



(12) 发明专利

(10) 授权公告号 CN 105142675 B

(45) 授权公告日 2021.08.20

(21) 申请号 201480018071.2  
 (22) 申请日 2014.03.14  
 (65) 同一申请的已公布的文献号  
 申请公布号 CN 105142675 A  
 (43) 申请公布日 2015.12.09  
 (30) 优先权数据  
 13159484.8 2013.03.15 EP  
 61/787,371 2013.03.15 US  
 61/939,754 2014.02.14 US  
 (85) PCT国际申请进入国家阶段日  
 2015.09.24  
 (86) PCT国际申请的申请数据  
 PCT/EP2014/055173 2014.03.14  
 (87) PCT国际申请的公布数据  
 WO2014/140317 EN 2014.09.18  
 (73) 专利权人 恩比伊治疗股份公司  
 地址 瑞士巴塞尔  
 (72) 发明人 乌尔夫·格劳乌德  
 罗格·伦佐·贝利  
 (74) 专利代理机构 北京坤瑞律师事务所 11494  
 代理人 封新琴

(51) Int.Cl.  
 A61K 47/68 (2017.01)  
 A61K 47/65 (2017.01)  
 A61P 37/02 (2006.01)  
 A61P 25/00 (2006.01)  
 A61P 31/04 (2006.01)  
 A61P 35/00 (2006.01)  
 A61P 25/28 (2006.01)  
 A61P 31/00 (2006.01)  
 C07K 16/40 (2006.01)

(56) 对比文件  
 WO 2012142659 A1, 2012.10.26  
 Younes et al..《Brentuximab Vedotin (SGN-35) for Relapsed CD30-Positive Lymphomas》.《The New England Journal of Medicine》.2010,第363卷(第19期),  
 Madej et al..《Engineering of an Anti-Epidermal Growth Factor Receptor Antibody to Single Chain Format and Labeling by Sortase A-Mediated Protein Ligation》.《Biotechnology and Bioengineering》.2012,第109卷(第6期),

审查员 孙迎琪

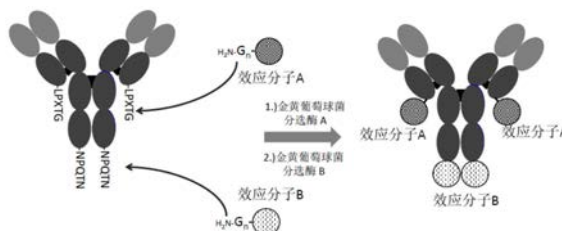
权利要求书1页 说明书50页  
序列表61页 附图17页

(54) 发明名称

通过序列特异性转氨酶制备免疫配体/效应分子结合物的方法

(57) 摘要

本发明涉及用于制备免疫载体/效应分子结合物的方法,该方法包括通过序列特异性转氨酶或其催化域将效应分子结合至免疫配体(图6b)。



1. 曲妥单抗/DM1结合物在制备用于治疗罹患Her-2/neu阳性的乳腺癌或卵巢癌的人对象的药物中的用途,其中曲妥单抗/DM1结合物由包括如下步骤的方法产生:

a) 产生曲妥单抗变体,所述曲妥单抗变体为Tras-HC-GS-LHS或Tras-LC-GS-LHS,其中Tras-HC-GS-LHS具有SEQ ID NO:36所示的未修饰的轻链,以及SEQ ID NO:34所示的C端具有GS连接子,LPETG分选酶基序,6×His标记和strep II标记的重链;Tras-LC-GS-LHS具有SEQ ID NO:32所示的未修饰的重链,以及SEQ ID NO:38所示的C端具有GS连接子,LPETG分选酶基序,6×His标记和strep II标记的轻链;其中,HC表示重链,LC表示轻链,GS表示甘氨酸-丝氨酸连接子,LHS表示LPETG标记+6×His标记+strep II标记;以及

b) 分选酶A介导的将步骤a)中的曲妥单抗变体与Gly<sub>5</sub>-修饰的DM1毒素结合;  
其中分选酶A具有SEQ ID NO:26所示的氨基酸序列。

## 通过序列特异性转肽酶制备免疫配体/效应分子结合物的方法

### 技术领域

[0001] 本发明涉及制备免疫配体/效应分子(payload)结合物的方法。

### 背景技术

[0002] 目前用于标记蛋白和/或将分子结合至蛋白(尤其是当涉及小分子效应分子或标记时)的主要方法,包括在化学结合(conjugation)中使用特异性连接子(linker)分子,将效应分子共价连接至蛋白质的自由赖氨酸和/或半胱氨酸。

[0003] 但是,许多蛋白质,例如免疫靶向策略中尤为感兴趣的抗体,都是较大的蛋白质,可能包含多个赖氨酸和半胱氨酸残基。由于连接子介导的化学结合是一种随机过程,因此,连接子介导的效应分子化学连接反应得到的是结合蛋白的不同种混合物,其在疗效和/或诊断用途上可能不同。显然,蛋白-效应分子结合物还给治疗性结合物的常规审批流程带来了显著困难,因为出于潜在安全顾虑,有关管理部门对其批次差异和/或活性药学成分(API)差异存在负面看法。

[0004] 此外,若期望限定蛋白负载的效应分子比例,则常需要根据期望的结合化学计量配比来纯化结合物。这不仅枯燥,而且可能限制提高生产过程中的商品成本,因为通常仅有部分连接子介导的结合蛋白具有期望的效应分子比例。对治疗相关的抗体/药物结合物(ADCs)而言尤为如此,其中根据所用的毒素,3-4个毒素分子似乎较为有利,但在典型的连接子介导的化学结合反应中,存在每个抗体结合0-8个毒素的情况(Panowski et al. (2014))。

[0005] 虽然存在上述缺陷,目前临床测试中所用的或由卫生部门批准用于疾病治疗的全部抗体/药物结合物,都是毒性小分子药物与抗体通过连接子介导的化学结合获得的(Lambert (2012) 或Mullard (2013))。

[0006] 本领域的工业和科学专家们普遍认同,分子效应分子(包括毒性或标记分子)与免疫配体的位点特异性化学计量配比结合,与连接子介导的化学结合反应相比,具有显著优势。因此,存在以蛋白质结构中特定氨基酸作为化学结合反应目标的尝试(Panowski et al. (2014))。

[0007] 一方面,尝试通过蛋白质结构中某些位置上的突变,来删除多余结合位点和/或提供作为连接子连接反应目标的期望结合位点(例如,赖氨酸和/或半胱氨酸残基)。

[0008] 另一方面,尝试通过在某些位点加入非天然氨基酸,例如硒代半胱氨酸、对叠氮基苯丙氨酸或乙酰苯丙氨酸(Hofer et al. (2009), Axup et al. (2012), 或Lemke (2011)),以控制化学结合反应。

[0009] 但是,所有这些途径都改变了将结合的蛋白质主要氨基酸序列,可能造成不期望的功能性质。此外,上述非天然氨基酸的加入通常效率低下,且不能够对蛋白定量加入特异性标记位点。

## 发明内容

[0010] 因此,存在克服随机结合方法的已知问题的迫切产业需求,尤其是对于治疗相关的免疫结合物的制备,包括但不限于ADCs。

[0011] 因此,本发明的目标之一,在于提供一种联合免疫配体和效应分子(例如药物、毒素、细胞因子、标记等)以制备位点特异结合抗体药物结合物的高效方法,优选为将全长单克隆抗体结合至小分子量毒素。

[0012] 本发明的又一目的在于,制备疗效更佳和/或生产再现性更高的免疫配体/效应分子结合物。

[0013] 本发明的又一目的在于,允许以位点特异性和/或序列特异性方式,将效应分子与免疫配体结合。

[0014] 本发明的又一目的在于,制备保存其组成部分的特征性质的免疫配体/效应分子结合物,例如靶向亲和性、靶向特异性、靶向敏感性、溶解性、药理学功能等。

[0015] 上述目的通过独立权利要求的主题达成,同时,其优选实施方式进一步公开于从属权利要求及说明书中。

[0016] 概念定义

[0017] 本申请所称术语“免疫配体”意在限定一种对指定目标(例如受体、细胞表面蛋白、细胞因子等)具有亲和性的实体、药剂或分子。可选地,所述免疫配体可以阻断或降低激动剂介导的应答,或抑制受体-激动剂相互作用。但是,最重要的是,该免疫配体可用作将效应分子递送到特异位点的穿梭载体,所述特异位点由所述免疫配体所识别的标靶所限定。因此,免疫配体以例如但不限于一种受体为标靶,将其效应分子递送至以含有大量所述受体为特征的位点。免疫配体包括但不限于,抗体、抗体片段、给予抗体的结合蛋白、拟抗体、受体、可溶性诱饵受体、对指定标靶具有亲和性的支架蛋白,以及受体的配体。

[0018] “抗体”,以及其同义词“免疫球蛋白”(Ig)通常包含4条多肽链,其中有2条重链(H)和2条轻链(L),因此是多亚基蛋白或其等同Ig同系物(例如,含有仅一条重链的骆驼科纳米抗体,可源于一条重链或轻链的单域抗体(dAbs));包括全长功能性突变体、变体或其衍生物(包括但不限于,鼠科抗体、嵌合抗体、人源化抗体和全人抗体,其能够保存Ig分子的必要表位结合特征),还包括双特异性、多特异性和双可变域免疫球蛋白,免疫球蛋白分子可以是任何种类(例如,IgG,IgE,IgM,IgD,IgA,和IgY)或子类(例如IgG1,IgG2,IgG3,IgG4,IgA1,和IgA2)和同种异型体。

[0019] 本申请所称的“基于抗体的结合蛋白”可以指其他非免疫球蛋白或非抗体衍生成分的情况下,任何包含至少一个抗体衍生的 $V_H$ , $V_L$ ,或 $C_H$ 免疫球蛋白域的蛋白。所述基于抗体的蛋白包括但不限于,(i)结合蛋白的 $F_c$ -融合蛋白,包括具有部分或全部免疫球蛋白 $C_H$ 域的受体或受体成分了;(ii) $V_H$ 和/或 $V_L$ 域连接替代性分子骨架的结合蛋白;或(iii)含有以通常不存在于天然抗体或抗体片段中的方式结合和/或组装的免疫球蛋白 $V_H$ 和/或 $V_L$ 域分子。

[0020] 本申请所称的“抗体药物结合物”(ADC)指的是与小分子量活性药学成分(API)连接的抗体或抗体片段或基于抗体的结合蛋白,其中API包括但不限于毒素(包括,例如,但不限于,微管蛋白抑制剂、肌动蛋白粘合剂、RNA聚合酶抑制剂、DNA插入和修饰/破坏药物)、激酶抑制剂或任何干扰细胞生存必需的和/或特定生理细胞通路必需的特定细胞通路的API。

[0021] 本申请所称的“抗体衍生物或片段”指的是包含至少一种源于非全长抗体的多肽



链的分子,包括但不限于:(i) Fab片段,其为由轻链可变域( $V_L$ )、重链可变域( $V_H$ )、轻链恒定域( $C_L$ )和重链恒定域1( $C_{H1}$ )组成的单价片段;(ii) F(ab')<sub>2</sub>片段,其为由包含由铰链区二硫键连接的2个Fab片段组成的二价片段;(iii) F<sub>ab</sub>(F<sub>d</sub>)片段的重链部分,由 $V_H$ 和 $C_{H1}$ 域组成;(iv) 可变片段( $F_v$ )片段,由抗体单臂上的 $V_L$ 和 $V_H$ 域组成;(v) 域抗体(dAb)片段,其包含单可变域;(vi) 分离的互补决定区(CDR);(vii) 单链 $F_v$ 片段(sc  $F_v$ );(viii) 双特异抗体,其为由在单一多肽链上表达但使用了其长度不允许同链两域配对的连接子的 $V_H$ 和 $V_L$ 域的双价、双特异性抗体,由此迫使所述域与另一条链的互补区配对形成2个抗原结合部位;以及(ix) 线形抗体,包含一对与互补轻链多肽共同形成一对抗原结合区的串联 $F_v$ 片段( $V_H$ - $C_{H1}$ - $V_H$ - $C_{H1}$ );和(x) 其他免疫球蛋白轻链和/或重链的非全长部分,或其突变体、变体或衍生物,或其组合。

[0022] 本申请所称的“修饰抗体形式”包含抗体-药物-结合物、聚亚烷基氧化物修饰的sc  $F_v$ 、单体抗体、双体抗体、骆驼科抗体、域抗体、双特异性或三特异性抗体、IgA,或由J链和分泌片连接的2个IgG结构、鲨鱼抗体、新世界灵长类骨架+非新世界灵长类CDR、移除铰链区的IgG4抗体、带有2个改造为CH3域的附加结合位点的IgG、Fc区改变以提高对Fc伽马受体亲和性的抗体、包含CH3+VL+VH的二聚结构,等等。

[0023] 本申请所称的“拟抗体”指的是不属于免疫球蛋白家族的抗体,甚至包括非蛋白,例如适体,或合成聚合物。某些种类具有类似抗体的 $\beta$ -片层结构。“拟抗体”或“替代性骨架”相对于抗体的潜在优势在于溶解性更好,组织渗透性更好,对热和酶的稳定性更高,且制备成本相对较低。

[0024] 某些拟抗体可以用大文库提供,其中提供对每一种可能标靶的特异性结合候选物。正如抗体,靶向特异性拟抗体可以使用高通量筛选(HTS)技术以及成熟的显示技术(例如噬菌体展示、西郡展示、酵母菌或哺乳类展示)来开发。目前开发出的拟抗体包括,例如,锚蛋白重复蛋白(称为DARPin),C型凝集素,金黄色葡萄球菌的A结构域蛋白,转铁蛋白,脂质运载蛋白,纤连蛋白第十型III结构域,Kunitz结构域的蛋白酶抑制剂,泛衍生粘合剂(称为affilins), $\gamma$ -晶状体衍生的粘合剂,半胱氨酸结,硫氧还蛋白A基支架粘合剂,核酸适体,由聚合物分子印迹制得的人造抗体,从细菌基因组肽库制得的多肽文库,SH-3域,stradobody,以二硫键和基于Ca<sup>2+</sup>、CTLA4-稳定化的膜受体的“A结构域”,Fyn SH3,以及适体(与特定靶分子结合的寡核苷酸或肽分子)。

[0025] 当免疫配体并非蛋白或肽时,例如,若其为适体时,应当优选地令其连接肽标签,以为本申请公开的酶促结合反应提供适当底物。

[0026] 本申请所称的“结合反应(conjugation)”指的是一个分子与另一个分子通过形成共价键而进行共价结合。

[0027] 本申请所称的“免疫毒素”指的是与代表毒素的蛋白或多肽结合的免疫配体,所述毒素包括,但不限于,细菌毒素(如白喉毒素A,假单胞菌外毒素,肉毒杆菌毒素),或如来自脊椎动物(例如,但不限于蛇),或非脊椎动物(例如但不限于蜘蛛,蝎子,软体动物,水母)的蛋白质毒液,或其功能片段。

[0028] 本申请所称的“低分子量效应分子”指代分子量不超过2500道尔顿的效应分子。

[0029] 本申请所称的“效应分子”指的是任何天然或合成制备的分子,包括可由化学合成的小分子量分子或化学实体,和需要通过宿主细胞发酵来制备且可对靶或抗原结合特异性免疫配体赋予新功能的较大分子或生物实体。

[0030] 本申请所称的“小分子量毒素”指的是分子量不超过2500道尔顿且对哺乳类细胞具有细胞毒性的小分子量细胞毒性化合物。

[0031] 本申请所称的“转肽酶”是一种酶或一种酶或蛋白的催化域,其能够催化多肽键断裂并继而直接或通过若干反应中间体形成新多肽键,由此在反应中保存第一多肽键的能量并将其传递给新的多肽键。优选地,所述转肽酶优选为将一个多肽或蛋白的C端连接至另一种肽或蛋白的N端。由于形成了新肽键,所述酶或功能域也称为“蛋白连接酶”,“肽连接酶”,或昵称为“蛋白或肽装订器”。所述蛋白连接酶包括,但不限于,分选酶、内含肽和分裂-内含肽。

[0032] 本申请所称的“序列特异性转肽酶”意在限定一种转肽酶,其需要至少一个具有指定多肽序列作为识别序列(N端和/或C端)的序列肽或蛋白,来连接所述底物肽或蛋白至另一肽或蛋白,或包含肽或蛋白元件的小分子量化合物。

[0033] 本申请所称的“位点特异性转肽酶”意在限定一种在至少一个底物肽或蛋白中具有特异性位点的转肽酶,其使用该特异位点来结合另一肽或蛋白,或包含肽或蛋白元件的小分子量化合物。

[0034] 发明背景及一般说明

[0035] 本申请公开了使用位点特异性转肽酶(例如分选酶和分裂-内含肽)来位点特异性且选择性地效应分子(优选为小分子量的毒素)结合至免疫配体(优选为抗体),以生成免疫配体效应分子结合物,优选为抗体药物结合物(ADCs)。优选的效应分子为以短链(优选为小于13(十三)个氨基酸长度)合成氨基酸序列修饰过并由此成为分选酶或分裂内含肽介导的免疫配体N端或C端共价结合底物的小分子量毒素(图1和图3)。传统上的效应分子和免疫配体的化学结合反应是如上所述的随机过程,与之相比,此述的结合反应的区别性特征在于,其是以位点特异性和限定化学计量配比的形式达成的。

[0036] 本申请还公开了位点特异性转肽酶(例如分选酶或分裂内含肽)介导的多聚体免疫配体结合反应,优选为,抗体与2种不同的毒性分子或其他标记,使用多聚体蛋白亚基(例如抗体的重链和轻链)的不同修饰,和不同转肽酶特异性的不同短氨基酸链修饰后的效应分子,进行结合,从而将至少两个不同的功能效应分子与多聚体免疫配体结合(图6)。

[0037] 本申请还公开了用于向免疫配体的N端或C端添加亲和性纯化和/或检测标记的方法,其经过酶介导的转肽作用,从而能够借由亲和性纯化和/或检测标记的移除,使用亲和性树脂来截留未完全结合因而仍保留了附加的亲和性纯化和/或检测标记的免疫配体,来选择效应分子与经修饰的结合蛋白完全(100%)结合的免疫配体(图4)。

[0038] 本申请还公开了免疫配体,其中转肽酶催化域直接融合至待结合蛋白的N端或C端,由此转肽活性成为待结合的免疫配体的组成部分,在转肽酶介导的结合反应期间无需再提供附加的可溶分选酶。

[0039] 所有上述实施例对任何效应分子均可进行位点特异性且化学计量配比受控的免疫配体结合反应,所述效应分子包括包括小分子毒素(化学实体)、毒性蛋白或荧光标记,优选为小分子量毒素,所述免疫配体优选地包括抗体;本申请实施例优于使用化学连接子化学方法将效应分子与蛋白结合的标准化学结合反应,因其无法控制结合比例和位点。因此,对于制备抗体药物结合物(ADCs),使用转肽酶(优选为分选酶或分裂内含肽)将毒性效应分子与抗体结合,能够得到更均一的产物,提高对癌症治疗的预期疗效(图12)。

[0040] 使用分选酶或分裂内含肽进行效应分子与免疫配体的酶促结合,允许效应分子与蛋白和免疫配体以位点特异性和化学计量配比进行结合,降低了商品成本,并提供了均一的免疫配体-效应分子结合物,尤其是,转肽酶的选择性允许在粗细胞培养上清液中进行效应分子与免疫配体的结合,且无需像传统连接子介导的化学结合反应那样需要纯化成分。因此,使用序列特异性转肽酶进行效应分子与免疫配体的结合,能够显著降低免疫配体-效应分子商品的成本,尤其是ADC生产。

[0041] 本申请所公开的第一种转肽酶,即分选酶,已在多种革兰氏阳性菌中发现,例如葡萄球菌、链球菌和肺炎球菌,且能催化毒力因子与细胞壁蛋白聚酶相连接,从而改变细菌用于回避被感染宿主免疫应答的表面特征标记(Mazmanian et al. (1999))。革兰氏阳性菌金黄色葡萄球菌(*Staphylococcus aureus*)的分选酶A(sortase A)首先得到表征(Ton-That et al. (1999)),其后进一步成为多种蛋白修饰中的工具(Tsukiji (2009))。分选酶的吸引力在于,两个待结合的分子仅需修饰或表达长度为5个氨基酸的短链肽标记(分选酶标记,对于*Staphylococcus aureus*分选酶A的情况为LPXTG,其中X为20种天然氨基酸中的任一种)和优选长度为3-5个氨基酸的甘氨酸链(Antos et al. (2009a)),而这些都能容易地添加到每个分子上,来达成与蛋白N端或C端的结合。该体系可用于两个蛋白的连接或结合,也可以用于小分子与蛋白的结合。

[0042] 第二种能得到肽键断裂和形成的转肽酶成为内含肽,其最初的发现是作为蛋白内含子,能够通过断开肽键并形成新的肽键,将自身从前体蛋白中移除(剪接)(Xu et al. (1993)) (图2a)。内含肽还能分离为N-内含肽域和C-内含肽域(称为分裂内含肽)并连接到后续催化外显肽域反式剪接的独立蛋白上。分裂内含肽已被用于N-外显肽和C-外显肽部分的共价连接,以及蛋白的纯化和/或环化(Elleuche (2010))。但是,为了将分裂内含肽也用于小分子效应分子的结合,必须使用的分裂内含肽能够在N-内含肽域或C-内含肽域可降为极少氨基酸个数时仍发挥作用,能够容易地通过化学合成添加到任何大小的分子上,类似于分选酶介导的转肽作用所需的优选为3个甘氨酸的短链。随着人造Ssp GyrB S11分裂内含肽的开发,该内含肽中的C-内含肽域仅包含6个氨基酸(Sun et al. (2004)),该条件得以达成,且该分裂内含肽被用于携带生物素的蛋白的C端标记(Volkman et al. (2009)) (图3a)。类似地,从Ssp DnaX分裂内含肽开发出的长度为11个氨基酸的短N-内含肽允许蛋白N-端与任何分子结合,只要该11个氨基酸长度的短链能够通过化学合成添加到所选效应分子上(图3b)。

[0043] 因此,本申请的一个方面为,将含有Ssp GyrB的6氨基酸C-内含肽域或Ssp DnaX的11氨基酸N-内含肽域、长度优选为3-5个甘氨酸的甘氨酸短链或含12个氨基酸GVFVHNSXXXXX(X为任何天然或人造氨基酸)的氨基酸短链,添加到效应分子上,这足以允许转肽酶分别将修饰后的效应分子结合到免疫配体上,所述免疫配体分别优选为包含分选酶识别基序,例如对于使用*Staphylococcus aureus*分选酶A的情况为LPXTG,或对于使用Ssp GyrB分裂内含肽的情况为150aa的N-内含肽域,或对于使用Ssp DnaX分裂内含肽的情况为139aa的C-内含肽域(参见图1和图3)。

[0044] 根据分选酶介导的结合反应的需要将氨基酸短链(例如3或5个甘氨酸残基)添加至小分子量毒素,或根据分裂内含肽介导的结合反应将12个氨基酸的短链添加至小分子量毒素,据发现能够提高某些疏水毒素分子的水溶性(数据未显示),由此该氨基酸-毒素加合

物可以在生理性缓冲液中溶解,保证了最优的分解酶或分裂内含肽的结合反应。这防止了大蛋白分子的结构完整性压力,尤其是对于接触到传统连接子化学和结合相关的有机溶剂和非生理性pH即容易降解的抗体而言。此外,疏水毒性分子与大蛋白尤其是抗体的结合,能够引起某种程度的蛋白聚集。这些也可以通过使用转肽酶加以改善,尤其是分选酶,因为其他亲水氨基酸停留在酶促生成的结合物中,降低了大蛋白或抗体药物结合物聚集的倾向。

[0045] 蛋白-蛋白或蛋白-肽连接(Mao et al. (2004), Parthasarathy et al. (2007) 或W02011/133704A2),甚至包括蛋白质环化(Antos et al. (2009b))的现有技术中均广泛记载了分选酶。分选酶蛋白或肽连接的应用还包括了使用抗体片段(例如带蛋白或肽标记的Fab-和scFv-片段(Möhlmann et al. (2011), Madej et al. (2012), 或US2010/0055761A1和W02012/142659A1))进行的蛋白或肽链接。甚至在两篇已发表的现有技术文件中,使用分选酶将全长抗体连接到了蛋白上(Levary et al. (2011), 例如EGFP,白蛋白,白树毒素被结合到了抗体轻链上),或使用分选酶将全长抗体连接到短肽上(Swee et al. (2013))。但是,尚无现有技术文件展示分选酶介导的小分子量毒素(例如阿里他汀或美登素等)与全长抗体或抗体片段的结合。尤其是,尚无现有技术中公开,如本申请所公开的,小分子量毒素的ADCs的生成能够得到小分子量毒素与IgH或IgL链(药物与抗体比例为2)或与IgH和IgL链(药物与抗体比例为4)均一结合的ADCs。

[0046] 虽然现有技术还公开了,以甘氨酸残基修饰非蛋白底物,令其可用于简单的单一亚基蛋白或肽的分选酶修饰(分别为Tsukiji (2009) 或W02007/108013A3),但对于更具挑战性的非蛋白底物(优选为小分子量毒素)与多聚体蛋白(优选为抗体)的均一结合,则此前尚无记载,虽然事实上,分选酶介导的蛋白或肽连接反应已存在于现有技术中多年。

[0047] 此外,本申请公开了多聚体蛋白的结合,尤其是全长单克隆抗体与2种不同的效应分子(优选为两种不同的小分子量毒素)的结合,此前也未记载于现有技术中,虽然分选酶介导的蛋白或肽连接反应已存在于现有技术中多年(Panowski et al. (2014))。

[0048] 现有技术中已知,分选酶可以接受含有最少3个甘氨酸的底物(Parthasarathy et al. (2007)),因此本申请可以包含含有至少三个(3个)添加到感兴趣的效应分子上的甘氨酸残基的效应分子,即使1个或2个甘氨酸残基就已经足够,且包含在本申请所公开的方法之中。对于小分子量效应分子而言,添加少数甘氨酸分子可以通过传统的合成肽化学来完成,如本申请所述。对于蛋白,要添加甘氨酸残基,可以通过添加数个位于蛋白开放阅读框内的甘氨酸残基(优选为至少3个)的密码子,或通过传统合成肽化学使得重组蛋白包含至少3个N端甘氨酸残基。

[0049] 由文献可知,不同的分选酶,例如Staphylococcus aureus的分选酶B,或来自其他革兰氏阳性菌的分选酶,可识别不同的五肽基序,这不同于Staphylococcus aureus的分选酶A识别基序LPXTG(X=任何氨基酸)(Spirig et al. (2011))。因此,本申请还包括将不同于Staphylococcus aureus的分选酶A识别基序LPXTG的其他分选酶识别基序添加到蛋白和免疫配体(包括优选为抗体)中的概念,以令其可用于与不同革兰氏阳性菌的不同同源分选酶进行分选酶结合。因此,蛋白和免疫配体,优选为抗体,还可以用不同的分选酶识别基序进行表达,例如来自Staphylococcus aureus的分选酶B的特异性五肽基序NPQTN,其可以与甘氨酸基序的效应分子相结合。

[0050] 在本申请的又一方面中,多聚体免疫配体,优选为但不限于由免疫球蛋白重链和

轻链组成的抗体,允许使用添加到所述多聚体蛋白不同肽上的不同分选酶识别序列(在抗体将不同分选酶识别序列添加到抗体重链和轻链的情况下),以能够在各自分选酶的存在下,通过执行与Gly<sub>n</sub>-标记的效应分子(n>2)的连续结合,将不同效应分子结合到所述不同多肽上(图6b)。为此,需要令抗体在重链和轻链上表达不同的C端修饰,以包含用于表达不同分选酶的不同分选酶识别基序。所述抗体随后可以与两种包含如上所述的甘氨酸修饰的不同效应分子结合。

[0051] 该形式的优点在于,ADCs特异性地载有两种不同的毒素,优选为能干扰不同细胞通路的毒素,如此在癌细胞杀灭中能够更为有效,因为被靶定的癌细胞较难同时回避ADCs中所含的两种毒素的攻击。

[0052] 对于本领域技术人员而言,显然,分选酶五肽识别基序,例如Staphylococcus aureus分选酶A LPXTG基序,可以选择性地添加到多聚体免疫配体的个体多肽上,以提供期望的结合位点。例如,对于抗体而言,这允许生成修饰抗体,要么仅含有添加到重链的分选酶识别基序(得到每个抗体结合2个效应分子),要么仅含有添加到轻链的分选酶识别基序(得到每个抗体结合2个效应分子),要么含有添加到重链和轻链的分选酶识别基序(得到每个抗体结合4个效应分子)。这些设计变型允许效应分子通过分选酶与抗体的特异性结合,要么仅添加到重链(生成药物与抗体比例为2的ADCs,即DAR2),或仅添加到轻链(生成药物与抗体比例为2个ADCs,即DAR2),或同时添加到重链和轻链(生成药物与抗体比例为4个ADCs,即DAR4)。由此一来,可以有控制地改变抗体的结合位点和化学计量配比,要么生成每抗体重链或轻链2个效应分子的结合,或通过将效应分子同时添加到重链和轻链,生成每抗体4个效应分子的结合。

[0053] 以上描述的是使用多聚体蛋白或免疫配体中的不同分选酶识别基序和分选酶,得到结合位点和化学计量配比的变型,类似地,本申请的又一方面在于,综合分选酶介导和分裂内含肽介导的结合反应,将不同效应分子与多聚体蛋白的不同多肽链结合。这一概念允许在单一步骤中,同时将不同效应分子与多聚体蛋白和免疫配体的不同多肽链结合,因为使用了不同转肽酶和底物(图6a)。

[0054] 应当理解,2种不同的效应分子与多聚体蛋白(优选为抗体,由二硫键连接的两条重链和轻链组成)的结合,可以通过结合分选酶介导的结合反应和分裂内含肽介导的结合反应来达成(如图6a所示),但是,也可能通过使用2种识别不同分选酶识别基序的不同分选酶(例如来自Staphylococcus aureus的分选酶A和分选酶B,如上所述),将两个不同的效应分子与多聚体蛋白(优选为抗体)结合。但是,这也可能包括来自其他分选酶类别(例如分选酶C、D、E、F)或来自其他细菌种类、具有不同分选酶基序特异性的分选酶。

[0055] 分选酶介导的效应分子与蛋白和免疫配体的结合,可以通过提供分选酶识别基序标记的蛋白和至少以3甘氨酸标记的效应分子,并添加酶促活性的分选酶或其功能性片段作为可溶性酶。本申请的又一方面汇总,分选酶的酶活性域还可以提供为与蛋白的N端或C端融合的域。此变型中,优选地,但非必须地,可以将N端分选酶域添加到N端分选酶识别基序,或将C端分选酶域添加到C端分选酶识别基序。两种可能性均保证了,与甘氨酸标记的效应分子反应后,分选酶域在反应期间从蛋白上移除(图5)。

[0056] 将分选酶介导的效应分子与蛋白的结合的该变型,在概念上类似于分裂内含肽介导的效应分子结合,其中分裂内含肽的酶促活性N-内含肽域连接待结合的蛋白,以限定该

蛋白中的结合位点。

[0057] 文献中已报告了大量不同底物特异性的各种分选酶转肽酶(Spirig et al. (2011)),类似地,从所谓的InBase数据库中能得到大量且数量还在增加的不同物种的分裂内含肽,和具有转肽所需的不同N-内含肽和C-内含肽序列的蛋白。因此,分裂内含肽介导的效应分子与免疫配体的结合示例公开了优选的Ssp GyrB S11分裂内含肽(Volkman et al. (2009)),因为C-内含肽域能够被降为线形的6氨基酸短链,但是,分裂内含肽介导的效应分子与蛋白和免疫配体的结合还可以通过InBase数据库的其他分裂内含肽达成,只要N-内含肽域或C-内含肽域足够短(优选为小于13个氨基酸),易于允许多肽合成并添加任何选定效应分子即可。但是,对本领域技术人员,显然,对于蛋白效应分子,可以通过对期望蛋白效应分子的ORF遗传融合,将任何大小的C-内含肽域融合到蛋白效应分子,由此使用小N-内含肽域或小C-内含肽域(<13个氨基酸)的分裂内含肽不存在理论优势。

[0058] 但是,如果要合成小分子效应分子与蛋白和免疫配体结合,则在优选的Ssp GyrB S11分裂内含肽的C-内含肽的情况下,或在Ssp DnaX的N-内含肽的情况下,本申请公开的小于13个氨基酸的小N-内含肽域或小C-内含肽域具有优势,因为所述短肽能够通过标准合成化学合成添加到任何合成小分子量效应分子上。

[0059] 分选酶介导和分裂内含肽介导的效应分子结合,能够在蛋白和免疫配体的N端或C端进行。这仅取决于分选酶基序/甘氨酸短链和N-内含肽域/C-内含肽域在蛋白和免疫配体的位置形式(图1)。

[0060] 对于抗体(作为优选的免疫配体),优选地将效应分子结合到抗体的C端,因为这令效应分子的位置离抗体的抗原结合位点最远。但是,这一优选形式不应看做对本申请的限定,对于其他免疫配体分子,例如拟抗体,其功能性结合域不位于分子的N端,则可以优选为将效应分子结合至其N端。

[0061] 本申请的又一方面中,通过向感兴趣的蛋白或免疫配体的分选酶识别基序或分裂内含肽域的远端添加亲和性纯化或检测标记(例如但不限于小肽标记(如组氨酸标记、strep-标记、MYC-标记或HA-标记)或较大蛋白亲和性纯化标记(如麦芽糖结合蛋白(MBP)标记、谷胱甘肽-S-转移酶(GST)标记,或几丁质结合标记)),提高分选酶和分裂内含肽介导的效应分子与蛋白和免疫配体的结合效率。

[0062] 在本申请的此方面中,转肽反应中,将亲和纯化标记从待结合的免疫配体上移除。这可以用于富集完全结合效应分子的免疫配体,因为仍包含亲和纯化标记的未反应蛋白和免疫配体可以通过与适当的亲和树脂结合来移除,而与效应分子完全结合的蛋白和免疫配体则不再含有亲和纯化标记,因此能够从未反应的免疫配体底物中特异分离。本申请的这一方面对于需要结合数种效应分子的多聚体蛋白和免疫配体(例如优选的抗体)尤为有效。使用位于分选酶或内含肽转肽酶结合位点远端的亲和纯化标记,保证移除因未完全效应分子结合反应仍含有亲和纯化标记的蛋白和免疫配体。

[0063] 与化学结合相比,这对获得均一的免疫配体/效应分子结合物,优选为小分子量毒素特异结合至抗体重链和/或轻链的C端的ADCs,具有显著优势。

[0064] 综上所述,本申请公开的方法提供了一种新颖高效的方法来位点特异性且按照化学计量配比地进行效应分子结合,优选为小分子量毒素与免疫配体(优选为抗体)结合,由此生成限定的免疫抗体/效应分子结合物,优选为ADCs,用于疾病治疗,优选为癌症治疗。该

方法还可以用于生成用于疾病诊断(优选为肿瘤诊断)的免疫抗体/效应分子结合物。该新颖方法能够通过利用肽键断裂和形成酶(转肽酶),包括分选酶和分裂内含肽或其催化活性片段,生成共价免疫抗体/效应分子结合物。所述酶能够催化含有氨基酸短链(优选为小于13个氨基酸)的效应分子与免疫配体的N端或C端的位点特异性共价结合,其中免疫配体经过适当修饰,允许分选酶和分裂内含肽在反应中断开和形成肽键。免疫配体优选为抗体,与小分子量毒素进行位点特异性结合,以生成具有限定的抗体效应分子或限定的药物抗体比例的抗体药物结合物(ADCs)。

### 附图说明

[0065] 图1:本图展示了分选酶A介导的效应分子与免疫配体(或结合蛋白)的位点特异性结合反应的原理,该反应可以在蛋白的N端(a)或C端(b)。为了达成N端结合反应,该效应分子需要包含分选酶五肽识别基序(此处为LPXTG,来自*Staphylococcus aureus*的分选酶A的识别基序,X代表20种天然氨基酸中的任一种),而待标记的免疫配体/结合蛋白的N端需表达具有自由N端氨基基团(此处表示为较小的 $H_2N$ -符号)的至少3个甘氨酸残基的N端延长,在此表示为 $G_n$  ( $n>2$ )。通常使用3-5个甘氨酸来修饰底物以进行分选酶介导的结合反应。添加如此所述的*Staphylococcus aureus*的重组分选酶A,然后催化LPXTG五肽基序中的T残基和C端G残基之间的肽键断开,并在 $G_n$  ( $n>2$ )短链的N端甘氨酸和T残基之间形成新肽键。转肽反应中,LPXTG基序中的C端G残基(粗体强调显示)被移除。(b)相反,为了达成效应分子与蛋白质的C端结合,这也是效应分子(尤其是毒素)与抗体结合的优选方法(参见图6),LPXTG分选酶识别五肽基序需要添加到免疫抗体/结合蛋白的C端(例如,通过实施例中所描述的重组蛋白表达技术),且效应分子需要以甘氨酸短链( $G_n$ ,  $n>2$ ,通常为3-5个甘氨酸)进行修饰。如(a)所述,添加来自*Staphylococcus aureus*的分选酶A能够催化 $G_n$ -短链对LPXTG基序的转肽作用,从而移除LPXTG基序(粗体)的末端G残基。

[0066] 图2:本图展示了内含肽(a)和分裂内含肽(b)介导的转肽作用的原理。(a)内含肽可以作为“蛋白内含子”存在于前体蛋白中,它们将成熟蛋白的N端和C端部分(通常称为N-外显肽和C-外显肽)分隔开。内含肽“蛋白内含子”能够在转肽反应中,催化内含肽和C-外显肽之间的肽键断裂,以及通过将C-外显肽的N端氨基酸转移到N-外显肽的C端氨基酸,催化N-外显肽和C-外显肽之间新肽键的形成。该反应的结果是从前体蛋白移除了内含肽“蛋白内含子”,并生成具有N-外显肽域和C-外显肽域之间的新肽键的成熟蛋白。(b)内含肽活性据记载还能分离为不同的域,可连接到不同蛋白,此种内含肽变型称为分裂内含肽。分裂内含肽的N-内含肽域和C-内含肽域在所连接的N-外显肽域和C-外显肽域上形成非共价结构复合体,该复合体可作为连续的内含肽进行相同的转肽反应,所述N-和C-外显肽域空间相邻,且属于复合体的一部分。N-内含肽和C-内含肽分裂内含肽反应的结果是“蛋白反式剪接”,本质上亦即N-外显肽域和C-外显肽域之间形成新肽键的蛋白连接反应。

[0067] 图3:本图展示了以极短C-内含肽域或极短N-内含肽域为特征的分裂内含肽如何用于将任何效应分子与免疫配体(或结合蛋白)结合,包括小分子实体,因为氨基酸短链可以通过化学合成,且易于通过共价化学连接与小分子实体结合。(a)本图展示了使用Ssp GyrB S11分裂内含肽(如Appleby et al. (2009)所述)进行效应分子与免疫配体/结合蛋白的C端结合。在此,该C-内含肽域仅为6个氨基酸长度,且包含如上所述的GVFVHN氨基酸序



列。但是,然而,由于需要一些等同于C-外显肽结构域的肽,还需加入另外的氨基酸,其中首先需要是丝氨酸或半胱氨酸残基,其余氨基酸可以选择。这表明在 $SX_n$ 符号中,其表示的是前有丝氨酸且后有n个氨基酸( $n>2$ ,优选为5,该氨基酸可以是20种天然氨基酸中的任一种(因此表示为X))的N端氨基酸短链。因此,如实施例所述,含有长6个氨基酸的迷你C-内含肽结构域的长12个氨基酸的短链和长6个氨基酸的C-外显肽氨基酸链足以允许N-内含肽/C-内含肽复合体催化GVFVHN- $SX_n$ (X是任意氨基酸, $n>2$ ,优选为5)中的天冬氨酸-丝氨酸肽键经转肽变为N-外显肽和N-内含肽之间的肽键。这将得到C端结合的免疫配体/结合蛋白,其中效应分子通过C-外显肽氨基酸短链链接。(b)本图展示了Ssp DnaX分裂内含肽的使用(如Song et al. (2012)所述),该分裂内含肽可以分裂为极短的长11个氨基酸的N-内含肽结构域和长139aa的C-内含肽域,用于效应分子与免疫配体/结合蛋白的N端结合。如上所述,这仅需要合成和长11个氨基酸的端N-内含肽域与任何效应分子的合成和链接(或通过重组蛋白技术添加),其然后允许效应分子与任何具有融合了长139aa的Ssp DnaX C-内含肽域的N端的免疫配体或蛋白质的N端特异结合。该反应得到N端结合的免疫配体/结合蛋白。因此,类似于分选酶转肽作用的情况,其中N-或C-端结合仅依赖于LPXTG和 $G_n$ 的肽基序相对于蛋白质和效应分子的设置情况,分裂内含肽也可以分别通过利用迷你短C-内含肽或迷你短N-内含肽结构域,例如Ssp GyrB和SSP DNAX的分裂内含肽,介导蛋白与短肽修饰的效应分子之间的位点特异性N端和C端结合。

[0068] 图4: (a)本图展示了在效应分子和免疫配体的结合反应中,除分选酶标记外,增加附加的亲纯化和/或检测标记的效用。(a)本图展示了,如何在金黄色葡萄球菌分选酶A转肽酶介导的C-端效应分子结合反应中,去除如实施例中附加添加的作为6xHis纯化标记(HHHHHH)、Myc蛋白检测用标记(EQKLISEEDL)和strepII亲和纯化标记(WSHPQFEK)的氨基酸。当使用Ni-NTA亲和树脂(对于6xHis标记)或strep-tactin亲和树脂(对于链球菌II-标记)来从结合产物中分离非结合物底物时,这允许对结合产物进行选择。这样的标记组合仅作为举例说明。

[0069] (b)本图说明的是,使用亲和纯化标记选择/纯化多聚体蛋白(如此述抗体)完全结合产物是特别有用的。实施例中还提供了,抗体可以通过重链和轻链的特异性结合位点来修饰,并且如果所述修饰靶向IgH和IgL链的C端,那么最多四个效应分子可被结合至抗体。其他亲和纯化标记(a)的加入,如图4(a)所示,使仅有1、2或3个(如此所述)效应分子结合到抗体形成的不完全结合产物结合到各自的亲和纯化树脂,由此可以容易地将其从完全效应分子结合产物中分离。这种方法当然也适用于内含肽修饰的免疫配体,而不仅限于此述的分选酶基序修饰的免疫配体。

[0070] 图5:该图展示了可以应用分选酶介导的结合反应的变形,其中,分选酶不作为单独的重组蛋白添加至分选酶标记的免疫配体和甘氨酸短链修饰的效应分子,而是令酶活性分选酶结构域被表达为LPXTG分选酶标记C端的融合蛋白。分选酶域无活性,除非将其与甘氨酸短链修饰的效应分子(或底物)共培养。只要甘氨酸短链的修饰底物(或效应分子)被添加到所述结构中,则所述融合分选酶域将通过剪切LPXTG标记的苏氨酸-4和甘氨酸-5位置之间的蛋白,并由此除去上述可选择添加的具有附加亲和纯化标记分选酶结构域,催化甘氨酸-效应分子底物到LPXTG分选酶标记的转肽反应。这个过程的优势在于,类似于具有催化活性的分裂内含肽结构域的添加,可通过重组蛋白技术,将分选酶结构域表达为待结合



免疫配体的一个组成部分。

[0071] 图6: (a) 本图展示了使用不同转肽酶(在此为分选酶和分裂内含肽)来同时将不同的效应分子结合到多聚体蛋白的不同亚单位,例如此处所述的抗体重链和轻链。在所选实施例中,以Ssp GyrB的N-内含肽域(例如由实施例2提供)修饰重链的C端,而用分选酶A五肽基序LPXTG(例如由实施例2提供,为简单起见省略附加标记)来修饰轻链。与甘氨酸短链修饰的效应分子A和与C-内含肽域修饰的效应分子B和分选酶共同培养,将允许效应分子B与重链和效应分子A与轻链的同时选择性结合。如果效应分子A和B是靶向不同细胞通路的毒素,则这种策略可以产生更有效的抗癌药物,而常规的ADCs只包含单个毒素部分。(b) 本图展示了使用不同的分选酶(在此为来自金黄色葡萄球菌的分选酶A和分选酶B)中,以同时将不同的效应分子结合到多聚体蛋白的不同亚单位,例如此处所述的抗体重链和轻链。在所选实施例中,以分选酶B的五肽识别基序LPXTG修饰重链的C端,而用分选酶A的五肽识别基序NPQTN修饰轻链。甘氨酸短链修饰的效应分子A和B与分选酶A和分选酶B的顺序结合将允许效应分子B与重链和效应分子A与轻链的同时选择性结合。(为简单起见,在结合结构中省略了来自LPXTG和NPQTN的其余肽序列)。如果效应分子A和B是靶向不同细胞通路的毒素,则这种策略可以产生更有效的抗癌药物,而常规的ADCs只包含单个毒素部分。

[0072] 图7:金黄色葡萄球菌重组酶活性分选酶A片段的(a) SDS-PAGE和(b) 免疫印迹分析。(a) SDS-PAGE的第1道中包含BSA (ca. 66.4kD),第M<sub>1</sub>道包含Genscript蛋白分子量标准(Cat.-NR:MO0505),第2道包含组氨酸标记纯化的金黄色葡萄球菌重组分选酶片段。SDS-PAGE中的蛋白以考马斯亮蓝染色。(b) 用抗组氨酸抗体(Genscript Cat.-Nr.:A00186)进行免疫印迹分析。第3道中包含组氨酸标记纯化的金黄色葡萄球菌重组分选酶A片段。第M<sub>2</sub>道包含Genscript蛋白分子量标准(Cat.-Nr:MM0908)。

[0073] 图8:DM1毒素结合的Tras-HC-GS-LHS (A) 和Tras-LC-GS-LHS (B) 的疏水作用色谱(HIC)分析。DAR1表示药物与抗体的比例为1;DAR2表示药物与抗体的比例为2。

[0074] 图9:所示ADC对过量表达HER2的SKBR3 (A) 和HER2阴性的T47D-5R $\alpha$ 细胞 (B) 的剂量依赖性细胞毒性作用。细胞与ADC连续稀释物共培养3天,之后使用CellTiter-Glo®发光溶液(Promega)检测细胞活力。LC:结合DM1-分选酶A的Tras-LC-GS-LHS;HC:结合DM1-分选酶A的Tras-HC-GS-LHS。

[0075] 图10:分选酶A介导的Gly<sub>5</sub>-FITC与带或不带GS间隔肽的mAb AC10变体的结合。在相同的其它条件下,使用分选酶A的系列稀释液将Gly<sub>5</sub>-FITC与mAb AC10-HC-GS-LHS/LC-GS-LHS (A) 和mAb AC10-HC-LS/LC-LS (B) 结合。反应产物在变性还原性SDS-PAGE凝胶上按大小进行分离。通过将凝胶置于UV箱中来令FITC可视化。所用的分选酶A浓度为:第1、9道:50 $\mu$ M;第2、10道:25 $\mu$ M;第3、11道:12.5 $\mu$ M;第4、12道:6.25 $\mu$ M;第5、13道:3.13 $\mu$ M;第6、14道:1.56 $\mu$ M;第7、15道:0.78 $\mu$ M;第8、16道:0.39 $\mu$ M。

[0076] 图11:间隔肽长度对轻链结合效率的影响。在相同的其它条件下,使用分选酶A的连续稀释液来令Gly<sub>5</sub>-FITC与mAbs AC10-HC-LS/LC-GS-LS (A,左),AC10-HC-LS/LC-GGS-LS (A,右),Ac10-HC-LS/LC-GGG-LS (B,左) 和AC10-HC-LS/LC-GGGGS-LS (B,右) 结合。反应产物在变性还原性SDS-PAGE凝胶上按大小进行分离。通过将凝胶置于UV箱中来令FITC可视化。所用的分选酶A浓度为:第1、8、15、22道:25 $\mu$ M;第2、9、16、23道:12.5 $\mu$ M;第3、10、17、24道:6.25 $\mu$ M;第4、11、18、25道:3.13 $\mu$ M;第5、12、19、26道:1.56 $\mu$ M;第6、13、20、27道:0.78 $\mu$ M;第7、

14、21、28道:0.39 $\mu$ M。

[0077] 图12:分选酶A介导的mAb AC10与vc-PAB-MMAE毒素的重链结合ADC的疏水作用色谱(HIC)分析,该分析能够区分未反应的底物(药物抗体比例DAR0=0),两个重链之一被结合(药物抗体比例DAR1=1)的底物,和两个经修饰的重链均被结合(药物抗体比例DAR2=2),如图所示。图A显示了标准的分选酶A介导的HC修饰AC10mAb的结合反应后的HIC分布,其中可检测仍为DAR0和DAR1的种类,旁边是期望的DAR2产物。图B显示了图A中的ADC制剂经过4轮StrepTactin®亲和纯化柱分析后的HIC分布。

[0078] 图13:对以合成Gly<sub>5</sub>修饰的 $\alpha$ -鹅膏蕈碱毒素(a)和以Gly<sub>5</sub>修饰的美登素毒素(b)的分析。图a和b中,所述合成结构显示于上方,所述五个甘氨酸由框高亮显示。下方提供了各化合物的质谱法和反相HPLC分析结果。a) Gly<sub>5</sub>修饰的 $\alpha$ -鹅膏蕈碱毒素的预期分子量为1302.07道尔顿,所观察到的分子量为1325.38道尔顿,对应于Ms+Na<sup>+</sup>。RP-HPLC分布表明纯度>95%。b) Gly<sub>5</sub>修饰的美登素毒素的预期分子量为991.41道尔顿,所观察到的分子量为957.69道尔顿,对应于Ms+Na<sup>+</sup>。RP-HPLC分布表明纯度>95%。

[0079] 图14:已由Concortis公司(San Diego,CA,U.S.)合成(结构1-6、9)的或可以合成(结构7&8)的5xGlycine(Gly<sub>5</sub>)修饰的毒素的结构,表明任何毒素均可被功能化并用于分选酶介导的酶促结合,只要毒素连接5个甘氨酸(如此处所示),或大于或等于一个甘氨酸的任何数目甘氨酸残基。甘氨酸修饰的毒素既可以合成为含有附加的经验证的连接子/间隔子结构,如图14a中的结构1-3所示,从而可能加入某些附加功能(例如,在某些亚细胞区室中的剪切能力),也可以合成为不具有附加的连接子,如图14b中的结构4-6所示。如果指定毒素处有几个可用的反应性基团,例如,在 $\alpha$ -鹅膏蕈碱毒素的情况下,甘氨酸残基可被添加到所述不同基团,如图14c中的结构7-9所示。

[0080] 图15:实施例12中测定的肿瘤体积。所述结果表明,采用相同的抗体和毒素部分的分选酶结合ADC,具有与市售的化学结合的Kadcyla®相仿的肿瘤杀伤活性。

[0081] 图16:实施例13中所述的用考马斯兰染色的凝胶。该数据显示了意想不到的发现,即,由分选酶A介导的在粗细胞培养物上清液中的抗体结合反应与使用纯化抗体具有同样的效率。

## 具体实施方式

[0082] 本申请公开了一种用于制备免疫配体/效应分子结合物的方法,包括将效应分子通过序列-特异性转肽酶或其催化域与免疫配体结合。

[0083] 在本申请的一个优选实施例中,所述效应分子和/或免疫配体

[0084] a) 完全由蛋白或肽组成

[0085] b) 包含至少一种蛋白或肽结构域,或

[0086] c) 包含至少一种肽链,

[0087] 且进一步地,蛋白或肽或域优选地包含能够由序列-特异性转肽酶或其催化域检测到的氨基酸序列。

[0088] 这意味着,例如,在效应分子和/或免疫配体是蛋白的情况下,这意味着所述蛋白的N-或C-端包含能被序列-特异性转肽酶检测的氨基酸序列。若原蛋白缺乏所述氨基酸序列,可以通过本领域已知重组方法将其融合到所述蛋白的N-或C-端。

[0089] 在效应分子和/或免疫配体不是蛋白的情况下,所述能被序列-特异性转肽酶检测的氨基酸序列可以通过本领域已知的传统化学交联法与前者结合。

[0090] 可以在特异性转肽酶识别序列和效应分子之间插入其他功能。这可以通过内吞入细胞后分类为可断键(例如包含脎,或二硫化物化学,或胞内蛋白酶特异性肽序列)或不可断键(例如,包含硫醚化学)的化学结构来实现。

[0091] 包含脎化学的化学结构可以在胞内结构溶酶体中选择性地断键(pH低于全身血液循环)。

[0092] 肽连接子具有被溶酶体蛋白酶(例如组织蛋白酶-B)选择性断裂的潜力,且与脎连接子对比,其已经展现出增加的血清稳定性和提高的抗肿瘤效果。缬氨酸-脯氨酸(Val-Cit)对是最常用的肽连接子,并理想地适合用于阿里他汀类药物,例如单甲基阿里他汀E(MMAE)。

[0093] 不可断键连接肽连接子长期受到忽视,因为研究者们相信断开连接子是释放药物的最合理方法。但是,结合物可以在与膜受体结合时快速内吞,且一旦内吞,免疫配体可以分解到可令效应分子,例如药物,暴露出来的程度。举一个突出的例子,硫醚连接子,使用SMCC(N-琥珀酰亚胺基-4-(N-马来酰亚胺甲基)-环己烷-1-羧酸酯)连接子(图)。

[0094] 所有这些方法的共同之处在于不存在真正的连接反应位点特异性。因为连接子介导的化学结合时随机过程,连接子介导的效应分子化学连接会得到结合蛋白的不均一混合物,其在疗效和/或诊断潜力上均可能不同。显然,蛋白-效应分子结合物的混合物也对治疗结合物的监管审批过程带来了显著挑战,因为批次间差异和/或活性药学成分差异(API)由于潜在安全隐患而导致了监管部门的负面看法。

[0095] 不可断键连接肽连接子长期受到忽视,因为研究者们相信断开连接子是释放药物的最合理方法。但是,结合物可以在与膜受体结合时快速内吞,且一旦内吞,免疫配体可以分解到可令效应分子,例如药物,暴露出来的程度。举一个突出的例子,硫醚连接子,使用SMCC(N-琥珀酰亚胺基-4-(N-马来酰亚胺甲基)-环己烷-1-羧酸酯)连接子(参见图14a,结构2)。

[0096] 所有这些方法的共同之处在于不存在真正的连接反应位点特异性。因为连接子介导的化学结合时随机过程,连接子介导的效应分子化学连接会得到结合蛋白的不均一混合物,其在疗效和/或诊断潜力上均可能不同。显然,蛋白-效应分子结合物的混合物也对治疗结合物的监管审批过程带来了显著挑战,因为批次间差异和/或活性药学成分差异(API)由于潜在安全隐患而导致了监管部门的负面看法。

[0097] 本申请的又一优选实施例中,免疫配体/效应分子结合物中包含的免疫配体至少选自以下一种:至少选自以下一种

[0098] 抗体,修饰抗体形式,抗体衍生物或片段,和/或

[0099] 拟抗体。

[0100] 优选地,在此实施例中,通过小于13个氨基酸的肽与小分子效应分子结合,得到用作序列-特异性转肽酶底物的小分子效应分子,由此其可通过转肽酶结合到含有可由所述转肽酶识别的C端修饰的单克隆抗体的C端。所述C端修饰可以包含在两条重链中,或两条轻链中,或全长抗体的重链和轻链中,从而允许以药物-抗体比例2或4(DAR2或DAR4)生成位点特异性结合的ADC。

[0101] 本申请的又一优选实施例中,免疫配体与至少一种选自以下的实体结合:

[0102] 受体,

[0103] 抗原,

[0104] 生长因子,

[0105] 细胞因子,和/或

[0106] 激素。

[0107] 本申请所称的术语“受体”指的是细胞表面分子,优选地为如下细胞表面分子(i)结合特异性的信号分子或信号分子群(例如受体,例如,血管内皮生长因子受体),和/或(ii)没有已知配体(即孤儿受体,例如,HER2/neu)。天然受体表达在细胞群表面上,或者它们仅仅具有所述分子的胞外域(无论所述形式是否天然存在),或在血浆或细胞或器官内执行天然结合功能的可溶性分子。优选地,所述受体是涉及特定致病过程的信号级联成员(例如,参与生长因子信号级联的受体),或者表达在参与病理过程的细胞或粒子(例如癌细胞)表面上。

[0108] 本申请所称的术语“抗原”是指具有诱导特异性免疫应答且可包含表面蛋白或蛋白复合物(例如,离子通道)的能力的物质。很多时候,抗原与致病实体相关,例如癌细胞。

[0109] 本申请所称的术语“细胞因子”是指小细胞信号传导蛋白质分子,其由许多细胞分泌,且是一类广泛用于胞内通讯的信号传导分子。细胞因子可以被分类为蛋白质,肽,或糖蛋白;术语“细胞因子”包含了大量和多样化的调控因子家族,其在全身由多样化胚胎起源细胞产生。

[0110] 本申请所称的术语“生长因子”涉及天然存在的能够刺激细胞生长、增殖和细胞分化的物质。通常,生长因子是蛋白质或类固醇激素。生长因子是用于调节多种细胞过程的重要因子。

[0111] 本申请所称的术语“激素”涉及由体内局部发出影响生物体其他部分细胞的信息的细胞、腺或器官释放的化学物质。该术语包括肽激素,脂质和磷脂衍生的激素(包括类固醇激素),和单胺。

[0112] 若免疫配体结合受体或抗原,所述免疫配体效应分子结合物可以例如被引导到一个特异的位点,例如,到一个致病实体,例如癌细胞,向其递送所述效应分子,例如毒素或化疗剂。因此,毒素或化疗剂的全身性毒性降低,而后者在作用位点的局部浓度增加,从而提供更好的疗效,而副作用降低。此外,相应的信号传导级联可被免疫配体的结合抑制。在效应分子是标记的情况下,后者可以用来标记特异位点,例如,通过免疫配体检测以特定表面抗原为特征癌细胞,用于诊断。

[0113] 若免疫配体结合生长因子,细胞因子,和/或激素,则所述免疫配体/效应分子结合物可以例如被引导到生长因子、细胞因子或激素通常结合的位点,从而位点特异性地递送效应分子。此外,相应的信号传导级联可由免疫配体的结合抑制。

[0114] 本申请所称的术语“结合”是指本领域熟知的免疫配体(例如,抗体或抗体片段)与它们的靶标之间的相互作用或其它非随机结合。优选地,这种结合反应的特征在于其高特异性和/或靶标灵敏度。优选地,该结合反应的特征在于解离常数(Kd)  $\leq 10^{-3}M$ ,优选为  $\leq 10^{-4}M$ ,  $\leq 10^{-5}M$ ,  $\leq 10^{-6}M$ ,  $\leq 10^{-7}M$ ,  $\leq 10^{-8}M$ ,  $\leq 10^{-9}M$ ,且最优选为  $\leq 10^{-10}$ 。

[0115] 本发明的一个优选实施例中,所述序列特异性转肽酶的至少一个催化域融合到免

疫配体或效应分子的N端或C端。

[0116] 所述融合的发生可以通过重组工程或通过化学连接。在本实施例中,促成免疫配体到效应分子的位点特异性结合的酶活性,并不需要作为单独的重组酶被加入到反应中,相反却是待结合的蛋白底物的一部分。

[0117] 优选地,所述序列特异性转肽酶是选自下组中的至少一种:

[0118] 分选酶,或其一种或多种片段或衍生物;

[0119] 分裂内含肽,或其一个或多个片段或衍生物。

[0120] 在一个优选的实施例中,该转肽酶是分选酶,作为分选酶结合底物的该效应分子(例如,毒素)优选通过加入少量的甘氨酸氨基酸残基,优选3或5个甘氨酸残基,来得到。

[0121] 在另一个优选的实施例中,其中该转肽酶是分裂内含肽,例如Ssp GyrB的分裂内含肽,作为分选酶结合底物的所述效应分子,例如毒素,是通过加入少于13个氨基酸残基的GVFVHN-SX<sub>n</sub>序列得到的,X是任何氨基酸,n是0 $\geq$ 和 $\leq$ 5之间的整数。

[0122] 现有技术中(Panowski et al. (2014))尚未记载使用转肽酶(优选为分选酶和分裂内含肽)令小分子量毒素与全长抗体结合来制备抗体药物结合物。

[0123] 分选酶已在各种革兰氏阳性菌,如金黄色葡萄球菌,链球菌和肺炎球菌物种中被鉴定发现,并在体内催化致病因子与细胞壁蛋白聚糖的结合,以改变细菌表面特征从而逃避被感染宿主的有效免疫应答(Mazmanian et al. (1999))。

[0124] 革兰氏阳性菌金黄色葡萄球菌(*Staphylococcus aureus*)的分选酶A(sortase A)首先得到表征(Ton-That et al. (1999)),其后进一步成为多种蛋白修饰中的工具(Tsukiji (2009))。

[0125] 分选酶的有益特征之一在于,两个待结合分子可以仅需要短肽标记(“分选酶标记”),其中对于金黄色葡萄球菌的分选酶A是,例如,在一个分子的C端(例如,效应分子)的LPXTG,和在另一分子的N端(例如,免疫配体,参见图1)的一个3至5个氨基酸甘氨酸短链。这些肽标记可通过常规的交联化学方式融合或结合到分子。这一方面允许利用该系统连接两种蛋白,另一方面也可用于结合小分子量化合物,优选小分子量的毒素蛋白。对于金黄色葡萄球菌的分选酶B的情况,各分选酶基序是NPQTN。

[0126] 内含肽最初作为蛋白内含子被发现,它能够通过断开肽键并形成新肽键来将自己移除(剪接)出前体蛋白(Xu et al. (1993)) (图2a)。

[0127] 天然和人造分裂内含肽包括,该内含肽编码区已经被分成N-内含肽域和C-内含肽域,其可以连接到不同的蛋白或肽,其连接方式使得随后的外显肽域的反式剪接(图2b)导致两种蛋白的结合。

[0128] 分裂内含肽就此被用于N-外显肽和C-外显肽部分的共价结合,也可用于蛋白的纯化和/或环化(Elleuche (2010))。本文所公开的一个实施例中,利用分裂内含肽用于小分子量化合物的结合,优选为小分子量毒素和其他小分子标记,其中小于13个氨基酸的C-外显肽短肽序列被连接到任何尺寸的分子上,类似于分选酶介导的转肽所要求的甘氨酸短链。

[0129] 对于分选酶而言,将甘氨酸短链的(>2个甘氨酸残基)添加到所选分子足以使分子被结合到含有五肽分选酶识别基序(例如像金黄色葡萄球菌分选酶A的LPXTG)的免疫配体。对于分裂内含肽而言,包含来自Ssp GyrB的6个氨基酸的短C-内含肽(GVFVHN)和短C-外显

肽(这里:SAGSGK)的最低限度12氨基酸GVFVHNSAGSGK氨基酸短链,足以修饰任何效应分子,优选为小分子量毒素,以用于通过分裂内含肽介导结合到含有Ssp GyrB分裂内含肽N-内含肽域等免疫配体上(Volkman et al. (2009))。还可以使用其他功能性内含肽域可以降低到<13个氨基酸长的小肽段的分裂内含肽。

[0130] 即使在文献中分裂酶并不总是称为酶,但其却具有此功能,因为其催化的反应导致肽键的断裂和新肽键的形成,因此可以被看作是转肽酶,因为现有肽键的能量被转移到了新的肽键。

[0131] 除了化学结合,转肽酶介导的结合发生在生理缓冲水溶液和生理温度的条件下,从而最低限度地影响结合反应中蛋白或抗体的完整性。此特征可确保所制得的结合物的最佳功能。

[0132] 本发明的另一个优选的实施例中,包含在免疫配体/效应分子结合物的效应分子是选自下组中的至少一种:

[0133] 标记物;

[0134] 加工标签,和/或

[0135] 药物。

[0136] 本申请所称术语“标记物”(也称为“检测标记”),可以指包含一种或多种适当的化学物质或酶的任何分子或部分,其在化学,物理或酶促反应中直接或间接产生可检测的化合物或信号。

[0137] 本申请所称术语“加工标签”,可以包括亲和标记,增溶标记,色谱标记和表位标记。亲和标记(也用作纯化标记)被附加到蛋白上,从而能够使用亲和技术从其粗生物源中纯化被标记分子。这些包括几丁质结合蛋白(CBP),麦芽糖结合蛋白(MBP),谷胱甘肽-S-转移酶(GST)。聚(组氨酸)标记,优选为6xHis标记,是一种广泛使用的加工标签,它与金属基质结合。增溶标记用于,特别是对于表达分子伴侣缺失物种(如大肠杆菌)中的重组蛋白,以协助蛋白的正确折叠并防止其析出。这些包括硫氧还蛋白(TRX)和聚(NANP)。某些亲和标记有作为增溶剂(如MBP)和GST的双重作用。

[0138] 色谱标记用于改变蛋白的色谱特性,以提供特异性分离技术的不同精度。通常,这些包括聚阴离子氨基酸,如FLAG标记。

[0139] 表位标记是因为能够在许多不同的物种中可靠地产生高亲和力抗体而被选定的短肽序列。表位标记通常来自病毒基因,这解释了它们的高免疫反应。表位标记包括,例如,V5标记,MYC标记和HA标记。这些标记在免疫印迹,免疫荧光和免疫沉淀实验中特别有用,但它们也可用于蛋白纯化中。

[0140] 加工标签还有许多其他用途,如特异性酶修饰(如生物素连接酶标记)和化学修饰(FLAsH)标记。通常,使用组合标记以对蛋白进行多功能修饰。

[0141] 优选地,所述标记物是选自下组中的至少一种:

[0142] 放射性标记,优选为放射性标记的肽或蛋白;

[0143] 荧光标记,优选为荧光肽或蛋白,和/或

[0144] 酶标记,优选过氧化物酶。

[0145] 所述对作为效应分子的潜在标记物的列举绝不是限制性的。在另一个优选的实施例中,所述药物是选自下组中的至少一种:

- [0146] 细胞因子;
- [0147] 放射性试剂;
- [0148] 抗炎药;
- [0149] 毒素,和/或
- [0150] 化疗剂。

[0151] 所述对作为效应分子的潜在药物的列举绝不是限制性的。本申请所称“细胞因子”是指由许多细胞分泌并且作为一类广泛用于细胞间通信的信号传导分子的细胞信号传导小蛋白分子。细胞因子可以被分类为蛋白质,肽,或糖蛋白;术语“细胞因子”包含了大量和多样化的调控因子家族,其在全身由多样化胚胎起源细胞产生。在本申请上下文中,细胞因子用于例如削弱甚至杀死致病实体,例如癌细胞。

[0152] 本申请所称术语“放射性试剂”涉及具有至少一个含不稳定核的原子的实体,该实体容易因此经历放射性衰变,放射 $\gamma$ 射线和/或亚原子粒子(如 $\alpha$ 或 $\beta$ 粒子),从而具有细胞杀伤效果。在本申请上下文中,放射性药剂是为了削弱甚至杀死致病实体,例如癌细胞。

[0153] 本申请所称术语“抗炎药”涉及减少炎症的化合物。这可以是,例如,类固醇,如特异性糖皮质激素(通常被称为糖皮质激素),其通过结合糖皮质激素受体减少炎症或肿胀。该术语还包括非甾体抗炎药(NSAID),其拮抗环氧合酶(COX)的酶。就其本身而言,COX酶合成前列腺素,产生炎症。在总体上,NSAIDs的防止前列腺素合成,从而减少或消除疼痛。该术语还涵盖免疫选择性抗炎衍生物(ImSAIDs),它是一类改变炎症细胞的激活和迁移的肽,炎症细胞是负责放大炎症反应的免疫细胞。

[0154] 本申请所称“毒素”术语涉及一种对活细胞或有机体有毒的分子。毒素可以是肽,或蛋白或优选为小分子量化合物,其用于削弱甚至杀死致病实体,例如癌细胞。毒素,如本文所意指,尤其包括蜂窝毒素。优选地,所述毒素是一种小分子毒素,即,具有 $\leq 2500$ 道尔顿的分子量。

[0155] 本申请所称术语“化疗剂”涉及具有抑制肿瘤特别是恶性(癌性)病变(例如癌,肉瘤,淋巴瘤,白血病)的发展或进展的功能性质的分子。抗癌剂或化疗剂通常具有抑制转移或血管生成的性质。化疗剂可以是细胞毒性剂或化疗剂。优选地,所述化疗剂是小分子量细胞抑制剂,其抑制或压制癌细胞的生长和/或繁殖。

[0156] 将细胞因子、放射性剂、毒素或化学治疗剂与免疫配体结合,可以帮助减少副作用和给药相关风险,因为:

[0157] a) 该免疫配体引导结合物到特异性位点,例如,到致病实体,如癌细胞,其中效应分子影响其毒性作用。因此,效应分子的全身性毒性降低,而后者则在作用部位的局部浓度增加,从而提供更好的疗效和较低的副作用。

[0158] b) 可以令该结合物内吞入致病实体中,由此效应分子在内吞后释放,才开始发展其期望的细胞毒作用,即不影响周围的细胞或组织。

[0159] 下表是潜在标靶/抗原的非限制性列表(第一列)和用于靶向前者的现有免疫配体的实施例(第二列)。第三列显示了潜在的毒素、细胞因子或化疗药物的非限制性列表。注意,第一和第三列中的实施例可以彼此随意组合,且还存在成百种其他标靶和效应分子。本发明的范围涵盖未在表中明确地提到的各个标靶/效应分子组合。

[0160]

标靶/抗原	现有免疫配体实施例	效应分子
表皮生长因子受体 (EGFR)	西妥昔单抗	Maytansinoides, 例如 Mertansine, 安丝菌素, Ravtansin, DM4, DM1
CD20	利妥昔单抗, 替伊莫单 抗, 托西莫单抗(mAb)	刺孢霉素, 例如奥佐米星
CD44		多柔比星
MUC1	莫坎妥珠单抗(mAb)	细菌绿脓杆菌外毒素 PE38
CD30	Brentuximab 单 抗 (mAb)	单甲基阿里他汀 F (MMAF); 单甲基阿里他汀 E (MMAE)
CD22	奥英妥珠单抗 (mAb)	吡咯开苯并吡庚三烯(PBD)
跨膜糖蛋白 NMB (GPNMB)	Glembatumumab 单 抗 (mAb)	白细胞介素 n-10 (IL10) (抗 炎)



[0161]

CD56	Lorvotuzumab 单抗 (mAb)	白喉毒素
CanAg	huC242 (mAb)	肿瘤坏死因子(TNF)
促黄体激素释放激 素(LHRH) 受体	[D-Lys(6)] LHRH	核糖核酸酶
前列腺特异性膜抗 原 (PSMA)		钇 <sup>90</sup>
CD74	Milatuzumab 单抗 (mAb)	碘 <sup>131</sup>
CD70		镱 <sup>177</sup>
AGS-16		环孢素
整合蛋白		甲氨蝶呤
CD19		紫杉烷类, 例如, 紫杉醇或 多西紫杉醇
粘连蛋白-4		
白细胞介素 2 受体	白 细 胞 介 素 -2 (Proleukin)	
CD3	UCHT1 (mAb)	
纤连蛋白胞外域 B	L19-SIP (与恒定域 CH4 融合的 scFv )	
SLAMF7 (CD319)	Elotuzumab 单抗 (mAb)	
SDC1	Indatuximab 单抗 (mAb)	
Her-2/neu	曲妥单抗(mAb)	
CD33	吉姆单抗(mAb)	

[0162] 本发明的又一实施例中,免疫配体包括至少两个分别与效应分子结合的亚单位。

[0163] 优选地,至少两个不同的效应分子可以结合到所述至少两个亚单位。此选项提供了可产生大量各种不同的免疫配体-效应分子结构的多功能工具库。例如,双特异性双域免疫配体可以与两个不同的效应分子结合,例如,一个标记和一个毒素。

[0164] 优选地,该至少两个不同的效应分子是干扰一种或多种细胞通路的毒性效应分子。

[0165] 上述实施方式的实现,可以通过,例如,通过利用两种不同的分选酶识别不同分选酶识别基序,并在所用抗体的轻链和重链上包含具有针对所述不同分选酶的识别基序的C端修饰,令全长抗体的两个轻链和两个重链各结合两个不同的效应分子。

[0166] 由此,可以得到由每两个全长Ig轻链和Ig重链组成的抗体药物结合物,其含有共价连接到所述重链和轻链的不同效应分子。

[0167] 所述实施例,优选地,使得所述至少两个亚单位的同步结合,生成每个所述亚单位结合相等效应分子的免疫配体-效应分子。

[0168] 在另一个优选的实施例中,所述具有至少两个亚单位的免疫配体的结合效率为每个结合位点至少80%。

[0169] 在另一个优选的实施例中,所述具有至少两个亚单位的免疫配体包含至少两个氨基酸组成的肽间隔序列,优选2-5个氨基酸,附加到两个亚单位中至少一个的C端。

[0170] 优选地,所述方法使得所述至少两个亚单位的同步结合,生成每个所述亚单位结合相等效应分子的免疫配体-效应分子。本发明的又一实施例中,该方法允许免疫配体和效应分子之间具有确定的化学计量配比。根据本实施例,可以在免疫配体和效应分子之间提供严格的定量关系,从而提高了各免疫配体/效应分子结合物的再现性和整体性能,特别是用于临床和/或治疗应用。这是通过所用转肽酶的序列特异性和/或位点特异性所带来的。

[0171] 在一个特别优选的实施例中,所述免疫配体和效应分子之间的化学计量配比限定关系,是通过除去部分反应的C端修饰免疫配体底物来实现的。这种移除可以,例如,通过亲和纯化来进行。所述方法,优选地,得到均一的药物免疫配体比例。

[0172] 优选地,所述移除是通过利用位于转肽酶识别基序或结构域的C端的亲和标记进行亲和纯化。本领域技术人员已知的标准方法可以用于这一目的,例如,His标记,CBP标记,CYD(共价但尚未解离NorpD肽)标记,链球菌II标记,FLAG标记,HPC(蛋白C的重链)标记,以及GST和MBP蛋白融合标记。

[0173] 本发明的又一实施方式中,该方法允许效应分子与免疫配体的位点特异性结合。根据本实施例,可以确保结合过程不干扰免疫配体或效应分子本身的活性,从而提高了各免疫配体/效应分子结合物的再现性和整体性能,尤其是用于临床和/或治疗应用。这是通过所述转肽酶的序列特异性和/或位点特异性达成的。除了在大多数情况下没有或仅有有限位点特异性(例如,当效应分子结合至游离氨基,如精氨酸,赖氨酸,天冬酰胺或谷氨酰胺时)的常规结合化学,可以精确地确定所述结合位点,从而使免疫配体(例如,靶特异性)或效应分子(例如,毒性)的特征功能不受影响。

[0174] 本发明还提供根据上述实施例的方法获得的免疫配体/效应分子结合物。

[0175] 优选地,所述免疫配体/效应分子结合物选自抗体/药物结合和/或抗体/标记物结合物。

[0176] 本发明还提供上述实施例的免疫配体/效应分子结合物的以下用途:

[0177] 指定病理情况的体外或体内诊断;

[0178] 对指定病理情况的体内或体外预测或预后;

[0179] 罹患或有风险罹患指定病理情况的人或动物对象的治疗;和/或

[0180] 研究和/或开发目的。

[0181] 优选地,所述病理情况选自以下至少一种:

[0182] 肿瘤疾病;

[0183] 自身免疫疾病;

[0184] 神经变性疾病;和/或

[0185] 传染病。

[0186] 在所有这些情况下,本发明的免疫配体/效应分子结合物可以具有有益效果,例如,通过引导后者到特异性位点,例如癌细胞,神经病理学部位,或自身免疫反应部位。

[0187] 所述效应分子,例如,毒素,化疗剂,细胞因子或药物被递送至所述部位,例如,以耗尽癌细胞,抗癌细胞增殖,溶解菌斑,抑制自身抗体等。

[0188] 在所有这些情况下,本发明的免疫配体/效应分子结合物可以具有有益效果,例如,通过引导后者到特异性位点,例如癌细胞,其中递送所述效应分子(例如毒素或化疗剂),例如,以耗尽癌细胞,抗癌细胞增殖。

[0189] 由此,毒素或化疗剂的全身性毒性降低,且后者在作用部位的局部浓度增加,提供更好的疗效和较低的副作用。此外,相应的信号传导级联可被免疫配体的结合抑制。在效应分子是标记的情况下,后者可以用来标记特异位点,例如,通过免疫配体检测以特定表面抗原为特征癌细胞,用于诊断。

[0190] 结合过程的位点特异性保证了相应免疫配体/效应分子结合物的高再现性和整体性能,特别是用于临床和/或治疗应用。

[0191] 本申请所称术语“肿瘤性疾病”,是指细胞或组织的异常状态或症状,其特征是快速增殖的细胞生长或肿瘤。更具体而言,该术语涉及癌性进程,例如,肿瘤和/或白血病。

[0192] 术语“神经病理学疾病”包括神经变性疾病,神经炎性疾病或痉挛疾病等。

[0193] 神经变性疾病的特征是神经元的结构或功能的渐进性丧失,包括神经元死亡。许多神经变性疾病,包括帕金森症,阿尔茨海默氏症,亨廷顿舞蹈症,肌萎缩性脊髓侧索硬化症和多发性硬化症的发生,都是神经变性进程的结果。不同的神经退行性疾病之间有很多相似之处,包括不规则蛋白装配以及诱导性细胞死亡。神经变性可以进一步发现于从分子到全身的许多不同水平的神经元系统。

[0194] 术语“神经变性疾病”和“神经炎性疾病”具有部分重叠的范围。炎症反应是神经变性疾病的标志,并通过不同神经元细胞死亡机制参与或促进该病症。沿犬尿氨酸路径(KP)的色氨酸分解代谢代表了这些机制之一。

[0195] 痉挛疾病是以脑细胞之间的异常信号传导为特征的脑功能障碍。痉挛疾病可以影响部分脑(部分性发作)或整个脑(全身性发作)。最突出的痉挛疾病是癫痫。

[0196] 本申请所称术语“自身免疫性疾病”,包括器官特异性自身免疫性疾病,其中自身免疫反应针对单一组织,如克罗恩病和溃疡性结肠炎,I型糖尿病,重症肌无力,白癜风,格雷夫斯疾病,桥本氏病,阿狄森氏病和自身免疫性胃炎和自身免疫性肝炎。该术语还包括非器官特异性自身免疫性疾病,其中自身免疫反应针对的是存在于全身几个或许多器官的部件。

[0197] 所述自身免疫疾病包括,例如,类风湿性关节炎,疾病,全身性红斑狼疮,全身性进行性硬化症及其变型,多发性肌炎和皮炎。

[0198] 其他自身免疫性疾病包括恶性贫血(包括一些自身免疫性胃炎),原发性胆汁性肝硬化,自身免疫性血小板减少症,干燥综合症,多发性硬化和牛皮癣。本领域技术人员能够

理解,本发明的方法可根据需要应用于上述或其他自体免疫疾病。

[0199] 本文所称术语“感染性疾病”,包括但不限于,感染性生物体所引起的任何疾病。感染性生物体可以包括病毒,(例如,单链RNA病毒,单链DNA病毒,人类免疫缺陷病毒(HIV),肝炎A、B和C病毒,单纯疱疹病毒(HSV),巨细胞病毒(CMV),EB病毒(EBV),人乳头状瘤病毒(HPV)),寄生虫(例如,原生动物病原体和后生动物病原体如疟原虫物种,利什曼原虫属物种,血吸虫物种,锥虫物种),细菌(例如,分枝杆菌,特别是结核分枝杆菌,沙门氏菌,链球菌,大肠杆菌,葡萄球菌),真菌(例如,念珠菌属,曲霉属物种),卡氏肺囊虫,以及朊病毒。

[0200] 本发明进一步提供了一种Gly<sub>n</sub>-修饰(n>1,优选n=3或n=5)的低分子量效应分子。

[0201] 本申请所称术语“Gly<sub>n</sub>-修饰”是指由n个甘氨酸残基组成的寡肽或肽添加到所述效应分子。本申请所称术语“低分子量效应分子化合物”应包括具有2500道尔顿或以下分子量的效应分子。

[0202] 所述效应分子,优选地,包含选自下列的至少一种:

[0203] 标记物,

[0204] 加工标签,和/或

[0205] 药物。

[0206] 所述标记物包含选自下列的至少一种:

[0207] 放射性标记,优选为放射性标记肽或蛋白

[0208] 荧光标记,优选为荧光肽或蛋白,和/或

[0209] 酶标记,优选为过氧化物酶。

[0210] 所述药物包含选自下列的至少一种:

[0211] 细胞因子

[0212] 放射性试剂

[0213] 毒素,和/或

[0214] 化疗剂。

[0215] 如上所述,所述毒素优选为小分子毒素,即具有≤2500道尔顿的分子量。优选地,所述毒素是选自下列的至少一种:

[0216] 美登素,

[0217] 单甲基阿里他汀,和/或

[0218] α-鹅膏蕈碱,

[0219] 或前者的衍生物。上述Gly<sub>n</sub>-修饰毒素的实施例如图14A-14C的结构1-9所示。

[0220] 本发明进一步提供了甘氨酸修饰的低分子量效应分子用于与免疫配体结合的用途。

[0221] 优选地,如上所述,所述结合是转肽酶(优选为分选酶和/或分裂内含肽)介导的结合。同样优选地,所述免疫配体是抗体。

[0222] 优选地,所述免疫配体是抗体。由此,可以提供抗体药物结合物(ADC)。

[0223] 优选地,所述免疫配体-效应分子结合反应在粗细胞培养物上清液中进行。这意味着,优选地,所述结合反应可用未纯化或仅部分纯化成分来进行。

[0224] 实验与附图

[0225] 虽然本发明已经在附图和上文中进行了详细说明描述,这些说明和描述应被认为是说明性或示例性而非限制性的。本发明并不限于所公开的实施例。本领域技术人员可以通过研究附图、说明书和权利要求,实施本发明权利要求中的方案,理解并制作上述公开实施例的其它变型。在权利要求中,词语“包括”不排除其他元件或步骤,并且不定冠词“一”或“一个”并不排除多个。虽然不同从属权利要求中列举了某些措施,但这并不表示不能被有利地使用这些措施的组合。权利要求中的任何附图标记不应当被解释为限制本发明的范围。

[0226] 本文所公开的所有氨基酸序列分别显示为从N端到C端的序列;本文所公开的所有核酸序列分别显示5'→3'的序列。

[0227] 实施例1:带有C端LPETG分选酶标记和附加的6x-His和strepII亲和纯化标记的CD19单克隆抗体的表达载体克隆与表达

[0228] 为了进行效应分子与抗体的C端结合,首先,需要表达含有包括识别基序(例如,用于金黄色葡萄球菌分选酶A的识别基序)的C端修饰的重组抗体。

[0229] 为了,可以基因合成用于抗人CD19特异性抗体的重链和轻链ORF,例如,利用提供所述基因合成的合同研究组织(CROs),例如Genscript(www.genscript.com,Piscataway, NJ, USA)。举例而言,可以在美国专利US8242252B2的序列53(变体HF)和58(变体LG)中找到人源化抗人CD19抗体hBU12的重链和轻链序列。所述抗人CD19抗体的V<sub>H</sub>和V<sub>L</sub>区域如下:

[0230] SEQ ID NO 1(人源化抗人CD19抗体hBU12的V<sub>H</sub>编码区域):

[0231] ATGGGATGGAGCTGGATCTTTCTTTCTCTCTCTGTCAGGAAGTGCAGGTGCCATTGTCAGGTTTCAGCTGCAAGAGTCTGGCCCTGGGTTGGTTAAGCCCTCCAGACCCTCAGTCTGACTTGTACTGTGTCTGGGGGTTCAATCAGCACTTCTGGTATGGGTGTAGGCTGGATTAGGCAGCACCCAGGGAAGGGTCTGGAGTGGATTGGACACATTTGGTGGGATGATGACAAGAGATATAACCCAGCCCTGAAGAGCAGAGTGACAATCTCTGTGGATACCTCCAAGAACCAGTTTAGCCTCAAGCTGTCCAGTGTGACAGCTGCAGATACTGCTGTCTACTACTGTGCTAGAAATGGAACCTTTGGTCCTACTATTTGACTACTGGGGCCAAGGCACCCTTGTCACAGTCTCCTCA

[0232] SEQ ID NO 1可翻译为如下氨基酸序列(SEQ ID NO 2):

[0233] MGWSWIFLFLLSGTAGVHCQVQLQESGPGLVKPSQTLSTCTVSGGSISTSGMGVWIRQHPGKGLEWIGHIWWDDDKRYNPALKSRVTISVDTSKNQFSLKLSSVTAADTAVYYCARMELWSYFFDYWGQGLVTVSS

[0234] SEQ ID NO 3(人源化抗人CD19抗体hBU12的V<sub>L</sub>编码区域):

[0235] ATGAAGTTGCCTGTTAGGCTGTTGGTGCTGATGTTCTGGATTCTGCTTCCAGCAGTGAAATTGTTCTCACCCAGTCTCCAGCAACCCTGTCTCTCTCTCCAGGGGAAAGGGCTACCCTGAGCTGCAGTGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAGAAGCCAGGGCAGGCTCCAGACTCCTGATTTATGACACATCCAACTGGCTTCTGTATTCCAGCAAGGTTTCAGTGGCAGTGGGCTGGAACAGATTTTACTCACAAATCAGCAGCCTGGAGCCAGAGGATGTTGCTGTCTATTACTGTTTTTCAGGGGAGTGATACCCATTCATTTTGGCCAAGGGACAAAGTTGAAATCAAA

[0236] SEQ ID NO 3可翻译为如下氨基酸序列(SEQ ID NO 4):

[0237] MKLPVRLVLVLMFWIPASSSEIVLTQSPATLSLSPGERATLSCSASSSVSYMHWYQQKPGQAPRLLIYDTSKLASGIPARFSGSGSTDFLTITSSLEPEDVAVYYCFQGSVYPFTFGGQGTKLEIK

[0238] 上述序列可以融合到含有附加C端标记的人IgG<sub>1</sub>恒定重链和恒定轻链区域中,以实现上述方法。

[0239] 为了实现本发明,可以合成具有编码LPETC金黄色葡萄球菌分选酶A识别标记、6xHis

标记(HHHHHH)、MYC标记(EQKLISEEDL)和strep II标记(WSHPQFEK)的附加3'-密码子的人恒定IgG<sub>1</sub>重链区域。

[0240] SEQ ID NO 5 (带有编码LPETG分选酶标记、6xHis标记和strepII标记的框内3'延长段的人IgG<sub>1</sub>重链恒定编码区域):

[0241] AGCACCAAGGGCCCATCTGTCTTCCCCCTGGCACCCCTCCTCCAAGAGCACCTCTGGGGGCACAGCTGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCTGAACCTGTGACAGTGTCTGGAACCTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGTGTCTACAGTCTCAGGACTTACTCCCTCAGCAGCGTGGTGACCGTGCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAACTCTGTGACAAAACCTCACACATGCCACCGTGGCCAGCACCTGAACCTCTGGGGGGACCGTCAGTCTTCTCTTCCCACCAAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCTCACCGTCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTCAAGGTCTCCAACAAAGCCCTCCAGCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGAGAACAACACTACAAGACCACGCTCCCGTGTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACACAGAAGAGCCTCTCCCTGTCTCCGGTAAACTGCCGAGACCGCCACCACCACCACCACCACGGCGAGCAGAAGCTGATCAGCGAGGAGACCTGGGCTGGAGCCACCCCA GTTCGAGAAGTAG

[0242] SEQ ID NO 5可翻译为如下氨基酸序列(SEQ ID NO 6,标记中的氨基酸以下划线表示):

[0243] STKGPSVFLPAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHFTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFNFSVMHEALHNHYTQKSLSLSPGKLPETGHHHHHHHGEQKLISEEDLWSHPQFEK •

[0244] 此外,可以合成具有编码LPETC金黄色葡萄球菌分选酶A识别标记、6xHis标记(HHHHHH)和strep II标记(WSHPQFEK)的附加3'-密码子的人恒定IgG<sub>1</sub>-k轻链区域,得到如下序列:

[0245] SEQ ID NO 7 (带有编码LPETG分选酶标记、6xHis标记、Myc标记和strepII标记的框内3'延长段的人IgG<sub>1</sub>-k轻链恒定编码区域):

[0246] ACGGTGGCTGCACCATCTGTCTTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACCTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGAAGGTGGATAACGCCCTCCAATCGGGTAACCTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTACCCATCAGGGCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTCTGCCCGAGACCGGCCACCACCACCACCACCACGGCGAGCAGAAGCTGATCAGCGAGGAGGACCTGGGCTGGAGCCACCCCAAGTTCGAGAAGTAG

[0247] SEQ ID NO 7可翻译为如下氨基酸序列(SEQ ID NO 8,标记中的氨基酸以下划线表示):

[0248] TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSTYLSLSTLTLKADYEEKHKVYACEVTHQGLSSPVTKSFNRGECLPETGHHHHHHGGEQKLISEEDLGWSHPQFEK •

[0249] 编码LPETC分选酶标记、6xHis标记和strep II标记的人源化抗人CD19抗体hBU12重链和轻链的完整编码区域如下：

[0250] SEQ ID NO 9 (带有C端LPETG分选酶标记、6xHis标记、Myc标记和strepII标记的抗体hBU12完整人IgG1V<sub>H</sub>-C<sub>H</sub>重链编码区域)：

[0251] ATGGGATGGAGCTGGATCTTTCTTTTCTCTCTGTCAGGAAGTGCAGGTGCCATTGTCAGGTTTCAGCTGCAAGAGTCTGGCCCTGGGTTGGTTAAGCCCTCCAGACCCTCAGTCTGACTTGTACTGTGTCTGGGGTTCAATCAGCACTTCTGGTATGGGTGTAGGCTGGATTAGGCAGCACCCAGGGAAGGCTCTGGAGTGGATTGGACACATTTGGTGGGATGATGACAAGAGATATAACCCAGCCCTGAAGAGCAGAGTGACAATCTCTGTGGATACCTCCAAGAACCAGTTTAGCCTCAAGCTGTCCAGTGTGACAGCTGCAGATACTGTGTCTACTACTGTGCTAGAATGGAACCTTTGGTCTACTATTTGACTACTGGGGCCAAGGCACCCTTGTACAGTCTCCTCAGCTAGCACCAAGGGCCCATCTGTCTTCCCCCTGGCACCTCTCTCCAAGAGCACCTCTGGGGGCACAGCTGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCTGAACCTGTGACAGTGTCTGGAACCTCAGGCGCCCTGACCAGCGCGTGCACACCTTCCCGCTGTCTACAGTCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCTCCAGCAGCTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAACAACACCAAGGTGGACAAGAAAGTTGAGCCAAATCTGTGACAAAACCTCACACATGCCACCGTGCCAGCACCTGAACCTCTGGGGGACCGTCACTCTTCTCTTCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAAGGCTCCTCACCCTCCTGACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCATCGAGAAAACCATCTCCAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCATCCCGGATGAGCTGACCAAGAACAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGAGAACAACTACAAGACCACGCCTCCCGTGTGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACACAGAAGAGCCTCTCCCTGTCTCCGGTAAACTGCCGAGACCGGCCACCACCACCACCACCACGGCGAGCAGAAGCTGATCAGCGAGGAGGACCTGGGCTGGAGCCACCCCGAGTTTCGAGAAGTAG

[0252] SEQ ID NO 9可翻译为如下氨基酸序列 (SEQ ID NO 10, 标记中的氨基酸以下划线表示)：

[0253] MGWSWIFLFLLSGTAGVHCQVQLQESGPGLVKPSQTLSTCTVSGGSI<sup>1</sup>STSGMGVGVIRQHPGKGLEWIGHIWDDDDKRYNPALKSRVTISVDTSKNQFSLKLSVTAADTAVYYCARMELWSY<sup>2</sup>YFDYWGQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDK<sup>3</sup>VEPKSCDKHTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK<sup>4</sup>TPPVLDSDGSFFLYSKLTVDKSRWQQGNV<sup>5</sup>FSCSVMHEALHNHYTQKSLSLSPGKLPETGHHHHHHGGEQKLISEEDLGWSHPQFEK •

[0254] SEQ ID NO 11 (带有C端LPETG分选酶标记、6xHis标记、Myc标记和strepII标记的抗体hBU12完整人IgG1V<sub>L</sub>-C<sub>L</sub> k轻链编码区域)：

[0255] ATGAAGTTGCTGTTAGGCTGTTGGTGCTGATGTTCTGGATTCTGCTCCAGCAGTCAAATTGTTCTCACCCAGTCTCCAGCAACCCTGTCTCTCTCTCCAGGGAAAGGGCTACCCTGAGCTGCAGTGCCAGCTCAAGTGTAAAG

TTACATGCACTGGTACCAGCAGAAGCCAGGGCAGGCTCCCAGACTCCTGATTTATGACACATCCAAACTGGCTTCTG  
 GTATTCCAGCAAGGTTTCAGTGGCAGTGGGTCTGGAACAGATTTTACACTACAATCAGCAGCCTGGAGCCAGAGGAT  
 GTTGCTGTCTATTACTGTTTTTCAGGGGAGTGATACCCATTCACCTTTTGGCCAAGGGACAAAGTTGGAATCAAAG  
 AACTGTGGCTGCACCATCTGTCTTCATCTTCCC GCCATCTGATGAGCAGTTGAAATCTGGA ACTGCCTCTGTTGTGT  
 GCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAAGTACAGTGGAAAGTGGATAACGCCCTCCAATCGGGTAACTCC  
 CAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGA  
 CTACGAGAAACACAAAGTCTACGCTGCGAAGTACCCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACA  
 GGGGAGAGTGTCTGCCCAGACCGGCCACCACCACCACCACCACGGCGAGCAGAAGCTGATCAGCGAGGAGGACCTG  
 GGCTGGAGCCACCCCCAGTTTCGAGAAGTAG

[0256] SEQ ID NO 11可翻译为如下氨基酸序列(SEQ ID NO 12,标记中的氨基酸以下划线表示):

[0257] MKLPVRLLVLMFWIPASSSEIVLTQSPATLSLSPGERATLSCSASSSVSYMHWYQQKPGQAPRLLIYD  
 TSKLASGIPARFSGSGSGTDFLTITSSLEPEDVAVYYCFQGSVYPFTFGQGTKLEIKRTVAAPSVFIFPPSDEQLK  
 SGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKSTYLSSTLTLSKADYEEKHKVYACEVTHQGLS  
 SPVTKSFNRGECLPETGHHHHHHGEQKLISEEDLGWSHPQFEK •

[0258] 如SEQ ID NO 9和11所公开的分别编码抗人CD19特异抗体的重链和轻链的编码区域,可以合成侧翼限制性酶位点(例如HindIII和NotI),由此可通过本领域已知的标准分子生物学方法,将其克隆到标准哺乳类表达载体中,例如pCDNA3.1-hygro(+)(Invitrogen)。

[0259] 用于带标记的hBU12抗人CD19抗体的pCDNA3.1-hygro(+)-IgH链表达载体的完整DNA序列如下:

[0260] SEQ ID NO 13(带有C端LPETG分选酶标记、6xHis标记和strepII标记(下划线)以及HindIII和NotI克隆位点(阴影显示)的抗体hBU12完整人IgG1V<sub>H</sub>-C<sub>H</sub>重链编码区域):

[0261] GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGT  
 TAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGC GCGAGCAAATTTAAGCTACAACAAGG  
 CAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAG  
 ATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA  
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCCCTGGCTGACCGCCAACGACCCCGCCATTGACGTCA  
 ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGACTATTTACGGTAAAC  
 TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG  
 CCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT  
 ACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTACTCACGGGGATTTCCAAGTCTCCA  
 CCCCATTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAAATCAACGGGACTTTCAAAATGTCGTAACAACCTCCGCC  
 CATTGACGCAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCC  
 ACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTT  
 CCATGGGATGGAGCTGGATCTTTCTTTTCTCCTCTGTCAGGAACTGCAGGTGCCATTGTCAGGTTTCAGCTGCAAGAG  
 TCTGGCCCTGGGTTGGTTAAGCCCTCCAGACCCTCAGTCTGACTTGTACTGTGTCTGGGGGTTCAATCAGCACTTC  
 TGGTATGGGTGTAGGCTGGATTAGGCAGCACCCAGGGAAGGGTCTGGAGTGGATTGGACACATTTGGTGGGATGATG  
 ACAAGAGATATAACCCAGCCCTGAAGAGCAGAGTGACAATCTCTGTGGATACCTCAAGAACCAGTTTAGCCTCAAG  
 CTGTCCAGTGTGACAGCTGCAGATACTGCTGTCTACTACTGTGCTAGAATGGAACCTTGGTCCTACTATTTTGACTA



CTGGGGCCAAGGCACCCTTGTACAGTCTCCTCAGCTAGCACCAAGGGCCCATCTGTCTTCCCCCTGGCACCCCTCT  
 CCAAGAGCACCTCTGGGGGCACAGCTGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCTGAACCTGTGACAGTGTCC  
 TGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCTACAGTCTCAGGACTCTACTCCCTCAG  
 CAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACA  
 CCAAGGTGGACAAGAAAGTTGAGCCAAATCTTGTGACAAAACCTCACACATGCCACCCTGCCAGCACCTGAACTC  
 CTGGGGGGACCGTCAGTCTTCTCTTCCCCCAAAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTAC  
 ATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATA  
 ATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCTCACCGTCTGCACCAG  
 GACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCATCGAGAAAACCATCTC  
 CAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCATCCCGGGATGAGCTGACCAAGAACCAGG  
 TCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAG  
 AACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAA  
 GAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACACAGAAGA  
 GCCTCTCCCTGTCTCCGGGTAAACTGCCCGAGACCGGCCACCACCACCACCACGGCGAGCAGAAGCTGATCAGC  
 GAGGAGGACCTGGGCTGGAGCCACCCCAAGTTCGAGAAGTAGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCC  
 GCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCCTGCCTTCTTGACCCTG  
 GAAGGTGCCACTCCCCTGTCTTCTTAATAAAAATGAGGAAATGCATCGCATTGTCTGAGTAGGTGTCATTCTAT  
 TCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGG  
 GCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGTATCCCCACGCGCCCTGTAGCGGCGCATT  
 AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTT  
 CTTCCCTTCCCTTCTCGCCACGTTCCGCGGCTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCGGAT  
 TTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAG  
 ACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTAAATAGTGGACTCTGTTCCAAACTGGAACAACACTCAA  
 CCCTATCTCGGTCTATTCTTTTATTATAAGGGATTTTGGGGATTTCCGGCCTATTGGTTAAAAAATGAGCTGATTT  
 AACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAGTCCCCAGGCTCCCCAGGCA  
 GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAGTCCCCAGGCTCCCCAGCAGGCAG  
 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCATCCCGCCCTAACTC  
 CGCCAGTTCGCCCATTCTCCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCTGCC  
 TCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGAGCTTGTAT  
 ATCCATTTTCGGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTGAGAAGTTTCTGATCG  
 AAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGA  
 GGGCGTGGATATGTCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGC  
 ATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCC  
 GTGCACAGGGTGTACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGTGTCTGCAGCCGGTCCGCGAGGCCATG  
 GATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCATTCGGACCGCAAGGAATCGGTCAATACAC  
 TACATGGCGTGATTTTATATGCGCGATTGCTGATCCCATGTGTATCACTGGCAAACCTGTGATGGACGACACCGTCA  
 GTGCGTCCGTGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCAC  
 GCGGATTTCCGGTCCAACAATGTCTGACGGACAATGGCCGATAACAGCGGTCAATGACTGGAGCGAGGCGATGTT  
 CGGGGATTTCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCT

ACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAA  
CTCTATCAGAGCTTGGTTGACGGCAATTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATC  
CGGAGCCGGGACTGTCGGGCGTACACAAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTAC  
TCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAGATTTTCGATTCC  
ACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGA  
TCTCATGCTGGAGTTCTTCGCCACCCCAACTGTTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA  
CAAATTTACAAATAAAGCATTTTTTTTACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCAT  
GTCTGTATAACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATTGTTATCC  
GCTCACAATTCACACAACATAACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA  
CATTAAATTGCGTTGCGCTCACTGCCCCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAA  
CGCGCGGGGAGAGGGGTTTTGCGTATTGGGCGCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCC  
GCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA  
ACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGC  
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGC  
GTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTTCTCC  
CTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTG  
GGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT  
AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAG  
AGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT  
ACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAA  
GCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA  
ACGAAAACCTCACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAA  
TGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC  
TATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGG  
GCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC  
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG  
GGAAGCTAGAGTAAGTAGTTCCGCAGTTAATAGTTTTCGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTAC  
GCTCGTCTTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC  
AAAAAAGCGGTTAGCTCCTTCGGTCCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTAT  
GGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGTGACTCAACCAAGT  
CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGTCAATACGGGATAATACCGCGCCACATAGC  
AGAACTTTAAAAGTGCTCATCATTGAAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCGCTGTTGAGATC  
CAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAA  
AAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTTCTCCTTTTT  
CAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACA  
AATAGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

[0262] 用于带标记的hBU12抗人CD19抗体的pCDNA3.1-hygro (+) IgL链表达载体的完整DNA序列如下:

[0263] SEQ ID NO 14 (带有C端LPETG分选酶标记、6xHis标记、Myc标记和strepII标记(下

划线)以及HindIII和NotI克隆位点(阴影显示)的抗体hBU12人IgG1V<sub>L</sub>-C<sub>L</sub> k轻链编码区域):

[0264] GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGT  
TAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGG  
CAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAG  
ATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA  
TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCCGCCATTGACGTCA  
ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGACTATTTACGGTAAAC  
TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG  
CCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT  
ACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTACTCACGGGGATTTCCAAGTCTCCA  
CCCCATTGACGTCAATGGGAGTTTGTGGCACCAAAAATCAACGGGACTTTCAAAATGTCGTAACAACCTCCGCCC  
CATTGACGCAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCC  
ACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTT  
CCATGAAGTTGCCTGTTAGGCTGTTGGTGCTGATGTTCTGGATTCTGCTTCCAGCAGTCAAATTGTTCTCACCCAG  
TCTCCAGCAACCCTGTCTCTCTCTCCAGGGGAAAGGGCTACCCTGAGCTGCAGTGCCAGCTCAAGTGAAGTTACAT  
GCACTGGTACCAGCAGAAGCCAGGGCAGGCTCCAGACTCCTGATTTATGACACATCCAAACTGGCTTCTGGTATTC  
CAGCAAGGTTTCAGTGGCAGTGGGTCTGGAACAGATTTTACACTACAATCAGCAGCCTGGAGCCAGAGGATGTTGCT  
GTCTATTACTGTTTTTCAGGGGAGTGATACCCATTCACTTTTGGCCAAGGGACAAAGTTGGAATCAAAGAAGTGT  
GGCTGCACCATCTGTCTTCATCTTCCC GCCATCTGATGAGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGC  
TGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAG  
AGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCTGACGCTGAGCAAAGCAGACTACGA  
GAAACACAAAGTCTACGCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAG  
AGTGTCTGCCCGAGACCGGCCACCACCACCACCACGGCGAGCAGAAGCTGATCAGCGAGGAGGACCTGGGCTGG  
AGCCACCCCGAGTTCGAGAAGTAGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTG  
TGCCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTCCTTCCCTTGACCCTGGAAGGTGCCACTCCCCT  
GTCCTTTCCTAATAAAATGAGGAAATGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGG  
GCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGG  
CGGAAAGAACCAGCTGGGGCTCTAGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTG  
GTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTCTTCCCTTCCCTTCTCGC  
CACGTTCCGCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTCCGATTTAGTGCTTTACGGCACC  
TCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTG  
ACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTGTTCAAAAGTGAACAACACTCAACCCTATCTCGGTCTATTC  
TTTTGATTTATAAGGGATTTTGGGATTTTCGGCCTATTGGTTAAAAATGAGCTGATTTAACAAAAATTTAACGCGA  
ATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGAAAAGTCCCCAGGCTCCCCAGGCAGGCAGAAGTATGCAAAGC  
ATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCA  
TCTCAATTAGTCAGCAACCATAGTCCC GCCCTAACTCCGCCATCCCCGCCCTAACTCCGCCAGTTCGCCCAT  
CTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCCTCTGCCTCTGAGCTATTCCAGAAG  
TAGTGAGGAGGCTTTTTTGGAGGCTAGGCTTTGCAAAAAGTCCCCGGGAGCTGTATATCCATTTTCGGATCTGA  
TCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCT

CCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTG  
CGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCC  
TCCGGAAGTGCTTGACATTGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTACGT  
TGCAAGACCTGCCTGAAACCGAACTGCCCCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCC  
GATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTTCAT  
ATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACCTGTGATGGACGACACCGTCAGTGCCTCCGTGCGCAGG  
CTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTTCGGCTCCAAC  
AATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTTCGGGGATTCCCAATACGA  
GGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATC  
CGGAGCTTGACAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGTT  
GACGGCAATTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCCG  
GCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGAAACC  
GACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAGATTTTCGATTCCACCGCCGCTTCTATGAA  
AGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTT  
CGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAG  
CATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATAACCGTCGACC  
TCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCACACAA  
CATAACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCT  
CACTGCCCCTTTCCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGT  
TTGCGTATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATC  
AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCC  
AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAC  
AAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC  
CCTCGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGC  
TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC  
CCCCTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC  
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC  
CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTT  
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG  
AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAG  
GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATC  
TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCT  
ATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCA  
GTGCTGCAATGATAACCGGAGACCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC  
GAGCGCAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAG  
TTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTTGGTATGG  
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCC  
TTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTC  
TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGACTCAACCAAGTCATTCTGAGAATAGTGTA

TGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTC  
 ATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC  
 TCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATG  
 CCGCAAAAAGGGAATAAGGGCGACACGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT  
 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAC  
 ATTTCCCGAAAAGTGCCACCTGACGTC

[0265] 所述结构能够通过转染进入哺乳类细胞,例如但不限于,通常用于重组抗体表达、带有添加至IgH和IgL链C端的分选酶A标记、6xHis标记、Myc标记和strepII的抗人CD19特异性人源化抗体hBU12的CHO细胞。

[0266] 实施例2:带有C端Ssp GyrB 11分裂内含肽和C端附加6x-His和strepII亲和纯化标记的单克隆抗体的表达载体克隆

[0267] 与金黄葡萄球菌分选酶A标记的IgG1重链和轻链的表达组件和载体相似,用于将Ssp GyrB 11分裂内含肽的N-内含肽结构域对IgH和IgL链进行C端融合的编码区域可以如下设计,以通过合格CRO(例如,Genscript(www.genscript.com,Piscataway,NJ,USA)对与如上所公开的抗人CD19携带相同元件的该基因进行基因合成。

[0268] Ssp GyrB 11分裂内含肽的N-内含肽结构域的150aa序列如下所示,可参见论文Appleby et al.(2009):

[0269] SEQ ID NO 15(Ssp GyrB 11分裂内含肽的N-内含肽结构域)

[0270] CFSGDTLVALTDGRSVSFEQLVVEEEKQGKQNFICYTIRHDGSIGVEKI INARKTKTNAKVIKVTLDNGES  
 I ICTPDHKFMLRDSYKCAMDLTDDSLMPLHRKISTTEDSGHMEAVLNYNHRIVNIEAVSETIDVYDIEVPHTHNF  
 ALAS

[0271] 使用哺乳类密码子对该氨基酸序列进行反向翻译,得到如下的Ssp GyrB 11分裂内含肽的N-内含肽结构域编码序列:

[0272] SEQ ID NO 16(Ssp GyrB 11分裂内含肽的N-内含肽结构域编码序列)

[0273] TGCTTCAGCGGCGACACCCTGGTGGCCCTGACCGACGGCAGAAGCGTGAGCTTCGAGCAGCTGGTGGAG  
 GAGGAGAAGCAGGGCAAGCAGAACTTCTGCTACACCATCAGACACGACGGCAGCATCGGCGTGGAGAAGATCATCAA  
 CGCCAGAAAGACCAAGACCAACGCCAAGGTGATCAAGGTGACCCTGGACAACGGCGAGAGCATCATCTGCACCCCCG  
 ACCACAAGTTCATGCTGAGAGACGGCAGCTACAAGTGCGCCATGGACCTGACCCTGGACGACAGCCTGATGCCCTG  
 CACAGAAAGATCAGCACCACCGAGGACAGCGGCCACATGGAGGCCGTGCTGAACTACAACCACAGAATCGTGAACAT  
 CGAGGCCGTGAGCGAGACCATCGACGTGTACGACATCGAGGTGCCCCACACCCACAACCTTCGCCCTGGCCAGC

[0274] 利用所述序列信息,可以设计带有包含Ssp GyrB 11分裂内含肽的N-内含肽结构域、6xHis标记和strepII标记的C端延长段的抗人CD19抗体hBU12的完整IgG1重链编码区域,如下SEQ ID NO 17所示:

[0275] ATGAATTTTGGACTGAGGCTGATTTTCTGGTGCTGACCCTGAAAGGCGTCCAGTGTGAGTTTCAGCTG  
 CAAGAGTCTGGCCCTGGGTTGGTTAAGCCCTCCAGACCCTCAGTCTGACTTGTACTGTGTCTGGGGTTCAATCAG  
 CACTTCTGGTATGGGTGTAGGCTGGATTAGGCAGCACCCAGGGAAGGTCTGGAGTGGATTGGACACATTTGGTGGG  
 ATGATGACAAGAGATATAACCCAGCCCTGAAGAGCAGAGTGACAATCTCTGTGGATACCTCCAAGAACCAGTTTAGC  
 CTCAAGCTGTCCAGTGTGACAGCTGCAGATACTGTCTACTACTGTGCTAGAATGGAACCTTGGTCTACTATTT  
 TGACTACTGGGGCCAAGGCACCCTTGTACAGTCTCCTCAGCTAGCACCAGGGCCATCTGTCTTCCCCCTGGCAC

CCTCCTCCAAGAGCACCTCTGGGGGCACAGCTGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCTGAACCTGTGACA  
 GTGTCCTGGAACCTCAGGCGCCCTGACCAGCGCGTGCACACCTCCCCGGCTGTCTACAGTCCTCAGGACTCTACTC  
 CCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCA  
 GCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACCTCACACATGCCACCGTGCCACGACCT  
 GAACTCCTGGGGGGACCGTCAGTCTTCCCTCTTCCCCCAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGA  
 GGTACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGG  
 TGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTACGCGTCTCACCCTCCTG  
 CACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAAC  
 CATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCATCCCGGGATGAGCTGACCAAGA  
 ACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAG  
 CCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGT  
 GGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACAC  
 AGAAGAGCCTCTCCCTGTCTCCGGGTAAATGCTTCAGCGGCGACACCCTGGTGGCCCTGACCGACGGCAGAAGCGTG  
 AGCTTCGAGCAGCTGGTGGAGGAGGAGAAGCAGGGCAAGCAGAATTCTGCTACACCATCAGACACGACGGCAGCAT  
 CGGCGTGGAGAAGATCATCAACGCCAGAAAGACCAAGACCAACGCCAAGGTGATCAAGGTGACCCTGGACAACGGCG  
 AGAGCATCATCTGCACCCCCGACCACAAGTTCATGCTGAGAGACGGCAGCTACAAGTGCGCCATGGACCTGACCCTG  
 GACGACAGCCTGATGCCCTGCACAGAAAGATCAGCACCACCGAGGACAGCGCCACATGGAGGCCGTGCTGAACTA  
 CAACCACAGAATCGTGAACATCGAGGCCGTGAGCGAGACCATCGACGTGTACGACATCGAGGTGCCCCACACCACA  
 ACTTCGCCCTGGCCAGCCACCATCACCATCACCATGGCTGGAGCCACCCCAAGTTCGAGAAGTAG

[0276] SEQ ID NO 17可翻译为如下氨基酸序列(SEQ ID NO 18,N-内含肽结构域中的氨基酸以下划线表示,6xHis标记和strepII标记以阴影表示):

[0277] MNFGLRLIFLVLTLKGVQCQVQLQESGPGLVKPSQTLSTCTVSGGSIISTSGMGVWIRQHPGKGLEW  
 IGHIWDDDKRYNPALKSRVTISVDTSKNQFSLKLSVTAADTAVYYCARMELWSYFYFDYWGQGLTVTVSSASTKG  
 PSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSVHTFPAVLQSSGLYSLSVVTVPSSSLGTQTY  
 ICNVNHNKPSNTKVDKKEPKSCDKHTHTCPPCPAPELGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK  
 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVY  
 LPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSGDSFFLYSKLTVDKSRWQQGNVFS  
 CSVMHEALHNHYTQKSLSLSPGKCFSGDTLVALTDGRSVSFEQLVEEEKQKQNFICYTIRHDGSIGVEKIIINARKTKT  
NAKVIKVTLDNGESIICTPDHKFMLRDGSYKCAMDLTDDSLMPLHRKISTTEDSGHMEAVLNYNHRIVNIEAVSE  
TIDVYDIEVPHTHNFALASHHHHHHGWSHPQFEK •

[0278] 类似地,可以设计带有包含Ssp GyrB 11分裂内含肽的N-内含肽结构域、6xHis标记和strepII标记的C端延长段的抗人CD19抗体hBU12的完整IgG1k-轻链编码区域,如下SEQ ID NO 19所示:

[0279] ATGAATTTTGGACTGAGGCTGATTTTCTGGTGCTGACCCTGAAAGGCGTCCAGTGTGACATTGTGCTG  
 ACCCAATCTCCAGCTTCTTTGGCTGTGTCTCTAGGGCAGAGGGCCACCATCTCCTGCAAGGCCAGCCAAAGTGTGA  
 TTTTGATGGTGATAGTTATATGAACTGGTACCAACAGAAACCAGGACAGCCACCCAAAGTCCCTCATCTATGCTGCAT  
 CCAATCTAGAATCTGGGATCCCAGCCAGGTTTAGTGGCAGTGGGTCTGGGACAGACTTCACCCTCAACATCCATCCT  
 GTGGAGGAGGAGGATGCTGCAACCTATTACTGTCAGCAAAGTAATGAGGATCCGTGGACGTTCCGGTGGAGGCACCAA  
 GCTGGAAATCAAACGTACGGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAA

CTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGAAGGTGGATAACGCCCTC  
 CAATCGGGTAACTCCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCTGAC  
 GCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTACCCATCAGGGCCTGAGCTCGCCCGTCA  
 CAAAGAGCTTCAACAGGGGAGAGTGCTTCAGCGGGACACCCTGGTGGCCCTGACCGACGGCAGAAGCGTGAGCTTC  
 GAGCAGCTGGTGGAGGAGAGAAGCAGGGCAAGCAGAACTTCTGCTACACCATCAGACACGACGGCAGCATCGGCGT  
 GGAGAAGATCATCAACGCCAGAAAGACCAAGACCAACGCCAAGGTGATCAAGGTGACCCTGGACAACGGCGAGAGCA  
 TCATCTGCACCCCCGACCACAAGTTCATGCTGAGAGACGGCAGCTACAAGTGCGCCATGGACCTGACCCTGGACGAC  
 AGCCTGATGCCCCTGCACAGAAAGATCAGCACCACCGAGGACAGCGGCCACATGGAGGCCGTGCTGAACTACAACCA  
 CAGAATCGTGAACATCGAGGCCGTGAGCGAGACCATCGACGTGTACGACATCGAGGTGCCCCACACCCACAACCTTCG  
 CCCTGGCCAGCCACCATCACCATCACCATGGCTGGAGCCACCCCCAGTTCGAGAAGTAG

[0280] SEQ ID NO 19可翻译为如下氨基酸序列(SEQ ID NO 20,N-内含肽结构域中的氨基酸以下划线表示,6xHis标记和strepII标记以阴影表示):

[0281] MNFGLRLIFLVLTLKGVQCDIVLTQSPASLAVSLGQRATISCKASQSVDFDGSYMNWYQQKPGQPPK  
 VLIYAASNLESGIPARFSGSGSDFTLNIHPVEEEDAATYYCQQSNEDPWTFGGGTKLEIKRTVAAPSVFIFPPS  
 DEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKSTYLSSTLTLSKADYEKHKVYACEVT  
 HQGLSSPVTKSFNRGECFSGDTLVALTDGRSVSFEQLVEEEKQKQNFICYTIRHDGSIGVEKIINARKTKTNAKVI  
 KVTLDNGESI ICTPDHKFMLRDGSYKAMDLLDSDLMLHRKISTTEDSGHMEAVLNYNHRIVNIEAVSETIDVY  
 DIEVPHTHNFALASHHHHHHGWSHPQFEK •

[0282] 如SEQ ID NO 17和19所公开的分别编码抗人CD19特异抗体的N-内含肽修饰的重链和轻链的编码区域,可以合成侧翼限制性酶位点(例如HindIII和NotI),由此可通过本领域已知的标准分子生物学方法,将其克隆到标准哺乳类表达载体中,例如pCDNA3.1-hygro(+)(Invitrogen)。

[0283] 用于N-内含肽标签hBU12抗人CD19抗体的pCDNA3.1-hygro(+)-IgH链表达载体的完整DNA序列如下所示:

[0284] SEQ ID NO 21(带有C端Ssp GyrB 11分裂内含肽N-内含肽结构域、6xHis标记和strepII标记(下划线)以及HindIII和NotI克隆位点(阴影显示)的抗体hBU12人IgG1V<sub>H</sub>-C<sub>H</sub>重链编码区域):

[0285] GACGGATCGGGAGATCTCCCGATCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGT  
 TAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGC CGGAGCAAAATTTAAGCTACAACAAGG  
 CAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGCCAG  
 ATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA  
 TGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCCCGCCATTGACGTCA  
 ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAAC  
 TGCCCACTTGGCAGTACATCAAGTGATCATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCG  
 CCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT  
 ACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTACTCACGGGATTTCCAAGTCTCCA  
 CCCCATTGACGTCAATGGGAGTTTGTGTTTGGCACAAAATCAACGGGACTTCCAAAATGTCGTAACAACCTCCGCC  
 CATTGACGCAAATGGGCGGTAGGCGGTACGGTGGGAGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCC  
 ACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTT

CCATGAATTTTGGACTGAGGCTGATTTTCCTGGTGCTGACCCTGAAAGGCGTCCAGTGTGAGTTTCAGCTGCAAGAG  
TCTGGCCCTGGGTTGGTTAAGCCCTCCCAGACCCTCAGTCTGACTTGTACTGTGTCTGGGGTTCAATCAGCACTTC  
TGGTATGGGTGTAGGCTGGATTAGGCAGCACCCAGGGAAGGGTCTGGAGTGGATTGGACACATTTGGTGGGATGATG  
ACAAGAGATATAACCCAGCCCTGAAGAGCAGAGTGACAATCTCTGTGGATACCTCCAAGAACCAGTTTAGCCTCAAG  
CTGTCCAGTGTGACAGCTGCAGATACTGTGTCTACTACTGTGCTAGAATGGAACCTTGGTCCTACTATTTTGACTA  
CTGGGGCCAAGGCACCCTTGTACAGTCTCCTCAGCTAGCACCAAGGGCCCATCTGTCTTCCCCCTGGCACCCCTCT  
CCAAGAGCACCTCTGGGGGCACAGCTGCCCTGGGCTGCCCTGGTCAAGGACTACTTCCCTGAACCTGTGACAGTGTCC  
TGGAACTCAGGCGCCCTGACCAGCGGCTGCACACCTTCCCGGCTGTCTACAGTCTCAGGACTCTACTCCCTCAG  
CAGCGTGGTGACCGTGCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACA  
CCAAGGTGGACAAGAAAGTTGAGCCAAATCTTGTGACAAAACCTCACACATGCCACCCTGCCAGCACCTGAACTC  
CTGGGGGGACCGTCACTCTTCTTCCCCCAAAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTAC  
ATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATA  
ATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCTCACCCTGCTGCACCAG  
GACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAACCATCTC  
CAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCATCCCGGGATGAGCTGACCAAGAACCAGG  
TCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAG  
AACAACCTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAA  
GAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACACAGAAGA  
GCCTCTCCCTGTCTCCGGGTAAATGCTTACGCGGGACACCCTGGTGGCCCTGACCGACGGCAGAAGCGTGAGCTTC  
GAGCAGCTGGTGGAGGAGGAGAAGCAGGGCAAGCAGAACTTCTGCTACACCATCAGACACGACGGCAGCATCGGCGT  
GGAGAAGATCATCAACGCCAGAAAGACCAAGACCAACGCCAAGGTGATCAAGGTGACCCTGGACAACGGCGAGAGCA  
TCATCTGCACCCCCGACCACAAGTTCATGCTGAGAGACGGCAGCTACAAGTGCGCCATGGACCTGACCCTGGACGAC  
AGCCTGATGCCCCTGCACAGAAAGATCAGCACACCAGGACAGCGCCACATGGAGGCCGTGCTGAACTACAACCA  
CAGAATCGTGAACATCGAGGCCGTGAGCGAGACCATCGACGTGTACGACATCGAGGTGCCCCACACCCACAACCTCG  
CCCTGGCCAGCCACCATCACCATCACCATGGCTGGAGCCACCCCAAGTTCGAGAAGTAGGCGGCCGCTCGAGTCTAG  
AGGGCCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCCTTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCG  
TGCCTTCTTGACCCTGGAAGGTGCCACTCCCCTGTCCCTTCCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTG  
AGTAGGTGTCAATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA  
TGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGC  
CCTGTAGCGGCGATTAAGCGCGGGGTTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCG  
CCCCTCCTTTGCTTTCTTCCCTTCTTCTCGCCACGTTCCGCGGCTTCCCCGTCAAGCTCTAAATCGGGGCAT  
CCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACCTGATTAGGGTGTGGTTACAGTAGTG  
GGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTC  
ACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTATTATAAGGGATTTTGGGGATTTCCGGCCTATTGGTT  
AAAAAATGAGCTGATTTAACAATAAATTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTC  
CCCAGGCTCCCCAGGCAGGAGGATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCA  
GGCTCCCCAGCAGGAGGAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCAGCCCTAACTCCGCC  
CATCCCGCCCTAACTCCGCCCAGTTCGCCCATTCTCCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGG  
CCGAGGCCGCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCCTAGGCTTTTGCAAAAAG



CTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACC GCGACGTCTGT  
CGAGAAGTTTCTGATCGAAAAGTTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTT  
TCAGCTTCGATGTAGGAGGGCGTGGATATGTCCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTAT  
GTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGAC  
CTATTGCATCTCCCGCCGTGCACAGGGTGCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGC  
CGGTTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCCGACCGCAA  
GGAATCGGTCAATACACTACATGGCGTGATTTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACCTGT  
GATGGACGACACCGTCAGTGCGTCCGTGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCGAAG  
TCCGGCACCTCGTGACGCGGATTTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGAC  
TGGAGCGAGGCGATGTTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTAT  
GGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGACAGGATCGCCGCGGCTCCGGGCGTATATGCTCC  
GCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGC  
GACGCAATCGTCCGATCCGGAGCCGGGACTGTGCGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGA  
TGGCTGTGTAGAAGTACTCGCCGATAGTGAAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGC  
TACGAGATTTTCGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACCGCGGCTGGATG  
ATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCAACTTGTATTGTCAGCTTATAATGGTTACAA  
ATAAAGCAATAGCATCACAAATTTCAAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTTGTCCAAACTCA  
TCAATGTATCTTATCATGTCTGTATAACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCT  
GTGTGAAATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTA  
ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC  
ATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTCCGCTTCTCGCTCACTGACTCG  
CTGCGCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGG  
GGATAACGCAGGAAAGAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT  
TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGA  
CTATAAAGATAACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA  
CCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG  
TCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCGACCGCTGCGCCTTATCCGTAACCTATCGT  
CTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA  
TGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCG  
CTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT  
GGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG  
GTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGA  
TCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGC  
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGCTCGTGTAGAT  
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAG  
ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCTCCATCCAG  
TCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTAC  
AGGCATCGTGGTGTACGCTCGTTCGTTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACAT  
GATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCCGATCGTTGTCAGAAGTAAGTTGGCCGCGAGT

TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG  
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGGCGACCGAGTTGCTCTTGCCCCGGCGTCAATACGGGATA  
ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATC  
TTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTTACCAG  
CGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC  
TCATACTCTTCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGT  
ATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

[0286] 用于带Ssp GyrB 11分裂内含肽的N-内含肽结构域标记的hBU12抗人CD19抗体的pCDNA3.1-hygro (+) IgL链表达载体的完整DNA序列如下：

[0287] SEQ ID NO 22 (带有C端Ssp GyrB 11分裂内含肽N-内含肽结构域、6xHis标记和strepII标记(下划线)以及HindIII和NotI克隆位点(阴影显示)的抗体hBU12人IgG1V<sub>L</sub>-C<sub>L</sub> k轻链编码区域)：

[0288] GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGT  
TAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGC GCGAGCAAAATTTAAGCTACAACAAGG  
CAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAG  
ATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA  
TGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCCGCCATTGACGTCA  
ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAAC  
TGCCCACTTGGCAGTACATCAAGTGATCATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCG  
CCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT  
ACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTACTCACGGGATTTCCAAGTCTCCA  
CCCCATTGACGTCAATGGGAGTTTGTGGTGGCACCAAAATCAACGGGACTTTCAAAATGTGTAACAACCTCCGCC  
CATTGACGCAAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCC  
ACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTT  
CCATGAATTTTGGACTGAGGCTGATTTTCTGGTGCTGACCCTGAAAGGCGTCCAGTGTGACATTGTGCTGACCCAA  
TCTCCAGCTTCTTTGGCTGTGTCTCTAGGGCAGAGGGCCACCATCTCTGCAAGGCCAGCCAAAGTGTTGATTTTGA  
TGGTGATAGTTATATGAACTGGTACCAACAGAAACCAGGACAGCCACCCAAAGTCCTCATCTATGCTGCATCCAATC  
TAGAATCTGGGATCCAGCCAGGTTTGTAGTGGCAGTGGGTCTGGGACAGACTTCACCCTCAACATCCATCCTGTGGAG  
GAGGAGGATGCTGCAACCTATTACTGTCAGCAAAGTAATGAGGATCCGTGGACGTTCCGGTGGAGGCACCAAGCTGGA  
AATCAAACGTACGGTGGCTGCACCATCTGTCTTTCATCTTCCC GCCATCTGATGAGCAGTTGAAATCTGGAAGTGCCT  
CTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCG  
GGTAACTCCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCTGACGCTGAG  
CAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTACCCATCAGGGCCTGAGCTCGCCCGTCACAAAGA  
GCTTCAACAGGGGAGAGTGCTTACGCGGCACACCCTGGTGGCCCTGACCGACGGCAGAAGCGTGAGCTTCGAGCAG  
CTGGTGGAGGAGGAGAAGCAGGGCAAGCAGAACTTCTGCTACACCATCAGACACGACGGCAGCATCGGCGTGGAGAA  
GATCATCAACGCCAGAAAGACCAAGACCAACGCCAAGGTGATCAAGGTGACCCTGGACAACGGCGAGAGCATCATCT  
GCACCCCGACCACAAGTTCATGCTGAGAGACGGCAGCTACAAGTGCGCCATGGACCTGACCCTGGACGACAGCCTG  
ATGCCCTGCACAGAAAGATCAGCACCACCGAGGACAGCGGCCACATGGAGGCCGTGCTGAACTACAACCACAGAAT  
CGTGAACATCGAGGCCGTGAGCGAGACCATCGACGTGTACGACATCGAGGTGCCCCACACCACAACCTTCGCCCTGG

CCAGCCACCATCACCATCACCATGGCTGGAGCCACCCCCAGTTCGAGAAGTAGGCCGGCCGCTCGAGTCTAGAGGGCC  
CGTTTAAACCCGCTGATCAGCCTCGACTGTGCCCTTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCGTGCCTT  
CCTTGACCCTGGAAGGTGCCACTCCCACTGTCCCTTCCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGG  
TGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGG  
GGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTA  
GCGGCGCATTAAAGCGCGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCT  
CCTTTTCGCTTTCTCCCTTCTTTCTCGCCACGTTCCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTT  
AGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCAT  
CGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCAAACTGGA  
ACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGGGGATTTCCGGCCTATTGGTTAAAAAA  
TGAGCTGATTTAAACAAAAATTTAACGCAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGG  
CTCCCCAGGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCC  
CCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCC  
GCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGG  
CCGCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGAAAAAGCTCCCG  
GGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTGAGAA  
GTTTCTGATCGAAAAGTTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTAGCT  
TCGATGTAGGAGGGCGTGGATATGTCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTAT  
CGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTG  
CATCTCCCGCCGTGCACAGGGTGTACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCCG  
CGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATC  
GGTCAATACTACTACATGGCGTGATTTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACCTGTGATGGA  
CGACACCGTCAGTGGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGC  
ACCTCGTGACGCGGATTTCCGGCTCCAACAATGTCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGC  
GAGGCGATGTTCCGGGATTCCCAATACGAGGTGCGCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCA  
GCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTG  
GTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTTCGATGATGCAGCTTGGGCGCAGGGTTCGATGCGACGCA  
ATCGTCCGATCCGGAGCCGGGACTGTGCGGGCTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTG  
TGTAGAAGTACTCGCCGATAGTGGAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAG  
ATTTTCGATTCACCGCCGCTTCTATGAAAGGTGGGCTTCGGAATCGTTTTCCGGGACCGCGGCTGGATGATCCTC  
CAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCAACTGTTTTATTGCAGCTTATAATGGTTACAAATAAAG  
CAATAGCATCACAATTTACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATG  
TATCTTATCATGTCTGTATAACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGA  
AATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGT  
GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT  
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGATTTGGGCGCTTCCCGCTTCTCGCTCACTGACTCGCTGCGC  
TCGGTCGTTCCGGTTCGCGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAA  
CGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGGTTGCTGGCGTTTTTCC  
ATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA

AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC  
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTT  
GCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAG  
TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG  
CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC  
TGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT  
TTTGTTCGAAAGCAGGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGA  
CGCTCAGTGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTT  
TAAATTAATAATGAAGTTTTAAATCAATCTAAAATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATC  
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTAC  
GATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGGAGACCCACGCTCACCAGGCTCCAGATTTAT  
CAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATT  
AATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCACAACGTTGTTGCCATTGCTACAGGCAT  
CGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATGATCCC  
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTATCA  
CTCATGGTTATGGCAGCACTGCATAATTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA  
CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCG  
CGCCACATAGCAGAACTTTAAAAGTGTCTCATCATTGAAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCG  
CTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTACCCAGCGTTTC  
TGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGAAAAAAGGGAATAAGGGCGACACGAAATGTTGAATACTCATA  
TCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG  
AAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

[0289] SEQ ID NO 21和22所述的pcDNA3.1-hygro (+) 表达载体能够通过转染进入哺乳类细胞,例如但不限于,通常用于重组抗体表达、带有添加至IgH和IgL链C端的分选酶A标记、6xHis标记、Myc标记和strepII的抗人CD19特异性人源化抗体hBU12的CHO细胞。

[0290] 实施例3:重组金黄葡萄球菌分选酶A的克隆与表达

[0291] 金黄葡萄球菌分选酶A的ORF发表于Genbank数据库中,可以索引号AF162687.1查找。该记录下的氨基酸序列显示为SEQ ID NO 23 (金黄葡萄球菌分选酶A的氨基酸序列)。

[0292] MKKWTNRLMTIAGVVLILVAAYLFAKPHIDNYLHDKDKDEKIEQYDKNVKEQASKDKKQAKPQIPKDK  
SKVAGYIEIPDADIKEPVYPGPATPEQLNRGVSFAEENESLDDQNIS IAGHTF IDRPNYQFTNLKAAKKGSMVYFKV  
GNETRKYKMTSIRDVKPTDVGVLDEQKGDQKQLTLITCDDYNEKTGVWEKRKIFVATEVK

[0293] 该Genbank索引对应序列的核苷酸序列如SEQ ID NO 24所示:

[0294] ATGAAAAATGGACAAATCGATTAATGACAAATCGCTGGTGTGGTACTTATCCTAGTGGCAGCATATTTG  
TTTGCTAAACCACATATCGATAATTATCTTCACGATAAAGATAAAGATGAAAAGATTGAACAATATGATAAAAATGT  
AAAAGAACAGGCGAGTAAAGATAAAAAGCAGCAAGCTAAACCTCAAATTCGAAAGATAAATCGAAAGTGGCAGGCT  
ATATTGAAATTCAGATGCTGATATTAAGAACCAGTATATCCAGGACCAGCAACACCTGAACAATTAATAGAGGT  
GTAAGCTTTGCAGAAGAAAATGAATCACTAGATGATCAAAAATTTCAATTGCAGGACACACTTTCATTGACCGTCC  
GAACTATCAATTTACAAATCTTAAAGCAGCCAAAAAAGGTAGTATGGTGTACTTTAAAGTTGGTAATGAAACACGTA  
AGTATAAAATGACAAGTATAAGAGATGTTAAGCCTACAGATGTAGGAGTTCTAGATGAACAAAAAGGTAAAGATAAA

CAATTAACATTAATTACTTGTGATGATTACAATGAAAAGACAGGCGTTTGGGAAAAACGTAAAATCTTTGTAGCTAC  
AGAAGTCAAATAA

[0295] 重组分选酶A酶促活性片段在大肠杆菌中的表达的相关技术信息,包括携带6xHis标记的氨基酸60-205(公开于参考文献W02007/108013A2中)。携带6xHis标记的金黄葡萄球菌分选酶A(aa60-205)的编码区域如下SEQ ID NO 25所示:

[0296] ATGCAAGCTAAACCTCAAATTCGAAAAGATAAAATCGAAAAGTGGCAGGCTATATTGAAATTCAGATGCT  
GATATTAAGAACCAGTATATCCAGGACCAGCAACACCTGAACAATTAATAGAGGTGTAAGCTTTGCAGAAGAAAA  
TGAATCACTAGATGATCAAAAATATTTCAATTGCAGGACACACTTTCATTGACCGTCCGAACTATCAATTTACAAATC  
TTAAAGCAGCCAAAAAGGTAGTATGGTGTACTTTAAAGTTGGTAATGAAACACGTAAGTATAAAATGACAAGTATA  
AGAGATGTAAAGCCTACAGATGTAGGAGTTCTAGATGAACAAAAAGGTAAAGATAAACAATTAACATTAATTACTTG  
TGATGATTACAATGAAAAGACAGGCGTTTGGGAAAAACGTAAAATCTTTGTAGCTACAGAAGTCAAACACCATCACC  
ATCACCATTAA

[0297] SEQ ID NO 25翻译为氨基酸序列SEQ ID NO 26:

[0298] MQAKPQIPKDKSKVAGYIEIPDADIKEPVYGPATPEQLNRGVSF AEENESLDDQNISIAGHTFIDRP  
NYQFTNLKAAKKGSMVYFKVGNETRKYKMTSIRDVKPTDVGLVLEQKQKDKQLTLITCDDYNEKTGVWEKRKIFVA  
TEVKHHHHHH •

[0299] SEQ ID NO 25所示的6xHis标记的金黄色葡萄球菌分选酶A片段的编码区域,可以被克隆到标准细菌表达载体中,例如像pET29(Novagen)中,以转化大肠杆菌菌株BL21(DE3)(Novagen)中,并根据本领域中已知的标准方法生成可以用于细菌生产重组分选酶A的大肠杆菌克隆。总之,分选酶A pET29表达质粒转化的大肠杆菌BL21(DE3)可以在37°C下以LB培养基与50μg/mL卡那霉素培养,直到OD<sub>600</sub>为0.5-0.8为止。然后可加入IPTG至0.4mM终浓度,并在30°C下诱导蛋白表达三小时。然后通过离心收获细胞,重悬于裂解缓冲液(50mM Tris,pH值8.0,300mM NaCl,补充有1mM的氯化镁,2单位/mL DNA酶I(NEB),260nM纳米抑肽酶,1.2μM亮肽素和1mM PMSF)。然后通过超声处理来裂解细胞,并可以按照制造商的说明进行操作,用Ni-NTA琼脂糖纯化澄清的上清液。

[0300] 根据SDS-PAGE判断纯度>90%的部分可以合并并用Tris缓冲盐水(25mM Tris pH值7.5,150mM NaCl)透析,并使用已发表的消光系数17,420M<sup>-1</sup>cm<sup>-1</sup>从A<sub>280</sub>测量值计算酶浓度。根据上述实验方案,制备ca.20mg>90%纯度的金黄色葡萄球菌分选酶A重组酶促活性片段(ca.17kD),重组蛋白的SDS-PAGE和免疫印迹分析如图7所示。

[0301] 实施例4:分选酶标记或N-内含肽标记的重组抗体在CHO细胞中的表达和纯化

[0302] a) CHO细胞表达:实施例2和3中公开了,从表达结构表达重组IgG1抗体,可以使用例如市售的CHO表达体系通过瞬时转染来实现,如使用购自Invitrogen的FreeStyle CHO体系,根据FreeStyle CHO厂商说明的指示进行。

[0303] 简言之,转染前约1天,CHO细胞应在摇瓶中的FreeStyle CHO培养基中以5-6×10<sup>6</sup>个细胞/ml接种,令其在加湿的接种孵化器中,轨道摇动器120rpm、37°C和7.5%二氧化碳氛围的条件下扩增。次日,当所述细胞到达1.2-1.5×10<sup>6</sup>/ml密度时,可进行转染。然后根据需要将细胞稀释到1×10<sup>6</sup>细胞/ml。随后需要取30ml所述细胞悬浮液,加入到125ml摇瓶中,并向600μl OptiPro SF-培养基(Invitrogen)中加入40μg 1:1混合的IgH和IgL表达质粒DNA。同时,需要向600μl OptiPro SF-培养基(Invitrogen)中加入40μl FreeStyle MAX转染试

剂,轻轻混合两个试样,并在室温下培养10分钟,以允许形成DNA-转染试剂复合物。然后,可以从上方将DNA-转染试剂混合物缓慢加入到125ml CHO细胞培养基中,随后令转染细胞在轨道摇动器120rpm、37°C和7.5%二氧化碳氛围的条件下生长6天。

[0304] 此后,可收集细胞培养物上清液并通过本领域(ELISA,Luminex公司等)已知适当方法进行抗体表达滴度分析。

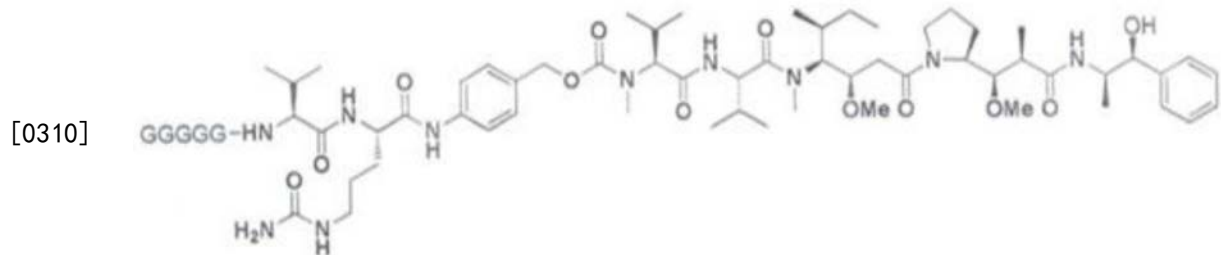
[0305] b) 蛋白A纯化:可使用市售蛋白A琼脂糖柱(Thermo Fisher,Pierce),按照厂商说明,从CHO细胞上清液进行重组抗体的蛋白A纯化。

[0306] 简言之,取澄清的细胞培养上清液,过适当大小和容量并用PBS平衡的蛋白A柱。残留培养基用PBS洗涤,且最终结合的IgG可以用低pH缓冲液洗脱,如pH为3.0的0.1M柠檬酸-氢氧化钠。洗脱的IgG应立即用1/10th体积的pH7.41M Tris/Cl中和。含IgG的合并组分随后可以在4°C下以PBS透析过夜。

[0307] 实施例4中提供的实验方案向本领域技术人员提供了从实施例1和2中公开的结构中制备足量的纯化重组抗体的方法。

[0308] 实施例5:通过分选酶和分裂内含肽介导的转肽作用生成位点特异性C端MMAE毒性效应分子结合单克隆抗体

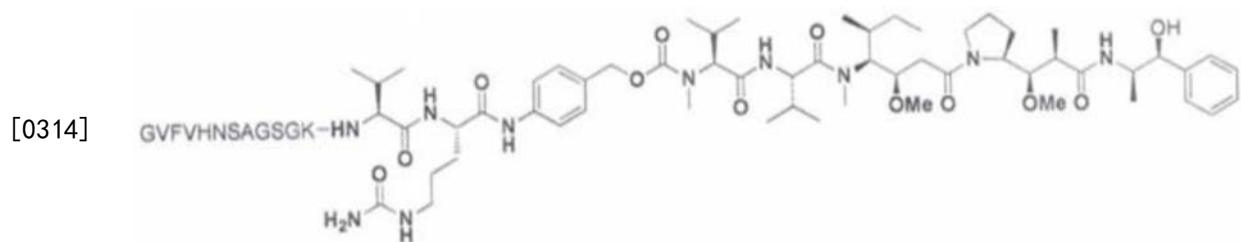
[0309] 如下式所示的,连接到长5个氨基酸的甘氨酸短链和长6个氨基酸的SSP GyrB S11C-内含肽分裂内含肽的单甲基阿里他汀毒素,可以从合格的化学CRO公司定制订购。



[0311] 式1:以vcPAB连接子进行5-甘氨酸修饰的MMAE:

[0312] Me=甲基(-CH<sub>3</sub>)基团

[0313] G=甘氨酸残基



[0315] 式2:携带vcPAB连接子以6氨基酸C-内含肽结构域GVFVHN和6氨基酸SAGSGK C-外显肽修饰的MMAE:

[0316] Me=甲基(-CH<sub>3</sub>)基团

[0317] GVFVHNSAGSGK=Gly-Val-Phe-Val-His-Asn-Ser-Ala-Gly-Ser-Gly-Lys短链

[0318] a) LPETG分选酶A基序标记的重组IgG抗体的毒性MMAE效应分子结合反应

[0319] 五甘氨酸修饰的MMAE毒性效应分子与LPETG分选酶A标记IgG1抗体(可以根据实施例1和4制备)的结合,可以用生理培养缓冲液(例如5mMTris/Cl,15mM NaCl,6mM CaCl<sub>2</sub>,pH 8.0),通过混合适当比例的LPETG标记的IgG1抗体和式1公开的甘氨酸修饰的MMAE毒素(例

如,1:1比例和50 $\mu$ M的浓度)和重组分选酶A(根据实施例3制备)(例如,5 $\mu$ M浓度),并在37 $^{\circ}$ C至40 $^{\circ}$ C下培养至少2小时。

[0320] 可以在停止反应后,用抗His标记和/或抗strepII标记抗体通过免疫印迹分析或ELISA分析6xHis标记和/或strepII标记来监测结合效率。

[0321] 可以通过镍-NTA柱或streptactin柱结合来富集完全结合产物,,所述柱分别结合6xHis标记或strepII标记,所述标记仅可存在于不完全反应的IgG1底物中。最终的IgG-效应分子结合物最后可以如上所述使用蛋白A纯化来进行纯化。

[0322] b) SSP GyrB S11N-内含肽标记的重组IgG抗体与的毒性MMAE效应分子的结合

[0323] SSP GyrB S11C-内含肽氨基酸修饰的MMAE毒性效应分子与N-内含肽标记的IgG1抗体(可以通过以下实施例2和4制备)的结合,可用生理培养缓冲液(20mM Tris/Cl,250mM NaCl,1mM EDTA,pH 8.5),混合适当比例的N-内含肽标记的IgG1抗体与式2公开的C-内含肽氨基酸修饰的MMAE毒素(例如,比例为1:10或1:25,IgG抗体浓度为5 $\mu$ M),并在室温下或在37 $^{\circ}$ C下培养至少4个小时。

[0324] 可以在停止反应后,用抗His标记和/或抗strepII标记抗体通过免疫印迹分析或ELISA分析6xHis标记和/或strepII标记来监测结合效率。

[0325] 可以通过镍-NTA柱或streptactin柱结合来富集完全结合产物,,所述柱分别结合6xHis标记或strepII标记,所述标记仅可存在于不完全反应的IgG1底物中。最终的IgG-效应分子结合物最后可以如上所述使用蛋白A纯化来进行纯化。

[0326] 总之,以上公开的实施例1-5允许本领域技术人员能够实践本发明中所述的,通过分选酶A介导或裂内含肽介导的转肽作用,进行毒性效应分子与C端的位点特异性酶促结合。

[0327] 实施例6:制备在轻链或重链上携带C端GS(甘氨酸-丝氨酸)连接子、LPETG分选酶基序和附加的6x-His和Strep II亲和纯化标记的曲妥单抗

[0328] 基本根据实施例1所述,生成编码单克隆抗体曲妥单抗(Tras)重链和轻链的抗体表达结构,其中抗体无标记(SEQ ID NOs:31-34)或C端携带GS(甘氨酸-丝氨酸)连接子、LPETG分选酶基、和6x-His和Strep II标记(SEQ ID NOs:35-38)。使用上述表达结构,通过与相应表达结构的共转染,在CHO细胞中生产Tras-HC-GS-LHS和Tras-LC-GS-LHS(HC=重链,LC=轻链,GS=甘氨酸-丝氨酸,LHS=LPETG标记+6xHis标记+strepII标记)。Tras-HC-GS-LHS是轻链未修饰(SEQ ID NO:35-36)且重链C端具有GS(甘氨酸-丝氨酸)连接子、LPETG分选酶基序、6xHis标记和strepII标记(SEQ ID NO:33-34)的曲妥单抗变体。Tras-LC-GS-LHS是重链未修饰(SEQ ID NOs:31-32)且轻链C端具有GS(甘氨酸-丝氨酸)连接子、LPETG分选酶基序、6xHis标记和strepII标记(SEQ ID NOs:37-38)的曲妥单抗变体。CHO细胞的转染和通过蛋白A琼脂糖层析法对抗体的亲和纯化,基本如实施例4所述。

[0329] 实施例7:曲妥单抗轻链或重链与Gly5修饰DM1毒素的分选酶A介导结合

[0330] 在含每种单克隆抗体(mAb)各10.5mg的1x分选酶缓冲液(25mM Tris-HCl,pH8.2;150mM NaCl;7.5mM CaCl<sub>2</sub>)中,进行包含Gly<sub>5</sub>-修饰的DM1毒素(购自Concortis,San Diego,CA,U.S.,结构如图14a所示)和金黄葡萄球菌的17kD重组分选酶片段(参见实施例3)的结合反应,如下表II所示。Tras-HC-GS-LHS在25 $^{\circ}$ C下培养2小时;Tras-LC-GS-LHS结合反应在25 $^{\circ}$ C下培养18小时。每种反应混合物随后通过Strep-Tactin<sup>®</sup>琼脂糖柱(IBA Life-Sciences,

Göttingen, Germany)。为此,取1ml Strep-Tactin琼脂糖通过重力装入多孔柱中,并以2柱体积的平衡缓冲液(100mM Tris-HCl, pH 8.0; 150mM NaCl; 1mM EDTA)进行平衡。每种结合混合物利用重力流经过同一柱两次(以增加树脂上的停留时间)。然后用又一柱体积的平衡缓冲液洗涤树脂,以最大化结合物产率,然后将合并组分立即装入蛋白A柱。为此,使用10柱体积缓冲液(25mM磷酸钠, pH 7.5)来平衡1ml蛋白A HiTrap柱。令每种结合反应物通过平衡后的柱,然后用5柱体积的缓冲液洗涤柱。使用5柱体积的洗脱液(0.1M琥珀酸, pH 2.8)来洗脱结合的结合物,收集其中1柱体积部分(收集在含有25%v/v 1M Tris Base的试管中以中和酸)并分析其蛋白含量。将含蛋白质组分合并,并通过G25柱色谱法进行制剂。为此,使用适用于每一种生产规模NAP 25柱,来对结合物制剂以供长期存储。平衡所述柱,装载,并以根据厂商说明,以10mM pH 5.0琥珀酸钠, 100mg/mL海藻糖, 0.1%w/v聚山梨醇酯20 (Kadcyla<sup>®</sup> (T-DM1)的制剂缓冲液,由Roche/Genentech销售)洗脱。

[0331] Tras-HC-GS-LHS和Tras-LC-GS-LHS的DM1结合体产率分别为8.0mg (76.2%) 和5.9mg (56.2%)。主要的过程损失发生在蛋白A和G25纯化中,大多因为为了保证用于每个后续步骤或存储的最大产物浓度而进行的削峰。

反应成分	HC	LC	终浓度
Tras-HC-GS-LHS (5.3mg/ml)	1981 $\mu$ l	-	5 $\mu$ M
Tras-LC-GS-LHS (5.5mg/ml)	-	1911 $\mu$ l	5 $\mu$ M
H <sub>2</sub> O	7775.25 $\mu$ l	7714 $\mu$ l	-
[0332] Gly <sub>5</sub> -DM1 (1mM)	1400 $\mu$ l	1400 $\mu$ l	100 $\mu$ M
分选酶 A (0.85mg/ml = ca. 50 $\mu$ M)	43.75 $\mu$ l	175 $\mu$ l	0.156/0.625 $\mu$ M
5x 分选酶缓冲液*	2800 $\mu$ l	2800 $\mu$ l	1x

[0333] 表2: Tras-HC-GS-LHS和Tras-LC-GS-LHS的结合条件

[0334] 通过疏水作用色谱(HIC)分析药物负载,其中使用了TOSOH Butyl-NPR4.6mm x 3.5cm, 2.5 $\mu$ m柱(流速0.8mL/min, A-1.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 25mM NaPi, pH=6.95 $\pm$ 0.05和B-75% 25mM NaPi, pH=6.95 $\pm$ 0.05, 25% IPA之间的12分钟线性梯度)。HIC数据表明,对于Tras-



HC-GS-LHS和Tras-LC-GS-LHS而言,均无可检测的剩余未结合单克隆抗体,且每种单克隆抗体中的大部分均负载了2个药物(参见图8)。

[0335] 实施例8:分选酶A介导的曲妥单抗-DM1结合物的体外毒性测试

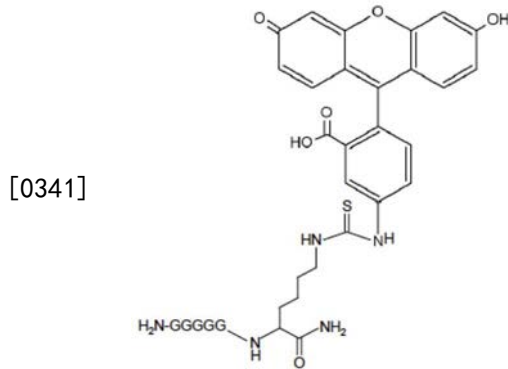
[0336] 下面研究DM1-分选酶A介导结合Tras-HC-GS-LHS和DM1-分选酶A介导结合Tras-LC-GS-LHS的毒性,并使用SKBR3细胞和T47D-5R细胞与**Kadcyla**<sup>®</sup> (Roche/Genentech) 进行比较;SKBR3细胞是一种人乳腺癌细胞系,其过度表达曲妥单抗同源抗原(Tras)HER-2/neu;T47D-5R细胞是一种乳腺癌细胞系,其对HER-2/neu的天然表达水平低,并改造为缺失细胞表面HER-2/neu (Graus-Porta et al. (1995)。将细胞接种在96孔板中的100 $\mu$ l完全DMEM培养基中(每孔10000个细胞)。培养一天后,小心地从每个孔中移除50 $\mu$ l的培养基,,并替换50 $\mu$ L以完全DMEM培养基稀释的每种ADC的3.5倍系列稀释液,得到的ADC浓度范围为20 $\mu$ g/ml至0.25ng/ml。每种稀释液一式两份或一式三份。在加湿培养箱中以37 $^{\circ}$ C、5%CO<sub>2</sub>额外培养3天后,将板从培养箱中取出,并平衡至室温。约30分钟后,每孔中加入100 $\mu$ l **CellTiter-Glo**<sup>®</sup> 发光溶液(Promega公司,Cat.No G7570),以450rpm摇动5分钟后,不摇动地培养10分钟,然后使用Tecan Infinity F200以每孔1秒的积分时间测量发光。所有三种ADC对过量表达HER-2/neu的SKBR3乳腺癌细胞系均具有高度细胞毒性,但对HER-2/neu阴性的T47D-5R乳腺癌细胞系则不然(参照图9)。HER-2/neu阳性乳腺癌细胞系SKBR3的EC<sub>50</sub>值分别为:**Kadcyla**<sup>®</sup>,32.4ng/ml;DM1结合的Tras-HC-GS-LHS,45.6ng/ml;Tras-LC-GS-LHS,51.4ng/ml,因此在体外肿瘤细胞杀伤实验中的有效范围相仿。与之相比,HER-2/neu阴性的乳腺癌细胞系T47D-5R中没有可检测的特异性细胞毒性,表明当比较相同靶向抗体和相同毒素(DM1)时,分选酶A酶促结合的ADC与传统的化学结合ADC的功能等同(图9)。

[0337] 但是,Tras-HC-GS-LHS和Tras-LC-GS-LHS的分选酶A介导结合ADCs,其与**Kadcyla**<sup>®</sup>的DAR(ca.3-4)相比相对较低的药物-抗体比例(ca.1.80,通过图8中的DAR1和DAR2峰值整合推导得来),并不等同在体外肿瘤细胞杀灭实验中得到成比例差异的细胞毒性(图9)。这一出乎意料的发现可能的原因在于,分选酶A介导的毒素-抗体结合更具有限定性和位点特异性,而与之相比,通过随机化学结合得到的**Kadcyla**<sup>®</sup>限定性较低。

[0338] 实施例9:通过改变插入抗体重链和轻链C端与分选酶A识别基序之间的不同间隔肽长度对抗体重链和轻链与效应分子的分选酶A介导结合进行同步性优化

[0339] 下面将研究抗体重链或轻链C端和LPETG分选酶A识别基序之间的间隔肽长度的影响。为此,基本根据实施例1所述,克隆抗体重链和轻链的表达结构,其中编码了在C-端包含LPETG分选酶A识别基序和strep- II纯化标记(SEQ ID N0s:39-46)且包含或不包含二氨基酸GS(甘氨酸-丝氨酸)间隔子的嵌合CD30特异性单克隆抗体Ac10的重链和轻链(HC序列来自US 2008213289A1,Seq1;LC序列来自US 2008213289A1,Seq9)。使用所述表达结构,通过与相应质粒的共转染,在CHO细胞中生产单克隆抗体Ac10-HC-GS-LHS/LC-GS-LHS和Ac10-HC-LS/LC-LS。Ac10-HC-GS-LHS/LC-GS-LHS是Ac10变体,其重链和轻链的C端各以GS间隔肽、LPETG分选酶A基序、6xHis标记和strep- II标记修饰(SEQ ID N0s:39-42;表3)。Ac10-HC-LS/LC-LS是Ac10变体,其重链和轻链的C端各以LPETG分选酶A基序、6xHis标记和strep- II标记修饰,但不含所述二肽GS连接子(SEQ ID N0s:43-46;表3)。CHO细胞的转染和通过蛋白A琼脂糖层析法对抗体的亲和纯化,基本如实施例4所述。

[0340] 为了研究结合效率,使用分选酶A的系列稀释液,进行五甘氨酸修饰的FITC (Gly<sub>5</sub>-FITC,参见如下式3)的结合反应。



[0342] 式3:五甘氨酸修饰的FITC (Gly<sub>5</sub>-FITC)

[0343] G=甘氨酸残基

[0344] 为此,使用分选酶A在1x分选酶缓冲液(25mM Tris-HCl,pH8.2;150mM NaCl;7.5mM CaCl<sub>2</sub>)中,将Gly<sub>5</sub>-FITC与2种Ac10变体结合,如表4所示。42°C下反应4小时后,通过变性还原性SDS-PAGE凝胶电泳分析反应物。将凝胶置于UV箱中以可视化FITC(图10)。与重链的结合效率高,无论重链C端和LPETG分选酶识别基序之间是否存在GS-连接子。出乎预料地,分选酶A介导的轻链结合效率显著低于分选酶A介导的重链结合效率。此外,令人吃惊地,发现位于抗体轻链C端和LPETG分选酶识别基序之间的二肽GS(甘氨酸-丝氨酸)间隔子,显著影响了连接效率。当存在GS-连接子时,与轻链的结合效率比重链低约5-10倍,而不存在连接子时,效率则降低50-100倍。因此,可以推断,增加轻链和LPETG分选酶识别基序之间的间隔肽长度可能进一步提高结合效率。

[0345] 因此,下面将研究增加轻链和LPETG分选酶识别基序之间的间隔肽长度对结合效能的影响。基本根据实施例1所述制备表达结构,其中编码了以2至5个氨基酸的连接子(SEQ ID NO:47-54)在C-端标记了LPETG分选酶识别基序和strep- II 纯化标记的单克隆抗体Ac10轻链。使用所述表达结构,通过与相应表达结构的共转染,在CHO细胞中生产单克隆抗体AC10-HC-LS/LC-GS-LS,AC10-HC-LS/LC-GGS-LS,AC10-HC-LS/LC-GGGs-LS和AC10-HC-LS/LC-GGGGS-LS。在每个所述抗体中,重链含LPETG分选酶识别基序和strep- II 纯化标记C端修饰(SEQ ID NO:43-44;表3),轻链则具有LPETG分选酶标记和strep- II 标记以及位于LPETG基序之前的GS,GGS,GGGs或GGGGS间隔肽任意之一的C-端修饰(SEQ ID NO:47-54;表3)。CHO细胞的转染和通过蛋白A琼脂糖层析法对抗体的亲和纯化,基本如实施例4所述。

[0346] 为了研究结合效率,使用分选酶A的系列稀释液,将五甘氨酸修饰的FITC (Gly<sub>5</sub>-FITC,参见如上式3)与4种不同Ac10mAb变体在1x分选酶缓冲液(25mM Tris-HCl,pH8.2;150mM NaCl;7.5mM CaCl<sub>2</sub>)中结合,如图5所示。42°C下反应4小时后,通过变性还原性SDS-PAGE凝胶电泳分析反应物。将凝胶置于UV箱中以可视化FITC(图11)。正如预期,所有4种抗体变体的重链结合均具有相同效率。与之相比,增加间隔肽长度显著促进了与轻链的结合。值得注意的是,最长的间隔肽(GGGGS)的分析结果显示,轻链结合效率与重链结合效率相同,由此允许重链和轻链C端均有修饰的抗体进行轻链和重链的同步结合。由此推断,该抗体形式有助于分选酶A介导生产每抗体载4种药物的均一ADC(DAR4)。

抗体	重链修饰	SEQ ID NOs	轻链修饰	SEQ ID NOs
[0347] Ac10-HC-GS-LHS/LC-GS-LHS	GS-LPETG-G-HHHHHH-G-WSHPQFEK	39,40	GS-LPETG-G-HHHHHH-G-WSHPQFEK	41, 42
Ac10-HC-LS/LC-LS	LPETG-G-WSHPQFEK	43, 44	LPETG-G-WSHPQFEK	45, 46
Ac10-HC-LS/LC-GS-LS	LPETG-G-WSHPQFEK	43, 44	GS-LPETG-G-WSHPQFEK	47, 48
Ac10-HC-LS/LC-GGS-LS	LPETG-G-WSHPQFEK	43, 44	GGG-LPETG-G-WSHPQFEK	49, 50
Ac10-HC-LS/LC-GGGS-LS	LPETG-G-WSHPQFEK	43, 44	GGGS-LPETG-G-WSHPQFEK	51, 52
[0348] Ac10-HC-LS/LC-GGGGS-LS	LPETG-G-WSHPQFEK	43, 44	GGGGS-LPETG-G-WSHPQFEK	53, 54
Ac10-HC-LS/LC-GGGGS-LS	LPETG-G-WSHPQFEK	43, 44	GGGGS-LPETG-G-WSHPQFEK	53, 54

[0349] 表3:制备所得的C端修饰单克隆抗体Ac10

反应成分	1-8	9-16	终浓度
[0350] Ac10-HC-GS-LHS/LC-GS-LHS (3.75mg/ml = 25 $\mu$ M)	10	-	5 $\mu$ M
Ac10-HC-LS/LC-LS (3.75mg/ml = 25 $\mu$ M)	-	10	5 $\mu$ M
H <sub>2</sub> O	20	20	-
Gly <sub>5</sub> -FITC (1mM)	5	5	100 $\mu$ M
Sortase A (2x serial dil. of ca. 50 $\mu$ M)	5	5	5 $\rightarrow$ 0.039 $\mu$ M
5x Sortase buffer	10	10	1x

[0351] 表4:单克隆抗体Ac10-HC-GS-LHS/LC-GS-LHS和Ac10-HC-LS/LC-LS的结合条件

反应成分	1-7	8-14	15-21	22-28	终浓度
Ac10-HC-LS/LC-GS-LS (3.75mg/ml=25 $\mu$ M)	10	-	-	-	5 $\mu$ M
Ac10-HC-LS/LC-GGS-LS (3.75mg/ml=25 $\mu$ M)	-	10	-	-	5 $\mu$ M
Ac10-HC-LS/LC-GGGS-LS	-	-	10	-	5 $\mu$ M

(3.75mg/ml=25 $\mu$ M)					
Ac10-HC-LS/LC-GGGGS-L S (3.75mg/ml=25 $\mu$ M)	-	-		10	5 $\mu$ M
H <sub>2</sub> O	20	20	20	20	-
Gly <sub>5</sub> -FITC (1mM)	5	5	5	5	100 $\mu$ M
分选酶 A (2x serial dil. of ca. 25 $\mu$ M)	5	5	5	5	2.5 $\rightarrow$ 0.0 39 $\mu$ M
5x 分选酶缓冲液	10	10	10	10	1x

[0354] 表5:单克隆抗体Ac10-HC-LS/LC-GS-LS,Ac10-HC-LS/LC-GGS-LS,Ac10-HC-LS/LC-GGGS-LS和Ac10-HC-LS/LC-GGGGS-LS的结合条件

[0355] 实施例10:利用strepII-标记亲和纯化制备均一的ADC

[0356] 使用如表6所示具有LPETG分选酶A基序和strepII亲和纯化标记的C端修饰的抗CD30抗体Ac10重链或轻链序列,与Gly<sub>5</sub>-标记的vc-PAB-MMAE(参见实施例5,式1)进行分选酶A介导结合。

抗体	重链修饰	SEQ ID NOs	轻链修饰	SEQ ID NOs
Ac10-HC-LS Ac-10-LC	LPETG-G- WSHPQFEK	43,44	none	29, 30
Ac10-HC Ac10-LC-GS- LHS	无	27, 28	GS-LPETG-G- HHHHHH-G- WSHPQFEK	41,42

[0358] 表6:HC或LC修饰的C端修饰抗体Ac10

[0359] 基本根据实施例1所公开的内容,构建了表4中编码Ac10重链或轻链序列的表达载体。基本根据实施例4所述,进行CHO细胞转染,并以蛋白A-琼脂糖层析进行抗体亲和纯化。

[0360] 基本根据实施例7所述的实验方法,进行重链或轻链分选酶基序标记的抗CD30抗

体与Gly<sub>5</sub>-标记的vc-PAB-MMAE(参见实施例5,式1)的分选酶A介导结合。

[0361] 如本申请的详细描述中的进一步描述,未反应的抗体将保留C-端的strep- II亲和纯化标记,可用于富集完全反应ADC(DAR2)。通过疏水作用色谱(HIC)对重链与vc-PAB-MMAE毒素的分选酶A结合进行分析(图12A),分析显示,大部分分选酶基序修饰的重链已得到结合,但HIC仍能检测到一定百分比的未反应底物(DAR0=药物抗体比例=零)或部分反应底物(DAR1=药物抗体比例=1)(图12A)。

[0362] 因此,基本如实施例7所述,蛋白A纯化的vc-PAB-MMAE结合物通过StrepTactin®亲和柱(IBA Sciences, Göttingen, Germany)4次,以出去未反应或部分反应的分选酶A-修饰抗体。

[0363] 图12显示,数轮后的多相vc-PAB-MMAE抗体药物结合体中,完全反应的DAR2ADCs(DAR2=药物-抗体比例=2)能够得到高度富集。该实验证明,利用向分选酶ALPETG识别基序C端附加亲和纯化标记,从而产生具有限定药物抗体比例(此处为DAR2)的均匀ADC,是可行的。

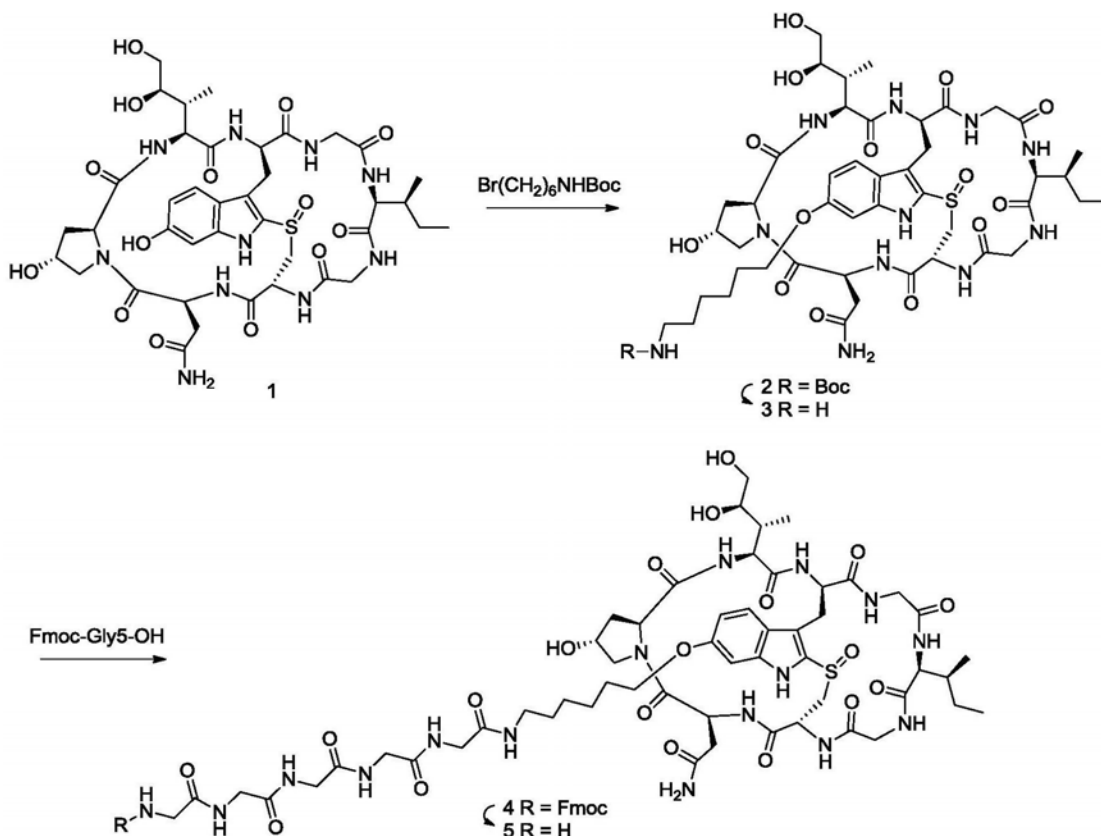
[0364] 实施例11:5x甘氨酸修饰的美登素和 $\alpha$ -鹅膏蕈碱毒素的合成

[0365] 为了允许两个不同的效应分子(优选为毒性效应分子)与重链、轻链C端分别以不同分选酶基序修饰的单个抗体结合,需要用甘氨酸残基修饰两个不同的毒素,优选为具有不同作用模式的毒素,由此可以对双重效应分子结合ADC靶定的癌细胞使用两种不同且潜在具有协同效应的途径进行攻击。对满足该要求的两个不同的甘氨酸修饰毒性效应分子(此处为美登素和 $\alpha$ -鹅膏蕈碱)进行合成,并在此进行描述。

[0366] 11.1 甘氨酸修饰的 $\alpha$ -鹅膏蕈碱的合成

[0367] 取30mg $\alpha$ -鹅膏蕈碱(结构1)(Sigma-Aldrich公司,订单号#A2263)溶解在1ml无水DMSO中。向该溶液加入19mg NH-Boc-氨基己基溴,随后加入叔丁醇钾(在THF中的1M溶液,35 $\mu$ L)。将反应混合物在室温下搅拌6小时,并再次加入叔丁醇钾(在THF中的1M溶液,20 $\mu$ L)。将反应物在室温下保持16小时。加入乙酸(10 $\mu$ L),并通过RP-HPLC直接纯化粗混合物(Sunfire C185 $\mu$ 3cm x 10cm column,50mL/min,5-50%乙腈/水15分钟梯度)。收集期望组分并冻干,得到结构式2的白色粉末(15mg),用TFA/DCM溶液(1/1,v/v,1ml)在室温下处理30分钟。减压除去挥发物,得到结构3的淡黄色胶状物,其不经进一步纯化就用于下一步骤。

[0368] 将Fmoc-Gly5-OH(8mg)溶解在无水DMF(0.5ml)中。加入HATU(Sigma-Aldrich公司,订单号#445460)(6mg),接着加入DIEA(10ml)(Sigma-Aldrich公司,订单号#496219)。将混合物在室温下轻轻搅拌30秒,然后转移到含化合物3的DMF溶液(0.5ml)中。30分钟后,LC/MS分析表明,化合物3全部消耗。加入哌啶(30 $\mu$ L),并通过LC/MS监测该反应的进程。1小时后,加入乙酸以中和反应,通过RP-HPLC(Sunfire C185 $\mu$ 3cm x 25cm column,50mL/min,2-40%乙腈/水30分钟梯度)纯化该混合物。合并组分并冻干,得到结构5的白色粉末(12mg)。化合物5的分析数据如图13A所示。



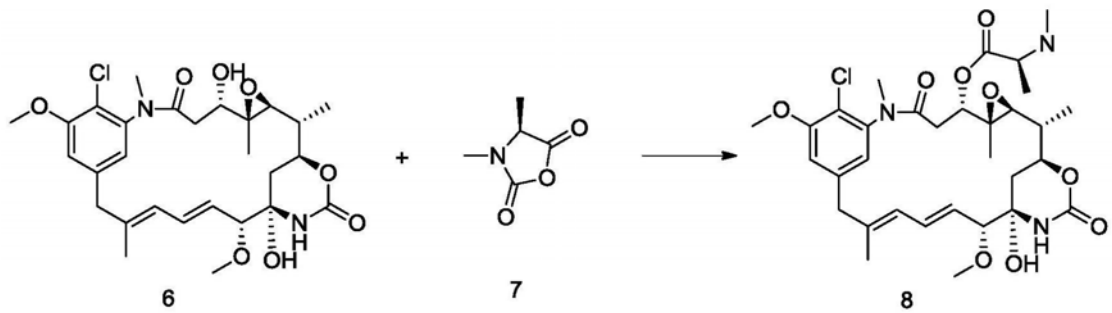
[0369]

[0370] 方案1: 甘氨酸修饰的 $\alpha$ -鹅膏蕈碱的合成

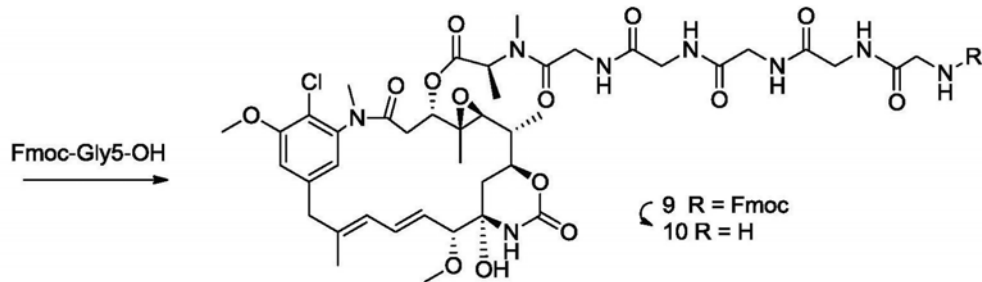
[0371] 11.2. 甘氨酸修饰的美登素的合成:

[0372] 美登醇 (0.6g, 1.1mmol) (Clearsynth Labs, Mumbai, India) 溶解在无水THF (6ml) 和无水DMF (3ml) 中, 然后加入1, 2ml DIEA (Sigma-Aldrich, 订单号#496219)。将溶液置于氩气氛下。一次性加入三氟甲磺酸锌 (1.2g) 和NMeAla NCA (0.7g)。将混合物超声处理, 直至固体溶解。将反应混合物在室温下搅拌2天, 然后用乙酸乙酯 (100ml) 稀释。用饱和的 $\text{NaHCO}_3$  (水溶液,  $2 \times 50\text{ml}$ ) 和盐水 (50ml) 洗涤。将有机层干燥 ( $\text{MgSO}_4$  干燥) 并浓缩, 得到粗美登醇-3- (S) - $\alpha$ -N-二甲基氨基丙 (8), 其不经进一步纯化即直接用于下一步骤。

[0373] 将Fmoc-Gly5-OH (26mg) 溶解在无水DMF (1ml) 中。加入HATU (Sigma-Aldrich公司, 订单号#445460) (19mg), 接着加入DIEA (18 $\mu\text{L}$ )。将混合物轻轻搅拌在室温下30秒, 然后转移到化合物8的THF溶液 (1ml) 中。30分钟后, LC/MS分析表明, 化合物8全部耗尽。加入哌啶 (40 $\mu\text{L}$ ) 并通过LC/MS监测反应进程。2小时后, 向反应中加入醚 (40ml), 收集沉淀的固体并用乙醚洗涤。通过RP-HPLC (Sunfire C185 $\mu\text{m}$  3cm x 10cm柱, 50ml/min, 10-60% 乙腈/水20分钟梯度) 纯化粗化合物。合并各部分并冻干得到化合物10, 其为白色粉末 (33mg)。化合物10的分析数据如图13B所示。



[0374]



[0375] 方案2:甘氨酸修饰的美登素的合成

[0376] 重要的是,应当注意,原则上,任何毒素可功能化并用于分选酶介导的酶促结合,只要该毒素上连接有5甘氨酸(如下所示)或大于或等于1甘氨酸的任何数目甘氨酸残基(参照图14)即可。

[0377] 实施例12:分选酶A结合的曲妥单抗-DM1在SKOV3卵巢癌异种移植模型中的体内肿瘤抑制

[0378] 取含 $5 \times 10^6$  SKOV3肿瘤细胞的 $200 \mu\text{l}$  PBS/基底胶(1:1比例),皮下植入5-6周龄雌性NMRI裸小鼠的左侧。通过卡规监控原发性肿瘤体积。平均肿瘤体积达到 $100\text{-}200\text{mm}^3$ 后,根据肿瘤大小将荷瘤动物随机分成3组(每组10只动物)。随机化之日(第0天)和第21天,对组1,2和3的动物分别静脉注射 $5\text{ml/kg}$  PBS, $15\text{mg/kg}$  **Kadcyla**<sup>®</sup>,或 $15\text{mg/kg}$ 分选酶A结合的曲妥单抗-DM1。每两周通过卡规测量肿瘤体积(图15)。39天后,停止实验,并根据公认的动物实验准则实施安乐死。

[0379] 在研究的过程中,模拟注射PBS的对照动物体内的肿瘤稳步增长到大约 $600\text{mm}^3$ 的体积。与之相比,**Kadcyla**<sup>®</sup>处理的动物的肿瘤萎缩,第39天基本上检测不到肿瘤。分选酶A结合的曲妥单抗-DM1的抗肿瘤活性没有显著不同于市售**Kadcyla**<sup>®</sup>,但分选酶-结合的T-DM1的药物-抗体比例较低,约为2,而对比**Kadcyla**<sup>®</sup>的DAR报告为3.5。与实施例8中的数据相结合,其结果表明,采用相同的抗体和毒素部分的分选酶结合的ADC,与市售的化学结合**Kadcyla**<sup>®</sup>相比,体外和体内肿瘤杀伤活性相仿,而药物-抗体比例更低。

[0380] 实施例13:粗CHO细胞上清液中的分选酶A介导的结合反应

[0381] 基本如实施例4所述地,在CHO细胞中生产曲妥单抗变体Tras-HC-LS/LC-GGGGS-LS,该变体由C端携带LPETG分选酶基序和Strep II纯化标记(SEQ ID NOs:055-056)的重链和C端携带五氨基酸Gly<sub>4</sub>-Ser间隔子(GGGGS)、LPETG分选酶基序和Strep II标记(SEQ ID NOs:057-058)的轻链组成。所得无血清粗细胞上清液含有约 $157\text{mg/L}$  Tras-HC-LS/LC-

GGGS-LS,并基本如实施例9所述,向该上清液中添加分选酶缓冲液、Gly<sub>5</sub>-FITC和分选酶A的系列稀释液,从而直接用于结合反应。同时,在相同的其他条件下,使用蛋白A亲和层析纯化后的Tras-HC-LS/LC-GGGS-LS进行结合反应。42℃下反应4小时后,通过变性还原性SDS-PAGE凝胶电泳分析反应物。将凝胶置于UV箱中以可视化FITC,并以考马斯亮蓝进行蛋白染色(图16)。数据显示,出乎意料地,粗细胞培养上清液中分选酶A介导的抗体结合反应与纯化抗体同样搞笑。此外,该结合反应具有高特异性,且粗CHO细胞上清液中的蛋白污染物无一非特异性结合。综上所述,数据表明,分选酶反应的稳定性能够有助于允许在CHO细胞生产后,纯化和下游加工之前,直接进行结合反应,从而促进ADC生产。

[0382] 参考文献

- [0383] Antos et al. (2009a) J. Am. Chem. Soc. 131, pp. 10800-10801
- [0384] Antos et al. (2009b) J. Biol. Chem. 284, 16028-16036
- [0385] Appleby et al. (2009) JBC 284, 6194-99
- [0386] Axup et al. (2012) Proc. Natl. Acad. Sci USA 109, 16102-16106
- [0387] Elleuche (2010) Appl. Microbiol. Biotechnol. 87, 479-489
- [0388] Graus-Porta et al. (1995) Mol. Cell. Biol. 15, p1182ff
- [0389] Hofer et al. (2009) Biochemistry 48, 12047-57
- [0390] Junutula et al. (2008) Nat. Biotechnol., 26, 925-932
- [0391] Lambert (2012) British J Clin Pharmacol 76, 248-262,
- [0392] Lemke (2011) Methods Mol. Biol. 751, 3-15
- [0393] Levary et al. (2011) PLoS One 6, e18342
- [0394] Madej et al. (2012) Biotechnol. Bioeng. 109, 1461-1470
- [0395] Mao et al. (2004) J. Am. Chem. Soc. 126, 2670-2671,
- [0396] Mazmanian et al. (1999) Science 285, 760-763
- [0397] McDonagh et al. (2006) Prot. Engin. Design Selection 19, 299-307
- [0398] **Möhlmann** et al. (2011) Chembiochem. 12, 1774-1780,
- [0399] Mullard (2013) Nature Rev. Drug Discov. 12, 329-332) .
- [0400] Parthasarathy et al. (2007) Bioconjugate Chem. 18, 469-476
- [0401] Perler (2002) Nucl. Acids Res. 30, 383-384
- [0402] Song et al. (2012) PLoS One 7, e45355
- [0403] Spirig et al. (2011) Molecular Microbiol. 82, 1044-1059
- [0404] Sun et al. (2004) J. Biol. Chem. 279, 35281-35286
- [0405] Swee et al. (2013) Proc. Natl. Acad. Sci USA 110, 1428-1433
- [0406] Ton-That et al. (1999) Proc. Natl. Acad. Sci USA 96, 12424-12429
- [0407] Tsukiji (2009) Chembiochem. 10, 787-798)
- [0408] Volkmann et al. (2009) PLoS One 4, e8381
- [0409] Xu et al. (1993) Cell 75, 1371-1377



Sequence Listings\_CN.txt

序列表

<110> 恩比伊治疗股份公司  
 <120> 制备免疫配体/效应分子结合物的方法  
 <130> ND40306  
 <160> 58  
 <170> PatentIn version 3.5  
 <210> 1  
 <211> 417  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> 人源化抗人CD19抗体hBU12的VH编码区域  
 <400> 1  
 atgggatgga gctggatctt tcttttctc ctgtcaggaa ctgcagggtg ccattgtcag 60  
 gttcagctgc aagagtctgg ccctgggttg gttaagccct cccagaccct cagtctgact 120  
 tgtactgtgt ctgggggttc aatcagcact tctggtatgg gtgtaggctg gattaggcag 180  
 caccagggga agggctctgga gtggattgga cacatttggg gggatgatga caagatat 240  
 aaccagccc tgaagagcag agtgacaatc tctgtggata cctccaagaa ccagtttagc 300  
 ctcaagctgt ccagtgtgac agctgcagat actgctgtct actactgtgc tagaatggaa 360  
 ctttgtctct actattttga ctactggggc caaggcacc ttgtcacagt ctctca 417

[0001]

<210> 2  
 <211> 139  
 <212> PRT  
 <213> 人工序列  
 <220>  
 <223> 人源化抗人CD19抗体hBU12的VH区域的VH区域  
 <400> 2  
 Met Gly Trp Ser Trp Ile Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly  
 1 5 10 15  
 Val His Cys Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys  
 20 25 30  
 Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile  
 35 40 45  
 Ser Thr Ser Gly Met Gly Val Gly Trp Ile Arg Gln His Pro Gly Lys  
 50 55 60  
 Gly Leu Glu Trp Ile Gly His Ile Trp Trp Asp Asp Asp Lys Arg Tyr  
 65 70 75 80  
 Asn Pro Ala Leu Lys Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys  
 85 90 95  
 Asn Gln Phe Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala  
 100 105 110  
 Val Tyr Tyr Cys Ala Arg Met Glu Leu Trp Ser Tyr Tyr Phe Asp Tyr

115 120 125

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
130 135

<210> 3  
<211> 375  
<212> DNA  
<213> 人工序列

<220>  
<223> 人源化抗人CD19抗体hBU12的VL编码区域

<400> 3  
atgaagtgc ctgtaggct gttggtctg atgttctgga ttctgcttc cagcagtga 60  
attgttctca ccagctctcc agcaacctg tctctctctc caggggaaag ggctaccctg 120  
agctgcagtg ccagctcaag tgtaagttac atgcactggt accagcagaa gccagggcag 180  
gtctccagac tctctgattta tgacacatcc aaactggctt ctggtattcc agcaaggttc 240  
agtggcagtg ggtctggaac agattttaca ctcacaatca gcagcctgga gccagaggat 300  
gttgctgtct attactgttt tcaggggagt gtatacccat tcacttttgg ccaagggaca 360  
aagttgaaa tcaaa 375

<210> 4  
<211> 125  
<212> PRT  
<213> 人工序列

[0002]

<220>  
<223> 人源化抗人CD19抗体hBU12的VH区域的氨基酸序列

<400> 4

Met Lys Leu Pro Val Arg Leu Leu Val Leu Met Phe Trp Ile Pro Ala  
1 5 10 15

Ser Ser Ser Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu  
20 25 30

Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Ser Ala Ser Ser Val  
35 40 45

Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu  
50 55 60

Leu Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Ile Pro Ala Arg Phe  
65 70 75 80

Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu  
85 90 95

Glu Pro Glu Asp Val Ala Val Tyr Tyr Cys Phe Gln Gly Ser Val Tyr  
100 105 110

Pro Phe Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys  
115 120 125

<210> 5  
<211> 1083



Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys  
 85 90 95  
 Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro  
 100 105 110  
 Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys  
 115 120 125  
 Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val  
 130 135 140  
 Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr  
 145 150 155 160  
 Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu  
 165 170 175  
 Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His  
 180 185 190  
 Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys  
 195 200 205  
 Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln  
 210 215 220  
 [0004] Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu  
 225 230 235 240  
 Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro  
 245 250 255  
 Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn  
 260 265 270  
 Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu  
 275 280 285  
 Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val  
 290 295 300  
 Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln  
 305 310 315 320  
 Lys Ser Leu Ser Leu Ser Pro Gly Lys Leu Pro Glu Thr Gly His His  
 325 330 335  
 His His His His Gly Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Gly  
 340 345 350  
 Trp Ser His Pro Gln Phe Glu Lys  
 355 360  
 <210> 7  
 <211> 414

<212> DNA  
<213> 人工序列

<220>

<223> 具有编码LPETG分选酶标记, 6xHis标记和strepII标记的框内3' 延长部的人IgG1 k轻链恒定编码区域

<400> 7

```
acggtggctg caccatctgt cttcatcttc cgcctatctg atgagcagtt gaaatctgga      60
actgcctctg ttgtgtgctt gctgaataac ttctatccca gagaggccaa agtacagtgg      120
aaggtggata acgcccctcca atcgggtaac tcccaggaga gtgtcacaga gcaggacagc      180
aaggacagca cctacagcct cagcagcacc ctgacgctga gcaaagcaga ctacgagaaa      240
cacaagaagt acgectgcga agtcacccat cagggcctga gctcgccegt cacaagagac      300
ttcaacaggg gagagtgtct gcccgagacc ggccaccacc accaccacca cggcgagcag      360
aagctgatca gcgaggagga cctgggctgg agccaccccc agttcgagaa gtag          414
```

<210> 8

<211> 137

<212> PRT

<213> 人工序列

<220>

<223> 具有LPETG分选酶标记, 6xHis标记和strepII标记的人IgG1 k轻链恒定区域的氨基酸序列

<400> 8

```
Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
1           5           10          15
```

[0005]

```
Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
20          25          30
```

```
Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
35          40          45
```

```
Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
50          55          60
```

```
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
65          70          75          80
```

```
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
85          90          95
```

```
Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Leu Pro Glu Thr Gly His
100         105         110
```

```
His His His His His Gly Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
115         120         125
```

```
Gly Trp Ser His Pro Gln Phe Glu Lys
130         135
```

<210> 9

<211> 1503

<212> DNA

<213> 人工序列

<220>

<223> 带有C端LPETG分选酶标记, 6xHis标记和strepII标记的hBU12全人IgG1 VH-CH重链编码区域

	<400> 9	
	atgggatgga gctggatctt tcttttctc ctgtcaggaa ctgcaggtgt ccattgtcag	60
	gttcagctgc aagagtctgg ccctgggttg gttaaagcct cccagaccct cagtctgact	120
	tgtacttgtt ctgggggttc aatcagcact tctggtatgg gtgtaggctg gattaggcag	180
	caccagggga aggtcttggg gtggattgga cacatttggg gggatgatga caagagatat	240
	aaccagccc tgaagagcag agtgacaatc tctgtggata cctccaagaa ccagttttagc	300
	ctcaagctgt ccagtgtgac agctgcagat actgctgtct actactgtgc tagaatggaa	360
	ctttgtctct actattttga ctactggggc caaggcacc ttgtcacagt ctctcagct	420
	agcaccaagg gcccatctgt ctccccctg gcaccctct ccaagagcac ctctgggggc	480
	acagctgcc tgggctgctt ggtcaaggac tacttccctg aacctgtgac agtgtcctgg	540
	aactcaggcg ccctgaccag cggcgtgcac accttcccg ctgtcttaca gtcctcagga	600
	ctctactccc tcagcagcgt ggtgaccgtg ccctccagca gcttgggcac ccagacctac	660
	atctgcaacg tgaatcaca gccagcaac accaaggtgg acaagaaagt tgagcccaaa	720
	tcttgtgaca aaactcacac atgcccaccg tgcccagcac ctgaactcct ggggggaccg	780
	tcagtcttcc tcttcccc aaacccaag gacaccctca tgatctccc gaccctgag	840
	gtcacatgcg tgggtgtgga cgtgagccac gaagaccctg aggtcaagtt caactgttac	900
	gtggacggcg tggaggtgca taatccaag acaaagccgc gggaggagca gtacaacagc	960
	acgtaccgtg tggtcagcgt cctcaccgtc ctgcaccagg actggtgaa tggcaaggag	1020
	tacaagtgca aggtctcaa caaagccctc ccagccccca tcgagaaaac catctcaaa	1080
[0006]	gccaaaggcg agccccgaga accacaggtg tacaccctgc cccatcccg ggatgagctg	1140
	accaagaacc aggtcagcct gacctgcctg gtcaaaggct tctatcccag cgacatcgcc	1200
	gtggagtggg agagcaatg gcagccggag aacaactaca agaccacgcc tcccgtgtg	1260
	gactccgacg gctccttctt cctctacagc aagctcaccg tggacaagag caggtggcag	1320
	caggggaacg tcttctcatg ctccgtgatg catgaggctc tgcacaacca ctacacacag	1380
	aagagcctct ccctgtctcc gggtaaacctg cccgagaccg gccaccacca ccaccaccac	1440
	ggcagcaga agctgatcag cgaggaggac ctgggctgga gccaccccc gttcgagaag	1500
	tag	1503
	<210> 10	
	<211> 500	
	<212> PRT	
	<213> 人工序列	
	<220>	
	<223> 具有C端LPETG分选酶标记, 6xHis标记和strepII标记的hBU12全人IgG1 VH-CH重链区域的氨基酸序列	
	<400> 10	
	Met Gly Trp Ser Trp Ile Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly	
	1 5 10 15	
	Val His Cys Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys	
	20 25 30	
	Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile	
	35 40 45	

Ser Thr Ser Gly Met Gly Val Gly Trp Ile Arg Gln His Pro Gly Lys  
 50 55 60

Gly Leu Glu Trp Ile Gly His Ile Trp Trp Asp Asp Asp Lys Arg Tyr  
 65 70 75 80

Asn Pro Ala Leu Lys Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys  
 85 90 95

Asn Gln Phe Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala  
 100 105 110

Val Tyr Tyr Cys Ala Arg Met Glu Leu Trp Ser Tyr Tyr Phe Asp Tyr  
 115 120 125

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly  
 130 135 140

Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly  
 145 150 155 160

Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val  
 165 170 175

Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe  
 180 185 190

[0007] Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val  
 195 200 205

Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val  
 210 215 220

Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys  
 225 230 235 240

Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu  
 245 250 255

Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr  
 260 265 270

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val  
 275 280 285

Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val  
 290 295 300

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser  
 305 310 315 320

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu  
 325 330 335

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala





caccaccacg gcgagcagaa gctgatcagc gaggaggacc tgggctggag ccacccccag 780  
ttcgagaagt ag 792

<210> 12  
<211> 263  
<212> PRT  
<213> 人工序列

<220>  
<223> 具有C端LPETG分选酶标记, 6xHis标记和strepII标记的hBU12全人IgG1 VL-CL k链区域的氨基酸序列

<400> 12

Met Lys Leu Pro Val Arg Leu Leu Val Leu Met Phe Trp Ile Pro Ala  
1 5 10 15

Ser Ser Ser Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu  
20 25 30

Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Ser Ala Ser Ser Ser Val  
35 40 45

Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu  
50 55 60

Leu Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Ile Pro Ala Arg Phe  
65 70 75 80

[0009]

Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu  
85 90 95

Glu Pro Glu Asp Val Ala Val Tyr Tyr Cys Phe Gln Gly Ser Val Tyr  
100 105 110

Pro Phe Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr Val  
115 120 125

Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys  
130 135 140

Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg  
145 150 155 160

Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn  
165 170 175

Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser  
180 185 190

Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys  
195 200 205

Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr  
210 215 220

Lys Ser Phe Asn Arg Gly Glu Cys Leu Pro Glu Thr Gly His His His  
225 230 235 240

His His His Gly Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Gly Trp  
245 250 255

Ser His Pro Gln Phe Glu Lys  
260

- <210> 13
- <211> 7045
- <212> DNA
- <213> 人工序列

<220>  
<223> 具有C端LPETG分选酶标记, 6xHis标记和strepII标记以及HindIII和NotI克隆位点的hBU12人IgG1 VH-CH 重链编码区域

[0010]

```

<400> 13
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgtc gagtagtgcg      120
cgagcaaaat ttaagtaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtctat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
cccgccatt  gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc      420
attgacgtca atgggtggac tatttacggt aaactgccca ctggcagta catcaagtgt      480
atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgccagta  catgacctta tgggactttc ctaactggca gtacatctac gtattagtca      600
tcgctattac catggtgatg cgggtttggc agtacatcaa tggcgtgga tagcggtttg      660
actcacgggg atttcaagt  ctccacccca ttgacgtcaa tgggagtttg ttttggcacc      720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg      780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca      840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc      900
gtttaaactt aagcttccat gggatggagc tggatcttcc ttttctctct gtcaggaact      960
gcaggtgtcc attgtcaggt tcagctgcaa gagtctggcc ctgggttggg taagccctcc     1020
cagacctca  gctgacttgg tactgtgtct gggggttcaa tcagcacttc tggatgggt      1080
gtaggcctgga ttaggcagca cccaggaag  ggtctggagt ggattggaca catttgggtg      1140
gatgatgaca agagatataa cccagccctg aagagcagag tgacaatctc tgtggatacc     1200
tccaagaacc agtttagcct caagctgtcc agtgtgacag ctgcagatac tgctgtctac     1260
tactgtgcta gaatggaact ttggtcctac tattttgact actggggcca aggcaccctt     1320
gtcacagtct cctcagtag  caccaagggc ccatctgtct tccccctggc accctctcc     1380
aagagcacct ctgggggcac agctgccttg ggctgcctgg tcaaggacta ctccctgaa     1440
cctgtgacag tgtcctggaa ctcaggcgc  ctgaccagcg gcgtgcacac cttccggct     1500
gtcctacagt cctcaggact ctactcctc agcagcgtgg tgaccgtgcc ctccagcagc     1560
ttgggcaccc agacctacat ctgcaacgtg aatcacaagc ccagcaaac caaggtggac     1620
aagaaagttg agcccaaatc ttgtgacaaa actcacacat gccaccgtg cccagcacct     1680

```

	gaactcctgg ggggaccgtc agtcttctc ttccccca aaccaagga caccctcatg	1740
	atctcccga cccctgaggt cacatgctg gtggtggac tgagccacga agaccctgag	1800
	gtcaagtta actggtactg ggacggcgtg gaggtgcata atgccaagac aaagcccgcg	1860
	gaggagcagt acaacagcac gtaccgtgtg gtcagcgtcc tcaccgtect gcaccaggac	1920
	tggtgtaat gcaaggagta caagtcaag gtctccaaca aagccctccc agccccatc	1980
	gagaaaacca tctccaaagc caaaggcag ccccgagaac cacaggtgta caccctgccc	2040
	ccatccggg atgagetgac caagaaccag gtcagcctga cctgcttggc caaaggcttc	2100
	tatcccagcg acatcgccgt ggagtgggag agcaatgggc agccggagaa caactacaag	2160
	accagcctc ccgtgctgga ctccgacggc tccttctcc tctacagca gtcaccgtg	2220
	gacaagagca ggtggcagca ggggaacgtc ttctcatgct ccgtgatgca tgaggctctg	2280
	cacaaccact acacacagaa gagcctctcc ctgtctccgg gtaaactgcc cgagaccggc	2340
	caccaccacc accaccacgg cgagcagaag ctgatcagcg aggaggacct gggctggagc	2400
	cacccccagt tcgagaagta ggcggccgtc cgagtctaga gggcccgttt aaaccctctg	2460
	atcagcctc actgtgcctt ctagtggcca gccatctgtt gtttgcctc cccccgtgcc	2520
	ttccttgacc ctggaaggtg ccaactccac tgtccttcc taataaatg aggaaattgc	2580
	atgcattgt ctgagttagt gtcattctat tctgggggtt ggggtggggc aggacagcaa	2640
	gggggagat tgggaagaca atagcaggca tgctggggat gcggtgggct ctatggcttc	2700
	tgaggcggaa agaaccagct ggggctctag ggggtatccc cacgcgcctc gtagcggcgc	2760
	attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgcctc	2820
[0011]	agcggccgtc ctttgcctt tcttccctc ctttctgcc acgttcgcc gctttccccg	2880
	tcaagctcta aatcggggca tcccttagg gttecgattt agtgctttac ggcacctcga	2940
	ccccaaaaa cttgattagg gtgatggtc acgtagtggg ccacgcctc gatagacggc	3000
	ttttgcctc ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaaactgg	3060
	aacaacactc aaccctatct cggctctatc ttttgattta taaggattt tggggatttc	3120
	ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt aattctgtg	3180
	aatgtgtgtc agttagggtg tggaaagtcc ccaggctccc caggcaggca gaagtatgca	3240
	aagcatgcat ctcaattagt cagcaaccag gtgtggaaag tcccaggct cccagcagg	3300
	cagaagtatg caaagatgc atctcaatta gtcagcaacc atagtcccgc ccctaactcc	3360
	gccatcccc cccctaactc cgcccagttc cgcccattct ccgcccctg gctgactaat	3420
	ttttttatt tatgcagagg ccgaggccgc ctctgcctc gagctattcc agaagtagtg	3480
	aggagcttt tttggaggcc taggcctttg caaaaagctc ccgggagctt gtatatccat	3540
	tttcgatct gatcagcac tgatgaaaaa gcctgaactc accgcgacgt ctgtcgagaa	3600
	gtttctgac gaaaagtctg acagcgtctc cgacctgatg cagctctcgg agggcgaaga	3660
	atctcgtct tcagcttcg atgtaggagg gcgtggatat gtctgcggg taaatagctg	3720
	cgccgatggt ttctacaaag atcgttatgt ttatcggcac tttgcatcgg ccgcgctccc	3780
	gattccggaa gtgcttgaca ttggggaatt cagcgagagc ctgacctatt gcattctccc	3840
	ccgtgcacag ggtgtcacgt tgcaagacct gcctgaaacc gaactgcccg ctgttctgca	3900
	gccggtcgcg gaggccatgg atgcgatcgc tgcggccgat cttagccaga cgagcgggtt	3960

cgGCCcattc ggaccgcaag gaatcggTca atactactaca tggcgtgatt tcatatgcgc	4020
gattgctgat ccccatgtgt atcactggca aactgtgatg gacgacaccg tcagtgcgtc	4080
cgTcgcgcag gctctcgaTg agctgatgct ttgggccgag gactgccccg aagtccggca	4140
cctcgtgcac gcggatttcg gctccaacaa tgtcctgacg gacaatggcc gcataacagc	4200
ggTcattgac tggagcgagg cgatgttcgg ggattcccaa tacgaggtcg ccaacatctt	4260
cttctggagg ccgtggttgg ctTgtatgga gcagcagacg cgctacttcg agcggaggca	4320
tccggagctt gcaggatcgc cgcggctccg ggcgtatatg ctccgcattg gtcttgacca	4380
actctatcag agcttggttg acggcaatTt cgatgatgca gcttgggcgc agggTcgaTg	4440
cgacgcaatc gtccgatccg gagccgggac tGtcgggcgt acacaaatcg cccgcagaag	4500
cgcgcccgTc tggaccgatg gctgtgtaga agtactcgcc gatagTggaa accgacgccc	4560
cagcactcgt ccgagggcaa aggaatagca cgtgctacga gatttcgatt ccaccgccgc	4620
cttctatgaa aggttgggct tcggaatcgt tttccgggac gccggctgga tgatcctcca	4680
gcgcggggat ctcatgctgg agTtcttcgc ccacccaac ttgtttattg cagcttataa	4740
tggttacaaa taaagcaata gcatcacaaa tttcacaaat aaagcatttt tttcactgca	4800
ttctagttgt ggtttgtcca aactcatcaa tGtatcttat catgtctgta taccgtgcac	4860
ctctagctag agcttggcgt aatcatggTc atagctgttt cctgtgtgaa attgttatcc	4920
gTcacaatT ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta	4980
atgagTgagc taactcacat taattgcgtt gcgctcactg cccgctttcc agTcgggaaa	5040
[0012] cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat	5100
tgggcgctct tccgcttccT cgctcactga ctcgtcgcgc tcggtcgttc ggctcggcg	5160
agcggtatca gTcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc	5220
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt	5280
gTcggcgttt ttccatagc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag	5340
tcagaggtgg cgaaaccgga caggactata aagataccag gcgtttcccc ctggaagctc	5400
cctcgtgcgc Tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cttttctccc	5460
ttcgggaagc gtggcgcttt ctcaatgctc acgctgtagg tatctcagtt cggTgtaggt	5520
cgTtcgctcc aagctgggct gTgtgcacga accccccgtt cagcccgacc gctgcgcctt	5580
atccggtaac tatcgtcttg agtccaacc ggtaagacac gacttatcgc cactggcagc	5640
agccactggt aacaggatta gcagagcgag gTatgtaggc ggtgctacag agTtcttgaa	5700
gtggtggcct aactacggct aactagaag gacagtattt ggtatctgcg ctctgctgaa	5760
gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg	5820
tagcggTggt tttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga	5880
agatcctttg atcttttcta cggggtctga cgctcagTgg aacgaaaact cacgttaagg	5940
gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg	6000
aagttttaaa tcaatctaaa gtatatatga gTaaactTgg tctgacagtt accaatgctt	6060
aatcagTgag gcacctatct cagcgatctg tctatttcgt tcacctag ttgcctgact	6120
ccccgctgTg tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat	6180

gataccgga gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg 6240  
 aaggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg 6300  
 ttgccggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgecat 6360  
 tgctacagge atcgtggtgt cagctcgtc gtttggatg gcttcattca gctccggttc 6420  
 ccaacgatca aggcgagtta catgatcccc catgttggtgc aaaaaagcgg ttagctcctt 6480  
 cggctctcgg atcgttgtca gaagtaagt ggcgcagtg ttatcactca tggttatggc 6540  
 agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcctttctg tgactggtga 6600  
 gtactcaacc aagtcattct gagaatagt tatgcggcga ccgagttgct cttgcccgcc 6660  
 gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtca tcattggaaa 6720  
 acgttcttcg gggcgaaaac tcicaaggat cttaccgctg ttgagatcca gttcagatga 6780  
 acccactcgt gcaccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg 6840  
 agcaaaaaca ggaagcaaaa atgccgcaaa aaagggaata agggcgacac ggaatgttg 6900  
 aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat 6960  
 gagcggatc atatttgaat gtatttagaa aataaaca ataggggttc cgccacatt 7020  
 tccccgaaa gtgccacctg acgtc 7045

<210> 14  
 <211> 6334  
 <212> DNA  
 <213> 人工序列

[0013]

<220>  
 <223> 具有C端LPETG分选酶标记, 6xHis标记和strepII标记以及HindIII和NotI克隆位点的hBU12人IgG1 VL-CL k轻链编码区域

<400> 14  
 gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaate tgctctgatg 60  
 ccgatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgt gagtagtgcg 120  
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180  
 ttagggttag gcgttttgcg ctgcttcgag atgtacgggc cagatatacg cgttgacatt 240  
 gattattgac tagttattaa tagtaataca ttacggggtc attagttcat agcccatata 300  
 tggagtccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc 360  
 cccgccatt gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc 420  
 attgacgtca atgggtggac tatttacggt aaactgcca cttggcagta catcaagtgt 480  
 atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
 atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca 600  
 tcgctattac catggtgatg cggttttggc agtacatcaa tggcggtgga tagcggtttg 660  
 actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc 720  
 aaaatcaacg ggactttcca aaatgtcgt acaactccgc ccattgacg caaatggcg 780  
 gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840  
 ctgcttactg gcttatcgaa attaatacga ctactatag ggagaccaa gctggctagc 900  
 gtttaaacct aagcttccat gaagttgcct gttaggctgt tgggtgctgat gttctggatt 960  
 cctgcttcca gcagtgaat tgttctcacc cagtctccag caacctgtc tctcttcca 1020

	ggggaaaggg ctaccctgag ctgcagtgcc agctcaagtg taagttacat gcaactggtac	1080
	cagcagaagc cagggcaggc tcccagactc ctgatttatg acacatccaa actggcttct	1140
	ggtattccag caaggttcag tggcagtggg tctggaacag attttact cacaatcagc	1200
	agcctggagc cagaggatgt tgctgtctat tactgttttc aggggagtg ataccattc	1260
	acttttggcc aagggacaaa gttggaate aaaagaactg tggtgcacc atctgtcttc	1320
	atcttcccgc catctgatga gcagttgaaa tctggaactg cctctgttgt gtgcctgctg	1380
	aataacttct atcccagaga ggccaaagta cagtgaaggg tggataacgc cctccaatcg	1440
	ggtaactccc aggagagtgt cacagagcag gacagcaagg acagcaccta cagcctcagc	1500
	agcaccttga cgctgagcaa agcagactac gaaaaacaca aagtctacgc ctgcgaagtc	1560
	accatcagg gcctgagctc gcccgtcaca aagagcttca acaggggaga gtgtctgccc	1620
	gagaccggcc accaccacca ccaccacggc gagcagaagc tgatcagcga ggaggacctg	1680
	ggctggagcc acccccagtt cgagaagtag gcggccgctc gagtctagag ggcccgttta	1740
	aaccctgctg tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttgcccctc	1800
	ccccgtgctt tcttggacce tgggaaggtc cactcccact gtcccttctc aataaaatga	1860
	ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctggggggtg ggggtgggca	1920
	ggacagcaag ggggaggatt ggggaacaaa tagcaggcat gctggggatg cgggtgggctc	1980
	tatggcttct gaggcgaaa gaaccagctg gggctctagg gggatcccc acgcgccctg	2040
	tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc	2100
	cagcgcctca ggcgccctc tttcgtctt tttcccttc tttctcgcca cgttcgccgg	2160
[0014]	ctttcccctg caagctctaa atcggggcat cccttaggg ttccgattta gtgctttacg	2220
	gcacctcagc cccaaaaaac ttgattaggg tgatggttca cgtagtgggc catcgccctg	2280
	atagacggtt tttcgcctt tgacgttga gtccacgttc tttaatagtg gactcttgtt	2340
	ccaaactgga acaaacctca accctatctc ggtctatctt tttgatttat aagggatitt	2400
	ggggatttgc gcctatttgt taaaaaatga gctgatttaa caaaaattta acgcgaatta	2460
	attctgtgga atgtgtgca gttagggtgt ggaaagtccc caggctcccc aggcagcagc	2520
	aagtatgcaa agcatgcatc tcaattagtc agcaaccagg tgtggaaagt ccccagctc	2580
	cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca tagtcccgcc	2640
	cctaactccg cccatcccgc ccctaactcc gcccagttcc gccattctc cgecccatgg	2700
	ctgactaatt tttttattt atgcagagge cgaggccgcc tctgcctctg agctattcca	2760
	gaagtagtga ggagctttt ttgaggcct aggcctttgc aaaaagctcc cgggagcttg	2820
	tatatccatt ttcggatctg atcagcacgt gatgaaaaag cctgaaactca ccgcgacgtc	2880
	tgctgagaag tttctgatgc aaaagttcga cagcgtctcc gacctgatgc agctctcgga	2940
	ggcggaagaa tctcgtgctt tcagcttcca tgtaggaggc cgtggatatg tctgcgggt	3000
	aaatagctgc gccgatggtt tctacaaaga tcgttatgtt tatcggcact ttgcatcggc	3060
	cgcgctcccg attccggaag tgcttgacat tggggaattc agcgagagcc tgacctattg	3120
	catctcccgc cgtgcacagg gtgtcacgtt gcaagacctg cctgaaaccg aactgcccgc	3180
	gtttctgcag ccggtcgcgg aggccatgga tgcatcgtc gcggccgata ttagccagac	3240
	gagcgggttc ggcccattcg gaccgcaagg aatcgggtcaa tacactacat ggctgtattt	3300

catatgcgcg attgctgata cccatgtgta tcaactggcaa actgtgatgg acgacaccgt	3360
cagtgcgtcc gtcgcgagg ctctcgatga gctgatgctt tgggccgagg actgccccga	3420
agtccggcac ctctgacag cggatttcgg ctccaacaat gtcctgacgg acaatggccg	3480
cataacagcg gtcattgact ggagcgaggc gatgttcggg gattcccaat acgaggctgc	3540
caacatcttc ttctggaggc cgtggttggc ttgtatggag cagcagacgc gctacttcga	3600
gcgaggcat ccggagcttg caggatcgcc gcggtccgg gcgtatatgc tccgattgg	3660
tcttgaccaa ctctatcaga gcttggttga cggcaatttc gatgatgcag cttgggcgca	3720
gggtcgatgc gacgcaatcg tccgatccgg agccgggact gtcggcgta cacaaatgc	3780
ccgcagaagc gcggccgtct ggaccgatgg ctgtgtagaa gtactcgccg atagtggaaa	3840
ccgacgccc agcactcgtc cgagggcaaa ggaatagcac gtgctacgag atttcgattc	3900
caccgcccc ttctatgaaa ggttgggctt cggaaatcgtt ttccgggacg ccggtggat	3960
gatcctccag cgcggggatc tcatgctgga gttcttcgcc caccccaact tgtttattgc	4020
agcttataat ggttacaaat aaagcaatag catcacaat ttacaaaata aagcattttt	4080
ttcactgcat tctagtgtg gtttgcctaa actcatcaat gtatcttata atgtctgtat	4140
accgtcgacc tctagctaga gcttggcgta atcatggtca tagctgttcc ctgtgtgaaa	4200
ttgttatecg ctcaaatc cacacaacat acgagccgga agcataaagt gtaaagcctg	4260
gggtgcctaa tgagtgagct aactcacatt aattgcgttgc cgtcactgc ccgctttcca	4320
gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg	4380
[0015] tttgcgtatt gggcgtctt ccgcttctc gctcactgac tcgctgcgt cggtcgttcg	4440
gctcggcga gcggtatcag ctcaactcaa ggcggttaata cggttatcca cagaatcagg	4500
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa	4560
ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcacc acaaaaatcg	4620
acgtcaagt cagagtggtc gaaaccgac agactataa agataccagg cgtttcccc	4680
tggaagctcc ctctgctgct ctctgttcc gacctgccg cttaccgat acctgtccgc	4740
ctttctcct tcggaagcg tggcgcttcc tcaatgctca cgtctaggt atctcagttc	4800
ggtgtaggtc gttcctcca agctgggctg tgtgcacgaa cccccgttc agcccaccg	4860
ctgcgcctta tccgtaact atcgtcttga gtccaaccg gtaagacacg acttatcgcc	4920
actggcagca gccactggtc acagattag cagagcgagg tatgtaggcg gtgctacaga	4980
gttcttgaag tgggtgccta actacggcta cactagaagg acagtatttg gtatctgcgc	5040
tctgctgaag ccagttacct tcgaaaaag agttggtagc tcttgatccg gcaaaaaac	5100
caccgctggt agcgggtggt tttttgttg caagcagcag attacgcga gaaaaaagg	5160
atctcaagaa gatccttga tctttctac ggggtctgac gctcagtga acgaaaactc	5220
acgtaagggt attttggca tgagattatc aaaaaggatc ttcacctaga tctttttaa	5280
ttaaaaatga agttttaa atcaatctaaag tatatatgag taaacttggc ctgacagtta	5340
ccaatgctta atcagtggc cacctatctc agcgatctgt ctatttcgtt catcatagt	5400
tgctgactc cccgtcgtg agataactac gatacgggag ggcttaccat ctggccccag	5460
tgctgcaatg ataccgag acccagctc accggctcca gatttatcag caataacca	5520

```

gccagccgga agggccgagc gcagaagtgg tctgcaact ttatccgcct ccatccagtc 5580
tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt 5640
tgttgccatt gctacagcca tcgtggtgtc acgctcgtcg ttggtatgg cttcattcag 5700
ctccggttcc caacgatcaa ggcgagtac atgatcccc atgttgtgca aaaaagcgg 5760
tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcaactat 5820
ggttatggca gcactgcata attctcttac tgatcatgcca tccgtaagat gcttttctgt 5880
gactggtgag tactcaacca agtcattctg agaatagttg atgcccgcac cgagttgctc 5940
ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa aagtgtctat 6000
cattgaaaa cgttcttcgg ggcgaaaa ctcaaggatc ttaccgctgt tgagatccag 6060
ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt 6120
ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg 6180
gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta 6240
ttgtctcatg agcgatatac tatttgaatg tatttagaaa aataaaciaa taggggttcc 6300
gcgcacattt ccccgaaaag tgccacctga cgtc 6334
    
```

<210> 15  
 <211> 150  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> Ssp GyrB 11分裂内含肽的N-内含肽结构域的氨基酸序列

[0016]

<400> 15

```

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



Asn Phe Ala Leu Ala Ser  
145 150

<210> 16  
<211> 450  
<212> DNA  
<213> 人工序列

<220>  
<223> 使用哺乳类密码子对SEQ ID NO 15反向翻译得到的Ssp GyrB 11 分裂内含肽的N-内含肽结构域的编码区域

<400> 16  
tgcttcagcg gcgacaccct ggtggccctg accgacggca gaagcgtgag cttcgagcag 60  
ctggtggagg aggagaagca gggcaagcag aacttctgct acaccatcag acacgacggc 120  
agcatcgcg tggagaagat catcaacgcc agaaagacca agaccaacgc caagtgatc 180  
aaggtgaccc tggacaacgg cgagagcadc atctgcaccc ccgaccacaa gttcatgctg 240  
agagacggca gctacaagtg cgccatggac ctgaccctgg acgacagcct gatgccctg 300  
cacagaaaga tcagcaccac cgaggacagc ggccacatgg aggccgtgct gaactacaac 360  
cacagaatcg tgaacatcga gcccgtgagc gagaccatcg acgtgtacga catcgaggtg 420  
ccccacacc acaacttcgc cctggccagc 450

<210> 17  
<211> 1905  
<212> DNA  
<213> 人工序列

[0017]

<220>  
<223> 具有包含Ssp GyrB 11 分裂内含肽的N-内含肽结构域及其后的6xHis标记和strepII标记的C端延长部的抗人CD19抗体hBU12的全IgG1重链编码区域

<400> 17  
atgaattttg gactgaggct gattttcctg gtgctgaccc tgaagagcgt ccagtgctcag 60  
gttcagctgc aagagtctgg ccctggggtg gttaaaccct cccagaccct cagtctgact 120  
tgtactgtgt ctgggggttc aatcagcact tctggtatgg gtgtaggctg gattaggcag 180  
caccaggga agggctctga gtggattgga cacatttggg gggatgatga caagagatat 240  
aaccagccc tgaagagcag agtgacaatc tctgtgata cctccaagaa ccagtttagc 300  
ctcaagctgt ccagtgtagc agctgcagat actgctgtct actactgtgc tagaatggaa 360  
ctttggtcct actattttga ctactggggc caaggcacc ttgtcagct ctctcagct 420  
agcaccaagg gccatctgt ctccccctg gcaeccctct ccaagagcac ctctgggggc 480  
acagctgcc tgggctgect ggtaaggac tacttccctg aacctgtgac agtgctctgg 540  
aactcaggcg cctgaccag cggcgtgac acctcccgg ctgtctaca gtctcagga 600  
ctctactccc tcagcagct ggtgaccgtg ccctccagca gcttggggcac ccagacctac 660  
atctgcaacg tgaatcacia gccagcaac accaaggtgg acaagaaagt tgagcccaaa 720  
tcttgtgaca aaactcaac atgcccaccg tgcccagcac ctgaactcct ggggggaccg 780  
tcagttctcc tcttcccccc aaaaccaag gacacctca tgatctccc gaccctgag 840  
gtccatgctg tgggtgtgga cgtgagccac gaagaccctg aggtcaagtt caactgttac 900  
gtggacggcg tggaggtgca taatgccaag acaaaagccgc gggaggagca gtacaacagc 960  
acgtaccgtg tggtagcgt cctcaccgtc ctgcaccag actgctgaa tggcaaggag 1020

```

tacaagtga aggtctccaa caaagccctc ccagccccc tcgagaaaac catctccaaa    1080
gccaaagggc agccccgaga accacaggtg tacaccctgc ccccatcccg ggatgagctg    1140
accaagaacc aggtcagcct gacctgcctg gtcaaaggct tctatcccag cgacatcgcc    1200
gtggagtggg agagcaatgg gcagccggag aacaactaca agaccacgcc tcccgtgetg    1260
gactccgacg gctccttctt cctctacagc aagctcaccg tggacaagag caggtggcag    1320
caggggaacg tcttctcatg ctccgtgatg catgaggctc tgcacaacca ctacacacag    1380
aagacctctt cctgtctccc ggtataatgc ttcagcggcg acacctggt ggcctgacc    1440
gacggcagaa gcgtgagctt cgagcagctg gtggaggagg agaagcaggg caagcagaac    1500
ttctgtctac ccatcagaca cgacggcagc atcggcgtgg agaagatcat caacccaga    1560
aagaccaaga ccaacgcca ggtgatcaag gtgaccctgg acaacggcga gagcatcatc    1620
tgacccccg accacaagtt catgctgaga gacggcagct acaagtgcgc catggacctg    1680
acctggagc acagcctgat gccctgcac agaaagatca gcaccaccga ggacagcggc    1740
cacatggagg ccgtgctgaa ctacaaccac agaatcgtga acatcgagge cgtgagcgag    1800
accatcgacg tgtacgacat cgaggtgccc cacaccaca acttcgccct ggccagccac    1860
caccaccatc accatggctg gagccacccc cagttcgaga agtag                    1905
    
```

```

<210> 18
<211> 634
<212> PRT
<213> 人工序列
    
```

[0018]

```

<220>
<223> SEQ ID NO 17翻译所得的氨基酸序列
<400> 18
    
```

```

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly
1           5           10          15
    
```

```

Val Gln Cys Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys
                20           25           30
    
```

```

Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile
                35           40           45
    
```

```

Ser Thr Ser Gly Met Gly Val Gly Trp Ile Arg Gln His Pro Gly Lys
50           55           60
    
```

```

Gly Leu Glu Trp Ile Gly His Ile Trp Trp Asp Asp Asp Lys Arg Tyr
65           70           75           80
    
```

```

Asn Pro Ala Leu Lys Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys
                85           90           95
    
```

```

Asn Gln Phe Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala
                100          105          110
    
```

```

Val Tyr Tyr Cys Ala Arg Met Glu Leu Trp Ser Tyr Tyr Phe Asp Tyr
                115          120          125
    
```

```

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly
                130          135          140
    
```

Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly  
 145 150 155 160  
 Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val  
 165 170 175  
 Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe  
 180 185 190  
 Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val  
 195 200 205  
 Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val  
 210 215 220  
 Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys  
 225 230 235 240  
 Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu  
 245 250 255  
 Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr  
 260 265 270  
 Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val  
 275 280 285  
 [0019] Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val  
 290 295 300  
 Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser  
 305 310 315 320  
 Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu  
 325 330 335  
 Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala  
 340 345 350  
 Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro  
 355 360 365  
 Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln  
 370 375 380  
 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala  
 385 390 395 400  
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr  
 405 410 415  
 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu  
 420 425 430  
 Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser

435 440 445

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser  
450 455 460

Leu Ser Pro Gly Lys Cys Phe Ser Gly Asp Thr Leu Val Ala Leu Thr  
465 470 475 480

Asp Gly Arg Ser Val Ser Phe Glu Gln Leu Val Glu Glu Lys Gln  
485 490 495

Gly Lys Gln Asn Phe Cys Tyr Thr Ile Arg His Asp Gly Ser Ile Gly  
500 505 510

Val Glu Lys Ile Ile Asn Ala Arg Lys Thr Lys Thr Asn Ala Lys Val  
515 520 525

Ile Lys Val Thr Leu Asp Asn Gly Glu Ser Ile Ile Cys Thr Pro Asp  
530 535 540

His Lys Phe Met Leu Arg Asp Gly Ser Tyr Lys Cys Ala Met Asp Leu  
545 550 555 560

Thr Leu Asp Asp Ser Leu Met Pro Leu His Arg Lys Ile Ser Thr Thr  
565 570 575

Glu Asp Ser Gly His Met Glu Ala Val Leu Asn Tyr Asn His Arg Ile  
580 585 590

Val Asn Ile Glu Ala Val Ser Glu Thr Ile Asp Val Tyr Asp Ile Glu  
595 600 605

Val Pro His Thr His Asn Phe Ala Leu Ala Ser His His His His His  
610 615 620

His Gly Trp Ser His Pro Gln Phe Glu Lys  
625 630

[0020]

<210> 19  
<211> 1206  
<212> DNA  
<213> 人工序列

<220>  
<223> 具有包含Ssp GyrB 11 分裂内含肽的N-内含肽结构域及其后的6xHis标记和strepII标记的C端延长部的抗人CD19抗体hBU12全IgG1 k轻链编码区域

<400> 19  
atgaattttg gactgaggct gattttcctg gtgctgacct tgaaggcgt ccagtgtgac 60  
attgtgctga cccaatctcc agcttctttg gctgtgtctc tagggcagag ggccaccatc 120  
tcctgcaagg ccagccaaag tgttgatttt gatggtgata gttatatgaa ctggtaccaa 180  
cagaaaccag gacagccacc caaagtctct atctatgctg catccaatct agaatctggg 240  
atcccagcca ggtttagtgg cagtgggtct gggacagact tcacctcaa catccatcct 300  
gtggaggagg aggatgctgc aacctattac tgtcagcaaa gtaatgagga tccgtggagc 360  
ttcgtgggag gcaccaagct gaaatcaaa cgtacggtgg ctgcaccatc tgtcttcac 420

```

ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat      480
aacttctatc ccagagaggc caaagtacag tggaaggtgg ataacgcctt ccaatcgggt      540
aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc      600
accttgacgc tgagcaaagc agactacgag aaacacaaag tctacgcctg cgaagtcacc      660
catcagggcc tgagctcgcc cgtcacaaag agcttcaaca ggggagagtg cttcagcggc      720
gacaccctgg tggccctgac cgacggcaga agcgtgagct tcgagcagct ggtggaggag      780
gagaagcagg gcaagcagaa cttctgctac accatcagac acgacggcag catcggcgtg      840
gagaagatca tcaacgccag aaagaccaag accaacgcca aggtgatcaa ggtgaccctg      900
gacaacggcg agagcatcat ctgcaccccc gaccacaagt tcatgctgag agacggcagc      960
tacaagtgcg ccatggacct gaccctggac gacagcctga tgcccctgca cagaaagatc     1020
agcaccaccg aggacagcgg ccacatggag gccgtgctga actacaacca cagaatcgtg     1080
aacatcgagg ccgtgagcga gaccatcgac gtgtacgaca tcgaggtgcc ccacaccac     1140
aacttcgecc tggccagcca ccatcaccat caccatggct ggagccaccc ccagttcgag     1200
aagtag                                             1206
    
```

```

<210> 20
<211> 401
<212> PRT
<213> 人工序列
    
```

[0021]

```

<220>
<223> SEQ ID NO 19翻译所得的氨基酸序列
    
```

```

<400> 20
    
```

```

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly
1           5           10           15
    
```

```

Val Gln Cys Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val
                20           25           30
    
```

```

Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val
          35           40           45
    
```

```

Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly
          50           55           60
    
```

```

Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly
65           70           75           80
    
```

```

Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu
          85           90           95
    
```

```

Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln
          100          105          110
    
```

```

Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu
          115          120          125
    
```

```

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser
          130          135          140
    
```

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160  
 Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
 165 170 175  
 Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
 180 185 190  
 Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
 195 200 205  
 Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
 210 215 220  
 Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Phe Ser Gly  
 225 230 235 240  
 Asp Thr Leu Val Ala Leu Thr Asp Gly Arg Ser Val Ser Phe Glu Gln  
 245 250 255  
 Leu Val Glu Glu Glu Lys Gln Gly Lys Gln Asn Phe Cys Tyr Thr Ile  
 260 265 270  
 Arg His Asp Gly Ser Ile Gly Val Glu Lys Ile Ile Asn Ala Arg Lys  
 275 280 285  
 Thr Lys Thr Asn Ala Lys Val Ile Lys Val Thr Leu Asp Asn Gly Glu  
 290 295 300  
 Ser Ile Ile Cys Thr Pro Asp His Lys Phe Met Leu Arg Asp Gly Ser  
 305 310 315 320  
 Tyr Lys Cys Ala Met Asp Leu Thr Leu Asp Asp Ser Leu Met Pro Leu  
 325 330 335  
 His Arg Lys Ile Ser Thr Thr Glu Asp Ser Gly His Met Glu Ala Val  
 340 345 350  
 Leu Asn Tyr Asn His Arg Ile Val Asn Ile Glu Ala Val Ser Glu Thr  
 355 360 365  
 Ile Asp Val Tyr Asp Ile Glu Val Pro His Thr His Asn Phe Ala Leu  
 370 375 380  
 Ala Ser His His His His His His Gly Trp Ser His Pro Gln Phe Glu  
 385 390 395 400

[0022]

Lys

<210> 21  
 <211> 7447  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> C端具有Ssp GyrB S11 分裂内含肽N-内含肽结构域及其后的6xHis标记, strepII标记, 以及HindIII和NotI克隆位点

的hBU12人IgG1 VH-CH 重链编码区域

<400> 21  
gacggatcgg gagatctccc gatcccctat ggctgactct cagtacaatc tgctctgatg 60  
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgtc gagtagtgcg 120  
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180  
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240  
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300  
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360  
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc 420  
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480  
atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540  
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600  
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660  
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720  
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780  
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840  
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc 900  
gtttaaactt aagcttccat gaattttgga ctgaggctga ttttctggt gctgacctg 960  
aaaggcgtcc agtgtcaggt tcagctgcaa gactctggcc ctgggttggg taagccctcc 1020  
[0023] cagaccctca gtctgactg tactgtgtct gggggttcaa tcagcacttc tggatgggt 1080  
gtaggctgga ttaggcagca cccaggggag ggtctggagt ggattggaca catttgggtg 1140  
gatgatgaca agagatataa cccagccctg aagagcagag tgacaatctc tgtggatacc 1200  
tccaagaacc agtttagcct caagctgtcc agtgtgacag ctgcagatac tgctgtctac 1260  
tactgtgcta gaatggaact ttggtcctac tattttgact actggggcca aggcaccctt 1320  
gtcacagtct cctcagctag caccaagggc ccatctgtct tccccctggc accctctcc 1380  
aagagcacct ctgggggcac agctgccctg ggctgcctgg tcaaggacta cttccctgaa 1440  
cctgtgacag tgtcctggaa ctcagggccc ctgaccagcg gcgtgcacac cttcccggct 1500  
gtctctacagt cctcaggact ctactcctc agcagcgtgg tgaccgtgcc ctccagcage 1560  
ttgggcaccc agacctacat ctgcaacgtg aatcacaagc ccagcaaac caaggtggac 1620  
aagaaagtgt agccc aaatc ttgtgacaaa actcacacat gccaccgtg cccagcacct 1680  
gaactctgg ggggaccgtc agtcttctc ttcccccaa aaccaagga caccctcatg 1740  
atctcccga cccctgaggt cacatgcgtg gtggtggacg tgagccacga agaccctgag 1800  
gtcaagttca actggtacgt ggacggcgtg gaggtgcata atgccaagac aaagcccgcg 1860  
gaggagcagt acaacagcac gtaccgtgtg gtcagcgtcc tcaccgtcct gcaccaggac 1920  
tggctgaatg gcaaggagta caagtgaag gtctccaaca aagccctccc agcccccatc 1980  
gagaaaacca tctccaaagc caaaggcag ccccgagaac cacaggtgta caccctgcc 2040  
ccatcccggg atgagctgac caagaaccag gtcagcctga cctgcctggt caaaggcttc 2100  
tatcccagcg acatcgccgt ggagtgggag agcaatgggc agccggagaa caactacaag 2160

accacgcctc ccgtgctgga ctccgacggc tccttcttcc tctacagcaa gctcaccgtg	2220
gacaagagca ggtggcagca ggggaacgtc ttctcatgct ccgigtatgca tgaggctctg	2280
cacaaccact acacacagaa gagcctctcc ctgtctccgg gtaaagtctt cagcggcgac	2340
accctggtgg ccctgaccga cggcagaagc gtgagcttcg agcagctggt ggaggaggag	2400
aagcagggca agcagaactt ctgctacacc atcagacacg acggcagcat cggcgtggag	2460
aagatcatca acgccagaaa gaccaagacc aacgccaagg tgatcaaggt gaccctggac	2520
aacggcgaga gcatcatctg cacccccacg cacaagtcca tgctgagaga cggcagctac	2580
aagtgcgcca tggacctgac cctggacgac agcctgatgc ccctgcacag aaagatcagc	2640
accaccgagg acagcggcca catggaggcc gtgtgtaact acaaccacag aatcgtgaac	2700
atcgaggccc tgagcgagac catcgacgtg tacgacatcg aggtgcccca caccacaac	2760
ttcggcctgg ccagccacca tcaccatcac catggctgga gccaccccca gttcgagaag	2820
taggcggccc ctcgagtcta gagggcccgt ttaaaccgcg tgatcagcct cgactgtgcc	2880
ttctagtgc cagccatctg ttgtttgccc ctccccgtg ccttccttga ccctggaagg	2940
tgccactccc actgtccttt cctaataaaa tgaggaaatt gcatcgatt gtctgagtag	3000
gtgtcattct attctggggg gtgggtggg gcaggacagc aagggggagg attggaaga	3060
caatagcagg catgctgggg atgcggtggg ctctatggct tctgaggcgg aaagaaccag	3120
ctggggctct agggggatc cccacgcgcc ctgtagcggc gcattaagcg cggcgggtgt	3180
ggtggttacg cgcagcgtga ccgctacact tgccagcggc cttagcggcc ctcctttcgc	3240
tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg	3300
catcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta	3360
gggtgatggt tcacgtatg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt	3420
ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat	3480
ctcgtctat tcttttgatt tataagggat tttggggatt tcggcctatt ggttaaaaa	3540
tgagctgatt taacaaaaa ttaacgcgaa ttaattctgt ggaatgtgtg tcagttaggg	3600
tgtgaaagt cccagcctc cccagcgag cagaagtatg caaagcatgc atctcaatta	3660
gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaagcat	3720
gcatctcaat tagtcagcaa ccatagtcgc gccctaact ccgcccaccc cggcccctaac	3780
tcgcccagc tcgcccatt ctccgcccc tggtgacta atttttttta tttatgcaga	3840
ggccgaggcc gcctctgct ctgagctatt ccagaagtag tgaggaggct tttttggagg	3900
cctagctctt tgcaaaaage tcccgggagc ttgtatatcc attttcggat ctgatcagca	3960
cgtgatgaaa aagcctgaac tcaccgcgac gtctgtcgag aagtttctga tcgaaaagtt	4020
cgacagcgtc tccgacctga tgcagctctc ggagggcgaa gaatctctgt ctttcagctt	4080
cgatgtagga gggcgtggat atgtcctcgc ggtaaatagc tgcgccgatg gtttctacaa	4140
agatcgttat gtttatcgcc actttgcacg gcccgcgctc ccgattccgg aagtcttga	4200
cattggggaa ttcagcgaga gcctgacctt ttgcatctcc cggcgtgcac aggtgtcac	4260
gttcaagac ctgcctgaaa ccgaactgcc cgctgttctg cagccggctc cggaggccat	4320
ggatgcgac gctgcggccc atcttagcca gacgagcggg ttcggcccat tcggaccgca	4380
aggaatcggc caatacacta catggcgtga tttcatatgc gcgattgctg atccccatgt	4440

[0024]



gtatcactgg caaactgtga tggacgacac cgtcagtgcg tccgtcgcgc aggctctcga	4500
tgagctgatg ctttgggccg aggactgccc cgaagtccgg cacctcgtgc acgcggattt	4560
cggctccaac aatgtcctga cggacaatgg ccgcataaca gcggtcattg actggagcga	4620
ggcgatgttc ggggattccc aatacagagt cgccaacatc ttcttctgga ggccgtggtt	4680
ggcttgtatg gagcagcaga cgcgctactt cgagcggagg catccggagc ttgcaggatc	4740
gcccgggctc cgggcgtata tgctccgcat tggctttgac caactctatc agagcttgg	4800
tgacggcaat ttcgatgatg cagcttgggc gcagggtcga tgcgacgcaa tcgtccgatc	4860
cggagccggg actgtcgggc gtacacaaat cgcccgcaga agcgcggccg tctggaccga	4920
tggtctgtga gaagtactcg ccgatagtgg aaaccgacgc cccagcactc gtccgagggc	4980
aaaggaatag cacgtgctac gagatttcca ttccaccgcc gccttctatg aaagttggg	5040
cttcggaatc gttttccggg acgccggtg gatgatctc cagcgcgggg atctcatgct	5100
ggagttcttc gcccaccca acttgtttat tgcagcttat aatggttaca aataaagcaa	5160
tagcatcaca aatttcacaa ataaagcatt ttttctactg cattctagtt gtggtttgtc	5220
caaactcadc aatgtatctt atcatgtctg tataccgtcg acctctagct agagcttggc	5280
gtaatcatg tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia	5340
catacagacc ggaagcataa agtgtaaagc ctgggggtgc taatgagtga gctaactcac	5400
attaattcgg ttgcgtcac tgcccgttt ccagtcggga aacctgtcgt gccagctgca	5460
ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgt cttccgcttc	5520
[0025] ctgctcact gactcgtgc gctcgtcgt tcggctcgg cgagcggat cagctcactc	5580
aaagcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc	5640
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag	5700
gctccgccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaacc	5760
gacaggacta taaagatacc aggcgttcc ccctggaagc tccctcgtgc gctctctgt	5820
tccaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct	5880
tttcaatg tcacgtgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg	5940
ctgtgtgac gaacccccg ttcagcccga ccgctgcgcc ttatccgta actatcgtct	6000
tgagtccaac ccgtaagac acgacttadc gccactgca gcagccactg gtaacaggat	6060
tagcagagcg aggtatgtag gcggtgctac agagtcttg aagtgggtgc ctaactacgg	6120
ctacactaga aggacagtat ttggtatctg cgtctgctg aagccagta ccttcgaaa	6180
aagagttggt agctcttgat ccggcaaaaca aaccaccgct ggtagcggtg gttttttgt	6240
ttgcaagcag cagattacgc gcagaaaaa aggatctcaa gaagatcctt tgatctttc	6300
tacgggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt	6360
atcaaaaagg atcttcacct agatcctttt aaattaataa tgaagtttta aatcaatcta	6420
aagtatatat gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg aggcacctat	6480
ctcagcgtac tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac	6540
tacgatacgg gagggcttac catctggccc cagtgtgca atgataccgc gagaccacg	6600
ctcaccgctc ccagatttat cagcaataaa ccagccagcc ggaaggccg agcgcagaag	6660

```

tggtcctgca acittatecg cctccatcca gtctattaat tgttgccggg aagctagagt 6720
aagtagttcg ccagttaata gtttgcgcaa cgttggtgcc attgctacag gcatcgtggt 6780
gtcacgctcg tcgtttgta tggcttcatt cagctccggt tccaacgat caaggcgagt 6840
tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggctctc cgatcgttgt 6900
cagaagtaag ttggcccgag tgttatcaact catggttatg gcagcactgc ataattctct 6960
tactgtcatg ccaatecgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt 7020
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac 7080
cgcgccacat agcagaactt taaaagtctt catcattgga aaacgttctt cggggcgaaa 7140
actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa 7200
ctgatcttca gcactcttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7260
aaatgccgca aaaaagggaa taagggcgac acgaaaatgt tgaatactca tactcttctt 7320
tttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 7380
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 7440
tgacgtc 7447
    
```

```

<210> 22
<211> 6748
<212> DNA
<213> 人工序列
    
```

[0026]

```

<220>
<223> 具有C端Ssp GyrB S11 N-内含肽结构域, 6xHis标记和strepII标记, 以及HindIII和NotI克隆位点的hBU12的人IgG1 VL-CL k轻链编码区域
    
```

```

<400> 22
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggctcgt gagtagtgcg 120
cgagcaaaat ttaagtaca acaaggcaag gcttgaccga caattgatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc 360
cccgccatt gacgtcaata atgacgtatg tcccattagt aacgccaata gggactttcc 420
attgacgta atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gectggcatt 540
atgcccgta catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgicgta acaactccgc ccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc 900
gtttaaactt aagcttccat gaattttgga ctgaggctga ttttctggt gctgaccctg 960
aaaggcgtcc agtgtgacat tgtgctgacc caatctccag cttctttggc tgtgtcteta 1020
gggcagaggg ccaccatctc ctgcaaggcc agccaaagt ttgattttga tggatgatag 1080
    
```

tatatgaact ggtaccaaca gaaaccagga cagccacca aagtcctcat ctatgctgca	1140
tccaatctag aatctgggat cccagccagg tttagtggca gtgggtctgg gacagacttc	1200
accctcaaca tccatectgt ggaggaggag gatgctgcaa cctattactg tcagcaaagt	1260
aatgaggatc cgtggacgtt cgggtggaggc accaagctgg aatcaaacg tacggtggct	1320
gcaccatctg tcttcatctt cccccatct gatgagcagt tgaaatctgg aactgcctct	1380
gttgtgtgcc tgctgaataa cttctatccc agagaggcca aagtacagtg gaaggtggat	1440
aacgccctcc aatcgggtaa ctcccaggag agtgtcacag agcaggacag caaggacagc	1500
acctacagcc tcagcagcac cctgacgctg agcaaagcag actacgagaa acacaaagtc	1560
tacgcctgcg aagtcaacca tcagggcctg agctcgccc tcacaaagag cttcaacagg	1620
ggagagtgtc tcagcggcga caccctggtg gccctgaccg acggcagaag cgtgagcttc	1680
gagcagctgg tggaggagga gaagcagggc aagcagaact tctgctacac catcagacac	1740
gacggcagca tcggcgtgga gaagatcacc aacgccagaa agaccaagac caacccaag	1800
gtgatcaagg tgaccctgga caacggcgag agcatcatct gcacccccga ccacaagttc	1860
atgctgagag acggcagcta caagtgcgcc atggacctga ccctggacga cagcctgatg	1920
cccctgcaca gaaagatcag caccaccgag gacagcgccc acatggaggc cgtgctgaac	1980
tacaaccaca gaatcgtgaa catcgaggcc gtgagcgaga ccatcgactg gtacgacatc	2040
gaggtgcccc acaccacaa cttcgcctg gccagccacc atcacatca ccatggctgg	2100
agccaccccc agttcgagaa gtaggcggcc gctcgagtct agagggccc tttaaacccg	2160
ctgatcagcc tcgactgtgc ctctagtgt ccagccatct gttgtttgcc cctccccgt	2220
gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat	2280
tgcatcgcac tgtctgagta ggtgtcatc tattctgggg ggtggggtgg ggcaggacag	2340
caagggggag gattgggaag acaatagcag gcatgctggg gatgcgggtg gctctatggc	2400
ttctgagggc gaaagaacca gctggggctc tagggggtat ccccacgcgc cctgtagcgg	2460
cgcattaagc gcggcgggtg tgggtgttac gcgcagcgtg accgctacac ttgccagcgc	2520
cctagcggcc gctcctttcg ctttcttccc ttcttttctc gccacgttcg ccggtttcc	2580
ccgtaagct ctaaactcggg gcatcccttt agggttccga tttagtgtt tacggcacct	2640
cgaccccaaa aaacttgatt agggtagtgg ttcacgtagt gggccatcgc cctgatagac	2700
ggtttttcgc ctttgacgt tggagtccac gttctttaat agtggactct ttttccaaac	2760
tggaacaaca ctcaacccta tctcggctca ttcttttgat ttataaggga ttttggggat	2820
ttcgccctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attaattctg	2880
tggaatgtgt gtcagttagg gtgtggaag tccccaggct cccagcagc gcagaagtat	2940
gcaaagcatg catctcaatt agtcagcaac caggtgtgga aagtcccag gctcccagc	3000
aggcagaagt atgcaaagca tgcatctcaa ttagtcagca accatagtcc cgcacctaac	3060
tccgcccatc cgcacctaa ctccgccag ttccgccat tctccggccc atggctgact	3120
aattttttt atttatgcag aggccgaggc cgcctctgcc tctgagctat tccagaagta	3180
gtgaggagcc tttttggag gcctaggctt ttgcaaaaag ctcccgggag cttgtatatc	3240
cattttcggg tctgatcagc acgtgatgaa aaagcctgaa ctaccgcga cgtctgtcga	3300
gaagtttctg atcgaagaat tcgacagcgt ctccgacctg atgcagctct cggagggcga	3360

[0027]

	agaattctcgt gctttcagct tcgatgtagg agggcgtgga tatgtcctgc gggtaaataag	3420
	ctgcccgat ggtttctaca aagatcgta tgtttatcgg cactttgcat cggccgcgct	3480
	cccgattccg gaagtgcttg acattgggga attcagcgag agcctgacct attgcatctc	3540
	ccgcccgtca cagggtgtca cgttgcaaga cctgcctgaa accgaactgc ccgctgttct	3600
	gcagccggtc gcggaggcca tggatcgat cgctgcggcc gatcttagcc agacgagcgg	3660
	gttcggccca ttcggaccgc aaggaatcgg tcaatacact acatggcgtg atttcatatg	3720
	cgcgattgct gatccccatg tgtatcactg gcaaaactgtg atggacgaca ccgtcagtgc	3780
	gtccgtcgcg caggctctcg atgagctgat gctttgggcc gaggactgcc ccgaagtccg	3840
	gcacctcgtg cacgcggatt tcggctccaa caatgtcctg acggacaatg gccgcataac	3900
	agcggctcatt gactggagcg aggcgatgtt cggggattcc caatacagag tcgccaacat	3960
	cttcttctg aggccctggt tggcttgtat ggagcagcag acgcgctact tcgagcggag	4020
	gcattccggag cttgcaggat cggccggct cggggctat atgctccgca ttggtcttga	4080
	ccaactctat cagagcttgg ttgacggcaa tttcgatgat gcagcttggg cgcagggtcg	4140
	atgcgacgca atcgtccgat ccggagccgg gactgtcggg cgtacacaaa tcgcccgcag	4200
	aagcgcggcc gtctggaccg atggctgtgt agaagtactc gccgatagtg gaaaccgacg	4260
	ccccagcact cgtccgaggg caaaggaata gcacgtgcta cgagatttcg attccaccgc	4320
	cgcttctat gaaaggttgg gcttcggaat cgttttcggg gacgccgctt ggatgatcct	4380
	ccagcgcggg gatctcatgc tggagtctt cgeccacccc aacttgttta ttgcagctta	4440
[0028]	taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat ttttttact	4500
	gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct gtataccgtc	4560
	gacctctagc tagagcttgg cgtaatcatg gtcatactg tttctgtgt gaaattgtta	4620
	tccgctcaca attccacaca acatacagc cggaagcata aagtgtaaag cctggggtgc	4680
	ctaatagagt agctaactca cattaattgc gttgcgctca ctgcccgtt tccagtcggg	4740
	aaacctgtc tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg	4800
	tattgggcgc tcttcgctt cctcgtctac tgactcgtg cgctcgtg ttcgctgcg	4860
	gcgagcggta tcagctcact caaagcggg aatacgtta tccacagaat caggggataa	4920
	cgcaggaag aacatgtgag caaaagcca gaaaaggcc aggaaccgta aaaagcccgc	4980
	gttgctggcg tttttcata ggctccgcc ccctgacgag catcacaaaa atcgacgctc	5040
	aagtcagagg tggcgaacc cgacaggact ataaagatac caggcgtttc ccctggaag	5100
	ctcctcgtg cgctctctg ttcgacct gccgcttacc ggatactgt ccgcctttct	5160
	cccttcggga agcgtggcgc tttctcaatg ctacagctgt aggtatctca gttcgggtga	5220
	ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagccc accgctgcgc	5280
	cttatccgtt aactatcgtc ttgagtcaca cccgtaaga cacgacttat cgccactggc	5340
	agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt	5400
	gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct	5460
	gaagccagtt accttcgaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc	5520
	tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca	5580

agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta	5640
agggattttg gtcgatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa	5700
atgaagtfff aaatcaatct aaagtatata tgagtaaact tggictgaca gttaccaatg	5760
cttaatcagt gaggcaccta tctcagcgat ctgtctatff cgttcaccca tagttgcctg	5820
actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc ccagtgctgc	5880
aatgataccg cgagaccac gctcaccggc tccagattta tcagcaataa accagccagc	5940
cggaagggcc gagcgagaa gtggctctgc aactttatcc gcctccatcc agtctattaa	6000
ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttg	6060
cattgctaca ggcacgtgg tgcacgctc gtcgtttggt atggcttcat tcagctccgg	6120
ttccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc	6180
cttcggtect ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat	6240
ggcagcactg cataattctc ttactgtcat gccatccgta agatgctfff ctgtgactgg	6300
tgagtactca accaagtcac tetgagaata gtgtatcgcg cgaccgagtt gctcttgccc	6360
ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg	6420
aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat	6480
gtaacccact cgtgcacca actgatcttc agcatctfff actttacca gcgtttctgg	6540
gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagga ataagggcga cacggaatg	6600
ttgaatactc ataactttcc tttttcaata ttattgaagc atttatcagg gttattgtct	6660
[0029] catgagcggg tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac	6720
atttccccga aaagtccac ctgacgtc	6748

- <210> 23
  - <211> 206
  - <212> PRT
  - <213> 人工序列
  - <220>
  - <223> 来自金黄色葡萄球菌 (Staphylococcus aureus) 的分选酶A氨基酸序列
  - <400> 23
- Met Lys Lys Trp Thr Asn Arg Leu Met Thr Ile Ala Gly Val Val Leu  
 1                    5                    10                    15
- Ile Leu Val Ala Ala Tyr Leu Phe Ala Lys Pro His Ile Asp Asn Tyr  
                   20                    25                    30
- Leu His Asp Lys Asp Lys Asp Glu Lys Ile Glu Gln Tyr Asp Lys Asn  
                   35                    40                    45
- Val Lys Glu Gln Ala Ser Lys Asp Lys Lys Gln Gln Ala Lys Pro Gln  
                   50                    55                    60
- Ile Pro Lys Asp Lys Ser Lys Val Ala Gly Tyr Ile Glu Ile Pro Asp  
 65                    70                    75                    80
- Ala Asp Ile Lys Glu Pro Val Tyr Pro Gly Pro Ala Thr Pro Glu Gln  
                   85                    90                    95

Leu Asn Arg Gly Val Ser Phe Ala Glu Glu Asn Glu Ser Leu Asp Asp  
 100 105 110

Gln Asn Ile Ser Ile Ala Gly His Thr Phe Ile Asp Arg Pro Asn Tyr  
 115 120 125

Gln Phe Thr Asn Leu Lys Ala Ala Lys Lys Gly Ser Met Val Tyr Phe  
 130 135 140

Lys Val Gly Asn Glu Thr Arg Lys Tyr Lys Met Thr Ser Ile Arg Asp  
 145 150 155 160

Val Lys Pro Thr Asp Val Gly Val Leu Asp Glu Gln Lys Gly Lys Asp  
 165 170 175

Lys Gln Leu Thr Leu Ile Thr Cys Asp Asp Tyr Asn Glu Lys Thr Gly  
 180 185 190

Val Trp Glu Lys Arg Lys Ile Phe Val Ala Thr Glu Val Lys  
 195 200 205

<210> 24  
 <211> 621  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 编码SEQ ID NO 23的核苷酸序列

[0030]

<400> 24  
 atgaaaaaat ggacaaatcg attaatgaca atcgcctggtg tggacttat cctagtgcca 60  
 gcatatttgt ttgctaaacc acatatcgat aattatcttc acgataaaga taaagatgaa 120  
 aagattgaac aatatgataa aaatgtaaaa gaacaggcga gtaaagataa aaagcagcaa 180  
 gctaaacctc aaattccgaa agataaatcg aaagtggcag gctatattga aattccagat 240  
 gctgatatta aagaaccagt atatccagga ccagcaacac ctgaacaatt aaatagaggt 300  
 gtaagctttg cagaagaaaa tgaatcacta gatgatcaaa atatttcaat tgcaggacac 360  
 actttcattg accgtccgaa ctatcaattt acaaatctta aagcagccaa aaaaggtagt 420  
 atggtgtact ttaaagttgg taatgaaaca cgtaaagtata aaatgacaag tataagagat 480  
 gttaagccta cagatgtagg agttctagat gaacaaaaag gtaaagataa acaattaaca 540  
 ttaattactt gtgatgatta caatgaaaag acagcgcttt gggaaaaacg taaaatcttt 600  
 gtagctacag aagtcaaata a 621

<210> 25  
 <211> 465  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 具有6xHis标记的金黄色葡萄球菌 (Staphylococcus aureus) 分选酶A (aa60-205)版本的编码区域

<400> 25  
 atgcaagcta aacctcaaat tccgaaagat aaatcgaaag tggcaggcta tattgaaatt 60  
 ccagatgctg atattaaaga accagtatat ccaggaccag caacacctga acaattaaat 120  
 agagggtgaa gctttgcaga agaaaatgaa tcactagatg atcaaaatat ttcaattgca 180

ggacacactt tcattgaccg tccgaactat caatttaca atcttaaagc agccaaaaaa	240
ggtagtattg tgtactttaa agttggtaat gaaacacgta agtataaaat gacaagtata	300
agagatgta agcctacaga tgtaggagtt ctatgatgaac aaaaaggtaa agataaacaa	360
ttaacattaa ttacttgtga tgattacaat gaaaagacag gcgtttggga aaaacgtaaa	420
atctttgtag ctacagaagt caaacacat caccatcacc attaa	465
<210> 26	
<211> 154	
<212> PRT	
<213> 人工序列	
<220>	
<223> 氨基酸序列翻译所得的SEQ ID NO 25	
<400> 26	
Met Gln Ala Lys Pro Gln Ile Pro Lys Asp Lys Ser Lys Val Ala Gly	
1 5 10 15	
Tyr Ile Glu Ile Pro Asp Ala Asp Ile Lys Glu Pro Val Tyr Pro Gly	
20 25 30	
Pro Ala Thr Pro Glu Gln Leu Asn Arg Gly Val Ser Phe Ala Glu Glu	
35 40 45	
Asn Glu Ser Leu Asp Asp Gln Asn Ile Ser Ile Ala Gly His Thr Phe	
50 55 60	
[0031] Ile Asp Arg Pro Asn Tyr Gln Phe Thr Asn Leu Lys Ala Ala Lys Lys	
65 70 75 80	
Gly Ser Met Val Tyr Phe Lys Val Gly Asn Glu Thr Arg Lys Tyr Lys	
85 90 95	
Met Thr Ser Ile Arg Asp Val Lys Pro Thr Asp Val Gly Val Leu Asp	
100 105 110	
Glu Gln Lys Gly Lys Asp Lys Gln Leu Thr Leu Ile Thr Cys Asp Asp	
115 120 125	
Tyr Asn Glu Lys Thr Gly Val Trp Glu Lys Arg Lys Ile Phe Val Ala	
130 135 140	
Thr Glu Val Lys His His His His His His	
145 150	
<210> 27	
<211> 1401	
<212> DNA	
<213> 人工序列	
<220>	
<223> mAb Ac10重链的核苷酸序列	
<400> 27	
atgaattttg gactgaggct gattttcctg gtgctgacce tgaaggcgt ccagtgtcag	60
atccagetgc agcagtctgg ccccagggtc gtgaaacctg gcgcctccgt gaagatctcc	120
tgcaaggcct cggctacac cttcaccgac tactacatca cctgggtcaa gcagaagccc	180

ggccagggcc tggaatggat cggctggatc tateccggct cggcaacac caagtacaac	240
gagaagtcca agggcaaggc caccctgacc gtggacacct cctcttcac cgccttcacg	300
cagctgtcct ccctgacctc cgaggatacc gccgtgtact tctgcgcaa ctacggcaac	360
tattggttcg cctactgggg ccagggcaca caagtgaccg tgtctgtctc tagcaccaag	420
ggcccatcgg tcttccccct ggcaccctcc tccaagagca cctctggggg cacagcagcc	480
ctgggctgcc tggcaagga ctacttccc gaaccgggta cgggtctctg gaactcaggc	540
gccctgacca gggcgtgca caccttccc gctgtcttac agtctcagg actctactcc	600
ctcagcagcg tggtgaccgt gccctccagc agcttgggca cccagacctc catctgcaac	660
gtgaatcaca agcccagcaa caccaagggt gacaagaaag ttgagcccaa atcttgtgac	720
aaaactcaca catgcccacc gtgcccagca cctgaactcc tggggggacc gtcagtcttc	780
ctcttcccc caaaacccaa ggacacctc atgatctccc ggaccctga ggtcacatgc	840
gtggtggtgg acgtgagcca cgaagacct gaggtcaagt tcaactggtc cgtggacggc	900
gtggaggtgc ataatgcaa gacaagccg cgggaggagc agtacaacag cacgtacct	960
gtggtcagcg tcctaccct cctgcaccag gactggctga atggcaagga gtacaagtgc	1020
aaggtctcca acaagccct cccagcccc atcgagaaaa ccatctcaa agccaaagg	1080
cagccccgag aaccacaggt gtacacctg ccccatccc gggatgagct gaccaagaac	1140
caggtcagcc tgacctgct ggtcaaaggc ttctatccca gcgacatgc cgtggagtgg	1200
gagagcaatg ggcagccgga gaacaactac aagaccagc ctcccgtgct ggactccgac	1260
ggctccttct tcctctacag caagctcacc gtggacaaga gcaggtggca gcaggggaac	1320
gtcttctcat gctccgtgat gcatgaggct etgcacaacc actacacaca gaagagcctc	1380
tcctgtctc cgggtaaata g	1401

[0032]

<210> 28  
 <211> 466  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> MAB AC10重链的氨基酸序列

<400> 28

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly  
 1                    5                    10                    15

Val Gln Cys Gln Ile Gln Leu Gln Gln Ser Gly Pro Glu Val Val Lys  
                   20                    25                    30

Pro Gly Ala Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe  
                   35                    40                    45

Thr Asp Tyr Tyr Ile Thr Trp Val Lys Gln Lys Pro Gly Gln Gly Leu  
                   50                    55                    60

Glu Trp Ile Gly Trp Ile Tyr Pro Gly Ser Gly Asn Thr Lys Tyr Asn  
 65                    70                    75                    80

Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Thr Ser Ser Ser



	85	90	95
	Thr Ala Phe Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Thr Ala Val 100	105	110
	Tyr Phe Cys Ala Asn Tyr Gly Asn Tyr Trp Phe Ala Tyr Trp Gly Gln 115	120	125
	Gly Thr Gln Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val 130	135	140
	Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala 145	150	155
	Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser 165	170	175
	Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val 180	185	190
	Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro 195	200	205
	Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys 210	215	220
[0033]	Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp 225	230	235
	Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly 245	250	255
	Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile 260	265	270
	Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu 275	280	285
	Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His 290	295	300
	Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg 305	310	315
	Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys 325	330	335
	Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu 340	345	350
	Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr 355	360	365
	Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu 370	375	380

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp  
 385 390 395 400  
 Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val  
 405 410 415  
 Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp  
 420 425 430  
 Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His  
 435 440 445  
 Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro  
 450 455 460

Gly Lys  
 465

<210> 29  
 <211> 714  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> mAb Ac10轻链的核苷酸序列

[0034]

<400> 29  
 atgaattttg gactgagct gattttcttg gtgctgacc tgaaggcgt ccagtgtgac 60  
 attgtgctga cccaatctcc agcttctttg getgtgtctc tagggcagag ggccaccatc 120  
 tcctgcaagg ccagccaaag tgttgatttt gatggtgata gttatatgaa ctggtaccaa 180  
 cagaaaccag gacagccacc caaagtctc atctatgctg catccaatct agaactctggg 240  
 atcccagcca ggtttagtgg cagtgggtct gggacagact tcaccctcaa catccatcct 300  
 gtggaggagg aggatgctgc aacctattac tgtcagcaaa gtaatgagga tccgtggacg 360  
 ttcggtggag gcaccaagct ggaaatcaaa cgtacggtgg ctgcaccatc tgtcttcac 420  
 ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctggtgtgtg cctgctgaat 480  
 aactctatc ccagagagge caaagtacag tggaaggtgg ataacgcect ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 accctgacgc tgagcaaaag agactacgag aaacacaaaag tctacgctg cgaagtcacc 660  
 catcagggcc tgagctgcc cgtcacaaaag agcttcaaca ggggagagtg ttag 714

<210> 30  
 <211> 237  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> mAb Ac10轻链的氨基酸序列

<400> 30

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly  
 1 5 10 15  
 Val Gln Cys Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val  
 20 25 30

Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val  
 35 40 45

Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly  
 50 55 60

Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly  
 65 70 75 80

Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu  
 85 90 95

Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln  
 100 105 110

Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu  
 115 120 125

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
 130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
 165 170 175

[0035] Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
 180 185 190

Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
 195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
 210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys  
 225 230 235

<210> 31

<211> 1422

<212> DNA

<213> 人工序列

<220>

<223> 曲妥单抗重链的核苷酸序列

<400> 31

atgaaaaaaaa acattgcctt tctgctggcc tccatgttcg tgttctctat cgccacaaac 60

gcttacgctg aagtccagct ggtcgaatct ggtggtggcc tggccagcc tggatgatca 120

ctgagactgt cctgtgctgc ttctggtttc aacatcaagg acacctacat ccattgggtc 180

agacaggcac ctggcaaggg actggaatgg gtcgcccga tctaccctac aaacggctac 240

actcgctacg ccgactccgt caagggacgc tttaccatct ccgcgacac ctctaaaaac 300

accgcctacc tgcagatgaa tagtctgagg gccgaggata ctgctgtgta ctactgctca 360

cgatggggag gcgacggctt ttacgctatg gattactggg gacagggaac tctgtcact 420

gtgtctagcg cttagcacc aa gggcccatcg gtcttcccc tggcacccctc ctccaagagc 480  
 acctctgggg gcacagcagc cctgggctgc ctggtaagg actacttccc cgaaccggtg 540  
 acgggtgicgt ggaactcagg cgcctgacc agcggcgtgc acaccttccc ggctgiccta 600  
 cagtccctcag gactctactc cctcagcagc gtggtgaccg tgccctccag cagcttgggc 660  
 acccagacct acatctgcaa cgtgaatcac aagcccagca acaccaaggt ggacaagaaa 720  
 gttgagccca aatcttgta caaaactcac acatgcccac cgtgcccagc acctgaactc 780  
 ctggggggac cgtcagtctt cctcttcccc ccaaaacca aggacacct catgatctcc 840  
 cggacccttg aggtcacatg cgtgggtgtg gacgtgagcc acgaagacce tgaggtcaag 900  
 ttcaactggt acgtggacgg cgtggaggtg cataatgcca agacaaagcc gcgggaggag 960  
 cagtacaaca gcacgtaccg tgtggtcagc gtcctcaccg tctgcacca ggactggctg 1020  
 aatggcaagg agtacaagtg caaggtctcc aacaaagccc tcccagcccc catcgagaaa 1080  
 accatctcca aagccaaagg gcagccccga gaaccacagg tgtacacct gccccatcc 1140  
 cgggatgagc tgaccaagaa ccaggtcagc ctgacctgcc tgggtcaaagg ctctatccc 1200  
 agcgacatcg ccgtggagtg ggagagcaat gggcagccgg agaacaacta caagaccacg 1260  
 cctcccgtgc tggactccga cggctccttc ttcctctaca gcaagctcac cgtggacaag 1320  
 agcaggtagc agcaggggaa cgtcttctca tgctccgtga tgcattgagc tctgcacaac 1380  
 cactacacac agaagagcct ctcctgtct cgggtaaat ag 1422

[0036]

<210> 32  
 <211> 473  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 曲妥单抗重链的氨基酸序列

<400> 32

Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15

Ile Ala Thr Asn Ala Tyr Ala Glu Val Gln Leu Val Glu Ser Gly Gly  
 20 25 30

Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser  
 35 40 45

Gly Phe Asn Ile Lys Asp Thr Tyr Ile His Trp Val Arg Gln Ala Pro  
 50 55 60

Gly Lys Gly Leu Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr  
 65 70 75 80

Thr Arg Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Ala Asp  
 85 90 95

Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu  
 100 105 110

Asp Thr Ala Val Tyr Tyr Cys Ser Arg Trp Gly Gly Asp Gly Phe Tyr

115	120	125
Ala Met Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala 130	135	140
Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser 145	150	155
Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe 165	170	175
Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly 180	185	190
Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu 195	200	205
Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr 210	215	220
Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys 225	230	235
Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro 245	250	255
[0037] Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys 260	265	270
Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val 275	280	285
Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr 290	295	300
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu 305	310	315
Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His 325	330	335
Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys 340	345	350
Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln 355	360	365
Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu 370	375	380
Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro 385	390	395
Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn 405	410	415

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu  
 420 425 430

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val  
 435 440 445

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln  
 450 455 460

Lys Ser Leu Ser Leu Ser Pro Gly Lys  
 465 470

<210> 33  
 <211> 1491  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和附加的6x-His和strep-II亲和纯化标记的曲妥单抗重链的核苷酸序列

[0038]

<400> 33  
 atgaaaaaaaa acattgcctt tctgctggcc tccatgttcg tgttctctat cgccacaaac 60  
 gcttacgetg aagtccagct ggtcgaatct ggtggctggcc tgggtccagcc tgggtggatca 120  
 ctgagactgt cctgtgctgc ttctggtttc aacatcaagg acacctacat ccattgggctc 180  
 agacaggcac ctggcaaggg actggaatgg gtcgcccga tctaccctac aaacggctac 240  
 actcgtctag ccgactccgt caagggacgc tttaccatct ccgccgacac ctctaaaaac 300  
 accgcctacc tgcagatgaa tagtctgagg gccgaggata ctgctgtgta ctactgctca 360  
 cgatggggag gcgacggctt ttacgctatg gattactggg gacagggaa tctgtgtcact 420  
 gtgtctagcg ctagcaccaa gggcccatcg gtcttcccc tggcacccctc ctccaagagc 480  
 acctctgggg gcacagcagc cctgggctgc ctgggtcaagg actacttccc cgaaccgggtg 540  
 acggtgtcgt ggaactcagg cgcctgacc agcggcgtgc acaccttccc ggctgtccta 600  
 cagtctctag gactctactc cctcagcagc gtgggtgacc tgccctccag cagcttgggc 660  
 acccagacct acatctgcaa cgtgaatcac aagcccagca acaccaaggt ggacaagaaa 720  
 gttgagccca aatcttgtga caaaactcac acatgcccac cgtgcccagc acctgaactc 780  
 ctggggggac cgtcagtctt cctcttcccc ccaaaacca aggacacct catgatctcc 840  
 cggacccttg aggtcacatg cgtgggtggtg gacgtgagcc acgaagacc tgaggtcaag 900  
 ttcaactggt acgtggacgg cgtggagggtg cataatgcca agacaaagcc gcgggaggag 960  
 cagtacaaca gcacgtaccg tgtggtcagc gtcctcaccg tctgcacca ggactggctg 1020  
 aatggcaagg agtacaagtg caaggtctcc aacaaagccc tcccagcccc catcgagaaa 1080  
 accatctcca aagccaaagg gcagccccga gaaccacagg tgtacacct gccccatcc 1140  
 cgggatgagc tgaccaagaa ccaggtcagc ctgacctgcc tgggtcaaagg cttctatccc 1200  
 agcgacatcg ccgtggagtg ggagagcaat gggcagccgg agaacaacta caagaccagc 1260  
 cctcccgtgc tggactccga cgctccttc ttctctaca gcaagctcac cgtggacaag 1320  
 agcagggtggc agcaggggaa cgtcttctca tgcctcgtga tgcattgagc tctgcacaac 1380  
 cactacacac agaagagcct ctcctgtct cgggtaaag gatccctgcc cgagaccggc 1440  
 ggacaccatc accatcaca tggctggagc caccccagc tcgagaagta g 1491

<210> 34  
 <211> 496  
 <212> PRT  
 <213> 人工序列

<220>

<223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和附加的6x-His和strep-II亲和纯化标记的曲妥单抗重链的氨基酸序列

<400> 34

Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15

Ile Ala Thr Asn Ala Tyr Ala Glu Val Gln Leu Val Glu Ser Gly Gly  
 20 25 30

Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser  
 35 40 45

Gly Phe Asn Ile Lys Asp Thr Tyr Ile His Trp Val Arg Gln Ala Pro  
 50 55 60

Gly Lys Gly Leu Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr  
 65 70 75 80

Thr Arg Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Ala Asp  
 85 90 95

[0039]

Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu  
 100 105 110

Asp Thr Ala Val Tyr Tyr Cys Ser Arg Trp Gly Gly Asp Gly Phe Tyr  
 115 120 125

Ala Met Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala  
 130 135 140

Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser  
 145 150 155 160

Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe  
 165 170 175

Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly  
 180 185 190

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu  
 195 200 205

Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr  
 210 215 220

Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys  
 225 230 235 240

Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro  
 245 250 255

	Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys 260 265 270	
	Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val 275 280 285	
	Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr 290 295 300	
	Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu 305 310 315 320	
	Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His 325 330 335	
	Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys 340 345 350	
	Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln 355 360 365	
	Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu 370 375 380	
	Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro 385 390 395 400	
[0040]	Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn 405 410 415	
	Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu 420 425 430	
	Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val 435 440 445	
	Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln 450 455 460	
	Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Ser Leu Pro Glu Thr Gly 465 470 475 480	
	Gly His His His His His His Gly Trp Ser His Pro Gln Phe Glu Lys 485 490 495	
	<210> 35	
	<211> 714	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 曲妥单抗轻链的核苷酸序列	
	<400> 35	
	atgaaaaaaa acattgcctt tctgctggcc tccatgttcg tgttctctat cgccacaaat	60
	gcttatgccg acatccagat gacacagtct ccctcttccc tgtecgcttc tgtggcgat	120
	cgagtgacaa tcacctgtag ggctagtcag gatgtgaata ctgctgttgc ttgtaccag	180



cagaaaccag gaaaagcccc taaactgctg atctactctg cctcattect gtactctggg 240  
 gtgccttctc gattcagtggt ttctagatct ggcaccgatt tcacactgac cattttctca 300  
 ctgcaacctg aggattttgc cacctactac tgtcagcagc actacacaac acctcccaca 360  
 tttggccagg gcacaaaagt ggagatcaaa cgtacgggtg ctgcaccatc tgtcttcatc 420  
 ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat 480  
 aactttatc ccagagaggc caaagtacag tggaaggtgg ataacgccct ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 accctgacgc tgagcaaagc agactacgag aaacacaaaag tctacgctg cgaagtcacc 660  
 catcagggcc tgagctcgcc cgtcacaaaag agcttcaaca ggggagagtg ttag 714

<210> 36  
 <211> 237  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 曲妥单抗轻链的氨基酸序列

<400> 36

Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15

Ile Ala Thr Asn Ala Tyr Ala Asp Ile Gln Met Thr Gln Ser Pro Ser  
 20 25 30

[0041]

Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala  
 35 40 45

Ser Gln Asp Val Asn Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly  
 50 55 60

Lys Ala Pro Lys Leu Leu Ile Tyr Ser Ala Ser Phe Leu Tyr Ser Gly  
 65 70 75 80

Val Pro Ser Arg Phe Ser Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu  
 85 90 95

Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln  
 100 105 110

Gln His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu  
 115 120 125

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
 130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
 165 170 175

Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys

180 185 190  
 Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
 195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
 210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys  
 225 230 235

<210> 37  
 <211> 783  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和附加的6x-His和strep-II亲和纯化标记的曲妥单抗轻链的核苷酸序列

[0042]

<400> 37  
 atgaaaaaa acattgcctt tctgctggcc tccatgttcg tgttctctat cgccacaaat 60  
 gcttatgccg acatccagat gacacagtct ccctcttccc tgtccgcttc tgtggcgcat 120  
 cgagtgacaa tcacctgtag ggctagtcag gatgtgaata ctgctgttgc ttggtaccag 180  
 cagaaaccag gaaaagcccc taaactgctg atctactctg cctcattcet gtactctggg 240  
 gtgccttctc gattcagtgg ttctagatct ggcaccgatt tcacactgac catttcttca 300  
 ctgcaacctg aggattttgc cacctactac tgtcagcagc actacacaac acctcccaca 360  
 ttggccagg gcacaaaagt ggagatcaaa cgtacgggtg ctgcaccate tgtcttcatc 420  
 ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat 480  
 aacttctatc ccagagaggc caaagtacag tggaagggtg ataagccct ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 accctgacgc tgagcaaacg agactacgag aaacacaaag tctacgcctg cgaagtcacc 660  
 catcagggcc tgagctcgcc cgtcacaaaag agettcaaca ggggagagtg tggatccctg 720  
 cccgagaccg gcggacacca tcacatcac catggctgga gccaccccca gttcgagaag 780  
 tag 783

<210> 38  
 <211> 260  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和附加的6x-His和strep-II亲和纯化标记的曲妥单抗轻链的氨基酸序列

<400> 38  
 Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15  
 Ile Ala Thr Asn Ala Tyr Ala Asp Ile Gln Met Thr Gln Ser Pro Ser  
 20 25 30  
 Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala  
 35 40 45

Ser Gln Asp Val Asn Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly  
 50 55 60

Lys Ala Pro Lys Leu Leu Ile Tyr Ser Ala Ser Phe Leu Tyr Ser Gly  
 65 70 75 80

Val Pro Ser Arg Phe Ser Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu  
 85 90 95

Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln  
 100 105 110

Gln His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu  
 115 120 125

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
 130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
 165 170 175

Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
 180 185 190

Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
 195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
 210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Gly Ser Leu  
 225 230 235 240

Pro Glu Thr Gly Gly His His His His His His Gly Trp Ser His Pro  
 245 250 255

Gln Phe Glu Lys  
 260

[0043]

<210> 39  
 <211> 1470  
 <212> DNA  
 <213> 人工序列

<220>

<223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和附加的6x-His和strep-II亲和纯化标记的mAb Ac10 重链的核苷酸序列

<400> 39  
 atgaattttg gactgaggct gattttcctg gtgctgacct tgaaggcgt ccagtgtcag 60  
 atccagctgc agcagctctgg ccccagggtc gtgaaacctg gcgctccgt gaagatctcc 120  
 tgcaaggcct cggctacac ctaccacgac tactacatca cctgggtca gcagaagccc 180  
 ggccagggcc tggaaatgat cggctggatc tatcccggt cggcaaac caagtacaac 240  
 gagaagtta agggcaaggc cacctgacc gtggacacct cctcttccac cgccttcag 300

cagctgtcct cctgacctc cgaggatacc gccgtgtact tctgcgcaa ctacggcaac	360
tattggttcg cctactgggg ccagggcaca caagtgaccg tgtctgtctg tagcaccaag	420
ggcccatcgg tcttccccct ggaccctcc tccaagagca cctctggggg cacagcagcc	480
ctgggctgcc tggteaagga ctacttcccc gaaccgggta cgggtgtctg gaactcagc	540
gccttgacca gggcgtgca caccttccc gctgtcctac agtcctcagg actctactcc	600
ctcagcagcg tggtagcctg gcctccagc agcttgggca ccagacctc catctgcaac	660
gtgaatcaca agcccagcaa caccaaggtg gacaagaaag ttgagcccaa atcttgtgac	720
aaaactcaca catgccacc gtgccagca cctgaactcc tggggggacc gtcagtcttc	780
cttttcccc caaaacccaa ggacaccctc atgatctccc ggaccctga ggtcacatgc	840
gtggtggtgg acgtgagcca cgaagaccct gaggtcaagt tcaactggta cgtggacggc	900
gtggaggtgc ataatccaa gacaaagccg cgggaggagc agtacaacag cacgtaccgt	960
gtggtcagcg tectaccct cctgcaccag gactggctga atggcaagga gtacaagtgc	1020
aaggctctca acaaaacct cccagcccc atcgagaaaa ccatctccaa agccaaaggg	1080
cagccccgag aaccacaggt gtaccacctg cccccatccc gggatgaget gaccaagaac	1140
caggctagcc tgacctgect ggtcaaagc ttctatccca gcgacatcgc cgtggagtgg	1200
gagagcaatg ggcagccgga gaacaactac aagaccacgc ctcccgtgct ggactccgac	1260
ggctctttct tctctacag caagctcacc gtggacaaga gcaggtggca gcagggaac	1320
gtcttctcat gctccgtgat gcatgagct ctgcacaacc actacacaca gaagaccctc	1380
tcctgtctc cgggtaaagg atccctgccc gagaccggcg gacaccatca ccatcaccat	1440
ggctggagcc acccccagtt cgagaagtag	1470

[0044]

- <210> 40
  - <211> 489
  - <212> PRT
  - <213> 人工序列
  
  - <220>
  - <223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和附加的 6x-His和strep II亲和纯化标记的mAb Ac10 重链的氨基酸序列
  
  - <400> 40
- ```

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly
1           5           10          15

Val Gln Cys Gln Ile Gln Leu Gln Gln Ser Gly Pro Glu Val Val Lys
                20           25           30

Pro Gly Ala Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe
          35           40           45

Thr Asp Tyr Tyr Ile Thr Trp Val Lys Gln Lys Pro Gly Gln Gly Leu
    50           55           60

Glu Trp Ile Gly Trp Ile Tyr Pro Gly Ser Gly Asn Thr Lys Tyr Asn
65           70           75           80

Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Thr Ser Ser Ser
          85           90           95
    
```

Thr Ala Phe Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Thr Ala Val  
 100 105 110

Tyr Phe Cys Ala Asn Tyr Gly Asn Tyr Trp Phe Ala Tyr Trp Gly Gln  
 115 120 125

Gly Thr Gln Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val  
 130 135 140

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala  
 145 150 155 160

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser  
 165 170 175

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val  
 180 185 190

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro  
 195 200 205

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys  
 210 215 220

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp  
 225 230 235 240

[0045] Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly  
 245 250 255

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile  
 260 265 270

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu  
 275 280 285

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His  
 290 295 300

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg  
 305 310 315 320

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys  
 325 330 335

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu  
 340 345 350

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr  
 355 360 365

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu  
 370 375 380

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp



Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly  
1 5 10 15

Val Gln Cys Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val  
20 25 30

Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val  
35 40 45

Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly  
50 55 60

Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly  
65 70 75 80

Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu  
85 90 95

Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln  
100 105 110

Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu  
115 120 125

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
130 135 140

[0047] Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
165 170 175

Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
180 185 190

Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Gly Ser Leu  
225 230 235 240

Pro Glu Thr Gly Gly His His His His His His Gly Trp Ser His Pro  
245 250 255

Gln Phe Glu Lys  
260

<210> 43  
<211> 1443  
<212> DNA  
<213> 人工序列

<220>  
<223> 具有C端LPETG分选酶A识别基序和strep II亲和纯化标记的mAb Ac10 重链核苷酸序列

|                                                                        |      |
|------------------------------------------------------------------------|------|
| <400> 43                                                               |      |
| atgaattttg gactgagget gattttctcg gtgetgaccc tgaaggcgt ccagtgtcag       | 60   |
| atccagctgc agcagctctgg ccccgaggtc gtgaaacctg gcgcctccgt gaagatctcc     | 120  |
| tgcaaggcct ccggetacac cttaccgac tactacatca cctgggtcaa gcagaagccc       | 180  |
| ggccagggcc tggaatggat eggctggatc tatcccgct ccggcaacac caagtacaac       | 240  |
| gagaagtcca agggcaagge caccctgacc gtggacacct cctcttccac cgccttcag       | 300  |
| cagctgtcct ccctgacctc egaggatacc gccgtgtact tctgcgcaa ctacggcaac       | 360  |
| tattggttcg cctaetgggg ccagggcaca caagtaccg tgtctgtgc tagcaccaag        | 420  |
| ggcccatcgg tcttccccct ggcacctcc tccaagagca cctctggggg cacagcagcc       | 480  |
| ctgggctgcc tggtaagga ctacttcccc gaaccgggta cgggtgtctg gaactcagge       | 540  |
| gccctgacca gcggcgtgca caccttcccg gctgtcttac agtcctcagg actctactcc      | 600  |
| ctcagcagcg tggtgaccgt gccctccagc agcttgggca cccagaccta catctgcaac      | 660  |
| gtgaatcaca agcccagcaa caccaagggt gacaagaaag ttgagcccaa atcttgtgac      | 720  |
| aaaactcaca catgccacc gtgccagca cctgaactcc tggggggacc gtcagtcttc        | 780  |
| ctcttcccc caaaacccaa ggacacctc atgatctccc ggaccttga ggtcacatgc         | 840  |
| gtggtggtgg acgtgagcca cgaagacct gaggtcaagt tcaactggta cgtggacgge       | 900  |
| gtggaggtgc ataatgcaa gacaaagccg cgggaggagc agtacaacag cacgtaccgt       | 960  |
| gtggtcagcg tctcaccgt cctgcaccag gactggctga atggcaagga gtacaagtgc       | 1020 |
| [0048] aaggtctcca acaaagccct cccagccccc atcgagaaaa ccatctcaa agccaaagg | 1080 |
| cagccccgag aaccacaggt gtacacctg ccccatccc gggatgagct gaccaagaac        | 1140 |
| caggtcagcc tgacctgct ggtcaaagge ttctatccca gcgacatcgc cgtggagtgg       | 1200 |
| gagagcaatg ggcagccgga gaacaactac aagaccacgc ctcccgtgct ggactccgac      | 1260 |
| ggctccttct tctctacag caagctcacc gtggacaaga gcaggtggca gcaggggaac       | 1320 |
| gtcttctcat gctccgtgat gcatgagget ctgcacaacc actacacaca gaagagctc       | 1380 |
| tcctgtctc cgggtaaact gcccagacc ggcggatgga gccaccccca gttcgagaag        | 1440 |
| tag                                                                    | 1443 |
| <210> 44                                                               |      |
| <211> 480                                                              |      |
| <212> PRT                                                              |      |
| <213> 人工序列                                                             |      |
| <220>                                                                  |      |
| <223> 具有C端LPETG分选酶A识别基序和strep-II亲和纯化标记的mAb Ac10 重链的氨基酸序列               |      |
| <400> 44                                                               |      |
| Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly        |      |
| 1                    5                    10                    15     |      |
| Val Gln Cys Gln Ile Gln Leu Gln Gln Ser Gly Pro Glu Val Val Lys        |      |
| 20                    25                    30                         |      |
| Pro Gly Ala Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe        |      |
| 35                    40                    45                         |      |



Thr Asp Tyr Tyr Ile Thr Trp Val Lys Gln Lys Pro Gly Gln Gly Leu  
 50 55 60  
 Glu Trp Ile Gly Trp Ile Tyr Pro Gly Ser Gly Asn Thr Lys Tyr Asn  
 65 70 75 80  
 Glu Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Thr Ser Ser Ser  
 85 90 95  
 Thr Ala Phe Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Thr Ala Val  
 100 105 110  
 Tyr Phe Cys Ala Asn Tyr Gly Asn Tyr Trp Phe Ala Tyr Trp Gly Gln  
 115 120 125  
 Gly Thr Gln Val Thr Val Ser Ala Ala Ser Thr Lys Gly Pro Ser Val  
 130 135 140  
 Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala  
 145 150 155 160  
 Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser  
 165 170 175  
 Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val  
 180 185 190  
 [0049] Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro  
 195 200 205  
 Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys  
 210 215 220  
 Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp  
 225 230 235 240  
 Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly  
 245 250 255  
 Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile  
 260 265 270  
 Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu  
 275 280 285  
 Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His  
 290 295 300  
 Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg  
 305 310 315 320  
 Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys  
 325 330 335  
 Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu  
 340 345 350

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr  
 355 360 365

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu  
 370 375 380

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp  
 385 390 395 400

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val  
 405 410 415

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp  
 420 425 430

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His  
 435 440 445

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro  
 450 455 460

Gly Lys Leu Pro Glu Thr Gly Gly Trp Ser His Pro Gln Phe Glu Lys  
 465 470 475 480

<210> 45  
 <211> 756  
 <212> DNA  
 <213> 人工序列

[0050]

<220>  
 <223> 具有C端LPETG分选酶A识别基序和strep-II亲和纯化标记的mAb Ac10轻链的核苷酸序列

<400> 45  
 atgaatttg gactgaggct gattttcctg gtgctgacc tgaaggcgt ccagtgtgac 60  
 attgtgctga cccaatctcc agcttctttg gctgtgtctc tagggcagag ggccaccatc 120  
 tcctgcaagg ccagccaaag tgttgatttt gatggtgata gttatatgaa ctggtaccaa 180  
 cagaaccagg gacagccacc caaagtctc atctatgctg catccaatct agaatctggg 240  
 atcccagcca ggtttagtgg cagtgggtct gggacagact tcaccctcaa catccatcct 300  
 gtggaggagg aggatgctgc aacctattac tgtcagcaaa gtaatgagga tccgtggacg 360  
 ttcggtagg gacccaagct ggaaatcaaa cgtacggtgg ctgcaccatc tgccttcac 420  
 ttcccgcat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat 480  
 aacttctatc ccagagaggc caaagtacag tggaaggtgg ataacgccct ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 acctgacgc tgagcaaaag agactacgag aaacacaaag tctacgctg cgaagtcacc 660  
 catcagggcc tgagctcgcc cgtcaciaag agcttcaaca ggggagagtg tetgcccag 720  
 accggcggat ggagccaccc ccagttcgag aagtag 756

<210> 46  
 <211> 251  
 <212> PRT  
 <213> 人工序列

<220>

<223> 具有C端LPETG分选酶A识别基序和strep-II亲和纯化标记的mAb Ac10 轻链的氨基酸序列

<400> 46

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly  
1 5 10 15

Val Gln Cys Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val  
20 25 30

Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val  
35 40 45

Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly  
50 55 60

Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly  
65 70 75 80

Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu  
85 90 95

Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln  
100 105 110

Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu  
115 120 125

[0051]

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
165 170 175

Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
180 185 190

Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Leu Pro Glu  
225 230 235 240

Thr Gly Gly Trp Ser His Pro Gln Phe Glu Lys  
245 250

<210> 47

<211> 762

<212> DNA

<213> 人工序列

<220>

<223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和strep II亲和纯化标记的mAb Ac10轻链的核苷酸序列

<400> 47  
 atgaattttg gactgagget gattttcctg gtgctgacce tgaaggcgt ccagtgtgac 60  
 attgtgctga cccaatctcc agcttctttg gctgtgtctc tagggcagag ggccaccatc 120  
 tcctgcaagg ccagccaaag tgttgatttt gatggtgata gttatatgaa ctggtaccaa 180  
 cagaaaccag gacagccacc caaagtctc atctatgctg catccaatct agaatctggg 240  
 atcccagcca ggtttagtgg cagtgggtct gggacagact tcaccctcaa catccatcct 300  
 gtggaggagg aggatgctgc aacctattac tgtcagcaaa gtaatgagga tccgtggacg 360  
 ttcggtggag gcaccaagct ggaaatcaaa cgtacgggtg ctgcaccatc tgtcttcac 420  
 ttcccgcat ctgatgagca gttgaaatct ggaactgect ctgttgtgtg cctgctgaat 480  
 aacttctatc ccagagaggc caaagtacag tggaaggtg ataacgcct ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 acctgacgc tgagcaaagc agactacgag aaacacaaag tctacgctg cgaagtcacc 660  
 catcagggcc tgagctgcc cgtcacaag agcttcaaca ggggagagtg tggatccctg 720  
 cccgagaccg gcggatggag ccacccccag ttcgagaagt ag 762

<210> 48  
 <211> 253  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 具有C端GS间隔肽, LPETG 分选酶A识别基序和strep II亲和纯化标记的mAb Ac10轻链的氨基酸序列

[0052]

<400> 48  
 Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly  
 1 5 10 15  
 Val Gln Cys Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val  
 20 25 30  
 Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val  
 35 40 45  
 Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly  
 50 55 60  
 Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly  
 65 70 75 80  
 Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu  
 85 90 95  
 Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln  
 100 105 110  
 Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu  
 115 120 125  
 Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
 130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
 165 170 175

Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
 180 185 190

Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
 195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
 210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Gly Ser Leu  
 225 230 235 240

Pro Glu Thr Gly Gly Trp Ser His Pro Gln Phe Glu Lys  
 245 250

<210> 49  
 <211> 765  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 具有C端GGG间隔肽, LPETG 分选酶A识别基序和strep II亲和纯化标记的mAb Ac10轻链的核苷酸序列

[0053]

<400> 49  
 atgaattttg gactgagget gattttcctg gtgctgacct tgaaggcgt ccagtgtgac 60  
 attgtgctga cccaatctcc agcttcittg gctgtgtctc tagggcagag gccaccatc 120  
 tctgcaagg ccagccaaag tgttgatttt gatggtgata gttatatgaa ctgtaccaa 180  
 cagaaaccag gacagccacc caaagtctc atctatgctg catccaatct agaatctggg 240  
 atccagcca ggttttagtg cagtgggtct gggacagact tcacctcaa catccatcct 300  
 gtggaggagg aggatgctgc aacctattac tgtcagcaaa gtaatgagga tccgtggacg 360  
 ttcggtggag gcaccaagct ggaatcaaa cgtacgggtg ctgcaccatc tgtcttcate 420  
 ttccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat 480  
 aactctatc ccagagagge caaagtacag tggaaggtgg ataacgcctt ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 accctgacgc tgagcaaagc agactacgag aaacacaaag tctacgcctg cgaagtcacc 660  
 catcagggcc tgagctgcc cgtcacaaag agcttcaaca ggggagagtg tgggtgatcc 720  
 ctgcccgaga ccggcggatg gagccacccc cagttcgaga agtag 765

<210> 50  
 <211> 254  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 具有C端GGG间隔肽, LPETG 分选酶A识别基序和strep II亲和纯化标记的mAb Ac10 轻链的氨基酸序列

<400> 50  
 Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly



tcctgcaagg ccagccaaag tgttgatttt gatggtgata gttatatgaa ctggtaccaa 180  
 cagaaaccag gacagccacc caaagtcctc atctatgctg catccaatct agaactctggg 240  
 atcccagcca ggitttagtg cagtgggtct gggacagact tcacctcaa catccatcct 300  
 gtggaggagg aggatgetgc aacctattac tgcagcaaa gtaatgagga tccgtggacg 360  
 ttcgtgtggg gcaccaagct ggaatcaaa cgtacggtg ctgcaccatc tgtcttcate 420  
 ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat 480  
 aacttctatc ccagagaggc caaagtacag tggaaggtgg ataacgcctt ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 acctgacgc tgagcaaagc agactacgag aaacacaaag tctacgcctg cgaagtcacc 660  
 catcagggcc tgagctgcc cgtcacaaag agcttcaaca ggggagagtg tgggggtgga 720  
 tccttgcctg agaccggcgg atggagccac ccccagtctg agaagtag 768

<210> 52

<211> 255

<212> PRT

<213> 人工序列

<220>

<223> 具有C端GGGS间隔肽, LPETG 分选酶A识别基序和strep II亲和纯化标记的mAb Ac10 轻链的氨基酸序列

<400> 52

Met Asn Phe Gly Leu Arg Leu Ile Phe Leu Val Leu Thr Leu Lys Gly  
 1 5 10 15

[0055]

Val Gln Cys Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val  
 20 25 30

Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val  
 35 40 45

Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly  
 50 55 60

Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly  
 65 70 75 80

Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu  
 85 90 95

Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln  
 100 105 110

Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu  
 115 120 125

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
 130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala





Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val  
 35 40 45

Asp Phe Asp Gly Asp Ser Tyr Met Asn Trp Tyr Gln Gln Lys Pro Gly  
 50 55 60

Gln Pro Pro Lys Val Leu Ile Tyr Ala Ala Ser Asn Leu Glu Ser Gly  
 65 70 75 80

Ile Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu  
 85 90 95

Asn Ile His Pro Val Glu Glu Glu Asp Ala Ala Thr Tyr Tyr Cys Gln  
 100 105 110

Gln Ser Asn Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu  
 115 120 125

Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser  
 130 135 140

Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn  
 145 150 155 160

Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala  
 165 170 175

[0057]

Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys  
 180 185 190

Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp  
 195 200 205

Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu  
 210 215 220

Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Gly Gly Gly  
 225 230 235 240

Gly Ser Leu Pro Glu Thr Gly Gly Trp Ser His Pro Gln Phe Glu Lys  
 245 250 255

- <210> 55
  - <211> 1464
  - <212> DNA
  - <213> 人工序列
  - <220>
  - <223> 具有C端LPETG分选酶标记和strep II亲和纯化标记的曲妥单抗重链的核苷酸序列
  - <400> 55
- |                                                                    |     |
|--------------------------------------------------------------------|-----|
| atgaaaaaaaa acattgcctt tctgctggcc tccatgttcg tgttctctat cgccacaaac | 60  |
| gcttacgctg aagtccagct ggtcgaatct ggtggtggcc tggcaccg cc tggatgatca | 120 |
| ctgagactgt cctgtgctgc tctgtgttcc aacatcaagg acacctacat ccattgggctc | 180 |
| agacaggcac ctggcaaggg actggaatgg gtcgcccga tctaccctac aaacggctac   | 240 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| actcgtctacg ccgactccgt caagggacgc tttaccatct ccgccgacac ctctaaaaac | 300  |
| accgcctacc tgcagatgaa tagtctgagg gccgaggata ctgctgtgta ctactgtctca | 360  |
| cgatggggag gcgacggctt ttacgctatg gattactggg gacagggaac tctggctact  | 420  |
| gtgtctagcg ctacaccaa gggcccatcg gtcttcccc tggcaccetc ctccaagagc    | 480  |
| acctctgggg gcacagcagc cctgggctgc ctgggtcaagg actacttccc cgaaccggtg | 540  |
| acggtgtcgt ggaactcagg cgccctgacc agcggcgtgc acaccttccc ggctgtccta  | 600  |
| cagtctcag gactctactc cctcagcagc gtggtgaccg tgcctccag cagcttgggc    | 660  |
| accagacct acatctgcaa cgtgaatcac aagcccagca acaccaaggt ggacaagaaa   | 720  |
| gttgagccca aatcttgtga caaaactcac acatgcccac cgtgcccagc acctgaactc  | 780  |
| ctggggggac cgtcagctct cctcttcccc ccaaaaccca aggacacct catgatctcc   | 840  |
| cggaccctg aggtcacatg cgtgggtggtg gacgtgagcc acgaagacc tgaggtcaag   | 900  |
| ttcaactggt acgtggacgg cgtggagggtg cataatgcca agacaaagcc gcgggaggag | 960  |
| cagtacaaca gcacgtaccg tgtggtcagc gtctctaccg tctgcacca ggactggetg   | 1020 |
| aatggcaagg agtacaagtg caaggtctcc aacaaagccc tcccagcccc catcgagaaa  | 1080 |
| accatctcca aagccaaagg gcagccccga gaaccacagg tgtacacct gccccatcc    | 1140 |
| cgggatgagc tgaccaagaa ccaggtcagc ctgacctgcc tggtaaagg cttctatccc   | 1200 |
| agcgacatcg ccgtggagtg ggagagcaat gggcagccgg agaacaacta caagaccacg  | 1260 |
| cctccccgtc tggactccga cggctccttc ttccttaca gcaagctcac cgtggacaag   | 1320 |
| agcaggtggc agcaggggaa cgtcttctca tgcctcgtga tgcattaggc tctgcacaac  | 1380 |
| cactacacac agaagagcct ctcctgtct cgggtaaac tgcccagagc cggcggatgg    | 1440 |
| agccaccccc agttcgagaa gtag                                         | 1464 |

[0058]

- <210> 56
- <211> 487
- <212> PRT
- <213> 氨基酸序列 mAb Ac10 重链 具有C端LPETG分选酶标记和strep II亲和纯化标记
- <400> 56

Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15

Ile Ala Thr Asn Ala Tyr Ala Glu Val Gln Leu Val Glu Ser Gly Gly  
 20 25 30

Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser  
 35 40 45

Gly Phe Asn Ile Lys Asp Thr Tyr Ile His Trp Val Arg Gln Ala Pro  
 50 55 60

Gly Lys Gly Leu Glu Trp Val Ala Arg Ile Tyr Pro Thr Asn Gly Tyr  
 65 70 75 80

Thr Arg Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Ala Asp  
 85 90 95

Thr Ser Lys Asn Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu

|        | 100                                                                    | 105 | 110 |
|--------|------------------------------------------------------------------------|-----|-----|
|        | Asp Thr Ala Val Tyr Tyr Cys Ser Arg Trp Gly Gly Asp Gly Phe Tyr<br>115 | 120 | 125 |
|        | Ala Met Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala<br>130 | 135 | 140 |
|        | Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser<br>145 | 150 | 155 |
|        | Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe<br>165 | 170 | 175 |
|        | Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly<br>180 | 185 | 190 |
|        | Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu<br>195 | 200 | 205 |
|        | Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr<br>210 | 215 | 220 |
|        | Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys<br>225 | 230 | 235 |
| [0059] | Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro<br>245 | 250 | 255 |
|        | Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys<br>260 | 265 | 270 |
|        | Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val<br>275 | 280 | 285 |
|        | Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr<br>290 | 295 | 300 |
|        | Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu<br>305 | 310 | 315 |
|        | Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His<br>325 | 330 | 335 |
|        | Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys<br>340 | 345 | 350 |
|        | Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln<br>355 | 360 | 365 |
|        | Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu<br>370 | 375 | 380 |
|        | Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro<br>385 | 390 | 395 |
|        |                                                                        |     | 400 |

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn  
 405 410 415

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu  
 420 425 430

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val  
 435 440 445

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln  
 450 455 460

Lys Ser Leu Ser Leu Ser Pro Gly Lys Leu Pro Glu Thr Gly Gly Trp  
 465 470 475 480

Ser His Pro Gln Phe Glu Lys  
 485

<210> 57  
 <211> 771  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 具有C端GGGGS连接器, LPETG分选酶标记和strep II亲和纯化标记的曲妥单抗轻链的核苷酸序列

[0060]

<400> 57  
 atgaaaaaaa acattgcctt tctgctggcc tccatgttcg tgttctctat cgcacaaaat 60  
 gcttatgccg acatccagat gacacagtct cctcttcccg tgtccgcttc tgtggcgat 120  
 cgagtgacaa tcacctgtag ggctagtcag gatgtgaata ctgctgttgc ttggtaccag 180  
 cagaaaccag gaaaagcccc taaactgctg atctactctg cctcattcct gtactctggg 240  
 gtgccttctc gattcagtgg ttctagatct ggcaccgatt tcacactgac catttctca 300  
 ctgcaacctg aggattttgc cacctactac tgtcagcagc actacacaac acctcccaca 360  
 tttggccagg gcacaaaagt ggagatcaaa cgtacgggtg ctgcaccatc tgtcttcac 420  
 ttcccgccat ctgatgagca gttgaaatct ggaactgcct ctgttgtgtg cctgctgaat 480  
 aacttctatc ccagagaggc caaagtacag tggaaggtgg ataacgcctt ccaatcgggt 540  
 aactcccagg agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc 600  
 acctgacgc tgagcaaagc agactacgag aaacacaaag tctacgcctg cgaagtcacc 660  
 catcagggcc tgagctcgcg cgtcacaagg agcttcaaca ggggagagtg tggagggggg 720  
 ggatccctgc ccgagaccgg cggatggagc caccgccagt tcgagaagta g 771

<210> 58  
 <211> 256  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 具有C端GGGGS连接器, LPETG分选酶标记和strep II亲和纯化标记的mAb Ac10轻链的氨基酸序列

<400> 58  
 Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15

Ile Ala Thr Asn Ala Tyr Ala Asp Ile Gln Met Thr Gln Ser Pro Ser

|        | 20                                                                                 | 25 | 30 |
|--------|------------------------------------------------------------------------------------|----|----|
|        | Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala<br>35 40 45        |    |    |
|        | Ser Gln Asp Val Asn Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly<br>50 55 60        |    |    |
|        | Lys Ala Pro Lys Leu Leu Ile Tyr Ser Ala Ser Phe Leu Tyr Ser Gly<br>65 70 75 80     |    |    |
|        | Val Pro Ser Arg Phe Ser Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu<br>85 90 95        |    |    |
|        | Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln<br>100 105 110     |    |    |
|        | Gln His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu<br>115 120 125     |    |    |
| [0061] | Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser<br>130 135 140     |    |    |
|        | Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn<br>145 150 155 160 |    |    |
|        | Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala<br>165 170 175     |    |    |
|        | Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys<br>180 185 190     |    |    |
|        | Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp<br>195 200 205     |    |    |
|        | Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu<br>210 215 220     |    |    |
|        | Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys Gly Gly Gly<br>225 230 235 240 |    |    |
|        | Gly Ser Leu Pro Glu Thr Gly Gly Trp Ser His Pro Gln Phe Glu Lys<br>245 250 255     |    |    |

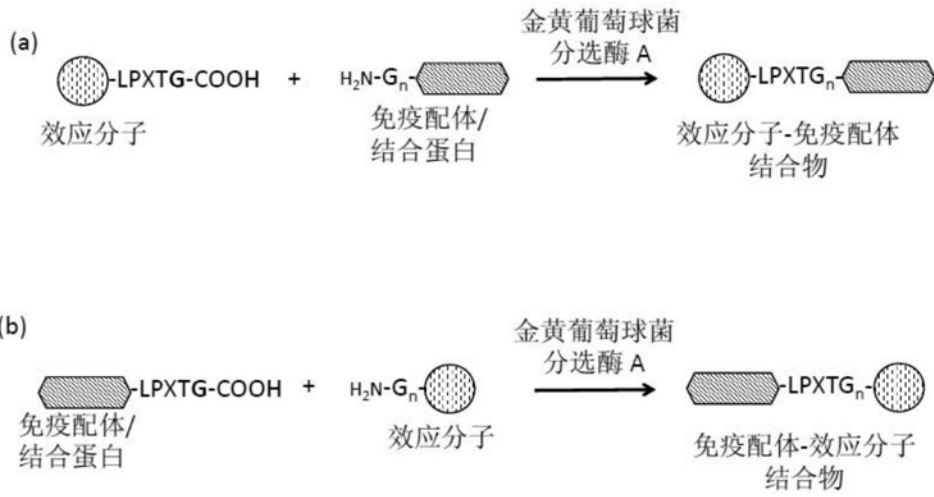


图1

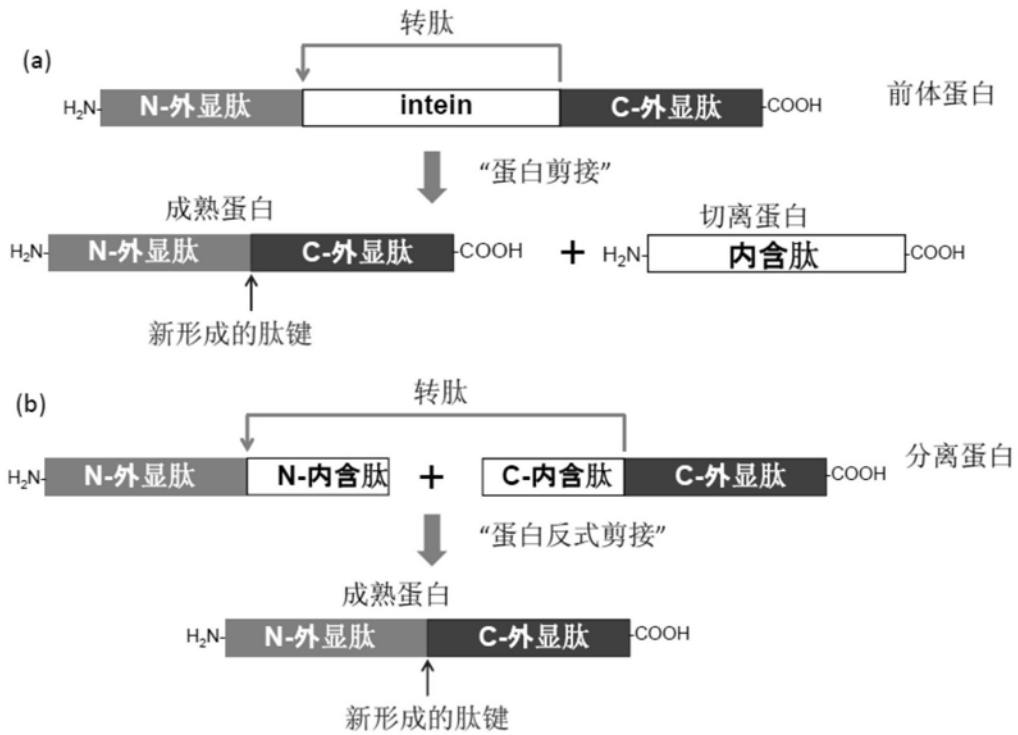


图2

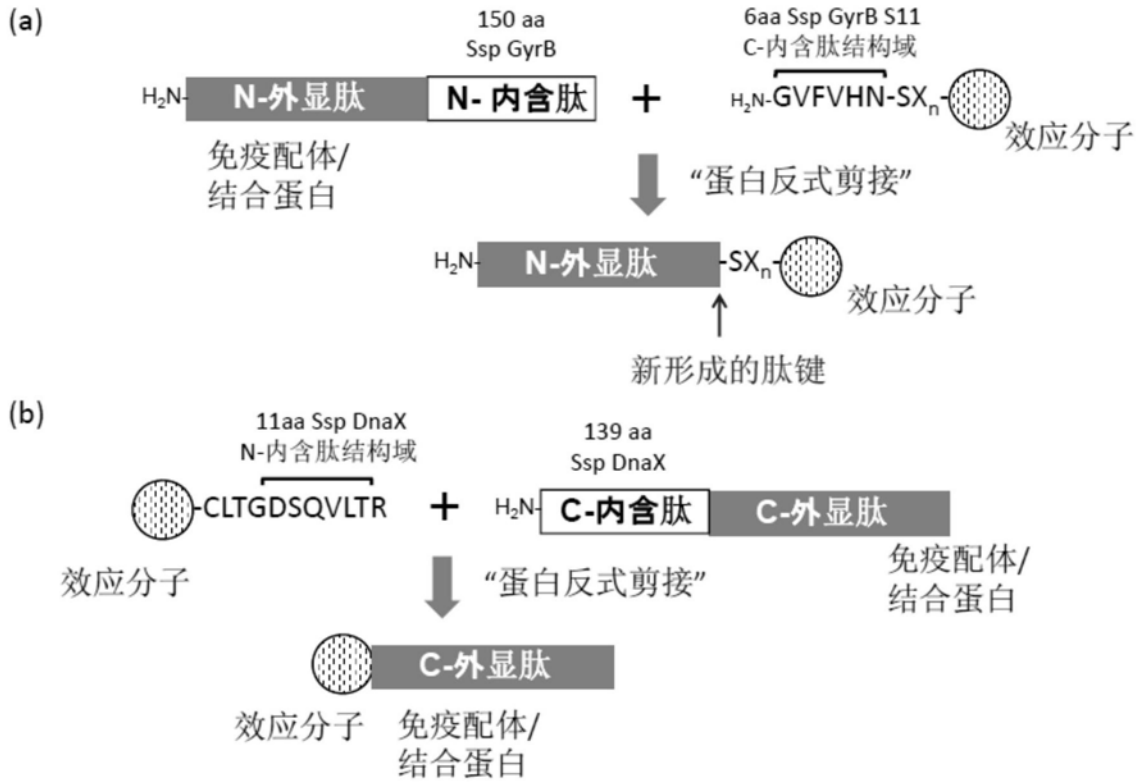


图3

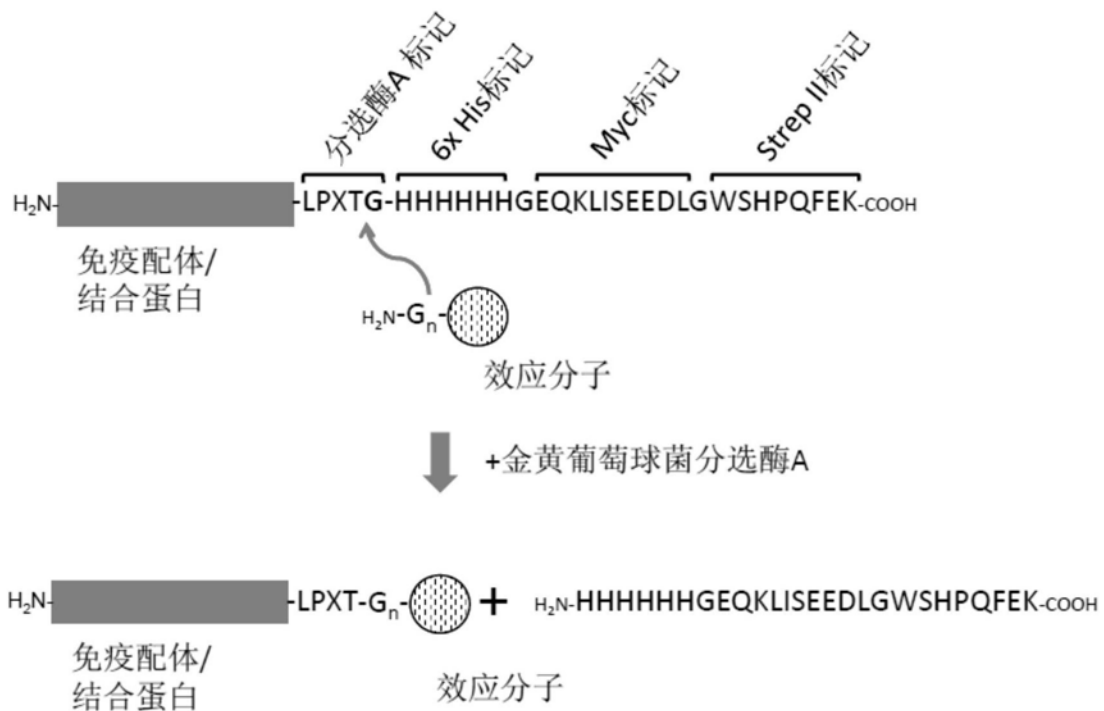


图4a

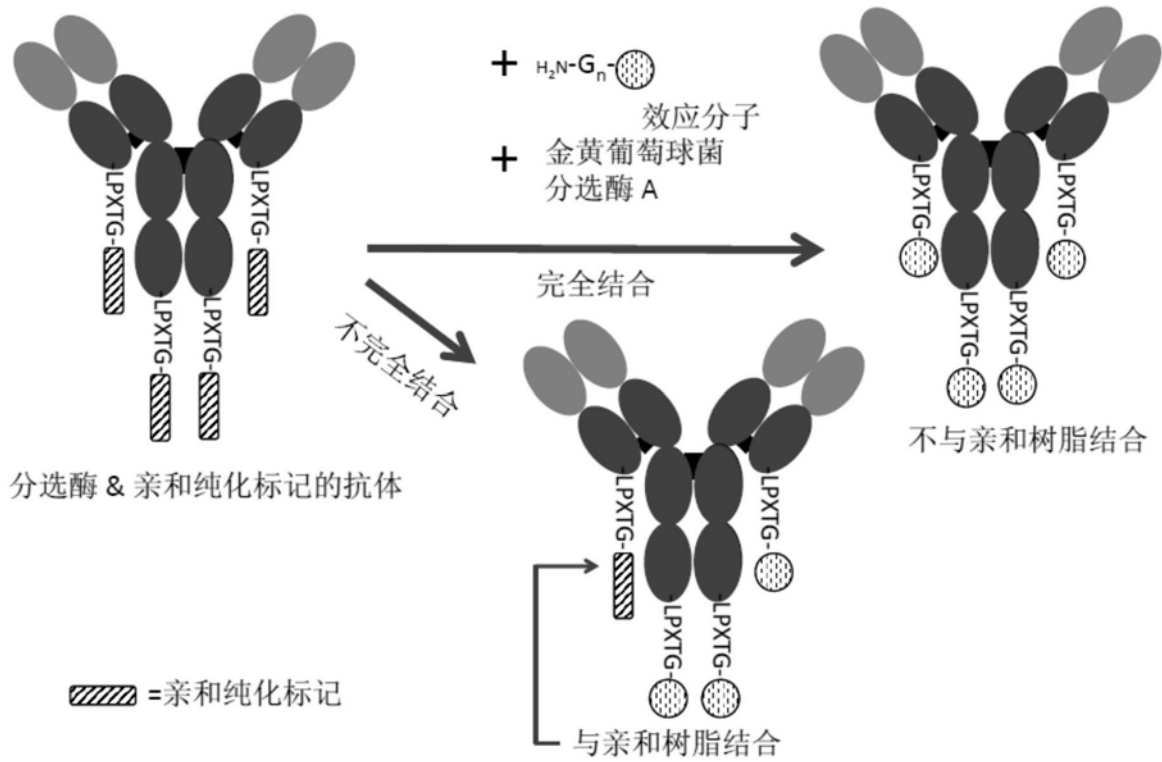


图4b

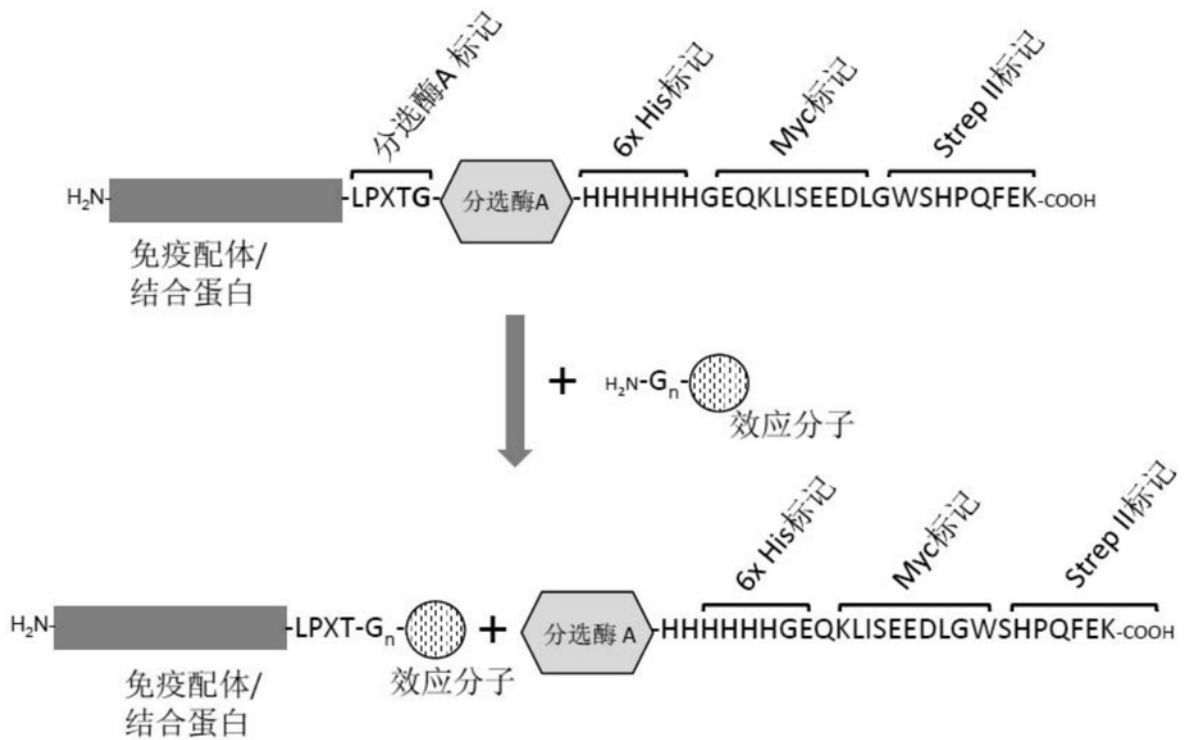
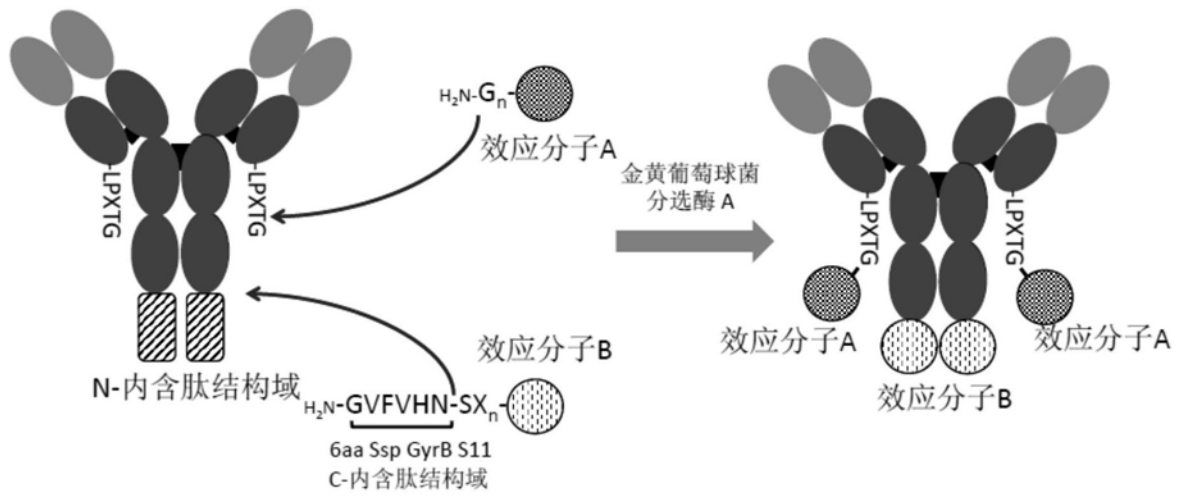


图5



(a)



(b)

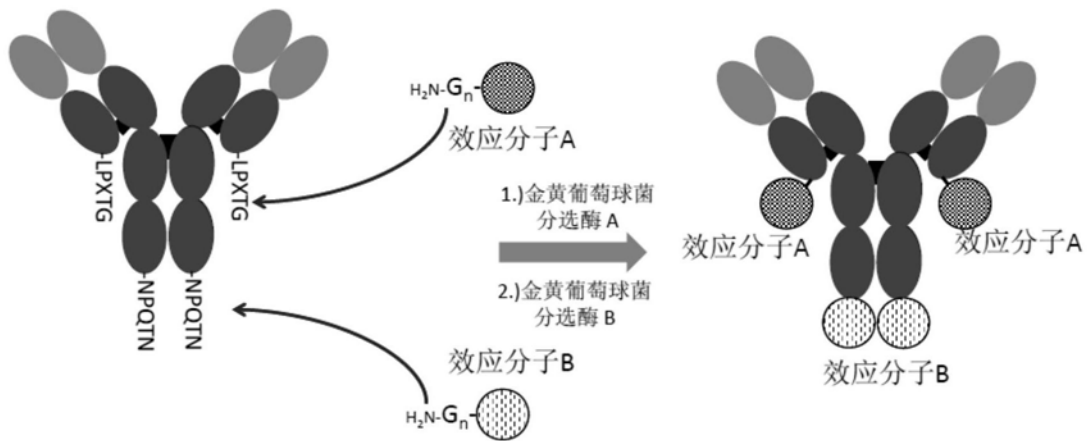


图6

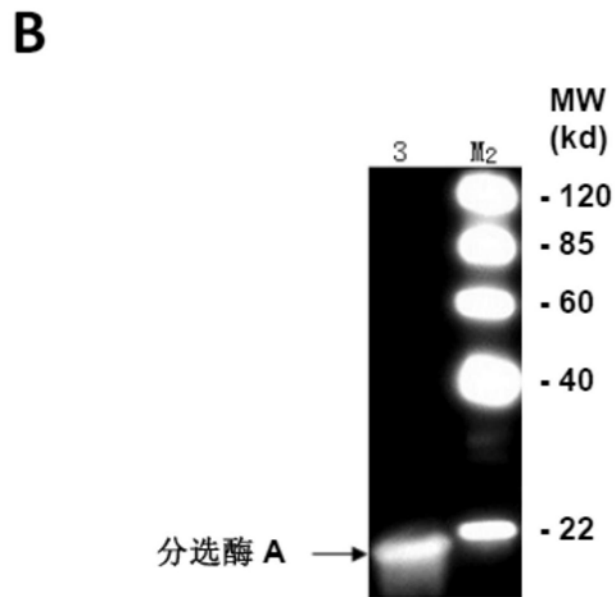
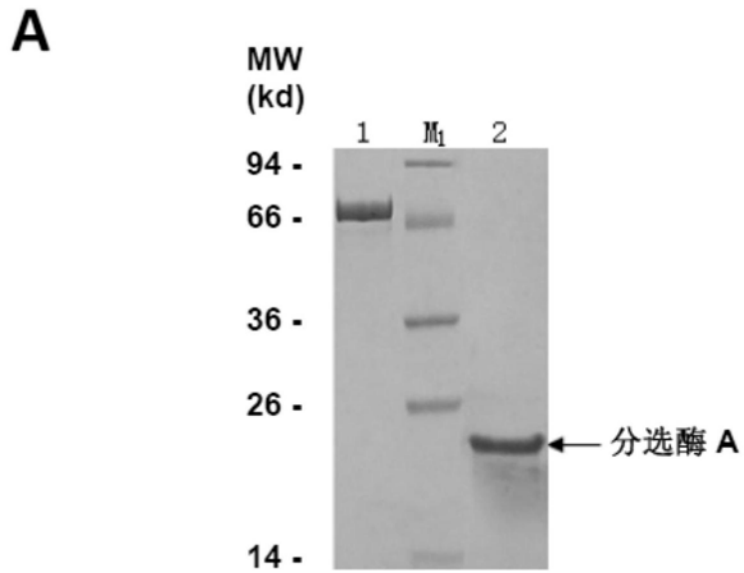


图7

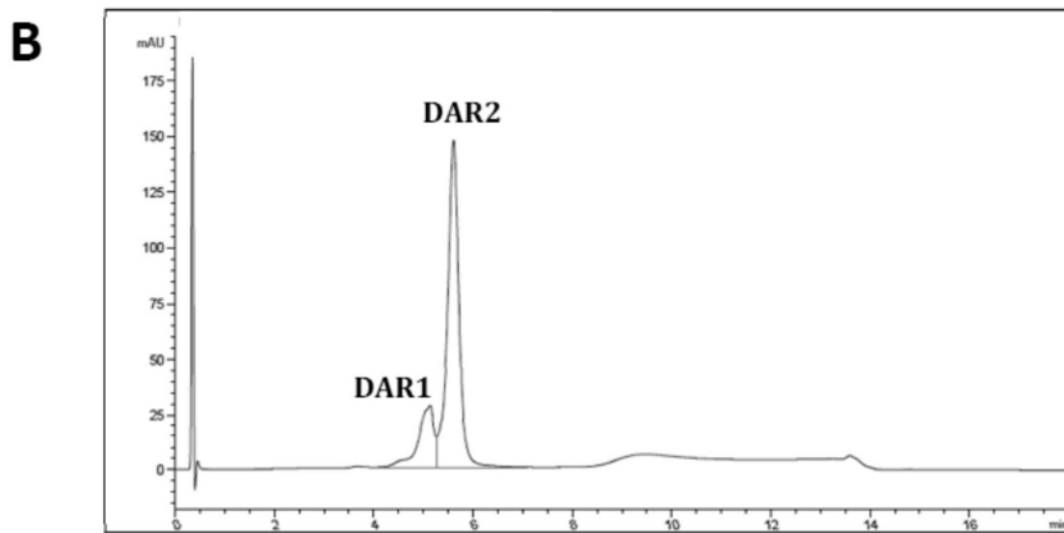
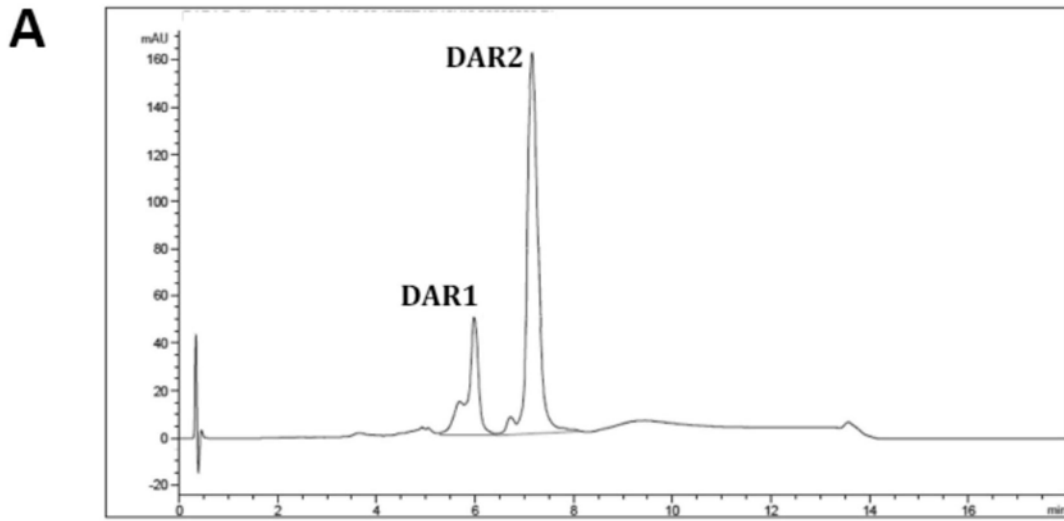


图8

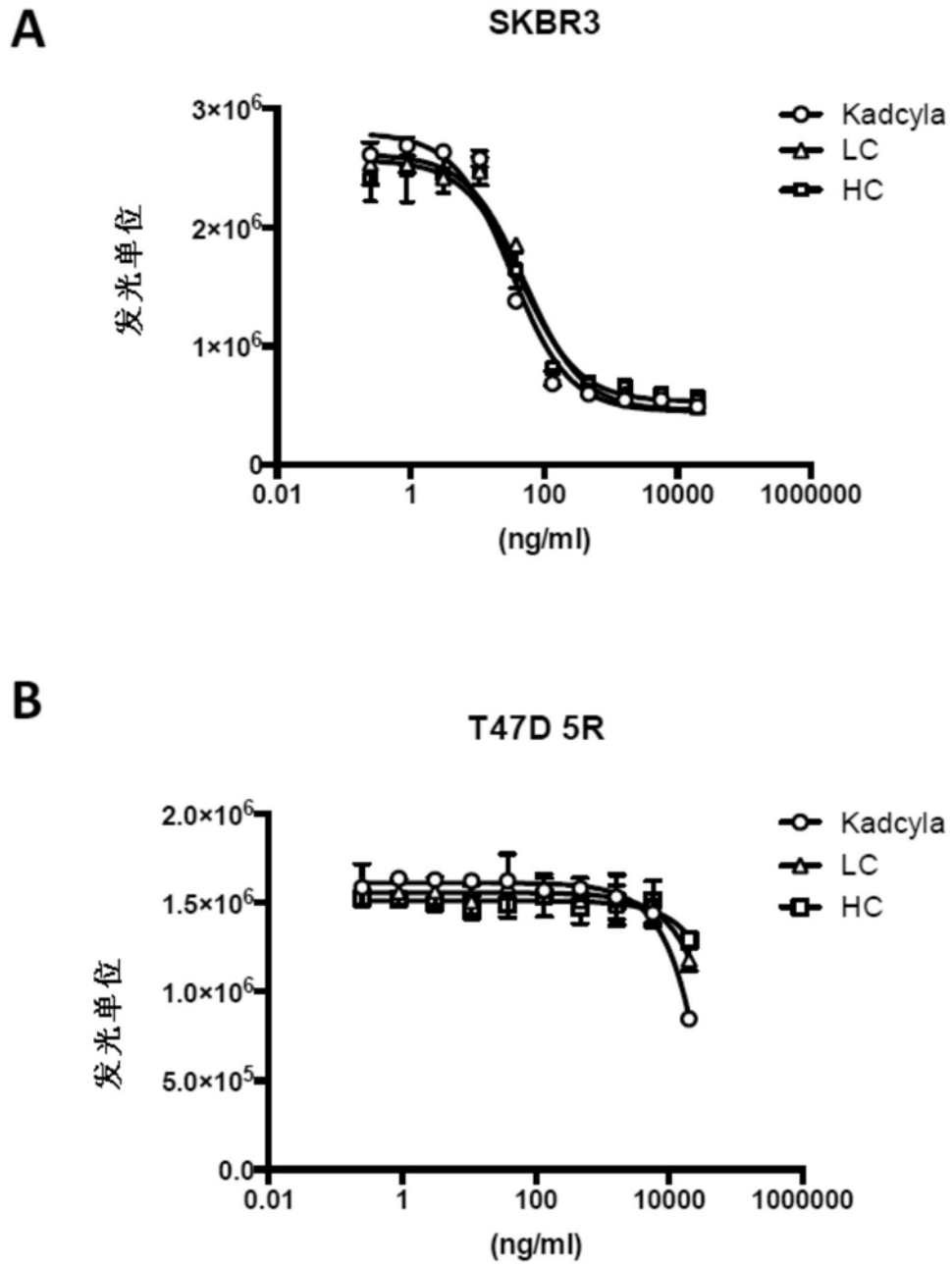


图9

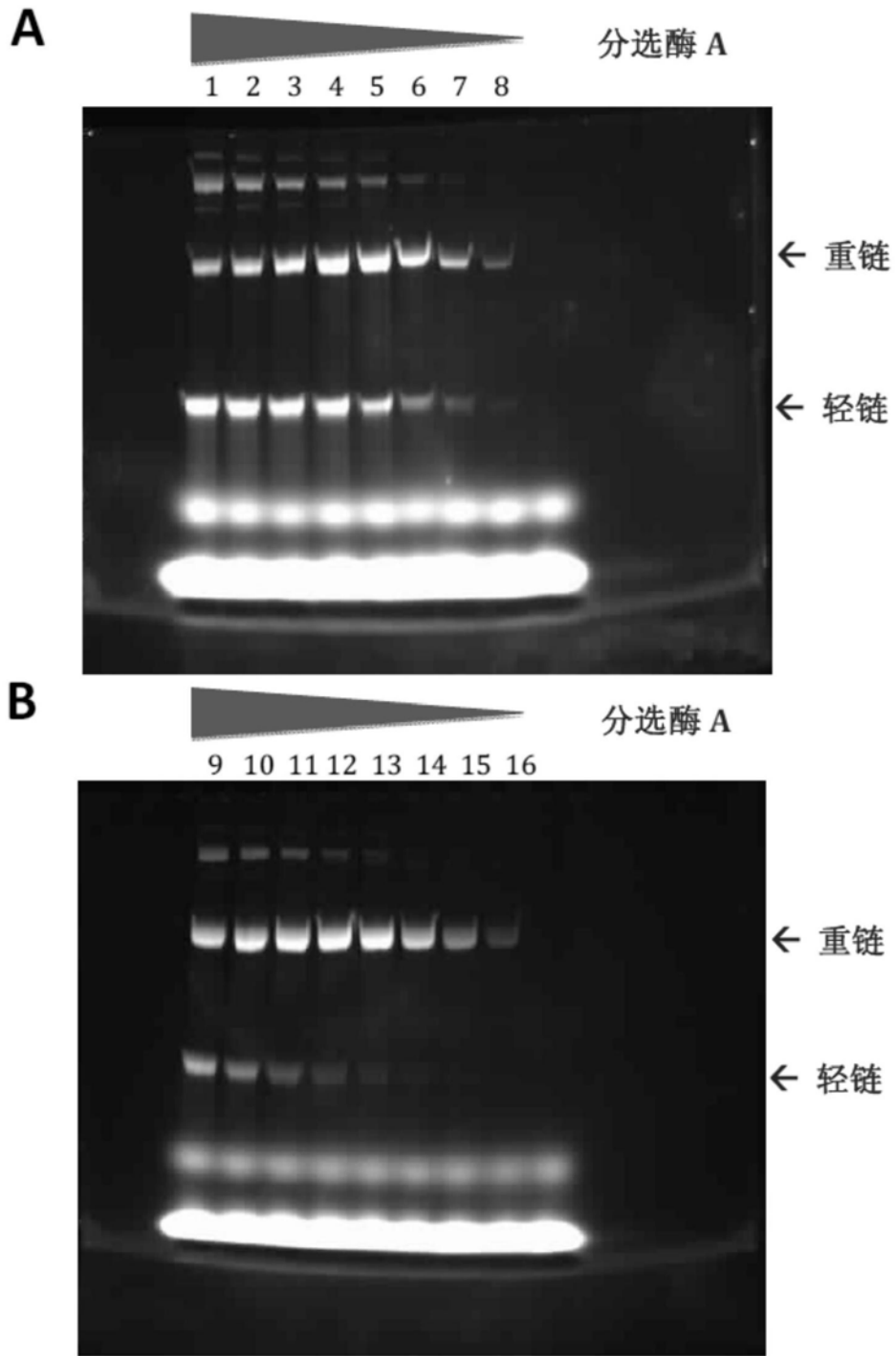


图10

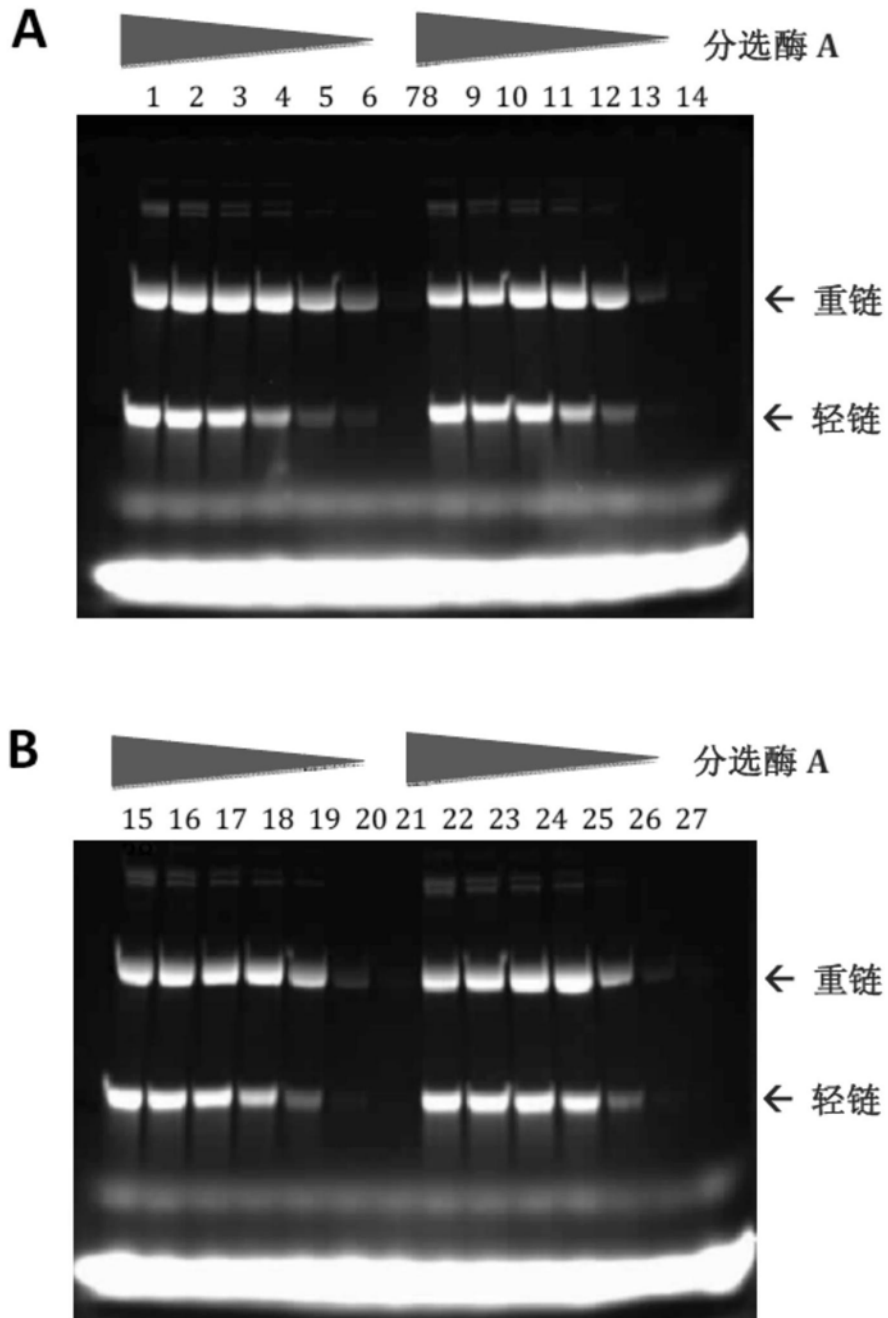


图11

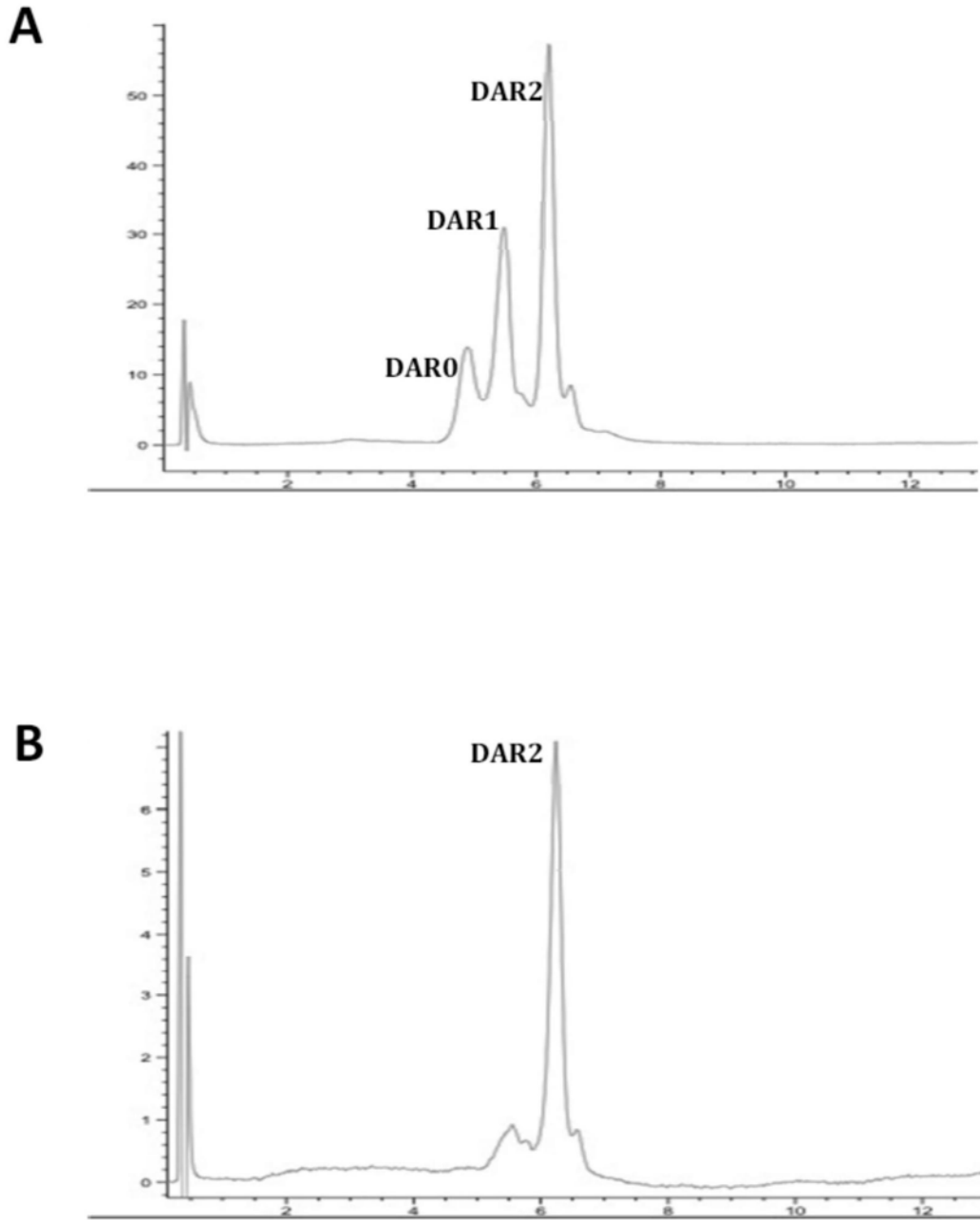
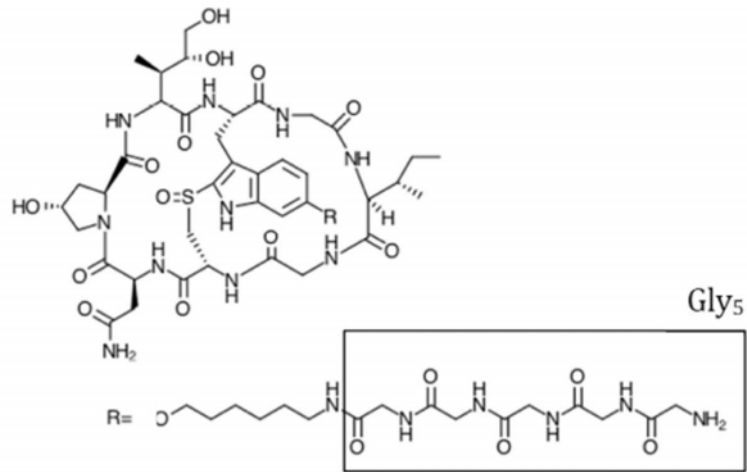
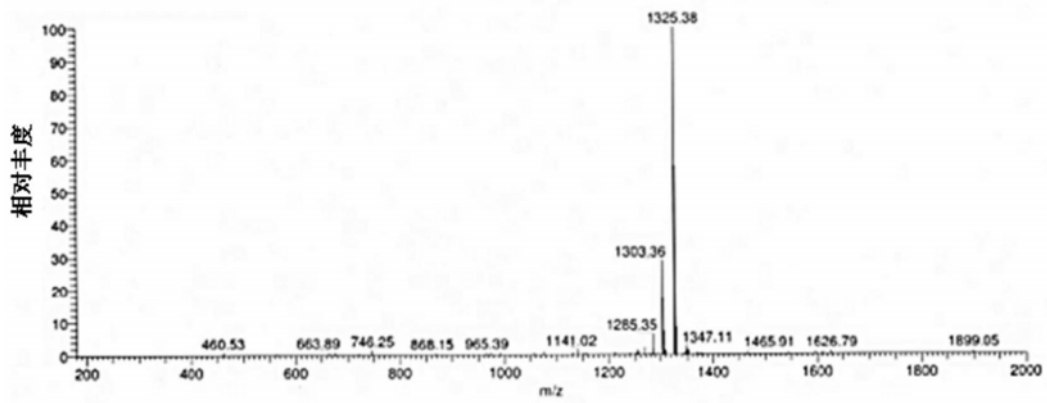


图12



Gly<sub>5</sub>-修饰的 α-鹅膏蕈碱

Gly<sub>5</sub>-修饰的 α-鹅膏蕈碱的质谱分析:



Gly<sub>5</sub>-修饰的 α-鹅膏蕈碱的 RP-HPLC 分析:

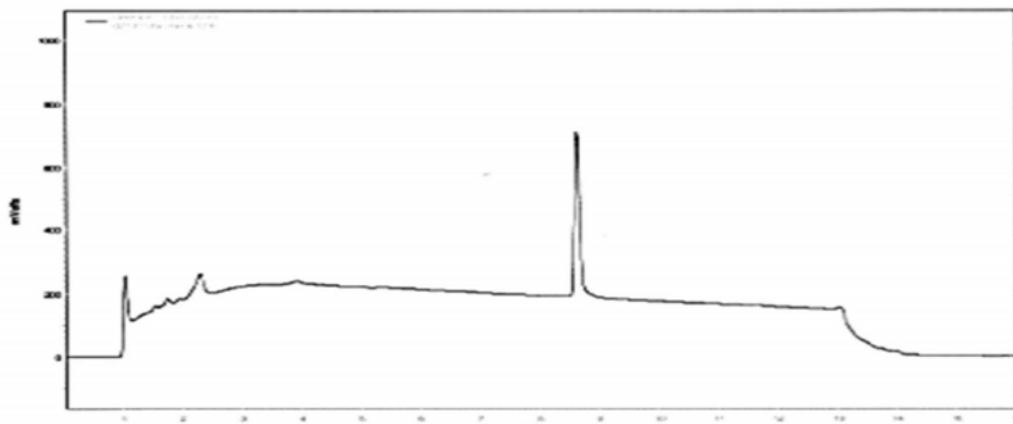
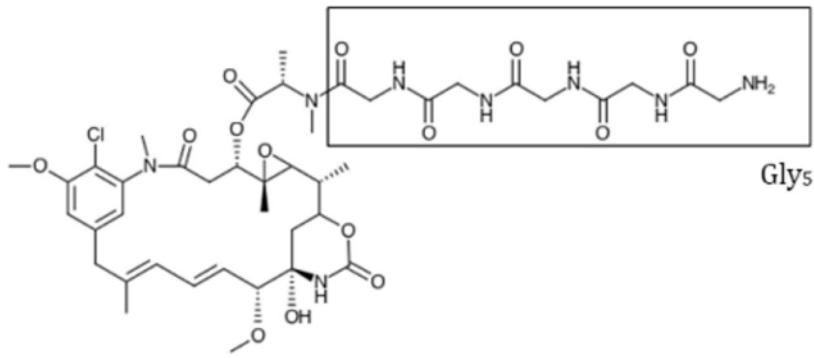


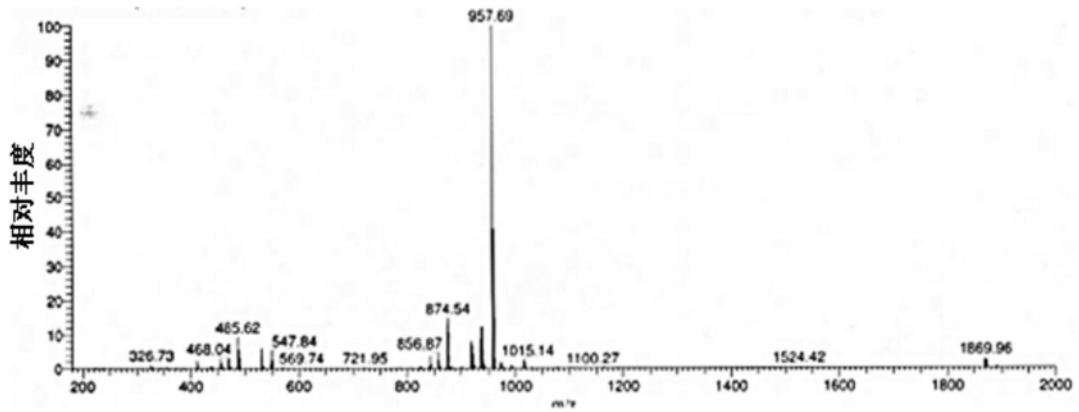
图13a





Gly<sub>5</sub>-修饰的美登素

Gly<sub>5</sub>-修饰的美登素的质谱分析



Gly<sub>5</sub>-修饰的美登素的 RP-HPLC 分析

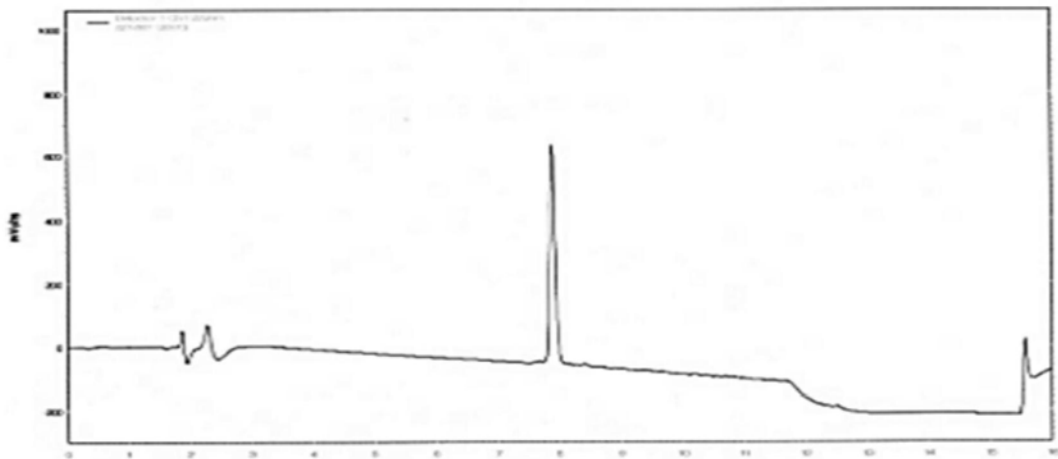
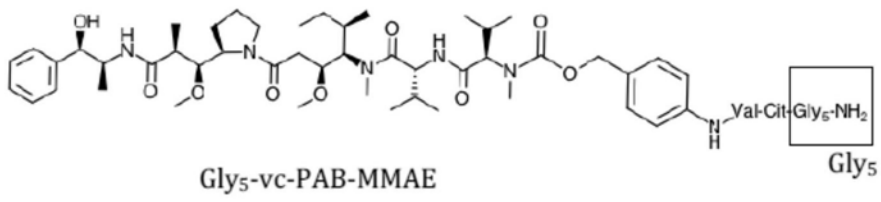
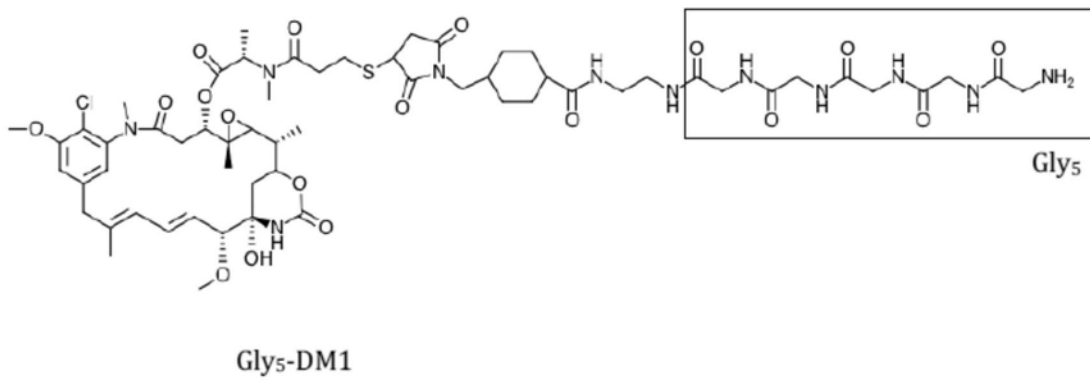


图13b

结构 1:



结构 2:



结构 3:

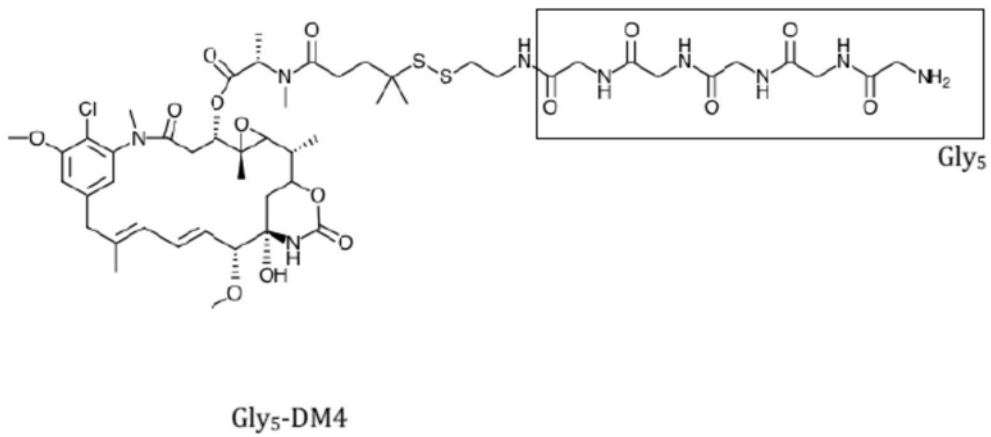
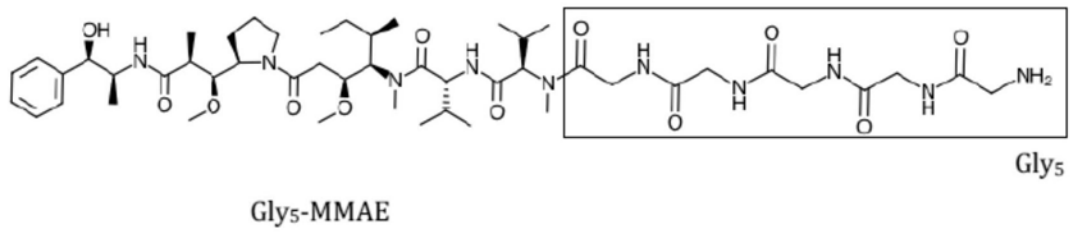
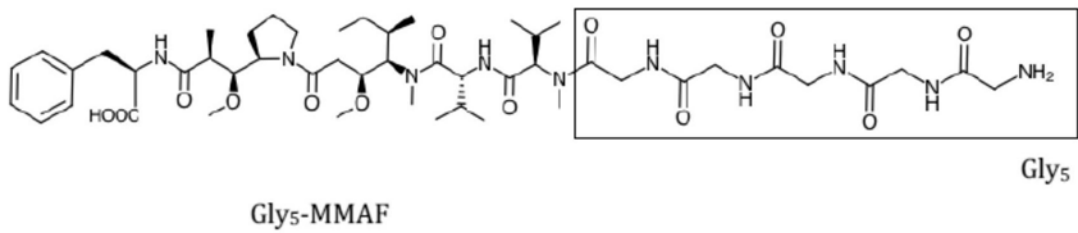


图14a

结构 4:



结构 5:



结构 6:

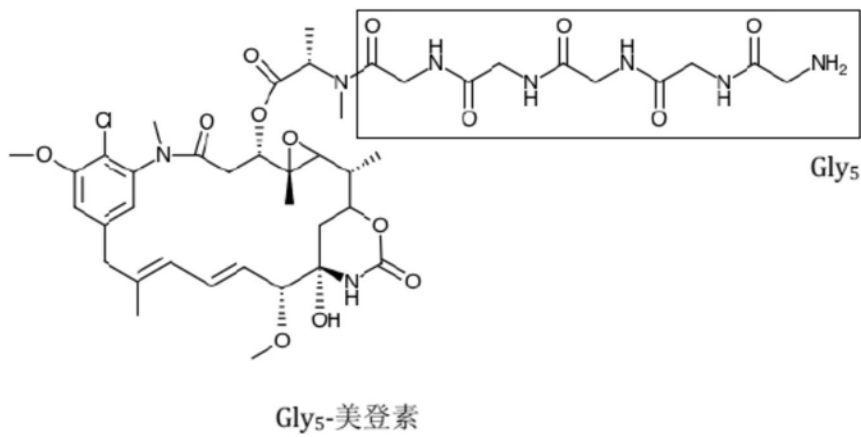
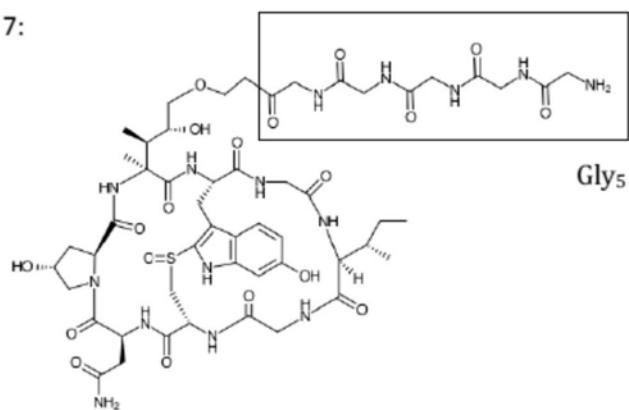
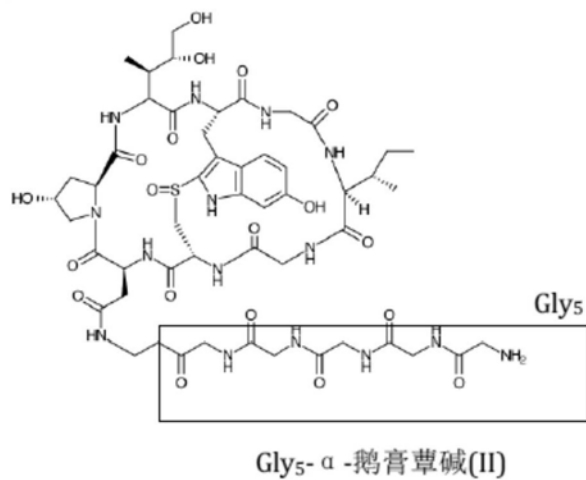


图14b

结构 7:



结构 8:



结构 9:

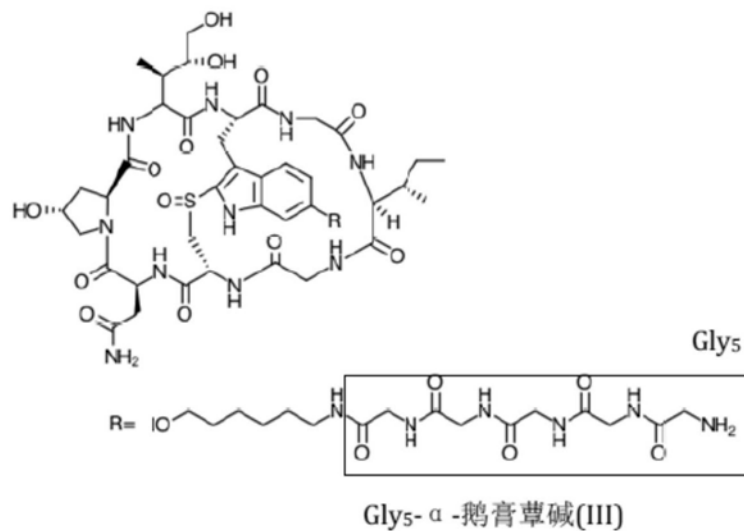


图14c

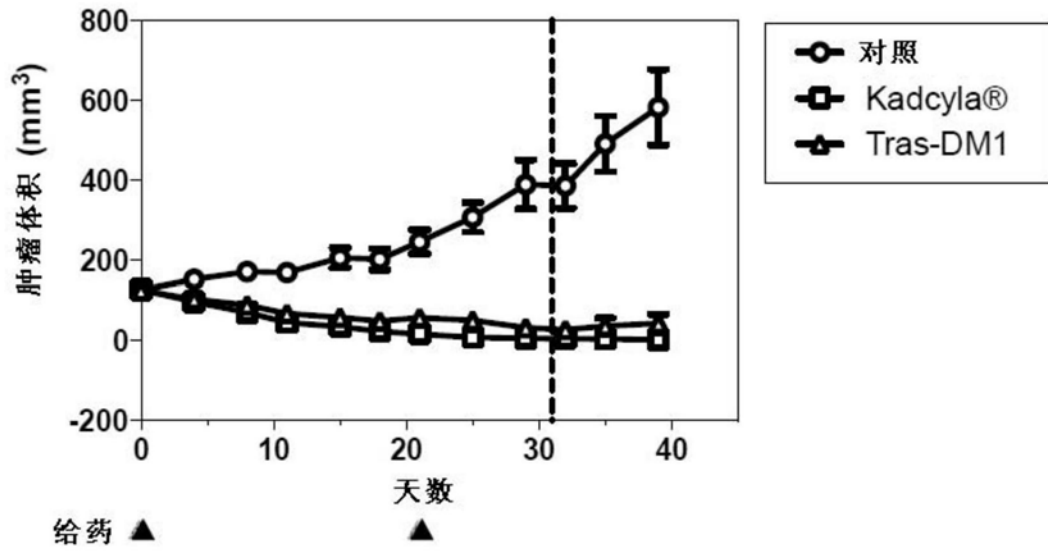


图15

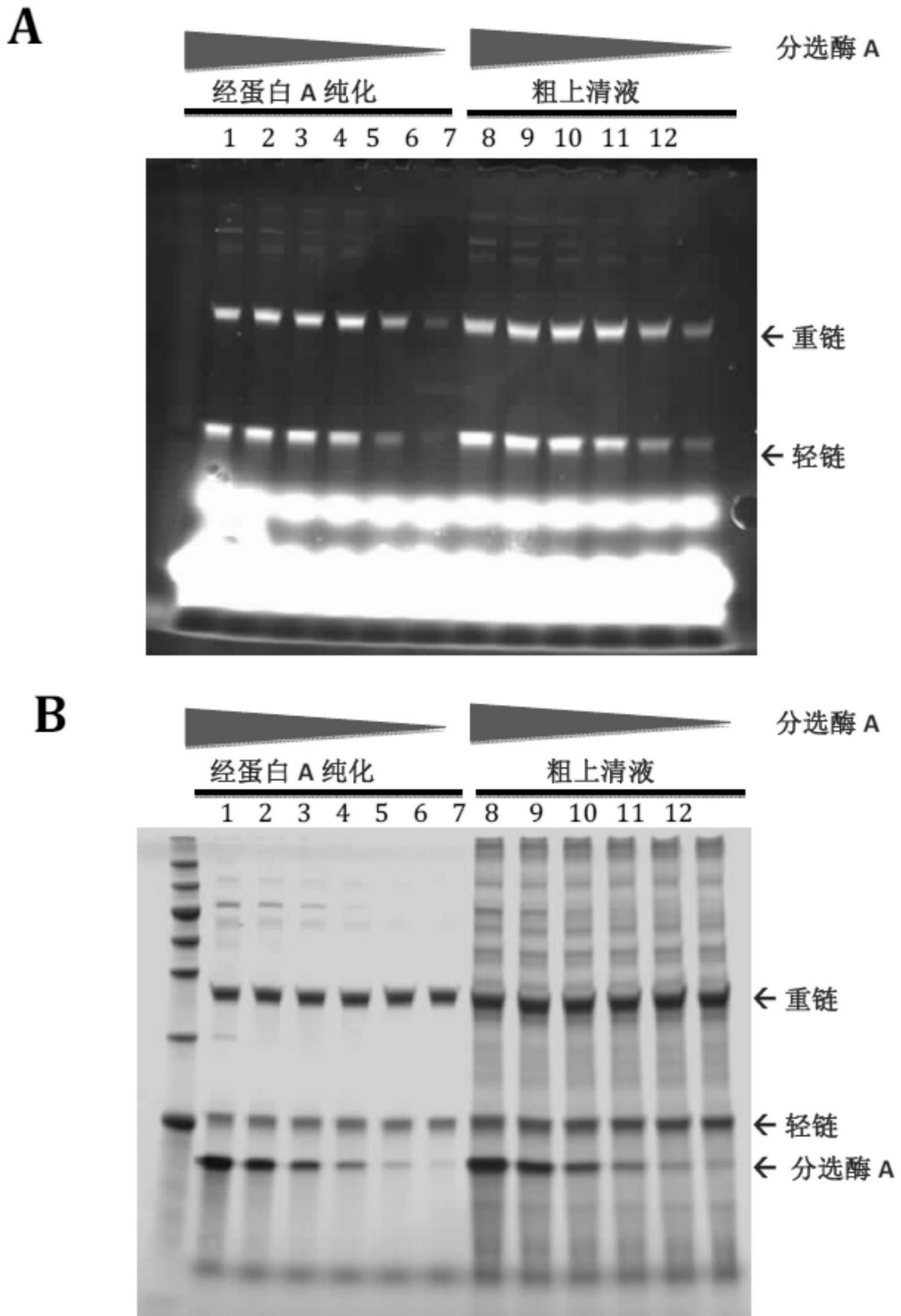


图16