(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 109111526 A
(43)申请公布日 2019.01.01

(21)申请号 201810666645.7

(22)申请日 2013.01.12

(30)优先权数据
61/586,099 2012.01.12 US
61/586,654 2012.01.13 US
61/667,901 2012.07.03 US
61/734,954 2012.12.07 US

(62)分案原申请数据
201380013452.7 2013.01.12

(71)申请人 比奥贝拉蒂治疗公司
地址 美国马萨诸塞州

(72)发明人 伊克塔·赛思·切布拉 刘童瑶
罗伯特·彼得斯 江海燕

(74)专利代理机构 上海胜康律师事务所 31263
代理人 樊英如 张静

(51)Int.Cl.
C07K 19/00(2006.01)
C12N 15/62(2006.01)
A61K 38/37(2006.01)
A61K 47/64(2017.01)
A61P 7/04(2006.01)

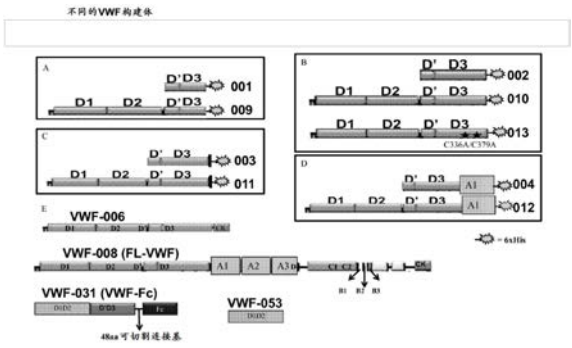
权利要求书1页 说明书131页
序列表18页 附图32页

(54)发明名称

嵌合因子VIII多肽及其用途

(57)摘要

本发明提供包含VWF的D'结构域和D3结构域的VWF片段、包含所述VWF片段和异源部分的嵌合蛋白或包含所述VWF片段和FVIII蛋白的嵌合蛋白及其使用方法。包含本发明的VWF片段的多肽链结合包含FVIII蛋白的多肽链或与之缔合，并且包含所述VWF片段的所述多肽链可阻止或抑制内源性VWF与所述FVIII蛋白的结合。通过阻止或抑制作为FVIII的半衰期限制因子的内源性VWF与所述FVIII的结合，所述VWF片段可诱导所述FVIII蛋白的半衰期延长。本发明还包括核苷酸、载体、宿主细胞、使用所述VWF片段或所述嵌合蛋白的方法。



CN 109111526 A

1.一种包含通过共价键连接的因子VIII("FVIII")蛋白和辅助部分(AM)的嵌合蛋白，其中所述辅助部分抑制或阻止内源性VWF结合所述FVIII蛋白。

2.一种编码根据权利要求1所述的嵌合蛋白的多核苷酸。

3.一种或多种载体，其包含根据权利要求2所述的多核苷酸和一个或多个可操作地连接至所述多核苷酸或所述多核苷酸组的启动子。

4.一种宿主细胞，其包含根据权利要求2所述的多核苷酸或根据权利要求3所述的载体。

5.一种药物组合物，其包含根据权利要求1所述的嵌合蛋白、根据权利要求2所述的多核苷酸、根据权利要求3所述的载体或根据权利要求4所述的宿主细胞以及可药用载体。

6.一种阻止或抑制FVIII蛋白与内源性VWF的相互作用的方法，其包括向需要其的受试者加入有效量的根据权利要求1所述的嵌合蛋白、根据权利要求2所述的多核苷酸、根据权利要求3所述的载体、根据权利要求4所述的宿主细胞或根据权利要求5所述的组合物，其中所述VWF片段抑制或阻止所述FVIII蛋白与内源性VWF的相互作用。

7.一种移除或减少FVIII蛋白的半衰期限制因子的方法，其中所述方法包括加入有效量的根据权利要求1所述的嵌合蛋白、根据权利要求2所述的多核苷酸、根据权利要求3所述的载体、根据权利要求4所述的宿主细胞或根据权利要求5所述的组合物，其中所述嵌合蛋白或由所述多核苷酸、所述载体编码的或由所述宿主细胞表达的所述嵌合蛋白阻止或抑制所述FVIII蛋白与内源性VWF的相互作用。

8.一种延长或增加FVIII蛋白的所述半衰期的方法，其中所述方法包括加入有效量的根据权利要求1所述的嵌合蛋白、根据权利要求2所述的多核苷酸、根据权利要求3所述的载体、根据权利要求4所述的宿主细胞或根据权利要求5所述的组合物，其中所述嵌合蛋白的所述VWF片段阻止或抑制所述FVIII蛋白与内源性VWF的相互作用。

9.一种治疗需要其的受试者的出血性疾病或病状的方法，其包括施用有效量的根据权利要求1所述的嵌合蛋白、根据权利要求2所述的多核苷酸、根据权利要求3所述的载体、根据权利要求4所述的宿主细胞或根据权利要求5所述的组合物，其中所述出血性疾病或病症选自由以下组成的组：出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血、髂腰肌鞘出血以及它们的任何组合。

10.一种制备嵌合蛋白的方法，其包括用根据权利要求2所述的多核苷酸、或根据权利要求3所述的载体转染一种或多种宿主细胞，以及在所述宿主细胞中表达所述VWF片段或所述嵌合蛋白。

嵌合因子VIII多肽及其用途

[0001] 本申请是申请日为2013年1月12日、中国专利申请号为201380013452.7(国际申请号为PCT/US2013/021330)、发明名称为“嵌合因子VIII多肽及其用途”的发明专利申请的分案申请。

[0002] 发明背景

[0003] 凝血是血液形成凝块的复杂过程。它是止血即阻止受损血管的失血的重要部分，其中受损血管壁被含血小板和血纤维蛋白的凝块覆盖，以终止流血并开始修复受损血管。凝血障碍可导致出血(溢血)或阻塞性凝血(血栓形成)的风险增加。

[0004] 凝血几乎在血管损伤破坏血管的内皮细胞内衬后的瞬间开始。血液暴露于蛋白质例如组织因子引发了血小板和血浆蛋白凝血因子血纤维蛋白原的变化。血小板立即在损伤部位形成血塞;这称为初期止血。二期止血与此同时进行:血浆中称为凝血因子或凝固因子的蛋白质在复杂的级联反应中响应，形成增强血小板栓的血纤维蛋白链。非限制性凝血因子包括但不限于:因子I(血纤维蛋白原)、因子II(凝血酶原)、组织因子、因子V(促凝血球蛋白原、不稳定因子)、因子VII(稳定因子、前转变素)、因子VIII(抗血友病因子A)、因子IX(抗血友病因子B或克雷司马斯因子(Christmas factor))、因子X(斯图亚特因子(Stuart-Prower factor))、因子XI(血浆促凝血酶原激酶先质)、因子XII(接触因子(Hageman factor))、因子XIII(血纤维蛋白稳定因子)、VWF、前激肽释放酶(弗莱彻因子(Flectcher factor))、高分子量激肽原(HMWK)(菲茨杰拉德因子)、纤连蛋白、抗凝血酶III、肝素辅因子II、蛋白C、蛋白S、蛋白Z、血纤维蛋白溶酶原、α2-抗血纤维蛋白溶酶、组织血纤维蛋白溶酶原激活物(tPA)、尿激酶、血纤维蛋白溶酶原激活物抑制剂-1(PAI1)和血纤维蛋白溶酶原激活物抑制剂-2(PAI2)。

[0005] 甲型血友病是编码凝血因子VIII(FVIII)的基因缺陷导致的出血性病症，并且在10,000例出生男性中有1-2例受其影响。Graw等,Nat.Rev.Genet.6(6):488-501(2005)。受甲型血友病影响的患者可通过输注纯化的或重组产生的FVIII治疗。然而，所有市售FVIII产品已知具有约8-12小时的半衰期，因此需要频繁向患者静脉内施用。参见Weiner M.A.和Cairo,M.S.,Pediatric Hematology Secrets,Lee,M.T.,12.Disorders of Coagulation,Elsevier Health Sciences,2001;Lillicrap,D.Thromb.Res.122增刊4:S2-8(2008)。此外，为延长FVIII半衰期已尝试过多种方法。例如，在开发中延长凝血因子半衰期的方法包括聚乙二醇化、糖聚乙二醇化以及与白蛋白缀合作用。参见Dumont等,Blood.119(13):3024-3030(2012年1月13日在线发表)。然而，不考虑所用的蛋白质工程，目前正在开发的长效FVIII产品具有提高的半衰期，但据报道半衰期有限—在临床前动物模型中仅提高约1.5至2倍。出处同上。在人类中显示出一致的结果，例如，据报道在甲型血友病患者中与ADVATE®相比，rFVIIIFc提高半衰期最多约1.7倍。出处同上。因此，尽管提高量很小，半衰期增加可表明存在其它T1/2限制因素。参见Liu,T.等,2007ISTH会议，摘要#P-M-035;Henrik,A.等,2011ISTH会议，摘要#P＝MO-181;Liu,T.等,2011ISTH会议，摘要#P-WE-131。

[0006] 血浆血管性血友病因子(VWF)具有大约12小时(在9至15小时的范围内)的半衰期。http://www.nhlbi.nih.gov/guidelines/vwd/2_scientificoverview.htm(最近访问2011

年10月22日)。VWF半衰期可受到多种因素的影响:糖基化模式、ADAMTS-13(具有血小板反应蛋白基序-13的解聚素和金属蛋白酶)以及VWF中的各种突变。

[0007] 在血浆中,95-98%的FVIII在具有全长VWF的紧密非共价复合物中循环。该复合物的形成对于在体内维持适当的FVIII血浆水平是重要的。Lenting等,Blood.92(11):3983-96(1998);Lenting等,J.Thromb.Haemost.5(7):1353-60(2007)。全长野生型FVIII主要以具有重链(MW 200kd)和轻链(MW 73kd)的异源二聚体存在。当FVIII由于重链第372和740位和轻链第1689位的蛋白水解而活化时,结合到FVIII的VWF从活化的FVIII移除。活化的FVIII连同活化的因子IX、钙和磷脂("因子X酶复合物(tenase complex)")一起参与因子X的活化,从而生成大量凝血酶。凝血酶继而切割血纤维蛋白原,形成可溶性血纤维蛋白单体,所述可溶性血纤维蛋白单体然后自发聚合形成可溶性血纤维蛋白聚合物。凝血酶还活化因子XIII,其连同钙一起用于交联并且稳定可溶性血纤维蛋白聚合物,形成交联(不溶性)血纤维蛋白。活化的FVIII通过蛋白水解从循环迅速清除。

[0008] 由于频繁给药以及给药方案造成的不便,仍需要开发不需要频繁施用的FVIII产品,即半衰期长于1.5至2倍半衰期限制的FVIII产品。

[0009] 发明概述

[0010] 本发明涉及包含因子VIII("FVIII")蛋白和辅助部分("AM")的嵌合蛋白,其中所述辅助部分抑制或阻止内源性VWF结合所述FVIII蛋白。所述FVIII蛋白和所述辅助部分通过共价键彼此连接,以阻止在内源性VWF存在下所述辅助部分的解离。在一个实施方案中,所述共价键是肽键、二硫键或连接基(linker),其强度足以阻止在内源性VWF存在下所述辅助部分从所述FVIII蛋白解离。在另一个实施方案中,所述辅助部分阻止所述FVIII蛋白通过VWF清除途径被清除。在其它实施方案中,所述辅助部分通过屏蔽或阻断所述FVIII蛋白上的VWF结合位点抑制或阻止内源性VWF结合所述FVIII蛋白。例如,VWF结合位点位于所述FVIII蛋白的所述A3结构域或所述C2结构域或所述A3结构域和所述C2结构域二者。

[0011] 在一些实施方案中,所述嵌合蛋白包括包含通过共价键彼此连接的FVIII蛋白和辅助部分的构建体,其中所述嵌合蛋白不包括引起所述FVIII蛋白的半衰期限制的FVIII半衰期限制因子,如全长VWF蛋白或成熟VWF蛋白。因此,在一些实施方案中,在内源性VWF存在下,所述嵌合蛋白的所述FVIII蛋白的所述半衰期可延长超过所述FVIII蛋白的所述半衰期限制。

[0012] 在某些实施方案中,所述辅助部分具有至少一种VWF样FVIII保护特性。所述VWF样FVIII保护特性的实例包括但不限于:保护所述FVIII蛋白不被一种或多种蛋白酶切割、保护所述FVIII蛋白不被活化、稳定所述FVIII蛋白的所述重链和/或所述轻链或防止所述FVIII蛋白被一种或多种清除剂受体清除。在一个实施方案中,所述辅助部分包括多肽、非多肽部分或它们二者。在另一个实施方案中,所述辅助部分可以是包含长度为至少约40个、至少约50个、至少约60个、至少约70个、至少约80个、至少约90个、至少约100个、至少约110个、至少约120个、至少约130个、至少约140个、至少约150个、至少约200个、至少约250个、至少约300个、至少约350个、至少约400个、至少约450个、至少约500个、至少约550个、至少约600个、至少约650个、至少约700个、至少约750个、至少约800个、至少约850个、至少约900个、至少约950个或至少约1000个氨基酸的氨基酸序列的多肽。在某些实施方案中,所述辅助部分包括VWF片段、免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS

序列、HAP序列、转铁蛋白或其片段或它们的任何组合。在其它实施方案中，所述辅助部分是包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合的非多肽部分。

[0013] 在某些实施方案中，所述辅助部分包括包含VWF的D'结构域和D3结构域的VWF片段，其中所述VWF片段通过除所述FVIII蛋白和所述辅助部分(VWF片段)之间的所述共价键之外的非共价键与所述FVIII蛋白缔合。在一个实例中，所述VWF片段是单体。在另一个实例中，所述VWF片段包括彼此之间一个或多个连接的两个、三个、四个、五个或六个VWF片段。

[0014] 在一个方面，所述嵌合蛋白包含辅助部分如VWF片段和至少一个异源部分(H1)以及所述辅助部分如VWF片段和所述异源部分(H1)之间任选的连接基。在一个实施方案中，所述异源部分(H1)可包括延长所述FVIII蛋白的所述半衰期的部分，如选自由以下组成的组的多肽：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合；或选自自由以下组成的组的非多肽部分：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在一个实施方案中，所述异源部分(H1)包括第一Fc区。在另一个实施方案中，所述异源部分(H1)包括包含至少约50个氨基酸、至少约100个氨基酸、至少约150个氨基酸、至少约200个氨基酸、至少约250个氨基酸、至少约300个氨基酸、至少约350个氨基酸、至少约400个氨基酸、至少约450个氨基酸、至少约500个氨基酸、至少约550个氨基酸、至少约600个氨基酸、至少约650个氨基酸、至少约700个氨基酸、至少约750个氨基酸、至少约800个氨基酸、至少约850个氨基酸、至少约900个氨基酸、至少约950个氨基酸或至少约1000个氨基酸的氨基酸序列。在其它实施方案中，所述嵌合蛋白包含所述辅助部分如VWF片段和所述异源部分(H1)之间的连接基，所述连接基是可切割的连接基。

[0015] 在另一个方面，所述嵌合蛋白中的所述FVIII蛋白包含FVIII和至少一个异源部分(H2)。在一个实施方案中，所述异源部分(H2)能够延长所述FVIII蛋白的所述半衰期，如选自由免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合组成的组的多肽或包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合的非多肽部分。在具体实施方案中，所述异源部分(H2)包括第二Fc区。

[0016] 在一些实施方案中，所述嵌合蛋白包含第一多肽链和第二多肽链，所述第一多肽链包含所述VWF片段、第一异源部分和连接基，所述第二多肽链包含所述FVIII蛋白和第二异源部分，其中所述第一多肽链和所述第二多肽链通过共价键彼此连接。在一个实例中，所述第一异源部分和所述第二异源部分通过所述共价键，如二硫键、肽键或连接基彼此连接，其中所述共价键防止体内所述第一多肽链中的所述VWF片段被内源性VWF替换。在一些实施方案中，所述FVIII蛋白和所述第二异源部分之间的所述连接基是可切割的连接基。

[0017] 在某些实施方案中，连接至所述VWF片段的所述第一异源部分(H1)和连接至所述FVIII蛋白的所述第二异源部分(H2)通过连接基，如scFc连接基连接，所述连接基是可加工的连接基。

[0018] 在其它实施方案中，所述嵌合蛋白中的所述FVIII蛋白还包含第三异源部分(H3)、第四异源部分(H4)、第五异源部分(H5)、第六异源部分(H6)或它们的任何组合。在一个实施方案中，所述第三异源部分(H3)、所述第四异源部分(H4)、所述第五异源部分(H5)、所述第

六异源部分(H6)中的一者或多者能够延长所述FVIII蛋白的所述半衰期。在另一个实施方案中,所述第三异源部分(H3)、所述第四异源部分(H4)、所述第五异源部分(H5)和所述第六异源部分(H6)连接至FVIII的所述C末端或N末端或插入FVIII的两个氨基酸之间。在其它实施方案中,所述第三异源部分(H3)、所述第四异源部分(H4)、所述第五异源部分(H5)或所述第六异源部分(H6)中的一者或多者包括包含至少约50个氨基酸、至少约100个氨基酸、至少约150个氨基酸、至少约200个氨基酸、至少约250个氨基酸、至少约300个氨基酸、至少约350个氨基酸、至少约400个氨基酸、至少约450个氨基酸、至少约500个氨基酸、至少约550个氨基酸、至少约600个氨基酸、至少约650个氨基酸、至少约700个氨基酸、至少约750个氨基酸、至少约800个氨基酸、至少约850个氨基酸、至少约900个氨基酸、至少约950个氨基酸或至少约1000个氨基酸的氨基酸序列。

[0019] 在一些实施方案中,所述FVIII蛋白和所述第二异源部分之间的所述连接基或所述VWF片段和所述第一异源部分之间的所述连接基还包含位于所述连接基的所述N-末端区的第一切割位点(P1)、位于所述连接基的所述C-末端区的第二切割位点(P2)或它们二者。在其它实施方案中,所述FVIII蛋白和所述辅助部分之间的所述连接基、所述FVIII蛋白和所述第二异源部分之间的所述连接基以及所述VWF片段和所述第一异源部分之间的所述连接基中的一者或多者具有约1至约2000个氨基酸的长度。

[0020] 在其它实施方案中,所述嵌合蛋白包含FVIII蛋白和辅助部分,它们通过所述FVIII蛋白和所述辅助部分之间的连接基连接,其中所述连接基还包含分选酶识别基序,如所述序列LPXTG(SEQ ID NO:106)。

[0021] 本发明涉及血管性血友病因子(von Willebrand Factor,VWF)片段,所述血管性血友病因子片段包含VWF的所述D'结构域和所述D3结构域,其中所述VWF片段结合因子VIII(FVIII)并且抑制内源性VWF与FVIII蛋白结合。在一个实施方案中,本发明的所述VWF片段不是SEQ ID NO:2的第764至1274位氨基酸。在一个实施方案中,无所述VWF片段的所述FVIII蛋白具有相当于野生型FVIII的半衰期。在另一个实施方案中,所述FVIII蛋白是包含FVIII和能够延长FVIII的半衰期的异源部分的融合蛋白。所述异源部分可以是多肽、非多肽部分或它们二者。所述异源多肽部分可选自由以下组成的组:免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合。在其它实施方案中,所述异源部分是免疫球蛋白恒定区或其部分,如Fc区。在其它实施方案中,所述非多肽部分选自由以下组成的组:聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在某些实施方案中,所述FVIII蛋白包含第一多肽链和第二多肽链,其中所述第一多肽链包含FVIII和第一Fc区,并且所述第二多肽链包含无FVIII的第二Fc区。

[0022] 在另一个实施方案中,所述VWF片段延长FVIII的半衰期。所述D'结构域的所述氨基酸序列与SEQ ID NO:2的第764至866位氨基酸可具有至少90%、95%、96%、97%、98%、99%或100%的同一性。另外,所述D3结构域的所述氨基酸序列与SEQ ID NO:2的第867至1240位氨基酸可具有至少90%、95%、96%、97%、98%、99%或100%的同一性。在某些实施方案中,所述VWF片段包含在对应于SEQ ID NO:2的第1099位残基、第1142位残基或它们二者的残基处的至少一个氨基酸取代。在具体实施方案中,VWF片段包含、基本上由或由SEQ ID NO:2的第764至1240位氨基酸组成。所述VWF片段还可包含VWF的所述D1结构域、所述D2

结构域或所述D1和D2结构域。在一些实施方案中，所述VWF片段还包含选自由以下组成的组的VWF结构域：所述A1结构域、所述A2结构域、所述A3结构域、所述D4结构域、所述B1结构域、所述B2结构域、所述B3结构域、所述C1结构域、所述C2结构域、所述CK结构域、它们的一个或多个片段以及它们的任何组合。在其它实施方案中，所述VWF片段是聚乙二醇化的、糖基化的、羟乙基淀粉化的(hesylated)或聚唾液酸化的。

[0023] 本发明还涉及包含本文所述的VWF片段、异源部分以及所述VWF片段和所述异源部分之间的任选的连接基的嵌合蛋白。所述异源部分可以是多肽、非多肽部分或它们二者。在一个实施方案中，所述异源多肽部分选自由以下组成的组：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合。在另一个实施方案中，所述异源非多肽部分选自由以下组成的组：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在具体实施方案中，所述异源部分是第一Fc区。所述嵌合蛋白还可包含第二Fc区，其中所述第二Fc区连接至所述第一Fc区或与所述第一Fc区缔合，或连接至所述VWF片段或与所述VWF片段缔合。

[0024] 在一个方面，本发明的嵌合蛋白包含选自由以下组成的组的式：

[0025] (aa)V-L1-H1-L2-H2，

[0026] (bb)H2-L2-H1-L1-V，

[0027] (cc)H1-L1-V-L2-H2，和

[0028] (dd)H2-L2-V-L1-H1，

[0029] 其中V是本文所述的一个或多个VWF片段，

[0030] L1和L2中的每个是任选的连接基；

[0031] H1是第一异源部分；

[0032] (-)是肽键或一个或多个氨基酸；并且

[0033] H2是任选的第二异源部分。

[0034] 在一个实施方案中，H1是第一异源部分，如本领域已知的半衰期延长分子。在一个实施方案中，所述第一异源部分是多肽。所述第一异源多肽部分选自由以下组成的组：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合。在另一个实施方案中，H1是选自由以下组成的组的非多肽部分：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。H2是任选的第二异源部分，如本领域已知的半衰期延长分子。在一个实施方案中，所述第二异源部分可选自由以下组成的组：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合。在另一个实施方案中，H2是选自由以下组成的组的非多肽部分：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在某些实施方案中，H1是第一Fc区，并且H2是第二Fc区。所述第一Fc区和所述第二Fc区可以是相同的或不同的，并且可以通过连接基或共价键如二硫键彼此连接。在另一个实施方案中，所述第二Fc区连接至因子VIII蛋白或与因子VIII蛋白缔合。任选地，可存在作为半衰期延长因子的第三异源部分H3，其连接至所述VWF片段、所述第一异源部分或所述第二异源部分。所述第三异源部分的非限制性实例可包括多肽或非多肽部分或它们二者。在一个实施方案中，所述第三异源多肽部分可选自由以下组成的组：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段

或它们的任何组合。在另一个实施方案中，H2是选自由以下组成的组的非多肽部分：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在一些实施方案中，H3通过可切割的连接基如凝血酶可切割连接基连接至所述VWF片段或所述第一或所述第二异源部分。所述连接基的非限制性实例在本文别处有所公开。

[0035] 在另一个方面，本发明提供包含本文所述的VWF片段、FVIII蛋白以及所述VWF片段和所述FVIII蛋白之间的任选的连接基的嵌合蛋白。所述VWF片段与所述FVIII蛋白结合。在一个实施方案中，嵌合蛋白包含本文所述的VWF片段，其连接至异源部分。所述异源部分可以是延长所述蛋白质的所述半衰期的部分，所述异源部分包括多肽、非多肽部分或它们二者。此类异源多肽部分的实例包括，如免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、其任何衍生物或变体或它们的任何组合。非多肽部分的实例包括，如聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。在另一个实施方案中，所述异源部分是连接至所述VWF片段的第一Fc区。在其它实施方案中，所述嵌合蛋白还包含连接至所述FVIII蛋白的第二Fc区。所述VWF片段或所述FVIII蛋白可通过连接基分别连接至所述第一Fc区或所述第二Fc区。在其它实施方案中，嵌合蛋白包含连接至第一异源部分如第一Fc区的本文所述的VWF片段，以及连接至第二异源部分如第二Fc区的FVIII蛋白，其中所述VWF片段还通过连接基或通过共价键连接至所述第二异源部分(如，第二Fc区)或所述FVIII蛋白，或所述第一异源部分(如，Fc区)还通过连接基或通过共价键连接至所述FVIII蛋白或所述第二异源部分(如，第二Fc区)。在一些实施方案中，所述嵌合蛋白的所述FVIII具有部分B-结构域。在一些实施方案中，具有部分B-结构域的所述FVIII蛋白是FVIII198(SEQ ID NO:105)。在其它实施方案中，所述嵌合蛋白还包含分选酶识别基序。

[0036] 在一些实施方案中，作为本发明的结果，与无所述VWF片段的FVIII蛋白或野生型FVIII相比，所述FVIII蛋白的所述半衰期延长。所述FVIII蛋白的所述半衰期比无所述VWF片段的FVIII蛋白的所述半衰期长至少约1.5倍、至少约2倍、至少约2.5倍、至少约3倍、至少约4倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约11倍或至少约12倍。在一个实施方案中，FVIII的所述半衰期比野生型FVIII的所述半衰期长约1.5倍至约20倍、约1.5倍至约15倍或约1.5倍至约10倍。在另一个实施方案中，与野生型FVIII或无所述VWF片段的FVIII蛋白相比，所述FVIII的所述半衰期延长约2倍至约10倍、约2倍至约9倍、约2倍至约8倍、约2倍至约7倍、约2倍至约6倍、约2倍至约5倍、约2倍至约4倍、约2倍至约3倍、约2.5倍至约10倍、约2.5倍至约9倍、约2.5倍至约8倍、约2.5倍至约7倍、约2.5倍至约6倍、约2.5倍至约5倍、约2.5倍至约4倍、约2.5倍至约3倍、约3倍至约10倍、约3倍至约9倍、约3倍至约8倍、约3倍至约7倍、约3倍至约6倍、约3倍至约5倍、约3倍至约4倍、约4倍至约6倍、约5倍至约7倍或约6倍至约8倍。在其它实施方案中，FVIII的所述半衰期为至少约17小时、至少约18小时、至少约19小时、至少约20小时、至少约21小时、至少约22小时、至少约23小时、至少约24小时、至少约25小时、至少约26小时、至少约27小时、至少约28小时、至少约29小时、至少约30小时、至少约31小时、至少约32小时、至少约33小时、至少约34小时、至少约35小时、至少约36小时、至少约48小时、至少约60小时、至少约72小时、至少约84小时、至少约96小时或至少约108小时。在其它实施方案中，FVIII的所述半衰期为约15小时至约两周、约16小时至约一周、约17小时至约一周、约18小时至约一周、约19小时至约一周、

约20小时至约一周、约21小时至约一周、约22小时至约一周、约23小时至约一周、约24小时至约一周、约36小时至约一周、约48小时至约一周、约60小时至约一周、约24小时至约六天、约24小时至约五天、约24小时至约四天、约24小时至约三天或约24小时至约两天。

[0037] 在一些实施方案中，每个受试者的所述FVIII蛋白的所述平均半衰期为约15小时、约16小时、约17小时、约18小时、约19小时、约20小时、约21小时、约22小时、约23小时、约24小时(1天)、约25小时、约26小时、约27小时、约28小时、约29小时、约30小时、约31小时、约32小时、约33小时、约34小时、约35小时、约36小时、约40小时、约44小时、约48小时(2天)、约54小时、约60小时、约72小时(3天)、约84小时、约96小时(4天)、约108小时、约120小时(5天)、约六天、约七天(一周)、约八天、约九天、约10天、约11天、约12天、约13天或约14天。

[0038] 在另一个方面，本发明的嵌合蛋白包含选自由以下组成的组的式：

[0039] (a)V-L1-H1-L3-C-L2-H2，

[0040] (b)H2-L2-C-L3-H1-L1-V，

[0041] (c)C-L2-H2-L3-V-L1-H1，

[0042] (d)H1-L1-V-L3-H2-L2-C，

[0043] (e)H1-L1-V-L3-C-L2-H2，

[0044] (f)H2-L2-C-L3-V-L1-H1，

[0045] (g)V-L1-H1-L3-H2-L2-C，

[0046] (h)C-L2-H2-L3-H1-L1-V，

[0047] (i)H2-L3-H1-L1-V-L2-C，

[0048] (j)C-L2-V-L1-H1-L3-H2，

[0049] (k)V-L2-C-L1-H1-L3-H2，和

[0050] (l)H2-L3-H1-L1-C-L2-V，

[0051] 其中V是本文所述的VWF片段；

[0052] L1或L2中的每个是任选的连接基，如凝血酶可切割连接基；

[0053] L3是任选的连接基，如scFc连接基，如可加工的连接基；

[0054] H1或H2中的每个是任选的异源部分；并且

[0055] C是FVIII蛋白；并且

[0056] (-)是肽键或一个或多个氨基酸。

[0057] 在其它方面，本发明的嵌合蛋白包含选自由以下组成的组的式：

[0058] (m)V-L1-H1:H2-L2-C，

[0059] (n)V-L1-H1:C-L2-H2，

[0060] (o)H1-L1-V:H2-L2-C，

[0061] (p)H1-L1-V:C-L2-H2，

[0062] (q)V:C-L1-H1:H2，

[0063] (r)V:H1-L1-C:H2，

[0064] (s)H2:H1-L1-C:V，

[0065] (t)C:V-L1-H1:H2，和

[0066] (u)C:H1-L1-V:H2，

[0067] 其中V是本文所述的VWF片段；

[0068] L1或L2中的每个是任选的连接基，如凝血酶可切割连接基；

[0069] H1或H2中的每个是任选的异源部分；并且

[0070] C是FVIII蛋白；

[0071] (-)是肽键或一个或多个氨基酸；并且

[0072] (:)是H1和H2之间、V和C之间以及V和H1和C和H2之间的化学或物理缔合。(:)表示化学缔合，如至少一个非肽键。在某些实施方案中，所述化学缔合即(:)是共价键。在一些实施方案中，H1和H2之间的所述缔合是共价键，如二硫键。在其它实施方案中，所述化学缔合即(:)是非共价相互作用，如离子相互作用、疏水相互作用、亲水相互作用、范德华相互作用、氢键。在某些实施方案中，所述FVIII蛋白和所述VWF片段之间的所述缔合是非共价键。在其它实施方案中，(:)是非肽共价键。在其它实施方案中，(:)是肽键。在一个实施方案中，H1是第一异源部分。在一个实施方案中，所述第一异源部分能够延长所述FVIII活性的半衰期。在另一个实施方案中，所述第一异源部分是多肽、非多肽部分或它们二者。在一个实施方案中，所述第一异源多肽部分可选自由以下组成的组：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合。在另一个实施方案中，所述非多肽部分选自：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在一些实施方案中，H2是第二异源部分。所述第二异源部分也可以是本领域已知的半衰期延长因子(half-life extender)，并且可以是多肽、非多肽部分或它们二者的组合。在一个实施方案中，所述第二异源部分选自由以下组成的组：免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段以及它们的任何组合。在某些实施方案中，所述非多肽部分选自由以下组成的组：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物以及它们的任何组合。在具体实施方案中，H1是第一Fc区。在一些实施方案中，H2是第二Fc区。任选地，可存在半衰期延长因子第三异源部分H3。H3可通过任选的连接基，如可切割的连接基，如凝血酶可切割连接基连接至V、C、H1或H2中的一者或多者。所述第三异源部分的非限制性实例可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、聚乙二醇(PEG)、PAS序列和羟乙基淀粉(HES)或其衍生物。

[0073] 在某些实施方案中，用于使式(a)至(u)的所述VWF片段、所述FVIII蛋白、所述第一异源部分和/或所述第二异源部分彼此连接的一个或多个所述连接基是可切割的连接基。用于所述嵌合蛋白的一个或多个所述切割位点可被选自由以下组成的组的蛋白酶切割：因子XIa、因子XIIa、激肽释放酶、因子VIIa、因子IXa、因子Xa、因子IIa(凝血酶)、弹性蛋白酶-2、粒酶-B、TEV、肠激酶、蛋白酶3C、分选酶A、MMP-12、MMP-13、MMP-17和MMP-20。在其它实施方案中，用于式(a)至(l)的一个或多个连接基(如，L3)包括可加工的连接基。所述可加工的连接基可在分泌时被胞内酶切割。所述可加工的连接基可包含位于所述连接基的所述N-末端区的第一切割位点(P1)、位于所述连接基的所述C-末端区的第二切割位点(P2)或它们二者。

[0074] 在一些实施方案中，一个或多个用于本发明的所述连接基具有至少约1至2000个氨基酸的长度。在具体实施方案中，一个或多个用于本发明的所述连接基具有至少约20、35、42、48、73、98、144、288、324、576或864个氨基酸的长度。在具体实施方案中，一个或多个所述连接基包含gly/ser肽。所述gly/ser肽可以是$(Gly4Ser)_3$或$(Gly4Ser)_4$。

[0075] 在其它方面，嵌合蛋白中的FVIII蛋白是功能性因子VIII蛋白。所述FVIII蛋白可

包含一个或多个其选自由以下组成的组的FVIII结构域:所述A1结构域、所述A2结构域、所述B结构域、所述A3结构域、所述C1结构域、所述C2结构域、一个或多个其片段以及它们的任何组合。在一个实施方案中,所述FVIII蛋白包含所述B结构域或其部分。在另一个实施方案中,所述FVIII蛋白是SQ B结构域缺失的FVIII。在其它实施方案中,所述FVIII蛋白包含单链FVIII。在其它实施方案中,所述FVIII蛋白包含FVIII的重链和因子VIII的轻链,其中所述重链和所述轻链通过金属键彼此缔合。在某些实施方案中,所述FVIII蛋白具有对低密度脂蛋白受体相关蛋白(LRP)的低亲合力或不与低密度脂蛋白受体相关蛋白(LRP)结合。例如,用于本发明的FVIII蛋白可包含降低对LRP的所述亲合力或消除与所述LRP的结合的至少一个氨基酸取代。所述至少一个氨基酸取代的非限制性实例位于与全长成熟FVIII的第471位残基、第484位残基、第487位残基、第490位残基、第497位残基、第2092位残基、第2093位残基或两个或更多个其组合对应的残基处。在一些实施方案中,本发明的嵌合蛋白中的所述FVIII蛋白包含至少一个氨基酸取代,其使得所述FVIII蛋白比无所述取代的FVIII蛋白更稳定。在其它实施方案中,所述FVIII蛋白包含所述A2结构域中的至少一个氨基酸取代和所述A3结构域中的至少一个氨基酸取代,其中所述A2结构域和所述A3结构域通过共价键彼此缔合。所述A2结构域中的所述氨基酸取代的非限制性实例位于与全长成熟FVIII的第662或664位残基对应的残基处。此外,所述A3结构域中的所述氨基酸取代的非限制性实例位于与全长成熟的聚唾液酸化FVIII的第1826或1828位残基对应的残基处。

[0076] 在另外的方面,本发明提供编码本文所述的VWF片段或本文所述的嵌合蛋白的多核苷酸,或一组包含第一核苷酸链和第二核苷酸链的多核苷酸,其中所述第一核苷酸链编码所述VWF片段,并且所述第二核苷酸链编码所述第二Fc区或所述嵌合蛋白的所述凝血因子或其片段。在一个实施方案中,所述多核苷酸组还包含第三多核苷酸链,其编码属于所述枯草杆菌蛋白酶样前蛋白转化酶家族的前蛋白转化酶。所述前蛋白转化酶的非限制性实例包括前蛋白转化酶枯草杆菌蛋白酶/kexin 3型(PACE或PCSK3)、前蛋白转化酶枯草杆菌蛋白酶/kexin 5型(PCSK5或PC5)、前蛋白转化酶枯草杆菌蛋白酶/kexin 7型(PCSK7或PC7)或酵母Kex 2。在其它方面,本发明包括包含所述多核苷酸或所述多核苷酸组和一个或多个可操作地连接至所述多核苷酸或所述多核苷酸组的启动子的载体,或一组包括第一载体和第二载体的载体,其中所述第一载体编码所述多核苷酸组的所述第一多核苷酸链,并且所述第二载体编码所述多核苷酸组的所述第二多核苷酸链。所述载体组还可包含第三载体,其包含编码PC5或PC7的第三多核苷酸链。在一些实施方案中,所述载体还包含PACE。在一些实施方案中,PACE切割所述VWF片段的所述D1D2结构域。

[0077] 在一些方面,本发明涉及包含所述VWF片段、所述嵌合蛋白、所述多核苷酸、所述多核苷酸组、所述载体或所述载体组以及可药用载体的药物组合物。本发明的组合物可延长因子VIII的半衰期。在其它方面,本发明包括包含所述多核苷酸、所述多核苷酸组、所述载体或所述载体组的宿主细胞。

[0078] 在其它方面,本发明涉及包含FVIII蛋白、辅助部分和任选的连接基的嵌合蛋白,其中所述辅助部分抑制或阻止内源性VWF结合所述FVIII蛋白,并且具有至少一种VWF样FVIII保护特性。所述VWF样FVIII保护特性包括保护所述FVIII蛋白不被一种或多种蛋白酶切割,保护所述FVIII蛋白不被活化,稳定所述FVIII蛋白的所述重链和/或所述轻链或防止所述FVIII蛋白被一种或多种清除剂受体清除。

[0079] 所述嵌合蛋白中的所述辅助部分可通过屏蔽或阻断所述FVIII蛋白上的VWF结合位点抑制或阻止内源性VWF结合所述FVIII蛋白。在一些实施方案中，所述VWF结合位点位于所述FVIII蛋白的所述A3结构域或所述C2结构域或所述FVIII蛋白的A3结构域和C2结构域二者。在另一个实施方案中，所述VWF结合位点是与SEQ ID NO:16的第1669至1689位和第2303至2332位氨基酸对应的所述氨基酸序列。在一些实施方案中，所述辅助部分是多肽、非多肽部分或它们二者。用作所述辅助部分的所述多肽可包含长度为至少40个、50个、60个、70个、80个、90个、100个、110个、120个、130个、140个、150个、200个、250个、300个、350个、400个、450个、500个、550个、600个、650个、700个、750个、800个、850个、900个、950个或1000个氨基酸的氨基酸序列。例如，用作辅助部分的所述多肽可选自由以下组成的组：VWF片段、免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、其它半衰期延长技术以及它们的任何组合。用作辅助部分的所述非多肽部分可选自由以下组成的组：聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)或其衍生物以及它们的任何组合。在一个实施方案中，所述辅助部分是本文描述的所述VWF片段。所述辅助部分和所述FVIII蛋白可例如通过连接基连接或彼此缔合。所述连接基可包括可切割的连接基，如凝血酶可切割连接基。

[0080] 在一个方面，本发明提供阻止或抑制FVIII蛋白与内源性VWF结合的方法，所述方法包括将有效量的所述VWF片段、所述嵌合蛋白、所述多核苷酸或所述多核苷酸组加入到包含FVIII蛋白或编码所述FVIII蛋白的多核苷酸的细胞，其中所述VWF片段结合所述FVIII蛋白。在另一个方面，本发明包括阻止或抑制所述FVIII蛋白与内源性VWF结合的方法，所述方法包括将有效量的所述嵌合蛋白、所述多核苷酸或所述多核苷酸组加入到需要其的受试者，其中所述VWF片段结合所述FVIII蛋白，并且从而阻止或抑制所述FVIII蛋白的结合。在一些方面，本发明包括延长或增加FVIII蛋白的半衰期的方法，其中所述方法包括将有效量的所述VWF片段、所述嵌合蛋白、所述多核苷酸或所述多核苷酸组加入到包含FVIII蛋白或编码所述FVIII蛋白的多核苷酸的细胞或加入到需要其的受试者，其中所述VWF片段结合所述FVIII蛋白。在其它方面，本发明涉及阻止或抑制FVIII蛋白从细胞清除的方法，其中所述方法包括将有效量的所述VWF片段、所述嵌合蛋白、所述多核苷酸或所述多核苷酸组加入到包含FVIII蛋白或编码所述FVIII蛋白的多核苷酸的细胞或加入到需要其的受试者，其中所述VWF片段结合所述FVIII蛋白。

[0081] 在另一个方面，本发明涉及治疗需要其的受试者的出血性疾病或病症的方法，所述方法包括施用有效量的所述VWF片段、所述嵌合蛋白、所述多核苷酸或所述多核苷酸组，其中所述出血性疾病或病症选自由以下组成的组：出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血以及髂腰肌鞘出血。在其它实施方案中，所述治疗是预防性的或按需的。在其它实施方案中，本发明涉及治疗需要其的受试者的与2N型血管性血友病相关的疾病或病症的方法，所述方法包括施用有效量的所述VWF片段、所述嵌合蛋白、所述多核苷酸或所述多核苷酸组，其中所述疾病或病症得以治疗。

[0082] 本发明提供以下实施方式：

[0083] 实施方式1：一种包含通过共价键连接的因子VIII("FVIII")蛋白和辅助部分(AM)

的嵌合蛋白，其中所述辅助部分抑制或阻止内源性VWF结合所述FVIII蛋白。

[0084] 实施方式2：根据实施方式1所述的嵌合蛋白，其中所述共价键在内源性VWF存在下阻止所述辅助部分从所述FVIII蛋白的解离。

[0085] 实施方式3：根据实施方式1或实施方式2所述的嵌合蛋白，其中所述共价键是肽键。

[0086] 实施方式4：根据实施方式1至3中任一项所述的嵌合蛋白，其中所述共价键是二硫键。

[0087] 实施方式5：根据实施方式1至4中任一项所述的嵌合蛋白，其中所述共价键是所述FVIII蛋白和所述辅助部分之间的连接基。

[0088] 实施方式6：根据实施方式1至5中任一项所述的嵌合蛋白，其中所述辅助部分阻止所述FVIII蛋白通过VWF清除途径被清除。

[0089] 实施方式7：根据实施方式1至6中任一项所述的嵌合蛋白，其中所述辅助部分通过屏蔽或阻断所述FVIII蛋白上的VWF结合位点抑制或阻止内源性VWF结合所述FVIII蛋白。

[0090] 实施方式8：根据实施方式7所述的嵌合蛋白，其中所述VWF结合位点位于所述FVIII蛋白的所述A3结构域或所述C2结构域或所述A3结构域和所述C2结构域二者。

[0091] 实施方式9：根据实施方式8所述的嵌合蛋白，其中所述VWF结合位点是与SEQ ID NO:16的第1669至1689位和第2303至2332位氨基酸对应的所述氨基酸序列。

[0092] 实施方式10：根据实施方式1至9中任一项所述的嵌合蛋白，其中所述嵌合蛋白不包含FVIII半衰期限制因子。

[0093] 实施方式11：根据实施方式10所述的嵌合蛋白，其中所述FVIII半衰期限制因子包括全长VWF蛋白或成熟VWF蛋白。

[0094] 实施方式12：根据实施方式1至11中任一项所述的嵌合蛋白，其中在内源性VWF存在下，所述FVIII蛋白的所述半衰期可延长超过所述FVIII蛋白的所述半衰期限制。

[0095] 实施方式13：根据实施方式1至12中任一项所述的嵌合蛋白，其中所述辅助部分具有至少一种VWF样FVIII保护特性。

[0096] 实施方式14：根据实施方式13所述的嵌合蛋白，其中所述VWF样FVIII保护特性包括保护所述FVIII蛋白不被一种或多种蛋白酶切割、保护所述FVIII蛋白不被活化、稳定所述FVIII蛋白的所述重链和/或所述轻链或防止所述FVIII蛋白被一种或多种清除剂受体清除。

[0097] 实施方式15：根据实施方式1至14中任一项所述的嵌合蛋白，其中所述辅助部分包括多肽、非多肽部分或它们二者。

[0098] 实施方式16：根据实施方式15所述的嵌合蛋白，其中所述多肽包含长度为至少约40个、至少约50个、至少约60个、至少约70个、至少约80个、至少约90个、至少约100个、至少约110个、至少约120个、至少约130个、至少约140个、至少约150个、至少约200个、至少约250个、至少约300个、至少约350个、至少约400个、至少约450个、至少约500个、至少约550个、至少约600个、至少约650个、至少约700个、至少约750个、至少约800个、至少约850个、至少约900个、至少约950个或至少约1000个氨基酸的氨基酸序列。

[0099] 实施方式17：根据实施方式1至16中任一项所述的嵌合蛋白，其中所述辅助部分包括VWF片段、免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP

序列、转铁蛋白或其片段或它们的任何组合。

[0100] 实施方式18:根据实施方式17所述的嵌合蛋白,其中所述非多肽部分包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0101] 实施方式19:根据实施方式17所述的嵌合蛋白,其中所述辅助部分包括包含VWF的D'结构域和D3结构域的VWF片段,其中所述VWF片段通过除所述共价键之外的非共价键与所述FVIII蛋白缔合。

[0102] 实施方式20:根据实施方式17或实施方式19所述的嵌合蛋白,其中所述VWF片段是单体。

[0103] 实施方式21:根据实施方式19或实施方式20所述的嵌合蛋白,其中所述VWF片段包括彼此之间一个或多个连接的两个、三个、四个、五个或六个VWF片段。

[0104] 实施方式22:根据实施方式1至21中任一项所述的嵌合蛋白,其中所述VWF片段包含至少一个异源部分(H1)以及所述VWF片段和所述异源部分(H1)之间的任选的连接基。

[0105] 实施方式23:根据实施方式22所述的嵌合蛋白,其中连接至所述VWF片段的所述至少一个异源部分(H1)包括多肽、非多肽部分或它们二者。

[0106] 实施方式24:根据实施方式22或实施方式23所述的嵌合蛋白,其中所述异源部分(H1)包括延长所述FVIII蛋白的所述半衰期的部分。

[0107] 实施方式25:根据实施方式24所述的嵌合蛋白,其中所述异源部分(H1)包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的任何组合。

[0108] 实施方式26:根据实施方式24所述的嵌合蛋白,其中所述非多肽部分包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0109] 实施方式27:根据实施方式25所述的嵌合蛋白,其中所述异源部分(H1)包括第一Fc区。

[0110] 实施方式28:根据实施方式24所述的嵌合蛋白,其中所述异源部分(H1)包括包含至少约50个氨基酸、至少约100个氨基酸、至少约150个氨基酸、至少约200个氨基酸、至少约250个氨基酸、至少约300个氨基酸、至少约350个氨基酸、至少约400个氨基酸、至少约450个氨基酸、至少约500个氨基酸、至少约550个氨基酸、至少约600个氨基酸、至少约650个氨基酸、至少约700个氨基酸、至少约750个氨基酸、至少约800个氨基酸、至少约850个氨基酸、至少约900个氨基酸、至少约950个氨基酸或至少约1000个氨基酸的氨基酸序列。

[0111] 实施方式29:根据实施方式22所述的嵌合蛋白,其中所述嵌合蛋白包含所述VWF片段和所述异源部分(H1)之间的连接基,所述连接基是可切割的连接基。

[0112] 实施方式30:根据实施方式29所述的嵌合蛋白,其中所述可切割的连接基包含一个或多个可切割位点。

[0113] 实施方式31:根据实施方式29或实施方式30中任一项所述的嵌合蛋白,其中所述可切割的连接基能够被选自由以下组成的组的蛋白酶切割:因子XIa、因子XIIa、激肽释放酶、因子VIIa、因子IXa、因子Xa、因子IIa(凝血酶)、弹性蛋白酶-2、粒酶-B、TEV、肠激酶、蛋白酶3C、分选酶A、MMP-12、MMP-13、MMP-17和MMP-20。

[0114] 实施方式32:根据实施方式29所述的嵌合蛋白,其中所述可切割的连接基包含TLDPRSFLLRNPNDKYEPFWEDEEK(SEQ ID NO:56)。

[0115] 实施方式33:根据实施方式29至32中任一项所述的嵌合蛋白,其中所述可切割的连接基包含一个或多个切割位点,所述切割位点包含选自由以下组成的组的氨基酸序列:RRRR(SEQ ID NO:52)、RKRRKR(SEQ ID NO:53)、RRRRS(SEQ ID NO:54)、TQSFNDFTR(SEQ ID NO:47)、SVSQTSKLTR(SEQ ID NO:48)、DFLAEGGGVR(SEQ ID NO:49)、TTKIKPR(SEQ ID NO:50)、LVPRG(SEQ ID NO:55)、ALRPRVVGGA(SEQ ID NO:51)、KLTRAET(SEQ ID NO:29)、DFTRVVG(SEQ ID NO:30)、TMTRIVGG(SEQ ID NO:31)、SPFRSTGG(SEQ ID NO:32)、LQVRIVGG(SEQ ID NO:33)、PLGRIVGG(SEQ ID NO:34)、IEGRTVGG(SEQ ID NO:35)、LTPRSLLV(SEQ ID NO:36)、LGPVSGVP(SEQ ID NO:37)、VAGDSLEE(SEQ ID NO:38)、GPAGLGGA(SEQ ID NO:39)、GPAGLRGA(SEQ ID NO:40)、APLGLRLR(SEQ ID NO:41)、PALPLVAQ(SEQ ID NO:42)、ENLYFQG(SEQ ID NO:43)、DDDKIVGG(SEQ ID NO:44)、LEVLFQGP(SEQ ID NO:45)和LPKTGSES(SEQ ID NO:46)。

[0116] 实施方式34:根据实施方式1至33中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含FVIII和至少一个异源部分(H2)。

[0117] 实施方式35:根据实施方式34所述的嵌合蛋白,其中所述异源部分(H2)能够延长所述FVIII蛋白的所述半衰期。

[0118] 实施方式36:根据实施方式34或实施方式35所述的嵌合蛋白,其中所述异源部分(H2)包括多肽、非多肽部分或它们二者。

[0119] 实施方式37:根据实施方式34或实施方式35所述的嵌合蛋白,其中所述异源部分(H2)包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的任何组合。

[0120] 实施方式38:根据实施方式34或实施方式35所述的嵌合蛋白,其中所述非多肽部分包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0121] 实施方式39:根据实施方式34所述的嵌合蛋白,其中所述异源部分(H2)包括第二Fc区。

[0122] 实施方式40:根据实施方式1至39中任一项所述的嵌合蛋白,所述嵌合蛋白包含第一多肽链和第二多肽链,所述第一多肽链包含所述VWF片段、第一异源部分和连接基,所述第二多肽链包含所述FVIII蛋白和第二异源部分,其中所述第一多肽链和所述第二多肽链通过共价键彼此连接。

[0123] 实施方式41:根据实施方式40所述的嵌合蛋白,其中所述第一异源部分和所述第二异源部分通过所述共价键彼此连接,其中所述共价键防止体内所述第一多肽链中的所述VWF片段被内源性VWF替换。

[0124] 实施方式42:根据实施方式41所述的嵌合蛋白,其中所述共价键是二硫键。

[0125] 实施方式43:根据实施方式34至42中任一项所述的嵌合蛋白,其中所述FVIII蛋白通过连接基连接至所述第二异源部分(H2)。

[0126] 实施方式44:根据实施方式43所述的嵌合蛋白,其中所述FVIII蛋白和所述第二异源部分之间的所述连接基是可切割的连接基。

[0127] 实施方式45:根据实施方式34至44所述的嵌合蛋白,其中所述第一异源部分(H1)和所述第二异源部分(H2)通过连接基连接。

[0128] 实施方式46:根据实施方式45所述的嵌合蛋白,其中所述连接基是scFc连接基。

[0129] 实施方式47:根据实施方式46所述的嵌合蛋白,其中所述scFc连接基是可加工的连接基。

[0130] 实施方式48:根据实施方式1至47所述的嵌合蛋白包含选自由以下组成的组的式:

[0131] (a)V-L1-H1-L3-C-L2-H2,

[0132] (b)H2-L2-C-L3-H1-L1-V,

[0133] (c)C-L2-H2-L3-V-L1-H1,

[0134] (d)H1-L1-V-L3-H2-L2-C,

[0135] (e)H1-L1-V-L3-C-L2-H2,

[0136] (f)H2-L2-C-L3-V-L1-H1,

[0137] (g)V-L1-H1-L3-H2-L2-C,

[0138] (h)C-L2-H2-L3-H1-L1-V,

[0139] (i)H2-L3-H1-L1-V-L2-C,

[0140] (j)C-L2-V-L1-H1-L3-H2,

[0141] (k)V-L2-C-L1-H1-L3-H2,和

[0142] (l)H2-L3-H1-L1-C-L2-V,

[0143] 其中V包括包含VWF的所述D'结构域和所述D3结构域的VWF片段;

[0144] L1是任选的连接基;

[0145] L2是任选的连接基;

[0146] (a)至(f)中的L3是任选的连接基,

[0147] (g)至(l)中的L3是任选的scFc连接基,

[0148] H1和H2中的每个包括任选的异源部分;

[0149] C包括FVIII蛋白;并且

[0150] (-)是肽键或一个或多个氨基酸。

[0151] 实施方式49:根据实施方式1至47所述的嵌合蛋白包含选自由以下组成的组的式:

[0152] (m)V-L1-H1:H2-L2-C,

[0153] (n)V-L1-H1:C-L2-H2,

[0154] (o)H1-L1-V:H2-L2-C,

[0155] (p)H1-L1-V:C-L2-H2,

[0156] (q)V:C-L1-H1:H2,

[0157] (r)V:H1-L1-C:H2,

[0158] (s)H2:H1-L1-C:V,

[0159] (t)C:V-L1-H1:H2,和

[0160] (u)C:H1-L1-V:H2,

[0161] 其中V是包含VWF的所述D'结构域和所述D3结构域的VWF片段;

[0162] L1是任选的连接基;

[0163] L2是任选的连接基;

[0164] H1是第一异源部分;

[0165] H2是第二异源部分;

[0166] C是FVIII蛋白;

[0167] (-)是肽键或一个或多个氨基酸;并且

[0168] (:)是所述H1和所述H2之间的共价键。

[0169] 实施方式50:.根据实施方式48和实施方式49所述的嵌合蛋白,其中所述VWF片段和所述FVIII蛋白通过除所述共价键、所述肽主链或所述一个或多个氨基酸之外的非共价键彼此缔合。

[0170] 实施方式51:根据实施方式48和实施方式49所述的嵌合蛋白,其中所述VWF片段抑制或阻止内源性VWF与所述FVIII蛋白的结合。

[0171] 实施方式52:根据实施方式49至51中任一项所述的嵌合蛋白,其中所述H1和所述H2之间的所述共价键是二硫键。

[0172] 实施方式53:根据实施方式48至52中任一项所述的嵌合蛋白,其中H1包括多肽、非多肽部分或它们二者。

[0173] 实施方式54:根据实施方式53所述的嵌合蛋白,其中H1包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的任何组合。

[0174] 实施方式55:根据实施方式53或实施方式54所述的嵌合蛋白,其中H1包括第一Fc区。

[0175] 实施方式56:根据实施方式53所述的嵌合蛋白,其中所述非多肽部分包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0176] 实施方式57:根据实施方式48至56中任一项所述的嵌合蛋白,其中H2包括多肽、非多肽部分或它们二者。

[0177] 实施方式58:根据实施方式57所述的嵌合蛋白,其中H2包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的任何组合。

[0178] 实施方式59:根据实施方式48至58中任一项所述的嵌合蛋白,其中H2包括第二Fc区。

[0179] 实施方式60:根据实施方式59所述的嵌合蛋白,其中所述非多肽部分包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0180] 实施方式61:根据实施方式60所述的嵌合蛋白,其中所述共价键是二硫键。

[0181] 实施方式62:根据实施方式1至61中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含第三异源部分(H3)。

[0182] 实施方式63:根据实施方式1至62中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含第四异源部分(H4)。

[0183] 实施方式64:根据实施方式1至63中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含第五异源部分(H5)。

[0184] 实施方式65:根据实施方式1至64中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含第六异源部分(H6)。

[0185] 实施方式66:根据实施方式61至65中任一项所述的嵌合蛋白,其中所述第三异源部分(H3)、所述第四异源部分(H4)、所述第五异源部分(H5)、所述第六异源部分(H6)中的一者或多者能够延长所述FVIII蛋白的所述半衰期。

[0186] 实施方式67:根据实施方式1至66中任一项所述的嵌合蛋白,其中所述第三异源部分(H3)、所述第四异源部分(H4)、所述第五异源部分(H5)和所述第六异源部分(H6)连接至FVIII的所述C末端或N末端或插入FVIII的两个氨基酸之间。

[0187] 实施方式68:根据实施方式1至67中任一项所述的嵌合蛋白,其中所述第三异源部分(H3)、所述第四异源部分(H4)、所述第五异源部分(H5)或所述第六异源部分(H6)中的一者或多者包括包含至少约50个氨基酸、至少约100个氨基酸、至少约150个氨基酸、至少约200个氨基酸、至少约250个氨基酸、至少约300个氨基酸、至少约350个氨基酸、至少约400个氨基酸、至少约450个氨基酸、至少约500个氨基酸、至少约550个氨基酸、至少约600个氨基酸、至少约650个氨基酸、至少约700个氨基酸、至少约750个氨基酸、至少约800个氨基酸、至少约850个氨基酸、至少约900个氨基酸、至少约950个氨基酸或至少约1000个氨基酸的氨基酸序列。

[0188] 实施方式69:根据实施方式1至68中任一项所述的嵌合蛋白,其中所述FVIII的所述半衰期比野生型FVIII延长至少约1.5倍、至少约2倍、至少约2.5倍、至少约3倍、至少约4倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约11倍或至少约12倍。

[0189] 实施方式70:根据实施方式1至69中任一项所述的嵌合蛋白,其中所述FVIII蛋白的所述半衰期为至少约10小时、至少约11小时、至少约12小时、至少约13小时、至少约14小时、至少约15小时、至少约16小时、至少约17小时、至少约18小时、至少约19小时、至少约20小时、至少约21小时、至少约22小时、至少约23小时、至少约24小时、至少约36小时、至少约48小时、至少约60小时、至少约72小时、至少约84小时、至少约96小时或至少约108小时。

[0190] 实施方式71:根据实施方式22至70中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述第二异源部分之间的所述连接基或所述VWF片段和所述第一异源部分之间的所述连接基还包含位于所述连接基的所述N-末端区的第一切割位点(P1)、位于所述连接基的所述C-末端区的第二切割位点(P2)或它们二者。

[0191] 实施方式72:根据实施方式22至71中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述第二异源部分之间的所述连接基、所述VWF片段和所述第一异源部分之间的所述连接基或它们二者包含TLDPRSFLLRNPNDKYEPFWEDEEK(SEQ ID NO:56)。

[0192] 实施方式73:根据实施方式22至71中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述第二异源部分之间的所述连接基、所述VWF片段和所述第一异源部分之间的所述连接基或它们二者被选自由以下组成的组的蛋白酶切割:因子XIa、因子XIIa、激肽释放酶、因子VIIa、因子IXa、因子Xa、因子IIa(凝血酶)、弹性蛋白酶-2、粒酶-B、TEV、肠激酶、蛋白酶3C、分选酶A、MMP-12、MMP-13、MMP-17和MMP-20。

[0193] 实施方式74:根据实施方式22至73中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述第二异源部分之间的所述连接基、所述VWF片段和所述第一异源部分之间的所述连接基或它们二者包含选自由以下组成的组的氨基酸序列:RRRR(SEQ ID NO:52)、RKRRKR(SEQ ID NO:53)、RRRRS(SEQ ID NO:54)、TQSFNDFTR(SEQ ID NO:47)、SVSQTSKLTR(SEQ ID NO:48)、DFLAEGGGVR(SEQ ID NO:49)、TTKIKPR(SEQ ID NO:50)、LVPRG(SEQ ID NO:55)、ALRPRVVGGA(SEQ ID NO:51)、KLTRAET(SEQ ID NO:29)、DFTRVVG(SEQ ID NO:30)、TMTRIVGG(SEQ ID NO:31)、SPFRSTGG(SEQ ID NO:32)、LQVRIVGG(SEQ ID NO:33)、PLGRIVGG(SEQ ID

NO:34)、IEGRTVGG(SEQ ID NO:35)、LTPRSLLV(SEQ ID NO:36)、LGPVSGVP(SEQ ID NO:37)、VAGDSLEE(SEQ ID NO:38)、GPAGLGGA(SEQ ID NO:39)、GPAGLRGA(SEQ ID NO:40)、APLGLRLR(SEQ ID NO:41)、PALPLVAQ(SEQ ID NO:42)、ENLYFQG(SEQ ID NO:43)、DDDKIVGG(SEQ ID NO:44)、LEVLFQGP(SEQ ID NO:45)和LPKTGSES(SEQ ID NO:46)。

[0194] 实施方式75:根据实施方式71至74中任一项所述的嵌合蛋白,其中所述第一酶切割位点和所述第二酶切割位点是相同的或不同的。

[0195] 实施方式76:根据实施方式5至75中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述辅助部分之间的所述连接基、所述FVIII蛋白和所述第二异源部分之间的所述连接基以及所述VWF片段和所述第一异源部分之间的所述连接基中的一者或多者具有约1至约2000个氨基酸的长度。

[0196] 实施方式77:根据实施方式5至75中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述辅助部分之间的所述连接基、所述FVIII蛋白和所述第二异源部分之间的所述连接基以及所述VWF片段和所述第一异源部分之间的所述连接基中的一者或多者具有至少约10个、20个、30个、40个、50个、60个、70个、80个、90个、100个、110个、120个、130个、140个、150个、160个、170个、180个、190个、200个、210个、220个、230个、240个、250个、300个、350个、400个、450个、500个、550个、600个、650个、700个、750个、800个、850个、900个、950个、1000个、1200个、1400个、1600个、1800个或2000个氨基酸的长度。

[0197] 实施方式78:根据实施方式5至77中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述辅助部分之间的所述连接基、所述FVIII蛋白和所述第二异源部分之间的所述连接基以及所述VWF片段和所述第一异源部分之间的所述连接基中的一者或多者包含gly/ser肽。

[0198] 实施方式79:根据实施方式78所述的嵌合蛋白,其中所述gly/ser肽具有式(Gly4Ser)n或S(Gly4Ser)n,其中n为选自由以下组成的组的正整数:1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80或100。

[0199] 实施方式80:根据实施方式79所述的嵌合蛋白,其中所述(Gly4Ser)n连接基是(Gly4Ser)3或(Gly4Ser)4。

[0200] 实施方式81:根据实施方式5至80中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述辅助部分之间的所述连接基是可切割的连接基。

[0201] 实施方式82:根据实施方式81所述的嵌合蛋白,其中所述可切割的连接基包含一个或多个凝血酶切割位点。

[0202] 实施方式83:根据实施方式81或实施方式82所述的嵌合蛋白,其中所述可切割的连接基包含TLDPRSFLLRNPNDKYEPFWEDEEK(SEQ ID NO:56)。

[0203] 实施方式84:根据实施方式81至83中任一项所述的嵌合蛋白,其中可切割的连接基被选自由以下组成的组的蛋白酶切割:因子XIa、因子XIIa、激肽释放酶、因子VIIa、因子IXa、因子Xa、因子IIa(凝血酶)、弹性蛋白酶-2、粒酶-B、TEV、肠激酶、蛋白酶3C、分选酶A、MMP-12、MMP-13、MMP-17和MMP-20。

[0204] 实施方式85:根据实施方式81至83中任一项所述的嵌合蛋白,其中所述可切割的连接基包含选自由以下组成的组的氨基酸序列:RRRR(SEQ ID NO:52)、RKRRKR(SEQ ID NO:53)、RRRRS(SEQ ID NO:54)、TQSFNDFTR(SEQ ID NO:47)、SVSQTSKLTR(SEQ ID NO:48)、

DFLAEGGGVR(SEQ ID NO:49)、TTKIKPR(SEQ ID NO:50)、LVPRG(SEQ ID NO:55)、ALRPRVVGGA(SEQ ID NO:51)、KLTRAET(SEQ ID NO:29)、DFTRVVG(SEQ ID NO:30)、TMTRIVGG(SEQ ID NO:31)、SPFRSTGG(SEQ ID NO:32)、LQVRIVGG(SEQ ID NO:33)、PLGRIVGG(SEQ ID NO:34)、IEGRTVGG(SEQ ID NO:35)、LTPRSLLV(SEQ ID NO:36)、LGPVSGVP(SEQ ID NO:37)、VAGDSLEE(SEQ ID NO:38)、GPAGLGGA(SEQ ID NO:39)、GPAGLRGA(SEQ ID NO:40)、APLGLRLR(SEQ ID NO:41)、PALPLVAQ(SEQ ID NO:42)、ENLYFQG(SEQ ID NO:43)、DDDKIVGG(SEQ ID NO:44)、LEVLFQGP(SEQ ID NO:45)和LPKTGSES(SEQ ID NO:46)。

[0205] 实施方式86:根据实施方式5至85中任一项所述的嵌合蛋白,其中所述FVIII蛋白和所述辅助部分之间的所述连接基还包含分选酶识别基序。

[0206] 实施方式87:根据实施方式86所述的嵌合蛋白,其中所述分选酶识别基序包含所述序列LPXTG(SEQ ID NO:106)。

[0207] 实施方式88:根据实施方式19至87中任一项所述的嵌合蛋白,其中所述VWF片段包含VWF的所述D'结构域和所述D3结构域。

[0208] 实施方式89:根据实施方式88所述的嵌合蛋白,其中所述VWF片段抑制或阻止内源性VWF与FVIII蛋白的结合。

[0209] 实施方式90:根据实施方式88或实施方式89所述的嵌合蛋白,其中所述VWF片段的所述D'结构域的所述氨基酸序列与SEQ ID NO:2的第764至866位氨基酸具有至少90%、95%、96%、97%、98%、99%或100%的同一性。

[0210] 实施方式91:根据实施方式88至90中任一项所述的嵌合蛋白,其中所述VWF片段的所述D3结构域的所述氨基酸序列与SEQ ID NO:2的第867至1240位氨基酸具有至少90%、95%、96%、97%、98%、99%或100%的同一性。

[0211] 实施方式92:根据实施方式88至91中任一项所述的嵌合蛋白,其中所述VWF片段包含在对应于SEQ ID NO:2的第1099位残基、第1142位残基或第1099和1142位残基二者的残基处的至少一个氨基酸取代。

[0212] 实施方式93:根据实施方式88至91中任一项所述的嵌合蛋,其中在所述VWF片段的所述序列中,除半胱氨酸之外的氨基酸被对应于SEQ ID NO:2的第1099位残基、第1142位残基或第1099和1142位残基二者的残基取代。

[0213] 实施方式94:根据实施方式88至93中任一项所述的嵌合蛋白,其中所述VWF片段的所述序列包含SEQ ID NO:2的第764至1240位氨基酸。

[0214] 实施方式95:根据实施方式88至94中任一项所述的嵌合蛋白,其中所述VWF片段还包含VWF的所述D1结构域、所述D2结构域或所述D1和D2结构域。

[0215] 实施方式96:根据实施方式88至95中任一项所述的嵌合蛋白,其中所述VWF片段还包含选自由以下组成的组的VWF结构域:所述A1结构域、所述A2结构域、所述A3结构域、所述D4结构域、所述B1结构域、所述B2结构域、所述B3结构域、所述C1结构域、所述C2结构域、所述CK结构域、它们的一个或多个片段以及它们的任何组合。

[0216] 实施方式97:根据实施方式88至95中任一项所述的嵌合蛋白,其中所述VWF片段基本上由或由以下组成:(1)VWF的所述D'和D3结构域或其片段;(2)VWF的所述D1、D'和D3结构域或其片段;(3)VWF的所述D2、D'和D3结构域或其片段;(4)VWF的所述D1、D2、D'和D3结构域或其片段;或(5)VWF的所述D1、D2、D'、D3和A1结构域或其片段。

[0217] 实施方式98:根据实施方式88至97中任一项所述的嵌合蛋白,还包含VWF的与其可操作地连接的信号肽。

[0218] 实施方式99:根据实施方式19至98中任一项所述的嵌合蛋白,其中所述VWF片段是聚乙二醇化的、糖基化的、羟乙基淀粉化的或聚唾液酸化的。

[0219] 实施方式100:根据实施方式1至99中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含一个或多个选自由以下组成的组的FVIII的结构域:所述A1结构域、所述A2结构域、所述B结构域、所述A3结构域、所述C1结构域、所述C2结构域、它们的一个或多个片段以及它们的任何组合。

[0220] 实施方式101:根据实施方式100所述的嵌合蛋白,其中所述FVIII蛋白包含所述A1结构域、所述A2结构域、所述A3结构域和所述C1结构域以及所述任选的C2结构域。

[0221] 实施方式102:根据实施方式100或101所述的嵌合蛋白,其中所述FVIII蛋白包含所述B结构域或其部分。

[0222] 实施方式103:根据实施方式100至102中任一项所述的嵌合蛋白,其包含与SEQ ID NO:16或SEQ ID NO:18具有至少60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%的同一性的氨基酸序列。

[0223] 实施方式104:根据实施方式100至103中任一项所述的嵌合蛋白,其中所述FVIII蛋白是SQ B结构域缺失的FVIII。

[0224] 实施方式105:根据实施方式100至104中任一项所述的嵌合蛋白,其中所述FVIII蛋白包括单链FVIII。

[0225] 实施方式106:根据实施方式105所述的嵌合蛋白,其中所述单链FVIII包含在对应于全长成熟因子VIII多肽(SEQ ID NO:16)的第1648位残基、第1645位残基或它们二者或SQ BDD因子VIII(SEQ ID NO:18)的第754位残基、第751位残基或它们二者的残基处的至少一个氨基酸取代。

[0226] 实施方式107:根据实施方式106所述的嵌合蛋白,其中所述氨基酸取代是除精氨酸之外的氨基酸。

[0227] 实施方式108:根据实施方式1至107中任一项所述的嵌合蛋白,其中所述FVIII蛋白包含第一链和第二链,所述第一链包括FVIII的重链,并且所述第二链包括因子VIII的轻链,其中所述重链和所述轻链通过金属键缔合。

[0228] 实施方式109:一种编码根据实施方式1至108中任一项所述的嵌合蛋白的多核苷酸。

[0229] 实施方式110:根据实施方式110所述的多核苷酸,其还包含编码PC5、PC7或弗林蛋白酶的另外多核苷酸序列。

[0230] 实施方式111:根据实施方式109或实施方式110所述的多核苷酸,其还包含编码VWF的D1结构域和D2结构域的另外多核苷酸序列。

[0231] 实施方式112:一种或多种载体,其包含根据实施方式109至111中任一项所述的多核苷酸和一个或多个可操作地连接至所述多核苷酸或所述多核苷酸组的启动子。

[0232] 实施方式113:根据实施方式112所述的一种或多种载体,其还包括另外的载体,所述另外的载体包含编码PC5、PC7或弗林蛋白酶的第二多核苷酸链。

[0233] 实施方式114:根据实施方式112或实施方式113所述的一种或多种载体,其还包括

另外的载体，所述另外的载体包含编码VWF的D1结构域和D2结构域的多核苷酸序列。

[0234] 实施方式115：一种宿主细胞，其包含根据实施方式109至111中任一项所述的多核苷酸或根据实施方式112至114中任一项所述的载体。

[0235] 实施方式116：根据实施方式115所述的宿主细胞，其包含编码PC5、PC7或弗林蛋白酶的另外载体。

[0236] 实施方式117：根据实施方式115或实施方式116所述的宿主细胞，其还包含另外的载体，所述另外的载体包含编码VWF的D1结构域和D2结构域的多核苷酸序列。

[0237] 实施方式118：根据实施方式115或实施方式118所述的宿主细胞，其为哺乳动物细胞。

[0238] 实施方式119：根据实施方式118所述的宿主细胞，其中所述哺乳动物细胞选自由以下组成的组：HEK 293细胞、CHO细胞和BHK细胞。

[0239] 实施方式120：一种药物组合物，其包含根据实施方式1至108中任一项所述的嵌合蛋白、根据实施方式109或实施方式111所述的多核苷酸、根据实施方式112或实施方式114所述的载体或根据实施方式115和实施方式118中任一项所述的宿主细胞以及可药用载体。

[0240] 实施方式121：根据实施方式120中任一项所述的组合物，其中与无所述VWF片段的所述嵌合蛋白的所述FVIII蛋白的所述半衰期相比，所述嵌合蛋白的所述FVIII蛋白的所述半衰期在FVIII/VWF双敲除(“DKO”)小鼠中延长。

[0241] 实施方式122：根据实施方式120或实施方式121所述的组合物，其中所述FVIII的所述半衰期比野生型FVIII延长至少约1.5倍、至少约2倍、至少约2.5倍、至少约3倍、至少约4倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约11倍、至少约12倍、至少20倍、至少30倍或至少40倍。

[0242] 实施方式123：根据实施方式120或实施方式122所述的组合物，其中因子VIII的所述半衰期为至少6小时、至少7小时、至少9小时、至少10小时、至少11小时、至少12小时、至少15小时、至少约17小时、至少约18小时、至少约19小时、至少约20小时、至少约21小时、至少约22小时、至少约23小时、至少约24小时、至少约25小时、至少约26小时、至少约27小时、至少约28小时、至少约29小时、至少约30小时、至少约31小时、至少约32小时、至少约33小时、至少约34小时、至少约35小时、至少约36小时、至少约48小时、至少约60小时、至少约72小时、至少约84小时、至少约96小时或至少约108小时。

[0243] 实施方式124：根据实施方式120至123中任一项所述的组合物，其通过选自由以下组成的组的途径施用：局部施用、眼内施用、肠胃外施用、鞘内施用、硬膜下施用和口服。

[0244] 实施方式125：根据实施方式124所述的组合物，其中所述肠胃外施用为静脉内或皮下施用。

[0245] 实施方式126：根据实施方式120至125中任一项所述的组合物，其用于治疗需要其的受试者的出血性疾病或病状。

[0246] 实施方式127：根据实施方式126所述的组合物，其中所述出血性疾病或病状选自由以下组成的组：出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血、髂腰肌鞘出血以及它们的任何组合。

[0247] 实施方式128：根据实施方式126或实施方式127所述的组合物，其中所述受试者被

安排经历外科手术。

[0248] 实施方式129:根据实施方式126或实施方式127中任一项所述的组合物,其中所述治疗是预防性的或发作期的。

[0249] 实施方式130:一种阻止或抑制FVIII蛋白与内源性VWF的相互作用的方法,其包括向需要其的受试者加入有效量的根据实施方式1至108中任一项所述的嵌合蛋白、根据实施方式109至111中任一项所述的多核苷酸、根据实施方式112至114中任一项所述的载体、根据实施方式115至119中任一项所述的宿主细胞或根据实施方式120至129中任一项所述的组合物,其中所述VWF片段抑制或阻止所述FVIII蛋白与内源性VWF的相互作用。

[0250] 实施方式131:一种移除或减少FVIII蛋白的半衰期限制因子的方法,其中所述方法包括加入有效量的根据实施方式1至108中任一项所述的嵌合蛋白、根据实施方式109至111中任一项所述的多核苷酸、根据实施方式112至114中任一项所述的载体、根据实施方式115至119中任一项所述的宿主细胞或根据实施方式120至129中任一项所述的组合物,其中所述嵌合蛋白或由所述多核苷酸、所述载体编码的或由所述宿主细胞表达的所述嵌合蛋白阻止或抑制所述FVIII蛋白与内源性VWF的相互作用。

[0251] 实施方式132:一种延长或增加FVIII蛋白的所述半衰期的方法,其中所述方法包括加入有效量的根据实施方式1至108中任一项所述的嵌合蛋白、根据实施方式109至111中任一项所述的多核苷酸、根据实施方式112至114中任一项所述的载体、根据实施方式115至119中任一项所述的宿主细胞或根据实施方式120至129中任一项所述的组合物,其中所述嵌合蛋白的所述VWF片段阻止或抑制所述FVIII蛋白与内源性VWF的相互作用。

[0252] 实施方式133:根据实施方式132所述的方法,其中所述FVIII蛋白的所述半衰期比野生型FVIII延长至少约1.5倍、至少约2倍、至少约2.5倍、至少约3倍、至少约4倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约11倍或至少约12倍。

[0253] 实施方式134:根据实施方式133所述的方法,其中因子VIII的所述半衰期为至少约17小时、至少约18小时、至少约19小时、至少约20小时、至少约21小时、至少约22小时、至少约23小时、至少约24小时、至少约26小时、至少约27小时、至少约28小时、至少约29小时、至少约30小时、至少约31小时、至少约32小时、至少约33小时、至少约34小时、至少约35小时、至少约36小时、至少约48小时、至少约60小时、至少约72小时、至少约84小时、至少约96小时或至少约108小时。

[0254] 实施方式135:一种治疗需要其的受试者的出血性疾病或病状的方法,其包括施用有效量的根据实施方式1至108中任一项所述的嵌合蛋白、根据实施方式109至111中任一项所述的多核苷酸、根据实施方式112至114中任一项所述的载体、根据实施方式115至119中任一项所述的宿主细胞或根据实施方式120至129中任一项所述的组合物,其中所述出血性疾病或病症选自由以下组成的组:出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血、髂腰肌鞘出血以及它们的任何组合。

[0255] 实施方式136:根据实施方式135所述的方法,其中所述治疗是预防性的或按需的(发作期)。

[0256] 实施方式137:根据实施方式130至136中任一项所述的方法,其中所述有效量为

0.1μg/kg至500mg/kg。

[0257] 实施方式138:根据实施方式130至137中任一项所述的方法,其中所述嵌合蛋白、所述多核苷酸、所述宿主细胞或所述组合物通过选自由以下组成的组的途径施用:局部施用、眼内施用、肠胃外施用、鞘内施用、硬膜下施用和口服。

[0258] 实施方式139:根据实施方式138所述的方法,其中所述肠胃外施用选自由以下组成的组:静脉内施用、皮下施用、肌肉内施用和真皮内施用。

[0259] 实施方式140:根据实施方式130至139中任一项所述的方法,其中所述受试者是人。

[0260] 实施方式141:根据实施方式140所述的方法,其中所述受试者患有甲型血友病。

[0261] 实施方式142:一种制备嵌合蛋白的方法,其包括用根据实施方式109至111中任一项所述的多核苷酸、或根据实施方式112至114中任一项所述的载体转染一种或多种宿主细胞,以及在所述宿主细胞中表达所述VWF片段或所述嵌合蛋白。

[0262] 实施方式143:根据实施方式142所述的方法,其中所述载体还包含编码加工酶的多核苷酸。

[0263] 实施方式144:根据实施方式143所述的方法,其中所述加工酶是PACE。

[0264] 实施方式145:根据实施方式144所述的方法,其中PACE切割所述VWF片段的所述D1D2结构域。

[0265] 实施方式146:根据实施方式142和实施方式143所述的方法,其中还包括用表达VWF的D1结构域和D2结构域的多核苷酸序列转染一种或多种宿主细胞。

[0266] 实施方式147:一种构建根据实施方式1至108中任一项所述的嵌合蛋白的方法,其包括在分选酶存在下通过共价键将所述辅助部分与所述FVIII蛋白连接。

[0267] 附图简要说明

[0268] 图1A-F.VWF蛋白的示意图。图1A示出了包含SEQ ID NO:73的第1至276位氨基酸(SEQ ID NO:2的第764至1039位氨基酸)的两个VWF片段。VWF-001在不存在VWF的前/原肽序列的情况下合成,而VWF-009在存在前/原肽序列(D1和D2结构域)的情况下合成。VWF-009的前肽在合成期间被切割,并且VWF-009包含具有D'和D3结构域序列的原肽。图1B示出了包含SEQ ID NO:73的第1至477位氨基酸(SEQ ID NO:2的第764至1240位氨基酸)的三个VWF片段。VWF-002在不存在前/原肽序列的情况下合成。除D'D3结构域之外,VWF-010还包含D1D2结构域。除在SEQ ID NO:72的第336和379位残基处取代半胱氨酸的丙氨酸残基之外,VWF-013还包含D1D2D'D3结构域。图1C示出了包含D'D3结构域和A1结构域的一部分的两个VWF片段。VWF-003具有SEQ ID NO:2的第764至1274位氨基酸。除D'D3结构域之外,VWF-011还包含D1D2结构域。图1D示出了两个构建体VWF-004和VWF-012。VWF-004包含D'D3结构域和A1结构域的完整序列。VWF-012包含D1D2D'D3结构域和A1结构域的完整序列。图1E示出了三个构建体。VWF-006包含D1D2D'D3结构域和VWF的CK结构域(半胱氨酸结结构域)。VWF-008是全长VWF。VWF-031(VWF-Fc)示出了包含通过可切割的连接基连接至单Fc区的D1D2D'D3结构域的构建体。VWF-053是D1D2结构域。图1F示出了包含原肽(D1和D2结构域)和成熟亚基(D'、D3、A1、A2、A3、D4、B1-3、C1-2结构域)的全长VWF蛋白。VWF蛋白为约250kDa蛋白质,并且通过二硫键形成多聚体(>20MDa)。在非共价复合物中VWF蛋白与FVIII(95-98%)缔合,然后通过保护FVIII不被蛋白酶切割/活化,稳定重链和轻链以及阻止FVIII被清除剂受体清除来延长

FVIII的半衰期。VWF蛋白还可通过VWF受体清除FVIII-VWF复合物以及阻止胞饮和rFVIIIFc循环来限制FVIII的半衰期。

[0269] 图2.VWF:FVIII异源二聚体构建体的实例的示意图。左边构建体示出了具有全长VWF的D'D3结构域(SEQ ID NO:73的第1-477位氨基酸)并且包含SEQ ID NO:72的第336和379位残基处的丙氨酸取代的VWF片段。嵌合蛋白构建体(FVIII 064/065)包含通过连接基连接至第一Fc区的VWF片段的C-末端,并且FVIII连接至第二Fc区,其中第二Fc区还通过连接基连接至VWF片段的N-末端(如,式C-H1-L1-V-L2-H2,其中V为VWF片段,C为FVIII,H1和H2为Fc区,并且L1和L2为可切割的连接基)。图2b中的构建体是在细胞内加工的VWF:FVIII异源二聚体构建体,其中第二Fc和VWF片段的N-末端之间的连接基被切割。FVIII-064包含VWF的D'D3结构域(具有C336A和C379取代的SEQ ID NO:73的第1至477位氨基酸)。FVIII-065包含VWF的D'D3结构域(SEQ ID NO:73的第1至276位氨基酸)。FVIII-136包含通过可被胞内蛋白酶加工的连接基连接至D'D3片段-Fc的FVIIIFc。当FVIII-136表达时,酶切割第二Fc(融合至FVIII-LC)和VWF D'D3片段(融合至第一Fc)之间的连接基,而融合至(或连接至)FVIII-LC的Fc区与融合至(或连接至)VWF片段的第一Fc形成共价键(如,二硫键)。FVIII-148是具有D'D3片段的单链FVIIIFc(通过将R1645A/R1648A突变引入FVIII基因而形成的单链FVIII)。

[0270] 图3.包含VWF和Fc之间的可变连接基实例的VWF:FVIII异源二聚体构建体实例的示意图。构建体(FVIII-064、FVIII-159、FVIII-160、FVIII-178和FVIII-179)具有如式C-H1-L1-V-L2-H2表示的通用结构,但包含不同连接基或氨基酸取代的实例。所示构建体包含相同的VWF片段,该片段为VWF的D'和D3结构域(即,具有氨基酸取代C336A和C379A的SEQ ID NO:73的第1至477位氨基酸)。构建体FVIII 64具有VWF片段和Fc(即,H2)之间的凝血酶可切割连接基(即,L2),该连接基具有20个氨基酸。构建体FVIII 159具有VWF片段和Fc(即,H2)之间的凝血酶可切割连接基(即,L2),该连接基具有35个氨基酸。构建体FVIII 160具有VWF片段和Fc(即,H2)之间的凝血酶可切割连接基(即,L2),该连接基具有48个氨基酸。构建体FVIII-180、FVIII-181和FVIII-182分别为包含FVIII C1结构域中的K2092A突变、FVIII C1结构域中的K2093A突变以及FVIII C1结构域中的K2092A/K2093A突变的FVIII-160的衍生物。构建体FVIII 178具有VWF片段和Fc(即,H2)之间的凝血酶可切割连接基(即,L2),该连接基具有73个氨基酸。构建体FVIII 179具有VWF片段和Fc(即,H2)之间的凝血酶可切割连接基(即,L2),该连接基具有98个氨基酸。

[0271] 图4:FVIII-VWF构建体的实例的示意图,其中VWF是VWF的D1D2D'D3片段,连接基是包含切割位点如凝血酶切割位点的可变长度连接基,SC FVIII是包含R1645A/R1648A取代的单链FVIII,H是异源部分,如免疫球蛋白恒定区或其部分、缀合聚乙二醇(PEG)和/或PEG的部分、白蛋白或白蛋白片段、白蛋白结合部分、HAP序列、聚唾液酸化和/或聚唾液酸部分、羟乙基淀粉(HES)和/或HES部分或PAS序列等,HC FVIII是FVIII的重链,LC FVIII是FVIII的轻链,并且Fc是免疫球蛋白恒定区的Fc区。图4A具有式VWF-连接基-SC FVIII。图4B具有式VWF-连接基-H-连接基-SC FVIII。连接基(VWF和H之间的第一连接基和H和SC FVIII之间的第二连接基)可以是相同的或不同的。图4C具有式VWF-连接基-SC FVIII-连接基-H。连接基(VWF和SC FVIII之间的第一连接基和SC FVIII和H之间的第二连接基)可以是相同的或不同的。图4D具有式VWF-连接基-HC FVIII-H-连接基-LC FVIII。连接基(VWF和HC FVIII之

间的第一连接基和H和LC FVIII之间的第二连接基)可以是相同的或不同的。图4E具有式HC FVIII-H-LC FVIII-连接基-第一Fc-连接基-VWF-连接基-第二Fc。连接基(LC FVIII和第一Fc之间的第一连接基、第一Fc和VWF之间的第二连接基和VWF和第二Fc之间的第三连接基)可以是相同的或不同的。连接基可以是可切割的连接基。例如,第一Fc和VWF之间的连接基可以是包含在连接基的N-末端和/或C-末端的切割位点的可切割的连接基。第一Fc和第二Fc可以是相同的或不同的。图4F具有式HC FVIII-H-LC FVIII-连接基-第一Fc-连接基-VWF-连接基-第二Fc。连接基(LC FVIII和第一Fc之间的第一连接基、第一Fc和VWF之间的第二连接基和VWF和第二Fc之间的第三连接基)可以是相同的或不同的。一个或多个连接基可以是可切割的连接基。例如,第一Fc和VWF之间的连接基可以是包含在连接基的N-末端和/或C-末端的切割位点的可切割的连接基。第一Fc和第二Fc可以是相同的或不同的。图4G具有式SC FVIII-连接基-Fc-连接基-VWF-H-连接基-Fc。图4H具有式聚乙二醇化或羟乙基淀粉化SC FVIII-连接基-Fc-连接基-VWF-H-连接基-Fc。连接基(SC FVIII和第一Fc之间的第一连接基、第一Fc和VWF之间的第二连接基以及H和第二Fc之间的第三连接基)可以是相同的或不同的。一个或多个连接基可以是可切割的连接基。例如,第一Fc和VWF之间的连接基可以是包含在连接基的N-末端和/或C-末端的切割位点的可切割的连接基。第一Fc和第二Fc可以是相同的或不同的。

[0272] 图5.FVIII-VWF异源二聚体共转染系统的示意图。构建体FVIII-155包含连接至Fc区的全长FVIII序列(丙氨酸残基取代第1645和1648位的精氨酸残基)。VWF-031包含D1D2D'D3片段(丙氨酸残基取代第336和379位的半胱氨酸残基),其连接至具有48凝血酶可切割连接基的另一个Fc区。在细胞内加工后,构建体FVIII-155生成融合至一个Fc片段的全长单链FVIII(SCFVIII),构建体VWF-031生成连接至另一个Fc片段的477个氨基酸的D'D3片段。两个共价键可在连接至SC FVIII或D'D3片段的Fc片段之间形成,这继而允许FVIII和D'D3的共价缔合,这是所需最终产物的主要特征。

[0273] 图6是VWF-009(D1D2D'D3 1-276aa×6HIS)的非还原和还原SDS PAGE,其显示出VWF-009以单体存在。未加工意指VVF-009具有原肽(D1D2结构域)。

[0274] 图7是VWF-002(D'D3 1-477aa×6his)或VWF-010(D1D2D'D31-477aa×6his)的非还原和还原SDS PAGE,其显示出VWF-002以单体存在,且VWF-010以二聚体存在。

[0275] 图8示出了图2(b)中示出的FVIII-VWF异源二聚体的凝血酶消化。第1泳道示出分子量标准。第2泳道是无凝血酶的rFVIII-Fc。第3泳道是有凝血酶的rFVIII-Fc。第5泳道是FVIIIFc-VWF。第6泳道示出了FVIIIFc-VWF和凝血酶。A1表示FVIII的A1结构域,A2表示FVIII的A2结构域,并且Δa3LC表示FVIII的轻链。

[0276] 图9A-B示出了通过FVIII显色测定法测量的FVIII活性。图9A示出了HemA小鼠中rFVIII和rFVIIIFc的药代动力学曲线。图9B示出了FVIII/VWF双敲除(DKO)小鼠中rFVIII和rFVIIIFc的PK曲线。Y轴示出FVIII活性,单位为mIU/mL,X轴示出时间。

[0277] 图10A-B示出了D'D3片段对FVIII的保护,其通过质粒注射48小时后mFVIII血浆水平(mIU/mL)和VWF表达水平(nM/mL)示出。用于示出FVIII保护的VWF片段为VWF-001(276aa,单体)、VWF-009(276aa,单体)、VWF-002(477aa,单体)、VWF-010(477aa,二聚体)、VWF-003(511aa,单体)、VWF-011(511aa,二聚体)、VWF-004(716aa,单体)、VWF-012(716aa,二聚体)、VWF-006和VWF-008。

[0278] 图11示出了共施用D'D3片段时FVIII-VWF DKO小鼠中rBDD-FVIII的药代动力学曲线。图11A示出了rBDD-FVIII和VWF-002或rBDD-FVIII和VWF-010共施用或rBDD-FVIII单独施用后FVIII/VWF DKO小鼠中通过FVIII显色测定测量的FVIII活性(mIU/mL)。图11B示出了施用后VWF-002和VWF-010的血浆水平(ng/mL)。X轴表示时间,单位为小时。

[0279] 图12示出了表达VWF D'D3的小鼠中rFVIIIFc的药代动力学曲线。图12A示出了编码D'D3结构域质粒DNA的高压注射(hydrodynamic injection,HDI)(第-5天)、rFVIIIFc的静脉内给药(第0天)和PK样品收集(第0天-第3天)的时间轴。图12B示出了通过FVIII显色测定测量的D1D2D'D3结构域(477aa)HDI(圆形)的FVIII/VWF DKO小鼠中和具有半胱氨酸取代的D1D2D'D3结构域(477aa)(矩形)的FVIII/VWF DKO小鼠中rFVIIIFc注入后血浆FVIII活性(mIU/mL)。无D'D3结构域HDI的对照小鼠中FVIII活性以三角形示出。图10C示出了D1D2D'D3二聚体或D1D2D'D3单体DNA构建体的HDI施用后D'D3血浆水平(ng/mL)。X轴表示时间,单位为小时。

[0280] 图13示出了FVIII/VWF DKO小鼠中通过HDI的D'D3-Fc连接基选择。将不同长度连接基(20aa(FVIII-064)、35aa(FVIII-159)或48aa(FVIII-160))插入D'D3结构域和Fc区之间。FVIII/VWF DKO小鼠中HDI后的FVIII活性(mIU/ml)通过FVIII显色测定测量。

[0281] 图14示出了FVIII/VWF DKO小鼠中单链FVIIIFc/D'D3异源二聚体的HDI。加工(双链)rFVIIIFc-D'D3(pSYN-FVIII-136)和单链rFVIIIFc-D'D3(pSYN-FVIII-148)的FVIII活性在HDI 24小时和48小时后测量。

[0282] 图15示出了通过Octet测定测量的FVIII-155/VWF-031异源二聚体与固定化hVWF的结合亲合力。FVIIIFc、FVIII和IgG也用作对照。x-轴显示时间,单位为秒,并且y-轴显示结合,单位为纳米(nm)。

[0283] 图16示出了FVIII/VWF缺陷(FVIII/VWF DKO)小鼠中FVIII-155/VWF-031的药代动力学。x-轴表示时间,单位为小时,并且y-轴表示相对于输入的FVIII回收率,单位为%。

[0284] 图17:VWF片段构建体实例的示意图,其中VWF是VWF的D1D2D'D3片段;连接基是包含切割位点如凝血酶切割位点的可变长度连接基;H是异源部分,如免疫球蛋白恒定区或其部分、缀合聚乙二醇(PEG)的部分和/或PEG、白蛋白或白蛋白片段、白蛋白结合部分、HAP序列、聚唾液酸化部分和/或聚唾液酸、羟乙基淀粉(HES)部分和/或HES或PAS序列等;并且Fc是免疫球蛋白的Fc区。图17A具有式D1D2-D'部分D3-H-部分D3-连接基-Fc。图17B具有式D1D2-部分D'-H-部分D'D3-连接基-Fc。图17C具有式D1D2-聚乙二醇化或羟乙基淀粉化D'D3-连接基-Fc。连接基可任选地被切割。

[0285] 图18:A)示出了HemA(菱形)和DKO(正方形)血浆二者中FVIIIFc随时间推移丧失FVIII活性。FVIII活性通过显色测定测量。X-轴显示时间,单位为小时,y-轴显示相对活性。B)示出了FVIII活性由于重链(HC)的解离或降解而丧失。左图示出了使用绵羊抗-FVIII多克隆抗体在Bio-rad 4-15%凝胶中进行的免疫沉淀测定。凝胶为还原的,并且通过Bio-rad系统成像。第1泳道示出Bio-rad未染色的分子量标准;第2泳道示出FVIIIFc和PBS;第3泳道示出FVIIIFc和DKO血浆;第5泳道示出单独的绵羊抗-FVIII多克隆抗体。右图示出了使用FVIII抗重链抗体(GMA012)进行的凝胶蛋白质印迹分析。第1泳道示出Bio-rad未染色的分子量标准;第2泳道示出FVIIIFc和PBS;第3泳道示出FVIIIFc和DKO血浆;第4泳道示出单独的绵羊抗-FVIII多克隆抗体。

[0286] 图19示出了通过显色测定测量的DKO小鼠血浆(左图)和HemA小鼠血浆(右图)中野生型FVIIIFc(圆形)、scFVIIIFc(单链FVIII)(实心三角形)或FVIII:VWF异源二聚体(如,FVIII155/VWF31)(空心三角形)随时间变化的FVIII活性。Y轴示出了相对FVIII活性。野生型FVIIIFc包含FVIII的双链(即,非共价结合的FVIII重链和FVIII轻链),从而具有三条链,即FVIII重链、融合至Fc的FVIII轻链以及单独Fc。ScFVIIIFc包含FVIII单链,并且从而具有两条链,一条具有融合至Fc的单链FVIII,另一条具有单独Fc。FVIII:VWF异源二聚体(如,FVIII155/VWF031)包含融合至Fc的单链FVIII和融合至Fc的VWF片段(D'D3)。

[0287] 图20示出了通过不同浓度的PC5或PACE(弗林蛋白酶)进行的由VWF片段(如,VWF-031(D1D2D'D3Fc))的D1D2结构域的加工。D1D2加工在还原条件的Bio-rad 4-15%凝胶上通过Bio-rad成像仪示出。第1泳道示出单独VWF031;第2泳道示出单独PC5;第3泳道示出单独PACE;第4泳道示出2.5%的VWF031和PC5;第5泳道示出5%的VWF031和PC5;第6泳道示出7.5%的VWF031和PC5;第7泳道示出10%的VWF031和PC5;第8泳道示出2.5%的VWF031和PACE;第9泳道示出5%的VWF031和PACE;第10泳道示出7.5%的VWF031;并且第11泳道示出10%的VWF031和PACE。

[0288] 图21:A)示出了通过ForteBio octet仪器进行的FVIII:VWF异源二聚体(如,FVIII-155/VWF-031)的结合测定。对于该测定,全长VWF使用APS传感器捕集。FVIIIFc和FVIII与全长VWF的结合在左下左图示出。未结合的FVIIIY1680(不具有对VWF的亲合力的突变体)和FVIII:VWF异源二聚体(FVIII155/VWF031)在右下图示出。B)示出了FVIII:VWF异源二聚体(如,FVIII-155/VWF-031)的另一个结合测定。在该测定中,构建体(VWF031构建体、FVIII-155/VWF031或FVIII)固定在蛋白G传感器上。测量构建体与FVIII的结合。

[0289] 图22示出了通过表面等离子共振实验测量的VWF D'D3结构域与FVIII分子的结合亲合力。VWF031构建体(100RU)通过1000RU抗人IgG捕集。B-结构域缺失的FVIII以1:1拟合以单周期动力学模式施加。总数为4。

[0290] 图23示出了当施用于FVIII/VWF DKO小鼠时,FVIIIFc/VWF异源二聚体构建体中不同连接基长度对药代动力学的影响。三个不同的连接基(48aa、73aa或98aa),即VWF031、VWF035和VWF036插入D'D3和Fc之间。归一化为5分钟值(%)的FVIII活性在Y-轴中示出。

[0291] 图24示出了VWF片段与FVIII的分选酶连接实例。A)示出了两个连接构建体,(1)C-末端融合至分选酶识别基序(如,LPXTG)的VWF片段以及(2)N-末端具有甘氨酸(n)的FVIII。在与分选酶反应后,VWF片段和分选酶识别基序连接至FVIII的N-末端。B)示出了两个连接构建体,(1)其C-末端融合至分选酶识别基序的FVIII以及(2)其N-末端具有甘氨酸(n)的VWF片段。在与分选酶反应后,FVIII和分选酶识别基序在VWF片段的N-末端融合至VWF片段。C)示出了两个连接构建体,(1)通过可变长度连接基融合至分选酶识别基序的VWF片段以及(2)其N-末端融合至甘氨酸(n)的FVIII。在与分选酶反应后,通过连接基融合至分选酶识别基序的VWF连接至FVIII的N-末端。D)示出了两个连接构建体,(1)通过可变长度连接基融合至分选酶识别基序的FVIII以及(2)其N-末端融合至甘氨酸(n)的VWF。在与分选酶反应后,通过连接基融合至分选酶识别基序的FVIII连接至VWF片段的N-末端。E)示出了包含通过可变长度连接基融合至分选酶识别基序的VWF片段的连接构建体,所述分选酶识别基序还融合至蛋白酶切割位点(如,凝血酶切割位点),该位点通过可变长度连接基融合至Fc。

[0292] 图25示出了FVIII155和FVIII198的比较示意图。FVIII155编码单链FVIIIFc蛋白。

FVIII198是包含部分B-结构域的单链FVIIIFc分子-226N6。226表示FVIII B-结构域的N-末端226个氨基酸，并且N6表示B-结构域中的六个N-糖基化位点。

[0293] 图26A)示出了测量DKO血浆中FVIII155和FVIII198随时间变化的相对活性的稳定性测定。如图所示，部分B-结构域存在于FVIII198中增加了单链FVIIIFc相对于FVIII155的稳定性;B)示出了DKO小鼠中FVIII198、FVIII155和双链(dcFVIIIFc)的半衰期比较。如图所示，与双链FVIII相比，单链FVIII(FVIII155)半衰期增加1.5倍。具有266N6B-结构域的单链FVIII(FVIII198)半衰期也增加1.5倍。曲线图示出了FVIII回收率对5分钟值(%)随时间的变化。

[0294] 发明详述

[0295] 定义

[0296] 值得注意的是，术语“一个”或“一种”实体是指该实体的一者或多者;例如，“核苷酸序列”应理解为表示一个或多个核苷酸序列。因此，术语“一个”(或“一种”)、“一个或多个”和“至少一个”在本文中可互换使用。

[0297] 术语“多核苷酸”或“核苷酸”旨在涵盖单个核酸以及多个核酸，并且指分离的核酸分子或构建体，如信使RNA(mRNA)或质粒DNA(pDNA)。在某些实施方案中，多核苷酸包含常规磷酸二酯键或非常规键(如，酰胺键，例如可见于肽核酸(PNA)中)。术语“核酸”是指任何一种或多种存在于多核苷酸中的核酸片段，如DNA或RNA片段。所谓“分离的”核酸或多核苷酸意指从其天然环境移除的核酸分子、DNA或RNA。例如，出于本发明的目的，包含于载体中的编码因子VIII多肽的重组多核苷酸被认为是分离的。分离的多核苷酸的另外实例包括保持在异源宿主细胞中或从溶液中的其它多核苷酸纯化(部分或基本上)的重组多核苷酸。分离的RNA分子包括本发明多核苷酸的体内或体外RNA转录物。根据本发明的分离的多核苷酸或核酸还包括此类合成产生的分子。此外，多核苷酸或核酸可包括调控元件，例如启动子、增强子、核糖体结合位点或转录终止信号。

[0298] 如本文所用，“编码区”或“编码序列”是由可翻译为氨基酸的密码子组成的多核苷酸的一部分。虽然“终止密码子”(TAG、TGA或TAA)通常不翻译为氨基酸，但其可视为编码区的部分，但任何侧翼序列，例如启动子、核糖体结合位点、转录终止子、内含子等不是编码区的部分。编码区的边界通常通过编码所得多肽的氨基末端的处于5’末端的起始密码子，和编码所得多肽的羧基末端的处于3’末端的翻译终止密码子确定。本发明的两个或更多个编码区可存在于单个多核苷酸构建体中，如单个载体上，或单独的多核苷酸构建体中，如单独的(不同的)载体上。然后，它遵循单个载体可以仅包含单个编码区，或包含两个或更多个编码区，如单个载体可独立地编码结合结构域-A和结合结构域-B，如下所述。此外，本发明的载体、多核苷酸或核酸可编码融合或未融合至编码本发明的结合结构域的核酸的异源编码区。异源编码区包括不限于特化元件或基序，例如分泌信号肽或异源功能结构域。

[0299] 哺乳动物细胞分泌的某些蛋白质与分泌信号肽缔合，一旦延长的蛋白质链开始跨过糙面内质网输出，所述信号肽即从成熟蛋白质切割。本领域的普通技术人员认识到，信号肽通常融合至多肽的N-末端，并且从完整的或“全长”多肽切割，以产生分泌或“成熟”形式的多肽。在某些实施方案中，使用天然信号肽，如免疫球蛋白重链或轻链信号肽，或保持指导与其可操作地缔合的多肽分泌的能力的该序列功能衍生物。或者，可使用异源哺乳动物信号肽，如人组织血纤维蛋白溶酶原激活物(TPA)或小鼠β-葡糖醛酸酶信号肽，或其功能衍

生物。

[0300] 术语“下游”是指位于参考核苷酸序列的3’端的核苷酸序列。在某些实施方案中，下游核苷酸序列涉及沿着转录的起始点的序列。例如，基因的翻译起始密码子位于转录的起始位点下游。

[0301] 术语“上游”是指位于参考核苷酸序列的5’端的核苷酸序列。在某些实施方案中，上游核苷酸序列涉及位于编码区的5’侧或转录的起始点的序列。例如，大多数启动子位于转录的起始位点上游。

[0302] 如本文所用，术语“调控区”是指位于编码区的上游(5’非编码序列)内，或下游(3’非编码序列)的核苷酸序列，并且其影响转录、RNA加工、稳定性或缔合编码区的翻译。调控区可包括启动子、翻译前导序列、内含子、多腺苷酸化识别序列、RNA加工位点、效应子结合位点和茎-环结构。如果预期编码区在真核细胞中表达，则多腺苷酸化信号和转录终止序列将通常位于编码序列的3’端。

[0303] 编码基因产物如多肽的多核苷酸可包括可操作地与一个或多个编码区缔合的启动子和/或其它转录或翻译控制元件。在可操作缔合中，基因产物如多肽的编码区与一个或多个调控区缔合，这样将基因产物的表达置于调控区的影响或控制下。例如，如果启动子功能的诱导引起编码基因产物的mRNA的转录，该基因产物由编码区编码，并且如果启动子和编码区之间的连接的性质不干扰启动子指导基因产物表达的能力，或干扰DNA模板被转录的能力，则编码区和启动子是“可操作地缔合的”。除启动子之外的其它转录控制元件，例如增强子、操纵子、阻遏子和转录终止信号也可操作地与编码区缔合，以指导基因产物表达。

[0304] 多个转录控制区是本领域技术人员已知的。这些包括但不限于在脊椎动物细胞发挥功能的转录控制区，例如但不限于来自巨细胞病毒(与内含子-A结合的立即早期启动子)、猿猴病毒40(早期启动子)和逆转录病毒(例如劳氏肉瘤病毒)的启动子和增强子片段。其它转录控制区包括来源于脊椎动物基因，例如肌动蛋白、热休克蛋白、牛生长激素和兔β-球蛋白的那些，以及能够控制真核细胞中基因表达的其它序列。另外的合适转录控制区包括组织特异性启动子和增强子以及淋巴因子诱导型启动子(如，干扰素或白介素诱导的启动子)。

[0305] 相似地，多种翻译控制元件是本领域普通技术人员已知的。这些包括但不限于核糖体结合位点、翻译起始和终止密码子，以及来源于小RNA病毒的元件(特别是内部核糖体进入位点或IRES，也称为CITE序列)。

[0306] 如本文所用，术语“表达”是指多核苷酸产生基因产物，例如RNA或多肽的过程。其包括但不限于多核苷酸转录为信使RNA(mRNA)、转运RNA(tRNA)、小发夹RNA(shRNA)、小干扰RNA(siRNA)或任何其它RNA产物，以及mRNA翻译为多肽。表达生成“基因产物”。如本文所用，基因产物可为核酸，如基因转录生成的信使RNA，或从转录物翻译的多肽。本文所述的基因产物还包括具有转录后修饰，如多腺苷酸化或剪接的核酸，或具有翻译后修饰，如甲基化、糖基化、添加脂质、与其它蛋白亚基缔合或蛋白酶切割的多肽。

[0307] “载体”是指用于将核酸克隆和/或转移至宿主细胞的任何媒介物。载体可以是另一个核酸片段可与其连接，以便形成连接片段的复制的复制子。“复制子”是指作为体内自主复制单元起作用，即能够在其自身控制下复制的任何遗传元件(如，质粒、噬菌体、粘粒、染色体、病毒)。术语“载体”包括用于在体外、离体或在体内将核酸引入细胞的病毒和非病

毒媒介物二者。大量载体是本领域已知和使用的，包括例如质粒、修饰的真核病毒或修饰的细菌病毒。多核苷酸插入合适的载体可通过将适当的多核苷酸片段连接至所选择的具有互补粘性末端的载体实现。

[0308] 载体可被工程化以编码选择性标记或报告基因，所述选择性标记或报告基因提供掺入载体的细胞的选择或识别。选择性标记或报告基因的表达允许识别和/或选择宿主细胞，该宿主细胞掺入并且表达包含于载体的其它编码区。本领域已知和使用的选择性标记基因的实例包括：提供对氨苄青霉素、链霉素、庆大霉素、卡那霉素、潮霉素、双丙氨膦除草剂、磺酰胺等抗性的基因；以及用作表型标记的基因，即花色素苷调控基因、异戊烯基转移酶基因等。本领域已知和使用的报告基因的实例包括：荧光素酶(Luc)、绿色荧光蛋白(GFP)、氯霉素乙酰转移酶(CAT)、-半乳糖苷酶(LacZ)、-葡糖醛酸酶(Gus)等。选择性标记也可视为报告基因。

[0309] 术语“质粒”是指通常携带不作为细胞的中央代谢部分的基因，并且通常为环状双链DNA分子形式的染色体外元件。此类元件可为自主复制序列、基因组整合序列、噬菌体或核苷酸序列、任何来源的线性、环状或超螺旋的单链或双链DNA或RNA，其中多个核苷酸序列连接或重组为能够将所选择基因产物的启动子片段和DNA序列以及适当的3’未翻译序列引入细胞的独特结构。

[0310] 可使用的真核病毒载体包括但不限于：腺病毒载体、逆转录病毒载体、腺相关病毒载体、痘病毒如牛痘病毒载体、杆状病毒载体或疱疹病毒载体。非病毒载体包括质粒、脂质体、带电荷脂质(细胞转染剂(cytofectin))、DNA-蛋白质复合物和生物聚合物。

[0311] “克隆载体”是指“复制子”，它是连续复制的单位长度核酸，并且包含复制起点，例如质粒、噬菌体或粘粒，另一个核酸片段可连接到其上，以便导致连接片段的复制。某些克隆载体能够在一种细胞类型，如细菌中复制，并且在另一种细胞类型，如真核细胞中表达。克隆载体通常包含可用于选择包含载体的细胞一个或多个序列和/或用于插入所关注的核酸序列的一个或多个多克隆位点。

[0312] 术语“表达载体”是指设计为使得插入的核酸序列在插入宿主细胞后表达的载体。插入的核酸序列设置为与调控区可操作缔合，如上所述。

[0313] 载体通过本领域已知的方法，如转染、电穿孔、显微注射、转导、细胞融合、DEAE葡聚糖、磷酸钙沉淀、脂质转染(溶酶体融合)、使用基因枪或DNA载体转运蛋白引入宿主细胞。

[0314] 如本文所用，“培养”意指在允许细胞生长或分裂或保持细胞活动状态的体外条件下温育细胞。如本文所用，“培养的细胞”意指体外繁殖的细胞。

[0315] 如本文所用，术语“多肽”旨在涵盖单个“多肽”以及多个“多肽”，并且是指由通过酰胺键(也称为肽键)线性连接的单体(氨基酸)构成的分子。术语“多肽”是指两个或更多个氨基酸的任何一条或多条链，并且不是指具体长度的产物。因此，肽、二肽、三肽、寡肽、“蛋白质”、“氨基酸链”或用于指两个或更多个氨基酸的一条或多条链的任何其它术语，包括在“多肽”的定义内，并且术语“多肽”可用于代替任何这些术语，或可与任何这些术语互换。术语“多肽”还意指多肽的表达后修饰产物，所述表达后修饰包括但不限于糖基化、乙酰化、磷酸化、酰胺化、通过已知保护/封端基团的衍生化、蛋白酶裂解或通过非天然存在的氨基酸的修饰。多肽可衍生自天然生物来源或重组制备技术，但不必从指定的核酸序列翻译。它可以任何方式生成，包括化学合成。

[0316] “分离的”多肽或片段、变体、或其衍生物是指不处于其天然环境的多肽。不需要具体水平的纯化。例如，分离的多肽可以仅仅从其天然或自然环境移除。出于本发明的目的，在宿主细胞中表达的重组产生的多肽和蛋白质被视为分离的，通过任何合适的技术分离、分级或部分或基本上纯化的天然或重组多肽也一样。

[0317] 本发明还包括多肽的片段或变体，以及它们的任何组合。当涉及本发明的多肽结合结构域或结合分子时，术语“片段”或“变体”包括保持参考多肽的至少一些特性(如，对FcRn结合结构域或Fc变体的FcRn结合亲合力、对FVIII变体的凝血活性或VWF片段的FVIII结合活性)的任何多肽。除本文别处讨论的具体抗体片段之外，多肽片段还包括蛋白酶水解片段，以及缺失片段，但不包括天然存在的全长多肽(或成熟多肽)。本发明的多肽结合结构域或结合分子的变体包括上述片段，以及具有由于氨基酸取代、缺失或插入而改变的氨基酸序列的多肽。变体可以是天然或非天然存在的。非天然存在的变体可使用本领域已知的诱变技术产生。变体多肽可包含保守或非保守氨基酸取代、缺失或添加。

[0318] 如本文所用，术语一个或多个“VWF片段”意指与FVIII相互作用并且保持全长VWF通常提供给FVIII的至少一种或多种特性的任何VWF片段，所述特性如防止FVIIIa的过早活化，防止过早蛋白水解，防止可导致过早清除的与磷脂膜的缔合，防止与可结合裸露FVIII而非VWF结合FVIII的FVIII清除受体的结合，和/或稳定FVIII重链和轻链相互作用。如本文所用，术语“VWF片段”不包括全长或成熟VWF蛋白。在具体实施方案中，如本文所用，“VWF片段”包含VWF蛋白的D’结构域和D3结构域，但不包括VWF蛋白的A1结构域、A2结构域、A3结构域、D4结构域、B1结构域、B2结构域、B3结构域、C1结构域、C2结构域和CK结构域。

[0319] 如本文所用，术语“半衰期限制因子”或“FVIII半衰期限制因子”表示防止FVIII蛋白的半衰期长于野生型FVIII(如，ADVATE®或REFACTO®)的1.5倍或2倍的因子。例如，全长或成熟VWF可通过诱导FVIII和VWF复合物通过一条或多条VWF清除通路从系统清除而充当FVIII半衰期限制因子。在一个实例中，内源性VWF是FVIII半衰期限制因子。在另一个实例中，非共价结合至FVIII蛋白的全长重组VWF分子是FVIII半衰期限制因子。

[0320] 如本文所用，术语“内源性VWF”表示天然存在于血浆中的VWF分子。内源性VWF分子可以是多聚体，但可以是单体或二聚体。血浆中的内源性VWF结合至FVIII，并且与FVIII形成非共价复合物。

[0321] “保守氨基酸取代”是其中氨基酸残基被具有类似侧链的氨基酸残基替代的取代。具有类似侧链的氨基酸残基家族在本领域中有所定义，包括碱性侧链(如，赖氨酸、精氨酸、组氨酸)、酸性侧链(如，天冬氨酸、谷氨酸)、不带电极性侧链(如，甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸)、非极性侧链(如，丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸)、β-支链侧链(如，苏氨酸、缬氨酸、异亮氨酸)和芳基侧链(如，酪氨酸、苯丙氨酸、色氨酸、组氨酸)。因此，如果多肽中的氨基酸被来自相同侧链家族的另一个氨基酸替代，则取代被视为保守的。在另一个实施方案中，氨基酸链可被侧链家族成员的顺序和/或组成不同、结构类似的链保守替代。

[0322] 如本领域已知，两个多肽之间的“序列同一性”通过比较一个多肽的氨基酸序列与第二个多肽的序列确定。在本文中讨论时，任何特定多肽是否与另一条多肽具有至少约50％、60％、70％、75％、80％、85％、90％、95％、99％或100％同一性可使用本领域已知的方法和计算机程序/软件确定，例如但不限于BESTFIT程序(Wisconsin Sequence Analysis

Package,Version 8 for Unix,Genetics Computer Group,University Research Park,575 Science Drive,Madison,WI 53711)。BESTFIT使用Smith and Waterman,Advances in Applied Mathematics 2:482-489(1981)的局部同源性算法查找两个序列之间的最佳同源性片段。当根据本发明使用BESTFIT或任何其它序列比对程序确定特定序列是否与参考序列具有例如95％同一性时,参数的设置毫无疑问使得同一性的百分比针对参考多肽序列全长计算,并且允许参考序列中氨基酸总数的最多5％的同源性空位。

[0323] 如本文所用,VWF序列或FVIII蛋白序列中的“对应氨基酸”或“等同氨基酸”通过比对识别以使第一VWF或FVIII序列和第二VWF或FVIII序列之间的同一性或相似性最大化。用于识别第二VWF或FVIII序列中的等同氨基酸的数量基于用于识别第一VWF或FVIII序列中的对应氨基酸的数量。

[0324] “融合”或“嵌合”蛋白包含连接至第二氨基酸序列的第一氨基酸序列,其在自然状态下不会天然连接。通常存在于单独蛋白质中的氨基酸序列可集合于融合多肽中,或通常存在于相同蛋白质中的氨基酸序列在融合多肽中可置于新排列,如本发明的因子VIII结构域与免疫球蛋白Fc结构域的融合。例如,通过化学合成或通过创建和翻译其中肽区以所需关系编码的多核苷酸来创建融合蛋白。嵌合蛋白还可包含通过共价非肽键或非共价键与第一氨基酸序列缔合的第二氨基酸序列。

[0325] 如本文所用,术语“半衰期”是指特定多肽在体内的生物半衰期。半衰期可通过施用给受试者的量的一半从动物中的循环和/或其它组织清除所需的时间来表示。当给定多肽的清除曲线构建为时间的函数时,曲线通常具有快速α-相和较长β-相两相。α-相通常表示所施用的Fc多肽在血管内腔和血管外间隙之间的平衡,并且部分通过多肽的大小确定。β-相通常表示血管内腔中多肽的分解代谢。在一些实施方案中,FVIII和包含FVIII的嵌合蛋白是单相的,并且因此不具有α-相,而只有单个β-相。因此,在某些实施方案中,如本文所用,术语半衰期是指β-相中多肽的半衰期。人抗体在人体中的典型β相半衰期为21天。

[0326] 应用于多核苷酸或多肽的术语“异源”意指多核苷酸或多肽来源于与其进行比较的实体不同的实体。因此,连接至VWF片段的异源多肽意指连接至VWF片段的多肽链,并且不是VWF片段的天然存在部分。例如,异源多核苷酸或抗原可来源于不同的物种、个体的不同细胞类型或不同个体的相同或不同类型细胞。

[0327] 如本文所用,术语“连接”是指分别共价或非共价连接至第二氨基酸序列或核苷酸序列的第一氨基酸序列或核苷酸序列。术语“共价连接”或“共价键合”是指共价键,如二硫键、肽键,或连接在一起的两个部分之间的一个或多个氨基酸,如连接基。第一氨基酸或核苷酸序列可直接连接第二氨基酸或核苷酸序列或与其并置,或者插入序列可将第一序列共价连接至第二序列。术语“连接”不仅意指第一氨基酸序列C-末端或N-末端融合至第二氨基酸序列,而且包括整个第一氨基酸序列(或第二氨基酸序列)插入第二氨基酸序列(或分别地第一氨基酸序列)中的任何两个氨基酸。在一个实施方案中,第一氨基酸序列可通过肽键或连接基连接至第二氨基酸序列。第一核苷酸序列可通过磷酸二酯键或连接基连接至第二核苷酸序列。连接基可以是肽或多肽(对于多肽链)或核苷酸或核苷酸链(对于核苷酸链)或任何化学部分(对于多肽和多核苷酸链二者)。共价键合有时以(-)或连字符表示。

[0328] 如本文所用,术语“与…缔合”指第一氨基酸链和第二氨基酸链之间形成的共价或非共价键。在一个实施方案中,术语“与…缔合”意指共价非肽键或非共价键。在一些实施方

案中，该缔合以冒号即(:)表示。在另一个实施方案中，其意指除肽键之外的共价键。在其它实施方案中，如本文所用，术语“共价缔合”意指两个部分之间通过共价键，如二硫键、肽键或一个或多个氨基酸(如，连接基)缔合。例如，氨基酸半胱氨酸包含可形成二硫键或与第二半胱氨酸残基上的巯基桥接的巯基。在大多数天然存在的IgG分子中，CH1和CL区通过二硫键缔合，并且两条重链通过两个二硫键在对应于使用Kabat编号系统的第239和242位(第226或229位，EU编号系统)的位置结合。共价键的实例包括但不限于：肽键、金属键、氢键、二硫键、σ键、π键、δ键、糖苷键、抓氢键(agnostic bond)、弯键、偶极键、π主链、双键、三键、四键、五键、六键、缀合作用、超缀合作用、芳香作用、哈普托数(hapticity)或反键。非共价键的非限制性实例包括离子键(如，阳离子-π键或盐键)、金属键、氢键(如，二氢键、二氢配合(dihydrogen complex)、低能障氢键或对称氢键)、范德华力、色散力(London dispersion force)、机械结合、卤键、亲金作用、嵌入、堆叠、熵力或化学极性。

[0329] 如本文所用，术语“单体-二聚体杂合物”是指包含通过二硫键彼此缔合的第一多肽链和第二多肽链的嵌合蛋白，其中第一链包含凝血因子，如因子VIII和Fc区，并且第二链包含、基本上由或由无凝血因子的Fc区组成。因此，单体-二聚体杂合构建体是包含只具有一个凝血因子的单体方面和具有两个Fc区的二聚体方面的杂合物。

[0330] 如本文所用，术语“切割位点”或“酶切割位点”是指酶识别的位点。某些酶切割位点包含细胞内加工位点。在一个实施方案中，多肽具有由在凝血级联反应期间活化的酶切割的酶切割位点，使得此类位点的切割在凝血形成部位进行。此类示例性位点包括例如由凝血酶、因子XIa或因子Xa识别的那些位点。示例性FXIa切割位点包括例如TQSFNDFTR(SEQ ID NO:47)和SVSQTSKLTR(SEQ ID NO:48)。示例性凝血酶切割位点包括例如DFLAEGGGVR(SEQ ID NO:49)、TTKIKPR(SEQ ID NO:50)、LVPRG(SEQ ID NO:55)和ALRPR(SEQ ID NO:51的第1至5位氨基酸)。其它酶切割位点是本领域已知的。

[0331] 如本文所用，术语“加工位点”或“细胞内加工位点”是指作为在多肽翻译后发挥功能的酶的靶标的多肽中的一种类型的酶切割位点。在一个实施方案中，此类酶在从高尔基腔至反面高尔基隔室的转运期间发挥功能。细胞内加工酶在蛋白质从细胞分泌之前切割多肽。此类加工位点的实例包括例如由PACE/弗林蛋白酶(其中PACE是成对碱性氨基酸裂解酶的缩写)家族内切肽酶靶向的那些。这些酶定位于高尔基体膜，并且切割序列基序Arg-[任何残基]-(Lys或Arg)-Arg的羧基末端侧的蛋白质。如本文所用，“弗林蛋白酶”家族酶包括例如PCSK1(也称为PC1/Pc3)、PCSK2(也称为PC2)、PCSK3(也称为弗林蛋白酶或PACE)、PCSK4(也称为PC4)、PCSK5(也称为PC5或PC6)、PCSK6(也称为PACE4)或PCSK7(也称为PC7/LPC、PC8或SPC7)。其它加工位点是本领域已知的。

[0332] 术语“弗林蛋白酶”是指对应于EC No.3.4.21.75的酶。弗林蛋白酶是枯草杆菌蛋白酶样前蛋白转化酶，其也称为PACE(成对碱性氨基酸裂解酶)。弗林蛋白酶缺失了失活前体蛋白的部分以将它们转化为生物活性蛋白。在其细胞内转运期间，原肽通过高尔基体中的弗林蛋白酶从成熟VWF分子切割。

[0333] 在包括超过一个加工或切割位点的构建体中，应当理解此类位点可以是相同的或不同的。

[0334] 如本文所用，止血病症意指特征在于由于形成血纤维蛋白凝块的能力受损或无能，具有自发或作为创伤结果的溢血趋势的基因遗传或获得的病症。此类疾病的实例包括

血友病。三种主要形式是甲型血友病(因子VIII缺失)、乙型血友病(因子IX缺失或“克雷司马疾病”)和丙型血友病(因子XI缺失,轻度出血趋势)。其它止血病症包括例如血管性血友病、因子XI缺失(PTA缺失)、因子XII缺失、血纤维蛋白原、凝血酶原、因子V、因子VII、因子X或因子XIII缺失或结构异常、GPIb缺陷或缺失的巨血小板综合征(Bernard-Soulier)。VWF的受体GPIb可为缺陷的,并导致初期凝块形成(初期止血)缺乏和出血趋势增加,以及格-尼二氏血小板无力症(thrombasthenia of Glanzman and Naegeli)(Glanzmann thrombasthenia)。在肝功能衰竭(急性和慢性形式)中,肝脏产生的凝血因子不足;这可增加出血风险。

[0335] 本发明的嵌合分子可预防地使用。如本文所用,术语“预防性治疗”是指在出血发作之前施用分子。在一个实施方案中,需要一般止血剂的受试者正在经历或将要经历外科手术。本发明的嵌合蛋白可在手术之前或之后作为预防即施用。本发明的嵌合蛋白可在外科手术期间或之后施用以控制急性出血发作。外科手术可包括但不限于肝脏移植、肝脏切除、牙科手术或干细胞移植。

[0336] 本发明的嵌合蛋白还用于按需(也称为“发病期”)治疗。术语“按需治疗”或“发病期治疗”是指响应出血发作的症状或在可引起出血的活动之前施用嵌合分子。在一个方面,当出血开始时,例如受伤后,或当预期要出血时,例如在外科手术之前,可给予受试者按需(发病期)治疗。在另一个方面,在增加出血风险的活动,例如接触运动之前可给予按需治疗。

[0337] 如本文所用,术语“急性出血”是指与根本原因无关的出血发作。例如,受试者可具有创伤、尿毒症、遗传性出血疾病(如,因子VII缺失)、血小板疾病或由于凝血因子抗体的发展而产生的抗性。

[0338] 如本文所用,治疗是指例如疾病或病症的严重性减少;病程的持续时间减少;与疾病或病症相关的一个或多个症状改善;为患有疾病或病症的受试者提供有益效果,而不必治愈疾病或病症,或与疾病或病症相关的一个或多个症状的预防。在一个实施方案中,术语“治疗”意指通过施用本发明的嵌合蛋白或VWF片段保持受试者中至少约1IU/dL、2IU/dL、3IU/dL、4IU/dL、5IU/dL、6IU/dL、7IU/dL、8IU/dL、9IU/dL、10IU/dL、11IU/dL、12IU/dL、13IU/dL、14IU/dL、15IU/dL、16IU/dL、17IU/dL、18IU/dL、19IU/dL或20IU/dL的FVIII谷底水平。在另一个实施方案中,治疗意指保持约1和约20IU/dL、约2和约20IU/dL、约3和约20IU/dL、约4和约20IU/dL、约5和约20IU/dL、约6和约20IU/dL、约7和约20IU/dL、约8和约20IU/dL、约9和约20IU/dL或约10和约20IU/dL之间的FVIII谷底水平。疾病或病症的治疗也可包括保持受试者中FVIII活性相当于非血友病受试者中FVIII活性的至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%或20%的水平。治疗所需的最小谷底水平可通过一种或多种已知方法测量,并且可为每个人调整(增加或减少)。

[0339] 嵌合蛋白

[0340] 本发明涉及通过阻止或抑制FVIII半衰期限制因子(如内源性VWF)在体内与FVIII蛋白缔合而延长因子VIII蛋白的半衰期。内源性VWF与非共价复合物中约95%至约98%的FVIII缔合。结合至FVIII蛋白的内源性VWF已知以各种方式保护FVIII。例如,全长VWF(作为具有约250kDa的多聚体)可保护FVIII不被蛋白酶切割和FVIII活化,稳定FVIII重链和/或

轻链以及防止FVIII被清除剂受体清除。然而，同时，内源性VWF通过阻止胞饮以及通过VWF清除途径从系统清除FVIII-VWF复合物限制FVIII半衰期。据信，如实例中所示，内源性VWF是半衰期限制因子，其防止融合至半衰期延长因子的FVIII蛋白的半衰期长于野生型FVIII的约两倍。因此，本发明使用辅助部分阻止或抑制内源性VWF和FVIII蛋白之间的相互作用，从而阻止FVIII蛋白被通过VWF清除途径清除和/或诱导胞饮。在一个实施方案中，辅助部分能够阻止或抑制FVIII蛋白与内源性VWF结合，并且具有至少一种VWF样FVIII保护特性。此外，辅助部分通过阻止或抑制与内源性VWF的相互作用减少FVIII从系统清除。本发明的辅助部分结合或缔合(如，通过非共价键合)FVIII蛋白和/或物理或化学阻断FVIII蛋白上的VWF结合位点。因此，与野生型FVIII或不与辅助部分缔合的FVIII相比，与辅助部分缔合的FVIII蛋白通过一个或多个VWF清除受体更缓慢地从循环清除。

[0341] 本发明辅助部分的实例包括例如多肽或FVIII蛋白的化学或物理修饰、添加、缺失或变型。用于本发明的辅助部分可包括多肽、非多肽部分或它们二者。作为辅助部分的多肽的非限制性实例包括例如本文所述的VWF片段、免疫球蛋白恒定区或其部分、转铁蛋白或其片段、白蛋白或其片段、白蛋白结合部分、HAP序列、PAS序列或它们的任何组合。非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。本发明使用的其它此类部分也是本领域已知的。

[0342] 在一个实施方案中，辅助部分通过共价或非共价键与FVIII蛋白缔合(或连接)。然而，在一些情况下，在内源性VWF存在下，辅助部分和FVIII蛋白之间的物理封闭或化学缔合(如，非共价键合)的强度可能不足以提供包含FVIII蛋白和辅助部分的稳定复合物。例如，在内源性VWF存在下，与FVIII蛋白形成无任何其它连接的非共价键的VWF片段可易于在体内从FVIII蛋白解离，从而用内源性VWF替代VWF片段(如，重组VWF，即rVWF)。因此，非共价结合内源性VWF的FVIII蛋白可经历VWF清除途径并从系统清除。为了阻止辅助部分与FVIII蛋白解离，在一些实施方案中，FVIII蛋白和辅助部分之间的键合是共价键，如肽键、一个或多个氨基酸或二硫键。在某些实施方案中，辅助部分和FVIII蛋白之间的缔合(即，键合)是肽键或FVIII蛋白和辅助部分之间的连接基(“FVIII/AM连接基”)。连接基的非限制性实例在本文别处有所描述。在一些实施方案中，辅助部分是包含、基本上由或由至少约10个、100个、200个、300个、400个、500个、600个、700个、800个、900个、1000个、1100个、1200个、1300个、1400个、1500个、1600个、1700个、1800个、1900个、2000个、2500个、3000个或4000个氨基酸组成的多肽。在其它实施方案中，辅助部分是包含、基本上由或由约100至约200个氨基酸、约200至约300个氨基酸、约300至约400个氨基酸、约400至约500个氨基酸、约500至约600个氨基酸、约600至约700个氨基酸、约700至约800个氨基酸、约800至约900个氨基酸或约900至约1000个氨基酸组成的多肽。在一些实施方案中，与FVIII蛋白共价缔合的辅助部分是本文别处所述的VWF片段。

[0343] 在某些实施方案中，辅助部分化学(如，非共价)结合或物理阻断FVIII蛋白上的一个或多个VWF结合位点。FVIII蛋白上的VWF结合位点位于FVIII蛋白的A3结构域或C2结构域内。在其它实施方案中，FVIII蛋白上的VWF结合位点位于A3结构域和C2结构域内。例如，FVIII蛋白上的VWF结合位点可对应于SEQ ID NO:16[全长成熟FVIII]的第1669至1689位和/或第2303至2332位氨基酸。

[0344] 在其它实施方案中，本发明的嵌合蛋白包含连接至辅助部分的FVIII蛋白，其中辅

助部分是VWF分子，如包含D'结构域和D3结构域，但不包含VWF清除受体结合位点的VWF片段，并且屏蔽或保护FVIII蛋白上的VWF结合位点，从而抑制或阻止FVIII蛋白与内源性VWF的相互作用。在某些实施方案中，辅助部分是VWF片段。用于本发明的VWF片段包含D'结构域和D3结构域，仍为FVIII蛋白提供VWF样特性的一个或多个优势，但VWF片段不经历VWF清除途径。FVIII蛋白和辅助部分可通过连接基(如，FVIII/AM连接基)共价缔合。在一个实施方案中，连接基可以是可切割的连接基。连接基的非限制性实例在本文别处有所公开。

[0345] 在其它实施方案中，本发明的嵌合蛋白包含FVIII蛋白和免疫球蛋白恒定区或其部分(即，辅助部分)，其中免疫球蛋白恒定区或其部分屏蔽或保护FVIII蛋白上的VWF结合位点，从而抑制或阻止FVIII蛋白与内源性VWF的相互作用。在其它实施方案中，免疫球蛋白恒定区或其部分是Fc区。

[0346] 在一个方面，本发明涉及嵌合或融合蛋白或包含本文所公开的一个或多个VWF片段的杂合物及其用途。嵌合或融合蛋白可融合或连接至一个或多个异源部分(本文有时以H或H1表示)。在一个实施方案中，异源部分(H1)是天然不与VWF片段一起存在和/或连接至VWF片段的异源肽或异源多肽。在另一个实施方案中，异源部分(H1)是非多肽部分，如肽或多肽与非多肽部分的化学修饰或组合。在一些实施方案中，VWF片段通过连接基(本文也称为“VWF连接基”)连接或键合至异源部分(H1)。在一个实施方案中，VWF连接基是可切割的连接基。VWF片段和异源部分(H1)之间的连接基的非限制性实例在本文别处有所公开。

[0347] 在一个实施方案中，用于本发明的异源部分(H1)改善了VWF片段的一种或多种药代动力学特性，而不显著影响VWF片段的生物学活性或功能(如，其与FVIII蛋白的结合或缔合)。在另一个实施方案中，连接至VWF片段的异源部分(H1)可延长VWF片段的半衰期。异源多肽部分的非限制性实例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的两种或多种组合。异源非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0348] 在一些实施方案中，异源部分(H1)可用于通过共价键连接VWF片段和FVIII蛋白。可提供共价键合的异源部分的实例包括但不限于包含铰链区如Fc区或FcRn结合伴侣的免疫球蛋白恒定区或其部分。在具体实例中，FVIII蛋白连接至第一Fc区，并且VWF片段连接至第二Fc区，其中第一Fc区和第二Fc区形成一个或多个二硫键。

[0349] 在一些实施方案中，异源部分(在本文中有时以“H”或“H1”表示)是免疫球蛋白恒定区或其部分。免疫球蛋白恒定区或其部分的非限制性实例可选自由以下组成的组：CH1结构域、CH2结构域、CH3结构域、CH4结构域、铰链结构域以及它们的两种或多种组合。在一个实施方案中，免疫球蛋白恒定区或其部分包含至少一个CH1结构域、至少一个CH2结构域、至少一个CH3结构域、至少一个CH4结构域或其功能片段。在另一个实施方案中，免疫球蛋白恒定区或其部分包含至少一个铰链结构域或其部分，以及至少一个CH2结构域或其部分(如，铰链-CH2方向)。在其它实施方案中，免疫球蛋白恒定结构域或其部分包含至少一个CH2结构域或其部分，以及至少一个CH3结构域或其部分(如，CH2-CH3方向)。组合的实例包括但不限于：CH2结构域、CH3结构域和铰链结构域，它们也称为Fc区(或Fc结构域)，如第一Fc区。在其它实施方案中，异源部分(H1)通过连接基连接至VWF片段。在某些实施方案中，异源部分(H1)是本文别处所述的FcRn结合伴侣。在其它实施方案中，异源部分(H1)是铰链区。

[0350] 在某些实施方案中，嵌合蛋白还包含第二(或另外的)异源部分(在本文中有时以“H2”表示)。值得注意的是，第一异源部分(H1)和第二异源部分(H2)可互换使用，并且可以是相同的或不同的。第二异源部分(H2)可通过肽键、一个或多个氨基酸，或通过连接基(如，FVIII连接基(如果连接至FVIII))连接至FVIII蛋白或嵌合蛋白的别处。此类构建体有时可称为FVIII/VWF异源二聚体。在一个实施方案中，异源部分(H2)包含异源多肽。在另一个实施方案中，异源部分(H2)包括非多肽部分。在其它实施方案中，异源部分(H2)包括异源部分和非多肽部分的组合。第二异源部分(H2)可以是半衰期延长因子。第二异源多肽部分(H2)的非限制性实例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的两种或多种组合。异源非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。在某些实施方案中，第一异源部分(H1)和第二异源部分是相同的或不同的。第一异源部分(H1)和第二异源部分(H2)的任一者或两者可赋予嵌合蛋白中FVIII蛋白的半衰期延长，提供比非共价缔合更强的连接，即通过嵌合蛋白中FVIII蛋白和VWF片段或它们二者之间的一个或多个共价键。一旦融合或连接至第一异源部分(H1)的VWF片段通过阻止或抑制FVIII蛋白和内源性VWF蛋白之间的相互作用移除半衰期上限，则融合至异源部分的FVIII蛋白可发挥其全部潜力，并且可具有长于野生型FVIII两倍的半衰期。

[0351] 在某些实施方案中，连接至VWF片段的第一异源部分(如，第一Fc区)和连接至FVIII蛋白的第二异源部分(如，第二Fc区)彼此缔合，使得该缔合阻止VWF片段在体内被内源性VWF替代。在一个实施方案中，第二异源部分是第二Fc区，其中第二Fc区通过共价键，如二硫键、肽键或连接基(一个或多个氨基酸)连接至第一异源部分，如第一Fc区或与之缔合。例如，在一个末端连接至FVIII蛋白的第二异源部分(如，第二Fc区)还可连接至第一异源部分(如，第一Fc区)，所述第一异源部分通过连接基(如，scFc连接基)连接至VWF片段，或通过共价或非共价键与第一异源部分缔合。在另一个实施方案中，第二异源部分(如，第二Fc区)连接至VWF片段，所述VWF片段已连接至第一异源部分。在一些实施方案中，嵌合蛋白包含第一多肽链和第二多肽链，所述第一多肽链包含VWF片段和第一异源部分，所述第二多肽链包含FVIII蛋白和第二异源部分，其中第一多肽链和第二多肽链缔合，其中包含第一异源部分的第一多肽链和包含第二异源部分的第二多肽链之间的缔合是共价键，因此允许VWF片段和FVIII蛋白保持其彼此相互作用。同时，可与FVIII蛋白形成非共价键的内源性VWF不能替代包含VWF片段的共价连接多肽链。

[0352] 第一异源部分(H1)和VWF片段(如，VWF连接基)之间的连接基可以是可切割的连接基，如凝血酶可切割连接基。可切割的连接基可被选自由以下组成的组的蛋白酶切割：因子XIa、因子XIIa、激肽释放酶、因子VIIa、因子IXa、因子Xa、因子IIa(凝血酶)、弹性蛋白酶-2、粒酶-B、TEV、肠激酶、蛋白酶3C、分选酶A、MMP-12、MMP-13、MMP-17、MMP-20以及它们的任何组合。在凝血级联反应活化时，这些可切割的连接基允许切割VWF片段，并且从FVIII蛋白解离，从而产生具有完全活性潜力的FVIII蛋白。

[0353] 在其它实施方案中，嵌合蛋白作为包含任何顺序的VWF片段、可切割的连接基、第一异源部分(H1)、可加工的连接基、FVIII蛋白和第二异源部分(H2)的单多肽链产生。在合成后，可加工的连接基可在分泌前被胞内蛋白酶切割，从而产生上述两条多肽链。在分泌前的单链构建体中，第二异源部分(如，第二Fc区)可通过可加工的连接基连接至VWF片段。在

某些实施方案中，一个或多个连接基可包含一个或多个切割位点。

[0354] 在一些实施方案中，本发明的嵌合蛋白还包含第三异源部分(本文有时以“H3”表示)。第三异源部分(H3)可以是半衰期延长因子。异源部分(H3)可包括异源多肽、非多肽部分或二者的组合。第三异源部分(H3)的非限制性实例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段、其任何衍生物或变体或它们的两种或多种组合。非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。连接至VWF片段的第一异源部分(H1)、连接至FVIII蛋白的第二异源部分(H2)和第三异源部分(H3)可以是相同的或不同的。在一个实施方案中，第一异源部分(H1)与第二异源部分(H2)相同，但不同于第三异源部分(H3)。在另一个实施方案中，第三异源部分(H3)融合或连接至嵌合蛋白的FVIII蛋白或VWF片段。在一些实施方案中，第三异源部分插入FVIII蛋白的一个或多个结构域中或FVIII蛋白的两个结构域之间。

[0355] 在一个实施方案中，嵌合蛋白包含第一多肽链和第二多肽链，其中第一链包含通过任选的连接基(如，FVIII连接基)连接至第一异源部分(H1)如第一Fc区的FVIII蛋白，并且第二链包含通过任选的连接基(如，VWF连接基)连接至第二异源部分(H2)如第二Fc区的VWF片段。FVIII蛋白还可包含第三异源部分(H3)，如任何半衰期延长部分，如白蛋白，或FVIII重链和FVIII轻链之间的PAS序列(即，SEQ ID NO:16的第1648位氨基酸残基)，从而作为单链FVIII蛋白。或者，FVIII蛋白可以是双链蛋白，即通过共价或非共价键(如，金属键)彼此缔合的FVIII重链和FVIII轻链，其中重链还连接至第三异源部分(H3)，如非结构半衰期延长多肽、白蛋白或其片段或PAS序列。在另一个实施方案中，嵌合蛋白包含第一多肽链和第二多肽链，其中第一链包含通过任选的连接基(如，FVIII连接基)连接至第一异源部分(H1)如第一Fc区的FVIII蛋白，并且第二链包含连接至第三异源部分(H3)如非结构半衰期延长多肽、白蛋白或PAS序列的VWF片段，其通过任选的连接基连接至第二异源部分(H2)如第二Fc区。在一些实施方案中，第三异源部分(H3)(如，半衰期延长多肽)可连接至FVIII蛋白的C-末端或N-末端或插入FVIII蛋白的两个结构域之间或FVIII蛋白的结构域中的两个氨基酸之间。

[0356] 在其它实施方案中，本发明的嵌合蛋白还包含第四异源部分(本文有时以“H4”表示)和/或第五异源部分(本文有时以“H5”表示)。第四或第五异源部分也可以是半衰期延长因子。第四异源部分和/或第五异源部分可与第三异源部分相同或不同。异源部分可包括异源多肽、非多肽部分或二者的组合。第四或第五异源部分的非限制性实例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段、其任何衍生物或变体或它们的两种或多种组合。非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。第一异源部分、第二异源部分、第三异源部分、第四异源部分和第五异源部分可以是相同的或不同的。在一些实施方案中，第四异源部分(如，半衰期延长多肽)可连接至FVIII蛋白的C-末端或N-末端或插入FVIII蛋白的两个结构域之间或FVIII蛋白的结构域中的两个氨基酸之间。在其它实施方案中，第五异源部分(如，半衰期延长多肽)也可连接至FVIII蛋白的C-末端或N-末端或插入FVIII蛋白的两个结构域之间或FVIII蛋白的结构域中的两个氨基酸之间。

[0357] 在某些实施方案中，嵌合蛋白包含FVIII蛋白、VWF片段、第一异源部分、第二异源

部分、第三异源部分、第四异源部分和第五异源部分，其中第一异源部分和第二异源部分形成包含FVIII蛋白的链和包含VWF片段的链之间的键(如，共价键)，并且第三异源部分、第四异源部分和第五异源部分是半衰期延长因子，并且其中包含FVIII蛋白的链和包含VWF片段的链之间的键强于FVIII和VWF片段之间的非共价相互作用，从而阻止内源性VWF在体内、体外或离体与FVIII蛋白结合。

[0358] 在其它实施方案中，嵌合蛋白包含FVIII蛋白、VWF片段、第一异源部分、第二异源部分、第三异源部分、第四异源部分、第五异源部分和第六异源部分(本文有时以“H6”表示)，其中第一异源部分和第二异源部分形成包含FVIII蛋白的链和包含VWF片段的链之间的键，并且第三异源部分、第四异源部分、第五异源部分和第六异源部分是半衰期延长因子，并且其中包含FVIII蛋白的链和包含VWF片段的链之间的键强于FVIII和VWF片段之间的相互作用，从而阻止内源性VWF在体内、体外或离体与FVIII蛋白结合。

[0359] 在一些实施方案中，嵌合蛋白包含选自由以下组成的组的式：

[0360] (aa)V-L1-H1-L2-H2，

[0361] (bb)H2-L2-H1-L1-V，

[0362] (cc)H1-L1-V-L2-H2，和

[0363] (dd)H2-L2-V-L1-H1，

[0364] 其中V包含本文所述的VWF片段；

[0365] L1和L2中的每个包含任选的连接基；并且

[0366] H1包含第一异源部分；并且

[0367] H2包含任选的第二异源部分。第一异源部分和第二异源部分的任一者或两者可以是半衰期延长部分。在一个实施方案中，H1包括多肽、非多肽部分或它们二者。用作H1的多肽可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、任何衍生物或变体或它们的任何组合。非多肽部分可包括聚乙二醇(PEG)、聚唾液酸、和羟乙基淀粉(HES)、其衍生物或变体或它们的任何组合。在另一个实施方案中，H2包括多肽、非多肽部分或它们二者。用作H2的多肽可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、任何衍生物或变体或它们的任何组合。非多肽部分可包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或变体或它们的任何组合。在某些实施方案中，式(aa)和(bb)中H1和H2之间的连接基是可加工的连接基。在其它实施方案中，式(aa)和(bb)中VWF片段和H1之间的连接基是可切割的连接基，如可被凝血酶切割的凝血酶可切割连接基。

[0368] 本文的式中多肽的方向从N-末端(左)至C-末端(右)列出。例如，式H-L-V意指式NH2-H-L-V-COOH。在一个实施方案中，本文所述的式可包含两个部分之间的另外序列。例如，式V-L1-H1-L2-H2还可包含在V的N-末端、V和L1之间、L1和H1之间、H1或L2之间、L2或H2之间或在H2的C-末端的序列，除非另外指明。在另一个实施方案中，连字符(-)表示肽键或一个或多个氨基酸。

[0369] 在具体实施方案中，嵌合蛋白包含、基本上由或由选自由以下组成的组的式中的一者或多者组成：(a1)V-H、(a2)H-V、(a3)V-L-H、(a4)H-L-V、(a5)V-L1-H1-H2、(a6)H2-H1-L1-V、(a7)V-L1-H1:H2、(a8)H2:H1-L1-V、(a9)V-H1:H2、(b1)H2:H1-V、(b2)V-L1-H1-L2-H2、(b3)H2-L2-H1-L1-V、(b4)H1-V-H2、(b5)H1-L1-V-L2-H2和(b6)H2-L2-V-L1-H1，其中V包含

本文所述的VWF片段中的一者或多者，L、L1或L2包含连接基，H或H1包含第一异源部分。在一个实施方案中，第一异源部分(H1)可以是多肽、非多肽部分或它们二者。异源多肽部分可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列或它们的任何组合。用作H1的非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。在另一个实施方案中，H2包含第二异源部分。第二异源部分可以是多肽、非多肽部分或它们二者。异源多肽部分可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列或它们的任何组合。用作H1的非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。在某些实施方案中，第一异源部分和第二异源部分之间的连接基是可加工的连接基。在其它实施方案中，VWF片段和第一异源部分或第二异源部分之间的连接基是可切割的连接基，其包含一个或多个切割位点，如凝血酶可切割连接基。

[0370] 本发明的嵌合蛋白包含选自由以下组成的组的式：(aa)、(bb)、(cc)、(dd)、(a1)、(a2)、(a3)、(a4)、(a5)、(a6)、(a7)、(a8)、(a9)、(b1)、(b2)、(b3)、(b4)、(b5)和(b6)以及FVIII蛋白，其共价连接至式中的VWF片段、第一异源部分(如，第一Fc区)、或第二异源部分(如，第二Fc区)或与其共价缔合。在一个实施方案中，FVIII蛋白通过共价或非共价键或通过连接基连接至VWF片段或与其缔合。在另一个实施方案中，FVIII蛋白可通过共价或非共价键或通过连接基连接至第一异源部分或第二异源部分。

[0371] 在一个实施方案中，本发明的嵌合蛋白包含共价连接至FVIII蛋白或与其共价缔合的本文所述VWF片段。例如，嵌合蛋白可包含VWF片段和FVIII蛋白，其中VWF片段和FVIII蛋白通过共价非肽键、肽键、非共价键，或通过连接基如可切割的连接基结合。在具体实施方案中，VWF片段和FVIII蛋白通过一个或多个二硫键结合或彼此相互作用。在另一个具体实施方案中，VWF片段通过非共价键结合位于FVIII的A3结构域、FVIII的C2结构域或FVIII的A3结构域和C2结构域二者的FVIII蛋白或与其相互作用。在另一个实施方案中，结合FVIII蛋白或与其相互作用的VWF片段连接或融合至第一异源部分。在其它实施方案中，结合VWF片段或与其相互作用的FVIII蛋白还连接至第二异源部分。在一些实施方案中，结合FVIII蛋白或与其相互作用的VWF片段还连接至第一异源部分，并且FVIII蛋白还连接至第二异源部分。在某些实施方案中，包含VWF片段和第一异源部分的第一多肽链以及包含FVIII蛋白和第二异源部分的第二多肽链彼此缔合，使得缔合不允许FVIII蛋白与其它部分如内源性VWF的相互作用。在一个实施方案中，缔合是共价键，如二硫键。

[0372] 每个VWF片段或FVIII蛋白可通过连接基，如可切割的连接基，如凝血酶可切割连接基结合或连接至第一和第二异源部分。VWF片段和第一异源部分之间的连接基在本文中可以VWF连接基表示。FVIII蛋白和第二异源部分之间的连接基在本文中可以FVIII连接基表示。或者，VWF片段或FVIII蛋白二者可通过连接基，如可切割的连接基，如凝血酶可切割连接基结合或连接至第一和第二异源部分。在某些实施方案中，连接至VWF片段的第一异源部分包括多肽、非多肽部分或它们二者。第一异源多肽部分的非限制性实例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的两种或多种组合。非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES或HAES)、其衍生物或变体或它们的任何组合。在其它实施方案中，连接至FVIII蛋白的第二异源部分包括多肽、非多肽部分或它们二者。第二异源部分的非限制性实

例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段或它们的两种或多种组合。非多肽部分的非限制性实例包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES或HAES)、其衍生物或变体或它们的任何组合。在一些实施方案中,VWF片段使用分选酶介导的体外蛋白连接而连接至FVIII。在一些实施方案中,使用分选酶识别基序。

[0373] 在一个实施方案中,第一异源部分是免疫球蛋白恒定区或其部分。在具体实施方案中,第一异源部分是第一Fc区。在一些实施方案中,第二异源部分是免疫球蛋白恒定区或其部分。在具体实施方案中,第二异源部分是第二Fc区。在具体实施方案中,嵌合蛋白包含本文所述的VWF片段和FVIII蛋白,其中VWF片段连接至免疫球蛋白恒定区或其部分,其为Fc区。在另一个实施方案中,嵌合蛋白包含本文所述的VWF片段和FVIII蛋白,其中FVIII蛋白连接至免疫球蛋白恒定区或其部分,其为Fc区。在其它实施方案中,嵌合蛋白包含本文所述的VWF片段和FVIII蛋白,其中VWF片段连接至第一免疫球蛋白恒定区,其为第一Fc区,并且FVIII蛋白连接至第二免疫球蛋白恒定区,其为第二Fc区,并且其中VWF片段和FVIII蛋白通过非共价键彼此结合或相互作用,或第一Fc区或第二Fc区通过共价键彼此缔合。在其它实施方案中,连接至第一异源部分的VWF片段还通过连接基如可加工的连接基连接至第二异源部分如第二Fc区。在一个方面,VWF片段通过连接基如VWF连接基如可切割的连接基连接至第一异源部分。在另一个方面,FVIII蛋白通过连接基如FVIII连接基如可切割的连接基连接至第二异源部分。异源部分的非限制性实例在本文别处有所公开,如第[0165]-[0193]段的免疫球蛋白恒定区或其部分、第[0194]-[0198]段的白蛋白片段或其变体、第[0293]段的HAP序列、第[0204]-[0205]段的转铁蛋白片段或其变体、第[0206]-[0213]段的聚合物如聚乙二醇、第[0214]-[0219]段的HES或第[0220]段的PSA以及第[0199]-[0202]段的PAS序列。

[0374] 在一些实施方案中,本发明的嵌合蛋白包含、基本上由或由选自由以下组成的组的式组成:

[0375] (a)V-L1-H1-L3-C-L2-H2,

[0376] (b)H2-L2-C-L3-H1-L1-V,

[0377] (c)C-L2-H2-L3-V-L1-H1,

[0378] (d)H1-L1-V-L3-H2-L2-C,

[0379] (e)H1-L1-V-L3-C-L2-H2,

[0380] (g)H2-L2-C-L3-V-L1-H1,

[0381] (g)V-L1-H1-L3-H2-L2-C,

[0382] (g)C-L2-H2-L3-H1-L1-V,

[0383] (i)H2-L3-H1-L1-V-L2-C,

[0384] (j)C-L2-V-L1-H1-L3-H2,

[0385] (k)V-L2-C-L1-H1-L3-H2,和

[0386] (l)H2-L3-H1-L1-C-L2-V,

[0387] 其中V是本文所述的VWF片段;

[0388] L1或L2中的每个是任选的连接基,如可切割的连接基,如凝血酶可切割连接基;

[0389] L3是任选的连接基,如可加工的连接基

[0390] H1和H2中的每个是任选的异源部分；

[0391] C是FVIII蛋白；并且

[0392] (-)是肽键或一个或多个氨基酸。

[0393] 在其它方面，本发明的嵌合蛋白包含选自由以下组成的组的式：

[0394] (m)V-L1-H1:H2-L2-C，

[0395] (n)V-L1-H1:C-L2-H2；

[0396] (o)H1-L1-V:H2-L2-C；

[0397] (p)H1-L1-V:C-L2-H2；

[0398] (q)V:C-L1-H1:H2；

[0399] (r)V:H1-L1-C:H2；

[0400] (s)H2:H1-L1-C:V，

[0401] (t)C:V-L1-H1:H2，和

[0402] (u)C:H1-L1-V:H2。

[0403] 其中V是本文所述的VWF片段；

[0404] L1或L2中的每个是任选的连接基，如凝血酶可切割连接基；

[0405] H1或H2中的每个是任选的异源部分；

[0406] (-)是肽键或一个或多个氨基酸；并且

[0407] C是FVIII蛋白；并且(:)是H1和H2之间的化学或物理缔合。

[0408] 在一个实施方案中，一个或多个异源部分是半衰期延长因子。半衰期延长因子是本领域已知的，并且此类半衰期延长因子的非限制性实例包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、转铁蛋白或其片段、其衍生物或变体或它们的两种或多种组合。非多肽部分可包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。

[0409] 在一个实施方案中，式(m)至(u)中的(:)表示化学缔合，如至少一个非肽键。在某些实施方案中，所述化学缔合即(:)是共价键。在其它实施方案中，所述化学缔合即(:)是非共价相互作用，如离子相互作用、疏水相互作用、亲水相互作用、范德华相互作用、氢键。在其它实施方案中，(:)是非肽共价键。在其它实施方案中，(:)是肽键。在其它实施方案中，式(m)至(u)中的(:)表示两个序列之间的物理缔合，其中第一序列的一部分靠近第二序列，使得第一序列屏蔽或阻断第二序列的一部分与另一部分的相互作用，并且还保持该物理缔合，而不允许第二序列与其它部分相互作用。

[0410] 本文中包括的式(a)至(u)仅仅作为本发明构建体的非限制性实例。式中多肽的方向从N-末端(左)至C-末端(右)示出。例如，式V-L1-H1-L3-C-L2-H2意指式NH2-V-L1-H1-L3-C-L2-H2-COOH。此外，(:)可以是两条多肽链通过第一链的任何部分和第二链的任何部分之间的共价键或非共价键之间的缔合或相互作用，除非另外指明。例如，式V-H1:H2-C具有两条多肽链，第一链是V-H1，并且第二链是C-H2，其中第一链中的V与第二链中的C相互作用或缔合和/或第一链中的H1与第二链中的H2相互作用或缔合。在一些实施方案中，(:)意指共价非肽键或非共价键。

[0411] 在某些实施方案中，嵌合蛋白包含、基本上由或由选自由以下组成的组的式组成：

[0412] (1)V:C，(2)H-V:C或C:V-H，

[0413] (3)V:C-H或H-C:V,(4)V-H1:H2-C或H1-V:C-H2,

[0414] (5)V:C-H1:H2或H2:H1-C:V,(6)H2:H1-V:C或C:V-H1:H2,

[0415] (7)H-L-V:C或C:V-L-H,(8)V:C-L-H或H-L-C:V,

[0416] (9)V-C或C-V,(10)H-V-C或C-V-H,

[0417] (11)V-H-C或C-H-V,(12)V-C-H或H-C-V,

[0418] (13)V-H1-C-H2或H2-C-H1-V,(14)H1-V-C-H2或H2-C-V-H1,

[0419] (15)H1-V-H2-C或C-H2-V-H1,(16)V-H1-H2-C或C-H2-H1-V,

[0420] (17)V-L-C或C-L-V,(18)H-L-V-C或C-V-L-H,

[0421] (19)H-V-L-C或C-L-V-H,(20)V-L-H-C或C-H-L-V,

[0422] (21)V-H-L-C或C-L-H-V,(22)V-L-C-H或H-C-L-V,

[0423] (23)V-C-L-H或H-L-C-V,(24)H-L1-V-L2-C或C-L2-V-L1-H,

[0424] (25)V-L-H1:H2-C或C-H2:H1-L-V,

[0425] (26)V-H1:H2-L-C或C-L-H2:H1-V,

[0426] (27)V:C-H1-H2或H2-H1-C:V,

[0427] (28)H2-H1-V:C或C:V-H1-H2,

[0428] (29)V:C-L-H1:H2或H2:H1-L-C:V,

[0429] (30)H2:H1-L-V:C或C:V-L-H1:H2,

[0430] (31)V-L1-H1:H2-L2-C或L-L2-H2:H1-L1-V,

[0431] (32)V:C-L-H1-H2或H2-H1-L-C:V,

[0432] (33)V:C-H1-L-H2或H2-L-H1-C:V,

[0433] (34)V:C-L1-H1-L2-H2或H2-L2-H1-L1-C:V,

[0434] (35)H2-H1-V:C或C:V-H1-H2,

[0435] (36)H2-H1-L-V:C或C:V-L-H1-H2,

[0436] (37)H2-L-H1-V:C或C:V-H1-L-H2,

[0437] (38)H2-L2-H1-L1-V:C或C:V-L1-H1-L2-H2,

[0438] (39)V-L1-H-L2-C或C-L2-H-L1-V,

[0439] (40)V-L1-C-L2-H或H-L2-C-L1-V,

[0440] (41)V-L-H1-C-H2或H2-C-H1-L-V,

[0441] (42)V-H1-C-L-H2或H2-L-C-H1-V,

[0442] (43)V-H1-L-C-H2或H2-C-L-H1-V,

[0443] (44)H1-L-V-C-H2或H2-C-V-L-H1,

[0444] (45)H1-V-L-C-H2或H2-C-L-V-H1,

[0445] (46)H1-V-C-L-H或H-L-C-V-H1,

[0446] (47)H1-L-V-H2-C或C-H2-V-L-H1,

[0447] (48)H1-V-L-H2-C或C-H2-L-V-H1,

[0448] (49)H1-V-H2-L-C或C-L-H2-V-H1,

[0449] (50)V-L-H1-H2-C或C-H2-H1-L-V,

[0450] (51)V-H1-L-H2-C或C-H2-L-H1-V,

[0451] (52)V-H1-H2-L-C或C-L-H2-H1-V,

[0452] (53)V-L1-H1-L2-C-H2或H2-C-L2-H1-L1-V，

[0453] (54)V-L1-H1-C-L2-H2或H2-L2-C-H1-L1-V，

[0454] (55)V-L1-H1-L2-C-L3-H2或H2-L3-C-L2-H1-L1-V，

[0455] (56)V-H1-L1-C-L2-H2或H2-L2-C-L1-H1-V，

[0456] (57)H1-L1-V-L2-C-H2或H2-C-L2-V-L1-H1，

[0457] (58)H1-L1-V-C-L2-H2或H2-L2-C-V-L1-H1，

[0458] (59)H1-L1-V-L2-C-L3-H2或H2-L3-C-L2-V-L1-H1，

[0459] (60)H1-V-L1-C-L2-H2或H2-L2-C-L1-V-H1，

[0460] (61)H1-L1-V-L2-H2-C或C-H2-L2-V-L1-H1，

[0461] (62)H1-L1-V-H2-L2-C或C-L2-H2-V-L1-H1，

[0462] (63)H1-L1-V-L2-H2-L3-C或C-L3-H2-L2-V-L1-H1，

[0463] (64)H1-V-L1-H2-L2-C或C-L2-H2-L1-V-H1，

[0464] (65)V-L1-H1-L2-H2-C或C-H2-L2-H1-L1-V，

[0465] (66)V-L1-H1-H2-L2-C或C-L2-H2-H1-L1-V，

[0466] (67)V-L1-H1-L2-H2-L3-C或C-L3-H2-L2-H1-L1-V,以及

[0467] (68)V-H1-L1-H2-L2-C或C-L2-H2-L1-H1-V，

[0468] V是本文所述的VWF片段；

[0469] C是FVIII蛋白；

[0470] H或H1是异源部分或第一异源部分；

[0471] H2是第二异源部分;第一和第二异源部分可以是相同的或不同的；

[0472] L、L1或L2中的每个是任选的连接基；

[0473] (-)是肽键或一个或多个氨基酸;并且

[0474] (:)是化学或物理缔合。每个连接基可以是相同的或不同的,并且每个可以是包含一个或多个酶切割位点的可切割的连接基。异源部分可以是本领域已知的半衰期延长技术、多肽、非多肽部分或它们二者。多肽部分可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、PAS序列、HAP序列、其任何衍生物或变体或它们的任何组合(如,Fc区)。非多肽部分可包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或变体或它们的任何组合。H、H1或H2中的每个可独立地根据特性选择,并且可以是全部相同的,或每一个是不同的。异源部分的非限制性实例在本文别处有所公开,如第[0126]-[0153]段的免疫球蛋白恒定区或其部分、第[0154]-[0157]段的白蛋白或其片段或变体、第[0166]-[0173]段的聚合物如聚乙二醇以及第[0159]-[0162]段的PAS序列。本文中包括式(1)至(68)仅仅作为本发明构建体的非限制性实例。

[0475] 在一个实施方案中,(:)表示化学缔合,如至少一个非肽键。在某些实施方案中,所述化学缔合即(:)是共价键。在其它实施方案中,所述化学缔合即(:)是非共价相互作用,如离子相互作用、疏水相互作用、亲水相互作用、范德华相互作用、氢键。在其它实施方案中,(:)是非肽共价键。在其它实施方案中,(:)是肽键。在其它实施方案中,(:)表示两个序列之间的物理缔合,其中第一序列的一部分靠近第二序列,使得第一序列屏蔽或阻断第二序列的一部分与另一部分的相互作用,并且还保持该物理缔合,而不允许第二序列与其它部分相互作用。

[0476] 在一个实施方案中，连接至嵌合蛋白中的VWF片段的第一异源部分(H或H1)是第一Fc区。在另一个实施方案中，连接至嵌合蛋白中的FVIII蛋白的第二异源部分(或H2)是第二Fc区。

[0477] 在某些实施方案中，本发明的嵌合蛋白包含两条多肽链，第一链包含、基本上由或由编码FVIII(如，单链FVIII)的氨基酸序列和第一异源部分(如，第一Fc区)组成，并且第二链包含、基本上由或由编码包含D'结构域和D3结构域的VWF片段的氨基酸序列、第二异源部分(如，第二Fc区)以及VWF片段和第二Fc结构域之间的连接基(如，VWF连接基)组成。VWF片段和第二Fc结构域之间的连接基可以是凝血酶可切割连接基。在一些实施方案中，单链FVIII蛋白包含第三异源部分，如半衰期延长因子，其连接至N-末端、C-末端或FVIII序列内的一个或多个位点。

[0478] 在其它实施方案中，本发明的嵌合蛋白包含三条多肽链，其中第一链包含、基本上由或由FVIII的重链组成，第二链包含、基本上由或由融合至第一异源部分(如，第一Fc区)的FVIII的轻链组成，并且第三多肽链包含、基本上由或由包含D'结构域和D3结构域的VWF片段、第二异源部分(如，第二Fc区)和连接基组成。VWF片段和第二异源部分之间的连接基可以是凝血酶可切割连接基。在一些实施方案中，重链FVIII连接至第三异源部分，如半衰期延长因子，其可连接至N-末端、C-末端或FVIII序列内的一个或多个位点。

[0479] 在其它实施方案中，本发明的嵌合蛋白包含两条多肽链，第一链包含、基本上由或由FVIII的重链组成，并且第二链包含、基本上由或由FVIII的轻链、第一异源部分(如，第一Fc区)、第一连接基(如，包含一个或多个细胞内加工位点的蛋白酶切割位点)、VWF片段、第二连接基(如，凝血酶可切割连接基)和第二异源部分(如，第二Fc区)组成，其中FVIII的轻链连接至第一异源部分(如，第一Fc区)，其还通过第一连接基(如，具有包含一个或多个细胞内加工位点的蛋白酶切割位点的可加工的连接基)连接至VWF片段，并且其中VWF片段通过第二连接基(如，凝血酶可切割连接基)连接至第二Fc区。在某些实施方案中，第一连接基和第二连接基是相同的或不同的。

[0480] 在某些实施方案中，本发明的嵌合蛋白包含一条多肽链，其包含单链FVIII蛋白、第一异源部分(如，第一Fc区)、第一连接基(如，凝血酶可切割连接基)、VWF片段、第二连接基(如，凝血酶可切割连接基)和第二异源部分(如，第二Fc区)，其中单链FVIII蛋白连接至第一异源部分，所述第一异源部分还通过第一连接基连接至VWF片段，并且VWF片段通过第二连接基连接至第二Fc区。在一个实施方案中，第一连接基是包含第一可切割位点和第二可切割位点的可切割的连接基。在另一个实施方案中，第二连接基是包含一个或两个可切割位点的可切割的连接基。在具体实施方案中，第二连接基是凝血酶可切割连接基。用于本发明的连接基可具有任何长度，如至少10个、50个、100个、200个、300个、400个、500个、600个或700个氨基酸。例如，连接基可具有20个氨基酸、35个氨基酸、42个氨基酸、73个氨基酸或98个氨基酸。

[0481] 在某些实施方案中，VWF片段通过肽键或连接基直接连接至FVIII蛋白。作为直接或通过连接基连接VWF片段和FVIII蛋白的一种方式，可利用酶连接(如，分选酶)。例如，分选酶是指通过识别和切割羧基末端分选信号来修饰表面蛋白的一组原核生物酶。对于分选酶的大多数底物，识别信号由基序LPXTG(Leu-Pro-任何残基-Thr-Gly(SEQ ID NO:106)和高度疏水的跨膜序列以及碱性残基簇例如精氨酸组成。切割在Thr和Gly之间进行，其中瞬

间结合从连接伴侣的Thr残基至活化位点Cys残基进行，然后是将蛋白质共价结合至细胞壁的转肽作用。在一些实施方案中，连接伴侣包含Gly(n)。

[0482] 在一个实施方案中，通过任选的连接基连接至分选酶识别基序的VWF片段可通过分选酶融合至连接至Gly(n)的FVIII蛋白，其中n可以是任何整数。连接构建体包含VWF片段(构建体的N-末端部分)和FVIII蛋白(构建体的C-末端部分)，其中分选酶识别基序插入其间。示例性构建体如图24(A)中所示。另一个连接构建体包含VWF片段(构建体的N-末端部分)、连接基、分选酶识别基序和FVIII蛋白(构建体的C-末端部分)(如，图24(C))。在另一个实施方案中，通过任选的连接基连接至分选酶识别基序的FVIII蛋白可通过分选酶融合至连接至Gly(n)的VWF片段，其中n是任何整数。所得的连接构建体包含FVIII蛋白(构建体的N-末端部分)和VWF片段(构建体的C-末端部分)，其中分选酶识别基序插入其间。示例性构建体如图24(B)中所示。另一个所得的连接构建体包含FVIII蛋白(构建体的N-末端部分)、连接基、分选酶识别基序和VWF片段(构建体的C-末端部分)(如，图24(D))。在其它实施方案中，通过第一任选的连接基连接至分选酶识别基序的VWF片段可融合至异源部分，如免疫球蛋白恒定区或其部分，如通过第二任选的连接基连接至凝血酶切割位点的Fc区。所得的构建体可包含VWF片段(N-末端部分)、第一连接基、分选酶识别基序、蛋白酶切割位点、第二任选的连接基和异源部分(如，图24(E))。在某些实施方案中，该所得的构建体是包含FVIII蛋白和第二异源部分，如免疫球蛋白恒定区或其部分，如第二Fc区的嵌合蛋白的部分。在一个实例中，在另一个实例中，嵌合包含三条多肽链，包含VWF片段、第一连接基、分选酶识别基序、蛋白酶切割位点、第二任选的连接基、第一异源部分的第一链，包含FVIII蛋白的轻链和第二异源部分的第二链，以及包含FVIII蛋白的重链的第三链。

[0483] 在其它实施方案中，本发明的嵌合蛋白包含VWF片段和FVIII蛋白，其中彼此共价缔合或彼此共价连接的VWF片段和FVIII蛋白具有的免疫原性小于无VWF片段的FVIII蛋白。免疫原性减小包括但不限于体液免疫响应减少，如中和抗体滴定度减小，或针对FVIII的细胞介导免疫响应，如生成各种细胞因子减少。

[0484] 在其它实施方案中，作为本发明的结果，与无VWF片段的FVIII蛋白或野生型FVIII相比，FVIII蛋白(或嵌合蛋白)的半衰期延长。FVIII蛋白的半衰期比无VWF片段的FVIII蛋白的半衰期长至少约1.5倍、至少约2倍、至少约2.5倍、至少约3倍、至少约4倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约11倍或至少约12倍。在一个实施方案中，FVIII的半衰期比野生型FVIII的半衰期长约1.5倍至约20倍、约1.5倍至约15倍或约1.5倍至约10倍。在另一个实施方案中，与野生型FVIII或无VWF片段的FVIII蛋白相比，FVIII的半衰期延长约2倍至约10倍、约2倍至约9倍、约2倍至约8倍、约2倍至约7倍、约2倍至约6倍、约2倍至约5倍、约2倍至约4倍、约2倍至约3倍、约2.5倍至约10倍、约2.5倍至约9倍、约2.5倍至约8倍、约2.5倍至约7倍、约2.5倍至约6倍、约2.5倍至约5倍、约2.5倍至约4倍、约2.5倍至约3倍、约3倍至约10倍、约3倍至约9倍、约3倍至约8倍、约3倍至约7倍、约3倍至约6倍、约3倍至约5倍、约3倍至约4倍、约4倍至约6倍、约5倍至约7倍或约6倍至约8倍。在其它实施方案中，FVIII的半衰期为至少约17小时、至少约18小时、至少约19小时、至少约20小时、至少约21小时、至少约22小时、至少约23小时、至少约24小时、至少约25小时、至少约26小时、至少约27小时、至少约28小时、至少约29小时、至少约30小时、至少约31小时、至少约32小时、至少约33小时、至少约34小时、至少约35小时、至少约36小时、至少约48小时、至

少约60小时、至少约72小时、至少约84小时、至少约96小时或至少约108小时。在其它实施方案中，FVIII的半衰期为约15小时至约两周、约16小时至约一周、约17小时至约一周、约18小时至约一周、约19小时至约一周、约20小时至约一周、约21小时至约一周、约22小时至约一周、约23小时至约一周、约24小时至约一周、约36小时至约一周、约48小时至约一周、约60小时至约一周、约24小时至约六天、约24小时至约五天、约24小时至约四天、约24小时至约三天或约24小时至约两天。

[0485] 在一些实施方案中，每个受试者的FVIII蛋白的平均半衰期为约15小时、约16小时、约17小时、约18小时、约19小时、约20小时、约21小时、约22小时、约23小时、约24小时(1天)、约25小时、约26小时、约27小时、约28小时、约29小时、约30小时、约31小时、约32小时、约33小时、约34小时、约35小时、约36小时、约40小时、约44小时、约48小时(2天)、约54小时、约60小时、约72小时(3天)、约84小时、约96小时(4天)、约108小时、约120小时(5天)、约六天、约七天(一周)、约八天、约九天、约10天、约11天、约12天、约13天或约14天。

[0486] 在某些实施方案中，与由FVIII或FVIII单体-二聚体杂合物组成的多肽相比，共价连接至VWF片段的FVIII蛋白的半衰期在FVIII/VWF双敲除(“DKO”)小鼠中可延长。

[0487] A)血管性血友病因子(VWF)片段

[0488] VWF(也称为F8VWF)是存在于血浆中的大多聚体糖蛋白，并且主要组成型地产生于内皮细胞(Weibel-Palade小体中)、巨核细胞(血小板的α颗粒)和内皮下结缔组织中。碱性VWF单体是具有2813个氨基酸的蛋白质。每个单体包含多个具有特定功能的特定结构域，D’和D3结构域(共同结合因子VIII)、A1结构域(结合血小板GPIb-受体、肝素和/或可能结合胶原)、A3结构域(结合胶原)、C1结构域(其中RGD结构域在活化时结合血小板整联蛋白αIIbβ3)和蛋白质C-末端的“半胱氨酸结”结构域(VWF与血小板衍生生长因子(PDGF)、转化生长因子-β(TGFβ)和β-人绒毛膜促性腺素(βHCG)共有)。

[0489] 人VWF的2813个单体氨基酸序列在Genbank中以登录号_NP_000543.2_报告。编码人VWF的核苷酸序列在Genbank中以登录号_NM_000552.3_报告。人VWF的核苷酸序列指定为SEQ ID NO:1。SEQ ID NO:2是由SEQ ID NO:1编码的氨基酸序列。VWF的每个结构域在表1中列出。

[0490] 表1

[0491]

VWF 结构域	氨基酸序列
VWF 信号肽 (SEQ ID NO: 2的第1至22位氨基酸)	1 MIPARFAGVL LALALILPGT LC 22
VWF D1D2区 (SEQ ID NO: 2的第23至763位氨基酸)	23 AEGTRGRS STARCSLFGS DFVNTFDGSM 51 YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE FFDIHLFVNG 101 TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL 151 SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS WALSSGEQWC 201 ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC 251 EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME 301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC 351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD 401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG 451 LHNSLVKLKH GAGVAMDGQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM 501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG 551 NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS 601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL 651 NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGCFCP PGLYMDERGD 701 CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD 751 AVLSSPLSHR SKR 763
VWF D'结构域 (SEQ ID NO:2的第764至866位氨基酸)	764 SLSCRPP MVKLVCPADN LRAEGLECTK TCQNYDLECM 801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV 851 CRDRKWNCTD HVCDAT 866
VWF D3结构域 (SEQ ID NO: 2的第867至1240位氨基酸)	867 CSTI GMAHYLTFDG LKYLFPGECQ YVLVQDYCGS 901 NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE 951 THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD 1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI

[0492]

	1051 MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS CESIGDCACF 1101 CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY ECEWRYNSCA 1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP 1240
VWF A1结构域 (SEQ ID NO: 2的第1241至1479位氨基酸)	1241 GGLVVPPTDA 1251 PVSPTTLYVE DISEPPLHDF YCSRLLDLVF LLDGSSRLSE AEFEVLKAFV 1301 VDMMERLRIS QKWVRVAVVE YHDGSHAYIG LKDRKRPSEL RRIASQVKYA 1351 GSQVASTSEV LKYTLFQIFS KIDRPEASRI ALLLMASQEP QRMSRNFVRY 1401 VQGLKKKKVI VIPVGIGPHA NLKQIRLIEK QAPENKAFVL SSVDELEQQR 1451 DEIVSYLCDL APEAPPPTLP PDMAQVTVG 1479
	1480 P GLLGVSTLGP KRNSMVLDVA 1501 FVLEGSDKIG EADFNRSKEF MEEVIQRMDV GQDSIHVTVL QYSYMVTVEY 1551 PFSEAQSKGD ILQRVREIRY QGGNRTNTGL ALRYLSDHSF LVSQGDREQA 1600 1601 PNLVYMVTGN PASDEIKRLP GDIQVVPIGV GPNANVQELE RIGWPNAPIL 1651 IQDFETLPRE APDLVLQRCC SGEGLQIPTL SPAPDCSQPL DVILLLDGSS 1701 SFPASYFDEM KSFAKAFISK ANIGPRLTQV SVLQYGSITT IDVPWNVVPE 1751 KAHLLSLVDV MQREGGPSQI GDALGFAVRY LTSEMHGARP GASKAVVILV 1801 TDVSVDSVDA AADAARSNRV TVFPIGIGDR YDAAQLRILA GPAGDSNVVK 1851 LQRIEDLPTM VTLGNSFLHK LCSGFVRICM DEDGNEKRPG DVWTLPDQCH 1901 TVTCQPDGQT LLKSHRVNCD RGLRPSCPNS QSPVKVEETC GCRWTCPCVC 1951 TGSSTRHIVT FDGQNFKLTG SCSYVLFQNK EQDLEVILHN GACSPGARQG 2001 CMKSIEVKHS ALSVEXHSDM EVTVNGRLVS VPYVGGNMEV NVYGAIMHEV 2051 RFNHLGHIFT FTPQNNEFQL QLSPKTFASK TYGLCGICDE NGANDFMLRD 2101 GTVTTDWKTL VQEWTVQRPG QTCQPILEEQ CLVPDSSHCQ VLLLPLFAEC 2151 HKVLAPATFY AICQQDSCHQ EQVCEVIASY AHLCRTNGVC VDWRTPDFCA 2201 MSCPPSLVYN HCEHGCPRHC DGNVSSCGDH PSEGCFCPPD KVMLEGSCVP 2251 EEACTQCIGE DGVQHQFLEA WVPDHQPCQI CTCLSGRKVN CTTQPCPTAK 2301 APTCGLCEVA RLRQNADQCC PEYECVCDPV SCDLPPVPHC ERGLQPTLTN 2351 PGECRPNFTC ACRKEECKRV SPPSCPPHRL PTLRKTQCCD EYECACNCVN 2401 STVSCPLGYL ASTATNDCGC TTTTCLPDKV CVHRSTIYPV GQFWEEGCDV

	2451 CTCTDMEDAV MGLRVAQCSQ KPCEDSCRSG FTYVLHEGEC CGRCLPSACE 2501 VVTGSPRGDS QSSWKSVGSQ WASPENPCLI NECVRVKEEV FIQQRNVSCP 2551 QLEVPVCPSG FQLSCKTSAC CPSCRCERME ACMLNGTVIG PGKTVMIDVC 2601 TTCRCMVQVG VISGFKLECR KTTCNPCPLG YKEENNTGEC CGRCLPTACT 2651 IQLRGGQIMT LKRDETLQDG CDTHFCKVNE RGEYFWEKRV TGCPPFDEHK 2701 CLAEGGKIMK IPGTCCDTCE EPECNDITAR LQYVKVGSCK SEVEVDIHYC 2751 QGKCASKAMY SIDINDVQDQ CSCCSPTRTE PMQVALHCTN GSVVYHEVLN 2801 AMECKCSPRK CSK
	核苷酸序列
全长 VWF (SEQ ID NO: 1)	ATGATTCCTG CCAGATTTGC CGGGGTGCTG CTTGCTCTGG CCCTCATTTT GCCAGGGACC CTTTGTGCAG AAGGAACTCG CGGCAGGTCA TCCACGGCCC TACTAAGGAC GGTCTAAACG GCCCCACGAC GAACGAGACC GGGAGTAAAA CGGTCCCTGG GAAACACGTC TTCCTTGAGC GCCGTCCAGT AGGTGCCGGG GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG TACAGCTTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA CTACGTCGGA AAAGCCTTCA CTGAAGCAGT TGTGGAAACT ACCCTCGTAC ATGTCGAAAC GCCCTATGAC GTCAATGGAG GACCGTCCCC CGACGGTCTT ACGCTCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC TCTCCGTGTA TCTTGGGGAA TTTTTTGACA TCCATTTGTT TGTCAATGGT TGCGAGGAAG AGCTAATAAC CCCTGAAGGT CTTACCGTTC TCTCACTCGG AGAGGCACAT AGAACCCCTT AAAAAACTGT AGGTAAACAA ACAGTTACCA ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCCT TGGCACTGTG TCCCCCTGGT TTCTCAGAGG TACGGGATAC GGAGGTTTCC CGACATAGAT CTTTGACTCC GACCCATGAT GTTCGACAGG CCACTCCGGA ATGGCTTTGT GGCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTGGCA ACTTTAACAT TACCGAAACA CCGGTCCTAG CTACCGTCGC CGTTGAAAGT TCAGGACGAC AGTCTGTCTA TGAAGTTGTT CTGGACGCCC GACACACCGT TGAAATTGTA CTTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT GAAACGACTT CTACTGAAAT ACTGGGTTCT TCCCTGGAAC TGGAGCCTGG GAATACTGAA ACGGTTGAGT ACCCGAGACT CGTCACCTCT TGTCACCACA GAACGGGCAT CTCCTCCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT CTTGCCCGTA GAGGAGGGTC GTCGAGTACG TTGTAGAGGA GACCCCTTTA CGTCTTCCCG GACACCCTCG TCACGGTCGA AGACTTCTCG TGGAGCCACA TTGCCCGCTG CCACCCTCTG GTGGACCCCG AGCCTTTTGT GGCCCTGTGT GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CCTGCCCTGC AACGGGCGAC GGTGGGAGAC CACCTGGGGC TCGGAAAACA CCGGGACACA CTCTTCTGAA ACACACTCAC ACGACCCCCC GACCTCACGC GGACGGGACG CCTCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAATG GTGCTGTACG GCTGGACCGA CCACAGCGCG TGCAGCCCAG TGTGCCCTGC TGGTATGGAG GGAGGACCTC ATGCGGGCCT GGACACGGGT CCTCCCTTAC CACGACATGC CGACCTGGCT GGTGTCGCGC ACGTCGGGTC ACACGGGACG ACCATACCTC TATAGGCAGT GTGTGTCCCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT CAATGAAATG TGTCAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG ATATCCGTCA CACACAGGGG AACGCGGTCC TGGACGGTCT CGGACGTGTA GTTACTTTAC ACAGTCCTCG CTACGCACCT ACCGACGTCG ACGGGACTCC GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGTCCCTGC GTGCATTCCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG CTGTCGAGGA CCTACTTCCG GAGACGCACC TCTCGTGGCT CACAGGGACG CACGTAAGGC CTTTCGCGAT GGGAGGGCCG TGGAGGGAGA GAGCTCTGAC CAACACCTGC ATTTGCCGAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT GTCCAGGGGA GTGCCTTGTC ACTGGTCAAT CCCACTTCAA GAGCTTTGAC

[0493]

[0494]

GTTGTGGACG TAAACGGCTT TGTCGGTCAC CTAGACGTCG TTACTTCTTA
CAGGTCCCCT CACGGAACAG TGACCAGTTA GGGTGAAGTT CTCGAAACTG
AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGGA
TTGCCAGGAC CACTCCTTCT CCATTGTCAT TGAGACTGTC CAGTGTGCTG
TTGTCTATGA AGTGGAAGTC ACCCTAGACG GTCATGGACG ACCGGGCCCT
AACGGTCCTG GTGAGGAAGA GGTAACAGTA ACTCTGACAG GTCACACGAC
ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGTCCG GCTGCCTGGC
CTGCACAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA
TACTGGCGCT GCGACACACG TGGGCGAGGC AGTGGCAGGC CGACGGACCG
GACGTGTTGT CGGAACACTT TGACTTCGTA CCCCGTCCTC AACGGTACCT
TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC
ATACAGTGAC GGCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG
ACCGGTCCTG TAGGTCGAGG GGGAGGACTT TCCACTGGAG GCGTAGGTCG
TATGTCACTG CCGGAGGCAC GCGGAGTCGA TGCCCCTCCT GGACGTCTAC
GACTGGGATG GCCGCGGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC
CGGGAAGACC TGCGGCCTGT GTGGGAATTA CAATGGCAAC CAGGGCGACG
CTGACCCTAC CGGCGCCCTC CGACGACCAC TTCGACAGGG GGCAGATACG
GCCCTTCTGG ACGCCGGACA CACCCTTAAT GTTACCGTTG GTCCCGCTGC
ACTTCCTTAC CCCCTCTGGG CTGGCRGAGC CCCGGGTGGA GGACTTCGGG
AACGCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG
TGAAGGAATG GGGGAGACCC GACCGYCTCG GGGCCCACCT CCTGAAGCCC
TTGCGGACCT TCGACGTGCC CCTGACGGTC CTGGACGTCT TCGTCGTGTC
CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGGTTCTCC GAGGAGGCGT
GCGCGGTCCT GACGTCCCCC ACATTCGAGG CCTGCCATCG TGCCGTCAGC
GCTAGGGACG CGGGAGTTGG GCGCGTACTG GTCCAAGAGG CTCCTCCGCA
CGCGCCAGGA CTGCAGGGGG TGTAAGCTCC GGACGGTAGC ACGGCAGTCG
CCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACGTGTGCT CCTGCTCGGA
CGGCCGCGAG TGCCTGTGCG GCGCCCTGGC CAGCTATGCC GCGGCCTGCG
GGCGACGGGA TGGACGCCTT GACGGCGATG CTGCACACGA GGACGAGCCT
GCCGGCGCTC ACGGACACGC CGCGGGACCG GTCGATACGG CGCCGGACGC
CGGGGAGAGG CGTGCGCGTC GCGTGGCGCG AGCCAGGCCG CTGTGAGCTG
AACTGCCCGA AAGGCCAGGT GTACCTGCAG TGCGGGACCC CCTGCAACCT
GCCCCTCTCC GCACGCGCAG CGCACCGCGC TCGGTCCGGC GACACTCGAC
TTGACGGGCT TTCCGGTCCA CATGGACGTC ACGCCCTGGG GGACGTTGGA
GACCTGCCGC TCTCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCTGCC
TGGAGGGCTG CTTCTGCCCC CCAGGGCTCT ACATGGATGA GAGGGGGGAC
CTGGACGGCG AGAGAGAGAA TGGGCCTACT CCTTACGTTA CTCCGGACGG
ACCTCCCGAC GAAGACGGGG GGTCCCGAGA TGTACCTACT CTCCCCCCTG
TGCGTGCCCA AGGCCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA
GCCAGAAGAC ATCTTCTCAG ACCATCACAC CATGTGCTAC TGTGAGGATG
ACGCACGGGT TCCGGGTCAC GGGGACAATG ATACTGCCAC TCTAGAAGGT
CGGTCTTCTG TAGAAGAGTC TGGTAGTGTG GTACACGATG ACACTCCTAC
GCTTCATGCA CTGTACCATG AGTGGAGTCC CCGGAAGCTT GCTGCCTGAC
GCTGTCCTCA GCAGTCCCCT GTCTCATCGC AGCAAAAGGA GCCTATCCTG
CGAAGTACGT GACATGGTAC TCACCTCAGG GGCCTTCGAA CGACGGACTG
CGACAGGAGT CGTCAGGGGA CAGAGTAGCG TCGTTTTCCT CGGATAGGAC
TCGGCCCCCC ATGGTCAAGC TGGTGTGTCC CGCTGACAAC CTGCGGGCTG
AAGGGCTCGA GTGTACCAAA ACGTGCCAGA ACTATGACCT GGAGTGCATG
AGCCGGGGGG TACCAGTTCG ACCACACAGG GCGACTGTTG GACGCCCGAC
TTCCCGAGCT CACATGGTTT TGCACGGTCT TGATACTGGA CCTCACGTAC
AGCATGGGCT GTGTCTCTGG CTGCCTCTGC CCCCCGGGCA TGGTCCGGCA
TGAGAACAGA TGTGTGGCCC TGGAAAGGTG TCCCTGCTTC CATCAGGGCA
TCGTACCCGA CACAGAGACC GACGGAGACG GGGGGCCCGT ACCAGGCCGT
ACTCTTGTCT ACACACCGGG ACCTTTCCAC AGGGACGAAG GTAGTCCCGT
AGGAGTATGC CCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC
TGTCGGGACC GGAAGTGGAA CTGCACAGAC CATGTGTGTG ATGCCACGTG
TCCTCATACG GGGACCTCTT TGTCACTTCT AACCGACGTT GTGAACACAG
ACAGCCCTGG CCTTCACCTT GACGTGTCTG GTACACACAC TACGGTGCAC
CTCCACGATC GGCATGGCCC ACTACCTCAC CTTCGACGGG CTCAAATACC
TGTTCCCCGG GGAGTGCCAG TACGTTCTGG TGCAGGATTA CTGCGGCAGT
GAGGTGCTAG CCGTACCGGG TGATGGAGTG GAAGCTGCCC GAGTTTATGG
ACAAGGGGCC CCTCACGGTC ATGCAAGACC ACGTCCTAAT GACGCCGTCA
AACCCTGGGA CCTTTCGGAT CCTAGTGGGG AATAAGGGAT GCAGCCACCC
CTCAGTGAAA TGCAAGAAAC GGGTCACCAT CCTGGTGGAG GGAGGAGAGA
TTGGGACCCT GGAAAGCCTA GGATCACCCC TTATTCCCTA CGTCGGTGGG

[0495]

GAGTCACTTT ACGTTCTTTG CCCAGTGGTA GGACCACCTC CCTCCTCTCT
TTGAGCTGTT TGACGGGGAG GTGAATGTGA AGAGGCCCAT GAAGGATGAG
ACTCACTTTG AGGTGGTGGA GTCTGGCCGG TACATCATTC TGCTGCTGGG
AACTCGACAA ACTGCCCCTC CACTTACACT TCTCCGGGTA CTTCCTACTC
TGAGTGAAAC TCCACCACCT CAGACCGGCC ATGTAGTAAG ACGACGACCC
CAAAGCCCTC TCCGTGGTCT GGGACCGCCA CCTGAGCATC TCCGTGGTCC
TGAAGCAGAC ATACCAGGAG AAAGTGTGTG GCCTGTGTGG GAATTTTGAT
GTTTCGGGAG AGGCACCAGA CCCTGGCGGT GGACTCGTAG AGGCACCAGG
ACTTCGTCTG TATGGTCCTC TTTCACACAC CGGACACACC CTTAAAACTA
GGCATCCAGA ACAATGACCT CACCAGCAGC AACCTCCAAG TGGAGGAAGA
CCCTGTGGAC TTTGGGAACT CCTGGAAAGT GAGCTCGCAG TGTGCTGACA
CCGTAGGTCT TGTTACTGGA GTGGTCGTCG TTGGAGGTTC ACCTCCTTCT
GGGACACCTG AAACCCTTGA GGACCTTTCA CTCGAGCGTC ACACGACTGT
CCAGAAAAGT GCCTCTGGAC TCATCCCCTG CCACCTGCCA TAACAACATC
ATGAAGCAGA CGATGGTGGA TTCCTCCTGT AGAATCCTTA CCAGTGACGT
GGTCTTTTCA CGGAGACCTG AGTAGGGGAC GGTGGACGGT ATTGTTGTAG
TACTTCGTCT GCTACCACCT AAGGAGGACA TCTTAGGAAT GGTCACTGCA
CTTCCAGGAC TGCAACAAGC TGGTGGACCC CGAGCCATAT CTGGATGTCT
GCATTTACGA CACCTGCTCC TGTGAGTCCA TTGGGGACTG CGCCTGCTTC
GAAGGTCCTG ACGTTGTTCG ACCACCTGGG GCTCGGTATA GACCTACAGA
CGTAAATGCT GTGGACGAGG ACACTCAGGT AACCCCTGAC GCGGACGAAG
TGCGACACCA TTGCTGCCTA TGCCCACGTG TGTGCCCAGC ATGGCAAGGT
GGTGACCTGG AGGACGGCCA CATTGTGCCC CCAGAGCTGC GAGGAGAGGA
ACGCTGTGGT AACGACGGAT ACGGGTGCAC ACACGGGTCG TACCGTTCCA
CCACTGGACC TCCTGCCGGT GTAACACGGG GGTCTCGACG CTCCTCTCCT
ATCTCCGGGA GAACGGGTAT GAGTGTGAGT GGCGCTATAA CAGCTGTGCA
CCTGCCTGTC AAGTCACGTG TCAGCACCCT GAGCCACTGG CCTGCCCTGT
TAGAGGCCCT CTTGCCCATA CTCACACTCA CCGCGATATT GTCGACACGT
GGACGGACAG TTCAGTGCAC AGTCGTGGGA CTCGGTGACC GGACGGGACA
GCAGTGTGTG GAGGGCTGCC ATGCCCACTG CCCTCCAGGG AAAATCCTGG
ATGAGCTTTT GCAGACCTGC GTTGACCCTG AAGACTGTCC AGTGTGTGAG
CGTCACACAC CTCCCGACGG TACGGGTGAC GGGAGGTCCC TTTTAGGACC
TACTCGAAAA CGTCTGGACG CAACTGGGAC TTCTGACAGG TCACACACTC
GTGGCTGGCC GGCGTTTTGC CTCAGGAAAG AAAGTCACCT TGAATCCCAG
TGACCCTGAG CACTGCCAGA TTTGCCACTG TGATGTTGTC AACCTCACCT
CACCGACCGG CCGCAAAACG GAGTCCTTTC TTTCAGTGGA ACTTAGGGTC
ACTGGGACTC GTGACGGTCT AAACGGTGAC ACTACAACAG TTGGAGTGGA
GTGAAGCCTG CCAGGAGCCG GGAGGCCTGG TGGTGCCTCC CACAGATGCC
CCGGTGAGCC CCACCACTCT GTATGTGGAG GACATCTCGG AACCGCCGTT
CACTTCGGAC GGTCCTCGGC CCTCCGGACC ACCACGGAGG GTGTCTACGG
GGCCACTCGG GGTGGTGAGA CATACACCTC CTGTAGAGCC TTGGCGGCAA
GCACGATTTC TACTGCAGCA GGCTACTGGA CCTGGTCTTC CTGCTGGATG
GCTCCTCCAG GCTGTCCGAG GCTGAGTTTG AAGTGCTGAA GGCCTTTGTG
CGTGCTAAAG ATGACGTCGT CCGATGACCT GGACCAGAAG GACGACCTAC
CGAGGAGGTC CGACAGGCTC CGACTCAAAC TTCACGACTT CCGGAAACAC
GTGGACATGA TGGAGCGGCT GCGCATCTCC CAGAAGTGGG TCCGCGTGGC
CGTGGTGGAG TACCACGACG GCTCCCACGC CTACATCGGG CTCAAGGACC
CACCTGTACT ACCTCGCCGA CGCGTAGAGG GTCTTCACCC AGGCGCACCG
GCACCACCTC ATGGTGCTGC CGAGGGTGCG GATGTAGCCC GAGTTCCTGG
GGAAGCGACC GTCAGAGCTG CGGCGCATTG CCAGCCAGGT GAAGTATGCG
GGCAGCCAGG TGGCCTCCAC CAGCGAGGTC TTGAAATACA CACTGTTCCA
CCTTCGCTGG CAGTCTCGAC GCCGCGTAAC GGTCGGTCCA CTTCATACGC
CCGTCGGTCC ACCGGAGGTG GTCGCTCCAG AACTTTATGT GTGACAAGGT
AATCTTCAGC AAGATCGACC GCCCTGAAGC CTCCCGCATC GCCCTGCTCC
TGATGGCCAG CCAGGAGCCC CAACGGATGT CCCGGAACTT TGTCCGCTAC
TTAGAAGTCG TTCTAGCTGG CGGGACTTCG GAGGGCGTAG CGGGACGAGG
ACTACCGGTC GGTCCTCGGG GTTGCCTACA GGGCCTTGAA ACAGGCGATG
GTCCAGGGCC TGAAGAAGAA GAAGGTCATT GTGATCCCGG TGGGCATTGG
GCCCCATGCC AACCTCAAGC AGATCCGCCT CATCGAGAAG CAGGCCCCTG
CAGGTCCCGG ACTTCTTCTT CTTCCAGTAA CACTAGGGCC ACCCGTAACC
CGGGGTACGG TTGGAGTTCG TCTAGGCGGA GTAGCTCTTC GTCCGGGGAC
AGAACAAGGC CTTCGTGCTG AGCAGTGTGG ATGAGCTGGA GCAGCAAAGG
GACGAGATCG TTAGCTACCT CTGTGACCTT GCCCCTGAAG CCCCTCCTCC
TCTTGTTCCG GAAGCACGAC TCGTCACACC TACTCGACCT CGTCGTTTCC
CTGCTCTAGC AATCGATGGA GACACTGGAA CGGGGACTTC GGGGAGGAGG

[0496]

TACTCTGCCC CCCGACATGG CACAAGTCAC TGTGGGCCCG GGGCTCTTGG
GGGTTTCGAC CCTGGGGCCC AAGAGGAACT CCATGGTTCT GGATGTGGCG
ATGAGACGGG GGGCTGTACC GTGTTCAGTG ACACCCGGGC CCCGAGAACC
CCCAAAGCTG GGACCCCGGG TTCTCCTTGA GGTACCAAGA CCTACACCGC
TTCGTCCTGG AAGGATCGGA CAAAATTGGT GAAGCCGACT TCAACAGGAG
CAAGGAGTTC ATGGAGGAGG TGATTCAGCG GATGGATGTG GGCCAGGACA
AAGCAGGACC TTCCTAGCCT GTTTTAACCA CTTCGGCTGA AGTTGTCCTC
GTTCCTCAAG TACCTCCTCC ACTAAGTCGC CTACCTACAC CCGGTCCTGT
GCATCCACGT CACGGTGCTG CAGTACTCCT ACATGGTGAC CGTGGAGTAC
CCCTTCAGCG AGGCACAGTC CAAAGGGGAC ATCCTGCAGC GGGTGCGAGA
CGTAGGTGCA GTGCCACGAC GTCATGAGGA TGTACCACTG GCACCTCATG
GGGAAGTCGC TCCGTGTCAG GTTTCCCCTG TAGGACGTCG CCCACGCTCT
GATCCGCTAC CAGGGCGGCA ACAGGACCAA CACTGGGCTG GCCCTGCGGT
ACCTCTCTGA CCACAGCTTC TTGGTCAGCC AGGGTGACCG GGAGCAGGCG
CTAGGCGATG GTCCCGCCGT TGTCCTGGTT GTGACCCGAC CGGGACGCCA
TGGAGAGACT GGTGTCGAAG AACCAGTCGG TCCCACTGGC CCTCGTCCGC
CCCAACCTGG TCTACATGGT CACCGGAAAT CCTGCCTCTG ATGAGATCAA
GAGGCTGCCT GGAGACATCC AGGTGGTGCC CATTGGAGTG GGCCCTAATG
GGGTTGGACC AGATGTACCA GTGGCCTTTA GGACGGAGAC TACTCTAGTT
CTCCGACGGA CCTCTGTAGG TCCACCACGG GTAACCTCAC CCGGGATTAC
CCAACGTGCA GGAGCTGGAG AGGATTGGCT GGCCCAATGC CCCTATCCTC
ATCCAGGACT TTGAGACGCT CCCCCGAGAG GCTCCTGACC TGGTGCTGCA
GGTTGCACGT CCTCGACCTC TCCTAACCGA CCGGGTTACG GGGATAGGAG
TAGGTCCTGA AACTCTGCGA GGGGGCTCTC CGAGGACTGG ACCACGACGT
GAGGTGCTGC TCCGGAGAGG GGCTGCAGAT CCCCACCCTC TCCCCTGCAC
CTGACTGCAG CCAGCCCCTG GACGTGATCC TTCTCCTGGA TGGCTCCTCC
CTCCACGACG AGGCCTCTCC CCGACGTCTA GGGGTGGGAG AGGGGACGTG
GACTGACGTC GGTCGGGGAC CTGCACTAGG AAGAGGACCT ACCGAGGAGG
AGTTTCCCAG CTTCTTATTT TGATGAAATG AAGAGTTTCG CCAAGGCTTT
CATTTCAAAA GCCAATATAG GGCCTCGTCT CACTCAGGTG TCAGTGCTGC
TCAAAGGGTC GAAGAATAAA ACTACTTTAC TTCTCAAAGC GGTTCCGAAA
GTAAAGTTTT CGGTTATATC CCGGAGCAGA GTGAGTCCAC AGTCACGACG
AGTATGGAAG CATCACCACC ATTGACGTGC CATGGAACGT GGTCCCGGAG
AAAGCCCATT TGCTGAGCCT TGTGGACGTC ATGCAGCGGG AGGGAGGCCC
TCATACCTTC GTAGTGGTGG TAACTGCACG GTACCTTGCA CCAGGGCCTC
TTTCGGGTAA ACGACTCGGA ACACCTGCAG TACGTCGCCC TCCCTCCGGG
CAGCCAAATC GGGGATGCCT TGGGCTTTGC TGTGCGATAC TTGACTTCAG
AAATGCATGG TGCCAGGCCG GGAGCCTCAA AGGCGGTGGT CATCCTGGTC
GTCGGTTTAG CCCCTACGGA ACCCGAAACG ACACGCTATG AACTGAAGTC
TTTACGTACC ACGGTCCGGC CCTCGGAGTT TCCGCCACCA GTAGGACCAG
ACGGACGTCT CTGTGGATTC AGTGGATGCA GCAGCTGATG CCGCCAGGTC
CAACAGAGTG ACAGTGTTCC CTATTGGAAT TGGAGATCGC TACGATGCAG
TGCCTGCAGA GACACCTAAG TCACCTACGT CGTCGACTAC GGCGGTCCAG
GTTGTCTCAC TGTCACAAGG GATAACCTTA ACCTCTAGCG ATGCTACGTC
CCCAGCTACG GATCTTGGCA GGCCCAGCAG GCGACTCCAA CGTGGTGAAG
CTCCAGCGAA TCGAAGACCT CCCTACCATG GTCACCTTGG GCAATTCCTT
GGGTCGATGC CTAGAACCGT CCGGGTCGTC CGCTGAGGTT GCACCACTTC
GAGGTCGCTT AGCTTCTGGA GGGATGGTAC CAGTGGAACC CGTTAAGGAA
CCTCCACAAA CTGTGCTCTG GATTTGTTAG GATTTGCATG GATGAGGATG
GGAATGAGAA GAGGCCCGGG GACGTCTGGA CCTTGCCAGA CCAGTGCCAC
GGAGGTGTTT GACACGAGAC CTAAACAATC CTAAACGTAC CTACTCCTAC
CCTTACTCTT CTCCGGGCCC CTGCAGACCT GGAACGGTCT GGTCACGGTG
ACCGTGACTT GCCAGCCAGA TGGCCAGACC TTGCTGAAGA GTCATCGGGT
CAACTGTGAC CGGGGGCTGA GGCCTTCGTG CCCTAACAGC CAGTCCCCTG
TGGCACTGAA CGGTCGGTCT ACCGGTCTGG AACGACTTCT CAGTAGCCCA
GTTGACACTG GCCCCCGACT CCGGAAGCAC GGGATTGTCG GTCAGGGGAC
TTAAAGTGGA AGAGACCTGT GGCTGCCGCT GGACCTGCCC CTGYGTGTGC
ACAGGCAGCT CCACTCGGCA CATCGTGACC TTTGATGGGC AGAATTTCAA
AATTTCACCT TCTCTGGACA CCGACGGCGA CCTGGACGGG GACRCACACG
TGTCCGTCGA GGTGAGCCGT GTAGCACTGG AAACTACCCG TCTTAAAGTT
GCTGACTGGC AGCTGTTCTT ATGTCCTATT TCAAAACAAG GAGCAGGACC
TGGAGGTGAT TCTCCATAAT GGTGCCTGCA GCCCTGGAGC AAGGCAGGGC
CGACTGACCG TCGACAAGAA TACAGGATAA AGTTTTGTTC CTCGTCCTGG
ACCTCCACTA AGAGGTATTA CCACGGACGT CGGGACCTCG TTCCGTCCCG
TGCATGAAAT CCATCGAGGT GAAGCACAGT GCCCTCTCCG TCGAGSTGCA

[0497]

CAGTGACATG GAGGTGACGG TGAATGGGAG ACTGGTCTCT GTTCCTTACG
ACGTACTTTA GGTAGCTCCA CTTCGTGTCA CGGGAGAGGC AGCTCSACGT
GTCACTGTAC CTCCACTGCC ACTTACCCTC TGACCAGAGA CAAGGAATGC
TGGGTGGGAA CATGGAAGTC AACGTTTATG GTGCCATCAT GCATGAGGTC
AGATTCAATC ACCTTGGTCA CATCTTCACA TTCACTCCAC AAAACAATGA
ACCCACCCTT GTACCTTCAG TTGCAAATAC CACGGTAGTA CGTACTCCAG
TCTAAGTTAG TGGAACCAGT GTAGAAGTGT AAGTGAGGTG TTTTGTTACT
GTTCCAACTG CAGCTCAGCC CCAAGACTTT TGCTTCAAAG ACGTATGGTC
TGTGTGGGAT CTGTGATGAG AACGGAGCCA ATGACTTCAT GCTGAGGGAT
CAAGGTTGAC GTCGAGTCGG GGTTCTGAAA ACGAAGTTTC TGCATACCAG
ACACACCCTA GACACTACTC TTGCCTCGGT TACTGAAGTA CGACTCCCTA
GGCACAGTCA CCACAGACTG GAAAACACTT GTTCAGGAAT GGACTGTGCA
GCGGCCAGGG CAGACGTGCC AGCCCATCCT GGAGGAGCAG TGTCTTGTCC
CCGTGTCAGT GGTGTCTGAC CTTTTGTGAA CAAGTCCTTA CCTGACACGT
CGCCGGTCCC GTCTGCACGG TCGGGTAGGA CCTCCTCGTC ACAGAACAGG
CCGACAGCTC CCACTGCCAG GTCCTCCTCT TACCACTGTT TGCTGAATGC
CACAAGGTCC TGGCTCCAGC CACATTCTAT GCCATCTGCC AGCAGGACAG
GGCTGTCGAG GGTGACGGTC CAGGAGGAGA ATGGTGACAA ACGACTTACG
GTGTTCCAGG ACCGAGGTCG GTGTAAGATA CGGTAGACGG TCGTCCTGTC
TTGCCACCAG GAGCAAGTGT GTGAGGTGAT CGCCTCTTAT GCCCACCTCT
GTCGGACCAA CGGGGTCTGC GTTGACTGGA GGACACCTGA TTTCTGTGCT
AACGGTGGTC CTCGTTCACA CACTCCACTA GCGGAGAATA CGGGTGGAGA
CAGCCTGGTT GCCCCAGACG CAACTGACCT CCTGTGGACT AAAGACACGA
ATGTCATGCC CACCATCTCT GGTCTACAAC CACTGTGAGC ATGGCTGTCC
CCGGCACTGT GATGGCAACG TGAGCTCCTG TGGGGACCAT CCCTCCGAAG
TACAGTACGG GTGGTAGAGA CCAGATGTTG GTGACACTCG TACCGACAGG
GGCCGTGACA CTACCGTTGC ACTCGAGGAC ACCCCTGGTA GGGAGGCTTC
GCTGTTTCTG CCCTCCAGAT AAAGTCATGT TGGAAGGCAG CTGTGTCCCT
GAAGAGGCCT GCACTCAGTG CATTGGTGAG GATGGAGTCC AGCACCAGTT
CGACAAAGAC GGGAGGTCTA TTTCAGTACA ACCTTCCGTC GACACAGGGA
CTTCTCCGGA CGTGAGTCAC GTAACCACTC CTACCTCAGG TCGTGGTCAA
CCTGGAAGCC TGGGTCCCGG ACCACCAGCC CTGTCAGATC TGCACATGCC
TCAGCGGGCG GAAGGTCAAC TGCACAACGC AGCCCTGCCC CACGGCCAAA
GGACCTTCGG ACCCAGGGCC TGGTGGTCGG GACAGTCTAG ACGTGTACGG
AGTCGCCCGC CTTCCAGTTG ACGTGTTGCG TCGGGACGGG GTGCCGGTTT
GCTCCCACGT GTGGCCTGTG TGAAGTAGCC CGCCTCCGCC AGAATGCAGA
CCAGTGCTGC CCCGAGTATG AGTGTGTGTG TGACCCAGTG AGCTGTGACC
CGAGGGTGCA CACCGGACAC ACTTCATCGG GCGGAGGCGG TCTTACGTCT
GGTCACGACG GGGCTCATAC TCACACACAC ACTGGGTCAC TCGACACTGG
TGCCCCCAGT GCCTCACTGT GAACGTGGCC TCCAGCCCAC ACTGACCAAC
CCTGGCGAGT GCAGACCCAA CTTCACCTGC GCCTGCAGGA AGGAGGAGTG
ACGGGGGTCA CGGAGTGACA CTTGCACCGG AGGTCGGGTG TGACTGGTTG
GGACCGCTCA CGTCTGGGTT GAAGTGGACG CGGACGTCCT TCCTCCTCAC
CAAAAGAGTG TCCCCACCCT CCTGCCCCCC GCACCGTTTG CCCACCCTTC
GGAAGACCCA GTGCTGTGAT GAGTATGAGT GTGCCTGCAA CTGTGTCAAC
GTTTTCTCAC AGGGGTGGGA GGACGGGGGG CGTGGCAAAC GGGTGGGAAG
CCTTCTGGGT CACGACACTA CTCATACTCA CACGGACGTT GACACAGTTG
TCCACAGTGA GCTGTCCCCT TGGGTACTTG GCCTCAACCG CCACCAATGA
CTGTGGCTGT ACCACAACCA CCTGCCTTCC CGACAAGGTG TGTGTCCACC
AGGTGTCACT CGACAGGGGA ACCCATGAAC CGGAGTTGGC GGTGGTTACT
GACACCGACA TGGTGTTGGT GGACGGAAGG GCTGTTCCAC ACACAGGTGG
GAAGCACCAT CTACCCTGTG GGCCAGTTCT GGGAGGAGGG CTGCGATGTG
TGCACCTGCA CCGACATGGA GGATGCCGTG ATGGGCCTCC GCGTGGCCCA
CTTCGTGGTA GATGGGACAC CCGGTCAAGA CCCTCCTCCC GACGCTACAC
ACGTGGACGT GGCTGTACCT CCTACGGCAC TACCCGGAGG CGCACCGGGT
GTGCTCCCAG AAGCCCTGTG AGGACAGCTG TCGGTCGGGC TTCACTTACG
TTCTGCATGA AGGCGAGTGC TGTGGAAGGT GCCTGCCATC TGCCTGTGAG
CACGAGGGTC TTCGGGACAC TCCTGTCGAC AGCCAGCCCG AAGTGAATGC
AAGACGTACT TCCGCTCACG ACACCTTCCA CGGACGGTAG ACGGACACTC
GTGGTGACTG GCTCACCGCG GGGGGACTCC CAGTCTTCCT GGAAGAGTGT
CGGCTCCCAG TGGGCCTCCC CGGAGAACCC CTGCCTCATC AATGAGTGTG
CACCACTGAC CGAGTGGCGC CCCCCTGAGG GTCAGAAGGA CCTTCTCACA
GCCGAGGGTC ACCCGGAGGG GCCTCTTGGG GACGGAGTAG TTACTCACAC
TCCGAGTGAA GGAGGAGGTC TTTATACAAC AAAGGAACGT CTCCTGCCCC
CAGCTGGAGG TCCCTGTCTG CCCCTCGGGC TTTCAGCTGA GCTGTAAGAC

[0498]

AGGCTCACTT CCTCCTCCAG AAATATGTTG TTTCCTTGCA GAGGACGGGG
GTCGACCTCC AGGGACAGAC GGGGAGCCCG AAAGTCGACT CGACATTCTG
CTCAGCGTGC TGCCCAAGCT GTCGCTGTGA GCGCATGGAG GCCTGCATGC
TCAATGGCAC TGTCATTGGG CCCGGGAAGA CTGTGATGAT CGATGTGTGC
GAGTCGCACG ACGGGTTCGA CAGCGACACT CGCGTACCTC CGGACGTACG
AGTTACCGTG ACAGTAACCC GGGCCCTTCT GACACTACTA GCTACACACG
ACGACCTGCC GCTGCATGGT GCAGGTGGGG GTCATCTCTG GATTCAAGCT
GGAGTGCAGG AAGACCACCT GCAACCCCTG CCCCCTGGGT TACAAGGAAG
TGCTGGACGG CGACGTACCA CGTCCACCCC CAGTAGAGAC CTAAGTTCGA
CCTCACGTCC TTCTGGTGGA CGTTGGGGAC GGGGGACCCA ATGTTCCTTC
AAAATAACAC AGGTGAATGT TGTGGGAGAT GTTTGCCTAC GGCTTGCACC
ATTCAGCTAA GAGGAGGACA GATCATGACA CTGAAGCGTG ATGAGACGCT
TTTTATTGTG TCCACTTACA ACACCCTCTA CAAACGGATG CCGAACGTGG
TAAGTCGATT CTCCTCCTGT CTAGTACTGT GACTTCGCAC TACTCTGCGA
CCAGGATGGC TGTGATACTC ACTTCTGCAA GGTCAATGAG AGAGGAGAGT
ACTTCTGGGA GAAGAGGGTC ACAGGCTGCC CACCCTTTGA TGAACACAAG
GGTCCTACCG ACACTATGAG TGAAGACGTT CCAGTTACTC TCTCCTCTCA
TGAAGACCCT CTTCTCCCAG TGTCCGACGG GTGGGAAACT ACTTGTGTTC
TGTCTTGCTG AGGGAGGTAA AATTATGAAA ATTCCAGGCA CCTGCTGTGA
CACATGTGAG GAGCCTGAGT GCAACGACAT CACTGCCAGG CTGCAGTATG
ACAGAACGAC TCCCTCCATT TTAATACTTT TAAGGTCCGT GGACGACACT
GTGTACACTC CTCGGACTCA CGTTGCTGTA GTGACGGTCC GACGTCATAC
TCAAGGTGGG AAGCTGTAAG TCTGAAGTAG AGGTGGATAT CCACTACTGC
CAGGGCAAAT GTGCCAGCAA AGCCATGTAC TCCATTGACA TCAACGATGT
AGTTCCACCC TTCGACATTC AGACTTCATC TCCACCTATA GGTGATGACG
GTCCCGTTTA CACGGTCGTT TCGGTACATG AGGTAACTGT AGTTGCTACA
GCAGGACCAG TGCTCCTGCT GCTCTCCGAC ACGGACGGAG CCCATGCAGG
TGGCCCTGCA CTGCACCAAT GGCTCTGTTG TGTACCATGA GGTTCTCAAT
CGTCCTGGTC ACGAGGACGA CGAGAGGCTG TGCCTGCCTC GGGTACGTCC
ACCGGGACGT GACGTGGTTA CCGAGACAAC ACATGGTACT CCAAGAGTTA
GCCATGGAGT GCAAATGCTC CCCCAGGAAG TGCAGCAAGT GA

[0499] 本发明涉及血管性血友病因子(VWF)片段,所述血管性血友病因子片段包含VWF的D'结构域和D3结构域,其中VWF片段抑制内源性VWF(全长VWF)与FVIII蛋白结合。在一个实施方案中,VWF片段结合或缔合FVIII蛋白。通过结合或缔合FVIII蛋白,本发明的VWF片段保护FVIII不被蛋白酶切割而且FVIII不被活化,稳定FVIII的重链和轻链,并且阻止FVIII被清除剂受体清除。在另一个实施方案中,VWF片段结合FVIII蛋白或与之缔合,并且阻断或防止FVIII蛋白与磷脂和活化的蛋白C结合。通过阻止或抑制FVIII蛋白与内源性、全长VWF结合,本发明的VWF片段减少VWF清除受体对FVIII的清除,并且从而延长FVIII的半衰期。因此,FVIII蛋白的半衰期延长是由于缺乏VWF清除受体结合位点的VWF片段与FVIII蛋白的结合或缔合,以及通过VWF片段屏蔽或保护FVIII蛋白不被包含VWF清除受体结合位点的内源性VWF清除。结合至VWF片段或受其保护的FVIII蛋白也可允许FVIII蛋白的循环。因此,VWF片段可不为全长成熟VWF。通过消除全长VWF分子中包含的VWF清除途径受体结合位点,本发明的FVIII/VWF异源二聚体从VWF清除途径解离,这允许FVIII半衰期进一步延长。

[0500] 包含D'结构域和D3结构域的VWF片段还可包含选自由以下组成的组的VWF结构域:A1结构域、A2结构域、A3结构域、D1结构域、D2结构域、D4结构域、B1结构域、B2结构域、B3结构域、C1结构域、C2结构域、CK结构域、它们的一个或多个片段以及它们的任何组合。在一个实施方案中,VWF片段包含、基本上由或由(1)VWF的D'和D3结构域或其片段;(2)VWF的D1、D'和D3结构域或其片段;(3)VWF的D2、D'和D3结构域或其片段;(4)VWF的D1、D2、D'和D3结构域或其片段;或(5)VWF的D1、D2、D'、D3和A1结构域或其片段组成。本文所述的VWF片段不包含结合VWF清除受体的位点。在另一个实施方案中,本文所述的VWF片段不是SEQ ID NO:2第764至1274位氨基酸。本发明的VWF片段可包含连接至或融合至VWF片段的任何其它序列,但不是全长VWF。例如,本文所述的VWF片段还可包含信号肽。

[0501] 在一个实施方案中,本发明的VWF片段包含VWF的D'结构域和D3结构域,其中D'结构域与SEQ ID NO:2的第764至866位氨基酸具有至少60%、70%、80%、85%、90%、95%、

96％、97％、98％、99％或100％的同一性，其中VWF片段结合FVIII蛋白，屏蔽、抑制或阻止内源性VWF片段与FVIII蛋白结合。在另一个实施方案中，VWF片段包含VWF的D'结构域和D3结构域，其中D3结构域与SEQ ID NO:2的第867至1240位氨基酸具有至少60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％的同一性，其中VWF片段结合FVIII蛋白或抑制或阻止内源性VWF片段与FVIII蛋白结合。在一些实施方案中，本文所述的VWF片段包含、基本上由或由VWF的D'结构域和D3结构域组成，所述结构域与SEQ ID NO:2的第764至1240位氨基酸具有至少60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％的同一性，其中VWF片段结合FVIII蛋白，或抑制或阻止内源性VWF片段与FVIII蛋白结合。在其它实施方案中，VWF片段包含、基本上由或由D1、D2、D'和D3结构域组成，所述结构域与SEQ ID NO:2的第23至1240位氨基酸具有至少60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％的同一性，其中VWF片段结合FVIII蛋白，或抑制或阻止内源性VWF片段与FVIII蛋白结合。在其它实施方案中，VWF片段还包含与其可操作地连接的信号肽。

[0502] 在一些实施方案中，本发明的VWF片段基本上由或由(1)D'D3结构域、D1D'D3结构域、D2D'D3结构域或D1D2D'D3结构域以及(2)最多约10个氨基酸(如，从SEQ ID NO:2的第764至1240位氨基酸至SEQ ID NO:2的第764至1250位氨基酸的任何序列)、最多约15个氨基酸(如，从SEQ ID NO:2的第764至1240位氨基酸至SEQ ID NO:2的第764至1255位氨基酸的任何序列)、最多约20个氨基酸(如，从SEQ ID NO:2的第764至1240位氨基酸至SEQ ID NO:2的第764至1260位氨基酸的任何序列)、最多约25个氨基酸(如，从SEQ ID NO:2的第764至1240位氨基酸至SEQ ID NO:2的第764至1265位氨基酸的任何序列)或最多约30个氨基酸(如，从SEQ ID NO:2的第764至1240位氨基酸至SEQ ID NO:2的第764至1260位氨基酸的任何序列)的另外的VWF序列组成。在具体实施方案中，包含或基本上由D'结构域和D3结构域组成的VWF片段既不是SEQ ID NO:2的第764至1274位氨基酸也不是全长成熟VWF。

[0503] 在其它实施方案中，包含连接至D1D2结构域的D'D3结构域的VWF片段还包含细胞内切割位点，(如，PACE或PC5的切割位点)，以允许在表达时D1D2结构域从D'D3结构域切割。细胞内切割位点的非限制性实例在本文别处有所公开。

[0504] 在其它实施方案中，VWF片段包含D'结构域和D3结构域，但不包含选自由以下组成的组的氨基酸序列：(1)SEQ ID NO:2的第1241至2813位氨基酸、(2)SEQ ID NO:2的第1270位氨基酸至第2813位氨基酸、(3)SEQ ID NO:2的第1271位氨基酸至第2813位氨基酸、(4)SEQ ID NO:2的第1272位氨基酸至第2813位氨基酸、(5)SEQ ID NO:2的第1273位氨基酸至第2813位氨基酸以及(6)SEQ ID NO:2的第1274位氨基酸至第2813位氨基酸。

[0505] 在其它实施方案中，本发明的VWF片段包含、基本上由或由与D'结构域、D3结构域和A1结构域对应的氨基酸序列组成，其中氨基酸序列与SEQ ID NO:2的第764至1479位氨基酸具有至少60％、70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％的同一性，其中VWF结合FVIII。在具体实施方案中，VWF片段不是SEQ ID NO:2的第764至1274位氨基酸。

[0506] 在一些实施方案中，本发明的VWF片段包含D'结构域和D3结构域，但不包含选自由以下组成的组的至少一个VWF结构域：(1)A1结构域、(2)A2结构域、(3)A3结构域、(4)D4结构域、(5)B1结构域、(6)B2结构域、(7)B3结构域、(8)C1结构域、(9)C2结构域、(10)CK结构域、(11)CK结构域和C2结构域、(12)CK结构域、C2结构域和C1结构域、(13)CK结构域、C2结构域、

C1结构域、B3结构域、(14)CK结构域、C2结构域、C1结构域、B3结构域、B2结构域、(15)CK结构域、C2结构域、C1结构域、B3结构域、B2结构域和B1结构域、(16)CK结构域、C2结构域、C1结构域、B3结构域、B2结构域、B1结构域和D4结构域、(17)CK结构域、C2结构域、C1结构域、B3结构域、B2结构域、B1结构域、D4结构域和A3结构域、(18)CK结构域、C2结构域、C1结构域、B3结构域、B2结构域、B1结构域、D4结构域、A3结构域和A2结构域、(19)CK结构域、C2结构域、C1结构域、B3结构域、B2结构域、B1结构域、D4结构域、A3结构域、A2结构域和A1结构域以及(20)它们的任何组合。

[0507] 在其它实施方案中,VWF片段包含D'D3结构域和一个或多个结构域或模块。此类结构域或模块的实例包括但不限于Zhour等,Blood published online 2012年4月6日:DOI 10.1182/blood-2012-01-405134中公开的结构域和模块。例如,VWF片段可包含D'D3结构域和一个或多个选自由以下组成的组的结构域或模块:A1结构域、A2结构域、A3结构域、D4N模块、VWD4模块、C8-4模块、TIL-4模块、C1模块、C2模块、C3模块、C4模块、C5模块、C5模块、C6模块以及它们的任何组合。

[0508] 在其它实施方案中,VWF片段连接至异源部分,其中异源部分连接至VWF片段的N-末端或C-末端,或插入VWF片段中的两个氨基酸之间。例如,VWF片段中异源部分的插入位点可位于D'结构域、D3结构域或它们二者中。异源部分可以是半衰期延长因子。

[0509] 在某些实施方案中,本发明的VWF片段形成多聚体,如二聚体、三聚体、四聚体、五聚体、六聚体、七聚体或高阶多聚体。在其它实施方案中,VWF片段是仅具有一个VWF片段的单体。在一些实施方案中,本发明的VWF片段可具有一个或多个氨基酸取代、缺失、添加或修饰。在一个实施方案中,VWF片段可包括氨基酸取代、缺失、添加或修饰,使得VWF片段不能形成二硫键或形成二聚体或多聚体。在另一个实施方案中,氨基酸取代在D'结构域和D3结构域内。在具体实施方案中,本发明的VWF片段包含在于SEQ ID NO:2的第1099位残基、第1142位残基或第1099和1142位残基二者对应的残基处的至少一个氨基酸取代。至少一个氨基酸取代可以是非天然存在于野生型VWF中的任何氨基酸。例如,氨基酸取代可以是除半胱氨酸之外的任何氨基酸,如异亮氨酸、丙氨酸、亮氨酸、天冬酰胺、赖氨酸、天冬氨酸、甲硫氨酸、苯丙氨酸、谷氨酸、苏氨酸、谷氨酰胺、色氨酸、甘氨酸、缬氨酸、脯氨酸、丝氨酸、酪氨酸、精氨酸或组氨酸。在另一个实例中,氨基酸取代具有阻止或抑制VWF片段形成多聚体的一个或多个氨基酸。

[0510] 在某些实施方案中,可进一步修饰本文所用的VWF片段,以提高其与FVIII的相互作用,如提高与FVIII的结合亲合力。作为非限制性实例,VWF片段包含在与SEQ ID NO:2的第764位氨基酸对应的残基处的丝氨酸残基以及在与SEQ ID NO:2的第773位氨基酸对应的残基处的赖氨酸残基。第764和/或773位残基可有助于VWF片段与FVIII的结合亲合力。在其它实施方案中,VWF片段可具有其它修饰,如片段可以是聚乙二醇化的、糖基化的、羟乙基淀粉化的或聚唾液酸化的。

[0511] B)异源部分

[0512] 异源部分可以是异源多肽或异源非多肽部分。在某些实施方案中,异源部分是本领域已知的半衰期延长分子,并且包括多肽、非多肽部分或它们二者的组合。异源多肽部分可包括免疫球蛋白恒定区或其部分、白蛋白或其片段、白蛋白结合部分、转铁蛋白或其片段、PAS序列、HAP序列、它们的衍生物或变体或它们的任何组合。在一些实施方案中,非多肽

结合部分包括聚乙二醇(PEG)、聚唾液酸、羟乙基淀粉(HES)、其衍生物或它们的任何组合。在某些实施方案中,可存在一个、两个、三个或更多个异源部分,每个可以是相同的或不同的分子。

[0513] 1)免疫球蛋白恒定区或其部分

[0514] 免疫球蛋白恒定区由CH(恒定重)结构域表示的结构域(CH1、CH2等)组成。取决于同种型(即IgG、IgM、IgA IgD或IgE),恒定区可包含三个或四个CH结构域。一些同种型(如IgG)恒定区也包含铰链区。参见Janeway等2001,Immunobiology,Garland Publishing,N.Y.,N.Y.。

[0515] 用于产生本发明嵌合蛋白的免疫球蛋白恒定区或其部分可从多个不同的来源获得。在优选的实施方案中,免疫球蛋白恒定区或其部分来源于人免疫球蛋白。然而,应当理解免疫球蛋白恒定区或其部分可来源于另一个哺乳动物物种,包括例如啮齿类(如,小鼠、大鼠、兔、豚鼠)或非人类灵长类(如,黑猩猩、猕猴)物种的免疫球蛋白。此外,免疫球蛋白恒定区或其部分可来源于任何免疫球蛋白种类,包括IgM、IgG、IgD、IgA和IgE,和任何免疫球蛋白同种型,包括IgG1、IgG2、IgG3和IgG4。在一个实施方案中,使用人同种型IgG1。

[0516] 多种免疫球蛋白恒定区基因序列(如,人恒定区基因序列)可以可公开获取的保藏物的形式获得。可选择具有特定效应子功能(或缺乏特定效应子功能)或具有特定修饰的恒定区域序列,以减小免疫原性。已公布多个抗体和抗体编码基因的序列,并且合适的Ig恒定区序列(如,铰链、CH2和/或CH3序列或其部分)可使用本领域已知的技术来源于这些序列。然后,可改变或合成使用任何上述法获得的遗传物质,以获得本发明的多肽。还应当理解,本发明的范围涵盖恒定区DNA序列的等位基因、变体和突变。

[0517] 可例如使用聚合酶链式反应和引物克隆免疫球蛋白恒定区或其部分的序列,选择该引物以扩增所关注的结构域。为从抗体克隆免疫球蛋白恒定区或其部分的序列,可从杂交瘤、脾脏或淋巴细胞分离mRNA,逆转录为DNA,并且通过PCR扩增抗体基因。PCR扩增法在美国专利号4,683,195、4,683,202、4,800,159、4,965,188;以及例如"PCR Protocols:A Guide to Methods and Applications"Innis等编著,Academic Press,San Diego,CA(1990);Ho等1989.Gene 77:51;Horton等1993.Methods Enzymol.217:270)中详细描述。PCR可根据公布的重链和轻链DNA和氨基酸序列,通过共有恒定区引物或通过更多具体引物引发。如上所讨论,PCR也可用于分离编码抗体轻链和重链的DNA克隆。在这种情况下,可通过共有引物或较大同源探针,例如小鼠恒定区探针筛选文库。适于抗体基因扩增的多组引物是本领域已知的(如,基于纯化抗体的N-末端序列的5'引物(Benhar和Pastan.1994.Protein Engineering 7:1509)、cDNA末端的快速扩增(Ruberti,F.等1994.J.Immunol.Methods 173:33)、抗体前导序列(Larrick等1989 Biochem.Biophys.Res.Commun.160:1250))。抗体序列的克隆在Newman等提交于1995年1月25日的美国专利号5,658,570中进一步描述,该专利以引用方式并入本文。

[0518] 本文所用的免疫球蛋白恒定区可包括所有结构域和铰链区或其部分。在一个实施方案中,免疫球蛋白恒定区或其部分包含CH2结构域、CH3结构域和铰链区,即Fc区或FcRn结合伴侣。

[0519] 如本文所用,术语"Fc区"定义为对应于天然免疫球蛋白Fc区的多肽的部分,即通过其两条重链的各自Fc结构域的二聚体缔合形成。天然Fc区与另一个Fc区形成同源二聚

体。相比之下，如本文所用，术语“遗传融合的Fc区”或“单链Fc区”(scFc区)是指合成二聚体Fc区，其包含单条多肽链内遗传连接的Fc结构域(即，在单个邻接基因序列中编码)。

[0520] 在一个实施方案中，“Fc区”是指始于铰链区、恰好在木瓜蛋白酶切割位点(即，IgG中的第216位残基，重链恒定区的第一个残基为第114位)上游并且结束于抗体C-末端的单个免疫球蛋白重链的部分。因此，完整的Fc结构域包含至少铰链结构域、CH2结构域和CH3结构域。

[0521] 取决于免疫球蛋白同种型，免疫球蛋白恒定区的Fc区可包括CH2、CH3和CH4结构域以及铰链区。包含免疫球蛋白的Fc区的嵌合蛋白赋予嵌合蛋白多个所需特性，包括增加的稳定性、增加的血清半衰期(参见Capon等，1989，Nature 337:525)以及结合Fc受体例如新生儿Fc受体(FcRn)(美国专利号6,086,875、6,485,726、6,030,613、WO 03/077834、US2003-0235536A1)，这些文献和专利全文以引用方式并入本文中。

[0522] 免疫球蛋白恒定区或其部分可以是FcRn结合伴侣。FcRn在成体上皮组织中具有活性，并且在肠道内腔、肺部气道、鼻面、阴道表面、结肠和直肠表面中表达(美国专利号6,485,726)。FcRn结合伴侣是结合FcRn的免疫球蛋白的一部分。

[0523] FcRn受体已从多个哺乳动物物种包括人分离。人FcRn、猴FcRn、大鼠FcRn和小鼠FcRn的序列是已知的(Story等1994，J.Exp.Med.180:2377)。FcRn受体在相对较低pH下结合IgG(而非其它免疫球蛋白种类，例如IgA、IgM、IgD和IgE)，在内腔至浆膜方向穿过细胞主动运输IgG，然后在间质液中发现的相对较高pH下释放IgG。其在成体上皮组织(美国专利号6,485,726、6,030,613、6,086,875、WO 03/077834、US2003-0235536A1)，包括肺和肠上皮细胞(Israel等1997，Immunology 92:69)、肾脏近端小管上皮细胞(Kobayashi等2002，Am.J.Physiol.Renal Physiol.282:F358)以及鼻腔上皮细胞、阴道表面和胆管树表面中表达。

[0524] 用于本发明的FcRn结合伴侣涵盖可被FcRn受体特异性结合的分子，包括完整IgG、IgG的Fc片段以及包括FcRn受体的完整结合区的其它片段。结合FcRn受体的IgG的Fc部分的区域根据X-射线结晶学进行了描述(Burmeister等1994，Nature 372:379)。Fc与FcRn的主要接触区域靠近CH2和CH3结构域的接合处。Fc-FcRn接触均在单Ig重链内。FcRn结合伴侣包括完整IgG、IgG的Fc片段以及包括FcRn的完整结合区的IgG其它片段。主要接触位点包括CH2结构域的第248、250-257、272、285、288、290-291、308-311和314位氨基酸残基和CH3结构域的第385-387、428和433-436位氨基酸残基。对免疫球蛋白或免疫球蛋白片段或区域的氨基酸编号的参考均根据Kabat等1991，Sequences of Proteins of Immunological Interest，U.S.Department of Public Health，Bethesda，Md进行。

[0525] 结合至FcRn的Fc区或FcRn结合伴侣可通过FcRn有效穿梭跨过上皮屏障，从而提供全身施用所需治疗分子的无创方式。另外，包含Fc区或FcRn结合伴侣的融合蛋白被表达FcRn的细胞内吞。但除被标记为降解之外，这些融合蛋白被回收再次进入循环，从而增加这些蛋白质的体内半衰期。在某些实施方案中，免疫球蛋白恒定区的部分是通常通过二硫键和其它非特异性相互作用与另一个Fc区或另一个FcRn结合伴侣缔合以形成二聚体和高阶多聚体的Fc区或FcRn结合伴侣。

[0526] 两种FcRn受体可结合单个Fc分子。晶体学数据表明每个FcRn分子结合Fc同源二聚体的单个多肽。在一个实施方案中，使FcRn结合伴侣如IgG的Fc片段连接至生物活性分子提

供了口服、口腔含化、舌下、直肠、阴道递送，作为气雾剂鼻腔施用或通过肺部途径，或通过眼部途径递送生物活性分子的方式。在另一个实施方案中，嵌合蛋白可侵入性施用，如皮下、静脉内施用。

[0527] FcRn结合伴侣区是可被FcRn受体特异性结合，随后通过Fc区的FcRn受体主动运输的分子或其部分。特异性结合是指形成在生理条件下相对稳定的复合物的两个分子。与通常具有低亲合力和中等至高容量的非特异结合不同，特异性结合的特征在于高亲合力和低至中等容量。通常，当亲合力常数KA大于10^6M^{-1}，或大于10^8M^{-1}时，结合被认为是特异性的。如有需要，可通过改变结合条件来减少非特异性结合，而基本上不影响特异性结合。技术人员可使用常规技术优化适当的结合条件，例如分子的浓度、溶液的离子强度、温度、结合时间、阻断剂(如，血清白蛋白、牛奶酪蛋白)的浓度等。

[0528] 在某些实施方案中，本发明的嵌合蛋白包含一个或多个截短Fc区，但足以将Fc受体(FcR)结合特性赋予Fc区。例如，结合FcRn的Fc区的部分(即，FcRn结合部分)包含IgG1的约第282-438位氨基酸，EU编号，主要接触位点是CH2结构域的第248、250-257、272、285、288、290-291、308-311和314位氨基酸和CH3结构域的第385-387、428和433-436位氨基酸残基。因此，本发明的Fc区可包含或由FcRn结合部分组成。FcRn结合部分可来源于任何同种型，包括IgG1、IgG2、IgG3和IgG4的重链。在一个实施方案中，使用来自人同种型IgG1的抗体的FcRn结合部分。在另一个实施方案中，使用来自人同种型IgG4的抗体的FcRn结合部分。

[0529] 在另一个实施方案中，“Fc区”包括Fc结构域或来源于Fc结构域的氨基酸序列。在某些实施方案中，Fc区包含以下中的至少一者：铰链(如，上、中和/或下铰链区)结构域(根据EU编号抗体Fc区的约第216-230位氨基酸)、CH2结构域(根据EU编号抗体Fc区的约第231-340位氨基酸)、CH3结构域(根据EU编号抗体Fc区的约第341-438位氨基酸)、CH4结构域中，或它们的变体、部分或片段。在其它实施方案中，Fc区包含完整Fc结构域(即，铰链结构域、CH2结构域和CH3结构域)。在一些实施方案中，Fc区包含、基本上由或由融合至CH3结构域(或其部分)的铰链结构域(或其部分)、融合至CH2结构域(或其部分)的铰链结构域(或其部分)、融合至CH3结构域(或其部分)的CH2结构域(或其部分)、融合至铰链结构域(或其部分)和CH3结构域(或其部分)二者的CH2结构域(或其部分)组成。在其它实施方案中，Fc区缺乏CH2结构域的至少一部分(如，CH2结构域的所有或部分)。在具体实施方案中，Fc区包含或由与EU编号第221至447位对应的氨基酸组成。

[0530] 本文中以F、F1或F2表示的Fc区可得自多个不同的来源。在一个实施方案中，多肽的Fc区来源于人免疫球蛋白。然而，应当理解Fc区可来源于另一个哺乳动物物种，包括例如啮齿类(如，小鼠、大鼠、兔、豚鼠)或非人类灵长类(如，黑猩猩、猕猴)物种的免疫球蛋白。此外，Fc结构域的多肽或其部分可来源于任何免疫球蛋白种类，包括IgM、IgG、IgD、IgA和IgE，和任何免疫球蛋白同种型，包括IgG1、IgG2、IgG3和IgG4。在另一个实施方案中，使用人同种型IgG1。

[0531] 在某些实施方案中，Fc变体赋予由包含所述野生型Fc结构域的Fc区所给予的至少一个效应子功能的变化(如，Fc区结合Fc受体(如FcγRI、FcγRII或FcγRIII)或补体蛋白(如C1q)，或触发抗体依赖性细胞毒性(ADCC)、吞噬作用或补体依赖性细胞毒性(CDCC)的能力增加或减少)。在其它实施方案中，Fc变体提供工程化的半胱氨酸残基。

[0532] 本发明的Fc区可利用本领域公认的Fc变体，所述变体已知赋予效应子功能和/或

FcR或FcRn结合的变化(如,增加或减少)。具体地讲,本发明的结合分子可包括例如国际PCT公布WO88/07089A1、WO96/14339A1、WO98/05787A1、WO98/23289A1、WO99/51642A1、WO99/58572A1、WO00/09560A2、WO00/32767A1、WO00/42072A2、WO02/44215A2、WO02/060919A2、WO03/074569A2、WO04/016750A2、WO04/029207A2、WO04/035752A2、WO04/063351A2、WO04/074455A2、WO04/099249A2、WO05/040217A2、WO04/044859、WO05/070963A1、WO05/077981A2、WO05/092925A2、WO05/123780A2、WO06/019447A1、WO06/047350A2和WO06/085967A2、美国专利公布号US2007/0231329、US2007/0231329、US2007/0237765、US2007/0237766、US2007/0237767、US2007/0243188、US20070248603、US20070286859、US20080057056或美国专利5,648,260、5,739,277、5,834,250、5,869,046、6,096,871、6,121,022、6,194,551、6,242,195、6,277,375、6,528,624、6,538,124、6,737,056、6,821,505、6,998,253、7,083,784、7,404,956和7,317,091中公开的一个或多个氨基酸位置处的变化(如,取代),每个专利以引用方式并入本文。在一个实施方案中,具体变化(如,本领域公开的一个或多个氨基酸的具体取代)可在一个或多个所公开的氨基酸位置处进行。在另一个实施方案中,可进行一个或多个所公开的氨基酸位置处的不同变化(如,本领域公开的一个或多个氨基酸位置的不同取代)。

[0533] 可根据熟知的程序,例如定点诱变等修饰IgG的Fc区或FcRn结合伴侣,以产生将由FcRn结合的修饰IgG或Fc片段或其部分。此类修饰包括远离FcRn接触位点的修饰,以及保持或甚至增强与FcRn的结合的接触位点内的修饰。例如,可取代人IgG1Fc(Fcγ1)中的如下单个氨基酸残基,而不导致Fc对FcRn的结合亲合力显著丧失:P238A、S239A、K246A、K248A、D249A、M252A、T256A、E258A、T260A、D265A、S267A、H268A、E269A、D270A、E272A、L274A、N276A、Y278A、D280A、V282A、E283A、H285A、N286A、T289A、K290A、R292A、E293A、E294A、Q295A、Y296F、N297A、S298A、Y300F、R301A、V303A、V305A、T307A、L309A、Q311A、D312A、N315A、K317A、E318A、K320A、K322A、S324A、K326A、A327Q、P329A、A330Q、P331A、E333A、K334A、T335A、S337A、K338A、K340A、Q342A、R344A、E345A、Q347A、R355A、E356A、M358A、T359A、K360A、N361A、Q362A、Y373A、S375A、D376A、A378Q、E380A、E382A、S383A、N384A、Q386A、E388A、N389A、N390A、Y391F、K392A、L398A、S400A、D401A、D413A、K414A、R416A、Q418A、Q419A、N421A、V422A、S424A、E430A、N434A、T437A、Q438A、K439A、S440A、S444A和K447A,其中例如P238A表示第238位的野生型脯氨酸被丙氨酸取代。例如,具体实施方案包括N297A突变,其移除了高度保守的N-糖基化位点。除丙氨酸之外,上述指定位置的其它氨基酸可被野生型氨基酸取代。突变可单独引入Fc,从而产生超过一百个不同于天然Fc的Fc区。另外,两个、三个或更多个这些单个突变的组合可一起引入,从而产生数百个Fc区。此外,可突变本发明的构建体的一个Fc区,而构建体的其它Fc区完全不突变,或它们二者均突变,但突变是不同。

[0534] 某些上述突变可赋予Fc区或FcRn结合伴侣新功能。例如,一个实施方案包括N297A,其移除了高度保守的N-糖基化位点。该突变的效应是减小免疫原性,从而增加Fc区的循环半衰期,并且使得Fc区不能结合FcγRI、FcγRIIA、FcγRIIB和FcγRIIIA,但不损坏FcRn的亲合力(Routledge等1995,Transplantation 60:847;Friend等1999,Transplantation 68:1632;Shields等1995,J.Biol.Chem.276:6591)。作为从上述突变产生的新功能的另一个实例,可增加对FcRn的亲合力,超过一些情况下的野生型。所述亲合力

增加可反映“结合”速率增加,“解离”速率降低,或“结合”速率增加和“解离”速率降低同时出现。据信突变的实例赋予对FcRn的亲合力增加,包括但不限于T256A、T307A、E380A和N434A(Shields等2001,J.Biol.Chem.276:6591)。

[0535] 另外,至少三个人Fcγ受体似乎识别低铰链区内IgG上的结合位点,通常为第234-237位氨基酸。因此,新功能和免疫原性潜在减小的另一个实例可来源于该区域的突变,例如通过替换人IgG1″ELLG″的第233-236位氨基酸为IgG2″PVA″(具有一个氨基酸缺失)的对应序列。据显示,当引入此类突变时,介导各种效应子功能的FcγRI、FcγRII和FcγRIII不结合IgG1。Ward和Ghetie 1995,Therapeutic Immunology 2:77和Armour等1999,Eur.J.Immunol.29:2613。

[0536] 在一个实施方案中,免疫球蛋白恒定区或其部分如Fc区是包括序列PKNSSMISNTP(SEQ ID NO:3),并且任选地还包括选自HQSLGTQ(SEQ ID NO:4)、HQNLSDGK(SEQ ID NO:5)、HQNISDGK(SEQ ID NO:6)或VISSHLGQ(SEQ ID NO:7)的序列的多肽(美国专利号5,739,277)。

[0537] 在另一个实施方案中,免疫球蛋白恒定区或其部分包含铰链区中的氨基酸序列或其部分,其与另一个免疫球蛋白恒定区或其部分形成一个或多个二硫键。免疫球蛋白恒定区或其部分的二硫键将包含FVIII的第一多肽和包含VWF片段的第二多肽放置在一起,使得内源性VWF不替代VWF片段,并且不结合FVIII。因此,第一免疫球蛋白恒定区或其部分和第二免疫球蛋白恒定区或其部分之间的二硫键阻止内源性VWF和FVIII蛋白之间的相互作用。所述VWF和FVIII蛋白之间的相互作用的抑制允许FVIII蛋白的半衰期超过两倍极限。铰链区或其部分还可连接至CH1、CH2、CH3的一个或多个结构域,或其片段以及它们的任何组合。在具体实例中,免疫球蛋白恒定区或其部分包含铰链区和CH2区(如,Fc区的第221-340位氨基酸)。

[0538] 在某些实施方案中,免疫球蛋白恒定区或其部分是半糖基化的。例如,包含两个Fc区或FcRn结合伴侣的嵌合蛋白可包含第一糖基化Fc区(如,糖基化CH2区)或FcRn结合伴侣和第二糖基化Fc区(如,糖基化CH2区)或FcRn结合伴侣。在一个实施方案中,连接基可插入糖基化和非糖基化Fc区之间。在另一个实施方案中,Fc区或FcRn结合伴侣是完全糖基化的,即所有Fc区均为糖基化的。在其它实施方案中,Fc区可以是非糖基化的,即Fc部分均不糖基化。

[0539] 在某些实施方案中,本发明的嵌合蛋白包含免疫球蛋白恒定区或其部分(如,Fc变体)的氨基酸取代,其改变了Ig恒定区的抗原非依赖性效应子功能,尤其是蛋白质的循环半衰期。

[0540] 当与缺乏这些取代的蛋白质比较时,此类蛋白质表现出与FcRn结合的增加或减少,因此在血清中半衰期分别增加或减少。预期对FcRn的具有增加的亲合力的Fc变体具有更长的血清半衰期,并且此类分子应用于治疗其中所施用多肽的长半衰期为所期望的哺乳动物的方法,如治疗慢性疾病或病症(参见如美国专利7,348,004、7,404,956和7,862,820)。相比之下,预期具有减少的FcRn的结合亲合力的Fc变体具有更短的半衰期,并且此类分子还用于例如施用至其中循环时间缩短可为有利的哺乳动物,例如用于体内诊断成像或当长期存在于循环中时起始多肽具有毒性副作用的情况。具有减少的FcRn结合亲合力的Fc变体也不太可能穿过胎盘,并且因此也用于治疗孕妇的疾病或病症。此外,其中FcRn的结合

亲合力减少的其它应用可为所期望的，包括其中定位脑、肾脏和/或肝脏为所期望的那些应用。在一个示例性实施方案中，本发明的嵌合蛋白表现出从脉管系统跨过肾小球上皮运输的减少。在另一个实施方案中，本发明的嵌合蛋白表现出从大脑跨过血脑屏障(BBB)至血管间隙的运输减少。在一个实施方案中，FcRn的结合改变的蛋白质包含至少一个Fc区或FcRn结合伴侣(如，一个或两个Fc区或FcRn结合伴侣)，Ig恒定区的“FcRn结合环”内具有一个或多个氨基酸取代。FcRn结合环由野生型全长Fc区的第280-299位氨基酸(根据EU编号)组成。在其它实施方案中，本发明的具有改变的FcRn的结合亲合力的嵌合蛋白中的Ig恒定区或其部分包含至少一个Fc区或FcRn结合伴侣，15Å FcRn“接触区”内具有一个或多个氨基酸取代。如本文所用，术语15Å FcRn“接触区”包括在野生型全长Fc部分以下位置的残基：第243-261位、第275-280位、第282-293位、第302-319位、第336-348位、第367位、第369位、第372-389位、第391位、第393位、第408位、第424位、第425-440位(EU编号)。在其它实施方案中，本发明的具有改变的FcRn结合亲合力的Ig恒定区或其部分包含至少一个Fc区或FcRn结合伴侣，其在与以下EU位置中的任何一者对应的氨基酸位置处具有一个或多个氨基酸取代：第256位、第277-281位、第283-288位、第303-309位、第313位、第338位、第342位、第376位、第381位、第384位、第385位、第387位、第434位(如，N434A或N434K)和第438位。改变FcRn结合活性的示例性氨基酸取代在国际PCT公布号WO05/047327中有所公开，该专利公布以引用方式并入本文。

[0541] 用于本发明的Fc区或FcRn结合伴侣也可包含本领域公认的氨基酸取代，其改变了嵌合蛋白的糖基化。例如，连接至VWF片段或FVIII蛋白的嵌合蛋白的Fc区或FcRn结合伴侣可包含具有导致糖基化(如，N-或O-连接糖基化)减少的突变的Fc区，或可包含野生型Fc部分的改变糖型(如，低岩藻糖或无岩藻糖聚糖)。

[0542] 在一个实施方案中，本发明的未加工嵌合蛋白可包含遗传融合的Fc区(即，scFc区)，其具有独立地选自本文所述的Ig恒定区或其部分的两个或更多个其组成型Ig恒定区或其部分。在一个实施方案中，二聚体Fc区的Fc区是相同的。在另一个实施方案中，至少两个Fc区是不同的。例如，本发明的蛋白质的Fc区或FcRn结合伴侣包含相同数量的氨基酸残基，或它们的长度可具有一个或多个氨基酸残基(如，约5个氨基酸残基(如，1个、2个、3个、4个或5个氨基酸残基)、约10个残基、约15个残基、约20个残基、约30个残基、约40个残基或约50个残基)的差异。在其它实施方案中，本发明的蛋白质的Fc区或FcRn结合伴侣可在一个或多个氨基酸位置处具有序列差异。例如，至少两个Fc区或FcRn结合伴侣可在约5个氨基酸位置(如，1个、2个、3个、4个或5个氨基酸位置)、约10个位置、约15个位置、约20个位置、约30个位置、约40个位置或约50个位置)处具有差异。

[0543] 2)白蛋白或片段或其变体

[0544] 在某些实施方案中，连接至VWF片段或连接至FVIII蛋白的异源部分是白蛋白或其功能片段。在其它实施方案中，本发明的嵌合蛋白包含FVIII蛋白和白蛋白或其片段，其中白蛋白或其片段屏蔽或保护FVIII蛋白上的VWF结合位点，从而抑制或阻止FVIII蛋白与内源性VWF的相互作用。

[0545] 人血清白蛋白(HSA或HA)是全长形式为609个氨基酸的蛋白质，负责大部分血清渗透压，并且还作为内源性和外源性配体的载体起作用。如本文所用，术语“白蛋白”包括全长

白蛋白或其功能片段、变体、衍生物或类似物。

[0546] 在一个实施方案中，嵌合蛋白包含本文所述的VWF片段和白蛋白、其片段或变体，其中VWF片段连接至白蛋白或其片段或变体。在另一个实施方案中，嵌合蛋白包含彼此结合的VWF片段和FVIII蛋白，其中VWF片段连接至白蛋白或其片段或变体，具有VIII活性的蛋白质连接至白蛋白或其片段或变体，或VWF片段和具有VIII活性的蛋白质二者连接至白蛋白或其片段或变体。在其它实施方案中，包含连接至白蛋白或其片段或变体的VWF片段的嵌合蛋白还连接至选自由以下组成的组的异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和PEG。在其它实施方案中，嵌合蛋白包含彼此结合的VWF片段和FVIII蛋白，其中FVIII蛋白连接至白蛋白或其片段或变体，并且还连接至选自由以下组成的组的异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和PEG。在其它实施方案中，嵌合蛋白包含连接至白蛋白或其片段或变体的VWF片段和连接至白蛋白或其片段或变体的FVIII蛋白，所述VWF片段和所述FVIII蛋白彼此结合，其中VWF片段活性还连接至选自由以下组成的组的第一异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和PEG，并且其中FVIII蛋白活性还连接至选自由以下组成的组的第二异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和PEG。

[0547] 在其它实施方案中，连接至VWF片段或FVIII蛋白的异源部分是白蛋白或其片段或变体，其延长(或能够延长)VWF片段或FVIII蛋白的半衰期。白蛋白或其片段或变体的另外实例在美国专利公布号2008/0194481A1、2008/0004206A1、2008/0161243A1、2008/0261877A1或2008/0153751A1或PCT申请公布号2008/033413A2、2009/058322A1或2007/021494A2中有所公开。

[0548] 3)白蛋白结合部分

[0549] 在某些实施方案中，连接至VWF片段或FVIII蛋白的异源部分是白蛋白结合部分，其包含白蛋白结合肽、细菌白蛋白结合结构域、白蛋白结合抗体片段或它们的任何组合。例如，白蛋白结合蛋白可以是细菌白蛋白结合蛋白、包括结构域抗体的抗体或抗体片段(参见美国专利号6,696,245)。例如，白蛋白结合蛋白可以是细菌白蛋白结合结构域，例如链球菌蛋白G的白蛋白结合结构域(Konig,T.和Skerra,A.(1998)J.Immunol.Methods 218,73-83)。可用作缀合伴侣的白蛋白结合肽的其它实例为例如具有Cys-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Cys共有序列的那些，其中Xaa_1为Asp、Asn、Ser、Thr或Trp，Xaa_2为Asn、Gln，H为Ile、Leu或Lys，Xaa_3为Ala、Asp、Phe、Trp或Tyr，并且Xaa_4为Asp、Gly、Leu、Phe、Ser或Thr，如美国专利申请2003/0069395或Dennis等(Dennis等(2002)J.Biol.Chem.277,35035-35043)所述。

[0550] 4)PAS序列

[0551] 在其它实施方案中，连接至VWF片段或FVIII蛋白的异源部分是PAS序列。在一个实施方案中，嵌合蛋白包含本文所述的VWF片段和PAS序列，其中VWF片段连接至PAS序列。在另一个实施方案中，本发明的嵌合蛋白包含FVIII蛋白和PAS序列，其中PAS序列屏蔽或保护FVIII蛋白上的VWF结合位点，从而抑制或阻止FVIII蛋白与内源性VWF的相互作用。

[0552] 如本文所用，PAS序列意指主要包含丙氨酸和丝氨酸残基或主要包含丙氨酸、丝氨酸和脯氨酸残基的氨基酸序列，所述氨基酸序列在生理条件下形成无规卷曲构象。因此，PAS序列是包含、基本上由或由丙氨酸、丝氨酸和脯氨酸组成的结构单元(building block)、氨基酸聚合物或序列盒，其可用作嵌合蛋白中异源部分的一部分。但是，技术人员

知道当加入除丙氨酸、丝氨酸和脯氨酸之外的残基，作为PAS序列中的微量组分时，氨基酸聚合物也可形成无规卷曲构象。如本文所用，术语“微量组分”意指除丙氨酸、丝氨酸和脯氨酸之外的氨基酸可加入PAS序列以达到某种程度，如最多约12％即约12/100个PAS序列的氨基酸，最多约10％即约10/100个PAS序列的氨基酸，最多约9％即约9/100个氨基酸，最多约8％即约8/100个氨基酸，约6％即约6/100个氨基酸，约5％即约5/100个氨基酸，约4％即约4/100个氨基酸，约3％即约3/100个氨基酸，约2％即约2/100个氨基酸，约1％即约1/100个氨基酸。不同于丙氨酸、丝氨酸和脯氨酸的氨基酸可选自由以下组成的组：Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Thr、Trp、Tyr和Val。

[0553] 在生理条件下，PAS序列段形成无规卷曲构象，并且从而可介导VWF因子或凝血活性蛋白增加的体内和/或体外稳定性。由于无规卷曲结构域不采取稳定结构，或通过自身起作用，由与其融合的VWF片段或FVIII蛋白介导的生物活性基本上得以保持。在其它实施方案中，形成无规卷曲结构域的PAS序列是生物惰性的，尤其是对于血浆中的蛋白水解、免疫原性、等电点/静电行为、结合细胞表面受体或内化，但仍然是生物可降解的，这提供了相对于合成聚合物例如PEG的显著优势。

[0554] 形成无规卷曲构象的PAS序列的非限制性实例包括选自由以下组成的组的氨基酸序列：ASPAAPAPASPAAPAPSAPA(SEQ ID NO:8)、AAPASPAPAAPSAPAPAAPS(SEQ ID NO:9)、APSSPSPSAPSSPSPASPSS(SEQ ID NO:10)、APSSPSPSAPSSPSPASPS(SEQ ID NO:11)、SSPSAPSPSSPASPSPSSPA(SEQ ID NO:12)、AASPAAPSAPPAAASPAAPSAPPA(SEQ ID NO:13)和ASAAAPAAASAAASAPSAAA(SEQ ID NO:14)或它们的任何组合。PAS序列的另外实例已知来自例如美国专利公布号2010/0292130 A1和PCT申请公布号WO 2008/155134 A1。

[0555] 5)HAP序列

[0556] 在某些实施方案中，连接至VWF片段或FVIII蛋白的异源部分是富含甘氨酸的均聚氨基酸聚合物(HAP)。HAP序列可包含甘氨酸的重复序列，具有的长度为至少50个氨基酸、至少100个氨基酸、120个氨基酸、140个氨基酸、160个氨基酸、180个氨基酸、200个氨基酸、250个氨基酸、300个氨基酸、350个氨基酸、400个氨基酸、450个氨基酸或500个氨基酸。在一个实施方案中，HAP序列能够延长融合至或连接至HAP序列的部分的半衰期。HAP序列的非限制性实例包括但不限于：$(Gly)_n$、$(Gly_4Ser)_n$或$S(Gly_4Ser)_n$，其中n为1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20。在一个实施方案中，n为20、21、22、23、24、25、26、26、28、29、30、31、32、33、34、35、36、37、38、39或40。在另一个实施方案中，n为50、60、70、80、90、100、110、120、130、140、150、160、170、180、190或200。参见例如Schlapschy M等，Protein Eng.Design Selection,20:273-284(2007)。

[0557] 6)转铁蛋白或其片段

[0558] 在某些实施方案中，连接至VWF片段或FVIII蛋白的异源部分是转铁蛋白或其片段。任何转铁蛋白均可用于制备本发明的嵌合蛋白。例如，野生型人Tf(Tf)是679个氨基酸的蛋白质，大约75KDa(不计算糖基化)，具有两个主要结构域N(约330个氨基酸)和C(约340个氨基酸)，它们似乎起源于基因复制。参见GenBank登录号NM001063、XM002793、M12530、XM039845、XM 039847和S95936(www.ncbi.nlm.nih.gov/)，它们全文均以引用方式并入本文。转铁蛋白包含两个结构域：N结构域和C结构域。N结构域包含两个亚结构域：N1结构域和N2结构域，并且C结构域包含两个亚结构域：C1结构域和C2结构域。

[0559] 在一个实施方案中，嵌合蛋白的转铁蛋白部分包括转铁蛋白剪接变体。在一个实例中，转铁蛋白剪接变体可以是人转铁蛋白的剪接变体，如Genbank登录号AAA61140。在另一个实施方案中，嵌合蛋白的转铁蛋白部分包括转铁蛋白序列的一个或多个结构域，如N结构域、C结构域、N1结构域、N2结构域、C1结构域、C2结构域或它们的任何组合。

[0560] 7)聚合物，如聚乙二醇(PEG)

[0561] 在其它实施方案中，连接至VWF片段的异源部分或具有凝血活性如FVIII活性的蛋白质是本领域已知的可溶性聚合物，包括但不限于：聚乙二醇、乙二醇/丙二醇共聚物、羧甲基纤维素、葡聚糖或聚乙烯醇。异源部分例如可溶性聚合物可连接至VWF片段或FVIII蛋白内或N-或C-末端的任何位置。在其它实施方案中，本发明的嵌合蛋白包含FVIII蛋白和PEG，其中PEG屏蔽或保护FVIII蛋白上的VWF结合位点，从而抑制或阻止FVIII蛋白与内源性VWF的相互作用。

[0562] 在某些实施方案中，嵌合蛋白包含本文所述的VWF片段和PEG，其中VWF片段连接至PEG。在另一个实施方案中，嵌合蛋白包含彼此结合的VWF片段和FVIII蛋白，其中VWF片段连接至PEG，FVIII蛋白连接至PEG，或VWF片段和FVIII蛋白二者均连接至PEG。在其它实施方案中，包含连接至PEG的VWF片段的嵌合蛋白还连接至选自由以下组成的组的异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES以及白蛋白、其片段或变体。在其它实施方案中，嵌合蛋白包含彼此结合的VWF片段和FVIII蛋白，其中FVIII蛋白还连接至选自由以下组成的组的异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和白蛋白、其片段或变体。在其它实施方案中，嵌合蛋白包含连接至PEG的VWF片段和连接至PEG的FVIII蛋白，所述VWF片段和所述FVIII蛋白彼此结合，其中VWF片段活性还连接至选自由以下组成的组的第一异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和白蛋白、其片段或变体，并且其中FVIII蛋白活性还连接至选自由以下组成的组的第二异源部分：免疫球蛋白恒定区或其部分(如，Fc区)、PAS序列、HES和白蛋白、其片段或变体。

[0563] 本发明还提供本发明嵌合蛋白的化学修饰衍生物，该衍生物可提供另外的优势，例如多肽的溶解度、稳定性和循环时间增加，或免疫原性减小(参见美国专利号4,179,337)。修饰的化学部分可选自由以下组成的组：水溶性聚合物包括但不限于聚乙二醇、乙二醇/丙二醇共聚物、羧甲基纤维素、葡聚糖和聚乙烯醇。嵌合蛋白可在分子内的随机位置或N-或C-末端，或分子内的预定位置修饰，并且可包括一个、两个、三个或更多个连接的化学部分。

[0564] 聚合物可具有任何分子量，并且可以是支化的或非支化的。对于聚乙二醇，在一个实施方案中，为便于处理和制备，分子量在约1kDa和约100kDa之间。可使用其它大小，取决于所需的特征(如，所需的持续释放的持续时间、对生物活性的影响(如果有的话)、易于处理、抗原性程度或缺乏抗原性以及聚乙二醇对蛋白质或类似物的其它已知影响)。例如，聚乙二醇可具有约200、500、1000、1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、10,000、10,500、11,000、11,500、12,000、12,500、13,000、13,500、14,000、14,500、15,000、15,500、16,000、16,500、17,000、17,500、18,000、18,500、19,000、19,500、20,000、25,000、30,000、35,000、40,000、45,000、50,000、55,000、60,000、65,000、70,000、75,000、80,000、85,000、90,000、95,000或100,000kDa的平均分子量。

[0565] 在一些实施方案中，聚乙二醇可具有支化结构。支化的聚乙二醇例如在美国专利号5,643,575;Morpurgo等,Appl.Biochem.Biotechnol.56:59-72(1996);Vorobjev等,Nucleosides Nucleotides 18:2745-2750(1999);和Caliceti等,Bioconjug.Chem.10:638-646(1999)中有所描述，这些专利和文献全文各自以引用方式并入本文。

[0566] 连接至本发明的每个嵌合蛋白、VWF片段或FVIII蛋白的聚乙二醇部分的数量(即，取代度)也可变化。例如，本发明的聚乙二醇化蛋白质可连接至平均1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、12个、15个、17个、20个或更多个聚乙二醇分子。相似地，平均取代度在例如1-3个、2-4个、3-5个、4-6个、5-7个、6-8个、7-9个、8-10个、9-11个、10-12个、11-13个、12-14个、13-15个、14-16个、15-17个、16-18个、17-19个或18-20个聚乙二醇部分/蛋白质分子的范围内。确定取代度的方法在例如Delgado等,Crit.Rev.Thera.Drug Carrier Sys.9:249-304(1992)中讨论。

[0567] 在一些实施方案中，FVIII蛋白可以是聚乙二醇化的。聚乙二醇化的因子VIII可以指因子VIII和至少一个聚乙二醇(PEG)分子之间形成的缀合物。

[0568] 在其它实施方案中，用于本发明的FVIII蛋白缀合至一种或多种聚合物。聚合物可以是水溶性的，并且共价或非共价连接至因子VIII或缀合至因子VIII的其它部分。聚合物的非限制性实例可以是聚(环氧烷)、聚(乙烯基吡咯烷酮)、聚(乙烯醇)、聚噁唑啉或聚(丙烯酰吗啉)。另外类型的聚合物缀合FVIII在美国专利号7,199,223中有所公开。

[0569] 8)羟乙基淀粉(HES)

[0570] 在某些实施方案中，连接至VWF片段或FVIII蛋白的异源部分是聚合物，如羟乙基淀粉(HES)或其衍生物。在一个实施方案中，嵌合蛋白包含本文所述的VWF片段和HES，其中VWF片段连接至HES。在其它实施方案中，本发明的嵌合蛋白包含融合至羟乙基淀粉(HES)的FVIII蛋白，其中羟乙基淀粉或其衍生物屏蔽或保护FVIII蛋白上的VWF结合位点免受内源性VWF的作用，从而抑制或阻止FVIII蛋白与内源性VWF的相互作用。

[0571] 羟乙基淀粉(HES)是天然存在的支链淀粉的衍生物，并且在体内被α-淀粉酶降解。HES是碳水化合物聚合物支链淀粉的取代衍生物，以最高95重量％的浓度存在于玉米淀粉中。HES表现出有利的生物特性，并且在临床血液稀释治疗中用作血浆代用剂(blood volume replacement agent)(Sommermeyer等,Krankenhauspharmazie,8(8),271-278(1987)和Weidler等,Arzneim.-Forschung/Drug Res.,41,494-498(1991))。

[0572] 支链淀粉包含葡萄糖部分，其中α-1,4-糖苷键存在于主链中，α-1,6-糖苷键存在于支化部位。该分子的物理-化学性质主要通过糖苷键的类型确定。由于切口α-1,4-糖苷键，生成具有约六个葡萄糖单体/转的螺旋结构。聚合物的物理-化学以及生物化学特性可通过取代修改。羟乙基的引入可通过碱性羟乙基化作用实现。通过修改反应条件，可采用相对于羟乙基化作用的未取代葡萄糖单体中相应羟基的不同反应性。由于这个事实，技术人员能够在有限范围内影响取代类型。

[0573] HES的主要特征在于分子量分布和取代度。取代度以DS表示，涉及技术人员已知的摩尔取代。参见如上文，尤其是第273页引用的Sommermeyer等,Krankenhauspharmazie,8(8),271-278(1987)。

[0574] 在一个实施方案中，羟乙基淀粉具有从1至300kD、从2至200kD、从3至100kD或从4至70kD的平均分子量(分子量平均值)。羟乙基淀粉还可表现出从0.1至3，优选地0.1至2，更

优选地0.1至0.9,优选地0.1至0.8的摩尔取代度,以及相对于羟乙基从2至20范围内的C2:C6取代比率。具有约130kD的平均分子量的HES的非限制性实例是具有取代度为0.2至0.8,例如0.2、0.3、0.4、0.5、0.6、0.7或0.8,优选地0.4至0.7,例如0.4、0.5、0.6或0.7的HES。在具体实施方案中,具有约130kD的平均分子量的HES是得自Fresenius的VOLUVEN®。VOLUVEN®是人工胶体,作为例如血浆代用剂,用于血容量不足治疗和预防的治疗指示。VOLUVEN®的特性是平均分子量为130,000+/-20,000D、摩尔取代为0.4,以及C2:C6比率为约9:1。在其它实施方案中,羟乙基淀粉的平均分子量范围为例如约4至70kD或10至70kD或12至70kD或18至70kD或50至70kD或4至50kD或10至50kD或12至50kD或18至50kD或4至18kD或10至18kD或12至18kD或4至12kD或10至12kD或4至10kD。在其它实施方案中,所用的羟乙基淀粉的平均分子量在大于4kD和小于70kD的范围内例如约10kD,或在从9至10kD或从10至11kD或从9至11kD的范围内或约12kD,或在从11至12kD)或从12至13kD或从11至13kD的范围内或约18kD,或在从17至18kD或从18至19kD或从17至19kD的范围内或约30kD,或在从29至30或从30至31kD的范围内或约50kD,或在从49至50kD或从50至51kD或从49至51kD的范围内。

[0575] 在某些实施方案中,异源部分可以是具有不同平均分子量和/或不同取代度和/或不同C2:C6取代比率的羟乙基淀粉混合物。因此,可使用具有不同平均分子量和不同取代度和不同C2:C6取代比率,或具有不同平均分子量和不同取代度和相同或大约相同C2:C6取代比率,或具有不同平均分子量和相同或大约相同取代度和不同C2:C6取代比率,或具有相同或大约相同平均分子量和不同取代度和不同C2:C6取代比率,或具有不同平均分子量和相同或大约相同取代度和相同或大约相同C2:C6取代比率,或具有相同或大约相同平均分子量和不同取代度和相同或大约相同C2:C6取代比率,或具有相同或大约相同平均分子量和相同或大约相同取代度和不同C2:C6取代比率,或具有大约相同平均分子量和大约相同取代度和大约相同C2:C6取代比率的羟乙基淀粉的混合物。

[0576] 9)聚唾液酸(PSA)

[0577] 在某些实施方案中,连接至VWF片段或FVIII蛋白的非多肽异源部分是聚合物,如聚唾液酸(PSA)或其衍生物。聚唾液酸(PSA)是由某些细菌菌株和哺乳动物某些细胞中产生的天然存在的唾液酸的非支链聚合物,Roth J.,等(1993)Polysialic Acid:From Microbes to Man,Roth J.,Rutishauser U.,Troy F.A.编辑(Birkhäuser Verlag,Basel,Switzerland),335-348页。它们可以从n=约80或更多的唾液酸残基至n=2的各种聚合度通过有限酸水解或通过神经氨酸酶消化,或通过天然、细菌来源形式的聚合物分离产生。不同聚唾液酸的组合物可变化,使得存在均聚形式,即包含大肠杆菌(E.coli)菌株K1和B-群脑膜炎球菌的荚膜多糖的α-2,8-连接聚唾液酸,其还见于神经细胞粘附分子(N-CAM)的雏形。还存在杂聚形式,例如大肠杆菌菌株K92和脑膜炎双球菌(N.meningitidis)的C群多糖的交替α-2,8α-2,9-聚唾液酸。唾液酸也可见于具有除唾液酸之外的单体的交替共聚物,例如脑膜炎双球菌的W135群或Y群。聚唾液酸具有重要生物学功能,包括通过病原菌逃离免疫和补体系统,以及胎儿发育期间未成熟神经元的胶质粘附调节(其中聚合物具有抗粘附功能),Cho和Troy,P.N.A.S.,USA,91(1994)11427-11431,但哺乳动物中不存在聚唾液酸的已知受体。大肠杆菌菌株K1的α-2,8-连接聚唾液酸也称为“多聚乙酰神经氨酸”,并且用于(以

各种长度)举例说明本发明。将聚唾液酸连接或缀合至多肽的各种方法已有所描述(例如参见美国专利号5,846,951、WO-A-0187922和US 2007/0191597 A1,这些专利全文以引用方式并入本文中。

[0578] C)FVIII蛋白

[0579] 如本文所用,除非另外指明,“FVIII蛋白”意指具有正常凝血作用的功能FVIII多肽。术语FVIII蛋白包括保持凝血途径中全长野生型因子VIII的功能的功能片段、其变体、类似物或衍生物。“FVIII蛋白”与FVIII多肽(或蛋白质)或FVIII可互换使用。FVIII功能的实例包括但不限于:活化凝血的能力、充当因子IX的辅因子的能力或在Ca2+和磷脂存在下与因子IX形成因子X酶复合物,然后将因子X转化为活化形式Xa的能力。FVIII蛋白可以是人、猪、狗、大鼠或鼠FVIII蛋白。此外,来自人和其它物种的FVIII之间的比较识别出可能为功能必须的保守残基(Cameron等,Thromb.Haemost.79:317-22(1998);US 6,251,632)。

[0580] 多个测试可用于评估凝血系统的功能:活化部分促凝血酶原激酶时间(aPTT)测试、显色测定、ROTEM测定、凝血酶原时间(PT)测试(还用于确定INR)、血纤维蛋白原测试(通常通过Clauss法进行)、血小板计数、血小板功能测试(通常通过PFA-100进行)、TCT、出血时间、混合测试(如果患者的血浆与正常血浆混合,是否纠正异常)、凝血因子测定、抗磷脂抗体、D-二聚体、基因测试(如因子V Leiden、凝血酶原突变G20210A)、稀释Russell蝰蛇毒时间(dRVVT)、血小板多种功能测试、凝血弹性描记(TEG或Sonoclot)、凝血弹性测量(thromboelastometry)(TEM®,如ROTEM®)或优球蛋白裂解时间(ELT)。

[0581] aPTT测试是测量“内因性”(也称为接触活化途径)和共同凝血途径二者功效的性能指标。该测试通常用于测量市售重组凝血因子如FVIII或FIX的凝血活性。它结合测量外因性途径的凝血酶原时间(PT)使用。

[0582] ROTEM分析提供关于止血:凝血时间的完整动力学、凝块形成、凝块稳定性和裂解的信息。凝血弹性测量的不同参数取决于血浆凝血系统的活性、血小板功能、血纤维蛋白溶解或影响这些相互作用的多个因素。该测定可提供二期止血的完整观察。

[0583] FVIII多肽和多核苷酸序列是已知的,正如多个功能片段、突变和修饰型式。人FVIII序列(全长)的实例以SEQ ID NO:16或18的子序列示出。

[0584] 表2.全长FVIII(FVIII信号肽以下划线表示;FVIII重链以双下划线表示;B结构域以斜体表示;并且FVIII轻链以纯文本表示)

[0585] 信号肽:(SEQ ID NO:15)

[0586] MQIELSTCFFLCLLRFCFS

[0587] 成熟因子VIII(SEQ ID NO:16)*

ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLL
GPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKEN
GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSL
MQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEI
SPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRF
DDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYT
DETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPIL
PGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILF
SVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDF
LSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYE
DSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIPENDIEKTDPWFAHRTPMPKIQNVSSSDLLM
LLRQSPTPHGLSLSDLQEAKYETFSDDPSPGAIDSNNSLSEMTHFRPQLHHSGDMVFTPESGLQLRLNEKLG
TTAATELKKLDFKVSSTSNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLDTTLFGKKSSPLTESGGPL
SLSEENNDSKLLESGLMNSQESSWGKNVSSTESGRLFKGKRAHGPALLTKDNALFKVSISLLKTNKTSNNSA
TNRKTHIDGPSLLIENSPSVWQNILESDTEFKKVTPLIHDRMLMDKNATALRLNHMSNKTTSSKNMEMVQQK
KEGPIPPDAQNPDMSFFKMLFLPESARWIQRTHGKNSLNSGQGPSPKQLVSLGPEKSVEGQNFLSEKNKVVV
GKGEFTKDVGLKEMVFPSSRNLFLTNLDNLHENNTHNQEKKIQEEIEKKETLIQENVVLPQIHTVTGTKNFM
KNLFLLSTRQNVEGSYDGAYAPVLQDFRSLNDSTNRTKKHTAHFSKKGEEENLEGLGNQTKQIVEKYACTTR
ISPNTSQQNFVTQRSKRALKQFRLPLEETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQS
PLSDCLTRSHSIPQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQGAKK
NNLSLAILTLEMTGDQREVGSLGTSATNSVTYKKVENTVLPKPDLPKTSGKVELLPKVHIYQKDLFPTETSN
GSPGHLDLVEGSLLQGTEGAIKWNEANRPGKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQE
KSPEKTAFKKKDTILSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKRHQREITRTTLQ
SDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVP
QFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGA
EPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVT
VQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYL
LSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLV
YSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQG
ARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRS
TLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQV
DFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTR
YLRIHPQSWVHQIALRMEVLGCEAQDLY

[0588]

[0589] 表3.编码全长FVIII的核苷酸序列(SEQ ID NO:17)*

```
 661                                                  ATG CAAATAGAGC TCTCCACCTG
 721    CTTCTTTCTG TGCCTTTTGC GATTCTGCTT TAGTGCCACC AGAAGATACT ACCTGGGTGC
 781    AGTGGAACTG TCATGGGACT ATATGCAAAG TGATCTCGGT GAGCTGCCTG TGGACGCAAG
 841    ATTTCCTCCT AGAGTGCCAA AATCTTTTCC ATTCAACACC TCAGTCGTGT ACAAAAAGAC
 901    TCTGTTTGTA GAATTCACGG ATCACCTTTT CAACATCGCT AAGCCAAGGC CACCCTGGAT
 961    GGGTCTGCTA GGTCCTACCA TCCAGGCTGA GGTTTATGAT ACAGTGGTCA TTACACTTAA
1021    GAACATGGCT TCCCATCCTG TCAGTCTTCA TGCTGTTGGT GTATCCTACT GGAAAGCTTC
1081    TGAGGGAGCT GAATATGATG ATCAGACCAG TCAAAGGGAG AAAGAAGATG ATAAAGTCTT
1141    CCCTGGTGGA AGCCATACAT ATGTCTGGCA GGTCCTGAAA GAGAATGGTC CAATGGCCTC
1201    TGACCCACTG TGCCTTACCT ACTCATATCT TTCTCATGTG GACCTGGTAA AAGACTTGAA
1261    TTCAGGCCTC ATTGGAGCCC TACTAGTATG TAGAGAAGGG AGTCTGGCCA AGGAAAAGAC
1321    ACAGACCTTG CACAAATTTA TACTACTTTT TGCTGTATTT GATGAAGGGA AAAGTTGGCA
1381    CTCAGAAACA AAGAACTCCT TGATGCAGGA TAGGGATGCT GCATCTGCTC GGGCCTGGCC
1441    TAAAATGCAC ACAGTCAATG GTTATGTAAA CAGGTCTCTG CCAGGTCTGA TTGGATGCCA
```

[0590]

1501 CAGGAAATCA GTCTATTGGC ATGTGATTGG AATGGGCACC ACTCCTGAAG TGCACTCAAT
1561 ATTCCTCGAA GGTCACACAT TTCTTGTGAG GAACCATCGC CAGGCGTCCT TGGAAATCTC
1621 GCCAATAACT TTCCTTACTG CTCAAACACT CTTGATGGAC CTTGGACAGT TTCTACTGTT
1681 TTGTCATATC TCTTCCCACC AACATGATGG CATGGAAGCT TATGTCAAAG TAGACAGCTG
1741 TCCAGAGGAA CCCCAACTAC GAATGAAAAA TAATGAAGAA GCGGAAGACT ATGATGATGA
1801 TCTTACTGAT TCTGAAATGG ATGTGGTCAG GTTTGATGAT GACAACTCTC CTTCCTTTAT
1861 CCAAATTCGC TCAGTTGCCA AGAAGCATCC TAAAACTTGG GTACATTACA TTGCTGCTGA
1921 AGAGGAGGAC TGGGACTATG CTCCCTTAGT CCTCGCCCCC GATGACAGAA GTTATAAAAG
1981 TCAATATTTG AACAATGGCC CTCAGCGGAT TGGTAGGAAG TACAAAAAAG TCCGATTTAT
2041 GGCATACACA GATGAAACCT TTAAGACTCG TGAAGCTATT CAGCATGAAT CAGGAATCTT
2101 GGGACCTTTA CTTTATGGGG AAGTTGGAGA CACACTGTTG ATTATATTTA AGAATCAAGC
2161 AAGCAGACCA TATAACATCT ACCCTCACGG AATCACTGAT GTCCGTCCTT TGTATTCAAG
2221 GAGATTACCA AAAGGTGTAA AACATTTGAA GGATTTTCCA ATTCTGCCAG GAGAAATATT
2281 CAAATATAAA TGGACAGTGA CTGTAGAAGA TGGGCCAACT AAATCAGATC CTCGGTGCCT
2341 GACCCGCTAT TACTCTAGTT TCGTTAATAT GGAGAGAGAT CTAGCTTCAG GACTCATTGG
2401 CCCTCTCCTC ATCTGCTACA AAGAATCTGT AGATCAAAGA GGAAACCAGA TAATGTCAGA
2461 CAAGAGGAAT GTCATCCTGT TTTCTGTATT TGATGAGAAC CGAAGCTGGT ACCTCACAGA
2521 GAATATACAA CGCTTTCTCC CCAATCCAGC TGGAGTGCAG CTTGAGGATC CAGAGTTCCA
2581 AGCCTCCAAC ATCATGCACA GCATCAATGG CTATGTTTTT GATAGTTTGC AGTTGTCAGT
2641 TTGTTTGCAT GAGGTGGCAT ACTGGTACAT TCTAAGCATT GGAGCACAGA CTGACTTCCT
2701 TTCTGTCTTC TTCTCTGGAT ATACCTTCAA ACACAAAATG GTCTATGAAG ACACACTCAC
2761 CCTATTCCCA TTCTCAGGAG AAACTGTCTT CATGTCGATG GAAAACCCAG GTCTATGGAT
2821 TCTGGGGTGC CACAACTCAG ACTTTCGGAA CAGAGGCATG ACCGCCTTAC TGAAGGTTTC
2881 TAGTTGTGAC AAGAACACTG GTGATTATTA CGAGGACAGT TATGAAGATA TTTCAGCATA
2941 CTTGCTGAGT AAAAACAATG CCATTGAACC AAGAAGCTTC TCCCAGAATT CAAGACACCC
3001 TAGCACTAGG CAAAAGCAAT TTAATGCCAC CACAATTCCA GAAAATGACA TAGAGAAGAC
3061 TGACCCTTGG TTTGCACACA GAACACCTAT GCCTAAAATA CAAAATGTCT CCTCTAGTGA
3121 TTTGTTGATG CTCTTGCGAC AGAGTCCTAC TCCACATGGG CTATCCTTAT CTGATCTCCA
3181 AGAAGCCAAA TATGAGACTT TTTCTGATGA TCCATCACCT GGAGCAATAG ACAGTAATAA
[0591] 3241 CAGCCTGTCT GAAATGACAC ACTTCAGGCC ACAGCTCCAT CACAGTGGGG ACATGGTATT
3301 TACCCCTGAG TCAGGCCTCC AATTAAGATT AAATGAGAAA CTGGGGACAA CTGCAGCAAC
3361 AGAGTTGAAG AAACTTGATT TCAAAGTTTC TAGTACATCA AATAATCTGA TTTCAACAAT
3421 TCCATCAGAC AATTTGGCAG CAGGTACTGA TAATACAAGT TCCTTAGGAC CCCCAAGTAT
3481 GCCAGTTCAT TATGATAGTC AATTAGATAC CACTCTATTT GGCAAAAAGT CATCTCCCCT
3541 TACTGAGTCT GGTGGACCTC TGAGCTTGAG TGAAGAAAAT AATGATTCAA AGTTGTTAGA
3601 ATCAGGTTTA ATGAATAGCC AAGAAAGTTC ATGGGGAAAA AATGTATCGT CAACAGAGAG
3661 TGGTAGGTTA TTTAAAGGGA AAAGAGCTCA TGGACCTGCT TTGTTGACTA AAGATAATGC
3721 CTTATTCAAA GTTAGCATCT CTTTGTTAAA GACAAACAAA ACTTCCAATA ATTCAGCAAC
3781 TAATAGAAAG ACTCACATTG ATGGCCCATC ATTATTAATT GAGAATAGTC CATCAGTCTG
3841 GCAAAATATA TTAGAAAGTG ACACTGAGTT TAAAAAAGTG ACACCTTTGA TTCATGACAG
3901 AATGCTTATG GACAAAAATG CTACAGCTTT GAGGCTAAAT CATATGTCAA ATAAAACTAC
3961 TTCATCAAAA AACATGGAAA TGGTCCAACA GAAAAAAGAG GGCCCCATTC CACCAGATGC
4021 ACAAAATCCA GATATGTCGT TCTTTAAGAT GCTATTCTTG CCAGAATCAG CAAGGTGGAT
4081 ACAAAGGACT CATGGAAAGA ACTCTCTGAA CTCTGGGCAA GGCCCCAGTC CAAAGCAATT
4141 AGTATCCTTA GGACCAGAAA AATCTGTGGA AGGTCAGAAT TTCTTGTCTG AGAAAAACAA
4201 AGTGGTAGTA GGAAAGGGTG AATTTACAAA GGACGTAGGA CTCAAAGAGA TGGTTTTTCC
4261 AAGCAGCAGA AACCTATTTC TTACTAACTT GGATAATTTA CATGAAAATA ATACACACAA
4321 TCAAGAAAAA AAAATTCAGG AAGAAATAGA AAAGAAGGAA ACATTAATCC AAGAGAATGT
4381 AGTTTTGCCT CAGATACATA CAGTGACTGG CACTAAGAAT TTCATGAAGA ACCTTTTCTT
4441 ACTGAGCACT AGGCAAAATG TAGAAGGTTC ATATGACGGG GCATATGCTC CAGTACTTCA
4501 AGATTTTAGG TCATTAAATG ATTCAACAAA TAGAACAAAG AAACACACAG CTCATTTCTC
4561 AAAAAAAGGG GAGGAAGAAA ACTTGGAAGG CTTGGGAAAT CAAACCAAGC AAATTGTAGA
4621 GAAATATGCA TGCACCACAA GGATATCTCC TAATACAAGC CAGCAGAATT TTGTCACGCA
4681 ACGTAGTAAG AGAGCTTTGA AACAATTCAG ACTCCCACTA GAAGAAACAG AACTTGAAAA
4741 AAGGATAATT GTGGATGACA CCTCAACCCA GTGGTCCAAA AACATGAAAC ATTTGACCCC
4801 GAGCACCCTC ACACAGATAG ACTACAATGA GAAGGAGAAA GGGGCCATTA CTCAGTCTCC
4861 CTTATCAGAT TGCCTTACGA GGAGTCATAG CATCCCTCAA GCAAATAGAT CTCCATTACC
4921 CATTGCAAAG GTATCATCAT TTCCATCTAT TAGACCTATA TATCTGACCA GGGTCCTATT
4981 CCAAGACAAC TCTTCTCATC TTCCAGCAGC ATCTTATAGA AAGAAAGATT CTGGGGTCCA

[0592]

```
5041    AGAAAGCAGT CATTTCTTAC AAGGAGCCAA AAAAAATAAC CTTTCTTTAG CCATTCTAAC
5101    CTTGGAGATG ACTGGTGATC AAAGAGAGGT TGGCTCCCTG GGGACAAGTG CCACAAATTC
5161    AGTCACATAC AAGAAAGTTG AGAACACTGT TCTCCCGAAA CCAGACTTGC CCAAAACATC
5221    TGGCAAAGTT GAATTGCTTC CAAAAGTTCA CATTTATCAG AAGGACCTAT TCCCTACGGA
5281    AACTAGCAAT GGGTCTCCTG GCCATCTGGA TCTCGTGGAA GGGAGCCTTC TTCAGGGAAC
5341    AGAGGGAGCG ATTAAGTGGA ATGAAGCAAA CAGACCTGGA AAAGTTCCCT TTCTGAGAGT
5401    AGCAACAGAA AGCTCTGCAA AGACTCCCTC CAAGCTATTG GATCCTCTTG CTTGGGATAA
5461    CCACTATGGT ACTCAGATAC CAAAAGAAGA GTGGAAATCC CAAGAGAAGT CACCAGAAAA
5521    AACAGCTTTT AAGAAAAAGG ATACCATTTT GTCCCTGAAC GCTTGTGAAA GCAATCATGC
5581    AATAGCAGCA ATAAATGAGG GACAAAATAA GCCCGAAATA GAAGTCACCT GGGCAAAGCA
5641    AGGTAGGACT GAAAGGCTGT GCTCTCAAAA CCCACCAGTC TTGAAACGCC ATCAACGGGA
5701    AATAACTCGT ACTACTCTTC AGTCAGATCA AGAGGAAATT GACTATGATG ATACCATATC
5761    AGTTGAAATG AAGAAGGAAG ATTTTGACAT TTATGATGAG GATGAAAATC AGAGCCCCCG
5821    CAGCTTTCAA AAGAAAACAC GACACTATTT TATTGCTGCA GTGGAGAGGC TCTGGGATTA
5881    TGGGATGAGT AGCTCCCCAC ATGTTCTAAG AAACAGGGCT CAGAGTGGCA GTGTCCCTCA
5941    GTTCAAGAAA GTTGTTTTCC AGGAATTTAC TGATGGCTCC TTTACTCAGC CCTTATACCG
6001    TGGAGAACTA AATGAACATT TGGGACTCCT GGGGCCATAT ATAAGAGCAG AAGTTGAAGA
6061    TAATATCATG GTAACTTTCA GAAATCAGGC CTCTCGTCCC TATTCCTTCT ATTCTAGCCT
6121    TATTTCTTAT GAGGAAGATC AGAGGCAAGG AGCAGAACCT AGAAAAAACT TTGTCAAGCC
6181    TAATGAAACC AAAACTTACT TTTGGAAAGT GCAACATCAT ATGGCACCCA CTAAAGATGA
6241    GTTTGACTGC AAAGCCTGGG CTTATTTCTC TGATGTTGAC CTGGAAAAAG ATGTGCACTC
6301    AGGCCTGATT GGACCCCTTC TGGTCTGCCA CACTAACACA CTGAACCCTG CTCATGGGAG
6361    ACAAGTGACA GTACAGGAAT TTGCTCTGTT TTTCACCATC TTTGATGAGA CCAAAAGCTG
6421    GTACTTCACT GAAAATATGG AAAGAAACTG CAGGGCTCCC TGCAATATCC AGATGGAAGA
6481    TCCCACTTTT AAAGAGAATT ATCGCTTCCA TGCAATCAAT GGCTACATAA TGGATACACT
6541    ACCTGGCTTA GTAATGGCTC AGGATCAAAG GATTCGATGG TATCTGCTCA GCATGGGCAG
6601    CAATGAAAAC ATCCATTCTA TTCATTTCAG TGGACATGTG TTCACTGTAC GAAAAAAAGA
6661    GGAGTATAAA ATGGCACTGT ACAATCTCTA TCCAGGTGTT TTTGAGACAG TGGAAATGTT
6721    ACCATCCAAA GCTGGAATTT GGCGGGTGGA ATGCCTTATT GGCGAGCATC TACATGCTGG
6781    GATGAGCACA CTTTTTCTGG TGTACAGCAA TAAGTGTCAG ACTCCCCTGG GAATGGCTTC
6841    TGGACACATT AGAGATTTTC AGATTACAGC TTCAGGACAA TATGGACAGT GGGCCCCAAA
6901    GCTGGCCAGA CTTCATTATT CCGGATCAAT CAATGCCTGG AGCACCAAGG AGCCCTTTTC
6961    TTGGATCAAG GTGGATCTGT TGGCACCAAT GATTATTCAC GGCATCAAGA CCCAGGGTGC
7021    CCGTCAGAAG TTCTCCAGCC TCTACATCTC TCAGTTTATC ATCATGTATA GTCTTGATGG
7081    GAAGAAGTGG CAGACTTATC GAGGAAATTC CACTGGAACC TTAATGGTCT TCTTTGGCAA
7141    TGTGGATTCA TCTGGGATAA AACACAATAT TTTTAACCCT CCAATTATTG CTCGATACAT
7201    CCGTTTGCAC CCAACTCATT ATAGCATTCG CAGCACTCTT CGCATGGAGT TGATGGGCTG
7261    TGATTTAAAT AGTTGCAGCA TGCCATTGGG AATGGAGAGT AAAGCAATAT CAGATGCACA
7321    GATTACTGCT TCATCCTACT TTACCAATAT GTTTGCCACC TGGTCTCCTT CAAAAGCTCG
7381    ACTTCACCTC CAAGGGAGGA GTAATGCCTG GAGACCTCAG GTGAATAATC CAAAAGAGTG
7441    GCTGCAAGTG GACTTCCAGA AGACAATGAA AGTCACAGGA GTAACTACTC AGGGAGTAAA
7501    ATCTCTGCTT ACCAGCATGT ATGTGAAGGA GTTCCTCATC TCCAGCAGTC AAGATGGCCA
7561    TCAGTGGACT CTCTTTTTTC AGAATGGCAA AGTAAAGGTT TTTCAGGGAA ATCAAGACTC
7621    CTTCACACCT GTGGTGAACT CTCTAGACCC ACCGTTACTG ACTCGCTACC TTCGAATTCA
7681    CCCCCAGAGT TGGGTGCACC AGATTGCCCT GAGGATGGAG GTTCTGGGCT GCGAGGCACA
7741    GGACCTCTAC
```

[0593] *加下划线的核酸编码信号肽。

[0594] FVIII多肽包括全长FVIII、N-末端减去Met的全长FVIII、成熟FVIII(减去信号序列)、N-末端具有另外Met的成熟FVIII和/或B结构域完全或部分缺失的FVIII。在某些实施方案中,FVIII变体包括部分或完全缺失的B结构域缺失。

[0595] 分离人FVIII基因并在哺乳动物细胞中表达(Toole,J.J.,等,Nature 312:342-347(1984);Gitschier,J.,等,Nature 312:326-330(1984);Wood,W.I.,等,Nature 312:330-337(1984);Vehar,G.A.,等,Nature 312:337-342(1984);WO 87/04187、WO 88/08035、WO 88/03558和美国专利号4,757,006)。FVIII氨基酸序列从cDNA推导,如美国专利号4,965,199中所示。此外,部分或完全B-结构域缺失的FVIII如美国专利号4,994,371和4,868,112中所示。在一些实施方案中,人FVIII B-结构域被人因子V B-结构域替代,如美国专利号5,004,803中所示。编码人因子VIII和氨基酸序列的cDNA序列分别如美国申请公布号2005/0100990的SEQ ID NO:17和16所示。

[0596] 猪FVIII序列在Toole,J.J.,等,Proc.Natl.Acad.Sci.USA 83:5939-5942(1986)中公开。另外,从猪脾脏cDNA文库PCR扩增FVIII序列得到的完整猪cDNA序列在Healey,J.F.,等,Blood 88:4209-4214(1996)中报道。具有所有结构域、所有亚基和具体氨基酸序列的取代的杂合人/猪FVIII在Lollar和Runge的美国专利号5,364,771和WO 93/20093中有所公开。最近,猪FVIII和猪A1和/或A2结构域被对应的人结构域取代的嵌合FVIII的A1和A2结构域的核苷酸和对应氨基酸序列在WO 94/11503中报道。Lollar,J.S.的美国专利号5,859,204还公开了猪cDNA和推导氨基酸序列。美国专利号6,458,563公开了B-结构域缺失的猪FVIII。

[0597] Lollar,J.S.的美国专利号5,859,204报道了抗原性减少和免疫反应性减少的FVIII的功能突变体。Lollar,J.S.的美国专利号6,376,463还报道了免疫反应性减少的FVIII的突变体。Saenko等的美国申请公布号2005/0100990报道了FVIII的A2结构域的功能突变。

[0598] 在一个实施方案中,FVIII(或嵌合蛋白的FVIII部分)与SEQ ID NO:18的第1至1438位氨基酸或SEQ ID NO:16(不含信号序列)的第1至2332位氨基酸的FVIII氨基酸序列或SEQ ID NO:15和SEQ ID NO:18的第-19至1438位氨基酸或SEQ ID NO:15和SEQ ID NO:16(含信号序列)的第-19至2332位氨基酸的FVIII氨基酸序列可具有至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%的同一性,其中FVIII具有凝血活性,如活化作为辅因子的因子IX,以将因子X转化为活化的因子X。FVIII(或嵌合蛋白的FVIII部分)可与SEQ ID NO:18的第1至1438位氨基酸或SEQ ID NO:16(不含信号序列)的第1至2332位氨基酸的FVIII氨基酸序列相同。FVIII还可包含信号序列。

[0599] 如本文所用,FVIII的“B-结构域”与本领域已知的由内部氨基酸序列同一性和蛋白酶裂解位点,如全长人FVIII的Ser741-Arg1648残基确定的B-结构域相同。其它人FVIII结构域通过如下氨基酸残基确定:A1,Ala1-Arg372残基;A2,Ser373-Arg740残基;A3,Ser1690-Asn2019残基;C1,Lys2020-Asn2172残基;C2,Ser2173-Tyr2332残基。A3-C1-C2序列包括Ser1690-Tyr2332残基。其余序列Glu1649-Arg1689残基通常称为a3酸性区。所有结构域包括猪、小鼠和狗FVIII的B-结构域的边界位置也是本领域已知的。在一个实施方案中,FVIII的B结构域是缺失的(“B-结构域缺失的因子VIII”或“BDD FVIII”)。BDD FVIII的实例是REFACTO®(重组BDD FVIII),其具有与表4中序列的因子VIII部分相同的序列。(BDD FVIII重链以双下划线表示;B结构域以斜体表示;并且BDD FVIII轻链以纯文本表示)。

[0600] 表4

[0601] BDD FVIII(SEQ ID NO:18)

[0602] ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLL
GPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKEN
GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSL
MQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEI
SPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRF
DDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYT
DETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPIL
PGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILF
SVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDF
LSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYE
DSYEDISAYLLSKNNAIEPR*SFSQNPPVLKRHQR*EITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQ
SPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGL
LGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEF
DCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAP
CNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMA
LYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW
APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGN
STGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQI
TASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLI
SSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

[0603] 表5.编码BDD FVIII的核苷酸序列(SEQ ID NO:19)*

[0604]

```
661                                    A TGCAAATAGA GCTCTCCACC TGCTTCTTTC
721  TGTGCCTTTT GCGATTCTGC TTTAGTGCCA CCAGAAGATA CTACCTGGGT GCAGTGGAAC
781  TGTCATGGGA CTATATGCAA AGTGATCTCG GTGAGCTGCC TGTGGACGCA AGATTTCCTC
841  CTAGAGTGCC AAAATCTTTT CCATTCAACA CCTCAGTCGT GTACAAAAAG ACTCTGTTTG
901  TAGAATTCAC GGATCACCTT TTCAACATCG CTAAGCCAAG GCCACCCTGG ATGGGTCTGC
961  TAGGTCCTAC CATCCAGGCT GAGGTTTATG ATACAGTGGT CATTACACTT AAGAACATGG
1021 CTTCCCATCC TGTCAGTCTT CATGCTGTTG GTGTATCCTA CTGGAAAGCT TCTGAGGGAG
1081 CTGAATATGA TGATCAGACC AGTCAAAGGG AGAAAGAAGA TGATAAAGTC TTCCCTGGTG
1141 GAAGCCATAC ATATGTCTGG CAGGTCCTGA AAGAGAATGG TCCAATGGCC TCTGACCCAC
1201 TGTGCCTTAC CTACTCATAT CTTTCTCATG TGGACCTGGT AAAAGACTTG AATTCAGGCC
1261 TCATTGGAGC CCTACTAGTA TGTAGAGAAG GGAGTCTGGC CAAGGAAAAG ACACAGACCT
1321 TGCACAAATT TATACTACTT TTTGCTGTAT TTGATGAAGG GAAAAGTTGG CACTCAGAAA
1381 CAAAGAACTC CTTGATGCAG GATAGGGATG CTGCATCTGC TCGGGCCTGG CCTAAAATGC
1441 ACACAGTCAA TGGTTATGTA AACAGGTCTC TGCCAGGTCT GATTGGATGC CACAGGAAAT
1501 CAGTCTATTG GCATGTGATT GGAATGGGCA CCACTCCTGA AGTGCACTCA ATATTCCTCG
1561 AAGGTCACAC ATTTCTTGTG AGGAACCATC GCCAGGCGTC CTTGGAAATC TCGCCAATAA
1621 CTTTCCTTAC TGCTCAAACA CTCTTGATGG ACCTTGGACA GTTTCTACTG TTTTGTCATA
1681 TCTCTTCCCA CCAACATGAT GGCATGGAAG CTTATGTCAA AGTAGACAGC TGTCCAGAGG
1741 AACCCCAACT ACGAATGAAA AATAATGAAG AAGCGGAAGA CTATGATGAT GATCTTACTG
1801 ATTCTGAAAT GGATGTGGTC AGGTTTGATG ATGACAACTC TCCTTCCTTT ATCCAAATTC
1861 GCTCAGTTGC CAAGAAGCAT CCTAAAACTT GGGTACATTA CATTGCTGCT GAAGAGGAGG
1921 ACTGGGACTA TGCTCCCTTA GTCCTCGCCC CCGATGACAG AAGTTATAAA AGTCAATATT
1981 TGAACAATGG CCCTCAGCGG ATTGGTAGGA AGTACAAAAA AGTCCGATTT ATGGCATACA
2041 CAGATGAAAC CTTTAAGACT CGTGAAGCTA TTCAGCATGA ATCAGGAATC TTGGGACCTT
2101 TACTTTATGG GGAAGTTGGA GACACACTGT TGATTATATT TAAGAATCAA GCAAGCAGAC
2161 CATATAACAT CTACCCTCAC GGAATCACTG ATGTCCGTCC TTTGTATTCA AGGAGATTAC
2221 CAAAAGGTGT AAAACATTTG AAGGATTTTC CAATTCTGCC AGGAGAAATA TTCAAATATA
2281 AATGGACAGT GACTGTAGAA GATGGGCCAA CTAAATCAGA TCCTCGGTGC CTGACCCGCT
2341 ATTACTCTAG TTTCGTTAAT ATGGAGAGAG ATCTAGCTTC AGGACTCATT GGCCCTCTCC
2401 TCATCTGCTA CAAAGAATCT GTAGATCAAA GAGGAAACCA GATAATGTCA GACAAGAGGA
2461 ATGTCATCCT GTTTTCTGTA TTTGATGAGA ACCGAAGCTG GTACCTCACA GAGAATATAC
2521 AACGCTTTCT CCCCAATCCA GCTGGAGTGC AGCTTGAGGA TCCAGAGTTC CAAGCCTCCA
2581 ACATCATGCA CAGCATCAAT GGCTATGTTT TTGATAGTTT GCAGTTGTCA GTTTGTTTGC
2641 ATGAGGTGGC ATACTGGTAC ATTCTAAGCA TTGGAGCACA GACTGACTTC CTTTCTGTCT
2701 TCTTCTCTGG ATATACCTTC AAACACAAAA TGGTCTATGA AGACACACTC ACCCTATTCC
2761 CATTCTCAGG AGAAACTGTC TTCATGTCGA TGGAAAACCC AGGTCTATGG ATTCTGGGGT
2821 GCCACAACTC AGACTTTCGG AACAGAGGCA TGACCGCCTT ACTGAAGGTT TCTAGTTGTG
2881 ACAAGAACAC TGGTGATTAT TACGAGGACA GTTATGAAGA TATTTCAGCA TACTTGCTGA
2941 GTAAAAACAA TGCCATTGAA CCAAGAAGCT TCTCTCAAAA CCCACCAGTC TTGAAACGCC
3001 ATCAACGGGA AATAACTCGT ACTACTCTTC AGTCAGATCA AGAGGAAATT GACTATGATG
3061 ATACCATATC AGTTGAAATG AAGAAGGAAG ATTTTGACAT TTATGATGAG GATGAAAATC
3121 AGAGCCCCCG CAGCTTTCAA AAGAAAACAC GACACTATTT TATTGCTGCA GTGGAGAGGC
3181 TCTGGGATTA TGGGATGAGT AGCTCCCCAC ATGTTCTAAG AAACAGGGCT CAGAGTGGCA
3241 GTGTCCCTCA GTTCAAGAAA GTTGTTTTCC AGGAATTTAC TGATGGCTCC TTTACTCAGC
3301 CCTTATACCG TGGAGAACTA AATGAACATT TGGGACTCCT GGGGCCATAT ATAAGAGCAG
3361 AAGTTGAAGA TAATATCATG GTAACTTTCA GAAATCAGGC CTCTCGTCCC TATTCCTTCT
3421 ATTCTAGCCT TATTTCTTAT GAGGAAGATC AGAGGCAAGG AGCAGAACCT AGAAAAAACT
3481 TTGTCAAGCC TAATGAAACC AAAACTTACT TTTGGAAAGT GCAACATCAT ATGGCACCCA
3541 CTAAAGATGA GTTTGACTGC AAAGCCTGGG CTTATTTCTC TGATGTTGAC CTGGAAAAAG
3601 ATGTGCACTC AGGCCTGATT GGACCCCTTC TGGTCTGCCA CACTAACACA CTGAACCCTG
3661 CTCATGGGAG ACAAGTGACA GTACAGGAAT TTGCTCTGTT TTTCACCATC TTTGATGAGA
3721 CCAAAAGCTG GTACTTCACT GAAAATATGG AAAGAAACTG CAGGGCTCCC TGCAATATCC
3781 AGATGGAAGA TCCCACTTTT AAAGAGAATT ATCGCTTCCA TGCAATCAAT GGCTACATAA
3841 TGGATACACT ACCTGGCTTA GTAATGGCTC AGGATCAAAG GATTCGATGG TATCTGCTCA
3901 GCATGGGCAG CAATGAAAAC ATCCATTCTA TTCATTTCAG TGGACATGTG TTCACTGTAC
3961 GAAAAAAAGA GGAGTATAAA ATGGCACTGT ACAATCTCTA TCCAGGTGTT TTTGAGACAG
```

[0605]

```
4021    TGGAAATGTT ACCATCCAAA GCTGGAATTT GGCGGGTGGA ATGCCTTATT GGCGAGCATC
4081    TACATGCTGG GATGAGCACA CTTTTTCTGG TGTACAGCAA TAAGTGTCAG ACTCCCCTGG
4141    GAATGGCTTC TGGACACATT AGAGATTTTC AGATTACAGC TTCAGGACAA TATGGACAGT
4201    GGGCCCCAAA GCTGGCCAGA CTTCATTATT CCGGATCAAT CAATGCCTGG AGCACCAAGG
4261    AGCCCTTTTC TTGGATCAAG GTGGATCTGT TGGCACCAAT GATTATTCAC GGCATCAAGA
4321    CCCAGGGTGC CCGTCAGAAG TTCTCCAGCC TCTACATCTC TCAGTTTATC ATCATGTATA
4381    GTCTTGATGG GAAGAAGTGG CAGACTTATC GAGGAAATTC CACTGGAACC TTAATGGTCT
4441    TCTTTGGCAA TGTGGATTCA TCTGGGATAA AACACAATAT TTTTAACCCT CCAATTATTG
4501    CTCGATACAT CCGTTTGCAC CCAACTCATT ATAGCATTCG CAGCACTCTT CGCATGGAGT
4561    TGATGGGCTG TGATTTAAAT AGTTGCAGCA TGCCATTGGG AATGGAGAGT AAAGCAATAT
4621    CAGATGCACA GATTACTGCT TCATCCTACT TTACCAATAT GTTTGCCACC TGGTCTCCTT
4681    CAAAAGCTCG ACTTCACCTC CAAGGGAGGA GTAATGCCTG GAGACCTCAG GTGAATAATC
4741    CAAAAGAGTG GCTGCAAGTG GACTTCCAGA AGACAATGAA AGTCACAGGA GTAACTACTC
4801    AGGGAGTAAA ATCTCTGCTT ACCAGCATGT ATGTGAAGGA GTTCCTCATC TCCAGCAGTC
4861    AAGATGGCCA TCAGTGGACT CTCTTTTTTC AGAATGGCAA AGTAAAGGTT TTTCAGGGAA
4921    ATCAAGACTC CTTCACACCT GTGGTGAACT CTCTAGACCC ACCGTTACTG ACTCGCTACC
4981    TTCGAATTCA CCCCCAGAGT TGGGTGCACC AGATTGCCCT GAGGATGGAG GTTCTGGGCT
5041    GCGAGGCACA GGACCTCTAC
```

[0606] *加下划线的核酸编码信号肽。

[0607] “B-结构域缺失的FVIII”可具有美国专利号6,316,226、6,346,513、7,041,635、5,789,203、6,060,447、5,595,886、6,228,620、5,972,885、6,048,720、5,543,502、5,610,278、5,171,844、5,112,950、4,868,112和6,458,563中公开的完全或部分缺失。在一些实施方案中,本发明的B-结构域缺失的FVIII序列包含美国专利号6,316,226(以及US 6,346,513)的第4段第4行至第5段第28行和实施例1-5中公开的缺失的任一者。在另一个实施方案中,B-结构域缺失的因子VIII是S743/Q1638B-结构域缺失的因子VIII(SQ BDD FVIII)(如,具有第744位氨基酸至第1637位氨基酸的缺失的因子VIII,如具有SEQ ID NO:16,即SEQ ID NO:18的第1-743位氨基酸和第1638-2332位氨基酸的因子VIII)。在一些实施方案中,本发明的B-结构域缺失的FVIII具有美国专利号5,789,203(以及US 6,060,447、US 5,595,886和US 6,228,620)的第2段第26-51行和实施例5-8中公开的缺失。在一些实施方案中,B-结构域缺失的因子VIII具有美国专利号5,972,885的第1段第25行至第2段第40行、美国专利号6,048,720的第6段第1-22行和实施例1、美国专利号5,543,502的第2段第17-46行、美国专利号5,171,844的第4段第22行至第5段第36行、美国专利号5,112,950的第2段第55-68行、图2和实施例1、美国专利号4,868,112的第2段第2行至第19段第21行和表2、美国专利号7,041,635的第2段第1行至第3段第19行、第3段第40行至第4段第67行、第7段第43行至第8段第26行和第11段第5行至第13段第39行或美国专利号6,458,563的第4段第25-53行中描述的缺失。

[0608] 在一些实施方案中,B-结构域缺失的FVIII具有大多数B结构域的缺失,但仍包含对于体内将初级翻译产物蛋白酶加工为两条多肽链必须的B结构域的氨基末端序列,如WO 91/09122中所公开。在一些实施方案中,B-结构域缺失的FVIII通过第747-1638位氨基酸缺失,即实际上B结构域的完整缺失而构建。Hoeben R.C.,等J.Biol.Chem.265(13):7318-7323(1990)。B-结构域缺失的因子VIII也可包含FVIII的第771-1666位氨基酸或第868-1562位氨基酸缺失。Meulien P.,等Protein Eng.2(4):301-6(1988)。作为本发明部分的另外B结构域缺失包括:第982至1562位或第760至1639位(Toole等,Proc.Natl.Acad.Sci.U.S.A.(1986)83,5939-5942)、第797至1562位(Eaton,等Biochemistry(1986)25:8343-8347)、第741至1646位(Kaufman(PCT公布专利申请号WO 87/04187))、第747-1560位(Sarver,等,DNA(1987)6:553-564)、第741至1648位(Pasek(PCT专

利申请号88/00831))或第816至1598位或第741至1648位(Lagner(Behring Inst.Mitt.(1988)No 82:16-25,EP 295597))氨基酸缺失。在其它实施方案中,BDD FVIII包括包含保持一个或多个N-连接糖基化位点,如对应于全长FVIII序列的氨基酸序列的第757、784、828、900、963或任选地943位残基的B-结构域的片段的FVIII多肽。B-结构域片段的实例包括B-结构域的226个氨基酸或163个氨基酸,如Miao,H.Z.,等,Blood 103(a):3412-3419(2004)、Kasuda,A,等,J.Thromb.Haemost.6:1352-1359(2008)和Pipe,S.W.,等,J.Thromb.Haemost.9:2235-2242(2011)中所公开(即,保持B结构域的前226个氨基酸或163个氨基酸)。在一些实施方案中,具有部分B-结构域的FVIII是FVIII198(SEQ ID NO:105)。FVIII198是包含部分B-结构域的单链FVIIIFc分子-226N6。226表示FVIII B-结构域的N-末端226个氨基酸,并且N6表示B-结构域中的六个N-糖基化位点。在其它实施方案中,BDD FVIII还包含第309位残基的点突变(从Phe突变为Ser),以提高BDD FVIII蛋白的表达。参见Miao,H.Z.,等,Blood 103(a):3412-3419(2004)。在其它实施方案中,BDD FVIII包括包含B-结构域的一部分,但不包含一个或多个弗林蛋白酶切割位点(如,Arg1313和Arg1648)的FVIII多肽。参见Pipe,S.W.,等,J.Thromb.Haemost.9:2235-2242(2011)。每个上述缺失可在任何FVIII序列中制备。

[0609] 用于本发明的FVIII蛋白可包括其中具有不影响FVIII凝血活性的一个或多个另外的异源序列或化学或物理修饰的FVIII。此类异源序列或化学或物理修饰可融合至FVIII蛋白的C-末端或N-末端或插入FVIII蛋白中两个氨基酸残基的一个或多个之间。此类插入FVIII蛋白不影响FVIII凝血活性或或FVIII功能。在一个实施方案中,插入改善了FVIII蛋白的药代动力学特性(如,半衰期)。在另一个实施方案中,插入可以是多于两个、三个、四个、五个或六个位点。

[0610] 在一个实施方案中,FVIII在第1648位氨基酸(全长因子VIII或SEQ ID NO:16中)、第754位氨基酸(S743/Q1638B-结构域缺失的因子VIII或SEQ ID NO:16中)的精氨酸或对应的精氨酸残基(其它变体中)之后切割,从而产生重链和轻链。在另一个实施方案中,FVIII包含通过金属离子介导的非共价键连接或缔合的重链和轻链。

[0611] 在其它实施方案中,FVIII是未在第1648位氨基酸(全长FVIII或SEQ ID NO:16中)、第754位氨基酸(S743/Q1638B-结构域缺失的FVIII或SEQ ID NO:18中)的精氨酸或对应的精氨酸残基(其它变体中)之后切割的单链FVIII。单链FVIII可包含一个或多个氨基酸取代。在一个实施方案中,氨基酸取代位于与全长成熟因子VIII多肽(SEQ ID NO:16)的第1648位残基、第1645位残基或它们二者或SQ BDD因子VIII(SEQ ID NO:18)的第754位残基、第751位残基或它们二者对应的残基。氨基酸取代可以是除精氨酸之外的任何氨基酸,如异亮氨酸、亮氨酸、赖氨酸、甲硫氨酸、苯丙氨酸、苏氨酸、色氨酸、缬氨酸、丙氨酸、天冬酰胺、天冬氨酸、半胱氨酸、谷氨酸、谷氨酰胺、甘氨酸、脯氨酸、硒半胱氨酸、丝氨酸、酪氨酸、组氨酸、鸟氨酸、吡咯赖氨酸或牛磺酸。

[0612] FVIII还可被凝血酶切割,然后活化为FVIIIa,作为活化因子IX(FIXa)的辅因子。并且活化的FIX与活化的FVIII一起形成因子X酶复合物,并且将因子X转化为活化的因子X(FXa)。对于活化而言,FVIII在三个精氨酸残基后的第372、740和1689位氨基酸(对应于B-结构域缺失的FVIII序列中的第372、740和795位氨基酸)被凝血酶切割,切割生成具有50kDa A1、43kDa A2和73kDa A3-C1-C2链的FVIIIa。在一个实施方案中,用于本发明的

FVIII蛋白是非活化的FVIII。在另一个实施方案中,FVIII蛋白是活化的FVIII。

[0613] 具有连接至VWF片段或与之缔合的FVIII多肽的蛋白质可包含与SEQ ID NO:16或18具有至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%的同一性的序列,其中该序列具有FVIII凝血活性,如活化因子IX,作为将因子X转化为活化因子X(FXa)的辅因子。

[0614] 如本文所用,"杂合"多肽和蛋白质意指第一多肽链如VWF片段(任选地融合至第一异源部分)与第二多肽链如FVIII蛋白(任选地融合至第二异源部分)组合,从而形成的异源二聚体。在一个实施方案中,杂合物中的第一多肽和第二多肽通过蛋白质-蛋白质相互作用,例如电荷-电荷或疏水相互作用彼此缔合。在另一个实施方案中,杂合物中的第一多肽和第二多肽通过二硫键或其它共价键彼此缔合。杂合物在例如US 2004/101740和US2006/074199中有所描述。第二多肽可以是第一多肽的相同拷贝或不同的多肽。在一个实施方案中,第一多肽是VWF片段-Fc融合蛋白,第二多肽是包含、基本上由或由FcRn结合结构域组成的多肽,其中第一多肽和第二多肽彼此缔合。在另一个实施方案中,第一多肽包含VWF片段-Fc融合蛋白,并且第二多肽包含FVIII-Fc融合蛋白,从而使杂合物为异源二聚体。第一多肽和第二多肽可通过共价键,如第一Fc区和第二Fc区之间的二硫键缔合。第一多肽和第二多肽还可通过VWF片段和FVIII蛋白之间的结合彼此缔合。

[0615] D)连接基

[0616] 本发明的嵌合蛋白还包含连接基。一个或多个连接基可存在于任何两个蛋白质之间,如辅助部分和FVIII蛋白之间(有时也称为"FVIII/AM连接基")、VWF片段和第一异源部分之间(有时也称为"VWF连接基")如第一Fc区、FVIII蛋白和第二异源部分之间(有时也称为"FVIII连接基")如第二Fc区、VWF片段和FVIII蛋白之间(如,FVIII/AM连接基)、VWF片段和第二异源部分之间和/或FVIII蛋白和第一异源部分之间。每个连接基可具有相同或不同的序列。在一个实施方案中,连接基是多肽连接基。在另一个实施方案中,连接基是非多肽连接基。

[0617] 用于本发明的连接基可包括任何有机分子。在一个实施方案中,连接基是聚合物,如聚乙二醇(PEG)或羟乙基淀粉(HES)。在另一个实施方案中,连接基是氨基酸序列(如,多肽连接基)。多肽连接基可包含至少约10个、20个、30个、40个、50个、60个、70个、80个、90个、100个、150个、200个、300个、400个、500个、600个、700个、800个、900个、1000个、1100个、1200个、1300个、1400个、1500个、1600个、1700个、1800个、1900个或2000个氨基酸。连接基可包含1-5个氨基酸、1-10个氨基酸、1-20个氨基酸、10-50个氨基酸、50-100个氨基酸、100-200个氨基酸、200-300个氨基酸、300-400个氨基酸、400-500个氨基酸、500-600个氨基酸、600-700个氨基酸、700-800个氨基酸、800-900个氨基酸或900-1000个氨基酸。

[0618] 多肽连接基的实例是本领域熟知的。在一个实施方案中,连接基包含序列G_n。连接基可包含序列$(GA)_n$。连接基可包含序列$(GGS)_n$。在其它实施方案中,连接基包含$(GGGS)_n$(SEQ ID NO:20)。在其它实施方案中,连接基包含序列$(GGS)_n(GGGGS)_n$(SEQ ID NO:21)。在这些实例中,n可以是1-100的整数。在其它实例中,n可以是1-20的整数,即1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20。连接基的实例包括但不限于:GGG、SGGSGGS(SEQ ID NO:22)、GGSGGSGGSGGSGGG(SEQ ID NO:23)、GGSGGSGGGGSGGGGS(SEQ ID NO:24)、GGSGGSGGSGGSGGSGGS(SEQ ID NO:25)、GGGGSGGGGSGGGGS(SEQ ID NO:26)、表13中的连接基

(SEQ ID NO:92、93和94)和表14A中的连接基(SEQ ID NO:95、96和97)。连接基不消除或减弱VWF片段活性或因子VIII的凝血活性。任选地,连接基增强VWF片段活性或因子VIII蛋白的凝血活性,如通过进一步减弱空间位阻效应并且使VWF片段或因子VIII部分更易于接触其靶结合位点。

[0619] 在一个实施方案中,用于嵌合蛋白的连接基的长度为15-25个氨基酸。在另一个实施方案中,用于嵌合蛋白连接基的长度为15-20个氨基酸。在一些实施方案中,用于嵌合蛋白的连接基的长度为10-25个氨基酸。在其它实施方案中,用于嵌合蛋白的连接基的长度为15个氨基酸。在其它实施方案中,用于嵌合蛋白的连接基是$(GGGGS)_n$(SEQ ID NO:27),其中G表示甘氨酸,S表示丝氨酸,并且n为1-20的整数。

[0620] E)切割位点

[0621] 连接基也可结合能够被化学(如,酯键的水解)、酶促(即,结合蛋白酶切割序列)或光解(如,发色团例如3-氨基-3-(2-硝基苯基)丙酸(ANP))切割,以便从另一个分子释放一个分子的部分。

[0622] 在一个实施方案中,连接基是可切割的连接基。可切割的连接基可包含N-末端或C-末端或它们二者的一个或多个切割位点。在另一个实施方案中,可切割的连接基基本上由或由一个或多个可切割位点组成。在其它实施方案中,可切割的连接基包含本文所述的异源氨基酸连接基序列或聚合物以及一个或多个可切割位点。

[0623] 在某些实施方案中,可切割的连接基包含一个或多个可在宿主细胞(即,细胞内加工位点)中切割的切割位点。切割位点的非限制性实例包括RRRR(SEQ ID NO:52)、RKRRKR(SEQ ID NO:53)和RRRRS(SEQ ID NO:54)。

[0624] 在其它实施方案中,可切割的连接基包含一个或多个在包含可切割的连接基的嵌合蛋白施用至受试者之后被蛋白酶切割的切割位点。在一个实施方案中,切割位点被选自由以下组成的组的蛋白酶切割:因子XIa、因子XIIa、激肽释放酶、因子VIIa、因子IXa、因子Xa、因子IIa(凝血酶)、弹性蛋白酶-2、MMP-12、MMP-13、MMP-17和MMP-20。在另一个实施方案中,切割位点选自由以下组成的组:FXIa切割位点(如,KLTR↓AET(SEQ ID NO:29))、FXIa切割位点(如,DFTR↓VVG(SEQ ID NO:30))、FXIIa切割位点(如,TMTR↓IVGG(SEQ ID NO:31))、激肽释放酶切割位点(如,SPFR↓STGG(SEQ ID NO:32))、FVIIa切割位点(如,LQVR↓IVGG(SEQ ID NO:33))、FIXa切割位点(如,PLGR↓IVGG(SEQ ID NO:34))、FXa切割位点(如,IEGR↓TVGG(SEQ ID NO:35))、FIIa(凝血酶)切割位点(如,LTPR↓SLLV(SEQ ID NO:36))、弹性蛋白酶-2切割位点(如,LGPV↓SGVP(SEQ ID NO:37))、粒酶-B切割(如,VAGD↓SLEE(SEQ ID NO:38))、MMP-12切割位点(如,GPAG↓LGGA(SEQ ID NO:39))、MMP-13切割位点(如,GPAG↓LRGA(SEQ ID NO:40))、MMP-17切割位点(如,APLG↓LRLR(SEQ ID NO:41))、MMP-20切割位点(如,PALP↓LVAQ(SEQ ID NO:42))、TEV切割位点(如,ENLYFQ↓G(SEQ ID NO:43))、肠激酶切割位点(如,DDDK↓IVGG(SEQ ID NO:44))、蛋白酶3C(PRESCISSION™)切割位点(如,LEVLFQ↓GP(SEQ ID NO:45))和分选酶A切割位点(如,LPKT↓GSES(SEQ ID NO:46))。在某些实施方案中,FXIa切割位点包括但不限于例如TQSFNDFTR(SEQ ID NO:47)和SVSQTSKLTR(SEQ ID NO:48)。非限制性示例性凝血酶切割位点包括例如DFLAEGGGVR(SEQ ID NO:49)、TTKIKPR(SEQ ID NO:50)或LVPRG(SEQ ID NO:55)以及包含、基本上由或由ALRPR组成的序列(如,ALRPRVVGGA(SEQ ID NO:51))。

[0625] 在具体实施方案中，切割位点是TLDPRSFLLRNPNDKYEPFWEDEEK(SEQ ID NO:56)。

[0626] 多核苷酸、载体、宿主细胞及其制备方法

[0627] 本发明还提供编码本文所述的VWF片段的多核苷酸、包含VWF片段和异源部分的嵌合蛋白、包含FVIII蛋白和辅助部分的嵌合蛋白或包含VWF片段和FVIII蛋白的嵌合蛋白。当VWF片段作为单多肽链连接至嵌合蛋白中的异源部分或FVIII蛋白时，本发明涉及编码连接至异源部分或FVIII蛋白的VWF片段的多核苷酸。当嵌合蛋白包含第一和第二多肽链，第一多肽链包含VWF片段和第一异源部分(如，第一Fc区)，并且第二多肽链包含第二异源部分(如，第二Fc区)，其中第一多肽链和第二多肽链彼此缔合时，多核苷酸可包含第一核苷酸序列和第二核苷酸序列。在一个实施方案中，第一核苷酸序列和第二核苷酸序列在相同的多核苷酸上。在另一个实施方案中，第一核苷酸序列和第二核苷酸序列在两个不同的多核苷酸上(如，不同的载体)。在某些实施方案中，本发明涉及一组包含第一核苷酸链和第二核苷酸链的的多核苷酸，其中第一核苷酸链编码嵌合蛋白的VWF片段，并且第二核苷酸链编码FVIII蛋白。

[0628] 在其它实施方案中，所述多核苷酸组还包含编码蛋白质转化酶的另外核苷酸链(如，当嵌合多肽被单条多核苷酸链编码时为第二核苷酸链或当嵌合蛋白被两条多核苷酸链编码时为第三核苷酸链)。蛋白质转化酶可选自由以下组成的组：前蛋白转化酶枯草杆菌蛋白酶/kexin 5型(PCSK5或PC5)、前蛋白转化酶枯草杆菌蛋白酶/kexin 7型(PCSK7或PC5)、酵母Kex 2、前蛋白转化酶枯草杆菌蛋白酶/kexin 3型(PACE或PCSK3)以及它们的两种或多种组合。在一些实施方案中，蛋白质转化酶是PACE、PC5或PC7。在具体实施方案中，蛋白质转化酶是PC5或PC7。参见国际专利申请号PCT/US2011/043568，该申请以引用方式并入本文。在另一个实施方案中，蛋白质转化酶是PACE/弗林蛋白酶。

[0629] 在某些实施方案中，本发明包括一组包含第一核苷酸序列、第二核苷酸序列和第三核苷酸序列的多核苷酸，所述第一核苷酸序列编码包含VWF的D'结构域和D3结构域的VWF片段，所述第二核苷酸序列编码FVIII蛋白，并且所述第三核苷酸序列编码VWF的D1结构域和D2结构域。在该实施方案中，D1结构域和D2结构域独立地表达(不连接至VWF片段的D'D3结构域)以便形成正确的二硫键以及D'D3结构域折叠。D1D2结构域表达可以是顺式的或反式的。

[0630] 如本文所用，表达载体是指包含用于插入的编码序列的转录和翻译的必要元件，或在RNA病毒载体的情况下用于引入适当的宿主细胞时复制和翻译的必要元件的任何核酸构建体。表达载体可包括质粒、噬菌粒、病毒及其衍生物。

[0631] 本发明的表达载体将包括编码VWF片段的多核苷酸或包含VWF片段的嵌合蛋白。

[0632] 在一个实施方案中，VWF片段的编码序列、第二异源部分(如，第二Fc区)或FVIII蛋白可操作地连接至表达调控序列。如本文所用，当两个核酸序列共价连接以允许每个组件核酸序列保持其功能时它们是可操作地连接的。当编码序列和基因表达控制序列共价连接时，以将编码序列的表达或转录和/或翻译置于基因表达控制序列的影响或控制下，据说它们是可操作地连接的。据说，如果5'基因表达序列中的启动子的诱导导致编码序列的转录，并且如果两个DNA序列之间的连接的性质不(1)导致引入移码突变，(2)妨碍启动子区指导编码序列转录的能力或(3)妨碍对应的RNA转录物翻译为蛋白质的能力，则两个DNA序列是可操作地连接的。因此，如果基因表达序列能够影响该编码核酸序列的转录，使得所得的转

录物翻译为所需蛋白质或多肽,则基因表达序列可操作地连接至编码核酸序列。

[0633] 如本文所用,基因表达控制序列是任何调控核苷酸序列,例如启动子序列或启动子-增强子组合,它们促进与其可操作连接的编码核酸的有效转录和翻译。基因表达控制序列可以是例如哺乳动物或病毒启动子,例如组成型或诱导型启动子。组成型哺乳动物启动子包括但不限于如下基因的启动子:次黄嘌呤磷酸核糖基转移酶(HPRT)、腺苷脱氨酶、丙酮酸激酶、β-肌动蛋白启动子以及其它组成型启动子。在真核细胞中发挥组成型作用的示例性病毒启动子包括例如来自巨细胞病毒(CMV)、猿猴病毒(如,SV40)、乳头状瘤病毒、腺病毒、人免疫缺陷病毒(HIV)、劳氏肉瘤病毒、巨细胞病毒、莫洛尼氏白血病病毒的长末端重复(LTR)以及其它逆转录病毒的启动子,以及单纯疱疹病毒的胸苷激酶启动子。其它组成型启动子是本领域普通技术人员已知的。用作本发明的基因表达序列的启动子还包括诱导型启动子。诱导型启动子在诱导剂存在下表达。例如,金属硫蛋白启动子在某些金属离子存在下被诱导以促进转录和翻译。其它诱导型启动子是本领域普通技术人员已知的。

[0634] 一般来讲,基因表达控制序列应包括(如果必要)分别参与转录和翻译起始的5'非转录和5'非翻译序列,例如TATA框、加帽序列、CAAT序列等。特别地,此类5'非转录序列将包括启动子区,所述启动子区包括用于可操作地连接的编码核酸的转录控制的启动子序列。基因表达序列任选地根据需要包括增强子序列或上游激活序列。

[0635] 病毒载体包括但不限于来自如下病毒的核酸序列:逆转录病毒例如莫洛尼氏鼠白血病病毒、哈维鼠肉瘤病毒、鼠乳腺肿瘤病毒和劳氏肉瘤病毒;腺病毒、腺相关病毒;SV40-型病毒;多瘤病毒;爱泼斯坦-巴尔病毒(Epstein-Barr virus);乳头状瘤病毒、疱疹病毒;牛痘病毒;脊髓灰质炎病毒;以及RNA病毒例如逆转录病毒。可轻松地采用本领域熟知的其它载体。某些病毒载体基于非细胞病变真核细胞病毒,其中非必需基因被所关注的基因替代。非细胞病变病毒包括逆转录病毒,其生活周期涉及基因组病毒RNA逆转录为DNA,随后原病毒整合到宿主细胞DNA。逆转录病毒已经批准用于人类基因治疗试验。最有用的是那些复制缺陷型逆转录病毒(即,能够指导所需蛋白质的合成,但不能制备感染性颗粒)。此类遗传改变的逆转录病毒表达载体通常用于基因的体内高效转导。用于制备复制缺陷型逆转录病毒的标准方案(包括如下步骤:将外源性遗传物质掺入质粒、用质粒转染包装细胞系、通过包装细胞系产生重组逆转录病毒、从组织培养基收集病毒颗粒以及用病毒颗粒感染靶细胞)在Kriegler,M.,Gene Transfer and Expression,A Laboratory Manual,W.H.Freeman Co.,New York(1990)和Murry,E.J.,Methods in Molecular Biology,第7卷,Humana Press,Inc.,Cliffton,N.J.(1991)中提供。

[0636] 在一个实施方案中,病毒是腺相关病毒,其为双链DNA病毒。腺相关病毒可工程化为复制缺陷型,并且能够感染多种细胞类型和物种。它还具有多个优点,例如热和脂溶剂稳定性;各种谱系细胞包括造血细胞中的高转导频率;以及缺乏重复感染抑制从而允许多个系列转导。据报道,腺相关病毒可以位点特异性方式整合进人细胞DNA,从而使逆转录病毒感染特征性的插入诱变可能性和插入基因表达可变性降至最小。此外,野生型腺相关病毒感染在不存在选择压力的情况下在组织培养基中传代大于100次,意味着腺相关病毒基因组整合是相对稳定的事件。腺相关病毒也可以染色体外的方式发挥功能。

[0637] 其它载体包括质粒载体。质粒载体在本领域中广泛描述,并且是本领域技术人员熟知的。参见例如Sambrook等,Molecular Cloning:A Laboratory Manual,第二版,Cold

Spring Harbor Laboratory Press,1989。最近几年发现,质粒载体对于在体内将基因递送至细胞特别有利,因为其不能在宿主基因组内复制并整合进宿主基因组。然而,这些具有与宿主细胞兼容的启动子的质粒可由在质粒内可操作地编码的基因表达肽。一些可商购获得的常用质粒包括pBR322、pUC18、pUC19、各种pcDNA质粒、pRC/CMV、各种pCMV质粒、pSV40和pBlueScript。具体质粒的另外实例包括pcDNA3.1,目录号V79020;pcDNA3.1/hygro,目录号V87020;pcDNA4/myc-His,目录号V86320;以及pBudCE4.1,目录号V53220,它们均得自Invitrogen(Carlsbad,CA)。其它质粒是本领域普通技术人员熟知的。另外,质粒可使用标准分子生物学技术定制设计,以移除和/或添加特定DNA片段。

[0638] 在一种可用于制备本发明的蛋白质的昆虫表达系统中,苜蓿银纹夜蛾(Autographa californica)核多角体病毒(AcNPV)用作表达外源基因的载体。病毒在草地贪夜蛾(Spodoptera frugiperda)细胞中生长。编码序列可克隆至病毒的非必需区(例如,多角体基因),并且置于ACNPV启动子(例如,多角体启动子)的控制下。编码序列的成功插入将导致多角体基因的失活和非包涵型重组病毒(即,缺乏多角体基因编码的蛋白质性外壳的病毒)的生成。然后,这些重组病毒用于感染插入基因在其中表达的草地贪夜蛾细胞。(参见例如Smith等(1983)J Virol 46:584;美国专利号4,215,051)。该表达系统的其它实例可见于Ausubel等编著(1989)Current Protocols in Molecular Biology,第2卷,Greene Publish.Assoc.&Wiley Interscience。

[0639] 可用于表达本发明的蛋白质的另一个系统是谷氨酰胺合成酶基因表达系统,也称为"GS表达系统"(Lonza Biologics PLC,Berkshire UK)。该表达系统在美国专利号5,981,216中详细描述。

[0640] 在哺乳动物宿主细胞中,可使用多个基于病毒的表达系统。在其中腺病毒用作表达载体的情况下,编码序列可连接至腺病毒转录/翻译控制复合物,如晚期启动子和三重前导序列。然后,该嵌合基因可通过体外或体内重组插入腺病毒基因组。插入病毒基因组的非必需区(如,E1或E3区)将产生可存活并且能够在感染宿主中表达肽的重组病毒。参见例如Logan&Shenk(1984)Proc Natl Acad Sci USA 81:3655。或者,可使用牛痘7.5K启动子。参见例如Mackett等(1982)Proc Natl Acad Sci USA 79:7415;Mackett等(1984)J Virol 49:857;Panicali等(1982)Proc Natl Acad Sci USA 79:4927。

[0641] 为增加制备效率,多核苷酸可设计为编码本发明蛋白质的多个由酶切割位点分隔的单元。可切割所得的多肽(如,用适当的酶处理)以回收多肽单元。这可增加单个启动子驱动的多肽的产量。当用于适当病毒表达系统时,由mRNA编码的每条多肽的翻译在转录物内部引导;如通过内部核糖体进入位点IRES。因此,多顺反子构建体指导单个大多顺反子mRNA的转录,继而指导多个单条多肽的翻译。该方法消除了多蛋白的产生和酶促加工,并且可显著增加由单个启动子驱动的多肽的产量。

[0642] 用于转化的载体将通常包含用于识别转化子的选择性标记。在细菌系统中,这可包括抗生素抗性基因,例如氨苄青霉素或卡那霉素。用于培养的哺乳动物细胞的选择性标记包括赋予对药物例如新霉素、潮霉素和甲氨蝶呤抗性的基因。选择性标记可以是可扩增的选择性标记。一个可扩增的选择性标记是二氢叶酸还原酶(DHFR)基因。Simonsen C C等(1983)Proc Natl Acad Sci USA 80:2495-9。选择性标记在Thilly(1986)Mammalian Cell Technology,Butterworth Publishers,Stoneham,Mass.中有所评述,并且选择性标记的选

择在本领域的普通技术人员的水平内。

[0643] 选择性标记可在单独的质粒上与所关注的基因同时引入细胞，或它们可在相同的质粒上引入。如果在相同的质粒上，选择性标记和所关注的基因可在不同启动子或相同启动子的控制下，后一个排列产生双顺反子信息。该类型的构建体是本领域已知的(例如，美国专利号4,713,339)。

[0644] 表达载体可编码允许易于纯化重组产生的蛋白质的标签。实例包括但不限于：载体pUR278(Ruther等(1983)EMBO J 2:1791)，其中要表达的蛋白质的编码序列可连接至载体的lac z编码区框内，以生成标签融合蛋白；pGEX载体可用于表达本发明的具有谷胱甘肽S-转移酶(GST)标签的蛋白质。这些蛋白质通常是可溶性的，并且可通过吸附到谷胱甘肽-琼脂糖珠粒，然后在谷胱甘肽存在下洗脱而轻松地从细胞纯化。载体包括纯化后易于移除标签的切割位点(凝血酶或因子Xa蛋白酶或PRESCISSION PROTEASE™(Pharmacia,Peapack,N.J.))。

[0645] 然后，表达载体转染或共转染至合适的靶细胞，该靶细胞表达多肽。本领域已知的转染技术包括但不限于：磷酸钙沉淀(Wigler等(1978)Cell 14:725)、电穿孔(Neumann等(1982)EMBO J 1:841)和基于脂质体的试剂。可使用多种宿主表达载体系统表达本文所述的蛋白质，包括原核和真核细胞二者。这些包括但不限于：微生物例如转化有包含适当编码序列的重组噬菌体DNA或质粒DNA表达载体的细菌(如，大肠杆菌)；转化有包含适当编码序列的重组酵母或真菌表达载体的酵母或丝状真菌；转染有包含适当编码序列的重组病毒表达载体(如，杆状病毒)的昆虫细胞系统；转染有重组病毒表达载体(如，花椰菜花叶病毒或烟草花叶病毒)或转化有包含适当编码序列的重组质粒表达载体(如，Ti质粒)的植物细胞系统；或动物细胞系统，包括哺乳动物细胞(如，HEK 293、CHO、Cos、HeLa、HKB11和BHK细胞)。

[0646] 在一个实施方案中，宿主细胞是真核细胞。如本文所用，真核细胞是指具有明确细胞核的任何动物或植物细胞。动物的真核细胞包括脊椎动物如哺乳动物细胞和无脊椎动物如昆虫的细胞。植物的真核细胞可特别地包括但不限于酵母细胞。真核细胞不同于原核细胞，如细菌。

[0647] 在某些实施方案中，真核细胞是哺乳动物细胞。哺乳动物细胞是来源于哺乳动物的任何细胞。哺乳动物细胞特别地包括但不限于哺乳动物细胞系。在一个实施方案中，哺乳动物细胞是人细胞。在另一个实施方案中，哺乳动物细胞是HEK 293细胞，它是人胚肾细胞系。HEK 293细胞以CRL-1533得自美国菌种保藏中心(American Type Culture Collection,Manassas,VA)，以及以293-H细胞(目录号11631-017)或293-F细胞(目录号11625-019)得自Invitrogen(Carlsbad,Calif.)。在一些实施方案中，哺乳动物细胞是PER.C6®细胞，它是来源于视网膜的人细胞系。PER.C6®细胞得自Crucell(Leiden,The Netherlands)。在其它实施方案中，哺乳动物细胞是中国仓鼠卵巢(CHO)细胞。CHO细胞得自美国菌种保藏中心(American Type Culture Collection,Manassas,VA.)(如，CHO-K1、CCL-61)。在其它实施方案中，哺乳动物细胞是幼仓鼠肾(BHK)细胞。BHK细胞得自美国菌种保藏中心(American Type Culture Collection,Manassas,Va.)(如，CRL-1632)。在一些实施方案中，哺乳动物细胞是HKB11细胞，它是HEK 293细胞和人B细胞系的杂合细胞系。Mei等,Mol.Biotechnol.34(2):165-78(2006)。

[0648] 在一个实施方案中，编码本发明的VWF片段或嵌合蛋白的质粒还包括选择性标记，

如zeocin抗性,并且转染至HEK 293细胞,用于纯化VWF片段或嵌合蛋白。

[0649] 在另一个实施方案中,将包含因子VIII-Fc融合编码序列和第一选择性标记如zeocin抗性基因的第一质粒和包含VWF片段-Fc编码序列和第二选择性标记如新霉素抗性基因的第二质粒共转染至HEK 293细胞,用于纯化因子VIII-Fc和VWF-Fc杂合物。第一和第二质粒可等量(即,1:1比例)引入,或它们可不等量引入。

[0650] 在一些实施方案中,包括因子VIII-Fc融合编码序列和第一选择性标记如zeocin抗性基因的第一质粒和包括VWF片段-Fc编码序列和第二选择性标记如新霉素抗性基因的第二质粒以及包括蛋白质转化酶编码序列(如,PC5或弗林蛋白酶)和第三选择性标记如潮霉素抗性基因的第三质粒共转染至HEK 293细胞,用于纯化因子VIII-VWF片段杂合物。第一和第二质粒可等量(即,1:1摩尔比)引入,或它们可不等量引入。在某些实施方案中,包括因子VIII-Fc融合编码序列、VWF片段-Fc编码序列和第一选择性标记如zeocin抗性基因的第一质粒和包括蛋白质转化酶编码序列(如,PC5或弗林蛋白酶)和第二选择性标记如潮霉素抗性基因的第二质粒共转染至HEK 293细胞,用于纯化因子VIII-VWF片段杂合物。在一个实施方案中,编码FVIII-Fc序列和VWF片段-Fc序列的核苷酸序列可连接,以编码一个单条多肽。在另一个实施方案中,编码FVIII-Fc序列和VWF片段-Fc序列的核苷酸序列可编码为两条多肽链。因子VIII-Fc融合编码序列和VWF片段-Fc编码序列的启动子可以是不同的,或它们可以是相同的。

[0651] 在一些实施方案中,包含弗林蛋白酶的质粒与包含因子VIII-Fc编码序列和/或VWF片段-Fc编码序列的质粒共转染。在一些实施方案中,弗林蛋白酶蛋白质在包含因子VIII-Fc融合编码序列的相同质粒上。在一些实施方案中,弗林蛋白酶蛋白质在包含VWF片段-Fc编码序列的相同质粒上。在一些实施方案中,弗林蛋白酶蛋白质在单独的质粒上。

[0652] 在其它实施方案中,转染细胞是稳定转染的。可使用本领域技术人员已知的常规技术选择这些细胞并保持为稳定的细胞系。

[0653] 使包含蛋白质的DNA构建体的宿主细胞在适当的生长培养基中生长。如本文所用,术语“适当的生长培养基”意指包含细胞生长所需的营养物质的培养基。细胞生长所需的营养物质可包括碳源、氮源、必需氨基酸、维生素、矿物质和生长因子。任选地,培养基可包含一种或多种选择因子。任选地,培养基可包含牛血清或胎牛血清(FCS)。在一个实施方案中,培养基基本上不包含IgG。生长培养基将通常通过例如药物选择或必需营养物质的缺失来选择包含DNA构建体的细胞,所述缺失被DNA构建体上的选择性标记补充或用DNA构建体共转染。培养的哺乳动物细胞通常在市售含血清或无血清培养基(如,MEM、DMEM、DMEM/F12)中生长。在一个实施方案中,培养基是CD293(Invitrogen,Carlsbad,CA.)。在另一个实施方案中,培养基是CD17(Invitrogen,Carlsbad,CA.)。适于所用的具体细胞系的培养基的选择在本领域的普通技术人员的水平内。

[0654] 为了共表达VWF片段和第二异源部分或FVIII蛋白,将宿主细胞在允许VWF片段和第二异源部分或FVIII蛋白二者表达的条件下培养。如本文所用,培养是指在体外保持活细胞至少一定的时间。保持可以是但不必包括活细胞群的增加。例如,保持在培养物中的细胞在群中可以是静止的,但仍然是活的,并且能够生成所需的产物,如重组蛋白或重组融合蛋白。用于培养真核细胞的合适条件是本领域熟知的,并且包括培养基、培养基补充剂、温度、pH、氧饱和度等的适当选择。出于商业目的,培养可包括使用各种类型的放大系统中的任何

一种，包括摇瓶、滚瓶、中空纤维生物反应器、搅拌槽生物反应器、气升式生物反应器、Wave生物反应器等。

[0655] 也可选择细胞培养条件以允许VWF片段与第二异源部分或FVIII蛋白缔合。允许VWF片段和/或FVIII蛋白表达的条件可包括存在维生素K源。例如，在一个实施方案中，将稳定转染的HEK 293细胞在补充有4mM谷氨酰胺的CD293培养基(Invitrogen,Carlsbad,CA)或OptiCHO培养基(Invitrogen,Carlsbad,CA)中培养。

[0656] 在一个方面，本发明涉及表达、制备或生成本发明的VWF片段的方法，其包括a)用编码VWF片段的多核苷酸转染宿主细胞，以及b)在培养基中适于表达VWF片段的条件下培养宿主细胞，其中VWF片段得以表达。在一个实施方案中，本发明涉及生成成熟VWF蛋白或其片段的方法，其包括a)用编码融合至VWF的原肽的VWF蛋白或其片段的第一多核苷酸和编码蛋白质转化酶如PC5、PC7或弗林蛋白酶的第二多核苷酸转染宿主细胞，以及b)在培养基中适于表达成熟VWF蛋白或其片段的条件下培养宿主细胞。编码VWF蛋白或其片段的多核苷酸也可融合至VWF的前肽。前肽序列可在分泌前插入内质网期间被切割。

[0657] 在另一个方面，本发明涉及表达、制备或生成包含连接至异源部分或FVIII蛋白或与之缔合的VWF片段的嵌合蛋白的方法，其包括a)用编码嵌合蛋白的多核苷酸或多核苷酸组转染一个或多个宿主细，胞以及b)在培养基中适于表达嵌合蛋白的条件下培养宿主细胞。在一个实施方案中，本发明涉及表达、制备或生成嵌合蛋白的方法，其包括a)用编码连接至异源部分的VWF片段的第一多核苷酸和编码连接至异源部分的FVIII蛋白的第二多核苷酸转染宿主细胞，以及b)在培养基中适于表达嵌合蛋白的条件下培养宿主细胞。第一多核苷酸和第二多核苷酸可在一个载体或两个载体中。在另一个实施方案中，本发明涉及表达、制备或生成嵌合蛋白的方法，其包括a)用编码连接至异源部分的VWF片段的第一多核苷酸、编码连接至异源部分的FVIII蛋白的第二多核苷酸以及编码蛋白质转化酶的第三多核苷酸转染宿主细胞，以及b)在培养基中适于表达嵌合蛋白的条件下培养宿主细胞。在其它实施方案中，本发明涉及表达、制备或生成嵌合蛋白的方法，其包括a)用编码包含连接至异源部分的D'结构域和D3结构域的VWF片段的第一多核苷酸、编码连接至异源部分的FVIII蛋白的第二多核苷酸以及编码VWF的D1结构域和D2结构域的第三多核苷酸转染宿主细胞，以及b)在培养基中适于表达嵌合蛋白的条件下培养宿主细胞。在一个实施方案中，第一多核苷酸、第二多核苷酸和第三多核苷酸可在一个载体或单独的载体中。在另一个实施方案中，第一多核苷酸和第二多核苷酸可在一个载体中，并且第三多核苷酸可在另一个载体中。在其它实施方案中，第一多核苷酸和第三多核苷酸可在一个载体中，并且第二多核苷酸可在另一个载体中。在一些实施方案中，第二多核苷酸和第三多核苷酸可在一个载体中，并且第一多核苷酸可在另一个载体中。

[0658] 在另外的实施方案中，包含VWF片段的蛋白质产物或包含VWF片段的嵌合蛋白分泌至培养基。培养基从细胞分离、浓缩、过滤，并且然后通过两个或三个亲合柱如蛋白A柱和一个或两个阴离子交换柱。

[0659] 在某些方面，本发明涉及由本文所述的方法制备的VWF片段或嵌合多肽。

[0660] 体外制备允许放大得到大量所需的本发明改变的多肽。在组织培养条件下培养哺乳动物细胞的技术是本领域已知的，并且包括均匀悬浮培养，如在气升式反应器或连续搅拌反应器中，或固定化或包埋细胞培养，如在中空纤维、微胶囊中、琼脂糖微珠粒或陶瓷料

筒上。如有必要和/或需要，可通过常规色谱法，例如凝胶过滤、离子交换色谱法、疏水相互作用色谱法(HIC、DEAE-纤维素色谱法或亲合色谱法纯化多肽的溶液。

[0661] 药物组合物

[0662] 包含VWF片段或本发明的嵌合蛋白的组合物可包含合适的可药用载体。例如，它们可包含促进活性化合物加工为制剂的赋形剂和/或辅助剂，所述制剂设计用于递送至作用位点。

[0663] 药物组合物可配制为通过弹丸式注射(bolus injection)肠胃外施用(即，静脉内、皮下或肌肉内)。注射制剂可以单位剂型存在，如在添加防腐剂的安瓿瓶或多剂量容器中。组合物可采取例如悬浮液、溶液或油性或水性媒介物中的乳液的形式，并且包含配制剂，例如悬浮、稳定和/或分散剂。或者，活性成分可以是用合适的媒介物如无热原水配制的粉末形式。

[0664] 合适的肠胃外施用制剂还包括水溶性形式的活性化合物例如水溶性盐的水溶液。此外，可施用活性化合物的悬浮液，作为适当的油性注射悬浮液。合适的亲脂性的溶剂或媒介物包括脂油例如芝蔴油或合成脂肪酸酯例如油酸乙酯或三酸甘油酯。水性注射悬浮液可包含增加悬浮液的粘度的物质，包括例如羧甲基纤维素钠、山梨糖醇和葡聚糖。任选地，悬浮液还可包含稳定剂。脂质体也可用于包封本发明的分子，用于递送至细胞或组织间隙。示例性可药用载体是生理相容性溶剂、分散介质、涂料、抗细菌剂和抗真菌剂、等渗剂和吸收延迟剂、水、盐水、磷酸盐缓冲盐水、右旋糖、甘油、乙醇等。在一些实施方案中，组合物包含等渗剂，例如糖、多元醇例如甘露糖醇、山梨糖醇或氯化钠。在其它实施方案中，组合物包含可药用物质，例如润湿剂或微量辅助性物质例如润湿或乳化剂、防腐剂或缓冲剂，它们提高了活性成分的储存寿命或有效性。

[0665] 本发明的组合物可具有各种形式，包括例如液体(如，可注射的和不可灌注的溶液)、分散体、悬浮液、半固体和固体剂型。优选的形式取决于施用和治疗应用模式。

[0666] 组合物可配制为溶液、微乳液、分散体、脂质体或其它适于高药物浓度的有序结构。无菌注射溶液可通过将所需量溶于适当溶剂的活性成分掺入上面列举的一种成分或成分的组合(根据需要)，然后过滤除菌来制备。一般来讲，分散体通过将活性成分掺入包含基本分散介质和所需的上面列举的其它成分的无菌媒介物来制备。就用于制备无菌注射溶液的无菌粉末而言，优选的制备方法是真空干燥和冷冻干燥，产生活性成分加上预先无菌过滤溶液的任何另外所需成分的粉末。可以例如通过使用涂料例如卵磷脂，就分散体而言通过保持所需的粒度以及通过使用表面活性剂来保持溶液的适当流动性。可通过将吸收延迟剂例如单硬脂酸盐和明胶包括于组合物中而产生注射组合物的长期吸收。

[0667] 活性成分可用控释制剂或装置配制。此类制剂和装置的实例包括植入物、透皮贴剂和微包封递送系统。可使用生物可降解的生物相容性聚合物，例如乙烯-醋酸乙烯酯、聚酸酐、聚乙醇酸、胶原、聚原酸酯和聚乳酸。制备此类制剂和装置的方法是本领域已知的。参见例如Sustained and Controlled Release Drug Delivery Systems,J.R.Robinson,编著,Marcel Dekker,Inc.,New York,1978。

[0668] 注射储存制剂可通过在生物可降解聚合物例如聚交酯-聚乙交酯中形成药物的微包封基质来制备。取决于药物与聚合物的比率以及所用的聚合物的性质，可控制药物释放速率。其它示例性生物可降解聚合物为聚原酸酯和聚酸酐。储存注射制剂也可通过将药物

包封于脂质体或微乳液中来制备。

[0669] 补充活性化合物可掺入组合物中。在一个实施方案中，本发明的VWF片段或嵌合蛋白与另一种凝血因子或其变体、片段、类似物或衍生物一起配制。例如，凝血因子包括但不限于：因子V、因子VII、因子VIII、因子IX、因子X、因子XI、因子XII、因子XIII、凝血酶原、血纤维蛋白原、血管性血友病因子或重组可溶性组织因子(rsTF)或任何上述的活性形式。止血剂的凝血因子也可包括抗纤溶药物，如ε-氨基己酸、氨甲环酸。

[0670] 可调整给药方案以提供最佳所需响应。例如，可施用单次推注，可随时间推移施用多个分剂量，或可根据治疗情况紧急性指示按比例减少或增加剂量。出于易于施用和剂量均匀的目的，以单位剂型配制肠胃外组合物是有利的。参见例如Remington's Pharmaceutical Sciences(Mack Pub.Co.,Easton,Pa.1980)。

[0671] 除活性化合物之外，液体剂型可包含惰性成分，例如水、乙醇、碳酸乙酯、乙酸乙酯、苄醇、苯甲酸苄酯、丙二醇、1,3-丁二醇、二甲基甲酰胺、油、甘油、四氢糠醇、聚乙二醇和脱水山梨糖醇的脂肪酸酯。

[0672] 合适的药用载体的非限制性实例还在E.W.Martin的Remington's Pharmaceutical Sciences中有所描述。赋形剂的一些实例包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、大米、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石粉、氯化钠、脱脂奶粉、甘油、丙烯、乙二醇、水、乙醇等。组合物也可包含pH缓冲剂和润湿或乳化剂。

[0673] 对于口服施用，药物组合物可采取通过常规方式制备的片剂或胶囊剂的形式。组合物也可以液体例如糖浆或悬浮液制备。液体可包括悬浮剂(如，山梨糖醇糖浆、纤维素衍生物或氢化食用脂)、乳化剂(卵磷脂或阿拉伯树胶)、非水性媒介物(如，杏仁油、油酯、乙醇或分馏植物油)和防腐剂(如，对羟基苯甲酸甲酯或丙酯或山梨酸甲)。制剂也可包括调味剂、着色剂和甜味剂。或者，组合物可以通过水或另一种合适的媒介物配制的干燥产品存在。

[0674] 对于口腔含化施用，组合物可采取根据常规方案的片剂或锭剂的形式。

[0675] 对于通过吸入施用，根据本发明使用的化合物以含或不含赋形剂的雾化气雾剂的形式或来自加压包或喷雾器的气雾喷剂的形式，任选地与推进剂如二氟二氯甲烷、一氟三氯甲烷、四氟二氯甲烷、二氧化碳或其它合适的气体一起便利地递送。就加压气雾剂而言，剂量单位可通过提供递送定量的阀门确定。用于吸入器或吹药器的例如明胶的胶囊剂和料筒可配制为包含化合物的粉末混合物和合适的粉末基料例如乳糖或淀粉。

[0676] 药物组合物也可配制为用于直肠施用的栓剂或保留灌肠剂，如包含常规栓剂基料例如可可油或其它甘油酯。

[0677] 基因治疗

[0678] 本发明的VWF片段或其嵌合蛋白可在哺乳动物如人患者体内生成，使用基因治疗方法治疗选自由以下组成的组的出血性疾病或病症：出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血以及髂腰肌鞘出血将有有益的治疗效果。在一个实施方案中，出血性疾病或病症是血友病。在另一个实施方案中，出血性疾病或病症是甲型血友病。这涉及可操作地连接至合适的表达控制序列的合适的VWF片段或嵌合蛋白编码核酸的施用。在某些实施方案中，这些序列掺入病毒载体。此

类基因治疗的合适病毒载体包括腺病毒载体、慢病毒载体、杆状病毒载体、爱泼斯坦-巴尔病毒载体、乳头多瘤空泡病毒载体、牛痘病毒载体、单纯疱疹病毒载体和腺相关病毒(AAV)载体。病毒载体可以是复制缺陷型病毒载体。在其它实施方案中，腺病毒载体具有其E1基因或E3基因的缺失。当使用腺病毒载体时，哺乳动物可以不暴露至编码选择性标记基因的核酸。在其它实施方案中，序列掺入本领域技术人员已知的非病毒载体。

[0679] VWF片段或嵌合蛋白的使用方法

[0680] 本发明的一个方面涉及通过阻断或屏蔽内源性VWF结合FVIII上的VWF结合位点来阻止或抑制FVIII与内源性VWF的相互作用。在一个实施方案中，本发明涉及构建半衰期长于野生型FVIII或FVIII单体-二聚体杂合物的FVIII蛋白的方法，所述方法包括使辅助部分与FVIII蛋白共价缔合，从而生成包含FVIII蛋白和辅助部分的嵌合蛋白，其中辅助部分屏蔽或抑制FVIII蛋白与内源性VWF的相互作用。用于该方法的嵌合蛋白包括本文所述的任何一种或多种嵌合蛋白。

[0681] 本发明的另一个方面包括给需要其的受试者施用半衰期长于野生型FVIII或FVIII单体-二聚体杂合物的FVIII蛋白的方法，所述FVIII蛋白由两条多肽链组成，第一链由编码FVIII和Fc区的氨基酸序列组成，并且第二链由Fc区组成，其中所述方法包括给受试者施用本文所述的VWF片段或本文所述的嵌合蛋白。单体-二聚体杂合物中的FVIII氨基酸序列可以是SQ FVIII或野生型FVIII。

[0682] 在一个实施方案中，本发明涉及使用辅助部分，如本文所述的VWF片段或包含VWF片段的嵌合蛋白阻止或抑制内源性VWF与FVIII蛋白相互作用的方法。在另一个实施方案中，能够与VWF片段相互作用的FVIII蛋白是内源性FVIII。在其它实施方案中，能够与VWF片段相互作用的FVIII蛋白是FVIII组合物，其在之前或之后独立地施用给受试者或与VWF片段或包含VWF片段的嵌合蛋白同时施用给受试者。在其它实施方案中，能够结合至VWF片段的FVIII蛋白是FVIII组合物，其与VWF片段或嵌合蛋白一起施用给受试者。在其它实施方案中，能够结合至VWF片段的FVIII蛋白是与VWF片段一起存在或与嵌合蛋白中的VWF片段缔合的FVIII。VWF片段或包含VWF片段的嵌合蛋白连接至FVIII蛋白或与之缔合，从而延长结合至VWF片段或嵌合蛋白的FVIII蛋白的半衰期。结合至VWF片段或嵌合蛋白的FVIII蛋白被屏蔽或保护不被从VWF的清除途径清除，并且因此与未结合至VWF片段或嵌合蛋白的FVIII蛋白相比清除减少。因此，屏蔽的FVIII蛋白具有比未连接至VWF片段或嵌合蛋白或与之缔合的FVIII蛋白更长的半衰期。在某些实施方案中，与本发明的VWF片段或嵌合蛋白缔合或受其保护的FVIII蛋白不被VWF清除受体清除。在其它实施方案中，与VWF片段或嵌合蛋白缔合或受其保护的FVIII蛋白从系统清除速度慢于不与VWF片段缔合或不受其保护的FVIII蛋白。

[0683] 在一个方面，本发明的VWF片段或包含其的嵌合蛋白与不包含VWF清除受体结合位点的VWF片段或嵌合蛋白相比，减少了从循环的清除。VWF片段阻止或抑制通过VWF清除途径从系统清除连接至VWF片段或与之缔合的FVIII。用于本发明的VWF片段还可提供由内源性VWF提供的至少一种或多种VWF样FVIII保护特性。在某些实施方案中，VWF片段还可掩蔽一个或多个FVIII清除受体结合位点，从而阻止FVIII被其自身清除途径清除。

[0684] 在另一个方面，本发明的VWF片段或嵌合蛋白可用于治疗或阻止与2N型血管性血友病(VWD)相关的疾病或病症。2N型VWD是由结合FVIII的缺陷型VWF导致的定性VWF缺陷，并

且从而导致循环FVIII的低水平。因此，本发明的VWF片段或嵌合蛋白通过结合至FVIII蛋白或被其结合不仅稳定了FVIII蛋白，而且阻止了FVIII蛋白从循环的清除。

[0685] 在一些实施方案中，通过VWF片段或嵌合蛋白防止或抑制FVIII蛋白结合内源性VWF可以是体外的或体内的。

[0686] 还提供了增加FVIII蛋白的半衰期的方法，其包括给需要其的受试者施用包含VWF片段和FVIII蛋白的VWF片段或嵌合蛋白。结合至全长VWF或与之缔合的非活化FVIII在血浆中的半衰期为约12至14小时。在其中循环中几乎不存在VWF的3型VWD中，FVIII的半衰期仅为约六小时，导致由于FVIII浓度的降低此类患者中存在轻度至中度甲型血友病症状。与结合至全长VWF或与之缔合的非活化FVIII的半衰期相比，连接至本发明的VWF片段或与之缔合的FVIII蛋白的半衰期可增加至少约1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2.0倍、2.1倍、2.2倍、2.3倍、2.4倍、2.6倍、2.7倍、2.8倍、2.9倍、3.0倍、3.1倍、3.2倍、3.3倍、3.4倍、3.5倍、3.6倍、3.7倍、3.8倍、3.9倍或4.0倍。在一个实施方案中，与结合至全长VWF或与之缔合的非活化FVIII的半衰期相比，连接至嵌合蛋白中VWF片段或与之缔合的FVIII蛋白的半衰期增加至少约2倍、2.5倍、3.0倍、3.5倍、4.0倍、4.5倍、5.0倍、5.5倍、6.0倍、7倍、8倍、9倍或10倍。在另一个实施方案中，与结合至全长VWF或与之缔合的非活化FVIII的半衰期相比，连接至嵌合蛋白中VWF片段或与之缔合的FVIII蛋白的半衰期增加约2至约5倍、约3至约10倍、约5至约15倍、约10至约20倍、约15至约25倍、约20至约30倍、约25至约35倍、约30至约40倍、约35至约45倍。在具体实施方案中，与FVIII和VWF双敲除小鼠中野生型FVIII的半衰期相比，连接至嵌合蛋白中VWF片段或与之缔合的FVIII蛋白的半衰期增加至少约30倍、31倍、32倍、33倍、34倍、35倍、36倍、37倍、38倍、39倍或40倍。在一些实施方案中，包含融合至第一异源部分如第一Fc区的VWF片段和连接至第二异源部分如第二Fc区的FVIII蛋白的嵌合蛋白的半衰期长于包含FVIII蛋白和两个Fc区的嵌合蛋白的半衰期，其中FVIII蛋白连接至两个Fc区中的一者(即，FVIII单体-二聚体杂合物)。在其它实施方案中，包含融合至第一异源部分如第一Fc区的VWF片段和连接至第二异源部分如第二Fc区的FVIII蛋白的嵌合蛋白的半衰期是包含FVIII蛋白和两个Fc区的嵌合蛋白的半衰期的至少约1.5倍、2倍、2.5倍、3.5倍、3.6倍、3.7倍、3.8倍、3.9倍、4.0倍、4.5倍或5.0倍，其中FVIII蛋白连接至两个Fc区中的一者(即，FVIII单体-二聚体杂合物)。

[0687] 在一些实施方案中，作为本发明的结果，与无VWF片段的FVIII蛋白或野生型FVIII相比，FVIII蛋白的半衰期延长。FVIII蛋白的半衰期比无VWF片段的FVIII蛋白的半衰期长至少约1.5倍、至少约2倍、至少约2.5倍、至少约3倍、至少约4倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约11倍或至少约12倍。在一个实施方案中，FVIII的半衰期比野生型FVIII的半衰期长约1.5倍至约20倍、约1.5倍至约15倍或约1.5倍至约10倍。在另一个实施方案中，与野生型FVIII或无VWF片段的FVIII蛋白相比，FVIII的半衰期延长约2倍至约10倍、约2倍至约9倍、约2倍至约8倍、约2倍至约7倍、约2倍至约6倍、约2倍至约5倍、约2倍至约4倍、约2倍至约3倍、约2.5倍至约10倍、约2.5倍至约9倍、约2.5倍至约8倍、约2.5倍至约7倍、约2.5倍至约6倍、约2.5倍至约5倍、约2.5倍至约4倍、约2.5倍至约3倍、约3倍至约10倍、约3倍至约9倍、约3倍至约8倍、约3倍至约7倍、约3倍至约6倍、约3倍至约5倍、约3倍至约4倍、约4倍至约6倍、约5倍至约7倍或约6倍至约8倍。在其它实施方案中，FVIII的半衰期为至少约17小时、至少约18小时、至少约19小时、至少约20小时、至少约

21小时、至少约22小时、至少约23小时、至少约24小时、至少约25小时、至少约26小时、至少约27小时、至少约28小时、至少约29小时、至少约30小时、至少约31小时、至少约32小时、至少约33小时、至少约34小时、至少约35小时、至少约36小时、至少约48小时、至少约60小时、至少约72小时、至少约84小时、至少约96小时或至少约108小时。在其它实施方案中，FVIII的半衰期为约15小时至约两周、约16小时至约一周、约17小时至约一周、约18小时至约一周、约19小时至约一周、约20小时至约一周、约21小时至约一周、约22小时至约一周、约23小时至约一周、约24小时至约一周、约36小时至约一周、约48小时至约一周、约60小时至约一周、约24小时至约六天、约24小时至约五天、约24小时至约四天、约24小时至约三天或约24小时至约两天。

[0688] 在一些实施方案中，每个受试者的FVIII蛋白的平均半衰期为约15小时、约16小时、约17小时、约18小时、约19小时、约20小时、约21小时、约22小时、约23小时、约24小时(1天)、约25小时、约26小时、约27小时、约28小时、约29小时、约30小时、约31小时、约32小时、约33小时、约34小时、约35小时、约36小时、约40小时、约44小时、约48小时(2天)、约54小时、约60小时、约72小时(3天)、约84小时、约96小时(4天)、约108小时、约120小时(5天)、约六天、约七天(一周)、约八天、约九天、约10天、约11天、约12天、约13天或约14天。

[0689] 在具体实施方案中，本发明的嵌合蛋白的半衰期比野生型FVIII或BDD FVIII的半衰期长约两倍。在另一个实施方案中，嵌合蛋白的半衰期比野生型FVIII或BDD FVIII的半衰期长约三倍。

[0690] 此外，本发明提供治疗或阻止出血性疾病或病症的方法，其包括施用有效量的VWF片段或嵌合蛋白(如，包含连接至第一异源部分如第一Fc区的VWF片段和连接至第二异源部分如第二Fc区的FVIII蛋白的嵌合蛋白，其中VWF片段结合至FVIII蛋白或与之缔合)。在一个实施方案中，出血性疾病或病症选自由以下组成的组：出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血以及髂腰肌鞘出血。在具体实施方案中，出血性疾病或病症是甲型血友病。

[0691] 如本领域技术人员所认识，VWF片段和包含辅助部分的嵌合蛋白，如本文所述的VWF片段和本发明制备的FVIII蛋白有多种用途，包括但不限于治疗患有止血病症的受试者的方法和治疗需要一般止血剂的受试者的方法。在一个实施方案中，本发明涉及治疗患有止血病症的受试者的方法，其包括施用治疗有效量的VWF片段或嵌合蛋白。

[0692] 嵌合蛋白中的FVIII蛋白部分通过作为带负电的磷脂表面上的因子IX的辅因子，从而形成因子X酶复合物来治疗或抑制止血病症。活化的凝血因子结合至磷脂表面将该过程定位至血管受损部位。在磷脂表面上，因子VIIIa通过因子IXa增加因子X活化的最大速度大约200,000倍，从而导致第二次大量生成凝血酶。

[0693] 包含辅助部分如VWF片段和FVIII蛋白的嵌合蛋白可用于治疗任何止血病症。可通过施用本发明的嵌合蛋白治疗的止血病症包括但不限于甲型血友病，以及因子VIII相关的缺陷或结构异常。在一个实施方案中，止血病症是甲型血友病。

[0694] 包含辅助部分VWF片段和FVIII蛋白的嵌合蛋白可用于预防性治疗患有止血病症的受试者。本发明的嵌合蛋白可用于治疗患有止血病症的受试者的急性出血发作。在另一个实施方案中，止血病症可以是凝血因子如血管性血友病因子缺陷的结果。在一个实施方

案中，止血病症是遗传性病症。在另一个实施方案中，止血病症是获得性病症。获得性病症可以是由潜在的继发性疾病或病状引起的。不相关病状可以是例如但不限于癌症、自身免疫疾病或妊娠。获得性病症可以是由年老或治疗潜在的继发性病症的药物治疗(如，癌症化学疗法)引起的。

[0695] 本发明还涉及治疗不患有先天性止血病症，但患有继发性疾病或病状的受试者的方法，所述继发性疾病或病状导致获得止血病症，如由于抗FVIII抗体的发展或治疗。因此，本发明涉及治疗需要一般止血剂的受试者的方法，其包括施用治疗有效量的通过本发明方法制备的包含VWF片段和FVIII蛋白的嵌合蛋白。

[0696] 本发明还涉及减小FVIII的免疫原性或诱导对FVIII的免疫原性减小的方法，其包括施用有效量的VWF片段、本文所述的嵌合蛋白或编码其的多核苷酸。

[0697] 在一个实施方案中，需要一般止血剂的受试者正在经历或将要经历外科手术。包含VWF片段和FVIII蛋白的嵌合蛋白可在作为预防性方案的手术之前、期间或之后施用。包含VWF片段和FVIII蛋白的嵌合蛋白可在控制急性出血发作的手术之前、期间或之后施用。

[0698] 包含VWF片段和FVIII蛋白的嵌合蛋白可用于治疗不患有止血病症的受试者的急性出血发作。急性出血发作可以是严重的创伤，如手术、车祸、伤害、枪击裂伤或导致出血无法控制的任何其它创伤事件引起的。出血发作的非限制性实例包括：出血性凝血障碍、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血和髂腰肌鞘出血以及它们的任何组合。

[0699] 在预防性应用中，将包含本发明的嵌合蛋白或VWF片段或其混合物的一种或多种组合物施用给不处于疾病状态的患者，以增强患者的抵抗力或减少疾病或病症相关的症状。此量定义为“预防有效剂量”。在治疗应用中，在相对较短的间隔内相对较高的剂量(如，从约1至400mg/kg多肽/剂，其中从5至25mg的剂量更常用于放射性免疫缀合物，并且高剂量用于细胞毒素药物修饰的多肽)有时是必须的，直到疾病的进展减少或停止，并且直到患者显示出疾病症状的部分或完全改善。然后，患者可施用预防性方案。

[0700] 在一些实施方案中，本发明的嵌合蛋白、VWF片段或组合物用于按需治疗，其包括治疗出血发作、关节出血症、肌肉出血、口腔出血、溢血、溢血进入肌肉、口腔溢血、创伤、头部创伤(头部外伤)、胃肠出血、颅内溢血、腹腔内溢血、胸腔溢血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血或髂腰肌鞘出血。受试者可需要手术预防、围手术期处理或手术治疗。此类手术包括如小手术、大手术、拔牙术、扁桃体切除术、腹股沟疝切开术、滑膜切除术、全膝关节置换术、颅骨切开术、骨缝合术、创伤手术、颅内手术、腹腔内手术、胸内手术或关节置换术。

[0701] 在一个实施方案中，包含VWF片段和FVIII蛋白的嵌合蛋白静脉内、皮下、肌肉内或通过任何粘膜表面，如口服、舌下、口腔含化、鼻腔、直肠、阴道或通过肺部途径施用。包含VWF片段和FVIII蛋白的嵌合蛋白可植入或连接至生物聚合物固相支持物，所述生物聚合物固相支持物允许嵌合蛋白缓慢释放至出血部位或植入绷带/敷料。包含VWF片段和FVIII蛋白的嵌合蛋白的剂量根据受试者和所用的具体施用途径变化。剂量可在0.1至100,000μg/kg体重的范围内。在一个实施方案中，剂量范围为0.1-1,000μg/kg。在另一个实施方案中，剂量范围为0.1-500μg/kg。蛋白质可连续或以具体时间间隔施用。体外测定可用于确定最

佳剂量范围和/或施用方案。测量凝血因子活性的体外测定是本领域已知的，如STA-CLOT VIIa-rTF凝血测定或ROTEM凝血测定。另外，有效剂量可从得自动物模型如血友病狗的剂量响应曲线推断(Mount等2002,Blood 99(8):2670)。

[0702] 现在对本发明进行了详细描述，通过参考下面的实施例将更清晰地理解本发明，这些实施例据此仅出于示例性目的纳入，并且并非意图对本发明进行限制。本文引用的所有专利和公布明确地以引用方式并入。

实施例

[0703] 在全部实施例中，使用以下材料和方法，除非另外指明。

[0704] 材料和方法

[0705] 一般来讲，除非另外指明，本发明的实施采用化学、生物物理学、分子生物学、重组DNA技术、免疫学(尤其是例如抗体技术)和标准电泳技术的常规技术。参见例如Sambrook,Fritsch和Maniatis,Molecular Cloning:Cold Spring Harbor Laboratory Press(1989);Antibody Engineering Protocols(Methods in Molecular Biology),510,Paul,S.,Humana Pr(1996);Antibody Engineering:A Practical Approach(Practical Approach Series,169),McCafferty,编著,Irl Pr(1996);Antibodies:A Laboratory Manual,Harlow等,CS.H.L.Press,Pub.(1999);和Current Protocols in Molecular Biology,Ausubel等编著,John Wiley&Sons(1992)。

[0706] 实施例1:克隆不同的VWF结构域(图1)

[0707] (a)克隆pSYN-VWF-001、002、003和004

[0708] pSYN-VWF-001至004包含编码VWF片段的核苷酸序列，它们是VWF-D'D3A蛋白质序列的第1-276位氨基酸(001)、第1-477位氨基酸(002)、第1-511位氨基酸(003)和第1-716位氨基酸(004)。氨基酸编号表示无原肽的成熟VWF序列，并且分别对应于SEQ ID NO:2的第764-1039位氨基酸(001)、第764-1240位氨基酸(002)、第764-1274位氨基酸(003)和第764-1479位氨基酸(004)。所有四个构建体在N-末端具有FVIII信号肽，其允许合成蛋白质的正确分泌，FVIII信号肽后面是用于蛋白质纯化的C-末端6×His标签。上述构建体通过使用如下引物组合合成：

[0709] pSYN VWF-001：

[0710] 具有VIII信号和BsiW1位点的ESC48-Fwd-VWF-D'D3

[0711] TCGCGACGTACGGCCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGC CTTTTGCGATTCTGCTTTAGCCTATCCTGTCGGCCCCCCATG (SEQ ID NO: 57)

[0712] 具有6个His和Not1位点的ESC50-Rev-VWF-部分D'D3(第1-276位氨基酸)

[0713] TGACCTCGAGCGGCCGCTCAGTGGTGATGGTGATGATGCAGAGGCACTTTTCTGGTG TCAGCACACTG (SEQ ID NO: 58)

[0714] pSYN VWF-002：

[0715] 具有VIII信号和BsiW1位点的ESC48-Fwd-VWF-D'D3

[0716] TCGCGACGTACGGCCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGC CTTTTGCGATTCTGCTTTAGCCTATCCTGTCGGCCCCCCATG (SEQ ID NO: 59)

[0717] 具有6个His和Not1位点的ESC51-Rev-VWF D'D3(第1-477位氨基酸)

[0718] TGACCTCGAGCGGCCGCTCAGTGGTGATGGTGATGATGCGGCTCCTGGCAGGCTTCA CAGGTGAGGTTGACAAC (SEQ ID NO: 60)

[0719] pSYN VWF-003：

[0720] 具有VIII信号和BsiW1位点的ESC48-Fwd-VWF-D'D3

[0721] TCGCGACGTACGGCCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGC CTTTTGCGATTCTGCTTTAGCCTATCCTGTCGGCCCCCCATG (SEQ ID NO: 61)

[0722] 具有6个His和Not1位点的ESC52-Rev-VWF-D'D3部分A1(第1-511位氨基酸)

[0723] TGACCTCGAGCGGCCGCTCAGTGGTGATGGTGATGATGCCTGCTGCAGTAGAAATCG TGCAACGGCGGTTC (SEQ ID NO: 62)

[0724] pSYN VWF-004：

[0725] 具有VIII信号和BsiW1位点的ESC48-Fwd-VWF-D'D3

[0726] TCGCGACGTACGGCCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGC CTTTTGCGATTCTGCTTTAGCCTATCCTGTCGGCCCCCCATG (SEQ ID NO:63)

[0727] 具有6个His和Not1位点的ESC53-Rev-VWF-D'D3A1(第1-716位氨基酸)

[0728] TGACCTCGAGCGGCCGCTCAGTGGTGATGGTGATGATGGCCCACAGTGACTTGTGCC ATGTGGGG (SEQ ID NO: 64)

[0729] 假设来自VWF-001、002、003和004构建体的蛋白质以单体存在。

[0730] 50μl PCR反应使用ESC48/ESC50、ESC48/ESC51、ESC48/ESC52、ESC48/ESC53引物组合和全长VWF质粒作为模板进行，使用2步PCR扩增循环：94℃2分钟；21个(96℃30秒，68℃2分钟)循环。正确大小条带(VWF 001为约960bp、VWF 002为1460bp、VWF 003为1520bp以及VWF 004为2150bp)通过凝胶提取试剂盒(Gel Extraction kit)(Qiagen,Valencia,Calif.)进行凝胶纯化，并克隆至pcDNA 4的BsiWI和Not1限制性酶切位点，以分别生成pSYN-VWF 001、002、003和004。

[0731] (b)克隆pSYN-VWF-006

[0732] pSYN-VWF-006包含VWF的D1D2D'D3-CK(半胱氨酸结)结构域。为克隆该构建体，包含一部分D3结构域和CK结构域的DNA片段的合成外包进行(Genscript-序列标识号122026，如下所示)。将Genscript构建体的片段亚克隆至BamH1/EcoRV消化的pSYN-VWF 008，即编码全长VWF的载体。

[0733] Genscript-序列号-122026(SEQ ID NO:65)

[0734]
GGATCCTAGTGGGGAATAAGGGATGCAGCCACCCCTCAGTGAAATGCAAGAAACGGGTCACCATCCTGGTGG
AGGGAGGAGAGATTGAGCTGTTTGACGGGGAGGTGAATGTGAAGAGGCCCATGAAGGATGAGACTCACTTTG
AGGTGGTGGAGTCTGGCCGGTACATCATTCTGCTGCTGGGCAAAGCCCTCTCCGTGGTCTGGGACCGCCACC
TGAGCATCTCCGTGGTCCTGAAGCAGACATACCAGGAGAAAGTGTGTGGCCTGTGTGGGAATTTTGATGGCA
TCCAGAACAATGACCTCACCAGCAGCAACCTCCAAGTGGAGGAAGACCCTGTGGACTTTGGGAACTCCTGGA
AAGTGAGCTCGCAGTGTGCTGACACCAGAAAAGTGCCTCTGGACTCATCCCCTGCCACCTGCCATAACAACA
TCATGAAGCAGACGATGGTGGATTCCTCCTGTAGAATCCTTACCAGTGACGTCTTCCAGGACTGCAACAAGC
TGGTGGACCCCGAGCCATATCTGGATGTCTGCATTTACGACACCTGCTCCTGTGAGTCCATTGGGGACTGCG
CCTGCTTCTGCGACACCATTGCTGCCTATGCCCACGTGTGTGCCCAGCATGGCAAGGTGGTGACCTGGAGGA
CGGCCACATTGTGCCCCCAGAGCTGCGAGGAGAGGAATCTCCGGGAGAACGGGTATGAGTGTGAGTGGCGCT
ATAACAGCTGTGCACCTGCCTGTCAAGTCACGTGTCAGCACCCTGAGCCACTGGCCTGCCCTGTGCAGTGTG
TGGAGGGCTGCCATGCCCACTGCCCTCCAGGGAAAATCCTGGATGAGCTTTTGCAGACCTGCGTTGACCCTG
AAGACTGTCCAGTGTGTGAGGTGGCTGGCCGGCGTTTTGCCTCAGGAAAGAAAGTCACCTTGAATCCCAGTG
ACCCTGAGCACTGCCAGATTTGCCACTGTGATGTTGTCAACCTCACCTGTGAAGCCTGCCAGGAGCCGGGAG
GCCTGGTGGTGCCTCCCACAGATGCCCCGGTGAGCCCCACCACTCTGTATGTGGATGAGACGCTCCAGGATG
GCTGTGATACTCACTTCTGCAAGGTCAATGAGAGAGGAGAGTACTTCTGGGAGAAGAGGGTCACAGGCTGCC
CACCCTTTGATGAACACAAGTGTCTTGCTGAGGGAGGTAAAATTATGAAAATTCCAGGCACCTGCTGTGACA
CATGTGAGGAGCCTGAGTGCAACGACATCACTGCCAGGCTGCAGTATGTCAAGGTGGGAAGCTGTAAGTCTG
AAGTAGAGGTGGATATC

[0735] (c)克隆pSYN-VWF-009、010、011、012和013

[0736] pSYN VWF 008构建体包含pcDNA 3.1中的全长VWF序列(SEQ ID NO:2的第1-2813位氨基酸)。其包括763个氨基酸的原肽(即,D1D2结构域),然后是其余的2050个氨基酸成熟VWF序列。pSYN-VWF-009、010、011和012包含分别与VWF 001、002、003和004相同的编码序列,但还具有N-末端的D1D2结构域(VWF原肽),而不是FVIII信号肽。pSYN-VWF-008具有Arg907处的BamH1位点和编码区末端(在终止密码子之后)的Not1位点。pSYN-VWF-008、001、002、003和004用BamH1和Not1限制性内切酶消化。来自pSYN-VWF-001(423bp)、pSYN-VWF-002(1026bp)、pSYN-VWF-003(1128bp)和pSYN-VWF-004(1743bp)的插入子连接至bamH1/Not1消化的pSYN-VWF-008(8242bp),以得到pSYN-VWF-009(D1D2D'D3:SEQ ID NO:2的第1-1039位氨基酸)、pSYN-VWF-010(D1D2D'D3:SEQ ID NO:2的第1-1240位氨基酸)、pSYN-VWF-011(D1D2D'D3:SEQ ID NO:2的第1-1274位氨基酸)、pSYN-VWF-012(D1D2D'D3:第1-1479位氨基酸)。所有4个构建体具有C-末端的6×His标签。在转染的细胞中,pSYN-VWF-009、010、011和012用原肽合成,但由于细胞内加工,分泌产物不包含任何原肽(D1D2)。从VWF-009构建体表达的蛋白质以单体存在,并且假设从VWF-010、011和012构建体表达的蛋白质以二聚体存在,如图6和图7中所示,其分别使用VWF-009和VWF-010作为实例。

[0737] pSYN-VWF-010用于生成pSYN-VWF-013,其在对应于SEQ ID NO:73的C336A和C379A处具有两个点突变(氨基酸编号表示无D1D2结构域-VWF序列2的成熟VWF序列)。预期这些突变阻止VWF D'D3结构域的二聚化。

[0738] (d)克隆pSYN-VWF-025和029

[0739] pSYN-VWF-025包含pLIVE载体中全长VWF的野生型D1D2D'D3序列,而pSYN-VWF-029包含具有pLIVE载体中的C336A/C379A突变的D1D2D'D3结构域。对于克隆pSYN-VWF-025和029,使用如下引物组合:

[0740] 具有Nhe1位点的ESC89-fwd=CTCACTATAGGGAGACCCAAGCTGGCTAGCCG(SEQ ID NO:66)

[0741] 具有Sal1的ESC91-rev=

[0742] CTGGATCCCGGGAGTCGACTCGTCAGTGGTGATGGTGATGATG(SEQ ID NO:67)

[0743] 50μl PCR反应使用ESC 89/ESC91引物组合和pSYN-VWF-010(对于pSYN-VWF-025)

或pSYN-VWF-013(对于pSYN-VWF-029)作为质粒模板进行,使用3步PCR扩增循环:94℃2分钟;21个(96℃30秒,55℃30秒,68℃4分钟)循环。预期大小条带(约3800bp)通过凝胶提取试剂盒(Qiagen,Valencia,Calif.)进行凝胶纯化并克隆至pLIVE-Mirus载体(Invitrogen,Carlsbad,Calif.)的Nhe1和Sal1限制性酶切位点,以生成pSYN-VWF-025和029。

[0744] (e)克隆pSYN-VWF-031

[0745] pSYN-VWF-031是D1D2D'D3(C336A/C379A)-Fc构建体,其在VWF D1D2D'D3(C336A/C379A)和Fc序列之间具有48个氨基酸长度的凝血酶可切割连接基(8×GGGGS(SEQ ID NO:110)+凝血酶位点)。为制备该构建体,将VWF-Fc区从构建体pSYN-FVIII-064(以下称为FVIII-VWF构建体)扩增。将pSYN-FVIII-VWF用Xba1和Nhe1消化。所得的4165bp包含VWF片段和Fc区的插入区域用作通过引物组合LW22/LW23扩增VWF和Fc区的模板。

[0746] 具有FVIII信号序列和BsiW1位点的LW22-FWD-VWF-D'D3

[0747] GCGCCGGCCGTACGATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGCCTATCCTGTCGGCCCCCCATG (SEQ ID NO: 68)

[0748] 具有终止密码子和Not1位点的LW23-Rev-Fc

[0749] TCATCAATGTATCTTATCATGTCTGAATTCGCGGCCGCTCATTTACC (SEQ ID NO: 69)

[0750] VWF 031的核苷酸序列(SEQ ID NO:108)

[0751]

```
   1  ATGATTCCTG CCAGATTTGC CGGGGTGCTG CTTGCTCTGG CCCTCATTTT
  51  GCCAGGGACC CTTTGTGCAG AAGGAACTCG CGGCAGGTCA TCCACGGCCC
 101  GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG
 151  TACAGCTTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA
 201  ACGCTCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC
 251  TCTCCGTGTA TCTTGGGGAA TTTTTTGACA TCCATTTGTT TGTCAATGGT
 301  ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG
 351  GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCCT
 401  ATGGCTTTGT GGCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG
 451  TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTGGCA ACTTTAACAT
 501  CTTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC
 551  CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT
 601  GAACGGGCAT CTCCTCCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT
 651  GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT
 701  TTGCCCGCTG CCACCCTCTG GTGGACCCCG AGCCTTTTGT GGCCCTGTGT
 751  GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CCTGCCCTGC
 801  CCTCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAATG GTGCTGTACG
 851  GCTGGACCGA CCACAGCGCG TGCAGCCCAG TGTGCCCTGC TGGTATGGAG
 901  TATAGGCAGT GTGTGTCCCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT
 951  CAATGAAATG TGTCAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG
1001  GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGTCCCTGC
1051  GTGCATTCCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG
1101  CAACACCTGC ATTTGCCGAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT
1151  GTCCAGGGGA GTGCCTTGTC ACTGGTCAAT CCCACTTCAA GAGCTTTGAC
1201  AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGGA
1251  TTGCCAGGAC CACTCCTTCT CCATTGTCAT TGAGACTGTC CAGTGTGCTG
1301  ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGTCCG GCTGCCTGGC
1351  CTGCACAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA
1401  TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC
1451  ATACAGTGAC GGCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG
1501  GACTGGGATG GCCGCGGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC
1551  CGGGAAGACC TGCGGCCTGT GTGGGAATTA CAATGGCAAC CAGGGCGACG
1601  ACTTCCTTAC CCCCTCTGGG CTGGCGGAGC CCCGGGTGGA GGACTTCGGG
1651  AACGCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG
1701  CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGGTTCTCC GAGGAGGCGT
1751  GCGCGGTCCT GACGTCCCCC ACATTCGAGG CCTGCCATCG TGCCGTCAGC
1801  CCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACGTGTGCT CCTGCTCGGA
1851  CGGCCGCGAG TGCCTGTGCG GCGCCCTGGC CAGCTATGCC GCGGCCTGCG
1901  CGGGGAGAGG CGTGCGCGTC GCGTGGCGCG AGCCAGGCCG CTGTGAGCTG
1951  AACTGCCCGA AAGGCCAGGT GTACCTGCAG TGCGGGACCC CCTGCAACCT
2001  GACCTGCCGC TCTCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCTGCC
2051  TGGAGGGCTG CTTCTGCCCC CCAGGGCTCT ACATGGATGA GAGGGGGGAC
```

```
2101    TGCGTGCCCA AGGCCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA
2151    GCCAGAAGAC ATCTTCTCAG ACCATCACAC CATGTGCTAC TGTGAGGATG
2201    GCTTCATGCA CTGTACCATG AGTGGAGTCC CCGGAAGCTT GCTGCCTGAC
2251    GCTGTCCTCA GCAGTCCCCT GTCTCATCGC AGCAAAAGGA GCCTATCCTG
2301    TCGGCCCCCC ATGGTCAAGC TGGTGTGTCC CGCTGACAAC CTGCGGGCTG
2351    AAGGGCTCGA GTGTACCAAA ACGTGCCAGA ACTATGACCT GGAGTGCATG
2401    AGCATGGGCT GTGTCTCTGG CTGCCTCTGC CCCCCGGGCA TGGTCCGGCA
2451    TGAGAACAGA TGTGTGGCCC TGGAAAGGTG TCCCTGCTTC CATCAGGGCA
2501    AGGAGTATGC CCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC
2551    TGTCGGGACC GGAAGTGGAA CTGCACAGAC CATGTGTGTG ATGCCACGTG
2601    CTCCACGATC GGCATGGCCC ACTACCTCAC CTTCGACGGG CTCAAATACC
2651    TGTTCCCCGG GGAGTGCCAG TACGTTCTGG TGCAGGATTA CTGCGGCAGT
2701    AACCCTGGGA CCTTTCGGAT CCTAGTGGGG AATAAGGGAT GCAGCCACCC
2751    CTCAGTGAAA TGCAAGAAAC GGGTCACCAT CCTGGTGGAG GGAGGAGAGA
2801    TTGAGCTGTT TGACGGGGAG GTGAATGTGA AGAGGCCCAT GAAGGATGAG
2851    ACTCACTTTG AGGTGGTGGA GTCTGGCCGG TACATCATTC TGCTGCTGGG
2901    CAAAGCCCTC TCCGTGGTCT GGGACCGCCA CCTGAGCATC TCCGTGGTCC
2951    TGAAGCAGAC ATACCAGGAG AAAGTGTGTG GCCTGTGTGG GAATTTTGAT
3001    GGCATCCAGA ACAATGACCT CACCAGCAGC AACCTCCAAG TGGAGGAAGA
3051    CCCTGTGGAC TTTGGGAACT CCTGGAAAGT GAGCTCGCAG TGTGCTGACA
3101    CCAGAAAAGT GCCTCTGGAC TCATCCCCTG CCACCTGCCA TAACAACATC
3151    ATGAAGCAGA CGATGGTGGA TTCCTCCTGT AGAATCCTTA CCAGTGACGT
3201    CTTCCAGGAC TGCAACAAGC TGGTGGACCC CGAGCCATAT CTGGATGTCT
3251    GCATTTACGA CACCTGCTCC TGTGAGTCCA TTGGGGACTG CGCCGCATTC
[0752] 3301    TGCGACACCA TTGCTGCCTA TGCCCACGTG TGTGCCCAGC ATGGCAAGGT
3351    GGTGACCTGG AGGACGGCCA CATTGTGCCC CCAGAGCTGC GAGGAGAGGA
3401    ATCTCCGGGA GAACGGGTAT GAGGCTGAGT GGCGCTATAA CAGCTGTGCA
3451    CCTGCCTGTC AAGTCACGTG TCAGCACCCT GAGCCACTGG CCTGCCCTGT
3501    GCAGTGTGTG GAGGGCTGCC ATGCCCACTG CCCTCCAGGG AAAATCCTGG
3551    ATGAGCTTTT GCAGACCTGC GTTGACCCTG AAGACTGTCC AGTGTGTGAG
3601    GTGGCTGGCC GGCGTTTTGC CTCAGGAAAG AAAGTCACCT TGAATCCCAG
3651    TGACCCTGAG CACTGCCAGA TTTGCCACTG TGATGTTGTC AACCTCACCT
3701    GTGAAGCCTG CCAGGAGCCG ATATCTGGCG GTGGAGGTTC CGGTGGCGGG
3751    GGATCCGGCG GTGGAGGTTC CGGCGGTGGA GGTTCCGGTG GCGGGGGATC
3801    CGGTGGCGGG GGATCCCTGG TCCCCCGGGG CAGCGGCGGT GGAGGTTCCG
3851    GTGGCGGGGG ATCCGACAAA ACTCACACAT GCCCACCGTG CCCAGCTCCA
3901    GAACTCCTGG GCGGACCGTC AGTCTTCCTC TTCCCCCCAA AACCCAAGGA
3951    CACCCTCATG ATCTCCCGGA CCCCTGAGGT CACATGCGTG GTGGTGGACG
4001    TGAGCCACGA AGACCCTGAG GTCAAGTTCA ACTGGTACGT GGACGGCGTG
4051    GAGGTGCATA ATGCCAAGAC AAAGCCGCGG GAGGAGCAGT ACAACAGCAC
4101    GTACCGTGTG GTCAGCGTCC TCACCGTCCT GCACCAGGAC TGGCTGAATG
4151    GCAAGGAGTA CAAGTGCAAG GTCTCCAACA AAGCCCTCCC AGCCCCCATC
4201    GAGAAAACCA TCTCCAAAGC CAAAGGGCAG CCCCGAGAAC CACAGGTGTA
4251    CACCCTGCCC CCATCCCGGG ATGAGCTGAC CAAGAACCAG GTCAGCCTGA
4301    CCTGCCTGGT CAAAGGCTTC TATCCCAGCG ACATCGCCGT GGAGTGGGAG
4351    AGCAATGGGC AGCCGGAGAA CAACTACAAG ACCACGCCTC CCGTGTTGGA
4401    CTCCGACGGC TCCTTCTTCC TCTACAGCAA GCTCACCGTG GACAAGAGCA
4451    GGTGGCAGCA GGGGAACGTC TTCTCATGCT CCGTGATGCA TGAGGCTCTG
4501    CACAACCACT ACACGCAGAA GAGCCTCTCC CTGTCTCCGG GTAAATGA
```

[0753] VWF 031的蛋白质序列(SEQ ID NO:109)

[0754]

```
   1    MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNTFDGSM
  51    YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE FFDIHLFVNG
 101    TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL
 151    SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS WALSSGEQWC
 201    ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC

 251    EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME
 301    YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC
 351    VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD
 401    NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG
 451    LHNSLVKLKH GAGVAMDGQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM
 501    DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG
 551    NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS
 601    PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL
 651    NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGCFCP PGLYMDERGD
 701    CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD
 751    AVLSSPLSHR SKRSLSCRPP MVKLVCPADN LRAEGLECTK TCQNYDLECM
 801    SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV
 851    CRDRKWNCTD HVCDATCSTI GMAHYLTFDG LKYLFPGECQ YVLVQDYCGS
 901    NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE
 951    THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD
1001    GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI
1051    MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS CESIGDCAAF
1101    CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY EAEWRYNSCA
1151    PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE
1201    VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP ISGGGGSGGG
1251    GSGGGGSGGG GSGGGGSGGG GSLVPRGSGG GGSGGGGSDK THTCPPCPAP
1301    ELLGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV
1351    EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI
1401    EKTISKAKGQ PREPQVYTLP PSRDELTKNQ VSLTCLVKGF YPSDIAVEWE
1451    SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL
1501    HNHYTQKSLS LSPGK*
```

[0755]

DNA构建体	VWF和Fc之间的连接基
VWF035	73 aa = IS{11×(GGGGS)} LVPRGSGGGGSGGGGS (SEQ ID NO: 96)
VWF036	98 aa = IS{16×(GGGGS)} LVPRGSGGGGSGGGGS (SEQ ID NO: 97)
VWF = D'D3（具有C336A/C379A的1-477aa）	

[0756] 将得自LW22/LW23扩增的PCR产物(约2300bp)克隆至BsiW1/Not1消化的pSYN-VWF-002以得到pSYN-VWF-014中间体。pSYN-VWF-014包含FVIII信号肽-D'D3-20个氨基酸的凝血酶可切割连接基,然后是Fc区。

[0757] 为生成D1D2D'D3-Fc构建体,D1D2D'D3区使用引物组合LW24/LW27通过标准PCR方法从pSYN-VWF-013扩增。

[0758] 具有BsiW1位点的LW24-Fwd-VWF D1D2D'D3克隆寡核苷酸

[0759] GCGCCGGCCGTACGATGATTCCTGCCAGATTTGCCGGGGTG (SEQ ID NO: 70)

[0760] 具有EcoRV的LW27-Rev-VWF D'D3寡核苷酸

[0761] CCACCGCCAGATATCGGCTCCTGGCAGGCTTCACAGGTGAG (SEQ ID NO:71)

[0762] 将得自LW22/LW23扩增的PCR产物(约3750bp)克隆至BsiW1/EcoRV消化的pSYN-VWF-014以得到pSYN-VWF-015中间体。改变VWF片段和Fc区之间的连接基长度以得到pSYN-VWF-031。

[0763] 全长VWF蛋白序列在表1中示出。

[0764] VWF-D1D2D'D3蛋白质序列1b(SEQ ID NO:72)

[0765]
```
   1    MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNTFDGSM
  51    YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE FFDIHLFVNG
 101    TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL
 151    SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS WALSSGEQWC
 201    ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC
 251    EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME
 301    YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC
 351    VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD
 401    NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG
 451    LHNSLVKLKH GAGVAMDGQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM
 501    DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG
 551    NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS
 601    PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL
 651    NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGCFCP PGLYMDERGD
 701    CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD
 751    AVLSSPLSHR SKRSLSCRPP MVKLVCPADN LRAEGLECTK TCQNYDLECM
 801    SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV
 851    CRDRKWNCTD HVCDATCSTI GMAHYLTFDG LKYLFPGECQ YVLVQDYCGS
 901    NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE
 951    THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD
1001    GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI
1051    MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS CESIGDCACF
1101    CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY ECEWRYNSCA
1151    PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE
1201    VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP*
```

[0766] VWF-D'D3蛋白质序列2(SEQ ID NO:73)

[0767]
```
   1    SLSCRPPMVK LVCPADNLRA EGLECTKTCQ NYDLECMSMG CVSGCLCPPG
  51    MVRHENRCVA LERCPCFHQG KEYAPGETVK IGCNTCVCRD RKWNCTDHVC
 101    DATCSTIGMA HYLTFDGLKY LFPGECQYVL VQDYCGSNPG TFRILVGNKG
 151    CSHPSVKCKK RVTILVEGGE IELFDGEVNV KRPMKDETHF EVVESGRYII
 201    LLLGKALSVV WDRHLSISVV LKQTYQEKVC GLCGNFDGIQ NNDLTSSNLQ
 251    VEEDPVDFGN SWKVSSQCAD TRKVPLDSSP ATCHNNIMKQ TMVDSSCRIL
 301    TSDVFQDCNK LVDPEPYLDV CIYDTCSCES IGDCACFCDT IAAYAHVCAQ
 351    HGKVVTWRTA TLCPQSCEER NLRENGYECE WRYNSCAPAC QVTCQHPEPL
 401    ACPVQCVEGC HAHCPPGKIL DELLQTCVDP EDCPVCEVAG RRFASGKKVT
 451    LNPSDPEHCQ ICHCDVVNLT CEACQEP
```

[0768] 实施例2:包含第二Fc链氨基末端的FVIII-Fc和VWF-D'D3结构域的异源二聚体构建体(FVIII-VWF-Fc异源二聚体,图2)

[0769] (a)pSYN-FVIII-064的克隆

[0770] FVIII-064质粒包含单链FC(scFc)支架,其具有在细胞中合成期间加工的酶切位点。构建体具有全长VWF的FVIII结合结构域(D'D3)。

[0771] 质粒(pSYN-FVIII-064)设计为表达FVIII-Fc和VWF-Fc异源二聚体,其中D'D3结构域结合FVIII并且阻止FVIII与磷脂和活化蛋白C的相互作用和/或阻止或抑制与内源性VWF的结合。来自pSYN-FVIII-064的蛋白质作为单个多肽在细胞中表达,其中FVIII-Fc亚基的C-末端通过6×(GGGGS)多肽连接基(SEQ ID NO:74)连接至VWF D'D3-Fc亚基的N-末端。此外,RRRRS(SEQ ID NO:75)和RKRRKR(SEQ ID NO:76)序列分别插入多肽连接基的5'和3'末端,通过前蛋白转化酶在每个序列最后的Arg之后在细胞内切割。因此,细胞可表达双链

FVIII-Fc/D'D3-Fc异源二聚体,其中FVIII-Fc链具有C-末端的RRRRS序列(SEQ ID NO:75),但连接基序列的其余部分被移除。另一个3×(GGGGS)多肽连接基(SEQ ID NO:28)以及凝血酶切割位点被引入VWF结构域和Fc区之间,一旦FVIII-VWF异源二聚体蛋白被凝血酶活化,使得FVIII与其它凝血因子相互作用,即促进VWF片段从FVIII释放。

[0772] 包含一部分第一Fc区,然后是6×(GGGGS)(SEQ ID NO:74)、VWF-D'D3结构域(1-477aa;C336A/C379A突变)、3×(GGGGS)(SEQ ID NO:28)、凝血酶切割位点和一部分第二Fc的DNA片段的合成外包进行(Genscript-序列号103069,如下所示)。将Genscript构建体的片段亚克隆至SalI/RsRII消化的pSYN-FVIII-049,pSYN-FVIII-049是可切割的连接基在两个Fc结构域之间的FVIII-Fc构建体。

[0773] Genscript-序列号103069(SEQ ID NO:82):

[0774]
CCGTCGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACC
ACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAACGGCGCCGCCGGAGCGGTGGCGGCGGATCAGGTG
GGGGTGGATCAGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGTGGGGGTGGATCAA
GGAAGAGGAGGAAGAGAAGCCTATCCTGTCGGCCCCCCATGGTCAAGCTGGTGTGTCCCGCTGACAACCTGC

[0775]
GGGCTGAAGGGCTCGAGTGTACCAAAACGTGCCAGAACTATGACCTGGAGTGCATGAGCATGGGCTGTGTCT
CTGGCTGCCTCTGCCCCCCGGGCATGGTCCGGCATGAGAATCGATGTGTGGCCCTGGAAAGGTGTCCCTGCT
TCCATCAGGGCAAGGAGTATGCCCCTGGAGAAACAGTGAAGATTGGCTGCAACACTTGTGTCTGTCGGGACC
GGAAGTGGAACTGCACAGACCATGTGTGTGATGCCACGTGCTCCACGATCGGCATGGCCCACTACCTCACCT
TCGACGGGCTCAAATACCTGTTCCCCGGGGAGTGCCAGTACGTTCTGGTGCAGGATTACTGCGGCAGTAACC
CTGGGACCTTTCGGATCCTAGTGGGGAATAAGGGATGCAGCCACCCCTCAGTGAAATGCAAGAAACGGGTCA
CCATCCTGGTGGAGGGAGGAGAGATTGAGCTGTTTGACGGGGAGGTGAATGTGAAGAGGCCCATGAAGGATG
AGACTCACTTTGAGGTGGTGGAGTCTGGCCGGTACATCATTCTGCTGCTGGGCAAAGCCCTCTCCGTGGTCT
GGGACCGCCACCTGAGCATCTCCGTGGTCCTGAAGCAGACATACCAGGAGAAAGTGTGTGGCCTGTGTGGGA
ATTTTGATGGCATCCAGAACAATGACCTCACCAGCAGCAACCTCCAAGTGGAGGAAGACCCTGTGGACTTTG
GGAACTCCTGGAAAGTGAGCTCGCAGTGTGCTGACACCAGAAAAGTGCCTCTGGACTCATCCCCTGCCACCT
GCCATAACAACATCATGAAGCAGACGATGGTGGATTCCTCCTGTAGAATCCTTACCAGTGACGTCTTCCAGG
ACTGCAACAAGCTGGTGGACCCCGAGCCATATCTGGATGTCTGCATTTACGACACCTGCTCCTGTGAGTCCA
TTGGGGACTGCGCCGCATTCTGCGACACCATTGCTGCCTATGCCCACGTGTGTGCCCAGCATGGCAAGGTGG
TGACCTGGAGGACGGCCACATTGTGCCCCCAGAGCTGCGAGGAGAGGAATCTCCGGGAGAACGGGTATGAGG
CTGAGTGGCGCTATAACAGCTGTGCACCTGCCTGTCAAGTCACGTGTCAGCACCCTGAGCCACTGGCCTGCC
CTGTGCAGTGTGTGGAGGGCTGCCATGCCCACTGCCCTCCAGGGAAAATCCTGGATGAGCTTTTGCAGACCT
GCGTTGACCCTGAAGACTGTCCAGTGTGTGAGGTGGCTGGCCGGCGTTTTGCCTCAGGAAAGAAAGTCACCT
TGAATCCCAGTGACCCTGAGCACTGCCAGATTTGCCACTGTGATGTTGTCAACCTCACCTGTGAAGCCTGCC
AGGAGCCGATCGATGGCGGTGGAGGTTCCGGTGGCGGGGGATCCCTGGTCCCCCGGGGCAGCGGAGGCGACA
AAACTCACACATGCCCACCGTGCCCAGCTCCAGAACTCCTGGGCGGACCGTCA

[0776] (b)pSYN-FVIII-065的克隆

[0777] FVIII-065质粒包含连接至第二Fc区的VWF的D'D3结构域的前276个氨基酸。将VWF片段通过使用引物组合ESC17和ESC41从全长VWF质粒pSYN-VWF-008进行PCR扩增。

[0778] 具有Cla1的ESC17-Fwd-VWF克隆寡核苷酸

[0779] GTCCGGCATGAGAATCGATGTGTG (SEQ ID NO: 77)

[0780] 具有EcoRV的ESC41-Rev-VWF

[0781] CCTCCACCGCCAGATATCAGAGGCACTTTTC (SEQ ID NO: 78)

[0782] 预期大小条带(约692bp)通过凝胶提取试剂盒(Qiagen,Valencia,Calif.)进行凝胶纯化并克隆至pSYN-FVIII-064的Cla1和EcoRV位点,生成pSYN-FVIII-065。

[0783] 实施例3:pSYN-FVIII-159、160、178、179的克隆(图3)

[0784] 为了改变VWF片段和Fc区之间的连接基长度,将EcoRV位点引入pSYN-FVIII-064中VWF和20个氨基酸连接基开始的接合处,然后可变大小连接基用于替代PSYN-FVIII-064中的20aa连接基。新DNA构建体为:pSYN-FVIII-159、160、178和179,其分别包含35aa、48aa、73aa和98aa连接基。

[0785] 为了将35个氨基酸的连接基插入pSYN-FVIII-159，从Integrated DNA Technologies,Inc(Coralville,IA)订购两个寡核苷酸(ESC78-105bp和ESC79-107bp)。将寡核苷酸使用标准PCR方法进行退火和延伸：

[0786] 引物：

[0787] 具有EcoRV位点的ESC78-Fwd

AAAGTGCCTCTGATATCTGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGTGGCGGG

[0788] GGATCCGGTGGCGGGGGATCCGGTGGCGGGGGATCCCTGGTCCCCCGG (SEQ ID NO: 79)

[0789] 具有RsRII位点的ESC79-Rev

GAAGAGGAAGACTGACGGTCCGCCCAGGAGTTCTGGAGCTGGGCACGGTGGGCATGT

[0790] GTGAGTTTTGTCGCCTCCGCTGCCCCGGGGGACCAGGGATCCCCCGCCAC (SEQ ID NO: 80)

[0791] 50μl PCR寡核苷酸退火和延伸反应使用ESC78/ESC79引物组合进行，使用3步PCR扩增循环：25个(96℃30秒，55℃30秒，68℃30秒)循环。将预期大小条带(约186bp)用凝胶提取试剂盒(Qiagen,Valencia,Calif.)进行凝胶纯化并克隆至pSYN-FVIII-064的EcoRV和RsRII限制性位点，生成pSYN-FVIII-159。

[0792] (b)克隆pSYN-FVIII-160、178和179

[0793] pSYN-VIII-160具有VWF片段和Fc区之间的48个氨基酸的连接基。编码48个氨基酸的连接基(ISGG GGSGGGGSGGGGSGGGGSGGGGSGGGGSLVPRGSGGGGSGGGGS)(SEQ ID NO:81)和一部分Fc区的DNA片段的合成外包进行(Genscript-序列号-132601，如下所示)。将Genscript构建体的片段亚克隆至EcoRV/RsRII消化的pSYN-FVIII-0159(如上所述)。

[0794] Genscript-序列号-132601(SEQ ID NO:83)

[0795] AAAGTGCCTCTGATATCTGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGCGGTGGA GGTTCCGGTGGCGGGGGATCCGGTGGCGGGGGATCCCTGGTCCCCCGGGGCAGCGGCGGTGGAGGTTCCGGT GGCGGGGGATCCGACAAAACTCACACATGCCCACCGTGCCCAGCTCCAGAACTCCTGGGCGGACCGTCAGTC TTCC

[0796] pSYN-VIII-178具有VWF片段和Fc区之间的73个氨基酸的连接基。编码73个氨基酸的连接基(ISGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSLVPRGSGGGGSGGGGS)(SEQ ID NO:84)和一部分Fc区的DNA片段的合成外包进行(Genscript-序列号-144849，如下所示)。将Genscript构建体的片段亚克隆至EcoRV/RsRII消化的pSYN-FVIII-0159(如上所述)。

[0797] Genscript-序列#-144849(SEQ ID NO:85)

GCCTGCCAGGAGCCGATATCTGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGCGGT

[0798] GGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGC GGTGGAGGTTCCGGTGGCGGGGGATCCGGTGGCGGGGGATCCCTGGTCCCCCGGGGCAGCGGCGGTGGAGGT TCCGGTGGCGGGGGATCCGACAAAACTCACACATGCCCCCGTGCCCAGCTCCAGAACTCCTGGGCGGACCGT CAGTCTTCCTC

[0799] pSYN-VIII-179具有VWF片段和Fc区之间的98个氨基酸的连接基。编码98个氨基酸的连接基(ISGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSLVPRGSGGGGSGGGGS)(SEQ ID NO:86)和一部分Fc区的DNA片段的合成外包进行(Genscript-序列号-144849，如下所示)。将Genscript构建体的片段亚克隆至EcoRV/

RsRII消化的pSYN-FVIII-0159(如上所述)。

[0800] Genscript-序列#-144849(SEQ ID NO:87)

[0801] GCCTGCCAGGAGCCGATATCTGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGCGGT
GGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGC
GGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCC
GGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGTGGCGGGGGATCCCTGGTCCCCCGGGGCAGCGGCGGTGGA
GGTTCCGGTGGCGGGGGATCCGACAAAACTCACACATGCCCACCGTGCCCAGCTCCAGAACTCCTGGGCGGA
CCGTCAGTCTTCCTCTTCCC

[0802] pSYN-FVIII-180、181和182的克隆

[0803] pSYN-FVIII-180、181和182从pSYN-FVIII-160构建。将K2093A或F2093A或K2093A/F2093A突变引入pSYN-FVIII-160中FVIII的C1结构域,分别形成pSYN-FVIII-180、pSYN-FVIII-181和pSYN-FVIII-182。

[0804] FVIII-VWF-Fc异源二聚体蛋白质序列(SEQ ID NO:88)

[0805] (FVIII序列第1-1457位氨基酸;下划线区表示Fc区;波浪下划线表示第一Fc和VWF片段之间的可切割的连接基;双下划线区表示VWF片段;粗体区表示VWF片段和Fc之间的可变长度可切割的连接基。FVIII-064、159、160、178和179构建体中的连接基长度可变化。)

```
   1  MQIELSTCFF LCLLRFCFSA TRRYYLGAVE LSWDYMQSDL GELPVDARFP
  51  PRVPKSFPFN TSVVYKKTLF VEFTDHLFNI AKPRPPWMGL LGPTIQAEVY
 101  DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR EKEDDKVFPG
 151  GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG LIGALLVCRE
 201  GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD AASARAWPKM
 251  HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL EGHTFLVRNH
 301  RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME AYVKVDSCPE
 351  EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI RSVAKKHPKT
 401  WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR KYKKVRFMAY
 451  TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR PYNIYPHGIT
 501  DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP TKSDPRCLTR
 551  YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR NVILFSVFDE
 601  NRSWYLTENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV FDSLQLSVCL
 651  HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVYEDTLTLF PFSGETVFMS
 701  MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED SYEDISAYLL
 751  SKNNAIEPRS FSQNPPVLKR HQREITRTTL QSDQEEIDYD DTISVEMKKE
 801  DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP HVLRNRAQSG
 851  SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA EVEDNIMVTF
 901  RNQASRPYSF YSSLISYEED QRQGAEPRKN FVKPNETKTY FWKVQHHMAP
 951  TKDEFDCKAW AYFSDVDLEK DVHSGLIGPL LVCHTNTLNP AHGRQVTVQE
1001  FALFFTIFDE TKSWYFTENM ERNCRAPCNI QMEDPTFKEN YRFHAINGYI
1051  MDTLPGLVMA QDQRIRWYLL SMGSNENIHS IHFSGHVFTV RKKEEYKMAL
1101  YNLYPGVFET VEMLPSKAGI WRVECLIGEH LHAGMSTLFL VYSNKCQTPL
1151  GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK EPFSWIKVDL
1201  LAPMIIHGIK TQGARQKFSS LYISQFIIMY SLDGKKWQTY RGNSTGTLMV
[0806] 1251  FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME LMGCDLNSCS
1301  MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR SNAWRPQVNN
1351  PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISSS QDGHQWTLFF
1401  QNGKVKVFQG NQDSFTPVVN SLDPPLLTRY LRIHPQSWVH QIALRMEVLG
1451  CEAQDLYDKT HTCPPCPAPE LLGGPSVFLF PPKPKDTLMI SRTPEVTCVV
1501  VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW
1551  LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SRDELTKNQV
1601  SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD
1651  KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKRRRRSG GGGSGGGGSG
1701  GGGSGGGGSG GGGSGGGGSR KRRKRSLSCR PPMVKLVCPA DNLRAEGLEC
1751  TKTCQNYDLE CMSMGCVSGC LCPPGMVRHE NRCVALERCP CFHQGKEYAP
1801  GETVKIGCNT CVCRDRKWNC TDHVCDATCS TIGMAHYLTF DGLKYLFPGE
1851  CQYVLVQDYC GSNPGTFRIL VGNKGCSHPS VKCKKRVTIL VEGGEIELFD
1901  GEVNVKRPMK DETHFEVVES GRYIILLLGK ALSVVWDRHL SISVVLKQTY
1951  QEKVCGLCGN FDGIQNNDLT SSNLQVEEDP VDFGNSWKVS SQCADTRKVP
2001  LDSSPATCHN NIMKQTMVDS SCRILTSDVF QDCNKLVDPE PYLDVCIYDT
2051  CSCESIGDCA AFCDTIAAYA HVCAQHGKVV TWRTATLCPQ SCEERNLREN
2101  GYEAEWRYNS CAPACQVTCQ HPEPLACPVQ CVEGCHAHCP PGKILDELLQ
2151  TCVDPEDCPV CEVAGRRFAS GKKVTLNPSD PEHCQICHCD VVNLTCEACQ
2201  EPIDGGGGSG GGGSLVPRGS GGDKTHTCPP CPAPELLGGP SVFLFPPKPK
2251  DTLMISRTPE VTCVVVDVSH EDPEVKFNWY VDGVEVHNAK TKPREEQYNS
2301  TYRVVSVLTV LHQDWLNGKE YKCKVSNKAL PAPIEKTISK AKGQPREPQV
2351  YTLPPSRDEL TKNQVSLTCL VKGFYPSDIA VEWESNGQPE NNYKTTPPVL

2401  DSDGSFFLYS KLTVDKSRWQ QGNVFSCSVM HEALHNHYTQ KSLSLSPGK
```

[0807] 实施例4:FVIII-VWF DNA构建体的实施例(图4)

[0808] VWF片段和FVIII蛋白可使用常规重组DNA技术通过连接基或另一个蛋白质或多肽连接在一起,如图4中所示。在图4A中,VWF的D1D2D'D3结构域通过48aa连接基-ISGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSLVPRGSGGGGSGGGGS(SEQ ID NO:89)连接至FVIII蛋白,并且保护

FVIII不被过早清除。为了进一步增加D'D3的FVIII保护活性，另一个具有半衰期延长潜力的蛋白质或多肽，例如白蛋白或PAS序列(异源部分)可掺入构建体。异源部分如白蛋白或PAS序列可掺入FVIII分子的不同位置；图4B-4D中示出了几个实例：在FVIII的N-末端(4B)、在FVIII的C-末端(4C)或在B区中(4D)。在那些构建体中，另外的蛋白质序列可增加D'D3保护活性并且进一步延长FVIII半衰期。

[0809] 此外，异源部分如白蛋白或PAS序列也可掺入FVIII/VWF异源二聚体构建体，如图4E-4G中所示。在图4E中，异源部分如白蛋白或PAS序列掺入FVIII-148的FVIII B结构域；在图4F中，异源部分如白蛋白或PAS序列掺入FVIII-136的FVIII B结构域区；在图4G中，异源部分如白蛋白或PAS序列用作使D'D3片段和Fc连接的连接基。在那些构型中，预期D'D3、Fc以及作为半衰期延长因子的异源部分(如，白蛋白/PAS序列)的协同效应使FVIII半衰期延长。

[0810] 实施例5：FVIIIFc-VWF异源二聚体共转染系统的质粒构建(图5)

[0811] 制备产生FVIIIFc-VWF异源二聚体的共转染系统，该系统包含三个DNA构建体。第一DNA构建体-pSYN-FVIII-155编码FVIII-Fc融合蛋白，其中单链FVIII蛋白直接融合至单个Fc片段，并且第二DNA构建体是pSYN-VWF-031，其编码D'D3-Fc融合蛋白(在上面实施例1中有所描述)。HEK293F细胞用80:15:5的比率两个质粒和第三质粒(PC5)转染。用PC5共转染用于确保D1和D2区的完全原肽加工，以得到成熟D'D3结构域。合成蛋白质以FVIIIFc/D'D3Fc异源二聚体和D'D3Fc同源二聚体分泌，并且将FVIIIFc/D'D3Fc异源二聚体通过蛋白质纯化与D'D3Fc同源二聚体分离。

[0812] pSYN-FVIII-155成熟蛋白质测序(SEQ ID NO:90)：

[0813]
ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLL
GPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKEN
GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSL
MQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEI
SPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRF
DDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYT
DETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPIL
PGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILF
SVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDF
LSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYE
DSYEDISAYLLSKNNAIEPRSFSQNPPVLKAHQAEITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQ
SPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGL
LGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEF
DCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAP
CNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMA
LYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW
APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGN
STGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQI
TASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLI
SSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLYDK
THTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE
QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT
CLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQ
KSLSLSPGK

[0814] pSYN-FVIII-155DNA测序(SEQ ID NO:91)：

[0815]

ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGTGCCACCAGAAGATAC
TACCTGGGTGCAGTGGAACTGTCATGGGACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGA
TTTCCTCCTAGAGTGCCAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAA
TTCACGGATCACCTTTTCAACATCGCTAAGCCAAGGCCACCCTGGATGGGTCTGCTAGGTCCTACCATCCAG
GCTGAGGTTTATGATACAGTGGTCATTACACTTAAGAACATGGCTTCCCATCCTGTCAGTCTTCATGCTGTT
GGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGAAGAT
GATAAAGTCTTCCCTGGTGGAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAATGGTCCAATGGCCTCT
GACCCACTGTGCCTTACCTACTCATATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATT
GGAGCCCTACTAGTATGTAGAGAAGGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAAATTTATACTA
CTTTTTGCTGTATTTGATGAAGGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGGATAGGGAT
GCTGCATCTGCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCAGGTCTG
ATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCTGAAGTGCACTCAATA
TTCCTCGAAGGTCACACATTTCTTGTGAGGAACCATCGCCAGGCGTCCTTGGAAATCTCGCCAATAACTTTC
CTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTTGTCATATCTCTTCCCACCAACAT
GATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAATAATGAA
GAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGGTTTGATGATGACAACTCT
CCTTCCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCCTAAAACTTGGGTACATTACATTGCTGCTGAA
GAGGAGGACTGGGACTATGCTCCCTTAGTCCTCGCCCCCGATGACAGAAGTTATAAAAGTCAATATTTGAAC

AATGGCCCTCAGCGGATTGGTAGGAAGTACAAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAG
ACTCGTGAAGCTATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTG
TTGATTATATTTAAGAATCAAGCAAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCCGTCCT
TTGTATTCAAGGAGATTACCAAAAGGTGTAAAACATTTGAAGGATTTTCCAATTCTGCCAGGAGAAATATTC
AAATATAAATGGACAGTGACTGTAGAAGATGGGCCAACTAAATCAGATCCTCGGTGCCTGACCCGCTATTAC
TCTAGTTTCGTTAATATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCCTCATCTGCTACAAAGAA
TCTGTAGATCAAAGAGGAAACCAGATAATGTCAGACAAGAGGAATGTCATCCTGTTTTCTGTATTTGATGAG
AACCGAAGCTGGTACCTCACAGAGAATATACAACGCTTTCTCCCCAATCCAGCTGGAGTGCAGCTTGAGGAT
CCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTTGATAGTTTGCAGTTGTCAGTT
TGTTTGCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCACAGACTGACTTCCTTTCTGTCTTCTTC
TCTGGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTCACCCTATTCCCATTCTCAGGAGAAACT
GTCTTCATGTCGATGGAAAACCCAGGTCTATGGATTCTGGGGTGCCACAACTCAGACTTTCGGAACAGAGGC
ATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGAT
ATTTCAGCATACTTGCTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCTCAAAACCCACCAGTCTTG
AAAGCCCATCAGGCGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTATGATGATACC
ATATCAGTTGAAATGAAGAAGGAAGATTTTGACATTTATGATGAGGATGAAAATCAGAGCCCCCGCAGCTTT
CAAAAGAAAACACGACACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCA
CATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACT
GATGGCTCCTTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGGCCATATATA
AGAGCAGAAGTTGAAGATAATATCATGGTAACTTTCAGAAATCAGGCCTCTCGTCCCTATTCCTTCTATTCT
AGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGAA
ACCAAAACTTACTTTTGGAAAGTGCAACATCATATGGCACCCACTAAAGATGAGTTTGACTGCAAAGCCTGG
GCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGGTCTGCCAC
ACTAACACACTGAACCCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCACCATCTTT
GATGAGACCAAAAGCTGGTACTTCACTGAAAATATGGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATG
GAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGC
[0816] TTAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATCCATTCT
ATTCATTTCAGTGGACATGTGTTCACTGTACGAAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTAT
CCAGGTGTTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAATTTGGCGGGTGGAATGCCTTATTGGC
GAGCATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAGTGTCAGACTCCCCTGGGAATG
GCTTCTGGACACATTAGAGATTTTCAGATTACAGCTTCAGGACAATATGGACAGTGGGCCCCAAAGCTGGCC
AGACTTCATTATTCCGGATCAATCAATGCCTGGAGCACCAAGGAGCCCTTTTCTTGGATCAAGGTGGATCTG
TTGGCACCAATGATTATTCACGGCATCAAGACCCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCT
CAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCCACTGGAACCTTA
ATGGTCTTCTTTGGCAATGTGGATTCATCTGGGATAAAACACAATATTTTTAACCCTCCAATTATTGCTCGA
TACATCCGTTTGCACCCAACTCATTATAGCATTCGCAGCACTCTTCGCATGGAGTTGATGGGCTGTGATTTA
AATAGTTGCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCCTAC
TTTACCAATATGTTTGCCACCTGGTCTCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATGCCTGG
AGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACAGGAGTA
ACTACTCAGGGAGTAAAATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGCAGTCAAGAT
GGCCATCAGTGGACTCTCTTTTTTCAGAATGGCAAAGTAAAGGTTTTTCAGGGAAATCAAGACTCCTTCACA
CCTGTGGTGAACTCTCTAGACCCACCGTTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTTGGGTGCAC
CAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTCTACGACAAAACTCACACATGCCCA
CCGTGCCCAGCTCCAGAACTCCTGGGCGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTC
ATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTC
AACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACG
TACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTC
TCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAG
GTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGC
TTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCT
CCCGTGTTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAG
GGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTG
TCTCCGGGTAAA

[0817] 所构建的另外VWF片段和FVIIIFC-VWF异源二聚体在下面列出。

[0818] 表6:VWF片段和FVIII/VWF异源二聚体构建体

[0819]

构建体	描述	载体
VWF		
pSYN-VWF-001	FVIII信号肽D'D3区（长度为1-276个氨基酸6×His）	pcDNA 4
pSYN-VWF-002	FVIII信号肽D'D3区（长度为1-477个氨基酸6×His）	pcDNA 4
pSYN-VWF-003	FVIII信号肽D'D3区部分A1（长度为1-511个氨基酸6×His）	pcDNA 4
pSYN-VWF-004	FVIII信号肽D'D3A1区（长度为1-716个氨基酸6×His）	pcDNA 4
pSYN-VWF-006	D1D2D'D3-连接基-CK1	pcDNA 3.1
pSYN-VWF-008	全长WT-VWF	pcDNA 3.1
pSYN-VWF-009	D1D2D'D3区(1-276aa, 6×His)	pcDNA 3.1
pSYN-VWF-010	D1D2D'D3区(1-477aa, 6×His)	pcDNA 3.1
pSYN-VWF-011	D1D2D'D3区部分A1 (1-511aa, 6×His)	pcDNA 3.1
pSYN-VWF-012	D1D2D'D3A1区(1-716aa, 6×His)	pcDNA 3.1
pSYN-VWF-013	D1D2D'D3区(1-477aa, C336A/C379A, 6×His)	pcDNA 3.1
pSYN-VWF-014	FVIII信号肽-D'D3(1-477aa, C336A/C379A)-具有包含凝血酶位点的20aa连接基的单链Fc	pcDNA 4
pSYN-VWF-015	D1D2D'D3(1-477aa, C336A/C379A)-具有包含凝血酶位点的20aa连接基的单链Fc	pcDNA 4
pSYN-VWF-016	FVIII信号肽-D'D3(1-477aa, WT)-具有包含凝血酶位点的20aa连接基的单链Fc	pcDNA 4
pSYN-VWF-017	D1D2D'D3 (1-477aa, WT)-具有包含凝血酶位点的20aa连接基的单链Fc	pcDNA 4
pSYN-VWF-025	pLIVE中的D1D2D'D3区(1-477aa, 6×His)	pLIVE
pSYN-VWF-029	pLIVE中的D1D2D'D3区(1-477aa, C336A/C379A, 6×His)	pLIVE
pSYN-VWF-030	FVIII信号肽-D'D3(1-477aa, C336A/C379A)-具有包含凝血酶位点的48aa连接基的单链Fc	pcDNA 4
pSYN-VWF-031	D1D2D'D3(1-477aa, C336A/C379A)-具有包含凝血酶位点的48aa连接基的单链Fc	pcDNA 4
pSYN-VWF-032	FVIII信号肽-D'D3(1-477aa, WT)-具有包含凝血酶位点的48aa连接基的单链Fc	pcDNA 4
pSYN-VWF-033	FVIII信号肽-D'D3(1-477aa, WT)-具有包含凝血酶位点的35aa连接基的单链Fc	pcDNA 4
FVIII		
pSYN-FVIII-055	具有R336I和Y1680F的BDD-FVIII scFc	pBUD
pSYN-FVIII-056	具有R336I、R562和Y1680F的BDD-FVIII scFc	pBUD
pSYN-FVIII-057	具有Y1680F的BDD-FVIII scFc	pBUD
pSYN-FVIII-058	具有S488A的BDD-FVIII scFc	pBUD
pSYN-FVIII-059	具有R336I、R562K、S488A的BDD-FVIII scFc	pBUD
pSYN-FVIII-060	具有R336I、R562K、Y1680F的BDD-FVIII scFc	pBUD
pSYN-FVIII-061	具有R336I、R562K、S488A、Y1680F的BDD-FVIII scFc	pBUD
pSYN-FVIII-064	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有20aa凝血酶可切割连接基	pBUD
pSYN-FVIII-065	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-276aa)，并且其间具有20aa凝血酶可切割连接基	pBUD
pSYN-FVIII-083	具有R336I、S488A、R562K、Y1680F、E1984V的BDD-FVIII scFc	pBUD
pSYN-FVIII-086	BDD-FVIII scFc，6×(GGGGS)连接基在FVIII的C2和Fc之间	pBUD
pSYN-FVIII-095	具有S104C、R562K、Y1680F、G1960C的BDD-FVIII scFc	pBUD
pSYN-FVIII-101	来自FVIII-041的BDD-FVIII scFc克隆至pcDNA 3.3 Topo	pcDNA 3.3 Topo
pSYN-FVIII-102	BDD-FVIII（用于二硫键结合的M662C/D1828C；APC切割突变R336I/R562K；以及用于VWF结合的Y1680F突变）	pBUD
pSYN-FVIII-103	BDD-FVIII scFc (Y662C/T1828C)	pBUD
pSYN-FVIII-104	BDD-FVIII scFc (G655C/ST1788C)	pBUD
pSYN-FVIII-113	BDD-FVIII (R490A/H497A)可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有20aa凝血酶可切割连接基	pBUD

[0820]

pSYN-FVIII-114	BDD-FVIII (R490A/H497A)可切割scFc，第二Fc上具有VWF D'D3 (1-276)，并且其间具有20aa凝血酶可切割连接基	pBUD
pSYN-FVIII-126	BDD-FVIII scFc (M662C/D1828C)	pcDNA 3.3 Topo
pSYN-FVIII-127	BDD-FVIII scFc（用于二硫键结合的M662C/D1828C；APC切割突变R336I/R562K；以及用于VWF结合的Y1680F突变）	pcDNA 3.3 Topo
pSYN-FVIII-128	BDD-FVIII scFc (Y664C/T1826C)	pcDNA 3.3 Topo
pSYN-FVIII-129	pSYN-VIII-64背景中的突变R336I、R562K、R490A、H497A、N1224A	pBUD
pSYN-FVIII-130	pSYN-VIII-65背景中的突变R336I、R562K、R490A、H497A、N1224A	pBUD
pSYN-FVIII-131	pSYN-VIII-64背景中的突变R471A、Y487A、R490A、H497A、N1224A	pBUD
pSYN-FVIII-132	pSYN-VIII-65背景中的突变R471A、Y487A、R490A、H497A、N1224A	pBUD
pSYN-FVIII-135	具有R1645A/R1648A的BDD-FVIII scFc	pcDNA 3.3 Topo
pSYN-FVIII-136	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-137	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-276aa)，并且其间具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-145	具有R471A/Y487A、R490A/H497A的BDD-FVIII scFc	pcDNA 3.3 Topo
pSYN-FVIII-146	BDD-FVIII可切割scFc(R471A/Y487A)，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-147	BDD-FVIII可切割scFc(R471A/Y487A)，第二Fc上具有VWF D'D3 (1-276aa)，并且其间具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-148	BDD-FVIII可切割scFc (R1645A/R1648A)，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-149	BDD-FVIII可切割scFc (R1645A/R1648A)，第二Fc上具有VWF D'D3 (1-276aa)，并且具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-155	融合至单链Fc (R1645A/R1648A)的BDD-FVIII	pcDNA 4
pSYN-FVIII-159	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有35aa凝血酶可切割连接基	pBUD
pSYN-FVIII-160	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有48aa凝血酶可切割连接基	pBUD
pSYN-FVIII-164	BDD-FVIII可切割scFc (R490A/H497A, R1645A/R1648A)，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-165	BDD-FVIII可切割scFc (R336I/R562K, R490A/H497A, R1645A/R1648A)，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且具有20aa凝血酶可切割连接基	pcDNA 3.3 Topo
pSYN-FVIII-178	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有73aa凝血酶可切割连接基	pBUD
pSYN-FVIII-179	BDD-FVIII可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有98aa凝血酶可切割连接基	pBUD
pSYN-FVIII-180	BDD-FVIII (K2092A)可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有48aa凝血酶可切割连接基	pBUD
pSYN-FVIII-181	BDD-FVIII (F2093A)可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有48aa凝血酶可切割连接基	pBUD
pSYN-FVIII-182	BDD-FVIII (K2092A/F2093A)可切割scFc，第二Fc上具有VWF D'D3 (1-477aa, C336A/C379A)，并且其间具有48aa凝血酶可切割连接基	pBUD

[0821] 实施例6:蛋白质纯化

[0822] VWF片段的蛋白质纯化

[0823] VWF片段通过两步纯化法纯化。硫酸镍带电的IMAC(固定化金属亲合色谱)柱用于初次纯化,Fractogel DEAE离子交换柱用于最终纯化。详细纯化方法如下所述。

[0824] (a)在Nickel IMAC上进行的VWF片段的初次纯化

[0825] 将14mL Nickel IMAC Sepharose HP柱[XK26/3]用25mM HEPES、500mM NaCl、10mM咪唑和0.05％Tween-20(pH7.5)进行平衡。用100mL 1M HEPES(pH7.5)和600mL 5M NaCl调整大约7.2L VWF条件培养基。然后加入80mL 1M咪唑(pH7.5)达到10mM终浓度。然后将7.8L调整的VWF条件培养基在2-8℃以10mL/分钟[113cm/小时]上柱。洗涤步骤以13.3mL/分钟[150cm/小时]进行。首先，用25mM HEPES、500mM NaCl、10mM咪唑和0.05％Tween-20(pH7.5)以正向流{“向下流(DownFlow)”}进行2×柱体积(CV)洗涤。然后，用25mM HEPES、500mM NaCl、10mM咪唑和0.05％Tween-20(pH7.5)以反向流{“向上流(UpFlow)”}进行3×CV洗涤。最后，用25mM HEPES、500mM NaCl、10mM咪唑和0.05％Tween-20(pH7.5)以正向流{“向下流”}进行3×CV洗涤。洗脱以10×CV梯度进行至50％B1(25mM HEPES、500mM NaCl、500mM咪唑和0.05％Tween-20(pH7.5))。级分体积设定为10mL。然后，柱用100％B1剥离(strip)。然后是用25mM HEPES、500mM NaCl、10mM咪唑和0.05％Tween-20(pH7.5)洗涤。二次剥离用1N NaOH进行。然后柱用1M TRIS、1M NaCl(pH7.8)，然后是25mM HEPES、500mM NaCl、10mM咪唑和0.05％Tween-20(pH7.5)冲洗。最后，柱用5CV的DPBS+20％乙醇冲洗并储存在4℃。

[0826] (b)在Fractogel DEAE上进行的VWF片段的二次纯化

[0827] VWF片段的二次纯化在Fractogel DEAE(pH7.5)上进行。首先，用200mg Zwittergent 3-14两性离子洗涤剂调整20mL VWF Nickel IMAC洗出液(对应于VWF片段峰)，尝试在不使用变性或还原赋形剂的情况下破坏聚集物质。在洗涤剂溶解后，将蛋白质在室温下放置大约15分钟。然后，用4克海藻糖、1mL 10％Tween-20、5mL 1M HEPES(pH7.5)和174mL“Milli-Q”水调整蛋白质。平衡缓冲液“A12”为25mM HEPES、50mM NaCl、1％海藻糖、0.05％Tween-20(pH7.5)。洗脱缓冲液“B1”为25mM HEPES、1000mM NaCl、1％海藻糖、0.05％Tween-20(pH7.5)。洗脱以10CV梯度进行至50％B1，保持5+CV，然后一步至100％B1。然后柱用0.85％磷酸，然后是1M TRIS、1M NaCl(pH7.5)剥离。然后柱用1N NaOH、2M NaCl，然后是1M TRIS、1M NaCl(pH7.5)解吸。然后柱用25mM HEPES、100mM NaCl+20％乙醇(pH7.5)冲洗，以用于储存。

[0828] (c)FVIII-VWF异源二聚体的蛋白质纯化

[0829] FVIII-VWF异源二聚体首先用亲合柱(GE VIIISelect)，然后是Fractogal TMAE离子交换柱纯化。(McCue JT,Selvitelli K,Walker J,J Chromatogr A.2009年11月6日;1216(45):7824-30.Epub 2009年9月23日.)

[0830] 对于FVIII-155/VWF-31的纯化，将切向流过滤(TFF)步骤用于缓冲更换澄清的条件培养基。然后使用亲合色谱法捕集滤液中的靶蛋白。然后进行弱阴离子交换色谱步骤以减少HMW物质。分子的纯度和大小二者通过HPLC-SEC和SDS-PAGE评估。FVIII-155/VWF-31的不同结构域的存在通过蛋白质印迹进一步确认。分子的比活性相当于B-结构域缺失的FVIII。

[0831] (d)FVIII-VWF异源二聚体的凝血酶消化(图8)

[0832] FVIII-VWF-Fc异源二聚体或FVIII-Fc(对照)在凝血酶切割缓冲液(50mM Tris(pH7.4)、150mM NaCl、2mM CaCl2、5％甘油)中以1:10的比率与凝血酶混合。将反应在37℃下温育20分钟。将消化产物在4-12％还原tris-甘氨酸凝胶上跑电泳。未消化的蛋白质用作对照。条带通过考马斯亮蓝染色观察。

[0833] (e)通过Octet测定评估VWF结合FVIII-155/VWF-031的能力

[0834] VWF结合FVIII-155/VWF-031的能力通过基于生物膜层干涉(BLI)的测量法(Octet测定)在25℃下用ForteBio Octet 384仪器、使用Tris结合缓冲液(50mM Tris(pH7.2)、150mM NaCl、5mM $CaCl_2$)测定。用于确定FVIII结合的Octet测定基于人血管性血友病因子(hVWF)(Haematologic Technologies目录号HCVWF-0191)疏水固定化至APS生物传感器,然后结合1.0％牛血清白蛋白(Jackson ImmunoResearch目录号001-000-161)。简而言之,hVWF(38.5nM)在Tris缓冲液中稀释,并加载到APS生物传感器600秒,反应探针上产生大约3.0-3.5nm结合。对照APS探针在不存在hVWF的情况下用1.0％BSA加载,以进行参考扣除(reference subtraction)。在加载后,所有探针在Tris缓冲液中温育300秒以建立新的基线。随后,生物传感器探针在FVIII-155/VWF-031、FVIIIFc药物物质或rFVIII(60nM)的溶液中在室温下温育5分钟,然后是5分钟解离步骤。使用Octet数据分析软件,从扣除数据(反应探针减去参考探针)得到结合响应(nm)。如图15中所示,与rFVIIIFc和rFVIII的VWF结合亲合力相比,VWF与FVIII-155/VWF-031的结合亲合力严重受损。这表明通过FVIIIFc/VWF异源二聚体内的D'D3片段成功屏蔽FVIII免于全长VWF的结合。

[0835] 实施例7:VWF-FVIII相互作用是FVIII半衰期延长的限制因素

[0836] 大多数循环FVIII以FVIII-VWF复合物存在(>95％的血浆FVIII)。该FVIII-VWF相互作用促进通过VWF清除途径的FVIII清除,从而使VWF半衰期(T1/2)成为FVIII半衰期延长的限制。为评估该假说,在FVIII缺陷小鼠(HemA小鼠,其具有完整的VWF基因)和FVIII/VWF缺陷(FVIII-VWF双敲除(DKO))小鼠中测试通过Fc技术限制FVIII半衰期延长。

[0837] HemA小鼠或FVIII-VWF DKO小鼠用单静脉内剂量的rFVIII或rFVIIIFc(HemA小鼠中125IU/kg或DKO小鼠中200IU/kg)治疗。收集HemA小鼠中的血液样品最多72小时,或FVIII/VWF DKO小鼠中的血液样品最多8小时。然后血浆样品的FVIII活性通过FVIII显色测定测量。两种rFVIII变体的药代动力学(PK)曲线使用WinNonline程序分析。

[0838] 如表7和图9中所示,在FVIII/VWF DKO小鼠中,与rFVIII的$T_{1/2}$(即,$T_{1/2}$为0.25小时)相比,rFVIIIFc显示出是其4.8倍的$T_{1/2}$(即,$T_{1/2}$为1.2小时)。相比之下,在HemA小鼠中测试时,与rFVIII相比,rFVIIIFc的$T_{1/2}$仅是其1.8倍。rFVIIIFc的$T_{1/2}$为13.7小时,符合内源性鼠VWF半衰期。这表明FVIII-VWF相互作用是FVIII半衰期延长的限制因素。为了实现超过2倍的FVIII半衰期延长,FVIII-VWF相互作用必须消除。

[0839] 表7:HemA和FVIIII/VWF DKO小鼠中的FVIII PK

[0840]

测试分子	FVIII缺陷小鼠		FVIII/VWF缺陷小鼠	
	$T_{1/2}$(小时)	相比rFVIII的$T_{1/2}$比率	$T_{1/2}$(小时)	$T_{1/2}$比率
rFVIII	7.6	1	0.25	1
rFVIIIFc	13.7	1.8	1.2	4.8

[0841] FVIII显色测定

[0842] FVIII活性使用得自DiaPharma(lot#N089019)的COATEST SP FVIII试剂盒测量,并且所有温育在37℃板式加热器上震荡进行。

[0843] rFVIII标准品的范围为从100mIU/mL至0.78mIU/mL。将混合的正常人血浆测定对

照和血浆样品(用1×Coatest缓冲液稀释)加入Immulon 2HB 96孔板,一式两份(25μL/孔)。将新鲜制备的IXa/FX/磷脂混合物(50μL)、25μL 25mM $CaCl_2$和50μL FXa底物按顺序加入每个孔,每次加入之间温育5分钟。在温育底物后,加入25μL 20%乙酸终止颜色反应,并且OD405的吸光度用SpectraMAX plus(Molecular Devices)仪器测量。数据使用SoftMax Pro软件(5.2版)分析。最低定量限(LLOQ)为7.8mIU/mL。

[0844] 实施例8:VWF D'D3二聚体保护FVIII不被FVIII蛋白水解和清除(图10)

[0845] VWF片段的FVIII保护活性通过其保护内源性鼠FVIII在VWF缺陷小鼠中不被清除的能力评估。如表8第1列中列出的不同VWF片段(图1,实施例1)通过以100μg/小鼠其对应DNA构建体的高压注射引入VWF缺陷小鼠的血液循环。收集注射后48小时的血浆样品,并且鼠FVIII血浆活性通过FVIII显色测定测量。VWF表达水平通过VWF ELISA测量。

[0846] 测试的四个不同长度的VWF片段为276个、477个、511个和716个氨基酸。测试276至716个氨基酸范围以查找无VWF清除受体的结合结构域(716aa)的FVIII结合所需的VWF片段的长度(276aa)。全长VWF和D1D2D'D3CK多聚体用作FVIII保护的阳性对照。在血液循环中,通过D1D2结构域合成的VWF片段以二聚体存在,当它们在无D1D2结构域的情况下合成时以单体存在。

[0847] 高压注射后血浆中鼠FVIII活性的增加衡量VWF片段的FVIII保护效果。如表8和图10A-B中所示,D'D3片段的前276aa无FVIII保护活性,如由类似的注射前/后FVIII血浆水平所证明(图10A)。然而,其它VWF片段的引入导致FVIII血浆水平显著增加,表明那些VWF片段可保护FVIII不被其清除途径清除。

[0848] 表8:VWF片段引入前/后FVIII/VWF DKO小鼠的鼠FVIII血浆水平(DNA构建体在图1中示出)

[0849]

DNA构建体	编码VWF片段	FVIII活性-前(mIU/mL)		FVIII活性-48小时(mIU/mL)		VWF抗原-48小时(nM/mL)	
		平均值	标准偏差	平均值	标准偏差	平均值	标准偏差
VWF-001	$D'D3_{276aa}$	53	31	86	16	2.8	1.9
VWF-009	$D1D2D'D3_{276aa}$	45	20	65	17	1.8	1.3
VWF-002	$D'D3_{477aa}$	56	3	257	38	17.0	0.5
VWF-010	$D1D2D'D3_{477aa}$	42	11	387	22	8.2	1.6
VWF-003	$D'D3A1_{511aa}$	88	21	253	47	12.9	2.2
VWF-011	$D1D2D'D3A1_{511aa}$	63	42	360	15	9.3	2.3
VWF-004	$D'D3A1_{716aa}$	87	8	239	56		
VWF-012	$D1D2D'D3A1_{716aa}$	64	22	307	29		
VWF-006	D1D2D'D3CK	38	10	249	20	2.4	1.0
VWF-008	全长VWF	51	8	380	41	10.6	2.3

[0850] 注射后血浆FVIII活性和包含全长VWF的D'D3结构域的VWF片段的血浆抗原水平的比率在表8中列出。类似的注射后FVIII/VWF比率从全长VWF和VWF片段的两种二聚体形式观察,意指这两种VWF片段二聚体提供与全长VWF相同的FVIII保护。此外,从VWF片段二聚体同种型与其对应的单体相比观察到FVIII/VWF比率高三倍:D'D3(477aa)二聚体具有38.7mIU/nmol的FVIII/VWF比率;D'D3(477aa)单体具有11.6mIU/nmol的FVIII/VWF比率;D'D3A1(511aa)二聚体具有32.9mIU/nmol的FVIII/VWF比率;并且D'D3(511aa)单体具有13.8mIU/nmol的FVIII/VWF比率,表明与其对应的单体相比,VWF片段的二聚体同种型提供更好的FVIII保护。

[0851] 表9:全长D'D3片段的FVIII保护效果

[0852]

DNA构建体	编码VWF片段	多聚体状态	FVIII/VWF (mIU/nmol)	
			平均值	(标准偏差)
VWF-002	$D'D3_{477aa}$	单体	11.6	(4.4)
VWF-010	$D1D2D'D3_{477aa}$	二聚体	38.7	(11.7)
VWF-003	$D'D3A1_{511aa}$	单体	13.8	(1.3)
VWF-011	$D1D2D'D3A1_{511aa}$	二聚体	32.9	(5.5)
VWF-008	全长VWF	多聚体	31.1	(6.7)

[0853] 高压注射:

[0854] 高压注射是非病毒基因递送至小型动物例如小鼠和大鼠的肝脏的高效和安全方

法。它最初描述为动物体重十倍体积的不含内毒素的裸露质粒DNA/盐水溶液在约5-7秒内的快速注射。裸露质粒DNA包含所关注的基因，并且肝脏产生的注射DNA的靶蛋白可在注射后24小时内检测。然后收集血浆样品研究表达蛋白的治疗性质。

[0855] 对于本专利申请进行的所有高压注射，将溶于0.9％无菌盐水溶液的2ml质粒DNA通过在约4-7秒内尾静脉注射递送至体重为20-35克的小鼠。密切监测小鼠前两个小时，直到恢复正常活性。通过眶后采血收集血液样品后，然后得到血浆样品并储存在-80℃，用于进一步分析。

[0856] VWF ELISA：

[0857] 以0.5ug/孔将山羊抗人VWF抗体(亲合纯化，Affinity Biological，GAVWF-AP)用作捕获抗体，并且VWF-EIA-D(Affinity Biologicals，VWF-EIA-D，1:100稀释)用作检测抗体以用于VWF ELISA。ELISA测定在标准ELISA程序后进行，TMB用作HRP底物，PBST/1.5％BSA/0.5M NaCl缓冲液用作阻断和结合缓冲液。测定标准范围为100ng至0.78ng，并且测定的最低定量限(LLOQ)为7.8ng/mL。

[0858] 实施例9：全长VWF D'D3片段的共施用延长FVIII-VWF DKO小鼠中的rBDD-FVIII半衰期(图11)

[0859] 实施例8已示出全长D'D3片段可保护内源性FVIII不被其清除途径清除。为了进一步评估D'D3蛋白的FVIII保护活性，通过静脉注射给FVIII-VWF DKO小鼠共施用B结构域缺失的FVIII(rBDD-FVIII)和D'D3二聚体(VWF-010)或rBDD-FVIII和D'D3单体(VWF-002)(200IU/kg rBDD-FVIII、770μg/kgD'D3二聚体以及590μg/kg D'D3单体)。然后通过其注射后血浆活性监测rBDD-FVIII的PK曲线。由于D'D3片段的较短体内半衰期，在初始共注射后三小时，通过相同途径施用另外剂量的D'D3，以保持所需的D'D3血浆水平。

[0860] 对于PK分析，通过注射后5分钟、30分钟、1小时、2小时、4小时和6小时眶后采血获得血浆样品，血浆FVIII活性和D'D3抗原水平通过FVIII显色测定和VWF ELISA进行分析。

[0861] 如图11和表10所示，D'D3单体延长rBDD-FVIII半衰期2.5倍，并且提高其回收率1.8倍。D'D3二聚体延长rBDD-FVIII半衰期4.1倍，并且提高其回收率3.5倍。提高的平均驻留时间、清除和AUC也在两种D'D3同种型中观察到。然而，与其单体形式相比，D'D3二聚体在所有PK参数中实现更好的结果。

[0862] 总而言之，全长D'D3的共注射保护FVIII不被其清除途径清除，如rBDD-FVIII的改善PK曲线所示。这些发现的潜在临床价值需要进一步评估。

[0863] 表10：与D'D3片段共施用时，FVIII-VWF DKO小鼠中的BDD-FVIII PK参数

[0864]

处理	5分钟回收率(%)	$T_{1/2}$(小时)	MRT(小时)	Cl (mL/小时/kg)	Vss (mL/kg)	AUC_D(小时*kg*mIU/mL/mIU)	$T_{1/2}$倍数增加	回收率倍数增加
rBDD-FVIII	25	0.23	0.24	407.72	133.14	0.0025		
rBDD-FVIII VWF-002	44	0.57	0.58	151.93	124.63	0.0066	2.5	1.8
rBDD-FVIII VWF-010	87	0.95	0.98	71.48	97.54	0.014	4.1	3.5

[0865] 实施例10:用D1D2结构域及其二聚体同种型合成的D'D3单体具有相同的FVIII保护活性,并且在FVIII-VWF DKO小鼠中还延长FVIIIFc的半衰期约4倍(图12)。

[0866] 为了定量D'D3结构域的FVIII保护能力,并且确定D'D3二聚化是否为其FVIII保护活性所必须的,两个DNA构建体(即,VWF-025(包含编码D1D2D'D3的DNA序列)和VWF-029(包含具有C336A和C379A突变的D1D2D'D3密码子DNA))中的每个通过高压注射施用至FVIII/VWF DKO小鼠。该注射导致FVIII/VWF DKO小鼠中的D'D3二聚体(VWF-025)或单体表达(VWF-029)。在高压注射后第5天,单静脉内剂量的rFVIIIFc以200IU/kg施用,并且血浆样品在rFVIIIFc静脉内注射后5分钟、4小时、8小时、16小时、24小时、31小时、40小时、55小时、66小时收集。以相同的剂量在初次受试的FVIII-VWF DKO小鼠中进行的rFVIIIFc PK研究用作rFVIIIFc半衰期基线。血浆FVIII活性通过FVIII显色测定分析。血浆D'D3水平通过VWF ELISA测量,并且rFVIIIFc PK曲线使用WinNonlin程序分析。

[0867] 如表11和图12中所示,通过循环中的VWF D'D3片段,rFVIIIFc的初始回收率从42%增加至D'D3二聚体的75%和D'D3单体的60%。rFVIIIFc的$T_{1/2}$也从2.5小时分别增加至9.3小时和9.2小时。类似于$T_{1/2}$,提高平均驻留时间、清除率和体积分布也在表达D'D3单体和二聚体的小鼠中观察到。总体上,观察到rFVIIIFc的半衰期提高约8倍,并且表达D'D3单体和二聚体二者的小鼠中AUC提高6倍。与其二聚体形式相同,使用VWF的原肽(D1D2)合成的全长VWF的D'D3单体足以提供与全长VWF分子相同的完全FVIII保护效果。

[0868] 在FVIII/VWF DKO小鼠中,WT-FVIII具有0.25小时的$T_{1/2}$。Fc融合技术使FVIII的$T_{1/2}$增加至1.2小时,增加约4.8倍。当Fc融合技术与D'D3结构域组合时,FVIII的$T_{1/2}$增加至9.3小时(D'D3二聚体)和9.2小时(D'D3单体),总共增加约37倍。(表10)该结果显示了Fc融合和D'D3VWF片段对FVIII半衰期延长的协同效应。

[0869] 表11:血液循环中含/不含D'D3片段的rFVIIIFc PK参数

[0870]

处理	5分钟回收率(%)	$T_{1/2}$(小时)	MRT(小时)	Cl (mL/小时/kg)	Vss (mL/kg)	AUC_D(小时*kg*mIU/mL/mIU)	$T_{1/2}$倍数增加	AUC D倍数增加
rFVIIIFc	43	1.2	0.76	39.5	67.0	0.025		
rFVIIIFc VWF-025	75	9.3	11.1	6.1	67.6	0.164	7.8	6.6
rFVIIIFc VWF-029	60	9.2	11.3	6.7	75.7	0.149	7.7	6.0

[0871] 实施例11:HemA小鼠中的FVIII-VWF异源二聚体PK

[0872] FVIII-VWF异源二聚体的领先候选物(例如FVIII-155/VWF-031)的PK曲线在HemA小鼠中测试,以评估其屏蔽FVIII免受内源性VWF作用的能力,及其延长FVIII半衰期的能力。

[0873] HemA小鼠用200IU/kg的单静脉内剂量的领先候选物处理,然后在5分钟、4小时、8小时、24小时、48小时、72小时、96小时和120小时收集血浆样品,血浆活性通过FVIII显色测定测试,并且FVIII变体的半衰期通过WinNonlin程序计算。

[0874] 在最佳FVIII/VWF异源二聚体构型中,FVIII结合至内源性VWF被完全抑制,因此rFVIII的基线半衰期从7.6小时减少至0.25小时,如实施例7中所示。当D'D3片段与FVIII非共价缔合时,观察到约8倍的半衰期益处(实施例9)。在FVIII/VWF异源二聚体的领先候选物中,VWF片段与FVIII分子共价缔合,能够实现更好的FVIII保护。本发明打开了进一步延长FVIII半衰期超过两倍上限的途径,通过可用的半衰期延长技术的组合,在不久将来HemA患者可期待更好的长期作用FVIII变体。

[0875] FVIII-155/VWF-031的PK曲线在HemA和FVIII/VWF DKO小鼠中测试,以评估D'D3片段屏蔽FVIII部分免受内源性VWF作用的能力。HemA或FVIII/VWF DKO小鼠用200IU/kg的单静脉内剂量的FVIII-155/VWF-031处理,然后在给药后5分钟、8小时、24小时和48小时收集血浆样品。血浆样品的FVIII活性通过FVIII显色测定测试,并且FVIII-155/VWF-031的半衰期使用WinNonlin程序计算。

[0876] 与rFVIIIFc和rFVIII相比,与固定化VWF结合的严重受损通过FVIII-155/VWF-031的生物膜层干涉(图15,Octet,ForteBio Inc.,Menlo Park,CA)检测。这显示分子中的D'D3结构域成功阻断FVIII结合至天然的VWF分子。因此,预期在两个不同的小鼠品系中rFVIII-155/VWF-031的类似半衰期。研究结果在图16和表12A中列出。根据预测,rFVIII-155/VWF-031在HemA和FVIII/VWF DKO小鼠中具有相当的PK曲线,表明FVIIIFc/VWF异源二聚体的半衰期与内源性VWF的半衰期无关。结果显示,通过VWF D'D3结构域抑制rFVIIIFc与内源性VWF之间的相互作用允许消除FVIII半衰期上限,并且打开了延长FVIII半衰期的可能性,使其超过可在无VWF D'D3结构域时实现的半衰期(约野生型FVIII的两倍)。

[0877] 表12A:FVIII/VWF DKO小鼠和HemA小鼠中的FVIII-155/VWF-031PK

[0878]

处理	5分钟回收率(%)	$T_{1/2}$(小时)	MRT(小时)	Cl(mL/小时/kg)	Vss(mL/kg)	AUC_D(小时*kg*mIU/mL/mIU)
FVIII-155/VWF-031 DKO	49	9.9	6.9	11.6	80.5	0.09
FVIII-155/VWF-031	69	10.8	707	11.9	92.1	0.08

[0879]

HemA						

[0880] D'D3结构域的FVIII保护能力通过比较FVIII/VWF DKO小鼠中FVIII-155/VWF-031与FVIIIFc的$t_{1/2}$进行评估。在单静脉内施用后,在5分钟、8小时、24小时和48小时收集FVIII-155/VWF-031的血液样品,并且在5分钟、1小时、2小时、4小时、6小时和8小时收集FVIIIFc的血液样品。血浆样品的FVIII活性通过FVIII显色测定测试,并且FVIII-155/VWF-031的半衰期使用WinNonlin程序计算。

[0881] 图16B和表12B显示,与DKO小鼠中的rFVIIIFc相比,FVIII-155/VWF-031的PK曲线显著增加:$t_{1/2}$增加约6倍;并且清除率和AUC增加约5倍。该结果显示,FVIIIFc/VWF异源二聚体中的D'D3结构域保护FVIII部分免受一些清除途径的清除,从而提供全长VWF通常提供的一些保护。该结论还在HemA小鼠中确认。当与HemA小鼠中的rFVIIIFc比较时,rFVIII-155/VWF-031显示出更短的$t_{1/2}$和更小的AUC,意指在该构型中,D'D3结构域(VWF-031)成功阻止FVIII蛋白(rFVIII-155)结合内源性VWF,该结合具有一定程度的半衰期延长特性,以及FVIII半衰期限制性质。全长VWF为250kDa,并且形成多聚体,使得内源性VWF可达最多2MDa,因此符合VWF的55kDa D'D3区不提供与该情形中非常大的内源性VWF通常所提供相同的保护的假说。由于VWF片段阻止内源性VWF结合rFVIII-155/VWF-031,因此在该具体构建体中,HemA小鼠中的半衰期减少。因此,表12B中的结果表明rFVIII-155/VWF-031分子能够阻止FVIII半衰期延长因子(内源性VWF)结合rFVIII-155/VWF-031。然而,实验显示移除FVIII半衰期限制因子会打开延长FVIII蛋白的半衰期超过前面所示的1.5倍或2倍的可能性。当FVIII与图4中所示的其它半衰期延长元件组合时,可实现FVIII的2倍半衰期延长上限的突破。

[0882] 表12B:FVIII/VWF DKO小鼠中的FVIII-155/VWF-031和FVIIIFc P K

[0883]

处理	5分钟回收率(%)	$T_{1/2}$(小时)	MRT(小时)	Cl(mL/小时/kg)	Vss(mL/kg)	AUC_D(小时*kg*mIU/mL/mIU)
FVIIIFc DKO	43	1.6	1.9	63.9	123.2	0.02
FVIII-155/VWF-031 DKO	49	9.9	6.9	11.6	80.5	0.09
倍数增加		6.2	3.6	5.5		4.5
FVIII-155/VWF-031 HemA	69	10.8	7.7	11.9	92.1	0.08
FVIIIFc HemA	86	16.4	20.3	2.9	57.7	0.35

[0884]

[0885] 实施例12:FVIII/D'D3异源二聚体的D'D3-Fc连接基的优化(图13)

[0886] 为允许rFVIIIFc逃脱VWF清除途径并且消除2倍FVIII半衰期延长上限,将VWF D'D3片段掺入到rFVIIIFc分子(图2),得到FVIIIFc/VWF异源二聚体。为了消除rFVIIIFc和内源性VWF之间的相互作用,并且使D'D3FVIII保护潜力最大化,调整D'D3结构域和Fc区之间的连接基,以允许最佳的FVIII/D'D3结合。更佳的连接基将允许D'D3结构域具有比次佳的

连接基构建体更强的FVIII保护。这可通过FVIII/VWF DKO小鼠中DNA构建体的高压注射测试。更佳的构建体将产生更高的FVIIIFc/D'D3异源二聚体的稳态蛋白质表达。

[0887] 将三种不同的FVIIIFc/D'D3异源二聚体(图3,实施例3)工程化以用于优化连接基选择。D'D3结构域和Fc区之间的可能连接基在表13中列出。那些DNA构建体通过以100μg/小鼠的高压注射("HDI")施用至FVIII/VWF DKO小鼠,并且在HDI后48小时收集血浆样品。循环FVIIIFc/D'D3异源二聚体活性通过FVIII显色测定分析。

[0888] 研究结果在图13中示出。在HDI后48小时,FVIII-064和FVIII-159达到类似的表达水平,表明20aa连接基和35aa连接基促进类似水平的FVIII/D'D3相互作用。在另一方面,FVIII-160显示出表达显著高于FVIII-064,意指与20aa和35aa连接基相比,48aa连接基允许更好的FVIII/D'D3结合。

[0889] VWF片段和Fc区之间的最佳连接基是FVIIIFc/VWF异源二聚体的关键要素之一。寻找最佳连接基将允许FVIII和VWF片段之间的最佳相互作用,阻止FVIII结合内源性VWF,使FVIII逃脱VWF清除途径,并且延长FVIII半衰期超过血浆VWF半衰期。

[0890] 表13:D'D3和Fc片段之间的不同连接基

[0891]

DNA 构建体	D'D3和Fc之间的连接基
FVIII-064 (SEQ ID NO: 92)	20 aa=IDGGGGSGGGGSLVPRGSGG
FVIII-159 (SEQ ID NO: 93)	35 aa=ISGGGGSGGGGSGGGGSGGGGSGGGGSLVPRGSGG
FVIII-160 (SEQ ID NO: 94)	48 aa=ISGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSLVPRGSGGGGSGGGGS

[0892] 实施例13:单链FVIII稳定性

[0893] 单链FVIII蛋白可比其双链同种型更稳定。为了测试该假说,制备两个DNA构建体:FVIII-136(含D'D3结构域的可加工FVIIIFc)和FVIII-148(含D'D3结构域的单链(SC)FVIIIFc,其包含R1645A/R1648A突变,以阻止FVIII重链和轻链之间的切割)。

[0894] 将两种质粒通过高压注射施用至FVIII/VWF DKO小鼠。在注射后24小时和48小时收集血浆样品,以测量两种FVIIIFc/D'D3同种型的表达水平。如图14中所示,在两个时间点观察到SC-FVIIIFc/D'D3构建体(FVIII-148)的更好表达趋势(p=0.12,p=0.19),表明单链FVIII可比其双链同种型(FVIII-136)更稳定或更好地表达。将对两种FVIII同种型的PK曲线,及其细胞培养物表达水平进一步研究。单链FVIII同种型可能用于替代常规的双链同种型,以实现更好的蛋白质生成和更长的体内FVIII半衰期。

[0895] 实施例14:聚乙二醇化

[0896] 一个或多个聚乙二醇(PEG)分子可连接至FVIII蛋白、VWF片段或它们二者的任何区域内。由于根据晶体结构FVIII的表面没有游离的半胱氨酸(PDB:2R7E,Shen等,Blood 111:1240(2008);PDB:3CDZ,Ngo,Structure,16:597-606(2008)),因此一种方法是将含半胱氨酸的肽(如,GGGSGCGGGS)(SEQ ID NO:107)插入或连接至FVIII蛋白、VWF片段或它们二者。然后可将包含马来酰亚胺的PEG分子特异性缀合至重组FVIII蛋白上引入的半胱氨酸。简而言之,包含Cys插入的重组FVIII蛋白可通过标准分子技术构建,并且在哺乳动物表达系统(如,HEK 293、CHO、BHK21、PER.C6和CAP细胞)中表达的重组FVIII蛋白可通过亲合和离

子交换色谱纯化。纯化的重组FVIII蛋白通过三(2-羧乙基)膦(TCEP)还原,以暴露引入的半胱氨酸的巯基,然后与马来酰亚胺PEG反应。测试所得重组FVIII蛋白的促凝血活性和延长的半衰期。

[0897] 将PEG连接至美国专利申请号61/670,553中公开的位置中至少一个或其它合适的插入位点,该专利申请全文以引用方式并入本文。聚乙二醇化重组FVIII蛋白的FVIII活性使用FVIII显色测定进行分析。聚乙二醇化重组FVIII蛋白的PK在HemA小鼠和FVIII-VWF DKO小鼠中分析,如上所述。

[0898] 实施例15:HemA和FVIII/VWF双敲除(DKO)血浆中的FVIII稳定性

[0899] 不同FVIIIFc融合的血浆稳定性在HemA或FVIII/VWF双敲除(DKO)血浆中测试。对于稳定性测定,将5IU/ml各种FVIIIFc蛋白与小鼠HemA或DKO血浆在37℃下温育。在不同时间点收集等分试样以通过FVIII显色测定测量活性。每个时间点的活性测量两次,并且将平均活性绘制为时间的函数。

[0900] 对于FVIIIFc免疫沉淀测定,将5μg FVIIIFc与250μl PBS或小鼠DKO血浆在37℃下温育24小时。通过在室温下加入5μg绵羊抗FVIII多克隆抗体(ab61370)1小时以及100μl蛋白A珠粒使FVIIIFc免疫沉淀。在4×1ml PBS洗涤后,将珠粒重悬于50μl 1×还原SDS-PAGE缓冲液中。煮沸后,将20μl样品(即约1μg FVIIIFc)上样到4-15%Bio-Rad无染色凝胶上。凝胶通过Bio-rad系统成像,然后通过FVIII抗重链抗体(GMA012)进行蛋白质印迹分析。

[0901] HemA和DKO血浆中FVIIIFc(双链FVIII分子,其具有单独的FVIII重链和轻链,通过非共价相互作用结合在一起)的活性随时间推移而减少(图18A)。由于缺乏VWF介导的保护,FVIIIFc活性损失在DKO血浆中更明显。这种FVIII活性损失主要是由于FVIII重链(HC)的解离或降解。在DKO血浆中温育24小时后观察到FVIIIFc重链减少约75%(图18B)。对于轻链(LC)(数据未示出)或未加工/单链FVIIIFc(即其中轻链和重链仍然共价结合在一起的FVIII分子,凝胶图片中的顶部条带)未观察到显著减少(图18B)。

[0902] 对于VWF增加FVIII的体内稳定性的假设,我们测试了嵌合蛋白-FVIII-VWF异源二聚体(FVIII155:VWF31,其中VWF D'D3通过Fc共价结合至FVIII)是否在Hem A和DKO血浆中更稳定。如图19中的血浆稳定性数据所示,D'D3的存在增加了HemA和DKO血浆中FVIIIFc的稳定性。无D'D3的单链FVIIIFc用作这些实验的对照(scFVIII)。从图19可知,单链FVIII比双链FVIIIFc更稳定;然而D'D3的存在显著地进一步增加了单链FVIIIFc分子的血浆稳定性。这暗示D'D3不仅通过将重链和轻链保持在一起,而且通过一些其它未知机制来稳定FVIII。

[0903] 实施例16:弗林蛋白酶/PACE用于VWF加工

[0904] VWF是独特的蛋白质,在这个意义上其包含非常大的原肽(即VWF的D1D2结构域,约85kDa)。VWF原肽作为VWF分子正确折叠的内部分子伴侣。测试两种酶PC5和弗林蛋白酶(PACE)的VWF加工。将VWF031构建体(D1D2D'D3Fc)与各种浓度的PC5或PACE瞬时共转染至HEK293细胞中。在四天后,收集组织培养基并经过蛋白A沉降。对于从D'D3Fc移除原肽(D1D2),即使在低浓度(2.5%)下,弗林蛋白酶(PACE)也比10%PC5高效(图20)。移除D1D2是重要的,因为D1D2的存在牵涉于阻止D'D3与FVIII的相互作用。

[0905] 实施例17:FVIII-VWF异源二聚体中的VWF片段阻止FVIII与全长VWF的相互作用

[0906] ForteBio octet仪器用于测试FVIII构建体155/VWF31异源二聚体结合全长VWF

(图21A)。对于结合测定,全长VWF通过使用APS传感器,然后用1％BSA阻断来捕集。在阻断后,测试不同FVIII构建体的VWF结合。根据预测,野生型FVIII和FVIIIFc强结合至VWF传感器。已知具有对VWF的低亲合力或无亲合力的FVIII Y1680F突变体,显示出VWF结合显著减少。FVIII155/VWF31异源二聚体完全不结合全长VWF,确认屏蔽FVIII免受FVIII-VWF异源二聚体中D'D3的作用。

[0907] 以相反方向进行相同的实验,以确定FVIII-VWF异源二聚体中的D'D3部分是否可与非共价结合至D'D3的其它FVIII分子相互作用。如图21B中所示,当固定化在蛋白G传感器上时,VWF31(D'D3Fc)构建体单独可强烈结合FVIII,然而FVIII155:VWF31异源二聚体中的D'D3不显示出与FVIII的任何结合。蛋白G单独结合FVIII作为对照。这些结合实验确认异源二聚体中的D'D3可仅与一个与其共价结合的FVIII分子相互作用,并且阻止FVIII与全长野生型VWF分子相互作用。

[0908] 为确定VWF D'D3与FVIII分子的准确结合亲合力,用VWF031进行表面等离子共振实验(图22)。使用抗人IgG捕集VWF031构建体(D'D3Fc),并且使B-结构域缺失的FVIII通过包含D'D3Fc的芯片。观察到FVIII约10nM的K_D。与全长野生型VWF分子相比,该亲合力低约25倍,并且类似于前面的文献所报道。

[0909] 实施例18:D'D3和Fc之间不同连接基长度对异源二聚体活性和PK的影响

[0910] 为检验D'D3和Fc之间凝血酶可切割连接基长度的变化是否对FVIII-VWF异源二聚体的PK和活性具有任何影响,将不同的VWF构建体与FVIII 155一起共表达。测试表14A中列出的三个不同连接基长度的构建体(VWF031、VWF035和VWF036)。将每种质粒与FVIII155质粒混合(实施例5)并转染至HEK 293细胞。在转染后第四天,收集细胞培养基并且浓缩至10IU/ml FVIII显色活性。

[0911] 然后将浓缩的细胞培养基以100IU/10mL/kg的剂量施用至8-12周龄的FVIII/VWF DKO小鼠。在给药后5分钟、8小时、16小时、24小时、32小时和48小时收集血浆样品。血浆样品的FVIII活性通过FVIII显色测定分析,并且半衰期使用WinNonlin-Phoenix程序计算。

[0912] 如图23中所示,当D'D3和Fc片段之间的连接基长度从48aa增加至73aa或98aa时,对应的FVIIIFc/VWF异源二聚体的半衰期增加并且分别达到12.2小时和13.3小时。这表示48aa长变体增加1.5至1.6倍。迄今为止,98aa连接基是利用D'D3片段的FVIII保护活性的最佳连接基,并且将其掺入到FVIIIFc/VWF异源二聚体,以进一步增加其半衰期。

[0913] 为比较连接基对FVIII活性的影响,对表达不同FVIII-VWF异源二聚体的细胞的组织培养基进行FVIII显色和aPTT测定。虽然与异源二聚体构建体的显色活性相比,aPTT活性减少2倍,但未发现各种连接基之间的显著差异,当连接基还包含凝血酶位点附近的PAR1位点时除外(表14B)。

[0914] 表14A:VWF D'D3和Fc之间的可变连接基的序列

[0915]

DNA 构建体	D'D3 和 Fc 之间的连接基
VWF031	48 aa = I S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S L V P R G S G G G G S G G G G S (SEQ ID NO: 95)
VWF035	73aa= I S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S L V P R G S G G G G S G G G G S (SEQ ID NO: 96)
VWF036	98aa=I S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S G G G G S L V P R G S G G G G S G G G G S (SEQ ID NO: 97)

[0916] 表14B:具有不同连接基长度的异源二聚体的活性

[0917]

样品ID	样品描述	D'D3和Fc之间的连接基长度(aa)	显色IU/mL	aPTT IU/mL	显色/aPTT
1	FVIII Fc 155+VWF15	20	1.81	0.85	2.14
2	FVIII Fc 155+VWF31	48	2.32	1.05	2.21
3	FVIII Fc 155+VWF33	35	2.21	1.02	2.16
4	FVIII Fc 155+VWF35	73	2.65	1.24	2.14
5	FVIII Fc 155+VWF36	98	2.75	1.11	2.47
6	FVIII Fc 155+VWF39	26 (凝血酶+PAR1)	1.85	1.21	1.53

[0918] 实施例19:使用分选酶将FVIII与VWF片段连接

[0919] 在另一个方面,通过使用分选酶介导的体外蛋白连接法将VWF片段(如D1D2D'D3或D'D3结构域)连接至FVIII。在一个实例中,将金黄色葡萄球菌(Staphylococcus aureus)分选酶A(LPXTG)识别基序引入VWF片段的C-末端并且将Gly(n)残基引入FVIII的N-末端(其中甘氨酸残基的数量可变)。所用的FVIII分子可以是单链的或双链的。催化转肽反应的分选酶将VWF片段共价连接至FVIII。相反方向的识别基序可用于连接这两种蛋白质,其中使在FVIII在N-末端具有LPXTG基序,并且使VWF片段在C-末端具有Gly(n)(参见图24:分选酶连接的实例)。LPXTG基序和甘氨酸残基可被其它分选酶识别序列替代。

[0920] 还制备了包含分选酶A识别序列Fc融合蛋白的VWF片段。对于Fc融合构建体,将VWF D1D2D'D3片段通过包含分选酶识别序列和凝血酶切割位点的GS连接基与IgG的Fc区融合(表15和16)。一旦蛋白质表达并在蛋白A柱上纯化,Fc区即可通过凝血酶切割移除。然后所得的具有分选酶A识别位点的VWF片段可用于与FVIII分子连接(图24-分选酶连接的实例-E列)。

[0921] pSYN-VWF-051具有54个氨基酸的连接基,其中VWF片段和Fc区之间具有分选酶和凝血酶位点。编码54个氨基酸的连接基(ISGGGGSGGG GSGGGGSGGG GSGGGGSGGG GSLPETGALR PRVVGGGGSG GGGS)(SEQ ID NO:98)和一部分Fc区的DNA片段的合成通过外包进行(Genewiz序列号10-210746313,如下所示)。将Genewiz构建体的片段亚克隆至EcoRV/RsRII消化的pSYN-VWF-031。

[0922] Genewiz-序列号-10-210746313(SEQ ID NO:99)

[0923]
AGGAGCCGATATCTGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGGCGGTGGAGGTTCCGGCG
GTGGAGGTTCCGGTGGCGGGGGATCCGGTGGCGGGGGATCCTTACCTGAAACTGGAGCCCTGCGGCCCC
GGGTCGTCGGCGGTGGAGGTTCCGGTGGCGGGGGATCCGACAAAACTCACACATGCCCACCGTGCCCAG
CTCCAGAACTCCTGGGCGGACCGTCAGTCTT

[0924] 包含N-末端五甘氨酸的单链FVIII的序列在表17和18中示出。

[0925] 表15:pSYN-VWF051(具有分选酶A识别基序和VWF片段和Fc之间的凝血酶可切割连接基的VWF D1D2D'D3Fc)的核苷酸序列(SEQ ID NO:100)

```
      1  ATGATTCCTG CCAGATTTGC CGGGGTGCTG CTTGCTCTGG CCCTCATTTT
     51  GCCAGGGACC CTTTGTGCAG AAGGAACTCG CGGCAGGTCA TCCACGGCCC
    101  GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG
    151  TACAGCTTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA
    201  ACGCTCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC
    251  TCTCCGTGTA TCTTGGGGAA TTTTTTGACA TCCATTTGTT TGTCAATGGT
[0926] 301  ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG
    351  GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCCT
    401  ATGGCTTTGT GGCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG
    451  TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTGGCA ACTTTAACAT
    501  CTTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC
    551  CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT
    601  GAACGGGCAT CTCCTCCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT
```

```
651    GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT
701    TTGCCCGCTG CCACCCTCTG GTGGACCCCG AGCCTTTTGT GGCCCTGTGT
751    GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CCTGCCCTGC
801    CCTCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAATG GTGCTGTACG
851    GCTGGACCGA CCACAGCGCG TGCAGCCCAG TGTGCCCTGC TGGTATGGAG
901    TATAGGCAGT GTGTGTCCCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT
951    CAATGAAATG TGTCAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG
1001   GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGTCCCTGC
1051   GTGCATTCCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG
1101   CAACACCTGC ATTTGCCGAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT
1151   GTCCAGGGGA GTGCCTTGTC ACTGGTCAAT CCCACTTCAA GAGCTTTGAC
1201   AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGGA
1251   TTGCCAGGAC CACTCCTTCT CCATTGTCAT TGAGACTGTC CAGTGTGCTG
1301   ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGTCCG GCTGCCTGGC
1351   CTGCACAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA
1401   TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC
1451   ATACAGTGAC GGCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG
1501   GACTGGGATG GCCGCGGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC
1551   CGGGAAGACC TGCGGCCTGT GTGGGAATTA CAATGGCAAC CAGGGCGACG
1601   ACTTCCTTAC CCCCTCTGGG CTGGCGGAGC CCCGGGTGGA GGACTTCGGG
1651   AACGCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG
1701   CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGGTTCTCC GAGGAGGCGT
1751   GCGCGGTCCT GACGTCCCCC ACATTCGAGG CCTGCCATCG TGCCGTCAGC
1801   CCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACGTGTGCT CCTGCTCGGA
1851   CGGCCGCGAG TGCCTGTGCG GCGCCCTGGC CAGCTATGCC GCGGCCTGCG
1901   CGGGGAGAGG CGTGCGCGTC GCGTGGCGCG AGCCAGGCCG CTGTGAGCTG
1951   AACTGCCCGA AAGGCCAGGT GTACCTGCAG TGCGGGACCC CCTGCAACCT
2001   GACCTGCCGC TCTCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCTGCC
2051   TGGAGGGCTG CTTCTGCCCC CCAGGGCTCT ACATGGATGA GAGGGGGGAC
[0927] 2101   TGCGTGCCCA AGGCCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA
2151   GCCAGAAGAC ATCTTCTCAG ACCATCACAC CATGTGCTAC TGTGAGGATG
2201   GCTTCATGCA CTGTACCATG AGTGGAGTCC CCGGAAGCTT GCTGCCTGAC
2251   GCTGTCCTCA GCAGTCCCCT GTCTCATCGC AGCAAAAGGA GCCTATCCTG
2301   TCGGCCCCCC ATGGTCAAGC TGGTGTGTCC CGCTGACAAC CTGCGGGCTG
2351   AAGGGCTCGA GTGTACCAAA ACGTGCCAGA ACTATGACCT GGAGTGCATG
2401   AGCATGGGCT GTGTCTCTGG CTGCCTCTGC CCCCCGGGCA TGGTCCGGCA
2451   TGAGAACAGA TGTGTGGCCC TGGAAAGGTG TCCCTGCTTC CATCAGGGCA
2501   AGGAGTATGC CCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC
2551   TGTCGGGACC GGAAGTGGAA CTGCACAGAC CATGTGTGTG ATGCCACGTG
2601   CTCCACGATC GGCATGGCCC ACTACCTCAC CTTCGACGGG CTCAAATACC
2651   TGTTCCCCGG GGAGTGCCAG TACGTTCTGG TGCAGGATTA CTGCGGCAGT
2701   AACCCTGGGA CCTTTCGGAT CCTAGTGGGG AATAAGGGAT GCAGCCACCC
2751   CTCAGTGAAA TGCAAGAAAC GGGTCACCAT CCTGGTGGAG GGAGGAGAGA
2801   TTGAGCTGTT TGACGGGGAG GTGAATGTGA AGAGGCCCAT GAAGGATGAG
2851   ACTCACTTTG AGGTGGTGGA GTCTGGCCGG TACATCATTC TGCTGCTGGG
2901   CAAAGCCCTC TCCGTGGTCT GGGACCGCCA CCTGAGCATC TCCGTGGTCC
2951   TGAAGCAGAC ATACCAGGAG AAAGTGTGTG GCCTGTGTGG GAATTTTGAT
3001   GGCATCCAGA ACAATGACCT CACCAGCAGC AACCTCCAAG TGGAGGAAGA
3051   CCCTGTGGAC TTTGGGAACT CCTGGAAAGT GAGCTCGCAG TGTGCTGACA
3101   CCAGAAAAGT GCCTCTGGAC TCATCCCCTG CCACCTGCCA TAACAACATC
3151   ATGAAGCAGA CGATGGTGGA TTCCTCCTGT AGAATCCTTA CCAGTGACGT
3201   CTTCCAGGAC TGCAACAAGC TGGTGGACCC CGAGCCATAT CTGGATGTCT
3251   GCATTTACGA CACCTGCTCC TGTGAGTCCA TTGGGGACTG CGCCGCATTC
3301   TGCGACACCA TTGCTGCCTA TGCCCACGTG TGTGCCCAGC ATGGCAAGGT
3351   GGTGACCTGG AGGACGGCCA CATTGTGCCC CCAGAGCTGC GAGGAGAGGA
3401   ATCTCCGGGA GAACGGGTAT GAGGCTGAGT GGCGCTATAA CAGCTGTGCA
3451   CCTGCCTGTC AAGTCACGTG TCAGCACCCT GAGCCACTGG CCTGCCCTGT
3501   GCAGTGTGTG GAGGGCTGCC ATGCCCACTG CCCTCCAGGG AAAATCCTGG
3551   ATGAGCTTTT GCAGACCTGC GTTGACCCTG AAGACTGTCC AGTGTGTGAG
```

[0928]

```
3601    GTGGCTGGCC GGCGTTTTGC CTCAGGAAAG AAAGTCACCT TGAATCCCAG
3651    TGACCCTGAG CACTGCCAGA TTTGCCACTG TGATGTTGTC AACCTCACCT
3701    GTGAAGCCTG CCAGGAGCCG ATATCTGGCG GTGGAGGTTC CGGTGGCGGG
3751    GGATCCGGCG GTGGAGGTTC CGGCGGTGGA GGTTCCGGTG GCGGGGGATC
3801    CGGTGGCGGG GGATCCTTAC CTGAAACTGG AGCCCTGCGG CCCCGGGTCG
3851    TCGGCGGTGG AGGTTCCGGT GGCGGGGGAT CCGACAAAAC TCACACATGC
3901    CCACCGTGCC CAGCTCCAGA ACTCCTGGGC GGACCGTCAG TCTTCCTCTT
3951    CCCCCCAAAA CCCAAGGACA CCCTCATGAT CTCCCGGACC CCTGAGGTCA
4001    CATGCGTGGT GGTGGACGTG AGCCACGAAG ACCCTGAGGT CAAGTTCAAC
4051    TGGTACGTGG ACGGCGTGGA GGTGCATAAT GCCAAGACAA AGCCGCGGGA
4101    GGAGCAGTAC AACAGCACGT ACCGTGTGGT CAGCGTCCTC ACCGTCCTGC
4151    ACCAGGACTG GCTGAATGGC AAGGAGTACA AGTGCAAGGT CTCCAACAAA
4201    GCCCTCCCAG CCCCCATCGA GAAAACCATC TCCAAAGCCA AAGGGCAGCC
4251    CCGAGAACCA CAGGTGTACA CCCTGCCCCC ATCCCGGGAT GAGCTGACCA
4301    AGAACCAGGT CAGCCTGACC TGCCTGGTCA AAGGCTTCTA TCCCAGCGAC
4351    ATCGCCGTGG AGTGGGAGAG CAATGGGCAG CCGGAGAACA ACTACAAGAC
4401    CACGCCTCCC GTGTTGGACT CCGACGGCTC CTTCTTCCTC TACAGCAAGC
4451    TCACCGTGGA CAAGAGCAGG TGGCAGCAGG GGAACGTCTT CTCATGCTCC
4501    GTGATGCATG AGGCTCTGCA CAACCACTAC ACGCAGAAGA GCCTCTCCCT
4551    GTCTCCGGGT AAATGA
```

[0929] 表16:VWF051(具有分选酶A识别基序和VWF片段和Fc之间的凝血酶可切割连接基的VWF D1D2D'D3Fc;分选酶A位点以粗体示出)的蛋白质序列(SEQ ID NO:101)

[0930]

```
   1    MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNTFDGSM
  51    YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE FFDIHLFVNG
 101    TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL
 151    SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS WALSSGEQWC
 201    ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC
 251    EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME
 301    YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC
 351    VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD
 401    NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG
 451    LHNSLVKLKH GAGVAMDGQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM
 501    DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG
 551    NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS
 601    PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL
 651    NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGCFCP PGLYMDERGD
 701    CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD
 751    AVLSSPLSHR SKRSLSCRPP MVKLVCPADN LRAEGLECTK TCQNYDLECM
 801    SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV
 851    CRDRKWNCTD HVCDATCSTI GMAHYLTFDG LKYLFPGECQ YVLVQDYCGS
 901    NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE
 951    THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD
1001    GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI
1051    MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS CESIGDCAAF
1101    CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY EAEWRYNSCA
1151    PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE
1201    VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP ISGGGGSGGG
1251    GSGGGGSGGG GSGGGGSGGG GS<b>LPETG</b>ALR PRVVGGGGSG GGGSDKTHTC
1301    PPCPAPELLG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN
1351    WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK
1401    ALPAPIEKTI SKAKGQPREP QVYTLPPSRD ELTKNQVSLT CLVKGFYPSD
1451    IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS
1501    VMHEALHNHY TQKSLSLSPG K*
```

[0931] 表17:FVIII 265(在N-末端具有五甘氨酸的FVIII单链分子)的核苷酸序列(SEQ ID NO:102)

```
          1    ATGCAAATAG AGCTCTCCAC CTGCTTCTTT CTGTGCCTTT TGCGATTCTG
         51    CTTTAGTGGA GGAGGAGGAG GAGCCACCAG AAGATACTAC CTGGGTGCAG
        101    TGGAACTGTC ATGGGACTAT ATGCAAAGTG ATCTCGGTGA GCTGCCTGTG
        151    GACGCAAGAT TTCCTCCTAG AGTGCCAAAA TCTTTTCCAT TCAACACCTC
        201    AGTCGTGTAC AAAAAGACTC TGTTTGTAGA ATTCACGGAT CACCTTTTCA
        251    ACATCGCTAA GCCAAGGCCA CCCTGGATGG GTCTGCTAGG TCCTACCATC
        301    CAGGCTGAGG TTTATGATAC AGTGGTCATT ACACTTAAGA ACATGGCTTC
        351    CCATCCTGTC AGTCTTCATG CTGTTGGTGT ATCCTACTGG AAAGCTTCTG
        401    AGGGAGCTGA ATATGATGAT CAGACCAGTC AAAGGGAGAA AGAAGATGAT
        451    AAAGTCTTCC CTGGTGGAAG CCATACATAT GTCTGGCAGG TCCTGAAAGA
        501    GAATGGTCCA ATGGCCTCTG ACCCACTGTG CCTTACCTAC TCATATCTTT
        551    CTCATGTGGA CCTGGTAAAA GACTTGAATT CAGGCCTCAT TGGAGCCCTA
        601    CTAGTATGTA GAGAAGGGAG TCTGGCCAAG GAAAAGACAC AGACCTTGCA
        651    CAAATTTATA CTACTTTTTG CTGTATTTGA TGAAGGGAAA AGTTGGCACT
        701    CAGAAACAAA GAACTCCTTG ATGCAGGATA GGGATGCTGC ATCTGCTCGG
        751    GCCTGGCCTA AAATGCACAC AGTCAATGGT TATGTAAACA GGTCTCTGCC
        801    AGGTCTGATT GGATGCCACA GGAAATCAGT CTATTGGCAT GTGATTGGAA
        851    TGGGCACCAC TCCTGAAGTG CACTCAATAT TCCTCGAAGG TCACACATTT
        901    CTTGTGAGGA ACCATCGCCA GGCGTCCTTG GAAATCTCGC CAATAACTTT
        951    CCTTACTGCT CAAACACTCT TGATGGACCT TGGACAGTTT CTACTGTTTT
       1001    GTCATATCTC TTCCCACCAA CATGATGGCA TGGAAGCTTA TGTCAAAGTA
       1051    GACAGCTGTC CAGAGGAACC CCAACTACGA ATGAAAAATA ATGAAGAAGC
       1101    GGAAGACTAT GATGATGATC TTACTGATTC TGAAATGGAT GTGGTCAGGT
       1151    TTGATGATGA CAACTCTCCT TCCTTTATCC AAATTCGCTC AGTTGCCAAG
       1201    AAGCATCCTA AAACTTGGGT ACATTACATT GCTGCTGAAG AGGAGGACTG
       1251    GGACTATGCT CCCTTAGTCC TCGCCCCCGA TGACAGAAGT TATAAAAGTC
       1301    AATATTTGAA CAATGGCCCT CAGCGGATTG GTAGGAAGTA CAAAAAAGTC
       1351    CGATTTATGG CATACACAGA TGAAACCTTT AAGACTCGTG AAGCTATTCA
[0932] 1401    GCATGAATCA GGAATCTTGG GACCTTTACT TTATGGGGAA GTTGGAGACA
       1451    CACTGTTGAT TATATTTAAG AATCAAGCAA GCAGACCATA TAACATCTAC
       1501    CCTCACGGAA TCACTGATGT CCGTCCTTTG TATTCAAGGA GATTACCAAA
       1551    AGGTGTAAAA CATTTGAAGG ATTTTCCAAT TCTGCCAGGA GAAATATTCA
       1601    AATATAAATG GACAGTGACT GTAGAAGATG GGCCAACTAA ATCAGATCCT
       1651    CGGTGCCTGA CCCGCTATTA CTCTAGTTTC GTTAATATGG AGAGAGATCT
       1701    AGCTTCAGGA CTCATTGGCC CTCTCCTCAT CTGCTACAAA GAATCTGTAG
       1751    ATCAAAGAGG AAACCAGATA ATGTCAGACA AGAGGAATGT CATCCTGTTT
       1801    TCTGTATTTG ATGAGAACCG AAGCTGGTAC CTCACAGAGA ATATACAACG
       1851    CTTTCTCCCC AATCCAGCTG GAGTGCAGCT TGAGGATCCA GAGTTCCAAG
       1901    CCTCCAACAT CATGCACAGC ATCAATGGCT ATGTTTTTGA TAGTTTGCAG
       1951    TTGTCAGTTT GTTTGCATGA GGTGGCATAC TGGTACATTC TAAGCATTGG
       2001    AGCACAGACT GACTTCCTTT CTGTCTTCTT CTCTGGATAT ACCTTCAAAC
       2051    ACAAAATGGT CTATGAAGAC ACACTCACCC TATTCCCATT CTCAGGAGAA
       2101    ACTGTCTTCA TGTCGATGGA AAACCCAGGT CTATGGATTC TGGGGTGCCA
       2151    CAACTCAGAC TTTCGGAACA GAGGCATGAC CGCCTTACTG AAGGTTTCTA
       2201    GTTGTGACAA GAACACTGGT GATTATTACG AGGACAGTTA TGAAGATATT
       2251    TCAGCATACT TGCTGAGTAA AAACAATGCC ATTGAACCAA GAAGCTTCTC
       2301    TCAAAACCCA CCAGTCTTGA AGGCCCATCA GGCCGAAATA ACTCGTACTA
       2351    CTCTTCAGTC AGATCAAGAG GAAATTGACT ATGATGATAC CATATCAGTT
       2401    GAAATGAAGA AGGAAGATTT TGACATTTAT GATGAGGATG AAAATCAGAG
       2451    CCCCCGCAGC TTTCAAAAGA AAACACGACA CTATTTTATT GCTGCAGTGG
       2501    AGAGGCTCTG GGATTATGGG ATGAGTAGCT CCCCACATGT TCTAAGAAAC
       2551    AGGGCTCAGA GTGGCAGTGT CCCTCAGTTC AAGAAAGTTG TTTTCCAGGA
       2601    ATTTACTGAT GGCTCCTTTA CTCAGCCCTT ATACCGTGGA GAACTAAATG
       2651    AACATTTGGG CCTCCTCGGC CCATATATAA GAGCAGAAGT TGAAGATAAT
       2701    ATCATGGTAA CTTTCAGAAA TCAGGCCTCT CGTCCCTATT CCTTCTATTC
       2751    TAGCCTTATT TCTTATGAGG AAGATCAGAG GCAAGGAGCA GAACCTAGAA
       2801    AAAACTTTGT CAAGCCTAAT GAAACCAAAA CTTACTTTTG GAAAGTGCAA
```

```
2851    CATCATATGG CACCCACTAA AGATGAGTTT GACTGCAAAG CCTGGGCTTA
2901    TTTCTCTGAT GTTGACCTGG AAAAAGATGT GCACTCAGGC CTGATTGGAC
2951    CCCTTCTGGT CTGCCACACT AACACACTGA ACCCTGCTCA TGGGAGACAA
3001    GTGACAGTAC AGGAATTTGC TCTGTTTTTC ACCATCTTTG ATGAGACCAA
3051    AAGCTGGTAC TTCACTGAAA ATATGGAAAG AAACTGCAGG GCTCCCTGCA
3101    ATATCCAGAT GGAAGATCCC ACTTTTAAAG AGAATTATCG CTTCCATGCA
3151    ATCAATGGCT ACATAATGGA TACACTACCT GGCTTAGTAA TGGCTCAGGA
3201    TCAAAGGATT CGATGGTATC TGCTCAGCAT GGGCAGCAAT GAAAACATCC
3251    ATTCTATTCA TTTCAGTGGA CATGTGTTCA CTGTACGAAA AAAAGAGGAG
3301    TATAAAATGG CACTGTACAA TCTCTATCCA GGTGTTTTTG AGACAGTGGA
3351    AATGTTACCA TCCAAAGCTG GAATTTGGCG GGTGGAATGC CTTATTGGCG
3401    AGCATCTACA TGCTGGGATG AGCACACTTT TTCTGGTGTA CAGCAATAAG
3451    TGTCAGACTC CCCTGGGAAT GGCTTCTGGA CACATTAGAG ATTTTCAGAT
3501    TACAGCTTCA GGACAATATG GACAGTGGGC CCCAAAGCTG GCCAGACTTC
3551    ATTATTCCGG ATCAATCAAT GCCTGGAGCA CCAAGGAGCC CTTTTCTTGG
3601    ATCAAGGTGG ATCTGTTGGC ACCAATGATT ATTCACGGCA TCAAGACCCA
3651    GGGTGCCCGT CAGAAGTTCT CCAGCCTCTA CATCTCTCAG TTTATCATCA
3701    TGTATAGTCT TGATGGGAAG AAGTGGCAGA CTTATCGAGG AAATTCCACT
3751    GGAACCTTAA TGGTCTTCTT TGGCAATGTG GATTCATCTG GGATAAAACA
3801    CAATATTTTT AACCCTCCAA TTATTGCTCG ATACATCCGT TTGCACCCAA
3851    CTCATTATAG CATTCGCAGC ACTCTTCGCA TGGAGTTGAT GGGCTGTGAT
3901    TTAAATAGTT GCAGCATGCC ATTGGGAATG GAGAGTAAAG CAATATCAGA
3951    TGCACAGATT ACTGCTTCAT CCTACTTTAC CAATATGTTT GCCACCTGGT
4001    CTCCTTCAAA AGCTCGACTT CACCTCCAAG GGAGGAGTAA TGCCTGGAGA
4051    CCTCAGGTGA ATAATCCAAA AGAGTGGCTG CAAGTGGACT TCCAGAAGAC
4101    AATGAAAGTC ACAGGAGTAA CTACTCAGGG AGTAAAATCT CTGCTTACCA
4151    GCATGTATGT GAAGGAGTTC CTCATCTCCA GCAGTCAAGA TGGCCATCAG
4201    TGGACTCTCT TTTTTCAGAA TGGCAAAGTA AAGGTTTTTC AGGGAAATCA
4251    AGACTCCTTC ACACCTGTGG TGAACTCTCT AGACCCACCG TTACTGACTC
4301    GCTACCTTCG AATTCACCCC CAGAGTTGGG TGCACCAGAT TGCCCTGAGG
4351    ATGGAGGTTC TGGGCTGCGA GGCACAGGAC CTCTACTGA
```

[0933]

[0934] 表18:FVIII 265(在N-末端具有五甘氨酸的FVIII单链分子:五甘氨酸以粗体示出)的蛋白质序列(SEQ ID NO:103)

```
          1    MQIELSTCFF LCLLRFCFSG GGGGATRRYY LGAVELSWDY MQSDLGELPV
         51    DARFPPRVPK SFPFNTSVVY KKTLFVEFTD HLFNIAKPRP PWMGLLGPTI
        101    QAEVYDTVVI TLKNMASHPV SLHAVGVSYW KASEGAEYDD QTSQREKEDD
        151    KVFPGGSHTY VWQVLKENGP MASDPLCLTY SYLSHVDLVK DLNSGLIGAL
        201    LVCREGSLAK EKTQTLHKFI LLFAVFDEGK SWHSETKNSL MQDRDAASAR
        251    AWPKMHTVNG YVNRSLPGLI GCHRKSVYWH VIGMGTTPEV HSIFLEGHTF
        301    LVRNHRQASL EISPITFLTA QTLLMDLGQF LLFCHISSHQ HDGMEAYVKV
        351    DSCPEEPQLR MKNNEEAEDY DDDLTDSEMD VVRFDDDNSP SFIQIRSVAK
        401    KHPKTWVHYI AAEEEDWDYA PLVLAPDDRS YKSQYLNNGP QRIGRKYKKV
        451    RFMAYTDETF KTREAIQHES GILGPLLYGE VGDTLLIIFK NQASRPYNIY
        501    PHGITDVRPL YSRRLPKGVK HLKDFPILPG EIFKYKWTVT VEDGPTKSDP
        551    RCLTRYYSSF VNMERDLASG LIGPLLICYK ESVDQRGNQI MSDKRNVILF
        601    SVFDENRSWY LTENIQRFLP NPAGVQLEDP EFQASNIMHS INGYVFDSLQ
        651    LSVCLHEVAY WYILSIGAQT DFLSVFFSGY TFKHKMVYED TLTLFPFSGE
        701    TVFMSMENPG LWILGCHNSD FRNRGMTALL KVSSCDKNTG DYYEDSYEDI
[0935]  751    SAYLLSKNNA IEPRSFSQNP PVLKAHQAEI TRTTLQSDQE EIDYDDTISV
        801    EMKKEDFDIY DEDENQSPRS FQKKTRHYFI AAVERLWDYG MSSSPHVLRN
        851    RAQSGSVPQF KKVVFQEFTD GSFTQPLYRG ELNEHLGLLG PYIRAEVEDN
        901    IMVTFRNQAS RPYSFYSSLI SYEEDQRQGA EPRKNFVKPN ETKTYFWKVQ
        951    HHMAPTKDEF DCKAWAYFSD VDLEKDVHSG LIGPLLVCHT NTLNPAHGRQ
       1001    VTVQEFALFF TIFDETKSWY FTENMERNCR APCNIQMEDP TFKENYRFHA
       1051    INGYIMDTLP GLVMAQDQRI RWYLLSMGSN ENIHSIHFSG HVFTVRKKEE

       1101    YKMALYNLYP GVFETVEMLP SKAGIWRVEC LIGEHLHAGM STLFLVYSNK
       1151    CQTPLGMASG HIRDFQITAS GQYGQWAPKL ARLHYSGSIN AWSTKEPFSW
       1201    IKVDLLAPMI IHGIKTQGAR QKFSSLYISQ FIIMYSLDGK KWQTYRGNST
       1251    GTLMVFFGNV DSSGIKHNIF NPPIIARYIR LHPTHYSIRS TLRMELMGCD
       1301    LNSCSMPLGM ESKAISDAQI TASSYFTNMF ATWSPSKARL HLQGRSNAWR
       1351    PQVNNPKEWL QVDFQKTMKV TGVTTQGVKS LLTSMYVKEF LISSSQDGHQ
       1401    WTLFFQNGKV KVFQGNQDSF TPVVNSLDPP LLTRYLRIHP QSWVHQIALR
       1451    MEVLGCEAQD LY*
```

[0936] 实施例20:HemA和FVIII/VWF双敲除(DKO)血浆中FVIII198的血浆稳定性和PK

[0937] 将FVIII 198(它是包含部分B-结构域的单链FVIIIFc分子-226N6;其中226表示FVIII B-结构域的N-末端226个氨基酸,并且N6表示B-结构域中的六个N-糖基化位点)的血浆稳定性与FVIII/VWF双敲除(DKO)血浆中的单链FVIIIFc(FVIII 155/Fc)进行比较。FVIII155和FVIII198的示意图在图25中可见。

[0938] 对于稳定性测定,将5IU/ml FVIII 198或FVIIIFc蛋白质与小鼠或DKO血浆在37℃下温育。在不同时间点收集等分试样以用于通过FVIII显色测定的活性测量。每个时间点的活性测量两次,并且将平均活性绘制为时间的函数。在稳定性测定中,部分B-结构域的存在增加了单链FVIIIFc的稳定性(图26A)。

[0939] 还将FVIII 198(单链-B226N6)的半衰期还与DKO小鼠中的FVIII155(单链B-结构域缺失的FVIII)进行比较。FVIII 198的半衰期比FVIII155长至少约1.5倍(图26B)。这些实验暗示FVIII稳定性与其体内半衰期之间可能存在相关性。

[0940] FVIII198核苷酸序列(具有部分B-结构域的FVIIIFc,226N6)(SEQ ID NO:104)

[0941]

```
   1  ATGCAAATAG AGCTCTCCAC CTGCTTCTTT CTGTGCCTTT TGCGATTCTG
  51  CTTTAGTGCC ACCAGAAGAT ACTACCTGGG TGCAGTGGAA CTGTCATGGG
 101  ACTATATGCA AAGTGATCTC GGTGAGCTGC CTGTGGACGC AAGATTTCCT
 151  CCTAGAGTGC CAAAATCTTT TCCATTCAAC ACCTCAGTCG TGTACAAAAA
 201  GACTCTGTTT GTAGAATTCA CGGATCACCT TTTCAACATC GCTAAGCCAA
 251  GGCCACCCTG GATGGGTCTG CTAGGTCCTA CCATCCAGGC TGAGGTTTAT
 301  GATACAGTGG TCATTACACT TAAGAACATG GCTTCCCATC CTGTCAGTCT
 351  TCATGCTGTT GGTGTATCCT ACTGGAAAGC TTCTGAGGGA GCTGAATATG
 401  ATGATCAGAC CAGTCAAAGG GAGAAAGAAG ATGATAAAGT CTTCCCTGGT
 451  GGAAGCCATA CATATGTCTG GCAGGTCCTG AAAGAGAATG GTCCAATGGC
 501  CTCTGACCCA CTGTGCCTTA CCTACTCATA TCTTTCTCAT GTGGACCTGG
 551  TAAAAGACTT GAATTCAGGC CTCATTGGAG CCCTACTAGT ATGTAGAGAA
 601  GGGAGTCTGG CCAAGGAAAA GACACAGACC TTGCACAAAT TTATACTACT
 651  TTTTGCTGTA TTTGATGAAG GGAAAAGTTG GCACTCAGAA ACAAAGAACT
 701  CCTTGATGCA GGATAGGGAT GCTGCATCTG CTCGGGCCTG GCCTAAAATG
 751  CACACAGTCA ATGGTTATGT AAACAGGTCT CTGCCAGGTC TGATTGGATG
 801  CCACAGGAAA TCAGTCTATT GGCATGTGAT TGGAATGGGC ACCACTCCTG
```

```
      851    AAGTGCACTC AATATTCCTC GAAGGTCACA CATTTCTTGT GAGGAACCAT
      901    CGCCAGGCGT CCTTGGAAAT CTCGCCAATA ACTTTCCTTA CTGCTCAAAC
      951    ACTCTTGATG GACCTTGGAC AGTTTCTACT GTTTTGTCAT ATCTCTTCCC
     1001    ACCAACATGA TGGCATGGAA GCTTATGTCA AAGTAGACAG CTGTCCAGAG
     1051    GAACCCCAAC TACGAATGAA AAATAATGAA GAAGCGGAAG ACTATGATGA
     1101    TGATCTTACT GATTCTGAAA TGGATGTGGT CAGGTTTGAT GATGACAACT
     1151    CTCCTTCCTT TATCCAAATT CGCTCAGTTG CCAAGAAGCA TCCTAAAACT
     1201    TGGGTACATT ACATTGCTGC TGAAGAGGAG GACTGGGACT ATGCTCCCTT
     1251    AGTCCTCGCC CCCGATGACA GAAGTTATAA AAGTCAATAT TTGAACAATG
     1301    GCCCTCAGCG GATTGGTAGG AAGTACAAAA AAGTCCGATT TATGGCATAC
     1351    ACAGATGAAA CCTTTAAGAC TCGTGAAGCT ATTCAGCATG AATCAGGAAT
     1401    CTTGGGACCT TTACTTTATG GGGAAGTTGG AGACACACTG TTGATTATAT
     1451    TTAAGAATCA AGCAAGCAGA CCATATAACA TCTACCCTCA CGGAATCACT
     1501    GATGTCCGTC CTTTGTATTC AAGGAGATTA CCAAAAGGTG TAAAACATTT
     1551    GAAGGATTTT CCAATTCTGC CAGGAGAAAT ATTCAAATAT AAATGGACAG
     1601    TGACTGTAGA AGATGGGCCA ACTAAATCAG ATCCTCGGTG CCTGACCCGC
     1651    TATTACTCTA GTTTCGTTAA TATGGAGAGA GATCTAGCTT CAGGACTCAT
     1701    TGGCCCTCTC CTCATCTGCT ACAAAGAATC TGTAGATCAA AGAGGAAACC
     1751    AGATAATGTC AGACAAGAGG AATGTCATCC TGTTTTCTGT ATTTGATGAG
     1801    AACCGAAGCT GGTACCTCAC AGAGAATATA CAACGCTTTC TCCCCAATCC
     1851    AGCTGGAGTG CAGCTTGAGG ATCCAGAGTT CCAAGCCTCC AACATCATGC
     1901    ACAGCATCAA TGGCTATGTT TTTGATAGTT TGCAGTTGTC AGTTTGTTTG
     1951    CATGAGGTGG CATACTGGTA CATTCTAAGC ATTGGAGCAC AGACTGACTT
     2001    CCTTTCTGTC TTCTTCTCTG GATATACCTT CAAACACAAA ATGGTCTATG
     2051    AAGACACACT CACCCTATTC CCATTCTCAG GAGAAACTGT CTTCATGTCG
     2101    ATGGAAAACC CAGGTCTATG GATTCTGGGG TGCCACAACT CAGACTTTCG
     2151    GAACAGAGGC ATGACCGCCT TACTGAAGGT TTCTAGTTGT GACAAGAACA
     2201    CTGGTGATTA TTACGAGGAC AGTTATGAAG ATATTTCAGC ATACTTGCTG
     2251    AGTAAAAACA ATGCCATTGA ACCAAGAAGC TTCTCTCAGA ATTCAAGACA
     2301    CCCTAGCACT AGGCAAAAGC AATTTAATGC CACCACAATT CCAGAAAATG
     2351    ACATAGAGAA GACTGACCCT TGGTTTGCAC ACAGAACACC TATGCCTAAA
     2401    ATACAAAATG TCTCCTCTAG TGATTTGTTG ATGCTCTTGC GACAGAGTCC
     2451    TACTCCACAT GGGCTATCCT TATCTGATCT CCAAGAAGCC AAATATGAGA
[0942] 2501  CTTTTTCTGA TGATCCATCA CCTGGAGCAA TAGACAGTAA TAACAGCCTG
     2551    TCTGAAATGA CACACTTCAG GCCACAGCTC CATCACAGTG GGGACATGGT
     2601    ATTTACCCCT GAGTCAGGCC TCCAATTAAG ATTAAATGAG AAACTGGGGA
     2651    CAACTGCAGC AACAGAGTTG AAGAAACTTG ATTTCAAAGT TTCTAGTACA
     2701    TCAAATAATC TGATTTCAAC AATTCCATCA GACAATTTGG CAGCAGGTAC
     2751    TGATAATACA AGTTCCTTAG GACCCCCAAG TATGCCAGTT CATTATGATA
     2801    GTCAATTAGA TACCACTCTA TTTGGCAAAA AGTCATCTCC CCTTACTGAG
     2851    TCTGGTGGAC CTCTGAGCTT GAGTGAAGAA AATAATGATT CAAAGTTGTT
     2901    AGAATCAGGT TTAATGAATA GCCAAGAAAG TTCATGGGGA AAAAATGTAT
     2951    CGTCAGAAAT AACTCGTACT ACTCTTCAGT CAGATCAAGA GGAAATTGAC
     3001    TATGATGATA CCATATCAGT TGAAATGAAG AAGGAAGATT TTGACATTTA
     3051    TGATGAGGAT GAAAATCAGA GCCCCCGCAG CTTTCAAAAG AAAACACGAC
     3101    ACTATTTTAT TGCTGCAGTG GAGAGGCTCT GGGATTATGG GATGAGTAGC
     3151    TCCCCACATG TTCTAAGAAA CAGGGCTCAG AGTGGCAGTG TCCCTCAGTT
     3201    CAAGAAAGTT GTTTTCCAGG AATTTACTGA TGGCTCCTTT ACTCAGCCCT
     3251    TATACCGTGG AGAACTAAAT GAACATTTGG GACTCCTGGG GCCATATATA
     3301    AGAGCAGAAG TTGAAGATAA TATCATGGTA ACTTTCAGAA ATCAGGCCTC
     3351    TCGTCCCTAT TCCTTCTATT CTAGCCTTAT TTCTTATGAG GAAGATCAGA
     3401    GGCAAGGAGC AGAACCTAGA AAAAACTTTG TCAAGCCTAA TGAAACCAAA
     3451    ACTTACTTTT GGAAAGTGCA ACATCATATG GCACCCACTA AAGATGAGTT
     3501    TGACTGCAAA GCCTGGGCTT ATTTCTCTGA TGTTGACCTG GAAAAAGATG
     3551    TGCACTCAGG CCTGATTGGA CCCCTTCTGG TCTGCCACAC TAACACACTG
     3601    AACCCTGCTC ATGGGAGACA AGTGACAGTA CAGGAATTTG CTCTGTTTTT
     3651    CACCATCTTT GATGAGACCA AAAGCTGGTA CTTCACTGAA AATATGGAAA
     3701    GAAACTGCAG GGCTCCCTGC AATATCCAGA TGGAAGATCC CACTTTTAAA
     3751    GAGAATTATC GCTTCCATGC AATCAATGGC TACATAATGG ATACACTACC
     3801    TGGCTTAGTA ATGGCTCAGG ATCAAAGGAT TCGATGGTAT CTGCTCAGCA
     3851    TGGGCAGCAA TGAAAACATC CATTCTATTC ATTTCAGTGG ACATGTGTTC
     3901    ACTGTACGAA AAAAAGAGGA GTATAAAATG GCACTGTACA ATCTCTATCC
     3951    AGGTGTTTTT GAGACAGTGG AAATGTTACC ATCCAAAGCT GGAATTTGGC
     4001    GGGTGGAATG CCTTATTGGC GAGCATCTAC ATGCTGGGAT GAGCACACTT
     4051    TTTCTGGTGT ACAGCAATAA GTGTCAGACT CCCCTGGGAA TGGCTTCTGG
     4101    ACACATTAGA GATTTTCAGA TTACAGCTTC AGGACAATAT GGACAGTGGG
     4151    CCCCAAAGCT GGCCAGACTT CATTATTCCG GATCAATCAA TGCCTGGAGC
```

[0943]

```
4201    ACCAAGGAGC CCTTTTCTTG GATCAAGGTG GATCTGTTGG CACCAATGAT
4251    TATTCACGGC ATCAAGACCC AGGGTGCCCG TCAGAAGTTC TCCAGCCTCT
4301    ACATCTCTCA GTTTATCATC ATGTATAGTC TTGATGGGAA GAAGTGGCAG
4351    ACTTATCGAG GAAATTCCAC TGGAACCTTA ATGGTCTTCT TTGGCAATGT
4401    GGATTCATCT GGGATAAAAC ACAATATTTT TAACCCTCCA ATTATTGCTC
4451    GATACATCCG TTTGCACCCA ACTCATTATA GCATTCGCAG CACTCTTCGC
4501    ATGGAGTTGA TGGGCTGTGA TTTAAATAGT TGCAGCATGC CATTGGGAAT
4551    GGAGAGTAAA GCAATATCAG ATGCACAGAT TACTGCTTCA TCCTACTTTA
4601    CCAATATGTT TGCCACCTGG TCTCCTTCAA AAGCTCGACT TCACCTCCAA
4651    GGGAGGAGTA ATGCCTGGAG ACCTCAGGTG AATAATCCAA AAGAGTGGCT
4701    GCAAGTGGAC TTCCAGAAGA CAATGAAAGT CACAGGAGTA ACTACTCAGG
4751    GAGTAAAATC TCTGCTTACC AGCATGTATG TGAAGGAGTT CCTCATCTCC
4801    AGCAGTCAAG ATGGCCATCA GTGGACTCTC TTTTTTCAGA ATGGCAAAGT
4851    AAAGGTTTTT CAGGGAAATC AAGACTCCTT CACACCTGTG GTGAACTCTC
4901    TAGACCCACC GTTACTGACT CGCTACCTTC GAATTCACCC CCAGAGTTGG
4951    GTGCACCAGA TTGCCCTGAG GATGGAGGTT CTGGGCTGCG AGGCACAGGA
5001    CCTCTACGAC AAAACTCACA CATGCCCACC GTGCCCAGCT CCAGAACTCC
5051    TGGGCGGACC GTCAGTCTTC CTCTTCCCCC CAAAACCCAA GGACACCCTC
5101    ATGATCTCCC GGACCCCTGA GGTCACATGC GTGGTGGTGG ACGTGAGCCA
5151    CGAAGACCCT GAGGTCAAGT TCAACTGGTA CGTGGACGGC GTGGAGGTGC
5201    ATAATGCCAA GACAAAGCCG CGGGAGGAGC AGTACAACAG CACGTACCGT
5251    GTGGTCAGCG TCCTCACCGT CCTGCACCAG GACTGGCTGA ATGGCAAGGA
5301    GTACAAGTGC AAGGTCTCCA ACAAAGCCCT CCCAGCCCCC ATCGAGAAAA
5351    CCATCTCCAA AGCCAAAGGG CAGCCCCGAG AACCACAGGT GTACACCCTG
5401    CCCCCATCCC GGGATGAGCT GACCAAGAAC CAGGTCAGCC TGACCTGCCT
5451    GGTCAAAGGC TTCTATCCCA GCGACATCGC CGTGGAGTGG GAGAGCAATG
5501    GGCAGCCGGA GAACAACTAC AAGACCACGC CTCCCGTGTT GGACTCCGAC
5551    GGCTCCTTCT TCCTCTACAG CAAGCTCACC GTGGACAAGA GCAGGTGGCA
5601    GCAGGGGAAC GTCTTCTCAT GCTCCGTGAT GCATGAGGCT CTGCACAACC
5651    ACTACACGCA GAAGAGCCTC TCCCTGTCTC CGGGTAAATG A
```

[0944] FVIII 198蛋白质序列(SEQ ID NO:105)

```
   1    MQIELSTCFF LCLLRFCFSA TRRYYLGAVE LSWDYMQSDL GELPVDARFP
  51    PRVPKSFPFN TSVVYKKTLF VEFTDHLFNI AKPRPPWMGL LGPTIQAEVY
 101    DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR EKEDDKVFPG
 151    GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG LIGALLVCRE
 201    GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD AASARAWPKM
 251    HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL EGHTFLVRNH
 301    RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME AYVKVDSCPE
 351    EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI RSVAKKHPKT
 401    WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR KYKKVRFMAY
 451    TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR PYNIYPHGIT
 501    DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP TKSDPRCLTR
 551    YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR NVILFSVFDE
 601    NRSWYLTENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV FDSLQLSVCL
 651    HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVYEDTLTLF PFSGETVFMS
 701    MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED SYEDISAYLL
 751    SKNNAIEPRS FSQNSRHPST RQKQFNATTI PENDIEKTDP WFAHRTPMPK
 801    IQNVSSSDLL MLLRQSPTPH GLSLSDLQEA KYETFSDDPS PGAIDSNNSL
 851    SEMTHFRPQL HHSGDMVFTP ESGLQLRLNE KLGTTAATEL KKLDFKVSST
 901    SNNLISTIPS DNLAAGTDNT SSLGPPSMPV HYDSQLDTTL FGKKSSPLTE
[0945]
 951    SGGPLSLSEE NNDSKLLESG LMNSQESSWG KNVSSEITRT TLQSDQEEID
1001    YDDTISVEMK KEDFDIYDED ENQSPRSFQK KTRHYFIAAV ERLWDYGMSS
1051    SPHVLRNRAQ SGSVPQFKKV VFQEFTDGSF TQPLYRGELN EHLGLLGPYI
1101    RAEVEDNIMV TFRNQASRPY SFYSSLISYE EDQRQGAEPR KNFVKPNETK
1151    TYFWKVQHHM APTKDEFDCK AWAYFSDVDL EKDVHSGLIG PLLVCHTNTL
1201    NPAHGRQVTV QEFALFFTIF DETKSWYFTE NMERNCRAPC NIQMEDPTFK
1251    ENYRFHAING YIMDTLPGLV MAQDQRIRWY LLSMGSNENI HSIHFSGHVF
1301    TVRKKEEYKM ALYNLYPGVF ETVEMLPSKA GIWRVECLIG EHLHAGMSTL
1351    FLVYSNKCQT PLGMASGHIR DFQITASGQY GQWAPKLARL HYSGSINAWS
1401    TKEPFSWIKV DLLAPMIIHG IKTQGARQKF SSLYISQFII MYSLDGKKWQ
1451    TYRGNSTGTL MVFFGNVDSS GIKHNIFNPP IIARYIRLHP THYSIRSTLR
1501    MELMGCDLNS CSMPLGMESK AISDAQITAS SYFTNMFATW SPSKARLHLQ
1551    GRSNAWRPQV NNPKEWLQVD FQKTMKVTGV TTQGVKSLLT SMYVKEFLIS
1601    SSQDGHQWTL FFQNGKVKVF QGNQDSFTPV VNSLDPPLLT RYLRIHPQSW
1651    VHQIALRMEV LGCEAQDLYD KTHTCPPCPA PELLGGPSVF LFPPKPKDTL

1701    MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR
1751    VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL
1801    PPSRDELTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD
1851    GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGK*
```

[0946] 实施例21:VWF的D1D2蛋白的表达

[0947] D'D3结构域的正确折叠对其结合FVIII是必要的。VWF原肽(D1D2-第1-763位氨基酸)对于D'D3的二硫键有效形成和折叠是必要的。其作为D'D3折叠的内部分子伴侣。制备VWF构建体的VWF片段可在VWF原肽(即D1D2结构域)直接连接至D'D3结构域时表达,并且在D'D3的常规细胞内加工期间移除(即顺式),或它可从其它质粒表达(即反式)。我们以其中D1D2以顺式或反式表达的方式设计FVIII-VWF异源二聚体。

[0948] 克隆VWF 053:VWF 053克隆表达VWF原肽(D1D2结构域),从而以反式表达D1D2。VWF原肽使用ESC54和ESC124从全长进行PCR扩增。

[0949] 具有BsiW1位点的ESC54-VWF-正向引物(SEQ ID NO:111)

[0950] (CGCTTCGCGACGTACGGCCGCCACCATGATTCCTGCCAGATTTGCCGGGGTGCTGCTTGCTC)

[0951] 具有Not1位点的ESC 124-D1D2克隆寡核苷酸-反向引物(SEQ ID NO:112)

[0952] (CTAGACTCGAGCGGCCGCTCACCTTTTGCTGCGATGAGACAGGGGACTGCTGAGGACAGC)

[0953] PCR产物用BsiW1和Not1消化并且连接至BsiW1/Not1消化的pCDNA 4。

[0954] VWF 053(VWF D1D2-原肽)的核苷酸序列(SEQ ID NO:113)

[0955]

```
   1  ATGATTCCTG CCAGATTTGC CGGGGTGCTG CTTGCTCTGG CCCTCATTTT
  51  GCCAGGGACC CTTTGTGCAG AAGGAACTCG CGGCAGGTCA TCCACGGCCC
 101  GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG
 151  TACAGCTTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA
 201  ACGCTCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC
 251  TCTCCGTGTA TCTTGGGGAA TTTTTTGACA TCCATTTGTT TGTCAATGGT
 301  ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG
 351  GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCCT
 401  ATGGCTTTGT GGCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG
 451  TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTGGCA ACTTTAACAT
 501  CTTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC
 551  CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT
 601  GAACGGGCAT CTCCTCCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT
 651  GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT
 701  TTGCCCGCTG CCACCCTCTG GTGGACCCCG AGCCTTTTGT GGCCCTGTGT
 751  GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CCTGCCCTGC
 801  CCTCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAATG GTGCTGTACG
 851  GCTGGACCGA CCACAGCGCG TGCAGCCCAG TGTGCCCTGC TGGTATGGAG
 901  TATAGGCAGT GTGTGTCCCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT
 951  CAATGAAATG TGTCAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG
1001  GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGTCCCTGC
1051  GTGCATTCCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG
1101  CAACACCTGC ATTTGCCGAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT
1151  GTCCAGGGGA GTGCCTTGTC ACTGGTCAAT CCCACTTCAA GAGCTTTGAC
1201  AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGGA
1251  TTGCCAGGAC CACTCCTTCT CCATTGTCAT TGAGACTGTC CAGTGTGCTG
1301  ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGTCCG GCTGCCTGGC
1351  CTGCACAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA
1401  TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC
1451  ATACAGTGAC GGCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG
1501  GACTGGGATG GCCGCGGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC
1551  CGGGAAGACC TGCGGCCTGT GTGGGAATTA CAATGGCAAC CAGGGCGACG
1601  ACTTCCTTAC CCCCTCTGGG CTGGCGGAGC CCCGGGTGGA GGACTTCGGG
1651  AACGCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG
1701  CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGGTTCTCC GAGGAGGCGT
1751  GCGCGGTCCT GACGTCCCCC ACATTCGAGG CCTGCCATCG TGCCGTCAGC
1801  CCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACGTGTGCT CCTGCTCGGA
1851  CGGCCGCGAG TGCCTGTGCG GCGCCCTGGC CAGCTATGCC GCGGCCTGCG
1901  CGGGGAGAGG CGTGCGCGTC GCGTGGCGCG AGCCAGGCCG CTGTGAGCTG
1951  AACTGCCCGA AAGGCCAGGT GTACCTGCAG TGCGGGACCC CCTGCAACCT
2001  GACCTGCCGC TCTCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCTGCC
2051  TGGAGGGCTG CTTCTGCCCC CCAGGGCTCT ACATGGATGA GAGGGGGGAC
2101  TGCGTGCCCA AGGCCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA
2151  GCCAGAAGAC ATCTTCTCAG ACCATCACAC CATGTGCTAC TGTGAGGATG
2201  GCTTCATGCA CTGTACCATG AGTGGAGTCC CCGGAAGCTT GCTGCCTGAC
2251  GCTGTCCTCA GCAGTCCCCT GTCTCATCGC AGCAAAAGG
```

[0956]

[0957] VWF 053(VWF D1D2-原肽)的蛋白质序列(SEQ ID NO:114)

[0958]

```
  1    MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNTFDGSM
 51    YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE FFDIHLFVNG
101    TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL
151    SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS WALSSGEQWC
201    ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC
251    EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME
301    YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC
351    VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD
401    NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG
451    LHNSLVKLKH GAGVAMDGQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM
501    DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG
551    NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS
601    PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL
651    NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGCFCP PGLYMDERGD
701    CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD
751    AVLSSPLSHR SKR
```

[0959] 上述具体实施方案的描述完全揭示了本发明的一般特性，使得其他人可通过应用本领域技能范围内的知识，无需过度实验即可容易地修改和/或调整此类具体实施方案，以用于各种应用，而不脱离本发明的一般概念。因此，基于本文提供的教导和引导，此类调整和修改旨在处于所公开的实施方案的等同形式的意义和范围内。应当理解，本文的措辞或术语出于描述而不是限制目的，使得技术人员可根据教导和引导解释本说明书的术语或措辞。

[0960] 考虑到本文所公开的本发明的描述和实施，本发明的其它实施方案对于本领域的技术人员将显而易见。预期描述和实施例仅被视为示例性的，以下权利要求书将指明本发明的真实范围和精神。

[0961] 本文引用的所有专利和专利公布全文以引用方式并入本文。

序列表

<110> 比奥贝拉蒂治疗公司

<120> 嵌合因子VIII多肽及其用途

<130> 2159.359PC04/EKS/C-K/E-H

<140> 待指定

<141> 2013-01-11

<150> US 61/586,099

<151> 2012-01-12

<150> US 61/586,654

<151> 2012-01-13

<150> US 61/667,901

<151> 2012-07-03

<150> US 61/734,954

<151> 2012-12-07

<160> 114

<170> SIPOSequenceListing 1.0

<210> 1

<211> 16842

<212> DNA

<213> 智人()

<400> 1

atgattcctg ccagatttgc cggggtgctg cttgctctgg ccctcatttt gccagggacc 60

ctttgtgcag aaggaactcg cggcaggtca tccacggccc tactaaggac ggtctaaacg 120

gccccacgac gaacgagacc gggagtaaaa cggtccctgg gaaacacgtc ttccttgagc 180

gccgtccagt aggtgccggg gatgcagcct tttcggaagt gacttcgtca acacctttga 240

tgggagcatg tacagctttg cgggatactg cagttacctc ctggcagggg gctgccagaa 300

ctacgtcgga aaagccttca ctgaagcagt tgtggaaact accctcgtac atgtcgaaac 360

gccctatgac gtcaatggag gaccgtcccc cgacggtctt acgctccttc tcgattattg 420

gggacttcca gaatggcaag agagtgagcc tctccgtgta tcttggggaa ttttttgaca 480

tccatttgtt tgtcaatggt tgcgaggaag agctaataac ccctgaaggt cttaccgttc 540

tctcactcgg agaggcacat agaaccccct aaaaaactgt aggtaaacaa acagttacca 600

accgtgacac agggggacca aagagtctcc atgccctatg cctccaaagg gctgtatcta 660

gaaactgagg ctgggtacta caagctgtcc ggtgaggcct tggcactgtg tccccctggt 720

ttctcagagg tacgggatac ggaggtttcc cgacatagat ctttgactcc gacccatgat 780

gttcgacagg ccactccgga atggctttgt ggccaggatc gatggcagcg gcaactttca 840

agtcctgctg tcagacagat acttcaacaa gacctgcggg ctgtgtggca actttaacat 900

taccgaaaca ccggtcctag ctaccgtcgc cgttgaaagt tcaggacgac agtctgtcta 960

tgaagttgtt ctggacgccc gacacaccgt tgaaattgta ctttgctgaa gatgacttta 1020

tgacccaaga agggaccttg acctcggacc cttatgactt tgccaactca tgggctctga 1080
gcagtggaga acagtggtgt gaaacgactt ctactgaaat actgggttct tccctggaac 1140
tggagcctgg gaatactgaa acggttgagt acccgagact cgtcacctct tgtcaccaca 1200
gaacgggcat ctcctcccag cagctcatgc aacatctcct ctggggaaat gcagaagggc 1260
ctgtgggagc agtgccagct tctgaagagc acctcggtgt cttgcccgta gaggagggtc 1320
gtcgagtacg ttgtagagga gaccccttta cgtcttcccg gacaccctcg tcacggtcga 1380
agacttctcg tggagccaca ttgcccgctg ccaccctctg gtggaccccg agccttttgt 1440
ggccctgtgt gagaagactt tgtgtgagtg tgctgggggg ctggagtgcg cctgccctgc 1500
aacgggcgac ggtgggagac cacctggggc tcggaaaaca ccgggacaca ctcttctgaa 1560
acacactcac acgacccccc gacctcacgc ggacgggacg cctcctggag tacgcccgga 1620
cctgtgccca ggagggaatg gtgctgtacg gctggaccga ccacagcgcg tgcagcccag 1680
tgtgccctgc tggtatggag ggaggacctc atgcgggcct ggacacgggt cctcccttac 1740
cacgacatgc cgacctggct ggtgtcgcgc acgtcgggtc acacgggacg accatacctc 1800
tataggcagt gtgtgtcccc ttgcgccagg acctgccaga gcctgcacat caatgaaatg 1860
tgtcaggagc gatgcgtgga tggctgcagc tgccctgagg atatccgtca cacacagggg 1920
aacgcggtcc tggacggtct cggacgtgta gttactttac acagtcctcg ctacgcacct 1980
accgacgtcg acgggactcc gacagctcct ggatgaaggc ctctgcgtgg agagcaccga 2040
gtgtccctgc gtgcattccg gaaagcgcta ccctcccggc acctccctct ctcgagactg 2100
ctgtcgagga cctacttccg gagacgcacc tctcgtggct cacagggacg cacgtaaggc 2160
ctttcgcgat gggagggccg tggagggaga gagctctgac caacacctgc atttgccgaa 2220
acagccagtg gatctgcagc aatgaagaat gtccagggga gtgccttgtc actggtcaat 2280
cccacttcaa gagctttgac gttgtggacg taaacggctt tgtcggtcac ctagacgtcg 2340
ttacttctta caggtcccct cacggaacag tgaccagtta gggtgaagtt ctcgaaactg 2400
aacagatact tcaccttcag tgggatctgc cagtacctgc tggcccggga ttgccaggac 2460
cactccttct ccattgtcat tgagactgtc cagtgtgctg ttgtctatga agtggaagtc 2520
accctagacg gtcatggacg accgggccct aacggtcctg gtgaggaaga ggtaacagta 2580
actctgacag gtcacacgac atgaccgcga cgctgtgtgc acccgctccg tcaccgtccg 2640
gctgcctggc ctgcacaaca gccttgtgaa actgaagcat ggggcaggag ttgccatgga 2700
tactggcgct gcgacacacg tgggcgaggc agtggcaggc cgacggaccg gacgtgttgt 2760
cggaacactt tgacttcgta ccccgtcctc aacggtacct tggccaggac atccagctcc 2820
ccctcctgaa aggtgacctc cgcatccagc atacagtgac ggcctccgtg cgcctcagct 2880
acggggagga cctgcagatg accggtcctg taggtcgagg gggaggactt tccactggag 2940
gcgtaggtcg tatgtcactg ccggaggcac gcggagtcga tgcccctcct ggacgtctac 3000
gactgggatg gccgcgggag gctgctggtg aagctgtccc ccgtctatgc cgggaagacc 3060
tgcggcctgt gtgggaatta caatggcaac cagggcgacg ctgaccctac cggcgccctc 3120
cgacgaccac ttcgacaggg ggcagatacg gcccttctgg acgccggaca cacccttaat 3180
gttaccgttg gtcccgctgc acttccttac cccctctggg ctggcrgagc cccgggtgga 3240
ggacttcggg aacgcctgga agctgcacgg ggactgccag gacctgcaga agcagcacag 3300
tgaaggaatg ggggagaccc gaccgyctcg gggcccacct cctgaagccc ttgcggacct 3360

tcgacgtgcc cctgacggtc ctggacgtct tcgtcgtgtc cgatccctgc gccctcaacc 3420
cgcgcatgac caggttctcc gaggaggcgt gcgcggtcct gacgtccccc acattcgagg 3480
cctgccatcg tgccgtcagc gctagggacg cgggagttgg gcgcgtactg gtccaagagg 3540
ctcctccgca cgcgccagga ctgcaggggg tgtaagctcc ggacggtagc acggcagtcg 3600
ccgctgccct acctgcggaa ctgccgctac gacgtgtgct cctgctcgga cggccgcgag 3660
tgcctgtgcg gcgccctggc cagctatgcc gcggcctgcg ggcgacggga tggacgcctt 3720
gacggcgatg ctgcacacga ggacgagcct gccggcgctc acggacacgc cgcgggaccg 3780
gtcgatacgg cgccggacgc cggggagagg cgtgcgcgtc gcgtggcgcg agccaggccg 3840
ctgtgagctg aactgcccga aaggccaggt gtacctgcag tgcgggaccc cctgcaacct 3900
gcccctctcc gcacgcgcag cgcaccgcgc tcggtccggc gacactcgac ttgacgggct 3960
ttccggtcca catggacgtc acgccctggg ggacgttgga gacctgccgc tctctctctt 4020
acccggatga ggaatgcaat gaggcctgcc tggagggctg cttctgcccc ccagggctct 4080
acatggatga gaggggggac ctggacggcg agagagagaa tgggcctact ccttacgtta 4140
ctccggacgg acctcccgac gaagacgggg ggtcccgaga tgtacctact ctccccccctg 4200
tgcgtgccca aggcccagtg cccctgttac tatgacggtg agatcttcca gccagaagac 4260
atcttctcag accatcacac catgtgctac tgtgaggatg acgcacgggt tccgggtcac 4320
ggggacaatg atactgccac tctagaaggt cggtcttctg tagaagagtc tggtagtgtg 4380
gtacacgatg acactcctac gcttcatgca ctgtaccatg agtggagtcc ccggaagctt 4440
gctgcctgac gctgtcctca gcagtcccct gtctcatcgc agcaaaagga gcctatcctg 4500
cgaagtacgt gacatggtac tcacctcagg ggccttcgaa cgacggactg cgacaggagt 4560
cgtcagggga cagagtagcg tcgttttcct cggataggac tcggcccccc atggtcaagc 4620
tggtgtgtcc cgctgacaac ctgcgggctg aagggctcga gtgtaccaaa acgtgccaga 4680
actatgacct ggagtgcatg agccgggggg taccagttcg accacacagg gcgactgttg 4740
gacgcccgac ttcccgagct cacatggttt tgcacggtct tgatactgga cctcacgtac 4800
agcatgggct gtgtctctgg ctgcctctgc cccccgggca tggtccggca tgagaacaga 4860
tgtgtggccc tggaaaggtg tccctgcttc catcagggca tcgtacccga cacagagacc 4920
gacggagacg gggggcccgt accaggccgt actcttgtct acacaccggg acctttccac 4980
agggacgaag gtagtcccgt aggagtatgc ccctggagaa acagtgaaga ttggctgcaa 5040
cacttgtgtc tgtcgggacc ggaagtggaa ctgcacagac catgtgtgtg atgccacgtg 5100
tcctcatacg gggacctctt tgtcacttct aaccgacgtt gtgaacacag acagccctgg 5160
ccttcacctt gacgtgtctg gtacacacac tacggtgcac ctccacgatc ggcatggccc 5220
actacctcac cttcgacggg ctcaaatacc tgttccccgg ggagtgccag tacgttctgg 5280
tgcaggatta ctgcggcagt gaggtgctag ccgtaccggg tgatggagtg gaagctgccc 5340
gagtttatgg acaaggggcc cctcacggtc atgcaagacc acgtcctaat gacgccgtca 5400
aaccctggga cctttcggat cctagtgggg aataagggat gcagccaccc ctcagtgaaa 5460
tgcaagaaac gggtcaccat cctggtggag ggaggagaga ttgggaccct ggaaagccta 5520
ggatcacccc ttattcccta cgtcggtggg gagtcacttt acgttctttg cccagtggta 5580
ggaccacctc cctcctctct ttgagctgtt tgacggggag gtgaatgtga agaggcccat 5640
gaaggatgag actcactttg aggtggtgga gtctggccgg tacatcattc tgctgctggg 5700

aactcgacaa actgcccctc cacttacact tctccgggta cttcctactc tgagtgaaac 5760
tccaccacct cagaccggcc atgtagtaag acgacgaccc caaagccctc tccgtggtct 5820
gggaccgcca cctgagcatc tccgtggtcc tgaagcagac ataccaggag aaagtgtgtg 5880
gcctgtgtgg gaattttgat gtttcgggag aggcaccaga ccctggcggt ggactcgtag 5940
aggcaccagg acttcgtctg tatggtcctc tttcacacac cggacacacc cttaaaacta 6000
ggcatccaga acaatgacct caccagcagc aacctccaag tggaggaaga ccctgtggac 6060
tttgggaact cctggaaagt gagctcgcag tgtgctgaca ccgtaggtct tgttactgga 6120
gtggtcgtcg ttggaggttc acctccttct gggacacctg aaacccttga ggacctttca 6180
ctcgagcgtc acacgactgt ccagaaaagt gcctctggac tcatcccctg ccacctgcca 6240
taacaacatc atgaagcaga cgatggtgga ttcctcctgt agaatcctta ccagtgacgt 6300
ggtcttttca cggagacctg agtaggggac ggtggacggt attgttgtag tacttcgtct 6360
gctaccacct aaggaggaca tcttaggaat ggtcactgca cttccaggac tgcaacaagc 6420
tggtggaccc cgagccatat ctggatgtct gcatttacga cacctgctcc tgtgagtcca 6480
ttggggactg cgcctgcttc gaaggtcctg acgttgttcg accacctggg gctcggtata 6540
gacctacaga cgtaaatgct gtggacgagg acactcaggt aacccctgac gcggacgaag 6600
tgcgacacca ttgctgccta tgcccacgtg tgtgcccagc atggcaaggt ggtgacctgg 6660
aggacggcca cattgtgccc ccagagctgc gaggagagga acgctgtggt aacgacggat 6720
acgggtgcac acacgggtcg taccgttcca ccactggacc tcctgccggt gtaacacggg 6780
ggtctcgacg ctcctctcct atctccggga gaacgggtat gagtgtgagt ggcgctataa 6840
cagctgtgca cctgcctgtc aagtcacgtg tcagcaccct gagccactgg cctgccctgt 6900
tagaggccct cttgcccata ctcacactca ccgcgatatt gtcgacacgt ggacggacag 6960
ttcagtgcac agtcgtggga ctcggtgacc ggacgggaca gcagtgtgtg gagggctgcc 7020
atgcccactg ccctccaggg aaaatcctgg atgagctttt gcagacctgc gttgaccctg 7080
aagactgtcc agtgtgtgag cgtcacacac ctcccgacgg tacgggtgac gggaggtccc 7140
ttttaggacc tactcgaaaa cgtctggacg caactgggac ttctgacagg tcacacactc 7200
gtggctggcc ggcgttttgc ctcaggaaag aaagtcacct tgaatcccag tgaccctgag 7260
cactgccaga tttgccactg tgatgttgtc aacctcacct caccgaccgg ccgcaaaacg 7320
gagtcctttc tttcagtgga acttagggtc actgggactc gtgacggtct aaacggtgac 7380
actacaacag ttggagtgga gtgaagcctg ccaggagccg ggaggcctgg tggtgcctcc 7440
cacagatgcc ccggtgagcc ccaccactct gtatgtggag gacatctcgg aaccgccgtt 7500
cacttcggac ggtcctcggc cctccggacc accacggagg gtgtctacgg ggccactcgg 7560
ggtggtgaga catacacctc ctgtagagcc ttggcggcaa gcacgatttc tactgcagca 7620
ggctactgga cctggtcttc ctgctggatg gctcctccag gctgtccgag gctgagtttg 7680
aagtgctgaa ggcctttgtg cgtgctaaag atgacgtcgt ccgatgacct ggaccagaag 7740
gacgacctac cgaggaggtc cgacaggctc cgactcaaac ttcacgactt ccggaaacac 7800
gtggacatga tggagcggct gcgcatctcc cagaagtggg tccgcgtggc cgtggtggag 7860
taccacgacg gctcccacgc ctacatcggg ctcaaggacc cacctgtact acctcgccga 7920
cgcgtagagg gtcttcaccc aggcgcaccg gcaccacctc atggtgctgc cgagggtgcg 7980
gatgtagccc gagttcctgg ggaagcgacc gtcagagctg cggcgcattg ccagccaggt 8040

gaagtatgcg ggcagccagg tggcctccac cagcgaggtc ttgaaataca cactgttcca 8100
ccttcgctgg cagtctcgac gccgcgtaac ggtcggtcca cttcatacgc ccgtcggtcc 8160
accggaggtg gtcgctccag aactttatgt gtgacaaggt aatcttcagc aagatcgacc 8220
gccctgaagc ctcccgcatc gccctgctcc tgatggccag ccaggagccc caacggatgt 8280
cccggaactt tgtccgctac ttagaagtcg ttctagctgg cgggacttcg gagggcgtag 8340
cgggacgagg actaccggtc ggtcctcggg gttgcctaca gggccttgaa acaggcgatg 8400
gtccagggcc tgaagaagaa gaaggtcatt gtgatcccgg tgggcattgg gccccatgcc 8460
aacctcaagc agatccgcct catcgagaag caggcccctg caggtcccgg acttcttctt 8520
cttccagtaa cactagggcc acccgtaacc cggggtacgg ttggagttcg tctaggcgga 8580
gtagctcttc gtccggggac agaacaaggc cttcgtgctg agcagtgtgg atgagctgga 8640
gcagcaaagg gacgagatcg ttagctacct ctgtgacctt gcccctgaag cccctcctcc 8700
tcttgttccg gaagcacgac tcgtcacacc tactcgacct cgtcgtttcc ctgctctagc 8760
aatcgatgga gacactggaa cggggacttc ggggaggagg tactctgccc cccgacatgg 8820
cacaagtcac tgtgggcccg gggctcttgg gggtttcgac cctggggccc aagaggaact 8880
ccatggttct ggatgtggcg atgagacggg gggctgtacc gtgttcagtg acacccgggc 8940
cccgagaacc cccaaagctg ggaccccggg ttctccttga ggtaccaaga cctacaccgc 9000
ttcgtcctgg aaggatcgga caaaattggt gaagccgact tcaacaggag caaggagttc 9060
atggaggagg tgattcagcg gatggatgtg ggccaggaca aagcaggacc ttcctagcct 9120
gttttaacca cttcggctga agttgtcctc gttcctcaag tacctcctcc actaagtcgc 9180
ctacctacac ccggtcctgt gcatccacgt cacggtgctg cagtactcct acatggtgac 9240
cgtggagtac cccttcagcg aggcacagtc caaaggggac atcctgcagc gggtgcgaga 9300
cgtaggtgca gtgccacgac gtcatgagga tgtaccactg gcacctcatg gggaagtcgc 9360
tccgtgtcag gtttcccctg taggacgtcg cccacgctct gatccgctac cagggcggca 9420
acaggaccaa cactgggctg gccctgcggt acctctctga ccacagcttc ttggtcagcc 9480
agggtgaccg ggagcaggcg ctaggcgatg gtcccgccgt tgtcctggtt gtgacccgac 9540
cgggacgcca tggagagact ggtgtcgaag aaccagtcgg tcccactggc cctcgtccgc 9600
cccaacctgg tctacatggt caccggaaat cctgcctctg atgagatcaa gaggctgcct 9660
ggagacatcc aggtggtgcc cattggagtg ggccctaatg gggttggacc agatgtacca 9720
gtggccttta ggacggagac tactctagtt ctccgacgga cctctgtagg tccaccacgg 9780
gtaacctcac ccgggattac ccaacgtgca ggagctggag aggattggct ggcccaatgc 9840
ccctatcctc atccaggact ttgagacgct cccccgagag gctcctgacc tggtgctgca 9900
ggttgcacgt cctcgacctc tcctaaccga ccgggttacg gggataggag taggtcctga 9960
aactctgcga gggggctctc cgaggactgg accacgacgt gaggtgctgc tccggagagg 10020
ggctgcagat ccccaccctc tcccctgcac ctgactgcag ccagcccctg gacgtgatcc 10080
ttctcctgga tggctcctcc ctccacgacg aggcctctcc ccgacgtcta ggggtgggag 10140
aggggacgtg gactgacgtc ggtcggggac ctgcactagg aagaggacct accgaggagg 10200
agtttcccag cttcttattt tgatgaaatg aagagtttcg ccaaggcttt catttcaaaa 10260
gccaatatag ggcctcgtct cactcaggtg tcagtgctgc tcaaagggtc gaagaataaa 10320
actactttac ttctcaaagc ggttccgaaa gtaaagtttt cggttatatc ccggagcaga 10380

gtgagtccac agtcacgacg agtatggaag catcaccacc attgacgtgc catggaacgt 10440
ggtcccggag aaagcccatt tgctgagcct tgtggacgtc atgcagcggg agggaggccc 10500
tcataccttc gtagtggtgg taactgcacg gtaccttgca ccagggcctc tttcgggtaa 10560
acgactcgga acacctgcag tacgtcgccc tccctccggg cagccaaatc ggggatgcct 10620
tgggctttgc tgtgcgatac ttgacttcag aaatgcatgg tgccaggccg ggagcctcaa 10680
aggcggtggt catcctggtc gtcggtttag cccctacgga acccgaaacg acacgctatg 10740
aactgaagtc tttacgtacc acggtccggc cctcggagtt tccgccacca gtaggaccag 10800
acggacgtct ctgtggattc agtggatgca gcagctgatg ccgccaggtc caacagagtg 10860
acagtgttcc ctattggaat tggagatcgc tacgatgcag tgcctgcaga gacacctaag 10920
tcacctacgt cgtcgactac ggcggtccag gttgtctcac tgtcacaagg gataacctta 10980
acctctagcg atgctacgtc cccagctacg gatcttggca ggcccagcag gcgactccaa 11040
cgtggtgaag ctccagcgaa tcgaagacct ccctaccatg gtcaccttgg gcaattcctt 11100
gggtcgatgc ctagaaccgt ccgggtcgtc cgctgaggtt gcaccacttc gaggtcgctt 11160
agcttctgga gggatggtac cagtggaacc cgttaaggaa cctccacaaa ctgtgctctg 11220
gatttgttag gatttgcatg gatgaggatg ggaatgagaa gaggcccggg gacgtctgga 11280
ccttgccaga ccagtgccac ggaggtgttt gacacgagac ctaaacaatc ctaaacgtac 11340
ctactcctac ccttactctt ctccgggccc ctgcagacct ggaacggtct ggtcacggtg 11400
accgtgactt gccagccaga tggccagacc ttgctgaaga gtcatcgggt caactgtgac 11460
cgggggctga ggccttcgtg ccctaacagc cagtcccctg tggcactgaa cggtcggtct 11520
accggtctgg aacgacttct cagtagccca gttgacactg gcccccgact ccggaagcac 11580
gggattgtcg gtcaggggac ttaaagtgga agagacctgt ggctgccgct ggacctgccc 11640
ctgygtgtgc acaggcagct ccactcggca catcgtgacc tttgatgggc agaatttcaa 11700
aatttcacct tctctggaca ccgacggcga cctggacggg gacrcacacg tgtccgtcga 11760
ggtgagccgt gtagcactgg aaactacccg tcttaaagtt gctgactggc agctgttctt 11820
atgtcctatt tcaaaacaag gagcaggacc tggaggtgat tctccataat ggtgcctgca 11880
gccctggagc aaggcagggc cgactgaccg tcgacaagaa tacaggataa agttttgttc 11940
ctcgtcctgg acctccacta agaggtatta ccacggacgt cgggacctcg ttccgtcccg 12000
tgcatgaaat ccatcgaggt gaagcacagt gccctctccg tcgagstgca cagtgacatg 12060
gaggtgacgg tgaatgggag actggtctct gttccttacg acgtacttta ggtagctcca 12120
cttcgtgtca cgggagaggc agctcsacgt gtcactgtac ctccactgcc acttaccctc 12180
tgaccagaga caaggaatgc tgggtgggaa catggaagtc aacgtttatg gtgccatcat 12240
gcatgaggtc agattcaatc accttggtca catcttcaca ttcactccac aaaacaatga 12300
acccaccctt gtaccttcag ttgcaaatac cacggtagta cgtactccag tctaagttag 12360
tggaaccagt gtagaagtgt aagtgaggtg ttttgttact gttccaactg cagctcagcc 12420
ccaagacttt tgcttcaaag acgtatggtc tgtgtgggat ctgtgatgag aacggagcca 12480
atgacttcat gctgagggat caaggttgac gtcgagtcgg ggttctgaaa acgaagtttc 12540
tgcataccag acacacccta gacactactc ttgcctcggt tactgaagta cgactcccta 12600
ggcacagtca ccacagactg gaaaacactt gttcaggaat ggactgtgca gcggccaggg 12660
cagacgtgcc agcccatcct ggaggagcag tgtcttgtcc ccgtgtcagt ggtgtctgac 12720

cttttgtgaa caagtcctta cctgacacgt cgccggtccc gtctgcacgg tcgggtagga 12780
cctcctcgtc acagaacagg ccgacagctc ccactgccag gtcctcctct taccactgtt 12840
tgctgaatgc cacaaggtcc tggctccagc cacattctat gccatctgcc agcaggacag 12900
ggctgtcgag ggtgacggtc caggaggaga atggtgacaa acgacttacg gtgttccagg 12960
accgaggtcg gtgtaagata cggtagacgg tcgtcctgtc ttgccaccag gagcaagtgt 13020
gtgaggtgat cgcctcttat gcccacctct gtcggaccaa cggggtctgc gttgactgga 13080
ggacacctga tttctgtgct aacggtggtc ctcgttcaca cactccacta gcggagaata 13140
cgggtggaga cagcctggtt gccccagacg caactgacct cctgtggact aaagacacga 13200
atgtcatgcc caccatctct ggtctacaac cactgtgagc atggctgtcc ccggcactgt 13260
gatggcaacg tgagctcctg tggggaccat ccctccgaag tacagtacgg gtggtagaga 13320
ccagatgttg gtgacactcg taccgacagg ggccgtgaca ctaccgttgc actcgaggac 13380
acccctggta gggaggcttc gctgtttctg ccctccagat aaagtcatgt tggaaggcag 13440
ctgtgtccct gaagaggcct gcactcagtg cattggtgag gatggagtcc agcaccagtt 13500
cgacaaagac gggaggtcta tttcagtaca accttccgtc gacacaggga cttctccgga 13560
cgtgagtcac gtaaccactc ctacctcagg tcgtggtcaa cctggaagcc tgggtcccgg 13620
accaccagcc ctgtcagatc tgcacatgcc tcagcgggcg gaaggtcaac tgcacaacgc 13680
agccctgccc cacggccaaa ggaccttcgg acccagggcc tggtggtcgg gacagtctag 13740
acgtgtacgg agtcgcccgc cttccagttg acgtgttgcg tcgggacggg gtgccggttt 13800
gctcccacgt gtggcctgtg tgaagtagcc cgcctccgcc agaatgcaga ccagtgctgc 13860
cccgagtatg agtgtgtgtg tgacccagtg agctgtgacc cgagggtgca caccggacac 13920
acttcatcgg gcggaggcgg tcttacgtct ggtcacgacg gggctcatac tcacacacac 13980
actgggtcac tcgacactgg tgcccccagt gcctcactgt gaacgtggcc tccagcccac 14040
actgaccaac cctggcgagt gcagacccaa cttcacctgc gcctgcagga aggaggagtg 14100
acgggggtca cggagtgaca cttgcaccgg aggtcgggtg tgactggttg ggaccgctca 14160
cgtctgggtt gaagtggacg cggacgtcct tcctcctcac caaaagagtg tccccaccct 14220
cctgcccccc gcaccgtttg cccacccttc ggaagaccca gtgctgtgat gagtatgagt 14280
gtgcctgcaa ctgtgtcaac gttttctcac aggggtggga ggacgggggg cgtggcaaac 14340
gggtgggaag ccttctgggt cacgacacta ctcatactca cacggacgtt gacacagttg 14400
tccacagtga gctgtcccct tgggtacttg gcctcaaccg ccaccaatga ctgtggctgt 14460
accacaacca cctgccttcc cgacaaggtg tgtgtccacc aggtgtcact cgacagggga 14520
acccatgaac cggagttggc ggtggttact gacaccgaca tggtgttggt ggacggaagg 14580
gctgttccac acacaggtgg gaagcaccat ctaccctgtg ggccagttct gggaggaggg 14640
ctgcgatgtg tgcacctgca ccgacatgga ggatgccgtg atgggcctcc gcgtggccca 14700
cttcgtggta gatgggacac ccggtcaaga ccctcctccc gacgctacac acgtggacgt 14760
ggctgtacct cctacggcac tacccggagg cgcaccgggt gtgctcccag aagccctgtg 14820
aggacagctg tcggtcgggc ttcacttacg ttctgcatga aggcgagtgc tgtggaaggt 14880
gcctgccatc tgcctgtgag cacgagggtc ttcgggacac tcctgtcgac agccagcccg 14940
aagtgaatgc aagacgtact tccgctcacg acaccttcca cggacggtag acggacactc 15000
gtggtgactg gctcaccgcg gggggactcc cagtcttcct ggaagagtgt cggctcccag 15060

tgggcctccc cggagaaccc ctgcctcatc aatgagtgtg caccactgac cgagtggcgc 15120
ccccctgagg gtcagaagga ccttctcaca gccgagggtc acccggaggg gcctcttggg 15180
gacggagtag ttactcacac tccgagtgaa ggaggaggtc tttatacaac aaaggaacgt 15240
ctcctgcccc cagctggagg tccctgtctg cccctcgggc tttcagctga gctgtaagac 15300
aggctcactt cctcctccag aaatatgttg tttccttgca gaggacgggg gtcgacctcc 15360
agggacagac ggggagcccg aaagtcgact cgacattctg ctcagcgtgc tgcccaagct 15420
gtcgctgtga gcgcatggag gcctgcatgc tcaatggcac tgtcattggg cccgggaaga 15480
ctgtgatgat cgatgtgtgc gagtcgcacg acgggttcga cagcgacact cgcgtacctc 15540
cggacgtacg agttaccgtg acagtaaccc gggcccttct gacactacta gctacacacg 15600
acgacctgcc gctgcatggt gcaggtgggg gtcatctctg gattcaagct ggagtgcagg 15660
aagaccacct gcaacccctg ccccctgggt tacaaggaag tgctggacgg cgacgtacca 15720
cgtccacccc cagtagagac ctaagttcga cctcacgtcc ttctggtgga cgttggggac 15780
gggggaccca atgttccttc aaaataacac aggtgaatgt tgtgggagat gtttgcctac 15840
ggcttgcacc attcagctaa gaggaggaca gatcatgaca ctgaagcgtg atgagacgct 15900
ttttattgtg tccacttaca acaccctcta caaacggatg ccgaacgtgg taagtcgatt 15960
ctcctcctgt ctagtactgt gacttcgcac tactctgcga ccaggatggc tgtgatactc 16020
acttctgcaa ggtcaatgag agaggagagt acttctggga gaagagggtc acaggctgcc 16080
cacccttga tgaacacaag ggtcctaccg acactatgag tgaagacgtt ccagttactc 16140
tctcctctca tgaagaccct cttctcccag tgtccgacgg gtgggaaact acttgtgttc 16200
tgtcttgctg agggaggtaa aattatgaaa attccaggca cctgctgtga cacatgtgag 16260
gagcctgagt gcaacgacat cactgccagg ctgcagtatg acagaacgac tccctccatt 16320
ttaatacttt taaggtccgt ggacgacact gtgtacactc ctcggactca cgttgctgta 16380
gtgacggtcc gacgtcatac tcaaggtggg aagctgtaag tctgaagtag aggtggatat 16440
ccactactgc cagggcaaat gtgccagcaa agccatgtac tccattgaca tcaacgatgt 16500
agttccaccc ttcgacattc agacttcatc tccacctata ggtgatgacg gtcccgttta 16560
cacggtcgtt tcggtacatg aggtaactgt agttgctaca gcaggaccag tgctcctgct 16620
gctctccgac acggacggag cccatgcagg tggccctgca ctgcaccaat ggctctgttg 16680
tgtaccatga ggttctcaat cgtcctggtc acgaggacga cgagaggctg tgcctgcctc 16740
gggtacgtcc accgggacgt gacgtggtta ccgagacaac acatggtact ccaagagtta 16800
gccatggagt gcaaatgctc ccccaggaag tgcagcaagt ga 16842

<210> 2
<211> 2813
<212> PRT
<213> 智人()
<220>
<221> misc_feature
<222> (1)..(22)
<223> VWF信号肽
<220>

<221> misc_feature
<222> (23)..(763)
<223> VWF D1D2区
<220>
<221> misc_feature
<222> (764)..(866)
<223> VWF D'结构域
<220>
<221> misc_feature
<222> (867)..(1240)
<223> VWF D3结构域
<220>
<221> misc_feature
<222> (1241)..(1479)
<223> VWF A1结构域
<220>
<221> misc_feature
<222> (2016)..(2016)
<223> Xaa可以是任何天然存在的氨基酸
<220>
<221> UNSURE
<222> (2016)..(2016)
<223> The 'Xaa' at location 2016 stands for Gln, Arg, Pro, or Leu.
<400> 2

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

115 120 125

Leu Ser Gly Glu Ala Tyr Gly Phe Val Ala Arg Ile Asp Gly Ser Gly

130 135 140

Asn Phe Gln Val Leu Leu Ser Asp Arg Tyr Phe Asn Lys Thr Cys Gly

145 150 155 160

Leu Cys Gly Asn Phe Asn Ile Phe Ala Glu Asp Asp Phe Met Thr Gln

165 170 175

Glu Gly Thr Leu Thr Ser Asp Pro Tyr Asp Phe Ala Asn Ser Trp Ala

180 185 190

Leu Ser Ser Gly Glu Gln Trp Cys Glu Arg Ala Ser Pro Pro Ser Ser

195 200 205

Ser Cys Asn Ile Ser Ser Gly Glu Met Gln Lys Gly Leu Trp Glu Gln

210 215 220

Cys Gln Leu Leu Lys Ser Thr Ser Val Phe Ala Arg Cys His Pro Leu

225 230 235 240

Val Asp Pro Glu Pro Phe Val Ala Leu Cys Glu Lys Thr Leu Cys Glu

245 250 255

Cys Ala Gly Gly Leu Glu Cys Ala Cys Pro Ala Leu Leu Glu Tyr Ala

260 265 270

Arg Thr Cys Ala Gln Glu Gly Met Val Leu Tyr Gly Trp Thr Asp His

275 280 285

Ser Ala Cys Ser Pro Val Cys Pro Ala Gly Met Glu Tyr Arg Gln Cys

290 295 300

Val Ser Pro Cys Ala Arg Thr Cys Gln Ser Leu His Ile Asn Glu Met

305 310 315 320

Cys Gln Glu Arg Cys Val Asp Gly Cys Ser Cys Pro Glu Gly Gln Leu

325 330 335

Leu Asp Glu Gly Leu Cys Val Glu Ser Thr Glu Cys Pro Cys Val His

340 345 350

Ser Gly Lys Arg Tyr Pro Pro Gly Thr Ser Leu Ser Arg Asp Cys Asn

355 360 365

Thr Cys Ile Cys Arg Asn Ser Gln Trp Ile Cys Ser Asn Glu Glu Cys

370 375 380

Pro Gly Glu Cys Leu Val Thr Gly Gln Ser His Phe Lys Ser Phe Asp

385 390 395 400

Asn Arg Tyr Phe Thr Phe Ser Gly Ile Cys Gln Tyr Leu Leu Ala Arg

405 410 415

Asp Cys Gln Asp His Ser Phe Ser Ile Val Ile Glu Thr Val Gln Cys

420 425 430

Ala Asp Asp Arg Asp Ala Val Cys Thr Arg Ser Val Thr Val Arg Leu
435 440 445
Pro Gly Leu His Asn Ser Leu Val Lys Leu Lys His Gly Ala Gly Val
450 455 460
Ala Met Asp Gly Gln Asp Ile Gln Leu Pro Leu Leu Lys Gly Asp Leu
465 470 475 480
Arg Ile Gln His Thr Val Thr Ala Ser Val Arg Leu Ser Tyr Gly Glu
485 490 495
Asp Leu Gln Met Asp Trp Asp Gly Arg Gly Arg Leu Leu Val Lys Leu
500 505 510
Ser Pro Val Tyr Ala Gly Lys Thr Cys Gly Leu Cys Gly Asn Tyr Asn
515 520 525
Gly Asn Gln Gly Asp Asp Phe Leu Thr Pro Ser Gly Leu Ala Glu Pro
530 535 540
Arg Val Glu Asp Phe Gly Asn Ala Trp Lys Leu His Gly Asp Cys Gln
545 550 555 560
Asp Leu Gln Lys Gln His Ser Asp Pro Cys Ala Leu Asn Pro Arg Met
565 570 575
Thr Arg Phe Ser Glu Glu Ala Cys Ala Val Leu Thr Ser Pro Thr Phe
580 585 590
Glu Ala Cys His Arg Ala Val Ser Pro Leu Pro Tyr Leu Arg Asn Cys
595 600 605
Arg Tyr Asp Val Cys Ser Cys Ser Asp Gly Arg Glu Cys Leu Cys Gly
610 615 620
Ala Leu Ala Ser Tyr Ala Ala Ala Cys Ala Gly Arg Gly Val Arg Val
625 630 635 640
Ala Trp Arg Glu Pro Gly Arg Cys Glu Leu Asn Cys Pro Lys Gly Gln
645 650 655
Val Tyr Leu Gln Cys Gly Thr Pro Cys Asn Leu Thr Cys Arg Ser Leu
660 665 670
Ser Tyr Pro Asp Glu Glu Cys Asn Glu Ala Cys Leu Glu Gly Cys Phe
675 680 685
Cys Pro Pro Gly Leu Tyr Met Asp Glu Arg Gly Asp Cys Val Pro Lys
690 695 700
Ala Gln Cys Pro Cys Tyr Tyr Asp Gly Glu Ile Phe Gln Pro Glu Asp
705 710 715 720
Ile Phe Ser Asp His His Thr Met Cys Tyr Cys Glu Asp Gly Phe Met
725 730 735
His Cys Thr Met Ser Gly Val Pro Gly Ser Leu Leu Pro Asp Ala Val

740 745 750

Leu Ser Ser Pro Leu Ser His Arg Ser Lys Arg Ser Leu Ser Cys Arg

755 760 765

Pro Pro Met Val Lys Leu Val Cys Pro Ala Asp Asn Leu Arg Ala Glu

770 775 780

Gly Leu Glu Cys Thr Lys Thr Cys Gln Asn Tyr Asp Leu Glu Cys Met

785 790 795 800

Ser Met Gly Cys Val Ser Gly Cys Leu Cys Pro Pro Gly Met Val Arg

805 810 815

His Glu Asn Arg Cys Val Ala Leu Glu Arg Cys Pro Cys Phe His Gln

820 825 830

Gly Lys Glu Tyr Ala Pro Gly Glu Thr Val Lys Ile Gly Cys Asn Thr

835 840 845

Cys Val Cys Arg Asp Arg Lys Trp Asn Cys Thr Asp His Val Cys Asp

850 855 860

Ala Thr Cys Ser Thr Ile Gly Met Ala His Tyr Leu Thr Phe Asp Gly

865 870 875 880

Leu Lys Tyr Leu Phe Pro Gly Glu Cys Gln Tyr Val Leu Val Gln Asp

885 890 895

Tyr Cys Gly Ser Asn Pro Gly Thr Phe Arg Ile Leu Val Gly Asn Lys

900 905 910

Gly Cys Ser His Pro Ser Val Lys Cys Lys Lys Arg Val Thr Ile Leu

915 920 925

Val Glu Gly Gly Glu Ile Glu Leu Phe Asp Gly Glu Val Asn Val Lys

930 935 940

Arg Pro Met Lys Asp Glu Thr His Phe Glu Val Val Glu Ser Gly Arg

945 950 955 960

Tyr Ile Ile Leu Leu Leu Gly Lys Ala Leu Ser Val Val Trp Asp Arg

965 970 975

His Leu Ser Ile Ser Val Val Leu Lys Gln Thr Tyr Gln Glu Lys Val

980 985 990

Cys Gly Leu Cys Gly Asn Phe Asp Gly Ile Gln Asn Asn Asp Leu Thr

995 1000 1005

Ser Ser Asn Leu Gln Val Glu Glu Asp Pro Val Asp Phe Gly Asn Ser

1010 1015 1020

Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro Leu Asp

1025 1030 1035 1040

Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln Thr Met Val

1045 1050 1055

Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe Gln Asp Cys Asn
1060 1065 1070
Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val Cys Ile Tyr Asp Thr
1075 1080 1085
Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala Cys Phe Cys Asp Thr Ile
1090 1095 1100
Ala Ala Tyr Ala His Val Cys Ala Gln His Gly Lys Val Val Thr Trp
1105 1110 1115 1120
Arg Thr Ala Thr Leu Cys Pro Gln Ser Cys Glu Glu Arg Asn Leu Arg
1125 1130 1135
Glu Asn Gly Tyr Glu Cys Glu Trp Arg Tyr Asn Ser Cys Ala Pro Ala
1140 1145 1150
Cys Gln Val Thr Cys Gln His Pro Glu Pro Leu Ala Cys Pro Val Gln
1155 1160 1165
Cys Val Glu Gly Cys His Ala His Cys Pro Pro Gly Lys Ile Leu Asp
1170 1175 1180
Glu Leu Leu Gln Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu
1185 1190 1195 1200
Val Ala Gly Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro
1205 1210 1215
Ser Asp Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn Leu
1220 1225 1230
Thr Cys Glu Ala Cys Gln Glu Pro Gly Gly Leu Val Val Pro Pro Thr
1235 1240 1245
Asp Ala Pro Val Ser Pro Thr Thr Leu Tyr Val Glu Asp Ile Ser Glu
1250 1255 1260
Pro Pro Leu His Asp Phe Tyr Cys Ser Arg Leu Leu Asp Leu Val Phe
1265 1270 1275 1280
Leu Leu Asp Gly Ser Ser Arg Leu Ser Glu Ala Glu Phe Glu Val Leu
1285 1290 1295
Lys Ala Phe Val Val Asp Met Met Glu Arg Leu Arg Ile Ser Gln Lys
1300 1305 1310
Trp Val Arg Val Ala Val Val Glu Tyr His Asp Gly Ser His Ala Tyr
1315 1320 1325
Ile Gly Leu Lys Asp Arg Lys Arg Pro Ser Glu Leu Arg Arg Ile Ala
1330 1335 1340
Ser Gln Val Lys Tyr Ala Gly Ser Gln Val Ala Ser Thr Ser Glu Val
1345 1350 1355 1360
Leu Lys Tyr Thr Leu Phe Gln Ile Phe Ser Lys Ile Asp Arg Pro Glu

1365 1370 1375

Ala Ser Arg Ile Ala Leu Leu Leu Met Ala Ser Gln Glu Pro Gln Arg
1380 1385 1390

Met Ser Arg Asn Phe Val Arg Tyr Val Gln Gly Leu Lys Lys Lys Lys
1395 1400 1405

Val Ile Val Ile Pro Val Gly Ile Gly Pro His Ala Asn Leu Lys Gln
1410 1415 1420

Ile Arg Leu Ile Glu Lys Gln Ala Pro Glu Asn Lys Ala Phe Val Leu
1425 1430 1435 1440

Ser Ser Val Asp Glu Leu Glu Gln Gln Arg Asp Glu Ile Val Ser Tyr
1445 1450 1455

Leu Cys Asp Leu Ala Pro Glu Ala Pro Pro Pro Thr Leu Pro Pro Asp
1460 1465 1470

Met Ala Gln Val Thr Val Gly Pro Gly Leu Leu Gly Val Ser Thr Leu
1475 1480 1485

Gly Pro Lys Arg Asn Ser Met Val Leu Asp Val Ala Phe Val Leu Glu
1490 1495 1500

Gly Ser Asp Lys Ile Gly Glu Ala Asp Phe Asn Arg Ser Lys Glu Phe
1505 1510 1515 1520

Met Glu Glu Val Ile Gln Arg Met Asp Val Gly Gln Asp Ser Ile His
1525 1530 1535

Val Thr Val Leu Gln Tyr Ser Tyr Met Val Thr Val Glu Tyr Pro Phe
1540 1545 1550

Ser Glu Ala Gln Ser Lys Gly Asp Ile Leu Gln Arg Val Arg Glu Ile
1555 1560 1565

Arg Tyr Gln Gly Gly Asn Arg Thr Asn Thr Gly Leu Ala Leu Arg Tyr
1570 1575 1580

Leu Ser Asp His Ser Phe Leu Val Ser Gln Gly Asp Arg Glu Gln Ala
1585 1590 1595 1600

Pro Asn Leu Val Tyr Met Val Thr Gly Asn Pro Ala Ser Asp Glu Ile
1605 1610 1615

Lys Arg Leu Pro Gly Asp Ile Gln Val Val Pro Ile Gly Val Gly Pro
1620 1625 1630

Asn Ala Asn Val Gln Glu Leu Glu Arg Ile Gly Trp Pro Asn Ala Pro
1635 1640 1645

Ile Leu Ile Gln Asp Phe Glu Thr Leu Pro Arg Glu Ala Pro Asp Leu
1650 1655 1660

Val Leu Gln Arg Cys Cys Ser Gly Glu Gly Leu Gln Ile Pro Thr Leu
1665 1670 1675 1680

Ser Pro Ala Pro Asp Cys Ser Gln Pro Leu Asp Val Ile Leu Leu Leu
1685 1690 1695
Asp Gly Ser Ser Ser Phe Pro Ala Ser Tyr Phe Asp Glu Met Lys Ser
1700 1705 1710
Phe Ala Lys Ala Phe Ile Ser Lys Ala Asn Ile Gly Pro Arg Leu Thr
1715 1720 1725
Gln Val Ser Val Leu Gln Tyr Gly Ser Ile Thr Thr Ile Asp Val Pro
1730 1735 1740
Trp Asn Val Val Pro Glu Lys Ala His Leu Leu Ser Leu Val Asp Val
1745 1750 1755 1760
Met Gln Arg Glu Gly Gly Pro Ser Gln Ile Gly Asp Ala Leu Gly Phe
1765 1770 1775
Ala Val Arg Tyr Leu Thr Ser Glu Met His Gly Ala Arg Pro Gly Ala
1780 1785 1790
Ser Lys Ala Val Val Ile Leu Val Thr Asp Val Ser Val Asp Ser Val
1795 1800 1805
Asp Ala Ala Ala Asp Ala Ala Arg Ser Asn Arg Val Thr Val Phe Pro
1810 1815 1820
Ile Gly Ile Gly Asp Arg Tyr Asp Ala Ala Gln Leu Arg Ile Leu Ala
1825 1830 1835 1840
Gly Pro Ala Gly Asp Ser Asn Val Val Lys Leu Gln Arg Ile Glu Asp
1845 1850 1855
Leu Pro Thr Met Val Thr Leu Gly Asn Ser Phe Leu His Lys Leu Cys
1860 1865 1870
Ser Gly Phe Val Arg Ile Cys Met Asp Glu Asp Gly Asn Glu Lys Arg
1875 1880 1885
Pro Gly Asp Val Trp Thr Leu Pro Asp Gln Cys His Thr Val Thr Cys
1890 1895 1900
Gln Pro Asp Gly Gln Thr Leu Leu Lys Ser His Arg Val Asn Cys Asp
1905 1910 1915 1920
Arg Gly Leu Arg Pro Ser Cys Pro Asn Ser Gln Ser Pro Val Lys Val
1925 1930 1935
Glu Glu Thr Cys Gly Cys Arg Trp Thr Cys Pro Cys Val Cys Thr Gly
1940 1945 1950
Ser Ser Thr Arg His Ile Val Thr Phe Asp Gly Gln Asn Phe Lys Leu
1955 1960 1965
Thr Gly Ser Cys Ser Tyr Val Leu Phe Gln Asn Lys Glu Gln Asp Leu
1970 1975 1980
Glu Val Ile Leu His Asn Gly Ala Cys Ser Pro Gly Ala Arg Gln Gly

1985 1990 1995 2000

Cys Met Lys Ser Ile Glu Val Lys His Ser Ala Leu Ser Val Glu Xaa
2005 2010 2015

His Ser Asp Met Glu Val Thr Val Asn Gly Arg Leu Val Ser Val Pro
2020 2025 2030

Tyr Val Gly Gly Asn Met Glu Val Asn Val Tyr Gly Ala Ile Met His
2035 2040 2045

Glu Val Arg Phe Asn His Leu Gly His Ile Phe Thr Phe Thr Pro Gln
2050 2055 2060

Asn Asn Glu Phe Gln Leu Gln Leu Ser Pro Lys Thr Phe Ala Ser Lys
2065 2070 2075 2080

Thr Tyr Gly Leu Cys Gly Ile Cys Asp Glu Asn Gly Ala Asn Asp Phe
2085 2090 2095

Met Leu Arg Asp Gly Thr Val Thr Thr Asp Trp Lys Thr Leu Val Gln
2100 2105 2110

Glu Trp Thr Val Gln Arg Pro Gly Gln Thr Cys Gln Pro Ile Leu Glu
2115 2120 2125

Glu Gln Cys Leu Val Pro Asp Ser Ser His Cys Gln Val Leu Leu Leu
2130 2135 2140

Pro Leu Phe Ala Glu Cys His Lys Val Leu Ala Pro Ala Thr Phe Tyr
2145 2150 2155 2160

Ala Ile Cys Gln Gln Asp Ser Cys His Gln Glu Gln Val Cys Glu Val
2165 2170 2175

Ile Ala Ser Tyr Ala His Leu Cys Arg Thr Asn Gly Val Cys Val Asp
2180 2185 2190

Trp Arg Thr Pro Asp Phe Cys Ala Met Ser Cys Pro Pro Ser Leu Val
2195 2200 2205

Tyr Asn His Cys Glu His Gly Cys Pro Arg His Cys Asp Gly Asn Val
2210 2215 2220

Ser Ser Cys Gly Asp His Pro Ser Glu Gly Cys Phe Cys Pro Pro Asp
2225 2230 2235 2240

Lys Val Met Leu Glu Gly Ser Cys Val Pro Glu Glu Ala Cys Thr Gln
2245 2250 2255

Cys Ile Gly Glu Asp Gly Val Gln His Gln Phe Leu Glu Ala Trp Val
2260 2265 2270

Pro Asp His Gln Pro Cys Gln Ile Cys Thr Cys Leu Ser Gly Arg Lys
2275 2280 2285

Val Asn Cys Thr Thr Gln Pro Cys Pro Thr Ala Lys Ala Pro Thr Cys
2290 2295 2300

Gly Leu Cys Glu Val Ala Arg Leu Arg Gln Asn Ala Asp Gln Cys Cys
2305 2310 2315 2320
Pro Glu Tyr Glu Cys Val Cys Asp Pro Val Ser Cys Asp Leu Pro Pro
2325 2330 2335
Val Pro His Cys Glu Arg Gly Leu Gln Pro Thr Leu Thr Asn Pro Gly
2340 2345 2350
Glu Cys Arg Pro Asn Phe Thr Cys Ala Cys Arg Lys Glu Glu Cys Lys
2355 2360 2365
Arg Val Ser Pro Pro Ser Cys Pro Pro His Arg Leu Pro Thr Leu Arg
2370 2375 2380
Lys Thr Gln Cys Cys Asp Glu Tyr Glu Cys Ala Cys Asn Cys Val Asn
2385 2390 2395 2400
Ser Thr Val Ser Cys Pro Leu Gly Tyr Leu Ala Ser Thr Ala Thr Asn
2405 2410 2415
Asp Cys Gly Cys Thr Thr Thr Thr Cys Leu Pro Asp Lys Val Cys Val
2420 2425 2430
His Arg Ser Thr Ile Tyr Pro Val Gly Gln Phe Trp Glu Glu Gly Cys
2435 2440 2445
Asp Val Cys Thr Cys Thr Asp Met Glu Asp Ala Val Met Gly Leu Arg
2450 2455 2460
Val Ala Gln Cys Ser Gln Lys Pro Cys Glu Asp Ser Cys Arg Ser Gly
2465 2470 2475 2480
Phe Thr Tyr Val Leu His Glu Gly Glu Cys Cys Gly Arg Cys Leu Pro
2485 2490 2495
Ser Ala Cys Glu Val Val Thr Gly Ser Pro Arg Gly Asp Ser Gln Ser
2500 2505 2510
Ser Trp Lys Ser Val Gly Ser Gln Trp Ala Ser Pro Glu Asn Pro Cys
2515 2520 2525
Leu Ile Asn Glu Cys Val Arg Val Lys Glu Glu Val Phe Ile Gln Gln
2530 2535 2540
Arg Asn Val Ser Cys Pro Gln Leu Glu Val Pro Val Cys Pro Ser Gly
2545 2550 2555 2560
Phe Gln Leu Ser Cys Lys Thr Ser Ala Cys Cys Pro Ser Cys Arg Cys
2565 2570 2575
Glu Arg Met Glu Ala Cys Met Leu Asn Gly Thr Val Ile Gly Pro Gly
2580 2585 2590
Lys Thr Val Met Ile Asp Val Cys Thr Thr Cys Arg Cys Met Val Gln
2595 2600 2605
Val Gly Val Ile Ser Gly Phe Lys Leu Glu Cys Arg Lys Thr Thr Cys

2610 2615 2620

Asn Pro Cys Pro Leu Gly Tyr Lys Glu Glu Asn Asn Thr Gly Glu Cys
2625 2630 2635 2640

Cys Gly Arg Cys Leu Pro Thr Ala Cys Thr Ile Gln Leu Arg Gly Gly
2645 2650 2655

Gln Ile Met Thr Leu Lys Arg Asp Glu Thr Leu Gln Asp Gly Cys Asp
2660 2665 2670

Thr His Phe Cys Lys Val Asn Glu Arg Gly Glu Tyr Phe Trp Glu Lys
2675 2680 2685

Arg Val Thr Gly Cys Pro Pro Phe Asp Glu His Lys Cys Leu Ala Glu
2690 2695 2700

Gly Gly Lys Ile Met Lys Ile Pro Gly Thr Cys Cys Asp Thr Cys Glu
2705 2710 2715 2720

Glu Pro Glu Cys Asn Asp Ile Thr Ala Arg Leu Gln Tyr Val Lys Val
2725 2730 2735

Gly Ser Cys Lys Ser Glu Val Glu Val Asp Ile His Tyr Cys Gln Gly
2740 2745 2750

Lys Cys Ala Ser Lys Ala Met Tyr Ser Ile Asp Ile Asn Asp Val Gln
2755 2760 2765

Asp Gln Cys Ser Cys Cys Ser Pro Thr Arg Thr Glu Pro Met Gln Val
2770 2775 2780

Ala Leu His Cys Thr Asn Gly Ser Val Val Tyr His Glu Val Leu Asn
2785 2790 2795 2800

Ala Met Glu Cys Lys Cys Ser Pro Arg Lys Cys Ser Lys
2805 2810

<210> 3

<211> 11

<212> PRT

<213> 人工序列()

<220>

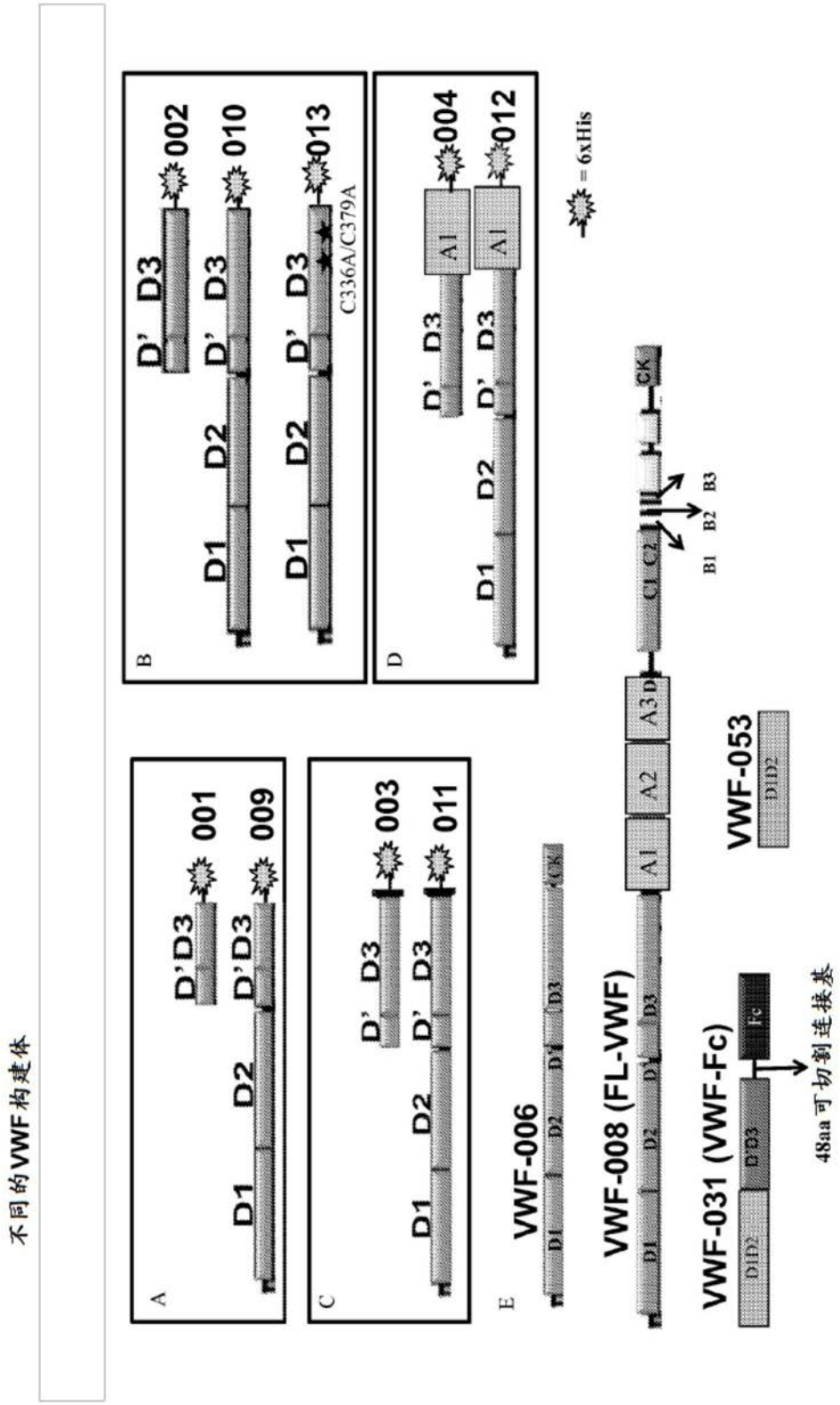

不同的VWF构建体
A
D'D3
001
D1
D2
D'D3
009
B
D' D3
002
D1
D2
D' D3
010
C336A/C379A
D1
D2
D' D3
013
C
D' D3
003
D1
D2
D' D3
011
D
D' D3
004
A1
D1
D2
D' D3
012
A1
= 6xHis
E
VWF-006
D1
D2
D'
D3
CK
VWF-008 (FL-VWF)
D1
D2
D'
D3
A1
A2
A3
D4
C1 C2
CK
B1 B2 B3
VWF-031 (VWF-Fc)
D1D2
D'D3
Fc
48aa可切割连接基
VWF-053
D1D2

图1

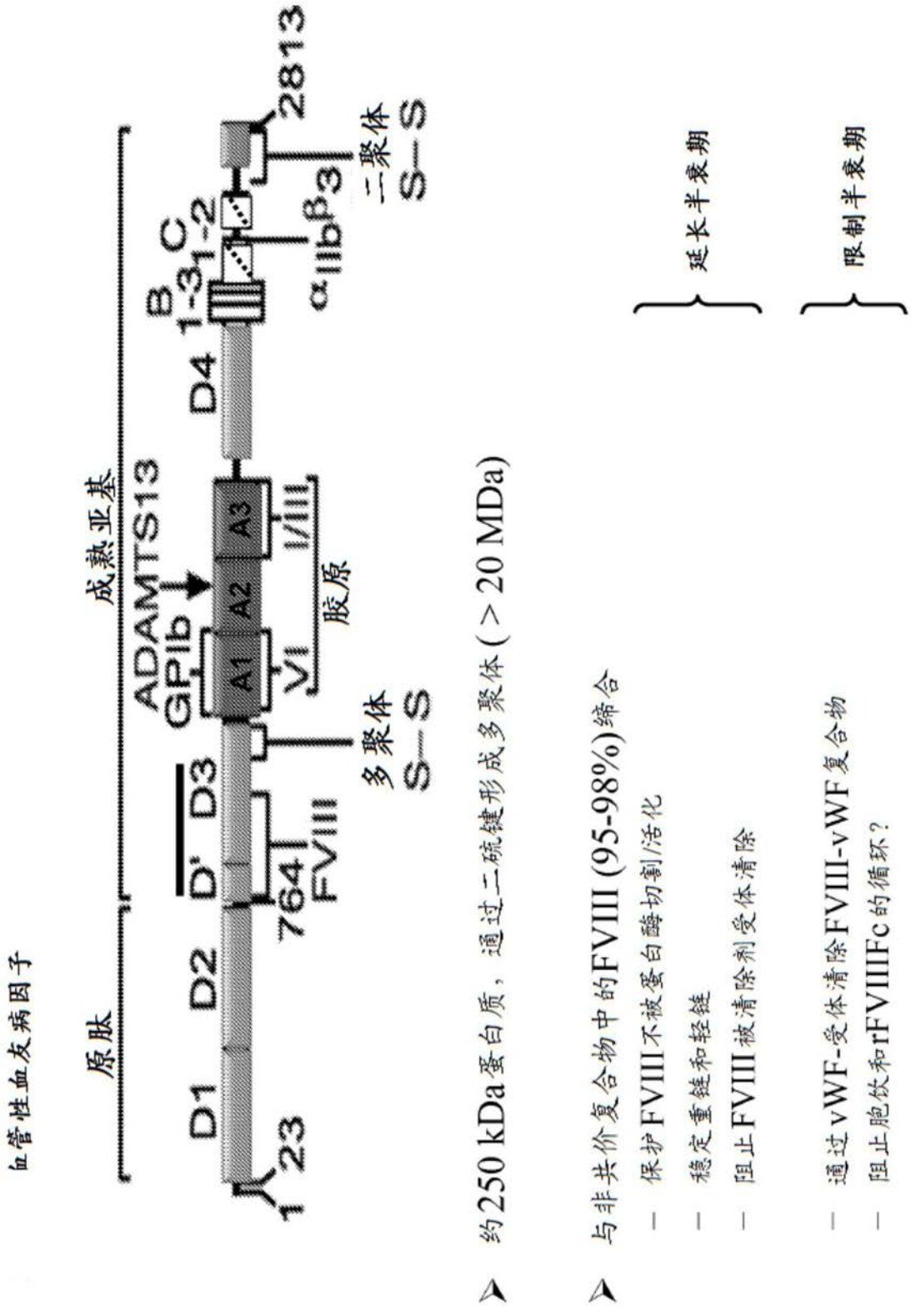

血管性血友病因子
原肽
成熟亚基
ADAMTS13
GPIb
D1
D2
D' D3
A1
A2
A3
D4
B
1-3
C
1-2
2813
1
23
764
FVIII
V1
VIII
胶原
多聚体
S—S
二聚体
S—S
αIIbβ3
约250 kDa蛋白质，通过二硫键形成多聚体（>20 MDa）
与非共价复合物中的FVIII (95-98%)缔合
— 保护FVIII不被蛋白酶切割/活化
— 稳定重链和轻链
— 阻止FVIII被清除剂受体清除
— 通过vWF-受体清除FVIII-vWF复合物
— 阻止胞饮和rFVIIIFc的循环？
延长半衰期
限制半衰期

图1F

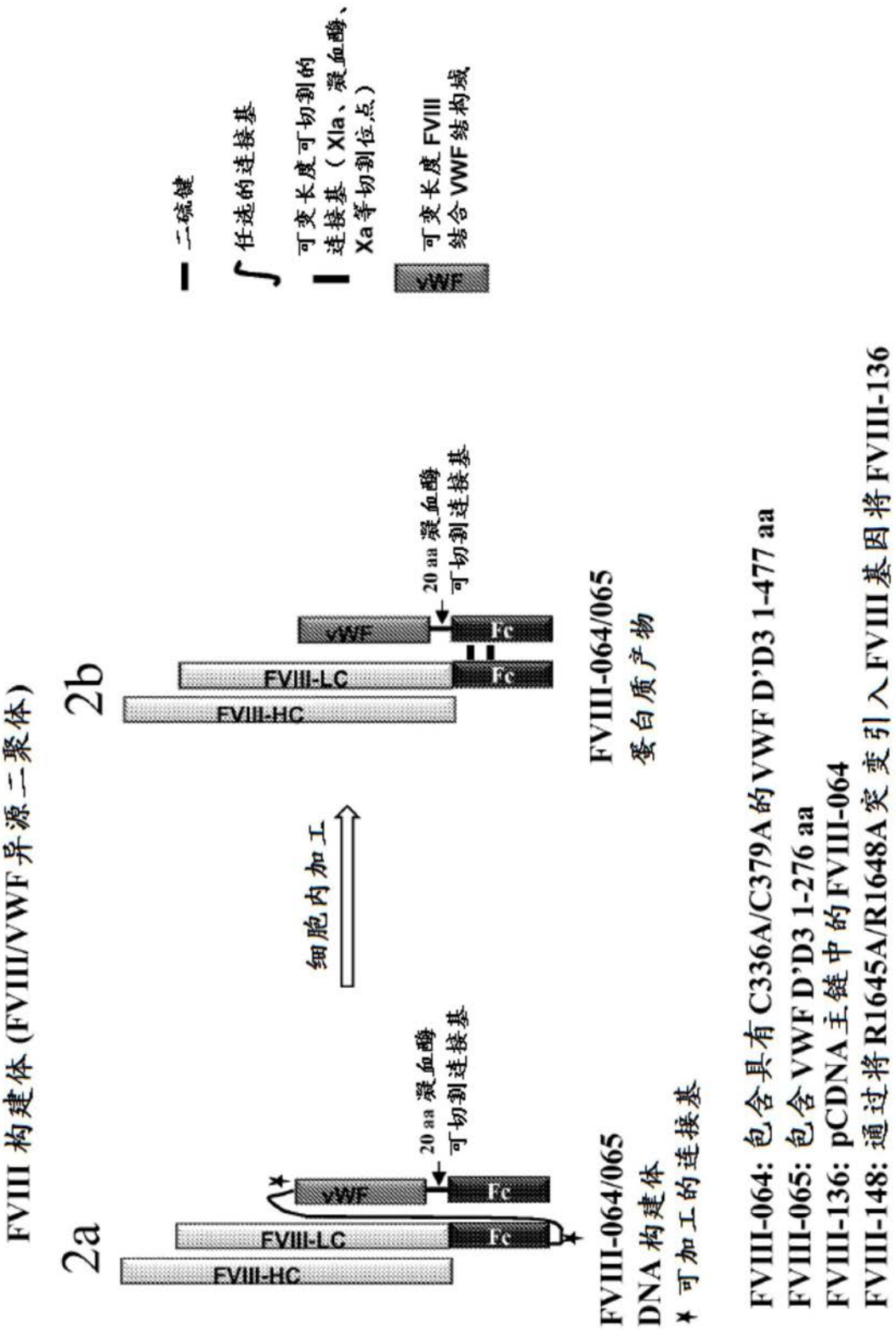

FVIII 构建体 (FVIII/VWF 异源二聚体)
2a
FVIII-HC
FVIII-LC
Fc
vWF
20 aa 凝血酶
可切割连接基
细胞内加工
2b
FVIII-HC
FVIII-LC
Fc
Fc
vWF
20 aa 凝血酶
可切割连接基
FVIII-064/065
蛋白质产物
FVIII-064/065
DNA 构建体
* 可加工的连接基
FVIII-064: 包含具有 C336A/C379A 的 VWF D'D3 1-477 aa
FVIII-065: 包含 VWF D'D3 1-276 aa
FVIII-136: pCDNA 主链中的 FVIII-064
FVIII-148: 通过将 R1645A/R1648A 突变引入 FVIII 基因将 FVIII-136
中的双链 FVIII 分子改为单链 FVIII
二硫键
任选的连接基
可变长度可切割的
连接基 (Xla、凝血酶、
Xa 等切割位点)
可变长度 FVIII
结合 VWF 结构域
vWF

图2

FVIII 构建体（可变连接基）

DNA 构建体	VWF 和 Fc 之间的连接基
FVIII-064	20 aa= ID {2X(GGGGS)}LVPRGSGG
FVIII-159	35 aa= IS{5X(GGGGS)}LVPRGSGG
FVIII-160 及其衍生物	FVIII-160: 48 aa= IS{6X(GGGGS)}LVPRGSGGGGSGGGGS
	FVIII-180: FVIII C1 结构域中具有 K2092A 突变的 FVIII-160
	FVIII-181: FVIII C1 结构域中具有 F2093A 突变的 FVIII-160
	FVIII-182: FVIII C1结构域中具有 K2092A/F2093A 突变的 FVIII-160
FVIII-178	73 aa= IS{11X(GGGGS)}LVPRGSGGGGSGGGGS
FVIII-179	98 aa= IS{16X(GGGGS)}LVPRGSGGGGSGGGGS
VWF= D'D3 (具有 C336A/C379A 的 1-477aa)	

FVIII-064, 159, 160, 178, 179

* 体内可加工的连接基

图3

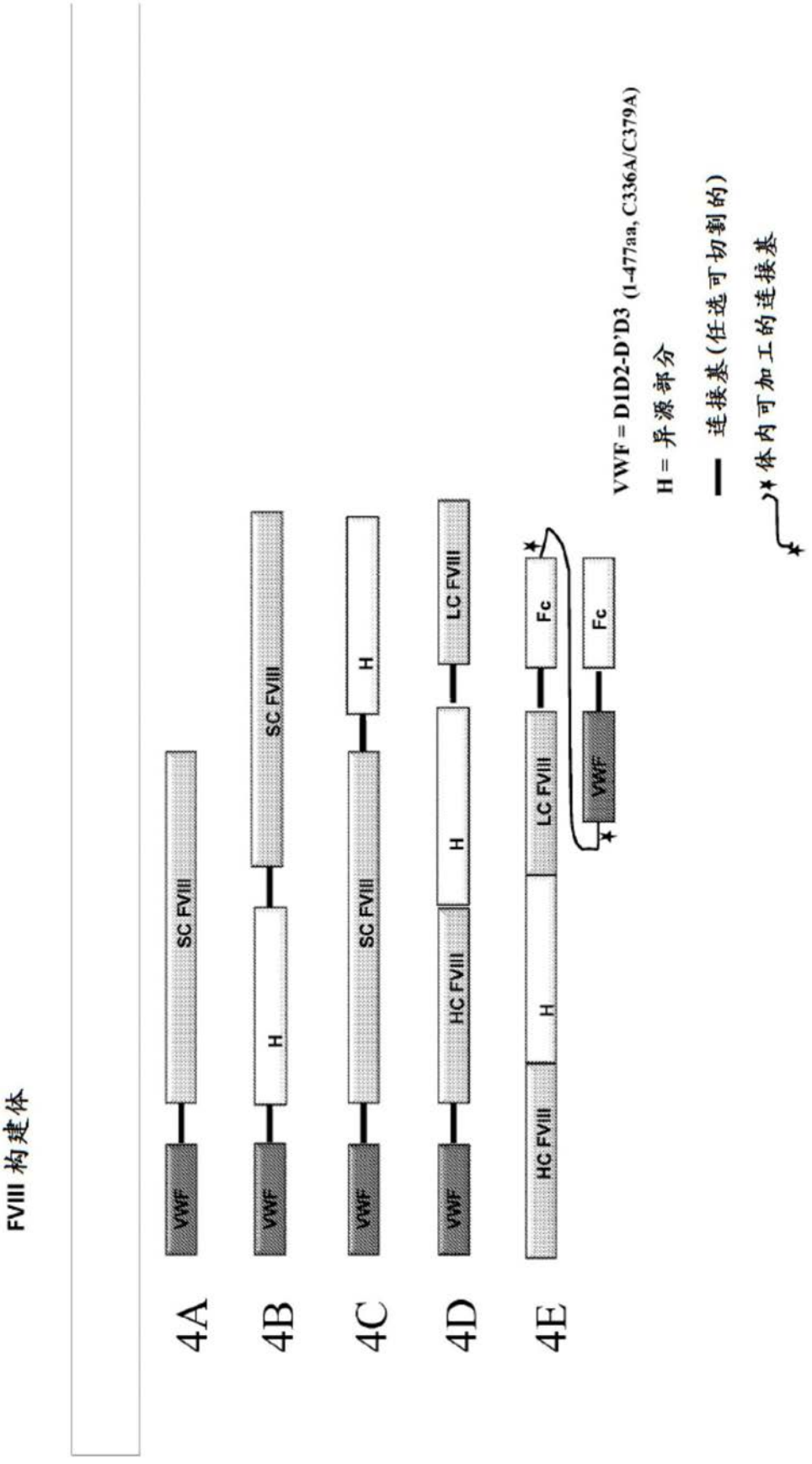

FVIII 构建体
4A
4B
4C
4D
4E
SC FVIII
H
HC FVIII
LC FVIII
VWF
Fc
VWF = D1D2-D'D3 (1-477aa, C336A/C379A)
H = 异源部分
连接基(任选可切割的)
体内可加工的连接基

图4

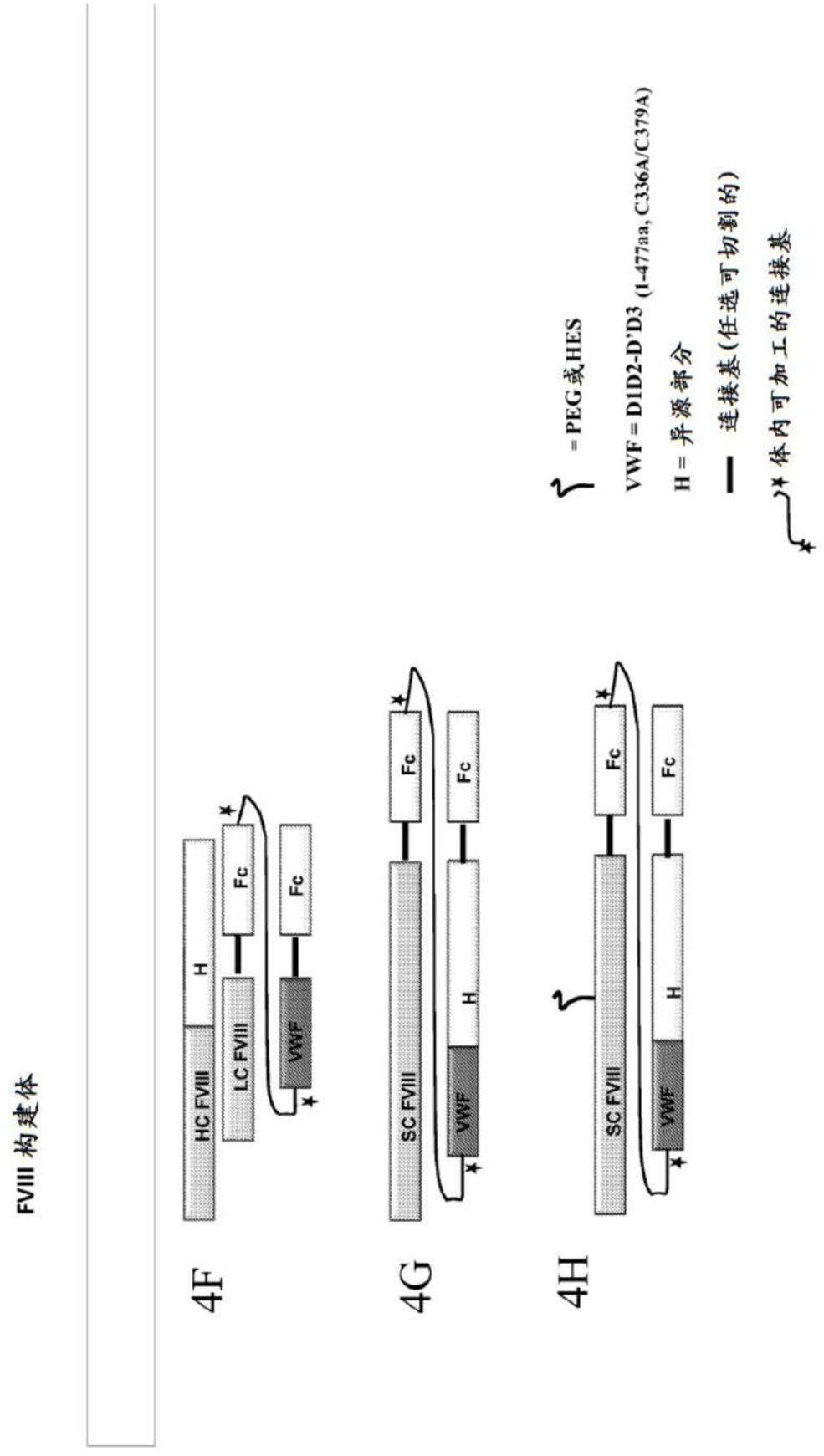
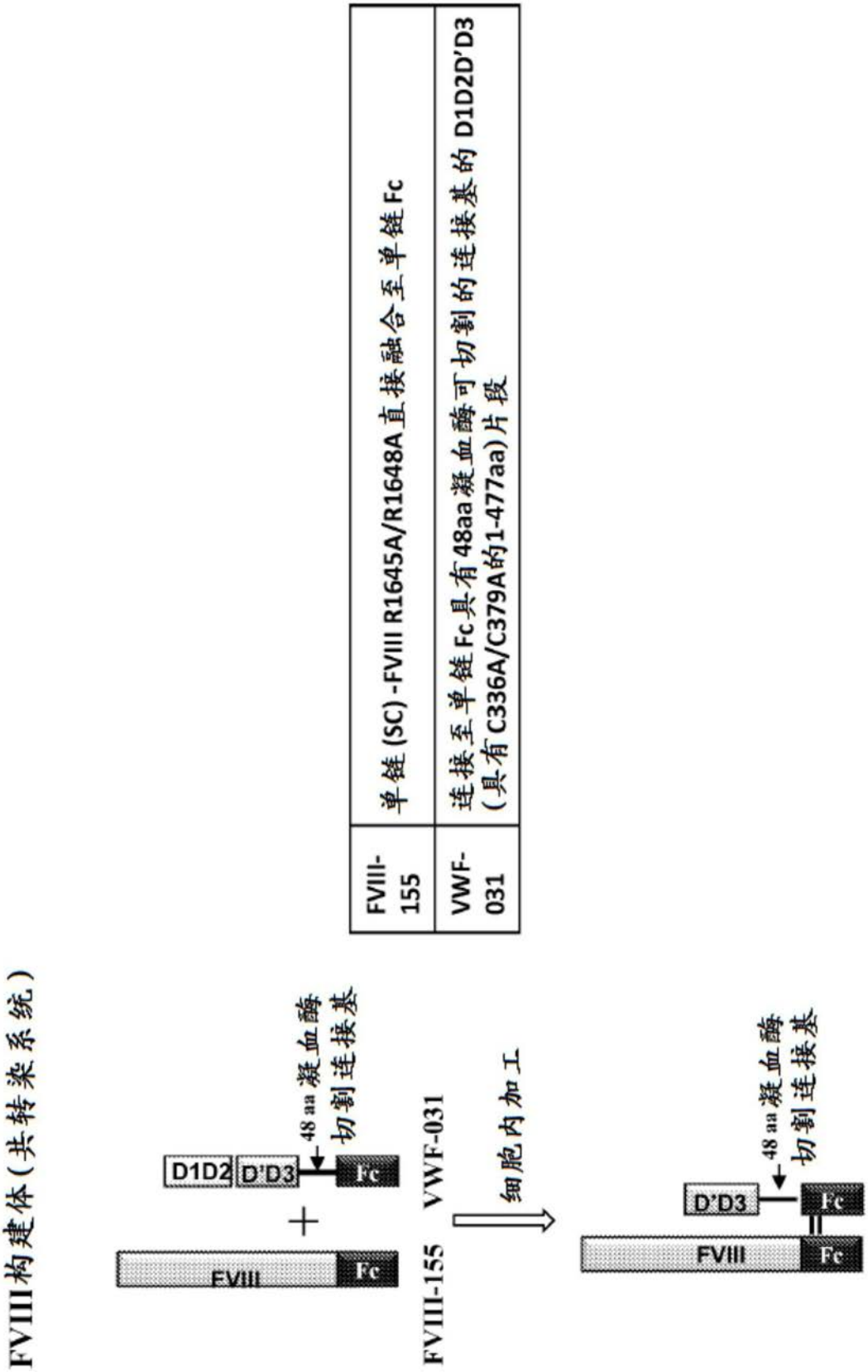

FVIII 构建体
4F
HC FVIII
H
LC FVIII
Fc
VWF
Fc
4G
SC FVIII
Fc
VWF
H
Fc
4H
SC FVIII
Fc
VWF
H
Fc
= PEG或HES
VWF = D1D2-D'D3 (1-477aa, C336A/C379A)
H = 异源部分
连接基(任选可切割的)
体内可加工的连接基

图4(续)



FVIII-155	单链 (SC) -FVIII R1645A/R1648A 直接融合至单链 Fc
VWF-031	连接至单链 Fc 具有 48aa 凝血酶可切割的连接基的 D1D2D'D3 (具有 C336A/C379A 的 1-477aa) 片段

图5

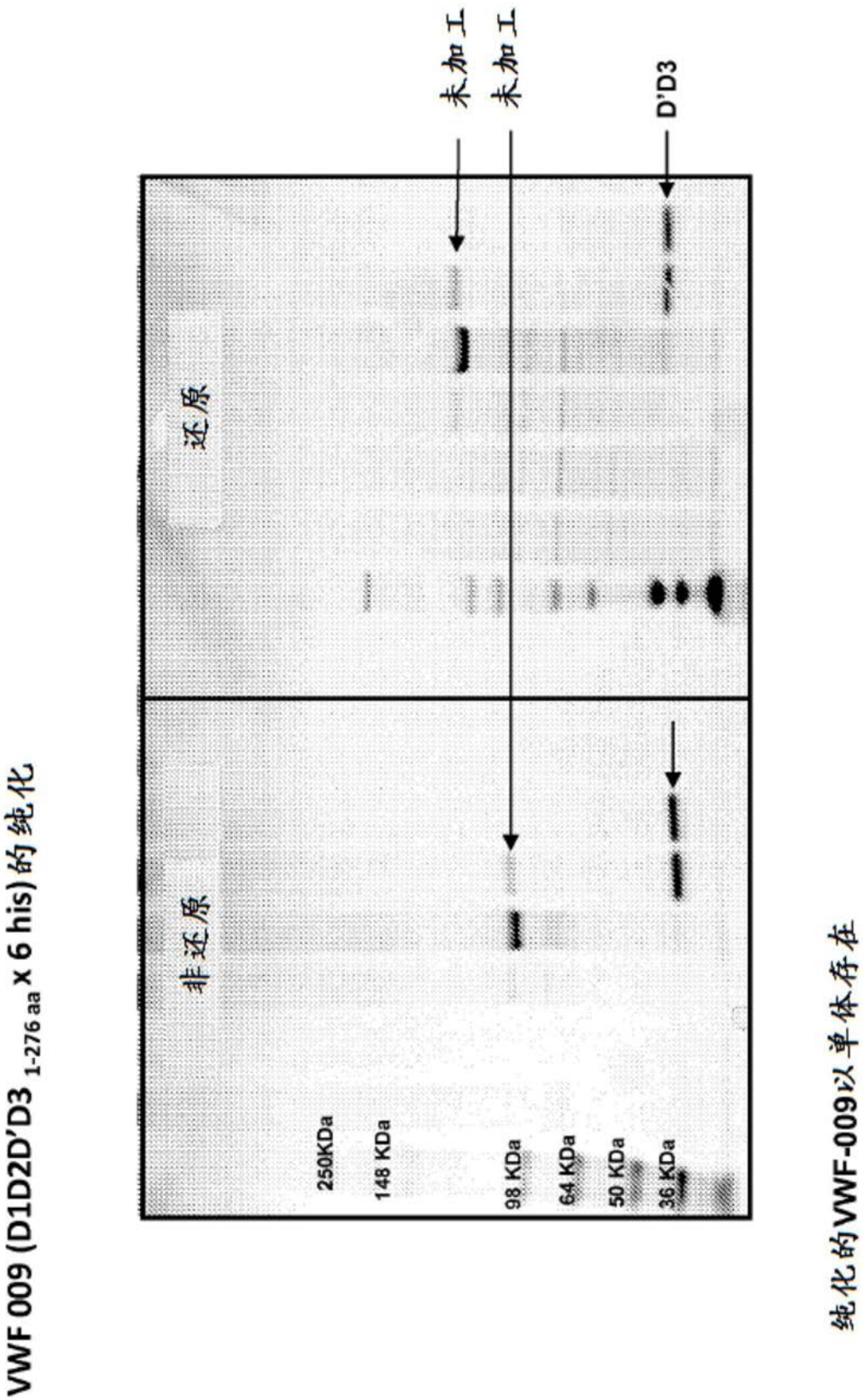

VWF 009 (D1D2D'D3 1-276 aa x 6 his)的纯化
非还原
还原
250KDa
148 KDa
98 KDa
64 KDa
50 KDa
36 KDa
未加工
未加工
D'D3
纯化的VWF-009以单体存在

图6

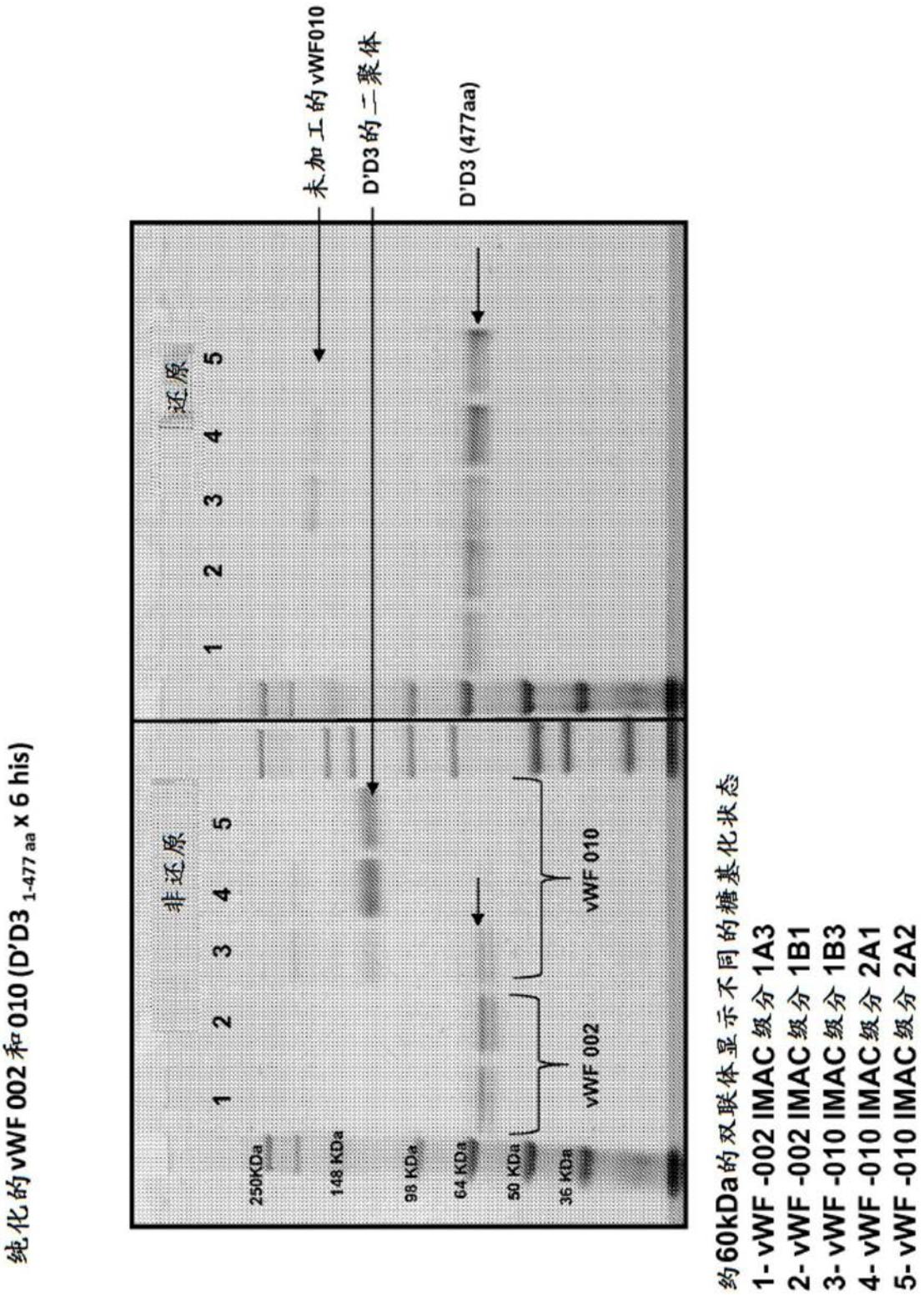

纯化的vWF 002和010 (D'D3 1-477 aa x 6 his)
非还原
还原
1 2 3 4 5
1 2 3 4 5
250KDa
148 KDa
98 KDa
64 KDa
50 KDa
36 KDa
vWF 002
vWF 010
未加工的vWF010
D'D3的二聚体
D'D3 (477aa)
约60kDa的双联体显示不同的糖基化状态
1- vWF -002 IMAC级分1A3
2- vWF -002 IMAC级分1B1
3- vWF -010 IMAC级分1B3
4- vWF -010 IMAC级分2A1
5- vWF -010 IMAC级分2A2

图7

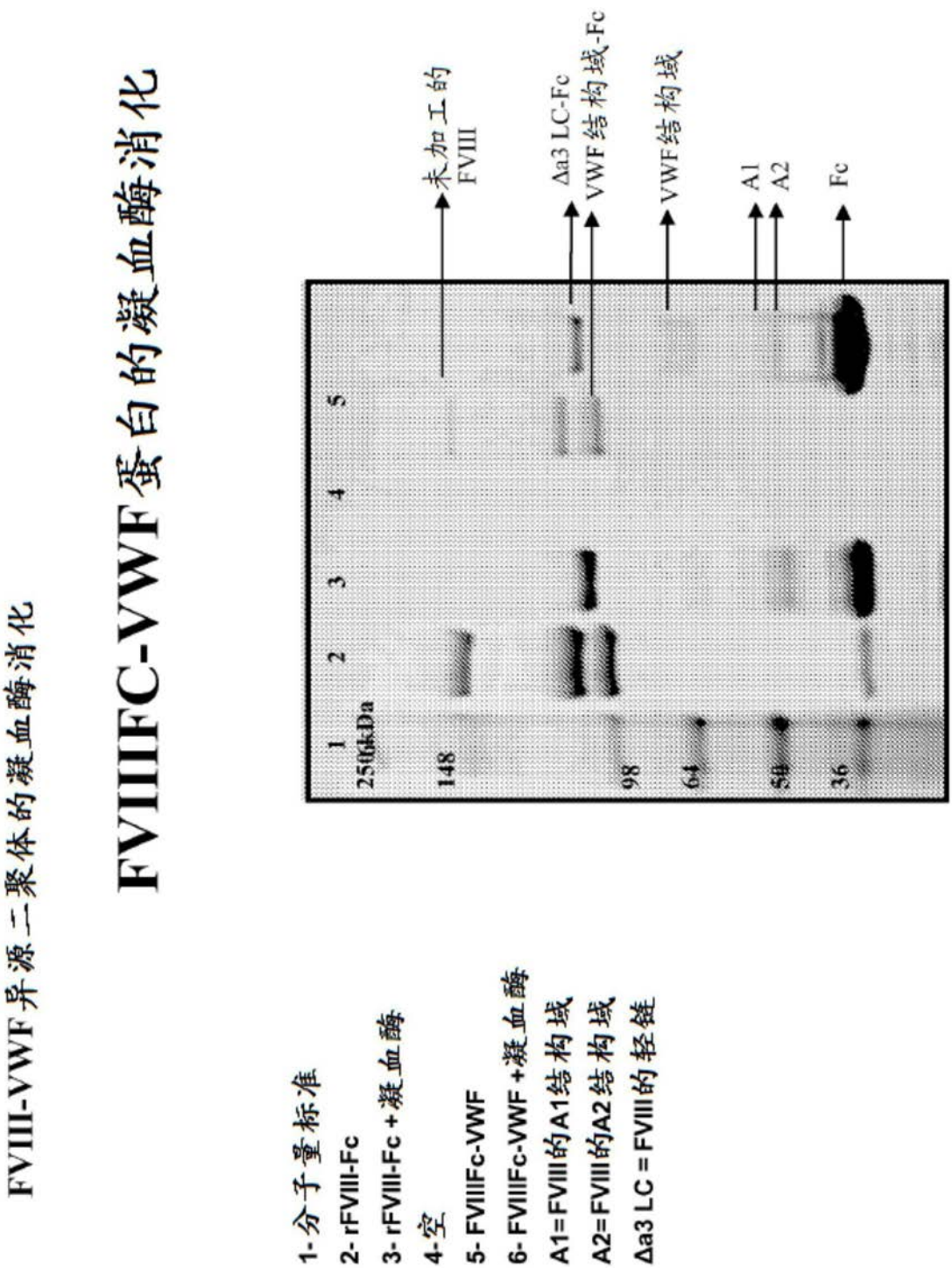

FVIII-VWF 异源二聚体的凝血酶消化
FVIIIFC-VWF 蛋白的凝血酶消化
1-分子量标准
2- rFVIII-Fc
3- rFVIII-Fc +凝血酶
4-空
5- FVIIIFc-VWF
6- FVIIIFc-VWF +凝血酶
A1=FVIII的A1结构域
A2=FVIII的A2结构域
Δa3 LC = FVIII的轻链
1
2
3
4
5
250kDa
148
98
64
50
36
未加工的 FVIII
Δa3 LC-Fc
VWF结构域-Fc
VWF结构域
A1
A2
Fc

图8

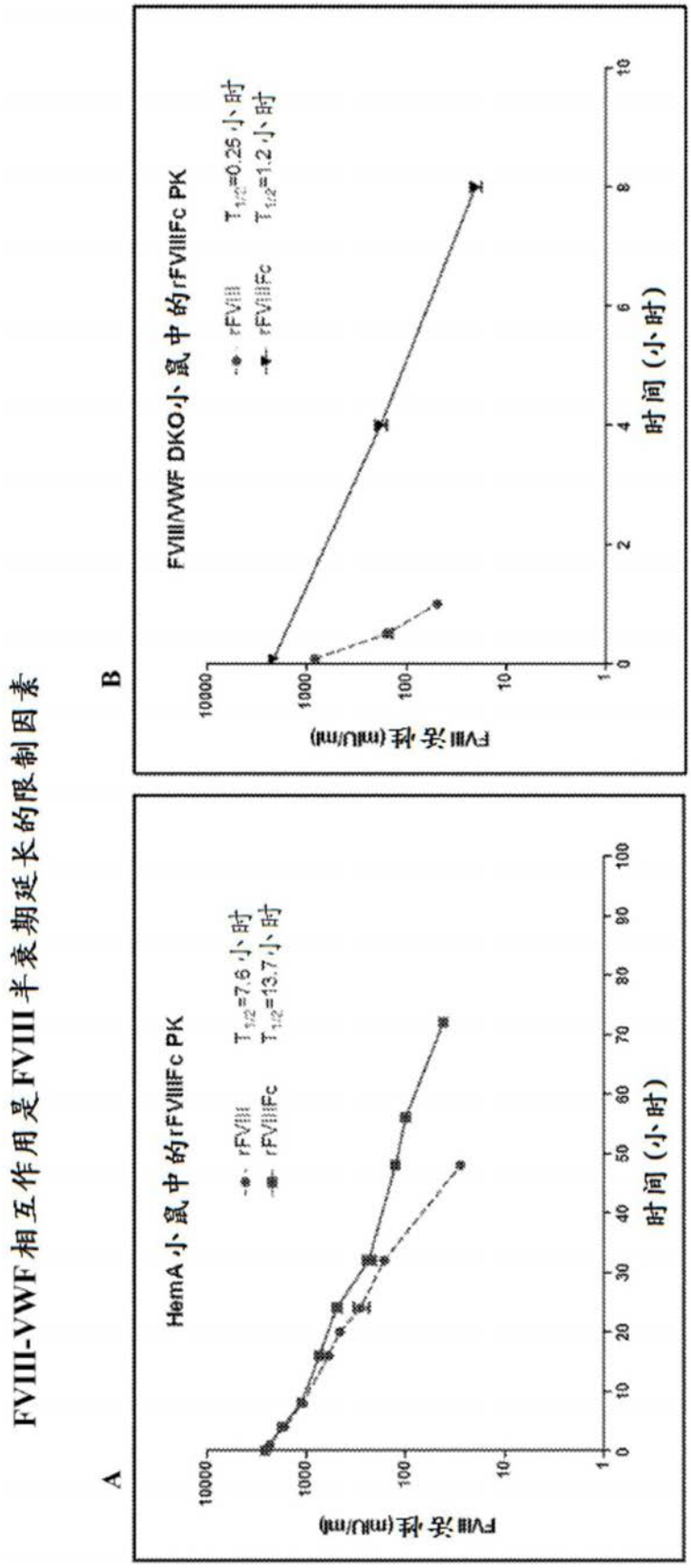

FVIII-VWF 相互作用是 FVIII 半衰期延长的限制因素
A
HemA 小鼠中的rFVIIIFc PK
rFVIII T1/2=7.6 小时
rFVIIIFc T1/2=13.7 小时
FVIII 活性 (mIU/ml)
时间 (小时)
B
FVIII/VWF DKO小鼠中的rFVIIIFc PK
rFVIII T1/2=0.25 小时
rFVIIIFc T1/2=1.2 小时
FVIII 活性 (mIU/ml)
时间 (小时)

图9

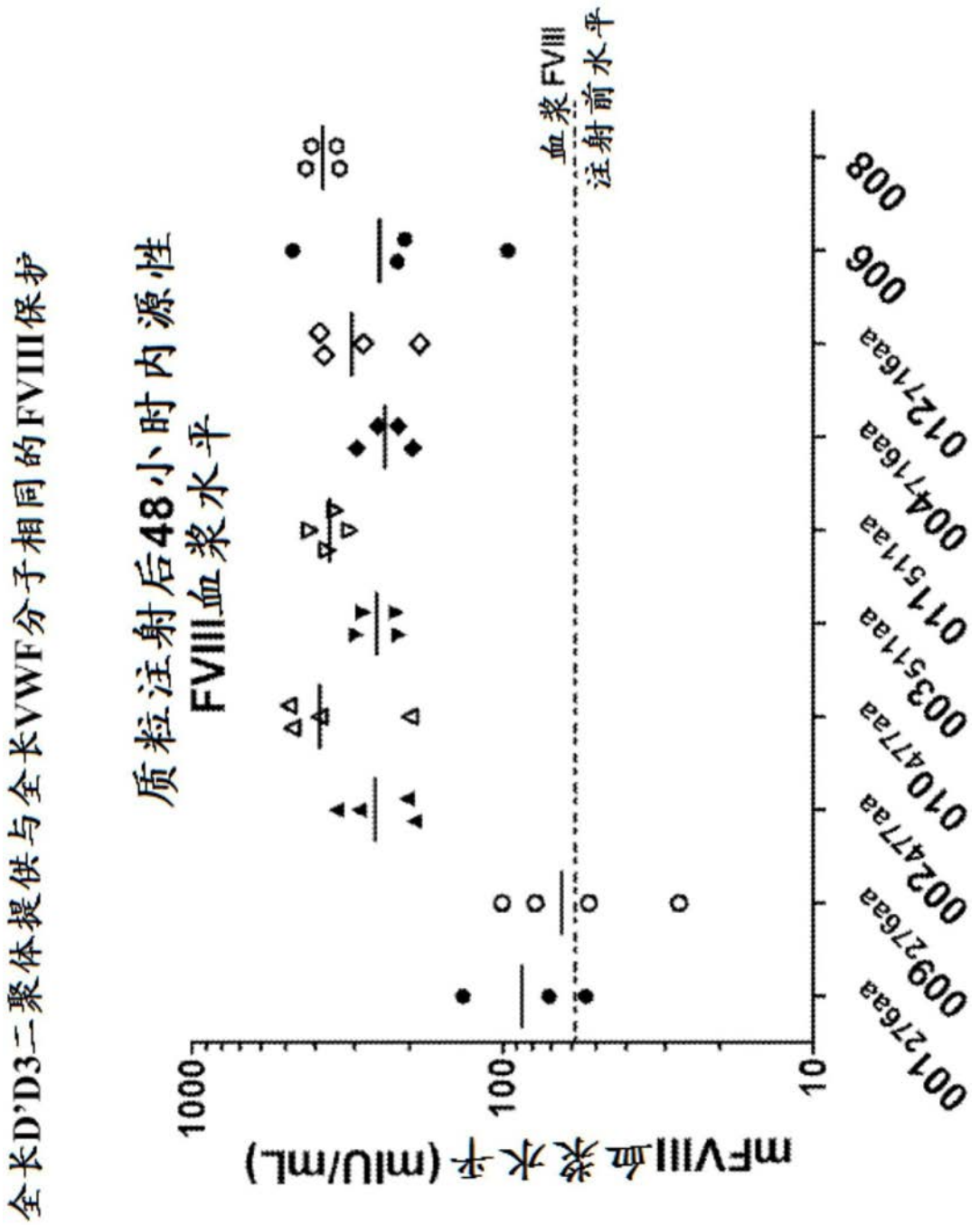

全长D'D3二聚体提供与全长VWF分子相同的FVIII保护
质粒注射后48小时内源性
FVIII血浆水平
mFVIII血浆水平 (mIU/mL)
1000
100
10
血浆FVIII
注射前水平
001276aa
009276aa
002477aa
010477aa
003511aa
011511aa
004716aa
012716aa
800
900

图10A

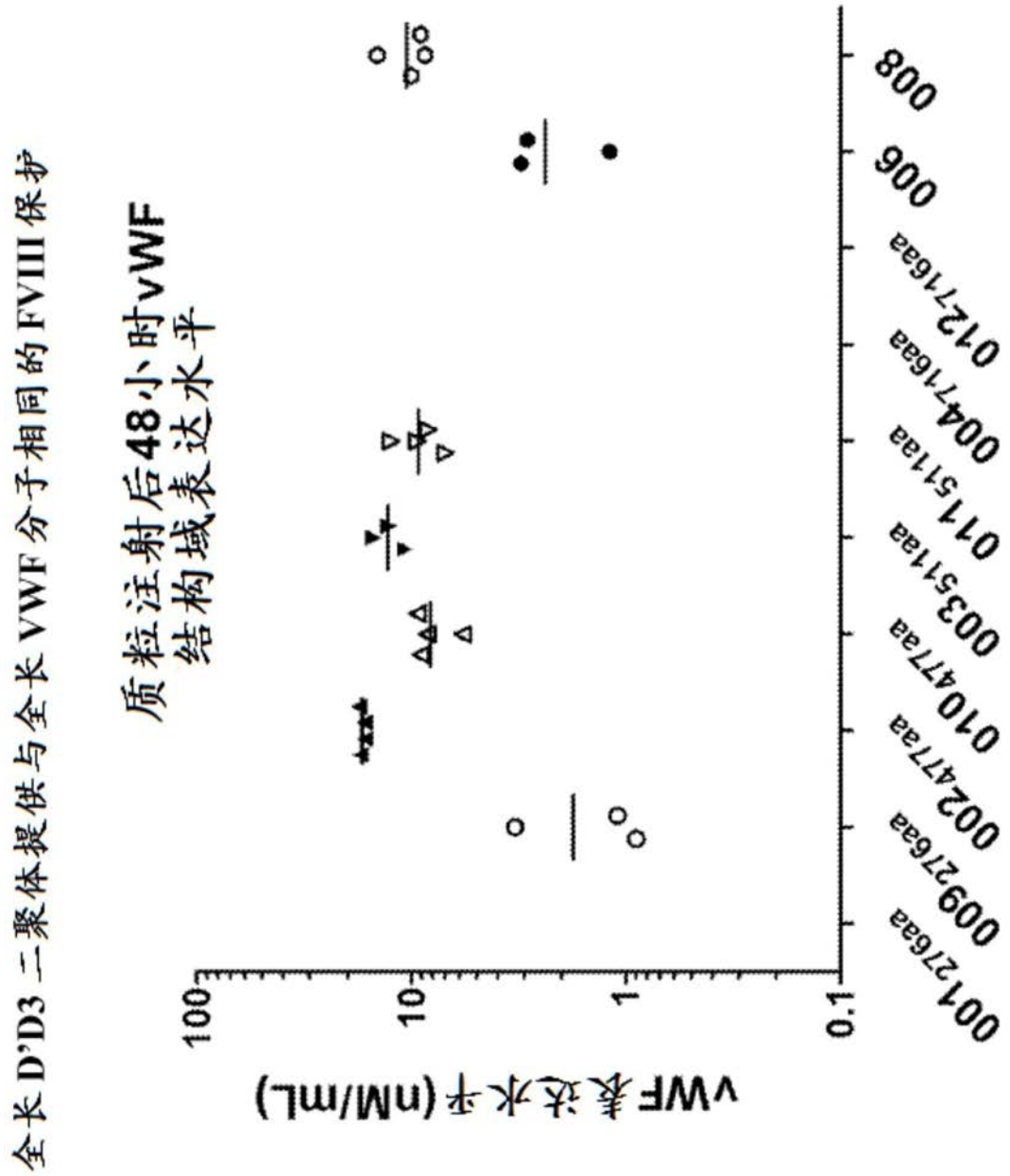

全长 D'D3 二聚体提供与全长 VWF 分子相同的 FVIII 保护
质粒注射后48小时vWF
结构域表达水平
vWF表达水平 (nM/mL)
100
10
1
0.1
001276aa
009276aa
002477aa
010477aa
003511aa
011511aa
004716aa
012716aa
006
008

图10B

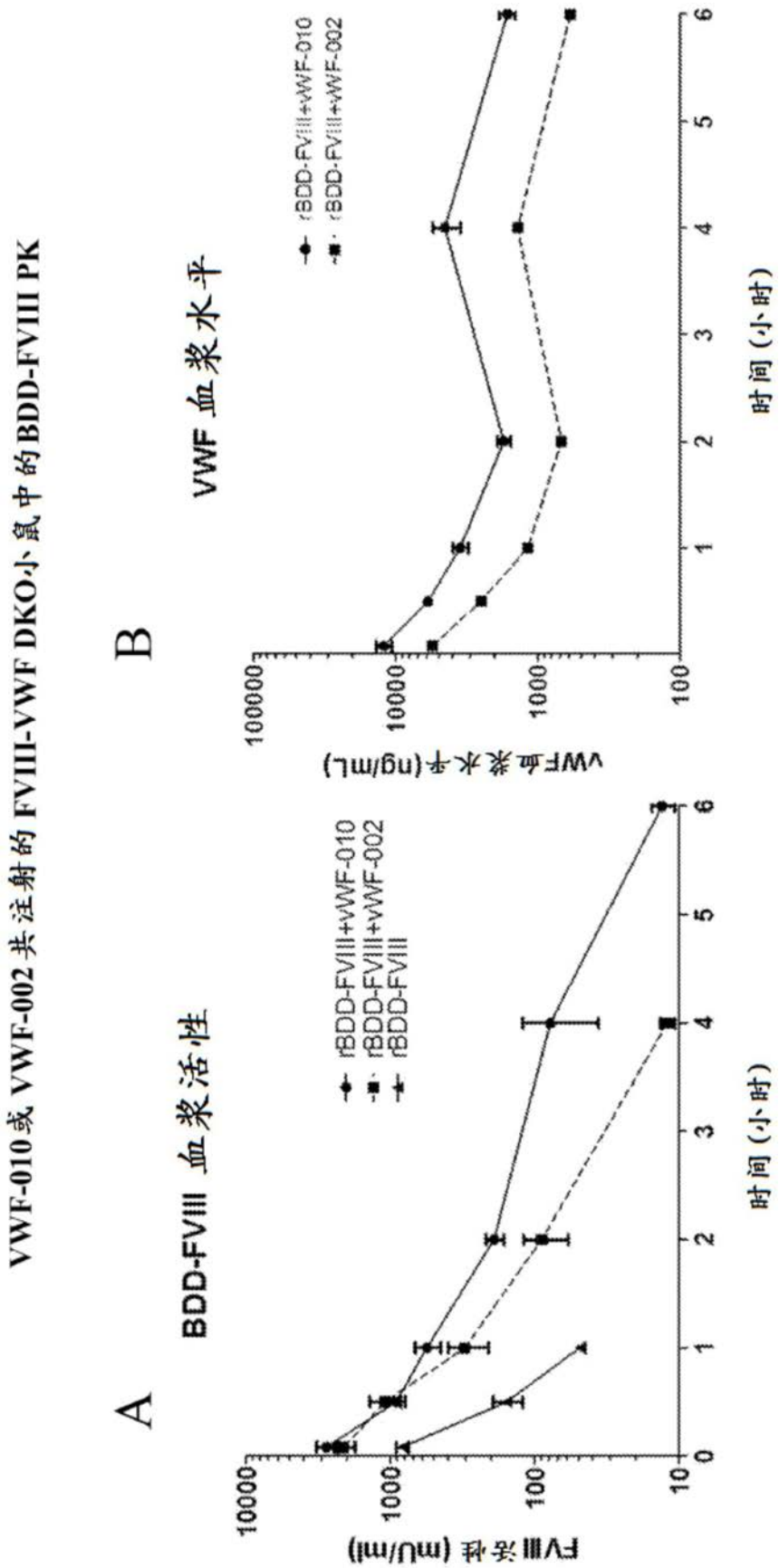

VWF-010或 VWF-002 共注射的 FVIII-VWF DKO小鼠中的BDD-FVIII PK
A
BDD-FVIII 血浆活性
rBDD-FVIII+vWF-010
rBDD-FVIII+vWF-002
rBDD-FVIII
FVIII活性 (mU/ml)
时间 (小时)
B
VWF 血浆水平
rBDD-FVIII+vWF-010
rBDD-FVIII+vWF-002
vWF血浆水平 (ng/mL)
时间 (小时)

图11

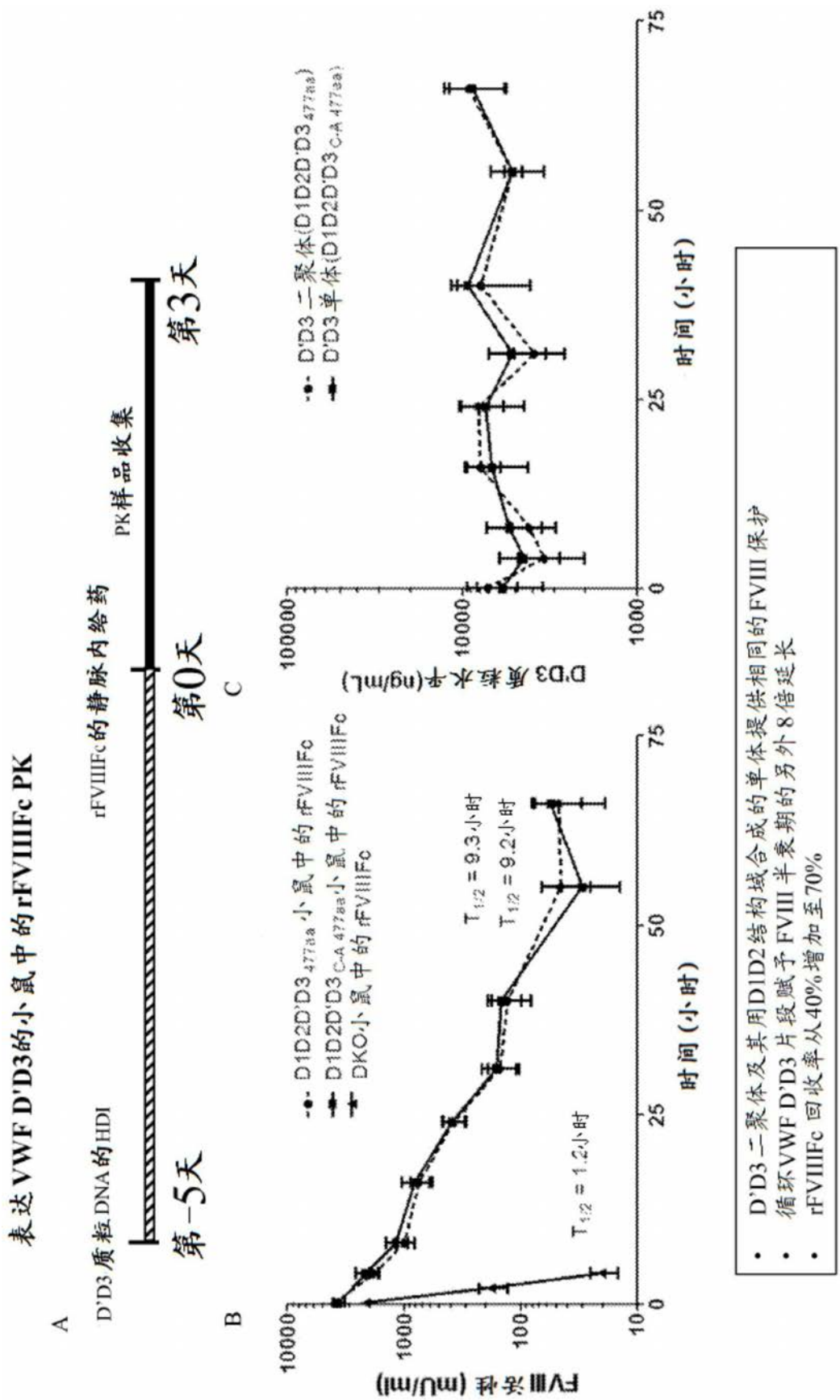

表达 VWF D'D3的小鼠中的rFVIIIFc PK
A
D'D3质粒DNA的HDI
rFVIIIFc的静脉内给药
PK样品收集
第-5天
第0天
第3天
B
C
D1D2D'D3 477aa小鼠中的rFVIIIFc
D1D2D'D3 C-A 477aa小鼠中的rFVIIIFc
DKO小鼠中的rFVIIIFc
T1/2 = 1.2小时
T1/2 = 9.3小时
T1/2 = 9.2小时
FVIII活性 (mU/ml)
时间 (小时)
D'D3 二聚体(D1D2D'D3 477aa)
D'D3 单体(D1D2D'D3 C-A 477aa)
D'D3 蛋白水平(ng/mL)
时间 (小时)
D'D3二聚体及其用D1D2结构域合成的单体提供相同的FVIII保护
循环VWF D'D3片段赋予FVIII半衰期的另外8倍延长
rFVIIIFc回收率从40%增加至70%

图12

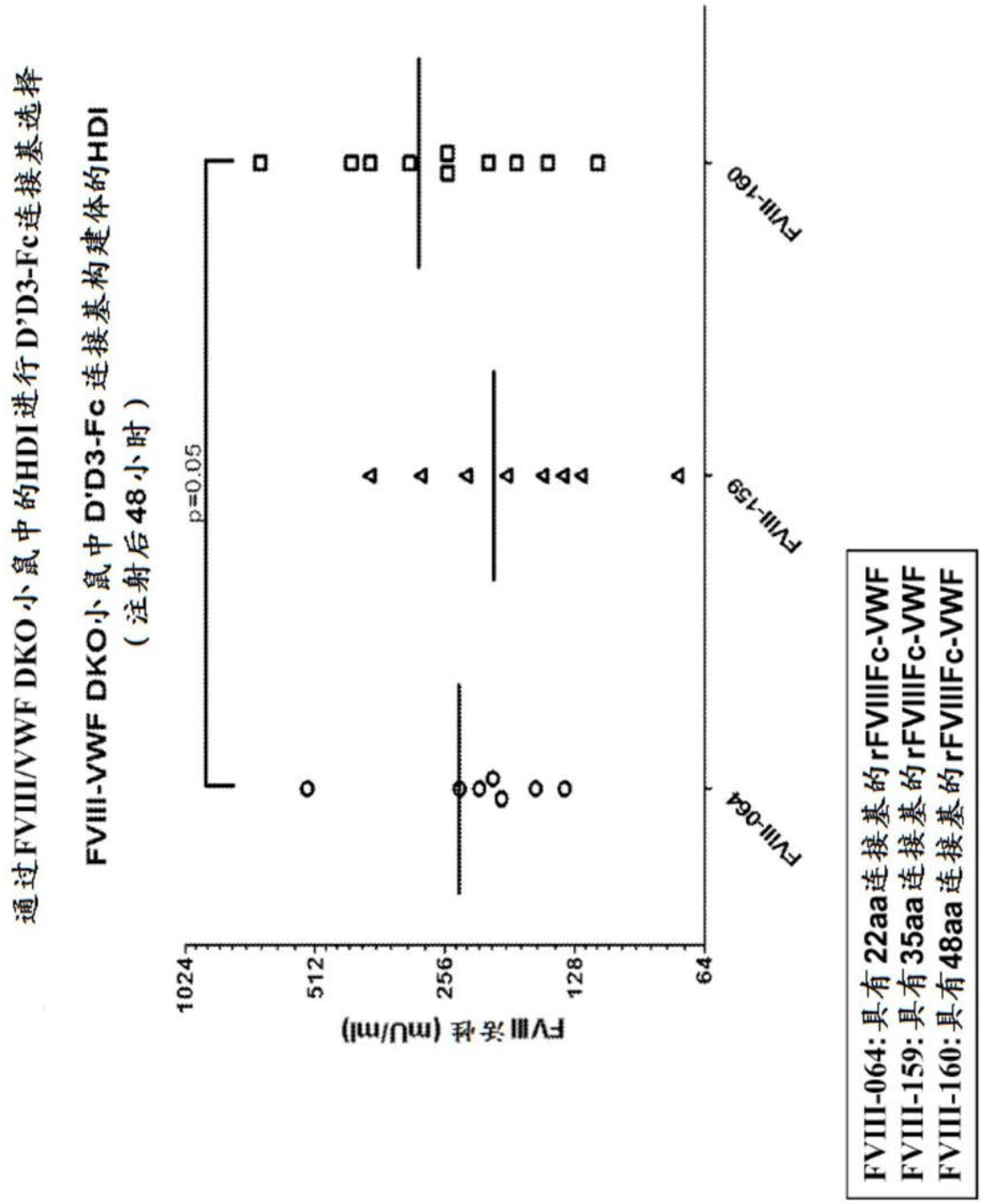

通过FVIII/VWF DKO小鼠中的HDI进行D'D3-Fc连接基选择
FVIII-VWF DKO小鼠中D'D3-Fc连接基构建体的HDI
（注射后48小时）
p=0.05
FVIII活性 (mU/ml)
1024
512
256
128
64
FVIII-064
FVIII-159
FVIII-160
FVIII-064: 具有22aa连接基的rFVIIIFc-VWF
FVIII-159: 具有35aa连接基的rFVIIIFc-VWF
FVIII-160: 具有48aa连接基的rFVIIIFc-VWF

图13

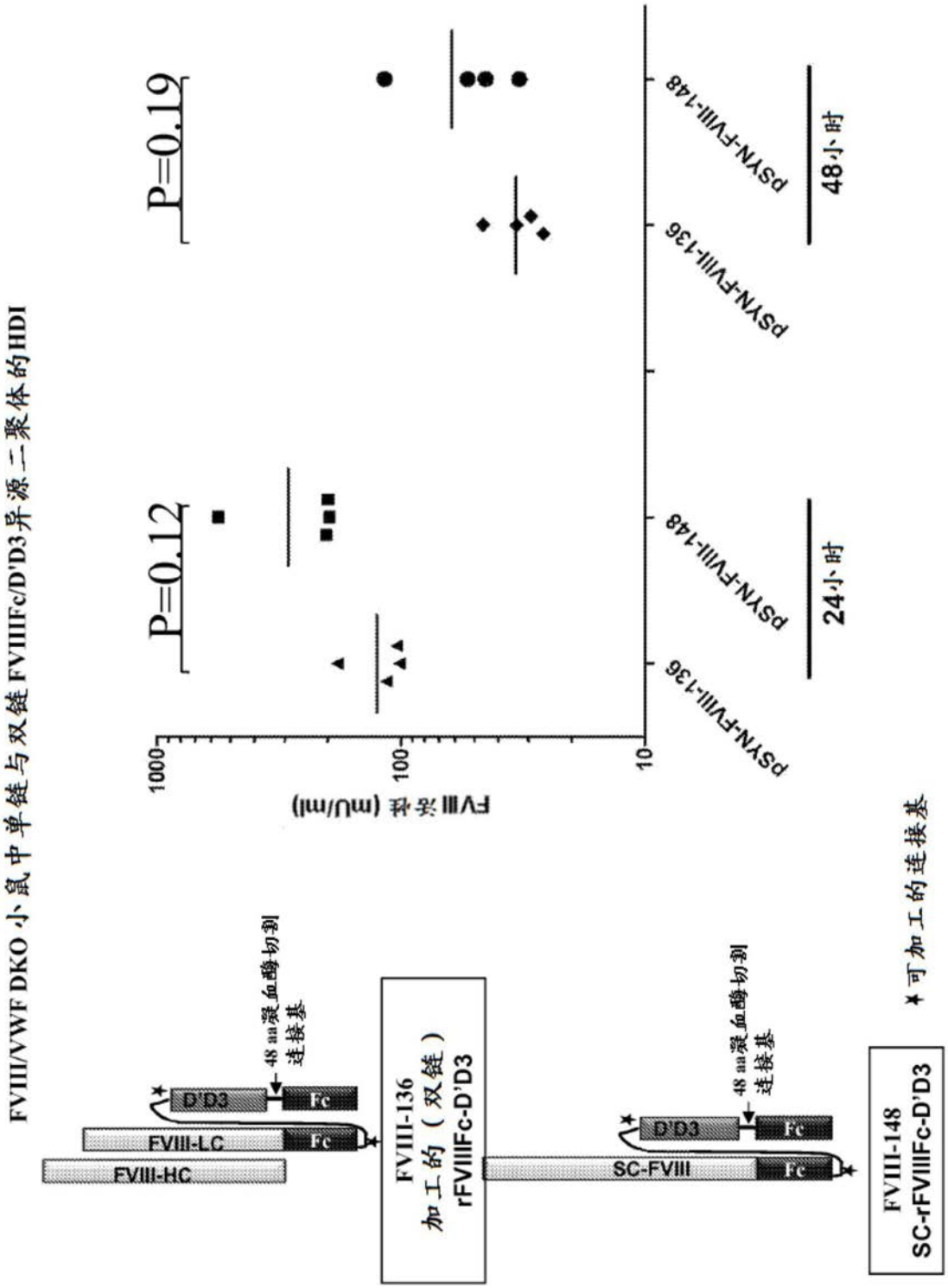

FVIII/VWF DKO 小鼠中单链与双链 FVIIIFc/D'D3 异源二聚体的 HDI
P=0.19
P=0.12
FVIII 活性 (mU/ml)
1000
100
10
pSYN-FVIII-136
pSYN-FVIII-148
pSYN-FVIII-136
pSYN-FVIII-148
24小时
48小时
FVIII-LC
FVIII-HC
D'D3
Fc
48 aa凝血酶切割连接基
FVIII-136
加工的（双链）
rFVIIIFc-D'D3
SC-FVIII
FVIII-148
SC-rFVIIIFc-D'D3
★可加工的连接基

图14

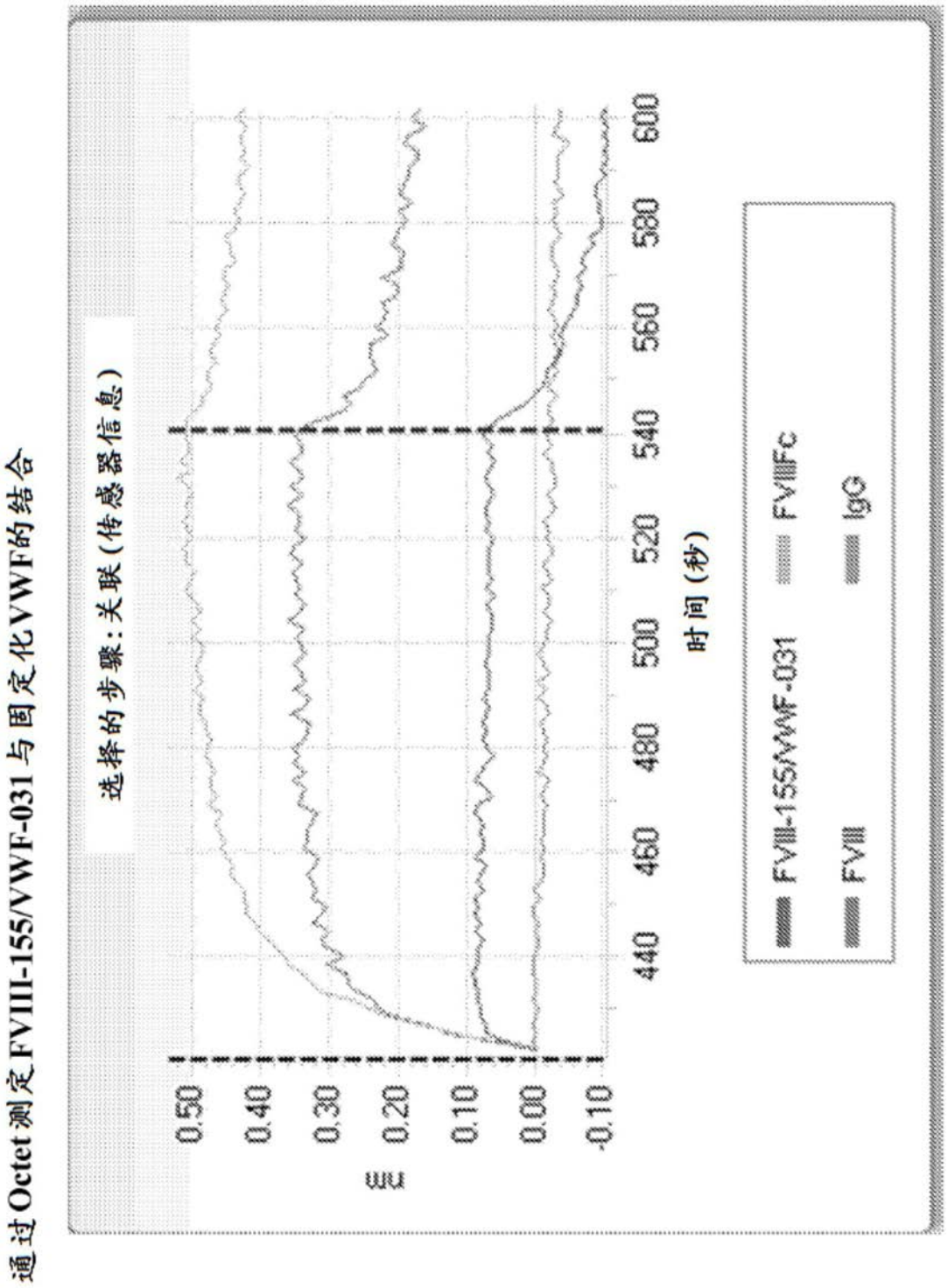

通过Octet测定FVIII-155/VWF-031与固定化VWF的结合
选择的步骤:关联(传感器信息)
nm
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
440
460
480
500
520
540
560
580
600
时间(秒)
FVIII-155/VWF-031
FVIIIFc
FVIII
IgG

图15

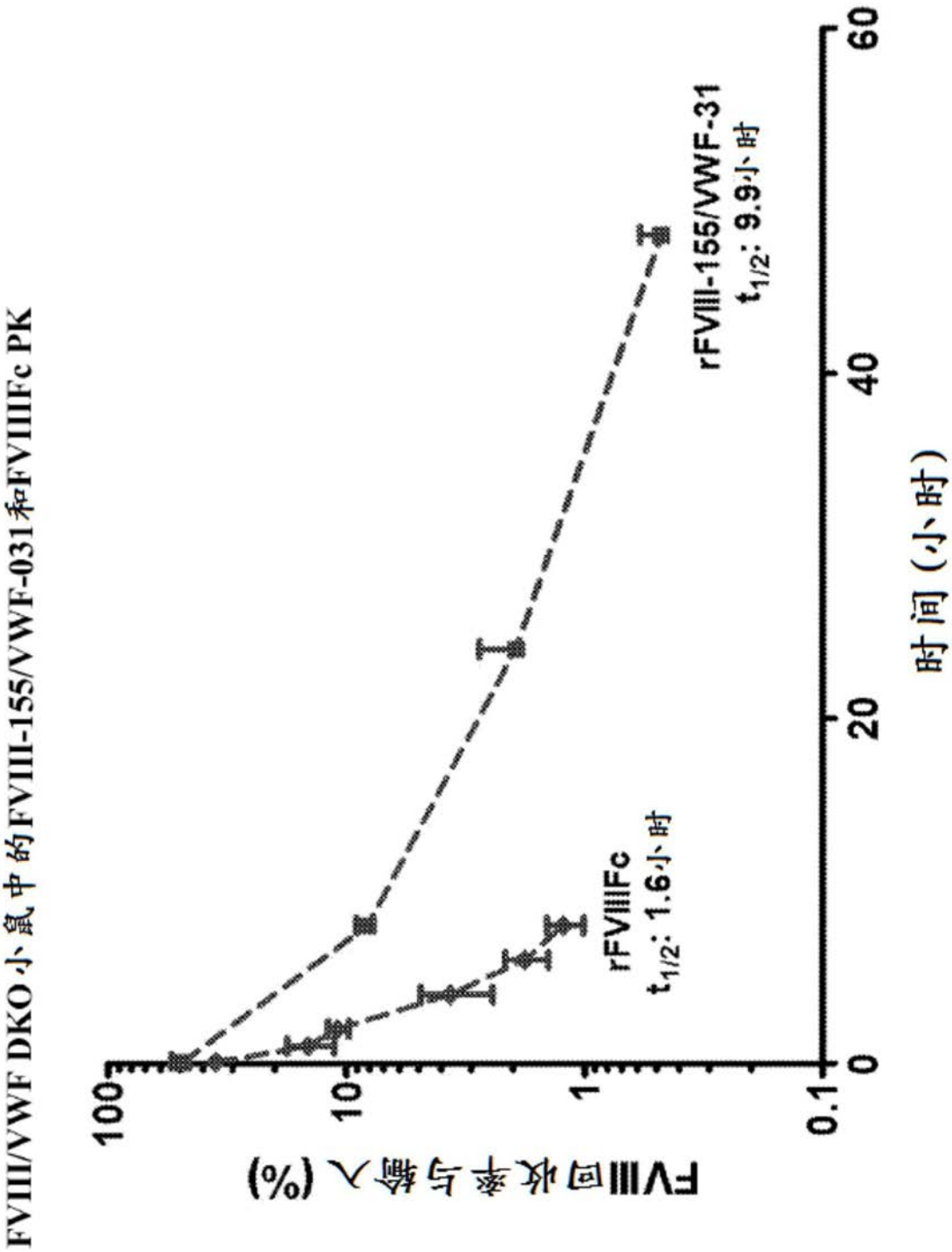

FVIII/VWF DKO小鼠中的FVIII-155/VWF-031和FVIIIFc PK
FVIII回收率与输入 (%)
100
10
1
0.1
0
20
40
60
时间 (小时)
rFVIII-155/VWF-31
t1/2: 9.9小时
rFVIIIFc
t1/2: 1.6小时

图16

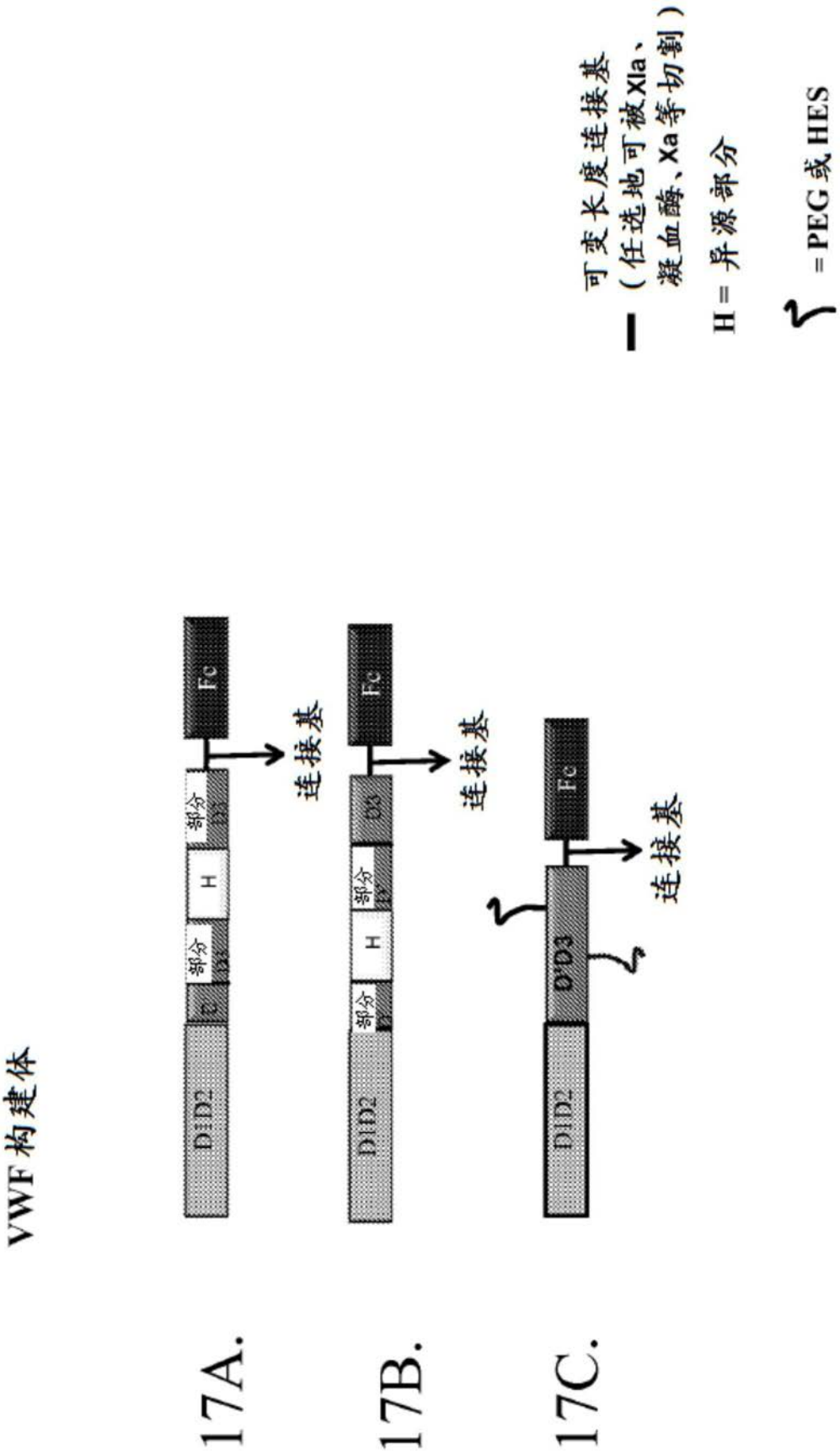

VWF 构建体
D1D2
部分
H
Fc
连接基
D3
D'D3
17A.
17B.
17C.
可变长度连接基
（任选地可被Xla、
凝血酶、Xa等切割）
H = 异源部分
= PEG或HES

图17

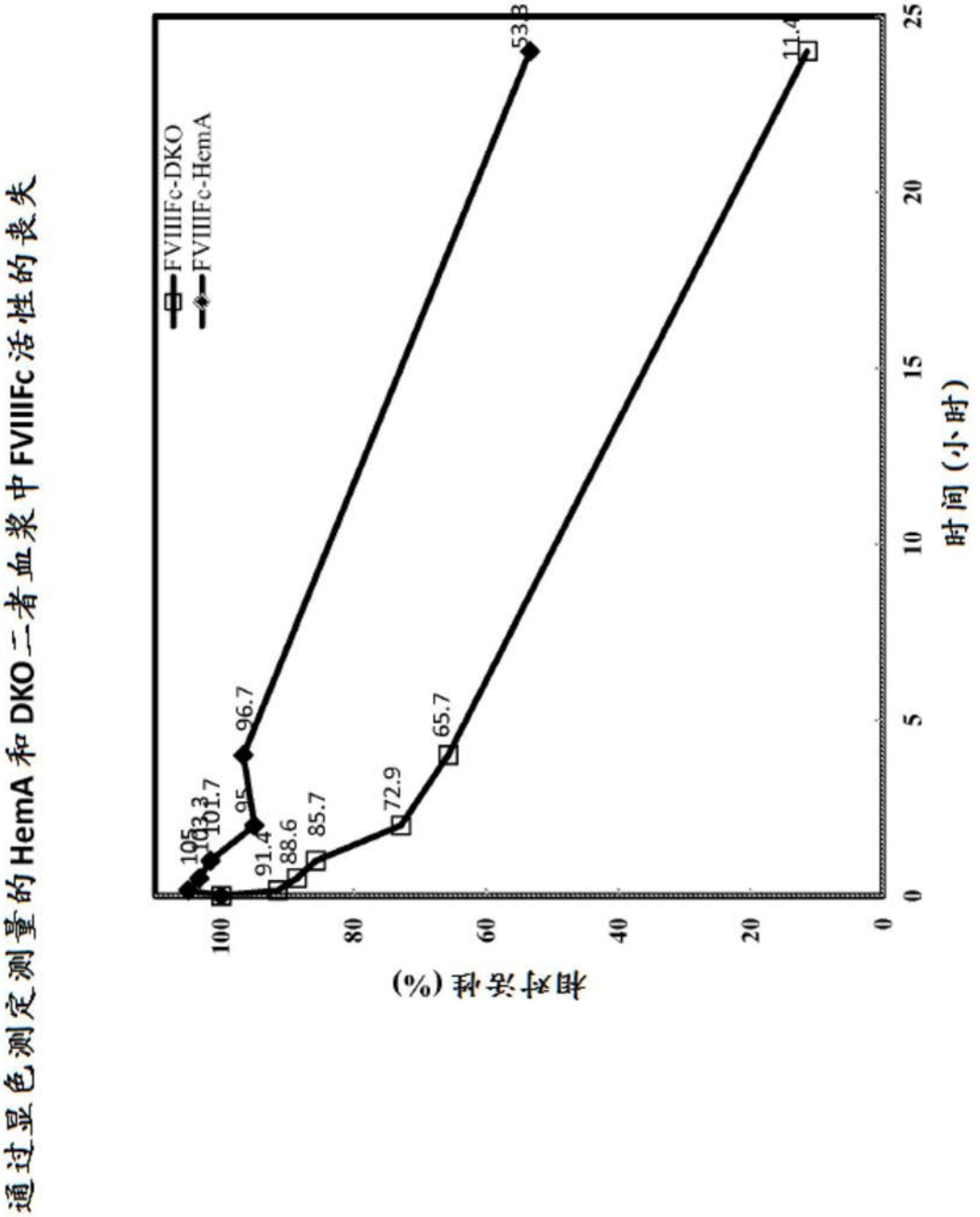

通过显色测定测量的HemA和DKO二者血浆中FVIIIFc活性的丧失
FVIIIFc-DKO
FVIIIFc-HemA
相对活性 (%)
时间 (小时)
0
20
40
60
80
100
0
5
10
15
20
25
53.3
11.4
96.7
65.7
72.9
85.7
88.6
91.4
95
101.7
105

图18A

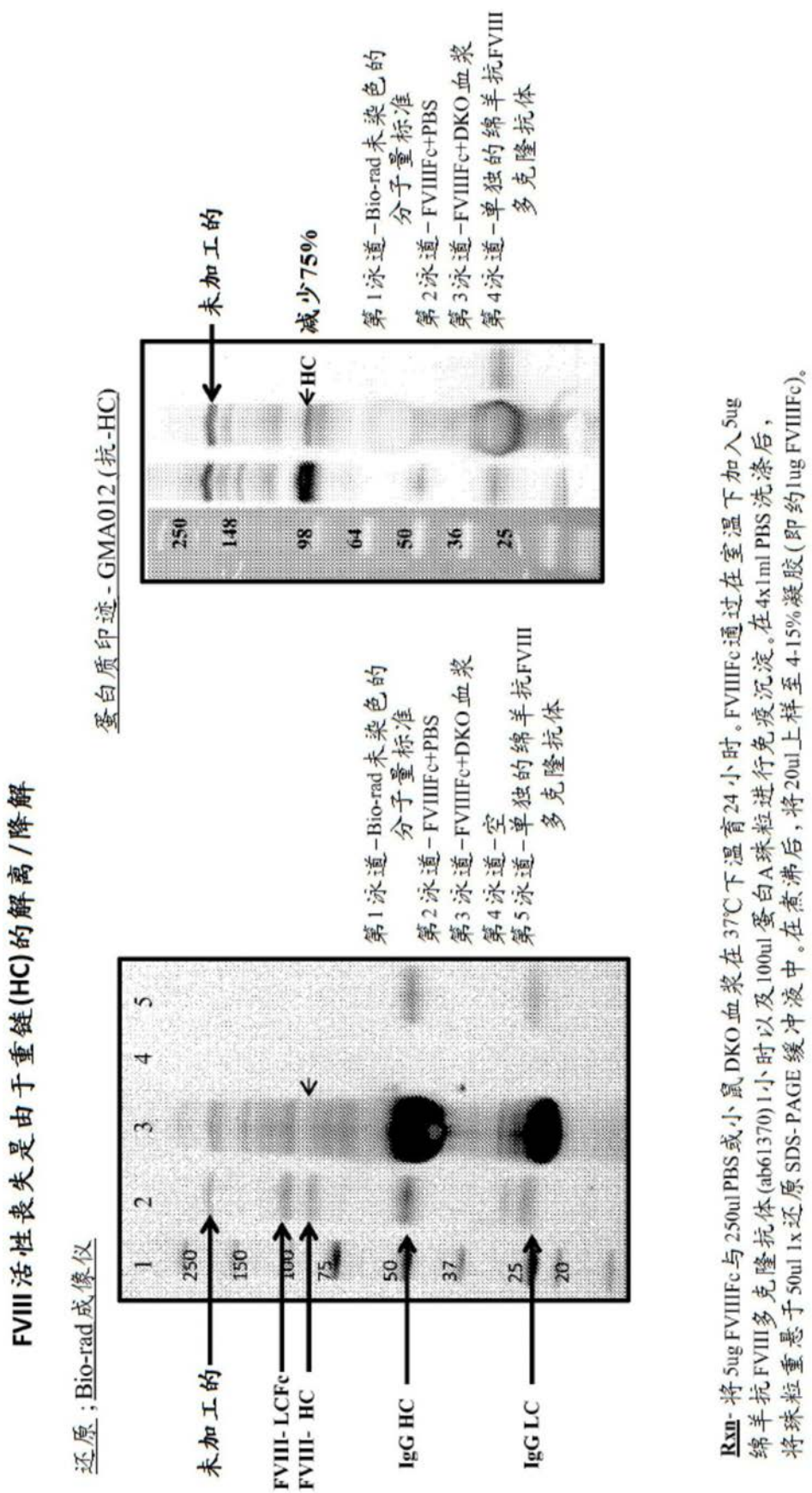

FVIII 活性丧失是由于重链(HC)的解离/降解
还原；Bio-rad成像仪
1 2 3 4 5
250
150
100
75
50
37
25
20
未加工的
FVIII- LCFc
FVIII- HC
IgG HC
IgG LC
第1泳道–Bio-rad未染色的分子量标准
第2泳道–FVIIIFc+PBS
第3泳道–FVIIIFc+DKO血浆
第4泳道–空
第5泳道–单独的绵羊抗FVIII多克隆抗体
蛋白质印迹 - GMA012 (抗-HC)
250
148
98
64
50
36
25
未加工的
减少75%
HC
第1泳道–Bio-rad未染色的分子量标准
第2泳道–FVIIIFc+PBS
第3泳道–FVIIIFc+DKO血浆
第4泳道–单独的绵羊抗FVIII多克隆抗体
Rxn- 将5ug FVIIIFc与250ul PBS或小鼠DKO血浆在37℃下温育24小时。FVIIIFc通过在室温下加入5ug绵羊抗FVIII多克隆抗体(ab61370)1小时以及100ul蛋白A珠粒进行免疫沉淀。在4x1ml PBS洗涤后，将珠粒重悬于50ul 1x还原SDS-PAGE缓冲液中。在煮沸后，将20ul上样至4-15%凝胶(即约1ug FVIIIFc)。

图18B

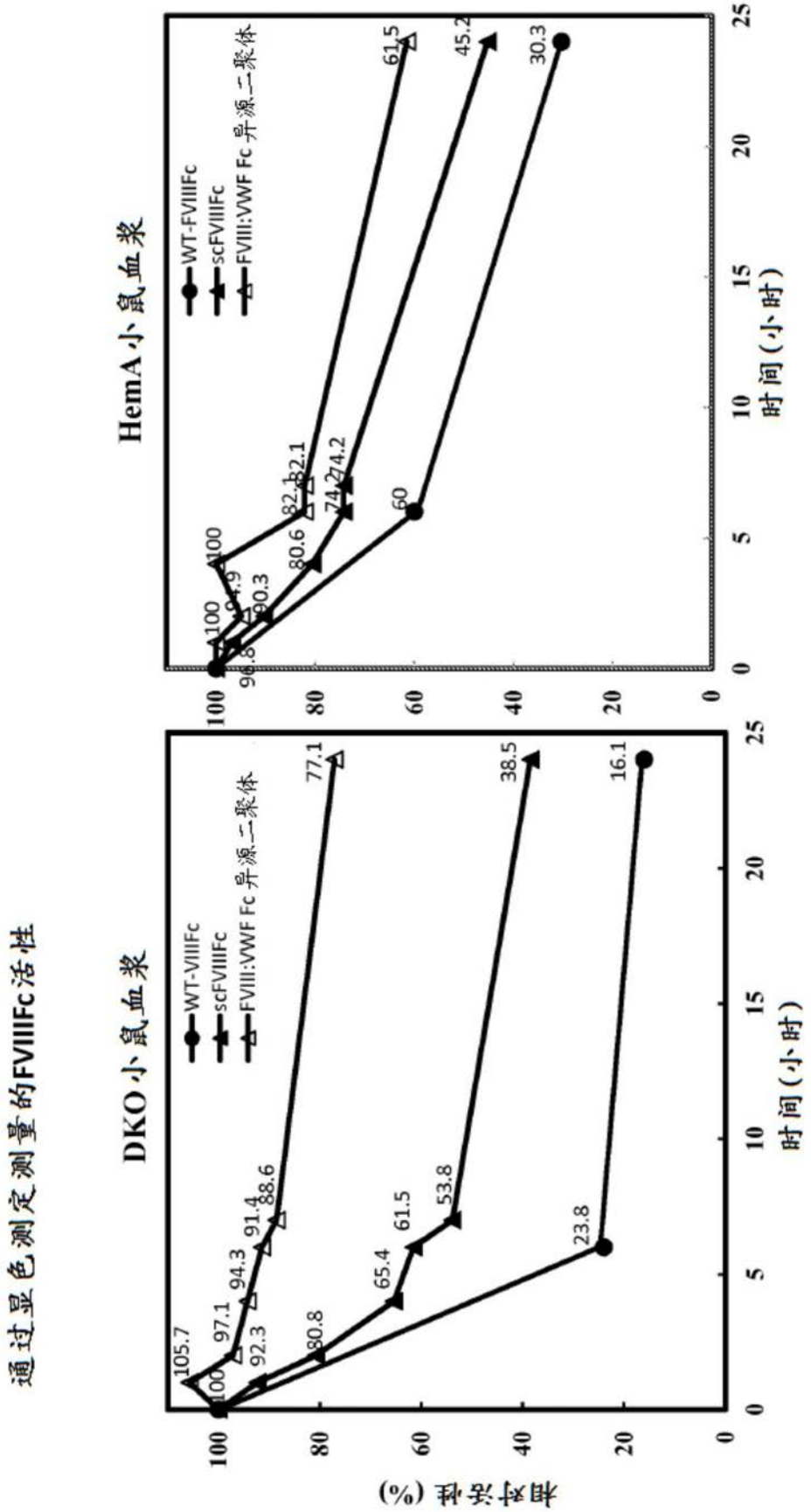

通过显色测定测量的FVIIIFc活性
DKO小鼠血浆
HemA小鼠血浆
WT-FVIIIFc
scFVIIIFc
FVIII:VWF Fc异源二聚体
相对活性 (%)
时间 (小时)
105.7
100
97.1
92.3
94.3
91.4
88.6
80.8
77.1
65.4
61.5
53.8
38.5
23.8
16.1
100
100
100
94.9
90.3
96.8
82.6
82.1
80.6
74.2
74.2
61.5
60
45.2
30.3

图19

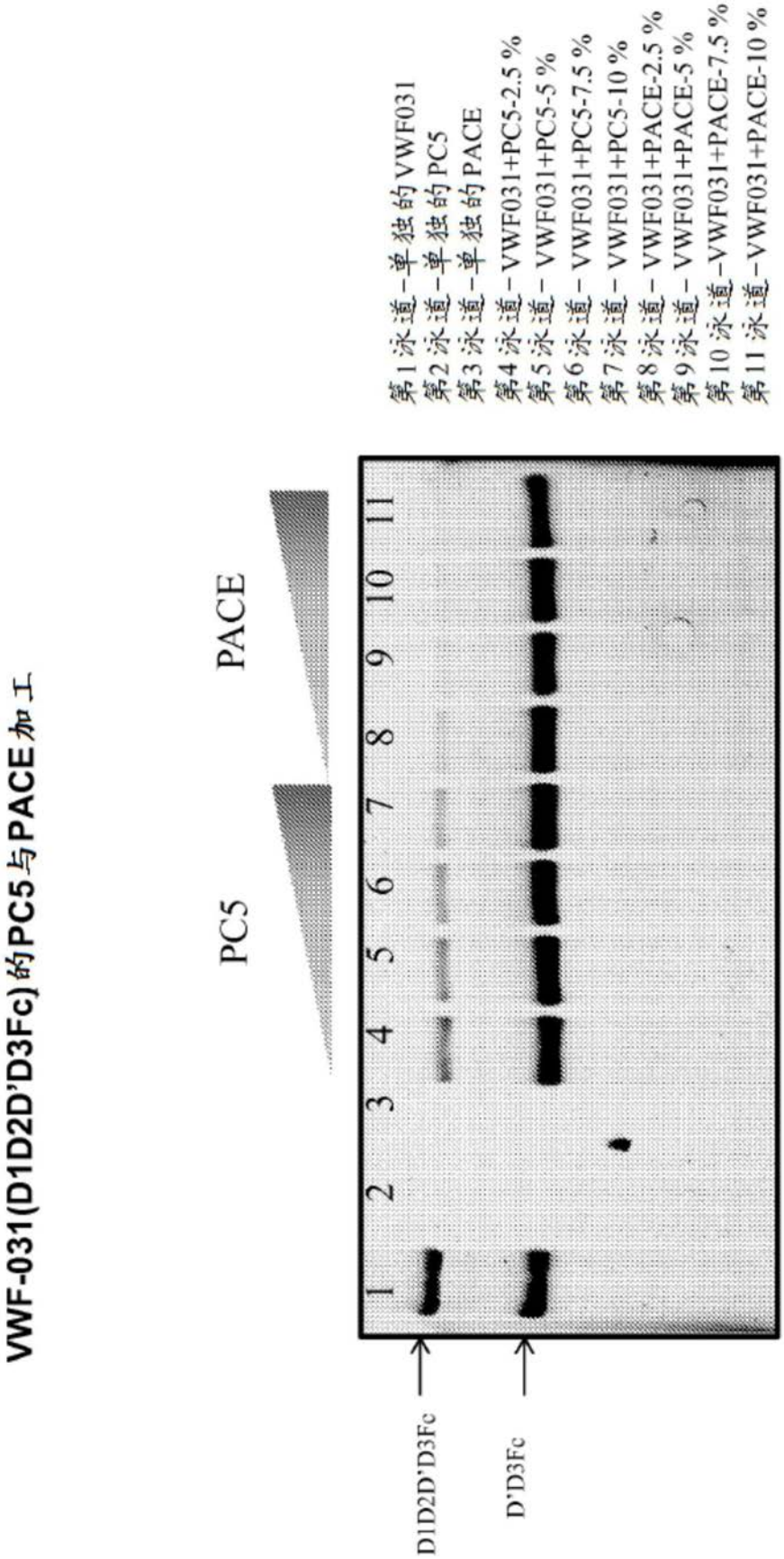

VWF-031(D1D2D'D3Fc)的PC5与PACE加工
PC5
PACE
1 2 3 4 5 6 7 8 9 10 11
D1D2D'D3Fc
D'D3Fc
第1泳道–单独的VWF031
第2泳道–单独的PC5
第3泳道–单独的PACE
第4泳道– VWF031+PC5-2.5 %
第5泳道– VWF031+PC5-5 %
第6泳道– VWF031+PC5-7.5 %
第7泳道– VWF031+PC5-10 %
第8泳道– VWF031+PACE-2.5 %
第9泳道– VWF031+PACE-5 %
第10泳道–VWF031+PACE-7.5 %
第11泳道–VWF031+PACE-10 %

图20

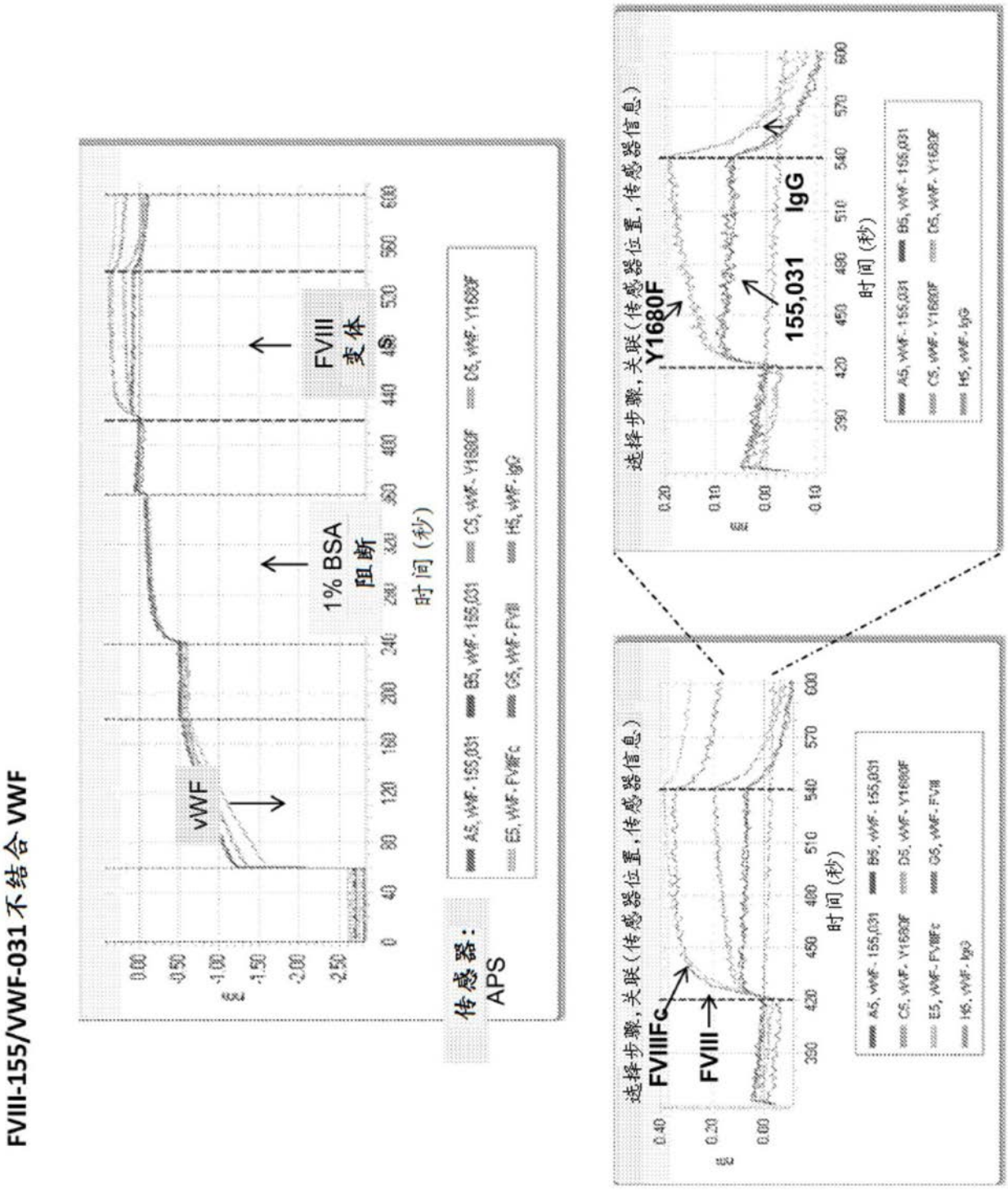

FVIII-155/VWF-031 不结合 VWF
vWF
1% BSA 阻断
FVIII 变体
时间 (秒)
传感器: APS
选择步骤，关联（传感器位置，传感器信息）
FVIIIFc
FVIII
Y1680F
155,031
IgG
时间 (秒)

图21A

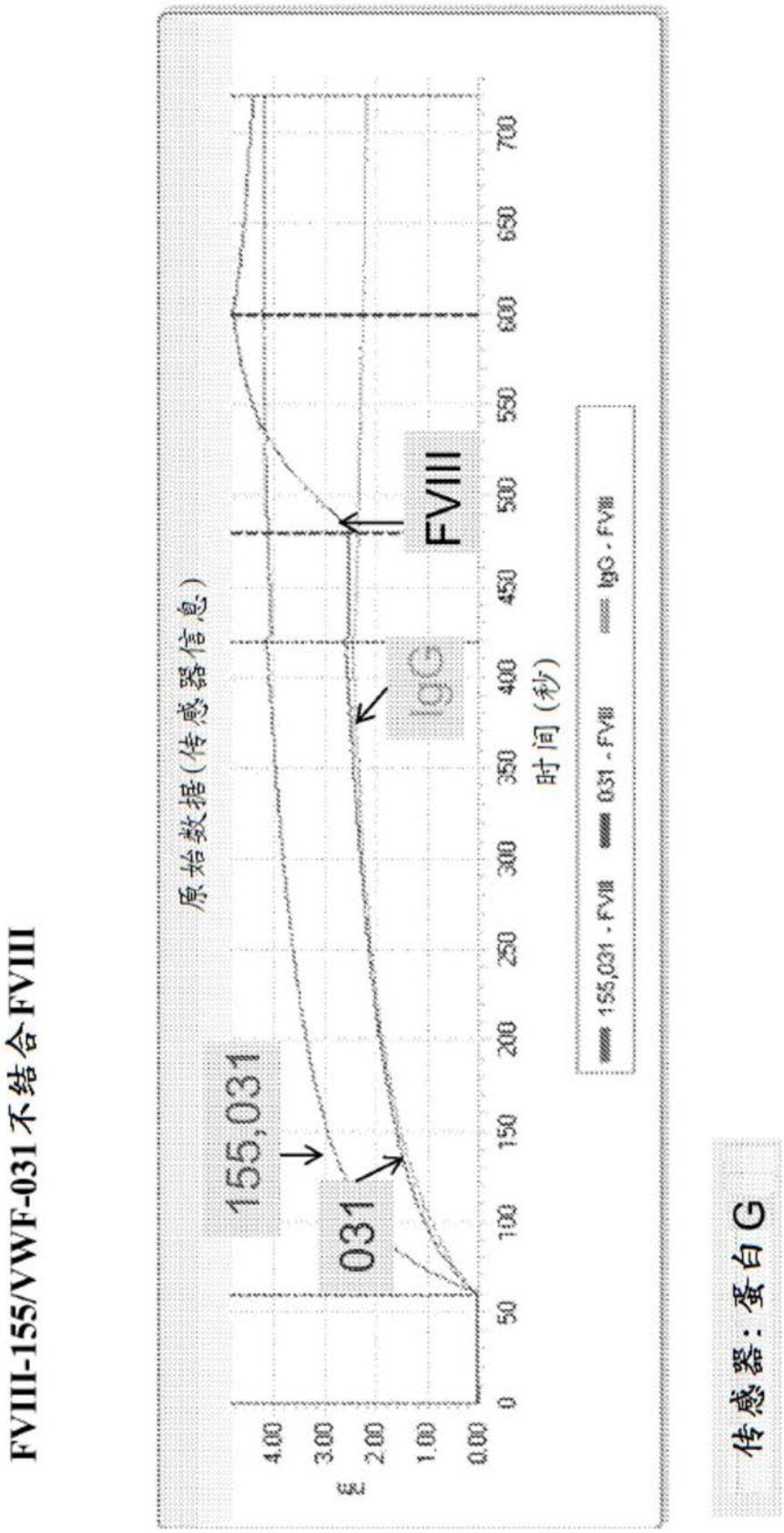

FVIII-155/VWF-031 不结合 FVIII
原始数据(传感器信息)
155,031
031
IgG
FVIII
RU
4.00
3.00
2.00
1.00
0.00
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
时间(秒)
155,031 - FVIII
031 - FVIII
IgG - FVIII
传感器:蛋白G

图21B

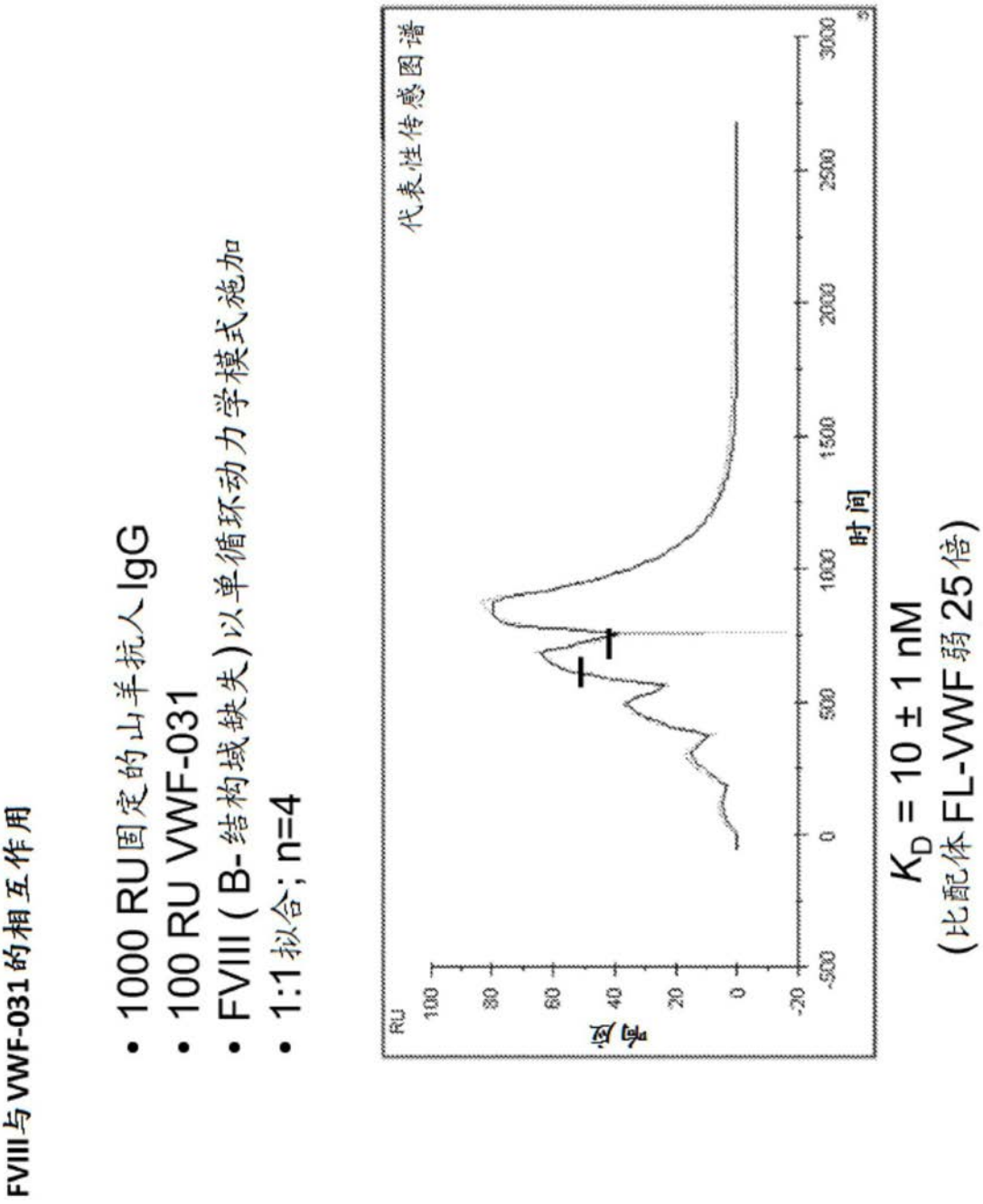

FVIII与VWF-031的相互作用
• 1000 RU固定的山羊抗人IgG
• 100 RU VWF-031
• FVIII（B-结构域缺失）以单循环动力学模式施加
• 1:1拟合；n=4
代表性传感图谱
RU
100
80
60
40
20
0
-20
-500
0
500
1000
1500
2000
2500
3000
时间
K_D = 10 ± 1 nM
（比配体FL-VWF弱25倍）

图22

FVIII-155	单链 (SC) - 直接融合至单链Fc的FVIII R1645A/R1648A
VWF-031	连接至单链Fc、具有48aa凝血酶可切割连接基的D1D2D'D3（具有C336A/C379A的1-477aa）片段
VWF-035	连接至单链Fc、具有73aa凝血酶可切割连接基的D1D2D'D3（具有C336A/C379A的1-477aa）片段
VWF-036	连接至单链Fc、具有98aa凝血酶可切割连接基的D1D2D'D3（具有C336A/C379A的1-477aa）片段

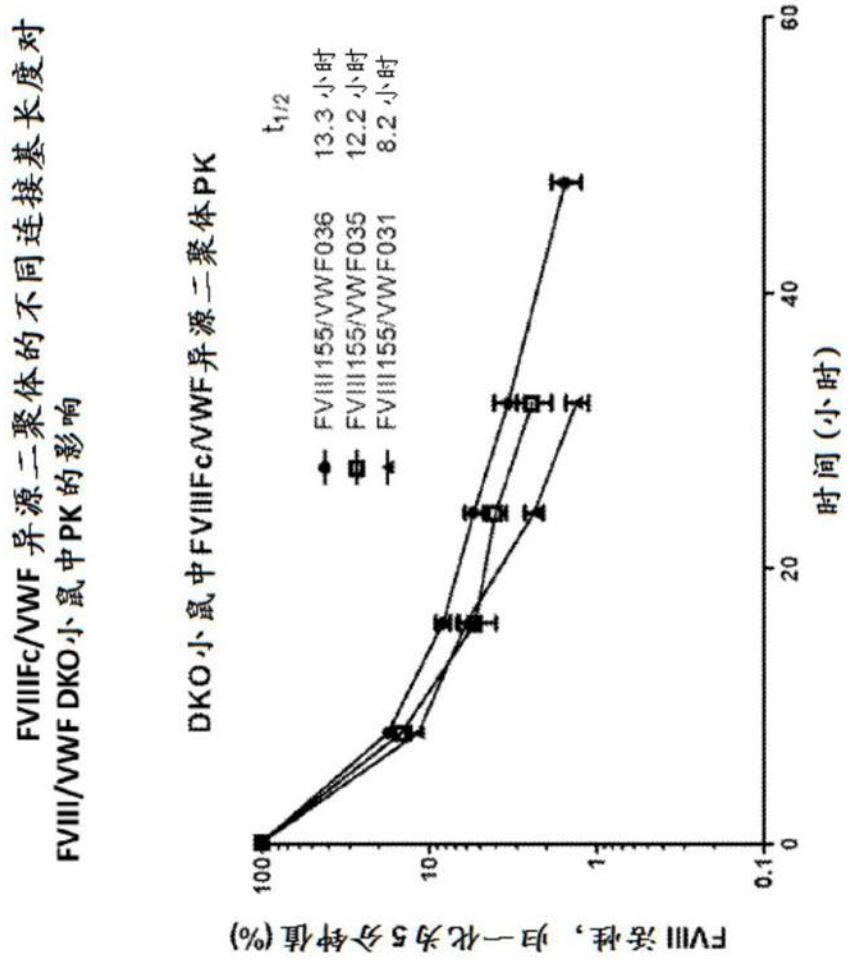


图23

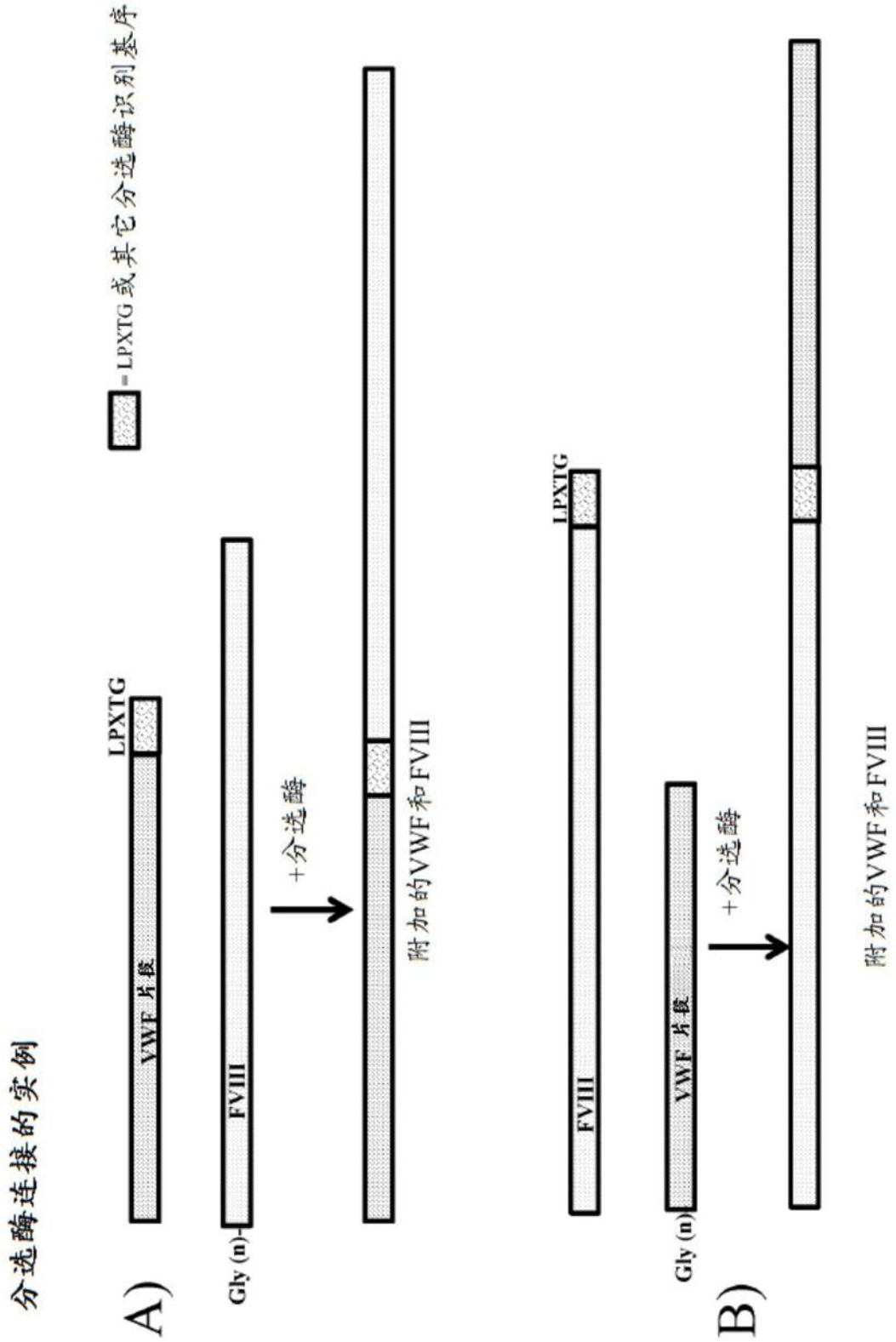

分选酶连接的实例
= LPXTG 或其它分选酶识别基序
A)
LPXTG
VWF 片段
FVIII
Gly (n)-
+分选酶
附加的VWF和FVIII
B)
LPXTG
FVIII
VWF 片段
Gly (n)
+分选酶
附加的VWF和FVIII

图24

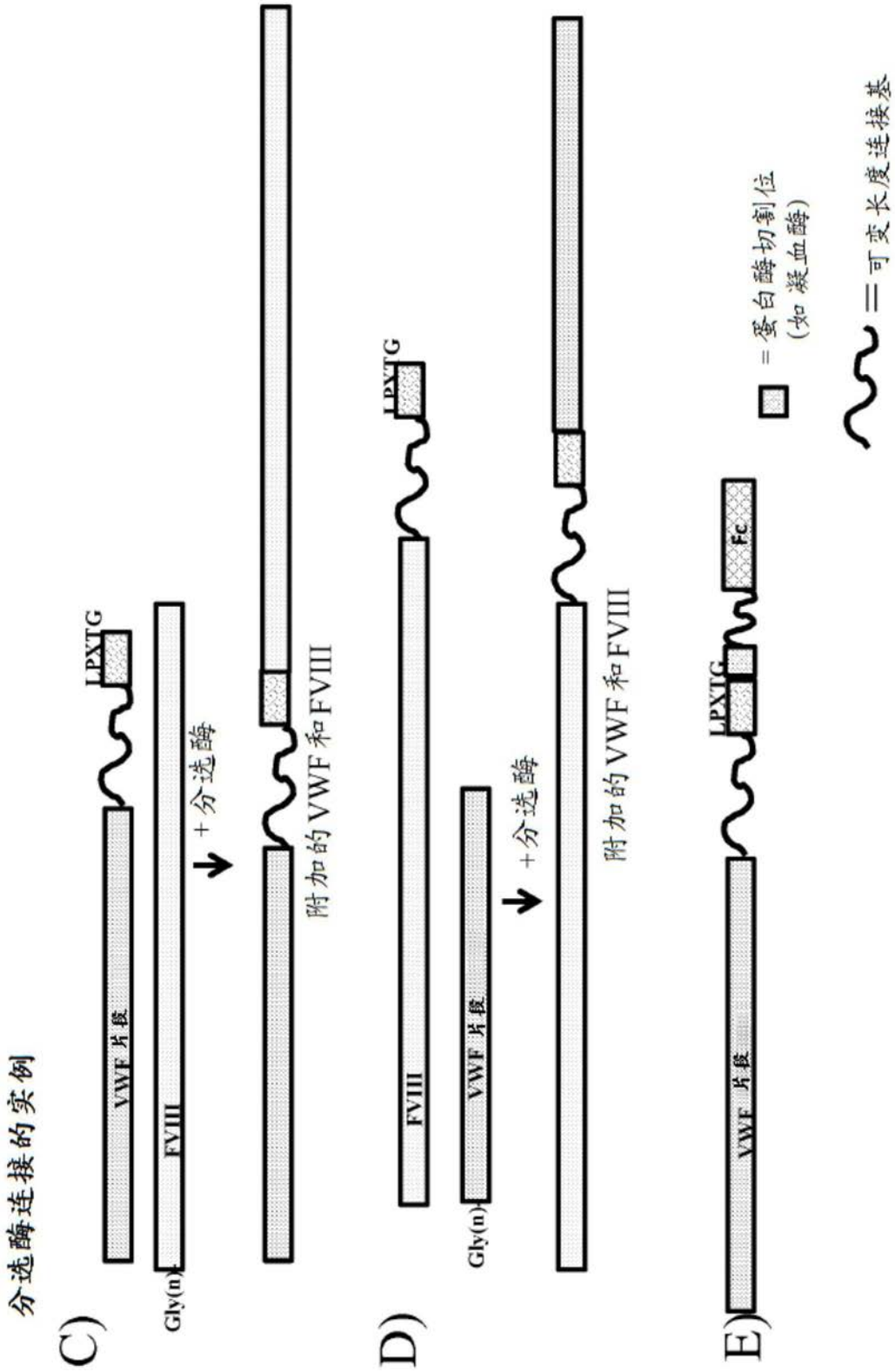

分选酶连接的实例
C)
VWF 片段
LPXTG
FVIII
Gly(n)
+分选酶
附加的 VWF 和 FVIII
D)
FVIII
LPXTG
VWF 片段
Gly(n)
+分选酶
附加的 VWF 和 FVIII
E)
VWF 片段
LPXTG
Fc
= 蛋白酶切割位
(如凝血酶)
= 可变长度连接基

图24(续)

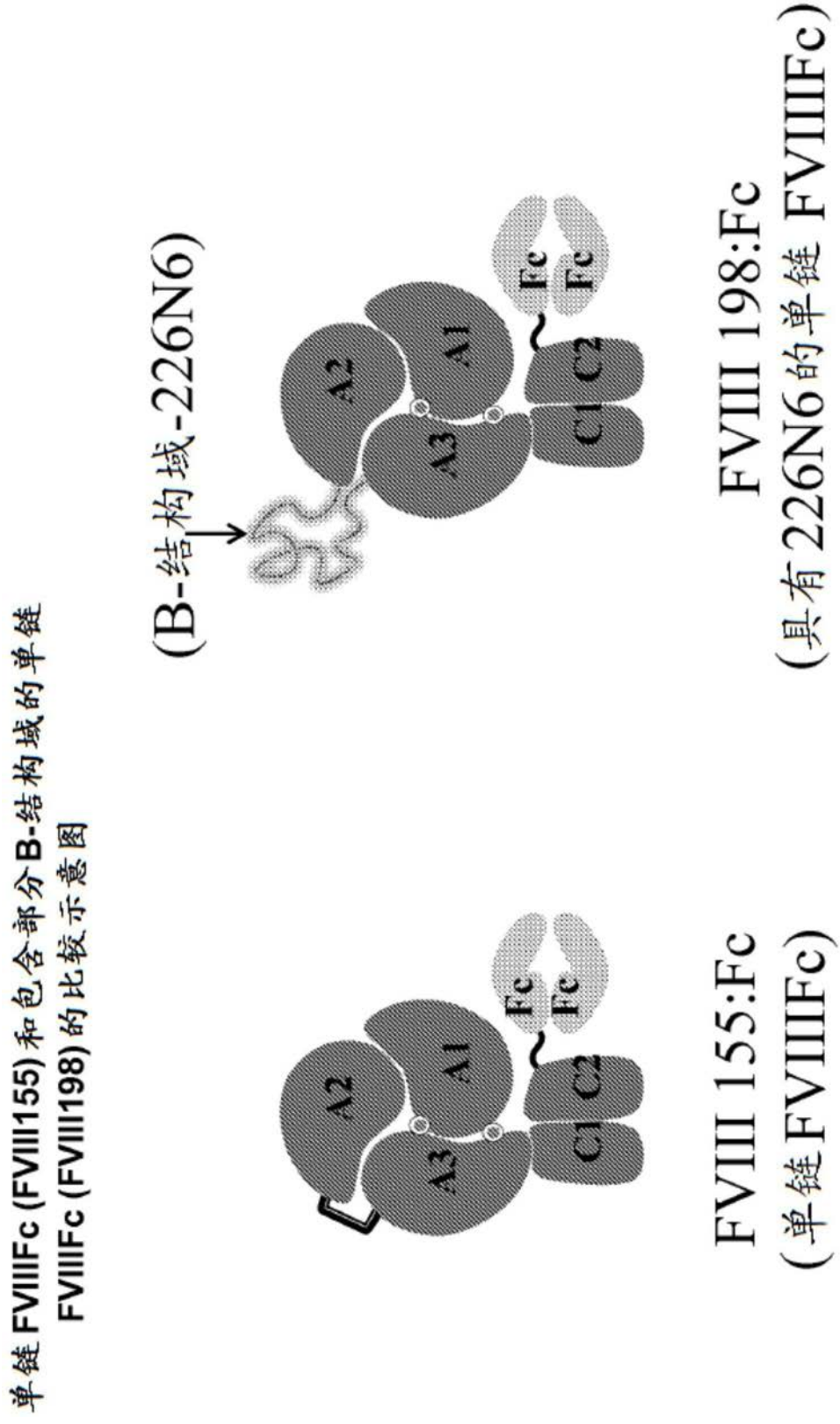

单链 FVIIIFc (FVIII155) 和包含部分 B-结构域的单链
FVIIIFc (FVIII198) 的比较示意图
(B-结构域-226N6)
A2
A1
A3
C1 C2
Fc
Fc
FVIII 155:Fc
(单链FVIIIFc)
FVIII 198:Fc
(具有226N6的单链 FVIIIFc)

图25

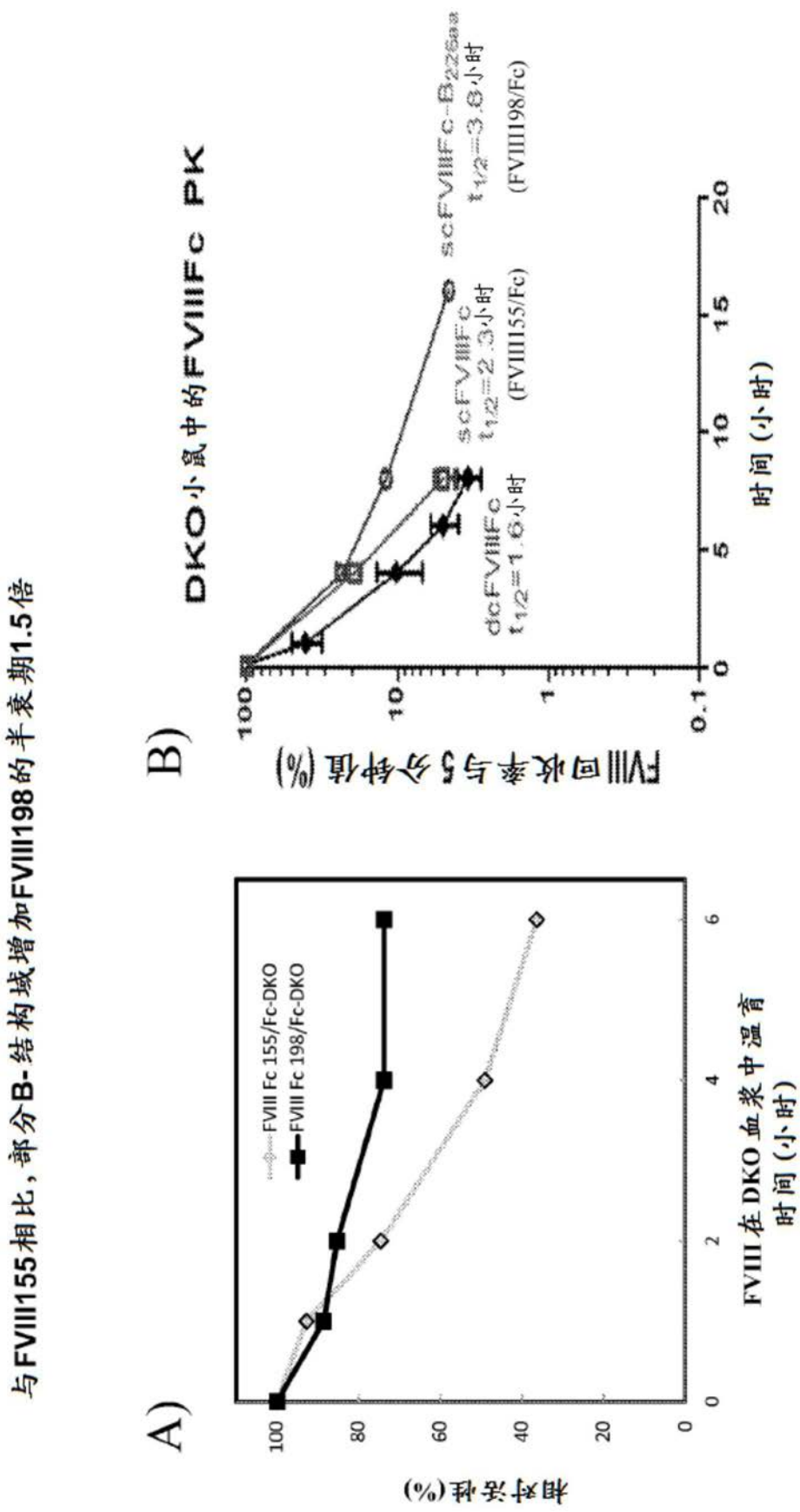

与FVIII155相比，部分B-结构域增加FVIII198的半衰期1.5倍
A)
相对活性(%)
FVIII Fc 155/Fc-DKO
FVIII Fc 198/Fc-DKO
FVIII在DKO血浆中温育
时间(小时)
B)
DKO小鼠中的FVIIIFc PK
FVIII回收率与5分钟值(%)
scFVIIIFc-B226aa
t1/2=3.8小时
(FVIII198/Fc)
scFVIIIFc
t1/2=2.3小时
(FVIII155/Fc)
dcFVIIIFc
t1/2=1.6小时
时间(小时)

图26