(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 104822837 B
(45)授权公告日 2018.06.05

(21)申请号 201380044460.8
(22)申请日 2013.08.20
(65)同一申请的已公布的文献号
申请公布号 CN 104822837 A
(43)申请公布日 2015.08.05
(30)优先权数据
61/684,974 2012.08.20 US
(85)PCT国际申请进入国家阶段日
2015.02.25
(86)PCT国际申请的申请数据
PCT/US2013/055731 2013.08.20
(87)PCT国际申请的公布数据
WO2014/031602 EN 2014.02.27

(73)专利权人 中央研究院
地址 中国台湾台北市南港区研究院路二段128号
(72)发明人 翁启惠 吴宗益 蔡宗义
(74)专利代理机构 上海专利商标事务所有限公司 31100
代理人 陈静

(51)Int.Cl.
C12P 19/04(2006.01)
C12N 9/10(2006.01)

审查员 张娜

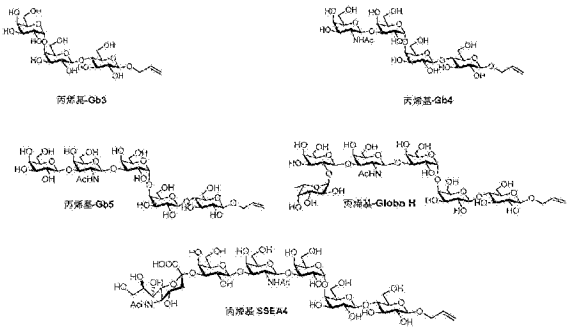
权利要求书4页 说明书62页
序列表7页 附图13页

(54)发明名称

寡糖的大规模酶合成的方法

(57)摘要

本发明提供一种UDP-Gal再生的方法及其与半乳糖基转移酶的组合使用，将半乳糖添加至适当的受体底物。本发明进一步揭露大规模生产Globo系列寡糖的合成方法，其中该方法可涉及糖基转移酶反应与核苷酸糖再生方法的组合。



CN 104822837 B

1.一种以酶性方式合成寡糖的方法,其特征在于,包含:

(i)在UDP-Gal再生酶组的存在下,将半乳糖制成UDP-Gal,其中该UDP-Gal再生酶组包含半乳糖激酶、UDP-糖焦磷酸化酶及丙酮酸激酶;及

(ii)在UDP-Gal及α-1,4-半乳糖转移酶存在下,将乳糖-OR^{1A}(Lac-OR^{1A})转换为Globo三糖-OR^{1A}(Gb3-OR^{1A}),其中R^{1A}为氢、经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经取代的炔基、经取代或未经取代的环碳基、经取代或未经取代的杂环基、经取代或未经取代的芳基、经取代或未经取代的杂芳基或氧保护基团、生物素或神经酰胺。

2.如权利要求1所述的方法,其特征在于,该UDP-Gal再生酶组包含焦磷酸化酶。

3.如权利要求1所述的方法,其特征在于,(i)及(ii)发生于Gb3-合成反应混合物中,该Gb3-合成反应混合物包括半乳糖、PEP、ATP、UTP、Lac-OR^{1A}、α-1,4-半乳糖转移酶及UDP-Gal再生酶组。

4.如权利要求3所述的方法,其特征在于,在任何酶反应发生前,Lac-OR^{1A}与半乳糖在Gb3-合成反应混合物中的摩尔比为1:1。

5.如权利要求2所述的方法,其特征在于,该α-1,4-半乳糖转移酶是来自脑膜炎双球菌的LgtC,该半乳糖激酶是来自大肠杆菌,该UDP-糖焦磷酸化酶是来自拟南芥,该丙酮酸激酶是来自大肠杆菌,或该焦磷酸化酶是来自大肠杆菌。

6.如权利要求1所述的方法,其特征在于,进一步包含分离Gb3-OR^{1A}。

7.如权利要求1所述的方法,其特征在于,进一步包含:

(iii)在UDP-GalNAc及β-1,3-N-乙酰半乳糖胺转移酶存在下,将Gb3-OR^{1A}转换成Globo四糖-OR^{1A}(Gb4-OR^{1A})。

8.如权利要求7所述的方法,其特征在于,进一步包含:

(iv)在UDP-GalNAc再生酶组存在下,将GalNAc制成UDP-GalNAc,其中该UDP-GalNAc再生酶组包含N-乙酰己糖胺1-激酶、N-乙酰己糖胺1-磷酸尿苷酰转移酶及丙酮酸激酶。

9.如权利要求8所述的方法,其特征在于,所述的UDP-GalNAc再生酶组还包含焦磷酸化酶。

10.如权利要求8所述的方法,其特征在于,(iii)及(iv)在Gb-4合成反应混合物中进行,该Gb-4合成反应混合物包含GalNAc、PEP、ATP、UTP、Gb3-OR^{1A}、β-1,3-N-乙酰半乳糖胺转移酶及UDP-GalNAc再生酶组。

11.如权利要求10所述的方法,其特征在于,该Gb-4合成反应混合物是通过将Gb-3合成反应混合物至少与GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶混合制得。

12.如权利要求9所述的方法,其特征在于,该β-1,3-N-乙酰半乳糖胺转移酶是来自流感嗜血杆菌的LgtD,N-乙酰己糖胺1-激酶是来自龙根菌,N-乙酰己糖胺1-磷酸尿苷酰转移酶是来自大肠杆菌,该丙酮酸激酶是来自大肠杆菌,或该焦磷酸化酶是来自大肠杆菌。

13.如权利要求7所述的方法,其特征在于,进一步包含分离Gb4-OR^{1A}。

14.如权利要求7所述的方法,其特征在于,进一步包含:

(v)在UDP-Gal及β-1,3-半乳糖转移酶存在下,将Gb4-OR^{1A}转换为Globo五糖-OR^{1A}(Gb5-OR^{1A})。

15.如权利要求14所述的方法,其特征在于,(v)是在Gb5-合成反应混合物中进行,该

Gb5-合成反应混合物包含半乳糖、PEP、ATP、UTP、Gb4-OR^{1A}、β-1,3-半乳糖转移酶及UDP-Gal再生酶组。

16.如权利要求14所述的方法,其特征在于,该β-1,3-半乳糖转移酶是来自流感嗜血杆菌的LgtD。

17.如权利要求14所述的方法,其特征在于,进一步包含分离Gb5-OR^{1A}。

18.如权利要求14所述的方法,其特征在于,进一步包含:

(vi)在GDP-Fuc及α-1,2-岩藻糖转移酶存在下,将Gb5-OR^{1A}转换为岩藻糖-Gb5-OR^{1A}。

19.如权利要求18所述的方法,其特征在于,进一步包含:

(vii)在GDP-Fuc再生酶组存在下,自岩藻糖生成GDP-Fuc,其中该GDP-Fuc再生酶组包含L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶、丙酮酸激酶。

20.如权利要求19所述的方法,其特征在于,所述的GDP-Fuc再生酶组还包含焦磷酸化酶。

21.如权利要求19所述的方法,其特征在于,(vi)及(vii)是在岩藻糖-Gb5-合成反应混合物中进行,该岩藻糖-Gb5-合成反应混合物包含岩藻糖、ATP、GTP、PEP、Gb5-OR^{1A}、α-1,2-岩藻糖转移酶及GDP-Fuc再生酶组。

22.如权利要求21所述的方法,其特征在于,该岩藻糖-Gb5-合成反应混合物是将Gb5-合成反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合而得。

23.如权利要求20所述的方法,其特征在于,该L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶是来自脆弱拟杆菌,该α-1,2-岩藻糖转移酶是来自幽门螺旋杆菌,该丙酮酸激酶是来自大肠杆菌,或该焦磷酸化酶是来自大肠杆菌。

24.如权利要求18所述的方法,其特征在于,进一步包含分离岩藻糖-Gb5-OR^{1A}。

25.如权利要求14所述的方法,其特征在于,进一步包含:

(viii)在CMP-Neu5Ac及α-2,3-唾液酸转移酶存在下,将Gb5-OR^{1A}转换为唾液酸-Gb5-OR^{1A}。

26.如权利要求25所述的方法,其特征在于,进一步包含:

(ix)在CMP-Neu5Ac再生酶组存在下,自Neu5Ac生成CMP-Neu5Ac,其中该CMP-Neu5Ac再生酶组包含胞苷单磷酸激酶、CMP-唾液酸合成酶、丙酮酸激酶。

27.如权利要求26所述的方法,其特征在于,所述的CMP-Neu5Ac再生酶组还包含:焦磷酸化酶。

28.如权利要求26所述的方法,其特征在于,(viii)及(ix)是在唾液酸-Gb5-合成反应混合物中进行,该唾液酸-Gb5-合成反应混合物包含Neu5Ac、CTP、PEP、Gb5-OR^{1A}、α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组。

29.如权利要求28所述的方法,其特征在于,该唾液酸-Gb5-合成反应混合物是将Gb5-合成反应混合物与至少Neu5Ac、CTP、α-2,3-唾液酸转移酶、胞苷单磷酸激酶及CMP-唾液酸合成酶混合而得。

30.如权利要求27所述的方法,其特征在于,该α-2,3-唾液酸转移酶是来自海洋细菌,该胞苷单磷酸激酶是来自大肠杆菌,该CMP-唾液酸合成酶是来自多杀性巴氏杆菌,该丙酮酸激酶是来自大肠杆菌,或该焦磷酸化酶是来自大肠杆菌。

31.如权利要求25所述的方法，其特征在于，进一步包含分离唾液酸-Gb5-OR^{1A}。

32.如权利要求3所述的方法，其特征在于，进一步包含：

(a)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶予以混合以形成Gb4-合成反应混合物；

(b)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；

(c)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，在β-1,3-半乳糖转移酶存在下，进一步培养Gb4-合成反应混合物；

(d)将(c)中所制备含有Gb5-OR^{1A}的反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合来形成岩藻糖-Gb5-OR^{1A}反应混合物；及

(e)在允许转换Gb5-OR^{1A}为岩藻糖-Gb5-OR^{1A}的条件下，培养岩藻糖-Gb5-OR^{1A}反应混合物。

33.如权利要求3所述的方法，其特征在于，进一步包含：

(a)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶混合来形成Gb4-合成反应混合物；

(b)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；

(c)分离Gb4-OR^{1A}；

(d)将Gb4-OR^{1A}与β-1,3-半乳糖转移酶及UDP-Gal再生酶组混合来形成Gb5-合成反应混合物；

(e)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，培养Gb5-合成反应混合物；

(f)将Gb5-合成反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合以形成岩藻糖-Gb5-OR^{1A}反应混合物；及

(g)在允许转换Gb5-OR^{1A}为岩藻糖-Gb5-OR^{1A}的条件下，培养岩藻糖-Gb5-OR^{1A}反应混合物。

34.如权利要求3所述的方法，其特征在于，进一步包含：

(a)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶予以混合以形成Gb4-合成反应混合物；

(b)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；

(c)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，在β-1,3-半乳糖转移酶存在下，进一步培养Gb4-合成反应混合物；

(d)将(c)中所制备含有Gb5-OR^{1A}的反应混合物与至少Neu5Ac、CTP、α-2,3-唾液酸转移酶、胞苷单磷酸激酶及CMP-唾液酸合成酶予以混合以形成唾液酸-Gb5-OR^{1A}反应混合物；及

(e)在允许转换Gb5-OR^{1A}为唾液酸-Gb5-OR^{1A}的条件下，培养唾液酸-Gb5-OR^{1A}反应混合物。

35.如权利要求3所述的方法，其特征在于，进一步包含：

(a)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶予以混合以形成Gb4-合成反应混合物；

(b)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；

(c)分离Gb4-OR^{1A}；

(d)将Gb4-OR^{1A}与β-1,3-半乳糖转移酶及UDP-Gal再生酶组混合以形成Gb5-合成反应混

合物；

(e)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，培养Gb5-合成反应混合物；

(f)将Gb5-OR^{1A}与α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组混合以形成唾液酸-Gb5-合成反应混合物，其中该CMP-Neu5Ac再生酶组包含胞苷单磷酸激酶、CMP-唾液酸合成酶、丙酮酸激酶；及

(g)在允许转换Gb4-OR^{1A}为唾液酸-Gb5-OR^{1A}的条件下，培养唾液酸-Gb5-合成反应混合物。

36.如权利要求35所述的方法，其特征在于，(f)中，CMP-Neu5Ac再生酶组还包含：焦磷酸化酶。

37.一种将半乳糖残基添加至底物中的方法，其特征在于，包含：

(i)在UDP-Gal再生酶组存在下，自半乳糖生成UDP-Gal，其中该UDP-Gal再生酶组包含半乳糖激酶、UDP-糖焦磷酸化酶及丙酮酸激酶；及

(ii)将UDP-Gal与底物分子通过半乳糖转移酶进行反应，在该底物分子上添加半乳糖残基。

38.如权利要求37所述的方法，其特征在于，该UDP-Gal再生酶组进一步包含焦磷酸化酶。

39.如权利要求37所述的方法，其特征在于，(i)及(ii)可在包含该UDP-Gal再生酶组、该半乳糖转移酶、该底物分子、半乳糖、ATP及UTP的反应混合物中进行。

40.如权利要求37所述的方法，其特征在于，该底物分子为多糖、寡糖、糖蛋白、糖脂、配糖基、神经酰胺或鞘糖脂。

41.如权利要求37所述的方法，其特征在于，该半乳糖转移酶为α-1,4-半乳糖转移酶、β-1,4-半乳糖转移酶、α-1,3-半乳糖转移酶或β-1,3-半乳糖转移酶。

寡糖的大规模酶合成的方法

技术领域

[0001] 本发明提供一种UDP-Gal再生的方法及其与半乳糖基转移酶的组合使用，将半乳糖添加至适当的受体底物。本发明进一步揭露大规模生产Globo系列寡糖的合成方法，其中该方法可涉及糖基转移酶反应与核苷酸糖再生方法的组合。

背景技术

[0002] Globo五糖(Globopentaose(Gb5))、岩藻糖-Gb5(Globo H)及唾液酸-Gb5(SSEA4)是Globo系列的鞘糖脂(glycosphingolipid)，其是于1983年首次在经培养人类畸胎癌细胞株中被发现[1]，其后陆续在数种恶性肿瘤中被发现[2,3]。报告显示，在乳癌患者身上Globo H的过度表达高达61％，Gb5的过度表达达77.5％及SSEA4的过度表达达95％。[4]另一方面，基因HER2(治疗性单株抗体Trastuzumab(Herceptin)的标靶，会干扰HER2/neu受体)在乳癌患者身上的过度表达仅达25％。[5]此比较清楚显示鞘糖脂抗原(Gb5及其衍生物，Globo H及SSEA4)乃作为发展癌症疫苗的更佳选择。因此，目前在一些国家已将Globo H共轭至血蓝蛋白(KLH)作为癌症疫苗，并已进入第三期临床试验。[6]

[0003] 现行用来合成Gb5、Globo H及SSEA4的方法存在一些缺点。传统化学合成法过于繁琐且费力，需要一些保护及去保护步骤才能获得高纯度及正确的型式，因此往往导致产量极低。迄今虽有诸多关于化学合成Globo H的相关报告。[7][8][9][10][11][12][13][14]然而，仅有两篇报告是关于SSEA4的合成。Hsu等人发表1锅法(one-pot)化学合成法来制备SSEA-4的多糖部分，其产率为24％[15]Zhen等人发表利用化学酶方法合成毫克规模的SSEA-4。[16]另一方面，基于Leloir-type糖基转移酶的酶性合成Globo H仅需将活化核苷酸糖为供体来催化糖化反应。然而，糖核苷酸成本过高，不适合大规模的合成。再者，副产物焦糖酸盐(pyrophosphate)及核苷二磷酸盐(nucleoside diphosphate)会抑制焦糖酸化酶及Leloir-type糖基转移酶在核苷酸糖的形成；因此，需要发展新的再生方法以克服现行技术的限制，并且可将核苷酸再充电来达到核苷酸糖产物的构建，进而使反应继续进行。过去数年间，有许多团队致力于解决糖核苷酸再生的主要问题，且大部分有关糖核苷酸再生的问题已获得解决。然而，关于糖核苷酸再生的技术仍有进步改善的空间，特别在于UDP-Gal的再生非常困难。例如，UDP-Gal的再生首次于1982年由Wong以及Whiteside所提出，通过UDP-Glc C4表异构酶(epimerase)来将UDP-Glc与UDP-Gal交互转换([17])。十年过后，本发明团队发展出二级UDP-Gal再生方法。其并不使用UDP-Glc C4表异构酶，而是利用位于大肠杆菌乳糖操纵子中的Glc-1-磷酸尿苷酰转移酶(uridylyltransferase)将Gal-1-磷酸盐及UDP-Glc交换为Glc-1-磷酸盐及UDP-Gal。[18]然而，在缺少适合的酶下，最后直接将UTP及Gal-1-磷酸盐缩合形成UDP-Gal的步骤确无法完成。由于目标化合物Gb5、Globo H及SSEA4为Gal有关分子，故如何克服UDP-Gal再生的主要困难并增加其效率即为大规模酶合成Gb5、Globo H及SSEA4的关键。

[0004] 综上所述，现行大规模合成Gb5、Globo H及SSEA4的方法仍存在诸多限制。因此，仍需要提供一种新颖的合成方法，以更有效率地制备Gb5、Globo H、SSEA4及其中间物。

发明内容

[0005] 本发明是关于一种核苷酸糖再生的方法及其于糖类合成的应用。该糖类合成方法涉及至少一种核苷酸糖再生系统(例如本文所描述的UDP-Gal再生系统)与至少一种糖基转移酶(例如半乳糖基转移酶)的组合,且该糖类合成方法以不可预期的高效率被用于合成各种寡糖(尾式(tailed)),包括丙烯基-尾式Gb3、Gb4、Gb5(也被称为SSEA3)、岩藻糖-Gb5(也被称为Globo H)及唾液酸-Gb5(也被称为SSEA4)。更具体的,本发明的合成方法允许链式反应生成诸如Globo H及SSEA4等终产物,而无纯化中间物的必要。

[0006] 因此,本发明的一个方面是关于一种将半乳糖残基通过连结半乳糖转移酶与UDP-Gal再生方法的反应添加至底物中的方法。该方法包括:(i)在UDP-Gal再生酶组存在下,自半乳糖生成UDP-Gal,其中该UDP-Gal再生酶组包含半乳糖激酶、UDP-糖焦磷酸化酶(UDP-sugar pyrophosphorylase)、丙酮酸激酶及视需要的焦磷酸化酶;(ii)将UDP-Gal与底物分子(例如多糖、寡糖、糖蛋白、糖脂或配糖基(aglycone))通过半乳糖转移酶(例如α-1,4-半乳糖转移酶、β-1,4-半乳糖转移酶、α-1,3-半乳糖转移酶或β-1,3-半乳糖转移酶)进行反应,在该底物分子上添加半乳糖残基;及视需要(iii)分离所生成的半乳糖化产物。步骤(i)及(ii)可在一包含该UDP-Gal再生酶组、该半乳糖转移酶、该底物分子、半乳糖、ATP及UTP的反应混合物中进行。在某些实施例中,该底物分子为神经酰胺或鞘糖脂。

[0007] 本发明的另一个方面是关于合成寡糖的方法,其是涉及至少一种核苷酸糖再生方法(例如UDP-Gal再生)及至少一种通过糖基转移酶(例如半乳糖转移酶、岩藻糖转移酶、唾液酸转移酶或N-乙酰半乳糖胺转移酶)的活性将单糖(例如半乳糖(Gal)、N-乙酰半乳糖胺(GalNAc)、岩藻糖(Fuc)及唾液酸(Neu5Ac))添加至合适的受体的反应。

[0008] 在某些实施例中,本发明方法是利用乳糖(例如尾式)作为起始物,以酶性方式合成寡糖。该方法包括:(i)在UDP-Gal再生酶组的存在下,自半乳糖生成UDP-Gal,其中该UDP-Gal再生酶组包含半乳糖激酶(例如来自大肠杆菌)、UDP-糖焦磷酸化酶(例如来自拟南芥(A.thaliana))、丙酮酸激酶(例如来自大肠杆菌)及视需要的焦磷酸化酶(例如来自大肠杆菌);(ii)在UDP-Gal及α-1,4-半乳糖转移酶(例如LgtC(诸如来自脑膜炎双球菌(N.meningitides))存在下,将Lac-OR^{1A}转换为Gb3-OR^{1A},其中R^{1A}为氢、经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经取代的炔基、经取代或未经取代的环碳基(carbocyclyl)、经取代或未经取代的杂环基、经取代或未经取代的芳基、经取代或未经取代的杂芳基或氧保护基团。

[0009] R^{1A}的实施例包括,但不限于氢、丙烯基、生物素、神经酰胺、或非氢基团(如烷基),其是进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如,羧酸)、迭氮、烯基(例如,丙烯基)、炔基(例如,丙炔基)、生物素或神经酰胺所取代。某些特定的实施例中,R^{1A}是氢、丙烯基、经取代的烷基、生物素或神经酰胺。

[0010] 如需要,Gb3-OR^{1A}可自反应混合物中分离。

[0011] 步骤(i)及(ii)可在Gb3-合成反应混合物中进行,该Gb3-合成反应混合物包括半乳糖、PEP、ATP、UTP、Lac-OR^{1A}、α-1,4-半乳糖转移酶及UDP-Gal再生酶组。一实施例中,在任何酶反应发生前,Lac-OR^{1A}与半乳糖在Gb3-合成反应混合物中的摩尔比为1:1。

[0012] 上述任一方法可进一步包括:(iii)在UDP-GalNAc及β-1,3-N-乙酰半乳糖胺转移

酶(例如LgtD,其是来自适当生物体如流感嗜血杆菌(H.influenza))存在下将Gb3-OR^{1A}转换成Gb4-OR^{1A},该步骤可与步骤(iv)结合,步骤(iv)包括在UDP-GalNAc再生酶组存在下,自GalNAc生成UDP-GalNAc,其中该UDP-GalNAc再生酶组包含N-乙酰己糖胺1-激酶(N-acetylhexosamine 1-kinase)(如来自龙根菌(B.longum))、N-乙酰己糖胺1-磷酸尿苷酰转移酶(例如来自大肠杆菌)及丙酮酸激酶(例如来自大肠杆菌),及视需要的焦磷酸化酶(例如来自大肠杆菌)。步骤(iii)及(iv)可在Gb-4合成反应混合物中进行,该反应混合物包含GalNAc、PEP、ATP、UTP、Gb3-OR^{1A}、β-1,3-N-乙酰半乳糖胺转移酶及UDP-GalNAc再生酶组。本发明的一实施例中,Gb-4合成反应混合物是通过将Gb-3合成反应混合物至少与GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶混合制得。如需要,Gb4-OR^{1A}可自反应混合物中分离。

[0013] Gb4-OR^{1A}合成之后,上述方法可进一步包括:(v)在UDP-Gal及β-1,3-半乳糖转移酶(例如LgtD,其是来自如流感嗜血杆菌)存在下,将Gb4-OR^{1A}转换为Gb5-OR^{1A},该步骤可与步骤(vi)结合,步骤(vi)包括在本发明所述UDP-Gal再生酶组存在下,自半乳糖生成UDP-Gal。本发明的一实施例中,步骤(v)及(vi)是在Gb5-合成反应混合物中进行,该反应混合物包含半乳糖、PEP、ATP、UTP、Gb4-OR^{1A}、β-1,3-半乳糖转移酶及UDP-Gal再生酶组。产物Gb5-OR^{1A}可自反应混合物中分离。

[0014] 上述方法可进一步包括转换Gb5-OR^{1A}的步骤,以获得岩藻糖-Gb5-OR^{1A}(Globo H)或唾液酸-Gb5-OR^{1A}(SSEA4)。

[0015] 有关Globo H的合成,该方法可进一步包括:(vii)在GDP-Fuc及α-1,2-岩藻糖转移酶(如来自幽门螺旋杆菌(H.pylor))存在下,将Gb5-OR^{1A}转换为岩藻糖-Gb5-OR^{1A},该步骤可与步骤(viii)结合,步骤(viii)包括在GDP-Fuc再生酶组存在下,自岩藻糖生成GDP-Fuc,其中该GDP-Fuc再生酶组包含L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶(例如来自脆弱拟杆菌(B.fragilis))、丙酮酸激酶(例如来自大肠杆菌)及焦磷酸化酶(例如来自大肠杆菌)。本发明的一实施例中,步骤(vii)及(viii)是在岩藻糖-Gb5-合成反应混合物中进行,该反应混合物包含岩藻糖、ATP、GTP、PEP、Gb5-OR^{1A}、α-1,2-岩藻糖转移酶及GDP-Fuc再生酶组。岩藻糖-Gb5-合成反应混合物是将Gb5-合成反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合而得。如需要,产物岩藻糖-Gb5-OR^{1A}可自反应混合物中分离。

[0016] 有关SSEA4的合成,本发明方法可进一步包括:(ix)在CMP-Neu5Ac及α-2,3-唾液酸转移酶(例如来自海洋细菌(M.bacteria))存在下,将Gb5-OR^{1A}转换为唾液酸-Gb5-OR^{1A},及(x):在CMP-Neu5Ac再生酶组存在下,自Neu5Ac生成CMP-Neu5Ac,其中该CMP-Neu5Ac再生酶组包含胞苷单磷酸激酶(例如来自大肠杆菌)、CMP-唾液酸合成酶(例如来自多杀性巴氏杆菌(P.multocida))、丙酮酸激酶(例如来自大肠杆菌)及视需要的焦磷酸化酶(例如来自大肠杆菌)。步骤(ix)及(x)可在唾液酸-Gb5-合成反应混合物中实施,该反应混合物包含Neu5Ac、CTP、PEP、Gb5-OR^{1A}、α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组。唾液酸-Gb5-合成反应混合物是将Gb5-合成反应混合物与至少Neu5Ac、CTP、α-2,3-唾液酸转移酶、胞苷单磷酸激酶及CMP-唾液酸合成酶混合而得。唾液酸-Gb5-OR^{1A}可自反应混合物中分离。

[0017] 本发明的一实施例中,Globo H合成方法的实施方式如下:(i)在本发明所述UDP-Gal再生酶存在下,自半乳糖生成UDP-Gal;(ii)在Gb3-合成反应混合物中将本发明所述的

Lac-OR^{1A}转换为Gb3-OR^{1A}，其中该反应混合物包含至少UDP-Gal、α-1,4-半乳糖基转移酶及UDP-Gal再生酶；(iii)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶混合以形成Gb4-合成混合物；(iv)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；(v)进一步在允许Gb4-OR^{1A}转换为Gb5-OR^{1A}的条件下，在1,3-半乳糖转移酶存在下培养Gb4-合成反应混合物；(vi)将含有Gb5-OR^{1A}的反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合以形成岩藻糖-Gb5-OR^{1A}反应混合物；(vii)在允许转换Gb5-OR^{1A}为岩藻糖-Gb5-OR^{1A}的条件下，培养岩藻糖-Gb5-OR^{1A}反应混合物；及视需要(viii)分离岩藻糖-Gb5-OR^{1A}。

[0018] 本发明另一实施例中，Globo H的合成方法的实施方式如下：(i)如本发明所述，在UDP-Gal再生酶存在下，自半乳糖生成UDP-Gal；(ii)如本发明所述，在Gb3-合成反应混合物中将Lac-OR^{1A}转换为Gb3-OR^{1A}，其中该反应混合物包含至少UDP-Gal、α-1,4-半乳糖转移酶及UDP-Gal再生酶；(iii)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶混合以形成Gb4-合成混合物；(iv)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；(v)分离Gb4-OR^{1A}；(vi)将Gb4-OR^{1A}与β-1,3-半乳糖转移酶及UDP-Gal再生酶组予以混合以形成Gb5-合成反应混合物；(vii)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，培养Gb5-合成反应混合物；(viii)将Gb5-合成反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合以形成岩藻糖-Gb5-OR^{1A}反应混合物；(ix)在允许转换Gb5-OR^{1A}为岩藻糖-Gb5-OR^{1A}的条件下，培养岩藻糖-Gb5-OR^{1A}反应混合物；及视需要(x)分离岩藻糖-Gb5-OR^{1A}。

[0019] 一种合成SSEA4方法的实施方式是如下：(i)如本发明所述，在UDP-Gal再生酶存在下，自半乳糖生成UDP-Gal；(ii)如本发明所述，在Gb3-合成反应混合物中，将Lac-OR^{1A}转换为Gb3-OR^{1A}，其中该反应混合物包含至少UDP-Gal、α-1,4-半乳糖转移酶及UDP-Gal再生酶；(iii)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶予以混合以形成Gb4-合成反应混合物；(iv)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；(v)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，在1,3-半乳糖转移酶存在下，进一步培养Gb4-合成反应混合物；(vi)将Gb4-合成反应混合物与至少Neu5Ac、CTP、α-2,3-唾液酸转移酶、胞苷单磷酸激酶及CMP-唾液酸合成酶予以混合以形成唾液酸-Gb5-OR^{1A}反应混合物；(vii)在允许转换Gb5-OR^{1A}为唾液酸-Gb5-OR^{1A}的条件下，培养唾液酸-Gb5-OR^{1A}反应混合物；及视需要(viii)分离唾液酸-Gb5-OR^{1A}。

[0020] 另一方面，一种合成SSEA4方法的实施方式是如下：(i)如本发明所述，在UDP-Gal再生酶存在下，自半乳糖生成UDP-Gal；(ii)如本发明所述，在Gb3-合成反应混合物中将Lac-OR^{1A}转换为Gb3-OR^{1A}，其中该反应混合物包含至少UDP-Gal、α-1,4-半乳糖转移酶及UDP-Gal再生酶；(iii)将Gb3-合成反应混合物与至少GalNAc、β-1,3-N-乙酰半乳糖胺转移酶、N-乙酰己糖胺1-激酶及N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶予以混合以形成Gb4-合成反应混合物；(iv)在允许转换Gb3-OR^{1A}为Gb4-OR^{1A}的条件下，培养Gb4-合成反应混合物；(v)分离Gb4-OR^{1A}；(vi)将Gb4-OR^{1A}与β-1,3-半乳糖转移酶及UDP-Gal再生酶组予以混合以形成

Gb5-合成反应混合物；(vii)在允许转换Gb4-OR^{1A}为Gb5-OR^{1A}的条件下，培养Gb5-合成反应混合物；(viii)将Gb5-OR^{1A}与α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组予以混合以形成唾液酸-Gb5-合成反应混合物，其中该CMP-Neu5Ac再生酶组包含胞苷单磷酸激酶、CMP-唾液酸合成酶、丙酮酸激酶及焦磷酸化酶；(ix)在允许转换Gb4-OR^{1A}为唾液酸-Gb5-OR^{1A}的条件下，培养唾液酸-Gb5-合成反应混合物；及视需要(x)分离唾液酸-Gb5-OR^{1A}。

[0021] 在一些实施例中，本发明所述的方法是使用Gb3(例如，尾式)作为酶性合成寡糖的起始材料。该方法包括：(i)如前所述，在UDP-GalNAc再生酶组存在下，自GalNAc生成UDP-GalNAc，并在UDP-GalNAc及β-1,3-N-乙酰半乳糖胺转移酶存在下，将Gb3-OR^{1A}转换为Gb4-OR^{1A}，其中R^{1A}为氢、经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经取代的炔基、经取代或未经取代的环碳基、经取代或未经取代的杂环基、经取代或未经取代的芳基、经取代或未经取代的杂芳基或氧保护基团。R^{1A}的实施例包括，但不限于氢、丙烯基、生物素、神经酰胺或一非氢基团(如烷基)，其是进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代。某些特定的实施例中，R^{1A}是氢、丙烯基、经取代的烷基、生物素或神经酰胺。步骤(i)及(ii)可在Gb4-合成反应混合物中进行，其中该反应混合物包含GalNAc、PEP、ATP、UTP、Gb3-OR^{1A}、β-1,3-N-乙酰半乳糖胺转移酶及UDP-GalNAc再生酶组。如需要，可分离Gb4-OR^{1A}。

[0022] 上述方法可进一步包括：(iii)在UDP-Gal及β-1,3-半乳糖转移酶存在下，将Gb4-OR^{1A}转换为Gb5-OR^{1A}，该步骤可与步骤(iv)结合，步骤(iv)包括如本发明所述的，在UDP-Gal再生酶存在下，自半乳糖生成UDP-Gal。步骤(iii)及(iv)可在Gb5-合成反应混合物中进行，其中该反应混合物包含半乳糖、PEP、ATP、UTP、Gb4-OR^{1A}、β-1,3-半乳糖转移酶及UDP-Gal再生酶组。产物Gb5-OR^{1A}可自反应混合物中分离而得。

[0023] 根据本发明的一实施例，可将Gb5-OR^{1A}再转换为岩藻糖-Gb5-OR^{1A}，其方法如下：(v)在GDP-Fuc及α-1,2-岩藻糖转移酶存在下，将Gb5-OR^{1A}转换为岩藻糖-Gb5-OR^{1A}，该步骤可与步骤(vi)结合，步骤(vi)包括在本发明所述的GDP-Fuc再生酶组存在下，自岩藻糖生成GDP-Fuc。步骤(v)及(vi)可在岩藻糖-Gb5-合成反应混合物中进行，该反应混合物包含岩藻糖、ATP、GTP、PEP、Gb5-OR^{1A}、α-1,2-岩藻糖转移酶及GDP-Fuc再生酶组。如需要，岩藻糖-Gb5-合成反应混合物是将Gb5-合成反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合而得。本发明方法可进一步包括分离岩藻糖-Gb5-OR^{1A}。

[0024] 本发明的另一实施例，可将Gb5-OR^{1A}再转换为唾液酸-Gb5-OR^{1A}，其方法如下：(vii)在CMP-Neu5Ac及α-2,3-唾液酸转移酶存在下，将Gb5-OR^{1A}转换为唾液酸-Gb5-OR^{1A}，该步骤可与步骤(viii)结合，步骤(viii)包括在本发明所述的CMP-Neu5Ac再生酶组存在下，自Neu5Ac生成CMP-Neu5Ac。步骤(vii)及(viii)可在唾液酸-Gb5-合成反应混合物中进行，该反应混合物包含Neu5Ac、CTP、PEP、Gb5-OR^{1A}、α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组。在特定实施例中，唾液酸-Gb5-生成反应混合物是将Gb5-合成反应混合物与至少Neu5Ac、CTP、α-2,3-唾液酸转移酶、胞苷单磷酸激酶及CMP-唾液酸合成酶混合而得。产物唾液酸-Gb5-OR^{1A}可分离自反应混合物中。

[0025] 根据本发明的其它实施例，此处所述的方法是关于使用Gb4(例如，尾式)作为起始

物以合成寡糖。该方法包括：(i)如本发明所述的UDP-Gal再生酶组存在下，自半乳糖生成UDP-Gal；及(ii)在UDP-Gal及β-1,3-半乳糖转移酶存在下，将Gb4-OR^{1A}转换为Gb5-OR^{1A}，其中R^{1A}为氢、经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经取代的炔基、经取代或未经取代的环碳基、经取代或未经取代的杂环基、经取代或未经取代的芳基、经取代或未经取代的杂芳基或氧保护基团。R^{1A}的实施例包括，但不限于氢、丙烯基、生物素、神经酰胺或一非氢基团(如烷基)，其是进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代。某些特定的实施例中，R^{1A}是氢、丙烯基、经取代的烷基、生物素或神经酰胺。根据本发明方法，步骤(i)及(ii)可在Gb5-合成反应混合物中实施，该反应混合物包含半乳糖、PEP、ATP、UTP、Gb4-OR^{1A}、β-1,3-半乳糖转移酶及UDP-Gal再生酶组。另一方面或选择性的，可分离所获得的Gb5-OR^{1A}。

[0026] 上述方法可进一步包括：(iii)在GDP-Fuc及α-1,2-岩藻糖转移酶存在下，将Gb5-OR^{1A}转换为岩藻糖-Gb5-OR^{1A}，该步骤可与步骤(iv)结合，步骤(iv)包括如本发明所述的GDP-Fuc再生酶组存在下，自岩藻糖生成GDP-Fuc。步骤(iii)及(iv)可在岩藻糖-Gb5-合成反应混合物中进行，该反应混合物包含岩藻糖、ATP、GTP、PEP、Gb5-OR^{1A}、α-1,2-岩藻糖转移酶及GDP-Fuc再生酶组。岩藻糖-Gb5-合成反应混合物是将Gb5-合成反应混合物与至少岩藻糖、GTP、α-1,2-岩藻糖转移酶及L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶混合而得。产物岩藻糖-Gb5-OR^{1A}可自反应混合物中分离。

[0027] 另一方面，上述方法可进一步包括：(v)在CMP-Neu5Ac及α-2,3-唾液酸转移酶存在下，将Gb5-OR^{1A}转换为唾液酸-Gb5-OR^{1A}，该步骤可与步骤(vi)结合，步骤(vi)包括在如本发明所述的CMP-Neu5Ac再生酶套组存在下，自Neu5Ac生成CMP-Neu5Ac。步骤(v)及(vi)可在唾液酸-Gb5-合成反应混合物中进行，该反应混合物包含Neu5Ac、CTP、PEP、Gb5-OR^{1A}、α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组。唾液酸-Gb5-合成反应混合物是将Gb5-合成应混合物与至少Neu5Ac、CTP、α-2,3-唾液酸转移酶、胞苷单磷酸激酶及CMP-唾液酸合成酶混合而得。根据本发明方法所得的唾液酸-Gb5-OR^{1A}可自反应混合物中分离。

[0028] 根据本发明其它实施例，此处所述方法是关于自Gb5合成岩藻糖-Gb5寡糖(Globo H)。该方法包括：(i)在如本发明所述的GDP-Fuc再生酶套组存在下，自岩藻糖生成GDP-Fuc；(ii)在GDP-Fuc及α-1,2-岩藻糖转移酶存在下，将Gb5-OR^{1A}转换为岩藻糖-Gb5-OR^{1A}；及视需要(iii)分离岩藻糖-Gb5-OR^{1A}。步骤(i)及(ii)可在岩藻糖-Gb5-合成反应混合物中进行，该反应混合物包含岩藻糖、ATP、GTP、PEP、Gb5-OR^{1A}、α-1,2-岩藻糖转移酶及GDP-Fuc再生酶组。

[0029] 根据本发明的其它实施例，此处所述的方法是关于自Gb5合成唾液酸-Gb5寡糖(Globo H)。该方法包括：(i)在如本发明所述的CMP-Neu5Ac再生酶套组存在下，自Neu5Ac生成CMP-Neu5Ac；(ii)在CMP-Neu5Ac及α-2,3-唾液酸转移酶存在下，将Gb5-OR^{1A}转换为唾液酸-Gb5-OR^{1A}；及视需要(iii)分离唾液酸-Gb5-OR^{1A}。步骤(i)及(ii)可在唾液酸-Gb5-合成反应混合物中进行，该反应混合物包含Neu5Ac、CTP、PEP、Gb5-OR、α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组。

[0030] 前述任一合成方法，不论所涉及的任一酶或反应中的任一底物(例如，乳糖、Gb3、Gb4或Gb5)，皆可固定于一支持体(support member)上。

[0031] 本发明的另一方面是有关于用于本发明寡糖合成方法的酶反应器(emzymatic

reactor)。该酶反应器可包含一或多个下述的反应室(reaction chamber):

[0032] (a)合成Gb3-OR^{1A}的反应室,其中该室包含α-1,4-半乳糖转移酶及UDP-Gal再生酶组,其包含半乳糖激酶、UDP-糖焦磷酸化酶、丙酮酸激酶及视需要的焦磷酸化酶;

[0033] (b)合成Gb4-OR^{1A}的反应室,其中该室包含β-1,3-N-乙酰半乳糖胺转移酶及UDP-GalAc再生酶组,其包含N-乙酰己糖胺1-激酶、N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶、丙酮酸激酶及视需要的焦磷酸化酶;

[0034] (c)合成Gb5-OR^{1A}的反应室,其中该室包含β-1,3-半乳糖转移酶及UDP-Gal再生酶组;

[0035] (d)合成岩藻糖-Gb5-OR^{1A}的反应室,其中该室包含α-1,2-岩藻糖转移酶及GDP-Fuc再生酶组,其包含L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶、丙酮酸激酶及视需要的焦磷酸化酶;及

[0036] (e)合成唾液酸-Gb5-OR^{1A}的反应室,其中该室包含α-2,3-唾液酸转移酶及CMP-Neu5Ac再生酶组,其包含胞苷单磷酸激酶、CMP-唾液酸合成酶、丙酮酸激酶及视需要的焦磷酸化酶。

[0037] 根据本发明的一些实施例,该酶反应器包含反应室如下:(a)及(b);(a)、(b)及(c);(a)、(b)、(c)及(d);(a)、(b)、(c)及(e);(b)及(c);(b)、(c)及(d);(b)、(c)及(e);(c)及(d)或(c)及(e)。

[0038] 根据本发明的其它实施例,此处所述的酶反应器包含一反应室,该反应室包含半乳糖转移酶(例如,α-1,4-半乳糖转移酶、β-1,4-半乳糖转移酶、α-1,3-半乳糖转移酶或β-1,3-半乳糖转移酶)及本发明所述的UDP-Gal再生酶组,其可包含半乳糖激酶、UDP焦磷酸化酶、丙酮酸激酶及视需要的焦磷酸化酶。

[0039] 在任一反应室中,一或多种酶可固定于一支持体上。根据本发明的特定实施例,该一或多个UDP-Gal再生酶组、UDP-GalNAc再生酶组、GDP-Fuc再生酶组及CMP-Neu5Ac再生酶组是各自固定于一支持体上。根据本发明的其它实施例,一反应室中的所有酶是固定于一支持体上。

[0040] 本发明也涵盖根据上述合成方法所制得的寡糖。

[0041] 本发明的一或多种实施例是详载于以下具体实施方式中。本发明的其它特征、目的及优势将清楚揭露于以下的定义、图式、实施例及权利要求书中。

[0042] 化学式的定义

[0043] 本发明相关的特定功能基团及化学名词的定义说明如后。化学元素的鉴别是根据周期表(CAS version,Handbook of Chemistry and Physics,75th Ed.)封页内部,而特定功能基团也据此作定义。此外,有机化学的一般原则以及特殊官能基团及其活性如有机化学,托马斯·索瑞尔,大学科学书籍,索萨利托,1999;史密斯和马区的高等有机化学,第5版,约翰威立公司,纽约,2001年;Larock,综合有机转化,VCH出版公司,纽约,1989;以及卡拉瑟斯,有机合成的一些现代方法,第3版,剑桥大学出版社,剑桥,1987(Organic Chemistry,Thomas Sorrell,University Science Books,Sausalito,1999;Smith and March March's Advanced Organic Chemistry,5th Edition,John Wiley&Sons,Inc.,New York,2001;Larock,Comprehensive Organic Transformations,VCH Publishers,Inc.,New York,1989;以及Carruthers,Some Modern Methods of Organic Synthesis,3rd

Edition,Cambridge University Press,Cambridge,1987)所描述。

[0044] 本发明所述化合物可包含一或多个非对称中心,因而存在多种异构体形式,诸如对映体及/或非对映异构体。例如,本发明所述化合物可呈个别对映体、非对映异构体或几何异构体等形式,或可呈混合异构体形式,包括外消旋混合物及富含一或多种异构体的混合物。可利用包括掌性高效液相层析(HPLC)以及掌性盐类的形成与结晶等公知技术,自混合物中分离同分异构物;或利用非对称合成法制备较佳的异构物。请参见Jacques等人,对映异构体,外消旋体和分辨率(威利跨学科,纽约,1981);维伦等人,四面体期刊33:2725(1977);以列,E.L.,碳化合物的立体化学(麦格劳-希尔,纽约,1962);以及维伦,S.H.,拆分剂和光学分辨率表P.268(E.L.以列编,圣母院大学出版社,巴黎圣母院,1972)(Enantiomers,Racemates and Resolutions(Wiley Interscience,New York,1981);Wilen等人,Tetrahedron 33:2725(1977);Eliel,E.L.Stereochemistry of Carbon Compounds(McGraw-Hill,NY,1962);以及Wilen,S.H.Tables of Resolving Agents and Optical Resolutions p.268(E.L.Eliel,Ed.,Univ.of Notre Dame Press,Notre Dame,IN 1972))。本发明另包括实质上不含其它异构物的个别异构物;另一方面,本发明包括多种异构物的混合物。

[0045] 根据本案说明书所载数值的范围,其是意指包括每一各别数值及该范围内的次范围。例如,“C_{1-6}烷基”意指包含C_1、C_2、C_3、C_4、C_5、C_6、C_{1-6}、C_{1-5}、C_{1-4}、C_{1-3}、C_{1-2}、C_{2-6}、C_{2-5}、C_{2-4}、C_{2-3}、C_{3-6}、C_{3-5}、C_{3-4}、C_{4-6}、C_{4-5}及C_{5-6}烷基。

[0046] 本文所指“烷基”为一具有1至30个碳原子的直链或支链的饱和烃基团(“C_{1-30}烷基”)。某些特定实施例,该烷基具有1至20个碳原子(“C_{1-20}烷基”)。某些特定实施例,该烷基具有1至10个碳原子(“C_{1-10}烷基”)。某些特定实施例,该烷基具有1至9个碳原子(“C_{1-9}烷基”)。某些特定实施例,该烷基具有1至8个碳原子(“C_{1-8}烷基”)。某些特定实施例,该烷基具有1至7个碳原子(“C_{1-7}烷基”)。某些特定实施例,该烷基具有1至6个碳原子(“C_{1-6}烷基”)。某些特定实施例,该烷基具有1至5个碳原子(“C_{1-5}烷基”)。某些特定实施例,该烷基具有1至4个碳原子(“C_{1-4}烷基”)。某些特定实施例,该烷基具有1至3个碳原子(“C_{1-3}烷基”)。某些特定实施例,该烷基具有1至2个碳原子(“C_{1-2}烷基”)。某些特定实施例,该烷基具有1个碳原子(“C_1烷基”)。某些特定实施例,该烷基具有2至6个碳原子(“C_{2-6}烷基”)。C_{1-6}烷基的实施例包括甲基(C_1)、乙基(C_2)、正丙基(C_3)、异丙基(C_3)、正丁基(C_4)、第三丁基(C_4)、第二丁基(C_4)、异丁基(C_4)、正戊基(C_5)、3-戊基(C_5)、戊基(C_5)、新戊基(C5)、3-甲基-2-丁基(C_5)、第三戊基(C_5)及正己基(C_6)。烷基的其它实施例包括正庚基(C_7)、正辛基(C_8)及其类似物。除非另有定义,否则如本文所指的烷基是各自独立地未经取代(未经取代的烷基)或经一或多个取代基所取代(经取代的烷基)。某些特定实施例,烷基是未经取代的C_{1-10}烷基(例如-CH_3)。其它特定实施例,烷基是经取代的C_{1-10}烷基。

[0047] 本文所指“烯基”或“烯类”为一具有2至30个碳原子及一或多个双键(例如1、2、3或4个双键)的直链或支链烃基团。某些特定实施例,该烯基具有2至20个碳原子(“C_{2-20}烯基”)。某些特定实施例,该烯基具有2至10个碳原子(“C_{2-10}烯基”)。某些特定实施例,该烯基具有2至9个碳原子(“C_{2-9}烯基”)。某些特定实施例,该烯基具有2至8个碳原子(“C_{2-8}烯基”)。某些特定实施例,该烯基具有2至7个碳原子(“C_{2-7}烯基”)。某些特定实施例,该烯基具有2至6个碳原子(“C_{2-6}烯基”)。某些特定实施例,该烯基具有2至5个碳原子(“C_{2-5}烯基”)。某些特

定实施例，该烯基具有2至4个碳原子（“C_{2-4}烯基”）。某些特定实施例，该烯基具有2至3个碳原子（“C_{2-3}烯基”）。某些特定实施例，该烯基具有2个碳原子（“C_2烯基”）。其中一或多个C＝C双键可存在于内部（如2-丁烯基）或末端（如1-丁烯基）。C_{2-4}烯基包括乙烯基（C_2）、1-丙烯基（C_3）、2-丙烯基（C_3）、1-丁烯基（C_4）、2-丁烯基（C_4）、丁间二烯基（C_4）及其类似物。C_{2-6}烯基的实施例包括前述的C_{2-4}烯基以及戊烯基（C_5）、戊间二烯基（C_5）、己烯基（C_6）及其类似物。烯基的其它实施例包括庚烯基（C_7）、辛烯基（C_8）、辛间三烯基（C_8）及其类似物。除非另有定义，否则如本文所指的烯基是各自独立地未经取代（未经取代的烯基）或经一或多个取代基所取代（经取代的烯基）。某些特定实施例，烷基是未经取代的C_{2-10}烯基。其它特定实施例，烯基是经取代的C_{2-10}烯基。

[0048] 本文所指“炔基”或“炔类”为一具有2至30个碳原子及一或多个参键（例如1、2、3或4个参键）的直链或支链烃基团（“C_{2-10}炔基”）。某些特定实施例，该炔基具有2至20个碳原子（“C_{2-20}炔基”）。某些特定实施例，该炔基具有2至10个碳原子（“C_{2-10}炔基”）。某些特定实施例，该炔基具有2至9个碳原子（“C_{2-9}炔基”）。某些特定实施例，该炔基具有2至8个碳原子（“C_{2-8}炔基”）。某些特定实施例，该炔基具有2至7个碳原子（“C_{2-7}炔基”）。某些特定实施例，该炔基具有2至6个碳原子（“C_{2-6}炔基”）。某些特定实施例，该炔基具有2至5个碳原子（“C_{2-5}炔基”）。某些特定实施例，该炔基具有2至4个碳原子（“C_{2-4}炔基”）。某些特定实施例，该炔基具有2至3个碳原子（“C_{2-3}炔基”）。某些特定实施例，该炔基具有2个碳原子（“C_2炔基”）。其中一或多个C≡C参键可存在于内部（如2-炔基）或末端（如1-炔基）。C_{2-4}炔基的实施例包括，但不限于乙炔基（C2）、1-丙炔基（C3）、2-丙炔基（C3）、1-丁炔基（C4）、2-丁炔基（C4）及其类似物。C2-6炔基的实施例包括前述的C_{2-4}炔基以及戊炔基（C_5）、己炔基（C_6）及其类似物。炔基的其它实施例包括庚炔基（C_7）、辛炔基（C_8）及其类似物。除非另有定义，否则如本文所指的炔基是各自独立地未经取代（未经取代的炔基）或经一或多个取代基所取代（经取代的炔基）。某些特定实施例，炔基是未经取代的C_{2-10}炔基。其它特定实施例，炔基是经取代的C_{2-10}炔基。

[0049] 本文所指“环碳基”为一具有3至10个环碳原子的非芳族环烃基团（“C_{3-10}环碳基”），且该非芳族环系统中不含杂原子。某些特定实施例，该环碳基具有3至8个环碳原子（“C_{3-8}环碳基”）。某些特定实施例，该环碳基具有3至7个环碳原子（“C_{3-7}环碳基”）。某些特定实施例，该环碳基具有3至6个环碳原子（“C_{3-6}环碳基”）。某些特定实施例，该环碳基具有5至10个环碳原子（“C_{5-10}环碳基”）。C_{3-6}环碳基的实施例包括，但不限于环丙基（C_3）、环丙烯基（C_3）、环丁基（C_4）、环丁烯基（C_4）、环戊基（C_5）、环戊烯基（C_5）、环己基（C_6）、环己烯基（C_6）、环己间二烯基（C_6）及其类似物。C_{3-8}环碳基的例包括，但不限于前述的C_{3-6}环碳基及环庚基（C_7）、环庚烯基（C_7）、环庚间二烯基（C_7）、环庚间三烯基（C_7）、环辛基（C_8）、环辛烯基（C_8）、二环[2.2.1]庚基（C_7）、二环[2.2.2]辛基（C_8）及其类似物。C_{3-10}环碳基的实施例包括，但不限于前述的C_{3-8}环碳基及环壬基（C_9）、环壬烯基（C_9）、环癸基（C_{10}）、环癸烯基（C10）、八氢-1H-茚基（C_9）、十氢萘基（C_{10}）、螺环[4.5]癸基（C_{10}）及其类似物。如前述实施例所例示者，某些特定实施例，环碳基是单环（“单环环碳基”）或多环（例如包含一稠环、螺环或桥环系统，诸如双环系统（“双环环碳基”）或三环系统（“三环环碳基”）），且其是呈饱和或可包含一或多个碳-碳双键或三键。“环碳基”也包含环系统，其中如前所述的环碳与一或多个芳基或杂芳基稠合，且其中连接点是位于该环碳上，在此情形下，在该环碳系统下继续标记其碳数。除非另有定义，否则如本文所指的环碳基是各自独立地未经取代（未经取代的环碳基）或经一或

多个取代基所取代(经取代的环碳基)。某些特定实施例,环碳基是未经取代的C_{3-10}环碳基。其它特定实施例,环碳基是经取代的C_{3-10}环碳基。

[0050] 某些特定实施例,"环碳基"是一具有3-10个环碳原子("C_{3-10}环烷基")的饱和单环。某些特定实施例,环烷基具有3至8个环碳原子("C_{3-8}环烷基")。某些特定实施例,环烷基具有3至6个环碳原子("C_{3-6}环烷基")。某些特定实施例,环烷基具有5至6个环碳原子("C_{5-6}环烷基")。某些特定实施例,环烷基具有5至10个环碳原子("C_{5-10}环烷基")。C_{5-6}环烷基的实施例包括环戊基(C_5)及环己基(C_6)。C_{3-6}环烷基的实施例包括前述的C_{5-6}环烷基及环丙基(C_3)及环丁基(C_4)。C_{3-8}环烷基的实施例包括前述的C3-6环烷基及环庚基(C_7)及环辛基(C_8)。除非另有定义,否则如本文所指的环烷基是各自独立地未经取代(未经取代的环烷基)或经一或多个取代基所取代(经取代的环烷基)。某些特定实施例,环烷基是未经取代的C_{3-10}环烷基。其它特定实施例,环烷基是经取代的C_{3-10}环烷基。

[0051] 本文所指"杂环基"为一具有1至4个杂原子的3-至14-元非芳族环系统,其中每一杂原子各自独立地选自氮、氧及硫("3-14元杂环基")。在包含一或多个氮原子的杂环基中,其连接点可为碳原子或氮原子,如价数允许。杂环基可为单环或多环(例如稠环、桥环或螺环形成双环系统或三环系统),且该杂环基可为饱和或包含一或多个C=C双键或C≡C参键。多环杂环系统在其中一或两个环中包含一或多个杂原子。杂环基也包括如上所定义的环系统,其是与一或多个环碳基稠合,其中连接点可位于环碳基或杂环上。多环的杂环基系统,在其一或两个环系统上可包括一或多个杂原子。"杂环基"也包含环系统,其中如前所述的杂环与一或多个环碳基稠合,且其中连接点是位于环碳上或杂环基环上,或如前所定义的环系统,其中该杂环与一或多个芳基或杂芳基稠合,且其中连接点是位于杂环上,在此情形下,在该杂环系统下继续标记其碳数。除非另有定义,否则如本文所指的杂环基是各自独立地未经取代(未经取代的杂环基)或经一或多个取代基所取代(经取代的杂环基)。某些特定实施例,杂环基是未经取代的3-14元杂环基。其它特定实施例,杂环基是经取代的3-14元杂环基。

[0052] 某些特定实施例中,该杂环基为具有环碳原子及1-4个杂环原子的5-10元非芳族环系统,其中该杂环原子是独立地选自氮、氧及硫。某些特定实施例中,该杂环基为具有环碳原子及1-4个杂环原子的5-8元非芳族环系统,其中该杂环原子是独立地选自氮、氧及硫。某些特定实施例中,该杂环基为具有环碳原子及1-4个杂环原子的5-6元非芳族环系统,其中该杂环原子是独立地选自氮、氧及硫。某些特定实施例中,该5-6元杂环基具有1-3个选自氮、氧及硫的杂环原子。某些特定实施例中,该5-6元杂环基具有1-2个选自氮、氧及硫的杂环原子。某些特定实施例中,该5-6元杂环基具有1个选自氮、氧及硫的杂环原子。

[0053] 包含一个杂原子的3元杂环基的实施例包括,但不限于氮杂环丙烯基(azirdinyl)、环氧乙烷基(oxiranyl)、环硫乙烷基(thiorenyl)。4元杂环基(包含一杂原子)的实施例包括,但不限于氮杂环丁烷基(azetidinyl)、氧杂环丁烷基(oxetanyl)及硫杂环丁烷基(thietanyl)。5元杂环基(包含一杂原子)的实施例包括,但不限于四氢呋喃基、二氢呋喃基、四氢苯硫基、二氢苯硫基、吡咯啶基、二氢吡咯基及吡咯基-2,5-二酮。5元杂环基(包含2个杂原子)的实施例包括,但不限于二氧戊环基(dioxolanyl)、氧硫环己烷基(oxathiolanyl)及二硫戊环(dithiolanyl)。5元杂环基(包含2个杂原子)的实施例包括,但不限于三唑啉基(triazolinyl)、恶二唑啉基(oxadiazolinyl)、噻二唑啉基

(thiadiazolinyl)。6元杂环基(包含1个杂原子)的实施例包括,但不限于哌啶基,四氢哌喃基、二氢哌喃基及噻烷基(thianyl)。6元杂环基(包含2个杂原子)的实施例包括,但不限于哌嗪基、吗啉基、二噻烷基(dithianyl)及二氧杂环己烷基(dioxanyl)。6元杂环基(包含2个杂原子)的实施例包括,但不限于三嗪基。7元杂环基(包含1个杂原子)的实施例包括,但不限于氮杂环庚烷基(azepanyl)、氧杂环丁烷基(oxepanyl)及硫杂环庚烷基(thiepanyl)。8元杂环基(包含1个杂原子)的实施例包括,但不限于氮杂环辛烷基(azocanyl)、氧杂环辛烷基(oxecanyl)及硫杂环辛烷基(thiocanyl)。双环杂环基的实施例包括,但不限于吲哚啉基(indolinyl)、异吲哚啉基、二氢苯并呋喃基、二氢苯并噻吩基、四氢苯并噻吩基、四氢苯并呋喃基、四氢吲哚基、四氢喹啉基、四氢异喹啉基、十氢喹啉基、十氢异喹啉基、八氢苯并哌喃基、八氢异苯并哌喃基、十氢萘啶基、十氢-1,8-萘啶基、八氢吡咯并[3,2-b]吡咯、吲哚基、邻苯二甲酰亚胺基、萘二甲酰亚胺基、苯并二氢吡喃基(chromanyl)、色满基(chromenyl)、1H-苯并[e][1,4]二氮呯(1H-benzo[e][1,4]diazepinyl)、1,4,5,7-四氢-哌喃并[3,4-b]吡咯基、5,6-二氢-4H-呋喃并[3,2-b]吡咯基、6,7-二氢-5H-呋喃并[3,2-b]哌喃基、5,7-二氢-4H-噻吩并[2,3-c]哌喃基、2,3-二氢-1H-吡咯并[2,3-b]吡啶基、2,3-二氢呋喃并[2,3-b]吡啶基、4,5,6,7-四氢-1H-吡咯并[2,3-b]吡啶基、4,5,6,7-四氢呋喃并[3,2-c]吡啶基、4,5,6,7-四氢噻吩并[3,2-b]吡啶基、1,2,3,4-四氢-1,6-萘啶基及其类似物。

[0054] 本文所指“芳基”是指一具有6-14个环碳原子且不含杂原子的单或多(例如双环或三环)的4n+2芳香环系统(例如6、10或14个电子分享在一个环状数组中)“C_{6-14}芳基”。某些特定实施例中,该芳基具有6个环碳原子(“C_6芳基”;例如苯基)。某些特定实施例中,该芳基具有10个环碳原子(“C_{10}芳基”;例如萘基,特定如1-萘基或2-萘基)。某些特定实施例中,该芳基具有14个环碳原子(“C_{14}芳基”;如蒽基(anthracyl))。“芳基”也包括环系统,如前所定义者,该芳基环是与一或多个环碳基或杂环基绸合,其中该连接点是位于芳基环上,在此情形下,在该芳环系统下继续标记其碳数。除非另有定义,否则如本文所指的芳基是各自独立地未经取代(未经取代的芳基)或经一或多个取代基所取代(经取代的芳基)。某些特定实施例中,芳基是未经取代的C_{6-14}芳基。其它特定实施例中,芳基是经取代的C_{6-14}芳基。

[0055] 本文所指“杂芳基”是指一5-14元单或多环(例如双环)的4n+2芳族环系统(例如6、10或14个电子分享在一个环状数组中),该芳族环系统具有碳原子及1-4个杂原子,其中该杂原子是各自独立地选自氮、氧及硫(“5-14元杂芳基”)。在包含一或多个氮原子的杂芳基中,如价数允许者,其连接点可为碳原子或氮原子。多环杂芳基环系统在其中一或两个环中包含一或多个杂原子。“杂芳基”也包括环系统,其中如上所述的杂芳环是与一或多个碳环或杂环基稠合,其中连接点是位于杂芳基环上,且在此情形下,在该杂芳环系统下继续标记其碳数。“杂芳基”也包括环系统,其中如上所述的杂芳环是与一或多个芳基稠合,其中连接点是位于芳基或杂芳基环上,且在此情形下,在该经稠合多环(芳基/杂芳基)系统下继续标记其碳数。其中一环不包含杂原子多环杂芳基团(例如吲哚基、喹啉基、咔唑基(carbazolyl)及其类似物),其连接点可位于环上,也就是,该环包含杂原子(如2-吲哚基)或该环不包含杂原子(如5-吲哚基)。

[0056] 某些特定实施例中,该杂芳基为5-10元芳族环系统,具有环碳原子及1-4个杂环原子,其中该杂环原子是各自独立地选自氮、氧及硫(“5-10元杂芳基”)。某些特定实施例中,该杂芳基为5-8元芳族环系统,具有环碳原子及1-4个杂环原子,其中该杂环原子是各自独

立地选自氮、氧及硫(“5-8元杂芳基”)。某些特定实施例中,该杂芳基为5-6元芳族环系统,具有环碳原子及1-4个杂环原子,其中该杂环原子是各自独立地选自氮、氧及硫(“5-6元杂芳基”)。某些特定实施例中,5-6元杂芳基具有1-3个各自独立地选自氮、氧及硫的杂环原子。某些特定实施例,5-6元杂芳基具有1-2个各自独立地选自氮、氧及硫的杂环原子。某些特定实施例,5-6元杂芳基具有1个选自氮、氧及硫的杂环原子。除非另有定义,否则如本文所指的杂芳基是各自独立地未经取代(未经取代的杂芳基)或经一或多个取代基所取代(经取代的杂芳基)。某些特定实施例,杂芳基是未经取代的5-14元杂芳基。其它特定实施例,杂芳基是经取代的5-14元杂芳基。

[0057] 5元杂芳基(包含一杂原子)的实施例包括,但不限于吡咯基、呋喃基及噻吩基。含有2个杂子的5元杂芳基的实施例包括,但不限于咪唑基、吡唑基、恶唑基、异恶唑基、噻唑基及异噻唑基。含有3个杂原子的5元杂芳基的实施例包括,但不限于三唑、恶二唑及噻二唑基。含有4个杂原子的5元杂芳基的实施例包括,但不限于四唑基。含有1个杂原子的6元杂芳基的实施例包括,但不限于吡啶基。含有2个杂原子的6元杂芳基的实施例包括,但不限于嗒嗪基、嘧啶基及吡嗪基。含有3或4个杂子的6元杂芳基的实施例分别包括,但不限于三嗪基及四嗪基。含有1个杂子的7元杂芳基的实施例包括,但不限于氮呯基(azepinyl)、恶呯基(oxepinyl)及噻呯基(thiepinyl)。5,6-双环杂芳基的实施例包括,但不限于吲哚基、异吲哚基、吲唑基、苯并三唑基、苯并噻吩基、异苯并噻吩基、苯并呋喃基、苯并异呋喃基、苯并咪唑基、苯并恶唑基、苯并异恶唑基、苯并恶二唑基、苯并噻唑基、苯并异噻唑基、苯并噻二唑基、吲哚嗪基(indolizinyl)及嘌呤基。6,6-双环杂芳基的实施例包括,但不限于萘啶基、喋啶基、喹啉基、异喹啉基、噌啉基、喹恶啉基、酞嗪基及喹唑啉基。三环杂芳基的实施例包括,但不限于啡啶基、二苯并呋喃基、咔唑基、吖啶基、啡噻嗪基、啡恶嗪基及啡嗪基。

[0058] 本文所用的术语“部份不饱和”是指一环部份包含至少一个双键或一个三键,该术语“部份不饱和”意图涵括具有多处不饱和位置的环,但不包括如本文所定义的芳基(例如苯基或杂芳基部份)。

[0059] 本文所用的术语“饱和”是指一环部份不具有任一双键或三键,也就是,该环结构全为单键。

[0060] 本文所指的烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是视需要可经取代(例如,经取代或未经取代烷基,经取代或未经取代烯基,经取代或未经取代炔基,经取代或未经取代杂烷基,经取代或未经取代杂烯基,经取代或未经取代杂炔基,经取代或未经取代环碳基,经取代或未经取代杂环基,经取代或未经取代芳基或经取代或未经取代杂芳基)。一般而言,本文所用“取代”的术语(不论前面是否带有“视需要”此一术语),意指基团中至少一个氢原子经一允许的取代基所取代,例如在取代的过程中取代基会导致一稳定化合物,该稳定化合物不致自动进行诸如重排、环化、脱除或其它转化反应。除非另有定义,否则一经取代的基团在一或多个适当位置具有取代基,且当超过一个位置经取代时,每一取代位置的取代基可为相同或不同。本文所用“取代”的术语包括所有可允许的有机化合物,任何本文所指导致形成稳定化合物的任何取代基。本发明包括任何及所有的组合以获得稳定化合物。基于本发明的目的,诸如氮的杂原子可被氢及/或任何如本文所述的适当的取代基所取代,其是满足该杂原子的价数,以致形成稳定基团。

[0061] 碳原子取代基的实施例包括,但不限于卤素、-CN、-NO_2、-N_3、-SO_2H、-SO_3H、-OH、-

OR^{aa}、$-ON(R^{bb})_2$、$-N(R^{bb})_2$、$-N(R^{bb})_3^+X^-$、$-N(OR^{cc})R^{bb}$、-SH、$-SR^{aa}$、$-SSR^{cc}$、$-C(=O)R^{aa}$、$-CO_2H$、-CHO、$-C(OR^{cc})_2$、$-CO_2R^{aa}$、$-OC(=O)R^{aa}$、$-OCO_2R^{aa}$、$-C(=O)N(R^{bb})_2$、$-OC(=O)N(R^{bb})_2$、$-NR^{bb}C(=O)R^{aa}$、$-NR^{bb}CO_2R^{aa}$、$-NR^{bb}C(=O)N(R^{bb})_2$、$-C(=NR^{bb})R^{aa}$、$-C(=NR^{bb})OR^{aa}$、$-OC(=NR^{bb})R^{aa}$、$-OC(=NR^{bb})OR^{aa}$、$-C(=NR^{bb})N(R^{bb})_2$、$-OC(=NR^{bb})N(R^{bb})_2$、$-NR^{bb}C(=NR^{bb})N(R^{bb})_2$、$-C(=O)NR^{bb}SO_2R^{aa}$、$-NR^{bb}SO_2R^{aa}$、$-SO_2N(R^{bb})_2$、$-SO_2R^{aa}$、$-SO_2OR^{aa}$、$-OSO_2R^{aa}$、$-S(=O)R^{aa}$、$-OS(=O)R^{aa}$、$-Si(R^{aa})_3$、$-OSi(R^{aa})_3$-C(=S)$N(R^{bb})_2$、$-C(=O)SR^{aa}$、$-C(=S)SR^{aa}$、$-SC(=S)SR^{aa}$、$-SC(=O)SR^{aa}$、$-OC(=O)SR^{aa}$、$-SC(=O)OR^{aa}$、$-SC(=O)R^{aa}$、$-P(=O)_2R^{aa}$、$-OP(=O)_2R^{aa}$、$-P(=O)(R^{aa})_2$、$-OP(=O)(R^{aa})_2$、$-OP(=O)(OR^{cc})_2$、$-P(=O)_2N(R^{bb})_2$、$-OP(=O)_2N(R^{bb})_2$、$-P(=O)(NR^{bb})_2$、$-OP(=O)(NR^{bb})_2$、$-NR^{bb}P(=O)(OR^{cc})_2$、$-NR^{bb}P(=O)(NR^{bb})_2$、$-P(R^{cc})_2$、$-P(R^{cc})_3$、$-OP(R^{cc})_2$、$-OP(R^{cc})_3$、$-B(R^{aa})_2$、$-B(OR^{cc})_2$、$-BR^{aa}(OR^{cc})$、$C_{1-10}$烷基、$C_{1-10}$全卤烷基、$C_{2-10}$烯基、$C_{2-10}$炔基、$C_{1-10}$杂烷基、$C_{2-10}$杂烯基、$C_{2-10}$杂炔基、$C_{3-14}$环碳基、3-14元杂环基、$C_{6-14}$芳基及5-14元杂芳基，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{dd}基团所取代；

[0062] 或在一碳原子上的两个相偕氢原子被以下基团所取代=O、=S、$=NN(R^{bb})_2$、$=NNR^{bb}C(=O)R^{aa}$、$=NNR^{bb}C(=O)OR^{aa}$、$=NNR^{bb}S(=O)_2R^{aa}$、$=NR^{bb}$或$=NOR^{cc}$；

[0063] 每一R^{aa}是独立地选自C_{1-10}烷基、C_{1-10}全卤烷基、C_{2-10}烯基、C_{2-10}炔基、C_{1-10}杂烷基、C_{2-10}杂烯基、C_{2-10}杂炔基、C_{3-10}环碳基、3-14元杂环基、C_{6-14}芳基及5-14元杂芳基，或两个R^{aa}基团结合形成一个3-14元杂环基或5-14元杂芳基环，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{dd}基团所取代；

[0064] 每一R^{bb}是独立地选自氢、-OH、$-OR^{aa}$、$-N(R^{cc})_2$、-CN、$-C(=O)R^{aa}$、$-C(=O)N(R^{cc})_2$、$-CO_2R^{aa}$、$-SO_2R^{aa}$、$-C(=NR^{cc})OR^{aa}$、$-C(=NR^{cc})N(R^{cc})_2$、$-SO_2N(R^{cc})_2$、$-SO_2R^{cc}$、$-SO_2OR^{cc}$、$-SOR^{aa}$、$-C(=S)N(R^{cc})_2$、$-C(=O)SR^{cc}$、$-C(=S)SR^{cc}$、$-P(=O)_2R^{aa}$、$-P(=O)(R^{aa})_2$、$-P(=O)_2N(R^{cc})_2$、$-P(=O)(NR^{cc})_2$、C_{1-10}烷基、C_{1-10}全卤烷基、C_{2-10}烯基、C_{2-10}炔基、C_{1-10}杂烷基、C_{2-10}杂烯基、C_{2-10}杂炔基、C_{3-10}环碳基、3-14元杂环基、C_{6-14}芳基及5-14元杂芳基，或两个R^{bb}基团结合形成一个3-14元杂环基或5-14元杂芳基环，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{dd}基团所取代；

[0065] 每一R^{cc}是独立地选自氢、C_{1-10}烷基、C_{1-10}全卤烷基、C_{2-10}烯基、C_{2-10}炔基、C_{1-10}杂烷基、C_{2-10}杂烯基、C_{2-10}杂炔基、C_{3-10}环碳基、3-14元杂环基、C_{6-14}芳基及5-14元杂芳基，或两个R^{cc}基团结合形成一个3-14元杂环基或5-14元杂芳基环，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{dd}基团所取代；

[0066] 每一R^{dd}是独立地选自卤素、-CN、$-NO_2$、$-N_3$、$-SO_2H$、$-SO_3H$、-OH、$-OR^{ee}$、$-ON(R^{ff})_2$、$-N(R^{ff})_2$、$-N(R^{ff})_3^+X^-$、$-N(OR^{ee})R^{ff}$、-SH、$-SR^{ee}$、$-SSR^{ee}$、$-C(=O)R^{ee}$、$-CO_2H$、$-CO_2R^{ee}$、$-OC(=O)R^{ee}$、$-OCO_2R^{ee}$、$-C(=O)N(R^{ff})_2$、$-OC(=O)N(R^{ff})_2$、$-NR^{ff}C(=O)R^{ee}$、$-NR^{ff}CO_2R^{ee}$、$-NR^{ff}C(=O)N(R^{ff})_2$、$-C(=NR^{ff})OR^{ee}$、$-OC(=NR^{ff})R^{ee}$、$-OC(=NR^{ff})OR^{ee}$、$-C(=NR^{ff})N(R^{ff})_2$、$-OC(=NR^{ff})N(R^{ff})_2$、$-NR^{ff}C(=NR^{ff})N(R^{ff})_2$,$-NR^{ff}SO_2R^{ee}$、$-SO_2N(R^{ff})_2$、$-SO_2R^{ee}$、$-SO_2OR^{ee}$、$-OSO_2R^{ee}$、$-S(=O)R^{ee}$、$-Si(R^{ee})_3$、$-OSi(R^{ee})_3$、$-C(=S)N(R^{ff})_2$、$-C(=O)SR^{ee}$、$-C(=S)SR^{ee}$、$-SC(=S)SR^{ee}$、$-P(=O)_2R^{ee}$、$-P(=O)(R^{ee})_2$、$-OP(=O)(R^{ee})_2$、$-OP(=O)(OR^{ee})_2$、$C_{1-6}$烷基、$C_{1-6}$全卤烷基、$C_{2-6}$烯基、

C_{2-6}炔基、C_{1-6}杂烷基、C_{2-6}杂烯基、C_{2-6}杂炔基、C_{3-10}环碳基、3-10元杂环基、C_{6-10}芳基及5-10元杂芳基，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{gg}基团所取代，或两个相偕R^{dd}基团可结合形成=O或=S；

[0067] 每一R^{ee}是独立地选自C_{1-6}烷基、C_{1-6}全卤烷基、C_{2-6}烯基、C_{2-6}炔基、C_{1-6}杂烷基、C_{2-6}杂烯基、C_{2-6}杂炔基、C_{3-10}环碳基、3-10元杂环基、C_{6-10}芳基及3-10元杂芳基，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{gg}基团所取代；

[0068] 每一R^{ff}是独立地选自氢、C_{1-6}烷基、C_{1-6}全卤烷基、C_{2-6}烯基、C_{2-6}炔基、C_{1-6}杂烷基、C_{2-6}杂烯基、C_{2-6}杂炔基、C_{3-10}环碳基、3-10元杂环基、C_{6-10}芳基及5-10元杂芳基，或两个R^{ff}基团结合形成一个3-14元杂环基或5-14元杂芳基环，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{gg}基团所取代；及

[0069] 每一R^{gg}是独立地选自卤素、-CN、$-NO_2$、$-N_3$、$-SO_2H$、$-SO_3H$、-OH、$-OC_{1-6}$烷基、$-ON(C_{1-6}$烷基$)_2$、$-N(C_{1-6}$烷基$)_2$、$-N(C_{1-6}$烷基$)_3{}^+X^-$、$-NH(C_{1-6}$烷基$)_2{}^+X^-$、$-NH_2(C_{1-6}$烷基$){}^+X^-$、$-NH_3{}^+X^-$、$-N(OC_{1-6}$烷基$)(C_{1-6}$烷基$)$、$-N(OH)(C_{1-6}$烷基$)$、-NH(OH)、-SH、$-SC_{1-6}$烷基、$-SS(C_{1-6}$烷基$)$、$-C(=O)(C_{1-6}$烷基$)$、$-CO_2H$、$-CO_2(C_{1-6}$烷基$)$、$-OC(=O)(C_{1-6}$烷基$)$、$-OCO_2(C_{1-6}$烷基$)$、$-C(=O)NH_2$、$-C(=O)N(C_{1-6}$烷基$)_2$、$-OC(=O)NH(C_{1-6}$烷基$)$、$-NHC(=O)(C_{1-6}$烷基$)$、$-N(C_{1-6}$烷基$)C(=O)(C_{1-6}$烷基$)$、$-NHCO_2(C_{1-6}$烷基$)$、$-NHC(=O)N(C_{1-6}$烷基$)_2$、$-NHC(=O)NH(C_{1-6}$烷基$)$、$-NHC(=O)NH_2$、$-C(=NH)O(C_{1-6}$烷基$)$、$-OC(=NH)(C_{1-6}$烷基$)$、$-OC(=NH)OC_{1-6}$烷基、$-C(=NH)N(C_{1-6}$烷基$)_2$、$-C(=NH)NH(C_{1-6}$烷基$)$、$-C(=NH)NH_2$、$-OC(=NH)N(C_{1-6}$烷基$)_2$、$-OC(NH)NH(C_{1-6}$烷基$)$、$-OC(NH)NH_2$、$-NHC(NH)N(C_{1-6}$烷基$)_2$、$-NHC(=NH)NH_2$、$-NHSO_2(C_{1-6}$烷基$)$、$-SO_2N(C_{1-6}$烷基$)_2$、$-SO_2NH(C_{1-6}$烷基$)$、$-SO_2NH_2$，$-SO_2C_{1-6}$烷基、$-SO_2OC_{1-6}$烷基、$-OSO_2C_{1-6}$烷基、$-SOC_{1-6}$烷基、$-Si(C_{1-6}$烷基$)_3$、$-OSi(C_{1-6}$烷基$)_3$、$-C(=S)N(C_{1-6}$烷基$)_2$、$C(=S)NH(C_{1-6}$烷基$)$、$C(=S)NH_2$、$-C(=O)S(C_{1-6}$烷基$)$、$-C(=S)SC_{1-6}$烷基、$-SC(=S)SC_{1-6}$烷基、$-P(=O)_2(C_{1-6}$烷基$)$、$-P(=O)(C_{1-6}$烷基$)_2$、$-OP(=O)(C_{1-6}$烷基$)_2$、$-OP(=O)(OC_{1-6}$烷基$)_2$、C_{1-6}烷基、C_{1-6}全卤烷基、C_{2-6}烯基、C_{2-6}炔基、C_{1-6}杂烷基、C_{2-6}杂烯基、C_{2-6}杂炔基、C_{3-10}环碳基、3-10元杂环基、C_{6-10}芳基及5-10元杂芳基；或两个相偕R^{gg}取代基可结合形成=O或=S；其中X^-为相对离子。

[0070] 本文所指的“卤基”或“卤素”是指氟基(-F)、氯基(-Cl)、溴基(-Br)或碘基(-I)。

[0071] 本文所指的“相对离子”是指一负电荷基团，其与正电荷四级胺连结来维持其电中性。相对离子的实施例包括卤离子(如，F^-、Cl^-、Br^-、I^-)、NO_3^-、ClO_4^-、OH^-、$H_2PO_4^-$、HSO_4^-、磺酸离子(如，甲磺酸盐、三氟甲磺酸盐、p-甲苯磺酸盐、苯磺酸盐、10-樟脑磺酸盐、萘-2-磺酸盐、萘-1-磺酸-5-磺酸盐、乙烷-1-磺酸-2-磺酸盐及其类似物)及羧酸离子(例如，乙酸盐、醋酸盐、丙酸盐、苯甲酸盐、甘油酸盐、酒石酸盐、甘醇酸盐及其类似物)。

[0072] 本文所指“羰基”是指一基团，其中直接连结至母分子的碳为sp2混成(hybridized)，且可经氧、氮或硫原子取代，例如选自酮类($-C(=O)R^{aa}$)、羧酸类($-CO_2H$)、醛类(-CHO)、酯类($-CO2R^{aa}$、$-C(=O)SR^{aa}$、$-C(=S)SR^{aa}$)、酰胺类($-C(=O)N(R^{bb})_2$、$-C(=O)NR^{bb}SO_2R^{aa}$、$-C(=S)N(R^{bb})2$)及亚胺类($-C(=NR^{bb})R^{aa}$、$-C(=NR^{bb})OR^{aa}$)、$-C(=NR^{bb})N(R^{bb})_2$)的基团，其中R^{aa}及R^{bb}是如本文所定义者。

[0073] 本文所指“迭氮化物”或“迭氮基”是指$-N_3$基团。

[0074] 本文所指“硫化物”或“硫基”是指-SH基团。“经取代的硫化物”或“经取代的硫基”延伸所指为一硫基团，其中直接附着于母分子上的硫原子是经取代(但取代基不为氢)，且包含选自$-SR^{aa}$、$-S=SR^{cc}$、$-SC(=S)SR^{aa}$、$-SC(=O)SR^{aa}$、$-SC(=O)OR^{aa}$及$-SC(=O)R^{aa}$的基团，其中R^{aa}及R^{cc}是如本文所定义者。

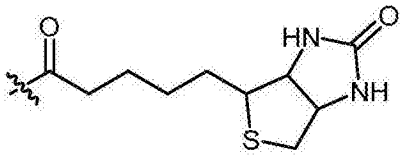
[0075] 本文所指“氨基”或“胺”是指$-NH_2$基团。“经取代的氨基”或“经取代的胺”延伸所指为一单一取代的氨基，双取代的氨基或三取代的氨基。某些特定实施例中，“经取代的氨基”为单一取代的氨基或双取代的氨基基团。

[0076] 本文所指“经单取代的氨基”或“经单取代的胺”是指一氨基基团，其中直接附着于母分子上的氮原子是经一个氢原子及一个非为氢的基团所取代，其包括选自$-NH(R^{bb})$、$-NHC(=O)R^{aa}$、$-NHCO_2R^{aa}$、$-NHC(=O)N(R^{bb})_2$、$-NHC(=NR^{bb})N(R^{bb})_2$、$-NHSO_2R^{aa}$、$-NHP(=O)(OR^{cc})_2$及$-NHP(=O)(NR^{bb})_2$的基团，其中$R^{aa}$、$R^{bb}$及$R^{cc}$是如本文所定义者，且其中$-NH(R^{bb})$基团的$R^{bb}$部份不为氢。

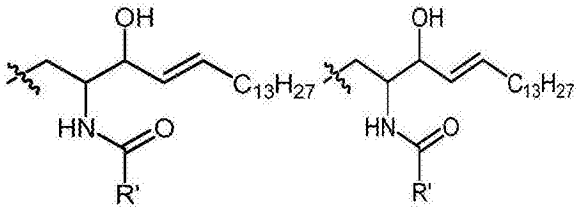
[0077] 本文所指“经双取代的氨基”是指一氨基基团，其中直接附着于母分子上的氮原子是经两个不为氢的基团所取代，其包括选自$-N(R^{bb})_2$、$-NR^{bb}C(=O)R^{aa}$、$-NR^{bb}CO_2R^{aa}$、$-NR^{bb}C(=O)N(R^{bb})_2$、$-NR^{bb}C(=NR^{bb})N(R^{bb})_2$、$-NR^{bb}SO_2R^{aa}$、$-NR^{bb}P(=O)(OR^{cc})_2$及$-NR^{bb}P(=O)(NR^{bb})_2$的基团，其中$R^{aa}$、$R^{bb}$及$R^{cc}$是如本文所定义者，其限制条件为该直接附着于母分子上的氮原子不被氢所取代。

[0078] 本文所指“经三取代的氨基”是指一氨基基团，其中直接附着于母分子上的氮原子是经三个基团所取代，其包括选自$-N(R^{bb})_3$及$-N(R^{bb})_3{}^+X^-$的基团，其中R^{bb}及X^-是如本文所定义者。

[0079] 本文所指“生物素”，例如R^{1A}基团，其包含以下结构：

[0080] 

[0081] 本文所指“神经酰胺”，例如R^{1A}基团，其包含以下结构：

[0082] 

[0083] 其中R′是视需要经取代的C_6-C_{30}烷基(例如C_6、C_7、C_8、C_9、C_{10}、C_{11}、C_{12}、C_{13}、C_{14}、C_{15}、C_{16}、C_{17}、C_{18}、C_{19}、C_{20}、C_{21}、C_{22}、C_{23}、C_{24}、C_{25}、C_{26}、C_{27}、C_{28}、C_{29}或C_{30}烷基)、视需要经取代的C_6-C_{30}烯基(例如C_6、C_7、C_8、C_9、C_{10}、C_{11}、C_{12}、C_{13}、C_{14}、C_{15}、C_{16}、C_{17}、C_{18}、C_{19}、C_{20}、C_{21}、C_{22}、C_{23}、C_{24}、C_{25}、C_{26}、C_{27}、C_{28}、C_{29}或C_{30}烯基)、或视需要经取代的C_6-C_{30}炔基(例如C_6、C_7、C_8、C_9、C_{10}、C_{11}、C_{12}、C_{13}、C_{14}、C_{15}、C_{16}、C_{17}、C_{18}、C_{19}、C_{20}、C_{21}、C_{22}、C_{23}、C_{24}、C_{25}、C_{26}、C_{27}、C_{28}、C_{29}或C_{30}炔基)基团。

[0084] 在价数允许下，氮原子可经取代或未经取代，且其包含初级、二级、三级及四级的氮原子。氮原子取代基的实施例包括，但不限于氢、-OH、$-OR^{aa}$、$-N(R^{cc})_2$、-CN、$-C(=O)R^{aa}$、$-C(=O)N(R^{cc})_2$、$-CO_2R^{aa}$、$-SO_2R^{aa}$、$-C(=NR^{bb})R^{aa}$、$-C(=NR^{cc})OR^{aa}$、$-C(=NR^{cc})N(R^{cc})_2$、$-SO_2N(R^{cc})_2$、$-SO_2R^{cc}$、$-SO_2OR^{cc}$、$-SOR^{aa}$、$-C(=S)N(R^{cc})_2$、$-C(=O)SR^{cc}$、$-C(=S)SR^{cc}$、$-P(=O)_2R^{aa}$、-P

(=O)$(R^{aa})_2$、-P(=O)$_2$N$(R^{cc})_2$、-P(=O)$(NR^{cc})_2$、C_{1-10}烷基、C_{1-10}全卤烷基、C_{2-10}烯基、C_{2-10}炔基、C_{1-10}杂烷基、C_{2-10}杂烯基、C_{2-10}杂炔基、C_{3-10}环碳基、3-14元杂环基、C_{6-14}芳基及5-14元杂芳基，或两个R^{cc}基团结合形成一个3-14元杂环基或5-14元杂芳基环，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{dd}基团所取代，且其中R^{aa}、R^{bb}、R^{cc}及R^{dd}是如本文定义者。

[0085] 某些特定实施例中，位于氮原子上的取代基是氮保护基团(本文中也称为氨基保护基团)。氮保护基团包括，但不限于-OH、$-OR^{aa}$、$-N(R^{cc})_2$、$-C(=O)R^{aa}$、$-C(=O)N(R^{cc})_2$、$-CO_2R^{aa}$、$-SO_2R^{aa}$、$-C(=NR^{cc})R^{aa}$、$-C(=NR^{cc})OR^{aa}$、$-C(=NR^{cc})N(R^{cc})_2$、$-SO_2N(R^{cc})_2$、$-SO_2R^{cc}$、$-SO_2OR^{cc}$、$-SOR^{aa}$、$-C(=S)N(R^{cc})_2$、$-C(=O)SR^{cc}$、$-C(=S)SR^{cc}$、C_{1-10}烷基(例如，芳烷基或杂芳烷基)，C_{2-10}烯基、C_{2-10}炔基、C_{1-10}杂烷基、C_{2-10}杂烯基、C_{2-10}杂炔基、C_{3-10}环碳基、3-14元杂环基、C_{6-14}芳基及5-14元杂芳基，其中每一烷基、烯基、炔基、杂烷基、杂烯基、杂炔基、环碳基、杂环基、芳基及杂芳基是各自独立地经0、1、2、3、4或5个R^{dd}基团所取代，且其中R^{aa}、R^{bb}、R^{cc}及R^{dd}是如本文定义者。氮保护基团是本领域技术人员所熟知，且揭露于Protecting Groups in Organic Synthesis,T.W.Greene and P.G.M.Wuts,3rd edition,John Wiley&Sons,其细节也并入本文中作为参考。

[0086] 例如，氮保护基团可为诸如酰胺基团(如$-C(=O)R^{aa}$)，其包括(但不限定为)甲酰胺、乙酰胺、氯乙酰胺、三氯乙酰胺、三氟乙酰胺、苯乙酰胺、3-苯丙酰胺、吡啶酰胺(picolinamide)、3-吡啶羧酰胺、N-苯甲酰基苯丙胺酰基衍生物、苯甲酰胺、p-苯甲酰苯胺、o-硝基苯乙酰胺、o-硝基苯氧基乙酰胺、乙酰乙酰胺、(N'-二硫苄氧胺酰基)乙酰胺、3-(p-羟苯基)丙酰胺、3-(o-硝基苯基)丙酰胺、2-甲基-2-(o-硝基苯氧基)丙酰胺、2-甲基-2-(o-苯偶氮苯氧基)丙酰胺、4-氯丁酰胺、3-甲基-3-硝基丁酰胺、o-硝基桂皮酰胺、N-乙酰基甲硫胺酸衍生物、o-硝基苯甲酰胺及o-(苯甲酰氧基甲基)苯甲酰胺。

[0087] 氮保护基团可为诸如胺甲酸酯基团(如$-C(=O)OR^{aa}$)，其包括(但不限定为)氨甲酸甲酯、氨甲酸乙酯、9-芴甲氧羰酰胺(9-fluorenylmethyl carbamate;(Fmoc))、9-(2-磺基)芴甲氧羰酰胺、9-(2,7-二溴)芴甲氧羰酰胺、2,7-二-t-丁基-[9-(10,10-二氧-10,10,10,10-四氢噻吨基)]甲基氨甲酸酯(DBD-Tmoc)、4-甲氧基苯甲酰甲基氨甲酸酯(Phenoc)、2,2,2-三氯乙基氨甲酸酯(Troc)、2-三甲基硅烷基乙基氨甲酸酯(Teoc)、2-苯基乙基氨甲酸酯(hZ)、1-(1-金刚烷)-1-甲基乙基氨甲酸酯(Adpoc)、1,1-二甲基-2-卤乙基氨甲酸酯、1,1-二甲基-2,2-二溴乙基氨甲酸酯(DB-t-BOC)、1,1-二甲基-2,2,2-三氯乙基氨甲酸酯(TCBOC)、1-甲基-1-(4-联苯基)乙基氨甲酸酯(Bpoc)、1-(3,5-二-t-丁基苯基)-1-甲基乙基氨甲酸酯(t-Bumeoc)、2-(2'-4'-吡啶基)乙基氨甲酸酯(Pyoc)、2-(N,N-二环已烷甲酰胺基)乙基氨甲酸酯、t-丁基氨甲酸酯(BOC)、1-金刚烷氨甲酸酯(Adoc)、乙烯基氨甲酸酯(Voc)、丙烯基氨甲酸酯(Alloc)、1-异丙基丙烯基氨甲酸酯(Ipaoc)、桂酰基氨甲酸酯(Coc)、4-硝基桂酰基氨甲酸酯(Noc)、8-喹啉基氨甲酸酯、N-羟基六氢吡啶基氨甲酸酯、烷基二硫代氨甲酸酯、苯甲基氨甲酸酯(Cbz)、p-甲氧基苯甲基氨甲酸酯(Moz)、p-硝基苯甲基氨甲酸酯、p-溴苯甲基氨甲酸酯、p-氯苯甲基氨甲酸酯、2,4-二氯苯甲基氨甲酸酯、4-甲基亚磺酰基苯甲基氨甲酸酯(Msz)、9-蒽基甲基氨甲酸酯、二苯基甲基氨甲酸酯、2-甲基硫乙基氨甲酸酯、2-甲基磺酰基乙基氨甲酸酯、2-(p-甲苯磺酰基)乙基氨甲酸酯、[2-(1,3-二硫杂环已基)]甲基氨甲酸酯(Dmoc)、4-甲基苯硫基氨甲酸酯(Mtpc)、2,4-二甲基苯硫基氨甲

酸酯(Bmpc)、2-磷鎓基乙基氨甲酸酯(Peoc)、2-三苯基磷鎓基异丙基氨甲酸酯(Ppoc)、1,1-二甲基-2-氰基乙基氨甲酸酯、m-氯-p-酰氧基苯甲基氨甲酸酯、p-(二羟氧硼基)苯甲基氨甲酸酯、5-苯异恶唑基甲基氨甲酸酯、2-(三氟甲基)-6-色酮基甲基氨甲酸酯(Tcroc)、m-硝基苯基氨甲酸酯、3,5-二甲氧基苯甲基氨甲酸酯、o-硝基苯甲基氨甲酸酯、3,4-二甲氧基-6-硝基苯甲基氨甲酸酯、苯基(o-硝基苯基)甲基氨甲酸酯、t-戊基氨甲酸酯、S-苯甲基硫氨甲酸酯、p-氰基苯甲基氨甲酸酯、环丁基氨甲酸酯、环己基氨甲酸酯、环戊基氨甲酸酯、环丙基甲基氨甲酸酯、p-癸氧基苯甲基氨甲酸酯、2,2-二甲氧基酰基乙烯基氨甲酸酯、o-(N,N-二甲基甲酰氨基)苯甲基氨甲酸酯、1,1-二甲基-3-(N,N-二甲基甲酰氨基)丙基氨甲酸酯、1,1-二甲基丙炔基氨甲酸酯、二(2-吡啶基)甲基氨甲酸酯、2-呋喃基甲基氨甲酸酯、2-碘乙基氨甲酸酯、异茨醇基氨甲酸酯、异丁基氨甲酸酯、异烟碱基氨甲酸酯、p-(p'-甲氧基苯偶氮基)苯甲基氨甲酸酯、1-甲基环丁基氨甲酸酯、1-甲基环己基氨甲酸酯、1-甲基-1-环丙基甲基氨甲酸酯、1-甲基-1-(3,5-二甲氧基二甲氧基苯基)乙基氨甲酸酯、1-甲基-1-(p-苯基偶氮苯基)乙基氨甲酸酯、1-甲基-1-苯基乙基氨甲酸酯、1-甲基-1-(4-吡啶基)乙基氨甲酸酯、苯基氨甲酸酯、p-(苯基偶氮)苯甲基氨甲酸酯、2,4,6-三-t-丁基苯基氨甲酸酯、4-(三甲基铵基)苯甲基氨甲酸酯及2,4,6-三甲基苯甲基氨甲酸酯。

[0088] 氮保护基团可为诸如磺酰胺基团(如$-S(=O)_2R^{aa}$),其包括(但不限定为)p-甲苯磺酰胺(Ts)、苯磺酰胺、2,3,6,-三甲基-4-甲氧基苯磺酰胺(Mtr)、2,4,6-三甲氧基苯磺酰胺(Mtb)、2,6-二甲基-4-甲氧基苯磺酰胺(Pme)、2,3,5,6-四甲基-4-甲氧基苯磺酰胺(Mte)、4-甲氧基苯磺酰胺(Mbs)、2,4,6-三甲基苯磺酰胺(Mts)、2,6-二甲氧基-4-甲基苯磺酰胺(iMds)、2,2,5,7,8-五甲基色满(pentamethylchroman)-6-磺酰胺(Pmc)、甲烷磺酰胺(Ms)、β-三甲基硅烷乙烷磺酰胺(SES)、9-蒽磺酰胺、4-(4',8'-二甲氧基萘基甲基)苯磺酰胺(DNMBS)、苯甲基磺酰胺、三氟甲基磺酰胺、及苯甲基甲酰磺酰胺。

[0089] 其它氮保护基团可为包括(但不限定为)啡噻嗪基-(10)-酰基衍生物、N'-p-甲苯磺酰基胺酰基衍生物、N'-苯基氨基硫酰基衍生物、N-苯甲酰基苯基丙胺酰基衍生物、N-乙酰基甲硫氨酸衍生物、4,5-二苯基-3-恶唑啉-2-酮、N-邻苯二甲酰亚胺、N-二硫杂丁二酰亚胺(Dts)、N-2,3-二苯基顺丁烯二酰亚胺、N-2,5-二甲基吡咯、N-1,1,4,4-四甲基二硅烷氮杂环戊烷加合物(STABASE)、5-经取代1,3-二甲基-1,3,5-三氮杂环己-2-酮、5-经取代1,3-二苯甲基-1,3,5-三氮杂环己-2-酮、1-经取代3,5-二硝基-4-吡啶酮、N-甲基胺、N-丙烯基胺、N-[2-(三甲基硅烷)乙氧基]甲基胺(SEM)、N-3-乙酰氧基丙基胺、N-(1-异丙基-4-硝基-2-氧基-3-吡咯啉-3-基)胺、四级铵基盐类、N-苯甲基胺、N-二(4-甲氧基苯基)甲基胺、N-5-二苯并环庚基胺、N-三苯基甲基胺(Tr)、N-[(4-甲氧基苯基)二苯基甲基]胺(MMTr)、N-9-苯基芴基胺(PhF)、N-2,7-二氯-9-芴基甲烯胺、N-二茂铁基甲基胺(Fcm)、N-2-氨甲基吡啶基N'-氧化物、N-1,1-二甲基硫甲烯胺、N-苯甲基亚胺(ideneamine)、N-p-甲氧基苯甲基亚胺、N-二苯基甲烯胺、N-[(2-吡啶基)异亚丙基]甲烯胺、N-(N',N'-二甲基氨基亚甲基)胺、N,N'-异亚丙基二胺、N-p-硝基苯甲基亚胺、N-柳基亚胺、N-5-氯柳基亚胺、N-(5-氯-2-羟苯基)苯基甲烯胺、N-环己基亚胺、N-(5,5-二甲基-3-氧基-1-环己烯基)胺、N-硼烷衍生物、N-二苯基硼酸衍生物、N-[苯基(五酰基铬-或钨)酰基]胺、N-铜螯合物、N-锌螯合物、N-硝基胺、N-硝基对胺苯砷酸钠、胺N-氧化物、二苯基膦酰胺(Dpp)、二甲基硫基膦酰胺(Mpt)、二苯基硫基膦酰胺(Ppt)、二烷基磷酰胺盐、二苯甲基磷酰胺盐、二苯基磷酰胺盐、苯亚磺酰胺、

o-硝基苯亚磺酰胺(Nps)、2,4-二硝基苯亚磺酰胺、五氯苯亚磺酰胺、2-硝基-4-甲氧基苯亚磺酰胺、三苯基甲基亚磺酰胺及3-硝基吡啶亚磺酰胺(Npys)。

[0090] 某些特定实施例,位于氧原子上的取代基是氧保护基团(本文中也称的羟基保护基团)。氧保护基团包括,但不限于$-R^{aa}$、$-N(R^{bb})_2$、$-C(=O)SR^{aa}$、$-C(=O)R^{aa}$、$-CO_2R^{aa}$、$-C(=O)N(R^{bb})_2$、$-C(=NR^{bb})R^{aa}$、$-C(=NR^{bb})OR^{aa}$、$-C(=NR^{bb})N(R^{bb})_2$、$-S(=O)R^{aa}$、$-SO_2R^{aa}$、$-Si(R^{aa})_3$、$-P(R^{cc})_2$、$-P(R^{cc})_3$、$-P(=O)_2R^{aa}$、$-P(=O)(R^{aa})_2$、$-P(=O)(OR^{cc})_2$、$-P(=O)_2N(R^{bb})_2$及$-P(=O)(NR^{bb})_2$,其中R^{aa}、R^{b}、及R^{cc}是如本文定义者。氧保护基团是本领域技术人员所熟知,且揭露于Protecting Groups in Organic Synthesis,T.W.Greene and P.G.M.Wuts,3rd edition,John Wiley & Sons,其细节并入本文中作为参考。

[0091] 氧保护基团的实施例包括,但不限于甲基、甲氧基甲基(MOM)、甲基硫基甲基(MTM)、t-丁基硫基甲基、(苯基二甲基硅烷)甲氧基甲基(SMOM)、苯甲氧基甲基(BOM)、p-甲氧基苯甲氧基甲基(PMBM)、(4-甲氧基苯氧基)甲基(p-AOM)、愈伤木酚甲基(GUM)、t-丁氧基甲基、4-戊烯氧基甲基(POM)、硅烷氧基甲基、2-甲氧基乙氧基甲基(MEM)、2,2,2-三氯乙氧基甲基、双(2-氯乙氧基)甲基、2-(三甲基硅烷)乙氧基甲基(SEMOR)、四氢哌喃基(THP)、3-溴四氢哌喃基、四氢硫基哌喃基、1-甲氧基环己基、4-甲氧基四氢哌喃基(MTHP)、4-甲氧基四氢硫基哌喃基、4-甲氧基四氢硫基哌喃基S,S-二氧化物、1-[(2-氯-4-甲基)苯基]-4-甲氧基哌啶-4-基(CTMP)、1,4-二氧己环-2-基、四氢呋喃基、四氢硫基呋喃基、2,3,3a,4,5,6,7,7a-八氢-7,8,8-三甲基-4,7-甲醇苯并呋喃-2-基、1-乙氧基乙基、1-(2-氯乙氧基)乙基、1-甲基-1-甲氧基乙基、1-甲基-1-苯甲氧基乙基、1-甲基-1-苯甲氧基-2-氟乙基、2,2,2-三氯乙基、2-三甲基硅烷乙基、2-(苯硒基)乙基、t-丁基、丙烯基、p-氯苯基、p-甲氧基苯基、2,4-二硝基苯基、苯甲基(Bn)、p-甲氧基苯甲基、3,4-二甲氧基苯甲基、o-硝基苯甲基、p-硝基苯甲基、p-卤苯甲基、2,6-二氯苯甲基、p-氰基苯甲基、p-苯基苯甲基、2-甲基吡啶、4-甲基吡啶、3-甲基-2-甲基吡啶N-氧化物、二苯基甲基、p,p'-二硝基二苯甲基、5-二苯并环庚基、三苯基甲基、α-萘基二苯基甲基、p-甲氧基苯基二苯基甲基、二(p-甲氧基苯基)苯基甲基、三(p-甲氧基苯基)甲基、4-(4'-溴苯甲基酰氧苯基)二苯基甲基、4,4′,4″-参(4,5-二氯酞酰亚氨基苯基)甲基、4,4′,4″-参(乙酰丙酰氧基苯基)甲基、4,4′,4″-参(苯甲酰氧基苯基)甲基、3-(咪唑-1-基)双(4′,4″-二甲氧基苯基)甲基、1,1-双(4-甲氧基苯基)-1′-芘基甲基、9-蒽基、9-(9-苯基)氧杂蒽基(xanthenyl)、9-(9-苯基-10-氧基)蒽基、1,3-苯并二硫戊环-2-基、苯并异噻唑S,S-二氧化物、三甲基硅烷(TMS)、三乙基硅烷(TES)、三异丙基硅烷(TIPS)、二甲基异丙基硅烷(IPDMS)、二乙基异丙基硅烷(DEIPS)、二甲基叔己基硅烷、t-丁基二甲基硅烷(TBDMS)、t-丁基二苯基硅烷(TBDPS)、三苯甲基硅烷、三-p-二甲苯基硅烷、三苯基硅烷、二苯基甲基硅烷(DPMS)、t-丁基甲氧基苯基硅烷(TBMPS)、甲酸盐、苯甲酰基甲酸盐、乙酸盐、氯乙酸盐、二氯乙酸盐、三氯乙酸盐、三氟乙酸盐、甲氧基乙酸盐、三苯基甲氧基乙酸盐、苯氧基乙酸盐、p-氯苯氧基乙酸盐、3-苯基丙酸盐、4-氧基戊酸盐(戊酮酸盐)、4,4-(乙烯二硫基)戊酸盐(戊酮基二硫基缩醛)、新戊酸盐(pivaloate)、金刚酸盐、巴豆酸盐(crotonate)、4-甲氧基巴豆酸盐、苯甲酸盐、p-苯基苯甲酸盐、2,4,6-三甲基苯甲酸盐(mesitoate)、烷基甲基碳酸盐、9-芴基甲基碳酸盐(Fmoc)、烷基乙基碳酸盐、烷基2,2,2-三氯乙基碳酸盐(Troc)、2-(三甲基硅烷)乙基碳酸盐(TMSEC)、2-(苯基磺酰基)乙基碳酸盐(Psec)、2-(三苯基磷鎓基)乙基碳酸盐(Peoc)、烷基异丁基碳酸盐、烷基乙烯基碳酸盐、烷

基丙烯基碳酸盐、烷基p-硝基苯基碳酸盐、烷基苯甲基碳酸盐、烷基p-甲氧基苯甲基碳酸盐、烷基3,4-二甲氧基苯甲基碳酸盐、烷基o-硝基苯甲基碳酸盐、烷基p-硝基苯甲基碳酸盐、烷基S-苯甲基硫基碳酸盐、4-乙氧基-1-萘基碳酸盐、甲基二硫基碳酸盐、2-碘基苯甲酸盐、4-迭氮丁酸盐、4-硝基-4-甲基戊酸盐、o-(二溴甲基)苯甲酸盐、2-甲酰基苯磺酸盐、2-(甲基硫基甲氧基)乙基、4-(甲基硫基甲氧基)丁酸盐、2-(甲基硫基甲氧基甲基)苯甲酸盐、2,6-二氯-4-甲基苯氧基乙酸盐、2,6-二氯-4-(1,1,3,3-四甲基丁基)苯氧基乙酸盐、2,4-双(1,1-二甲基丙基)苯氧基乙酸盐、氯二苯基乙酸盐、异丁酸盐、单琥珀酸盐、(E)-2-甲基-2-丁烯酸盐、o-(甲氧基酰基)苯甲酸盐、α-萘甲酸盐、硝酸盐、烷基N,N,N',N'-四甲基偶磷基二酰胺盐、烷基N-苯基胺甲酸盐、硼酸盐、二甲基硫膦基、烷基2,4-二硝基苯基次磺酸盐、硫酸盐、甲烷磺酸盐(甲磺酸盐)、苯甲基磺酸盐及甲苯磺酸盐(Ts)。

[0092] 取代基的其它实施例是更详细地揭露于本文的发明说明、实施例及权利要求书。不应将上述所例示的取代基以任何方式限制本发明的范畴。

[0093] 本文所指"盐"是指任何及所有盐类。

[0094] 本文所指"医药学上可接受的盐"是指在良好医药评估下,适用于人类及低等动物组织的接触而不具过度的毒性、刺激性、过敏反应及类似反应的盐类,且在合理利益/风险比例上具有相称性。医药学上可接受的盐属公知技术的范畴,例如Berge等人,详述药学上可接受的盐,药学科学期刊(1977)66:1-19(describes pharmaceutically acceptable salts in detail in J.Pharmaceutical Sciences(1977)66:1-19)。根据本发明的化合物的医药学上可接受的盐包括那些衍生自适当无机及有机酸及碱。医药学上可接受的盐的实施例包括无毒酸加成盐,其是氨基基团的盐与诸如盐酸、氢溴酸、磷酸、硫酸及过氯酸的无机酸形成或与诸如乙酸、草酸、顺丁烯二酸、酒石酸、柠檬酸、琥珀酸或丙二酸的有机酸形成,或利用公知技术的其它方法,例如离子交换法。其它医药学上可接受的盐包括己二酸盐、海藻酸盐、抗坏血酸盐、天冬胺酸盐、苯磺酸盐、苯甲酸盐、重硫酸盐、硼酸盐、丁酸盐、樟脑酸盐、樟脑磺酸盐、柠檬酸盐、环戊烷丙酸盐、二葡萄糖酸盐、十二烷基磺酸盐、乙烷磺酸盐、甲酸盐、反丁烯二酸盐、葡萄庚酸盐、甘油磷酸盐、葡萄糖酸盐、半硫酸盐、庚酸盐、己酸盐、氢碘酸盐、2-羟-乙烷磺酸盐、乳糖酸盐、乳酸盐、月桂酸盐、月桂基硫酸盐、苹果酸盐、马来酸盐、丙二酸盐、甲烷磺酸盐、2-萘磺酸盐、烟碱酸盐、亚硝酸盐、十八烯酸盐、草酸盐、棕榈酸盐、双羟萘酸盐(pamoate)、果冻酸盐、过硫酸盐、3-苯基丙酸盐、磷酸盐、苦味酸盐、三甲基乙酸酸盐、丙酸盐、硬脂酸盐、琥珀酸盐、硫酸盐、酒石酸盐、硫基氰酸盐、p-甲苯磺酸盐、十一酸盐、戊酸盐及其类似物。衍生自适当碱的医药学上可接受的盐包括碱金属、碱土金属、铵基及$N^+(C_{1-4}烷基)_4$盐类。代表性的碱金属或碱土金属包括钠、锂、钾、钙、镁及其类似物。在适当的情况下,医药学上可接受的盐可进一步包括无毒性铵盐、季铵盐及胺阳离子,其是使用诸如卤化物、氢氧化物、羧酸盐、硫酸盐、磷酸盐、亚硝酸盐、低级烷基磺酸盐及芳基磺酸盐的相对离子来形成。

附图说明

[0095] 图1表示尾式丙烯基Gb3、Gb4、Gb5、Globo H及SSEA4的化学结构式。

[0096] 图2表示糖基化反应结合核苷酸糖再生及合成的TLC监测结果。A:结合半乳糖基化反应及核苷酸糖再生的合成,例如丙烯基-Gb3。B:结合乙酰半乳糖氨基化反应及UDP-Gal再

生的合成，例如丙烯基-Gb4。C:结合半乳糖基化反应及UDP-Gal再生的合成，例如丙烯基-Gb5。D:结合岩藻糖基化反应及GDP-Fuc再生的合成，例如丙烯基-Globo H。E:结合唾液酸反应及CMP-Neu5Ac再生的合成，例如丙烯基-SSEA4。

[0097] 图3表示鞘糖脂的生物合成途径，其涉及半乳糖残基的加成，可经由本发明所述组合半乳糖转移酶与UDP-Gal再生方法所催化而得。

[0098] 图4表示利用核苷酸糖再生系统，以酶性合成方法经由Lac→Gb3→Gb4→Gb5途径制备Globo H。

[0099] 图5表示利用核苷酸糖再生系统，以酶性合成方法经由Lac→Gb3→Gb4→Gb5途径制备SSEA4。

[0100] 图6表示利用核苷酸糖再生系统，以酶性合成方法经由丙烯基-Lac→丙烯基-Gb3→丙烯基-Gb4→丙烯基-Gb5途径制备丙烯基-Globo H。

[0101] 图7表示利用核苷酸糖再生系统，以酶性合成方法经由丙烯基-Lac→丙烯基-Gb3→丙烯基-Gb4→丙烯基-Gb5途径制备丙烯基-SSEA4。

[0102] 图8表示生化合成过程中所得高纯度的中间物丙烯基-Gb3、丙烯基-Gb4及丙烯基-Gb5。

[0103] 图9表示利用未经修饰及经修饰的FutC经由生化合成而得高纯度的丙烯基-Globo H。

[0104] 图10表示利用JT-FAJ-16经由生化合成而得高纯度的丙烯基-SSEA4。

具体实施方式

[0105] 以下是关于一种新发展的核苷酸糖再生方法，及其通过一合适糖基转移酶的反应，将糖基载入一适当受体的用途。本发明的做法在无须纯化中间物下，允许链反应来合成糖基化分子，例如Gb3、Gb4、Gb5、Globo H及SSEA4等寡糖，进而快速且大量生成糖基化产物。此外，本发明的合成方法可用于大规模制备所需的寡糖及糖共轭物(glycoconjugates)。

[0106] UDP-Gal再生系统及其半乳糖化的用途

[0107] UDP-Gal再生系统是例示于图2A，且所涉及的酶则如下表1所例：

[0108] 表1:用于UDP-Gal再生系统的酶

[0109]

酶	活性	实施例
半乳糖激酶(GalK)	在ATP存在下，催化α-D-半乳糖的磷酸化反应，以生成半乳糖-1-磷酸(Gal-1-P)	大肠杆菌 (如GenBank accession no. AP012306.1 智人(如GenBank accession no. NP_000145) 热水海嗜热菌(*M. hydrothermalis*) (如GenBank accession no. YP_004368991)) 生痰二氧化碳嗜纤维菌(*S. sputigena*) (如GenBank accession no. AEC00832) 软髮菌(*H. hydrossis*) (如GenBank accession no. YP_004451189)
UDP-糖焦磷酸化酶(USP)	在UTP存在下，催化Gal-1-P 转换为UDP-Gal	拟南芥(如GenBank accession no. AF360236.1 利什曼原虫(*L. major*) (如GenBank accession no. ABY79093) 克氏锥虫(*T. cruzi*) (如GenBank accession no. ADD10758) 杜氏利什曼原虫(*L. donovani*) (如GenBank accession no. XP_00385998) 大豆 (如GenBank accession no. NP_001237434)
丙酮酸激酶(PykF)	催化将源自磷酸烯醇丙酮酸盐(PEP)的磷酸基团转移成ADP的过程，以生成丙酮酸及ATP或UTP	大肠杆菌(如GenBank accession no. U00096.2) 汉堡硝化杆菌(*N. hamburgensis*) (如GenBank accession no. YP_576506) 紫色非硫光合菌(*R. palustris*) (如GenBank accession no. YP_7830161) 威廉港鼠尾菌(*M. ruestringensis*) (如GenBank accession no. YP_004787669) 软髮菌(如GenBank accession no. YP_004450514) 球状螺旋体(*S. coccoides*) (如GenBank accession no. YP_00441096)
焦磷酸化酶	作用于二磷	大肠杆菌 (如GenBank accession no. U00096.2

[0110]

酶	活性	实施例
(PPA) (选择性)	酸盐键的酸酐水解酶	蓝隐藻(*G. theta*) (如GenBank accession no. CAI77906) 丁酸梭状芽孢杆菌(*C. butyricum*) (如GenBank accession no. ZP_04525837) 胚芽乳酸杆菌(*L. plantarum*) (如GenBank accession no. EFK28054) 猪双臼乳酸杆菌(*L. suebicus*) (如GenBan accession no. ZP_09451344)

[0111] 如本文所定义用于UDP-Gal再生系统的酶可以是野生型酶。本文所指野生型酶是一见于适当物种的自然存在的酶。某些特定实施例中，GalK、USP、PykF及PPA酶可分别见于大肠杆菌、拟南芥、大肠杆菌及大肠杆菌。来自这些物种的酶的实施例已见于表1。其它则可由本领域技术员容易地确定，例如搜寻公开的基因数据库，例如GenBank。根据本发明的其它实施例，酶可以是上述酶的同系物(homologs)，也属本领域技术员所熟知的范围。例如，该等同系物可利用一实施例酶的氨基酸序列或编码核苷酸序列作为搜寻索引，比对基因库而得。

[0112] 另一方面，涉及UDP-Gal再生系统的酶可以是一对应野生型的功能性变异体。本文所指对应野生型的功能性变异体具有与对应野生型相同的酶活性，且基本上与该对应野生型具有高度氨基酸序列相似性，例如至少与对应野生型具有80％、85％、90％、95或98％的氨基酸序列同一性。两氨基酸序列的“同一性百分比”是利用Karlin及Altschul的运算方法来获得(Proc.Natl.Acad.Sci.USA 87:2264-68,1990)，或经Karlin及Altschul修饰的运算方法来获得(Proc.Natl.Acad.Sci.USA 90:5873-77,1993)。此种运算方法已被并入NBLAST及XBLAST程序(2.0版) (Altschul,等人J.Mol.Biol.215:403-10,1990)。可使用XBLAST程序(score＝50、wordlength＝3)下进行BLAST蛋白质搜寻，来获得所欲氨基酸序列的类似物。当两个序列间有间断时，可使用Altschul等人(Nucleic Acids Res.25(17):3389-3402,1997)所揭示的Gapped BLAST。当使用BLAST及Gapped BLAST程序时，可使用个别程序(如，XBLAST and NBLAST)所提供的预设参数。

[0113] 功能性变异体可以有多种突变，其包括一或多个氨基酸残基的插入、删除或取代。该变异体的突变通常位于对野生型酶的活性不具关键影响性的部位，即在功能性区域未产生突变或仅包含保守氨基酸的取代。本领域技术员能了解，可对硫辛酸连接酶突变体进行保守氨基酸取代能获得功效相等的变异体，也就是，该等变异体保有该特定硫辛酸连接酶突变体的功效。本文所指“保守氨基酸取代”是指一氨基酸取代不致改变该蛋白质的相对电荷或其大小特征。本领域技术员可利用公知的方法来改变多肽序列以制备变异体，制备方法可参考以下文献：例如，分子克隆：实验室手册，J.Sambrook等编，第二版，冷泉港实验室出版社，冷泉港，纽约，1989或当前的分子生物学协议，FM奥苏贝尔等编，约翰威立公司，纽

约(Molecular Cloning:A Laboratory Manual,J.Sambrook,等人,eds.,Second Edition,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,New York,1989或Current Protocols in Molecular Biology,F.M.Ausubel,等人,eds.,John Wiley & Sons,Inc.,New York)。氨基酸的保守性取代包括在以下氨基酸的间的取代:(a)M、I、L、V;(b)F、Y、W;(c)K、R、H;(d)A、G;(e)S、T;(f)Q、N及(g)E、D。

[0114] 任何涉及UDP-Gal再生系统的酶均可通过一般技术而制得。根据本发明的实施例,酶是分离自天然资源。根据本发明的其它实施例,酶是通过一般重组技术而制得。如需要,标靶酶的编码序列可以根据其寄主细胞,作最适化的编码调整。例如,通过大肠杆菌细胞作为寄主制备酶时,编码该酶的基因可经修饰使所含密码适用于大肠杆菌细胞中。

[0115] 如图2A所示,UDP-Gal再生系统可与半乳糖化反应共轭,通过半乳糖转移酶的活性将半乳糖残基添加至适当的底物上。半乳糖转移酶的实施例是例示于表2。

[0116] 表2半乳糖转移酶

[0117]

半乳糖转移酶	酶活性	实施例
β-1,4-半乳糖转移酶(β4GalT),包括异构型1-7 (β-1,4-半乳糖转移	自UDP-Gal催化转移半乳糖至适当受体,例如在β-1,4连接方式下,	*智人*[如, GI: 903740] 褐鼠[如GI: 3258653] 海洋杆菌(*Zobellia*

[0118]

半乳糖转移酶	酶活性	实施例
酶1-7)	末端具有2-乙酰胺-2-脱氧-D-葡糖-基团的糖蛋白或糖脂质受体	*galactanivorans*) [如GI: 340619721] *产气荚膜梭状芽孢杆菌*[如GI: 18309463]
β-1,3-半乳糖转移酶(β3GalT)	自UDP-Gal催化转移半乳糖至适当的受体，例如在β-1,3-连接方式下，末端具有2-乙酰胺-2-脱氧-D-葡糖-基团或GalNAc残基的糖蛋白或糖脂质受体	致倦库蚊(*Culex quinquefasciatus*) [如GI:167873909] 热带家蚊(*Zea mays*) [如GI: 195643406] 肠道螺旋体(*Brachyspira pilosicoli*) [如GI: 300871377] 尿肠球菌(*Enterococcus faecium*) [如GI: 257822935] 来自如流行性感冒嗜血杆菌[L42023.1]的LgtD
α-1,4-半乳糖转移酶(α4GalT) 例如： 乳糖酰基鞘胺醇(Lactosylceramide) 4-α-半乳糖转移酶	自UDP-Gal催化转移半乳糖至适当的受体，例如在α-1,4连接方式下，具有末端半乳糖残基或GlcNAc残基上的糖蛋白或糖脂质受体	智人[如GI:55956926] 地中海雪貂(*Mustela putorius furo*) [如GI: 355666115] 家鼷鼠(*Mus musculus*) [如GI: 51921295] 沟鼠(*Rattus norvegicus*) [如GI: 67677925] 来自如脑膜炎双球菌[如AF355193.1]的LgtC
α-1,3-半乳糖转移酶(α3GalT) **例如：** α-1,3-半乳糖转移酶1 α-1,3-半乳糖转移酶2	自UDP-Gal催化转移半乳糖至适当的受体，例如在α-1,3连接方式下，末端具有半乳糖残基或GlcNAc残基的糖蛋白或糖脂质受体	家鼷鼠[如GI:224922807] 地中海雪貂 [如GI: 355690122] 僧帽猴(*Cebus paella*) [如GI: 19698748] 沟鼠[如GI: 28625949]

[0119] 如前所述的野生型半乳糖转移酶及功能性变异体均属于本案说明的范畴。该等半

乳糖转移酶可通过一般方法而制得。

[0120] UDP-Gal再生系统与一或多种半乳糖转移酶的组合可在高产率下将半乳糖残基添加至适当的底物(受体)。用于半乳糖转移酶的底物为本领域技术员所熟知者(如上述表2所示)。较佳者,该底物具有一个末端糖残基(例如,Gal、GalNAc或GlcNAc),故半乳糖残基可被添加至其上。某些实施例中,该底物为多糖(具有大于50个单糖单元)、寡糖(具有2-50个单糖单元)、糖蛋白或糖多肽或糖脂质。用于本文所述半乳糖化方法的半乳糖转移酶的形式是根据所欲终产物及合成该终产物的底物而定,其是本领域技术员所熟知的范畴。本文所述有关于组合UDP-Gal再生系统/半乳糖转移酶的做法可用于糖鞘脂的合成。实施例是例示于图3。

[0121] 其它实施例中,该等组合UDP-Gal再生系统/半乳糖转移酶的做法可用于合成Globo系列寡糖,例如自乳糖合成Gb3,或自Gb4合成Gb5。请参见图2A及2C。也请参见以下说明。

[0122] UDP-GalNAc再生系统及其N-乙酰半乳糖胺化的用途

[0123] UDP-GalNAc再生系统可与N-乙酰半乳糖胺转移酶(GalNAcT)(例如β-1,3-N-乙酰半乳糖胺转移酶)合并使用,在适当受体上加入GalNAc残基。

[0124] UDP-GalNAc再生系统所涉及的酶的例子是如表3所示:

[0125] 表3:用于UDP-GalNAc再生系统的酶

[0126]

酶	活性	实施例
N-乙酰己糖胺1-激酶(GalNAcK)	藉连续两底物-两产物机制,将ATP及N-乙酰己糖胺转换为ADP及N-乙酰-α-D-己糖胺1-磷酸	来自龙根菌(如GenBank accession no. CP000246.1)的GalNAcK 短双歧杆菌(如GenBank accession no. ZP_06596651) 血溶性隐密杆菌(*A. haemolyticum*)(如GenBank accession no. YP_003696399 双歧乳杆菌(如GenBank accession no. YP_003938776)
N-乙酰葡萄	催化UTP及N-乙	来自大肠杆菌 (如GenBank accession no.

[0127]

酶	活性	实施例
糖胺1-磷酸尿苷酰转移酶(GlmU)	酰-α-D-葡萄糖胺1-磷酸转换为二磷酸及UDP-N-乙酰-D-葡萄糖胺	U00096.2)的GlmU 拟南芥(如GenBank accession no. AEE31311) 比米琴吉欧菌(*G. bemidjiensis*) (如GenBank accession no. ACH37122) 幽门螺旋菌(如GenBank accession no. YP_003728906)
丙酮酸激酶(PykF)	自磷酸烯醇丙酮酸(PEP) 催化转移磷酸基团至ADP，以生成丙酮酸及ATP或UTP	大肠杆菌(如GenBank accession no. U00096.2) 汉堡硝化杆菌(如GenBank accession no. YP_576506) 紫色非硫光合菌(如GenBank accession no. YP_7830161) 威廉港鼠尾菌(如GenBank accession no. YP_004787669) 软髮菌(如GenBank accession no. YP_004450514) 球状螺旋体(如GenBank accession no. YP_00441096)
焦磷酸化酶(PPA) (选择性)	作用于二磷酸盐键的酸酐水解酶	大肠杆菌(如GenBank accession no. U00096.2 蓝隐藻(如GenBank accession no. CAI77906) 丁酸梭状芽孢杆菌(如GenBank accession no. ZP_04525837) 胚芽乳酸杆菌(如GenBank accession no. EFK28054) 猪双臼乳酸杆菌(如GenBan accession no. ZP_09451344)

[0128] N-乙烯半乳糖胺转移酶(例如,β-1,3-GalNAcT或β-1,4-GalNAcT)是用于催化反应的酶,该反应是将GalNAc残基添加至适当的受体,例如肽或寡糖。实施例包括来自流感嗜血杆菌(GenBank accession no.L42023.1.)的LgtD。其它实施例包括,但不限于莱姆病螺旋体(B.garinii)(如GenBank accession no.AEW68905)的LgtD、乳糖奈瑟菌(N.lactamica)(如GenBank accession no.AAN08512)的LgtD及猫立克次体(R.felis)(如GenBank accession no.YP_246702)的LgtD。

[0129] UDP-GalNAc再生系统/GalNAcT组合做法中所使用的任何酶可为如本文所述的野

生型酶或其功能性变异体。任何传统方法可用于制备这种酶。根据本发明的一实施例，该做法是用于自Gb3合成Gb4，请参见图2B。

[0130] GDP-Fuc再生系统及其岩藻糖化的用途

[0131] GDP-Fuc再生系统可与岩藻糖转移酶(例如，α-1,2-岩藻糖转移酶、α-1,3-岩藻糖转移酶或α-2,6-岩藻糖转移酶)合并使用，将岩藻糖残基添加至适当的受体(例如寡糖)，该受体可共轭至另一分子，例如脂肪或多肽。

[0132] GDP-Fuc再生系统所涉及的酶的例子是如表4所示：

[0133] 表4：用于GDP-Fuc再生系统的酶

[0134]

酶	活性	实施例
L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶(FKP)	可将岩藻糖及ATP产生Fuc-1-P及GDP-Fuc的生物功能性酶	脆弱拟杆菌 (如GenBank accession no. CR626927.1) 智人(如GenBank accession no. NP_003829) 沟鼠(如GenBank accession no. NP_955788)
丙酮酸激酶(PykF)	自磷酸烯醇丙酮酸(PEP) 催化转移磷酸基团至ADP，以生成丙酮酸及ATP或UTP	大肠杆菌(如GenBank accession no. U00096.2) 汉堡硝化杆菌(如GenBank accession no. YP_576506) 紫色非硫光合菌(如GenBank accession no. YP_7830161) 威廉港鼠尾菌(如GenBank accession no. YP_004787669) 软髮菌(如GenBank accession no. YP_004450514) 球状螺旋体(如GenBank accession no. YP_00441096)
焦磷酸化酶(PPA) (选择	作用于二磷酸盐键的酸酐水解酶	大肠杆菌(如GenBank accession no. U00096.2 蓝隐藻(如GenBank accession no. CAI77906)

[0135]

酶	活性	实施例
性)		丁酸梭状芽孢杆菌(如GenBank accession no. ZP_04525837) 胚芽乳酸杆菌(如GenBank accession no. EFK28054) 猪双臼乳酸杆菌(如GenBan accession no. ZP_09451344)

[0136] 岩藻糖转移酶可将来自GDP-岩藻糖(鸟苷二磷酸盐-岩藻糖)供体底物的L-岩藻糖转移至适当的受体,该受体可为另一糖。在N-连结糖化反应下,岩藻糖转移酶可将岩藻糖残基添加至核心GlcNAc(N-乙酰葡萄糖胺)糖,或在O-连结糖化反应下,将该岩藻糖添加至蛋白质。岩藻糖转移酶包括α-1,3-岩藻糖转移酶、α-1,2-岩藻糖转移酶及α-1,6-岩藻糖转移酶。根据本发明的实施例包括来自大肠杆菌(如GenBank accession no.U00096.2)的α-1,2-岩藻糖转移酶、来自脆弱拟杆菌(如GenBank accession no.YP_213404)的α-1,3-岩藻糖转移酶及来自有爪蟾蜍(X.laevis)(如GenBank accession no.NP_001083664)的α-1,6-岩藻糖转移酶。

[0137] GDP-Fuc再生系统/FucT组合做法中所使用的任何酶可为如本文所述的野生型酶或其功能性变异体。该等酶可用任何传统方法制备。根据本发明的实施例,该做法是用于自Gb3合成Gb4,请参见图2D。

[0138] CMP-Neu5Ac再生系统及其唾液酸化(Sialylation)的用途

[0139] CMP-Neu5Ac再生系统可与唾液酸转移酶(如α-2,3-唾液酸转移酶)组合以将唾液酸残基(Neu5Ac)加至适当受体底物(如寡糖)。

[0140] CMP-Neu5Ac再生系统所涉及的酶例示于表5:

[0141] 表5:用于CMP-Neu5Ac再生系统的酶

[0142]

酶	活性	实施例
胞苷单磷酸激酶(CMK)	催化CMP的磷酸化反应以生成CDP	大肠杆菌 (如GenBank accession no. U00096.2) 芽孢枯草杆菌(*B. amyloliquefaciens*) (如GenBank accession no. ABS74466) 麻疯分枝杆菌(*M. leprae*) (如, GenBank accession no. CAB08279)

[0143]

酶	活性	实施例
		鸟结核杆菌(*M. avium*) (如GenBank accession no. AAS03731) 莱姆病螺旋体 (如GenBank accession no. AEW68468)
CMP-唾液酸合成酶(Css)	将CTP及唾液酸催化合成CMP唾液酸	多杀性巴氏杆菌 (如GenBank accession no. AE004439.1 脑膜炎双球菌 (如GenBank accession no. AAB60780) 虹鳟(*O. mykiss*) (如GenBank accession no. BAB47150) 艾欧翰氏海洋菌(*I. ioihiensis*) (如GenBank accession no. AAV81361) 空肠曲杆菌(*C. jejuni*) (如GenBank accession no. ABI32334)
丙酮酸激酶(PykF)	自磷酸烯醇丙酮酸(PEP)催化转移磷酸基团至ADP，以生成丙酮酸及ATP或UTP	大肠杆菌 (如GenBank accession no. U00096.2) 汉堡硝化杆菌 (如GenBank accession no. YP_576506) 紫色非硫光合菌 (如GenBank accession no. YP_7830161) 威廉港鼠尾菌 (如GenBank accession no. YP_004787669) 软髮菌 (如GenBank accession no. YP_004450514) 球状螺旋体 (如GenBank accession no. YP_00441096)
焦磷酸化酶(PPA) (选择性)	作用于二磷酸盐键的酸酐水解酶	大肠杆菌 (如GenBank accession no. U00096.2 蓝隐藻 (如GenBank accession no. CAI77906) 丁酸梭状芽孢杆菌 (如GenBank accession no. ZP_04525837) 胚芽乳酸杆菌 (如GenBank accession no. EFK28054)

[0144]

酶	活性	实施例
		猪双曰乳酸杆菌 (如GenBan accession no. ZP_09451344)

[0145] 唾液酸转移酶是将唾液酸转移至初生(nascent)寡糖的酶。该酶家族将唾液酸添加至唾液酸化糖脂(神经节苷脂)的末端部位添加至糖蛋白的N-或O-联糖链。约有20种不同的唾液酸转移酶,包括以α-2,3键结唾液酸至半乳糖的唾液酸转移酶(例如,α-2,3-唾液酸转移酶)及以α-2,6键结唾液酸至半乳糖或N-乙酰半乳糖胺的唾液酸转移酶(例如,α-2,6-唾液酸转移酶)。根据本发明的实施例包括来自海洋细菌(M.bacteria)(GenBank accession no.AB308042.1)、小家鼠(如GenBank accession no.BAA06068)或多杀性巴氏杆菌(如GenBank accession no.AET17056)的α-2,3-唾液酸转移酶以及来自家牛(B.Taurus)(如GenBank accession no.NP_001008668)、灰仓鼠(C.griseus)(如GenBank accession no.NP_001233744)或沟鼠(如GenBank accession no.AAC42086)的α-2,6-唾液酸转移酶。

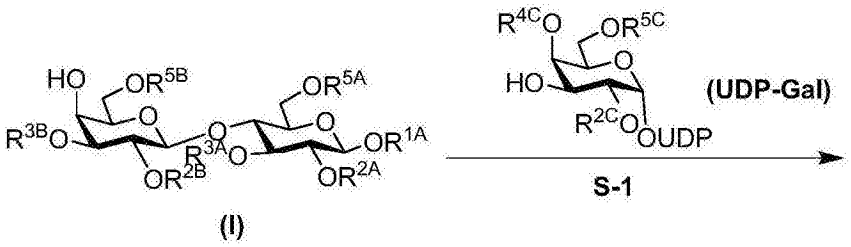
[0146] CMP-Neu5Ac再生系统/唾液酸转移酶组合做法所使用的酶可为如本文所述的野生型酶或其功能性变异体。这种酶可用任何传统方法制备。根据本发明的实施例,该做法是用于自Gb3合成Gb4,请参见图2E。

[0147] Globo系列寡糖的合成

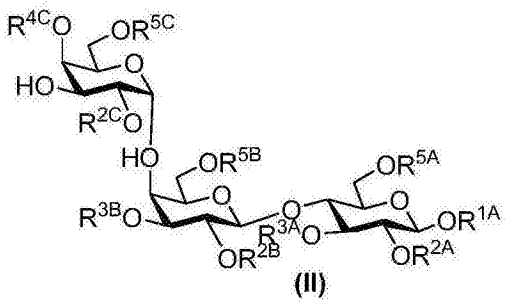
[0148] 上述组合做法包括UDP-Gal再生/半乳糖转移酶、UDP-GalNAc再生/GalNAcT、GDP-Fuc再生/岩藻糖转移酶及CMP-Neu5Ac再生/唾液酸转移酶,其可单独地或经组合地应用,以合成Globo系列寡糖,包括Gb3、Gb4、Gb5、Globo H(岩藻糖-Gb5)及SSEA4(唾液酸-Gb5)。如同下述更详细的说明,所有Globo-系列寡糖可以是未经取代或经取代的。

[0149] 步骤S-1

[0150] 生物合成方法中的第一步骤(S-1)是将式(I)化合物或其盐经酶性反转为式(II)化合物或其盐:



[0151]



[0152] 其中R^{1A}为氢、经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经

取代的炔基、经取代或未经取代的环碳基、经取代或未经取代的杂环基、经取代或未经取代的芳基、经取代或未经取代的杂芳基或氧保护基团。R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}及R^{5C}是各自独立地为氢、经取代或或经取代的C_{1-6}烷基或氧保护基团。

[0153] 因此，本发明的一方面是提供一种自式(I)化合物或其盐酶性合成式(II)化合物或其盐的方法，其包括在尿苷二磷酸-Gal(UDP-Gal)及α-1,4-半乳糖转移酶存在下，转换式(I)化合物为式(II)化合物，以及在上述表1所示的酶组存在下，自半乳糖生成UDP-Gal，请见图2A。为实施该酶性反应，可将必要成份(如半乳糖、半乳糖转移酶、UDP-Gal再生酶组、ATP、UTP及其它(例如Mg^{++}))混合成一反应混合物，置于允许生成式(II)化合物的条件下培养。这种条件是本领域技术人员所熟知的。请参见以下实施例说明。

[0154] R^{1A}基团可作为一功能性基团，允许Globo-系列寡糖共轭至另一分子，例如蛋白质或脂质。另一方面，其可作为一保护基团。

[0155] 某些特定实施例中，R^{1A}为氢。

[0156] 在其它实施例中，R^{1A}为经取代或未经取代的烷基，例如经取代或未经取代的C_{1-6}烷基、经取代或未经取代的C_{2-6}烷基、经取代或未经取代的C_{3-6}烷基、经取代或未经取代的C_{4-6}烷基、经取代或未经取代的C_{5-6}烷基、经取代或未经取代的C_{2-5}烷基、经取代或未经取代的C_{2-4}烷基、经取代或未经取代的C_{2-3}烷基、经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。如本文所定义的生物素及神经酰胺是被经取代的烷基所涵括。某些特定实施例中，R^{1A}为未经取代的烷基。例如在一些特定实施例中，R^{1A}为甲基、乙基、丙基、异丙基、第二丁基、异丁基或第三丁基。另一方面，某些特定实施例中，R^{1A}为经取代的烷基。在一些特定实施例中，R^{1A}是进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代。某些特定实施例中，该等取代基的取代位置是位于烷基的末端(最后一个碳原子)。某些特定实施例中，R^{1A}为经一或多个氨基(-NH2)基团所取代的烷基。某些特定实施例中，R^{1A}为在末端位置(最后一个碳原子)经氨基($-NH_2$)基团取代的烷基。某些特定实施例中，R^{1A}为$-(CH_2)_n-NH_2$，其中n为1、2、3、4、5或6。某些特定实施例中，R^{1A}为5-戊基氨基($-(CH_2)_5-NH_2$)。

[0157] 某些特定实施例中，R^{1A}为经取代或未经取代的烯基，例如经取代或未经取代的C_{2-6}烯基、经取代或未经取代的C_{3-6}烯基、经取代或未经取代的C_{4-6}烯基、经取代或未经取代的C_{5-6}烯基、经取代或未经取代的C_{2-5}烯基、经取代或未经取代的C_{2-4}烯基、经取代或未经取代的C_{2-3}烯基、经取代或未经取代的C_2烯基、经取代或未经取代的C_3烯基、经取代或未经取代的C_4烯基、经取代或未经取代的C_5烯基或经取代或未经取代的C_6烯基。某些特定实施例中，R^{1A}为$-(CH_2)_m-CH=CH_2$，其中n为1、2或3。某些特定实施例中，R^{1A}为丙烯基($-CH_2CH=CH_2$)。某些特定实施例中，R^{1A}为进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代的烯基。某些特定实施例中，该等取代基的取代位置是位于烯基的末端(最后一个碳原子)。

[0158] 某些特定实施例中，R^{1A}为经取代或未经取代的炔基，例如经取代或未经取代的C_{2-6}炔基、经取代或未经取代的C_{3-6}炔基、经取代或未经取代的C_{4-6}炔基、经取代或未经取代的C_{5-6}炔基、经取代或未经取代的C_{2-5}炔基、经取代或未经取代的C_{2-4}炔基、经取代或未经取代

的C_{2-3}炔基、经取代或未经取代的C_2炔基、经取代或未经取代的C_3炔基、经取代或未经取代的C_4炔基、经取代或未经取代的C_5炔基或经取代或未经取代的C_6炔基。某些特定实施例中，R^{1A}为进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代的炔基。某些特定实施例中，这些取代基的取代位置是位于炔基的末端(最后一个碳原子)。

[0159] 某些特定实施例中，R^{1A}为经取代或未经取代的杂环基，例如经取代或未经取代的5-至8-元杂环基、经取代或未经取代的5-至7-元杂环基、经取代或未经取代的5-至6-元杂环基、经取代或未经取代的5-元杂环基、经取代或未经取代的6-元杂环基、经取代或未经取代的7-元杂环基或经取代或未经取代的8-元杂环基。某些特定实施例中，R^{1A}为进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代的杂环基。

[0160] 某些特定实施例中，R^{1A}为经取代或未经取代的环碳基，例如经取代或未经取代的C_{3-6}碳环基、经取代或未经取代的C_{3-5}碳环基、经取代或未经取代的C_{3-4}碳环基、经取代或未经取代的C_3碳环基、经取代或未经取代的C_4碳环基、经取代或未经取代的C_5碳环基或经取代或未经取代的C_6碳环基。某些特定实施例中，R^{1A}为进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代的碳环基。

[0161] 某些特定实施例中，R^{1A}为经取代或未经取代的芳基，例如经取代或未经取代的C_6芳基(苯基)或经取代或未经取代的C_{10}芳基(萘基)。某些特定实施例中，R^{1A}为进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代的芳基。

[0162] 某些特定实施例中，R^{1A}为经取代或未经取代的杂芳基，例如经取代或未经取代的5-元杂芳基或经取代或未经取代的6-元杂芳基。某些特定实施例中，R^{1A}为进一步被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代的杂芳基。

[0163] 某些特定实施例中，R^{1A}为氢、丙烯基、经取代的烷基、生物素或神经酰胺。

[0164] 进一步而言，R^{1A}可为上述任一非氢基团的混合物，例如经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经取代的炔基、经取代或未经取代的碳环基、经取代或未经取代的杂环基、经取代或未经取代的芳基、经取代或未经取代的杂芳基，以提供由包含1、2、3、4、5、6、7、8、9或10种不同基团所组合的键结基团。作为一非限制性的实施例，R^{1A}可为由包含烷基及芳基基团所组合的键结基团，例如烷基-芳基-烷基，且其是视需要进一步在该键结基团的任一位置(例如末端位置)被经取代或未经取代的硫、经取代或未经取代的氨基、羰基(例如，羧酸)、迭氮、烯基(例如，丙烯基)、炔基(例如，丙炔基)、生物素或神经酰胺基团所取代。

[0165] 某些特定实施例中，R^{1A}是如本文所定义的氧保护基团。

[0166] 某些特定实施例中，R^{2A}为氢。某些特定实施例中，R^{2A}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{2A}为氧保护基团。

[0167] 某些特定实施例中，R^{3A}为氢。某些特定实施例中，R^{3A}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{3A}为氧保护基团。

[0168] 某些特定实施例中，R^{5A}为氢。某些特定实施例中，R^{5A}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{5A}为氧保护基团。

[0169] 某些特定实施例中，R^{2B}为氢。某些特定实施例中，R^{2B}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{2B}为氧保护基团。

[0170] 某些特定实施例中，R^{3B}为氢。某些特定实施例中，R^{3B}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{3B}为氧保护基团。

[0171] 某些特定实施例中，R^{5B}为氢。某些特定实施例中，R^{5B}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{5B}为氧保护基团。

[0172] 某些特定实施例中，R^{2C}为氢。某些特定实施例中，R^{2C}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{2C}为氧保护基团。

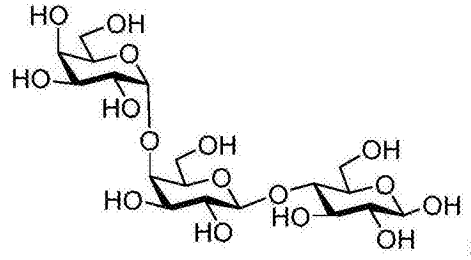
[0173] 某些特定实施例中，R^{4C}为氢。某些特定实施例中，R^{4C}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{4C}为氧保护基团。

[0174] 某些特定实施例中，R^{5C}为氢。某些特定实施例中，R^{5C}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{5C}为氧保护基团。

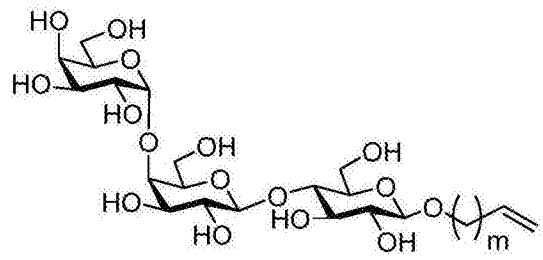
[0175] 某些特定实施例中，R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}及R^{5C}是各自独立地为氢。某些特定实施例中，R^{1A}为经取代或未经取代的烯基，且R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}及R^{5C}是各自独立地为氢。某些特定实施例中，R^{1A}为经取代或未经取代的烷基，且R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}及R^{5C}是各自独立地为氢。

[0176] 式(I)化合物的实施例包括，但不限于其盐类。

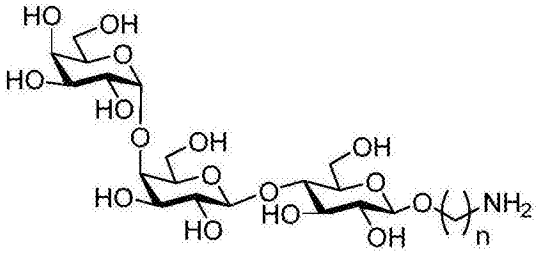
[0177] 式(II)化合物的实施例包括，但不限于以下：



，

[0178] 

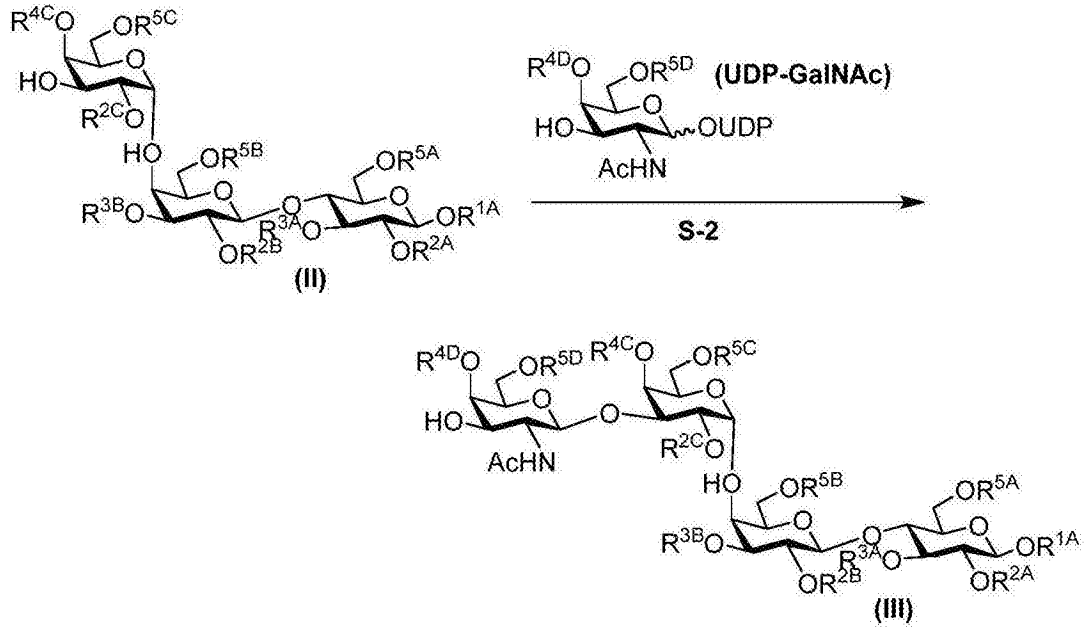
，及



[0179] 及其盐类。

[0180] 步骤S-2

[0181] 生物合成方法的第二步骤(S-2)是将式(II)化合物或其盐经酶性转换为式(III)化合物或其盐：

[0182] 

[0183] 其中R^{1A}、R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}及R^{5C}是如本文所定义的，且其中R^{4D}及R^{5D}各自独立地为氢、经取代或未经取代的C_{1-6}烷基或氧保护基团。

[0184] 某些特定实施例中，R^{2D}为氢。某些特定实施例中，R^{2D}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3

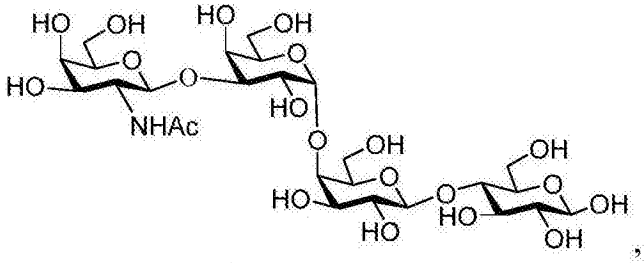
烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{2D}为氮保护基团，例如乙酰基(Ac；-C＝OCH_3)。

[0185] 某些特定实施例中，R^{4D}为氢。某些特定实施例中，R^{4D}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{4D}为氧保护基团。

[0186] 某些特定实施例中，R^{5D}为氢。某些特定实施例中，R^{5D}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{5D}为氧保护基团。

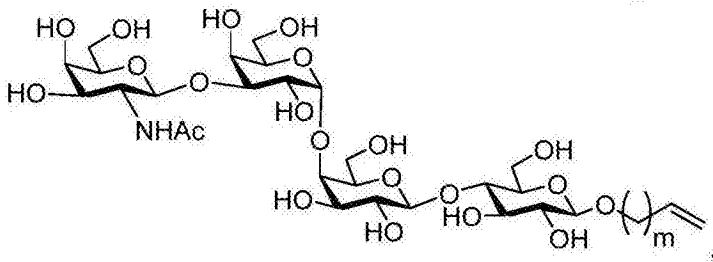
[0187] 某些特定实施例中，R^{4D}及R^{5D}均为氢。某些特定实施例中，R^{2D}为氮保护基团，例如乙酰基(Ac；-C＝OCH_3)，且R^{4D}及R^{5D}各自为氢。

[0188] 式(III)化合物的实施例包括，但不限于以下：



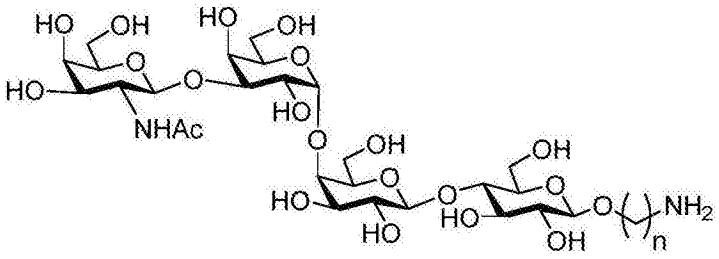
，

[0189]



，及

[0190]

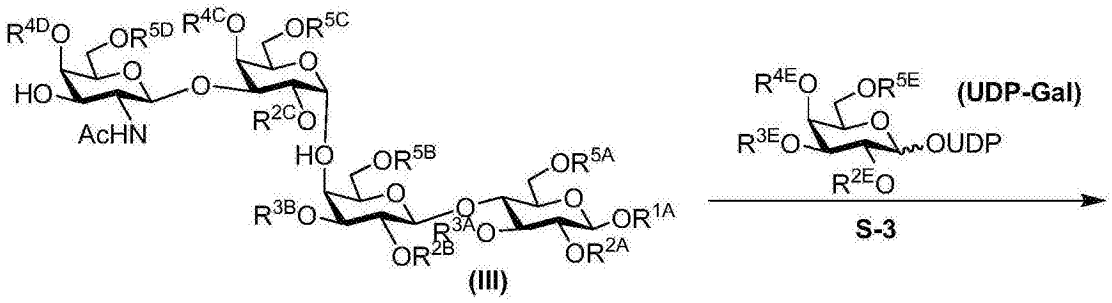


[0191] 及其盐。

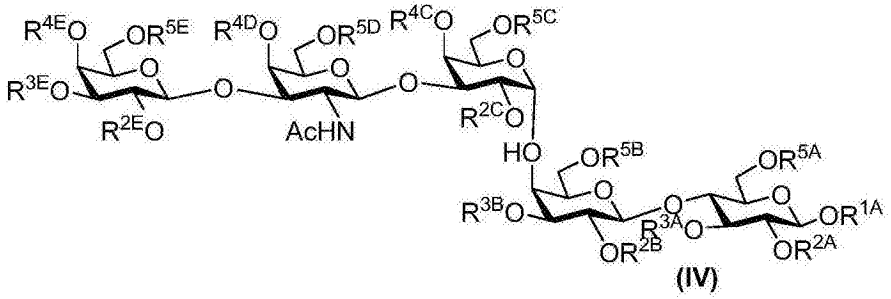
[0192] 步骤(S-2)的方法，是在合适的条件下，将式(II)化合物或其盐进行酶性合成式(III)化合物或其盐。式(II)化合物的底物可由任何本领域公知的方法或本文所述的方法制得。某些实施例中，式(II)化合物是分离自上述步骤S-1中的反应混合物。其它实施例中，在无须纯化式(II)化合物的情形下，可直接使用步骤S-1的整个反应混合物。在GalNAcT(例如，β-1,3-GalNAcT)存在及下，式(II)化合物可与UDP-GalNAc培养在允许式(II)化合物转换为式(III)化合物的条件下。某些实施例中，该GalNAcT-催化反应是与如本文所述的UDP-GalNAc再生方法予以结合。如图2B所示。请见以下实施例说明。

[0193] 步骤S-3

[0194] 生物合成方法的第三步骤(S-3)是将式(III)化合物或其盐经酶性转换为式(IV)化合物或其盐：



[0195]



[0196] 其中R^{1A}、R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}、R^{5C}、R^{2D}、R^{4D}及R^{5D}是如本文所定义者，且R^{2E}、R^{3E}、R^{4E}及R^{5E}是各自独立地为氢、经取代或未经取代的C_{1-6}烷基或氧保护基团。

[0197] 某些特定实施例中，R^{2E}为氢。某些特定实施例中，R^{2E}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例，R^{2E}为氧保护基团。

[0198] 某些特定实施例中，R^{3E}为氢。某些特定实施例中，R^{3E}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例，R^{3E}为氧保护基团。

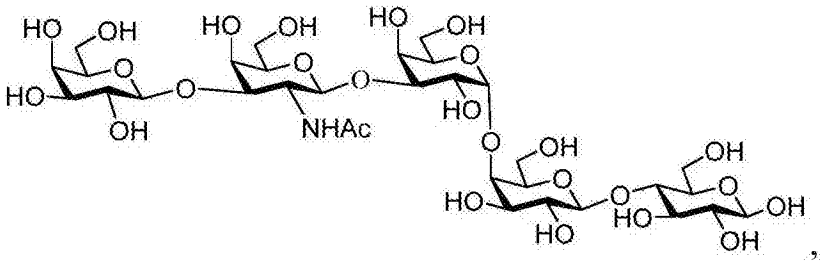
[0199] 某些特定实施例中，R^{4E}为氢。某些特定实施例中，R^{4E}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{4E}为氧保护基团。

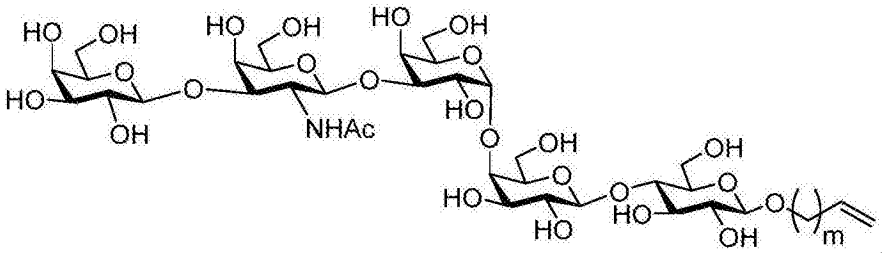
[0200] 某些特定实施例中，R^{5E}为氢。某些特定实施例中，R^{5E}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{5E}为氧保护基团。

[0201] 某些特定实施例中，R^{3E}为氢。某些特定实施例中，R^{3E}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例，R^{3E}为氧保护基团。

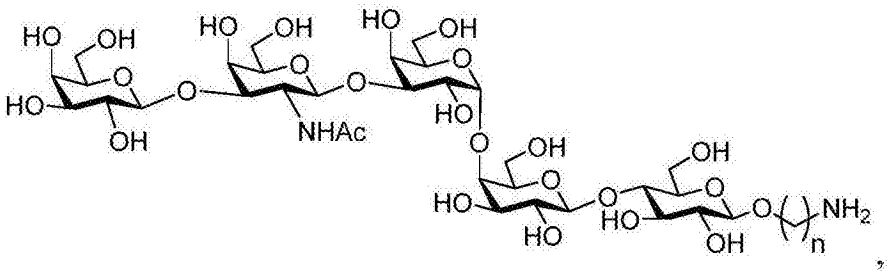
[0202] 某些特定实施例中，R^{2E}、R^{3E}、R^{4E}及R^{5E}各自为氢。

[0203] 式(IV)化合物的实施例包括,但不限于以下:

[0204] 



及

[0205] 

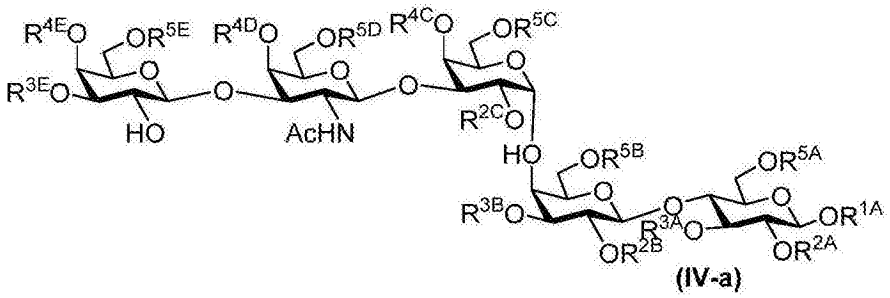
[0206] 及其盐。

[0207] 步骤S-3是有关β-1,3-半乳糖转移酶活性所致的酶反应,其是在本领域技术人员所熟知的条件下进行。式(III)化合物的底物,例如Gb4,可由任何本领域公知的方法或本文所述的方法制得。某些实施例中,式(III)化合物是分离自上述步骤S-2中的反应混合物。其它实施例中,在无须纯化式(III)化合物的情形下,可直接使用步骤S-2的整个反应混合物。在β-1,3-半乳糖转移酶存在下,式(III)化合物可与UDP-Gal培养在允许转换式(III)化合物为式(IV)化合物的条件下。某些实施例中,该GalT-催化反应是与如本文所述的UDP-Gal再生方法予以结合。如图2A所示。请见以下实施例说明。

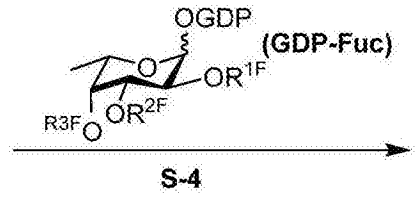
[0208] 某些特定实施例中,β-1,3-GalNAcT/β-1,3-GalT双功能性酶(例如来自如流感嗜血杆菌的LgtD)可用于步骤S-2及步骤S-3。

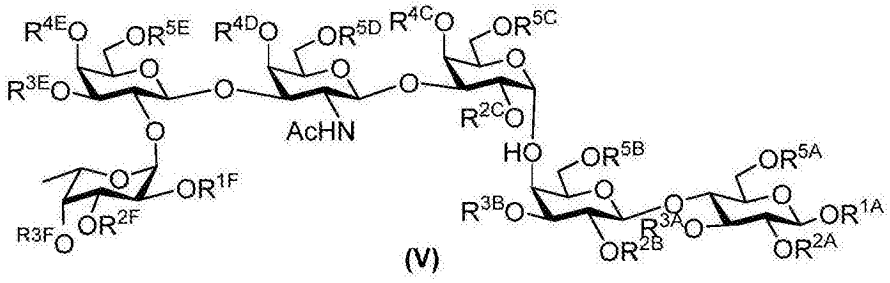
[0209] 步骤S-4

[0210] 式(IV)化合物可在其末端环E的不同位置上进行取代。例如,式(IV)化合物的某些特定方面,其中R^{2E}为氢,在步骤S-4生物合成过程的一视需要步骤,涉及转换式(IV-a)化合物或其盐为式(V)化合物或其盐。



[0211]





[0212] 其中R^{1A}、R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}、R^{5C}、R^{2D}、R^{4D}、R^{5D}、R^{3E}、R^{4E}及R^{5E}如本文所定义，且R^{1F}、R^{2F}及R^{3F}各自独立地为氢、经取代或未经取代的C_{1-6}烷基或氧保护基团。

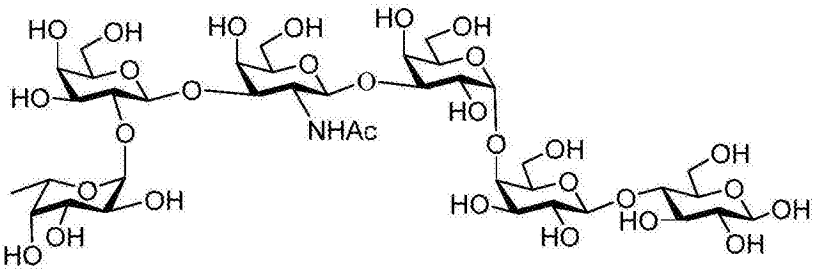
[0213] 某些特定实施例中，R^{1F}为氢。某些特定实施例中，R^{1F}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{1F}为氧保护基团。

[0214] 某些特定实施例中，R^{2F}为氢。某些特定实施例中，R^{2F}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{2F}为氧保护基团。

[0215] 某些特定实施例中，R^{3F}为氢。某些特定实施例中，R^{3F}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{3F}为氧保护基团。

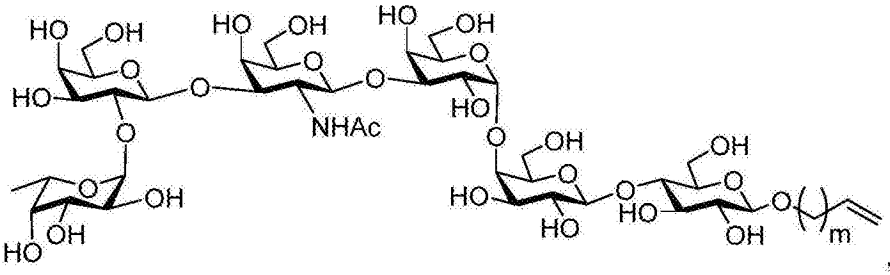
[0216] 某些特定实施例中，R^{1F}、R^{2F}及R^{3F}各自为氢。

[0217] 式(V)化合物的实施例包括，但不限于以下：

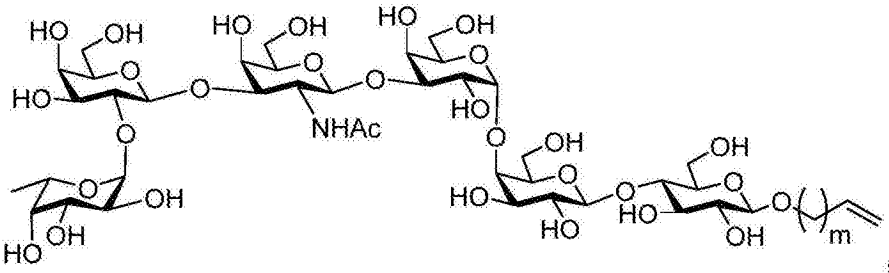


，

[0218]



，及



，

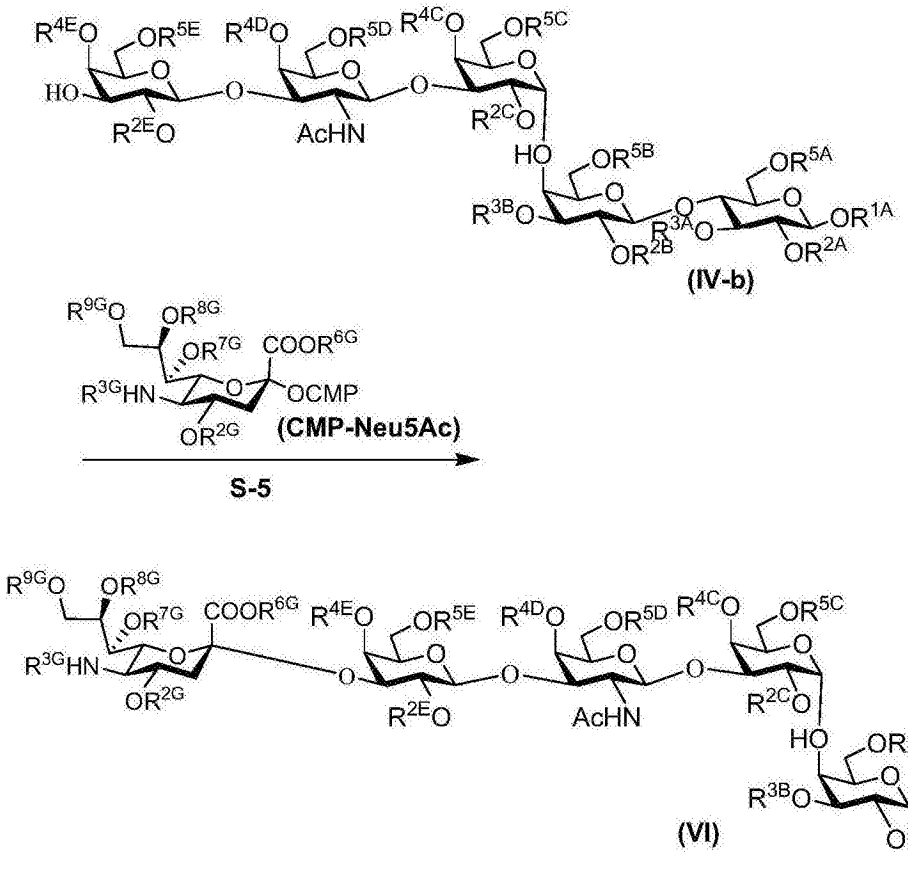
[0219] 及其盐。

[0220] 步骤S-4涉及有关α-1,2-岩藻糖转移酶活性所致的酶反应，其是在本领域技术人员所熟知的条件下进行。式(IV)化合物的底物，例如Gb5，可由任何本领域公知的方法或本文所述的方法制得。某些实施例中，式(IV)化合物是分离自上述步骤S-3中的反应混合物。其它实施例中，在无须纯化式(V)化合物的情形下，可直接使用步骤S-3的整个反应混合物。在岩藻糖转移酶存在下，式(IV)化合物可与GDP-Fuc培养在允许转换式(IV)化合物为式(V)化合物的条件下。某些实施例中，该FucT-催化反应是与如本文所述的UDP-Fuc再生方法予以结合，如图2D所示。请见以下实施例说明。

[0221] 步骤S-5

[0222] 在式(IV)化合物的其它方面中，R^{3E}为氢，在步骤S-5生物合成过的一视需要步骤，涉及转换式(IV-b)化合物或其盐为式(VI)化合物或其盐：

[0223]



[0224] 其中R^{1A}、R^{2A}、R^{3A}、R^{5A}、R^{2B}、R^{3B}、R^{5B}、R^{2C}、R^{4C}、R^{5C}、R^{2D}、R^{4D}、R^{5D}、R^{2E}、R^{4E}及R^{5E}如本文所定义者;R^{3G}为氢、经取代或未经取代的C_{1-6}烷基或氮保护基团;且R^{6G}、R^{7G}、R^{8G}及R^{9G}各自独立地为氢、经取代或未经取代的C_{1-6}烷基或氧保护基团。

[0225] 某些特定实施例中,R^{3G}为氢。某些特定实施例中,R^{3G}为经取代或未经取代的C_{1-6}烷基,例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中,R^{3G}为氮保护基团,例如乙酰基(Ac;$-C(=O)CH_3$)。

[0226] 某些特定实施例中,R^{6G}为氢。某些特定实施例中,R^{6G}为经取代或未经取代的C_{1-6}烷基,例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中,R^{6G}为氧保护基团。

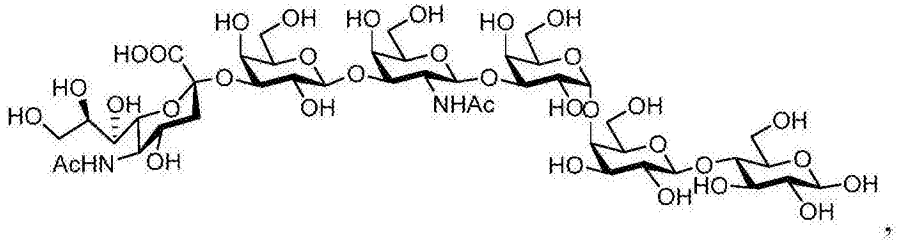
[0227] 某些特定实施例中,R^{7G}为氢。某些特定实施例中,R^{7G}为经取代或未经取代的C_{1-6}烷基,例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中,R^{7G}为氧保护基团。

[0228] 某些特定实施例中,R^{8G}为氢。某些特定实施例中,R^{8G}为经取代或未经取代的C_{1-6}烷基,例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中,R^{8G}为氧保护基团。

[0229] 某些特定实施例中，R^{9G}为氢。某些特定实施例中，R^{9G}为经取代或未经取代的C_{1-6}烷基，例如经取代或未经取代的C_1烷基、经取代或未经取代的C_2烷基、经取代或未经取代的C_3烷基、经取代或未经取代的C_4烷基、经取代或未经取代的C_5烷基或经取代或未经取代的C_6烷基。某些特定实施例中，R^{9G}为氧保护基团。

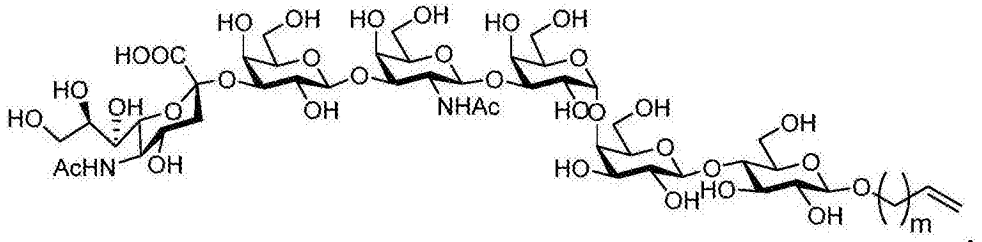
[0230] 某些特定实施例中，R^{6G}、R^{7G}、R^{8G}及R^{9G}各自为氢。某些特定实施例中，R^{3G}为氮保护基团，例如乙酰基(Ac；-C=OCH$_3$)；且R^{6G}、R^{7G}、R^{8G}及R^{9G}各自为氢。

[0231] 式(V)化合物的实施例包括，但不限于以下：



，

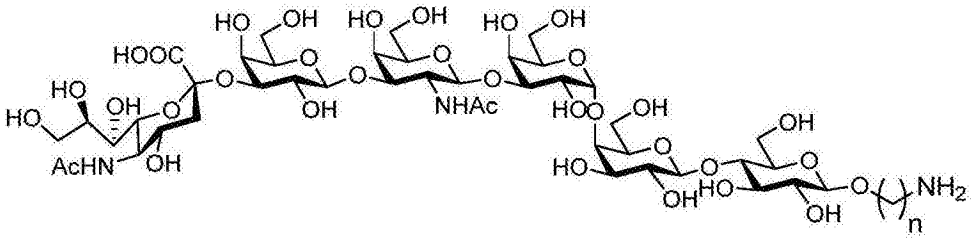
[0232]



，

及

[0233]



[0234] 及其盐。

[0235] 步骤S-5有关α-2,3-唾液酸转移酶活性所致的酶反应，其是在本领域技术人员所熟知的条件下进行。式(IV)化合物的底物，例如Gb5，可由任何本领域公知的方法或本文所述的方法制得。某些实施例中，式(IV)化合物是分离自上述步骤S-3中的反应混合物。其它实施例中，在无须纯化式(IV)化合物的情形下，可直接使用步骤S-3的整个反应混合物。在唾液酸转移酶存在下，式(IV)化合物可与CMP-Neu5Ac培养在允许转换式(IV)化合物为式(V)化合物的条件下。某些实施例中，该唾液酸转移酶-催化反应是与如本文所述的CMP-Neu5Ac再生方法予以结合。如图2E所示。请见本文以下的实施例说明。

[0236] 步骤S-1至S-5的个别及如上所述的任何组合均属本发明所揭示的范畴。本发明也包括以上述任何合成方法所制得的任何化合物，例如以上所述的化合物。

[0237] 根据本发明某些特定实施例，本发明揭示一种自乳糖经由包括上述步骤S-1、S-2、S-3及S-4的链反应，或经由上述步骤S-1、S-2、S-3、S-4或S-5来合成Globo H或SSEA4。该Globo H或SSEA4可以是非尾式(untailed)(R^{1A}为氢；请见图3及4)或尾式(例如，R^{1A}为丙烯基；请见图5及6)。在每一步骤中，半乳糖转移酶反应可与对应的核苷酸糖再生方法结合。请见图3-6。根据本发明的一实施例，上述方法是以一炉式(one-pot manner)来实施，也就是，

每一个前反应混合物直接用于下一步反应，而无需纯化前一反应所得的底物。也就是说，该一炉式方法免除纯化任何中间物的步骤。或者，步骤S-1及步骤S-2是以一炉式在无须纯化任何中间物的情形下实施。步骤S-2之后，Gb4是分离自反应混合物，且将该纯化的GB4用于接着的步骤S-3、S-4及/或S-5。无须进一步纯化其它中间物。

[0238] 用于每一反应步骤的酶可各自溶解于反应混合物中，或固定于一或多个支持体上。如需要，在链反应过程中可添加额外的酶。

[0239] 酶反应器

[0240] 包括上述二或以上步骤的任何组合的酶性链反应，可以在酶反应器中进行，该反应器包括一或多个反应室。每个反应室是设计为进行该链反应的一个步骤。更具体的，每个反应室各自包含上述步骤S-1至S-5中所涉及的酶。

[0241] 某些实施例中，每个反应室中的一或多种酶，或全部的酶是固定于适当的支持体上(例如支持膜)。如需要，连续反应步骤的反应室可以连结，以至于当前反应室的酶反应完成后，所得的反应混合物可流动至下一反应室来进行下个反应步骤。某些实施例中，前反应的产物是未经纯化，将包括产物的整个反应混合物加入至下个反应室中，来进行下个酶性反应。请见图3及4。

[0242] 例如，步骤S-1的反应可在酶反应器中的第一个反应室中进行，其中涉及步骤S-1的一或多种酶是固定于支持体上。步骤S-1结束后，将第一反应室中的反应混合物(包括Gb3产物)置入第二反应室(包含步骤S-2合成Gb4所需的所有酶及试剂)。在一实施例中，将Gb4纯化后用于下个反应步骤。在另一实施例中，将第二反应室中的整个反应混合物(包括Gb4)置入第三反应室(包含步骤S-3合成Gb5所需的所有酶及试剂)。此后，可将第三反应室中的反应混合物置入第四反应室(包含步骤S-4所需的所有酶及试剂)，或置入第五反应室(包含步骤S-5所需的所有酶及试剂)。

[0243] 其它实施例中，酶反应器包含一个反应室，该反应室含有上述合成步骤的一所需的酶、试剂及适当的底物。该底物是固定于支持体上。根据本发明的特定实施例，反应室含有步骤S-1所需的酶及试剂，其中底物(Lac-丙烯基)是固定于支持体上。Gb3-丙烯基在步骤S-1中被合成后，反应室中的反应混合物被第二反应混合物取代，该第二反应混合物包含步骤S-2所需的酶及试剂。Gb4-丙烯基在步骤S-2中被合成后，第二反应混合物被包含步骤S-3(合成Gb5-丙烯基)所需的所有酶及试剂的第三反应混合物所取代。此后，第三反应混合物是被包含步骤S-4(合成Globo H-丙烯基)所需的酶及试剂的第四反应物，或包含步骤S-5(合成SSEA4-丙烯基)所需的酶及试剂的第五反应物所取代。

[0244] 无须进一步的阐述，本领域技术员即可根据上述的说明充分的实施本发明。因此，以下特定实施例是用以说明，而非以任何方式限制本发明。本文中所揭示的所有参考文献是根据其目的或内容以引用的方式并入本文中。

[0245] 实施例

[0246] 以下提供的实施例将有助于了解本发明的各种方面，该等实施例是说明本发明的特定实施方面，并非用于限制本案所请的权利要求书。

[0247] 实施例1:Globo系列寡糖的合成

[0248] UDP-Gal再生的新方法

[0249] 在2004年，Kotake氏等人自豌豆苗发现UDP-糖焦磷酸化酶，该酶对于不同单糖-1-

磷酸具有宽广的底物特异性，可形成UDP-糖。[19]两年后，Kotake氏等人及Somers氏等人分别发表类似功能的酶，USP，其存在于阿拉伯芥(Arabidopsis)中。[20],[21]最近，同源酶也被证实存在于寄生生物利什曼原虫(Leishmania)及锥虫(Trypanosoma)中。[22],[23]USP酶备受瞩目，因为它不仅保有使UTP与Glc-1-磷酸及Gal-1-磷酸浓缩(condense)的能力，尚具有可与其它单糖-1-P、GlcA-1-磷酸及Xyl-1-磷酸浓缩的能力。因此，我们选择USP来直接将Gal-1-磷酸与UTP浓缩后进行UDP-Gal再生，进而达成UDP-Gal合成的第三再生。

[0250] 丙烯基-Gb3的合成

[0251] 反应混合物(200mL)包含溶在100mM Tris-HCl缓冲液(pH 7.0)中的丙烯基-Lac(10mmol)、半乳糖(10mmol)、磷酸烯醇丙酮酸(PEP)(22mmol)、ATP(0.05mmol)、UTP(0.125mmol)及10mM $MgCl_2$。将1,4-半乳糖转移酶(LgtC)(100U)、半乳糖激酶(GalK)(50U)、UDP-糖焦磷酸化酶(USP)(150U)、丙酮酸激酶(PykF)(200U)及焦磷酸化酶(PPA)(200U)加入其中以启动反应。将烧瓶置于25℃下培养，并以TLC监测反应过程，以对-茴香醛染色标记。如任一反应尚未完成，则加入更多酶直到反应完成，并以TLC及ESI-MS确定其产物。

[0252] 丙烯基-Gb4的合成

[0253] 在丙烯基-Gb3合成之后，再加入其它成分，该成分包括溶在220mL溶液中的N-乙酰半乳糖胺(GalNAc)(9.9mmol)、PEP(22mmol)、1,3-N-乙酰半乳糖胺转移酶(1,3-GalNAcT,LgtD)(100U)、N-乙酰己糖胺1-激酶(GalNAcK)(50U)、N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶(GlmU)(200U)、PykF(100U)及PPA(100U)。将混合物置于25℃下培养，并以TLC及ESI-MS监测直到反应完全。进一步以C-18凝胶管柱纯化该产物并以NMR确定其特征。

[0254] 丙烯基-Gb5的合成

[0255] 反应混合物(250mL)包含溶在100mM Tris-HCl缓冲液(pH 7.0)中的丙烯基-Gb4(9mmol)、半乳糖(9mmol)、PEP(22mmol)、ATP(0.05mmol)、UTP(0.125mmol)及10mM $MgCl_2$。在反应混合物中加入1,3-半乳糖转移酶(1,3-GalT,LgtD)(200U)、GalK(50U)、USP(150U)、PykF(100U)及PPA(100U)以启动反应，并置于25℃下培养，直到反应完全。

[0256] 丙烯基-Globo H的合成

[0257] 将丙烯基-Gb5反应产物的一半量(～4.5mmol)，在未经额外纯化下即直接用来制备丙烯基-globo H。将包含岩藻糖(5mmol)、ATP(0.05mmol)、GTP(0.5mmol)及PEP(11mmol)(溶于100mM Tris-HCl缓冲液(pH 7.0))的溶液加入L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶(FKP)(200U)、PykF(200U)、PPA(200U)及1,2-岩藻糖转移酶(FutC)(200U)中，并置于25℃下培养直到反应完全，另以C-18凝胶层析纯化该产物并确定其特征。

[0258] 丙烯基-SSEA4的合成

[0259] 将丙烯基-Gb5反应混合物的另一半的量(4.5mmol)用于合成丙烯基-SSEA4。于其中加入溶于100mM Tris-HCl缓冲液(pH 8.0)中的N-乙酰基神经胺糖酸(Neu5Ac)(5mmol)、ATP(0.05mmol)、CTP(0.25mmol)、PEP(11mmol)及10mM $MgCl_2$，接着加入胞苷单磷酸激酶(CMK)(50U)、CMP-唾液酸合成酶(Css)(120U)、PykF(100U)、PPA(100U)及α-2,3-唾液酸转移酶(JT-FAJ-16)(150U)。其过程由TLC监控，且如上述方法纯化并确定该产物的特征。

[0260] 寡糖的纯化与特征确定

[0261] 加热反应混合物至90℃并维持30分钟，继续进行离心(5000rpm，20分钟)以去除反应混合物中的蛋白质。滤液续以C-18凝胶层析法纯化并以梯度液(自100％水至10％甲醇

(在水中))洗涤。收集馏分并以TLC[丁醇/氢氧化铵/水＝5:3:2(v/v/v)]监测,收集丙烯基-寡糖的馏分并予以冻干。通过HPLC(使用Cosmosil 5SL-II管柱,在(H_2O/乙腈＝19/81)下,在等度模式(isocratic mode)下)可获得纯度高于99%的产物。以1H NMR、^{13}C NMR及质谱仪(Avance 600及APEX-ultra 9.4T FTICR-MS,Bruker Daltonics)分析丙烯基-Lac、丙烯基-Gb3、丙烯基-Gb4、丙烯基-Gb5、烯基丙基-Globo H及丙烯基-SSEA4的结构。

[0262] 用于核苷酸糖合成、糖基转移酶及ATP再生的基因的克隆

[0263] 所有基因是通过基因组DNA或cDNA库,以个别引物(表6)经PCR而获得,且将该PCR产物接入经修饰的pET47b载体。ATG之后,接着为His-tag、AcTEV蛋白酶切割位置、及侧接特别限制识别酶的ccdB正选择基因(positive selection gene),或在C-末端His-tag的pET28a。为了增加基因表达量,使用对大肠杆菌最优化的密码来合成该四个糖基转移酶。通过化学转化方法,将带有正确序列的质粒转到ArcticExpress/RIL感受态细胞。挑选单一菌落,并将其接种至具有卡那霉素的TB培养基,隔夜培养后使用新鲜TB培养基更新细胞培养物。当OD600达到0.5时,以终浓度为0.1mM的IPTG诱发目标蛋白质的表达。接着继续在16℃下生长24小时候,收集该大肠杆菌,并在含有50mM磷酸钠缓冲液(pH8.0)、300mM氯化纳及10mM咪唑的缓冲液中,以微液化器将细胞打破。将该细胞在4℃及10,000rpm下离心30分钟。将悬浮液倒入经平衡的Ni-NTA琼脂,且将沉淀物丢弃。使用相同的缓冲液(但所含咪唑的浓度较高(250mM))将结合的蛋白质洗提出来。使用Qubit Protein Quantitation(Invitrogen,CA)测量蛋白质浓度,并用SDS-PAGE确认其纯度。

[0264] 表6:用于在大肠杆菌中表达唾液酸酶的引物

[0265]

SEQ ID NO	引物[a]	序列 (5′ →3′)	限制酶位置	来自基因组或 cDNA 库的基因源
SEQ ID NO: 1	*galK*-F	CTGTATTTTCAGGGA**GCGATCGC**TATGAGTCTG AAAGAAAAAACA[b]	AsiSI	大肠杆菌 MG1655 ATCC 700926
SEQ ID NO: 2	*galK*-R	GCCTCGAGTCATTAC**GTTTAAAC**TCAGCACTGT CCTGCTCCTTG	PmeI	
SEQ ID NO: 3	*usp*-F	CTGTATTTTCAGGGA**GCGATCGC**TATGGCTTCT ACGGTTGATTC	AsiSI	拟南芥的 cDNA 库
SEQ ID NO: 4	*usp*-R	GCCTCGAGTCATTAC**GTTTAAAC**TCAATCTTCA ACAGAAAATTTGC	PmeI	
SEQ ID NO: 5	*lgtC*-F[b]	GATATA**CCATGG**AAATGGACATCGTTTTCGCGG CG	NcoI	最优化基因
SEQ ID NO: 6	*lgtC*-R[b]	GTGGTG**CTCGAG**GTAGATTTTACGCAGGAAAC G	XhoI	
SEQ ID NO: 7	*galnacK-F*	CTGTATTTTCAGGGAGCGATCGCTATGAACAAG ACTTATGATTTTAAAAG	AsiSI	龙根菌 ATCC 15697
SEQ ID NO: 8	*galnacK-R*	GCCTCGAGTCATTACGTTTAAACTTAAATGTATG AATATACTATCTTC	PmeI	
SEQ ID NO: 9	*glmU*-F	CTGTATTTTCAGGGA**GCGATCGC**TATGTTGAAT AATGCTATGAGC	AsiSI	大肠杆菌 MG1655 ATCC 700926
SEQ ID NO: 10	*glmU*-R	GCCTCGAGTCATTAC**GTTTAAAC**TCACTTTTTCT TTACCGGACG	PmeI	
SEQ ID NO: 11	*lgtD*-F[b]	GATATA**CCATGG**AAAACTGCCCGCTGGTTTCT	NcoI	最优化基因
SEQ ID NO: 12	*lgtD*-R[b]	GTGGTG**CTCGAG**GAAGATAACGTTGATTTTACG G	XhoI	
SEQ ID NO: 13	*fkp*-F	CAGGGA**GCGATCGC**TATGCAAAAACTACTATCT TTA	AsiSI	脆弱拟杆菌 9343

[0266]

SEQ ID NO	引物[a]	序列 (5′ →3′)	限制酶位置	来自基因组或 cDNA 库的基因源
SEQ ID NO: 14	*fkp*-R	CATTAC**GTTTAAAC**TTATGATCGTGATACTTGGAA	PmeI	ATCC 25285
SEQ ID NO: 15	*futC*-F[b]	CTGTATTTTCAGGGA**GCGATCGC**TATGGCGTTCAAAGTTGTTCAG	AsiSI	最优化基因
SEQ ID NO: 16	*futC*-R[b]	GCCTCGAGTCATTAC**GTTTAAAC**TTACGCGTTGTATTTCTGAGAT	PmeI	
SEQ ID NO: 17	*cmk*-F	CAGGGA**GCGATCGC**TATGACGGCAATTGCCCCGGTT	AsiSI	大肠杆菌 MG1655 ATCC 700926
SEQ ID NO: 18	*cmk*-R	CATTAC**GTTTAAAC**TTATGCGAGAGCCAATTTCTG	PmeI	
SEQ ID NO: 19	*css*-F	GATATA**CCATGG**AAACAAATATTGCGATCATTCCTG	NcoI	多杀性巴氏杆菌 ATCC BAA-1113
SEQ ID NO: 20	*css*-R	GTGGTG**CTCGAG**TTTATTGGATAAAATTTCCGCGAG	XhoI	
SEQ ID NO: 21	*jt-faj-16*-F[b]	GATATA**CCATGG**AAATGAACAACGACAACTCTACC	NcoI	最优化基因
SEQ ID NO: 22	*jt-faj-16*-R[b]	GTGGTG**CTCGAG**GATGTCAGAGATCAGTTTGATG	XhoI	
SEQ ID NO: 23	*pykF*-F	CTGTATTTTCAGGGA**GCGATCGC**TATGAAAAAGACCAAAATTGTTTG	AsiSI	大肠杆菌 MG1655 ATCC 700926
SEQ ID NO: 24	*pykF*-R	GCCTCGAGTCATTAC**GTTTAAAC**TTACAGGACGTGAACAGATG	PmeI	
SEQ ID NO: 25	*ppa*-F	CAGGGA**GCGATCGC**TATGAGCTTACTCAACGTCCCT	AsiSI	大肠杆菌 MG1655 ATCC 700926
SEQ ID NO: 26	*ppa*-R	CATTAC**GTTTAAAC**TTATTTATTCTTTGCGCGCTC	PmeI	

[0267] [a]用于正向 (F) 及逆向 (R) PCR反应以放大各基因的编码序列的引物对。

[0268] [b]粗底线表示限制酶识别位置。

[0269] [c]用于大肠杆菌的最优化密码。见如Puigbò等人,Nucleic Acids Research(核酸研究期刊)(2007)35(S2):W126-W130。

[0270] 酶分析

[0271] 为维持恒定的分析条件,所有活性测定均在37℃,以10mM $MgCl_2$、100mM Tris以及pH为7.5的条件下进行。

[0272] (i)半乳糖激酶(GalK)、N-乙酰己糖胺1-激酶(GalNAcK)、岩藻糖激酶(FKP)及胞苷单磷酸激酶(CMK)活性测定

[0273] 荧光分析法是基于ADP产量(ATP消耗量)的监测,其是使用测量NADH消耗量的丙酮酸激酶/乳酸脱氢酶结合酶分析。见如,Murray等人,"人类α-1,3-岩藻糖转移酶V的机制:糖苷切割前发生的亲核攻击"生物化学期刊(1997)36:823-831("Mechanism of Human α-1,3-Fucosyltransferase V:Glycosidic Cleavage Occurs Prior to Nucleophilic Attack"Biochemistry(1997)36:823-831);及Gosselin等人,"一种糖基转移酶的连续分光光度分析法"分析生物化学期刊(1994)220:92-97("A Continuous Spectrophotometric Assay for Glycosyltransferases"Analytical Biochemistry(1994)220:92-97)。NADH的荧光特性为具有340nm的激发波长及450nm的放射波长。调配100uL的反应混合物,其含有在100mM Tris(pH 7.5)中的该结合酶(来自兔子肌肉的丙酮酸激酶(5单位)及乳酸脱氢酶(7单位))、底物及辅助因子(0.2mM NADH、0.8mM PEP及10mM $MgCl_2$)。通过加入个别的糖来启动反应。Kcat及Km动力参数是使用Grafit version 7(Erithacus Software,Middlesex,UK)的理论方程式,通过适用该实验数据的曲线来计算。一单位糖激酶活性的定义为,在25℃下,每分钟可形成1umol糖-1-P的量。

[0274] (ii)UDP-糖焦磷酸化酶(USP)、N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶(GlmU)、GDP-L-岩藻糖焦磷酸化酶(FKP)及CMP-唾液酸合成酶(Css)活性的测量

[0275] 焦糖酸盐的产量是使用EnzCheck Pyrophosphate Assay Kit(Invitrogen,CA,USA)来测量。分析成分包括:在UV-Star微孔板(Greiner Bio One,Germany)中的200uM 2-氨基-6-巯基-7-甲基嘌呤核糖核苷、1单位核苷磷酸化酶、0.03单位无机焦磷酸化酶、10mM $MgCl_2$及50mM Tris(pH 7.5于100uL规模)。除FKP外,所有的成分皆在微孔板中混合并使其平衡,直到形成平基础线。通过加入酶来启动反应。除CMP-唾液酸合成酶外,一单位酶活性的定义为,在25℃下,每分钟可由个别糖-1-P形成1umol核苷糖的量;至于CMP-唾液酸合成酶,其一单位酶活性是定义为,在25℃下,每分钟可形成1umol焦糖酸盐的量。

[0276] (iii)糖基转移酶的测量:α-1,4-半乳糖转移酶(LgtC)、β-1,3-N-乙酰半乳糖胺转移酶(β-1,3-GalNAcT,LgtD)、β-1,3-半乳糖转移酶(LgtD),α-1,2-岩藻糖转移酶(FutC)及α-2,3-唾液酸转移酶(JT-FAJ-16)。

[0277] 使用荧光分析法监测UDP、GDP或CDP的生产,其是使用测量NADH消耗量的丙酮酸激酶/乳酸脱氢酶结合酶分析。见如,Murray等人,"人类α-1,3-岩藻糖转移酶V的机制:糖苷切割前发生的亲核攻击"生物化学期刊(1997)36:823-831("Mechanism of Humanα-1,3-Fucosyltransferase V:Glycosidic Cleavage Occurs Prior to Nucleophilic Attack"Biochemistry(1997)36:823-831);及Gosselin等人,"一种糖基转移酶的连续分光光度分析法"分析生物化学期刊(1994)220:92-97("A Continuous Spectrophotometric Assay

for Glycosyltransferases"Analytical Biochemistry(1994)220:92-97)。除了核苷糖，所有的分析成分是在25℃下，同时置于多孔板荧光计(SpectraMax M2 Readers,Molecular Devices)中培养。通过加入对应的核苷糖来启动反应。Kcat及Km动力参数是使用Grafit version 7(Erithacus Software,Middlesex,UK)的理论方程式，通过适用该实验数据的曲线来计算。一单位活性的定义为，在25℃下，每分钟可自各别核苷糖中，催化转移1umol糖至受体的酶量。

[0278] (iv)丙酮酸激酶(PyrK)的量测

[0279] 丙酮酸激酶分析法是将前述的糖基酶稍作修改而来，且也为基于NADH的消耗量。制备100uL的反应混合物，其包含在100mM Tris(pH7.5)中的0.8mM ADP、0.8mM PEP、0.2mM NADH、10mM $MgCl_2$及来自兔子肌肉的乳酸脱氢酶(5单位)，该混合物是置于黑色多孔板。NADH具有340nm的激发波长及450nm的放射波长。通过加入适当量的重组大肠杆菌丙酮酸激酶来启动反应。一单位酶活性的定义为，在25℃下，每分钟可由个别糖-1-P形成1umol核苷糖的量；至于CMP-唾液酸合成酶，其一单位的丙酮酸激酶是定义为，在25℃下，每分钟能将1umol磷酸(烯醇)丙酮酸转成丙酮酸的量。

[0280] (v)焦磷酸化酶(PPA)的量测

[0281] 焦磷酸化酶分析法是将来自EnzCheck Pyrophosphate Assay Kit(Invitrogen,CA,USA)套组所提供的磷酸化酶测量法，稍作修改而来。分析成分包括：1mM焦糖酸盐、200uM 2-氨基-6-巯基-7-甲基嘌呤核糖核苷、1单位核苷磷酸化酶、10mM $MgCl_2$及50mM Tris(pH of 7.5)，其是与适当量的重组大肠杆菌焦磷酸化酶，以100uL规模置于UV-Star微孔板(Greiner Bio One,Germany)中。一单位的焦磷酸化酶活性是定义为，在25℃下，每分钟能释放1.0umole无机焦糖酸盐的量。

[0282] (vi)最佳pH的量测

[0283] 酶活性的最佳pH是在4.0-10.0的范围内，使用前述的标准酶分析方法来量测，其包括醋酸钠、MES、MOPS、HEPES、Tris-HCl及CHES缓冲液。该缓冲液的pH值是在培养温度下调整。所有的反应皆进行三重复以利统计评估。

[0284] (vii)最佳二价金属离子的量测

[0285] 所需金属的分析是在标准分析条件下进行。在EDTA的存在或不存在下，将酶与金属离子(Mg^{2+}、Mn^{2+}、Mg^{2+}、Mn^{2+}、Ca^{2+}、Zn^{2+}、Co^{2+}或Ni^{2+})混合(最终浓度10mM)。所有的反应皆进行三重复以利统计评估。

[0286] (viii)最佳温度的量测

[0287] 温度对酶活性的影响力的评估，是将适当量的纯酶培养在MOPS缓冲液(pH 7.0)、10mM $MgCl_2$及个别底物中。为维持分析的一致性，先将所有的成分混和好，且在分析温度下预热5分钟，再于恒定的温度下，将酶加入来启动反应，并以多模态平板读取机(SpectraMax M5,Molecular Devices)来记录。温度范围在20至60℃间。所有的反应皆进行三重复以利统计评估。

[0288] 酶组合物

[0289] UDP-Gal再生/半乳糖基化

[0290] 1.GalK：源自大肠杆菌的半乳糖激酶

[0291] 2.USP：源自拟南芥的UDP-糖焦磷酸化酶

[0292] 3.LgtC:源自脑膜炎双球菌的α-1,4-半乳糖转移酶,但其密码对大肠杆菌最优化

[0293] 4.PykF:源自大肠杆菌的丙酮酸激酶

[0294] 5.PPA:源自大肠杆菌的焦磷酸化酶

[0295] 密码最优化的LgtC酶的编码序列是如下所示(SEQ ID NO:27):

[0296] ATGGACATCGTTTTCGCGGCGGACGACAACTACGCGGCGTACCTGTGCGTTGCGGCGAAATCTGTTGAA
GCGGCGCACCCGGACACCGAAATCCGTTTCCACGTTCTGGACGCGGGTATCTCTGAAGCGAACCGTGCGGCGGTTGC
GGCGAACCTGCGTGGTGGTGGTGGTAACATCCGTTTCATCGACGTTAACCCGGAAGACTTCGCGGGTTTCCCGCTGA
ACATCCGTCACATCTCTATCACCACCTACGCGCGTCTGAAACTGGGTGAATACATCGCGGACTGCGACAAAGTTCTG
TACCTGGACATCGACGTTCTGGTTCGTGACTCTCTGACCCCGCTGTGGGACACCGACCTGGGTGACAACTGGCTGGG
TGCGTGCATCGACCTGTTCGTTGAACGTCAGGAAGGTTACAAACAGAAAATCGGTATGGCGGACGGTGAATACTACT
TCAACGCGGGTGTTCTGCTGATCAACCTGAAAAAATGGCGTCGTCACGACATCTTCAAAATGTCTTGCGAATGGGTT
GAACAGTACAAAGACGTTATGCAGTACCAGGACCAGGACATCCTGAACGGTCTGTTCAAAGGTGGTGTTTGCTACGC
GAACTCTCGTTTCAACTTCATGCCGACCAACTACGCGTTCATGGCGAACCGTTTCGCGTCTCGTCACACCGACCCGC
TGTACCGTGACCGTACCAACACCGTTATGCCGGTTGCGGTTTCTCACTACTGCGGTCCGGCGAAACCGTGGCACCGT
GACTGCACCGCGTGGGGTGCGGAACGTTTCACCGAACTGGCGGGTTCTCTGACCACCGTTCCGGAAGAATGGCGTGG
TAAACTGGCGGTTCCGCACCGTATGTTCTCTACCAAACGTATGCTGCAGCGTTGGCGTCGTAAACTGTCTGCGCGTT
TCCTGCGTAAAATCTACTGA

[0297] UDP-GalNAc再生/乙酰半乳糖胺化

[0298] 1.GalNAcK:源自龙根菌的N-乙酰己糖胺1-激酶

[0299] 2.GlmU:源自大肠杆菌的N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶

[0300] 3.LgtD:源自流感嗜血杆菌的β-1,3-半乳糖转移酶,但其密码对大

[0301] 肠杆菌最优化

[0302] 4.PykF:源自大肠杆菌的丙酮酸激酶

[0303] 5.PPA:源自大肠杆菌的焦磷酸化酶

[0304] 密码最优化的LgtD酶的编码序列是如下所示(SEQ ID NO:28):

[0305] ATGGAAAACTGCCCGCTGGTTTCTGTTATCGTTTGCGCGTACAACGCGGAACAGTACATCGACGAATCT
ATCTCTTCTATCATCAACCAGACCTACGAAAACCTGGAAATCATCGTTATCAACGACGGTTCTACCGACCTGACCCT
GTCTCACCTGGAAGAAATCTCTAAACTGGACAAACGTATCAAAATCATCTCTAACAAATACAACCTGGGTTTCATCA
ACTCTCTGAACATCGGTCTGGGTTGCTTCTCTGGTAAATACTTCGCGCGTATGGACGCGGACGACATCGCGAAACCG
TCTTGGATCGAAAAAATCGTTACCTACCTGGAAAAAAACGACCACATCACCGCGATGGGTTCTTACCTGGAAATCAT
CGTTGAAAAAGAATGCGGTATCATCGGTTCTCAGTACAAAACCGGTGACATCTGGAAAAACCCGCTGCTGCACAACG
ACATCTGCGAAGCGATGCTGTTCTACAACCCGATCCACAACAACACCATGATCATGCGTGCGAACGTTTACCGTGAA
CACAAACTGATCTTCAACAAAGACTACCCGTACGCGGAAGACTACAAATTCTGGTCTGAAGTTTCTCGTCTGGGTTG
CCTGGCGAACTACCCGGAAGCGCTGGTTAAATACCGTCTGCACGGTAACCAGACCTCTTCTGTTTACAACCACGAAC
AGAACGAAACCGCGAAAAAAATCAAACGTGAAAACATCACCTACTACCTGAACAAAATCGGTATCGACATCAAAGTT
ATCAACTCTGTTTCTCTGCTGGAAATCTACCACGTTGACAAATCTAACAAAGTTCTGAAATCTATCCTGTACGAAAT
GTACATGTCTCTGGACAAATACACCATCACCTCTCTGCTGCACTTCATCAAATACCACCTGGAACTGTTCGACCTGA
AACAGAACCTGAAAATCATCAAAAAATTCATCCGTAAAATCAACGTTATCTTCTAG

[0306] GDP-FKP再生/岩藻糖基化

[0307] 1.FKP:源自脆弱拟杆菌的L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶

[0308] 2.FutC:源自幽门螺旋杆菌的α-1,2-岩藻糖转移酶,但其密码对大肠杆菌最优化

[0309] 3.PykF:源自大肠杆菌的丙酮酸激酶

[0310] 4.PPA:源自大肠杆菌的焦磷酸化酶

[0311] 密码最优化的FutC酶的编码序列是如下所示(SEQ ID NO:29):

[0312] ATGGCGTTCAAAGTTGTTCAGATCTGCGGTGGTCTGGGTAACCAGATGTTCCAGTACGCGTTCGCGAAA
TCTCTGCAGAAACACTCTAACACCCCGGTTCTGCTGGACATCACCTCTTTCGACTGGTCTGACCGTAAAATGCAGCT
GGAACTGTTCCCGATCGACCTGCCGTACGCGTCTGCGAAAGAAATCGCGATCGCGAAAATGCAGCACCTGCCGAAAC
TGGTTCGTGACGCGCTGAAATGCATGGGTTTCGACCGTGTTTCTCAGGAAATCGTTTTCGAATACGAACCGAAACTG
CTGAAACCGTCTCGTCTGACCTACTTCTTCGGTTACTTCCAGGACCCGCGTTACTTCGACGCGATCTCTCCGCTGAT
CAAACAGACCTTCACCCTGCCGCCGCCGCCGGAAAACAACAAAAACAACAACAAAAAAGAAGAAGAATACCAGTGCA
AACTGTCTCTGATCCTGGCGGCGAAAAACTCTGTTTTCGTTCACATCCGTCGTGGTGACTACGTTGGTATCGGTTGC
CAGCTGGGTATCGACTACCAGAAAAAAGCGCTGGAATACATGGCGAAACGTGTTCCGAACATGGAACTGTTCGTTTT
CTGCGAAGACCTGGAATTCACCCAGAACCTGGACCTGGGTTACCCGTTCATGGACATGACCACCCGTGACAAAGAAG
AAGAAGCGTACTGGGACATGCTGCTGATGCAGTCTTGCCAGCACGGTATCATCGCGAACTCTACCTACTCTTGGTGG
GCGGCGTACCTGATCGAAAACCCGGAAAAAATCATCATCGGTCCGAAACACTGGCTGTTCGGTCACGAAAACATCCT
GTGCAAAGAATGGGTTAAAATCGAATCTCACTTCGAAGTTAAATCTCAGAAATACAACGCGTAA

[0313] CMP-Neu5Ac再生/唾液酸化

[0314] 1.CMK:源自大肠杆菌的胞苷单磷酸激酶

[0315] 2.Css:源自多杀性巴氏杆菌的CMP-唾液酸合成酶

[0316] 3.JT-FAJ-16:源自海洋细菌的α-2,3-唾液酸转移酶,但其密码对大肠杆菌最优化

[0317] 4.PykF:源自大肠杆菌的丙酮酸激酶

[0318] 5.PPA:源自大肠杆菌的焦磷酸化酶

[0319] 密码最优化的JT-FAJ-16酶的编码序列是如下所示(SEQ ID NO:30):

[0320] ATGAACAACGACAACTCTACCACCACCAACAACAACGCGATCGAAATCTACGTTGACCGTGCGACCCTG
CCGACCATCCAGCAGATGACCAAAATCGTTTCTCAGAAAACCTCTAACAAAAAACTGATCTCTTGGTCTCGTTACCC
GATCACCGACAAATCTCTGCTGAAAAAAATCAACGCGGAATTCTTCAAAGAACAGTTCGAACTGACCGAATCTCTGA
AAAACATCATCCTGTCTGAAAACATCGACAACCTGATCATCCACGGTAACACCCTGTGGTCTATCGACGTTGTTGAC
ATCATCAAAGAAGTTAACCTGCTGGGTAAAAACATCCCGATCGAACTGCACTTCTACGACGACGGTTCTGCGGAATA
CGTTCGTATCTACGAATTCTCTAAACTGCCGGAATCTGAACAGAAATACAAAACCTCTCTGTCTAAAAACAACATCA
AATTCTCTATCGACGGTACCGACTCTTTCAAAAACACCATCGAAAACATCTACGGTTTCTCTCAGCTGTACCCGACC
ACCTACCACATGCTGCGTGCGGACATCTTCGACACCACCCTGAAAATCAACCCGCTGCGTGAACTGCTGTCTAACAA
CATCAAACAGATGAAATGGGACTACTTCAAAGACTTCAACTACAAACAGAAAGACATCTTCTACTCTCTGACCAACT
TCAACCCGAAAGAAATCCAGGAAGACTTCAACAAAAACTCTAACAAAAACTTCATCTTCATCGGTTCTAACTCTGCG
ACCGCGACCGCGGAAGAACAGATCAACATCATCTCTGAAGCGAAAAAAGAAAACTCTTCTATCATCACCAACTCTAT
CTCTGACTACGACCTGTTCTTCAAAGGTCACCCGTCTGCGACCTTCAACGAACAGATCATCAACGCGCACGACATGA
TCGAAATCAACAACAAAATCCCGTTCGAAGCGCTGATCATGACCGGTATCCTGCCGGACGCGGTTGGTGGTATGGGT
TCTTCTGTTTTCTTCTCTATCCCGAAAGAAGTTAAAAACAAATTCGTTTTCTACAAATCTGGTACCGACATCGAAAA
CAACTCTCTGATCCAGGTTATGCTGAAACTGAACCTGATCAACCGTGACAACATCAAACTGATCTCTGACATCTAA

[0321] 材料及化学品

[0322] 所有的核苷酸、糖、核苷糖及化学品皆是购自Sigma-Aldrich(St.Louis,MO)。限制酶及T4 DNA接合酶是或自NEB(Beverly,MA)。引物是订购自Proligo Singapore Pte Ltd(Singapore)。Ni-NTA Agarose来自Qiagen(Santa Clarita,CA)。Bio-Gel P2凝胶是购自Bio-Rad(Hercules,CA)。质粒pET28a、pET47b及经保护的TLC玻璃平板(被厚度为0.25 mm的Silica Gel 60及F254所覆盖)是购自EMD Chemicals Inc(Carlsbad,CA)。ArcticExpress/RIL感受态细胞是获自Agilent Genomics(La Jolla,CA)。以上未提及的其它物质皆是采购尽可能高质量者。

[0323] 所有的反应皆以薄膜层析法(TLC)监控(流动相:丁醇:乙酸盐:水=5:3:2)。TLC是以p-甲氧苯甲醛来染色。

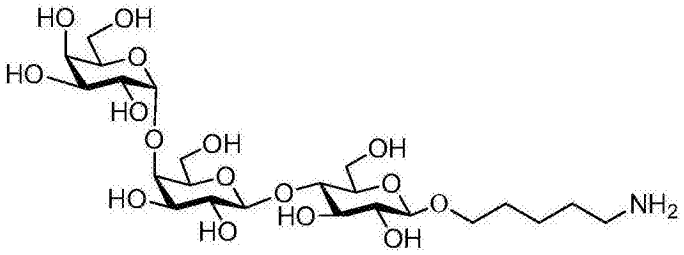
[0324] 丙烯基-Lac的合成

[0325] 带有连接符(linker)的不同乳糖是根据学术报导的方法[Carbohydrate Research 2004,339,2415-2424.]来合成。^{1}H NMR(600MHz,D2O)δ6.01(m,1H),5.40-5.37(dd,J=17.3,1.4 Hz,1H),5.30-5.28(d,J=10.3 Hz,1H),4.54(d,J=8.1 Hz,1H),4.46(d,J=7.8Hz,1H),4.41-4.38(m,1H),4.25-4.22(m,1H),4.00-3.97(dd,J=12.3,2.1 Hz,1H),3.93(d,J=3.3 Hz,1H),3.80-3.71(m,4 H),3.67-3.53(m,5H),3.35-3.33(m,1H);^{13}C NMR(150 MHz,D2O)δ133.3,118.7,102.9,101.0,78.3,75.3,74.7,74.4,72.8,72.5,70.9,70.6,68.5,60.9,60.1;HRMS(ESI-TOF,MNa^{+})$C_{15}H_{26}O_{11}Na^{+}$计算值405.1367,测量值405.1346。

[0326] 带有连接符的Gb3的大规模制备

[0327] 将5mmol带有连接符的乳糖、5mmol半乳糖、12mmol磷酸烯醇丙酮酸(PEP)、0.25mmol ATP、0.25mmol UTP及10mM $MgCl_2$加入100mM Tris-HCl缓冲溶液(pH 7.5)。通过加入适当量的α-1,4-半乳糖转移酶(LgtC)、半乳糖激酶(GalK)、UDP-糖焦磷酸化酶(USP)、丙酮酸激酶(PykF)及焦磷酸化酶(PPA)来启动反应。将锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC观察反应。如果反应停止,则加入更多的酶。当TLC发现其中没有起始物时,则表示该反应停止了。使用^{18}C逆相管柱分离该Gb3产物,其产率为99%。

[0328] 丙烯基-Gb3:^{1}H NMR(600 MHz,D2O)δ6.00(m,1H),5.42(d,J=17.2 Hz,1H),5.32(d,J=10.4 Hz,1H),4.97(d,J=3.3 Hz,1H),4.56(d,J=7.9 Hz,1H),4.53(d,J=7.7 Hz,1H)、4.43-4.37(m,2H),4.27-4.24(m,1H),4.06-3.58(m,16H),3.37-3.34(t,J=8.0Hz,1H);^{13}C NMR(150MHz,D2O)δ133.3,118.7,103.3,100.9,100.3,78.6,77.3,75.4,74.8,74.5,72.9,72.2,70.9,70.8,70.6,69.1,68.9,68.5,60.5,60.4,60.0;HRMS(ESI-TOF,MNa^{+})$C_{21}H_{36}O_{16}Na^{+}$计算值567.1896,测量值567.1858。

[0329] 

[0330] 5-氨基戊基-Gb3

[0331] ^{1}H NMR(600MHz,D2O)δ4.97(d,J=3.3Hz,1H),4.56(d,J=7.9Hz,1H)、4.54-4.50

(m,2H),4.37(dd,J=6.0Hz,J=0.6Hz,1H),4.27-4.24(m,1H),4.06-3.59(m,18H),3.35(t,J=8.0Hz,1H),3.03(t,J=7.4Hz,2H),1.75-1.68(m,4H),1.51-1.46(m,2H);^{13}C NMR(150MHz,D2O)δ103.3,101.9,100.3,78.7,77.3,75.4,74.8,74.5,72.9,72.2,70.9,70.8,70.6,69.1,68.9,68.5,60.5,60.4,60.0,39.3,28.1,26.4,22.1。

[0332] 带有连接符的Gb4的大规模制备

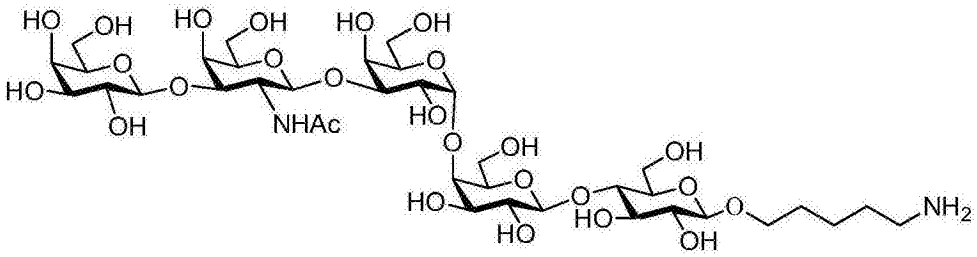
[0333] 将5mmol带有连接符的Gb3、5mmol N-乙酰基半乳糖胺(GalNAc)、12mmol磷酸烯醇丙酮酸(PEP)、0.25mmol ATP、0.25mmol UTP及10mM $MgCl_2$加入100mM Tris-HCl缓冲溶液(pH 7.5)。通过加入适当量的β-1,3-N-乙酰半乳糖胺转移酶(LgtD)、N-乙酰己糖胺1-激酶(GalNAcK)、N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶(GlmU)、丙酮酸激酶(PykF)及焦磷酸化酶(PPA)来启始反应。将锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC观察反应。如果反应停止,则加入更多的酶。当TLC发现其中没有起始物时,则表示该反应停止了。使用^{18}C逆相管柱分离该Gb4产物,其产率为96%。

[0334] 丙烯基-Gb4:^{1}H NMR(600MHz,D2O)δ6.01(m,1H),5.40-5.38(dd,J=17.3,1.4Hz,1H),5.30(d,J=10.5Hz,1H),4.92(d,J=3.9Hz,1H),4.64(d,J=8.5Hz,1H),4.54(d,J=7.9Hz,1H),4.53(d,J=7.8Hz,1H),4.42-4.38(m,2H),4.26-4.22(m,2H),4.05(d,J=2.9Hz,1H),4.01-3.99(dd,J=12.3,1.8Hz,1H),3.98-3.89(m,5H),3.86-3.74(m,7H),3.72-3.57(m,7H),3.37-3.34(t,J=8.6Hz,1H),2.05(s,3H);^{13}C NMR(150MHz,D2O)δ133.2,118.7,103.3,103.2,100.9,100.4,78.7,78.6,77.2,75.4,74.9,74.8,74.5,72.9,72.1,70.9,70.8,70.6,70.2,68.9,67.7,67.6,60.9,60.5,60.3,60.2,60.0,52.6,22.2;HRMS(MALDI,MNa^+)$C_{29}H_{49}NO_{21}Na^+$计算值770.2689,测量值770.2654。

[0335] 带有连接符的Gb5的大规模制备

[0336] 将5mmol丙烯基-Gb4、5mmo半乳糖、12mmol磷酸烯醇丙酮酸(PEP)、0.25mmol ATP、0.25mmol UTP与10mM $MgCl_2$加入100mM Tris-HCl缓冲液(pH 7.5)。通过加入适当量的β-1,3-半乳糖转移酶、半乳糖激酶(GalK)、UDP-糖焦磷酸化酶(USP)、丙酮酸激酶(PykF)及焦磷酸化酶(PPA)来启始反应。将锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC观察反应。如果反应停止,则加入更多的酶。当TLC发现其中没有起始物时,则表示该反应停止了。使用^{18}C逆相管柱分离该Gb5产物,其产率为95%。

[0337] 丙烯基-Gb5:^{1}H NMR(600MHz,D2O)δ6.01(m,1H),5.41-5.38(dd,J=17.3,1.4Hz,1H),5.31(d,J=10.6Hz,1H),4.93(d,J=4.0Hz,1H),4.71(d,J=8.5Hz,1H),4.55(d,J=8.1Hz,1H),4.53(d,J=7.8Hz,1H),4.47(d,J=7.7Hz,1H),4.42-4.39(m,2H),4.27-4.23(m,2H),4.20(d,J=3.2Hz,1H),4.09-3.90(m,8H),3.87-3.59(m,17H),3.55-3.52(m,1H),3.36-3.33(t,J=8.6Hz,1H),2.04(s,3H);^{13}C NMR(150MHz,D2O)δ175.1,133.2,118.7,104.8,103.3,102.9,100.9,100.4,79.6,78.7,78.6,77.2,75.4,74.9,74.8,74.6,74.5,72.9,72.4,72.1,70.9,70.6(2C),70.2,68.9,68.5,67.9,67.6,60.9(2C),60.33,60.28,60.0,51.5,22.2;HRMS(ESI-TOF,MNa^+)$C_{35}H_{59}NO_{26}Na^+$计算值932.3218,测量值932.3235。

[0338] 

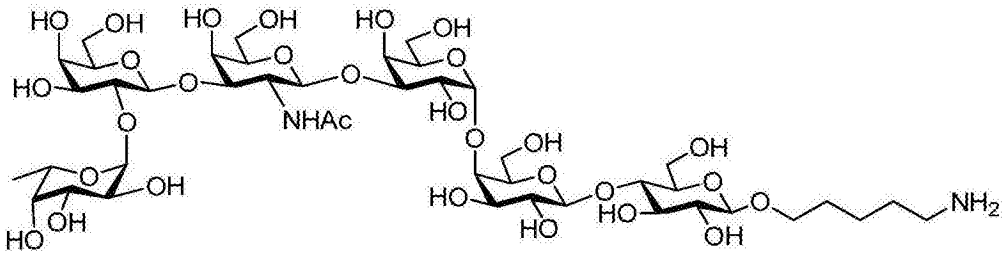
[0339] 5-氨基戊基-Gb5

[0340] ^{1}H NMR(600MHz,D2O),4.47(d,1H,J=8.42Hz),4.30(d,1H,J=7.9Hz),4.28(d,1H,J=8.1Hz),4.24(d,1H,J=7.7Hz),4.19(t,1H J=7.0Hz),4.04(d,1H,J=2.8Hz),3.97(d,1H,J=2.98Hz),3.87-3.35(m,32H),3.30(t,1H,J=7.7Hz),3.09(t,1H,J=8.5Hz),2.79(t,2H,J=7.6Hz),1.82(s,3H),1.51-1.43,(m,4H),1.28-1.21(m,2H)^{13}C NMR(150MHz,D2O),ä175.0,104.7,103.1,102.8,101.8,100.2,79.4,78.5,78.4,76.9,75.3,74.8,74.7,74.4,74.3,72.8,72.2,71.9,70.6,70.4,70.0,69.9,68.7,68.4,67.8,67.4,60.82,60.77,60.13,60.1,59.8,51.3,39.1,28.0,26.3,22.1,21.9MALDI-TOF:$C_{37}H_{66}N_2O_{26}$[M+H]$^{+}$计算值955.3904;测量值955.3972。

[0341] 带有连接符的Globo H的大规模制备

[0342] 将5mmol带有连接符的Gb5、5mmol岩藻糖、12mmol磷酸烯醇丙酮酸(PEP)、0.25mmol ATP、0.25mmol GTP与10mM $MgCl_2$加入100mM Tris-HCl缓冲液(pH 7.5)。通过加入α-1,2-岩藻糖转移酶、L-岩藻糖激酶/GDP-岩藻糖焦磷酸化酶(FKP)、丙酮酸激酶(PykF)及焦磷酸化酶(PPA)来启始反应。将锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC观察反应。如果反应停止,则加入更多的酶。当TLC发现其中没有起始物时,则表示该反应停止了。使用^{18}C逆相管柱分离该Globo H产物,其产率为94%。

[0343] 丙烯基-Globo H:^{1}H NMR(600MHz,D2O)δ6.01(m,1H),5.41-5.38(dd,J=17.3,1.4Hz,1H),5.31(d,J=10.7Hz,1H),5.24(d,J=4.0Hz,1H),4.91(d,J=3.9Hz,1H),4.63(d,J=7.7Hz,1H),4.56-4.52(m,3H),4.42-4.40(m,2H),4.26-4.23(m,3H),4.12(d,J=2.2Hz,1H),4.05(d,J=3.0Hz,1H),4.03-3.59(m,28H),3.36-3.33(t,J=8.2Hz,1H),2.06(s,3H),1.24(d,J=6.5Hz,3H);^{13}C NMR(150MHz,D2O)δ174.3,133.2,118.7,103.9,103.2,102.0,100.9,100.4,99.3,78.7,78.3,77.1,76.3,76.1,75.5,75.0,74.8,74.6,74.5,73.5,72.9,72.1,71.8,70.8,70.6,70.1,69.5,69.2,69.1,68.5,68.0,67.8,66.8,60.95,60.93,60.3(2C),60.0,51.6,22.2,15.3;HRMS(MALDI,MNa^{+})$C_{41}H_{70}NO_{30}Na^{+}$计算值1079.3875,测量值1078.4145.

[0344] 

[0345] 5-氨基戊基-Globo H

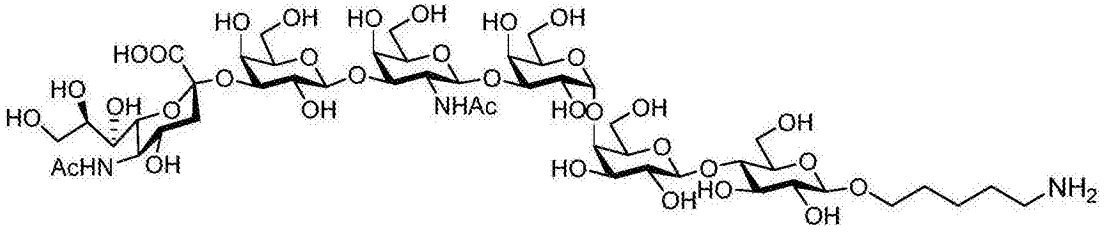
[0346] ^{1}H NMR(600 MHz,D2O)ä5.12(d,1H,J=3.9 Hz),4.78(d,1H,J=3.6 Hz),4.50(d,1H,J=7.7 Hz),4.43(d,1H,J=7.5 Hz),4.40(d,1H,J=7.7 Hz),4.37(d,1H,J=8.0 Hz),

4.30(t,1H,J=6.2 Hz),4.15-4.10(m,2H),3.99(d,1H,J=1.8 Hz),3.92(d,1H,J=2.2 Hz),3.90-3.47(m,33H),3.19(t,1H,J=8.3 Hz),2.89(t,2H,J=7.5 Hz),1.94(s,3H),1.60-1.55(m,4H),1.38-1.31(m,2H),1.11(d,3H,J=6.4 Hz).^{13}C NMR(150 MHz,D2O)ä176.1,105.7,105.0,103.74,103.65,102.1,100.97,80.5,79.9,78.8,78.0,77.8,77.2,76.76,76.5,76.3,76.2,75.3,74.6,73.8,73.5,72.5,71.81,71.78,71.2,71.1,70.9,70.8,70.2,69.7,69.5,68.5,62.66,62.64,62.0,61.7,53.3,41.0,29.9,28.1,23.9,23.8,17.0MALDI-TOF:$C_{43}H_{76}N_2O_{30}$[M+Na]$^+$计算值1123.4381,测量值1123.4385

[0347] 带有连接符的SSEA4的大规模制备

[0348] 将5mmol带有连接符的Gb5、5mmol岩藻糖、12mmol磷酸烯醇丙酮酸(PEP)、0.25mmol ATP、0.25mmol CTP与10mM $MgCl_2$加入100mM Tris-HCl缓冲液(pH 7.5)。通过加入适当量的α-2,3-唾液酸转移酶、胞苷单磷酸激酶(CMK)、CMP-唾液酸合成酶(Css)、丙酮酸激酶(PykF)及焦磷酸化酶(PPA)来启始反应。将锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC观察反应。如果反应停止,则加入更多的酶。当TLC发现其中没有起始物时,则表示该反应停止了。使用^{18}C逆相管柱分离该SSEA4产物,其产率为45%。

[0349] 丙烯基-SSEA4:^{1}H NMR(600 MHz,D2O)δ6.00(m,1H),5.40-5.37(d,J=17.3 Hz,1H),5.30-5.28(d,J=10.4 Hz,1H),4.92(d,J=3.9 Hz,1H),4.70(d,J=8.5 Hz,1H),4.54-4.51(m,3H)、4.40-4.38(m,2H),4.25-4.18(m,3H),4.10-3.52(m,34 H),3.35-3.32(t,J=8.6 Hz,1H),2.77(dd,J=12.5,4.6Hz,1H),2.03(s,6H),1.80(t,J=12.1Hz,1H);^{13}C NMR(150MHz,D2O)δ175.2,175.1,174.1,133.4,121.6,118.9,104.7,103.4,103.1,101.1,100.5,99.8,79.9,78.9,78.8,77.3,75.7,75.5,75.0,74.7,74.6,73.0,72.9,72.2,72.1,71.9,71.0,70.8,70.4,69.1,69.0,68.5,68.2,68.0,67.7,67.5,62.6,61.1,60.5,60.4,60.1,51.7,51.4,39.8,22.4,22.1;HRMS(ESI-TOF,M-H)$C_{46}H_{75}N_2O_{34}$-计算值1199.4196,测量值1199.4208。

[0350] 

[0351] 5-氨基戊基-SSEA4

[0352] ^{1}H NMR(600MHz,D2O)δ4.94(d,J=3.8Hz,1H),4.72(d,J=8.5Hz,1H),4.54-4.50(m,3H),4.40(t,J=6.4Hz,1H),4.27(d,J=2.0Hz,1H),4.20(d,J=2.8Hz,1H),4.10-3.54(m,37H),3.34-3.31(m,1H),3.02(t,J=7.6Hz,2H),2.78(dd,J=12.4,4.6Hz,1H),2.05(m,6H),1.80(t,12.2Hz,1H),1.74-1.67(m,4H),1.51-1.45(m,2H);^{13}C NMR(150MHz,D_2O)δ175.0,174.9,173.9,104.5,103.2,102.9,101.9,100.3,99.6,79.7,78.8、78.7,77.1,75.5,75.4,74.8,74.7,74.6,74.5,72.9,72.7,72.1,71.8,70.8,70.2,70.0,68.9,68.9,68.3,68.0,67.8,67.5,67.3,62.4,60.9,60.3,60.3,60.0,51.6,51.3,39.7,39.3,28.1,26.5,22.3,22.0,22.0;HRMS(ESI-TOF,MNa^+)计算值$C_{48}H_{83}N_3O_{34}Na$ 1268.4756,测量值1268.4760.

[0353]

表 7 鞘糖脂类的基本组合物

	Gal	Glc	GalNAc	GlcNAc	Neu5Ac	Fuc
Globo 系列						
Globo 四糖 (Gb4)	2	1	1	0	0	0
Globo 五糖 (Gb5)	3	1	1	0	0	0
Globo H	3	1	1	0	0	1

[0354]

表 7 鞘糖脂类的基本组合物

(岩藻糖 -Gb5)						
SSEA4 (唾液酸 -Gb5)	3	1	1	0	1	0
异 globo 四糖	2	1	1	0	0	0
新乳糖系列	2	1	0	1	1	0
乳糖系列	2	1	0	1	1	0
鞘糖脂系列	2	1	1	0	2	0

[0355]

表8 糖基化反应与再生反应的各步骤的产率			
	涉及的酶	产物	产率
步骤1.	GalK, USP, PykF, PPA, LgtC*	丙烯基-Gb3	99%
步骤2.	GalNAcK, GlmU, PykF, PPA, LgtD*	丙烯基-Gb4	96%
步骤3.	GalK, USP, PykF, PPA, LgtD*	丙烯基-Gb5	95%**
步骤4a.	FKP, PykF, PPA, FutC*	丙烯基-Globo H	94%
步骤4b.	Css, CMK, PykF, PPA, JT-FAJ-16*	丙烯基-SSEA4	45%

[0356] *DNA序列对大肠杆菌表达最优化.

[0357] **当使用纯丙烯基-Gb4作为受体.

[0358] 实施例2:从丙烯基-乳糖以一步骤(One-Step)合成丙烯基-Gb5(SSEA3)

[0359] 在无需纯化任何中间物下,自丙烯基-乳糖一步骤(one-step)合成丙烯基-Gb5的链反应是例示于图6。

[0360] 在锥形瓶中,将5mmol丙烯基-乳糖、5mmol半乳糖、12mmol PEP、0.25mmol ATP、0.25mmol UTP及10mM $MgCl_2$在100mM Tris-HCl缓冲液(pH7.5)中混合。在锥形瓶中加入适当的α-1,4-半乳糖转移酶(LgtC)、GalK、USP、PykF及PPA来启动酶反应并合成丙烯基-Gb3。将含有反应混合物的锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC分析法监测合成过程。当发现不再合成丙烯基-Gb3时,则加入更多的酶。

[0361] 于合成丙烯基-Gb3后,将其它组的组分(包括5mmol GalNAc、12mmol PEP及适当量的N-乙酰己糖胺1-激酶(GalNAcK)、N-乙酰葡萄糖胺1-磷酸尿苷酰转移酶(GlmU)、PykF、PPA及β-1,3-N-乙酰半乳糖胺转移酶(LgtD))加入锥形瓶中。在与合成丙烯基-Gb3的相同条件下培养所形成的反应混合物。当发现不合成丙烯基-Gb4时,则加入更多量的酶。

[0362] 于合成丙烯基-Gb4后,于未将丙烯基-Gb4纯化的情况下,将5mmol半乳糖及12mmol PEP加入锥形瓶中。通过加入适当的β-1,3-半乳糖转移酶(LgtD)、GalK、USP、PykF及PPA来启动下个半乳糖苷化反应来合成丙烯基-Gb5。将含有反应混合物的锥形瓶在温和摇荡下,置于16～50℃的培养箱中。利用TLC分析法监测合成过程。若发现不再合成丙烯基-Gb5时,可加入更多的酶。此由丙烯基-乳糖一步骤合成丙烯基-Gb5的产率为约40%。

[0363] 参考文献:

[0364] 1.卡纳吉,R等人,阶段特异性胚胎抗原(SSEA-3及SSEA-4))为从人类畸胎样瘤细胞分离出的球系列神经节苷脂的表位。欧洲分子生物期刊,1983.2(12):p.2355-61

(Kannagi,R.,et al.,Stage-specific embryonic antigens(SSEA-3and-4)are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells.EMBO J,1983.2(12):p.2355-61.)

[0365] 2.先戈,Z等人,利用免疫组织化学选择作为用于免疫攻击的标靶的癌症抗原:I.关注神经节苷脂。国际癌症杂志,1997.73(1):p.42-49(Shengle,Z.,et al.,Selection of tumor antigens as targets for immune attack using immunohistochemistry: I.Focus on gangliosides.International Journal of Cancer,1997.73(1):p.42-49.)

[0366] 3.先戈,Z等人,利用免疫组织化学选择作为用于免疫攻击的标靶的癌症抗原:II.血型相关抗原。国际癌症杂志,1997.73(1):p.50-56(Shengle,Z.,et al.,Selection of tumor antigens as targets for immune attack using immunohistochemistry: II.Blood group-related antigens.International Journal of Cancer,1997.73(1): p.50-56.)

[0367] 4.张,W.-W等人,在乳癌干细胞中的Globo H和SSEA3的表达,以及Globo H合成中的岩藻糖转移酶1和2的参与。美国国家科学院院刊,2008.105(33):p.11667-11672(Chang, W.-W.,et al.,Expression of Globo H and SSEA3in breast cancer stem cells and the involvement of fucosyl transferases 1 and 2in Globo H synthesis.Proceedings of the National Academy of Sciences,2008.105(33): p.11667-11672.)

[0368] 5.史莱门,D.J等人,化疗加上抗HER2的单株抗体于过度表达HER2转移性乳癌上的用途。新英格兰医学杂志,2001.344(11):p.783-792(Slamon,D.J.,et al.,Use of Chemotherapy plus a Monoclonal Antibody against HER2for Metastatic Breast Cancer That Overexpresses HER2.New England Journal of Medicine,2001.344(11): p.783-792.)

[0369] 6.王,Z.-G等人,来自以globo H钥孔虫戚血蓝蛋白疫苗免疫的病患的多克隆抗体:通过使用全合成的免疫肿瘤抗原作免疫反应的分离、定量、和定性。美国国家科学院院刊,2000.97(6):p.2719-2724(Wang,Z.-G.,et al.,Polyclonal antibodies from patients immunized with a globo H-keyhole limpet hemocyanin vaccine: Isolation,quantification,and characterization of immune responses by using totally synthetic immobilized tumor antigens.Proceedings of the National Academy of Sciences of the United States of America,2000.97(6):p.2719-2724.)

[0370] 7.毕多洛,M.T等人,人类乳癌肿瘤相关抗原的全合成。美国化学学会期刊,1995.117(29):p.7840-7841.(Bilodeau,M.T.,et al.,Total Synthesis of a Human Breast Tumor Associated Antigen.Journal of the American Chemical Society, 1995.117(29):p.7840-7841.)

[0371] 8.帕克,T.K等人,人类乳癌肿瘤(Globo-H)抗原的全合成和结构证明。美国化学学会期刊,1996.118(46):p.11488-11500(Park,T.K.,et al.,Total Synthesis and Proof of Structure of a Human Breast Tumor(Globo-H)Antigen.Journal of the American Chemical Society,1996.118(46):p.11488-11500.)

[0372] 9.斯洛芬,S.F等人,癌症中的碳水化合物疫苗:全合成globo H六糖共轭物在人类

上的免疫原性。美国国家科学院院刊,1999.96(10):p.5710-5715(Slovin,S.F.,et al.,Carbohydrate vaccines in cancer:Immunogenicity of a fully synthetic globo H hexasaccharide conjugate in man.Proceedings of the National Academy of Sciences of the United States of America,1999.96(10):p.5710-5715.)

[0373] 10.博塞,F、L.A.马可罗、和P.H.希博格,肿瘤相关碳水化合物抗原Globo-H、SSEA-3和Gb3的线性合成。有机化学期刊,2002.67(19):p.6659-6670(Bosse,F.,L.A.Marcaurelle,and P.H.Seeberger,Linear Synthesis of the Tumor-Associated Carbohydrate Antigens Globo-H,SSEA-3,and Gb3.The Journal of Organic Chemistry,2002.67(19):p.6659-6670.)

[0374] 11.沃兹,D.B.、B.肯斯达格诺、和P.H.希博格,肿瘤相关碳水化合物抗原Gb-3和Globo-H的自动化合成:á-半乳糖苷链接的结合。美国化学学会期刊,2007.129(10):p.2770-2771(Werz,D.B.,B.Castagner,and P.H.Seeberger,Automated Synthesis of the Tumor-Associated Carbohydrate Antigens Gb-3and Globo-H:Incorporation ofá-Galactosidic Linkages.Journal of the American Chemical Society,2007.129(10):p.2770-2771.)

[0375] 12.王,Z等人,基于硫代糖苷供体的预激活的肿瘤相关碳水化合物抗原Globo-H的多组份一锅合成。有机化学期刊,2007.72(17):p.6409-6420(Wang,Z.,et al.,Multi-Component One-Pot Synthesis of the Tumor-Associated Carbohydrate Antigen Globo-H Based on Preactivation of Thioglycosyl Donors.The Journal of Organic Chemistry,2007.72(17):p.6409-6420.)

[0376] 13.简,I、K.艾尔和S.J.狄尼薛夫斯基,用于抗癌疫苗的Globo-H的实用全合成。有机化学期刊,2009.74(21):p.8452-8455(Jeon,I.,K.Iyer,and S.J.Danishefsky,A Practical Total Synthesis of Globo-H for Use in Anticancer Vaccines.The Journal of Organic Chemistry,2009.74(21):p.8452-8455.)

[0377] 14.弗雷,B等人,使用可程序化反应基础一锅法进行Globo-H六糖的合成。应用化学杂志国际版,2001.40(7):p.1274-1277(Fred,B.,et al.,Synthesis of the Globo H Hexasaccharide Using the Programmable Reactivity-Based One-Pot Strategy.Angewandte Chemie International Edition,2001.40(7):p.1274-1277.)

[0378] 15.许,C.-H等人,用于有效制备天然唾液酸糖苷的高度α选择性磷酸唾液酸供体。欧洲化学,2010.16(6):p.1754-1760(Hsu,C.-H.,et al.,Highly Alpha-Selective Sialyl Phosphate Donors for Efficient Preparation of Natural Sialosides.Chemistry-A European Journal,2010.16(6):p.1754-1760.)

[0379] 16.甄,W等人,肿瘤相关碳水化合物抗原Globo-H和阶段特异性胚胎抗原4的化学酶合成。高级合成与催化,2008.350(11-12):p.1717-1728(Zhen,W.,et al.,Chemoenzymatic Syntheses of Tumor-Associated Carbohydrate Antigen Globo-H and Stage-Specific Embryonic Antigen 4.Advanced Synthesis & Catalysis,2008.350(11-12):p.1717-1728.)

[0380] 17.皇,C.H.,S.L.海涅和G.M.怀德赛,二磷酸尿苷葡萄糖和二磷酸尿苷半乳糖原位再生的N-乙酰基乳糖胺的酶催化合成。有机化学期刊,1982.47(27):p.5416-5418(Wong,

C.H.,S.L.Haynie,and G.M.Whitesides,Enzyme-catalyzed synthesis of N-acetyllactosamine with in situ regeneration of uridine 5'-diphosphate glucose and uridine 5'-diphosphate galactose.The Journal of Organic Chemistry,1982.47(27):p.5416-5418.)

[0381] 18.皇,C.H.,R.王和Y.市川,用于酶寡糖合成的糖核苷酸再生:UDP-半乳糖、UDP-2-脱氧半乳糖和UDP-半乳糖胺再生中的半乳糖-1-磷酸尿苷酸转移酶的用途。有机化学期刊,1992.57(16):p.4343-4344(Wong,C.H.,R.Wang,and Y.Ichikawa,Regeneration of sugar nucleotide for enzymic oligosaccharide synthesis:use of Gal-1-phosphate uridyltransferase in the regeneration of UDP-galactose,UDP-2-deoxygalactose, and UDP-galactosamine.The Journal of Organic Chemistry,1992.57(16):p.4343-4344.)

[0382] 19.小竹,T等人,具有针对豌豆苗各种单糖1-磷酸的广底物特异性的UDP-糖类焦磷酸化酶。生物化学杂志,2004.279(44):p.45728-45736(Kotake,T.,et al.,UDP-sugar Pyrophosphorylase with Broad Substrate Specificity Toward Various Monosaccharide 1-Phosphates from Pea Sprouts.Journal of Biological Chemistry, 2004.279(44):p.45728-45736.)

[0383] 20.理特勒,L.A.等人,拟南芥的UDP-糖类焦磷酸化酶的表征和表达。植物生理与生物化学,2006.44(4):p.171-180(Litterer,L.A.,et al.,Characterization and expression of Arabidopsis UDP-sugar pyrophosphorylase.Plant Physiology and Biochemistry,2006.44(4):p.171-180.)

[0384] 21.小竹,T等人,拟南芥的UDP-糖类焦磷酸化酶的性质和生理功能。生物科学生技及生化,2007.71(3):p.761-771(Kotake,T.,et al.,Properties and Physiological Functions of UDP-Sugar Pyrophosphorylase in Arabidopsis.Bioscience, Biotechnology,and Biochemistry,2007.71(3):p.761-771.)

[0385] 22.丹姆诺,S等人,利什曼原虫UDP-糖类焦磷酸化酶。生物化学杂志,2010.285(2):p.878-887(Damerow,S.,et al.,Leishmania UDP-sugar Pyrophosphorylase.Journal of Biological Chemistry,2010.285(2):p.878-887.)

[0386] 23.杨,T.和M.巴佩尔,在克氏锥虫中具有广底物特异性的新颖UDP-糖类焦磷酸化酶的鉴定。生物化学杂志,2010.429(3):p.533-543(Yang,T.and M.Bar-Peled, Identification of a novel UDP-sugar pyrophosphorylase with a broad substrate specificity in Trypanosoma cruzi.Biochemical Journal,2010.429(3):p.533-543.)

[0387] 24.佩森,K等人,脑膜炎双球菌来源的保留半乳糖转移酶LgtC为供体和受体糖类似物的复合物的结晶结构。自然结构与分子生物学杂志,2001.8(2):p.166-175(Persson, K.,et al.,Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs.Nat Struct Mol Biol,2001.8(2):p.166-175.)

[0388] 25.江东,S等人,简易制备麦芽糖、纤维二糖、和乳糖的烷基和环烷基á-糖苷。碳水化合物研究,2004.339(14):p.2415-2424(Koto,S.,et al.,Simple preparations of alkyl and cycloalkylá-glycosides of maltose,cellobiose,and

lactose.Carbohydrate Research,2004.339(14):p.2415-2424.)

[0389] 26.安托,T等人,通过代谢工程大肠杆菌的高细胞密度培养来作格罗博三糖和红细胞糖苷脂四糖的大规模体内合成。生物化学,2005.87(2):p.197-203(Antoine,T.,et al.,Large scale in vivo synthesis of globotriose and globotetraose by high cell density culture of metabolically engineered Escherichia coli.Biochimie,2005.87(2):p.197-203.)

[0390] 27.张,J等人,具重组的α-1,4-半乳糖基转移酶的格罗博三糖和其衍生物的有效化学酶合成。碳水化合物研究,2002.337(11):p.969-976(Zhang,J.,et al.,Efficient chemoenzymatic synthesis of globotriose and its derivatives with a recombinant[alpha]-(1-->4)-galactosyltransferase.Carbohydrate Research,2002.337(11):p.969-976.)

[0391] 28.邵,J等人,通过使用â(1→3)N-乙酰氨基半乳糖基转移酶/UDP-乙酰葡萄糖胺C4表异构酶融合蛋白的红细胞糖苷脂和异红细胞糖苷脂的有效合成。化学通讯,2003(12):p.1422-1423(Shao,J.,et al.,Efficient synthesis of globoside and isogloboside tetrasaccharides by using â(1 →3)N-acetylgalactosaminyltransferase/UDP-N-acetylglucosamine C4 epimerase fusion protein.Chemical Communications,2003(12):p.1422-1423.)

[0392] 29.邵,J等人,有效合成红细胞糖苷脂四糖和异红细胞糖苷脂四糖的供体底物再生。应用与环境微生物学,2002.68(11):p.5634-5640(Shao,J.,et al.,Donor Substrate Regeneration for Efficient Synthesis of Globotetraose and Isoglobotetraose.Appl.Environ.Microbiol.,2002.68(11):p.5634-5640.)

[0393] 30.凡妮莎,B等人,GalNAc和UTP来源的肽类和寡糖的有效酶糖基化。生物化学,2007.8(1):p.37-40(Vanessa,B.,et al.,Efficient Enzymatic Glycosylation of Peptides and Oligosaccharides from GalNAc and UTP.ChemBioChem,2007.8(1):p.37-40.)

[0394] 31.西本,M和M.北冈,比菲德氏-龙根菌的完整乳-N-二糖1/半乳-N-二糖代谢途径中的N-乙酰六碳糖激酶的鉴定。应用与环境微生物学,2007.73(20):p.6444-6449(Nishimoto,M.and M.Kitaoka,Identification of N-Acetylhexosamine 1-Kinase in the Complete Lacto-N-Biose I/Galacto-N-Biose Metabolic Pathway in Bifidobacterium longum.Appl.Environ.Microbiol.,2007.73(20):p.6444-6449.)

[0395] 32.方,J等人,对于大肠杆菌K12来源的N-乙酰葡萄糖胺-1-磷酸尿苷转移酶的焦磷酸化酶结构域的广核苷三磷酸特异性的系统研究。生物有机化学与医药化学通讯,2009.19(22):p.6429-6432(Fang,J.,et al.,Systematic study on the broad nucleotide triphosphate specificity of the pyrophosphorylase domain of the N-acetylglucosamine-1-phosphate uridyltransferase from Escherichia coli K12.Bioorganic & Medicinal Chemistry Letters,2009.19(22):p.6429-6432.)

[0396] 33.胡德,D.W.等人,流感嗜血杆菌菌株Rd(RM118)的主要含红细胞糖苷脂四糖的脂多糖类的表达基因基础。糖生物学,2001.11(11):p.957-967(Hood,D.W.,et al.,Genetic basis for expression of the major globotetraose-containing

lipopolysaccharide from H.influenzae strain Rd(RM118).Glycobiology,2001.11(11):p.957-967.)

[0397] 34.邵,J等人,流感嗜血杆菌菌株Rd的β-1,3-N-乙酰氨基半乳糖转移酶LgtD的过度表达和生化特性。生物化学与生物物理研究通讯,2002.295(1):p.1-8(Shao,J.,et al.,Overexpression and biochemical characterization of[beta]-1,3-N-acetylgalactosaminyltransferase LgtD from Haemophilus influenzae strain Rd.Biochemical and Biophysical Research Communications,2002.295(1):p.1-8.)

[0398] 35.瑞吉安索,M等人,使用流感嗜血杆菌β-1,3-N-乙酰氨基半乳糖转移酶LgtD的新颖β-1,3-半乳糖转移酶活性的格罗博五糖合成。FEBS通讯,2007.581(14):p.2652-2656(Randriantsoa,M.,et al.,Synthesis of globopentaose using a novel[beta]1,3-galactosyltransferase activity of the Haemophilus influenzae[beta]1,3-N-acetylgalactosaminyltransferase LgtD.FEBS Letters,2007.581(14):p.2652-2656.)

[0399] 36.柯尼,M.J.等人,人类共生体对于表面岩藻糖化作用使用类宿主途径。科学,2005.307(5716):p.1778-1781(Coyne,M.J.,et al.,Human Symbionts Use a Host-Like Pathway for Surface Fucosylation.Science,2005.307(5716):p.1778-1781.)

[0400] 37.王,W等人,GDP-l-岩藻糖和刘易斯X聚糖衍生物的化学酶合成。美国国家科学院院刊,2009.106(38):p.16096-16101(Wang,W.,et al.,Chemoenzymatic synthesis of GDP-l-fucose and the Lewis X glycan derivatives.Proceedings of the National Academy of Sciences,2009.106(38):p.16096-16101.)

[0401] 38.小泉,S等人,GDP-岩藻糖和刘易斯X通过细菌耦合的大规模合成。工业微生物学与生物技术杂志,2000.25(4):p.213-217(Koizumi,S.,et al.,Large-scale production of GDP-fucose and Lewis X by bacterial coupling.Journal of Industrial Microbiology and Biotechnology,2000.25(4):p.213-217.)

[0402] 39.史坦,丹尼尔B、Y.-N.林和C.-H.林,用于肿瘤相关抗原的酶合成的幽门螺杆菌α-1,2-岩藻糖转移酶的特性。高级合成与催化,2008.350(14-15):p.2313-2321(Stein,Daniel B.,Y.-N.Lin,and C.-H.Lin,Characterization of Helicobacter pylori alpha1,2-Fucosyltransferase for Enzymatic Synthesis of Tumor-Associated Antigens.Advanced Synthesis & Catalysis,2008.350(14-15):p.2313-2321.)

[0403] 40.丹纳,M.E.,唾液酸生物合成的酶。生物有机化学,2005.33(3):p.216-228(Tanner,M.E.,The enzymes of sialic acid biosynthesis.Bioorganic Chemistry,2005.33(3):p.216-228.)

[0404] 41.李,Y等人,具改善底物杂交性的多巴性巴氏杆菌CMP-唾液酸合成酶和突变体脑膜炎双球菌CMP-唾液酸合成酶。应用微生物学与生物技术,2011:p.1-13(Li,Y.,et al.,Pasteurella multocida CMP-sialic acid synthetase and mutants of Neisseria meningitidis CMP-sialic acid synthetase with improved substrate promiscuity.Applied Microbiology and Biotechnology,2011:p.1-13.)

[0405] 42.库什,Y等人,海洋细菌的唾液酸转移酶有效地利用甘胺酸脂质底物。糖类生物学,2010.20(2):p.187-198(Kushi,Y.,et al.,Sialyltransferases of marine bacteria efficiently utilize glycosphingolipid substrates.Glycobiology,2010.20(2):

p.187-198.)

[0406] 43.山本,T,海洋细菌唾液酸转移酶。海洋药物,2010.8(11):p.2781-2794(Yamamoto,T.,Marine Bacterial Sialyltransferases.Marine Drugs,2010.8(11):p.2781-2794.)

[0407] 44.高仓,Y、H.冢本和T.山本,弧菌类JT-FAJ-16来源的á/â-半乳糖苷α-2,3-唾液酸转移酶的分子转殖、表达和特性。生物化学期刊,2007.142(3):p.403-412(Takakura,Y.,H.Tsukamoto,and T.Yamamoto,Molecular Cloning,Expression and Properties of an α/β-Galactosideα2,3-Sialyltransferase from Vibrio sp.JT-FAJ-16.Journal of Biochemistry,2007.142(3):p.403-412.)

[0408] 其它实施方式

[0409] 除非上下文中另有相反或其它说明,权利要求书中所述的“一”、“一个”及“该”可表示一或大于一个。除非上下文中另有关于产物或方法的相反或其它说明,当权利要求书或说明书中,在一组的一或多个成员间出现“或”时,则视为其中出现或使用一个、大于一个或所有的成员。关于本发明所提供的产物或方法中的方式包含出现或使用刚好一个成员。关于本发明所提供的产物或方法中的方式包含出现或使用大于一个或所有的成员。

[0410] 再者,本发明包含所有的变异物、组合及替换,所列权利要求中的一或多个限制、组件、子句及描述的术语也被引入另一权利要求中。例如,依附至另一权利要求的任何权利要求可经修饰,进而包括依附至相同基础权利要求的其它权利要求中的一或多个限制。当权利要求中列有组件时(如马库什型式),所有组件的各次群也包括于内,且任意组件皆可自该群中移除。应理解的是,一般而言,当发明或发明方式是包含特定组件及/或特征、该发明的特定方式或由该组件及/或特征所组成或实质上组成的发明方式。为达简化的目的,这些未特定揭示的方式也包含于本发明中。另须注意者,“包含”及“包括”为开放性用语,其容许含有额外的组件或步骤。当提供一范围时,其也包含端点值。再者,除非上下文中另有相反或其它说明,本领域技术人员能了解,在不同发明方式中,该数值范围可包含特定值或于所述范围内的次范围,且除非上下文中另有相反或其它说明,该范围最低值的单位的十分的一也包含于其中。

[0411] 本案中所提及的不同专利案件、公开的专利申请案、文献资料及其它的公开,皆以全文的方式并入本案。若所并入的参考文件与本案说明书有冲突时,以本案说明书的内容为准。再者,本发明中落入现有技术的任何特定方式,可自本案的一或多个权利要求中排除。因该方式被视为本领域技术人员所知,故即使其未明确地被排除,其可被排除。不论是基于任何理由或是否与现有技术有关,本发明中的任何特定方式皆可移除。

[0412] 本领域人员经由例行实验即可了解或推知的对等物皆为本发明的一部份。上述说明书的内容并非用于局限本发明的范围,权利要求请求保护的内容方为本发明的范围。本领域技术人员在不悖离所附权利要求书所定义本发明的精神或范畴的情况下可对本文所揭露的方式进行修改及改变。

按照专利合作条约第 28 或 41 条修改的替换页

序列表

<110> 中央研究院

<120> 寡糖的大规模酶合成的方法

<130> 148826 PWCN

<150> US 61/684,974

<151> 2012-08-20

<160> 30

<170> PatentIn version 3.5

<210> 1
<211> 45
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 1
ctgtattttc agggagcgat cgctatgagt ctgaaagaaa aaaca 45

[0001]

<210> 2
<211> 44
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 2
gcctcgagtc attacgttta aactcagcac tgtcctgctc cttg 44

<210> 3
<211> 44
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 3
ctgtattttc agggagcgat cgctatggct tctacggttg attc 44

<210> 4
<211> 46
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 4
gcctcgagtc attacgttta aactcaatct tcaacagaaa atttgc 46

<210> 5
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 5
gatataccat ggaaatggac atcgttttcg cggcg 35

<210> 6
<211> 33
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 6
gtggtgctcg aggtagattt tacgcaggaa acg 33

<210> 7
<211> 50
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 7
[0002] ctgtatttte agggagcgat cgctatgaac aagacttatg attttaaaag 50

<210> 8
<211> 48
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 8
gcctcgagtc attacgttta aacttaaatg tatgaatata ctatcttc 48

<210> 9
<211> 45
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 9
ctgtatttte agggagcgat cgctatgttg aataatgcta tgagc 45

<210> 10
<211> 44
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸

<400> 10
gcctcgagtc attacgttta aactcacttt ttctttaccg gacg 44

<210> 11
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 11
gatataccat ggaaaactgc ccgctggttt ct 32

<210> 12
<211> 34
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 12
gtggtgctcg aggaagataa cgttgatttt acgg 34

<210> 13
<211> 36
<212> DNA
<213> 人工序列
[0003] <220>
<223> 合成寡核苷酸
<400> 13
cagggagcga tcgctatgca aaaactacta tcttta 36

<210> 14
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 14
cattacgttt aaacttatga tcgtgatact tggaa 35

<210> 15
<211> 45
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 15
ctgtatttte agggagcgat cgctatggcg ttcaaagttg ttcag 45

<210> 16
<211> 45
<212> DNA

<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 16
gcctcgagtc attacgttta aacttacgcg ttgtatttct gagat 45

<210> 17
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 17
cagggagcga tcgctatgac ggcaattgcc ccggtt 36

<210> 18
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 18
cattacgttt aaacttatgc gagagccaat ttctg 35

[0004]

<210> 19
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 19
gatataccat ggaaacaaat attgcgatca ttcctg 36

<210> 20
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 20
gtggtgctcg agtttattgg ataaaatttc cgcgag 36

<210> 21
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 21
gatataccat ggaaatgaac aacgacaact ctacc 35

<210> 22
<211> 34
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 22
gtggtgctcg aggatgtcag agatcagttt gatg 34

<210> 23
<211> 47
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 23
ctgtatttte agggagcgat cgctatgaaa aagaccaaaa ttgtttg 47

<210> 24
<211> 43
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 24
[0005] gcctcgagtc attacgttta aacttacagg acgtgaacag atg 43

<210> 25
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 25
cagggagcga tcgctatgag cttactcaac gtccct 36

<210> 26
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成寡核苷酸
<400> 26
cattacgttt aaacttattt attctttgcg cgctc 35

<210> 27
<211> 936
<212> DNA
<213> 大肠杆菌 (Escherichia coli)
<400> 27
atggacatcg ttttcgcggc ggacgacaac tacgcggcgt acctgtgcgt tgcggcgaaa 60

tctgttgaag cggcgcaccc ggacaccgaa atccgtttcc acgttctgga cgcgggtatc 120
tctgaagcga accgtgcggc ggttgcggcg aacctgcgtg gtggtggtgg taacatccgt 180
ttcatcgacg ttaacccgga agacttcgcg ggtttcccgc tgaacatccg tcacatctct 240
atcaccacct acgcgcgtct gaaactgggt gaatacatcg cggactgcga caaagttctg 300
tacctggaca tcgacgttct ggttcgtgac tctctgaccc cgctgtggga caccgacctg 360
ggtgacaact ggctgggtgc gtgcatcgac ctgttcgttg aacgtcagga aggttacaaa 420
cagaaaatcg gtatggcgga cggtgaatac tacttcaacg cgggtgttct gctgatcaac 480
ctgaaaaaat ggcgtcgtca cgacatcttc aaaatgtctt gcgaatgggt tgaacagtac 540
aaagacgtta tgcagtacca ggaccaggac atcctgaacg gtctgttcaa aggtggtgtt 600
tgctacgcga actctcgttt caacttcatg ccgaccaact acgcgttcat ggcgaaccgt 660
ttcgcgtctc gtcacaccga cccgctgtac cgtgaccgta ccaacaccgt tatgccggtt 720
gcggtttctc actactgcgg tccggcgaaa ccgtggcacc gtgactgcac cgcgtggggt 780
gcggaacgtt tcaccgaact ggcgggttct ctgaccaccg ttccggaaga atggcgtggt 840
aaactggcgg ttccgcaccg tatgttctct accaaacgta tgctgcagcg ttggcgtcgt 900
aaactgtctg cgcgtttcct gcgtaaaatc tactga 936

<210> 28
<211> 972
<212> DNA
<213> 大肠杆菌（Escherichia coli）
<400> 28
atggaaaact gcccgctggt tctgttatc gttgcgcgt acaacgcgga acagtacatc 60
gacgaatcta tctcttctat catcaaccag acctacgaaa acctggaaat catcgttatc 120
aacgacggtt ctaccgacct gaccctgtct cacctggaag aaatctctaa actggacaaa 180
cgtatcaaaa tcatctctaa caaatacaac ctgggtttca tcaactctct gaacatcggt 240
[0006] ctgggttgct tctctggtaa atacttcgcg cgtatggacg cggacgacat cgcgaaaccg 300
tcttggatcg aaaaaatcgt tacctacctg gaaaaaaacg accacatcac cgcgatgggt 360
tcttacctgg aaatcatcgt tgaaaaagaa tgcggtatca tcggttctca gtacaaaacc 420
ggtgacatct ggaaaaaccc gctgctgcac aacgacatct gcgaagcgat gctgttctac 480
aacccgatcc acaacaacac catgatcatg cgtgcgaacg tttaccgtga acacaaactg 540
atcttcaaca aagactaccc gtacgcggaa gactacaaat tctggtctga agtttctcgt 600
ctgggttgcc tggcgaacta cccggaagcg ctggttaaat accgtctgca cggtaaccag 660
acctcttctg tttacaacca cgaacagaac gaaaccgcga aaaaaatcaa acgtgaaaac 720
atcacctact acctgaacaa aatcggtatc gacatcaaag ttatcaactc tgtttctctg 780
ctggaaatct accacgttga caaatctaac aaagttctga aatctatcct gtacgaaatg 840
tacatgtctc tggacaaata caccatcacc tctctgctgc acttcatcaa ataccacctg 900
gaactgttcg acctgaaaca gaacctgaaa atcatcaaaa aattcatccg taaaatcaac 960
gttatcttct ag 972

<210> 29
<211> 903
<212> DNA
<213> 大肠杆菌（Escherichia coli）
<400> 29
atggcgttca aagttgttca gatctgcggt ggtctgggta accagatgtt ccagtacgcg 60
ttcgcgaaat ctctgcagaa acactctaac accccggttc tgctggacat cacctctttc 120
gactggtctg accgtaaaat gcagctggaa ctgttcccga tcgacctgcc gtacgcgtct 180
gcgaaagaaa tcgcgatcgc gaaaatgcag cacctgccga aactggttcg tgacgcgctg 240
aaatgcatgg gtttcgaccg tgtttctcag gaaatcgttt tcgaatacga accgaaactg 300
ctgaaaccgt ctcgtctgac ctacttcttc ggttacttcc aggacccgcg ttacttcgac 360
gcgatctctc cgctgatcaa acagaccttc accctgccgc cgccgccgga aaacaacaaa 420

aacaacaaca aaaaagaaga agaataccag tgcaaactgt ctctgatcct ggcggcgaaa 480
aactctgttt tcgttcacat ccgtcgtggt gactacgttg gtatcggttg ccagctgggt 540
atcgactacc agaaaaaagc gctggaatac atggcgaaac gtgttccgaa catggaactg 600
ttcgttttct gcgaagacct ggaattcacc cagaacctgg acctgggtta cccgttcatg 660
gacatgacca cccgtgacaa agaagaagaa gcgtactggg acatgctgct gatgcagtct 720
tgccagcacg gtatcatcgc gaactctacc tactcttggt gggcggcgta cctgatcgaa 780
aacccggaaa aaatcatcat cggtccgaaa cactggctgt tcggtcacga aaacatcctg 840
tgcaaagaat gggttaaaat cgaatctcac ttcgaagtta aatctcagaa atacaacgcg 900
taa 903

<210> 30
<211> 1146
<212> DNA
<213> 大肠杆菌（Escherichia coli）
<400> 30

atgaacaacg acaactctac caccaccaac aacaacgcga tcgaaatcta cgttgaccgt 60
gcgaccctgc cgaccatcca gcagatgacc aaaatcgttt ctcagaaaac ctctaacaaa 120
[0007] aaactgatct cttggtctcg ttacccgatc accgacaaat ctctgctgaa aaaaatcaac 180
gcggaattct tcaaagaaca gttcgaactg accgaatctc tgaaaaacat catcctgtct 240
gaaaacatcg acaacctgat catccacggt aacaccctgt ggtctatcga cgttgttgac 300
atcatcaaag aagttaacct gctgggtaaa aacatcccga tcgaactgca cttctacgac 360
gacggttctg cggaatacgt tcgtatctac gaattctcta aactgccgga atctgaacag 420
aaatacaaaa cctctctgtc taaaaacaac atcaaattct ctatcgacgg taccgactct 480
ttcaaaaaca ccatcgaaaa catctacggt ttctctcagc tgtacccgac cacctaccac 540
atgctgcgtg cggacatctt cgacaccacc ctgaaaatca acccgctgcg tgaactgctg 600
tctaacaaca tcaaacagat gaaatgggac tacttcaaag acttcaacta caaacagaaa 660
gacatcttct actctctgac caacttcaac ccgaaagaaa tccaggaaga cttcaacaaa 720
aactctaaca aaaacttcat ctttatcggt tctaactctg cgaccgcgac cgcggaagaa 780
cagatcaaca tcatctctga agcgaaaaaa gaaaactctt ctatcatcac caactctatc 840
tctgactacg acctgttctt caaaggtcac ccgtctgcga ccttcaacga acagatcatc 900
aacgcgcacg acatgatcga aatcaacaac aaaatcccgt tcgaagcgct gatcatgacc 960
ggtatcctgc cggacgcggt tggtggtatg ggttcttctg ttttcttctc tatcccgaaa 1020
gaagttaaaa acaaattcgt tttctacaaa tctggtaccg acatcgaaaa caactctctg 1080
atccaggtta tgctgaaact gaacctgatc aaccgtgaca acatcaaact gatctctgac 1140
atctaa 1146

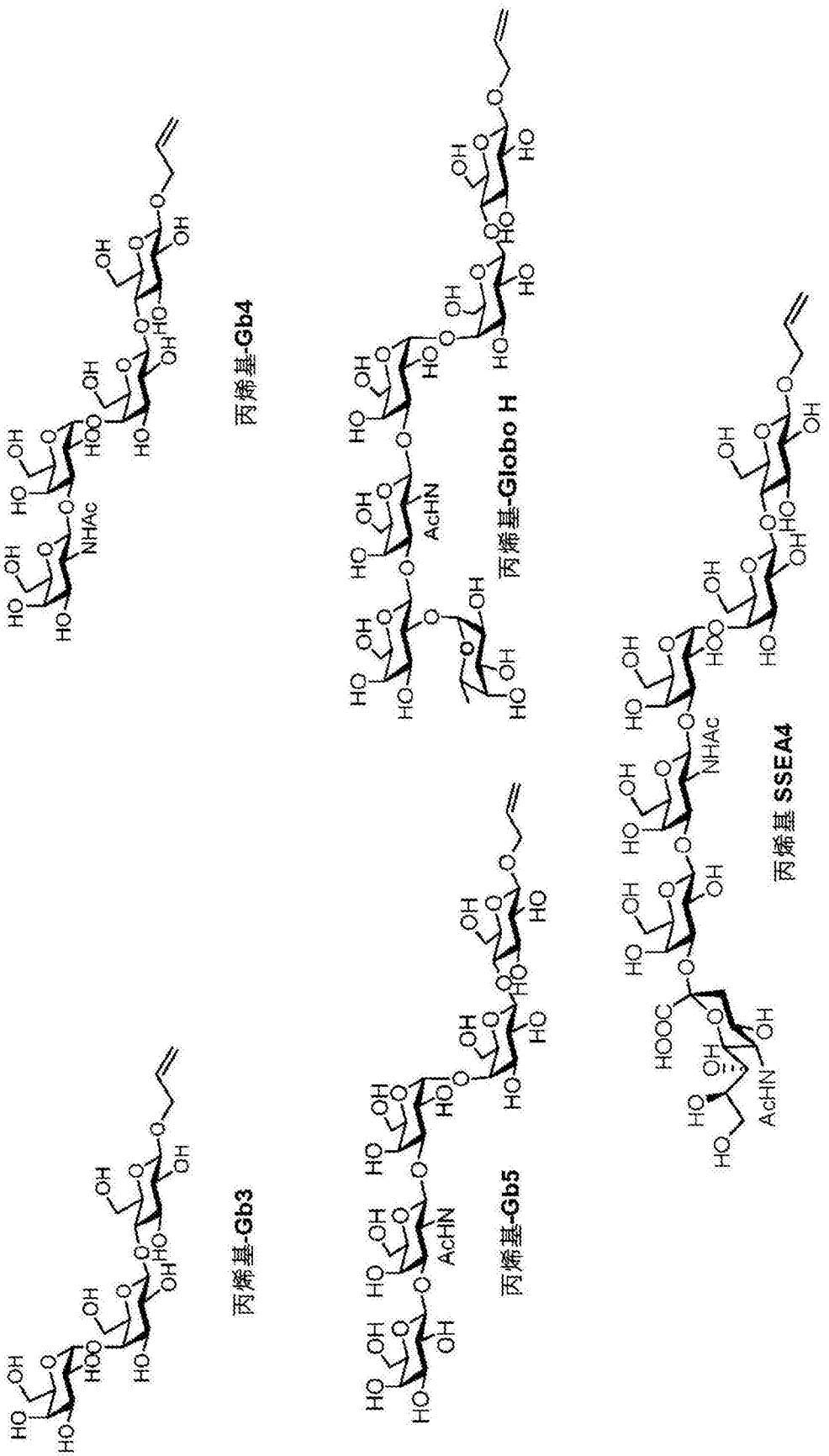
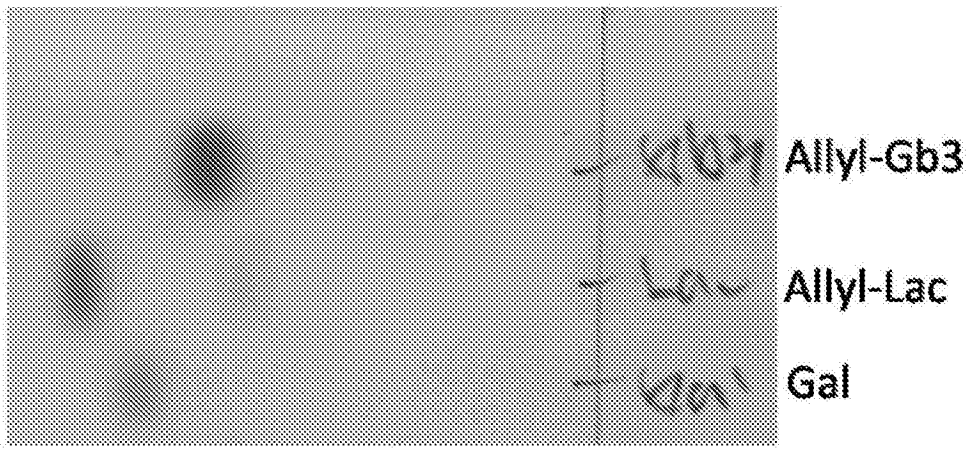


图1


Allyl-Gb3
Allyl-Lac
Gal

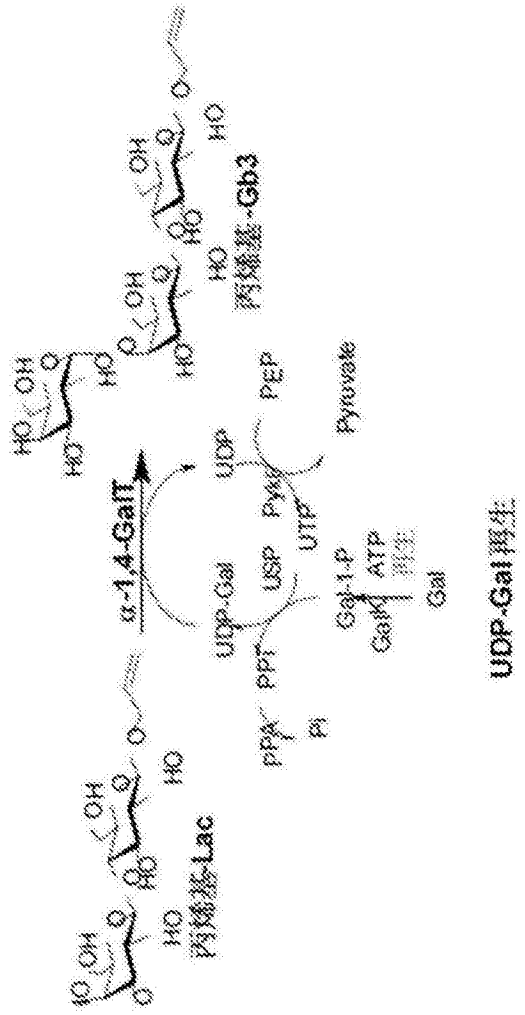
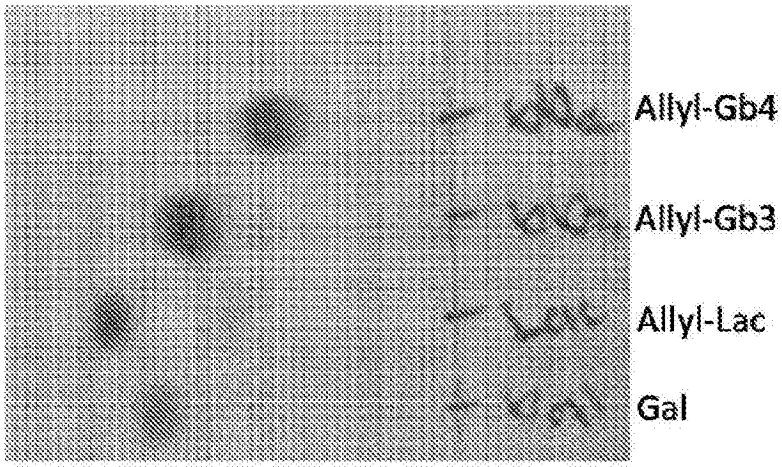

α-1,4-GalT
UDP-Gal
UDP
UTP
PEP
Pyruvate
PPi
Gal-1-P
ATP
Gal
UDP-Gal 再生
丙烯基-Gb3
丙烯基-Lac

图2A


Allyl-Gb4
Allyl-Gb3
Allyl-Lac
Gal

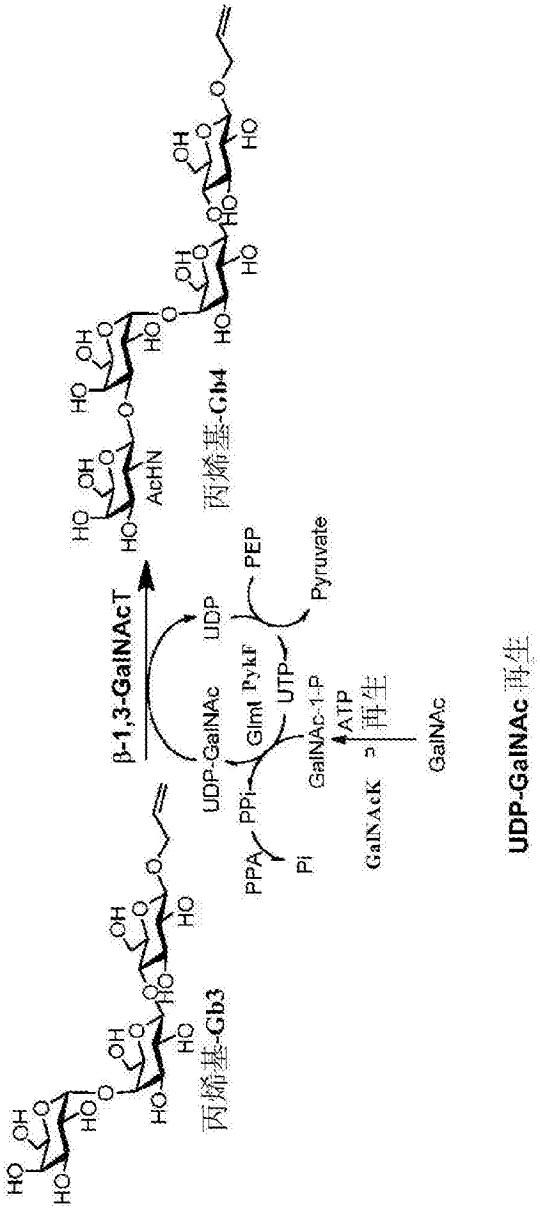

丙烯基-Gb4
β-1,3-GalNAcT
UDP
PEP
Pyruvate
UDP-GalNAc
GlmU PykF
UTP
GalNAc-1-P
ATP
再生
PPi
PPA
Pi
GalNAcK
GalNAc
UDP-GalNAc 再生
丙烯基-Gb3

图2B

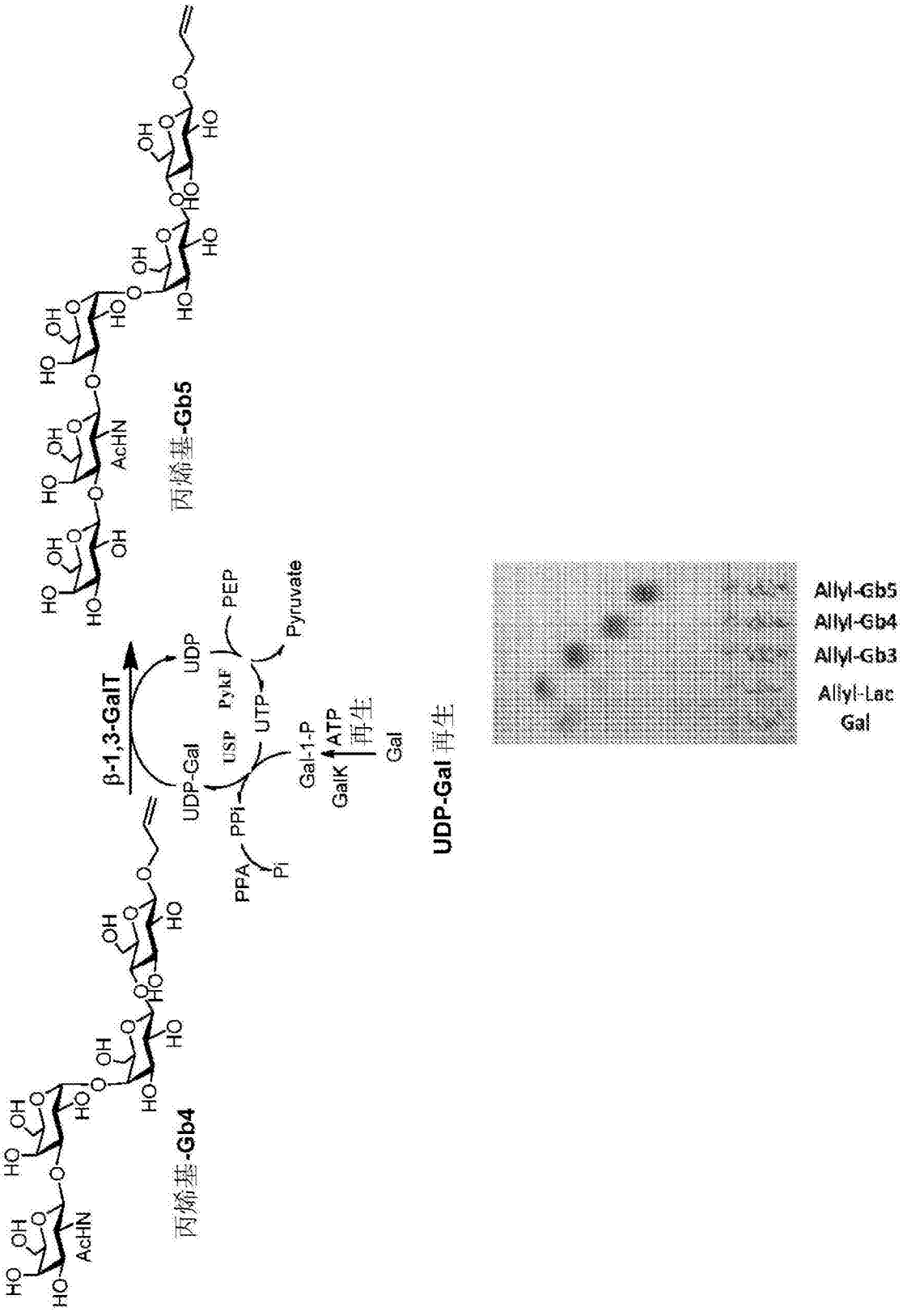

丙烯基-Gb4
丙烯基-Gb5
β-1,3-GalT
UDP-Gal
UDP
USP
PykF
PEP
Pyruvate
UTP
Gal-1-P
GalK
ATP
再生
Gal
PPA
PPi
Pi
UDP-Gal 再生
Allyl-Gb5
Allyl-Gb4
Allyl-Gb3
Allyl-Lac
Gal

图2C

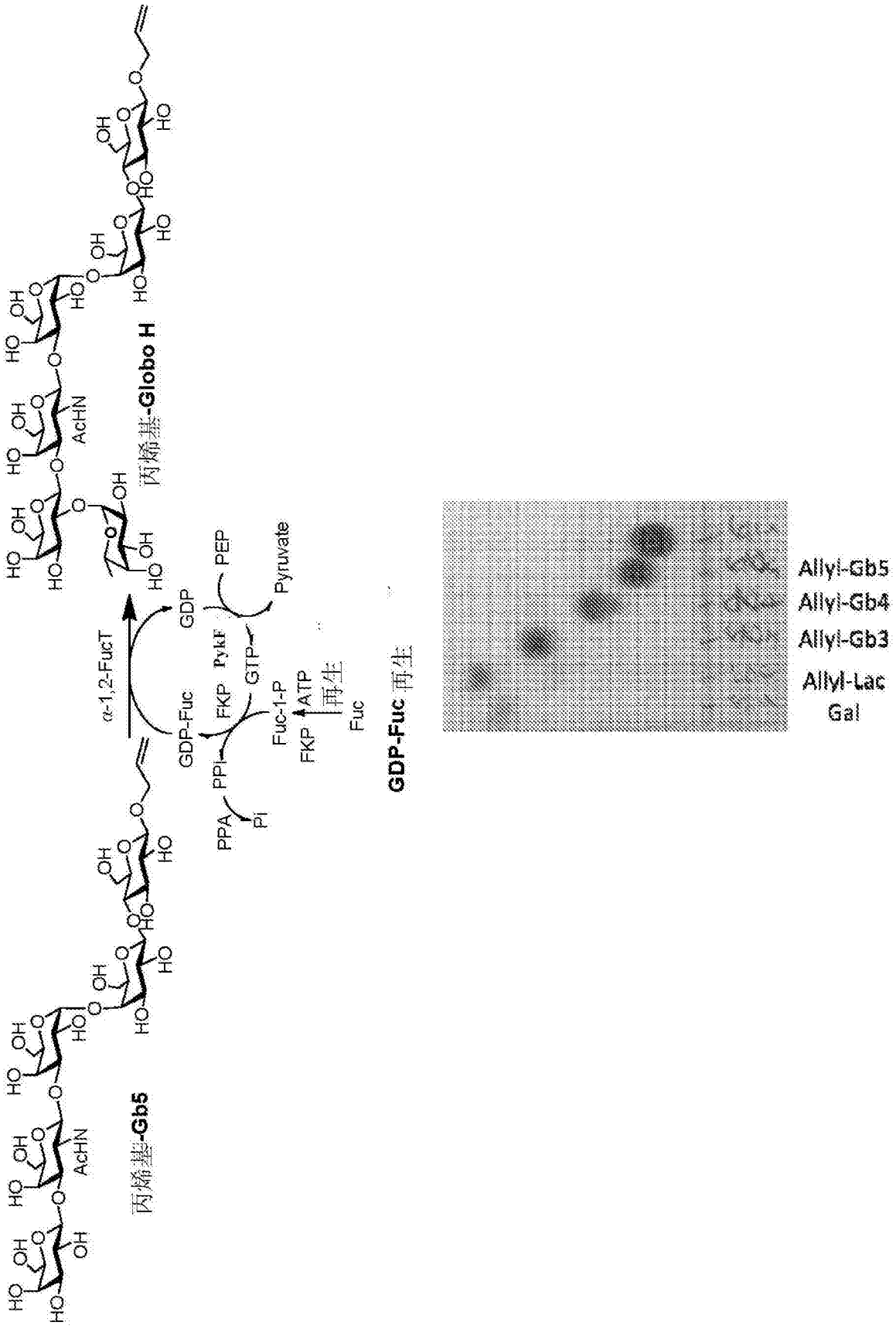

丙烯基-Gb5
丙烯基-Globo H
α-1,2-FucT
GDP-Fuc
GDP
FKP
PykF
PEP
Pyruvate
GTP
Fuc-1-P
ATP
FKP
再生
Fuc
PPA
PPi
Pi
GDP-Fuc 再生
Allyl-Gb5
Allyl-Gb4
Allyl-Gb3
Allyl-Lac
Gal

图2D

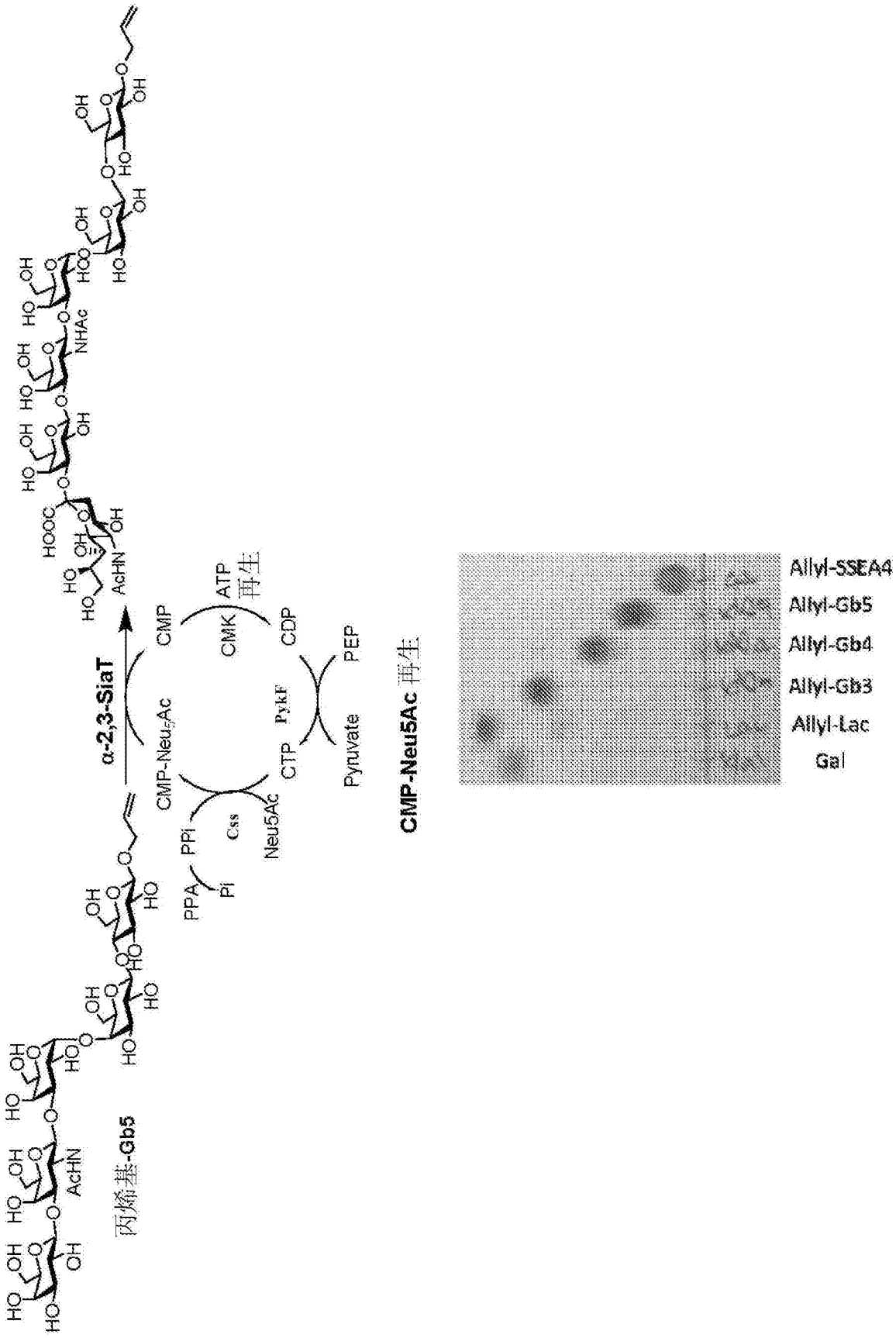

丙烯基-Gb5
α-2,3-SiaT
CMP-Neu5Ac
CMP
ATP
CMK
再生
CDP
PykF
PEP
Pyruvate
CTP
Neu5Ac
Css
PPi
PPA
Pi
CMP-Neu5Ac 再生
Allyl-SSEA4
Allyl-Gb5
Allyl-Gb4
Allyl-Gb3
Allyl-Lac
Gal

图2E

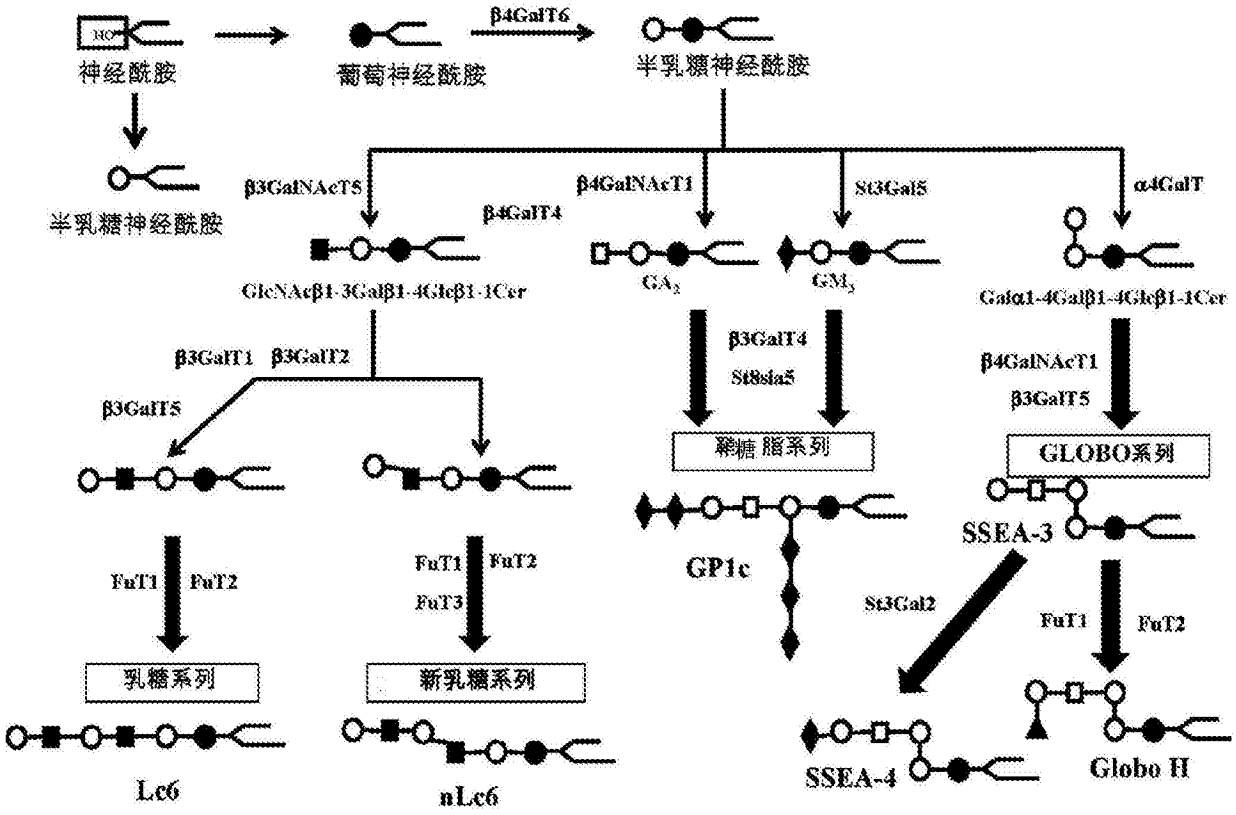

神经酰胺
葡萄神经酰胺
β4GalT6
半乳糖神经酰胺
半乳糖神经酰胺
β3GalNAcT5
β4GalT4
β4GalNAcT1
St3Gal5
α4GalT
GlcNAcβ1-3Galβ1-4Glcβ1-1Cer
GA2
GM3
Galα1-4Galβ1-4Glcβ1-1Cer
β3GalT1 β3GalT2
β3GalT5
β3GalT4
St8sia5
β4GalNAcT1
β3GalT5
神经节脂系列
GLOBO系列
GP1c
SSEA-3
FuT1 FuT2
FuT1 FuT2
FuT3
St3Gal2
FuT1 FuT2
乳糖系列
新乳糖系列
Lc6
nLc6
SSEA-4
Globo H

图3

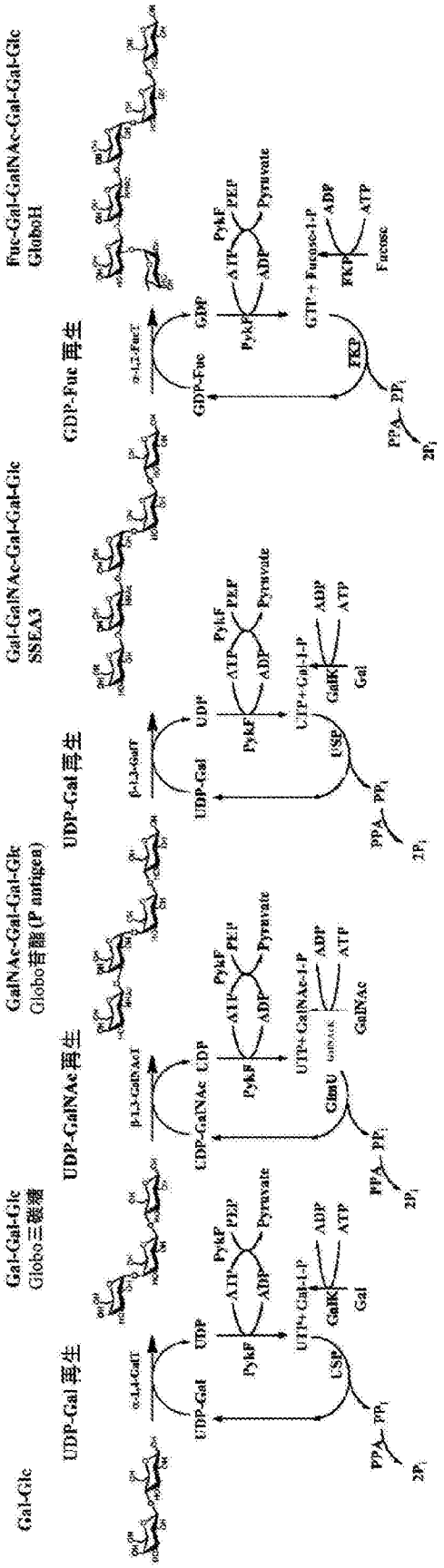

Gal-Glc
Gal-Gal-Glc
Globo三碳糖
GalNAc-Gal-Gal-Glc
Globo苷酯(P antigen)
Gal-GalNAc-Gal-Gal-Glc
SSEA3
Fuc-Gal-GalNAc-Gal-Gal-Glc
GloboH
UDP-Gal 再生
UDP-GalNAc 再生
UDP-Gal 再生
GDP-Fuc 再生
α-1,4-GalT
β-1,3-GalNAcT
β-1,3-GalT
α-1,2-FucT
UDP-Gal
UDP-GalNAc
GDP-Fuc
UDP
GDP
PykF
PEP
Pyruvate
ATP
ADP
UTP+Gal-1-P
GalK
Gal
USP
UTP+GalNAc-1-P
GalNAc
GlmU
PPA
PPi
2Pi
GTP+Fucose-1-P
FKP
Fucose

图4

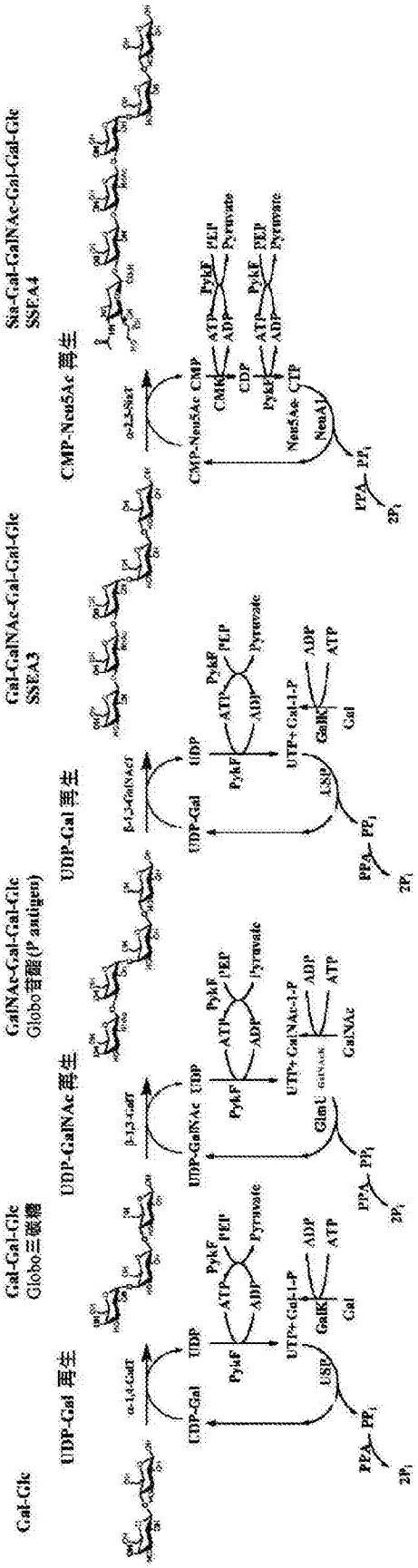

Gal-Glc
UDP-Gal 再生
Gal-Gal-Glc
Globo三糖糖
UDP-GalNAc 再生
GalNAc-Gal-Gal-Glc
Globo四糖(P antigen)
UDP-Gal 再生
Gal-GalNAc-Gal-Gal-Glc
SSEA3
CMP-Neu5Ac 再生
Sia-Gal-GalNAc-Gal-Gal-Glc
SSEA4

图5

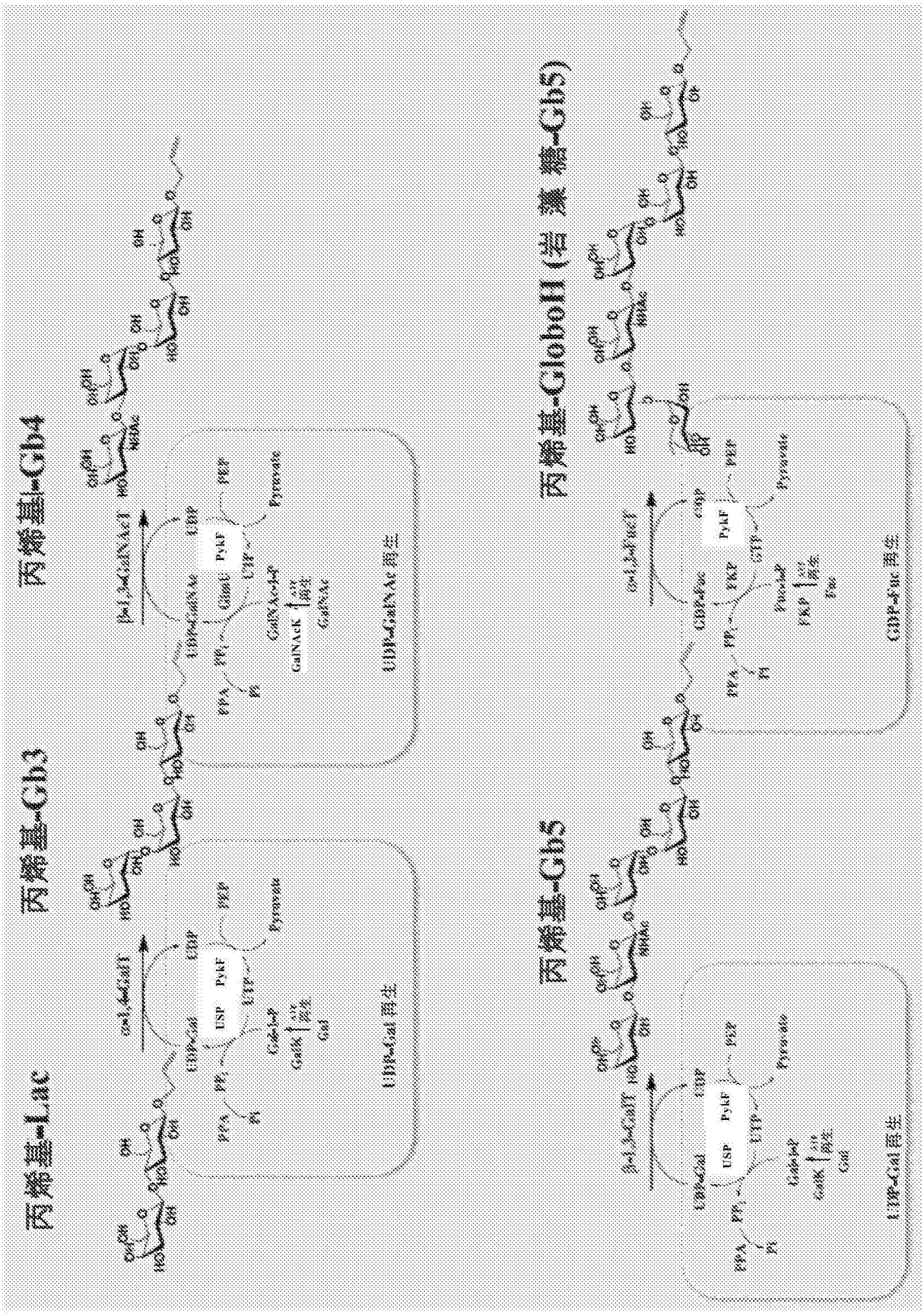

丙烯基-Lac
丙烯基-Gb3
丙烯基-Gb4
丙烯基-Gb5
丙烯基-GloboH (岩 藻 糖-Gb5)
UDP-Gal 再生
UDP-GalNAc 再生
GDP-Fuc 再生
USP
PykF
UTP
PEP
Pyruvate
GalNAcK
GalNAc
GalNAc-1-P
GalK
Gal
Gal-1-P
Fuc
FKP
GTP
Fuc-1-P
PPA
Pi
PP1
UDP-Gal
UDP-GalNAc
GDP-Fuc
UDP-Gal 再生
GDP-Fuc 再生

图6

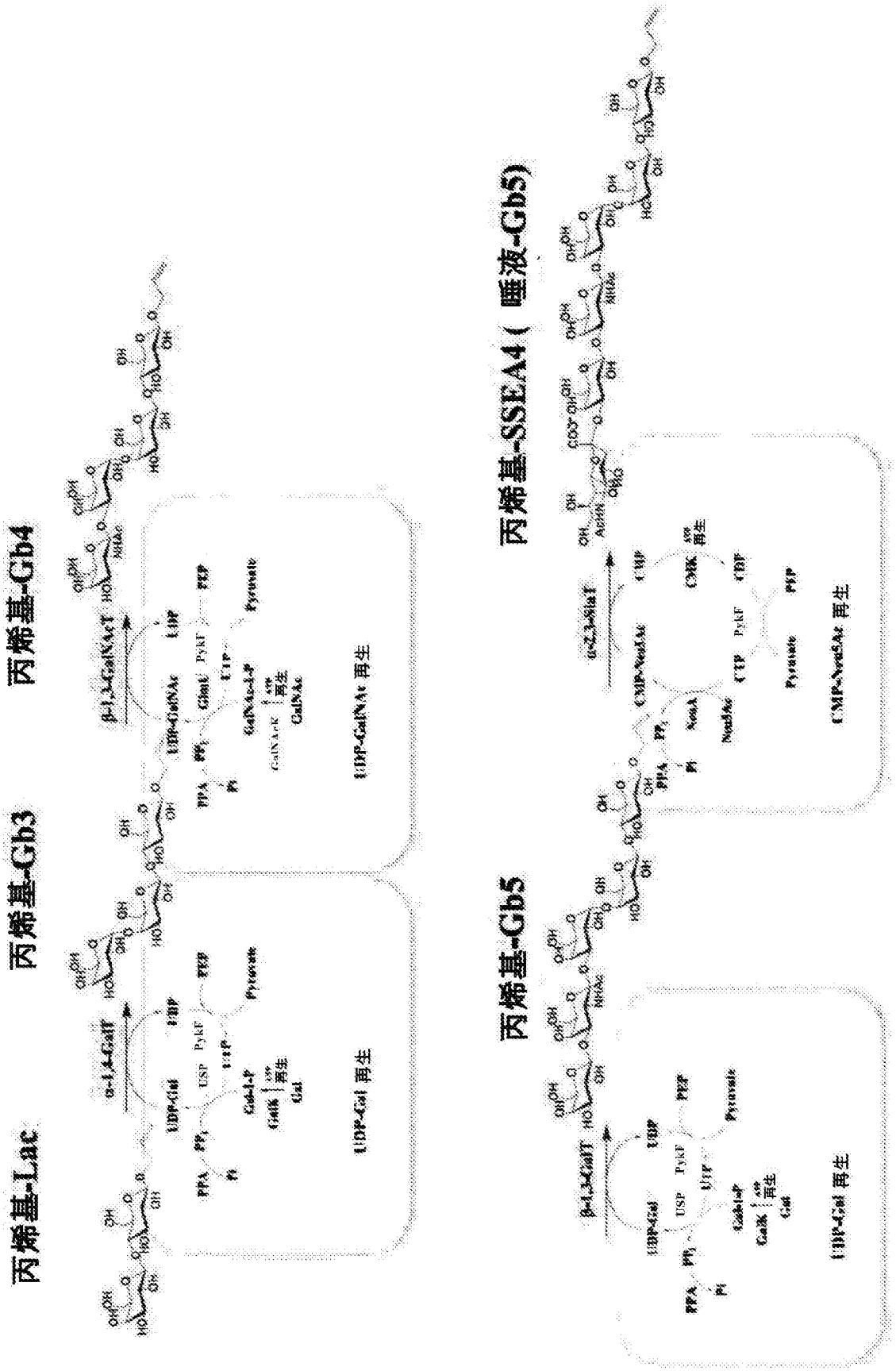

丙烯基-Lac
丙烯基-Gb3
丙烯基-Gb4
丙烯基-Gb5
丙烯基-SSEA4（唾液-Gb5）
UDP-Gal 再生
UDP-GalNAc 再生
CMP-Neu5Ac 再生

图7

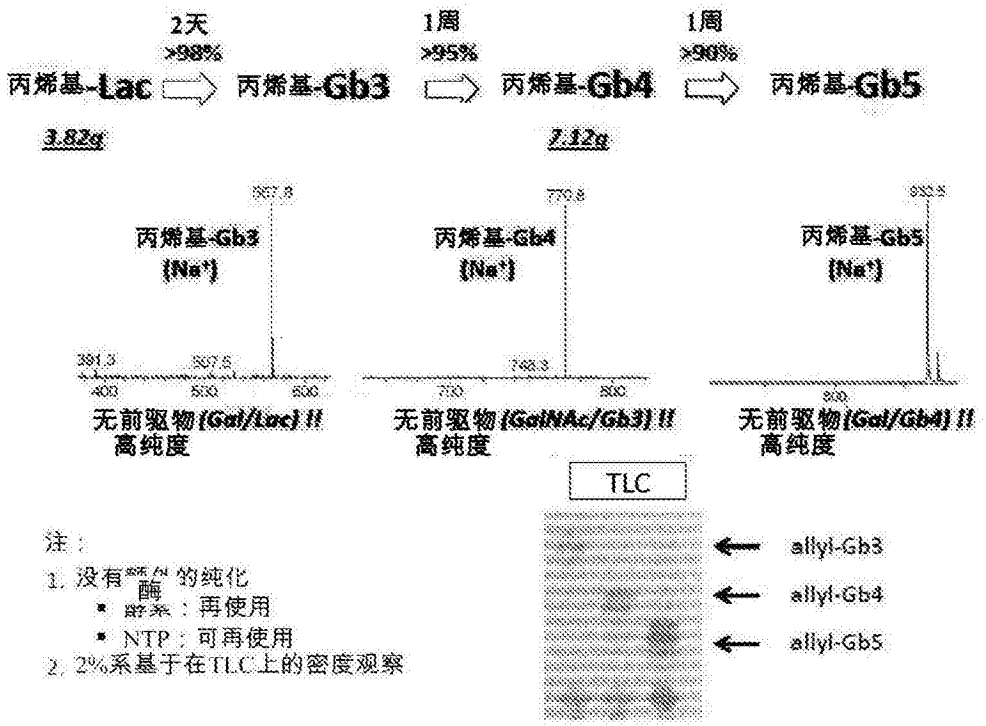

2天
>98%
丙烯基-Lac
丙烯基-Gb3
1周
>95%
丙烯基-Gb4
1周
>90%
丙烯基-Gb5
3.82g
7.12g
丙烯基-Gb3
(Na+)
丙烯基-Gb4
(Na+)
丙烯基-Gb5
(Na+)
无前驱物(Gal/Lac) !!
高纯度
无前驱物(GalNAc/Gb3) !!
高纯度
无前驱物(Gal/Gb4) !!
高纯度
TLC
注：
1. 没有酶的纯化
酶：再使用
NTP：可再使用
2. 2%系基于在TLC上的密度观察
allyl-Gb3
allyl-Gb4
allyl-Gb5

图8

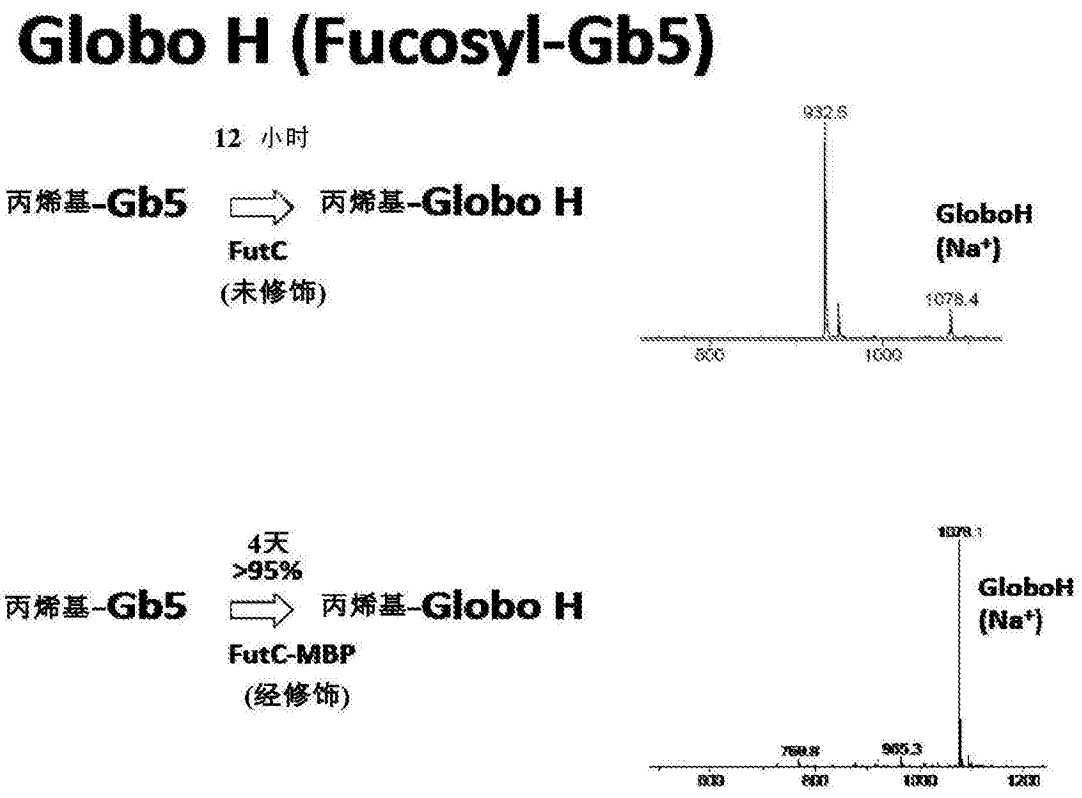

Globo H (Fucosyl-Gb5)
12 小时
丙烯基-Gb5
丙烯基-Globo H
FutC
(未修饰)
GloboH
(Na+)
4天
>95%
丙烯基-Gb5
丙烯基-Globo H
FutC-MBP
(经修饰)
GloboH
(Na+)

图9

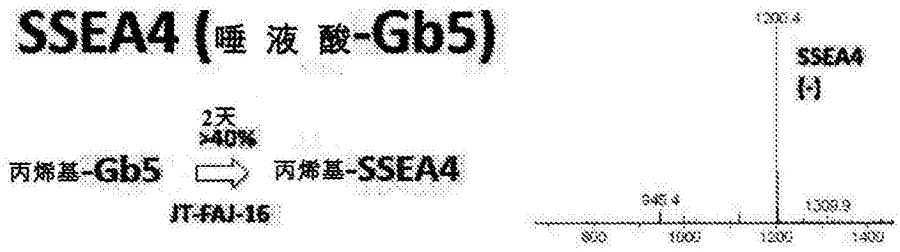


图10