



(12)发明专利申请

(10)申请公布号 CN 110054699 A

(43)申请公布日 2019. 07. 26

(21)申请号 201811639711.8

约翰·库曼

(22)申请日 2013.07.10

(74)专利代理机构 上海胜康律师事务所 31263

代理人 樊英如 邱晓敏

(30)优先权数据

61/670,401 2012.07.11 US

61/759,819 2013.02.01 US

61/801,504 2013.03.15 US

61/801,544 2013.03.15 US

61/827,158 2013.05.24 US

61/840,811 2013.06.28 US

(51)Int.Cl.

C07K 19/00(2006.01)

A61K 38/37(2006.01)

A61K 47/68(2017.01)

A61P 7/04(2006.01)

(62)分案原申请数据

201380046914.5 2013.07.10

(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

常佩云 罗伯特·T·彼得斯

权利要求书3页 说明书164页

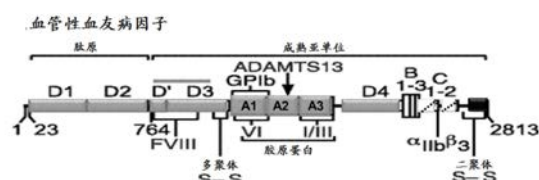
序列表240页 附图23页

(54)发明名称

具有XTEN和血管性血友病因子蛋白的因子VIII复合物、及其用途

(57)摘要

本发明包括一种嵌合蛋白,其包含含有VWF的D'结构域和D3结构域的VWF蛋白、一个或多个XTEN序列和FVIII蛋白,其中所述VWF片段、所述XTEN序列或所述FVIII蛋白彼此连接或缔合。所述嵌合蛋白还可包含一个或多个Ig恒定区或其一部分(例如,Fc区)。VWF片段的多肽链与连接于XTEN序列的FVIII多肽链缔合。所述VWF片段多肽链可阻止或抑制内源性VWF与连接于所述XTEN序列的FVIII蛋白结合。通过阻止或抑制内源性VWF与所述FVIII蛋白的结合,所述VWF片段可延长包含FVIII蛋白的所述嵌合蛋白的半衰期。本发明包括核苷酸、载体、宿主细胞、所述VWF片段的用途或所述嵌合蛋白。



➤ 约 250 kDa 的蛋白质,通过二硫键形成多聚体 (> 20 MDa)

➤ 在非共价复合物中与FVIII (95-98%)缔合

- 保护FVIII免遭蛋白酶切割/活化
- 稳定重链和轻链
- 阻止清除受体对FVIII的清除

延长半衰期

- 通过VWF受体对FVIII-vWF复合物的清除
- 阻止rFVIII-Fc的胞饮和再循环?

限制半衰期

1. 一种嵌合蛋白,其包含(i)第一多肽和(ii)第二多肽,所述第一多肽包含(1)血管性血友病因子(VWF)片段,(2)延长长度多肽(XTEN)和(3)第一Fc区,所述VWF片段包含VWF的D'结构域和D3结构域,且所述第二多肽包含(4)因子VIII(FVIII)蛋白和(5)第二Fc区,

其中所述XTEN包含与SEQ ID NO:141-169或170所示氨基酸序列具有至少80%同一性的氨基酸序列,

其中所述VWF片段不包含对应于SEQ ID NO:2的氨基酸1274至2813,和

其中所述第一Fc区通过二硫键与所述第二Fc区缔合。

2. 如权利要求1所述的嵌合蛋白,其中所述第一多肽和所述第二多肽不通过肽键连接。

3. 如权利要求1所述的嵌合蛋白,其中所述第一Fc区与所述XTEN连接。

4. 如权利要求3所述的嵌合蛋白,其中所述第一Fc区通过可切割接头与所述XTEN连接。

5. 如权利要求1所述的嵌合蛋白,其中所述第二Fc区通过可加工的接头进一步与所述第一Fc区连接。

6. 如权利要求1所述的嵌合蛋白,

其中所述第一多肽包含与SEQ ID NO:101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116或117所示氨基酸序列具有至少80%同一性的氨基酸序列;和

其中所述VWF片段包含与SEQ ID NO:118,119或120所示氨基酸序列具有至少80%同一性的氨基酸序列。

7. 如权利要求1所述的嵌合蛋白,其中所述FVIII蛋白包含B结构域或其部分。

8. 如权利要求7所述的嵌合蛋白,其中所述FVIII蛋白是缺失SQ B结构域的FVIII。

9. 如权利要求7所述的嵌合蛋白,其中额外的XTEN插入所述B结构域或其部分内。

10. 如权利要求1所述的嵌合蛋白,其中所述VWF片段包含与SEQ ID NO:119的氨基酸764-1240具有至少90%同一性的氨基酸序列。

11. 如权利要求10所述的嵌合蛋白,其中所述VWF片段基本上由SEQ ID NO:119的氨基酸764至1240组成。

12. 如权利要求1所述的嵌合蛋白,其中所述XTEN是与SEQ ID NO:39,SEQ ID NO:40,SEQ ID NO:41,SEQ ID NO:42,SEQ ID NO:43,或SEQ ID NO:44具有至少95%同一性的氨基酸序列。

13. 如权利要求1所述的嵌合蛋白,其中所述XTEN是与SEQ ID NO:39,SEQ ID NO:40,SEQ ID NO:41,SEQ ID NO:42,SEQ ID NO:43,或SEQ ID NO:44具有至少80%同一性的氨基酸序列。

14. 一种药物组合物,其包含如权利要求1所述的嵌合蛋白和药学上可接受的载体。

15. 如权利要求1所述的嵌合蛋白,其中所述FVIII蛋白包含(i)重链和轻链,其中所述FVIII蛋白的重链包含与SEQ ID NO:4的氨基酸1-740具有至少90%同一性的氨基酸序列,以及所述FVIII蛋白的轻链包含与SEQ ID NO:4的氨基酸1649-2332具有至少90%同一性的氨基酸序列。

16. 如权利要求15所述的嵌合蛋白,其中所述FVIII蛋白包含B结构域或其部分。

17. 如权利要求1所述的嵌合蛋白,其中所述第二多肽包含与SEQ ID NO:103所示氨基酸序列具有至少90%同一性的氨基酸序列。

18. 如权利要求1所述的嵌合蛋白,其中所述第二多肽包含与SEQ ID NO:103所示氨基

酸序列具有至少95%同一性的氨基酸序列。

19. 如权利要求1所述的嵌合蛋白,其中所述第二多肽包含SEQ ID NO:103所示的氨基酸序列。

20. 如权利要求1所述的嵌合蛋白,其中所述第二多肽包含与SEQ ID NO:103所示的氨基酸序列具有至少90%同一性的氨基酸序列,并且所述第一多肽包含与SEQ ID NO:119所示的氨基酸序列具有至少90%同一性的氨基酸序列。

21. 如权利要求1所述的嵌合蛋白,其中所述第二多肽包含与SEQ ID NO:103所示氨基酸序列具有至少95%同一性的氨基酸序列,并且所述第一多肽包含与SEQ ID NO:119所示的氨基酸序列具有至少95%同一性的氨基酸序列。

22. 如权利要求4所述的嵌合蛋白,其中所述可切割接头包含一个或多个可被蛋白酶切割的切割位点。

23. 如权利要求22所述的嵌合蛋白,其中所述蛋白酶是因子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,或MMP-20。

24. 如权利要求4所述的嵌合蛋白,其中所述可切割接头包含一个或多个凝血酶切割位点。

25. 如权利要求9所述的嵌合蛋白,其中所述额外的XTEN包含长度为288个氨基酸的氨基酸序列。

26. 如权利要求1所述的嵌合蛋白,其中所述XTEN包含与SEQ ID NO:145,146,147或148所示的氨基酸序列具有至少90%同一性的氨基酸序列。

27. 如权利要求11所述的嵌合蛋白,其中所述VWF片段由SEQ ID NO:119的氨基酸764至1240组成。

28. 如权利要求1所述的嵌合蛋白,其中所述XTEN包含SEQ ID NO:145,146,147或148所示的氨基酸序列。

29. 嵌合蛋白,其包含(i)第一多肽和(ii)第二多肽,所述第一多肽包含(1)VWF片段,(2)第一XTEN,(3)凝血酶可切割接头,和(4)第一Fc区,且所述第二多肽包含(5)FVIII蛋白,(6)第二XTEN,和(7)第二Fc区,

其中所述VWF片段由SEQ ID NO:119的氨基酸764至1240组成;

其中所述第一XTEN包含SEQ ID NO:145,146,147或148所示的氨基酸序列;

其中所述FVIII蛋白包含完全或部分缺失的所述B结构域;

其中所述第二XTEN包含SEQ ID NO:145,146,147或148所示的氨基酸序列,其中所述第二XTEN插入所述B结构域或其部分内,并且其中所述XTEN的长度为约36至约3000个氨基酸;和

其中所述第一Fc区和所述第二Fc区形成二硫键。

30. 如权利要求1所述的嵌合蛋白,其中所述XTEN的长度为约36至约3000个氨基酸。

31. 嵌合蛋白,其包含(i)第一多肽,(ii)第二多肽和(iii)第三多肽,

其中所述第一多肽包含(1)VWF片段,(2)第一XTEN,(3)凝血酶可切割接头,和(4)第一Fc区;

其中所述第二多肽包含(5)FVIII蛋白的重链和(6)第二XTEN;和

其中所述第三多肽包含(7)FVIII蛋白的轻链和(8)第二Fc区；

其中所述FVIII蛋白的重链和所述FVIII蛋白的轻链形成金属键；

其中所述第一XTEN包含与SEQ ID NO:141-169或170中所示的氨基酸序列具有至少80%同一性的氨基酸序列；

其中所述第二XTEN包含与SEQ ID NO:141-169或170中所示的氨基酸序列具有至少80%同一性的氨基酸序列；

其中所述VWF片段不包含对应于SEQ ID NO:2的氨基酸1274至2813；和

其中所述第一Fc区和所述第二Fc区形成二硫键。

32.如权利要求31所述的嵌合蛋白,其中所述第一XTEN和所述第二XTEN的长度为约36至约3000个氨基酸。

33.如权利要求1所述的嵌合蛋白,其中所述第一多肽中的所述VWF片段包含与SEQ ID NO:119的氨基酸764-1240具有至少90%同一性的氨基酸序列,并且所述第二多肽包含与SEQ ID NO:103中所示的氨基酸序列具有至少90%同一性的氨基酸序列。

34.如权利要求1所述的嵌合蛋白,其中所述第一多肽中的所述VWF片段包含与SEQ ID NO:119的氨基酸764-1240具有至少95%同一性的氨基酸序列,并且所述第二多肽包含与SEQ ID NO:103中所示的氨基酸序列具有至少95%同一性的氨基酸序列。

35.如权利要求31所述的嵌合蛋白,其中所述第一多肽中的所述VWF片段包含与SEQ ID NO:119的氨基酸764-1240具有至少90%同一性的氨基酸序列。

36.如权利要求31所述的嵌合蛋白,其中所述第一多肽中的所述VWF片段包含与SEQ ID NO:119的氨基酸764-1240具有至少95%同一性的氨基酸序列。

具有XTEN和血管性血友病因子蛋白的因子VIII复合物、及其用途

[0001] 本申请是申请日为2013年7月10日、中国专利申请号为201380046914.5 (国际申请号为PCT/US2013/049989)、发明名称为“具有XTEN和血管性血友病因子蛋白的因子VIII复合物、及其用途”的发明专利申请的分案申请。

[0002] 发明背景

[0003] 血友病A是一种由于编码凝血因子VIII (FVIII) 的基因缺陷所导致的出血性疾病并且影响着10,000个出生男孩中的1-2个。Graw等人, Nat. Rev. Genet. 6 (6) :488-501 (2005)。患有血友病A的患者可通过输注纯化的或重组生成的FVIII来治疗。但是, 已知所有商购可得的FVIII产品具有约8-12小时的半衰期, 从而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半衰期已尝试许多方法。例如, 延长凝血因子的半衰期的研发方法包括聚乙二醇化、糖基聚乙二醇化和与白蛋白的缀合。参见Dumont等人, Blood. 119 (13) :3024-3030 (2012年1月13日在线公布)。但是, 不管所用的蛋白质工程如何, 据报道目前正在开发的长效FVIII产品在临床前动物模型中具有仅约1.5至2小时的有限半衰期。参见同上。已在人中显示一致性结果, 例如, 据报道rFVIII-Fc在血友病A患者中相比于ADVATE®改善半衰期高达约1.7倍。参见同上。因此, 尽管改善较小, 但是半衰期增加可能指示其它T_{1/2}限制因素的存在。参见Liu, T. 等人, 2007 ISTH会议, 摘要#P-M-035; Henrik, A. 等人, 2011 ISTH会议, 摘要#P-M-181; Liu, T. 等人, 2011 ISTH会议摘要#P-WE-131。

[0004] 血浆血管性血友病因子 (von Willebrand Factor, VWF) 具有约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 (最近一次访问为2011年10月22日)。VWF半衰期可受许多因素影响: 糖基化模式、ADAMTS-13 (具有血小板反应蛋白基序-13的解联蛋白和金属蛋白酶) 和VWF中的各种突变。

[0005] 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。这种复合物的形成对于体内维持适当的FVIII血浆水平是重要的。Lenting等人, Blood. 92 (11) : 3983-96 (1998); Lenting等人, J. Thromb. Haemost. 5 (7) :1353-60 (2007)。全长野生型FVIII主要作为具有重链 (MW 200kD) 和轻链 (MW 73kD) 的异二聚体存在。当FVIII由于在重链的位置372和740处和在轻链的位置1689处的蛋白质水解而被活化, 结合于FVIII的VWF从活化FVIII去除。活化FVIII与活化因子IX、钙和磷脂 (“X酶复合物 (tenase complex)”) 一起诱导因子X的活化, 从而产生大量凝血酶。反过来, 凝血酶然后切割纤维蛋白原以形成可溶性纤维蛋白单体, 其然后自发地聚合以形成可溶性纤维蛋白聚合物。凝血酶还活化因子XIII, 该因子与钙一起用来交联并稳定可溶性纤维蛋白聚合物, 从而形成交联 (不可溶性) 纤维蛋白。活化FVIII通过蛋白质水解从循环被快速清除。

[0006] 由于频繁给药和由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倍半衰期限度的FVIII产品。

[0007] 发明简述

[0008] 本发明涉及一种嵌合蛋白,其包含(i)含有VWF的D'结构域和D3结构域的血管性血友病因子(VWF)片段、(ii)XTEN序列和(iii)FVIII蛋白,其中VWF片段和XTEN序列通过任选的接头连接并且其中VWF片段或XTEN序列连接于FVIII蛋白或与FVIII蛋白缔合。嵌合蛋白可包含含有VWF片段、XTEN序列和FVIII蛋白的单一多肽链、或两条多肽链,第一链包含VWF片段且第二链包含FVIII蛋白,其中XTEN多肽连接于VWF片段或FVIII蛋白。

[0009] 在一个实施方案中,本发明的嵌合蛋白包含包括以下的式:

[0010] (a) V-X-FVIII、

[0011] (b) FVIII-X-V、

[0012] (c) V-X:FVIII、

[0013] (d) X-V:FVIII、

[0014] (e) FVIII:V-X或

[0015] (f) FVIII:X-V,

[0016] 其中V包括VWF片段,

[0017] X包括一个或多个XTEN序列,且

[0018] FVIII包括FVIII蛋白。连字号(-)可以为肽键或接头例如可切割的接头,而冒号(:)表示多肽之间的化学缔合或物理缔合,例如共价或非共价键。

[0019] 在另一个实施方案中,嵌合蛋白还包含(iv)连接于VWF片段、XTEN序列、FVIII蛋白或其任意组合的免疫球蛋白(Ig)恒定区或其一部分(也表示为F1或第一Ig恒定区或其一部分)。在其它实施方案中,嵌合蛋白还包含另外的Ig恒定区或其一部分(也表示为F2或第二Ig恒定区或其一部分)。第一Ig恒定区或其一部分可连接于VWF片段或XTEN序列,并且第二Ig恒定区可连接于FVIII蛋白。第一Ig恒定区、第二Ig恒定区或其一部分或两者可延长FVIII蛋白的半衰期。

[0020] 在一些实施方案中,第二Ig恒定区或其一部分(F2)通过接头例如可加工的接头连接于VWF片段。在其它实施方案中,第二Ig恒定区或其一部分(F2)与(第一)Ig恒定区或其一部分(F1)缔合。第二Ig恒定区或其一部分(F2)和第一Ig恒定区或其一部分(F1)可以相同或不同。第二Ig恒定区或其一部分可通过共价键例如二硫键与Ig恒定区或其一部分缔合。连接于第一Ig恒定区或其一部分的VWF片段也可与通过非共价键连接于第二Fc区的FVIII蛋白缔合。在某些实施方案中,FVIII蛋白还可包含一个或多个另外的XTEN序列,所述另外的XTEN序列连接于FVIII蛋白的C-末端或N-末端或紧邻FVIII蛋白中的一个或多个氨基酸(例如,一个或多个XTEN插入位点)的下游插入。在一些实施方案中,FVIII蛋白的半衰期相较于不含VWF片段的野生型FVIII或FVIII蛋白有所延长。

[0021] 在一些实施方案中,嵌合蛋白包含包括以下的式:

[0022] (g) V-L2-X-L1-F1:FVIII-L3-F2;

[0023] (h) V-L2-X-L1-F1:F2-L3-FVIII;

[0024] (i) F1-L1-X-L2-V:FVIII-L3-F2;

[0025] (j) F1-L1-X-L2-V:F2-L3-FVIII;

[0026] (k) V-L2-X-L1-F1-L4-FVIII-L3-F2;

[0027] (1) F2-L3-FVIII-L4-F1-L1-X-L2-V;

[0028] (m) FVIII-L3-F2-L4-V-L2-X-L1-F1;或

[0029] (n) F1-L1-X-L2-V-L4-F2-L3-FVIII,

[0030] 其中V包括VWF片段,

[0031] L1、L2和L3中的每一个包括任选的接头例如可切割的接头,

[0032] L4为任选的接头,例如可加工的接头,

[0033] FVIII包括FVIII蛋白,

[0034] X包括一个或多个XTEN序列,

[0035] F1包括任选的第一Ig恒定区或其一部分,

[0036] F2包括任选的第二Ig恒定区或其一部分,且

[0037] (:) 为共价键或非共价键。

[0038] 本发明也涉及一种嵌合蛋白,其包含(i)FVIII蛋白、(ii)XTEN序列和(iii)Ig恒定区或其一部分,其中XTEN序列在FVIII蛋白的N-末端或C末端通过任选的接头连接于FVIII蛋白或紧邻FVIII蛋白中的一个或多个氨基酸(例如,一个或多个插入位点)的下游插入并且其中Ig恒定区或其一部分连接于FVIII蛋白或XTEN序列或与FVIII蛋白或XTEN序列缔合。在一个实施方案中,可用于嵌合蛋白的Ig恒定区或其一部分包含第一Fc区。在另一个实施方案中,嵌合蛋白还包含另外的Ig恒定区或其一部分。可用于本发明的另外的Ig恒定区或其一部分可包含第二Fc区,第二Fc区例如通过共价键连接于第一Fc区或与第一Fc区缔合。在其它实施方案中,第一Fc区通过接头例如可加工的接头连接于第二Fc区。

[0039] 在其它方面,嵌合蛋白包含(i)FVIII蛋白、(ii)XTEN序列、(iii)VWF片段和(iv)Ig恒定区或其一部分,其包含VWF的D'结构域和D3结构域,其中XTEN序列在FVIII蛋白的N-末端或C末端通过任选的接头连接于FVIII蛋白或紧邻FVIII蛋白中的一个或多个氨基酸(例如,一个或多个插入位点)的下游插入,VWF片段连接于FVIII蛋白或XTEN序列或与FVIII蛋白或XTEN序列缔合,并且Ig恒定区或其一部分连接于FVIII蛋白、XTEN序列、VWF片段或其任意组合。嵌合蛋白的非限制性实例可以包含包括以下的式:

[0040] (1) FVIII(X1)-L1-F1:V-L2-X2-L3-F2;

[0041] (2) FVIII(X1)-L1-F1:F2-L3-X2-L2-V;

[0042] (3) F1-L1-FVIII(X1):V-L2-X2-L3-F2;

[0043] (4) F1-L1-FVIII(X1):F2-L3-X2-L2-V;

[0044] (5) FVIII(X1)-L1-F1-L4-V-L2-X2-L3-F2;

[0045] (6) FVIII(X1)-L1-F1-L4-F2-L3-X2-L2-V;

[0046] (7) F1-L1-FVIII(X1)-L4-V-L2-X2-L3-F2或

[0047] (8) F1-L1-FVIII(X1)-L4-F2-L3-X2-L2-V,

[0048] 其中FVIII(X1)包括FVIII蛋白和一个或多个XTEN序列,其中XTEN序列中的一个或多个连接于FVIII蛋白的N-末端或C-末端或紧邻FVIII蛋白中的一个或多个氨基酸(例如,一个或多个XTEN插入位点)的下游插入;

[0049] L1、L2或L3中的每一个包括任选的接头,例如可切割的接头;

[0050] L4为接头、可加工的接头;

[0051] X2包括一个或多个XTEN序列;

[0052] F1包括Ig恒定区或其一部分；

[0053] F2包括任选的另外的Ig恒定区或其一部分，且

[0054] V包括VWF片段；

[0055] (-) 为肽键或一个或多个氨基酸；且

[0056] (:) 包括共价键或非共价键。

[0057] 本发明的一个方面是可用于嵌合蛋白的VWF片段未结合于VWF清除受体，其阻止或抑制FVIII蛋白与内源性VWF的相互作用。包含VWF片段的嵌合蛋白由此具有减少的清除率或未通过VWF清除途径被清除。本发明的另一方面是VWF片段能够保护FVIII蛋白免遭一个或多个蛋白酶切割，保护FVIII蛋白免遭活化，稳定FVIII蛋白的重链和/或轻链，或阻止FVIII蛋白被一个或多个清除剂受体清除。

[0058] 因为VWF片段能够阻止或抑制FVIII蛋白与内源性VWF之间的相互作用，所以FVIII蛋白的半衰期相较于不含VWF片段的FVIII蛋白有所延长。在一个实施方案中，FVIII蛋白的半衰期比野生型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。在另一个实施方案中，FVIII蛋白的半衰期为至少约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。

[0059] 可用于嵌合蛋白的Ig恒定区或其一部分包含第一Fc区，第一Fc区通过任选的接头例如可切割的接头连接于VWF片段。嵌合蛋白还可包含另外的Ig恒定区或其一部分，其通过任选的接头连接于FVIII蛋白或XTEN序列、Ig恒定区或其一部分、VWF片段或其任意组合。在一个实施方案中，另外的Ig恒定区或其一部分通过任选的接头连接于FVIII蛋白。另外的Ig恒定区或其一部分可包含第二Fc区。

[0060] 可用于本发明的Ig恒定区或其一部分和可用于本发明的另外的Ig恒定区或其一部分相同或不同。

[0061] 在一些方面，FVIII蛋白在FVIII蛋白的C-末端或N-末端连接于XTEN序列或紧邻成熟的天然人FVIII中的一个或多个氨基酸(例如，一个或多个插入位点)的下游插入，或其任意组合。FVIII蛋白中的一个或多个插入位点可位于FVIII蛋白的一个或多个选自由以下组成的组的结构域内：A1结构域、a1酸性区域、A2结构域、a2酸性区域、A3结构域、B结构域、C1结构域、C2结构域及其任意组合，或可位于FVIII蛋白的一个或多个选自由以下组成的组的结构域之间：A1结构域和a1酸性区域、a1酸性区域和A2结构域、A2结构域和a2酸性区域、a2酸性区域和B结构域、B结构域和A3结构域、A3结构域和C1结构域、C1结构域和C2结构域及其任意组合，或可位于FVIII蛋白的两个选自由以下组成的组的结构域之间：A1结构域和a1酸性区域、a1酸性区域和A2结构域、A2结构域和a2酸性区域、a2酸性区域和B结构域、B结构域和A3结构域、A3结构域和C1结构域、C1结构域和C2结构域及其任意组合。

[0062] 在一个实施方案中，一个或多个插入位点位于紧邻成熟的天然人FVIII(例如，SEQ ID NO:4[成熟FVIII序列-全长])中的一个或多个选自由以下组成的组的氨基酸的下游：表7、8、9、10、11中的氨基酸残基或其任意组合。

[0063] 在另一个实施方案中,一个或多个插入位点位于成熟的天然人FVIII的一个或多个受纳环(permissive loop)内。在其它实施方案中,一个或多个插入位点位于成熟的天然人FVIII的a3区域内。例如,XTEN序列可紧邻对应于SEQ ID NO:4(全长成熟FVIII)的氨基酸1656的下游插入。在其它实施方案中,FVIII蛋白连接于至少两个XTEN序列,在a3区域内插入的第一XTEN序列,和在FVIII蛋白的受纳环(例如,A1-1、A1-2、A2-1、A2-2、A3-1或A3-2)内插入的第二XTEN序列。在又其它实施方案中,FVIII蛋白连接于至少三个XTEN序列,在a3区域内插入的第一XTEN序列,和在FVIII蛋白中的一个或两个受纳环(例如,A1-1、A1-2、A2-1、A2-2、A3-1或A3-2)内插入第二XTEN序列和第三XTEN序列。

[0064] 在某些实施方案中,供一个或多个XTEN插入的一个或多个插入位点紧邻一个或多个选自由以下组成的组的氨基酸(相对于成熟FVIII序列编号)的下游:

- [0065] (1) 氨基酸3、(2) 氨基酸18、(3) 氨基酸22、
[0066] (4) 氨基酸26、(5) 氨基酸32、(6) 氨基酸40、
[0067] (7) 氨基酸60、(8) 氨基酸65、(9) 氨基酸81、
[0068] (10) 氨基酸116、(11) 氨基酸119、(12) 氨基酸130、
[0069] (13) 氨基酸188、(14) 氨基酸211、(15) 氨基酸216、
[0070] (16) 氨基酸220、(17) 氨基酸224、(18) 氨基酸230、
[0071] (19) 氨基酸333、(20) 氨基酸336、(21) 氨基酸339、
[0072] (22) 氨基酸375、(23) 氨基酸399、(24) 氨基酸403、
[0073] (25) 氨基酸409、(26) 氨基酸416、(26) 氨基酸442、
[0074] (28) 氨基酸487、(29) 氨基酸490、(30) 氨基酸494、
[0075] (31) 氨基酸500、(32) 氨基酸518、(33) 氨基酸599、
[0076] (34) 氨基酸603、(35) 氨基酸713、(36) 氨基酸745、
[0077] (37) 氨基酸1656、(38) 氨基酸1711、(39) 氨基酸1720、
[0078] (40) 氨基酸1725、(41) 氨基酸1749、(42) 氨基酸1796、
[0079] (43) 氨基酸1802、(44) 氨基酸1827、(45) 氨基酸1861、
[0080] (46) 氨基酸1896、(47) 氨基酸1900、(48) 氨基酸1904、
[0081] (49) 氨基酸1905、(50) 氨基酸1910、(51) 氨基酸1937、
[0082] (52) 氨基酸2019、(53) 氨基酸2068、(54) 氨基酸2111、
[0083] (55) 氨基酸2120、(56) 氨基酸2171、(57) 氨基酸2188、
[0084] (58) 氨基酸2227、(59) 氨基酸2277和
[0085] (60) 其两个或更多的组合。

[0086] 在一些实施方案中,一个XTEN插入到FVIII蛋白内。在一些实施方案中,两个XTEN插入到FVIII蛋白内。在一些实施方案中,3个XTEN插入到FVIII蛋白内。

[0087] 在一个特定实例中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,并且第二XTEN紧邻对应于SEQ ID NO:4(全长成熟FVIII)的氨基酸1720的下游插入。在另一实例中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸403的下游插入,并且第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入。在一些实例中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入,并且第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入。在其它实例中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,第二XTEN紧

邻对应于SEQ ID NO:4的氨基酸1656的下游插入,并且第三XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入。在另外其它实施方案中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸403的下游插入,第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入,并且第三XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入。在又其它实施方案中,第一XTEN在对应于SEQ ID NO:4的氨基酸403与404之间插入,第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入,并且第三XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入。在某些实施方案中,第一XTEN紧邻对应于SEQ ID NO:4(全长成熟FVIII)的氨基酸26的下游插入,第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入,并且第三XTEN紧邻对应于SEQ ID NO:4的氨基酸1900的下游插入。在一些实施方案中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,第二XTEN紧邻对应于SEQ ID NO:2的氨基酸1656的下游插入,第三XTEN紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入,并且第四XTEN紧邻对应于SEQ ID NO:4的氨基酸1900的下游插入。在另一实例中,XTEN紧邻对应于SEQ ID NO:4的氨基酸745的下游插入。在另外的实例中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入并且第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1900的下游插入。在一些实施方案中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入,并且第三XTEN紧邻对应于SEQ ID NO:4的氨基酸1900的下游插入。在另一实例中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸403的下游插入并且第二XTEN紧邻对应于SEQ ID NO:4的氨基酸745的下游插入。在一些实施方案中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸745的下游插入,并且第二XTEN紧邻对应于SEQ ID NO:4的氨基酸1900的下游插入。在一些实施方案中,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸18的下游插入,并且第二XTEN紧邻对应于SEQ ID NO:4的氨基酸745的下游插入。

[0088] 在一些实施方案中,FVIII蛋白为双链FVIII同工型。在一些实施方案中,FVIII蛋白为单链FVIII同工型。

[0089] 在一些实施方案中,插入的XTEN为SEQ ID NO:39(AE288)。在一些实例中,插入的XTEN为SEQ ID NO:38和37(AG144和AE144)。在一些实例中,插入的XTEN为SEQ ID NO:37、38和37(AE144、AG144和AE144)。在一些实施方案中,插入的XTEN为SEQ ID NO:37和40(AE144和AE288)。在一些实施方案中,插入的XTEN为AE42(SEQ ID NO:36)、AE72(SEQ ID NO:127)、AE144_2A(SEQ ID NO:128)、AE144_3B(SEQ ID NO:129)、AE144_4A(SEQ ID NO:130)、AE144_5A(SEQ ID NO:131)、AE144_6B(SEQ ID NO:132)、AG144_A(SEQ ID NO:133)、AG144_B(SEQ ID NO:134)、AG144_C(SEQ ID NO:135)、AG144_F(SEQ ID NO:136)、AE864(SEQ ID NO:43)、AE576(SEQ ID NO:41)、AE288(SEQ ID NO:39)、AE288_2(SEQ ID NO:137)、AE144(SEQ ID NO:37)、AG864(SEQ ID NO:44)、AG576(SEQ ID NO:42)、AG288(SEQ ID NO:40)、AG144(SEQ ID NO:38)及其任意组合。

[0090] 可用于本发明的FVIII蛋白可包含B结构域或其一部分,例如,SQ B结构域缺失的FVIII。在一个实施方案中,FVIII蛋白包括单链FVIII。在另一个实施方案中,单链FVIII含有对应于全长成熟因子VIII多肽(SEQ ID NO:4)的残基1648、残基1645或两者或SQ BDD因子VIII(SEQ ID NO:6)的残基754、残基751或两者的残基处的至少一个氨基酸替代。在其它实施方案中,氨基酸替代是除精氨酸之外的氨基酸。在一些实施方案中,FVIII蛋白包含FVIII的重链和FVIII的轻链,其中重链和轻链通过金属键彼此缔合。

[0091] FVIII蛋白可例如通过含有降低与LRP的亲合力或消除与LRP的结合的至少一个氨基酸替代而具有与低密度脂蛋白受体相关蛋白 (LRP) 的低亲合力或未结合于低密度脂蛋白受体相关蛋白 (LRP)。此类至少一个氨基酸替代可以在对应于全长成熟FVIII的残基471、残基484、残基487、残基490、残基497、残基2092、残基2093或其两个或更多的组合的残基处。在特定实施方案中,在残基471、484或497处的氨基酸替代是除精氨酸之外的氨基酸,在残基487处的氨基酸替代是除酪氨酸之外的氨基酸,在残基2092处的氨基酸替代是除赖氨酸之外的氨基酸,或在残基2093处的氨基酸替代是除苯丙氨酸之外的氨基酸。

[0092] 在一些实施方案中,FVIII蛋白含有至少一个氨基酸替代,其诱导FVIII蛋白比没有替代的FVIII蛋白更稳定。此类替代可位于FVIII蛋白的A2结构域和A3结构域中,例如,在对应于全长成熟FVIII的残基664、残基1826、残基662、残基1828或其两个或更多的组合的残基处。

[0093] 可用于本发明的VWF片段包含D'结构域和D3结构域,其一起能够结合于FVIII。VWF片段可包含与SEQ ID NO:2的氨基酸764至866具有至少90%、95%、96%、97%、98%、99%或100%同一性的D'结构域的氨基酸序列和/或与SEQ ID NO:2的氨基酸867至1240具有至少90%、95%、96%、97%、98%、99%或100%同一性的D3结构域的氨基酸序列。在一个实施方案中,VWF片段为单体。在另一个实施方案中,VWF片段包括至少两个VWF片段、至少三个VWF片段、至少四个VWF片段、至少五个VWF片段或至少六个VWF片段。在一个实施方案中,两个或更多个VWF片段可以相同或它们可以不同。VWF片段可包含与SEQ ID NO:2的氨基酸764至1240具有至少90%、95%、96%、97%、98%、99%或100%同一性的氨基酸。VWF片段可基本上由以下组成或由以下组成:SEQ ID NO:2的氨基酸764至1240。在某些实施方案中,VWF片段可含有在对应于SEQ ID NO:2的残基1099、残基1142、或残基1099和1142两者的残基处的至少一个氨基酸替代。在其它实施方案中,VWF片段还包含VWF的D1结构域、D2结构域或D1和D2结构域。

[0094] VWF片段还可以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、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、其一个或多个片段及其任意组合。例如,VWF片段可基本上由以下组成或由以下组成:(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。在一些实施方案中,VWF片段还包含可操作地连接于VWF片段的VWF或FVIII的信号肽。

[0095] 可用于本发明的一个或多个接头具有至少约10、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1200、1400、1600、1800或2000个氨基酸的长度。在一些实施方案中,一个或多个接头具有约1至约2000个氨基酸的长度。在一个实施方案中,一个或多个接头具有至少约20、35、42、48、73、75、95、98、144、288、324、333、576或864个氨基酸的长度。在另一个实施方案中,一个或多个接头包含gly/ser肽、XTEN序列或两者。gly/ser肽的实例包括但不限于,式 $(\text{Gly}_4\text{Ser})_n$ (SEQ ID NO:139) 或 $\text{S}(\text{Gly}_4\text{Ser})_n$ (SEQ ID NO:140),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和10。例如, $(\text{Gly}_4\text{Ser})_n$ 接头可以为 $(\text{Gly}_4\text{Ser})_3$ (SEQ ID NO:63) 或 $(\text{Gly}_4\text{Ser})_4$ (SEQ ID NO:138)。在一个实施方案中,接头包含在接头的N-末端的至少一个第一切割位点,在接头的C-末端的至少

一个第二切割位点或两者。在另一个实施方案中,接头包括20个氨基酸、35个氨基酸、48个氨基酸、73个氨基酸或95个氨基酸的凝血酶可切割的接头。可切割的接头可包含被选自由以下组成的组的蛋白酶切割的一个或多个切割位点: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、颗粒酶-B、TEV、肠激酶、蛋白酶3C、分选酶A、MMP-12、MMP-13、MMP-17和MMP-20,例如,TLDPRSFLLRNPNDKYEPFWEDEEK (SEQ ID NO:8)。一个或多个切割位点的非限制性实例包括选自由以下组成的组的氨基酸序列:RRRR (SEQ ID NO:9)、RKRRKR (SEQ ID NO:10)、RRRRS (SEQ ID NO:11)、TQSFNDFTR (SEQ ID NO:12)、SVSQTSKLTR (SEQ ID NO:13)、DFLAEGGGVR (SEQ ID NO:14)、TTKIKPR (SEQ ID NO:15)、LVPRG (SEQ ID NO:16)、ALRPR (SEQ ID NO:17)、KLTRAET (SEQ ID NO:18)、DFTRVVG (SEQ ID NO:19)、TMTRIVGG (SEQ ID NO:20)、SPFRSTGG (SEQ ID NO:21)、LQVRIVGG (SEQ ID NO:22)、PLGRIVGG (SEQ ID NO:23)、IEGRTVGG (SEQ ID NO:24)、LTPRSLV (SEQ ID NO:25)、LGPVSGVP (SEQ ID NO:26)、VAGDSLEE (SEQ ID NO:27)、GPAGLGGA (SEQ ID NO:28)、GPAGLRGA (SEQ ID NO:29)、APLGLRLR (SEQ ID NO:30)、PALPLVAQ (SEQ ID NO:31)、ENLYFQG (SEQ ID NO:32)、DDDKIVGG (SEQ ID NO:33)、LEVLFQGP (SEQ ID NO:34)和LPKTGSES (SEQ ID NO:35)。在一些实施方案中,第一切割位点和第二切割位点相同或不同。

[0096] 可用于本发明的XTEN序列可以选自由以下组成的组:AE42 (SEQ ID NO:36)、AE144 (SEQ ID NO:37)、AG144 (SEQ ID NO:38)、AE288 (SEQ ID NO:39)、AG288 (SEQ ID NO:40)、AE576 (SEQ ID NO:41)、AG576 (SEQ ID NO:42)、AE864 (SEQ ID NO:43)、AE72 (SEQ ID NO:127)、AE144_2A (SEQ ID NO:128)、AE144_3B (SEQ ID NO:129)、AE144_4A (SEQ ID NO:130)、AE144_5A (SEQ ID NO:131)、AE144_6B (SEQ ID NO:132)、AG144_A (SEQ ID NO:133)、AG144_B (SEQ ID NO:134)、AG144_C (SEQ ID NO:135)、AG144_F (SEQ ID NO:136)、AE288_2 (SEQ ID NO:137)或AG864 (SEQ ID NO:44)。在特定实施方案中,XTEN序列包括AE288或AG288。

[0097] 本发明的嵌合蛋白可以是聚唾液酸化的、聚乙二醇化的或羟乙基淀粉化的。

[0098] 本发明也涉及编码嵌合蛋白的多核苷酸或一组多核苷酸。多核苷酸还可包括编码PC5或PC7的多核苷酸链。本发明也涉及载体包含多核苷酸或一组多核苷酸和可操作地连接于多核苷酸或一组多核苷酸的一个或多个启动子的载体。载体还可包括另外的载体,其包含编码PC5或PC7的多核苷酸链。本发明也涉及包含多核苷酸或载体的宿主细胞。宿主细胞可以为哺乳动物细胞,例如,HEK293细胞、CHO细胞或BHK细胞。在一些实施方案中,宿主细胞的PC5或PC7切割VWF的D1D2结构域。

[0099] 本发明还涉及包含嵌合蛋白、多核苷酸、载体或宿主细胞和药学上可接受的载体的药物组合物。本发明的组合物因而具有相较于野生型FVIII蛋白延长的半衰期。FVIII蛋白的半衰期比野生型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。因子VIII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。

[0100] 本发明的组合物可通过选自由以下组成的组的途径来施用:局部施用、眼内施用、

胃肠外施用、鞘内施用、硬膜下施用和口服施用。在一个实施方案中,组合物经由胃肠外施用,例如,静脉内或皮下施用来施用。本发明的组合物可用于治疗需要其的受试者中的出血性疾病或病患。出血性疾病或病患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、胸腔内溢血、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。在一个实施方案中,用嵌合蛋白治疗的受试者被安排进行手术。在另一个实施方案中,治疗为防治性的或按需治疗。

[0101] 本发明也涉及阻止或抑制FVIII蛋白与内源性VWF结合的方法,所述方法包括向需要其的受试者添加有效量的嵌合蛋白、多核苷酸载体、宿主细胞或组合物,其中VWF片段结合于FVIII蛋白并且因而阻止或抑制内源性VWF的结合。本发明还涉及延长或增加FVIII蛋白的半衰期的方法,其中所述方法包括向需要其的受试者施用有效量的嵌合蛋白、多核苷酸、载体、宿主细胞或组合物,其中VWF片段结合于FVIII蛋白并且因而延长或增加FVIII蛋白的半衰期。还提供阻止或抑制FVIII蛋白从细胞的清除的方法,其中所述方法包括向包含FVIII蛋白或编码FVIII蛋白的多核苷酸的细胞施用有效量的嵌合蛋白、多核苷酸、载体、宿主细胞或组合物,其中所述蛋白质具有结合于FVIII蛋白的VWF活性。可用于本方法的受试者是动物,例如,人,例如罹患血友病A的患者。

[0102] 本发明也提供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嵌合蛋白、多核苷酸、载体、宿主细胞或组合物,其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、胸腔内溢血、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。治疗可以为防治性的或按需治疗。在一个实施方案中,有效量为0.1 μ g/kg至500mg/kg。

[0103] 本发明也包括制备嵌合蛋白的方法,其包括用多核苷酸或载体转染一个或多个宿主细胞以及在宿主细胞中表达嵌合蛋白。

[0104] 附图简述

[0105] 图1A-D.VWF片段的示意图。图1A示出可用于本发明的三个示例性VWF片段,例如,VWF-002、VWF-010和VWF-013。VWF-002含有SEQ ID NO:124的氨基酸1至477(SEQ ID NO:2的氨基酸764至1240)并且在没有前肽/肽原序列的情况下合成。VWF-010除了D'D3结构域之外还含有D1D2结构域。VWF-013除了在SEQ ID NO:123的残基336和379处替代半胱氨酸的丙氨酸残基之外还含有D1D2D'D3结构域。图1B示出VWF-031,其含有通过可切割的接头例如48个氨基酸的凝血酶可切割的接头融合于Ig恒定区或其一部分例如Fc区的D1D2D'D3结构域。图1C示出VWF-025(其为编码在pLIVE载体中含有的D1D2D'D3结构域的核苷酸序列),和VWF-029(其为编码在pLIVE载体中具有两个氨基酸替代C336A和C379A的D1D2D'D3结构域的核苷酸序列)。图1D示出包含肽原(D1和D2结构域)和成熟亚单位(D',D3、A1、A2、A3、D4、B1-3、C1-2结构域)的全长VWF片段。VWF片段为约250kDa的蛋白质并且通过二硫键形成多聚体(>20MDa)。VWF片段与非共价复合物中的FVIII(95-98%)缔合,然后通过保护FVIII免遭蛋白酶切割/活化,稳定重链和轻链以及阻止清除剂受体对FVIII的清除来延长FVIII的半衰期。VWF片段也可通过VWF受体对FVIII-VWF复合物的清除以及阻止rFVIIIIFc的胞饮和再循环来限制FVIII的半衰期。

[0106] 图2.在VWF D'D3表达小鼠中或在FVIII和VWF双敲除(DKO)小鼠中rFVIII-XTEN(rFVIII-AE288或rFVIII-288AE)的药动学曲线。图2A示出编码质粒DNA(VWF-025)的D'D3结构域的流体动力学注射(HDI)(第-5天)、rFVIII-XTEN AE288的静脉内给药(第0天)和PK样本收集(第5天)的时间轴。图2B示出于D1D2D'D3小鼠中静脉内(IV)给药rFVIII-XTEN288(倒三角形)和于DKO小鼠中静脉内给药rFVIII-XTEN288(菱形)之后通过FVIII显色测定测量的FVIII活性。图2C示出在施用VWF-025后的D'D3血浆水平(ng/mL)。X轴表示以小时计的时间。

[0107] 图3.示例性的VWF:FVIII异二聚体构建体的示意图。构建体具有以式FVIII-F1-L1-V-X-L2-F2表示的共同结构,但含有不同可变接头的实例。示出的构建体(FVIII-161)含有连接于第一Fc区和VWF片段的异二聚体FVIII(重链和轻链通过金属键缔合),VWF片段为连接于XTEN序列的VWF的D'和D3结构域(即,具有氨基酸替代C336A和C379A的SEQ ID NO:2的氨基酸1至477),XTEN序列进一步连接于可切割的接头和第二Fc区。FVIII-161中含有的XTEN序列是XTEN AE288序列,并且接头为凝血酶可切割的接头,其具有35个氨基酸。在FVIII-161中,连接于第一Fc区的FVIII蛋白通过可加工的接头连接于VWF片段。在表达时,可加工的接头可通过胞内加工酶切割,因而制备彼此缔合的三个多肽链的构建体。

[0108] 图4为FVIII-VWF异二聚体或单体实例的示意图。FVIII-168、FVIII-175、FVIII-172、FVIII-174和FVIII170。构建体FVIII-168包含连接于第一Fc区的单链FVIII序列(具有在残基1645和1648处替代精氨酸残基的丙氨酸残基),第一Fc区然后融合于通过凝血酶可切割的接头连接于第二Fc区的VWF片段,所述接头具有48个氨基酸。AE288XTEN插入到单链FVIII序列的B结构域中。第一Fc区与VWF片段之间的连接包含能够通过胞内加工酶,即可加工的接头切割的接头。构建体FVIII-175包含连接于AE288XTEN和第一Fc区的单链FVIII(具有在残基1645和1648处替代精氨酸残基的丙氨酸残基),第一Fc区通过接头例如可加工的接头连接于第二Fc区。AE288XTEN插入到单链FVIII序列的B结构域中。构建体FVIII-172包含两条多肽链,第一链包含融合于AE288XTEN的重链FVIII序列,第二链包含轻链FVIII序列、第一Fc区、接头(例如,可加工的接头)、VWF片段、凝血酶可切割的接头(例如,48个氨基酸)和第二Fc区。构建体FVIII-174包含两条多肽链,第一链包含融合于AE288XTEN的重链FVIII序列并且第二链包含轻链FVIII、第一Fc区、接头(例如,可加工的接头)和第二Fc区。构建体FVIII-170包含VWF片段、AE288XTEN、接头(例如,凝血酶可切割的接头,其为35个氨基酸长度),和单链FVIII序列。

[0109] 图5.含有与Fc区组合的XTEN序列的FVIII/VWF异二聚体的药动学曲线。通过流体动力学注射(HDI)以100ug/小鼠的剂量向FVIII:VWF双敲除(DKO)小鼠施用构建体FVIII-161、FVIII-168和FVIII-172。通过HDI以50μg/小鼠的剂量向FVIII:VWF DKO小鼠施用构建体FVIII-170。通过HDI后FVIII显色测定24小时分析HDI后血浆FVIII活性。将含有XTEN序列和Fc结构域的FVIII:VWF异二聚体的FVIII活性与不含VWF片段、XTEN序列和Fc结构域的BDD-FVIII的FVIII活性进行比较。

[0110] 图6.FVIII-VWF异二聚体示例共转染系统的示意图。图6A.构建体FVIII-169含有连接于Fc区的全长FVIII序列(具有在1645和1648处替代精氨酸残基的丙氨酸残基和具有插入到单链FVIII序列中的XTEN序列)。VWF-031含有D1D2D'D3片段(具有在336和379处替代半胱氨酸残基的丙氨酸残基),所述片段连接于具有48凝血酶可切割的接头的另一Fc区。在胞内加工后,构建体FVIII-169生成融合于一个Fc片段和XTEN序列的全长单链FVIII

(SCFVIII), 并且构建体VWF-031生成连接于另一Fc片段的477个氨基酸的D'D3片段。两个共价键可在连接于SC FVIII或D'D3片段的Fc片段之间形成, 这反过来允许FVIII和D'D3的非共价缔合。图6B. 构建体FVIII-173含有异二聚体FVIII序列、连接于XTEN序列的重链FVIII序列和连接于Fc区的轻链FVIII序列。VWF-031如上文所述。在胞内加工后, 构建体FVIII-173生成异二聚体蛋白, 融合于XTEN序列的重链FVIII、融合于一个Fc片段的轻链FVIII, 并且构建体VWF-031生成连接于另一Fc片段的477个氨基酸的D'D3片段。两个共价键可在连接于轻链FVIII或D'D3片段的Fc片段之间形成, 这反过来允许FVIII和D'D3的非共价缔合。

[0111] 图7. 在Octet测定中含有XTEN序列和Fc结构域的示例性FVIII:VWF与固定化hVWF的结合亲和力。融合的FVIII-169/VWF-031和FVIII-057 (rFVIIIIFc) 与固定化hVWF的结合亲和力使用基于生物层干涉测量法的测量 (Octet测定) 进行测试。图7A示出FVIII169和FVIIIIFc原料药 (阳性对照) 与固定化hVWF的结合响应 (以纳摩尔计)。图7B示出人IgG1 (阴性对照) 与固定化hVWF的结合响应。

[0112] 图8. 在HemA和FVIII:VWF双敲除 (DKO) 小鼠中FVIII-169的药动学 (PK) 曲线。图8A示出在HemA小鼠中FVIII-169/VWF-031和FVIIIIFc的PK曲线。HemA小鼠用单次静脉内剂量的FVIII-169/VWF-031以200IU/kg治疗。从小鼠收集的血浆样本通过FVIII显色测定进行测试。FVIII-169/VWF-031的半衰期使用WinNonlin程序计算。图8B示出在FVIII/VWF DKO小鼠中FVIII-169/VWF-031、FVIII-169/Fc和FVIIIIFc的PK曲线。

[0113] 图9. 在表达D'D3的FVIII/VWF DKO小鼠中FVIII-XTEN变体的PK曲线。图9A示出FVIII-XTEN变体、具有一个XTEN的FVIII、具有两个XTEN的FVIII和具有三个XTEN的FVIII的PK曲线的比较。将一个、两个或三个XTEN插入到包括C-末端和B-结构域的FVIII的各个部分中。CT指示XTEN连接于FVIII的C-末端。插入位点B/CT指示一个XTEN在FVIII蛋白的氨基酸残基745和氨基酸残基746之间插入并且另一XTEN连接于FVIII蛋白的C-末端。氨基酸残基编号对应于SQ BDD FVIII蛋白序列。插入位点1900/B/CT指示第一XTEN在氨基酸残基1900和FVIII的氨基酸残基1901之间插入, 第二XTEN在FVIII的氨基酸残基745和氨基酸残基746之间插入, 并且第三XTEN连接于FVIII的C-末端。用于施用FVIII-XTEN变体的小鼠品系为表达D'D3结构域的DKO小鼠品系。图9B示出具有三个XTEN插入的FVIII-XTEN的PK曲线。向FVIII/VWF DKO小鼠或HemA小鼠施用FVIII-XTEN (1900/B/CT) 变体。比较FVIII-XTEN (1900/B/CT) 的半衰期。

[0114] 图10. 通过显色测定测量的小鼠DKO血浆中FVIIIIFc (空心三角形)、FVIII169:Fc (实心圆形) 和FVIII169:VWF31 (空心三角形) 的FVIII活性。FVIII:Fc含有融合于Fc二聚体 (即, 单体-二聚体杂合体) 的双链FVIII (重链和轻链)。FVIII169如上文所述 (含有B结构域中的AE288, 紧邻对应于成熟FVIII序列的氨基酸745的下游)。FVIII169:Fc含有融合于Fc二聚体的FVIII169。FVIII169:VWF31除了Fc二聚体、融合于第一Fc区的FVIII169和融合于第二Fc区的VWF31之外还含有VWF31, 其中第一Fc区和第二Fc区形成共价键, 例如, 一个或多个二硫键。

[0115] 图11. Fc、XTEN和VWF-D'D3片段对FVIII半衰期延长的效应。向FVIII和VWF双敲除 (DKO) 小鼠施用BDD-FVIII (REFACTO®) (正方形)、FVIIIIFc (圆形)、FVIII169/Fc (三角形) 和FVIII169/VWF031 (倒三角形)。FVIII活性通过显色测定来测量, 并且半衰期使用WinNonlin-Phoenix程序来计算。X轴示出时间, 并且Y轴示出以mU/mL计的FVIII血浆活性。

[0116] 图12A-C. rFVIII-XTEN/VWF异二聚体中的不同XTEN在HemA小鼠中的效应。图12A示出相较于含有紧邻对应于成熟FVIII序列的残基745的下游的XTEN的FVIII-169, 紧邻对应于成熟FVIII序列(即, FVIII-195(双链FVIII同工型)和FVIII-199(单链FVIII同工型))的残基1900和1656的下游插入的两个XTEN的标准化至5分钟值(%)的FVIII血浆活性。在HemA小鼠中施用FVIII-169/VWF-031(实心圆形)、FVIII-199/VWF-031(实心正方形)和FVIII-195/VWF031(空心正方形)以测量FVIII血浆活性。图12B示出相较于FVIII-169(仅XTEN插入到B结构域中), 紧邻对应于成熟FVIII序列的残基403(A2结构域)和745(B结构域)(即, FVIII-203)和残基745(B结构域)和1900(A3结构域)(FVIII-204)的下游插入的第二XTEN的半衰期延长效应。向HemA小鼠施用FVIII-204/VWF031(实心三角形)、FVIII-169/VWF-031(实心圆形)、FVIII-203/VWF-031(实心正方形)和scBDD-FVIII(空心菱形)。X轴示出标准化至5分钟值(%)的FVIII血浆活性, 并且Y轴示出以小时计的时间。图12C示出相较于FVIII-169(单个XTEN插入到B结构域中)和不含任何Fc区或任何XTEN的单链FVIII(即, FVIII-207), 紧邻残基18(A1结构域)和745(B结构域)(即, FVIII-205)的下游插入的两个XTEN的半衰期延长效应。图12C另外示出相较于FVIII-169(单个XTEN紧邻残基745的下游插入), 紧邻残基26(A1结构域)、1656(A3结构域)和1900(A3结构域)(即, FVIII-201)的下游并入的三个XTEN插入的半衰期延长效应。向HemA小鼠施用FVIII-205/VWF-031(实心正方形)、FVIII-201/VWF-031(倒三角形)、FVIII-169/VWF-031(实心圆形)和FVIII-207(空心菱形)。标准化至5分钟值(%)的FVIII血浆活性(X轴)随着以小时计的时间(Y轴)测量。

[0117] 图13. 在FVIII/VWF DKO小鼠rFVIII-XTEN/VWF-XTEN异二聚体的FVIII活性。通过FVIII显色测定分析血浆样本的FVIII活性, 并且对血浆FVIII活性(X轴)作为时间的函数(Y轴)的回归曲线作图。FVIII-155(不含任何XTEN的scFVIIIIFc)与VWF-034(具有AE 288XTEN加上35个残基的凝血酶可切割的接头的VWF-Fc)共表达。将FVIII-155/VWF-034的半衰期与FVIII-169/VWF-031的半衰期比较, 后者具有插入到FVIII的B结构域接合点(紧邻对应于成熟FVIII多肽的残基745的下游)的AE 288XTEN。

[0118] 图14A-H. 各种rFVIII-XTEN/VWF构建体的示意图。这些构建体也在本文其它部分中描述。图14A示出单链B结构域缺失的FVIII蛋白(有时在本文被表示为scBDD-FVIII)。scBDD-FVIII构建体含有在残基1645和1648处从Arg至Ala的两个替代。图14B示出两条多肽链构建体(FVIII155/VWF031), 第一链包含连接于不含任何XTEN的Fc区的单链FVIII和第二链包含连接于Fc区的VWF D'D3片段。该构建体用作对照。图14C示出两条多肽链构建体(FVIII199/VWF031), 第一链包含连接于Fc区的单链FVIII, 其中第一XTEN紧邻对应于成熟FVIII序列的残基1900的下游插入并且第二XTEN紧邻对应于成熟FVIII序列的残基1656的下游插入, 并且第二链包含连接于Fc区的VWF D'D3片段。图14D示出两条多肽链构建体(FVIII201/VWF031), 第一链包含连接于Fc区的单链FVIII蛋白, 其中第一XTEN紧邻对应于成熟FVIII序列的残基26的下游插入, 第二XTEN紧邻对应于成熟FVIII序列的残基1656的下游插入, 并且第三XTEN紧邻对应于成熟FVIII序列的残基1900的下游插入, 并且第二链包含连接于Fc区的VWF D'D3片段。图14E示出两条多肽链构建体(FVIII169/VWF031), 第一链包含连接于Fc区的单链FVIII蛋白, 其中XTEN紧邻对应于成熟FVIII序列的残基745(表示为"B")的下游插入, 并且第二链包含连接于Fc区的VWF D'D3片段。图14F示出两条多肽链构建体(FVIII203/VWF031), 第一链包含单链FVIII蛋白, 其中第一XTEN在对应于成熟FVIII序

列的残基745("B")处插入且第二XTEN在对应于成熟FVIII序列的残基1900插入,并且第二链包含连接于Fc区的VWF D'D3片段。图14G示出两条多肽链构建体(FVIII204/VWF031),第一链包含连接于Fc区的单链FVIII蛋白,其中第一XTEN紧邻对应于成熟FVIII序列的残基403的下游插入且第二XTEN紧邻对应于成熟FVIII序列的残基745("B")的下游插入,并且第二链包含连接于Fc区的VWF D'D3片段。图14H示出两条多肽链构建体(FVIII205/VWF031),第一链包含单链FVIII,其中第一XTEN紧邻对应于成熟FVIII序列的残基18的下游插入且第二XTEN紧邻对应于成熟FVIII序列的残基745("B")的下游插入,并且第二链包含连接于Fc区的VWF D'D3片段。

[0119] 图15.在FVIII/VWF DK0小鼠中rFVIII-XTEN/VWF和BDD-FVIII的FVIII活性。通过FVIII显色测定分析血浆样本的FVIII活性,并且对血浆FVIII活性(X轴)作为时间的函数(Y轴)的回归曲线作图。将rFVIII-XTEN/VWF(FVIII-205/VWF-031)的半衰期与BDD-FVIII和rFVIIIIFc的半衰期比较。

[0120] 图16.使用尾夹出血模型在HemA小鼠中FVIII-XTEN-Fc:VWF-Fc异二聚体的功效。HemA小鼠尾夹出血模型用于比较FVIII169/VWF034、FVIII205/VWF031和BDD-FVIII的功效。将200IU/kg的FVIII169/VWF034和FVIII205/VWF031的中值失血(以ml计)与200IU/kg的BDD-FVIII、65IU/kg的BDD-FVIII、20IU/kg的BDD-FVIII和媒介物比较。

[0121] 发明详述

[0122] 定义

[0123] 值得注意的是,术语"一(a)"或"一(an)"实体是指一个或多个该实体;例如,"一个核苷酸序列"应理解为表示一个或多个核苷酸序列。因此,术语"一(a)"(或"一(an)")、"一个或多个"和"至少一个"在本文可交换地使用。

[0124] 术语"多核苷酸"或"核苷酸"旨在涵盖单个核酸以及多个核酸,且是指分离的核酸分子或构建体,例如,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(例如酰胺键,诸如于肽核酸(PNA)中发现的)。术语"核酸"是指存在于多核苷酸中的任何一个或多个核酸区段,例如,DNA或RNA片段。所谓"分离的"核酸或多核苷酸意指已从其天然环境去除的核酸分子、DNA或RNA。例如,出于本发明的目的,编码载体中含有的因子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于异源宿主细胞中或(部分或基本上)从溶液中的其它多核苷酸纯化的重组多核苷酸。分离的RNA分子包括本发明的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合成地生成的此类分子。此外,多核苷酸或核酸可包括调控元件诸如启动子、增强子、核糖体结合位点或转录终止信号。

[0125] 如本文所使用,"编码区"或"编码序列"为多核苷酸的一部分,其由可翻译成氨基酸的密码子组成。尽管"终止密码子"(TAG、TGA或TAA)通常未被翻译成氨基酸,其可以被认为编码区的一部分,但是任何侧翼序列,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一部分。编码区的边界通常由编码所得的多肽的氨基末端的在5'末端的起始密码子和编码所得多肽的羧基末端的在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,例如在单个载体上,或存在于单独的多核苷酸构建体中,例如在单独(不同)的载体上。然后可得出结论,单个载体可含有仅仅单个编码区、或包含两个或更多个编码区,例如,单个载体可分别编码如下所述的结合结构域-A

和结合结构域-B。此外,本发明的载体、多核苷酸或核酸可编码异源编码区,所述异源编码区融合于或未融合于编码本发明的结合结构域的核酸。异源编码区包括但不限于专门的元件或基序,诸如分泌信号肽或异源功能结构域。

[0126] 由哺乳动物细胞分泌的某些蛋白质与分泌信号肽缔合,一旦启动粗面内质网中生长蛋白质链的输出,就从成熟蛋白切割所述分泌信号肽。本领域普通技术人员意识到,信号肽一般融合于多肽的N-末端,并且从完整或“全长”多肽切割以生成多肽的分泌或“成熟”形式。在某些实施方案中,该序列的天然信号肽或功能衍生物保留指导可操作地与其缔合的多肽的分泌的能力。可选地,可以使用异源哺乳动物信号肽,例如人组织纤溶酶原活化剂(TPA)或小鼠 β -葡萄糖苷酸酶信号肽、或其功能衍生物。

[0127] 术语“下游”是指位于参考核苷酸序列的3'的核苷酸序列。在某些实施方案中,下游核苷酸序列涉及跟随转录起始点的序列。例如,基因的翻译起始密码子位于转录起始点的下游。

[0128] 术语“上游”是指位于参考核苷酸序列的5'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'或转录起始点的序列。例如,大多数启动子位于转录起始点的上游。

[0129] 如本文所使用,术语“调控区”是指位于编码区上游(5'非编码序列)、之内或下游(3'非编码序列)的核苷酸序列,并且其影响缔合编码区的转录、RNA加工、稳定性或翻译。调控区可以包括启动子、翻译前导序列、内含子、聚腺苷酸化识别序列、RNA加工位点、效应子结合位点和茎环结构。如果编码区用来在真核细胞中表达,则聚腺苷酸化信号和转录终止序列通常将位于编码序列的3'。

[0130] 编码基因产物例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,基因产物例如多肽的编码区与一个或多个调控区缔合以使在调控区的影响或控制下进行基因产物的表达。例如,如果诱导启动子功能导致编码由编码区编码的基因产物的mRNA转录,并且如果启动子与编码区之间的连接的性质未干扰启动子指导基因产物表达的能力或干扰转录DNA模板的能力,则编码区和启动子“可操作地缔合”。除启动子以外,其它转录控制元件例如增强子、操作子、阻遏子和转录终止信号还可操作地与编码区缔合以指导基因产物表达。

[0131] 多种转录控制区域为本领域技术人员已知的。这些包括但不限于,在脊椎动物细胞中发挥作用的转录控制区域,诸如但不限于来自巨细胞病毒(立即早期启动子,与内含子-A结合)、猿猴病毒40(早期启动子)和逆转录病毒(诸如劳斯肉瘤病毒)的启动子和增强子区段。其它转录控制区域包括源自脊椎动物基因诸如肌动蛋白、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的基因表达的其它序列。另外的适合的转录控制区域包括组织特异性启动子和增强子以及淋巴因子诱导性启动子(例如,可通过干扰素或白介素诱导的启动子)。

[0132] 类似地,多种翻译控制元件为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,以及源自小核糖核酸病毒的元件(特别是内部核糖体进入位点或IRES,也被称为CITE序列)。

[0133] 如本文所使用的术语“表达”是指多核苷酸借此生成基因产物,例如RNA或多肽的过程。其包括但不限于多核苷酸至信使RNA(mRNA)、转运RNA(tRNA)、小发夹RNA(shRNA)、小

干扰RNA (siRNA) 或任何其它RNA产物的转录以及mRNA至多肽的翻译。表达生成“基因产物”。如本文所使用,基因产物可以或者为核酸,例如通过基因的转录而生成的信使RNA,或者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例如聚腺苷酸化或剪接的核酸,或具有后翻译修饰例如甲基化、糖基化、脂质的添加、与其它蛋白亚基的缔合或蛋白水解切割的多肽。

[0134] “载体”是指用于将核酸克隆进宿主细胞和/或将核酸转移到宿主细胞中的任何媒介物。载体可以为复制子,另一核酸区段可连接于复制子以使引起连接区段的复制。“复制子”是指作为体内复制的自主单位发挥作用,即,能够在其自身控制下复制的任何基因元件(例如,质粒、噬菌体、粘粒、染色体、病毒)。术语“载体”包括体外、离体或体内将核酸引入到细胞中的病毒和非病毒媒介物两者。大量载体为本领域已知并使用,包括例如质粒、修饰的真核细胞病毒或修饰的细菌病毒。将多核苷酸插入到适合的载体中可通过将适当的多核苷酸片段连接到具有互补的粘性末端的所选择的载体来完成。

[0135] 载体可被工程改造为编码可选择标记或报告蛋白,其提供对已并入载体中的细胞的选择或鉴定。可选择标记或报告蛋白的表达允许鉴定和/或选择并入和表达在载体上含有的其它编码区的宿主细胞。本领域中已知且使用的可选择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、磺酰胺等的抗性的基因;和用作表型标记的基因,即花青素调控基因、异戊烷基转移酶基因等。本领域中已知且使用的报告蛋白的实例包括: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)、-葡萄糖苷酸酶 (Gus) 等。可选择标记也可被认为是报告蛋白。

[0136] 术语“质粒”是指通常携带不是细胞中心代谢的一部分的基因的染色体外元件,并且通常呈圆形双链DNA分子形式。此类元件可以是自主复制序列、基因组整合序列、噬菌体或核苷酸序列,线性、圆形或超螺旋,具有单链或双链DNA或RNA,源自任何来源,其中许多核苷酸序列已联接或重组到独特构造中,所述独特构造能够将用于将选定基因产物的启动子片段和DNA序列以及适当的3'未翻译序列一起引入细胞中。

[0137] 可使用的真核细胞病毒载体包括但不限于,腺病毒载体、逆转录病毒载体、腺相关病毒载体和痘病毒,例如,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、带电脂质 (cytofectin)、DNA-蛋白质复合物和生物聚合物。

[0138] “克隆载体”是指“复制子”,其为按顺序复制的单位长度的核酸,并且其包含复制起点,诸如质粒、噬菌体或粘粒,另一核酸区段可连接至复制起点以使引起连接区段的复制。某些克隆载体能够在一种细胞类型例如细菌中复制并且在另一种细胞类型例如真核细胞中表达。克隆载体通常包含可用于选择包含载体的细胞的一个或多个序列和/或用于插入目标核酸序列的一个或多个多克隆位点。

[0139] 术语“表达载体”是指被设计为能够在插入到宿主细胞中后实现插入核酸序列的表达的媒介物。插入核酸序列被布置成与如上所述的调控区可操作缔合。

[0140] 通过本领域熟知的方法,例如转染、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载体转运蛋白将载体引入宿主细胞中。

[0141] 如本文所使用的“培养 (culture)”、“培养 (to culture)”和“培养 (culturing)”意指在允许细胞生长或分裂或者保持细胞处于存活状态的体外条件下温育细胞。如本文所使

用的“培养细胞”意指体外繁殖的细胞。

[0142] 如本文所使用,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)组成的分子。术语“多肽”是指两个或更多个氨基酸的任何链,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”或任何其它术语用于指两个或更多个氨基酸的链,包括在“多肽”的定义范围内,并且术语“多肽”可与这些术语中的任一种替代或交换使用。术语“多肽”也旨在是指多肽的表达后修饰的产物,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、用已知保护/封闭基团的衍生化、蛋白水解切割或用非天然存在氨基酸的修饰。多肽可源自自然生物源或由重组技术生成,但不一定从指定核酸序列翻译。其可以任何方式产生,包括通过化学合成来产生。

[0143] “分离的”多肽或其片段、变体或衍生物是指未处于其自然环境的多肽。无特定的纯化水平是需要的。例如,分离的多肽可简单地从其天然或自然环境去除。出于本发明目的,在宿主细胞中表达的重组生成的多肽和蛋白质被认为是分离的,已通过任何适合的技术而分离、分级分离、或者部分或大致上纯化的天然或重组多肽也被认为是分离的。

[0144] 还包括在本发明中的是多肽的片段或变体、及其任何组合。术语“片段”或“变体”当是指本发明的多肽结合结构域或结合分子时包括保留参考多肽的至少一些性质(例如,对于FcRn结合结构域或Fc变体的FcRn结合亲和力、FVIII变体的凝结活性或VWF片段的FVIII结合活性)的任何多肽。除了在本文别处论述的特异性抗体片段之外,多肽的片段还包括蛋白水解片段以及缺失片段,但是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如上所述的片段,并且也包括具有由于氨基酸替代、缺失或插入引起的改变的氨基酸序列的多肽。变体可以天然或非天然存在。非天然存在的变体可使用本领域已知的诱变技术生成。变体多肽可包含保守性或非保守性氨基酸替代、缺失或添加。

[0145] 本文所使用的术语“VWF片段”或“VWF片段”意指与FVIII相互作用并且保留通常由全长VWF提供给FVIII的至少一种或多种性质的任何VWF片段,所述性质例如阻止对FVIIIa的过早活化、阻止过早蛋白质水解、阻止可导致过早清除的与磷脂膜的缔合、阻止结合于FVIII清除受体(其可结合裸FVIII但不是VWF结合的FVIII)和/或稳定FVIII重链和轻链相互作用。

[0146] “保守性氨基酸替代”是其中氨基酸残基被替换为具有类似侧链的氨基酸残基的替代。已在本领域中定义具有类似侧链的氨基酸残基的家族,包括碱性侧链(例如,赖氨酸、精氨酸、组氨酸)、酸性侧链(例如,天冬氨酸、谷氨酸)、未荷电极性侧链(例如,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(例如,丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、蛋氨酸、色氨酸)、 β -支链侧链(例如,苏氨酸、缬氨酸、异亮氨酸)和芳族侧链(例如,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替换为来自相同侧链家族的另一氨基酸,则替代被认为是保守的。在另一个实施方案中,一串氨基酸可被保守地替换为侧链家族成员的次序和/或组成不同的结构上类似的串。

[0147] 如本领域中已知的,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多肽的序列来确定。当在本文论述时,任何特定多肽是否与另一多肽具有至少约50%、60%、70%、75%、80%、85%、90%、95%、99%或100%同一性可使用本领域中已知的

方法和计算机程序/软件来确定,诸如但不限于BESTFIT程序(Wisconsin序列分析包,Unix第8版,Genetics Computer Group,University Research Park,575Science Drive, Madison,WI 53711)。BESTFIT使用Smith和Waterman,Advances in Applied Mathematics 2:482-489 (1981)的局部同源算法来找出两个序列之间的同源性的最佳区段。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,设定参数,当然,使得在参考多肽序列的全长范围内计算同一性百分比以及允许参考序列中的氨基酸总数的同源性高达5%的空位。

[0148] 如本文所使用,VWF序列或FVIII蛋白序列中“对应于……的氨基酸”或“等价氨基酸”通过比对使第一VWF或FVIII序列和第二VWF或FVIII序列之间的同一性或类似性最大化来鉴定。鉴定第二VWF或FVIII序列中的等价氨基酸所用的编号是基于鉴定第一VWF或FVIII序列中的对应氨基酸所用的编号。

[0149] 如本文所使用,术语“插入位点”是指FVIII多肽或其片段、变体或衍生物的位置,其紧邻异源部分可插入的位置的上游。“插入位点”被指定为编号,所述编号为成熟的天然FVIII (SEQ ID NO:4) 中插入位点对应的氨基酸的编号,其紧邻插入位置的N-端。例如,短语“a3在对应于SEQ ID NO:4的氨基酸1656的插入位点包含XTEN”表明异源部分位于对应于SEQ ID NO:4的氨基酸1656和氨基酸1657的两个氨基酸之间。

[0150] 如本文所使用的短语“紧邻氨基酸的下游”是指正好挨着氨基酸的末端羧基的位置。类似地,短语“紧邻氨基酸上游”是指正好挨着氨基酸的末端胺基的位置。因此,如本文所使用的短语“在插入位点的两个氨基酸之间”是指其中XTEN或任何其它多肽在两个相邻氨基酸之间插入的位置。因此,短语“紧邻氨基酸的下游插入”和“在插入位点的两个氨基酸之间插入”与“在插入位点插入”同义使用。

[0151] 如本文所使用,术语“插入”、“被插入”、“插入到”或语法相关的术语,是指嵌合多肽中的XTEN相对于天然的成熟人FVIII中的类似位置的位置。如本文所使用,术语是指重组FVIII多肽相对于天然的成熟人FVIII的特征,且不指示、暗示或推断制备嵌合多肽的任何方法或过程。例如,关于本文提供的嵌合多肽,短语“XTEN插入到紧邻FVIII多肽的残基745的下游”意指嵌合多肽包含紧邻对应于天然的成熟人FVIII中的氨基酸745的氨基酸的下游的XTEN,例如,受对应于天然的成熟人FVIII的氨基酸745和746的氨基酸限制。

[0152] “融合”或“嵌合”蛋白包含连接于在自然界中非天然连接的第二氨基酸序列的第一氨基酸序列。通常以单独蛋白质存在的氨基酸序列可一起集中于融合多肽中,或通常存在于同一蛋白质中的氨基酸序列可在融合多肽中以新排列布置,例如,本发明的因子VIII结构域与Ig Fc结构域的融合。融合蛋白例如通过化学合成、或通过产生和翻译其中肽区域以所需关系编码的多核苷酸来产生。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。

[0153] 如本文所使用,术语“半衰期”是指体内特定多肽的生物半衰期。半衰期可由向受试者施用的数量半数从动物中的循环和/或其它组织清除所需的时间表示。当构建作为时间的函数的给定多肽的清除曲线时,曲线通常为具有快速 α -相和较长 β -相的两相。 α -相通常表示所施用的Fc多肽在血管内和血管外空间之间的平衡,并且部分通过多肽的大小来确定。 β -相通常表示多肽在血管内空间中的分解代谢。在一些实施方案中,FVIII和包含FVIII的嵌合蛋白为单相的,并且因而不具有 α 相,但仅仅单一 β 相。因此,在某些实施方案中,如本

文所使用的术语半衰期是指 β -相中多肽的半衰期。人抗体在人中的典型 β -相半衰期为21天。

[0154] 如本文所使用的术语“连接”是指分别共价或非共价联接于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。第一氨基酸或核苷酸序列可以直接联接于或并置于第二氨基酸或核苷酸序列或可选地间插序列可共价联接第一序列与第二序列。术语“连接”不仅意指第一氨基酸序列与第二氨基酸序列在C-末端或N-末端的融合,而且包括将整个第一氨基酸序列(或第二氨基酸序列)插入到第二氨基酸序列(分别地,或第一氨基酸序列)的任意两个氨基酸中。在一个实施方案中,第一氨基酸序列可通过肽键或接头连接于第二氨基酸序列。第一核苷酸序列可通过磷酸二酯键或接头连接于第二核苷酸序列。接头可以为肽或多肽(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两者)。术语“连接”也通过连字号(-)表示。

[0155] 如本文所使用的术语“与……缔合”是指在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”意指共价键、非肽键或非共价键。这种缔合可通过冒号即(:)表示。在另一个实施方案中,其意指除肽键之外的共价键。例如,氨基酸半胱氨酸包含可与第二半胱氨酸残基上的硫醇基形成二硫键或桥的硫醇基。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(位置226或229,EU编号系统)的239和242的位置处缔合。共价键的实例包括但不限于,肽键、金属键、氢键、二硫键、S键、 π 键、 δ 键、糖苷键、不可知键、弯键、偶极键、 π 反向键、双键、三键、四重键、五重键、六重键、共轭、超共轭、芳香性、哈普托数(hapticity)或反键。非共价键的非限制性实例包括离子键(例如,阳离子 π 键或盐键)、金属键、氢键(例如,双氢键、双氢复合物、低势垒氢键或对称氢键)、范德华力、伦敦色散力、机械键、卤素键、亲金性、嵌入、堆叠、熵力或化学极性。

[0156] 本文所使用的术语“单体-二聚体杂合体”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例如因子VIII和第一Fc区并且第二链包含以下、基本上由以下组成或由以下组成:不含凝血因子的第二Fc区。单体-二聚体杂合体构建体因而为包含仅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体。

[0157] 如本文所使用,术语“切割位点”或“酶切割位点”是指被酶识别的位点。某些酶切割位点包括胞内加工位点。在一个实施方案中,多肽具有被在凝血级联期间活化的酶切割的酶切割位点,以使此类位点的切割在血块形成的位点发生。示例性的此类位点包括,例如由凝血酶、因子XIa或因子Xa识别的那些。示例性的FXIa切割位点包括,例如,TQSFNDFTR (SEQ ID NO:45)和SVSQTSKLTR (SEQ ID NO:46)。示例性的凝血酶切割位点包括,例如DFLAEGGGVR (SEQ ID NO:47)、TTKIKPR (SEQ ID NO:48)、LVPRG (SEQ ID NO:49)和ALRPR (SEQ ID NO:50)的氨基酸1至5)。其它酶切割位点是本领域已知的。

[0158] 如本文所使用,术语“加工位点”或“胞内加工位点”是指多肽中的一类酶切割位点,其为在翻译多肽后发挥作用的酶的靶标。在一个实施方案中,此类酶在从高尔基体腔至转运高尔基区室转运期间发挥作用。胞内加工酶在从细胞分泌蛋白质之前切割多肽。此类加工位点的实例包括,例如由内肽酶的PACE/弗林蛋白酶(其中PACE是成对碱性氨基酸切割酶(Paired basic Amino acid Cleaving Enzyme)首字母缩略词)家族靶向的加工位点。这

些酶被定位至高尔基膜并且切割序列基序Arg-[任意残基]- (Lys或Arg)-Arg的羧基端侧上的蛋白。如本文所使用,酶的“弗林蛋白酶”家族包括,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。

[0159] 在包括超过一个加工或切割位点的构建体中,应理解此类位点可以相同或不同。

[0160] 术语“弗林蛋白酶”是指对应于EC No.3.4.21.75的酶。弗林蛋白酶为枯草杆菌蛋白酶样蛋白原转化酶,其也称为PACE (成对碱性氨基酸切割酶)。弗林蛋白酶删除无活性前体蛋白的部分以将它们转化成生物活性蛋白。在其细胞内转运期间,可通过弗林蛋白酶从成熟VWF分子切割VWF的肽原。在一些实施方案中,弗林蛋白酶从VWF的D'D3切割D1D2。在其它实施方案中,编码弗林蛋白酶的核苷酸序列可与编码VWF片段的核苷酸序列一起被表达以使可通过弗林蛋白酶在细胞内切下D1D2结构域。

[0161] 在包括超过一个加工或切割位点的构建体中,应理解此类位点可以相同或不同。

[0162] 如本文所使用的“可加工的接头”是指包含至少一个胞内加工位点的接头,其在本文别处描述。

[0163] 如本文所使用的止血障碍意指基因遗传性或获得性疾患,其特征不在于由于形成纤维蛋白凝块的能力受损或丧失,自发地或由于创伤造成的出血的倾向。此类病症的实例包括血友病。三种主要形式为血友病A (因子VIII缺乏症)、血友病B (因子IX缺乏症或“Christmas病”)和血友病C (因子XI缺乏症,轻微出血倾向)。其它止血障碍包括,例如冯维勒布兰德病、因子XI缺乏症 (PTA缺乏症)、因子XII缺乏症、纤维蛋白原、凝血酶原、因子V、因子VII、因子X或因子XIII的缺乏症或结构异常、Bernard-Soulier综合征 (其为GPIb缺少或缺乏症)。VWF的受体GPIb可以是缺乏的并且导致缺少初级血块形成 (初级止血) 和增加的出血倾向,和Glanzmann和Naegeli的血小板机能不全 (Glanzmann血小板机能不全)。在肝衰竭 (急性和慢性形式) 中,肝脏生成凝结因子不足;这可增加出血危险。

[0164] 可防治性地使用本发明的嵌合分子。如本文所使用的术语“防治性治疗”是指在流血事件之前施用分子。在一个实施方案中,需要一般止血剂的受试者正经受或即将经受手术。本发明的嵌合蛋白可在手术之前或之后作为防治剂施用。本发明的嵌合蛋白可在手术期间或之后施用以控制急性出血事件。手术可包括但不限于,肝移植、肝切除、牙科手术或干细胞移植。

[0165] 本发明的嵌合蛋白还用于按需治疗。术语“按需治疗”是指响应于流血事件的症状或在可导致出血的活动之前施用嵌合分子。一方面,当出血开始时,诸如在手术后,或当预计出血时,诸如在手术之前可向受试者给予按需治疗。在另一方面,可在增加出血风险的活动,诸如接触性运动之前给予按需治疗。

[0166] 如本文所使用的术语“急性出血”是指出血事件,而不管根本原因如何。例如,受试者可具有创伤、尿毒症、遗传性出血性病症 (例如因子VII缺乏症)、血小板病症或归因于抗体的开发而对凝血因子的抗性。

[0167] 如本文所使用的治疗 (treat, treatment, treating) 是指,例如降低疾病或疾患的严重性;减少病程的持续时间;改善与疾病或疾患相关的一种或多种症状;向患有疾病或疾患的受试者提供有益效果而不一定治愈疾病或疾患、或防治与疾病或疾患相关的一种或多

种症状。在一个实施方案中,术语“治疗(treating)”或“治疗(treatment)”意指通过施用本发明的嵌合蛋白或VWF片段使受试者中的FVIII低谷水平维持为至少约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。在另一个实施方案中,治疗(treating)或治疗(treatment)意指维持FVIII低谷水平在约1至约20IU/dL、约2至约20IU/dL、约3至约20IU/dL、约4至约20IU/dL、约5至约20IU/dL、约6至约20IU/dL、约7至约20IU/dL、约8至约20IU/dL、约9至约20IU/dL或约10至约20IU/dL之间。疾病或疾患的治疗也可包括使受试者中的FVIII活性维持在与非血友病受试者中至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%或20%的FVIII活性类似的水平。治疗所需的最小低谷水平可通过一种或多种已知方法测量并且可针对每位人进行调节(增加或减少)。

[0168] 嵌合蛋白

[0169] 本发明涉及使用VWF片段和XTEN序列通过阻止或抑制FVIII半衰期限制因子即内源性VWF与FVIII蛋白缔合来延长因子VIII蛋白的半衰期。内源性VWF与非共价复合物中的约95%至约98%的FVIII缔合。虽然内源性VWF为FVIII半衰期限制因子,但是结合于FVIII蛋白的内源性VWF也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免遭蛋白酶切割和FVIII活化,稳定FVIII重链和/或轻链,并且阻止清除剂受体对FVIII的清除。但是,同时内源性VWF通过阻止胞饮和通过从系统经由VWF清除途径清除FVIII-VWF复合物来限制FVIII半衰期。尽管不受理论限制,但是据信内源性VWF为半衰期限制因子,其阻止融合于半衰期延长剂的FVIII蛋白的半衰期比野生型FVIII的半衰期长约两倍。因此,本发明涉及使用VWF片段来阻止或抑制内源性VWF与FVIII蛋白之间的相互作用,从而通过使用单独的XTEN序列或与Ig恒定区或其一部分组合的XTEN序列来增加FVIII蛋白的半衰期。XTEN序列可连接于FVIII蛋白或VWF片段。与VWF片段缔合的FVIII蛋白因而更缓慢地从循环被一个或多个VWF清除受体清除,然后相较于野生型FVIII或不含VWF片段的FVIII蛋白,可具有XTEN序列或与Ig恒定区组合的XTEN序列的完整的半衰期延长。

[0170] 在一个实施方案中,VWF片段通过共价或非共价键与FVIII蛋白缔合(或连接)。然而,在某些情况下,VWF片段和FVIII蛋白之间的物理封闭或化学缔合(例如,非共价键合)可能不足够强以在内源性VWF存在下提供包含FVIII蛋白和VWF片段的稳定复合物。例如,在无任何其它连接下与FVIII蛋白形成非共价键的VWF片段可在内源性VWF存在下容易地从FVIII蛋白体内解离,用内源性VWF替换VWF片段(例如,重组VWF,即rVWF)。因此,非共价结合于内源性VWF的FVIII蛋白将经受VWF清除途径并且容易地从系统清除。为了阻止VWF片段与FVIII蛋白的解离,在一些实施方案中,FVIII蛋白和VWF片段之间的缔合或连接为共价键,例如,肽键、一个或多个氨基酸或二硫键。在某些实施方案中,加成物部分和FVIII蛋白之间的缔合(即,连接)为FVIII蛋白和VWF片段(“FVIII/VWF接头”)之间的肽键或接头。接头的非限制性实例在本文别处描述。在一些实施方案中,VWF片段为包含以下、基本上由以下组成或由以下组成的多肽:至少约10、100、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2500、3000或4000个氨基酸。VWF片段的非限制性实例在本文别处描述。

[0171] 在某些实施方案中,VWF片段以化学方法(例如,非共价)结合于或以物理方法封闭

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。在又其它实施方案中，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:4[全长成熟FVIII]的氨基酸1669至1689和/或2303至2332。

[0172] 本发明也提供还包含一个或多个XTEN序列的嵌合蛋白(包含FVIII蛋白和VWF片段)，所述XTEN序列提供另外的半衰期延长性质。一个或多个XTEN序列可在FVIII蛋白或VWF片段内插入或连接于FVIII蛋白或VWF片段的N-末端或C-末端。本发明还包括连接于XTEN序列(第一半衰期延长部分)和Ig恒定区或其一部分(第二半衰期延长部分)的FVIII蛋白以使两个半衰期延长部分通过两种不同机制延长FVIII蛋白的半衰期。

[0173] 在一些实施方案中，嵌合蛋白包含连接于第一Ig恒定区或其一部分(例如，第一FcRn结合配偶体)的FVIII蛋白、连接于第二Ig恒定区或其一部分(例如，第二FcRn结合配偶体)的VWF片段和插入或连接于FVIII蛋白或VWF片段的一个或多个XTEN序列，其中VWF片段阻止FVIII半衰期限制因子(例如，内源性VWF)结合于FVIII蛋白，其中第一和第二Ig恒定区或其部分形成共价键，例如二硫键，并且一个或多个XTEN序列延长FVIII蛋白的半衰期。

[0174] 在某些实施方案中，本发明的嵌合蛋白包含通过任选的接头(即，FVIII/VWF接头)连接于VWF片段的FVIII蛋白和一个或多个插入或连接于FVIII蛋白或VWF片段的XTEN序列，其中VWF片段阻止FVIII半衰期限制因子(例如，内源性VWF)结合于FVIII蛋白，并且一个或多个XTEN序列延长FVIII蛋白的半衰期。一方面，任选的接头(FVIII/VWF接头)包含分选酶识别基序。在另一方面，任选的接头(FVIII/VWF接头)包含可切割的位点。切割接头(即，含有一个或多个切割位点的接头)的实例在本文别处描述。

[0175] 本发明的嵌合蛋白包括但不限于：

[0176] (1) 包含D' 结构域和D3结构域的VWF片段、XTEN序列和FVIII，其中XTEN序列连接于VWF片段；

[0177] (2) FVIII蛋白、XTEN序列和Ig恒定区或其一部分，其中FVIII蛋白连接于XTEN序列和Ig恒定区或其一部分，或

[0178] (3) FVIII蛋白、XTEN序列和VWF片段，其中XTEN序列在C-末端或N-末端连接于FVIII蛋白或紧邻FVIII的一个或多个氨基酸(例如，一个或多个XTEN插入位点)的下游插入，并且VWF片段和FVIII蛋白彼此缔合。

[0179] (1) 连接于XTEN和FVIII的血管性血友病因子(VWF) 片段

[0180] 本发明涉及嵌合蛋白，其包含(i) 含有VWF的D' 结构域和D3结构域的VWF片段，(ii) XTEN序列，和(iii) FVIII蛋白，其中(i)、(ii) 和(iii) 彼此连接或缔合。连接于XTEN序列的VWF片段，作为本发明中嵌合蛋白的一部分，与FVIII蛋白缔合，因而阻止或抑制内源性VWF和FVIII蛋白之间的相互作用。在某些实施方案中，能够阻止或抑制FVIII蛋白与内源性VWF结合的VWF片段可同时具有至少一种VWF样FVIII保护性质。VWF样FVIII保护性质的实例包括但不限于，保护FVIII免遭蛋白酶切割和FVIII活化、稳定FVIII重链和/或轻链以及阻止清除剂受体对FVIII的清除。因此，VWF片段可通过VWF清除途径阻止FVIII蛋白的清除，因而减少FVIII从循环系统的清除。在一些实施方案中，本发明的VWF片段结合于FVIII蛋白或与FVIII蛋白缔合和/或以物理方法或以化学方法封闭FVIII蛋白上的VWF结合位点。相较于野生型FVIII或不与VWF片段缔合的FVIII，与VWF片段缔合的FVIII蛋白因而更缓慢地从循环

清除。

[0181] 在一个实施方案中,本发明涉及嵌合蛋白,其包含(i)含有VWF的D'结构域和D3结构域的VWF片段,(ii)XTEN序列,和(iii)FVIII蛋白,其中XTEN序列连接于VWF片段(例如,(a1)V-X或(a2)X-V,其中V包括VWF片段和X包括XTEN序列),并且VWF片段连接于或与FVIII蛋白缔合。在另一个实施方案中,VWF片段和XTEN序列可通过接头(例如,(a3)V-L-X或(a4)X-L-V)或肽键连接。接头可以为可切割的接头例如凝血酶可切割的接头,所述接头可在凝固位点被切割。在其它实施方案中,VWF片段、XTEN序列和FVIII蛋白被置于单一多肽链中。在又其它实施方案中,嵌合蛋白包含两条多肽链,第一链包含VWF片段和XTEN序列并且第二链包含FVIII蛋白。在另外其它实施方案中,嵌合蛋白包含三条多肽链,第一链包含VWF片段和XTEN序列,第二链包含FVIII的轻链且第三链包含FVIII的重链,其中第一链和第二链彼此缔合(例如,共价键,例如二硫键),并且第二链和第三链彼此缔合(例如,金属键)。在又其它实施方案中,XTEN序列可连接于VWF片段的N-末端或C-末端或紧邻VWF片段中一个或多个氨基酸的下游插入。

[0182] 在某些实施方案中,本发明的嵌合蛋白包含包括以下的式:

[0183] (a) V-X-FVIII、

[0184] (b) FVIII-X-V、

[0185] (c) V-X:FVIII、

[0186] (d) X-V:FVIII、

[0187] (e) FVIII:V-X、

[0188] (f) FVIII:X-V或

[0189] (a5) X-V-FVIII、

[0190] 其中V包括VWF片段,

[0191] X包括一个或多个XTEN序列,

[0192] FVIII包括FVIII蛋白;

[0193] (-)表示肽键或一个或多个氨基酸;且

[0194] (:)为化学缔合或物理缔合。在一个实施方案中,(:)表示化学缔合,例如,至少一个非肽键。在另一个实施方案中,化学缔合即(:)为共价键。在其它实施方案中,化学缔合即(:)为非共价相互作用,例如离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在其它实施方案中,(:)为非肽共价键。在又其它实施方案中,(:)为肽键。在另外其它实施方案中,(:)表示在两个序列之间的物理缔合,其中第一序列的一部分与第二序列紧密相邻以使第一序列掩蔽或阻断第二序列的一部分与另一部分相互作用,以及进一步在不允许第二序列与其它部分相互作用的情况下保持这种物理缔合。本文多肽式的取向从N-末端(左)至C-末端(右)列出。例如,式V-X-FVIII意指式NH₂-V-X-FVIII-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任何另外的序列。例如,式V-X-FVIII还可包含在V与X之间、X与FVIII之间的V的N-末端、或在FVIII的C-末端的任何序列,除非另有规定。在另一个实施方案中,连字号(-)指示肽键。

[0195] 在其它实施方案中,本发明的嵌合蛋白包含包括以下的式:

[0196] (a) V(X₁)-X₂-FVIII、

[0197] (b) FVIII-X₂-V(X₁)、

- [0198] (c) V(X1):FVIII、
- [0199] (d) FVIII:V(X1) 或
- [0200] (a5) X2-V(X1)-FVIII，
- [0201] 其中V(X1)包括VWF片段和第一XTEN序列(X1)，其中XTEN序列紧邻VWF片段中一个或多个氨基酸的下游插入，
- [0202] X2包括一个或多个任选的XTEN序列，
- [0203] FVIII包括FVIII蛋白；
- [0204] (-) 为肽键或一个或多个氨基酸；且
- [0205] (:) 为化学缔合或物理缔合。
- [0206] 在一些实施方案中，嵌合蛋白包含(i)含有VWF的D'结构域和D3结构域的VWF片段，(ii)XTEN序列，(iii)FVIII蛋白，(iv)第一任选的接头，和(v)第二任选的接头，其中XTEN序列通过接头连接于VWF片段和/或FVIII蛋白。在某些实施方案中，嵌合蛋白包含包括以下的式：
- [0207] (b1) V-L1-X-L2-FVIII、
- [0208] (b2) FVIII-L2-X-L1-V、
- [0209] (b3) V-L1-X:FVIII、
- [0210] (b4) X-L1-V:FVIII、
- [0211] (b5) FVIII:V-L1-X、
- [0212] (b6) FVIII:X-L1-V、
- [0213] (b7) X-L1-V-L2-FVIII或
- [0214] (b8) FVIII-L2-V-L1-X，
- [0215] 其中V包括VWF片段，
- [0216] X包括一个或多个XTEN序列，
- [0217] FVIII包括FVIII蛋白，
- [0218] L1包括第一任选的接头例如，第一可切割的接头，
- [0219] L2包括第二任选的接头例如，第二可切割的接头或任选的可加工的接头；
- [0220] (-) 为肽键或一个或氨基酸；且
- [0221] (:) 为化学缔合或物理缔合。在一个实施方案中，(:)表示化学缔合，例如，至少一个非肽键。在另一个实施方案中，化学缔合即(:)为共价键。在其它实施方案中，化学缔合即(:)为非共价相互作用，例如离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在其它实施方案中，(:)为非肽共价键。在又其它实施方案中，(:)为肽键。在另外其它实施方案中，(:)表示在两个序列之间的物理缔合，其中第一序列的一部分与第二序列紧密相邻以使第一序列掩蔽或阻断第二序列的一部分与另一部分相互作用，以及进一步在不允许第二序列与其它部分相互作用的情况下保持这种物理缔合。本文多肽式的取向从N-末端(左)至C-末端(右)列出。例如，式(b1)V-L1-X-L2-FVIII意指式NH₂-V-L1-X-L2-FVIII-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任何另外的序列。在另一个实施方案中，连字号(-)指示肽键。
- [0222] 本公开的另一方面是提供这样的FVIII嵌合蛋白，其与FVIII半衰期限制因子例如内源性VWF的相互作用减少或无相互作用，并且同时使用XTEN序列(第一半衰期延长剂)与

在FVIII蛋白和VWF片段之间提供共价键的第二半衰期延长剂或部分(例如Ig恒定区或其一部分)的组合使FVIII蛋白的半衰期最大化。在一个实施方案中,本发明的嵌合蛋白包含(i)含有VWF的D'结构域和D3结构域的VWF片段,(ii)XTEN序列,(iii)FVIII蛋白,和(iv)Ig恒定区或其一部分(在本文也被称为F),其中(1)VWF片段通过任选的接头例如可切割的接头连接于XTEN序列,(2)VWF片段通过另外的任选的接头例如可切割的接头与FVIII蛋白缔合或连接,并且(3)Ig恒定区或其一部分连接于VWF片段、XTEN序列或FVIII蛋白。在另一个实施方案中,本发明的嵌合蛋白包含(i)含有VWF的D'结构域和D3结构域的VWF片段,(ii)XTEN序列,(iii)FVIII蛋白,(iv)Ig恒定区或其一部分(F1或第一Ig恒定区或其一部分),和(v)另外的Ig恒定区或其一部分(F2或第二Ig恒定区或其一部分),其中(1)VWF片段通过任选的接头例如可切割的接头连接于XTEN序列,(2)XTEN序列或VWF片段连接于Ig恒定区或其一部分,(3)FVIII连接于另外的Ig恒定区或其一部分,和(4)Ig恒定区或其一部分与另外的Ig恒定区或其一部分缔合或连接。在一个实施方案中,两个Ig恒定区或其一部分之间的缔合或连接为共价键,例如二硫键。在另一个实施方案中,两个Ig恒定区或其一部分之间的缔合或连接为可加工的接头,其中可加工的接头通过蛋白酶细胞内加工。例如,嵌合蛋白包含包括以下的:

- [0223] (g) V-L2-X-L1-F1:FVIII-L3-F2;
- [0224] (h) V-L2-X-L1-F1:F2-L3-FVIII;
- [0225] (i) F-L1-X-L2-V:FVIII-L3-F2;
- [0226] (j) F-L1-X-L2-V:F2-L3-FVIII;
- [0227] (k) V-L2-X-L1-F1-L4-FVIII-L3-F2;
- [0228] (l) F2-L3-FVIII-L4-F1-L1-X-L2-V;
- [0229] (m) FVIII-L2-F2-L4-V-L2-X-L1-F1;或
- [0230] (n) F1-L1-X-L2-V-L4-F2-L2-FVIII,
- [0231] 其中V包括VWF片段,
- [0232] L1和L3中的每一个包括任选的接头,
- [0233] L2包括任选的接头例如可切割的接头,
- [0234] L4为任选的接头,例如可加工的接头,
- [0235] FVIII包括FVIII蛋白,
- [0236] X包括一个或多个XTEN序列,
- [0237] F1包括任选的Ig恒定区或其一部分,
- [0238] F2包括任选的另外的Ig恒定区或其一部分;
- [0239] (-) 为肽键或一个或多个氨基酸;且
- [0240] (:) 为化学缔合或物理缔合。

[0241] 在一些实施方案中,呈本文公开的任何构建体或式的FVIII蛋白还可包含至少一个、至少两个、至少三个、至少四个、至少五个或至少六个XTEN序列,每一个XTEN序列紧邻FVIII蛋白中一个或多个氨基酸的下游插入或连接于FVIII蛋白的N-末端或C-末端。本文别处公开了XTEN插入位点的非限制性实例。

[0242] 在一个实施方案中,(:)表示化学缔合,例如,至少一个非肽键。在另一个实施方案中,化学缔合即(:)为共价键。在其它实施方案中,化学缔合即(:)为非共价相互作用,例如

离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在其它实施方案中, (:) 为非肽共价键。在又其它实施方案中, (:) 为肽键。在另外其它实施方案中, (:) 表示在两个序列之间的物理缔合, 其中第一序列的一部分与第二序列紧密相邻以使第一序列掩蔽或阻断第二序列的一部分与另一部分相互作用, 以及进一步在不允许第二序列与其它部分相互作用的情况下保持这种物理缔合。本文多肽式的取向从N-末端(左)至C-末端(右)列出。例如, 式(n) F1-L1-X-L2-V-L4-F2-L2-FVIII意指式NH₂-F1-L1-X-L2-V-L4-F2-L2-FVIII-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任何另外的序列。在另一个实施方案中, 连字号(-)指示肽键。

[0243] 在一个实施方案中, 连接于VWF片段或FVIII蛋白的Ig恒定区或其一部分(有时在本文由“F”或“F1”表示)和另外的Ig恒定区或其一部分(有时在本文由“F2”表示)中的任一个或两个可延长VWF片段、FVIII蛋白或两者的半衰期。在另一个实施方案中, 成对的Ig恒定区或其一部分(有时在本文由“F”或“F1”表示)和另外的Ig恒定区或其一部分(有时在本文由“F2”表示)(其每一个连接于VWF片段和FVIII蛋白)提供比FVIII蛋白和VWF片段之间的非共价键更强的键, 即共价键, 例如二硫键, 从而阻止内源性VWF体内替换VWF片段。F1或F2可包括Fc区或FcRn结合配偶体。在其它实施方案中, 连接于VWF片段和/或FVIII蛋白的F1和F2中的任一个或两个在F1和F2之间形成共价键(例如, 二硫键), 从而替换紧密相邻的VWF片段和FVIII蛋白以阻止FVIII蛋白与VWF片段的相互作用。在一些实施方案中, F1和F2相同或不同。F1和F2的非限制性实例可以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、其任何功能片段、衍生物或类似物和其两个或更多个的组合。在一个实施方案中, F1、F2或两者包括至少一个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、衍生物或类似物。在另一个实施方案中, F1、F2或两者包括至少一个铰链结构域或其部分和至少一个CH2结构域或其部分(例如, 以铰链-CH2取向)。在其它实施方案中, F1、F2或两者包括至少一个CH2结构域或其部分和至少一个CH3结构域或其部分(例如, 以CH2-CH3取向)。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, 其也称为Fc区(或Fc结构域), 例如, F1的第一Fc区或第一FcRn结合配偶体和F2的第二Fc区或第二FcRn结合配偶体。在其它实施方案中, F1通过接头连接于VWF片段, 和/或F2通过接头连接于FVIII蛋白。在一些实施方案中, F1和/或F2包含以下、基本上由以下组成或由以下组成: 铰链区。Fc区或FcRn结合配偶体的另外的非限制性实例在本文别处描述。

[0244] 在某些实施方案中, 本发明的嵌合蛋白包含两条多肽链, 第一多肽链包含以下、基本上由以下组成或由以下组成: 包含D'结构域和D3结构域的VWF片段、XTEN序列、第一Ig恒定区或其一部分(例如, 第一Fc区)、和VWF片段与XTEN序列或XTEN序列或第一Ig恒定区或其一部分之间的任选的接头, 并且第二多肽链包含以下、基本上由以下组成或由以下组成: FVIII蛋白和第二Ig恒定区或其一部分(例如, 第二Fc区)。VWF片段和第一Ig恒定区或其一部分之间的接头可以为可切割的接头, 例如凝血酶可切割的接头, 其可在凝固位点被切割。在一些实施方案中, 第一多肽链和第二多肽链彼此缔合。第一链和第二链之间的缔合阻止包含VWF片段的第一链体内替换为内源性VWF。在一个实施方案中, 第一链和第二链之间的缔合可以为共价键。在特定实施方案中, 共价键为二硫键。在一些实施方案中, 第二链中的FVIII蛋白还包含连接于FVIII蛋白的C-末端或N-末端或紧邻FVIII蛋白中一个或多个氨基酸(例如, 本文公开的至少一个插入位点)的下游插入的一个或多个XTEN序列。插入位点的

非限制性实例在本文别处描述。

[0245] 在其它实施方案中,本发明的嵌合蛋白包含三条多肽链,其中第一多肽链包含以下、基本上由以下组成或由以下组成:FVIII蛋白的重链;第二多肽链包含以下、基本上由以下组成或由以下组成:融合于第一Ig恒定区或其一部分(例如第一Fc区)的FVIII蛋白的轻链;并且第三多肽链包含以下、基本上由以下组成或由以下组成:包含D'结构域和D3结构域的VWF片段、XTEN序列、第二Ig恒定区或其一部分(例如,第二Fc区)、和XTEN序列和第二Ig恒定区或其一部分或VWF片段和XTEN序列之间的任选的接头。第三链中的接头可以为可切割的接头,其在凝固位点被切割,例如凝血酶切割位点。在一些实施方案中,重链FVIII或轻链FVIII连接于一个或多个XTEN序列,所述XTEN序列可连接于N-末端、C-末端或在FVIII序列中的一个或多个插入位点内插入。本文别处公开了插入位点的非限制性实例。

[0246] 在另外其它实施方案中,本发明的嵌合蛋白包含两条多肽链,第一多肽链包含以下、基本上由以下组成或由以下组成:FVIII蛋白的重链;且第二多肽链包含以下、基本上由以下组成或由以下组成:FVIII蛋白的轻链、第一Ig恒定区或其一部分(例如第一Fc区)、第一接头(例如,可加工的接头,其含有包含一个或多个胞内加工位点的一个或多个蛋白酶切割位点)、VWF片段、第二接头(例如,凝血酶可切割的接头)、XTEN序列和第二Ig恒定区或其一部分(例如,第二Fc区),其中FVIII蛋白的轻链连接于第一Ig恒定区或其一部分(例如,第一Fc区),第一Ig恒定区或其一部分通过第一接头进一步连接于VWF片段,并且其中VWF片段连接于XTEN序列,XTEN序列通过第二接头进一步连接于第二Ig恒定区或其一部分。在某些实施方案中,第一接头为可加工的接头,并且第二接头为可切割的接头。在表达时,嵌合蛋白可通过胞内加工酶加工,其切割可加工的接头,并且因此嵌合蛋白可包含以下、基本上由以下组成或由以下组成:三条多肽链。此外,VWF片段可由于可切割的接头在凝固位点被切下。

[0247] 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Ig恒定区或其一部分(例如,第一Fc区)、第一接头(例如,可加工的接头)、VWF片段、XTEN序列、第二接头(例如,凝血酶可切割的接头),和第二Ig恒定区或其一部分(例如,第二Fc区),其中单链FVIII蛋白连接于第一Ig恒定区或其一部分,第一Ig恒定区或其一部分通过第一接头还连接于VWF片段,并且VWF片段连接于XTEN序列,XTEN序列进一步连接于第二Ig恒定区或其一部分。在一个实施方案中,VWF片段和XTEN序列通过第二接头连接。在另一个实施方案中,XTEN序列和第二Ig恒定区或其一部分通过第二接头连接。在其它实施方案中,第二链还包含第三接头。单一多肽链可因而包含通过第二接头连接于XTEN序列的VWF片段和通过第三接头连接于第二Ig恒定区或其一部分的XTEN。第二接头和第三接头可以相同或不同。在一个实施方案中,第一接头为可加工的接头。在另一个实施方案中,第二接头或第三接头为包含一个或两个可切割的位点的可切割的接头。在具体实施方案中,第二接头为凝血酶可切割的接头。可用于本发明的接头在本文别处描述。

[0248] (2) FVIII、XTEN和Fc

[0249] 本发明的嵌合蛋白还包含(i) FVIII蛋白,(ii) XTEN序列(第一半衰期延长剂),和(iii) Ig恒定区或其一部分(第二半衰期延长剂),其中XTEN序列通过任选的接头连接于FVIII蛋白和通过另外的任选的接头连接于Ig恒定区或其一部分。可以一起使用XTEN序列和Ig恒定区或其一部分来延长FVIII蛋白的半衰期。在一个实施方案中,嵌合蛋白为单体。

在另一个实施方案中,嵌合蛋白为二聚体(同二聚体或异二聚体)。

[0250] 本发明还涉及嵌合蛋白,其包含(i) FVIII蛋白,(ii) XTEN序列,(iii) Ig恒定区或其一部分(即,第一Ig恒定区或其一部分,“F”或“F1”),和(iv) 另外的Ig恒定区或其一部分(即,第二Ig恒定区或其一部分或“F2”)。在一个实施方案中,XTEN序列在C-末端或N-末端连接于FVIII蛋白或紧邻FVIII蛋白中的一个或多个氨基酸(例如,一个或多个XTEN插入位点)下游插入,FVIII蛋白连接于第一Ig恒定区或其一部分,并且第一Ig恒定区或其一部分和第二Ig恒定区或其一部分通过任选的接头彼此缔合或连接。在某些方面,嵌合蛋白为单体-二聚体杂合体,其包含第一多肽链和第二多肽链,其中第一多肽链包含FVIII蛋白、XTEN序列和第一Ig恒定区或其一部分,并且第二多肽链包含以下、基本上由以下组成或由以下组成:不含FVIII蛋白的第二Ig恒定区或其一部分,并且其中第一链和第二链彼此缔合。Ig恒定区或其一部分(例如第一Fc区)和另外的Ig恒定区或其一部分(例如,第二Fc区)之间的缔合为化学缔合或物理缔合。在某些实施方案中,化学缔合为共价键。在其它实施方案中,化学缔合为非共价相互作用,例如离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在其它实施方案中,缔合为非肽共价键。在又其它实施方案中,缔合为肽键。

[0251] 在其它方面,嵌合蛋白为单一多肽链,其包含FVIII蛋白、XTEN序列、第一Ig恒定区或其一部分、接头例如可加工的接头、和第二Ig恒定区或其一部分,其中单一多肽链在表达后通过胞内酶被加工且变成两条多肽链。

[0252] 在一个实施方案中,连接于FVIII蛋白的Ig恒定区或其一部分(有时在本文由“F”或“F1”表示)可与XTEN序列一起延长FVIII蛋白的半衰期。在另一个实施方案中,Ig恒定区或其一部分(“F”或“F1”)是在本文别处描述的Fc区或FcRn结合配偶体。

[0253] 在其它实施方案中,与第一Ig恒定区或其一部分缔合或连接的另外的Ig恒定区或其一部分(有时在本文由“F2”表示或第二Ig恒定区或其一部分)还可延长FVIII蛋白的半衰期。在其它实施方案中,第二Ig恒定区或其一部分(“F2”)与第一Ig恒定区或其一部分和XTEN序列一起可延长FVIII蛋白的半衰期。另外的Ig恒定区或其一部分可以为在本文别处描述的Fc区或FcRn结合配偶体。

[0254] 在某些实施方案中,与第一Ig恒定区或其一部分缔合的第二Ig恒定区或其一部分进一步连接于在本文别处描述的VWF片段和任选的XTEN序列。

[0255] 在一些实施方案中,Ig恒定区或其一部分(“F”或“F1”或第一Ig恒定区或其一部分)和另外的Ig恒定区或其一部分(即,第二Ig恒定区或其一部分或“F2”)(在该段落中以“Ig恒定区或其部分”表示)中的任一个或两个可包括但不限于,CH1结构域、CH2结构域、CH3结构域、CH4结构域、铰链结构域、其任何功能片段、衍生物或类似物或其两个或更多个组合。在一个实施方案中,Ig恒定区或其一部分包含至少一个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、衍生物或类似物。在另一个实施方案中,Ig恒定区或其一部分包含至少一个铰链结构域或其部分和至少一个CH2结构域或其部分(例如,以铰链-CH2取向)。在其它实施方案中,Ig恒定结构域或其部分包含至少一个CH2结构域或其部分和至少一个CH3结构域或其部分(例如,以CF2-CH3取向)。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,其也称为Fc区(或Fc结构域),例如,第一Fc区。Ig恒定区或其部分的另外实例在本文别处描述。

[0256] 本发明的嵌合蛋白可具有相较于野生型FVIII延长的FVIII蛋白的半衰期。在一个

实施方案中,FVIII蛋白的半衰期比野生型FVIII的半衰期延长至少约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。在另一个实施方案中,FVIII蛋白的半衰期为至少约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。

[0257] (3) FVIII、XTEN和VWF

[0258] 一方面,本发明的嵌合蛋白包含(i) FVIII蛋白,(ii) XTEN序列,和(iii) 包含VWF的D' 结构域和D3结构域的VWF片段,其中FVIII蛋白连接于XTEN序列并且其中FVIII蛋白与VWF片段缔合或连接。在一个实施方案中,本文所述的嵌合蛋白的VWF片段不能够结合于VWF清除受体。在另一个实施方案中,VWF片段能够保护FVIII蛋白免遭一种或多种蛋白酶切割,保护FVIII蛋白免遭活化,稳定FVIII蛋白的重链和/或轻链,或阻止一个或多个清除剂受体对FVIII蛋白的清除。在其它实施方案中,VWF片段阻止或抑制内源性VWF与FVIII蛋白中的VWF结合位点结合。VWF结合位点可位于FVIII蛋白的A3结构域或C2结构域或者A3结构域和C2结构域两者。在具体实施方案中,VWF结合位点包含对应于SEQ ID NO:2的氨基酸1669至1689和/或氨基酸2303至2332的氨基酸序列。

[0259] 在另一方面,嵌合蛋白包含(i) FVIII蛋白,(ii) XTEN序列,(iii) 包含VWF的D' 结构域和D3结构域的VWF片段,和(iv) Ig恒定区或其一部分,其中XTEN序列在C-末端或N-末端连接于FVIII蛋白或紧邻FVIII蛋白中一个或多个氨基酸(例如,本文公开的一个或多个XTEN插入位点)的下游插入,VWF片段连接于FVIII蛋白或XTEN序列或与FVIII蛋白或XTEN序列缔合,并且Ig恒定区或其一部分连接于FVIII蛋白、XTEN序列、VWF片段或其任意组合。可用于本发明的嵌合蛋白的Ig恒定区或其一部分在本文别处描述。在一个实施方案中,Ig恒定区或其一部分能够延长FVIII蛋白的半衰期。在另一个实施方案中,Ig恒定区或其一部分包含第一Fc区或第一FcRn结合配偶体。在另外其它实施方案中,Ig恒定区或其一部分通过任选的接头连接于FVIII蛋白。在又其它实施方案中,接头包括可切割的接头。嵌合蛋白可以为含有(i)、(ii)、(iii)和(iv)的单一多肽链,即单体(即,单链),或为含有包含(i)和(ii)的第一链和包含(iii)和(iv)的第二链的两条链。在其它方面,嵌合蛋白为二聚体(例如,同二聚体或异二聚体)。在一个实施方案中,嵌合蛋白包含两条链,各包含(i)、(ii)、(iii)和(iv)。

[0260] 在某些实施方案中,嵌合蛋白包含(i) FVIII蛋白,(ii) XTEN序列,(iii) 包含VWF的D' 结构域和D3结构域的VWF片段,(iv) Ig恒定区或其一部分(有时也表示为"F"、“第一Ig恒定区或其一部分”、或“F2”),和(v) 另外的Ig恒定区或其一部分(有时也表示为“F2”或“第二Ig恒定区或其一部分”),其中(1) FVIII蛋白在FVIII蛋白的C-末端或N-末端连接于XTEN序列或紧邻FVIII蛋白中一个或多个氨基酸(例如,本文公开的一个或多个XTEN插入位点)的下游插入,(2) XTEN序列或FVIII蛋白连接于Ig恒定区或其一部分,(3) VWF片段连接于第二Ig恒定区或其一部分,和(4) Ig恒定区或其一部分与第二Ig恒定区或其一部分缔合。在一个实施方案中,连接于FVIII蛋白或XTEN序列的Ig恒定区或其一部分通过接头例如可加工的接头进一步连接于VWF片段。在另一个实施方案中,可用于本发明的嵌合蛋白的另外的Ig恒定

区或其一部分可通过任选的接头例如可加工的接头进一步连接于FVIII蛋白或Ig恒定区或其一部分。在一些实施方案中,成对的Ig恒定区或其一部分和另外的Ig恒定区或其一部分(其每一个连接于VWF片段和FVIII蛋白)提供比FVIII蛋白和VWF片段之间的非共价键更强的键,即共价键,例如二硫键,从而阻止内源性VWF体内替换VWF片段。在其它实施方案中,Ig恒定区或其一部分和另外的Ig恒定区或其一部分中的任一个或两个能够延长FVIII蛋白或VWF片段的半衰期。在其它实施方案中,另外的Ig恒定区或其一部分包含第二Fc区或FcRn结合配偶体。嵌合蛋白中的Ig恒定区或其一部分和另外的Ig恒定区或其一部分相同或不同。

[0261] 在某些实施方案中,Ig恒定区或其一部分和另外的Ig恒定区或其一部分通过化学缔合或物理缔合而缔合。在一个实施方案中,化学缔合即(:)为至少一个非肽键。在某些实施方案中,化学缔合即(:)为共价键。在其它实施方案中,化学缔合即(:)为非共价相互作用,例如离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在其它实施方案中,(:)为非肽共价键。在又其它实施方案中,(:)为肽键。在另外其它实施方案中,(:)表示在两个序列之间的物理缔合,其中第一序列的一部分与第二序列紧密相邻以使第一序列掩蔽或阻断第二序列的一部分与另一部分相互作用。在一些实施方案中,Ig恒定区或其一部分和另外的Ig恒定区或其一部分之间的缔合可以为共价键,例如二硫键,其阻止VWF片段或含有VWF片段的多肽替换为内源性VWF。因此,阻止FVIII蛋白与内源性VWF之间的相互作用减少或消除FVIII蛋白的这个半衰期限制因子,并且因而FVIII蛋白的半衰期相较于不含VWF蛋白的FVIII蛋白或野生型FVIII有所延长。

[0262] 在其它方面,嵌合蛋白包含包括以下的式:

[0263] (1) FVIII (X1) -L1-F1:V-L2-X2-L3-F2;

[0264] (2) FVIII (X1) -L1-F1:F2-L3-X2-L2-V;

[0265] (3) F1-L1-FVIII (X1) :V-L2-X2-L3-F2;

[0266] (4) F1-L1-FVIII (X1) :F2-L3-X2-L2-V;

[0267] (5) FVIII (X1) -L1-F1-L4-V-L2-X2-L3-F2;

[0268] (6) FVIII (X1) -L1-F1-L4-F2-L3-X2-L2-V;

[0269] (7) F1-L1-FVIII (X1) -L4-V-L2-X2-L3-F2或

[0270] (8) F1-L1-FVIII (X1) -L4-F2-L3-X2-L2-V,

[0271] 其中FVIII (X1) 包含FVIII蛋白和一个或多个XTEN序列,其中一个或多个XTEN序列连接于FVIII蛋白的N-末端或C-末端或紧邻FVIII蛋白中一个或多个氨基酸(例如,本文公开的一个或多个XTEN插入位点)的下游插入;

[0272] L1、L2或L3中的每一个包括任选的接头,例如可切割的接头;

[0273] L4为接头,例如可加工的接头;

[0274] X2包括一个或多个任选的XTEN序列;

[0275] F1包括Ig恒定区或其一部分;

[0276] F2包括任选的另外的Ig恒定区或其一部分,且

[0277] V包括VWF片段;

[0278] (-) 为肽键或一个或多个氨基酸;且

[0279] (:) 包括化学缔合或物理缔合。在一个实施方案中,(:) 表示化学缔合,例如,至少一个非肽键。在另一个实施方案中,化学缔合即(:) 为共价键。在其它实施方案中,化学缔合

即(:)为非共价相互作用,例如离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在其它实施方案中,(:)为非肽共价键。在又其它实施方案中,(:)为肽键。在另外其它实施方案中,(:)表示在两个序列之间的物理缔合,其中第一序列的一部分与第二序列紧密相邻以使第一序列掩蔽或阻断第二序列的一部分与另一部分相互作用,以及进一步在不允许第二序列与其它部分相互作用的情况下保持这种物理缔合。本文多肽式的取向以N-末端(左)至C-末端(右)列出。例如,式V-X-FVIII意指式NH₂-V-X-FVIII-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任何另外的序列。例如,式V-X-FVIII还可包含在V与X之间、X与FVIII之间的V的N-末端、或在FVIII的C-末端的任何序列,除非另有规定。在另一个实施方案中,连字号(-)指示肽键。

[0280] 一方面,嵌合蛋白包含两条多肽链,(A)第一链包含(i)单链FVIII蛋白(ii)XTEN序列,和(iii)第一Ig恒定区或其一部分,例如第一Fc区或FcRn结合配偶体,其中XTEN序列在N-末端或C-末端连接于FVIII蛋白或紧邻FVIII蛋白的一个或多个氨基酸(例如,本文公开的一个或多个XTEN插入位点)的下游插入并且第一Ig恒定区或其一部分当XTEN序列在N-末端或C-末端连接于FVIII蛋白时连接于XTEN序列或当XTEN序列在FVIII蛋白内插入时连接于FVIII蛋白,和(B)第二链包含(iv)含有D'结构域和D3结构域的VWF片段,(v)接头,和(vi)第二Ig恒定区或其一部分,例如第二Fc区或第二FcRn结合配偶体,其中VWF片段连接于接头例如可切割的接头,所述接头进一步连接于第二Ig恒定区或其一部分,并且其中第一多肽链和第二多肽链彼此缔合,例如共价键,例如二硫键。在一个实施方案中,接头为在本文别处描述的可切割的接头,例如凝血酶可切割的接头。在一些实施方案中,第二链包含一个或多个在(iv)和(v)或(v)和(vi)之间的XTEN序列。

[0281] 在其它方面,嵌合蛋白包含一条多肽链,其包含(i)单链FVIII蛋白(ii)XTEN序列,(iii)第一Ig恒定区或其一部分,例如第一Fc区或第一FcRn结合配偶体,(iv)第一接头,(v)包含D'结构域和D3结构域的VWF片段、(vi)第二接头和(vii)第二Ig恒定区或其一部分,例如,第二Fc区或第二FcRn结合配偶体,其中(i)至(vii)以该顺序或以任何顺序连接。在一个实施方案中,第一接头为可加工的接头,其可在表达后被细胞内加工或切割且将单一多肽链制成两条多肽链。在另一个实施方案中,第二接头为本文所述的可切割的接头,例如凝血酶可切割的接头。本文所使用的XTEN序列可通过任选的接头在FVIII蛋白的N-末端或C末端连接于FVIII蛋白或紧邻FVIII蛋白中一个或多个氨基酸(例如,一个或多个XTEN插入位点)的下游插入。

[0282] 在某些方面,嵌合蛋白包含三条多肽链,(A)第一多肽链包含彼此连接的(i)FVIII蛋白的重链和(ii)XTEN序列,并且(B)第二多肽链包含彼此连接的(iii)FVIII蛋白的轻链和(iv)第一Ig恒定区或其一部分,例如第一Fc区或第一FcRn结合配偶体,和(C)第三多肽链包含(v)包含D'结构域和D3结构域的VWF片段、(vi)接头和(vii)第二Ig恒定区或其一部分,例如,第二Fc区或第二FcRn结合配偶体,其中第二链与第一链和第三链缔合。在一个实施方案中,第一链和第二链之间的缔合为化学缔合或物理缔合。例如,第一链和第二链之间的缔合可以为金属键。在另一个实施方案中,第二链和第三链之间的缔合也为化学缔合或物理缔合,例如共价键或非共价键。在某些实施方案中,第二链和第三链之间的缔合是通过两个Ig恒定区或其一部分且为二硫键。第二链和第三链之间的键合阻止或抑制FVIII蛋白与内源性VWF结合,因而阻止FVIII蛋白通过VWF清除途径被清除。在一些实施方案中,接头为可

加工的接头,其在宿主细胞中表达后被细胞内切割。本文所使用的XTEN序列通过任选的接头在FVIII蛋白的N-末端或C末端连接于FVIII蛋白或紧邻FVIII蛋白中一个或多个氨基酸(例如,一个或多个XTEN插入位点)的下游插入。

[0283] 在某些实施方案中,VWF片段通过肽键或接头直接连接于FVIII蛋白,所述FVIII蛋白包含一个或多个XTEN。作为通过直接连接(例如肽键)或接头连接VWF片段和其中一个或多个XTEN插入或连接的FVIII蛋白的一种方式,可采用酶连接(例如,分选酶)。例如,分选酶是指一组通过识别并切割羧基末端分选信号来修饰表面蛋白的原核酶。对于分选酶的大多数底物,识别信号由基序LPXTG (Leu-Pro-any-Thr-Gly (SEQ ID NO:51)、然后高疏水性跨膜序列、然后一簇碱性残基诸如精氨酸组成。切割在Thr与Gly之间发生,其中通过Thr残基短暂连接至连接配偶体的活性位点Cys残基,然后进行将蛋白质共价连接至细胞壁的转肽作用。在一些实施方案中,连接配偶体含有Gly (n)。在其它实施方案中,嵌合蛋白还包含分选酶识别基序。在一些实施方案中,VWF片段连接至FVIII,所述FVIII包含一个或多个在蛋白连接插入或使用体外介导的分选酶连接于蛋白连接的XTEN。

[0284] 在一个实施方案中,通过任选的接头连接于分选酶识别基序的VWF片段可融合于通过分选酶连接于Gly (n)的FVIII蛋白,其中n可以为任意整数并且其中一个或多个XTEN在FVIII蛋白内插入或连接于FVIII蛋白。连接构建体包含VWF片段(构建体的N-端部分)和其中一个或多个XTEN被插入或连接的FVIII蛋白(构建体的C-端部分),其中分选酶识别基序插入于两者之间。另一连接构建体包含VWF片段(构建体的N-端部分、接头、分选酶识别基序和其中一个或多个XTEN被插入或连接的FVIII蛋白(构建体的C-端部分)。在另一个实施方案中,通过任选的接头连接于分选酶识别基序的FVIII蛋白可融合于通过分选酶连接于Gly (n)的VWF片段,其中n为任意整数。所得的连接构建体包含其中一个或多个XTEN被插入或连接的FVIII蛋白(构建体的N-端部分),和VWF片段(构建体的C-端部分),其中分选酶识别基序插入于两者之间。另一所得的连接构建体包含其中一个或多个XTEN被插入或连接的FVIII蛋白(构建体的N-端部分)、接头、分选酶识别基序和VWF片段(构建体的C-端部分)。在其它实施方案中,通过第一任选的接头连接于分选酶识别基序的VWF片段可融合于通过第二任选的接头连接于凝血酶切割位点的异源部分,例如免疫球蛋白恒定区或其一部分,例如Fc区。所得的构建体可包含VWF片段(N-端部分)、第一接头、分选酶识别基序、蛋白酶切割位点、第二任选的接头和异源部分。

[0285] 在一些实施方案中,VWF片段与FVIII蛋白缔合。VWF片段和FVIII蛋白之间的缔合可以为化学缔合或物理缔合。化学缔合可以为非共价相互作用,例如离子相互作用、疏水性相互作用、亲水性相互作用、范德华相互作用或氢键。在另外其它实施方案中,FVIII蛋白和VWF片段之间的缔合为在两个序列之间的物理缔合,例如,由于具有FVIII蛋白的序列与具有VWF片段的序列的另外缔合,其中第一序列的一部分与第二序列紧密相邻以使第一序列掩蔽或阻断第二序列的一部分与另一部分相互作用。

[0286] 由于通过VWF片段阻止或抑制内源性VWF与FVIII蛋白相互作用,本文所述的嵌合蛋白相较于野生型FVIII或不含VWF片段的相应的嵌合蛋白具有延长的半衰期。在一个实施方案中,FVIII蛋白的半衰期比不含VWF片段的FVIII蛋白延长至少约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。在另一个实施方案中,FVIII蛋白的半衰期为至

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。在特定实施方案中，在HemA小鼠中FVIII蛋白的半衰期延长至少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25小时、至少约26小时或至少约27小时。

[0287] A) 血管性血友病因子 (VWF) 片段

[0288] VWF (也称为F8VWF) 为一种存在于血浆中的大的多聚体糖蛋白并且在内皮 (于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组成性生成。基本VWF单体为2813个氨基酸的蛋白质。每一个单体含有许多具有特定功能的特定结构域，D'/D3结构域 (其结合于因子VIII)、A1结构域 (其结合于血小板GPIb-受体、肝素和/或可能结合于胶原蛋白)、A3结构域 (其结合于胶原蛋白)、C1结构域 (其中RGD结构域当其被活化时结合于血小板整合素 α IIb β 3) 和在蛋白质的C-端的“半胱氨酸结”结构域 (其VWF共有血小板衍生生长因子 (PDGF)、转化生长因子- β (TGF β) 和 β -人绒毛膜促性腺激素 (hCG))。

[0289] 如本文所使用的术语“VWF片段”包括但不限于，包含D'结构域和D3结构域的功能性VWF片段，其能够抑制内源性VWF与FVIII结合。在一个实施方案中，VWF片段结合于FVIII蛋白。在另一个实施方案中，VWF片段封闭FVIII蛋白上的VWF结合位点，从而抑制FVIII蛋白与内源性VWF的相互作用。VWF片段包括保留VWF的这些活性的衍生物、变体、突变体或类似物。

[0290] 人VWF的2813单体氨基酸序列在Genbank中以登录号_NP_000543.2_报道。编码人VWF的核苷酸序列在Genbank中以登录号__NM_000552.3_报道。人VWF的核苷酸序列被指定为SEQ ID NO:1。SEQ ID NO:2为由SEQ ID NO:1编码的氨基酸序列。VWF的各结构域于表1中列出。

[0291] 表1.VWF序列

[0292]

VWF 结构域	氨基酸序列		
VWF 信号肽 (SEQ ID NO: 2 的 氨基酸 1 至 22)	1	<u>MIPARFAGVL LALALILPGT LC</u>	
	22		
VWF D1D2 区 (SEQ ID NO: 2 的 氨基酸 23 至 763)	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

[0293]

	451	QCADDRDAVC TRSVTVRLPG LHNSLVKLKH GAGVAMDGD IQLPLLKGD RIQHTVTASV RLSYGEDLQM	
	501	DWDGRGRLLV KLSFVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG	
	551	NAWKLHGDCQ DLQKQHSDDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS	
	601	PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL	
	651	NCPKGQVYLQ CGTPCNLTCT SLSYPDEECN EACLEGCFCP PGLYMDERGD	
	701	CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD	
	751	AVLSSPLSHR	SKR
	763		
VWF D' 结构域	764	<u>SLSCRPP MVKLVCADN</u> <u>LRAEGLECTK</u> <u>TCONYDLECM</u>	
	801	<u>SMGCVSGCLC PPGMVRHENR CVALERCPCF</u> <u>HQKEYAPGE TVKIGCNTCV</u>	
	851	<u>CRDRKWNCTD</u>	HVCDAT
	866		
VWF D3 结构域	867	<u>CSTI GMAHYLTFDG</u> <u>LKYLFPGEQ YVLVQDYCGS</u>	
	901	<u>NPGTFRILVG NKGCSHPSVK CKKRVTLVE</u> <u>GGEIELFDGE VNVKRPMKDE</u>	
	951	<u>THFEVVESGR YIILLGKAL SVVWDRHLSI</u> <u>SVVLKQTYQE KVCGLCGNFD</u>	
	1001	<u>GIQNDLTSS NLQVEEDPVD FGNSWKVSSQ</u> <u>CADTRKVPLD SSPATCHNNI</u>	
	1051	<u>MKQTMVDSSC RILTSDFQD CNKLVDPEPY</u> <u>LDVCIYDTCS CESIGDCACF</u>	
	1101	<u>CDTIAAYAHV CAQHKGKVVW RTATLCPQSC</u> <u>EERNLRENGY ECEWRYNSCA</u>	
	1151	<u>PACQVTCQHP EPLACPVQCV EGCHAHCPPG</u> <u>KILDELLQTC VDPEDCPVCE</u>	
	1201	<u>VAGRRFASGK KVTLNPSDPE HCQICHCDVV</u> <u>NLTCEACQEP</u>	
	1240		
VWF A1 结构域	1241	GGLVVPPTDA	
	1251	PVSPTTLYVE DISEPPLHDF YCSRLLDLVF LLDGSSRLSE AEFEVLKAFV	
	1301	VDMMERLRIS QKWVRVAVVE YHDGSHAYIG LKDRKRPSSEL RRIASQVKYA	
	1351	GSQVASTSEV LKYTLFQIFS KIDRPEASRI ALLMASQEP QRMSRNFVRY	
	1401	VQGLKKKKVI VIPVGIGPHA NLKQIRLIEK	

[0294]

	1451	QAPENKAFVL DEIVSYLCDL 1479	SSVDELEQQR APEAPPPTLP PDMAQVTVG
	1480		P
	1501	GLLGVSTLGP FVLEGSDKIG GQDSIHVTVL	KRNSMVLDVA EADFNRSEKF QYSYMTVEY
	1551	PFSEAQSKGD ALRYLSDHSF 1600	ILQRVREIRY QGGNRTNTGL LVSQGDREQA
	1601	PNLVYMTGN GPNANVQELE	PASDEIKRLP RIGWPNAPIL
	1651	IQDFETLPRE SPAPDCSQPL	APDLVLQRCC DVILLLDGSS
	1701	SFPASYFDEM SVLQYGSITT	KSFAKAFISK IDVPWNVPE
	1751	KAHLLSLVDV LTSEMHGARP	MQREGGPSQI GASKAVVILV
	1801	TDVSVDSVDA YDAAQLRILA	AADAARSNRV GPAGDSNVVK
	1851	LQRIEDLPTM DEDGNEKRP	VTLGNSFLHK DVWTLPDQCH
	1901	TVTCQPDGQT QSPVKVEETC	LLKSHRVNCD GCRWTCPCVC
	1951	TGSSTRHIVT EQDLEVILHN	FDGQNFKLGT GACSPGARQG
	2001	CMKSIEVKHS VPYVGGNMEV	ALSVEXHSDM NVYGAIMHEV
	2051	RFNHLGHIFT TYGLCGICDE	FTPQNNEFQL NGANDFMLRD
	2101	GTVTTDWCTL CLVPDSSHQ	VQEWTVQRP VLLLPLFAEC
	2151	HKVLAPATFY AHLCRTNGVC	AICQQDSCHQ VDWRTPDFCA
	2201	MSCPPSLVYN PSEGCFPCPD	HCEHGCPRHC KVMLEGSCVP
	2251	EEACTQCIGE CTCLSGRKVN	DGVQHQFLEA CTTQPCPTAK
	2301	APTCGLCEVA SCDLPPVPHC	RLRQNAQDCC ERGLQPTLTN
	2351	PGECRPNFTC PTLRKTQCCD	ACRKEECKRV EYECACNCVN
	2401	STVSCPLGYL CVHRSTIYPV	ASTATNDCGC GQFWEEGCDV
	2451	CTCTDMEDAV FTYVLHEGEC	MGLRVAQCSQ CGRCLPSACE
	2501	VVTGSPRGDS NECVRVKEEV	QSSWKSVMGSQ FIQQRNVSCP
	2551	QLEVPVCPSP ACMLNGTVIG	FQLSCKTSAC PGKTMIDVC

	2601 TTCRCMVQVG VISGFKLECR KTTNCPCLG 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(SEQ ID NO: 1)
全长VWF	1 ATGATTCCTG CCAGATTTCG 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 TTGCCCCTG CCACCCTCTG GTGGACCCG 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

[0295]

[0296]

```

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 TCGGCCTGT
GTGGGAATTA CAATGGCAAC CAGGGCGACG
1601 ACTTCCTTAC CCCCTCTGGG
CTGGCRGAGC CCCGGGTGGA GGA CTTCGGG
1651 AACGCCTGGA AGCTGCACGG
GGA CTGCCAG 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
GCGTGCGCG AGCCAGGCCG CTGTGAGCTG
1951 AACTGCCCCA AAGGCCAGGT
GTACCTGCAG TCGGGGACCC CCTGCAACCT
2001 GACCTGCCGC TCTCTCTCTT
ACCCGGATGA GGAATGCAAT GAGGCCTGCC
2051 TGGAGGGCTG CTTCTGCCCC
CCAGGGCTCT ACATGGATGA GAGGGGGGAC
2101 TGCGTGCCCA AGGCCAGTG
CCCCTGTTAC TATGACGGTG AGATCTTCCA
2151 GCCAGAAGAC ATCTTCTCAG
ACCATCACAC CATGTGCTAC TGTGAGGATG
2201 GCTTCATGCA CTGTACCATG
AGTGGAGTCC CCGGAAGCTT GCTGCCTGAC

```

[0297]

```

2251  GCTGTCTCA GCAGTCCCCT
      GTCTCATCGC AGCAAAAGGA GCCTATCCTG
2301  TCGGCCCCCC ATGGTCAAGC
      TGGTGTGTCC CGCTGACAAC CTGCGGGCTG
2351  AAGGGCTCGA GTGTACCAA
      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 CTTGACGGG CTCAAATACC
2651  TGTTCCCCGG GGAGTGCCAG
      TACGTTCTGG TGCAGGATTA CTGCGGCAGT
2701  AACCTGGGA 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 CGCCTGCTTC
3301  TGCACACCA TTGCTGCCTA
      TGCCACGTG TGTGCCCAGC ATGGCAAGGT
3351  GGTGACCTGG AGGACGGCCA
      CATTGTGCCC CCAGAGCTGC GAGGAGAGGA
3401  ATCTCCGGGA GAACGGGTAT
      GAGTGTGAGT GGCCTATAA CAGCTGTGCA
3451  CCTGCCTGTC AAGTCACGTG
      TCAGCACCTT GAGCCACTGG CCTGCCCTGT
3501  GCAGTGTGTG GAGGGCTGCC

```


[0298]

```

ATGCCCCTG CCCTCCAGGG AAAATCCTGG
3551 ATGAGCTTTT GCAGACCTGC
GTTGACCCTG AAGACTGTCC AGTGTGTGAG
3601 GTGGCTGGCC GCGTTTTTGC
CTCAGGAAAG AAAGTCACCT TGAATCCCAG
3651 TGACCCTGAG CACTGCCAGA
TTTGCCACTG TGATGTTGTC AACCTCACCT
3701 GTGAAGCCTG CCAGGAGCCG
GGAGGCCTGG TGGTGCCTCC CACAGATGCC
3751 CCGGTGAGCC CCACCACTCT
GTATGTGGAG GACATCTCGG AACCGCCGTT
3801 GCACGATTTT TACTGCAGCA
GGCTACTGGA CCTGGTCTTC CTGCTGGATG
3851 GCTCCTCCAG GCTGTCCGAG
GCTGAGTTTG AAGTGCTGAA GGCCTTTGTG
3901 GTGGACATGA TGGAGCGGCT
GCGCATCTCC CAGAAGTGGG TCCGCGTGGC
3951 CGTGGTGGAG TACCACGACG
GCTCCACGC CTACATCGGG CTCAAGGACC
4001 GGAAGCGACC GTCAGAGCTG
CGGCGCATTG CCAGCCAGGT GAAGTATGCG
4051 GGCAGCCAGG TGGCCTCCAC
CAGCGAGGTC TTGAAATACA CACTGTTCCA
4101 AATCTTCAGC AAGATCGACC
GCCCTGAAGC CTCCCGCATC GCCCTGCTCC
4151 TGATGGCCAG CCAGGAGCCC
CAACGGATGT CCCGGAACCT TGTCCGCTAC
4201 GTCCAGGGCC TGAAGAAGAA
GAAGGTCATT GTGATCCCGG TGGGCATTGG
4251 GCCCCATGCC AACCTCAAGC
AGATCCGCCT CATCGAGAAG CAGGCCCTG
4301 AGAACAAGGC CTTCGTGCTG
AGCAGTGTGG ATGAGCTGGA GCAGCAAAGG
4351 GACGAGATCG TTAGCTACCT
CTGTGACCTT GCCCTGAAG CCCCTCCTCC
4401 TACTCTGCCC CCCGACATGG
CACAAGTCAC TGTGGGCCCC GGGCTCTTGG
4451 GGGTTTCGAC CCTGGGGCCC
AAGAGGAACT CCATGGTTCT GGATGTGGCG
4501 TTCGTCTTGG AAGGATCGGA
CAAAATTGGT GAAGCCGACT TCAACAGGAG
4551 CAAGGAGTTC ATGGAGGAGG
TGATTACGCG GATGGATGTG GGCCAGGACA
4601 GCATCCACGT CACGGTGCTG
CAGTACTCCT ACATGGTGAC CGTGGAGTAC
4651 CCCTTCAGCG AGGCACAGTC
CAAAGGGGAC ATCCTGCAGC GGGTGCGAGA
4701 GATCCGCTAC CAGGGCGGCA
ACAGGACCAA CACTGGGCTG GCCCTGCGGT
4751 ACCTCTCTGA CCACAGCTTC
TTGGTCAGCC AGGGTGACCG GGAGCAGGCG

```

[0299]

```

4801  CCCAACCTGG TCTACATGGT
      CACCGGAAAT CCTGCCTCTG ATGAGATCAA
4851  GAGGCTGCCT GGAGACATCC
      AGGTGGTGCC CATTGGAGTG GGCCCTAATG
4901  CCAACGTGCA GGAGCTGGAG
      AGGATTGGCT GGCCCAATGC CCCTATCCTC
4951  ATCCAGGACT TTGAGACGCT
      CCCCCGAGAG GCTCCTGACC TGGTGCTGCA
5001  GAGGTGCTGC TCCGGAGAGG
      GGCTGCAGAT CCCACCCCTC TCCCCTGCAC
5051  CTGACTGCAG CCAGCCCCTG
      GACGTGATCC TTCTCCTGGA TGGCTCCTCC
5101  AGTTTCCCAG CTTCTTATTT
      TGATGAAATG AAGAGTTTCG CCAAGGCTTT
5151  CATTTCAAAA GCCAATATAG
      GGCCTCGTCT CACTCAGGTG TCAGTGCTGC
5201  AGTATGGAAG CATCACCACC
      ATTGACGTGC CATGGAACGT GGTCCCGGAG
5251  AAAGCCCATT TGCTGAGCCT
      TGTGGACGTC ATGCAGCGGG AGGGAGGCCC
5301  CAGCCAAATC GGGGATGCCT
      TGGGCTTTGC TGTGCGATAC TTGACTTCAG
5351  AAATGCATGG TGCCAGGCCG
      GGAGCCTCAA AGGCGGTGGT CATCCTGGTC
5401  ACGGACGTCT CTGTGGATTG
      AGTGGATGCA GCAGCTGATG CCGCCAGGTC
5451  CAACAGAGTG ACAGTGTTCC
      CTATTGGAAT TGGAGATCGC TACGATGCAG
5501  CCCAGCTACG GATCTTGGCA
      GGCCCAGCAG GCGACTCCAA CGTGGTGAAG
5551  CTCCAGCGAA TCGAAGACCT
      CCCTACCATG GTCACCTTGG GCAATTCCTT
5601  CCTCCACAAA CTGTGCTCTG
      GATTTGTTAG GATTGTCATG GATGAGGATG
5651  GGAATGAGAA GAGGCCCGGG
      GACGTCTGGA CCTTGCCAGA CCAGTGCCAC
5701  ACCGTGACTT GCCAGCCAGA
      TGGCCAGACC TTGCTGAAGA GTCATCGGGT
5751  CAACTGTGAC CGGGGGCTGA
      GGCCTTCGTG CCCTAACAGC CAGTCCCCTG
5801  TTAAAGTGGA AGAGACCTGT
      GGCTGCCGCT GGACCTGCCC CTGYGTGTGC
5851  ACAGGCAGCT CCACTCGGCA
      CATCGTGACC TTTGATGGGC AGAATTTCAA
5901  GCTGACTGGC AGCTGTTCTT
      ATGTCCTATT TCAAAACAAG GAGCAGGACC
5951  TGGAGGTGAT TCTCCATAAT
      GGTGCCTGCA GCCCTGGAGC AAGGCAGGGC
6001  TGCATGAAAT CCATCGAGGT
      GAAGCACAGT GCCCTCTCCG TCGAGSTGCA
6051  CAGTGACATG GAGGTGACGG

```


[0300]

```

TGAATGGGAG ACTGGTCTCT GTTCCTTACG
6101 TGGGTGGGAA CATGGAAGTC
AACGTTTATG GTGCCATCAT GCATGAGGTC
6151 AGATTCAATC ACCTTGGTCA
CATCTTCACA TTCACTCCAC AAAACAATGA
6201 GTTCCAACCTG CAGCTCAGCC
CCAAGACTTT TGCTTCAAAG ACGTATGGTC
6251 TGTGTGGGAT CTGTGATGAG
AACGGAGCCA ATGACTTCAT GCTGAGGGAT
6301 GGCACAGTCA CCACAGACTG
GAAAACACTT GTTCAGGAAT GGAAGTGCA
6351 GCGGCCAGGG CAGACGTGCC
AGCCCATCCT GGAGGAGCAG TGTCTTGTC
6401 CCGACAGCTC CCACTGCCAG
GTCCTCCTCT TACCACTGTT TGCTGAATGC
6451 CACAAGGTCC TGGCTCCAGC
CACATTCTAT GCCATCTGCC AGCAGGACAG
6501 TTGCCACCAG GAGCAAGTGT
GTGAGGTGAT CGCCTCTTAT GCCCACCTCT
6551 GTCGGACCAA CGGGGTCTGC
GTTGACTGGA GGACACCTGA TTTCTGTGCT
6601 ATGTCATGCC CACCATCTCT
GGTCTACAAC CACTGTGAGC ATGGCTGTCC
6651 CCGGCACTGT GATGGCAACG
TGAGCTCCTG TGGGGACCAT CCCTCCGAAG
6701 GCTGTTTCTG CCCTCCAGAT
AAAGTCATGT TGGAAGGCAG CTGTGTCCCT
6751 GAAGAGGCCT GCACTCAGTG
CATTGGTGAG GATGGAGTCC AGCACCAGTT
6801 CCTGGAAGCC TGGGTCCCGG
ACCACCAGCC CTGTGAGATC TGCACATGCC
6851 TCAGCGGGCG GAAGGTCAAC
TGCACAACGC AGCCCTGCCC CACGGCCAAA
6901 GCTCCCACGT GTGGCCTGTG
TGAAGTAGCC CGCCTCCGCC AGAATGCAGA
6951 CCAGTGCTGC CCCGAGTATG
AGTGTGTGTG TGACCCAGTG AGCTGTGACC
7001 TGCCCCCAGT GCCTCACTGT
GAACGTGGCC TCCAGCCCAC ACTGACCAAC
7051 CCTGGCGAGT GCAGACCCAA
CTTCACCTGC GCCTGCAGGA AGGAGGAGTG
7101 CAAAAGAGTG TCCCCACCCT
CCTGCCCCCC GCACCGTTTG CCCACCCTTC
7151 GGAAGACCCA GTGCTGTGAT
GAGTATGAGT GTGCCTGCAA CTGTGTCAAC
7201 TCCACAGTGA GCTGTCCCCT
TGGGTACTTG GCCTCAACCG CCACCAATGA
7251 CTGTGGCTGT ACCACAACCA
CCTGCCTTCC CGACAAGGTG TGTGTCCACC
7301 GAAGCACCAT CTACCCTGTG
GGCCAGTTCT GGGAGGAGGG CTGCGATGTG

```

[0301]

```

7351  TGCACCTGCA CCGACATGGA
      GGATGCCGTG ATGGGCCTCC GCGTGGCCCA
7401  GTGCTCCCAG AAGCCCTGTG
      AGGACAGCTG TCGGTCGGGC TTCACTTACG
7451  TTCTGCATGA AGGCGAGTGC
      TGTGGAAGGT GCCTGCCATC TGCCTGTGAG
7501  GTGGTGACTG GCTCACCGCG
      GGGGGACTCC CAGTCTTCCT GGAAGAGTGT
7551  CGGCTCCCAG TGGGCCTCCC
      CGGAGAACCC CTGCCTCATC AATGAGTGTG
7601  TCCGAGTGAA GGAGGAGGTC
      TTTATACAAC AAAGGAACGT CTCCTGCCCC
7651  CAGCTGGAGG TCCCTGTCTG
      CCCCTCGGGC TTTCAGCTGA GCTGTAAGAC
7701  CTCAGCGTGC TGCCCAAGCT
      GTCGCTGTGA GCGCATGGAG GCCTGCATGC
7751  TCAATGGCAC TGTCATTGGG
      CCCGGGAAGA CTGTGATGAT CGATGTGTGC
7801  ACGACCTGCC GCTGCATGGT
      GCAGGTGGGG GTCATCTCTG GATTCAAGCT
7851  GGAGTGCAGG AAGACCACCT
      GCAACCCCTG CCCCCTGGGT TACAAGGAAG
7901  AAAATAACAC AGGTGAATGT
      TGTGGGAGAT GTTTGCCTAC GGCTTGCACC
7951  ATTCAGCTAA GAGGAGGACA
      GATCATGACA CTGAAGCGTG ATGAGACGCT
8001  CCAGGATGGC TGTGATACTC
      ACTTCTGCAA GGTCAATGAG AGAGGAGAGT
8051  ACTTCTGGGA GAAGAGGGTC
      ACAGGCTGCC CACCCTTTGA TGAACACAAG
8101  TGTCTTGCTG AGGGAGGTAA
      AATTATGAAA ATTCCAGGCA CCTGCTGTGA
8151  CACATGTGAG GAGCCTGAGT
      GCAACGACAT CACTGCCAGG CTGCAGTATG
8201  TCAAGGTGGG AAGCTGTAAG
      TCTGAAGTAG AGGTGGATAT CCACTACTGC
8251  CAGGGCAAAT GTGCCAGCAA
      AGCCATGTAC TCCATTGACA TCAACGATGT
8301  GCAGGACCAG TGCTCCTGCT
      GCTCTCCGAC ACGGACGGAG CCCATGCAGG
8351  TGGCCCTGCA CTGCACCAAT
      GGCTCTGTTG TGTACCATGA GGTCTCAAT
8401  GCCATGGAGT GCAAATGCTC
      CCCCAGGAAG TGCAGCAAGT GA

```

[0302] 如本文所使用的VWF片段可以是包含VWF的D' 结构域和D3结构域的VWF片段,其中VWF片段结合于因子VIII (FVIII) 并且抑制内源性VWF (全长VWF) 与FVIII结合。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、其一个或多个片段及其任意组合。在一个实施方案中,VWF片段包含以下、

基本上由以下组成或由以下组成：(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。本文所述的VWF片段不含结合于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的氨基酸764至1274。本发明的VWF片段可包含连接于或融合于VWF片段的任何其它序列。例如，本文所述的VWF片段还可包含信号肽。

[0303] 在一个实施方案中，VWF片段结合于FVIII蛋白或与FVIII蛋白缔合。通过结合于FVIII蛋白或与FVIII蛋白缔合，本发明的VWF片段保护FVIII免遭蛋白酶切割和FVIII活化，稳定FVIII的重链和轻链，以及阻止清除剂受体对FVIII的清除。在另一个实施方案中，VWF片段结合于FVIII蛋白或与FVIII蛋白缔合并且阻断或阻止FVIII蛋白与磷脂和活化蛋白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的清除并且因而延长FVIII蛋白的半衰期。在一个实施方案中，FVIII蛋白的半衰期延长因而归因于缺少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片段掩蔽或保护FVIII蛋白免遭含有VWF清除受体结合位点的内源性VWF。结合于或受VWF片段保护的FVIII蛋白还可允许FVIII蛋白的再循环。通过消除全长VWF分子中含有的VWF清除途径受体结合位点，掩蔽本发明的FVIII/VWF异二聚体免遭VWF清除途径，进一步延长FVIII半衰期。

[0304] 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的氨基酸764至866具有至少60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性，其中VWF片段阻止内源性VWF与FVIII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的氨基酸867至1240具有至少60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性，其中VWF片段阻止内源性VWF与FVIII结合。在一些实施方案中，本文所述的VWF片段包含以下、基本上由以下组成或由以下组成：VWF的D' 结构域和D3结构域，其与SEQ ID NO:2的氨基酸764至1240具有至少60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性，其中VWF片段阻止内源性VWF与FVIII结合。在其它实施方案中，VWF片段包含以下、基本上由以下组成或由以下组成：与SEQ ID NO:2的氨基酸23至1240具有至少60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的D1、D2、D' 和D3结构域，其中VWF片段阻止内源性VWF与FVIII结合。在又其它实施方案中，VWF片段还包含与其可操作地连接的信号肽。

[0305] 在一些实施方案中，本发明的VWF片段基本上由以下组成或由以下组成：(1) D'D3结构域、D1D'D3结构域、D2D'D3结构域或D1D2D'D3结构域和(2) 多达约10个氨基酸(例如，来自SEQ ID NO:2的氨基酸764至1240至SEQ ID NO:2的氨基酸764至1250的任何序列)、多达约15个氨基酸(例如，来自SEQ ID NO:2的氨基酸764至1240至SEQ ID NO:2的氨基酸764至1255的任何序列)、多达约20个氨基酸(例如，来自SEQ ID NO:2的氨基酸764至1240至SEQ ID NO:2的氨基酸764至1260的任何序列)、多达约25个氨基酸(例如，来自SEQ ID NO:2的氨基酸764至1240至SEQ ID NO:2的氨基酸764至1265的任何序列)、或多达约30个氨基酸(例如，来自SEQ ID NO:2的氨基酸764至1240至SEQ ID NO:2的氨基酸764至1260的任何序列)的另外的VWF序列。在特定实施方案中，VWF片段包含以下或基本上由以下组成：D' 结构域和

D3结构域既不是SEQ ID NO:2的氨基酸764至1274也不是全长成熟VWF。在一些实施方案中，D1D2结构域与D' D3结构域反式表达。在一些实施方案中，D1D2结构域与D' D3结构域顺式表达。

[0306] 在其它实施方案中，包含连接于D1D2结构域的D' D3结构域的VWF片段还包含胞内切割位点，例如(PACE(弗林蛋白酶)或PC5切割的位点)，从而允许在表达时从D' D3结构域切割D1D2结构域。本文别处公开了胞内切割位点的非限制性实例。

[0307] 在另外其它实施方案中，VWF片段包含D' 结构域和D3结构域，但不包含选自以下组成的组的氨基酸序列：(1) SEQ ID NO:2的氨基酸1241至2813、(2) SEQ ID NO:2的氨基酸1270至氨基酸2813、(3) SEQ ID NO:2的氨基酸1271至氨基酸2813、(4) SEQ ID NO:2的氨基酸1272至氨基酸2813、(5) SEQ ID NO:2的氨基酸1273至氨基酸2813、(6) SEQ ID NO:2的氨基酸1274至氨基酸2813及其任意组合。

[0308] 在又其它实施方案中，本发明的VWF片段包含以下、基本上由以下组成或由以下组成：对应于D' 结构域、D3结构域和A1结构域的氨基酸序列，其中所述氨基酸序列与SEQ ID NO:2的氨基酸764至1479具有至少60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的同一性，其中VWF片段阻止内源性VWF与FVIII结合。在特定实施方案中，VWF片段不是SEQ ID NO:2的氨基酸764至1274。

[0309] 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，但不包含至少一个选自以下组成的组的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，和(20) 其任意组合。

[0310] 在另外其它实施方案中，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，Zhou等人，Blood 2012年4月6日在线公布：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模块及其任意组合。

[0311] 在又其它实施方案中，VWF片段连接于异源部分，其中异源部分连接于VWF片段的N-末端或C-末端或紧邻VWF片段中FVIII蛋白中一个或多个氨基酸(例如，一个或多个XTEN插入位点)的下游插入。例如，VWF片段中异源部分的插入位点可以于D' 结构域、D3结构域或两者中。异源部分可以为半衰期延长剂。

[0312] 在某些实施方案中，本发明的VWF片段形成多聚体，例如二聚体、三聚体、四聚体、五聚体、六聚体、七聚体或更高等级的多聚体。在其它实施方案中，VWF片段为具有仅仅一个

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替代、缺失、添加或修饰。在一个实施方案中,VWF片段可包括氨基酸替代、缺失、添加或修饰以使VWF片段不能够形成二硫键或形成二聚体或多聚体。在另一个实施方案中,氨基酸替代在D'结构域和D3结构域内。在特定实施方案中,本发明的VWF片段含有在对应于SEQ ID NO:2的残基1099、残基1142、或残基1099和1142两者的残基处的至少一个氨基酸替代。至少一个氨基酸替代可以为非天然存在于野生型VWF中的任意氨基酸。例如,氨基酸替代可以为除半胱氨酸之外的任意氨基酸,例如,异亮氨酸、丙氨酸、亮氨酸、天冬酰胺、赖氨酸、天冬氨酸、蛋氨酸、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。在另一实例中,氨基酸替代具有一个或多个阻止或抑制VWF片段形成多聚体的氨基酸。

[0313] 在某些实施方案中,本文有用的VWF片段可被进一步修饰以改善其与FVIII的相互作用,例如以改善与FVIII的结合亲和力。作为非限制性实例,VWF片段包含在对应于SEQ ID NO:2的氨基酸764的残基处的丝氨酸残基和在对应于SEQ ID NO:2的氨基酸773的残基处的赖氨酸残基。残基764和/或773可有助于VWF片段与FVIII的结合亲和力。在其它实施方案中,可用于本发明的VWF片段可具有其它修饰,例如蛋白质可以为聚乙二醇化的、糖基化的、羟乙基淀粉化的或聚唾液酸化的。

[0314] B) XTEN序列

[0315] 如本文所用的“XTEN序列”是指具有非天然存在的大致上非重复序列的延长长度多肽,其主要由小亲水性氨基酸组成,其中该序列在生理条件下具有低程度的二级或三级结构或无二级或三级结构。作为嵌合蛋白配偶体,XTEN可用作当连接于本发明的VWF片段或FVIII序列以产生嵌合蛋白时赋予某些期望药动学、物化和药学性质的载体。此类期望性质包括但不限于增强的药动学参数和溶解度特征。如本文所使用,“XTEN”明确排除抗体或抗体片段诸如单链抗体或轻链或重链的Fc片段。

[0316] 在一些实施方案中,本发明的XTEN序列为具有大于约20、30、40、50、60、70、80、90、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1200、1400、1600、1800或2000个氨基酸残基的肽或多肽。在某些实施方案中,XTEN为具有大于约20至约3000个氨基酸残基、大于30至约2500个残基、大于40至约2000个残基、大于50至约1500个残基、大于60至约1000个残基、大于70至约900个残基、大于80至约800个残基、大于90至约700个残基、大于100至约600个残基、大于110至约500个残基或大于120至约400个残基的肽或多肽。

[0317] 本发明的XTEN序列可包含9至14个氨基酸残基的一个或多个序列基序或与该序列基序具有至少80%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%同一性的氨基酸序列,其中所述基序包含以下、基本上由以下组成或由以下组成:4至6种选自以下组成的组的氨基酸:甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)、谷氨酸(E)和脯氨酸(P)。参见US2010-0239554A1。

[0318] 在一些实施方案中,XTEN包含非重叠序列基序,其中约80%、或至少约85%、或至少约90%、或约91%、或约92%、或约93%、或约94%、或约95%、或约96%、或约97%、或约98%、或约99%或约100%的序列由选自表2A的单个基序家族的非重叠序列的多个单元组成,从而生成家族序列。如本文所使用,“家族”意指XTEN具有仅仅选自表2A的单一

基序分类的基序；即AD、AE、AF、AG、AM、AQ、BC或BD XTEN，以及非来自家族基序的XTEN中的任何其它氨基酸被选择以实现所需特性，诸如以允许通过编码核苷酸并入限制位点、并入切割序列或以实现与FVIII或VWF的更好连接。在XTEN家族的一些实施方案中，XTEN序列包含AD基序家族的、或AE基序家族的、或AF基序家族的、或AG基序家族的、或AM基序家族的、或AQ基序家族的、或BC家族的、或BD家族的非重叠序列基序的多个单元，其中所得XTEN展现出上文所述的同源性范围。在其它实施方案中，XTEN包含来自两个或更多个表2A的基序家族的基序序列的多个单元。这些序列可被选择以实现所需物理/化学特征，包括由基序的氨基酸组成赋予的诸如净电荷、亲水性、缺少二级结构或缺少重复性的性质，如在下文更充分地描述。在上文该段落中所述的实施方案中，并入XTEN中的基序可被选择并且使用本文所述的方法组装以实现约36个至约3000个氨基酸残基的XTEN。

[0319] 表2A. 12个氨基酸的XTEN序列基序和基序家族

[0320]

基序家族 *	基序序列
AD	GESPGGSSGSES
AD	GSEGSSGPGESS
AD	GSSSEGSSEGGP
AD	GSGGEPSESGSS
AE, AM	GSPAGSPTSTEE
AE, AM, AQ	GSFPATSGSETP
AE, AM, AQ	GTSESATPESGP
AE, AM, AQ	GTSTEPSEGSAP
AF, AM	GSTSESPSGTAP
AF, AM	GTSTPESGSASP
AF, AM	GTSPSGESSTAP
AF, AM	GSTSSTAESPGP
AG, AM	GTPGSGTASSSP
AG, AM	GSSTPSGATGSP
AG, AM	GSSPSASTGTGP
AG, AM	GASPGTSSTGSP
AQ	GEPAGSPTSTSE
AQ	GTGEPSTPASE
AQ	GSGPSTESAPTE
AQ	GSETPSGPSETA
AQ	GPSETSTSEPGA
AQ	GSPSEPTTEG TSA
BC	GSGASEPTSTEP
BC	GSEPATSGTEPS
BC	GTSEPSTSEPGA
BC	GTSTEPSEPGSA

BD	GSTAGSETSTEA
BD	GSETATSGSETA
BD	GTSESATSESGA
BD	GTSTEASEGSAS

[0321] • 指示当以不同排列一起使用时产生“家族序列”的个别基序序列

[0322] XTEN可具有用于插入到或连接于FVIII或VWF的不同长度。在一个实施方案中,基于融合蛋白中实现的特性或功能选择XTEN序列的长度。根据期望特性或功能,XTEN可以为可用作载体的短或中等长度序列或较长序列。在某些实施方案中,XTEN包括约6至约99个氨基酸残基的短区段、约100至约399个氨基酸残基的中等长度、和约400至约1000个和多达约3000氨基酸残基的较长长度。因此,插入到或连接于FVIII或VWF的XTEN可具有约6、约12、约36、约40、约42、约72、约96、约144、约288、约400、约500、约576、约600、约700、约800、约864、约900、约1000、约1500、约2000、约2500或多达约3000个氨基酸残基长的长度。在其它实施方案中,XTEN序列为约6至约50、约50至约100、约100至约150、约150至约250、约250至约400、约400至约500、约500至约900、约900至约1500、约1500至约2000、或约2000至约3000氨基酸残基的长度。插入到或连接于FVIII或VWF的XTEN的精确长度可以在没有不利地影响FVIII或VWF的活性的情况下变化。在一个实施方案中,本文所使用的一个或多个XTEN具有36个氨基酸、42个氨基酸、72个氨基酸、144个氨基酸、288个氨基酸、576个氨基酸、或864个氨基酸的长度且可以选自一个或多个XTEN家族序列;即,AD、AE、AF、AG、AM、AQ、BC或BD。

[0323] 在一些实施方案中,本发明中使用的XTEN序列与选自由以下组成的组的序列具有至少60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性:AE42、AG42、AE48、AM48、AE72、AG72、AE108、AG108、AE144、AF144、AG144、AE180、AG180、AE216、AG216、AE252、AG252、AE288、AG288、AE324、AG324、AE360、AG360、AE396、AG396、AE432、AG432、AE468、AG468、AE504、AG504、AF504、AE540、AG540、AF540、AD576、AE576、AF576、AG576、AE612、AG612、AE624、AE648、AG648、AG684、AE720、AG720、AE756、AG756、AE792、AG792、AE828、AG828、AD836、AE864、AF864、AG864、AM875、AE912、AM923、AM1318、BC864、BD864、AE948、AE1044、AE1140、AE1236、AE1332、AE1428、AE1524、AE1620、AE1716、AE1812、AE1908、AE2004A、AG948、AG1044、AG1140、AG1236、AG1332、AG1428、AG1524、AG1620、AG1716、AG1812、AG1908和AG2004。参见US 2010-0239554A1。

[0324] 在一个实施方案中,XTEN序列与选自由以下组成的组的氨基酸序列具有至少60%、70%、80%、90%、95%、96%、97%、98%、99%或100%同一性:AE42 (SEQ ID NO:36)、AE72 (SEQ ID NO:127)、AE144_2A (SEQ IDNO:128)、AE144_3B (SEQ ID NO:129)、AE144_4A (SEQ ID NO:130)、AE144_5A (SEQ IDNO:131)、AE144_6B (SEQ IDNO:132)、AG144_A (SEQ ID NO:133)、AG144_B (SEQ IDNO:134)、AG144_C (SEQ ID NO:135)、AG144_F (SEQ IDNO:136)、AE864 (SEQ ID NO:43)、AE576 (SEQ ID NO:41)、AE288 (SEQ IDNO:39)、AE288_2 (SEQ ID NO:137)、AE144 (SEQ ID NO:37)、AG864 (SEQ ID NO:44)、AG576 (SEQ ID NO:42)、AG288 (SEQ ID NO:40)、AG144 (SEQ ID NO:38) 及其任意组合。

[0325] 在一些实施方案中,XTEN的少于100%的氨基酸选自甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)、谷氨酸(E)和脯氨酸(P),或少于100%的序列由来自表2A的序列基序或表2B的XTEN序列组成。在此类实施方案中,XTEN的剩余氨基酸残基选自其它14种天然L-氨

氨酸中的任一种,但可优先选自亲水性氨基酸以使XTEN序列含有至少约90%、91%、92%、93%、94%、95%、96%、97%、98%或至少约99%亲水性氨基酸。缀合构建体中所用的XTEN中的疏水性氨基酸的含量可少于5%、或少于2%、或少于1%疏水性氨基酸的含量。XTEN构建中相对冷门的疏水性残基包括色氨酸、苯丙氨酸、酪氨酸、亮氨酸、异亮氨酸、缬氨酸和蛋氨酸。此外,XTEN序列可含有少于5%或少于4%或少于3%或少于2%或少于1%下列氨基酸或不含有下列氨基酸:蛋氨酸(例如,以避免氧化)、或天冬酰胺和谷氨酰胺(以避免脱酰胺化)。

[0326] 在另一个实施方案中,XTEN序列选自由以下组成的组:AE42 (SEQ ID NO:36)、AE72 (SEQ ID NO:127)、AE144_2A (SEQ ID NO:128)、AE144_3B (SEQ ID NO:129)、AE144_4A (SEQ ID NO:130)、AE144_5A (SEQ ID NO:131)、AE144_6B (SEQ ID NO:132)、AG144_A (SEQ ID NO:133)、AG144_B (SEQ ID NO:134)、AG144_C (SEQ ID NO:135)、AG144_F (SEQ ID NO:136)、AE864 (SEQ ID NO:43)、AE576 (SEQ ID NO:41)、AE288 (SEQ ID NO:39)、AE288_2 (SEQ ID NO:137)、AE144 (SEQ ID NO:37)、AG864 (SEQ ID NO:44)、AG576 (SEQ ID NO:42)、AG288 (SEQ ID NO:40)、AG144 (SEQ ID NO:38) 及其任意组合。在具体实施方案中,XTEN序列为AE288。本发明的某些XTEN序列的氨基酸序列示于表2B中。

[0327] 表2B.XTEN序列

[0328]

XTEN	氨基酸序列
AE42 SEQ ID NO: 36	GAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPASS
AE72 SEQ ID NO: 127	GAPTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPA TSGSETPGTSESATPESGPGTSTEPSEGSAPGASS
AE144 SEQ ID NO:37	GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTE EGTSTEPSEG SAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSE GSAPGTSESA PESGPGSEPATSGSETPGTSTEPSEGSAP
AE144_2A (SEQ ID NO: 128)	TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSET PGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPES GPG
AE144_3B (SEQ ID NO: 129)	SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS PGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APG
AE144_4A (SEQ ID NO: 130)	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APG
AE144_5A (SEQ ID NO:	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETP GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTST

[0329]

131)	EEG
AE144_6B (SEQ ID NO: 132)	TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGP GSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS PGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APG
AG144 SEQ ID NO:38	GTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTG PGASPGTSST GSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTS TGSPGSSPSA STGTGPGTPGSGTASSSPGSSTPSGATGSP
AG144_A (SEQ ID NO: 133)	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASS PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASS SPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSST GSP
AG144_B (SEQ ID NO: 134)	GTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASS PGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATG SPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSST GSP
AG144_C (SEQ ID NO: 135)	GTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTG PGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTG SPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSST GSP
AG144_F (SEQ ID NO: 136)	GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTG PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGT GPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSST GSP
AE288 SEQ ID NO:39	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSET PGTSESATPESG PGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGTSESATPE SGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPT STEEGTSTEPSE GSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEP EGSAP
AE288_2 (SEQ ID NO: 137)	GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAP
AG288	PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATG

[0330]

SEQ NO:40	ID	SPGTPGSGTASS SPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGAT GSPGSSPSASTG TGPSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGA TGSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSG ATGSPGSSPSAS TGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPS GATGS
AE576 SEQ NO:41	ID	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSA PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSE TPGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEG SAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSEPATSG SETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGSPAGSP TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTE PSEGSAPGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSP AGSPTSTEEGSP AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAP
AG576 SEQ NO:42	ID	PGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGT GPGSSTPSGATG SPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSST GSPGTPGSGTAS SSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTA SSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSG ATGSPGSSTPSG ATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPS GATGSPGSSTPS GATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGS GTASSSPGASPG TSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPG SGTASSSPGSST PSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSS TPSGATGSPGSS TPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGT PGSGTASSSPGS STPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGS
AE864		GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTE EGTSTEPSEGSA

[0331]

SEQ NO:43	ID	PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSE TPGSPAGSPTST EEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEG SAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSEPATSG SETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGSPAGSP TSTEETSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTE PSEGSAPGSPAG SPTSTEETSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEP ATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSP AGSPTSTEEGSP AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGT SESATPESGPGS EPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPG SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSPAGSPTSTEE GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESG PGTSESATPESG PGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGS APGTSTEPSEGS APGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP
AG864 SEQ NO:44	ID	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSS PGSSTPSGATGS PGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATG SPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGAT GSPGASPGTSST GSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGSSTPSGA TGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTS STGSPGTPGSGT ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSA STGTGPGTPGSG TASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTP SGATGSPGSSTP SGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSST PSGATGSPGSST PSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTP GSGTASSSPGAS PGTSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGT PGSGTASSSPGS STPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPG SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSP GTPGSGTASSSP

[0332]

GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGS PGASPGTSSTGS PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGT GPGTPGSGTASS SPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP
--

[0333] 在另外的实施方案中,本发明中使用的XTEN序列影响本发明的嵌合蛋白的物理或化学性质,例如药动学。本发明中使用的XTEN序列可展现出一种或多种下列有利性质:构象柔性、增强的水溶解性、高程度的蛋白酶抗性、低免疫原性、与哺乳动物受体的低结合或增加的流体动力学(或Stoke)半径。在具体实施方案中,本发明中连接于FVIII蛋白的XTEN序列增加药动学性质诸如较长终末半衰期或增加的曲线下面积(AUC),以使本文所述的嵌合蛋白体内驻留的时间相较于野生型FVIII有所增加。在另外的实施方案中,本发明中使用的XTEN序列增加药动学性质诸如较长终末半衰期或增加的曲线下面积(AUC),以使FVIII蛋白体内驻留的时间相较于野生型FVIII有所增加。

[0334] 可采用多种方法和测定来确定包含XTEN序列的蛋白质的物理/化学性质。此类方法包括但不限于分析离心、EPR、HPLC离子交换、HPLC-尺寸排阻、HPLC-反相、光散射、毛细管电泳、圆二色性、差示扫描量热法、荧光、HPLC-离子交换、HPLC-尺寸排阻、IR、NMR、拉曼光谱学、折射测量法和UV/可见光谱学。另外的方法于Amau等人,Prot Expr and Purif 48,1-13 (2006)中公开。

[0335] 可根据本发明使用的XTEN序列的另外实例于US专利公布号2010/0239554A1、2010/0323956A1、2011/0046060A1、2011/0046061A1、2011/0077199A1、或2011/0172146A1、或国际专利公布号WO 2010091122A1、WO 2010144502A2、WO 2010144508A1、WO 2011028228A1、WO 2011028229 A1或WO 2011028344A2中公开。

[0336] C) 因子VIII (FVIII) 蛋白

[0337] 如本文所使用的“FVIII蛋白”意指在凝结方面有其正常作用的功能性FVIII多肽,除非另有规定。术语FVIII蛋白包括保留全长野生型因子VIII在凝结途径中的功能的功能片段、其变体、类似物或衍生物。“FVIII蛋白”可与FVIII多肽(或蛋白)或FVIII交换使用。FVIII功能的实例包括但不限于,活化凝血的能力、用作因子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蛋白可以为人、猪、犬、大鼠或鼠FVIII蛋白。此外,来自人的FVIII与其它种类之间的比较已鉴定可能为功能所需的保守性残基(Cameron等人,Thromb.Haemost.79:317-22 (1998);US 6,251,632)。

[0338] 许多试验可用于评估凝血系统的功能:活化部分凝血活酶时间(aPTT)试验、显色测定、ROTEM测定、凝血酶原时间(PT)试验(也用于确定INR)、纤维蛋白原测试(常通过Clauss方法)、血小板计数、血小板功能测试(常通过PFA-100)、TCT、出血时间、混合试验(在患者的血浆与正常血浆混合时异常是否校正)、凝结因子测定、抗磷脂抗体、D-二聚体、基因试验(例如,因子V Leiden、凝血酶原突变G20210A)、稀释度蝰蛇毒时间(dRVVT)、杂项血小板功能测试、凝血弹性描记法(TEG或Sonoclot)、凝血弹性测量法(TEM[®],例如,ROTEM[®])或优球蛋白溶解时间(ELT)。

[0339] aPTT试验是测量“固有的”(也是指接触活化途径)和常见凝结途径两者的功效的

性能指示剂。该试验常用于测量商购可得的重组凝血因子,例如FVIII或FIX的凝血活性。其与测量外源途径的凝血酶原时间 (PT) 结合使用。

[0340] ROTEM分析提供关于整个止血动力学的信息:凝血时间、血块形成、血块稳定性和溶胞。凝血弹性测量法中的不同参数取决于血浆凝血系统的活性、血小板功能、纤维蛋白溶解或影响这些相互作用的很多因素。该测定可提供次级止血的完整视图。

[0341] FVIII多肽和多核苷酸序列是已知的,很多任何功能片段、突变体和修饰形式也是已知的。人FVIII序列(全长)的实例在下面示出。

[0342] 表3. 全长因子VIII的氨基酸序列

[0343] (全长FVIII (FVIII信号肽为加下划线的;FVIII重链为加双下划线的;B结构域为斜体的;且FVIII轻链为纯文本的)

信号肽: (SEQ ID NO: 3)
 MQIELSTCFFLCLLRFCFS

成熟因子VIII (SEQ ID NO: 4)*

ATTRYYLGAVELSWDYMOSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLEVEFT
 DHLFENIAKPRPPWMGLLGPTIQAEVYD TVVITLKNMASHPVSLHAVGVSYWKASE
 GAEYDDQTSOREKEDDKVPFGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLV
 KDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDR
 DAASARAWPKMHTVNGYVNRSLPGLIGCHRKS VYWHVIGMGTTPEVHSIFLEGHT
 FLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCP
 EEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDDNSPSFIQIRSVAKKHPKTWVHY
 IAAEEEDWDYAPLV LAPDDRSYKSQYLNNGPQORIGRKYKKVRFMAYTDETFKTR
 AIQHESGILGPLYGEVGD TLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGV
 KHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSFVNMERDLASGLIGP
 LLICYKESVDQRGNOIMSDKRNVLFSVFDENRSWYLTENIQRF LNPAGVOLED
 PEFQASNIMHSINGYVFD SLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKH
 KMVYEDTLTLFPFSGETVFM MENPGLWILGCHNSDFRNRGMTALLKVSSCDKNT
 GDYYEDSYEDISAYLLSKNNAIEPRSFSONSRHPSTRQKQFNATTIPENDIEKTD
 PWFARHTFMPKIQNVSSDLLMLLRQSPTPHGLSLSDLQEAKEYTFSDDPSPGAI
 DSNNSLSEMTHFRPQLHHSGDMVFTPESEGLQRLNEKLGTAAATELKKLDFKVSS
 TSNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLD TTLFGKKSSPLTESGGP
 LSLSEENND SKLLESGLMNSQE3SWGKNVSSSTESGRLFKGKRAHGPA LLTKDNAL
 FKVSISLLKTNKTSNNSATNRKTHIDGPSLLIENSPSVWQNI LESDTEFKKV TPL
 IHDRMLMDKNATALRLNHMSNKT TSSKNMEMVQKKEGP IPPDAQNPDM SFFKML
 FLPE SARWIQORTHGKNSLNSGQGPS PKQLVSLGPEKSVEGQNFLSEKNKV VVGKG
 EFTKDVGLKEMVFPSSRNLF LTNLDNLHENNTHNQEKKIQEEIEKKETLIQENNV
 LPQIHTVTGTKNFMKNL FLLSTRQNVESGYDGAYAPVLQDFRSLNDSTNR TKKHT
 AHFSKKGEEENLEGLGNQTKQIVEKYACTTRISPNTSQQNFVTQR SKRALKQFRL
 PLEETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQSP LSDCLT
 RSHSIPQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRK KDSGVQE
 SSHFLQGAKKNNLSLAILTLEMTGDQREVGS LGTSATNSVTYKKVENTVLPK PDL
 PKTSGKVELLPKVHIYQKDLFPTETSN GSPGHLDLVEGSLLQGTEGAIKWNEANR
 PGKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSP EKTAFKKK
 DTILSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKR HQREI
 TRTTLQSDQEEIDYDDTISVEMKKEDFDIYDE DENQSPRSFQKKTRHYFIAAVER
 LWDYGMSSSPHVLNRNAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLG

[0344]

PYIRAEVEDNIMVTFRNQASRPYSFYSSLSIYEEDQRQGAEPKRFVKNPNETKTY
 FWKVQH HMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLL VCHTNTLNPAHGRQ
 VTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYI
 MDTLPLGLVMAQDQIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYP
 GVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTP LGMASGHIRDF
 QITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMI IHGIKTQGAR
 QKFSSLYISQFIIMYSLDGKKWQTYRGNSTG TLMVFFGNVDSSGIKHNI FNPPII
 ARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMF
 ATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTVGTTQGVKSLTSM
 YVKEFLISSSQDGHQWTLFFQNGKVVFQGNQDSFTP VVNSLDPLLLTRYLR IHP
 QSWVHQIALRMEVLGCEAQDLY

[0345] 表4. 编码全长FVIII的核苷酸序列 (SEQ ID NO:5)*

[0346]

```
661                                     ATG  
    CAAATAGAGC TCTCCACCTG  
721  CTTCTTTCTG TGCCTTTTGC GATTCTGCTT TAGTGCCACC  
    AGAAGATACT ACCTGGGTGC  
781  AGTGGAAGCTG TCATGGGACT ATATGCAAAG TGATCTCGGT  
    GAGCTGCCTG TGGACGCAAG  
841  ATTTCCCTCCT AGAGTGCCAA AATCTTTTCC ATTCAACACC  
    TCAGTCGTGT AAAAAAGAC  
901  TCTGTTTGTA GAATTCACGG ATCACCTTTT CAACATCGCT  
    AAGCCAAGGC CACCCTGGAT  
961  GGGTCTGCTA GGTCTACCA TCCAGGCTGA GGTTTATGAT  
    ACAGTGGTCA TTACACTTAA  
1021 GAACATGGCT TCCCATCCTG TCAGTCTTCA TGCTGTTGGT  
    GTATCCTACT GGAAAGCTTC  
1081 TGAGGGAGCT GAATATGATG ATCAGACCAG TCAAAGGGAG  
    AAAGAAGATG ATAAAGTCTT  
1141 CCCTGGTGGA AGCCATACAT ATGTCTGGCA GGTCTGAAA  
    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  
1501 CAGGAAATCA GTCTATTGGC ATGTGATTGG AATGGGCACC  
    ACTCCTGAAG TGCACTCAAT
```


[0347]

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 CTCCTTTAT
1861 CCAAATTCGC TCAGTTGCCA AGAAGCATCC TAAAACTTGG
GTACATTACA TTGCTGCTGA
1921 AGAGGAGGAC TGGGACTATG CTCCCTTAGT CCTCGCCCCC
GATGACAGAA GTTATAAAAG
1981 TCAATATTTG AACAAATGGCC 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 AAAGTGTCTT CATGTCGATG
GAAAACCCAG GTCTATGGAT
2821 TCTGGGGTGC CACAACCTCAG 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

[0348]

3061 TGACCCTTGG TTTGCACACA GAACACCTAT GCCTAAAATA
CAAAATGTCT CCTCTAGTGA
3121 TTTGTTGATG CTCTTGCGAC AGAGTCCTAC TCCACATGGG
CTATCCTTAT CTGATCTCCA
3181 AGAAGCCAAA TATGAGACTT TTTCTGATGA TCCATCACCT
GGAGCAATAG ACAGTAATAA
3241 CAGCCTGTCT GAAATGACAC ACTTCAGGCC ACAGCTCCAT
CACAGTGGGG ACATGGTATT
3301 TACCCCTGAG TCAGGCCTCC AATTAAGATT AAATGAGAAA
CTGGGGACAA CTGCAGCAAC
3361 AGAGTTGAAG AAACCTGATT TCAAAGTTTC TAGTACATCA
AATAATCTGA TTTCAACAAT
3421 TCCATCAGAC AATTGCGCAG 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 TTAAAGGGA 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 ATAAACTAC
3961 TTCATCAAAA AACATGGAAA TGGTCCAACA GAAAAAGAG
GGCCCCATTC CACCAGATGC
4021 ACAAATCCA 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 AACCTATTTT TTAATACTT GGATAATTTA
CATGAAAATA ATACACACAA
4321 TCAAGAAAAA AAAATTTCAGG 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

[0349]

4561 AAAAAAAGGG GAGGAAGAAA ACTTGGAAGG CTTGGGAAAT
CAAACCAAGC AAATTGTAGA
4621 GAAATATGCA TGCACCACAA GGATATCTCC TAATACAAGC
CAGCAGAATT TTGTCACGCA
4681 ACGTAGTAAG AGAGCTTTGA AACAAATTCAG ACTCCCCTA
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
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 TCTCGTGGA
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
CAGAGTGGA GTGTCCCTCA
5941 GTTCAAGAAA GTTGTTTTCC AGGAATTTAC TGATGGCTCC
TTTACTCAGC CTTATACCG
6001 TGGAGAACTA AATGAACATT TGGGACTCCT GGGGCCATAT
ATAAGAGCAG AAGTTGAAGA

[0350]

6061 TAATATCATG GTAAC TTTC A GAAATCAGGC CTCTCGTCCC
TATTCCTTCT ATTCTAGCCT
6121 TATTTCTTAT GAGGAAGATC AGAGGCAAGG AGCAGAACCT
AGAAAAA ACT TTGTCAAGCC
6181 TAATGAAACC AAAACTTACT TTTGGAAAGT GCAACATCAT
ATGGCACCCA CTAAAGATGA
6241 GTTGACTGCG AAAGCCTGGG CTTATTTCTC TGATGTTGAC
CTGGAAAAAG ATGTGCACTC
6301 AGGCCTGATT GGACCCCTTC TGGTCTGCCA CACTAACACA
CTGAACCCTG CTCATGGGAG
6361 ACAAGTGACA GTACAGGAAT TTGCTCTGTT TTTCACCATC
TTTGATGAGA CCAAAGCTG
6421 GTACTTCACT GAAAATATGG AAAGAACTG CAGGGCTCCC
TGCAATATCC AGATGGAAGA
6481 TCCCACTTTT AAAGAGAATT ATCGCTTCCA TGCAATCAAT
GGCTACATAA TGGATACACT
6541 ACCTGGCTTA GTAATGGCTC AGGATCAAAG GATTCGATGG
TATCTGCTCA GCATGGGCAG
6601 CAATGAAAAC ATCCATTCTA TTCATTTT CAG TGGACATGTG
TTCATGTAC GAAAAAAGA
6661 GGAGTATAAA ATGGCACTGT ACAATCTCTA TCCAGGTGTT
TTTGAGACAG TGGAAATGTT
6721 ACCATCCAAA GCTGGAATTT GGCGGGTGGA ATGCCTTATT
GGCGAGCATC TACATGCTGG
6781 GATGAGCACA CTTTTTCTGG TGTACAGCAA TAAGTGT CAG
ACTCCCCTGG GAATGGCTTC
6841 TGGACACATT AGAGATTTTC AGATTACAGC TTCAGGACAA
TATGGACAGT GGGCCCCAAA
6901 GCTGGCCAGA CTTTATTATT 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
GTAATACTC AGGGAGTAAA
7501 ATCTCTGCTT ACCAGCATGT ATGTGAAGGA GTTCCTCATC
TCCAGCAGTC AAGATGGCCA

```
7561 TCAGTGGACT CTCTTTTTC AGAATGGCAA AGTAAAGGTT
      TTTCAGGGAA ATCAAGACTC
7621 CTTACACCT GTGGTGAAC CTCTAGACCC ACCGTTACTG
      ACTCGCTACC TTCGAATTCA
[0351] 7681 CCCCCAGAGT TGGGTGCACC AGATTGCCCT GAGGATGGAG
      GTTCTGGGCT GCGAGGCACA
      7741 GGACCTCTAC
```

[0352] *加下划线的核酸编码信号肽。

[0353] FVIII多肽包括全长FVIII、在N-末端缺少Met的全长FVIII、成熟FVIII(缺少信号序列)、在N-末端有附加的Met的成熟FVIII和/或结构域B全部或部分缺失的FVIII。在某些实施方案中,FVIII变体包括B结构域缺失,不管是否部分或全部缺失。

[0354] 天然的成熟人FVIII的序列以SEQ ID NO:4呈现。天然FVIII蛋白具有下式:A1-a1-A2-a2-B-a3-A3-C1-C2,其中A1、A2和A3为结构相关的“A结构域”,B为“B结构域”,C1和C2为结构相关的“C结构域”以及a1、a2和a3为酸性间隔区。提及SEQ ID NO:4中的一级氨基酸序列位置,人FVIII的A1结构域从Ala1延伸至约Arg336,a1间隔区从约Met337延伸至约Val374,A2结构域从约Ala375延伸至约Tyr719,a2间隔区从约Glu720延伸至约Arg740,B结构域从约Ser741延伸至约Arg 1648,a3间隔区从约Glu1649延伸至约Arg1689,A3结构域从约Ser1690延伸至约Leu2025,C1结构域从约Gly2026延伸至约Asn2072,且C2结构域从约Ser2073延伸至Tyr2332。除特定蛋白水解切割位点之外,结构域和FVIII区域之间的边界的位置的指定在不同参考文献中可变化。因此,本文所示的边界通过使用术语“约”被指定为近似的。

[0355] 人FVIII基因被分离且在哺乳动物细胞中表达(Toole,J.J.,等人,Nature312:342-347(1984);Gitschier,J.,等人,Nature 312:326-330(1984);Wood,W.I.,等人,Nature 312:330-337(1984);Vehar,G.A.,等人,Nature 312:337-342(1984);WO 87/04187;WO 88/08035;WO 88/03558;和美国专利号4,757,006)。如美国专利号4,965,199中所示,从cDNA推导出FVIII氨基酸序列。此外,部分或完全B-结构域缺失的FVIII示于美国专利号4,994,371和4,868,112中。在一些实施方案中,人FVIII B-结构域被替换为人因子VB-结构域,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序列分别示于US申请公布号2005/0100990的SEQ ID NO:4和5中。

[0356] 猪FVIII序列公布于Toole,J.J.,等人,Proc.Natl.Acad.Sci.USA83:5939-5942(1986)中。此外,获自来自猪脾cDNA文库的FVIII序列的PCR扩增的完整猪cDNA序列已报道于Healey,J.F.,等人,Blood 88:4209-4214(1996)中。具有所有结构域、所有亚单位和特定氨基酸序列的替代的杂合人/猪FVIII于Lollar和Runge的美国专利号5,364,771、以及WO 93/20093中公开。近来,猪A1和/或A2结构域被替代为相应的人结构域的猪FVIII和嵌合FVIII的A1和A2结构域的核苷酸和相应的氨基酸序列报道于WO 94/11503中。美国专利号5,859,204,Lollar,J.S.也公开了猪cDNA和推导的氨基酸序列。美国专利号6,458,563公开了B结构域缺失的猪FVIII。

[0357] Lollar,J.S.的美国专利号5,859,204报道了具有减少的抗原性和减少的免疫反应性的FVIII的功能性突变体。Lollar,J.S.的美国专利号6,376,463也报道了具有减少的免疫反应性的FVIII的突变体。Saenko等人的US申请公布号2005/0100990报道了FVIII的A2

结构域中的功能性突变。

[0358] 在一个实施方案中,FVIII(或嵌合蛋白的FVIII部分)可与SEQ ID NO:6的氨基酸1至1438或SEQ ID NO:4的氨基酸1至2332的FVIII氨基酸序列(不含信号序列)或SEQ ID NO:3的氨基酸1至19和SEQ ID NO:6的氨基酸1至1438或SEQ ID NO:3的氨基酸1至19和SEQ ID NO:4的氨基酸1至2332的FVIII氨基酸序列(具有信号序列)具有至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%同一性,其中FVIII具有凝血活性,例如活化作为辅因子的因子IX以使因子X转化为活化因子X。FVIII(或嵌合蛋白的FVIII部分)可以与SEQ ID NO:6的氨基酸1至1438或SEQ ID NO:4的氨基酸1至2332的FVIII氨基酸序列(不含信号序列)相同。FVIII还可以包含信号序列。

[0359] 如本文所使用,FVIII的“B-结构域”与本领域中已知的B-结构域相同,其由内部氨基酸序列同一性和蛋白水解切割位点定义,例如全长人FVIII的残基Ser741-Arg1648。其它人FVIII结构域由下列氨基酸残基定义:A1,残基Ala1-Arg372;A2,残基Ser373-Arg740;A3,残基Ser1690-Asn2019;C1,残基Lys2020-Asn2172;C2,残基Ser2173-Tyr2332。A3-C1-C2序列包括残基Ser1690-Tyr2332。剩余序列、残基Glu1649-Arg1689通常被称为a3酸性区域。猪、小鼠和犬FVIII的所有结构域(包括B-结构域)的边界的位置也是本领域已知的。在一个实施方案中,FVIII的B结构域缺失(“B-结构域缺失的因子VIII”或“BDD FVIII”)。BDD FVIII的实例为REFACTO[®](重组BDD FVIII),其具有与表5中的序列的因子VIII部分相同的序列。(BDD FVIII重链为加双下划线的;B结构域为斜体的;和BDD FVIII轻链为纯文本的)。编码表5中列出的编码氨基酸序列的核苷酸序列(SEQ ID NO:7)于表6中示出。

[0360] 表5.B-结构域缺失的因子VIII(BDD FVIII)的氨基酸序列

BDD FVIII (SEQ ID NO: 6)

ATRRYYLGAVELSWDYMQSDLGELPVDARFPFPRVPKSFPFNTSVVYKKTLEFVEFT
DHLFNIAPRPPWMGLLGPTIOAEVYDVTVVITLKNMASHPVSLHAVGVSYWKASE
GAEYDDQTSOREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLV
KDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDR
DAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHT
FLVRNHRQASLEISPITFLTAOTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCP
EELQLRMKNNEEAEDYDDDLTDSEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHY
IAAEEEDWDYAPLV LAPDDRSYKSQYLNNGPQIRGRKYKKVRFMAYTDETFKTR
AIQHESGILGPLYGEVGDTLIIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGV
KHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFVNMERDLASGLIGP
LLICYKESVDQGRNQIMSDKRNVLFSVFENRSWYLTENIQRFPLNPAGVQLED
PEFQASNIMHSINGYVFDLSQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKH
KMVYEDTLTLFPFSGETVFMSENPLWILGCHNSDFRNRMGTALLKVSSCDKNT
GDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDY
DDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVL

[0361]

NRAQSGSVPQFKKVVFEFTDGSFTQPLYRGELNEHLGGLGPYIRAEVEDNIMVT
FRNQASRPYSFYSSLISYEEDQRQGAEPKRFVKNPNETKTYFWKVQHMAPTKDE
FDCKAWAYFSDVDLEKDVHSGLIGPLLCHTNTLNPAHGRQVTVQEFALFFTIFD
ETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQR
IRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPKAG
IWRVECLIGEHLHAGMSTLFLVYSNKCQTPMGASGHIRDFQITASGQYGQWAPK
LARLHYSGSINAWSTKEPFSWIKVDLLAPMI IHGIKTQGARQKFSSLYISQFIIM
YSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNI FNPPIIARYIRLHPHTYSISIR
STLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQG
RSNAWRPQVNNPKEWLQVDFQKTMKVTGVTQGVKSLTSMYVKEFLISSSQDGH
QWTLFFQNGKVVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVL
GCEAQDLY

[0362] 表6. 编码BDD FVIII的核苷酸序列 (SEQ ID NO:7)*

[0363]

```

661                                     A TGCAAATAGA
   GCTCTCCACC TGCTTCTTTC
721  TGTGCCTTTT GCGATTCTGC TTTAGTGCCA CCAGAAGATA
   CTACCTGGGT GCAGTGGAAC
781  TGTCATGGGA CTATATGCAA AGTGATCTCG GTGAGCTGCC
   TGTGGACGCA AGATTTCTC
841  CTAGAGTGCC AAAATCTTTT CCATTCAACA CCTCAGTCGT
   GTACAAAAAG ACTCTGTTG
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
   GTTCTACTG TTTTGTCTA

```


[0364]

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 ATGGCATAACA
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 GCCACAATC AGACTTTCGG AACAGAGGCA TGACCGCCTT
ACTGAAGGTT TCTAGTTGTG
2881 ACAAGAACAC TGGTGATTAT TACGAGGACA GTTATGAAGA
TATTTTCAGCA 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

[0365]

3181 TCTGGGATTA TGGGATGAGT AGCTCCCCAC ATGTTCTAAG
AAACAGGGCT CAGAGTGGCA
3241 GTGTCCCTCA GTTCAAGAAA GTTGTTTTCC AGGAATTTAC
TGATGGCTCC TTTACTCAGC
3301 CCTTATACCG TGGAGAATAA AATGAACATT TGGGACTCCT
GGGGCCATAT ATAAGAGCAG
3361 AAGTTGAAGA TAATATCATG GTAACCTTCA 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 AAAGAACTG
CAGGGCTCCC TGCAATATCC
3781 AGATGGAAGA TCCCACCTTT AAAGAGAATT ATCGCTTCCA
TGCAATCAAT GGCTACATAA
3841 TGGATACACT ACCTGGCTTA GTAATGGCTC AGGATCAAAG
GATTCGATGG TATCTGCTCA
3901 GCATGGGCAG CAATGAAAAC ATCCATTCTA TTCATTTCTCAG
TGGACATGTG TTTACTGTAC
3961 GAAAAAAGA GGAGTATAAA ATGGCACTGT ACAATCTCTA
TCCAGGTGTT TTTGAGACAG
4021 TGGAAATGTT ACCATCCAAA GCTGGAATTT GGCGGGTGGG
ATGCCTTATT GGCGAGCATC
4081 TACATGCTGG GATGAGCACA CTTTTTCTGG TGTACAGCAA
TAAGTGTCAG ACTCCCCTGG
4141 GAATGGCTTC TGGACACATT AGAGATTTTC AGATTACAGC
TTCAGGACAA TATGGACAGT
4201 GGGCCCCAAA GCTGGCCAGA CTTTATTATT 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
TTTTAACCTT 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 GTAAC TACTC
4801 AGGGAGTAAA ATCTCTGCTT ACCAGCATGT ATGTGAAGGA
GTTCCCTCATC TCCAGCAGTC
[0366] 4861 AAGATGGCCA TCAGTGGACT CTCTTTTTTTC AGAATGGCAA
AGTAAAGGTT TTTCAGGGAA
4921 ATCAAGACTC CTTACACCT GTGGTGAACCT CTCTAGACCC
ACCGTTACTG ACTCGCTACC
4981 TTCGAATTCA CCCCCAGAGT TGGGTGCACC AGATTGCCCT
GAGGATGGAG GTTCTGGGCT
5041 GCGAGGCACA GGACCTCTAC

[0367] *加下划线的核酸编码信号肽。

[0368] “B-结构域缺失的FVIII”可以具有全部或部分缺失，公开于美国专利号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。在一些实施方案中，本发明的B-结构域缺失的FVIII序列包含在美国专利号6,316,226(也于US 6,346,513中)的第4栏第4行至第5栏第28行和实施例1-5公开的缺失中的任一种。在另一个实施方案中，B-结构域缺失的因子VIII为S743/Q1638B-结构域缺失的因子VIII(SQ BDD FVIII)(例如，具有从氨基酸744至氨基酸1637的缺失的因子VIII，例如具有SEQ ID NO:4的氨基酸1-743和氨基酸1638-2332的因子VIII，即SEQ ID NO:6)。在一些实施方案中，本发明的B-结构域缺失的FVIII具有在美国专利号5,789,203(也于US 6,060,447、US 5,595,886和US 6,228,620中)的第2栏第26-51行和实施例5-8公开的缺失。在一些实施方案中，B-结构域缺失的因子VIII具有以下中所述的缺失：美国专利号5,972,885的第1栏第25行至第2栏第40行；美国专利号6,048,720的第6栏第1-22行和实施例1；美国专利号5,543,502的第2栏第17-46行；美国专利号5,171,844的第4栏第22行至第5栏第36行；美国专利号5,112,950的第2栏第55-68行、图2和实施例1；美国专利号4,868,112的第2栏第2行至第19栏第21行和表2；美国专利号7,041,635的第2栏第1行至第3栏第19行、第3栏第40行至第4栏第67行、第7栏第43行至第8栏第26行；和第11栏第5行至第13栏第39行；或美国专利号6,458,563的第4栏第25-53行。在一些实施方案中，B-结构域缺失的FVIII具有大部分B结构域的缺失，但仍含有对于初级翻译产物的体内蛋白水解加工成两条多肽链而言是必要的B结构域的氨基端序列，如于W0 91/09122中公开。在一些实施方案中，B-结构域缺失的FVIII通过缺失氨基酸747-1638，即几乎完全缺失B结构域来构建。Hoeben R.C.等人J.Biol.Chem.265(13):7318-7323(1990)。B-结构域缺失的因子VIII也可含有FVIII的氨基酸771-1666或氨基酸868-1562的缺失。Meulien P.等人Protein Eng.2(4):301-6(1988)。为本发明部分的另外的B结构域缺失包括：氨基酸982至1562或760至1639(Toole等人,Proc.Natl.Acad.Sci.U.S.A.(1986)83,5939-5942))、797至1562(Eaton等人Biochemistry(1986)25:8343-8347))、741至1646(Kaufman(PCT公布申请号W0 87/04187))、747-1560(Sarver,等人,DNA(1987)6:553-564))、741至1648(Pasek(PCT申请号88/00831))、或816至1598或741至1648(Lagner(Behring Inst.Mitt.(1988)82期:16-25,EP 295597))的缺失。在其它实施方案中，BDD

FVIII包括含有B-结构域的片段的FVIII多肽,所述片段保留一个或多个N连接的糖基化位点,例如残基757、784、828、900、963或任选地943,其对应于全长FVIII序列的氨基酸序列。B-结构域片段的实例包括公开于Miao,H.Z.,等人,Blood 103(a):3412-3419(2004),Kasuda,A,等人,J.Thromb.Haemost.6:1352-1359(2008)和Pipe,S.W.,等人,J.Thromb.Haemost.9:2235-2242(2011)中的B-结构域的226个氨基酸或163个氨基酸(即,B-结构域的前226个氨基酸或163个氨基酸被保留)。在又其它实施方案中,BDD FVIII还包含在残基309处的点突变(从Phe至Ser)以改善BDD FVIII蛋白的表达。参见Miao,H.Z.,等人,Blood 103(a):3412-3419(2004)。在又其它实施方案中,BDD FVIII包括含有B-结构域的一部分但不含一个或多个弗林蛋白酶切割位点(例如,Arg1313和Arg 1648)的FVIII多肽。参见Pipe,S.W.,等人,J.Thromb.Haemost.9:2235-2242(2011)。可在任何FVIII序列中制备前述缺失中的每一个。

[0369] 在一些实施方案中,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FVIII蛋白为FVIII198(SEQ ID NO:89)。FVIII198为含有单链FVIIIIFc分子-226N6的部分B-结构域。226表示FVIII B-结构域的N-末端226氨基酸,并且N6表示B-结构域中的6个N-糖基化位点。

[0370] 在一个实施方案中,FVIII恰好在于氨基酸1648(于全长因子VIII或SEQ ID NO:4中)、氨基酸754(于S743/Q1638B-结构域缺失的因子VIII或SEQ ID NO:6中)处的精氨酸、或相应的精氨酸残基(于其它变体中)之后被切割,从而产生重链和轻链。在另一个实施方案中,FVIII包含重链和轻链,其通过金属离子介导的非共价键连接或缔合。

[0371] 在其它实施方案中,FVIII为单链FVIII,其恰好在于氨基酸1648(于全长FVIII或SEQ ID NO:4中)、氨基酸754(于S743/Q1638B-结构域缺失的FVIII或SEQ ID NO:6中)处的精氨酸、或相应的精氨酸残基(于其它变体中)之后还未被切割。单链FVIII可以包含一个或多个氨基酸替代。在一个实施方案中,氨基酸替代处于对应于全长成熟因子VIII多肽(SEQ ID NO:4)的残基1648、残基1645、或两者或者SQ BDD因子VIII(SEQ ID NO:6)的残基754、残基751、或两者的残基处。氨基酸替代可以为除精氨酸之外的任意氨基酸,例如,异亮氨酸、亮氨酸、赖氨酸、蛋氨酸、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、硒代半胱氨酸、丝氨酸、酪氨酸、组氨酸、鸟氨酸、吡咯赖氨酸或牛磺酸。

[0372] FVIII可进一步被凝血酶切割,然后被活化为FVIIIa,用作活化因子IX的辅因子(FIXa)。并且活化FIX与活化FVIII一起形成X酶复合物并且将因子X转化为活化因子X(FXa)。对于活化,FVIII在三个精氨酸残基之后、在氨基酸372、740和1689(对应于B-结构域缺失的FVIII序列中的氨基酸372、740和795)处被凝血酶切割,而切割产生具有50kDa A1、43kDa A2和73kDa A3-C1-C2链的FVIIIa。在一个实施方案中,可用于本发明的FVIII蛋白为非活性FVIII。在另一个实施方案中,FVIII蛋白为活化FVIII。

[0373] 具有连接于VWF片段或与VWF片段缔合的FVIII多肽的蛋白质可包含与SEQ ID NO:4或6具有至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%同一性的序列,其中该序列具有FVIII凝血活性,例如将因子IX活化为辅因子以使因子X转化为活化因子X(FXa)。

[0374] 如本文所使用的“杂合”或“嵌合”多肽和蛋白质包括任选地融合于第一Ig恒定区

或其一部分的第一多肽链(例如VWF片段)与任选地融合于第二Ig恒定区或其一部分的第二多肽链(例如连接于XTEN序列的FVIII蛋白)的组合,从而形成异二聚体。在一个实施方案中,杂合体中的第一多肽和第二多肽经由蛋白质-蛋白质相互作用诸如电荷-电荷或疏水性相互作用彼此缔合。在另一个实施方案中,杂合体中的第一多肽和第二多肽经由二硫化物或其它共价键彼此缔合。杂合体例如于US 2004/101740和US 2006/074199中描述。第二多肽可以为第一多肽的相同拷贝或非相同多肽。在一个实施方案中,第一多肽为FVIII蛋白(X)-Fc融合蛋白,并且第二多肽为包含以下、基本上由以下组成或由以下组成的多肽:Fc区,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-XTEN-Fc融合蛋白,并且第二多肽包含FVIII-Fc融合蛋白,从而制备杂合异二聚体。在其它实施方案中,第一多肽包含VWF片段-Fc融合蛋白,并且第二多肽包含FVIII(X)-Fc融合蛋白,从而制备杂合异二聚体。在另外其它实施方案中,第一多肽包含VWF片段-XTEN-Fc融合蛋白,并且第二多肽包含FVIII(X)-Fc融合蛋白。第一多肽和第二多肽可通过第一Fc区和第二Fc区之间的共价键例如二硫键缔合。第一多肽和第二多肽可进一步通过VWF片段和FVIII蛋白之间的结合而彼此缔合。

[0375] 可用于本发明的FVIII蛋白可包括具有一个或多个另外的XTEN序列的FVIII,所述另外的XTEN序列未影响FVIII凝结活性。此类XTEN序列可融合于FVIII蛋白的C-末端或N-末端或在FVIII蛋白中的两个氨基酸残基中的一个或多个之间插入,其中插入不影响FVIII凝结活性或FVIII功能。在一个实施方案中,插入改善FVIII蛋白的药动力学性质(例如,半衰期)。在另一个实施方案中,插入可以为多次插入,例如多于两次、三次、四次、五次、六次、七次、八次、九次或十次插入。插入位点的实例包括但不限于,表7、8、9、10、11、12、13、14、15中列出的位点或其任意组合。

[0376] 连接于一个或多个XTEN序列的FVIII蛋白可以FVIII(X)、FVIII(X1)、FVIII(a→b)-X-FVIII(c→d)表示,其中FVIII(a→b)包括以下、基本上由以下组成或由以下组成:从氨基酸残基"a"至氨基酸残基"b"的FVIII蛋白的第一部分;X或X1包括以下、基本上由以下组成或由以下组成:一个或多个XTEN序列;FVIII(c→d)包括以下、基本上由以下组成或由以下组成:从氨基酸残基"c"至氨基酸残基"d"的FVIII蛋白的第二部分;

[0377] a为FVIII蛋白的第一部分的N-端氨基酸残基,

[0378] b为FVIII蛋白的第一部分的C-端氨基酸残基但也为XTEN序列插入于其中的插入位点的两个氨基酸的N-端氨基酸残基,

[0379] c为FVIII蛋白的第二部分的N-端氨基酸残基但也为XTEN序列插入于其中的插入位点的两个氨基酸的C-端氨基酸残基,且

[0380] d为FVIII蛋白的C-端氨基酸残基,并且

[0381] 其中FVIII蛋白的第一部分和FVIII蛋白的第二部分彼此不相同并且一起具有足够长度以使FVIII蛋白具有FVIII凝结活性。

[0382] 在一个实施方案中,FVIII蛋白的第一部分和FVIII蛋白的第二部分分别为SEQ ID NO:4[全长成熟FVIII序列]或SEQ ID NO:6[B-结构域缺失的FVIII]的片段,例如N-端部分和C-端部分。在某些实施方案中,FVIII蛋白的第一部分包含FVIII蛋白的A1结构域和A2结构域。FVIII蛋白的第二部分包含A3结构域、C1结构域和任选地C2结构域。在另外其它实施方案中,FVIII蛋白的第一部分包含A1结构域和A2结构域,并且FVIII蛋白的第二部分包含B

结构域的一部分、A3结构域、C1结构域和任选地C2结构域。在又其它实施方案中，FVIII蛋白的第一部分包含FVIII蛋白的A1结构域、A2结构域和B结构域的一部分，并且FVIII蛋白的第二部分包含A3结构域、C1结构域和任选地C2结构域。在又其它实施方案中，FVIII蛋白的第一部分包含FVIII蛋白的A1结构域、A2结构域和B结构域的第一部分。FVIII蛋白的第二部分包含B结构域的第二部分、A3结构域、C1结构域和任选地C2结构域。在一些实施方案中，两个氨基酸(“b”和“c”)可以为表7、8、9、10、11、12、13、14和15中所示的氨基酸残基插入位点中的任何一个或多个。例如，“b”可以为紧邻一个或多个XTEN序列插入或连接于其中的位点的上游的氨基酸残基，并且“c”可以为紧邻一个或多个XTEN序列插入或连接于其中的位点的下游的氨基酸残基。在一些实施方案中，“a”为FVIII蛋白的第一成熟氨基酸序列，并且“d”为FVIII蛋白的最后一个氨基酸序列。例如，FVIII_(a→b)可以为与SEQ ID NO:6[B结构域缺失的FVIII氨基酸序列]或SEQ ID NO:4[全长FVIII]的氨基酸1至745具有至少70%、80%、90%、95%、96%、97%、98%、99%或100%同一性的氨基酸序列并且FVIII_(c→d)可以分别为SEQ ID NO:6的氨基酸746至1438或SEQ ID NO:4的氨基酸1641至2332。

[0383] 在一些方面，FVIII蛋白中的插入位点位于FVIII蛋白的一个或多个结构域(其为N-末端、A1结构域、A2结构域、A3结构域、B结构域、C1结构域、C2结构域、C-末端或其两个或更多的组合)中，或位于FVIII蛋白的两个结构域(其为A1结构域和a1酸性区域、以及a1酸性区域和A2结构域、A2结构域和a2酸性区域、a2酸性区域和B结构域、B结构域和A3结构域、和A3结构域和C1结构域、C1结构域和C2结构域或其任意组合)之间。例如，XTEN序列可插入于其中的插入位点选自由以下组成的组：N-末端和A1结构域、N-末端和A2结构域、N-末端和A3结构域、N-末端和B结构域、N-末端和C1结构域、N-末端和C2结构域、N-末端和C-末端、A1和A2结构域、A1和A3结构域、A1和B结构域、A1和C1结构域、A1和C2结构域、A1结构域和C-末端、A2和A3结构域、A2和B结构域、A2和C1结构域、A2和C2结构域、A2结构域和C-末端、A3和B结构域、A3和C1结构域、A3和C2结构域、A3结构域和C-末端、B和C1结构域、B和C2结构域、B结构域和C-末端、C1和C2结构域、C1和C-末端、C2结构域和C-末端以及其两个或更多的组合。插入位点的非限制性实例于表7、8、9、10、11、12、13、14和15中列出。

[0384] FVIII蛋白，其中XTEN序列紧邻FVIII蛋白中一个或多个氨基酸(例如，一个或多个XTEN插入位点)的下游插入或在C-末端或N-末端连接，在连接于XTEN序列或被XTEN序列插入后保留FVIII活性。XTEN序列可插入到FVIII蛋白中一次或多于一次，两次、三次、四次、五次或六次以使插入不影响FVIII活性(即，FVIII蛋白仍保留凝结性质)。

[0385] 可用于本发明的FVIII蛋白可在FVIII蛋白的N-末端或C-末端通过任选的接头连接于一个或多个XTEN多肽或被一个或多个任选的接头紧邻FVIII蛋白中一个或多个氨基酸(例如，一个或多个XTEN插入位点)的下游插入。在一个实施方案中，XTEN序列插入于其中的两个氨基酸残基或与XTEN序列连接的氨基酸残基对应于SEQ ID NO:4[全长成熟FVIII]的两个或一个氨基酸残基，其选自由以下组成的组：在表7、表8、表9和表10中的残基及其任意组合。

[0386] 在其它实施方案中，至少一个XTEN序列被插入到任何一个或多个本文公开的XTEN插入位点或其任意组合中。一方面，至少一个XTEN序列被插入到表7中公开的一个或多个氨基酸中公开的一个或多个XTEN插入位点中。

[0387] 表7: 示例性的XTEN插入位点

[0388]

No.	XTEN 插入点*	插入残基	FVIII BDD 下游序列	FVIII 结构域
1	0	(N- 末端)	ATR	A1
2	3	R	RYY	A1
3	17	M	QSD	A1
4	18	Q	SDL	A1
5	22	G	ELP	A1
6	24	L	PVD	A1
7	26	V	DAR	A1
8	28	A	RFP	A1
9	32	P	RVP	A1
10	38	F	PFN	A1
11	40	F	NTS	A1
12	41	N	TSV	A1
13	60	N	IAK	A1
14	61	I	AKP	A1
15	65	R	PPW	A1
16	81	Y	DTV	A1
17	111	G	AEY	A1
18	116	D	QTS	A1
19	119	S	QRE	A1
20	120	Q	REK	A1
21	128	V	FPG	A1
22	129	F	PGG	A1
23	130	P	GGG	A1
24	182	G	SLA	A1
25	185	A	KEK	A1
26	188	K	TQT	A1
27	205	G	KSW	A1
28	210	S	ETK	A1
29	211	E	TKN	A1
30	216	L	MQD	A1
31	220	R	DAA	A1
32	222	A	ASA	A1
33	223	A	SAR	A1
34	224	S	ARA	A1
35	230	K	MHT	A1
36	243	P	GLI	A1
37	244	G	LIG	A1
38	250	R	KSV	A1
39	318	D	GME	A1
40	333	P	QLR	A1
42	334	Q	LRM	A1
43	336	R	MKN	a1

[0389]

No.	XTEN 插入点*	插入残基	FVIII BDD 下游序列	FVIII 结构域
44	339	N	NEE	a1
45	345	D	YDD	a1
46	357	V	VRF	a1
47	367	S	FIQ	a1
48	370	S	RPY	a1
49	375	A	KKH	A2
50	376	K	KHP	A2
51	378	H	PKT	A2
52	399	V	LAP	A2
53	403	D	DRS	A2
54	405	R	SYK	A2
55	409	S	QYL	A2
56	416	P	QRI	A2
57	434	E	TFK	A2
58	438	T	REA	A2
59	441	A	IQH	A2
60	442	I	QHE	A2
61	463	I	IFK	A2
62	487	Y	SRR	A2
63	490	R	LPK	A2
64	492	P	KGV	A2
65	493	K	GVK	A2
66	494	G	VKH	A2
67	500	D	FPI	A2
68	506	G	EIF	A2
69	518	E	DGP	A2
70	556	K	ESV	A2
71	565	Q	IMS	A2
72	566	I	MSD	A2
73	598	P	AGV	A2
74	599	A	GVQ	A2
75	603	L	EDP	A2
76	616	S	ING	A2
77	686	G	LWI	A2
78	713	K	NTG	A2
79	719	Y	EDS	A2
80	730	L	LSK	A2
81	733	K	NNA	A2
82	745	N	PPV**	B
83	1640	P	PVL	B
84	1652	R	TTL	B
85	1656	Q	SDQ	A3
86	1685	N	QSP	A3
87	1711	M	SSS	A3
88	1713	S	SPH	A3

[0390]

No.	XTEN 插入点*	插入残基	FVIII BDD 下游序列	FVIII 结构域
89	1720	N	RAQ	A3
90	1724	S	GSV	A3
91	1725	G	SVP	A3
92	1726	S	VPQ	A3
93	1741	G	SFT	A3
94	1744	T	QPL	A3
95	1749	R	GEL	A3
96	1773	V	TFR	A3
97	1792	Y	EED	A3
98	1793	E	EDQ	A3
99	1796	Q	RQG	A3
100	1798	Q	GAE	A3
101	1799	G	AEP	A3
102	1802	P	RKN	A3
103	1803	R	KNF	A3
104	1807	V	KPN	A3
105	1808	K	PNE	A3
106	1827	K	DEF	A3
107	1844	E	KDV	A3
108	1861	N	TLN	A3
109	1863	L	NPA	A3
110	1896	E	RNC	A3
111	1900	R	APC	A3
112	1904	N	IQM	A3
113	1905	I	QME	A3
114	1910	P	TFK	A3
115	1920	A	ING	A3
116	1937	D	QRI	A3
117	1981	G	VFE	A3
118	2019	N	KCQ	A3
119	2020	K	CQT	C1
120	2044	G	QWA	C1
121	2068	F	SWI	C1
122	2073	V	DLL	C1
123	2090	R	QKF	C1
124	2092	K	FSS	C1
125	2093	F	SSL	C1
126	2111	K	WQT	C1
127	2115	Y	RGN	C1
128	2120	T	GTL	C1
129	2125	V	FFG	C1
130	2171	L	NSC	C1
131	2173	S	CSM	C2
132	2188	A	QIT	C2
133	2223	V	NNP	C2

[0391]

No.	XTEN 插入点*	插入残基	FVIII BDD 下游序列	FVIII 结构域
134	2224	N	NPK	C2
135	2227	K	EWL	C2
136	2268	G	HQW	C2
137	2277	N	GKV	C2
138	2278	G	KVK	C2
139	2290	F	TPV	C2
140	2332	Y	FVIII的C末端	CT

[0392] *指示基于成熟全长人FVIII的氨基酸数目的XTEN的插入点,其中插入可在所指示氨基酸的N-或C-末端侧上。

[0393] 在一些实施方案中,一个或多个XTEN序列被插入在从对应于SEQ ID NO:4的氨基酸32、220、224、336、339、399、416、603、1656、1711、1725、1905或1910向上或向下的约六个氨基酸之内,或其任意组合。

[0394] 表8.示例性的XTEN插入范围

[0395]

No.	XTEN 插入点	插入残基	FVIII BDD 下游序列	FVIII 结构域	距插入 残基的 距离*
9	32	P	RVP	A1	-3, +6
31	220	R	DAA	A1	-
34	224	S	ARA	A1	+5
43	336	R	MKN	a1	-1, +6
44	339	N	NEE	a1	-4, +5
52	399	V	LAP	A2	-6, +3
56	416	P	QRI	A2	+6
75	603	L	EDP	A2	6, +6
85	1656	Q	SDQ	B	-3, +6
87	1711	M	SSS	A3	-6, +1
91	1725	G	SVP	A3	+6
113	1905	I	QME	A3	+6
114	1910	P	TFK	A3	-5, +6

[0396] *距插入残基的距离是指距可进行插入的指定插入残基(残基"0")的N-末端(负数)或C-末端(正数)的氨基酸相对数目。名称"-x"是指在指定插入残基的N-末端侧上相隔x个氨基酸的插入位点。类似地,名称"+x"是指在指定插入残基的C-末端侧上相隔x个氨基酸的插入位点。

[0397] 例如,“-1,+2”指示插入在标记为-1、0、+1或+2的氨基酸残基的N-末端或C-末端处进行。

[0398] 在其它实施方案中,一个或多个XTEN序列紧邻一个或多个对应于全长成熟人FVIII的氨基酸(选自由表9中的一个或多个插入位点组成的组)的下游插入。

[0399] 表9.示例性的XTEN插入位点或范围

[0400]

No.	XTEN插入点范围*	第一插入残基	FVIII结构域
-----	------------	--------	----------

3	18-32	Q	A1
8	40	F	A1
18	211-224	E	A1
27	336-403	R	A1,A2
43	599	A	A2
47	745-1640	N	B
50	1656-1728	Q	B,a3,A3
57	1796-1804	R	A3
65	1900-1912	R	A3
81	2171-2332	L	C1,C2

[0401] *指示相对于成熟人FVIII的氨基酸编号进行编号的插入位点的范围

[0402] 在另外其它实施方案中,一个或多个XTEN插入到FVIII的B结构域。在一个实例中,XTEN在对应于SEQ ID NO:4的氨基酸740和1640之间插入,其中在氨基酸740和1640之间的FVIII序列任选地不存在。在另一实例中,XTEN在对应于SEQ ID NO:4的氨基酸741和1690之间插入,其中在氨基酸740和1690之间的FVIII序列任选地不存在。在其它实例中,XTEN在对应于SEQ ID NO:4的氨基酸741和1648之间插入,其中在氨基酸741和1648之间的FVIII序列任选地不存在。在另外其它实例中,XTEN在对应于SEQ ID NO:4的氨基酸743和1638之间插入,其中在氨基酸743和1638之间的FVIII序列任选地不存在。在又其它实例中,XTEN在对应于SEQ ID NO:4的氨基酸745和1656之间插入,其中在氨基酸745和1656之间的FVIII序列任选地不存在。在一些实例中,XTEN在对应于SEQ ID NO:4的氨基酸745和1657之间插入,其中在氨基酸745和1657之间的FVIII序列任选地不存在。在某些实例中,XTEN在对应于SEQ ID NO:4的氨基酸745和1667之间插入,其中在氨基酸745和1667之间的FVIII序列任选地不存在。在又其它实例中,XTEN在对应于SEQ ID NO:4的氨基酸745和1686之间插入,其中在氨基酸745和1686之间的FVIII序列任选地不存在。在一些其它实例中,XTEN在对应于SEQ ID NO:4的氨基酸747和1642之间插入,其中在氨基酸747和1642之间的FVIII序列任选地不存在。在又其它实例中,XTEN在对应于SEQ ID NO:4的氨基酸751和1667之间插入,其中在氨基酸751和1667之间的FVIII序列任选地不存在。

[0403] 在一些实施方案中,一个或多个XTEN插入到一个或多个紧邻选自表10中的氨基酸残基组成的组的插入位点的氨基酸的下游的氨基酸。

[0404] 表10:FVIII XTEN插入位点和构建体名称

[0405]

构建体 编号	结构域	上游残基 No.*	下游残基 No.*	上游序列	下游序列
F8X-1	A1	3	4	ATR	RYY
F8X-2	A1	18	19	YMQ	SDL
F8X-3	A1	22	23	DLG	ELP
F8X-4	A1	26	27	LPV	DAR
F8X-5	A1	40	41	FPF	NTS
F8X-6	A1	60	61	LFN	IAK
F8X-7	A1	116	117	YDD	QTS
F8X-8	A1	130	131	VFP	GGs
F8X-9	A1	188	189	KEK	TQT
F8X-10	A1	216	217	NSL	MQD
F8X-11	A1	230	231	WPK	MHT
F8X-12	A1	333	334	EEP	QLR
F8X-13	A2	375	376	SVA	KKH
F8X-14	A2	403	404	APD	DRS
F8X-15	A2	442	443	EAI	QHE
F8X-16	A2	490	491	RRL	PKG
F8X-17	A2	518	519	TVE	DGP
F8X-18	A2	599	600	NPA	GVQ
F8X-19	A2	713	714	CDK	NTG
F8X-20	BD	745	746	SQN	PPV
F8X-21	BD	745	746	SQN	PPV
F8X-22	BD**	745	746	SQN	PPV
F8X-23	A3	1720	1721	APT	KDE
F8X-24	A3	1796	1797	EDQ	RQG
F8X-25	A3	1802	1803	AEP	RKN
F8X-26	A3	1827	1828	PTK	DEF
F8X-27	A3	1861	1862	HTN	TLN
F8X-28	A3	1896	1897	NME	RNC
F8X-29	A3	1900	1901	NCR	APC
F8X-30	A3	1904	1905	PCN	IQM
F8X-31	A3	1937	1938	AQD	QRI
F8X-32	C1	2019	2020	YSN	KCQ
F8X-33	C1	2068	2069	EPF	SWI
F8X-34	C1	2111	2112	GKK	WQT
F8X-35	C1	2120	2121	NST	GTL
F8X-36	C2	2171	2172	CDL	NSC
F8X-37	C2	2188	2189	SDA	QIT
F8X-38	C2	2227	2228	NPK	EWL
F8X-39	C2	2277	2278	FQN	GKV
F8X-40	CT	2332	NA	DLY	NA
F8X-41	CT	2332	NA	DLY	NA
F8X-42	A1	3	4	ATR	ATR
pSD0001	A2	403	404		
pSD0002	A2	599	600		

[0406]

pSD0021	N-末端	0	1		
pSD0022	A1	32	33		
pSD0023	A1	65	66		
pSD0024	A1	81	82		
pSD0025	A1	119	120		
pSD0026	A1	211	212		
pSD0027	A1	220	221		
pSD0028	A1	224	225		
pSD0029	A1	336	337		
pSD0030	A1	339	340		
pSD0031	A2	378	379		
pSD0032	A2	399	400		
pSD0033	A2	409	410		
pSD0034	A2	416	417		
pSD0035	A2	487	488		
pSD0036	A2	494	495		
pSD0037	A2	500	501		
pSD0038	A2	603	604		
pSD0039	A3	1656	1657		
pSD0040	A3	1711	1712		
pSD0041	A3	1725	1726		
pSD0042	A3	1749	1750		
pSD0043	A3	1905	1906		
pSD0044	A3	1910	1911		
pDS0062	A3	1900	1901		

[0407] *指示成熟FVIII蛋白的氨基酸数目

[0408] 在一个实施方案中,一个或多个XTEN插入位点位于FVIII蛋白的一个或多个表面暴露的柔性环结构(例如,受纳环)内。例如,至少一个XTEN序列可插入到包含至少两个“受纳环”的各FVIII“A”结构域中,至少一个XTEN多肽可插入所述受纳环中而不消除重组蛋白的促凝血活性、或待在宿主细胞中体内或体外表达的重组蛋白的能力。受纳环为允许插入至少一个XTEN序列的区域,其除其他属性之外还具有高表面或溶剂暴露和高构象柔性。A1结构域包括受纳环-1(A1-1)区域和受纳环-2(A1-2)区域、A2结构域包括受纳环-1(A2-1)区域和受纳环-2(A2-2)区域、A3结构域包括受纳环-1(A3-1)区域和受纳环-2(A3-2)区域。

[0409] 一方面,FVIII A1结构域(A1-1)中的第一受纳环位于β链1和β链2之间,并且FVIII A2结构域(A1-2)中的第二受纳环位于β链11和β链12之间。FVIII A2结构域(A2-1)中的第一受纳环位于β链22和β链23之间,并且FVIII A2结构域(A2-2)中的第二受纳环位于β链32和β链33之间。FVIII A3结构域(A3-1)中的第一受纳环位于β链38和β链39之间,并且FVIII A3(A3-2)中的第二受纳环位于β链45和β链46之间。在某些方面,包含A1-1的表面暴露的柔性环结构对应于天然的成熟FVIII中SEQ ID NO:4的约氨基酸15至约氨基酸45,例如SEQ ID NO:4的约氨基酸18至约氨基酸41的区域。在其它方面,包含A1-2的表面暴露的柔性环结构对应于天然的成熟FVIII中SEQ ID NO:4的约氨基酸201至约氨基酸232,例如SEQ ID NO:4的约氨基酸218至约氨基酸229的区域。在又其它方面,包含A2-1的表面暴露的柔性环结构对应于天然的成熟FVIII中SEQ ID NO:4的约氨基酸395至约氨基酸421,例如SEQ ID NO:4的约氨基酸397至约氨基酸418的区域。在又其它实施方案中,包含A2-2的表面暴露的柔性

环结构对应于天然的成熟人FVIII中SEQ ID NO:4的约氨基酸577至约氨基酸635,例如SEQ ID NO:4的约氨基酸595至约氨基酸607的区域。在某些方面,包含A3-1的表面暴露的柔性环结构对应于天然的成熟人FVIII中SEQ ID NO:4约氨基酸1705至约氨基酸1732,例如SEQ ID NO:4的约氨基酸1711至约氨基酸1725的区域。在又其它方面,包含A3-2的表面暴露的柔性环结构对应于天然的成熟人FVIII中SEQ ID NO:4的约氨基酸1884至约氨基酸1917,例如SEQ ID NO:4的约氨基酸1899至约氨基酸1911的区域。

[0410] 在另一个实施方案中,至少一个XTEN序列插入于其中的一个或多个氨基酸位于a3结构域内,例如对应于全长成熟FVIII多肽的氨基酸1649至1689。在特定实施方案中,XTEN序列在SEQ ID NO:4(全长成熟FVIII)的氨基酸1656至1657之间插入。在具体实施方案中,包含紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入的XTEN序列的FVIII蛋白还包含对应于SEQ ID NO:4的氨基酸745至氨基酸1656的缺失。

[0411] 在一些实施方案中,供一个或多个XTEN插入的一个或多个插入位点紧邻一个或多个选自自由以下组成的组的氨基酸的下游:

- [0412] (1) 氨基酸3、(2) 氨基酸18、(3) 氨基酸22、
[0413] (4) 氨基酸26、(5) 氨基酸32、(6) 氨基酸40、
[0414] (7) 氨基酸60、(8) 氨基酸65、(9) 氨基酸81、
[0415] (10) 氨基酸116、(11) 氨基酸119、(12) 氨基酸130、
[0416] (13) 氨基酸188、(14) 氨基酸211、(15) 氨基酸216、
[0417] (16) 氨基酸220、(17) 氨基酸224、(18) 氨基酸230、
[0418] (19) 氨基酸333、(20) 氨基酸336、(21) 氨基酸339、
[0419] (22) 氨基酸375、(23) 氨基酸399、(24) 氨基酸403、
[0420] (25) 氨基酸409、(26) 氨基酸416、(26) 氨基酸442、
[0421] (28) 氨基酸487、(29) 氨基酸490、(30) 氨基酸494、
[0422] (31) 氨基酸500、(32) 氨基酸518、(33) 氨基酸599、
[0423] (34) 氨基酸603、(35) 氨基酸713、(36) 氨基酸745、
[0424] (37) 氨基酸1656、(38) 氨基酸1711、(39) 氨基酸1720、
[0425] (40) 氨基酸1725、(41) 氨基酸1749、(42) 氨基酸1796、
[0426] (43) 氨基酸1802、(44) 氨基酸1827、(45) 氨基酸1861、
[0427] (46) 氨基酸1896、(47) 氨基酸1900、(48) 氨基酸1904、
[0428] (49) 氨基酸1905、(50) 氨基酸1910、(51) 氨基酸1937、
[0429] (52) 氨基酸2019、(53) 氨基酸2068、(54) 氨基酸2111、
[0430] (55) 氨基酸2120、(56) 氨基酸2171、(57) 氨基酸2188、
[0431] (58) 氨基酸2227、(59) 氨基酸2277和
[0432] (60) 其两个或更多个的组合。

[0433] 在一个实施方案中,可用于本发明的FVIII蛋白包含两个XTEN序列,第一XTEN序列插入到第一XTEN插入位点且第二XTEN插入到第二XTEN插入位点。第一XTEN插入位点和第二XTEN插入位点的非限制性实例于表11中列出。

[0434] 表11. 两个XTEN的示例性插入位点

[0435]

插入 1		插入 2	
插入位点	结构域	插入位点	结构域
745	B	2332	CT
26	A1	403	A2
40	A1	403	A2
18	A1	403	A2
26	A1	599	A2
40	A1	599	A2
18	A1	599	A2
1720	A3	1900	A3
1725	A3	1900	A3
1711	A3	1905	A3
1720	A3	1905	A3
1725	A3	1905	A3
1656	A3	26	A1
1656	A3	18	A1
1656	A3	40	A1
1656	A3	399	A2
1656	A3	403	A2
1656	A3	1725	A3
1656	A3	1720	A3
1900	A3	18	A1
1900	A3	26	A1
1900	A3	40	A1
1905	A3	18	A1
1905	A3	40	A1
1905	A3	26	A1
1910	A3	26	A1
18	A1	399	A2
26	A1	399	A2
40	A1	399	A2
18	A1	403	A2

	插入 1		插入 2	
	插入位点	结构域	插入位点	结构域
[0436]	1656	A3	1900	A3
	1656	A3	1905	A3
	1711	A3	40	A1
	1711	A3	26	A1
	1720	A3	26	A1
	1720	A3	40	A1
	1720	A3	18	A1
	1725	A3	26	A1
	1725	A3	40	A1
	1725	A3	18	A1
	1720	A3	403	A2
	1720	A3	399	A2
	1711	A3	403	A2
	1720	A3	403	A2
	1725	A3	403	A2
	1725	A3	399	A2
	1711	A3	403	A2
	1900	A3	399	A2
	1900	A3	403	A2
	1905	A3	403	A2
	1905	A3	399	A2
	1910	A3	403	A2

[0437] 插入或连接于FVIII蛋白的两个XTEN可以相同或不同。在一些实施方案中,可用于本发明的FVIII蛋白包含插入到FVIII蛋白的两个XTEN序列,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸745的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸2332的下游(C-末端)插入。在其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26、40、1656、或1720的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸403的下游插入。在另外其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26或40的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸599的下游插入。在又其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26、40、399、403、1725、1720、1900、1905或2332的下游插入。在某些实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸1900的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26或40的下游插入。在一些实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26或40的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸399的下游插入。在其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26或40的下游插入。在又其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸1720的下游插入,并且第二XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18的下游插入。在特定实施方案中,FVIII蛋白包含两个XTEN序列,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸745的下游插入并且第二XTEN序列紧

邻对应于SEQ ID NO:4的氨基酸2332的下游插入,其中FVIII蛋白还具有对应于SEQ ID NO:4的氨基酸745至对应于SEQ ID NO:4的氨基酸1685的缺失、在对应于SEQ ID NO:4的氨基酸1680处的突变或替代(例如Y1680F)、在对应于SEQ ID NO:4的氨基酸1648处的突变或替代(例如R1648A)、或在对应于SEQ ID NO:4的氨基酸1648(例如R1648A)和对应于SEQ ID NO:4的氨基酸1680(例如Y1680F)处的至少两种突变或替代。在具体实施方案中,FVIII蛋白包含两个XTEN序列,第一XTEN紧邻对应于SEQ ID NO:4的氨基酸1656的下游插入并且第二XTEN序列紧邻SEQ ID NO:4的氨基酸2332的下游插入,其中FVIII蛋白还具有对应于SEQ ID NO:4的氨基酸745至氨基酸1656的缺失。

[0438] 在某些实施方案中,FVIII蛋白包含三个XTEN序列,第一XTEN序列插入到第一XTEN插入位点,第二XTEN序列插入到第二XTEN序列,并且第三XTEN序列插入到第三XTEN插入位点。第一、第二或第三XTEN序列可以相同或不同。第一、第二和第三插入位点可以选自本文公开的插入位点中的任一个的组。在一些实施方案中,包含三个XTEN序列的FVIII蛋白还可包含突变或替代,例如对应于SEQ ID NO:4的氨基酸1648,例如R1648A。例如,第一、第二和第三XTEN插入位点的非限制性实例于表12中列出。

[0439] 表12. 三个XTEN的示例性插入位点

[0440]

插入 1		插入 2		插入 3	
插入位点	结构域	插入位点	结构域	插入位点	结构域
26	A1	403	A2	1656	A3
26	A1	403	A2	1720	A3
26	A1	403	A2	1900	A3
26	A1	1656	A3	1720	A3
26	A1	1656	A3	1900	A3
26	A1	1720	A3	1900	A3
403	A2	1656	A3	1720	A3
403	A2	1656	A3	1900	A3
403	A2	1720	A3	1900	A3
1656	A3	1720	A3	1900	A3
745	B	1900		2332	CT
18	A1	745	B	2332	CT
26	A1	745	B	2332	CT
40	A1	745	B	2332	CT
18	A1	745	B	2332	CT
40	A1	745	B	2332	CT
403	A2	745	B	2332	CT
399	A2	745	B	2332	CT
1725	A3	745	B	2332	CT
1720	A3	745	B	2332	CT
1711	A3	745	B	2332	CT
1900	A3	745	B	2332	CT
1905	A3	745	B	2332	CT
1910	A3	745	B	2332	CT

[0441] 在一些实施方案中,FVIII蛋白包含三个XTEN序列,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,第二XTEN序列插入对应于SEQ ID NO:4的氨基酸403的下

游,并且第三XTEN序列插入对应于SEQ ID NO:4的氨基酸1656、1720或1900的下游。在其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,第二XTEN序列插入对应于SEQ ID NO:4的氨基酸1656的下游,并且第三XTEN序列插入对应于SEQ ID NO:4的氨基酸1720或1900的下游。在另外其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸26的下游插入,第二XTEN序列插入对应于SEQ ID NO:4的氨基酸1720的下游,并且第三XTEN序列插入对应于SEQ ID NO:4的氨基酸1900的下游。在又其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸403的下游插入,第二XTEN序列插入对应于SEQ ID NO:4的氨基酸1656的下游,并且第三XTEN序列插入对应于SEQ ID NO:4的氨基酸1720或1900的下游。在其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸403或1656的下游插入,第二XTEN序列插入对应于SEQ ID NO:4的氨基酸1720的下游,并且第三XTEN序列插入对应于SEQ ID NO:4的氨基酸1900的下游。在其它实施方案中,第一XTEN序列紧邻对应于SEQ ID NO:4的氨基酸18、26、40、399、403、1711、1720、1725、1900、1905或1910的下游插入,第二XTEN序列插入对应于SEQ ID NO:4的氨基酸745的下游,并且第三XTEN序列插入对应于SEQ ID NO:4的氨基酸2332的下游。

[0442] 在其它实施方案中,本发明中的FVIII蛋白包含四个XTEN序列,第一XTEN序列插入到第一插入位点,第二XTEN序列插入到第二插入位点,第三XTEN序列插入到第三插入位点,并且第四XTEN序列插入到第四插入位点。第一、第二、第三和第四XTEN序列可以是相同、不同或其组合。在一些实施方案中,包含第四XTEN序列的FVIII蛋白还可包含突变或替代,例如对应于SEQ ID NO:4的氨基酸1648,例如R1648A。第一、第二、第三和第四XTEN插入位点的非限制性于表13中列出。

[0443] 表13.四个XTEN的示例性插入位点

[0444]

插入 1		插入 2		插入 3		插入 4	
插入位点	结构域	插入位点	结构域	插入位点	结构域	插入位点	结构域
26	A1	403	A2	1656	a3	1720	A3
26	A1	403	A2	1656	a3	1900	A3
26	A1	403	A2	1720	A3	1900	A3
26	A1	1656	a3	1720	A3	1900	A3
403	A2	1656	a3	1720	A3	1900	A3
0040	A1	0403	A2	745	B	2332	CT
0040	A1	0403	A2	745	B	2332	CT
0018	A1	0409	A2	745	B	2332	CT
0040	A1	0409	A2	745	B	2332	CT
0040	A1	0409	A2	745	B	2332	CT
0018	A1	0409	A2	745	B	2332	CT
0040	A1	1720	A3	745	B	2332	CT
0026	A1	1720	A3	745	B	2332	CT
0018	A1	1720	A3	745	B	2332	CT
0018	A1	1720	A3	745	B	2332	CT
0018	A1	1720	A3	745	B	2332	CT
0026	A1	1720	A3	745	B	2332	CT
0018	A1	1720	A3	745	B	2332	CT
0018	A1	1900	A3	745	B	2332	CT
0018	A1	1900	A3	745	B	2332	CT
0026	A1	1900	A3	745	B	2332	CT
0040	A1	1900	A3	745	B	2332	CT
0040	A1	1905	A3	745	B	2332	CT
0018	A1	1905	A3	745	B	2332	CT
0040	A1	1905	A3	745	B	2332	CT
0026	A1	1905	A3	745	B	2332	CT
0018	A1	1905	A3	745	B	2332	CT
0018	A1	1905	A3	745	B	2332	CT
0018	A1	1910	A3	745	B	2332	CT
0018	A1	1910	A3	745	B	2332	CT
0040	A1	1910	A3	745	B	2332	CT
0026	A1	1910	A3	745	B	2332	CT
0018	A1	1910	A3	745	B	2332	CT
0026	A1	1910	A3	745	B	2332	CT
0040	A1	1910	A3	745	B	2332	CT
0018	A1	1910	A3	745	B	2332	CT
0409	A2	1720	A3	745	B	2332	CT
0403	A2	1720	A3	745	B	2332	CT
0409	A2	1720	A3	745	B	2332	CT
0403	A2	1720	A3	745	B	2332	CT

[0445]

0403	A2	1720	A3	745	B	2332	CT
0403	A2	1900	A3	745	B	2332	CT
0403	A2	1900	A3	745	B	2332	CT
0409	A2	1900	A3	745	B	2332	CT
0403	A2	1900	A3	745	B	2332	CT
0403	A2	1900	A3	745	B	2332	CT
0409	A2	1900	A3	745	B	2332	CT
0409	A2	1905	A3	745	B	2332	CT
0403	A2	1905	A3	745	B	2332	CT
0403	A2	1905	A3	745	B	2332	CT
0403	A2	1905	A3	745	B	2332	CT
0409	A2	1905	A3	745	B	2332	CT
0403	A2	1905	A3	745	B	2332	CT
0409	A2	1910	A3	745	B	2332	CT
0403	A2	1910	A3	745	B	2332	CT
0403	A2	1910	A3	745	B	2332	CT
0403	A2	1910	A3	745	B	2332	CT
0403	A2	1910	A3	745	B	2332	CT
0403	A2	1910	A3	745	B	2332	CT
1720	A3	1900	A3	745	B	2332	CT
1720	A3	1905	A3	745	B	2332	CT
1720	A3	1910	A3	745	B	2332	CT
1720	A3	1910	A3	745	B	2332	CT
0403	A2	1656	a3	1720	A3	2332	CT
0403	A2	1656	a3	1900	A3	2332	CT
0403	A2	1720	A3	1900	A3	2332	CT
1656	a3	1720	A3	1900	A3	2332	CT
0018	A1	0403	A2	1656	a3	2332	CT
0018	A1	0403	A2	1720	A3	2332	CT
0018	A1	0403	A2	1900	A3	2332	CT
0018	A1	1656	a3	1720	A3	2332	CT
0018	A1	1656	a3	1900	A3	2332	CT
0018	A1	1720	A3	1900	A3	2332	CT
0018	A1	0403	A2	0745	B	2332	CT
0018	A1	0745	B	1720	A3	2332	CT
0018	A1	0745	B	1900	A3	2332	CT
0403	A2	0745	B	1720	A3	2332	CT
0403	A2	0745	B	1900	A3	2332	CT
0745	B	1720	A3	1900	A3	2332	CT
0188	A1	1900	A3	0745	B	2332	CT
0599		1900	A3	0745	B	2332	CT
2068		1900	A3	0745	B	2332	CT
2171		1900	A3	0745	B	2332	CT
2227		1900	A3	0745	B	2332	CT
2277		1900	A3	0745	B	2332	CT

[0446] 在一些实施方案中, FVIII蛋白包含五个XTEN序列, 第一XTEN序列插入到第一插入位点, 第二XTEN序列插入到第二插入位点, 第三XTEN序列插入到第三XTEN插入位点, 第四XTEN序列插入到第四XTEN插入位点且第五XTEN序列插入到第五XTEN插入位点。第一、第二、第三、第四或第五XTEN序列可以相同、不同或其组合。第一、第二、第三、第四和第五插入位点的非限制性实例于表14中列出。

[0447] 表14. 五个XTEN的示例性插入位点

XTEN 插入 1	XTEN 插入 2	XTEN 插入 3	XTEN 插入 4	XTEN 插入 5
0403	1656	1720	1900	2332
0018	0403	1656	1720	2332
0018	0403	1656	1900	2332
0018	0403	1720	1900	2332
0018	1656	1720	1900	2332
0018	0403	0745	1720	2332
0018	0403	0745	1900	2332
0018	0745	1720	1900	2332
0403	0745	1720	1900	2332

[0449] 在某些实施方案中, FVIII蛋白包含六个XTEN序列, 第一XTEN序列插入到第一XTEN插入位点, 第二XTEN序列插入到第二XTEN插入位点, 第三XTEN序列插入到第三XTEN插入位点, 第四XTEN序列插入到第四XTEN插入位点, 第五XTEN序列插入到第五XTEN插入位点, 并且第六XTEN序列插入到第六XTEN插入位点。第一、第二、第三、第四、第五或第六XTEN序列可以相同、不同或其组合。六个XTEN插入位点的实例包括但不限于于表15中列出的插入位点。

[0450] 表15. 六个XTEN的示例性XTEN插入位点

XTEN 插入 1	XTEN 插入 2	XTEN 插入 3	XTEN 插入 4	XTEN 插入 5	XTEN 插入 6
0018	0403	1656	1720	1900	2332
0018	0403	0745	1720	1900	2332

[0452] 在一个特定实例中, 第一XTEN在对应于SEQ ID NO:4的氨基酸26和27之间插入, 并且第二XTEN在对应于SEQ ID NO:4 (全长成熟FVIII) 的氨基酸1720和1721之间插入。在另一实例中, 第一XTEN在对应于SEQ ID NO:4的氨基酸403和404之间插入, 并且第二XTEN在对应于SEQ ID NO:4的氨基酸1720和1721之间插入。在一些实例中, 第一XTEN在对应于SEQ ID NO:4的氨基酸1656和1657之间插入, 并且第二XTEN在对应于SEQ ID NO:4的氨基酸1720和1721之间插入。在其它实例中, 第一XTEN在对应于SEQ ID NO:4的氨基酸26和27之间插入, 第二XTEN在对应于SEQ ID NO:4的氨基酸1656和1657之间插入, 并且第三XTEN在对应于SEQ ID NO:4的氨基酸1720和1721之间插入。在另外其它实施方案中, 第一XTEN在对应于SEQ ID NO:4的氨基酸403和404之间插入, 第二XTEN在对应于SEQ ID NO:4的氨基酸1656和1657之间插入, 并且第三XTEN在对应于SEQ ID NO:4的氨基酸1720和1721之间插入。在又其它实施方案中, 第一XTEN在对应于SEQ ID NO:4的氨基酸403和404之间插入, 第二XTEN在对应于SEQ ID NO:4的氨基酸1656和1657之间插入, 并且第三XTEN在对应于SEQ ID NO:4的氨基酸1720和1721之间插入。在某些实施方案中, 第一XTEN在对应于SEQ ID NO:4的氨基酸26和27之间插入, 第二XTEN在对应于SEQ ID NO:4的氨基酸1720和1721的之间插入, 并且第三XTEN在对应于SEQ ID NO:4的氨基酸1900和1901之间插入。在一些实施方案中, 第一XTEN在对应于SEQ ID NO:4的氨基酸26和27之间插入, 第二XTEN在对应于SEQ ID NO:4的氨基酸1656和1657之间插入, 第三XTEN在对应于SEQ ID NO:4的氨基酸1720和1721之间插入, 并且第四XTEN在对应于SEQ ID NO:4的1900和1901之间插入。

[0453] 在特定实施方案中,XTEN序列在全长因子VIII的氨基酸745和746之间或B-结构域缺失的因子VIII的相应的插入位点插入。

[0454] 在一些实施方案中,本发明的嵌合蛋白包含两个多肽序列,第一多肽序列包含与选自FVIII-161 (SEQ ID NO:101)、FVIII-169 (SEQ ID NO:103)、FVIII-170 (SEQ ID NO:102)、FVIII-173 (SEQ ID NO:104)、FVIII-195 (SEQ ID NO:105)、FVIII-196 (SEQ ID NO:106)、FVIII-199 (SEQ ID NO:107)、FVIII-201 (SEQ ID NO:108)、FVIII-203 (SEQ ID NO:109)、FVIII-204 (SEQ ID NO:110)、FVIII-205 (SEQ ID NO:111)、FVIII-266 (SEQ ID NO:112)、FVIII-267 (SEQ ID NO:113)、FVIII-268 (SEQ ID NO:114)、FVIII-269 (SEQ ID NO:115)、FVIII-271 (SEQ ID NO:116)或FVIII-272 (SEQ ID NO:117)的序列具有至少约80%、90%、95%或100%同一性的氨基酸序列并且第二多肽序列包含与选自VWF031 (SEQ ID NO:118)、VWF034 (SEQ ID NO:119)、或VWF-036 (SEQ ID NO:120)的序列具有至少约80%、90%、95%或100%同一性的氨基酸序列。

[0455] D) Ig恒定区或其一部分

[0456] 连接于本发明中的XTEN序列的VWF片段或FVIII蛋白还可包含Ig恒定区或其一部分。Ig恒定区或其一部分可改善与XTEN序列组合的VWF片段或FVIII蛋白的药理学或药效学性质。在某些实施方案中,Ig恒定区或其一部分延长融合于Ig恒定区或其一部分的分子的半衰期。

[0457] Ig恒定区为由称为CH(恒定重)结构域(CH1、CH2等)的结构域组成。根据同种型(即IgG、IgM、IgA、IgD或IgE),恒定区可由三个或四个CH结构域组成。一些同种型(例如IgG)恒定区还含有铰链区。参见Janeway等人2001,Immunobiology,Garland Publishing,N.Y.,N.Y。

[0458] 用于生成本发明的嵌合蛋白的Ig恒定区或其一部分可获自许多不同来源。在一些实施方案中,Ig恒定区或其一部分源自人Ig。但是,应理解,Ig恒定区或其一部分可源自另一哺乳动物物种的Ig,包括例如啮齿类动物(例如小鼠、大鼠、兔、豚鼠)或非人灵长类动物(例如黑猩猩、猕猴)物种。而且,Ig恒定区或其一部分可源自任何Ig类别(包括IgM、IgG、IgD、IgA和IgE),以及任何Ig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。

[0459] 多种Ig恒定区基因序列(例如,人恒定区基因序列)以公开访问的保藏形式可得。可以选择具有特定效应功能(或缺少特定效应功能)或具有特定修饰以减少免疫原性的恒定区结构域序列。抗体和抗体编码基因的很多序列已被公布并且适合的Ig恒定区序列(例如,铰链、CH2和/或CH3序列或其部分)可使用本领域认可的技术衍生自这些序列。然后可改变或合成使用任何前述方法获得的遗传物质以获得本发明的多肽。还应理解,本发明的范围涵盖恒定区DNA序列的等位基因、变体和突变。

[0460] 例如使用聚合酶链式反应和被选择以扩增目标结构域的引物可克隆Ig恒定区或其一部分的序列。为了克隆来自抗体的Ig恒定区或其一部分的序列,mRNA可从杂交瘤、脾或淋巴细胞分离,被逆转转录为DNA中,并且抗体基因由PCR扩增。PCR扩增方法详细地描述于美国专利号4,683,195;4,683,202;4,800,159;4,965,188;和于例如"PCR Protocols: A Guide to Methods and Applications" Innis等人编,Academic Press,San Diego,CA (1990);Ho等人1989.Gene 77:51;Horton等人1993.Methods Enzymol.217:270)中。PCR可

基于公布的重链和轻链DNA和氨基酸序列由共有恒定区引物或由更具特异性的引物启动。如上所论述,PCR还可用于分离编码抗体轻链和重链的DNA克隆。在这种情况下,文库可通过共有引物或较大的同源探针诸如小鼠恒定区探针进行筛选。适用于抗体基因扩增的许多引物组是本领域已知的(例如,基于纯化抗体的N-端序列的5'引物(Benhar和Pastan.1994.Protein Engineering 7:1509);cDNA端的快速扩增(Ruberti,F.等人1994.J.Immunol.Methods 173:33);抗体前导序列(Larrick等人1989Biochem.Biophys.Res.Commun.160:1250)。抗体序列的克隆进一步描述于Newman等人,美国专利号5,658,570,1995年1月25日,其通过引用并入本文中。

[0461] 本文所使用的Ig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,Ig恒定区或其一部分包含CH2结构域、CH3结构域和铰链区,即Fc区或FcRn结合配偶体。

[0462] 如本文所使用,术语“Fc区”被定义为多肽中对应于天然Ig的Fc区的部分,即,如通过其两条重链的各自Fc结构域的二聚缔合形成。天然Fc区与另一Fc区形成同二聚体。与此相反,术语“基因融合的Fc区”或“单链Fc区”(scFc区域),如本文所使用,是指由在单一多肽链(即,于单个连续基因序列中编码)内基因连接的Fc结构域组成的合成二聚Fc区。

[0463] 在一个实施方案中,“Fc区”是指单个Ig重链的部分,从铰链区中的正好木瓜蛋白酶切割位点的上游(即IgG中的残基216,取重链恒定区的第一残基为114)开始且在抗体的C-末端结束。因此,完整Fc结构域包含至少铰链结构域、CH2结构域和CH3结构域。

[0464] Ig恒定区的Fc区取决于Ig同种型可包括CH2、CH3和CH4结构域、以及铰链区。包含Ig的Fc区的嵌合蛋白赋予嵌合蛋白数种期望性质,包括增加的稳定性、增加的血清半衰期(参见Capon等人,1989,Nature 337:525)以及与Fc受体诸如新生Fc受体(FcRn)的结合(美国专利号6,086,875、6,485,726、6,030,613;WO 03/077834;US2003-0235536A1),其以引用方式整体并入本文。

[0465] Ig恒定区或其一部分可以为FcRn结合配偶体。FcRn在成人上皮组织中为活性的并且在肠、肺气道、鼻腔表面、阴道表面、结肠和直肠表面的内腔中表达(美国专利号6,485,726)。FcRn结合配偶体为结合于FcRn的Ig的一部分。

[0466] FcRn受体已从数种哺乳动物物种(包括人)分离。人FcRn、猴FcRn、大鼠FcRn和小鼠FcRn的序列是已知的(Story等人1994,J.Exp.Med.180:2377)。FcRn受体以相对低pH结合IgG(但不结合其它Ig类诸如IgA、IgM、IgD和IgE),使内腔中的IgG跨细胞地主动转运至浆膜方向,然后在相对较高的pH下释放于间质流体中发现的IgG。其在成人上皮组织中表达(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(Israel等人1997,Immunology 92:69)、肾脏近端管状上皮(Kobayashi等人2002,Am.J.Physiol.Renal Physiol.282:F358)以及鼻腔上皮、阴道表面和胆管树表面。

[0467] 可用于本发明的FcRn结合配偶体涵盖可被FcRn受体特异性地结合的分子,包括整个IgG、IgG的Fc片段和包括FcRn受体的完整结合区的其它片段。已基于X射线结晶学描述结合于FcRn受体的IgG的Fc部分的区域(Burmeister等人1994,Nature 372:379)。Fc与FcRn的主要接触区邻近CH2和CH3结构域的接合点。Fc-FcRn接触均在单个Ig重链内。FcRn结合配偶体包括整个IgG、IgG的Fc片段和IgG中包括FcRn的完整结合区的其它片段。主要接触位点包括CH2结构域的氨基酸残基248、250-257、272、285、288、290-291、308-311和314和CH3结构域的氨基酸残基385-387、428和433-436。参考的Ig或Ig片段、或者区域的氨基酸编号均基

于Kabat等人1991, Sequences of Proteins of Immunological Interest, U.S. Department of Public Health, Bethesda, Md.

[0468] 结合于FcRn的Fc区或FcRn结合配偶体可通过FcRn有效地跨过上皮屏障穿梭, 因而提供全身性施用所需治疗性分子的非侵入性方式。此外, 包含Fc区或FcRn结合配偶体的融合蛋白被表达FcRn的细胞胞饮。但是代替被标记用于降解, 这些融合蛋白被再循环以再次进入循环, 因而增加这些蛋白质的体内半衰期。在某些实施方案中, Ig恒定区的部分为Fc区或FcRn结合配偶体, 其通常经由二硫键和其它非特异性相互作用与另一Fc区或另一FcRn结合配偶体缔合以形成二聚体和更高等级的多聚体。

[0469] 两个FcRn受体可结合单个Fc分子。晶体学数据表明, 各FcRn分子结合Fc同二聚体的单一多肽。在一个实施方案中, FcRn结合配偶体, 例如IgG的Fc片段与生物活性分子的连接提供经口、经颊、舌下、直肠、阴道、作为经鼻施用气溶胶或经由肺途径、或经由眼内途径递送生物活性分子的手段。在另一个实施方案中, 嵌合蛋白可被侵入性施用, 例如皮下、静脉内。

[0470] FcRn结合配偶体区域为可被FcRn受体特异性结合、随后被Fc区的FcRn受体主动转运的分子或其一部分。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。特异性结合的特征在于高亲和力和低到中等程度的容量, 区别于通常具有低亲和力和中到高等程度的容量的非特异性结合。通常, 当亲和力常数 K_A 高于 $10^6 M^{-1}$ 、或高于 $10^8 M^{-1}$ 时结合被认为是特异性的。如有必要, 可在不显著影响特异性结合的情况下通过改变结合条件而减少非特异性结合。适当的结合条件诸如分子浓度、溶液离子强度、温度、允许结合的时间、封闭剂(例如血清白蛋白、乳酪蛋白)浓度等可由本领域技术人员使用常规技术进行优化。

[0471] 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的Fc区, 尽管如此其足以向Fc区赋予Fc受体(FcR)结合性质。例如, 结合于FcRn的Fc区的部分(即, FcRn结合部分)包含根据EU编号的IgG1的约氨基酸282-438(其中一级接触位点为CH2结构域的氨基酸248、250-257、272、285、288、290-291、308-311和314, 和CH3结构域的氨基酸残基385-387、428和433-436。因此, 本发明的Fc区可以包含以下或由以下组成: FcRn结合部分。FcRn结合部分可源自任何同种型(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。

[0472] 在另一个实施方案中, “Fc区”包括Fc结构域或源自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个: 铰链(例如, 较高、居中和/或较低的铰链区)结构域(根据EU编号的抗体Fc区的约氨基酸216-230)、CH2结构域(根据EU编号的抗体Fc区的约氨基酸231-340)、CH3结构域(根据EU编号的抗体Fc区的约氨基酸341-438)、CH4结构域、或其变体部分或片段。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区包含以下、基本上由以下组成或由以下组成: 融合于CH3结构域(或其一部分)的铰链结构域(或其一部分)、融合于CH2结构域(或其一部分)的铰链结构域(或其一部分)、融合于CH3结构域(或其一部分)的CH2结构域(或其一部分)、融合于铰链结构域(或其一部分)和CH3结构域(或其一部分)两者的CH2结构域(或其一部分)。在又其它实施方案中, Fc区缺少CH2结构域的至少一部分(例如, 所有或部分的CH2结构域)。在特定实施方案中, Fc区包含以下或由以下组成: 对应于EU编号221至447的氨基酸。

[0473] 在本文被表示为F、F1或F2的Fc区可获自许多不同来源。在一个实施方案中,多肽的Fc区源自人Ig。但是,应理解Fc区可源自另一哺乳动物物种的Ig,包括例如,啮齿类动物(例如小鼠、大鼠、兔或豚鼠)或非人灵长类动物(例如黑猩猩、猕猴)种类。而且,Fc结构域或其部分的多肽可源自任何Ig类别(包括IgM、IgG、IgD、IgA和IgE)以及任何Ig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。

[0474] 在某些实施方案中,Fc变体使由包含所述野生型Fc结构域的Fc区赋予的至少一种效应功能发生变化(例如,Fc区结合于Fc受体(例如Fc γ RI、Fc γ RII或Fc γ RIII)或补体蛋白(例如C1q)、或触发抗体-依赖性细胞毒性(ADCC)、吞噬或补体依赖性细胞毒性(CDCC)的能力改善或降低)。在其它实施方案中,Fc变体提供工程改造的半胱氨酸残基。

[0475] 本发明的Fc区可采用本领域公认的Fc变体,其已知可使效应功能和/或FcR或FcRn结合发生变化(例如增强或减少)。具体而言,本发明的结合分子可以包括,例如在以下文献公开的氨基酸位置中的一个或多个处的变化(例如,替代):国际PCT公布W088/07089A1、W096/14339A1、W098/05787A1、W098/23289A1、W099/51642A1、W099/58572A1、W000/09560A2、W000/32767A1、W000/42072A2、W002/44215A2、W002/060919A2、W003/074569A2、W004/016750A2、W004/029207A2、W004/035752A2、W004/063351A2、W004/074455A2、W004/099249A2、W005/040217A2、W004/044859、W005/070963A1、W005/077981A2、W005/092925A2、W005/123780A2、W006/019447A1、W006/047350A2和W006/085967A2;US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,所述文献各自通过引用并入本文。在一个实施方案中,可在一个或多个公开的氨基酸位置处进行特定变化(例如,本领域公开的一个或多个氨基酸的特定替代)。在另一个实施方案中,可在一个或多个公开的氨基酸位置处进行不同变化(例如,本领域公开的一个或多个氨基酸位置的不同替代)。

[0476] IgG的Fc区或FcRn结合配偶体可根据公认的程序诸如定向诱变等进行修饰以得到将由FcRn结合的经修饰的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在接触位点内的修饰,它们保持或甚至增强与FcRn的结合。例如,人IgG1Fc(Fc γ 1)中的下列单个氨基酸残基可在不显著损失Fc对于FcRn的结合亲和力的情况下被替代: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表示在位置编号238处被丙氨酸替代的野生型脯氨酸。例如,特定实施方案并入N297A突变,去除高度保守的N-糖基化位点。除了丙氨酸之

外,其它氨基酸可在上文指明的位置处被替代为野生型氨基酸。可简单地将突变引入Fc中,从而产生超过100个不同于天然Fc的Fc区。此外,可一起引入两个、三个或更多个这些个别突变的组合,从而产生超过一百个的Fc区。而且,本发明的构建体的Fc区中的一个可以发生突变并且构建体的其它Fc区根本不发生突变,或它们均可发生突变但具有不同突变。

[0477] 某些上述突变可对Fc区或FcRn结合配偶体赋予新功能性。例如,一个实施方案并入N297A,去除高度保守的N-糖基化位点。这种突变的效应为减少免疫原性,从而增强Fc区的循环半衰期,以及在不损害对FcRn的亲力的情况下使Fc区不能结合于Fc γ RI、Fc γ RIIA、Fc γ RIIB和Fc γ RIIIA(Routledge等人1995,Transplantation 60:847;Friend等人1999,Transplantation 68:1632;Shields等人1995,J.Biol.Chem.276:6591)。作为由上文所述的突变产生的新功能性的另一实例,在某些情况下对FcRn的亲力可比野生型有所增加。这种增加的亲力可反映增加的“结合(on)”速率,下降的“解离(off)”速率或者增加的“结合(on)”速率和下降的“解离(off)”速率两者。据信赋予对FcRn的亲力增加的突变的实例包括但不限于,T256A、T307A、E380A和N434A(Shields等人2001,J.Biol.Chem.276:6591)。

[0478] 此外,至少三个人Fc γ 受体似乎识别IgG上在较低铰链区内的结合位点,一般为氨基酸234-237。因此,新功能性和潜在的降低的免疫原性的另一实例可由这个区域的突变引起,如(例如)通过将人IgG1“ELLG”的氨基酸233-236替换为来自IgG2“PVA”的相应序列(具有一个氨基酸缺失)。已显示,当引入此类突变时介导各种效应功能的Fc γ RI、Fc γ RII和Fc γ RIII将不结合于IgG1.Ward和Ghetie 1995,Therapeutic Immunology 2:77和Armour等人1999,Eur.J.Immunol.29:2613。

[0479] 在一个实施方案中,Ig恒定区或其一部分例如Fc区为包括序列PKNSSMISNTP(SEQ ID NO:52)且任选地还包括选自HQSLGTQ(SEQ ID NO:53)、HQNLSDGK(SEQ ID NO:54)、HQNISDGK(SEQ ID NO:55)或VISSHLGQ(SEQ ID NO:56)的序列的多肽(美国专利号5,739,277)。

[0480] 在另一个实施方案中,免疫球蛋白恒定区或其一部分包含在铰链区或其一部分中的与另一免疫球蛋白恒定区或其一部分形成一个或多个二硫键的氨基酸序列。免疫球蛋白恒定区或其一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以使内源性VWF不替换VWF片段并且不结合于FVIII。因此,第一免疫球蛋白恒定区或其一部分和第二免疫球蛋白恒定区或其一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。VWF和FVIII蛋白之间的相互作用的这种抑制允许FVIII蛋白的半衰期超出两倍限值。铰链区或其一部分可进一步连接于CH1、CH2、CH3的一个或多个结构域,其片段及其任意组合。在特定实施方案中,免疫球蛋白恒定区或其一部分为铰链区和CH2。

[0481] 在某些实施方案中,Ig恒定区或其一部分为半糖基化的。例如,包含两个Fc区或FcRn结合配偶体的嵌合蛋白可含有第一、糖基化的Fc区(例如,糖基化的CH2区域)或FcRn结合配偶体和第二、未糖基化的Fc区(例如,未糖基化的CH2区域)或FcRn结合配偶体。在一个实施方案中,接头可在糖基化的Fc区和未糖基化的Fc区之间插入。在另一个实施方案中,Fc区或FcRn结合配偶体被完全糖基化,即所有Fc区被糖基化。在其它实施方案中,Fc区可以未被糖基化,即无一Fc部分被糖基化。

[0482] 在某些实施方案中,本发明的嵌合蛋白包含Ig恒定区或其一部分(例如,Fc变体)

的氨基酸替代,其改变Ig恒定区的抗原依赖性效应功能,特别地是蛋白质的循环半衰期。

[0483] 此类蛋白质展现出与FcRn的结合当相较于缺少这些替代的蛋白质时增加或减少,并且因此血清中的半衰期分别增加或减少。对FcRn的亲合力有所改善的Fc变体预期具有较长血清半衰期,并且此类分子应用于需要所施用多肽的长半衰期的哺乳动物的方法例如以治疗治疗慢性疾病或病症(参见,例如美国专利7,348,004、7,404,956和7,862,820)。与此相反,具有降低的FcRn结合亲和力的Fc变体预期具有较短半衰期,并且此类分子例如也用于向这样的哺乳动物施用:在所述哺乳动物中缩短的循环时间例如可对于体内诊断成像或在起始多肽当存在于循环中达延长的时间时具有毒性副作用的情况中是有益的。具有降低的FcRn结合亲和力的Fc变体也不太可能跨过胎盘,且因此还可用于治疗孕妇的疾病或病症。而且,可能需要降低的FcRn结合亲和力的其它应用包括需要定位脑、肾和/或肝的那些应用。在一个示例性实施方案中,本发明的嵌合蛋白展现出从脉管系统跨过肾小球上皮的转运减少。在另一个实施方案中,本发明的嵌合蛋白展现出从脑跨过血脑屏障(BBB)进入血管空间的转运减少。在一个实施方案中,具有改变的FcRn结合的蛋白质包含至少一个Fc区或FcRn结合配偶体(例如,一个或两个Fc区或FcRn结合配偶体),其具有一个或多个在Ig恒定区的“FcRn结合环”内的氨基酸替代。FcRn结合环由野生型全长Fc区的氨基酸残基280-299(根据EU编号)组成。在其它实施方案中,具有改变的FcRn结合亲和力的本发明的嵌合蛋白中的Ig恒定区或其一部分包含具有一个或多个在15 Å FcRn“接触区”内的氨基酸替代的至少一个Fc区或FcRn结合配偶体。如本文所使用,术语15 Å FcRn“接触区”包括在野生型全长Fc部分的下列位置处的残基:243-261、275-280、282-293、302-319、336-348、367、369、372-389、391、393、408、424、425-440(EU编号)。在其它实施方案中,具有改变的FcRn结合亲和力的本发明的Ig恒定区或其一部分包含具有一个或多个在对应于下列EU位置中的任一个的氨基酸位置处的氨基酸替代的至少一个Fc区或FcRn结合配偶体:256、277-281、283-288、303-309、313、338、342、376、381、384、385、387、434(例如,N434A或N434K)和438。示例性的改变FcRn结合活性的氨基酸替代公开于国际PCT公布号W005/047327中,其通过引用并入本文。

[0484] 用于本发明的Fc区或FcRn结合配偶体还可包含本领域认可的改变嵌合蛋白的糖基化的氨基酸替代。例如,连接于VWF片段或FVIII蛋白的嵌合蛋白的Fc区或FcRn结合配偶体可以包含具有导致糖基化(例如,N-或O-连接的糖基化)减少的突变的Fc区或可以包含野生型Fc部分的改变的糖型(例如,低岩藻糖或无岩藻糖的聚糖)。

[0485] 在一个实施方案中,本发明的未加工的嵌合蛋白可以包含具有两个或更多个其组成性Ig恒定区或其一部分(独立地选自本文所述的Ig恒定区或其一部分)的基因融合的Fc区(即,scFc区域)。在一个实施方案中,二聚Fc区的Fc区是相同的。在另一个实施方案中,Fc区中的至少两个是不同的。例如,本发明蛋白质的Fc区或FcRn结合配偶体包含相同数量的氨基酸残基或它们在长度上可相差一个或多个氨基酸残基(例如,约5个氨基酸残基(例如,1、2、3、4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)。在另外其它实施方案中,本发明蛋白质的Fc区或FcRn结合配偶体的序列可在一个或多个氨基酸位置不同。例如,Fc区或FcRn结合配偶体中的至少两个在约5个氨基酸位置(例如,1、2、3、4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)不同。

[0486] E) 接头

[0487] 本发明的嵌合蛋白还包含一个或多个接头。一种类型的接头为可切割的接头,其当向受试者体内施用时可被各种蛋白酶例如在凝固位点切割。在一个实施方案中,可切割的接头允许部分例如VWF片段从嵌合蛋白在凝血级联位点切割,因而允许活化FVIII(FVIIIa)以具有其FVIIIa活性。另一类型的接头为可加工的接头,其含有胞内切割位点并且因而可通过宿主细胞中的胞内加工酶被切割,允许方便表达多肽以及形成嵌合蛋白。

[0488] 一个或多个接头可存在于嵌合蛋白中的任意两个蛋白之间。在一个实施方案中,嵌合蛋白包含(i) VWF片段, (ii) XTEN序列和 (iii) FVIII蛋白,其中VWF片段通过接头例如可切割的接头连接于XTEN序列,并且XTEN序列进一步连接于FVIII蛋白(即,V-L-X-FVIII)。在另一个实施方案中,嵌合蛋白包含(i) VWF片段, (ii) XTEN序列,和 (iii) FVIII蛋白,其中VWF片段连接于XTEN序列,并且XTEN序列通过接头例如可切割的接头(即,V-X-L-FVIII)连接于FVIII蛋白。

[0489] 在某些实施方案中,嵌合蛋白包含(i) VWF片段, (ii) XTEN序列, (iii) 第一Ig恒定区或其一部分(例如第一Fc区), (iv) FVIII蛋白,和 (v) 第二Ig恒定区或其一部分(例如,第二Fc区),其中VWF片段通过任选的接头例如可切割的接头连接于XTEN序列。XTEN序列可通过接头例如可切割的接头进一步连接于第一Ig恒定区或其一部分。FVIII蛋白(含有或不含XTEN序列)也可通过任选的接头例如可切割的接头连接于第二Ig恒定区或其一部分。在某些实施方案中,嵌合蛋白还包含在第一Ig恒定区或其一部分(例如第一Fc区)和第二Ig恒定区或其一部分(例如,第二Fc区)之间,在VWF片段和第二Ig恒定区或其一部分之间、或在FVIII蛋白和第一Ig恒定区或其一部分(例如第一Fc区)之间的一个或多个接头例如可加工的接头。

[0490] 在一些实施方案中,本发明包括嵌合蛋白,其包含(i) FVIII蛋白, (ii) XTEN序列, (iii) 第一Ig恒定区或其一部分,和 (iv) 第二Ig恒定区或其一部分,其中第一Ig恒定区或其一部分和第二Ig恒定区或其一部分通过可加工的接头连接。

[0491] 可用于本发明的接头可包括任何有机分子。在一个实施方案中,接头包括聚合物,例如,聚乙二醇(PEG)或羟乙基淀粉(HES)。在另一个实施方案中,接头包含氨基酸序列。接头可包含至少约10、20、30、40、50、60、70、80、90、100、150、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900或2000个氨基酸。接头可包含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。在一个实施方案中,接头包含XTEN序列。XTEN的另外实例可根据本发明使用并且公开于美国专利公布号2010/0239554A1、2010/0323956A1、2011/0046060A1、2011/0046061A1、2011/0077199A1或2011/0172146A1或国际专利公布号WO 2010091122A1、WO 2010144502A2、WO 2010144508A1、WO 2011028228A1、WO 2011028229A1或WO 2011028344A2中。在另一个实施方案中,接头为PAS序列。

[0492] 可用于本发明的接头可包括任何有机分子。在一个实施方案中,接头为聚合物,例如,聚乙二醇(PEG)或羟乙基淀粉(HES)。在另一个实施方案中,接头为氨基酸序列。接头可

包含至少约10、20、30、40、50、60、70、80、90、100、150、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900或2000个氨基酸。接头可包含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。

[0493] 接头的实例为本领域熟知的。在一个实施方案中,接头包含序列 G_n 。接头可包含序列 $(GA)_n$ 。接头可包含序列 $(GGS)_n$ 。在其它实施方案中,接头包含 $(G\ GGS)_n$ (SEQ ID NO:57)。在又其它实施方案中,接头包含序列 $(GGS)_n$ (GGG GS) $_n$ (SEQ ID NO:58)。在这些情况下,n可以为1-100的整数。在其它情况下,n可以为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。接头的实例包括但不限于,GGG,SGGS GGS (SEQ ID NO:59)、GGSGGSGGSGGSGGG (SEQ ID NO:60)、GGS GGSGGGSGGGGS (SEQ ID NO:61)、GGSGGSGGSGGSGGSGGS (SEQ ID NO:62)或GGGGSGGGSGGGGS (SEQ ID NO:63)。接头不消除或减少因子VIII的VWF片段活性或凝血活性。任选地,接头增强因子VIII蛋白的VWF片段活性或凝血活性,例如通过进一步减少空间位阻效应以及使VWF片段或因子VIII蛋白更易接近其靶结合位点。

[0494] 在一个实施方案中,可用于嵌合蛋白的接头为15-25个氨基酸长。在另一个实施方案中,可用于嵌合蛋白的接头为15-20个氨基酸长。在一些实施方案中,嵌合蛋白的接头为10-25个氨基酸长。在其它实施方案中,嵌合蛋白的接头为15个氨基酸长。在又其它实施方案中,嵌合蛋白的接头为 $(GGGS)_n$ (SEQ ID NO:64),其中G表示甘氨酸,S表示丝氨酸且n为1-20的整数。

[0495] F) 切割位点

[0496] 接头还可并入能够以化学方法(例如,水解酯键)、用酶法(即,并入蛋白酶切割序列)或通过光解(例如,生色团诸如3-氨基-3-(2-硝基苯基)丙酸(ANP))被切割的部分以便从另一个分子释放一个分子。

[0497] 在一个实施方案中,接头为可切割的接头。可切割的接头可包含在N-末端或C-末端或两者的一个或多个切割位点。在另一个实施方案中,可切割的接头基本上由以下组成或由以下组成:一个或多个可切割的位点。在其它实施方案中,可切割的接头包含本文所述的异源氨基酸接头序列或聚合物和一个或多个可切割的位点。

[0498] 在某些实施方案中,可切割的接头包含可在宿主细胞中被切割的一个或多个切割位点(即,胞内加工位点)。切割位点的非限制性实例包括RRRR (SEQ ID NO:9)、RKRRKR (SEQ ID NO:10)和RRRRS (SEQ ID NO:11)。

[0499] 在其它实施方案中,可切割的接头包含在向受试者施用包含可切割的接头的嵌合蛋白后被蛋白酶切割的一个或多个切割位点。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(例如,KLTR↓AET (SEQ ID NO:65))、FXIa切割位点(例如,DFTR↓VVG (SEQ ID NO:66))、FXIIa切割位点(例如,TMTR↓IVGG (SEQ ID NO:67))、激肽释放酶切割位点(例如,SPFR↓STGG (SEQ ID NO:68))、FVIIa切割位点(例如,LQVR↓IVGG (SEQ ID NO:69))、FIXa切割位点(例如,PLGR↓IVGG (SEQ ID NO:70))、FXa切割位点(例如,

IEGR↓TVGG (SEQ ID NO:71))、FIIa (凝血酶) 切割位点 (例如, LTPR↓SLLV (SEQ ID NO:72))、弹性蛋白酶-2 切割位点 (例如, LGPV↓SGVP (SEQ ID NO:73))、颗粒酶-B 切割 (例如, VAGD↓SLEE (SEQ ID NO:74))、MMP-12 切割位点 (例如, GPAG↓LGGA (SEQ ID NO:75))、MMP-13 切割位点 (例如, GPAG↓LRGA (SEQ ID NO:76))、MMP-17 切割位点 (例如, APLG↓LRLR (SEQ ID NO:77))、MMP-20 切割位点 (例如, PALP↓LVAQ (SEQ ID NO:78))、TEV 切割位点 (例如, ENLYFQ↓G (SEQ ID NO:79))、肠激酶切割位点 (例如, DDDK↓IVGG (SEQ ID NO:80))、蛋白酶3C (PRESCISSINTM) 切割位点 (例如, LEVLFQ↓GP (SEQ ID NO:81)) 和分选酶A 切割位点 (例如, LPKT↓GSES) (SEQ ID NO:82)。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:83) 和SVSQTSKLTR (SEQ ID NO:84)。非限制性示例性凝血酶切割位点包括, 例如DFLAEGGGVR (SEQ ID NO:85)、TTKIKPR (SEQ ID NO:86) 或LVPRG (SEQ ID NO:87), 和包含以下、基本上由以下组成或由以下组成的序列: ALRPR (SEQ ID NO:17) (例如, ALRPRVVGGA (SEQ ID NO:88))。

[0500] 在具体实施方案中, 切割位点为TLDPRSFLLRNPNDKYEPFWEDEEK (SEQ ID NO:8)。

[0501] 多核苷酸、载体和宿主细胞

[0502] 本发明也提供编码以下的多核苷酸: (a) 连接于XTEN序列和FVIII蛋白的VWF片段, (b) 连接于XTEN序列和Fc的FVIII蛋白, 或 (c) 连接于本文所述的XTEN序列和VWF片段的FVIII蛋白。当嵌合蛋白为单一多肽链 (例如, F2-L2-X-V-L1-F1-FVIII, 其中FVIII包括FVIII蛋白, F1包括第一Ig恒定区或其一部分, 例如第一Fc区, L1包括第一接头, V包括VWF片段, X包括XTEN序列, L2包括第二接头, 并且F2包括第二Ig恒定区或其一部分, 例如第二Fc区) 时, 本发明涉及编码单一多肽链的单个多核苷酸链。当嵌合蛋白包含第一和第二多肽链 (F2-L2-X-V:FVIII-F1) 时, 第一多肽链包含连接于XTEN序列的VWF片段, XTEN序列通过可切割的接头 (例如, F2-L2-X-V) 进一步连接于第一Ig恒定区或其一部分 (例如, 第一Fc区), 并且第二多肽链包含FVIII蛋白和第二Ig恒定区或其一部分 (例如, 第二Fc区) (例如, FVIII-F1), 其中第一多肽链和第二多肽链彼此缔合, 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, 第一多肽链和第二多肽链可由单个多核苷酸链编码。在另一个实施方案中, 第一多肽链和第二多肽链通过两个不同的多核苷酸即第一核苷酸序列和第二核苷酸序列编码。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 (例如, 不同载体) 上。在某些实施方案中, 本发明涉及包含第一核苷酸链和第二核苷酸链的多核苷酸组, 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。在一些实施方案中, 包含两条多肽链或三条多肽链的嵌合蛋白可通过单个多核苷酸链编码, 然后被加工成两条或三条 (或更多条) 多肽链。在另外其它实施方案中, 包含这些多肽链的嵌合蛋白可由两条或三条多核苷酸链编码。

[0503] 在其它实施方案中, 多核苷酸组还包含编码蛋白转化酶的另外的核苷酸链 (例如, 当嵌合多肽由单条多核苷酸链编码时, 为第二核苷酸链或当嵌合蛋白由两条多核苷酸链编码时, 为第三核苷酸链)。蛋白转化酶可以选自由以下组成的组: 蛋白原转化酶枯草杆菌蛋白酶/kexin5型 (PCSK5或PC5)、蛋白原转化酶枯草杆菌蛋白酶/kexin7型 (PCSK7或PC5)、酵母Kex 2、蛋白原转化酶枯草杆菌蛋白酶/kexin3型 (PACE或PCSK3) 及其两个或更多的组合。在一些实施方案中, 蛋白转化酶为PACE、PC5或PC7。在具体实施方案中, 蛋白转化酶为PC5或PC7。参见国际申请号PCT/US2011/043568。

[0504] 如本文所使用,表达载体是指任何核酸构建体,其含有用于插入的编码序列的转录和翻译的必要元件,或在RNA病毒载体的情况下含有当引入适当的宿主细胞时用于复制和翻译的必要元件。表达载体可包括质粒、噬菌粒、病毒及其衍生物。

[0505] 本发明的表达载体将包括编码本文所述的嵌合蛋白的多核苷酸。在一个实施方案中,VWF片段和XTEN、FVIII蛋白和XTEN或两者的一个或多个编码序列可操作地连接于表达控制序列。如本文所使用,当两个核酸序列以允许每种组分核酸序列保留其功能性的方式共价连接时,则它们可操作地连接。当编码序列和基因表达控制序列以在影响或控制基因表达控制序列下进行编码序列的表达或转录和/或翻译的方式共价连接时,则它们据说可操作地连接。如果对5'基因表达序列中启动子的诱导导致编码序列的转录以及如果两个DNA序列之间的连接的性质不(1)导致移码突变的引入,(2)干扰启动子区域指导编码序列的转录的能力,或(3)干扰相应的RNA转录物被翻译成蛋白质的能力,则两个DNA序列据说可操作地连接。因此,如果基因表达序列能够影响编码核酸序列的转录以使所得转录物被翻译成所需蛋白质或多肽,则基因表达序列将可操作地连接于编码核酸序列。

[0506] 如本文所使用的基因表达控制序列为任何调控核苷酸序列,诸如启动子序列或启动子-增强组合,其促进与其可操作地连接的编码核酸的有效转录和翻译。基因表达控制序列可例如为哺乳动物或病毒启动子,诸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,下列基因的启动子:次黄嘌呤磷酸核糖基转移酶(HPRT)、腺苷脱氨酶、丙酮酸激酶、 β -肌动蛋白启动子和其它组成型启动子。在真核细胞中组成型发挥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(例如,SV40)、乳头状瘤病毒、腺病毒、人免疫缺陷病毒(HIV)、劳斯肉瘤病毒、巨细胞病毒、莫洛尼白血病病毒的长末端重复序列(LTR)和其它逆转录病毒的启动子,以及单纯性疱疹病毒的胸苷激酶启动子。其它组成型启动子为本领域普通技术人员已知的。可用作本发明的基因表达序列的启动子也包括诱导型启动子。在诱导剂存在下表达诱导型启动子。例如,在某些金属离子存在下诱导金属硫蛋白启动子以促进转录和翻译。其它诱导型启动子为本领域普通技术人员已知的。

[0507] 通常,基因表达控制序列应包括,如有必要,分别牵涉转录和翻译的启动的5'非转录和5'非翻译序列,诸如TATA盒、加帽序列、CAAT序列等。尤其,此类5'非转录序列将包括启动子区,其包括用于转录控制可操作地联接的编码核酸的启动子序列。基因表达序列按需任选地包括增强子序列或上游活化剂序列。

[0508] 病毒载体包括但不限于,来自下列病毒的核酸序列:逆转录病毒,诸如莫洛尼鼠白血病病毒、哈维鼠肉瘤病毒、鼠乳房肿瘤病毒和劳斯肉瘤病毒;腺病毒、腺相关病毒;SV40型病毒;多瘤病毒;埃-巴二氏病毒(Epstein-Barr virus);乳头状瘤病毒;疱疹病毒;牛痘病毒;脊髓灰质炎病毒;和RNA病毒诸如逆转录病毒。可容易地采用本领域熟知的其它载体。某些病毒载体是基于非细胞病变真核细胞病毒,其中非必需基因已被替换为目标基因。非细胞病变病毒包括逆转录病毒,其生命周期牵涉将基因组病毒RNA逆转转录成DNA,随后将原病毒整合到宿主细胞DNA中。逆转录病毒已被批准用于人基因治疗试验。大多数有用的是为复制缺陷型的那些逆转录病毒(即,能够指导所需蛋白质的合成但不能制造传染性微粒)。此类经基因改变的逆转录病毒表达载体具有用于体内高效率转导基因的一般效用。用于产生复制缺陷型逆转录病毒的标准方案(包括以下步骤:将外源性的遗传物质并入质粒中,用质粒转染包装细胞系,通过包装细胞系生成重组逆转录病毒,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., Clifton, N.J. (1991) 中提供。

[0509] 在一个实施方案中,病毒为腺相关病毒、双链DNA病毒。腺相关病毒可被工程改造为复制缺陷型并且能够感染范围广泛的细胞类型和种类。其进一步具有优势,诸如热稳定性和脂质溶剂稳定性;不同谱系细胞(包括造血细胞)中的高转导频率;以及缺少重复感染抑制,因而允许多个系列的转导。据报道,腺相关病毒可以位点特异性方式被整合到人细胞DNA,从而使插入诱变的可能性和逆转录病毒感染的插入基因表达特征的变化性最小化。而且,在不存在选择压力下,跟踪组织培养物中的野生型腺相关病毒感染大于100代,暗示腺相关病毒基因组整合为相对稳定事件。腺相关病毒也可以染色体外的方式发挥作用。

[0510] 其它载体包括质粒载体。质粒载体已在本领域中广泛描述并且为本领域技术人员熟知。参见,例如Sambrook等人, Molecular Cloning: A Laboratory Manual, 第二版, Cold Spring Harbor Laboratory Press, 1989。最近几年,已发现质粒载体对于体内递送基因至细胞而言是特别有利的,因为它们不能在宿主基因组内复制和整合到宿主基因组中。但是,具有与宿主细胞相容的启动子的这些质粒可从在质粒内可操作地编码的基因表达肽。得自商业供应商的一些常用质粒包括pBR322、pUC18、pUC19、各种pcDNA质粒、pRC/CMV、各种pCMV质粒、pSV40和pBlueScript。特定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; 和pBudCE4.1, 目录号V53220, 均来自Invitrogen (Carlsbad, CA.)。其它质粒为本领域普通技术人员熟知的。此外,可使用标准分子生物学技术定制设计质粒以去除和/或添加DNA的特定片段。

[0511] 在可用于生成本发明蛋白质的一种昆虫表达系统中,苜蓿银纹夜蛾(*Autographa californica*)核型多角体病毒(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。编码序列可被克隆到病毒的非必需区域(例如,多角体蛋白基因)中并且置于ACNPV启动子(例如,多角体蛋白启动子)的控制。成功地插入编码序列将导致多角体蛋白基因的活化和非闭塞性重组病毒(即,缺少由多角体蛋白基因编码的蛋白质包被的病毒)的生成。这些重组病毒然后用于感染插入基因于其中表达的草地贪夜蛾细胞(参见,例如Smith等人(1983) *J Virol* 46:584;美国专利号4,215,051)。这种表达系统的另外实例可见于Ausubel等人编(1989) *Current Protocols in Molecular Biology*, 第2卷, Greene Publish. Assoc. & Wiley Interscience中。

[0512] 可用于表达本发明蛋白质的另一系统为谷氨酰胺合成酶基因表达系统,也被称为“GS表达系统”(Lonza Biologics PLC, Berkshire UK)。这种表达系统详细地描述于美国专利号5,981,216中。

[0513] 在哺乳动物宿主细胞中,可利用许多基于病毒的表达系统。在腺病毒用作表达载体的情况下,编码序列可连接于腺病毒转录/翻译控制复合物,例如后期启动子和三重前导序列。这种嵌合基因然后可通过体外或体内重组被插入到腺病毒基因组。插入到病毒基因组的非必需区域(例如,区域E1或E3)将产生能存活的并且能够表达感染宿主中的肽的重组病毒。参见,例如Logan & Shenk (1984) *Proc Natl Acad Sci USA* 81:3655)。可选地,可使用牛痘7.5K启动子。参见,例如Mackett等人(1982) *Proc Natl Acad Sci USA* 79:7415; Mackett等人(1984) *J Virol* 49:857; Panicali等人(1982) *Proc Natl Acad Sci USA* 79:

4927。

[0514] 为了增加生产效率,可将多核苷酸设计成编码本发明蛋白质的通过酶切割位点隔开的多个单元。可(例如,通过用适当的酶处理)切割所得多肽以便回收多肽单元。这可增加通过单个启动子驱动的多肽的产率。当于适当的病毒表达系统中使用时,在转录物中内部指导由mRNA编码的每种多肽的翻译;例如通过内部核糖体进入位点IRES。因此,多顺反子性构建体指导单个、大的多顺反子性mRNA的转录,反过来所述多顺反子性mRNA指导多个、个别多肽的翻译。这种方法消除多蛋白的生成和酶促加工并且可显著地增加通过单个启动子驱动的多肽的产率。

[0515] 转化中使用的载体将通常含有用于鉴定转化体的可选择标记。在细菌系统中,这可包括抗生素抗性基因诸如氨苄青霉素或卡那霉素。培养的哺乳动物细胞中使用的可选择标记包括赋予对药物诸如新霉素、潮霉素和甲氨蝶呤的抗性的基因。可选择标记可以为可扩增的可选择标记。一种可扩增的可选择标记为二氢叶酸还原酶(DHFR)基因。Simonsen C C等人(1983)Proc Natl Acad Sci USA 80:2495-9。可选择标记由Thilly(1986)Mammalian Cell Technology, Butterworth Publishers, Stoneham, Mass. 综述,并且对可选择标记的选择完全在本领域普通技术人员的水平之内。

[0516] 可选择标记与目标基因可在单独质粒上同时被引入细胞,或它们可被引入到同一质粒上。如果在同一质粒上,则可选择标记和目标基因可处于不同启动子或同一启动子的控制之下,后者排列产生双顺反子信息。这种类型的构建体是本领域已知的(例如,美国专利号4,713,339)。

[0517] 表达载体可编码允许容易纯化重组生成的蛋白质的标记。实例包括但不限于,载体pUR278(Ruther等人(1983)EMBO J 2:1791),其中可将待表达的蛋白质的编码序列连接到框内具有lac z编码区的载体以使生成标记的融合蛋白;pGEX载体可用于表达具有谷胱甘肽S-转移酶(GST)标记的本发明蛋白质。这些蛋白质通常为可溶的并且可通过吸附至谷胱甘肽-琼脂糖珠粒、然后在游离谷胱甘肽存在下洗脱而容易地从细胞纯化。载体包括用于在纯化后容易去除标记的切割位点(凝血酶或因子Xa蛋白酶或PRESCISSION PROTEASE™(Pharmacia, Peapack, N.J.))。

[0518] 表达载体然后被转染或共转染到适合的靶细胞中,其将表达多肽。本领域中已知的转染技术包括但不限于,磷酸钙沉淀(Wigler等人(1978)Cell 14:725)、电穿孔(Neumann等人(1982)EMBO J 1:841)和基于脂质体的试剂。可利用多种宿主-表达载体系统以表达本文所述的蛋白质,包括原核细胞和真核细胞两者。这些包括但不限于,微生物诸如用含有适当的编码序列的重组噬菌体DNA或质粒DNA表达载体转化的的细菌(例如,大肠杆菌);用含有适当的编码序列的重组酵母或真菌表达载体转化的酵母或丝状真菌;用含有适当的编码序列的重组病毒表达载体(例如,杆状病毒)感染的昆虫细胞系统;用重组病毒表达载体(例如,花椰菜花叶病毒或烟草花叶病毒)感染或用含有适当的编码序列的重组质粒表达载体(例如,Ti质粒)转化的植物细胞系统;或动物细胞系统,包括哺乳动物细胞(例如,HEK 293、CHO、Cos、HeLa、HKB11和BHK细胞)。

[0519] 在一个实施方案中,宿主细胞为真核细胞。如本文所使用,真核细胞是指具有定形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例如哺乳动物的细胞,以及无脊椎动物例如昆虫的细胞。植物的真核细胞具体地可包括但不限于,酵母细胞。真核细胞

不同于原核细胞,例如细菌。

[0520] 在某些实施方案中,真核细胞为哺乳动物细胞。哺乳动物细胞为源自哺乳动物的任何细胞。哺乳动物细胞具体地包括但不限于,哺乳动物细胞系。在一个实施方案中,哺乳动物细胞为人细胞。在另一个实施方案中,哺乳动物细胞为HEK 293细胞,其为人胚胎肾细胞系。HEK 293细胞从美国菌种保藏中心(American Type Culture Collection), Manassas, VA以CRL-1533,和从Invitrogen (Carlsbad, Calif.) 以293-H细胞(目录号11631-017)或293-F细胞(目录号11625-019)可得。在一些实施方案中,哺乳动物细胞为**PER.C6[®]**细胞,其为源自视网膜的人细胞系。**PER.C6[®]**细胞得自Crucell (Leiden, The Netherlands)。在其它实施方案中,哺乳动物细胞为中国仓鼠卵巢(CHO)细胞。CHO细胞得自美国菌种保藏中心, Manassas, VA. (例如, CHO-K1; CCL-61)。在又其它实施方案中,哺乳动物细胞为幼仓鼠肾(BHK)细胞。BHK细胞得自美国菌种保藏中心, Manassas, Va. (例如, CRL-1632)。在一些实施方案中,哺乳动物细胞为HKB11细胞,其为HEK293细胞和人B细胞系的杂种细胞系。Mei等人, Mol. Biotechnol. 34 (2): 165-78 (2006)。

[0521] 在一个实施方案中,将包括FVIII(X)-Fc融合编码序列、VWF片段-L-Fc融合编码序列、或两者和可选择标记(例如博莱霉素抗性)的质粒转染到HEK293细胞中用于生成嵌合蛋白。

[0522] 在另一个实施方案中,将包括FVIII-Fc融合编码序列、VWF片段-XTEN-L-Fc融合编码序列、或两者和可选择标记(例如博莱霉素抗性)的质粒转染到HEK 293细胞中用于生成嵌合蛋白。

[0523] 在其它实施方案中,将包括FVIII(X)-Fc融合编码序列、Fc编码序列、或两者和可选择标记(例如博莱霉素抗性)的质粒转染到HEK 293细胞中用于生成嵌合蛋白。

[0524] 在一些实施方案中,将包括FVIII(X)-Fc融合编码序列和第一可选择标记(例如博莱霉素抗性基因)的第一质粒、和包括Fc编码序列或VWF片段-L-Fc编码序列和第二可选择标记(例如新霉素抗性基因)的第二质粒、以及包括蛋白转化酶编码序列和第三可选择标记(例如潮霉素抗性基因)的第三质粒共转染到HEK 293细胞中用于生成嵌合蛋白。第一质粒和第二质粒可以等量(即, 1:1摩尔比率)引入,或它们可以不等量引入。

[0525] 在又其它实施方案中,将包括FVIII-Fc融合编码序列和第一可选择标记(例如博莱霉素抗性基因)的第一质粒、和包括VWF片段-XTEN-L-Fc编码序列和第二可选择标记(例如新霉素抗性基因)的第二质粒、以及包括蛋白转化酶编码序列和第三可选择标记(例如潮霉素抗性基因)的第三质粒共转染到HEK 293细胞中用于生成嵌合蛋白。第一质粒和第二质粒可以等量(即, 1:1摩尔比率)引入,或它们可以不等量引入。

[0526] 在另外其它实施方案中,将包括FVIII(X)-Fc融合编码序列和第一可选择标记(例如博莱霉素抗性基因)的第一质粒、和包括VWF片段-XTEN-L-Fc编码序列和第二可选择标记(例如新霉素抗性基因)的第二质粒、以及包括蛋白转化酶编码序列和第三可选择标记(例如潮霉素抗性基因)的第三质粒共转染到HEK 293细胞中用于生成嵌合蛋白。第一质粒和第二质粒可以等量(即, 1:1摩尔比率)引入,或它们可以不等量引入。

[0527] 在某些实施方案中,将包括嵌合蛋白编码FVIII(含有或不含XTEN)-F1-L1-V-XTEN-L2-F2编码序列和第一可选择标记(例如博莱霉素抗性基因)的第一质粒、和包括蛋白

转化酶编码序列和第二可选择标记(例如潮霉素抗性基因)的第二质粒共转染到HEK 293细胞中用于生成嵌合蛋白。FVIII(X)-Fc编码序列和VWF-XTEN-Fc编码序列的启动子可以不同或它们可以是相同的。

[0528] 在又其它实施方案中,转染细胞为稳定转染的。可选择这些细胞并使用本领域技术人员已知的常规技术使其保持为稳定细胞系。

[0529] 使含有蛋白质的DNA构建体的宿主细胞生长于适当的生长培养基中。如本文所使用,术语“适当的生长培养基”意指含有细胞生长所需的营养素的培养基。细胞生长所需的营养素可以包括碳源、氮源、必需氨基酸、维生素、矿物和生长因子。任选地,培养基可含有一个或多个选择因子。任选地,任选地培养基可含有小牛血清或胎牛血清(FCS)。在一个实施方案中,培养基基本上不含IgG。通过例如用DNA构建体上的可选择标记补充的或用DNA构建体共转染的必需营养素中的药物选择或缺陷,生长培养基通常选择含有DNA构建体的细胞。培养的哺乳动物细胞一般生长于商购可得的含血清或无血清培养基(例如, MEM、DMEM、DMEM/F12)中。在一个实施方案中,培养基为CD293(Invitrogen, Carlsbad, CA.)。在另一个实施方案中,培养基为CD17(Invitrogen, Carlsbad, CA.)。对适合用于所用的特定细胞系的培养基的选择在本领域普通技术人员的水平之内。

[0530] 为了共表达嵌合蛋白的两条多肽链,在允许表达两条链的条件下培养宿主细胞。如本文所使用,培养是指体外保持活细胞至少一段确定的时间。保持可以但不一定包括活细胞群体的增加。例如,保持于培养物中的细胞可以在群体中为静态的,但是仍能存活的并且能够生成所需产物,例如重组蛋白或重组融合蛋白。适合用于培养真核细胞的条件为本领域熟知的并且包括对培养基、培养基补充剂、温度、pH、氧饱和度等的适当选择。出于商业用途,培养可包括使用各种类型的放大系统中的任一种,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、波浪生物反应器等。

[0531] 也选择细胞培养条件以允许VWF片段与FVIII蛋白的缔合。允许表达VWF片段和/或FVIII蛋白的条件可以包括维生素K的来源的存在。例如,在一个实施方案中,稳定转染的HEK 293细胞于补充有4mM谷氨酰胺的CD293培养基(Invitrogen, Carlsbad, CA)或OptiCHO培养基(Invitrogen, Carlsbad, CA)中培养。

[0532] 一方面,本发明涉及表达、制备或生产本发明的嵌合蛋白的方法,所述方法包括a)转染包含编码嵌合蛋白的多核苷酸的宿主细胞和b)在适用于表达嵌合蛋白的条件下在培养基中培养宿主细胞,于其中表达嵌合蛋白。

[0533] 在另外的实施方案中,将含有连接于XTEN序列的VWF片段或连接于XTEN序列的FVIII蛋白的蛋白质产物分泌到培养基中。从细胞分离培养基,浓缩,过滤,然后通过两个或三个亲和柱,例如蛋白质A柱和一个或两个阴离子交换柱。

[0534] 在某些方面,本发明涉及通过本文所述的方法产生的嵌合蛋白。

[0535] 体外产生允许放大以得到大量所需的本发明的改变多肽。用于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且包括均质悬浮液培养(例如于气升式反应器或连续搅拌反应器中)或固定化或捕获式细胞培养(例如于中空纤维、微囊中,在琼脂糖微珠或陶瓷滤芯上)。如有必要和/或需要,多肽溶液可通过常规色谱方法(例如凝胶过滤、离子交换色谱、疏水相互作用色谱(HIC,在DEAE-纤维素上的色谱或亲和力色谱)纯化。

[0536] 药物组合物

[0537] 含有本发明的嵌合蛋白的组合物可含有适合的药学上可接受的载体。例如，它们可含有促进活性化合物加工成被设计成用于递送至作用位点的制剂的赋形剂和/或助剂。

[0538] 药物组合物可被配制用于通过弹丸注射进行胃肠外施用(即静脉内、皮下或肌内)。用于注射的制剂可以单位剂型存在，例如于安瓿或添加有防腐剂的多剂量容器中。组合物可采用诸如于油性或水性媒介物中的悬浮液、溶液或乳液的形式，并且含有诸如助悬剂、稳定剂和/或分散剂的配制剂。可选地，活性成分可以呈粉末形式用于用适合的媒介物例如无热源水复原。

[0539] 适合用于胃肠外施用的制剂还包括呈水溶性形式，例如水溶性盐的活性化合物的水溶液。此外，可以施用呈适当的油性注射悬浮液形式的活性化合物的悬浮液。适合的亲脂性溶剂或媒介物包括脂肪油，例如芝麻油或合成脂肪酸酯，例如，油酸乙酯或三甘油酯。水性注射悬浮液可含有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，悬浮液还可含有稳定剂。脂质体也可用于封装本发明的分子以用于递送至细胞或间质空间。示例性的药学上可接受的载体为生理上相容的溶剂、分散介质、包衣剂、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂、水、盐水、磷酸盐缓冲盐水、葡萄糖、甘油、乙醇等。在一些实施方案中，组合物包含等渗剂，例如糖、多元醇诸如甘露醇、山梨糖醇或氯化钠。在其它实施方案中，组合物包含药学上可接受的物质诸如润湿剂或少量辅助物质诸如润湿剂或乳化剂、防腐剂或缓冲液，其增强活性成分的货架期或有效性。

[0540] 本发明的组合物可以呈多种形式，包括例如液体(例如，可注射和可输注的溶液)、分散液、悬浮液、半固体和固体剂型。优选的形式取决于施用模式和治疗应用。

[0541] 组合物可被配制为溶液、微乳、分散液、脂质体或适于高药物浓度的其它有序结构。无菌注射溶液可通过将需要量的活性成分按需与一种上文提到的成分或成分组合并入适当的溶剂中，然后过滤灭菌来制备。一般而言，分散液通过将活性成分并入含有碱性分散介质和来自上文提到的那些的所需其它成分的无菌媒介物中来制备。在用于制备无菌注射溶液的无菌粉末的情况下，优选的制备方法为真空干燥和冻干，得到活性成分加上来自先前经过无菌过滤的溶液的任何另外的所需成分的粉末。例如通过使用诸如卵磷脂的包衣剂，通过在分散液的情况下维持所需颗粒大小和通过使用表面活性剂，可以维持溶液的适当流动性。可注射组合物的延长吸收可通过将延迟吸收的试剂例如单硬脂酸盐和明胶纳入组合物中而产生。

[0542] 活性成分可用控制释放的制剂或装置进行配制。此类制剂和装置的实例包括植入物、经皮贴剂和微囊化递送系统。可以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、聚原酯和聚乳酸。用于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, Marcel Dekker, Inc., New York, 1978。

[0543] 可注射贮库制剂可通过于生物可降解的聚合物诸如聚丙交酯-聚乙交酯中形成药物的微囊化基质来制备。根据药物与聚合物的比率，以及所用聚合物的性质，可控制药物释放的速率。其它示例性的生物可降解的聚合物为聚原酯和聚酸酐。贮库可注射制剂还可以通过于脂质体或微乳中捕获药物来制备。

[0544] 补充性活性化合物可被并入组合物中。在一个实施方案中，本发明的嵌合蛋白与另一凝血因子或其变体、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、纤维蛋白原、血管性血友病因子或重组可溶性组织因子(rsTF)或任何前述因子的活化形式。止血剂的凝血因子也可包括抗纤维蛋白溶解药物,例如 ϵ -氨基-己酸、氨甲环酸。

[0545] 可以调节给药方案以提供最佳所需响应。例如,可施用单次推注,可随着时间推移施用数个分剂量,或可按比例减少或增加剂量,如通过治疗情况的紧急状态指示。以剂量单位形式配制胃肠外组合物对于易施用性和剂量均匀性而言是有利的。参见,例如Remington's Pharmaceutical Sciences (Mack Pub.Co., Easton, Pa. 1980)。

[0546] 除了活性化合物之外,液体剂型可含有惰性成分诸如水、乙醇、碳酸乙酯、乙酸乙酯、苯甲醇、苯甲酸苄酯、丙二醇、1,3-丁二醇、二甲基甲酰胺、油类、甘油、四氢糠醇、聚乙二醇和脱水山梨糖醇脂肪酸酯。

[0547] 适合的药用载体的非限制性实例也描述于E.W.Martin的Remington's Pharmaceutical Sciences中。赋形剂的一些实例包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、稻米、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石、氯化钠、脱脂乳粉、甘油、丙烯、二醇、水、乙醇等。组合物也可含有pH缓冲剂和润湿剂或乳化剂。

[0548] 对于口服施用,药物组合物可采用通过常规方法制备的片剂或胶囊的形式。组合物也可以液体例如糖浆或悬浮液形式制备。液体可包括助悬剂(例如,山梨糖醇糖浆、纤维素衍生物或氢化食用脂肪)、乳化剂(卵磷脂或阿拉伯胶)、非水性媒介物(例如,杏仁油、油酯、乙醇或分级分离的蔬菜油)和防腐剂(例如,对羟基苯甲酸甲酯或对羟基苯甲酸丙酯或山梨酸)。制剂也可包含调味剂、着色剂和甜味剂。可选地,组合物可以呈现为用水或另一种适合的媒介物复原的干燥产品。

[0549] 对于经颊施用,组合物可根据常规方案采用片剂或锭剂的形式。

[0550] 对于通过吸入施用,根据本发明使用的化合物以含有或不含赋形剂的雾化气溶胶形式或呈来自加压包或喷雾器、任选地含有推进剂(例如二氯二氟甲烷、三氯氟甲烷、二氯四氟甲烷、二氧化碳或其它适合的气体)的气溶胶喷雾的形式方便递送。在加压气溶胶的情况下,剂量单位可通过提供阀以递送计量的量来确定。例如吸入器或吹入器中使用的明胶的胶囊及其盒,可被配制成含有化合物的粉末混合物和适合的粉末基质诸如乳糖或淀粉。

[0551] 药物组合物也可被配制成用于以例如含有常规栓剂基质诸如可可脂或其它甘油酯的栓剂或保留灌肠剂形式直肠施用。

[0552] 在一个实施方案中,药物组合物包含嵌合蛋白、编码嵌合蛋白的多核苷酸、包含多核苷酸的载体或包含载体的宿主细胞、和药学上可接受的载体。相较于野生型FVIII蛋白或相应的不含VWF片段的FVIII蛋白,嵌合蛋白中的FVIII蛋白具有延长的半衰期。在一个实施方案中,FVIII蛋白的半衰期比野生型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。在另一个实施方案中,因子VIII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。

[0553] 在一些实施方案中,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、胃肠外施用、鞘内施用、硬膜下施用和口服施用。胃肠外施用可以为静脉内或皮下施用。

[0554] 在其它实施方案中,组合物用于治疗需要其的受试者中的出血性疾病或病患。出血性疾病或病患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、胸腔内溢血、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。在又其它实施方案中,受试者被安排进行手术。在另外其它实施方案中,治疗为防治性的或按需治疗。

[0555] 基因治疗

[0556] 本发明的其嵌合蛋白可在治疗上有益的治疗选自由以下组成的组的出血性疾病或病症的基因治疗方法。哺乳动物例如人患者中体内生成: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、胸腔内溢血、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。在一个实施方案中,出血性疾病或病症为血友病。在另一个实施方案中,出血性疾病或病症为血友病A。这牵涉施用可操作地连接于适合的表达控制序列的适合的嵌合蛋白编码核酸。在某些实施方案中,这些序列并入病毒载体中。用于此类基因治疗的适合的病毒载体包括腺病毒载体、慢病毒载体、杆状病毒载体、埃-巴二氏病毒载体、乳多空病毒载体、牛痘病毒载体、单纯性疱疹病毒载体和腺相关病毒(AAV)载体。病毒载体可以为复制缺陷型病毒载体。在其它实施方案中,腺病毒载体在其E1基因或E3基因中具有缺失。当使用腺病毒载体时,哺乳动物可以不暴露于编码可选择标记基因的核酸。在其它实施方案中,序列并入本领域技术人员已知的非病毒载体中。

[0557] 使用嵌合蛋白的方法

[0558] 本发明涉及使用本文所述的嵌合蛋白阻止或抑制内源性VWF与FVIII蛋白结合的方法。本发明还涉及使用具有连接于XTEN和Ig恒定区或其一部分的FVIII蛋白的嵌合蛋白的方法。

[0559] 本发明的一方面涉及通过封闭或掩蔽FVIII上的VWF结合位点免遭内源性VWF以及同时使用与Ig恒定区或其一部分组合的XTEN序列(其也可以为半衰期延长剂)延长FVIII蛋白的半衰期来阻止或抑制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的FVIII蛋白的方法。在一个实施方案中,XTEN序列抑制或阻止嵌合蛋白中的FVIII蛋白与内源性VWF的相互作用。在另一个实施方案中,Ig恒定区或其一部分抑制或阻止FVIII蛋白与内源性VWF的相互作用。可用于该方法的嵌合蛋白包括任何一个或多个本文所述的嵌合蛋白。

[0560] 本发明的另一方面包括向需要其的受试者施用包含半衰期长于野生型FVIII的FVIII蛋白的嵌合蛋白的方法,其中所述方法包括向受试者施用本文所述的嵌合蛋白。

[0561] 在一个实施方案中,本发明涉及使用XTEN序列和Ig恒定区或其一部分延长FVIII蛋白和VWF片段的半衰期来阻止或抑制内源性VWF与FVIII蛋白相互作用。连接于XTEN序列(例如,FVIII(X))且然后结合于VWF片段或与VWF片段缔合的FVIII蛋白被掩蔽或保护免遭VWF的清除途径,因而相较于未结合于VWF片段的FVIII蛋白而具有减少的清除率。相较于未

结合于XTEN序列和VWF片段或与XTEN序列和VWF片段缔合的FVIII蛋白,经掩蔽的FVIII蛋白因而具有半衰期的最大延长。在某些实施方案中,与VWF片段缔合或受VWF片段保护的并且连接于XTEN序列的FVIII蛋白未被VWF清除受体清除。在其它实施方案中,与VWF片段缔合或受VWF片段保护的并且连接于XTEN序列的FVIII蛋白从系统被清除的速率慢于不与VWF片段缔合或未受VWF片段保护的并且连接于XTEN序列的FVIII蛋白。

[0562] 一方面,因为VWF片段不含VWF清除受体结合位点,包含连接于XTEN序列的FVIII蛋白或结合于连接于XTEN的VWF片段或与连接于XTEN的VWF片段缔合的FVIII蛋白的嵌合蛋白从循环的清除减少。VWF片段阻止或抑制结合于VWF片段或与VWF片段缔合的FVIII通过VWF清除途径从系统的清除。可用于本发明的VWF片段也可提供由内源性VWF提供的至少一种或多种VWF样FVIII保护性质。在某些实施方案中,VWF片段或XTEN序列还可遮蔽一个或多个FVIII清除受体结合位点,从而通过其自身清除途径阻止FVIII的清除。

[0563] 在一些实施方案中,可在体外或体内通过VWF片段或XTEN序列阻止或抑制结合于内源性VWF的FVIII蛋白。

[0564] 也提供了增加FVIII蛋白的半衰期的方法,所述方法包括向需要其的受试者施用本文所述的嵌合蛋白。在血浆中,结合于全长VWF或与全长VWF缔合的未活化FVIII的半衰期为约12至14小时。在VWD 3型中,其中循环中基本无VWF,FVIII的半衰期仅为约六小时,由于FVIII浓度降低而在此类患者中导致轻至中度血友病A的症状。与结合于全长VWF或与全长VWF缔合的未活化FVIII的半衰期相比,连接于本发明的VWF片段或XTEN序列或与本发明的VWF片段或XTEN序列缔合的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。

[0565] 在一个实施方案中,在包含XTEN序列的嵌合蛋白中连接于VWF片段或与VWF片段缔合或者连接于Ig恒定区或其一部分的FVIII蛋白的半衰期相较于结合于全长VWF或与全长VWF缔合的未活化FVIII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,在包含XTEN序列的嵌合蛋白中连接于VWF片段或Ig恒定区或其一部分或与VWF片段或Ig恒定区或其一部分缔合的FVIII蛋白的半衰期相较于结合于全长VWF或与全长VWF缔合的未活化FVIII或野生型FVIII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,在包含XTEN序列的嵌合蛋白中连接于VWF片段或与VWF片段缔合或者连接于Ig恒定区的FVIII蛋白的半衰期相较于在FVIII和VWF双敲除小鼠中野生型FVIII的半衰期增加至少约30、31、32、33、34、35、36、37、38、39或40倍。

[0566] 在一些实施方案中,包含融合于第一Ig恒定区或其一部分(例如第一Fc区)的VWF片段和XTEN序列以及连接于XTEN序列和第二Ig恒定区或其一部分(例如第二Fc区)的FVIII蛋白的嵌合蛋白的半衰期比与内源性VWF缔合的FVIII的半衰期长。在其它实施方案中,嵌合蛋白的半衰期为野生型FVIII或与内源性VWF缔合的FVIII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。

[0567] 在一些实施方案中,作为本发明的结果,FVIII蛋白的半衰期相较于不含VWF片段的FVIII蛋白或野生型FVIII有所延长。本发明的嵌合蛋白的半衰期为不含VWF片段的FVIII

蛋白或野生型FVIII的半衰期的至少约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倍长。在一个实施方案中，FVIII的半衰期为野生型FVIII的半衰期的约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，相较于野生型FVIII或不含VWF片段的FVIII蛋白，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，本发明的嵌合蛋白的半衰期为至少约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。在又其它实施方案中，本发明的嵌合蛋白的半衰期为约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。

[0568] 在一些实施方案中，每位受试者的本发明的嵌合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、约十天、约11天、约12天、约13天或约14天。

[0569] 此外，本发明提供治疗或预防出血性疾病或病症的方法，所述方法包括施用有效量的嵌合蛋白。在一个实施方案中，出血性疾病或病症选自以下组成的组：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、胸腔内溢血、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。在具体实施方案中，出血性疾病或病症为血友病A。

[0570] 通过本发明制备的、阻止或抑制FVIII蛋白与内源性VWF的相互作用、包含与本文所述的VWF片段组合的XTEN序列和Ig恒定区或其一部分的嵌合蛋白具有许多将由本领域技术人员识别的用途，包括但不限于治疗具有止血障碍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，本发明涉及治疗具有止血障碍的受试者的方法，所述方法包括施用治疗有效量的嵌合蛋白。

[0571] 嵌合蛋白中的FVIII蛋白部分通过将带负电荷的磷脂表面上的因子IX用作辅因子、从而形成X酶复合物来治疗或预防止血障碍。活化凝结因子与磷脂表面的结合使这个过程集中于血管损伤的位点。在磷脂表面上，因子VIIIa通过因子IXa使因子X活化的最大速率增加约200,000倍，从而导致凝血酶产生的第二次大爆发。

[0572] 本发明的嵌合蛋白可用于治疗任何止血障碍。可通过施用本发明的嵌合蛋白进行治疗的止血障碍包括但不限于，血友病A、以及与因子VIII有关的缺乏症或结构异常。在一个实施方案中，止血障碍为血友病A。

[0573] 本发明的嵌合蛋白可防治性地用于治疗具有止血障碍的受试者。本发明的嵌合蛋白可用于治疗具有止血障碍的受试者中的急性出血事件。在另一个实施方案中，止血障碍可以为缺陷性凝血因子例如血管性血友病因子的结果。在一个实施方案中，止血障碍为遗传性疾病。在另一个实施方案中，止血障碍为获得性疾病。获得性疾病可由潜在的继发性疾病或病患产生。不相关疾患可以为，例如，但并不作为限制，癌症、自身免疫疾病或妊娠。获得性疾病可起因于暮年或治疗潜在的继发性病症（例如癌症化疗）的药物治疗。

[0574] 本发明也涉及治疗未患先天性止血障碍但患有继发性疾病或疾患、导致获得止血障碍（例如归因于对抗FVIII抗体的开发或手术）的受试者的方法。本发明因而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方法制备的嵌合蛋白。

[0575] 本发明也涉及减少FVIII的免疫原性或诱导更少的针对FVIII的免疫原性的方法，所述方法包括施用有效量的本文所述的嵌合蛋白或编码它们的多核苷酸。

[0576] 在一个实施方案中，需要一般止血剂的受试者正经受或即将经受手术。本发明的嵌合蛋白可在手术之前、期间或之后作为预防方案施用。本发明的嵌合蛋白可在手术之前、期间或之后施用以控制急性出血事件。

[0577] 本发明的嵌合蛋白可用于治疗具有急性出血事件的未患有止血障碍的受试者。急性出血事件可由严重创伤造成，例如手术、车祸、创伤、裂伤枪伤、或任何其它导致不受控制的出血的创伤事件。流血事件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、胸腔内溢血、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。

[0578] 在预防应用中，向尚未处于疾病状态的患者施用含有本发明的嵌合蛋白或其混合物的一种或多种组合物以增强患者的抵抗或减少与疾病或病症相关的症状。此类量被定义为“防治性有效剂量”。在治疗性应用中，有时需要以相对短间隔施用的相对高剂量（例如，每剂约1至400mg/kg的多肽，其中5至25mg的剂量对于放射免疫性缀合物而言更常用，而对于细胞毒素-药物修饰的多肽而言则剂量更高）直至疾病的进展减少或终止，并且直至患者显示疾病症状的部分或完全减轻。此后，患者被施用防治性方案。

[0579] 在一些实施方案中，本发明的嵌合蛋白或组合物用于按需治疗，其包括用于流血事件、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（头创伤）、胃肠出血、颅内溢血、腹腔内溢血、胸腔内溢血、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的治疗。受试者可以需要手术防治、围术期管理或手术治疗。此类手术包括，例如小手术、大手术、拔牙、扁桃体切除术、腹股沟疝切开术、滑膜切除术、全膝置换、开颅术、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手术。

[0580] 在一个实施方案中，本发明的嵌合蛋白静脉内、皮下、肌肉或经由任何粘膜表面，例如经口、舌下、经颊、经鼻、直肠、阴道或经由肺部途径施用。本发明的包含VWF片段和FVIII蛋白的嵌合蛋白可被植入或连接于生物聚合物固体支撑物，其允许嵌合蛋白缓慢释

放至出血位点或植入绷带/敷料中。嵌合蛋白的剂量将取决于受试者和所用的特定施用途径而变化。剂量范围可为0.1至100,000 μ g/kg体重。在一个实施方案中,剂量范围为0.1-1,000 μ g/kg。在另一个实施方案中,剂量范围为0.1-500 μ g/kg。蛋白质可以连续或以特定的时间间隔施用。可以采用体外测定以确定最佳剂量范围和/或施用时间表。测量凝血因子活性的体外测定是本领域已知的,例如STA-CLOT VIIa-rTF凝血测定或ROTEM凝血测定。此外,有效剂量可从获自动物模型例如血友病犬的剂量-响应曲线推算出(Mount等人2002,Blood 99(8):2670)。

[0581] 现在已详细地描述本发明,通过参考下列实施例将更清楚地理解本发明,这些实施例包括于此仅仅为了说明的目的,而不旨在限制本发明。本文提及的所有专利、出版物和论文被明确地且具体地以引用方式并入本文。

实施例

[0582] 除非另有说明,否则在通篇实施例中使用下列材料和方法。

[0583] 材料和方法

[0584] 除非另外指出,否则本发明的实践通常采用化学、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)的常规技术和电泳标准技术。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, Irl Pr (1996); Antibodies: A Laboratory Manual, Harlow 等人, CS. H. L. Press, Pub. (1999); 和 Current Protocols in Molecular Biology, Ausubel 等人编, John Wiley & Sons (1992)。

[0585] 实施例1: 克隆不同的VWF结构域(图1)

[0586] (a) 克隆pSYN-VWF-002

[0587] pSYN-VWF-002含有编码VWF片段的核苷酸序列,其为SEQ ID NO:100的氨基酸1-477。[VWF-D'D3蛋白序列]氨基酸编号表示不含肽原的成熟VWF序列并且对应于SEQ ID NO:2的氨基酸764-1240。pSYN-VWF-002构建体具有在N-末端的FVIII信号肽,其允许适当分泌合成蛋白质和随后在C-末端的6xHis标签(用于蛋白质纯化)。其通过使用下列引物组合进行合成:

[0588] 具有VIII信号和BsiW1位点的ESC48-Fwd-VWF-D'D3

TCGCGACGTACGGCCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGT

[0589] GCCTTTTGCATTCTGCTTTAGCCTATCCTGTCGGCCCCCATG (SEQ ID NO: 90)

[0590] 具有6His和Not1位点的ESC51-Rev-VWF D'D3 (1-477个氨基酸)

TGACCTCGAGCGGCCGCTCAGTGGTGATGGTGATGATGCGGCTCCTGGCAGGCTT

[0591] CACAGGTGAGGTTGACAAC (SEQ ID NO: 91)

[0592] 利用ESC48/ESC 51引物组合和全长VWF质粒作为模板、使用2步PCR扩增循环:94 $^{\circ}$ C 2分钟;21个循环的(96 $^{\circ}$ C 30秒,68 $^{\circ}$ C 2分钟)进行50 μ l PCR反应。1460bp条带用凝胶提取试剂

盒(Qiagen,Valencia,Calif.)进行凝胶纯化并且克隆到pcDNA 4的BsiWI和NotI限制位点以产生pSYN-VWF 002。

[0593] (b) 克隆pSYN-VWF-010和013

[0594] 使用pSYN-VWF-008和pSYN-VWF-002构建pSYN-VWF-010。pSYN-VWF-008含有在pcDNA3.1 (SEQ ID NO:2的氨基酸1-2813)中的全长VWF序列,其包括763个氨基酸的肽原(即,D1D2结构域)和紧随的成熟VWF的剩余2050个氨基酸的序列。pSYN-VWF-002中的FVIII信号肽被替换为来自pSYN-VWF-008的D1D2结构域,所得构建体为pSYN-VWF-010。pSYN-VWF-008具有在Arg907的BamHI位点和在编码区末端(在终止密码子之后)的NotI位点。pSYN-VWF-008和002用BamHI和NotI限制性内切酶消化。将来自pSYN-VWF-002(1026 bp)的插入物连接到经bamHI/NotI消化的pSYN-VWF-008(8242bp)以获得pSYN-VWF-010(D1D2D'D3:SEQ ID NO:2的氨基酸1-1240),也在C-末端加入6xHis标签。在转化的细胞中,用肽原合成pSYN-VWF-010但由于胞内加工,所以分泌的产物不含任何肽原(D1D2)。来自VWF-010的蛋白质以二聚体形式存在。

[0595] pSYN-VWF-010用于产生pSYN-VWF-013,其具有在对应于SEQ ID NO:100的C336A和C379A(氨基酸编号表示不含D1D2结构域-VWF序列2的成熟VWF序列)处的两个点突变。预测这些突变阻止VWF D'D3结构域的二聚化。

[0596] (c) 克隆pSYN-VWF-025和pSYN-VWF-029

[0597] pSYN-VWF-025含有在pLIVE载体中全长VWF的野生型D1D2D'D3序列,并且pSYN-VWF-029含有具有C336A和C379A突变的D1D2D'D3序列。对于克隆pSYN-VWF-025,使用下列引物组合:

[0598] 具有NheI位点的ESC 89-fwd=CTCACTATAGGGAGACCCAAGCTGGCTAGCCG (SEQ ID NO:92)

[0599] 具有SalI的ESC 91-rev=CTGGATCCCGGGAGTCGACTCGTCAGTGGTGATGGTGATGATG (SEQ ID NO:93)

[0600] 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)质粒作为模板,使用3步PCR扩增循环:94℃2分钟;21个循环的(96℃-30秒,55℃-30秒,68℃-4分钟)进行50μl PCR反应。期望大小的条带(约3800bp)用凝胶提取试剂盒(Qiagen,Valencia,Calif.)进行凝胶纯化并且克隆到pLIVE-Mirus载体(Invitrogen,Carlsbad,Calif.)的NheI和SalI限制位点以产生pSYN-VWF 025和029。

[0601] (d) 克隆pSYN-VWF-031

[0602] pSYN-VWF-031为D1D2D'D3(C336A/C379A)-Fc构建体,其在VWF D1D2D'D3(C336A/C379A)和Fc序列之间具有48个氨基酸长的凝血酶可切割的接头(8x GGGGS(SEQ ID NO 94)+凝血酶位点)。为了制备该构建体,从构建体pSYN-FVIII-064(参考下面FVIII-VWF构建体)扩增VWF-Fc区。用XbaI和NheI消化pSYN-FVIII-VWF。含有VWF片段和Fc区的4165bp的所得插入区域用作用于通过引物组合LW 22/LW23扩增VWF和Fc区的模板。

[0603] 具有FVIII信号序列和BsiWI位点的LW 22-FWD-VWF-D'D3

[0604] GCGCCGGCCGTACGATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTT
GCGATTCTGCTTTAGCCTATCCTGTCGGCCCCCATG (SEQ ID NO: 95)

[0605] 具有终止密码子和NotI位点的LW 23-Rev-Fc

[0606] TCATCAATGTATCTTATCATGTCTGAATTCGCGGCCGCTCATTACC (SEQ ID NO:96)

[0607] 将获自LW22/LW23扩增的PCR产物(约2300bp)克隆到经BsiW1/NotI消化的pSYN-VWF-002以获得pSYN-VWF-014中间体。pSYN-VWF-014含有FVIII信号肽-D'D3-20氨基酸的凝血酶可切割的接头和紧随的Fc区。

[0608] 为了产生D1D2D'D3-Fc构建体,通过标准PCR方法使用引物组合LW24/LW27从pSYN-VWF-013扩增D1D2D'D3区域。

[0609] 具有BsiW1位点的LW24-Fwd-VWF D1D2D'D3克隆寡核苷酸

[0610] GCGCCGGCCGTACGATGATTCCTGCCAGATTTGCCGGGGTG (SEQ ID NO:97)

[0611] 具有EcoRV的LW27-Rev-VWF D'D3寡核苷酸

[0612] CCACCGCCAGATATCGGCTCCTGGCAGGCTTCACAGGTGAG (SEQ ID NO:98)

[0613] 将获自LW22/LW23扩增的PCR产物(约3750bp)克隆到经BsiW1/EcoRV消化的pSYN-VWF-014以获得pSYN-VWF-015中间体。改变VWF片段和Fc区之间的接头长度以获得pSYN-VWF-031。

[0614] VWF-D1D2D'D3蛋白序列1 (SEQ ID NO:99)

[0615]

1 MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS
DFVNTFDGSM
51 YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE
FFDIHLFVNG
101 TVTQGDQRV MPYASKGLYL ETEAGYYKLS GEAYGFVARI
DGSGNFQVLL
151 SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS
WALSSGEQWC
201 ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL
VDPEPFVALC
251 EKTLCCEAGG LECACPALLE YARTCAQEGM VLYGWDHSA
CSPVCPAGME
301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG
LCVESTECPC
351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECVL
TGQSHFKSFD
401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC
TRSVTVRLPG
451 LHNSLVKLKH GAGVAMDGQD IQLPLLKGD LRIQHTVTASV
RLSYGEDLQM
501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG
LAEPRVEDFG
551 NAWKLHGDCQ DLQKQHS DPC ALNPRMTRES EEACAVLTSP
TFEACHRAVS
601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV
AWREPGRCEL
651 NCPKGQVYLQ CGTPCNLTCT SLSYPDEECN EACLEGCFCP
PGLYMDERGD
701 CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM
SGVPGSLLPD
751 AVLSSPLSHR SKRSLSCRPP MVKLVC PADN LRAEGLECTK
TCQNYDLECM
801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE
TVKIGCNTCV
851 CRDRKWNCTD HVCDATCSTI GMAHYLTFDG LKYLFPGEQ
YVLVQDYCGS
901 NPGTFRILVG NKGCSHPSVK CKKRVITLVE GGEIELFDGE
VNVKRPMKDE

951 THFEVVESGR YIILLGKAL SVVWDRHLSI SVVLKQTYQE
 KVCGLCGNFD
 1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD
 SSPATCHNNI
 1051 MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS
 CESIGDCACF
 1101 CDTIAAYAHV CAQHGVVVTW RTATLCPQSC EERNLRENGY
 ECEWRYNCA
 1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC
 VDPEDCPVCE
 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP*
 VWF-D'D3 蛋白序列2 (SEQ ID NO: 100)
 1 SLSCRPPMVK LVCPADNLRA EGLECTKTCQ NYDLECMMSG
 CVSGCLCPPG
 51 MVRHENRCVA LERCPCFHQG KEYAPGETVK IGCNTCVCRD
 RKWNCTDHVC
 101 DATCSTIGMA HYLTFDGLKY LFPGECQYVL VQDYCGSNPG
 TFRILVGNGK
 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 WRYNCAAPAC
 QVTCQHPEPL
 401 ACPVQCVEGC HAHCPPGKIL DELLQTCVDP EDCPVCEVAG
 RRFASGKKVT
 451 LNPSDPEHCQ ICHCDVVNLT CEACQEP

[0617] 实施例2:D'D3和XTEN融合物对FVIII半衰期延长的作用

[0618] 为了评价对rFVIII-XTEN融合蛋白的D'D3FVIII半衰期延长潜力,通过流体动力学注射其相应的DNA构建体VWF-025将VWF D'D3二聚体引入FVIII-VWF DK0小鼠(实施例1)。在D'D3已达到稳态表达(注射后第5天)后,通过以200IU/kg剂量静脉内注射施用单剂的rFVIII-XTEN。收集血样直到rFVIII-XTEN给药后120小时。通过FVIII显色测定分析血浆FVIII活性。通过VWF ELISA测量D'D3表达水平,并且使用WinNonlin程序分析rFVIII Fc PK曲线。

[0619] 研究结果示于图2中,并且循环中含有/不含D'D3的rFVIII-XTEN的PK参数列于表16中。D'D3二聚体进一步延长rFVIII-XTEN $t_{1/2}$,从3.4小时至17.8小时,增加5倍。除了半衰期之外,也观察到MRT增加5倍,AUC增加3.6倍,清除率减少3.8倍。

[0620] 我们已观察到D'D3片段和XTEN技术的协同效应,将评价一系列FVIII/VWF/XTEN构建体在血友病动物中的FVIII半衰期延长潜力。

[0621] 表16:血液循环中在含有/不含D'D3下的rFVIII-XTEN PK参数

	治疗	5 分钟回产率(%)	$t_{1/2}$ (小时)	MRT (小时)	Cl (mL/小时/kg)	V _{ss} (mL/kg)	AUC _D (小时*kg*mIU/mL/mIU)
[0622]	rFVIII _{XTEN} VWF-025	80	17.8	19.3	3.5	67.4	0.29
	rFVIII _{XTEN}	74	3.4	3.8	13.1	63.68	0.08
	改善核苷酸倍数	1.1	5.2	5.1	3.8	0.9	3.6

[0623] FVIII-XTEN的蛋白质纯化

[0624] 在BDD-FVIII的C-末端插入AE288XTEN用于该研究。为了纯化该蛋白质,首先使用切向流过滤(TFF)步骤对条件培养基进行缓冲液交换。然后使用强阴离子交换色谱捕获滤液中的产物,然后进一步使用亲和色谱进行纯化。分子的纯度为HPLC-SEC可接受的并且进一步通过免疫印迹确认。分子的比活性与B-结构域缺失的FVIII类似,如通过aPTT测定和ELISA所测量。

[0625] FVIII显色测定

[0626] FVIII活性使用来自DiaPharma的COATEST SP FVIII试剂盒(批号N089019)进行测量并且所有温育在37℃板式加热器上在振摇下执行。

[0627] rFVIII标准的范围为100mIU/mL至0.78mIU/mL。将合并的正常人血浆测定对照和血浆样本(用1X Coatest缓冲液稀释)加入到Immulon 2HB 96孔板,一式两份(25μL/孔)。新鲜制备的IXa/FX/磷脂混合物(50μL)、25μL 25mM CaCl₂和50μL FXa底物相继加入到各孔,在每次添加之间温育5分钟。在与底物温育后,将25μL 20%乙酸加入以终止显色反应,并且OD₄₀₅的吸光度用SpectraMAX+ (Molecular Devices) 仪器测量。用SoftMax Pro软件(5.2版)分析数据。最低定量水平(LLQ)为7.8mIU/mL。

[0628] VWF ELISA:

[0629] 山羊抗人VWF抗体(亲和力纯化的、生物亲和型,GAVWF-AP)以0.5ug/孔用作捕获抗体并且VWF-EIA-D(生物亲和型,VWF-EIA-D,1:100稀释)用作VWF ELISA的检测抗体。按照标准ELISA程序执行ELISA测定,TMB用作HRP底物,PBST/1.5%BSA/0.5M NaCl缓冲液用作封闭缓冲液和结合缓冲液。测定标准范围为100ng至0.78ng,并且测定的最低定量限(LLQ)为7.8ng/mL。

[0630] 实施例3:含有FVIII/VWF构建体的XTEN的质粒构建

[0631] (a) 克隆pSYN-FVIII-161(图3)

[0632] FVIII-161质粒包含具有酶切割位点的单链Fc(scFc)支架,所述酶切割位点在合成期间于细胞中进行加工。构建体具有全长VWF的FVIII结合结构域(D'D3)。

[0633] 质粒(pSYN-FVIII-161)被设计用于表达FVIII-Fc和VWF-Fc异二聚体,在其中D'D3结构域结合于FVIII且阻止FVIII与磷脂和活化蛋白C的相互作用。来自pSYN-FVIII-161的蛋白质在细胞中被表达为单一多肽,在单一多肽中FVIII-Fc亚单位的C-末端通过6x(GGGGS)多肽接头(SEQ ID NO:64)连接于VWF D'D3-Fc亚单位的N-末端。此外,RRRRS(SEQ ID NO:11)和RKRRKR(SEQ ID NO:10)序列分别在多肽接头的5'和3'端插入,用于通过每个序列中紧随最后一个Arg的蛋白原转化酶进行胞内切割。因此,细胞可表达双链FVIII-Fc/D'D3-Fc异二聚体,在其中FVIII-Fc链具有在C-末端的RRRRS序列(SEQ ID NO:11),但是接

头序列的其余部分已被去除。一旦FVIII-VWF异二聚体蛋白被凝血酶活化,允许FVIII与其它凝血因子相互作用,就在VWF结构域和Fc区之间引入紧随IS {5X (GGGS)} LVPRGSGG (SEQ ID NO:122) 多肽(含有凝血酶切割位点)的AE288XTEN片段以促进VWF片段从FVIII的释放。

[0634] pSYN-FVIII-161 (SEQ ID NO:101) 蛋白序列(FVIII序列氨基酸位置1-1457;加下划线的区域表示Fc区;波浪下划线表示在第一Fc和VWF片段之间的可切割的接头;加双下划线的区域表示VWF片段;粗体区域表示在VWF片段和Fc之间的可切割的接头。

```
1  MQIELSTCFF LCLLRFCFSA TRRYYLGAVE LSWDYMQSDL  
   GELPVDARFP  
51  PRVPKSFFPN TSVVYKKTLE 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  
[0635] 451 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR  
   PYNIYPHGIT  
501 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP  
   TKSDPRCLTR  
551 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR  
   NVILFSVFDE  
601 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV  
   FDSLQLSVCL  
651 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEYEDTLTLF  
   PFSGETVFMS  
701 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED  
   SYEDISAYLL  
751 SKNNAIEPRS FSQNPPVLKR HQREITRTTL QSDQEEIDYD  
   DTISVEMKKE  
801 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP  
   HVLNRNAQSG  
851 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA  
   EVEDNIMVTF  
901 RNQASRPYSF YSSLISYEED QRQGAEPRKN FVKPNETKTY  
   FWKVQHMAP
```

[0636]

951 TKDEFDCKAW AYFSDVDLEK DVHSGLIGPL LVCHTNTLNP
AHGRQVTVQE
1001 FALFFTIFDE TKSIFYTENM ERNCRAPCNI QMEDPTFKEN
YRFHAINGYI
1051 MDTLPGLVMA QDQIRWYLL SMGSNENIHS IHFSGHVFTV
RKKEEYKMAL
1101 YNLYPGVFET VEMLPSKAGI WRVECLIGEH LHAGMSTLFL
VYSNKCQTPL
1151 GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK
EPFSWIKVDL
1201 LAPMIIHGK TQGARQKFSS LYISQFIIMY SLDGKKWQTY
RGNSTGTLMV
1251 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
LMGCDLNSCS
1301 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
SNAWRPQVNN
1351 PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISS
QDGHQWTLFF
1401 QNGKVKVFQG NQDSFTPVVN SLDPPLLTRY LRIHPQSWVH
QIALRMEVLG
1451 CEAQDLYDKT HTCPCPAPE LLGGPSVFLF PPKPKDTLMI
SRTPEVTCVV
1501 VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
SVLTVLHQDW
1551 LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP
SRDELTKNQV
1601 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGGS
FFLYSKLTVD
1651 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKRRRRSG
GGSGGGGSG
1701 GGSGGGGSG GGGSGGGGSR KRRKRSLSR PPMVKLVCPA
DNLRAEGLEC
1751 TKTCQNYDLE CSMGCVSGC LCPPGMVRHE NRCVALERCP
CFHQGKEYAP
1801 GETVKIGCNT CVCRRKWNC TDHVCDATCS TIGMAHYLTF
DGLKYLFPGE
1851 COYVLVODYC GSNPGTFRIL VGNKGCSHPS VKCKKRVTL
VEGGEIELED
1901 GEVNVKRPMK DETHFEVVES GRYIILLGK ALSVVWDRHL
SISVVLKQTY
1951 QEKVCGLCGN FDGIQNNDLT SSNLQVEEDP VDFGNSWKVS
SQCADTRKVP
2001 LDSSPATCHN NIMKQTMVDS SCRILTSDFV QDCNKLVDPE
PYLDVCIYDT
2051 CSCESIGDCA AFCDTIAAYA HVCAQHGVV TWRTATLCPO
SCEERNLREN
2101 GYEAEWRYNS CAPACQVTCQ HPEPLACPVQ CVEGCHAHCP
PGKILDELLO
2151 TCVDPEDCPV CEVAGRRFAS GKKVTLNPSD PEHCQICHCD
VVNLTCACQ

	2201	<u>EPISGTSESA</u> <u>TPESGPGSEP</u> <u>ATSGSETPGT</u> <u>SESATPESGP</u>
	<u>GSEPATSGSE</u>	
	2251	<u>TPGTSESATP</u> <u>ESGPGTSTEP</u> <u>SEGSAPGSPA</u> <u>GSPTSTEEGT</u>
	<u>SESATPESGP</u>	
	2301	<u>GSEPATSGSE</u> <u>TPGTSESATP</u> <u>ESGPGSPAGS</u> <u>PTSTEEGSPA</u>
	<u>GSPTSTEEGT</u>	
	2351	<u>STEPSEGSAP</u> <u>GTSESATPES</u> <u>GPGTSESATP</u> <u>ESGPGTSESA</u>
	<u>TPESGPGSEP</u>	
	2401	<u>ATSGSETPGS</u> <u>EPATSGSETP</u> <u>GSPAGSPTST</u> <u>EEGTSTEPSE</u>
	<u>GSAPGTSTEP</u>	
	2451	<u>SEGSAPGSEP</u> <u>ATSGSETPGT</u> <u>SESATPESGP</u> <u>GTSTEPSEGS</u>
[0637]	<u>APDSGGGGSG</u>	
	2501	<u>GGGSGGGGSG</u> <u>GGGSGGGGSL</u> <u>VPRGSGG</u> <u>DKT</u> <u>HTCPPCPAPE</u>
	<u>LLGGPSVFLF</u>	
	2551	<u>PPKPKDTLMI</u> <u>SRTPEVTCVV</u> <u>VDVSHEDPEV</u> <u>KFNWYVDGVE</u>
	<u>VHNAKTKPRE</u>	
	2601	<u>EQYNSTYRVV</u> <u>SVLTVLHQDW</u> <u>LNGKEYKCKV</u> <u>SNKALPAPIE</u>
	<u>KTISKAKGQP</u>	
	2651	<u>REPQVYTLPP</u> <u>SRDELTKNQV</u> <u>SLTCLVKGFY</u> <u>PSDIAVEWES</u>
	<u>NGQPENNYKT</u>	
	2701	<u>TPPVLDSDGS</u> <u>FFLYSKLTVD</u> <u>KSRWQQGNVF</u> <u>SCSVMEALH</u>
	<u>NHYTQKSLSL</u>	
	2751	<u>SPGK</u>

[0638] (b) 克隆_pSYN-FVIII-168、175、172和174 (图4A-4D)

[0639] pSYN-FVIII-168、172、174和175为pSYN-FVIII-161的衍生物。将R1645A/R1648A突变引入pSYN-FVIII-161以形成pSYN-FVIII-168,其生成SC-FVIII同工型,并且将AE288XTEN直接融合到FVIII-HC的C-末端用于进一步延长半衰期。为了构建pSYN-FVIII-175,从pSYN-FVIII-168去除D'D3密码子序列,用于评价Fc和XTEN技术对FVIII半衰期延长的作用。

[0640] 为了构建pSYN-FVIII-172,将AE288XTEN片段直接到FVIII-HC的C-末端用于进一步延长半衰期,并且从pSYN-FVIII-172去除D'D3密码子序列以形成pSYN-FVIII-174,用于评价Fc和XTEN技术对FVIII半衰期延长的作用。

[0641] (c) 克隆pSYN-FVIII-170 (图4E)

[0642] 构建pSYN-FVIII-170以评价XTEN和D'D3片段对FVIII半衰期延长的作用。将密码子序列VWF-D1D2D'D3片段和BDD-FVIII引入表达盒的5'和3'端,使用紧随35aa凝血酶可切割的接头的AE288XTEN密码子序列来连接VWF和FVIII分子。在胞内加工后,分泌蛋白包含含有成熟VWF分子的D'D3片段的多肽,所述D'D3片段通过AE288XTEN/35aa凝血酶可切割的接头连接于成熟BDD-FVIII的N-末端。

[0643] pSYN-FVIII-170蛋白序列 (SEQ ID NO:102)

[0644]

1 SLSCRPPMVK LVCPADNLRA EGLECTKTCQ NYDLECMSMG
 CVSGCLCPPG
 51 MVRHENRCVA LERCPCFHQG KEYAPGETVK IGCNTCVCRD
 RKWNCTDHVC
 101 DATCSTIGMA HYLTFDGLKY LFPGECQYVL VQDYCGSNPG
 TFRILVGNGK
 151 CSHPSVKCKK RVTILVEGGE IELFDGEVNV KRPMKDETHF
 EVVESGRYII
 201 LLLGKALSVV WDRHLSISVV LKQTYQEKVC GLCGNFDGIQ
 NNDLTSSNLQ
 251 VEEDPVDFGN SWKVSSQCAD TRKVPLDSSP ATCHNNIMKQ
 TMVDSSCRIL
 301 TSDVFQDCNK LVDPEPYLDV CIYDTCSCES IGDCAAFCDT
 IAAYAHVCAQ
 351 HGKVVTWRTA TLCPQSCEER NLRENGYEAE WRYNSCAPAC
 QVTCQHPEPL
 401 ACPVQCVEGC HAHCPPGKIL DELLQTCVDP EDCPVCEVAG
 RRFASGKKVT
 451 LNPSDPEHCQ ICHCDVNLIT CEACQEPISG TSESATPESG
PGSEPATSGS
 501 ETPGTSESAT PESGPGSEPA TSGSETPGTS ESATPESGPG
TSTEPSEGS
 551 PGSPAGSPTS TEEGTSESAT PESGPGSEPA TSGSETPGTS
ESATPESGPG
 601 SPAGSPTSTE EGSPAGSPTS TEEGTSTEPS EGSAPGTSES
ATPESGPGTS
 651 ESATPESGPG TSESATPESG PGSEPATSGS ETPGSEPATS
GSETPGSPAG
 701 SPTSTEEGTS TEPSEGSAPG TSTEPSEGS PGSEPATSGS
ETPGTSESAT
 751 PESGPGTSTE PSEGSAPDSG GGSGGGGGSG GGSGGGGGSG
GGGSLVPRGS
 801 GGASATRRYY LGAVELSWDY MQSDLGELPV DARFPFRVPK
 SFPFNTSVVY
 851 KKTLEFVEFTD HLFNIAKPRP PWMGLLGPTI QAEVYDTVVI
 TLKNMASHPV
 901 SLHAVGVSYW KASEGAEYDD QTSQREKEDD KVFPGGSHY
 VWQVLKENG
 951 MASDPLCLTY SYLSHVDLVK DLNSGLIGAL LVCREGSLAK
 EKTQTLHKFI


```

1001 LLFAVFDEGK SWHSETKNSL MQDRDAASAR AWPKMHTVNG
YVNRSLPGLI
1051 GCHRKSVYWH VIGMGTTPPEV HSIFLEGHTF LVRNHRQASL
EISPITFLTA
1101 QTLLMDLGQF LLFCHISSHQ HDGMEAYVKV DSCPEEPQLR
MKNNEEAEDY
1151 DDDLTDSEMD VVRFDDDNSP SFIQIRSVAK KHPKTWVHYI
AAEEEDWDYA
1201 PLVLAPDDRS YKSQYLNGP QRIGRKYKKV RFMAYTDETF
KTREAIQHEs
1251 GILGPLLYGE VGDTLLIIFK NQASRPYNIY PHGITDVRPL
YSRRLPKGVK
1301 HLKDFPILPG EIFKYKWTVT VEDGPTKSDP RCLTRYSSFF
VNMERDLASG
1351 LIGPLLYCYK ESVDQQRNGI MSDKRNVLIF SVFDENRSWY
LTENIQRFLP
1401 NPAGVQLEDP EFQASNIMHS INGYVFDSLQ LSVCLHEVAY
WYILSIGAQT
1451 DFLSVFFSGY TFKHKMVYED TLTLFPFSGE TVFMSMENPG
LWILGCHNSD
1501 FRNRGMTALL KVSSCDKNTG DYYEDSYEDI SAYLLSKNNA
IEPRSFQNP
1551 PVLKRHQREI TRTTLQSDQE EIDYDDTISV EMKKEDFDIY
DEDENQSPRS
[0645] 1601 FQKKTRHYFI AAVERLWDYG MSSSPHVLRN RAQSGSVPQF
KKVVFQEFTD
1651 GSFTQPLYRG ELNEHLGLLG PYIRAEVEDN IMVTFRNQAS
RPYSFYSSLI
1701 SYEEDQRQGA EPRKNFVKPN ETKTYFWKVQ HHMAPTKDEF
DCKAWAYFSD
1751 VDLEKDVHSG LIGPLLVCHT NTLNPAHGRQ VTVQEFALFF
TIFDETKSWY
1801 FTENMERNCR APCNIQMEDP TFKENYRFHA INGYIMDTLP
GLVMAQDQRI
1851 RWYLLSMGSN ENIHSIHFSG HVFTVRKKEE YKMALYNLYP
GVFETVEMLP
1901 SKAGIWRVEC LIGEHLHAGM STLFLVYSNK CQTPLGMASG
HIRDFQITAS
1951 GQYGQWAPKL ARLHYSGSIN AWSTKEPFSW IKVDLLAPMI
IHGIKTQGAR
2001 QKFSSLYISQ FIIMYSLDGK KWQTYRGNST GTLMVFFGNV
DSSGIKHNI
2051 NPPIIARYIR LHPHYSIRS TLRMELMGCD LNSCSMPLGM
ESKAISDAQI
2101 TASSYFTNMF ATWSPSKARL HLQGRSNAWR PQVNNPKEWL
QVDFQKTMKV
2151 TGVTTQGVKS LLTSMYVKEF LISSSQDGHQ WTLFFQNGKV
KVFQGNQDSF
2201 TPVVNSLDPP LLTRYLRIHP QSWVHQIALR MEVLGCEAQD LY

```

[0646] 实施例4:在FVIII和VWF缺陷小鼠中流体动力学注射含有FVIIIIF/VWF构建体的XTEN

[0647] 图3和4中含有DNA构建体的XTEN已将2-3个半衰期延长元件组合在一起。为了评价它们的FVIII半衰期延长潜力,通过以100ug/小鼠剂量流体动力学注射(HDI)将图3和图4中

DNA构建体的选择组引入FVIII/VWF双敲除(DKO)小鼠中。然后通过HDI后24小时眶后血液采集来收集血样。HDI后血浆FVIII活性通过FVIII显色测定来分析,且结果于表17和图5中列出。相较于野生型BDD-FVIII,所有含有DNA构建体的XTEN在HDI后24小时得到显著更高的FVIII血浆活性,指示相应的分子具有比BDD-FVIII显著更长的循环蛋白半衰期。进一步在血友病动物中评价那些半衰期延长元件的组合的应用。

[0648] 表17:在FVIII/VWF DKO小鼠中HDI后24小时的FVIII血浆活性

[0649]	DNA 构建体	BDD- FVIII	FVIII-161	FVIII- 168	FVIII- 172	BDD- FVIII	FVIII- 170
	DNA 剂量 ($\mu\text{g}/\text{小鼠}$)	100	100	100	100	50	50
	FVIII 活性 (mU/mL)	219 $\pm$ 72	2446 $\pm$ 1012	2209 $\pm$ 609	1671 $\pm$ 223	197 $\pm$ 21	399 $\pm$ 30

[0650] 流体动力学注射:

[0651] 流体动力学注射是一种至诸如小鼠和大鼠的小型动物的肝脏的有效的且安全的非病毒基因递送方法。其最初可被描述成在约5-7秒内以动物体重的十分之一快速注射不含内毒素的裸质粒DNA/盐水溶液。裸质粒DNA含有目标基因并且肝脏以动物体重的十分之一生成。靶蛋白于肝脏中由经注射的DNA生成并且可在注射后24小时被检测。然后收集血浆样本以研究表达蛋白质的疗效。

[0652] 对于本文执行的所有流体动力学注射,将2ml于0.9%无菌盐水溶液中的质粒DNA在约4-7秒内经由静脉内尾静脉注射递送至称重为20-35克的小鼠。在最开始的前几个小时内密切监测小鼠直至正常活性恢复。在经由眶后血液采集收集血样后,然后获得血浆样本并将其贮存于-80℃用于进一步分析。

[0653] 实施例5:含有XTEN插入的FVIIIIFc-VWF异二聚体的共转染系统的质粒构建(图6)

[0654] 为了增加蛋白生产率,产生用于蛋白产生的两个共转染系统,其含有三个DNA构建体。第一DNA构建体编码FVIII-Fc融合蛋白,在其中AE288XTEN片段直接融合于FVIII重链的C-末端和然后野生型FVIII轻链片段(pSYN-FVIII-173,图6B)或具有R1645A/R1648A突变的FVIII轻链片段(pSYN-FVIII-169,图6A),然后将FVIII轻链直接融合于单个Fc片段。第二DNA构建体为pSYN-VWF-031,其编码D'D3-Fc融合蛋白(实施例1)。将HEK293F细胞用两个质粒连同第三质粒(PC5)以80:15:5比率转染。合成蛋白质以FVIII(XTEN)Fc/D'D3Fc异二聚体和D'D3Fc二聚体形式分泌并且通过蛋白质纯化从D'D3Fc二聚体分离FVIII(XTEN)Fc/D'D3Fc异二聚体。

[0655] pSYN-FVIII-169成熟蛋白序列(SEQ ID NO:103):

[0656]

1 ATRRYYLGAV ELSWDYMQSD LGELPVDARF PPRVPKSFPF
NTSVVYKKTL
51 FVEFTDHLFN IAKPRPPWMG LLGPTIQAEV YDTVVITLKN
MASHPVSLHA
101 VGVSYWKASE GAEYDDQTSQ REKEDDKVFP GGSHTYVWQV
LKENGPMASD
151 PLCLTYSYLS HVDLVKDLNS GLIGALLVCR EGSLAKEKTQ
TLHKFILLFA
201 VFDEGKSWHS ETKNSLMQDR DAASARAWPK MHTVNGYVNR
SLPGLIGCHR
251 KSVYWHVIGM GTTPEVHSIF LEGHTFLVRN HRQASLEISP
ITFLTAQTLL
301 MDLGQFLLFC HISSHQHDGM EAYVKVDSCP EEPQLRMKNN
EEAEDYDDDL
351 TDSEMDVVRF DDDNSPSFIQ IRSVAKKHPK TWVHYIAAEE
EDWDYAPLVL
401 APDDRSYKSQ YLNNGPQRIG RKYKKVRFMA YTDETFKTRE
AIQHESGILG
451 PLYGEVGD TLLIIFKNQAS RPYNIYPHGI TDVRPLYSR
LPKGVKHLKD

[0657]

501 FPILPGEIFK YKWTVTVEDG PTKSDPRCLT RYSSSFVNME
RDLASGLIGP
551 LLICYKESVD QRGNQIMSDK RNVILFSVFD ENRSWYLTEN
IQRFLPNPAG
601 VQLEDPEFQA SNIMHSINGY VFDSLQLSVC LHEVAYWYIL
SIGAQTDFLS
651 VFFSGYTFFKH KVMYEDTLTL FPFSGETVFM SMENPGLWIL
GCHNSDFRNR
701 GMTALLKVSS CDKNTGDYIE DSYEDISAYL LSKNNAIEPR
SFSQNGAPGT
751 SESATPESGP GSEPATSGSE TPGTSESATP ESGPGSEPAT
SGSETPGTSE
801 SATPESGPGT STEPSEGSAP GSPAGSPTST EEGTSESATP
ESGPGSEPAT
851 SGSETPGTSE SATPESGPGS PAGSPTSTEE GSPAGSPTST
EEGTSTEPSE
901 GSAPGTSESA TPESGPGTSE SATPESGPGT SESATPESGP
GSEPATSGSE
951 TPGSEPATSG SETPGSPAGS PTSTEEGTST EPSEGSAPGT
STEPSEGSAP
1001 GSEPATSGSE TPGTSESATP ESGPGTSTEP SEGSAPASSP
PVLKRHQAEI
1051 TRTTLQSDQE EIDYDDTISV EMKKEDFDIY DEDENQSPRS
FQKKTRHYFI
1101 AAVERLWDYG MSSSPHVLNR RAQSGSVPQF KKVVFQEFTD
GSFTQPLYRG
1151 ELNEHLGLLG PYIRAEVEDN IMVTFRNQAS RPYSFYSSLI
SYEEDQRQGA
1201 EPRKNFVKPN ETKTYFWKVQ HHMAPTKDEF DCKAWAYFSD
VDLEKDVHSG
1251 LIGPLLVCHT NTLNPAHGRQ VTVQEFALFF TIFDETKSWY
FTENMERNCR
1301 APCNIQMEDP TFKENYRFHA INGYIMDTLP GLVMAQDQRI
RWYLLSMGSN
1351 ENIHSIHFG HVFTVRKKEE YKMALYNLYP GVFETVEMLP
SKAGIWRVEC
1401 LIGEHLHAGM STLFLVYSNK CQTPLGMASG HIRDFQITAS
GQYGQWAPKL
1451 ARLHYSGSIN AWSTKEPFSW IKVDLLAPMI IHGIKTQGAR
QKFSSLYISQ
1501 FIIMYSLDGK KWQTYRGNST GTLMVFFGNV DSSGIKHNI
NPPIIARYIR
1551 LHPHYSIRS TLRMELMGCD LNCSMPLGM ESKAISDAQI
TASSYFTNMF
1601 ATWSPSKARL HLQGRSNAWR PQVNNPKEWL QVDFQKTMKV
TGVTTQGVKS
1651 LLTSMYVKEF LISSSQDGHQ WTLFFQNGKV KVFQGNQDSF
TPVNSLDPP
1701 LLTRYLRIHP QSWVHQIALR MEVLGCEAQD LYDKTHTCPP
CPAPELLGGP

1751 SVFLFPPKPK DTLMISRTPE VTCVVVDVSH EDPEVKFNWY
VDGVEVHNAK
1801 TKPREEQYNS TYRVVSVLTV LHQDWLNGKE YKCKVSNKAL
PAPIEKTISK
[0658] 1851 AKGQPREPQV YTLPPSRDEL TKNQVSLTCL VKGFYPSDIA
VEWESNGQPE
1901 NNYKTTTPVL DSDGSFFLYS KLTVDKSRWQ QGNVFSCSVM
HEALHNHYTQ
1951 KSLSLSPGK

[0659] pSYN-FVIII-173成熟蛋白测序 (SEQ ID NO:104):

1 ATRRYILGAV ELSWDYMQSD LGELPVDARF PPRVPKSFPF
NTSVVYKCTL
51 FVEFTDHLFN IAKPRPPWMG LLGPTIQAEV YDTVVITLKN
MASHPVSLHA
101 VGVSYWKASE GA EYDDQTSQ REKEDDKVFP GGSHTYVWQV
LKENGPMASD
151 PLCLTYSYLS HVDLVKDLNS GLIGALLVCR EGSLAKEKTQ
TLHKFILLFA
201 VFDEGKSWHS ETKNSLMQDR DAASARAWPK MHTVNGYVNR
SLPGLIGCHR
251 KSVYWHVIGM GTTPEVHSIF LEGHTFLVRN HRQASLEISP
ITFLTAQTLL
301 MDLGQFLLFC HISSHQHDGM EAYVKVDSCP EEPQLRMKNN
EEAEDYDDDL
351 TDSEMDVVRF DDDNSPSFIQ IRSVAKKHPK TWVHYIAAEE
EDWDYAPLVL
401 APDDRSYKSQ YLNNGPQRIG RKYKKVRFMA YTDETFKTRE
AIQHESGILG
451 PLYGEVGD T LLIIFKNQAS RPYNIYPHGI TDVRPLYSRR
LPKGVKHLKD
[0660] 501 FPILPGEIFK YKWTVTVEDG PTKSDPRCLT RYSSSFVNME
RDLASGLIGP
551 LLICYKESVD QRGNQIMSDK RNVILFSVFD ENRSWYL TEN
IQRFLPNPAG
601 VQLEDPEFQA SNIMHSINGY VFDSLQLSVC LHEVAYWYIL
SIGAQTDFLS
651 VFFSGYTFKH KMYEDTLTL FPFSGETVFM SMENPGLWIL
GCHNSDFRNR
701 GMTALLKVSS CDKNTGDYIE DSYEDISAYL LSKNNAIEPR
SFSQNGAPGT
751 SESATPESGP GSEPATSGSE TPGTSESATP ESGPGSEPAT
SGSETPGTSE
801 SATPESGPGT STEPSEGSAP GSPAGSPTST EEGTSESATP
ESGPGSEPAT
851 SGSETPGTSE SATPESGPGS PAGSPTSTEE GSPAGSPTST
EEGTSTEPSE
901 GSAPGTSESA TPESGPGTSE SATPESGPGT SESATPESGP
GSEPATSGSE
951 TPGSEPATSG SETPGSPAGS PTSTEEGTST EPSEGSAPGT
STEPSEGSAP

1001 GSEPATSGSE TPGTSESATP ESGPGTSTEP SEGSAPASSP
PVLKRHQREI
1051 TRTTLQSDQE EIDYDDTISV EMKKEDFDIY DEDENQSPRS
FQKKTRHYFI
1101 AAVERLWDYG MSSSPHVLN RAQSGSVPQF KKVVFQEFDT
GSFTQPLYRG
1151 ELNEHLGLLG PYIRAEVEDN IMVTFRNQAS RPYSFYSSLI
SYEEDQRQGA
1201 EPRKNFVKPN ETKTYFWKVQ HHMAPTKDEF DCKAWAYFSD
VDLEKDVHSG
1251 LIGPLLVCHT NTLNPAHGRQ VTVQEFALFF TIFDETKSWY
FTENMERNCR
1301 APCNIQMEDP TFKENYRFHA INGYIMDTLP GLVMAQDQRI
RWYLLSMGSN
1351 ENIHSIHFSG HVFTVRKKEE YKMALYNLYP GVFETVEMLP
SKAGIWRVEC
1401 LIGEHLHAGM STLFLVYSNK CQTPLGMASG HIRDFQITAS
GQYGQWAPKL
1451 ARLHYSGSIN AWSTKEPFSW IKVDLLAPMI IHGIKTQGAR
[0661] QKFSSLYISQ
1501 FIIMYSLDGK KWQTYRGNST GTLMVFFGNV DSSGIKHNI
NPPIIARYIR
1551 LHPHYSIRS TLRMELMGCD LNCSMPLGM ESKAISDAQI
TASSYFTNMF
1601 ATWSPSKARL HLQGRSNAWR PQVNNPKEWL QVDFQKTMKV
TGVTTQGVKS
1651 LLTSMYVKEF LISSSQDGHQ WTLFFQNGKV KVFQGNQDSF
TPVVNSLDPP
1701 LLTRYLRIHP QSWVHQIALR MEVLGCEAQD LYDKTHTCPP
CPAPELLGGP
1751 SVFLFPPKPK DTLMISRTPE VTCVVVDVSH EDPEVKFNWY
VDGVEVHNAK
1801 TKPREEQYNS TYRVVSVLTV LHQDWLNGKE YKCKVSNKAL
PAPIEKTISK
1851 AKGQPREPQV YTLPPSRDEL TKNQVSLTCL VKGFYPSDIA
VEWESNGQPE
1901 NNYKTPPVVL DSDGSFFLYS KLTVDKSRWQ QGNVFSCSVM
HEALHNHYTQ
1951 KSLSLSPGK

[0662] 实施例6.FVIII-169/VWF-031和FVIII-173/VWF-031的蛋白质纯化

[0663] 使用切向流过滤 (TFF) 步骤对澄明的条件培养基进行缓冲液交换。然后使用两步
骤色谱过程纯化FVIII-169/VWF-031或FVIII-173/VWF-031异二聚体。使用弱阴离子交换树
脂,然后使用亲和力色谱。最终纯化产物通过SEC-HPLC具有可接受的纯度。比活性与B-结构
域缺失的FVIII相容,如通过FVIII显色测定和A280浓度所测量。这个分子的纯度和各部分
的存在通过SDS-PAGE和蛋白质印迹来确认。

[0664] 实施例7.通过Octet测定评价FVIII-169/VWF-031的VWF结合能力

[0665] 在25℃用ForteBio Octet 384仪器,使用Tris结合缓冲液(50mM Tris,pH 7.2、
150mM NaCl、5mM CaCl₂),通过基于生物层干涉测量法(Bio-Layer Interferometry,BLI)
的测量结果(Octet测定)获得FVIII-169/VWF-031的VWF结合能力。用于确定FVIII结合的

Octet测定是基于将人血管性血友病因子 (Haematologic Technologies 目录号HCVWF-0191) 疏水性固定到APS生物传感器上, 然后结合1.0%牛血清白蛋白 (Jackson ImmunoResearch 目录号001-000-161)。简言之, 将hvWF (20 μ g/mL) 稀释于Tris缓冲液中并且加载到APS生物传感器上持续600秒, 得到在反应探针上的约3.0-3.5nm结合。对照APS探针在hvWF不存在的情况下用1.0%BSA加载用于参考扣除。在加载后, 将所有探针在Tris缓冲液中温育300秒以建立新基线。随后, 在室温下将生物传感器探针在FVIII-XTEN 169或FVIIIIFc原料药的溶液 (0、0.6、2、6、20、60、200、600IU/mL) 中温育5分钟, 随后进行5分钟解离步骤。使用Octet数据分析软件, 结合响应 (nm) 源自扣除的数据 (反应探针-参考探针)。对于FVIII-169/VWF-031没有检测到至固定化VWF的结合 (图7), 指示D'D3片段完全掩蔽FVIII免遭全长VWF分子。

[0666] 实施例8. 在HemA和FVIII/VWF DKO小鼠中的FVIII-169/VWF-031PK

[0667] 在HemA和FVIII/VWF DKO小鼠中测试FVIII-169/VWF-031的PK曲线以评价D'D3片段掩蔽FVIII部分免遭内源性VWF的能力。HemA或FVIII/VWF DKO小鼠用单个静脉内剂量的FVIII-169/VWF-031以200IU/kg治疗, 然后在给药后5分钟、8小时、24小时、48小时和72小时收集血浆样本。通过FVIII显色测定来测试血浆样本的FVIII活性, 并且使用WinNonlin程序计算FVIII-169/VWF-031的半衰期。

[0668] 通过针对FVIII-169/VWF-031的生物层干涉测量法 (图7) 显示完全抑制构建体与固定化VWF的结合。这指示分子中的D'D3片段已成功地封闭FVIII与天然VWF分子的结合, 因此在两个不同的小鼠品系中预测FVIII-169/VWF-031的类似的半衰期。如图8A和表18中示出, 正如预期的, FVIII-169/VWF-031在HemA和FVIII/VWF DKO小鼠中具有类似的PK曲线, 已表明FVIIIIFc/VWF异二聚体的半衰期与内源性VWF的半衰期无关。FVIIIIFc/VWF异二聚体半衰期从内源性VWF半衰期的分离消除FVIII延伸最高限度并且打开进一步延长FVIII半衰期超过由内源性VWF强加的2倍半衰期限值的可能性。

[0669] 表18: 在HemA和FVIII/VWF DKO小鼠中的FVIII-169/VWF-031 PK

小鼠品系	回收率 (%)	$t_{1/2}$ (小时)	MRT (小时)	Cl (mL/小时/kg)	Vss (mL/kg)	AUC (小时 * kg * mIU/mL/mIU)
FVIII/VWF DKO	69	17.94	20.1	4.06	81.69	0.2461
HemA	83	16.65	18.44	3.57	85.72	0.28

[0671] 通过在FVIII/VWF DKO小鼠中比较FVIII-169/VWF-031以及FVIII-169/Fc和FVIIIIFc的半衰期评价XTEN插入和D'D3片段的FVIII保护能力。在单次静脉内施用后, 对于FVIII-169/VWF-031在5分钟、8小时、24小时、48小时和72小时, 对于FVIII-169/Fc在5分钟、8小时、24小时、32小时、48小时和对于FVIIIIFc在5分钟、1、2、4、6和8小时收集血样。通过FVIII显色测定测试血浆样本的FVIII活性, 并且使用WinNonlin程序计算FVIII-155/VWF-031的半衰期。

[0672] 图8B和表19中总结了研究结果, 由于VWF保护的丧失, rFVIIIIFc在DKO小鼠中具有

1.6小时半衰期。当将XTEN插入引入到FVIIIIFc分子中时,所得FVIII-169/Fc分子具有7小时的半衰期,通过XTEN插入得到4倍半衰期延长。最终,当D'D3片段并入分子中以形成FVIII-169/VWF-031时,观察到17小时的半衰期。通过D'D3片段另外进一步增加2.5倍。除半衰期改善之外,也观察到改善的平均滞留时间(MRT)、清除率(CI)和AUC,如表19中所示。

[0673] 在HemA和FVIII/VWF DKO小鼠两者中FVIII-169/VWF-031已达到17-18小时 $t_{1/2}$,这是由VWF清除强加的 $t_{1/2}$ 延长最高限度的上限。更多的 $t_{1/2}$ 延长元件可进一步并入该分子中诸如在FVIII内的第二XTEN插入。D'D3片段和XTEN插入的协同效应提供完全保护FVIII免遭其清除途径的可能性,最终突破2倍FVIII $t_{1/2}$ 延长限度可能通过FVIIIIFc/XTEN/VWF变体来达到。

[0674] 表19. 在FVIII/VWF DKO小鼠中的FVIII-169/VWF-031PK

小鼠品系	治疗	回收率 (%)	$t_{1/2}$ (小时)	MRT (小时)	CI (mL/小时/kg)	Vss (mL/kg)	AUC/D (小时*kg*mIU/mL/mIU)
FVIII/VWF DKO	rFVIIIIFc	35	1.6	2.1	57.7	120.2	0.0173
	rFVIII-169/Fc	77	7.0	6.2	6.4	39.2	0.1573
	rFVIII-169/VWF-031	69	17.9	20.1	4.1	81.7	0.2461

[0675] 实施例9: 在D'D3表达FVIII/VWF DKO小鼠中的FVIII-XTEN变体细胞培养基浓缩物PK

[0677] 在D'D3表达FVIII/VWF DKO小鼠模型(于实施例2中描述)中评价D'D3片段延长FVII-XTEN的 $t_{1/2}$ 的能力。在该研究中,代替使用VWF-025以将D'D3二聚体引入到循环中,使用VWF-029构建体将D'D3单体引入循环中。为了制备FVIII-XTEN变体蛋白,在转染后第4天制备小规模(50-100mL)短暂转染培养基,收集细胞培养基并且浓缩以达到10-20IU/mL FVIII活性范围,这适用于PK研究。然后将经浓缩的细胞培养基用于在FVIII/VWF DKO小鼠中在循环中含有或不含D'D3的情况下的标准PK研究。

[0678] 在系统中测试共计6个含有1-3个XTEN插入的FVIII-XTEN变体,在表20中总结它们的 $t_{1/2}$ 并且在图9A中对来自代表性变体的数据进行作图。

[0679] 观察到在循环中存在D'D3片段下所有FVIII-XTEN变体的较长半衰期(表20),这证明D'D3保护FVIII-XTEN免遭其清除途径。此外,当相较于其在HemA小鼠中的14小时的半衰期时,LSD0055.021在D'D3表达DKO小鼠中具有20.4小时的 $t_{1/2}$ (图9B、表20),指示最终突破FVIII分子的2倍半衰期延长最高限度。通过进一步修饰FVIII(XTEN)/VWF分子,我们可能达到甚至更长的FVIII $t_{1/2}$,并且提供HemA患者仅需要每周一次或更不频繁的给药方案的FVIII蛋白。

[0680] 表20. 在D'D3表达FVIII/VWF DKO小鼠中的FVIII-XTEN $t_{1/2}$

[0681]

FVIII-XTEN ID	XTEN 插入编号	插入位点	XTEN 大小	t _{1/2} (小时) DKO 小鼠	t _{1/2} (小时) pLIVE-D'D3/DKO 小鼠	t _{1/2} (小时) HemA 小鼠
pSD-0013	1	CT	144	3.3	7.9	
LSD0003.009	2	B*/CT	144/288	9.7	16.4	
LSD0038.015	2	1656/26	144/144	7.8	17.2	
LSD0049.002	3	18/B*/CT	144/144/288	12.6	17.5	
LSD0051.002	3	403/B*/CT	144/144/288	11.1	19.9	
LSD0055.021	3	1900/B*/CT	144/144/288	16	20.4	14

[0682] *B指示XTEN序列(例如,144)紧邻对应于成熟FVIII序列的氨基酸残基745的下游插入。

[0683] 实施例10:含有VWF的FVIII变体和含有XTEN的FVIII变体在FVIII/VWF双敲除(DKO)血浆中的稳定性

[0684] 在FVIII/VWF双敲除(DKO)小鼠血浆中测试rFVIIIIFc蛋白变体的血浆稳定性。对于稳定性测定,用指导rFVIIIIFc或FVIII-169(288AE XTEN在B-结构域接合点插入的rFVIIIIFc)的质粒和指导IgG-Fc或VWF-031(融合于IgG-Fc的VWF D'D3区域)的质粒共转染HEK293细胞。在转染后第四天,收集细胞培养基并且基于FVIII显色活性浓缩至30IU/mL。然后将经浓缩的细胞培养基加入到DKO小鼠血浆中以得到5IU/mL的FVIII活性并且在37℃温育。在不同时间点收集等分试样以用于通过显色测定进行活性测量。一式两份测量在每个时间点的活性,并且将平均活性作为时间的函数作图。在DKO小鼠血浆中,其中重链和轻链通过非共价相互作用连在一起的FVIIIIFc(一种双链(dc)FVIII分子)的活性随时间下降(图10)。FVIII-169:Fc(其含有在B-结构域接合点处的288AE XTEN插入)的活性以相对于rFVIIIIFc减少的速率衰退,指示通过XTEN插入赋予增强的稳定性。假定VWF已被建议增强体内FVIII的稳定性,那么我们评价FVIII-169:VWF-031的血浆稳定性。这种异二聚体分子(其中FVIII元件和VWF D'D3元件融合于Fc的各自半结构域)相对于FVIII-169:Fc展现出另外的血浆稳定性,指示VWF D'D3结构域和XTEN对rFVIIIIFc的血浆稳定性具有协同效应。

[0685] 实施例11:Fc融合、XTEN插入和VWF的D'D3片段对FVIII半衰期的作用

[0686] 为了评估Fc融合、XTEN插入和VWF的D'D3片段对FVIII半衰期的作用,在FVIII/VWF双敲除(DKO)小鼠中评价B结构域缺失的重组FVIII(rBDD-FVIII)、rFVIIIIFc、FVIII-169:Fc和FVIII-169:VWF-031的药动学性质。

[0687] 通过单次静脉内施用200IU/kg的FVIII蛋白治疗DKO小鼠,并且在指定时间点收集血浆样本,如于图11中指示。通过FVIII显色测定分析血浆样本的FVIII活性并且使用WinNonlin-Phoenix程序计算半衰期。所测试分子的药动学参数于表21中列出。每个FVIII变体的血浆FVIII活性的时间回归曲线在图11中作图。

[0688] 未修饰的BDD-FVIII在DKO小鼠中具有0.23小时的半衰期,由于通过Fc:FcRn相互作用使FVIIIIFc蛋白再循环,FVIIIIFc融合蛋白在DKO小鼠中具有1.66小时的延长的半衰期。当将AEXTEN多肽的288个残基并入在FVIIIIFc分子内FVIII的B结构域区域时,在DKO小鼠中所得FVIII169/Fc蛋白质的半衰期进一步延长至7.41小时。最终,通过添加VWF的D'D3结构域,FVIII169/VWF031异二聚体的半衰期在DKO小鼠中已达到17.9小时(图11、表21)。除半衰期之外,所有其它PK参数在添加每个元件的情况下也按比例得以改善(表21)。FVIII可耐受多个半衰期延长元件,并且三个元件对FVIII半衰期延长的这种协同效应能够使得FVIII-

XTEN VWF异二聚体的半衰期进一步得以改善。

[0689] 表21:FVIII变体的PK参数

FVIII	FVIII 同工型	XTEN 插入		$T_{1/2}$ (小时)	MRT (小时)	Cl (mL/小时/kg)	Vss (mL/kg)	AUC _{D kg*} 小时/mL
		位点	XTEN 长度					
BDD-FVIII	dc			0.23	0.24	407.72	97.42	0.0025
FVIII-Fc	dc			1.66	2.06	62.66	128.82	0.0161
FVIII169/Fc	sc	B*	AE288	7.41	6.67	6.24	41.61	0.1603
FVIII169/VWF031	sc	B*	AE288	17.94	20.1	4.06	81.69	0.2463

[0691] *B指示XTEN序列(例如,144)紧邻对应于成熟FVIII序列的氨基酸残基745的下游插入。

[0692] 实施例12:不同FVIII-XTEN_VWF异二聚体的药动学性质

[0693] 为了评价VWF-D'D3片段和XTEN插入对FVIII半衰期的组合作用,在HemA小鼠中测试FVIII-XTEN-Fc:VWF-Fc异二聚体的药动学性质并且与BDD-FVIII(scBDD-FVIII)和FVIII-169:VWF-031的单链同工型的药动学性质(实施例10)相比较。产生七个新的FVIII-XTEN-Fc构建体(蛋白质序列于表24中列出)。那些构建体的示意图示于图14A-H中。FVIII-195和FVIII-199分别为FVIII双链和单链同工型,各含有在位置1900和1656处的两个XTEN插入。FVIII-196和FVIII-201分别为FVIII双链和单链同工型,各含有在位置26、1656和1900处的三个XTEN插入。FVIII-203、FVIII-204和FVIII-205为具有两个在B结构域接合点和分别在位置1900、403或18处的XTEN插入的sc-FVIII-Fc分子。每个FVIII-XTEN-Fc构建体与VWF-031在HEK293细胞中共表达以生成FVIII-XTEN-Fc/VWF异二聚体蛋白。在转染后四天,收集细胞培养基并且基于FVIII显色活性被浓缩至20IU/mL(FVIII-195:VWF-031、FVIII-196:VWF-031、FVIII-199:VWF-031、FVIII-203:VWF-031和FVIII-204:VWF-031)或被纯化(scBDD-FVIII、FVIII-169:VWF-031、FVIII-201:VWF-031和FVIII-205:VWF-031)。已证明通过在FVIII-XTEN-Fc:VWF-Fc异二聚体(FVIII-169:VWF-031,实施例5)中的D'D3片段来完全分子内掩蔽FVIII分子免遭内源性VWF,所以选择HemA小鼠用于PK评价。通过以200IU/10mL/kg的剂量静脉内施用,向8-12周龄HemA小鼠施用经纯化的蛋白质或经浓缩的细胞培养基。在给药后5分钟、8小时、16小时、24小时、32小时、48小时、72小时和96小时收集血浆样本。通过FVIII显色测定分析血浆样本的FVIII活性并且使用WinNonlin-Phoenix程序计算半衰期。所测试分子的药动学参数于表22中列出。将在选定时间点的FVIII-XTEN-Fc/VWF-Fc变体的血浆FVIII活性于图12A-C中作图。

[0694] 当XTEN插入到位置1900和1656(FVIII-195、FVIII-199)时,观察到scFVIII同工型(FVIII-199:VWF-031)的半衰期相较于FVIII-169:VWF-031得以适度改善。然而,dcFVIII同工型展现出比FVIII-169:VWF-031短的半衰期,指示单链同工型可能比相应的双链同工型显著更稳定(表22和图12A)。当第三XTEN插入在位置26并入FVIII-199时,所得分子FVIII-201:VWF-031的半衰期已达到24.6小时,其表示相对于scBDD-FVIII的超过三倍半衰期改善(表22和图12C)。我们也测试了各自与B结构域XTEN插入组合的在位置403(A2结构域)、1900(A3结构域)和18(A1结构域)处的第二XTEN插入的半衰期延长效应。尽管添加A2或A3XTEN插

入未赋予另外的半衰期益处(表22,图12b),但是添加A1插入则进一步延长FVIII-XTEN-Fc:VWF-Fc异二聚体的半衰期至29.4小时(表22,图12C),其比scBDD-FVIII延长超过三倍。

[0695] 当XTEN并入FVIII-Fc/VWF异二聚体构建体时,所得分子的半衰期改善的程度是可变的,并且在半衰期和XTEN插入的位点或数量之间没有观察到明显相关性,表明FVIII-XTEN-Fc/VWF异二聚体的半衰期通过整个分子的完整性而不是通过XTEN插入的数量或位置来确定。

[0696] 观察到的FVIII-XTEN-Fc:VWF-Fc异二聚体的24.6小时和29.4小时半衰期明显超过对FVIII半衰期延长的1.6至2倍限值。如果将这种发现转译给HemA患者,则为FVIII防治将允许每周一次或更不频繁的给药。

[0697] 表22:FVIII-XTEN-Fc/VWF-Fc异二聚体的PK参数

[0698]

FVIII	FVIII同工型	XTEN 插入		T _{1/2} (小时)	MRT (小时)	Cl (mL/小时/kg)	V _{ss} (mL/kg)	AUC D kg ⁻¹ 小时/mL
		位点	XTEN长度					
scBDD-FVIII	sc			7.16	10.16	4.38	44.44	0.23
FVIII169/VWF031	sc	B*	AE288	16.65	18.44	3.57	65.79	0.28
FVIII195/VWF031	dc	1656/1900	AG144/AE144	12.56	13.88	9.04	125.48	0.11
FVIII199/VWF031	sc	1656/1900	AG144/AE144	18.57	20.09	6.24	125.28	0.16
FVIII201/VWF031	sc	26/1656/1900	AG144/AG144/AE144	24.63	33.67	1.9	63.97	0.53
FVIII203/VWF031	sc	403/B*	AE144/AE288	15.52	18	3.65	65.61	0.27
FVIII204/VWF031	sc	1900/B*	AE144/AE288	16.3	20.63	2.87	59.14	0.35
FVIII205/VWF031	sc	18/B*	AE144/AE288	29.4	37.06	1.82	67.39	0.55

[0699] *B指示XTEN序列(例如,144)紧邻对应于成熟FVIII序列的氨基酸残基745的下游插入。

[0700] 除了将XTEN并入FVIII分子中之外,我们也评价并入XTEN作为D'D3和Fc片段之间的接头的潜在半衰期延长益处。FVIII-155(scFVIII-Fc)与VWF-034(具有AE 288XTEN加上35个残基的凝血酶可切割的接头的VWF-Fc)在HEK293细胞中共表达。在转染后第4天,收集细胞培养基并且基于FVIII活性测定被浓缩至20IU/mL。通过单次静脉内注射以200IU/10mL/kg用经浓缩的细胞培养基对FVIII/VWF DK0小鼠给药。在给药后5分钟、8小时、24小时、48小时、72小时和96小时收集血浆样本。通过FVIII显色测定分析血浆样本的FVIII活性,并且将血浆FVIII活性作为时间的函数的回归曲线作图(图13)。FVIII-155/VWF-034展现出与FVIII-169/VWF-031(其具有插入到FVIII的B结构域接合点的AE 288XTEN)相同的半衰期改善,如通过两个分子的重叠回归曲线所示(图13)。证明XTEN插入到VWF-Fc多肽赋予与通过在FVIII多肽的B结构域接合点处的XTEN插入赋予的半衰期改善类似的幅度的半衰期改善,这表明进一步半衰期改善可能在其中分子内XTEN插入到FVIII多肽与在VWF-Fc多肽的VWF

和Fc元件之间的结构域间XTEN插入相组合的异二聚体分子中是可能的。

[0701] 实施例13A: 另外的FVIII-XTEN_VWF异二聚体的药动学性质

[0702] 除了表22中列出的FVIII-XTEN VWF异二聚体之外, 含有XTEN插入、FVIII的单链和双链形式的不同组成的FVIII-XTEN VWF异二聚体(表23A)在HemA中被测试或将被测试它们的药动学性质。下面也公开了各种FVIII构建体(表23B)和VWF构建体(表23C)。通过以200IU/10mL/kg的单一剂量静脉内施用异二聚体蛋白质来治疗HemA小鼠。然后将在给药后5分钟、24、48、72、96和120小时收集血浆样本。将通过FVIII显色测定来分析血浆样本的FVIII活性并且将使用WinNonlin-Phoenix程序计算半衰期。所列出的异二聚体的蛋白质序列于表25中列出。

[0703] 表23A. 合理的用于PK和活性改善的FVIII-XTEN-Fc:VWF-Fc异二聚体组合。

	pSYN VWF-015	pSYN VWF-031	pSYN VWF-034 **	pSYN VWF-036
pSYN FVIII 010	-	$t_{1/2}$ 8.7 小时 DKO 小鼠	待测试	-
pSYN FVIII 155	$t_{1/2}$ 6.3 小时 DKO 小鼠	$t_{1/2}$ 10.8 小时 HemA 小鼠	$t_{1/2}$ 18.6 小时 HemA 小鼠	$t_{1/2}$ 13.3 小时 HemA 小鼠
pSYN FVIII 169 **	-	$t_{1/2}$ 16.7 小时 HemA 小鼠	$t_{1/2}$ 31.1 小时 HemA 小鼠	-
pSYN FVIII 173 **	-	$t_{1/2}$ 15.2 小时 DKO 小鼠	$t_{1/2}$ 28.9 小时 HemA 小鼠	待测试
pSYN FVIII 205	-	$t_{1/2}$ 29.4 小时 HemA 小鼠	$t_{1/2}$ 32.4 小时 HemA 小鼠	$t_{1/2}$ 29.7 小时 HemA 小鼠
pSYN FVIII 266	-	$t_{1/2}$ 24.5 小时 HemA 小鼠	$t_{1/2}$ 27.4 小时 HemA 小鼠	-
pSYN FVII 267	-	$t_{1/2}$ 23.0 小时 HemA 小鼠	$t_{1/2}$ 25.7 小时 HemA 小鼠	
pSYN FVIII 268	-	待测试	待测试	待测试
pSYN FVIII 268 的双链同 工型		待测试	待测试	待测试
** XTEN 的长度可被改变至 72、144、288、324、333、576 或 864。				

[0705] 表23B.FVIII构建体:

[0706]

pSYN FVIII 010	双链 FVIII _{FC}
pSYN FVIII 169	具有 B-结构域中的 288 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 173	具有 B-结构域中的 288 AE XTEN 的双链 FVIII _{FC}
pSYN FVIII 195	具有在氨基酸 1656 和 1900 处的两个 144 XTEN 的双链 FVIII _{FC}
pSYN FVIII 196	具有在氨基酸 26、1656 和 1900 处的三个 144 XTEN 的双链 FVIII _{FC}
pSYN FVIII 199	具有在氨基酸 1656 和 1900 处的两个 144 XTEN 的单链 FVIII _{FC}
pSYN FVIII 201	具有在氨基酸 26、1656 和 1900 处的三个 144 XTEN 的单链 FVIII _{FC}
pSYN FVIII 203	具有在氨基酸 1900 处的 144 AE XTEN 和 B-结构域中的 288 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 204	具有在氨基酸 403 处的 144 AE XTEN 和 B-结构域中的 288 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 205	具有在氨基酸 18 处的 144 AE XTEN 和 B-结构域中的 288 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 207	单链 FVIII (无 Fc, 无 XTEN)
pSYN FVIII 266	具有在氨基酸 18 处的 42 AE XTEN 和 B-结构域中的 288 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 267	具有在氨基酸 18 处的 72 AE XTEN 和 B-结构域中的 288 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 268	具有在氨基酸 18 处的 144 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 269	具有在氨基酸 18 处的 72 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 271	具有在氨基酸 18 处的 42 AE XTEN 的单链 FVIII _{FC}
pSYN FVIII 272	具有在氨基酸 18 处的 144 AE XTEN 和 B-结构域中的 288 AE XTEN 的单链 FVIII (无 Fc)

[0707] 表23C.VWF构建体:

[0708]

pSYN VWF031	具有 C1099A/C1142A 的 VWF-D1D2D'D3- 48aa 长的凝血酶可切割的 GS 接头-Fc
pSYN VWF034	具有 C1099A/C1142A 的 VWF-D1D2D'D3- 288AE XTEN +35aa 长的凝血酶可切割的 GS 接头-Fc
pSYN VWF035	具有 C1099A/C1142A 的 VWF-D1D2D'D3- 72aa 长的凝血酶可切割的 GS 接头-Fc
pSYN VWF036	具有 C1099A/C1142A 的 VWF-D1D2D'D3- 98aa 长凝血酶可切割的 GS 接头-Fc
pSYN VWF041	具有 D3 中的 288 AE XTEN 和在具有 C1099A/C1142A 的 D3-Fc 之后 48aa 长的凝血酶可切割的 GS 接头的 VWF-D1D2D'D3

[0709] 实施例13B: 另外的FVIII-XTEN_VWF异二聚体的药动学性质

[0710] FVIII-XTEN_VWF异二聚体在HemA小鼠中被测试它们的药动学性质。所测试的异二聚体为FVIII169/VWF034、FVIII205/VWF034、FVIII205/VWF036和FVIII266/VWF031。用200IU/10mL/kg的单次静脉内剂量的各种异二聚体蛋白向HemA小鼠施用。在给药后5分钟、24、48、72、96和120小时收集血浆样本。通过FVIII显色测定分析血浆样本的FVIII活性,并且使用WinNonlin-Phoenix程序计算半衰期。PK结果在下面表24中示出。

[0711] 表24. 在HemA小鼠中另外的FVIII-XTEN_VWF-PK

[0712]

治疗	5 分钟 回收率 (%)	HL (小时)	MRT (小时)	Cl(mL/ 小时/ kg)	Vss (mL/kg)	AUC_D (小时*kg *mIU/ mL/mIU)	t _{1/2} 增加的 倍数对 scBDD- FVIII
ScBDD- FVIII		7.16	10.16	4.83	44.44	0.23	-
FVIII169/V WF034	76	31.1	34.57	1.73	59.77	0.58	4.3
FVIII205/V WF034	68	32.41	39.79	1.55	61.73	0.64	4.6
FVIII205/V WF036	74	29.71	36.35	1.61	58.43	0.62	4.1
FVIII266/V WF031	66	24.45	22.75	2.67	60.83	0.37	3.4

[0713] pSYNFVIII010核苷酸序列-(双链FVIIIIFc) (SEQ ID NO:125)

[0714]

```
1   ATGCAAATAG AGCTCTCCAC CTGCTTCTTT CTGTGCCTTT
   TGCGATTCTG
51  CTTTAGTGCC ACCAGAAGAT ACTACCTGGG TGCAGTGGAA
   CTGTCATGGG
101 ACTATATGCA AAGTGATCTC GGTGAGCTGC CTGTGGACGC
   AAGATTTTCT
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 TGGAAATGGGC
   ACCACTCCTG
851 AAGTGCACTC AATATTCCTC GAAGGTCACA CATTTCTTGT
   GAGGAACCAT
901 CGCCAGGCGT CCTTGGAAT CTGCGCAATA ACTTTCCTTA
   CTGCTCAAAC
951 ACTCTTGATG GACCTTGAC AGTTTCTACT GTTTTGTCAT
   ATCTCTTCCC
1001 ACCAACATGA TGGCATGGAA GCTTATGTCA AAGTAGACAG
   CTGTCCAGAG
1051 GAACCCCAAC TACGAATGAA AAATAATGAA GAAGCGGAAG
   ACTATGATGA
1101 TGATCTTACT GATTCTGAAA TGGATGTGGT CAGGTTTGAT
   GATGACAACT
```

[0715]

1151 CTCCTTCCTT TATCCAAATT CGCTCAGTTG CCAAGAAGCA
TCCTAAAACT
1201 TGGGTACATT ACATTGCTGC TGAAGAGGAG GACTGGGACT
ATGCTCCCTT
1251 AGTCCTCGCC CCCGATGACA GAAGTTATAA AAGTCAATAT
TTGAACAATG
1301 GCCCTCAGCG GATTGGTAGG AAGTACAAAA AAGTCCGATT
TATGGCATACT
1351 ACAGATGAAA CCTTTAAGAC TCGTGAAGCT ATTCAGCATG
AATCAGGAAT
1401 CTTGGGACCT TTACTTTATG GGGGAAGTTGG AGACACACTG
TTGATTATAT
1451 TTAAGAATCA AGCAAGCAGA CCATATAACA TCTACCCTCA
CGGAATCACT
1501 GATGTCCGTC CTTTGTATTC AAGGAGATTA CCAAAGGTG
TAAAACATTT
1551 GAAGGATTTT CCAATTCTGC CAGGAGAAAT ATTCAAATAT
AAATGGACAG
1601 TGACTGTAGA AGATGGGCCA ACTAAATCAG ATCCTCGGTG
CCTGACCCGC
1651 TATTACTCTA GTTTCGTTAA TATGGAGAGA GATCTAGCTT
CAGGACTCAT
1701 TGGCCCTCTC CTCATCTGCT ACAAAGAATC TG TAGATCAA
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 GAGAACTGT
CTTCATGTCG
2101 ATGGAAAACC CAGGTCTATG GATTCTGGGG TGCCACAAC
CAGACTTTTCG
2151 GAACAGAGGC ATGACCGCCT TACTGAAGGT TTCTAGTTGT
GACAAGAACA
2201 CTGGTGATTA TTACGAGGAC AGTTATGAAG ATATTTTCAGC
ATACTTGCTG
2251 AGTAAAAACA ATGCCATTGA ACCAAGAAGC TTCTCTCAA
ACCCACCAGT
2301 CTTGAAACGC CATCAACGGG AAATAACTCG TACTACTCTT
CAGTCAGATC
2351 AAGAGGAAAT TGACTATGAT GATACCATAT CAGTTGAAAT
GAAGAAGGAA

[0716]

2401 GATTTTGACA TTTATGATGA GGATGAAAAT CAGAGCCCCC
GCAGCTTTCA
2451 AAAGAAAACA CGACACTATT TTATTGCTGC AGTGGAGAGG
CTCTGGGATT
2501 ATGGGATGAG TAGCTCCCCA CATGTTCTAA GAAACAGGGC
TCAGAGTGGC
2551 AGTGTCCTC AGTTCAAGAA AGTTGTTTTT CAGGAATTTA
CTGATGGCTC
2601 CTTTACTCAG CCCTTATACC GTGGAGAACT AAATGAACAT
TTGGGACTCC
2651 TGGGGCCATA TATAAGAGCA GAAGTTGAAG ATAATATCAT
GGTAACTTTC
2701 AGAAATCAGG CCTCTCGTCC CTATTCCTTC TATTCTAGCC
TTATTTCTTA
2751 TGAGGAAGAT CAGAGGCAAG GAGCAGAACC TAGAAAAAAC
TTTGTCAAGC
2801 CTAATGAAAC CAAAACCTAC TTTTGAAAG TGCAACATCA
TATGGCACCC
2851 ACTAAAGATG AGTTTGACTG CAAAGCCTGG GCTTATTTCT
CTGATGTTGA
2901 CCTGGAAAAA GATGTGCACT CAGGCCTGAT TGGACCCCTT
CTGGTCTGCC
2951 ACACTAACAC ACTGAACCCT GCTCATGGGA GACAAGTGAC
AGTACAGGAA
3001 TTTGCTCTGT TTTTCACCAT CTTTGATGAG ACCAAAAGCT
GGTACTTCAC
3051 TGAAAATATG GAAAGAACT GCAGGGCTCC CTGCAATATC
CAGATGGAAG
3101 ATCCCACCTT TAAAGAGAAT TATCGCTTCC ATGCAATCAA
TGGCTACATA
3151 ATGGATACAC TACCTGGCTT AGTAATGGCT CAGGATCAAA
GGATTGATG
3201 GTATCTGCTC AGCATGGGCA GCAATGAAAA CATCCATTCT
ATTCATTTCA
3251 GTGGACATGT GTTCACTGTA CGAAAAAAG AGGAGTATAA
AATGGCACTG
3301 TACAATCTCT ATCCAGGTGT TTTTGAGACA GTGGAAATGT
TACCATCCAA
3351 AGCTGGAATT TGGCGGGTGG AATGCCTTAT TGGCGAGCAT
CTACATGCTG
3401 GGATGAGCAC ACTTTTTCTG GTGTACAGCA ATAAGTGTC
GACTCCCCTG
3451 GGAATGGCTT CTGGACACAT TAGAGATTTT CAGATTACAG
CTTCAGGACA
3501 ATATGGACAG TGGGCCCCAA AGCTGGCCAG ACTTCATTAT
TCCGGATCAA
3551 TCAATGCCTG GAGACCAAG GAGCCCTTTT CTTGGATCAA
GGTGGATCTG
3601 TTGGCACCAA TGATTATTCA CGGCATCAAG ACCCAGGGTG
CCCGTCAGAA

[0717]

```
3651 GTTCTCCAGC CTCTACATCT CTCAGTTTAT CATCATGTAT
    AGTCTTGATG
3701 GGAAGAAGTG GCAGACTTAT CGAGGAAATT CCACTGGAAC
    CTTAATGGTC
3751 TTCTTTGGCA ATGTGGATTC ATCTGGGATA AAACACAATA
    TTTTAAACCC
3801 TCCAATTATT GCTCGATACA TCCGTTTGCA CCCAACTCAT
    TATAGCATTC
3851 GCAGCACTCT TCGCATGGAG TTGATGGGCT GTGATTTAAA
    TAGTTGCAGC
3901 ATGCCATTGG GAATGGAGAG TAAAGCAATA TCAGATGCAC
    AGATTACTGC
3951 TTCATCCTAC TTTACCAATA TGTTTGCCAC CTGGTCTCCT
    TCAAAAGCTC
4001 GACTTCACCT CCAAGGGAGG AGTAATGCCT GGAGACCTCA
    GGTGAATAAT
4051 CCAAAAGAGT GGCTGCAAGT GGAATTCCAG AAGACAATGA
    AAGTCACAGG
4101 AGTAACTACT CAGGGAGTAA AATCTCTGCT TACCAGCATG
    TATGTGAAGG
4151 AGTTCCTCAT CTCCAGCAGT CAAGATGGCC ATCAGTGGAC
    TCTCTTTTTT
4201 CAGAATGGCA AAGTAAAGGT TTTTCAGGGA AATCAAGACT
    CCTTCACACC
4251 TGTGGTGAAC TCTCTAGACC CACCGTTACT GACTCGCTAC
    CTTCGAATTC
4301 ACCCCCAGAG TTGGGTGCAC CAGATTGCCC TGAGGATGGA
    GGTTCCTGGGC
4351 TGCGAGGCAC AGGACCTCTA CGACAAAAC CACACATGCC
    CACCGTGCCC
4401 AGCTCCAGAA CTCCTGGGCG GACCGTCAGT CTTCTCTTTC
    CCCCCAAAC
4451 CCAAGGACAC CCTCATGATC TCCCGGACCC CTGAGGTCAC
    ATGCGTGGTG
4501 GTGGACGTGA GCCACGAAGA CCCTGAGGTC AAGTTCAACT
    GGTACGTGGA
4551 CGGCGTGGAG GTGCATAATG CCAAGACAAA GCCGCGGGAG
    GAGCAGTACA
4601 ACAGCACGTA CCGTGTGGTC AGCGTCCTCA CCGTCCTGCA
    CCAGGACTGG
4651 CTGAATGGCA AGGAGTACAA GTGCAAGGTC TCCAACAAAG
    CCCTCCCAGC
4701 CCCCATCGAG AAAACCATCT CCAAAGCCAA AGGGCAGCCC
    CGAGAACCAC
4751 AGGTGTACAC CCTGCCCCCA TCCCGGGATG AGCTGACCAA
    GAACCAGGTC
4801 AGCCTGACCT GCCTGGTCAA AGGCTTCTAT CCCAGCGACA
    TCGCCGTGGA
4851 GTGGGAGAGC AATGGGCAGC CGGAGAACAA CTACAAGACC
    ACGCCTCCCC
```

4901 TGTGACTC CGACGGCTCC TTCTTCCTCT ACAGCAAGCT
CACCGTGGAC
[0718] 4951 AAGAGCAGGT GGCAGCAGGG GAACGTCTTC TCATGCTCCG
TGATGCATGA
5001 GGCTCTGCAC AACCCTACA CGCAGAAGAG CCTCTCCCTG
TCTCCGGGTA
5051 AATGA

[0719] pSYNFVIII010蛋白序列-(双链FVIIIIFc) (SEQ ID NO:126)

1 MQIELSTCFF LCLLRFCFSA TRRYYLGA VE LSWDYMQSDL
GELPVDARFP
51 PRVPKSFPPN TSVVYKKTLE VEFTDHLFNI AKPRPPWMGL
LGPTIQAEVY
101 DTVVITLKNM ASHPVSLHAV GVS YWKASEG AEYDDQTSQR
EKEDDKVFP
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 LLYGEVGD TL LIIFKNQASR
PYNIYPHGIT
[0720] 501 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
551 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE
601 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
651 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVYEDTLTLF
PFSGETVFMS
701 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYIED
SYEDISAYLL
751 SKNNAIEPRS FSQNPPVLKR HQREITRTTL QSDQEEIDYD
DTISVEMKKE
801 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP
HVLNRNAQSG
851 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA
EVEDNIMVTF
901 RNQASRPYSF YSSLISYEED QRQGAEPKRN FVKPNETKTY
FWKVQHHMAP
951 TKDEFDCKAW AYFSDVDLEK DVHSGLIGPL LVCHTNTLNP
AHGRQVTVQE
1001 FALFFTIFDE TKS WYFTENM ERNCRAPCNI QMEDPTFKEN
YRFHAINGYI

```

1051 MDTLPGLVMA QDQIRIRWYLL SMGSNENIHS IHFSGHVFTV
    RKKEEYKMAL
1101 YNLYPGVFET VEMLPSKAGI WRVECLIGEH LHAGMSTLFL
    VYSNKCQTPL
1151 GMASGHIRDF QITASGQYQG WAPKLARLHY SGSINAWSTK
    EPFSWIKVDL
1201 LAPMIIHGK TQGARQKFSS LYISQFIIMY SLDGKKWQTY
    RGNSTGTLMV
1251 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
    LMGCDLNSCS
1301 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
    SNAWRPQVNN
[0721] 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 TPPVLDSGGS
    FFLYSKLTVD
1651 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK*

```

[0722] 实施例14:在血友病A小鼠中具有大于三倍半衰期延长的一类新的凝结因子VIII分子

[0723] 一类新的FVIII分子被设计成含有两个多肽;一个多肽由具有在FVIII序列内的一个或多个位置处的XTEN插入的单链B-结构域缺失的(BDD)FVIII组成,而一个多肽由VWF的D'D3区域组成。每个多肽也重组融合于IgG1的Fc区以使正确对齐的D'D3区域能够结合FVIII部分。通过短暂转染在HEK 293细胞中表达所得FVIII变体,并且从条件培养基纯化。通过FVIII显色测定评价FVIII活性并且在FVIII敲除(HemA)和FVIII/VWF双敲除(DKO)小鼠两者中评估药动学性质。

[0724] 将XTEN和VWF的D'D3区域并入rFVIII中导致融合蛋白从内源性VWF清除的解偶联的同时延长它们的循环半衰期。这种融合构型中的FVIII完全掩蔽与VWF相互作用,如通过生物层干涉测量法(Octet)分析所测量的。与此一致的是,它们在HemA和DKO小鼠中的药动学曲线发现是相同的,指示它们在小鼠中的清除速率有效地从VWF分离。对XTEN长度和用于插入XTEN的位置的最优化鉴定了已超过VWF限制(16小时)的蛋白质亚组,在HemA小鼠中达到高达30小时的循环半衰期,表示比BDD-FVIII改善4倍。重要的是,这些蛋白质保持它们的官能度,如通过FVIII显色测定来判断。

[0725] VWF依赖性已设定治疗性FVIII的半衰期的基本限制。从VWF清除途径解偶联FVIII的同时通过融合VWF的D'D3区域和XTEN来延长半衰期会产生具有4倍半衰期延长的新FVIII分子。这是已超过通过全行业在开发长效FVIII方面的努力观察到的半衰期限制的工程改造的FVIII的第一篇报导,表示防治性治疗血友病A的潜在显著进步。

[0726] 表25:FVIII-XTEN-Fc和VWF-Fc构建体的蛋白质序列

[0727] FVIII 195蛋白序列(具有在氨基酸 1656和1900处的两个144 AE XTEN 的双链 FVIIIFc) (SEQ ID NO:105)

1 MQIELSTCFF LCLLRFCFSA TRRYLGAVE LSWDYMQSDL
GELPVDARFP
51 PRVPKSFFFN TSVVYKKTLE VEFTDHLFNI AKPRPPWMGL
LGPTIQAEVY
101 DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR
EKEDDKVFP
151 GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG
LIGALLVCRE
201 GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD
AASARAWPKM
251 HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL
EGHTFLVRNH
[0728] 301 RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME
AYVKVDSCPE
351 EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI
RSVAKKHPKT
401 WWHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR
KYKKVRFMAY
451 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR
PYNIPHGIT
501 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
551 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE

[0729]

601 NRSWYLTEINI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
651 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVYEDTLTLF
PFSGETVFMS
701 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYED
SYEDISAYLL
751 SKNNAIEPRS FSQNPPVLKR HQREITRTTL QGAPGTPGSG
TASSSPGASP
801 GTSSTGSPGA SPGTSSTGSP GASPGTSSTG SPGSSPSAST
GTGPGTPGSG
851 TASSSPGASP GTSSTGSPGA SPGTSSTGSP GASPGTSSTG
SPGSSTPSGA
901 TGSPGSSTPS GATGSPGASP GTSSTGSPAS SSDQEEIDYD
DTISVEMKKE
951 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP
HVLNRNAQSG
1001 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA
EVEDNIMVTF
1051 RNQASRPYSF YSSLISYEED QRQGAEPKRN FVKPNETKTY
FWKVQHHMAP
1101 TKDEFDCKAW AYFSDVDLEK DVHSGLIGPL LVCHTNLNP
AHGRQVTVQE
1151 FALFFTIFDE TKSIFYTENM ERNCRGAPTS ESATPESGPG
SEPATSGSET
1201 PGTSESATPE SGPGSEPATS GSETPGTSES ATPESGPGTS
TEPSEGSAPG
1251 TSESATPESG PGSPAGSPTS TEEGSPAGSP TSTEEGSPAG
SPTSTEEGTS
1301 ESATPESGPG TSTEPSEGSA PGASSAPCNI QMEDPTFKEN
YRFHAINGYI
1351 MDTLPGLVMA QDQIRIRWYLL SMGSNENIHS IHFSGHVFTV
RKKEEYKMAL
1401 YNLYPGVFET VEMPLPSKAGI WRVECLIGEH LHAGMSTLFL
VYSNKCQTPL
1451 GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK
EPFSWIKVDL
1501 LAPMIIHGK TQGARQKFSS LYISQFIIMY SLDGKKWQTY
RGNSTGTLMV
1551 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
LMGCDLNSCS
1601 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
SNAWRPQVNN
1651 PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISS
QDGHQWTLFF
1701 QNGKVKVFQG NQDSFTPVVN SLDPPLLTRY LRIHPQSWVH
QIALRMEVLG
1751 CEAQDLYDKT HTCPPCPAPE LLGGPSVFLF PPKPKDTLMI
SRTPEVTCVV
1801 VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
SVLTVLHQDW

- 1851 LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP
SRDELTKNQV
- [0730] 1901 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGGS
FFLYSKLTVD
- 1951 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK*
- [0731] FVIII 196蛋白序列 (具有在氨基酸26、1656和1900处的三个144 AE XTEN的双链 FVIIIIFc) (SEQ ID NO:106)

1 MQIELSTCFF LCLLRFCFSA TRRYYLGAWE LSWDYMQSDL
GELPVGAPGS

51 SPSASTGTGP GSSPSASTGT GPGASPGTSS TGSPGASPGT
SSTGSPGSST

101 PSGATGSPGS SPSASTGTGP GASPGTSSTG SPGSSPSAST
GTGPGTPGSG

151 TASSSPGSST PSGATGSPGS STPSGATGSP GASPGTSSTG
SPASSDARFP

201 PRVPKSFPPN TSVVYKKTLE VEFTDHLFNI AKPRPPWMGL
LGPTIQAEVY

251 DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR
EKEDDKVFPG

301 GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG
LIGALLVCRE

351 GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD
AASARAWPKM

401 HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL
EGHTFLVRNH

451 RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME
AYVKVDSCPE

[0732] 501 EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI
RSVAKKHPKT

551 WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR
KYKKVRFMAY

601 TDETFKTREA IQHESGILGP LLYGEVGDIT LIIFKNQASR
PYNIIYPHGIT

651 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR

701 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE

751 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL

801 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEDTLTLF
PFSGETVFMS

851 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED
SYEDISAYLL

901 SKNNAIEPRS FSQNPPVLKR HQREITRTTL QGAPGTPGSG
TASSSPGASP

951 GTSSTGSPGA SPGTSSTGSP GASPGTSSTG SPGSSPSAST
GTGPGTPGSG

1001 TASSSPGASP GTSSTGSPGA SPGTSSTGSP GASPGTSSTG
SPGSSTPSGA

1051 TGSPGSSTPS GATGSPGASP GTSSTGSPAS SSDQEEIDYD
DTISVEMKKE
1101 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP
HVLNRNAQSG
1151 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA
EVEDNIMVTF
1201 RNQASRPYSF YSSLISYEED QRQGAEPRKN FVKPNETKTY
FWKVQHMAP
1251 TKDEFDCKAW AYFSDVDLEK DVHSGLIGPL LVCHTNTLNP
AHGRQVTVQE
1301 FALFFTIFDE TKSIFYTENM ERNCRGAPTS ESATPESGPG
SEPATSGSET
1351 PGTSESATPE SGPGSEPATS GSETPGTSES ATPESGPGTS
TEPSEGSAAG
1401 TSESATPESG PGSPAGSPTS TEEGSPAGSP TSTEESGPA
SPTSTEEGTS
1451 ESATPESGPG TSTEPSEGSA PGASSAPCNI QMEDPTFKEN
YRFHAINGYI
1501 MDTLPLVMA QDQIRWYLL SMGSNENIHS IHFSGHVFTV
RKKEEYKMA
1551 YNLYPGVFET VEMLPSKAGI WRVECLIGEH LHAGMSTLFL
VYSNKCQTPL
1601 GMASGHIRDF QITASGQYQG WAPKLARLHY SGSINAWSTK
EPFSWIKVDL
1651 LAPMIIHGK TQGARQKFSS LYISQFIIMY SLDGKKWQTY
RGNSTGTLMV
1701 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
LMGCDLNSCS
1751 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
SNAWRPQVNN
1801 PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISS
QDGHQWTLFF
1851 QNGKVVFQNG QNDSFTPVVN SLDPPLLTRY LRIHPQSWVH
QIALRMEVLG
1901 CEAQDLYDKT HTCPCPAPE LLGGPSVFLF PPKPKDTLMI
SRTPEVTCVV
1951 VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
SVLTVLHQDW
2001 LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP
SRDELTKNQV
2051 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGGS
FFLYSKLTVD
2101 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK*

[0734] FVIII 199蛋白序列(具有在氨基酸1656和1900处的三个144 AE XTEN的单链 FVIIIIfc) (SEQ ID NO:107)

[0735]

1 MQIELSTCFF LCLLRFCFSA TRRYYLGAWE LSWDYMQSDL
GELPVDARFP
51 PRVPKSFPPN TSVVYKKTLE 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 TDETFTKTREA IQHESGILGP LLYGEVGDIT LIIFKNQASR
PYNIIYPHGIT
501 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
551 YSSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE
601 NRSWYLTENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
651 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEYEDTLTLF
PFSGETVFMS
701 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED
SYEDISAYLL
751 SKNNAIEPRS FSQNPPVLKR HQAEITRTTL QGAPGTPGSG
TASSSPGASP
801 GTSSTGSPGA SPGTSSTGSP GASPSTGSPG SPGSSPSAST
GTGPGTPGSG
851 TASSSPGASP GTSSTGSPGA SPGTSSTGSP GASPSTGSPG
SPGSSPSAST
901 TGSPGSPGAS GATGSPGASP GTSSTGSPAS SSDQEEIDYD
DTISVEMKKE
951 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP
HVLNRNAQSG
1001 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA
EVEDNIMVTF
1051 RNQASRPYSF YSSLISYEED QRQGAEPKRN FVKPNETKTY
FWKVQHMAP
1101 TKDEFDCAW AYFSDVDLEK DVHSGGIGPL LVCHTNTLNP
AHGRQVTVQE
1151 FALFFTIFDE TKSIFYTENM ERNCRGAPTS ESATPESGPG
SEPATSGSET
1201 PGTSESATPE SGPGSEPATG GSETPGTSES ATPESGPGTS
TEPSEGSAAG
1251 TSESATPESG PGSPAGSPTS TEEGSPAGSP TSTEEGSPAG
SPTSTEEGTS
1301 ESATPESGPG TSTEPSEGSA PGASSAPCNI QMEDPTFKEN
YRFHAINGYI

1351 MDTLPGLVMA QDQRIRWYLL SMGSNENIHS IHFSGHVFTV
 RKKEEYKMAL
 1401 YNLYPGVFET VEMLPSKAGI WRVECLIGEH LHAGMSTLFL
 VYSNKCQTPL
 1451 GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK
 EPFSWIKVDL
 1501 LAPMIIHGIK TQGARQKFSS LYISQFIIMY SLDGKKWQTY
 RGNSTGTLMV
 1551 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
 LMGC DLNSCS
 1601 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
 SNAWRPQVNN
 [0736] 1651 PKEWLQVDFQ KTMKVTGVT T QGVKSLLTSM YVKEFLISS
 QDGHQWTLFF
 1701 QNGKVKVFQ NQDSFTP VVN SLDPPLLTRY LRIHPQSWVH
 QIALRMEVLG
 1751 CEAQDLYDKT HTPPCPAPE LLGGPSVFLF PPKPKDTLMI
 SRTPEVTCVV
 1801 VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
 SVLTVLHQDW
 1851 LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP
 SRDELTKNQV
 1901 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDS DGS
 FFLYSKLTVD
 1951 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK*

[0737] FVIII 201蛋白序列(具有在氨基酸26、1656和1900处的三个144AE XTEN的单链 FVIIIIFc) (SEQ ID NO:108)

1 MQIELSTCFF LCLLRFCFSA TRRYYLGA VE LSWDYMQSDL
 GELPVGAPGS
 51 SPSASTGTGP GSSPSASTGT GPGASPGTSS TGSPGASPGT
 SSTGSPGSST
 101 PSGATGSPGS SPSASTGTGP GASPGTSSTG SPGSSPSAST
 GTGPGTPGSG
 151 TASSSPGSST PSGATGSPGS STPSGATGSP GASPGTSSTG
 SPASSDARFP
 201 PRVPKSFPFN TSVVYKKTLE VEFTDHLFNI AKPRPPWMGL
 LGPTIQAEVY
 [0738] 251 DTVVITLKNM ASHPVSLHAV GVS YWKASEG AEYDDQTSQR
 EKEDDKVFPG
 301 GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG
 LIGALLVCRE
 351 GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD
 AASARAWPKM
 401 HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL
 EGHTFLVRNH
 451 RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME
 AYVKVDSCPE
 501 EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI
 RSVAKKHPKT

[0739]

551 WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR
KYKKVRFMAY
601 TDETFKTREA IQHESGILGP LLYGEVGD TL LIIFKNQASR
PYNIYPHGIT
651 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
701 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE
751 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
801 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEDTLTLF
PFSGETVFM S
851 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYIED
SYEDISAYLL
901 SKNNAIEPRS FSQNPPVLKR HQAEITRTTL QGAPGTPGSG
TASSSPGASP
951 GTSSTGSPGA SPGTSSTGSP GASPGTSSTG SPGSSPSAST
GTGPGTPGSG
1001 TASSSPGASP GTSSTGSPGA SPGTSSTGSP GASPGTSSTG
SPGSSTPSGA
1051 TGSPGSSTPS GATGSPGASP GTSSTGSPAS SSDQEEIDYD
DTISVEMKKE
1101 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP
HVLNRNAQSG
1151 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA
EVEDNIMVTF
1201 RNQASRPYSF YSSLISYEED QRQGAEP RKN FVKPNETKTY
FWKVQH HMAP
1251 TKDEFDC KAW AYFSDVDLEK DVHSGLIGPL LVCHTNTLNP
AHGRQVTVQE
1301 FALFFTIFDE TKS WYFTENM ERNCRGAPTS ESATPESGPG
SEPATSGSET
1351 PGTSESATPE SGPGSEPAT S GSETPGTSES ATPESGPGTS
TEPSEG SAPG
1401 TSESATPESG PGSPAGSPTS TEEGSPAGSP TSTEEGSPAG
SPTSTEEGTS
1451 ESATPESGPG TSTEPSEGSA PGASSAPCNI QMEDPTFKEN
YRFHAINGYI
1501 MDTLPGLVMA QDQIRIRWYLL SMGSNENIHS IHFSGHVFTV
RKKEEYK MAL
1551 YNLYPGVFET VEMLP SKAGI WRVECLIGEH LHAGMSTLFL
VYSNKCQTPL
1601 GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK
EPFSWIKVDL
1651 LAPMIIHG IKTQ GARQKFSS LYISQFIIMY SLDGKKWQTY
RGNSTGTLMV
1701 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
LMGCDLNSCS
1751 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
SNAWRPQVNN

- 1801 PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISSS
QDGHQWTLFF
1851 QNGKVVFQG NQDSFTPVVN SLDPPLLTRY LRIHPQSWVH
QIALRMEVLG
1901 CEAQDLYDKT HTCPCCPAPE LLGGPSVFLF PPKPKDTLMI
SRTPEVTCVV
[0740] 1951 VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
SVLTVLHQDW
2001 LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP
SRDELTKNQV
2051 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGDS
FFLYSKLTVD
2101 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK*
- [0741] FVIII 203蛋白序列 (具有两个AE XTEN的单链FVIIIIFc; 一个为B-结构域中的
288AE XTEN且一个为在氨基酸1900处的144AEXTEN) (SEQ ID NO:109)

- 1 MQIELSTCFF LCLLRFCFSA TRRYYLGAVE LSWDYMQSDL
GELPVDARFP
51 PRVPKSFPFN TSVVYKKTLE 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
[0742] 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 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
651 HEVAYWYILS IGAQTDFLSV FFSGYTFFKHK MVYEDTLTLF
PFSGETVFMS
701 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED
SYEDISAYLL
751 SKNNAIEPRS FSQNGAPGTS ESATPESGPG SEPATSGSET
PGTSESATPE
801 SGPGSEPATS GSETPGTSES ATPESGPGTS TEPSEGSAPG
SPAGSPTSTE

851 EGTSESATPE SGPGSEPATS GSETPGTSES ATPESGPGSP
 AGSPTSTEEG
 901 SPAGSPTSTE EGTSTEPSEG SAPGTSESAT PESGPGTSES
 ATPESGPGTS
 951 ESATPESGPG SEPATSGSET PGSEPATSGS ETPGSPAGSP
 TSTEEGTSTE
 1001 PSEGSAPGTS TEPSEGSAPG SEPATSGSET PGTSESATPE
 SGPGTSTEPS
 1051 EGSAPASSPP VLKRHQAEIT RTTLQSDQEE IDYDDTISVE
 MKKEDFDIYD
 1101 EDENQSPRSF QKKTRHYFIA AVERLWDYGM SSSPHVLRNR
 AQSGSVPQFK
 1151 KVVVFQFTDG SFTQPLYRGE LNEHLGLLGP YIRAEVEDNI
 MVTFRNQASR
 1201 PYSFYSSLIS YEEDQRQGA E PRKNFVKPNE TKTYFWKVQH
 HMAPTKDEFD
 1251 CKAWAYFSDV DLEKDVHSL IGPLLVCNTN TLNPAHGRQV
 TVQEFALFFT
 1301 IFDETKSWYF TENMERNCRG APTSESATPE SGPGSEPATS
 GSETPGTSES
 1351 ATPESGPGSE PATSGSETPG TSESATPESG PGTSTEPSEG
 SAPGTSESAT
 1401 PESGPGSPAG SPTSTEEGSP AGSPTSTEEG SPAGSPTSTE
 EGTSESATPE
 1451 SGPGTSTEPS EGSAPGASSA PCNIQMEDPT FKENYRFHAI
 NGYIMDTLPG
 1501 LVMAQDQRIR WYLLSMGSNE NIHSIHFSGH VFTVRKKEEY
 KMALYNLYPG
 1551 VFETVEMLPS KAGIWRVECL IGEHLHAGMS TLFLVYSNKC
 QTPLGMASGH
 1601 IRDFQITASG QYGQWAPKLA RLHYSGSINA WSTKEPFSWI
 KVDLLAPMII
 1651 HGIKTQGARQ KFSSLYISQF IIMYSLDGKK WQTYRGNSTG
 TLMVFFGNVD
 1701 SSGIKHNIFN PPIIARYIRL HPTHYSIRST LRMELMGCDL
 NSCSMPLGME
 1751 SKAISDAQIT ASSYFTNMFA TWSPSKARLH LQGRSNAWRP
 QVNNPKEWLQ
 1801 VDFQKTMKVT GVTTQGVKSL LTSMYVKEFL ISSSQDGHQW
 TLFFQNGKVK
 1851 VFQGNQDSFT PVVNSLDPPL LTRYLRIHPQ SWVHQIALRM
 EVLGCEAQDL
 1901 YDKTHTCPPC PAPELLGGPS VFLFPPKPKD TLMISRTPEV
 TCVVVDVSHE
 1951 DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL
 HQDWLNGKEY
 2001 KCKVSNKALP APIEKTISKA KGQPREPQVY TLPPSRDELT
 KNQVSLTCLV
 2051 KGFYPSDIAV EWESNGQPEN NYKTPPVLD SDGSFFLYSK
 LTVDKSRWQQ
 2101 GNVFSCSVMH EALHNHYTQK SLSLSPGK*

[0744] FVIII204蛋白序列(具有两个AE XTEN的单链FVIIIIFc;一个为B-结构域中的288AE XTEN且一个为在氨基酸403处的144AEXTEN)(SEQ ID NO:110)

[0745]

1 MQIELSTCFF LCLLRFCFSA TRRYYLGAWE LSWDYMQSDL
GELPVDARFP
51 PRVPKSFPPN TSVVYKKTLE 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 PDGAPTSTEP SEGSA PGSPA
GSPTSTEEGT
451 STEPSEGSAP GTSTEPSEGS APGTSESATP ESGPGTSTEP
SEGSAPGTSE
501 SATPESGPGS EPATSGSETP GTSTEPSEGS APGTSTEPSE
GSAPGTSESA
551 TPESGPGTSE SATPESGPGA SSDRSYKSQY LNNGPQRIGR
KYKKVRFMAY
601 TDETFKTREA IQHESGILGP LLYGEVGD TL LIIFKNQASR
PYNIYPHGIT
651 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
701 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE
751 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
801 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEYEDTLTLF
PFSGETVFMS
851 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYIED
SYEDISAYLL
901 SKNNAIEPRS FSQNGAPGTS ESATPESGPG SEPATSGSET
PGTSESATPE
951 SGPGSEPATS GSETPGTSES ATPESGPGTS TEPSEGSAPG
SPAGSPTSTE
1001 EGTSESATPE SGPGSEPATS GSETPGTSES ATPESGPGSP
AGSPTSTEEG
1051 SPAGSPTSTE EGTSTEPSEG SAPGTSESAT PESGPGTSES
ATPESGPGTS
1101 ESATPESGPG SEPATSGSET PGSEPATSGS ETPGSPAGSP
TSTEEGTSTE

1151 PSEGSAPGTS TEPSEGSAPG SEPATSGSET PGTSESATPE
 SGPGTSTEPS
 1201 EGSAPASSPP VLKRHQAEIT RTTLQSDQEE IDYDDTISVE
 MKKEDFDIYD
 1251 EDENQSPRSF QKKTRHYFIA AVERLWDYGM SSSPHVLRNR
 AQSGSVPQFK
 1301 KVVFEFTDGT SFTQPLYRGE LNEHLGLLGP YIRAEVEDNI
 MVTFRNQASR
 1351 PYSFYSSLIS YEEDQRQGA E PRKNFVKPNE TKTYFWKVQH
 HMAPTKDEFD
 1401 CKAWAYFSDV DLEKDVHSG L IGPLLVCHTN TLNPAHGRQV
 TVQEFALFFT
 1451 IFDETKSWYF TENMERNCR PCNIQMEDPT FKENYRFHAI
 NGYIMDTLPG
 1501 LVMAQDQIR WYLLSMGSNE NIHSIHFSGH VFTVRKKEEY
 KMALYNLYPG
 1551 VFETVEMLPS KAGIWRVECL IGEHLHAGMS TLFLVYSNKC
 QTPLGMA SGH
 [0746] 1601 IRDFQITASG QYGQWAPKLA RLHYSGSINA WSTKEPFSWI
 KVDLLAPMII
 1651 HGIKTQGARQ KFSSLYISQF IIMYSLDGKK WQTYRGNSTG
 TLMVFFGNVD
 1701 SSGIKHNIFN PPIIARYIRL HPTHYSIRST LRMELMGCDL
 NSCSMPLGME
 1751 SKAISDAQIT ASSYFTNMFA TWSPSKARLH LQGRSNAWRP
 QVNNPKEWLQ
 1801 VDFQKTMKVT GVTTQGVKSL LTSMYVKEFL ISSSQDGHQW
 TLFFQNGKVK
 1851 VFQGNQDSFT PVVNSLDPPL LTRYLRIHPQ SWVHQIALRM
 EVLGCEAQDL
 1901 YDKTHTCPPC PAPELLGGPS VFLFPPKPKD TLMISRTPEV
 TCVVVDVSHE
 1951 DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL
 HQDWLNGKEY
 2001 KCKVSNKALP APIEKTISKA KGQPREPQVY TLPPSRDELT
 KNQVSLTCLV
 2051 KGFYPSDIAV EWESNGQPEN NYKTTTPVLD SDGSFFLYSK
 LTVDKSRWQQ
 2101 GNVFSCSVMH EALHNHYTQK SLSLSPGK*

[0747] FVIII 205蛋白序列(具有两个AE XTEN的单链FVIIIIFc;一个为B-结构域中的288AEXTEN且一个为在氨基酸18处的144AEXTEN) (SEQ ID NO:111)

- 1 MQIELSTCFF LCLLRFCFSA TRRYLGAVE LSWDYMQGAP
 TSESATPESG
 [0748] 51 PGSEPATSGS ETPGTSESAT PESGPGSEPA TSGSETPGTS
 ESATPESGPG
 101 TSTEPSEGA PGSPAGSPTS TEEGTSESAT PESGPGSEPA
 TSGSETPGTS

[0749]

151 ESATPESGPG SPAGSPTSTE EGSPAGSPTS TEEGASSSDL
GELPVDARFP
201 PRVPKSFFFN TSVVYKKTLE VEFTDHLFNI AKPRPPWMGL
LGPTIQAEVY
251 DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR
EKEDDKVFPG
301 GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG
LIGALLVCRE
351 GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD
AASARAWPKM
401 HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL
EGHTFLVRNH
451 RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME
AYVKVDSCPE
501 EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI
RSVAKKHPKT
551 WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR
KYKKVRFMAY
601 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR
PYNIYPHGIT
651 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
701 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE
751 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
801 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVEYEDTLTLF
PFSGETVFMS
851 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYYED
SYEDISAYLL
901 SKNNAIEPRS FSQNGAPGTS ESATPESGPG SEPATSGSET
PGTSESATPE
951 SGPGSEPATS GSETPGTSES ATPESGPGTS TEPSEGSAPG
SPAGSPTSTE
1001 EGTSESATPE SGPGSEPATS GSETPGTSES ATPESGPGSP
AGSPTSTEEG
1051 SPAGSPTSTE EGTSTEPSEG SAPGTSESAT PESGPGTSES
ATPESGPGTS
1101 ESATPESGPG SEPATSGSET PGSEPATSGS ETPGSPAGSP
TSTEEGTSTE
1151 PSEGSAPGTS TEPSEGSAPG SEPATSGSET PGTSESATPE
SGPGTSTEPS
1201 EGSAPASSPP VLKRHQAEIT RTTLQSDQEE IDYDDTISVE
MKKEDFDIYD
1251 EDENQSPRSF QKKTRHYFIA AVERLWDYGM SSSPHVLRNR
AQSGSVPQFK
1301 KVVFEFTDG SFTQPLYRGE LNEHLGLLGP YIRAEVEDNI
MVTRNQASR
1351 PYSFYSSLIS YEEDQRQGA E PRKNFVKPNE TKTYFWKVQH
HMAPTKDEFD

1401 CKAWAYFSDV DLEKDVHSGL IGPLLVCHTN TLNPAHGRQV
 TVQEFALFFT
 1451 IFDETKSWYF TENMERN CRA PCNIQMEDPT FKENYRFHAI
 NGYIMDTLPG
 1501 LVMAQDQIR WYLLSMGSNE NIHSIHFSGH VFTVRKKEEY
 KMALYNLYPG
 1551 VFETVEMLPS KAGIWRVECL IGEHLHAGMS TLFLVYSNKC
 QTPLGMA SGH
 1601 IRDFQITASG QYGQWAPKLA RLHYSGSINA WSTKEPFSWI
 KVDLLAPMII
 1651 HGIKTQGARQ KFSSLYISQF IIMYSLDGKK WQTYRGNSTG
 TLMVFFGNVD
 1701 SSGIKHNIFN PPIIARYIRL HPTHYSIRST LRMELMGCDL
 NSCSMPLGME
 [0750] 1751 SKAISDAQIT ASSYFTNMFA TWSPSKARLH LQGRSNAWRP
 QVNNPKEWLQ
 1801 VDFQKTMKVT GVTTQGVKSL LTS MYVKEFL ISSSQDGHQW
 TLFFQNGKVK
 1851 VFQGNQDSFT PVVNSLDPPL LTRYLRIHPQ SWVHQIALRM
 EVLGCEAQDL
 1901 YDKTHTCPPC PAPELLGGPS VFLFPPKPKD TLMISRTPEV
 TCVVVDVSHE
 1951 DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL
 HQDWLNGKEY
 2001 KCKVSNKALP APIEKTISKA KGQPREPQVY TLPPSRDELT
 KNQVSLTCLV
 2051 KGFYPSDIAV EWESNGQPEN NYKTTPPVLD SDGSFFLYSK
 LTVDKSRWQQ
 2101 GNVFSCSVMH EALHNHYTQK SLSLSPGK*

[0751] pSYN FVIII 266蛋白序列(具有在氨基酸18处的42 AE-XTEN和B-结构域中的288
 AE XTEN的FVIIIIFc) SEQ ID NO:112

1 MQIELSTCFF LCLLRFCFSA TRRYYLGA VE LSWDYMQGAP
 GSPAGSPTST
 51 EEGTSESATP ESGPGSEPAT SGSETPASSS DLGELPVDAR
 FPPRVPKSFP
 101 FNTSVVYKKT LFVEFTDHLF NIAKPRPPWM GLLGPTIQAE
 VYDTVVITLK
 151 NMASHPVSLH AVGVSYWKAS EGA EYDDQTS QREKEDDKVF
 PGGSHTYVWQ
 [0752] 201 VLKENGPMAS DPLCLTYSYL SHVDLVKDLN SGLIGALLVC
 REGSLAKEKT
 251 QTLHKFILLF AVFDEGKSWH SETKNSLMQD RDAASARAWP
 KMHTVNGYVN
 301 RSLPGLIGCH RKS VYWHVIG MGTTP EVHSI FLEGHTFLVR
 NHRQASLEIS
 351 PITFLTAQTL LMDLGQFLLF CHISSHQHDG MEAYVKVDSC
 PEEPQLRMKN
 401 NEEAEDYDDD LTDSEMDVVR FDDDNSPSFI QIRSVAKKHP
 KTWVHYIAAE

[0753]

451 EEDWDYAPLV LAPDDRSYKS QYLNNGPQRI GRKYKKVRFM
AYTDETFKTR
501 EAIQHESGIL GPLLYGEVGD TLLIIFKNQA SRPYNIYPHG
ITDVRPLYSR
551 RLPKGVKHLK DFPILPGEIF KYKWTVTVED GPTKSDPRCL
TRYSSSFVNM
601 ERDLASGLIG PLLICYKESV DQRGNQIMSD KRNVLFSVF
DENRSWYLTE
651 NIQRFLPNPA GVQLEDPEFQ ASNIMHSING YVFDSLQLSV
CLHEVAYWYI
701 LSIGAQTDFL SVFFSGYTFK HKMVYEDTLT LFPFSGETVF
MSMENPGLWI
751 LGCHNSDFRN RGMTALLKVS SCDKNTGDYY EDSYEDISAY
LLSKNNAIEP
801 RSFSQNGAPG TSESATPESG PGSEPATSGS ETPGTSESAT
PESGPGSEPA
851 TSGSETPGTS ESATPESGPG TSTEPSEGS SA PGSPAGSPTS
TEEGTSESAT
901 PESGPGSEPA TSGSETPGTS ESATPESGPG SPAGSPTSTE
EGSPAGSPTS
951 TEEGTSTEPS EGSAPGTSES ATPESGPGTS ESATPESGPG
TSESATPESG
1001 PGSEPATSGS ETPGSEPATS GSETPGSPAG SPTSTEEGTS
TEPSEGSAPG
1051 TSTEPSEGS SA PGSEPATSGS ETPGTSESAT PESGPGTSTE
PSEGSAPASS
1101 PPVLKRHQAE ITRTTLQSDQ EEIDYDDTIS VEMKKEDFDI
YDEDENQSPR
1151 SFQKKTRHYF IAAVERLWDY GMSSSPHVLR NRAQSGSVPQ
FKKVVFQEF
1201 DGSFTQPLYR GELNEHLGLL GPYIRAEVED NIMVTFRNQA
SRPYSFYSSL
1251 ISYEEDQRQG AEPRKNFVKP NETKTYFWKV QHMAPTKDE
FDCKAWAYFS
1301 DVDLEKDVHS GLIGPLLCH TNTLNPAHGR QVTVQEFALF
FTIFDETKSW
1351 YFTENMERNR RAPCNQMED PTFKENYRFH AINGYIMDTL
PGLVMAQDQR
1401 IRWYLLSMGS NENIHSIHFS GHVFTVRKKE EYKMALYNLY
PGVFETVEML
1451 PSKAGIWRVE CLIGEHLHAG MSTLFLVYSN KCQTPLGMAS
GHIRDFQITA
1501 SGQYGQWAPK LARLHYSISI NAWSTKEPFS WIKVDLLAPM
IIHGIKTQGA
1551 RQKFSSLYIS QFIIMYSLDG KKWQTYRGNS TGTLMVFFGN
VDSSGIKHNI
1601 FNPPIIARYI RLHPHTHSIR STLRMELMGC DLNSCSMPLG
MESKAISDAQ
1651 ITASSYFTNM FATWSPSKAR LHLQGRSNAW RPQVNNPKEW
LQVDFQKTMK

1701 VTGVTTQGVK SLLTSMYVKE FLISSSQDGH QWTLFFQNGK
VKVFQGNQDS
1751 FTPVVNSLDP PLLTRYLRIH PQSWVHQIAL RMEVLGCEAQ
DLYDKTHTCP
1801 PCPAPELLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS
HEDPEVKFNW
[0754] 1851 YVDGVEVHNA KTKPREEQYN STYRVVSVLT VLHQDWLNGK
EYKCKVSNKA
1901 LPAPIEKTIS KAKGQPREPQ VYTLPPSRDE LTKNQVSLTC
LVKGFYPSDI
1951 AVEWESNGQP ENNYKTTTPV LDSDGSFFLY SKLTVDKSRW
QQGNVFSCSV
2001 MHEALHNHYT QKSLSLSPGK *

[0755] pSYN FVIII 267蛋白序列(具有在氨基酸18处的72 AE-XTEN和B-结构域中的288AE XTEN的FVIII Fc)SEQ ID NO:113

1 MQIELSTCFF LCLLRFCFSA TRRYILGAVE LSWDYMQGAP
TSESATPESG
51 PGSEPATSGS ETPGTSESAT PESGPGSEPA TSGSETPGTS
ESATPESGPG
101 TSTEPSEGS A PGASSSDLGE LPVDARFPPR VPKSFPFNTS
VVYKKTLEVE
151 FTDHLFNI AK PRPPWMGLLG PTIQAEVYDT VVITLKNMAS
HPVSLHAVGV
201 SYWKASEGAE YDDQTSQREK EDDKVFPFGS HTYVWQVLKE
NGPMASDPLC
251 LTYSYLSHVD LVKDLNSGLI GALLVCREGS LAKEKTQTLH
KFILLFAVFD
301 EGKSWHSETK NSLMQDRDAA SARAWPKMHT VNGYVNRSLP
GLIGCHRKSV
351 YWHVIGMGTT PEVHSIFLEG HTFLVRNHRQ ASLEISPITF
LTAQTLLMDL
[0756] 401 GQFLLFCHIS SHQHDGMEAY VKVDSCPEEP QLRMKNNEEA
EDYDDDLTDS
451 EMDVVRFDDD NSPSFIQIRS VAKKHPKTWV HYIAAEEEDW
DYAPLV LAPD
501 DRSYKSQYLN NGPQRIGRKY KKVRFMAYTD ETFKTREAIQ
HESGILGPLL
551 YGEVGD TLLI IFKNQASRPY NIYPHGITDV RPLYSRRLPK
GVKHLKDFPI
601 LPGEIFKYKW TVTVEDGPTK SDPRCLTRY SSFVNMERDL
ASGLIGPLLI
651 CYKESVDQRG NQIMSDKRN ILFSVFDENR SWYLTENIQR
FLPNPAGVQL
701 EDPEFQASNI MHSINGYVFD SLQLSVCLHE VAYWYILSIG
AQTDFLSVFF
751 SGYTFFKH KMV YEDTLTLFPF SGETVFMSME NPGLWILGCH
NSDFRNRGMT
801 ALLKVSSCDK NTGDYYEDSY EDISAYLLSK NNAIEPRSFS
QNGAPGTSES

851 ATPESGPGSE PATSGSETPG TSESATPESG PGSEPATSGS
 ETPGTSESAT
 901 PESGPGTSTE PSEGSAPGSP AGSPTSTEEG TSESATPESG
 PGSEPATSGS
 951 ETPGTSESAT PESGPGSPAG SPTSTEEGSP AGSPTSTEEG
 TSTEPSEGSA
 1001 PGTSESATPE SGPGTSESAT PESGPGTSES ATPESGPGSE
 PATSGSETPG
 1051 SEPATSGSET PGSPAGSPTS TEEGTSTEPS EGSAPGTSTE
 PSEGSAPGSE
 1101 PATSGSETPG TSESATPESG PGTSTEPSEG SAPASSPPVL
 KRHQAEITRT
 1151 TLQSDQEEID YDDTISVEMK KEDFDIYDED ENQSPRSFQK
 KTRHYFIAAV
 1201 ERLWDYGMSS SPHVLRNRAQ SGSVPPQFKKV VFQEFDTGSGF
 TQPLYRGELN
 1251 EHLGLLGPYI RAEVEDNIMV TFRNQASRPY SFYSSLISYE
 EDQRQGAEP
 1301 KNFVKPNETK TYFWKVQHMM APTKDEFDCK AWAYFSDVDL
 EKDVHSGLIG
 1351 PLLVCHTNTL NPAHGRQVTV QEFALFFTIF DETKSWYFTE
 NMERNCRAPC
 1401 NIQMEDPTFK ENYRFHAING YIMDTLPGLV MAQDQIRIRWY
 LLSMGSNENI
 [0757] 1451 HSIHFSGHVF TVRKKEEYKM ALYNLYPGVF ETVEMLPSKA
 GIWRVECLIG
 1501 EHLHAGMSTL FLVYSNKCQT PLGMASGHIR DFQITASGQY
 GQWAPKLARL
 1551 HYSGSINAWS TKEPFSWIKV DLLAPMIHIG IKTQGARQKF
 SSLYISQFII
 1601 MYSLDGKKWQ TYRGNSTGTL MVFFGNVDSS GIKHNIFNPP
 IIARYIRLHP
 1651 THYSIRSTLR MELMGCDLNS CSMPLGMESK AISDAQITAS
 SYFTNMFATW
 1701 SPSKARLHLQ GRSNAWRPQV NNPKEWLQVD FQKTMKVTGV
 TTQGVKSLLT
 1751 SMYVKEFLIS SSQDGHQWTL FFQNGKVKVF QGNQDSFTPV
 VNSLDPPLLT
 1801 RYLRIHPQSW VHQIALRMEV LGCEAQDLYD KTHTCPPCPA
 PELLGGPSVF
 1851 LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG
 VEVHNAKTKP
 1901 REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNAKALPAP
 IEKTISKAKG
 1951 QPREPQVYTL PPSRDELTKN QVSLTCLVKG FYPSDIAVEW
 ESNGQPENNY
 2001 KTTTPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA
 LHNHYTQKSL
 2051 SLSPGK*

[0758] pSYN FVIII 268蛋白序列(具有在氨基酸18处的144 AE-XTEN的FVIII Fc)SEQ ID
NO:114

[0759]

1 MQIELSTCFF LCLLRFCFSA TRRYYLGAWE LSWDYMQGAP
TSESATPESG
51 PGSEPATSGS ETPGTSESAT PESGPGSEPA TSGSETPGTS
ESATPESGPG
101 TSTEPSEGS A PGSPAGSPTS TEEGTSESAT PESGPGSEPA
TSGSETPGTS
151 ESATPESGPG SPAGSPTSTE EGSPAGSPTS TEEGASSSDL
GELPVDARFP
201 PRVPKSFFPN TSVVYKKTLE VEFTDHLFNI AKPRPPWMGL
LGPTIQAEVY
251 DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR
EKEDDKVFPG
301 GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG
LIGALLVCRE
351 GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD
AASARAWPKM
401 HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL
EGHTFLVRNH
451 RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME
AYVKVDSCPE
501 EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI
RSVAKKHPKT
551 WVHYIAAEEE DWDIAPLVLA PDDRSYKSQY LNNGPQRIGR
KYKKVRFMAY
601 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR
PYNIPHGIT
651 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
TKSDPRCLTR
701 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
NVILFSVFDE
751 NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
FDSLQLSVCL
801 HEVAYWYILS IGAQTDFLSV FFSGYTFKHK MVYEDTLTLF
PFSGETVFMS
851 MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYIED
SYEDISAYLL
901 SKNNAIEPRS FSQNPPVLKR HQAEITRTTL QSDQEEIDYD
DTISVEMKKE
951 DFDIYDEDEN QSPRSFQKKT RHYFIAAVER LWDYGMSSSP
HVLNRNAQSG
1001 SVPQFKKVVF QEFTDGSFTQ PLYRGELNEH LGLLGPYIRA
EVEDNIMVTF
1051 RNQASRPYSF YSSLISYEED QRQGAEPKRN FVKPNETKTY
FWKVQHMAP
1101 TKDEFDCAW AYFSDVDLEK DVHSGGLIGPL LVCHTNTLNP
AHGRQVTVQE
1151 FALFFTIFDE TKSIFYTENM ERNCRAPCNI QMEDPTFKEN
YRFHANGYI

1201 MDTLPGLVMA QDQIRIRWYLL SMGSNENIHS IHFSGHVFTV
RKKEEYKMAL
1251 YNLYPGVFET VEMPLPSKAGI WRVECLIGEH LHAGMSTLFL
VYSNKCQTPL
1301 GMASGHIRDF QITASGQYGQ WAPKLARLHY SGSINAWSTK
EPFSWIKVDL
1351 LAPMIIHGIK TQGARQKFSS LYISQFIIMY SLDGKKWQTY
RGNSTGTLMV
1401 FFGNVDSSGI KHNIFNPPII ARYIRLHPTH YSIRSTLRME
LMGCDLNSCS
1451 MPLGMESKAI SDAQITASSY FTNMFATWSP SKARLHLQGR
SNAWRPQVNN
[0760] 1501 PKEWLQVDFQ KTMKVTGVTT QGVKSLLTSM YVKEFLISS
QDGHQWTLFF
1551 QNGKVKVFQG NQDSFTPVVN SLDPPLLTRY LRIHPQSWVH
QIALRMEVLG
1601 CEAQDLYDKT HTCPPCPAPE LLGGPSVFLF PPKPKDTLMI
SRTPEVTCVV
1651 VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
SVLTVLHQDW
1701 LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP
SRDELTKNQV
1751 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGDS
FFLYSKLTVD
1801 KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK*

[0761] pSYN FVIII 269蛋白序列(具有在氨基酸18处的72AE-XTEN的FVIII Fc) SEQ ID
NO:115

1 MQIELSTCFF LCLLRFCFSA TRRYILGAVE LSWDYMQGAP
TSESATPESG
51 PGSEPATSGS ETPGTSESAT PESGPGSEPA TSGSETPGTS
ESATPESGPG
101 TSTEPSEGS A PGASSSDLGE LPVDARFPPR VPKSFPPNTS
VVYKKTLEVE
151 FTDHLFNIK PRPPWMGLLG PTIQAEVYDT VVITLKNMAS
HPVSLHAVGV
201 SYWKASEGAE YDDQTSQREK EDDKVFPFGS HTYVWQVLKE
NGPMASDPLC
[0762] 251 LTYSYLSHVD LVKDLNSGLI GALLVCREGS LAKEKTQTLH
KFILLFAVFD
301 EGKSWHSETK NSLMQDRDAA SARAWPKMHT VNGYVNRSLP
GLIGCHRKSV
351 YWHVIGMGT PEVHSIFLEG HTFLVRNHRQ ASLEISPITF
LTAQTLLMDL
401 GQFLLFCHIS SHQHDGMEAY VKVDSCPEEP QLRMKNNEEA
EDYDDDLTDS
451 EMDVVRFD D NSPSFIQIRS VAKKHPKTWV HYIAAEEEDW
DYAPLV LAPD
501 DRSYKSQYLN NGPQRIGRKY KKVRFMAYTD ETFKTREAIQ
HESGILGPLL

551 YGEVGDTHLLI IFKNQASRPY NIYPHGITDV RPLYSRRLPK
 GVKHLKDFPI
 601 LPGEIFKYKW TVTVEDGPTK SDPRCLTRY Y SSFVNMERDL
 ASGLIGPLLI
 651 CYKESVDQRG NQIMSDKRN ILFSVFDENR SWYL TENIQR
 FLPNPAGVQL
 701 EDPEFQASNI MHSINGYVFD SLQLSVCLHE VAYWYILSIG
 AQTDFLSVFF
 751 SGYTFKHK MV YEDTLTLFPF SGETVFMSME NPGLWILGCH
 NSDFRNRGMT
 801 ALLKVSSCDK NTGDY YEDSY EDISAYLLSK NNAIEPRSF S
 QNPPVLKRHQ
 851 AEITRTTLQS DQEEIDYDDT ISVEMKKEDF DIYDE DENQS
 PRSFQKKTRH
 901 YFIAAVERLW DYGMSSSPHV LRNRAQSGSV PQFKKVVFQE
 FTDGSFTQPL
 951 YRGELNEHLG LLGPYIRAEV EDNIMVTFRN QASRPYSFYS
 SLISYEEDQR
 1001 QGAEPKRFV KPNETKTYFW KVQHMAPTK DEFDC KAWAY
 FSDVDLEKDV
 1051 HSGLIGPLLV CHTNTLNPAH GRQVTVQEFA LFFTIFDETK
 SWYFTENMER
 1101 NCRAPCNIQM EDPTFKENYR FHAINGYIMD TLPGLVMAQD
 QRIRWYLLSM
 [0763] 1151 GSNENIHSIH FSGHVFTVRK KEEYKMALYN LYPGVFETVE
 MLPSKAGIWR
 1201 VECLIGEHLH AGMSTLFLVY SNKCQTPLGM ASGHIRDFQI
 TASGQYGQWA
 1251 PKLARLHYS G SINAWSTKEP FSWIKVDLLA PMIIHG IKTQ
 GARQKFSSLY
 1301 ISQFIIMYSL DGKKWQTYRG NSTGTLMVFF GNVDS SGIKH
 NIFNPPIIAR
 1351 YIRLHPHYS IRSTLRMELM GCDLN SC SMP LGMESKAISD
 AQITASSYFT
 1401 NMFATWSPSK ARLHLQGRSN AWRPQVNNPK EWLQVDFQKT
 MKVTGVTTQG
 1451 VKSLLTSMYV KEFLISSQD GHQWTLFFQN GKV KVFQGNQ
 DSFTPVVNSL
 1501 DPPLLTRYLR IHPQSWVHQI ALRMEVLGCE AQDLYDKTHT
 CPPCPAPELL
 1551 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF
 NWYVDGVEVH
 1601 NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN
 KALPAPIEKT
 1651 ISKAKGQPRE PQVYTLPPSR DELTKNQVSL TCLVKGFYPS
 DIAVEWESNG
 1701 QPENNYKTP PVLDS DGSFF LYSKLTVDKS RWQQGNVFSC
 SVMHEALHNH
 1751 YTQKSLSLSP GK*

[0764] pSYNFVIII271蛋白序列(具有在氨基酸18处的42 AE-XTEN的FVIII Fc) SEQ ID
 NO:116

[0765]

1 MQIELSTCFF LCLLRFCFSA TRRYYLGAWE LSWDYMQGAP
GSPAGSPTST
51 EEGTSESATP ESGPGSEPAT SGSETPASSS DLGELPVDAR
FPPRVPKSFP
101 FNTSVVYKKT LFVEFTDHLF NIAKPRPPWM GLLGPTIQAE
VYDTVVITLK
151 NMASHPVSLH AVGVSYWKAS EGAEYDDQTS QREKEDDKVF
PGGSHTYVWQ
201 VLKENGPMAS DPLCLTYSYL SHVDLVKDLN SGLIGALLVC
REGSLAKEKT
251 QTLHKFILLF AVFDEGKSWH SETKNSLMQD RDAASARAWP
KMHTVNGYVN
301 RSLPGLIGCH RKSVMYWHVIG MGTTPPEVHSI FLEGHTFLVR
NHRQASLEIS
351 PITFLTAQTL LMDLGQFLLF CHISSHQHDG MEAYVKVDSC
PEEPQLRMKN
401 NEEAEDYDDD LTDSEMDVVR FDDDNPSFI QIRSVAKKHP
KTWVHYIAAE
451 EEDWDYAPLV LAPDDRSYKS QYLNNGPQRI GRKYKKVRFM
AYTDETFKTR
501 EAIQHESGIL GPLLYGEVGD TLLIIFKNQA SRPYNIPYHG
ITDVRPLYSR
551 RLPKGVKHLK DFPILPGEIF KYKWTVTVED GPTKSDPRCL
TRYYSFVNM
601 ERDLASGLIG PLLICYKESV DQRGNQIMSD KRNVLFSVF
DENRSWYLTE
651 NIQRFLPNPA GVQLEDPEFQ ASNIMHSING YVFDSLQLSV
CLHEVAYWYI
701 LSIGAQTDFL SVFFSGYTFK HKMVEYEDTLT LFPFSGETVF
MSMENPGLWI
751 LGCHNSDFRN RGMTALLKVS SCDKNTGDYY EDSYEDISAY
LLSKNNAIEP
801 RSFSQNPPVL KRHQAEITRT TLQSDQEEID YDDTISVEMK
KEDFDIYDED
851 ENQSPRSFQK KTRHYFIAAV ERLWDYGMSS SPHVLRNRAQ
SGSVPQFKKV
901 VFQEFTDGSF TQPLYRGELN EHLGLLGPYI RAEVEDNIMV
TFRNQASRPY
951 SFYSSLISYE EDQRQGAEPK KNFVKPNETK TYFWKVQHMM
APTKDEFDCK
1001 AWAYFSDVDL EKDVHSGGLIG PLLVCHTNTL NPAHGRQVTV
QEFALFFTIF
1051 DETKSWYFTE NMERNCRAPC NIQMEDPTFK ENYRFHAING
YIMDTLPGLV
1101 MAQDQIRRWY LLSMGSNENI HSIHFSGHVF TVRKKEEYKM
ALYNLYPGVF
1151 ETVEMLPSKA GIWRVECLIG EHLHAGMSTL FLVYSNKCQT
PLGMASGHIR

1201 DFQITASGQY GQWAPKLARL HYSGSINAWs TKEPFSWIKV
 DLLAPMIIHG
 1251 IKTQGARQKF SSLYISQFII MYSLDGKKWQ TYRGNSTGTL
 MVFFGNVDSS
 1301 GIKHNIFNPP IIARYIRLHP THYSIRSTLR MELMGCDLNS
 CSMPLGMESK
 1351 AISDAQITAS SYFTNMFATW SPSKARLHLQ GRSNAWRPQV
 NNPKEWLQVD
 1401 FQKTMKVTGV TTQGVKSLLT SMYVKEFLIS SSQDGHQWTL
 FFQNGKVKVF
 [0766] 1451 QGNQDSFTPV VNSLDPPLLT RYLRIHPQSW VHQIALRMEV
 LGCEAQDLYD
 1501 KTHTCPPCPA PELLGGPSVF LFPPKPKDTL MISRTPEVTC
 VVVDVSHEDP
 1551 EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ
 DWLNGKEYKC
 1601 KVSNAKALPAP IEKTISKAKG QPREPQVYTL PPSRDELTKN
 QVSLTCLVKG
 1651 FYPSDIAVEW ESNQOPENNY KTTTPVLDSD GSFFLYSKLT
 VDKSRWQQGN
 1701 VFSCSVMHEA LHNHYTQKSL SLSPGK*

[0767] pSYN FVIII蛋白序列272 (具有在氨基酸18处的144AE XTEN和B-结构域中的244 AE XTEN的FVIII-无Fc) SEQ ID NO:117

1 MQIELSTCFF LCLLRFCFSA TRRYLGAVE LSWDYMQGAP
 TSESATPESG
 51 PGSEPATSGS ETPGTSESAT PESGPGSEPA TSGSETPGTS
 ESATPESGPG
 101 TSTEPSEGSa PGSPAGSPTS TEEGTSESAT PESGPGSEPA
 TSGSETPGTS
 151 ESATPESGPG SPAGSPTSTE EGSPAGSPTS TEEGASSSDL
 GELPVDARFP
 201 PRVPKSFPFN TSVVYKKTFL VEFTDHLFNI AKPRPPWMGL
 LGPTIQAEVY
 251 DTVVITLKNM ASHPVSLHAV GVSYWKASEG AEYDDQTSQR
 EKEDDKVFPG
 [0768] 301 GSHTYVWQVL KENGPMASDP LCLTYSYLSH VDLVKDLNSG
 LIGALLVCRE
 351 GSLAKEKTQT LHKFILLFAV FDEGKSWHSE TKNSLMQDRD
 AASARAWPKM
 401 HTVNGYVNRS LPGLIGCHRK SVYWHVIGMG TTPEVHSIFL
 EGHTFLVRNH
 451 RQASLEISPI TFLTAQTLLM DLGQFLLFCH ISSHQHDGME
 AYVKVDSCPE
 501 EPQLRMKNNE EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI
 RSVAKKHPKT
 551 WVHYIAAEEE DWDYAPLVLA PDDRSYKSQY LNNGPQRIGR
 KYKKVRFMAY
 601 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR
 PYNIPHGIT

[0769]

```

651  DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP
    TKSDPRCLTR
701  YYSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR
    NVILFSVFDE
751  NRSWYL TENI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV
    FDSLQLSVCL
801  HEVAYWYILS IGAQTDFLSV FFSGYTFFKHK MVYEDTLTLF
    PFSGETVFM
851  MENPGLWILG CHNSDFRNRG MTALLKVSSC DKNTGDYIED
    SYEDISAYLL
901  SKNNAIEPRS FSQNGAPGTS ESATPESGPG SEPATSGSET
    PGTSESATPE
951  SGPGSEPATS GSETPGTSES ATPESGPGTS TEPSEGSAPG
    SPAGSPTSTE
1001 EGTSESATPE SGPGSEPATS GSETPGTSES ATPESGPGSP
    AGSPTSTEEG
1051 SPAGSPTSTE EGTSTEPSEG SAPGTSESAT PESGPGTSES
    ATPESGPGTS
1101 ESATPESGPG SEPATSGSET PGSEPATSGS ETPGSPAGSP
    TSTEEGTSTE
1151 PSEGSAPGTS TEPSEGSAPG SEPATSGSET PGTSESATPE
    SGPGTSTEPS
1201 EGSAPASSPP VLKRHQAEIT RTTLQSDQEE IDYDDTISVE
    MKKEDFDIYD
1251 EDENQSPRSF QKKTRHYFIA AVERLWDYGM SSSPHVLRNR
    AQSGSVPQFK
1301 KVVVFQFTDG SFTQPLYRGE LNEHLGLLGP YIRAEVEDNI
    MVTFRNQASR
1351 PYSFYSSLIS YEEDQRQGA E PRKNFVKPNE TKTYFWKVQH
    HMAPTKDEFD
1401 CKAWAYFSDV DLEKDVHSG L IGPLLVCHTN TLNPAHGRQV
    TVQEFALFFT
1451 IFDETKSWYF TENMERN CRA PCNIQMEDPT FKENYRFHAI
    NGYIMDTLPG
1501 LVMAQDQRI R WYLLSMGSNE NIHSIHFSGH VFTVRKKEEY
    KMALYNLYPG
1551 VFETVEMLPS KAGIWRVECL IGEHLHAGMS TLFLVYSNKC
    QTPLGMASGH
1601 IRDFQITASG QYGQWAPKLA RLHYSGSINA WSTKEPFSWI
    KVDLLAPMII
1651 HGIKTQGARQ KFSSLYISQF IIMYSLDGKK WQTYRGNSTG
    TLMVFFGNVD
1701 SSGIKHNIFN PPIIARYIR L HPTHYSIRST LRMELMGCDL
    NSCSMPLGME
1751 SKAISDAQIT ASSYFTNMFA TWSPSKARLH LQGRSNAWRP
    QVNNPKEWLQ
1801 VDFQKTMKVT GVTTQGVKSL LTS MYVKEFL ISSSQDGHQW
    TLFFQNGKVK
1851 VFQGNQDSFT PVVNSLDPPL LTRYLRIHPQ SWVHQIALRM
    EVLGCEAQDL
1901 Y*

```

[0770] pSYN VWF 031蛋白序列(VWF D1D2D'D3-48aa长的凝血酶可切割的GS接头-Fc) SEQ
ID NO:118

[0771]

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 VLYGWDHSA
CSPVCPAGME
301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG
LCVESTECPC
351 VHS GKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV
TGQSHFKSFD
401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC
TRSVTVRLPG
451 LHNLSLVKLKH GAGVAMDGQD IQLPLLKGD L RIQHTVTASV
RLSYGEDLQM
501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTSPG
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 MVKLVC PADN LRAEGLECTK
TCQNYDLECM
801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE
TVKIGCNTCV
851 CRDRKWNCTD HVCDATCSTI GMAHYLTFDG LKYLFPGEQ
YVLVQDYCGS
901 NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE
VNVKRPMKDE
951 THFEVVESGR YIILLGKAL SVVWDRHLSI SVVLKQTYQE
KVCGLCGNFD
1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD
SSPATCHNNI
1051 MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS
CESIGDCAAF
1101 CDTIAAYAHV CAQH GKVV TW RTATLCPQSC EERNLRENGY
EAEWRYNSCA

1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC
 VDPEDCPVCE
 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP
 ISGGGSGGG
 1251 GSGGGSGGG GSGGGSGGG GSLVPRGSGG GSGGGGSDK
 THTCPAP
 [0772] 1301 ELLGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE
 VKFNWYVDGV
 1351 EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK
 VSNKALPAPI
 1401 EKTISKAKGQ PREPQVYTL PPSRDELTKNQ VSLTCLVKGF
 YPSDIAVEWE
 1451 SNGQPENNYK TTPVLDSDG SFFLYSKLTV DKSRWQQGNV
 FSCSVMHEAL
 1501 HNHYTQKSLS LSPGK*

[0773] pSYN VWF 034蛋白序列(VWF D1D2D'D3-288AE XTEN-35aa长的凝血酶可切割的GS
 接头-Fc) SEQ ID NO:119

1 MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS
 DFVNTFDGSM
 51 YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE
 FFDIHLFVNG
 101 TVTQGDQSVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI
 DGSGNFQVLL
 151 SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS
 WALSSGEQWC
 201 ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL
 VDPEPFVALC
 251 EKTLCCECAGG LECACPALLE YARTCAQEGM VLYGWDHSA
 CSPVCPAGME
 301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG
 LCVESTECPC
 351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV
 TGQSHFKSFD
 [0774] 401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC
 TRSVTVRLPG
 451 LHNSLVKLKH GAGVAMDGQD IQLPLLKGD LRIQHTVTASV
 RLSYGEDLQM
 501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG
 LAEPRVEDFG
 551 NAWKLHGDCQ DLQKQHS DPC ALNPRMTRFS EEACAVLTSP
 TFEACHRAVS
 601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV
 AWREPGRCEL
 651 NCPKGQVYLQ CGTPCNLTCT 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 YIILLGKAL SVVWDRHLSI SVVLKQTYQE
 KVCGLCGNFD
 1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD
 SSPATCHNNI
 1051 MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS
 CESIGDCAAF
 1101 CDTIAAYAHV CAQHKGKVVW RTATLCPQSC EERNLRENGY
 EAEWRYNCA
 1151 PACQVTCQHP EPLACPVCV EGCHAHCPPG KILDELLQTC
 VDPEDCPVCE
 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP
 ISGTSESATP
 1251 ESGPGSEPAT SGSETPGTSE SATPESGPGS EPATSGSETP
 GTSESATPES
 1301 GPGTSTEPSE GSAPGSPAGS PTSTEEGTSE SATPESGPGS
 EPATSGSETP
 1351 GTSESATPES GPGSPAGSPT STEEGSPAGS PTSTEEGTST
 EPSEGSAPGT
 1401 SESATPESGP GTSESATPES GPGTSESATP ESGPGSEPAT
 SGSETPGSEP
 1451 ATSGSETPGS PAGSPTSTEE GTSTEPSEGS APGTSTEPSE
 GSAPGSEPAT
 1501 SGSETPGTSE SATPESGPGT STEPSEGSAP DIGGGGGSGG
 GGSLVPRGSG
 1551 GDKTHTCPPC PAPELLGGPS VFLFPPKPKD TLMISRTPEV
 TCVVVDVSHE
 1601 DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL
 HQDWLNGKEY
 1651 KCKVSNKALP APIEKTISKA KGQPREPQVY TLPPSRDELT
 KNQVSLTCLV
 1701 KGFYPSDIAV EWESNGQPEN NYKTPPVLD SDGSFFLYSK
 LTVDKSRWQQ
 1751 GNVFSCSVMH EALHNHYTQK SLSLSPGK*

[0776] pSYNVWF 036蛋白序列 (VWF D1D2D'D-98aa长的凝血酶可切割的GS接头-Fc) SEQ
ID NO:120

1 MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS
 DFVNTFDGSM
 51 YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE
 FFDIHLFVNG
 101 TVTQGDQRVs MPYASKGLYL ETEAGYYKLS GEAYGFVARI
 DGSGNFQVLL
 151 SDRYFNKTCG LCGNFNIFAE DDFMTQEGTL TSDPYDFANS
 WALSSGEQWC

[0778]

201 ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL
VDPEPFVALC
251 EKTLCCECAGG LECACPALLE YARTCAQEGM VLYGWDHSA
CSPVCPAGME
301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG
LCVESTECPC
351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV
TGQSHFKSFD
401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC
TRSVTVRLPG
451 LHNSLVKLKH GAGVAMDGQD IQLPLLKGD LRIQHTVTASV
RLSYGEDLQM
501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG
LAEPRVEDFG
551 NAWKLHGDCQ DLQKQHSDFC ALNPRMTRFS EEACAVLTSP
TFEACHRAVS
601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV
AWREPGRCEL
651 NCPKGQVYLQ CGTPCNLTCT 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 CKKRVTLVE GGEIELFDGE
VNVKRPMKDE
951 THFEVVESGR YIILLGKAL SVVWDRHLSI SVVLKQTYQE
KVCGLCGNFD
1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD
SSPATCHNNI
1051 MKQTMVDSSC RILTSDVFQD CNKLVDPEPY LDVCIYDTCS
CESIGDCAAF
1101 CDTIAAYAHV CAQHKGKVVW RTATLCPQSC EERNLRENGY
EAEWRYNSCA
1151 PACQVTCQHP EPLACPVCV EGCHAHCPPG KILDELLQTC
VDPEDCPVCE
1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP
ISGGGGSGGG
1251 GSGGGGSGGG GSGGGGSGGG GSGGGGSGGG GSGGGGSGGG
GSGGGGSGGG
1301 GSGGGGSGGG GSGGGGSGGG GSLVPRGSGG GSGGGGSGGG
THTCPPCPAP
1351 ELLGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE
VKFNWYVDGV
1401 EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK
VSNKALPAPI

1451 EKTISKAKGQ PREPQVYTL PPSRDELTKNQ VSLTCLVKGF
YPSDIAVEWE

[0779] 1501 SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV
FSCSVMEAL

1551 HNHYTQKSLS LSPGK*

[0780] pSYN Fc-015蛋白序列(IgG-Fc结构域) SEQ ID NO:121

1 METDTLLLWV LLLWVPGSTG DKTHTCPPCP APELLGGPSV
FLFPPKPKDT

51 LMISRTPEVT CVVVDVSHED PEVKFNWYVD GVEVHNAKTK
PREEQYNSTY

[0781] 101 RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
GQPREPQVYT

151 LPPSRDELTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN
YKTTTPVLDS

201 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS
LSLSPGK*

[0782] 实施例15:FVIII-XTEN-Fc:VWF-Fc异二聚体相较于野生型BDD-FVIII已保持正常的FVIII比活性。

[0783] FVIII-XTEN-Fc:VWF-Fc异二聚体的FVIII比活性被确定。使用两步色谱过程纯化异二聚体。使用弱阴离子交换树脂,随后使用亲和力色谱。最终纯化产物通过SEC-HPLC具有可接受的纯度。将比活性与B-结构域缺失的FVIII (BDD-FVIII) 比较,如通过FVIII显色测定和A280浓度所测量的。数据呈现于表26中。所有测试分子已表明与BDD-FVIII类似的FVIII比活性。分子的纯度和各部分的存在通过SDS-PAGE和蛋白质印迹来确认。

[0784] 表26:FVIII-XTEN-Fc:VWF-Fc异二聚体的FVIII比活性

构建体	FVIII 207 scBDDFVIII	FVIII-66 dcBDD FVIII)	FVIII 155 / vWF31	FVIII 155 / vWF39	FVIII 169 / vWF31	FVIII 205 / vWF31	FVIII 169 / vWF34
[0785] 测量的 比活性 (IU/nmol)	1473	1592	1534	1796	1511	1345	1505

[0786] 在HemA小鼠中比较rFVIII-XTEN/D'D3和BDD-FVIII的半衰期(图15;表27)。如图15示出,rFVIII-XTEN/D'D3达到的半衰期为由BDD-FVIII达到的半衰期的四倍长。

[0787] 表27:在HemA小鼠中的rFVIII-XTEN/D'D3和BDD-FVIII

治疗	5 分钟回 收率(%)	HL (小时)	MRT (小时)	Cl (mL/小 时/kg)	Vss (mL/k g)	AUC_D (小时 *kg*mIU/mL/ mIU)
[0788] BDD-FV III	89	7.6	11	4.5	49.2	0.22
rFVIII-Fc	78	16	20	2.9	57.8	0.35
rFVIII-X TEN/D'D 3	86	30	36	1.8	63.4	0.57

[0789] 实施例16:如通过一阶段aPTT测定测量的FVIII-XTEN-Fc:VWF-Fc异二聚体在止血

方面的效能 (FVIII活性)

[0790] 通过它们的FVIII特异性aPTT活性评价FVIII-XTEN-Fc:VWF-Fc异二聚体在止血方面的效能,如于表28中总结。如由表28所示,尽管添加VWF D’D3片段和将XTEN插入到FVIII的结构域中会减少异二聚体的FVIII特异性aPTT活性(如由FVIII155/VWF031数据和FVIII205/VWF031数据指示),但是XTEN插入到FVIII B结构域区域或VWF D’D3片段的C-末端对FVIII特异性aPTT活性无负面影响(如通过FVIII169/VWF031数据和FVIII169/VWF034数据指示)。相较于双链BDD-FVIII (dcBDD-FVIII),FVIII155/VWF031、FVIII169/VWF031、FVIII169/VWF034和VWF205/VWF031显示特异性aPTT活性分别减少2.5倍、2.8倍、2.6倍和5.5倍。

[0791] 表28:FVIII-XTEN-Fc:VWF-Fc异二聚体的FVIII特异性aPTT活性

[0792]	构建体	FVIII 207 scBDD- FVIII	FVIII-66 dcBDD- FVIII	FVIII 155 / VWF31	FVIII 169 /VWF31	FVIII 205 / VWF31	FVIII 169 / VWF34
	测量的 特异性 aPTT 活性 (IU/nmol)	818 ± 153	1188 ± 213	448 ± 111	416 ± 70	214 ± 38	436 ± 189

[0793] FVIII特异性aPTT测定

[0794] 将FVIII变体用aPTT缓冲液 (0.15M NaCl、0.05M Tris-HCl、1%BSA,pH7.4) 稀释至线性测定范围 (200–1.6mU/mL)。将50μL所稀释的样本或标准相继与50μL 37℃的天然人HemA合并的血浆、50μL 37℃的aPTT试剂 (ACTIN® FSL活化头丝束蛋白试剂–Dade Behring,参照#B4219-2) 混合并在37℃温育4分钟。然后将50μl 20mM CaCl₂ (Dade Behring [参照#ORF037]) 加入至反应混合物以开始凝血反应。使用每个样本的凝血时间 (从添加CaCl₂直至血块形成发作的时间长度), 针对用第8种国际标准FVIII浓缩物产生的标准计算aPTT活性。针对通过OD280测量的每个分子的蛋白浓度计算特异性aPTT活性。

[0795] 实施例17:FVIII-XTEN-Fc:VWF-Fc异二聚体在HemA小鼠尾夹出血模型中的体内功效

[0796] 为了进一步利用异二聚体的止血效能,在HemA小鼠尾夹出血模型中相较于BDD-FVIII评价FVIII169/VWF034和FVIII205/VWF031的急性功效。HemA小鼠通过以200、65和20IU/kg单次静脉内注射BDD-FVIII进行治疗以产生尾夹后损伤失血对照水平。将来自用200IU/kg FVIII169/VWF034或FVIII205/VWF031治疗的小鼠的失血与BDD-FVIII治疗的对照组小鼠的失血相比较以评估它们对止血的效能。使用媒介物治疗的动物以产生模型的失血基线。如图16中所示,观察到来自所有FVIII治疗组的失血相较于媒介物治疗的动物的失血显著性减少 (p<0.05)。FVIII169/VWF034和FVIII205/VWF031两者在HemA小鼠尾夹模型中是有效的。相较于BDD-FVIII,对于FVIII169/VWF034观察到效能下降约3倍,如由通过65IU/kg BDD-FVIII和200IU/kg FVIII169/VWF034达到的类似失血减少所示。关于FVIII205/VWF034,已观察到效能下降10倍,如由通过20IU/kg BDD-FVIII和200IU/kg FVIII205/VWF031达到的类似的失血减少所示。

[0797] 即使FVIII169/VWF034和FVIII205/VWF031相较于rBDD-FVIII具有类似的特异性FVIII显色活性,但是由于对分子的修饰,它们的FVIII aPTT活性和体内效能两者均减少。

那些数据指示FVIII分子的aPTT活性关于预测其对止血的体内效能为比FVIII显色活性更准确的测量结果。

[0798] HemA小鼠尾夹出血模型

[0799] 8-10周龄雄性HemA小鼠用于该研究。在尾夹损伤之前,用50mg/kg氯胺酮/0.5mg/kg右美托咪啶混合物麻醉小鼠并置于37℃加热垫上以有助于保持体温。然后将小鼠的尾巴浸入37℃水中10分钟以扩张侧静脉。在静脉扩张后,经由尾静脉注射rFVIII或媒介物溶液并且5分钟后,使用具有直边的#11解剖刀切去尾巴远端1cm。将流血收集到13ml 37℃温盐水持续30分钟,然后将小鼠施以安乐死,尽管通过双侧胸廓切开术仍处于麻醉状态。通过在收集血液(以克计)之前和之后血液采集管的重量变化用重量测定法定量失血,其转换为失血体积的毫升(mL) (1g重量变化=1mL失血)。

[0800] 前面对特定实施方案的描述将充分地揭示本发明的一般性,他人可通过应用在本领域技能内的知识,在不偏离本发明的一般概念下无需过度实验,容易地修改和/或使之适应各种应用和此类特定实施方案。因此,基于本文呈现的教义和指导,此类改变和修饰旨在公开的实施方案的等效物的含义和范围之内。应理解,本文的措辞或术语是用于描述的目的而并非限制,以使本说明书的术语或措辞由熟练技术人员根据教义和指导来解释。

[0801] 根据考虑本文公开的说明书和本发明实践,本发明的其它实施方案将对本领域技术人员而言是明显的。说明书和实施例旨在被认为仅仅是示例性的,其中本发明的真实范围和精神由下列权利要求指示。

[0802] 本文所述的所有专利和公布通过引用整体并入本文。

序列表

- <110> 比奥贝拉蒂治疗公司
<120> 具有XTEN和血管性血友病因子蛋白的因子VIII复合物、及其用途
<130> 2159.365PC06/EKS/C-K/E-H
<140> 待指定
<141> 随附
<150> US 61/840,811
<151> 2013-06-28
<150> US 61/827,158
<151> 2013-05-24
<150> US 61/801,544
<151> 2013-03-15
<150> US 61/801,504
<151> 2013-03-15
<150> US 61/759,819
<151> 2013-02-01
<150> US 61/670,401
<151> 2012-07-11
<160> 140
<170> PatentIn 3.5版
<210> 1
<211> 8442
<212> DNA
<213> 智人
<400> 1

```
atgattcctg ccagatttgc cggggtgctg cttgctctgg ccctcattht gccaggacc 60
ctttgtgcag aaggaactcg cggcaggcca tccacggccc gatgcagcct tttcggaagt 120
gacttcgtca acacctttga tgggagcatg tacagctttg cgggatactg cagttacctc 180
ctggcagggg gctgccagaa acgctccttc tcgattattg gggacttcca gaatggcaag 240
agagtgaacc tctccgtgta tcttggggaa ttttttgaca tccatttggt tgtcaatggt 300
accgtgacac agggggacca aagagtctcc atgccctatg cctccaaagg gctgtatcta 360
gaaactgagg ctgggtacta caagctgtcc ggtgaggcct atggctttgt ggccaggatc 420
gatggcagcg gcaactttca agtcttctg tcagacagat acttcaacaa gacctgcggg 480
ctgtgtggca actttaacat ctttctgtga gatgacttta tgaccaaga agggaccttg 540
acctcggacc cttatgactt tgccaactca tgggctctga gcagtggaga acagtgggtg 600
gaacggggcat ctctctccag cagctcatgc aacatctcct ctggggaaat gcagaagggc 660
ctgtggggagc agtgccagct tctgaagagc acctcggtgt ttgcccgtg ccacctctg 720
gtggaccccg agccttttgt ggccctgtgt gagaagactt tgtgtgagtg tgctgggggg 780
```

ctggagtgcg cctgccctgc cctcctggag tacgcccga cctgtgcccga ggagggaatg 840
 gtgctgtacg gctggaccga ccacagcgcg tgcagcccag tgtgccctgc tggtatggag 900
 tataggcagt gtgtgtcccc ttgcgccagg acctgccaga gcctgcacat caatgaaatg 960
 tgtcaggagc gatgcgtgga tggctgcagc tgccctgagg gacagctcct ggatgaaggc 1020
 ctctgcgtgg agagcaccga gtgtccctgc gtgcattccg gaaagcgcta ccctcccggc 1080
 acctccctct ctcgagactg caacacctgc atttgccgaa acagccagtg gatctgcagc 1140
 aatgaagaat gtccagggga gtgccttgtc actggtcaat cccacttcaa gagctttgac 1200
 aacagatact tcaccttcag tgggatctgc cagtacctgc tggcccggga ttgccaggac 1260
 cactccttct ccattgtcat tgagactgtc cagtgtgctg atgaccgga cgctgtgtgc 1320
 acccgctccg tcaccgtccg gctgcctggc ctgcacaaca gccttgtgaa actgaagcat 1380
 ggggcaggag ttgccatgga tggccaggac atccagctcc cctcctgaa aggtgacctc 1440
 cgcattccagc atacagtgc ggctccgtg cgctcagct acggggagga cctgcagatg 1500
 gactgggatg gccgcgggag gctgtgtgtg aagctgtccc ccgtctatgc cgggaagacc 1560
 tgcggcctgt gtgggaatta caatggcaac cagggcgacg acttccttac cccctctggg 1620
 ctggcrgagc cccgggtgga ggacttcggg aacgcctgga agctgcacgg ggactgccag 1680
 gacctgcaga agcagcacag cgatccctgc gccctcaacc cgcgcatgac caggttctcc 1740
 gaggaggcgt gcgcggtcct gacgtcccc acattcgagg cctgccatcg tgccgtcagc 1800
 ccgtgccct acctgcgaa ctgccgtac gacgtgtgct cctgctcgga cggccgcgag 1860
 tgctgtgcg gcgccctggc cagctatgcc gcggcctgcg cggggagagg cgtgcgcgtc 1920
 gcgtggcgcg agccaggccg ctgtgagctg aactgcccga aaggccaggt gtacctgcag 1980
 tgcgggaccc cctgcaacct gacctgccgc tctctctctt acccgatga ggaatgcaat 2040
 gaggcctgcc tggagggtg cttctgcccc ccagggtctt acatggatga gaggggggac 2100
 tgctgtgccc agggccagt cccctgttac tatgacggtg agatcttcca gccagaagac 2160
 atcttctcag accatcacac catgtgctac tgtgaggatg gcttcatgca ctgtaccatg 2220
 agtggagtcc ccggaagctt gctgcctgac gctgtcctca gcagtcccct gtctcatcgc 2280
 agcaaaaagga gcctatcctg tcggccccc atggtcaagc tgggtgtgtc cgctgacaac 2340
 ctgcgggctg aagggtcga gtgtacaaa acgtgccaga actatgacct ggagtgcag 2400
 agcatgggct gtgtctctgg ctgcctctgc cccccggga tgggtccgga tgagaacaga 2460
 tgtgtggccc tgaaaagtg tcctgtctc catcaggga aggagtatgc ccctggagaa 2520
 acagtgaaga ttggctgcaa cacttgtgtc tgtcgggacc ggaagtggaa ctgcacagac 2580
 catgtgtgtg atgccacgtg ctccacgac ggcatggccc actacctcac cttcgacggg 2640
 ctcaaatacc tgttccccgg ggagtgccag tacgttctgg tgcaggatta ctgcggcagt 2700
 aaccctggga ctttcggat cctagtgggg aataaggat gcagccccc ctcaagtaaa 2760
 tgcaagaaac gggtcacat cctggtggag ggaggagaga ttgagctgtt tgacggggag 2820
 gtgaatgtga agaggcccat gaaggatgag actcactttg aggtggtgga gtctggccgg 2880
 tacatcattc tgctgtctgg caaagccctc tccgtggtct gggaccgcca cctgagcatc 2940
 tccgtggtcc tgaagcagac ataccaggag aaagtgtgtg gcctgtgtgg gaattttgat 3000
 ggcatccaga acaatgacct caccagcagc aacctccaag tggaggaaga ccctgtggac 3060
 tttgggaact cctggaaaagt gagctcgag tgtgtgaca ccagaaaagt gcctctggac 3120

tcatccccctg ccacctgccca taacaacatc atgaagcaga cgatggtgga ttcctcctgt 3180
agaatcctta ccagtgcagt cttccaggac tgcaacaagc tgggtggacc cgagccatat 3240
ctggatgtct gcatttacga cacctgctcc tgtgagtcca ttggggactg cgcctgcttc 3300
tgcgacacca ttgctgccta tgcccacgtg tgtgcccagc atggcaaggt ggtgacctgg 3360
aggacggcca cattgtgccc ccagagctgc gaggagagga atctccggga gaacgggtat 3420
gagtgtgagt ggcgctataa cagctgtgca cctgcctgtc aagtcacgtg tcagcacctt 3480
gagccactgg cctgccctgt gcagtgtgtg gagggctgcc atgcccactg ccctccaggg 3540
aaaatcctgg atgagctttt gcagacctgc gttgacctg aagactgtcc agtgtgtgag 3600
gtggctggcc ggcgttttgc ctgaggaaag aaagtcacct tgaatcccag tgacctgag 3660
cactgccaga tttgccactg tgatgttgtc aaacctacct gtgaagcctg ccaggagccg 3720
ggaggcctgg tggcgctcc cacagatgcc ccggtgagcc ccaccactct gtatgtggag 3780
gacatctcgg aaccgccgtt gcacgatttc tactgcagca ggctactgga cctggtcttc 3840
ctgctggatg gctcctccag gctgtccgag gctgagtttg aagtgtgaa ggcctttgtg 3900
gtggacatga tggagcggct gcgcctctcc cagaagtggg tccgcgtggc cgtggtggag 3960
taccacgacg gctcccacgc ctacatcggg ctcaaggacc ggaagcgacc gtcagagctg 4020
cggcgcattg ccagccaggt gaagtatgcg ggcagccagg tggcctccac cagcgaggtc 4080
ttgaaataca cactgttcca aatcttcagc aagatcgacc gccctgaagc ctccgcctc 4140
gccctgctcc tgatggccag ccaggagccc caacgatgt cccggaactt tgtccgtac 4200
gtccagggcc tgaagaagaa gaaggtcatt gtgatcccg tgggcattgg gccccatgcc 4260
aacctcaagc agatccgcct catcgagaag caggcccctg agaacaaggc ctctgtgctg 4320
agcagtgtgg atgagctgga gcagcaaagg gacgagatcg ttagctacct ctgtgacctt 4380
gccccgaag cccctcctcc tactctgccc cccgacatgg cacaagtcac tgtgggcccg 4440
gggctcttgg gggtttcgac cctggggccc aagaggaact ccatggttct ggatgtggcg 4500
ttcgtcctgg aaggatcgga caaaattggt gaagccgact tcaacaggag caaggagttc 4560
atggaggagg tgattcagcg gatggatgtg ggccaggaca gcatccactg cacggtgctg 4620
cagtactcct acatggtgac cgtggagtac cccttcagcg aggcacagtc caaaggggac 4680
atcctgcagc ggggtgcgaga gatccgctac cagggcggca acaggaccaa cactgggctg 4740
gccctgcggt acctctctga ccacagcttc ttggtcagcc agggtgaccg ggagcaggcg 4800
cccaacctgg tctacatggt caccggaaat cctgcctctg atgagatcaa gaggtgcct 4860
ggagacatcc aggtggtgcc cattggagtg ggccctaatt ccaacgtgca ggagctggag 4920
aggattggct ggcccaatgc ccctatctc atccaggact ttgagacgct cccccgagag 4980
gctcctgacc tgggtgctga gaggtgctgc tccggagagg ggctgcagat cccaccctc 5040
tcccctgcac ctgactgcag ccagcccctg gacgtgatec ttctcctgga tggctcctcc 5100
agtttcccag cttcttattt tgatgaaatg aagagtttcg ccaaggcttt catttcaaaa 5160
gccaatatag ggctcgtct cactcaggtg tcagtgtgc agtatggaag catcaccacc 5220
attgacgtgc catggaacgt ggtcccggag aaagcccatt tgctgagcct tgtggacgtc 5280
atgcagcggg agggaggccc cagccaaatc ggggatgcct tgggctttgc tgtgcgatac 5340
ttgacttcag aaatgcatgg tgccaggccg ggagcctcaa aggcggtggt catcctggtc 5400
acggacgtct ctgtggattc agtggatgca gcagctgatg ccgccaggtc caacagagtg 5460

acagtgttcc ctattggaat tggagatcgc tacgatgcag cccagctacg gatcttggca 5520
ggcccagcag gcgactccaa cgtggtgaag ctccagcgaa tcgaagacct ccctaccatg 5580
gtcaccttgg gcaattcctt cctccacaaa ctgtgctctg gatttgtagg gatttgcatg 5640
gatgaggatg ggaatgagaa gaggcccgagg gacgtctgga ccttgccaga ccagtgccac 5700
accgtgactt gccagccaga tggccagacc ttgctgaaga gtcacgggt caactgtgac 5760
cgggggctga ggccctcgtg ccctaacagc cagtccctg ttaaagtga agagacctgt 5820
ggctgccgct ggacctgccc ctgygtgtgc acaggcagct ccactcggca catcgtgacc 5880
tttgatgggc agaatttcaa gctgactggc agctgttctt atgtcctatt tcaaaacaag 5940
gagcaggacc tggaggtgat tctccataat ggtgcctgca gccctggagc aaggcagggc 6000
tgcataaaat ccatcgaggt gaagcacagt gccctctccg tcgagstgca cagtgcacatg 6060
gaggtgacgg tgaatgggag actggtctct gtctcttacg tgggtgggaa catggaagtc 6120
aacgtttatg gtgcatcat gcatgaggtc agattcaatc accttggta catcttcaca 6180
ttcactccac aaaacaatga gttccaactg cagctcagcc ccaagacttt tgcttcaaag 6240
acgtatggtc tgtgtgggat ctgtgatgag aacggagcca atgacttcat gctgagggat 6300
ggcacagtca ccacagactg gaaaacactt gttcaggaat ggactgtgca gcggccaggg 6360
cagacgtgcc agcccatcct ggaggagcag tgtcttgtcc ccgacagctc ccactgccag 6420
gtcctcctct taccactgtt tgctgaatgc cacaaggtcc tggctccagc cacattctat 6480
gccatctgcc agcaggacag ttgccaccag gagcaagtgt gtgaggtgat cgcctcttat 6540
gccacacctt gtcggaccaa cggggctctgc gttgactgga ggacacctga tttctgtgct 6600
atgtcatgcc caccatctct ggtctacaac cactgtgagc atggctgtcc ccggcactgt 6660
gatggcaacg tgagctcctg tggggacctt ccctccgaag gctgtttctg ccctccagat 6720
aaagtcatgt tggaaggcag ctgtgtccct gaagaggcct gcactcagtg cattgggtgag 6780
gatggagtcc agcaccagtt cctggaagcc tgggtcccg accaccagcc ctgtcagatc 6840
tgcacatgcc tcagcgggag gaaggtcaac tgcacaacgc agccctgccc caccgcaaaa 6900
gtccccacgt gtggcctgtg tgaagtagcc cgcctccgca agaatgcaga ccagtgtgct 6960
cccagatatg agtgtgtgtg tgaccacagt agctgtgacc tgccccagt gcctcactgt 7020
gaacgtggcc tccagcccac actgaccaac cctggcgagt gcagacccaa cttcacctgc 7080
gcctgcagga aggaggagt caaaagagt tccccacct cctgcccccc gcaccgtttg 7140
cccacccttc ggaagacca gtgctgtgat gattatgagt gtgcctgcaa ctgtgtcaac 7200
tccacagtga gctgtccct tgggtacttg gcctcaaccg ccaccaatga ctgtggctgt 7260
accacaacca cctgccttcc cgacaagggt tgtgtccacc gaagcaccat ctaccctgtg 7320
ggcagttct gggaggagg ctgcgatgtg tgcacctgca ccgacatgga ggatgccgtg 7380
atgggcctcc gcgtggccca gtgctcccag aagccctgtg aggacagctg tcggctgggc 7440
ttcacttacg ttctgcatga aggcgagtgc tgtggaaggt gcctgccatc tgctgtgag 7500
gtggtgactg gctcaccgag gggggactcc cagtcttctt ggaagagtgt cggtcccg 7560
tgggcctccc cggagaaccc ctgcctcatc aatgagtgtg tccgagtga ggaggaggtc 7620
tttatacaac aaaggaacgt ctctgcccc cagctggagg tccctgtctg cccctcgggc 7680
tttcagctga gctgtaagac ctacagctgc tgcccaagct gtcgctgtga gcgcatggag 7740
gcctgcatgc tcaatggcac tgtcattggg cccgggaaga ctgtgatgat cgatgtgtgc 7800

acgacctgcc gctgcatggt gcaggtgggg gtcattctctg gattcaagct ggagtgcagg 7860
aagaccacct gcaacccctg cccctgggt tacaaggaag aaaataacac aggtgaatgt 7920
tgtgggagat gtttgcctac ggcttgacc attcagctaa gaggaggaca gatcatgaca 7980
ctgaagcgtg atgagacgct ccaggatggc tgtgatactc acttctgcaa ggtcaatgag 8040
agaggagagt acttctggga gaagagggtc acaggctgcc caccctttga tgaacacaag 8100
tgtcttgctg agggaggtaa aattatgaaa attccaggca cctgctgtga cacatgtgag 8160
gagcctgagt gcaacgacat cactgccagg ctgcagtatg tcaaggtggg aagctgtaag 8220
tctgaagtag aggtggatat ccactactgc cagggcaaat gtgccagcaa agccatgtac 8280
tccattgaca tcaacgatgt gcaggaccag tgctcctgct gctctccgac acggacggag 8340
cccatgcagg tggccctgca ctgcaccaat ggctctgttg tgtaccatga ggttctcaat 8400
gccatggagt gcaaattgctc ccccaggaag tgcagcaagt ga 8442

<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区域

<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可以为任何天然存在的氨基酸

<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
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
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
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
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		
	1010 1015	1020
Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro		
	1025 1030	1035
Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln		
	1040 1045	1050
Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe		
	1055 1060	1065
Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val		
	1070 1075	1080
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala		
	1085 1090	1095
Cys Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln		
	1100 1105	1110
His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln		
	1115 1120	1125
Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Cys Glu		
	1130 1135	1140
Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln		
	1145 1150	1155
His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu Gly Cys		
	1160 1165	1170
His Ala His Cys Pro Pro Gly Lys Ile Leu Asp Glu Leu Leu Gln		
	1175 1180	1185
Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu Val Ala Gly		
	1190 1195	1200
Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro Ser Asp		
	1205 1210	1215
Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn Leu Thr		

1220	1225	1230
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
1250	1255	1260
Glu Pro Pro Leu His Asp	Phe Tyr Cys Ser Arg	Leu Leu Asp Leu
1265	1270	1275
Val Phe Leu Leu Asp Gly	Ser Ser Arg Leu Ser	Glu Ala Glu Phe
1280	1285	1290
Glu Val Leu Lys Ala Phe	Val Val Asp Met Met	Glu Arg Leu Arg
1295	1300	1305
Ile Ser Gln Lys Trp Val	Arg Val Ala Val Val	Glu Tyr His Asp
1310	1315	1320
Gly Ser His Ala Tyr Ile	Gly Leu Lys Asp Arg	Lys Arg Pro Ser
1325	1330	1335
Glu Leu Arg Arg Ile Ala	Ser Gln Val Lys Tyr	Ala Gly Ser Gln
1340	1345	1350
Val Ala Ser Thr Ser Glu	Val Leu Lys Tyr Thr	Leu Phe Gln Ile
1355	1360	1365
Phe Ser Lys Ile Asp Arg	Pro Glu Ala Ser Arg	Ile Ala Leu Leu
1370	1375	1380
Leu Met Ala Ser Gln Glu	Pro Gln Arg Met Ser	Arg Asn Phe Val
1385	1390	1395
Arg Tyr Val Gln Gly Leu	Lys Lys Lys Lys Val	Ile Val Ile Pro
1400	1405	1410
Val Gly Ile Gly Pro His	Ala Asn Leu Lys Gln	Ile Arg Leu Ile
1415	1420	1425
Glu Lys Gln Ala Pro Glu	Asn Lys Ala Phe Val	Leu Ser Ser Val
1430	1435	1440
Asp Glu Leu Glu Gln Gln	Arg Asp Glu Ile Val	Ser Tyr Leu Cys
1445	1450	1455
Asp Leu Ala Pro Glu Ala	Pro Pro Pro Thr Leu	Pro Pro Asp Met
1460	1465	1470
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
1490	1495	1500
Glu Gly Ser Asp Lys Ile	Gly Glu Ala Asp Phe	Asn Arg Ser Lys
1505	1510	1515

Glu Phe Met Glu Glu Val Ile Gln Arg Met Asp Val Gly Gln Asp		
1520	1525	1530
Ser Ile His Val Thr Val Leu Gln Tyr Ser Tyr Met Val Thr Val		
1535	1540	1545
Glu Tyr Pro Phe Ser Glu Ala Gln Ser Lys Gly Asp Ile Leu Gln		
1550	1555	1560
Arg Val Arg Glu Ile Arg Tyr Gln Gly Gly Asn Arg Thr Asn Thr		
1565	1570	1575
Gly Leu Ala Leu Arg Tyr Leu Ser Asp His Ser Phe Leu Val Ser		
1580	1585	1590
Gln Gly Asp Arg Glu Gln Ala Pro Asn Leu Val Tyr Met Val Thr		
1595	1600	1605
Gly Asn Pro Ala Ser Asp Glu Ile Lys Arg Leu Pro Gly Asp Ile		
1610	1615	1620
Gln Val Val Pro Ile Gly Val Gly Pro Asn Ala Asn Val Gln Glu		
1625	1630	1635
Leu Glu Arg Ile Gly Trp Pro Asn Ala Pro Ile Leu Ile Gln Asp		
1640	1645	1650
Phe Glu Thr Leu Pro Arg Glu Ala Pro Asp Leu Val Leu Gln Arg		
1655	1660	1665
Cys Cys Ser Gly Glu Gly Leu Gln Ile Pro Thr Leu Ser Pro Ala		
1670	1675	1680
Pro Asp Cys Ser Gln Pro Leu Asp Val Ile Leu Leu Leu Asp Gly		
1685	1690	1695
Ser Ser Ser Phe Pro Ala Ser Tyr Phe Asp Glu Met Lys Ser Phe		
1700	1705	1710
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		
1730	1735	1740
Pro Trp Asn Val Val Pro Glu Lys Ala His Leu Leu Ser Leu Val		
1745	1750	1755
Asp Val Met Gln Arg Glu Gly Gly Pro Ser Gln Ile Gly Asp Ala		
1760	1765	1770
Leu Gly Phe Ala Val Arg Tyr Leu Thr Ser Glu Met His Gly Ala		
1775	1780	1785
Arg Pro Gly Ala Ser Lys Ala Val Val Ile Leu Val Thr Asp Val		
1790	1795	1800
Ser Val Asp Ser Val Asp Ala Ala Ala Asp Ala Ala Arg Ser Asn		

1805	1810	1815
Arg Val Thr Val Phe Pro	Ile Gly Ile Gly Asp	Arg Tyr Asp Ala
1820	1825	1830
Ala Gln Leu Arg Ile Leu	Ala Gly Pro Ala Gly	Asp Ser Asn Val
1835	1840	1845
Val Lys Leu Gln Arg Ile	Glu Asp Leu Pro Thr	Met Val Thr Leu
1850	1855	1860
Gly Asn Ser Phe Leu His	Lys Leu Cys Ser Gly	Phe Val Arg Ile
1865	1870	1875
Cys Met Asp Glu Asp Gly	Asn Glu Lys Arg Pro	Gly Asp Val Trp
1880	1885	1890
Thr Leu Pro Asp Gln Cys	His Thr Val Thr Cys	Gln Pro Asp Gly
1895	1900	1905
Gln Thr Leu Leu Lys Ser	His Arg Val Asn Cys	Asp Arg Gly Leu
1910	1915	1920
Arg Pro Ser Cys Pro Asn	Ser Gln Ser Pro Val	Lys Val Glu Glu
1925	1930	1935
Thr Cys Gly Cys Arg Trp	Thr Cys Pro Cys Val	Cys Thr Gly Ser
1940	1945	1950
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
1970	1975	1980
Leu Glu Val Ile Leu His	Asn Gly Ala Cys Ser	Pro Gly Ala Arg
1985	1990	1995
Gln Gly Cys Met Lys Ser	Ile Glu Val Lys His	Ser Ala Leu Ser
2000	2005	2010
Val Glu Xaa His Ser Asp	Met Glu Val Thr Val	Asn Gly Arg Leu
2015	2020	2025
Val Ser Val Pro Tyr Val	Gly Gly Asn Met Glu	Val Asn Val Tyr
2030	2035	2040
Gly Ala Ile Met His Glu	Val Arg Phe Asn His	Leu Gly His Ile
2045	2050	2055
Phe Thr Phe Thr Pro Gln	Asn Asn Glu Phe Gln	Leu Gln Leu Ser
2060	2065	2070
Pro Lys Thr Phe Ala Ser	Lys Thr Tyr Gly Leu	Cys Gly Ile Cys
2075	2080	2085
Asp Glu Asn Gly Ala Asn	Asp Phe Met Leu Arg	Asp Gly Thr Val
2090	2095	2100

Thr Thr Asp Trp Lys Thr Leu Val Gln Glu Trp Thr Val Gln Arg		
2105	2110	2115
Pro Gly Gln Thr Cys Gln Pro Ile Leu Glu Glu Gln Cys Leu Val		
2120	2125	2130
Pro Asp Ser Ser His Cys Gln Val Leu Leu Leu Pro Leu Phe Ala		
2135	2140	2145
Glu Cys His Lys Val Leu Ala Pro Ala Thr Phe Tyr Ala Ile Cys		
2150	2155	2160
Gln Gln Asp Ser Cys His Gln Glu Gln Val Cys Glu Val Ile Ala		
2165	2170	2175
Ser Tyr Ala His Leu Cys Arg Thr Asn Gly Val Cys Val Asp Trp		
2180	2185	2190
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		
2210	2215	2220
Val Ser Ser Cys Gly Asp His Pro Ser Glu Gly Cys Phe Cys Pro		
2225	2230	2235
Pro Asp Lys Val Met Leu Glu Gly Ser Cys Val Pro Glu Glu Ala		
2240	2245	2250
Cys Thr Gln Cys Ile Gly Glu Asp Gly Val Gln His Gln Phe Leu		
2255	2260	2265
Glu Ala Trp Val Pro Asp His Gln Pro Cys Gln Ile Cys Thr Cys		
2270	2275	2280
Leu Ser Gly Arg Lys Val Asn Cys Thr Thr Gln Pro Cys Pro Thr		
2285	2290	2295
Ala Lys Ala Pro Thr Cys Gly Leu Cys Glu Val Ala Arg Leu Arg		
2300	2305	2310
Gln Asn Ala Asp Gln Cys Cys Pro Glu Tyr Glu Cys Val Cys Asp		
2315	2320	2325
Pro Val Ser Cys Asp Leu Pro Pro Val Pro His Cys Glu Arg Gly		
2330	2335	2340
Leu Gln Pro Thr Leu Thr Asn Pro Gly Glu Cys Arg Pro Asn Phe		
2345	2350	2355
Thr Cys Ala Cys Arg Lys Glu Glu Cys Lys Arg Val Ser Pro Pro		
2360	2365	2370
Ser Cys Pro Pro His Arg Leu Pro Thr Leu Arg Lys Thr Gln Cys		
2375	2380	2385
Cys Asp Glu Tyr Glu Cys Ala Cys Asn Cys Val Asn Ser Thr Val		

2390	2395	2400
Ser Cys Pro Leu Gly Tyr	Leu Ala Ser Thr Ala	Thr Asn Asp Cys
2405	2410	2415
Gly Cys Thr Thr Thr Thr	Cys Leu Pro Asp Lys	Val Cys Val His
2420	2425	2430
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
2450	2455	2460
Arg Val Ala Gln Cys Ser	Gln Lys Pro Cys Glu	Asp Ser Cys Arg
2465	2470	2475
Ser Gly Phe Thr Tyr Val	Leu His Glu Gly Glu	Cys Cys Gly Arg
2480	2485	2490
Cys Leu Pro Ser Ala Cys	Glu Val Val Thr Gly	Ser Pro Arg Gly
2495	2500	2505
Asp Ser Gln Ser Ser Trp	Lys Ser Val Gly Ser	Gln Trp Ala Ser
2510	2515	2520
Pro Glu Asn Pro Cys Leu	Ile Asn Glu Cys Val	Arg Val Lys Glu
2525	2530	2535
Glu Val Phe Ile Gln Gln	Arg Asn Val Ser Cys	Pro Gln Leu Glu
2540	2545	2550
Val Pro Val Cys Pro Ser	Gly Phe Gln Leu Ser	Cys Lys Thr Ser
2555	2560	2565
Ala Cys Cys Pro Ser Cys	Arg Cys Glu Arg Met	Glu Ala Cys Met
2570	2575	2580
Leu Asn Gly Thr Val Ile	Gly Pro Gly Lys Thr	Val Met Ile Asp
2585	2590	2595
Val Cys Thr Thr Cys Arg	Cys Met Val Gln Val	Gly Val Ile Ser
2600	2605	2610
Gly Phe Lys Leu Glu Cys	Arg Lys Thr Thr Cys	Asn Pro Cys Pro
2615	2620	2625
Leu Gly Tyr Lys Glu Glu	Asn Asn Thr Gly Glu	Cys Cys Gly Arg
2630	2635	2640
Cys Leu Pro Thr Ala Cys	Thr Ile Gln Leu Arg	Gly Gly Gln Ile
2645	2650	2655
Met Thr Leu Lys Arg Asp	Glu Thr Leu Gln Asp	Gly Cys Asp Thr
2660	2665	2670
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		
2690	2695	2700
Glu Gly Gly Lys Ile Met Lys Ile Pro Gly Thr Cys Cys Asp Thr		
2705	2710	2715
Cys Glu Glu Pro Glu Cys Asn Asp Ile Thr Ala Arg Leu Gln Tyr		
2720	2725	2730
Val Lys Val Gly Ser Cys Lys Ser Glu Val Glu Val Asp Ile His		
2735	2740	2745
Tyr Cys Gln Gly Lys Cys Ala Ser Lys Ala Met Tyr Ser Ile Asp		
2750	2755	2760
Ile Asn Asp Val Gln Asp Gln Cys Ser Cys Cys Ser Pro Thr Arg		
2765	2770	2775
Thr Glu Pro Met Gln Val Ala Leu His Cys Thr Asn Gly Ser Val		
2780	2785	2790
Val Tyr His Glu Val Leu Asn Ala Met Glu Cys Lys Cys Ser Pro		
2795	2800	2805
Arg Lys Cys Ser Lys		
2810		
<210> 3		
<211> 19		
<212> PRT		
<213> 智人		
<400> 3		
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10 15
Cys Phe Ser		
<210> 4		
<211> 2332		
<212> PRT		
<213> 智人		
<400> 4		
Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr		
1	5	10 15
Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro		
	20	25 30
Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys		
	35	40 45
Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys Pro		
50	55	60

Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln	Ala	Glu	Val
65					70					75					80
Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser	His	Pro	Val
			85						90					95	
Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser	Glu	Gly	Ala
			100						105					110	
Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp	Asp	Lys	Val
		115					120						125		
Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu	Lys	Glu	Asn
		130				135					140				
Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser	Tyr	Leu	Ser
145					150					155					160
His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile	Gly	Ala	Leu
				165					170					175	
Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr	Gln	Thr	Leu
			180						185					190	
His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly	Lys	Ser	Trp
		195					200						205		
His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp	Ala	Ala	Ser
		210					215				220				
Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr	Val	Asn	Arg
225					230					235					240
Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val	Tyr	Trp	His
			245						250					255	
Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile	Phe	Leu	Glu
			260						265					270	
Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser	Leu	Glu	Ile
		275					280						285		
Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met	Asp	Leu	Gly
		290					295					300			
Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His	Asp	Gly	Met
305					310					315					320
Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro	Gln	Leu	Arg
			325						330					335	
Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	Leu	Thr	Asp
			340						345				350		
Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	Phe
		355					360					365			
Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His

370	375	380
Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu		
385	390	395
Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly Pro		400
405	410	415
Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr Thr		
420	425	430
Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly Ile		
435	440	445
Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile Ile		
450	455	460
Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile		
465	470	475
Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys		
485	490	495
His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr Lys		
500	505	510
Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg Cys		
515	520	525
Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu Ala		
530	535	540
Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp		
545	550	555
Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe		
565	570	575
Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln		
580	585	590
Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe		
595	600	605
Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser		
610	615	620
Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu		
625	630	635
Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr		
645	650	655
Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro		
660	665	670
Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp		
675	680	685

Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala
690							695					700			
Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu
705						710				715					720
Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys	Asn	Asn	Ala
				725					730					735	
Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Ser	Arg	His	Pro	Ser	Thr	Arg
			740					745					750		
Gln	Lys	Gln	Phe	Asn	Ala	Thr	Thr	Ile	Pro	Glu	Asn	Asp	Ile	Glu	Lys
		755					760					765			
Thr	Asp	Pro	Trp	Phe	Ala	His	Arg	Thr	Pro	Met	Pro	Lys	Ile	Gln	Asn
770						775					780				
Val	Ser	Ser	Ser	Asp	Leu	Leu	Met	Leu	Leu	Arg	Gln	Ser	Pro	Thr	Pro
785					790					795					800
His	Gly	Leu	Ser	Leu	Ser	Asp	Leu	Gln	Glu	Ala	Lys	Tyr	Glu	Thr	Phe
				805					810					815	
Ser	Asp	Asp	Pro	Ser	Pro	Gly	Ala	Ile	Asp	Ser	Asn	Asn	Ser	Leu	Ser
			820					825					830		
Glu	Met	Thr	His	Phe	Arg	Pro	Gln	Leu	His	His	Ser	Gly	Asp	Met	Val
		835					840					845			
Phe	Thr	Pro	Glu	Ser	Gly	Leu	Gln	Leu	Arg	Leu	Asn	Glu	Lys	Leu	Gly
850						855					860				
Thr	Thr	Ala	Ala	Thr	Glu	Leu	Lys	Lys	Leu	Asp	Phe	Lys	Val	Ser	Ser
865					870					875					880
Thr	Ser	Asn	Asn	Leu	Ile	Ser	Thr	Ile	Pro	Ser	Asp	Asn	Leu	Ala	Ala
				885					890					895	
Gly	Thr	Asp	Asn	Thr	Ser	Ser	Leu	Gly	Pro	Pro	Ser	Met	Pro	Val	His
			900					905					910		
Tyr	Asp	Ser	Gln	Leu	Asp	Thr	Thr	Leu	Phe	Gly	Lys	Lys	Ser	Ser	Pro
		915					920					925			
Leu	Thr	Glu	Ser	Gly	Gly	Pro	Leu	Ser	Leu	Ser	Glu	Glu	Asn	Asn	Asp
930						935					940				
Ser	Lys	Leu	Leu	Glu	Ser	Gly	Leu	Met	Asn	Ser	Gln	Glu	Ser	Ser	Trp
945					950					955					960
Gly	Lys	Asn	Val	Ser	Ser	Thr	Glu	Ser	Gly	Arg	Leu	Phe	Lys	Gly	Lys
				965					970					975	
Arg	Ala	His	Gly	Pro	Ala	Leu	Leu	Thr	Lys	Asp	Asn	Ala	Leu	Phe	Lys
			980					985					990		
Val	Ser	Ile	Ser	Leu	Leu	Lys	Thr	Asn	Lys	Thr	Ser	Asn	Asn	Ser	Ala

995	1000	1005
Thr Asn Arg Lys Thr His	Ile Asp Gly Pro Ser	Leu Leu Ile Glu
1010	1015	1020
Asn Ser Pro Ser Val Trp	Gln Asn Ile Leu Glu	Ser Asp Thr Glu
1025	1030	1035
Phe Lys Lys Val Thr Pro	Leu Ile His Asp Arg	Met Leu Met Asp
1040	1045	1050
Lys Asn Ala Thr Ala Leu	Arg Leu Asn His Met	Ser Asn Lys Thr
1055	1060	1065
Thr Ser Ser Lys Asn Met	Glu Met Val Gln Gln	Lys Lys Glu Gly
1070	1075	1080
Pro Ile Pro Pro Asp Ala	Gln Asn Pro Asp Met	Ser Phe Phe Lys
1085	1090	1095
Met Leu Phe Leu Pro Glu	Ser Ala Arg Trp Ile	Gln Arg Thr His
1100	1105	1110
Gly Lys Asn Ser Leu Asn	Ser Gly Gln Gly Pro	Ser Pro Lys Gln
1115	1120	1125
Leu Val Ser Leu Gly Pro	Glu Lys Ser Val Glu	Gly Gln Asn Phe
1130	1135	1140
Leu Ser Glu Lys Asn Lys	Val Val Val Gly Lys	Gly Glu Phe Thr
1145	1150	1155
Lys Asp Val Gly Leu Lys	Glu Met Val Phe Pro	Ser Ser Arg Asn
1160	1165	1170
Leu Phe Leu Thr Asn Leu	Asp Asn Leu His Glu	Asn Asn Thr His
1175	1180	1185
Asn Gln Glu Lys Lys Ile	Gln Glu Glu Ile Glu	Lys Lys Glu Thr
1190	1195	1200
Leu Ile Gln Glu Asn Val	Val Leu Pro Gln Ile	His Thr Val Thr
1205	1210	1215
Gly Thr Lys Asn Phe Met	Lys Asn Leu Phe Leu	Leu Ser Thr Arg
1220	1225	1230
Gln Asn Val Glu Gly Ser	Tyr Asp Gly Ala Tyr	Ala Pro Val Leu
1235	1240	1245
Gln Asp Phe Arg Ser Leu	Asn Asp Ser Thr Asn	Arg Thr Lys Lys
1250	1255	1260
His Thr Ala His Phe Ser	Lys Lys Gly Glu Glu	Glu Asn Leu Glu
1265	1270	1275
Gly Leu Gly Asn Gln Thr	Lys Gln Ile Val Glu	Lys Tyr Ala Cys
1280	1285	1290

Thr Thr Arg Ile Ser Pro Asn Thr Ser Gln Gln Asn Phe Val Thr		
1295	1300	1305
Gln Arg Ser Lys Arg Ala Leu Lys Gln Phe Arg Leu Pro Leu Glu		
1310	1315	1320
Glu Thr Glu Leu Glu Lys Arg Ile Ile Val Asp Asp Thr Ser Thr		
1325	1330	1335
Gln Trp Ser Lys Asn Met Lys His Leu Thr Pro Ser Thr Leu Thr		
1340	1345	1350
Gln Ile Asp Tyr Asn Glu Lys Glu Lys Gly Ala Ile Thr Gln Ser		
1355	1360	1365
Pro Leu Ser Asp Cys Leu Thr Arg Ser His Ser Ile Pro Gln Ala		
1370	1375	1380
Asn Arg Ser Pro Leu Pro Ile Ala Lys Val Ser Ser Phe Pro Ser		
1385	1390	1395
Ile Arg Pro Ile Tyr Leu Thr Arg Val Leu Phe Gln Asp Asn Ser		
1400	1405	1410
Ser His Leu Pro Ala Ala Ser Tyr Arg Lys Lys Asp Ser Gly Val		
1415	1420	1425
Gln Glu Ser Ser His Phe Leu Gln Gly Ala Lys Lys Asn Asn Leu		
1430	1435	1440
Ser Leu Ala Ile Leu Thr Leu Glu Met Thr Gly Asp Gln Arg Glu		
1445	1450	1455
Val Gly Ser Leu Gly Thr Ser Ala Thr Asn Ser Val Thr Tyr Lys		
1460	1465	1470
Lys Val Glu Asn Thr Val Leu Pro Lys Pro Asp Leu Pro Lys Thr		
1475	1480	1485
Ser Gly Lys Val Glu Leu Leu Pro Lys Val His Ile Tyr Gln Lys		
1490	1495	1500
Asp Leu Phe Pro Thr Glu Thr Ser Asn Gly Ser Pro Gly His Leu		
1505	1510	1515
Asp Leu Val Glu Gly Ser Leu Leu Gln Gly Thr Glu Gly Ala Ile		
1520	1525	1530
Lys Trp Asn Glu Ala Asn Arg Pro Gly Lys Val Pro Phe Leu Arg		
1535	1540	1545
Val Ala Thr Glu Ser Ser Ala Lys Thr Pro Ser Lys Leu Leu Asp		
1550	1555	1560
Pro Leu Ala Trp Asp Asn His Tyr Gly Thr Gln Ile Pro Lys Glu		
1565	1570	1575
Glu Trp Lys Ser Gln Glu Lys Ser Pro Glu Lys Thr Ala Phe Lys		

1580	1585	1590
Lys Lys Asp Thr Ile Leu	Ser Leu Asn Ala Cys	Glu Ser Asn His
1595	1600	1605
Ala Ile Ala Ala Ile Asn	Glu Gly Gln Asn Lys	Pro Glu Ile Glu
1610	1615	1620
Val Thr Trp Ala Lys Gln	Gly Arg Thr Glu Arg	Leu Cys Ser Gln
1625	1630	1635
Asn Pro Pro Val Leu Lys	Arg His Gln Arg Glu	Ile Thr Arg Thr
1640	1645	1650
Thr Leu Gln Ser Asp Gln	Glu Glu Ile Asp Tyr	Asp Asp Thr Ile
1655	1660	1665
Ser Val Glu Met Lys Lys	Glu Asp Phe Asp Ile	Tyr Asp Glu Asp
1670	1675	1680
Glu Asn Gln Ser Pro Arg	Ser Phe Gln Lys Lys	Thr Arg His Tyr
1685	1690	1695
Phe Ile Ala Ala Val Glu	Arg Leu Trp Asp Tyr	Gly Met Ser Ser
1700	1705	1710
Ser Pro His Val Leu Arg	Asn Arg Ala Gln Ser	Gly Ser Val Pro
1715	1720	1725
Gln Phe Lys Lys Val Val	Phe Gln Glu Phe Thr	Asp Gly Ser Phe
1730	1735	1740
Thr Gln Pro Leu Tyr Arg	Gly Glu Leu Asn Glu	His Leu Gly Leu
1745	1750	1755
Leu Gly Pro Tyr Ile Arg	Ala Glu Val Glu Asp	Asn Ile Met Val
1760	1765	1770
Thr Phe Arg Asn Gln Ala	Ser Arg Pro Tyr Ser	Phe Tyr Ser Ser
1775	1780	1785
Leu Ile Ser Tyr Glu Glu	Asp Gln Arg Gln Gly	Ala Glu Pro Arg
1790	1795	1800
Lys Asn Phe Val Lys Pro	Asn Glu Thr Lys Thr	Tyr Phe Trp Lys
1805	1810	1815
Val Gln His His Met Ala	Pro Thr Lys Asp Glu	Phe Asp Cys Lys
1820	1825	1830
Ala Trp Ala Tyr Phe Ser	Asp Val Asp Leu Glu	Lys Asp Val His
1835	1840	1845
Ser Gly Leu Ile Gly Pro	Leu Leu Val Cys His	Thr Asn Thr Leu
1850	1855	1860
Asn Pro Ala His Gly Arg	Gln Val Thr Val Gln	Glu Phe Ala Leu
1865	1870	1875

Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu		
1880	1885	1890
Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu		
1895	1900	1905
Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly		
1910	1915	1920
Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln		
1925	1930	1935
Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile		
1940	1945	1950
His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys		
1955	1960	1965
Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe		
1970	1975	1980
Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val		
1985	1990	1995
Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu		
2000	2005	2010
Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala		
2015	2020	2025
Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr		
2030	2035	2040
Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser		
2045	2050	2055
Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val		
2060	2065	2070
Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly		
2075	2080	2085
Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile		
2090	2095	2100
Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn		
2105	2110	2115
Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser		
2120	2125	2130
Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr		
2135	2140	2145
Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg		
2150	2155	2160
Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu		

2165	2170	2175
Gly Met Glu Ser Lys Ala	Ile Ser Asp Ala Gln	Ile Thr Ala Ser
2180	2185	2190
Ser Tyr Phe Thr Asn Met	Phe Ala Thr Trp Ser	Pro Ser Lys Ala
2195	2200	2205
Arg Leu His Leu Gln Gly	Arg Ser Asn Ala Trp	Arg Pro Gln Val
2210	2215	2220
Asn Asn Pro Lys Glu Trp	Leu Gln Val Asp Phe	Gln Lys Thr Met
2225	2230	2235
Lys Val Thr Gly Val Thr	Thr Gln Gly Val Lys	Ser Leu Leu Thr
2240	2245	2250
Ser Met Tyr Val Lys Glu	Phe Leu Ile Ser Ser	Ser Gln Asp Gly
2255	2260	2265
His Gln Trp Thr Leu Phe	Phe Gln Asn Gly Lys	Val Lys Val Phe
2270	2275	2280
Gln Gly Asn Gln Asp Ser	Phe Thr Pro Val Val	Asn Ser Leu Asp
2285	2290	2295
Pro Pro Leu Leu Thr Arg	Tyr Leu Arg Ile His	Pro Gln Ser Trp
2300	2305	2310
Val His Gln Ile Ala Leu	Arg Met Glu Val Leu	Gly Cys Glu Ala
2315	2320	2325
Gln Asp Leu Tyr		
2330		

<210> 5

<211> 7053

<212> DNA

<213> 智人

<400> 5

atgcaaatag agctctccac ctgcttcttt ctgtgccttt tgcgattctg ctttagtgcc 60
accagaagat actacctggg tgcagtggaa ctgtcatggg actatatgca aagtgatctc 120
ggtagagctgc ctgtggacgc aagatttctt cctagagtgc caaatcttt tccattcaac 180
acctcagtcg tgtacaaaaa gactctgttt gtagaattca cggatcacct tttcaacatc 240
gctaagccaa ggccaccctg gatgggtctg ctaggtccta ccatccaggc tgaggtttat 300
gatacagtg tcatctacact taagaacatg gcttcccac ctgtcagtc tcatgctgtt 360
gggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtcaaagg 420
gagaaagaag atgataaagt cttccctggg ggaagccata catatgtctg gcaggtcctg 480
aaagagaatg gtccaatggc ctctgaccca ctgtgcctta cctactcata tctttctcat 540
gtggacctgg taaaagactt gaattcaggc ctcatggag ccctactagt atgtagagaa 600
gggagtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta 660

```

tttgatgaag ggaaaagttg gcactcagaa acaaagaact ccttgatgca ggatagggat 720
gctgcatctg ctcgggcctg gcctaaaatg cacacagtca atggttatgt aaacaggtct 780
ctgccagggtc tgattggatg ccacaggaaa tcagtctatt ggcatgtgat tggaatgggc 840
accactcctg aagtgcactc aatattcctc gaaggtcaca ctttcttgt gaggaacct 900
cgccaggcgt ccttggaat ctgcgaata actttccta ctgctcaaac actcttgatg 960
gaccttgac agtttctact gttttgtcat atctcttccc accaacaatga tggcatggaa 1020
gcttatgtca aagtagacag ctgtccagag gaacccaac tacgaatgaa aaataatgaa 1080
gaagcggaag actatgatga tgatcttact gattctgaaa tggatgtggt caggtttgat 1140
gatgacaact ctcttcctt tatccaaatt cgctcagttg ccaagaagca tcctaaaact 1200
tggttacatt acattgctgc tgaagaggag gactgggact atgctccctt agtctcgc 1260
cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg 1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa ctttaagac tcgtgaagct 1380
attcagcatg aatcaggaat cttgggacct ttactttatg gggaagttgg agacacactg 1440
ttgattatat ttaagaatca agcaagcaga ccatataaca tctaccctca cggaatcact 1500
gatgtccgtc ctttgtattc aaggagatta caaaaggtg taaaacattt gaaggatttt 1560
ccaattctgc caggagaaat attcaaatat aaatggacag tgactgtaga agatgggcca 1620
actaaatcag atcctcggtg cctgaccgcg tattactcta gtttcgttaa tatggagaga 1680
gatctagctt caggactcat tggccctctc ctcatctgct acaaagaatc tgtagatcaa 1740
agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag 1800
aaccgaagct ggtacctcac agagaatata caacgcttcc tcccaatcc agctggagt 1860
cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatgtt 1920
tttgatagtt tgcagttgtc agtttgtttg catgaggtgg catactggta cattctaagc 1980
attggagcac agactgactt ctttctgtc ttcttctctg gatatacctt caaacacaaa 2040
atggtctatg aagacacact caccctattc ccattctcag gagaaactgt cttcatgtcg 2100
atggaaaacc caggtctatg gattctgggg tgccacaact cagactttcg gaacagaggc 2160
atgaccgcct tactgaaggt ttctagtgtg gacaagaaca ctggtgatta ttacgaggac 2220
agttatgaag atatttcagc atacttgctg agtaaaaaca atgccattga accaagaagc 2280
ttctcccaga attcaagaca ccctagcact aggcaaaagc aatttaatgc caccacaatt 2340
ccagaaaatg acatagagaa gactgacct tggtttgcac acagaacacc tatgcctaaa 2400
atacaaatg tctcctctag tgatttgttg atgctcttgac gagagagtc tactccacat 2460
gggctatcct tatctgatct ccaagaagcc aaatatgaga ctttttctga tgatccatca 2520
cctggagcaa tagacagtaa taacagcctg tctgaaatga cacacttcag gccacagctc 2580
catcacagtg gggacatggt atttaccct gagtcaggcc tccaattaag attaaatgag 2640
aaactgggga caactgcagc aacagagttg aagaaacttg atttcaaagt ttctagtaca 2700
tcaaataatc tgatttcaac aattccatca gacaatttg cagcaggtac tgataataca 2760
agttccttag gaccccaag tatgccagtt cattatgata gtcaattaga taccactcta 2820
tttggaacaaa agtcatctcc cttactgag tctggtggac ctctgagctt gagtgaagaa 2880
aataatgatt caaagttgtt agaatcaggt ttaatgaata gccaaagaaag ttcatgggga 2940
aaaaatgtat cgtcaacaga gagtggtagg ttatttaaag ggaaaagagc tcatggacct 3000

```

gctttgttga ctaaagataa tgccttattc aaagttagca tctctttgtt aaagacaaac 3060
 aaaacttcca ataattcagc aactaataga aagactcaca ttgatggccc atcattatta 3120
 attgagaata gtccatcagt ctggcaaaat atattagaaa gtgacactga gtttaaaaaa 3180
 gtgacacctt tgattcatga cagaatgctt atggacaaaa atgctacagc tttgaggcta 3240
 aatcatatgt caaataaaac tacttcatca aaaaacatgg aaatgggtcca acagaaaaaa 3300
 gagggcccca ttccaccaga tgcacaaaat ccagatatgt cgttctttaa gatgctattc 3360
 ttgccagaat cagcaaggtg gatacaaagg actcatggaa agaactctct gaactctggg 3420
 caaggcccca gtccaaagca attagtatcc ttaggaccag aaaaatctgt ggaaggctcag 3480
 aatttcttgt ctgagaaaaa caaagtggta gtaggaaagg gtgaatttac aaaggacgta 3540
 ggactcaaaag agatggtttt tccaagcagc agaaacctat ttcttactaa cttggataat 3600
 ttacatgaaa ataatacaca caatcaagaa aaaaaattc aggaagaaat agaaaagaag 3660
 gaaacattaa tccaagagaa tgtagttttg cctcagatac atacagtac tggcactaag 3720
 aatttcatga agaacctttt ctactgagc actaggcaaa atgtagaagg ttcatatgac 3780
 ggggcatatg ctccagtact tcaagatttt aggtcattaa atgattcaac aaatagaaca 3840
 aagaaacaca cagctcattt ctcaaaaaaa ggggaggaag aaaacttga aggcttggga 3900
 aatcaaacca agcaaatgt agagaaatat gcatgcacca caaggatatc tcctaataca 3960
 agccagcaga attttgtcac gcaacgtagt aagagagctt tgaaacaatt cagactccca 4020
 ctagaagaaa cagaacttga aaaaaggata attgtggatg acacctcaac ccagtgtgcc 4080
 aaaaacatga aacatttgac cccgagcacc ctcacacaga tagactacaa tgagaaggag 4140
 aaaggggcca ttactcagtc tcccttatca gattgcctta cgaggagtca tagcatccct 4200
 caagcaaata gatctccatt acccattgca aaggtatcat catttccatc tattagacct 4260
 atatatctga ccagggtcct attccaagac aactcttctc atcttccagc agcatcttat 4320
 agaaagaaag attctggggt ccaagaaagc agtcatttct tacaaggagc caaaaaaaat 4380
 aacctttctt tagccattct aaccttggag atgactgggt atcaaagaga ggttggctcc 4440
 ctggggacaa gtgccacaaa ttcagtcaca tacaagaaag ttgagaacac tgttctccc 4500
 aaaccagact tgcccaaaac atctggcaaa gttgaattgc ttccaaaagt tcacatttat 4560
 cagaaggacc tattccctac ggaaactagc aatgggtctc ctggccatct ggatctcgtg 4620
 gaaggagacc ttcttcaggg aacagaggga gcgattaagt ggaatgaagc aaacagacct 4680
 ggaaaagttc ctttctgag agtagcaaca gaaagctctg caaagactcc ctccaagcta 4740
 ttggatcctc ttgcttggga taaccactat ggtactcaga taccaaaaga agagtggaaa 4800
 tccaagaga agtcaccaga aaaaacagct tttaagaaaa aggataccat tttgtccctg 4860
 aacgcttggtg aaagcaatca tgcaatagca gcaataaatg agggacaaaa taagcccga 4920
 atagaagtca cctgggcaaa gcaaggtagg actgaaaggc tgtgctctca aaaccacca 4980
 gtcttgaac gccatcaacg ggaaataact cgtactactc ttcagtcaga tcaagaggaa 5040
 attgactatg atgataccat atcagttgaa atgaagaagg aagattttga catttatgat 5100
 gaggatgaaa atcagagccc ccgcagcttt caaaagaaaa cagcacta ttttattgct 5160
 gcagtggaga ggctctggga ttatgggatg agtagctccc cacatgttct aagaaacagg 5220
 gctcagagtg gcagtgtccc tcagttcaag aaagttgttt tccaggaatt tactgatggc 5280
 tcctttactc agcccttata ccgtggagaa ctaaatgaac atttgggact cctggggcca 5340

tatataagag cagaagttga agataatatc atggtaactt tcagaaatca ggcctctcgt 5400
 ccctatttcct tctatttctag ccttatttct tatgaggaag atcagaggca aggagcagaa 5460
 cctagaaaaa actttgtcaa gcctaataa accaaaactt acttttggaa agtgcaacat 5520
 catatggcac ccactaaaga tgagtttgac tgcaaagcct gggcttattt ctctgatgtt 5580
 gacctggaaa aagatgtgca ctcaggcctg attggacccc ttctggctctg ccacactaac 5640
 acactgaacc ctgctcatgg gagacaagtg acagtacagg aatttgctct gtttttcacc 5700
 atctttgatg agaccaaaaag ctggtacttc actgaaaata tggaaagaaa ctgcagggt 5760
 ccctgcaata tccagatgga agatcccact tttaaagaga attatcgctt ccatgcaatc 5820
 aatggctaca taatggatac actacctggc ttagtaatgg ctcaggatca aaggattcga 5880
 tggtatctgc tcagcatggg cagcaatgaa aacatccatt ctattcattt cagtggacat 5940
 gtgttctactg tacgaaaaaa agaggagtat aaaatggcac tgtacaatct ctatccaggt 6000
 gtttttgaga cagtggaaat gttaccatcc aaagctggaa tttggcgggt ggaatgcctt 6060
 attggcgagc atctacatgc tgggatgagc acacttttct tgggtgtacag caataagtgt 6120
 cagactcccc tgggaatggc ttctggacac attagagatt ttcagattac agcttcagga 6180
 caatatggac agtgggcccc aaagctggcc agacttcatt attccgcatc aatcaatgcc 6240
 tggagcacca aggagccctt ttcttgatc aaggtggatc tgttggcacc aatgattatt 6300
 cacggcatca agaccagggt tgcccgctcag aagttctcca gcctctacat ctctcagttt 6360
 atcatcatgt atagtcttga tgggaagaag tggcagactt atcgaggaaa ttccactgga 6420
 accttaatgg tcttctttgg caatgtggat tcatctggga taaaacacaa tatttttaac 6480
 cctccaatta ttgctcgata catccgtttg cacccaactc attatagcat tcgcagcact 6540
 ctctcgatgg agttgatggg ctgtgattta aatagttgca gcatgccatt gggaatggag 6600
 agtaaagcaa tatcagatgc acagattact gcttcactct actttaccaa tatgtttgcc 6660
 acctgggtctc cttcaaaaagc tcgacttcac ctccaaggga ggagtaatgc ctggagacct 6720
 caggtgaata atccaaaaga gtggctgcaa gtggacttcc agaagacaat gaaagtcaca 6780
 ggagtaacta ctcagggagt aaaatctctg cttaccagca tgtatgtgaa ggagttcctc 6840
 atctccagca gtcaagatgg ccatcagtgg actctctttt ttcagaatgg caaagtaaag 6900
 gtttttcagg gaaatcaaga ctcttcaca cctgtggtga actctctaga cccaccgtta 6960
 ctgactcgct accttcgaat tcacccccag agttgggtgc accagattgc cctgaggatg 7020
 gaggttctgg gctgcgaggc acaggacctc tac 7053

<210> 6

<211> 1438

<212> PRT

<213> 人工序列

<220>

<223> BDDD FVIII

<400> 6

Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr

1

5

10

15

Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro

20	25	30
Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys		
35	40	45
Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys Pro		
50	55	60
Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val		
65	70	75
Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val		
85	90	95
Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala		
100	105	110
Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val		
115	120	125
Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn		
130	135	140
Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser		
145	150	155
His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu		
165	170	175
Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu		
180	185	190
His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp		
195	200	205
His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser		
210	215	220
Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg		
225	230	235
Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His		
245	250	255
Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu		
260	265	270
Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile		
275	280	285
Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly		
290	295	300
Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met		
305	310	315
Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg		
325	330	335

Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	Leu	Thr	Asp				
																340	345	350	
Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	Phe				
																355	360	365	
Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His				
																370	375	380	
Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	Leu				
																385	390	395	400
Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro				
																405	410	415	
Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr				
																420	425	430	
Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile				
																435	440	445	
Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile				
																450	455	460	
Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly	Ile				
																465	470	475	480
Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val	Lys				
																485	490	495	
His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys				
																500	505	510	
Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys				
																515	520	525	
Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala				
																530	535	540	
Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp				
																545	550	555	560
Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe				
																565	570	575	
Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile	Gln				
																580	585	590	
Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu	Phe				
																595	600	605	
Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp	Ser				
																610	615	620	
Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile	Leu				
																625	630	635	640
Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr				

				645					650					655			
Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro		
				660					665					670			
Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp		
				675					680					685			
Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala		
				690					695					700			
Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu		
705									710					715			720
Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys	Asn	Asn	Ala		
				725					730					735			
Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Pro	Pro	Val	Leu	Lys	Arg	His		
				740					745					750			
Gln	Arg	Glu	Ile	Thr	Arg	Thr	Thr	Leu	Gln	Ser	Asp	Gln	Glu	Glu	Ile		
				755					760					765			
Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys	Lys	Glu	Asp	Phe	Asp		
				770					775					780			
Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro	Arg	Ser	Phe	Gln	Lys	Lys		
785									790					795			800
Thr	Arg	His	Tyr	Phe	Ile	Ala	Ala	Val	Glu	Arg	Leu	Trp	Asp	Tyr	Gly		
				805					810					815			
Met	Ser	Ser	Ser	Pro	His	Val	Leu	Arg	Asn	Arg	Ala	Gln	Ser	Gly	Ser		
				820					825					830			
Val	Pro	Gln	Phe	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr	Asp	Gly	Ser		
				835					840					845			
Phe	Thr	Gln	Pro	Leu	Tyr	Arg	Gly	Glu	Leu	Asn	Glu	His	Leu	Gly	Leu		
				850					855					860			
Leu	Gly	Pro	Tyr	Ile	Arg	Ala	Glu	Val	Glu	Asp	Asn	Ile	Met	Val	Thr		
865									870					875			880
Phe	Arg	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Ser	Phe	Tyr	Ser	Ser	Leu	Ile		
				885					890					895			
Ser	Tyr	Glu	Glu	Asp	Gln	Arg	Gln	Gly	Ala	Glu	Pro	Arg	Lys	Asn	Phe		
				900					905					910			
Val	Lys	Pro	Asn	Glu	Thr	Lys	Thr	Tyr	Phe	Trp	Lys	Val	Gln	His	His		
				915					920					925			
Met	Ala	Pro	Thr	Lys	Asp	Glu	Phe	Asp	Cys	Lys	Ala	Trp	Ala	Tyr	Phe		
				930					935					940			
Ser	Asp	Val	Asp	Leu	Glu	Lys	Asp	Val	His	Ser	Gly	Leu	Ile	Gly	Pro		
945									950					955			960

Leu	Leu	Val	Cys	His	Thr	Asn	Thr	Leu	Asn	Pro	Ala	His	Gly	Arg	Gln	
Val	Thr	Val	Gln	Glu	Phe	Ala	Leu	Phe	Phe	Thr	Ile	Phe	Asp	Glu	Thr	
Lys	Ser	Trp	Tyr	Phe	Thr	Glu	Asn	Met	Glu	Arg	Asn	Cys	Arg	Ala	Pro	
Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	Thr	Phe	Lys	Glu	Asn	Tyr	Arg		
Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	Met	Asp	Thr	Leu	Pro	Gly	Leu		
Val	Met	Ala	Gln	Asp	Gln	Arg	Ile	Arg	Trp	Tyr	Leu	Leu	Ser	Met		
Gly	Ser	Asn	Glu	Asn	Ile	His	Ser	Ile	His	Phe	Ser	Gly	His	Val		
Phe	Thr	Val	Arg	Lys	Lys	Glu	Glu	Tyr	Lys	Met	Ala	Leu	Tyr	Asn		
Leu	Tyr	Pro	Gly	Val	Phe	Glu	Thr	Val	Glu	Met	Leu	Pro	Ser	Lys		
Ala	Gly	Ile	Trp	Arg	Val	Glu	Cys	Leu	Ile	Gly	Glu	His	Leu	His		
Ala	Gly	Met	Ser	Thr	Leu	Phe	Leu	Val	Tyr	Ser	Asn	Lys	Cys	Gln		
Thr	Pro	Leu	Gly	Met	Ala	Ser	Gly	His	Ile	Arg	Asp	Phe	Gln	Ile		
Thr	Ala	Ser	Gly	Gln	Tyr	Gly	Gln	Trp	Ala	Pro	Lys	Leu	Ala	Arg		
Leu	His	Tyr	Ser	Gly	Ser	Ile	Asn	Ala	Trp	Ser	Thr	Lys	Glu	Pro		
Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu	Leu	Ala	Pro	Met	Ile	Ile	His		
Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys	Phe	Ser	Ser	Leu	Tyr		
Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu	Asp	Gly	Lys	Lys	Trp		
Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr	Leu	Met	Val	Phe	Phe		
Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His	Asn	Ile	Phe	Asn	Pro		
Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His	Pro	Thr	His	Tyr	Ser		

1250	1255	1260
Ile Arg Ser Thr Leu Arg	Met Glu Leu Met Gly	Cys Asp Leu Asn
1265	1270	1275
Ser Cys Ser Met Pro Leu	Gly Met Glu Ser Lys	Ala Ile Ser Asp
1280	1285	1290
Ala Gln Ile Thr Ala Ser	Ser Tyr Phe Thr Asn	Met Phe Ala Thr
1295	1300	1305
Trp Ser Pro Ser Lys Ala	Arg Leu His Leu Gln	Gly Arg Ser Asn
1310	1315	1320
Ala Trp Arg Pro Gln Val	Asn Asn Pro Lys Glu	Trp Leu Gln Val
1325	1330	1335
Asp Phe Gln Lys Thr Met	Lys Val Thr Gly Val	Thr Thr Gln Gly
1340	1345	1350
Val Lys Ser Leu Leu Thr	Ser Met Tyr Val Lys	Glu Phe Leu Ile
1355	1360	1365
Ser Ser Ser Gln Asp Gly	His Gln Trp Thr Leu	Phe Phe Gln Asn
1370	1375	1380
Gly Lys Val Lys Val Phe	Gln Gly Asn Gln Asp	Ser Phe Thr Pro
1385	1390	1395
Val Val Asn Ser Leu Asp	Pro Pro Leu Leu Thr	Arg Tyr Leu Arg
1400	1405	1410
Ile His Pro Gln Ser Trp	Val His Gln Ile Ala	Leu Arg Met Glu
1415	1420	1425
Val Leu Gly Cys Glu Ala	Gln Asp Leu Tyr	
1430	1435	

<210> 7

<211> 4371

<212> DNA

<213> 人工序列

<220>

<223> BDDD FVIII

<400> 7

atgcaaatag agctctccac ctgtctcttt ctgtgccttt tgcgattctg ctttagtgcc 60
accagaagat actacctggg tgcagtggaa ctgtcatggg actatatgca aagtgatctc 120
ggtagagctgc ctgtggacgc aagatttcct cctagagtgc caaatcttt tccattcaac 180
acctcagtcg tgtacaaaaa gactctgttt gtagaattca cggatcacct ttccaacatc 240
gctaagccaa ggccaccctg gatgggtctg ctaggtccta ccatccaggc tgaggtttat 300
gatacagtggt tcattacact taagaacatg gcttcccac ctagtcagtct tcatgctgtt 360
gggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtcaaagg 420

gagaaagaag atgataaagt cttccctggt ggaagccata catatgtctg gcaggtcctg 480
 aaagagaatg gtccaatggc ctctgaccca ctgtgcctta cctactcata tctttctcat 540
 gtggacctgg taaaagactt gaattcaggc ctcatggag ccctactagt atgtagagaa 600
 gggagtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta 660
 tttgatgaag ggaaaagttg gcactcagaa acaaagaact cttgatgca ggatagggat 720
 gctgcatctg ctcgggcctg gcctaaaatg cacacagtca atggttatgt aaacaggtct 780
 ctgccaggtc tgattggatg ccacaggaaa tcagtctatt ggcatgtgat tggaatgggc 840
 accactcctg aagtgcactc aatattcctc gaaggtcaca ctttcttgt gaggaacct 900
 cgccaggcgt ctttgaaaat ctgcceaata actttcctta ctgctcaaac actcttgatg 960
 gaccttggac agtttctact gttttgtcat atctcttccc accaacaatga tggcatggaa 1020
 gcttatgtca aagtagacag ctgtccagag gaacccaac tacgaatgaa aaataatgaa 1080
 gaagcggaag actatgatga tgatcttact gattctgaaa tggatgtggt caggtttgat 1140
 gatgacaact ctcttctctt tatccaaatt cgctcagttg ccaagaagca tcctaaaact 1200
 tgggtacatt acattgctgc tgaagaggag gactgggact atgctccctt agtctctgcc 1260
 ccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg 1320
 aagtacaaaa aagtccgatt tatggcatac acagatgaaa cctttaagac tcgtgaagct 1380
 attcagcatg aatcaggaat cttgggacct ttactttatg gggaagttgg agacacactg 1440
 ttgattatat ttaagaatca agcaagcaga ccatataaca tctaccctca cggaatcact 1500
 gatgtccgtc ctttgtattc aaggagatta ccaaaaggtg taaaacattt gaaggatttt 1560
 ccaattctgc caggagaaat attcaaatat aaatggacag tgactgtaga agatgggcca 1620
 actaaatcag atctctgggt cctgaccgcg tattactcta gtttcgttaa tatggagaga 1680
 gatctagctt caggactcat tggccctctc ctcatctgct acaaagaatc tgtagatcaa 1740
 agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag 1800
 aaccgaagct ggtacctcac agagaatata caacgcttc tcccaatcc agctggagtg 1860
 cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatgtt 1920
 tttgatagtt tgcagttgtc agtttgttg catgaggtgg catactggta cattctaagc 1980
 attggagcac agactgactt ctttctgtc ttcttctctg gatatacctt caaacacaaa 2040
 atgggtctatg aagacacact caccctattc ccattctcag gagaaactgt cttcatgtcg 2100
 atggaaaacc caggtctatg gattctgggg tgccacaact cagactttcg gaacagaggc 2160
 atgaccgcct tactgaaggt ttctagttgt gacaagaaca ctggtgatta ttacgaggac 2220
 agttatgaag atatttcagc atacttgctg agtaaaaaca atgccattga accaagaagc 2280
 ttctctcaaa acccaccagt cttgaaacgc catcaacggg aaataactcg tactactctt 2340
 cagtcagatc aagaggaaat tgactatgat gataccatat cagttgaaat gaagaaggaa 2400
 gattttgaca tttatgatga ggatgaaaat cagagcccc gcagctttca aaagaaaaca 2460
 cgacactatt ttattgctgc agtggagagg ctctgggatt atgggatgag tagctcccca 2520
 catgttctaa gaaacagggc tcagagtggc agtgtccctc agttcaagaa agttgttttc 2580
 caggaattta ctgatggctc ctttactcag cccttatacc gtggagaact aaatgaacat 2640
 ttgggactcc tggggccata tataagagca gaagttgaag ataatatcat ggtaactttc 2700
 agaaatcagg cctctcgtcc ctattccttc tattctagcc ttatttctta tgaggaagat 2760

cagaggcaag gagcagaacc tagaaaaaac tttgtcaagc ctaatgaaac caaaacttac 2820
 ttttggaag tgcaacatca tatggcacc actaaagatg agtttgactg caaagcctgg 2880
 gcttatttct ctgatgttga cctggaaaaa gatgtgcact caggcctgat tggaccctt 2940
 ctggctctgcc acactaacac actgaaccct gctcatggga gacaagtac agtacaggaa 3000
 tttgctctgt ttttcacat ctttgatgag accaaaagct ggtacttcac tgaaaatatg 3060
 gaaagaaact gcagggtcc ctgcaatatac cagatggaag atcccacttt taaagagaat 3120
 tatcgcttcc atgcaatcaa tggctacata atggatacac tacctggctt agtaatggct 3180
 caggatcaaa ggattcgatg gtatctgctc agcatgggca gcaatgaaaa catccattct 3240
 attcatttca gtggacatgt gttcactgta cgaaaaaaag aggagtataa aatggcactg 3300
 tacaatctct atccagggtg ttttgagaca gtggaaatgt taccatccaa agctggaatt 3360
 tggcgggtgg aatgccttat tggcgagcat ctacatgctg ggatgagcac actttttctg 3420
 gtgtacagca ataagtgtca gactccctg ggaatggctt ctggacacat tagagatttt 3480
 cagattacag cttcaggaca atatggacag tgggccccaa agctggccag acttcattat 3540
 tccgatcaa tcaatgcctg gagcaccaag gagccctttt cttggatcaa ggtggatctg 3600
 ttggcaccaa tgattattca cggcatcaag acccagggtg cccgtcagaa gttctccagc 3660
 ctctacatct ctcagtttat catcatgtat agtcttgatg ggaagaagtg gcagacttat 3720
 cgaggaaatt ccactggaac cttaatggtc ttctttggca atgtggattc atctgggata 3780
 aaacacaata tttttaaccc tccaattatt gctcgataca tccgtttgca cccaactcat 3840
 tatagcattc gcagcactct tcgcatggag ttgatgggtg gtgatttaaa tagttgcagc 3900
 atgccattgg gaatggagag taaagcaata tcagatgcac agattactgc ttcactctac 3960
 tttaccaata tgtttgccac ctggctctct tcaaaagctc gacttcacct ccaaggagg 4020
 agtaatgcct ggagacctca ggtgaataat ccaaaagagt ggctgcaagt ggacttccag 4080
 aagacaatga aagtcacagg agtaactact caggagtaa aatctctgct taccagcatg 4140
 tatgtgaagg agttcctcat ctccagcagt caagatggcc atcagtggac tctctttttt 4200
 cagaatggca aagtaaaggt ttttcaggga aatcaagact cttcacacc tgtgtgtaac 4260
 tctctagacc caccgttact gactcgctac cttcgaattc acccccagag ttgggtgcac 4320
 cagattgccc tgaggatgga ggttctgggc tgcgaggcac aggacctcta c 4371

<210> 8

<211> 25

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 8

Thr Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro Asn Asp Lys Tyr

1

5

10

15

Glu Pro Phe Trp Glu Asp Glu Glu Lys

20

25

<210> 9

<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 9
Arg Arg Arg Arg
1
<210> 10
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 10
Arg Lys Arg Arg Lys Arg
1 5
<210> 11
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 11
Arg Arg Arg Arg Ser
1 5
<210> 12
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 12
Thr Gln Ser Phe Asn Asp Phe Thr Arg
1 5
<210> 13
<211> 10
<212> PRT
<213> 人工序列

<220>

<223> 切割位点

<400> 13

Ser Val Ser Gln Thr Ser Lys Leu Thr Arg

1 5 10

<210> 14

<211> 10

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 14

Asp Phe Leu Ala Glu Gly Gly Gly Val Arg

1 5 10

<210> 15

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 15

Thr Thr Lys Ile Lys Pro Arg

1 5

<210> 16

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 16

Leu Val Pro Arg Gly

1 5

<210> 17

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 17

Ala Leu Arg Pro Arg

1 5

<210> 18

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 18

Lys Leu Thr Arg Ala Glu Thr

1 5

<210> 19

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 19

Asp Phe Thr Arg Val Val Gly

1 5

<210> 20

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 20

Thr Met Thr Arg Ile Val Gly Gly

1 5

<210> 21

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 21

Ser Pro Phe Arg Ser Thr Gly Gly

1 5

<210> 22

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 22

Leu Gln Val Arg Ile Val Gly Gly

1 5

<210> 23

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 23

Pro Leu Gly Arg Ile Val Gly Gly

1 5

<210> 24

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 24

Ile Glu Gly Arg Thr Val Gly Gly

1 5

<210> 25

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 25

Leu Thr Pro Arg Ser Leu Leu Val

1 5

<210> 26

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 26

Leu Gly Pro Val Ser Gly Val Pro

1 5

<210> 27

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 27

Val Ala Gly Asp Ser Leu Glu Glu

1 5

<210> 28

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 28

Gly Pro Ala Gly Leu Gly Gly Ala

1 5

<210> 29

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 29

Gly Pro Ala Gly Leu Arg Gly Ala

1 5

<210> 30

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 30

Ala Pro Leu Gly Leu Arg Leu Arg

1 5

<210> 31

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 31

Pro Ala Leu Pro Leu Val Ala Gln

1 5

<210> 32

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 32

Glu Asn Leu Tyr Phe Gln Gly

1 5

<210> 33

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 33

Asp Asp Asp Lys Ile Val Gly Gly

1 5

<210> 34

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 34

Leu Glu Val Leu Phe Gln Gly Pro

1 5

<210> 35

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 35

Leu Pro Lys Thr Gly Ser Glu Ser

1 5

<210> 36

<211> 42

<212> PRT

<213> 人工序列

<220>

<223> AE42

<400> 36

Gly Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly

1 5 10 15

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala

20 25 30

Thr Ser Gly Ser Glu Thr Pro Ala Ser Ser

35 40

<210> 37

<211> 143

<212> PRT

<213> 人工序列

<220>

<223> AE144

<400> 37

Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu

1 5 10 15

Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly

20 25 30

Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu

35 40 45

Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro

50 55 60

Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser Gly

65 70 75 80

Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro

				85					90					95					
Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Thr	Ser	Glu				
				100					105					110					
Ser	Ala	Pro	Glu	Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser				
				115					120					125					
Glu	Thr	Pro	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro					
				130					135					140					

<210> 38

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AG144

<400> 38

Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ser	Ser	Thr				
1				5					10					15					
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr				
				20					25					30					
Gly	Thr	Gly	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro				
				35					40					45					
Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro				
				50					55					60					
Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala				
65					70					75					80				
Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro				
				85					90					95					
Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro				
				100					105					110					
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala				
				115					120					125					
Ser	Ser	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro				
				130					135					140					

<210> 39

<211> 288

<212> PRT

<213> 人工序列

<220>

<223> AE288

<400> 39

Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro			
1	5	10	15
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro			
20	25	30	
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro			
35	40	45	
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr			
50	55	60	
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr			
65	70	75	80
Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
85	90	95	
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu			
100	105	110	
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr			
115	120	125	
Ser Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu			
130	135	140	
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu			
145	150	155	160
Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro			
165	170	175	
Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
180	185	190	
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro			
195	200	205	
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr			
210	215	220	
Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
225	230	235	240
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro			
245	250	255	
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro			
260	265	270	
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
275	280	285	

<210> 40
 <211> 288
 <212> PRT

<213> 人工序列

<220>

<223> AG288

<400> 40

Pro Gly Ala Ser	Pro Gly Thr Ser Ser Thr Gly Ser	Pro Gly Ala Ser
1	5	10 15
Pro Gly Thr Ser Ser Thr Gly Ser	Pro Gly Thr Pro Gly Ser Gly Thr	
20	25	30
Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser		
35	40	45
Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser		
50	55	60
Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr		
65	70	75 80
Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser		
85	90	95
Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser		
100	105	110
Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser		
115	120	125
Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser		
130	135	140
Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser		
145	150	155 160
Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro Ser Ala Ser		
165	170	175
Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly		
180	185	190
Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser		
195	200	205
Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly		
210	215	220
Ala Thr Gly Ser Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly		
225	230	235 240
Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser		
245	250	255
Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Thr Pro Gly Ser Gly Thr		
260	265	270
Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser		

275	280	285
<210> 41		
<211> 576		
<212> PRT		
<213> 人工序列		
<220>		
<223> AE576		
<400> 41		
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu		
1 5 10 15		
Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu		
20 25 30		
Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu		
35 40 45		
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr		
50 55 60		
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro		
65 70 75 80		
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro		
85 90 95		
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Pro Ala		
100 105 110		
Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro		
115 120 125		
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro		
130 135 140		
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala		
145 150 155 160		
Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu		
165 170 175		
Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro		
180 185 190		
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr		
195 200 205		
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro		
210 215 220		
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro		
225 230 235 240		
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr		

	245		250		255
Glu Pro Ser	Glu Gly Ser Ala Pro	Gly Thr Ser	Glu Ser Ala Thr	Pro	
	260		265		270
Glu Ser Gly	Pro Gly Thr Ser	Glu Ser Ala Thr	Pro Glu Ser	Gly Pro	
	275		280		285
Gly Ser Pro	Ala Gly Ser Pro	Thr Ser Thr	Glu Glu Gly	Thr Ser	Glu
	290		295		300
Ser Ala Thr	Pro Glu Ser Gly	Pro Gly Ser	Glu Pro Ala	Thr Ser	Gly
305		310		315	320
Ser Glu Thr	Pro Gly Thr Ser	Glu Ser Ala Thr	Pro Glu Ser	Gly Pro	
	325		330		335
Gly Thr Ser	Thr Glu Pro Ser	Glu Gly Ser	Ala Pro Gly	Thr Ser	Thr
	340		345		350
Glu Pro Ser	Glu Gly Ser Ala	Pro Gly Thr	Ser Thr Glu	Pro Ser	Glu
	355		360		365
Gly Ser Ala	Pro Gly Thr Ser	Thr Glu Pro	Ser Glu Gly	Ser Ala	Pro
	370		375		380
Gly Thr Ser	Thr Glu Pro Ser	Glu Gly Ser	Ala Pro Gly	Thr Ser	Thr
385		390		395	400
Glu Pro Ser	Glu Gly Ser Ala	Pro Gly Ser	Pro Ala Gly	Ser Pro	Thr
	405		410		415
Ser Thr Glu	Glu Gly Thr Ser	Thr Glu Pro	Ser Glu Gly	Ser Ala	Pro
	420		425		430
Gly Thr Ser	Glu Ser Ala Thr	Pro Glu Ser	Gly Pro Gly	Ser Glu	Pro
	435		440		445
Ala Thr Ser	Gly Ser Glu Thr	Pro Gly Thr	Ser Glu Ser	Ala Thr	Pro
	450		455		460
Glu Ser Gly	Pro Gly Ser Glu	Pro Ala Thr	Ser Gly Ser	Glu Thr	Pro
465		470		475	480
Gly Thr Ser	Glu Ser Ala Thr	Pro Glu Ser	Gly Pro Gly	Thr Ser	Thr
	485		490		495
Glu Pro Ser	Glu Gly Ser Ala	Pro Gly Thr	Ser Glu Ser	Ala Thr	Pro
	500		505		510
Glu Ser Gly	Pro Gly Ser Pro	Ala Gly Ser	Pro Thr Ser	Thr Glu	Glu
	515		520		525
Gly Ser Pro	Ala Gly Ser Pro	Thr Ser Thr	Glu Glu Gly	Ser Pro	Ala
	530		535		540
Gly Ser Pro	Thr Ser Thr	Glu Glu Gly	Thr Ser Glu	Ser Ala	Thr
545		550		555	560

Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
	565	570	575
<210> 42			
<211> 576			
<212> PRT			
<213> 人工序列			
<220>			
<223> AG576			
<400> 42			
Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser			
1	5	10	15
Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro Ser Ala Ser			
	20	25	30
Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly			
	35	40	45
Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser			
	50	55	60
Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser			
65	70	75	80
Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser			
	85	90	95
Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Thr Pro			
	100	105	110
Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser			
	115	120	125
Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser			
	130	135	140
Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser			
145	150	155	160
Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Thr Pro Gly Ser Gly Thr			
	165	170	175
Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser			
	180	185	190
Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser			
	195	200	205
Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly			
	210	215	220
Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser			
225	230	235	240

Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Thr Pro		
245	250	255
Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly		
260	265	270
Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser		
275	280	285
Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser		
290	295	300
Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser		
305	310	315
Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser		
325	330	335
Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser		
340	345	350
Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser		
355	360	365
Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser		
370	375	380
Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Thr Pro		
385	390	395
Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly		
405	410	415
Ala Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser		
420	425	430
Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Thr Pro		
435	440	445
Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly		
450	455	460
Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser		
465	470	475
Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser		
485	490	495
Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser		
500	505	510
Ser Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser		
515	520	525
Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser		
530	535	540
Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser		

545					550					555					560
Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser
					565					570					575
<210> 43															
<211> 864															
<212> PRT															
<213> 人工序列															
<220>															
<223> AE864															
<400> 43															
Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser	Thr	Glu	Glu	Gly	Thr	Ser	Glu
1				5					10					15	
Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu
				20				25					30		
Gly	Ser	Ala	Pro	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser	Thr	Glu	Glu
				35				40					45		
Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Thr	Ser	Thr
				50				55				60			
Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro
65					70				75					80	
Glu	Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro
				85					90					95	
Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Ser	Pro	Ala
				100					105					110	
Gly	Ser	Pro	Thr	Ser	Thr	Glu	Glu	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro
				115				120					125		
Glu	Ser	Gly	Pro	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro
				130				135					140		
Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Ser	Pro	Ala
145					150				155					160	
Gly	Ser	Pro	Thr	Ser	Thr	Glu	Glu	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu
				165					170					175	
Gly	Ser	Ala	Pro	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro
				180					185					190	
Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Thr
				195				200					205		
Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro
				210					215					220	
Glu	Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro

225	230	235	240
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr			
	245	250	255
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro			
	260	265	270
Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
	275	280	285
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu			
	290	295	300
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly			
305	310	315	320
Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
	325	330	335
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr			
	340	345	350
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu			
	355	360	365
Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
	370	375	380
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr			
385	390	395	400
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr			
	405	410	415
Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
	420	425	430
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro			
	435	440	445
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro			
	450	455	460
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro			
465	470	475	480
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr			
	485	490	495
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro			
	500	505	510
Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu			
	515	520	525
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala			
530	535	540	

Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro			
545	550	555	560
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
	565	570	575
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro			
	580	585	590
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro			
	595	600	605
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro			
	610	615	620
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr			
625	630	635	640
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr			
	645	650	655
Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
	660	665	670
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu			
	675	680	685
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr			
	690	695	700
Ser Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu			
705	710	715	720
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu			
	725	730	735
Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro			
	740	745	750
Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
	755	760	765
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro			
	770	775	780
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr			
785	790	795	800
Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
	805	810	815
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro			
	820	825	830
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro			
	835	840	845
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			

850	855	860
<210> 44		
<211> 864		
<212> PRT		
<213> 人工序列		
<220>		
<223> AG864		
<400> 44		
Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser Pro		
1 5 10 15		
Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr		
20 25 30		
Gly Thr Gly Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro		
35 40 45		
Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro		
50 55 60		
Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser Ser		
65 70 75 80		
Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro		
85 90 95		
Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Thr Pro Gly		
100 105 110		
Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser		
115 120 125		
Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro		
130 135 140		
Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr		
145 150 155 160		
Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser		
165 170 175		
Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro		
180 185 190		
Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro		
195 200 205		
Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr		
210 215 220		
Gly Thr Gly Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro		
225 230 235 240		
Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro		

				245					250					255			
Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
				260					265					270			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
				275					280					285			
Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ala	Ser	Pro		
				290					295					300			
Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
305					310				315						320		
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
				325					330					335			
Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Thr	Pro	Gly		
				340					345					350			
Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
				355					360					365			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
				370					375					380			
Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr		
385					390				395						400		
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala		
				405					410					415			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
				420					425					430			
Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ser	Ser	Thr		
				435					440					445			
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala		
				450					455					460			
Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro		
465					470				475						480		
Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro		
				485					490					495			
Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
				500					505					510			
Thr	Gly	Ser	Pro	Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro		
				515					520					525			
Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro		
				530					535					540			
Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
545					550				555						560		

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro		
565	570	575
Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr		
580	585	590
Pro Ser Gly Ala Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala		
595	600	605
Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro		
610	615	620
Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr		
625	630	635
Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala		
645	650	655
Thr Gly Ser Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro		
660	665	670
Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro		
675	680	685
Gly Thr Ser Ser Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala		
690	695	700
Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro		
705	710	715
Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro		
725	730	735
Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser Ser		
740	745	750
Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro		
755	760	765
Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro		
770	775	780
Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser Ser		
785	790	795
Thr Gly Ser Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro		
805	810	815
Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr		
820	825	830
Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala		
835	840	845
Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro		
850	855	860

<210> 45

<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 45
Thr Gln Ser Phe Asn Asp Phe Thr Arg
1 5
<210> 46
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 46
Ser Val Ser Gln Thr Ser Lys Leu Thr Arg
1 5 10
<210> 47
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 47
Asp Phe Leu Ala Glu Gly Gly Gly Val Arg
1 5 10
<210> 48
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 48
Thr Thr Lys Ile Lys Pro Arg
1 5
<210> 49
<211> 5
<212> PRT
<213> 人工序列

<220>
<223> 切割位点
<400> 49
Leu Val Pro Arg Gly
1 5
<210> 50
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 50
Ala Leu Arg Pro Arg
1 5
<210> 51
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 分选酶识别基序
<220>
<221> misc_feature
<222> (3) .. (3)
<223> Xaa可以为任何天然存在的氨基酸
<400> 51
Leu Pro Xaa Thr Gly
1 5
<210> 52
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> 多肽
<400> 52
Pro Lys Asn Ser Ser Met Ile Ser Asn Thr Pro
1 5 10
<210> 53
<211> 7
<212> PRT

<213> 人工序列

<220>

<223> 多肽

<400> 53

His Gln Ser Leu Gly Thr Gln

1 5

<210> 54

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 多肽

<400> 54

His Gln Asn Leu Ser Asp Gly Lys

1 5

<210> 55

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 多肽

<400> 55

His Gln Asn Ile Ser Asp Gly Lys

1 5

<210> 56

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 多肽

<400> 56

Val Ile Ser Ser His Leu Gly Gln

1 5

<210> 57

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> 接头

<220>

<221> 重复

<222> (1) .. (4)

<223> 重复1至100次

<400> 57

Gly Gly Gly Ser

1

<210> 58

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 接头

<220>

<221> 重复

<222> (1) .. (3)

<223> 可以重复1至100次

<220>

<221> 重复

<222> (4) .. (8)

<223> 可以重复1至100次

<400> 58

Gly Gly Ser Gly Gly Gly Gly Ser

1

5

<210> 59

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 59

Ser Gly Gly Ser Gly Gly Ser

1

5

<210> 60

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 60

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Gly
1 5 10 15

<210> 61

<211> 16

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 61

Gly Gly Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> 62

<211> 18

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 62

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
1 5 10 15

Gly Ser

<210> 63

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 63

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> 64

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 接头

<220>

<221> 重复

<222> (1) .. (5)
<223> 可以重复1-20次
<400> 64
Gly Gly Gly Gly Ser
1 5
<210> 65
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 65
Lys Leu Thr Arg Ala Glu Thr
1 5
<210> 66
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 66
Asp Phe Thr Arg Val Val Gly
1 5
<210> 67
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE

<222> (4) .. (5)
<223> 切割位点
<400> 67
Thr Met Thr Arg Ile Val Gly Gly
1 5
<210> 68
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 68
Ser Pro Phe Arg Ser Thr Gly Gly
1 5
<210> 69
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 69
Leu Gln Val Arg Ile Val Gly Gly
1 5
<210> 70
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE

<222> (4) .. (5)

<223> 切割位点

<400> 70

Pro Leu Gly Arg Ile Val Gly Gly

1 5

<210> 71

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<220>

<221> MISC_FEATURE

<222> (4) .. (5)

<223> 切割位点

<400> 71

Ile Glu Gly Arg Thr Val Gly Gly

1 5

<210> 72

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<220>

<221> MISC_FEATURE

<222> (4) .. (5)

<223> 切割位点

<400> 72

Leu Thr Pro Arg Ser Leu Leu Val

1 5

<210> 73

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<220>

<221> MISC_FEATURE

<222> (4) .. (5)
<223> 切割位点
<400> 73
Leu Gly Pro Val Ser Gly Val Pro
1 5
<210> 74
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 74
Val Ala Gly Asp Ser Leu Glu Glu
1 5
<210> 75
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 75
Gly Pro Ala Gly Leu Gly Gly Ala
1 5
<210> 76
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE

<222> (4) .. (5)

<223> 切割位点

<400> 76

Gly Pro Ala Gly Leu Arg Gly Ala

1 5

<210> 77

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<220>

<221> MISC_FEATURE

<222> (4) .. (5)

<223> 切割位点

<400> 77

Ala Pro Leu Gly Leu Arg Leu Arg

1 5

<210> 78

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<220>

<221> MISC_FEATURE

<222> (4) .. (5)

<223> 切割位点

<400> 78

Pro Ala Leu Pro Leu Val Ala Gln

1 5

<210> 79

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<220>

<221> MISC_FEATURE

<222> (6) .. (7)
<223> 切割位点
<400> 79
Glu Asn Leu Tyr Phe Gln Gly
1 5
<210> 80
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (4) .. (5)
<223> 切割位点
<400> 80
Asp Asp Asp Lys Ile Val Gly Gly
1 5
<210> 81
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE
<222> (6) .. (7)
<223> 切割位点
<400> 81
Leu Glu Val Leu Phe Gln Gly Pro
1 5
<210> 82
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<220>
<221> MISC_FEATURE

<222> (4) .. (5)
<223> 切割位点
<400> 82
Leu Pro Lys Thr Gly Ser Glu Ser
1 5
<210> 83
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 83
Thr Gln Ser Phe Asn Asp Phe Thr Arg
1 5
<210> 84
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 84
Ser Val Ser Gln Thr Ser Lys Leu Thr Arg
1 5 10
<210> 85
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 85
Asp Phe Leu Ala Glu Gly Gly Gly Val Arg
1 5 10
<210> 86
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> 切割位点
<400> 86

Thr Thr Lys Ile Lys Pro Arg

1 5

<210> 87

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 87

Leu Val Pro Arg Gly

1 5

<210> 88

<211> 10

<212> PRT

<213> 人工序列

<220>

<223> 切割位点

<400> 88

Ala Leu Arg Pro Arg Val Val Gly Gly Ala

1 5 10

<210> 89

<211> 1896

<212> PRT

<213> 人工序列

<220>

<223> FVIII 198

<400> 89

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe

1 5 10 15

Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser

20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg

35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val

50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile

65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln

85 90 95

Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro

				405					410					415			
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn		
				420					425					430			
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met		
				435					440					445			
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu		
				450					455					460			
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu		
465									470					475			480
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro		
				485					490					495			
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys		
				500					505					510			
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe		
				515					520					525			
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp		
				530					535					540			
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg		
545									550					555			560
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu		
				565					570					575			
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val		
				580					585					590			
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu		
				595					600					605			
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp		
				610					615					620			
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val		
625									630					635			640
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp		
				645					650					655			
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe		
				660					665					670			
Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr		
				675					680					685			
Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro		
				690					695					700			
Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly		
705									710					715			720

Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp
				725				730				735			
Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys
				740				745				750			
Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Ser	Arg	His	Pro
				755				760				765			
Ser	Thr	Arg	Gln	Lys	Gln	Phe	Asn	Ala	Thr	Thr	Ile	Pro	Glu	Asn	Asp
				770				775				780			
Ile	Glu	Lys	Thr	Asp	Pro	Trp	Phe	Ala	His	Arg	Thr	Pro	Met	Pro	Lys
785				790				795				800			
Ile	Gln	Asn	Val	Ser	Ser	Ser	Asp	Leu	Leu	Met	Leu	Leu	Arg	Gln	Ser
				805				810				815			
Pro	Thr	Pro	His	Gly	Leu	Ser	Leu	Ser	Asp	Leu	Gln	Glu	Ala	Lys	Tyr
				820				825				830			
Glu	Thr	Phe	Ser	Asp	Asp	Pro	Ser	Pro	Gly	Ala	Ile	Asp	Ser	Asn	Asn
				835				840				845			
Ser	Leu	Ser	Glu	Met	Thr	His	Phe	Arg	Pro	Gln	Leu	His	His	Ser	Gly
				850				855				860			
Asp	Met	Val	Phe	Thr	Pro	Glu	Ser	Gly	Leu	Gln	Leu	Arg	Leu	Asn	Glu
865				870				875				880			
Lys	Leu	Gly	Thr	Thr	Ala	Ala	Thr	Glu	Leu	Lys	Lys	Leu	Asp	Phe	Lys
				885				890				895			
Val	Ser	Ser	Thr	Ser	Asn	Asn	Leu	Ile	Ser	Thr	Ile	Pro	Ser	Asp	Asn
				900				905				910			
Leu	Ala	Ala	Gly	Thr	Asp	Asn	Thr	Ser	Ser	Leu	Gly	Pro	Pro	Ser	Met
				915				920				925			
Pro	Val	His	Tyr	Asp	Ser	Gln	Leu	Asp	Thr	Thr	Leu	Phe	Gly	Lys	Lys
				930				935				940			
Ser	Ser	Pro	Leu	Thr	Glu	Ser	Gly	Gly	Pro	Leu	Ser	Leu	Ser	Glu	Glu
945				950				955				960			
Asn	Asn	Asp	Ser	Lys	Leu	Leu	Glu	Ser	Gly	Leu	Met	Asn	Ser	Gln	Glu
				965				970				975			
Ser	Ser	Trp	Gly	Lys	Asn	Val	Ser	Ser	Glu	Ile	Thr	Arg	Thr	Thr	Leu
				980				985				990			
Gln	Ser	Asp	Gln	Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu
				995				1000				1005			
Met	Lys	Lys	Glu	Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	
				1010				1015				1020			
Ser	Pro	Arg	Ser	Phe	Gln	Lys	Lys	Thr	Arg	His	Tyr	Phe	Ile	Ala	

1025	1030	1035
Ala Val Glu Arg Leu Trp	Asp Tyr Gly Met Ser	Ser Ser Pro His
1040	1045	1050
Val Leu Arg Asn Arg Ala	Gln Ser Gly Ser Val	Pro Gln Phe Lys
1055	1060	1065
Lys Val Val Phe Gln Glu	Phe Thr Asp Gly Ser	Phe Thr Gln Pro
1070	1075	1080
Leu Tyr Arg Gly Glu Leu	Asn Glu His Leu Gly	Leu Leu Gly Pro
1085	1090	1095
Tyr Ile Arg Ala Glu Val	Glu Asp Asn Ile Met	Val Thr Phe Arg
1100	1105	1110
Asn Gln Ala Ser Arg Pro	Tyr Ser Phe Tyr Ser	Ser Leu Ile Ser
1115	1120	1125
Tyr Glu Glu Asp Gln Arg	Gln Gly Ala Glu Pro	Arg Lys Asn Phe
1130	1135	1140
Val Lys Pro Asn Glu Thr	Lys Thr Tyr Phe Trp	Lys Val Gln His
1145	1150	1155
His Met Ala Pro Thr Lys	Asp Glu Phe Asp Cys	Lys Ala Trp Ala
1160	1165	1170
Tyr Phe Ser Asp Val Asp	Leu Glu Lys Asp Val	His Ser Gly Leu
1175	1180	1185
Ile Gly Pro Leu Leu Val	Cys His Thr Asn Thr	Leu Asn Pro Ala
1190	1195	1200
His Gly Arg Gln Val Thr	Val Gln Glu Phe Ala	Leu Phe Phe Thr
1205	1210	1215
Ile Phe Asp Glu Thr Lys	Ser Trp Tyr Phe Thr	Glu Asn Met Glu
1220	1225	1230
Arg Asn Cys Arg Ala Pro	Cys Asn Ile Gln Met	Glu Asp Pro Thr
1235	1240	1245
Phe Lys Glu Asn Tyr Arg	Phe His Ala Ile Asn	Gly Tyr Ile Met
1250	1255	1260
Asp Thr Leu Pro Gly Leu	Val Met Ala Gln Asp	Gln Arg Ile Arg
1265	1270	1275
Trp Tyr Leu Leu Ser Met	Gly Ser Asn Glu Asn	Ile His Ser Ile
1280	1285	1290
His Phe Ser Gly His Val	Phe Thr Val Arg Lys	Lys Glu Glu Tyr
1295	1300	1305
Lys Met Ala Leu Tyr Asn	Leu Tyr Pro Gly Val	Phe Glu Thr Val
1310	1315	1320

Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu		
1325	1330	1335
Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val		
1340	1345	1350
Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His		
1355	1360	1365
Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp		
1370	1375	1380
Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala		
1385	1390	1395
Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu		
1400	1405	1410
Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln		
1415	1420	1425
Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser		
1430	1435	1440
Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly		
1445	1450	1455
Thr Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys		
1460	1465	1470
His Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu		
1475	1480	1485
His Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu		
1490	1495	1500
Met Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu		
1505	1510	1515
Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe		
1520	1525	1530
Thr Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His		
1535	1540	1545
Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro		
1550	1555	1560
Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr		
1565	1570	1575
Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr		
1580	1585	1590
Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp		
1595	1600	1605
Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn		

1610	1615	1620
Gln Asp Ser Phe Thr Pro	Val Val Asn Ser Leu	Asp Pro Pro Leu
1625	1630	1635
Leu Thr Arg Tyr Leu Arg	Ile His Pro Gln Ser	Trp Val His Gln
1640	1645	1650
Ile Ala Leu Arg Met Glu	Val Leu Gly Cys Glu	Ala Gln Asp Leu
1655	1660	1665
Tyr Asp Lys Thr His Thr	Cys Pro Pro Cys Pro	Ala Pro Glu Leu
1670	1675	1680
Leu Gly Gly Pro Ser Val	Phe Leu Phe Pro Pro	Lys Pro Lys Asp
1685	1690	1695
Thr Leu Met Ile Ser Arg	Thr Pro Glu Val Thr	Cys Val Val Val
1700	1705	1710
Asp Val Ser His Glu Asp	Pro Glu Val Lys Phe	Asn Trp Tyr Val
1715	1720	1725
Asp Gly Val Glu Val His	Asn Ala Lys Thr Lys	Pro Arg Glu Glu
1730	1735	1740
Gln Tyr Asn Ser Thr Tyr	Arg Val Val Ser Val	Leu Thr Val Leu
1745	1750	1755
His Gln Asp Trp Leu Asn	Gly Lys Glu Tyr Lys	Cys Lys Val Ser
1760	1765	1770
Asn Lys Ala Leu Pro Ala	Pro Ile Glu Lys Thr	Ile Ser Lys Ala
1775	1780	1785
Lys Gly Gln Pro Arg Glu	Pro Gln Val Tyr Thr	Leu Pro Pro Ser
1790	1795	1800
Arg Asp Glu Leu Thr Lys	Asn Gln Val Ser Leu	Thr Cys Leu Val
1805	1810	1815
Lys Gly Phe Tyr Pro Ser	Asp Ile Ala Val Glu	Trp Glu Ser Asn
1820	1825	1830
Gly Gln Pro Glu Asn Asn	Tyr Lys Thr Thr Pro	Pro Val Leu Asp
1835	1840	1845
Ser Asp Gly Ser Phe Phe	Leu Tyr Ser Lys Leu	Thr Val Asp Lys
1850	1855	1860
Ser Arg Trp Gln Gln Gly	Asn Val Phe Ser Cys	Ser Val Met His
1865	1870	1875
Glu Ala Leu His Asn His	Tyr Thr Gln Lys Ser	Leu Ser Leu Ser
1880	1885	1890
Pro Gly Lys		
1895		

<210> 90
<211> 99
<212> DNA
<213> 人工序列
<220>
<223> 具有VIII信号和BsiW1位点的ESC48- Fwd - VWF-D'D3
<400> 90
tcgcgacgta cggccgccac catgcaaata gagctetcca cctgcttctt tctgtgcctt 60
ttgcgattct gcttttagcct atcctgtcgg ccccccattg 99
<210> 91
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 具有6His和Not 1位点的ESC51- Rev- VWF D'D3 (1-477个氨基酸)
<400> 91
tgacctcgag cggccgctca gtggtgatgg tgatgatgcg gctcctggca ggcttcacag 60
gtgaggttga caac 74
<210> 92
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 具有Nhe1位点的ESC 89-fwd
<400> 92
ctcactatag ggagacccaa gctggctagc cg 32
<210> 93
<211> 43
<212> DNA
<213> 人工序列
<220>
<223> 具有Sal1的ESC 91-rev
<400> 93
ctggatcccc ggagtcgact cgtcagtggg gatggtgatg atg 43
<210> 94
<211> 40
<212> PRT
<213> 人工序列
<220>

<223> 接头

<400> 94

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
 20 25 30
Gly Gly Ser Gly Gly Gly Gly Ser
 35 40

<210> 95

<211> 92

<212> DNA

<213> 人工序列

<220>

<223> 具有FVIII信号序列和BsiW1位点的LW 22-FWD-VWF-D'D3

<400> 95

gcgccggccg tacgatgcaa atagagctct ccacctgctt ctttctgtgc cttttgcgat 60
tctgcttttag cctatcctgt cgcccccca tg 92

<210> 96

<211> 47

<212> DNA

<213> 人工序列

<220>

<223> 具有终止密码子和Not1位点的LW 23-Rev- Fc

<400> 96

tcatcaatgt atcttatcat gtctgaattc gcggccgctc atttacc 47

<210> 97

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 具有BsiW1位点的LW24- Fwd- VWF D1D2D'D3克隆寡核苷酸

<400> 97

gcgccggccg tacgatgatt cctgccagat ttgccggggt g 41

<210> 98

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 具有EcoRV的LW27-Rev-VWF D'D3寡核苷酸

<400> 98

ccaccgccag atatcggctc ctggcaggct tcacaggtga g 41

<210> 99

<211> 1240

<212> PRT

<213> 人工序列

<220>

<223> VWF-D1D2D'D3蛋白序列1

<400> 99

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
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
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
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			
	1010	1015	1020
Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro			
	1025	1030	1035
Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln			
	1040	1045	1050
Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe			
	1055	1060	1065
Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val			
	1070	1075	1080
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala			
	1085	1090	1095
Cys Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln			
	1100	1105	1110
His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln			
	1115	1120	1125
Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Cys Glu			
	1130	1135	1140
Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln			
	1145	1150	1155
His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu Gly Cys			

1160	1165	1170
His Ala His Cys Pro Pro Gly Lys Ile Leu Asp Glu Leu Leu Gln		
1175	1180	1185
Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu Val Ala Gly		
1190	1195	1200
Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro Ser Asp		
1205	1210	1215
Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn Leu Thr		
1220	1225	1230
Cys Glu Ala Cys Gln Glu Pro		
1235	1240	
<210> 100		
<211> 477		
<212> PRT		
<213> 人工序列		
<220>		
<223> VWF-D'D3蛋白序列2		
<400> 100		
Ser Leu Ser Cys Arg Pro Pro Met Val Lys Leu Val Cys Pro Ala Asp		
1 5 10 15		
Asn Leu Arg Ala Glu Gly Leu Glu Cys Thr Lys Thr Cys Gln Asn Tyr		
20 25 30		
Asp Leu Glu Cys Met Ser Met Gly Cys Val Ser Gly Cys Leu Cys Pro		
35 40 45		
Pro Gly Met Val Arg His Glu Asn Arg Cys Val Ala Leu Glu Arg Cys		
50 55 60		
Pro Cys Phe His Gln Gly Lys Glu Tyr Ala Pro Gly Glu Thr Val Lys		
65 70 75 80		
Ile Gly Cys Asn Thr Cys Val Cys Arg Asp Arg Lys Trp Asn Cys Thr		
85 90 95		
Asp His Val Cys Asp Ala Thr Cys Ser Thr Ile Gly Met Ala His Tyr		
100 105 110		
Leu Thr Phe Asp Gly Leu Lys Tyr Leu Phe Pro Gly Glu Cys Gln Tyr		
115 120 125		
Val Leu Val Gln Asp Tyr Cys Gly Ser Asn Pro Gly Thr Phe Arg Ile		
130 135 140		
Leu Val Gly Asn Lys Gly Cys Ser His Pro Ser Val Lys Cys Lys Lys		
145 150 155 160		
Arg Val Thr Ile Leu Val Glu Gly Gly Glu Ile Glu Leu Phe Asp Gly		

				165						170					175		
Glu	Val	Asn	Val	Lys	Arg	Pro	Met	Lys	Asp	Glu	Thr	His	Phe	Glu	Val		
				180						185					190		
Val	Glu	Ser	Gly	Arg	Tyr	Ile	Ile	Leu	Leu	Leu	Gly	Lys	Ala	Leu	Ser		
				195						200					205		
Val	Val	Trp	Asp	Arg	His	Leu	Ser	Ile	Ser	Val	Val	Leu	Lys	Gln	Thr		
				210						215					220		
Tyr	Gln	Glu	Lys	Val	Cys	Gly	Leu	Cys	Gly	Asn	Phe	Asp	Gly	Ile	Gln		
225										230					240		
Asn	Asn	Asp	Leu	Thr	Ser	Ser	Asn	Leu	Gln	Val	Glu	Glu	Asp	Pro	Val		
				245						250					255		
Asp	Phe	Gly	Asn	Ser	Trp	Lys	Val	Ser	Ser	Gln	Cys	Ala	Asp	Thr	Arg		
				260						265					270		
Lys	Val	Pro	Leu	Asp	Ser	Ser	Pro	Ala	Thr	Cys	His	Asn	Asn	Ile	Met		
				275						280					285		
Lys	Gln	Thr	Met	Val	Asp	Ser	Ser	Cys	Arg	Ile	Leu	Thr	Ser	Asp	Val		
				290						295					300		
Phe	Gln	Asp	Cys	Asn	Lys	Leu	Val	Asp	Pro	Glu	Pro	Tyr	Leu	Asp	Val		
305										310					320		
Cys	Ile	Tyr	Asp	Thr	Cys	Ser	Cys	Glu	Ser	Ile	Gly	Asp	Cys	Ala	Cys		
				325						330					335		
Phe	Cys	Asp	Thr	Ile	Ala	Ala	Tyr	Ala	His	Val	Cys	Ala	Gln	His	Gly		
				340						345					350		
Lys	Val	Val	Thr	Trp	Arg	Thr	Ala	Thr	Leu	Cys	Pro	Gln	Ser	Cys	Glu		
				355						360					365		
Glu	Arg	Asn	Leu	Arg	Glu	Asn	Gly	Tyr	Glu	Cys	Glu	Trp	Arg	Tyr	Asn		
				370						375					380		
Ser	Cys	Ala	Pro	Ala	Cys	Gln	Val	Thr	Cys	Gln	His	Pro	Glu	Pro	Leu		
385										390					400		
Ala	Cys	Pro	Val	Gln	Cys	Val	Glu	Gly	Cys	His	Ala	His	Cys	Pro	Pro		
				405						410					415		
Gly	Lys	Ile	Leu	Asp	Glu	Leu	Leu	Gln	Thr	Cys	Val	Asp	Pro	Glu	Asp		
				420						425					430		
Cys	Pro	Val	Cys	Glu	Val	Ala	Gly	Arg	Arg	Phe	Ala	Ser	Gly	Lys	Lys		
				435						440					445		
Val	Thr	Leu	Asn	Pro	Ser	Asp	Pro	Glu	His	Cys	Gln	Ile	Cys	His	Cys		
				450						455					460		
Asp	Val	Val	Asn	Leu	Thr	Cys	Glu	Ala	Cys	Gln	Glu	Pro					
465										470					475		

<210> 101

<211> 2754

<212> PRT

<213> 人工序列

<220>

<223> pSYN-FVIII-161

<400> 101

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe

1 5 10 15

Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser

20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg

35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val

50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile

65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln

85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser

100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser

115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp

130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu

145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser

165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile

180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr

195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly

210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp

225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr

245 250 255

Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val			
			260						265					270				
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile			
			275						280					285				
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser			
			290						295					300				
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met			
305											310				320			
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His			
														325				
														330				
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro			
														340				
														345				
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp			
														355				
														360				
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser			
														370				
														375				
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr			
385														390				
														395				
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro			
														405				
														410				
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn			
														420				
														425				
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met			
														435				
														440				
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu			
														450				
														455				
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu			
465														470				
														475				
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro			
														485				
														490				
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys			
														500				
														505				
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe			
														515				
														520				
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp			
														530				
														535				
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg			
545														550				
														555				
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu			

	565		570		575
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val					
	580		585		590
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu					
	595		600		605
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp					
	610		615		620
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val					
625		630		635	640
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp					
	645		650		655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe					
	660		665		670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr					
	675		680		685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro					
	690		695		700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly					
705		710		715	720
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp					
	725		730		735
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys					
	740		745		750
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu					
	755		760		765
Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln					
	770		775		780
Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu					
785		790		795	800
Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe					
	805		810		815
Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp					
	820		825		830
Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln					
	835		840		845
Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr					
	850		855		860
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His					
865		870		875	880

Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile		
885	890	895
Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser		
900	905	910
Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg		
915	920	925
Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val		
930	935	940
Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp		
945	950	955
Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu		
965	970	975
Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His		
980	985	990
Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe		
995	1000	1005
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn		
1010	1015	1020
Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys		
1025	1030	1035
Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr		
1040	1045	1050
Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr		
1055	1060	1065
Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe		
1070	1075	1080
Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met		
1085	1090	1095
Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met		
1100	1105	1110
Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly		
1115	1120	1125
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1130	1135	1140
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1145	1150	1155
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1160	1165	1170
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		

1175	1180	1185
Thr Lys Glu Pro Phe Ser	Trp Ile Lys Val Asp	Leu Leu Ala Pro
1190	1195	1200
Met Ile Ile His Gly Ile	Lys Thr Gln Gly Ala	Arg Gln Lys Phe
1205	1210	1215
Ser Ser Leu Tyr Ile Ser	Gln Phe Ile Ile Met	Tyr Ser Leu Asp
1220	1225	1230
Gly Lys Lys Trp Gln Thr	Tyr Arg Gly Asn Ser	Thr Gly Thr Leu
1235	1240	1245
Met Val Phe Phe Gly Asn	Val Asp Ser Ser Gly	Ile Lys His Asn
1250	1255	1260
Ile Phe Asn Pro Pro Ile	Ile Ala Arg Tyr Ile	Arg Leu His Pro
1265	1270	1275
Thr His Tyr Ser Ile Arg	Ser Thr Leu Arg Met	Glu Leu Met Gly
1280	1285	1290
Cys Asp Leu Asn Ser Cys	Ser Met Pro Leu Gly	Met Glu Ser Lys
1295	1300	1305
Ala Ile Ser Asp Ala Gln	Ile Thr Ala Ser Ser	Tyr Phe Thr Asn
1310	1315	1320
Met Phe Ala Thr Trp Ser	Pro Ser Lys Ala Arg	Leu His Leu Gln
1325	1330	1335
Gly Arg Ser Asn Ala Trp	Arg Pro Gln Val Asn	Asn Pro Lys Glu
1340	1345	1350
Trp Leu Gln Val Asp Phe	Gln Lys Thr Met Lys	Val Thr Gly Val
1355	1360	1365
Thr Thr Gln Gly Val Lys	Ser Leu Leu Thr Ser	Met Tyr Val Lys
1370	1375	1380
Glu Phe Leu Ile Ser Ser	Ser Gln Asp Gly His	Gln Trp Thr Leu
1385	1390	1395
Phe Phe Gln Asn Gly Lys	Val Lys Val Phe Gln	Gly Asn Gln Asp
1400	1405	1410
Ser Phe Thr Pro Val Val	Asn Ser Leu Asp Pro	Pro Leu Leu Thr
1415	1420	1425
Arg Tyr Leu Arg Ile His	Pro Gln Ser Trp Val	His Gln Ile Ala
1430	1435	1440
Leu Arg Met Glu Val Leu	Gly Cys Glu Ala Gln	Asp Leu Tyr Asp
1445	1450	1455
Lys Thr His Thr Cys Pro	Pro Cys Pro Ala Pro	Glu Leu Leu Gly
1460	1465	1470

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu		
1475	1480	1485
Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val		
1490	1495	1500
Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly		
1505	1510	1515
Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr		
1520	1525	1530
Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln		
1535	1540	1545
Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys		
1550	1555	1560
Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly		
1565	1570	1575
Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp		
1580	1585	1590
Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly		
1595	1600	1605
Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln		
1610	1615	1620
Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp		
1625	1630	1635
Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg		
1640	1645	1650
Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala		
1655	1660	1665
Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly		
1670	1675	1680
Lys Arg Arg Arg Arg Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly		
1685	1690	1695
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly		
1700	1705	1710
Ser Gly Gly Gly Gly Ser Arg Lys Arg Arg Lys Arg Ser Leu Ser		
1715	1720	1725
Cys Arg Pro Pro Met Val Lys Leu Val Cys Pro Ala Asp Asn Leu		
1730	1735	1740
Arg Ala Glu Gly Leu Glu Cys Thr Lys Thr Cys Gln Asn Tyr Asp		
1745	1750	1755
Leu Glu Cys Met Ser Met Gly Cys Val Ser Gly Cys Leu Cys Pro		

1760	1765	1770
Pro Gly Met Val Arg His	Glu Asn Arg Cys Val	Ala Leu Glu Arg
1775	1780	1785
Cys Pro Cys Phe His Gln	Gly Lys Glu Tyr Ala	Pro Gly Glu Thr
1790	1795	1800
Val Lys Ile Gly Cys Asn	Thr Cys Val Cys Arg	Asp Arg Lys Trp
1805	1810	1815
Asn Cys Thr Asp His Val	Cys Asp Ala Thr Cys	Ser Thr Ile Gly
1820	1825	1830
Met Ala His Tyr Leu Thr	Phe Asp Gly Leu Lys	Tyr Leu Phe Pro
1835	1840	1845
Gly Glu Cys Gln Tyr Val	Leu Val Gln Asp Tyr	Cys Gly Ser Asn
1850	1855	1860
Pro Gly Thr Phe Arg Ile	Leu Val Gly Asn Lys	Gly Cys Ser His
1865	1870	1875
Pro Ser Val Lys Cys Lys	Lys Arg Val Thr Ile	Leu Val Glu Gly
1880	1885	1890
Gly Glu Ile Glu Leu Phe	Asp Gly Glu Val Asn	Val Lys Arg Pro
1895	1900	1905
Met Lys Asp Glu Thr His	Phe Glu Val Val Glu	Ser Gly Arg Tyr
1910	1915	1920
Ile Ile Leu Leu Leu Gly	Lys Ala Leu Ser Val	Val Trp Asp Arg
1925	1930	1935
His Leu Ser Ile Ser Val	Val Leu Lys Gln Thr	Tyr Gln Glu Lys
1940	1945	1950
Val Cys Gly Leu Cys Gly	Asn Phe Asp Gly Ile	Gln Asn Asn Asp
1955	1960	1965
Leu Thr Ser Ser Asn Leu	Gln Val Glu Glu Asp	Pro Val Asp Phe
1970	1975	1980
Gly Asn Ser Trp Lys Val	Ser Ser Gln Cys Ala	Asp Thr Arg Lys
1985	1990	1995
Val Pro Leu Asp Ser Ser	Pro Ala Thr Cys His	Asn Asn Ile Met
2000	2005	2010
Lys Gln Thr Met Val Asp	Ser Ser Cys Arg Ile	Leu Thr Ser Asp
2015	2020	2025
Val Phe Gln Asp Cys Asn	Lys Leu Val Asp Pro	Glu Pro Tyr Leu
2030	2035	2040
Asp Val Cys Ile Tyr Asp	Thr Cys Ser Cys Glu	Ser Ile Gly Asp
2045	2050	2055

Cys Ala Ala Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys		
2060	2065	2070
Ala Gln His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys		
2075	2080	2085
Pro Gln Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu		
2090	2095	2100
Ala Glu Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr		
2105	2110	2115
Cys Gln His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu		
2120	2125	2130
Gly Cys His Ala His Cys Pro Pro Gly Lys Ile Leu Asp Glu Leu		
2135	2140	2145
Leu Gln Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu Val		
2150	2155	2160
Ala Gly Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro		
2165	2170	2175
Ser Asp Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn		
2180	2185	2190
Leu Thr Cys Glu Ala Cys Gln Glu Pro Ile Ser Gly Thr Ser Glu		
2195	2200	2205
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser		
2210	2215	2220
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser		
2225	2230	2235
Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
2240	2245	2250
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr		
2255	2260	2265
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro		
2270	2275	2280
Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser		
2285	2290	2295
Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
2300	2305	2310
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala		
2315	2320	2325
Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro		
2330	2335	2340
Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser		

2345	2350	2355
Ala Pro Gly Thr Ser Glu	Ser Ala Thr Pro Glu	Ser Gly Pro Gly
2360	2365	2370
Thr Ser Glu Ser Ala Thr	Pro Glu Ser Gly Pro	Gly Thr Ser Glu
2375	2380	2385
Ser Ala Thr Pro Glu Ser	Gly Pro Gly Ser Glu	Pro Ala Thr Ser
2390	2395	2400
Gly Ser Glu Thr Pro Gly	Ser Glu Pro Ala Thr	Ser Gly Ser Glu
2405	2410	2415
Thr Pro Gly Ser Pro Ala	Gly Ser Pro Thr Ser	Thr Glu Glu Gly
2420	2425	2430
Thr Ser Thr Glu Pro Ser	Glu Gly Ser Ala Pro	Gly Thr Ser Thr
2435	2440	2445
Glu Pro Ser Glu Gly Ser	Ala Pro Gly Ser Glu	Pro Ala Thr Ser
2450	2455	2460
Gly Ser Glu Thr Pro Gly	Thr Ser Glu Ser Ala	Thr Pro Glu Ser
2465	2470	2475
Gly Pro Gly Thr Ser Thr	Glu Pro Ser Glu Gly	Ser Ala Pro Asp
2480	2485	2490
Ser Gly Gly Gly Gly Ser	Gly Gly Gly Gly Ser	Gly Gly Gly Gly
2495	2500	2505
Ser Gly Gly Gly Gly Ser	Gly Gly Gly Gly Ser	Leu Val Pro Arg
2510	2515	2520
Gly Ser Gly Gly Asp Lys	Thr His Thr Cys Pro	Pro Cys Pro Ala
2525	2530	2535
Pro Glu Leu Leu Gly Gly	Pro Ser Val Phe Leu	Phe Pro Pro Lys
2540	2545	2550
Pro Lys Asp Thr Leu Met	Ile Ser Arg Thr Pro	Glu Val Thr Cys
2555	2560	2565
Val Val Val Asp Val Ser	His Glu Asp Pro Glu	Val Lys Phe Asn
2570	2575	2580
Trp Tyr Val Asp Gly Val	Glu Val His Asn Ala	Lys Thr Lys Pro
2585	2590	2595
Arg Glu Glu Gln Tyr Asn	Ser Thr Tyr Arg Val	Val Ser Val Leu
2600	2605	2610
Thr Val Leu His Gln Asp	Trp Leu Asn Gly Lys	Glu Tyr Lys Cys
2615	2620	2625
Lys Val Ser Asn Lys Ala	Leu Pro Ala Pro Ile	Glu Lys Thr Ile
2630	2635	2640

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
 2645 2650 2655
 Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr
 2660 2665 2670
 Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
 2675 2680 2685
 Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 2690 2695 2700
 Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
 2705 2710 2715
 Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser
 2720 2725 2730
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
 2735 2740 2745

Ser Leu Ser Pro Gly Lys
 2750

<210> 102

<211> 2242

<212> PRT

<213> 人工序列

<220>

<223> pSYN-FVIII-170蛋白序列

<400> 102

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

Val	Leu	Val	Gln	Asp	Tyr	Cys	Gly	Ser	Asn	Pro	Gly	Thr	Phe	Arg	Ile
130						135					140				
Leu	Val	Gly	Asn	Lys	Gly	Cys	Ser	His	Pro	Ser	Val	Lys	Cys	Lys	Lys
145					150					155					160
Arg	Val	Thr	Ile	Leu	Val	Glu	Gly	Gly	Glu	Ile	Glu	Leu	Phe	Asp	Gly
				165					170					175	
Glu	Val	Asn	Val	Lys	Arg	Pro	Met	Lys	Asp	Glu	Thr	His	Phe	Glu	Val
				180				185					190		
Val	Glu	Ser	Gly	Arg	Tyr	Ile	Ile	Leu	Leu	Leu	Gly	Lys	Ala	Leu	Ser
				195				200				205			
Val	Val	Trp	Asp	Arg	His	Leu	Ser	Ile	Ser	Val	Val	Leu	Lys	Gln	Thr
				210				215				220			
Tyr	Gln	Glu	Lys	Val	Cys	Gly	Leu	Cys	Gly	Asn	Phe	Asp	Gly	Ile	Gln
225					230					235					240
Asn	Asn	Asp	Leu	Thr	Ser	Ser	Asn	Leu	Gln	Val	Glu	Glu	Asp	Pro	Val
				245					250					255	
Asp	Phe	Gly	Asn	Ser	Trp	Lys	Val	Ser	Ser	Gln	Cys	Ala	Asp	Thr	Arg
				260				265					270		
Lys	Val	Pro	Leu	Asp	Ser	Ser	Pro	Ala	Thr	Cys	His	Asn	Asn	Ile	Met
				275				280				285			
Lys	Gln	Thr	Met	Val	Asp	Ser	Ser	Cys	Arg	Ile	Leu	Thr	Ser	Asp	Val
				290				295				300			
Phe	Gln	Asp	Cys	Asn	Lys	Leu	Val	Asp	Pro	Glu	Pro	Tyr	Leu	Asp	Val
305					310					315					320
Cys	Ile	Tyr	Asp	Thr	Cys	Ser	Cys	Glu	Ser	Ile	Gly	Asp	Cys	Ala	Ala
				325					330					335	
Phe	Cys	Asp	Thr	Ile	Ala	Ala	Tyr	Ala	His	Val	Cys	Ala	Gln	His	Gly
				340				345					350		
Lys	Val	Val	Thr	Trp	Arg	Thr	Ala	Thr	Leu	Cys	Pro	Gln	Ser	Cys	Glu
				355				360				365			
Glu	Arg	Asn	Leu	Arg	Glu	Asn	Gly	Tyr	Glu	Ala	Glu	Trp	Arg	Tyr	Asn
				370				375				380			
Ser	Cys	Ala	Pro	Ala	Cys	Gln	Val	Thr	Cys	Gln	His	Pro	Glu	Pro	Leu
385					390					395					400
Ala	Cys	Pro	Val	Gln	Cys	Val	Glu	Gly	Cys	His	Ala	His	Cys	Pro	Pro
				405					410					415	
Gly	Lys	Ile	Leu	Asp	Glu	Leu	Leu	Gln	Thr	Cys	Val	Asp	Pro	Glu	Asp
				420				425				430			
Cys	Pro	Val	Cys	Glu	Val	Ala	Gly	Arg	Arg	Phe	Ala	Ser	Gly	Lys	Lys

435	440	445
Val Thr Leu Asn Pro Ser Asp	Pro Glu His Cys Gln Ile Cys His Cys	
450	455	460
Asp Val Val Asn Leu Thr Cys Glu Ala Cys Gln Glu Pro Ile Ser Gly		
465	470	475
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
485	490	495
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
500	505	510
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
515	520	525
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu		
530	535	540
Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser		
545	550	555
Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		
565	570	575
Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser		
580	585	590
Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser		
595	600	605
Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly		
610	615	620
Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser		
625	630	635
Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
645	650	655
Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		
660	665	670
Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala		
675	680	685
Thr Ser Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser		
690	695	700
Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		
705	710	715
Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro Ala		
725	730	735
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
740	745	750

Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Asp		
755	760	765
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser		
770	775	780
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Leu Val Pro Arg Gly Ser		
785	790	795
Gly Gly Ala Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu		
	805	810
		815
Ser Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala		
	820	825
		830
Arg Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val		
	835	840
		845
Val Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn		
	850	855
		860
Ile Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile		
865	870	875
		880
Gln Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala		
	885	890
		895
Ser His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala		
	900	905
		910
Ser Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu		
	915	920
		925
Asp Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val		
	930	935
		940
Leu Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr		
945	950	955
		960
Ser Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu		
	965	970
		975
Ile Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys		
	980	985
		990
Thr Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu		
	995	1000
		1005
Gly Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp		
	1010	1015
		1020
Arg Asp Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val		
	1025	1030
		1035
Asn Gly Tyr Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His		
	1040	1045
		1050
Arg Lys Ser Val Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro		

1055	1060	1065
Glu Val His Ser Ile Phe	Leu Glu Gly His Thr	Phe Leu Val Arg
1070	1075	1080
Asn His Arg Gln Ala Ser	Leu Glu Ile Ser Pro	Ile Thr Phe Leu
1085	1090	1095
Thr Ala Gln Thr Leu Leu	Met Asp Leu Gly Gln	Phe Leu Leu Phe
1100	1105	1110
Cys His Ile Ser Ser His	Gln His Asp Gly Met	Glu Ala Tyr Val
1115	1120	1125
Lys Val Asp Ser Cys Pro	Glu Glu Pro Gln Leu	Arg Met Lys Asn
1130	1135	1140
Asn Glu Glu Ala Glu Asp	Tyr Asp Asp Asp Leu	Thr Asp Ser Glu
1145	1150	1155
Met Asp Val Val Arg Phe	Asp Asp Asp Asn Ser	Pro Ser Phe Ile
1160	1165	1170
Gln Ile Arg Ser Val Ala	Lys Lys His Pro Lys	Thr Trp Val His
1175	1180	1185
Tyr Ile Ala Ala Glu Glu	Glu Asp Trp Asp Tyr	Ala Pro Leu Val
1190	1195	1200
Leu Ala Pro Asp Asp Arg	Ser Tyr Lys Ser Gln	Tyr Leu Asn Asn
1205	1210	1215
Gly Pro Gln Arg Ile Gly	Arg Lys Tyr Lys Lys	Val Arg Phe Met
1220	1225	1230
Ala Tyr Thr Asp Glu Thr	Phe Lys Thr Arg Glu	Ala Ile Gln His
1235	1240	1245
Glu Ser Gly Ile Leu Gly	Pro Leu Leu Tyr Gly	Glu Val Gly Asp
1250	1255	1260
Thr Leu Leu Ile Ile Phe	Lys Asn Gln Ala Ser	Arg Pro Tyr Asn
1265	1270	1275
Ile Tyr Pro His Gly Ile	Thr Asp Val Arg Pro	Leu Tyr Ser Arg
1280	1285	1290
Arg Leu Pro Lys Gly Val	Lys His Leu Lys Asp	Phe Pro Ile Leu
1295	1300	1305
Pro Gly Glu Ile Phe Lys	Tyr Lys Trp Thr Val	Thr Val Glu Asp
1310	1315	1320
Gly Pro Thr Lys Ser Asp	Pro Arg Cys Leu Thr	Arg Tyr Tyr Ser
1325	1330	1335
Ser Phe Val Asn Met Glu	Arg Asp Leu Ala Ser	Gly Leu Ile Gly
1340	1345	1350

Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn		
1355	1360	1365
Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe		
1370	1375	1380
Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe		
1385	1390	1395
Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe Gln		
1400	1405	1410
Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser		
1415	1420	1425
Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile		
1430	1435	1440
Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser		
1445	1450	1455
Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr		
1460	1465	1470
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn		
1475	1480	1485
Pro Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn		
1490	1495	1500
Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn		
1505	1510	1515
Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr		
1520	1525	1530
Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln		
1535	1540	1545
Asn Pro Pro Val Leu Lys Arg His Gln Arg Glu Ile Thr Arg Thr		
1550	1555	1560
Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile		
1565	1570	1575
Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp		
1580	1585	1590
Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr		
1595	1600	1605
Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser		
1610	1615	1620
Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro		
1625	1630	1635
Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe		

1640	1645	1650
Thr Gln Pro Leu Tyr Arg	Gly Glu Leu Asn Glu	His Leu Gly Leu
1655	1660	1665
Leu Gly Pro Tyr Ile Arg	Ala Glu Val Glu Asp	Asn Ile Met Val
1670	1675	1680
Thr Phe Arg Asn Gln Ala	Ser Arg Pro Tyr Ser	Phe Tyr Ser Ser
1685	1690	1695
Leu Ile Ser Tyr Glu Glu	Asp Gln Arg Gln Gly	Ala Glu Pro Arg
1700	1705	1710
Lys Asn Phe Val Lys Pro	Asn Glu Thr Lys Thr	Tyr Phe Trp Lys
1715	1720	1725
Val Gln His His Met Ala	Pro Thr Lys Asp Glu	Phe Asp Cys Lys
1730	1735	1740
Ala Trp Ala Tyr Phe Ser	Asp Val Asp Leu Glu	Lys Asp Val His
1745	1750	1755
Ser Gly Leu Ile Gly Pro	Leu Leu Val Cys His	Thr Asn Thr Leu
1760	1765	1770
Asn Pro Ala His Gly Arg	Gln Val Thr Val Gln	Glu Phe Ala Leu
1775	1780	1785
Phe Phe Thr Ile Phe Asp	Glu Thr Lys Ser Trp	Tyr Phe Thr Glu
1790	1795	1800
Asn Met Glu Arg Asn Cys	Arg Ala Pro Cys Asn	Ile Gln Met Glu
1805	1810	1815
Asp Pro Thr Phe Lys Glu	Asn Tyr Arg Phe His	Ala Ile Asn Gly
1820	1825	1830
Tyr Ile Met Asp Thr Leu	Pro Gly Leu Val Met	Ala Gln Asp Gln
1835	1840	1845
Arg Ile Arg Trp Tyr Leu	Leu Ser Met Gly Ser	Asn Glu Asn Ile
1850	1855	1860
His Ser Ile His Phe Ser	Gly His Val Phe Thr	Val Arg Lys Lys
1865	1870	1875
Glu Glu Tyr Lys Met Ala	Leu Tyr Asn Leu Tyr	Pro Gly Val Phe
1880	1885	1890
Glu Thr Val Glu Met Leu	Pro Ser Lys Ala Gly	Ile Trp Arg Val
1895	1900	1905
Glu Cys Leu Ile Gly Glu	His Leu His Ala Gly	Met Ser Thr Leu
1910	1915	1920
Phe Leu Val Tyr Ser Asn	Lys Cys Gln Thr Pro	Leu Gly Met Ala
1925	1930	1935

Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala	Ser Gly Gln Tyr
1940	1950
Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His	Tyr Ser Gly Ser
1955	1965
Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser	Trp Ile Lys Val
1970	1980
Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile	Lys Thr Gln Gly
1985	1995
Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser	Gln Phe Ile Ile
2000	2010
Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr	Tyr Arg Gly Asn
2015	2025
Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn	Val Asp Ser Ser
2030	2040
Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile	Ile Ala Arg Tyr
2045	2055
Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg	Ser Thr Leu Arg
2060	2070
Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys	Ser Met Pro Leu
2075	2085
Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln	Ile Thr Ala Ser
2090	2100
Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser	Pro Ser Lys Ala
2105	2115
Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp	Arg Pro Gln Val
2120	2130
Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe	Gln Lys Thr Met
2135	2145
Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys	Ser Leu Leu Thr
2150	2160
Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser	Ser Gln Asp Gly
2165	2175
His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys	Val Lys Val Phe
2180	2190
Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val	Asn Ser Leu Asp
2195	2205
Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His	Pro Gln Ser Trp
2210	2220
Val His Gln Ile Ala Leu Arg Met Glu Val Leu	Gly Cys Glu Ala

225	230	235	240
Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His			
	245	250	255
Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu			
	260	265	270
Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile			
	275	280	285
Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly			
	290	295	300
Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met			
305	310	315	320
Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg			
	325	330	335
Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp			
	340	345	350
Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro Ser Phe			
	355	360	365
Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr Trp Val His			
	370	375	380
Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu			
385	390	395	400
Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly Pro			
	405	410	415
Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr Thr			
	420	425	430
Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly Ile			
	435	440	445
Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile Ile			
	450	455	460
Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile			
465	470	475	480
Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys			
	485	490	495
His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr Lys			
	500	505	510
Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg Cys			
	515	520	525
Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu Ala			
530	535	540	

Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp			
545	550	555	560
Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe			
	565	570	575
Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln			
	580	585	590
Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe			
	595	600	605
Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser			
610	615	620	
Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu			
625	630	635	640
Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr			
	645	650	655
Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro			
	660	665	670
Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp			
	675	680	685
Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala			
690	695	700	
Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu			
705	710	715	720
Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala			
	725	730	735
Ile Glu Pro Arg Ser Phe Ser Gln Asn Gly Ala Pro Gly Thr Ser Glu			
	740	745	750
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly			
	755	760	765
Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro			
	770	775	780
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu			
785	790	795	800
Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu			
	805	810	815
Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu			
	820	825	830
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro			
	835	840	845
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro			

850	855	860
Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu		
865	870	875
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr		880
	885	890
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro		895
	900	905
Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro		910
	915	920
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro		925
	930	935
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser Gly		940
945	950	955
Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu		960
	965	970
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr		975
	980	985
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro Ala Thr Ser Gly		990
	995	1000
Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		1005
	1010	1015
Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Ala Ser		1020
	1025	1030
Ser Pro Pro Val Leu Lys Arg His Gln Ala Glu Ile Thr Arg Thr		1035
	1040	1045
Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile		1050
	1055	1060
Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp		1065
	1070	1075
Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr		1080
	1085	1090
Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser		1095
	1100	1105
Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro		1110
	1115	1120
Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe		1125
	1130	1135
Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu		1140
	1145	1150
		1155

Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val		
1160	1165	1170
Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser		
1175	1180	1185
Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg		
1190	1195	1200
Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys		
1205	1210	1215
Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys		
1220	1225	1230
Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His		
1235	1240	1245
Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu		
1250	1255	1260
Asn Pro Ala His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu		
1265	1270	1275
Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu		
1280	1285	1290
Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu		
1295	1300	1305
Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly		
1310	1315	1320
Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln		
1325	1330	1335
Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile		
1340	1345	1350
His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys		
1355	1360	1365
Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe		
1370	1375	1380
Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val		
1385	1390	1395
Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu		
1400	1405	1410
Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala		
1415	1420	1425
Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr		
1430	1435	1440
Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser		

1445	1450	1455
Ile Asn Ala Trp Ser Thr	Lys Glu Pro Phe Ser	Trp Ile Lys Val
1460	1465	1470
Asp Leu Leu Ala Pro Met	Ile Ile His Gly Ile	Lys Thr Gln Gly
1475	1480	1485
Ala Arg Gln Lys Phe Ser	Ser Leu Tyr Ile Ser	Gln Phe Ile Ile
1490	1495	1500
Met Tyr Ser Leu Asp Gly	Lys Lys Trp Gln Thr	Tyr Arg Gly Asn
1505	1510	1515
Ser Thr Gly Thr Leu Met	Val Phe Phe Gly Asn	Val Asp Ser Ser
1520	1525	1530
Gly Ile Lys His Asn Ile	Phe Asn Pro Pro Ile	Ile Ala Arg Tyr
1535	1540	1545
Ile Arg Leu His Pro Thr	His Tyr Ser Ile Arg	Ser Thr Leu Arg
1550	1555	1560
Met Glu Leu Met Gly Cys	Asp Leu Asn Ser Cys	Ser Met Pro Leu
1565	1570	1575
Gly Met Glu Ser Lys Ala	Ile Ser Asp Ala Gln	Ile Thr Ala Ser
1580	1585	1590
Ser Tyr Phe Thr Asn Met	Phe Ala Thr Trp Ser	Pro Ser Lys Ala
1595	1600	1605
Arg Leu His Leu Gln Gly	Arg Ser Asn Ala Trp	Arg Pro Gln Val
1610	1615	1620
Asn Asn Pro Lys Glu Trp	Leu Gln Val Asp Phe	Gln Lys Thr Met
1625	1630	1635
Lys Val Thr Gly Val Thr	Thr Gln Gly Val Lys	Ser Leu Leu Thr
1640	1645	1650
Ser Met Tyr Val Lys Glu	Phe Leu Ile Ser Ser	Ser Gln Asp Gly
1655	1660	1665
His Gln Trp Thr Leu Phe	Phe Gln Asn Gly Lys	Val Lys Val Phe
1670	1675	1680
Gln Gly Asn Gln Asp Ser	Phe Thr Pro Val Val	Asn Ser Leu Asp
1685	1690	1695
Pro Pro Leu Leu Thr Arg	Tyr Leu Arg Ile His	Pro Gln Ser Trp
1700	1705	1710
Val His Gln Ile Ala Leu	Arg Met Glu Val Leu	Gly Cys Glu Ala
1715	1720	1725
Gln Asp Leu Tyr Asp Lys	Thr His Thr Cys Pro	Pro Cys Pro Ala
1730	1735	1740

Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys
1745						1750						1755		
Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys
1760						1765						1770		
Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn
1775						1780						1785		
Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro
1790						1795						1800		
Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu
1805						1810						1815		
Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys
1820						1825						1830		
Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile
1835						1840						1845		
Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu
1850						1855						1860		
Pro	Pro	Ser	Arg	Asp	Glu	Leu	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr
1865						1870						1875		
Cys	Leu	Val	Lys	Gly	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp
1880						1885						1890		
Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro	Pro
1895						1900						1905		
Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr	Ser	Lys	Leu	Thr
1910						1915						1920		
Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn	Val	Phe	Ser	Cys	Ser
1925						1930						1935		
Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr	Gln	Lys	Ser	Leu
1940						1945						1950		

Ser Leu Ser Pro Gly Lys

1955

<210> 104

<211> 1959

<212> PRT

<213> 人工序列

<220>

<223> pSYN-FVIII-173成熟蛋白

<400> 104

Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr

1

5

10

15

Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg	Phe	Pro	Pro
20				25				30							
Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val	Tyr	Lys	Lys
35				40				45							
Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile	Ala	Lys	Pro
50				55				60							
Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln	Ala	Glu	Val
65				70				75				80			
Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser	His	Pro	Val
85				90				95							
Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser	Glu	Gly	Ala
100				105				110							
Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp	Asp	Lys	Val
115				120				125							
Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu	Lys	Glu	Asn
130				135				140							
Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser	Tyr	Leu	Ser
145				150				155				160			
His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile	Gly	Ala	Leu
165				170				175							
Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr	Gln	Thr	Leu
180				185				190							
His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly	Lys	Ser	Trp
195				200				205							
His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp	Ala	Ala	Ser
210				215				220							
Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr	Val	Asn	Arg
225				230				235				240			
Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val	Tyr	Trp	His
245				250				255							
Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile	Phe	Leu	Glu
260				265				270							
Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser	Leu	Glu	Ile
275				280				285							
Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met	Asp	Leu	Gly
290				295				300							
Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His	Asp	Gly	Met
305				310				315				320			
Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro	Gln	Leu	Arg

				325					330					335			
Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	Leu	Thr	Asp		
				340					345					350			
Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	Phe		
				355					360					365			
Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His		
				370					375					380			
Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	Leu		
385																	
				390					395					400			
Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro		
				405					410					415			
Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr		
				420					425					430			
Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile		
				435					440					445			
Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile		
				450					455					460			
Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly	Ile		
465																	
				470					475					480			
Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val	Lys		
				485					490					495			
His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys		
				500					505					510			
Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys		
				515					520					525			
Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala		
				530					535					540			
Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp		
545																	
				550					555					560			
Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe		
				565					570					575			
Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile	Gln		
				580					585					590			
Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu	Phe		
				595					600					605			
Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp	Ser		
				610					615					620			
Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile	Leu		
625																	
				630					635					640			

Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr
645 650 655
Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro
660 665 670
Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp
675 680 685
Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala
690 695 700
Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu
705 710 715 720
Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala
725 730 735
Ile Glu Pro Arg Ser Phe Ser Gln Asn Gly Ala Pro Gly Thr Ser Glu
740 745 750
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly
755 760 765
Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro
770 775 780
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu
785 790 795 800
Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu
805 810 815
Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu
820 825 830
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro
835 840 845
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro
850 855 860
Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu
865 870 875 880
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr
885 890 895
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro
900 905 910
Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro
915 920 925
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro
930 935 940
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser Gly

945	950	955	960
Ser Glu Thr Pro Gly	Ser Pro Ala Gly	Ser Pro Thr Ser Thr	Glu Glu
	965	970	975
Gly Thr Ser Thr Glu	Pro Ser Glu Gly	Ser Ala Pro Gly	Thr Ser Thr
	980	985	990
Glu Pro Ser Glu Gly	Ser Ala Pro Gly	Ser Glu Pro Ala	Thr Ser Gly
	995	1000	1005
Ser Glu Thr Pro Gly	Thr Ser Glu Ser Ala	Thr Pro Glu Ser	Gly
	1010	1015	1020
Pro Gly Thr Ser Thr	Glu Pro Ser Glu Gly	Ser Ala Pro Ala	Ser
	1025	1030	1035
Ser Pro Pro Val Leu	Lys Arg His Gln Arg	Glu Ile Thr Arg	Thr
	1040	1045	1050
Thr Leu Gln Ser Asp	Gln Glu Glu Ile Asp	Tyr Asp Asp Thr	Ile
	1055	1060	1065
Ser Val Glu Met Lys	Lys Glu Asp Phe Asp	Ile Tyr Asp Glu	Asp
	1070	1075	1080
Glu Asn Gln Ser Pro	Arg Ser Phe Gln Lys	Lys Thr Arg His	Tyr
	1085	1090	1095
Phe Ile Ala Ala Val	Glu Arg Leu Trp Asp	Tyr Gly Met Ser	Ser
	1100	1105	1110
Ser Pro His Val Leu	Arg Asn Arg Ala Gln	Ser Gly Ser Val	Pro
	1115	1120	1125
Gln Phe Lys Lys Val	Val Phe Gln Glu Phe	Thr Asp Gly Ser	Phe
	1130	1135	1140
Thr Gln Pro Leu Tyr	Arg Gly Glu Leu Asn	Glu His Leu Gly	Leu
	1145	1150	1155
Leu Gly Pro Tyr Ile	Arg Ala Glu Val Glu	Asp Asn Ile Met	Val
	1160	1165	1170
Thr Phe Arg Asn Gln	Ala Ser Arg Pro Tyr	Ser Phe Tyr Ser	Ser
	1175	1180	1185
Leu Ile Ser Tyr Glu	Glu Asp Gln Arg Gln	Gly Ala Glu Pro	Arg
	1190	1195	1200
Lys Asn Phe Val Lys	Pro Asn Glu Thr Lys	Thr Tyr Phe Trp	Lys
	1205	1210	1215
Val Gln His His Met	Ala Pro Thr Lys Asp	Glu Phe Asp Cys	Lys
	1220	1225	1230
Ala Trp Ala Tyr Phe	Ser Asp Val Asp Leu	Glu Lys Asp Val	His
	1235	1240	1245

Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu		
1250	1255	1260
Asn Pro Ala His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu		
1265	1270	1275
Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu		
1280	1285	1290
Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu		
1295	1300	1305
Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly		
1310	1315	1320
Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln		
1325	1330	1335
Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile		
1340	1345	1350
His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys		
1355	1360	1365
Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe		
1370	1375	1380
Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val		
1385	1390	1395
Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu		
1400	1405	1410
Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala		
1415	1420	1425
Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr		
1430	1435	1440
Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser		
1445	1450	1455
Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val		
1460	1465	1470
Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly		
1475	1480	1485
Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile		
1490	1495	1500
Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn		
1505	1510	1515
Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser		
1520	1525	1530
Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr		

1535	1540	1545
Ile Arg Leu His Pro Thr	His Tyr Ser Ile Arg	Ser Thr Leu Arg
1550	1555	1560
Met Glu Leu Met Gly Cys	Asp Leu Asn Ser Cys	Ser Met Pro Leu
1565	1570	1575
Gly Met Glu Ser Lys Ala	Ile Ser Asp Ala Gln	Ile Thr Ala Ser
1580	1585	1590
Ser Tyr Phe Thr Asn Met	Phe Ala Thr Trp Ser	Pro Ser Lys Ala
1595	1600	1605
Arg Leu His Leu Gln Gly	Arg Ser Asn Ala Trp	Arg Pro Gln Val
1610	1615	1620
Asn Asn Pro Lys Glu Trp	Leu Gln Val Asp Phe	Gln Lys Thr Met
1625	1630	1635
Lys Val Thr Gly Val Thr	Thr Gln Gly Val Lys	Ser Leu Leu Thr
1640	1645	1650
Ser Met Tyr Val Lys Glu	Phe Leu Ile Ser Ser	Ser Gln Asp Gly
1655	1660	1665
His Gln Trp Thr Leu Phe	Phe Gln Asn Gly Lys	Val Lys Val Phe
1670	1675	1680
Gln Gly Asn Gln Asp Ser	Phe Thr Pro Val Val	Asn Ser Leu Asp
1685	1690	1695
Pro Pro Leu Leu Thr Arg	Tyr Leu Arg Ile His	Pro Gln Ser Trp
1700	1705	1710
Val His Gln Ile Ala Leu	Arg Met Glu Val Leu	Gly Cys Glu Ala
1715	1720	1725
Gln Asp Leu Tyr Asp Lys	Thr His Thr Cys Pro	Pro Cys Pro Ala
1730	1735	1740
Pro Glu Leu Leu Gly Gly	Pro Ser Val Phe Leu	Phe Pro Pro Lys
1745	1750	1755
Pro Lys Asp Thr Leu Met	Ile Ser Arg Thr Pro	Glu Val Thr Cys
1760	1765	1770
Val Val Val Asp Val Ser	His Glu Asp Pro Glu	Val Lys Phe Asn
1775	1780	1785
Trp Tyr Val Asp Gly Val	Glu Val His Asn Ala	Lys Thr Lys Pro
1790	1795	1800
Arg Glu Glu Gln Tyr Asn	Ser Thr Tyr Arg Val	Val Ser Val Leu
1805	1810	1815
Thr Val Leu His Gln Asp	Trp Leu Asn Gly Lys	Glu Tyr Lys Cys
1820	1825	1830

Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
 1835 1840 1845
 Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
 1850 1855 1860
 Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr
 1865 1870 1875
 Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
 1880 1885 1890
 Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 1895 1900 1905
 Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
 1910 1915 1920
 Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser
 1925 1930 1935
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
 1940 1945 1950

Ser Leu Ser Pro Gly Lys
 1955

<210> 105

<211> 1984

<212> PRT

<213> 人工序列

<220>

<223> FVIII 195蛋白序列

<400> 105

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

His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
115				120				125							
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
130				135				140							
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
145				150				155				160			
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
165				170				175							
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
180				185				190							
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
195				200				205							
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
210				215				220							
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
225				230				235				240			
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
245				250				255							
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
260				265				270							
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
275				280				285							
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
290				295				300							
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met
305				310				315				320			
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His
325				330				335							
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro
340				345				350							
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp
355				360				365							
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser
370				375				380							
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr
385				390				395				400			
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro
405				410				415							
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn

420	425	430
Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met		
435	440	445
Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu		
450	455	460
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu		
465	470	475
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro		
485	490	495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys		
500	505	510
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe		
515	520	525
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp		
530	535	540
Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg		
545	550	555
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu		
565	570	575
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val		
580	585	590
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu		
595	600	605
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp		
610	615	620
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val		
625	630	635
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp		
645	650	655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe		
660	665	670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr		
675	680	685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro		
690	695	700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly		
705	710	715
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp		
725	730	735

Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys			
740	745	750	
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu			
755	760	765	
Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Gly Ala Pro			
770	775	780	
Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro			
785	790	795	800
Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser			
805	810	815	
Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro			
820	825	830	
Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Thr Pro Gly			
835	840	845	
Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser			
850	855	860	
Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro			
865	870	875	880
Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser Thr			
885	890	895	
Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala			
900	905	910	
Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro			
915	920	925	
Ala Ser Ser Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser			
930	935	940	
Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn			
945	950	955	960
Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala			
965	970	975	
Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser Pro His Val			
980	985	990	
Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val			
995	1000	1005	
Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr			
1010	1015	1020	
Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr Ile			
1025	1030	1035	
Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln			

1040	1045	1050
Ala Ser Arg Pro Tyr Ser	Phe Tyr Ser Ser Leu	Ile Ser Tyr Glu
1055	1060	1065
Glu Asp Gln Arg Gln Gly	Ala Glu Pro Arg Lys	Asn Phe Val Lys
1070	1075	1080
Pro Asn Glu Thr Lys Thr	Tyr Phe Trp Lys Val	Gln His His Met
1085	1090	1095
Ala Pro Thr Lys Asp Glu	Phe Asp Cys Lys Ala	Trp Ala Tyr Phe
1100	1105	1110
Ser Asp Val Asp Leu Glu	Lys Asp Val His Ser	Gly Leu Ile Gly
1115	1120	1125
Pro Leu Leu Val Cys His	Thr Asn Thr Leu Asn	Pro Ala His Gly
1130	1135	1140
Arg Gln Val Thr Val Gln	Glu Phe Ala Leu Phe	Phe Thr Ile Phe
1145	1150	1155
Asp Glu Thr Lys Ser Trp	Tyr Phe Thr Glu Asn	Met Glu Arg Asn
1160	1165	1170
Cys Arg Gly Ala Pro Thr	Ser Glu Ser Ala Thr	Pro Glu Ser Gly
1175	1180	1185
Pro Gly Ser Glu Pro Ala	Thr Ser Gly Ser Glu	Thr Pro Gly Thr
1190	1195	1200
Ser Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly	Ser Glu Pro Ala
1205	1210	1215
Thr Ser Gly Ser Glu Thr	Pro Gly Thr Ser Glu	Ser Ala Thr Pro
1220	1225	1230
Glu Ser Gly Pro Gly Thr	Ser Thr Glu Pro Ser	Glu Gly Ser Ala
1235	1240	1245
Pro Gly Thr Ser Glu Ser	Ala Thr Pro Glu Ser	Gly Pro Gly Ser
1250	1255	1260
Pro Ala Gly Ser Pro Thr	Ser Thr Glu Glu Gly	Ser Pro Ala Gly
1265	1270	1275
Ser Pro Thr Ser Thr Glu	Glu Gly Ser Pro Ala	Gly Ser Pro Thr
1280	1285	1290
Ser Thr Glu Glu Gly Thr	Ser Glu Ser Ala Thr	Pro Glu Ser Gly
1295	1300	1305
Pro Gly Thr Ser Thr Glu	Pro Ser Glu Gly Ser	Ala Pro Gly Ala
1310	1315	1320
Ser Ser Ala Pro Cys Asn	Ile Gln Met Glu Asp	Pro Thr Phe Lys
1325	1330	1335

Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr		
1340	1345	1350
Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr		
1355	1360	1365
Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe		
1370	1375	1380
Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met		
1385	1390	1395
Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met		
1400	1405	1410
Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly		
1415	1420	1425
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1430	1435	1440
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1445	1450	1455
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1460	1465	1470
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1475	1480	1485
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1490	1495	1500
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1505	1510	1515
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1520	1525	1530
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1535	1540	1545
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1550	1555	1560
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1565	1570	1575
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1580	1585	1590
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1595	1600	1605
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1610	1615	1620
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		

1625	1630	1635
Gly Arg Ser Asn Ala Trp	Arg Pro Gln Val Asn	Asn Pro Lys Glu
1640	1645	1650
Trp Leu Gln Val Asp Phe	Gln Lys Thr Met Lys	Val Thr Gly Val
1655	1660	1665
Thr Thr Gln Gly Val Lys	Ser Leu Leu Thr Ser	Met Tyr Val Lys
1670	1675	1680
Glu Phe Leu Ile Ser Ser	Ser Gln Asp Gly His	Gln Trp Thr Leu
1685	1690	1695
Phe Phe Gln Asn Gly Lys	Val Lys Val Phe Gln	Gly Asn Gln Asp
1700	1705	1710
Ser Phe Thr Pro Val Val	Asn Ser Leu Asp Pro	Pro Leu Leu Thr
1715	1720	1725
Arg Tyr Leu Arg Ile His	Pro Gln Ser Trp Val	His Gln Ile Ala
1730	1735	1740
Leu Arg Met Glu Val Leu	Gly Cys Glu Ala Gln	Asp Leu Tyr Asp
1745	1750	1755
Lys Thr His Thr Cys Pro	Pro Cys Pro Ala Pro	Glu Leu Leu Gly
1760	1765	1770
Gly Pro Ser Val Phe Leu	Phe Pro Pro Lys Pro	Lys Asp Thr Leu
1775	1780	1785
Met Ile Ser Arg Thr Pro	Glu Val Thr Cys Val	Val Val Asp Val
1790	1795	1800
Ser His Glu Asp Pro Glu	Val Lys Phe Asn Trp	Tyr Val Asp Gly
1805	1810	1815
Val Glu Val His Asn Ala	Lys Thr Lys Pro Arg	Glu Glu Gln Tyr
1820	1825	1830
Asn Ser Thr Tyr Arg Val	Val Ser Val Leu Thr	Val Leu His Gln
1835	1840	1845
Asp Trp Leu Asn Gly Lys	Glu Tyr Lys Cys Lys	Val Ser Asn Lys
1850	1855	1860
Ala Leu Pro Ala Pro Ile	Glu Lys Thr Ile Ser	Lys Ala Lys Gly
1865	1870	1875
Gln Pro Arg Glu Pro Gln	Val Tyr Thr Leu Pro	Pro Ser Arg Asp
1880	1885	1890
Glu Leu Thr Lys Asn Gln	Val Ser Leu Thr Cys	Leu Val Lys Gly
1895	1900	1905
Phe Tyr Pro Ser Asp Ile	Ala Val Glu Trp Glu	Ser Asn Gly Gln
1910	1915	1920

Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
 1925 1930 1935
 Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 1940 1945 1950
 Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
 1955 1960 1965
 Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly
 1970 1975 1980

Lys

<210> 106

<211> 2134

<212> PRT

<213> 人工序列

<220>

<223> FVIII 196蛋白序列

<400> 106

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

180	185	190
Ala Ser Ser Asp Ala Arg Phe Pro	Pro Arg Val Pro Lys Ser Phe Pro	
195	200	205
Phe Asn Thr Ser Val Val Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr		
210	215	220
Asp His Leu Phe Asn Ile Ala Lys Pro Arg Pro Pro Trp Met Gly Leu		
225	230	235
Leu Gly Pro Thr Ile Gln Ala Glu Val Tyr Asp Thr Val Val Ile Thr		
245	250	255
Leu Lys Asn Met Ala Ser His Pro Val Ser Leu His Ala Val Gly Val		
260	265	270
Ser Tyr Trp Lys Ala Ser Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser		
275	280	285
Gln Arg Glu Lys Glu Asp Asp Lys Val Phe Pro Gly Gly Ser His Thr		
290	295	300
Tyr Val Trp Gln Val Leu Lys Glu Asn Gly Pro Met Ala Ser Asp Pro		
305	310	315
Leu Cys Leu Thr Tyr Ser Tyr Leu Ser His Val Asp Leu Val Lys Asp		
325	330	335
Leu Asn Ser Gly Leu Ile Gly Ala Leu Leu Val Cys Arg Glu Gly Ser		
340	345	350
Leu Ala Lys Glu Lys Thr Gln Thr Leu His Lys Phe Ile Leu Leu Phe		
355	360	365
Ala Val Phe Asp Glu Gly Lys Ser Trp His Ser Glu Thr Lys Asn Ser		
370	375	380
Leu Met Gln Asp Arg Asp Ala Ala Ser Ala Arg Ala Trp Pro Lys Met		
385	390	395
His Thr Val Asn Gly Tyr Val Asn Arg Ser Leu Pro Gly Leu Ile Gly		
405	410	415
Cys His Arg Lys Ser Val Tyr Trp His Val Ile Gly Met Gly Thr Thr		
420	425	430
Pro Glu Val His Ser Ile Phe Leu Glu Gly His Thr Phe Leu Val Arg		
435	440	445
Asn His Arg Gln Ala Ser Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr		
450	455	460
Ala Gln Thr Leu Leu Met Asp Leu Gly Gln Phe Leu Leu Phe Cys His		
465	470	475
Ile Ser Ser His Gln His Asp Gly Met Glu Ala Tyr Val Lys Val Asp		
485	490	495

Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn Asn Glu Glu Ala		
500	505	510
Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met Asp Val Val Arg		
515	520	525
Phe Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile Arg Ser Val Ala		
530	535	540
Lys Lys His Pro Lys Thr Trp Val His Tyr Ile Ala Ala Glu Glu Glu		
545	550	555
Asp Trp Asp Tyr Ala Pro Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr		
565	570	575
Lys Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr		
580	585	590
Lys Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg		
595	600	605
Glu Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly		
610	615	620
Glu Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg		
625	630	635
Pro Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr		
645	650	655
Ser Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile		
660	665	670
Leu Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp		
675	680	685
Gly Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser		
690	695	700
Phe Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu		
705	710	715
Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met		
725	730	735
Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg		
740	745	750
Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala		
755	760	765
Gly Val Gln Leu Glu Asp Pro Glu Phe Gln Ala Ser Asn Ile Met His		
770	775	780
Ser Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu		
785	790	795
His Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp		

				805					810					815			
Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val		
				820					825					830			
Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe		
				835					840					845			
Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser		
				850					855					860			
Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys		
865									870					875			880
Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser		
				885					890					895			
Ala	Tyr	Leu	Leu	Ser	Lys	Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser		
				900					905					910			
Gln	Asn	Pro	Pro	Val	Leu	Lys	Arg	His	Gln	Arg	Glu	Ile	Thr	Arg	Thr		
				915					920					925			
Thr	Leu	Gln	Gly	Ala	Pro	Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser		
				930					935					940			
Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala		
945									950					955			960
Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr		
				965					970					975			
Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr		
				980					985					990			
Gly	Pro	Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ala		
				995					1000					1005			
Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly			
				1010					1015					1020			
Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser			
				1025					1030					1035			
Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser			
				1040					1045					1050			
Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ala			
				1055					1060					1065			
Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Ala	Ser	Ser	Ser	Asp			
				1070					1075					1080			
Gln	Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys			
				1085					1090					1095			
Lys	Glu	Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro			
				1100					1105					1110			

Arg Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val		
1115	1120	1125
Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu		
1130	1135	1140
Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val		
1145	1150	1155
Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr		
1160	1165	1170
Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr Ile		
1175	1180	1185
Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln		
1190	1195	1200
Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu		
1205	1210	1215
Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys		
1220	1225	1230
Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met		
1235	1240	1245
Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe		
1250	1255	1260
Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly		
1265	1270	1275
Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly		
1280	1285	1290
Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe		
1295	1300	1305
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn		
1310	1315	1320
Cys Arg Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		
1325	1330	1335
Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr		
1340	1345	1350
Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
1355	1360	1365
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro		
1370	1375	1380
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala		
1385	1390	1395
Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser		

1400	1405	1410
Pro Ala Gly Ser Pro Thr	Ser Thr Glu Glu Gly	Ser Pro Ala Gly
1415	1420	1425
Ser Pro Thr Ser Thr Glu	Glu Gly Ser Pro Ala	Gly Ser Pro Thr
1430	1435	1440
Ser Thr Glu Glu Gly Thr	Ser Glu Ser Ala Thr	Pro Glu Ser Gly
1445	1450	1455
Pro Gly Thr Ser Thr Glu	Pro Ser Glu Gly Ser	Ala Pro Gly Ala
1460	1465	1470
Ser Ser Ala Pro Cys Asn	Ile Gln Met Glu Asp	Pro Thr Phe Lys
1475	1480	1485
Glu Asn Tyr Arg Phe His	Ala Ile Asn Gly Tyr	Ile Met Asp Thr
1490	1495	1500
Leu Pro Gly Leu Val Met	Ala Gln Asp Gln Arg	Ile Arg Trp Tyr
1505	1510	1515
Leu Leu Ser Met Gly Ser	Asn Glu Asn Ile His	Ser Ile His Phe
1520	1525	1530
Ser Gly His Val Phe Thr	Val Arg Lys Lys Glu	Glu Tyr Lys Met
1535	1540	1545
Ala Leu Tyr Asn Leu Tyr	Pro Gly Val Phe Glu	Thr Val Glu Met
1550	1555	1560
Leu Pro Ser Lys Ala Gly	Ile Trp Arg Val Glu	Cys Leu Ile Gly
1565	1570	1575
Glu His Leu His Ala Gly	Met Ser Thr Leu Phe	Leu Val Tyr Ser
1580	1585	1590
Asn Lys Cys Gln Thr Pro	Leu Gly Met Ala Ser	Gly His Ile Arg
1595	1600	1605
Asp Phe Gln Ile Thr Ala	Ser Gly Gln Tyr Gly	Gln Trp Ala Pro
1610	1615	1620
Lys Leu Ala Arg Leu His	Tyr Ser Gly Ser Ile	Asn Ala Trp Ser
1625	1630	1635
Thr Lys Glu Pro Phe Ser	Trp Ile Lys Val Asp	Leu Leu Ala Pro
1640	1645	1650
Met Ile Ile His Gly Ile	Lys Thr Gln Gly Ala	Arg Gln Lys Phe
1655	1660	1665
Ser Ser Leu Tyr Ile Ser	Gln Phe Ile Ile Met	Tyr Ser Leu Asp
1670	1675	1680
Gly Lys Lys Trp Gln Thr	Tyr Arg Gly Asn Ser	Thr Gly Thr Leu
1685	1690	1695

Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1700	1705	1710
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1715	1720	1725
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1730	1735	1740
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1745	1750	1755
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1760	1765	1770
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1775	1780	1785
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1790	1795	1800
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1805	1810	1815
Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys		
1820	1825	1830
Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu		
1835	1840	1845
Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp		
1850	1855	1860
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr		
1865	1870	1875
Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala		
1880	1885	1890
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr Asp		
1895	1900	1905
Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly		
1910	1915	1920
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu		
1925	1930	1935
Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val		
1940	1945	1950
Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly		
1955	1960	1965
Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr		
1970	1975	1980
Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln		

1985	1990	1995
Asp Trp Leu Asn Gly Lys	Glu Tyr Lys Cys Lys	Val Ser Asn Lys
2000	2005	2010
Ala Leu Pro Ala Pro Ile	Glu Lys Thr Ile Ser	Lys Ala Lys Gly
2015	2020	2025
Gln Pro Arg Glu Pro Gln	Val Tyr Thr Leu Pro	Pro Ser Arg Asp
2030	2035	2040
Glu Leu Thr Lys Asn Gln	Val Ser Leu Thr Cys	Leu Val Lys Gly
2045	2050	2055
Phe Tyr Pro Ser Asp Ile	Ala Val Glu Trp Glu	Ser Asn Gly Gln
2060	2065	2070
Pro Glu Asn Asn Tyr Lys	Thr Thr Pro Pro Val	Leu Asp Ser Asp
2075	2080	2085
Gly Ser Phe Phe Leu Tyr	Ser Lys Leu Thr Val	Asp Lys Ser Arg
2090	2095	2100
Trp Gln Gln Gly Asn Val	Phe Ser Cys Ser Val	Met His Glu Ala
2105	2110	2115
Leu His Asn His Tyr Thr	Gln Lys Ser Leu Ser	Leu Ser Pro Gly
2120	2125	2130

Lys

<210> 107

<211> 1984

<212> PRT

<213> 人工序列

<220>

<223> FVIII 199蛋白序列

<400> 107

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg		
35	40	45
Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val		
50	55	60
Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile		
65	70	75
Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln		
85	90	95

Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro

				405					410					415			
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn		
				420					425					430			
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met		
				435					440					445			
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu		
				450					455					460			
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu		
465									470					475			480
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro		
				485					490					495			
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys		
				500					505					510			
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe		
				515					520					525			
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp		
				530					535					540			
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg		
545									550					555			560
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu		
				565					570					575			
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val		
				580					585					590			
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu		
				595					600					605			
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp		
				610					615					620			
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val		
625									630					635			640
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp		
				645					650					655			
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe		
				660					665					670			
Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr		
				675					680					685			
Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro		
				690					695					700			
Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly		
705									710					715			720

Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp
				725				730				735			
Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys
				740				745				750			
Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Pro	Pro	Val	Leu
				755				760				765			
Lys	Arg	His	Gln	Ala	Glu	Ile	Thr	Arg	Thr	Thr	Leu	Gln	Gly	Ala	Pro
				770				775				780			
Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ala	Ser	Pro
785				790				795				800			
Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser
				805				810				815			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro
				820				825				830			
Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Thr	Pro	Gly
				835				840				845			
Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser
				850				855				860			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro
865				870				875				880			
Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr
				885				890				895			
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala
				900				905				910			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro
				915				920				925			
Ala	Ser	Ser	Ser	Asp	Gln	Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser
				930				935				940			
Val	Glu	Met	Lys	Lys	Glu	Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn
945				950				955				960			
Gln	Ser	Pro	Arg	Ser	Phe	Gln	Lys	Lys	Thr	Arg	His	Tyr	Phe	Ile	Ala
				965				970				975			
Ala	Val	Glu	Arg	Leu	Trp	Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	His	Val
				980				985				990			
Leu	Arg	Asn	Arg	Ala	Gln	Ser	Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val
				995				1000				1005			
Val	Phe	Gln	Glu	Phe	Thr	Asp	Gly	Ser	Phe	Thr	Gln	Pro	Leu	Tyr	
				1010				1015				1020			
Arg	Gly	Glu	Leu	Asn	Glu	His	Leu	Gly	Leu	Leu	Gly	Pro	Tyr	Ile	

1025	1030	1035
Arg Ala Glu Val Glu Asp	Asn Ile Met Val Thr	Phe Arg Asn Gln
1040	1045	1050
Ala Ser Arg Pro Tyr Ser	Phe Tyr Ser Ser Leu	Ile Ser Tyr Glu
1055	1060	1065
Glu Asp Gln Arg Gln Gly	Ala Glu Pro Arg Lys	Asn Phe Val Lys
1070	1075	1080
Pro Asn Glu Thr Lys Thr	Tyr Phe Trp Lys Val	Gln His His Met
1085	1090	1095
Ala Pro Thr Lys Asp Glu	Phe Asp Cys Lys Ala	Trp Ala Tyr Phe
1100	1105	1110
Ser Asp Val Asp Leu Glu	Lys Asp Val His Ser	Gly Leu Ile Gly
1115	1120	1125
Pro Leu Leu Val Cys His	Thr Asn Thr Leu Asn	Pro Ala His Gly
1130	1135	1140
Arg Gln Val Thr Val Gln	Glu Phe Ala Leu Phe	Phe Thr Ile Phe
1145	1150	1155
Asp Glu Thr Lys Ser Trp	Tyr Phe Thr Glu Asn	Met Glu Arg Asn
1160	1165	1170
Cys Arg Gly Ala Pro Thr	Ser Glu Ser Ala Thr	Pro Glu Ser Gly
1175	1180	1185
Pro Gly Ser Glu Pro Ala	Thr Ser Gly Ser Glu	Thr Pro Gly Thr
1190	1195	1200
Ser Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly	Ser Glu Pro Ala
1205	1210	1215
Thr Ser Gly Ser Glu Thr	Pro Gly Thr Ser Glu	Ser Ala Thr Pro
1220	1225	1230
Glu Ser Gly Pro Gly Thr	Ser Thr Glu Pro Ser	Glu Gly Ser Ala
1235	1240	1245
Pro Gly Thr Ser Glu Ser	Ala Thr Pro Glu Ser	Gly Pro Gly Ser
1250	1255	1260
Pro Ala Gly Ser Pro Thr	Ser Thr Glu Glu Gly	Ser Pro Ala Gly
1265	1270	1275
Ser Pro Thr Ser Thr Glu	Glu Gly Ser Pro Ala	Gly Ser Pro Thr
1280	1285	1290
Ser Thr Glu Glu Gly Thr	Ser Glu Ser Ala Thr	Pro Glu Ser Gly
1295	1300	1305
Pro Gly Thr Ser Thr Glu	Pro Ser Glu Gly Ser	Ala Pro Gly Ala
1310	1315	1320

Ser Ser Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys		
1325	1330	1335
Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr		
1340	1345	1350
Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr		
1355	1360	1365
Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe		
1370	1375	1380
Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met		
1385	1390	1395
Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met		
1400	1405	1410
Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly		
1415	1420	1425
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1430	1435	1440
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1445	1450	1455
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1460	1465	1470
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1475	1480	1485
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1490	1495	1500
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1505	1510	1515
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1520	1525	1530
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1535	1540	1545
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1550	1555	1560
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1565	1570	1575
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1580	1585	1590
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1595	1600	1605
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		

1610	1615	1620
Met Phe Ala Thr Trp Ser	Pro Ser Lys Ala Arg	Leu His Leu Gln
1625	1630	1635
Gly Arg Ser Asn Ala Trp	Arg Pro Gln Val Asn	Asn Pro Lys Glu
1640	1645	1650
Trp Leu Gln Val Asp Phe	Gln Lys Thr Met Lys	Val Thr Gly Val
1655	1660	1665
Thr Thr Gln Gly Val Lys	Ser Leu Leu Thr Ser	Met Tyr Val Lys
1670	1675	1680
Glu Phe Leu Ile Ser Ser	Ser Gln Asp Gly His	Gln Trp Thr Leu
1685	1690	1695
Phe Phe Gln Asn Gly Lys	Val Lys Val Phe Gln	Gly Asn Gln Asp
1700	1705	1710
Ser Phe Thr Pro Val Val	Asn Ser Leu Asp Pro	Pro Leu Leu Thr
1715	1720	1725
Arg Tyr Leu Arg Ile His	Pro Gln Ser Trp Val	His Gln Ile Ala
1730	1735	1740
Leu Arg Met Glu Val Leu	Gly Cys Glu Ala Gln	Asp Leu Tyr Asp
1745	1750	1755
Lys Thr His Thr Cys Pro	Pro Cys Pro Ala Pro	Glu Leu Leu Gly
1760	1765	1770
Gly Pro Ser Val Phe Leu	Phe Pro Pro Lys Pro	Lys Asp Thr Leu
1775	1780	1785
Met Ile Ser Arg Thr Pro	Glu Val Thr Cys Val	Val Val Asp Val
1790	1795	1800
Ser His Glu Asp Pro Glu	Val Lys Phe Asn Trp	Tyr Val Asp Gly
1805	1810	1815
Val Glu Val His Asn Ala	Lys Thr Lys Pro Arg	Glu Glu Gln Tyr
1820	1825	1830
Asn Ser Thr Tyr Arg Val	Val Ser Val Leu Thr	Val Leu His Gln
1835	1840	1845
Asp Trp Leu Asn Gly Lys	Glu Tyr Lys Cys Lys	Val Ser Asn Lys
1850	1855	1860
Ala Leu Pro Ala Pro Ile	Glu Lys Thr Ile Ser	Lys Ala Lys Gly
1865	1870	1875
Gln Pro Arg Glu Pro Gln	Val Tyr Thr Leu Pro	Pro Ser Arg Asp
1880	1885	1890
Glu Leu Thr Lys Asn Gln	Val Ser Leu Thr Cys	Leu Val Lys Gly
1895	1900	1905

Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln
1910 1915 1920
Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
1925 1930 1935
Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
1940 1945 1950
Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
1955 1960 1965
Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly
1970 1975 1980

Lys

<210> 108

<211> 2134

<212> PRT

<213> 人工序列

<220>

<223> FVIII 201蛋白序列

<400> 108

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

				165					170					175			
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
				180					185					190			
Ala	Ser	Ser	Asp	Ala	Arg	Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro		
				195					200					205			
Phe	Asn	Thr	Ser	Val	Val	Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr		
				210					215					220			
Asp	His	Leu	Phe	Asn	Ile	Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu		
225					230					235					240		
Leu	Gly	Pro	Thr	Ile	Gln	Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr		
				245						250					255		
Leu	Lys	Asn	Met	Ala	Ser	His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val		
				260						265					270		
Ser	Tyr	Trp	Lys	Ala	Ser	Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser		
				275						280					285		
Gln	Arg	Glu	Lys	Glu	Asp	Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr		
				290						295					300		
Tyr	Val	Trp	Gln	Val	Leu	Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro		
305					310					315					320		
Leu	Cys	Leu	Thr	Tyr	Ser	Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp		
				325						330					335		
Leu	Asn	Ser	Gly	Leu	Ile	Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser		
				340						345					350		
Leu	Ala	Lys	Glu	Lys	Thr	Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe		
				355						360					365		
Ala	Val	Phe	Asp	Glu	Gly	Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser		
				370						375					380		
Leu	Met	Gln	Asp	Arg	Asp	Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met		
385					390					395					400		
His	Thr	Val	Asn	Gly	Tyr	Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly		
				405						410					415		
Cys	His	Arg	Lys	Ser	Val	Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr		
				420						425					430		
Pro	Glu	Val	His	Ser	Ile	Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg		
				435						440					445		
Asn	His	Arg	Gln	Ala	Ser	Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr		
				450						455					460		
Ala	Gln	Thr	Leu	Leu	Met	Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His		
465					470					475					480		

Ile Ser Ser His Gln His Asp Gly Met Glu Ala Tyr Val Lys Val Asp		
485	490	495
Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn Asn Glu Glu Ala		
500	505	510
Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met Asp Val Val Arg		
515	520	525
Phe Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile Arg Ser Val Ala		
530	535	540
Lys Lys His Pro Lys Thr Trp Val His Tyr Ile Ala Ala Glu Glu Glu		
545	550	555
Asp Trp Asp Tyr Ala Pro Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr		
565	570	575
Lys Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr		
580	585	590
Lys Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg		
595	600	605
Glu Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly		
610	615	620
Glu Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg		
625	630	635
Pro Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr		
645	650	655
Ser Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile		
660	665	670
Leu Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp		
675	680	685
Gly Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser		
690	695	700
Phe Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu		
705	710	715
Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met		
725	730	735
Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg		
740	745	750
Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala		
755	760	765
Gly Val Gln Leu Glu Asp Pro Glu Phe Gln Ala Ser Asn Ile Met His		
770	775	780
Ser Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu		

785	790	795	800
His Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp			
	805	810	815
Phe Leu Ser Val Phe Phe Ser Gly Tyr Thr Phe Lys His Lys Met Val			
	820	825	830
Tyr Glu Asp Thr Leu Thr Leu Phe Pro Phe Ser Gly Glu Thr Val Phe			
	835	840	845
Met Ser Met Glu Asn Pro Gly Leu Trp Ile Leu Gly Cys His Asn Ser			
	850	855	860
Asp Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys			
865	870	875	880
Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser			
	885	890	895
Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser			
	900	905	910
Gln Asn Pro Pro Val Leu Lys Arg His Gln Ala Glu Ile Thr Arg Thr			
	915	920	925
Thr Leu Gln Gly Ala Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser			
	930	935	940
Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala			
945	950	955	960
Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr			
	965	970	975
Ser Ser Thr Gly Ser Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr			
	980	985	990
Gly Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala			
	995	1000	1005
Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly			
	1010	1015	1020
Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser			
	1025	1030	1035
Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser			
	1040	1045	1050
Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala			
	1055	1060	1065
Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Ala Ser Ser Ser Asp			
	1070	1075	1080
Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys			
	1085	1090	1095

Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro		
1100	1105	1110
Arg Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val		
1115	1120	1125
Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu		
1130	1135	1140
Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val		
1145	1150	1155
Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr		
1160	1165	1170
Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr Ile		
1175	1180	1185
Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln		
1190	1195	1200
Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu		
1205	1210	1215
Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys		
1220	1225	1230
Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met		
1235	1240	1245
Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe		
1250	1255	1260
Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly		
1265	1270	1275
Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly		
1280	1285	1290
Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe		
1295	1300	1305
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn		
1310	1315	1320
Cys Arg Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		
1325	1330	1335
Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr		
1340	1345	1350
Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
1355	1360	1365
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro		
1370	1375	1380
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala		

1385	1390	1395
Pro Gly Thr Ser Glu Ser	Ala Thr Pro Glu Ser	Gly Pro Gly Ser
1400	1405	1410
Pro Ala Gly Ser Pro Thr	Ser Thr Glu Glu Gly	Ser Pro Ala Gly
1415	1420	1425
Ser Pro Thr Ser Thr Glu	Glu Gly Ser Pro Ala	Gly Ser Pro Thr
1430	1435	1440
Ser Thr Glu Glu Gly Thr	Ser Glu Ser Ala Thr	Pro Glu Ser Gly
1445	1450	1455
Pro Gly Thr Ser Thr Glu	Pro Ser Glu Gly Ser	Ala Pro Gly Ala
1460	1465	1470
Ser Ser Ala Pro Cys Asn	Ile Gln Met Glu Asp	Pro Thr Phe Lys
1475	1480	1485
Glu Asn Tyr Arg Phe His	Ala Ile Asn Gly Tyr	Ile Met Asp Thr
1490	1495	1500
Leu Pro Gly Leu Val Met	Ala Gln Asp Gln Arg	Ile Arg Trp Tyr
1505	1510	1515
Leu Leu Ser Met Gly Ser	Asn Glu Asn Ile His	Ser Ile His Phe
1520	1525	1530
Ser Gly His Val Phe Thr	Val Arg Lys Lys Glu	Glu Tyr Lys Met
1535	1540	1545
Ala Leu Tyr Asn Leu Tyr	Pro Gly Val Phe Glu	Thr Val Glu Met
1550	1555	1560
Leu Pro Ser Lys Ala Gly	Ile Trp Arg Val Glu	Cys Leu Ile Gly
1565	1570	1575
Glu His Leu His Ala Gly	Met Ser Thr Leu Phe	Leu Val Tyr Ser
1580	1585	1590
Asn Lys Cys Gln Thr Pro	Leu Gly Met Ala Ser	Gly His Ile Arg
1595	1600	1605
Asp Phe Gln Ile Thr Ala	Ser Gly Gln Tyr Gly	Gln Trp Ala Pro
1610	1615	1620
Lys Leu Ala Arg Leu His	Tyr Ser Gly Ser Ile	Asn Ala Trp Ser
1625	1630	1635
Thr Lys Glu Pro Phe Ser	Trp Ile Lys Val Asp	Leu Leu Ala Pro
1640	1645	1650
Met Ile Ile His Gly Ile	Lys Thr Gln Gly Ala	Arg Gln Lys Phe
1655	1660	1665
Ser Ser Leu Tyr Ile Ser	Gln Phe Ile Ile Met	Tyr Ser Leu Asp
1670	1675	1680

Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1685	1690	1695
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1700	1705	1710
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1715	1720	1725
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1730	1735	1740
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1745	1750	1755
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1760	1765	1770
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1775	1780	1785
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1790	1795	1800
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1805	1810	1815
Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys		
1820	1825	1830
Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu		
1835	1840	1845
Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp		
1850	1855	1860
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr		
1865	1870	1875
Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala		
1880	1885	1890
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr Asp		
1895	1900	1905
Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly		
1910	1915	1920
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu		
1925	1930	1935
Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val		
1940	1945	1950
Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly		
1955	1960	1965
Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr		

1970	1975	1980
Asn Ser Thr Tyr Arg Val	Val Ser Val Leu Thr	Val Leu His Gln
1985	1990	1995
Asp Trp Leu Asn Gly Lys	Glu Tyr Lys Cys Lys	Val Ser Asn Lys
2000	2005	2010
Ala Leu Pro Ala Pro Ile	Glu Lys Thr Ile Ser	Lys Ala Lys Gly
2015	2020	2025
Gln Pro Arg Glu Pro Gln	Val Tyr Thr Leu Pro	Pro Ser Arg Asp
2030	2035	2040
Glu Leu Thr Lys Asn Gln	Val Ser Leu Thr Cys	Leu Val Lys Gly
2045	2050	2055
Phe Tyr Pro Ser Asp Ile	Ala Val Glu Trp Glu	Ser Asn Gly Gln
2060	2065	2070
Pro Glu Asn Asn Tyr Lys	Thr Thr Pro Pro Val	Leu Asp Ser Asp
2075	2080	2085
Gly Ser Phe Phe Leu Tyr	Ser Lys Leu Thr Val	Asp Lys Ser Arg
2090	2095	2100
Trp Gln Gln Gly Asn Val	Phe Ser Cys Ser Val	Met His Glu Ala
2105	2110	2115
Leu His Asn His Tyr Thr	Gln Lys Ser Leu Ser	Leu Ser Pro Gly
2120	2125	2130

Lys

<210> 109

<211> 2128

<212> PRT

<213> 人工序列

<220>

<223> FVIII 203蛋白序列

<400> 109

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg		
35	40	45
Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val		
50	55	60
Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile		
65	70	75
		80

Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln
85				90				95							
Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
100				105				110							
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
115				120				125							
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
130				135				140							
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
145				150				155				160			
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
165				170				175							
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
180				185				190							
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
195				200				205							
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
210				215				220							
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
225				230				235				240			
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
245				250				255							
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
260				265				270							
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
275				280				285							
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
290				295				300							
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met
305				310				315				320			
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His
325				330				335							
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro
340				345				350							
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp
355				360				365							
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser
370				375				380							
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr

385	390	395	400
Trp Val His Tyr Ile	Ala Ala Glu Glu Glu	Asp Trp Asp Tyr Ala	Pro
	405	410	415
Leu Val Leu Ala Pro	Asp Asp Arg Ser Tyr Lys	Ser Gln Tyr Leu	Asn
	420	425	430
Asn Gly Pro Gln Arg	Ile Gly Arg Lys Tyr Lys	Lys Val Arg Phe	Met
	435	440	445
Ala Tyr Thr Asp Glu	Thr Phe Lys Thr Arg	Glu Ala Ile Gln His	Glu
	450	455	460
Ser Gly Ile Leu Gly	Pro Leu Leu Tyr Gly	Glu Val Gly Asp Thr	Leu
465	470	475	480
Leu Ile Ile Phe Lys	Asn Gln Ala Ser Arg	Pro Tyr Asn Ile Tyr	Pro
	485	490	495
His Gly Ile Thr Asp	Val Arg Pro Leu Tyr	Ser Arg Arg Leu Pro	Lys
	500	505	510
Gly Val Lys His Leu	Lys Asp Phe Pro Ile	Leu Pro Gly Glu Ile	Phe
	515	520	525
Lys Tyr Lys Trp Thr	Val Thr Val Glu Asp	Gly Pro Thr Lys Ser	Asp
	530	535	540
Pro Arg Cys Leu Thr	Arg Tyr Tyr Ser Ser	Phe Val Asn Met Glu	Arg
545	550	555	560
Asp Leu Ala Ser Gly	Leu Ile Gly Pro Leu	Leu Ile Cys Tyr Lys	Glu
	565	570	575
Ser Val Asp Gln Arg	Gly Asn Gln Ile Met	Ser Asp Lys Arg Asn	Val
	580	585	590
Ile Leu Phe Ser Val	Phe Asp Glu Asn Arg	Ser Trp Tyr Leu Thr	Glu
	595	600	605
Asn Ile Gln Arg Phe	Leu Pro Asn Pro Ala	Gly Val Gln Leu Glu	Asp
	610	615	620
Pro Glu Phe Gln Ala	Ser Asn Ile Met His	Ser Ile Asn Gly Tyr	Val
625	630	635	640
Phe Asp Ser Leu Gln	Leu Ser Val Cys Leu	His Glu Val Ala Tyr	Trp
	645	650	655
Tyr Ile Leu Ser Ile	Gly Ala Gln Thr Asp	Phe Leu Ser Val Phe	Phe
	660	665	670
Ser Gly Tyr Thr Phe	Lys His Lys Met Val	Tyr Glu Asp Thr Leu	Thr
	675	680	685
Leu Phe Pro Phe Ser	Gly Glu Thr Val Phe	Met Ser Met Glu Asn	Pro
	690	695	700

Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly
705					710					715					720
Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp
					725					730					735
Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys
					740					745					750
Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Gly	Ala	Pro	Gly
					755					760					765
Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala
					770					775					780
Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu
785					790					795					800
Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly
					805					810					815
Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Thr	Glu
					820					825					830
Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser
					835					840					845
Thr	Glu	Glu	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly
					850					855					860
Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Thr	Ser	Glu	Ser
865					870					875					880
Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser
					885					890					895
Thr	Glu	Glu	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser	Thr	Glu	Glu	Gly
					900					905					910
Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Thr	Ser	Glu	Ser
					915					920					925
Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu
					930					935					940
Ser	Gly	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly
945					950					955					960
Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Ser	Glu	Pro	Ala
					965					970					975
Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser
					980					985					990
Thr	Glu	Glu	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly
					995					1000					1005
Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Ser	Glu	Pro	

1010	1015	1020
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr		
1025	1030	1035
Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser		
1040	1045	1050
Ala Pro Ala Ser Ser Pro Pro Val Leu Lys Arg His Gln Ala Glu		
1055	1060	1065
Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr		
1070	1075	1080
Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile		
1085	1090	1095
Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys		
1100	1105	1110
Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr		
1115	1120	1125
Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser		
1130	1135	1140
Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr		
1145	1150	1155
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu		
1160	1165	1170
His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp		
1175	1180	1185
Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser		
1190	1195	1200
Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly		
1205	1210	1215
Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr		
1220	1225	1230
Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu		
1235	1240	1245
Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu		
1250	1255	1260
Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His		
1265	1270	1275
Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln		
1280	1285	1290
Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp		
1295	1300	1305

Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Gly Ala Pro Thr		
1310	1315	1320
Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
1325	1330	1335
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro		
1340	1345	1350
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr		
1355	1360	1365
Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr		
1370	1375	1380
Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser		
1385	1390	1395
Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr		
1400	1405	1410
Ser Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu		
1415	1420	1425
Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr		
1430	1435	1440
Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu		
1445	1450	1455
Pro Ser Glu Gly Ser Ala Pro Gly Ala Ser Ser Ala Pro Cys Asn		
1460	1465	1470
Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His		
1475	1480	1485
Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met		
1490	1495	1500
Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser		
1505	1510	1515
Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr		
1520	1525	1530
Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr		
1535	1540	1545
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly		
1550	1555	1560
Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly		
1565	1570	1575
Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro		
1580	1585	1590
Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala		

1595	1600	1605
Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu	Ala Arg Leu His	
1610	1615	1620
Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys	Glu Pro Phe Ser	
1625	1630	1635
Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile	Ile His Gly Ile	
1640	1645	1650
Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser	Leu Tyr Ile Ser	
1655	1660	1665
Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys	Lys Trp Gln Thr	
1670	1675	1680
Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val	Phe Phe Gly Asn	
1685	1690	1695
Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe	Asn Pro Pro Ile	
1700	1705	1710
Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His	Tyr Ser Ile Arg	
1715	1720	1725
Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp	Leu Asn Ser Cys	
1730	1735	1740
Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile	Ser Asp Ala Gln	
1745	1750	1755
Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe	Ala Thr Trp Ser	
1760	1765	1770
Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg	Ser Asn Ala Trp	
1775	1780	1785
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu	Gln Val Asp Phe	
1790	1795	1800
Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr	Gln Gly Val Lys	
1805	1810	1815
Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe	Leu Ile Ser Ser	
1820	1825	1830
Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe	Gln Asn Gly Lys	
1835	1840	1845
Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe	Thr Pro Val Val	
1850	1855	1860
Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr	Leu Arg Ile His	
1865	1870	1875
Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg	Met Glu Val Leu	
1880	1885	1890

Gly Cys Glu Ala Gln Asp Leu Tyr Asp Lys Thr His Thr Cys Pro		
1895	1900	1905
Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu		
1910	1915	1920
Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro		
1925	1930	1935
Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu		
1940	1945	1950
Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala		
1955	1960	1965
Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val		
1970	1975	1980
Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys		
1985	1990	1995
Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile		
2000	2005	2010
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln		
2015	2020	2025
Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln		
2030	2035	2040
Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile		
2045	2050	2055
Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys		
2060	2065	2070
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr		
2075	2080	2085
Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val		
2090	2095	2100
Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr		
2105	2110	2115
Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		
2120	2125	
<210> 110		
<211> 2128		
<212> PRT		
<213> 人工序列		
<220>		
<223> FVIII 204蛋白序列		
<400> 110		

Met	Gln	Ile	Glu	Leu	Ser	Thr	Cys	Phe	Phe	Leu	Cys	Leu	Leu	Arg	Phe
1			5					10						15	
Cys	Phe	Ser	Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser
			20					25					30		
Trp	Asp	Tyr	Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg
			35					40					45		
Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val
			50				55				60				
Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile
65				70						75				80	
Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln
				85					90					95	
Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
			100					105					110		
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
			115					120					125		
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
			130				135				140				
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
145				150						155				160	
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
				165					170					175	
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
			180					185					190		
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
			195				200					205			
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
			210				215				220				
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
225				230						235				240	
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
				245					250					255	
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
			260				265						270		
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
			275				280						285		
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
			290				295				300				
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met

305	310	315	320
Asp Leu Gly Gln Phe	Leu Leu Phe Cys His	Ile Ser Ser His Gln His	
325	330	335	
Asp Gly Met Glu Ala Tyr Val Lys Val	Asp Ser Cys Pro Glu Glu Pro		
340	345	350	
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp			
355	360	365	
Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser			
370	375	380	
Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr			
385	390	395	400
Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro			
405	410	415	
Leu Val Leu Ala Pro Asp Gly Ala Pro Thr Ser Thr Glu Pro Ser Glu			
420	425	430	
Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu			
435	440	445	
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr			
450	455	460	
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro			
465	470	475	480
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
485	490	495	
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro			
500	505	510	
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Thr Glu Pro Ser Glu			
515	520	525	
Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro			
530	535	540	
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Glu			
545	550	555	560
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ala Ser Ser Asp Arg Ser Tyr			
565	570	575	
Lys Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr			
580	585	590	
Lys Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg			
595	600	605	
Glu Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly			
610	615	620	

Glu Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg			
625	630	635	640
Pro Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr			
	645	650	655
Ser Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile			
	660	665	670
Leu Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp			
	675	680	685
Gly Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser			
	690	695	700
Phe Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu			
	705	710	715
Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met			
	725	730	735
Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg			
	740	745	750
Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala			
	755	760	765
Gly Val Gln Leu Glu Asp Pro Glu Phe Gln Ala Ser Asn Ile Met His			
	770	775	780
Ser Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu			
	785	790	795
His Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp			
	805	810	815
Phe Leu Ser Val Phe Phe Ser Gly Tyr Thr Phe Lys His Lys Met Val			
	820	825	830
Tyr Glu Asp Thr Leu Thr Leu Phe Pro Phe Ser Gly Glu Thr Val Phe			
	835	840	845
Met Ser Met Glu Asn Pro Gly Leu Trp Ile Leu Gly Cys His Asn Ser			
	850	855	860
Asp Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys			
	865	870	875
Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser			
	885	890	895
Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser			
	900	905	910
Gln Asn Gly Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly			
	915	920	925
Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser			

930	935	940
Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly	Ser Glu Pro Ala Thr Ser
945	950	955
Gly Ser Glu Thr Pro	Gly Thr Ser Glu Ser	Ala Thr Pro Glu Ser Gly
965	970	975
Pro Gly Thr Ser Thr	Glu Pro Ser Glu Gly	Ser Ala Pro Gly Ser Pro
980	985	990
Ala Gly Ser Pro Thr	Ser Thr Glu Glu Gly	Thr Ser Glu Ser Ala Thr
995	1000	1005
Pro Glu Ser Gly Pro	Gly Ser Glu Pro Ala	Thr Ser Gly Ser Glu
1010	1015	1020
Thr Pro Gly Thr Ser	Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly
1025	1030	1035
Ser Pro Ala Gly Ser	Pro Thr Ser Thr Glu	Glu Gly Ser Pro Ala
1040	1045	1050
Gly Ser Pro Thr Ser	Thr Glu Glu Gly Thr	Ser Thr Glu Pro Ser
1055	1060	1065
Glu Gly Ser Ala Pro	Gly Thr Ser Glu Ser	Ala Thr Pro Glu Ser
1070	1075	1080
Gly Pro Gly Thr Ser	Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly
1085	1090	1095
Thr Ser Glu Ser Ala	Thr Pro Glu Ser Gly	Pro Gly Ser Glu Pro
1100	1105	1110
Ala Thr Ser Gly Ser	Glu Thr Pro Gly Ser	Glu Pro Ala Thr Ser
1115	1120	1125
Gly Ser Glu Thr Pro	Gly Ser Pro Ala Gly	Ser Pro Thr Ser Thr
1130	1135	1140
Glu Glu Gly Thr Ser	Thr Glu Pro Ser Glu	Gly Ser Ala Pro Gly
1145	1150	1155
Thr Ser Thr Glu Pro	Ser Glu Gly Ser Ala	Pro Gly Ser Glu Pro
1160	1165	1170
Ala Thr Ser Gly Ser	Glu Thr Pro Gly Thr	Ser Glu Ser Ala Thr
1175	1180	1185
Pro Glu Ser Gly Pro	Gly Thr Ser Thr Glu	Pro Ser Glu Gly Ser
1190	1195	1200
Ala Pro Ala Ser Ser	Pro Pro Val Leu	Lys Arg His Gln Ala Glu
1205	1210	1215
Ile Thr Arg Thr Thr	Leu Gln Ser Asp	Gln Glu Glu Ile Asp Tyr
1220	1225	1230

Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile		
1235	1240	1245
Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys		
1250	1255	1260
Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr		
1265	1270	1275
Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser		
1280	1285	1290
Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr		
1295	1300	1305
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu		
1310	1315	1320
His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp		
1325	1330	1335
Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser		
1340	1345	1350
Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly		
1355	1360	1365
Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr		
1370	1375	1380
Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu		
1385	1390	1395
Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu		
1400	1405	1410
Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His		
1415	1420	1425
Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln		
1430	1435	1440
Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp		
1445	1450	1455
Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn		
1460	1465	1470
Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His		
1475	1480	1485
Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met		
1490	1495	1500
Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser		
1505	1510	1515
Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr		

1520	1525	1530
Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr		
1535	1540	1545
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly		
1550	1555	1560
Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly		
1565	1570	1575
Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro		
1580	1585	1590
Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala		
1595	1600	1605
Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His		
1610	1615	1620
Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser		
1625	1630	1635
Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile		
1640	1645	1650
Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser		
1655	1660	1665
Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr		
1670	1675	1680
Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn		
1685	1690	1695
Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile		
1700	1705	1710
Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg		
1715	1720	1725
Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys		
1730	1735	1740
Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln		
1745	1750	1755
Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser		
1760	1765	1770
Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp		
1775	1780	1785
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe		
1790	1795	1800
Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys		
1805	1810	1815

Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser		
1820	1825	1830
Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys		
1835	1840	1845
Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val		
1850	1855	1860
Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His		
1865	1870	1875
Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu		
1880	1885	1890
Gly Cys Glu Ala Gln Asp Leu Tyr Asp Lys Thr His Thr Cys Pro		
1895	1900	1905
Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu		
1910	1915	1920
Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro		
1925	1930	1935
Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu		
1940	1945	1950
Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala		
1955	1960	1965
Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val		
1970	1975	1980
Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys		
1985	1990	1995
Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile		
2000	2005	2010
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln		
2015	2020	2025
Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln		
2030	2035	2040
Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile		
2045	2050	2055
Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys		
2060	2065	2070
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr		
2075	2080	2085
Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val		
2090	2095	2100
Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr		

2105	2110	2115
Gln Lys Ser Leu Ser Leu	Ser Pro Gly Lys	
2120	2125	
<210> 111		
<211> 2128		
<212> PRT		
<213> 人工序列		
<220>		
<223> FVIII 205蛋白序列		
<400> 111		
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu		
35	40	45
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
50	55	60
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
65	70	75
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
85	90	95
Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		
100	105	110
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser		
115	120	125
Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser		
130	135	140
Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		
145	150	155
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala Gly		
165	170	175
Ser Pro Thr Ser Thr Glu Glu Gly Ala Ser Ser Ser Asp Leu Gly Glu		
180	185	190
Leu Pro Val Asp Ala Arg Phe Pro Pro Arg Val Pro Lys Ser Phe Pro		
195	200	205
Phe Asn Thr Ser Val Val Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr		
210	215	220
Asp His Leu Phe Asn Ile Ala Lys Pro Arg Pro Pro Trp Met Gly Leu		

225	230	235	240
Leu Gly Pro Thr Ile	Gln Ala Glu Val Tyr	Asp Thr Val Val Ile	Thr
245	250	255	
Leu Lys Asn Met Ala	Ser His Pro Val Ser	Leu His Ala Val Gly	Val
260	265	270	
Ser Tyr Trp Lys Ala	Ser Glu Gly Ala Glu	Tyr Asp Asp Gln Thr	Ser
275	280	285	
Gln Arg Glu Lys Glu	Asp Asp Lys Val Phe	Pro Gly Gly Ser His	Thr
290	295	300	
Tyr Val Trp Gln Val	Leu Lys Glu Asn Gly	Pro Met Ala Ser Asp	Pro
305	310	315	320
Leu Cys Leu Thr Tyr	Ser Tyr Leu Ser His	Val Asp Leu Val Lys	Asp
325	330	335	
Leu Asn Ser Gly Leu	Ile Gly Ala Leu Leu	Val Cys Arg Glu Gly	Ser
340	345	350	
Leu Ala Lys Glu Lys	Thr Gln Thr Leu His	Lys Phe Ile Leu Leu	Phe
355	360	365	
Ala Val Phe Asp Glu	Gly Lys Ser Trp His	Ser Glu Thr Lys Asn	Ser
370	375	380	
Leu Met Gln Asp Arg	Asp Ala Ala Ser Ala	Arg Ala Trp Pro Lys	Met
385	390	395	400
His Thr Val Asn Gly	Tyr Val Asn Arg Ser	Leu Pro Gly Leu Ile	Gly
405	410	415	
Cys His Arg Lys Ser	Val Tyr Trp His Val	Ile Gly Met Gly Thr	Thr
420	425	430	
Pro Glu Val His Ser	Ile Phe Leu Glu Gly	His Thr Phe Leu Val	Arg
435	440	445	
Asn His Arg Gln Ala	Ser Leu Glu Ile Ser	Pro Ile Thr Phe Leu	Thr
450	455	460	
Ala Gln Thr Leu Leu	Met Asp Leu Gly Gln	Phe Leu Leu Phe Cys	His
465	470	475	480
Ile Ser Ser His Gln	His Asp Gly Met Glu	Ala Tyr Val Lys Val	Asp
485	490	495	
Ser Cys Pro Glu Glu	Pro Gln Leu Arg Met	Lys Asn Asn Glu Glu	Ala
500	505	510	
Glu Asp Tyr Asp Asp	Asp Leu Thr Asp Ser	Glu Met Asp Val Val	Arg
515	520	525	
Phe Asp Asp Asp Asn	Ser Pro Ser Phe Ile	Gln Ile Arg Ser Val	Ala
530	535	540	

Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu
545						550				555					560
Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr
				565						570					575
Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr
				580						585					590
Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg
		595								600					605
Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly
		610								615					620
Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg
625										630					640
Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr
										645					655
Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile
										660					670
Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp
										675					685
Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser
										690					700
Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu
705										710					720
Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met
										725					735
Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg
										740					750
Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala
										755					765
Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His
										770					780
Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu
785										790					800
His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp
										805					815
Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val
										820					830
Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe
										835					845
Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser

850	855	860
Asp Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys		
865	870	875
Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser		880
	885	890
Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser		895
	900	905
Gln Asn Gly Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		910
	915	920
Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser		925
	930	935
Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser		940
945	950	955
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		960
	965	970
Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro		975
	980	985
Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr		990
	995	1000
Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu		1005
	1010	1015
Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		1020
	1025	1030
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala		1035
	1040	1045
Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser		1050
	1055	1060
Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser		1065
	1070	1075
Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		1080
	1085	1090
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro		1095
	1100	1105
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser		1110
	1115	1120
Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr		1125
	1130	1135
Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		1140
	1145	1150
		1155

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro		
1160	1165	1170
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr		
1175	1180	1185
Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser		
1190	1195	1200
Ala Pro Ala Ser Ser Pro Pro Val Leu Lys Arg His Gln Ala Glu		
1205	1210	1215
Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr		
1220	1225	1230
Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile		
1235	1240	1245
Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys		
1250	1255	1260
Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr		
1265	1270	1275
Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser		
1280	1285	1290
Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr		
1295	1300	1305
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu		
1310	1315	1320
His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp		
1325	1330	1335
Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser		
1340	1345	1350
Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly		
1355	1360	1365
Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr		
1370	1375	1380
Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu		
1385	1390	1395
Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu		
1400	1405	1410
Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His		
1415	1420	1425
Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln		
1430	1435	1440
Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp		

1445	1450	1455
Tyr Phe Thr Glu Asn Met	Glu Arg Asn Cys Arg	Ala Pro Cys Asn
1460	1465	1470
Ile Gln Met Glu Asp Pro	Thr Phe Lys Glu Asn	Tyr Arg Phe His
1475	1480	1485
Ala Ile Asn Gly Tyr Ile	Met Asp Thr Leu Pro	Gly Leu Val Met
1490	1495	1500
Ala Gln Asp Gln Arg Ile	Arg Trp Tyr Leu Leu	Ser Met Gly Ser
1505	1510	1515
Asn Glu Asn Ile His Ser	Ile His Phe Ser Gly	His Val Phe Thr
1520	1525	1530
Val Arg Lys Lys Glu Glu	Tyr Lys Met Ala Leu	Tyr Asn Leu Tyr
1535	1540	1545
Pro Gly Val Phe Glu Thr	Val Glu Met Leu Pro	Ser Lys Ala Gly
1550	1555	1560
Ile Trp Arg Val Glu Cys	Leu Ile Gly Glu His	Leu His Ala Gly
1565	1570	1575
Met Ser Thr Leu Phe Leu	Val Tyr Ser Asn Lys	Cys Gln Thr Pro
1580	1585	1590
Leu Gly Met Ala Ser Gly	His Ile Arg Asp Phe	Gln Ile Thr Ala
1595	1600	1605
Ser Gly Gln Tyr Gly Gln	Trp Ala Pro Lys Leu	Ala Arg Leu His
1610	1615	1620
Tyr Ser Gly Ser Ile Asn	Ala Trp Ser Thr Lys	Glu Pro Phe Ser
1625	1630	1635
Trp Ile Lys Val Asp Leu	Leu Ala Pro Met Ile	Ile His Gly Ile
1640	1645	1650
Lys Thr Gln Gly Ala Arg	Gln Lys Phe Ser Ser	Leu Tyr Ile Ser
1655	1660	1665
Gln Phe Ile Ile Met Tyr	Ser Leu Asp Gly Lys	Lys Trp Gln Thr
1670	1675	1680
Tyr Arg Gly Asn Ser Thr	Gly Thr Leu Met Val	Phe Phe Gly Asn
1685	1690	1695
Val Asp Ser Ser Gly Ile	Lys His Asn Ile Phe	Asn Pro Pro Ile
1700	1705	1710
Ile Ala Arg Tyr Ile Arg	Leu His Pro Thr His	Tyr Ser Ile Arg
1715	1720	1725
Ser Thr Leu Arg Met Glu	Leu Met Gly Cys Asp	Leu Asn Ser Cys
1730	1735	1740

Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln		
1745	1750	1755
Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser		
1760	1765	1770
Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp		
1775	1780	1785
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe		
1790	1795	1800
Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys		
1805	1810	1815
Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser		
1820	1825	1830
Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys		
1835	1840	1845
Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val		
1850	1855	1860
Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His		
1865	1870	1875
Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu		
1880	1885	1890
Gly Cys Glu Ala Gln Asp Leu Tyr Asp Lys Thr His Thr Cys Pro		
1895	1900	1905
Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu		
1910	1915	1920
Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro		
1925	1930	1935
Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu		
1940	1945	1950
Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala		
1955	1960	1965
Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val		
1970	1975	1980
Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys		
1985	1990	1995
Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile		
2000	2005	2010
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln		
2015	2020	2025
Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln		

2030	2035	2040
Val Ser Leu Thr Cys Leu	Val Lys Gly Phe Tyr	Pro Ser Asp Ile
2045	2050	2055
Ala Val Glu Trp Glu Ser	Asn Gly Gln Pro Glu	Asn Asn Tyr Lys
2060	2065	2070
Thr Thr Pro Pro Val Leu	Asp Ser Asp Gly Ser	Phe Phe Leu Tyr
2075	2080	2085
Ser Lys Leu Thr Val Asp	Lys Ser Arg Trp Gln	Gln Gly Asn Val
2090	2095	2100
Phe Ser Cys Ser Val Met	His Glu Ala Leu His	Asn His Tyr Thr
2105	2110	2115
Gln Lys Ser Leu Ser Leu	Ser Pro Gly Lys	
2120	2125	
<210> 112		
<211> 2020		
<212> PRT		
<213> 人工序列		
<220>		
<223> pSYN FVIII 266蛋白序列		
<400> 112		
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Gly Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr		
35	40	45
Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro		
50	55	60
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Ala Ser Ser Ser		
65	70	75
Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro Arg Val Pro		
85	90	95
Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys Thr Leu Phe		
100	105	110
Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys Pro Arg Pro Pro		
115	120	125
Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val Tyr Asp Thr		
130	135	140
Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val Ser Leu His		

145	150	155	160
Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala Glu Tyr Asp			
	165	170	175
Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val Phe Pro Gly			
	180	185	190
Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn Gly Pro Met			
	195	200	205
Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser His Val Asp			
	210	215	220
Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu Leu Val Cys			
225	230	235	240
Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu His Lys Phe			
	245	250	255
Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp His Ser Glu			
	260	265	270
Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser Ala Arg Ala			
	275	280	285
Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg Ser Leu Pro			
	290	295	300
Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His Val Ile Gly			
305	310	315	320
Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu Gly His Thr			
	325	330	335
Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile Ser Pro Ile			
	340	345	350
Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly Gln Phe Leu			
	355	360	365
Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met Glu Ala Tyr			
	370	375	380
Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn			
385	390	395	400
Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met			
	405	410	415
Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile			
	420	425	430
Arg Ser Val Ala Lys Lys His Pro Lys Thr Trp Val His Tyr Ile Ala			
	435	440	445
Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu Ala Pro Asp			
450	455	460	

Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro	Gln	Arg	Ile
465					470				475						480
Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr	Asp	Glu	Thr
				485					490					495	
Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile	Leu	Gly	Pro
			500					505					510		
Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile	Phe	Lys	Asn
		515					520					525			
Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly	Ile	Thr	Asp	Val
	530					535					540				
Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val	Lys	His	Leu	Lys
545				550					555						560
Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys	Trp	Thr	Val
			565						570					575	
Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys	Leu	Thr	Arg
		580						585					590		
Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala	Ser	Gly	Leu
	595					600						605			
Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp	Gln	Arg	Gly
	610					615						620			
Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe	Ser	Val	Phe
625				630					635						640
Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile	Gln	Arg	Phe	Leu
			645						650					655	
Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu	Phe	Gln	Ala	Ser
	660							665					670		
Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp	Ser	Leu	Gln	Leu
	675					680						685			
Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile	Leu	Ser	Ile	Gly
	690					695					700				
Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr	Thr	Phe	Lys
705				710					715						720
His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro	Phe	Ser	Gly
			725						730					735	
Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp	Ile	Leu	Gly
		740						745					750		
Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala	Leu	Leu	Lys
	755						760					765			
Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu	Asp	Ser	Tyr

770	775	780
Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro		
785	790	795
Arg Ser Phe Ser Gln Asn Gly Ala Pro Gly Thr Ser Glu Ser Ala Thr		800
	805	810
Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr		815
	820	825
Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu		830
	835	840
Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr		845
	850	855
Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala		860
865	870	875
Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser		880
	885	890
Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser		895
	900	905
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		910
	915	920
Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro		925
	930	935
Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser		940
945	950	955
Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		960
	965	970
Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser		975
	980	985
Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser		990
	995	1000
Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu		1005
	1010	1015
Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly		1020
	1025	1030
Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr		1035
	1040	1045
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro Ala Thr Ser		1050
	1055	1060
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser		1065
	1070	1075
		1080

Gly Pro Gly Thr Ser Thr	Glu Pro Ser Glu Gly	Ser Ala Pro Ala
1085	1090	1095
Ser Ser Pro Pro Val Leu	Lys Arg His Gln Ala	Glu Ile Thr Arg
1100	1105	1110
Thr Thr Leu Gln Ser Asp	Gln Glu Glu Ile Asp	Tyr Asp Asp Thr
1115	1120	1125
Ile Ser Val Glu Met Lys	Lys Glu Asp Phe Asp	Ile Tyr Asp Glu
1130	1135	1140
Asp Glu Asn Gln Ser Pro	Arg Ser Phe Gln Lys	Lys Thr Arg His
1145	1150	1155
Tyr Phe Ile Ala Ala Val	Glu Arg Leu Trp Asp	Tyr Gly Met Ser
1160	1165	1170
Ser Ser Pro His Val Leu	Arg Asn Arg Ala Gln	Ser Gly Ser Val
1175	1180	1185
Pro Gln Phe Lys Lys Val	Val Phe Gln Glu Phe	Thr Asp Gly Ser
1190	1195	1200
Phe Thr Gln Pro Leu Tyr	Arg Gly Glu Leu Asn	Glu His Leu Gly
1205	1210	1215
Leu Leu Gly Pro Tyr Ile	Arg Ala Glu Val Glu	Asp Asn Ile Met
1220	1225	1230
Val Thr Phe Arg Asn Gln	Ala Ser Arg Pro Tyr	Ser Phe Tyr Ser
1235	1240	1245
Ser Leu Ile Ser Tyr Glu	Glu Asp Gln Arg Gln	Gly Ala Glu Pro
1250	1255	1260
Arg Lys Asn Phe Val Lys	Pro Asn Glu Thr Lys	Thr Tyr Phe Trp
1265	1270	1275
Lys Val Gln His His Met	Ala Pro Thr Lys Asp	Glu Phe Asp Cys
1280	1285	1290
Lys Ala Trp Ala Tyr Phe	Ser Asp Val Asp Leu	Glu Lys Asp Val
1295	1300	1305
His Ser Gly Leu Ile Gly	Pro Leu Leu Val Cys	His Thr Asn Thr
1310	1315	1320
Leu Asn Pro Ala His Gly	Arg Gln Val Thr Val	Gln Glu Phe Ala
1325	1330	1335
Leu Phe Phe Thr Ile Phe	Asp Glu Thr Lys Ser	Trp Tyr Phe Thr
1340	1345	1350
Glu Asn Met Glu Arg Asn	Cys Arg Ala Pro Cys	Asn Ile Gln Met
1355	1360	1365
Glu Asp Pro Thr Phe Lys	Glu Asn Tyr Arg Phe	His Ala Ile Asn

1370	1375	1380
Gly Tyr Ile Met Asp Thr	Leu Pro Gly Leu Val	Met Ala Gln Asp
1385	1390	1395
Gln Arg Ile Arg Trp Tyr	Leu Leu Ser Met Gly	Ser Asn Glu Asn
1400	1405	1410
Ile His Ser Ile His Phe	Ser Gly His Val Phe	Thr Val Arg Lys
1415	1420	1425
Lys Glu Glu Tyr Lys Met	Ala Leu Tyr Asn Leu	Tyr Pro Gly Val
1430	1435	1440
Phe Glu Thr Val Glu Met	Leu Pro Ser Lys Ala	Gly Ile Trp Arg
1445	1450	1455
Val Glu Cys Leu Ile Gly	Glu His Leu His Ala	Gly Met Ser Thr
1460	1465	1470
Leu Phe Leu Val Tyr Ser	Asn Lys Cys Gln Thr	Pro Leu Gly Met
1475	1480	1485
Ala Ser Gly His Ile Arg	Asp Phe Gln Ile Thr	Ala Ser Gly Gln
1490	1495	1500
Tyr Gly Gln Trp Ala Pro	Lys Leu Ala Arg Leu	His Tyr Ser Gly
1505	1510	1515
Ser Ile Asn Ala Trp Ser	Thr Lys Glu Pro Phe	Ser Trp Ile Lys
1520	1525	1530
Val Asp Leu Leu Ala Pro	Met Ile Ile His Gly	Ile Lys Thr Gln
1535	1540	1545
Gly Ala Arg Gln Lys Phe	Ser Ser Leu Tyr Ile	Ser Gln Phe Ile
1550	1555	1560
Ile Met Tyr Ser Leu Asp	Gly Lys Lys Trp Gln	Thr Tyr Arg Gly
1565	1570	1575
Asn Ser Thr Gly Thr Leu	Met Val Phe Phe Gly	Asn Val Asp Ser
1580	1585	1590
Ser Gly Ile Lys His Asn	Ile Phe Asn Pro Pro	Ile Ile Ala Arg
1595	1600	1605
Tyr Ile Arg Leu His Pro	Thr His Tyr Ser Ile	Arg Ser Thr Leu
1610	1615	1620
Arg Met Glu Leu Met Gly	Cys Asp Leu Asn Ser	Cys Ser Met Pro
1625	1630	1635
Leu Gly Met Glu Ser Lys	Ala Ile Ser Asp Ala	Gln Ile Thr Ala
1640	1645	1650
Ser Ser Tyr Phe Thr Asn	Met Phe Ala Thr Trp	Ser Pro Ser Lys
1655	1660	1665

Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln		
1670	1675	1680
Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr		
1685	1690	1695
Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu		
1700	1705	1710
Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp		
1715	1720	1725
Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val		
1730	1735	1740
Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu		
1745	1750	1755
Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser		
1760	1765	1770
Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu		
1775	1780	1785
Ala Gln Asp Leu Tyr Asp Lys Thr His Thr Cys Pro Pro Cys Pro		
1790	1795	1800
Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro		
1805	1810	1815
Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr		
1820	1825	1830
Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe		
1835	1840	1845
Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys		
1850	1855	1860
Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val		
1865	1870	1875
Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys		
1880	1885	1890
Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr		
1895	1900	1905
Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr		
1910	1915	1920
Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu		
1925	1930	1935
Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu		
1940	1945	1950
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro		

1955	1960	1965
Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu		
1970	1975	1980
Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys		
1985	1990	1995
Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser		
2000	2005	2010
Leu Ser Leu Ser Pro Gly Lys		
2015	2020	
<210> 113		
<211> 2056		
<212> PRT		
<213> 人工序列		
<220>		
<223> pSYN FVIII 267蛋白序列		
<400> 113		
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1 5 10 15		
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20 25 30		
Trp Asp Tyr Met Gln Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu		
35 40 45		
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
50 55 60		
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
65 70 75 80		
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
85 90 95		
Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		
100 105 110		
Ala Ser Ser Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro		
115 120 125		
Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys		
130 135 140		
Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys		
145 150 155 160		
Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu		
165 170 175		
Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro		

180							185					190				
Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser	Glu	Gly	
195							200					205				
Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp	Asp	Lys	
210							215					220				
Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu	Lys	Glu	
225							230					235				
Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser	Tyr	Leu	
245							250					255				
Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile	Gly	Ala	
260							265					270				
Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr	Gln	Thr	
275							280					285				
Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly	Lys	Ser	
290							295					300				
Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp	Ala	Ala	
305							310					315				
Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr	Val	Asn	
325							330					335				
Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val	Tyr	Trp	
340							345					350				
His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile	Phe	Leu	
355							360					365				
Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser	Leu	Glu	
370							375					380				
Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met	Asp	Leu	
385							390					395				
Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His	Asp	Gly	
405							410					415				
Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro	Gln	Leu	
420							425					430				
Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	Leu	Thr	
435							440					445				
Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	
450							455					460				
Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	
465							470					475				
His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	
485							490					495				

Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly		
500	505	510
Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr		
515	520	525
Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly		
530	535	540
Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile		
545	550	555
Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly		
565	570	575
Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val		
580	585	590
Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr		
595	600	605
Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg		
610	615	620
Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu		
625	630	635
Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val		
645	650	655
Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu		
660	665	670
Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile		
675	680	685
Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu		
690	695	700
Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp		
705	710	715
Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile		
725	730	735
Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly		
740	745	750
Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe		
755	760	765
Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu		
770	775	780
Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr		
785	790	795
Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr		

	805		810		815
Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn					
	820		825		830
Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Gly Ala Pro Gly Thr Ser					
	835		840		845
Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser					
	850		855		860
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly					
865		870		875	880
Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser					
	885		890		895
Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser					
	900		905		910
Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu					
	915		920		925
Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu					
	930		935		940
Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr					
945		950		955	960
Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu					
	965		970		975
Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser					
	980		985		990
Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr					
	995		1000		1005
Pro Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser					
	1010		1015		1020
Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly					
	1025		1030		1035
Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro					
	1040		1045		1050
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro					
	1055		1060		1065
Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser					
	1070		1075		1080
Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly					
	1085		1090		1095
Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu					
	1100		1105		1110

Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser		
1115	1120	1125
Glu Gly Ser Ala Pro Ala Ser Ser Pro Pro Val Leu Lys Arg His		
1130	1135	1140
Gln Ala Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu		
1145	1150	1155
Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp		
1160	1165	1170
Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe		
1175	1180	1185
Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu		
1190	1195	1200
Trp Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg		
1205	1210	1215
Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln		
1220	1225	1230
Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu		
1235	1240	1245
Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu		
1250	1255	1260
Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg		
1265	1270	1275
Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln		
1280	1285	1290
Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu		
1295	1300	1305
Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr		
1310	1315	1320
Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val		
1325	1330	1335
Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu		
1340	1345	1350
Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val		
1355	1360	1365
Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr		
1370	1375	1380
Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala		
1385	1390	1395
Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr		

1400	1405	1410
Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly		
1415	1420	1425
Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser		
1430	1435	1440
Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His		
1445	1450	1455
Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr		
1460	1465	1470
Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser		
1475	1480	1485
Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu		
1490	1495	1500
His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys		
1505	1510	1515
Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln		
1520	1525	1530
Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala		
1535	1540	1545
Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu		
1550	1555	1560
Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile		
1565	1570	1575
His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu		
1580	1585	1590
Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys		
1595	1600	1605
Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe		
1610	1615	1620
Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn		
1625	1630	1635
Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr		
1640	1645	1650
Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu		
1655	1660	1665
Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser		
1670	1675	1680
Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala		
1685	1690	1695

Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser		
1700	1705	1710
Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln		
1715	1720	1725
Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln		
1730	1735	1740
Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu		
1745	1750	1755
Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln		
1760	1765	1770
Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr		
1775	1780	1785
Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu		
1790	1795	1800
Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met		
1805	1810	1815
Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr Asp Lys Thr His		
1820	1825	1830
Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser		
1835	1840	1845
Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser		
1850	1855	1860
Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu		
1865	1870	1875
Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val		
1880	1885	1890
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr		
1895	1900	1905
Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu		
1910	1915	1920
Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro		
1925	1930	1935
Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg		
1940	1945	1950
Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr		
1955	1960	1965
Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro		
1970	1975	1980
Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn		

1985	1990	1995
Asn Tyr Lys Thr Thr Pro	Pro Val Leu Asp Ser	Asp Gly Ser Phe
2000	2005	2010
Phe Leu Tyr Ser Lys Leu	Thr Val Asp Lys Ser	Arg Trp Gln Gln
2015	2020	2025
Gly Asn Val Phe Ser Cys	Ser Val Met His Glu	Ala Leu His Asn
2030	2035	2040
His Tyr Thr Gln Lys Ser	Leu Ser Leu Ser Pro	Gly Lys
2045	2050	2055
<210> 114		
<211> 1834		
<212> PRT		
<213> 人工序列		
<220>		
<223> pSYN FVIII 268蛋白序列		
<400> 114		
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu		
35	40	45
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
50	55	60
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
65	70	75
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
85	90	95
Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		
100	105	110
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser		
115	120	125
Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser		
130	135	140
Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		
145	150	155
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala Gly		
165	170	175
Ser Pro Thr Ser Thr Glu Glu Gly Ala Ser Ser Ser Asp Leu Gly Glu		

180	185	190
Leu Pro Val Asp Ala Arg Phe	Pro Pro Arg Val Pro Lys Ser Phe Pro	
195	200	205
Phe Asn Thr Ser Val Val Tyr	Lys Lys Thr Leu Phe Val Glu Phe Thr	
210	215	220
Asp His Leu Phe Asn Ile Ala	Lys Pro Arg Pro Pro Trp Met Gly Leu	
225	230	235
Leu Gly Pro Thr Ile Gln Ala	Glu Val Tyr Asp Thr Val Val Ile Thr	
245	250	255
Leu Lys Asn Met Ala Ser His	Pro Val Ser Leu His Ala Val Gly Val	
260	265	270
Ser Tyr Trp Lys Ala Ser Glu	Gly Ala Glu Tyr Asp Asp Gln Thr Ser	
275	280	285
Gln Arg Glu Lys Glu Asp Asp	Lys Val Phe Pro Gly Gly Ser His Thr	
290	295	300
Tyr Val Trp Gln Val Leu Lys	Glu Asn Gly Pro Met Ala Ser Asp Pro	
305	310	315
Leu Cys Leu Thr Tyr Ser Tyr	Leu Ser His Val Asp Leu Val Lys Asp	
325	330	335
Leu Asn Ser Gly Leu Ile Gly	Ala Leu Leu Val Cys Arg Glu Gly Ser	
340	345	350
Leu Ala Lys Glu Lys Thr Gln	Thr Leu His Lys Phe Ile Leu Leu Phe	
355	360	365
Ala Val Phe Asp Glu Gly Lys	Ser Trp His Ser Glu Thr Lys Asn Ser	
370	375	380
Leu Met Gln Asp Arg Asp Ala	Ala Ser Ala Arg Ala Trp Pro Lys Met	
385	390	395
His Thr Val Asn Gly Tyr Val	Asn Arg Ser Leu Pro Gly Leu Ile Gly	
405	410	415
Cys His Arg Lys Ser Val Tyr	Trp His Val Ile Gly Met Gly Thr Thr	
420	425	430
Pro Glu Val His Ser Ile Phe	Leu Glu Gly His Thr Phe Leu Val Arg	
435	440	445
Asn His Arg Gln Ala Ser Leu	Glu Ile Ser Pro Ile Thr Phe Leu Thr	
450	455	460
Ala Gln Thr Leu Leu Met Asp	Leu Gly Gln Phe Leu Leu Phe Cys His	
465	470	475
Ile Ser Ser His Gln His Asp	Gly Met Glu Ala Tyr Val Lys Val Asp	
485	490	495

Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn Asn Glu Glu Ala		
500	505	510
Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met Asp Val Val Arg		
515	520	525
Phe Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile Arg Ser Val Ala		
530	535	540
Lys Lys His Pro Lys Thr Trp Val His Tyr Ile Ala Ala Glu Glu Glu		
545	550	555
Asp Trp Asp Tyr Ala Pro Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr		
565	570	575
Lys Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr		
580	585	590
Lys Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg		
595	600	605
Glu Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly		
610	615	620
Glu Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg		
625	630	635
Pro Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr		
645	650	655
Ser Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile		
660	665	670
Leu Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp		
675	680	685
Gly Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser		
690	695	700
Phe Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu		
705	710	715
Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met		
725	730	735
Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg		
740	745	750
Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala		
755	760	765
Gly Val Gln Leu Glu Asp Pro Glu Phe Gln Ala Ser Asn Ile Met His		
770	775	780
Ser Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu		
785	790	795
His Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp		

				805				810				815			
Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val
				820				825				830			
Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe
				835				840				845			
Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser
				850				855				860			
Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys
865				870				875				880			
Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser
				885				890				895			
Ala	Tyr	Leu	Leu	Ser	Lys	Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser
				900				905				910			
Gln	Asn	Pro	Pro	Val	Leu	Lys	Arg	His	Gln	Ala	Glu	Ile	Thr	Arg	Thr
				915				920				925			
Thr	Leu	Gln	Ser	Asp	Gln	Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser
930				935				940							
Val	Glu	Met	Lys	Lys	Glu	Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn
945				950				955				960			
Gln	Ser	Pro	Arg	Ser	Phe	Gln	Lys	Lys	Thr	Arg	His	Tyr	Phe	Ile	Ala
				965				970				975			
Ala	Val	Glu	Arg	Leu	Trp	Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	His	Val
				980				985				990			
Leu	Arg	Asn	Arg	Ala	Gln	Ser	Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val
				995				1000				1005			
Val	Phe	Gln	Glu	Phe	Thr	Asp	Gly	Ser	Phe	Thr	Gln	Pro	Leu	Tyr	
1010				1015				1020							
Arg	Gly	Glu	Leu	Asn	Glu	His	Leu	Gly	Leu	Leu	Gly	Pro	Tyr	Ile	
1025				1030				1035							
Arg	Ala	Glu	Val	Glu	Asp	Asn	Ile	Met	Val	Thr	Phe	Arg	Asn	Gln	
1040				1045				1050							
Ala	Ser	Arg	Pro	Tyr	Ser	Phe	Tyr	Ser	Ser	Leu	Ile	Ser	Tyr	Glu	
1055				1060				1065							
Glu	Asp	Gln	Arg	Gln	Gly	Ala	Glu	Pro	Arg	Lys	Asn	Phe	Val	Lys	
1070				1075				1080							
Pro	Asn	Glu	Thr	Lys	Thr	Tyr	Phe	Trp	Lys	Val	Gln	His	His	Met	
1085				1090				1095							
Ala	Pro	Thr	Lys	Asp	Glu	Phe	Asp	Cys	Lys	Ala	Trp	Ala	Tyr	Phe	
1100				1105				1110							

Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly		
1115	1120	1125
Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly		
1130	1135	1140
Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe		
1145	1150	1155
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn		
1160	1165	1170
Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys		
1175	1180	1185
Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr		
1190	1195	1200
Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr		
1205	1210	1215
Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe		
1220	1225	1230
Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met		
1235	1240	1245
Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met		
1250	1255	1260
Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly		
1265	1270	1275
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1280	1285	1290
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1295	1300	1305
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1310	1315	1320
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1325	1330	1335
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1340	1345	1350
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1355	1360	1365
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1370	1375	1380
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1385	1390	1395
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		

1400	1405	1410
Ile Phe Asn Pro Pro Ile	Ile Ala Arg Tyr Ile	Arg Leu His Pro
1415	1420	1425
Thr His Tyr Ser Ile Arg	Ser Thr Leu Arg Met	Glu Leu Met Gly
1430	1435	1440
Cys Asp Leu Asn Ser Cys	Ser Met Pro Leu Gly	Met Glu Ser Lys
1445	1450	1455
Ala Ile Ser Asp Ala Gln	Ile Thr Ala Ser Ser	Tyr Phe Thr Asn
1460	1465	1470
Met Phe Ala Thr Trp Ser	Pro Ser Lys Ala Arg	Leu His Leu Gln
1475	1480	1485
Gly Arg Ser Asn Ala Trp	Arg Pro Gln Val Asn	Asn Pro Lys Glu
1490	1495	1500
Trp Leu Gln Val Asp Phe	Gln Lys Thr Met Lys	Val Thr Gly Val
1505	1510	1515
Thr Thr Gln Gly Val Lys	Ser Leu Leu Thr Ser	Met Tyr Val Lys
1520	1525	1530
Glu Phe Leu Ile Ser Ser	Ser Gln Asp Gly His	Gln Trp Thr Leu
1535	1540	1545
Phe Phe Gln Asn Gly Lys	Val Lys Val Phe Gln	Gly Asn Gln Asp
1550	1555	1560
Ser Phe Thr Pro Val Val	Asn Ser Leu Asp Pro	Pro Leu Leu Thr
1565	1570	1575
Arg Tyr Leu Arg Ile His	Pro Gln Ser Trp Val	His Gln Ile Ala
1580	1585	1590
Leu Arg Met Glu Val Leu	Gly Cys Glu Ala Gln	Asp Leu Tyr Asp
1595	1600	1605
Lys Thr His Thr Cys Pro	Pro Cys Pro Ala Pro	Glu Leu Leu Gly
1610	1615	1620
Gly Pro Ser Val Phe Leu	Phe Pro Pro Lys Pro	Lys Asp Thr Leu
1625	1630	1635
Met Ile Ser Arg Thr Pro	Glu Val Thr Cys Val	Val Val Asp Val
1640	1645	1650
Ser His Glu Asp Pro Glu	Val Lys Phe Asn Trp	Tyr Val Asp Gly
1655	1660	1665
Val Glu Val His Asn Ala	Lys Thr Lys Pro Arg	Glu Glu Gln Tyr
1670	1675	1680
Asn Ser Thr Tyr Arg Val	Val Ser Val Leu Thr	Val Leu His Gln
1685	1690	1695

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys		
1700	1705	1710
Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly		
1715	1720	1725
Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp		
1730	1735	1740
Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly		
1745	1750	1755
Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln		
1760	1765	1770
Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp		
1775	1780	1785
Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg		
1790	1795	1800
Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala		
1805	1810	1815
Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly		
1820	1825	1830

Lys

<210> 115

<211> 1762

<212> PRT

<213> 人工序列

<220>

<223> pSYN FVIII 269蛋白序列

<400> 115

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe			
1	5	10	15
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser			
20	25	30	
Trp Asp Tyr Met Gln Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu			
35	40	45	
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly			
50	55	60	
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala			
65	70	75	80
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu			
85	90	95	
Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly			

100	105	110
Ala Ser Ser Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro		
115	120	125
Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys		
130	135	140
Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys		
145	150	155
Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu		
165	170	175
Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro		
180	185	190
Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly		
195	200	205
Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys		
210	215	220
Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu		
225	230	235
Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu		
245	250	255
Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala		
260	265	270
Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr		
275	280	285
Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser		
290	295	300
Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala		
305	310	315
Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn		
325	330	335
Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp		
340	345	350
His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu		
355	360	365
Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu		
370	375	380
Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu		
385	390	395
Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly		
405	410	415

Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro	Gln	Leu			
																420	425	430
Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	Leu	Thr			
																435	440	445
Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser			
																450	455	460
Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val			
																465	470	475
His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val			
																485	490	495
Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly			
																500	505	510
Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr			
																515	520	525
Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly			
																530	535	540
Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile			
																545	550	555
Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly			
																565	570	575
Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val			
																580	585	590
Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr			
																595	600	605
Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg			
																610	615	620
Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu			
																625	630	635
Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val			
																645	650	655
Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu			
																660	665	670
Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile			
																675	680	685
Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu			
																690	695	700
Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp			
																705	710	715
Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile			

	725		730		735
Leu Ser Ile Gly Ala Gln Thr Asp Phe	Leu Ser Val Phe Phe Ser Gly				
740	745	750			
Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe					
755	760	765			
Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu					
770	775	780			
Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr					
785	790	795	800		
Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr					
805	810	815			
Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn					
820	825	830			
Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu Lys Arg					
835	840	845			
His Gln Ala Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu					
850	855	860			
Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe					
865	870	875	880		
Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys					
885	890	895			
Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr					
900	905	910			
Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly					
915	920	925			
Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly					
930	935	940			
Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly					
945	950	955	960		
Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val					
965	970	975			
Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu					
980	985	990			
Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn					
995	1000	1005			
Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln					
1010	1015	1020			
His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp					
1025	1030	1035			

Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly		
1040	1045	1050
Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro		
1055	1060	1065
Ala His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe		
1070	1075	1080
Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met		
1085	1090	1095
Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro		
1100	1105	1110
Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile		
1115	1120	1125
Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile		
1130	1135	1140
Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser		
1145	1150	1155
Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu		
1160	1165	1170
Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr		
1175	1180	1185
Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys		
1190	1195	1200
Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu		
1205	1210	1215
Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly		
1220	1225	1230
His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln		
1235	1240	1245
Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn		
1250	1255	1260
Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu		
1265	1270	1275
Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg		
1280	1285	1290
Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr		
1295	1300	1305
Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr		
1310	1315	1320
Gly Thr Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile		

1325	1330	1335
Lys His Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg		
1340	1345	1350
Leu His Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu		
1355	1360	1365
Leu Met Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met		
1370	1375	1380
Glu Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr		
1385	1390	1395
Phe Thr Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu		
1400	1405	1410
His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn		
1415	1420	1425
Pro Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val		
1430	1435	1440
Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met		
1445	1450	1455
Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln		
1460	1465	1470
Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly		
1475	1480	1485
Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro		
1490	1495	1500
Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His		
1505	1510	1515
Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp		
1520	1525	1530
Leu Tyr Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu		
1535	1540	1545
Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys		
1550	1555	1560
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val		
1565	1570	1575
Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr		
1580	1585	1590
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu		
1595	1600	1605
Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val		
1610	1615	1620

Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val		
1625	1630	1635
Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys		
1640	1645	1650
Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro		
1655	1660	1665
Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu		
1670	1675	1680
Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser		
1685	1690	1695
Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu		
1700	1705	1710
Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp		
1715	1720	1725
Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met		
1730	1735	1740
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu		
1745	1750	1755

Ser Pro Gly Lys

1760

<210> 116

<211> 1726

<212> PRT

<213> 人工序列

<220>

<223> pSYNFVIII 271蛋白序列

<400> 116

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Gly Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr		
35	40	45
Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro		
50	55	60
Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Ala Ser Ser Ser		
65	70	75
Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro Arg Val Pro		
85	90	95

Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys Thr Leu Phe			
100	105	110	
Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys Pro Arg Pro Pro			
115	120	125	
Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val Tyr Asp Thr			
130	135	140	
Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val Ser Leu His			
145	150	155	160
Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala Glu Tyr Asp			
165	170	175	
Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val Phe Pro Gly			
180	185	190	
Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn Gly Pro Met			
195	200	205	
Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser His Val Asp			
210	215	220	
Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu Leu Val Cys			
225	230	235	240
Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu His Lys Phe			
245	250	255	
Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp His Ser Glu			
260	265	270	
Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser Ala Arg Ala			
275	280	285	
Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg Ser Leu Pro			
290	295	300	
Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His Val Ile Gly			
305	310	315	320
Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu Gly His Thr			
325	330	335	
Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile Ser Pro Ile			
340	345	350	
Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly Gln Phe Leu			
355	360	365	
Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met Glu Ala Tyr			
370	375	380	
Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn			
385	390	395	400
Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met			

				405				410				415			
Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	Phe	Ile	Gln	Ile
				420				425				430			
Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His	Tyr	Ile	Ala
				435				440				445			
Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	Leu	Ala	Pro	Asp
				450				455				460			
Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro	Gln	Arg	Ile
465				470				475				480			
Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr	Asp	Glu	Thr
				485				490				495			
Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile	Leu	Gly	Pro
				500				505				510			
Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile	Phe	Lys	Asn
				515				520				525			
Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly	Ile	Thr	Asp	Val
530				535				540							
Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val	Lys	His	Leu	Lys
545				550				555				560			
Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys	Trp	Thr	Val
				565				570				575			
Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys	Leu	Thr	Arg
				580				585				590			
Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala	Ser	Gly	Leu
				595				600				605			
Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp	Gln	Arg	Gly
610				615				620							
Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe	Ser	Val	Phe
625				630				635				640			
Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile	Gln	Arg	Phe	Leu
				645				650				655			
Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu	Phe	Gln	Ala	Ser
				660				665				670			
Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp	Ser	Leu	Gln	Leu
675				680				685							
Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile	Leu	Ser	Ile	Gly
690				695				700							
Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr	Thr	Phe	Lys
705				710				715				720			

His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro Phe Ser Gly		
725	730	735
Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp Ile Leu Gly		
740	745	750
Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys		
755	760	765
Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr		
770	775	780
Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro		
785	790	800
Arg Ser Phe Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Ala Glu		
805	810	815
Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp		
820	825	830
Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp		
835	840	845
Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His		
850	855	860
Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser		
865	870	880
Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln		
885	890	895
Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln		
900	905	910
Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro		
915	920	925
Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn		
930	935	940
Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu		
945	950	960
Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys Pro		
965	970	975
Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met Ala Pro		
980	985	990
Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val		
995	1000	1005
Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu		
1010	1015	1020
Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val		

1025	1030	1035
Thr Val Gln Glu Phe Ala	Leu Phe Phe Thr Ile	Phe Asp Glu Thr
1040	1045	1050
Lys Ser Trp Tyr Phe Thr	Glu Asn Met Glu Arg	Asn Cys Arg Ala
1055	1060	1065
Pro Cys Asn Ile Gln Met	Glu Asp Pro Thr Phe	Lys Glu Asn Tyr
1070	1075	1080
Arg Phe His Ala Ile Asn	Gly Tyr Ile Met Asp	Thr Leu Pro Gly
1085	1090	1095
Leu Val Met Ala Gln Asp	Gln Arg Ile Arg Trp	Tyr Leu Leu Ser
1100	1105	1110
Met Gly Ser Asn Glu Asn	Ile His Ser Ile His	Phe Ser Gly His
1115	1120	1125
Val Phe Thr Val Arg Lys	Lys Glu Glu Tyr Lys	Met Ala Leu Tyr
1130	1135	1140
Asn Leu Tyr Pro Gly Val	Phe Glu Thr Val Glu	Met Leu Pro Ser
1145	1150	1155
Lys Ala Gly Ile Trp Arg	Val Glu Cys Leu Ile	Gly Glu His Leu
1160	1165	1170
His Ala Gly Met Ser Thr	Leu Phe Leu Val Tyr	Ser Asn Lys Cys
1175	1180	1185
Gln Thr Pro Leu Gly Met	Ala Ser Gly His Ile	Arg Asp Phe Gln
1190	1195	1200
Ile Thr Ala Ser Gly Gln	Tyr Gly Gln Trp Ala	Pro Lys Leu Ala
1205	1210	1215
Arg Leu His Tyr Ser Gly	Ser Ile Asn Ala Trp	Ser Thr Lys Glu
1220	1225	1230
Pro Phe Ser Trp Ile Lys	Val Asp Leu Leu Ala	Pro Met Ile Ile
1235	1240	1245
His Gly Ile Lys Thr Gln	Gly Ala Arg Gln Lys	Phe Ser Ser Leu
1250	1255	1260
Tyr Ile Ser Gln Phe Ile	Ile Met Tyr Ser Leu	Asp Gly Lys Lys
1265	1270	1275
Trp Gln Thr Tyr Arg Gly	Asn Ser Thr Gly Thr	Leu Met Val Phe
1280	1285	1290
Phe Gly Asn Val Asp Ser	Ser Gly Ile Lys His	Asn Ile Phe Asn
1295	1300	1305
Pro Pro Ile Ile Ala Arg	Tyr Ile Arg Leu His	Pro Thr His Tyr
1310	1315	1320

Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu		
1325	1330	1335
Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser		
1340	1345	1350
Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala		
1355	1360	1365
Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser		
1370	1375	1380
Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln		
1385	1390	1395
Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln		
1400	1405	1410
Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu		
1415	1420	1425
Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln		
1430	1435	1440
Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr		
1445	1450	1455
Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu		
1460	1465	1470
Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met		
1475	1480	1485
Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr Asp Lys Thr His		
1490	1495	1500
Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser		
1505	1510	1515
Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser		
1520	1525	1530
Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu		
1535	1540	1545
Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val		
1550	1555	1560
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr		
1565	1570	1575
Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu		
1580	1585	1590
Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro		
1595	1600	1605
Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg		

1610	1615	1620
Glu Pro Gln Val Tyr Thr	Leu Pro Pro Ser Arg	Asp Glu Leu Thr
1625	1630	1635
Lys Asn Gln Val Ser Leu	Thr Cys Leu Val Lys	Gly Phe Tyr Pro
1640	1645	1650
Ser Asp Ile Ala Val Glu	Trp Glu Ser Asn Gly	Gln Pro Glu Asn
1655	1660	1665
Asn Tyr Lys Thr Thr Pro	Pro Val Leu Asp Ser	Asp Gly Ser Phe
1670	1675	1680
Phe Leu Tyr Ser Lys Leu	Thr Val Asp Lys Ser	Arg Trp Gln Gln
1685	1690	1695
Gly Asn Val Phe Ser Cys	Ser Val Met His Glu	Ala Leu His Asn
1700	1705	1710
His Tyr Thr Gln Lys Ser	Leu Ser Leu Ser Pro	Gly Lys
1715	1720	1725
<210> 117		
<211> 1901		
<212> PRT		
<213> 人工序列		
<220>		
<223> pSYN FVIII蛋白序列272		
<400> 117		
Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe		
1	5	10
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser		
20	25	30
Trp Asp Tyr Met Gln Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu		
35	40	45
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly		
50	55	60
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala		
65	70	75
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu		
85	90	95
Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		
100	105	110
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser		
115	120	125
Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser		

130	135	140
Glu Thr Pro Gly Thr Ser	Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly
145	150	155
Ser Pro Ala Gly Ser Pro	Thr Ser Thr Glu Glu Gly	Ser Pro Ala Gly
165	170	175
Ser Pro Thr Ser Thr Glu	Glu Gly Ala Ser Ser Ser	Asp Leu Gly Glu
180	185	190
Leu Pro Val Asp Ala Arg	Phe Pro Pro Arg Val	Pro Lys Ser Phe Pro
195	200	205
Phe Asn Thr Ser Val Val	Tyr Lys Lys Thr Leu	Phe Val Glu Phe Thr
210	215	220
Asp His Leu Phe Asn Ile	Ala Lys Pro Arg Pro	Pro Trp Met Gly Leu
225	230	235
Leu Gly Pro Thr Ile Gln	Ala Glu Val Tyr Asp	Thr Val Val Ile Thr
245	250	255
Leu Lys Asn Met Ala Ser	His Pro Val Ser Leu	His Ala Val Gly Val
260	265	270
Ser Tyr Trp Lys Ala Ser	Glu Gly Ala Glu Tyr	Asp Asp Gln Thr Ser
275	280	285
Gln Arg Glu Lys Glu Asp	Asp Lys Val Phe Pro	Gly Gly Ser His Thr
290	295	300
Tyr Val Trp Gln Val Leu	Lys Glu Asn Gly Pro	Met Ala Ser Asp Pro
305	310	315
Leu Cys Leu Thr Tyr Ser	Tyr Leu Ser His Val	Asp Leu Val Lys Asp
325	330	335
Leu Asn Ser Gly Leu Ile	Gly Ala Leu Leu Val	Cys Arg Glu Gly Ser
340	345	350
Leu Ala Lys Glu Lys Thr	Gln Thr Leu His Lys	Phe Ile Leu Leu Phe
355	360	365
Ala Val Phe Asp Glu Gly	Lys Ser Trp His Ser	Glu Thr Lys Asn Ser
370	375	380
Leu Met Gln Asp Arg Asp	Ala Ala Ser Ala Arg	Ala Trp Pro Lys Met
385	390	395
His Thr Val Asn Gly Tyr	Val Asn Arg Ser Leu	Pro Gly Leu Ile Gly
405	410	415
Cys His Arg Lys Ser Val	Tyr Trp His Val Ile	Gly Met Gly Thr Thr
420	425	430
Pro Glu Val His Ser Ile	Phe Leu Glu Gly His	Thr Phe Leu Val Arg
435	440	445

Asn His Arg Gln Ala Ser Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr			
450	455	460	
Ala Gln Thr Leu Leu Met Asp Leu Gly Gln Phe Leu Leu Phe Cys His			
465	470	475	480
Ile Ser Ser His Gln His Asp Gly Met Glu Ala Tyr Val Lys Val Asp			
	485	490	495
Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn Asn Glu Glu Ala			
	500	505	510
Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met Asp Val Val Arg			
	515	520	525
Phe Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile Arg Ser Val Ala			
	530	535	540
Lys Lys His Pro Lys Thr Trp Val His Tyr Ile Ala Ala Glu Glu Glu			
545	550	555	560
Asp Trp Asp Tyr Ala Pro Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr			
	565	570	575
Lys Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr			
	580	585	590
Lys Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg			
	595	600	605
Glu Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly			
	610	615	620
Glu Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg			
625	630	635	640
Pro Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr			
	645	650	655
Ser Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile			
	660	665	670
Leu Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp			
	675	680	685
Gly Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser			
	690	695	700
Phe Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu			
705	710	715	720
Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met			
	725	730	735
Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg			
	740	745	750
Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala			

755	760	765
Gly Val Gln Leu Glu Asp Pro	Glu Phe Gln Ala Ser Asn Ile Met His	
770	775	780
Ser Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu		
785	790	795
His Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp		800
805	810	815
Phe Leu Ser Val Phe Phe Ser Gly Tyr Thr Phe Lys His Lys Met Val		
820	825	830
Tyr Glu Asp Thr Leu Thr Leu Phe Pro Phe Ser Gly Glu Thr Val Phe		
835	840	845
Met Ser Met Glu Asn Pro Gly Leu Trp Ile Leu Gly Cys His Asn Ser		
850	855	860
Asp Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys		
865	870	875
Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser		
885	890	895
Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser		
900	905	910
Gln Asn Gly Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		
915	920	925
Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser		
930	935	940
Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser		
945	950	955
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly		
965	970	975
Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro		
980	985	990
Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr		
995	1000	1005
Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu		
1010	1015	1020
Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		
1025	1030	1035
Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala		
1040	1045	1050
Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser		
1055	1060	1065

Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser		
1070	1075	1080
Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly		
1085	1090	1095
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro		
1100	1105	1110
Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala Thr Ser		
1115	1120	1125
Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr		
1130	1135	1140
Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly		
1145	1150	1155
Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro		
1160	1165	1170
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr		
1175	1180	1185
Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser		
1190	1195	1200
Ala Pro Ala Ser Ser Pro Pro Val Leu Lys Arg His Gln Ala Glu		
1205	1210	1215
Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr		
1220	1225	1230
Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile		
1235	1240	1245
Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys		
1250	1255	1260
Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr		
1265	1270	1275
Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser		
1280	1285	1290
Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr		
1295	1300	1305
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu		
1310	1315	1320
His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp		
1325	1330	1335
Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser		
1340	1345	1350
Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly		

1355	1360	1365
Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr		
1370	1375	1380
Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu		
1385	1390	1395
Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu		
1400	1405	1410
Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His		
1415	1420	1425
Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln		
1430	1435	1440
Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp		
1445	1450	1455
Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn		
1460	1465	1470
Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His		
1475	1480	1485
Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met		
1490	1495	1500
Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser		
1505	1510	1515
Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr		
1520	1525	1530
Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr		
1535	1540	1545
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly		
1550	1555	1560
Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly		
1565	1570	1575
Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro		
1580	1585	1590
Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala		
1595	1600	1605
Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His		
1610	1615	1620
Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser		
1625	1630	1635
Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile		
1640	1645	1650

Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser		
1655	1660	1665
Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr		
1670	1675	1680
Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn		
1685	1690	1695
Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile		
1700	1705	1710
Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg		
1715	1720	1725
Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys		
1730	1735	1740
Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln		
1745	1750	1755
Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser		
1760	1765	1770
Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp		
1775	1780	1785
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe		
1790	1795	1800
Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys		
1805	1810	1815
Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser		
1820	1825	1830
Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys		
1835	1840	1845
Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val		
1850	1855	1860
Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His		
1865	1870	1875
Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu		
1880	1885	1890
Gly Cys Glu Ala Gln Asp Leu Tyr		
1895	1900	
<210>	118	
<211>	1515	
<212>	PRT	
<213>	人工序列	
<220>		

<223> pSYN VWF 031蛋白序列

<400> 118

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
Cys Gln Glu Arg Cys Val Asp Gly Cys Ser Cys Pro Glu Gly Gln Leu		320
	325	330
Leu Asp Glu Gly Leu Cys Val Glu Ser Thr Glu Cys Pro Cys Val His		335
	340	345
Ser Gly Lys Arg Tyr Pro Pro Gly Thr Ser Leu Ser Arg Asp Cys Asn		350
	355	360
Thr Cys Ile Cys Arg Asn Ser Gln Trp Ile Cys Ser Asn Glu Glu Cys		365
370	375	380
Pro Gly Glu Cys Leu Val Thr Gly Gln Ser His Phe Lys Ser Phe Asp		
385	390	395
Asn Arg Tyr Phe Thr Phe Ser Gly Ile Cys Gln Tyr Leu Leu Ala Arg		400
	405	410
Asp Cys Gln Asp His Ser Phe Ser Ile Val Ile Glu Thr Val Gln Cys		415
	420	425
Ala Asp Asp Arg Asp Ala Val Cys Thr Arg Ser Val Thr Val Arg Leu		430
	435	440
Pro Gly Leu His Asn Ser Leu Val Lys Leu Lys His Gly Ala Gly Val		445
450	455	460
Ala Met Asp Gly Gln Asp Ile Gln Leu Pro Leu Leu Lys Gly Asp Leu		
465	470	475
Arg Ile Gln His Thr Val Thr Ala Ser Val Arg Leu Ser Tyr Gly Glu		480
	485	490
Asp Leu Gln Met Asp Trp Asp Gly Arg Gly Arg Leu Leu Val Lys Leu		495
	500	505
Ser Pro Val Tyr Ala Gly Lys Thr Cys Gly Leu Cys Gly Asn Tyr Asn		510
	515	520
Gly Asn Gln Gly Asp Asp Phe Leu Thr Pro Ser Gly Leu Ala Glu Pro		525
530	535	540
Arg Val Glu Asp Phe Gly Asn Ala Trp Lys Leu His Gly Asp Cys Gln		
545	550	555
Asp Leu Gln Lys Gln His Ser Asp Pro Cys Ala Leu Asn Pro Arg Met		560
	565	570
Thr Arg Phe Ser Glu Glu Ala Cys Ala Val Leu Thr Ser Pro Thr Phe		575
	580	585
Glu Ala Cys His Arg Ala Val Ser Pro Leu Pro Tyr Leu Arg Asn Cys		590
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 Val Glu Ser Gly Arg	
945	950	955
Tyr Ile Ile Leu Leu Leu Gly Lys Ala Leu Ser Val Val Trp Asp Arg		960
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		
1010	1015	1020
Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro		
1025	1030	1035
Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln		
1040	1045	1050
Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe		
1055	1060	1065
Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val		
1070	1075	1080
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala		
1085	1090	1095
Ala Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln		
1100	1105	1110
His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln		
1115	1120	1125
Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Ala Glu		
1130	1135	1140
Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln		
1145	1150	1155
His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu Gly Cys		
1160	1165	1170
His Ala His Cys Pro Pro Gly Lys Ile Leu Asp Glu Leu Leu Gln		
1175	1180	1185
Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu Val Ala Gly		
1190	1195	1200
Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro Ser Asp		
1205	1210	1215

Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn Leu Thr		
1220	1225	1230
Cys Glu Ala Cys Gln Glu Pro Ile Ser Gly Gly Gly Gly Ser Gly		
1235	1240	1245
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly		
1250	1255	1260
Gly Gly Gly Ser Gly Gly Gly Gly Ser Leu Val Pro Arg Gly Ser		
1265	1270	1275
Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Asp Lys Thr His Thr		
1280	1285	1290
Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val		
1295	1300	1305
Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg		
1310	1315	1320
Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp		
1325	1330	1335
Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His		
1340	1345	1350
Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr		
1355	1360	1365
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn		
1370	1375	1380
Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala		
1385	1390	1395
Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu		
1400	1405	1410
Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys		
1415	1420	1425
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser		
1430	1435	1440
Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn		
1445	1450	1455
Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe		
1460	1465	1470
Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly		
1475	1480	1485
Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His		
1490	1495	1500
Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		

1505	1510	1515
<210> 119		
<211> 1778		
<212> PRT		
<213> 人工序列		
<220>		
<223> pSYN VWF 034蛋白序列		
<400> 119		
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
Thr	Arg	Phe	Ser	Glu	Glu	Ala	Cys	Ala	Val	Leu	Thr	Ser	Pro	Thr	Phe
Glu	Ala	Cys	His	Arg	Ala	Val	Ser	Pro	Leu	Pro	Tyr	Leu	Arg	Asn	Cys
Arg	Tyr	Asp	Val	Cys	Ser	Cys	Ser	Asp	Gly	Arg	Glu	Cys	Leu	Cys	Gly
Ala	Leu	Ala	Ser	Tyr	Ala	Ala	Ala	Cys	Ala	Gly	Arg	Gly	Val	Arg	Val
Ala	Trp	Arg	Glu	Pro	Gly	Arg	Cys	Glu	Leu	Asn	Cys	Pro	Lys	Gly	Gln
Val	Tyr	Leu	Gln	Cys	Gly	Thr	Pro	Cys	Asn	Leu	Thr	Cys	Arg	Ser	Leu
Ser	Tyr	Pro	Asp	Glu	Glu	Cys	Asn	Glu	Ala	Cys	Leu	Glu	Gly	Cys	Phe
Cys	Pro	Pro	Gly	Leu	Tyr	Met	Asp	Glu	Arg	Gly	Asp	Cys	Val	Pro	Lys
Ala	Gln	Cys	Pro	Cys	Tyr	Tyr	Asp	Gly	Glu	Ile	Phe	Gln	Pro	Glu	Asp
Ile	Phe	Ser	Asp	His	His	Thr	Met	Cys	Tyr	Cys	Glu	Asp	Gly	Phe	Met
His	Cys	Thr	Met	Ser	Gly	Val	Pro	Gly	Ser	Leu	Leu	Pro	Asp	Ala	Val
Leu	Ser	Ser	Pro	Leu	Ser	His	Arg	Ser	Lys	Arg	Ser	Leu	Ser	Cys	Arg
Pro	Pro	Met	Val	Lys	Leu	Val	Cys	Pro	Ala	Asp	Asn	Leu	Arg	Ala	Glu
Gly	Leu	Glu	Cys	Thr	Lys	Thr	Cys	Gln	Asn	Tyr	Asp	Leu	Glu	Cys	Met
Ser	Met	Gly	Cys	Val	Ser	Gly	Cys	Leu	Cys	Pro	Pro	Gly	Met	Val	Arg
His	Glu	Asn	Arg	Cys	Val	Ala	Leu	Glu	Arg	Cys	Pro	Cys	Phe	His	Gln
Gly	Lys	Glu	Tyr	Ala	Pro	Gly	Glu	Thr	Val	Lys	Ile	Gly	Cys	Asn	Thr
Cys	Val	Cys	Arg	Asp	Arg	Lys	Trp	Asn	Cys	Thr	Asp	His	Val	Cys	Asp
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			
	1010	1015	1020
Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro			
	1025	1030	1035
Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln			
	1040	1045	1050
Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe			
	1055	1060	1065
Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val			
	1070	1075	1080
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala			
	1085	1090	1095
Ala Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln			
	1100	1105	1110
His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln			
	1115	1120	1125
Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Ala Glu			
	1130	1135	1140
Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln			
	1145	1150	1155
His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu Gly Cys			
	1160	1165	1170

His	Ala	His	Cys	Pro	Pro	Gly	Lys	Ile	Leu	Asp	Glu	Leu	Leu	Gln
1175						1180					1185			
Thr	Cys	Val	Asp	Pro	Glu	Asp	Cys	Pro	Val	Cys	Glu	Val	Ala	Gly
1190						1195					1200			
Arg	Arg	Phe	Ala	Ser	Gly	Lys	Lys	Val	Thr	Leu	Asn	Pro	Ser	Asp
1205						1210					1215			
Pro	Glu	His	Cys	Gln	Ile	Cys	His	Cys	Asp	Val	Val	Asn	Leu	Thr
1220						1225					1230			
Cys	Glu	Ala	Cys	Gln	Glu	Pro	Ile	Ser	Gly	Thr	Ser	Glu	Ser	Ala
1235						1240					1245			
Thr	Pro	Glu	Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser
1250						1255					1260			
Glu	Thr	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro
1265						1270					1275			
Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Thr	Ser
1280						1285					1290			
Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Thr	Glu	Pro
1295						1300					1305			
Ser	Glu	Gly	Ser	Ala	Pro	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser
1310						1315					1320			
Thr	Glu	Glu	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro
1325						1330					1335			
Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Thr	Ser
1340						1345					1350			
Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Ser	Pro	Ala	Gly	Ser
1355						1360					1365			
Pro	Thr	Ser	Thr	Glu	Glu	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser
1370						1375					1380			
Thr	Glu	Glu	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro
1385						1390					1395			
Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser
1400						1405					1410			
Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Glu	Ser	Ala
1415						1420					1425			
Thr	Pro	Glu	Ser	Gly	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser
1430						1435					1440			
Glu	Thr	Pro	Gly	Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro
1445						1450					1455			
Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser	Thr	Glu	Glu	Gly	Thr	Ser

1460	1465	1470
Thr Glu Pro Ser Glu Gly	Ser Ala Pro Gly Thr	Ser Thr Glu Pro
1475	1480	1485
Ser Glu Gly Ser Ala Pro	Gly Ser Glu Pro Ala	Thr Ser Gly Ser
1490	1495	1500
Glu Thr Pro Gly Thr Ser	Glu Ser Ala Thr Pro	Glu Ser Gly Pro
1505	1510	1515
Gly Thr Ser Thr Glu Pro	Ser Glu Gly Ser Ala	Pro Asp Ile Gly
1520	1525	1530
Gly Gly Gly Gly Ser Gly	Gly Gly Gly Ser Leu	Val Pro Arg Gly
1535	1540	1545
Ser Gly Gly Asp Lys Thr	His Thr Cys Pro Pro	Cys Pro Ala Pro
1550	1555	1560
Glu Leu Leu Gly Gly Pro	Ser Val Phe Leu Phe	Pro Pro Lys Pro
1565	1570	1575
Lys Asp Thr Leu Met Ile	Ser Arg Thr Pro Glu	Val Thr Cys Val
1580	1585	1590
Val Val Asp Val Ser His	Glu Asp Pro Glu Val	Lys Phe Asn Trp
1595	1600	1605
Tyr Val Asp Gly Val Glu	Val His Asn Ala Lys	Thr Lys Pro Arg
1610	1615	1620
Glu Glu Gln Tyr Asn Ser	Thr Tyr Arg Val Val	Ser Val Leu Thr
1625	1630	1635
Val Leu His Gln Asp Trp	Leu Asn Gly Lys Glu	Tyr Lys Cys Lys
1640	1645	1650
Val Ser Asn Lys Ala Leu	Pro Ala Pro Ile Glu	Lys Thr Ile Ser
1655	1660	1665
Lys Ala Lys Gly Gln Pro	Arg Glu Pro Gln Val	Tyr Thr Leu Pro
1670	1675	1680
Pro Ser Arg Asp Glu Leu	Thr Lys Asn Gln Val	Ser Leu Thr Cys
1685	1690	1695
Leu Val Lys Gly Phe Tyr	Pro Ser Asp Ile Ala	Val Glu Trp Glu
1700	1705	1710
Ser Asn Gly Gln Pro Glu	Asn Asn Tyr Lys Thr	Thr Pro Pro Val
1715	1720	1725
Leu Asp Ser Asp Gly Ser	Phe Phe Leu Tyr Ser	Lys Leu Thr Val
1730	1735	1740
Asp Lys Ser Arg Trp Gln	Gln Gly Asn Val Phe	Ser Cys Ser Val
1745	1750	1755

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 1760 1765 1770
 Leu Ser Pro Gly Lys
 1775
 <210> 120
 <211> 1565
 <212> PRT
 <213> 人工序列
 <220>
 <223> pSYN VWF 036 蛋白序列
 <400> 120
 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
Asp Leu Gln Lys Gln His Ser Asp Pro Cys Ala Leu Asn Pro Arg Met		560
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
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
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
Ser Met Gly Cys Val Ser Gly Cys Leu Cys Pro Pro Gly Met Val Arg		800
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		
	1010	1015 1020
Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro		
	1025	1030 1035
Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln		
	1040	1045 1050
Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe		
	1055	1060 1065
Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val		
	1070	1075 1080
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala		
	1085	1090 1095
Ala Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln		
	1100	1105 1110
His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln		
	1115	1120 1125
Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Ala Glu		
	1130	1135 1140
Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln		

1145	1150	1155
His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu Gly Cys		
1160	1165	1170
His Ala His Cys Pro Pro Gly Lys Ile Leu Asp Glu Leu Leu Gln		
1175	1180	1185
Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu Val Ala Gly		
1190	1195	1200
Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro Ser Asp		
1205	1210	1215
Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn Leu Thr		
1220	1225	1230
Cys Glu Ala Cys Gln Glu Pro Ile Ser Gly Gly Gly Gly Ser Gly		
1235	1240	1245
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly		
1250	1255	1260
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly		
1265	1270	1275
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly		
1280	1285	1290
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly		
1295	1300	1305
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Leu		
1310	1315	1320
Val Pro Arg Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser		
1325	1330	1335
Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu		
1340	1345	1350
Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr		
1355	1360	1365
Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp		
1370	1375	1380
Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp		
1385	1390	1395
Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln		
1400	1405	1410
Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His		
1415	1420	1425
Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn		
1430	1435	1440

Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys
 1445 1450 1455
 Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg
 1460 1465 1470
 Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys
 1475 1480 1485
 Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
 1490 1495 1500
 Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 1505 1510 1515
 Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
 1520 1525 1530
 Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu
 1535 1540 1545
 Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
 1550 1555 1560

Gly Lys

1565

<210> 121

<211> 247

<212> PRT

<213> 人工序列

<220>

<223> pSYN Fc-015 蛋白序列

<400> 121

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

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu
 115 120 125
 Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
 130 135 140
 Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys
 145 150 155 160
 Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
 165 170 175
 Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
 180 185 190
 Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
 195 200 205
 Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser
 210 215 220
 Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
 225 230 235 240
 Leu Ser Leu Ser Pro Gly Lys
 245

<210> 122

<211> 33

<212> PRT

<213> 人工序列

<220>

<223> 多肽

<400> 122

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
 1 5 10 15
 Gly Gly Gly Ser Gly Gly Gly Gly Ser Leu Val Pro Arg Gly Ser Gly
 20 25 30
 Gly

<210> 123

<211> 1240

<212> PRT

<213> 人工序列

<220>

<223> VWF-D1D2D'D3

<400> 123

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
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
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
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			
1010	1015	1020	
Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg Lys Val Pro			
1025	1030	1035	
Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met Lys Gln			
1040	1045	1050	
Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val Phe			
1055	1060	1065	
Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val			
1070	1075	1080	
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala			
1085	1090	1095	
Cys Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln			
1100	1105	1110	
His Gly Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln			
1115	1120	1125	
Ser Cys Glu Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Cys Glu			
1130	1135	1140	
Trp Arg Tyr Asn Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln			
1145	1150	1155	
His Pro Glu Pro Leu Ala Cys Pro Val Gln Cys Val Glu Gly Cys			
1160	1165	1170	
His Ala His Cys Pro Pro Gly Lys Ile Leu Asp Glu Leu Leu Gln			
1175	1180	1185	
Thr Cys Val Asp Pro Glu Asp Cys Pro Val Cys Glu Val Ala Gly			
1190	1195	1200	
Arg Arg Phe Ala Ser Gly Lys Lys Val Thr Leu Asn Pro Ser Asp			
1205	1210	1215	
Pro Glu His Cys Gln Ile Cys His Cys Asp Val Val Asn Leu Thr			
1220	1225	1230	
Cys Glu Ala Cys Gln Glu Pro			
1235	1240		

<210> 124

<211> 477

<212> PRT

<213> 人工序列

<220>

<223> VWF-D'D3

<400> 124

Ser	Leu	Ser	Cys	Arg	Pro	Pro	Met	Val	Lys	Leu	Val	Cys	Pro	Ala	Asp
1				5					10					15	
Asn	Leu	Arg	Ala	Glu	Gly	Leu	Glu	Cys	Thr	Lys	Thr	Cys	Gln	Asn	Tyr
			20					25					30		
Asp	Leu	Glu	Cys	Met	Ser	Met	Gly	Cys	Val	Ser	Gly	Cys	Leu	Cys	Pro
		35					40					45			
Pro	Gly	Met	Val	Arg	His	Glu	Asn	Arg	Cys	Val	Ala	Leu	Glu	Arg	Cys
		50				55					60				
Pro	Cys	Phe	His	Gln	Gly	Lys	Glu	Tyr	Ala	Pro	Gly	Glu	Thr	Val	Lys
65				70					75					80	
Ile	Gly	Cys	Asn	Thr	Cys	Val	Cys	Arg	Asp	Arg	Lys	Trp	Asn	Cys	Thr
			85					90					95		
Asp	His	Val	Cys	Asp	Ala	Thr	Cys	Ser	Thr	Ile	Gly	Met	Ala	His	Tyr
		100						105					110		
Leu	Thr	Phe	Asp	Gly	Leu	Lys	Tyr	Leu	Phe	Pro	Gly	Glu	Cys	Gln	Tyr
		115					120					125			
Val	Leu	Val	Gln	Asp	Tyr	Cys	Gly	Ser	Asn	Pro	Gly	Thr	Phe	Arg	Ile
		130					135					140			
Leu	Val	Gly	Asn	Lys	Gly	Cys	Ser	His	Pro	Ser	Val	Lys	Cys	Lys	Lys
145				150					155					160	
Arg	Val	Thr	Ile	Leu	Val	Glu	Gly	Gly	Glu	Ile	Glu	Leu	Phe	Asp	Gly
			165					170					175		
Glu	Val	Asn	Val	Lys	Arg	Pro	Met	Lys	Asp	Glu	Thr	His	Phe	Glu	Val
		180						185					190		
Val	Glu	Ser	Gly	Arg	Tyr	Ile	Ile	Leu	Leu	Leu	Gly	Lys	Ala	Leu	Ser
		195					200					205			
Val	Val	Trp	Asp	Arg	His	Leu	Ser	Ile	Ser	Val	Val	Leu	Lys	Gln	Thr
		210				215					220				
Tyr	Gln	Glu	Lys	Val	Cys	Gly	Leu	Cys	Gly	Asn	Phe	Asp	Gly	Ile	Gln
225				230					235					240	
Asn	Asn	Asp	Leu	Thr	Ser	Ser	Asn	Leu	Gln	Val	Glu	Glu	Asp	Pro	Val
			245					250					255		

Asp Phe Gly Asn Ser Trp Lys Val Ser Ser Gln Cys Ala Asp Thr Arg		
260	265	270
Lys Val Pro Leu Asp Ser Ser Pro Ala Thr Cys His Asn Asn Ile Met		
275	280	285
Lys Gln Thr Met Val Asp Ser Ser Cys Arg Ile Leu Thr Ser Asp Val		
290	295	300
Phe Gln Asp Cys Asn Lys Leu Val Asp Pro Glu Pro Tyr Leu Asp Val		
305	310	315
Cys Ile Tyr Asp Thr Cys Ser Cys Glu Ser Ile Gly Asp Cys Ala Cys		
325	330	335
Phe Cys Asp Thr Ile Ala Ala Tyr Ala His Val Cys Ala Gln His Gly		
340	345	350
Lys Val Val Thr Trp Arg Thr Ala Thr Leu Cys Pro Gln Ser Cys Glu		
355	360	365
Glu Arg Asn Leu Arg Glu Asn Gly Tyr Glu Cys Glu Trp Arg Tyr Asn		
370	375	380
Ser Cys Ala Pro Ala Cys Gln Val Thr Cys Gln His Pro Glu Pro Leu		
385	390	395
Ala Cys Pro Val Gln Cys Val Glu Gly Cys His Ala His Cys Pro Pro		
405	410	415
Gly Lys Ile Leu Asp Glu Leu Leu Gln Thr Cys Val Asp Pro Glu Asp		
420	425	430
Cys Pro Val Cys Glu Val Ala Gly Arg Arg Phe Ala Ser Gly Lys Lys		
435	440	445
Val Thr Leu Asn Pro Ser Asp Pro Glu His Cys Gln Ile Cys His Cys		
450	455	460
Asp Val Val Asn Leu Thr Cys Glu Ala Cys Gln Glu Pro		
465	470	475

<210> 125

<211> 5055

<212> DNA

<213> 人工序列

<220>

<223> pSYNFVIII 010

<400> 125

```

atgcaaatag agctctccac ctgcttcttt ctgtgccttt tgcgattctg ctttagtgcc 60
accagaagat actacctggg tgcagtggaa ctgtcatggg actatatgca aagtgatctc 120
ggtagactgc ctgtggacgc aagatttcct cctagagtgc caaaatcttt tccattcaac 180
acctcagtcg tgtacaaaaa gactctgttt gtagaattca cggatcacct tttcaacatc 240

```

gctaagccaa ggccaccctg gatgggtctg ctaggtccta ccatccaggc tgaggtttat 300
gatacagtgg tcattacact taagaacatg gcttcccatc ctgtcagtct tcatgtctgtt 360
gggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtcaaagg 420
gagaaagaag atgataaagt cttccctggg ggaagccata catatgtctg gcaggtcctg 480
aaagagaatg gtccaatggc ctctgaccca ctgtgcctta cctactcata tctttctcat 540
gtggacctgg taaaagactt gaattcaggc ctattggag ccctactagt atgtagagaa 600
gggagtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta 660
tttgatgaag ggaaaagt gactcagaa acaagaact ccttgatgca ggatagggat 720
gctgcatctg ctggggcctg gcctaaaatg cacacagtca atggttatgt aaacaggtct 780
ctgccaggtc tgattggatg ccacaggaaa tcagtctatt ggcatgtgat tggaatgggc 840
accactcctg aagtgcactc aatattcctc gaaggtcaca ctttcttgt gaggaacct 900
cgccaggcgt ccttggaat ctgccaata actttcctta ctgctcaaac actcttgatg 960
gaccttgga agtttctact gttttgtcat atctcttccc accaactga tggcatggaa 1020
gcttatgtca aagtagacag ctgtccagag gaacccaac tacgaatgaa aaataatgaa 1080
gaagcggaag actatgatga tgatcttact gattctgaaa tggatgtggg caggtttgat 1140
gatgacaact ctcttcctt tatccaaatt cgctcagttg ccaagaagca tcctaaaact 1200
tggttacatt acattgctgc tgaagaggag gactgggact atgctccctt agtctcgc 1260
cccgatgaca gaagtataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg 1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa cctttaagac tcgtgaagct 1380
attcagcatg aatcaggaat ctgggacct ttactttatg gggaagttgg agacacactg 1440
ttgattatat ttaagaatca agcaagcaga ccatataaca tctaccctca cggaatcact 1500
gatgtccgtc ctttgtattc aaggagatta ccaaagggtg taaaacattt gaaggatttt 1560
ccaattctgc caggagaaat attcaaatat aaatggacag tgactgtaga agatgggcca 1620
actaaatcag atcctcgggt cctgaccgcg tattactcta gtttcgttaa tatggagaga 1680
gatctagctt caggactcat tggccctctc ctcatctgct acaaagaatc tgtagatcaa 1740
agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag 1800
aaccgaagct ggtacctcac agagaatata caacgcttc tcccaatcc agctggagtg 1860
cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatgtt 1920
tttgatagtt tgcagttgtc agtttgtttg catgaggtgg cactactgga cattctaagc 1980
attggagcac agactgactt ctttctgtc ttcttctctg gatatacctt caaacacaaa 2040
atgggtctatg aagacacact caccctattc ccattctcag gagaaactgt cttcatgtcg 2100
atggaaaacc caggtctatg gattctgggg tgccacaact cagactttcg gaacagaggc 2160
atgaccgcct tactgaaggt ttctagttgt gacaagaaca ctggtgatta ttacaggagc 2220
agttatgaag atatttcagc atacttgtg agtaaaaaca atgccattga accaagaagc 2280
ttctctcaaa acccaccagt cttgaaacgc catcaacggg aaataactcg tactactctt 2340
cagtcagatc aagaggaaat tgactatgat gataccatat cagttgaaat gaagaaggaa 2400
gattttgaca tttatgatga ggatgaaaat cagagcccc gcagctttca aaagaaaaca 2460
cgacactatt ttattgctgc agtgagagg ctctgggatt atgggatgag tagctcccca 2520
catgttctaa gaaacagggc tcagagtggc agtgtccctc agttcaagaa agttgttttc 2580

caggaattta ctgatggctc ctttactcag cccttatacc gtggagaact aaatgaacat 2640
ttgggactcc tggggccata tataagagca gaagttgaag ataatatcat ggtaactttc 2700
agaaatcagg cctctcgtcc ctattccttc tattctagcc ttattttctta tgaggaagat 2760
cagaggcaag gagcagaacc tagaaaaaac tttgtcaagc ctaatgaaac caaaacttac 2820
ttttggaaaag tgcaacatca tatggcaccc actaaagatg agtttgactg caaagcctgg 2880
gcttattttct ctgatgttga cctggaaaaa gatgtgcact caggcctgat tggacccctt 2940
ctggctctgcc acactaacac actgaaccct gctcatggga gacaagtac agtacaggaa 3000
tttgctctgt ttttcacat ctttgatgag accaaaagct ggtacttcac tgaaaatatg 3060
gaaagaaact gcagggtcc ctgcaatata cagatggaag atcccacttt taaagagaat 3120
tatcgcttcc atgcaatcaa tggctacata atggatacac tacctggctt agtaatggct 3180
caggatcaaa ggattcgatg gtatctgctc agcatgggca gcaatgaaaa catccattct 3240
attcatttca gtggacatgt gttcactgta cgaaaaaag aggagtataa aatggcactg 3300
tacaatctct atccagggtg ttttgagaca gtggaaatgt taccatccaa agctggaatt 3360
tggcgggtgg aatgccttat tggcgagcat ctacatgctg ggatgagcac actttttctg 3420
gtgtacagca ataagtgtca gactcccctg ggaatggctt ctggacacat tagagatttt 3480
cagattacag cttcaggaca atatggacag tgggccccaa agctggccag acttcattat 3540
tccgatcaa tcaatgcctg gagcaccaag gagccctttt cttggatcaa ggtggatctg 3600
ttggcaccaa tgattattca cggcatcaag acccagggtg cccgtcagaa gttctccagc 3660
ctctacatct ctcagtttat catcatgtat agtcttgatg ggaagaagtg gcagacttat 3720
cgaggaaatt ccactggaac cttaatggtc ttctttggca atgtggattc atctgggata 3780
aaacacaata tttttaacct tccaattatt gctcgataca tccgtttgca cccaactcat 3840
tatagcattc gcagcactct tcgcatggag ttgatgggt gtgatttaaa tagttgcagc 3900
atgccattgg gaatggagag taaagcaata tcagatgcac agattactgc ttcactctac 3960
tttaccaata tgtttgccac ctggtctcct tcaaaagctc gacttcacct ccaagggagg 4020
agtaatgcct ggagacctca ggtgaataat ccaaaagagt ggctgcaagt ggacttcag 4080
aagacaatga aagtcacagg agtaactact caggagtaa aatctctgct taccagcatg 4140
tatgtgaagg agttcctcat ctccagcagt caagatggcc atcagtggac tctctttttt 4200
cagaatggca aagtaaaggt ttttcaggga aatcaagact cttcacacc tgttgtgaac 4260
tctctagacc caccgttact gactcgctac cttcgaattc acccccagag ttgggtgcac 4320
cagattgccc tgaggatgga ggttctgggc tgcgaggcac aggacctcta cgacaaaact 4380
cacacatgcc caccgtgccc agtccagaa ctctgggag gaccgtcagt cttcctcttc 4440
ccccaaaaac ccaaggacac cctcatgate tcccggacce ctgaggtcac atgcgtggtg 4500
gtggacgtga gccacgaaga ccctgaggtc aagttcaact ggtacgtgga cggcgtggag 4560
gtgcataatg ccaagacaaa gccgcgggag gagcagtaca acagcacgta ccgtgtggtc 4620
agcgtcctca ccgtcctgca ccaggactgg ctgaatggca aggagtacaa gtgcaaggtc 4680
tccaacaaag ccctcccagc ccccatcgag aaaaccatct ccaaagccaa agggcagccc 4740
cgagaaccac aggtgtacac cctgccccca tcccgggatg agctgaccaa gaaccaggtc 4800
agcctgacct gcctgtgcaa aggtttctat cccagcgaca tcgccgtgga gtgggagagc 4860
aatgggcagc cggagaacaa ctacaagacc acgcctcccg tgttgactc cgacggctcc 4920

ttcttcctct acagcaagct caccgtggac aagagcaggt ggcagcaggg gaacgtcttc 4980
tcatgctccg tgatgcatga ggctctgcac aaccactaca cgcagaagag cctctccctg 5040
tctccgggta aatga 5055

<210> 126

<211> 1684

<212> PRT

<213> 人工序列

<220>

<223> pSYNFVIII 010

<400> 126

Met	Gln	Ile	Glu	Leu	Ser	Thr	Cys	Phe	Phe	Leu	Cys	Leu	Leu	Arg	Phe
1				5					10					15	
Cys	Phe	Ser	Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser
			20					25					30		
Trp	Asp	Tyr	Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg
			35				40						45		
Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val
			50				55				60				
Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile
65				70					75					80	
Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln
				85				90						95	
Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
			100					105					110		
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
			115				120						125		
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
			130				135					140			
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
145				150						155				160	
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
				165					170					175	
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
			180					185					190		
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
			195				200						205		
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
			210				215						220		
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp

225	230	235	240
Ala Ala Ser Ala Arg	Ala Trp Pro Lys Met	His Thr Val Asn Gly Tyr	
	245	250	255
Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val			
	260	265	270
Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile			
	275	280	285
Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser			
	290	295	300
Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met			
305	310	315	320
Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His			
	325	330	335
Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro			
	340	345	350
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp			
	355	360	365
Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser			
	370	375	380
Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr			
385	390	395	400
Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro			
	405	410	415
Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn			
	420	425	430
Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met			
	435	440	445
Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu			
	450	455	460
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu			
465	470	475	480
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro			
	485	490	495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys			
	500	505	510
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe			
	515	520	525
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp			
	530	535	540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg			
545	550	555	560
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu			
	565	570	575
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val			
	580	585	590
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu			
	595	600	605
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp			
	610	615	620
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val			
	625	630	635
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp			
	645	650	655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe			
	660	665	670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr			
	675	680	685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro			
	690	695	700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly			
	705	710	715
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp			
	725	730	735
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys			
	740	745	750
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu			
	755	760	765
Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln			
	770	775	780
Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu			
	785	790	795
Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe			
	805	810	815
Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp			
	820	825	830
Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln			
	835	840	845
Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr			

850	855	860
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His		
865	870	875
Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile		880
885	890	895
Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser		
900	905	910
Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg		
915	920	925
Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val		
930	935	940
Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp		
945	950	955
Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu		960
965	970	975
Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His		
980	985	990
Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe		
995	1000	1005
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn		
1010	1015	1020
Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys		
1025	1030	1035
Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr		
1040	1045	1050
Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr		
1055	1060	1065
Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe		
1070	1075	1080
Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met		
1085	1090	1095
Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met		
1100	1105	1110
Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly		
1115	1120	1125
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1130	1135	1140
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1145	1150	1155

Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1160	1165	1170
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1175	1180	1185
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1190	1195	1200
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1205	1210	1215
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1220	1225	1230
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1235	1240	1245
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1250	1255	1260
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1265	1270	1275
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1280	1285	1290
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1295	1300	1305
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1310	1315	1320
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1325	1330	1335
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1340	1345	1350
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1355	1360	1365
Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys		
1370	1375	1380
Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu		
1385	1390	1395
Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp		
1400	1405	1410
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr		
1415	1420	1425
Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala		
1430	1435	1440
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr Asp		

1445	1450	1455
Lys Thr His Thr Cys Pro	Pro Cys Pro Ala Pro	Glu Leu Leu Gly
1460	1465	1470
Gly Pro Ser Val Phe Leu	Phe Pro Pro Lys Pro	Lys Asp Thr Leu
1475	1480	1485
Met Ile Ser Arg Thr Pro	Glu Val Thr Cys Val	Val Val Asp Val
1490	1495	1500
Ser His Glu Asp Pro Glu	Val Lys Phe Asn Trp	Tyr Val Asp Gly
1505	1510	1515
Val Glu Val His Asn Ala	Lys Thr Lys Pro Arg	Glu Glu Gln Tyr
1520	1525	1530
Asn Ser Thr Tyr Arg Val	Val Ser Val Leu Thr	Val Leu His Gln
1535	1540	1545
Asp Trp Leu Asn Gly Lys	Glu Tyr Lys Cys Lys	Val Ser Asn Lys
1550	1555	1560
Ala Leu Pro Ala Pro Ile	Glu Lys Thr Ile Ser	Lys Ala Lys Gly
1565	1570	1575
Gln Pro Arg Glu Pro Gln	Val Tyr Thr Leu Pro	Pro Ser Arg Asp
1580	1585	1590
Glu Leu Thr Lys Asn Gln	Val Ser Leu Thr Cys	Leu Val Lys Gly
1595	1600	1605
Phe Tyr Pro Ser Asp Ile	Ala Val Glu Trp Glu	Ser Asn Gly Gln
1610	1615	1620
Pro Glu Asn Asn Tyr Lys	Thr Thr Pro Pro Val	Leu Asp Ser Asp
1625	1630	1635
Gly Ser Phe Phe Leu Tyr	Ser Lys Leu Thr Val	Asp Lys Ser Arg
1640	1645	1650
Trp Gln Gln Gly Asn Val	Phe Ser Cys Ser Val	Met His Glu Ala
1655	1660	1665
Leu His Asn His Tyr Thr	Gln Lys Ser Leu Ser	Leu Ser Pro Gly
1670	1675	1680

Lys

<210> 127

<211> 78

<212> PRT

<213> 人工序列

<220>

<223> AE72 XTEN

<400> 127

Gly Ala Pro Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser
 1 5 10 15
 Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala
 20 25 30
 Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu
 35 40 45
 Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr
 50 55 60
 Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ala Ser Ser
 65 70 75

<210> 128

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AE144_2A

<400> 128

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly
 1 5 10 15
 Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly
 20 25 30
 Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly
 35 40 45
 Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu
 50 55 60
 Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu
 65 70 75 80
 Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly
 85 90 95
 Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu
 100 105 110
 Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu
 115 120 125
 Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly
 130 135 140

<210> 129

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AE144_3B

<400> 129

```

Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser
1          5          10          15
Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser
          20          25          30
Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly
          35          40          45
Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu
          50          55          60
Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly
65          70          75          80
Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly
          85          90          95
Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu
          100          105          110
Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser
          115          120          125
Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly
          130          135          140

```

<210> 130

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AE144_4A

<400> 130

```

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala
1          5          10          15
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu
          20          25          30
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly
          35          40          45
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu
          50          55          60
Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu
65          70          75          80
Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly

```

	85		90		95
Ser Pro Ala Gly	Ser Pro Thr Ser Thr Glu Glu Gly	Ser Pro Ala Gly			
100	105	110			
Ser Pro Thr Ser Thr Glu Glu Gly	Thr Ser Glu Ser Ala Thr Pro Glu				
115	120	125			
Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly	Ser Ala Pro Gly				
130	135	140			
<210> 131					
<211> 144					
<212> PRT					
<213> 人工序列					
<220>					
<223> AE144_5A					
<400> 131					
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala					
1	5	10	15		
Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu					
20	25	30			
Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly					
35	40	45			
Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu					
50	55	60			
Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser					
65	70	75	80		
Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly					
85	90	95			
Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser					
100	105	110			
Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser					
115	120	125			
Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly					
130	135	140			
<210> 132					
<211> 144					
<212> PRT					
<213> 人工序列					
<220>					
<223> AE144_6B					
<400> 132					

Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Thr	Ser	Glu	Ser
1				5					10					15	
Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu
			20					25					30		
Ser	Gly	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu	Ser	Gly	Pro	Gly
			35					40					45		
Ser	Glu	Pro	Ala	Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Ser	Glu	Pro	Ala
			50					55					60		
Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Ser	Pro	Ala	Gly	Ser	Pro	Thr	Ser
65					70					75					80
Thr	Glu	Glu	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly
					85					90				95	
Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly	Ser	Glu	Pro	Ala
			100						105				110		
Thr	Ser	Gly	Ser	Glu	Thr	Pro	Gly	Thr	Ser	Glu	Ser	Ala	Thr	Pro	Glu
			115						120				125		
Ser	Gly	Pro	Gly	Thr	Ser	Thr	Glu	Pro	Ser	Glu	Gly	Ser	Ala	Pro	Gly
			130						135				140		

<210> 133

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AG144_A

<400> 133

Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro
1				5					10					15	
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr
			20					25					30		
Gly	Thr	Gly	Pro	Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro
			35					40					45		
Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro
			50					55					60		
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser
65					70					75					80
Thr	Gly	Ser	Pro	Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro
					85					90				95	
Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Thr	Pro	Gly
			100						105				110		

Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser
 115 120 125

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro
 130 135 140

<210> 134

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AG144_B

<400> 134

Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr
 1 5 10 15

Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser
 20 25 30

Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro
 35 40 45

Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro
 50 55 60

Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr
 65 70 75 80

Gly Thr Gly Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro
 85 90 95

Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro
 100 105 110

Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser
 115 120 125

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro
 130 135 140

<210> 135

<211> 144

<212> PRT

<213> 人工序列

<220>

<223> AG144_C

<400> 135

Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro
 1 5 10 15

Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser

	20		25		30												
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
	35						40					45					
Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Thr	Pro	Gly		
	50					55					60						
Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
65				70					75						80		
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
				85					90					95			
Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr		
	100						105					110					
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala		
	115						120					125					
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
	130					135						140					
<210>	136																
<211>	144																
<212>	PRT																
<213>	人工序列																
<220>																	
<223>	AG144_F																
<400>	136																
Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ser	Ser	Pro		
1				5				10					15				
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
				20				25				30					
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		
	35						40					45					
Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro		
	50					55					60						
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser		
65				70					75						80		
Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro		
				85					90					95			
Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ser	Ser	Thr		
	100						105					110					
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala		
	115						120					125					
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro		

130	135	140
<210> 137		
<211> 288		
<212> PRT		
<213> 人工序列		
<220>		
<223> AE288_2		
<400> 137		
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu		
1 5 10 15		
Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly		
20 25 30		
Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro		
35 40 45		
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr		
50 55 60		
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu		
65 70 75 80		
Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro		
85 90 95		
Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr		
100 105 110		
Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr		
115 120 125		
Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro		
130 135 140		
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro		
145 150 155 160		
Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro		
165 170 175		
Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro		
180 185 190		
Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr		
195 200 205		
Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro		
210 215 220		
Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu		
225 230 235 240		
Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala		

<400> 140

Ser Gly Gly Gly Gly Ser

1

5

不同的VWF构建体

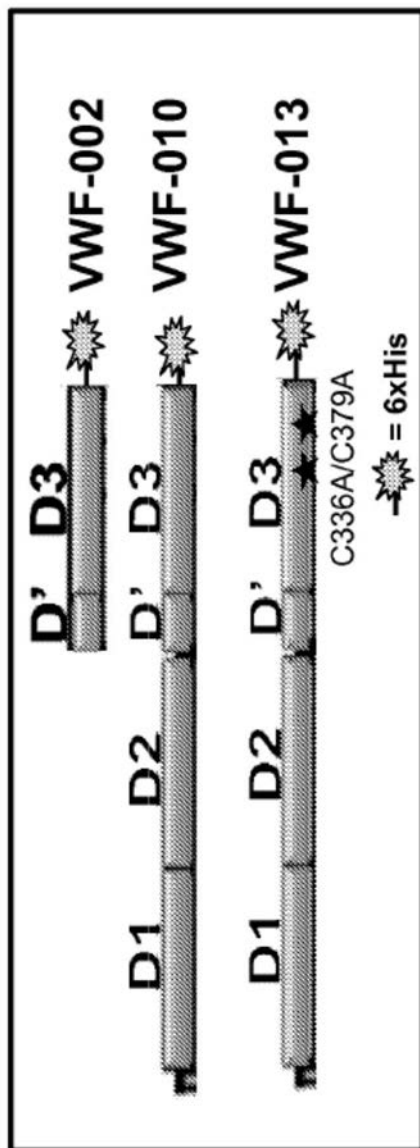


图1A

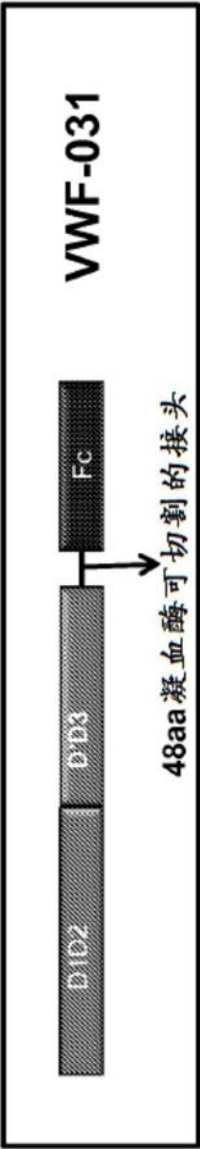


图1B

VWF-025: pLIVE 载体中的D1D2D'D3密码子序列
VWF-029: pLIVE 载体中的D1D2D'D3C_{336A/C379A}

图1C

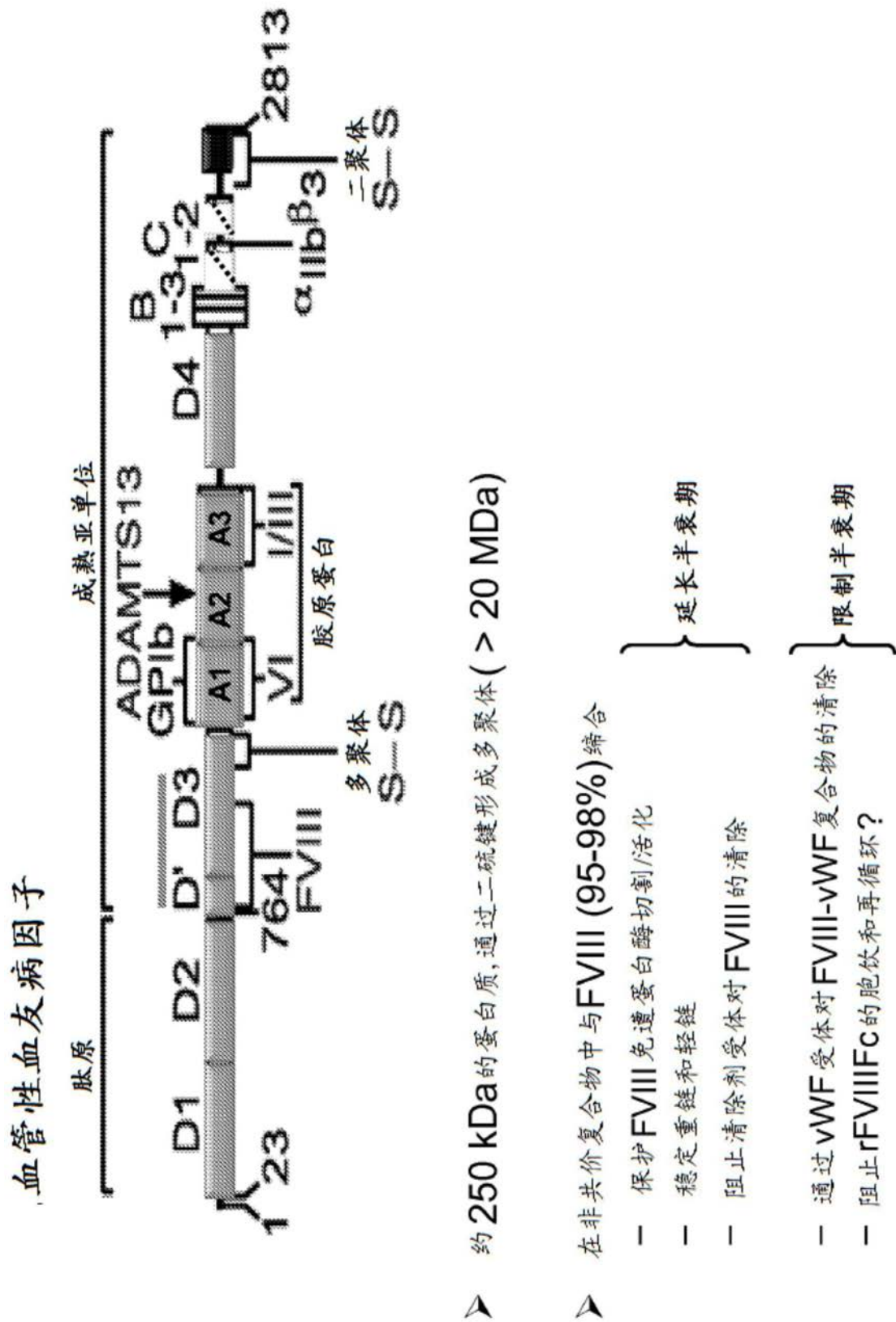
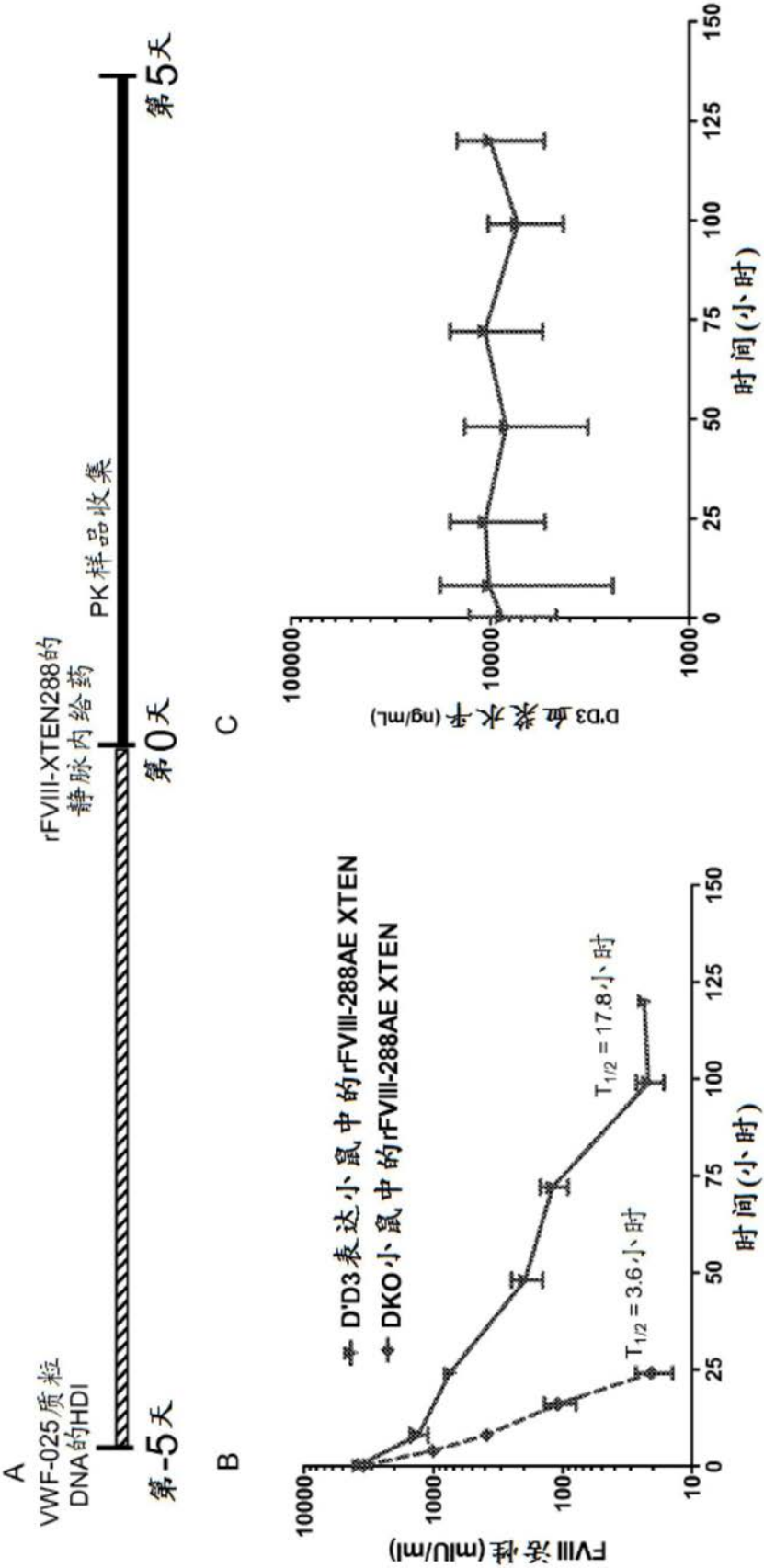


图1D

VWF D'D3表达小鼠中的rFVIII-XTEN PK



由于VWF D'D3片段FVIII-XTEN的5倍附加半衰期延长

图2

FVIII^{FC}/VWF 异二聚体构建体(可变接头)

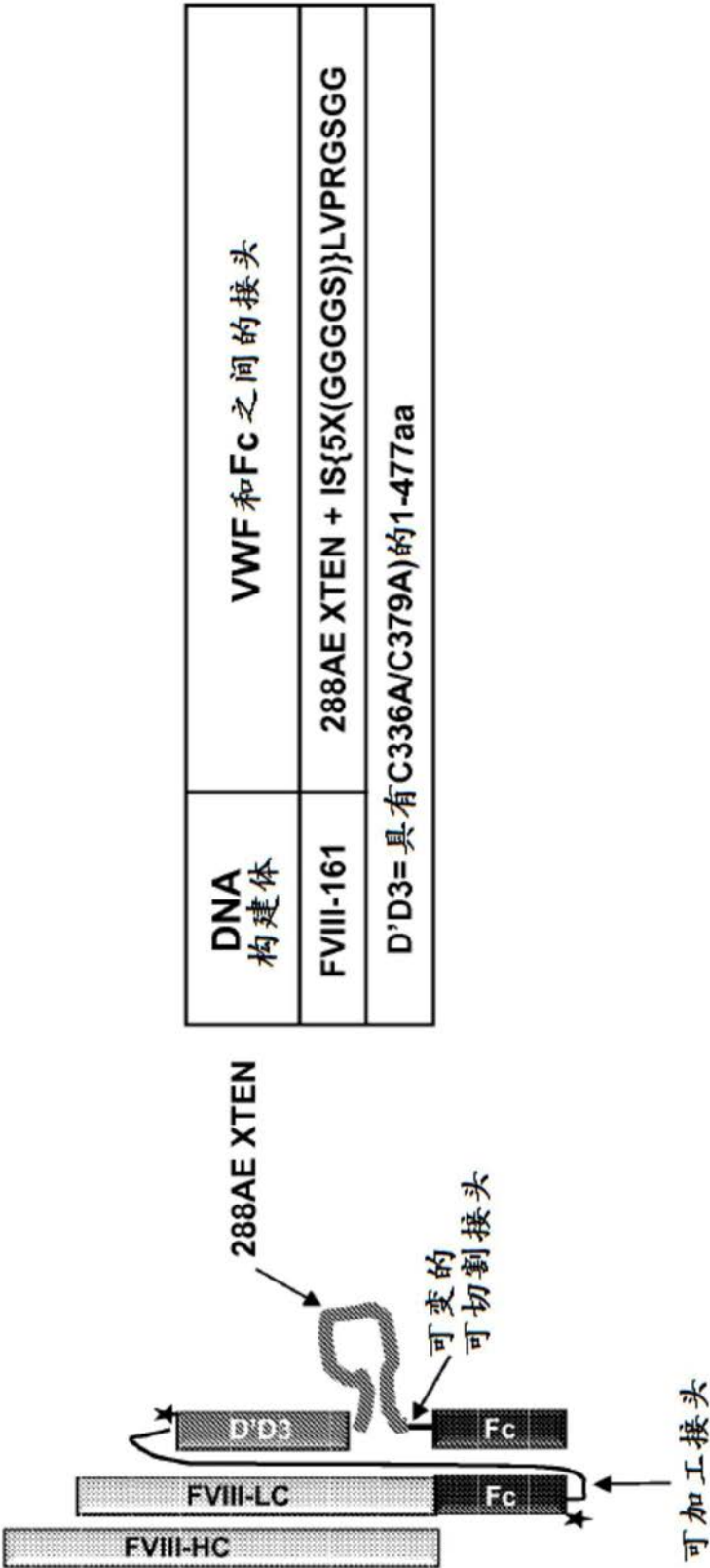


图3

具有XTEN插入的FVIII/VWF构建体

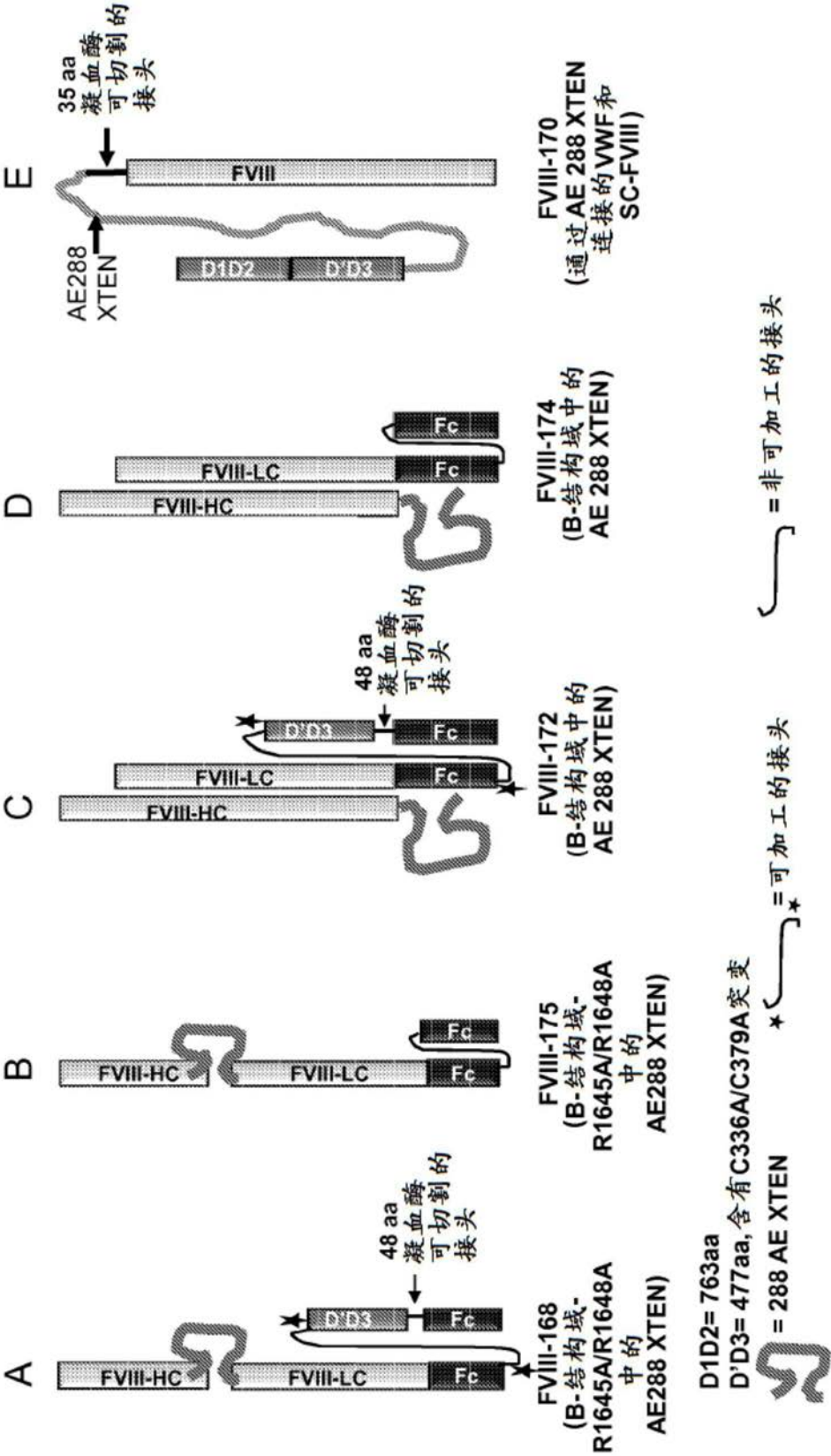


图4

XTEN插入改善FVIII/VWF异二聚体的PK

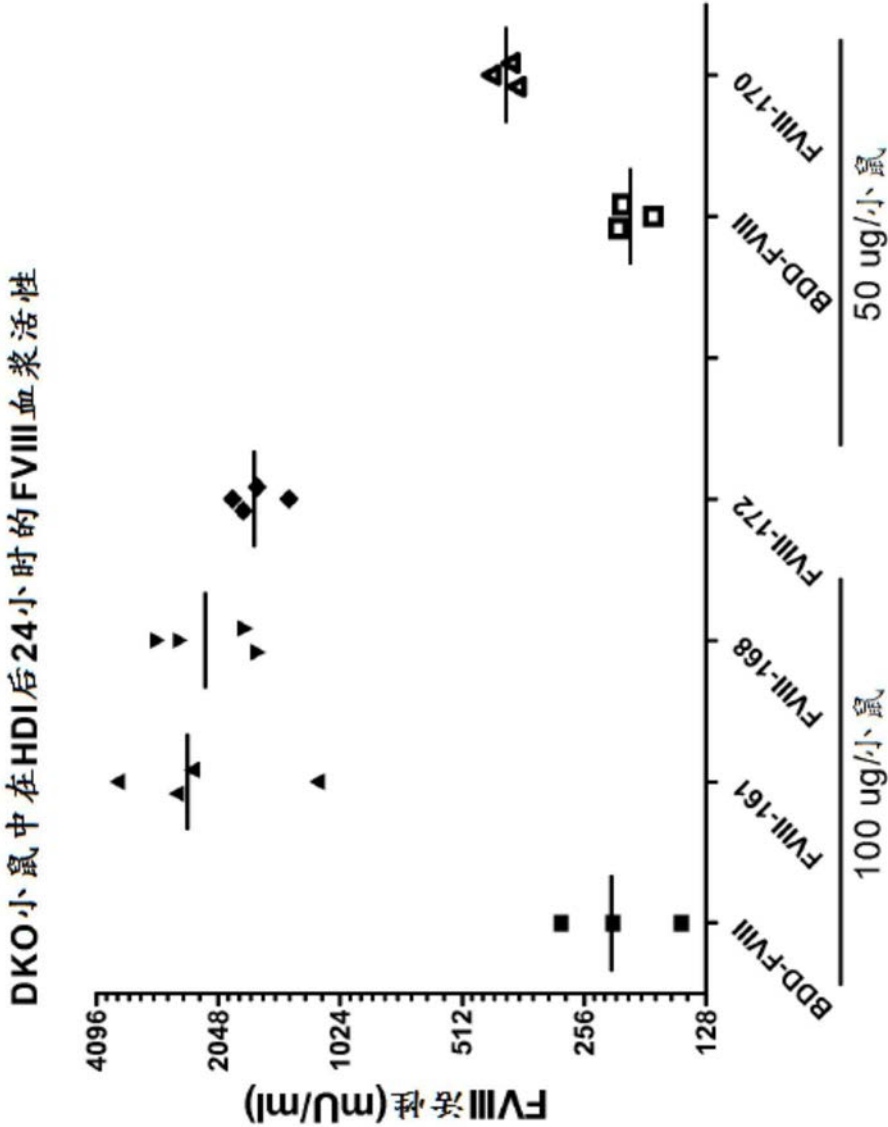


图5

FVIII 构建体(共转染系统)

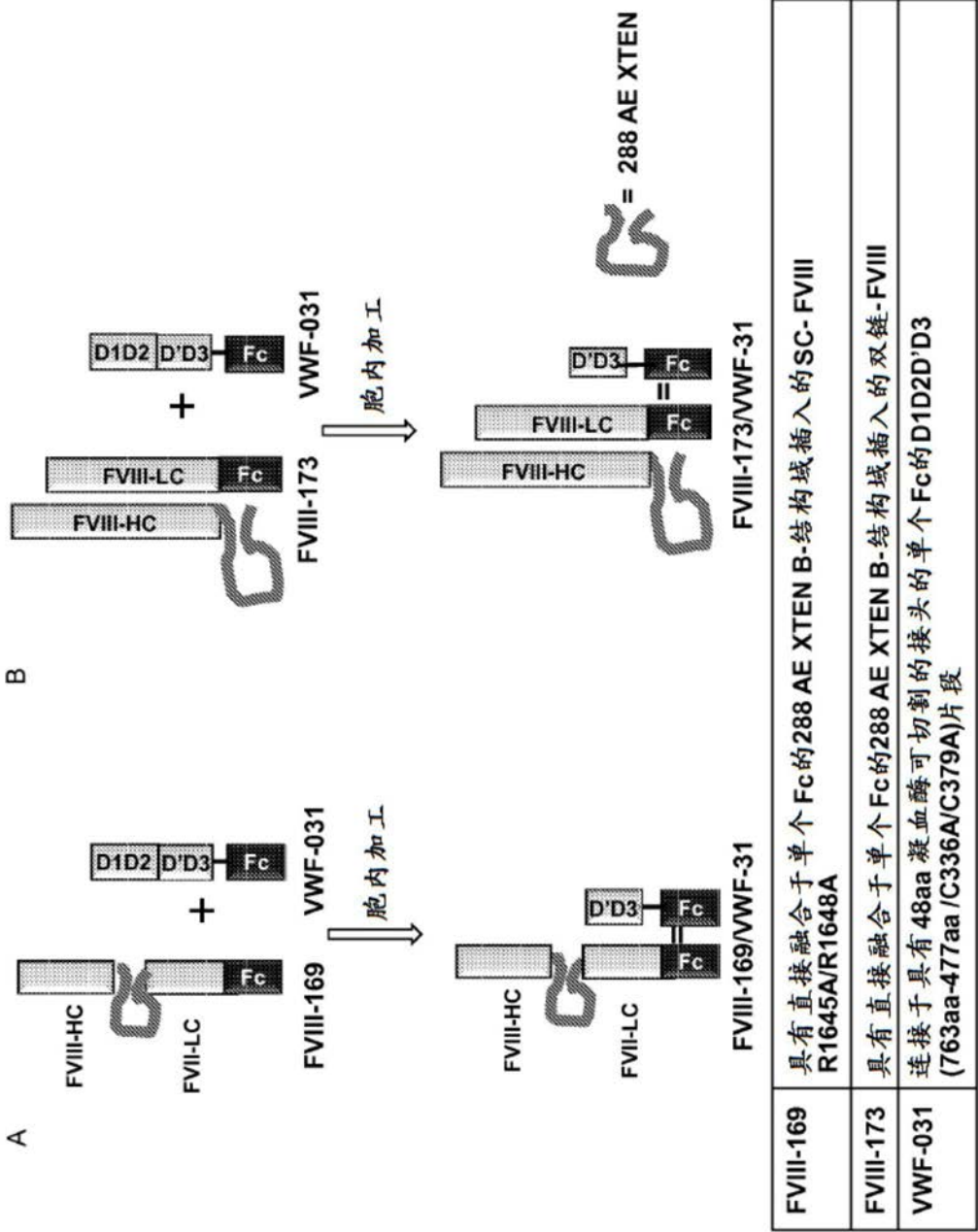


图6

FVIII-169/VWF-31不具有针对固定化hVWF的可检测的结合

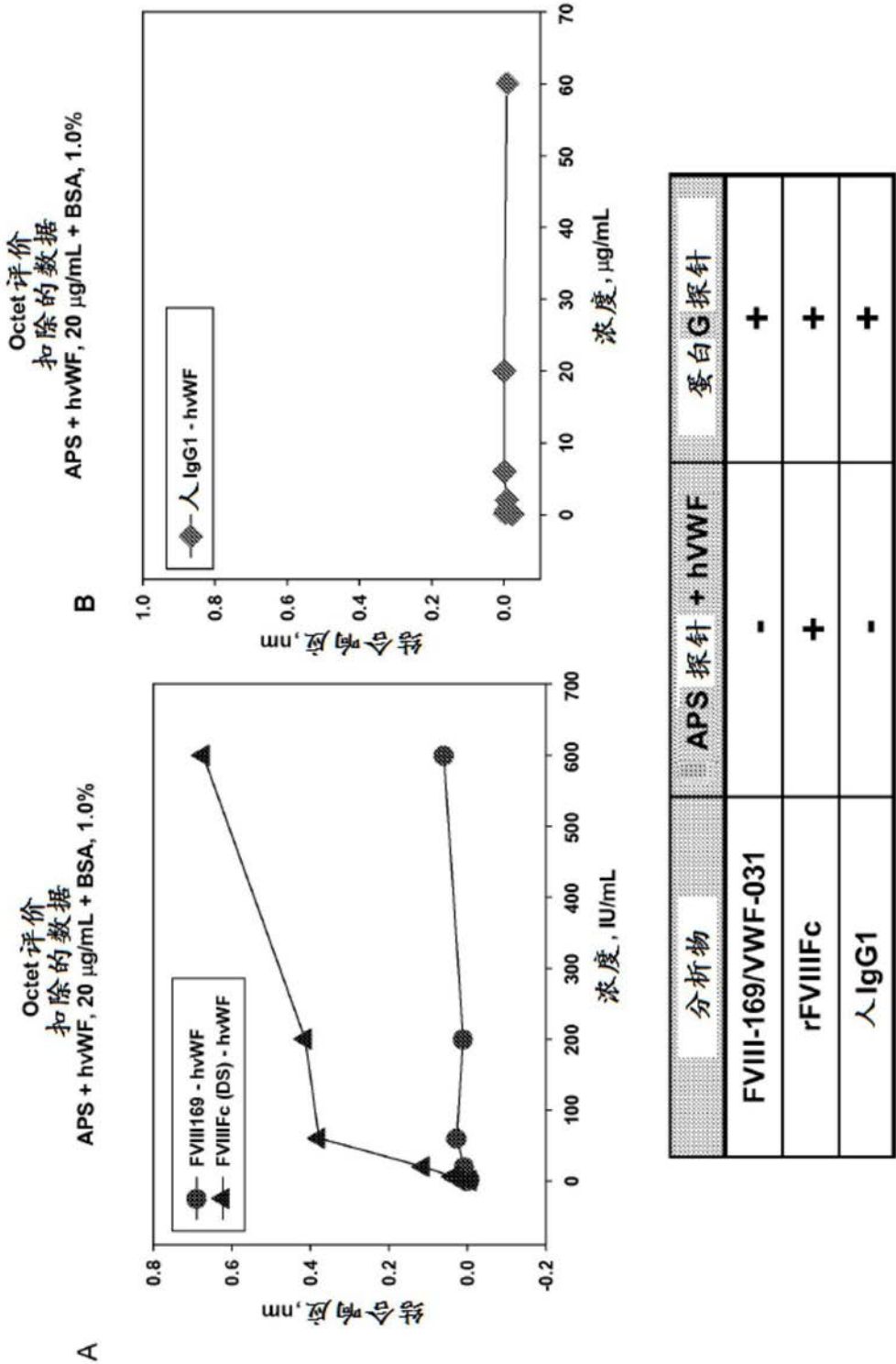


图7

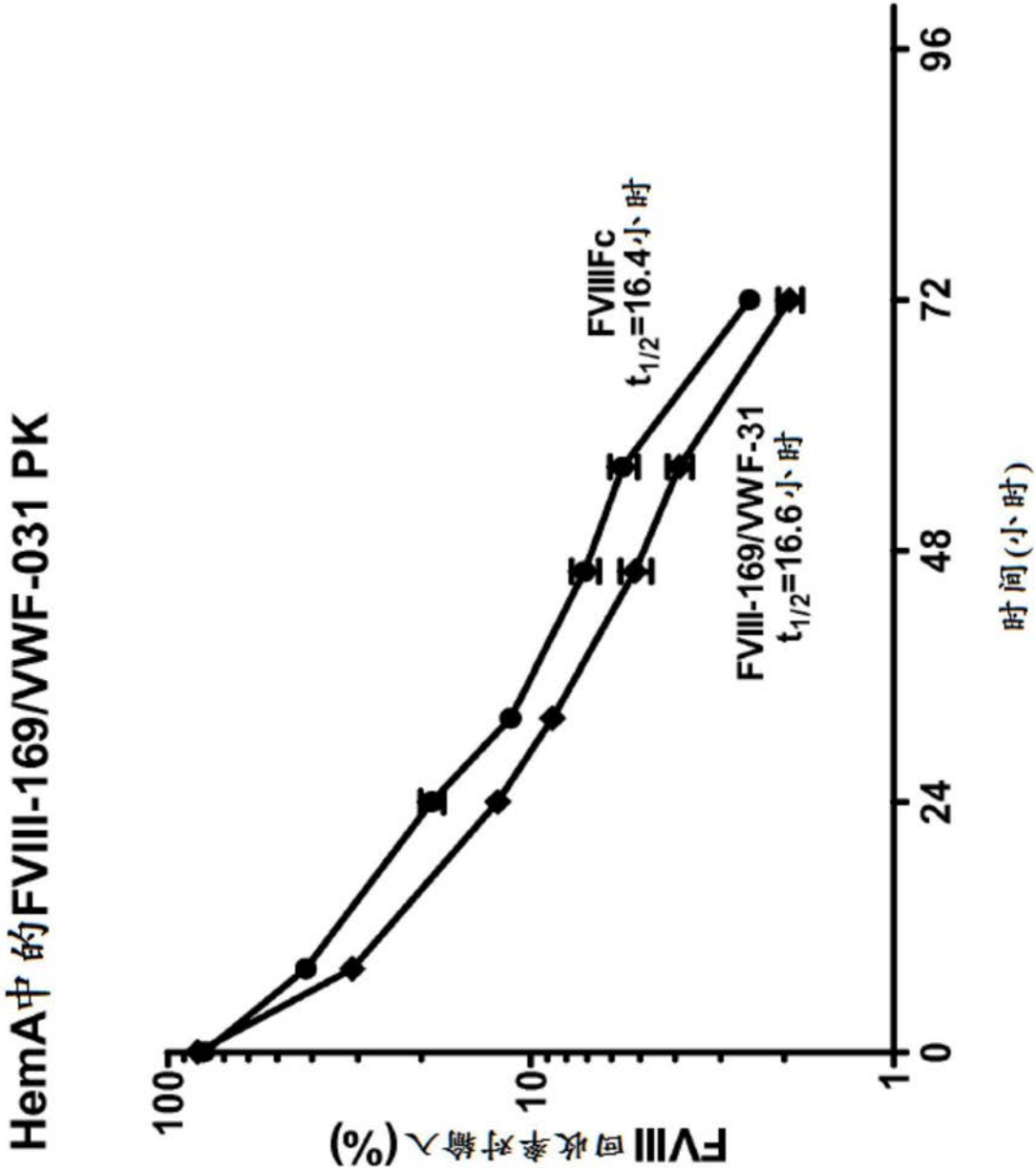


图8A

FVIII/VWF DKO小鼠中的FVIII-169/VWF-031 PK

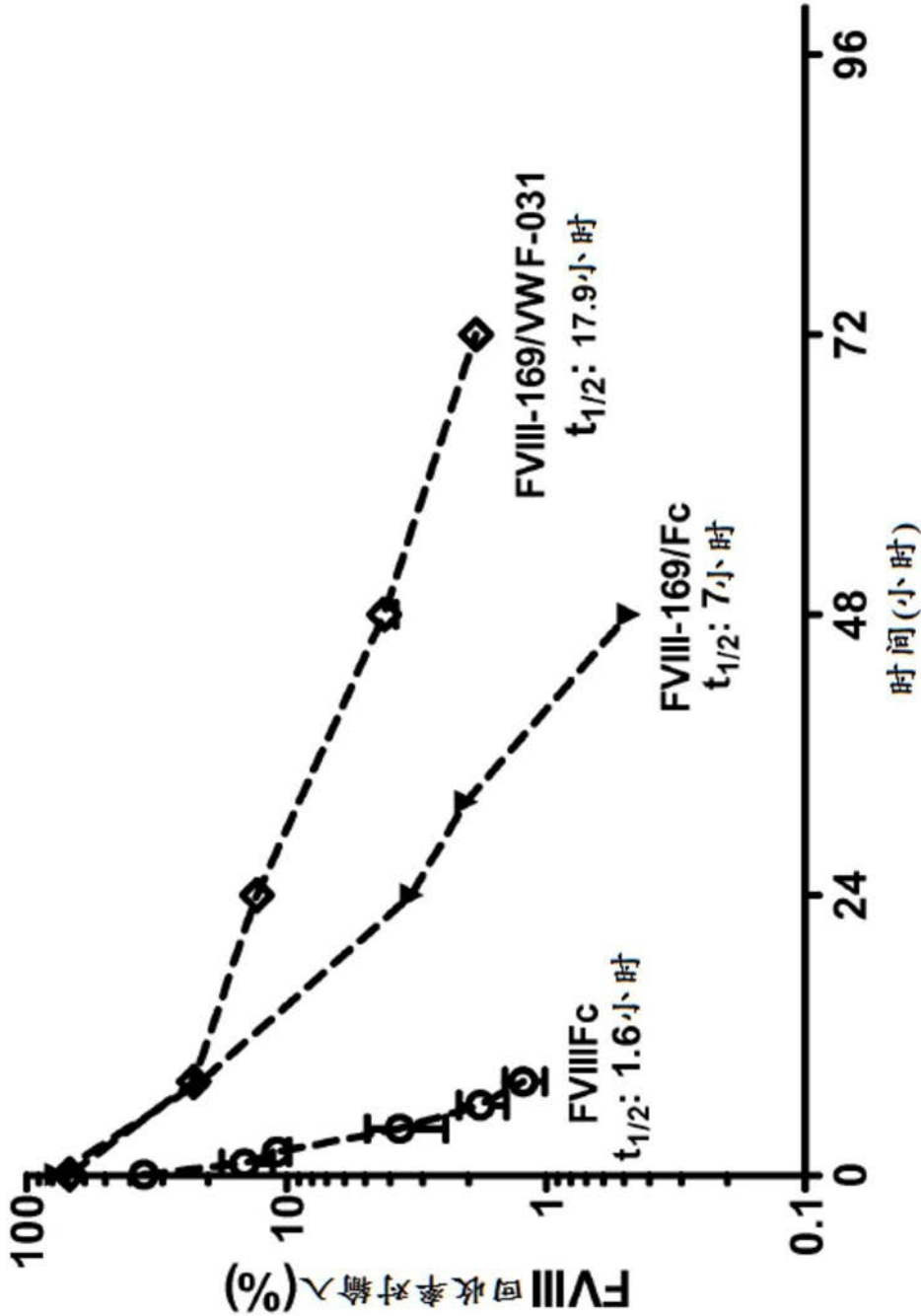


图8B

D'D3表达FVIII/VWF DKO小鼠中的FVIII-XTEN变体半衰期

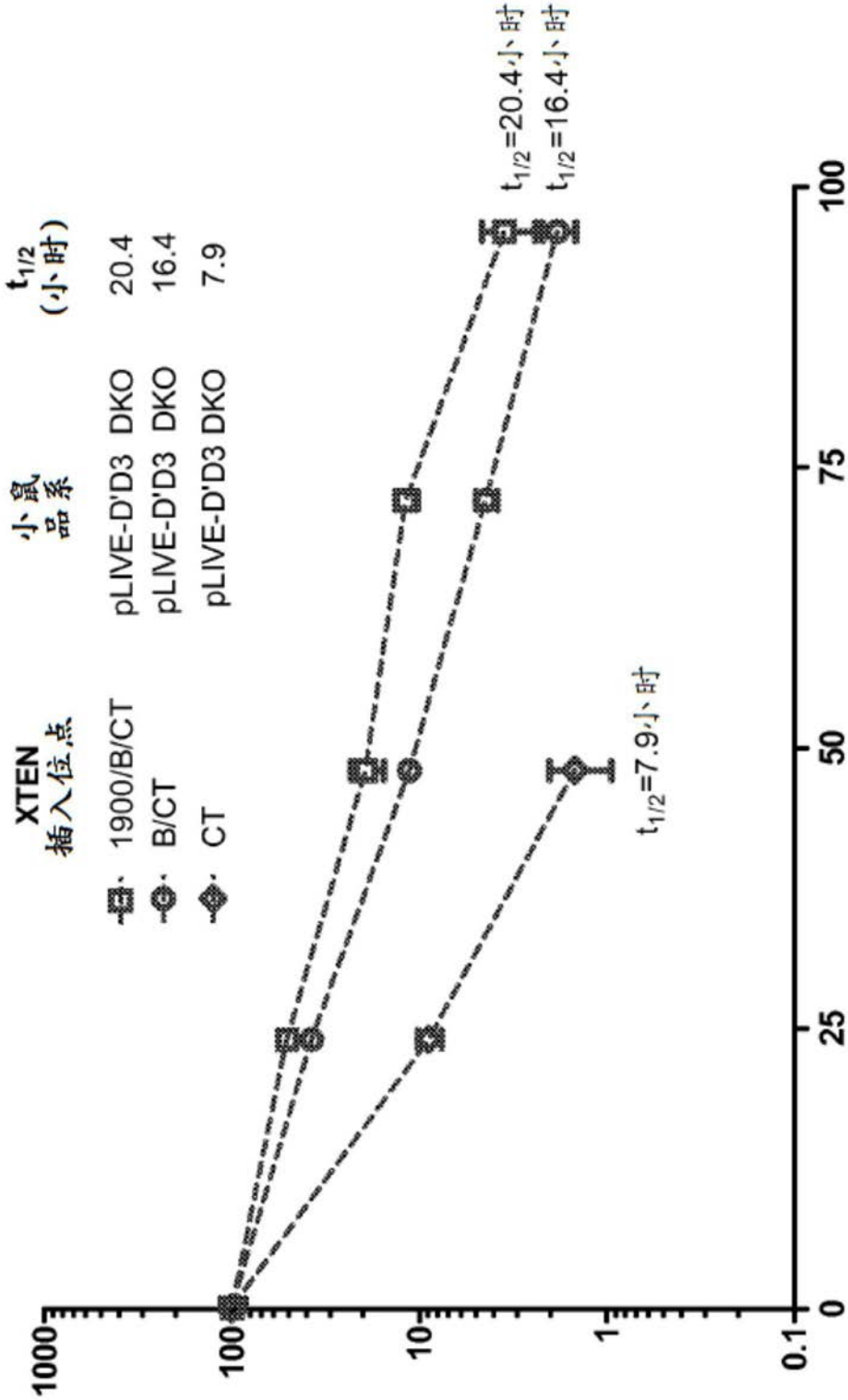


图9A

具有三个XTEN插入的FVIII-XTEN的PK

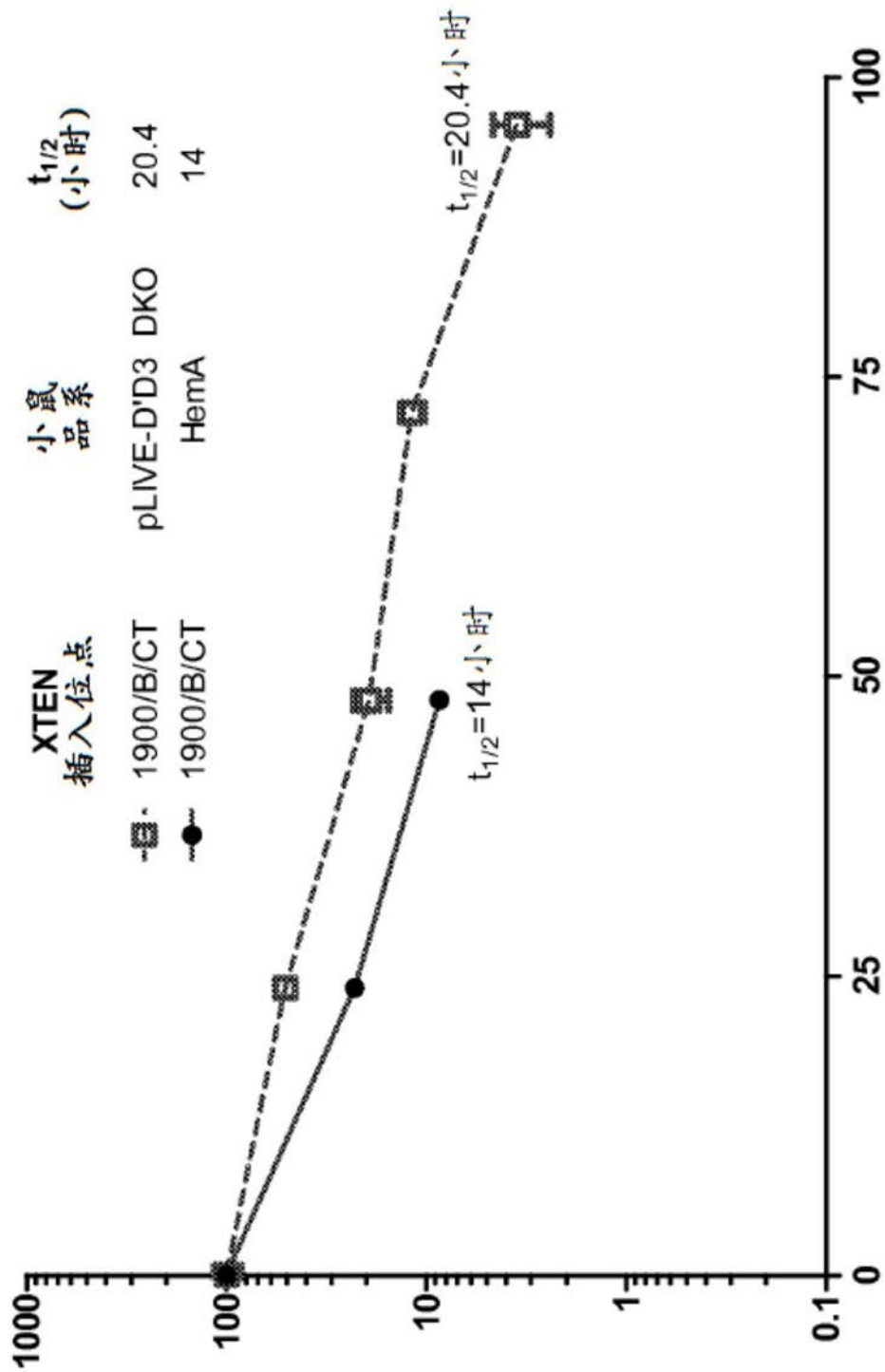


图9B

小鼠DKO血浆中的FVIII Fc活性

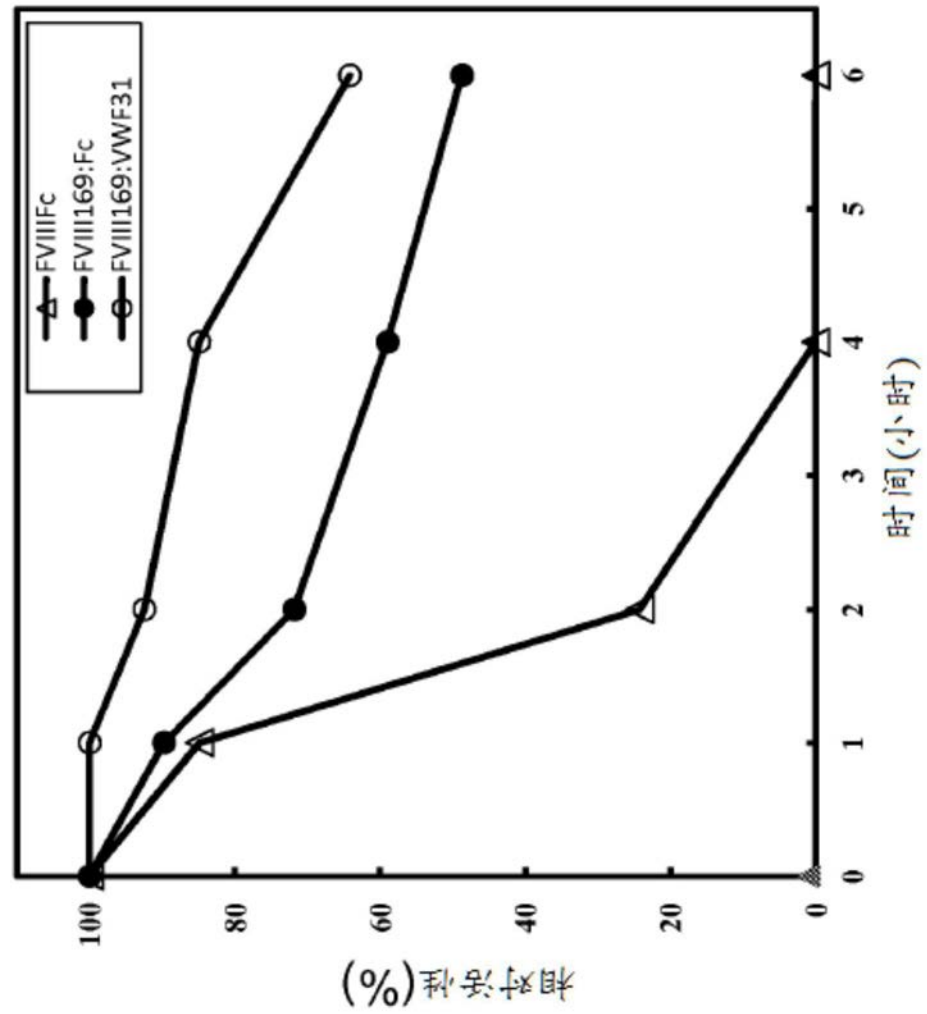


图10

Fc、XTEN和VWF-D'D3片段对FVIII半衰期延长的累加效应

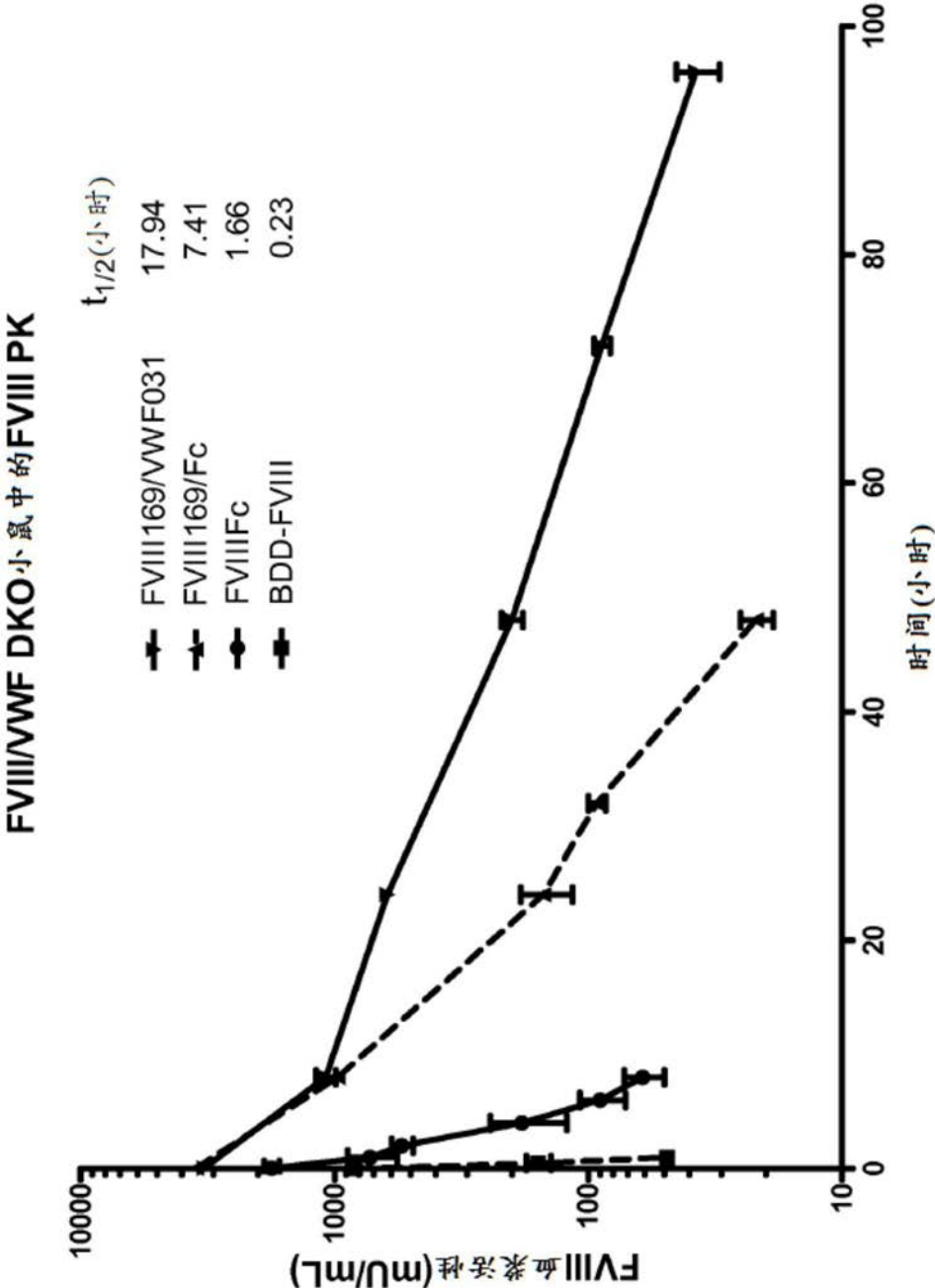


图11

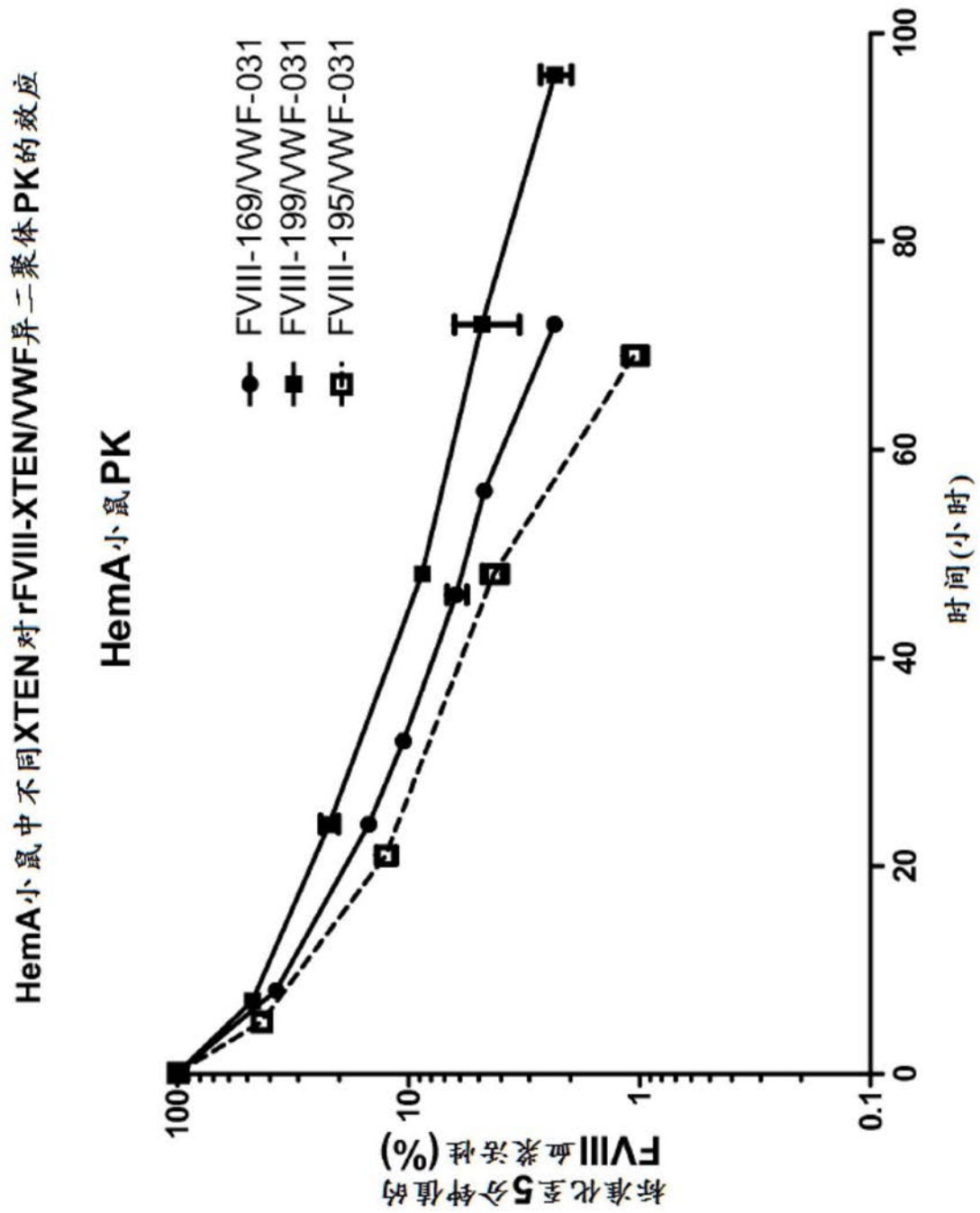


图12A

Hema小鼠中不同XTEN对rFVIII-XTEN/VWF异二聚体PK的效应

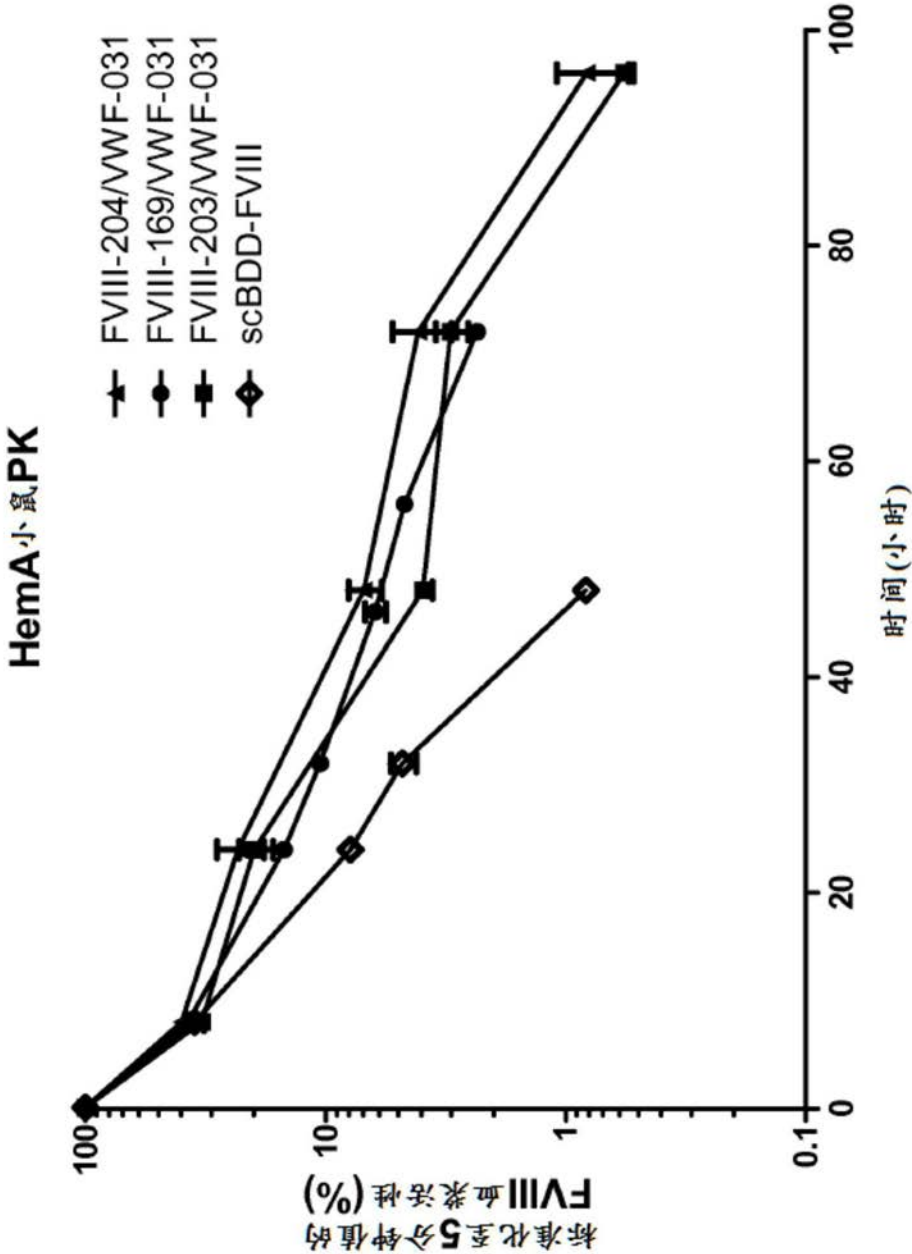


图12B

Hema小鼠中不同XTEN对rFVIII-XTEN/VWF异二聚体PK的效应

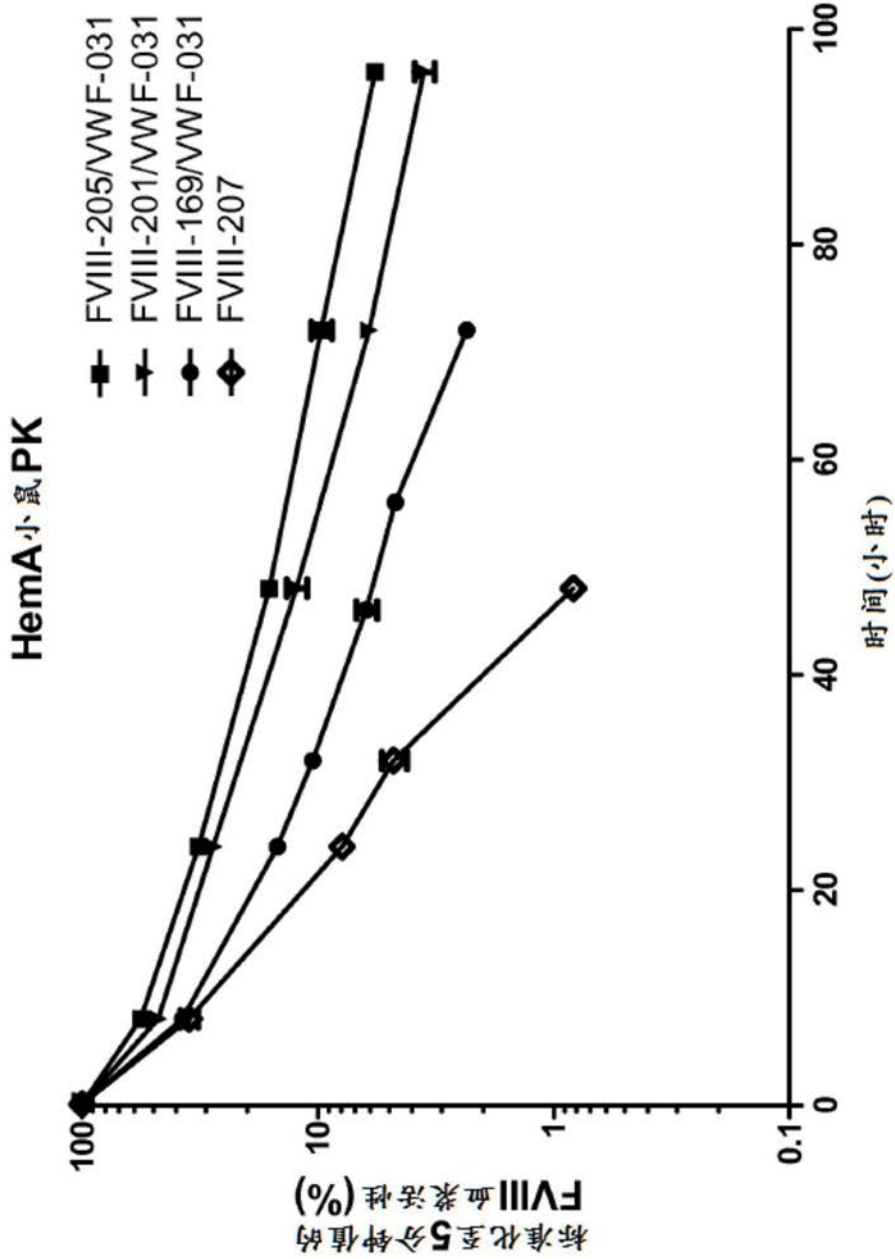


图12C

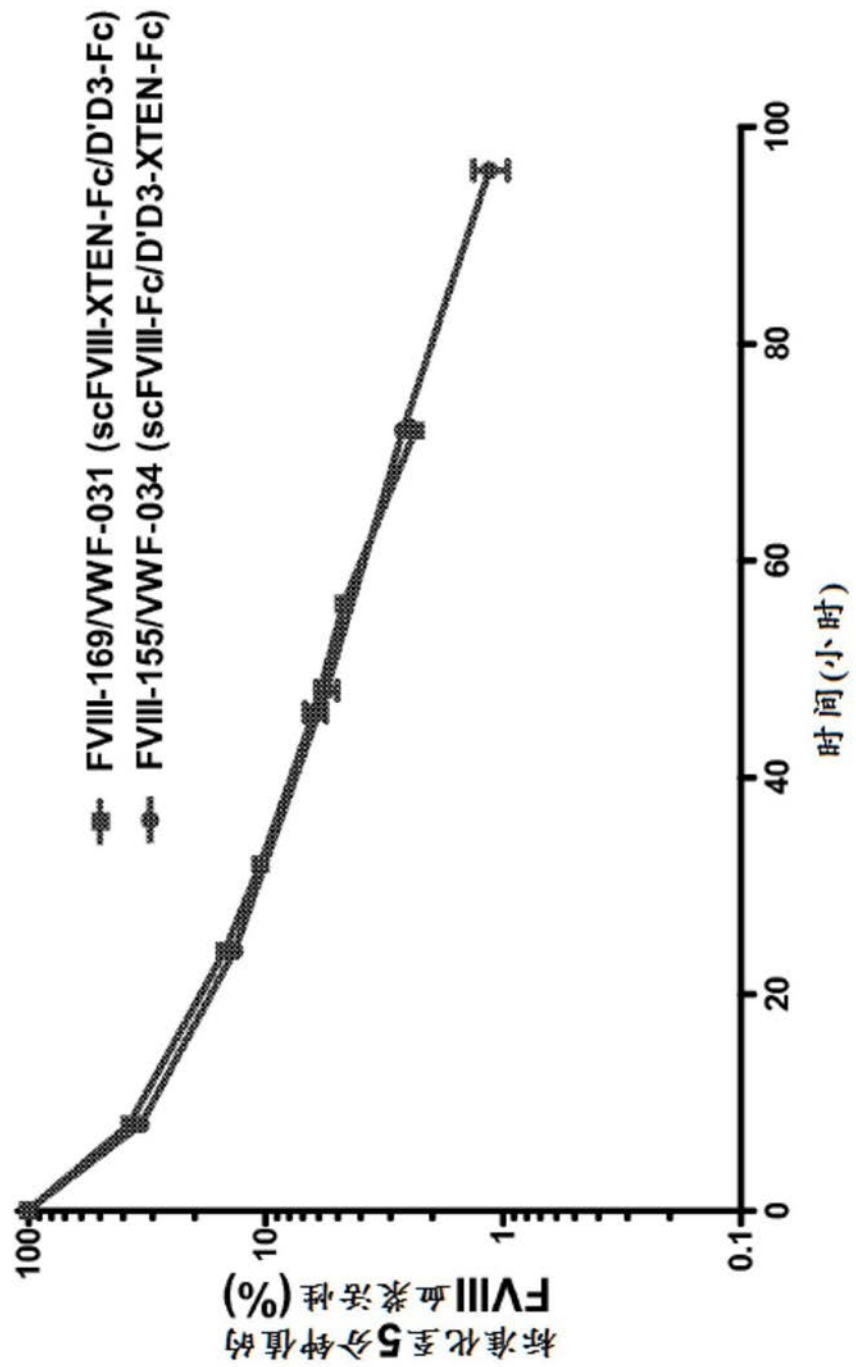
FVIII/VWF DKO小鼠中的rFVIII-XTEN/VWF-XTEN异二聚体PK

图13

HemA小鼠中的rFVIII-XTEN/VWF 异二聚体PK

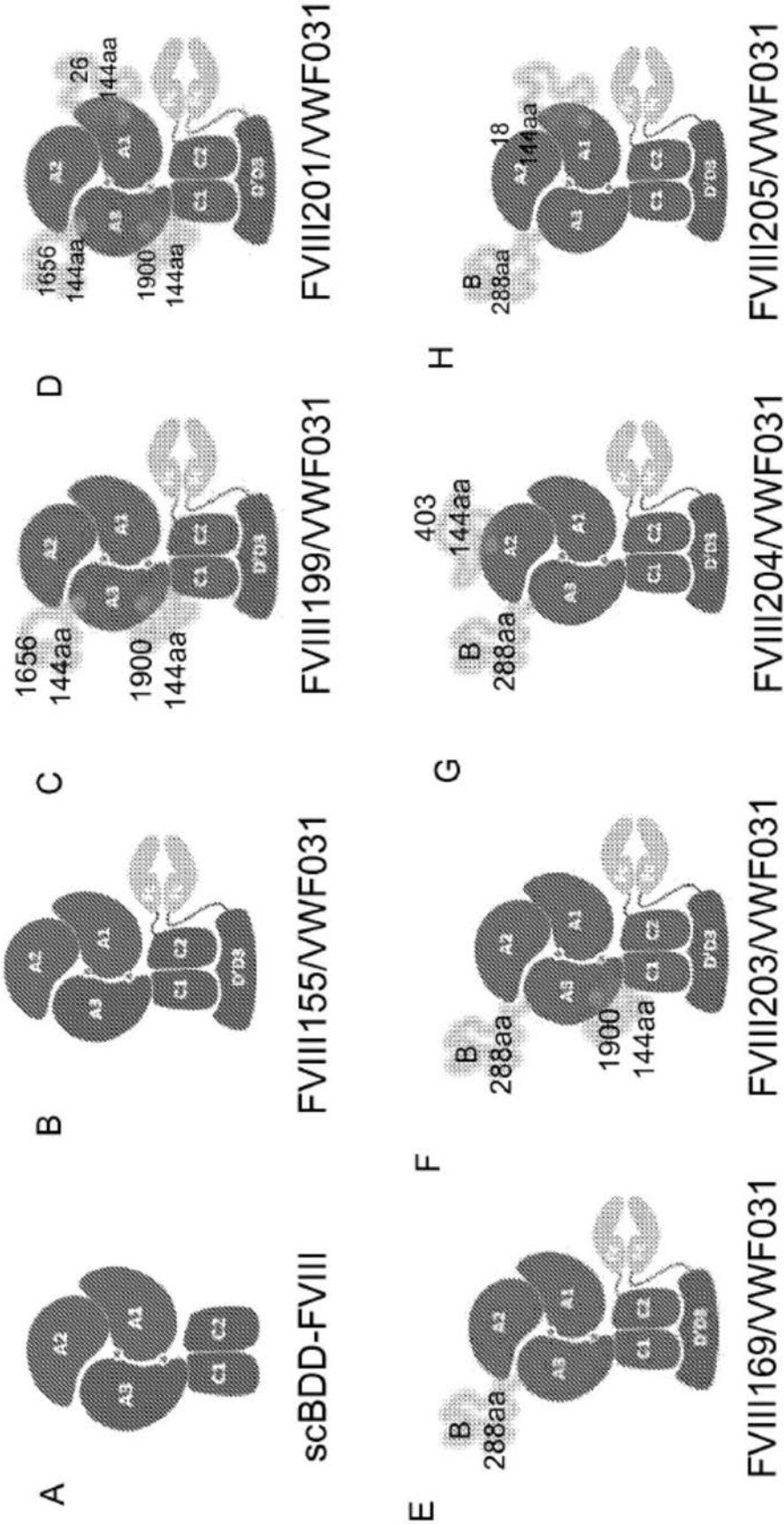


图14

通过rFVIII-XTEN/VWF引起HemA小鼠中的半衰期延长

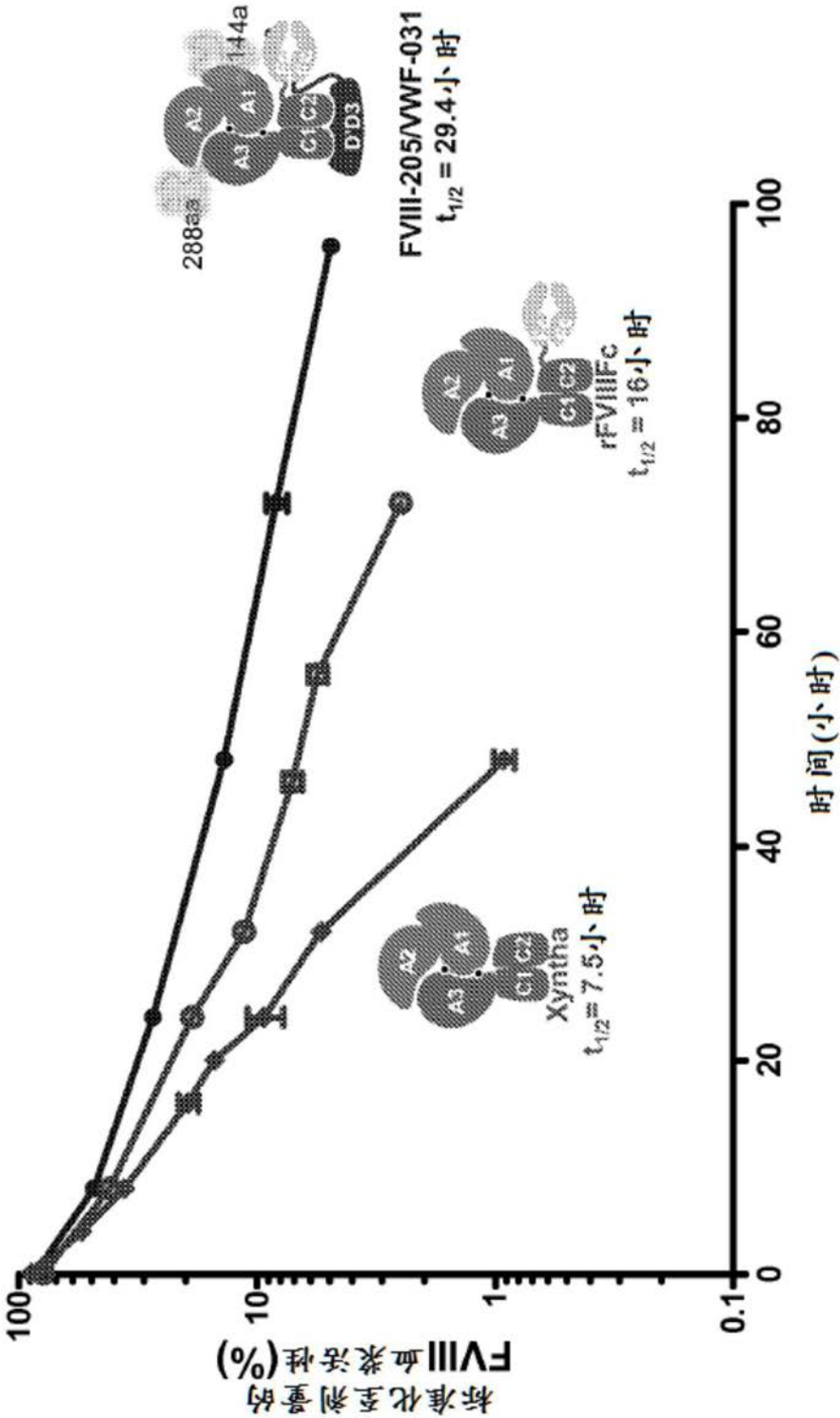


图15

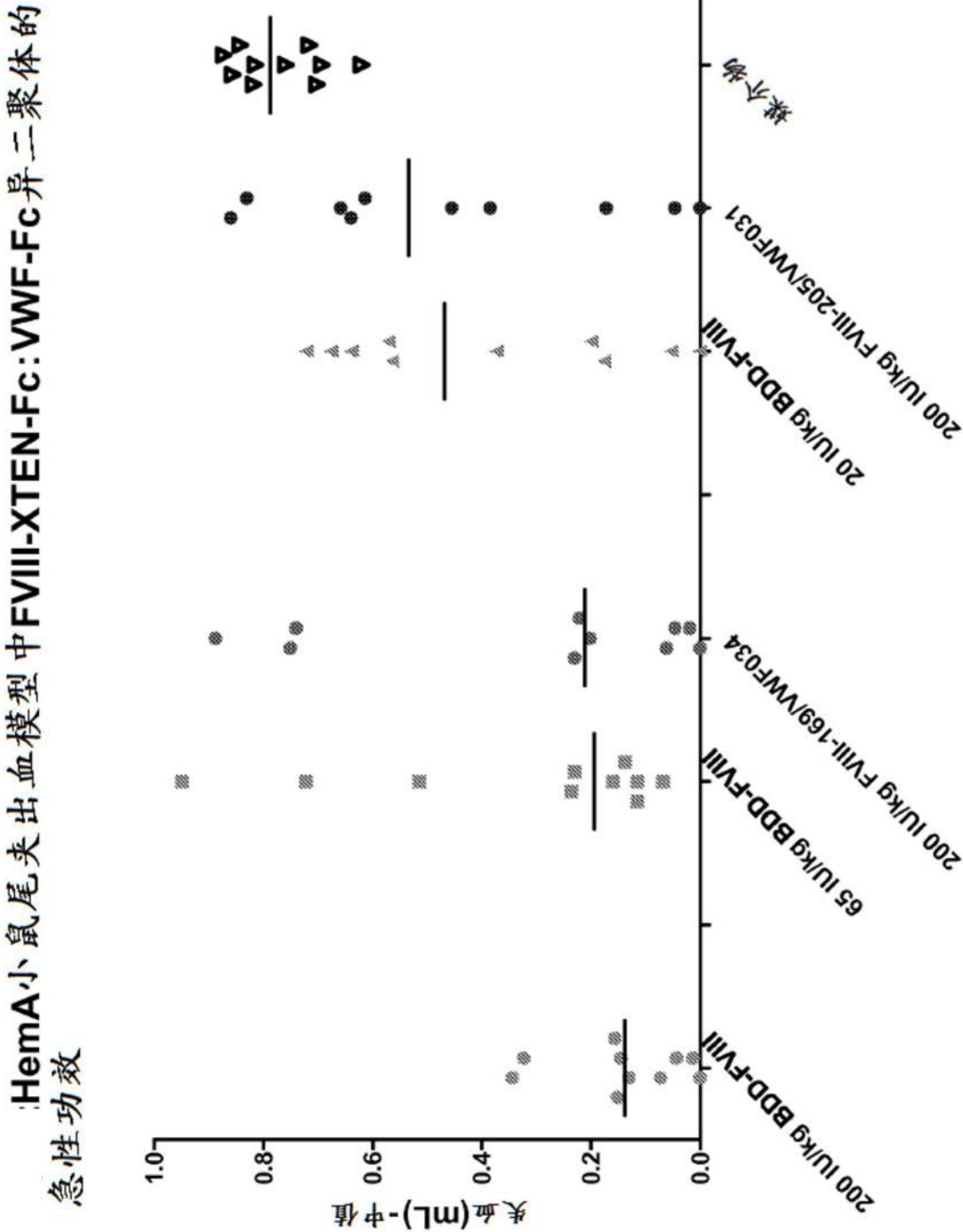


图16