

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.



[12] 发明专利申请公布说明书

[21] 申请号 200780033902.3

C07D 239/46 (2006.01)

C07K 14/435 (2006.01)

C12Q 1/48 (2006.01)

[43] 公开日 2009 年 8 月 26 日

[11] 公开号 CN 101516857A

[22] 申请日 2007.7.24

[21] 申请号 200780033902.3

[30] 优先权

[32] 2006.7.25 [33] EP [31] 06117779.6

[86] 国际申请 PCT/EP2007/057597 2007.7.24

[87] 国际公布 WO2008/012296 英 2008.1.31

[85] 进入国家阶段日期 2009.3.12

[71] 申请人 洛桑联邦理工学院

地址 瑞士洛桑

[72] 发明人 A·戈蒂尔 K·约翰松

M·金德曼 A·朱耶拉特

F·博菲尔斯

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 赵蓉民 陈建芳

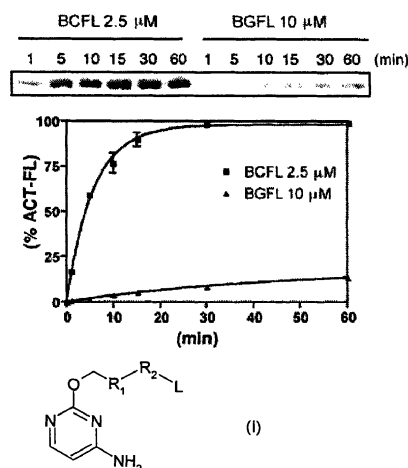
权利要求书 7 页 说明书 26 页 序列表 6 页
附图 3 页

[54] 发明名称

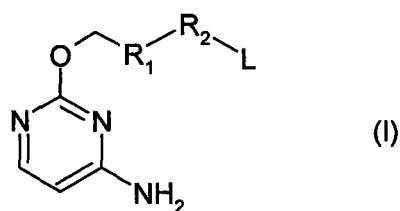
用合成探针标记融合蛋白

[57] 摘要

本申请的发明是用合成探针标记融合蛋白。本发明涉及被称为烷基胞嘧啶转移酶 (ACT) 的新蛋白质, 其衍生自 O^6 -烷基鸟嘌呤-DNA 烷基转移酶; ACT 的底物, 其用于特异性地将标记转移到这些 ACT; 以及包含这些的融合蛋白。根据本发明的底物是式 (I) 的取代胞嘧啶, 如右。其中, R_1 是芳族基或杂芳族基, 或者是双键连接于 OCH_2 - 的任选取代的不饱和烃基、环烃基或杂环基; R_2 是连接体; 并且 L 是一个标记或多个相同或不同的标记。本发明进一步涉及将标记 L 从这些式 (I) 底物转移到 ACT 和 ACT 融合蛋白的方法。ACT-式 (I) 化合物的系统与已知系统 O^6 -烷基鸟嘌呤-DNA 烷基转移酶 (AGT)-苄基鸟嘌呤一起特别适合于双重标记研究。



1. 式(I)化合物,



其中,

R_1 是芳族基或杂芳族基, 或者是双键连接于 OCH_2- 的任选取代的不饱和烃基、环烃基或杂环基;

R_2 是连接体; 并且

L 是一个标记或多个相同或不同的标记。

2. 根据权利要求 1 所述的式(I)化合物, 其中 R_1 是苯基。

3. 根据权利要求 2 所述的式(I)化合物, 其中 R_1 是对位取代的苯基。

4. 根据权利要求 1 所述的式(I)化合物, 其中 R_2 是具有 1 至 300 个碳原子的直链或支链亚烃基, 其中任选地

(a) 一个或多个碳原子被氧置换, 特别地, 其中每隔两个碳原子被氧置换, 例如具有 1 至 100 个乙烯氧基单元的聚乙烯氧基;

(b) 一个或多个碳原子被携带氢原子的氮置换, 并且相邻的碳原子被氧取代, 其代表酰胺官能 $-NH-CO-$;

(c) 一个或多个碳原子被氧置换并且相邻的碳原子被氧取代, 其代表酯官能 $-O-CO-$;

(d) 两个相邻碳原子之间的键是双键或三键, 其代表官能 $-CH=CH-$ 或 $-C\equiv C-$;

(e) 一个或多个碳原子被下列基团置换: 亚苯基、饱和或不饱和的亚环烃基、饱和或不饱和的亚双环烃基、桥连杂芳基或者桥连饱和或不饱和杂环基;

(f) 两个相邻碳原子被下列置换: 二硫键 $-S-S-$;

或者两个或多个、尤其两个或三个亚烃基和/或如上文(a)至(f)所定义的改性亚烃基的组合，其任选地含有取代基。

5. 根据权利要求 4 所述的式(I)化合物, 其中 R_2 是具有 1 至 25 个碳原子的直链亚烃基, 其中碳原子任选地被酰胺官能-NH-CO-置换。

6. 根据权利要求 1 所述的式(I)化合物, 其中 L 是光谱探针、代表特异性结合对的一部分的分子或者共价连接于固相支持体的分子。

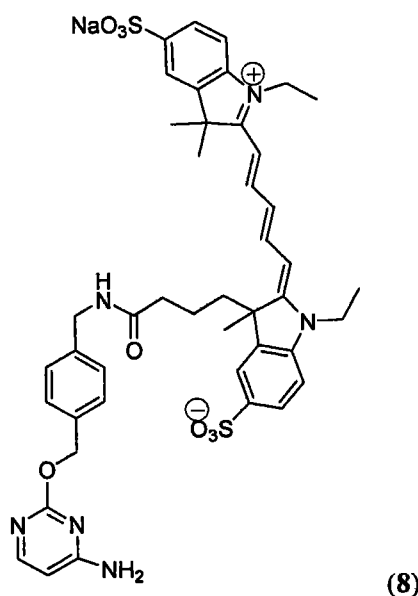
7. 根据权利要求 6 所述的式(I)化合物, 其中 L 是荧光团。

8. 根据权利要求 7 所述的化合物 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-四甲基若丹明-5-或 6-甲酰胺, 或其混合物。

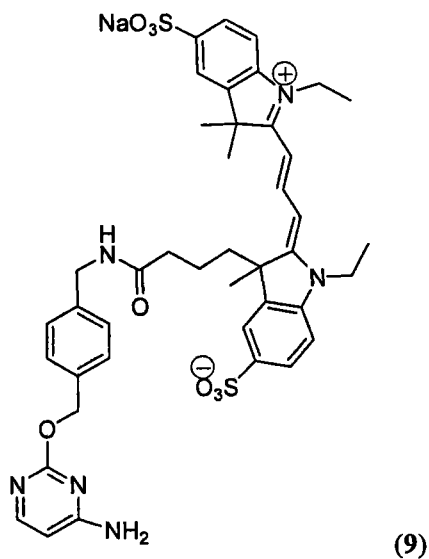
9. 根据权利要求 7 所述的化合物 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-荧光素-5-或 6-甲酰胺，或其混合物。

10. 根据权利要求 7 所述的化合物 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-二乙酰基荧光素-5-或 6-甲酰胺，或其混合物。

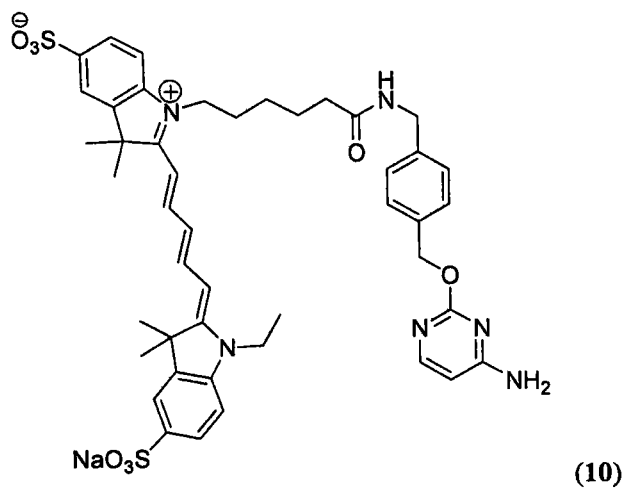
11. 根据权利要求 7 所述的化合物, 其为式 8 的 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-DY647-甲酰胺。



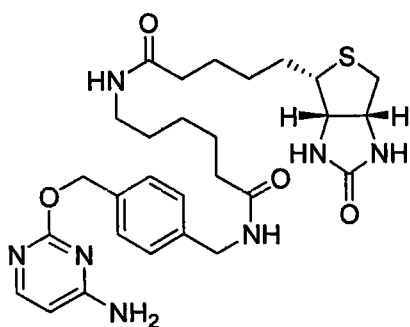
12. 根据权利要求 7 所述的化合物, 其为式 9 的 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-DY547-甲酰胺。



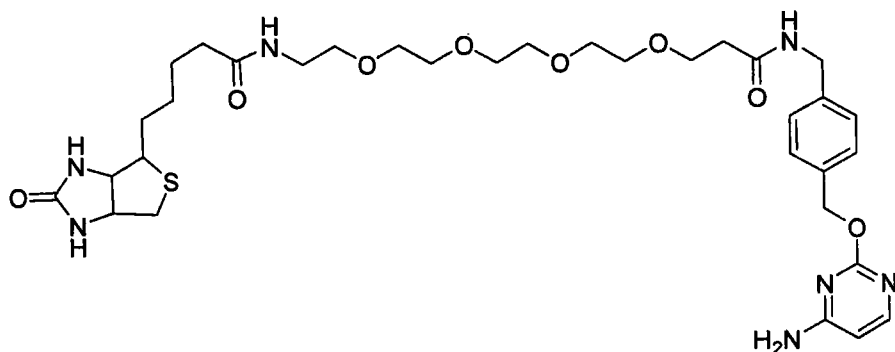
13. 根据权利要求 7 所述的化合物, 其为式 10 的 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-Cy5-甲酰胺。



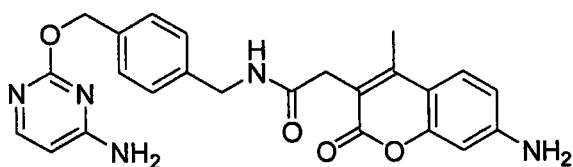
14. 式 12 的根据权利要求 7 所述的化合物 BC-生物素。



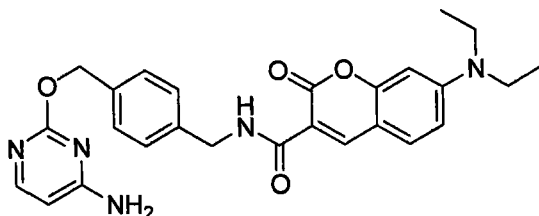
15. 式 13 的根据权利要求 7 所述的化合物 BC-PEG-生物素。



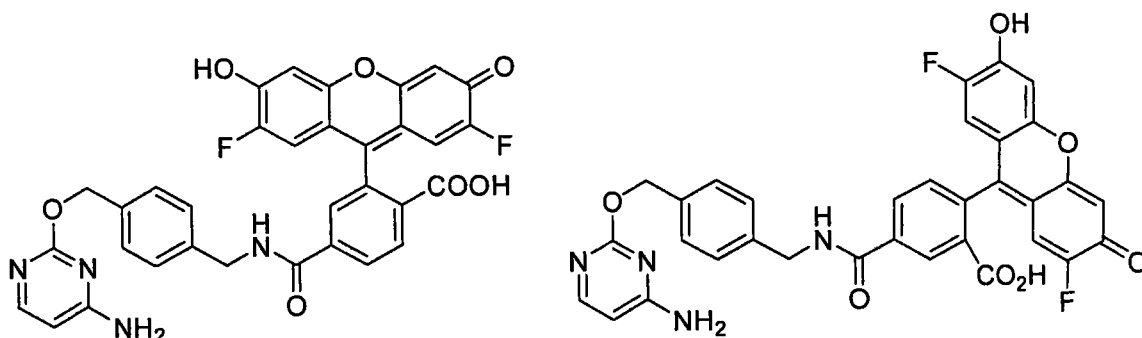
16. 式 14 的根据权利要求 7 所述的化合物 BC-360。



17. 式 15 的根据权利要求 7 所述的化合物 BC-430。

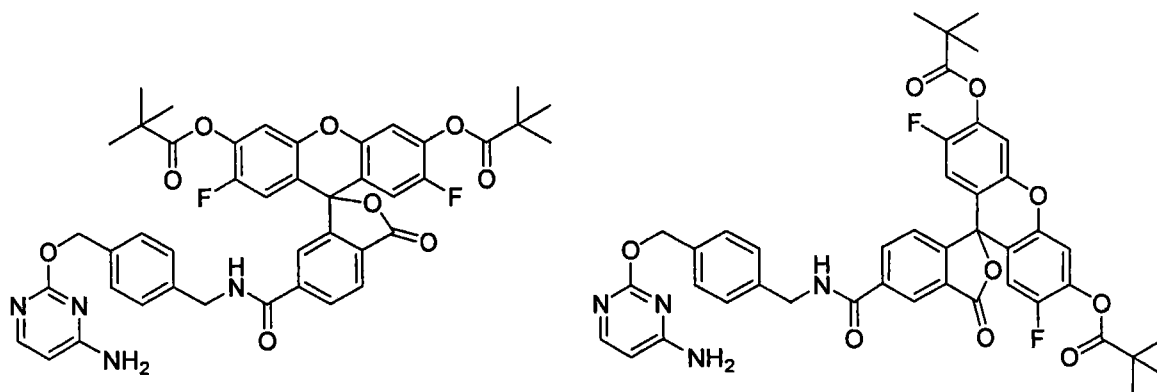


18. 式 16 之一的根据权利要求 7 所述的化合物 BC-俄勒岗绿或其混合物。

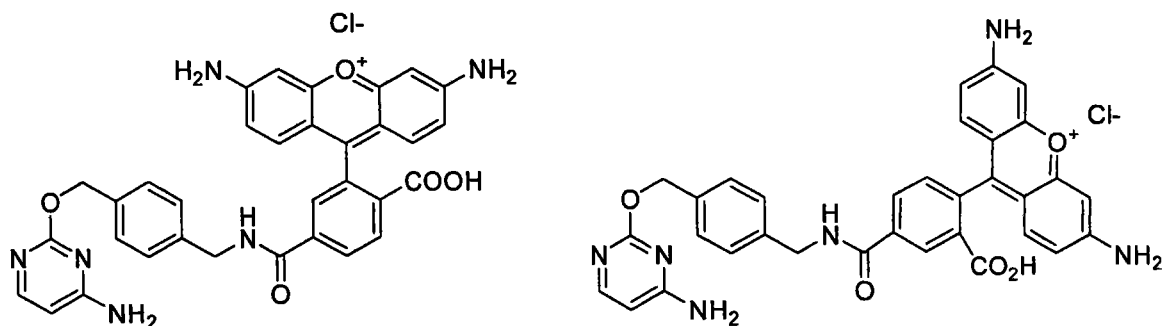


19. 式 17 之一的根据权利要求 7 所述的化合物 BC-俄勒岗绿二新

戊酰基酯或其混合物。



20. 式 18 之一的根据权利要求 7 所述的化合物 BC-505 或其混合物。



21. 烷基胞嘧啶转移酶(ACT), 其是这样的蛋白质:

(a) 由 170 至 220 个氨基酸组成;

(b) 包含至少一个半胱氨酸;

(c) 与 O²-苄基胞嘧啶进行反应, 由此至少与在相同条件下与 O⁶-苄基鸟嘌呤的反应一样快地将苄基取代基转移至 (b) 的半胱氨酸的巯基官能上。

22. 权利要求 21 所述的烷基胞嘧啶转移酶(ACT), 其包含 177 至 185 个氨基酸。

23. 权利要求 21 所述的烷基胞嘧啶转移酶(ACT), 其选自:

根据 SEQ ID NO:1 的蛋白质;

由于下列取代而与 SEQ ID NO:1 不同的蛋白质:

R114A、S131V、E148Q、G157W 和 M159R;

R114S、S131T、D135T、E148D、G157P 和 M159E;

R114N、S131N、G157A 和 M159S;

R114A、S131T、D135S、G157K 和 M159E;

R114E、S131R、D135A、G157E 和 M159E;

R114S、S131V、E148Q、G157L 和 M159R;

R114E、S131N、D135N、G157T 和 M159F;

以及这样的蛋白质,其在除了位置 114、131、135、148、157 和 159 之外的位置上的一个、两个或三个氨基酸上与之不同。

24. 序列 SEQ ID NO:1 的根据权利要求 21 所述的蛋白质。

25. 权利要求 21 所述的烷基胞嘧啶转移酶(ACT),其选自:
根据 SEQ ID NO:12 的蛋白质;

SEQ ID NO:12 的蛋白质,其中

位置 60 上的氨基酸为 M 或 I;

位置 114 上的氨基酸为 A、E、N、R 或 S;

位置 121 上的氨基酸为 A 或 V;

位置 131 上的氨基酸为 N、S、T 或 V;

位置 135 上的氨基酸为 D、N 或 T;

位置 148 上的氨基酸为 D、E、Q 或 V;

位置 153 上的氨基酸为 L 或 S;

位置 157 上的氨基酸为 A、G、L、T、P 或 W; 和

位置 159 上的氨基酸为 E、F、M、R、S 或 L;

以及这样的蛋白质,其在除了位置 60、114、121、131、135、148、153、157 和 159 之外的位置上的一个、两个或三个氨基酸上与之不同。

26. 权利要求 25 所述的烷基胞嘧啶转移酶(ACT),其选自根据 SEQ ID NO:12 的蛋白质以及这样的蛋白质,其在除了位置 60、114、121、131、135、148、153、157 和 159 之外的位置上的一个氨基酸上与之不同。

27. 序列 SEQ ID NO:12 的根据权利要求 25 的蛋白质。

28. 产生根据权利要求 21 所述的烷基胞嘧啶转移酶(ACT)的方法，其特征在于编码 O⁶-烷基鸟嘌呤-DNA 烷基转移酶、O⁶-烷基鸟嘌呤-DNA 烷基转移酶突变体或烷基胞嘧啶转移酶的 DNA 在可达 10 个的氨基酸位置上通过饱和诱变被随机化，所获得的文库被转化到合适的噬菌粒中，期望的噬菌体通过与携带标记的苄基胞嘧啶反应进行选择，然后，所述标记转移到其上的噬菌体采用覆盖有针对所述标记的抗体的磁珠进行分离。

29. 将标记 L 从根据权利要求 1 所述的式(I)化合物转移到根据权利要求 21 所述的烷基胞嘧啶转移酶或者包含根据权利要求 21 所述的烷基胞嘧啶转移酶的融合蛋白的方法。

30. 检测和/或操纵目标蛋白的方法，其中所述目标蛋白被掺入包含根据权利要求 21 所述的烷基胞嘧啶转移酶的融合蛋白中，该烷基胞嘧啶转移酶融合蛋白与根据权利要求 1 所述的式(I)化合物接触，并且所述烷基胞嘧啶转移酶融合蛋白在设计来识别和/或处理所述标记的系统中利用所述标记 L 进行检测并且任选地被进一步操纵。

用合成探针标记融合蛋白

技术领域

[1] 本发明涉及将标记从底物转移到融合蛋白的方法，所述融合蛋白具有特定设计来接受所述标记的蛋白质部分；新的特异性底物；以及新的蛋白质，其接受适合于这种方法的特异性底物的标记。

背景技术

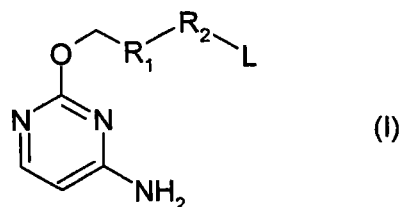
[2] 对于改良的标记技术存在持续需求，所述标记技术允许特异性地标记目标蛋白质以便在体外或体内条件下分离和/或跟踪这些目标蛋白质。一个具体的方法公开在 WO 02/083937 中，其描述了一种检测和/或操纵目标蛋白的方法，其中该蛋白质被融合于 O⁶-烷基鸟嘌呤-DNA 烷基转移酶 (AGT) 并且该 AGT 融合蛋白与携带标记的特异性 AGT 底物接触，由此该标记被转移到该融合蛋白。该 AGT 融合蛋白然后利用该标记进行检测并且任选地进一步进行操纵。野生型 AGT 的几种突变体显示出比野生型 AGT (WO 2004/031404; Juillerat, A.等, Chem. Biol. 10:313-317, 2003; Gronemeyer, T.等, Protein Eng. Des. Sel. 19:309-316, 2006) 更适合于这种标记方法，并且宽范围的取代苄基鸟嘌呤和相关的杂芳基甲基鸟嘌呤化合物被描述用于将标记转移到包含 AGT 和 AGT 突变体的融合蛋白(WO 2004/031405)。

[3] 简单的 O²-苄基-胞嘧啶是已知的。Freccero, M.等, J. Am. Chem. Soc. 125:3544-3553, 2003 通过胞嘧啶与邻醌甲基化物 (o-quinone methide) 的反应获得 O²-邻羟基苄基胞嘧啶。Ward, A.D.和 Baker, B. R., J. Med. Chem. 20:88-92, 1977 描述了从 2-氯-4-氨基嘧啶与苄醇的钠盐获得的 O²-苄基胞嘧啶。

发明内容

[4] 本发明涉及被称为烷基胞嘧啶转移酶(ACT)的新蛋白质，其衍

生自O⁶-烷基鸟嘌呤-DNA烷基转移酶；以及涉及ACT的底物，其特异性地将标记转移到这些ACT以及转移到包含这类ACT的融合蛋白。根据本发明的底物是式(I)的取代胞嘧啶，



其中，

R₁ 是芳族基或杂芳族基，或者是双键连接于 OCH₂-的任选取代的不饱和烃基、环烷基或杂环基；

R₂ 是连接体；并且

L 是一个标记或多个相同或不同的标记。

[5] 本发明进一步涉及将标记从这些式(I)底物转移到烷基胞嘧啶转移酶(ACT)和 ACT 融合蛋白的方法。

附图说明

[6] 图 1: 包含 AGT 和 ACT 的融合蛋白的体外双重标记实验，其利用携带荧光标记的特异性底物苄基鸟嘌呤 (BG) 和苄基胞嘧啶 (BC) 进行。

[7] 实验如下进行：用 BGFL、BGCy5、BCFL、BCCy5 或者 BGCy5/BCFL 或 BGFL/BCCy5 的等摩尔混合物温育 GST-ACT1 和 6xHis-^NAGT 的等摩尔混合物。荧光染料标记的蛋白质混合物通过 SDS-PAGE 进行分析，参见实施例 12。GST-ACT1: ACT 1 (SEQ ID NO:1) 与谷胱甘肽 S-转移酶 (GST) 的融合蛋白。6xHis-AGT: ^NAGT (SEQ ID NO:2) 与短肽 6xHis 的融合蛋白。BGFL 和 BGCy5: 分别用携带荧光素的连接体以及 Cy5 取代的苄基鸟嘌呤，参见 Juillerat, A.等, Chem. Biol. 10:313-317, 2003, 和化合物 11 (实施例 9)。BCFL: 化合物 5 (实施例 4)。BCCy5: 化合物 10 (实施例 8)。

[8] 图 2: 利用携带荧光标记 (FL) 的底物苄基鸟嘌呤 (BG) 和苄基胞嘧啶 (BC) 的 ACT10 标记实验，证明对 BC 的选择性。

[9] 温育不同时间后 SDS 凝胶实验的荧光读数被显示在上部，其

对应于底部的图。从 ACT 形成的 ACT-荧光素偶联物的百分比 (% ACT-FL) 对温育时间 (min) 被显示。通过用 BGFL 或 BCFL 温育 0.5 μ M 的 GST-ACT10 混合物进行该实验。GST-ACT10: ACT 10 (SEQ ID NO:12)与谷胱甘肽 S-转移酶 (GST) 的融合蛋白。BGFL 和 BCFL: 参见图 1 的说明。发现的结合常数: $k_{BC} = 1130 \pm 150 \text{ M}^{-1}\text{S}^{-1}$; $k_{BG} \sim 10 \text{ M}^{-1}\text{S}^{-1}$ 。

[10] 图 3: 尿素诱导的 ACT1、ACT9 和 ACT10 解折叠。

[11] 显示的是所形成的 ACT-荧光素的百分比 (%ACT-FL)。在 0 M 尿素[(NH₂)₂CO]下所获得的值被设定为 100%。蛋白质 (0.5 μ M) 在添加有尿素(0 至 8M)的动力学缓冲液(kinetic buffer)(50 mM HEPES, pH 7.2, 1 mM DTT)中温育 30 分钟。然后该溶液被调节至 20 μ M BCFL 并且温育 2 小时。样品然后在 SDS 缓冲液中在 95°C 下沸腾 5 分钟(min)。荧光素染料标记的蛋白质混合物通过 SDS-PAGE 进行分析。该数据组用 $Y = 100 / (1 + 10^{((\log C_{1/2} - X) \cdot \text{HillSlope}))}$ 进行拟合, 以得到半解折叠尿素浓度 $C_{1/2}$ 。对于不同突变体, 发现下列 $C_{1/2}$ 值: ACT1 $2.8 \pm 0.1 \text{ M}$; ACT9 $4.1 \pm 0.1 \text{ M}$; ACT10 $5.1 \pm 0.2 \text{ M}$ 。

具体实施方式

[12] 本发明涉及称为烷基胞嘧啶转移酶(ACT)的新蛋白质, 其衍生自 O⁶-烷基鸟嘌呤-DNA 烷基转移酶, 特别适合于从式(I)底物转移标记。

[13] 烷基胞嘧啶转移酶(ACT)被定义为这样的蛋白质: (a) 由 170 至 220 个氨基酸, 优选 175 至 190 个氨基酸, 最优选 177 至 185 个氨基酸组成; (b) 包含至少一个半胱氨酸; (c) 与 O²-苄基胞嘧啶进行反应, 由此至少与在相同条件下与 O⁶-苄基鸟嘌呤的反应一样快地将苄基取代基转移至 (b) 的半胱氨酸的巯基官能上。

[14] 本发明的 ACT 在基于 AGT 的噬菌体展示的定向进化方法中从编码 AGT 的 DNA 加以制备。作为基于噬菌体展示的定向进化的起始点, 编码突变体 ^NAGT (SEQ ID NO:2, Gronemeyer, T.等, Protein Eng. Des. Sel. 19:309-316, 2006) 的 DNA 被使用。^NAGT 对苄基鸟嘌呤衍生物所展示出的活性比野生型 ^{wt}AGT (Juillerat, A. 等, Chem. Biol. 10:313-317, 2003) 高大约 50 倍并且它本身已经通过利用噬菌体展示的定向进化和苄基鸟嘌呤底物获得。残基 Tyr114、Lys131、Ser135、Val148、

Gly156、Gly157、Glu159 的密码子经由饱和诱变被随机化。噬菌粒 pAK100 中的转化产生 10^7 个独立克隆的突变体 AGT 文库。为了选择，噬菌体文库与作为底物的 BCFL——携带荧光素的苄基胞嘧啶（化合物 5）——温育，以标记与苄基胞嘧啶进行反应的突变体。这允许随后通过利用覆盖有抗荧光素抗体的磁珠富集相应的噬菌体。经过 6 轮对 BCFL 活性的筛选后，克隆 ACT1 至克隆 ACT8 通过 DNA 测序进行分析（表 1，以单字母显示的氨基酸）。在随后的测试中，明显的是，来自克隆 ACT1 至克隆 ACT8 的蛋白质对 4 M 尿素变性仅显示有限的稳定性。为了改善这一点，对同一文库进行再筛选。来自最佳分离的克隆的蛋白质被列为 ACT9 并且在 4 M 尿素中显示出良好的稳定性，但是对苄基胞嘧啶具有有限的反应性。接着，如下获得 ACT10：通过编码 ACT9 的 DNA 的易错 PCR，接着进行随后的噬菌体选择，这导致残基 Met60Ile、Ala121Val 和 Leu153Ser 进一步的修饰，作为具有高反应性和对尿素变性的高稳定性的 ACT 变体。所述克隆被表达并且所述蛋白质被纯化为谷胱甘肽 S-转移酶（GST）融合蛋白。

[15] 对苄基胞嘧啶具有特异性的这 10 种蛋白质被称为烷基胞嘧啶转移酶（ACT）1 至 10，并且是本发明的主题。

[16] 在本发明中考虑的另外的蛋白质是：

-ACT，其是 ACT 1 至 8 的同源物并且在除了位置 114、131、135、148、157 和 159 之外的位置上的一个、两个或三个氨基酸不同于 ACT 1 至 8；以及

-ACT，其是 ACT 1 至 10 的同源物并且在除了位置 60、114、121、131、135、148、153、157 和 159 之外的位置上的一个、两个或三个氨基酸不同于 ACT 1 至 10。

[17] 也被考虑在内的是 ACT 1 (SEQ ID NO 1) 的类似物，其中，

位置 114 上的氨基酸为 A、E、N、R 或 S；

位置 131 上的氨基酸为 N、S、T 或 V；

位置 135 上的氨基酸为 D、N 或 T；

位置 148 上的氨基酸为 D、E 或 Q；

位置 157 上的氨基酸为 A、G、L、T、P 或 W；和

位置 159 上的氨基酸为 E、F、M、R 或 S。

[18] 同样被考虑在内的是ACT10 (SEQ ID NO 12)的类似物,其中,
位置 60 上的氨基酸为 M 或 I;
位置 114 上的氨基酸为 A、E、N、R 或 S;
位置 121 上的氨基酸为 A 或 V;
位置 131 上的氨基酸为 N、S、T 或 V;
位置 135 上的氨基酸为 D、N 或 T;
位置 148 上的氨基酸为 D、E、Q 或 V;
位置 153 上的氨基酸为 L 或 S;
位置 157 上的氨基酸为 A、G、L、T、P 或 W; 和
位置 159 上的氨基酸为 E、F、M、R、S 或 L。

[19] 优选被称为烷基胞嘧啶转移酶 (ACT) 1、2、7、8 和 10 的蛋白质。特别优选具有在 SEQ ID NO:12 中显示的氨基酸序列的蛋白质 ACT10。还优选 ACT10 的同源物,其在除了位置 60、114、121、131、135、148、153、157 和 159 之外的位置上的一个氨基酸上与 ACT10 不同。

[20] 表 1. ^NAGT 和 ACT 1 至 10 的氨基酸序列

	残基									
	60	114	121	131	135	148	153	157	159	
^N AGT	M	Y	A	K	S	V	L	G	E	SEQ ID NO:2
ACT1	M	R	A	S	D	E	L	G	M	SEQ ID NO:1
ACT2	M	A	A	V	D	Q	L	W	R	
ACT3	M	S	A	T	T	D	L	P	E	
ACT4	M	N	A	N	D	E	L	A	S	
ACT5	M	A	A	T	S	E	L	K	E	
ACT6	M	E	A	R	A	E	L	E	E	
ACT7	M	S	A	V	D	Q	L	L	R	
ACT8	M	E	A	N	N	E	L	T	F	
ACT9	M	E	A	N	D	V	L	P	F	
ACT10	I	E	V	N	D	V	S	P	L	SEQ ID NO:12

[21] 本发明进一步涉及产生烷基胞嘧啶转移酶的方法,其特征在于: 编码 AGT 或 ACT 的 DNA 在可达 10 个的氨基酸位置上通过饱和诱变被随机化, 所获得的文库被转化到合适的噬菌粒中, 期望的噬菌体通过与携带标记的苄基胞嘧啶反应进行选择, 然后, 所述标记转移到其上的噬菌体采用覆盖有针对所述标记的抗体的磁珠进行分离。本发明也涉及这种定向进化方法的产物。

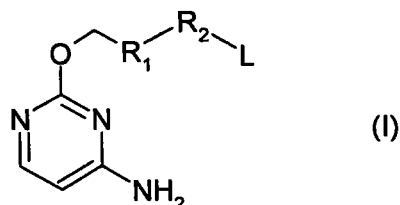
[22] 饱和诱变在本领域中是已知的并且例如如 Dube, D. K. 和 Loeb, L. A., *Biochemistry*, 28:5703-5707, 1989 所述进行。

[23] 采用噬菌体和噬菌粒的定向进化方法也是众所周知的, 并且被描述在例如 Smith, G. P. 和 Petrenko, V.A., *Chem. Rev.* 97:391-410, 1997; Hoess, R. H. et al., *Chem. Rev.* 101 :3205-3218, 2001 中。优选的方法在系统中利用噬菌粒 pAK100, 如 Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R. and Plückthun, A., *J. Immunol. Methods* 201 :35-55, 1997 所述。

[24] 适合于选择期望克隆的化合物是如式(I)所描述并且是本发明主题的底物, 尤其是其中 R_1 是对位取代苯基的式(I)底物。被转移到期望 ACT 的这种底物的标记可以是其抗体容易获得的任何标记, 并且不限于荧光染料标记或任何其它光谱标记。

[25] 利用携带抗特定标记的抗体的磁珠进行分离在本领域中也是众所周知的, 并且被描述在例如 Gronemeyer 等, *Protein Eng. Des. Sel.* 19:309-316, 2006 中。

[25A] 在进一步的方面, 本发明涉及式(I)化合物,



其中,

[26] R_1 是芳族基或杂芳族基, 或者是双键连接于 OCH_2- 的任选取代的不饱和烃基、环烃基或杂环基;

R_2 是连接体; 并且

L 是一个标记或多个相同或不同的标记。

作为芳族基的 R_1 优选是苯基或萘基, 尤其是对位或间位被 R_2 取代的苯基。

[27] 杂芳族基 R_1 是单环或双环杂芳基, 其包括零个、一个、两个、三个或四个环氮原子以及零个或一个氧原子以及零个或一个硫原子, 条件是至少一个环原子是氮、氧或硫原子, 并且其具有 5 个至 12 个环原子, 优选 5 个或 6 个环原子; 并且其除了携带取代基 R_2 外可以未被

取代或者被一个或多个尤其是一个选自下列的另外的取代基取代：低碳烷基，如甲基；低碳烷氧基，如甲氧基或乙氧基；卤素，如氯、溴或氟；卤代低碳烷基，如三氟甲基；或羟基。

[28] 优选地，单环或双环杂芳基 R_1 选自 2H-吡咯基、吡咯基、咪唑基、苯并咪唑基、吡唑基、吡啶基、吡嗪基、嘧啶基、哒嗪基、4H-喹啉基、异喹啉基、喹啉基(quinolyl)、2,3-二氮杂蔡基、1,5-二氮杂蔡基、喹喔啉基、喹唑啉基、喹啉基(quinoliny)、蝶啶基、中氮茛基(indoliziny)、3H-吡啶基、吡啶基、异吡啶基、噁唑基、异噁唑基、噻唑基、异噻唑基、三唑基、四唑基、呋喃基(furazany)、苯并[d]-吡唑基、噻吩基和呋喃基。更优选地，单环或双环杂芳基选自吡咯基；咪唑基，如 1H-咪唑-1-基；苯并咪唑基，如 1-苯并咪唑基；吡唑基，尤其是 5-吡唑基；吡啶基，例如 2-、3-或 4-吡啶基；嘧啶基，尤其是 2-嘧啶基；吡嗪基；异喹啉基，尤其是 3-异喹啉基；喹啉基，尤其是 4-或 8-喹啉基；吡啶基，尤其是 3-吡啶基；噻唑基；三唑基；四唑基；苯并[d]-吡唑基；噻吩基；和呋喃基。

[29] 在本发明的特别优选的实施方式中，杂芳基 R_1 是噻吩基，尤其是 2-噻吩基，其在 3-、4-或 5-位上，优选在 4-位上，携带另外的取代基 R_2 ；或者是 3-噻吩基，其在 4-位上携带另外的取代基 R_2 。

[30] 任选取代的不饱和烃基 R_1 是 1-烯基，其在 1-或 2-位上，优选在 2-位上携带另外的取代基 R_2 ，或者 1-炔基。被考虑在 1-烯基中的取代基是例如低碳烷基，如甲基；低碳烷氧基，如甲氧基；低碳酰氧基，如乙酰氧基；或卤代基，如氯代基。在本发明特别优选的实施方式中， R_1 是 1-炔基。

[31] 任选取代的不饱和环烷基是在 1-位上不饱和的、具有 5 至 7 个碳原子的环烯基，例如 1-环戊烯基或 1-环己烯基，其在任何位置上携带另外的取代基 R_2 。被考虑的取代基是例如低碳烷基，如甲基；低碳烷氧基，如甲氧基；低碳酰氧基，如乙酰氧基；或卤代基，如氯代基。

[32] 任选取代的不饱和杂环基具有 3 至 12 个原子，1 至 5 个选自氮、氧和硫的杂原子，以及在将该杂环基与 OCH_2 -中的亚甲基连接的位置上的双键。被考虑的取代基是例如低碳烷基，如甲基；低碳烷氧

基，如甲氧基；低碳酰氧基，如乙酰氧基；或卤代基，如氯代基。

[33] 特别地，任选取代的不饱和杂环基是部分饱和的杂芳族基，如上文对于杂芳族基 R_1 所定义。这种杂环基的实例是异噁唑烷基，尤其是 3-异噁唑烷基，其在 5-位上携带另外的取代基；或者 5-异噁唑烷基，其在 3-位置上携带另外的取代基。

[34] 连接体基团 R_2 优选是柔性连接体，其将标记 L 或多个相同或不同的标记 L 连接于底物。在预想的应用的背景下，即在将底物转移到包含 ACT 的融合蛋白的背景下，选择连接体单元。它们也增加了所述底物在适当的溶剂中的溶解性。所用的连接体在实际应用的条件下是化学稳定的。所述连接体既不干扰与 ACT 的反应也不干扰标记 L 的检测，但是可被构造以便在式(I)化合物与包含 ACT 的融合蛋白反应之后及时在某一点被切割。

[35] 连接体 R_2 是具有 1 至 300 个碳原子的直链或支链亚烃基，其中任选地

(a) 一个或多个碳原子被氧置换，特别地，其中每隔两个碳原子被氧置换，例如具有 1 至 100 个乙烯氧基单元的聚乙烯氧基；

(b) 一个或多个碳原子被携带氢原子的氮置换，并且相邻的碳原子被氧代(oxo)取代，其代表酰胺官能-NH-CO-；

(c) 一个或多个碳原子被氧置换并且相邻的碳原子被氧代取代，其代表酯官能-O-CO-；

(d) 两个相邻碳原子之间的键是双键或三键，其代表官能-CH=CH-或-C≡C-；

(e) 一个或多个碳原子被下列基团置换：亚苯基、饱和或不饱和的亚环烷基、饱和或不饱和亚双环烷基、桥连杂芳基或者桥连饱和或不饱和杂环基；

(f) 两个相邻碳原子被下列置换：二硫键-S-S-；

或者两个或多个、尤其是两个或三个亚烃基和/或上文(a)至(f)所定义的改性亚烃基的组合，其任选地含有取代基。

[36] 所考虑的取代基是例如低碳烷基，如甲基；低碳烷氧基，如甲氧基；低碳酰氧基，如乙酰氧基；或卤代基，如氯代基。

[37] 考虑的另外的取代基是例如当 α -氨基酸，特别是天然存在的

α -氨基酸被掺入连接体 R_2 时获得的那些取代基, 在所述连接体 R_2 中碳原子被酰胺官能-NH-CO-置换, 如(b)所定义。在这种连接体中, 亚烃基 R_2 的部分碳链被基团-(NH-CHR-CO) $_n$ -置换, 其中 n 在 1 至 100 之间并且 R 代表 α -氨基酸的不同残基。

[38] 另外的取代基是产生可光致断裂连接体 R_2 的取代基, 如邻硝基苯基。特别地, 该取代基邻硝基苯基位于与酰胺键相邻的碳原子上, 例如在基团-NH-CO-CH₂-CH(邻硝基苯基)-NH-CO-中与酰胺键相邻的碳原子上, 或者作为聚乙二醇链中的取代基, 例如在基团-O-CH₂-CH(邻硝基苯基)-O-中。考虑的其它可光致断裂连接体是例如苯甲酰甲基、烷氧基苯偶姻 (alkoxybenzoin)、苄硫醚和新戊酰二醇 (pivaloyl glycol) 衍生物。

[39] 如上文(e)所定义的置换碳原子的亚苯基是例如 1,2-、1,3-或优选 1,4-亚苯基。在具体的实施方式中, 亚苯基进一步被硝基取代并且与如上面(a)、(b)、(c)、(d)和(f)所提及的其它置换相组合; 代表可光致断裂基团; 并且例如是 4-硝基-1,3-亚苯基, 如在-CO-NH-CH₂-4-硝基-1,3-亚苯基-CH(CH₃)-O-CO-中的 4-硝基-1,3-亚苯基, 或者是 2-甲氧基-5-硝基-1,4-亚苯基, 如在-CH₂-O-2-甲氧基-5-硝基-1,4-亚苯基-CH(CH₃)-O-中的 2-甲氧基-5-硝基-1,4-亚苯基。代表可光致断裂连接体的其它具体实施方式是例如-1,4-亚苯基-CO-CH₂-O-CO-CH₂-(苯甲酰甲基)、-1,4-亚苯基-CH(OR)-CO-1,4-亚苯基-(烷氧基苯偶姻)或者-3,5-二甲氧基-1,4-亚苯基-CH₂-O-(二甲氧基苄基部分)。如上文(e)所定义的置换碳原子的饱和或不饱和亚环烷基衍生自具有 3 至 7 个碳原子的环烷基, 优选地衍生自环戊基或环己基, 并且例如是 1,2-或 1,3-亚环戊基; 1,2-、1,3-或优选 1,4-亚环己基; 或者也可以是例如在 1-或 2-位上不饱和的 1,4-亚环己基。如上文(e)所定义的置换碳原子的饱和或不饱和亚双环烷基衍生自具有 7 个或 8 个碳原子的双环烷基, 并且例如是双环[2.2.1]庚烯或双环[2.2.2]辛烯, 优选 1,4-双环[2.2.1]庚烯, 其任选地在 2-位上不饱和或者在 2-和 5-位上双重不饱和; 以及 1,4-双环[2.2.2]辛烯, 其任选地在 2-位上不饱和或者在 2-和 5-位上双重不饱和。如上文(e)所定义的置换碳原子的桥连杂芳族基例如是亚三唑烷基(triazolidene), 优选 1,4-亚三唑烷基, 或者是亚异噁唑烷基(isoxazolidene), 优选 3,5-亚异噁唑烷基。

如上文(e)所定义的置换碳原子的桥连饱和或不饱和杂环基例如衍生自如上面 R_1 中所定义的不饱和杂环基,例如亚异噁唑烷基(isoxazolidene), 优选 3,5-亚异噁唑烷基, 或者全饱和杂环基, 其具有 3 至 12 个原子, 其中的 1 至 3 个是选自氮、氧和硫的杂原子, 例如吡咯烷二基(pyrrolidinediyl)、哌啶二基、四氢呋喃二基、二噁烷二基、吗啉二基或四氢噻吩二基, 优选 2,5-四氢呋喃二基或 2,5-二噁烷二基。考虑的具体杂环基是糖部分, 例如 α -或 β -呋喃糖基或者 α -或 β -吡喃糖基部分。

[40] 连接体 R_2 中的环亚结构降低了分子柔性, 如通过 R_2 内可旋转键的数目所测定, 其导致更好的膜渗透速率, 这对于所有体内标记应用来说是重要的。

[41] 连接体 R_2 优选是具有 1 至 25 个碳原子的直链亚烃基或者具有 4 至 100 个乙烯氧基单元的直链聚乙二醇基团, 其任选地通过 $-\text{CH}=\text{CH}-$ 或 $-\text{C}\equiv\text{C}-$ 基团连接于基团 R_1 。进一步优选的是具有 1 至 25 个碳原子的直链亚烃基, 其中碳原子任选地被酰胺官能 $-\text{NH}-\text{CO}-$ 置换, 并且任选地携带可光致断裂亚单元, 例如邻硝基苯基。进一步优选的是支链连接体, 其包括具有 3 至 6 个乙二醇单元的聚乙二醇基团以及其中碳原子被酰胺键置换的亚烃基, 并且进一步携带取代的氨基和羟基官能。其它优选的支链连接体具有树枝状(树状)结构, 其中胺、酰胺和/或醚官能置换亚烃基的碳原子。

[42] 特别优选的连接体 R_2 是具有 2 至 20 个碳原子的直链亚烃基, 其中一个或两个碳原子被氮置换并且所述直链亚烃基在与所述氮相邻处任选地被氧代(oxo)取代。

[43] 另一特别优选的连接体 R_2 是具有 10 至 40 个碳原子的直链亚烃基, 其任选地被氧代取代, 其中 3 至 12 个碳原子被氧置换并且一个或两个碳原子被氮置换, 任选地在氮相邻处被氧代取代。

[44] 连接体 R_2 可携带一个或多个相同或不同的标记, 例如 1 至 100 个相同或不同的标记, 特别是 1 至 5 个, 优选一个、两个或三个, 特别是一个或两个相同或不同的标记。

[45] 低碳烷基是具有 1 至 7 个、优选 1 至 4 个 C 原子的烷基, 并且是直链的或支链的, 优选地, 低碳烷基是丁基, 如正丁基、仲丁基、异丁基、叔丁基, 丙基, 如正丙基或异丙基, 乙基或甲基。更优选地,

低碳烷基是甲基。

[46] 在低碳烷氧基中，低碳烷基如上文所定义。低碳烷氧基优选指正丁氧基、叔丁氧基、异丙氧基、乙氧基或甲氧基，特别是甲氧基。

[47] 在低碳酰氧基中，低碳酰基具有甲酰或低碳烷基羰基的含义，其中低碳烷基如上文所定义。低碳酰氧基优选指正丁酰氧基、正丙酰氧基、异丙酰氧基、乙酰氧基或甲酰氧基，特别是乙酰氧基。

[48] 卤素是氟、氯、溴或碘，特别是氯。

[49] 底物的标记 L 可以由本领域技术人员根据融合蛋白所预期的应用进行选择。标记使得携带标记 L 的标记融合蛋白容易进行检测或者与其环境分离。考虑的其它标记是那些能够感应并诱导标记融合蛋白和/或底物的环境变化的标记；或者这样的标记，其由于底物的物理和/或化学特性而有助于操纵融合蛋白并且被特异引入所述融合蛋白。在本发明的上下文中所理解的标记是这样的取代基，其不同于氢或标准官能团，特别是不同于氢、羟基、氨基、卤素、羧酸酯、酰胺、羧酸酯、腈、氰酸酯、异氰酸酯、磺酸酯、磺酰胺、磺酸酯、醛、酮、醚和硫醚取代基。

[50] 标记的实例包括光谱探针例如荧光团或生色团、磁探针或造影剂；放射性标记分子；这样的分子，其是能特异性地与配偶体结合的特异性结合对的一部分；被认为与其它生物分子相互作用的分子；被认为与其它生物分子相互作用的分子文库；能交联到其它分子上的分子；在暴露于 H_2O_2 和抗坏血酸盐后能产生羟基的分子，如系留金属螯合物 (tethered metal-chelate)；在光照射后能产生反应性基团的分子，例如孔雀绿；共价连接于固相支持体的分子，其中所述支持体可以是载玻片、微量滴定板或本领域技术人员所知的任何聚合物；核酸或其衍生物，其能与其互补链进行碱基配对；脂质或其它疏水性分子，具有膜插入特性；具有期望的酶促性质、化学性质或物理性质的生物分子；或者具有上述所列任何特性的组合的分子。

[51] 另外的标记 L 是带正电荷的线性或支化聚合物，其已知促进所连接的分子在活细胞质膜上的转移。这对于这样的底物特别重要，所述底物另外具有低的细胞膜渗透性或者对活细胞的细胞膜实际上是不可渗透的。非细胞可渗透的 ACT 底物在与这种基团 L 偶联后将变成

细胞可渗透。这种细胞膜转运增强基团 L 包括，例如，具有 6-15 个精氨酸残基的 D-和/或 L-精氨酸的线性聚（精氨酸）；具有 6-15 个亚单元的线性聚合物，所述亚单元每个都携带胍盐基团；具有 6 至高达 50 个亚单元的低聚物或短长度聚合物，其一部分已连接胍盐基团；和/或 HIV-tat 蛋白质的序列的部分，特别是亚基 Tat49-Tat57 (以一字母氨基酸码表示为 RKKRRQRRR)。ACT 底物通过如上文所定义的连接体 R₂ 共价连接于该基团 L，其优选在活细胞内是不稳定的并且可以被降解，例如通过被细胞内酯酶断裂酯基 R₂ 而降解，这直接导致或者在由酯官能的断裂所引起的进一步反应中导致 ACT 底物与增强细胞膜渗透性的单元 L 的分离。

[52] 优选作为标记 L 的是光谱探针，以及这样的分子，其是能特异性地与配偶体结合的特异性结合对的一部分，即所谓的亲和标记。还优选作为标记 L 的是共价连接于固相支持体的分子。优选的光谱探针是荧光团。

[53] 当标记 L 是荧光团、生色团、磁标记、放射性标记或类似物时，检测通过适合于所述标记的标准方法进行并且不论该方法被体外或体内应用。如果 L 是荧光团，则该方法可以与绿色荧光蛋白(GFP)的应用相媲美，所述绿色荧光蛋白被基因融合于目标蛋白，并且允许在活细胞中进行蛋白质研究。标记 L 的具体实例还是展示非线性光学特性的硼化合物。特别优选的是这样的标记，其使得一个底物的 L (L₁) 是两个相互作用的光谱探针 L₁/L₂ 的一个成员并且另一底物的 L (L₂) 是另一成员，其中能量可以在供体与受体（猝灭剂）之间当它们非常接近时（小于 10 纳米的距离）通过动态淬灭或静态淬灭而非辐射地转移。在携带 L₁ 的标记底物与 ACT 融合蛋白以及携带 L₂ 的另一种标记底物（例如，苜蓿基鸟嘌呤型）与另一相应的融合蛋白如 AGT 融合蛋白的反应后，这种标记对 L₁/L₂ 改变了其光谱特性。这种标记对 L₁/L₂ 的实例是下列详细解释的 FRET 对。

[54] 考虑的具体荧光团是：Alexa Fluor 染料，其包括 Alexa Fluor 350、488、532、546、555、635 和 647 (Invitrogen Corp., Carlsbad, CA 92008, USA, 还可参见 Panchuk-Voloshina, N.等, J. Histochem. & Cytochem. 47:1179-1188, 1999); 香豆素，例如 7-二甲氨基-香豆素-4-乙酸（由

Invitrogen Molecular Probes 作为产品 D374 供应的琥珀酰亚胺酯)、7-氨基-4-甲基-香豆素-3-乙酸和 7-二乙氨基-香豆素-3-甲酸; 花青 (Cyanine)-3 (Cy 3)、花青 5 (Cy5) 和花青 5.5 (Cy5.5) (Amersham - GE Healthcare, Solingen, 德国); ATTO 488、ATTO 532、ATTO 600 和 ATTO 655 (Atto-Tec, D57076 Siegen, 德国); DY-505、DY-547、DY-632 和 DY-647 (Dyomics, Jena, 德国); 以及 5(6)-羧基荧光素和二氟-5(6)-羧基荧光素(Oregon Green)。式(I)底物上的这些特定标记可以与已知的 AGT-苄基鸟嘌呤系统进行组合, 其中苄基鸟嘌呤上的标记是猝灭剂, 其与在与 ACT 反应的式(I)底物上的荧光团形成 FRET 对。这种猝灭剂是: QSY 35、QSY 9 和 QSY 21 (Invitrogen Molecular Probes); BHQ-1、BHQ-2 和 BHQ-3 (Biosearch Technologies, Inc., Novato, CA 94949, USA 的 Black Hole Quencher™); ATTO 540Q 和 ATTO 612Q (Atto-Tec, D57076 Siegen, 德国); 4-二甲氨基-偶氮苯-4'-磺酰基衍生物(Dabsyl)以及 4-二甲氨基偶氮苯-4'-羰基衍生物(Dabcyl)。

[55] 取决于标记 L 的特性, 包含目标蛋白和 ACT 的融合蛋白可以在与底物的反应后被结合于固相支持体上。与包含 ACT 的融合蛋白反应的底物的标记 L 当开始与 ACT 反应时可以已经连接于固相支持体, 或者可以随后即转移到 ACT 之后被用于将标记的 ACT 融合蛋白连接于固相支持体。所述标记可以是特定结合对的一个成员, 所述特定结合对的另一成员共价地或者通过任何其它方法连接于或可连接于所述固相支持体。考虑的特定结合对例如是生物素和抗生素蛋白或链霉抗生物素蛋白。所述结合对的任一成员可以是底物的标记 L, 另一个成员被连接于所述固相支持体。允许方便地结合于固相支持体的标记的进一步的例子是例如麦芽糖结合蛋白、糖蛋白、FLAG 标签或者反应性取代基, 所述反应性取代基允许在这种取代基与固相支持体表面上的互补官能团之间的化学选择性反应。反应性取代基与互补官能团的这种配对的实例例如是胺和活性羧基, 其形成酰胺; 叠氮化物和丙炔酸衍生物, 其经历 1,3-偶极环加成反应; 胺和另一胺官能团, 其与加入的活化性双二羧酸衍生物类型的双官能连接体试剂反应, 产生两个酰胺键; 或者本领域已知的其它组合。

[56] 合适的固相支持体的实例例如是玻璃表面如载玻片、微量滴

定板和合适的传感元件，尤其是官能化聚合物（例如以珠的形式）；化学改性的氧化物表面，例如二氧化硅、五氧化钽或二氧化钛；或者还有化学改性的金属表面，例如贵金属表面如金或银表面。然后，不可逆地连接和/或点滴 ACT 底物可被用于以空间分辨的方式、尤其是通过点滴将 ACT 融合蛋白连接在固相支持体上，形成蛋白质微阵列、DNA 微阵列或小分子阵列。

[57] 当标记 L 在暴露于外部刺激物后能产生反应性基团例如羟基时，则产生的基团可使 ACT 融合蛋白以及与该 ACT 融合蛋白非常接近的那些蛋白质失活，这允许研究这些蛋白质的作用。这种标记的实例是系留金属螯合物络合物，其经暴露于 H_2O_2 和抗坏血酸盐后产生羟基；以及生色团例如孔雀绿，其经激光照射后产生羟基。生色团和激光用于产生羟基作为生色团辅助性激光诱导失活（CALI）在本领域中也是已知的。在本发明中，用携带作为标记 L 的生色团例如孔雀绿的底物标记 ACT 融合蛋白以及随后的激光照射以时间控制和空间分辨的方式使得标记的 ACT 融合蛋白以及与所述 ACT 融合蛋白相互作用的那些蛋白质失活。该方法在体内或体外都可以被应用。此外，与该 ACT 融合蛋白非常接近的蛋白质同样可以如下进行鉴别：通过借助特异性抗体检测该蛋白质的片段，通过在高分辨率二维电泳凝胶上那些蛋白质的消失，或者通过经由分离和测序技术如质谱或借助 N-端降解的蛋白质测序来鉴定断裂的蛋白质片段。

[58] 当标记 L 是可交联于其它蛋白质的分子，例如，含有官能团如马来酰亚胺、活性酯或叠氮化物的分子以及本领域技术人员已知的其它分子时，使这种标记 ACT 底物与 ACT 融合蛋白——所述 ACT 融合蛋白与其它蛋白质相互作用——接触（体内或体外），导致 ACT 融合蛋白经由标记而与其的相互作用蛋白质发生共价交联。这允许鉴别与 ACT 融合蛋白相互作用的蛋白质。光致交联用标记 L 例如是二苯甲酮。在交联的特殊方面，标记 L 是本身为 ACT 底物的分子，这导致 ACT 融合蛋白的二聚化。这种二聚体的化学结构可以是对称的（同型二聚体）或非对称的（异型二聚体）。

[59] 考虑的其它标记是例如富勒烯（fullerenes）；用于中子俘获治疗（neutron capture treatment）的硼烷；核苷酸或低聚核苷酸，例如

用于自寻址芯片；肽核酸；和金属螯合物，例如与 DNA 特异性结合的铂螯合物。

[60] 具有期望的酶促性质、化学性质或物理性质的具体生物分子是氨甲喋呤。氨甲喋呤是酶二氢叶酸还原酶 (DHFR) 的紧密结合抑制剂。其中 L 是氨甲喋呤的式(I)化合物属于众所周知的被称为“二聚化化学诱导物 (CID)”的类别。利用 ACT 与 DNA 结合结构域 LexA 的融合蛋白以及将 DHFR 与转录激活结构域 B42 添加到用其中 L 是氨甲喋呤的式(I)化合物体内标记 ACT 融合蛋白中，诱导了 ACT-LexA 融合蛋白与 DHFR-B42 融合蛋白的偶联 (“二聚化”)，这导致 LexA 和 B42 空间上接近以及随后的转录刺激。

[61] 如果底物携带两种或多种标记，这些标记可以相同或不同。具体的优选组合是两种不同的亲和标记，或者一种亲和标记和一种生色团标记，尤其是一种亲和标记和一种荧光团标记，或者一对光谱相互作用标记 L_1/L_2 ，例如 FRET 对。

[62] 优选式(I)化合物，其中 R_1 是苯基，特别是对位取代的苯基。

[63] 优选式(I)化合物，其中 L 是光谱探针例如荧光团。同样优选这样的化合物，其中 L 是代表特异性结合对的一部分的分子；以及这样的化合物，其中 L 是共价连接于固相支持体的分子。

[64] 最优选的是实施例中的化合物。

[65] 本发明进一步涉及检测和/或操纵目标蛋白的方法，其中所述目标蛋白被掺入 ACT 融合蛋白中，所述 ACT 融合蛋白与上文所述的携带标记的式(I)化合物接触，并且所述 ACT 融合蛋白在设计来识别和/或处理所述标记的系统中利用所述标记进行检测并且任选地被进一步操纵。

[66] 在本发明的方法中，目标蛋白或肽被融合于 ACT。目标蛋白或肽可以具有任何长度并且具有或不具有二级、三级或四级结构，并且优选由至少 12 个氨基酸并可达 2000 个氨基酸组成。这种目标蛋白或肽的实例例如是酶、DNA 结合蛋白、转录调控蛋白、膜蛋白、核受体蛋白、核定位信号蛋白、蛋白辅因子、小单体 GTP 酶、ATP 结合盒蛋白、胞内结构蛋白、具有负责将蛋白质靶向特定细胞区室的序列的蛋白质、一般用作标记或亲和标签的蛋白质、以及上述蛋白质的结构

域或子结构域。目标蛋白或肽优选通过连接体被融合于 ACT，所述连接体可以被酶切割，例如在 DNA 阶段被合适的限制性酶切割和/或在蛋白质阶段可被合适的酶切割。

[67] ACT 具有将底物上即式(I)化合物上存在的标记转移到形成融合蛋白的一部分的 ACT 的半胱氨酸残基之一上的特性。在优选的实施方式中，ACT 选自上文所定义的 ACT 1 至 ACT 10 及其同源物。特别优选的是 ACT 10。

[68] 包含目标蛋白和 ACT 的融合蛋白与具体的式(I)底物接触。对反应条件进行选择，使得 ACT 与底物反应并且转移所述底物的标记。常用的条件是室温下例如约 25°C 下约 pH 7 的缓冲溶液。然而，可以理解，ACT 也在各种其它条件下反应，并且本文所提及的那些条件不限制本发明的范围。

[69] 所述底物的标记由本领域技术人员根据所述融合蛋白的预期应用进行选择。在包含 ACT 的融合蛋白与底物接触后，标记 L 被共价结合于融合蛋白。然后，标记 ACT 融合蛋白进一步借助转移的标记进行操纵和/或检测。标记 L 可以由多个相同或不同的标记组成。如果底物包含一个以上标记 L，相应的标记 ACT 融合蛋白也将包含一个以上标记，其给出了进一步操纵和/或检测标记融合蛋白的更多选择。

[70] 在本发明意义中的“检测 (Detected)”是指，具有携带标记 L 的 ACT 的融合蛋白由于所述标记的特性而可以进行定位，并且融合蛋白的量可以直接地或者通过参考标准物进行测定。在本发明意义中的“操纵 (Manipulated)”是指，具有携带标记 L 的 ACT 的融合蛋白由于所述标记的特性而可进一步进行反应，例如与体外或体内系统分离、富集和纯化，即与其它蛋白质和/或非蛋白质材料分离，引入到溶液中（例如，尤其是如果所述融合蛋白不可溶时），沉淀（例如，尤其是如果融合蛋白可溶时）或者另外地固定于固体上，以及还进一步进行处理，例如通过切割连接体，这分离所述标记并且再生所述融合蛋白。普通技术人员充分理解由于标记 L 的特性以及将 L 连接于融合蛋白的连接体 R₂ 的特性所带来的进一步处理携带标记 L 的融合蛋白的许多可能性。

[71] 在体外，ACT 融合蛋白与本发明的底物的反应一般可以在细

胞提取物中进行或者用提纯或富集形式的 ACT 融合蛋白进行。

[72] 如果关于本发明底物的试验在体内或在细胞提取物中进行，内源性 AGT 的反应将不会干扰 ACT 融合蛋白与式(I)底物的反应，因为所述底物不与内源性 AGT 相互作用（或者至少不可检测地与内源性 AGT 相互作用），只与 ACT 相互作用。这是本组合 ACT-式(I)底物相比于现有技术中所述的 AGT-苄基鸟嘌呤型底物的标准组合的实质性优点。

[73] 如上文所述的式(I)底物上的具体荧光团标记可以与已知的 AGT-苄基鸟嘌呤系统进行组合，其中苄基鸟嘌呤上的标记是猝灭剂，其与在与 ACT 反应的式(I)底物上的荧光团形成 FRET 对。优选的是上文所列的、作为式(I)底物的标记 L 的猝灭剂。

[74] 可选地，式(I)底物上的标记 L 可以是上述猝灭剂之一，并且与 ACT 的反应可以在含有 AGT 融合蛋白-苄基鸟嘌呤底物组合的混合物中进行，其中苄基鸟嘌呤上的标记是所提及的荧光团之一。

[75] 本发明的底物一般通过本领域已知的标准方法进行制备。有用的原材料是 2-氯嘧啶-4-胺或者在嘧啶环的位置 2 上具有活化离去基团的另一胞嘧啶衍生物。该化合物然后与式 $\text{HO-CH}_2\text{-R}_1\text{-R}_2\text{-L}$ 的醇或类似化合物 $\text{HO-CH}_2\text{-R}_1\text{-R}_2'$ 进行反应，其中 R_2' 是连接体 R_2 的母体，其允许标记 L 的引入并且由此完成残基 $\text{R}_2\text{-L}$ 。特别有用的母体 R_2' 是具有氨基官能的那些母体，所述氨基官能允许与携带羧基官能的标记进行缩合反应，例如以琥珀酰亚胺酯的形式，这产生掺入酰胺官能的连接体 R_2 。

[76] 对于预期的官能度的合适保护基可以由本领域技术人员进行选择，并且例如被概括在 Greene, T.W. 和 Wuts, P. G. M., "Protective Groups in Organic Synthesis", John Wiley & Sons, New York 1991 中。

[77] 本发明进一步涉及将标记从这些式(I)底物转移到烷基胞嘧啶转移酶(ACT)和 ACT 融合蛋白的方法。

[78] 组合 ACT-式(I)底物与已知的 AGT-苄基鸟嘌呤型底物一起特别适合于跟踪两种不同的蛋白质，这是因为式(I)底物和苄基鸟嘌呤型底物将它们的标记分别选择性地转移到 ACT 和 AGT 融合蛋白上。

[79] 为了证实利用 ACT 进行特异性双重标记的可行性，六组氨酸

标记的 ^NAGT ($6\text{xHis-}^N\text{AGT}$)和 GST-ACT1 融合蛋白的等摩尔混合物用 BGFL、BGCy5、BCFL、BCCy5 或者 BGCy5/BCFL 或 BGFL/BCCy5 的等摩尔混合物进行温育。与 BGFL 或 BGCy5 的温育导致 $6\text{xHis-}^N\text{AGT}$ 的完全标记以及 GST-ACT1 非常低的标记（在两种情况下估计低于 5%）。以相反的方式，与 BCFL 或 BCCy5 的温育导致 GST-ACT1 的完全标记以及 $6\text{xHis-}^N\text{AGT}$ 的部分标记（分别为 76%和 6%）。

[80] 然而，与 BGCy5/BCFL (或 BGFL/BCCy5)的等摩尔混合物的温育导致特异性双重彩色标记，即 $6\text{xHis-}^N\text{AGT}$ 用 Cy5 染料（或分别用荧光素）的特异性标记以及 GST-ACT1 用荧光素（或分别用 Cy5 染料）的特异性标记。在两种情况下，交叉标记已经被评价为小于 1%。

表 2. ^NAGT 和 ACT 1 至 10 作为与 BCFL 和 BGFL 的 GST-融合蛋白的活性

	$[\text{s}^{-1}.\text{M}^{-1}]$	
	k_{BCFL}	k_{BGFL}
^NAGT	26	28000
ACT 1	730	<10
ACT 2	330	<10
ACT 3	100	n.d.
ACT 4	<100	n.d.
ACT 5	350	210
ACT 6	460	160
ACT 7	260	<10
ACT 8	600	<10
ACT 9	90	<1
ACT 10	1130	10

n.d.: 未测定

[81] 实施例

缩写:

BC = 苄基胞嘧啶

BG = 苄基鸟嘌呤

DTT = 二硫苏糖醇

DMF = 二甲基甲酰胺

MPLC = 中压液相色谱

PBS = 磷酸缓冲盐溶液

RT = 室温

TFA = 三氟乙酸

TEA = 三乙胺

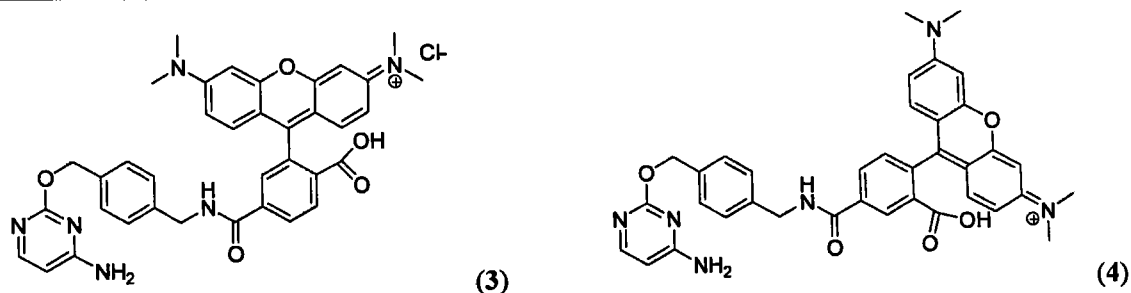
[82] 实施例 1: N-(4-((4-氨基嘧啶-2-基氧)甲基)苄基)-2,2,2-三氟乙酰胺 (1)

在氩气气氛下将 760 mg (3.25 mmol) 2,2,2-三氟-N-(4-羟甲基-苄基)-乙酰胺溶解于 3 mL 干燥的二甲基乙酰胺中, 并且将 273 mg (8.15 mmol) NaH 在 5 分钟 (min) 内加入。然后加入 211 mg (1.63 mmol) 2-氯嘧啶-4-胺并且该溶液在 90°C 下搅拌过夜。将 1 ml 水小心加入以猝灭所有过量的 NaH, 并且该混合物被倒入 50 ml 0.5 N HCl 中。粗产物用乙酸乙酯进行萃取, 合并的有机相用盐水进行洗涤并且经 MgSO_4 干燥。溶剂蒸发后, 产物通过急骤柱层析 (从 1:1 至 3:1 的乙酸乙酯:环己烷梯度) 进行纯化。产量: 350 mg (52%)。ESI-MS m/z 327 $[\text{M}+\text{H}]^+$ 。

[83] 实施例 2: 2-(4-(氨甲基)苄氧基)4-氨基嘧啶 (2)

将 150 mg (0.46 mmol) N-(4-((4-氨基嘧啶-2-基氧)甲基)苄基)-2,2,2-三氟乙酰胺(1)溶解于 2 ml 甲醇中并且用 5 ml 甲胺(33%乙醇溶液)进行处理。反应混合物在室温下搅拌过夜并且在真空中去除所有挥发物。产物被用于下一步, 而没有进一步提纯。ESI-MS m/z 231 $[\text{M}+\text{H}]^+$ 。

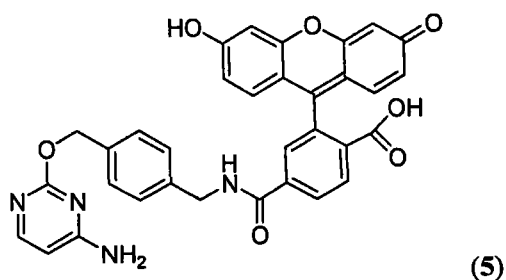
[84] 实施例 3: N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-四甲基若丹明-6-甲酰胺 (3)和 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-四甲基若丹明-5-甲酰胺(4)



将 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) (3.6 mg, 0.016 mmol)和 5(6)-羧基四甲基若丹明琥珀酰亚胺酯 (8.2 mg, 0.016 mmol) 与 2.4 μl TEA 一起溶解于 800 μl DMF 中并且在 31°C 下加热过夜。在真空中蒸发溶剂

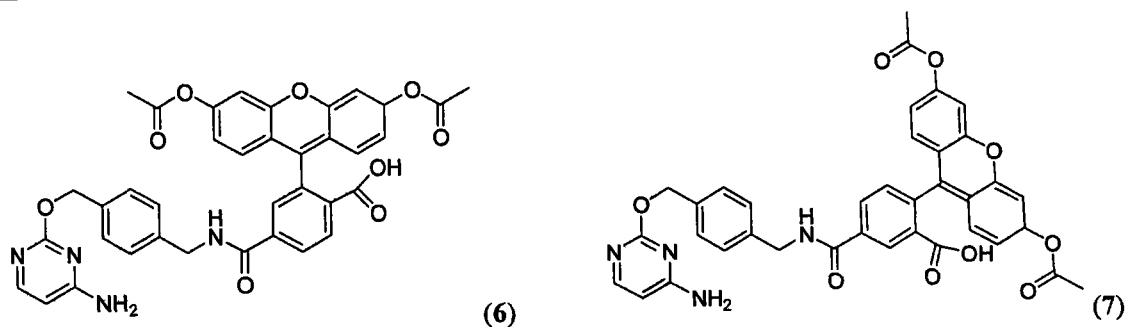
并且化合物利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上通过反相 MPLC (中压液相色谱) 进行分离。MS (ESI) m/z 643 $[M-Cl]^+$ 。

[85] 实施例 4: N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-荧光素-6-甲酰胺 (5), BCFL



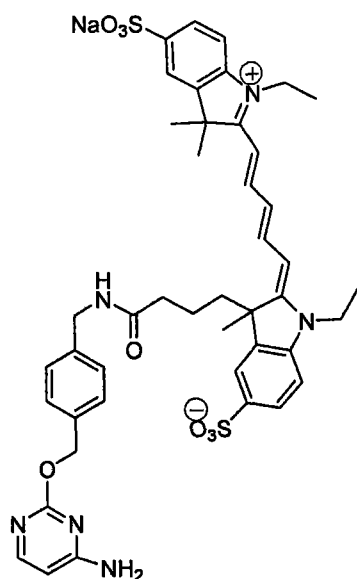
将 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) (3.9 mg, 0.017 mmol)和 5(6)-羧基荧光素琥珀酰亚胺酯 (8.3 mg, 0.017 mmol) 与 2.6 μ l TEA 一起溶解于 800 μ l DMF 中并且在 31°C 下加热过夜。在真空中蒸发溶剂并且产物利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上通过反相 MPLC 进行分离。MS (ESI) m/z 589 $[M+H]^+$ 。取决于纯化方法, 产物也可含有相应的异构体 5-甲酰胺。

[86] 实施例 5: N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-二乙酰基荧光素-6-甲酰胺 (6) 和 N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-二乙酰基荧光素-5-甲酰胺(7)



将 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) (4.1 mg, 0.018 mmol)和 5(6)-羧基荧光素二乙酸酯琥珀酰亚胺酯 (10 mg, 0.018 mmol) 与 2.7 μ l TEA 一起溶解于 800 μ l DMF 中并且在 31°C 下加热过夜。在真空中蒸发溶剂并且化合物利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上通过反相 MPLC 进行分离。MS (ESI) m/z 673 $[M+H]^+$ 。

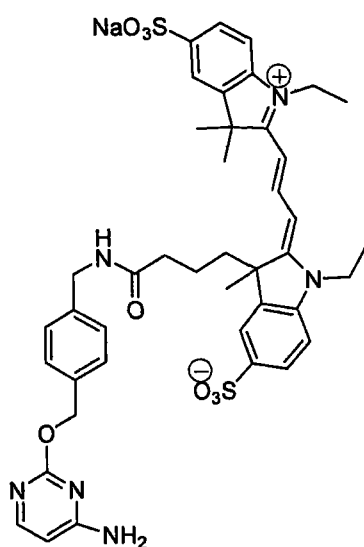
[87] 实施例 6: N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-DY647-甲酰

胺 (8)

(8)

将 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) (1.5 mg, 0.007 mmol)和 DY-547-NHS (Dyomics 染料) (5 mg, 0.007 mmol)与 1.0 μ l TEA 一起溶解于 500 μ l DMF 中并且在 31 $^{\circ}$ C 下加热过夜。在真空中蒸发溶剂并且产物利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上通过反相 MPLC 进行分离。MS (ESI) m/z 853 [M-Na]⁻。

[88] 实施例 7: N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-DY547-甲酰胺 (9)

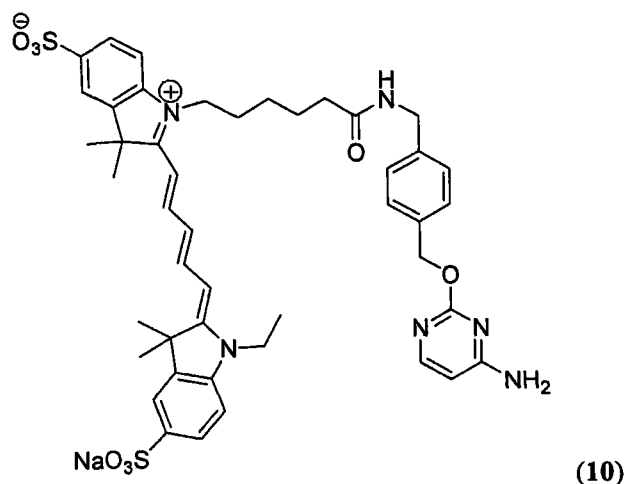


(9)

将 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) (1.5 mg, 0.007 mmol)和 DY-547-NHS (Dyomics 染料) (5 mg, 0.007 mmol)与 1.0 μ l TEA 一起溶解于 500 μ l DMF 中并且在 31 $^{\circ}$ C 下加热过夜。在真空中蒸发溶剂并且产物

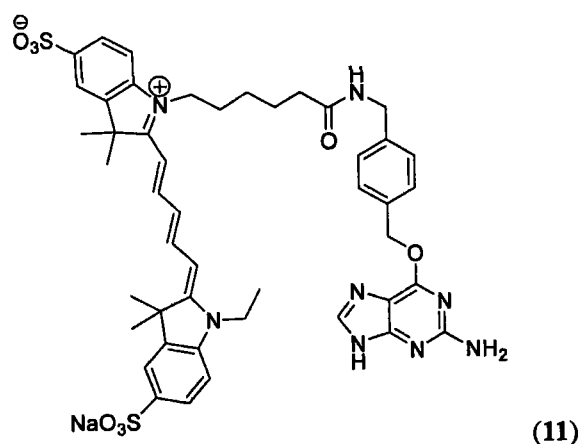
利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上通过反相 MPLC 进行分离。MS (ESI) m/z 827 $[M-Na]^+$ 。

[89] 实施例 8: N-[4-(4-氨基嘧啶-2-基氧甲基)-苄基]-Cy5-甲酰胺(10), BCCy5



将 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) (1.5 mg, 0.007 mmol)和 Cy5-NHS (GE 健康护理染料 (healthcare dye)) (5.4 mg, 0.007 mmol)与 1.0 μ l TEA 一起溶解于 500 μ l DMF 中并且在 31 $^{\circ}$ C 下加热过夜。在真空中蒸发溶剂并且产物利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上通过反相 MPLC 进行分离。MS (ESI) m/z 868 $[M-Na]^+$ 。

[90] 实施例 9: O⁶-[4-(Cy5)-氨甲基]-苄基鸟嘌呤(11), BGCy5



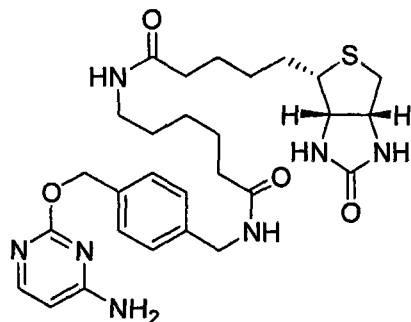
将 O⁶-4-氨甲基-苄基鸟嘌呤(1.9 mg, 0.007 mmol)和 Cy5-NHS (GE 健康护理染料) (5.4 mg, 0.007 mmol)与 1.0 μ l TEA 一起溶解于 500 μ l DMF 中并且在 31 $^{\circ}$ C 下加热过夜。在真空中蒸发溶剂并且产物利用水:乙腈的线性梯度(在 20 分钟内从 95:5 至 20:80, 0.08% TFA)在 C18 柱上

通过反相 MPLC 进行分离。MS (ESI) m/z 908 $[M-Na]^+$ 。

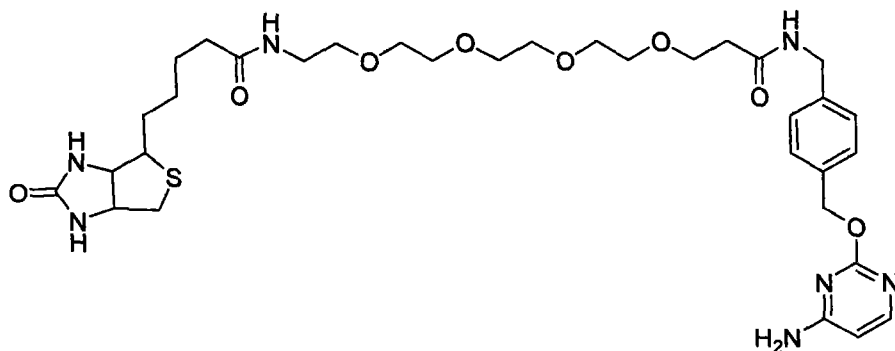
[91] 实施例 10:

下列化合物类似于实施例 3 通过 2-(4-(氨甲基)苄氧基)嘧啶-4-胺(2) 进行制备:

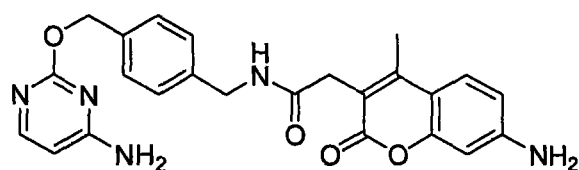
式 12 的 BC-生物素:



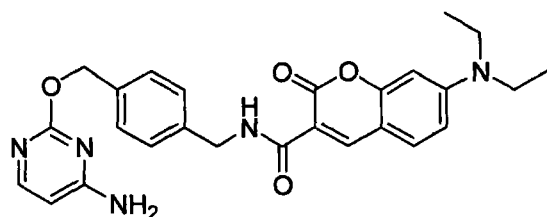
式 13 的 BC-PEG-生物素:



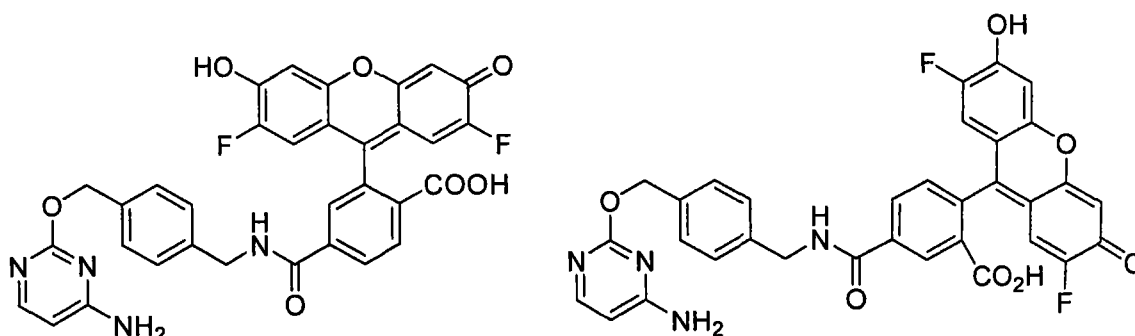
式 14 的 BC-360:



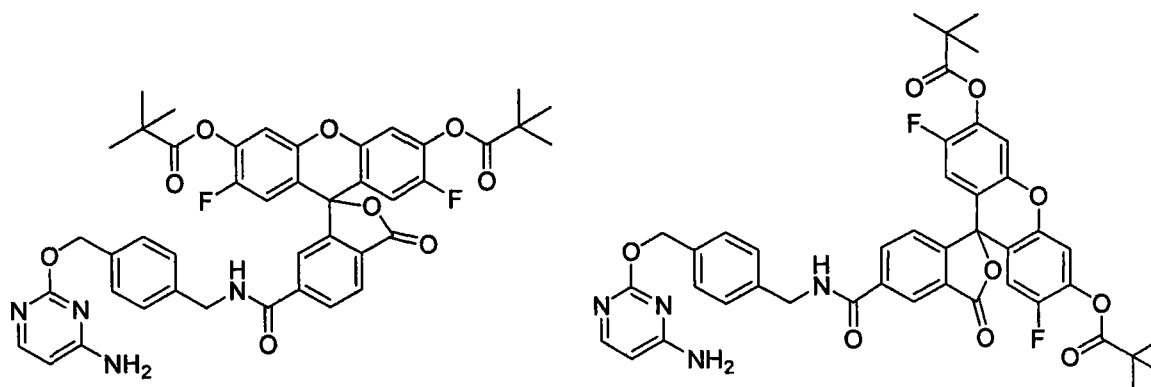
式 15 的 BC-430:



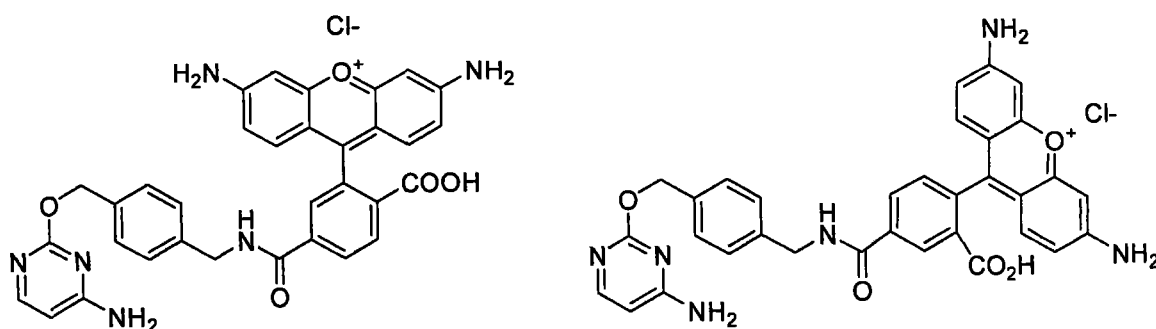
由式 16 化合物的混合物组成的 BC-俄勒岗绿 (Oregon Green):



由式 17 化合物的混合物组成的 BC-俄勒岗绿二新戊酰基酯:



由式 18 的化合物的混合物组成的 BC-505:



[92] 实施例 11: 文库构建和噬菌体选择

利用引物 1-5 和作为模板的^NAGT 基因 (SEQ ID NO:2) 的重叠序列延伸聚合酶链反应 (PCR) 允许随机化^NAGT 的残基 114、131、135、148、156、157 和 159。^NAGT 是^{wt}AGT 的 182 个氨基酸突变体, 其中后 25 个氨基酸被删除并且其具有突变 K32I、L33F、C62A、Q115S、Q116H、K125A、A127T、R128A、G131K、G132T、M134L、R135S、N150Q、I151G、N152D、G153L、A154D、N157G 和 S159E (Gronemeyer 等, Protein Eng. Des. Sel. 19:309-316, 2006)。引物 1 (N, AGT, SEQ ID NO:3) 和引物 2 (C, AGT, SEQ ID NO:4) 含有 *Sfi*I 限制性位点; 引物 3 (SEQ ID NO:5) 包含随机碱基, 用于在位置 114 上进行随机化; 引物 4

(SEQ ID NO:6) 包含随机碱基,用于位置 131 和 135 上进行随机化;引物 5 (SEQ ID NO:7) 包含随机碱基,用于在位置 148、156、157 和 159 上进行随机化。PCR 产物被连接到噬菌体展示载体 pAK100 并且所得构建物被电穿孔入大肠杆菌 (*E. coli*) XL1-蓝(Stratagene, USA)。这产生包含约 10^7 个克隆的文库。文库细胞在 37°C 下在 2YT 培养基 (25 $\mu\text{g/mL}$ 氯霉素, 1%葡萄糖, 1mM MgCl_2) 中生长直至光密度 OD_{600} 达到 0.6。然后将 2×10^{10} VCS M13 辅助噬菌体加入并且该培养物在无摇动下在 37°C 下温育 30 分钟并且在 170 rpm 下在 37°C 下温育 3 小时(h)。细胞通过离心 (4000 rpm, 5 min, RT) 进行收获, 重悬于 SB-MOPS (50 mM 3-吗啉丙磺酸, 25 $\mu\text{g/mL}$ 氯霉素, 70 $\mu\text{g/mL}$ 卡那霉素, 1 mM MgCl_2) 并且在 220 rpm 下在 37°C 下温育 1-4 小时, 然后在 220 rpm 下在 24°C 下温育过夜。细胞被沉淀并且含有噬菌体的上清液被调节至 1 mM DTT 并且在选择前储存于 4°C 下。对于选择来说, 将 BCFL (化合物 5) 加入到噬菌体溶液 (1 mL) 中直至终浓度 5 μM , 在室温下缓慢旋转 30 分钟。反应通过加入 8 μM BC 和 200 μM BG 进行猝灭。使用聚乙二醇 8000 (4% w/v) 和 NaCl (3% w/v) 沉淀噬菌体, 在台式离心机中进行离心 (13000 rpm, 4°C) 并且重悬于 500 μL PBS 中。向该溶液加入 500 μL PBSMM (含 4%脱脂奶粉的 PBS) 并且在室温下缓慢旋转该溶液 60 分钟。将 200 μL 覆盖有抗荧光素抗体的磁珠 (用 PBS 洗涤两次并且用 PBSMM 封闭 60 分钟) 加入到噬菌体制剂中并且在 4°C 下旋转 30 分钟。在标记噬菌体固定后, 该珠用 PBSMM 洗涤 3 次, 用 PBST (含 0.05% 吐温-20 的 PBS) 洗涤 5 次, 用 PBS 洗涤 2 次。通过用 100 μL 0.1 M 甘氨酸, pH 2.5 温育该珠 5 分钟来洗脱噬菌体, 并且该溶液用 50 μL 1 M Tris-HCl pH 8 进行中和。大肠杆菌 JM101 用洗脱的噬菌体进行感染, 在添加有 1%葡萄糖和 25 $\mu\text{g/mL}$ 氯霉素的 2YT 板上进行铺板, 然后在 37°C 下温育过夜。第二天, 从板上刮下克隆, 等分并且在下一轮选择前储存在 -80°C 下。进行六轮选择。

[93] 实施例 12: 选择的 ACT 作为 GST 融合蛋白的表征

根据实施例 9 的噬菌体选择之后分离的突变体的基因利用分别含有 *Bam*H1 和 *Eco*R1 限制性位点的引物 6 (N, AGT, SEQ ID NO:8) 和 7 (C, AGT, SEQ ID NO:9) 进行 PCR 扩增, 所述限制性位点用于随后亚克隆

到 pGEX-2T 载体 (Amersham Biosciences, Otelfingen, 瑞士)中。作为 GST-ACT 融合蛋白的蛋白质的表达和纯化如以前关于 AGT 融合蛋白的描述进行 (A. Juillerat, T. Gronemeyer, A. Keppler, S. Grendreizig, H. Pick, H. Vogel, K. Johnsson, Chem. Biol. 10:313, 2003)。用 BCFL (化合物 5)和携带荧光素的 BG (BGFL, WO 02/08397)进行标记的反应速率通过在 24°C下在反应缓冲液(50 mM HEPES, pH 7.2, 1 mM DTT, 200 µg/mL BSA)中用合适的荧光底物(2-20 µM)温育相应的 ACT 蛋白(0.2-0.4 µM)进行测定。在不同的时间进行取样, 并且标记反应通过加入 4×SDS 缓冲液(8% SDS, 10% β-巯基乙醇, 240 mM Tris pH 6.8, 40% 甘油)并且在 95°C下温育 5 分钟进行猝灭。通过在 SDS-PAGE 凝胶中检测荧光染料标记蛋白以及利用 Pharox FX™分子成像仪量化荧光强度来确定反应进程。数据被拟合成准一级反应模型。然后通过将准一级常数除以荧光底物浓度而获得二级反应速率常数。

[94] 实施例 13: 体外特异性双重标记试验

体外双重标记试验如下进行: 在 24°C下在反应缓冲液(50 mM HEPES, pH 7.2, 1 mM DTT, 200 µg/mL BSA)中用 BGFL(携带荧光素的苄基鸟嘌呤, WO 02/08397)、BGCy5(携带荧光染料 Cy5 的苄基鸟嘌呤, 化合物 11, 实施例 9)、BCFL(化合物 5, 实施例 4)、BCCy5 (化合物 10, 实施例 8) (5 µM)或者 BGCy5/BCFL 或 BGFL/BCCy5 的等摩尔混合物 (各 5 µM) 温育 GST-ACT1 和 6xHis-^NAGT 的等摩尔混合物(终浓度 0.5 µM) 60 分钟。标记反应通过加入 4×SDS 缓冲液(8% SDS, 10% β-巯基乙醇, 240 mM Tris pH 6.8, 40%甘油)并且在 95°C下温育 5 分钟进行猝灭。荧光染料标记蛋白混合物如上所述通过 SDS-PAGE 进行分析。

[95] 实施例 14: 哺乳动物细胞表达载体的构建

对于融合蛋白 ^NAGT-NLS3、^NAGT-βGal、ACT1-NLS3 和 ACT1 βGal 在哺乳动物细胞中的表达, ^NAGT 和 ACT1 基因利用引物 8 (N, AGT, SEQ ID NO:10)和 9 (C, AGT, SEQ ID NO:11)进行 PCR 扩增并且被插入到哺乳动物表达载体 pECFP-Nuc (Clontech)或含有 β-半乳糖苷酶基因的哺乳动物表达质粒的 *NheI*/*Bgl* II 限制性位点。

<110> 洛桑联邦理工学院

<120> 用合成探针标记融合蛋白

<130> P343A

<150> EP06117779.6

<151> 2006-07-25

<160> 12

<170> PatentIn version 3.3

<210> 1

<211> 182

<212> PRT

<213> 人工序列

<220>

<223> 衍生自06-烷基鸟嘌呤-DNA烷基转移酶的烷基胞嘧啶转移酶1

<400> 1

Met Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp Ser Pro Leu
1 5 10 15

Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His Glu Ile Ile
20 25 30

Phe Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu Val Pro Ala
35 40 45

Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln Ala Thr Ala
50 55 60

Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu Glu Phe Pro
65 70 75 80

Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser Phe Thr Arg
85 90 95

Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly Glu Val Ile
100 105 110

Ser Arg Ser His Leu Ala Ala Leu Ala Gly Asn Pro Ala Ala Thr Ala
115 120 125

Ala Val Ser Thr Ala Leu Asp Gly Asn Pro Val Pro Ile Leu Ile Pro
130 135 140

Cys His Arg Glu Val Gln Gly Asp Leu Asp Val Gly Gly Tyr Met Gly
145 150 155 160

Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly His Arg Leu
165 170 175

Gly Lys Pro Gly Leu Gly
180

<210> 2
<211> 182
<212> PRT
<213> 人工序列

<220>
<223> 06-烷基鸟嘌呤-DNA烷基转移酶 (AGT) 突变体
<400> 2

Met Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp Ser Pro Leu
1 5 10 15

Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His Glu Ile Ile
20 25 30

Phe Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu Val Pro Ala
35 40 45

Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln Ala Thr Ala
50 55 60

Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu Glu Phe Pro
65 70 75 80

Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser Phe Thr Arg
85 90 95

Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly Glu Val Ile
100 105 110

Ser Tyr Ser His Leu Ala Ala Leu Ala Gly Asn Pro Ala Ala Thr Ala
115 120 125

Ala Val Lys Thr Ala Leu Ser Gly Asn Pro Val Pro Ile Leu Ile Pro
130 135 140

Cys His Arg Val Val Gln Gly Asp Leu Asp Val Gly Gly Tyr Glu Gly
145 150 155 160

Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly His Arg Leu
165 170 175

Gly Lys Pro Gly Leu Gly
180

<210> 3
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 具有SfiI限制性位点的N, AGT

<400> 3	
ctactcgcgg cccagccggc catggcggac tacaaagaca tggacaagga ttgtgaaatg	60
<210> 4	
<211> 45	
<212> DNA	
<213> 人工序列	
<220>	
<223> 具有SfiI限制性位点的C, AGT	
<400> 4	
gcaatggaat tcggccccg aggccgctcc caagcctggc ttccc	45
<210> 5	
<211> 66	
<212> DNA	
<213> 人工序列	
<220>	
<223> 用于在位置114上随机化	
<220>	
<221> misc_feature	
<222> (50)..(51)	
<223> n 是a、c、g或 t	
<400> 5	
cactgctgcc gtggctgcgg ggttgccctgc cagggtgct aagtgagamn nagaaatcac	60
ttctcc	66
<210> 6	
<211> 69	
<212> DNA	
<213> 人工序列	
<220>	
<223> 用于在位置131和135上随机化	
<220>	
<221> misc_feature	
<222> (19)..(20)	
<223> n 是a、c、g或 t	
<220>	
<221> misc_feature	
<222> (43)..(44)	
<223> n 是a、c、g或 t	
<220>	
<221> misc_feature	
<222> (46)..(47)	
<223> n 是a、c、g或 t	
<220>	
<221> misc_feature	
<222> (52)..(53)	
<223> n 是a、c、g或 t	
<400> 6	
ctgatcccg gccacagann kgtccagggg gatttgatg tgnnknnkta cnnkggtgga	60

ctggccgtg	69
<210> 7	
<211> 69	
<212> DNA	
<213> 人工序列	
<220>	
<223> 用于在位置148、156、157 和159上随机化	
<220>	
<221> misc_feature	
<222> (19)..(20)	
<223> n 是a、c、g或 t	
<220>	
<221> misc_feature	
<222> (43)..(44)	
<223> n 是a、c、g或 t	
<220>	
<221> misc_feature	
<222> (46)..(47)	
<223> n是a, c, g 或 t	
<220>	
<221> misc_feature	
<222> (52)..(53)	
<223> n 是a、c、g或 t	
<400> 7	
ctgatcccggt gccacagann kgtccagggg gatttggatg tgnnknnkta cnnkggtgga	60
ctggccgtg	69
<210> 8	
<211> 33	
<212> DNA	
<213> 人工序列	
<220>	
<223> 具有限制性位点BamH1和EcoR1的N, AGT	
<400> 8	
cgaaatggat ccatggacaa ggattgtgaa atg	33
<210> 9	
<211> 30	
<212> DNA	
<213> 人工序列	
<220>	
<223> 具有限制性位点BamH1和EcoR1的C, AGT	
<400> 9	
gcctttgaat tccttatccc aagcctggct	30
<210> 10	
<211> 49	
<212> DNA	
<213> 人工序列	
<220>	
<223> 具有限制性位点NheI/BglIII的N, AGT	

<400> 10
gatcgagcta gcgctaccgg tcgccaccat ggacaaggat tgtgaaatg 49

<210> 11
<211> 48
<212> DNA
<213> 人工序列

<220>
<223> 具有限制性位点NheI/BglIII的C, AGT

<400> 11
ccaggcagat cttcccaagc ctggcttccc caaccggtgg ccttcattg 48

<210> 12
<211> 182
<212> PRT
<213> 人工序列

<220>
<223> 衍生自06-烷基鸟嘌呤-DNA烷基转移酶的烷基胞嘧啶转移酶10

<400> 12

Met	Asp	Lys	Asp	Cys	Glu	Met	Lys	Arg	Thr	Thr	Leu	Asp	Ser	Pro	Leu
1				5					10					15	
Gly	Lys	Leu	Glu	Leu	Ser	Gly	Cys	Glu	Gln	Gly	Leu	His	Glu	Ile	Ile
		20					25						30		
Phe	Leu	Gly	Lys	Gly	Thr	Ser	Ala	Ala	Asp	Ala	Val	Glu	Val	Pro	Ala
		35					40					45			
Pro	Ala	Ala	Val	Leu	Gly	Gly	Pro	Glu	Pro	Leu	Ile	Gln	Ala	Thr	Ala
	50				55					60					
Trp	Leu	Asn	Ala	Tyr	Phe	His	Gln	Pro	Glu	Ala	Ile	Glu	Glu	Phe	Pro
65				70					75					80	
Val	Pro	Ala	Leu	His	His	Pro	Val	Phe	Gln	Gln	Glu	Ser	Phe	Thr	Arg
			85					90						95	
Gln	Val	Leu	Trp	Lys	Leu	Leu	Lys	Val	Val	Lys	Phe	Gly	Glu	Val	Ile
		100					105					110			
Ser	Glu	Ser	His	Leu	Ala	Ala	Leu	Val	Gly	Asn	Pro	Ala	Ala	Thr	Ala
		115					120					125			
Ala	Val	Asn	Thr	Ala	Leu	Asp	Gly	Asn	Pro	Val	Pro	Ile	Leu	Ile	Pro
	130					135					140				
Cys	His	Arg	Val	Val	Gln	Gly	Asp	Ser	Asp	Val	Gly	Pro	Tyr	Leu	Gly
145					150					155				160	
Gly	Leu	Ala	Val	Lys	Glu	Trp	Leu	Leu	Ala	His	Glu	Gly	His	Arg	Leu
				165					170					175	

Gly Lys Pro Gly Leu Gly
180

GST-ACT1	-	-	+	+	+	+	+	+	+	+	+
6xHis-AGT	+	+	-	-	+	+	+	+	+	+	+
BGFL	+	-	-	-	-	+	+	-	-	-	-
BGCy5	-	+	-	-	+	-	-	+	-	-	-
BCFL	-	-	+	-	+	-	-	-	+	-	-
BCCy5	-	-	-	+	-	+	-	-	-	+	-



图 1

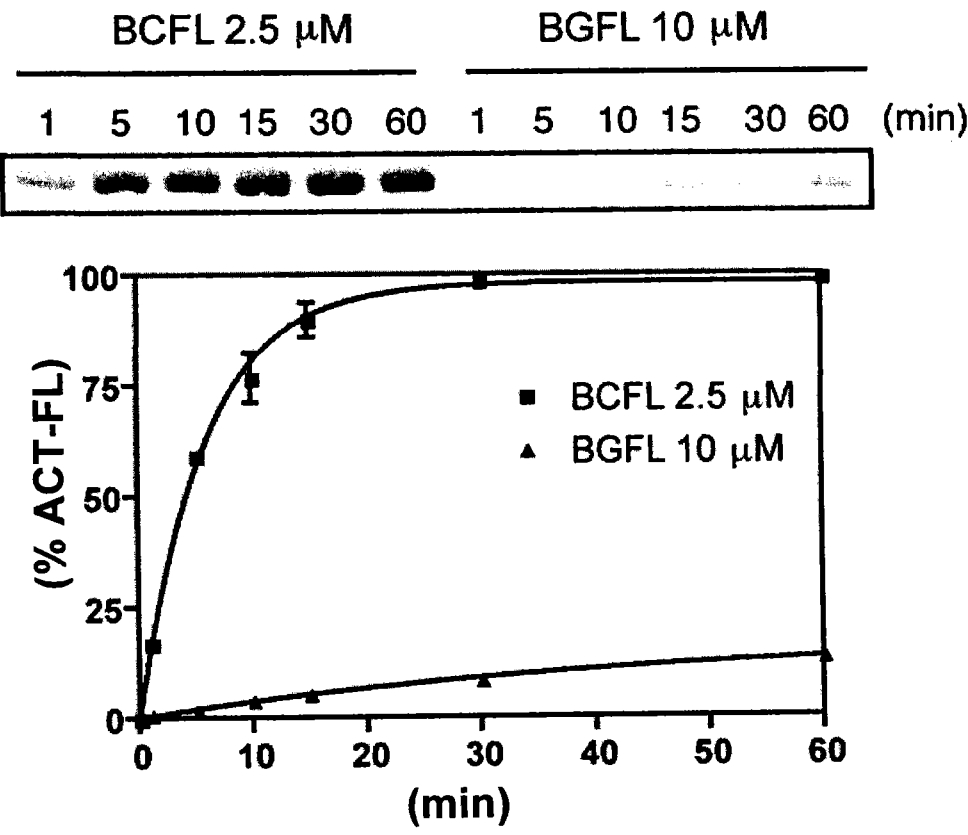


图 2

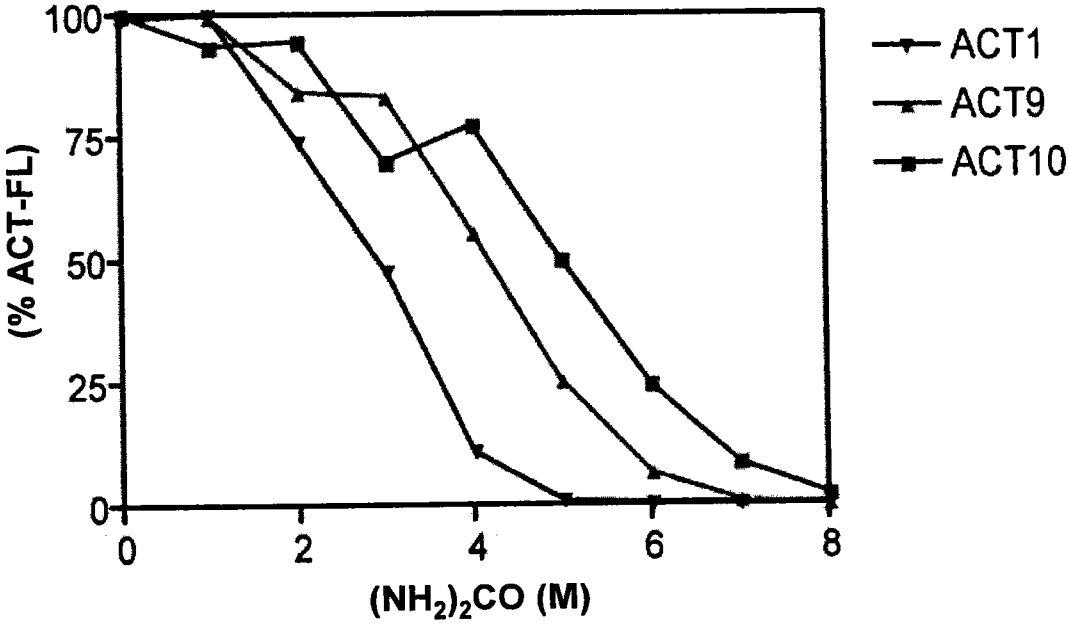


图 3