烷基）、F、I、Br、Cl、烷氧基（如 C1-C6 烷氧基）；

r^5 、 r^6 、 r^7 选自 1-6 之间的任意整数；

r^8 选自 0、1；

M^1 选自直接键、NH、O、S；

R^a 选自 H、脂肪族烷基（如 C1-C6 烷基，例如 Me、Et、iPr、tBu）、芳香族烷基（如苯基-C1-C6 烷基）、环烷基（如 C3-C6 环烷基）；

Z 选自 O、S、BH；

$base^2$ 选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；



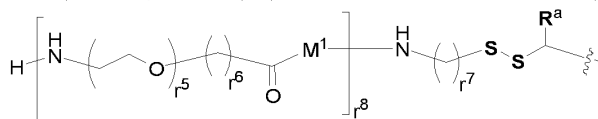
W—R 表示可逆阻断基团。

在一些实施方案中，X 选自 O、S。

在一些实施方案中，X 为 O。

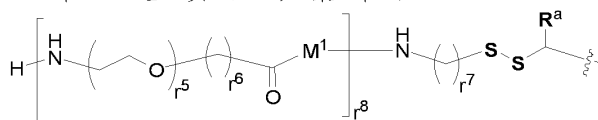
在一些实施方案中，W 为直接键。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个（如 R^1 或 R^5 ）为

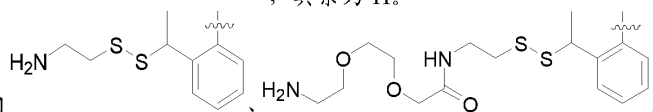


，其余为 H。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中， R^1 为



，其余为 H。



在一些实施方案中，R 选自

在一些实施方案中， r^5 选自 1、2、3。

在一些实施方案中， r^5 为 2。

在一些实施方案中， r^6 选自 1、2、3。

在一些实施方案中， r^6 为 1。

在一些实施方案中， r^7 选自 1、2、3。

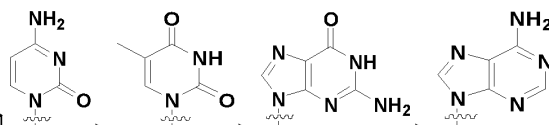
在一些实施方案中， r^7 为 2。

在一些实施方案中， M^1 为直接键。

在一些实施方案中， R^a 选自 C1-C6 烷基。

在一些实施方案中， R^a 为甲基。

在一些实施方案中，Z 为 O。



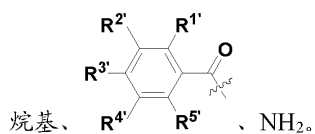
在一些实施方案中， $base^2$ 选自

在一些实施方案中，前述可逆阻断基团 R' 选自 N_3 -C1-C6 烷基、C1-C6 烷基-SS-C1-C6

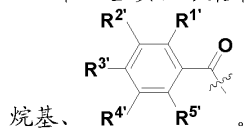


烷基、 R^4 、 R^5 、 NH_2 、 $-ONH_2$ 、 $-OCOR_z$ 、 $-OCONHR_z$ ，其中，各 R_z 独立地选自脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）或芳香族烷基（如苯基 C1-C6 烷基）。

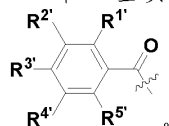
在一些实施方案中，所述可逆阻断基团 R' 选自 N_3 -C1-C6 烷基、C1-C6 烷基-SS-C1-C6



在一些实施方案中，所述可逆阻断基团 R' 选自 N_3 -C1-C6 烷基、C1-C6 烷基-SS-C1-C6



在一些实施方案中，所述可逆阻断基团 R' 选自 N_3 -CH₂-、CH₃-CH₂-S-S-CH₂-、



在一些实施方案中， $\text{R}^{1'}$ 、 $\text{R}^{2'}$ 、 $\text{R}^{3'}$ 、 $\text{R}^{4'}$ 、 $\text{R}^{5'}$ 中，任意一个（如 $\text{R}^{1'}$ 或 $\text{R}^{5'}$ ）为 , 其余各自独立地选自 H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）、芳香族烷基（如苯基-C1-C6 烷基）、F、I、Br、Cl、烷氧基（如 C1-C6 烷氧基）。

在一些实施方案中， $\text{R}^{1'}$ 、 $\text{R}^{2'}$ 、 $\text{R}^{3'}$ 、 $\text{R}^{4'}$ 、 $\text{R}^{5'}$ 中，任意一个（如 $\text{R}^{1'}$ 或 $\text{R}^{5'}$ ）为 , 其余为 H。

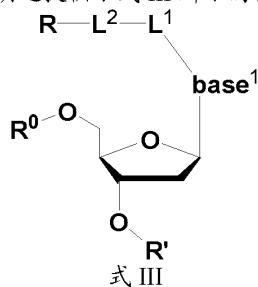
在一些实施方案中， $\text{R}^{1'}$ 、 $\text{R}^{2'}$ 、 $\text{R}^{3'}$ 、 $\text{R}^{4'}$ 、 $\text{R}^{5'}$ 中， $\text{R}^{1'}$ 为 , 其余为 H。

在一些实施方案中， $\text{R}^{b'}$ 、 $\text{R}^{c'}$ 中，任意一个选自 -N_3 、-SS-C1-C6 烷基（如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu）、-ONH₂、-OCOR_x、-OCONHR_x，另一个选自 H、脂肪族烷基（如 C1-C6 烷基，例如 Me、Et、iPr、tBu）、芳香族烷基（如苯基-C1-C6 烷基）、环烷基（如 C3-C6 环烷基），其中，各 R_x 独立地选自脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）或芳香族烷基（如苯基 C1-C6 烷基）。

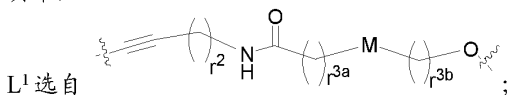
在一些实施方案中， $\text{R}^{b'}$ 、 $\text{R}^{c'}$ 中，任意一个选自 -N_3 、-SS-C1-C6 烷基（如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu），另一个选自 C1-C6 烷基。

在一些实施方案中， $\text{R}^{b'}$ 、 $\text{R}^{c'}$ 中，任意一个选自 -N_3 、-SS-Me，另一个为甲基。

在本发明的第六方面，本发明还提供了式 III 所示的化合物或其盐，



其中：



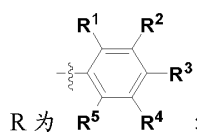
r^2 选自 1、2、3；

r^{3a} 选自 1、2、3；

r^{3b} 选自 0、1、2、3；

M 选自直接键、O；

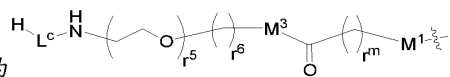
L^2 为 ；



R¹、R²、R³、R⁴、R⁵ 中,任意一个(如 R³)为



另一个(如 R¹ 或 R²)选自 , 剩余三个为 H;



L^c 选自直接键、

L^c 中的 NH 端与 H 相连, L^c 中的=O 端与 NH 相连;

r^m 选自 0、1、2、3;

r⁵ 选自 1、2、3;

r⁶ 选自 1、2、3;

r¹⁰ 选自 1-10 之间的任意整数;

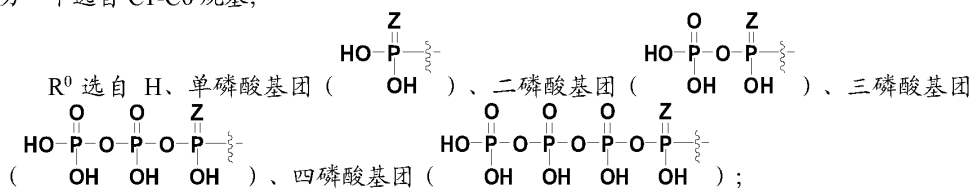
r¹¹ 选自 1-6 之间的任意整数;

M¹ 选自 NH、O;

M¹ 选自 NH、O;

M³ 选自直接键、NH;

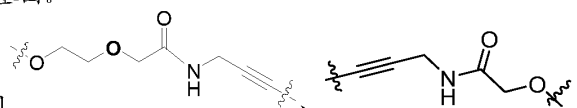
R^b、R^c 中,任意一个选自 -SS-C1-C6 烷基(如 -SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu), 另一个选自 C1-C6 烷基;



各 Z 独立地选自 O, S, BH;

base¹ 选自碱基、脱氮碱基或其互变异构体,例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;

式 III 中, R' 表示可逆阻断基团。



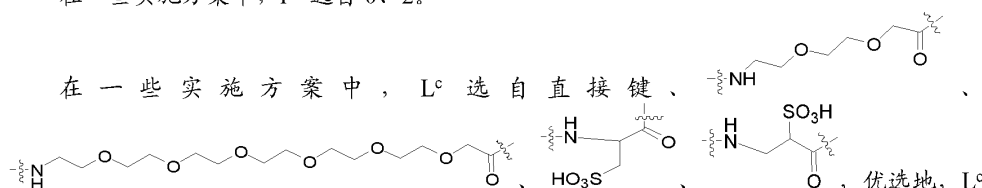
在一些实施方案中, L¹ 选自

在一些实施方案中, r² 为 1。

在一些实施方案中, r^{3a} 为 1。

在一些实施方案中, r^{3b} 选自 0、2。

在一些实施方案中, L^c 选自直接键、



中的 NH 端与 H 相连, L^c 中的=O 端与 NH 相连。

在一些实施方案中, r^m 选自 0、1。

在一些实施方案中, r⁵ 为 2。

在一些实施方案中, r⁶ 为 1 或 2。

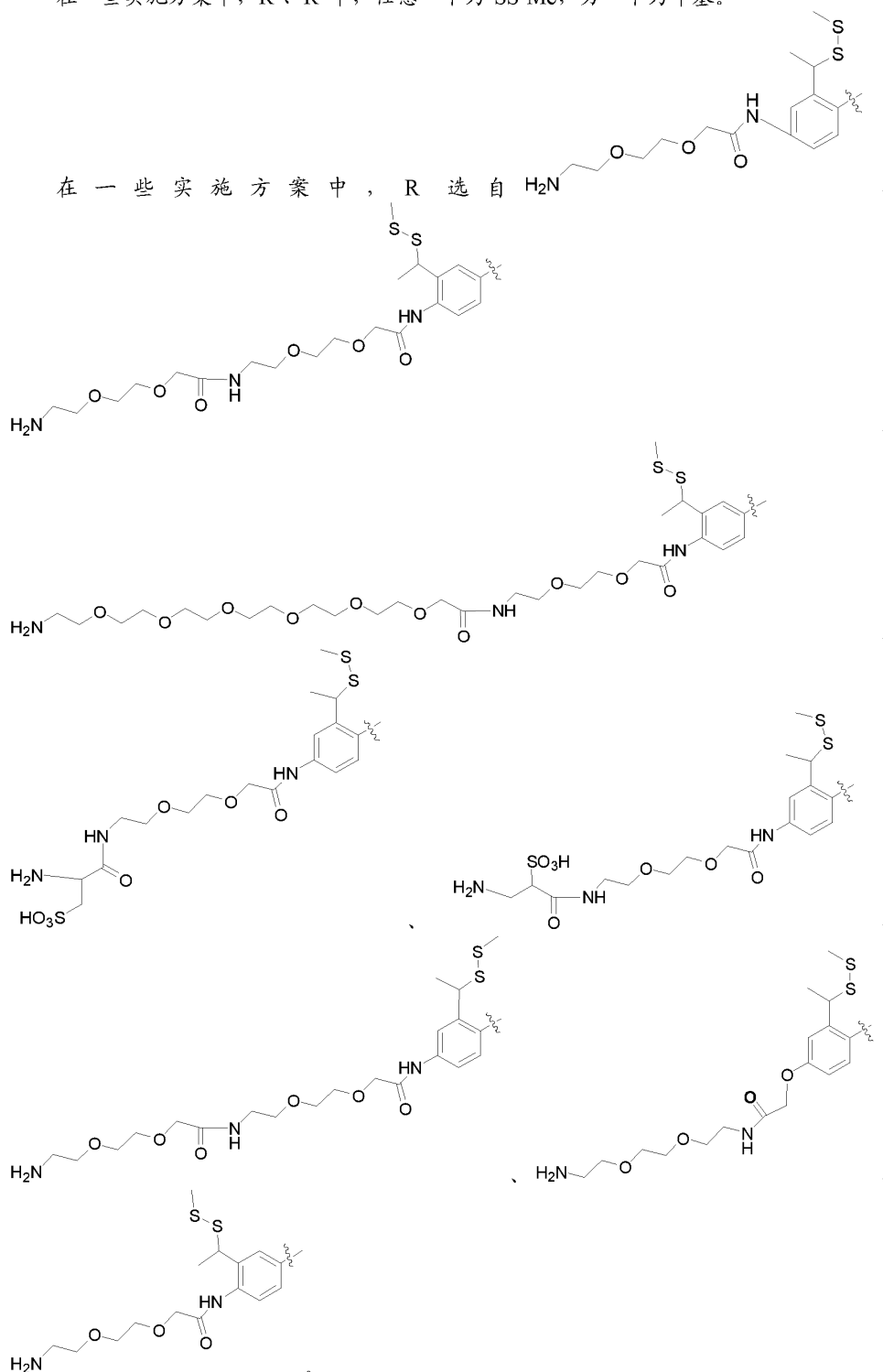
在一些实施方案中, r¹⁰ 选自 2-6 之间的任意整数。

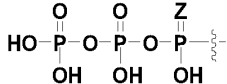
在一些实施方案中, r¹⁰ 为 2 或 6。

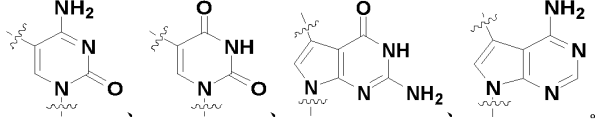
在一些实施方案中， r^{11} 选自 1、2、3。

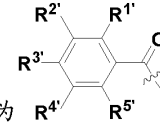
在一些实施方案中， r^{11} 为 1。

在一些实施方案中， R^b 、 R^c 中，任意一个为 -SS-Me，另一个为甲基。



在一些实施方案中，R⁰ 为三磷酸基团 ()。
 在一些实施方案中，Z 为 O。

在一些实施方案中，base¹ 选自 。

在一些实施方案中，所述可逆阻断基团 R' 为 ，R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}

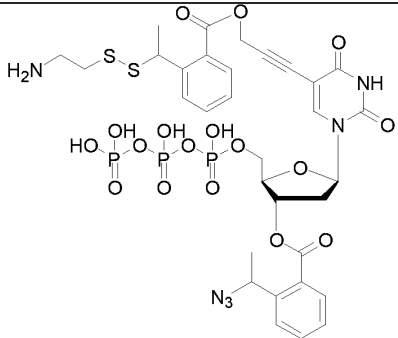
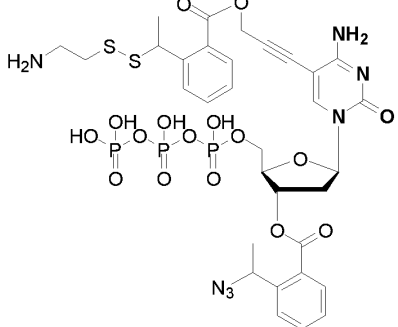
中，任意一个（如 R^{1'} 或 R^{5'}）为 ，另一个（如 R^{3'}）选自 H、C1-C6 烷氧基（如甲氧基），其余为 H；R^{b'}、R^{c'} 中，任意一个选自 -N₃、-SS-C1-C6 烷基（如 -SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu），另一个选自 C1-C6 烷基。

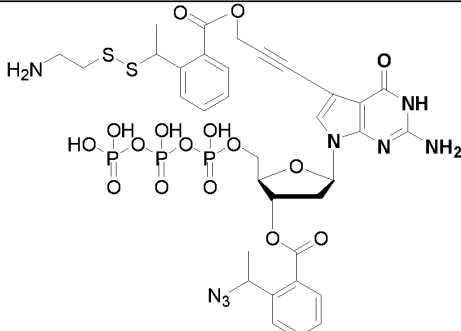
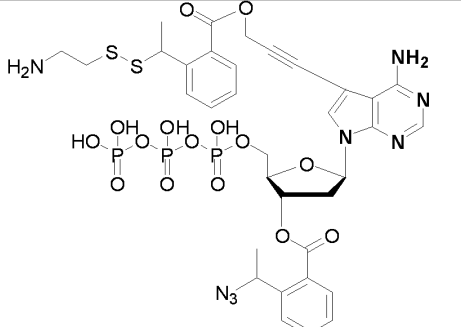
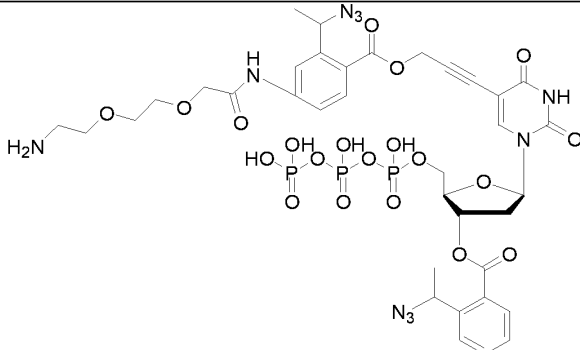
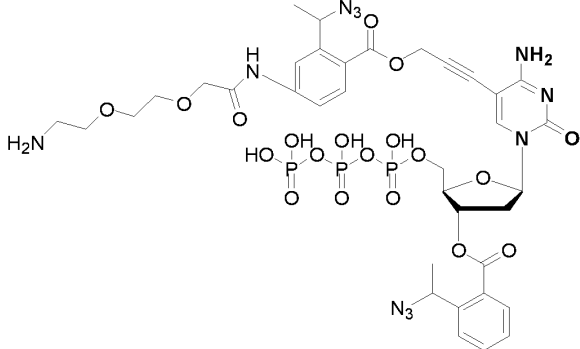
在一些实施方案中，R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'} 中，R^{1'} 为 ，R^{3'} 选自 H、C1-C6 烷氧基（如甲氧基），其余为 H。

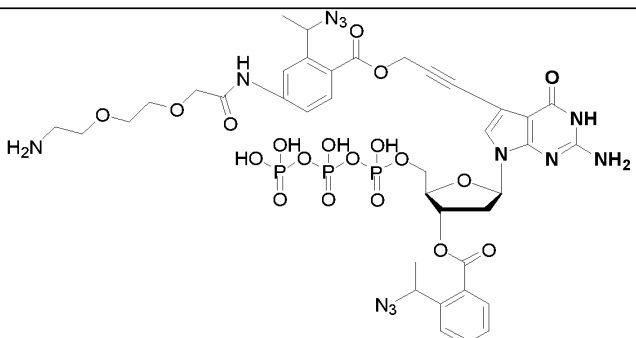
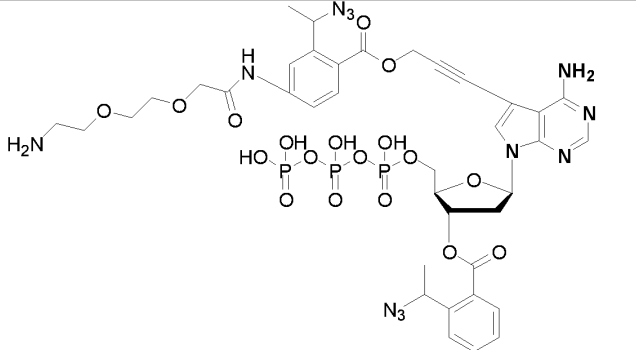
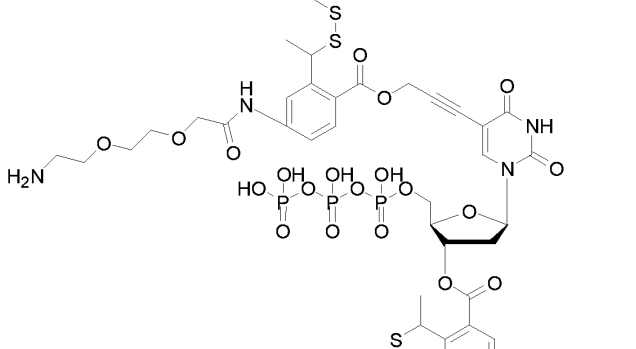
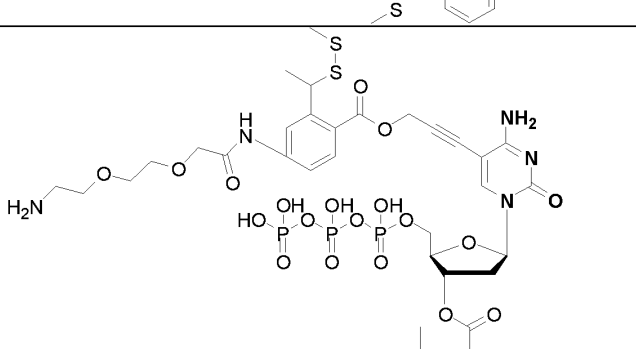
在一些实施方案中，R^{b'}、R^{c'} 中，任意一个为 -N₃、-SS-Me，另一个为甲基。

在一些实施方案中，前述化合物选自以下表 A：

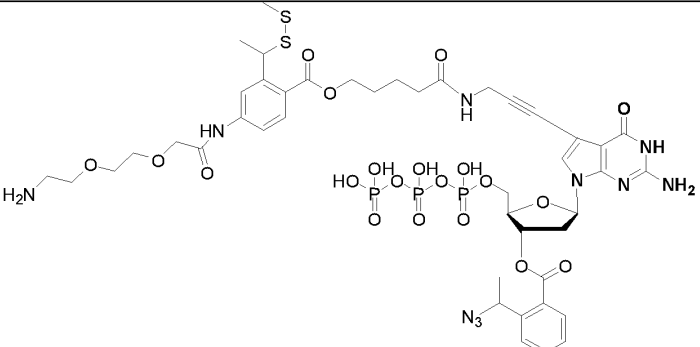
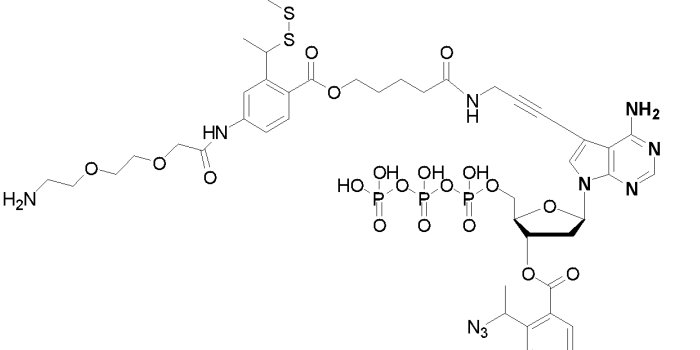
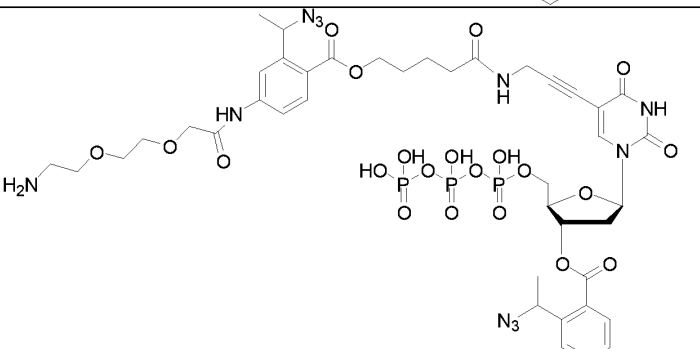
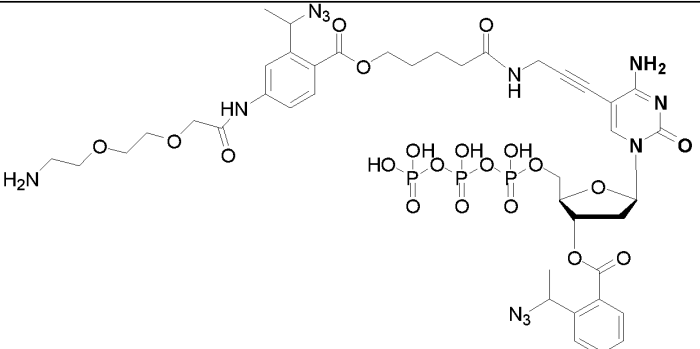
表 A:

编号	结构式
1A	
1B	

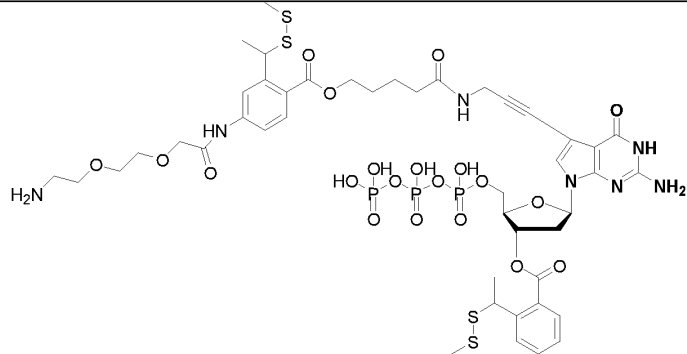
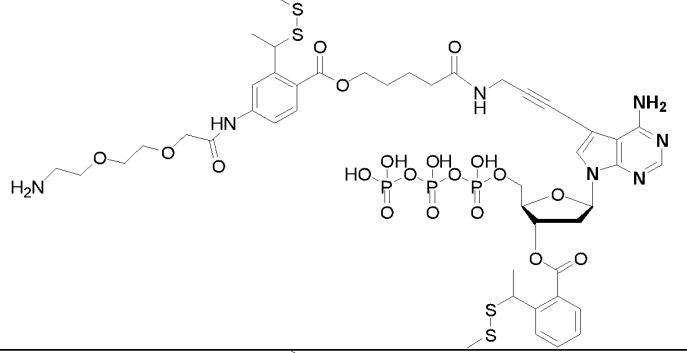
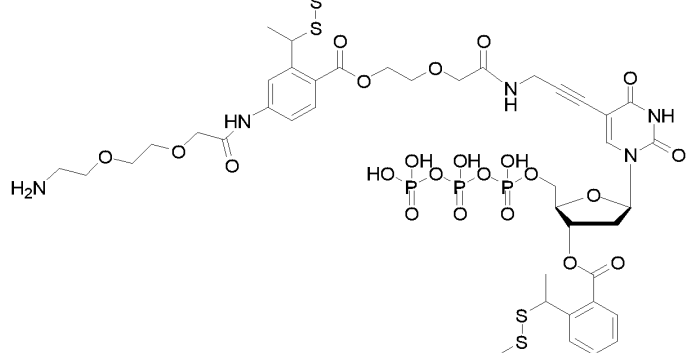
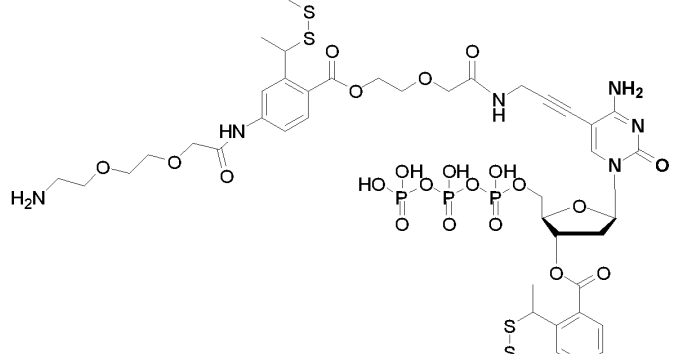
1C	
1D	
2A	
2B	

2C	
2D	
3A	
3B	

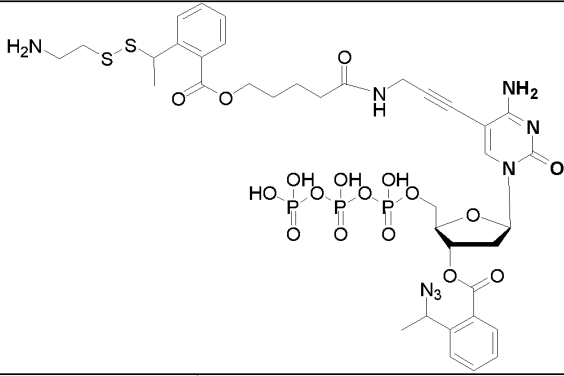
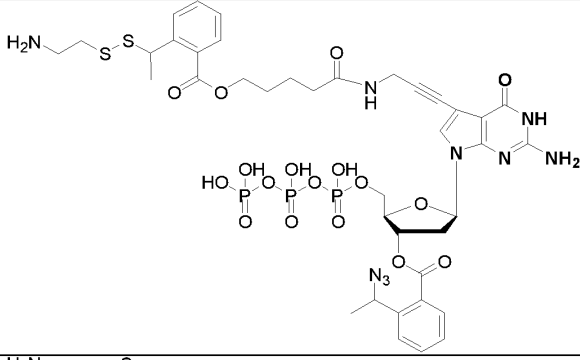
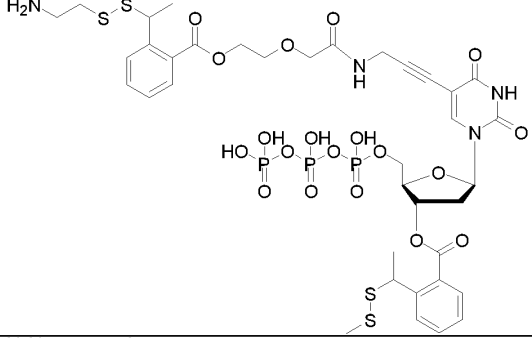
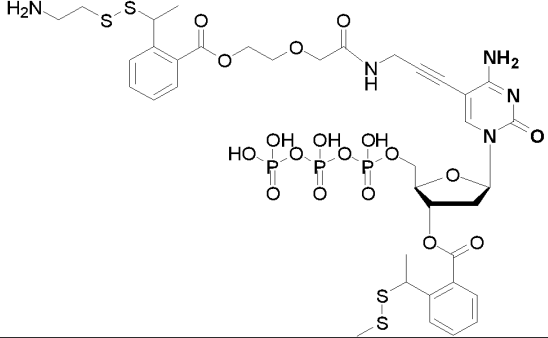
3C	
3D	
4A	
4B	

4C	
4D	
5A	
5B	

5C	
5D	
6A	
6B	

6C	
6D	
7A	
7B	

7C	
7D	
8-1A	
8-1B	

8-1C	
8-1D	
8-2A	
8-2B	

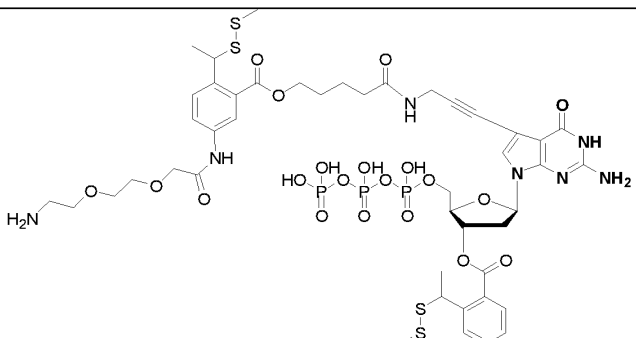
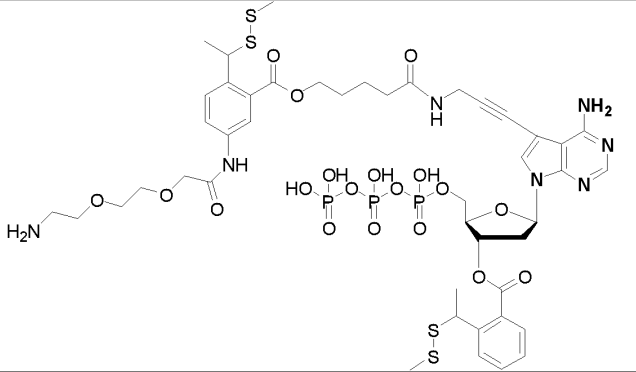
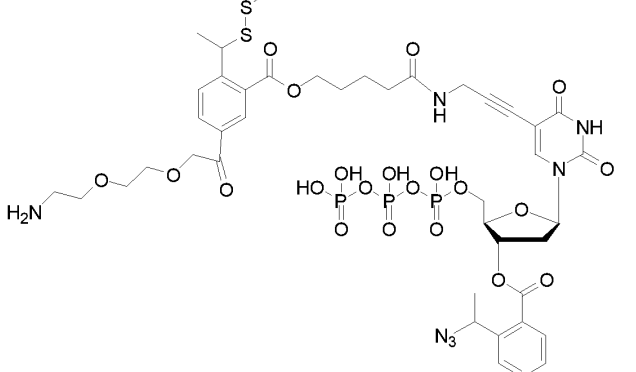
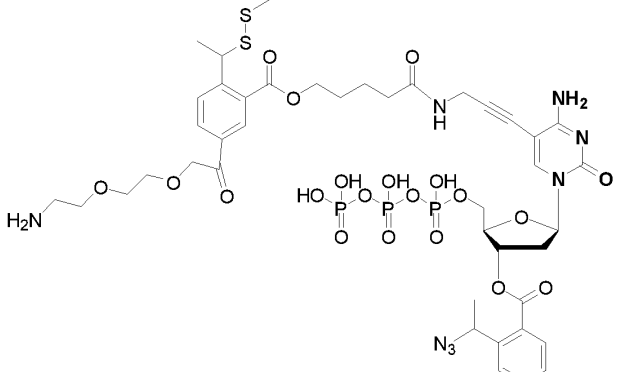
8-2C	 <chem>NCSCCSCC(=O)OCCOC(=O)NCC#CC1=CN2C(=O)NC(=N2)N1COP(=O)(O)OP(=O)(O)OP(=O)(O)OC3OC(COC(=O)C4=CC=C(C(C)S)C4)C3</chem>
8-2D	 <chem>NCSCCSCC(=O)OCCOC(=O)NCC#CC1=CN2C(=O)NC(=N2)N1COP(=O)(O)OP(=O)(O)OP(=O)(O)OC3OC(COC(=O)C4=CC=C(C(C)S)C4)C3</chem>
11A	 <chem>NCSCCSCC(=O)OCCOC(=O)NCC#CC1=CN2C(=O)NC(=N2)N1COP(=O)(O)OP(=O)(O)OP(=O)(O)OC3OC(COC(=O)C4=CC=C(C(C)S)C4)C3</chem>
11B	 <chem>NCSCCSCC(=O)OCCOC(=O)NCC#CC1=CN2C(=O)NC(=N2)N1COP(=O)(O)OP(=O)(O)OP(=O)(O)OC3OC(COC(=O)C4=CC=C(C(C)S)C4)C3</chem>

11C	
11D	
12A	
12B	

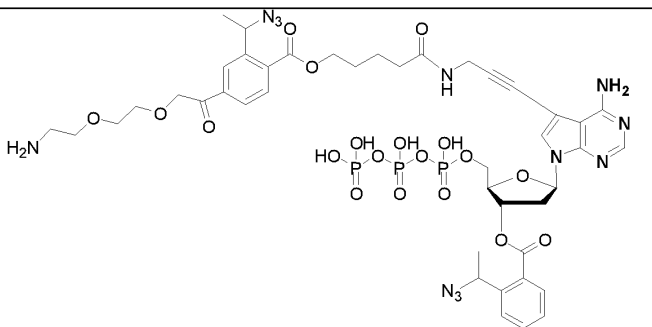
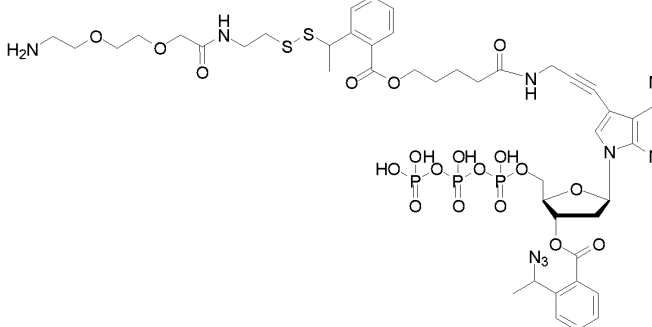
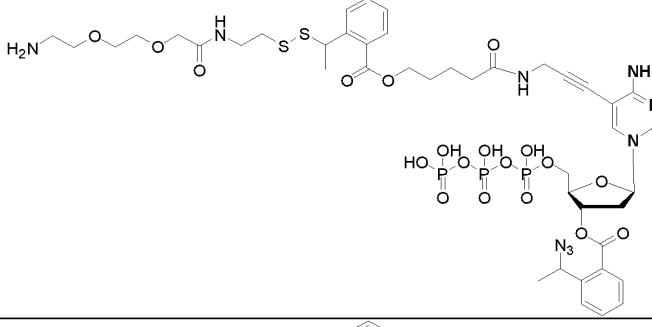
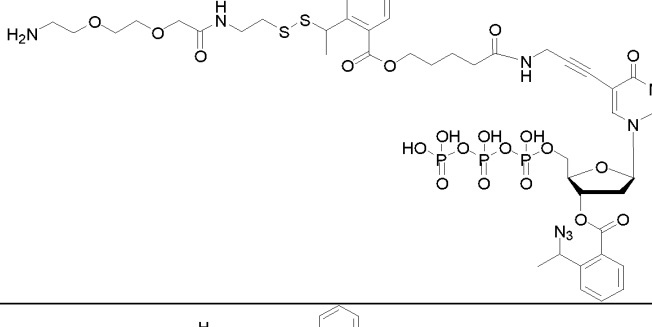
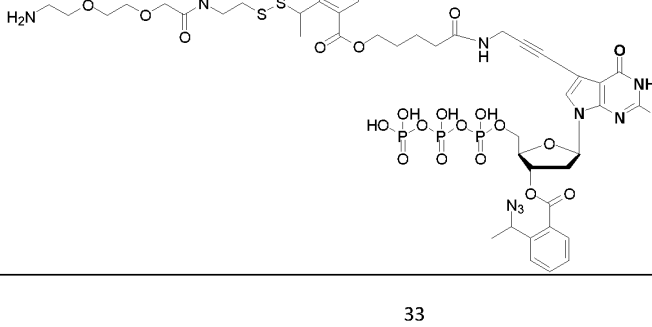
12C	
12D	
13A	
13B	

13C	
13D	
14A	
14B	

14C	
14D	
15A	
15B	

15C	
15D	
16A	
16B	

17C	<chem>Cc1cc(ccc1C#N)C(=O)OCCCCCNC#CC#CN2C(=O)NC(=O)N2[C@H]3O[C@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@@H](OC(=O)c4cc(C)c(cc4)N=[N+]=[N-])[C@H](O)[C@H]3O</chem>
17D	<chem>Cc1cc(ccc1C#N)C(=O)OCCCCCNC#CC#CN2C(=O)NC(=O)N2[C@H]3O[C@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@@H](OC(=O)c4cc(C)c(cc4)N=[N+]=[N-])[C@H](O)[C@H]3O</chem>
18A	<chem>Cc1cc(ccc1C#N)C(=O)OCCCCCNC#CC#CN2C(=O)NC(=O)N2[C@H]3O[C@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@@H](OC(=O)c4cc(C)c(cc4)N=[N+]=[N-])[C@H](O)[C@H]3O</chem>
18B	<chem>Cc1cc(ccc1C#N)C(=O)OCCCCCNC#CC#CN2C(=O)NC(=O)N2[C@H]3O[C@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@@H](OC(=O)c4cc(C)c(cc4)N=[N+]=[N-])[C@H](O)[C@H]3O</chem>
18C	<chem>Cc1cc(ccc1C#N)C(=O)OCCCCCNC#CC#CN2C(=O)NC(=O)N2[C@H]3O[C@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@@H](OC(=O)c4cc(C)c(cc4)N=[N+]=[N-])[C@H](O)[C@H]3O</chem>

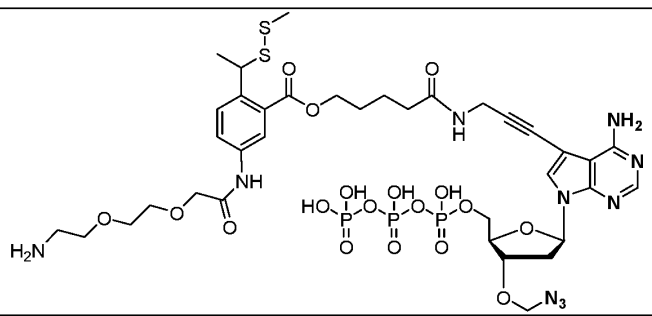
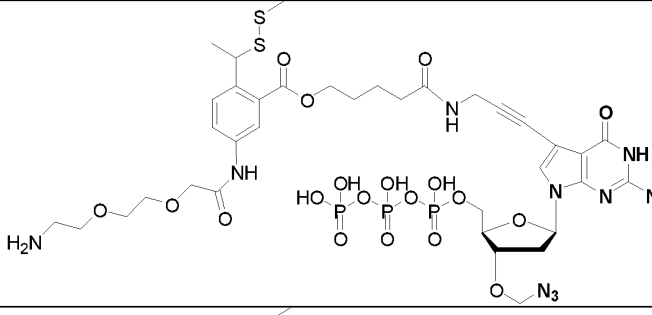
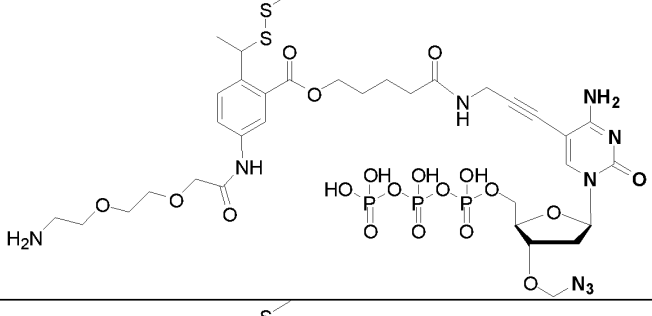
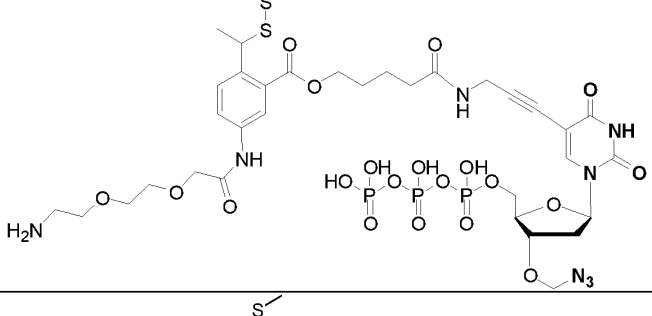
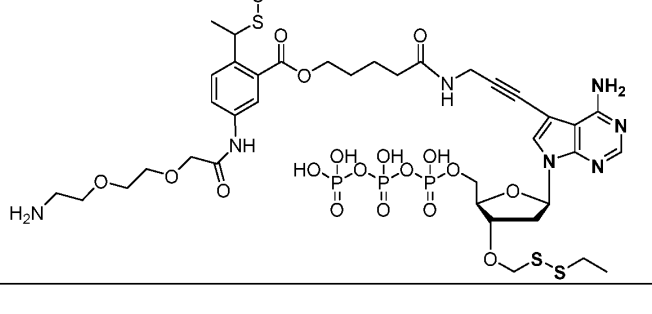
18D	
19A	
19B	
19C	
19D	

20A	
20B	
20C	
20D	

21A	Chemical structure 21A: A complex molecule featuring a disulfide-linked phenyl group, a linker, a nucleoside, and a purine derivative. The structure includes a disulfide bridge (H₂N-CH₂-CH₂-S-S-) attached to a phenyl ring, which is linked via an ester and an amide to a nucleoside. The nucleoside consists of a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), and a purine base with an amino group (NH₂).
21B	Chemical structure 21B: A complex molecule featuring a disulfide-linked phenyl group, a linker, a nucleoside, and a pyrazole derivative. The structure includes a disulfide bridge (H₂N-CH₂-CH₂-S-S-) attached to a phenyl ring, which is linked via an ester and an amide to a nucleoside. The nucleoside consists of a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), and a pyrazole base with an amino group (NH₂).
21C	Chemical structure 21C: A complex molecule featuring a disulfide-linked phenyl group, a linker, a nucleoside, and a pyrimidine derivative. The structure includes a disulfide bridge (H₂N-CH₂-CH₂-S-S-) attached to a phenyl ring, which is linked via an ester and an amide to a nucleoside. The nucleoside consists of a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), and a pyrimidine base with an amino group (NH₂).
21D	Chemical structure 21D: A complex molecule featuring a disulfide-linked phenyl group, a linker, a nucleoside, and a pyrimidine derivative. The structure includes a disulfide bridge (H₂N-CH₂-CH₂-S-S-) attached to a phenyl ring, which is linked via an ester and an amide to a nucleoside. The nucleoside consists of a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), and a pyrimidine base with an amino group (NH₂).

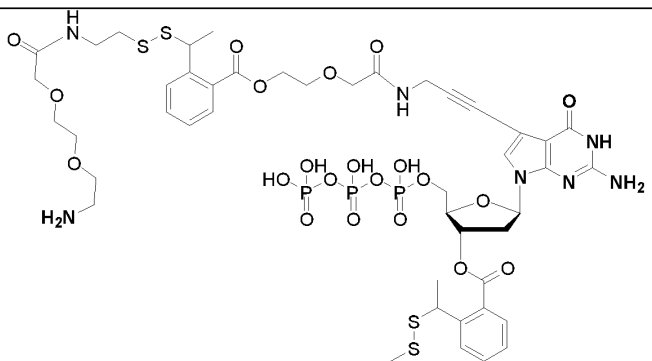
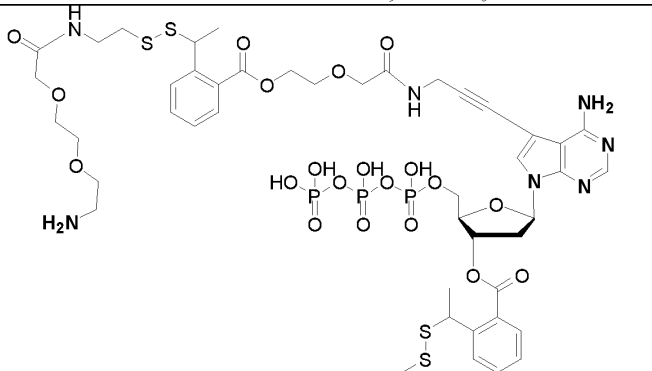
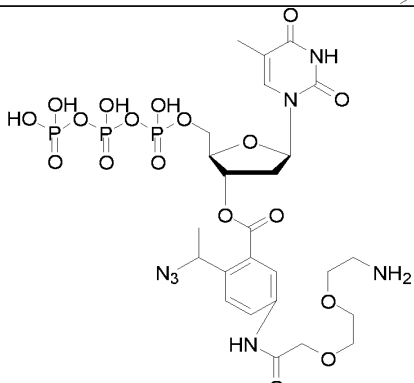
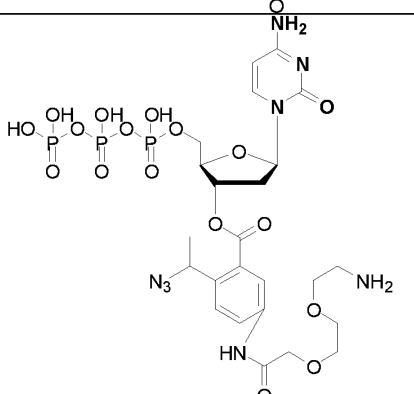
22A	 <chem>NCSCCSCC(=O)OCCCCC(=O)NCC#CC1=NC2=C(N1)NC(=O)N=C2N[C@H]3O[C@@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@H](OC(=O)C(C)Sc4ccccc4)[C@H](O)[C@H]3O</chem>
22B	 <chem>NCSCCSCC(=O)OCCCCC(=O)NCC#CC1=NC2=C(N1)N=CN=C2N[C@H]3O[C@@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@H](OC(=O)C(C)Sc4ccccc4)[C@H](O)[C@H]3O</chem>
22C	 <chem>NCSCCSCC(=O)OCCCCC(=O)NCC#CC1=NC(=O)N=C(N1)N[C@H]2O[C@@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@H](OC(=O)C(C)Sc3ccccc3)[C@H](O)[C@H]2O</chem>
22D	 <chem>NCSCCSCC(=O)OCCCCC(=O)NCC#CC1=NC(=O)NC(=O)N1[C@H]2O[C@@H](COP(=O)(O)OP(=O)(O)OP(=O)(O)O)[C@H](OC(=O)C(C)Sc3ccccc3)[C@H](O)[C@H]12</chem>

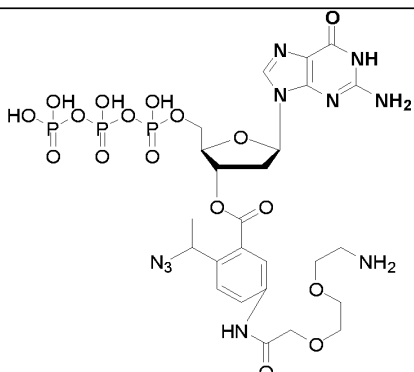
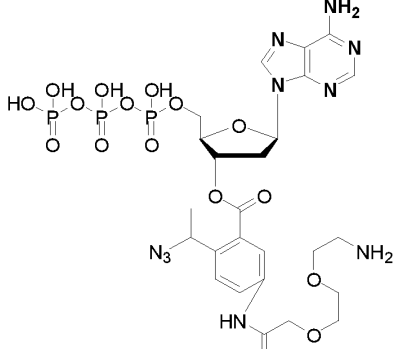
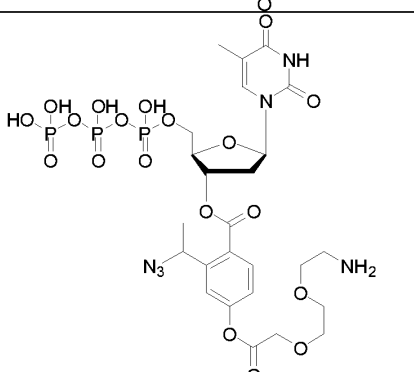
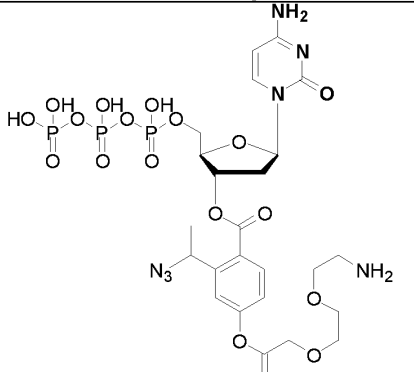
23A	
23B	
23C	
23D	

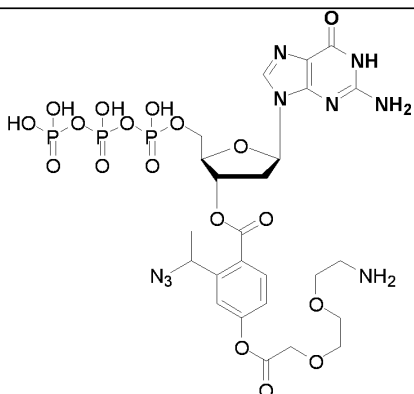
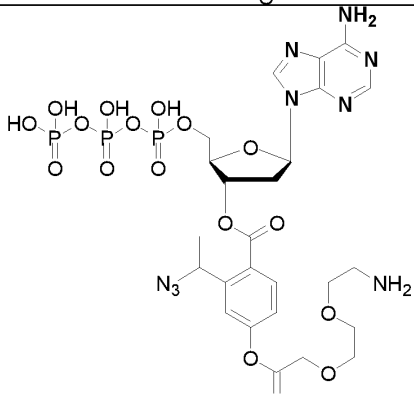
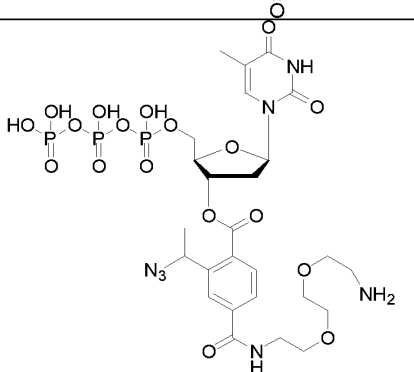
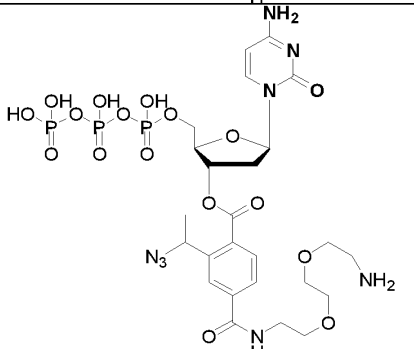
24A	
24B	
24C	
24D	
25A	

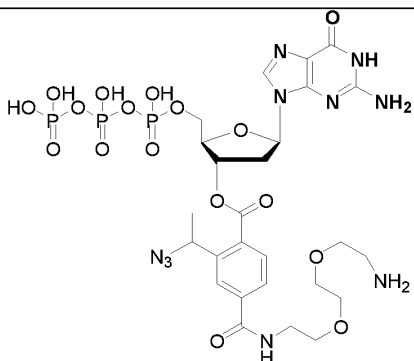
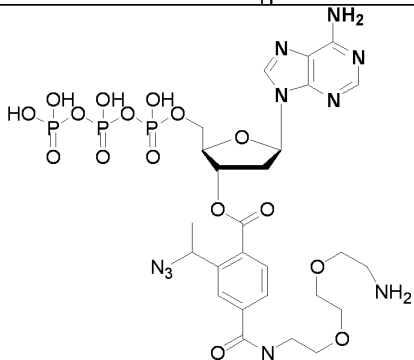
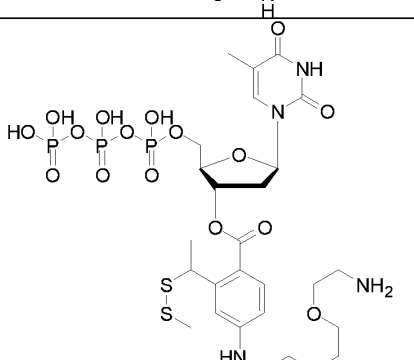
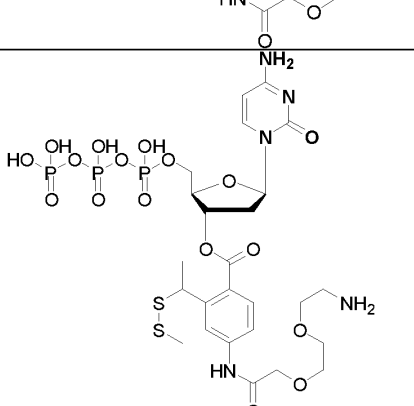
25B	
25C	
25D	
26A	
26B	

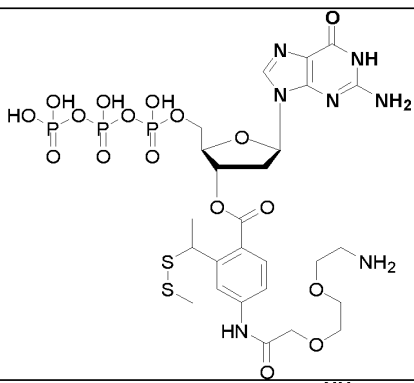
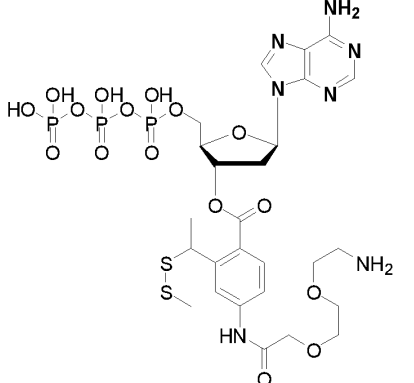
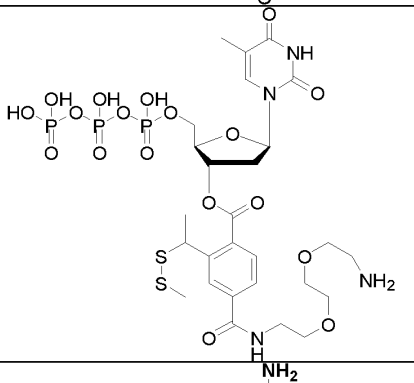
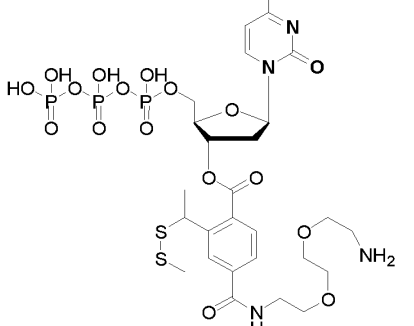
26C	
26D	
27A	
27B	

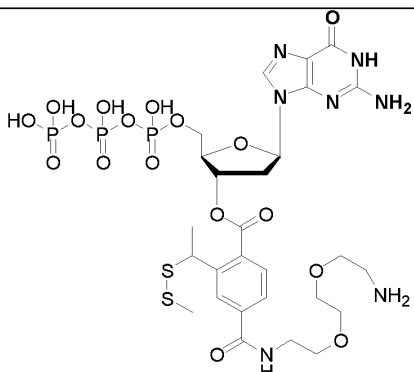
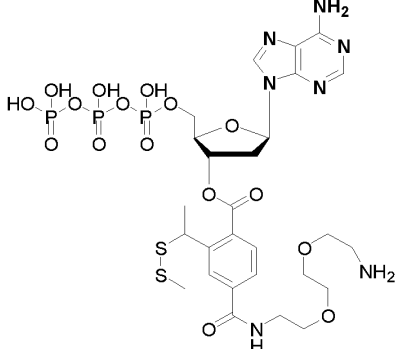
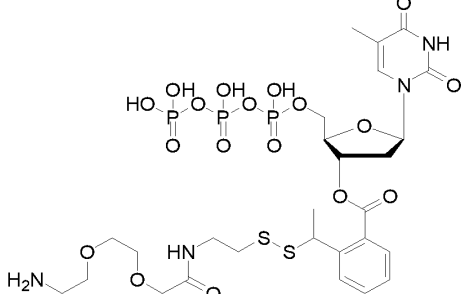
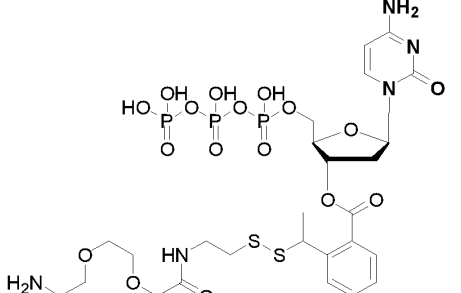
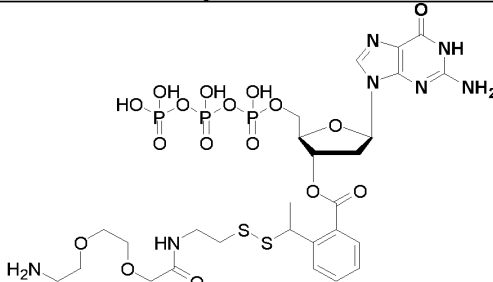
29C	
29D	
30A	
30B	

30C	
30D	
31A	
31B	

31C	
31D	
32A	
32B	

32C	
32D	
33A	
33B	

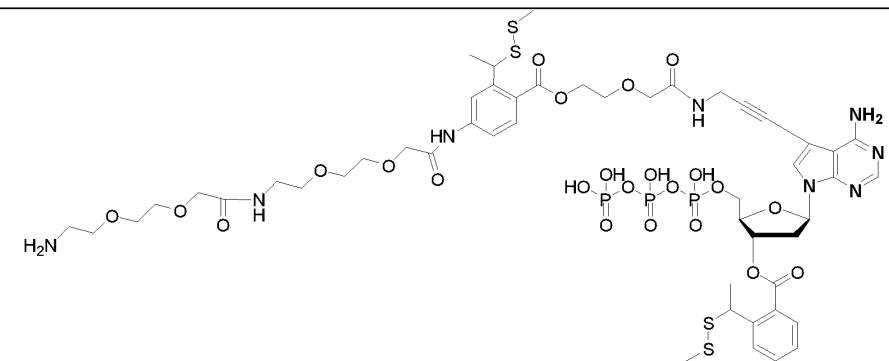
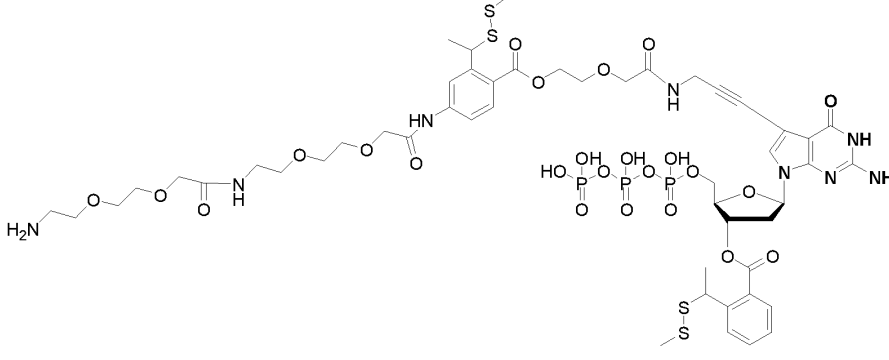
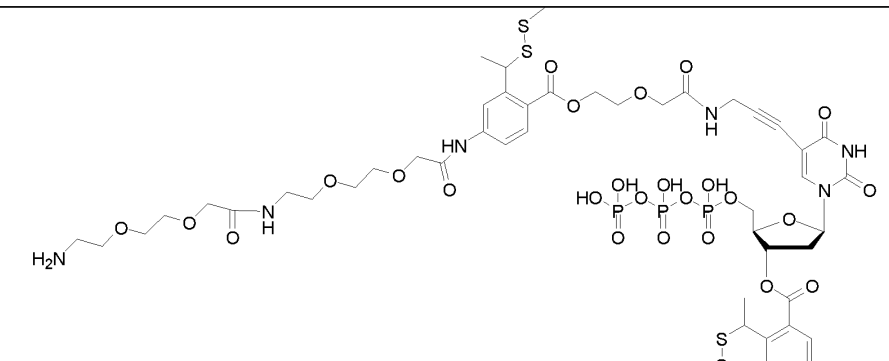
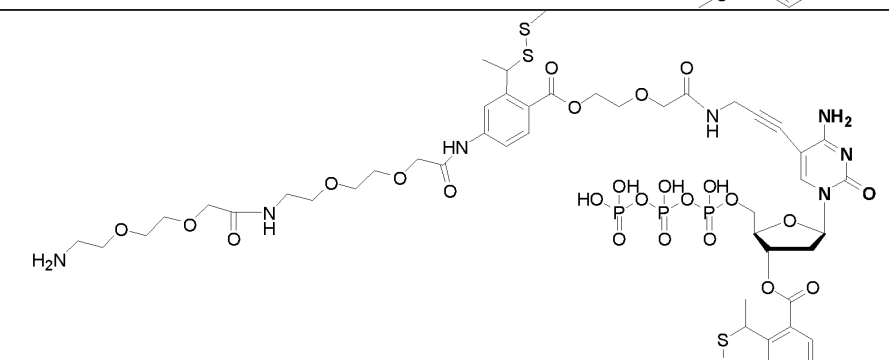
33C	
33D	
34A	
34B	

34C	
34D	
35A	
35B	
35C	

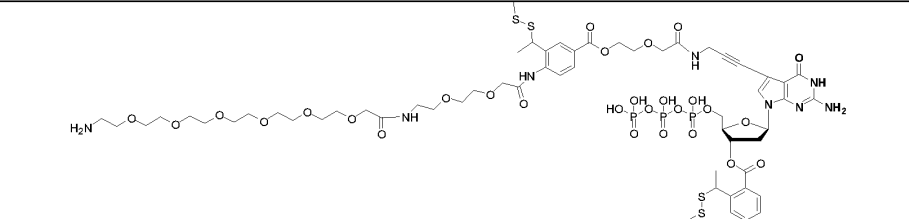
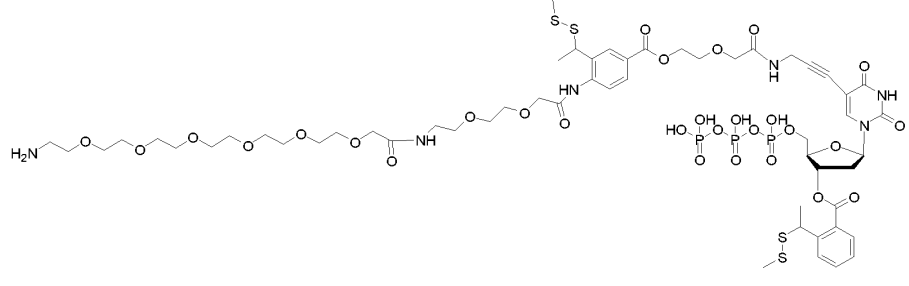
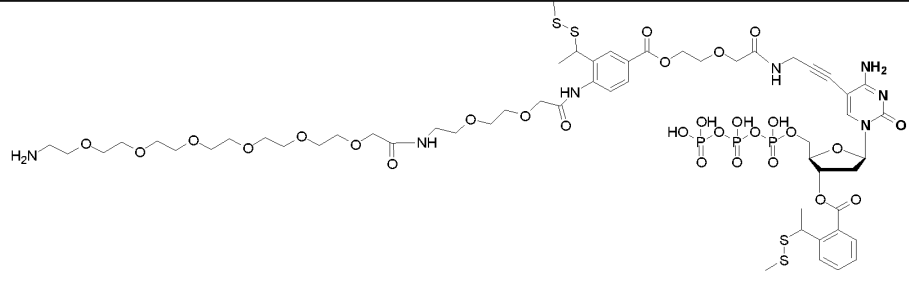
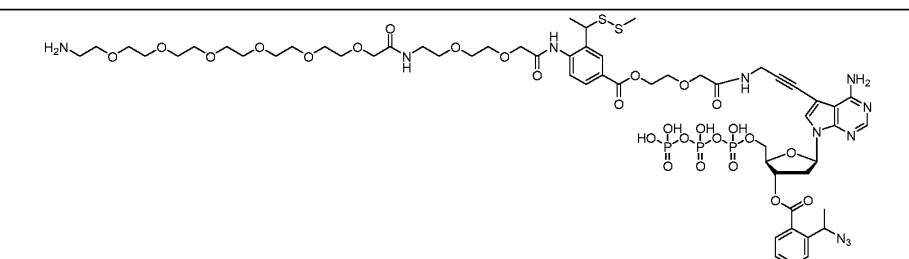
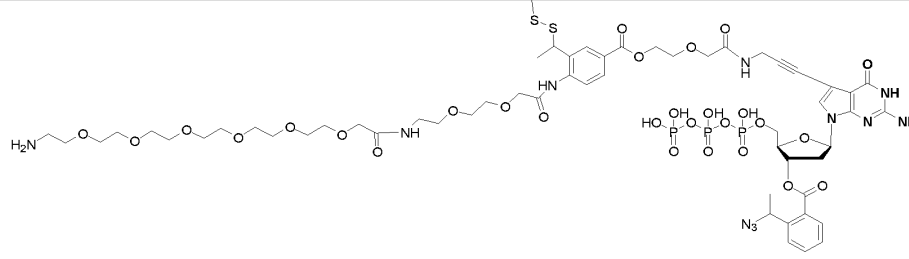
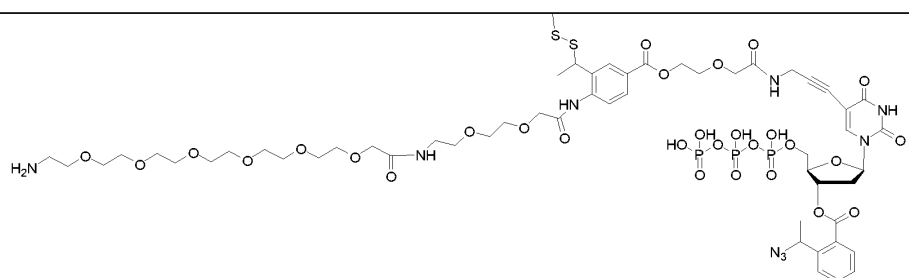
35D	
36A	
36B	
36C	

36D	
37A	
37B	
37C	

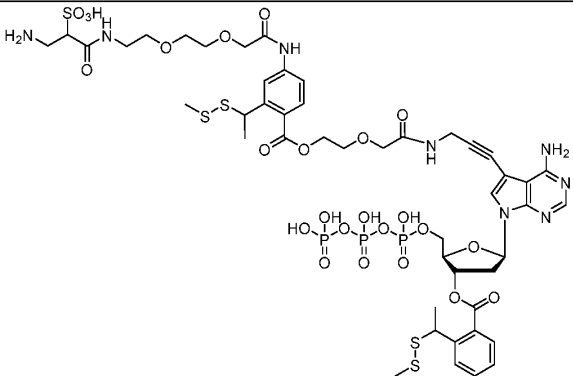
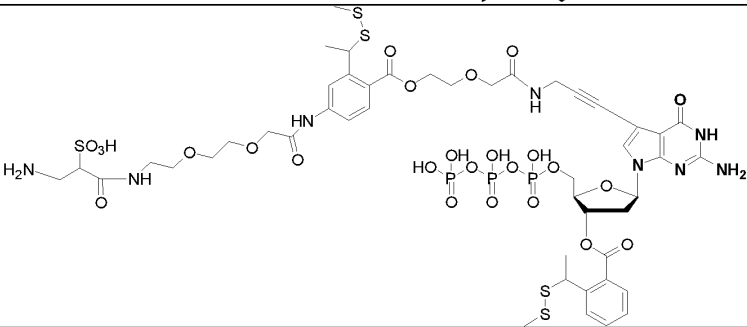
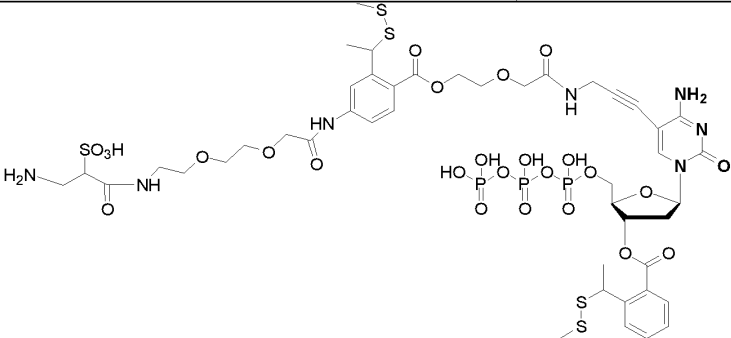
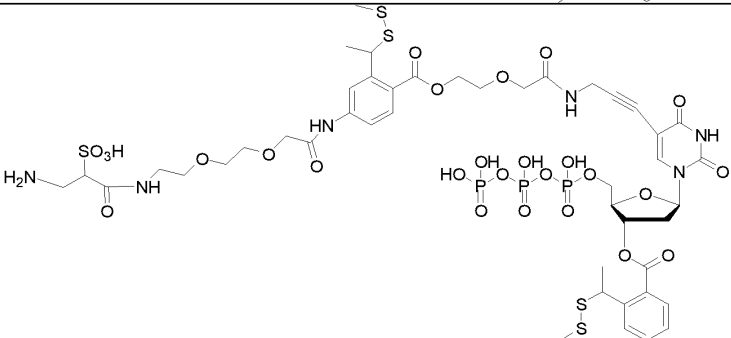
[illegible]

	 Chemical structure of compound 39A, featuring a long poly(ethylene glycol) (PEG) chain with a terminal amine group (H₂N) and a carboxylic acid group (COOH). The PEG chain is linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃). The benzene ring is also linked via an amide bond to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), which is connected to a ribose sugar moiety. The ribose sugar is linked via an amide bond to a pyrimidine ring, which is substituted with an amino group (NH₂) and a methyl group (CH₃). The pyrimidine ring is also linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃).
39B	 Chemical structure of compound 39B, featuring a long poly(ethylene glycol) (PEG) chain with a terminal amine group (H₂N) and a carboxylic acid group (COOH). The PEG chain is linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃). The benzene ring is also linked via an amide bond to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), which is connected to a ribose sugar moiety. The ribose sugar is linked via an amide bond to a pyrimidine ring, which is substituted with an amino group (NH₂) and a methyl group (CH₃). The pyrimidine ring is also linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃).
39C	 Chemical structure of compound 39C, featuring a long poly(ethylene glycol) (PEG) chain with a terminal amine group (H₂N) and a carboxylic acid group (COOH). The PEG chain is linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃). The benzene ring is also linked via an amide bond to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), which is connected to a ribose sugar moiety. The ribose sugar is linked via an amide bond to a pyrimidine ring, which is substituted with an amino group (NH₂) and a methyl group (CH₃). The pyrimidine ring is also linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃).
39D	 Chemical structure of compound 39D, featuring a long poly(ethylene glycol) (PEG) chain with a terminal amine group (H₂N) and a carboxylic acid group (COOH). The PEG chain is linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃). The benzene ring is also linked via an amide bond to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-), which is connected to a ribose sugar moiety. The ribose sugar is linked via an amide bond to a pyrimidine ring, which is substituted with an amino group (NH₂) and a methyl group (CH₃). The pyrimidine ring is also linked via an amide bond to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methyl group (CH₃).

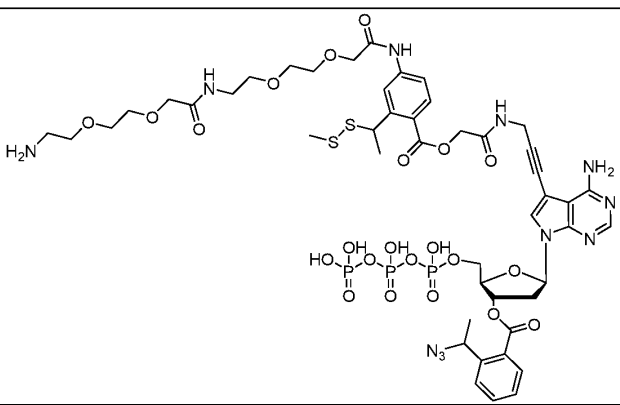
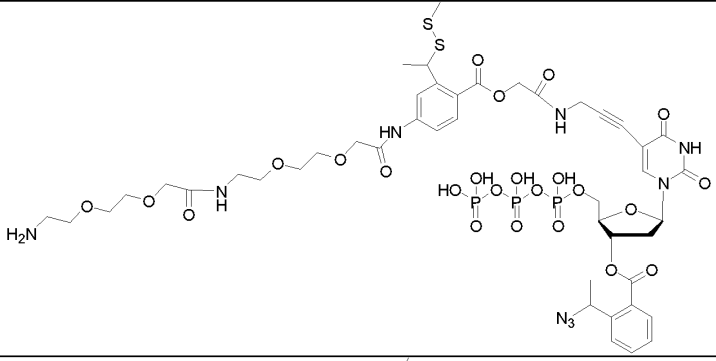
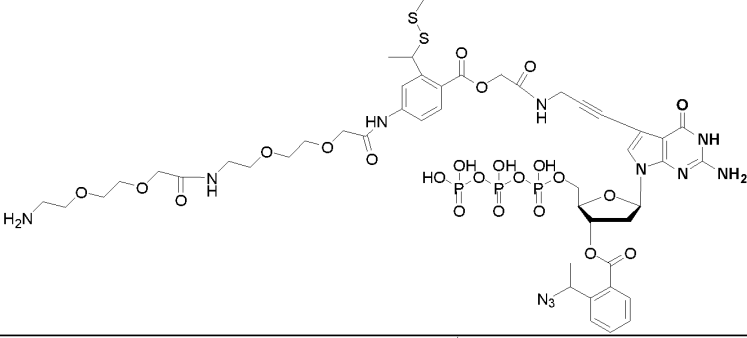
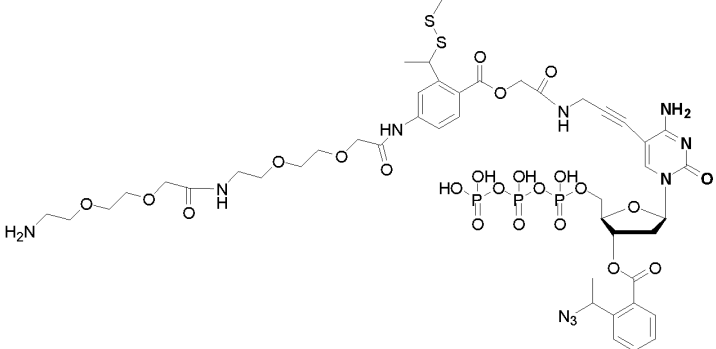
40A	
40B	
40C	
40D	
41A	

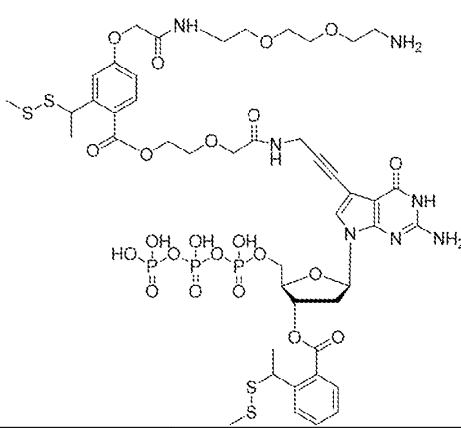
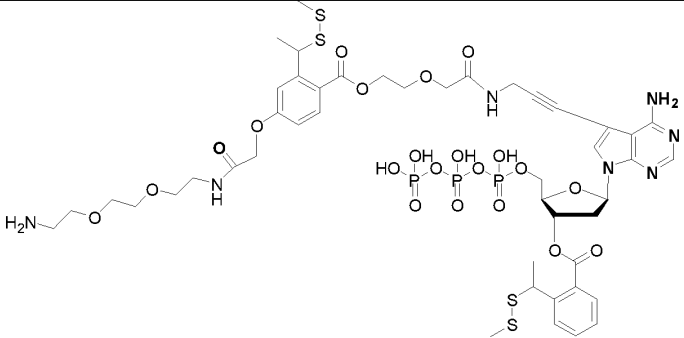
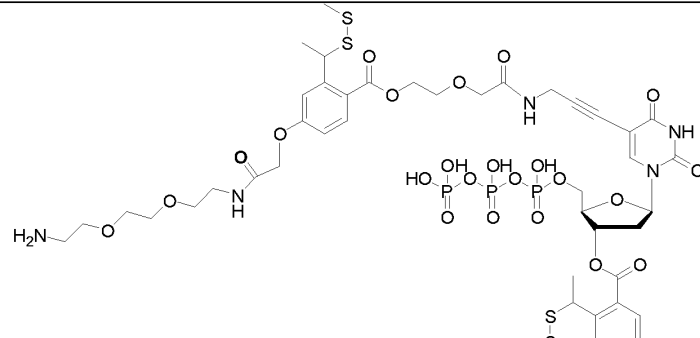
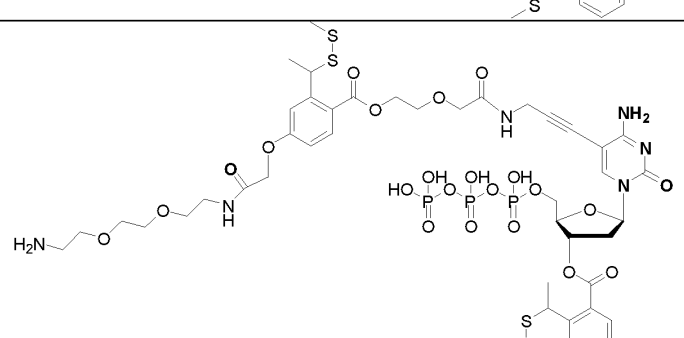
41B	
41C	
41D	
42A	
42B	
42C	

42D	
43A	
43B	
43C	
43D	

44A	
44B	
44C	
44D	

45A	
45B	
45C	
45D	

46A	
46B	
46C	
46D	

47A	
47B	
47C	
47D	

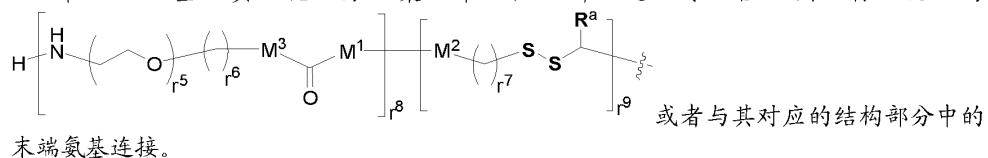
48A	
48B	
48C	
48D	

。在一些实施方案中，前述化合物或其盐携带额外的可检测标记。

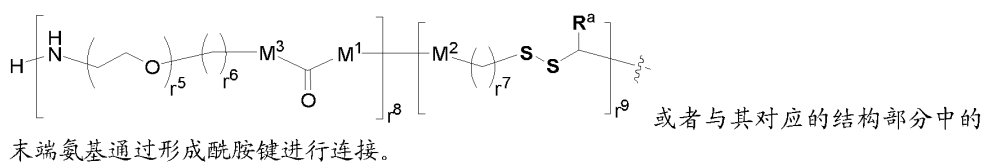
在一些实施方案中，所述化合物或其盐携带的额外的可检测标记是通过亲和试剂（如抗体、适配体、亲和物结蛋白）引入的，所述亲和试剂携带所述可检测标记，且所述亲和试剂可以特异性识别并结合所述化合物或其盐的表位。

在一些实施方案中，所述可检测标记与式 I、式 II、式 I-1、式 I-2、式 II-1 或式 II-2 所示化合物中的 R 或前述具体化合物中对应于 R 的结构部分连接。

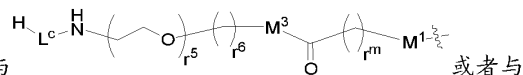
在一些实施方案中，所述可检测标记与



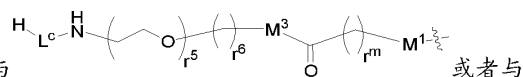
在一些实施方案中，所述可检测标记中的羧基与



在一些实施方案中,所述可检测标记与式 III 所示化合物中的 R 连接或前述具体化合物中对应于 R 的结构部分连接。



在一些实施方案中,所述可检测标记与其对应的结构部分中的末端氨基连接。



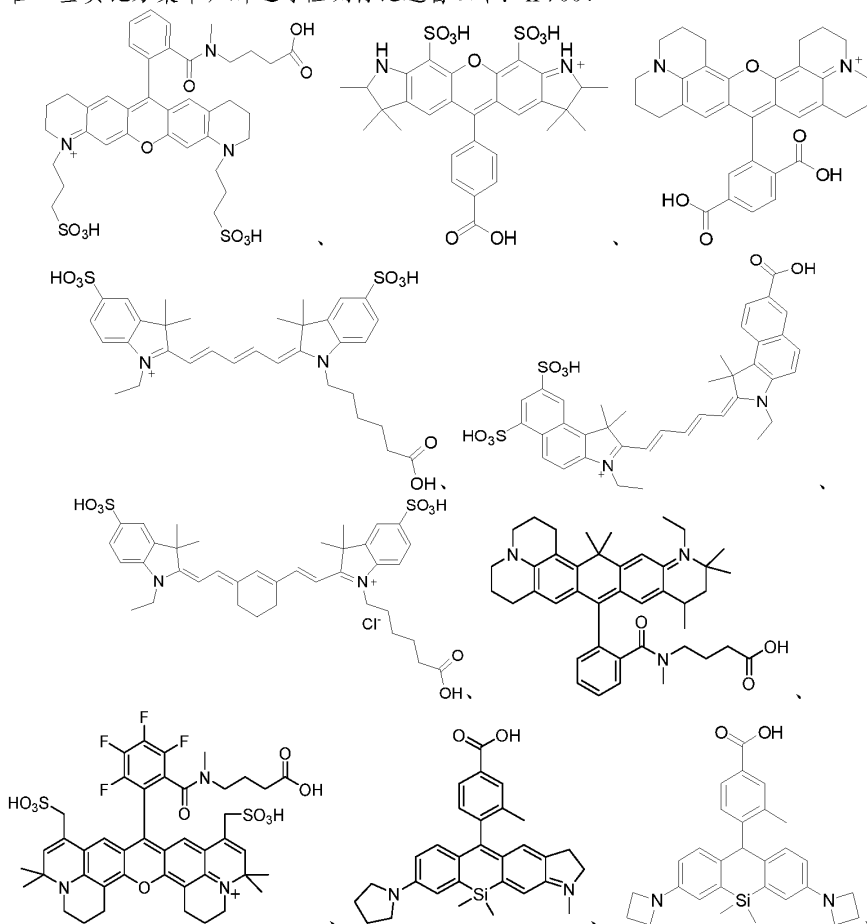
在一些实施方案中,所述可检测标记与对应的结构部分中的末端氨基通过形成酰胺键进行连接。

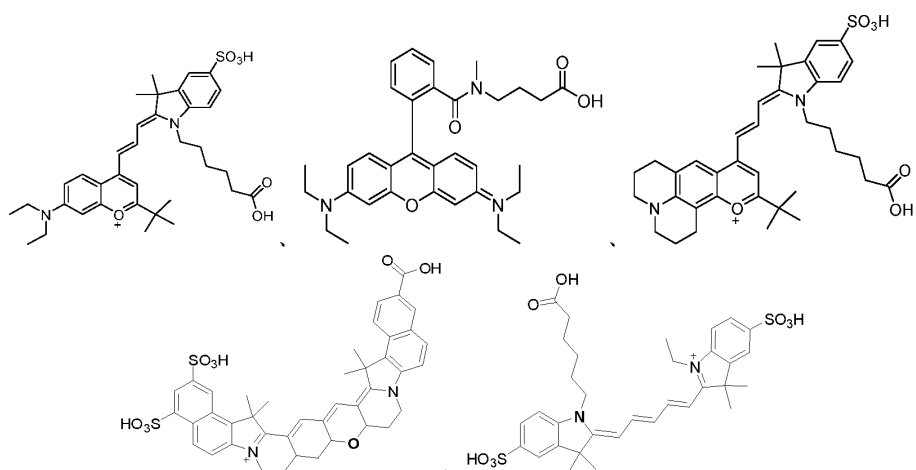
在一些实施方案中, base¹不同, 式 I 所示化合物携带的额外的可检测标记不同。

在一些实施方案中, base^2 不同, 式 II 所示化合物携带的额外的可检测标记不同。

在一些实施方案中，所述可检测标记为荧光标记。

在一些实施方案中,所述可检测标记选自以下: iF700、

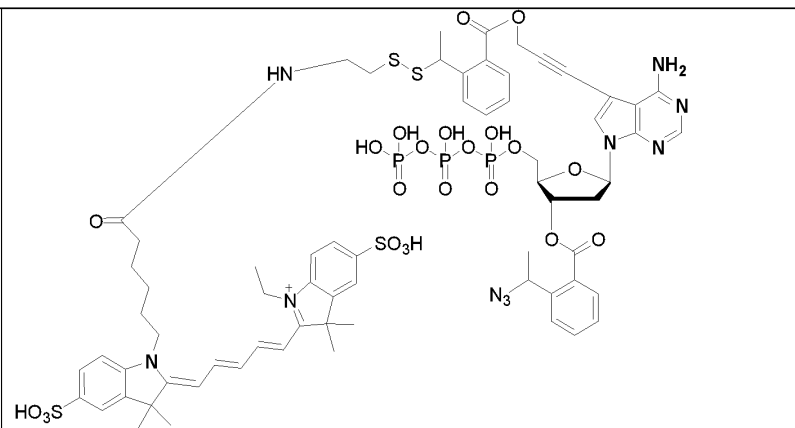
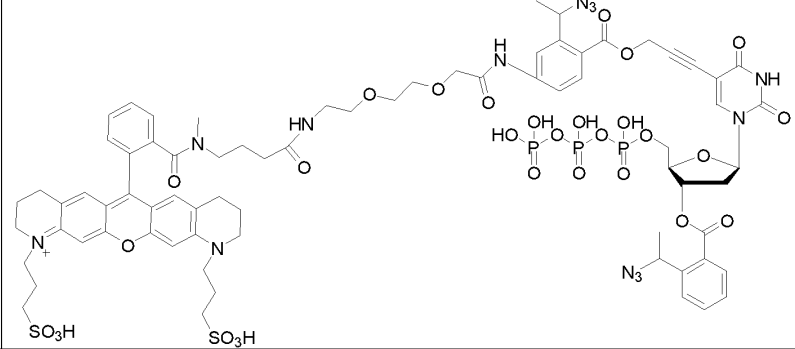
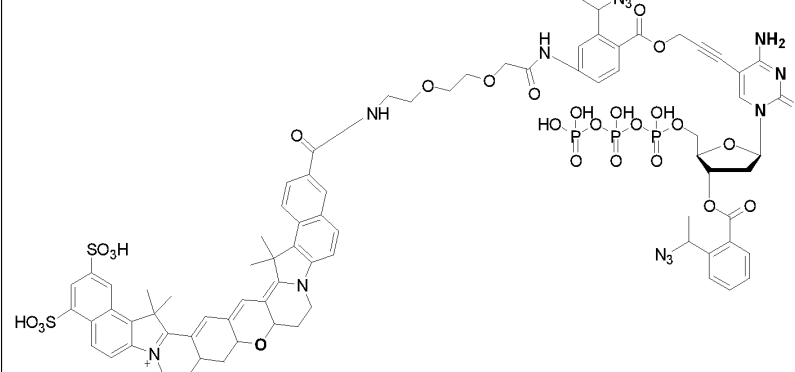
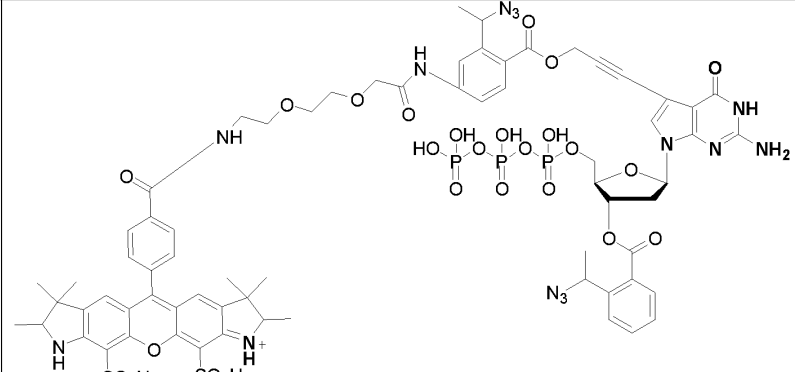




在一些实施方案中，前述化合物选自以下表 B:

表 B:

编号	结构式
1A'	
1B'	
1C'	

1D'	
2A'	
2B'	
2C'	

2D'	Chemical structure 2D' is a complex molecule featuring a central linker with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The linker is connected to a pyridine ring and a pyrazole ring. The pyridine ring is substituted with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The pyrazole ring is substituted with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The molecule also contains a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂).
3A'	Chemical structure 3A' is a complex molecule featuring a central linker with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The linker is connected to a pyridine ring and a pyrazole ring. The pyridine ring is substituted with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The pyrazole ring is substituted with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The molecule also contains a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂).
3B'	Chemical structure 3B' is a complex molecule featuring a central linker with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The linker is connected to a pyridine ring and a pyrazole ring. The pyridine ring is substituted with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The pyrazole ring is substituted with a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂). The molecule also contains a sulfonate group (SO₃H) and a sulfonamide group (SO₂NH₂).

3C'	Chemical structure 3C' is a complex molecule. It features a central pyrimidine ring system with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is further linked to a phosphate chain (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The phosphate chain is connected to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). This benzene ring is also linked to a long aliphatic chain containing an amide group (NH-CO-) and a hydroxyl group (OH). The aliphatic chain is connected to a pyridine ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). The pyridine ring is further linked to a complex polycyclic system, including a pyrazole ring and a sulfonamide group (SO₂NH₂).
3D'	Chemical structure 3D' is a complex molecule. It features a central pyrimidine ring system with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is further linked to a phosphate chain (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The phosphate chain is connected to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). This benzene ring is also linked to a long aliphatic chain containing an amide group (NH-CO-) and a hydroxyl group (OH). The aliphatic chain is connected to a pyridine ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). The pyridine ring is further linked to a complex polycyclic system, including a pyrazole ring and a sulfonamide group (SO₂NH₂).
4A'	Chemical structure 4A' is a complex molecule. It features a central pyrimidine ring system with a carbonyl group (C=O) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is further linked to a phosphate chain (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The phosphate chain is connected to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). This benzene ring is also linked to a long aliphatic chain containing an amide group (NH-CO-) and a hydroxyl group (OH). The aliphatic chain is connected to a pyridine ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). The pyridine ring is further linked to a complex polycyclic system, including a pyrazole ring and a sulfonamide group (SO₂NH₂).
4B'	Chemical structure 4B' is a complex molecule. It features a central pyrimidine ring system with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is further linked to a phosphate chain (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The phosphate chain is connected to a benzene ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). This benzene ring is also linked to a long aliphatic chain containing an amide group (NH-CO-) and a hydroxyl group (OH). The aliphatic chain is connected to a pyridine ring, which is substituted with a methylthio group (SCH₃) and a methylsulfonyl group (SO₂CH₃). The pyridine ring is further linked to a complex polycyclic system, including a pyrazole ring and a sulfonamide group (SO₂NH₂).

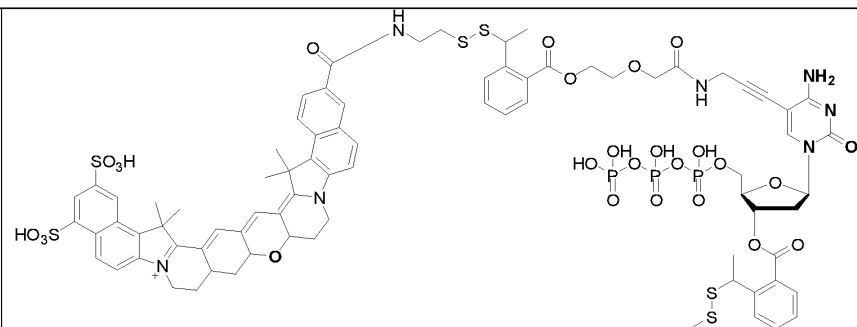
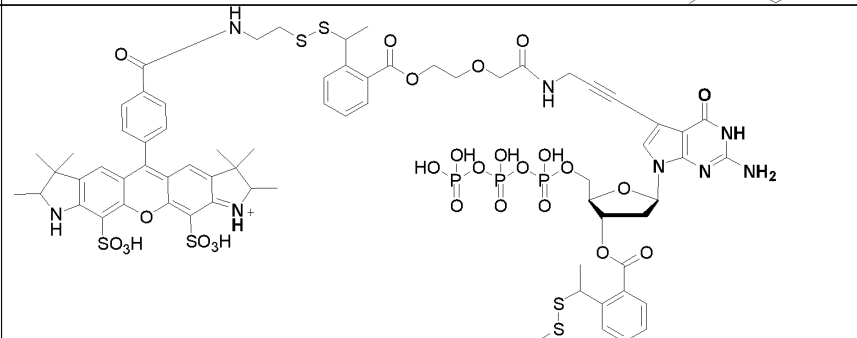
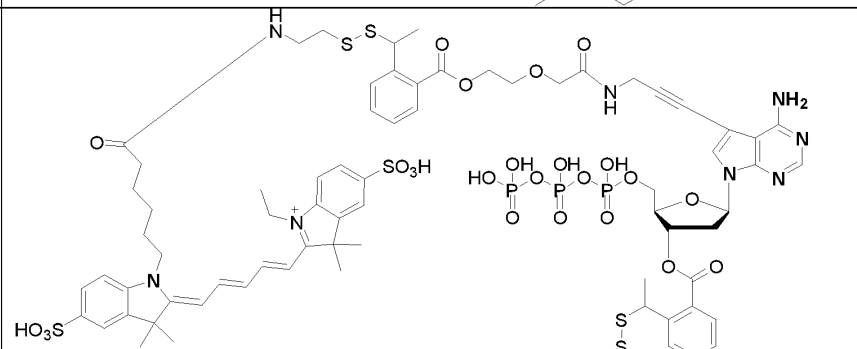
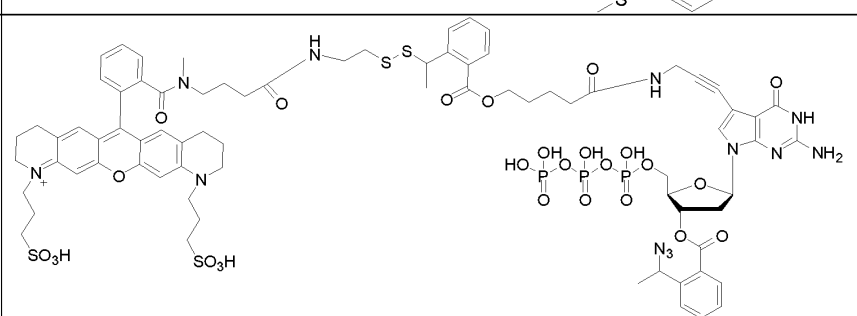
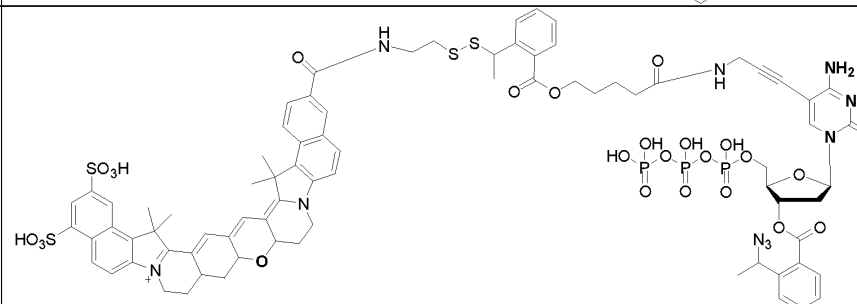
4C'	Chemical structure 4C' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃) and a long chain. The chain includes a carboxylate group (-COO-) linked to a poly(ethylene glycol) (PEG) chain, which is further connected to a carboxylate group (-COO-) and a long alkyl chain. The alkyl chain is terminated by a triazole ring system, which is linked to a sugar moiety (ribose) via an ester bond. The sugar moiety is further substituted with a phosphate group (-P(=O)(OH)-O-) and a 4-azido-2-methylphenyl group (-C(=O)-C₆H₄-N₃).
4D'	Chemical structure 4D' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃) and a long chain. The chain includes a carboxylate group (-COO-) linked to a poly(ethylene glycol) (PEG) chain, which is further connected to a carboxylate group (-COO-) and a long alkyl chain. The alkyl chain is terminated by a triazole ring system, which is linked to a sugar moiety (ribose) via an ester bond. The sugar moiety is further substituted with a phosphate group (-P(=O)(OH)-O-) and a 4-azido-2-methylphenyl group (-C(=O)-C₆H₄-N₃).
5A'	Chemical structure 5A' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃) and a long chain. The chain includes a carboxylate group (-COO-) linked to a poly(ethylene glycol) (PEG) chain, which is further connected to a carboxylate group (-COO-) and a long alkyl chain. The alkyl chain is terminated by a triazole ring system, which is linked to a sugar moiety (ribose) via an ester bond. The sugar moiety is further substituted with a phosphate group (-P(=O)(OH)-O-) and a 4-azido-2-methylphenyl group (-C(=O)-C₆H₄-N₃).
5B'	Chemical structure 5B' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃) and a long chain. The chain includes a carboxylate group (-COO-) linked to a poly(ethylene glycol) (PEG) chain, which is further connected to a carboxylate group (-COO-) and a long alkyl chain. The alkyl chain is terminated by a triazole ring system, which is linked to a sugar moiety (ribose) via an ester bond. The sugar moiety is further substituted with a phosphate group (-P(=O)(OH)-O-) and a 4-azido-2-methylphenyl group (-C(=O)-C₆H₄-N₃).

5C'	Chemical structure of compound 5C'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and a central nitrogen atom. The macrocycle is linked via a long, flexible chain containing an amide, ether, and azide group to a 2-azido-5-((2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl)methoxy)benzoate derivative. The benzotriazine moiety is further substituted with a 2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl group.
5D'	Chemical structure of compound 5D'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and a central nitrogen atom. The macrocycle is linked via a long, flexible chain containing an amide, ether, and azide group to a 2-azido-5-((2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl)methoxy)benzoate derivative. The benzotriazine moiety is further substituted with a 2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl group.
6A'	Chemical structure of compound 6A'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and a central nitrogen atom. The macrocycle is linked via a long, flexible chain containing an amide, ether, and azide group to a 2-azido-5-((2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl)methoxy)benzoate derivative. The benzotriazine moiety is further substituted with a 2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl group.
6B'	Chemical structure of compound 6B'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and a central nitrogen atom. The macrocycle is linked via a long, flexible chain containing an amide, ether, and azide group to a 2-azido-5-((2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl)methoxy)benzoate derivative. The benzotriazine moiety is further substituted with a 2-azido-1,2,3,4-tetrahydro-2H-benzotriazin-4-yl group.

6C'	Chemical structure 6C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). This core is linked via a sulfonate group (SO₃H) to a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The chain is further substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule also includes a sulfonate group (SO₃H) and a sulfonate group (SO₃H).
6D'	Chemical structure 6D' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). This core is linked via a sulfonate group (SO₃H) to a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The chain is further substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule also includes a sulfonate group (SO₃H) and a sulfonate group (SO₃H).
7A'	Chemical structure 7A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). This core is linked via a sulfonate group (SO₃H) to a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The chain is further substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule also includes a sulfonate group (SO₃H) and a sulfonate group (SO₃H).
7B'	Chemical structure 7B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). This core is linked via a sulfonate group (SO₃H) to a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The chain is further substituted with a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule also includes a sulfonate group (SO₃H) and a sulfonate group (SO₃H).

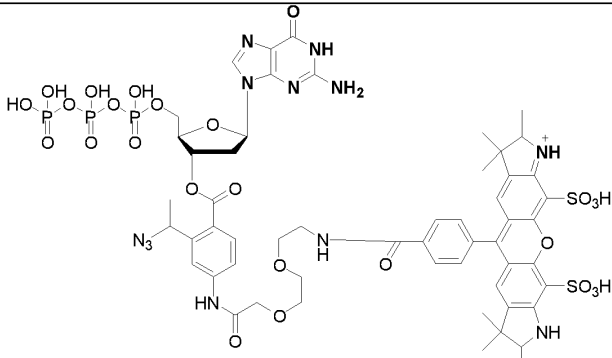
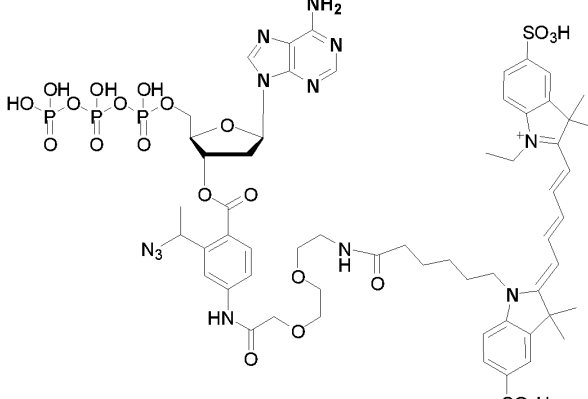
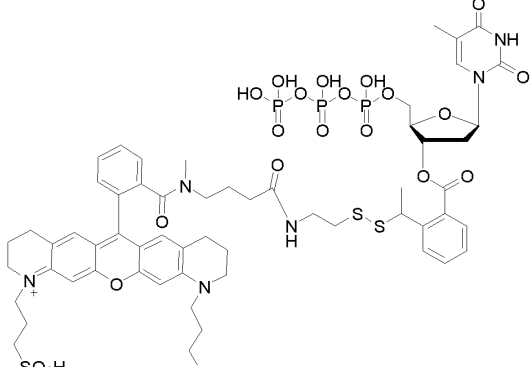
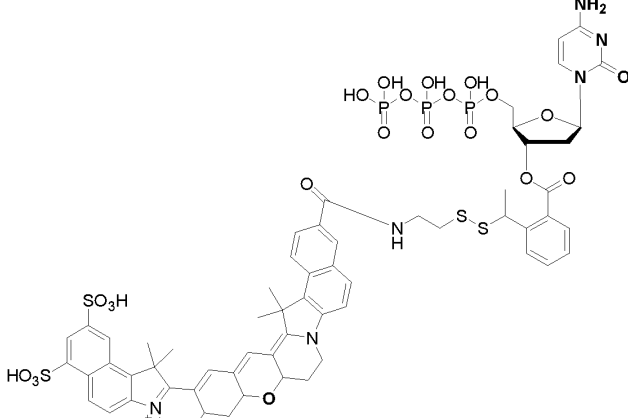
7B'	Chemical structure 7B' is a complex molecule. It features a central pyrimidine ring with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is substituted with a 4-methyl-5-(methylsulfenyl)phenyl group at the 2' position and a 4-methyl-5-(methylsulfenyl)phenyl group at the 3' position. The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 4' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 3' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).
7C'	Chemical structure 7C' is a complex molecule. It features a central pyrimidine ring with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is substituted with a 4-methyl-5-(methylsulfenyl)phenyl group at the 2' position and a 4-methyl-5-(methylsulfenyl)phenyl group at the 3' position. The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 4' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 3' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).
7D'	Chemical structure 7D' is a complex molecule. It features a central pyrimidine ring with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is substituted with a 4-methyl-5-(methylsulfenyl)phenyl group at the 2' position and a 4-methyl-5-(methylsulfenyl)phenyl group at the 3' position. The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 4' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 3' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).
8-1A'	Chemical structure 8-1A' is a complex molecule. It features a central pyrimidine ring with an amino group (NH₂) at position 2 and a carbonyl group (C=O) at position 4. This pyrimidine is linked via a triazole ring to a ribose sugar. The ribose is substituted with a 4-methyl-5-(methylsulfenyl)phenyl group at the 2' position and a 4-methyl-5-(methylsulfenyl)phenyl group at the 3' position. The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 4' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 3' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). The 5' position of the ribose is linked to a chain of four phosphate groups (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).

8-1B'	Chemical structure 8-1B' is a complex molecule. It features a large polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H) and a fused heterocyclic system with a nitrogen atom. This system is connected via a long chain containing a disulfide bridge (S-S) and a thioether linkage (-S-) to a central linker. The linker consists of a hexamethylene chain, an amide group (-NH-), and a terminal group that is a 2,4-diaminopyrimidin-5-yl group. The pyrimidine ring is further substituted with a 2,3,4,5-tetrahydroxybutyl group and a 2-azido-1-phenylethyl group.
8-1C'	Chemical structure 8-1C' is a complex molecule. It features a large polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H) and a fused heterocyclic system with a nitrogen atom. This system is connected via a long chain containing a disulfide bridge (S-S) and a thioether linkage (-S-) to a central linker. The linker consists of a hexamethylene chain, an amide group (-NH-), and a terminal group that is a 2,4-diaminopyrimidin-5-yl group. The pyrimidine ring is further substituted with a 2,3,4,5-tetrahydroxybutyl group and a 2-azido-1-phenylethyl group.
8-1D'	Chemical structure 8-1D' is a complex molecule. It features a large polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H) and a fused heterocyclic system with a nitrogen atom. This system is connected via a long chain containing a disulfide bridge (S-S) and a thioether linkage (-S-) to a central linker. The linker consists of a hexamethylene chain, an amide group (-NH-), and a terminal group that is a 2,4-diaminopyrimidin-5-yl group. The pyrimidine ring is further substituted with a 2,3,4,5-tetrahydroxybutyl group and a 2-azido-1-phenylethyl group.
8-2A'	Chemical structure 8-2A' is a complex molecule. It features a large polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H) and a fused heterocyclic system with a nitrogen atom. This system is connected via a long chain containing a disulfide bridge (S-S) and a thioether linkage (-S-) to a central linker. The linker consists of a hexamethylene chain, an amide group (-NH-), and a terminal group that is a 2,4-diaminopyrimidin-5-yl group. The pyrimidine ring is further substituted with a 2,3,4,5-tetrahydroxybutyl group and a 2-azido-1-phenylethyl group.

8-2B'	
8-2C'	
8-2D'	
9A'	
9B'	

9C'	Chemical structure 9C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO_3H) and a side chain containing a disulfide bridge (S-S) and a thioether linkage ($\text{S-CH}_2\text{CH}_2\text{NH-CO-}$). This side chain is connected to a long alkyl chain, which is further linked to a triphosphate group ($\text{HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-}$). The triphosphate group is attached to a ribose sugar, which is in turn linked to a pyridine ring substituted with a sulfonic acid group (SO_3H).
9D'	Chemical structure 9D' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO_3H) and a side chain containing a disulfide bridge (S-S) and a thioether linkage ($\text{S-CH}_2\text{CH}_2\text{NH-CO-}$). This side chain is connected to a long alkyl chain, which is further linked to a triphosphate group ($\text{HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-}$). The triphosphate group is attached to a ribose sugar, which is in turn linked to a pyridine ring substituted with a sulfonic acid group (SO_3H).
10A'	Chemical structure 10A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO_3H) and a side chain containing a disulfide bridge (S-S) and a thioether linkage ($\text{S-CH}_2\text{CH}_2\text{NH-CO-}$). This side chain is connected to a long alkyl chain, which is further linked to a triphosphate group ($\text{HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-}$). The triphosphate group is attached to a ribose sugar, which is in turn linked to a pyridine ring substituted with a sulfonic acid group (SO_3H).
10B'	Chemical structure 10B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO_3H) and a side chain containing a disulfide bridge (S-S) and a thioether linkage ($\text{S-CH}_2\text{CH}_2\text{NH-CO-}$). This side chain is connected to a long alkyl chain, which is further linked to a triphosphate group ($\text{HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-}$). The triphosphate group is attached to a ribose sugar, which is in turn linked to a pyridine ring substituted with a sulfonic acid group (SO_3H).

10C'	Chemical structure 10C' is a complex molecule. It features a central linker consisting of a hexamethylene chain with a disulfide bridge (-S-S-) and a triphosphate group (-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). This linker is connected to a 2-aminobenzimidazole moiety on the right and a 4-azido-2-methylphenyl ester group on the left. The ester group is further linked to a 4-sulfamoylphenyl group, which is connected to a 4-sulfamoyl-2,6-dimethyl-5,7-dihydroxy-1,4-dihydropyridine moiety.
10D'	Chemical structure 10D' is a complex molecule. It features a central linker consisting of a hexamethylene chain with a disulfide bridge (-S-S-) and a triphosphate group (-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). This linker is connected to a 2-aminobenzimidazole moiety on the right and a 4-azido-2-methylphenyl ester group on the left. The ester group is further linked to a 4-sulfamoylphenyl group, which is connected to a 4-sulfamoyl-2,6-dimethyl-5,7-dihydroxy-1,4-dihydropyridine moiety.
11A'	Chemical structure 11A' is a complex molecule. It features a central linker consisting of a hexamethylene chain with a disulfide bridge (-S-S-) and a triphosphate group (-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). This linker is connected to a 2-aminobenzimidazole moiety on the right and a 4-azido-2-methylphenyl ester group on the left. The ester group is further linked to a 4-sulfamoylphenyl group, which is connected to a 4-sulfamoyl-2,6-dimethyl-5,7-dihydroxy-1,4-dihydropyridine moiety.
11B'	Chemical structure 11B' is a complex molecule. It features a central linker consisting of a hexamethylene chain with a disulfide bridge (-S-S-) and a triphosphate group (-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-). This linker is connected to a 2-aminobenzimidazole moiety on the right and a 4-azido-2-methylphenyl ester group on the left. The ester group is further linked to a 4-sulfamoylphenyl group, which is connected to a 4-sulfamoyl-2,6-dimethyl-5,7-dihydroxy-1,4-dihydropyridine moiety.

11C'	
11D'	
12A'	
12B'	

12C'	Chemical structure 12C' is a complex molecule. It features a central porphyrin-like macrocycle with two sulfonic acid groups (SO₃H) and two methyl groups. Attached to the macrocycle is a side chain containing a benzyl group, a disulfide bridge, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further connected to an adenine base.
12D'	Chemical structure 12D' is a complex molecule. It features a central porphyrin-like macrocycle with two sulfonic acid groups (SO₃H) and two methyl groups. Attached to the macrocycle is a side chain containing a benzyl group, a disulfide bridge, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further connected to an adenine base.
13A'	Chemical structure 13A' is a complex molecule. It features a central porphyrin-like macrocycle with two sulfonic acid groups (SO₃H) and two methyl groups. Attached to the macrocycle is a side chain containing a benzyl group, a disulfide bridge, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further connected to an adenine base.
13B'	Chemical structure 13B' is a complex molecule. It features a central porphyrin-like macrocycle with two sulfonic acid groups (SO₃H) and two methyl groups. Attached to the macrocycle is a side chain containing a benzyl group, a disulfide bridge, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further connected to an adenine base.

13C'	Chemical structure of compound 13C'. It features a phthalocyanine-like macrocycle with two sulfonate groups (SO₃H) and a central nitrogen atom with a positive charge. The macrocycle is linked via a long chain containing a thioether, an amide, and a triphosphate group to a nucleoside derivative. The nucleoside consists of a ribose sugar, a phosphate group, and a pyrimidine base with an amino group (NH₂).
13D'	Chemical structure of compound 13D'. It features a phthalocyanine-like macrocycle with two sulfonate groups (SO₃H) and a central nitrogen atom with a positive charge. The macrocycle is linked via a long chain containing a thioether, an amide, and a triphosphate group to a nucleoside derivative. The nucleoside consists of a ribose sugar, a phosphate group, and a pyrimidine base with an amino group (NH₂).
14A'	Chemical structure of compound 14A'. It features a phthalocyanine-like macrocycle with two sulfonate groups (SO₃H) and a central nitrogen atom with a positive charge. The macrocycle is linked via a long chain containing a thioether, an amide, and a triphosphate group to a nucleoside derivative. The nucleoside consists of a ribose sugar, a phosphate group, and a pyrimidine base with an amino group (NH₂).
14B'	Chemical structure of compound 14B'. It features a phthalocyanine-like macrocycle with two sulfonate groups (SO₃H) and a central nitrogen atom with a positive charge. The macrocycle is linked via a long chain containing a thioether, an amide, and a triphosphate group to a nucleoside derivative. The nucleoside consists of a ribose sugar, a phosphate group, and a pyrimidine base with an amino group (NH₂).

14C'	Chemical structure of compound 14C'. It features a central porphyrin-like macrocycle with two sulfonate groups (SO₃H) at the 3 and 5 positions. The macrocycle is substituted with a long, flexible linker containing an amide bond, a polyether chain, and a thioether group. The thioether group is connected to a benzene ring, which is further linked to a long alkyl chain ending in an amide bond. This amide bond is connected to a triazole ring, which is substituted with a 2-azido-1-phenylethyl group and a 2-amino-1-phenylethyl group. The triazole ring is also connected to a sugar moiety (ribose) via a glycosidic bond. The sugar moiety is further linked to a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).
14D'	Chemical structure of compound 14D'. It features a central porphyrin-like macrocycle with two sulfonate groups (SO₃H) at the 3 and 5 positions. The macrocycle is substituted with a long, flexible linker containing an amide bond, a polyether chain, and a thioether group. The thioether group is connected to a benzene ring, which is further linked to a long alkyl chain ending in an amide bond. This amide bond is connected to a triazole ring, which is substituted with a 2-azido-1-phenylethyl group and a 2-amino-1-phenylethyl group. The triazole ring is also connected to a sugar moiety (ribose) via a glycosidic bond. The sugar moiety is further linked to a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).
15A'	Chemical structure of compound 15A'. It features a central porphyrin-like macrocycle with two sulfonate groups (SO₃H) at the 3 and 5 positions. The macrocycle is substituted with a long, flexible linker containing an amide bond, a polyether chain, and a thioether group. The thioether group is connected to a benzene ring, which is further linked to a long alkyl chain ending in an amide bond. This amide bond is connected to a triazole ring, which is substituted with a 2-azido-1-phenylethyl group and a 2-amino-1-phenylethyl group. The triazole ring is also connected to a sugar moiety (ribose) via a glycosidic bond. The sugar moiety is further linked to a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).
15B'	Chemical structure of compound 15B'. It features a central porphyrin-like macrocycle with two sulfonate groups (SO₃H) at the 3 and 5 positions. The macrocycle is substituted with a long, flexible linker containing an amide bond, a polyether chain, and a thioether group. The thioether group is connected to a benzene ring, which is further linked to a long alkyl chain ending in an amide bond. This amide bond is connected to a triazole ring, which is substituted with a 2-azido-1-phenylethyl group and a 2-amino-1-phenylethyl group. The triazole ring is also connected to a sugar moiety (ribose) via a glycosidic bond. The sugar moiety is further linked to a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-).

15C'	Chemical structure 15C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthioacetate group (-COOCH₃). This core is linked via an amide bond to a long chain containing a terminal amine group (-NH₂) and a terminal amide group (-CONH₂). The chain also includes a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is further substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is also substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂).
15D'	Chemical structure 15D' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthioacetate group (-COOCH₃). This core is linked via an amide bond to a long chain containing a terminal amine group (-NH₂) and a terminal amide group (-CONH₂). The chain also includes a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is further substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is also substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂).
16A'	Chemical structure 16A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthioacetate group (-COOCH₃). This core is linked via an amide bond to a long chain containing a terminal amine group (-NH₂) and a terminal amide group (-CONH₂). The chain also includes a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is further substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is also substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂).
16B'	Chemical structure 16B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthioacetate group (-COOCH₃). This core is linked via an amide bond to a long chain containing a terminal amine group (-NH₂) and a terminal amide group (-CONH₂). The chain also includes a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is further substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂). The molecule is also substituted with a terminal amide group (-CONH₂) and a terminal amide group (-CONH₂).

16C'	Chemical structure 16C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). This core is linked via a long, flexible chain containing an amide and an ether linkage to a nucleoside moiety. The nucleoside consists of a ribose sugar attached to a pyrimidine base (uracil) via an N-glycosidic bond. The pyrimidine base is further substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The ribose sugar is also substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The entire molecule is terminated with a sulfonic acid group (-SO₃H).
16D'	Chemical structure 16D' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). This core is linked via a long, flexible chain containing an amide and an ether linkage to a nucleoside moiety. The nucleoside consists of a ribose sugar attached to a pyrimidine base (uracil) via an N-glycosidic bond. The pyrimidine base is further substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The ribose sugar is also substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The entire molecule is terminated with a sulfonic acid group (-SO₃H).
17A'	Chemical structure 17A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). This core is linked via a long, flexible chain containing an amide and an ether linkage to a nucleoside moiety. The nucleoside consists of a ribose sugar attached to a pyrimidine base (uracil) via an N-glycosidic bond. The pyrimidine base is further substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The ribose sugar is also substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The entire molecule is terminated with a sulfonic acid group (-SO₃H).
17B'	Chemical structure 17B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). This core is linked via a long, flexible chain containing an amide and an ether linkage to a nucleoside moiety. The nucleoside consists of a ribose sugar attached to a pyrimidine base (uracil) via an N-glycosidic bond. The pyrimidine base is further substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The ribose sugar is also substituted with a methylthio group (-S-CH₃) and a methylthio carbonyl group (-C(=O)-S-CH₃). The entire molecule is terminated with a sulfonic acid group (-SO₃H).

17C'	Chemical structure 17C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO₃H) and a side chain containing a sulfonamide group (SO₂NH-) and a sulfonate group (SO₃Na). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group is a nucleoside derivative, specifically a 2'-azido-2'-deoxyribose-5-phosphate, which is further substituted with a 4-aminopyrimidin-2-yl group.
17D'	Chemical structure 17D' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO₃H) and a side chain containing a sulfonamide group (SO₂NH-) and a sulfonate group (SO₃Na). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group is a nucleoside derivative, specifically a 2'-azido-2'-deoxyribose-5-phosphate, which is further substituted with a 4-aminopyrimidin-2-yl group.
18A'	Chemical structure 18A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO₃H) and a side chain containing a sulfonamide group (SO₂NH-) and a sulfonate group (SO₃Na). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group is a nucleoside derivative, specifically a 2'-azido-2'-deoxyribose-5-phosphate, which is further substituted with a 4-aminopyrimidin-2-yl group.
18B'	Chemical structure 18B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonic acid group (SO₃H) and a side chain containing a sulfonamide group (SO₂NH-) and a sulfonate group (SO₃Na). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group is a nucleoside derivative, specifically a 2'-azido-2'-deoxyribose-5-phosphate, which is further substituted with a 4-aminopyrimidin-2-yl group.

18C'	Chemical structure of compound 18C'. It features a phthalocyanine core with two sulfonic acid groups (SO₃H) and two imine groups (NH). The phthalocyanine is linked via a long chain containing a terminal azide group (N₃) and a terminal amide group (NH). The amide is connected to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) which is linked to a ribose sugar. The ribose is further linked to a pyrazole ring, which is substituted with a 4-azido-2-methylphenyl group (N₃-C₆H₄-CH₃).
18D'	Chemical structure of compound 18D'. It features a phthalocyanine core with two sulfonic acid groups (SO₃H) and two imine groups (NH). The phthalocyanine is linked via a long chain containing a terminal azide group (N₃) and a terminal amide group (NH). The amide is connected to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) which is linked to a ribose sugar. The ribose is further linked to a pyrazole ring, which is substituted with a 4-azido-2-methylphenyl group (N₃-C₆H₄-CH₃).
19A'	Chemical structure of compound 19A'. It features a phthalocyanine core with two sulfonic acid groups (SO₃H) and two imine groups (NH). The phthalocyanine is linked via a long chain containing a terminal azide group (N₃) and a terminal amide group (NH). The amide is connected to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) which is linked to a ribose sugar. The ribose is further linked to a pyrazole ring, which is substituted with a 4-azido-2-methylphenyl group (N₃-C₆H₄-CH₃).
19B'	Chemical structure of compound 19B'. It features a phthalocyanine core with two sulfonic acid groups (SO₃H) and two imine groups (NH). The phthalocyanine is linked via a long chain containing a terminal azide group (N₃) and a terminal amide group (NH). The amide is connected to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) which is linked to a ribose sugar. The ribose is further linked to a pyrazole ring, which is substituted with a 4-azido-2-methylphenyl group (N₃-C₆H₄-CH₃).
19C'	Chemical structure of compound 19C'. It features a phthalocyanine core with two sulfonic acid groups (SO₃H) and two imine groups (NH). The phthalocyanine is linked via a long chain containing a terminal azide group (N₃) and a terminal amide group (NH). The amide is connected to a triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) which is linked to a ribose sugar. The ribose is further linked to a pyrazole ring, which is substituted with a 4-azido-2-methylphenyl group (N₃-C₆H₄-CH₃).

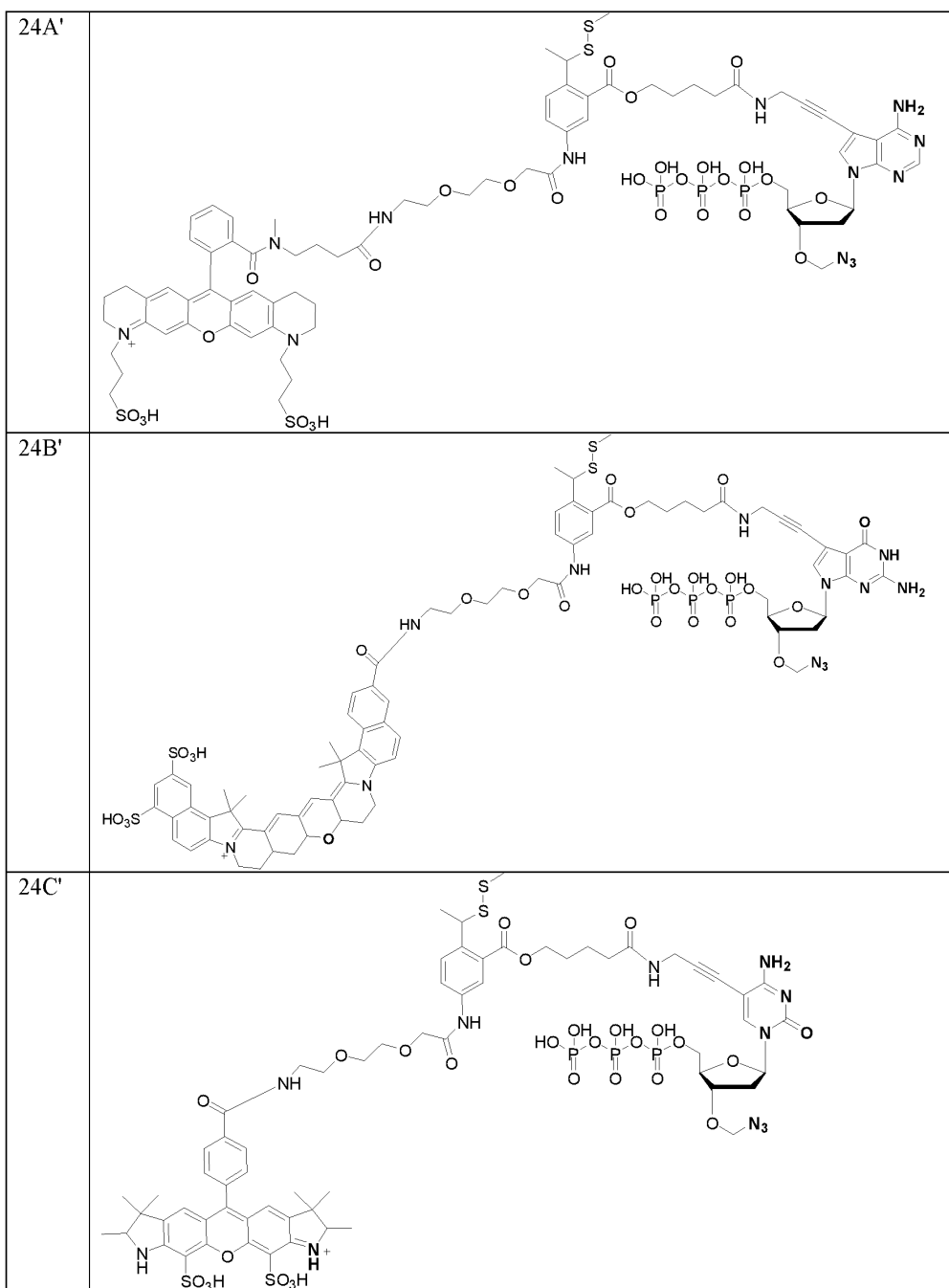
19D'	Chemical structure of compound 19D'. It features a central porphyrin-like macrocycle with a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃). The macrocycle is linked via a long, flexible chain containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃) to a side chain. This side chain includes a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃), and is further connected to a complex moiety containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃).
20A'	Chemical structure of compound 20A'. It features a central porphyrin-like macrocycle with a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃). The macrocycle is linked via a long, flexible chain containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃) to a side chain. This side chain includes a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃), and is further connected to a complex moiety containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃).
20B'	Chemical structure of compound 20B'. It features a central porphyrin-like macrocycle with a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃). The macrocycle is linked via a long, flexible chain containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃) to a side chain. This side chain includes a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃), and is further connected to a complex moiety containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃).
20C'	Chemical structure of compound 20C'. It features a central porphyrin-like macrocycle with a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃). The macrocycle is linked via a long, flexible chain containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃) to a side chain. This side chain includes a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃), and is further connected to a complex moiety containing a sulfonate group (SO₃H) and a trimethylammonium group (N⁺Me₃).

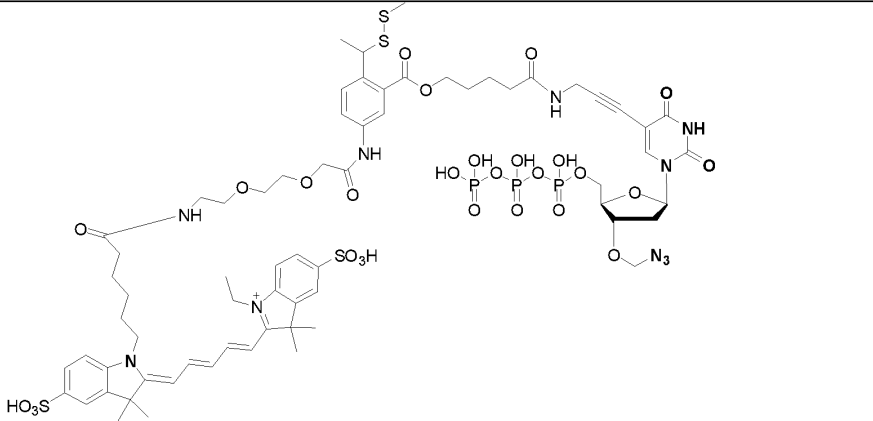
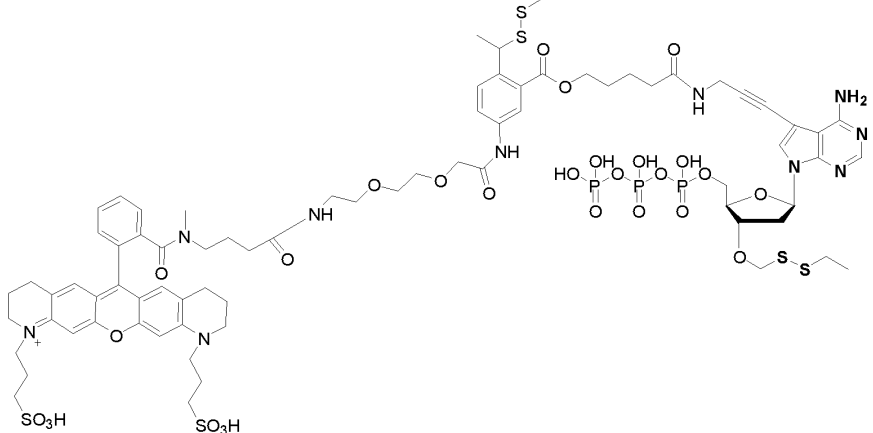
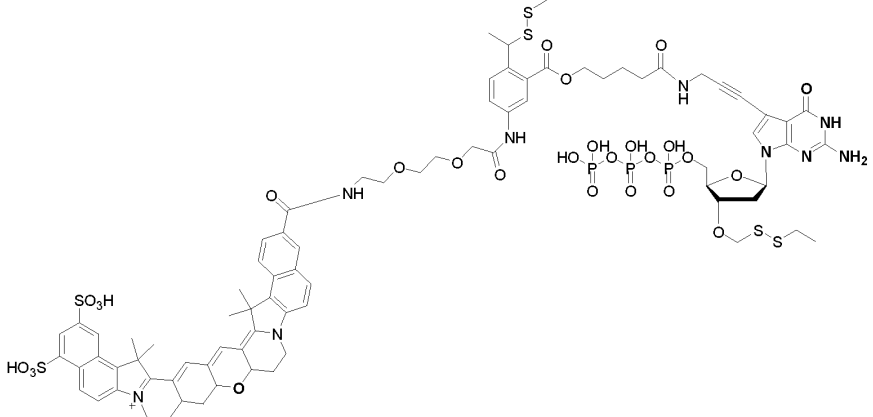
20D'	Chemical structure of compound 20D'. It features a complex molecule with a central chromophore consisting of a benzene ring fused to a five-membered ring containing a nitrogen atom with a positive charge. This is linked via a long chain containing a double bond and a sulfonate group (SO₃H) to a large, multi-ring system. A side chain includes a sulfonate group (SO₃H) and a long alkyl chain. Another side chain contains a sulfonate group (SO₃H) and a long alkyl chain. The molecule is terminated by a complex group including a sulfonate group (SO₃H), a long alkyl chain, and a complex heterocyclic system with a sulfonate group (SO₃H) and a long alkyl chain.
21A'	Chemical structure of compound 21A'. It features a complex molecule with a central chromophore consisting of a benzene ring fused to a five-membered ring containing a nitrogen atom with a positive charge. This is linked via a long chain containing a double bond and a sulfonate group (SO₃H) to a large, multi-ring system. A side chain includes a sulfonate group (SO₃H) and a long alkyl chain. Another side chain contains a sulfonate group (SO₃H) and a long alkyl chain. The molecule is terminated by a complex group including a sulfonate group (SO₃H), a long alkyl chain, and a complex heterocyclic system with a sulfonate group (SO₃H) and a long alkyl chain.
21B'	Chemical structure of compound 21B'. It features a complex molecule with a central chromophore consisting of a benzene ring fused to a five-membered ring containing a nitrogen atom with a positive charge. This is linked via a long chain containing a double bond and a sulfonate group (SO₃H) to a large, multi-ring system. A side chain includes a sulfonate group (SO₃H) and a long alkyl chain. Another side chain contains a sulfonate group (SO₃H) and a long alkyl chain. The molecule is terminated by a complex group including a sulfonate group (SO₃H), a long alkyl chain, and a complex heterocyclic system with a sulfonate group (SO₃H) and a long alkyl chain.
21C'	Chemical structure of compound 21C'. It features a complex molecule with a central chromophore consisting of a benzene ring fused to a five-membered ring containing a nitrogen atom with a positive charge. This is linked via a long chain containing a double bond and a sulfonate group (SO₃H) to a large, multi-ring system. A side chain includes a sulfonate group (SO₃H) and a long alkyl chain. Another side chain contains a sulfonate group (SO₃H) and a long alkyl chain. The molecule is terminated by a complex group including a sulfonate group (SO₃H), a long alkyl chain, and a complex heterocyclic system with a sulfonate group (SO₃H) and a long alkyl chain.

21D'	Chemical structure 21D' is a complex molecule. It features a central chain with a disulfide bridge (S-S) connecting a thiol group (-SH) to a thioether group (-S-). This chain is linked to a large, multi-ring system on the left, which includes a sulfonate group (-SO₃H) and a sulfonate group (-SO₃Na). The chain also connects to a nucleotide-like structure on the right, which includes a ribose sugar, a phosphate group, and a base (adenine derivative).
22A'	Chemical structure 22A' is a complex molecule. It features a central chain with a disulfide bridge (S-S) connecting a thiol group (-SH) to a thioether group (-S-). This chain is linked to a large, multi-ring system on the left, which includes a sulfonate group (-SO₃H) and a sulfonate group (-SO₃Na). The chain also connects to a nucleotide-like structure on the right, which includes a ribose sugar, a phosphate group, and a base (adenine derivative).
22B'	Chemical structure 22B' is a complex molecule. It features a central chain with a disulfide bridge (S-S) connecting a thiol group (-SH) to a thioether group (-S-). This chain is linked to a large, multi-ring system on the left, which includes a sulfonate group (-SO₃H) and a sulfonate group (-SO₃Na). The chain also connects to a nucleotide-like structure on the right, which includes a ribose sugar, a phosphate group, and a base (adenine derivative).
22C'	Chemical structure 22C' is a complex molecule. It features a central chain with a disulfide bridge (S-S) connecting a thiol group (-SH) to a thioether group (-S-). This chain is linked to a large, multi-ring system on the left, which includes a sulfonate group (-SO₃H) and a sulfonate group (-SO₃Na). The chain also connects to a nucleotide-like structure on the right, which includes a ribose sugar, a phosphate group, and a base (adenine derivative).

22D'	Chemical structure of compound 22D'. It features a central triphenylmethyl cation core. One phenyl ring is substituted with a sulfonate group (SO₃H). Another phenyl ring is linked via a long chain containing a disulfide bridge (S-S) to a thiol group (SH). A third phenyl ring is linked via a long chain containing a disulfide bridge (S-S) to a thiol group (SH). The central carbon is also bonded to a long chain containing a disulfide bridge (S-S) to a thiol group (SH). The structure is highly symmetrical and complex, with multiple sulfonate and disulfide groups.
23A'	Chemical structure of compound 23A'. It features a central triphenylmethyl cation core. One phenyl ring is substituted with a sulfonate group (SO₃H). Another phenyl ring is linked via a long chain containing a disulfide bridge (S-S) to a thiol group (SH). A third phenyl ring is linked via a long chain containing a disulfide bridge (S-S) to a thiol group (SH). The central carbon is also bonded to a long chain containing a disulfide bridge (S-S) to a thiol group (SH). The structure is highly symmetrical and complex, with multiple sulfonate and disulfide groups.
23B'	Chemical structure of compound 23B'. It features a central triphenylmethyl cation core. One phenyl ring is substituted with a sulfonate group (SO₃H). Another phenyl ring is linked via a long chain containing a disulfide bridge (S-S) to a thiol group (SH). A third phenyl ring is linked via a long chain containing a disulfide bridge (S-S) to a thiol group (SH). The central carbon is also bonded to a long chain containing a disulfide bridge (S-S) to a thiol group (SH). The structure is highly symmetrical and complex, with multiple sulfonate and disulfide groups.

23B''	
23C'	
23D'	



24D'	
25A'	
25B'	

25C'	Chemical structure 25C' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two N-methyl groups. Attached to the macrocycle is a long chain containing a disulfide bridge (S-S-CH₃), a carbonyl group, and a linker with a terminal amine group (NH₂). This linker is connected to a ribose sugar derivative, which is further linked to a pyrimidine ring system.
25D'	Chemical structure 25D' is similar to 25C' but with a different linker. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two N-methyl groups. Attached to the macrocycle is a long chain containing a disulfide bridge (S-S-CH₃), a carbonyl group, and a linker with a terminal amine group (NH₂). This linker is connected to a ribose sugar derivative, which is further linked to a pyrimidine ring system.
26A'	Chemical structure 26A' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two N-methyl groups. Attached to the macrocycle is a long chain containing a disulfide bridge (S-S-CH₃), a carbonyl group, and a linker with a terminal amine group (NH₂). This linker is connected to a ribose sugar derivative, which is further linked to a pyrimidine ring system.
26B'	Chemical structure 26B' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two N-methyl groups. Attached to the macrocycle is a long chain containing a disulfide bridge (S-S-CH₃), a carbonyl group, and a linker with a terminal amine group (NH₂). This linker is connected to a ribose sugar derivative, which is further linked to a pyrimidine ring system.

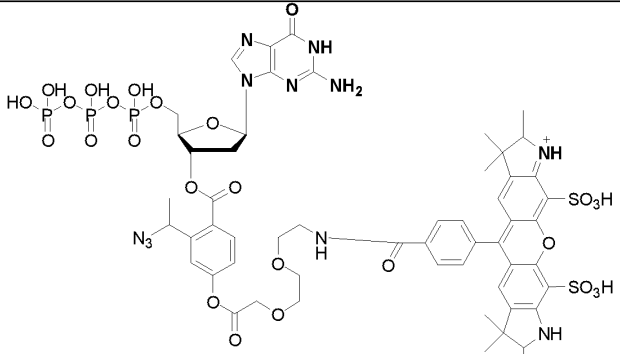
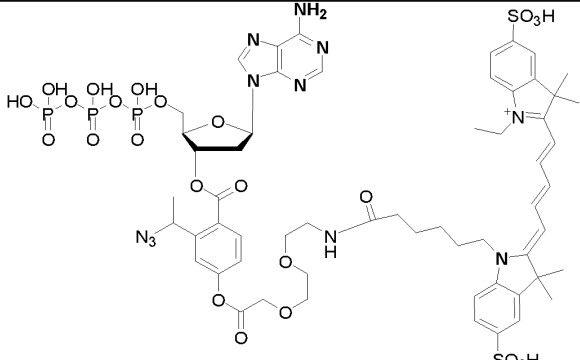
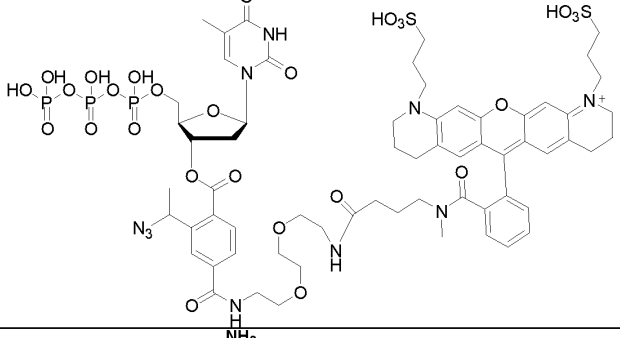
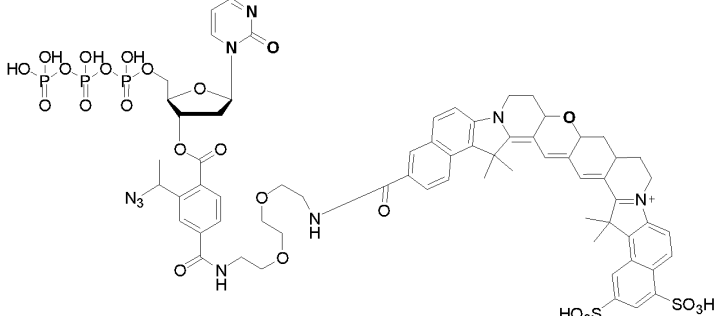
[illegible]

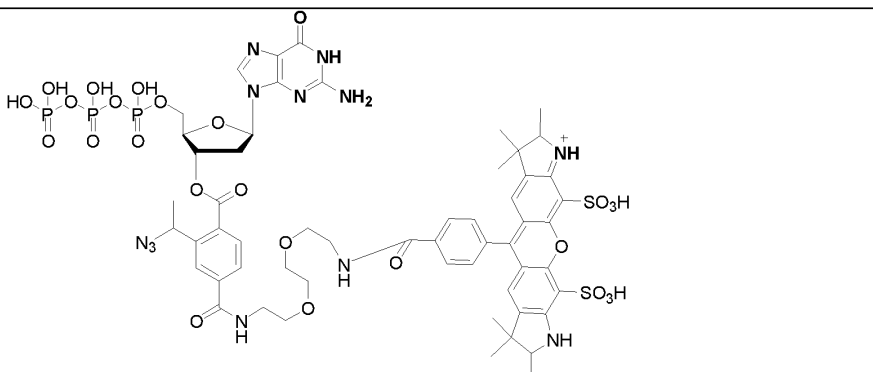
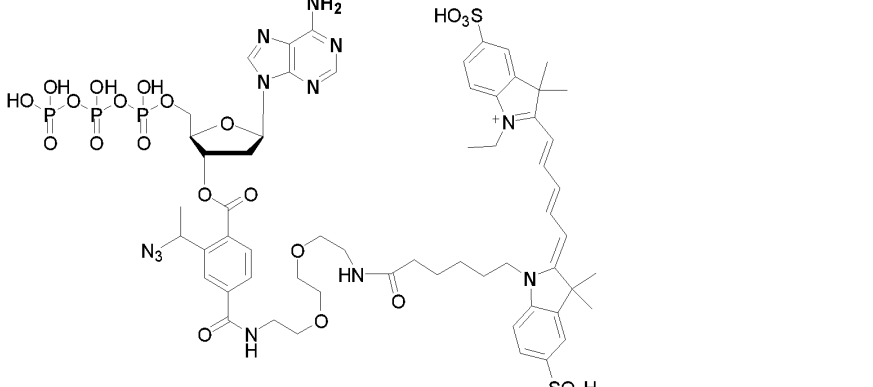
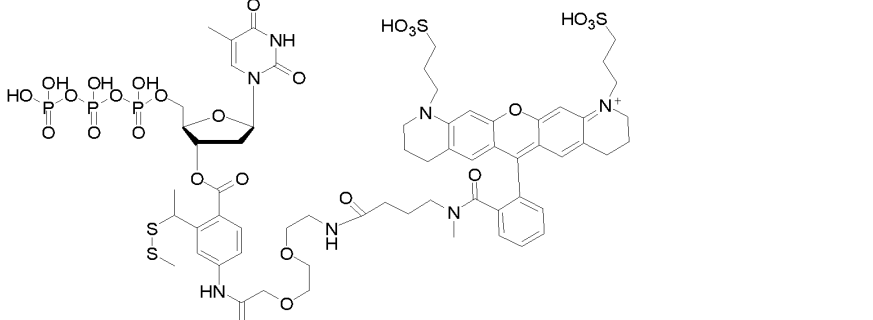
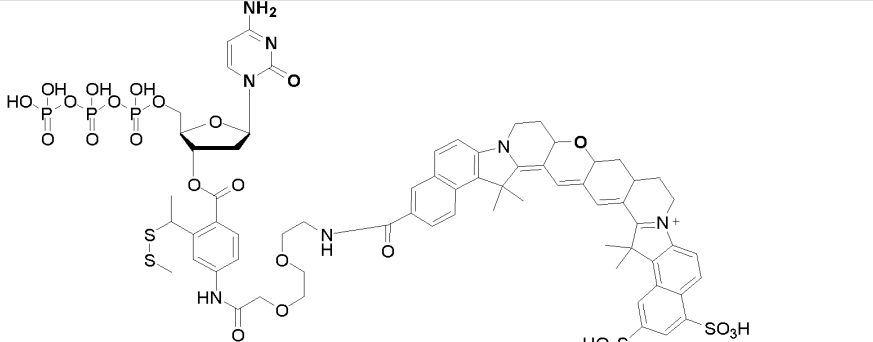
27C'	Chemical structure 27C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group includes a pyrimidine ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). The molecule also contains a sulfonate group (-SO₃H) and a sulfonamide group (-SO₂NH-).
27D'	Chemical structure 27D' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group includes a pyrimidine ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). The molecule also contains a sulfonate group (-SO₃H) and a sulfonamide group (-SO₂NH-).
28A'	Chemical structure 28A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group includes a pyrimidine ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). The molecule also contains a sulfonate group (-SO₃H) and a sulfonamide group (-SO₂NH-).
28B'	Chemical structure 28B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a terminal group. The terminal group includes a pyrimidine ring substituted with a sulfonamide group (-SO₂NH-) and a sulfonate group (-SO₃H). The molecule also contains a sulfonate group (-SO₃H) and a sulfonamide group (-SO₂NH-).

28C'	Chemical structure 28C' is a complex molecule featuring a central triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) linked to various functional groups. On the left, it connects to a long chain containing a benzamide derivative and a sulfonamide group. On the right, it connects to a ribose sugar derivative, which is further linked to a pyrimidine ring system. The molecule also includes a thiol group (-SH) and a disulfide bridge (-S-S-).
28D'	Chemical structure 28D' is a complex molecule featuring a central triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) linked to various functional groups. On the left, it connects to a long chain containing a benzamide derivative and a sulfonamide group. On the right, it connects to a ribose sugar derivative, which is further linked to a pyrimidine ring system. The molecule also includes a thiol group (-SH) and a disulfide bridge (-S-S-).
29A'	Chemical structure 29A' is a complex molecule featuring a central triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) linked to various functional groups. On the left, it connects to a long chain containing a benzamide derivative and a sulfonamide group. On the right, it connects to a ribose sugar derivative, which is further linked to a pyrimidine ring system. The molecule also includes a thiol group (-SH) and a disulfide bridge (-S-S-).
29B'	Chemical structure 29B' is a complex molecule featuring a central triphosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) linked to various functional groups. On the left, it connects to a long chain containing a benzamide derivative and a sulfonamide group. On the right, it connects to a ribose sugar derivative, which is further linked to a pyrimidine ring system. The molecule also includes a thiol group (-SH) and a disulfide bridge (-S-S-).

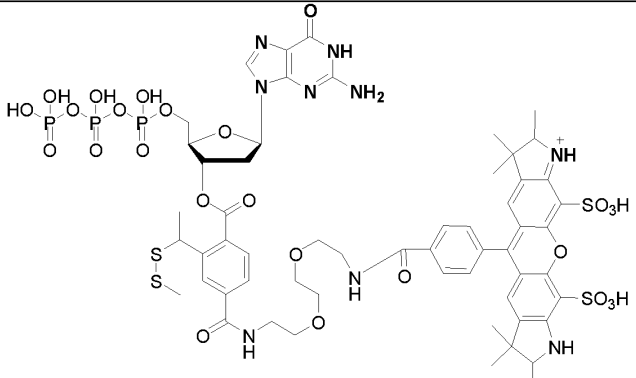
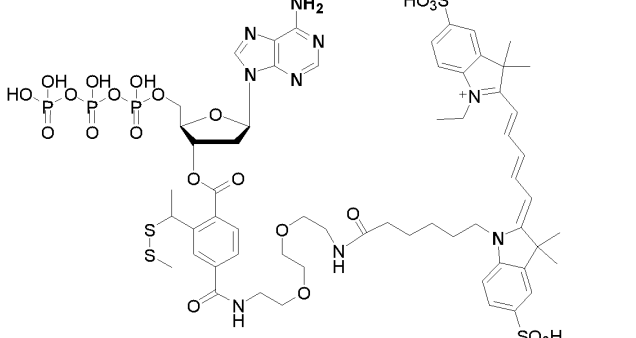
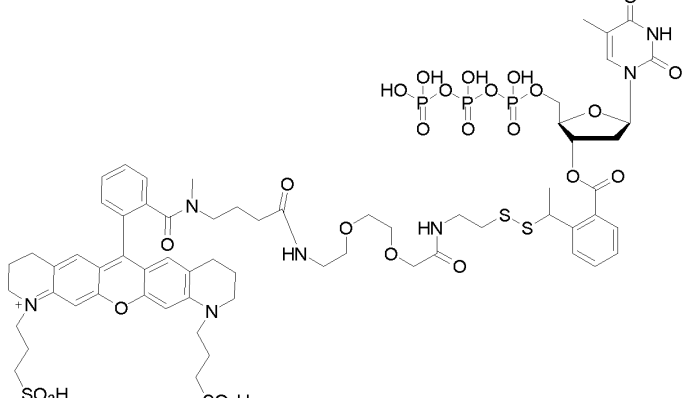
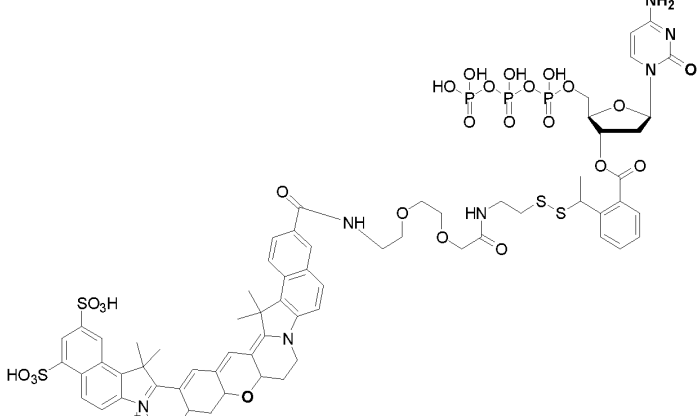
29C'	Chemical structure of compound 29C' is a complex molecule. It features a central xanthone core with two sulfonate groups (SO₃H) and a trimethylammonium group (N⁺Me₃). The xanthone is substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-). The triphosphate is linked to a ribose sugar, which is further connected to a purine base (adenine) via a glycosidic bond. The purine base is also substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-).
29D'	Chemical structure of compound 29D' is a complex molecule. It features a central xanthone core with two sulfonate groups (SO₃H) and a trimethylammonium group (N⁺Me₃). The xanthone is substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-). The triphosphate is linked to a ribose sugar, which is further connected to a purine base (adenine) via a glycosidic bond. The purine base is also substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-).
30A'	Chemical structure of compound 30A' is a complex molecule. It features a central xanthone core with two sulfonate groups (SO₃H) and a trimethylammonium group (N⁺Me₃). The xanthone is substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-). The triphosphate is linked to a ribose sugar, which is further connected to a purine base (adenine) via a glycosidic bond. The purine base is also substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-).
30B'	Chemical structure of compound 30B' is a complex molecule. It features a central xanthone core with two sulfonate groups (SO₃H) and a trimethylammonium group (N⁺Me₃). The xanthone is substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-). The triphosphate is linked to a ribose sugar, which is further connected to a purine base (adenine) via a glycosidic bond. The purine base is also substituted with a long chain containing a disulfide bridge (S-S), a benzoyl group, and a triphosphate group (HO-P(=O)(OH)-O-).

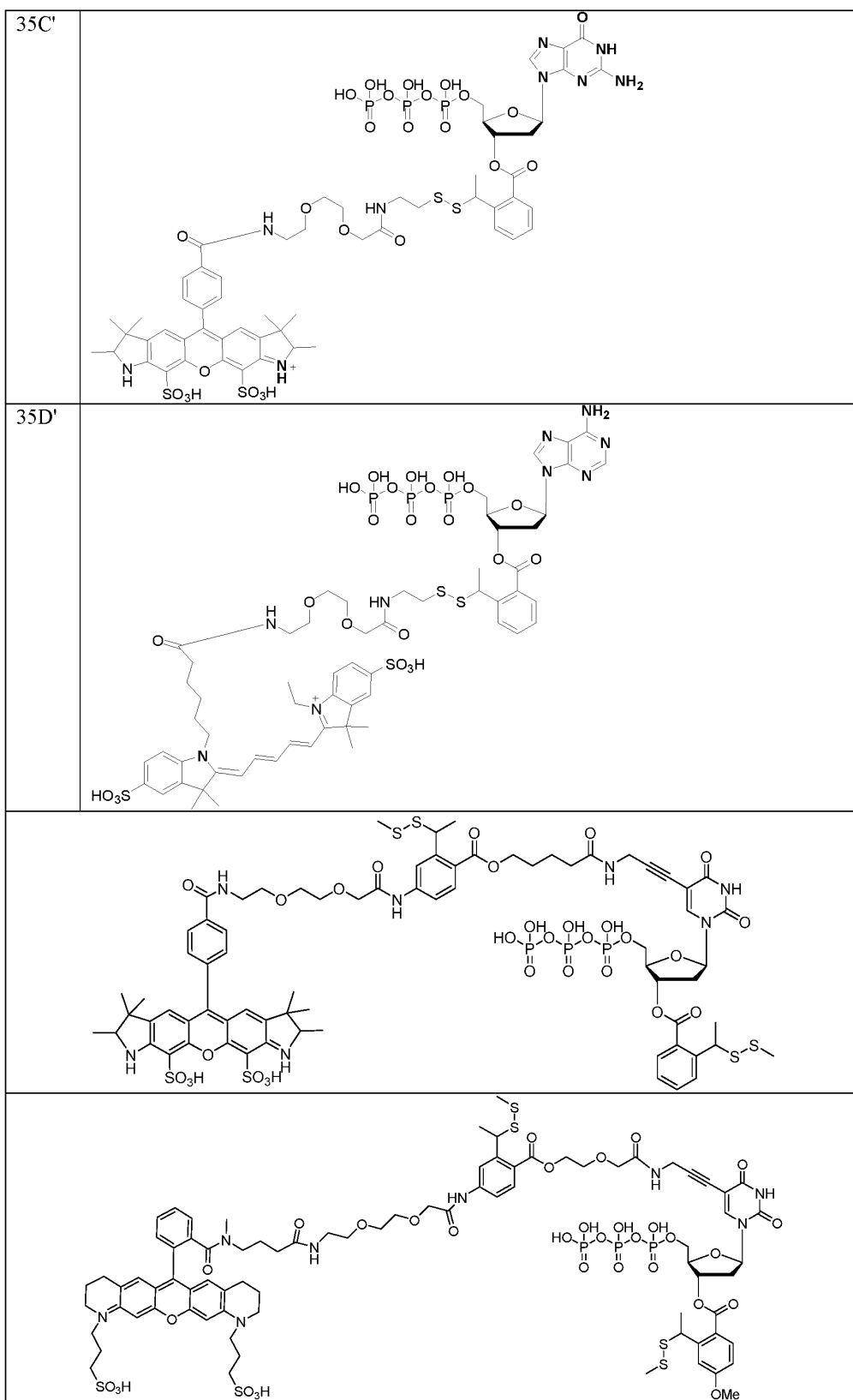
30C'	Chemical structure 30C' features a nucleotide backbone with a 5'-triphosphate group and a 3'-azido group. The sugar is linked to a pyrimidine base (uracil). The pyrimidine base is connected via a carbonyl group to a benzene ring, which is further linked to a side chain containing a sulfonamide group and a sulfonate group (SO₃H).
30D'	Chemical structure 30D' features a nucleotide backbone with a 5'-triphosphate group and a 3'-azido group. The sugar is linked to a pyrimidine base (uracil). The pyrimidine base is connected via a carbonyl group to a benzene ring, which is further linked to a side chain containing a sulfonamide group and a sulfonate group (SO₃H).
31A'	Chemical structure 31A' features a nucleotide backbone with a 5'-triphosphate group and a 3'-azido group. The sugar is linked to a pyrimidine base (uracil). The pyrimidine base is connected via a carbonyl group to a benzene ring, which is further linked to a side chain containing a sulfonamide group and a sulfonate group (SO₃H).
31B'	Chemical structure 31B' features a nucleotide backbone with a 5'-triphosphate group and a 3'-azido group. The sugar is linked to a pyrimidine base (uracil). The pyrimidine base is connected via a carbonyl group to a benzene ring, which is further linked to a side chain containing a sulfonamide group and a sulfonate group (SO₃H).

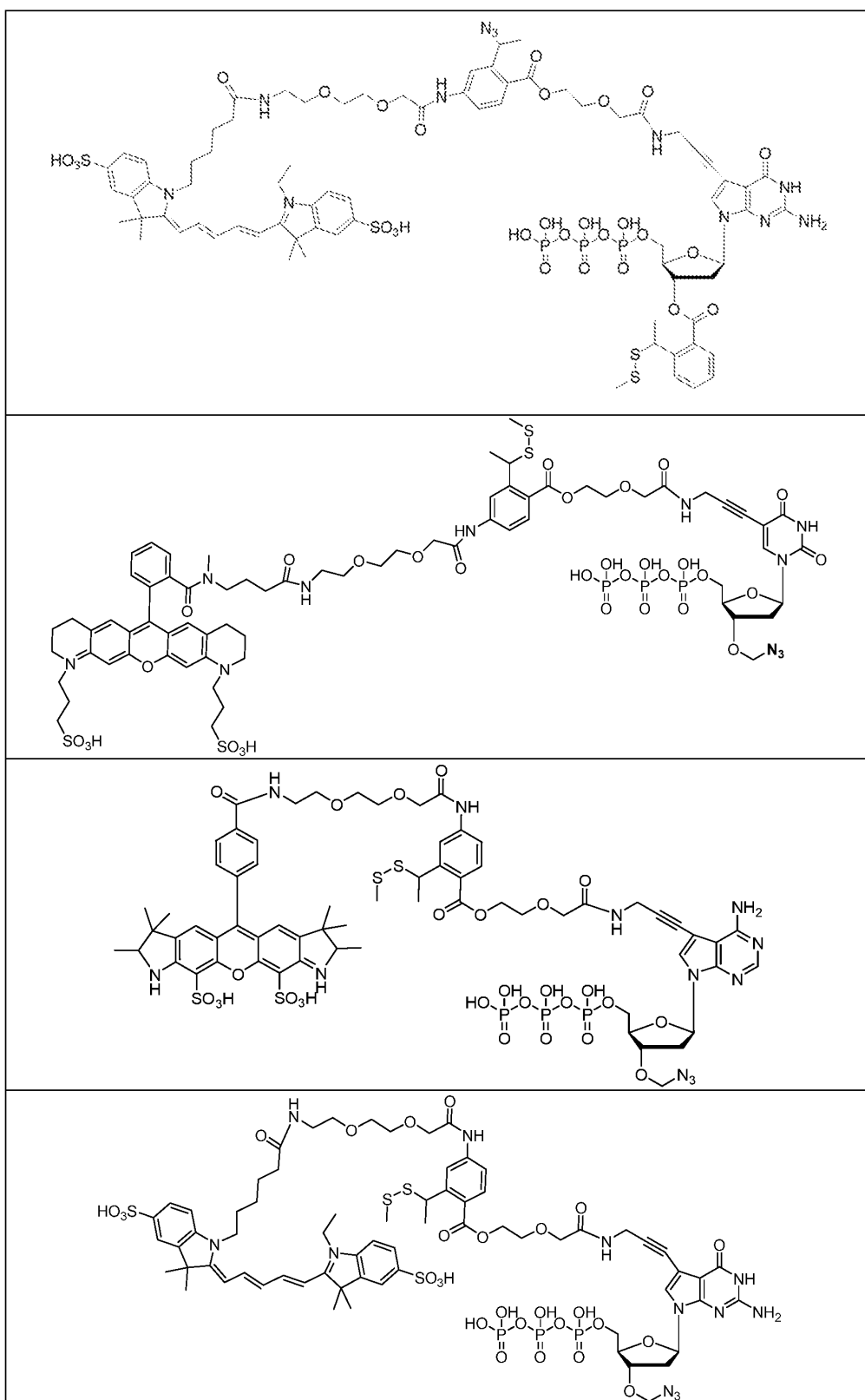
31C'	
31D'	
32A'	
32B'	

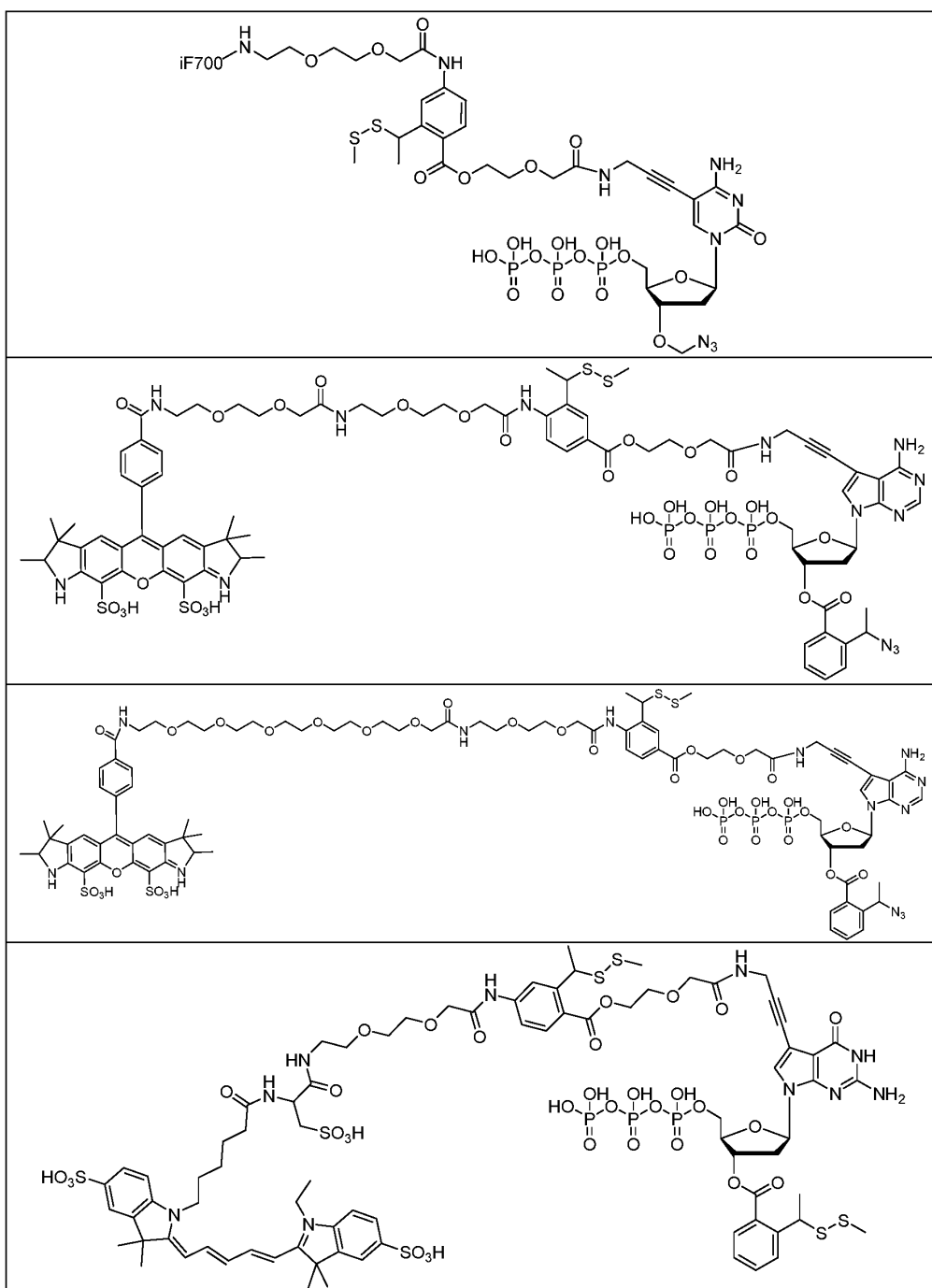
32C'	
32D'	
33A'	
33B'	

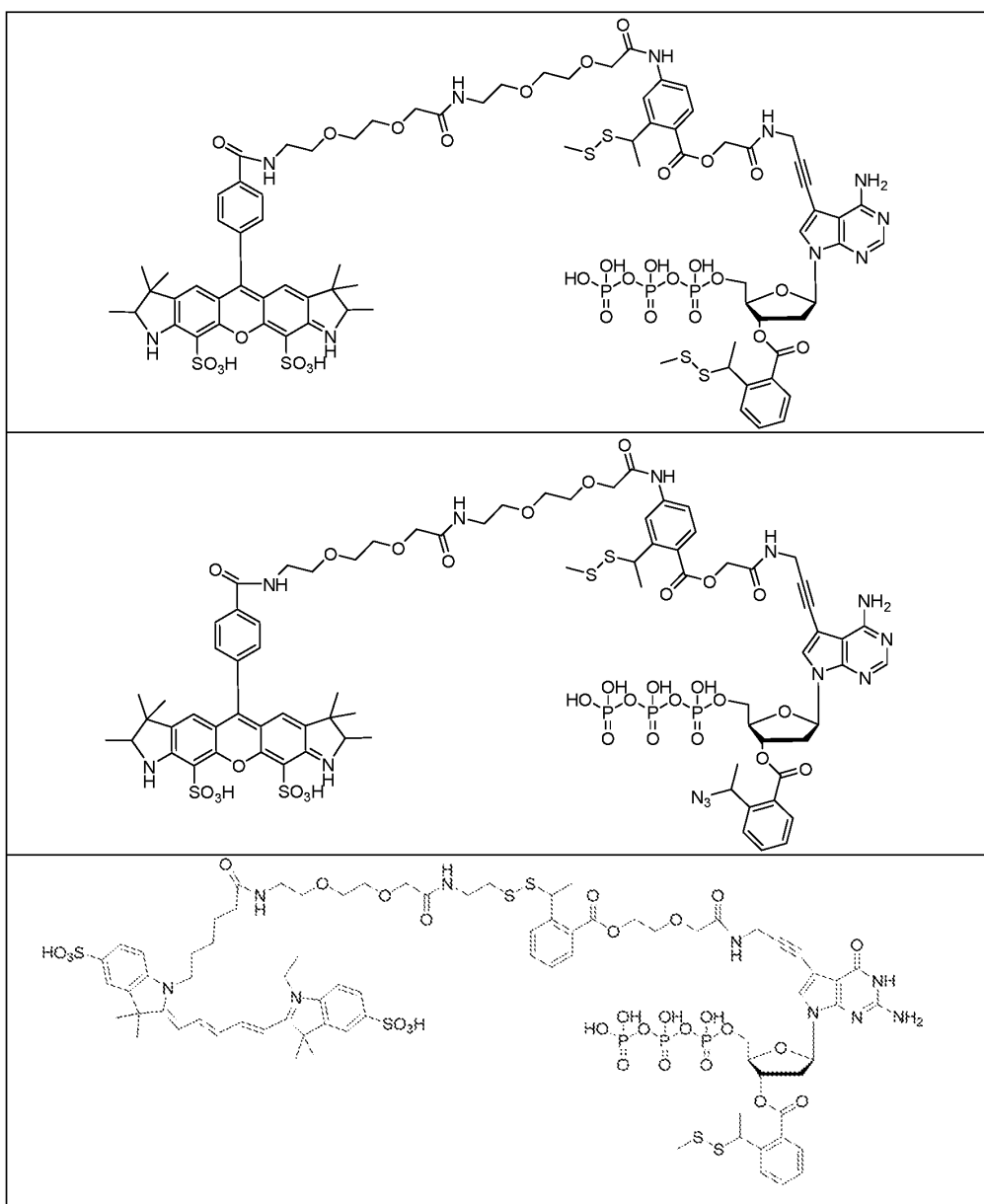
33C'	Chemical structure 33C' features a central nucleoside core. The sugar moiety is a ribose derivative with a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) attached to the 5' carbon. The base is a purine derivative with an amino group (NH₂) at the 6-position. The nucleoside is linked via an ester bond to a phenyl ring, which is further substituted with a thioether group (-S-S-) and a thioamide group (-NH-C(=O)-). The thioamide group is connected to a long chain containing a secondary amine and a carbonyl group, which is then linked to a complex polycyclic aromatic system. This system includes a naphthalene-like core with a quaternary ammonium group (N⁺) and two sulfonate groups (-SO₃H).
33D'	Chemical structure 33D' features a central nucleoside core. The sugar moiety is a ribose derivative with a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) attached to the 5' carbon. The base is a purine derivative with an amino group (NH₂) at the 6-position. The nucleoside is linked via an ester bond to a phenyl ring, which is further substituted with a thioether group (-S-S-) and a thioamide group (-NH-C(=O)-). The thioamide group is connected to a long chain containing a secondary amine and a carbonyl group, which is then linked to a complex polycyclic aromatic system. This system includes a naphthalene-like core with a quaternary ammonium group (N⁺) and two sulfonate groups (-SO₃H).
34A'	Chemical structure 34A' features a central nucleoside core. The sugar moiety is a ribose derivative with a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) attached to the 5' carbon. The base is a purine derivative with an amino group (NH₂) at the 6-position. The nucleoside is linked via an ester bond to a phenyl ring, which is further substituted with a thioether group (-S-S-) and a thioamide group (-NH-C(=O)-). The thioamide group is connected to a long chain containing a secondary amine and a carbonyl group, which is then linked to a complex polycyclic aromatic system. This system includes a naphthalene-like core with a quaternary ammonium group (N⁺) and two sulfonate groups (-SO₃H).
34B'	Chemical structure 34B' features a central nucleoside core. The sugar moiety is a ribose derivative with a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-) attached to the 5' carbon. The base is a purine derivative with an amino group (NH₂) at the 6-position. The nucleoside is linked via an ester bond to a phenyl ring, which is further substituted with a thioether group (-S-S-) and a thioamide group (-NH-C(=O)-). The thioamide group is connected to a long chain containing a secondary amine and a carbonyl group, which is then linked to a complex polycyclic aromatic system. This system includes a naphthalene-like core with a quaternary ammonium group (N⁺) and two sulfonate groups (-SO₃H).

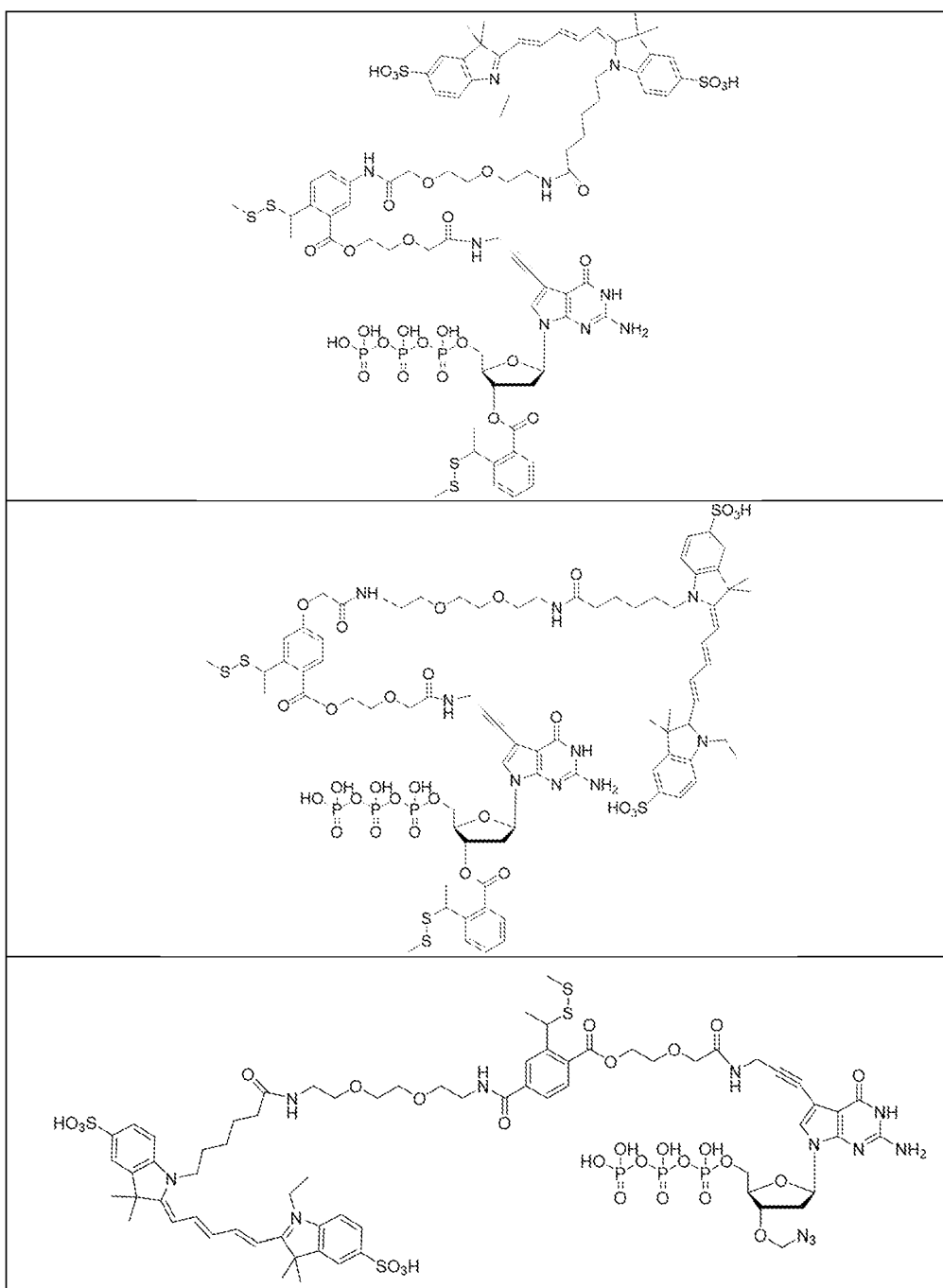
34C'	
34D'	
35A'	
35B'	











。在本发明的第七方面，本发明提供了终止核酸合成的方法，其包括：将前述的化合物或其盐掺入待终止的核酸分子中。

在一些实施方案中，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现。

在一些实施方案中，所述方法包括：使用聚合酶，将所述化合物或其盐掺入待终止的核酸分子中。

在一些实施方案中，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用

聚合酶进行核苷酸聚合反应，从而将所述化合物或其盐掺入待终止的核酸分子的3'端。

在本发明的第八方面，本发明提供了制备在测序反应中与目标单链多核苷酸互补的生长多核苷酸的方法，其包括将前述的化合物或其盐掺入所述生长的互补多核苷酸，其中，所述化合物或其盐的掺入防止了任何后续的核苷酸引入所述生长的互补多核苷酸中。

在一些实施方案中，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现。

在一些实施方案中，所述方法包括：使用聚合酶，将所述化合物或其盐掺入所述生长的互补多核苷酸。

在一些实施方案中，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述化合物或其盐掺入所述生长的互补多核苷酸的3'端。

在本发明的第九方面，本发明提供了核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是前述的化合物或其盐。

在本发明的第十方面，本发明提供了核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是前述的化合物或其盐，且所述生长的核酸链中预先掺入至少一个与目标单链多核苷酸互补的核苷酸，预先掺入的至少一个与目标单链多核苷酸互补的核苷酸是已被除去可逆阻断基团和任意的可检测标记的前述的化合物或其盐。

在本发明的第十一方面，本发明提供了测定目标单链多核苷酸的序列的方法，其包括：

1) 监测生长的核酸链中与目标单链多核苷酸互补的核苷酸的掺入，其中，掺入的至少一个互补核苷酸是前述的化合物或其盐，以及，

2) 确定掺入的核苷酸的类型。

在一些实施方案中，在引入下一个互补核苷酸之前，将所述可逆阻断基团和任意的可检测标记除去。

在一些实施方案中，所述可逆阻断基团和所述可检测标记被同时除去。

在一些实施方案中，所述可逆阻断基团和所述可检测标记被先后除去；例如，在所述可检测标记被除去之后，所述可逆阻断基团被除去，或者，在所述可逆阻断基团被除去之后，所述可检测标记被除去。

在一些实施方案中，所述测定目标单链多核苷酸的序列的方法包括以下步骤：

(a) 提供多种不同的核苷酸，其中至少一种核苷酸是前述的化合物或其盐，任选地其余的核苷酸是前述的化合物或其盐；

(b) 将所述多种不同的核苷酸掺入目标单链多核苷酸的互补序列中，其中，所述多种不同的核苷酸在检测时可以相互区分开；

(c) 检测(b)的核苷酸，从而确定掺入的核苷酸的类型；

(d) 除去(b)的核苷酸中的可逆阻断基团和任意的其携带的可检测标记；和

(e) 任选地重复步骤(a) - (d)一次或多次；

从而确定所述目标单链多核苷酸的序列。

在一些实施方案中，所述测定目标单链多核苷酸的序列的方法包括以下步骤：

(1) 提供第一核苷酸、第二核苷酸、第三核苷酸和第四核苷酸，四种核苷酸中的至少一种是前述的化合物或其盐，任选地其余的核苷酸是前述的化合物或其盐；

(2) 将所述四种核苷酸与目标单链多核苷酸进行接触；除去未掺入生长的核酸链中的所述核苷酸；检测掺入生长的核酸链中的所述核苷酸；除去掺入生长的核酸链中的所述核苷酸中的所述可逆阻断基团和任意的其携带的所述可检测标记；

任选地，还包括(3)：重复(1) - (2)一次或多次。

在一些实施方案中，所述测定目标单链多核苷酸的序列的方法包括以下步骤：

(a) 提供包含双链体、包含至少一种前述的化合物或其盐的核苷酸、聚合酶和切除试剂的混合物；所述双链体包含生长的核酸链以及待测序的核酸链；

(b) 进行包含以下步骤(i)、(ii)和(iii)的反应，任选地，重复一次或多次：

步骤 (i)：使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体；

步骤 (ii)：对所述核酸中间体进行检测；

步骤 (iii)：使用切除试剂将所述核酸中间体所包含的可逆阻断基团和任选的所述可检测标记切除。

在一些实施方案中，对所述可逆阻断基团的切除和对所述可检测标记的切除同时进行，或者，对所述可逆阻断基团的切除和对所述可检测标记的切除分步进行（例如，先切除所述可逆阻断基团，或者先切除所述可检测标记）。

在一些实施方案中，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是同样的试剂。

在一些实施方案中，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是不同的试剂。

在一些实施方案中，所述双链体连接于支持物上。

在一些实施方案中，所述生长的核酸链为引物。

在一些实施方案中，所述引物通过退火至待测序的核酸链上，形成所述双链体。

在一些实施方案中，所述双链体、所述化合物或其盐、以及所述聚合酶一起形成含有溶液相和固相的反应体系。

在一些实施方案中，在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体。

在一些实施方案中，所述聚合酶选自 KOD 聚合酶或其突变体（例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391）。

在一些实施方案中，在任意一个检测所述核酸中间体的步骤前，移除前一步骤的反应体系的溶液相，保留连接于支持物上的双链体。

在一些实施方案中，所述切除试剂与所述双链体或所述生长的核酸链在含有溶液相和固相的反应体系中接触。

在一些实施方案中，所述切除试剂能够切除掺入生长的核酸链的化合物中的可逆阻断基团和任选的其携带的可检测标记，并且不会影响双链体骨架上的磷酸二酯键。

在一些实施方案中，在任意一个切除所述核酸中间体所包含的可逆阻断基团和任选的可检测标记的步骤后，移除这一步骤反应体系的溶液相。

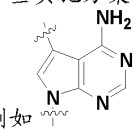
在一些实施方案中，在任意一个包含移除操作的步骤之后，进行洗涤操作。

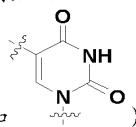
在一些实施方案中，步骤 (ii) 之后，进一步包括：根据步骤 (ii) 检测得到的信号，确定步骤 (i) 中掺入生长的核酸链的化合物的类型，并基于碱基互补配对原则，确定待测序的核酸链中相应位置处的核苷酸类型。

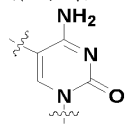
在本发明的第十二方面，本发明提供了试剂盒，其包含至少一个前述的化合物或其盐。

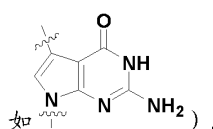
在一些实施方案中，所述试剂盒包含第一、第二、第三和第四化合物，所述第一、第二、第三和第四化合物各自独立地为前述的化合物或其盐。

在一些实施方案中，所述第一化合物中，base¹ 选自腺嘌呤、7-脱氧腺嘌呤或其互变异构

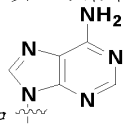
构体（例如 ）；所述第二化合物中，base¹ 选自胸腺嘧啶、尿嘧啶或其互变异构

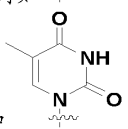
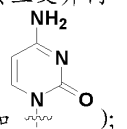
体（例如 ）；所述第三化合物中，base¹ 选自胞嘧啶或其互变异构体（例如

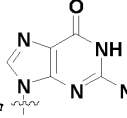
）；所述第四化合物中，base¹ 选自鸟嘌呤、7-脱氧鸟嘌呤或其互变异构体（例



在一些实施方案中,所述第一化合物中,base²选自腺嘌呤、7-脱氮腺嘌呤或其互变异构

构体(例如 ); 所述第二化合物中,base²选自胸腺嘧啶、尿嘧啶或其互变异构

体(例如 ); 所述第三化合物中,base²选自胞嘧啶或其互变异构体(例如 );

所述第四化合物中,base²选自鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体(例如 ).

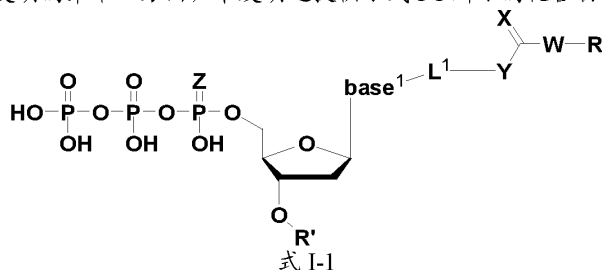
在一些实施方案中,所述第一、第二、第三和第四化合物包含的base¹或base²互不相同。

在一些实施方案中,所述第一、第二、第三和第四化合物携带的额外的可检测标记互不相同。

在一些实施方案中,所述试剂盒还包含: 用于预处理核酸分子的试剂; 用于连接待测序的核酸分子的支持物; 用于将待测序的核酸分子与支持物连接(例如, 共价或非共价连接)的试剂; 用于起始核苷酸聚合反应的引物; 用于进行核苷酸聚合反应的聚合酶; 一种或多种缓冲溶液; 一种或多种洗涤溶液; 或其任何组合。

在本发明的第十三方面,本发明提供了前述的化合物或其盐或者前述的试剂盒用于测定目标单链多核苷酸的序列的用途。

另外,在本发明的第十四方面,本发明还提供了式 I-1 所示的化合物或其盐,



其中:



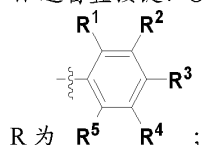
r¹、r²、r^{3a}、r^{3b}各自独立地选自 1-6 之间的任意整数;

M 选自直接键、CH₂、NH、O、S;

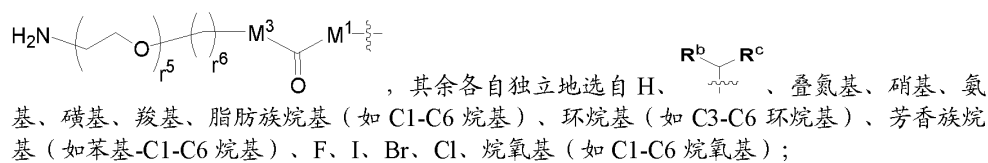
X 选自 O、S、NH;

Y 选自直接键、O、S、NH;

W 选自直接键、O、S、NH;



R¹、R²、R³、R⁴、R⁵中,任意一个(如 R²、R³或 R⁴)为



r^5 、 r^6 各自独立地选自 1-6 之间的任意整数；

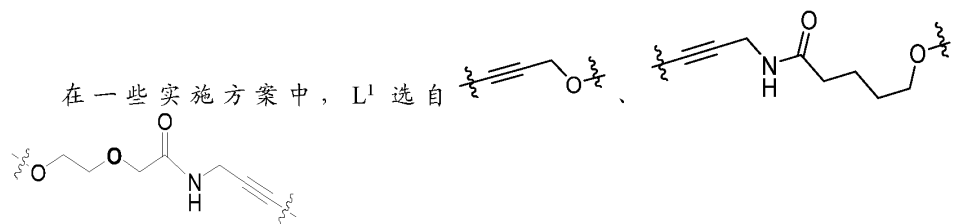
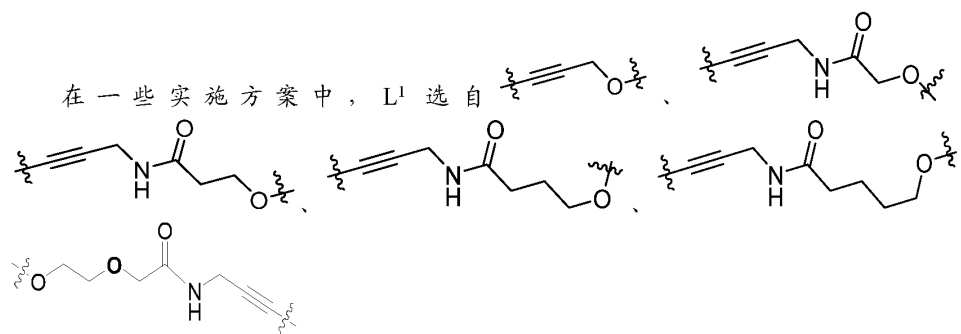
M^1 、 M^3 各自独立地选自直接键、NH、O、S；

R^b 、 R^c 中，任意一个选自 -N₃、-SS-C1-C6 烷基（如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu）、-ONH₂、-OCOR_x、-OCONHR_x，另一个选自 H、脂肪族烷基（如 C1-C6 烷基，例如 Me、Et、iPr、tBu）、芳香族烷基（如苯基-C1-C6 烷基）、环烷基（如 C3-C6 环烷基），其中，各 R_x 独立地选自脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）或芳香族烷基（如苯基 C1-C6 烷基）；

Z 选自 O，S，BH；

base¹ 选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；

R' 表示可逆阻断基团。



在一些实施方案中， r^1 选自 1、2、3。

在一些实施方案中， r^1 为 1。

在一些实施方案中， r^2 选自 1、2、3。

在一些实施方案中， r^2 为 1。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5。

在一些实施方案中， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3。

在一些实施方案中， r^{3a} 为 1。

在一些实施方案中， r^{3b} 为 2。

在一些实施方案中，M 选自 CH₂、O。

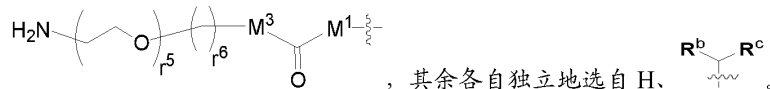
在一些实施方案中，X 选自 O、S。

在一些实施方案中，X 为 O。

在一些实施方案中，Y 为直接键。

在一些实施方案中，W 为直接键。

在一些实施方案中，R¹、R²、R³、R⁴、R⁵ 中，任意一个（如 R²、R³ 或 R⁴）为



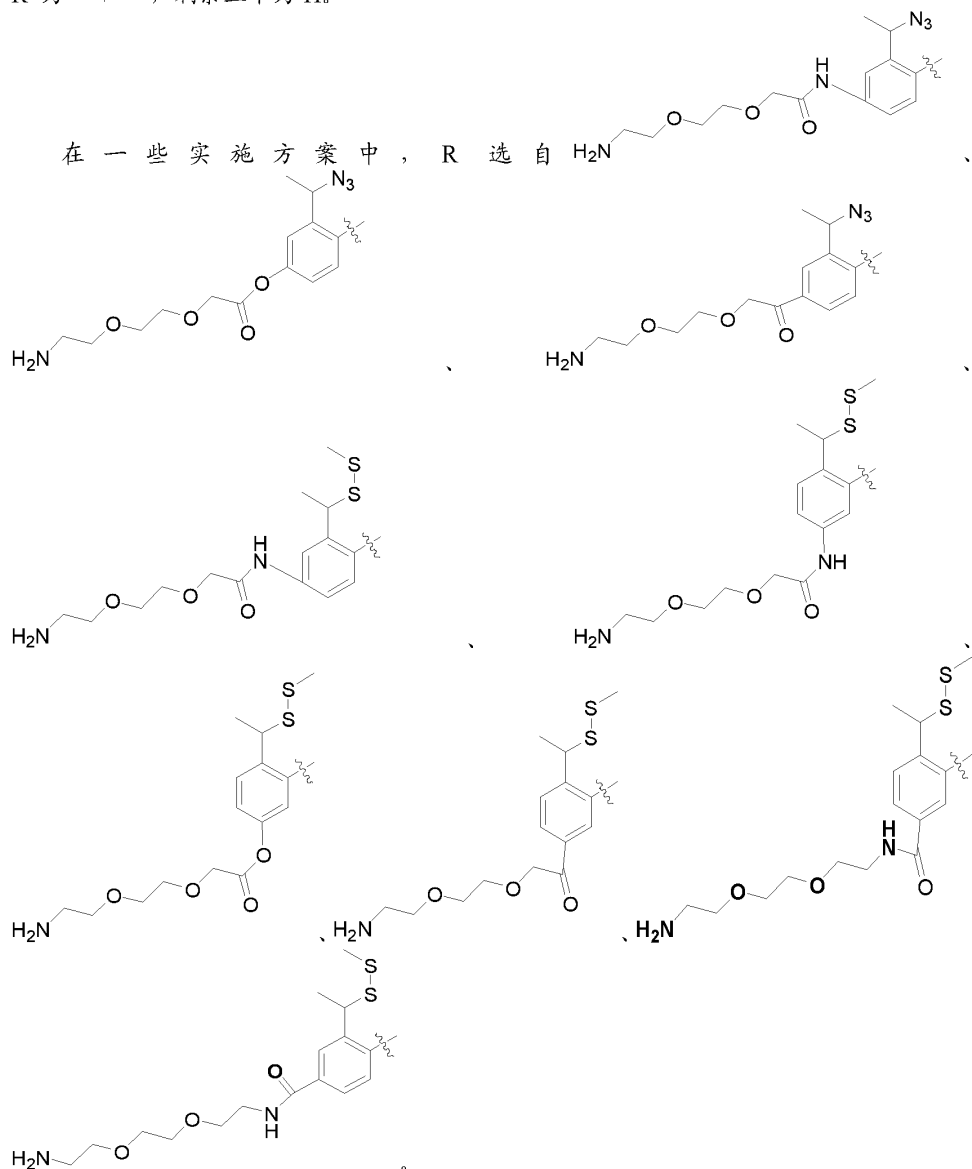
在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个（如 R^2 、 R^3 或 R^4 ）为

$$\text{H}_2\text{N}-(\text{CH}_2\text{CH}_2\text{O})_{r^5}-(\text{CH}_2\text{CH}_2\text{O})_{r^6}-\text{M}^3-\text{C}(=\text{O})-\text{M}^1-$$

，另一个（如 R^1 或 R^5 ）为 R^b 或 R^c ，剩余三个为 H。

在一些实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中， R^3 或 R^4 为

R^1 为 R^b 或 R^c ，剩余三个为 H。



在一些实施方案中， r^5 选自 1、2、3。

在一些实施方案中， r^5 为 2。

在一些实施方案中， r^6 选自 1、2、3。

在一些实施方案中， r^6 为 1 或 2。

在一些实施方案中， M^1 选自直接键、NH、O。

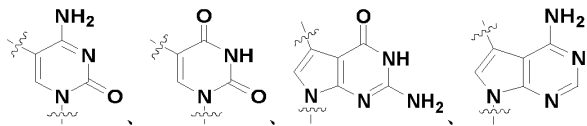
在一些实施方案中， M^3 选自直接键、NH。

在一些实施方案中, M^1 为 NH、O 或 S 时, M^3 为直接键, 且 M^3 为 NH、O 或 S 时, M^1 为直接键。

在一些实施方案中, R^b 、 R^c 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基 (如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$), 另一个选自 C1-C6 烷基。

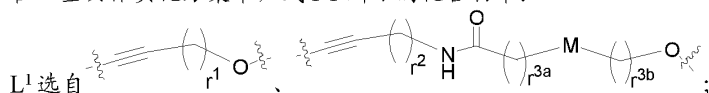
在一些实施方案中, R^b 、 R^c 中, 任意一个选自 $-N_3$ 、 $-SS-Me$, 另一个为甲基。

在一些实施方案中, Z 为 O。



在一些实施方案中, $base^1$ 选自

在一些具体实施方案中, 式 I-1 所示的化合物中:



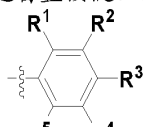
r^1 、 r^2 、 r^{3a} 、 r^{3b} 各自独立地选自 1-6 之间的任意整数;

M 选自直接键、 CH_2 、NH、O、S;

X 选自 O、S、NH;

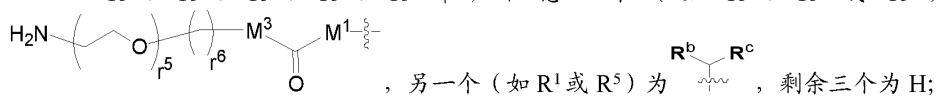
Y 选自直接键、O、S、NH;

W 选自直接键、O、S、NH;



R 为 R^5 ;

R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^2 、 R^3 或 R^4) 为



, 另一个 (如 R^1 或 R^5) 为

r^5 、 r^6 各自独立地选自 1-6 之间的任意整数;

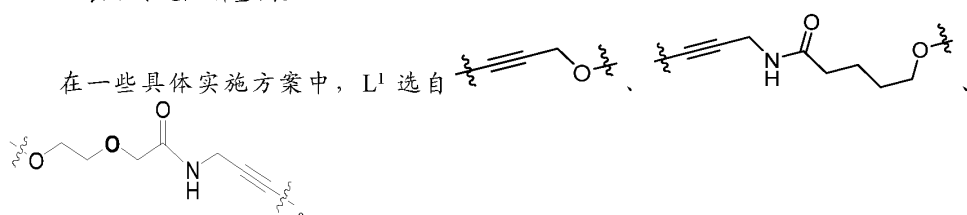
M^1 、 M^3 各自独立地选自直接键、NH、O、S;

R^b 、 R^c 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基 (如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$), 另一个选自 C1-C6 烷基;

Z 选自 O, S, BH;

$base^1$ 选自碱基、脱氮碱基或其互变异构体, 例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;

R' 表示可逆阻断基团。



在一些具体实施方案中, r^1 选自 1、2、3。

在一些具体实施方案中, r^1 为 1。

在一些具体实施方案中, r^2 选自 1、2、3。

在一些具体实施方案中, r^2 为 1。

在一些具体实施方案中, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5。

在一些具体实施方案中, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3。

在一些具体实施方案中, r^{3a} 为 1。

在一些具体实施方案中, r^{3b} 为 2。

在一些具体实施方案中, M 选自 CH_2 、O。

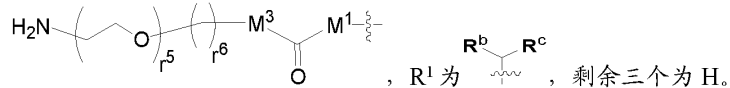
在一些具体实施方案中，X选自O、S。

在一些具体实施方案中，X为O。

在一些具体实施方案中，Y为直接键。

在一些具体实施方案中，W为直接键。

在一些具体实施方案中， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中， R^3 或 R^4 为



在一些具体实施方案中， r^5 选自1、2、3。

在一些具体实施方案中， r^5 为2。

在一些具体实施方案中， r^6 选自1、2、3。

在一些具体实施方案中， r^6 为1或2。

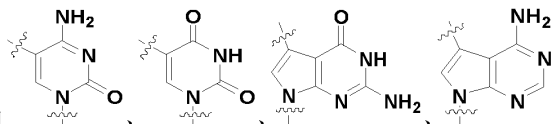
在一些具体实施方案中， M^1 选自直接键、NH、O。

在一些具体实施方案中， M^3 选自直接键、NH。

在一些具体实施方案中， M^1 为NH、O或S时， M^3 为直接键，且 M^3 为NH、O或S时， M^1 为直接键。

在一些具体实施方案中， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-Me$ ，另一个为甲基。

在一些具体实施方案中，Z为O。



在一些具体实施方案中， $base^1$ 选自

在一些实施方案中，前述可逆阻断基团 R' 选自 N_3 -C1-C6烷基、C1-C6烷基-SS-C1-C6

烷基、 R^2 、 $R^{1'}$ 、 R^3 、 $R^{4'}$ 、 R^5 、 NH_2 、 $-ONH_2$ 、 $-OCOR_z$ 、 $-OCONHR_z$ ，其中，各 R_z 独立地选自脂肪族烷基（如C1-C6烷基），环烷基（如C3-C6环烷基）或芳香族烷基（如苯基-C1-C6烷基）。

在一些实施方案中，所述可逆阻断基团 R' 选自 N_3 -C1-C6烷基、C1-C6烷基-SS-C1-C6

烷基、 R^2 、 $R^{1'}$ 、 R^3 、 $R^{4'}$ 、 R^5 、 NH_2 。

在一些实施方案中，所述可逆阻断基团 R' 选自 N_3 -C1-C6烷基、C1-C6烷基-SS-C1-C6

烷基、 R^2 、 $R^{1'}$ 、 R^3 、 $R^{4'}$ 、 R^5 。

在一些实施方案中，所述可逆阻断基团 R' 选自 N_3-CH_2- 、 $CH_3-CH_2-S-S-CH_2-$ 、

R^2 、 $R^{1'}$ 、 R^3 、 $R^{4'}$ 、 R^5 。

在一些实施方案中， $R^{1'}$ 、 R^2 、 R^3 、 $R^{4'}$ 、 R^5 中，任意一个（如 $R^{1'}$ 或 R^5 ）为 , 其余各自独立地选自H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基（如C1-C6烷基）、环烷基（如C3-C6环烷基）、芳香族烷基（如苯基-C1-C6烷基）、F、I、Br、Cl、烷氧基（如C1-C6烷氧基）。

在一些实施方案中，R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中，任意一个（如 R^{1'}或 R^{5'}）为 ，其余为 H。

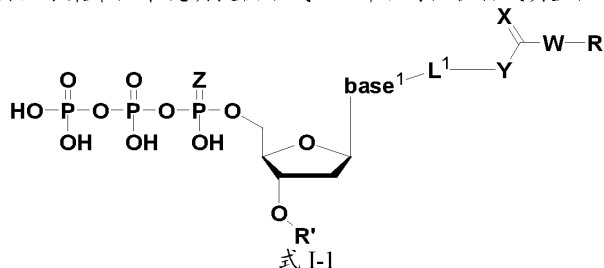
在一些实施方案中， $R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中， $R^{1'}$ 为 ，其余为H。

在一些实施方案中, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基 (如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$)、 $-ONH_2$ 、 $-OCOR_x$ 、 $-OCONHR_x$, 另一个选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me 、 Et 、 iPr 、 tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基 (如 C1-C6 烷基), 环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基)。

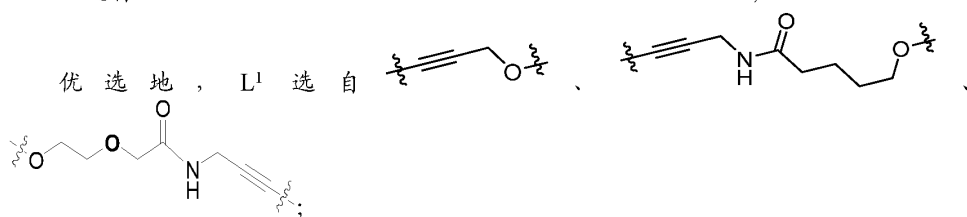
在一些实施方案中, R^{b'}、R^{c'}中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu), 另一个选自 C1-C6 烷基。

在一些实施方案中, R^{b'}、R^{c'}中, 任意一个选自-N₃、-SS-Me, 另一个为甲基。

在一些具体实施方案中, 本发明提供了式 I-1 所示的化合物或其盐,



其中：



r^1 、 r^2 、 r^{3a} 、 r^{3b} 各自独立地选自 1-6 之间的任意整数;

优选地, r^1 选自 1、2、3;

更优选地, r^1 为 1;

优选地, r^2 选自 1、2、3;

更优选地, r^2 为 1;

优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5;

更优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3;

最优选地, r^{3a} 为 1;

最优选地, r^{3b} 为 2;

M选自直接键、CH₂、NH、O、S;

优选地, M 选自 CH_2 、O;

X 选自 O、S、NH;

优选地, X 选自 O、S;

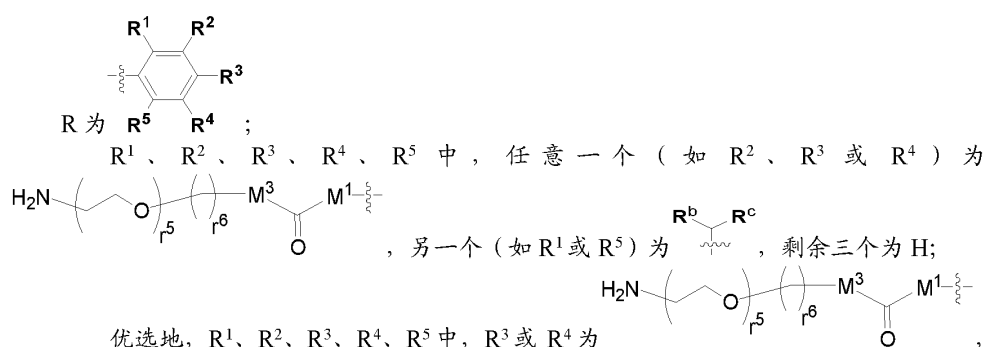
更优选地, X 为 0;

Y选自直接键、O、S、NH;

优选地, Y 为直接键;

W 选自直接键、O、S、NH;

优选地，W为直接键；



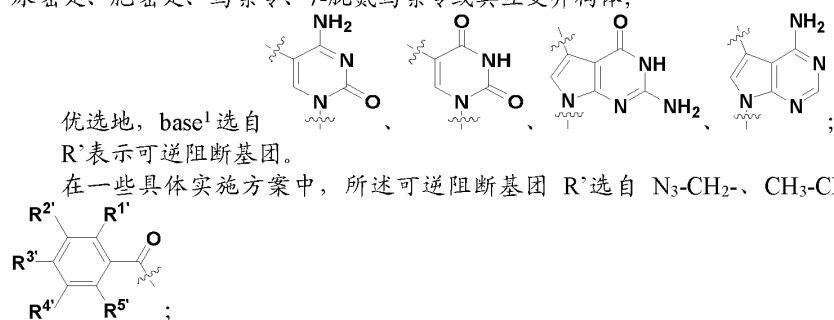
R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基（如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$ ），另一个选自 $C1-C6$ 烷基；

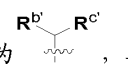
优选地， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-Me$ ，另一个为甲基；

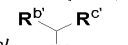
Z 选自 O、S、BH；

优选地， Z 为 O；

$base^1$ 选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；



R^1 、 R^2 、 R^3 、 R^4 、 R^5 中，任意一个（如 R^1 或 R^5 ）为 ，其余为 H，

优选地， R^1 、 R^2 、 R^3 、 R^4 、 R^5 中， R^1 为 ，其余为 H，

R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基（如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$ ），另一个选自 $C1-C6$ 烷基，

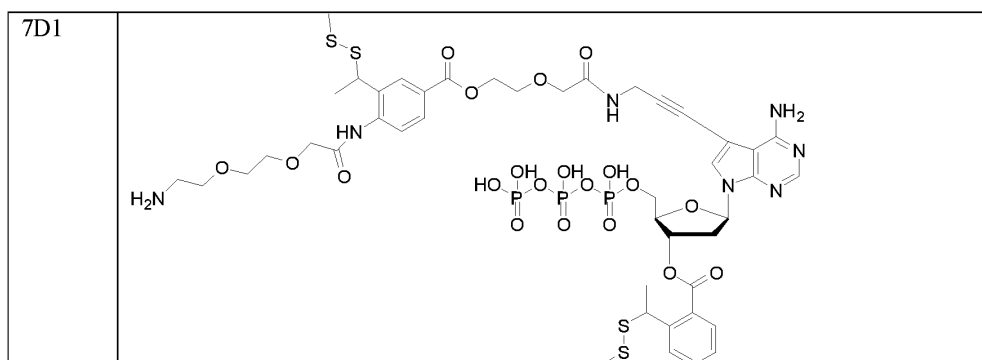
优选地， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-Me$ ，另一个为甲基。

在一些实施方案中，前述化合物选自以下表 C：

表 C：

编号	结构式
----	-----

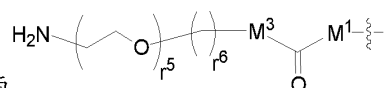
[illegible]



。在一些实施方案中，前述化合物或其盐携带额外的可检测标记。

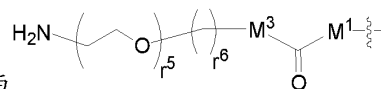
在一些实施方案中，所述化合物或其盐携带的额外的可检测标记是通过亲和试剂（如抗体、适体、亲和物、结蛋白）引入的，所述亲和试剂携带所述可检测标记，且所述亲和试剂可以特异性识别并结合所述化合物或其盐的表位。

在一些实施方案中，所述可检测标记与式 I-1 所示化合物中的 R 或前述具体化合物中对应于 R 的结构部分连接。



在一些实施方案中，所述可检测标记与
基连接。

中的末端氨



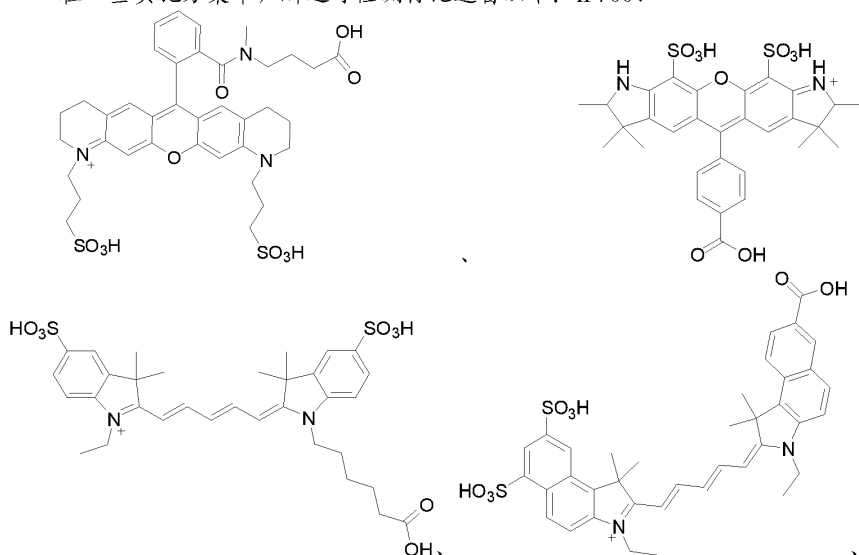
在一些实施方案中，所述可检测标记中的羧基与
的末端氨基通过形成酰胺键进行连接。

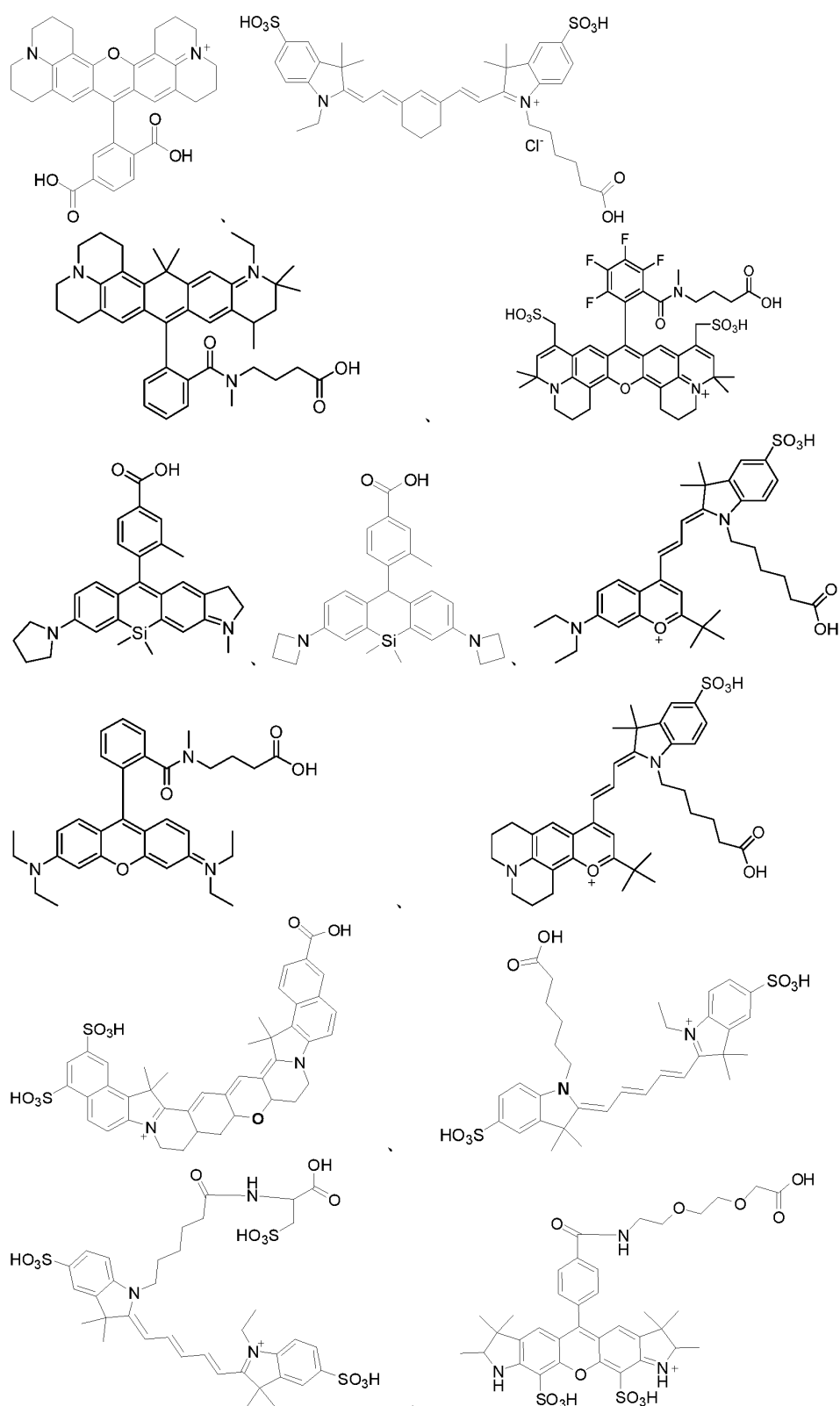
中

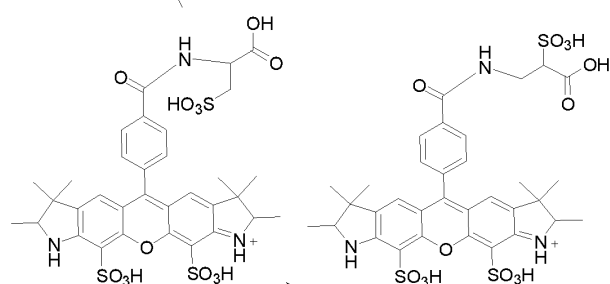
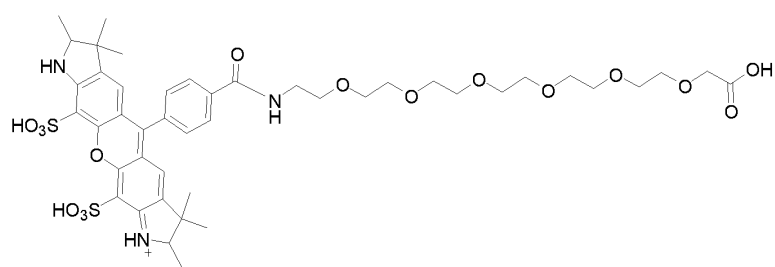
在一些实施方案中，base¹ 不同，式 I-1 所示化合物携带的额外的可检测标记不同。

在一些实施方案中，所述可检测标记为荧光标记。

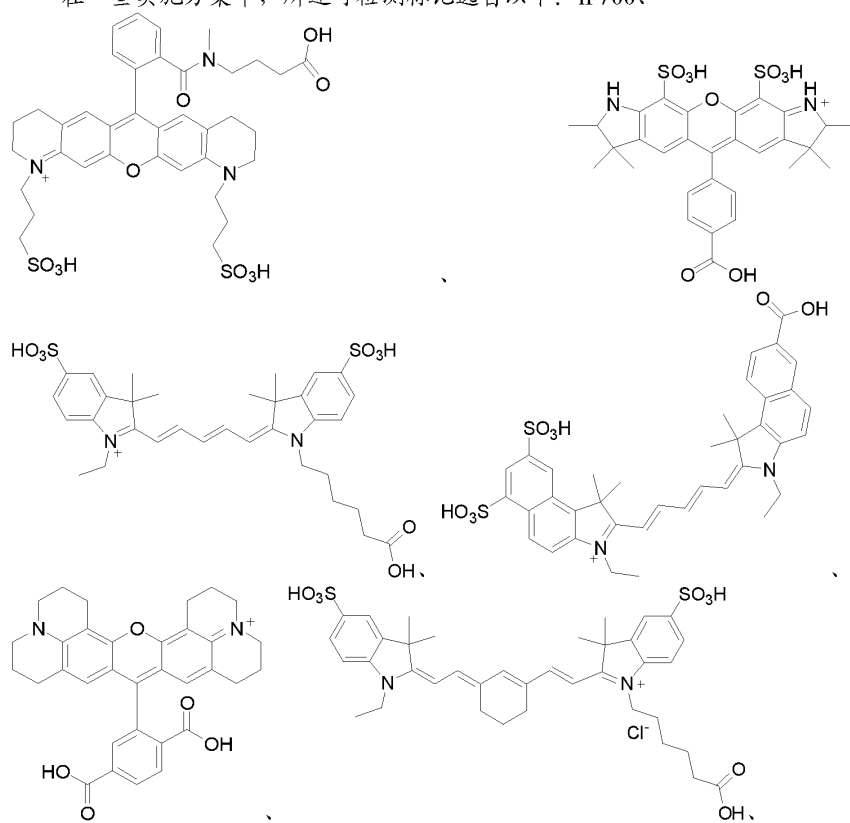
在一些实施方案中，所述可检测标记选自以下：iF700、

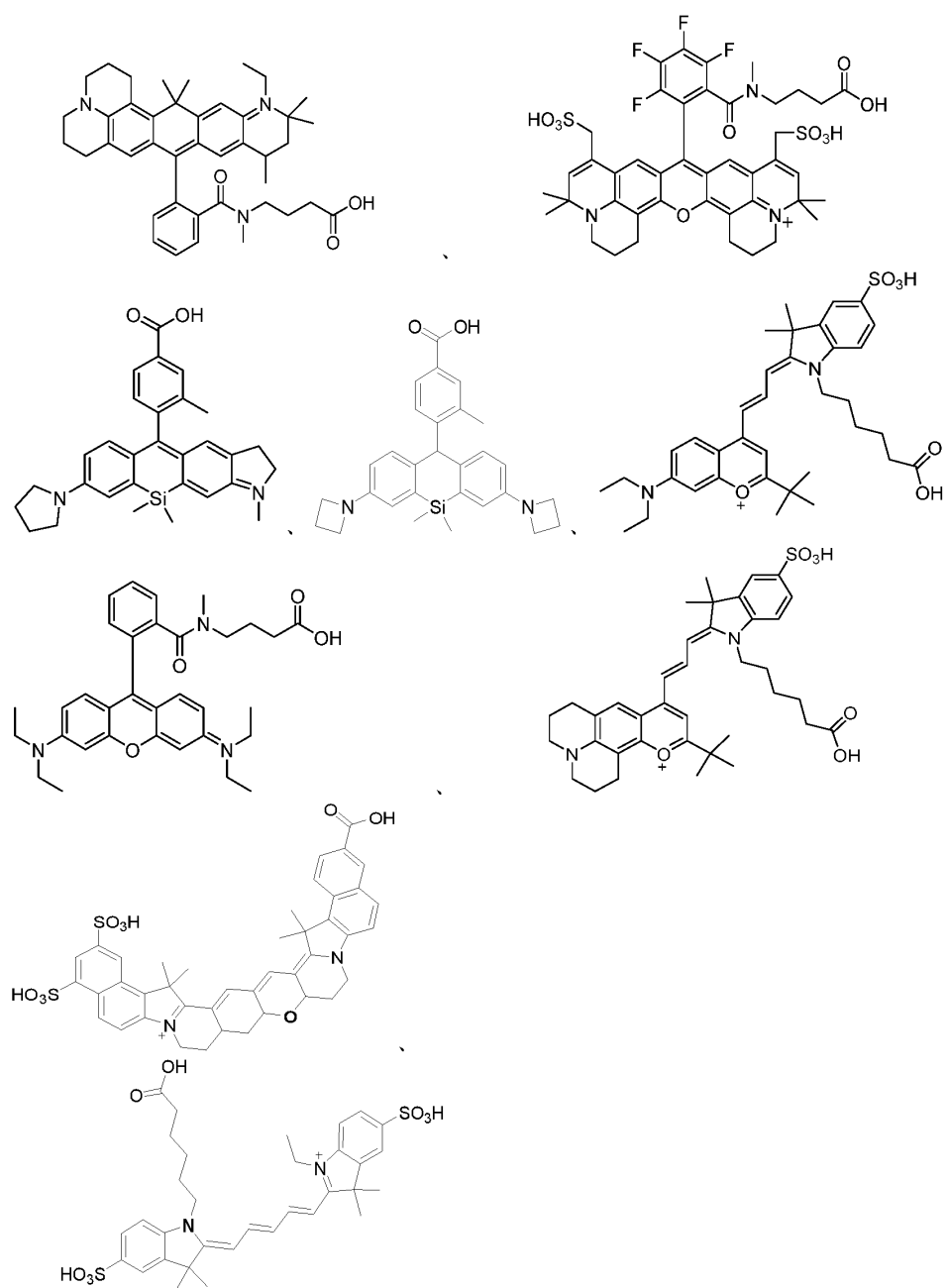






在一些实施方案中，所述可检测标记选自以下：iF700、

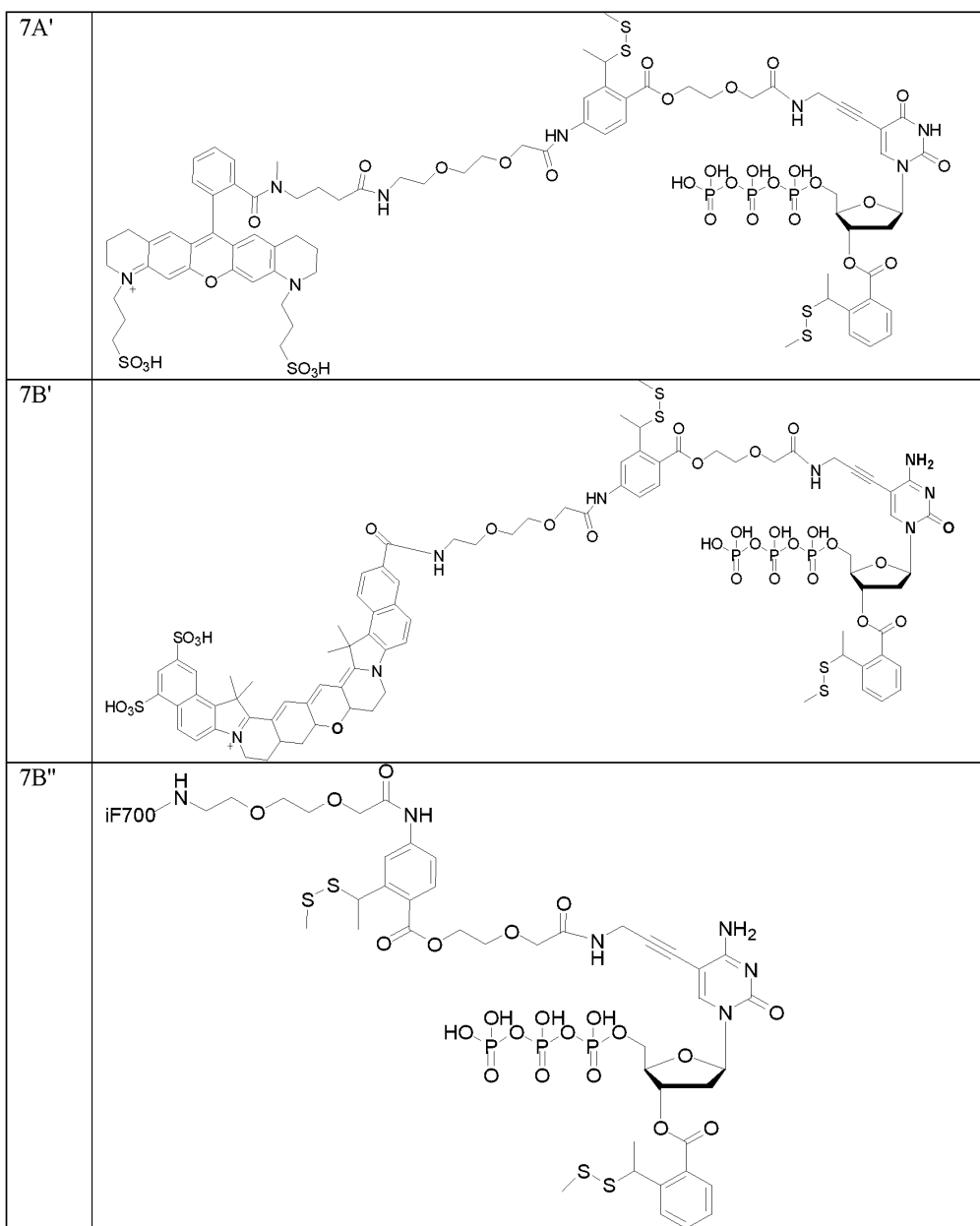




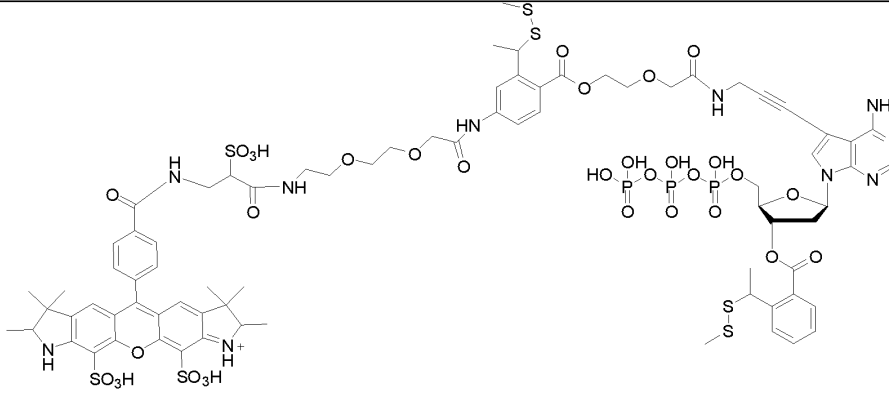
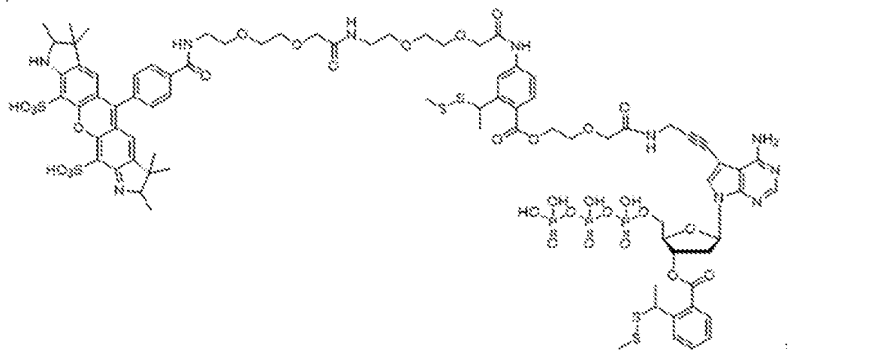
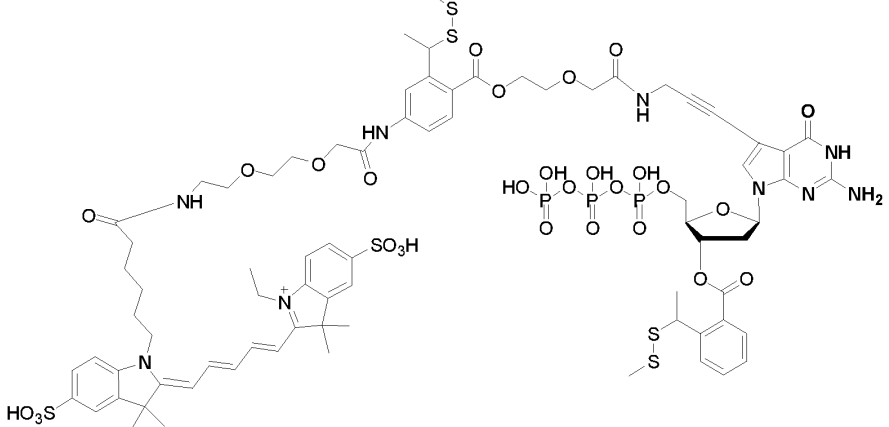
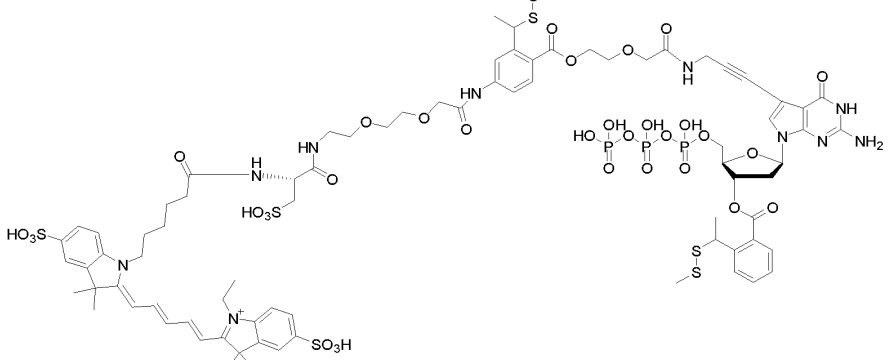
在一些实施方案中，前述化合物选自以下表 D:

表 D:

编号	结构式
----	-----



7C'	
7C''	
7C'''	
7C''''	

7C ^{''''}	
7C ^{'''''}	
7D'	
7D''	

。在本发明的第十五方面，本发明提供了终止核酸合成的方法，其包括：将前述的化合物或其盐掺入待终止的核酸分子中。

在一些实施方案中，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现。

在一些实施方案中，所述方法包括：使用聚合酶，将所述化合物或其盐掺入待终止的核酸分子中。

在一些实施方案中，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述化合物或其盐掺入待终止的核酸分子的3'端。

在本发明的第十六方面，本发明提供了制备在测序反应中与目标单链多核苷酸互补的生长的多核苷酸的方法，其包括将前述的化合物或其盐掺入所述生长的互补多核苷酸，其中，所述化合物或其盐的掺入防止了任何后续的核苷酸引入所述生长的互补多核苷酸中。

在一些实施方案中，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现。

在一些实施方案中，所述方法包括：使用聚合酶，将所述化合物或其盐掺入所述生长的互补多核苷酸。

在一些实施方案中，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述化合物或其盐掺入所述生长的互补多核苷酸的3'端。

在本发明的第十七方面，本发明提供了核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是前述的化合物或其盐。

在本发明的第十八方面，本发明提供了核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是前述的化合物或其盐，且所述生长的核酸链中预先掺入至少一个与目标单链多核苷酸互补的核苷酸，预先掺入的至少一个与目标单链多核苷酸互补的核苷酸是已被除去可逆阻断基团和任选的可检测标记的前述的化合物或其盐。

在本发明的第十九方面，本发明提供了测定目标单链多核苷酸的序列的方法，其包括：

1) 监测生长的核酸链中与目标单链多核苷酸互补的核苷酸的掺入，其中，掺入的至少一个互补核苷酸是前述的化合物或其盐，以及，

2) 确定掺入的核苷酸的类型。

在一些实施方案中，在引入下一个互补核苷酸之前，将所述可逆阻断基团和任选的可检测标记除去。

在一些实施方案中，所述可逆阻断基团和所述可检测标记被同时除去。

在一些实施方案中，所述可逆阻断基团和所述可检测标记被先后除去；例如，在所述可检测标记被除去之后，所述可逆阻断基团被除去，或者，在所述可逆阻断基团被除去之后，所述可检测标记被除去。

在一些实施方案中，所述测定目标单链多核苷酸的序列的方法包括以下步骤：

(a) 提供多种不同的核苷酸，其中至少一种核苷酸是前述的化合物或其盐，任选地其余的核苷酸是前述的化合物或其盐；

(b) 将所述多种不同的核苷酸掺入目标单链多核苷酸的互补序列中，其中，所述多种不同的核苷酸在检测时可以相互区分开；

(c) 检测(b)的核苷酸，从而确定掺入的核苷酸的类型；

(d) 除去(b)的核苷酸中的可逆阻断基团和任选的其携带的可检测标记；和

(e) 任选地重复步骤(a) - (d)一次或多次；

从而确定所述目标单链多核苷酸的序列。

在一些实施方案中，所述测定目标单链多核苷酸的序列的方法包括以下步骤：

(1) 提供第一核苷酸、第二核苷酸、第三核苷酸和第四核苷酸，四种核苷酸中的至少一种是前述的化合物或其盐，任选地其余的核苷酸是前述的化合物或其盐；

(2) 将所述四种核苷酸与目标单链多核苷酸进行接触；除去未掺入生长的核酸链中的所述核苷酸；检测掺入生长的核酸链中的所述核苷酸；除去掺入生长的核酸链中的所述核苷酸中的所述可逆阻断基团和任选的其携带的所述可检测标记；

任选地，还包括(3)：重复(1)-(2)一次或多次。

在一些实施方案中，所述测定目标单链多核苷酸的序列的方法包括以下步骤：

(a) 提供包含双链体、包含至少一种前述的化合物或其盐的核苷酸、聚合酶和切除试剂的混合物；所述双链体包含生长的核酸链以及待测序的核酸链；

(b) 进行包含以下步骤(i)、(ii)和(iii)的反应，任选地，重复一次或多次：

步骤(i)：使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体；

步骤(ii)：对所述核酸中间体进行检测；

步骤(iii)：使用切除试剂将所述核酸中间体所包含的可逆阻断基团和任选的所述可检测标记切除。

在一些实施方案中，对所述可逆阻断基团的切除和对所述可检测标记的切除同时进行，或者，对所述可逆阻断基团的切除和对所述可检测标记的切除分步进行（例如，先切除所述可逆阻断基团，或者先切除所述可检测标记）。

在一些实施方案中，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是同样的试剂。

在一些实施方案中，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是不同的试剂。

在一些实施方案中，所述双链体连接于支持物上。

在一些实施方案中，所述生长的核酸链为引物。

在一些实施方案中，所述引物通过退火至待测序的核酸链上，形成所述双链体。

在一些实施方案中，所述双链体、所述化合物或其盐、以及所述聚合酶一起形成含有溶液相和固相的反应体系。

在一些实施方案中，在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体。

在一些实施方案中，所述聚合酶选自 KOD 聚合酶或其突变体（例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391）。

在一些实施方案中，在任意一个检测所述核酸中间体的步骤前，移除前一步骤的反应体系的溶液相，保留连接于支持物上的双链体。

在一些实施方案中，所述切除试剂与所述双链体或所述生长的核酸链在含有溶液相和固相的反应体系中接触。

在一些实施方案中，所述切除试剂能够切除掺入生长的核酸链的化合物中的可逆阻断基团和任选的其携带的可检测标记，并且不会影响双链体骨架上的磷酸二酯键。

在一些实施方案中，在任意一个切除所述核酸中间体所包含的可逆阻断基团和任选的可检测标记的步骤后，移除这一步骤反应体系的溶液相。

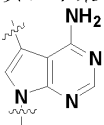
在一些实施方案中，在任意一个包含移除操作的步骤之后，进行洗涤操作。

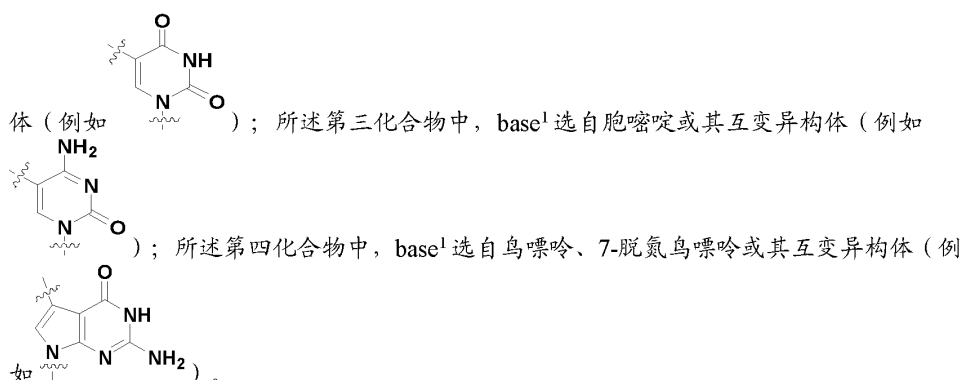
在一些实施方案中，步骤(ii)之后，进一步包括：根据步骤(ii)检测得到的信号，确定步骤(i)中掺入生长的核酸链的化合物的类型，并基于碱基互补配对原则，确定待测序的核酸链中相应位置处的核苷酸类型。

在本发明的第二十方面，本发明提供了试剂盒，其包含至少一个前述的化合物或其盐。

在一些实施方案中，所述试剂盒包含第一、第二、第三和第四化合物，所述第一、第二、第三和第四化合物各自独立地为前述的化合物或其盐。

在一些实施方案中，所述第一化合物中，base¹选自腺嘌呤、7-脱氧腺嘌呤或其互变异构

构体（例如 ）；所述第二化合物中，base¹选自胸腺嘧啶、尿嘧啶或其互变异构



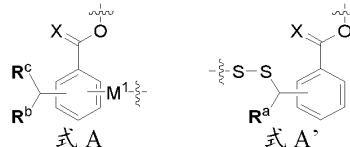
在一些实施方案中, 所述第一、第二、第三和第四化合物包含的 base¹ 互不相同。

在一些实施方案中, 所述第一、第二、第三和第四化合物携带的额外的可检测标记互不相同。

在一些实施方案中, 所述试剂盒还包含: 用于预处理核酸分子的试剂; 用于连接待测的核酸分子的支持物; 用于将待测的核酸分子与支持物连接(例如, 共价或非共价连接)的试剂; 用于起始核苷酸聚合反应的引物; 用于进行核苷酸聚合反应的聚合酶; 一种或多种缓冲溶液; 一种或多种洗涤溶液; 或其任何组合。

在本发明的第二十一方面, 本发明提供了前述的化合物或其盐或者前述的试剂盒用于测定目标单链多核苷酸的序列的用途。

另外, 在本发明的第二十二方面, 本发明还进一步提供了一种核苷酸类似物, 其由核糖或脱氧核糖、可逆阻断基团、碱基或脱氮碱基或其互变异构体、用于连接可检测标记的连接子和任选的磷酸基团形成, 其中, 所述连接子包含如下式 A 或式 A' 所示的结构:



其中:

R^b、R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu)、-ONH₂、-OCOR_x、-OCONHR_x、-S-SO₂R_x, 另一个选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基 (如 C1-C6 烷基), 环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基);

优选地, R^b、R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu), 另一个选自 C1-C6 烷基;

更优选地, R^b、R^c 中, 任意一个选自 -N₃、-SS-Me、-SS-Et, 另一个为甲基;

最优地, R^b、R^c 中, 任意一个选自 -N₃、-SS-Me, 另一个为甲基;

R^a 选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基);

优选地, R^a 选自 H、C1-C6 烷基;

更优选地, R^a 为甲基;

X 选自 O、S、NH;

优选地, X 选自 O、S;

更优选地, X 为 O;

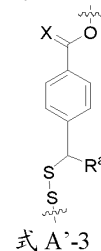
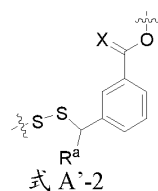
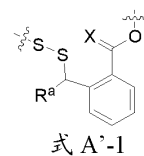
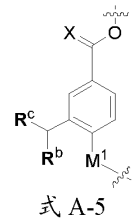
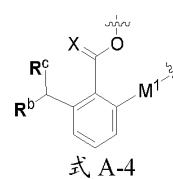
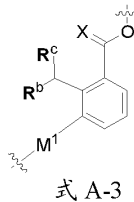
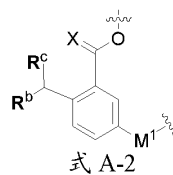
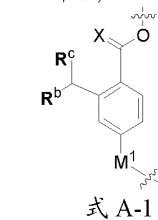
M¹ 选自直接键、NH、O、S、CH₂;

优选地, M¹ 选自直接键、NH、O;

优选地, 所述可逆阻断基团与所述核糖或脱氧核糖的 3'-OH 相连, 所述碱基或脱氮碱基或其互变异构体与所述核糖或脱氧核糖的 1'-C 相连, 所述任选的磷酸基团与所述核糖或

脱氧核糖的 5'-OH 相连, 所述连接子与所述碱基或脱氮碱基或其互变异构体相连。

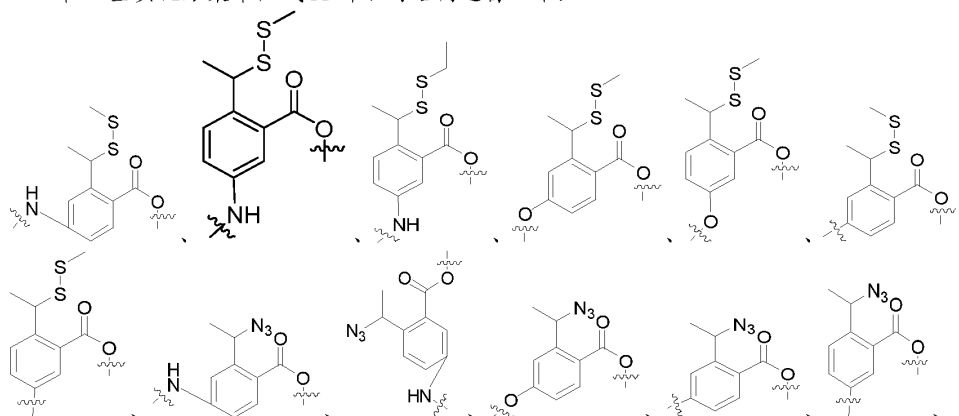
在一些实施方案中, 式 A 所示的结构如式 A-1、式 A-2、式 A-3、式 A-4 或式 A-5 所示, 优选如式 A-1、式 A-2 或式 A-5 所示; 或者, 式 A' 所示的结构如式 A'-1、式 A'-2 或式 A'-3 所示, 优选如式 A'-1 所示;

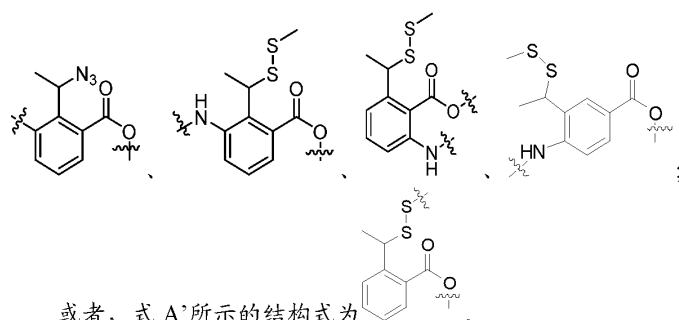


其中:

R^a、R^b、R^c、X、M¹ 的定义如前所述。

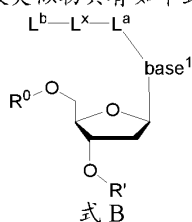
在一些实施方案中, 式 A 所示的结构选自以下:





或者, 式 A' 所示的结构式为

在一些实施方案中, 所述核苷酸类似物具有如下式 B 所示的结构:



L^x 具有前述的式 A、式 A'、式 A-1、式 A-2、式 A-3、式 A-4、式 A-5、式 A'-1、式 A'-2 或式 A'-3 所示的结构式, 且 L^x 的 M' 端或 S 端与 L^b 相连, L^x 的 O 端与 L^a 相连;

$L^b-L^x-L^a$ 表示用于连接可检测标记的连接子;

base¹ 表示碱基或脱氮碱基或其互变异构体;

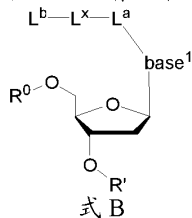
R' 表示可逆阻断基团;

R⁰ 表示 H 或磷酸基团;

L^a 表示所述连接子中用于连接碱基或脱氮碱基或其互变异构体的部分;

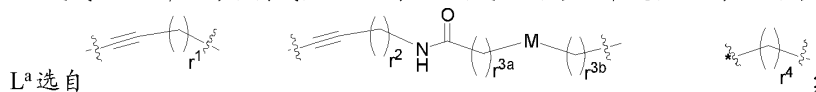
L^b 表示所述连接子中用于连接可检测标记的部分。

在一些实施方案中, 所述核苷酸类似物具有如下式 B 所示的结构:



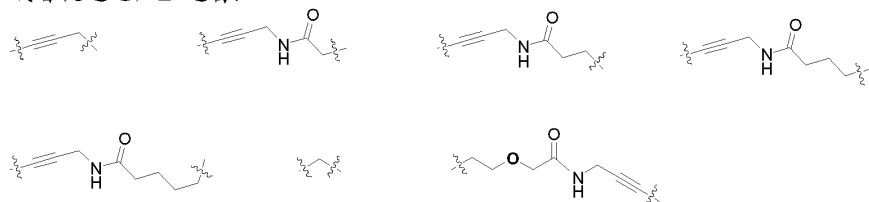
其中:

L^x 具有如前所述的式 A、式 A'、式 A-1、式 A-2、式 A-3、式 A-4、式 A-5、式 A'-1、式 A'-2 或式 A'-3 所示的结构式, 且 L^x 的 M' 端或 S 端与 L^b 相连, L^x 的 O 端与 L^a 相连;

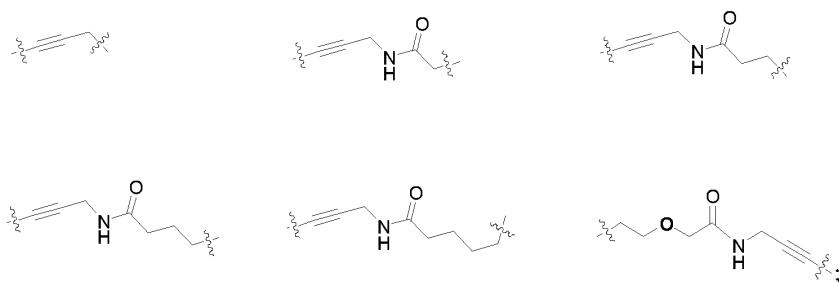


优选地, L^a 选自

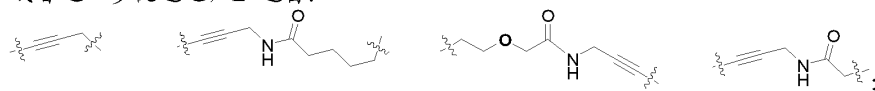
或者优选地, L^a 选自:



或者更优选地, L^a 选自:



或者进一步优选地, L^a 选自:



或者最优选地, L^a 选自:



r^1 、 r^2 、 r^{3a} 、 r^{3b} 、 r^4 各自独立地选自 1-6 之间的任意整数;

优选地, r^1 选自 1、2、3;

更优选地, r^1 为 1;

优选地, r^2 选自 1、2、3;

更优选地, r^2 为 1;

优选地, r^{3a} 、 r^{3b} 各自独立地选自 0、1、2、3、4、5, 且 r^{3a} 、 r^{3b} 不同时为 0;

更优选地, r^{3a} 、 r^{3b} 各自独立地选自 0、1、2、3, 且 r^{3a} 、 r^{3b} 不同时为 0;

进一步优选地, r^{3a} 选自 0、1, r^{3b} 选自 0、2, 且 r^{3a} 、 r^{3b} 不同时为 0;

最优选地, r^{3a} 为 1;

最优选地, r^{3b} 为 2;

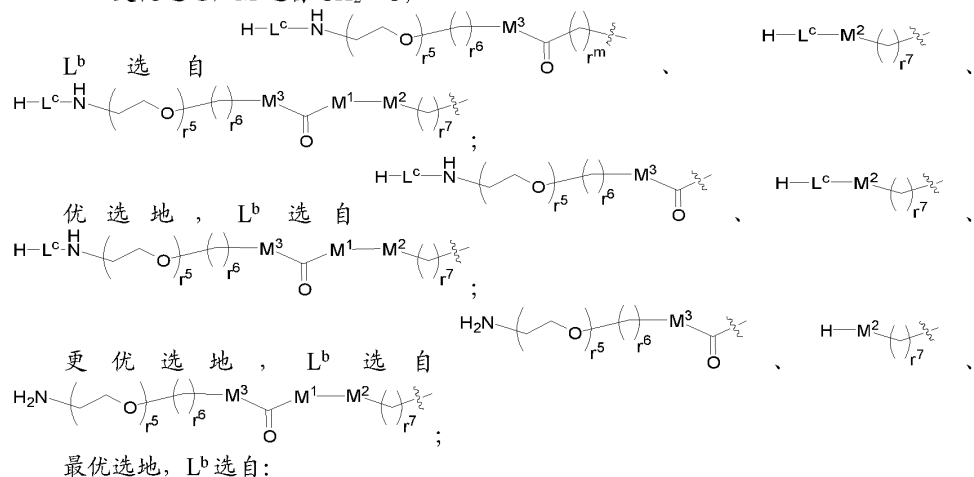
优选地, r^4 选自 1、2、3;

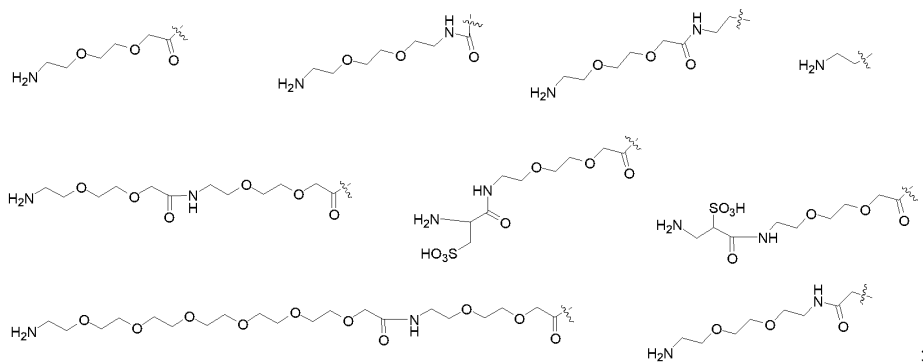
更优选地, r^4 为 1;

M 选自直接键、 CH_2 、NH、O、S;

优选地, M 选自直接键、 CH_2 、O;

更优选地, M 选自 CH_2 、O;





L^c 选自直接键、
 L^c 中的NH端与 L^b 中的H相连接， L^c 中的=O端与 L^b 中的NH相连接；

优选地， L^c 选自直接键、
 中的NH端与 L^b 中的H相连接， L^c 中的=O端与 L^b 中的NH相连接；

r^m 选自0-6之间的任意整数；

优选地， r^m 选自0、1、2、3；

更优选地， r^m 选自0、1；

r^5 、 r^6 、 r^7 各自独立地选自1-6之间的任意整数；

优选地， r^5 选自1、2、3；

更优选地， r^5 为2；

优选地， r^6 选自1、2、3；

更优选地， r^6 为1或2；

优选地， r^7 选自1、2、3；

更优选地， r^7 为2；

r^{10} 选自1-10之间的任意整数；

优选地， r^{10} 选自2-6之间的任意整数；

更优选地， r^{10} 为2或6；

r^{11} 选自1-6之间的任意整数；

优选地， r^{11} 选自1、2、3；

更优选地， r^{11} 为1；

M^1 、 M^2 、 M^3 各自独立地选自直接键、NH、O、S、 CH_2 ；

优选地， M^1 选自直接键、NH、O；

优选地， M^2 为NH；

优选地， M^3 选自直接键、NH；

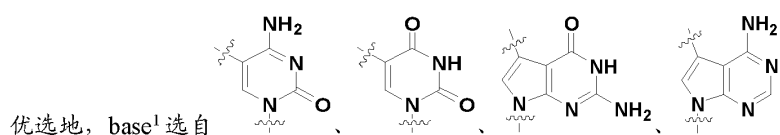
$base^1$ 表示碱基或脱氮碱基；

R' 表示可逆阻断基团；

$L^b-L^x-L^a$ 表示用于连接可检测标记的连接子；

R^0 表示H或磷酸基团。

在一些实施方案中，所述 $base^1$ 选自腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；



优选地, base¹ 选自

在一些实施方案中, R⁰ 选自 H、单磷酸基团()、二磷酸基团()、

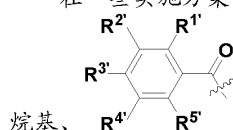
三磷酸基团()、四磷酸基团();

优选地, R⁰ 为三磷酸基团();

各 Z 独立地选自 O, S, BH;

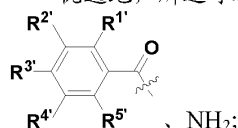
优选地, Z 为 O。

在一些实施方案中, 所述可逆阻断基团 R' 选自 N₃-C1-C6 烷基、C1-C6 烷基-SS-C1-C6



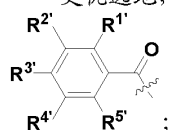
烷基、R^{4'}、R^{5'}、NH₂、-ONH₂、-OCOR_Z、-OCONHR_Z, 其中, 各 R_Z 独立地选自脂肪族烷基(如 C1-C6 烷基)、环烷基(如 C3-C6 环烷基)或芳香族烷基(如苯基 C1-C6 烷基);

优选地, 所述可逆阻断基团 R' 选自 N₃-C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、

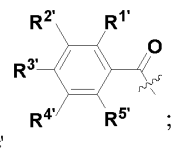


、NH₂;

更优选地, 所述可逆阻断基团 R' 选自 N₃-C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、



最优选地, 所述可逆阻断基团 R' 选自 N₃-CH₂-、CH₃-CH₂-S-S-CH₂-、



R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, 任意一个(如 R^{1'}或 R^{5'})为 , 其余各自独立地选自 H、叠氨基、硝基、氨基、磺基、羧基、脂肪族烷基(如 C1-C6 烷基)、环烷基(如 C3-C6 环烷基)、芳香族烷基(如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基(如 C1-C6 烷氧基),

优选地, R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, 任意一个(如 R^{1'}或 R^{5'})为 , 另一个(如 R^{3'})选自 H、C1-C6 烷氧基(如甲氧基), 其余为 H,

更优选地, R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, 任意一个(如 R^{1'}或 R^{5'})为 , 其余为 H,

最优选地, R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, R^{1'}为 , 其余为 H,

R^{b'}、R^{c'}中, 任意一个选自-N₃、-SS-C1-C6 烷基(如-SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu)、-ONH₂、-OCOR_X、-OCONHR_X, 另一个选自 H、脂肪族烷基(如 C1-C6 烷基, 例如 Me、Et、

iPr、tBu)、芳香族烷基(如苯基-C1-C6 烷基)、环烷基(如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基(如 C1-C6 烷基), 环烷基(如 C3-C6 环烷基)或芳香族烷基(如苯基 C1-C6 烷基),

优选地, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基(如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$), 另一个选自 C1-C6 烷基,

更优选地, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-Me$, 另一个为甲基。

在一些实施方案中, 所述的核苷酸类似物选自前述表 A、表 C 中的化合物。

在一些实施方案中, 所述核苷酸类似物携带额外的可检测标记, 且所述可检测标记与连接子(优选 L^b) 连接;

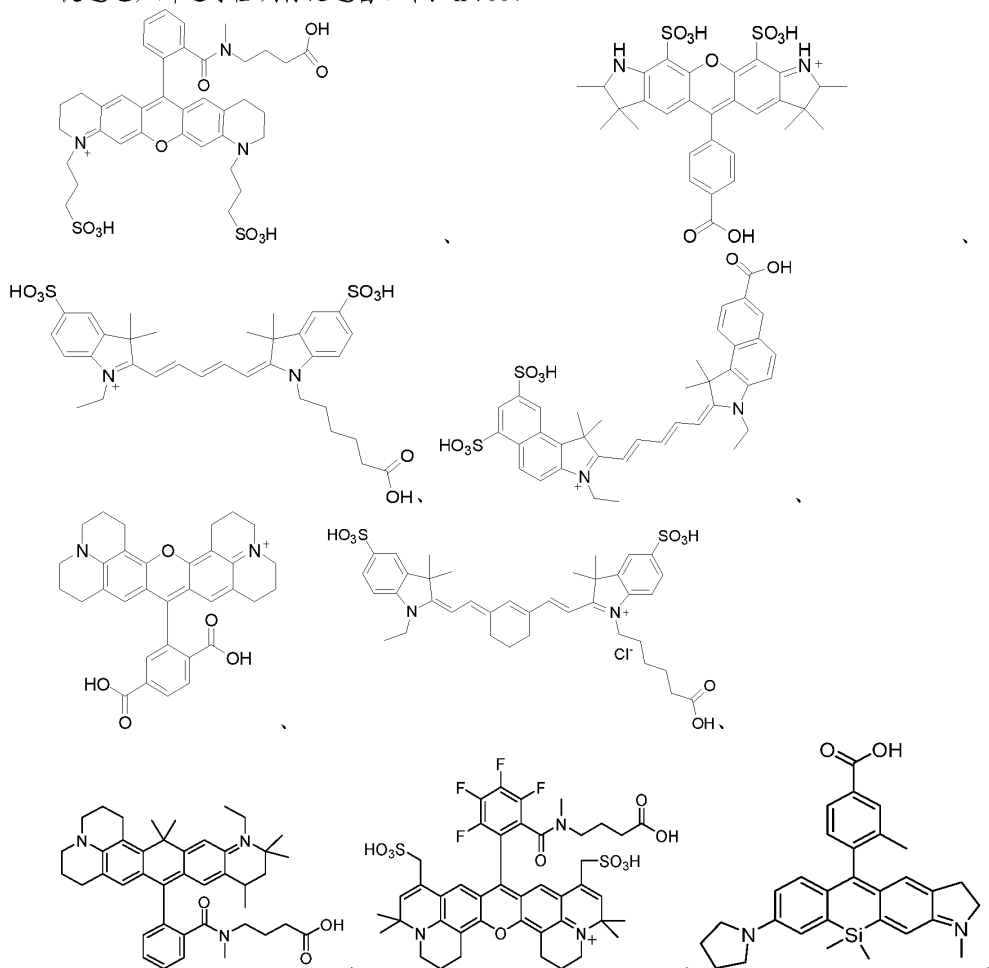
优选地, 所述核苷酸类似物携带的额外的可检测标记是通过亲和试剂(如抗体、适体、亲和物、结蛋白)引入的, 所述亲和试剂携带所述可检测标记, 且所述亲和试剂可以特异性识别并结合所述核苷酸类似物的表位; 优选地, 所述可检测标记与所述 L^b 的末端氨基连接;

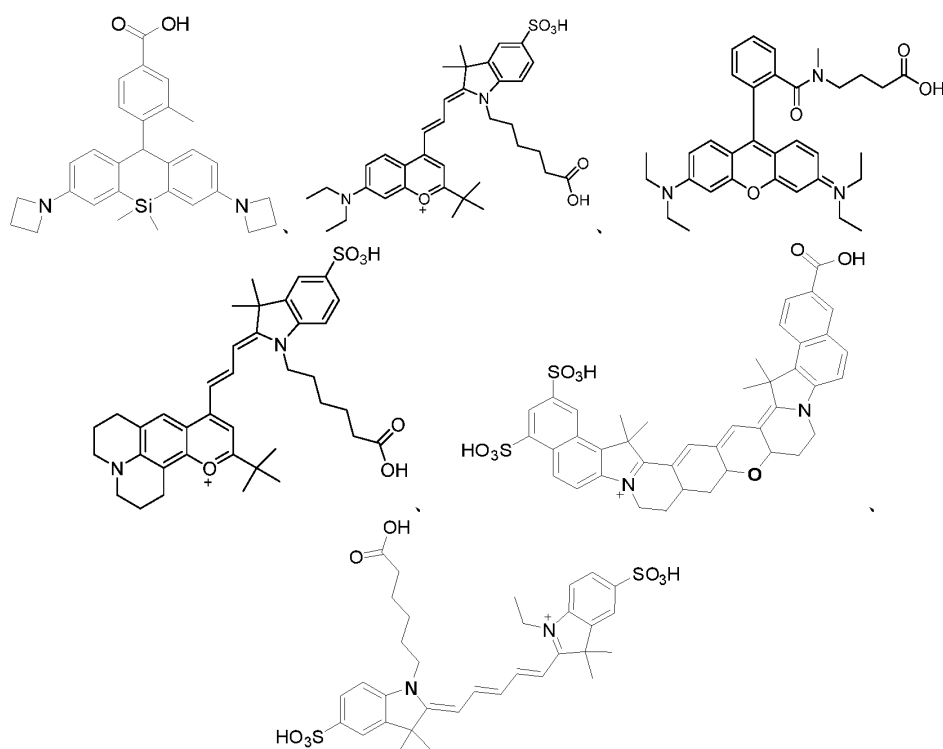
优选地, 所述可检测标记中的羧基与所述 L^b 的末端氨基通过形成酰胺键进行连接;

优选地, $base^I$ 不同, 所述核苷酸类似物携带的额外的可检测标记不同;

优选地, 所述可检测标记为荧光标记;

优选地, 所述可检测标记选自以下: iF700、





在一些实施方案中，所述核苷酸类似物选自前述表 B、表 D 中的化合物。

在本发明的第二十三方面，本发明提供了终止核酸合成的方法，其包括：将前述的核苷酸类似物掺入待终止的核酸分子中；

优选地，所述核苷酸类似物的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述核苷酸类似物掺入待终止的核酸分子中；

优选地，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述核苷酸类似物掺入待终止的核酸分子的 3' 端。

在本发明的第二十四方面，本发明提供了制备在测序反应中与目标单链多核苷酸互补的生长的多核苷酸的方法，其包括将前述的核苷酸类似物掺入所述生长的互补多核苷酸中，其中，所述核苷酸类似物的掺入防止了任何后续的核苷酸引入所述生长的互补多核苷酸中；

优选地，所述核苷酸类似物的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述核苷酸类似物掺入所述生长的互补多核苷酸中；

优选地，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述核苷酸类似物掺入所述生长的互补多核苷酸的 3' 端。

在本发明的第二十五方面，本发明提供了核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，

所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是前述的核苷酸类似物；

或者，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是前述的核苷酸类似物，且所述生长的核酸链中预先掺入至少一个与目标单链多核苷酸互补的核苷酸，预先掺入的至少一个与目标单链多核苷酸互补的核苷酸是已被除去可逆阻断基团和任选的可检测标记的前述的核苷酸类似物。

在本发明的第二十六方面，本发明提供了测定目标单链多核苷酸的序列的方法，其包括：

1) 监测生长的核酸链中与目标单链多核苷酸互补的核苷酸的掺入，其中，掺入的至少

一个互补核苷酸是前述的核苷酸类似物, 以及,

2) 确定掺入的核苷酸的类型;

优选地, 在引入下一个互补核苷酸之前, 将所述可逆阻断基团和任选的可检测标记除去;

优选地, 所述可逆阻断基团和所述可检测标记被同时除去;

优选地, 所述可逆阻断基团和所述可检测标记被先后除去; 例如, 在所述可检测标记被除去之后, 所述可逆阻断基团被除去, 或者, 在所述可逆阻断基团被除去之后, 所述可检测标记被除去。

在一些实施方案中, 所述方法包括以下步骤:

(a) 提供多种不同的核苷酸, 其中至少一种核苷酸是前述的核苷酸类似物, 任选地其余的核苷酸是前述的核苷酸类似物;

(b) 将所述多种不同的核苷酸掺入目标单链多核苷酸的互补序列中, 其中, 所述多种不同的核苷酸在检测时可以相互区分;

(c) 检测 (b) 的核苷酸, 从而确定掺入的核苷酸的类型;

(d) 除去 (b) 的核苷酸中的可逆阻断基团和任选的其携带的可检测标记; 和

(e) 任选地重复步骤 (a) - (d) 一次或多次;

从而确定所述目标单链多核苷酸的序列。

在一些实施方案中, 所述方法包括以下步骤:

(1) 提供第一核苷酸、第二核苷酸、第三核苷酸和第四核苷酸, 四种核苷酸中的至少一种是前述的核苷酸类似物, 任选地其余的核苷酸是前述的核苷酸类似物;

(2) 将所述四种核苷酸与目标单链多核苷酸进行接触; 除去未掺入生长的核酸链中的所述核苷酸; 检测掺入生长的核酸链中的所述核苷酸; 除去掺入生长的核酸链中的所述核苷酸中的所述可逆阻断基团和任选的其携带的所述可检测标记;

任选地, 还包括 (3): 重复 (1) - (2) 一次或多次。

在一些实施方案中, 所述方法包括以下步骤:

(a) 提供包含双链体、包含至少一种前述的核苷酸类似物的核苷酸、聚合酶和切除试剂的混合物; 所述双链体包含生长的核酸链以及待测序的核酸链;

(b) 进行包含以下步骤 (i)、(ii) 和 (iii) 的反应, 任选地, 重复一次或多次:

步骤 (i): 使用聚合酶, 使所述核苷酸类似物掺入生长的核酸链, 形成包含可逆阻断基团和任选的可检测标记的核酸中间体;

步骤 (ii): 对所述核酸中间体进行检测;

步骤 (iii): 使用切除试剂将所述核酸中间体所包含的可逆阻断基团和任选的可检测标记切除;

优选地, 对所述可逆阻断基团的切除和对所述可检测标记的切除同时进行, 或者, 对所述可逆阻断基团的切除和对所述可检测标记的切除分步进行 (例如, 先切除所述可逆阻断基团, 或者先切除所述可检测标记);

优选地, 对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是同样的试剂;

优选地, 对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是不同的试剂。

在一些实施方案中, 所述双链体连接于支持物上;

优选地, 所述生长的核酸链为引物;

优选地, 所述引物通过退火至待测序的核酸链上, 形成所述双链体;

优选地, 所述双链体、所述核苷酸类似物、以及所述聚合酶一起形成含有溶液相和固相的反应体系;

优选地, 在允许聚合酶进行核苷酸聚合反应的条件下, 使用聚合酶, 使所述核苷酸类似物掺入生长的核酸链, 形成包含可逆阻断基团和任选的可检测标记的核酸中间体;

优选地, 所述聚合酶选自 KOD 聚合酶或其突变体 (例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391);

优选地, 在任意一个检测所述核酸中间体的步骤前, 移除前一步骤的反应体系的溶液相, 保留连接于支持物上的双链体;

优选地, 所述切除试剂与所述双链体或所述生长的核酸链在含有溶液相和固相的反应

体系中接触;

优选地,所述切除试剂能够切除掺入生长的核酸链的核苷酸类似物中的可逆阻断基团和任选的其携带的可检测标记,并且不会影响双链体骨架上的磷酸二酯键;

优选地,在任意一个切除所述核酸中间体所包含的可逆阻断基团和任选的可检测标记的步骤后,移除这一步骤反应体系的溶液相;

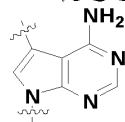
优选地,在任意一个包含移除操作的步骤之后,进行洗涤操作;

优选地,步骤(ii)之后,进一步包括:根据步骤(ii)检测得到的信号,确定步骤(i)中掺入生长的核酸链的核苷酸类似物的类型,并基于碱基互补配对原则,确定待测序的核酸链中相应位置处的核苷酸类型。

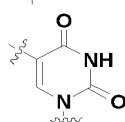
在本发明的第二十七方面,本发明提供了试剂盒,其包含至少一个前述的核苷酸类似物;

优选地,所述试剂盒包含第一、第二、第三和第四化合物,所述第一、第二、第三和第四化合物各自独立地为前述的核苷酸类似物;

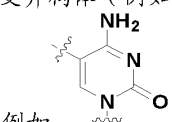
优选地,所述第一化合物中,base¹选自腺嘌呤、7-脱氮腺嘌呤或其互变异构体(例如



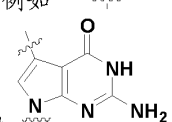
);所述第二化合物中,base¹选自胸腺嘧啶、尿嘧啶或其互变异构体(例如



);所述第三化合物中,base¹选自胞嘧啶或其互变异构体(例如



所述第四化合物中,base¹选自鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体(例如



优选地,所述第一、第二、第三和第四化合物包含的base¹互不相同;

优选地,所述第一、第二、第三和第四化合物携带的额外的可检测标记互不相同。

在一些实施方案中,所述试剂盒还包含:用于预处理核酸分子的试剂;用于连接待测序的核酸分子的支持物;用于将待测序的核酸分子与支持物连接(例如,共价或非共价连接)的试剂;用于起始核苷酸聚合反应的引物;用于进行核苷酸聚合反应的聚合酶;一种或多种缓冲溶液;一种或多种洗涤溶液;或其任何组合。

在本发明的第二十八方面,本发明提供了前述的核苷酸类似物或者前述的试剂盒用于测定目标单链多核苷酸的序列的用途。

附图说明

图1表示本发明实施例的碱基带有可切断(荧光)标记核苷酸类似物;

图2表示本发明实施例的3'-OH可逆阻断基团带有(荧光)标记核苷酸类似物。

具体实施方式

下面通过具体实施例详细描述本发明的实施方式,但是无论如何它们不能解释为对本发明的限制。

除非特殊说明,上述基团和取代基具有药物化学领域的普通含义。

在本说明书的各部分,本发明公开化合物的取代基按照基团种类或范围公开。特别指出,本发明包括这些基团种类和范围的各个成员的每一个独立的次级组合。例如,术语“C1-C6烷基”特别指独立公开的甲基、乙基、C3烷基、C4烷基、C5烷基和C6烷基。

另外,需要说明的是,除非以其他方式明确指出,在本文中通篇采用的描述方式“各...独立地为/选自”和“...各自独立地为/选自”可以互换,均应做广义理解,其既可以是指在不同基团中,相同或不同的符号之间所表达的具体选项之间互相不影响,也可以表示在相同的基团中,相同或不同的符号之间所表达的具体选项之间互相不影响。

术语“脂肪族烷基”指的是任意的含有1-20个碳原子的直链或支链饱和基团,例如,C₁-C₁₂烷基,优选C₁-C₆烷基。

术语“C₁-C₆烷基”指的是任意的含有 1-6 个碳原子的直链或支链饱和基团，例如甲基 (Me)、乙基 (Et)、正丙基、异丙基 (iPr)、正丁基、异丁基、叔丁基 (tBu)、仲丁基、正戊基、叔戊基、正己基等。

术语“烷氧基”指的是任意上述烷基 (例如 C₁-C₆烷基等)，其通过氧原子(-O-)连接到分子的其余部分。

术语“环烷基”是指具有饱和环的 3-10 元单环系统的烃，例如，C₃-C₈ 环烷基，优选 C₃-C₆ 环烷基。

术语“C₃-C₆ 环烷基”是指具有饱和环的 3-6 元单环系统的烃，C₃-C₆ 环烷基可以为环丙基、环丁基、环戊基、环己基等。

术语“芳香族烷基”指的是芳基烷基或杂芳基烷基，其中烷基如上文所定义。

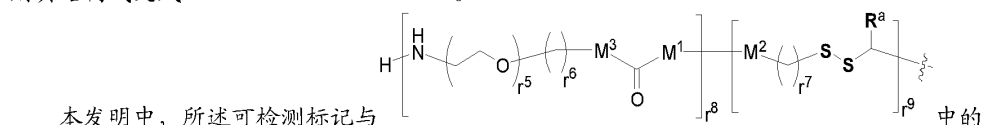
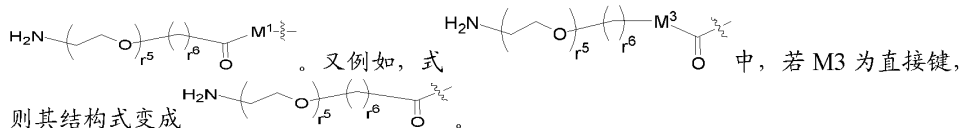
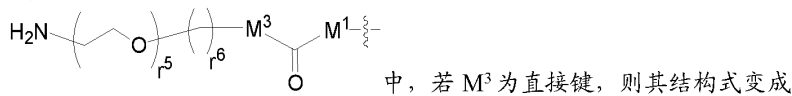
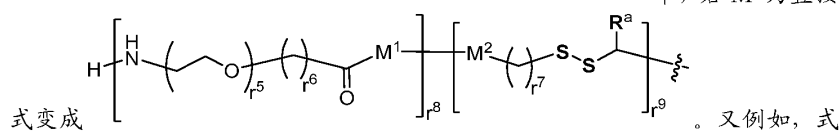
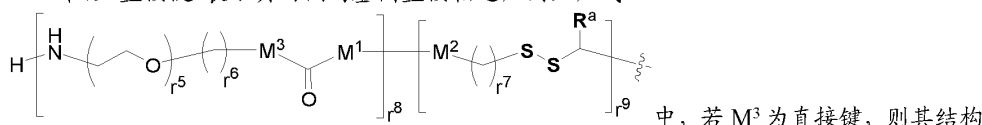
术语“杂芳基”是指芳族的杂环，通常为具有 1 至 3 个选自 N、O 或 S 的杂原子的 5-、6-、7-、8-元的杂环；杂芳基环可以任选地进一步稠合或连接于芳族和非芳族的碳环和杂环。所述杂芳基的非限制性的实例为例如吡啶基、吡嗪基、嘧啶基、哒嗪基、咪唑基、咪唑基、噻唑基、异噻唑基、噻吩基、吡咯基、苯基-吡咯基、呋喃基、苯基-呋喃基、噁唑基、异噁唑基、吡唑基、噻吩基、苯并呋喃基、苯并噻吩基、苯并 1,3-二氧戊环 (苯并二噁茂)、异二氢咪唑基、苯并咪唑基、吡唑基、喹啉基、异喹啉基、1,2,3-三唑基、1-苯基-1,2,3-三唑基、2,3-二氢咪唑基、2,3-二氢苯并呋喃基、2,3-二氢苯并噻吩基、苯并吡喃基、2,3-二氢苯并噻吩基、2,3-二氢喹啉基等。

术语“芳基”指的是具有 6-14 个碳原子的碳环芳族体系的基团，例如 C₆-C₁₀ 芳基，优选苯基。

从所有上述描述中，对本领域技术人员显而易见的是，其名称是复合名称的任意基团，例如“苯基 C₁-C₆ 烷基”，应该指的是常规地从其衍生的部分例如从被苯基取代的 C₁-C₆ 烷基来构建，其中 C₁-C₆ 烷基如上文所定义。

如本文所使用，术语“式 I、式 I-1、式 I-2、式 II、式 II-1、式 II-2、式 III 所示的化合物的盐”的例子是由形成阴离子的有机酸形成的有机酸加合盐，包括但不限于甲酸盐、乙酸盐、丙酸盐、苯甲酸盐、马来酸盐、富马酸盐、琥珀酸盐、酒石酸盐、柠檬酸盐、抗坏血酸盐、α-酮戊二酸盐、α-甘油磷酸盐、烷基磺酸盐或芳基磺酸盐；优选地，所述烷基磺酸盐为甲基磺酸盐或乙基磺酸盐；所述芳基磺酸盐为苯磺酸盐或对甲苯磺酸盐。也可形成合适的无机盐，包括但不限于盐酸盐、氢溴酸盐、氢碘酸盐、硝酸盐、碳酸氢盐和碳酸盐、硫酸盐或磷酸盐等。

术语“直接键”表示其两侧的基团直接相连，例如，式



本发明中，所述可检测标记与

末端氨基连接,其包含以下两种情况:1)当 r^8 不为0时,“末端氨基”指的是该结构式最末

尾的氨基;2)当 r^8 为0时,该结构式变为

$$\text{H}-\left[\text{M}^2\left(\text{CH}_2\right)_{r^7}\text{S}-\text{C}\left(\text{R}^a\right)\right]_{r^9}$$

M^2 可以为NH,此时,可以理解的是,“末端氨基”指的是该变更后的结构式最末尾的氨基。

本发明中,所述可检测标记与

$$\text{H}_2\text{N}-\left(\text{CH}_2\right)_{r^5}-\text{O}-\left(\text{CH}_2\right)_{r^6}-\text{M}^3-\text{C}(=\text{O})-\text{M}^1$$

中的末端氨基连接,

其中,“末端氨基”指的是该结构式最末尾的氨基。

本发明中,所述可检测标记与与所述 L^b 的末端氨基连接,且所述 L^b 例如为

$\text{H}_2\text{N}-\left(\text{CH}_2\right)_{r^5}-\text{O}-\left(\text{CH}_2\right)_{r^6}-\text{M}^3-\text{C}(=\text{O})-\text{CH}_2$ 时,“末端氨基”指的是该结构式最末尾的氨基。

在本发明的方法中,只要是由生长的核酸链和待测序的核酸链这两条链组成的物质,均称其为“双链体”,与生长的核酸链或待测序的核酸链的链长无关,待测序的核酸链可以比生长的核酸链的链长更长。

在本发明的方法中,待测序的核酸分子可以是任何目的核酸分子。在某些优选的实施方案中,所述待测序的核酸分子包含脱氧核糖核苷酸、核糖核苷酸、经修饰的脱氧核糖核苷酸、经修饰的核糖核苷酸、或其任何组合。在本发明的方法中,待测序的核酸分子不受其类型的限制。在某些优选的实施方案中,所述待测序的核酸分子为DNA或RNA。在某些优选的实施方案中,所述待测序的核酸分子可以为基因组DNA,线粒体DNA,叶绿体DNA,mRNA,cDNA,miRNA,或siRNA。在某些优选的实施方案中,所述待测序的核酸分子为线性的或者环状的。在某些优选的实施方案中,所述待测序的核酸分子为双链的或者单链的。例如,所述待测序的核酸分子可以为单链DNA(ssDNA),双链DNA(dsDNA),单链RNA(ssRNA),双链RNA(dsRNA),或者DNA和RNA的杂合体。在某些优选的实施方案中,所述待测序的核酸分子为单链DNA。在某些优选的实施方案中,所述待测序的核酸分子为双链DNA。

在本发明的方法中,待测序的核酸分子不受其来源的限制。在某些优选的实施方案中,待测序的核酸分子可以获自任何来源,例如,任何细胞、组织或生物体(例如,病毒,细菌,真菌,植物和动物)。在某些优选的实施方案中,待测序的核酸分子源自哺乳动物(例如,人、非人灵长类动物、啮齿类动物或犬科动物)、植物、鸟类、爬行类、鱼类、真菌、细菌或病毒。

从细胞、组织或生物体中提取或获得核酸分子的方法是本领域技术人员公知的。合适的方法包括但不限于乙醇沉淀法,氯仿抽提法等。关于此类方法的详细描述可参见例如,J.Sambrook等人,分子克隆:实验室手册,第2版,冷泉港实验室出版社,1989,以及F.M.Ausubel等人,精编分子生物学实验指南,第3版,John Wiley&Sons,Inc.,1995。另外,还可使用各种商业化的试剂盒来从各种来源(例如细胞、组织或生物体)提取核酸分子。

在本发明的方法中,待测序的核酸分子不受其长度的限制。在某些优选的实施方案中,待测序的核酸分子的长度可以为至少10bp,至少20bp,至少30bp,至少40bp,至少50bp,至少100bp,至少200bp,至少300bp,至少400bp,至少500bp,至少1000bp,或者至少2000bp。在某些优选的实施方案中,待测序的核酸分子的长度可以为10-20bp,20-30bp,30-40bp,40-50bp,50-100bp,100-200bp,200-300bp,300-400bp,400-500bp,500-1000bp,1000-2000bp,或者超过2000bp。在某些优选的实施方案中,待测序的核酸分子可具有10-1000bp的长度,以利于进行高通量测序。

在本发明的制备多核苷酸的方法或测序方法中,可使用合适的聚合酶来进行核苷酸聚合反应。在一些示例性实施方案中,所述聚合酶能够以DNA为模板合成新的DNA链(例如DNA聚合酶)。在一些示例性实施方案中,所述聚合酶能够以RNA为模板合成新的DNA链(例如反转录酶)。在一些示例性实施方案中,所述聚合酶能够以DNA或RNA为模板合成新的RNA链(例如RNA聚合酶)。因此,在某些优选的实施方案中,所述聚合酶选自DNA聚合酶,RNA聚合酶,和反转录酶。可根据实际需要,选择合适的聚合酶来进行核苷酸聚合反应。在某些优选的实施方案中,所述聚合反应为聚合酶链式反应(PCR)。在某些优选的实施方案中,所述聚合反应为反转录反应。

在本发明的方法中，可以使用 KOD 聚合酶或其突变体进行核苷酸聚合反应。KOD 聚合酶或其突变体(例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391)对本发明的修饰的核苷或核苷酸具有可接受的聚合效率。KOD POL391 和 KOD POL171 对本发明的修饰的核苷酸的具有可接受的聚合效率。在某些实施方案中，KOD POL391 或 KOD POL171 对本发明的修饰的核苷酸的聚合效率在 70% 以上，例如 70%-80%、80%-90% 或 90%-100%。

在本发明的制备多核苷酸的方法或测序方法中，核苷酸的聚合反应在适宜的条件下进行。适宜的聚合条件包括溶液相的组成以及各成分的浓度、溶液相的 pH、聚合温度等。在适宜的条件下进行聚合，有利于获得可接受的、甚至高的聚合效率。

在本发明的一些实施方案中，式 I、式 II、式 III、式 B 所示化合物中，脱氧核糖 3' 位置处的羟基(-OH)受保护(被 R' 所保护)，因此，它们能够终止聚合酶(例如 DNA 聚合酶)的聚合作用。例如，当式 I、式 II、式 III、式 B 所示化合物被引入生长的核酸链的 3' 端时，由于该化合物的脱氧核糖的 3' 位置处不存在游离的羟基(-OH)，聚合酶将无法继续进行下一轮的聚合反应，从而聚合反应将被终止。在这种情况下，在每一轮的聚合反应中，将有且只有一个碱基被掺入生长的核酸链。

此外，所述式 I、式 II、式 III、式 B 所示化合物的脱氧核糖 3' 位置处羟基(-OH)的保护基团(R')能够被去除，并转变为游离的羟基(-OH)。随后，可使用聚合酶和式 I、式 II、式 III、式 B 所示化合物对生长的核酸链进行下一轮的聚合反应，并再次引入一个碱基。

因此，所述式 I、式 II、式 III、式 B 所示化合物的脱氧核糖 3' 位置处羟基(-OH)可以被可逆阻断：当式 I、式 II、式 III、式 B 所示化合物被掺入生长的核酸链的 3' 端时，它们将终止聚合酶继续进行聚合作用，终止生长的核酸链的进一步延伸；并且，在式 I、式 II、式 III、式 B 所示化合物所包含的阻断基团被去除后，聚合酶将能够继续对生长的核酸链进行聚合作用，继续延伸核酸链。

本文描述的某些实施方案涉及常规可检测标记的使用。可通过任何适合的方法进行检测，包括荧光光谱学或其他光学手段。优选的标记为荧光标记即荧光团，该荧光团在吸收能量后发出限定波长的辐射。已知许多种适合的荧光标记。例如，Welch 等人(Chem.Eur.J.5(3):951-960,1999)公开了丹酰基-功能化的荧光部分，其可在本发明中使用。Zhu 等人(Cytometry28:206-211,1997)描述了荧光标记 Cy3 和 Cy5 的使用，其也可以在本发明中使用。Prober 等人(Science238:336-341,1987)、Connell 等人(BioTechniques5(4):342-384,1987)、Ansorge 等人(Nucl.AcidsRes. 15(11):4593-4602,1987)和 Smith 等人(Nature321:674,1986)也公开了适合使用的标记。其他可商业购得的荧光标记包括但不限于 iF700、荧光素、若丹明(包括 TMR、德克萨斯红和 Rox)、alexa、氟硼荧、吖啶、香豆素、芘、苯并蒽和花青苷。特别说明的是，iF700 是本领域常规使用的荧光标记，例如参见 US20180223358A1 的 Table 7，该荧光标记可以市购获得。

本申请中也可以使用多重标记，例如双荧光团 FRET 盒(Tet. Let. 46:8867-8871,2000)、也可以使用多荧光体树枝状系统(J.Am.Chem.Soc. 123:8101-8108,2001)。虽然优选荧光标记，对于本领域的普通技术人员来说其他形式的可检测标记也明显适用。例如微颗粒，包括量子点(Empodocles 等人, Nature 399:126-130,1999)、金纳米颗粒(Reichert 等人, Anal.Chem. 72:6025-6029,2000)和微珠(Lacoste 等人, Proc.Natl.Acad.Sci USA 97(17):9461-9466,2000)也都可以使用。

本申请也可以使用多组分标记。多组分标记是依赖于与用于检测的另外化合物的相互作用的标记。在生物学中最常用的多组分标记是生物素-链霉亲和素系统。生物素用作与核苷酸或修饰的核苷酸相连接的标记。然后单独加入链霉亲和素使检测发生。可以使用其他多组分系统。例如，二硝基苯酚具有可商业购得的可用于检测的荧光抗体。

在本文描述的某些实施方案中，可以通过亲和试剂(如抗体、适体、亲和物 Affimer、结蛋白 Knottin)的引入使得修饰的核苷酸或核苷分子携带上文描述的可检测标记，所述亲和试剂可以特异性识别并结合所述修饰的核苷酸或核苷分子的表位，具体原理详见 WO2018129214A1。WO2018129214A1 中的全部相关内容引入本申请中。

在本文描述的另外一些实施方案中，可以将修饰的核苷酸或核苷分子与上文描述的可检测标记相连接。在某些这类实施方案中，所用的连接基团可裂解。使用可裂解的连接基团确保了在需要时所述标记能够在检测后被除去，这避免了与随后并入的任何标记的核苷

酸或核苷的任何干扰信号。

可通过任何适合的方法裂解所述连接基团,包括暴露于酸、碱、亲核试剂、亲电试剂、自由基、金属、还原剂或氧化剂、光、温度、酶等。还可以使用用于断裂碱基处的保护基的相同催化剂裂解本文中讨论的连接基团。如

Greene&Wuts,Protective Groups in Organic Synthesis(有机合成中的保护基),John Wiley&Sons 中所公开的,合适的连接基团可修改自标准的化学保护基。Guillier 等人(Chem.Rev.100:2092-2157,2000)中还公开了用于固相合成的合适的可裂解的连接基团。

使用术语“可裂解的连接基团”并非意味着需要除去整个连接基团,例如,从核苷酸或修饰的核苷酸中除去。当可检测标记与核苷酸或修饰的核苷酸相连接时,核苷裂解位点可位于连接基团上的位置,该位置能够确保在裂解后一部分的连接基团仍与所述核苷酸或修饰的核苷酸保持连接。

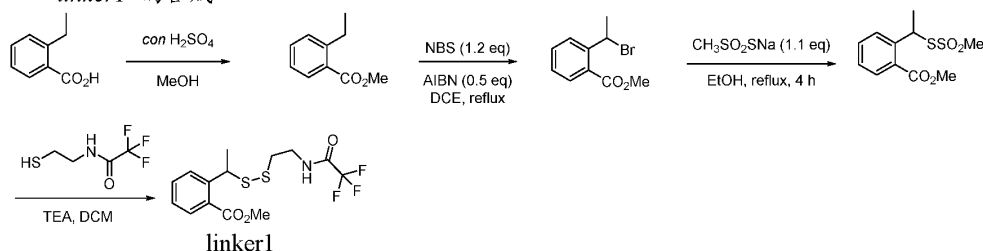
本领域技术人员可知,本申请中的核苷酸在 Sanger 法测序、第二代高通量测序(NGS 测序)和第三代测序(单分子测序)中均具有效用,因为通过使用本文描述的可逆阻断核苷酸类似物可实现逐个检测测序中掺入的核苷酸类型。

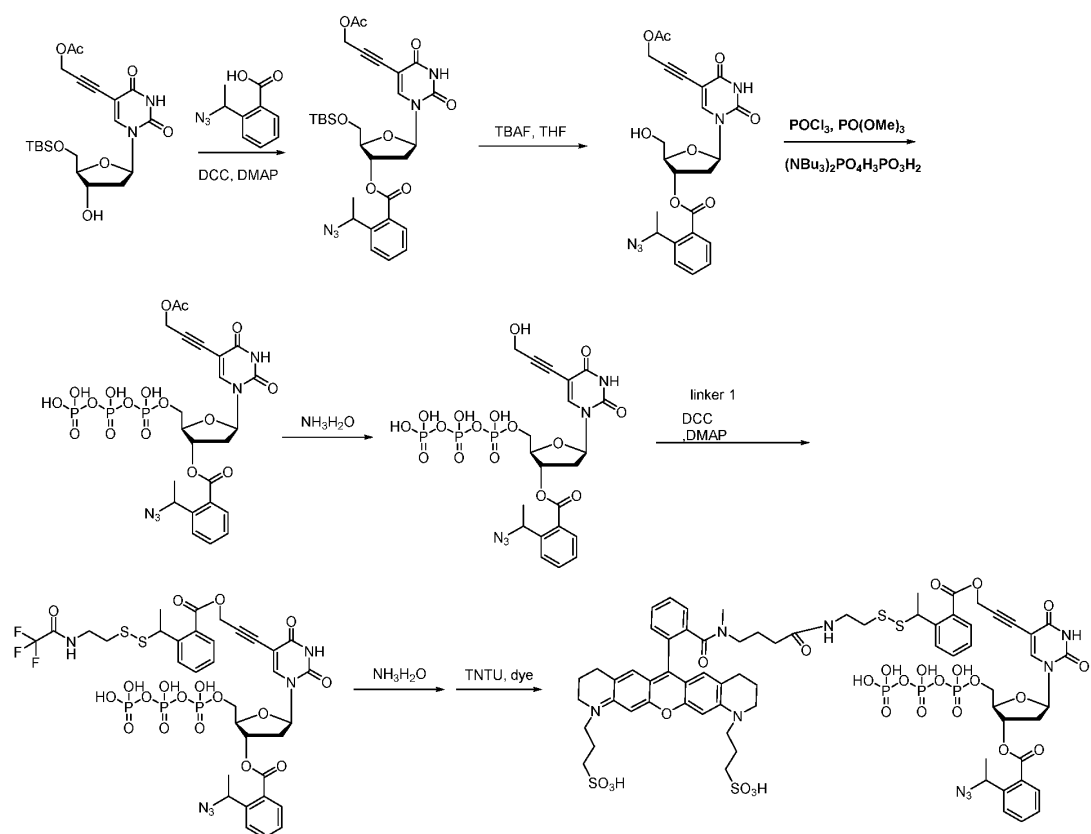
下面结合具体实施例对本发明进行进一步的解释说明。需要说明的是,以下各实施例仅列举了四种碱基中的一种碱基的核苷酸类似物的制备方法,本领域技术人员可以参照该方法,制备合成剩余三种碱基的核苷酸类似物。而且,以下各实施例化合物在最后一步与染料相连接时,可以与任意已知的染料进行连接,以下实施例各列举了一种染料,仅为示例,并非对染料选择范围的限制。另外,除非特别说明,否则各原料均可以市购获得。

一、化合物制备实施例

实施例 1

linker1 的合成

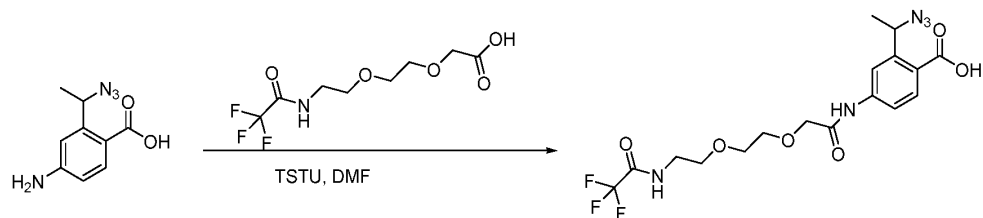




核苷与酸 DCC 缩合后, TBAF 脱去保护基, 再利用三磷酸化的三磷酸。氨解后与 linker 1 偶联, 在脱去三氟乙酰基, 最后与荧光染料偶联得到产品。

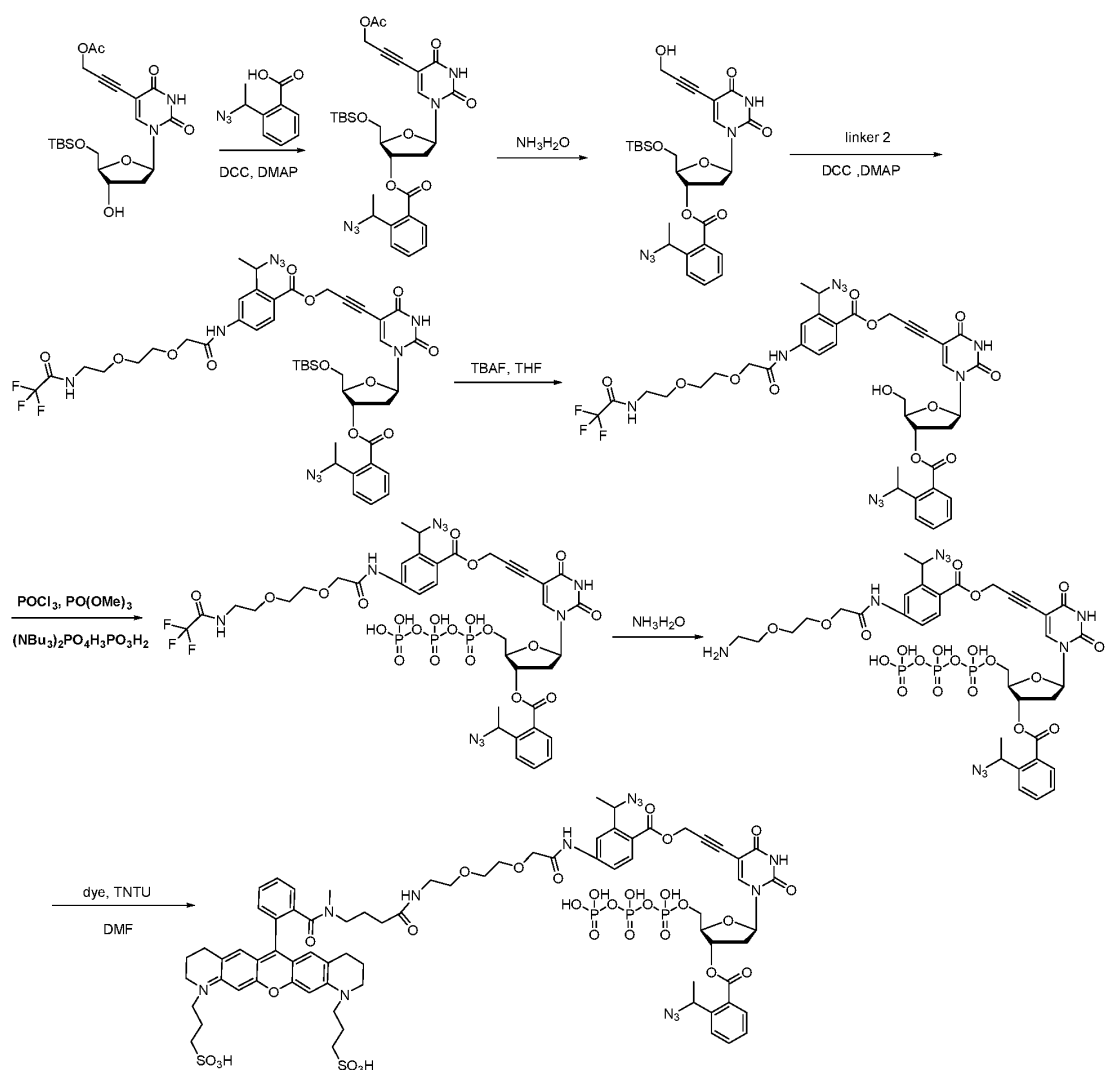
实施例 2

linker 2 的合成



linker 2

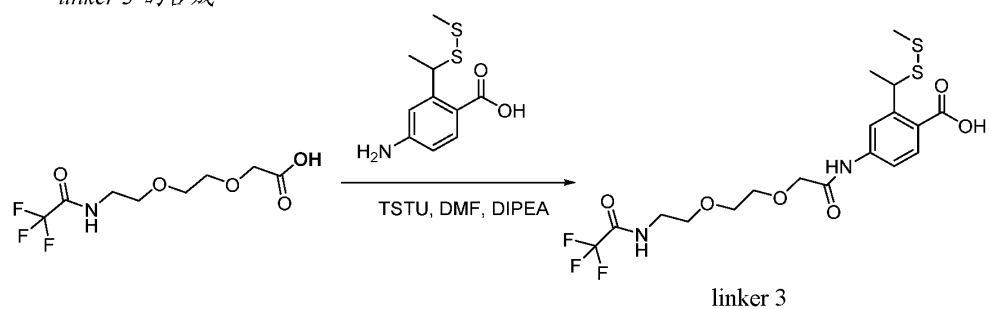
对氨基叠氮苯甲酸与三氟乙酰基保护的 PEG linker 缩合得到 linker 2。



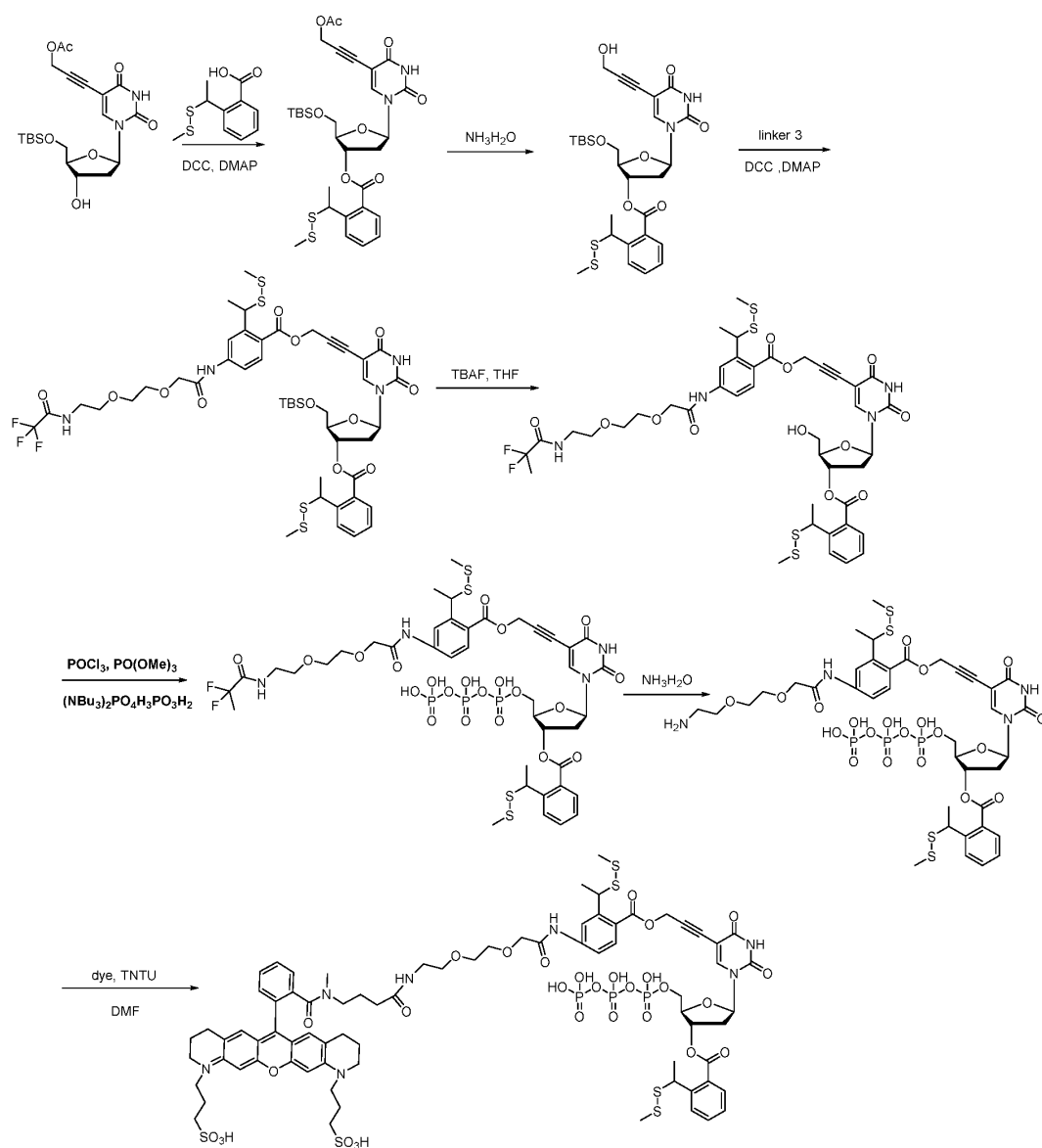
核苷与酸 DCC 缩合后, 氨基脱去乙酰基, 然后再 DCC 缩合。利用 TBAF 脱去保护基, 再利用三磷酸化的三磷酸。氨解脱去三氟乙酰基, 最后与荧光染料偶联得到产品。

实施例 3

linker 3 的合成



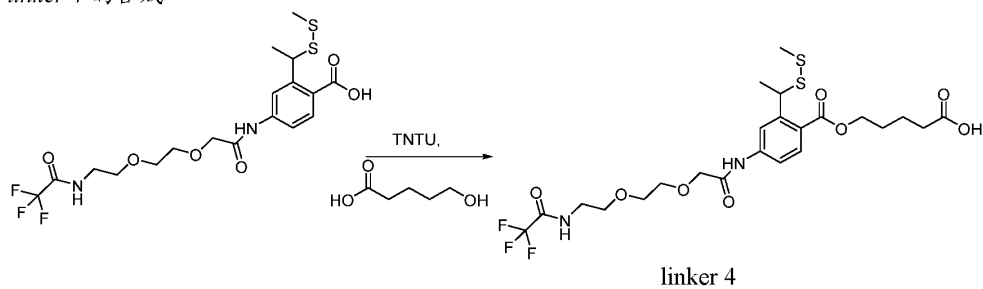
合成路线与鉴定结果参见实施例 7-1。

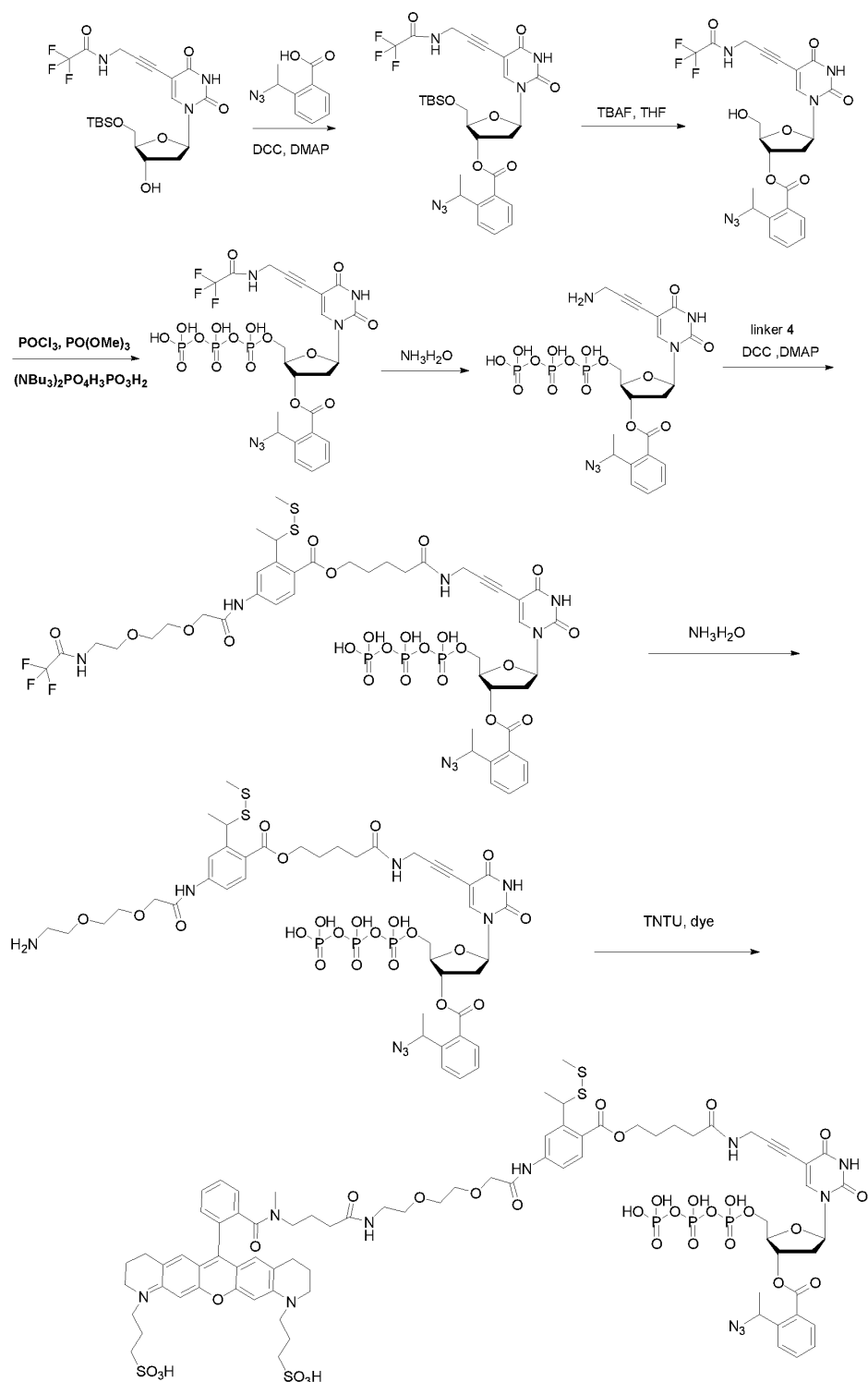


核苷与酸 DCC 缩合后，氨基脱去乙酰基，然后再 DCC 缩合。利用 TBAF 脱去保护基，再利用三磷酸化的三磷酸。氨解脱去三氟乙酰基，最后与荧光染料偶联得到产品。

实施例 4

linker 4 的合成

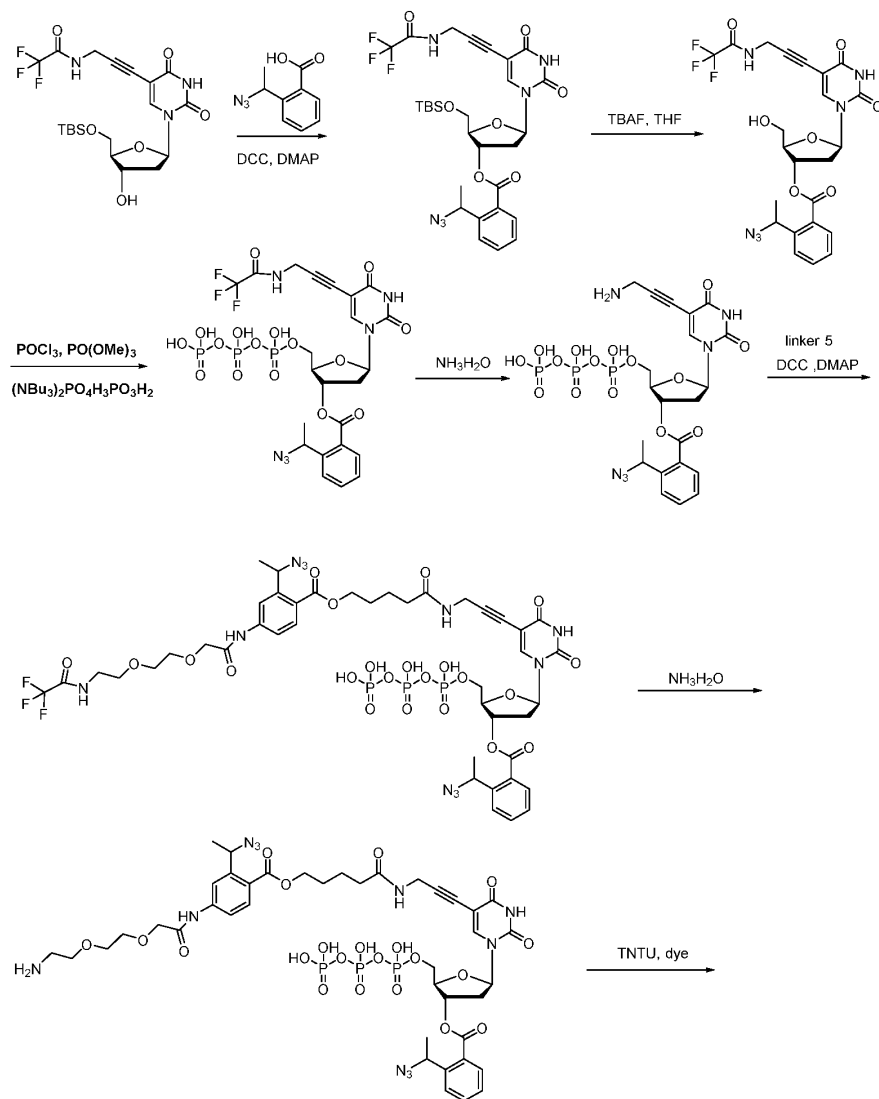
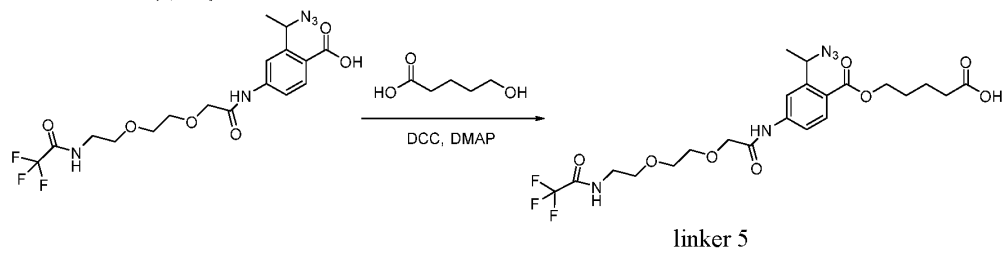


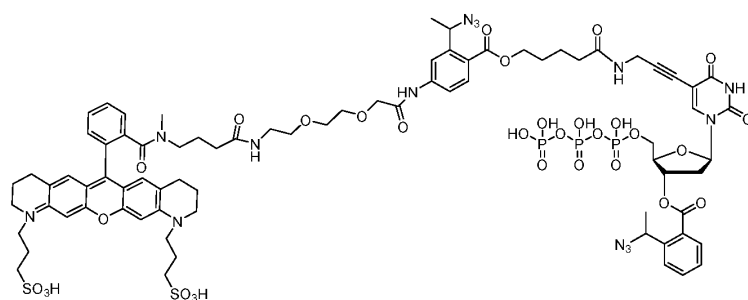


核苷与酸 DCC 缩合后, 利用 TBAF 脱去保护基, 再利用三磷酸化的三磷酸。氨解脱去三氟乙酰基, 与 linker 4 偶联, 再脱去 linker 中三氟乙酰基, 最后与荧光染料偶联得到产品。

实施例 5

linker 5 的合成



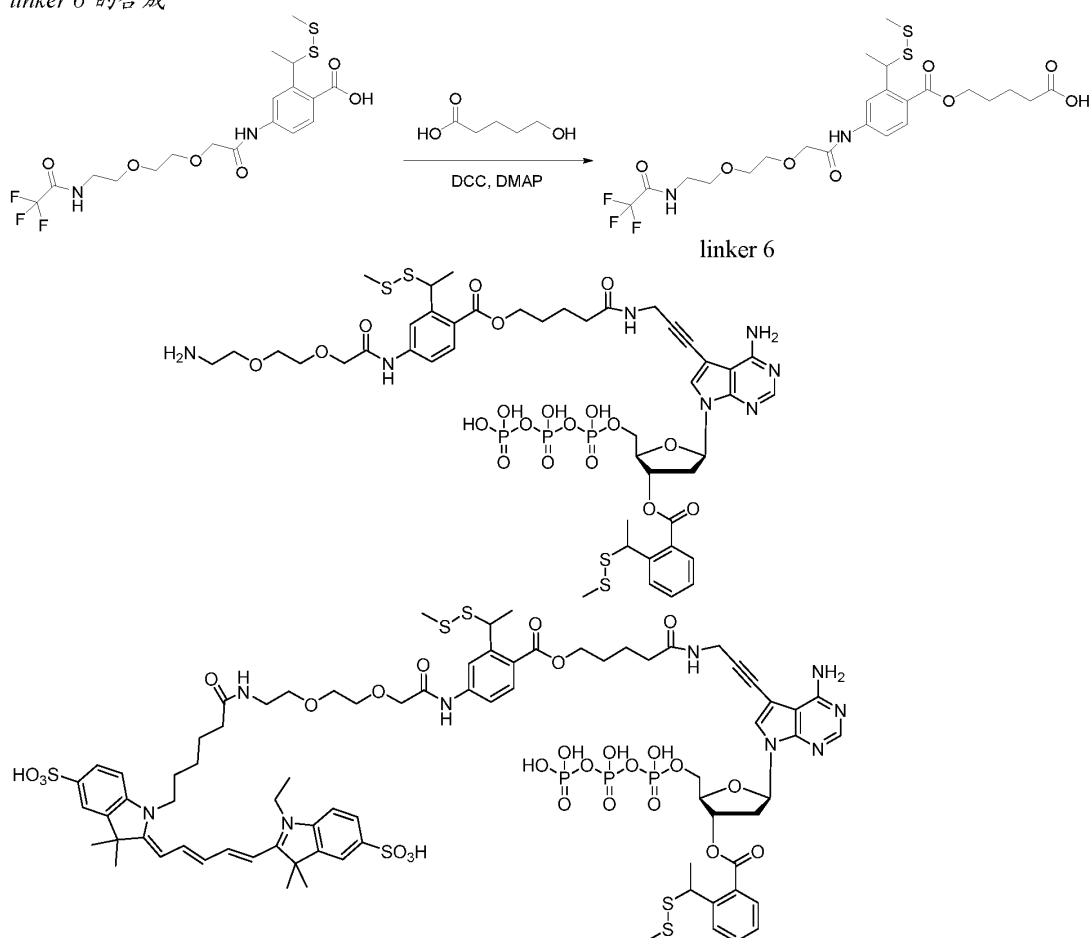


核

昔与酸 DCC 缩合后，利用 TBAF 脱去保护基，再利用三磷酸化的三磷酸。氨解脱去三氟乙酰基，与 linker 5 偶联，再脱去 linker 中三氟乙酰基，最后与荧光染料偶联得到产品。

实施例 6-1

linker 6 的合成



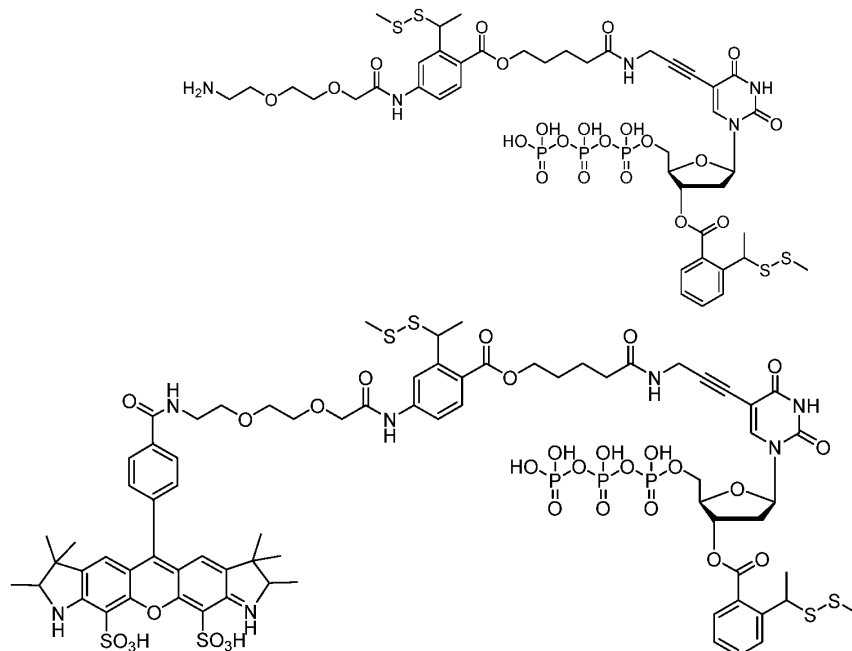
合成方法与实施例 7-2 类似。15mg 固体，HPLC 纯度 > 99%。MS[ES(-)], m/z 1861.4. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.27 (s, 1H), 8.63 - 8.58 (m, 1H), 8.36 (t, $J = 13.1$ Hz, 2H), 8.09 (d, $J = 6.3$ Hz, 1H), 8.00 (s, 1H), 7.92 - 7.88 (m, 2H), 7.84 (d, $J = 4.5$ Hz, 1H), 7.81 (s, 2H), 7.78 (d, $J = 8.6$ Hz, 1H), 7.71 - 7.67 (m, 1H), 7.65 - 7.62 (m, 4H), 7.49 - 7.43 (m, 1H), 7.31 (dd, $J = 8.3, 3.1$ Hz, 2H), 6.67 - 6.54 (m, 2H), 6.31 (dd, $J = 13.8, 5.3$ Hz, 2H), 5.68 (s, 1H), 5.30 (q, $J = 6.9$ Hz, 1H), 5.26 - 5.17 (m, 1H), 4.34 (dt, $J = 22.7, 3.7$ Hz, 1H), 4.22 (t, $J = 4.8$ Hz, 2H), 4.13 (s, 2H), 4.11 (d, $J = 5.0$ Hz, 2H), 4.08 - 4.00 (m, 2H), 3.67 - 3.65 (m, 1H), 3.59 - 3.56 (m, 2H), 3.22 - 3.15 (m, 2H), 2.21 (t, $J = 6.6$ Hz, 2H), 2.11 (d, $J = 2.3$ Hz, 4H), 2.09 - 2.04 (m, 4H), 1.66 (s, 3H),

1.60 (d, $J = 6.9$ Hz, 3H), 1.52 (dd, $J = 14.9, 7.4$ Hz, 2H), 1.37 – 1.29 (m, 2H), 1.26 (t, $J = 7.1$ Hz, 4H).

实施例 6-2

linker 的合成

linker 6 的合成方法同实施例 6-1。

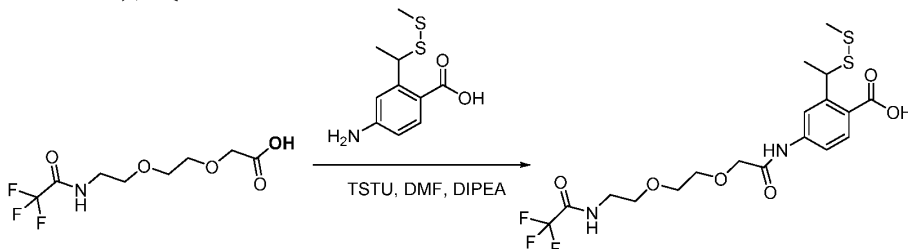


合成方法与实施例 7-1 类似, 16mg 固体, HPLC 纯度 > 99%。MS[ES(-)], m/z 1809.3。¹H NMR (400 MHz, DMSO-*d*₆) δ 10.16 (s, 1H), 8.95 (d, $J = 8.2$ Hz, 1H), 8.85 (t, $J = 5.5$ Hz, 1H), 8.80 (s, 2H), 8.23 (d, $J = 8.2$ Hz, 1H), 8.14 (d, $J = 8.4$ Hz, 2H), 7.99 (d, $J = 1.7$ Hz, 1H), 7.86 (d, $J = 7.6$ Hz, 1H), 7.82 (d, $J = 8.5$ Hz, 1H), 7.69 (dd, $J = 8.7, 1.6$ Hz, 1H), 7.62 (d, $J = 3.5$ Hz, 2H), 7.52 (d, $J = 7.9$ Hz, 2H), 7.45 – 7.37 (m, 1H), 6.72 (s, 2H), 6.30 – 6.20 (m, 1H), 5.57 – 5.54 (m, 1H), 5.31 (q, $J = 6.9$ Hz, 1H), 5.22 – 5.13 (m, 1H), 4.36 (d, $J = 20.9$ Hz, 1H), 4.24 – 4.19 (m, 2H), 4.14 (s, 2H), 4.08 (d, $J = 4.0$ Hz, 1H), 4.06 – 4.00 (m, 2H), 3.80 (dd, $J = 13.1, 6.8$ Hz, 2H), 3.73 – 3.69 (m, 2H), 3.67 (dd, $J = 7.0, 4.1$ Hz, 2H), 3.63 (d, $J = 5.6$ Hz, 2H), 2.22 (t, $J = 7.0$ Hz, 2H), 2.13 (s, 3H), 2.07 (t, $J = 7.8$ Hz, 5H), 1.73 – 1.66 (m, 2H), 1.65 (d, $J = 6.9$ Hz, 3H), 1.60 (d, $J = 7.0$ Hz, 3H), 1.17 (d, $J = 6.7$ Hz, 12H), 1.00 (d, $J = 2.0$ Hz, 6H)。

实施例 7-1

(1) 步骤 1

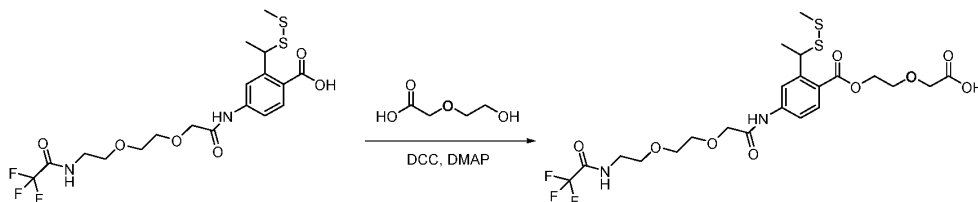
linker 7 的合成



PEG-CF₃ 底物 (OKeanos Tech, 货号 OK20A409) 1g 溶解在 DMF 中, 加入 TSTU (2 eq) 与 N,N-二异丙基乙胺 (3 eq), 反应在室温搅拌 2h 后, 加入苯胺底物 (OKeanos Tech, 货号 OK20A408) (1.2 eq)。反应 4h 后, 加入水淬灭反应。利用二氯甲烷萃取, 浓缩干燥后利用柱分离提纯得到无色油状化合物。MS[ES(-)], m/z 484.3。¹H NMR (400 MHz, DMSO-*d*₆) δ

12.92 (s, 1H), 9.98 (s, 1H), 9.53 (t, J = 5.2 Hz, 1H), 7.88 (d, J = 2.0 Hz, 1H), 7.81 (d, J = 8.4 Hz, 1H), 7.69 (dd, J = 8.4, 2.0 Hz, 1H), 5.48 (q, J = 6.8 Hz, 1H), 4.11 (s, 2H), 3.68 – 3.66 (m, 2H), 3.62 – 3.60 (m, 2H), 3.55 (t, J = 5.6 Hz, 2H), 3.37 (q, J = 5.6 Hz, 2H), 2.15 (s, 3H), 1.60 (d, J = 6.8 Hz, 3H).

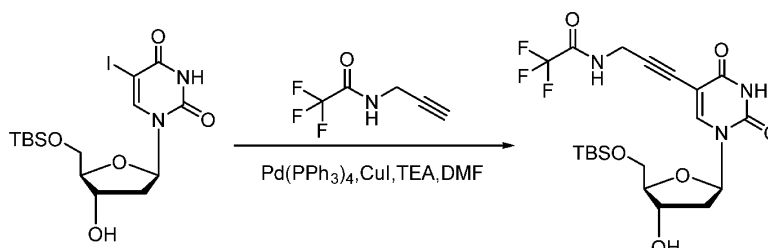
(2) 步骤 2



linker 7

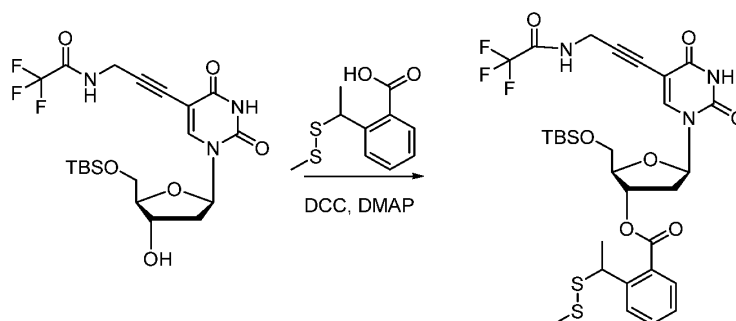
将步骤 1 获得的底物 100mg 溶解在 3mL 的 DMF 中, 加入 DCC (1.2 eq) 与 DMAP (10% mol), 反应搅拌 30 分钟后, 加入羧基酸底物 (毕得医药, 货号 BD628858) (2 eq)。反应搅拌 12 小时后, 直接利用制备 HPLC 进行分离 (C18 反相柱, 250*30 mm,) 得到白色固体 (linker 7) 85mg。MS[ES(-)], m/z 585.2. ¹H NMR (400 MHz, DMSO-d₆) δ 10.06 (s, 1H), 9.53 (t, J = 5.2 Hz, 1H), 7.92 (d, J = 2.0 Hz, 1H), 7.82 (d, J = 8.4 Hz, 1H), 7.73 (dd, J = 8.4, 2.0 Hz, 1H), 5.31 (q, J = 6.8 Hz, 1H), 4.40 – 4.30 (m, 3H), 4.11 (s, 2H), 3.88 (s, 2H), 3.78 (t, J = 4.8 Hz, 2H), 3.68 – 3.66 (m, 2H), 3.62 – 3.60 (m, 2H), 3.55 (t, J = 5.6 Hz, 2H), 3.37 (q, J = 5.6 Hz, 2H), 2.13 (s, 3H), 1.61 (d, J = 6.8 Hz, 3H).

(3) 步骤 3



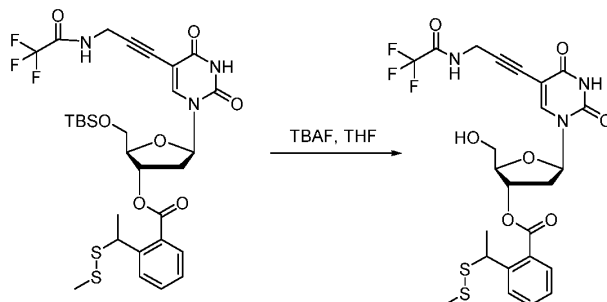
碘代 U 核苷底物 (OKeanos Tech, 货号 OK-N-16001) 1g 溶解于 DMF 中, 加入 Pd(PPh₃)₄ (10 mol%), CuI (15 mol%), 三乙胺 (3 eq) 与底物炔丙胺 (OKeanos Tech, 货号 OK20A410) (1.5eq) 在 60 度条件下反应 12h。加入水淬灭反应, DCM 萃取。浓缩后柱分离得到白色固体产品 1.1 g。MS[ES(-)], m/z 491.1. ¹H NMR (400 MHz, DMSO-d₆) δ 11.67 (s, 1H), 10.00 (t, J = 5.5 Hz, 1H), 7.94 (s, 1H), 6.12 (dd, J = 7.6, 5.9 Hz, 1H), 5.28 (d, J = 4.1 Hz, 1H), 4.26 – 4.12 (m, 3H), 3.88 (q, J = 2.8 Hz, 1H), 3.81 (dd, J = 11.5, 2.6 Hz, 1H), 3.73 (dd, J = 11.5, 3.1 Hz, 1H), 2.17 (ddd, J = 13.2, 6.0, 2.8 Hz, 1H), 2.05 (ddd, J = 13.3, 7.7, 5.8 Hz, 1H), 0.87 (s, 9H), 0.08 (d, J = 1.8 Hz, 6H).

(4) 步骤 4



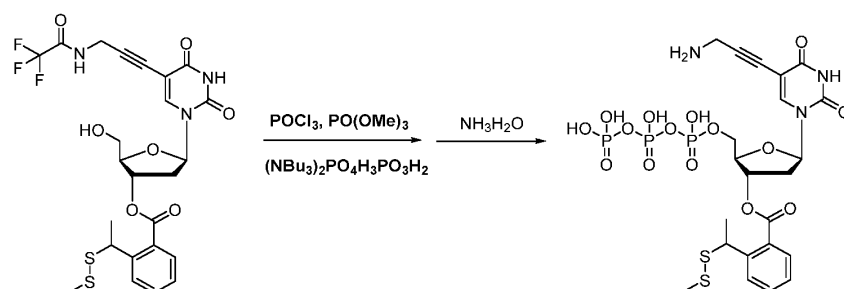
将步骤3获得的底物核苷300mg溶解在10mL的DMF中,加入DCC(1.2 eq)与DMAP(10% mol),反应搅拌30分钟后,加入二羧酸底物(OKeanos Tech, 货号OK20A420)(1.5 eq)。反应搅拌12小时后,直接柱分离得到白色固体359mg。MS[ES(-)],m/z 700.3。¹H NMR (400 MHz, DMSO-*d*₆) δ 11.74 (s, 1H), 10.03 (t, J = 5.5 Hz, 1H), 7.98 (s, 1H), 7.90 – 7.82 (m, 1H), 7.68 – 7.58 (m, 2H), 7.47 – 7.40 (m, 1H), 6.23 – 6.18 (m, 1H), 5.45 – 5.42 (m, 1H), 5.20 – 5.12 (m, 1H), 4.38 – 4.16 (m, 3H), 3.98 – 3.87 (m, 2H), 2.60 – 2.53 (m, 1H), 2.40 – 2.31 (m, 1H), 2.06 (d, J = 0.6 Hz, 3H), 1.65 (dd, J = 7.0, 1.0 Hz, 3H), 0.90 (s, 9H), 0.13 (d, J = 1.2 Hz, 6H)。

(5) 步骤5



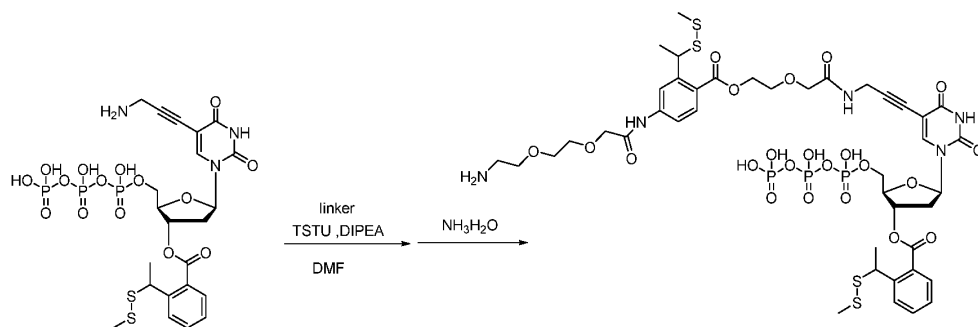
将步骤4获得的底物核苷300mg溶解在10mL的THF中,在0度下加入TBAF(2 eq, 1M in THF),反应在0度搅拌30分钟后,升至室温搅拌4小时。反应直接柱分离得到白色固体200mg。MS[ES(-)],m/z 587.2。¹H NMR (400 MHz, DMSO-*d*₆) δ 11.68 (d, J = 3.2 Hz, 1H), 10.06 (t, J = 5.6 Hz, 1H), 8.23 (d, J = 3.2 Hz, 1H), 7.88 – 7.78 (m, 1H), 7.66 – 7.56 (m, 2H), 7.43 – 7.39 (m, 1H), 6.25 – 6.13 (m, 1H), 5.48 – 5.45 (m, 1H), 5.31 (t, J = 5.2 Hz, 1H), 5.18 – 5.12 (m, 1H), 4.28 – 4.16 (m, 3H), 3.76 – 3.67 (m, 2H), 2.50 – 2.40 (m, 2H), 2.04 (d, J = 9.2 Hz, 3H), 1.68 – 1.59 (m, 3H)。

(6) 步骤6



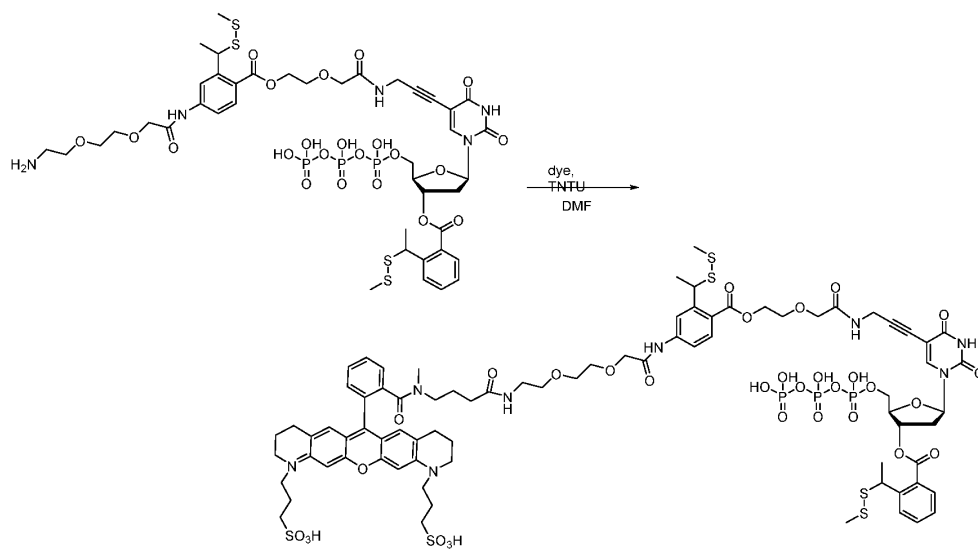
将步骤5获得的底物核苷200mg溶解在5mL的磷酸三甲酯中,在0度下加入三氯氧磷(1.5 eq),反应在0度搅拌120分钟后,将反应液加入到三丁基焦磷酸铵(2 eq)的DMF(5 mL)溶液中,反应在0度继续搅拌3h后,加入0.1 M TEAB缓冲液淬灭。反应利用制备HPLC反相柱(C18, 流动相: 0.1 M TEAB-乙腈)分离,浓缩后直接加入浓氨水3 mL,反应2h后利用制备HPLC反相柱(C18, 流动相: 0.1 M TEAB-乙腈)分离得到白色固体120mg。MS[ES(-)],m/z 745.5。¹H NMR (400 MHz, D₂O) δ 8.51 (s, 1H), 7.90 (dd, J = 7.9, 1.4 Hz, 1H), 7.73 (d, J = 7.8 Hz, 1H), 7.67 (td, J = 7.6, 1.4 Hz, 1H), 7.48 (t, J = 7.5 Hz, 1H), 6.45 (td, J = 9.2, 5.7 Hz, 1H), 5.75 (t, J = 4.7 Hz, 1H), 5.15 – 5.07 (m, 1H), 4.72 – 4.64 (m, 1H), 4.39 (d, J = 3.3 Hz, 2H), 4.07 (s, 2H), 2.82 – 2.76 (m, 1H), 2.64 – 2.54 (m, 1H), 2.02 (d, J = 6.4 Hz, 3H), 1.73 (dd, J = 7.1, 2.2 Hz, 3H); ³¹P NMR (162 MHz, D₂O) δ -9.79 (dd, J = 20.3, 9.6 Hz, 1P), -11.68 (d, J = 19.2 Hz, 1P), -22.82 (td, J = 19.2, 6.1 Hz, 1P)。

(7) 步骤7



将步骤2获得的 linker 7 底物 20mg 溶解在 1mL 的 DMF 中, 加入 TSTU (1.5 eq) 与二异丙基乙基胺 (3 eq), 反应在室温搅拌 120 分钟后, 加入步骤6获得的核苷酸底物 (2 eq), 反应在室温继续搅拌 3h 后, 加入 0.1 M TEAB 缓冲液淬灭。反应利用制备 HPLC 反相柱 (C18, 流动相: 0.1 M TEAB-乙腈) 分离, 浓缩后直接加入浓氨水 3 mL, 反应 2h 后利用制备 HPLC 反相柱 (C18, 流动相: 0.1 M TEAB-乙腈) 分离得到白色固体 15mg。MS[ES(-)], m/z 1201.7。¹H NMR (400 MHz, DMSO-d₆) δ 10.91 (s, 1H), 8.96–8.84 (m, 1H), 8.28 (d, J = 2.0 Hz, 1H), 8.18 (d, J = 11.6 Hz, 1H), 7.92 (dd, J = 8.6, 1.5 Hz, 1H), 7.86 (d, J = 7.1 Hz, 1H), 7.66 (d, J = 8.7 Hz, 1H), 7.63–7.57 (m, 2H), 7.46–7.38 (m, 1H), 6.21 (ddd, J = 14.8, 8.4, 6.2 Hz, 1H), 5.57 (s, 1H), 5.35 (q, J = 6.8 Hz, 1H), 5.21–5.13 (m, 1H), 4.39 (d, J = 12.5 Hz, 2H), 4.32 (s, 2H), 4.22 (s, 2H), 4.14 (d, J = 6.3 Hz, 2H), 4.09 (d, J = 3.7 Hz, 2H), 4.05 (s, 2H), 3.82–3.79 (m, 4H), 3.68–3.64 (m, 2H), 3.63–3.59 (m, 2H), 2.99–2.93 (m, 2H), 2.60–2.52 (m, 1H), 2.47–2.36 (m, 1H), 2.15 (s, 3H), 2.05 (d, J = 21.1 Hz, 3H), 1.64 (d, J = 7.0 Hz, 3H), 1.61 (d, J = 6.9 Hz, 3H)。

(8) 步骤 8



将染料 (北京澄道科技有限公司, 货号 CD0014) 10mg 溶解在 0.5 mL 的 DMF 中, 加入 TNTU (1.5 eq) 与二异丙基乙基胺 (3 eq), 反应在室温搅拌 120 分钟后, 加入步骤7获得的核苷酸底物 (2 eq), 反应在室温继续搅拌 3h 后, 加入 0.1 M TEAB 缓冲液淬灭。反应利用制备 HPLC 反相柱 (C18, 流动相: 0.1 M TEAB-乙腈) 分离得到红色固体 8 mg。HPLC 纯度 > 99%。

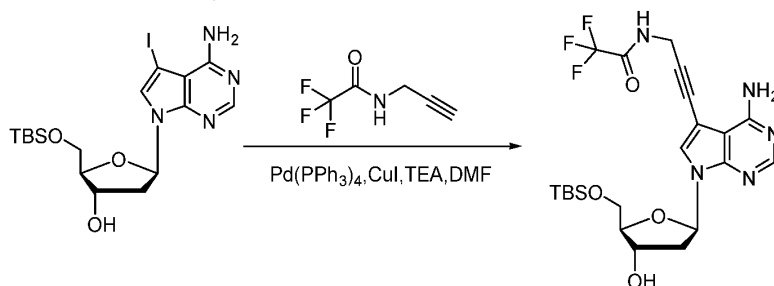
MS[ES(-)], m/z 1938.7。¹H NMR (400 MHz, DMSO-d₆) δ 10.39 (brs, 1H), 9.15–9.10 (m, 1H), 8.28 (d, J = 8.8 Hz, 1H), 8.03–8.00 (m, 1H), 7.88–7.79 (m, 3H), 7.76–7.67 (m, 3H), 7.66–7.60 (m, 3H), 7.58–7.39 (m, 3H), 6.99 (s, 2H), 6.80 (s, 2H), 6.27–6.20 (m, 1H), 5.59 (d, J = 5.6 Hz, 1H), 5.34–5.29 (m, 1H), 5.18 (p, J = 6.8 Hz, 1H), 4.44–4.29 (m, 4H), 4.14–4.04 (m, 9H), 3.81 (t, J = 4.8 Hz, 2H), 3.77–3.61 (m, 7H), 3.59–3.51 (m, 6H), 3.40 (t, J = 6.0 Hz, 2H), 3.19–3.14 (m, 2H), 3.11 (t, J = 7.6 Hz, 2H), 2.69–2.65 (d, J = 6.5 Hz, 4H), 2.59–2.55 (m, 5H), 2.46–

2.35 (m, 1H), 2.14 (d, J = 3.5 Hz, 3H), 2.08 (s, 1H), 2.04 (s, 1H), 2.00 – 1.91 (m, 4H), 1.87 – 1.80 (m, 4H), 1.70 (t, J = 7.6 Hz, 2H), 1.65 (d, J = 7.2 Hz, 3H), 1.60 (d, J = 6.8 Hz, 3H). ³¹P NMR (162 MHz, DMSO-d₆) δ -11.24 (d, J = 22.7 Hz, 1P), -12.05 (d, J = 22.9 Hz, 1P), -23.63 (t, J = 22.7 Hz, 1P).

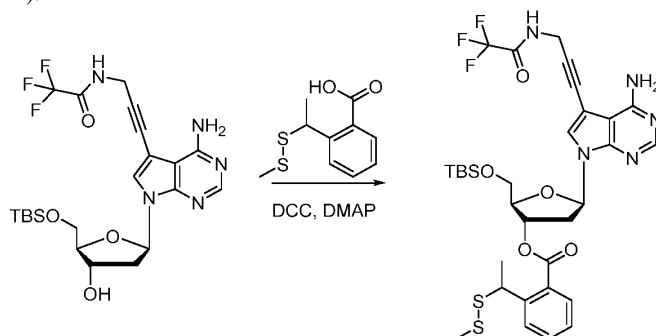
实施例 7-2

linker 的合成

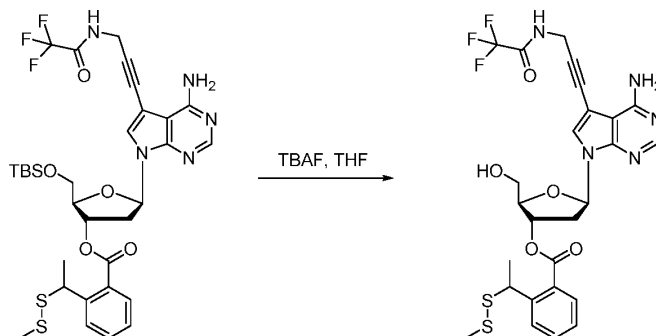
linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



合成方法与实施例 7-1 类似, 使用碘代 A 核苷底物 (Okeanos Tech, 货号 OK-N-16003) 替代碘代 U 核苷底物。MS[ES(-)], m/z 513.2. ¹H NMR (400 MHz, DMSO-d₆) δ 10.04 (t, J = 5.5 Hz, 1H), 8.09 (s, 1H), 7.67 (s, 1H), 6.85 (s, 2H), 6.47 (t, J = 6.7 Hz, 1H), 5.29 (d, J = 4.2 Hz, 1H), 4.33 – 4.29 (m, 1H), 4.27 (d, J = 5.4 Hz, 2H), 3.81 (q, J = 3.9 Hz, 1H), 3.75 (dd, J = 11.1, 4.1 Hz, 1H), 3.67 (dd, J = 11.1, 4.1 Hz, 1H), 2.43 – 2.36 (m, 1H), 2.24 – 2.18 (m, 1H), 0.85 (s, 9H), 0.03 (d, J = 1.8 Hz, 6H).

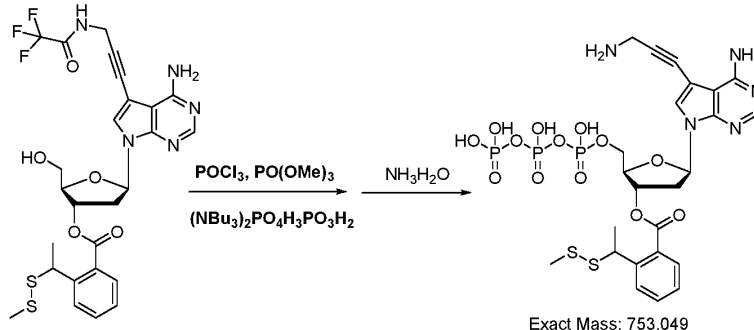


合成方法与实施例 7-1 类似。MS[ES(-)], m/z 723.1. ¹H NMR (400 MHz, DMSO-d₆) δ 10.05 (s, 1H), 8.10 (d, J = 4.4 Hz, 1H), 7.91 – 7.83 (m, 1H), 7.74 (d, J = 3.6 Hz, 1H), 7.65 – 7.56 (m, 2H), 7.46 – 7.40 (m, 1H), 6.84 (brs, 2H), 6.64 – 6.48 (m, 1H), 5.58 – 5.52 (m, 1H), 5.24 – 5.10 (m, 1H), 4.37 – 4.15 (m, 3H), 3.95 – 3.74 (m, 2H), 2.84 – 2.75 (m, 1H), 2.70 – 2.54 (m, 1H), 2.06 (d, J = 1.2 Hz, 3H), 1.64 (d, J = 7.0 Hz, 3H), 1.28 – 1.17 (m, 1H), 0.87 (s, 9H), 0.07 (dd, J = 4.8, 2.7 Hz, 6H).



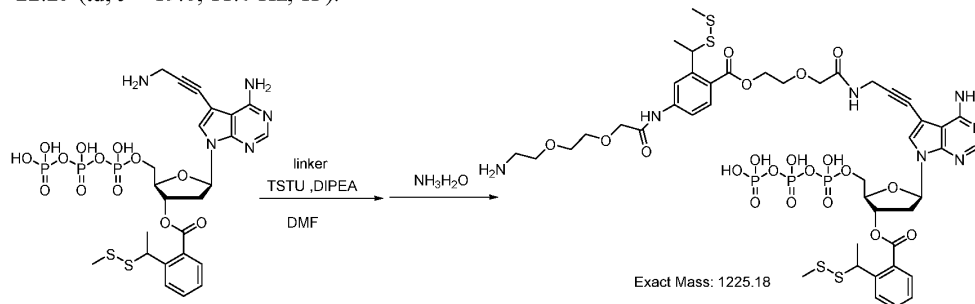
合成方法与实施例 7-1 类似。MS[ES(-)], m/z 609.2. ¹H NMR (400 MHz, DMSO-d₆) δ 10.10 (t, J = 5.2 Hz, 1H), 8.12 (d, J = 4.8 Hz, 1H), 7.90 (dd, J = 7.4, 5.0 Hz, 1H), 7.85 (s, 1H), 7.66 –

7.61 (m, 2H), 7.48 – 7.43 (m, 1H), 6.82 (brs, 2H), 6.60 – 6.51 (m, 1H), 5.59 (d, $J = 4.4$ Hz, 1H), 5.36 (td, $J = 5.5, 1.5$ Hz, 1H), 5.21 (qd, $J = 7.0, 4.1$ Hz, 1H), 4.32 (d, $J = 5.3$ Hz, 2H), 4.26 – 4.20 (m, 1H), 3.75 – 3.64 (m, 2H), 2.90 – 2.77 (m, 1H), 2.65 – 2.51 (m, 1H), 2.09 (d, $J = 8.6$ Hz, 3H), 1.67 (d, $J = 7.0$ Hz, 3H).



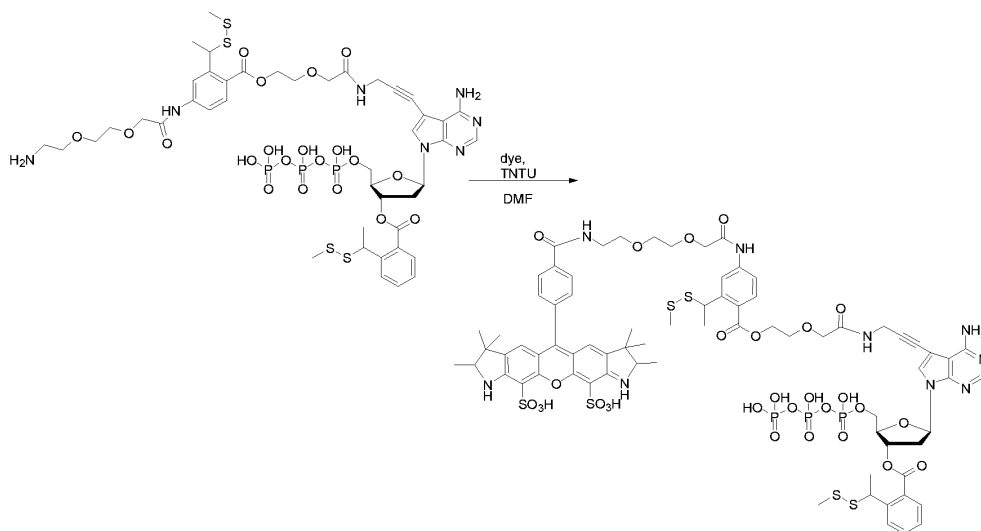
Exact Mass: 753.049

合成方法与实施例 7-1 类似。MS[ES(-)], m/z 753.0. ^1H NMR (400 MHz, D_2O) δ 8.04 (d, $J = 7.1$ Hz, 1H), 7.95 – 7.83 (m, 2H), 7.72 – 7.56 (m, 2H), 7.49 – 7.44 (m, 1H), 6.70 – 6.48 (m, 1H), 5.71 (dd, $J = 17.3, 4.9$ Hz, 1H), 5.12 – 5.00 (m, 1H), 4.62 (d, $J = 10.7$ Hz, 1H), 4.41 – 4.21 (m, 2H), 4.18 (d, $J = 8.9$ Hz, 2H), 2.78 – 2.31 (m, 2H), 2.03 – 1.83 (m, 3H), 1.70 (dd, $J = 10.3, 7.1$ Hz, 3H); ^{31}P NMR (162 MHz, D_2O) δ -7.18 (dd, $J = 20.1, 13.4$ Hz, 1P), -11.24 (d, $J = 19.0$ Hz, 1P), -22.25 (td, $J = 19.6, 11.0$ Hz, 1P).



Exact Mass: 1225.18

合成方法与实施例 7-1 类似。MS[ES(-)], m/z 1225.1. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 11.18 (d, $J = 6.1$ Hz, 1H), 8.37 (t, $J = 5.4$ Hz, 1H), 8.21 (dd, $J = 7.4, 1.7$ Hz, 1H), 8.09 (s, 1H), 7.84 (d, $J = 8.5$ Hz, 1H), 7.69 – 7.58 (m, 3H), 7.47 – 7.40 (m, 1H), 7.29 (s, 1H), 6.40 (s, 2H), 6.30 (dd, $J = 9.4, 5.7$ Hz, 1H), 5.66 (s, 1H), 5.35 (q, $J = 6.9$ Hz, 1H), 5.20 (q, $J = 6.9$ Hz, 1H), 4.41 (d, $J = 12.9$ Hz, 1H), 4.34 (dd, $J = 8.4, 4.2$ Hz, 1H), 4.30 (s, 2H), 4.13 – 4.08 (m, 2H), 4.06 – 4.03 (m, 2H), 4.01 (s, 2H), 3.83 – 3.80 (m, 2H), 3.67 (d, $J = 2.8$ Hz, 2H), 3.63 – 3.61 (m, 6H), 2.98 – 2.93 (m, 2H), 2.60 – 2.52 (m, 1H), 2.47 – 2.36 (m, 1H), 2.13 (s, 3H), 2.07 (s, 3H), 1.64 (d, $J = 7.0$ Hz, 3H), 1.61 (d, $J = 6.9$ Hz, 3H).

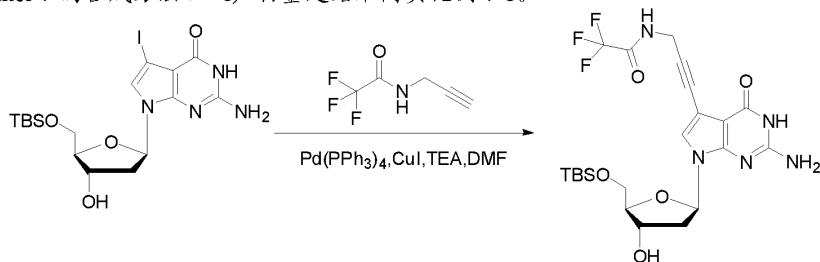


合成方法与实施例 7-1 类似, 将染料替换为 AF532(购自 OKeanos Tech, 货号 OK F 532)。13mg 固体, HPLC 纯度 > 99%。
 MS[ES(-)], m/z 1833.0. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.43 (t, $J = 4.8$ Hz, 1H), 8.91 (t, $J = 5.6$ Hz, 1H), 8.80 (s, 2H), 8.50 (t, $J = 5.6$ Hz, 1H), 8.15 (d, $J = 8.8$ Hz, 2H), 8.07 (d, $J = 6.4$ Hz, 1H), 8.03 (s, 1H), 7.90 (dd, $J = 7.6, 2.4$ Hz, 1H), 7.84 – 7.79 (m, 2H), 7.69 (d, $J = 8.4$ Hz, 1H), 7.63 (d, $J = 4.4$ Hz, 2H), 7.49 – 7.42 (m, 3H), 6.70 (s, 2H), 6.65 – 6.58 (m, 1H), 5.69 (d, $J = 2.8$ Hz, 1H), 5.33 – 5.27 (m, 1H), 5.21 (qd, $J = 7.2, 2.0$ Hz, 1H), 4.45 – 4.34 (m, 3H), 4.32 – 4.30 (m, 1H), 4.18 (s, 2H), 4.14 (d, $J = 5.6$ Hz, 3H), 4.01 (s, 3H), 3.81 – 3.77 (m, 7H), 3.73 – 3.69 (m, 4H), 3.68 – 3.63 (m, 7H), 3.52 (q, $J = 5.6$ Hz, 4H), 2.12 (d, $J = 2.4$ Hz, 3H), 2.10 (s, 1H), 2.05 (s, 1H), 1.65 (d, $J = 6.8$ Hz, 3H), 1.59 (d, $J = 7.2$ Hz, 3H), 0.99 (d, $J = 2.8$ Hz, 5H). ^{31}P NMR (162 MHz, $\text{DMSO}-d_6$) δ -10.99 (d, $J = 20.6$ Hz, 1P), -12.48 (d, $J = 24.0$ Hz, 1P), -23.41 (t, $J = 22.5$ Hz, 1P).

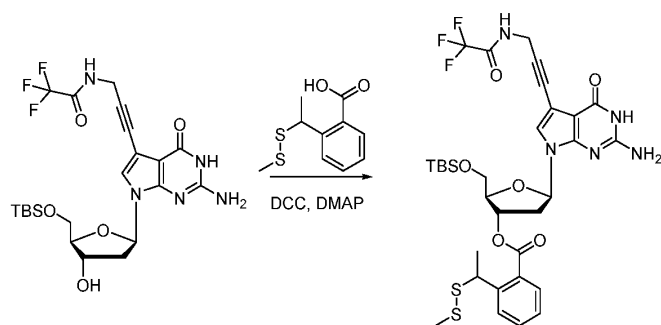
实施例 7-3

linker 的合成

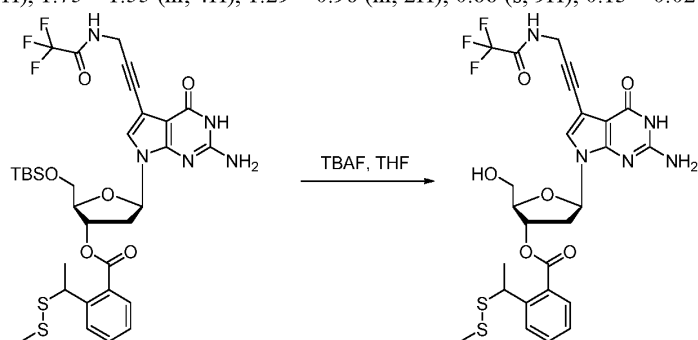
linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



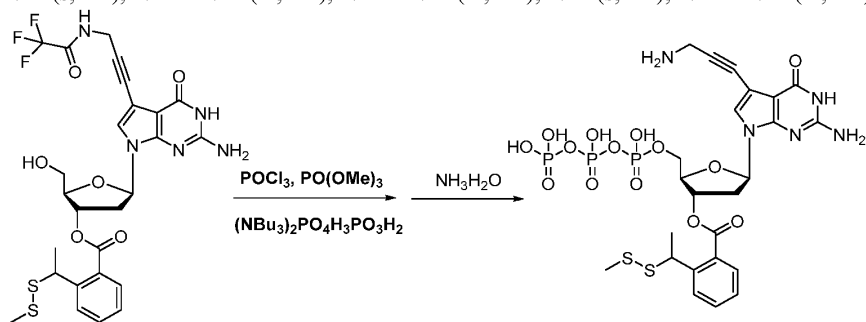
合成方法与实施例 7-1 类似, 使用碘代 G 核苷底物 (OKeanos Tech, 货号 OK-N-16004) 替代碘代 U 核苷底物。MS[ES(-)], m/z 529.3. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.48 (s, 1H), 9.98 (t, $J = 5.6$ Hz, 1H), 7.21 (s, 1H), 6.38 – 6.21 (m, 3H), 5.23 (d, $J = 3.8$ Hz, 1H), 4.26 (s, 1H), 4.18 (d, $J = 5.6$ Hz, 2H), 3.78 (q, $J = 4.0$ Hz, 1H), 3.73 – 3.60 (m, 2H), 2.32 – 2.19 (m, 1H), 2.13 – 2.08 (m, 1H), 2.04 (s, 1H), 0.85 (s, 9H), 0.03 (d, $J = 1.8$ Hz, 6H).



合成方法与实施例 7-1 类似。MS[ES(-)], m/z 739.2. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.53 (s, 1H), 10.00 (t, $J = 5.6$ Hz, 1H), 7.83 (dd, $J = 7.5, 3.6$ Hz, 1H), 7.61 – 7.58 (m, 2H), 7.48 – 7.39 (m, 1H), 7.29 (d, $J = 2.0$ Hz, 1H), 6.42 – 6.25 (m, 3H), 5.72 (s, 1H), 5.58 – 5.46 (m, 1H), 5.16 (p, $J = 7.2$ Hz, 1H), 4.25 – 4.14 (m, 3H), 3.82 (t, $J = 3.6$ Hz, 2H), 2.71 – 2.60 (m, 1H), 2.50 – 2.45 (m, 1H), 2.07 (s, 3H), 1.75 – 1.55 (m, 4H), 1.29 – 0.96 (m, 2H), 0.88 (s, 9H), 0.13 – 0.02 (m, 6H).

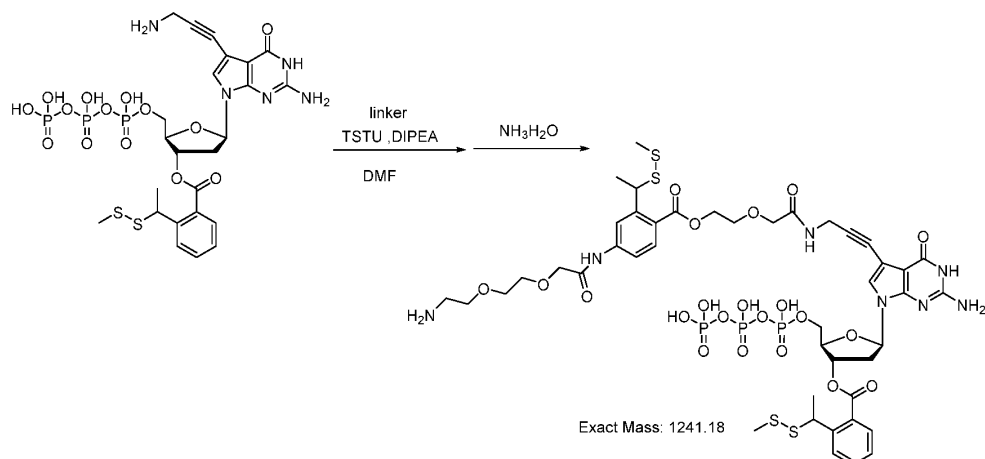


合成方法与实施例 7-1 类似。MS[ES(-)], m/z 625.7. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.53 (s, 1H), 10.00 (t, $J = 5.6$ Hz, 1H), 7.83 (dd, $J = 7.5, 3.6$ Hz, 1H), 7.61 – 7.58 (m, 2H), 7.48 – 7.39 (m, 1H), 7.29 (d, $J = 2.0$ Hz, 1H), 6.42 – 6.25 (m, 3H), 5.72 (s, 1H), 5.58 – 5.46 (m, 1H), 5.16 (p, $J = 7.2$ Hz, 1H), 4.25 – 4.14 (m, 3H), 3.82 (t, $J = 3.6$ Hz, 2H), 2.71 – 2.60 (m, 1H), 2.50 – 2.45 (m, 1H), 2.07 (s, 3H), 1.75 – 1.55 (m, 4H), 1.29 – 0.96 (m, 2H), 0.88 (s, 9H), 0.13 – 0.02 (m, 6H).

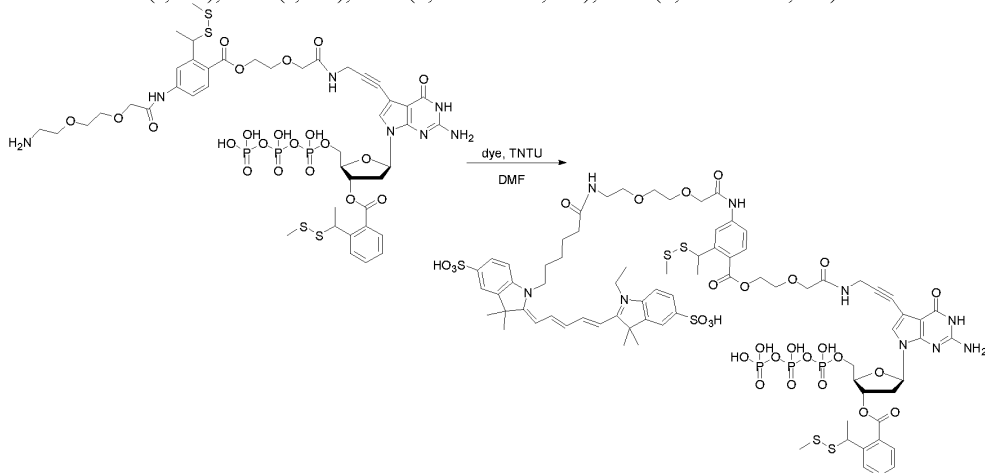


Exact Mass: 769.04

合成方法与实施例 7-1 类似。MS[ES(-)], m/z 769.2. ^1H NMR (400 MHz, D_2O) δ 7.81 – 7.74 (m, 1H), 7.51 – 7.40 (m, 3H), 7.31 (q, $J = 6.9$ Hz, 1H), 6.38 – 6.19 (m, 1H), 5.65 (dd, $J = 11.2, 5.2$ Hz, 1H), 5.07 – 4.96 (m, 1H), 4.53 (d, $J = 20.8$ Hz, 1H), 4.35 – 4.18 (m, 2H), 4.10 (s, 2H), 2.87 – 2.70 (m, 1H), 2.54 – 2.40 (m, 1H), 1.97 – 1.76 (m, 3H), 1.59 – 1.54 (m, 3H); ^{31}P NMR (162 MHz, D_2O) δ -9.27 (d, $J = 19.6$ Hz, 1P), -11.36 (d, $J = 18.5$ Hz, 1P), -22.79 (td, $J = 19.2, 4.7$ Hz, 1P).



合成方法与实施例 7-1 类似。MS[ES(-)],m/z 1241.6. ^1H NMR (400 MHz, DMSO- d_6) δ 11.25 (s, 1H), 8.29 (d, J = 7.7 Hz, 2H), 7.87 (d, J = 7.0 Hz, 2H), 7.69 – 7.58 (m, 3H), 7.47 – 7.40 (m, 1H), 7.29 (s, 1H), 6.40 (s, 2H), 6.30 (dd, J = 9.4, 5.7 Hz, 1H), 5.66 (s, 1H), 5.35 (q, J = 6.9 Hz, 1H), 5.20 (q, J = 6.9 Hz, 1H), 4.41 (d, J = 12.9 Hz, 1H), 4.34 (dd, J = 8.4, 4.2 Hz, 1H), 4.30 (s, 2H), 4.13 – 4.08 (m, 2H), 4.06 – 4.03 (m, 2H), 4.01 (s, 2H), 3.83 – 3.80 (m, 2H), 3.67 (d, J = 2.8 Hz, 2H), 3.63 – 3.61 (m, 6H), 2.98 – 2.93 (m, 2H), 2.47 – 2.42 (m, 1H), 2.41 – 2.30 (m, 1H), 2.13 (s, 3H), 2.07 (s, 3H), 1.66 (d, J = 7.0 Hz, 3H), 1.62 (d, J = 6.9 Hz, 3H).



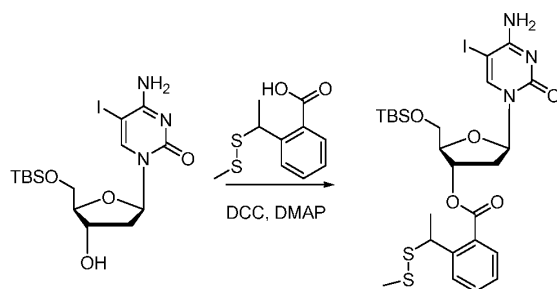
合成方法与实施例 7-1 类似,将染料替换为 CY5(购自 OKeanos Tech, 货号 OK-F-13103)。18mg 固体, HPLC 纯度 > 99%。

MS[ES(-)],m/z 1878.9. ^1H NMR (400 MHz, DMSO- d_6) δ 10.66 (brs, 1H), 10.43 (s, 1H), 8.43 – 8.32 (m, 2H), 8.03 (s, 1H), 7.92 – 7.82 (m, 2H), 7.82 (d, J = 1.6 Hz, 2H), 7.72 (dd, J = 8.8, 2.0 Hz, 1H), 7.66 – 7.60 (m, 3H), 7.46 – 7.42 (m, 1H), 7.31 (dd, J = 8.0, 3.6 Hz, 2H), 6.59 (t, J = 12.4 Hz, 1H), 6.58 – 6.47 (m, 1H), 6.35 – 6.27 (m, 2H), 5.66 (t, J = 5.2 Hz, 1H), 5.31 (q, J = 7.2 Hz, 1H), 5.20 (q, J = 7.2 Hz, 1H), 4.45 – 4.34 (m, 2H), 4.31 (t, J = 5.2 Hz, 1H), 4.25 (t, J = 5.2 Hz, 1H), 4.14 (s, 2H), 4.12 – 4.05 (m, 5H), 4.00 (s, 2H), 3.81 (t, J = 4.4 Hz, 2H), 3.66 – 3.64 (m, 2H), 3.58 – 3.55 (m, 2H), 3.41 (t, J = 5.6 Hz, 2H), 3.19 (q, J = 5.6 Hz, 2H), 2.13 (d, J = 3.6 Hz, 3H), 2.08 – 2.05 (m, 2H), 1.68 (s, 9H), 1.65 (d, J = 7.2 Hz, 2H), 1.61 (d, J = 7.2 Hz, 2H), 1.56 – 1.49 (m, 2H), 1.35 – 1.30 (m, 2H), 1.25 (t, J = 7.2 Hz, 3H). ^{31}P NMR (162 MHz, DMSO- d_6) δ -11.33 (d, J = 22.6 Hz, 1P), -11.92 (d, J = 22.8 Hz, 1P), -23.68 (t, J = 22.6 Hz, 1P).

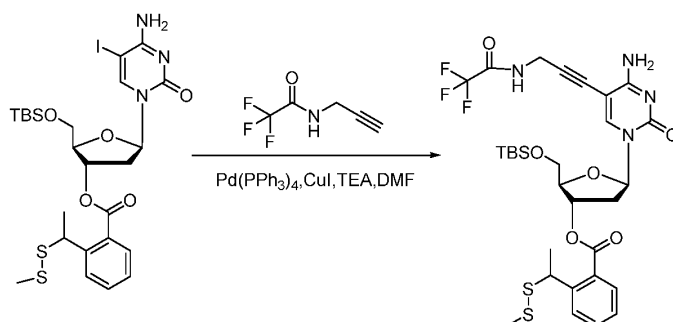
实施例 7-4

linker 的合成

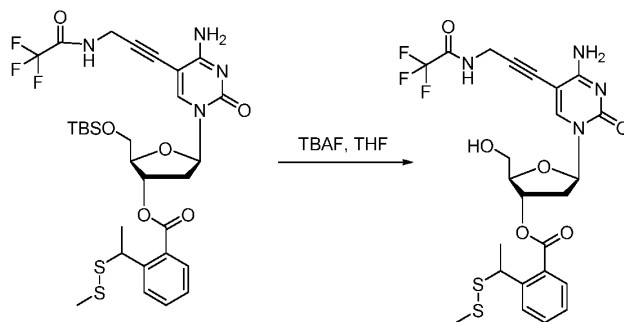
linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



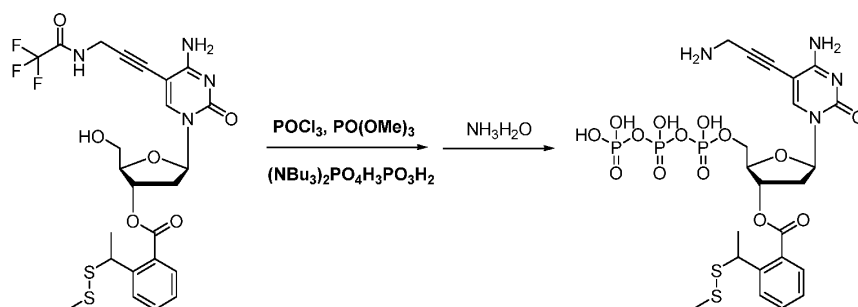
合成方法与实施例 7-1 类似, 使用碘代 C 核苷底物 (OKeanos Tech, 货号 OK-N-16002) 替代碘代 U 核苷底物。MS[ES(-)], m/z 678.1. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 7.97 (s, 1H), 7.94 – 7.79 (m, 2H), 7.65 – 7.54 (m, 2H), 7.46 – 7.33 (m, 1H), 6.68 (s, 1H), 6.18 – 6.14 (m, 1H), 5.40 (d, $J = 6.1$ Hz, 1H), 5.16 (qd, $J = 7.0, 2.1$ Hz, 1H), 4.35 – 4.24 (m, 1H), 3.97 – 3.85 (m, 2H), 2.55 – 2.47 (m, 1H), 2.24 – 2.16 (m, 1H), 2.04 (d, $J = 3.2$ Hz, 3H), 1.63 (d, $J = 7.0$ Hz, 3H), 0.89 (s, 9H), 0.12 (s, 6H).



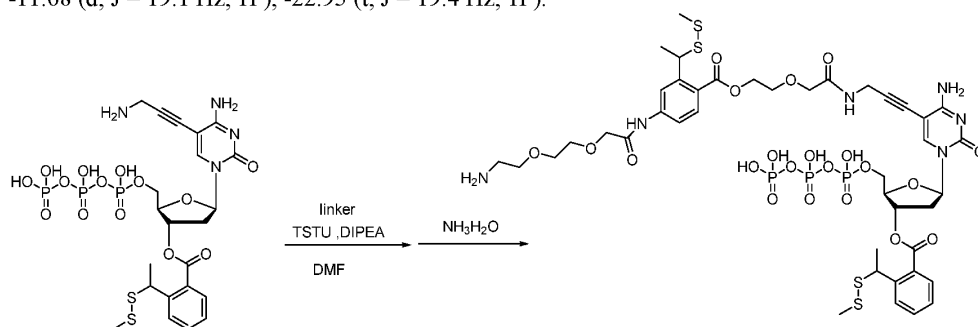
合成方法与实施例 7-1 类似。MS[ES(-)], m/z 700.4. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 9.93 (t, $J = 5.2$ Hz, 1H), 7.99 (s, 1H), 7.91 (s, 1H), 7.85 (d, $J = 7.6$ Hz, 1H), 7.62 (d, $J = 5.3$ Hz, 2H), 7.47 – 7.40 (m, 1H), 6.97 (s, 1H), 6.25 – 6.20 (m, 1H), 5.42 (d, $J = 5.9$ Hz, 1H), 5.17 (qd, $J = 7.0, 2.2$ Hz, 1H), 4.36 – 4.29 (m, 1H), 4.27 (d, $J = 5.2$ Hz, 2H), 3.95 – 3.87 (m, 2H), 2.65 – 2.52 (m, 1H), 2.27 – 2.17 (m, 1H), 2.06 (d, $J = 2.8$ Hz, 3H), 1.65 (d, $J = 7.0$ Hz, 3H), 0.88 (s, 9H), 0.11 (d, $J = 0.8$ Hz, 6H).



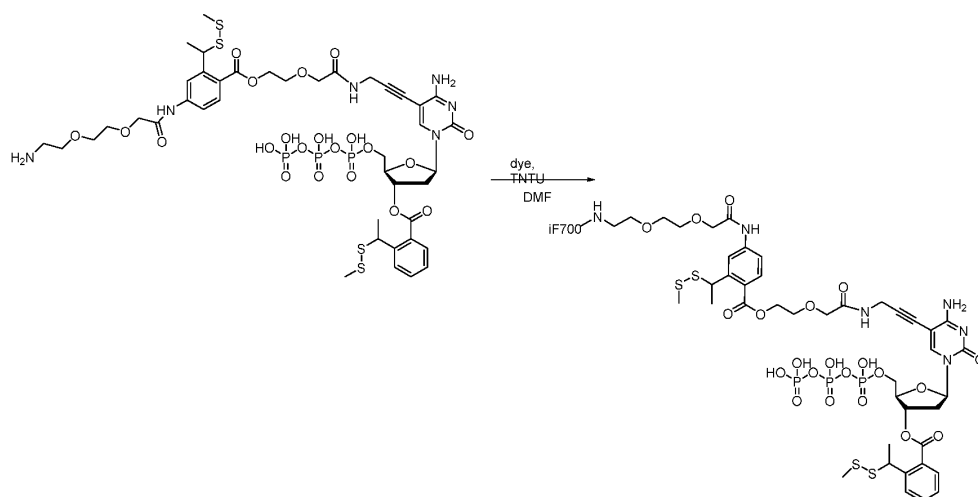
合成方法与实施例 7-1 类似。MS[ES(-)], m/z 580.5. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 9.94 (s, 1H), 8.18 (s, 1H), 7.89 (s, 1H), 7.82 (d, $J = 7.6$ Hz, 1H), 7.61 (q, $J = 4.1, 3.0$ Hz, 2H), 7.41 (t, $J = 8.1$ Hz, 1H), 6.91 (s, 1H), 6.26 – 6.17 (m, 1H), 5.45 (d, $J = 6.0$ Hz, 1H), 5.28 (t, $J = 5.2$ Hz, 1H), 5.15 (qd, $J = 7.0, 3.6$ Hz, 1H), 4.35 – 4.16 (m, 3H), 3.70 (dt, $J = 5.6, 3.3$ Hz, 2H), 2.54 – 2.44 (m, 1H), 2.33 – 2.26 (m, 1H), 2.04 (d, $J = 5.7$ Hz, 3H), 1.63 (d, $J = 7.0$ Hz, 3H).



合成方法与实施例 7-1 类似。MS[ES(-)],m/z 745.3. ^1H NMR (400 MHz, D_2O) δ 8.51 (s, 1H), 7.91 (d, J = 7.7 Hz, 1H), 7.74 (d, J = 7.9 Hz, 1H), 7.68 (t, J = 7.7 Hz, 1H), 7.48 (t, J = 7.5 Hz, 1H), 6.47 - 6.42 (m, 1H), 5.73 (d, J = 5.6 Hz, 1H), 5.18 - 5.06 (m, 1H), 4.71 (d, J = 9.0 Hz, 1H), 4.39 (q, J = 12.5 Hz, 2H), 4.09 (d, J = 1.6 Hz, 2H), 2.83 - 2.79 (m, 1H), 2.58 - 2.50 (m, 1H), 2.03 (dd, J = 3.0, 1.5 Hz, 3H), 1.74 (d, J = 7.0 Hz, 3H); ^{31}P NMR (162 MHz, D_2O) δ -9.99 (d, J = 19.4 Hz, 1P), -11.68 (d, J = 19.1 Hz, 1P), -22.93 (t, J = 19.4 Hz, 1P).



合成方法与实施例 7-1 类似。MS[ES(-)],m/z 1201.6. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.99 (s, 1H), 8.75 (d, J = 7.2 Hz, 1H), 8.23 (s, 1H), 8.07 (d, J = 7.3 Hz, 1H), 7.87 - 7.84 (m, 3H), 7.70 - 7.58 (m, 3H), 7.47 - 7.37 (m, 1H), 6.95 (s, 1H), 6.23 (dd, J = 16.7, 8.2 Hz, 1H), 5.55 (s, 1H), 5.33 (q, J = 6.8 Hz, 1H), 5.21 - 5.14 (m, 1H), 4.44 - 4.36 (m, 2H), 4.32 (dd, J = 7.8, 4.1 Hz, 2H), 4.24 (s, 2H), 4.13 (d, J = 3.3 Hz, 4H), 4.04 (s, 2H), 3.82 (s, 2H), 3.68 - 3.55 (m, 6H), 2.97 (s, 2H), 2.46 - 2.34 (m, 2H), 2.13 (d, J = 2.1 Hz, 3H), 2.05 (d, J = 16.9 Hz, 3H), 1.64 (d, J = 6.7 Hz, 3H), 1.61 (d, J = 6.9 Hz, 3H).



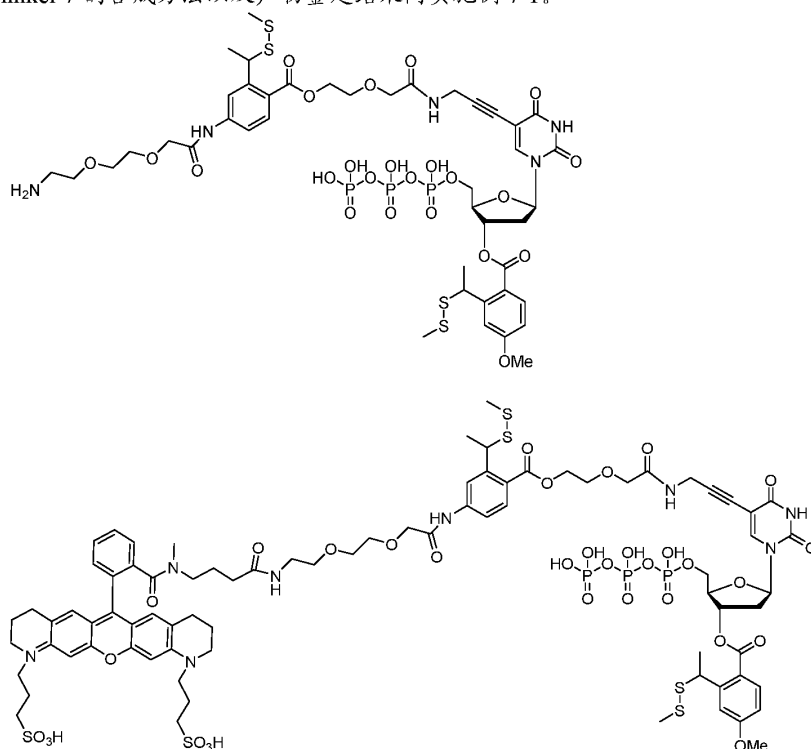
合成方法与实施例 7-1 类似, 其中, 将染料替换为 iF700 商品 (iFluor™ 700 succinimidyl ester, 购自 AAT Bioquest)。16mg 固体, HPLC 纯度 > 99%。MS[ES(-)],m/z 1962.8. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.30 (s, 1H), 8.91 - 8.84 (m, 1H), 8.80 - 8.74 (m, 1H), 8.41 - 8.35 (m,

3H), 8.25 (t, J = 13.2 Hz, 1H), 8.13 (d, J = 6.4 Hz, 1H), 7.99 (s, 1H), 7.90 – 7.78 (m, 4H), 7.71 (d, J = 8.8 Hz, 1H), 7.65 – 7.60 (m, 2H), 7.53 (dd, J = 8.0, 1.6 Hz, 1H), 7.48 – 7.37 (m, 3H), 6.94 (d, J = 8.8 Hz, 2H), 6.65 (d, J = 8.0 Hz, 2H), 6.54 (t, J = 12.6 Hz, 1H), 6.38 (d, J = 13.6 Hz, 1H), 6.28 – 6.22 (m, 1H), 6.08 (d, J = 13.6 Hz, 1H), 5.55 (s, 1H), 5.31 (q, J = 6.8 Hz, 1H), 5.18 (p, J = 7.2 Hz, 1H), 4.64 (s, 2H), 4.44 – 4.30 (m, 4H), 4.14 – 4.04 (m, 8H), 3.82 (t, J = 4.4 Hz, 2H), 3.68 – 3.63 (m, 3H), 3.60 – 3.43 (m, 8H), 3.39 – 3.31 (m, 4H), 3.17 (s, 3H), 2.56 (t, J = 7.2 Hz, 2H), 2.46 – 2.41 (m, 1H), 2.31 – 2.25 (m, 2H), 2.12 (s, 2H), 2.08 (s, 1H), 2.04 (s, 1H), 1.80 (s, 2H), 1.65 (d, J = 6.8 Hz, 2H), 1.60 (d, J = 7.2 Hz, 2H), 1.47 (s, 5H). ³¹P NMR (162 MHz, DMSO-*d*₆) δ -11.13 (d, J = 22.3 Hz, 1P), -11.90 (d, J = 22.6 Hz, 1P), -23.56 (t, J = 22.3 Hz, 1P).

实施例 8

linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。

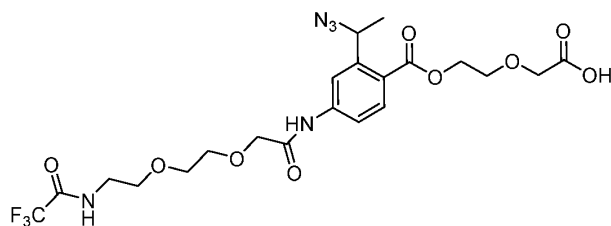


合成方法与实施例 7-1 类似。15mg 固体, HPLC 纯度 > 99%。

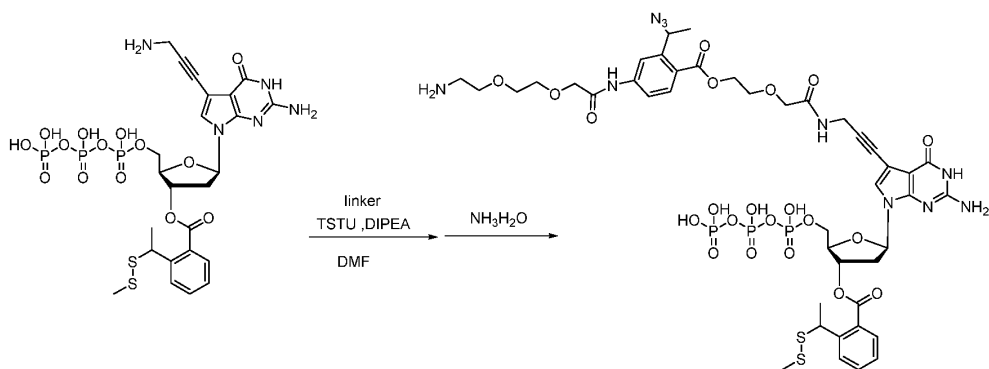
MS[ES(-)]*m/z* 1968.4. ¹H NMR (400 MHz, DMSO-*d*₆) δ 10.31 (d, *J* = 18.9 Hz, 1H), 9.00 (s, 1H), 8.23 (d, *J* = 8.9 Hz, 1H), 8.05 – 7.88 (m, 3H), 7.87 – 7.77 (m, 2H), 7.72 – 7.67 (m, 3H), 7.63 (dd, *J* = 5.2, 3.6 Hz, 1H), 7.57 – 7.43 (m, 1H), 7.09 (s, 1H), 6.99 (s, 3H), 6.79 (d, *J* = 5.3 Hz, 2H), 6.27 – 6.20 (m, 1H), 5.51 (s, 1H), 5.34 – 5.28 (m, 2H), 4.44 – 4.36 (m, 2H), 4.32 (d, *J* = 16.2 Hz, 2H), 4.13 (s, 2H), 4.09 (s, 2H), 4.03 (s, 2H), 3.84 (s, 3H), 3.82 – 3.79 (m, 2H), 3.67 – 3.64 (m, 2H), 3.42 – 3.38 (m, 2H), 3.19 – 3.14 (m, 2H), 3.13 – 3.08 (m, 2H), 2.89 (d, *J* = 3.8 Hz, 2H), 2.70 – 2.64 (m, 4H), 2.59 – 2.53 (m, 6H), 2.13 (s, 5H), 2.08 (d, *J* = 11.3 Hz, 2H), 1.97 (dt, *J* = 10.7, 7.0 Hz, 5H), 1.83 (dt, *J* = 12.2, 6.0 Hz, 4H), 1.71 (t, *J* = 7.2 Hz, 2H), 1.63 (d, *J* = 7.0 Hz, 3H), 1.60 (d, *J* = 6.9 Hz, 3H), 1.23 (s, 3H).

实施例 9

linker 的合成

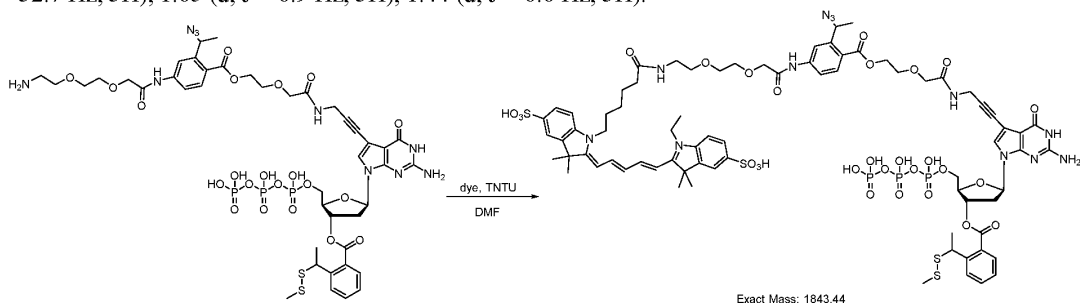


linker 8 的合成方法同实施例 7-1。MS[ES(-)], m/z 548.2. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 10.08 (s, 1H), 9.51 (s, 1H), 7.94 (d, $J = 2.0$ Hz, 1H), 7.86 (d, $J = 8.6$ Hz, 1H), 7.77 (dd, $J = 8.6, 2.0$ Hz, 1H), 5.70 (q, $J = 6.6$ Hz, 1H), 4.41 – 4.32 (m, 2H), 4.12 (s, 2H), 3.96 (s, 2H), 3.79 (t, $J = 4.6$ Hz, 2H), 3.67 (dd, $J = 5.7, 3.2$ Hz, 2H), 3.61 (dd, $J = 5.7, 3.1$ Hz, 2H), 3.55 (t, $J = 5.7$ Hz, 2H), 3.37 (dd, $J = 11.3, 5.6$ Hz, 2H), 2.07 (s, 1H), 1.44 (d, $J = 6.7$ Hz, 3H).



Exact Mass: 1204.22

合成方法与实施例 7-3 类似。MS[ES(-)], m/z 1203.2. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 11.18 (s, 1H), 8.28 (d, $J = 6.3$ Hz, 2H), 7.92 (d, $J = 8.5$ Hz, 1H), 7.87 (d, $J = 7.7$ Hz, 1H), 7.69 (d, $J = 7.1$ Hz, 1H), 7.65 – 7.58 (m, 2H), 7.43 (t, $J = 5.3$ Hz, 1H), 7.27 (d, $J = 4.3$ Hz, 1H), 6.47 – 6.32 (m, 2H), 6.32 – 6.24 (m, 1H), 5.70 (q, $J = 6.5$ Hz, 1H), 5.63 (s, 1H), 5.22 – 5.18 (m, 1H), 4.41 – 4.37 (m, 2H), 4.32 (t, $J = 4.9$ Hz, 1H), 4.28 (s, 2H), 4.26 (s, 1H), 4.11 (d, $J = 5.9$ Hz, 2H), 4.04 (s, 2H), 4.01 (s, 2H), 3.82 (s, 2H), 3.66 (d, $J = 13.7$ Hz, 2H), 3.63 – 3.59 (m, 4H), 2.97 (s, 2H), 2.10 (d, $J = 32.7$ Hz, 3H), 1.65 (d, $J = 6.9$ Hz, 3H), 1.44 (d, $J = 6.6$ Hz, 3H).



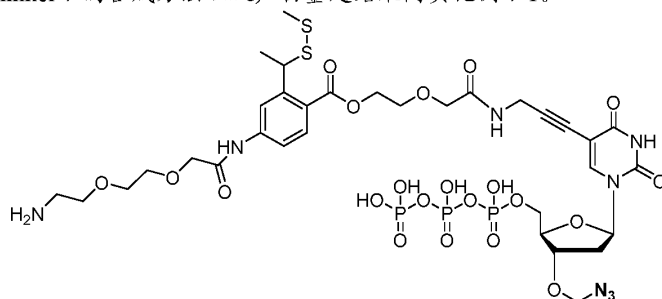
Exact Mass: 1843.44

合成方法与实施例 7-3 类似。20mg 固体, HPLC 纯度 > 99%。

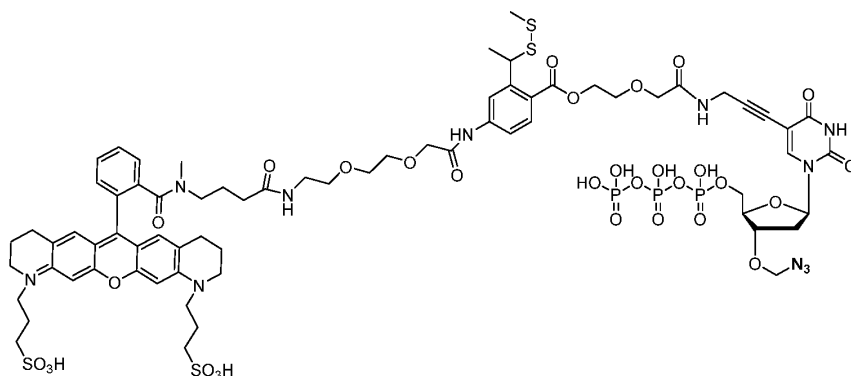
MS[ES(-)], m/z 1842.4. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 10.55 (s, 1H), 10.38 (s, 1H), 8.42 (s, 1H), 8.35 (t, $J = 13.1$ Hz, 2H), 8.03 (s, 1H), 7.88 (dd, $J = 10.9, 6.8$ Hz, 3H), 7.81 (s, 2H), 7.75 (d, $J = 8.8$ Hz, 1H), 7.68 – 7.58 (m, 4H), 7.47 – 7.40 (m, 1H), 7.35 – 7.25 (m, 3H), 6.59 (t, $J = 12.3$ Hz, 1H), 6.40 (s, 2H), 6.30 (dd, $J = 13.7, 7.2$ Hz, 3H), 5.69 (q, $J = 6.6$ Hz, 1H), 5.63 (s, 1H), 5.20 (q, $J = 6.8$ Hz, 1H), 4.44 – 4.36 (m, 2H), 4.32 (t, $J = 4.7$ Hz, 1H), 4.25 (t, $J = 4.4$ Hz, 1H), 4.16 – 4.09 (m, 6H), 4.06 (s, 2H), 4.00 (s, 2H), 3.84 – 3.79 (m, 2H), 3.65 (dd, $J = 5.4, 3.5$ Hz, 2H), 3.59 – 3.55 (m, 2H), 3.43 – 3.39 (m, 2H), 3.19 (dd, $J = 11.5, 5.8$ Hz, 2H), 2.45 (dd, $J = 11.6, 6.5$ Hz, 1H), 2.38 (dd, $J = 13.6, 5.4$ Hz, 1H), 2.10 (d, $J = 29.1$ Hz, 3H), 2.04 (dd, $J = 12.6, 6.9$ Hz, 2H), 1.68 (s, 12H), 1.65 (d, $J = 7.0$ Hz, 3H), 1.56 – 1.49 (m, 2H), 1.43 (d, $J = 6.6$ Hz, 3H), 1.32 (dd, $J = 14.2, 7.8$ Hz, 2H), 1.26 (d, $J = 7.0$ Hz, 3H), 1.23 (s, 2H).

实施例 10-1linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



合成方法与实施例 7-1 类似。MS[ES(-)],m/z 1047.2. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 11.06 (s, 1H), 8.81 (d, $J = 5.8$ Hz, 1H), 8.32 (s, 1H), 8.04 (s, 1H), 7.92 (d, $J = 8.6$ Hz, 1H), 7.65 (d, $J = 8.7$ Hz, 1H), 6.10 (t, $J = 7.0$ Hz, 1H), 5.36 (q, $J = 6.9$ Hz, 1H), 4.92 (d, $J = 8.9$ Hz, 1H), 4.81 (d, $J = 8.9$ Hz, 1H), 4.49 – 4.43 (m, 1H), 4.41 – 4.35 (m, 1H), 4.33 – 4.26 (m, 1H), 4.24 (s, 2H), 4.10 (dd, $J = 10.1, 3.4$ Hz, 2H), 4.04 (s, 2H), 4.00 (dd, $J = 11.0, 5.5$ Hz, 2H), 3.83 – 3.78 (m, 2H), 3.66 (d, $J = 4.8$ Hz, 2H), 3.64 – 3.57 (m, 6H), 2.99 – 2.91 (m, 2H), 2.37 – 2.27 (m, 2H), 2.15 (d, $J = 1.1$ Hz, 3H), 1.61 (d, $J = 6.9$ Hz, 3H).

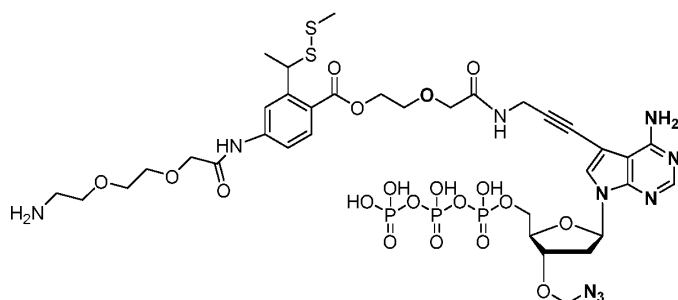


合成方法与实施例 7-1 类似。16mg 固体, HPLC 纯度 > 99%。

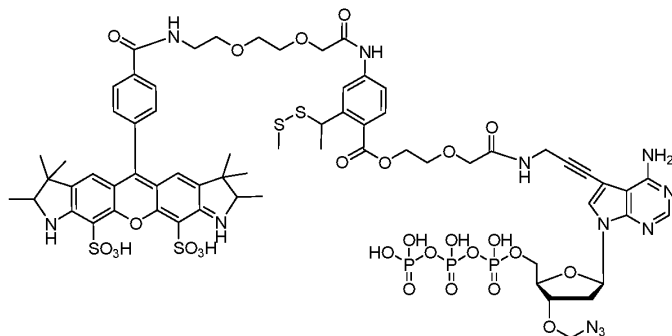
MS[ES(-)],m/z 1783.4. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.39 (d, $J = 18.2$ Hz, 1H), 9.00 (t, $J = 5.0$ Hz, 1H), 8.13 (s, 1H), 8.03 – 7.98 (m, 1H), 7.82 (d, $J = 8.6$ Hz, 2H), 7.75 – 7.66 (m, 3H), 7.66 – 7.60 (m, 1H), 7.58 – 7.41 (m, 1H), 6.99 (s, 2H), 6.79 (d, $J = 5.6$ Hz, 2H), 6.11 (t, $J = 6.9$ Hz, 1H), 5.30 (q, $J = 6.9$ Hz, 1H), 4.91 (d, $J = 8.9$ Hz, 1H), 4.83 (d, $J = 8.9$ Hz, 1H), 4.49 – 4.45 (m, 1H), 4.44 – 4.32 (m, 2H), 4.14 (s, 2H), 4.11 – 4.07 (m, 2H), 4.03 (s, 4H), 3.81 (t, $J = 4.4$ Hz, 2H), 3.70 (d, $J = 7.8$ Hz, 2H), 3.68 – 3.63 (m, 2H), 3.56 (dd, $J = 9.5, 5.3$ Hz, 6H), 3.40 (t, $J = 5.9$ Hz, 2H), 3.20 – 3.14 (m, 2H), 3.13 – 3.06 (m, 2H), 2.67 (s, 3H), 2.58 (t, $J = 7.1$ Hz, 4H), 2.40 – 2.26 (m, 2H), 2.13 (s, 3H), 2.02 – 1.91 (m, 4H), 1.84 (d, $J = 5.0$ Hz, 4H), 1.70 (t, $J = 7.2$ Hz, 1H), 1.60 (d, $J = 6.9$ Hz, 3H), 1.28 – 1.16 (m, 3H), 1.01 – 0.92 (m, 1H); ^{31}P NMR (162 MHz, $\text{DMSO}-d_6$) δ -11.29 (d, $J = 23.1$ Hz), -12.09 (d, $J = 23.0$ Hz), -23.70 (t, $J = 22.7$ Hz).

实施例 10-2linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



合成方法与实施例 7-2 类似。MS[ES(-)],m/z 1069.2. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 11.18 (d, $J = 6.1$ Hz, 1H), 8.37 (t, $J = 5.4$ Hz, 1H), 8.21 (dd, $J = 7.4, 1.7$ Hz, 1H), 8.09 (s, 1H), 7.84 (d, $J = 8.5$ Hz, 1H), 7.74 (d, $J = 2.4$ Hz, 1H), 7.68 – 7.60 (m, 1H), 6.50 (dd, $J = 8.5, 6.1$ Hz, 1H), 5.31 (q, $J = 6.9$ Hz, 1H), 4.93 (d, $J = 8.9$ Hz, 1H), 4.87 (d, $J = 8.9$ Hz, 1H), 4.60 (d, $J = 4.3$ Hz, 1H), 4.43 – 4.32 (m, 2H), 4.29 (s, 2H), 4.16 – 4.09 (m, 4H), 4.01 (s, 2H), 3.98 – 3.93 (m, 2H), 3.83 – 3.79 (m, 2H), 3.67 (d, $J = 3.0$ Hz, 2H), 3.62 (d, $J = 7.5$ Hz, 4H), 2.95 (t, $J = 4.9$ Hz, 2H), 2.69 – 2.62 (m, 1H), 2.39 (dd, $J = 13.0, 6.0$ Hz, 1H), 2.11 (d, $J = 5.6$ Hz, 3H), 1.59 (d, $J = 6.8$ Hz, 3H).

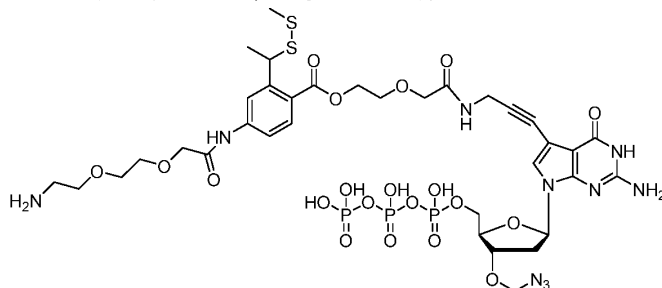


合成方法与实施例 7-2 类似。18mg 固体, HPLC 纯度 > 99%。MS[ES(-)],m/z 1678.3. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.52 (d, $J = 4.9$ Hz, 1H), 8.94 (t, $J = 5.0$ Hz, 1H), 8.80 (s, 2H), 8.48 (t, $J = 5.3$ Hz, 1H), 8.16 (d, $J = 8.4$ Hz, 2H), 8.08 (s, 1H), 8.04 (s, 1H), 7.82 – 7.74 (m, 2H), 7.70 (d, $J = 8.5$ Hz, 1H), 7.49 (d, $J = 7.9$ Hz, 2H), 6.71 (s, 2H), 6.49 (dd, $J = 8.7, 5.9$ Hz, 1H), 5.29 (q, $J = 7.0$ Hz, 1H), 4.92 (d, $J = 8.9$ Hz, 1H), 4.87 (d, $J = 8.9$ Hz, 1H), 4.60 (d, $J = 4.8$ Hz, 1H), 4.44 – 4.34 (m, 2H), 4.19 (s, 2H), 4.13 (d, $J = 5.4$ Hz, 2H), 4.00 (s, 2H), 3.97 – 3.92 (m, 2H), 3.82 – 3.77 (m, 4H), 3.74 – 3.70 (m, 2H), 3.68 – 3.63 (m, 4H), 3.53 (dd, $J = 10.6, 5.0$ Hz, 2H), 2.68 – 2.63 (m, 1H), 2.39 (dd, $J = 13.1, 6.1$ Hz, 1H), 2.12 (d, $J = 1.4$ Hz, 3H), 1.58 (d, $J = 6.9$ Hz, 3H), 1.16 (d, $J = 6.5$ Hz, 6H), 1.13 (s, 6H), 0.99 (d, $J = 2.0$ Hz, 6H); ^{31}P NMR (162 MHz, $\text{DMSO}-d_6$) δ -11.12 (d, $J = 22.8$ Hz), -12.12 (d, $J = 23.4$ Hz), -23.53 (t, $J = 22.7$ Hz).

实施例 10-3

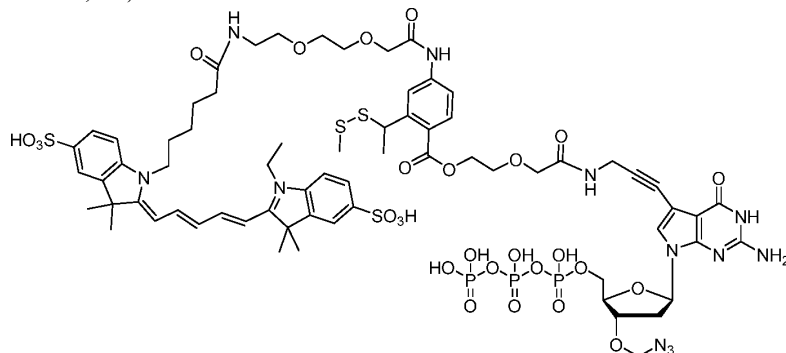
linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



合成方法与实施例 7-3 类似。MS[ES(-)],m/z 1085.2. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 11.24 (s, 1H), 8.25 (s, 2H), 7.86 (d, $J = 8.5$ Hz, 1H), 7.66 (d, $J = 8.7$ Hz, 1H), 7.18 (d, $J = 2.3$ Hz,

1H), 6.49 (s, 2H), 6.23 (dd, J = 8.9, 5.8 Hz, 1H), 5.34 (q, J = 6.9 Hz, 1H), 4.91 (d, J = 8.8 Hz, 1H), 4.85 (d, J = 8.8 Hz, 1H), 4.54 (d, J = 3.8 Hz, 1H), 4.42 - 4.31 (m, 2H), 4.29 (s, 2H), 4.12 - 4.06 (m, 2H), 4.00 (s, 2H), 3.92 (s, 2H), 3.84 - 3.78 (m, 2H), 3.66 (s, 2H), 3.61 (d, J = 4.9 Hz, 4H), 2.98 - 2.89 (m, 2H), 2.61 - 2.54 (m, 1H), 2.28 (dd, J = 13.1, 5.1 Hz, 1H), 2.13 (d, J = 1.6 Hz, 3H), 1.61 (d, J = 6.9 Hz, 3H).



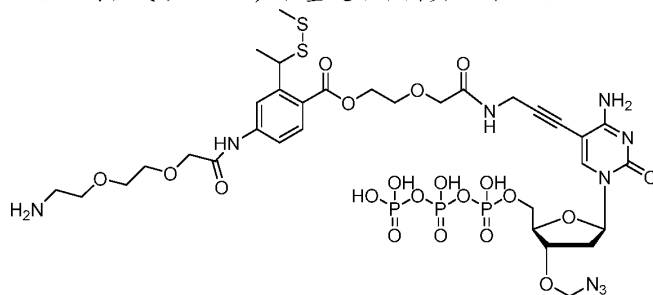
合成方法与实施例 7-3 类似。14mg 固体, HPLC 纯度 > 99%。

MS[ES(-)], m/z 1724.4. ¹H NMR (400 MHz, DMSO-*d*₆) δ 10.56 (s, 1H), 10.35 (s, 1H), 8.43 - 8.29 (m, 3H), 8.01 (d, J = 1.7 Hz, 1H), 7.89 (t, J = 5.5 Hz, 1H), 7.86 - 7.79 (m, 3H), 7.74 - 7.68 (m, 1H), 7.67 - 7.59 (m, 2H), 7.31 (dd, J = 8.3, 2.9 Hz, 2H), 7.21 (s, 1H), 6.59 (t, J = 12.3 Hz, 1H), 6.47 (s, 2H), 6.30 (dd, J = 13.8, 5.2 Hz, 2H), 6.22 (dd, J = 8.9, 5.8 Hz, 1H), 5.31 (q, J = 6.9 Hz, 1H), 4.91 (d, J = 8.9 Hz, 1H), 4.86 (d, J = 8.9 Hz, 1H), 4.53 (d, J = 4.5 Hz, 1H), 4.44 - 4.35 (m, 2H), 4.16 - 4.09 (m, 4H), 4.09 - 4.04 (m, 2H), 3.99 (s, 2H), 3.97 - 3.88 (m, 2H), 3.83 - 3.79 (m, 2H), 3.65 (dd, J = 5.7, 3.4 Hz, 2H), 3.59 - 3.55 (m, 2H), 3.41 (t, J = 5.8 Hz, 2H), 3.19 (dd, J = 11.6, 5.9 Hz, 2H), 2.62 (dd, J = 14.8, 5.7 Hz, 1H), 2.29 (dd, J = 13.3, 6.0 Hz, 1H), 2.13 (s, 3H), 2.06 (dd, J = 9.4, 5.0 Hz, 2H), 1.60 (d, J = 7.0 Hz, 3H), 1.52 (dt, J = 14.8, 7.3 Hz, 2H), 1.33 (d, J = 6.7 Hz, 2H), 1.25 (t, J = 7.1 Hz, 3H); ³¹P NMR (162 MHz, DMSO-*d*₆) δ -11.06 (d, J = 21.0 Hz), -12.40 (d, J = 23.7 Hz), -23.60 (t, J = 22.2 Hz).

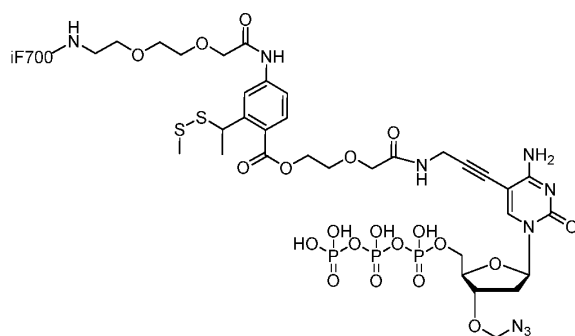
实施例 10-4

linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



合成方法与实施例 7-4 类似。MS[ES(-)], m/z 1047.2. ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.15 (s, 1H), 8.66 (d, J = 3.9 Hz, 1H), 8.29 (dd, J = 4.5, 1.8 Hz, 1H), 7.94 (s, 1H), 7.88 (d, J = 8.6 Hz, 1H), 7.78 (s, 1H), 7.69 - 7.62 (m, 1H), 6.90 (s, 1H), 6.11 (t, J = 6.3 Hz, 1H), 5.35 (q, J = 6.9 Hz, 1H), 4.91 (d, J = 8.9 Hz, 1H), 4.81 (d, J = 8.9 Hz, 1H), 4.46 - 4.36 (m, 2H), 4.26 (s, 2H), 4.11 (d, J = 3.7 Hz, 4H), 4.04 (s, 2H), 4.02 - 3.91 (m, 2H), 3.84 - 3.78 (m, 2H), 3.69 - 3.64 (m, 2H), 3.64 - 3.56 (m, 4H), 2.94 (t, J = 4.9 Hz, 2H), 2.31 (dd, J = 12.2, 4.5 Hz, 1H), 2.18 - 2.15 (m, 1H), 2.14 (s, 3H), 1.61 (d, J = 6.9 Hz, 3H).



合成方法与实施例 7-4 类似。20mg 固体，HPLC 纯度 > 99%。

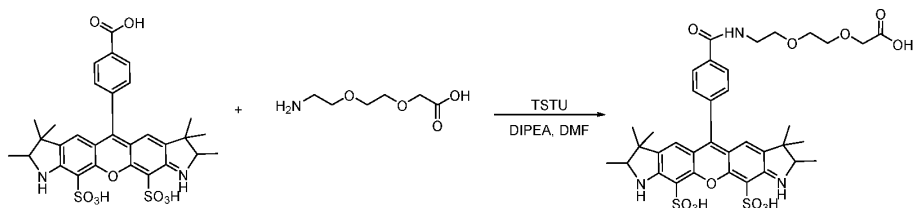
MS[ES(-)], m/z 1807.2. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 10.28 (s, 1H), 8.89 (s, 1H), 8.76 (d, J = 5.9 Hz, 2H), 8.39 (d, J = 18.8 Hz, 3H), 8.25 (t, J = 12.6 Hz, 1H), 7.99 (d, J = 9.1 Hz, 2H), 7.87 (s, 1H), 7.82 (d, J = 8.6 Hz, 1H), 7.78 (s, 1H), 7.71 (d, J = 9.7 Hz, 1H), 7.56 – 7.51 (m, 1H), 7.47 (d, J = 8.2 Hz, 2H), 6.99 – 6.87 (m, 2H), 6.65 (d, J = 8.2 Hz, 2H), 6.54 (t, J = 12.6 Hz, 1H), 6.38 (d, J = 12.7 Hz, 1H), 6.10 (t, J = 6.0 Hz, 1H), 5.30 (q, J = 6.9 Hz, 1H), 4.90 (d, J = 8.9 Hz, 1H), 4.82 (d, J = 8.9 Hz, 1H), 4.64 (s, 1H), 4.44 – 4.34 (m, 2H), 4.11 (d, J = 9.2 Hz, 4H), 4.02 (s, 2H), 3.81 (t, J = 4.3 Hz, 2H), 3.66 – 3.62 (m, 2H), 3.59 – 3.52 (m, 4H), 3.52 – 3.46 (m, 2H), 3.35 (dd, J = 11.9, 5.9 Hz, 2H), 3.17 (s, 2H), 2.55 (d, J = 7.2 Hz, 1H), 2.35 – 2.22 (m, 3H), 2.21 – 2.15 (m, 1H), 2.12 (s, 3H), 1.80 (s, 2H), 1.59 (d, J = 6.9 Hz, 3H), 1.47 (s, 5H); ^{31}P NMR (162 MHz, $\text{DMSO-}d_6$) δ -11.10 (d, J = 21.6 Hz), -12.24 (d, J = 23.1 Hz), -23.61 (t, J = 22.4 Hz).

实施例 11

(1) linker 的合成

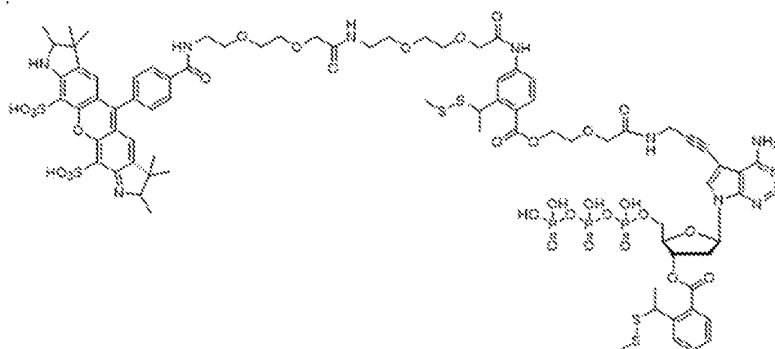
linker 7 的合成方法以及产物鉴定结果同实施例 7-1。

(2) 染料的合成



向 50 mL 干燥的圆底烧瓶中依次加入染料 AF532 (63 mg, 0.1 mmol, 1.0 eq), TSTU (1.5 eq, 45.2 mg), 加毕, 加入 3 ml DMF 溶解, 最后加入 DIPEA (2.5 eq, 52 μL), 然后置于 45 $^{\circ}\text{C}$ 油浴搅拌 3 h, 加入 PEG linker 2-(2-(2-aminoethoxy)ethoxy)acetic acid (40 mg, 2.0 eq), 45 $^{\circ}\text{C}$ 反应过夜, 加入水淬灭反应。使用制备液相纯化得目标产物 (65.6 mg, 85% 产率), 产物通过 MS 鉴定。MS[ES(-)], m/z 771.2.

(3) 终产物的合成



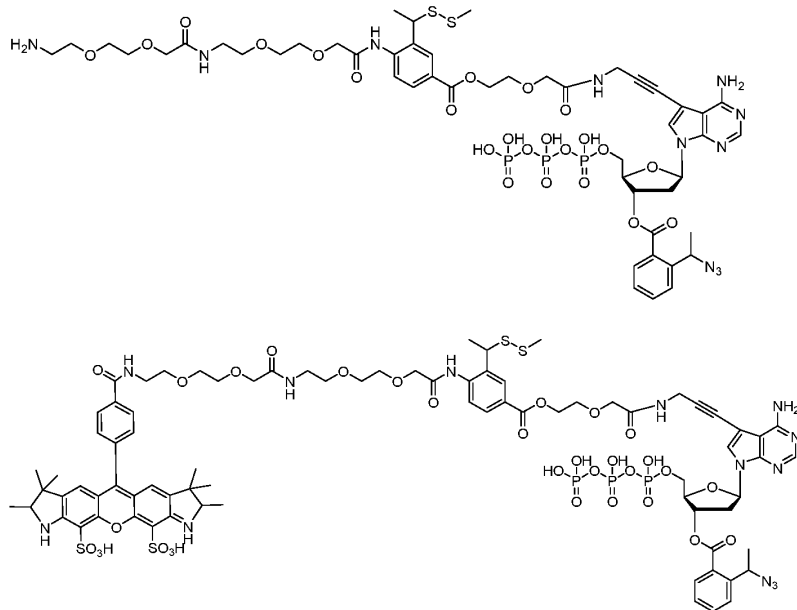
合成方法与实施例 7-2 类似，其中，将染料替换为步骤 (2) 获得的染料。21mg 固体，HPLC 纯度 > 99%。

MS[ES(-)], m/z 1978.4. ^1H NMR (400 MHz, DMSO- d_6) δ 10.46 (s, 1H), 8.88 (t, J = 5.2 Hz, 1H), 8.80 (s, 2H), 8.52 (t, J = 5.1 Hz, 1H), 8.14 (d, J = 8.4 Hz, 2H), 8.06 (d, J = 6.6 Hz, 1H), 8.04 (s, 1H), 7.90 (dd, J = 7.7, 3.0 Hz, 1H), 7.84 (dd, J = 11.5, 6.6 Hz, 2H), 7.74 (t, J = 5.7 Hz, 1H), 7.69 (d, J = 8.7 Hz, 1H), 7.64 (d, J = 3.8 Hz, 2H), 7.49 (d, J = 8.0 Hz, 2H), 7.47 - 7.41 (m, 1H), 6.70 (s, 2H), 6.66 - 6.56 (m, 1H), 5.70 (d, J = 4.2 Hz, 1H), 5.30 (q, J = 7.0 Hz, 1H), 5.21 (qd, J = 6.9, 2.1 Hz, 1H), 4.45 - 4.38 (m, 2H), 4.36 - 4.28 (m, 2H), 4.14 (s, 4H), 4.09 - 4.05 (m, 2H), 4.02 (s, 2H), 3.89 (s, 2H), 3.83 - 3.76 (m, 4H), 3.65 - 3.63 (m, 2H), 3.61 (s, 2H), 3.58 - 3.56 (m, 2H), 3.53 - 3.49 (m, 2H), 3.48 - 3.44 (m, 2H), 3.30 - 3.26 (m, 2H), 2.12 (d, J = 2.0 Hz, 3H), 2.08 (d, J = 19.2 Hz, 3H), 1.66 (d, J = 7.0 Hz, 3H), 1.60 (d, J = 6.9 Hz, 3H), 1.16 (d, J = 6.6 Hz, 6H), 1.13 (s, 6H), 1.00 (s, 6H).

实施例 12

linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



合成方法与实施例 11 类似。

实施例 13

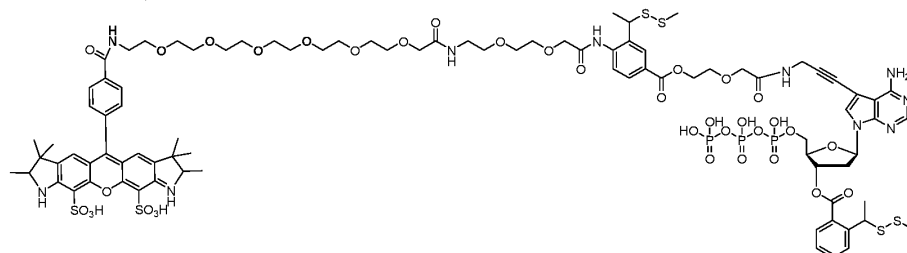
(1) linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。

(2) 染料的合成

染料合成方法与实施例 11 类似。

(3) 终产物的合成

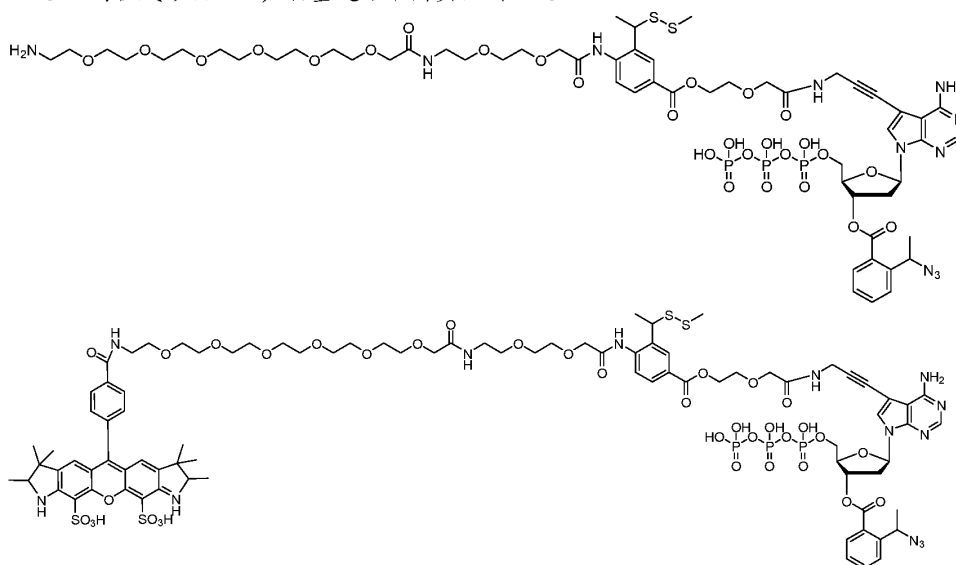


合成方法与实施例 11 类似，其中，将染料替换为步骤 (2) 获得的染料。23mg 固体，HPLC 纯度 > 99%。

MS[ES(-)], m/z 2154.5. ^1H NMR (400 MHz, DMSO- d_6) δ 8.94 (s, 1H), 8.88 – 8.83 (m, 1H), 8.80 (s, 2H), 8.56 (s, 1H), 8.13 (d, J = 8.4 Hz, 2H), 8.10 – 8.06 (m, 2H), 7.94 – 7.83 (m, 4H), 7.82 (d, J = 4.5 Hz, 1H), 7.63 (d, J = 3.9 Hz, 2H), 7.51 (d, J = 8.2 Hz, 2H), 7.47 – 7.43 (m, 1H), 6.71 (s, 2H), 6.65 – 6.57 (m, 1H), 5.67 (s, 1H), 5.26 – 5.11 (m, 2H), 4.53 – 4.41 (m, 2H), 4.40 – 4.28 (m, 2H), 4.15 (d, J = 5.2 Hz, 2H), 4.12 – 4.05 (m, 2H), 4.02 (s, 4H), 3.83 (dd, J = 8.2, 4.3 Hz, 4H), 3.80 – 3.75 (m, 2H), 3.59 (s, 4H), 3.56 (d, J = 5.9 Hz, 4H), 2.10 (d, J = 7.5 Hz, 5H), 2.06 (s, 2H), 1.66 (d, J = 6.3 Hz, 6H), 1.17 (s, 6H), 0.99 (s, 6H).

实施例 14linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。



染料以及终产物的合成方法与实施例 11 类似。

实施例 15-1

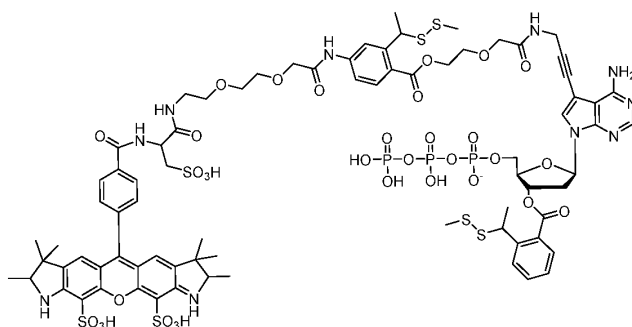
(1) linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。

(2) 染料的合成

染料合成方法与实施例 15-2 类似，使用染料 AF532 替代 CY5。

(3) 终产物的合成



合成方法与实施例 7-2 类似，其中，将染料替换为步骤 (2) 获得的染料。18mg 固体，HPLC 纯度 > 99%。

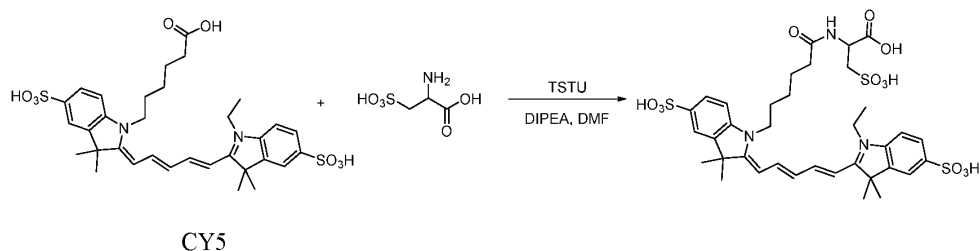
MS[ES(-)], m/z 1983.3. ^1H NMR (400 MHz, DMSO- d_6) δ 10.42 (d, J = 144.3 Hz, 1H), 8.97 - 8.80 (m, 2H), 8.50 (d, J = 30.7 Hz, 1H), 8.14 - 8.05 (m, 3H), 7.91 (d, J = 7.8 Hz, 1H), 7.84 (dd, J = 12.3, 8.6 Hz, 2H), 7.75 - 7.60 (m, 3H), 7.55 (d, J = 8.3 Hz, 1H), 7.49 - 7.42 (m, 1H), 7.39 - 7.25 (m, 2H), 6.97 - 6.89 (m, 1H), 6.71 (s, 1H), 6.64 - 6.58 (m, 1H), 5.70 - 5.64 (m, 1H), 5.34 - 5.26 (m, 1H), 5.22 (qd, J = 7.0, 1.8 Hz, 1H), 4.58 - 4.30 (m, 4H), 4.19 (s, 1H), 4.14 (s, 2H), 4.02 (s, 2H), 3.84 - 3.79 (m, 2H), 3.68 - 3.65 (m, 2H), 3.60 (d, J = 4.6 Hz, 2H), 3.27 (d, J = 4.5 Hz, 2H), 3.15 (d, J = 8.9 Hz, 2H), 2.74 (s, 6H), 2.12 (s, 3H), 2.09 (d, J = 19.8 Hz, 3H), 1.67 (d, J = 6.9 Hz, 3H), 1.60 (dd, J = 6.6, 2.4 Hz, 3H), 1.23 (s, 2H), 1.18 - 1.13 (m, 6H), 1.00 (d, J = 1.8 Hz, 3H).

实施例 15-2

(1) linker 的合成

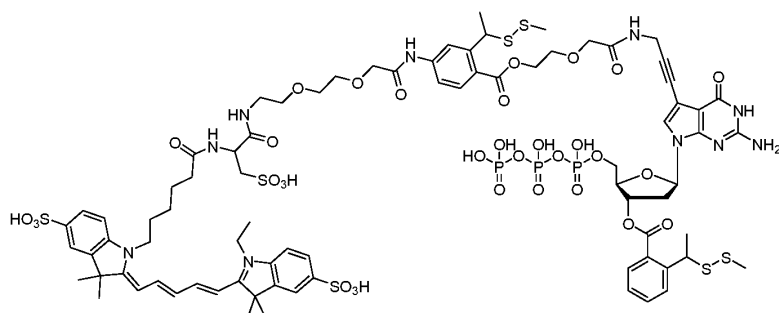
linker 7 的合成方法以及产物鉴定结果同实施例 7-1。

(2) 染料的合成



向 50 mL 干燥的圆底烧瓶中依次加入染料 CY5 (250 mg, 0.38 mmol, 1.0 eq), TSTU (1.2 eq, 110 mg), 加毕, 加入 5 mL DMF 溶解, 最后加入 DIPEA (3.0 eq, 200 μL), 然后室温搅拌 3h, 随后, 加入带磺酸的片段 (114 mg, 2.0 eq), 室温搅拌 6h, 反应完全后, 加入水淬灭反应。使用制备液相纯化获得目标产物 (325 mg, 产率 89%), 产物通过 MS 鉴定。MS[ES(-)], m/z 807.2.

(3) 终产物的合成



合成方法与实施例 7-3 类似，其中，将染料替换为步骤 (2) 获得的染料。25mg 固体，HPLC 纯度 >99%。

MS[ES(-)], m/z 2030.4. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.52 (s, 1H), 10.28 (s, 1H), 8.43 (d, $J = 2.5$ Hz, 1H), 8.34 (t, $J = 13.0$ Hz, 2H), 8.00 (s, 1H), 7.97 (d, $J = 6.3$ Hz, 1H), 7.87 (d, $J = 7.6$ Hz, 1H), 7.86–7.81 (m, 2H), 7.80 (s, 2H), 7.73 (dd, $J = 8.6, 1.9$ Hz, 1H), 7.67–7.58 (m, 4H), 7.47–7.41 (m, 1H), 7.34 (d, $J = 8.4$ Hz, 1H), 7.30 (dd, $J = 5.8, 2.4$ Hz, 2H), 6.70 (t, $J = 12.4$ Hz, 1H), 6.39 (s, 2H), 6.37–6.24 (m, 3H), 5.64 (t, $J = 4.0$ Hz, 1H), 5.31 (q, $J = 6.8$ Hz, 1H), 5.24–5.16 (m, 1H), 4.42–4.34 (m, 3H), 4.33–4.23 (m, 2H), 4.13 (s, 2H), 4.12 (d, $J = 5.4$ Hz, 2H), 4.08–4.02 (m, 2H), 4.00 (s, 2H), 3.83–3.79 (m, 2H), 3.66–3.62 (m, 2H), 3.58–3.55 (m, 2H), 3.43–3.39 (m, 2H), 3.23–3.16 (m, 2H), 2.76 (dd, $J = 13.9, 7.9$ Hz, 2H), 2.41–2.32 (m, 1H), 2.13 (d, $J = 1.9$ Hz, 5H), 2.08 (s, 1H), 1.66 (d, $J = 7.1$ Hz, 3H), 1.61 (d, $J = 6.9$ Hz, 3H), 1.59–1.50 (m, 2H), 1.40 (dd, $J = 16.1, 9.0$ Hz, 2H), 1.25 (t, $J = 7.1$ Hz, 4H).

实施例 16

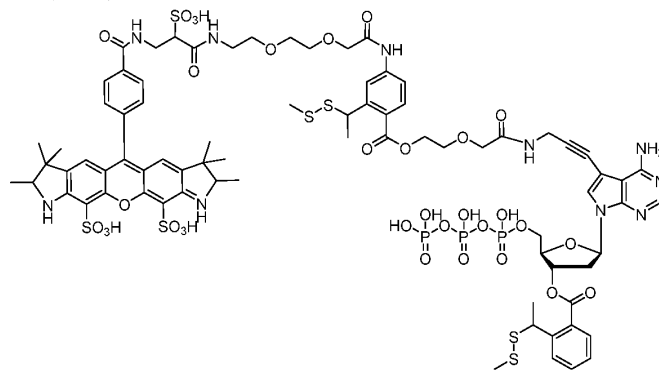
(1) linker 的合成

linker 7 的合成方法以及产物鉴定结果同实施例 7-1。

(2) 染料的合成

染料合成方法与实施例 15-2 类似，使用染料 AF532 替代 CY5。

(3) 终产物的合成

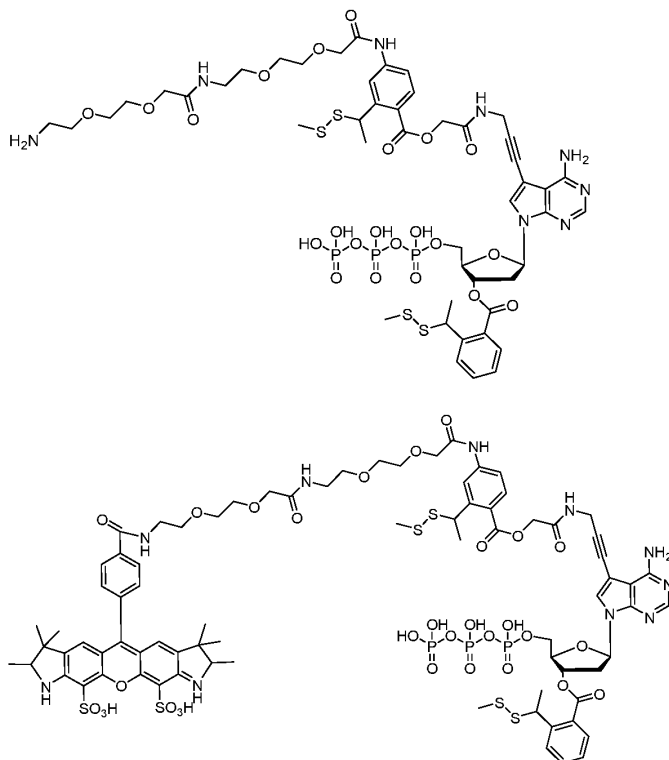


合成方法与实施例 7-2 类似，其中，将染料替换为步骤 (2) 获得的染料。22mg 固体，HPLC 纯度 >99%。

MS[ES(-)], m/z 1984.3. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 10.28 (s, 1H), 8.81 (s, 3H), 8.51 (s, 1H), 8.13 (d, $J = 8.4$ Hz, 2H), 8.08 (d, $J = 6.2$ Hz, 1H), 8.00 (s, 1H), 7.96–7.88 (m, 2H), 7.82 (dd, $J = 12.3, 6.3$ Hz, 2H), 7.67 (dd, $J = 12.5, 6.7$ Hz, 2H), 7.64 (s, 2H), 7.52 (d, $J = 8.0$ Hz, 2H), 7.48–7.41 (m, 1H), 6.72 (s, 2H), 6.64–6.57 (m, 1H), 5.67–5.64 (m, 1H), 5.30 (q, $J = 6.4$ Hz, 1H), 5.21 (t, $J = 6.1$ Hz, 1H), 4.47–4.27 (m, 4H), 4.13 (s, 4H), 4.11–4.06 (m, 2H), 4.02 (s, 2H), 3.89 (s, 2H), 3.80 (dd, $J = 12.6, 5.7$ Hz, 4H), 3.68–3.64 (m, 2H), 3.62 (s, 2H), 3.57 (dd, $J = 11.3, 5.1$ Hz, 2H), 3.49 (s, 4H), 3.47 (d, $J = 2.6$ Hz, 4H), 3.25 (d, $J = 5.7$ Hz, 2H), 3.21 (dd, $J = 11.6, 5.8$ Hz, 2H), 2.69–2.65 (m, 1H), 2.57 (dd, $J = 5.7, 2.7$ Hz, 1H), 2.31 (t, $J = 6.5$ Hz, 3H), 2.12 (s, 2H), 2.11 (d, $J = 3.4$ Hz, 3H), 2.07 (s, 2H), 2.06 (s, 2H), 1.66 (d, $J = 7.1$ Hz, 3H), 1.60 (d, $J = 7.1$ Hz, 3H), 1.17 (d, $J = 6.6$ Hz, 6H), 1.14 (s, 6H), 1.01 (d, $J = 1.9$ Hz, 6H).

实施例 17linker 的合成

linker 9 的合成方法与实施例 7-1 类似。

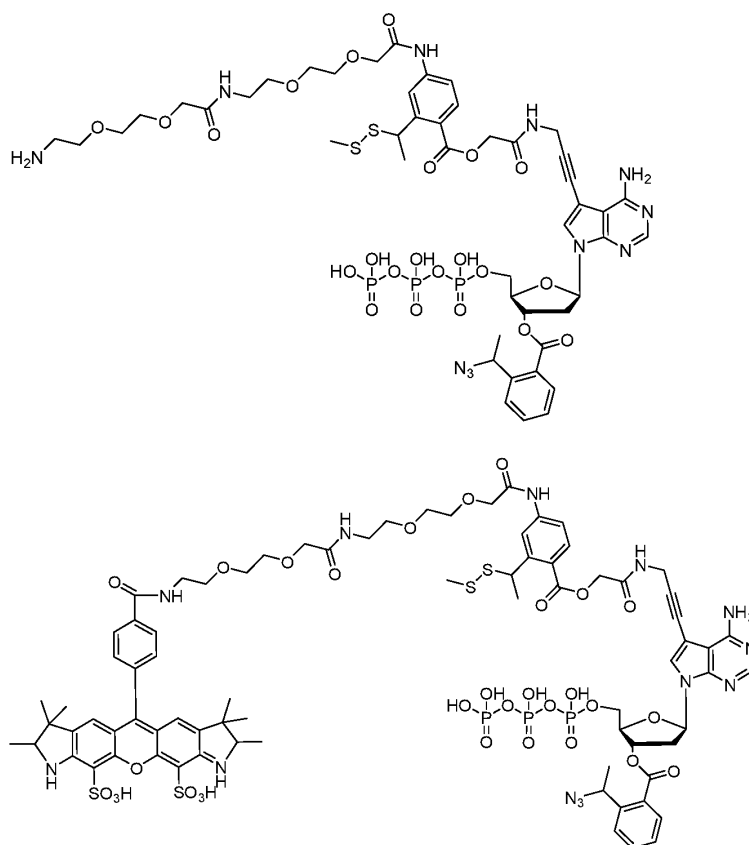


合成方法与实施例 7-2 类似，25mg 固体，HPLC 纯度 > 99%。

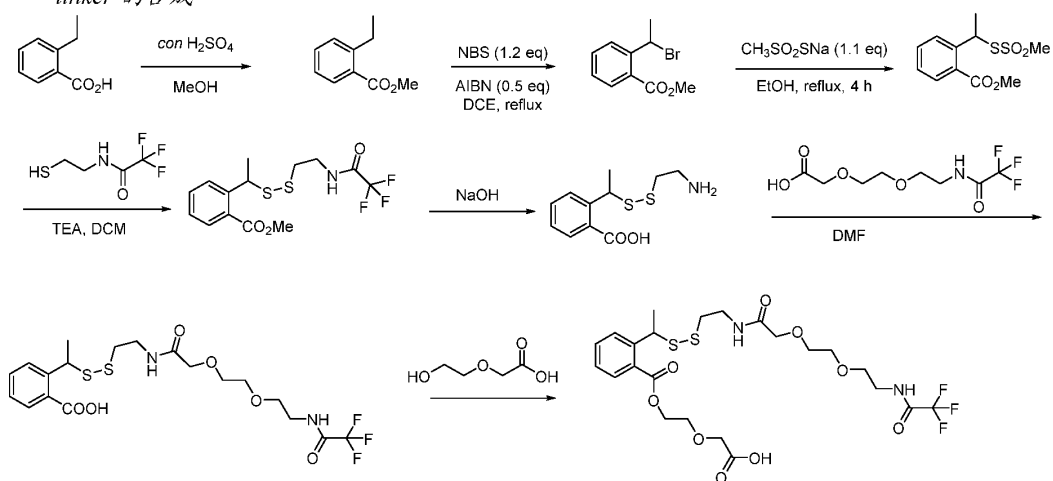
MS[ES(-)], m/z 1934.3. ¹H NMR (400 MHz, DMSO-*d*₆) δ 10.21 (d, J = 3.5 Hz, 1H), 9.05 (d, J = 4.9 Hz, 1H), 8.85 (t, J = 5.5 Hz, 1H), 8.81 (s, 2H), 8.13 (d, J = 8.4 Hz, 2H), 8.09 (d, J = 6.6 Hz, 1H), 7.99 (s, 1H), 7.95 – 7.88 (m, 3H), 7.72 (dd, J = 13.2, 7.5 Hz, 2H), 7.66 – 7.59 (m, 2H), 7.51 – 7.43 (m, 3H), 6.71 (s, 2H), 6.67 – 6.58 (m, 1H), 5.70 (d, J = 3.4 Hz, 1H), 5.30 (q, J = 6.6 Hz, 1H), 5.25 – 5.17 (m, 1H), 4.82 – 4.70 (m, 2H), 4.38 – 4.28 (m, 2H), 4.21 (d, J = 4.3 Hz, 3H), 4.12 (s, 2H), 4.09 – 4.01 (m, 2H), 3.89 (s, 2H), 3.82 – 3.77 (m, 2H), 3.64 (dd, J = 4.7, 2.1 Hz, 2H), 3.61 (s, 2H), 3.58 (dd, J = 5.4, 2.9 Hz, 2H), 3.53 – 3.49 (m, 2H), 3.49 – 3.43 (m, 2H), 3.29 (dd, J = 11.8, 5.9 Hz, 2H), 2.62 – 2.52 (m, 2H), 2.16 (s, 3H), 2.08 (d, J = 17.6 Hz, 3H), 1.66 (d, J = 7.0 Hz, 3H), 1.59 (d, J = 6.9 Hz, 3H), 1.16 (d, J = 6.5 Hz, 6H), 1.14 (d, J = 1.2 Hz, 6H), 1.00 (d, J = 2.1 Hz, 6H).

实施例 18linker 的合成

linker 9 的合成方法与实施例 7-1 类似。



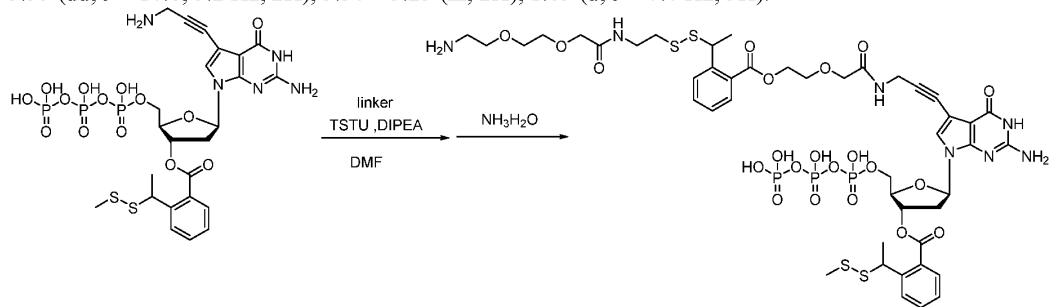
实施例 19 linker 的合成



linker 10

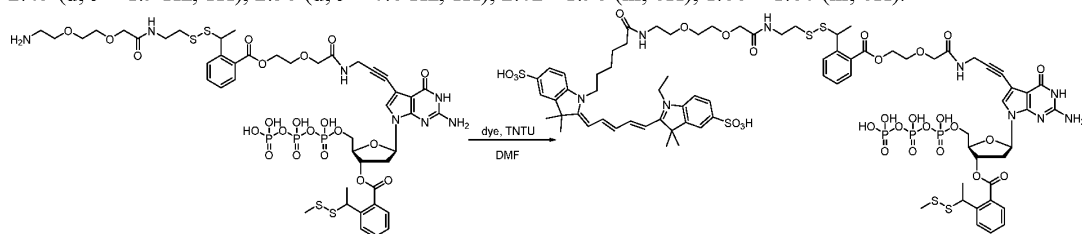
乙基苯甲酸在酸的条件下酯化，然后利用 NBS 溴代得到溴代产物。随后利用硫酸盐制备得到中间体，再利用硫醇胺制备二硫键化合物，水解后与 PEG 片段缩合得到 linker 10。
 MS[ES(-)], m/z 599.1. ¹H NMR (500 MHz, DMSO-*d*₆) δ 9.59 (s, 1H), 7.80 (dd, J = 14.3, 6.9 Hz, 2H), 7.59 (q, J = 8.1 Hz, 2H), 7.42 – 7.36 (m, 1H), 5.13 (q, J = 6.9 Hz, 1H), 4.44 – 4.33 (m, 2H), 3.89 (s, 2H), 3.85 (s, 2H), 3.81 – 3.75 (m, 2H), 3.61 (s, 2H), 3.55 (s, 4H), 3.51 (t, J = 5.7 Hz, 2H),

3.35 (dd, $J = 10.6, 5.2$ Hz, 2H), 3.30–3.23 (m, 2H), 1.63 (d, $J = 7.0$ Hz, 3H).



Exact Mass: 1255.19

合成方法与实施例 7-3 类似。MS[ES(-)], m/z 1254.2. ^1H NMR (500 MHz, DMSO- d_6) δ 8.54 (s, 1H), 7.95 (s, 1H), 7.89–7.83 (m, 1H), 7.76 (d, J = 5.1 Hz, 1H), 7.63 (d, J = 6.0 Hz, 2H), 7.58 (d, J = 6.7 Hz, 2H), 7.46–7.37 (m, 2H), 7.30–7.18 (m, 1H), 6.69–6.63 (m, 1H), 6.35 (s, 2H), 5.56 (dd, J = 6.2, 5.2 Hz, 1H), 5.32 (t, J = 4.8 Hz, 1H), 5.20 (dd, J = 9.6, 4.7 Hz, 1H), 5.14–5.07 (m, 1H), 4.44 (s, 2H), 4.37–4.25 (m, 2H), 4.12 (dd, J = 3.2, 2.3 Hz, 2H), 4.01 (s, 4H), 3.89–3.84 (m, 2H), 3.83 (dd, J = 6.8, 4.4 Hz, 2H), 3.63 (d, J = 7.7 Hz, 2H), 3.56 (s, 4H), 3.45–3.42 (m, 2H), 2.45 (d, J = 1.9 Hz, 1H), 2.36 (d, J = 0.6 Hz, 1H), 2.02–1.96 (m, 6H), 1.66–1.60 (m, 6H).



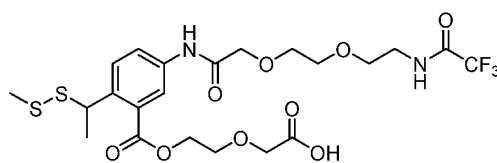
Exact Mass: 1894.41

合成方法与实施例 7-3 类似。16mg 固体, HPLC 纯度 > 99%。

MS[ES(-)], m/z 1893.4. ¹H NMR (500 MHz, DMSO-*d*₆) δ 10.53 (s, 1H), 8.46 – 8.30 (m, 2H), 7.87 (d, J = 6.8 Hz, 2H), 7.81 (s, 2H), 7.79 (d, J = 7.8 Hz, 1H), 7.65 - 7.62 (m, 3H), 7.58 (q, J = 7.8 Hz, 2H), 7.46 – 7.36 (m, 2H), 7.31 (d, J = 8.2 Hz, 2H), 7.29 (s, 1H), 6.59 (t, J = 12.3 Hz, 1H), 6.40 (s, 2H), 6.34 – 6.20 (m, 3H), 5.63 (s, 1H), 5.22 - 5.18 (m, 1H), 5.13 (q, J = 6.9 Hz, 1H), 4.49 - 4.44(m, 1H), 4.41 - 4.37 (m, 1H), 4.32 (s, 1H), 4.26 (s, 1H), 4.12 (d, J = 5.0 Hz, 2H), 4.07 (t, J = 8.0 Hz, 2H), 4.00 (s, 2H), 3.85 (s, 2H), 3.82 (t, J = 4.0 Hz, , 3H), 3.53 (s, 4H), 3.52 (s, 2H), 3.40 – 3.35 (m, 2H), 3.29 – 3.24 (m, 2H), 3.17 (d, J = 5.7 Hz, 2H), 2.13 (s, 1H), 2.06 (d, J = 4.0 Hz, 1H), 1.67 (s, 12H), 1.65 (d, J = 6.9 Hz, 3H), 1.62 (d, J = 6.9 Hz, 3H), 1.56 – 1.49 (m, 2H), 1.35 - 1.29 (m, 2H), 1.26 - 1.23 (m, 6H).

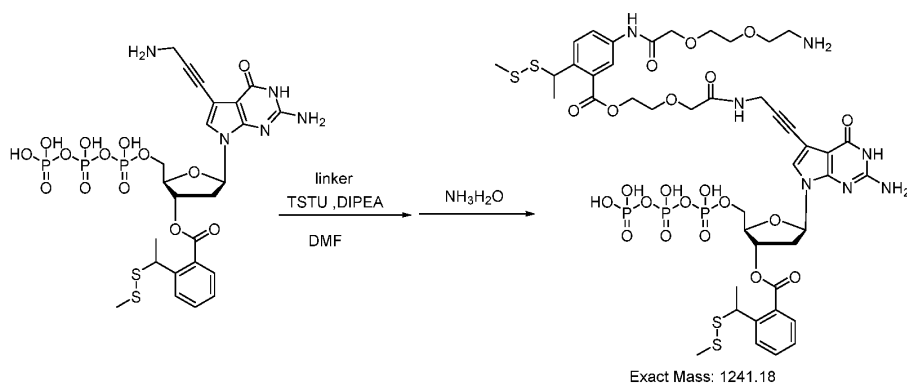
实施例 20-1

linker 的合成

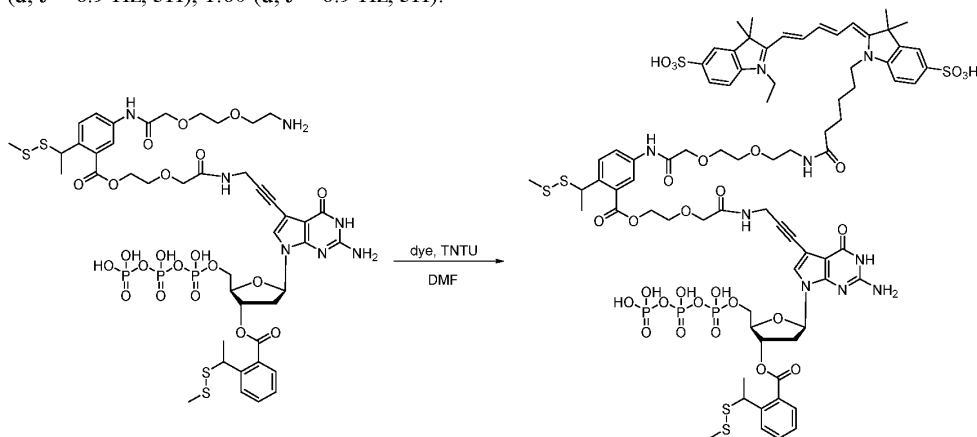


linker 11

linker 11 的合成方法与实施例 7-1 类似。MS[ES(-)], m/z 585.1. ¹H NMR (500 MHz, DMSO-*d*₆) δ 9.97 (s, 1H), 9.59 (s, 1H), 8.04 (d, J = 2.1 Hz, 1H), 7.89 (dd, J = 8.6, 2.1 Hz, 1H), 7.55 (d, J = 8.6 Hz, 1H), 5.13 (q, J = 6.9 Hz, 1H), 4.44 - 4.34 (m, 2H), 4.09 (s, 2H), 3.87 (s, 2H), 3.78 (s, 2H), 3.66 (dd, J = 5.7, 3.1 Hz, 2H), 3.61 (dd, J = 5.6, 3.1 Hz, 2H), 3.55 (t, J = 5.7 Hz, 2H), 3.37 (dd, J = 11.0, 5.4 Hz, 2H), 2.10 (s, 3H), 1.62 (d, J = 7.0 Hz, 3H).



合成方法与实施例 7-3 类似。MS[ES(-)], m/z 1240.2. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 10.83 (s, 1H), 8.44 – 8.36 (m, 1H), 8.25 (s, 1H), 8.03 – 7.91 (m, 1H), 7.87 (d, J = 7.8 Hz, 1H), 7.66 – 7.58 (m, 2H), 7.52 (d, J = 8.7 Hz, 1H), 7.48 – 7.40 (m, 1H), 7.33 (d, J = 8.2 Hz, 1H), 7.30 – 7.23 (m, 1H), 7.07 (d, J = 8.3 Hz, 1H), 6.41 – 6.37 (m, 2H), 6.35 – 6.27 (m, 1H), 5.64 (s, 1H), 5.22 – 5.16 (m, 2H), 4.51 – 4.41 (m, 1H), 4.41 – 4.35 (m, 1H), 4.34 (d, J = 4.7 Hz, 1H), 4.26 (s, 1H), 4.25 (s, 2H), 4.15 – 4.10 (m, 2H), 4.05 (s, 4H), 3.82 (s, 2H), 3.68 – 3.58 (m, 6H), 3.55 – 3.51 (m, 2H), 2.63 (t, J = 7.1 Hz, 1H), 2.37 (d, J = 13.0 Hz, 1H), 2.11 (d, J = 18.7 Hz, 3H), 2.08 (s, 3H), 1.66 (d, J = 6.9 Hz, 3H), 1.60 (d, J = 6.9 Hz, 3H).

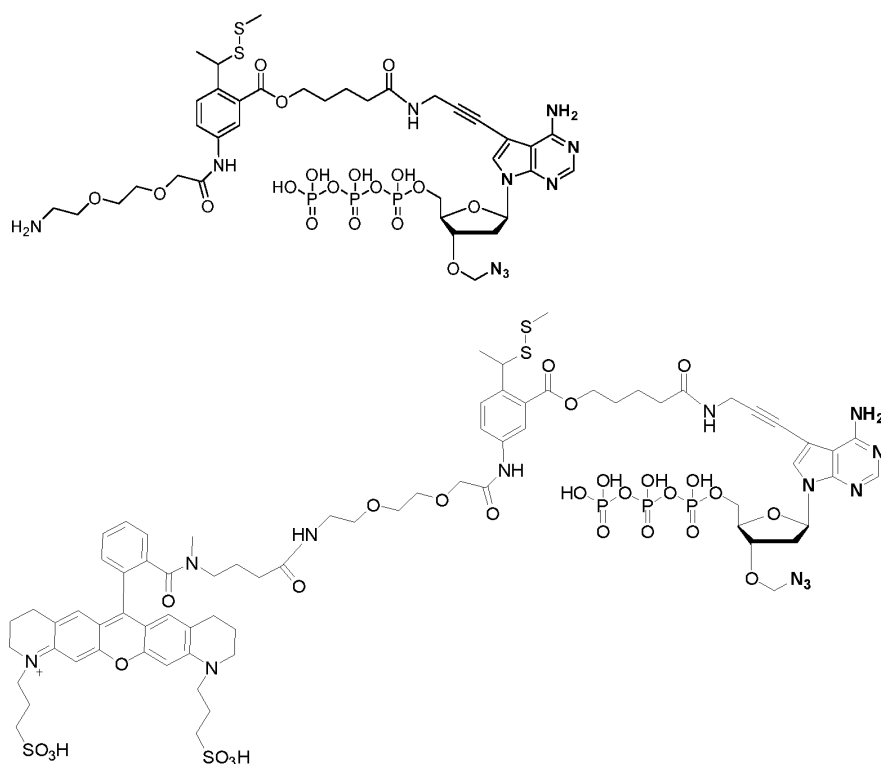


合成方法与实施例 7-3 类似。18mg 固体, HPLC 纯度 > 99%。
MS[ES(-)], m/z 1779.4. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 10.51 (s, 1H), 10.21 (s, 1H), 8.43 – 8.29 (m, 3H), 8.12 (s, 1H), 7.91 (dd, J = 10.0, 3.2 Hz, 2H), 7.87 (d, J = 7.4 Hz, 1H), 7.81 (s, 2H), 7.64 – 7.62 (m, 4H), 7.53 (d, J = 8.6 Hz, 1H), 7.44 (td, J = 7.2, 1.8 Hz, 1H), 7.35 – 7.28 (m, 3H), 7.27 (d, J = 4.8 Hz, 1H), 7.07 (d, J = 8.5 Hz, 1H), 6.59 (t, J = 12.3 Hz, 1H), 6.37 (s, 1H), 6.31 (dd, J = 14.2, 6.9 Hz, 3H), 5.63 (s, 1H), 5.23 – 5.16 (m, 1H), 5.13 (q, J = 7.0 Hz, 1H), 4.48 – 4.39 (m, 2H), 4.33 (t, J = 4.2 Hz, 1H), 4.26 (t, J = 3.5 Hz, 1H), 4.13 (s, 2H), 4.11 (s, 2H), 4.06 (s, 2H), 4.00 (s, 2H), 3.83 – 3.80 (m, 2H), 3.65 – 3.63 (m, 2H), 3.56 (dd, J = 7.8, 3.4 Hz, 2H), 3.53 (d, J = 7.3 Hz, 2H), 3.40 (d, J = 5.4 Hz, 2H), 3.19 (dd, J = 11.3, 5.7 Hz, 2H), 2.63 (t, J = 5.4 Hz, 1H), 2.41 – 2.34 (m, 1H), 2.13 (s, 1H), 2.10 – 2.06 (m, 6H), 1.68 (s, 12H), 1.65 (d, J = 7.0 Hz, 3H), 1.61 (d, J = 6.9 Hz, 3H), 1.54 – 1.49 (m, 2H), 1.32 (dd, J = 12.2, 6.8 Hz, 2H), 1.25 (dd, J = 11.3, 4.5 Hz, 9H).

实施例 20-2

linker 的合成

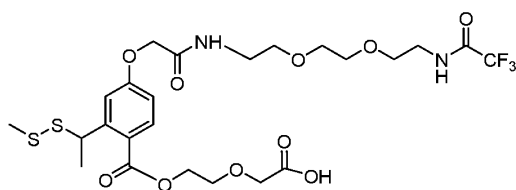
linker 的合成方法与实施例 20-1 类似。



合成路线与实施例 20-1 类似。

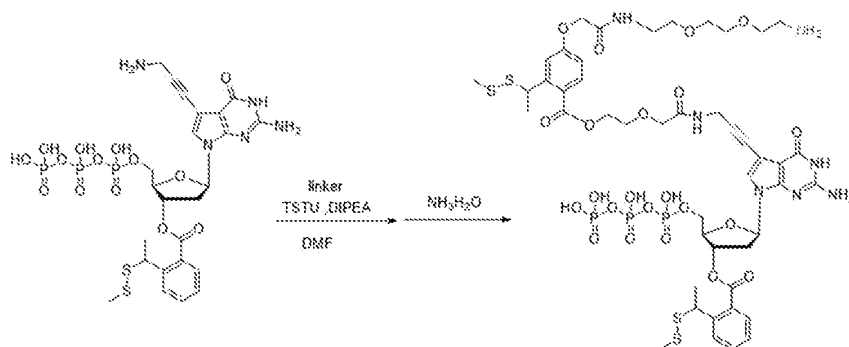
实施例 21

linker 的合成

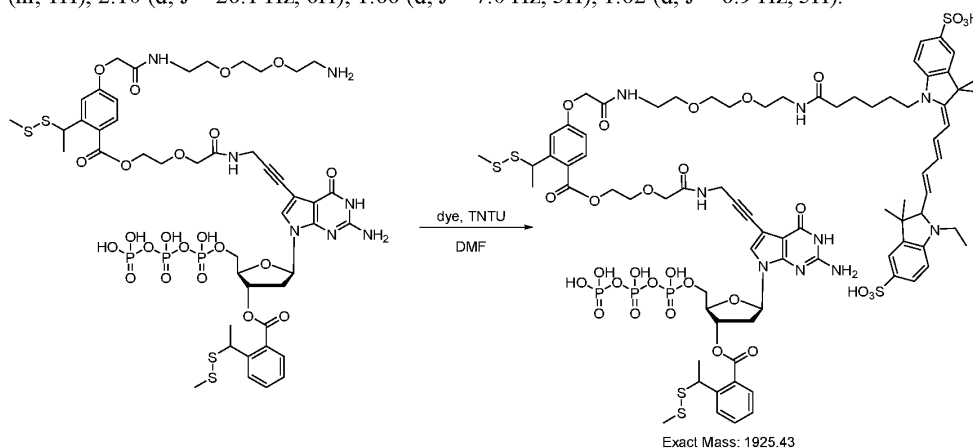


linker 12

linker 12 的合成方法与实施例 7-1 类似。MS[ES(-)], m/z 629.2. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 9.53 (s, 1H), 8.22 (t, $J = 5.7$ Hz, 1H), 7.85 (d, $J = 8.7$ Hz, 1H), 7.14 (d, $J = 2.5$ Hz, 1H), 6.96 (dd, $J = 8.8, 2.5$ Hz, 1H), 5.33 (q, $J = 6.8$ Hz, 1H), 4.63 – 4.53 (m, 2H), 4.40 – 4.30 (m, 2H), 3.94 (s, 2H), 3.79 (t, $J = 4.5$ Hz, 2H), 3.52 – 3.48 (m, 4H), 3.46 – 3.43 (m, 2H), 3.36 – 3.32 (m, 4H), 3.29 (dd, $J = 11.8, 5.9$ Hz, 2H), 2.13 (s, 3H), 1.62 (d, $J = 7.0$ Hz, 3H).



合成方法与实施例 7-3 类似。MS[ES(-)], m/z 1284.2. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.77 (s, 1H), 8.46 (dd, $J = 14.7, 6.2$ Hz, 1H), 7.89 (dd, $J = 17.9, 8.1$ Hz, 2H), 7.66 – 7.60 (m, 2H), 7.47 – 7.41 (m, 1H), 7.30 (d, $J = 4.2$ Hz, 1H), 7.09 (d, $J = 1.9$ Hz, 1H), 7.07 – 6.99 (m, 1H), 6.51 – 6.35 (m, 2H), 6.35 – 6.23 (m, 1H), 5.64 (s, 1H), 5.35 (d, $J = 6.9$ Hz, 1H), 5.20 (dd, $J = 6.9, 2.8$ Hz, 1H), 4.66 (s, 2H), 4.44 – 4.38 (m, 1H), 4.37 – 4.33 (m, 1H), 4.32 (d, $J = 4.8$ Hz, 1H), 4.25 (t, $J = 4.2$ Hz, 1H), 4.11 (d, $J = 5.2$ Hz, 2H), 4.05 – 4.02 (m, 2H), 4.01 (s, 2H), 3.83 – 3.80 (m, 2H), 3.62 (s, 2H), 3.51 (d, $J = 7.5$ Hz, 4H), 3.45 (d, $J = 5.5$ Hz, 4H), 3.29 – 3.25 (m, 2H), 2.49 – 2.43 (m, 1H), 2.41 – 2.36 (m, 1H), 2.10 (d, $J = 26.1$ Hz, 6H), 1.66 (d, $J = 7.0$ Hz, 3H), 1.62 (d, $J = 6.9$ Hz, 3H).

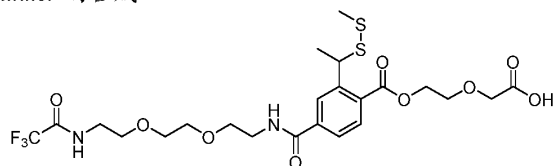


合成方法与实施例 7-3 类似。21mg 固体, HPLC 纯度 > 99%。

MS[ES(-)], m/z 1924.4. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 10.55 (s, 1H), 8.49 (d, $J = 6.0$ Hz, 1H), 8.36 (t, $J = 12.9$ Hz, 2H), 7.89 – 7.84 (m, 2H), 7.81 (s, 2H), 7.65 – 7.59 (m, 3H), 7.47 – 7.40 (m, 1H), 7.31 (dd, $J = 8.2, 1.4$ Hz, 2H), 7.12 (d, $J = 2.2$ Hz, 1H), 7.01 – 6.94 (m, 1H), 6.59 (t, $J = 12.2$ Hz, 1H), 6.40 (s, 2H), 6.30 (dd, $J = 13.2, 4.5$ Hz, 2H), 5.60 (d, $J = 58.2$ Hz, 1H), 5.33 (q, $J = 6.9$ Hz, 1H), 5.20 (dd, $J = 12.8, 5.8$ Hz, 1H), 4.58 (s, 2H), 4.44 – 4.33 (m, 2H), 4.33 – 4.28 (m, 1H), 4.28 – 4.22 (m, 1H), 4.13 (dd, $J = 11.7, 5.6$ Hz, 2H), 4.09 – 4.05 (m, 2H), 4.00 (s, 2H), 3.82 – 3.79 (m, 2H), 3.48 (s, 2H), 3.45 – 3.42 (m, 2H), 3.38 – 3.34 (m, 2H), 3.28 (dd, $J = 11.5, 5.8$ Hz, 2H), 3.16 (dd, $J = 11.3, 5.5$ Hz, 2H), 2.13 (s, 1H), 2.12 (s, 2H), 2.08 – 2.03 (m, 3H), 1.68 (s, 9H), 1.65 (d, $J = 6.2$ Hz, 3H), 1.61 (d, $J = 6.9$ Hz, 3H), 1.55 – 1.49 (m, 2H), 1.35 – 1.30 (m, 2H), 1.25 (dd, $J = 12.0, 4.4$ Hz, 6H).

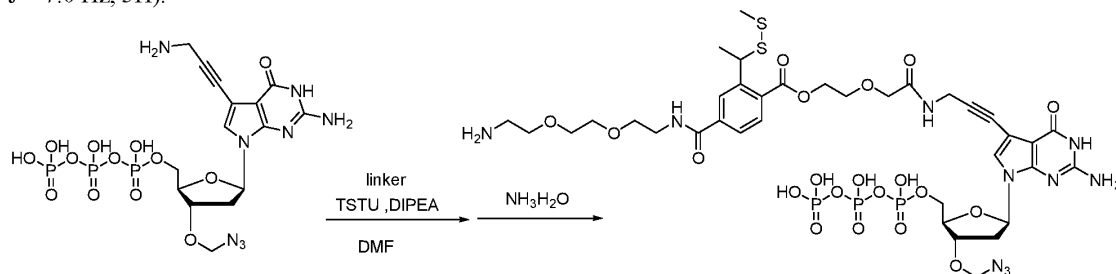
实施例 22

linker 的合成



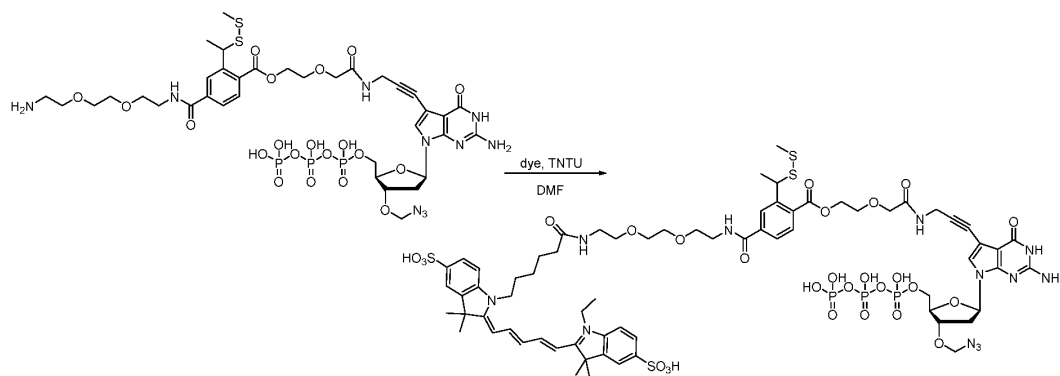
linker 13

linker 13 的合成方法与实施例 7-1 类似。MS[ES(-)], m/z 599.1. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 9.53 (s, 1H), 8.81 (t, $J = 5.5$ Hz, 1H), 8.05 (s, 1H), 7.84 (s, 2H), 5.17 (q, $J = 7.0$ Hz, 1H), 4.50 – 4.31 (m, 2H), 3.84 (s, 2H), 3.80 – 3.77 (m, 2H), 3.55 (d, $J = 6.1$ Hz, 2H), 3.53 (s, 4H), 3.49 (t, $J = 5.8$ Hz, 2H), 3.43 (dd, $J = 11.5, 5.7$ Hz, 2H), 3.32 (dd, $J = 10.7, 5.3$ Hz, 2H), 2.09 (s, 3H), 1.67 (d, $J = 7.0$ Hz, 3H).



Exact Mass: 1100.19

合成方法与实施例 7-3 类似。MS[ES(-)], m/z 1099.2. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 9.36 (s, 1H), 8.43 (s, 1H), 8.06 (d, $J = 7.9$ Hz, 2H), 7.92 (d, $J = 8.2$ Hz, 1H), 7.20 (s, 1H), 6.45 (s, 2H), 6.23 (dd, $J = 8.8, 6.0$ Hz, 1H), 5.18 (q, $J = 7.1$ Hz, 1H), 4.90 (d, $J = 8.8$ Hz, 1H), 4.85 (d, $J = 8.8$ Hz, 1H), 4.53 (d, $J = 4.0$ Hz, 1H), 4.51 – 4.36 (m, 2H), 4.14 – 4.05 (m, 4H), 4.01 (s, 2H), 3.90 (dd, $J = 11.1, 6.1$ Hz, 2H), 3.85 – 3.80 (m, 2H), 3.63 – 3.58 (m, 6H), 3.54 (s, 4H), 2.62 – 2.54 (m, 1H), 2.28 (dd, $J = 12.9, 5.5$ Hz, 1H), 2.09 (d, $J = 2.8$ Hz, 3H), 1.66 (d, $J = 7.0$ Hz, 3H).



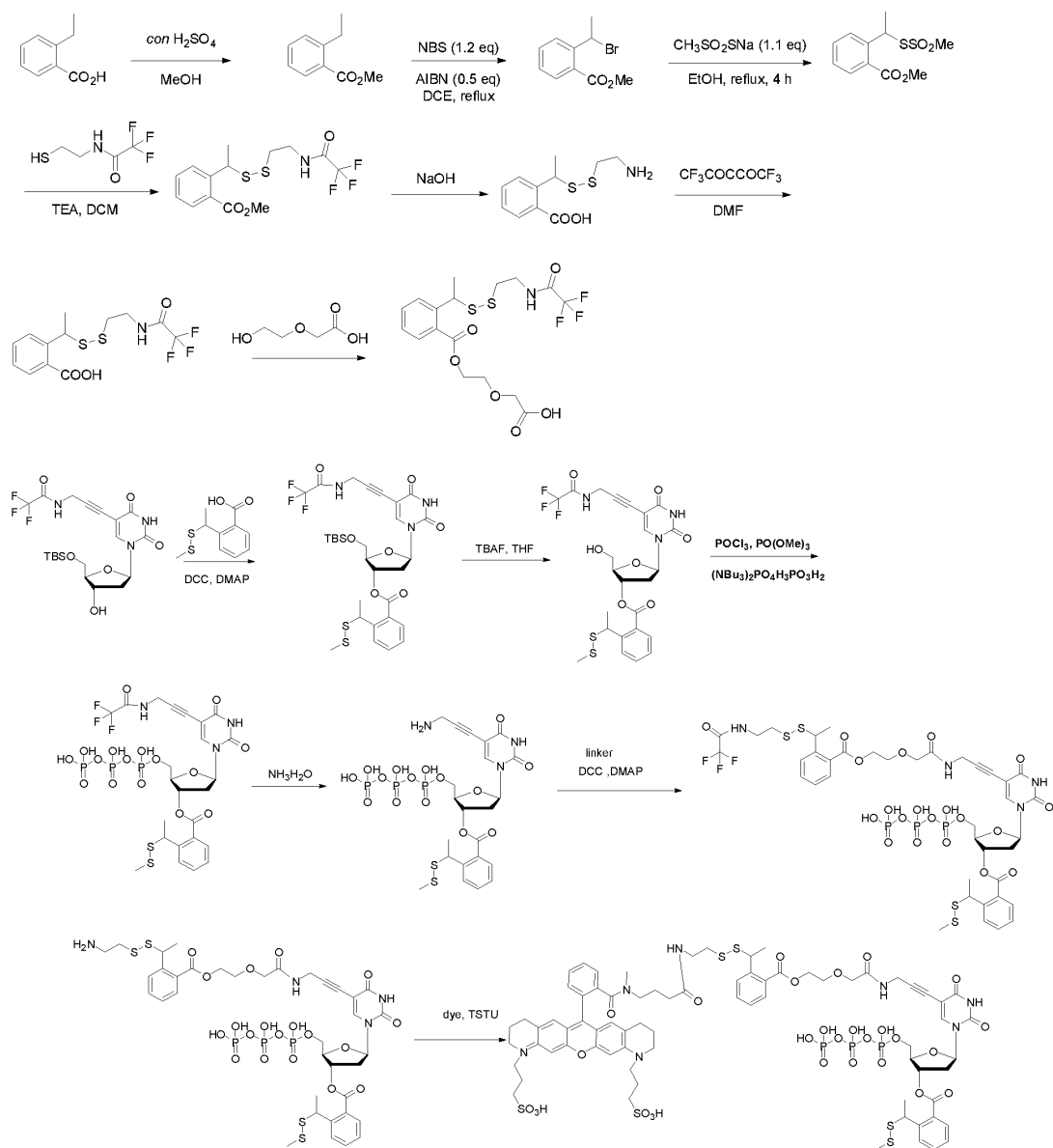
Exact Mass: 1739.41

合成方法与实施例 7-3 类似。13mg 固体, HPLC 纯度 > 99%。

MS[ES(-)], m/z 1738.4. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 10.55 (s, 1H), 8.91 (s, 1H), 8.45 (t, $J = 5.6$ Hz, 1H), 8.35 (t, $J = 13.0$ Hz, 2H), 8.04 (s, 1H), 7.92 – 7.84 (m, 3H), 7.80 (d, $J = 1.7$ Hz, 2H), 7.64 (ddd, $J = 8.2, 3.7, 1.5$ Hz, 2H), 7.31 (dd, $J = 8.3, 3.0$ Hz, 2H), 7.22 (s, 1H), 6.59 (t, $J = 12.3$ Hz, 1H), 6.46 (s, 2H), 6.30 (dd, $J = 13.7, 4.3$ Hz, 2H), 6.21 (dd, $J = 8.9, 5.8$ Hz, 1H), 5.15 (q, $J = 6.9$ Hz, 1H), 4.90 (d, $J = 8.9$ Hz, 1H), 4.85 (d, $J = 8.9$ Hz, 1H), 4.53 (d, $J = 3.9$ Hz, 1H), 4.51 – 4.44 (m, 1H), 4.44 – 4.38 (m, 1H), 4.15 – 4.09 (m, 4H), 4.09 – 4.03 (m, 2H), 3.99 (s, 2H), 3.97 – 3.87 (m, 2H), 3.84 – 3.80 (m, 2H), 3.76 – 3.72 (m, 2H), 3.44 – 3.41 (m, 2H), 3.38 – 3.34 (m, 4H), 3.15 (dd, $J = 11.6, 5.8$ Hz, 2H), 2.91 – 2.88 (m, 2H), 2.66 – 2.60 (m, 1H), 2.28 (dd, $J = 13.1, 5.6$ Hz, 1H), 2.08 (s, 3H), 2.07 – 2.02 (m, 2H), 1.67 (s, 12H), 1.65 (d, $J = 8.0$ Hz, 3H), 1.56 – 1.48 (m, 2H), 1.35 – 1.29 (m, 2H), 1.26 – 1.22 (m, 5H).

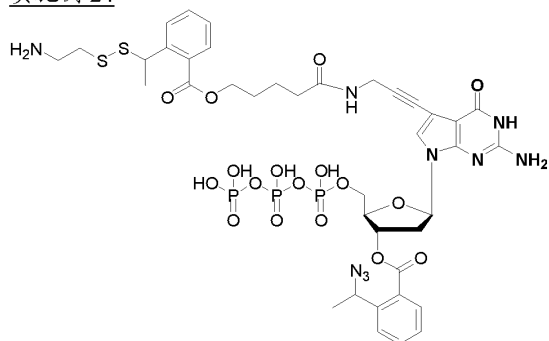
实施例 23

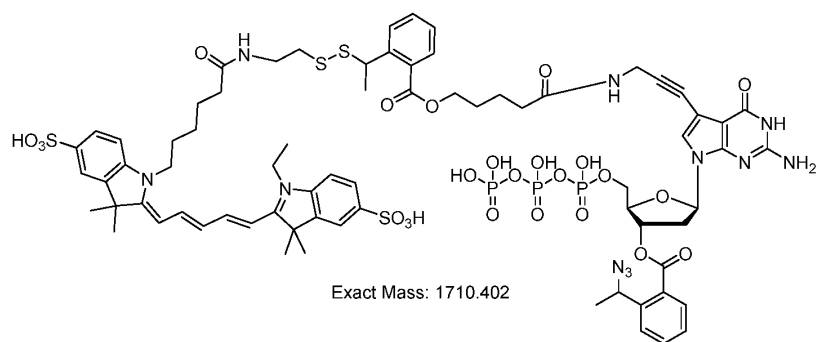
linker 的合成



合成路线与实施例 19 类似。

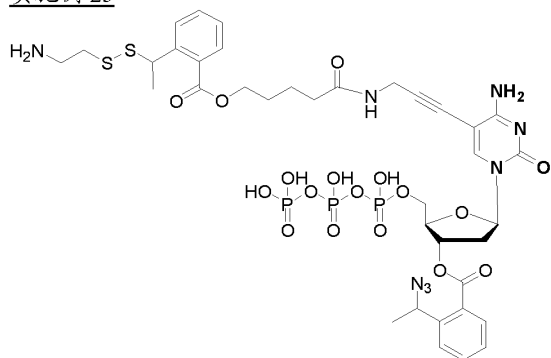
实施例 24



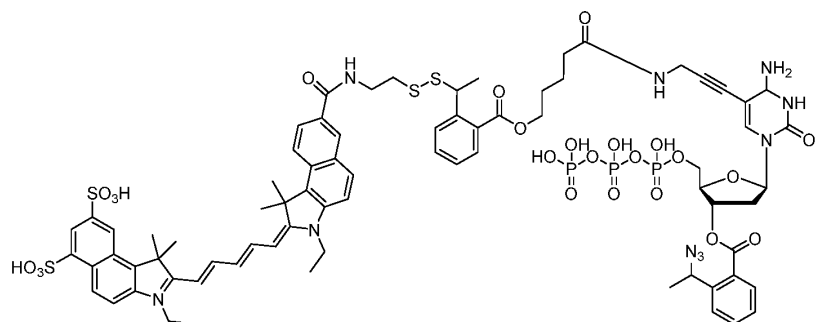


合成路线与实施例 7-3 类似。

实施例 25

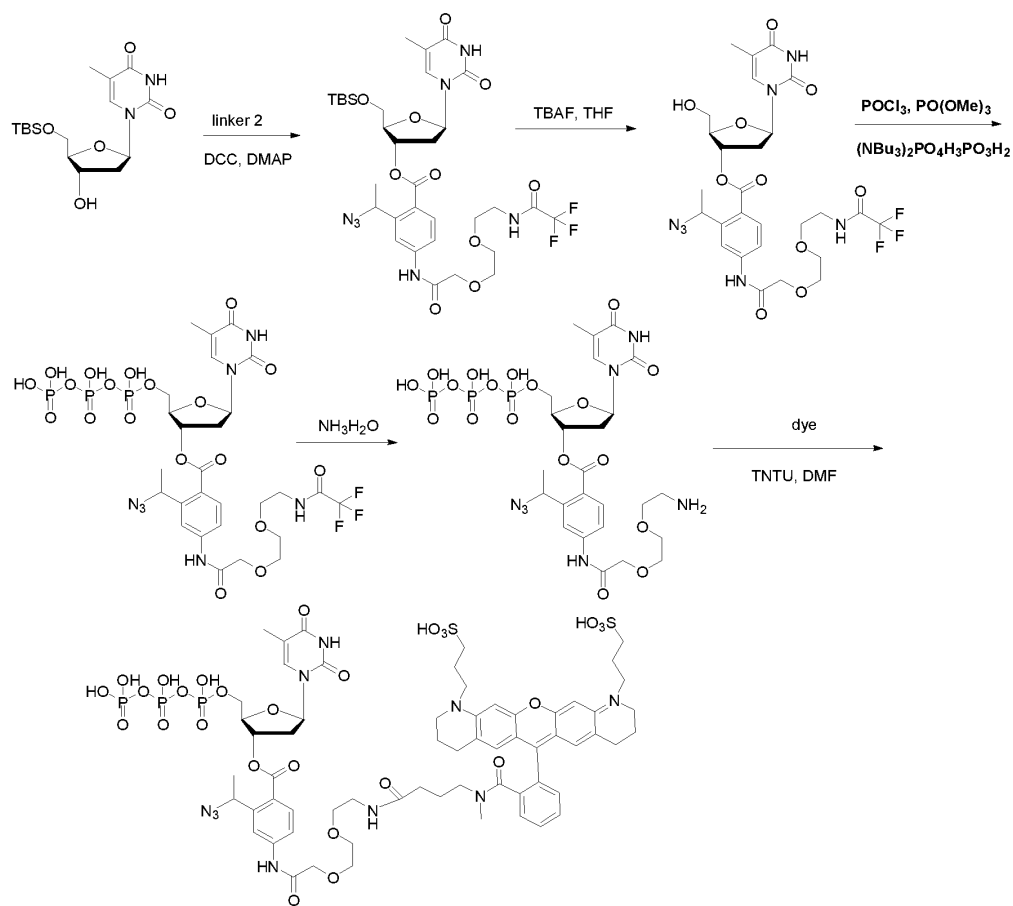


合成路线与实施例 7-3 类似。



合成路线与实施例 7-3 类似。

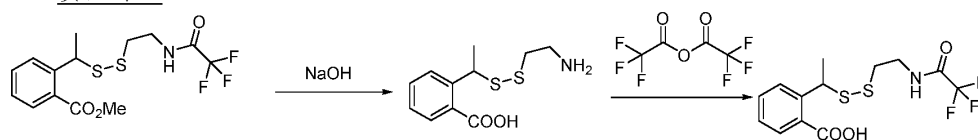
实施例 26

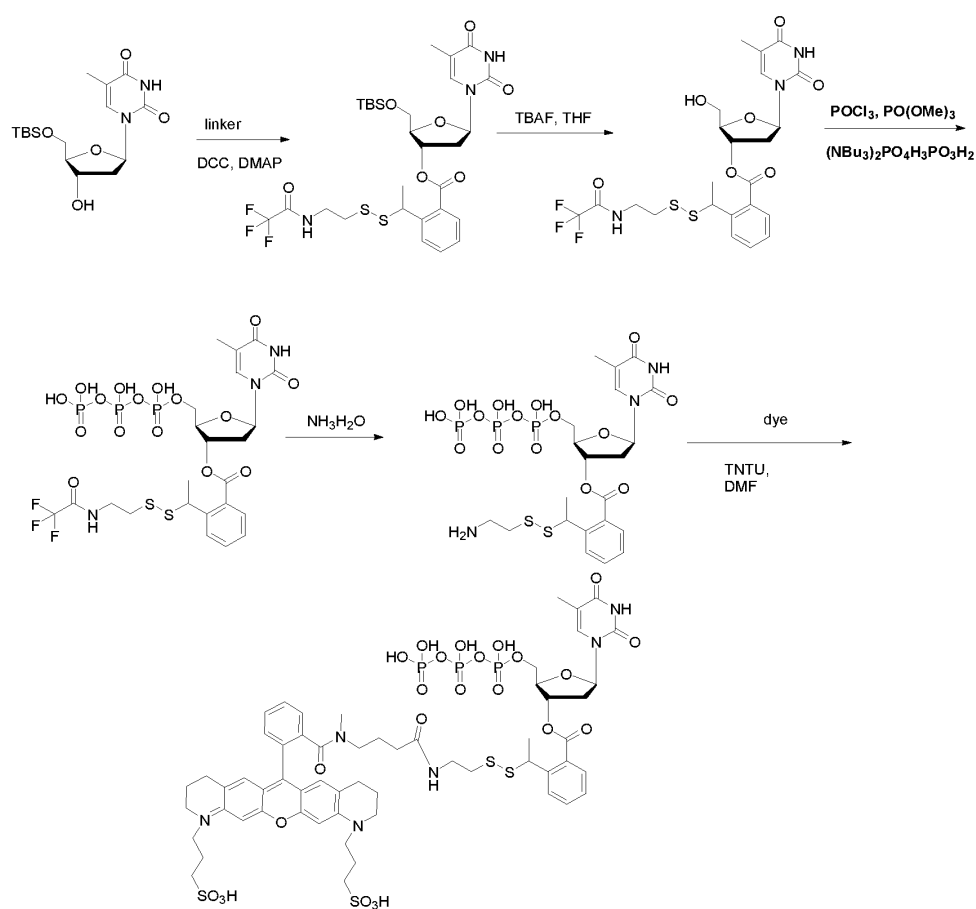


linker 2 的合成方法参见实施例 2。

核苷与 linker 2 DCC 缩合得到酯化产物，随后脱去 TBS 保护基、三磷酸化以及氨解后得到片段，片段与染料偶联得到产品。

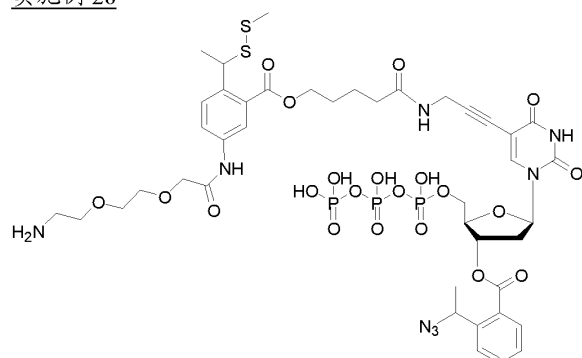
实施例 27

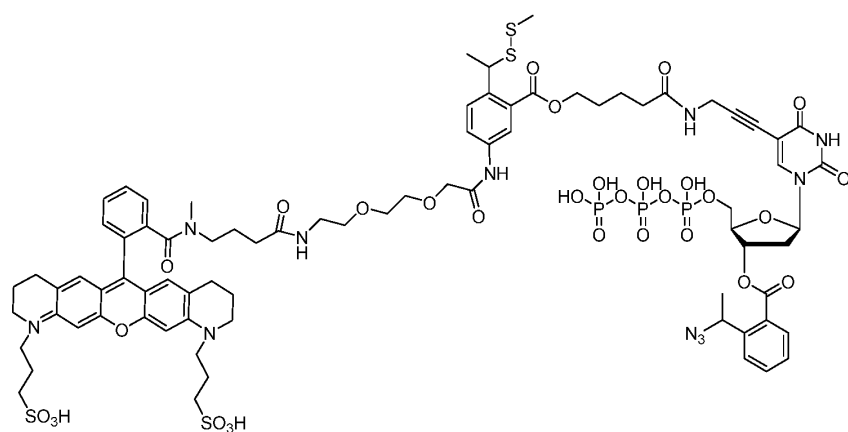




合成路线与实施例 19 类似。

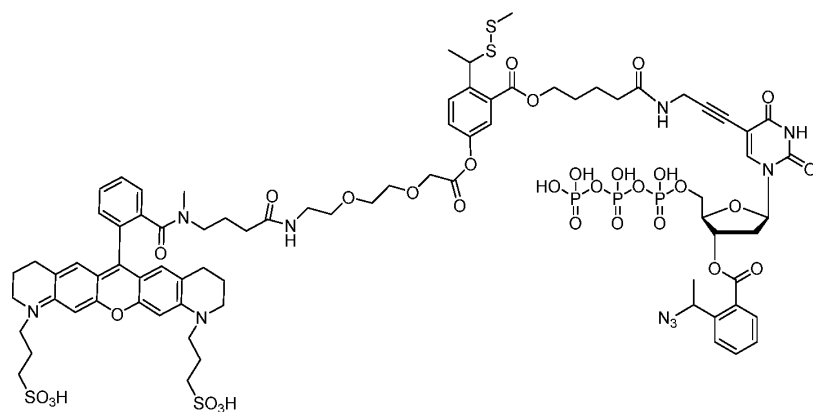
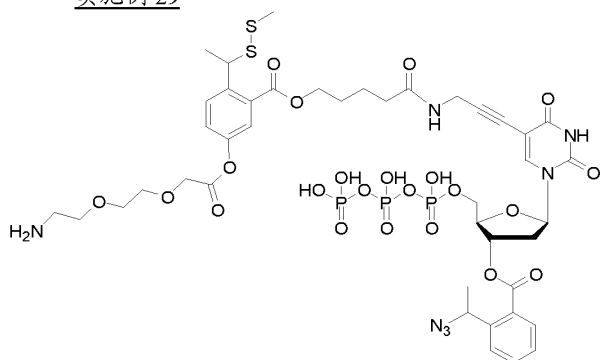
实施例 28





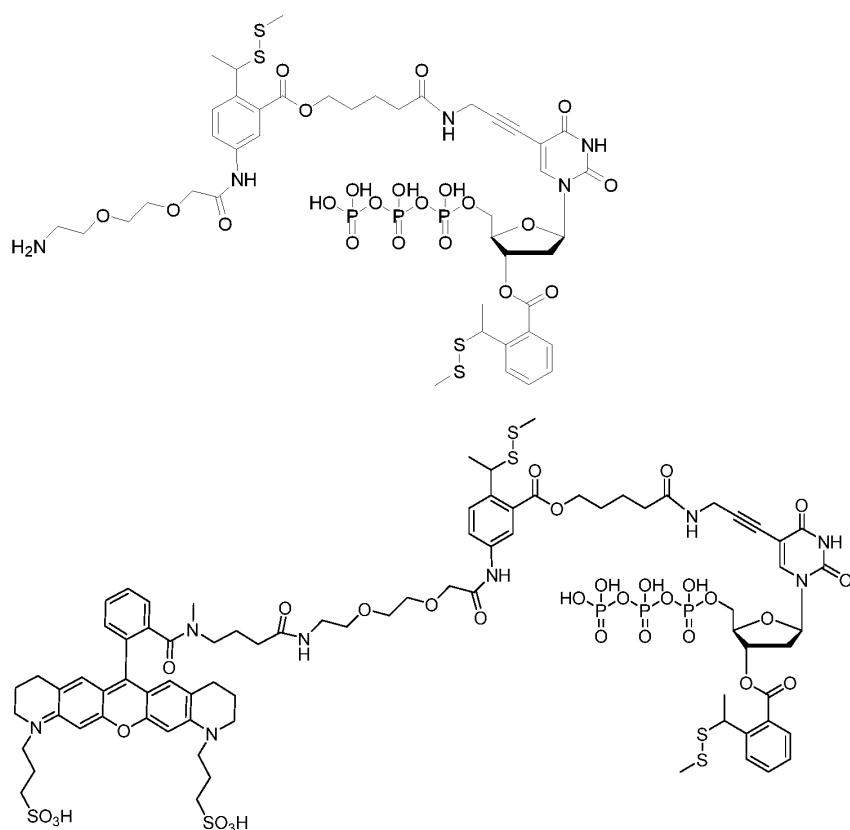
合成路线与实施例 7-1 类似。

实施例 29



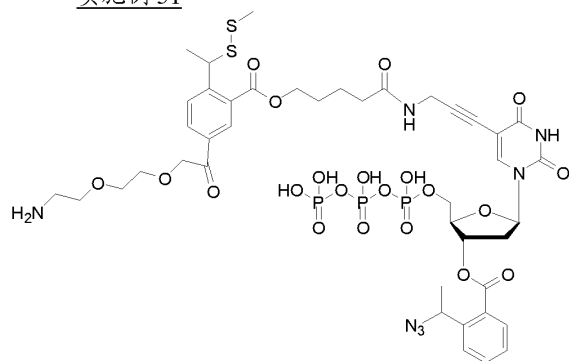
合成路线与实施例 7-1 类似。

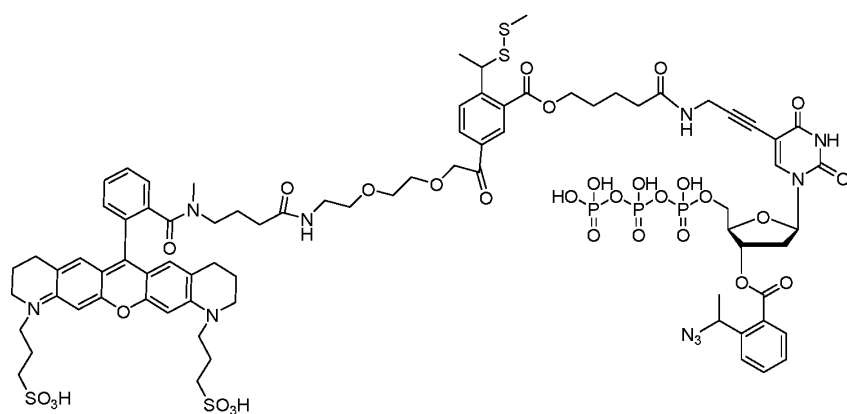
实施例 30



合成路线与实施例 7-1 类似。

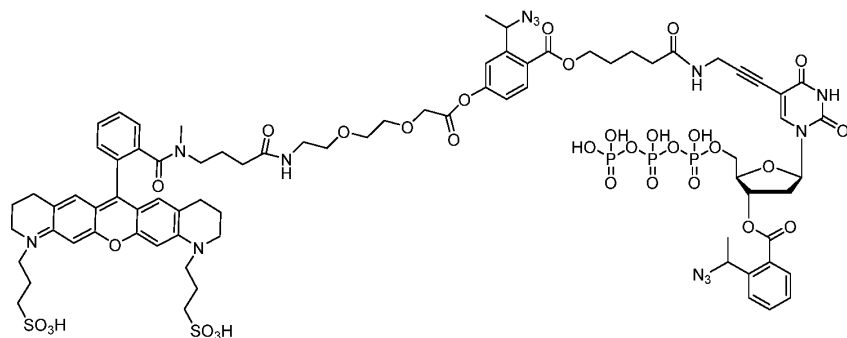
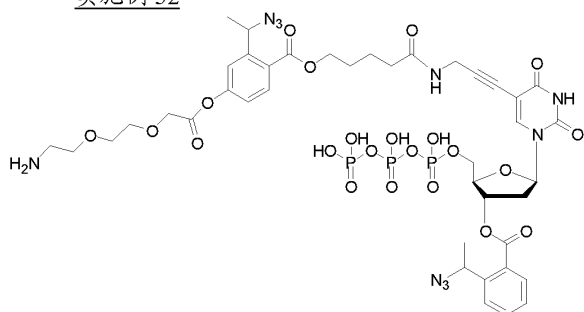
实施例 31





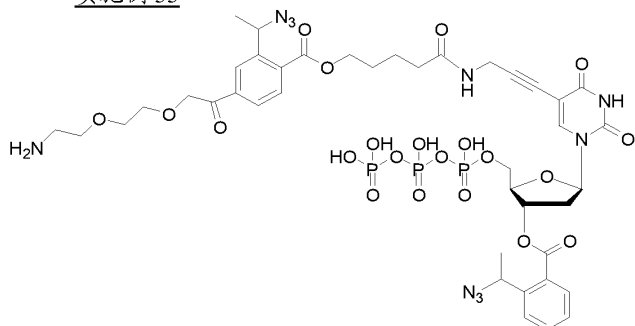
合成路线与实施例 7-1 类似。

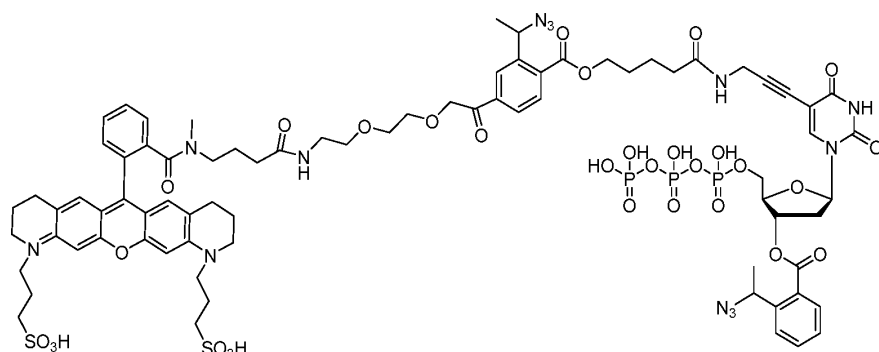
实施例 32



合成路线与实施例 7-1 类似。

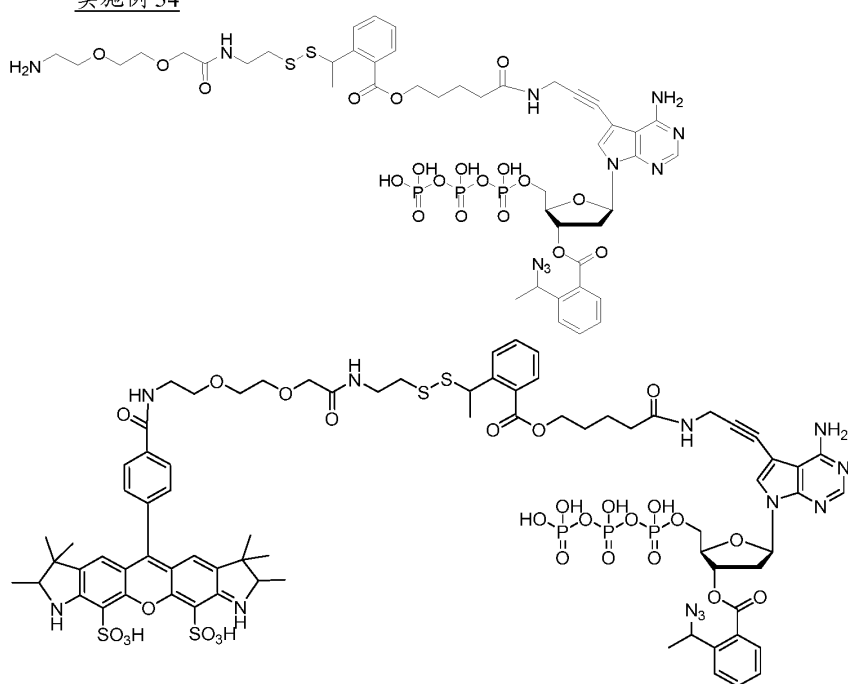
实施例 33





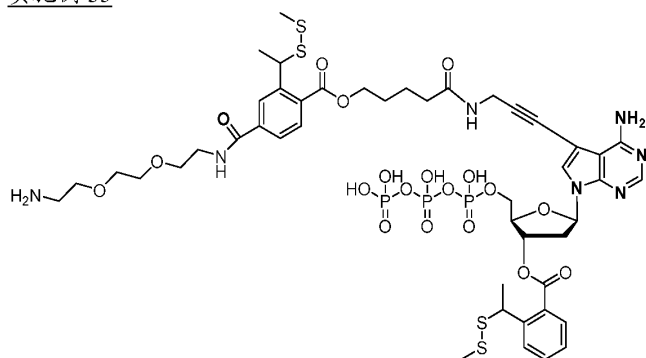
合成路线与实施例 7-1 类似。

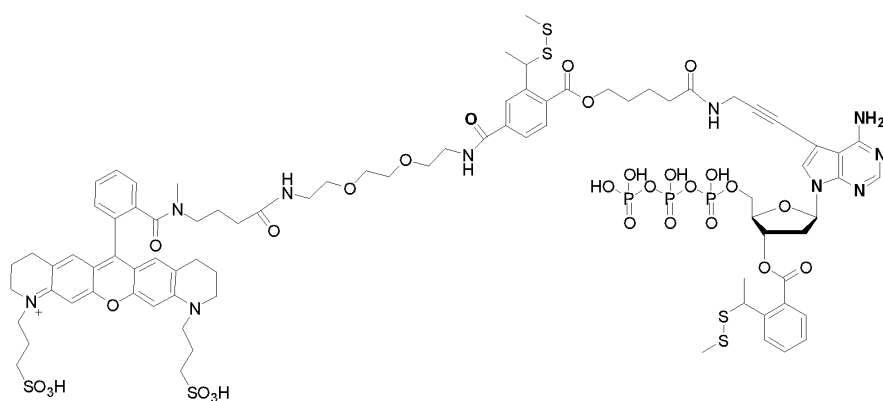
实施例 34



合成路线参考实施例 19。

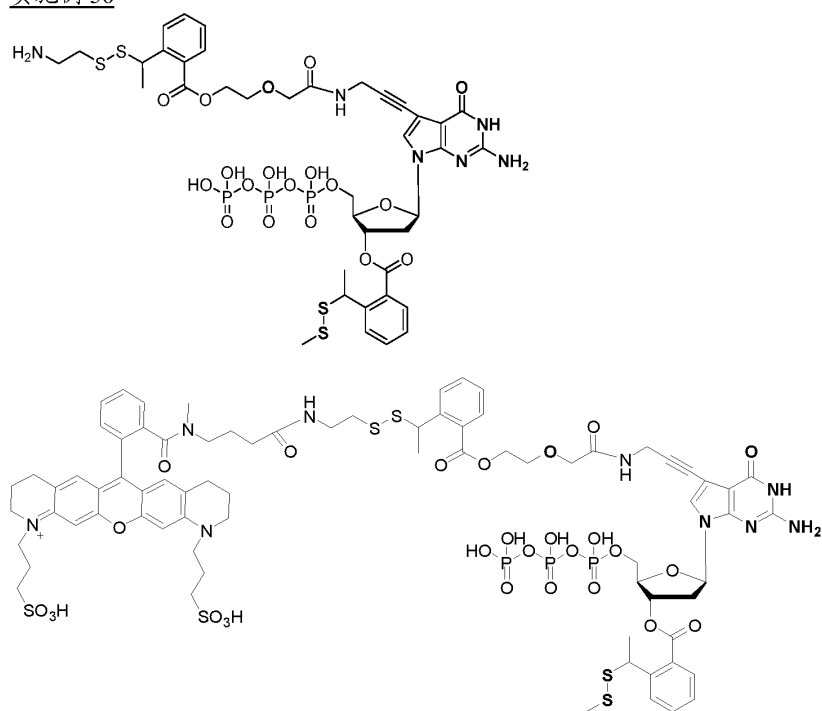
实施例 35





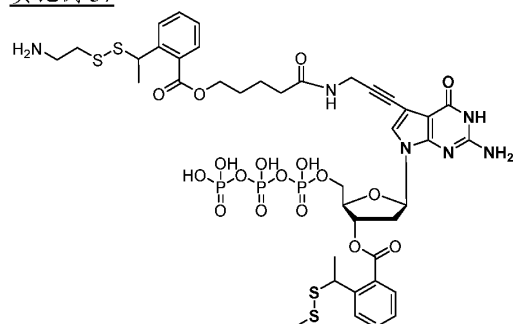
合成路线与实施例 22 类似。

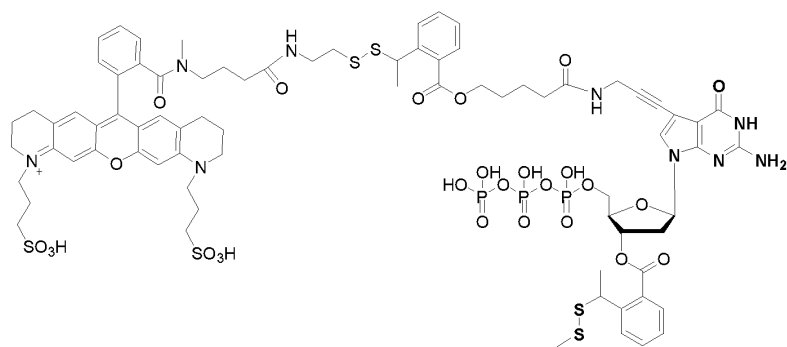
实施例 36



合成路线与实施例 19 类似。

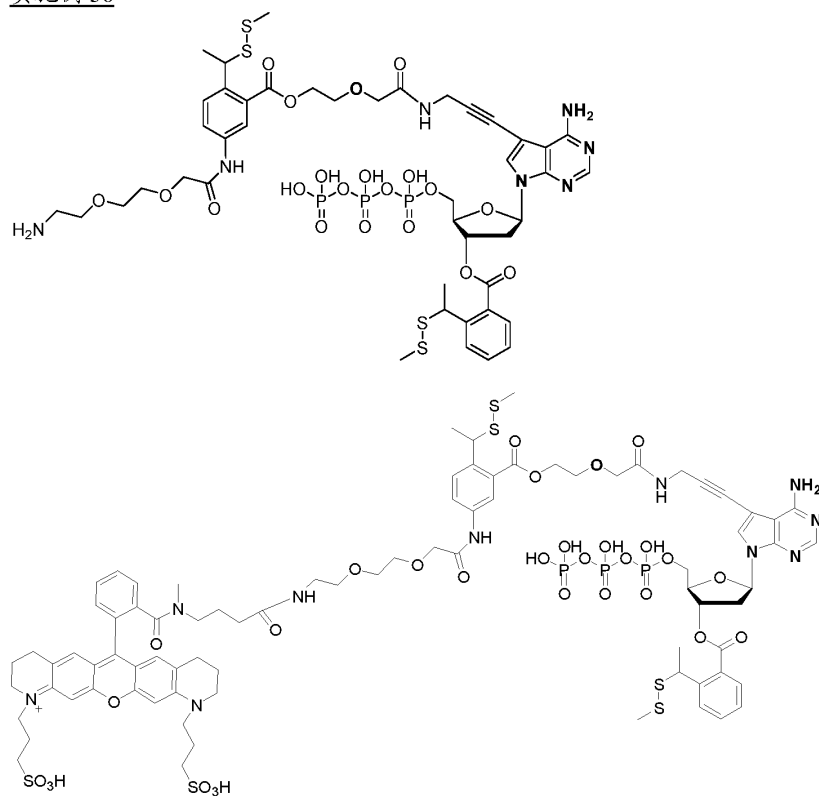
实施例 37





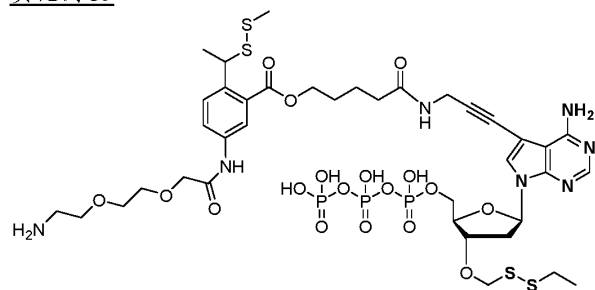
合成路线与实施例 8 类似。

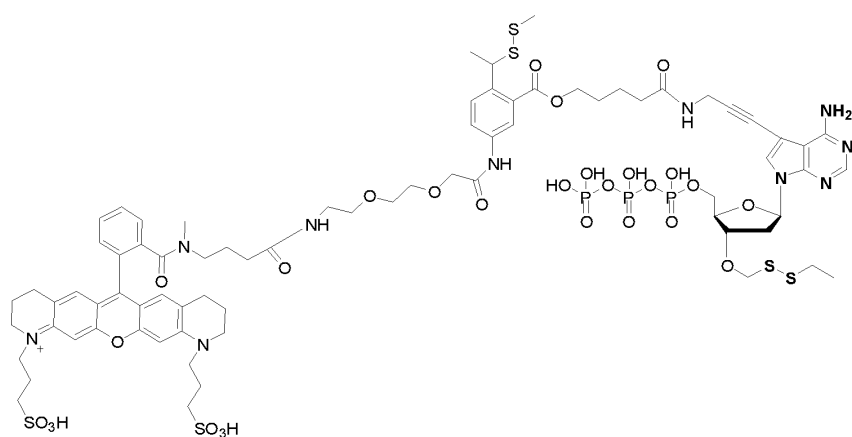
实施例 38



合成路线与实施例 7-2 类似。

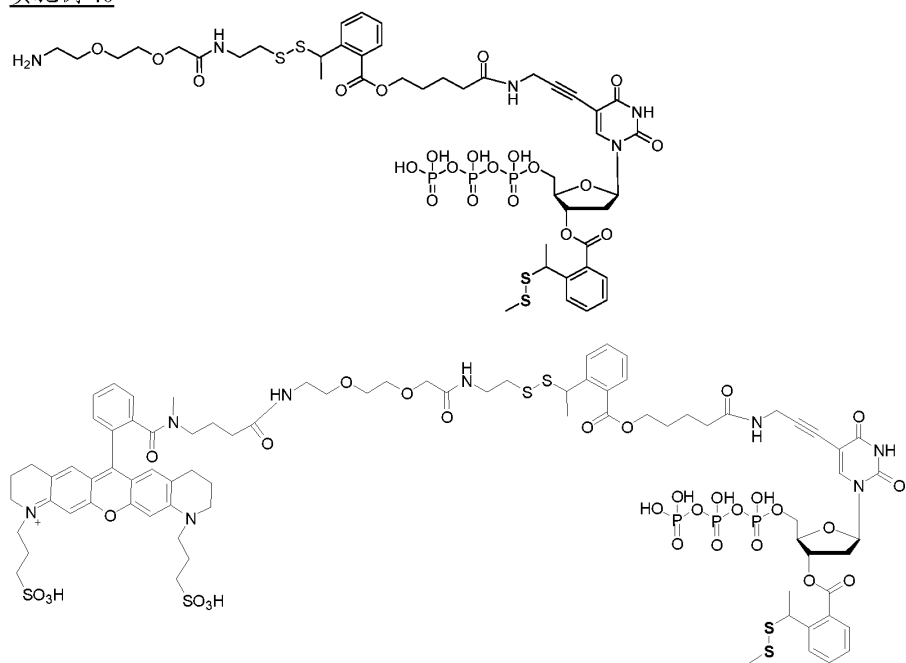
实施例 39





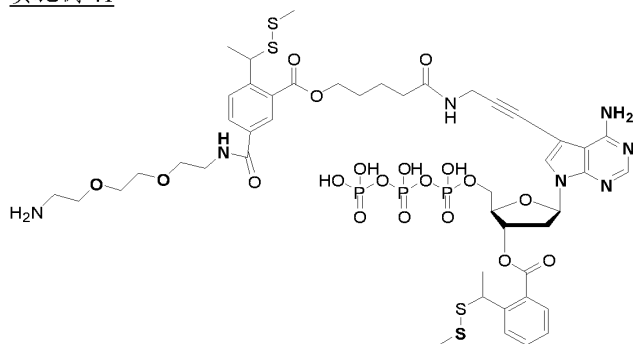
合成路线与实施例 6-2 类似。

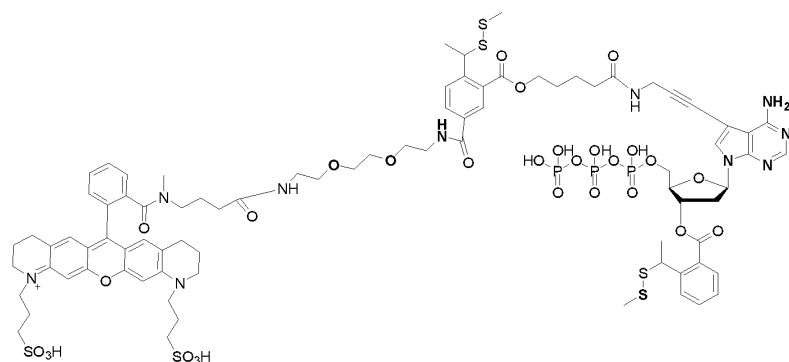
实施例 40



合成路线与实施例 19 类似。

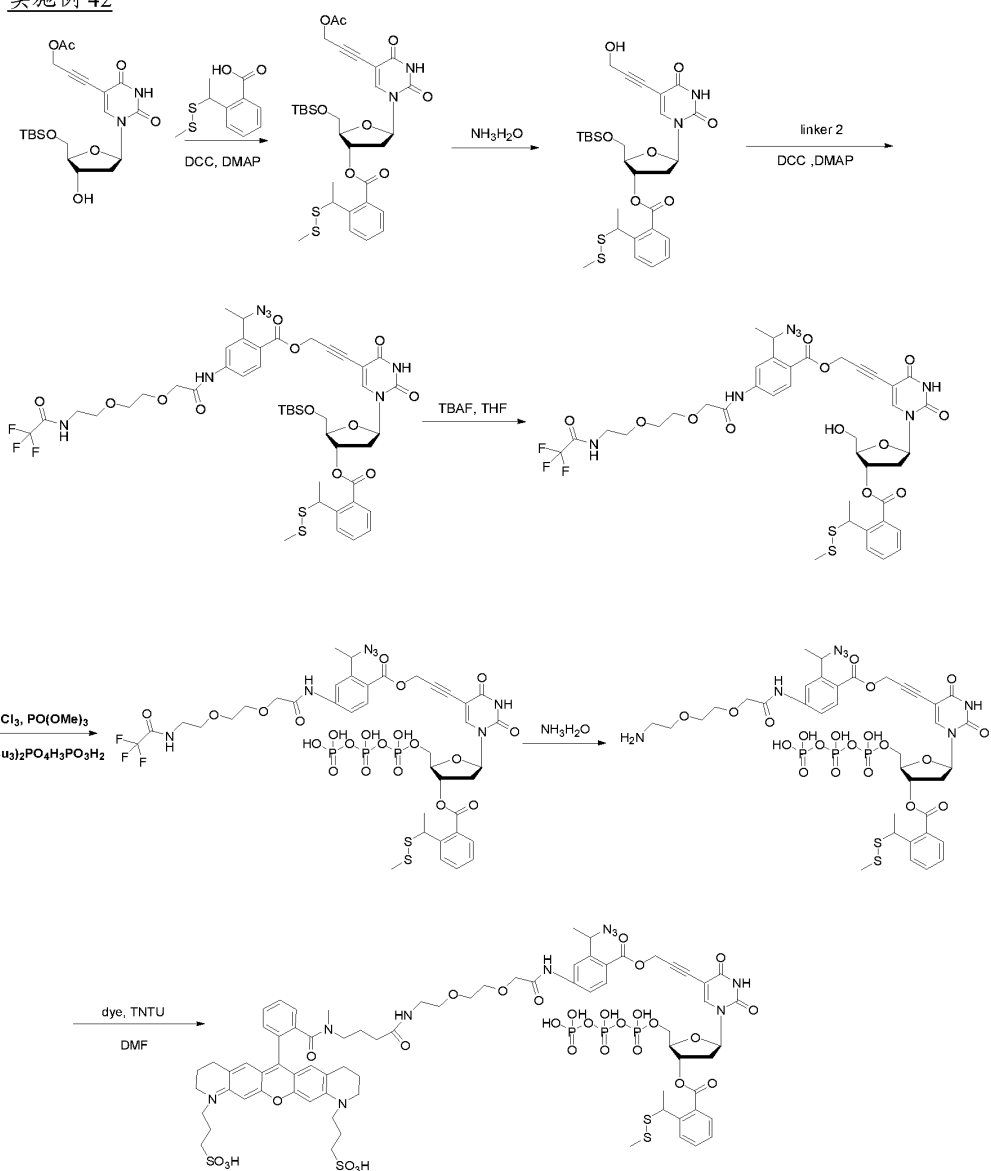
实施例 41





合成路线与实施例 22 类似。

实施例 42

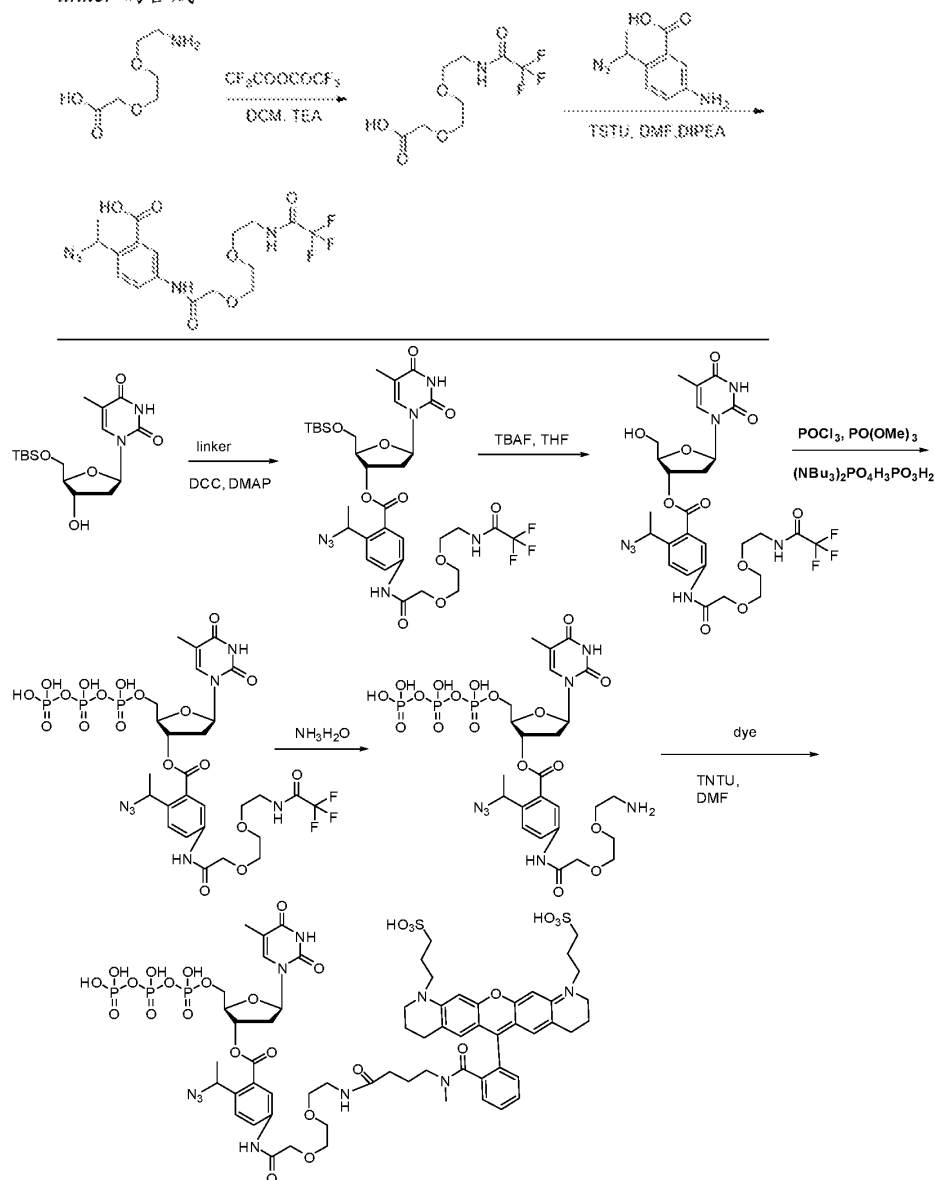


linker 2 的合成参见实施例 2。

化合物缩合后得到 TBS 保护核苷，氨解去掉乙酰基后，与 linker2 DCC 缩合，脱去 TBS 保护基，三磷酸化后再氨基得到氨基化合物，最后标记荧光染料得到产品。

实施例 43

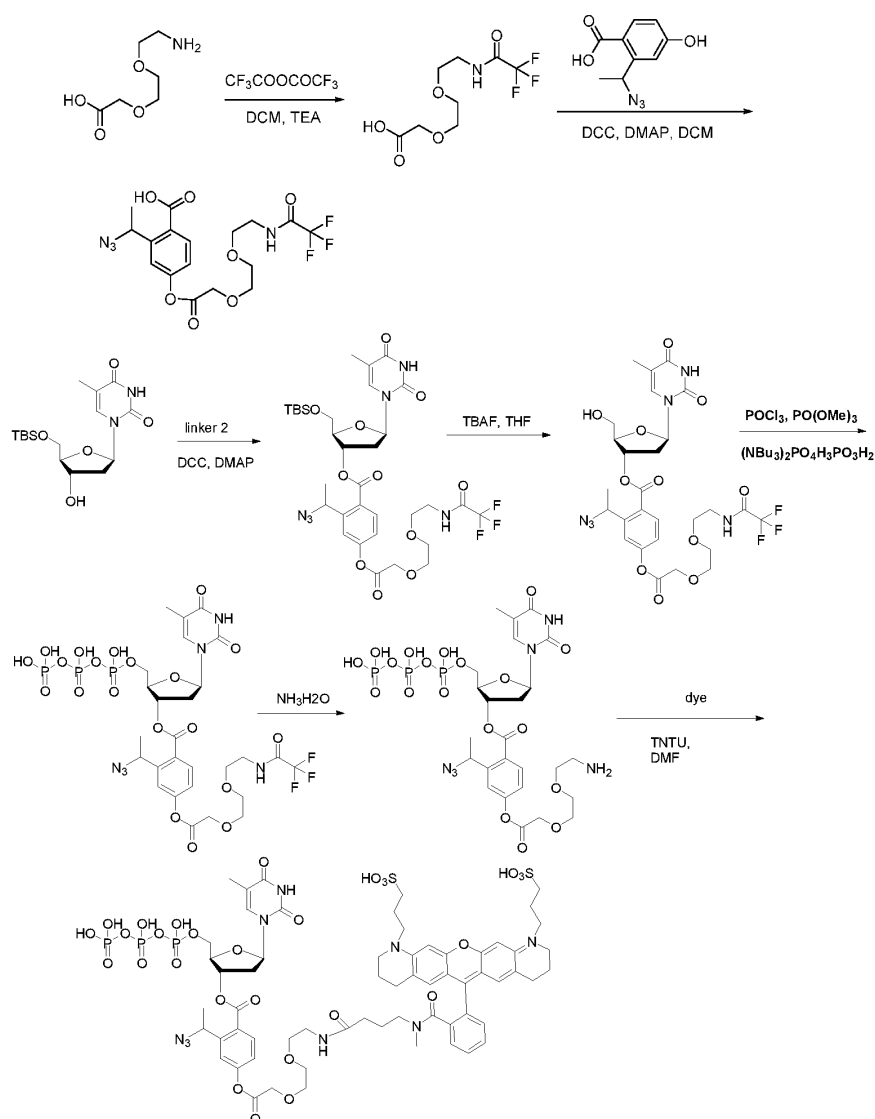
linker 的合成



Linker 在 DCC 缩合条件下反应，脱去 TBS 保护基，三磷酸化后得到三磷酸。氨基脱去三氟乙酰基，最后与染料缩合得到产品。

实施例 44

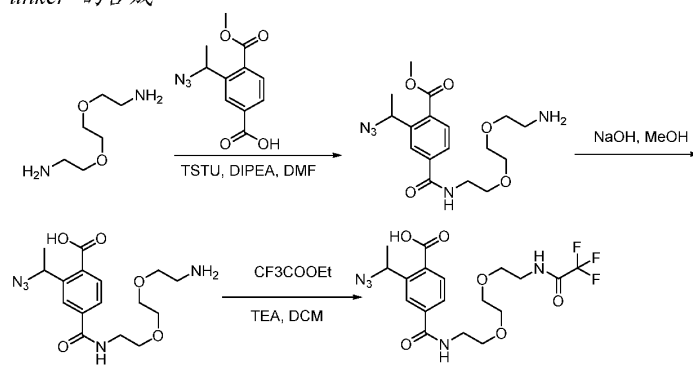
linker 2 的合成

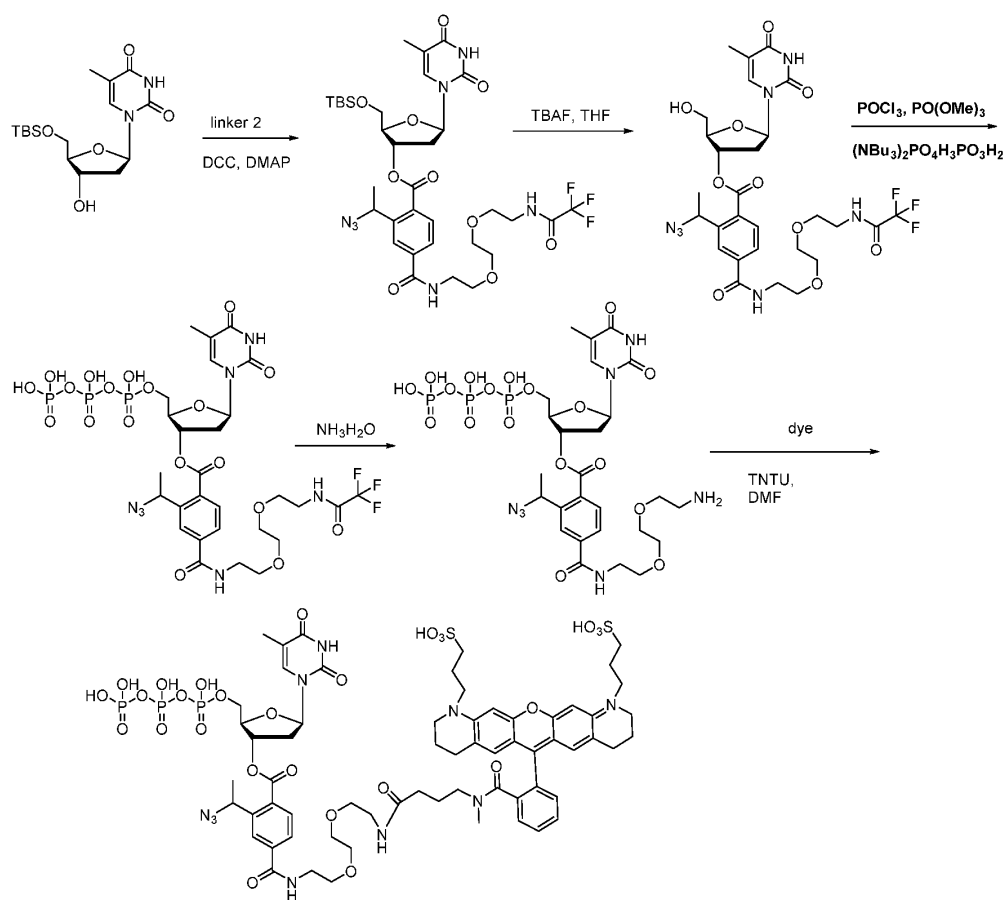


linker 2 在 DCC 缩合条件下反应，脱去 TBS 保护基，三磷酸化后得到三磷酸。氨基脱去三氟乙酰基，最后与染料缩合得到产品。

实施例 45

linker 的合成

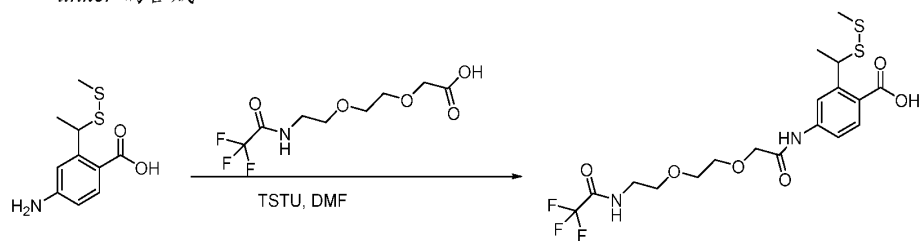


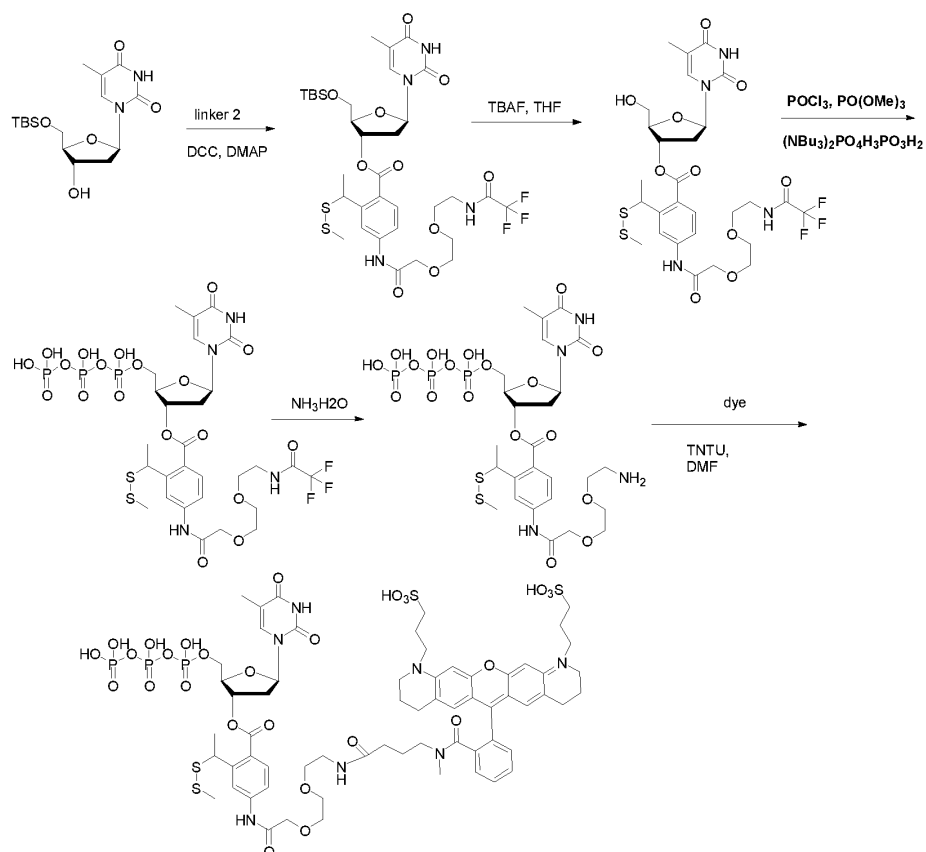


合成路线与实施例 2 类似。

实施例 46

linker 的合成

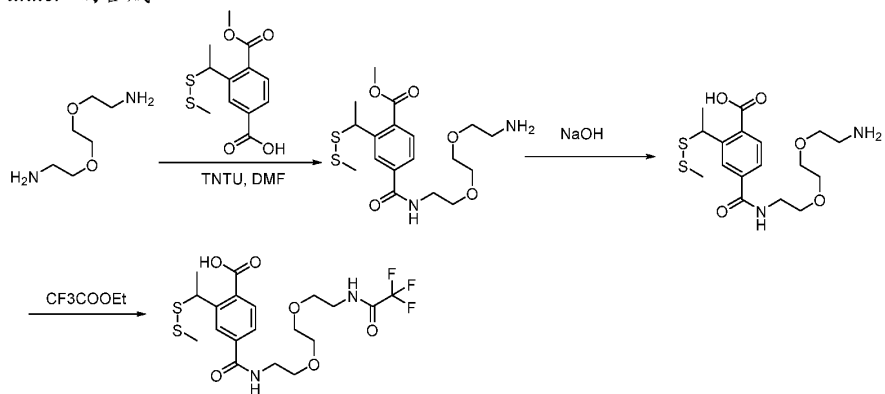


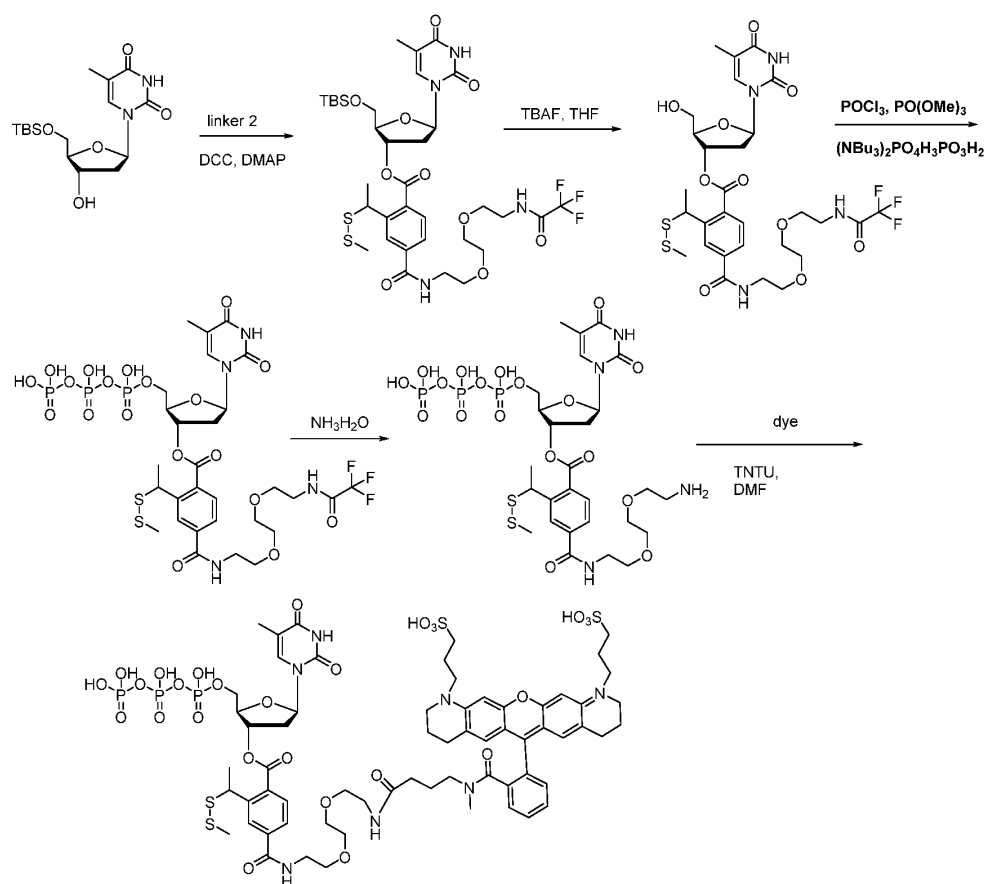


linker 的合成参见实施例 3。

实施例 47

linker 的合成

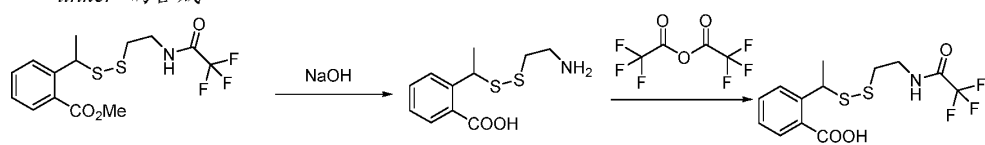




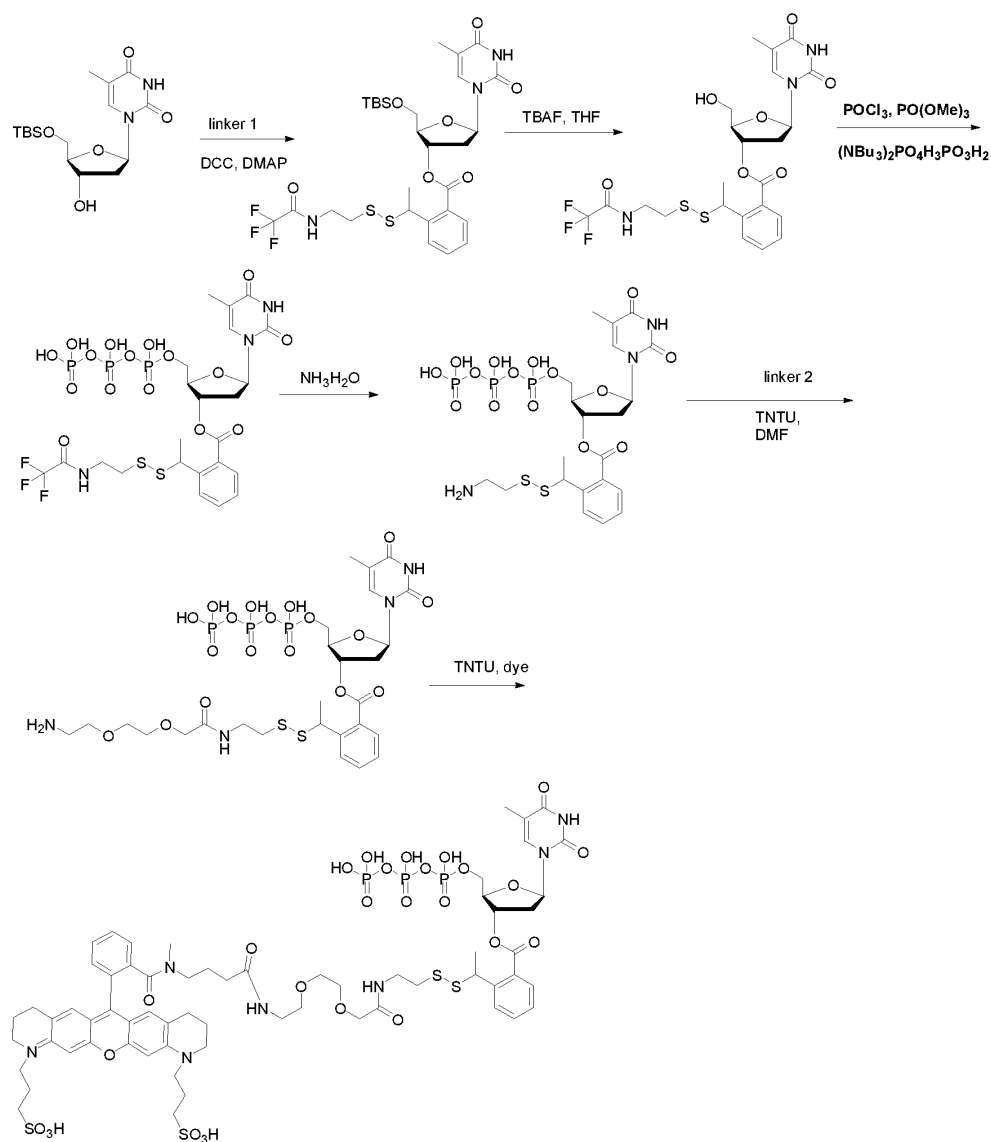
合成路线与实施例 2 类似。

实施例 48

linker 的合成

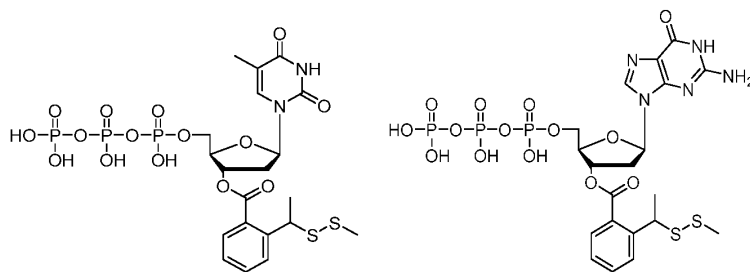


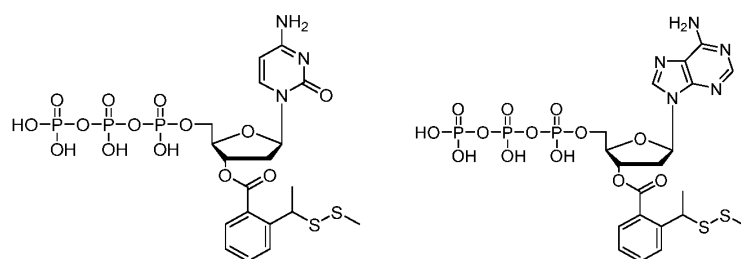
linker 是 2-(2-(2-氨基乙氧基)乙氧基)乙酸盐, 直接购买自安耐吉化学, 货号 A012027。



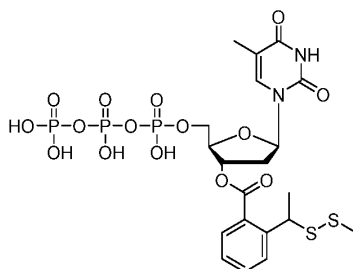
二、测序效果评估实施例

作为对照的以下四种 cold dNTP 可以参照 PCT/CN2021/125262 中实施例三十、实施例三十一、实施例三十二、实施例三十三的制备方法进行合成。

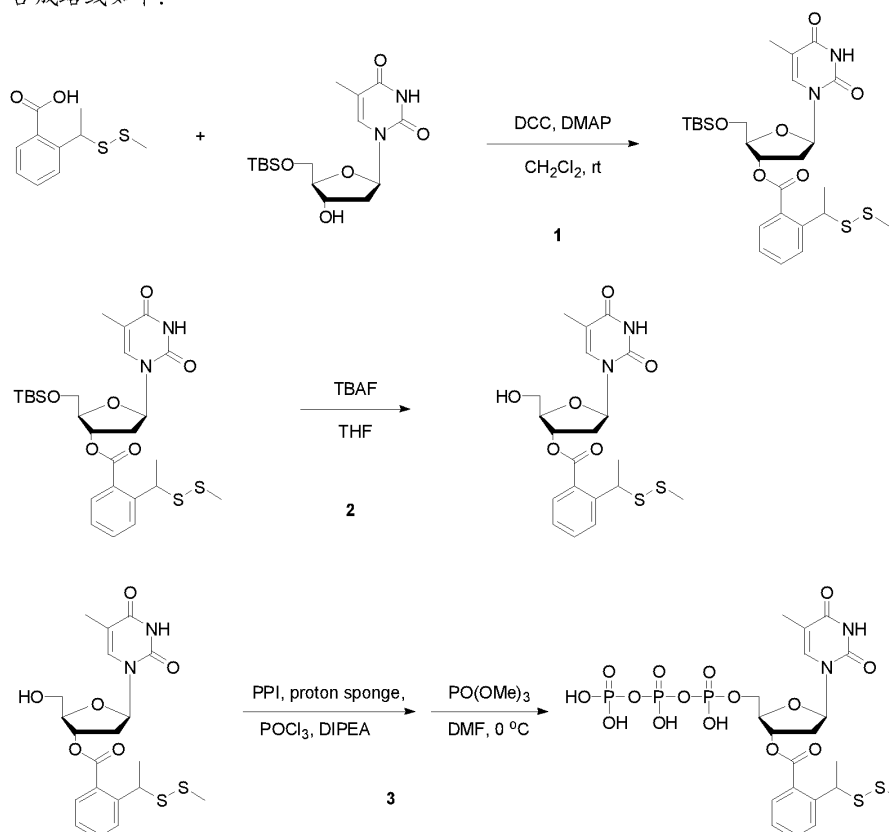




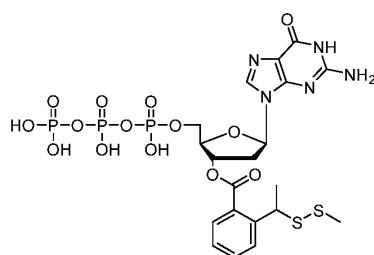
具体地:



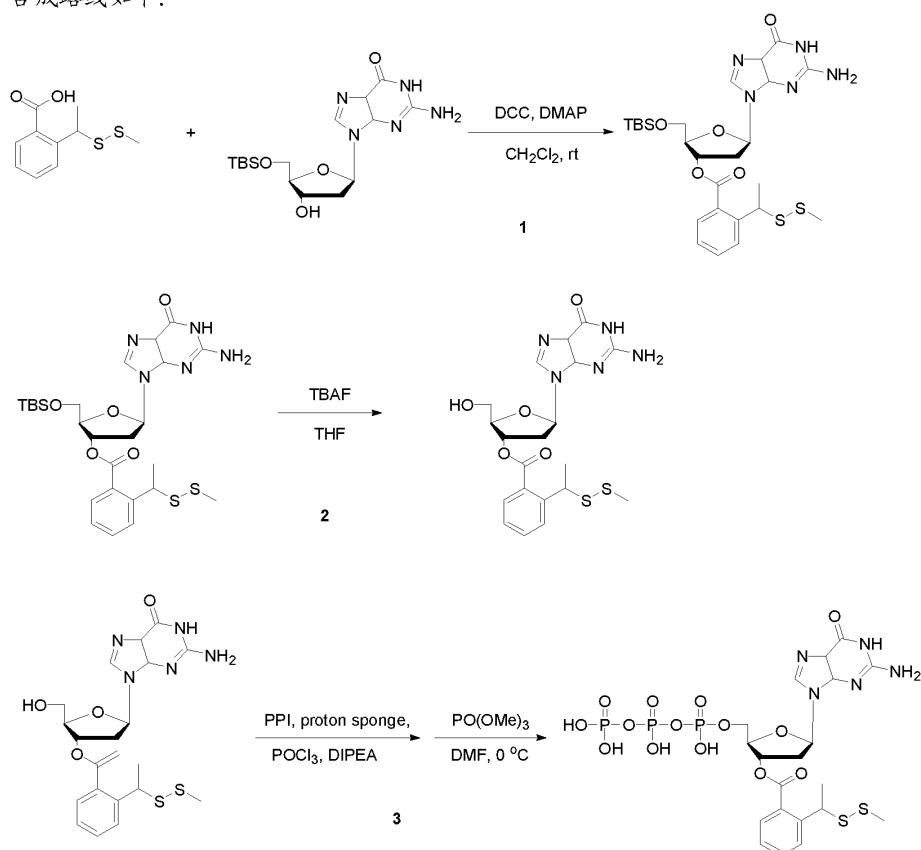
合成路线如下:



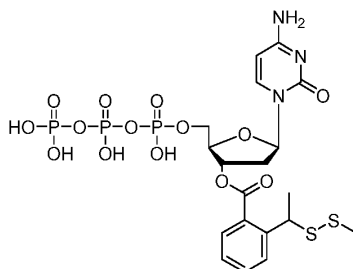
MW = 692. ^1H NMR (400 MHz, D_2O) δ 7.90 – 7.82 (m, 2H), 7.69 – 7.58 (m, 2H), 7.46 – 7.39 (m, 1H), 6.47 – 6.35 (m, 1H), 5.73 (s, 1H), 5.10 – 5.01 (m, 1H), 4.57 (dd, J = 13.6, 3.0 Hz, 1H), 4.41 – 4.24 (m, 2H), 2.65 – 2.52 (m, 2H), 1.98 (d, J = 15.2 Hz, 3H), 1.94 (s, 3H), 1.68 (d, J = 7.2 Hz, 3H).



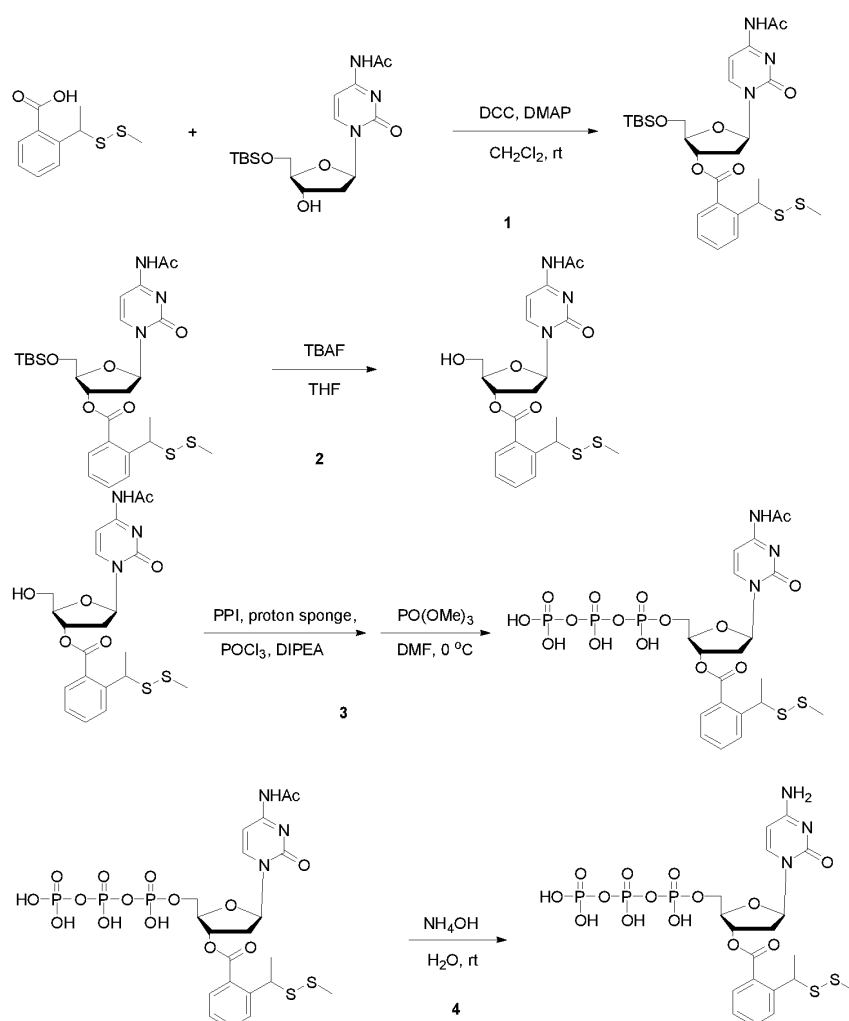
合成路线如下:



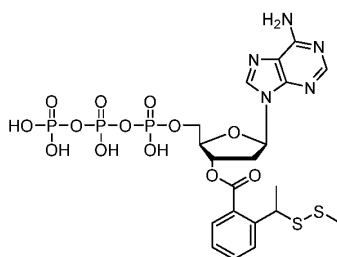
MW = 717. ^1H NMR (400 MHz, D_2O) δ 8.14 (t, $J = 3.1$ Hz, 1H), 7.82 (dd, $J = 8.1, 3.1$ Hz, 1H), 7.55 (q, $J = 3.2$ Hz, 2H), 7.38 - 7.34 (m, 1H), 6.31 - 6.22 (m, 1H), 5.76 (d, $J = 4.1$ Hz, 1H), 5.10 - 5.00 (m, 1H), 4.65 - 4.55 (m, 1H), 4.36 - 4.21 (m, 2H), 3.02 - 2.94 (m, 1H), 2.71 - 2.61 (m, 1H), 1.94 (dd, $J = 22.2, 3.2$ Hz, 3H), 1.62 (dd, $J = 7.2, 3.3$ Hz, 3H).



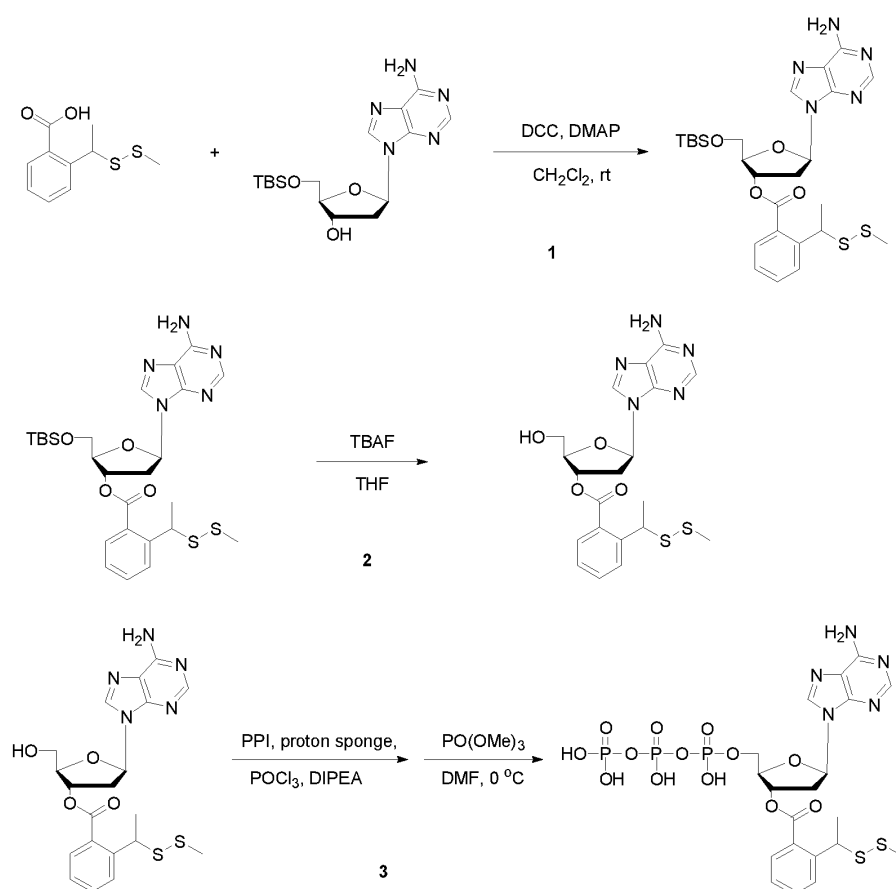
合成路线如下:



MW = 677. ^1H NMR (400 MHz, D_2O) δ 8.08 – 8.00 (m, 1H), 7.91 – 7.84 (m, 1H), 7.72 – 7.59 (m, 2H), 7.49 – 7.40 (m, 1H), 6.48 – 6.40 (m, 1H), 6.19 (dd, J = 7.7, 2.3 Hz, 1H), 5.73 – 5.64 (m, 1H), 5.13 – 5.02 (m, 1H), 4.60 (d, J = 12.6 Hz, 1H), 4.36 – 4.27 (m, 2H), 2.71 – 2.64 (m, 1H), 2.55 – 2.44 (m, 1H), 2.00 (dd, J = 9.5, 2.2 Hz, 3H), 1.69 (dd, J = 7.2, 2.1 Hz, 3H).



合成路线如下:

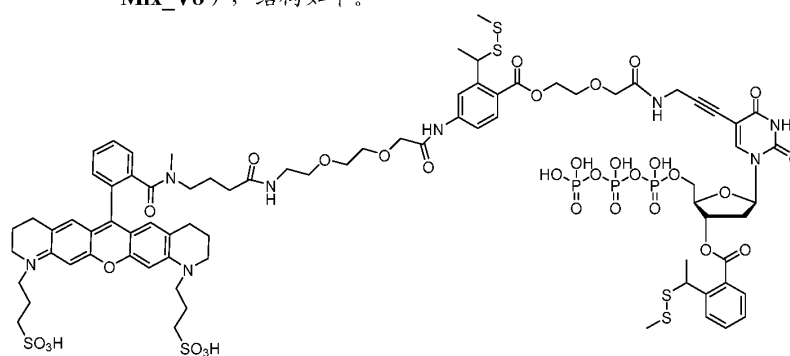


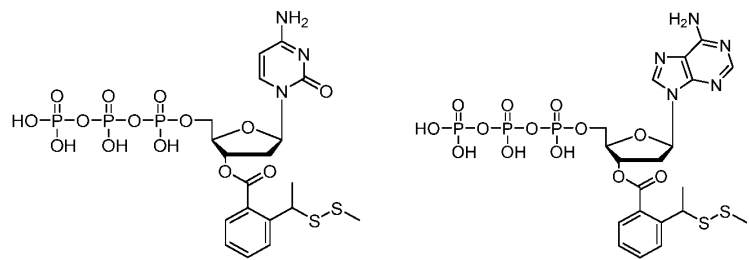
MW = 701. ^1H NMR (400 MHz, D_2O) δ 8.53 (dd, $J = 5.2, 1.8$ Hz, 1H), 8.18 – 8.10 (m, 1H), 7.83 (d, $J = 7.8$ Hz, 1H), 7.55 – 7.51 (m, 2H), 7.38 – 7.34 (m, 1H), 6.49 – 6.40 (m, 1H), 5.81 – 5.73 (m, 1H), 5.07 – 5.00 (m, 1H), 4.67 – 4.57 (m, 1H), 4.39 – 4.18 (m, 2H), 3.06 – 2.95 (m, 1H), 2.75 (ddd, $J = 21.2, 14.2, 5.6$ Hz, 1H), 1.93 (dd, $J = 25.1, 1.9$ Hz, 3H), 1.60 (dd, $J = 6.9, 4.0$ Hz, 3H).

测序例 1

核苷酸底物

(1) hot dNTP: 实施例 7-1 至 7-4 合成获得的四种终产物 (命名为 **SSEB Hot Mix_V8**), 结构如下。





按照 MGISEQ-2000 测序仪操作规程, 使用上述核苷酸底物以及 MGISEQ - 2000 高通量测序试剂套装 (FCL SE50) 进行测序。

1. 利用 Ecoli 测序文库制备 DNA 纳米球。
 2. 将 DNA 纳米球装载到 BGISEQ2000 测序芯片上。
 3. 将装载完成的测序芯片加载至 BGISEQ2000 测序仪上, 设置测序流程, hot dNTP 聚合: 60℃ 2min; cold dNTP 聚合: 60℃ 2min; 信号采集; 切除阻断基团: 65℃ 2min。
 4. 下机数据进行 basecall 分析, 输出测序指标、比对率、错误率、Q30 等, 结果见表 1。
- 结果表明, 此类核苷酸类似物在高通量测序中起到良好的可逆阻断效果, 可实现逐个检测测序中掺入的核苷酸类型。

表 1 Basecall 分析结果

核苷酸类似物 hot dNTP Mix	SSEB Hot Mix_V8
参考基因 (Reference)	Ecoli
循环数 (Cycle Number)	51
Q30 (%)	92.67
Lag (%)	0.43
Runon (%)	0.31
ESR (%)	82.60
比对率 (Mapping Rate, %)	98.15
平均错误率 (AvgError Rate, %)	0.50

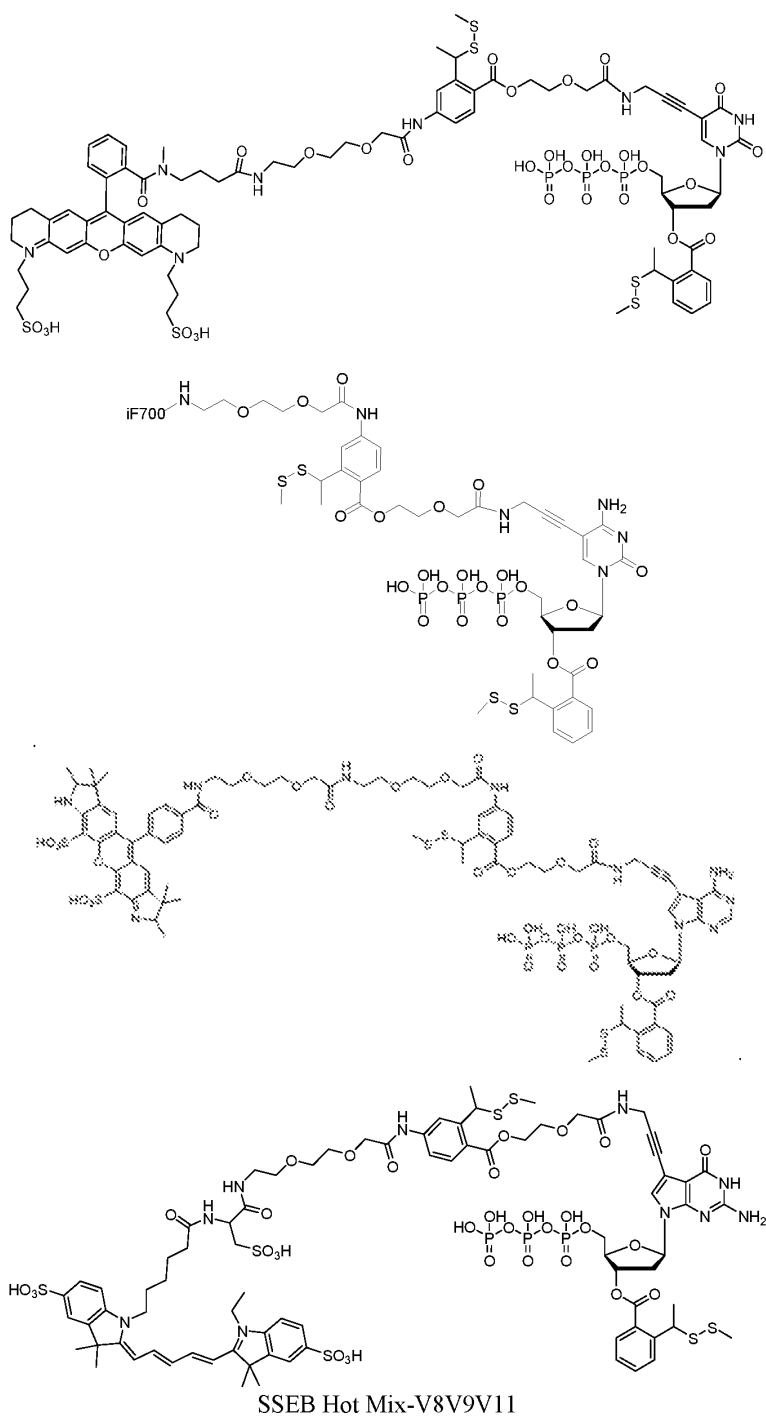
注:

- 1、Q30(%)指 Basecall 结果错误率低于 0.001(准确性高于 99.9%)的 base 占有比率;
- 2、Lag(%)指测序滞后所占比率;
- 3、Runon(%)指测序超前所占比率;
- 4、ESR(%)指经过过滤最终测序所得的 reads 数 (Total Reads) 占 Basecall 能识别出来的 DNB 数量 (DNB Number) 的比率。

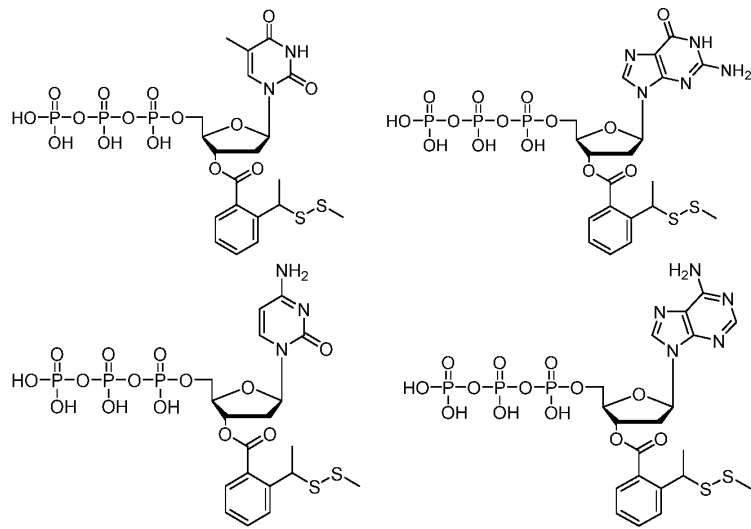
测序例 2

核苷酸底物

(1) hot dNTP: 实施例 7-1、7-4、11、15-2 合成获得的四种终产物 (命名为 SSEB Hot Mix-V8V9V11), 结构如下。



(2) cold dNTP: 通过常规方法合成四种 cold dNTP, 结构如下。



按照 MGISEQ-2000 测序仪操作规程, 使用上述核苷酸底物以及 MGISEQ - 2000 高通量测序试剂套装 (FCL PE100) 进行测序。

1. 利用 Ecoli 测序文库制备 DNA 纳米球;
2. 将 DNA 纳米球装载到 BGISEQ2000 测序芯片上;
3. 将装载完成的测序芯片加载至 BGISEQ2000 测序仪上, 设置测序流程, hot dNTP 聚合: 55℃ 30s; cold dNTP 聚合: 55℃ 60s; 信号采集; 切除阻断基团: 55℃ 8s。
4. 下机数据进行 basecall 分析, 输出测序指标、比对率、错误率、Q30 等, 结果见表 2。

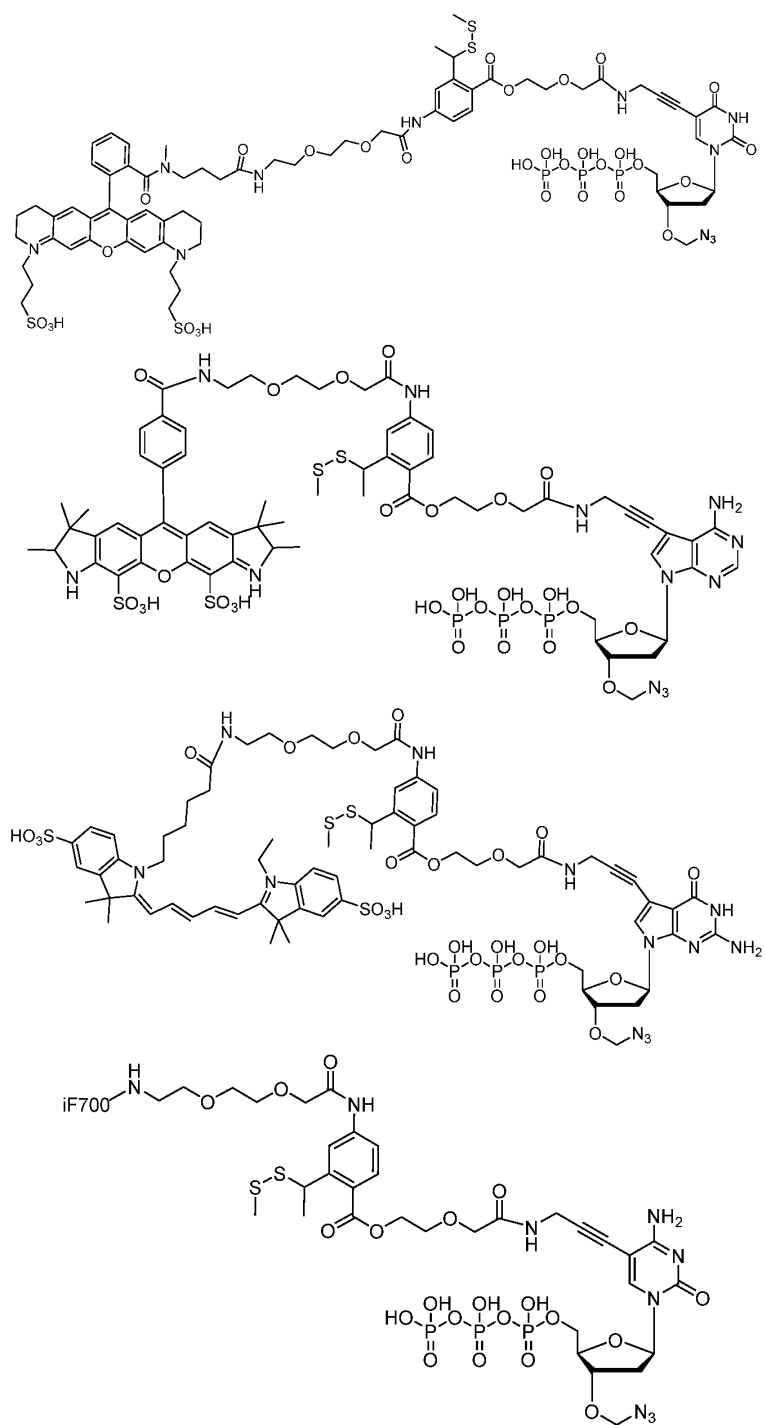
结果表明, 此类核苷酸类似物在高通量测序中起到良好的可逆阻断效果, 可实现逐个检测测序中掺入的核苷酸类型。

表 2 Basecall 分析结果

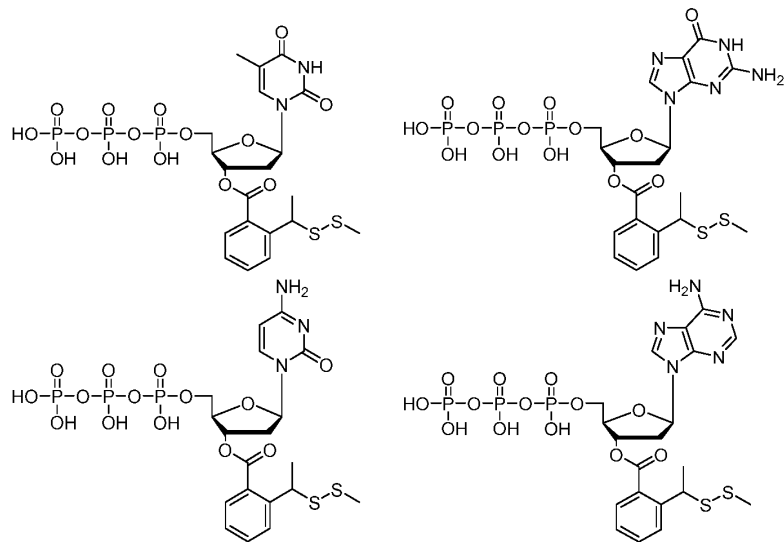
核苷酸类似物 hot dNTP Mix	SSEB Hot Mix-V8V9V11
参考基因 (Reference)	Ecoli
循环数 (Cycle Number)	212
Q30 (%)	95.79
Lag 1 (%)	0.08
Lag 2 (%)	0.11
Runon 1 (%)	0.12
Runon 2 (%)	0.26
ESR (%)	76.81
比对率 (Mapping Rate, %)	99.89
平均错误率 (AvgError Rate, %)	0.15

测序例 3
核苷酸底物

(1) hot dNTP: 实施例 10-1、10-2、10-3、10-4 合成获得的四种终产物 (命名为 SSEB Hot Mix-N₃), 结构如下。

SSEB Hot Mix-N₃

(2) cold dNTP: 通过常规方法合成四种 cold dNTP, 结构如下。



按照 MGISEQ-2000 测序仪操作规程, 使用上述核苷酸底物以及 MGISEQ - 2000 高通量测序试剂套装 (FCL PE100) 进行测序。

1. 利用 Ecoli 测序文库制备 DNA 纳米球;
2. 将 DNA 纳米球装载到 BGISEQ2000 测序芯片上;
3. 将装载完成的测序芯片加载至 BGISEQ2000 测序仪上, 设置测序流程, hot dNTP 聚合: 60℃ 2min; cold dNTP 聚合: 60℃ 2min; 信号采集; 切除阻断基团: 65℃ 2min。
4. 下机数据进行 basecall 分析, 输出测序指标、比对率、错误率、Q30 等, 结果见表 3。

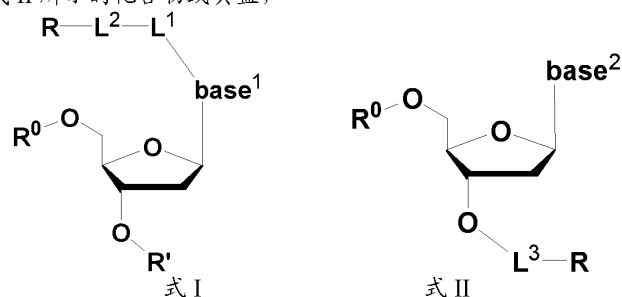
结果表明, 此类核苷酸类似物在高通量测序中起到良好的可逆阻断效果, 可实现逐个检测测序中掺入的核苷酸类型。

表 3 Basecall 分析结果

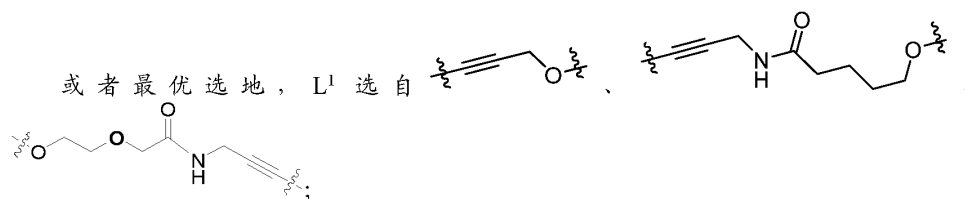
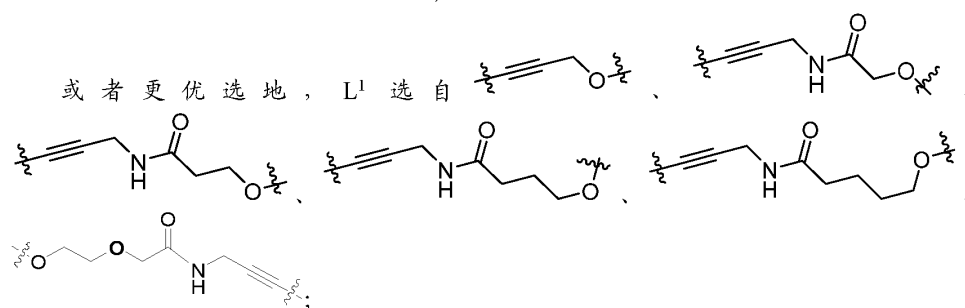
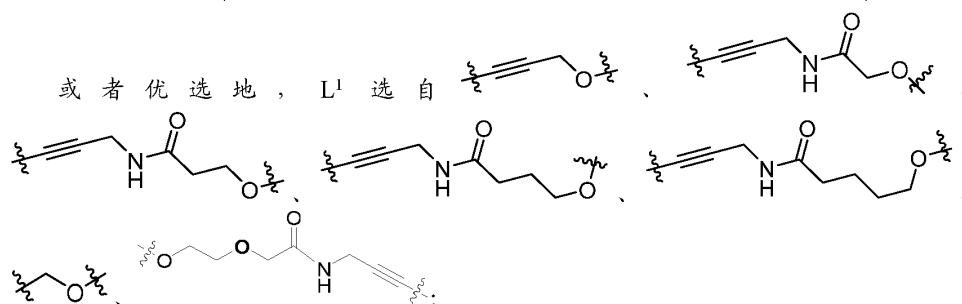
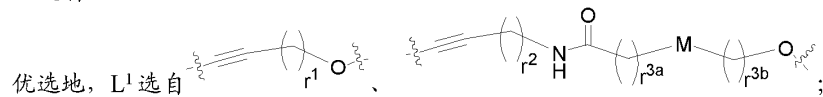
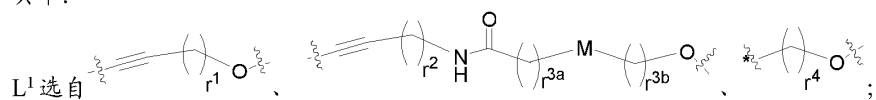
核苷酸类似物 hot dNTP Mix	SSEB Hot Mix-N ₃
参考基因 (Reference)	Ecoli
循环数 (Cycle Number)	53
Q30 (%)	95.45
Lag 1 (%)	0.20
Lag 2 (%)	0.22
Runon 1 (%)	0.43
Runon 2 (%)	0.45
ESR (%)	80.29
比对率 (Mapping Rate, %)	96.82
平均错误率 (AvgError Rate, %)	0.57

权利要求

1、式 I 或式 II 所示的化合物或其盐，



其中：

 r^1 、 r^2 、 r^{3a} 、 r^{3b} 、 r^4 各自独立地选自 1-6 之间的任意整数；优选地， r^1 选自 1、2、3；更优选地， r^1 为 1；优选地， r^2 选自 1、2、3；更优选地， r^2 为 1；优选地， r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5；

更优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3;

最优地, r^{3a} 为 1;

最优地, r^{3b} 为 2;

优选地, r^4 选自 1、2、3;

更优选地, r^4 为 1;

M 选自直接键、 CH_2 、NH、O、S;

优选地, M 选自 CH_2 、O;

L^2 为 , 优选地, W 与 R 相连;

L^3 为 , 优选地, W 与 R 相连;

各 X 独立地选自 O、S、NH;

优选地, 各 X 独立地选自 O、S;

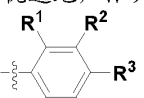
更优选地, X 为 O;

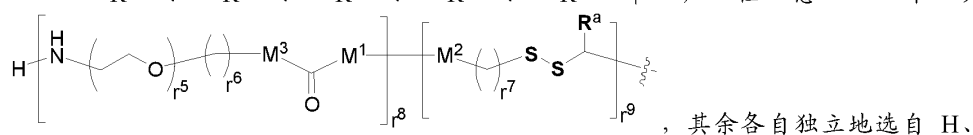
Y 选自直接键、O、S、NH;

优选地, Y 为直接键;

各 W 独立地选自直接键、O、S、NH;

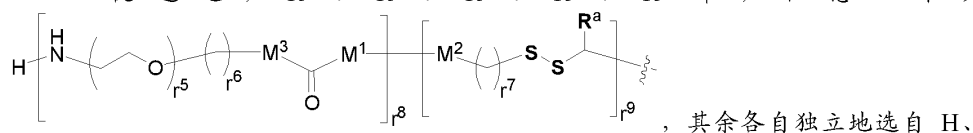
优选地, W 为直接键;

;
R 为 R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个为



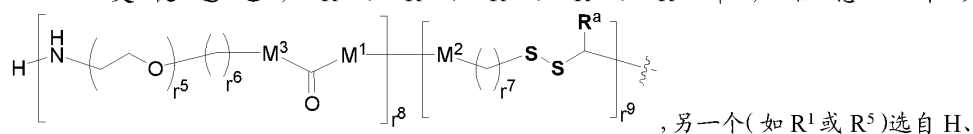
R^b 、 R^c 、叠氨基、硝基、氨基、磺基、羧基、脂肪族烷基(如 C1-C6 烷基)、环烷基(如 C3-C6 环烷基)、芳香族烷基(如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基(如 C1-C6 烷氧基);

优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个为



R^b 、 R^c ;

更优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个为



R^b 、 R^c ;

, 剩余三个为 H;

r^5 、 r^6 、 r^7 各自独立地选自 1-6 之间的任意整数;

优选地, r^5 选自 1、2、3;

更优选地, r^5 为 2;

优选地, r^6 选自 1、2、3;

更优选地, r^6 为 1 或 2;

优选地, r^7 选自 1、2、3;

更优选地, r^7 为 2;

r^8 、 r^9 各自独立地选自 0、1, 且 r^8 、 r^9 不同时为 0;

M^1 、 M^2 、 M^3 各自独立地选自直接键、NH、O、S;

优选地, M^1 选自直接键、NH、O;

优选地, M^2 为 NH;

优选地, M^3 选自直接键、NH;

R^a 选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基);

优选地, R^a 选自 H、C1-C6 烷基;

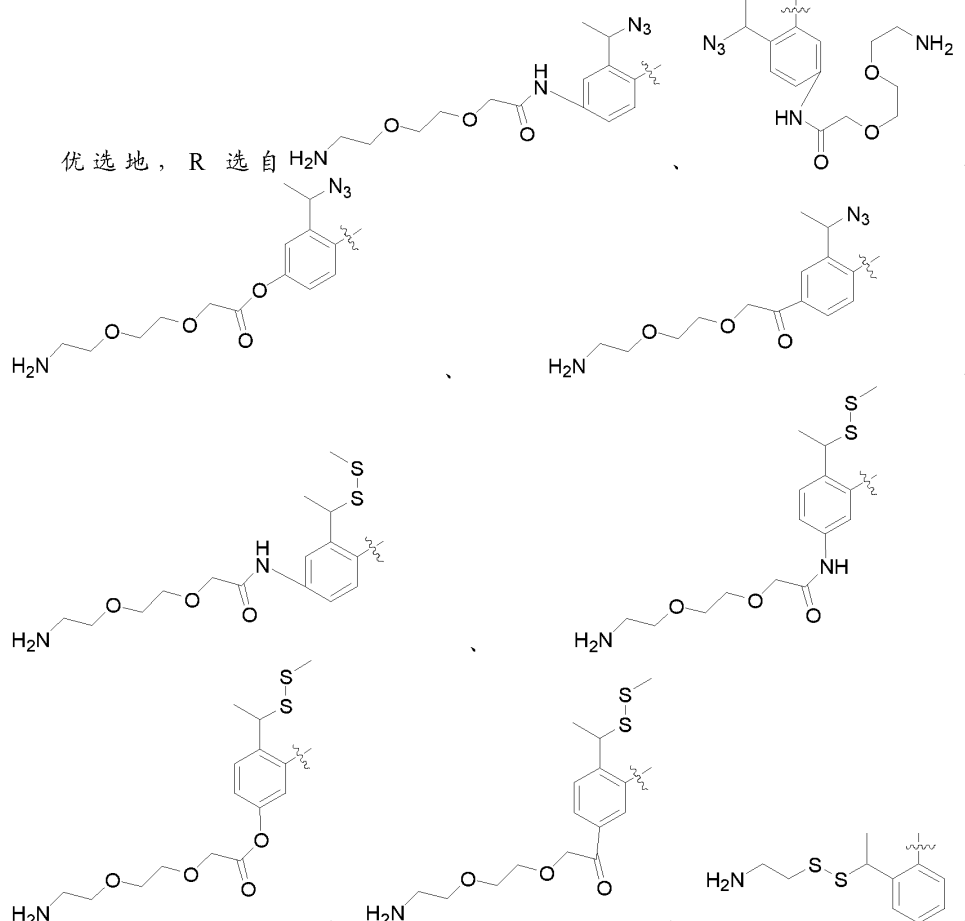
更优选地, R^a 为甲基;

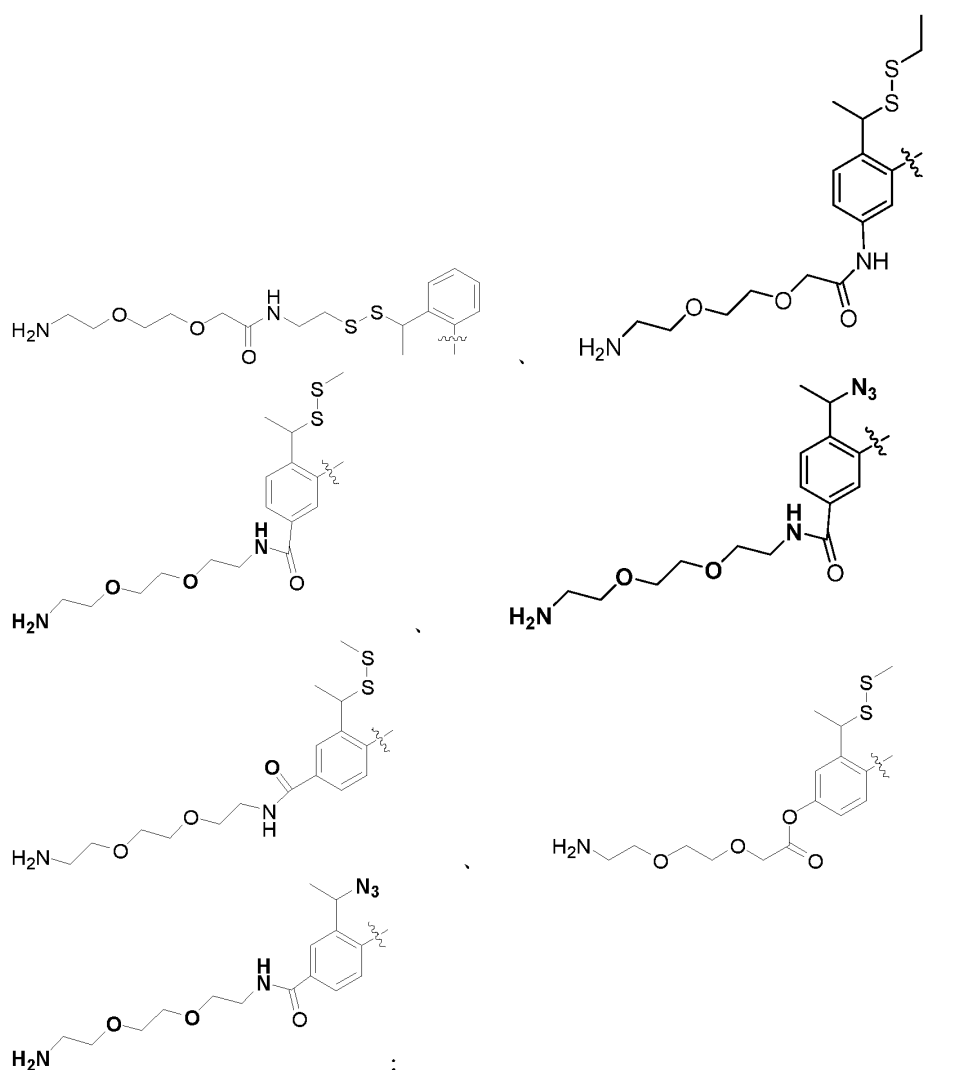
R^b 、 R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu)、-ONH₂、-OCOR_x、-OCONHR_x, 另一个选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基 (如 C1-C6 烷基), 环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基);

优选地, R^b 、 R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu), 另一个选自 C1-C6 烷基;

更优选地, R^b 、 R^c 中, 任意一个选自 -N₃、-SS-Me、-SS-Et, 另一个为甲基;

最优选地, R^b 、 R^c 中, 任意一个选自 -N₃、-SS-Me, 另一个为甲基;





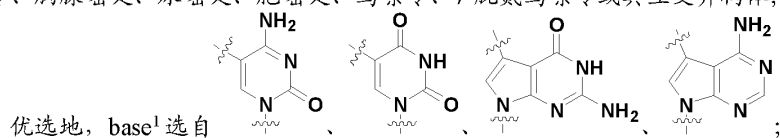
R^0 选自 H、单磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-$)、二磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)、三磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)、四磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)；

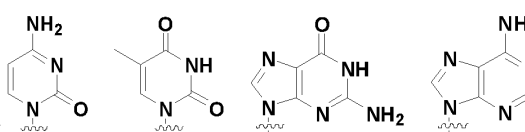
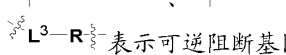
优选地， R^0 为三磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)；

各 Z 独立地选自 O, S, BH；

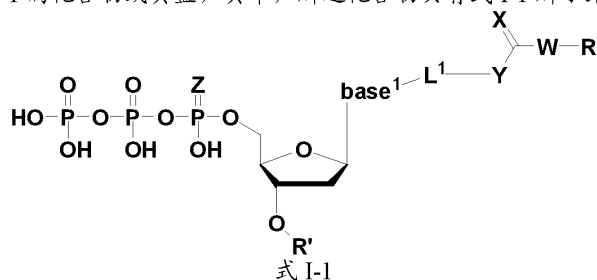
优选地，Z 为 O；

base^1 、 base^2 各自独立地选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；

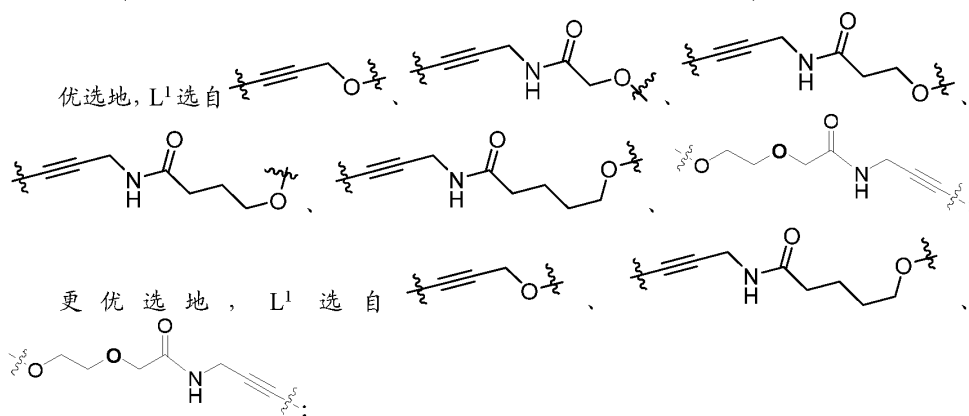
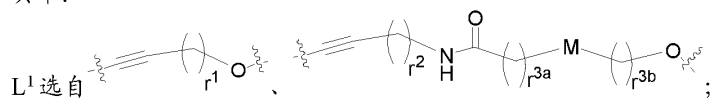


优选地, base² 选自  ;
式 I 中, R' 表示可逆阻断基团, 式 II 中,  表示可逆阻断基团。

2、权利要求 1 的化合物或其盐, 其中, 所述化合物具有式 I-1 所示结构,



其中:



r¹、r²、r^{3a}、r^{3b} 各自独立地选自 1-6 之间的任意整数;

优选地, r¹ 选自 1、2、3;

更优选地, r¹ 为 1;

优选地, r² 选自 1、2、3;

更优选地, r² 为 1;

优选地, r^{3a}、r^{3b} 各自独立地选自 1、2、3、4、5;

更优选地, r^{3a}、r^{3b} 各自独立地选自 1、2、3;

最优选地, r^{3a} 为 1;

最优选地, r^{3b} 为 2;

M 选自直接键、CH₂、NH、O、S;

优选地, M 选自 CH₂、O;

X 选自 O、S、NH;

优选地, X 选自 O、S;

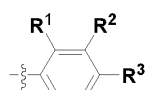
更优选地, X 为 O;

Y 选自直接键、O、S、NH;

优选地, Y 为直接键;

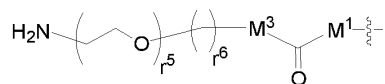
W 选自直接键、O、S、NH;

优选地, W 为直接键;



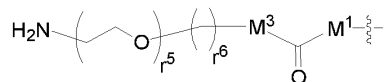
R 为 R^5 R^4 ;

R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^2 、 R^3 或 R^4) 为



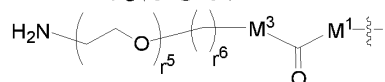
, 其余各自独立地选自 H、 R^b R^c 、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基)、芳香族烷基 (如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基 (如 C1-C6 烷氧基);

优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^2 、 R^3 或 R^4) 为



, 其余各自独立地选自 H、 R^b R^c ;

更优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^2 、 R^3 或 R^4) 为



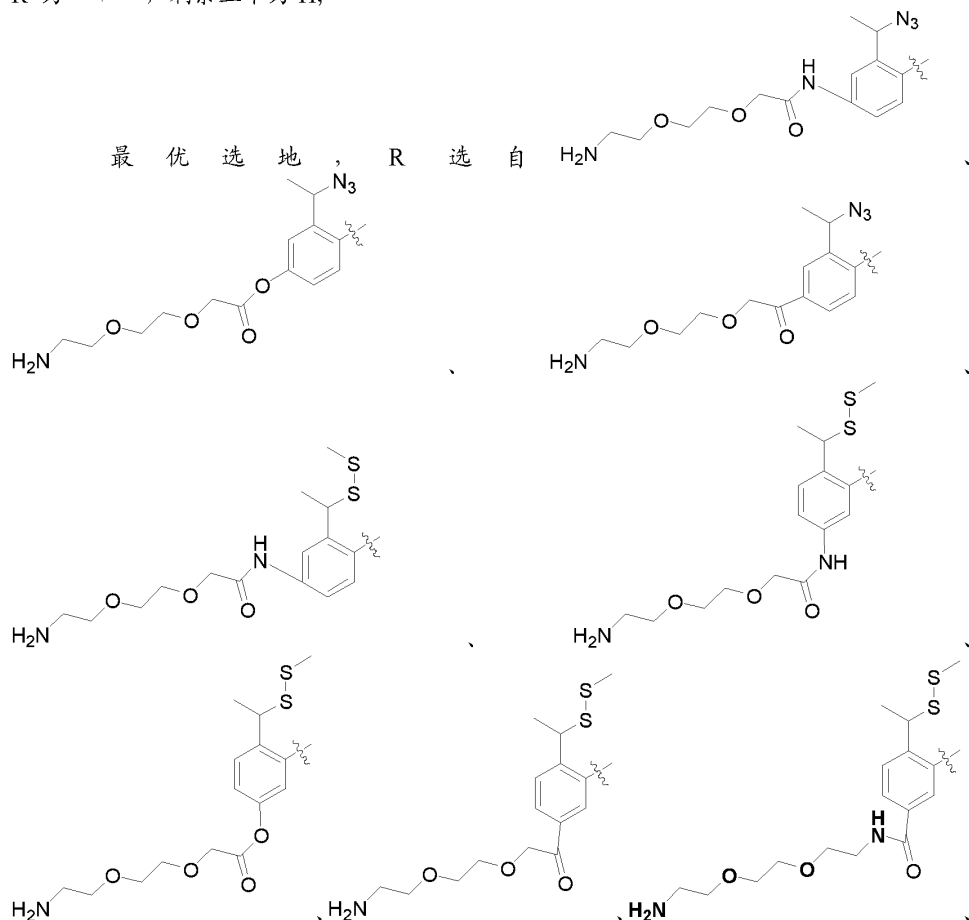
, 另一个 (如 R^1 或 R^5) 为 R^b R^c , 剩余三个为 H;

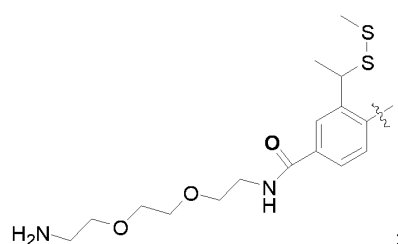
进一步优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, R^3 或 R^4 为



R^1 为 R^b R^c , 剩余三个为 H;

最优选地, R 选自 H_2N -(CH2)5-O-(CH2)6-M3-C(=O)-M1-、





r^5 、 r^6 各自独立地选自1-6之间的任意整数；

优选地， r^5 选自1、2、3；

更优选地， r^5 为2；

优选地， r^6 选自1、2、3；

更优选地， r^6 为1或2；

M^1 、 M^3 各自独立地选自直接键、NH、O、S；

优选地， M^1 选自直接键、NH、O；

优选地， M^3 选自直接键、NH；

更优选地， M^1 为NH、O或S时， M^3 为直接键，且 M^3 为NH、O或S时， M^1 为直接键；

R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基（如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$ ）、 $-ONH_2$ 、 $-OCOR_x$ 、 $-OCONHR_x$ ，另一个选自H、脂肪族烷基（如C1-C6烷基，例如Me、Et、iPr、tBu）、芳香族烷基（如苯基-C1-C6烷基）、环烷基（如C3-C6环烷基），其中，各 R_x 独立地选自脂肪族烷基（如C1-C6烷基），环烷基（如C3-C6环烷基）或芳香族烷基（如苯基C1-C6烷基）；

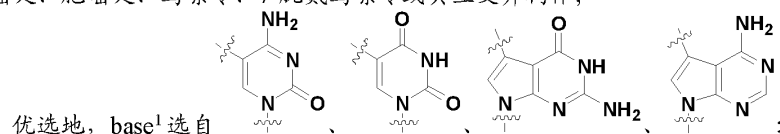
优选地， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基（如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$ ），另一个选自C1-C6烷基；

更优选地， R^b 、 R^c 中，任意一个选自 $-N_3$ 、 $-SS-Me$ ，另一个为甲基；

Z选自O、S、BH；

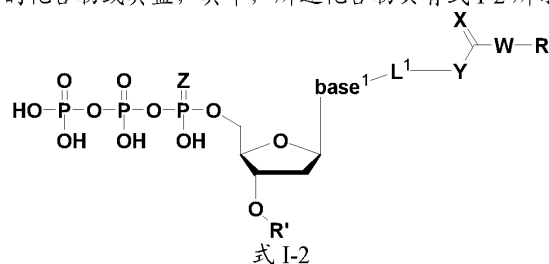
优选地，Z为O；

$base^1$ 选自碱基、脱氮碱基或其互变异构体，例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；



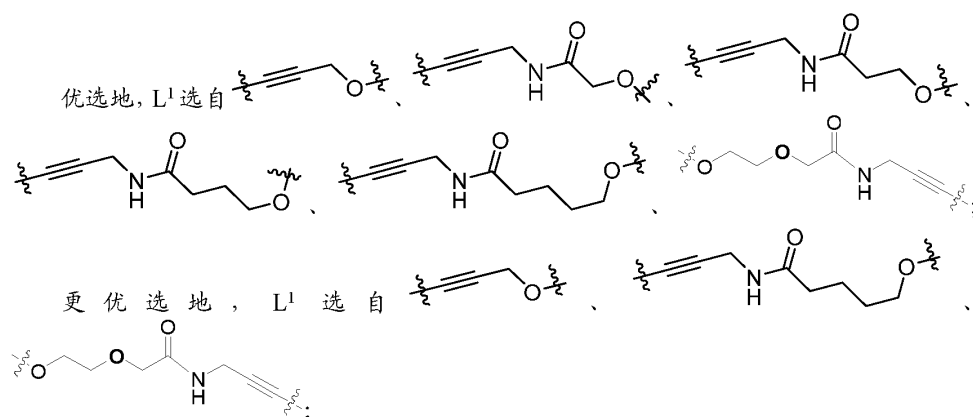
优选地， $base^1$ 选自
R'表示可逆阻断基团。

3、权利要求1的化合物或其盐，其中，所述化合物具有式I-2所示结构，



其中：





r^1 、 r^2 、 r^{3a} 、 r^{3b} 各自独立地选自 1-6 之间的任意整数;

优选地, r^1 选自 1、2、3;

更优选地, r^1 为 1;

优选地, r^2 选自 1、2、3;

更优选地, r^2 为 1;

优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5;

更优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3;

最优选地, r^{3a} 为 1;

最优选地, r^{3b} 为 2;

M 选自直接键、 CH_2 、NH、O、S;

优选地, M 选自 CH_2 、O;

X 选自 O、S、NH;

优选地, X 选自 O、S;

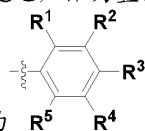
更优选地, X 为 O;

Y 选自直接键、O、S、NH;

优选地, Y 为直接键;

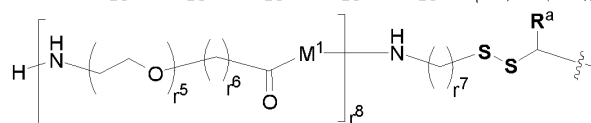
W 选自直接键、O、S、NH;

优选地, W 为直接键;



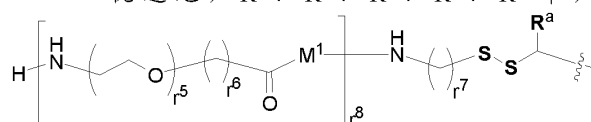
R 为 R^5 ;

R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^1 或 R^5) 为



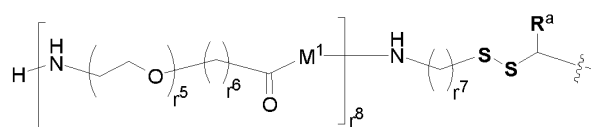
, 其余各自独立地选自 H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基)、芳香族烷基 (如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基 (如 C1-C6 烷氧基);

优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^1 或 R^5) 为

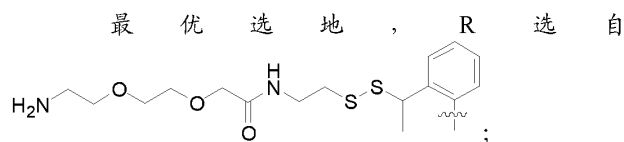


, 其余为 H;

更优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, R^1 为



, 其余为 H;



r^5 、 r^6 、 r^7 各自独立地选自 1-6 之间的任意整数;

优选地, r^5 选自 1、2、3;

更优选地, r^5 为 2;

优选地, r^6 选自 1、2、3;

更优选地, r^6 为 1;

优选地, r^7 选自 1、2、3;

更优选地, r^7 为 2;

r^8 选自 0、1;

M^1 选自直接键、NH、O、S;

优选地, M^1 为直接键;

R^a 选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基);

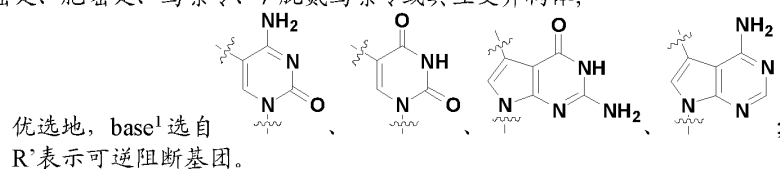
优选地, R^a 选自 C1-C6 烷基;

更优选地, R^a 为甲基;

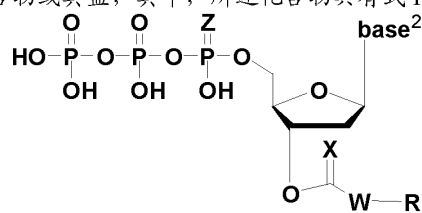
Z 选自 O、S、BH;

优选地, Z 为 O;

$base^1$ 选自碱基、脱氮碱基或其互变异构体, 例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;



4、权利要求 1 的化合物或其盐, 其中, 所述化合物具有式 II-1 所示结构,



式 II-1

其中:

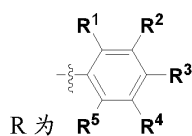
X 选自 O、S、NH;

优选地, X 选自 O、S;

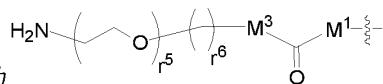
更优选地, X 为 O;

W 选自直接键、O、S、NH;

优选地, W 为直接键;

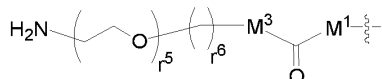


R^1, R^2, R^3, R^4, R^5 中, 任意一个(如 R^3 或 R^4)为



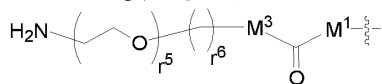
其余各自独立地选自 H、、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基（如 C1-C6 烷基）、环烷基（如 C3-C6 环烷基）、芳香族烷基（如苯基-C1-C6 烷基）、F、I、Br、Cl、烷氧基（如 C1-C6 烷氧基）；

优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^3 或 R^4) 为

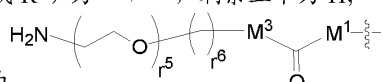


, 其余各自独立地选自 H、 ;

更优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个 (如 R^3 或 R^4) 为

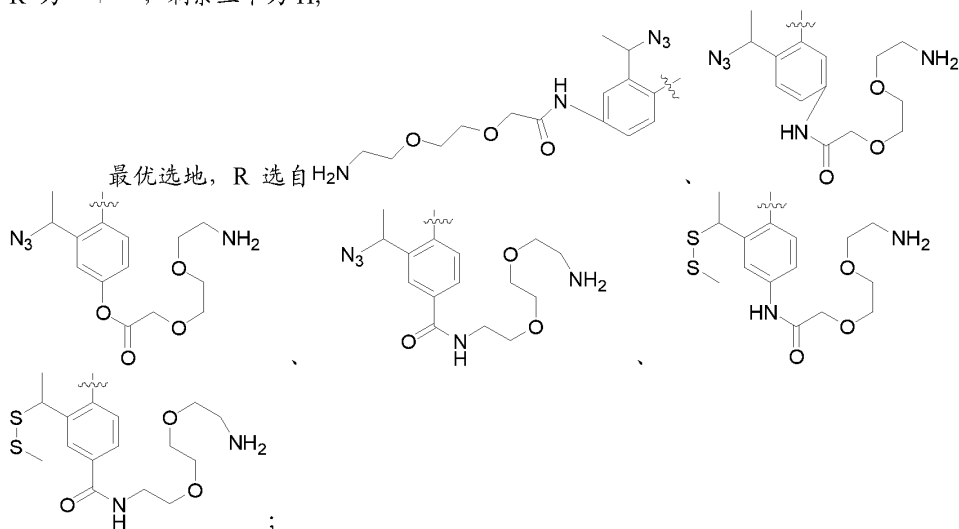


, 另一个 (如 R^1 或 R^5) 为 , 剩余三个为 H;



进一步优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, R^3 或 R^4 为

R^1 为 ，剩余三个为 H；



r^5 、 r^6 各自独立地选自 1-6 之间的任意整数;

优选地, r^5 选自 1、2、3;

更优选地, r^5 为 2;

优选地, r^6 选自 1、2、3;

更优选地, r^6 选自 1、2;

M¹、M³各自独立地选自直接键、NH、O、S;

优选地, M^1 选自直接键、NH、O;

优选地, M^3 选自直接键、NH;

R^b、R^c中,任意一个选自-N₃、-SS-C1-C6 烷基(如-SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu)、-ONH₂、-OCOR_x、-OCONHR_x,另一个选自H、脂肪族烷基(如C1-C6 烷基,例如Me、Et、iPr、tBu)、芳香族烷基(如苯基-C1-C6 烷基)、环烷基(如C3-C6 环烷基),其中,各

R_x独立地选自脂肪族烷基(如 C1-C6 烷基), 环烷基(如 C3-C6 环烷基)或芳香族烷基(如苯基 C1-C6 烷基);

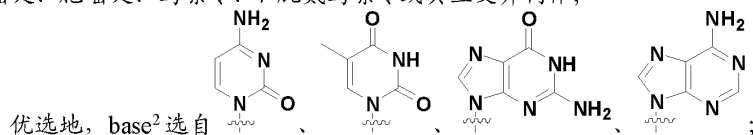
优选地, R^b、R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基(如 -SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu), 另一个选自 C1-C6 烷基;

更优选地, R^b、R^c 中, 任意一个为 -N₃ 或 -SS-Me, 另一个为甲基;

Z 选自 O, S, BH;

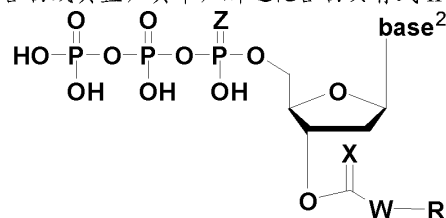
优选地, Z 为 O;

base² 选自碱基、脱氮碱基或其互变异构体, 例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;



$\begin{array}{c} \text{X} \\ \parallel \\ \text{---} \text{W} \text{---} \text{R} \end{array}$ 表示可逆阻断基团。

5、权利要求 1 的化合物或其盐, 其中, 所述化合物具有式 II-2 所示结构,



式 II-2

其中:

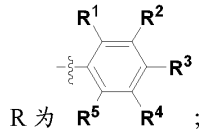
X 选自 O、S、NH;

优选地, X 选自 O、S;

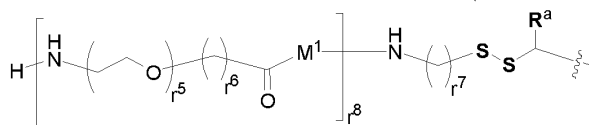
更优选地, X 为 O;

W 选自直接键、O、S、NH;

优选地, W 为直接键;

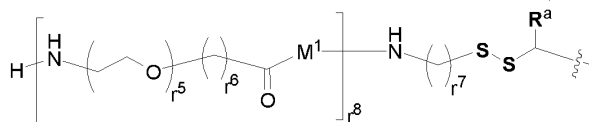


R¹、R²、R³、R⁴、R⁵ 中, 任意一个(如 R¹ 或 R⁵) 为



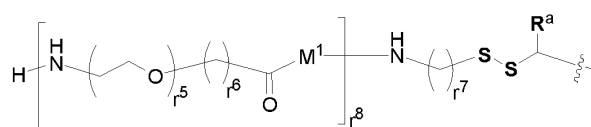
, 其余各自独立地选自 H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基(如 C1-C6 烷基)、环烷基(如 C3-C6 环烷基)、芳香族烷基(如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基(如 C1-C6 烷氧基);

优选地, R¹、R²、R³、R⁴、R⁵ 中, 任意一个(如 R¹ 或 R⁵) 为

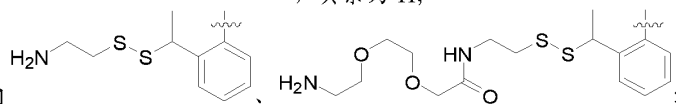


, 其余为 H;

更优选地, R¹、R²、R³、R⁴、R⁵ 中, R¹ 为



, 其余为 H;



最优选地, R 选自

r^5 、 r^6 、 r^7 选自 1-6 之间的任意整数;

优选地, r^5 选自 1、2、3;

更优选地, r^5 为 2;

优选地, r^6 选自 1、2、3;

更优选地, r^6 为 1;

优选地, r^7 选自 1、2、3;

更优选地, r^7 为 2;

r^8 选自 0、1;

M^1 选自直接键、NH、O、S;

优选地, M^1 为直接键;

R^a 选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基);

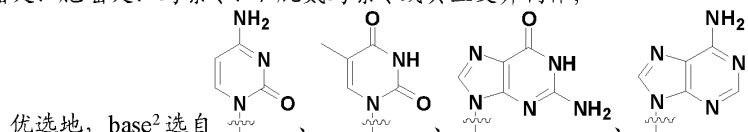
优选地, R^a 选自 C1-C6 烷基;

更优选地, R^a 为甲基;

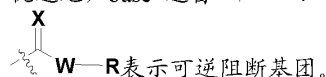
Z 选自 O、S、BH;

优选地, Z 为 O;

$base^2$ 选自碱基、脱氮碱基或其互变异构体, 例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;



优选地, $base^2$ 选自



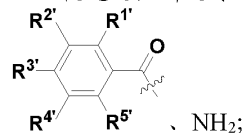
W—R 表示可逆阻断基团。

6、权利要求 1-3 任一项所述的化合物或其盐, 其中, 所述可逆阻断基团 R' 选自 N_3 -C1-C6



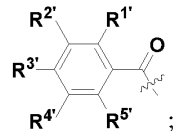
烷基、C1-C6 烷基-SS-C1-C6 烷基、 NH_2 、 $-ONH_2$ 、 $-OCOR_z$ 、 $-OCONHR_z$, 其中, 各 R_z 独立地选自脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基);

优选地, 所述可逆阻断基团 R' 选自 N_3 -C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、

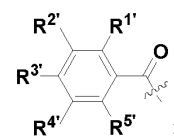


NH_2 ;

更优选地, 所述可逆阻断基团 R' 选自 N_3 -C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、



;



最优选地, 所述可逆阻断基团 R' 选自 N_3-CH_2- 、 $CH_3-CH_2-S-S-CH_2-$ 、 $R^{b'}$ 、 $R^{c'}$;

$R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中, 任意一个 (如 $R^{1'}$ 或 $R^{5'}$) 为 $R^{b'}$ 、 $R^{c'}$, 其余各自独立地选自 H、叠氨基、硝基、氨基、磺基、羧基、脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基)、芳香族烷基 (如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基 (如 C1-C6 烷氧基),



优选地, $R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中, 任意一个 (如 $R^{1'}$ 或 $R^{5'}$) 为 $R^{b'}$ 、 $R^{c'}$, 其余为 H,



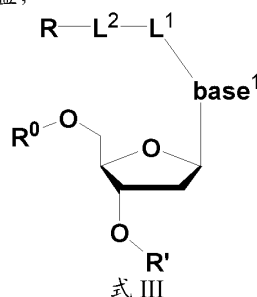
更优选地, $R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中, $R^{1'}$ 为 $R^{b'}$ 、 $R^{c'}$, 其余为 H,

$R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基 (如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$)、 $-ONH_2$ 、 $-OCOR_x$ 、 $-OCONHR_x$, 另一个选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基 (如 C1-C6 烷基), 环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基),

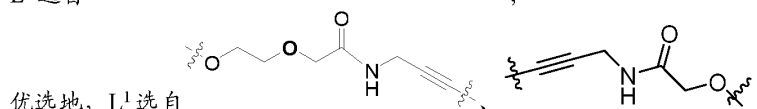
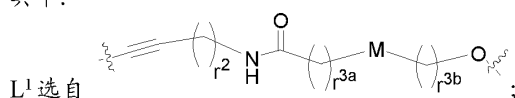
优选地, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-C1-C6$ 烷基 (如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$), 另一个选自 C1-C6 烷基,

更优选地, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-Me$, 另一个为甲基。

7、式 III 所示的化合物或其盐,



其中:



优选地, L^1 选自

r^2 选自 1、2、3;

优选地, r^2 为 1;

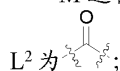
r^{3a} 选自 1、2、3;

优选地, r^{3a} 为 1;

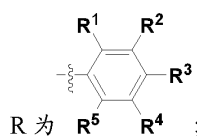
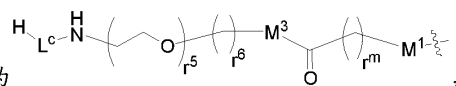
r^{3b} 选自 0、1、2、3;

优选地, r^{3b} 选自 0、2;

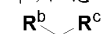
M 选自直接键、O;



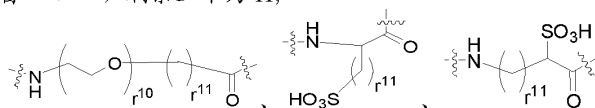
L^2 为 $R^{b'}$ 、 $R^{c'}$;

 R^1, R^2, R^3, R^4, R^5 中, 任意一个(如 R^3)为

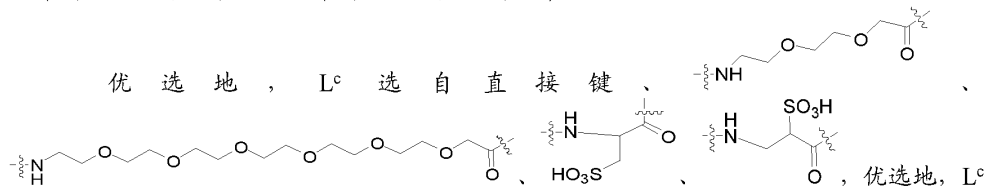
另一个 (如 R^1 或 R^2) 选自 ，剩余三个为 H;



L^c 选自直接键、 $\begin{array}{c} \text{NH} \\ | \\ \text{---} \end{array}$, $\begin{array}{c} \text{r}^{10} \quad \text{r}^{11} \\ \diagdown \quad \diagup \\ \text{O} \end{array}$ 、 HO_3S — $\begin{array}{c} \text{r}^{11} \\ | \\ \text{O} \end{array}$ 、 $\begin{array}{c} \text{r}^{11} \\ | \\ \text{O} \end{array}$ ，优选地， L^c 中的NH端与H相连， L^c 中的=O端与NH相连；



优选地， L^c 选自直接键、



中的 NH 端与 H 相连, L^c 中的=O 端与 NH 相连;

 r^m 选自 0、1、2、3;

优选地, r^m 选自 0、1;

r^5 选自 1、2、3;

优选地, r^5 为 2;

 r^6 选自 1、2、3;

优选地, r^6 为 1 或 2;

r^{10} 选自 1-10 之间的任意整数;

优选地, r^{10} 选自 2-6 之间的任意整数;

更优选地, r^{10} 为 2 或 6;

r^{11} 选自 1-6 之间的任意整数;

优选地, r^{11} 选自 1、2、3;

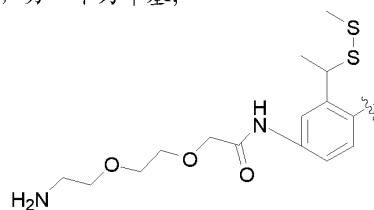
更优选地, r^{11} 为 1;

M^1 选自 NH 、 O ;

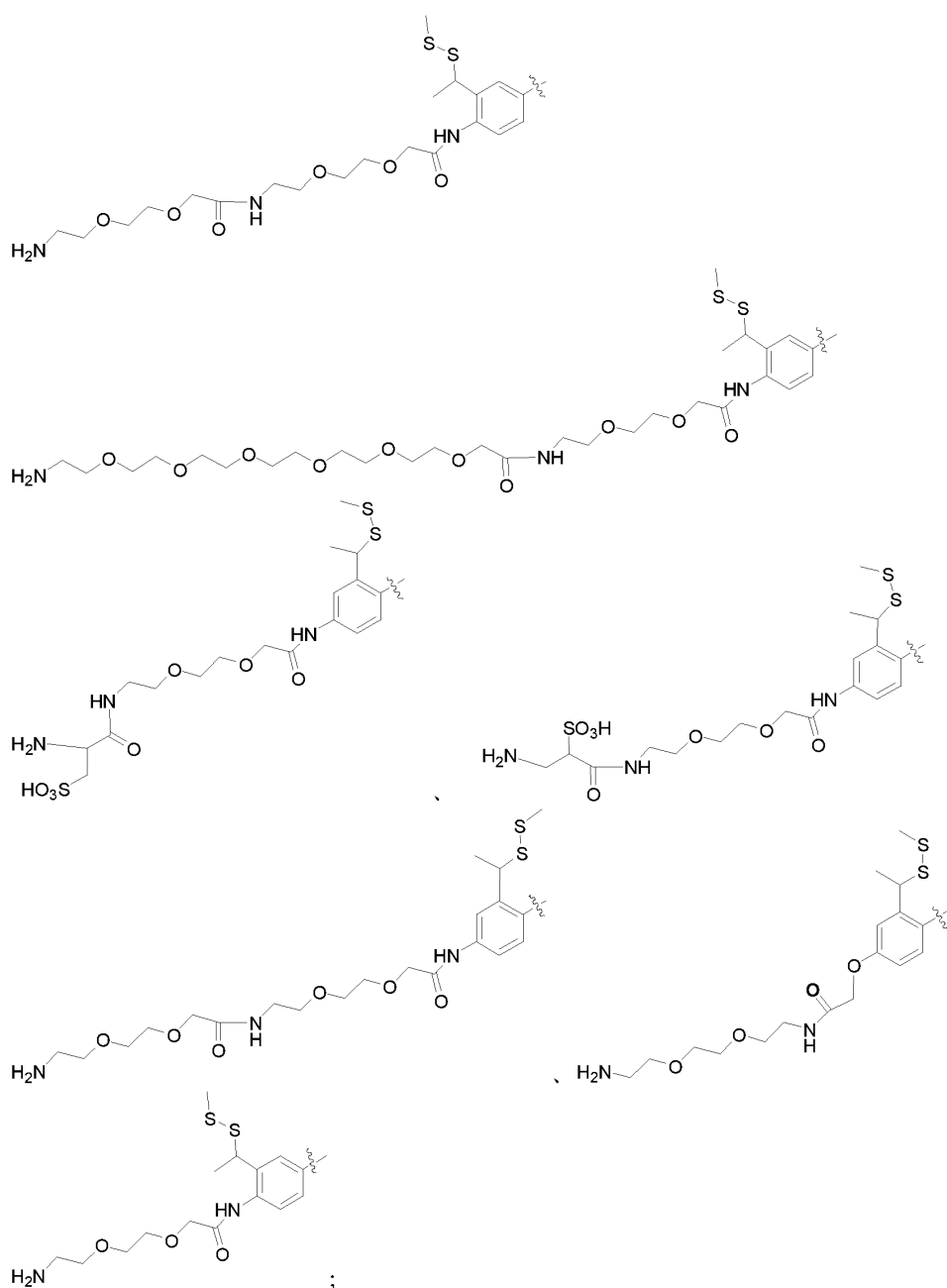
M³选自直接键、NH;

R^b、R^c 中，任意一个选自 -SS-C1-C6 烷基（如 -SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu），另一个选自 C1-C6 烷基；

优选地, R^b 、 R^c 中, 任意一个为 -SS-Me, 另一个为甲基;



优选地，R选自 H_2N



R^0 选自 H、单磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-$)、二磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)、三磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)、四磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)；

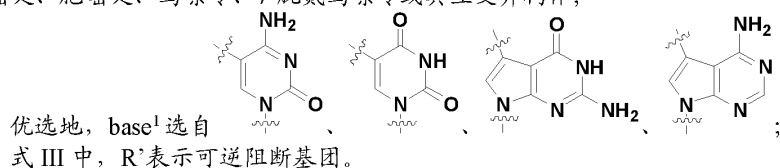
优选地， R^0 为三磷酸基团 ($\text{HO}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-\text{O}-\text{P}(\text{OH})(\text{Z})-$)；

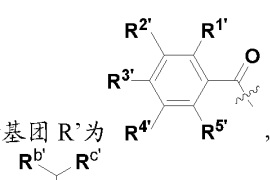
各 Z 独立地选自 O, S, BH；

优选地，Z 为 O；

base^1 选自碱基、脱氧碱基或其互变异构体，例如腺嘌呤、7-脱氧腺嘌呤、胸腺嘧啶、

尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体；



8、权利要求 7 所述的化合物或其盐，其中，所述可逆阻断基团 R' 为  ，

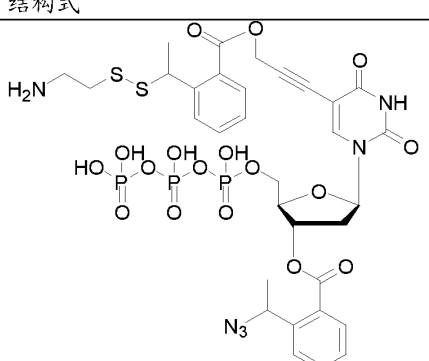
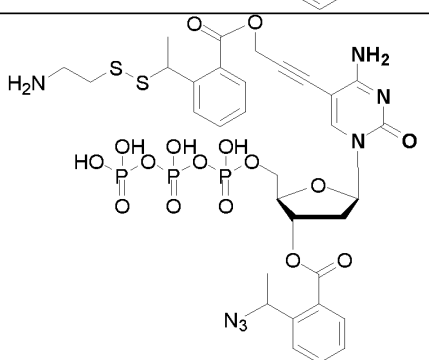
R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'} 中，任意一个（如 R^{1'} 或 R^{5'}）为  ，另一个（如 R^{3'}）选自 H、C1-C6 烷氧基（如甲氧基），其余为 H，

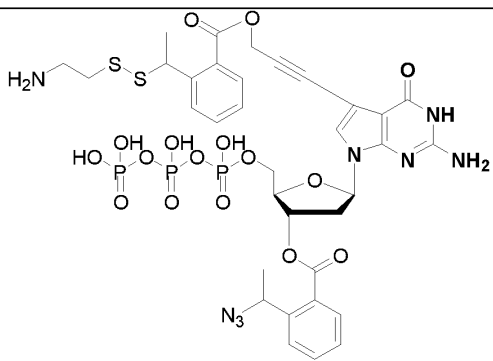
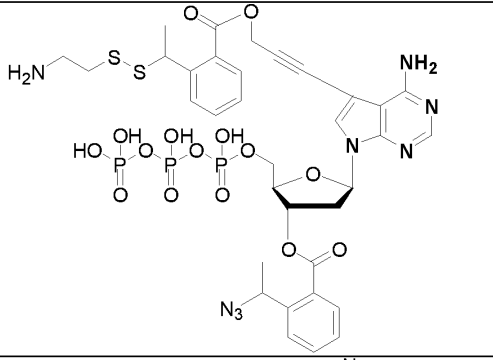
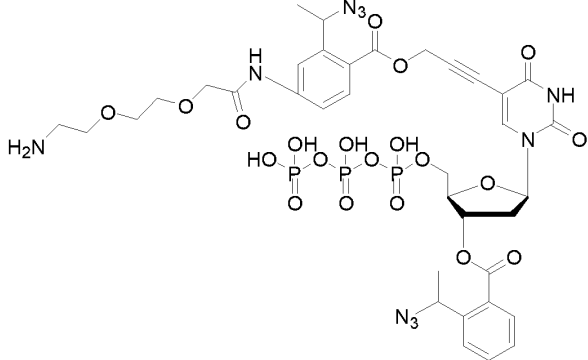
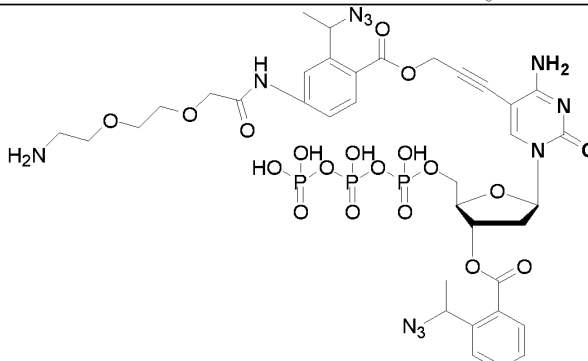
优选地，R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'} 中，R^{1'} 为  ，R^{3'} 选自 H、C1-C6 烷氧基（如甲氧基），其余为 H，

R^{b'}、R^{c'} 中，任意一个选自 -N₃、-SS-C1-C6 烷基（如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu），另一个选自 C1-C6 烷基，

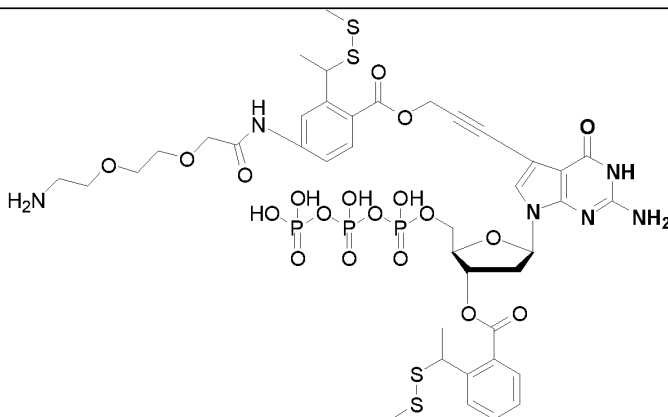
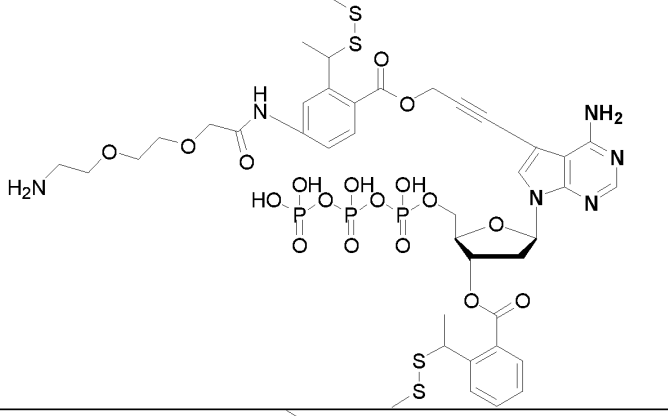
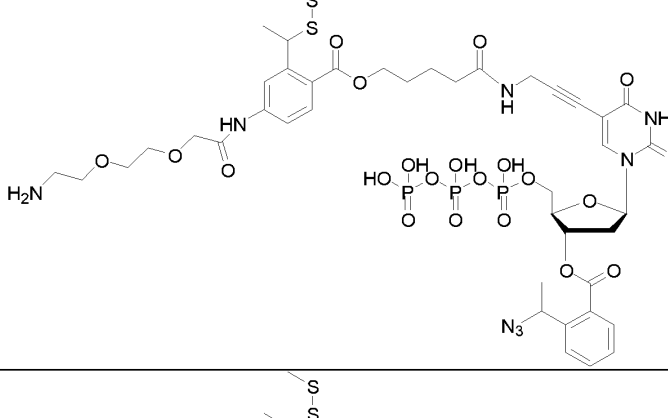
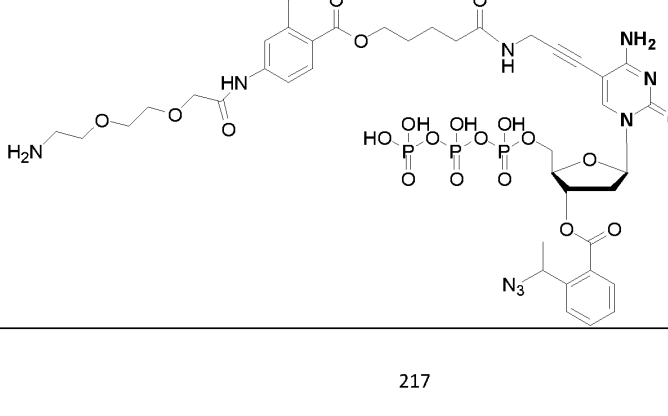
优选地，R^{b'}、R^{c'} 中，任意一个为 -N₃、-SS-Me，另一个为甲基。

9、权利要求 1-8 任一项所述的化合物或其盐，其中，所述化合物选自以下：

编号	结构式
1A	
1B	

1C	
1D	
2A	
2B	

2C	
2D	
3A	
3B	

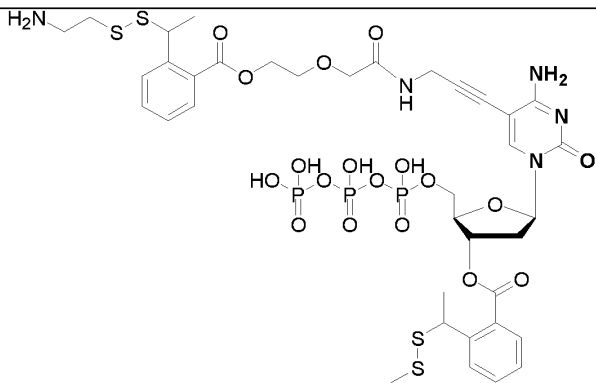
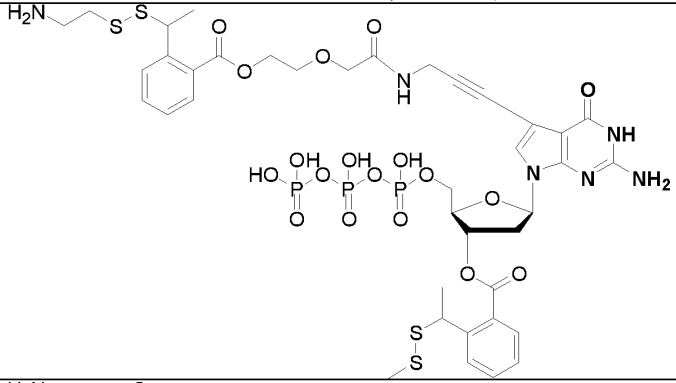
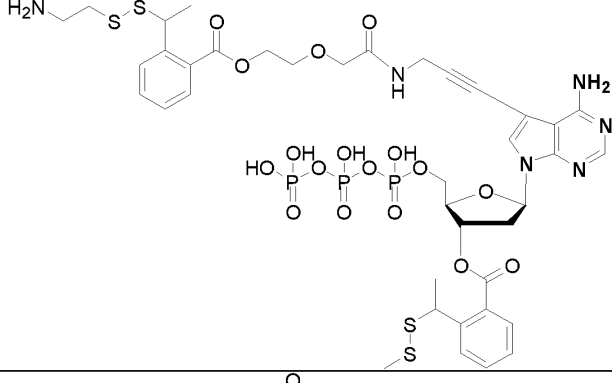
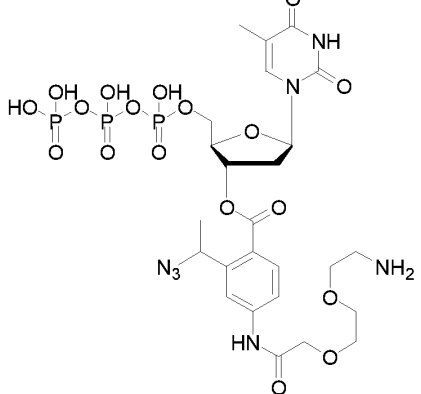
3C	
3D	
4A	
4B	

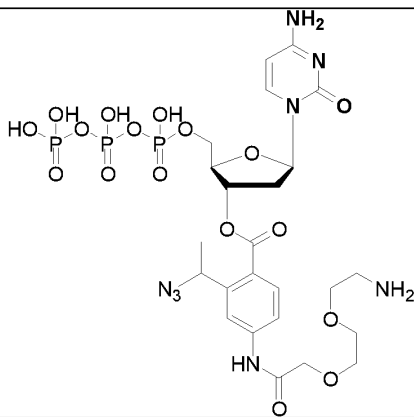
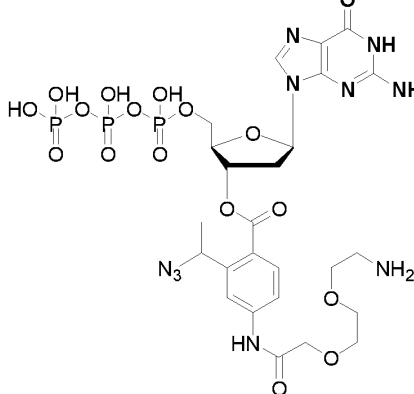
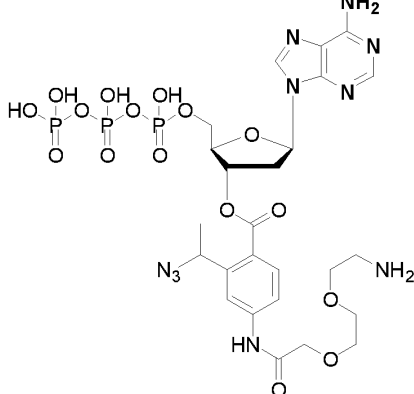
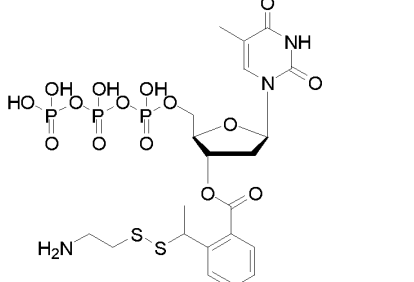
[illegible]

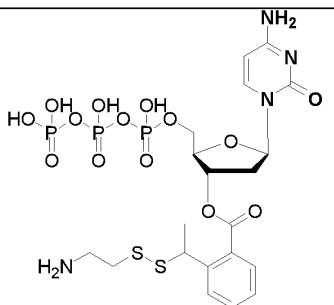
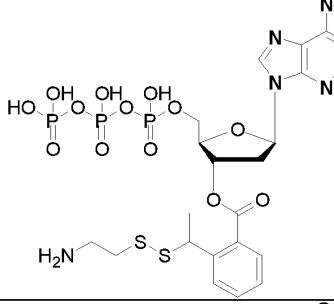
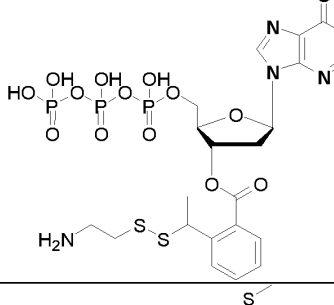
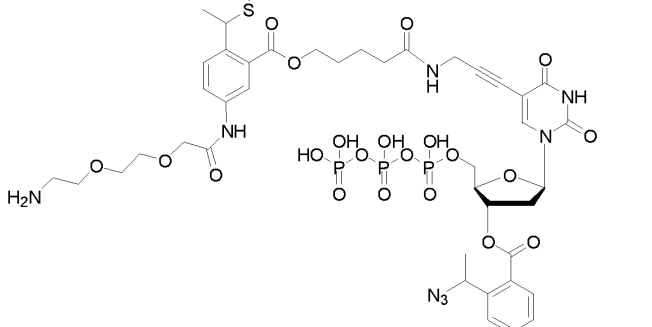
5C	
5D	
6A	
6B	

6C	
6D	
7A	
7B	

7C	
7D	
7D1	
8-1A	

8-2B	
8-2C	
8-2D	
11A	

11B	
11C	
11D	
12A	

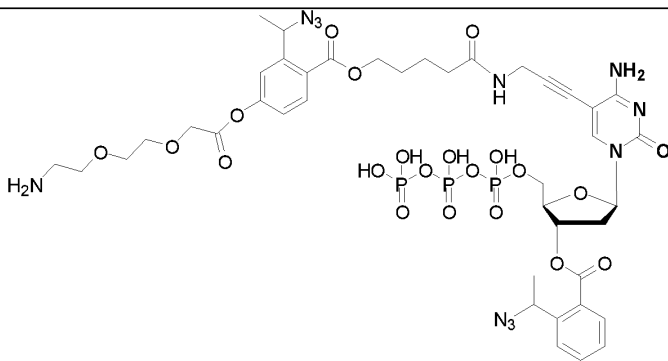
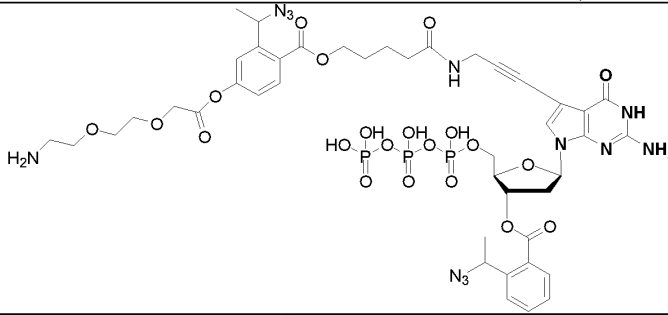
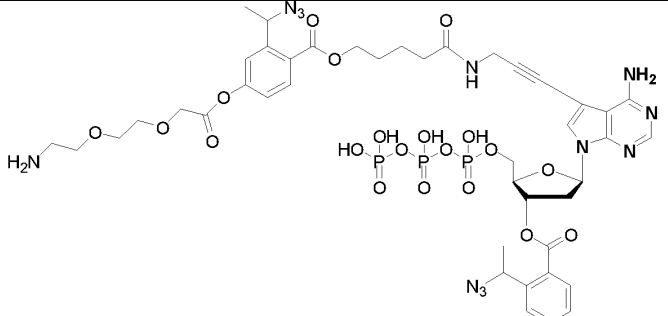
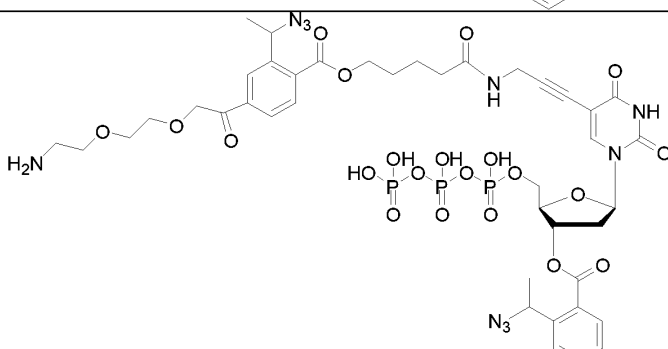
12B	
12C	
12D	
13A	

[illegible]

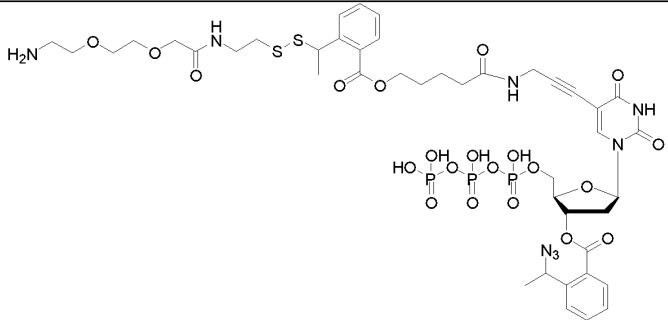
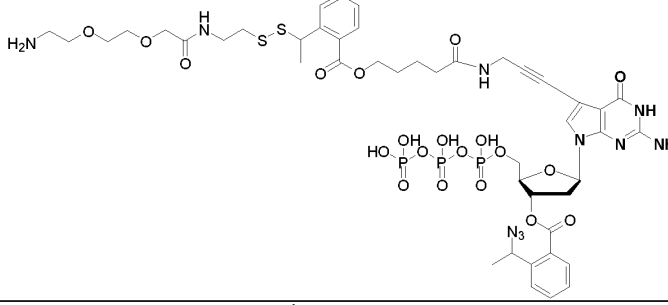
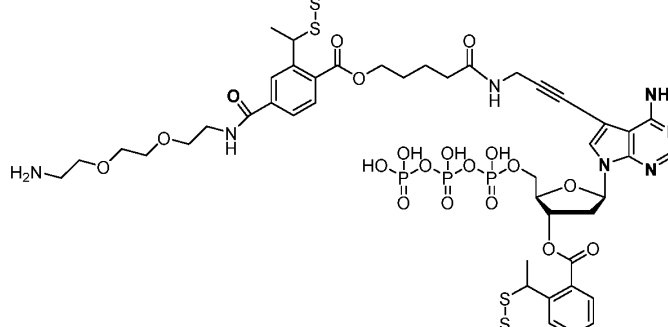
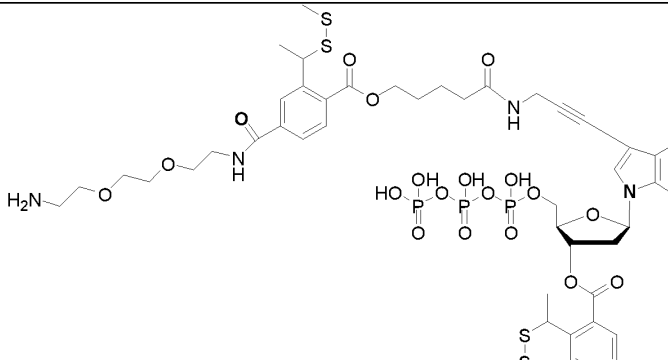
[illegible]

15B	Chemical structure 15B shows a complex molecule. It features a central sugar moiety (ribose) linked via a phosphate chain to a thiol-terminated lipid tail. The sugar is substituted with a benzoyl group and a carbamate-linked amine.
15C	Chemical structure 15C is similar to 15B, but with different substituents on the sugar moiety.
15D	Chemical structure 15D is similar to 15B, but with different substituents on the sugar moiety.
16A	Chemical structure 16A is similar to 15B, but with different substituents on the sugar moiety.

16B	The structure shows a central ribose-like sugar ring. Attached to the C2' position is a tetraphosphate chain (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-). The C3' position is linked via an ester bond to a long alkyl chain ending in a methyl disulfide group (-SS-Me). The C4' position is linked via an ester bond to a benzene ring substituted with a methyl disulfide group (-SS-Me) at the para position and a long alkyl chain ending in a methyl disulfide group (-SS-Me) at the ortho position. The C5' position is linked via an ester bond to a nucleoside derivative consisting of a pyrimidine ring with an amino group (-NH₂) at the C6 position and a carbonyl group (=O) at the C2 position.
16C	This structure is similar to 16B, but the nucleoside derivative attached to the C5' position of the sugar is a purine derivative. It consists of a fused bicyclic system (imidazole and pyrimidine rings) with an amino group (-NH₂) at the C6 position and a carbonyl group (=O) at the C2 position.
16D	This structure is similar to 16B, but the nucleoside derivative attached to the C5' position of the sugar is a pyrimidine derivative. It consists of a single six-membered ring with an amino group (-NH₂) at the C6 position and a carbonyl group (=O) at the C2 position.
17A	The structure shows a central ribose-like sugar ring. Attached to the C2' position is a tetraphosphate chain (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-). The C3' position is linked via an ester bond to a long alkyl chain ending in a methyl disulfide group (-SS-Me). The C4' position is linked via an ester bond to a benzene ring substituted with a methyl disulfide group (-SS-Me) at the para position and a long alkyl chain ending in a methyl disulfide group (-SS-Me) at the ortho position. The C5' position is linked via an ester bond to a nucleoside derivative consisting of a pyrimidine ring with a carbonyl group (=O) at the C6 position and a carbonyl group (=O) at the C2 position.

17B	
17C	
17D	
18A	

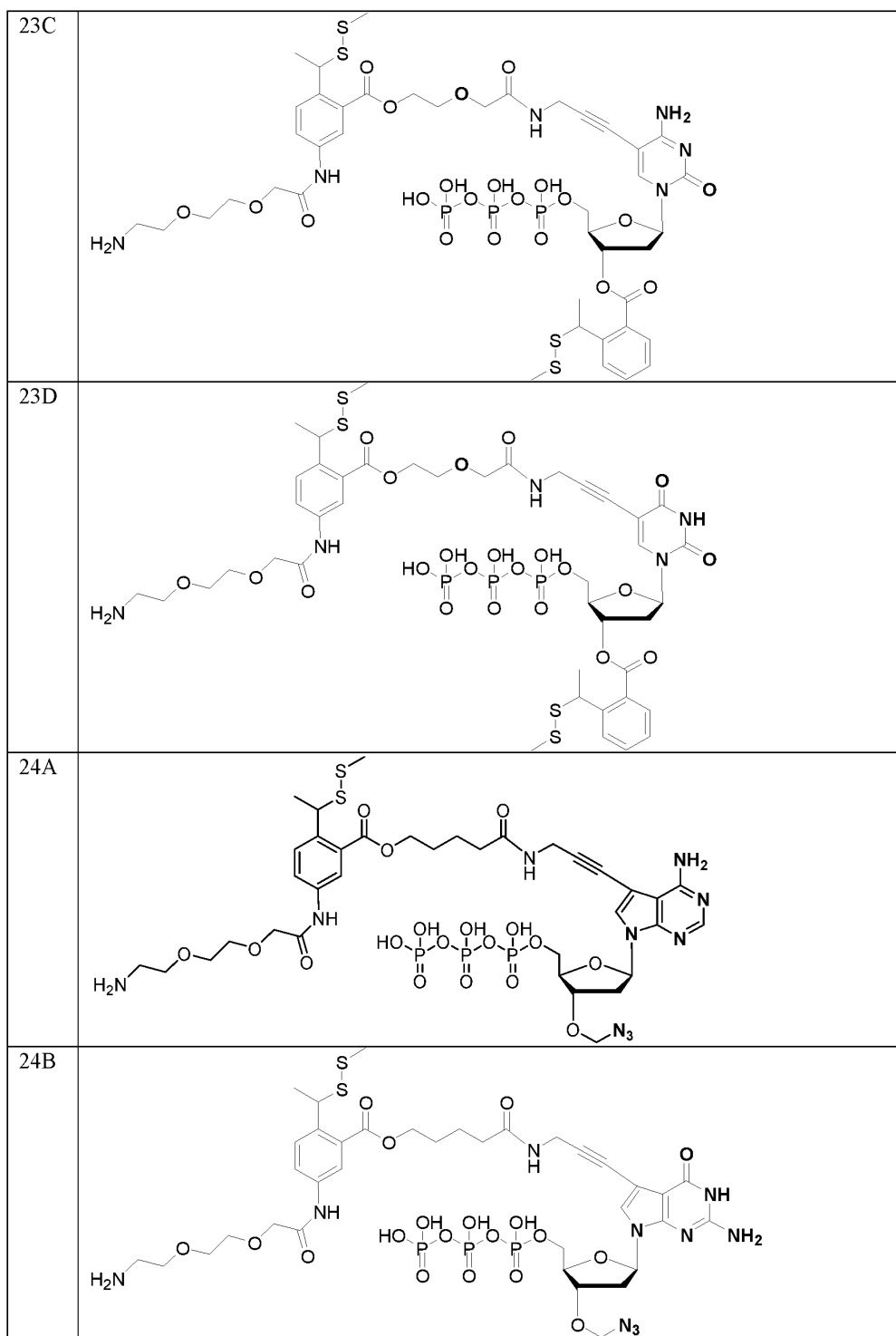
18B	
18C	
18D	
19A	
19B	

19C	
19D	
20A	
20B	

20C	
20D	
21A	
21B	

21C	
21D	
22A	
22B	

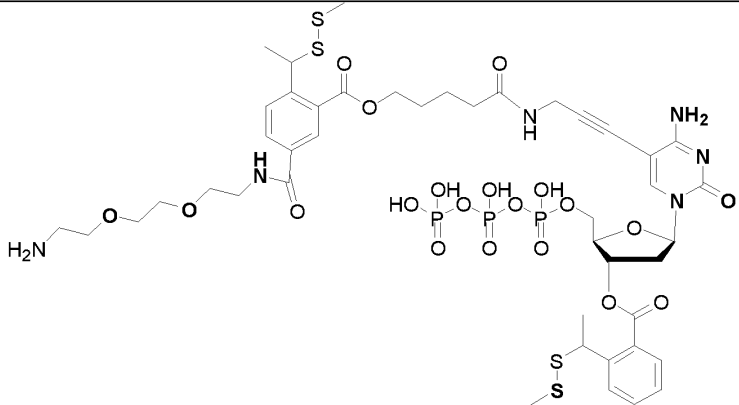
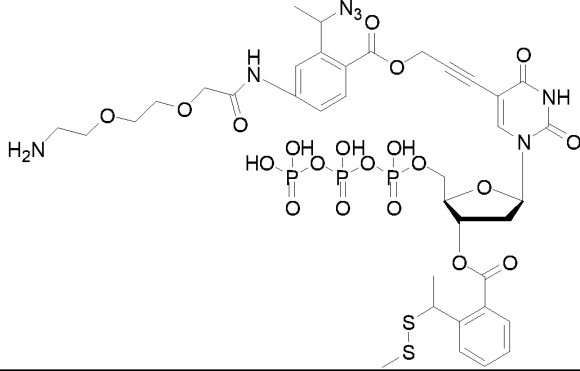
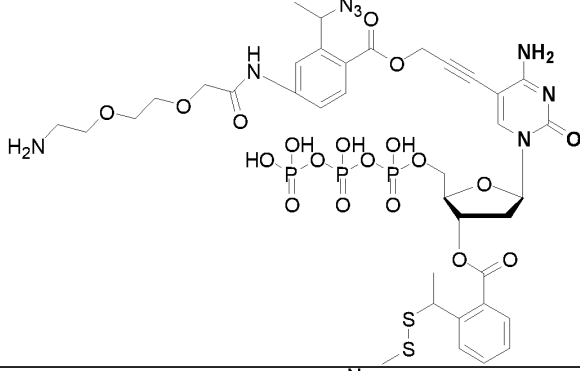
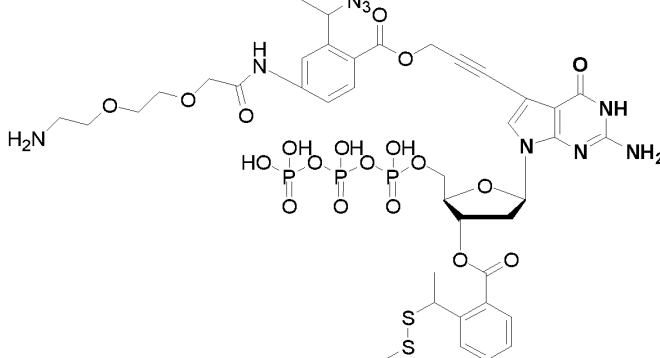
22C	
22D	
23A	
23B	



24C	
24D	
25A	
25B	
25C	

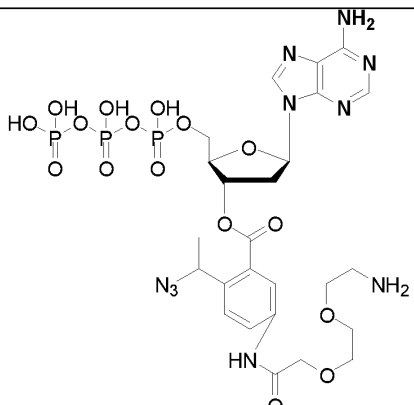
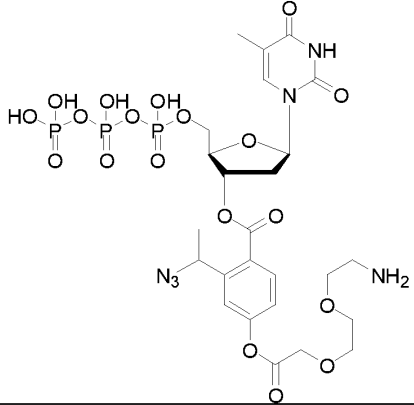
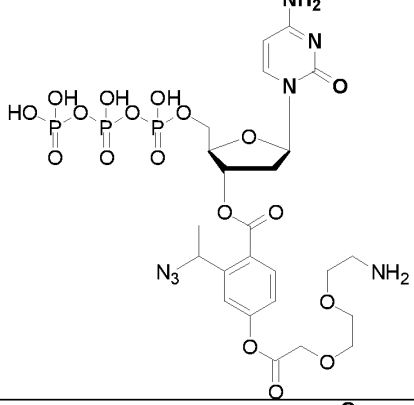
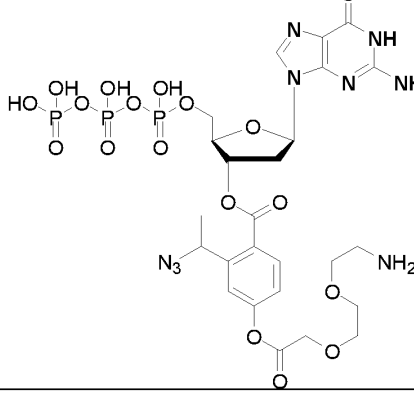
25D	
26A	
26B	
26C	

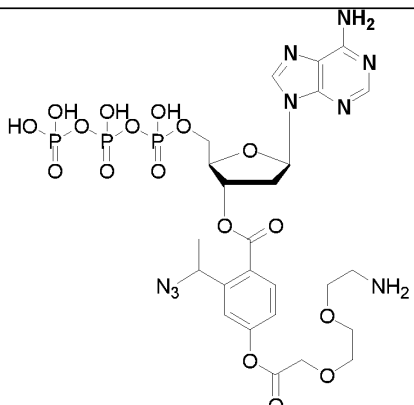
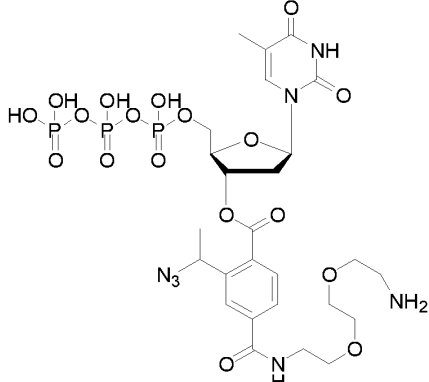
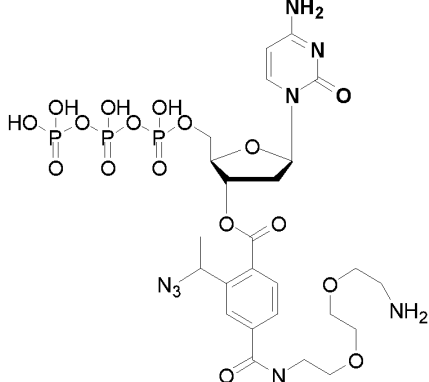
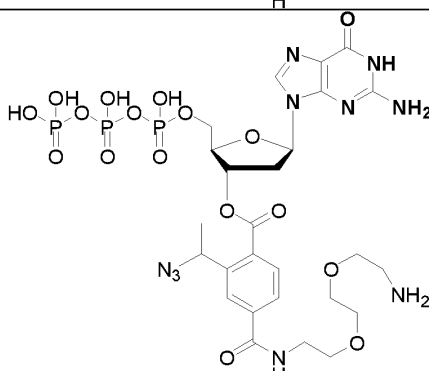
26D	
27A	
27B	
27C	

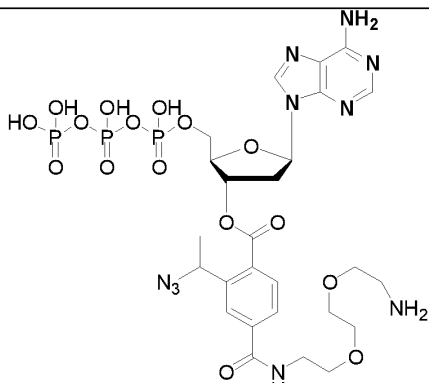
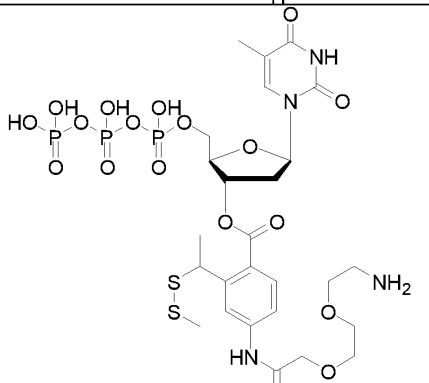
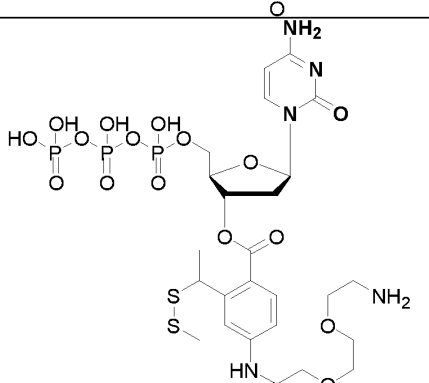
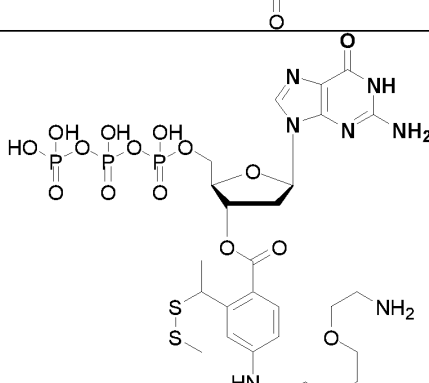
27D	
28A	
28B	
28C	

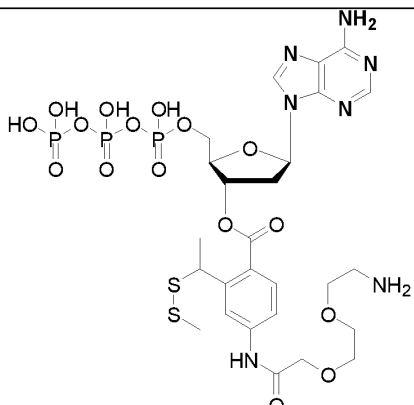
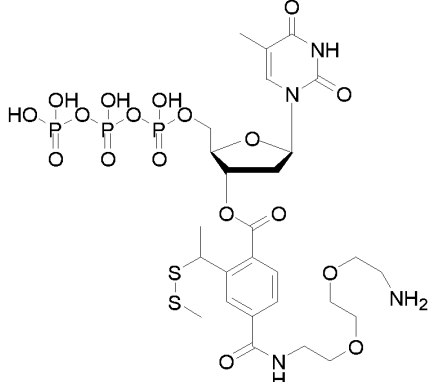
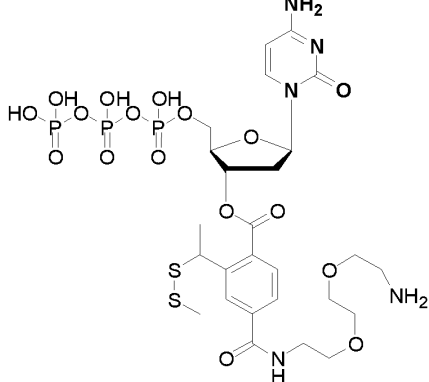
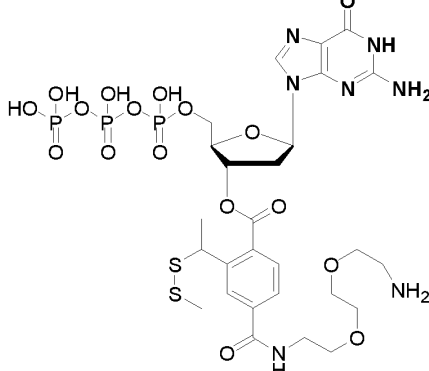
28D	
29A	
29B	
29C	

29D	
30A	
30B	
30C	

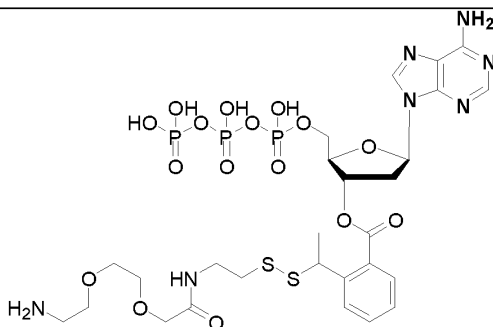
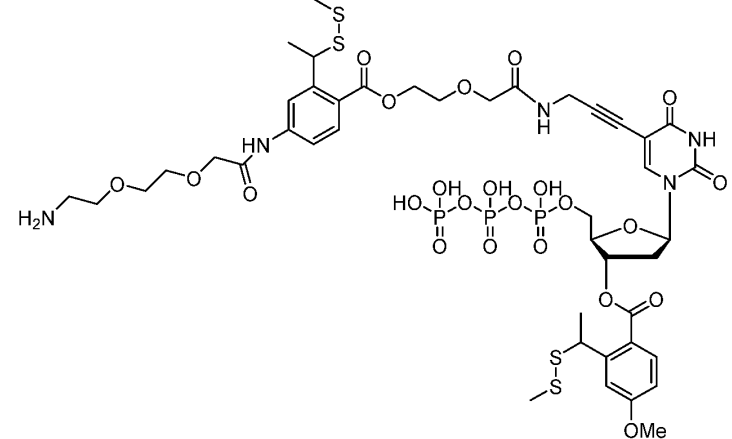
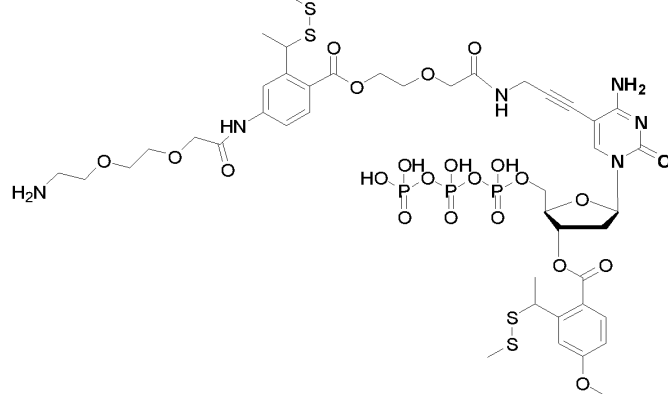
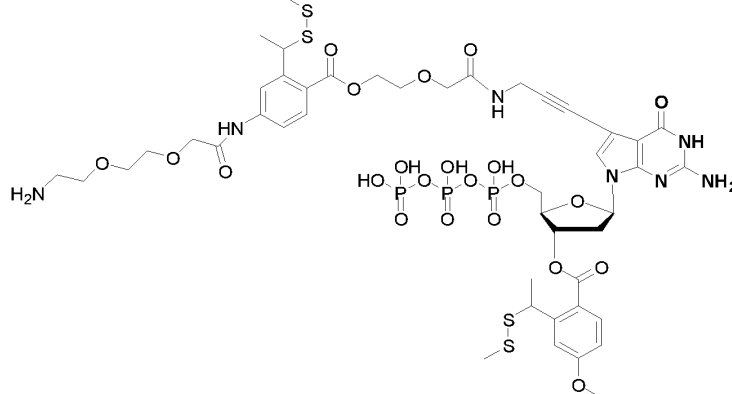
30D	
31A	
31B	
31C	

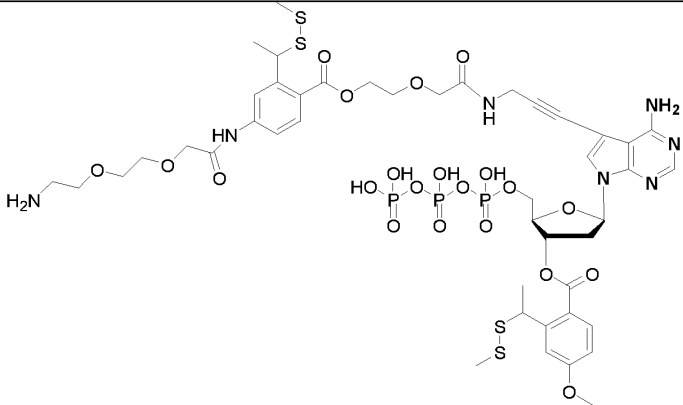
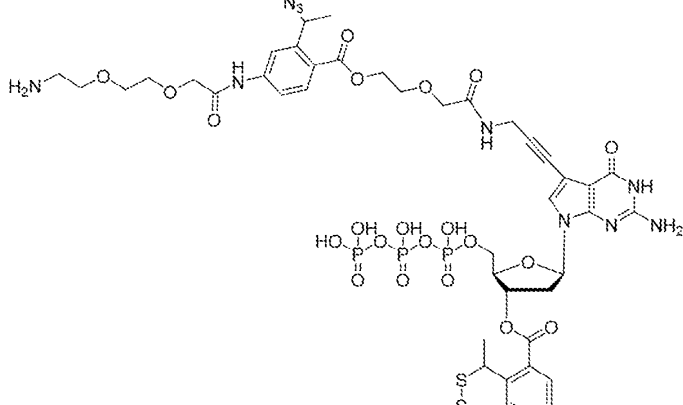
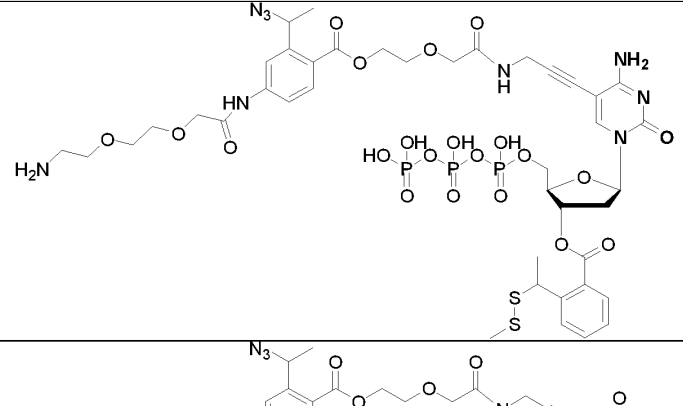
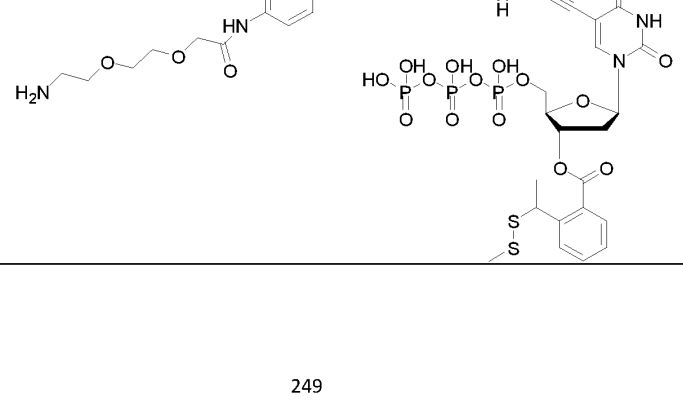
31D	
32A	
32B	
32C	

32D	
33A	
33B	
33C	

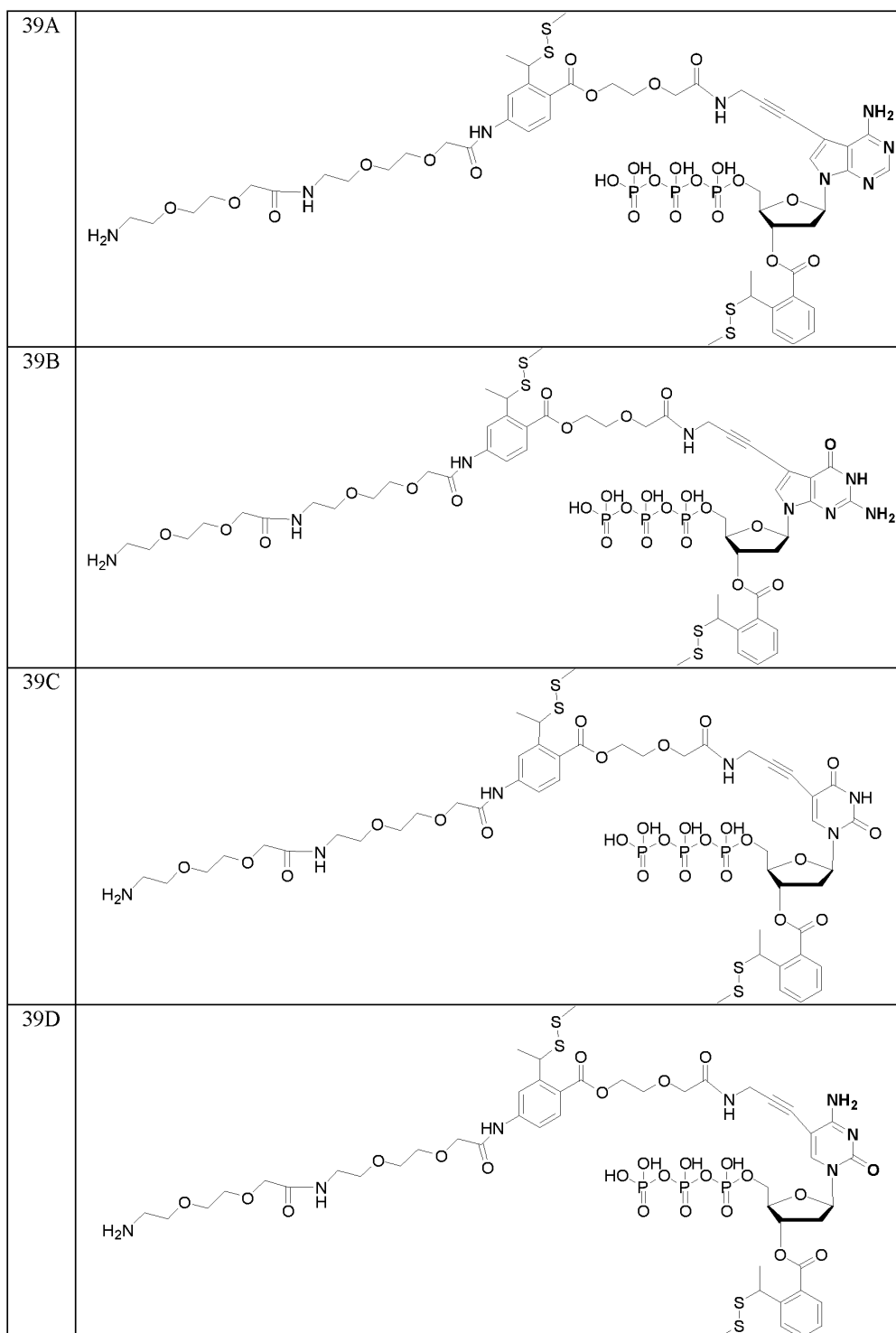
33D	
34A	
34B	
34C	

34D	
35A	
35B	
35C	

35D	
36A	
36B	
36C	

36D	
37A	
37B	
37C	

[illegible]



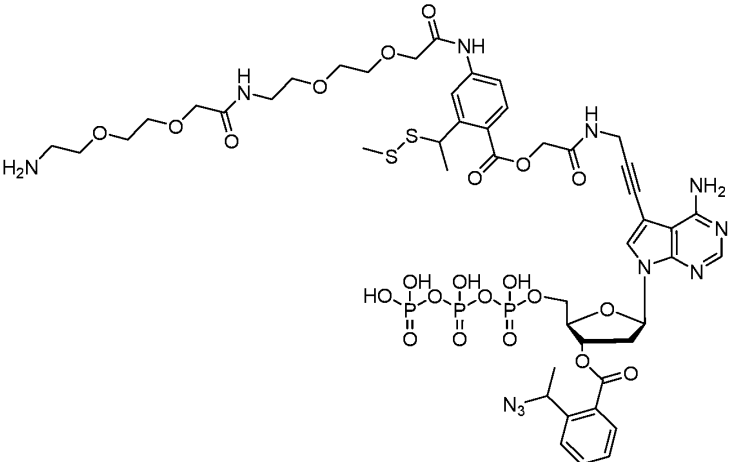
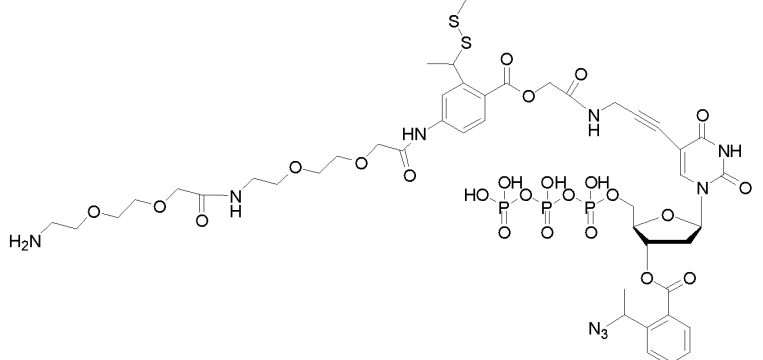
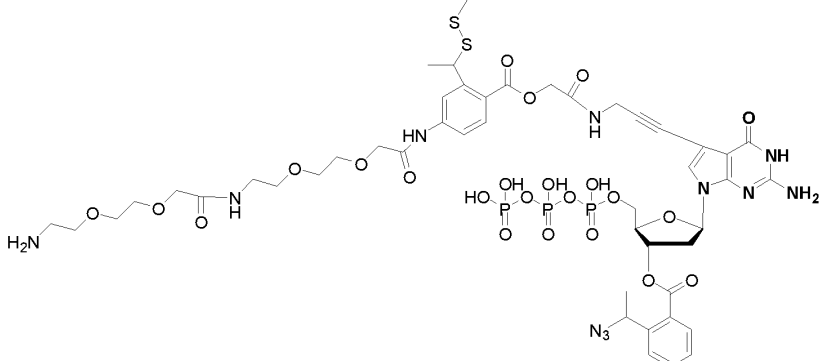
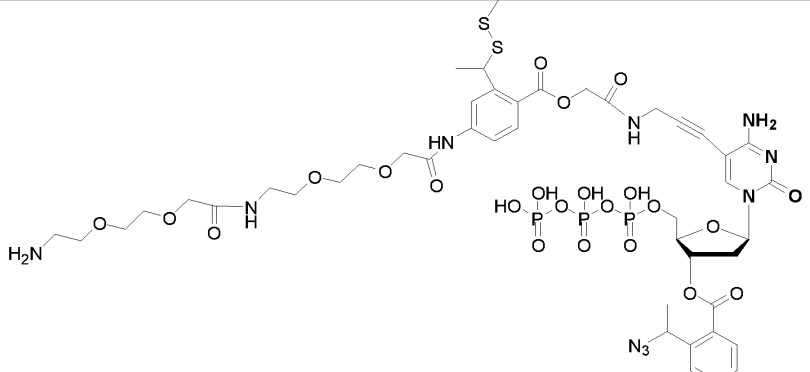
40A	
40B	
40C	
40D	
41A	

41B	
41C	
41D	
42A	
42B	
42C	
42D	

43A	
43B	
43C	
43D	

44A	
44B	
44C	
44D	

45A	
45B	
45C	
45D	

46A	
46B	
46C	
46D	

47A	
47B	
47C	
47D	

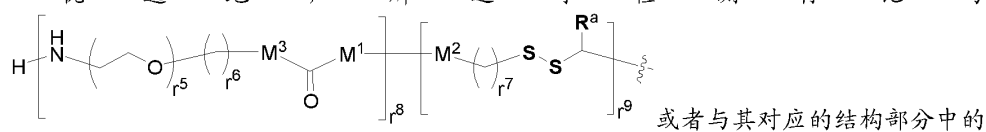
48A	
48B	
48C	
48D	

10、权利要求 1-9 任一项所述的化合物或其盐，其中，所述化合物或其盐携带额外的可检测标记；

优选地，所述化合物或其盐携带的额外的可检测标记是通过亲和试剂（如抗体、适体、亲和物、结蛋白）引入的，所述亲和试剂携带所述可检测标记，且所述亲和试剂可以特异性识别并结合所述化合物或其盐的表位；

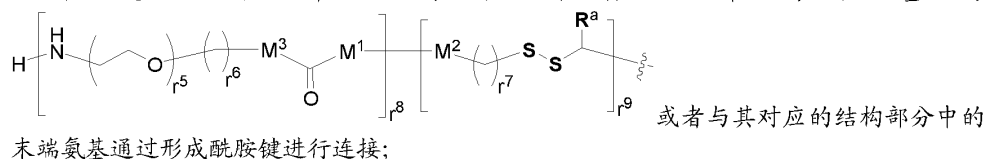
或者优选地，所述可检测标记与式 I、式 II、式 I-1、式 I-2、式 II-1 或式 II-2 所示化合物中的 R 或权利要求 9 的化合物中对应于 R 的结构部分连接；

优选地，所述可检测标记与

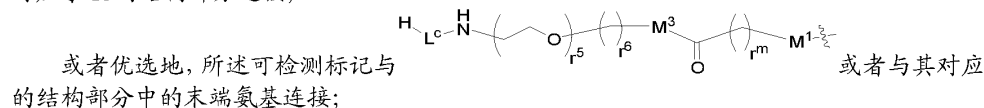


末端氨基连接;

优选地, 所述可检测标记中的羧基与



或者优选地, 所述可检测标记与式 III 所示化合物中的 R 连接或权利要求 9 的化合物中对应于 R 的结构部分连接;



或者优选地, 所述可检测标记与

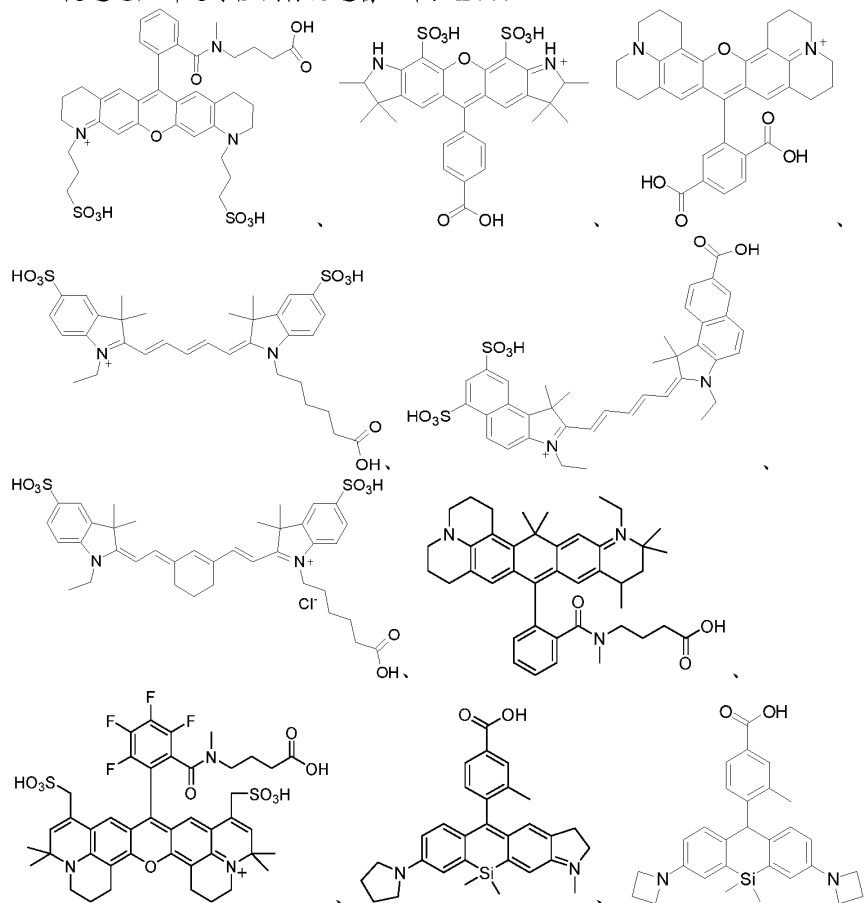
$$\text{H}_2\text{N}-\left(\text{CH}_2\text{CH}_2\text{O}\right)_{r^5}\left(\text{CH}_2\right)_{r^6}\text{M}^3\text{C}(=\text{O})\left(\text{CH}_2\right)_{r^m}\text{M}^1-$$
 或者与其对应的结构部分中的末端氨基通过形成酰胺键进行连接;

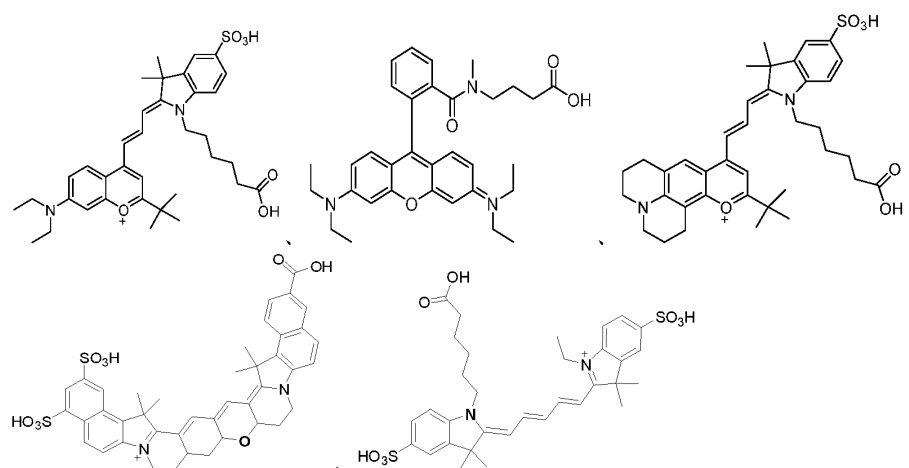
优选地, base¹ 不同, 式 I 所示化合物携带的额外的可检测标记不同;

优选地, base² 不同, 式 II 所示化合物携带的额外的可检测标记不同;

优选地, 所述可检测标记为荧光标记;

优选地, 所述可检测标记选自以下: iF700、

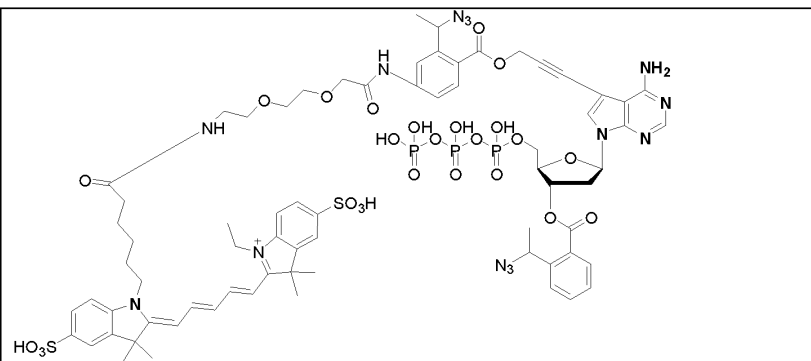
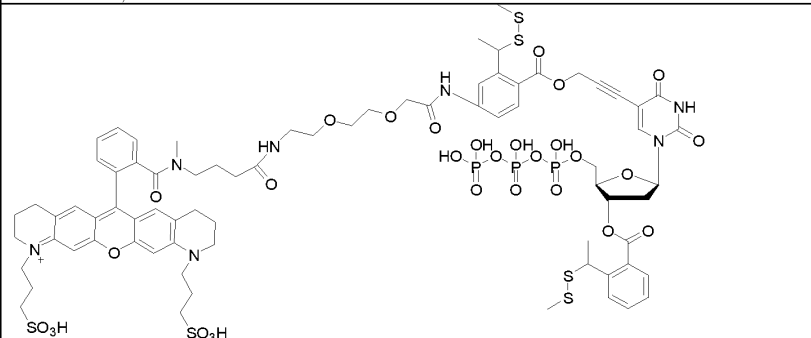
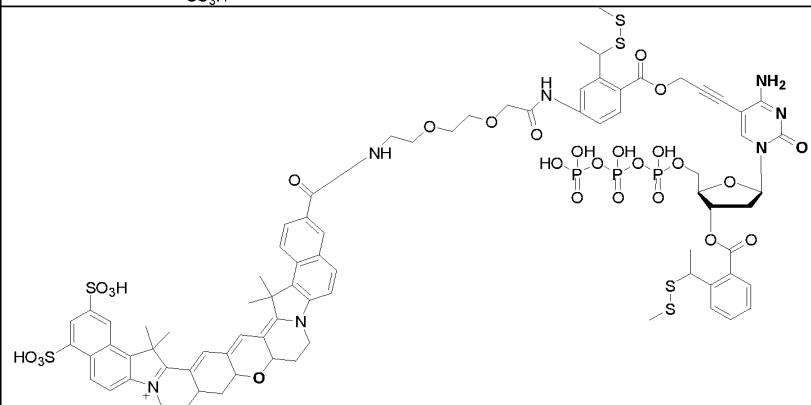
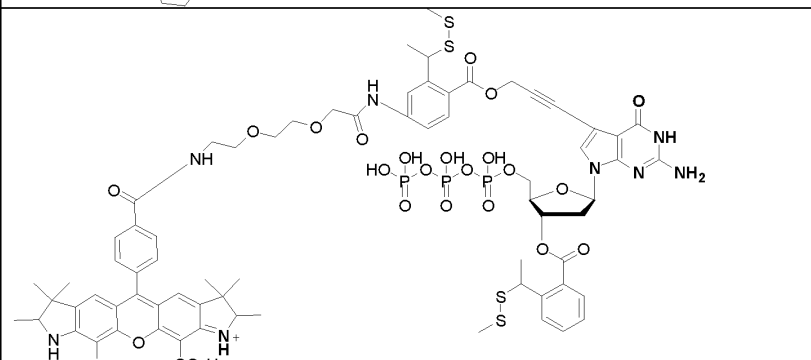




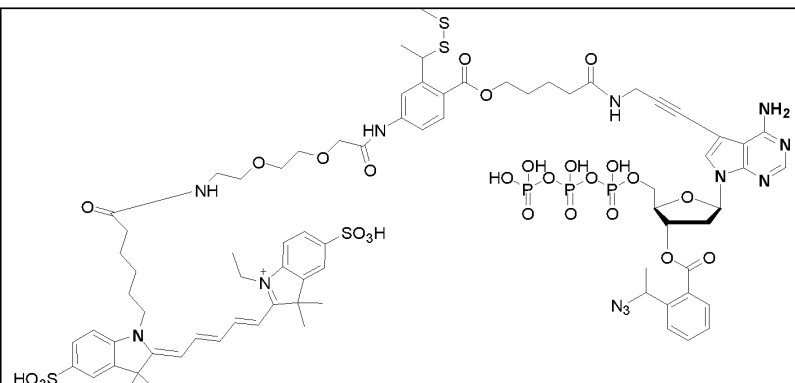
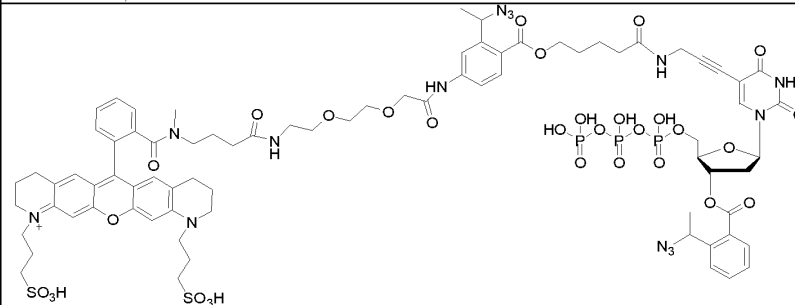
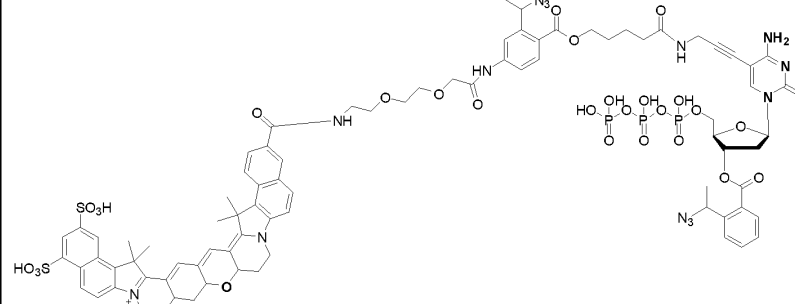
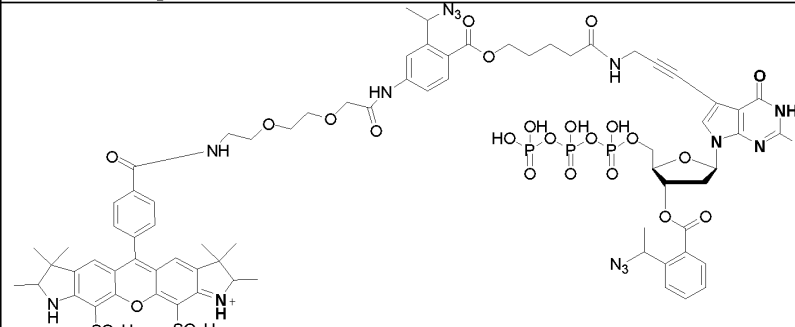
11、权利要求 10 所述的化合物或其盐，其中，所述化合物选自以下：

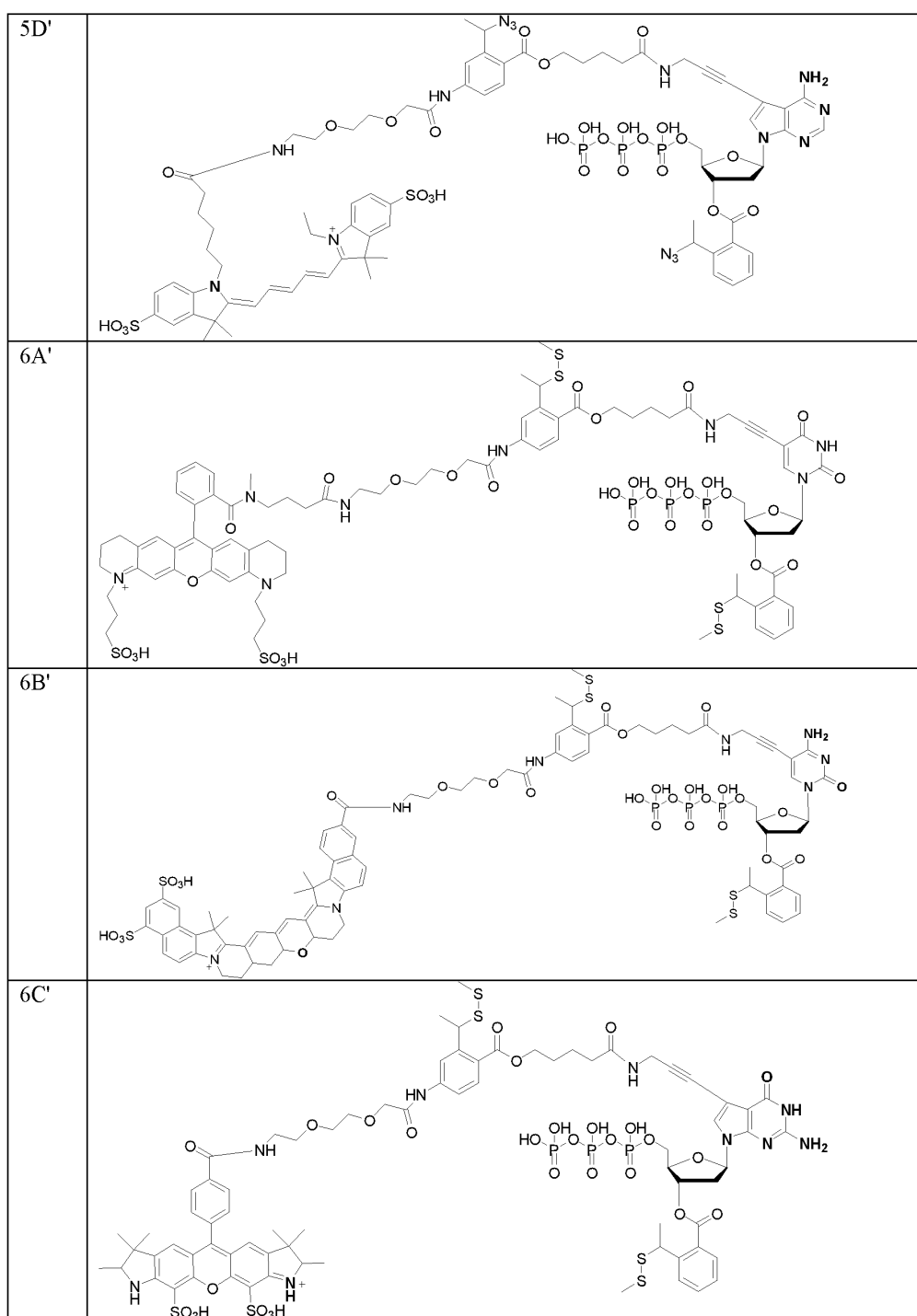
编号	结构式
1A'	
1B'	
1C'	

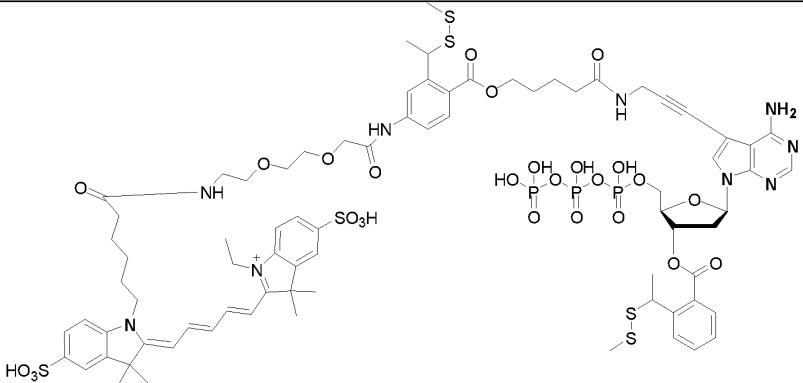
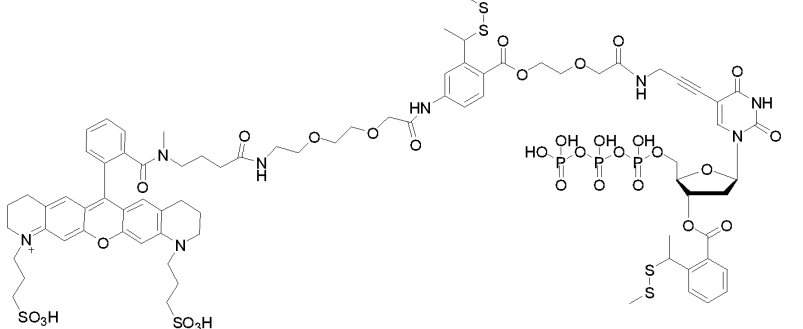
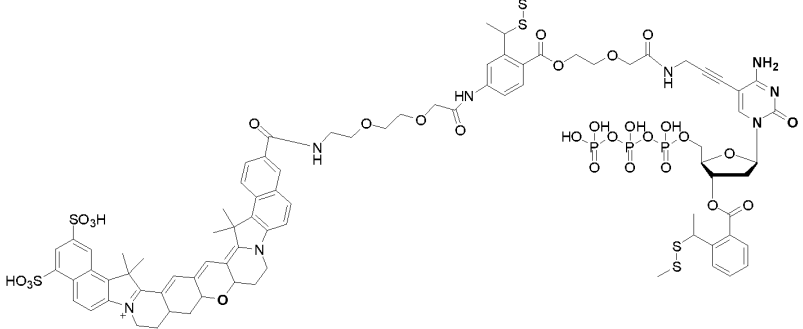
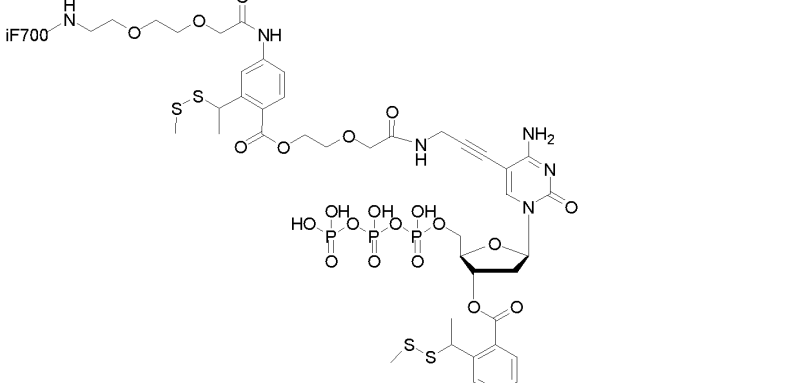
1D'	Chemical structure 1D' is a complex molecule featuring a central sugar moiety (ribose) linked via a phosphate group to a pyrimidine ring (cytosine). The cytosine is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H). The sugar moiety is also linked via a phosphate group to a pyrimidine ring (uracil). The uracil is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H).
2A'	Chemical structure 2A' is a complex molecule featuring a central sugar moiety (ribose) linked via a phosphate group to a pyrimidine ring (cytosine). The cytosine is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H). The sugar moiety is also linked via a phosphate group to a pyrimidine ring (uracil). The uracil is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H).
2B'	Chemical structure 2B' is a complex molecule featuring a central sugar moiety (ribose) linked via a phosphate group to a pyrimidine ring (cytosine). The cytosine is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H). The sugar moiety is also linked via a phosphate group to a pyrimidine ring (uracil). The uracil is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H).
2C'	Chemical structure 2C' is a complex molecule featuring a central sugar moiety (ribose) linked via a phosphate group to a pyrimidine ring (cytosine). The cytosine is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H). The sugar moiety is also linked via a phosphate group to a pyrimidine ring (uracil). The uracil is further linked via a triazole ring to a phenyl group. The phenyl group is substituted with a sulfonamide group (SO₂NH₂) and a sulfonate group (SO₃H).

2D'	
3A'	
3B'	
3C'	

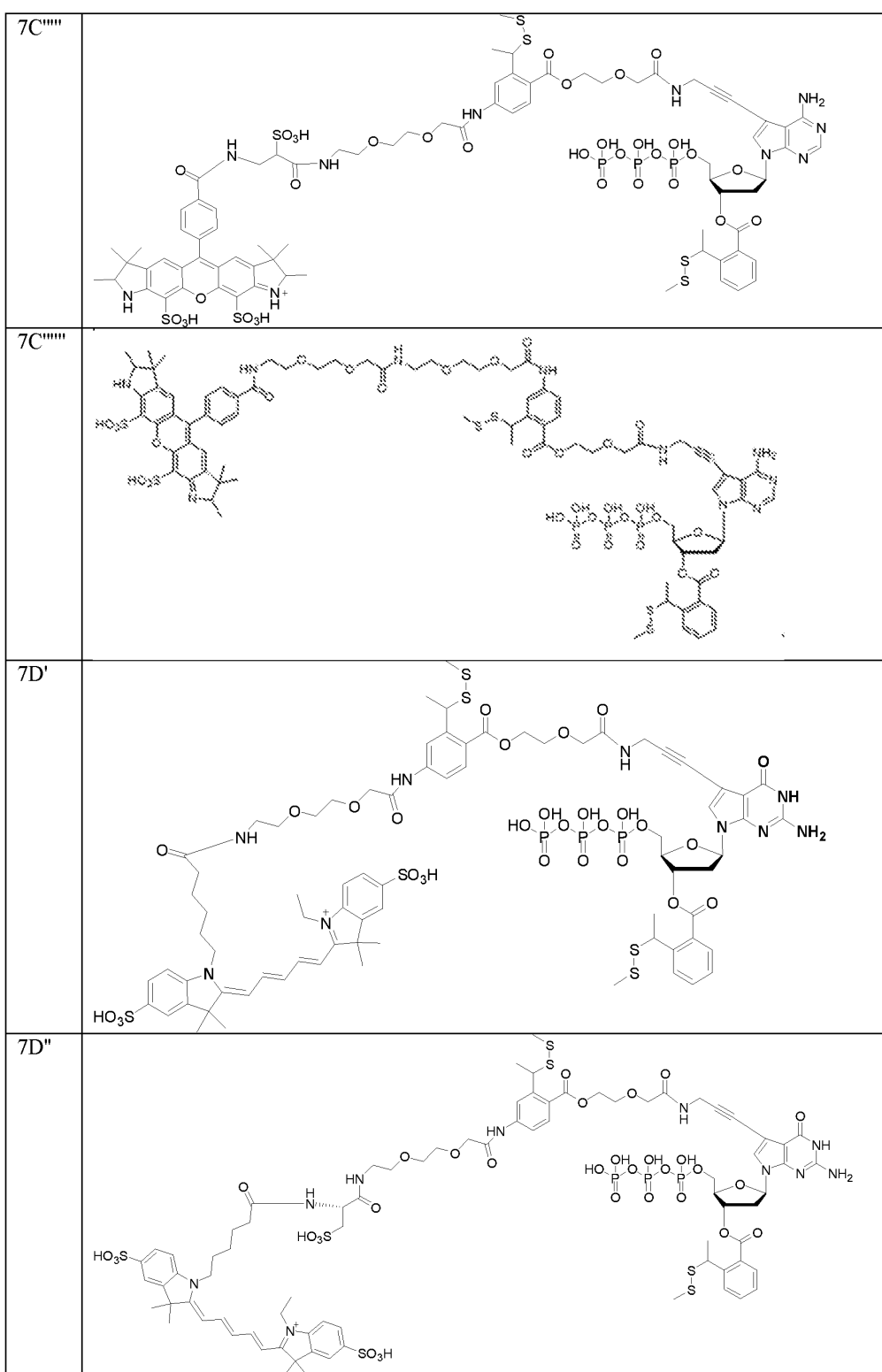
3D'	
4A'	
4B'	
4C'	

4D'	
5A'	
5B'	
5C'	

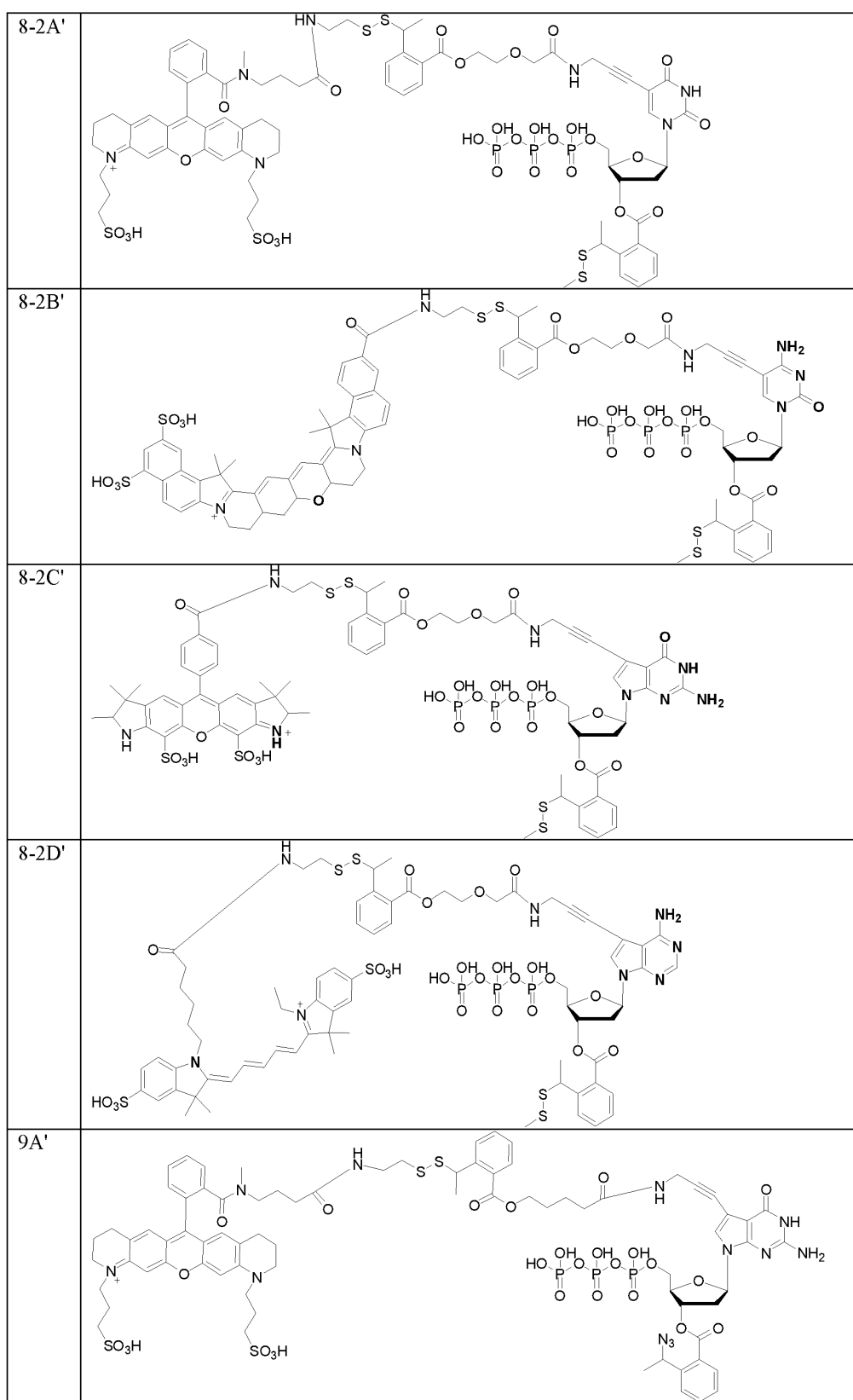


6D'	
7A'	
7B'	
7B''	

7C'	Chemical structure of compound 7C'. It features a central porphyrin core with two sulfonate groups (SO₃H) and two methyl groups. The core is substituted with a long, flexible linker containing multiple ether and amide bonds. The linker terminates in a thiol group (SH) and a complex side chain. The side chain includes a benzene ring with a thiol group (SH), a carbonyl group, and a triphosphate group (three phosphate units). The triphosphate group is linked to a ribose sugar, which is further linked to a pyrimidine base (cytosine) via a glycosidic bond.
7C''	Chemical structure of compound 7C''. It is similar to 7C' but features a different linker and side chain. The linker is shorter and contains fewer ether and amide bonds. The side chain includes a benzene ring with a thiol group (SH), a carbonyl group, and a triphosphate group (three phosphate units). The triphosphate group is linked to a ribose sugar, which is further linked to a pyrimidine base (cytosine) via a glycosidic bond.
7C'''	Chemical structure of compound 7C'''. It features a central porphyrin core with two sulfonate groups (SO₃H) and two methyl groups. The core is substituted with a long, flexible linker containing multiple ether and amide bonds. The linker terminates in a thiol group (SH) and a complex side chain. The side chain includes a benzene ring with a thiol group (SH), a carbonyl group, and a triphosphate group (three phosphate units). The triphosphate group is linked to a ribose sugar, which is further linked to a pyrimidine base (cytosine) via a glycosidic bond.
7C''''	Chemical structure of compound 7C''''. It features a central porphyrin core with two sulfonate groups (SO₃H) and two methyl groups. The core is substituted with a long, flexible linker containing multiple ether and amide bonds. The linker terminates in a thiol group (SH) and a complex side chain. The side chain includes a benzene ring with a thiol group (SH), a carbonyl group, and a triphosphate group (three phosphate units). The triphosphate group is linked to a ribose sugar, which is further linked to a pyrimidine base (cytosine) via a glycosidic bond.



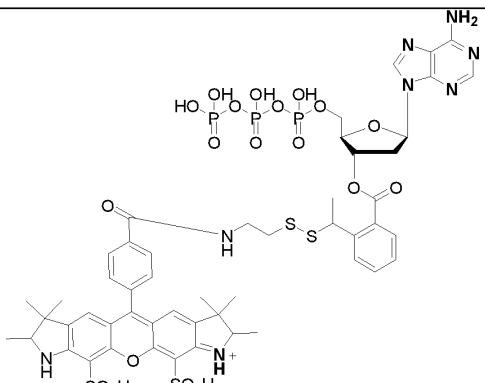
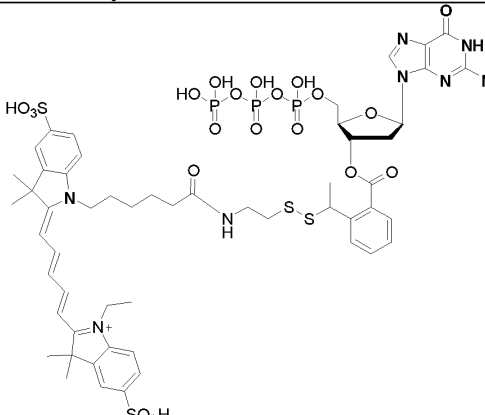
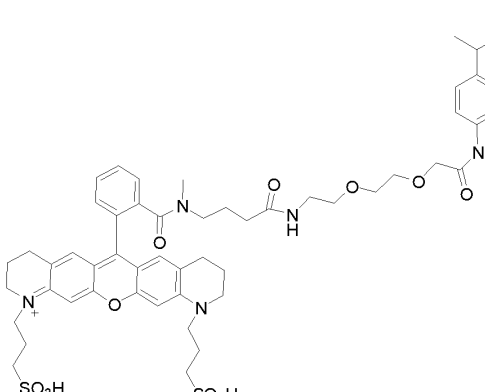
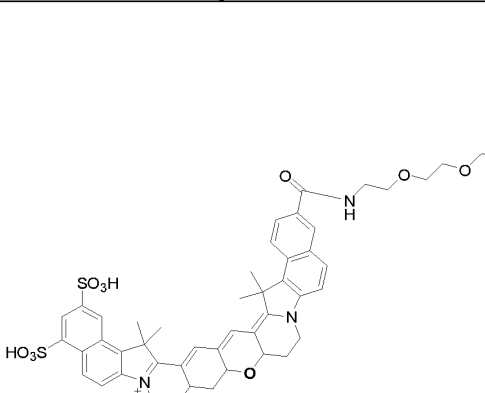
8-1A'	Chemical structure 8-1A' features a sulfonate group (SO₃H) attached to a complex polycyclic aromatic system. This system is linked via a disulfide bridge (S-S) to a chain containing a carboxylic acid group (COOH) and a nucleoside derivative. The nucleoside derivative includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O), and a pyrimidine base with an amino group (NH₂).
8-1B'	Chemical structure 8-1B' features a sulfonate group (SO₃H) attached to a complex polycyclic aromatic system. This system is linked via a disulfide bridge (S-S) to a chain containing a carboxylic acid group (COOH) and a nucleoside derivative. The nucleoside derivative includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O), and a pyrimidine base with a carbonyl group (C=O).
8-1C'	Chemical structure 8-1C' features a sulfonate group (SO₃H) attached to a complex polycyclic aromatic system. This system is linked via a disulfide bridge (S-S) to a chain containing a carboxylic acid group (COOH) and a nucleoside derivative. The nucleoside derivative includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O), and a pyrimidine base with an amino group (NH₂).
8-1D'	Chemical structure 8-1D' features a sulfonate group (SO₃H) attached to a complex polycyclic aromatic system. This system is linked via a disulfide bridge (S-S) to a chain containing a carboxylic acid group (COOH) and a nucleoside derivative. The nucleoside derivative includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O-P(=O)(OH)-O), and a pyrimidine base with an amino group (NH₂).



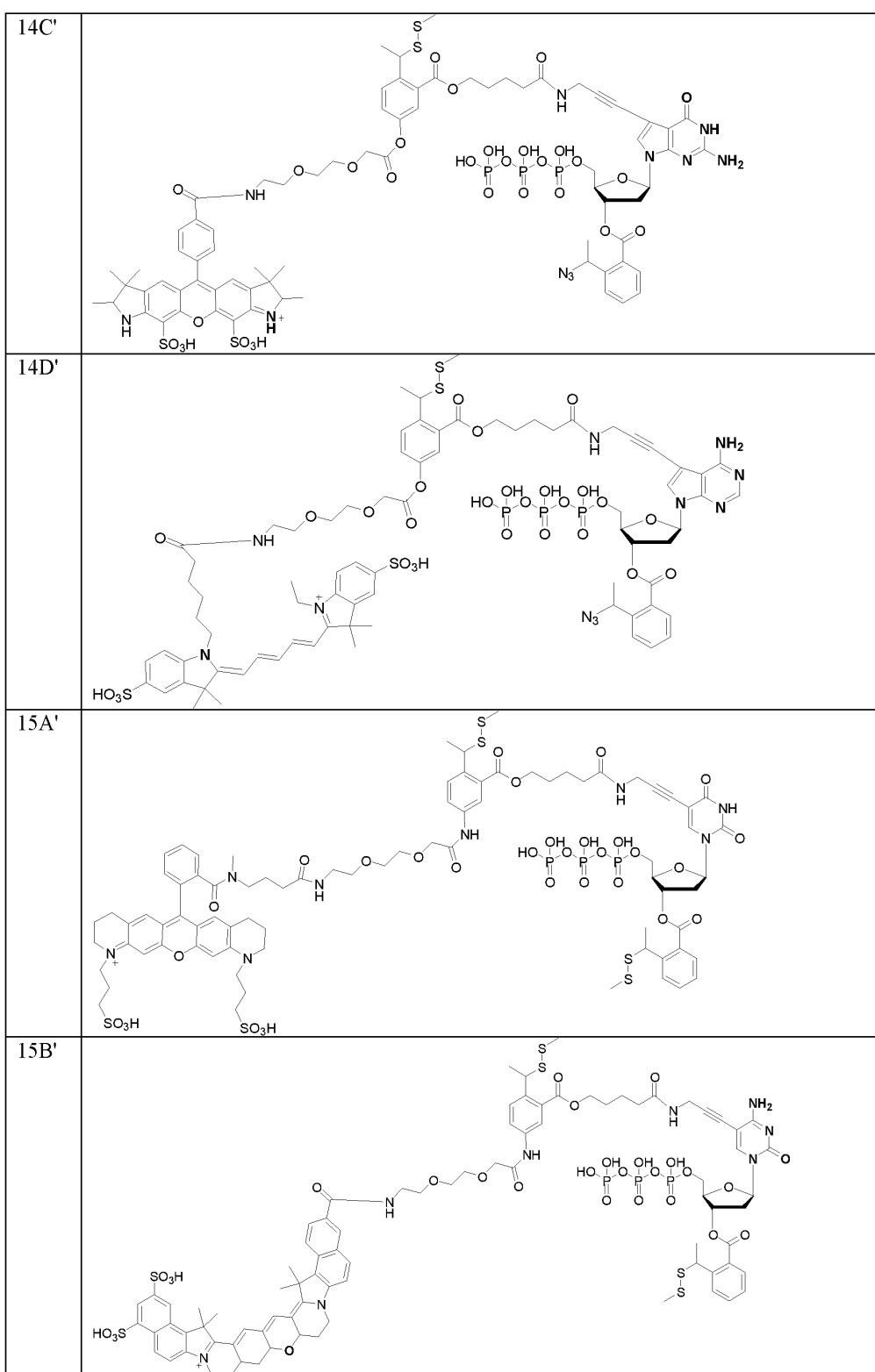
9B'	Chemical structure 9B' features a complex polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H). This system is linked via a long chain containing a disulfide bridge (S-S) and a carbonyl group to a central linker. The linker is further connected to a sugar moiety (ribose) which is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a nucleoside derivative (2'-azido-2'-deoxyribose-5'-phosphate) at the 3' position.
9C'	Chemical structure 9C' features a complex polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H). This system is linked via a long chain containing a disulfide bridge (S-S) and a carbonyl group to a central linker. The linker is further connected to a sugar moiety (ribose) which is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a nucleoside derivative (2'-azido-2'-deoxyribose-5'-phosphate) at the 3' position.
9D'	Chemical structure 9D' features a complex polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H). This system is linked via a long chain containing a disulfide bridge (S-S) and a carbonyl group to a central linker. The linker is further connected to a sugar moiety (ribose) which is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a nucleoside derivative (2'-azido-2'-deoxyribose-5'-phosphate) at the 3' position.
10A'	Chemical structure 10A' features a complex polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H). This system is linked via a long chain containing a disulfide bridge (S-S) and a carbonyl group to a central linker. The linker is further connected to a sugar moiety (ribose) which is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a nucleoside derivative (2'-azido-2'-deoxyribose-5'-phosphate) at the 3' position.
10B'	Chemical structure 10B' features a complex polycyclic aromatic system on the left, including a naphthalene ring with two sulfonic acid groups (SO₃H). This system is linked via a long chain containing a disulfide bridge (S-S) and a carbonyl group to a central linker. The linker is further connected to a sugar moiety (ribose) which is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a nucleoside derivative (2'-azido-2'-deoxyribose-5'-phosphate) at the 3' position.

10C'	
10D'	
11A'	
11B'	

[illegible]

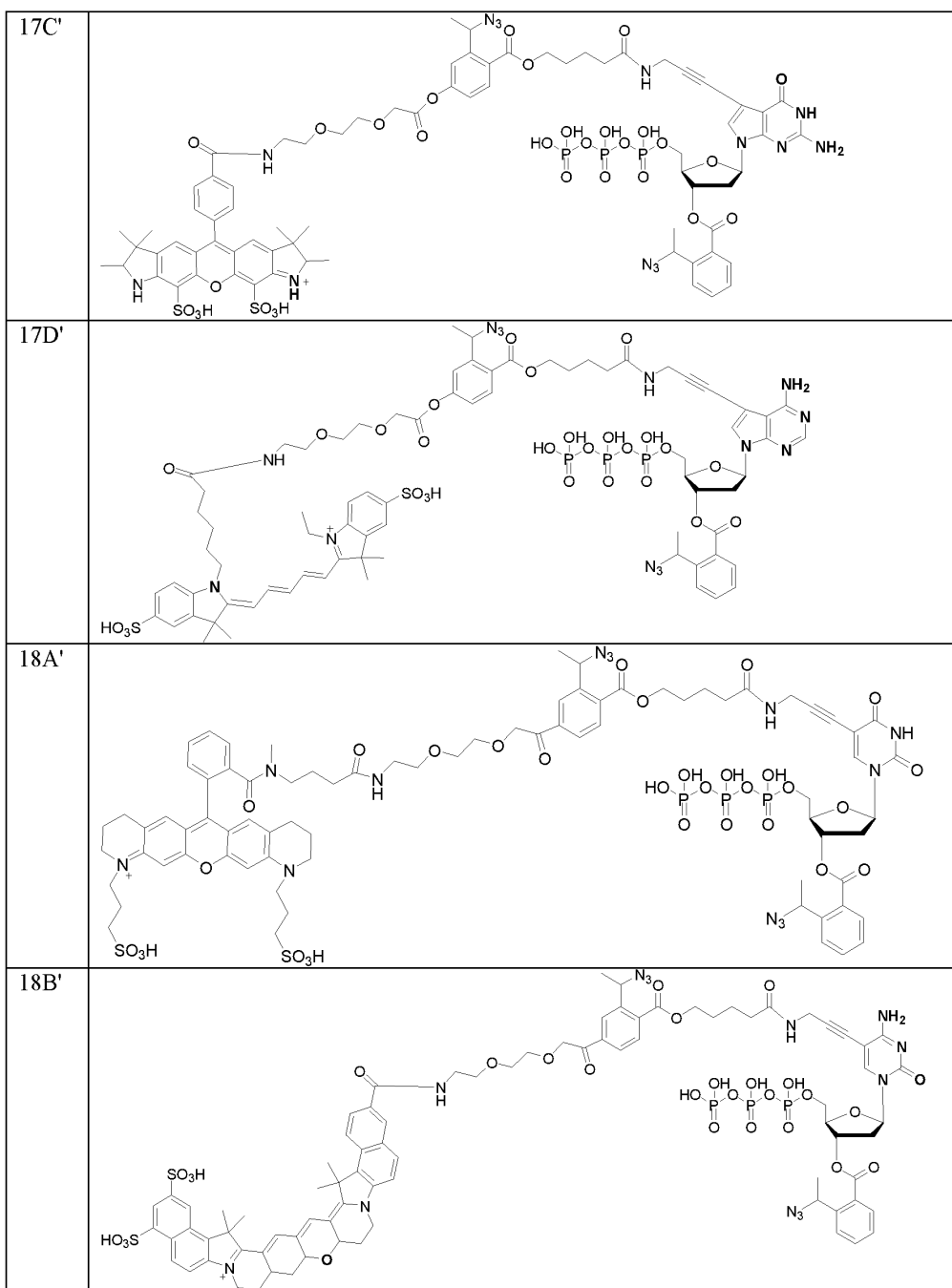
12C'	
12D'	
13A'	
13B'	

13C'	Chemical structure 13C' is a complex molecule. It features a central core consisting of a benzene ring substituted with a dimethylsulfoniomethyl group (-CH₂SO₂CH₃) and a dimethylaminomethyl group (-CH₂NHCH₃). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a nucleoside moiety. The nucleoside part includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-), and a pyrimidine base with an amino group (-NH₂) and a carbonyl group (=O). A 3-azido-4-methylphenyl group is also attached to the nucleoside moiety.
13D'	Chemical structure 13D' is similar to 13C' but with a different central core. The core is a benzene ring substituted with a dimethylsulfoniomethyl group (-CH₂SO₂CH₃) and a dimethylaminomethyl group (-CH₂NHCH₃). It is linked via a long, flexible chain containing multiple ether and amide linkages to a nucleoside moiety. The nucleoside part includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-), and a pyrimidine base with an amino group (-NH₂) and a carbonyl group (=O). A 3-azido-4-methylphenyl group is also attached to the nucleoside moiety.
14A'	Chemical structure 14A' is a complex molecule. It features a central core consisting of a benzene ring substituted with a dimethylsulfoniomethyl group (-CH₂SO₂CH₃) and a dimethylaminomethyl group (-CH₂NHCH₃). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a nucleoside moiety. The nucleoside part includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-), and a pyrimidine base with an amino group (-NH₂) and a carbonyl group (=O). A 3-azido-4-methylphenyl group is also attached to the nucleoside moiety.
14B'	Chemical structure 14B' is a complex molecule. It features a central core consisting of a benzene ring substituted with a dimethylsulfoniomethyl group (-CH₂SO₂CH₃) and a dimethylaminomethyl group (-CH₂NHCH₃). This core is linked via a long, flexible chain containing multiple ether and amide linkages to a nucleoside moiety. The nucleoside part includes a ribose sugar, a phosphate group (HO-P(=O)(OH)-O-), and a pyrimidine base with an amino group (-NH₂) and a carbonyl group (=O). A 3-azido-4-methylphenyl group is also attached to the nucleoside moiety.



15C'	Chemical structure 15C' is a complex molecule. It features a central core with two sulfonate groups (SO₃H) and a sulfonamide group (NH-SO₂-). This core is linked via a long chain containing a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) to a pyrimidine ring. The pyrimidine ring is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H). The molecule also includes a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) linked to a pyrimidine ring, which is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H).
15D'	Chemical structure 15D' is a complex molecule. It features a central core with two sulfonate groups (SO₃H) and a sulfonamide group (NH-SO₂-). This core is linked via a long chain containing a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) to a pyrimidine ring. The pyrimidine ring is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H). The molecule also includes a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) linked to a pyrimidine ring, which is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H).
16A'	Chemical structure 16A' is a complex molecule. It features a central core with two sulfonate groups (SO₃H) and a sulfonamide group (NH-SO₂-). This core is linked via a long chain containing a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) to a pyrimidine ring. The pyrimidine ring is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H). The molecule also includes a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) linked to a pyrimidine ring, which is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H).
16B'	Chemical structure 16B' is a complex molecule. It features a central core with two sulfonate groups (SO₃H) and a sulfonamide group (NH-SO₂-). This core is linked via a long chain containing a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) to a pyrimidine ring. The pyrimidine ring is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H). The molecule also includes a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H) linked to a pyrimidine ring, which is further substituted with a sulfonamide group (NH-SO₂-) and a sulfonate group (SO₃H).

16C'	Chemical structure 16C' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃), a methyl ester group (-COOCH₃), and a long chain containing a sulfonamide group (-SO₂NH-), a polyether chain (-O-CH₂-CH₂-O-CH₂-CH₂-O-CH₂-CH₂-O-), and a carboxylic acid group (-COOH). This chain is further substituted with a 2-amino-2H-imidazo[5,1-b]pyridine moiety and a 4-azido-2-methylphenyl group. The molecule also contains a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety and a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety.
16D'	Chemical structure 16D' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃), a methyl ester group (-COOCH₃), and a long chain containing a sulfonamide group (-SO₂NH-), a polyether chain (-O-CH₂-CH₂-O-CH₂-CH₂-O-CH₂-CH₂-O-), and a carboxylic acid group (-COOH). This chain is further substituted with a 2-amino-2H-imidazo[5,1-b]pyridine moiety and a 4-azido-2-methylphenyl group. The molecule also contains a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety and a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety.
17A'	Chemical structure 17A' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃), a methyl ester group (-COOCH₃), and a long chain containing a sulfonamide group (-SO₂NH-), a polyether chain (-O-CH₂-CH₂-O-CH₂-CH₂-O-CH₂-CH₂-O-), and a carboxylic acid group (-COOH). This chain is further substituted with a 2-amino-2H-imidazo[5,1-b]pyridine moiety and a 4-azido-2-methylphenyl group. The molecule also contains a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety and a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety.
17B'	Chemical structure 17B' is a complex molecule. It features a central benzene ring substituted with a methylthio group (-S-CH₃), a methyl ester group (-COOCH₃), and a long chain containing a sulfonamide group (-SO₂NH-), a polyether chain (-O-CH₂-CH₂-O-CH₂-CH₂-O-CH₂-CH₂-O-), and a carboxylic acid group (-COOH). This chain is further substituted with a 2-amino-2H-imidazo[5,1-b]pyridine moiety and a 4-azido-2-methylphenyl group. The molecule also contains a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety and a 2,6-dimethyl-4-sulfonamido-1,3,5-triazine moiety.

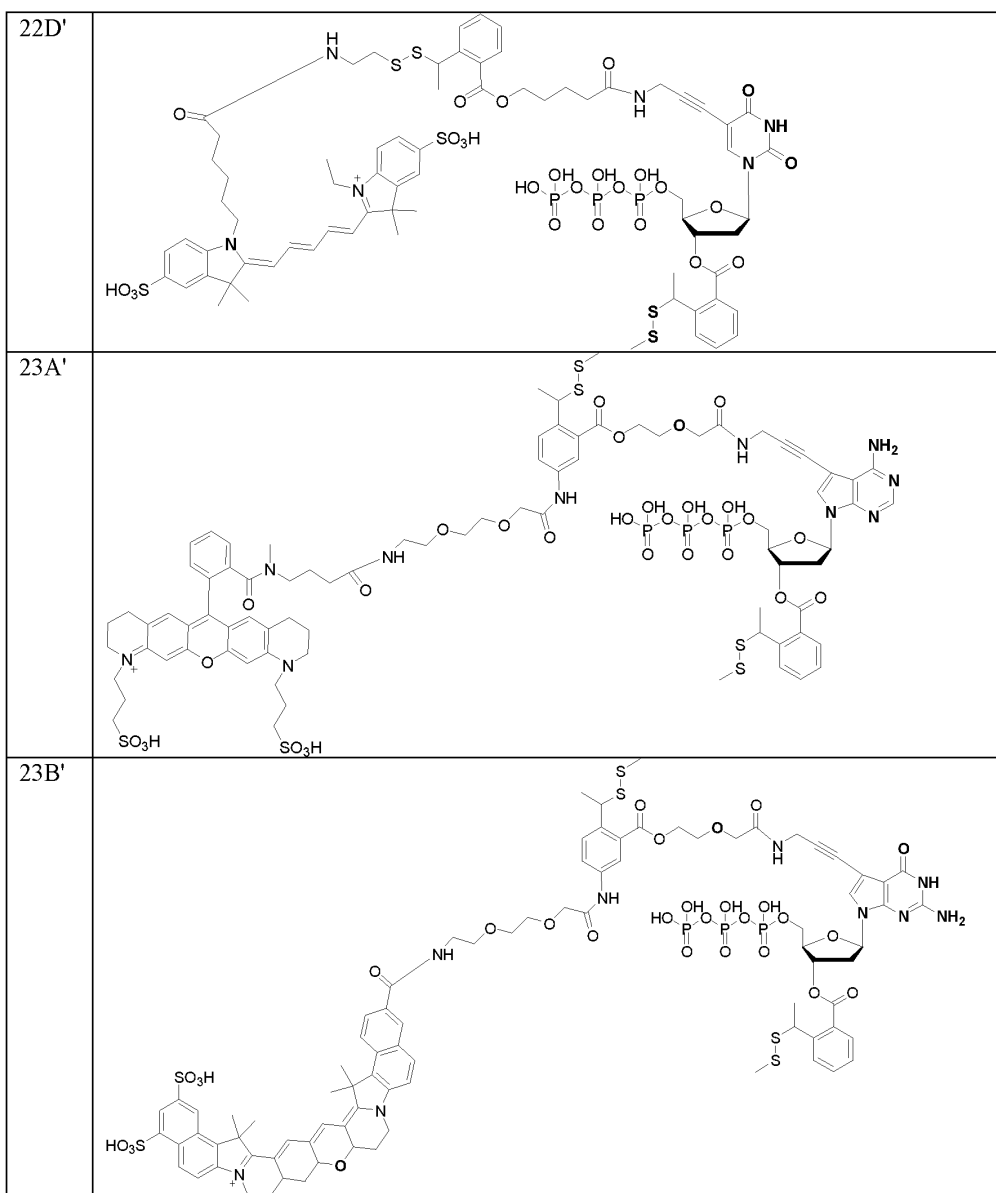


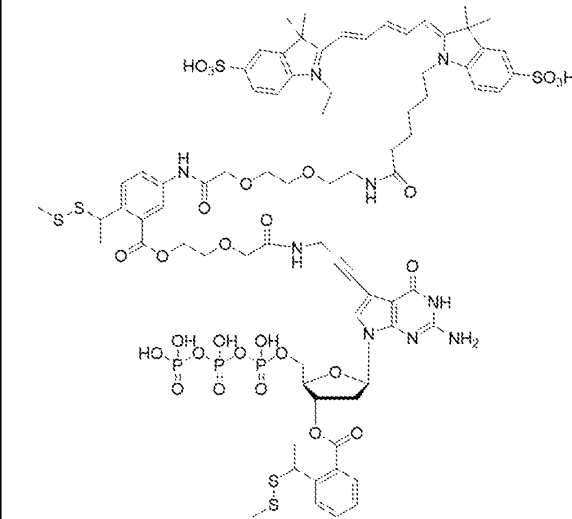
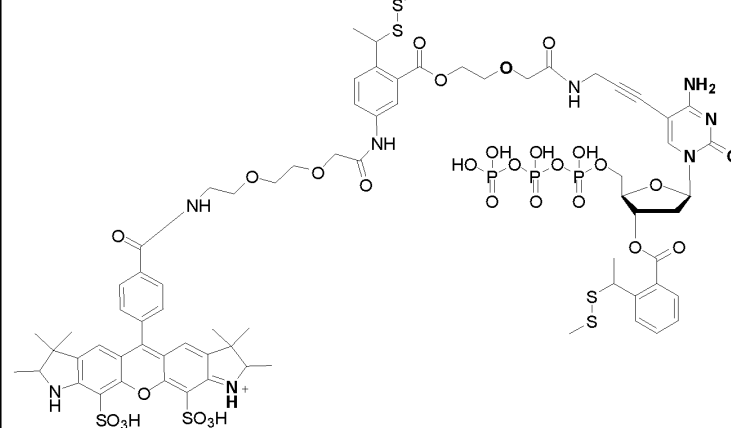
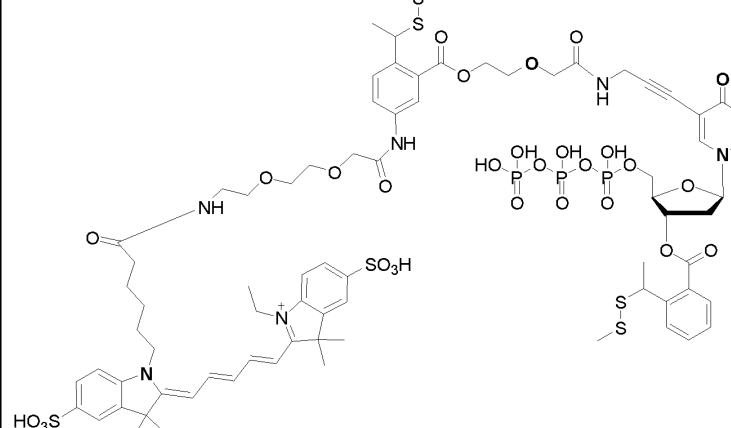
18C'	Chemical structure of 18C' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two trimethylammonium groups (N⁺(CH₃)₃). The macrocycle is linked via a long, flexible chain containing a terminal azide group (N₃) and a terminal amide group (NH). This chain is further connected to a side chain containing a terminal amide group (NH) and a terminal amine group (NH₂). The side chain also includes a terminal amide group (NH) and a terminal amine group (NH₂). The molecule is highly branched and contains multiple functional groups, including sulfonic acid, amine, amide, and azide.
18D'	Chemical structure of 18D' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two trimethylammonium groups (N⁺(CH₃)₃). The macrocycle is linked via a long, flexible chain containing a terminal azide group (N₃) and a terminal amide group (NH). This chain is further connected to a side chain containing a terminal amide group (NH) and a terminal amine group (NH₂). The side chain also includes a terminal amide group (NH) and a terminal amine group (NH₂). The molecule is highly branched and contains multiple functional groups, including sulfonic acid, amine, amide, and azide.
19A'	Chemical structure of 19A' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two trimethylammonium groups (N⁺(CH₃)₃). The macrocycle is linked via a long, flexible chain containing a terminal azide group (N₃) and a terminal amide group (NH). This chain is further connected to a side chain containing a terminal amide group (NH) and a terminal amine group (NH₂). The side chain also includes a terminal amide group (NH) and a terminal amine group (NH₂). The molecule is highly branched and contains multiple functional groups, including sulfonic acid, amine, amide, and azide.
19B'	Chemical structure of 19B' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two trimethylammonium groups (N⁺(CH₃)₃). The macrocycle is linked via a long, flexible chain containing a terminal azide group (N₃) and a terminal amide group (NH). This chain is further connected to a side chain containing a terminal amide group (NH) and a terminal amine group (NH₂). The side chain also includes a terminal amide group (NH) and a terminal amine group (NH₂). The molecule is highly branched and contains multiple functional groups, including sulfonic acid, amine, amide, and azide.
19C'	Chemical structure of 19C' is a complex molecule. It features a central phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two trimethylammonium groups (N⁺(CH₃)₃). The macrocycle is linked via a long, flexible chain containing a terminal azide group (N₃) and a terminal amide group (NH). This chain is further connected to a side chain containing a terminal amide group (NH) and a terminal amine group (NH₂). The side chain also includes a terminal amide group (NH) and a terminal amine group (NH₂). The molecule is highly branched and contains multiple functional groups, including sulfonic acid, amine, amide, and azide.

19D'	Chemical structure of compound 19D'. It features a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule is linked via a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H) to a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H).
20A'	Chemical structure of compound 20A'. It features a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule is linked via a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H) to a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H).
20B'	Chemical structure of compound 20B'. It features a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule is linked via a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H) to a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H).
20C'	Chemical structure of compound 20C'. It features a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H). The molecule is linked via a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H) to a complex molecule with a central indole-like core substituted with a sulfonate group (SO₃H) and a long chain containing a sulfonate group (SO₃H) and a sulfonate group (SO₃H).

20D'	
21A'	
21B'	
21C'	

21D'	
22A'	
22B'	
22C'	



23B''	
23C'	
23D'	

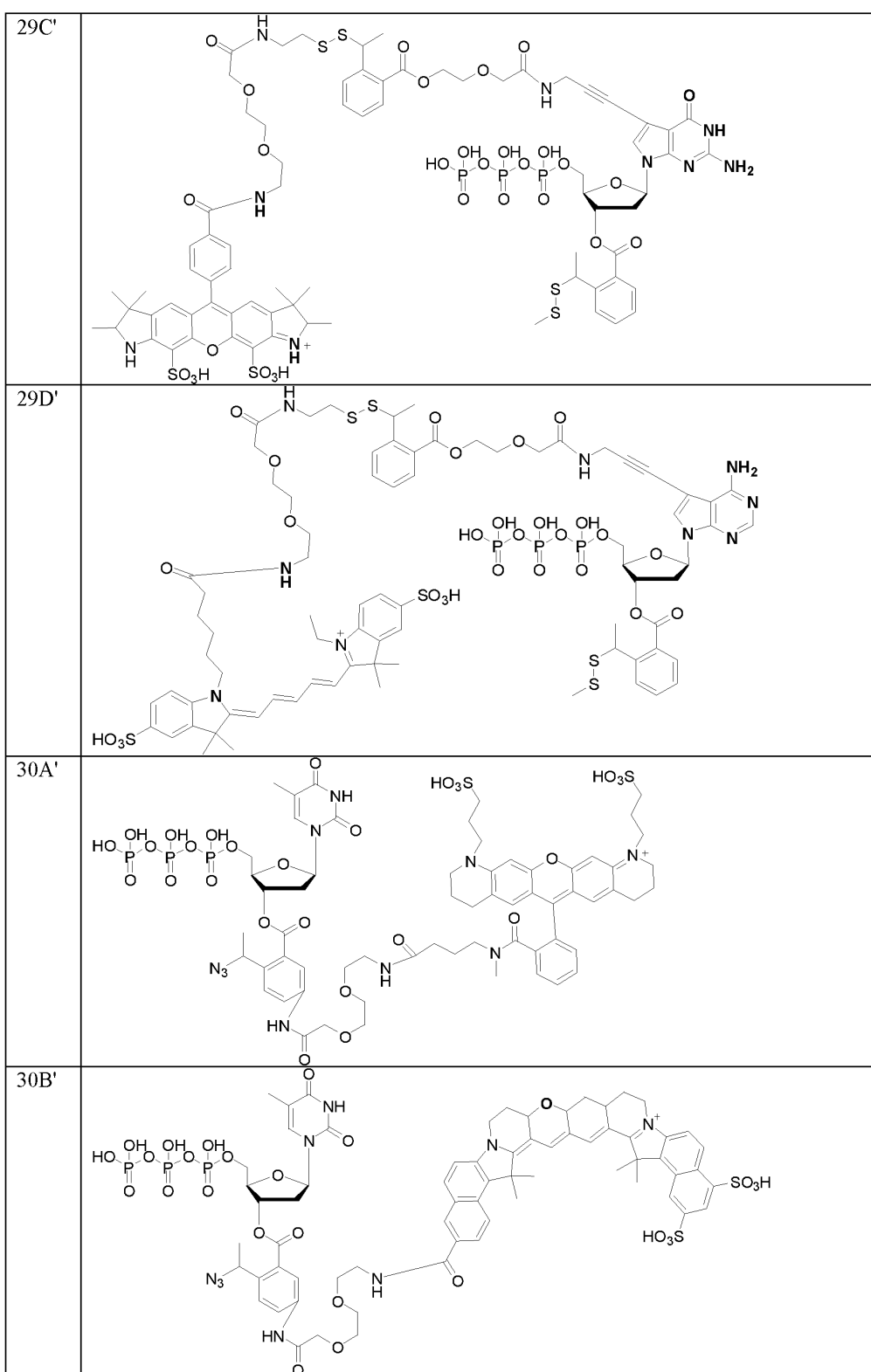
24D'	Chemical structure of compound 24D'. It features a central polyene chain. On the left, there is a sulfonamide group (SO₃H) and a long alkyl chain ending in a sulfonamide group (SO₃H). On the right, there is a sulfonamide group (SO₃H) and a long alkyl chain ending in a sulfonamide group (SO₃H). The structure also includes a pyridine ring, a sulfonamide group (SO₃H), and a sulfonamide group (SO₃H).
25A'	Chemical structure of compound 25A'. It features a central polyene chain. On the left, there is a sulfonamide group (SO₃H) and a long alkyl chain ending in a sulfonamide group (SO₃H). On the right, there is a sulfonamide group (SO₃H) and a long alkyl chain ending in a sulfonamide group (SO₃H). The structure also includes a pyridine ring, a sulfonamide group (SO₃H), and a sulfonamide group (SO₃H).
25B'	Chemical structure of compound 25B'. It features a central polyene chain. On the left, there is a sulfonamide group (SO₃H) and a long alkyl chain ending in a sulfonamide group (SO₃H). On the right, there is a sulfonamide group (SO₃H) and a long alkyl chain ending in a sulfonamide group (SO₃H). The structure also includes a pyridine ring, a sulfonamide group (SO₃H), and a sulfonamide group (SO₃H).

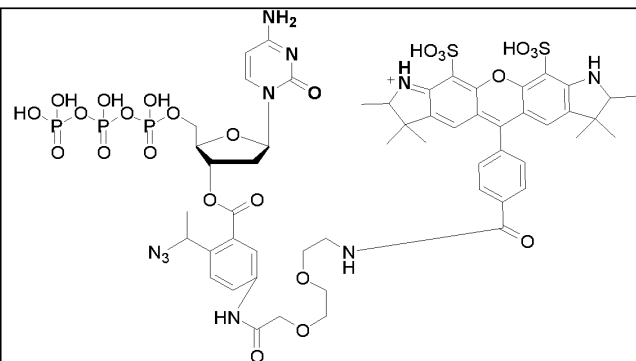
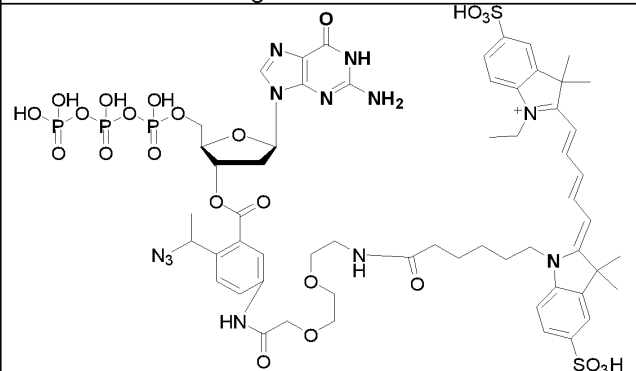
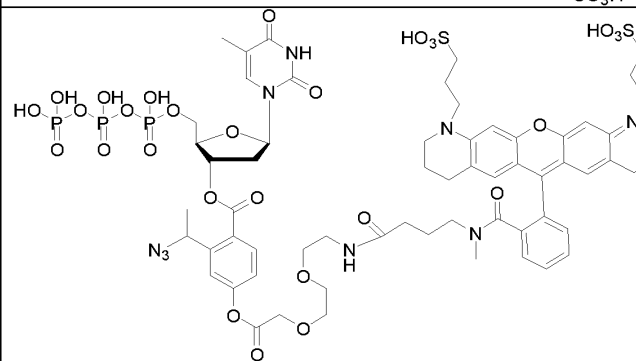
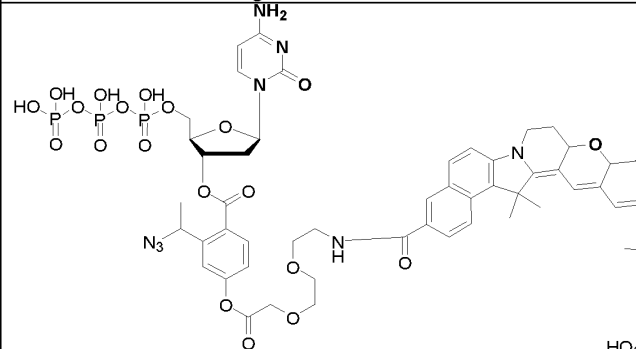
25C'	Chemical structure 25C' is a complex molecule. It features a central xanthine core with two sulfonate groups (SO₃H) at the 2 and 6 positions. The xanthine is linked via a methylene chain to a pyridine ring. The pyridine ring is substituted with a methylthio group (SCH₃) and a methyl group. A long, flexible linker consisting of an amide, an ether, and another amide connects the pyridine ring to a ribose sugar. The ribose sugar is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a thioether group (-S-S-CH₂-CH₃). The ribose is also linked to a pyrimidine ring, which is substituted with an amino group (NH₂) and a carbonyl group (C=O).
25D'	Chemical structure 25D' is a complex molecule. It features a central xanthine core with two sulfonate groups (SO₃H) at the 2 and 6 positions. The xanthine is linked via a methylene chain to a pyridine ring. The pyridine ring is substituted with a methylthio group (SCH₃) and a methyl group. A long, flexible linker consisting of an amide, an ether, and another amide connects the pyridine ring to a ribose sugar. The ribose sugar is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a thioether group (-S-S-CH₂-CH₃). The ribose is also linked to a pyrimidine ring, which is substituted with a carbonyl group (C=O) and a methyl group.
26A'	Chemical structure 26A' is a complex molecule. It features a central xanthine core with two sulfonate groups (SO₃H) at the 2 and 6 positions. The xanthine is linked via a methylene chain to a pyridine ring. The pyridine ring is substituted with a methylthio group (SCH₃) and a methyl group. A long, flexible linker consisting of an amide, an ether, and another amide connects the pyridine ring to a ribose sugar. The ribose sugar is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a thioether group (-S-S-CH₂-CH₃). The ribose is also linked to a pyrimidine ring, which is substituted with a carbonyl group (C=O) and a methyl group.
26B'	Chemical structure 26B' is a complex molecule. It features a central xanthine core with two sulfonate groups (SO₃H) at the 2 and 6 positions. The xanthine is linked via a methylene chain to a pyridine ring. The pyridine ring is substituted with a methylthio group (SCH₃) and a methyl group. A long, flexible linker consisting of an amide, an ether, and another amide connects the pyridine ring to a ribose sugar. The ribose sugar is substituted with a phosphate group (HO-P(=O)(OH)-O-) and a thioether group (-S-S-CH₂-CH₃). The ribose is also linked to a pyrimidine ring, which is substituted with a carbonyl group (C=O) and a methyl group.

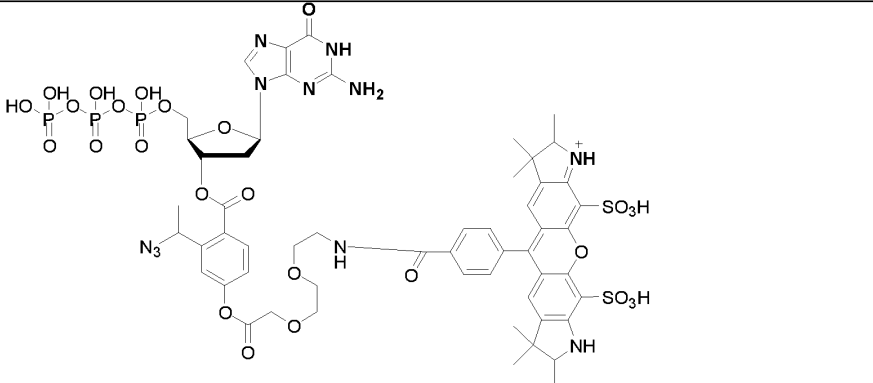
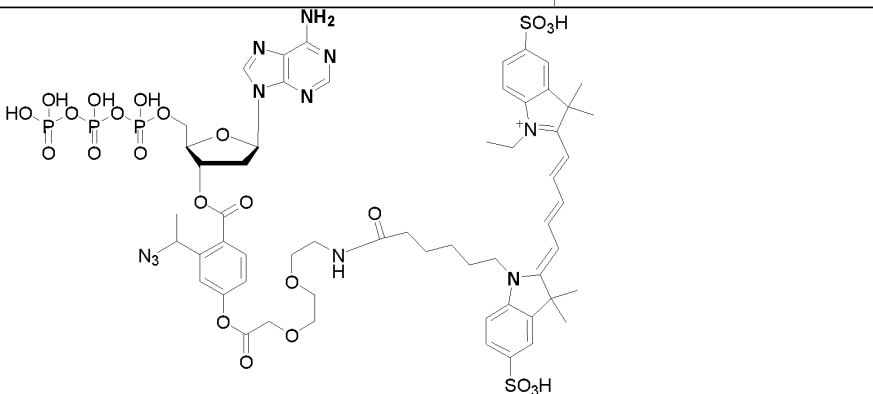
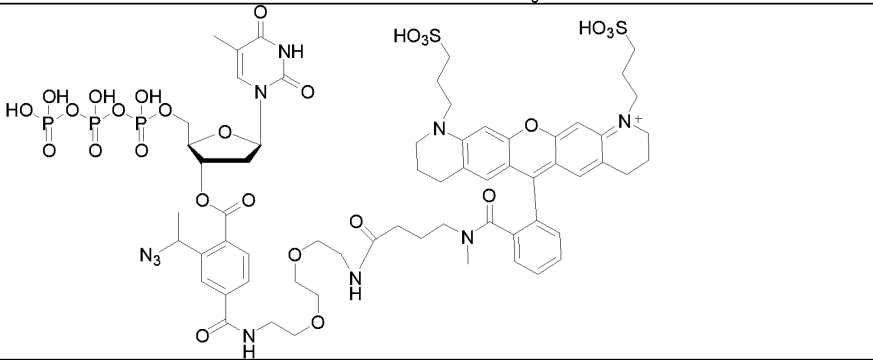
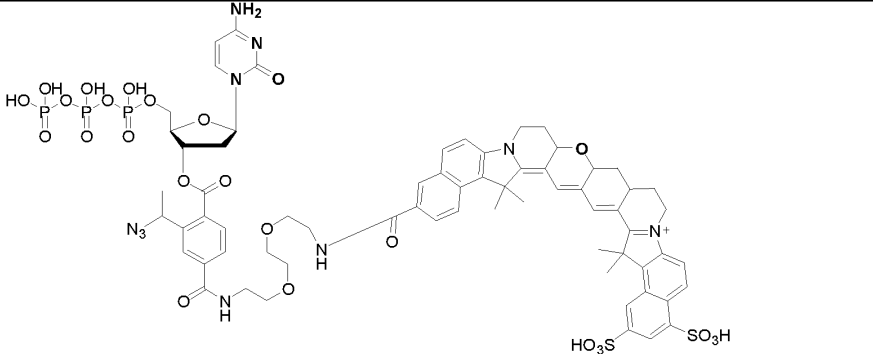
26C'	Chemical structure of compound 26C'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two imine groups. The macrocycle is linked via a long chain containing a disulfide bond, a thioether, and a carbonyl group to a side chain. The side chain includes a long alkyl chain, a carbonyl group, and a terminal group consisting of a triphosphate moiety, a ribose sugar, and a 2-aminobenzimidazole moiety.
26D'	Chemical structure of compound 26D'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two imine groups. The macrocycle is linked via a long chain containing a disulfide bond, a thioether, and a carbonyl group to a side chain. The side chain includes a long alkyl chain, a carbonyl group, and a terminal group consisting of a triphosphate moiety, a ribose sugar, and a 2-aminobenzimidazole moiety.
27A'	Chemical structure of compound 27A'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two imine groups. The macrocycle is linked via a long chain containing a disulfide bond, a thioether, and a carbonyl group to a side chain. The side chain includes a long alkyl chain, a carbonyl group, and a terminal group consisting of a triphosphate moiety, a ribose sugar, and a 2-aminobenzimidazole moiety.
27B'	Chemical structure of compound 27B'. It features a phthalocyanine-like macrocycle with two sulfonic acid groups (SO₃H) and two imine groups. The macrocycle is linked via a long chain containing a disulfide bond, a thioether, and a carbonyl group to a side chain. The side chain includes a long alkyl chain, a carbonyl group, and a terminal group consisting of a triphosphate moiety, a ribose sugar, and a 2-aminobenzimidazole moiety.

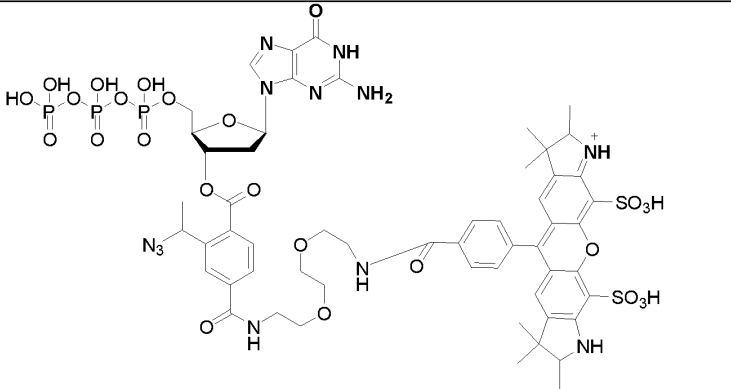
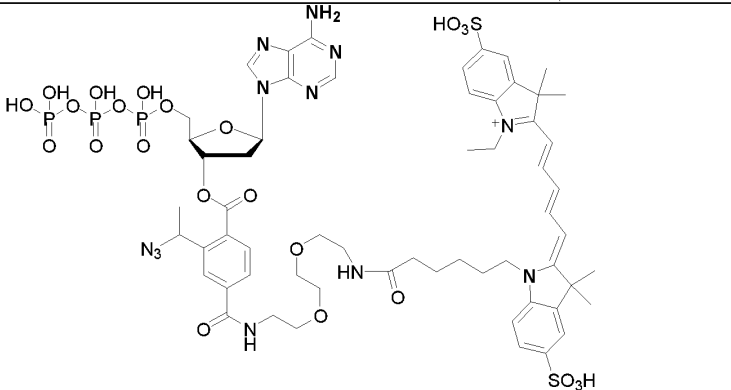
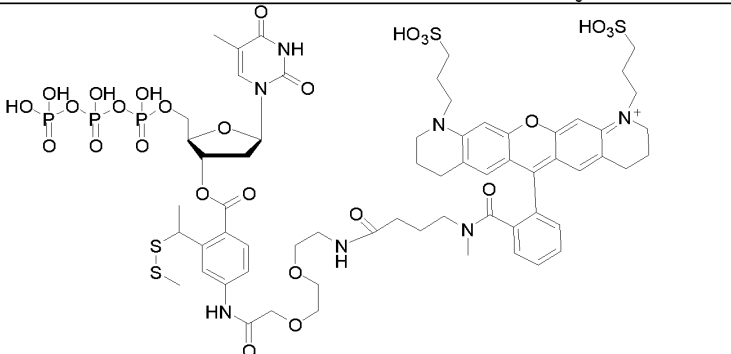
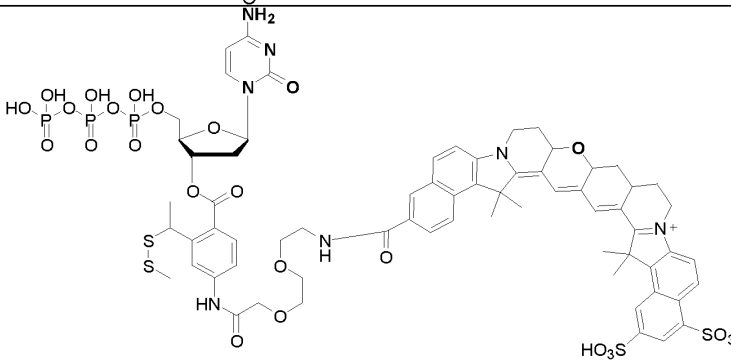
27C'	Chemical structure of compound 27C'. It features a central xanthone core with two sulfonate groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group.
27D'	Chemical structure of compound 27D'. It features a central xanthone core with two sulfonate groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group.
28A'	Chemical structure of compound 28A'. It features a central xanthone core with two sulfonate groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group.
28B'	Chemical structure of compound 28B'. It features a central xanthone core with two sulfonate groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group. The thioether is connected to a benzene ring, which is further substituted with a carbamate group and a long chain containing a carbamate group, a polyether linker, and a thioether group.

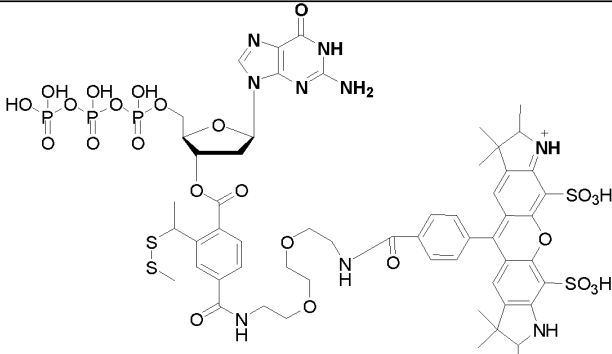
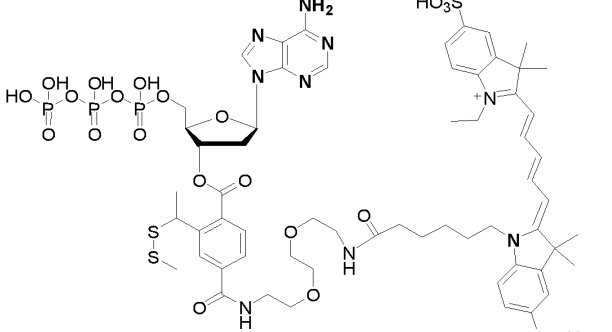
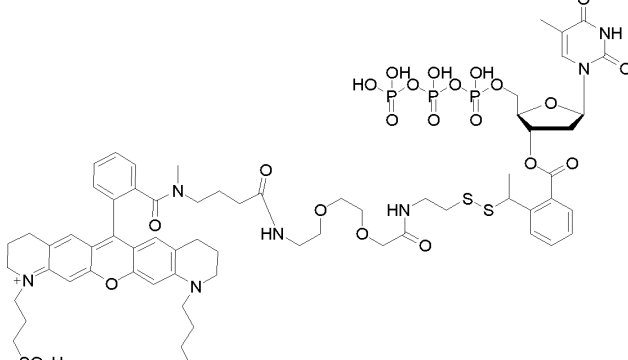
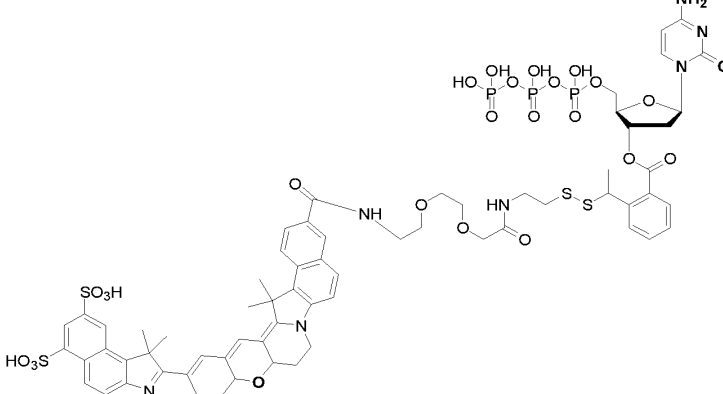
28C'	Chemical structure of compound 28C' is a complex molecule. It features a central xanthone core with two sulfonic acid groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a poly(ethylene glycol) (PEG) linker, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further substituted with a 2-aminopurine-9-yl group. The molecule also contains a thioether linkage and a sulfonamide group.
28D'	Chemical structure of compound 28D' is a complex molecule. It features a central xanthone core with two sulfonic acid groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a poly(ethylene glycol) (PEG) linker, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further substituted with a 2-aminopurine-9-yl group. The molecule also contains a thioether linkage and a sulfonamide group.
29A'	Chemical structure of compound 29A' is a complex molecule. It features a central xanthone core with two sulfonic acid groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a poly(ethylene glycol) (PEG) linker, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further substituted with a 2-aminopurine-9-yl group. The molecule also contains a thioether linkage and a sulfonamide group.
29B'	Chemical structure of compound 29B' is a complex molecule. It features a central xanthone core with two sulfonic acid groups (SO₃H) at the 1 and 8 positions. The xanthone is substituted with a long chain containing a carbamate group, a poly(ethylene glycol) (PEG) linker, and a phosphate group. The phosphate group is linked to a ribose sugar, which is further substituted with a 2-aminopurine-9-yl group. The molecule also contains a thioether linkage and a sulfonamide group.

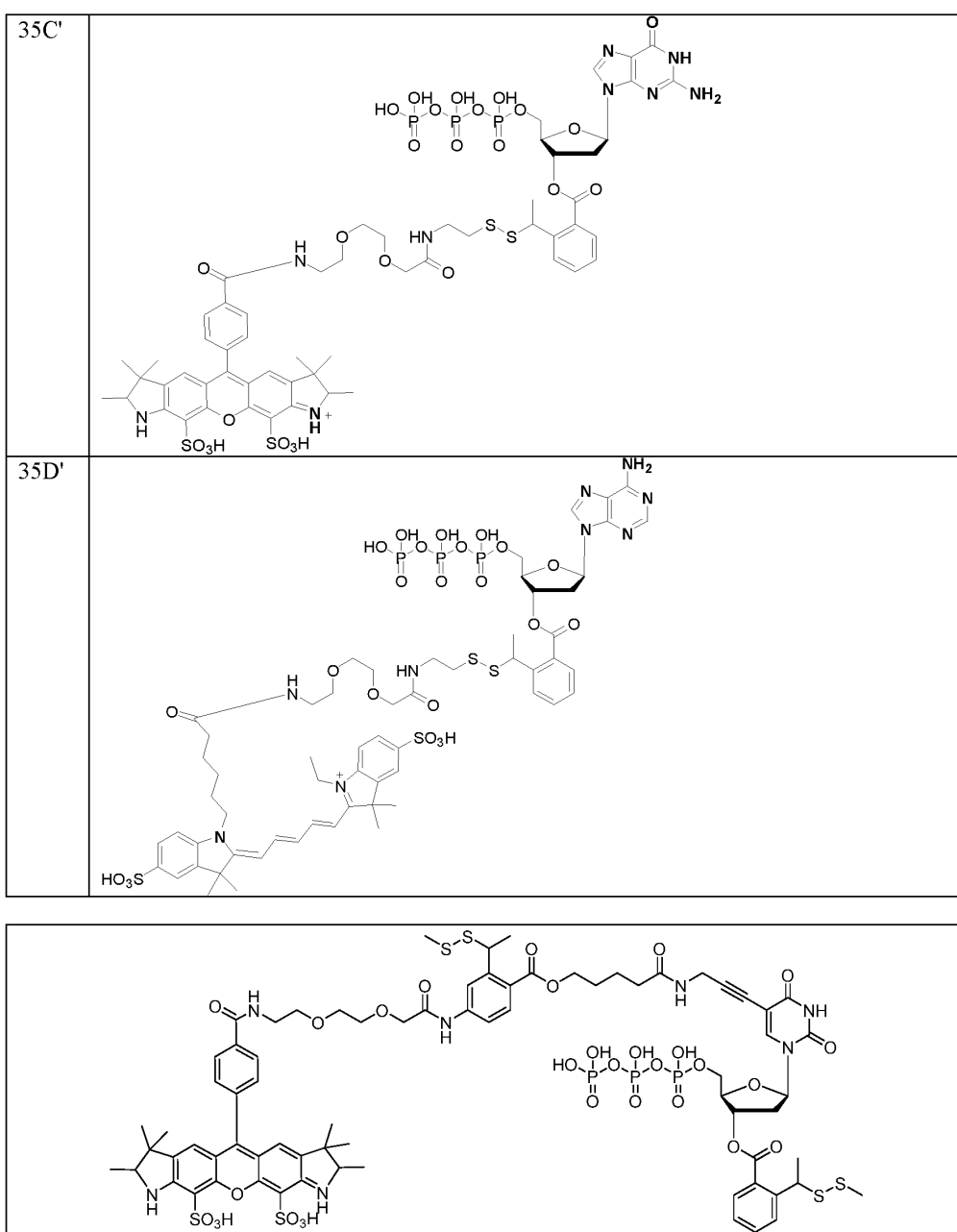


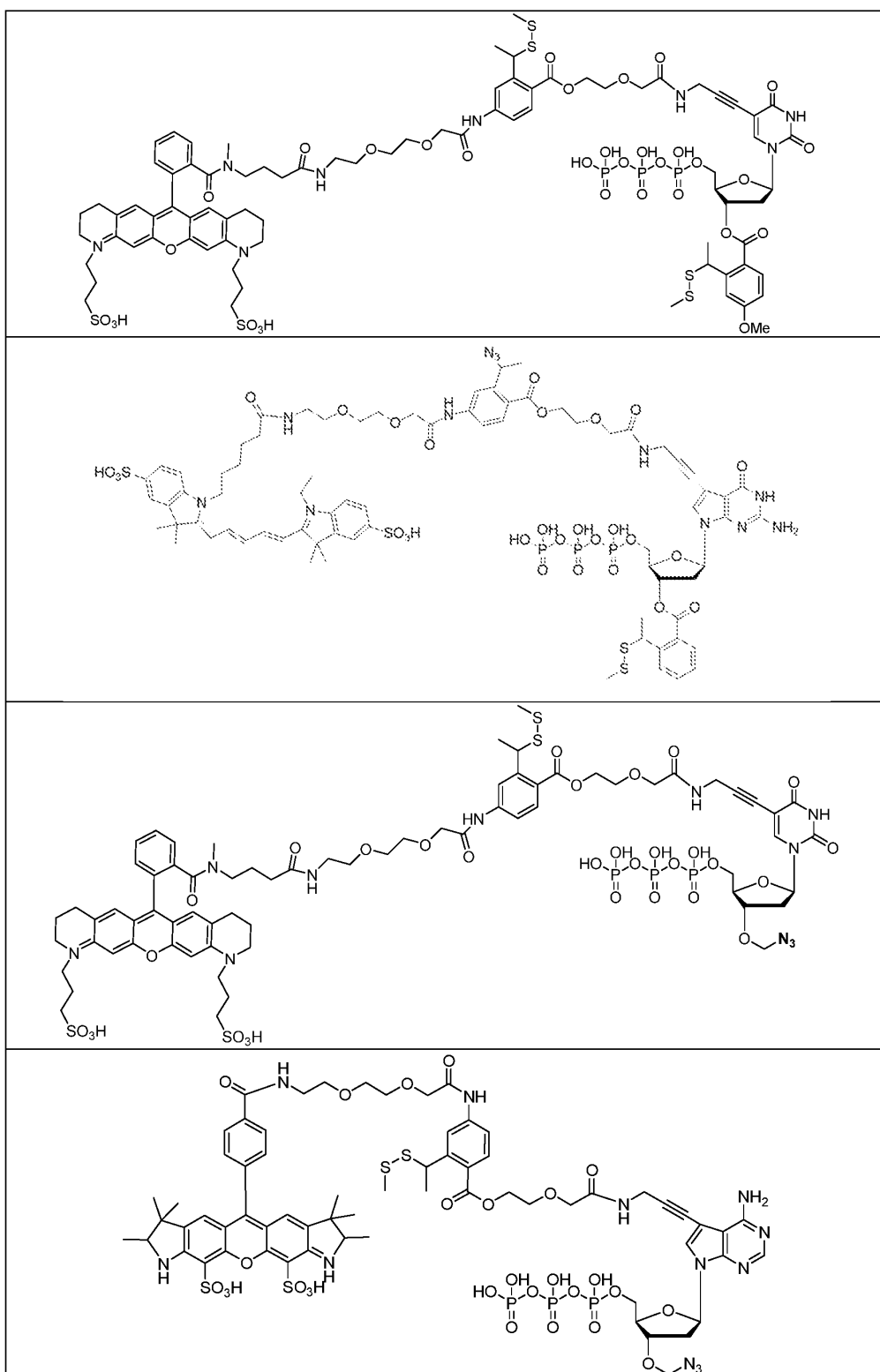
30C'	
30D'	
31A'	
31B'	

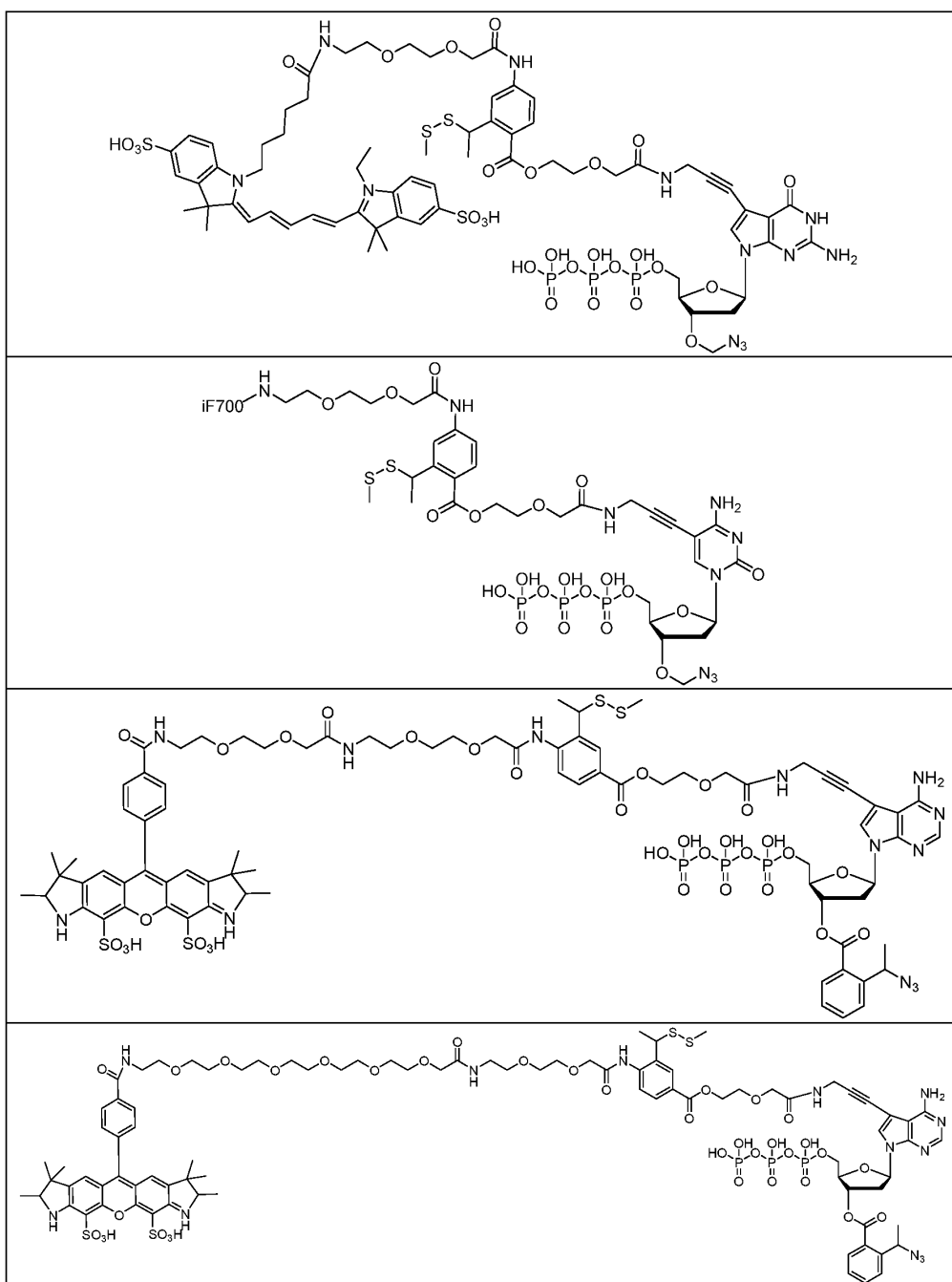
31C'	
31D'	
32A'	
32B'	

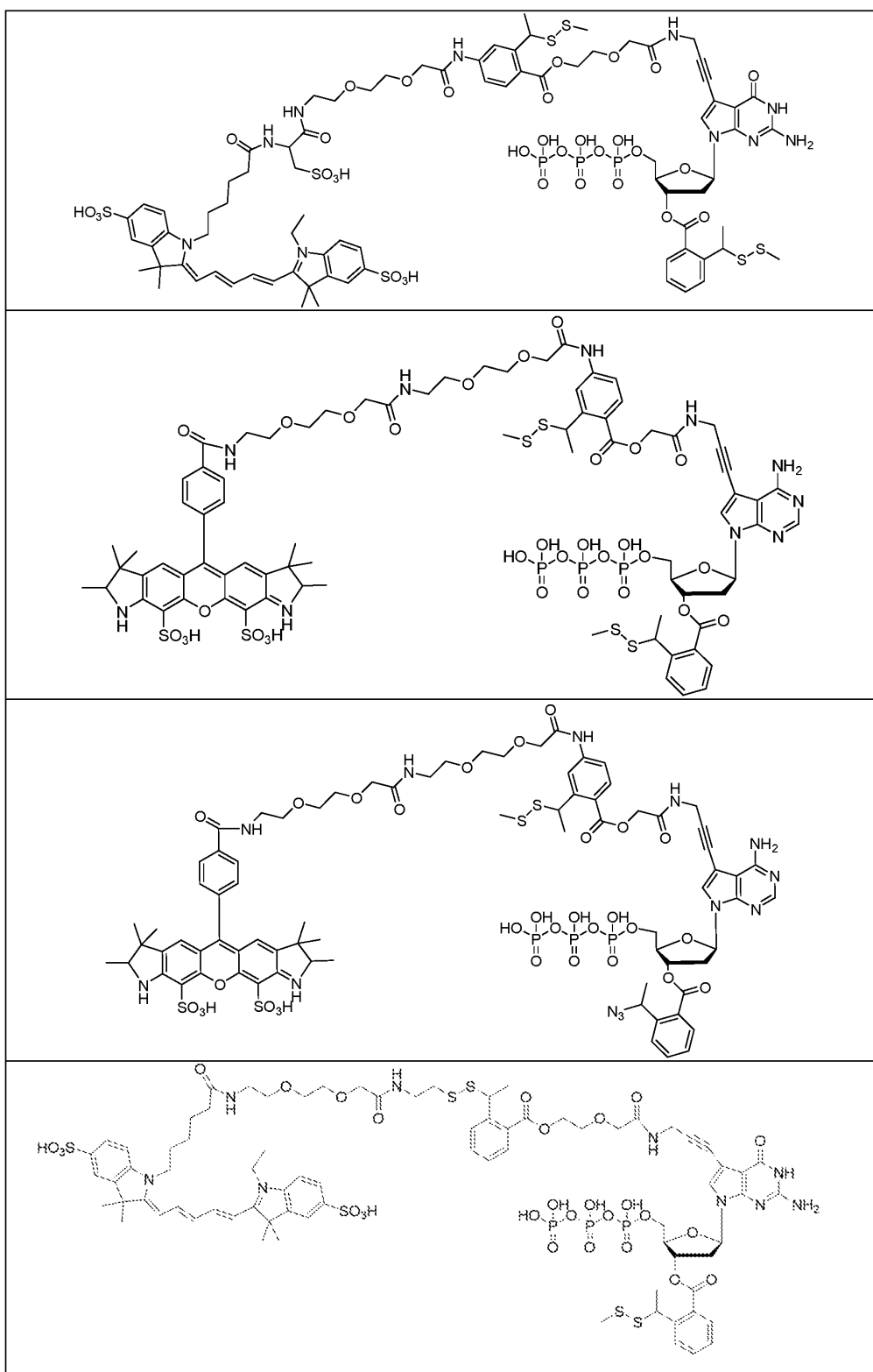
32C'	
32D'	
33A'	
33B'	

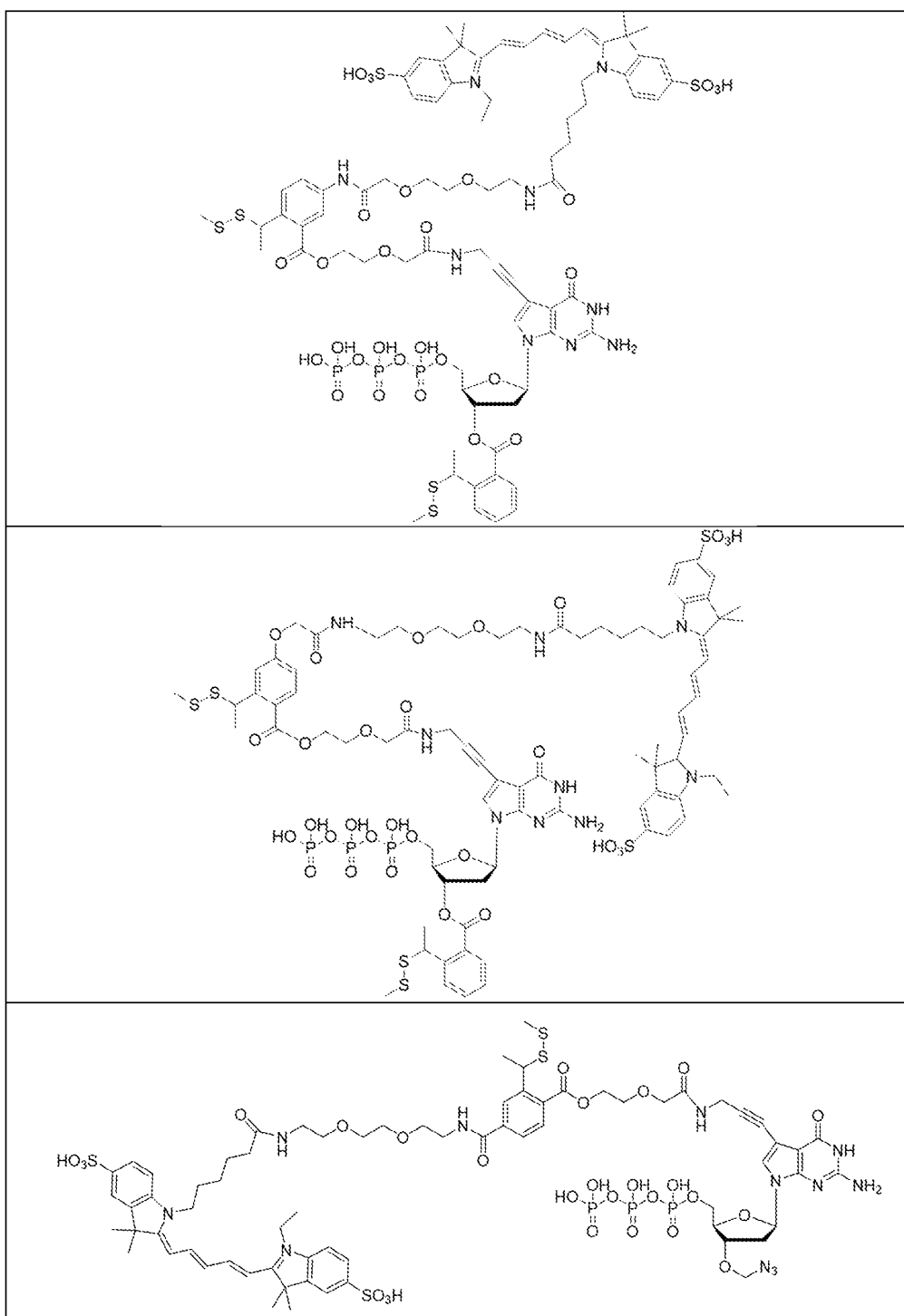
34C'	
34D'	
35A'	
35B'	











12、终止核酸合成的方法，其包括：将权利要求 1-11 任一项所述的化合物或其盐掺入待终止的核酸分子中；

优选地，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述化合物或其盐掺入待终止的核酸分子中；

优选地,所述方法包括:在允许聚合酶进行核苷酸聚合反应的条件下,使用聚合酶进行核苷酸聚合反应,从而将所述化合物或其盐掺入待终止的核酸分子的3'端。

13、制备在测序反应中与目标单链多核苷酸互补的生长的多核苷酸的方法,其包括将权利要求1-11中任一项所述的化合物或其盐掺入所述生长的互补多核苷酸,其中,所述化合物或其盐的掺入防止了任何后续的核苷酸引入所述生长的互补多核苷酸中;

优选地,所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现;

优选地,所述方法包括:使用聚合酶,将所述化合物或其盐掺入所述生长的互补多核苷酸;

优选地,所述方法包括:在允许聚合酶进行核苷酸聚合反应的条件下,使用聚合酶进行核苷酸聚合反应,从而将所述化合物或其盐掺入所述生长的互补多核苷酸的3'端。

14、核酸中间体,其是在测定目标单链多核苷酸的序列中形成的,其中,

所述核酸中间体是通过以下步骤形成的:

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸,形成所述核酸中间体,其中,掺入的一个互补核苷酸是权利要求1-11中任一项所述的化合物或其盐;

或者,所述核酸中间体是通过以下步骤形成的:

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸,形成所述核酸中间体,其中,掺入的一个互补核苷酸是权利要求1-11中任一项所述的化合物或其盐,且所述生长的核酸链中预先掺入至少一个与目标单链多核苷酸互补的核苷酸,预先掺入的至少一个与目标单链多核苷酸互补的核苷酸是已被除去可逆阻断基团和任选的可检测标记的权利要求1-11中任一项所述的化合物或其盐。

15、测定目标单链多核苷酸的序列的方法,其包括:

1) 监测生长的核酸链中与目标单链多核苷酸互补的核苷酸的掺入,其中,掺入的至少一个互补核苷酸是权利要求1-11中任一项所述的化合物或其盐,以及,

2) 确定掺入的核苷酸的类型;

优选地,在引入下一个互补核苷酸之前,将所述可逆阻断基团和任选的可检测标记除去;

优选地,所述可逆阻断基团和所述可检测标记被同时除去;

优选地,所述可逆阻断基团和所述可检测标记被先后除去;例如,在所述可检测标记被除去之后,所述可逆阻断基团被除去,或者,在所述可逆阻断基团被除去之后,所述可检测标记被除去。

16、权利要求15的方法,其包括以下步骤:

(a) 提供多种不同的核苷酸,其中至少一种核苷酸是权利要求10-11任一项所述的化合物或其盐,任选地其余的核苷酸是权利要求1-11任一项所述的化合物或其盐;

(b) 将所述多种不同的核苷酸掺入目标单链多核苷酸的互补序列中,其中,所述多种不同的核苷酸在检测时可以相互区分开;

(c) 检测(b)的核苷酸,从而确定掺入的核苷酸的类型;

(d) 除去(b)的核苷酸中的可逆阻断基团和任选的其携带的可检测标记;和

(e) 任选地重复步骤(a) - (d)一次或多次;

从而确定所述目标单链多核苷酸的序列。

17、权利要求15的方法,其包括以下步骤:

(1) 提供第一核苷酸、第二核苷酸、第三核苷酸和第四核苷酸,四种核苷酸中的至少一种是权利要求10-11中任一项所述的化合物或其盐,任选地其余的核苷酸是权利要求1-11任一项所述的化合物或其盐;

(2) 将所述四种核苷酸与目标单链多核苷酸进行接触;除去未掺入生长的核酸链中的所述核苷酸;检测掺入生长的核酸链中的所述核苷酸;除去掺入生长的核酸链中的所述核苷酸中的所述可逆阻断基团和任选的其携带的所述可检测标记;

任选地,还包括(3):重复(1) - (2)一次或多次。

18、权利要求 15 的方法，其包括以下步骤：

(a) 提供包含双链体、包含至少一种权利要求 10-11 中任一项所述的化合物或其盐的核苷酸、聚合酶和切除试剂的混合物；所述双链体包含生长的核酸链以及待测定的核酸链；

(b) 进行包含以下步骤 (i)、(ii) 和 (iii) 的反应，任选地，重复一次或多次：

步骤 (i)：使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体；

步骤 (ii)：对所述核酸中间体进行检测；

步骤 (iii)：使用切除试剂将所述核酸中间体所包含的可逆阻断基团和任选的可检测标记切除；

优选地，对所述可逆阻断基团的切除和对所述可检测标记的切除同时进行，或者，对所述可逆阻断基团的切除和对所述可检测标记的切除分步进行（例如，先切除所述可逆阻断基团，或者先切除所述可检测标记）；

优选地，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是同样的试剂；

优选地，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是不同的试剂。

19、权利要求 18 的方法，其中，所述双链体连接于支持物上；

优选地，所述生长的核酸链为引物；

优选地，所述引物通过退火至待测定的核酸链上，形成所述双链体；

优选地，所述双链体、所述化合物或其盐、以及所述聚合酶一起形成含有溶液相和固相的反应体系；

优选地，在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体；

优选地，所述聚合酶选自 KOD 聚合酶或其突变体（例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391）；

优选地，在任意一个检测所述核酸中间体的步骤前，移除前一步骤的反应体系的溶液相，保留连接于支持物上的双链体；

优选地，所述切除试剂与所述双链体或所述生长的核酸链在含有溶液相和固相的反应体系中接触；

优选地，所述切除试剂能够切除掺入生长的核酸链的化合物中的可逆阻断基团和任选的其携带的可检测标记，并且不会影响双链体骨架上的磷酸二酯键；

优选地，在任意一个切除所述核酸中间体所包含的可逆阻断基团和任选的可检测标记的步骤后，移除这一步骤反应体系的溶液相；

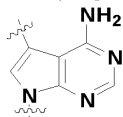
优选地，在任意一个包含移除操作的步骤之后，进行洗涤操作；

优选地，步骤 (ii) 之后，进一步包括：根据步骤 (ii) 检测得到的信号，确定步骤 (i) 中掺入生长的核酸链的化合物的类型，并基于碱基互补配对原则，确定待测定的核酸链中相应位置处的核苷酸类型。

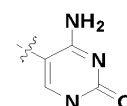
20、试剂盒，其包含至少一个权利要求 1-11 任一项所述的化合物或其盐；

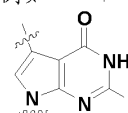
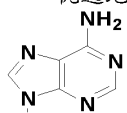
优选地，所述试剂盒包含第一、第二、第三和第四化合物，所述第一、第二、第三和第四化合物各自独立地为权利要求 1-11 任一项所述的化合物或其盐；

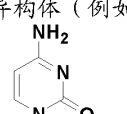
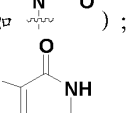
优选地，所述第一化合物中，base¹选自腺嘌呤、7-脱氮腺嘌呤或其互变异构体（例如

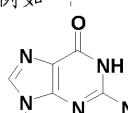


）；所述第二化合物中，base¹选自胸腺嘧啶、尿嘧啶或其互变异构体（例如

); 所述第三化合物中, base¹选自胞嘧啶或其互变异构体(例如 );

所述第四化合物中, base¹选自鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体(例如 );
优选地, 所述第一化合物中, base²选自腺嘌呤、7-脱氮腺嘌呤或其互变异构体(例如 );

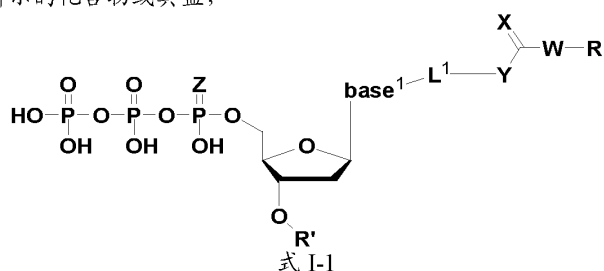
); 所述第二化合物中, base²选自胸腺嘧啶、尿嘧啶或其互变异构体(例如 ); 所
述第三化合物中, base²选自胞嘧啶或其互变异构体(例如 ); 所

述第四化合物中, base²选自鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体(例如 );
优选地, 所述第一、第二、第三和第四化合物包含的 base¹或 base²互不相同;
优选地, 所述第一、第二、第三和第四化合物携带的额外的可检测标记互不相同。

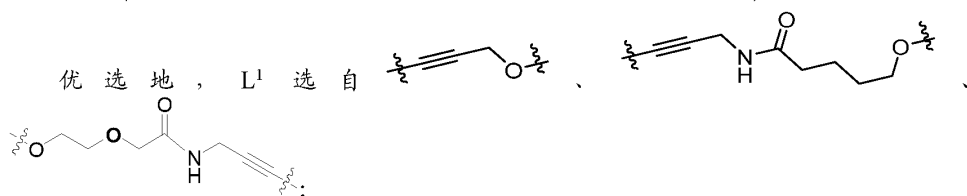
21、权利要求 20 的试剂盒, 其中, 所述试剂盒还包含: 用于预处理核酸分子的试剂;
用于连接待测序的核酸分子的支持物; 用于将待测序的核酸分子与支持物连接(例如, 共价
或非共价连接)的试剂; 用于起始核苷酸聚合反应的引物; 用于进行核苷酸聚合反应的聚合
酶; 一种或多种缓冲溶液; 一种或多种洗涤溶液; 或其任何组合。

22、权利要求 1-11 任一项所述的化合物或其盐或者权利要求 20-21 任一项所述的试剂
盒用于测定目标单链多核苷酸的序列的用途。

23、式 I-1 所示的化合物或其盐,



其中:



r^1 、 r^2 、 r^{3a} 、 r^{3b} 各自独立地选自 1-6 之间的任意整数;

优选地, r^1 选自 1、2、3;

更优选地, r^1 为 1;

优选地, r^2 选自 1、2、3;

更优选地, r^2 为 1;

优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3、4、5;

更优选地, r^{3a} 、 r^{3b} 各自独立地选自 1、2、3;

最优选地, r^{3a} 为 1;

最优选地, r^{3b} 为 2;

M 选自直接键、 CH_2 、NH、O、S;

优选地, M 选自 CH_2 、O;

X 选自 O、S、NH;

优选地, X 选自 O、S;

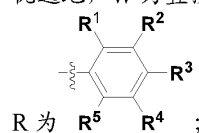
更优选地, X 为 O;

Y 选自直接键、O、S、NH;

优选地, Y 为直接键;

W 选自直接键、O、S、NH;

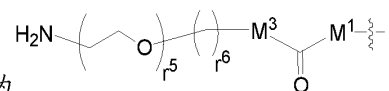
优选地, W 为直接键;



R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, 任意一个(如 R^2 、 R^3 或 R^4) 为



另一个(如 R^1 或 R^5) 为  , 剩余三个为 H;



优选地, R^1 、 R^2 、 R^3 、 R^4 、 R^5 中, R^3 或 R^4 为



R^1 为  , 剩余三个为 H;

r^5 、 r^6 各自独立地选自 1-6 之间的任意整数;

优选地, r^5 选自 1、2、3;

更优选地, r^5 为 2;

优选地, r^6 选自 1、2、3;

更优选地, r^6 为 1 或 2;

M^1 、 M^3 各自独立地选自直接键、NH、O、S;

优选地, M^1 选自直接键、NH、O;

优选地, M^3 选自直接键、NH;

更优选地, M^1 为 NH、O 或 S 时, M^3 为直接键, 且 M^3 为 NH、O 或 S 时, M^1 为直接键;

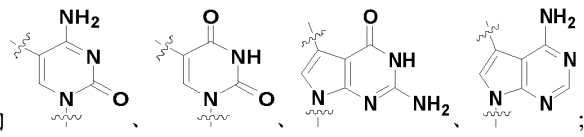
R^b 、 R^c 中, 任意一个选自 $-\text{N}_3$ 、 $-\text{SS}-\text{C}_1-\text{C}_6$ 烷基(如 $-\text{SS}-\text{Me}$ 、 $-\text{SS}-\text{Et}$ 、 $-\text{SS}-\text{iPr}$ 、 $-\text{SS}-\text{t-Bu}$), 另一个选自 C_1-C_6 烷基;

优选地, R^b 、 R^c 中, 任意一个选自 $-\text{N}_3$ 、 $-\text{SS}-\text{Me}$, 另一个为甲基;

Z 选自 O、S、BH;

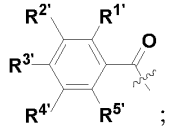
优选地, Z 为 O;

base^1 选自碱基、脱氮碱基或其互变异构体, 例如腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;

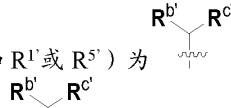


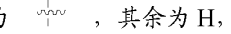
优选地, base¹选自
R'表示可逆阻断基团。

24、权利要求 23 所述的化合物或其盐, 其中, 所述可逆阻断基团 R'选自 N₃-CH₂-、



CH₃-CH₂-S-S-CH₂-、

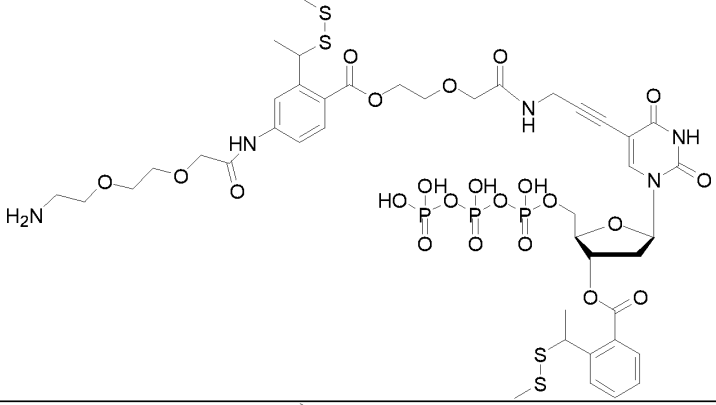
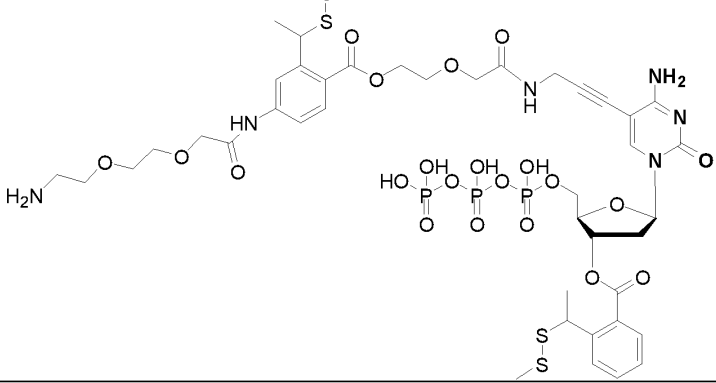
R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, 任意一个 (如 R^{1'}或 R^{5'}) 为 , 其余为 H,

优选地, R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, R^{1'}为 , 其余为 H,

R^b、R^c中, 任意一个选自 -N₃、-SS-C₁-C₆ 烷基 (如 -SS-Me, -SS-Et, -SS-iPr, -SS-t-Bu),
另一个选自 C₁-C₆ 烷基,

优选地, R^b、R^c中, 任意一个选自 -N₃、-SS-Me, 另一个为甲基。

25、权利要求 23-24 任一项所述的化合物或其盐, 其中, 所述化合物选自以下:

编号	结构式
7A	
7B	

7C	
7D	
7D1	

26、权利要求 23-25 任一项所述的化合物或其盐，其中，所述化合物或其盐携带额外的可检测标记；

优选地，所述化合物或其盐携带的额外的可检测标记是通过亲和试剂（如抗体、适体、Affimer、Knottin）引入的，所述亲和试剂携带所述可检测标记，且所述亲和试剂可以特异性识别并结合所述化合物或其盐的表位；

或者优选地，所述可检测标记与权利要求 23-24 任一项所述的式 I-1 所示化合物中的 R 或权利要求 25 的化合物中对应于 R 的结构部分连接；

优选地，所述可检测标记与

中的末端氨基连接；

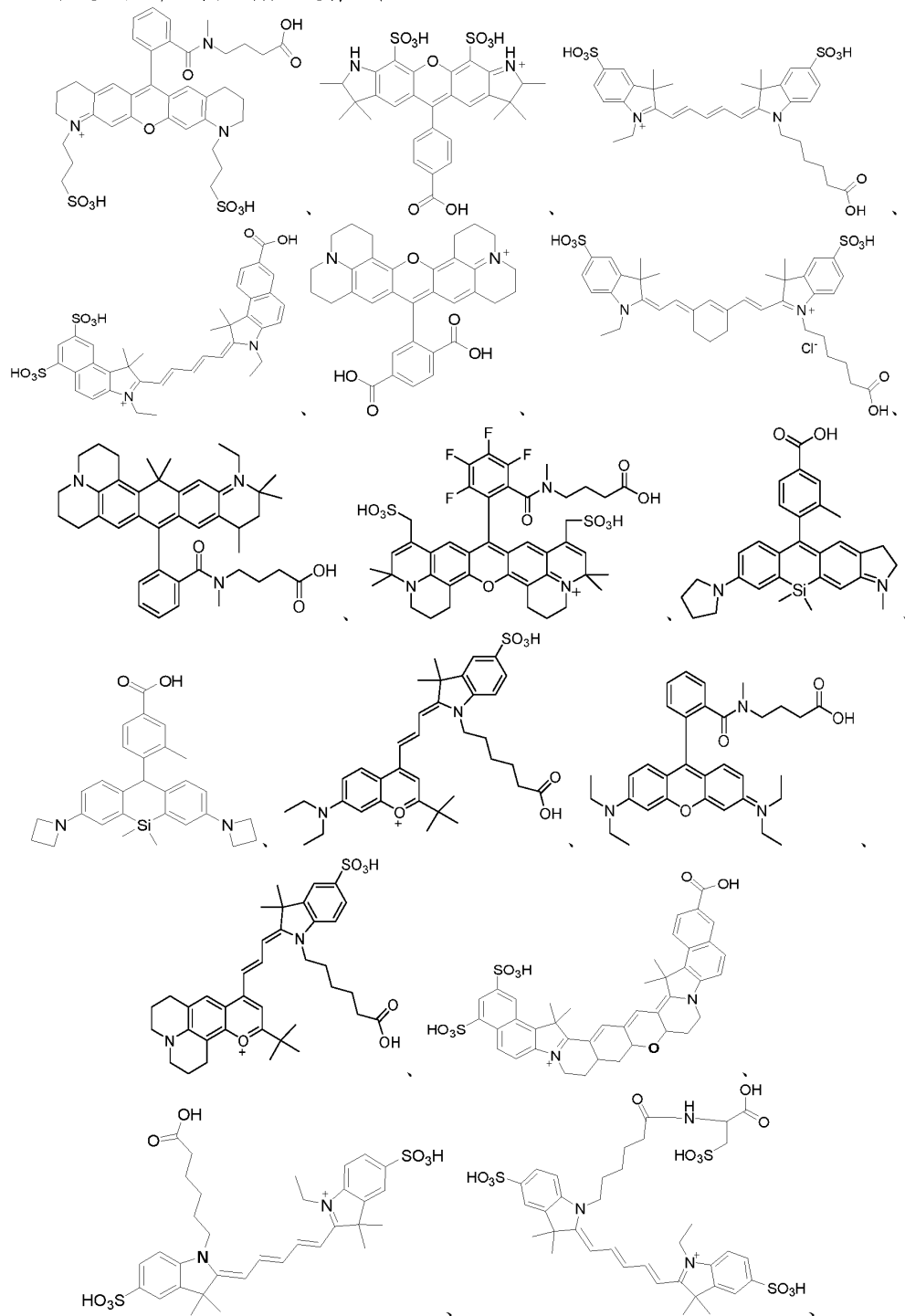
优选地，所述可检测标记中的羧基与

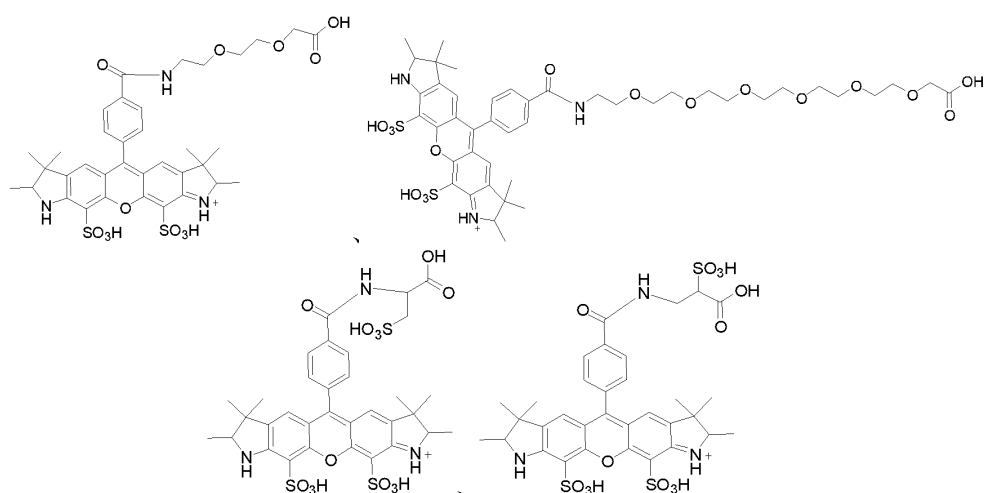
 中的末端氨基通过形成酰胺键进行连接；

优选地，base¹ 不同，式 I-1 所示化合物携带的额外的可检测标记不同；

优选地, 所述可检测标记为荧光标记;

优选地, 所述可检测标记选自以下: iF700、





27、权利要求 26 所述的化合物或其盐，其中，所述化合物选自以下：

编号	结构式
7A'	
7B'	

7B''	
7C'	
7C''	
7C'''	

7C ^{'''}	
7C ^{''''}	
7D'	
7D''	

°

28、终止核酸合成的方法，其包括：将权利要求 23-27 任一项所述的化合物或其盐掺入

待终止的核酸分子中；

优选地，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述化合物或其盐掺入待终止的核酸分子中；

优选地，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述化合物或其盐掺入待终止的核酸分子的3'端。

29、制备在测序反应中与目标单链多核苷酸互补的生长的多核苷酸的方法，其包括将权利要求23-27中任一项所述的化合物或其盐掺入所述生长的互补多核苷酸，其中，所述化合物或其盐的掺入防止了任何后续的核苷酸引入所述生长的互补多核苷酸中；

优选地，所述化合物或其盐的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述化合物或其盐掺入所述生长的互补多核苷酸；

优选地，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述化合物或其盐掺入所述生长的互补多核苷酸的3'端。

30、核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，

所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是权利要求23-27中任一项所述的化合物或其盐；

或者，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是权利要求23-27中任一项所述的化合物或其盐，且所述生长的核酸链中预先掺入至少一个与目标单链多核苷酸互补的核苷酸，预先掺入的至少一个与目标单链多核苷酸互补的核苷酸是已被除去可逆阻断基团和任选的可检测标记的权利要求23-27中任一项所述的化合物或其盐。

31、测定目标单链多核苷酸的序列的方法，其包括：

1) 监测生长的核酸链中与目标单链多核苷酸互补的核苷酸的掺入，其中，掺入的至少一个互补核苷酸是权利要求23-27中任一项所述的化合物或其盐，以及，

2) 确定掺入的核苷酸的类型；

优选地，在引入下一个互补核苷酸之前，将所述可逆阻断基团和任选的可检测标记除去；

优选地，所述可逆阻断基团和所述可检测标记被同时除去；

优选地，所述可逆阻断基团和所述可检测标记被先后除去；例如，在所述可检测标记被除去之后，所述可逆阻断基团被除去，或者，在所述可逆阻断基团被除去之后，所述可检测标记被除去。

32、权利要求31的方法，其包括以下步骤：

(a) 提供多种不同的核苷酸，其中至少一种核苷酸是权利要求26-27任一项所述的化合物或其盐，任选地其余的核苷酸是权利要求23-27任一项所述的化合物或其盐；

(b) 将所述多种不同的核苷酸掺入目标单链多核苷酸的互补序列中，其中，所述多种不同的核苷酸在检测时可以相互区分开；

(c) 检测(b)的核苷酸，从而确定掺入的核苷酸的类型；

(d) 除去(b)的核苷酸中的可逆阻断基团和任选的其携带的可检测标记；和

(e) 任选地重复步骤(a) - (d)一次或多次；

从而确定所述目标单链多核苷酸的序列。

33、权利要求31的方法，其包括以下步骤：

(1) 提供第一核苷酸、第二核苷酸、第三核苷酸和第四核苷酸，四种核苷酸中的至少一种是权利要求26-27中任一项所述的化合物或其盐，任选地其余的核苷酸是权利要求23-27任一项所述的化合物或其盐；

(2) 将所述四种核苷酸与目标单链多核苷酸进行接触；除去未掺入生长的核酸链中的

所述核苷酸；检测掺入生长的核酸链中的所述核苷酸；除去掺入生长的核酸链中的所述核苷酸中的所述可逆阻断基团和任选的其携带的所述可检测标记；

任选地，还包括 (3)：重复 (1) - (2) 一次或多次。

34、权利要求 31 的方法，其包括以下步骤：

(a) 提供包含双链体、包含至少一种权利要求 26-27 中任一项所述的化合物或其盐的核苷酸、聚合酶和切除试剂的混合物；所述双链体包含生长的核酸链以及待测序的核酸链；

(b) 进行包含以下步骤 (i)、(ii) 和 (iii) 的反应，任选地，重复一次或多次：

步骤 (i)：使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体；

步骤 (ii)：对所述核酸中间体进行检测；

步骤 (iii)：使用切除试剂将所述核酸中间体所包含的可逆阻断基团和任选的可检测标记切除；

优选地，对所述可逆阻断基团的切除和对所述可检测标记的切除同时进行，或者，对所述可逆阻断基团的切除和对所述可检测标记的切除分步进行（例如，先切除所述可逆阻断基团，或者先切除所述可检测标记）；

优选地，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是同样的试剂；

优选地，对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是不同的试剂。

35、权利要求 34 的方法，其中，所述双链体连接于支持物上；

优选地，所述生长的核酸链为引物；

优选地，所述引物通过退火至待测序的核酸链上，形成所述双链体；

优选地，所述双链体、所述化合物或其盐、以及所述聚合酶一起形成含有溶液相和固相的反应体系；

优选地，在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶，使所述化合物或其盐掺入生长的核酸链，形成包含可逆阻断基团和任选的可检测标记的核酸中间体；

优选地，所述聚合酶选自 KOD 聚合酶或其突变体（例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391）；

优选地，在任意一个检测所述核酸中间体的步骤前，移除前一步骤的反应体系的溶液相，保留连接于支持物上的双链体；

优选地，所述切除试剂与所述双链体或所述生长的核酸链在含有溶液相和固相的反应体系中接触；

优选地，所述切除试剂能够切除掺入生长的核酸链的化合物中的可逆阻断基团和任选的其携带的可检测标记，并且不会影响双链体骨架上的磷酸二酯键；

优选地，在任意一个切除所述核酸中间体所包含的可逆阻断基团和任选的可检测标记的步骤后，移除这一步骤反应体系的溶液相；

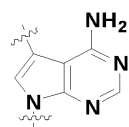
优选地，在任意一个包含移除操作的步骤之后，进行洗涤操作；

优选地，步骤 (ii) 之后，进一步包括：根据步骤 (ii) 检测得到的信号，确定步骤 (i) 中掺入生长的核酸链的化合物的类型，并基于碱基互补配对原则，确定待测序的核酸链中相应位置处的核苷酸类型。

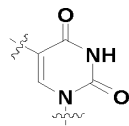
36、试剂盒，其包含至少一个权利要求 23-27 任一项所述的化合物或其盐；

优选地，所述试剂盒包含第一、第二、第三和第四化合物，所述第一、第二、第三和第四化合物各自独立地为权利要求 23-27 任一项所述的化合物或其盐；

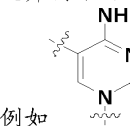
优选地，所述第一化合物中，base¹选自腺嘌呤、7-脱氮腺嘌呤或其互变异构体（例如



); 所述第二化合物中, base¹ 选自胸腺嘧啶、尿嘧啶或其互变异构体 (例如



); 所述第三化合物中, base¹ 选自胞嘧啶或其互变异构体 (例如



所述第四化合物中, base¹ 选自鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体 (例如

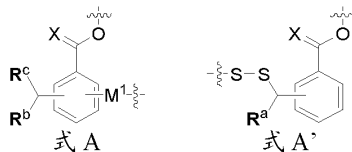
优选地, 所述第一、第二、第三和第四化合物包含的 base¹ 互不相同;

优选地, 所述第一、第二、第三和第四化合物携带的额外的可检测标记互不相同。

37、权利要求 36 的试剂盒, 其中, 所述试剂盒还包含: 用于预处理核酸分子的试剂; 用于连接待测序的核酸分子的支持物; 用于将待测序的核酸分子与支持物连接 (例如, 共价或非共价连接) 的试剂; 用于起始核苷酸聚合反应的引物; 用于进行核苷酸聚合反应的聚合酶; 一种或多种缓冲溶液; 一种或多种洗涤溶液; 或其任何组合。

38、权利要求 23-27 任一项所述的化合物或其盐或者权利要求 36-37 任一项所述的试剂盒用于测定目标单链多核苷酸的序列的用途。

39、一种核苷酸类似物, 其由核糖或脱氧核糖、可逆阻断基团、碱基或脱氮碱基或其互变异构体、用于连接可检测标记的连接子和任选的磷酸基团形成, 其中, 所述连接子包含如下式 A 或式 A' 所示的结构:



其中:

R^b、R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu)、-ONH₂、-OCOR_x、-OCONHR_x、-S-SO₂R_x, 另一个选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基);

优选地, R^b、R^c 中, 任意一个选自 -N₃、-SS-C1-C6 烷基 (如 -SS-Me、-SS-Et、-SS-iPr、-SS-t-Bu), 另一个选自 C1-C6 烷基;

更优选地, R^b、R^c 中, 任意一个选自 -N₃、-SS-Me、-SS-Et, 另一个为甲基;

最优选地, R^b、R^c 中, 任意一个选自 -N₃、-SS-Me, 另一个为甲基;

R^a 选自 H、脂肪族烷基 (如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基 (如苯基-C1-C6 烷基)、环烷基 (如 C3-C6 环烷基);

优选地, R^a 选自 H、C1-C6 烷基;

更优选地, R^a 为甲基;

X 选自 O、S、NH;

优选地, X 选自 O、S;

更优选地, X 为 O;

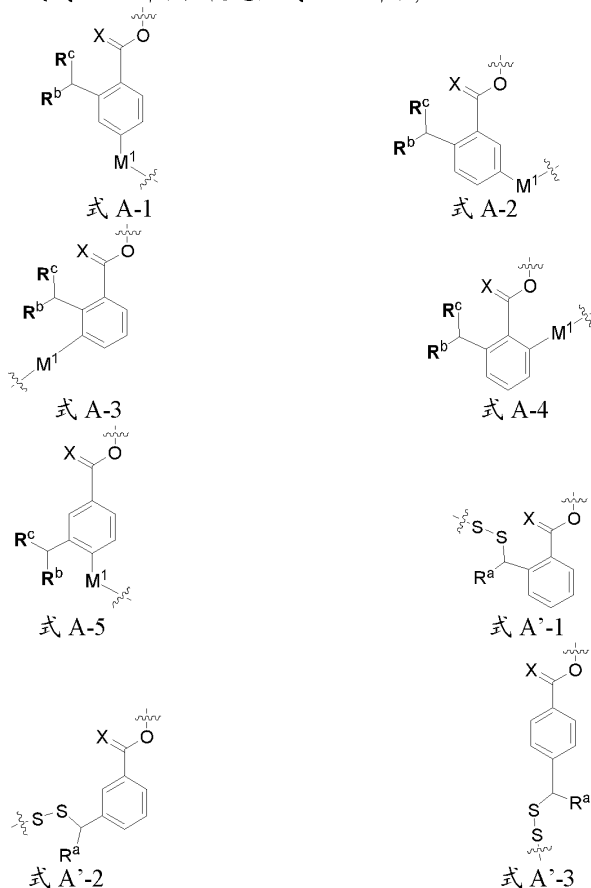
M¹ 选自直接键、NH、O、S、CH₂;

优选地, M¹ 选自直接键、NH、O;

优选地, 所述可逆阻断基团与所述核糖或脱氧核糖的 3'-OH 相连, 所述碱基或脱氮碱

基或其互变异构体与所述核糖或脱氧核糖的 1'-C 相连, 所述任选的磷酸基团与所述核糖或脱氧核糖的 5'-OH 相连, 所述连接子与所述碱基或脱氮碱基或其互变异构体相连。

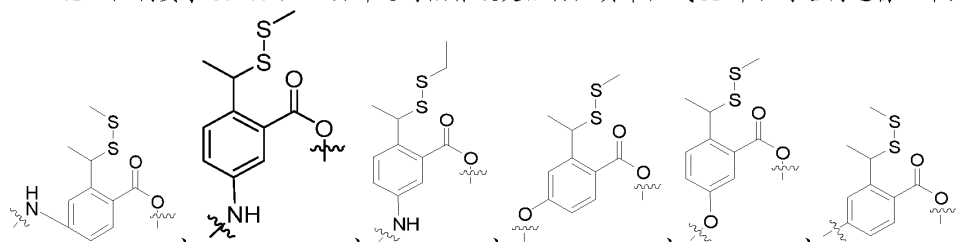
40、权利要求 39 所述的核苷酸类似物, 其中, 式 A 所示的结构如式 A-1、式 A-2、式 A-3、式 A-4 或式 A-5 所示, 优选如式 A-1、式 A-2 或式 A-5 所示; 或者, 式 A' 所示的结构如式 A'-1、式 A'-2 或式 A'-3 所示, 优选如式 A'-1 所示;

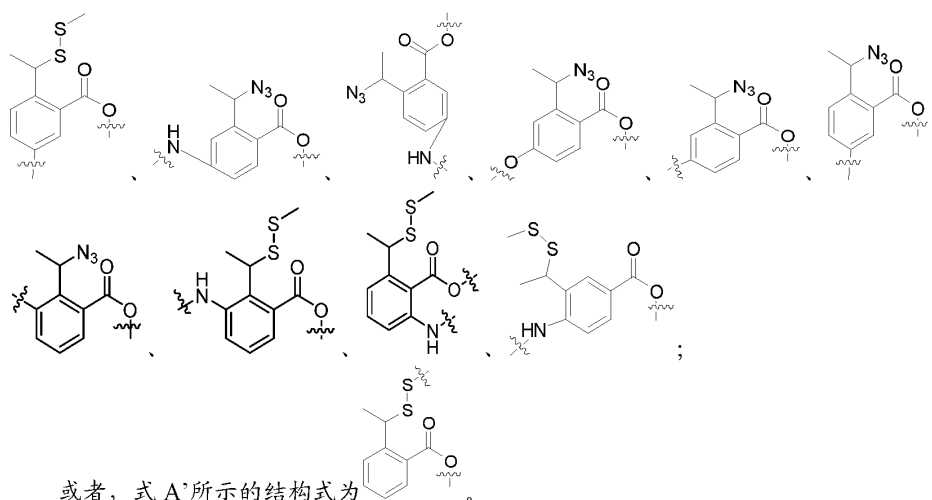


其中:

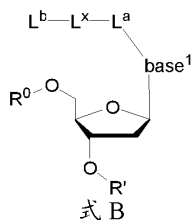
R^a 、 R^b 、 R^c 、X、 M^1 的定义如前所述。

41、权利要求 39-40 任一项所述的核苷酸类似物, 其中, 式 A 所示的结构选自以下:





42、权利要求 39-41 任一项所述的核苷酸类似物，其中，所述核苷酸类似物具有如下式 B 所示的结构：



L^x 具有权利要求 39-41 任一项所述的式 A、式 A'、式 A-1、式 A-2、式 A-3、式 A-4、式 A-5、式 A'-1、式 A'-2 或式 A'-3 所示的结构式，且 L^x 的 M¹ 端或 S 端与 L^b 相连， L^x 的 O 端与 L^a 相连；

$L^b-L^x-L^a$ 表示用于连接可检测标记的连接子；

base¹ 表示碱基或脱氮碱基或其互变异构体；

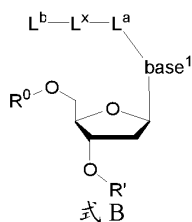
R' 表示可逆阻断基团；

R^0 表示 H 或磷酸基团；

L^a 表示所述连接子中用于连接碱基或脱氮碱基或其互变异构体的部分；

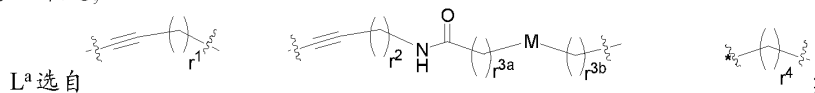
L^b 表示所述连接子中用于连接可检测标记的部分。

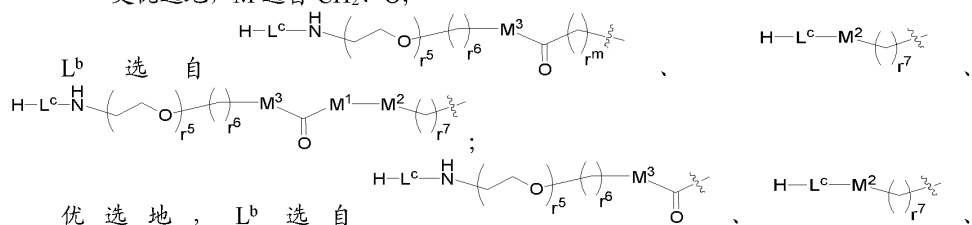
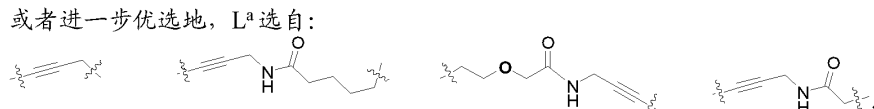
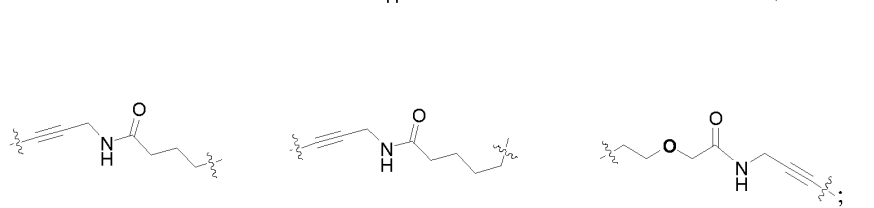
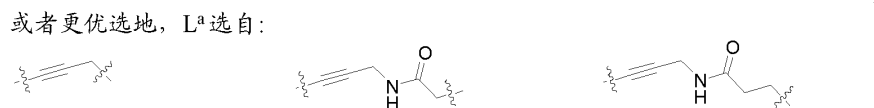
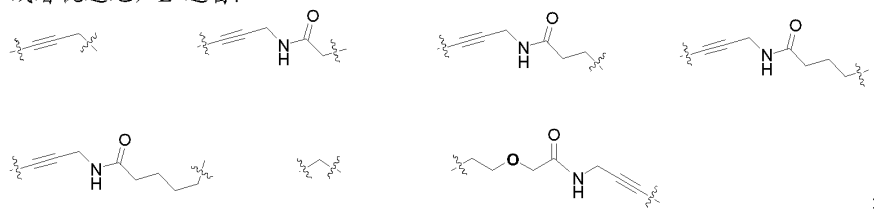
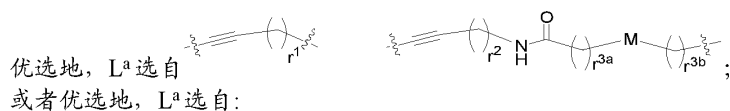
43、权利要求 39-42 任一项所述的核苷酸类似物，其中，所述核苷酸类似物具有如下式 B 所示的结构：

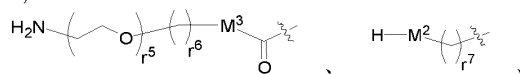
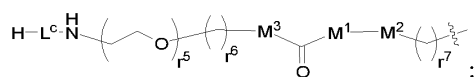


其中：

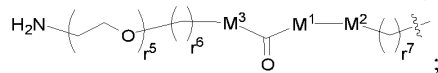
L^x 具有如权利要求 39-41 任一项的式 A、式 A'、式 A-1、式 A-2、式 A-3、式 A-4、式 A-5、式 A'-1、式 A'-2 或式 A'-3 所示的结构式，且 L^x 的 M¹ 端或 S 端与 L^b 相连， L^x 的 O 端与 L^a 相连；



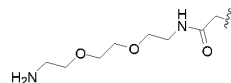
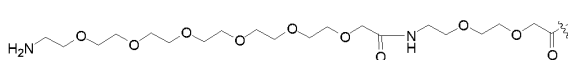
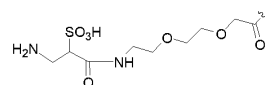
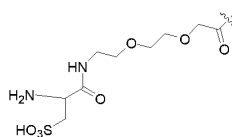
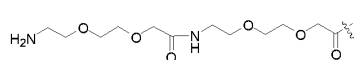
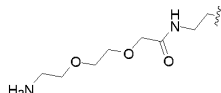
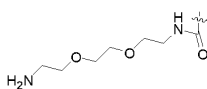
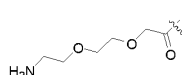




更优选地， L^b 选自

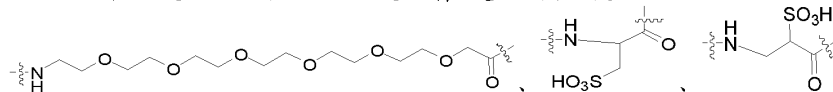


最优选地, L^b 选自:



L^c 选自直接键、 NH 、 $\text{C}(\text{R}^{10})\text{C}(\text{R}^{11})\text{C}(=\text{O})$ 、 HO_3S 、 $\text{C}(\text{R}^{11})\text{C}(=\text{O})$ ，优选地， L^c 中的 NH 端与 L^b 中的 H 相连接， L^c 中的 $=\text{O}$ 端与 L^b 中的 NH 相连接；

优选地， L^c 选自直接键、



中的 NH 端与 L^b 中的 H 相连接, L^c 中的 =O 端与 L^b 中的 NH 相连接;

r^m 选自 0-6 之间的任意整数;

优选地, r^m 选自 0、1、2、3;

更优选地, r^m 选自 0、1;

r^5, r^6, r^7 各自独立地选自 1-6 之间的任意整数:

优选地, r^5 选自 1、2、3;

更优选地, r^5 为 2;

优选地, r^6 选自 1、2、3;

更优选地, r^6 为 1 或 2;

优选地, r^7 选自 1、2、3;

更优选地, r^7 为 2;

r^{10} 选自 1-10 之间的任意整数;

优选地, r^{10} 选自 2-6 之间的任意整数;

更优选地, r^{10} 为 2 或 6;

r^{11} 选自 1-6 之间的任意整数;

优选地, r^{11} 选自 1、2、3;

更优选地, r^{11} 为 1;

M^1 、 M^2 、 M^3 各自独立地选自直接键、NH、O、S、 CH_2 ;

优选地, M^1 选自直接键、NH、O;

优选地, M^2 为 NH:

优选地, M^3 选自直接键、NH;

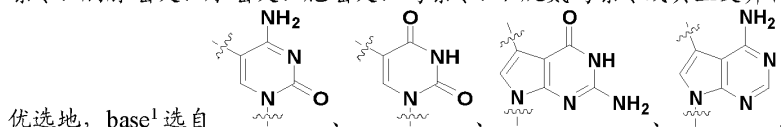
base¹表示碱基或脱氮碱基;

R'表示可逆阻断基团;

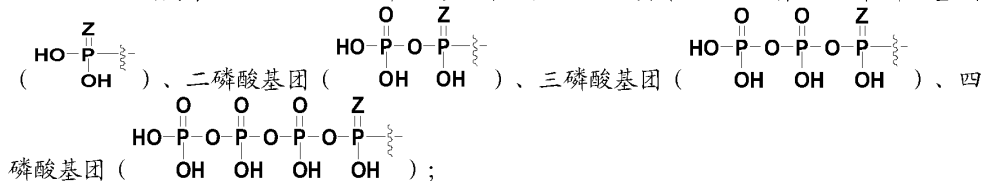
L^b-L^x-L^a表示用于连接可检测标记的连接子;

R⁰表示 H 或磷酸基团。

44、权利要求 42-43 任一项所述的核苷酸类似物, 其中, 所述 base¹选自腺嘌呤、7-脱氮腺嘌呤、胸腺嘧啶、尿嘧啶、胞嘧啶、鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体;



45、权利要求 42-44 任一项所述的核苷酸类似物, 其中, R⁰选自 H、单磷酸基团

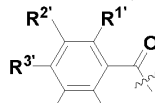


优选地, R⁰为三磷酸基团 ();

各 Z 独立地选自 O, S, BH;

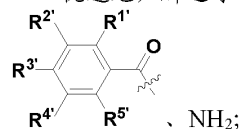
优选地, Z 为 O。

46、权利要求 42-45 任一项所述的核苷酸类似物, 其中, 所述可逆阻断基团 R'选自

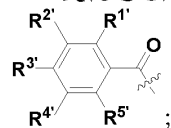


N₃-C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、NH₂、-ONH₂、-OCOR_Z、-OCONHR_Z, 其中, 各 R_Z独立地选自脂肪族烷基 (如 C1-C6 烷基), 环烷基 (如 C3-C6 环烷基) 或芳香族烷基 (如苯基 C1-C6 烷基);

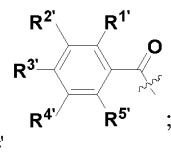
优选地, 所述可逆阻断基团 R'选自 N₃-C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、



更优选地, 所述可逆阻断基团 R'选自 N₃-C1-C6 烷基、C1-C6 烷基-SS-C1-C6 烷基、



最优选地, 所述可逆阻断基团 R'选自 N₃-CH₂-、CH₃-CH₂-S-S-CH₂-、



R^{1'}、R^{2'}、R^{3'}、R^{4'}、R^{5'}中, 任意一个 (如 R^{1'}或 R^{5'}) 为 , 其余各自独立地选自 H、叠氮基、硝基、氨基、磺基、羧基、脂肪族烷基 (如 C1-C6 烷基)、环烷基 (如 C3-C6 环烷基)、芳香族烷基 (如苯基-C1-C6 烷基)、F、I、Br、Cl、烷氧基 (如 C1-C6 烷氧基),

优选地, $R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中, 任意一个(如 $R^{1'}$ 或 $R^{5'}$)为 , 另一个(如 $R^{3'}$)选自 H、C1-C6 烷基(如甲氧基), 其余为 H,

更优选地, $R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中, 任意一个(如 $R^{1'}$ 或 $R^{5'}$)为 , 其余为 H,

最优选地, $R^{1'}$ 、 $R^{2'}$ 、 $R^{3'}$ 、 $R^{4'}$ 、 $R^{5'}$ 中, $R^{1'}$ 为 , 其余为 H,

$R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS$ -C1-C6 烷基(如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$)、 $-ONH_2$ 、 $-OCOR_x$ 、 $-OCONHR_x$, 另一个选自 H、脂肪族烷基(如 C1-C6 烷基, 例如 Me、Et、iPr、tBu)、芳香族烷基(如苯基-C1-C6 烷基)、环烷基(如 C3-C6 环烷基), 其中, 各 R_x 独立地选自脂肪族烷基(如 C1-C6 烷基), 环烷基(如 C3-C6 环烷基)或芳香族烷基(如苯基 C1-C6 烷基),

优选地, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS$ -C1-C6 烷基(如 $-SS-Me$ 、 $-SS-Et$ 、 $-SS-iPr$ 、 $-SS-t-Bu$), 另一个选自 C1-C6 烷基,

更优选地, $R^{b'}$ 、 $R^{c'}$ 中, 任意一个选自 $-N_3$ 、 $-SS-Me$, 另一个为甲基。

47、权利要求 39-46 任一项所述的核苷酸类似物, 其选自: 权利要求 9 所述的化合物。

48、权利要求 39-47 任一项所述的核苷酸类似物, 其中, 所述核苷酸类似物携带额外的可检测标记, 且所述可检测标记与连接子(优选 L^b)连接;

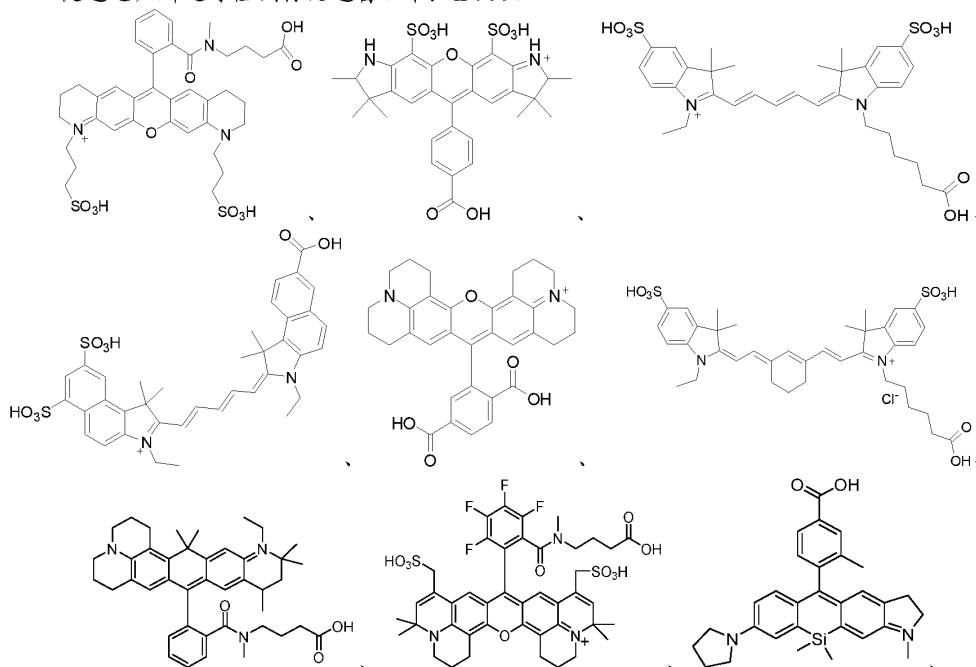
优选地, 所述核苷酸类似物携带的额外的可检测标记是通过亲和试剂(如抗体、适体、亲和物、结蛋白)引入的, 所述亲和试剂携带所述可检测标记, 且所述亲和试剂可以特异性识别并结合所述核苷酸类似物的表位; 优选地, 所述可检测标记与所述 L^b 的末端氨基连接;

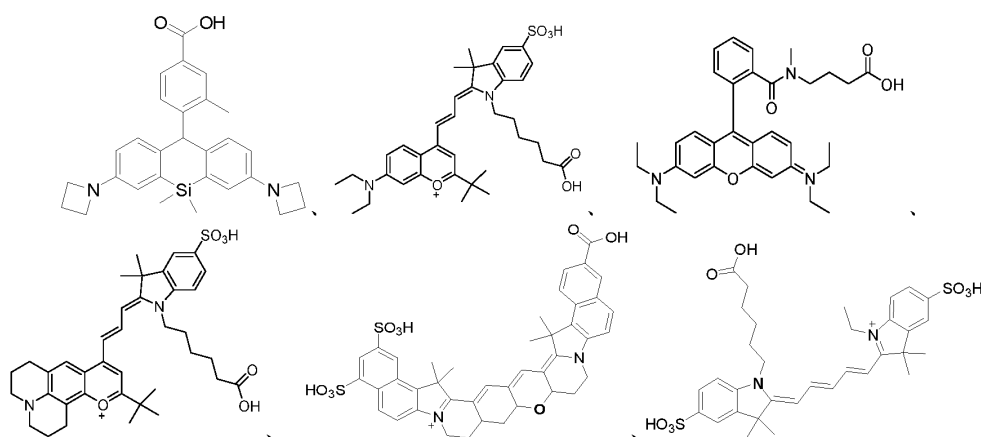
优选地, 所述可检测标记中的羧基与所述 L^b 的末端氨基通过形成酰胺键进行连接;

优选地, $base^1$ 不同, 所述核苷酸类似物携带的额外的可检测标记不同;

优选地, 所述可检测标记为荧光标记;

优选地, 所述可检测标记选自以下: iF700、





49、权利要求 48 所述的核苷酸类似物，其中，所述核苷酸类似物选自：权利要求 11 所述的化合物。

50、终止核酸合成的方法，其包括：将权利要求 39-49 任一项所述的核苷酸类似物掺入待终止的核酸分子中；

优选地，所述核苷酸类似物的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述核苷酸类似物掺入待终止的核酸分子中；

优选地，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述核苷酸类似物掺入待终止的核酸分子的 3'端。

51、制备在测序反应中与目标单链多核苷酸互补的生长的多核苷酸的方法，其包括将权利要求 39-49 任一项所述的核苷酸类似物掺入所述生长的互补多核苷酸，其中，所述核苷酸类似物的掺入防止了任何后续的核苷酸引入所述生长的互补多核苷酸中；

优选地，所述核苷酸类似物的掺入通过末端转移酶、末端聚合酶或逆转录酶来实现；

优选地，所述方法包括：使用聚合酶，将所述核苷酸类似物掺入所述生长的互补多核苷酸；

优选地，所述方法包括：在允许聚合酶进行核苷酸聚合反应的条件下，使用聚合酶进行核苷酸聚合反应，从而将所述核苷酸类似物掺入所述生长的互补多核苷酸的 3'端。

52、核酸中间体，其是在测定目标单链多核苷酸的序列中形成的，其中，

所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是权利要求 39-49 任一项所述的核苷酸类似物；

或者，所述核酸中间体是通过以下步骤形成的：

向生长的核酸链中掺入一个与目标单链多核苷酸互补的核苷酸，形成所述核酸中间体，其中，掺入的一个互补核苷酸是权利要求 39-49 任一项所述的核苷酸类似物，且所述生长的核酸链中预先掺入至少一个与目标单链多核苷酸互补的核苷酸，预先掺入的至少一个与目标单链多核苷酸互补的核苷酸是已被除去可逆阻断基团和任选的可检测标记的权利要求 39-49 任一项所述的核苷酸类似物。

53、测定目标单链多核苷酸的序列的方法，其包括：

1) 监测生长的核酸链中与目标单链多核苷酸互补的核苷酸的掺入，其中，掺入的至少一个互补核苷酸是权利要求 39-49 任一项所述的核苷酸类似物，以及，

2) 确定掺入的核苷酸的类型；

优选地，在引入下一个互补核苷酸之前，将所述可逆阻断基团和任选的可检测标记除去；

优选地，所述可逆阻断基团和所述可检测标记被同时除去；

优选地，所述可逆阻断基团和所述可检测标记被先后除去；例如，在所述可检测标记被

除去之后,所述可逆阻断基团被除去,或者,在所述可逆阻断基团被除去之后,所述可检测标记被除去。

54、权利要求 53 所述的方法,其包括以下步骤:

- (a) 提供多种不同的核苷酸,其中至少一种核苷酸是权利要求 48-49 任一项所述的核苷酸类似物,任选地其余的核苷酸是权利要求 39-49 任一项所述的核苷酸类似物;
 - (b) 将所述多种不同的核苷酸掺入目标单链多核苷酸的互补序列中,其中,所述多种不同的核苷酸在检测时可以相互区分;
 - (c) 检测 (b) 的核苷酸,从而确定掺入的核苷酸的类型;
 - (d) 除去 (b) 的核苷酸中的可逆阻断基团和任选的其携带的可检测标记;和
 - (e) 任选地重复步骤 (a) - (d) 一次或多次;
- 从而确定所述目标单链多核苷酸的序列。

55、权利要求 53 所述的方法,其包括以下步骤:

- (1) 提供第一核苷酸、第二核苷酸、第三核苷酸和第四核苷酸,四种核苷酸中的至少一种是权利要求 48-49 任一项所述的核苷酸类似物,任选地其余的核苷酸是权利要求 39-49 任一项所述的核苷酸类似物;
 - (2) 将所述四种核苷酸与目标单链多核苷酸进行接触;除去未掺入生长的核酸链中的所述核苷酸;检测掺入生长的核酸链中的所述核苷酸;除去掺入生长的核酸链中的所述核苷酸中的所述可逆阻断基团和任选的其携带的所述可检测标记;
- 任选地,还包括 (3): 重复 (1) - (2) 一次或多次。

56、权利要求 53 所述的方法,其包括以下步骤:

- (a) 提供包含双链体、包含至少一种权利要求 48-49 任一项所述的核苷酸类似物的核苷酸、聚合酶和切除试剂的混合物;所述双链体包含生长的核酸链以及待测序的核酸链;
 - (b) 进行包含以下步骤 (i)、(ii) 和 (iii) 的反应,任选地,重复一次或多次:
 - 步骤 (i): 使用聚合酶,使所述核苷酸类似物掺入生长的核酸链,形成包含可逆阻断基团和任选的可检测标记的核酸中间体;
 - 步骤 (ii): 对所述核酸中间体进行检测;
 - 步骤 (iii): 使用切除试剂将所述核酸中间体所包含的可逆阻断基团和任选的可检测标记切除;
- 优选地,对所述可逆阻断基团的切除和对所述可检测标记的切除同时进行,或者,对所述可逆阻断基团的切除和对所述可检测标记的切除分步进行(例如,先切除所述可逆阻断基团,或者先切除所述可检测标记);
- 优选地,对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是同样的试剂;
- 优选地,对所述可逆阻断基团的切除和对所述可检测标记的切除使用的切除试剂是不同的试剂。

57、权利要求 56 所述的方法,其中,所述双链体连接于支持物上;

优选地,所述生长的核酸链为引物;

优选地,所述引物通过退火至待测序的核酸链上,形成所述双链体;

优选地,所述双链体、所述核苷酸类似物、以及所述聚合酶一起形成含有溶液相和固相的反应体系;

优选地,在允许聚合酶进行核苷酸聚合反应的条件下,使用聚合酶,使所述核苷酸类似物掺入生长的核酸链,形成包含可逆阻断基团和任选的可检测标记的核酸中间体;

优选地,所述聚合酶选自 KOD 聚合酶或其突变体(例如 KOD POL151、KOD POL157、KOD POL171、KOD POL174、KOD POL376、KOD POL391);

优选地,在任意一个检测所述核酸中间体的步骤前,移除前一步骤的反应体系的溶液相,保留连接于支持物上的双链体;

优选地,所述切除试剂与所述双链体或所述生长的核酸链在含有溶液相和固相的反应

体系中接触;

优选地,所述切除试剂能够切除掺入生长的核酸链的核苷酸类似物中的可逆阻断基团和任选的其携带的可检测标记,并且不会影响双链体骨架上的磷酸二酯键;

优选地,在任意一个切除所述核酸中间体所包含的可逆阻断基团和任选的可检测标记的步骤后,移除这一步骤反应体系的溶液相;

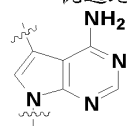
优选地,在任意一个包含移除操作的步骤之后,进行洗涤操作;

优选地,步骤(ii)之后,进一步包括:根据步骤(ii)检测得到的信号,确定步骤(i)中掺入生长的核酸链的核苷酸类似物的类型,并基于碱基互补配对原则,确定待测序的核酸链中相应位置处的核苷酸类型。

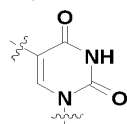
58、试剂盒,其包含至少一个权利要求39-49任一项所述的核苷酸类似物;

优选地,所述试剂盒包含第一、第二、第三和第四化合物,所述第一、第二、第三和第四化合物各自独立地为权利要求39-49任一项所述的核苷酸类似物;

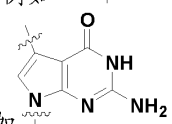
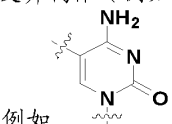
优选地,所述第一化合物中,base¹选自腺嘌呤、7-脱氮腺嘌呤或其互变异构体(例如



);所述第二化合物中,base¹选自胸腺嘧啶、尿嘧啶或其互变异构体(例如



);所述第三化合物中,base¹选自胞嘧啶或其互变异构体(例如



所述第四化合物中,base¹选自鸟嘌呤、7-脱氮鸟嘌呤或其互变异构体(例如

优选地,所述第一、第二、第三和第四化合物包含的base¹互不相同;

优选地,所述第一、第二、第三和第四化合物携带的额外的可检测标记互不相同。

59、权利要求58所述的试剂盒,其中,所述试剂盒还包含:用于预处理核酸分子的试剂;用于连接待测序的核酸分子的支持物;用于将待测序的核酸分子与支持物连接(例如,共价或非共价连接)的试剂;用于起始核苷酸聚合反应的引物;用于进行核苷酸聚合反应的聚合酶;一种或多种缓冲溶液;一种或多种洗涤溶液;或其任何组合。

60、权利要求39-49任一项所述的核苷酸类似物或者权利要求58-59任一项所述的试剂盒用于测定目标单链多核苷酸的序列的用途。