



등록특허 10-2714760



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2024년10월11일  
(11) 등록번호 10-2714760  
(24) 등록일자 2024년10월02일

- (51) 국제특허분류(Int. Cl.)  
*C07K 14/755* (2006.01)
- (52) CPC특허분류  
*C07K 14/755* (2013.01)  
*C07K 2319/30* (2013.01)
- (21) 출원번호 10-2024-7004922(분할)
- (22) 출원일자(국제) 2014년06월27일  
심사청구일자 2024년03월14일
- (85) 번역문제출일자 2024년02월13일
- (65) 공개번호 10-2024-0023705
- (43) 공개일자 2024년02월22일
- (62) 원출원 특허 10-2022-7045978  
원출원일자(국제) 2014년06월27일  
심사청구일자 2023년01월25일
- (86) 국제출원번호 PCT/US2014/044731
- (87) 국제공개번호 WO 2014/210558  
국제공개일자 2014년12월31일
- (30) 우선권주장  
61/840,872 2013년06월28일 미국(US)
- (56) 선행기술조사문헌  
Nature Biotech., Vol. 27, No. 12,  
pp.1186-1190 (2009. 11. 15.)  
(뒷면에 계속)

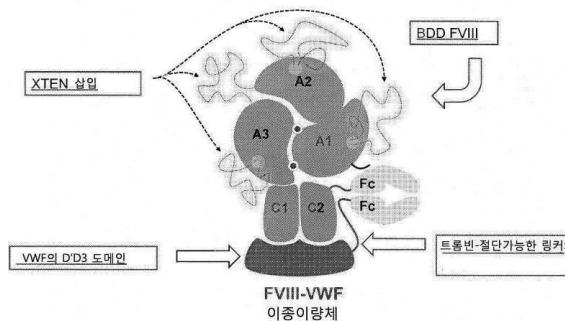
전체 청구항 수 : 총 27 항

심사관 : 김수진

(54) 발명의 명칭 XTEN을 지닌 트롬빈 절단가능한 링커 및 이의 용도

**(57) 요약**

본 발명은 VWF 링커를 통해 이종 잔기에 융합된 VWF 단백질을 포함하는 키메라 분자를 제공한다. 본 발명은 트롬빈의 존재하에서 절단될 수 있는 효율적인 VWF 링커를 제공한다. 키메라 분자는 FVIII 단백질 및 제2의 이종 잔기를 포함하는 폴리펩타이드를 추가로 포함할 수 있으며, 여기서 VWF 단백질을 포함하는 쇄 및 FVIII 단백질을 포함하는 쇄는 서로 연합되어 있다. 본 발명은 또한 뉴클레오타이드, 벡터, 숙주 세포, 키메라 단백질을 사용하는 방법을 포함한다.

**대 표 도**

(52) CPC특허분류

C07K 2319/50 (2013.01)

(56) 선행기술조사문현

미국 특허출원공개공보 US2011/0183907

호(2011.07.28.)

KR1020140115347 A

KR1020150036510 A

US20080146782 A1

US20110183907 A1

US20150266943 A1

---

## 명세서

### 청구범위

#### 청구항 1

지혈 장애(hemostatic disorder)의 치료에 사용하기 위한 약제학적 조성물이며,

상기 약제학적 조성물은 제1 및 제2의 폴리펩타이드 쇄를 포함하는 키메라 분자를 포함하고;

여기서 제1의 폴리펩타이드 쇄는 폰 빌레브란트 인자(von Willebrand Factor: VWF) 단백질, 제1의 연장된 재조합 폴리펩타이드(XTEN) 서열, VWF 링커 및 제1의 면역글로불린 불변 영역 또는 이의 일부를 포함하며;

VWF 단백질은 VWF 의 D' 도메인 및 D3 도메인을 포함하고, D' 도메인은 서열 번호 2의 764 내지 866번 아미노산과 동일한 아미노산 서열을 포함하고, VWF 단백질의 D3 도메인은 서열 번호 2의 1099번 및 1142번 잔기에서 아미노산 치환을 갖는 서열 번호 2의 867 내지 1240번 아미노산과 동일한 아미노산 서열을 포함하고;

VWF 링커는 서열 번호 16에 상응하는 Glu720 내지 Arg740과 동일한 아미노산 서열을 포함하는 인자 VIII(FVII I)로부터의 a2 영역을 포함하고, 여기서 a2 영역은 트롬빈에 의해 절단될 수 있으며;

제1의 XTEN 서열은 VWF 단백질을 VWF 링커와 연결시키고, VWF 링커는 제1의 XTEN 서열을 제1의 면역글로불린 불변 영역 또는 이의 일부에 연결시키고;

제2의 폴리펩타이드 쇄는 FVIII 단백질, 제2의 XTEN 서열, 및 제2의 면역글로불린 불변 영역 또는 이의 일부를 포함하며;

FVIII 단백질은 서열 번호 16의 Ala1 내지 Arg745와 동일한 아미노산 서열 및 서열 번호 16의 Glu1649 내지 Tyr2332와 동일한 아미노산 서열을 포함하고;

제1 및 제2의 폴리펩타이드 쇄는 공유결합에 의해 연합되어 있는 것인, 약제학적 조성물.

#### 청구항 2

제1항에 있어서, 상기 제2의 XTEN 서열이 FVIII 단백질의 N-말단 또는 C-말단에 연결되어 있거나 서로 인접한 2개의 FVIII 아미노산 사이에 삽입된 것인, 약제학적 조성물.

#### 청구항 3

제1항에 있어서, 적어도 하나의 XTEN 서열이 42개의 아미노산, 72개의 아미노산, 108개의 아미노산, 144개의 아미노산, 180개의 아미노산, 216개의 아미노산, 252개의 아미노산, 288개의 아미노산, 324개의 아미노산, 360개의 아미노산, 396개의 아미노산, 432개의 아미노산, 468개의 아미노산, 504개의 아미노산, 540개의 아미노산, 576개의 아미노산, 612개의 아미노산, 624개의 아미노산, 648개의 아미노산, 684개의 아미노산, 720개의 아미노산, 756개의 아미노산, 792개의 아미노산, 828개의 아미노산, 836개의 아미노산, 864개의 아미노산, 875개의 아미노산, 912개의 아미노산, 923개의 아미노산, 948개의 아미노산, 1044개의 아미노산, 1140개의 아미노산, 1236개의 아미노산, 1318개의 아미노산, 1332개의 아미노산, 1428개의 아미노산, 1524개의 아미노산, 1620개의 아미노산, 1716개의 아미노산, 1812개의 아미노산, 1908개의 아미노산, 또는 2004개의 아미노산을 포함하는 것인 약제학적 조성물.

#### 청구항 4

제1항에 있어서, 적어도 하나의 XTEN 서열이 AE42, AE72, AE864, AE576, AE288, AE144, AG864, AG576, AG288 또는 AG144로부터 선택되는 것인 약제학적 조성물.

#### 청구항 5

제1항에 있어서, 적어도 하나의 XTEN 서열이 서열 번호 39; 서열 번호 40; 서열 번호 47; 서열 번호 45; 서열 번호 44; 서열 번호 41; 서열 번호 48; 서열 번호 46; 서열 번호 44 또는 서열 번호 42로부터 선택되는 것인 약제학적 조성물.

**청구항 6**

제1항에 있어서, 상기 VWF 단백질이 서열 번호 2의 1099번 및 1142번 아미노산 잔기에서 시스테인으로부터 알라닌으로의 아미노산 치환을 함유하는 것인 약제학적 조성물.

**청구항 7**

제1항에 있어서, 상기 VWF 단백질이 A1 VWF 도메인, A2 VWF 도메인, A3 VWF 도메인, D4 VWF 도메인, B1 VWF 도메인, B2 VWF 도메인, B3 VWF 도메인, C1 VWF 도메인, C2 VWF 도메인 또는 CK VWF 도메인을 포함하지 않는 것인 약제학적 조성물.

**청구항 8**

제1항에 있어서, 상기 VWF 단백질이 VWF의 D' 및 D3 도메인으로 이루어진 것인 약제학적 조성물.

**청구항 9**

제1항에 있어서, 상기 제1의 면역글로불린 불변 영역 또는 이의 일부가 제1의 Fc 영역을 포함하고, 상기 제2의 면역글로불린 불변 영역 또는 이의 일부가 제2의 Fc 영역을 포함하는 것인 약제학적 조성물.

**청구항 10**

제1항에 있어서, 제1의 Fc 영역과 제2의 Fc 영역이 동일한 것인 약제학적 조성물.

**청구항 11**

제1항에 있어서, 상기 제1의 면역글로불린 불변 영역 또는 이의 일부 및 상기 제2의 면역글로불린 불변 영역 또는 이의 일부가 하나 이상의 이황화물 결합에 의해 서로 연합되어 있는 것인 약제학적 조성물.

**청구항 12**

제1항에 있어서, 키메라 분자가 다음 화학식을 포함하는 것인 약제학적 조성물:

V-X1-L1-H1:H2-L2-C(X2);

여기서, V는 VWF 단백질이고;

X1은 제1의 XTEN 서열이며;

L1은 VWF 링커이고;

L2는 임의의 FVIII 링커이며;

H1은 제1의 면역글로불린 불변 영역 또는 이의 일부이고;

H2는 제2의 면역글로불린 불변 영역 또는 이의 일부이며;

C는 FVIII 단백질이고;

C(X2)는 제2의 XTEN 서열에 융합된 FVIII 단백질이고, 여기서 제2의 XTEN 서열은 서로 인접한 2개의 FVIII 아미노산 사이에 삽입되며;

(-)는 팹타이드 결합 또는 하나 이상의 아미노산이고;

(:)는 H1과 H2 사이의 하나 이상의 공유 결합이다.

**청구항 13**

지혈 장애(hemostatic disorder)의 치료에 사용하기 위한 약제학적 조성물이며,

상기 약제학적 조성물은 제1 및 제2의 폴리팹타이드 쇄를 포함하는 키메라 분자를 포함하고;

여기서 제1의 폴리팹타이드 쇄는, N-말단으로부터 C-말단까지, 혼 빌레브란트 인자(von Willebrand Factor: VWF) 단백질, 제1의 연장된 재조합 폴리팹타이드(XTEN) 서열, VWF 링커 및 제1의 Fc 영역 또는 이의 일부를 포

함하며;

VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함하고;

VWF 단백질의 D' 도메인은 서열 번호 2의 764 내지 866번 아미노산과 동일한 아미노산 서열을 포함하며;

VWF 단백질의 D3 도메인은 서열 번호 2의 1099번 및 1142번 아미노산 잔기에서 시스테인으로부터 알라닌으로의 아미노산 치환을 갖는 서열 번호 2의 867 내지 1240번 아미노산과 동일한 아미노산 서열을 포함하고;

VWF 링커는 서열 번호 16에 상응하는 Glu720 내지 Arg740과 동일한 아미노산 서열을 포함하는 인자 VIII(FVII I)로부터의 a2 영역을 포함하고, 여기서 a2 영역은 트롬빈에 의해 절단될 수 있으며;

VWF 링커는 적어도 30개의 아미노산이고;

제1의 XTEN 서열은 VWF 단백질을 VWF 링커와 연결시키고, VWF 링커는 제1의 XTEN 서열을 제1의 Fc 영역 또는 이의 일부에 연결시키며;

제2의 폴리펩타이드 쇄는 FVIII 단백질, FVIII 단백질에 삽입된 제2의 XTEN 서열, 및 FVIII 단백질의 C-말단에 연결되는 제2의 Fc 영역 또는 이의 일부를 포함하고;

FVIII 단백질은 서열 번호 16의 Ala1 내지 Asn745와 동일한 아미노산 서열 및 서열 번호 16의 Ser1690 내지 Tyr2332와 동일한 아미노산 서열을 포함하며;

제1 및 제2의 폴리펩타이드 쇄는 제1의 Fc 영역 또는 이의 일부와 제2의 Fc 영역 또는 이의 일부 사이의 하나 이상의 이황화물 결합에 의해 연합되어 있는 것인, 약제학적 조성물.

#### 청구항 14

제13항에 있어서, 상기 VWF 링커는 서열 번호 4와 동일한 아미노산 서열을 포함하는 것인 약제학적 조성물.

#### 청구항 15

제13항에 있어서, 상기 제1의 Fc 영역 또는 이의 일부와 상기 제2의 Fc 영역 또는 이의 일부가 두 개의 이황화물 결합에 의해 연합되어 있는 것인 약제학적 조성물.

#### 청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서, 지혈 장애가 혈우병 A인 약제학적 조성물.

#### 청구항 17

제16항에 있어서, 지혈 장애가 중증 혈우병 A인 약제학적 조성물.

#### 청구항 18

제1항 내지 제15항 중 어느 한 항에 있어서, 지혈 장애가 폰 빌레브란트병(von Willebrand disease)인 약제학적 조성물.

#### 청구항 19

키메라 단백질을 암호화하는 폴리뉴클레오타이드 세트이며,

상기 폴리뉴클레오타이드 세트는 제1의 폴리뉴클레오타이드 및 제2의 폴리뉴클레오타이드를 포함하고, 제1의 폴리뉴클레오타이드 및 제2의 폴리뉴클레오타이드가 각각 제1의 폴리펩타이드 쇄 및 제2의 폴리펩타이드 쇄를 암호화하고;

여기서 제1의 폴리펩타이드 쇄는 폰 빌레브란트 인자(VWF) 단백질, 제1의 연장된 재조합 폴리펩타이드(XTEN) 서열, VWF 링커 및 제1 면역글로불린 불면 영역 또는 이의 일부를 포함하며;

VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함하고, D' 도메인은 서열 번호 2의 764 내지 866번 아미노산과 동일한 아미노산 서열을 포함하고, VWF 단백질의 D3 도메인은 서열 번호 2의 1099번 및 1142번 잔기에서 아미노산 치환을 갖는 서열 번호 2의 867 내지 1240번 아미노산과 동일한 아미노산 서열을 포함하고;

VWF 링커는 서열 번호 16에 상응하는 Glu720 내지 Arg740과 동일한 아미노산 서열을 포함하는 인자 VIII(FVII I)

I)로부터의 a2 영역을 포함하고, 여기서 a2 영역은 트롭빈에 의해 절단될 수 있으며;

제1의 XTEN 서열은 VWF 단백질을 VWF 링커와 연결시키고, VWF 링커는 제1의 XTEN 서열을 제1의 면역글로불린 불변 영역 또는 이의 일부에 연결시키고;

제2의 폴리펩타이드 쇄는 FVIII 단백질, 제2의 XTEN 서열 및 제2의 면역글로불린 불변 영역 또는 이의 일부를 포함하며;

FVIII 단백질은 서열 번호 16의 Ala1 내지 Asn745와 동일한 아미노산 서열 및 서열 번호 16의 Glu1649 내지 Tyr2332와 동일한 아미노산 서열을 포함하는 것인, 폴리뉴클레오타이드 세트.

#### 청구항 20

키메라 단백질을 암호화하는 폴리뉴클레오타이드 세트이며,

상기 폴리뉴클레오타이드 세트는 제1의 폴리뉴클레오타이드 및 제2의 폴리뉴클레오타이드를 포함하고, 제1의 폴리뉴클레오타이드 및 제2의 폴리뉴클레오타이드는 각각 제1의 폴리펩타이드 쇄 및 제2의 폴리펩타이드 쇄를 암호화하고;

여기서 제1의 폴리펩타이드 쇄는, N-말단에서 C-말단까지, 폰 빌레브란트 인자(VWF) 단백질, 제1의 연장된 재조합 폴리펩타이드(XTEN) 서열, VWF 링커 및 제1의 Fc 영역 또는 이의 일부를 포함하며;

VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함하고;

VWF 단백질의 D' 도메인은 서열 번호 2의 764 내지 866번 아미노산과 동일한 아미노산 서열을 포함하며;

VWF 단백질의 D3 도메인은 서열 번호 2의 1099번 및 1142번 아미노산 잔기에서 시스테인으로부터 알라닌으로의 아미노산 치환을 갖는 서열 번호 2의 867 내지 1240번 아미노산과 동일한 아미노산 서열을 포함하고;

VWF 링커는 서열 번호 16에 상응하는 Glu720 내지 Arg740과 동일한 아미노산 서열을 포함하는 인자 VIII(FVII I)로부터의 a2 영역을 포함하고, 여기서 a2 영역은 트롭빈에 의해 절단될 수 있으며;

VWF 링커는 적어도 30개의 아미노산이고;

제1의 XTEN 서열은 VWF 단백질을 VWF 링커와 연결시키고, VWF 링커는 제1의 XTEN 서열을 제1의 Fc 영역 또는 이의 일부에 연결시키며;

제2의 폴리펩타이드 쇄는 FVIII 단백질, FVIII 단백질에 삽입된 제2의 XTEN 서열, 및 FVIII 단백질의 C-말단에 연결되는 제2의 Fc 영역 또는 이의 일부를 포함하고;

FVIII 단백질은 서열 번호 16의 Ala1 내지 Asn745와 동일한 서열 및 서열 번호 16의 Ser1690 내지 Tyr2332와 동일한 아미노산 서열을 포함하는 것인, 폴리펩타이드 세트.

#### 청구항 21

제19항 또는 제20항에 있어서, 쌍을 이룬 염기성 아미노산 절단 효소(Paired basic Amino acid Cleaving Enzyme: PACE)를 암호화하는 폴리뉴클레오타이드를 추가로 포함하는 폴리뉴클레오타이드 세트.

#### 청구항 22

제19항 또는 제20항의 폴리뉴클레오타이드 세트를 포함하는 숙주 세포.

#### 청구항 23

제21항의 폴리뉴클레오타이드 세트를 포함하는 숙주 세포.

#### 청구항 24

제22항에 있어서, 포유동물 세포인 숙주 세포.

#### 청구항 25

제24항에 있어서, 포유동물 세포가 HEK293 세포, CHO 세포 및 BHK 세포로 이루어진 그룹으로부터 선택되는 것인

숙주 세포.

### 청구항 26

제25항에 있어서, HEK293 세포인 숙주 세포.

### 청구항 27

제26항에 있어서, HEK293F 세포인 숙주 세포.

## 발명의 설명

### 배경기술

[0001]

해모필라(Haemophilia) A는 응고 인자 VIII(FVIII)를 암호화하는 유전자내의 결함에 의해 유발된 출혈 질환이며 10,000명의 남성 출생시 1 내지 2명에게 영향을 미친다. Graw 등, Nat. Rev. Genet. 6(6): 488-501 (2005). 혈우병 A에 걸린 환자는 정제되거나 재조합적으로 생산된 FVIII의 주입으로 치료할 수 있다. 그러나, 모든 상업적으로 이용가능한 FVIII 생성물은, 반감기가 약 8 내지 12시간이고, 환자에게 빈번한 정맥내 투여를 필요로 한다. [참조: Weiner M.A. and Cairo, M.S., Pediatric Hematology Secrets, Lee, M.T., 12. Disorders of Coagulation, Elsevier Health Sciences, 2001; Lillicrap, D. Thromb. Res. 122 Suppl 4:S2-8 (2008)]. 또한, 다수의 시도가, FVIII 반감기를 연장시키기 위해 노력되어 왔다. 예를 들면, 응고 인자의 반감기를 연장시키기 위한 개발에 있어서의 시도는 폐길화(pegylation), 당폐길화, 및 알부민과의 접합을 포함한다. [참조: Dumont 등, Blood. 119(13): 3024-3030 (Published online Jan. 13, 2012)]. 그러나, 사용된 단백질 가공에도 불구하고, 현재 개발중인 장기 작용하는 FVIII 생성물은 반감기를 개선시켜 왔으나, 반감기는 전임상 동물 모델에서 - 단지 약 1.5 내지 2배의 개선으로 제한되는 것으로 보고되어 있다. (상기 참조). 일치하는 결과가 사람에서 입증되었는데, 예를 들어, rFVIIIFc는 혈우병 A 환자에서 ADVATE®과 비교하여 ~1.7배까지 반감기를 개선시키는 것으로 보고되었다. (상기 참조). 따라서, 약간의 개선에도 불구하고, 반감기는 다른  $t_{1/2}$  제한 인자의 존재를 나타낼 수 있다.

[0002]

투여 스케줄에 의해 유발된 빈번한 투여 및 불편함으로 인하여, 빈번하지 않은 투여를 필요로 하는 FVIII 생성물, 즉, 1.5 내지 2배 반감기 제한보다 긴 반감기를 갖는 FVIII 생성물을 개발할 필요성이 여전히 존재한다.

[0003]

### 발명의 간단한 요약

[0004]

본 발명은 폰 빌레브란트 인자(Von Willebrand Factor: VWF) 단백질, 이종 잔기(H1), XTEN 서열, 및 VWF 단백질과 이종 잔기를 연결시키는 VWF 링커를 포함하는 키메라 분자에 관한 것이며, 여기서 상기 VWF 링커는: (i) 인자 VIII로부터의 a2 영역(FVIII); (ii) FVIII로부터의 a1 영역; (iii) FVIII으로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다) (서열 번호 3) 및 PAR1 엑소사이트(exosite) 상호작용 모티프(motif)를 포함하는 트롭빈 절단 부위; 또는 (v) 이의 어떠한 조합으로부터 선택된 폴리펩타이드를 포함하며, 여기서 상기 XTEN 서열은 VWF 단백질, 이종 잔기(H1), VWF 링커, 또는 이의 어떠한 조합에 연결되어 있다. 일 구현예에서, XTEN 서열은 VWF 단백질과 VWF 링커 또는 VWF 링커와 이종 잔기를 연결시킨다. 다른 구현예에서, 키메라 분자는 FVIII 단백질을 포함하는 제2의 폴리펩타이드 쇄를 추가로 포함하며, 여기서 제1의 폴리펩타이드 쇄 및 제2의 폴리펩타이드 쇄는 서로 연합되어 있다. 다른 구현예에서, 키메라 분자내 FVIII 단백질은 추가의 XTEN 서열을 추가로 포함한다. 추가의 XTEN 서열은 FVIII 단백질의 N-말단 또는 C-말단에 연결될 수 있거나 서로 인접한 2개의 FVIII 아미노산 사이에 삽입된다. 여전히 다른 구현예에서, 제2의 폴리펩타이드는 제2의 이종 잔기(H2)를 추가로 포함한다.

[0005]

본 개재내용은 또한 VWF 단백질, 이종 잔기(H1), 및 VWF 단백질과 이종 잔기(H1)를 연결하는 VWF 링커(linker)를 포함하는 제1의 폴리펩타이드 쇄 및 FVIII 단백질 및 XTEN 서열을 포함하는 제2의 폴리펩타이드 쇄를 포함하는 키메라 분자를 포함하며, 여기서, 상기 제1의 폴리펩타이드 쇄내 VWF 링커는: (i) FVIII으로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서, X는 지방족 아미노산이다) (서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프(exocite interaction motif)를 포함하는 트롭빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함하고, 여기서 제1의 폴리펩타이드 쇄 및 제2의 폴리펩타이드 쇄는 서로 연합되어 있다. 일 구현예에서, XTEN 서열은 FVIII 단백질의 N-말단 또는 C-말단에 연결되어 있거나 서로 인접한 2개의 FVIII 아미노산 사이에 삽입되어 있다. 다른 구현예에서, 키메라 분자는 추가의 XTEN 서열을 추가로

포함하고, 이는 VWF 단백질, 이종 잔기, VWF 링커, 또는 이의 어떠한 조합에 연결되어 있다. 다른 구현예에서, 키메라 분자는 제2의 이종 잔기(H2)를 추가로 포함한다. 여전히 다른 구현예에서, 제2의 이종 잔기는 FVIII 단백질, XTEN 서열, 또는 둘 다에 연결되어 있다.

[0006] 본 개재내용의 키메라 분자의 경우, VWF 단백질, VWF 링커, FVIII 단백질, 또는 키메라 단백질내 어떠한 다른 성분에 연결된 XTEN 서열은 약 42개 아미노산, 약 72개 아미노산, 약 108개 아미노산, 약 144개 아미노산, 약 180개 아미노산, 약 216개 아미노산, 약 252개 아미노산, 약 288개 아미노산, 약 324개 아미노산, 약 360개 아미노산, 약 396개 아미노산, 약 432개 아미노산, 약 468개 아미노산, 약 504개 아미노산, 약 540개 아미노산, 약 576개 아미노산, 약 612개 아미노산, 약 624개 아미노산, 약 648개 아미노산, 약 684개 아미노산, 약 720개 아미노산, 약 756개 아미노산, 약 792개 아미노산, 약 828개 아미노산, 약 836개 아미노산, 약 864개 아미노산, 약 875개 아미노산, 약 912개 아미노산, 약 923개 아미노산, 약 948개 아미노산, 약 1044개 아미노산, 약 1140개 아미노산, 약 1236개 아미노산, 약 1318개 아미노산, 약 1332개 아미노산, 약 1428개 아미노산, 약 1524개 아미노산, 약 1620개 아미노산, 약 1716개 아미노산, 약 1812개 아미노산, 약 1908개 아미노산, 또는 약 2004개 아미노산을 포함한다. 일부 구현예에서, XTEN 폴리펩타이드는 AE42, AE72, AE864, AE576, AE288, AE144, AG864, AG576, AG288, 또는 AG144로부터 선택된다. 다른 구현예에서, XTEN 폴리펩타이드는 서열 번호 39; 서열 번호 40; 서열 번호 47; 서열 번호 45; 서열 번호 44; 서열 번호 41; 서열 번호 48; 서열 번호 46, 서열 번호 44, 또는 서열 번호 42로부터 선택된다.

[0007] 다른 측면에서, 키메라 분자내 추가의 XTEN 서열은 약 42개 아미노산, 약 72개 아미노산, 약 108개 아미노산, 약 144개 아미노산, 약 180개 아미노산, 약 216개 아미노산, 약 252개 아미노산, 약 288개 아미노산, 약 324개 아미노산, 약 360개 아미노산, 약 396개 아미노산, 약 432개 아미노산, 약 468개 아미노산, 약 504개 아미노산, 약 540개 아미노산, 약 576개 아미노산, 약 612개 아미노산, 약 624개 아미노산, 약 648개 아미노산, 약 684개 아미노산, 약 720개 아미노산, 약 756개 아미노산, 약 792개 아미노산, 약 828개 아미노산, 약 836개 아미노산, 약 864개 아미노산, 약 875개 아미노산, 약 912개 아미노산, 약 923개 아미노산, 약 948개 아미노산, 약 1044개 아미노산, 약 1140개 아미노산, 약 1236개 아미노산, 약 1318개 아미노산, 약 1332개 아미노산, 약 1428개 아미노산, 약 1524개 아미노산, 약 1620개 아미노산, 약 1716개 아미노산, 약 1812개 아미노산, 약 1908개 아미노산, 또는 약 2004개 아미노산을 포함한다. 일부 구현예에서, 추가의 XTEN 폴리펩타이드는 AE42, AE72, AE864, AE576, AE288, AE144, AG864, AG576, AG288, 또는 AG144로부터 선택된다. 특정의 구현예에서, 추가의 XTEN 폴리펩타이드는 서열 번호 39; 서열 번호 40; 서열 번호 47; 서열 번호 45; 서열 번호 43; 서열 번호 41; 서열 번호 48; 서열 번호 46, 서열 번호 44, or 서열 번호 42로부터 선택된다.

[0008] 일 구현예에서, 키메라 분자내에서 VWF 단백질과 이종 잔기를 연결시키는데 유용한 VWF 링커는 전장의 FVIII에 상응하는 Glu720 내지 Arg740에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 100% 동일한 아미노산 서열을 포함하는 a2 영역을 포함하며, 여기서 a2 영역은 트롬빈에 의해 절단될 수 있다. 특수한 구현예에서, a2 영역은 ISDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFS(서열 번호 4)을 포함한다. 다른 구현예에서, VWF 단백질 및 이종 잔기를 연결시키는데 유용한 VWF 링커는 전장의 FVIII에 상응하는 Met337 내지 Arg372에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 100% 동일한 아미노산 서열을 포함하는 a1 영역을 포함하고, 여기서 a1 영역은 트롬빈에 의해 절단될 수 있다. 일부 구현예에서, a1 영역은 ISMKNNNEAEDYDDDLTDSEMDVVRFDDNSPSFIQIRSV(서열 번호 5)를 포함한다.

[0009] 다른 구현예에서, VWF 단백질과 이종 잔기를 연결시키는데 유용한 VWF 링커는 전장의 FVIII에 상응하는 Glu1649 내지 Arg1689에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 또는 100% 동일한 a3 영역을 포함하고, 여기서 상기 a3 영역은 트롬빈에 의해 절단될 수 있다. 구체적인 구현예에서, a3 영역은 ISEITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQ(서열 번호 6)를 포함한다.

[0010] 여전히 다른 구현예에서, VWF 단백질과 이종 잔기를 연결시키는데 유용한 VWF 링커는 X-V-P-R(서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위를 포함하고, 여기서 PAR1 엑소사이트 상호작용 모티프는 S-F-L-L-R-N(서열 번호 7)을 포함한다. 일 구현예에서, PAR1 엑소사이트 상호작용 모티프는 P, P-N, P-N-D, P-N-D-K(서열 번호 8), P-N-D-K-Y(서열 번호 9), P-N-D-K-Y-E(서열 번호 10), P-N-D-K-Y-E-P(서열 번호 11), P-N-D-K-Y-E-P-F(서열 번호 12), P-N-D-K-Y-E-P-F-W(서열 번호 13), P-N-D-K-Y-E-P-F-W-E(서열 번호 14), P-N-D-K-Y-E-P-F-W-E-D(서열 번호 20), P-N-D-K-Y-E-P-F-W-E-D-E(서열 번호 21), P-N-D-K-Y-E-P-F-W-E-D-E(서열 번호 22), P-N-D-K-Y-E-P-F-W-E-D-E-S(서열 번호 23), 또는 이의 조합으로부터 선택된 서열을 추가로 포함한다. 다른 구현예에서, 여기서 지방족 아미노산은 글리신, 알라닌, 발린, 루이신, 또는 이소루이신으로부

터 선택된다. 특수한 구현예에서, VWF 링커는 GGLVPRSFLLRNPNDKYEPFWEDDEES(서열 번호 24)를 포함한다.

[0011] 특정의 구현예에서, 트롬빈은 트롬빈 절단 부위가 키메라 분자내 VWF 링커에 대해 치환된 경우 트롬빈이 트롬빈 절단 부위를 절단할 수 있는 것보다 더 빠르게 VWF 링커를 절단한다. 다른 구현예에서, 트롬빈은, 트롬빈 절단 부위가 키메라 분자내에서 VWF 링커에 대해 치환되는 경우 트롬빈이 트롬빈 절단 부위를 절단할 수 있는 것 보다 적어도 약 10배, 적어도 약 20배, 적어도 약 30배, 적어도 약 40배, 적어도 약 50배, 적어도 약 60배, 적어도 약 70배, 적어도 약 80배, 적어도 약 90배 또는 적어도 약 100배로 VWF 링커를 절단한다.

[0012] 일부 구현예에서, VWF 링커는 적어도 약 5, 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, 또는 2000개 아미노산의 길이를 갖는 하나 이상의 아미노산을 또한 포함한다. 하나의 예에서, 하나 이상의 아미노산은 gly 웨بت아이드를 포함한다. 다른 예에서, 하나 이상의 아미노산은 GlyGly를 포함한다. 다른 예에서, 하나 이상의 아미노산은 gly/ser 웨بت아이드를 포함한다. 일부 예에서, gly/ser 웨بت아이드는 화학식 (Gly<sub>4</sub>Ser)<sub>n</sub> 또는 S(Gly<sub>4</sub>Ser)<sub>n</sub>를 가지며, 여기서 n은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 또는 100으로부터 선택된 양의 정수이다. 특정의 예에서, (Gly<sub>4</sub>Ser)<sub>n</sub> 링커는 (Gly<sub>4</sub>Ser)<sub>3</sub>(서열 번호 89) 또는 (Gly<sub>4</sub>Ser)<sub>4</sub>(서열 번호 90)이다.

[0013] 본 발명의 키메라 분자에 유용한 VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함할 수 있고, 여기서 D' 도메인 및 D3 도메인은 FVIII 단백질에 결합할 수 있다. 하나의 구현예에서, VWF 단백질의 D' 도메인은 서열 번호 2의 764 내지 866번 아미노산에 대해 적어도 약 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 아미노산 서열을 포함한다. 다른 구현예에서, VWF 단백질의 D3 도메인은 서열 번호 2의 867 내지 1240번 아미노산에 대해 적어도 약 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 아미노산 서열을 포함한다. 다른 구현예에서, VWF 단백질은 서열 번호 2의 1099번 잔기, 1142번 잔기, 또는 1099 및 1142번 잔기 둘 다에 상응하는 잔기에서 적어도 하나의 아미노산 치환을 함유한다. 여전히 다른 구현예에서, VWF 단백질의 서열내에서, 시스테인 이외의 아미노산은 서열 번호 2의 1099번 잔기, 1142번 잔기, 또는 1099번 및 1142번 잔기 둘 다에 상응하는 잔기에 대해 치환된다. 여전히 다른 구현예에서, VWF 단백질의 서열은 서열 번호 2의 764 내지 1240번 아미노산을 포함한다. 특정의 구현예에서, VWF 단백질은 VWF의 D1 도메인, D2 도메인, 또는 D1 및 D2 도메인을 추가로 포함한다. 일부 구현예에서, VWF 단백질은 A1 도메인, A2 도메인, A3 도메인, D4 도메인, B1 도메인, B2 도메인, B3 도메인, C1 도메인, C2 도메인, CK 도메인, 이의 하나 이상의 단편, 또는 이의 어떠한 조합으로부터 선택된 VWF 도메인을 또한 포함한다. 다른 구현예에서, VWF 단백질은: (1) VWF 또는 이의 단편의 D' 및 D3 도메인; (2) VWF 또는 이의 단편의 D1, D', 및 D3 도메인; (3) VWF 또는 이의 단편의 D2, D', 및 D3 도메인; (4) VWF 또는 이의 단편의 D1, D2, D', 및 D3 도메인; 또는 (5) VWF 또는 이의 단편의 D1, D2, D', D3, 및 A1 도메인으로 필수적으로 이루어지거나 이로 이루어진다. 여전히 다른 구현예에서, VWF 단백질은 VWF의 단일 웨بت아이드를 추가로 포함한다. 여전히 다른 구현예에서, VWF 단백질은 폐길화되어 있거나, 글리코실화되어 있거나, 헤실화되어 있거나, 폴리실릴화되어 있다. 용어 "폐길화된"은 단백질 상에 폴리에틸렌 글리콜(PEG)을 갖는 것을 말하며, 용어 "글리코실화"는 단백질 상에 글리코실화를 갖는 것을 말하고; 용어 "헤실화된"은 단백질 상에 하이드록시에틸 전분(HES)을 갖는 것을 말하고; 용어 "폴리시알릴화"는 단백질 상에 폴리시알산(PSA)을 갖는 것을 말한다. PEG, HES, 및 PSA의 예는 본원의 어디에서도 나타나 있다.

[0014] 일부 측면에서, VWF 잔기를 통해 VWF 단백질에 융합된 이종 잔기(H1)는 키메라 분자의 반감기를 연장시킬 수 있다. 일 구현예에서, 이종 잔기(H1)는 면역글로불린 고정 영역 또는 이의 일부, 알부민, 알부민-결합 잔기, PAS, HAP, 트랜스페린 또는 이의 단편, 폴리에틸렌 글리콜(PEG), 하이드록시에틸 전분(HES), PSA, 사람 만성 고나도트로핀의  $\beta$  소단위의 C-말단 웨بت아이드(CTP), 또는 이의 어떠한 조합을 포함한다. 다른 구현예에서, 이종 잔기는 FcRn 결합 파트너를 포함한다. 다른 구현예에서, 이종 잔기는 Fc 영역을 포함한다. 다른 구현예에서, 이종 잔기(H1)는 클리어런스 수용체(clearance receptor), 또는 이의 단편을 포함하며, 여기서 상기 클리어런스 수용체는 FVIII 단백질의 FVIII 클리어런스 수용체에 대한 결합을 차단한다. 일부 구현예에서, 여기서, 클리어런스 수용체는 저-밀도 지단백질 수용체-관련된 단백질 1(LRP1) 또는 이의 FVIII-결합 단편이다.

[0015] 일부 구현예에서, 임의의 FVIII 링커를 경유하여 FVIII 단백질에 융합된 제2의 이종 잔기는 면역글로불린 불변 영역(constant region) 또는 이의 부위, 알부민, 알부민-결합 폴리웨بت아이드, PAS, 사람 만성 고나도트로핀의  $\beta$  소단위의 C-말단 웨بت아이드(CTP), 폴리에틸렌 글리콜(PEG), 하이드록시에틸 전분(HES), 알부민-결합 소 분자, 또는 이의 어떠한 조합을 포함한다. 일 구현예에서, 제2의 이종 잔기(H2)는 FVIII 단백질의 반감기를 연장시킬 수

있다. 다른 구현예에서, 제2의 이종 잔기(H2)는 폴리펩타이드, 비-폴리펩타이드 잔기, 또는 둘 다를 포함한다. 다른 구현예에서, 제2의 이종 잔기(H2)는 면역글로불린 불변 영역 또는 이의 일부를 포함한다. 여전히 다른 구현예에서, 제2의 이종 잔기는 FcRn 결합 파트너를 포함한다. 여전히 다른 구현예에서, 제2의 이종 잔기는 제2의 Fc 영역을 포함한다.

[0016] 일부 구현예에서, VWF 링커를 통해 VWF 단백질에 융합된 제1의 이종 잔기 및 임의의 링커를 통해 FVIII 단백질에 융합된 제2의 이종 잔기(여기서, XTEN 서열은 성분 중의 하나에 융합되어 있다)는 서로 연합되어 있다. 일구현예에서, 제1의 폴리펩타이드와 제2의 폴리펩타이드 사이의 연합은 공유 결합이다. 다른 구현예에서, 제1의 이종 잔기와 제2의 이종 잔기 사이의 연합은 이황화물 결합이다. 다른 구현예에서, 제1의 이종 잔기는 FcRn 결합 파트너이고 제2의 이종 잔기는 FcRn 결합 파트너이다. 여전히 다른 구현예에서, 제1의 이종 잔기는 Fc 영역이고, 제2의 이종 잔기는 Fc 영역이다.

[0017] 특정의 구현예에서, FVIII 단백질은 FVIII 링커에 의해 제2의 이종 잔기에 연결된다. 일 구현예에서, 제2의 링커는 절단가능한 링커이다. 다른 구현예에서, FVIII 링커는 VWF 링커와 동일하다. 다른 구현예에서, FVIII 링커는 VWF 링커와는 상이하다.

[0018] 일부 측면에서, 본 발명의 키메라 분자는: (a) V-L1-X1-H1:H2-L2-X2-C; (b) V-X1-L1-H1:H2-L2-X2-C; (c) V-L1-X1-H1:H2-X2-L2-C; (d) V-X1-L1-H1:H2-X2-L2-C; (e) V-L1-X1-H1: H2-L2-C(X2); (f) V-X1-L1-H1:H2-L2-C(X2); (g) C-X2-L2-H2:H1-X1-L1-V; (h) C-X2-L2-H2:H1-L1-X1-V; (i) C-L2-X2-H2:H1-L1-X1-V; (j) C-L2-X2-H2:H1-L1-X1-V; (k) C(X2)-L2-H2:H1-X1-L1-V; 또는 (l) C(X2)-L2-H2:H1-L1-X1-V로부터 선택된 화학식을 포함하며; 여기서, V는 VWF 단백질이고; L1은 VWF 링커이며; L2는 임의의 FVIII 링커이고; H1은 제1의 이종 잔기이며; H2는 제2의 이종 잔기이고; X1은 XTEN 서열이며; X2는 임의의 XTEN 서열이고; C는 FVIII 단백질이며; C(X2)는 XTEN 서열에 융합된 FVIII 단백질이고, 여기서, XTEN 서열은 서로 인접한 2개의 FVIII 아미노산 사이에 삽입되고; (-)은 웨პ타이드 결합 또는 하나 이상의 아미노산이며; (:)는 H1과 H2 사이의 공유 결합이다.

[0019] 다른 측면에서, 키메라 분자는: (a) V-L1-X1-H1: H2-L2-X2-C; (b) V-X1-L1-H1: H2-L2-X2-C; (c) V-L1-X1-H1: H2-X2-L2-C; (d) V-X1-L1-H1: H2-X2-L2-C; (e) V-L1-X1-H1: H2-L2-C(X2); (f) V-X1-L1-H1: H2-L2-C(X2); (g) C-X2-L2-H2: H1-X1-L1-V; (h) C-X2-L2-H2: H1-L1-X1-V; (i) C-L2-X2-H2:H1-L1-X1-V; (j) C-L2-X2-H2:H1-L1-X1-V; (k) C(X2)-L2-H2:H1-X1-L1-V; 또는 (l) C(X2)-L2-H2:H1-L1-X1-V로부터 선택된 화학식을 포함하고; 여기서, V는 VWF 단백질이고; L1은 VWF 링커이며; L2는 임의의 FVIII 링커이고; H1은 제1의 이종 잔기이며; H2는 제2의 이종 잔기이고; X1은 임의의 XTEN 서열이며; X2는 XTEN 서열이고; C는 FVIII 단백질이며; C(X2)는 XTEN 서열에 융합된 FVIII 단백질이고, 여기서, XTEN 서열은 서로 인접한 2개의 FVIII 아미노산 사이에 삽입되고; (-)은 웨პ타이드 결합 또는 하나 이상의 아미노산이며; (:)는 H1과 H2 사이의 공유 결합이다.

[0020] 본 발명의 키메라 분자에서, VWF 단백질은 내인성 VWF의 FVIII 단백질에 대한 결합을 억제하거나 방지할 수 있다.

[0021] 특정의 측면에서, 키메라 분자내 FVIII 단백질은 제3의 이종 잔기(H3)를 포함할 수 있다. 제3의 이종 잔기(H3)는 XTEN 서열일 수 있다. 다른 측면에서, FVIII 단백질은 제4의 이종 잔기(H4)를 포함한다. 제4의 이종 잔기(H4)는 XTEN 서열일 수 있다. 일부 측면에서, FVIII 단백질은 제5의 이종 잔기(H5)를 포함한다. 제5의 이종 잔기는 XTEN 서열일 수 있다. 다른 측면에서, FVIII 단백질은 제6의 이종 잔기(H6)를 포함한다. 제6의 이종 잔기는 XTEN 서열일 수 있다. 특정의 측면에서, 하나 이상의 제3의 이종 잔기(H3), 제4의 이종 잔기(H4), 제5의 이종 잔기(H5), 및 제6의 이종 잔기(H6)는 키메라 분자의 반감기를 연장시킬 수 있다. 다른 측면에서, 제3의 이종 잔기(H3), 제4의 이종 잔기(H4), 제5의 이종 잔기(H5), 및 제6의 이종 잔기(H6)는 FVIII의 C 말단 또는 N 말단에 연결되거나 FVIII 단백질의 2개의 아미노산 사이에 삽입될 수 있다. 여전히 다른 측면에서, 하나 이상의 제3의 이종 잔기, 제4의 이종 잔기, 제5의 이종 잔기, 및 제6의 이종 잔기는 약 42개 아미노산, 약 72개 아미노산, 약 108개 아미노산, 약 144개 아미노산, 약 180개 아미노산, 약 216개 아미노산, 약 252개 아미노산, 약 288개 아미노산, 약 324개 아미노산, 약 360개 아미노산, 약 396개 아미노산, 약 432개 아미노산, 약 468개 아미노산, 약 504개 아미노산, 약 540개 아미노산, 약 576개 아미노산, 약 612개 아미노산, 약 624개 아미노산, 약 648개 아미노산, 약 684개 아미노산, 약 720개 아미노산, 약 756개 아미노산, 약 792개 아미노산, 약 828개 아미노산, 약 836개 아미노산, 약 864개 아미노산, 약 875개 아미노산, 약 912개 아미노산, 약 923개 아미노산, 약 948개 아미노산, 약 1044개 아미노산, 약 1140개 아미노산, 약 1236개 아미노산, 약 1318개 아미노산, 약 1332개 아미노산, 약 1428개 아미노산, 약 1524개 아미노산, 약 1620개 아미노산, 약 1716개 아미노산, 약 1812개 아미노산, 약 1908개 아미노산, 또는 약 2004개 아미노산 중 하나 이상으로부터 선택된 길이를 포함한다. 예를

들면, 제3의 이종 잔기, 제4의 이종 잔기, 제5의 이종 잔기, 또는 제6의 이종 잔기의 XTEN 서열은 AE42, AE72, AE864, AE576, AE288, AE144, AG864, AG576, AG288, 또는 AG144로부터 선택될 수 있다. 보다 구체적으로, XTEN 서열은 서열 번호 39; 서열 번호 40; 서열 번호 47; 서열 번호 45; 서열 번호 43; 서열 번호 41; 서열 번호 48; 서열 번호 46, 서열 번호 44, 또는 서열 번호 42로부터 선택될 수 있다.

[0022] 일부 구현예에서, 키메라 단백질의 반감기는 야생형 FVIII보다 적어도 약 1.5배, 적어도 약 2배, 적어도 약 2.5배, 적어도 약 3배, 적어도 약 4배, 적어도 약 5배, 적어도 약 6배, 적어도 약 7배, 적어도 약 8배, 적어도 약 9배, 적어도 약 10배, 적어도 약 11배, 적어도 약 12배 더 길게 연장된다.

[0023] 본 개재내용은 또한 키메라 분자를 암호화하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드 세트 또는 이의 상보성 서열을 제공한다. 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트는 또한, PC5 또는 PC7을 암호화하는 폴리뉴클레오타이드 쇄를 포함할 수 있다.

[0024] 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트 및 당해 폴리뉴클레오타이드 또는 폴리뉴크레오타이드 세트에 작동적으로 연결된 하나 이상의 프로모터를 포함하는 벡터 또는 벡터의 세트가 또한 포함된다. 일부 구현예에서, 벡터 또는 벡터의 세트는 또한 PC5 또는 PC7을 암호화하는 추가의 폴리뉴클레오타이드를 추가로 포함할 수 있다.

[0025] 본 발명은 또한 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트 또는 벡터 또는 벡터의 세트를 포함하는 숙주 세포를 포함한다. 일 구현예에서, 숙주 세포는 포유동물 세포이다. 다른 구현예에서, 숙주 세포는 HEK293 세포, CHO 세포, 또는 BHK 세포로부터 선택된다.

[0026] 일부 측면에서, 본 발명은 본원에 개재된 키메라 분자, 당해 키메라 분자를 암호화하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트, 당해 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트를 포함하는 벡터 또는 벡터의 세트, 또는 본원에 개재된 숙주 세포, 및 약제학적으로 허용되는 담체를 포함하는 약제학적 조성물을 포함한다. 일 구현예에서, 조성을 속의 키메라 분자는 야생형 FVIII 단백질과 비교하여 연장된 반감기를 갖는다. 다른 구현예에서, 조성을 속의 키메라 분자의 반감기는 야생형 FVIII보다 적어도 약 1.5배, 적어도 약 2배, 적어도 약 2.5배, 적어도 약 3배, 적어도 약 4배, 적어도 약 5배, 적어도 약 6배, 적어도 약 7배, 적어도 약 8배, 적어도 약 9배, 적어도 약 10배, 적어도 약 11배, 적어도 약 12배로 연장된다.

[0027] 또한 이를 필요로 하는 피검자에게 유효량의 본원에 개재된 키메라 분자, 당해 키메라 분자를 암호화하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트, 본원에 개재된 벡터 또는 벡터의 세트, 본원에 개재된 숙주 세포, 또는 본원에 개재된 조성을 투여함을 포함하여, 출혈 사건의 빈도 또는 정도를 감소시키는 방법이 포함된다. 본 발명은 또한 유효량의 본원에 개재된 키메라 분자, 당해 키메라 분자를 암호화하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트, 본원에 개재된 벡터 또는 벡터의 세트, 본원에 개재된 숙주 세포, 또는 본원에 개재된 조성을 투여함을 포함하여, 이를 필요로 하는 피검자에서 출혈 사건의 발생을 예방하는 방법을 포함한다. 일 구현예에서, 출혈 사건은 출혈 응고 질환, 출혈성 관절증, 근육 출혈, 구강 출혈, 출혈, 근육내로의 출혈, 구강 출혈, 외상, 두개 외상(trauma capititis), 위장 출혈, 두개내 출혈, 복부내 출혈, 흉곽내 출혈, 골절, 중추신경계 출혈, 인후후강내 출혈, 복막후 출혈, 장요근 쉬쓰(iliopsoas sheath)내 출혈, 또는 이의 어떠한 조합으로부터 기인한다. 다른 구현예에서, 본원에 개재된 키메라 분자, 당해 키메라 분자를 암호화하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 세트, 본원에 개재된 벡터 또는 벡터의 세트, 본원에 개재된 숙주 세포, 또는 본원에 개재된 조성을 국소 투여, 안구내 투여, 비경구 투여, 척추강내 투여, 경막하 투여, 경구 투여, 또는 이의 어떠한 조합으로부터 선택된 경로에 의해 투여될 수 있다.

[0028] 본 개재내용은 또한 하나 이상의 숙주 세포를 본원에 개재된 폴리뉴클레오타이드로 또는 본원에 개재된 벡터로 형질감염시키는 단계 및 숙주 세포내에서 키메라 분자를 발현시키는 단계를 포함하여, 키메라 분자를 제조하는 방법을 포함한다. 당해 방법은 키메라 분자를 분리하는 단계를 추가로 포함한다. 일부 구현예에서, 키메라 분자의 FVIII 활성은 aPTT 검정 또는 ROTEM 검정에 의해 측정될 수 있다.

### 도면의 간단한 설명

[0029] 도 1은 2개의 폴리펩타이드 쇄를 포함하는 키메라 분자(FVIII-XTEN/VWF 이종이량체)의 예시적인 도해를 나타내며 상기 제1의 쇄는 트롬빈 절단가능한 VWF 링커를 통해 Fc 영역에 융합된 VWF 단백질(예를 들면, VWF의 D' 도메인 및 D3 도메인)을 포함하며 제2 쇄는 FVIII 링커를 통해 제2의 Fc 영역에 융합된 FVIII 단백질을 포함한다. FVIII 단백질은 FVIII의 다양한 도메인(domain)내 하나 이상의 XTEN을 포함한다.

도 2는 다양한 VWF 작제물을 나타내며, 각각의 각제물은 대조군(즉, VWF-052)을 제외하고는 트롬빈 절단가능한 VWF 링커를 통해 Fc 영역에 융합된 D' 도메인 및 D3 도메인을 포함한다. VWF-031은 L-V-P-R(서열 번호 25)의 트롬빈 절단 부위를 포함하는 48개 아미노산의 링커를 포함한다. VWF-034는 288개 아미노산을 갖는 XTEN 서열 및 L-V-P-R(서열 번호 25)의 트롬빈 절단 부위를 포함하는 35개 아미노산의 링커를 포함한다. VWF-035는 L-V-P-R(서열 번호 25)의 트롬빈 절단 부위를 포함하는 73개 아미노산의 링커를 포함한다. VWF-036은 L-V-P-R(서열 번호 25)의 트롬빈 절단 부위를 포함하는 98개 아미노산의 링커를 포함한다. VWF-039는 L-V-P-R(서열 번호 25)의 트롬빈 절단 부위를 포함하는 26개 아미노산의 VWF 링커 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 26개 아미노산의 VWF 링커를 포함한다. VWF-051은 A-L-R-P-R-V-V(서열 번호 26)의 트롬빈 절단 부위를 포함하는 54개 아미노산의 링커를 포함한다. VWF-052는 어떠한 트롬빈 절단 부위(대조군)의 부재하에서 48개 아미노산의 링커를 포함한다. VWF-054는 FVIII으로부터의 a1 영역을 포함하는 40개 아미노산의 VWF 링커를 포함한다. VWF-055는 FVIII으로부터의 a2 영역을 포함하는 34개 아미노산의 VWF 링커를 포함한다. VWF-056은 FVIII으로부터의 a3 영역을 포함하는 46개 아미노산의 VWF 링커를 포함한다.

도 3a는 VWF-Fc 기능 작제물, 즉, VWF-031, VWF-034, VWF-036, VWF-039, VWF-051, 및 VWF-052에 대한 RU의 단위의 포획 밀도의 함수로서 초당 공명 단위(RU/s)의 단위인 트롬빈-매개된 절단의 비율을 나타낸다. 도 3b는 VWF-Fc 융합 작제물, 즉, VWF-031, VWF-034, VWF-036, VWF-051, 및 VWF-052에 대한 RU의 단위의 포획 밀도의 함수로서 초당 공명 단위(RU/s)의 단위인 트롬빈-매개된 절단의 비율을 나타낸다. 이들 실시예에서, 각각의 VWF-Fc 융합 작제물은 다양한 밀도에서 포획되고 후속적으로 고정된 농도의 사람 알파-트롬빈에 노출된다. 도 3a 및 도 3b에서 각각의 곡선의 기울기는 각각의 작제물에 대한 트롬빈 절단에 대한 민감성을 직접적으로 반영한다.

도 4a는 VWF-Fc 기능 작제물, 즉, VWF-054, VWF-055, 및 VWF-056에 대한 RU의 단위의 포획 밀도의 함수로서 초당 공명 단위(RU/s)의 단위인 트롬빈-매개된 절단의 비율을 나타낸다. 도 4b는 VWF-Fc 기능 작제물, 즉, VWF-031, VWF-039, VWF-054, VWF-055, 및 VWF-056에 대한 RU의 단위의 포획 밀도의 함수로서 초당 공명 단위(RU/s)의 단위인 트롬빈-매개된 절단의 비율을 나타낸다. 이들 실시예에서, 각각의 VWF-Fc 융합 작제물은 다양한 밀도에서 포획되고 후속적으로 고정된 농도의 사람 알파-트롬빈에 노출된다. 도 4a 및 도 4b에서 각각의 곡선의 기울기는 각각의 작제물에 대해 트롬빈 절단에 대한 민감성을 직접적으로 반영한다.

도 5는 트롬빈-매개된 절단에 대한 다양한 VWF-Fc 작제물, VWF-031, VWF-034, VWF-036, VWF-039, VWF-051, VWF-052, VWF-054, VWF-055, 및 VWF-056의 민감성을 측정하기 위한 선형 회귀 분석의 결과를 나타낸다. 값은 초기의 역수의 단위로 나타내며 도 3 및 도 4에 나타낸 곡선의 기울기를 반영한다. 2개의 상이한 작제물의 상대적인 민감성은 이들 각각의 기울기의 몫으로부터 기원한다. 기울기<sub>VWF-039</sub>/기울기<sub>VWF-031</sub>는 71이고, 이는, VWF-Fc 융합 작제물 VWF-039가 VWF-031보다 트롬빈-매개된 절단에 대해 71배 더 민감성임을 나타낸다. 기울기<sub>VWF-055</sub>/기울기<sub>VWF-031</sub>는 65이고, 기울기<sub>VWF-051</sub>/기울기<sub>VWF-031</sub>는 1.8이다.

도 6은 전혈 ROTEM 검정에 의해 측정된 HemA 환자에서 다양한 키메라 분자의 응고 시간을 나타낸다. FVII155/VWF-031은 2개의 폴리펩타이드 쇄를 포함하고, 제1 쇄는 Fc 영역에 융합된 BDD FVIII를 포함하고 제2 쇄는 최소의 트롬빈 절단 부위(즉, L-V-P-R (서열 번호 25))를 통해 Fc 영역에 융합된 VWF의 D' 도메인 및 D3 도메인을 포함한다. FVII155/VWF-039는 2개의 폴리펩타이드 쇄를 포함하고, 제1 쇄는 Fc 영역에 융합된 BDD FVIII를 포함하고 제2 쇄는 L-V-P-R (서열 번호 25) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 VWF 링커를 통해 Fc 영역에 융합된 VWF의 D' 도메인 및 D3 도메인을 포함한다. FVII155/VWF-055는 2개의 폴리펩타이드 쇄를 포함하고, 제1 쇄는 Fc 영역에 융합된 BDD FVIII를 포함하고 제2 쇄는 FVIII로부터의 a2 영역을 포함하는 VWF 링커를 통해 Fc 영역에 융합된 VWF의 D' 도메인 및 D3 도메인을 포함한다.

도 7은 대표적인 FVIII-VWF 이종이량체 및 FVIII169, FVIII286, VWF057, VWF059, and VWF062 작제물의 도해를 나타낸다. 예를 들면, FVIII169 작제물은 Fc 영역에 융합된 R1648A 치환을 지닌 B 도메인 결실된 FVIII 단백질을 포함하며, 여기서, XTEN 서열(예를 들면, AE288)은 성숙한 전장의 FVIII (A1-a1-A2-a2-288XTEN-a3-A3-C1-C2-Fc)에 상응하는 745번 아미노산에서 삽입된다. FVIII286 작제물은 Fc 영역에 융합된 R1648 치환을 지닌 B 도메인 결실된 FVIII 단백질을 포함하며, 여기서, XTEN 서열(예를 들면, AE288)은 FVIII 와 Fc (A1-a1-A2-a2-288XTEN-a3-A3-C1-C2-a2-Fc)의 사이에 추가의 a2 영역을 지닌, 성숙한 전장의 FVIII에 상응하는 745번 아미노산에 삽입된다. VWF057은 VWF 링커를 통해 Fc 영역에 연결된 VWF 단백질의 D'D3 도메인(D'D3 도메인내에 2개의 아미노산 치환, 즉, C336A 및 C379A를 지님)을 포함하는 VWF-Fc 융합 작제물이며, 이는 LVPR 트롬빈 부위 ("LVPR") 및 GS 링커("GS")를 포함하고, 여기서, XTEN 서열(즉, 144XTEN)은 D'D3 도메인과 VWF 링커(D'D3-

144XTEN-GS+LVPR-Fc) 사이에 삽입된다. VWF059는 VWF 링커로서 산성 영역 2(a2) 영역을 통해 Fc 영역에 연결된 VWF 단백질의 D'D3 도메인(D'D3 도메인내에 2개의 아미노산 치환, 즉., C336A 및 C379A를 지님)을 포함하는 VWF-Fc 융합 작제물이며, 여기서 XTEN 서열은 D'D3 도메인과 VWF 링커 사이에 삽입된다. VWF062는 Fc 영역에 연결된 VWF 단백질의 D'D3 도메인(D'D3 도메인내에 2개의 아미노산 치환, 즉., C336A 및 C379A를 지님)을 포함하는 VWF-Fc 융합 작제물이며, 여기서 XTEN 서열은 D'D3 도메인과 Fc 영역(D'D3-144XTEN-Fc) 사이에 삽입된다.

도 8은 HemA 마우스 꼬리 클립 모델에서 B 도메인 결실된 FVIII("SQ BDD FVIII" 또는 "BDD-rFVIII") 또는 비히클 대조군과 비교하여, FVIII-XTEN-Fc/D'D3-링커-Fc 이종이량체(즉, FVIII169/VWF034, FVIII169/VWF059, 및 FVIII169/VWF057)의 급성 효능을 나타낸다. BDD-rFVIII는 원형으로 나타내는 반면 FVIII169/VWF034는 삼각형으로 나타내고, FVIII169/VWF059는 삼각형으로 나타내며, FVIII169/VWF057은 중공 원(hollow circle)으로 나타내고, 비히클은 역전된 삼각형으로 나타낸다. VWF034는 Fc 영역에 VWF 링커를 통해 융합된 VWF의 D' 도메인 및 D3 도메인을 포함하는 VWF-Fc 융합 작제물이며, 이는 LVPR을 포함하고, 여기서, XTEN 서열 (즉, 288XTEN)은 D'D3 도메인과 VWF 링커(D'D3-288XTEN-LVPR-Fc) 사이에 삽입된다. FVIII169, VWF059, 및 VWF057의 작제물 세부사항은 본원의 어느곳에 나타나 있다. 각각의 처리 그룹에서 75 IU/kg의 작제물을 투여한 후 마우스의 중간 혈액 손실(uL)은 수평선으로 나타낸다

## 발명을 실시하기 위한 구체적인 내용

### [0030] 발명의 상세한 설명

본 발명은 XTEN 서열 및 이종 잔기, 예를 들면, 반감기를 연장하는 잔기를 지닌 VWF 단백질 또는 FVIII 단백질을 연결하는 트롬빈 절단가능한 링커를 포함하는 키메라 분자에 관한 것이다. 본 발명은 또한 2개의 폴리펩타이드 쇄를 포함하는 키메라 분자를 제공하며, 제1 쇄는 이종 잔기에 융합된 VWF 단백질을 포함하고, 제2 쇄는 FVIII 단백질 및 제2의 이종 잔기를 포함하며, 여기서 키메라 분자는 제1 또는 제2의 폴리펩타이드 쇄내에 XTEN 서열을 포함하고, 여기서 VWF 단백질 또는 FVIII 단백질(또는 둘 다)는 VWF 링커 또는 FVIII 링커(또는 둘 다)를 통해 이종 잔기에 융합된다. 트롬빈 절단가능한 링커(VWF 링커 또는 FVIII 링커)는, 트롬빈이 용이하게 이용 가능한 손상 부위에서 트롬빈에 의해 효율적으로 절단될 수 있다. 예시적인 키메라 분자는 본 상세한 설명 및 도면에 설명되어 있다. 일부 구현예에서, 본 발명은 예를 들면, 도 1 내지 7에 설정된 구조를 갖는 키메라 분자에 관한 것이다. 다른 구현예에서, 본 발명은 본원에 개재된 키메라 분자를 암호화하는 폴리뉴클레오타이드에 관한 것이다.

[0032] 명세서 및 특허청구범위의 명확한 이해를 제공하기 위하여, 다음의 정의를 하기 제공한다.

### [0033] 1. 정의

[0034] 용어 단수("a" 또는 "an")의 실체는 하나 이상의 실체를 말하는데; 예를 들면, "뉴클레오타이드 서열"은 하나 이상의 뉴클레오타이드 서열을 나타내는 것으로 이해된다. 자체로서, 용어 "하나("a"(또는 "an")), "하나 이상" 및 "적어도 하나"는 본원에서 상호교환적으로 사용될 수 있다.

[0035] 용어 "약"은 대략, 거의, 주변, 또는 이의 범위를 의미하기 위해 본원에서 사용된다. 용어 "약"이 수치 범위와 관련하여 사용되는 경우, 이는 설정된 수치의 상 하 경계부를 연장함으로써 당해 범위를 변형시킨다. 일반적으로, 용어 "약"은 기술된 값을 상 하로 10 퍼센트의 변형까지 수치의 상 하를 변형시키기 위해 본원에서 사용된다.

[0036] 용어 "폴리뉴클레오타이드" 또는 "뉴클레오타이드"는 단수의 핵산 및 다수의 핵산을 포함하는 것으로 의도되며, 분리된 핵산 분자 또는 작제물, 예를 들면, 전령 RNA(mRNA) 또는 플라스미드 DNA(pDNA)를 말한다. 특정의 구현 예에서, 폴리뉴클레오타이드는 통상의 포스포디에스테르 결합 또는 비-통상적인 결합(예를 들면, 웨პ타이드 핵산(PNA)에서 발견된 것과 같은 아미드 결합)을 포함한다. 용어 "핵산"은 폴리뉴클레오타이드에 존재하는 어떠한 하나 이상의 핵산 분절, 예를 들면, DNA 또는 RNA 단편을 말한다. "분리된" 핵산 또는 폴리뉴클레오타이드는 핵산 분자, DNA 또는 RNA를 의도하며, 이는 이의 천연의 환경으로부터 제거된다. 예를 들면, 백터내에 함유된 인자 VIII 폴리펩타이드를 암호화하는 재조합체 폴리뉴클레오타이드는 본 발명의 목적을 위해 분리되는 것으로 고려된다. 분리된 폴리뉴클레오타이드의 추가의 예는 이종 숙주 세포 속에 유지된 재조합체 폴리뉴클레오타이드 또는 용액 속에서 다른 폴리뉴클레오타이드로부터 정제된(부분적으로 또는 실질적으로) 재조합체 폴리뉴클레오타이드를 포함한다. 분리된 RNA 분자는 본 발명의 폴리뉴클레오타이드의 생체내 또는 시험관내 RNA 전사체를 포함한다. 본 발명에 따른 분리된 폴리뉴클레오타이드 또는 핵산은 또한 합성적으로 생산된 이러한 분자를 포함한다. 또한, 폴리뉴클레오타이드 또는 핵산은 프로모터, 인핸서(enhaner), 리보소ーム 결합 부위, 또는 전사 종결

시그날과 같은 조절 성분을 포함할 수 있다.

[0037] 본원에 사용된 것으로서, "암호화 영역" 또는 "암호화 서열"은 아미노산으로 해독가능한 코돈으로 이루어진 폴리뉴클레오타이드의 일부이다. "종결 코돈"(TAG, TGA, 또는 TAA)이 아미노산으로 전형적으로 해독되지 않지만, 이는 암호화 영역의 부분인 것으로 고려될 수 있으나, 어떠한 플랜킹 서열(flanking sequence), 예를 들면, 프로모터, 리보소ーム 결합 부위, 전사 터미네이터, 인트론 등은 암호화 영역의 부분이 아니다. 암호화 영역의 경계는 전형적으로 생성되는 폴리펩타이드의 아미노 말단을 암호화하는 5' 말단에서의 출발 코돈, 및 생성되는 폴리펩타이드의 카복실 말단을 암호화하는, 3' 말단에서의 해독 정지 코돈에 의해 결정된다. 본 발명의 2개 이상의 암호화 영역은 단일의 폴리뉴클레오타이드 작제물내, 예를 들면, 단일 벡터 상에, 또는 별개의 폴리뉴클레오타이드 작제물, 예를 들면, 별도의(상이한) 벡터 상에 존재할 수 있다. 이는, 이후에 단일 벡터가 단지 1개의 암호화 영역을 함유할 수 있거나 2개 이상의 암호화 영역을 포함할 수 있음을 수반하는데, 예를 들면, 단일 벡터는 하기 기술된 바와 같이 키메라 분자의 제1의 폴리펩타이드 쇄 및 제2의 폴리펩타이드 쇄를 별도로 암호화할 수 있다. 또한, 본 발명의 벡터, 폴리뉴클레오타이드, 또는 핵산은 본 발명의 키메라 분자를 암호화하는 핵산에 융합되거나 융합되지 않은 이종 암호화 영역을 암호화할 수 있다. 이종 암호화 영역은 분비성 시그날 웨بت아이드 또는 이종 기능성 도메인과 같은 특수화된 성분 또는 모티프를 포함하나, 이에 한정되지 않는다.

[0038] 포유동물 세포에 의해 분비된 특정의 단백질은, 조면 소포체를 가로질러 성장하는 단백질 쇄의 배출이 개시되면 성숙한 단백질로부터 절단되는 분비성 시그날 웨بت아이드와 연합되어 있다. 당해 분야의 통상의 기술자들은, 단일의 웨بت아이드가 일반적으로 폴리펩타이드의 N-말단에 융합되어 있으며, 완전하거나 "전장"의 폴리펩타이드로부터 절단되어 분비되거나 "성숙한" 형태의 폴리펩타이드를 생산함을 인지하고 있다. 특정의 구현예에서, 천연의 시그날 웨بت아이드, 예를 들어, FVIII 시그날 웨بت아이드 또는 VWF 시그날 웨بت아이드가 사용되거나, 이와 작동적으로 연합된 폴리펩타이드의 분비를 지시하는 능력을 보유한 이러한 서열의 기능성 유도체가 사용된다. 대안적으로, 이종 포유동물 시그날 웨بت아이드, 예를 들면, 사람 조직 플라스미노겐 활성인자(TPA) 또는 마우스 ?-글루쿠로니다제 시그날 웨بت아이드, 또는 이의 기능성 유도체가 사용될 수 있다.

[0039] 용어 "하부(downstream)"는 참조 뉴클레오타이드 서열에 대해 3'에 위치하는 뉴클레오타이드 서열을 말한다. 특정의 구현예에서, 하부 뉴클레오타이드 서열은 전사 출발점을 따라가는 서열을 말한다. 예를 들어, 유전자의 해독 개시 코돈은 전사의 출발 부위의 하부에 위치한다.

[0040] 용어 "상부(upstream)"는 참조 뉴클레오타이드 서열에 대해 5'에 위치하는 뉴클레오타이드 서열을 말한다. 특정의 구현예에서, 상부 뉴클레오타이드 서열은 전사의 암호화 영역 또는 출발점의 5' 부위에 위치하는 서열에 관한 것이다. 예를 들면, 대부분의 프로모터는 전사의 출발 지점의 상부에 위치한다.

[0041] 본원에 사용된 것으로서, 용어 "조절 영역"은 암호화 영역의 상부(5' 비-암호화 서열), 내부, 또는 하부(3' 비-암호화 서열)에 위치한 뉴클레오타이드 서열을 말하며, 이는 연합된 암호화 영역의 전사, RNA 프로세싱, 안전성, 또는 해독에 영향을 미친다. 조절 영역은 프로모터, 해독 리더 서열, 인트론, 폴리아데닐화 인지 서열, RNA 프로세싱 부위, 효과기 결합 부위 및 스템-루프 구조(stem-loop structure)를 포함할 수 있다. 암호화 영역이 진핵 세포내에서의 발현용으로 의도된 경우, 폴리아데닐화 시그날 및 전사 종결 서열은 일반적으로 암호화 서열에 대해 3'로 위치할 것이다.

[0042] 유전자 생성물, 예를 들면, 폴리펩타이드를 암호화하는 폴리뉴클레오타이드는 하나 이상의 암호화 영역에 작동적으로 연합된 프로모터 및/또는 다른 전사 또는 해독 조절 성분을 포함할 수 있다. 작동가능한 연합에서 유전자 생성물, 예를 들면, 폴리펩타이드에 대한 암호화 영역은 조절 영역(들)의 영향 또는 조절 하에 유전자 생성물을 발현하도록 하는 방식으로 하나 이상의 조절 영역과 연합된다. 예를 들어, 암호화 영역 및 프로모터는, 프로모터 기능의 도입이 암호화 영역에 의해 암호화된 유전자 생성물을 암호화하는 mRNA의 전사를 초래하는 경우, 및 프로모터와 암호화 영역 사이의 연결의 특성이 유전자 생성물의 발현을 지시하거나 전사될 DNA 주형의 능력을 방해하는 프로모터의 능력을 방해하지 않는 경우, "작동적으로 연합되어" 있다. 프로모터 이외의 다른 전사 조절 성분, 예를 들면, 인핸서, 프로모터, 리프레서, 및 전사 종결 시그날은 또한 암호화 영역과 작동적으로 연합되어 유전자 생성물 발현을 지시할 수 있다.

[0043] 다양한 전사 조절 영역은 당해 분야의 숙련가에게 공지되어 있다. 이들은 사이토메갈로바이러스(인트론-A와 접합된 인터미디어트 얼리 프로모터(immediate early promoter)), 시미안 바이러스 40(얼리 프로모터), 및 레트로바이러스(예: 로우스 사코마 바이러스(Rous sarcoma virus))로부터의 프로모터 및 인핸서 분절과 같은, 그러나 이에 한정되지 않는 척추동물 세포내에서 기능하는 전사 조절 영역을 포함하나, 이에 한정되지 않는다. 다른 전사 조절 영역은 액틴, 열 쇼크 단백질, 소 성장 호르몬 및 토키 ?-글로빈과 같은 척추동물 유전자로부터 기원한

것들, 및 진핵 세포내에서 유전자 발현을 조절할 수 있는 다른 서열을 포함한다. 추가의 적합한 전사 조절 영역은 조직-특이적인 프로모터 및 또한 인핸서 및 램포카인-유도성 프로모터 (예를 들면., 인터페론 또는 인터루킨에 의해 유도가능한 프로모터)를 포함한다.

[0044] 유사하게, 다양한 해독 조절 성분이 당해 분야의 통상의 기술자에게 공지되어 있다. 이들은 리보소ーム 결합 부위, 해독 개시 및 종결 코돈, 및 피코르나바이러스로부터 기원한 성분(특히 내부 리보소ーム 도입 부위 또는 CITE 서열로 또한 언급된 IRES)을 포함하나, 이에 한정되지 않는다.

[0045] 본원에 사용된 것으로서 용어 "발현"은, 폴리뉴클레오타이드가 유전자 생성물, 예를 들면, RNA 또는 폴리펩타이드를 생산하는 공정을 말한다. 이는 폴리뉴클레오타이드의 전령 RNA(mRNA), 전달 RNA(tRNA), 소 헤어핀(small hairpin) RNA(shRNA), 소 방해 RNA(siRNA) 또는 어떠한 다른 RNA 생성물로의 전사, 및 mRNA의 폴리펩타이드내로의 해독을 포함하나, 이에 한정되지 않는다. 발현은 "유전자 생성물"을 생산한다. 본원에 사용된 것으로서, 유전자 생성물은 핵산, 예를 들면, 유전자의 전사에 의해 생산된 전령 RNA, 또는 전사체로부터 해독된 폴리펩타이드일 수 있다. 본원에 기술된 유전자 생성물은 전사 후 변형, 예를 들면, 폴리아데닐화 또는 스플라이싱(splicing)을 지닌 핵산, 또는 전사 후 변형, 예를 들면, 다른 단백질 소단위, 또는 단백질분해성 절단과 함께, 메틸화, 글리코실화, 및 지질의 첨가를 추가로 포함한다.

[0046] "벡터"는 핵산의 숙주 세포내로의 클로닝 및/또는 전달을 위한 어떠한 비히클을 말한다. 벡터는, 다른 핵산 분절이 부착하여 부착된 분절의 복제를 가져오도록 할 수 있는 레플리콘(replicon)일 수 있다. "레플리콘"은 생체내에서 자가 복제 단위로서 기능하는, 즉, 이의 자체 조절하에 복제할 수 있는 어떠한 유전 성분(예를 들면, 플라스미드, 파아지, 코스미드, 염색체, 바이러스)를 말한다. 용어 "벡터"는 시험관내, 생체외 또는 생체내에서 핵산을 세포내로 도입하기 위한 바이러스 및 비바이러스 비히클 둘 다를 포함한다. 예를 들면, 플라스미드, 변형된 진행 바이러스, 또는 변형된 세균 바이러스를 포함하는, 다수의 벡터가 공지되어 있으며 당해 분야에서 사용되고 있다. 폴리뉴클레오타이드의 적합한 벡터내로의 삽입은 적절한 폴리뉴클레오타이드 단편을 상보성의 점착성 말단(complementary cohesive termini)을 갖는 선택된 벡터내로 연결시킴으로써 달성할 수 있다.

[0047] 벡터는 당해 벡터내로 혼입된 세포의 선택 또는 확인을 제공하는 선택가능한 마커 또는 리포터를암호화하도록 가공될 수 있다. 선택가능한 마커 또는 리포터의 발현은 벡터 내에 함유된 다른 암호화 영역을 혼입하여 발현하는 숙주 세포를 확인 및/또는 선택하도록 한다. 당해 분야에 공지되고 사용된 선택가능한 마커 유전자의 예는 다음을 포함한다: 암피실린, 스트렙토마이신, 젠타마이신, 가나마이신, 하이그로마이신, 비알라포스 제초제(bialaphos herbicide), 셀폰아미드 등; 및 표현형 마커로서 사용된 유전자, 즉, 안토시아닌 조절 유전자, 이소펜타닐 트랜스퍼라제 유전자 등. 당해 분야에 공지되고 사용된 리포터의 예는 다음을 포함한다: 루시퍼라제(Luc), 녹색 형광성 단백질(GFP), 클로람페니콜 아세틸트랜스퍼라제(CAT), -갈락토시다제(LacZ), -글루쿠로니다제(Gus), 등. 선택가능한 마커는 또한 리포터인 것으로 고려될 수 있다.

[0048] 용어 "플라스미드"는 세포의 중심 물질대사의 일부가 아니며, 일반적으로 원형의 이본체 DNA 분자의 형태인 유전자를 흔히 수반하는 염색체외 성분을 말한다. 이러한 성분은 어느 공급원으로부터도 기원한 일본체 또는 이본체 DNA 또는 RNA의 자가 복제하는 서열, 계놈 통합 서열, 파아지 또는 뉴클레오타이드 서열, 선형, 환형, 또는 슈퍼코일(supercoiled)일 수 있으며, 여기서 다수의 뉴클레오타이드 서열은 적절한 3' 해독되지 않은 서열과 함께 선택된 유전자 생성물에 대한 프로모터 단편과 DNA 서열을 세포내로 도입시킬 수 있는 유일한 구조물로 결합되거나 조합된다.

[0049] 사용될 수 있는 진핵 세포 바이러스 벡터는 아데노바이러스 벡터, 레트로바이러스 벡터, 아데노-관련 바이러스 벡터, 폭스바이러스, 예를 들면, 박시니아 바이러스 벡터, 바클로바이러스 벡터, 또는 헤르페스바이러스 벡터를 포함하나, 이에 한정되지 않는다. 비-바이러스 벡터는 플라스미드, 리포좀, 전기적으로 하전된 지질(사이토펩틴), DNA-단백질 복합체, 및 생물중합체를 포함한다.

[0050] "클로닝 벡터"는 연속으로 복제하는 핵산의 단위 길이이며 다른 핵산 분절이 부착하여 부착된 분절의 복제를 가져오는 플라스미드, 파아지 또는 코스미드와 같은 복제 오리진을 포함하는 "레플리콘"을 말한다. 특정의 클로닝 벡터는 하나의 세포형, 예를 들면, 세균내에서 복제할 수 있고 다른, 예를 들면, 진핵 세포내에서 발현할 수 있다. 클로닝 벡터는 전형적으로 목적한 핵산 서열의 삽입을 위한 벡터 및/또는 하나 이상의 다른 클로닝 부위를 포함하는 세포의 선택에 사용될 수 있는 하나 이상의 서열을 포함한다.

[0051] 용어 "발현 벡터"는 숙주 세포내로 삽입 후 삽입된 핵산 서열의 발현이 가능하도록 설계된 비히클을 말한다. 삽입된 핵산 서열은 위에서 기술한 바와 같은 조절 영역과 작동 가능하게 연합하도록 위치한다.

- [0052] 벡터는 당해 분야에 잘 공지된 방법, 예를 들면, 형질감염, 전기천공, 미세주사, 형질도입, 세포 융합, DEAE 텍스트란, 인산칼슘 침전, 지질감염(리소좀 융합), 유전자 총의 사용, 또는 DNA 벡터 전달인자의 사용에 의해 숙주 세포내로 도입된다.
- [0053] 본원에 사용된 것으로서, "배양" "배양물에" 및 "배양하는"은 세포 성장 또는 분열을 허용하는 시험관내 조건에서 세포를 항온처리하거나 살아있는 상태로 세포를 유지함을 의미한다. 본원에 사용된 것으로서 "배양된 세포"는 시험관내에서 증식되는 세포를 의미한다.
- [0054] 본원에 사용된 것으로서, 용어 "폴리펩타이드"는 단수의 "폴리펩타이드" 뿐만 아니라 복수의 "폴리펩타이드"를 포함하는 것으로 의도되며, 아미드 결합(펩타이드 결합으로 또한 공지됨)에 의해 선형 연결된 단량체(아미노산)으로 구성된 분자를 말한다. 용어 "폴리펩타이드"는 2개 이상의 아미노산의 어떠한 쇄 또는 쇄들을 말하며, 생성물의 구체적인 길이를 말하지는 않는다. 따라서, 2개 이상의 아미노산의 쇄 또는 쇄들을 언급하기 위해 사용된 웹타이드, 디펩타이드, 트리펩타이드, 올리고펩타이드, "단백질", "아미노산 쇄", 또는 어떠한 다른 용어는 "폴리펩타이드"의 정의내에 포함되며, 용어 "폴리펩타이드"는 이들 형태 중의 어느 것 대신에 또는 이와 상호교환적으로 사용될 수 있다. 용어 "폴리펩타이드"는 또한 글리코실화, 아세틸화, 인산화, 아미드화, 공지된 보호/차단 그룹에 의한 유도체화, 단백질분해적 절단, 또는 비-천연적으로 존재하는 아미노산에 의한 변형을 제한없이 포함하는, 폴리펩타이드의 발현 후 변형의 생성물을 말하는 것으로 의도된다. 폴리펩타이드는 천연의 생물학적 공급원으로부터 기원할 수 있거나 재조합체 기술로 생산될 수 있으나, 지정된 핵산 서열로부터 필수적으로 해독되지 않는다. 이는 화학 합성을 포함하는, 어떠한 방식으로도 생성될 수 있다.
- [0055] "분리된 폴리펩타이드 또는 단편, 변이체, 또는 이의 유도체"는 이의 천연 환경(natural milieu)에서 존재하지 않는 폴리펩타이드를 말한다. 특별한 수준의 정제는 요구되지 않는다. 예를 들어, 분리된 폴리펩타이드는 이의 원래의 또는 천연 환경으로부터 단순하게 제거될 수 있다. 재조합적으로 생산된 폴리펩타이드 및 숙주 세포내에서 발현된 단백질은 어떠한 적합한 기술에 의해서 별도로, 분획화되거나, 부분적으로 또는 실질적으로 정제된 천연의 또는 재조합체 폴리펩타이드와 같이, 본 발명의 목적을 위해 분리되는 것으로 고려된다.
- [0056] 또한 폴리펩타이드의 단편 또는 변이체, 및 이의 어떠한 조합도 본 발명에 포함된다. 본 발명의 폴리펩타이드 결합 도메인 또는 결합 분자를 언급하는 경우 용어 "단편" 또는 "변이체"는 참조 폴리펩타이드의 적어도 일부의 특성(예를 들면, FcRn 결합 도메인 또는 Fc 변이체에 대한 FcRn 결합 친화성, FVIII 변이체에 대한 응고 활성, 또는 VWF 단백질에 대한 FVIII 결합 활성)을 보유한 어떠한 폴리펩타이드도 포함한다. 폴리펩타이드의 단편은 본원의 어디에서 논의된 특이적인 항체 단편 이외에, 단백질분해성 단편, 및 결실 단편을 포함하나, 천연적으로 존재하는 전장의 폴리펩타이드(또는 성숙한 폴리펩타이드)는 포함하지 않는다. 본 발명의 폴리펩타이드 결합 도메인 또는 결합 분자의 변이체는 위에서 논의된 바와 같은 단편, 및 또한 아미노산 치환, 결실, 또는 삽입으로 인하여 변경된 아미노산 서열을 지닌 폴리펩타이드를 포함한다. 변이체는 천연적으로 또는 비-천연적으로 존재할 수 있다. 비-천연적으로 존재하는 변이체는 당해 분야에 공지된 돌연변이유발 기술을 사용하여 생산할 수 있다. 변이체 폴리펩타이드는 보존적 또는 비-보존적 아미노산 치환, 결실 또는 침가를 포함할 수 있다.
- [0057] 본원에 사용된 것으로서 용어 "VWF 단편" 또는 "VWF 단편"은 FVIII와 상호작용하며 전장의 VWF에 의해 FVII에 일반적으로 제공된 적어도 하나 이상의 특성, 예를 들면, FVIIIa으로의 미성숙 활성화를 예방하고/하거나, 미성숙 단백질분해를 방지하고/하거나, 미성숙 정화를 초래할 수 있는 인지질 막과의 연합을 예방하고/하거나, 네이크드(naked) FVIII에 결합할 수 있으나 VWF-결합된 FVII에 결합할 수 없는 FVIII 정화 수용체에 대한 결합을 예방하고/하거나 FVIII 중쇄 및 경쇄 상호작용을 안정화시키는 특성을 보유하는 어떠한 VWF 단편도 의미한다. 특수한 구현예에서, 본원에 사용된 것으로서 "VWF 단편"은 VWF 단백질의 D' 도메인 및 D3 도메인을 포함하지만, VWF 단백질의 A1 도메인, A2 도메인, A3 도메인, D4 도메인, B1 도메인, B2 도메인, B3 도메인, C1 도메인, C2 도메인, 및 CK 도메인을 포함하지 않는다.
- [0058] 본원에 사용된 것으로서 용어 "반감기 제한 인자" 또는 "FVIII 반감기 제한 인자"는 야생형 FVIII (예를 들면, ADVATE<sup>®</sup> 또는 REFACTO<sup>®</sup>)과 비교하여 FVIII 단백질의 반감기를 1.5배 이하 또는 2배 이하로 방지하는 인자를 나타낸다. 예를 들면, 전장의 또는 성숙한 VWF는 시스템으로부터 정화될 FVIII 및 VWF 복합체를 하나 이상의 VWF 정화 경로에 의해 유도함으로써 FVIII 반감기 제한 인자로서 작용할 수 있다. 하나의 예에서, 내인성 VWF는 FVIII 반감기 제한 인자이다. 다른 예에서, FVIII 단백질에 비-공유결합으로 결합된 전장의 재조합체 VWF 분자는 FVIII-반감기 제한 인자이다.
- [0059] 본원에 사용된 것으로서 용어 "내인성 VWF"는 혈장 속에 천연적으로 존재하는 VWF 분자를 나타낸다. 내인성 VWF

분자는 다량체일 수 있지만, 단량체 또는 이량체일 수 있다. 혈장 속의 내인성 VWF는 FVIII에 결합하여 FVIII과 비-공유결합성 복합체를 형성한다.

[0060] "보존적 아미노산 치환"은, 아미노산 잔기가 유사한 측쇄를 갖는 아미노산 잔기로 치환된 것이다. 유사한 측쇄를 갖는 아미노산 잔기의 계열은 염기성 측쇄(예를 들면, 라이신, 아르기닌, 히스티딘), 산성 측쇄(예를 들면, 아스파르트산, 글루탐산), 하전되지 않는 극성 측쇄(예를 들면, 글리신, 아스파라긴, 글루탐산, 세린, 트레오닌, 타이로신, 시스테인), 비극성 측쇄(예를 들면, 알라닌, 발린, 루이신, 이소루이신, 프롤린, 페닐알라닌, 메티오닌, 트립토판), 베타-측쇄된 측쇄(예를 들면, 트레오닌, 발린, 이소루이신) 및 방향족 측쇄 (예를 들면, 타이로신, 페닐알라닌, 트립토판, 히스티딘)을 포함하여, 당해 분야에 확인되어 있다. 따라서, 폴리펩타이드내 아미노산이 동일한 측쇄 계열로부터의 다른 아미노산으로 치환되는 경우, 당해 치환은 보존적인 것으로 고려된다. 다른 구현예에서, 아미노산의 스트링(string)은 측쇄 계열 구성원의 순서 및/또는 구성에 있어서 상이한 구조적으로 유사한 스트링으로 보존적으로 치환될 수 있다

[0061] 당해 분야에 알려진 바와 같이, 2개의 폴리펩타이드 사이의 "서열 동질성"은 하나의 폴리펩타이드의 아미노산 서열을 제2의 폴리펩타이드의 서열과 비교함으로써 측정한다. 본원에서 논의하는 경우, 어떠한 특수한 폴리펩타이드가 다른 폴리펩타이드에 대해 적어도 약 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, 또는 100% 동일한지의 여부는 BESTFIT 프로그램(Wisconsin Sequence Analysis Package, 유닉스용 버전 8, 제조원: 미국 위스콘신주 53711 매디슨 사이언스 드라이드 575 유니버시티 리서치 파크 소재의 Genetics Computer Group)와 같은, 그러나 이에 한정되지 않는 당해 분야에 공지된 방법 및 컴퓨터 프로그램/소프트웨어를 사용하여 측정할 수 있다. BESTFIT는 스미쓰(Smith) 및 워터맨(Waterman)의 국소 상동성 알고리즘[참조: Advances in Applied Mathematics 2:482-489 (1981)]을 사용하여 2개의 서열 사이의 상동성의 가장 우수한 분절을 찾는다. 특수한 서열이 예를 들면, 본 발명에 따른 참조 서열과 95% 동일한지를 측정하기 위한 BESTFIT 또는 어떠한 다른 서열 정렬 프로그램을 사용하는 경우, 매개변수가 물론 설정됨으로써 동질성 퍼센트를 참조 폴리펩타이드 서열의 전체 길이에 걸쳐 계산하고 참조 서열내 아미노산의 총 수의 5% 이하의 상동성에 있어서의 갭(gap)을 허용한다.

[0062] 본원에 사용된 바와 같이, VWF 서열 또는 FVIII 단백질 서열에서 "에 상응하는 아미노산" 또는 "등가의 아미노산"은 제1의 VWF 또는 FVIII 서열과 제2의 VWF 또는 FVIII 서열 사이의 동질성 또는 유사성을 최대화하기 위한 정렬에 의해 확인된다. 제2의 VWF 또는 FVIII 서열에서 등가의 아미노산을 확인하는데 사용된 수는 제1의 VWF 또는 FVIII 서열에서 상응하는 아미노산을 확인하는데 사용된 수를 기본으로 한다.

[0063] "융합" 또는 "키메라" 분자는 자연에서 천연적으로 연결되지 않는 제2의 아미노산에 연결된 제1의 아미노산 서열을 포함한다. 별도의 단백질에 일반적으로 존재하는 아미노산 서열은 융합 폴리펩타이드 속에서 함께 존재할 수있거나, 동일한 단백질 속에 일반적으로 존재하는 아미노산 서열은 융합 폴리펩타이드, 예를 들면, 면역글로불린 Fc 도메인이 있는 본 발명의 인자 VIII 도메인의 융합체 속에 새로운 배열로 위치할 수 있다. 융합 단백질은 예를 들면, 화학 합성에 의해, 또는 웨타이드 영역이 요구된 관계로 암호화된 폴리뉴클레오타이드를 생성하여 해독시킴으로써 생성된다. 키메라 단백질은 공유결합성, 비-펩타이드 결합 또는 비-공유 결합과 연합된 제2의 아미노산 서열을 추가로 포함할 수 있다.

[0064] 본원에 사용된 것으로서, 용어 "반감기"는 생체내에서 특수한 폴리펩타이드의 생물학적 반감기를 말한다. 반감기는, 피검자에게 투여된 양의 1/2이 동물내 순환 및/또는 다른 조직으로부터 정화되는데 요구되는 시간으로 나타낼 수 있다. 제공된 폴리펩타이드의 정화 곡선이 시간의 함수로서 구성되는 경우, 당해 곡선은 일반적으로 신속한 ?-상 및 보다 긴 β-상을 지닌 이상(biphasic)이다. ?-상은 내부-와 외부-혈관 공간 사이에서 투여된 폴리펩타이드의 균형을 전형적으로 나타내며 부분적으로는, 폴리펩타이드의 크기에 의해 측정된다. β-상은 전형적으로 혈관내 공간 속에서의 폴리펩타이드의 이화작용을 나타낸다. 일부 구현예에서, 본 발명의 키메라 분자는 일상성(monophasic)이므로, 알파 상을 가지지 않지만 단지 단일의 베타 상이다. 따라서, 특정의 구현예에서, 본원에 사용된 것으로서 용어 반감기는 β-상의 폴리펩타이드의 반감기를 말한다. 사람에서 사람 항체의 대표적인 상 반감기는 21일이다.

[0065] 폴리뉴클레오타이드 또는 폴리펩타이드에 적용된 것으로서 용어 "이종"은, 폴리뉴클레오타이드 또는 폴리펩타이드가, 이것이 비교되는 실체의 것으로부터 구별되는 실체로부터 기원함을 의미한다. 따라서, VWF 단백질에 연결된 이종 폴리펩타이드는 VWF 단백질에 연결되고 VWF 단백질의 천연적으로 존재하는 부분이 아닌 폴리펩타이드 쇄를 의미한다. 예를 들면, 이종의 폴리뉴클레오타이드 또는 항원은 상이한 종, 개인의 상이한 세포형, 또는 구별되는 개인의 동일하거나 상이한 세포형으로부터 기원할 수 있다.

[0066] 본원에 사용된 것으로서 용어 "연결된(linked)", "융합된" 또는 "연결된(connected)"은 제2의 아미노산 서열 또

는 뉴클레오타이드 서열에 결합된 (예를 들면, 웨타이드 결합 또는 포스포디에스테르 결합 각각을 통해) 제1의 아미노산 서열 또는 뉴클레오타이드 서열을 말한다. 용어 "공유 연결된" 또는 "공유결합성 연결"은 서로 연결된 2개의 잔기 사이의 공유 결합, 예를 들면, 이황화물 결합, 웨타이드 결합, 또는 하나 이상의 아미노산, 예를 들면, 링커를 말한다. 제1의 아미노산 또는 뉴클레오타이드 서열은 제2의 아미노산 또는 뉴클레오타이드 서열에 직접결합될 수 있거나 달리는 개재 서열(intervening sequence)이 제1 서열을 제2 서열에 결합시킬 수 있다. 용어 "연결된", "융합된", 또는 "연결된"은 C-말단 또는 N-말단에서 제2의 아미노산 서열에 대한 제1의 아미노산 서열의 융합만을 의미하는 것이 아니라, 또한 제1의 전체 아미노산 서열(또는 제2의 아미노산 서열)의 제2의 아미노산 서열(또는 제1의 아미노산 서열 각각)내 어떠한 2개의 아미노산의 삽입을 포함한다. 하나의 구현예에서, 제1의 아미노산 서열은 웨타이드 결합 또는 링커에 의해 제2의 아미노산 서열에 결합될 수 있다. 제1의 뉴클레오타이드 서열은 포스포디에스테르 결합 또는 링커에 의해 제2의 뉴클레오타이드 서열에 결합될 수 있다. 링커는 웨타이드 또는 폴리웨타이드(폴리웨타이드 쇄의 경우) 또는 뉴클레오타이드 또는 뉴클레오타이드 쇄(뉴클레오타이드 쇄의 경우) 또는 어떠한 화학적 잔기(폴리웨타이드 및 폴리뉴클레오타이드 쇄 둘 다의 경우)일 수 있다. 공유결합성 연결은 때때로 (-) 또는 하이픈으로 나타낸다.

[0067]

본원에 사용된 것으로서 용어 "와 연합된"은 제1의 아미노산 쇄와 제2의 아미노산 쇄 사이에 형성된 공유결합성 또는 비-공유결합성 결합을 말한다. 하나의 구현예에서, 용어 "와 연합된"은 공유결합성, 비-웨타이드 결합 또는 비-공유결합성 결합을 의미한다. 일부 구현예에서, 당해 연합은 쌍점, 예를 들면, (:)로 나타낸다. 다른 구현예에서, 이는웨타이드 결합을 제외하고 공유결합성 결합을 의미한다. 다른 구현예에서, 본원에 사용된 것으로서 용어 "공유결합적으로 연합된"은, 공유 결합, 예를들면, 이황화물 결합, 웨타이드 결합, 또는 하나 이상의 아미노산(예를 들면, 링커)에 의한 2개의 잔기 사이의 연합을 의미한다. 예를 들면, 아미노산 시스테인은 제2의 시스테인 잔기 상의 이황화물 결합 또는 티올 그룹과의 브릿지를 형성할 수 있는 티올 그룹을 포함한다. 대부분의 천연적으로 존재하는 IgG 분자에서, CH1 및 CL 영역은 이황화물 결합에 의해 연합되어있으며 2개의 중쇄는 카밧 번호매김 시스템(Kbat numbering system)을 사용하여 239 및 242번에 상응하는 위치(226 또는 229, EU 번호매김 시스템)에서 2개의 이황화물 결합에 의해 연합된다. 공유 결합의 예는 웨타이드 결합, 금속 결합, 수소 결합, 이황화물 결합, 시그마 결합, pi 결합, 델타 결합, 글리코시드 결합, 아그노스틱 결합(agnostic bond), 벤트 결합(bent bond), 쌍극 결합(dipolar bond), Pi 골격, 이중 결합, 삼중 결합, 사중 결합, 오중 결합, 육중 결합, 접합, 하이퍼접합(hyperconjugation), 방향성, 촉각성, 또는 항결합(antibonding)을 포함하나, 이에 한정되지 않는다. 비-공유결합성 결합의 비-제한적 예는 이온 결합(예를 들면, 양이온-pi 결합 또는 염 결합), 금속 결합, 수소 결합(예를 들면, 이수소 결합, 이수소 복합체, 저-장벽 수소 결합, 또는 대칭 수소 결합), 반데르 바알스 힘(van der Walls force), 런던 분산력(London dispersion force), 기계적 결합, 할로겐 결합, 친금성(aurophilicity), 인터칼레이션(intercalation), 적층(stacking), 엔트로피력(entropic force), 또는 화학 극성을 포함한다.

[0068]

본원에 사용된 것으로서, 용어 "절단 부위" 또는 "효소적 절단 부위"는 효소에 의해 인식된 부위를 말한다. 하나의 구현예에서, 폴리웨타이드는 응고 캐스케이드(cascade) 동안에 활성화된 효소에 의해 절단된 효소절단 부위를 가짐으로써, 이러한 부위의 절단은 응고 형성 부위에서 일어난다. 다른 구현예에서, FVIII 단백질 및 제2의 이중 잔기를 연결시키는 FVIII 링커는 절단 부위를 포함할 수 있다. 예시적인 이러한 부위는 예를 들면, 트롬빈, 인자 XIa 또는 인자 Xa에 의해 인식된 것들을 포함한다. 예시적인 FXIa 절단 부위는 예를 들면, TQSFDFTFTR(서열 번호 27) 및 SVSQTSKLTR(서열 번호 28)을 포함한다. 예시적인 트롬빈 절단 부위는, 예를 들면, DFLAEGGGVR(서열 번호 29), TTKIKPR(서열 번호 30), LVPRG(서열 번호 31) 및 ALRPR(서열 번호 26의 1 내지 5번 아미노산)을 포함한다. 다른 효소 분해 부위는 당해 분야에 공지되어 있다. 트롬빈에 의해 절단될 수 있는 절단 부위는 "트롬빈 절단 부위"로서 본원에 언급된다.

[0069]

본원에 사용된 것으로서, 용어 "프로세싱 부위" 또는 "세포내 프로세싱 부위"는 폴리웨타이드의 해독 후 기능하는 효소에 대한 표적인 폴리웨타이드내 효소 절단 부위의 유형을 말한다. 하나의 구현예에서, 이러한 효소는 내강(Golgi lumen)으로부터 경 골지 구획까지의 수송 동안 기능한다. 세포내 프로세싱 효소는 세포로부터 단백질의 분비 전에 폴리웨타이드를 절단한다. 이러한 프로세싱 부위의 예는, 예를 들면, PACE/푸린(여기서, PACE는 쌍을 이룬 염기성 아미노산 절단 효소에 대한 머리글자이다) 계열의 엔도웨타이드를 포함한다. 이들 효소는 골지막에 국재화되어 서열 모티프 Arg-[어떠한 잔기]-(Lys 또는 Arg)-Arg의 카복시 말단 측면에서 단백질을 절단한다. 본원에 사용된 것으로서, "푸린" 계열의 효소는 예를 들면, PCSK1(또한 PC1/Pc3으로서 공지됨), PCSK2(또한 PC2로서 공지됨), PCSK3(또한 푸린 또는 PACE로서 공지됨), PCSK4(또한 PC4로서 공지됨), PCSK5(또한 PC5 또는 PC6로서 공지됨), PCSK6(또한 PACE4로서 공지됨), 또는 PCSK7(또한 PC7/LPC, PC8, 또는 SPC7로서 공지됨)

됨)를 포함한다. 다른 프로세시 부위는 당해 분야에 공지되어 있다. 본원에 언급된 용어 "프로세싱가능한 링커"는 세포내 프로세싱 부위를 포함하는 링커를 의미한다.

[0070] 용어 "푸린"은 EC 제3.4.21.75호에 상응하는 효소를 말한다. 푸린은 서브틸리신-유사 단백질 전환효소이며, 이는 또한 PACE(쌍을 이룬 염기성 아미노산 절단 효소)(Paired basic Amino acid Cleaving Enzyme)로 또한 공지되어 있다. 푸린은, 불활성의 전구체 단백질을 생물학적으로 활성인 단백질로 전환시키는 불활성 전구체 단백질의 부분을 결실하고 있다. 이의 세포내 수송 동안, 전구-펩타이드는 골지내에서 푸린 효소에 의해 성숙한 VWF 분자로부터 절단된다.

[0071] 하나 이상의 프로세싱 또는 절단 부위를 포함하는 작제물에서, 이러한 부위는 동일하거나 상이할수 있는 것으로 이해될 것이다.

[0072] 본원에 사용된 것으로서, 지혈성 질환은 피브린 응괴를 형성하는데 있어서의 손상된 능력 또는 불능으로 인하여, 자발적으로 또는 외상의 결과로서, 출혈 경향성에 의해 특징화되는 유전적으로 유전되거나 후천적인 상태를 의미한다. 이러한 질환의 예는 혈우병을 포함한다. 3개의 주요 형태는 혈우병 A(인자 VIII 결핍증), 혈우병 B(인자 IX 결핍증 또는 "크리스마스병") 및 혈우병 C(인자 XI 결핍증, 약한 출혈 경향성)이다. 다른 지혈성 질환은 예를 들면, 폰 빌레브란드병(von Willebrand disease), 인자 XI 결핍증(PTA 결핍증), 인자 XII 결핍증, 피브리노겐, 프로트롬빈, 인자 V, 인자 VII, 인자 X 또는 인자 XIII에 있어서 결핍증 또는 구조적 비정상, GPIb에 있어서 결함 또는 결핍증을 포함한다. VWF에 대한 수용체인 GPIb는 결합할 수 있으며 주요 응괴 형성(1차 지혈) 및 증가된 출혈 경향성을 가진다. 글란즈만(Glanzman) 및 내겔리(Naegeli)의 혈소판무력증[글란즈만 혈소판무력증(Glanzmann thrombasthenia)]을 초래할 수 있다. 간 부전(급성 및 만성형)에서, 간에 의한 응고 인자의 불충분한 생산이 존재하며; 이는 출혈 위험을 증가시킬 수 있다.

[0073] 본 발명의 키메라 분자는 예방학적으로 사용될 수 있다. 본원에 사용된 것으로서, 용어 "예방학적 치료"는 출혈 사건 이전의 분자의 투여를 말한다. 하나의 구현예에서, 일반적인 지혈제를 필요로 하는 피검자는 수술 중이거나, 수술 예정에 있다. 본 발명의 키메라 단백질은 예방제로서 수술 전에 또는 후에 투여될 수 있다. 본 발명의 키메라 단백질은 수술 동안 또는 후에 투여되어 급성 출혈 사건을 제어할 수 있다. 수술은 간 이식, 간 절제, 치과 수술, 또는 줄기 세포 이식을 포함하나, 이에 한정되지 않는다.

[0074] 본 발명의 키메라 분자는 또한 요구형(또는 "사건적") 치료를 위해 사용된다. 용어 "요구시 치료" 또는 "사건적 치료"는 출혈을 유발할 수 있는 활성 전에 또는 출혈 사건의 증상에 대한 반응시 키메라 분자의 투여를 말한다. 하나의 국면에서, 요구시(사건적) 치료는 손상 후와 같이 출혈이 시작되는 경우, 또는 수술 전과 같이 출혈이 예상되는 경우에 피검자에게 제공될 수 있다. 다른 측면에서, 요구시 치료는 접촉성 스포츠와 같이, 출혈 위험이 증가하는 활동 전에 제공될 수 있다.

[0075] 본원에 사용된 것으로서, 용어 "급성 출혈"은 직면한 위험에 무관한 출혈 사건을 말한다. 예를 들면, 피검자는 외상, 요독증, 유전성 출혈 질환(예를 들면, 인자 VII 결핍증) 혈소판 질환, 또는 응괴 인자에 대한 항체의 발달에 기인한 내성을 가질 수 있다.

[0076] 본원에 사용된 것으로서, 치료하다, 치료, 치료하는은 예를 들면, 질병 또는 상태의 중증도에 있어서의 감소; 질병 과정의 기간에 있어서의 감소; 질병 또는 상태와 관련된 하나 이상의 증상의 완화; 질병 또는 상태를 필수적으로 치유하지 않고, 질병 또는 상태를 지닌 피검자에 대한 유리한 효과의 제공, 또는 질병 또는 상태와 관련된 하나 이상의 증상의 예방을 말한다. 하나의 구현예에서, 용어 "치료하는" 또는 "치료"는 본 발명의 키메라 분자의 투여에 의해 피검자에서 FVIII 최저 농도(trough level)를 적어도 약 1 IU/dL, 2 IU/dL, 3 IU/dL, 4 IU/dL, 5 IU/dL, 6 IU/dL, 7 IU/dL, 8 IU/dL, 9 IU/dL, 10 IU/dL, 11 IU/dL, 12 IU/dL, 13 IU/dL, 14 IU/dL, 15 IU/dL, 16 IU/dL, 17 IU/dL, 18 IU/dL, 19 IU/dL, 또는 20 IU/dL로 유지함을 의미한다. 다른 구현예에서, '치료하는' 또는 '치료'는 FVIII 최저 수준을 약 1 내지 약 20 IU/dL, 약 2 내지 약 20 IU/dL, 약 3 내지 약 20 IU/dL, 약 4 내지 약 20 IU/dL, 약 5 내지 약 20 IU/dL, 약 6 내지 약 20 IU/dL, 약 7 내지 약 20 IU/dL, 약 8 내지 약 20 IU/dL, 약 9 내지 약 20 IU/dL, 또는 약 10 내지 약 20 IU/dL로 유지함을 의미한다. 질병 또는 상태의 '치료' 또는 '치료하는'은 또한 피검자에서 FVIII 활성을 비-혈우병 피검자에서 FVIII 활성의 적어도 약 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 또는 20%에 필적한 수준에서 유지함을 포함할 수 있다. 치료에 요구되는 최소의 최저 수준은 하나 이상의 공지된 방법으로 측정할 수 있으며 각각의 사람에 대해 조절(증가 또는 감소)할 수 있다.

## II. 키메라 분자

[0078]

본 발명의 키메라 분자를 설계하여, VWF 단백질 또는 FVIII 단백질이 융합된 다른 잔기로부터 VWF 단백질 또는 FVIII 단백질의 방출을 증진시킨다. 본 발명은 손상 부위에서 신속하고 효율적으로 제거할 수 있는 트롬빈 절단 가능한 링커를 제공한다. 본 발명의 하나의 측면에서, 키메라 분자는 폰 빌레브란트 인자(VWF) 단백질, 이종 잔기(H1), XTEN 서열, 및 VWF 단백질과 이종 잔기를 연결하는 VWF 링커를 포함할 수 있으며, 여기서 VWF 링커는: (i) 인자 VIII(FVIII)로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합으로부터 선택된 폴리펩타이드를 포함하며, 여기서 XTEN 서열은 VWF 단백질, 이종 잔기(H1), VWF 링커, 또는 이의 어떠한 조합에도 연결된다. 본 발명의 다른 측면에서, 키메라 분자는 VWF 단백질, 이종 잔기(H1), 및 VWF 단백질과 이종 잔기(H1)를 연결하는 VWF 링커를 포함하는 제1의 폴리펩타이드 쇄 및 FVIII 단백질 및 XTEN 서열을 포함하는 제2의 폴리펩타이드 서열을 포함할 수 있으며, 여기서 제1 폴리펩타이드 서열내 VWF 링커는: (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함하며, 여기서 제1 폴리펩타이드 쇄 및 제2 폴리펩타이드 쇄는 서로 연합되어 있다.

[0079]

본 발명의 다른 측면에서, 키메라 분자는 FVIII 링커를 통해 이종 잔기에 융합된 FVIII 단백질을 포함하는 폴리펩타이드 쇄를 포함하며, 여기서 FVIII 링커는: (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함한다.

[0080]

## II.A. VWF, XTEN, VWF 링커를 지닌 키메라 분자

[0081]

본 발명은 VWF 링커를 통해 XTEN 서열에 융합된 VWF 단백질을 포함하는 키메라 분자를 제공하며, 여기서 VWF 링커는 (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함한다.

[0082]

하나의 구현예에서, 키메라 분자는 VWF 단백질, 이종 잔기(H1), XTEN 서열, 및 VWF 단백질을 이종 잔기와 연결하는 VWF 링커를 포함하며, 여기서 XTEN 서열은 VWF 단백질과 VWF 링커 사이에 위치하고, 여기서 VWF 링커는 (i) 인자 VIII(FVIII)로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함한다. 다른 구현예에서, 키메라 분자는 VWF 단백질, 이종 잔기(H1), XTEN 서열, 및 VWF 단백질과 이종 잔기를 연결시키는 VWF 링커를 포함하고, 여기서 XTEN 서열은 VWF 링커와 이종 잔기 사이에 위치하며, 여기서 VWF 링커는: (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합으로부터 선택된 폴리펩타이드를 포함한다.

[0083]

다른 구현예에서, 키메라 분자는 폴리펩타이드 쇄를 추가로 포함하고, FVIII 단백질을 포함하며, 여기서 제1 쇄는 VWF 단백질을 포함하고 제2 쇄는 서로 연합된 FVIII 단백질을 포함한다. 하나의 예에서, 연합은 공유 결합, 예를 들면, 이황화물 결합, 연합일 수 있다. 여전히 다른 구현예에서, FVIII 단백질을 포함하는 폴리펩타이드 쇄는 추가의 XTEN 서열을 추가로 포함한다. 추가의 XTEN 서열은 FVIII 단백질의 N-말단 또는 C-말단에 연결될 수 있거나 서로 인접한 2개의 FVIII 아미노산 사이에 삽입된다. 여전히 다른 구현예에서, FVIII 단백질을 포함하는 쇄는 제2의 이종 잔기(H2)를 추가로 포함한다. 일부 구현예에서, FVIII 단백질은 FVIII 링커를 통해 제2의 이종 잔기에 융합된다. 특정의 구현예에서, FVIII 링커는 VWF 단백질과 이종 잔기를 연결하는 VWF 링커와 동일하다. 다른 구현예에서, FVIII 링커는 VWF 단백질과 이종 잔기를 연결하는 VWF 링커와는 상이하다.

[0084]

특정의 구현예에서, 키메라 분자는: (i) V-L1-X1-H1:H2-L2-X2-C; (ii) V-X1-L1-H1:H2-L2-X2-C; (iii) V-L1-X1-H1:H2-X2-L2-C; (iv) V-X1-L1-H1:H2-X2-L2-C; (v) V-L1-X1-H1:H2-L2-C(X2); (vi) V-X1-L1-H1:H2-L2-C(X2); (vii) C-X2-L2-H2:H1-X1-L1-V; (viii) C-X2-L2-H2:H1-L1-X1-V; (ix) C-L2-X2-H2:H1-L1-X1-V; (x) C-L2-X2-H2:H1-L1-X1-V; (xi) C(X2)-L2-H2:H1-X1-L1-V; 또는 (xii) C(X2)-L2-H2:H1-L1-X1-V로부터 선택된 화학식을 포함하고; 여기서 V는 VWF 단백질이고; L1은 VWF 링커이며; L2는 임의의 FVIII 링커이고; H1은 제1의 이종 잔기이며; H2는 제2의 이종 잔기이고; X1은 XTEN 서열이며; X2는 임의의 XTEN 서열이고; C는 FVIII 단백질이며; C(X2)는 XTEN 서열에 융합된 FVIII 단백질이고, 여기서 XTEN 서열은 서로 인접한 2개의 FVIII 아미노산 사이에

삽입되며; (-)는 웨타이드 결합 또는 하나 이상의 아미노산이고; (:)는 H1과 H2 사이의 공유 결합이다.

[0085] 일부 구현예에서, 키메라 단백질내 FVIII 단백질은 제3의 이종 잔기(H3)를 포함하고, 이는 XTEN 서열일 수 있다. 다른 구현예에서, 키메라 분자의 FVIII 단백질은 제4의 이종 잔기(H4)를 포함하며, 이는 XTEN 서열일 수 있다. 여전히 다른 구현예에서, 키메라 분자의 FVIII 단백질은 제5의 이종 잔기(H5)를 포함하고, 이는 XTEN 서열일 수 있다. 여전히 다른 구현예에서, 키메라 분자의 FVIII 단백질은 제6의 이종 잔기(H6)를 포함하고, 이는 XTEN 서열일 수 있다. 특정의 구현예에서, 하나 이상의 제3의 이종 잔기(H3), 제4의 이종 잔기(H4), 제5의 이종 잔기(H5), 및 제6의 이종 잔기(H6)는 키메라 분자의 반감기를 연장시킬 수 있다. 일부 구현예에서, 제3의 이종 잔기(H3), 제4의 이종 잔기(H4), 제5의 이종 잔기(H5), 및 제6의 이종 잔기(H6)는 FVIII 단백질 C-말단 또는 N-말단에 연결되어 있거나 FVIII 단백질의 2개의 아미노산 사이에 삽입되어 있다.

#### II.B. FVIII, XTEN, VWF 단백질, VWF 링커를 지닌 키메라 분자

[0087] 본 발명은 또한 VWF 단백질, 이종 잔기(H1), 및 VWF 단백질과 이종 잔기(H1)를 연결시키는 VWF 링커를 포함하는 제1의 폴리웨타이드 및 FVIII 단백질과 XTEN 서열을 포함하는 제2의 폴리웨타이드를 제공하며, 여기서, 제1의 폴리웨타이드내 VWF 링커는: (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합으로부터 선택된 폴리웨타이드를 포함하고, 여기서 제1의 폴리웨타이드 쇄 및 제2의 폴리웨타이드 쇄는 서로 연합되어 있다. 하나의 구현예에서, 여기서 XTEN 서열은 FVIII 단백질의 N-말단 또는 C-말단에 연결되거나 서로 인접한 2개의 FVIII 아미노산 사이에 삽입된다. 다른 구현예에서, 키메라 분자는 추가의 XTEN 서열을 추가로 포함하고, 이는 VWF 단백질, 이종잔기, VWF 링커, 또는 이의 어떠한 조합에 연결되어 있다. 다른 구현예에서, 키메라 분자는 제2의 이종 잔기(H2)를 추가로 포함한다. 추가의 구현예에서, 키메라 분자의 제2의 이종 잔기는 FVIII 단백질, XTEN 서열, 또는 둘 다에 연결되어 있다. 여전히 다른 구현예에서, 제2의 이종 잔기는 FVIII 단백질 또는 FVIII 링커를 통해 XTEN 서열에 연결되어 있다. 여전히 다른 구현예에서, FVIII 링커는 VWF 링커와 동일하다. 일부 구현예에서, FVIII 링커는 VWF 링커와는 상이하다.

[0088] 특정의 구현예에서, 키메라 분자는: (i) V-L1-X1-H1:H2-L2-X2-C; (ii) V-X1-L1-H1:H2-L2-X2-C; (iii) V-L1-X1-H1:H2-X2-L2-C; (iv) V-X1-L1-H1:H2-X2-L2-C; (v) V-L1-X1-H1:H2-L2-C(X2); (vi) V-X1-L1-H1:H2-L2-C(X2); (vii) C-X2-L2-H2:H1-X1-L1-V; (viii) C-X2-L2-H2:H1-L1-X1-V; (ix) C-L2-X2-H2:H1-L1-X1-V; (x) C-L2-X2-H2:H1-L1-X1-V; (xi) C(X2)-L2-H2:H1-X1-L1-V; 또는 (xii) C(X2)-L2-H2:H1-L1-X1-V로부터 선택된 화학식을 포함하고; 여기서 V는 VWF 단백질이고; L1은 VWF 링커이며; L2는 임의의 FVIII 링커이고; H1은 제1의 이종 잔기이며; H2는 제2의 이종 잔기이고; X1은 XTEN 서열이며; X2는 임의의 XTEN 서열이고; C는 FVIII 단백질이며; C(X2)는 XTEN 서열에 융합된 FVIII 단백질이고, 여기서 XTEN 서열은 서로 인접한 2개의 FVIII 아미노산 사이에 삽입되며; (-)는 웨타이드 결합 또는 하나 이상의 아미노산이고; (:)는 H1과 H2 사이의 공유 결합이다. 하나의 구현예에서, VWF 링커 및 FVIII 링커는 동일할 수 있다. 다른 구현예에서, VWF 링커 및 FVIII 링커는 상이하다.

[0089] 특정의 구현예에서, 키메라 분자의 FVIII 단백질은 제3의 이종 잔기 (H3)를 포함하고, 이는 XTEN 서열일 수 있다. 다른 구현예에서, 키메라 분자의 FVIII 단백질은 제4의 이종 잔기(H4)를 포함하며, 이는 XTEN 서열이다. 여전히 다른 구현예에서, 키메라 분자의 FVIII 단백질은 제5의 이종 잔기(H5)를 포함하고, 이는 XTEN 서열일 수 있다. 여전히 다른 구현예에서, FVIII 단백질은 제6의 이종 잔기(H6)를 포함하고, 이는 XTEN 서열일 수 있다. 특정의 구현예에서, 하나 이상의 제3의 이종 잔기(H3), 제4의 이종 잔기(H4), 제5의 이종 잔기(H5), 및 제6의 이종 잔기(H6)는 키메라 분자의 반감기를 연장시킬 수 있다. 일부 구현예에서, 제3의 이종 잔기(H3), 제4의 이종 잔기(H4), 제5의 이종 잔기(H5), 및/또는 제6의 이종 잔기(H6)는 FVIII 단백질의 C 말단 또는 N 말단에 연결되거나 FVIII 단백질의 2개의 아미노산 사이에 삽입된다.

#### II.C. FVIII, XTEN, 및 FVIII 링커를 지닌 키메라 분자

[0091] 본 발명의 키메라 분자는 FVIII 단백질, XTEN 서열, 및 FVIII 링커에 의해 융합된 이종 잔기를 포함할 수 있으며, 이는 (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함한다. 특정의 구현예에서, 키메라 분자는 2개의 폴리웨타이드 쇄를 포함하고, 제1의 쇄는 FVIII 링커를 통해 제1의 Fc 영역에 융합된 FVIII 링커를 포함하며, 제2의 쇄는 Fc 영역에 융합된 VWF 단백질 (예를 들면, VWF의 D' 도메인 및 D3 도메인)을 포함하고, 여기서 제1의 폴리웨타이드 쇄내의 FVIII 링커는: (i) FVIII로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3

영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위; 또는 (v) 이의 어떠한 조합을 포함하고, 여기서 제1의 폴리펩타이드 쇄 및 제2의 폴리펩타이드 쇄는 서로 연합되어 있으며 여기서 XTEN 서열은 제1의 폴리펩타이드(예를 들면, FVIII 단백질의 N 말단 또는 C 말단, 링커, 제1의 Fc 영역 또는 FVIII 단백질내), 제2의 폴리펩타이드(예를 들면 VWF 단백질의 N 말단 또는 C 말단, 또는 Fc 영역 또는 FVIII 단백질내), 또는 둘 다에 연결되어 있다. 구체적인 구현예에서, 제1의 폴리펩타이드 쇄내 링커는 FVIII로부터의 a2 영역을 포함한다.

[0092] 특정의 구현예에서, 키메라 분자는 (i) V-L2-X2-H2: H1-L1-X1-C; (ii) V-X2-L2-H2: H1-L1-X1-C; (iii) V-L2-X2-H2: H1-X1-L1-C; (iv) V-X2-L2-H2: H1-X1-L1-C; (v) V-L2-X2-H2: H1-L1-C(X1); (vi) V-X2-L2-H2: H1-L1-C(X1); (vii) C-X1-L1-H1: H2-X2-L2-V; (viii) C-X1-L1-H1: H2-L2-X2-V; (ix) C-L1-X1-H1:H2-L2-X2-V; (x) C-L1-X1-H1:H2-L2-X2-V; (xi) C(X1)-L1-H1:H2-X2-L2-V; 또는 (xii) C(X1)-L1-H1:H2-L2-X2-V로부터 선택된 화학식을 포함하고, 여기서 V는 VWF 단백질이고; L1은 FVIII 링커이며; L2는 임의의 VWF 링커이고; H1은 제1의 이종 잔기이며; H2는 제2의 이종 잔기이고; X1은 임의의 XTEN 서열이며; X2는 임의의 XTEN 서열이고; C는 FVIII 단백질이며; C(X1)은 XTEN 서열에 융합된 FVIII 단백질이고, 여기서 XTEN 서열은 서로 인접한 2개의 FVIII 아미노산 사이에 삽입되며; (-)는 펩타이드 결합 또는 하나 이상의 아미노산이고; (:)는 H1과 H2 사이의 공유 결합이며 여기서 적어도 하나의 XTEN 서열은 키메라 분자내에 존재한다. 하나의 구현예에서, VWF 링커 및 FVIII 링커는 동일하다. 다른 구현예에서, VWF 링커 및 FVIII 링커는 상이하다.

#### [0093] II.D. 키메라 분자의 성분

##### [0094] II.C.1. VWF 링커 또는 FVIII 링커

[0095] 본 발명의 키메라 분자에 유용한 VWF 링커 또는 FVIII 링커는 VWF 단백질과 이종 잔기 또는 FVIII 단백질과 이종 잔기를 융합시키는 트롬빈 절단가능한 링커이다. 하나의 구현예에서, VWF 링커 또는 FVIII 링커는 FVIII의 a1 링커를 포함한다. 다른 구현예에서, VWF 링커 또는 FVIII 링커는 FVIII의 a2 링커를 포함한다. 다른 구현예에서, VWF 링커 또는 FVIII 링커는 FVIII의 a3 링커를 포함한다. 여전히 다른 구현예에서, VWF 링커 또는 FVIII 링커는 X-V-P-R(여기서, X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위를 포함한다.

[0096] 하나의 구현예에서, VWF 링커 또는 FVIII 링커는 전장의 성숙한 FVIII에 상응하는 Met337 내지 Arg372에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99%, 또는 100% 동일한 아미노산 서열을 포함하는 a1 영역을 포함하고, 여기서 a1 영역은 트롬빈에 의해 절단될 수 있다. 다른 측면에서, VWF 링커 또는 FVIII 링커는 전장의 성숙한 FVIII에 상응하는 337 내지 372번 아미노산에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99%, 또는 100% 동일한 아미노산 서열을 포함하는 a1 영역을 포함하고, 여기서 a1 영역은 트롬빈에 의해 절단될 수 있다. 다른 구현예에서, VWF 링커 또는 FVIII 링커는 추가의, 예를 들면, 1개, 2개, 3개, 4개, 5개, 10개 이상의 아미노산을 추가로 포함한다. 특수한 구현예에서, VWF 링커 또는 FVIII 링커는 ISMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSV (서열 번호 5)를 포함한다.

[0097] 일부 구현예에서, VWF 링커 또는 FVIII 링커는 전장의 성숙한 FVIII에 상응하는 Glu720 내지 Arg740에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99%, 또는 100% 동일한 아미노산 서열을 포함하는 a2 영역을 포함하고, 여기서 a2 영역은 트롬빈에 의해 절단될 수 있다. 다른 구현예에서, VWF 링커 또는 FVIII 링커는 전장의 성숙한 FVIII에 상응하는 712 내지 743번 아미노산에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99%, 또는 100% 동일한 아미노산 서열을 포함하는 a2 영역을 포함하고, 여기서 a2 영역은 트롬빈에 의해 절단될 수 있다. 여전히 다른 구현예에서, VWF 링커 또는 FVIII 링커는 추가의, 예를 들면, 1개, 2개, 3개, 4개, 5개, 10개 이상의 아미노산을 추가로 포함한다. 특수한 구현예에서, VWF 링커는 ISDKNTGYYEDSYEDISAYLLSKNNAEPRSFS (서열 번호 4)를 포함한다.

[0098] 특정한 구현예에서, VWF 링커 또는 FVIII 링커는 전장의 성숙한 FVIII에 상응하는 Glu1649 내지 Arg1689에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99%, 또는 100% 동일한 아미노산 서열을 포함하는 a3 영역을 포함하고, 여기서 a3 영역은 트롬빈에 의해 절단될 수 있다. 일부 구현예에서, VWF 링커 또는 FVIII 링커는 전장의 성숙한 FVIII에 상응하는 1649 내지 1692번 아미노산에 대해 적어도 약 80%, 약 85%, 약 90%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99%, 또는 100% 동일한 아미노산 서열을 포함하는 a3 영역을 포함하고, 여기서 a3 영역은 트롬빈에 의해 절단될 수 있다. 다른 구현예에서, VWF 링커 또는 FVIII 링커는 추가의, 예를 들면, 1개, 2개, 3개, 4개, 5개, 10개 이상의 아미노산을 추가로 포함한다. 구체적인 구현예에서, VWF 링커 또는 FVIII 링커는 ISEITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQ (서열 번호 6)을 포함한다.

- [0099] 다른 구현예에서, VWF 링커 또는 FVIII 링커는 X-V-P-R(서열 번호 3) 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 트롬빈 절단 부위를 포함하고 여기서 PAR1 엑소사이트 상호작용 모티프는 S-F-L-L-R-N (서열 번호 7)를 포함한다. 일부 구현예에서, PAR1 엑소사이트 상호작용 모티프는 P, P-N, P-N-D, P-N-D-K(서열 번호 8), P-N-D-K-Y(서열 번호 9), P-N-D-K-Y-E(서열 번호 10), P-N-D-K-Y-E-P(서열 번호 11), P-N-D-K-Y-E-P-F(서열 번호 12), P-N-D-K-Y-E-P-F-W(서열 번호 13), P-N-D-K-Y-E-P-F-W-E(서열 번호 14), P-N-D-K-Y-E-P-F-W-E-D(서열 번호 20), P-N-D-K-Y-E-P-F-W-E-D-E(서열 번호 21), P-N-D-K-Y-E-P-F-W-E-D-E-E(서열 번호 22), P-N-D-K-Y-E-P-F-W-E-D-E-S(서열 번호 23), 또는 이의 어떠한 조합도 포함한다. 다른 구현예에서, X-V-P-R을 포함하는 트롬빈 절단 부위에 대한 지방족 아미노산은 글리신, 알라닌, 발린, 루이신, 또는 이소루이신으로부터 선택된다. 구체적인 구현예에서, 트롬빈 절단 부위는 L-V-P-R을 포함한다. 일부 구현예에서, 트롬빈은, 트롬빈 절단 부위(L-V-P-R)가 VWF 링커 또는 FVIII 링커 각각에 대해 치환되는 경우(즉, PAR1 엑소사이트 상호작용 모티프의 부재하여) 트롬빈이 트롬빈 절단 부위(예를 들면, L-V-P-R)를 절단하는 것보다 신속하게 VWF 링커 또는 FVIII 링커를 절단한다. 일부 구현예에서, 트롬빈은, 트롬빈 절단 부위(예를 들면, L-V-P-R)가 VWF 링커 또는 FVIII 링커에 대해 치환되는 경우 트롬빈이 트롬빈 절단 부위(예를 들면, L-V-P-R)를 절단할 수 있는 것 보다 적어도 약 10배, 적어도 약 20배, 적어도 약 30배, 적어도 약 40배, 적어도 약 50배, 적어도 약 60배, 적어도 약 70배, 적어도 약 80배, 적어도 약 90배 또는 적어도 약 100배 더 신속하게 VWF 링커 또는 FVIII 링커를 절단한다.
- [0100] 일부 구현예에서, (i) a1 영역, (ii) a2 영역, (iii) a3 영역 또는 (iv) 트롬빈 절단 부위 X-V-P-R 및 PAR1 엑소사이트 상호작용 모티프를 포함하는 VWF 링커 또는 FVIII 링커는, 길이가 적어도 약 2, 5, 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, 또는 2000개 아미노산인 하나 이상의 아미노산을 추가로 포함한다. 하나의 구현예에서, 하나 이상의 아미노산은 gly 웨타이드를 포함한다. 다른 구현예에서, 하나 이상의 아미노산은 GlyGly를 포함한다. 다른 구현예에서, 하나 이상의 아미노산은 IleSer을 포함한다. 여전히 다른 구현예에서, 하나 이상의 아미노산은 gly/ser 웨타이드를 포함한다. 여전히 다른 구현예에서, 하나 이상의 아미노산은 화학식 (Gly<sub>4</sub>Ser)<sub>n</sub> 또는 S(Gly<sub>4</sub>Ser)<sub>n</sub>(여기서 n은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 또는 100으로부터 선택된 양의 정수이다)의 gly/ser 웨타이드를 포함한다. 일부 구현예에서, 하나 이상의 아미노산은 (Gly<sub>4</sub>Ser)<sub>3</sub> (서열 번호 89) 또는 (Gly<sub>4</sub>Ser)<sub>4</sub> (서열 번호 90)을 포함한다.
- [0101] II.C.2. VWF 단백질
- [0102] VWF(F8VWF로서 또한 공지됨)는 혈액 혈장에 존재하고 내피[바이엘-펠라드 본체(Weibel-Palade bodies)에서], 다형핵 거대세포(혈소판의 ?-파립), 및 내피하층 연결 조직에 존재하는 거대한 다량체성 당단백질이다. 염기성 VWF 단량체는 2813개 아미노산 단백질이다. 모든 단량체는 특수 기능을 지닌 특이 도메인, D'/D3 도메인(이는 인자 VIII에 결합한다), A1 도메인(이는 혈소판 GPIb-수용체, 혜파린, 및/또는 가능하게는 콜라겐에 결합한다), A3 도메인(이는 콜라겐에 결합한다), C1 도메인(여기서 RGD 도메인은 활성화되는 경우 혈소판 인테그린 ?IIb?3에 결합한다), 및 단백질의 C-말단에서 "시스테인 노트("cysteine knot)" 도메인[당해 VWF는 혈소판-기원한 성장 인자(PDGF), 형질전환 성장 인자-β(TGFβ) 및 β-사람 만성 고나도트로핀(βHCG)과 공유한다] 다수를 함유한다.
- [0103] 본원에 사용된 것으로서 용어 "VWF 단백질"은 내인성 VWF의 FVIII에 대한 결합을 억제할 수 있는, D' 도메인 및 D3 도메인을 포함하는 전장의 VWF 단백질 또는 기능성 VWF 단편을 포함하나, 이에 한정되지 않는다. 하나의 구현예에서, VWF 단백질은 FVIII에 결합한다. 다른 구현예에서, VWF 단백질은 FVIII 상에서 VWF 결합 부위를 차단함으로써 FVIII와 내인성 VWF의 상호작용을 억제한다. 다른 구현예에서, VWF 단백질은 VWF 정화 경로(clearance pathway)에 의해 정화되지 않는다. VWF 단백질은 VWF의 이들 활성을 보유하는 유도체, 변이체, 돌연변이체, 또는 유사체를 포함한다.
- [0104] 사람 VWF에 대한 2813 단량체 아미노산 서열은 진뱅크(Genbank)에 수탁번호 제\_NP\_000543.2\_호로 보고되어 있다. 사람 VWF를 암호화하는 뉴클레오타이드 서열은 진뱅크에 수탁 번호 제\_NM\_000552.3\_호로 보고되어 있다. 사람 VWF의 뉴클레오타이드 서열은 서열 번호 1로 정의된다. 서열 번호 2는 서열 번호 1에 의해 암호화된 아미노산 서열이다. VWF의 각각의 도메인은 표 1에 나열되어 있다.

[0105]

## 표 1. VWF 서열

VWF 도메인	아미노산 서열	
VWF 시그날 웹타이드 (서열 번호 2 의 1 내지 22 번 아미노산)	1 <u>MIPARFAGVL LALALILPGT LC</u>	22
VWF D1D2 영역 (서열 번호 2 의 23 내지 763 번 아미노산)	23 <u>AEGTRGRS STARCSILFGS</u> 51 <u>DFVNNTFDGSM</u> <u>YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE</u> <u>FFDIHLFVNG</u> 101 <u>TVIQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI</u> <u>DGSGNFQVLL</u> 151 <u>SDRYFNKTCG LCGNFnIFAE DDFMTQEGTL TSDPYDFANS</u> <u>WALSSGEQWC</u> 201 <u>ERASPPSSC NISSGEMQKG LWEQCQLLKs TSVFARCHPL</u> <u>VDPEPFVALC</u> 251 <u>EKTLCCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA</u> <u>CSPVCPAGME</u> 301 <u>YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG</u> <u>LCVESTECPc</u> 351 <u>VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV</u> <u>TGQSHFKSFD</u> 401 <u>NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC</u> <u>TRSVTVRLPG</u> 451 <u>LHNSLVKLKH GAGVAMDQD IQPLLKGDL RIQHTVTASV</u> <u>RLSYGEDLQM</u> 501 <u>DWDGRGRLLV KLSPPVYAGKT CGLCGNYNGN QGDDFLTPSG</u> <u>LAEPRVEDFG</u> 551 <u>NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP</u> <u>TFEACHRAVS</u> 601 <u>PLEYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGRV</u> <u>AWREPGRC</u> 651 <u>NCPKGQVYLO CGTPCNLTCR SLSYPDEECN EACLEGCFCP</u> <u>PGLYMDERGD</u> 701 <u>CVPKAQCPY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM</u> <u>SGVVGSSLPD</u> 751 <u>AVLSSPLSHR SKR</u>	763
VWF D* 도메인	764 <u>SLSCRPP MVKLVCPADN LRAEGLCTK</u> 801 <u>TCQNYDLECM</u> <u>SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE</u> <u>TVKIGCNTV</u> 851 <u>CRDRKWNCTD HVCDAT</u>	866
VWF D3 도메인	867 <u>CSTI GMAHYLTFDG LKYLEPGEHQ</u> 901 <u>YVLLVQDYCGS</u> <u>NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE</u> <u>VNVKRPMKDE</u> 951 <u>THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE</u> <u>KVCGLCGNFD</u> 1001 <u>GIONNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD</u> <u>SSPATICNNI</u> 1051 <u>MKOTMVDSSC RILTSDFVFQD CNKLVDPEPY LDVCIYDTCS</u> <u>CESIGDCACF</u> 1101 <u>CDTIAAYAHV CAQHGKVVTW RTATLCPOQSC EERNLRENGY</u> <u>ECEWRYNSCA</u>	

[0106]

	1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLOTC VDPEDCPVCE 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP 1240
VWF A1 도메인	1241 GGLVVPPTDA 1251 PVSPTTLYVE DISEPPLHDF YCSRLLDLVF LLDGSSRLSE AEFEVLKAFV 1301 VDMMERLRLIS QKWVRVAVVE YHDGSHAYIG IKDRKRSEL RRIASQVKYA 1351 GSQVASTSEV LKYTLFQIFS KIDRPEASRI ALLLMASQEP QRMSRNFKVRY 1401 VQGLKKKKVI VIPVGIGPHA NLKQIRLIEK QAPENKAFVL SSVDELEQQR 1451 DEIVSYLCDL APEAPPPTLP PDMAQVTVG 1479
	1480 P GLLGVSTLGP KRNSMVLDVA 1501 FVLEGSDKIG EADFNRSKEF MEEVIQRMDV GDSDIHTVTL QYSYMTVVEY 1551 PFSEAQSKGD ILQRVREIRY QGGNRNTNTGL ALRYLSDHSF LVSQGDREQA 1600 1601 PNLYMVTGN PASDEIKRLP GDIQVVPIGV GPNANVQELE RIGWPNAPI 1651 IQDFETLPRE APDLVLQRCC SGEGLQIPTL SPAPDCSQPL DVILLLDGSS 1701 SFPASYFDEM KSFAKAFISK ANIGPRLTQV SVLQYGSITT IDVPWNVVP 1751 KAHLSSLVDV MQREGGPSQI GDALGFAVRY LTSEMHGARP GASKAVVILV 1801 TDVSVDSDVA AADAARSNRV TVFPIGIGDR YDAAQLRILA GPAGDSNVVK 1851 LQRIEDLPTM VTLGNSFLHK LCSGFVRICM DEDGNEKRP DVWTLPDQCH 1901 TVTCQPDGQT LLKSHRVNCD RGLRPSCPNS QSPVKVEETC GCRWTCPVC 1951 TGSSTRHIVT FDGQNFKLTG SCSYVLFQNK EQDLEVILHN GACSPGARQG 2001 CMKSIEVKHS ALSVEXHSDM EVTVNGRLVS VPYVGGNMEV NVYGAIMHEV 2051 RFNHIGHIFT FTPQNNEFQL QLSPKTFAST TYGLCGICDE NGANDFMLRD 2101 GTVITDWKTL VQEWTVQRPG QTCQPILEEQ CLVPDSSHQC VLLLPLFAEC 2151 HKVLIAPATFY AICQQDSCHQ EQVCEVIASY AHLCRTNGVC VDWRTPDFCA 2201 MSCPPSLVYN HCEHGCPRHC DGNVSSCGDH PSEGCFCPD KVMLEGSCVP 2251 EEACTQCIGE DGVQHQFLEA WVPDHQPCQI CTCLSGRKVN CTTQPCPTAK 2301 APTCGLCEVA RLQNADQCC PEYECVCDPV SCIDLPPVPHC ERGLQFTLTN 2351 PGECRPNFTC ACRKEECKRV SPPSCPPHRL PTLRKTQCCD EYECACNCVN 2401 STVSCPLGYL ASTATNDCGC TTTTCLPDFKV CVHRSTIYPV GQFWEEGCDV 2451 CTCTDMEDAV MGLRVAQCSQ KPCEDSCRSG FTYVLHEGEC CGRCLPSACE

[0107]

	2501 VVTGSPRGDS QSSWKSVGSQ WASPENPCLI NECVRVKEEV FIQQRNVS 2551 QLEVPCPSG FQLSCKTSAC CPSCRERME ACMLNGTIVG PGKTVMDVC 2601 TTCCRMMVQVG VISGFKLECR KTTCNPCPLG YKEENNNTGEC CGRCLPTACT 2651 IQLRGQQIMT LKRDETLQDG CDTHFCKVNE RGEYFWEKRV TGCPPFDEHK 2701 CLAEGGKIMK IPIGTCCDTCE EPECNDITAR LQYVKVGSK SEVEVDIHYC 2751 QGKCASKAMY SIDINDVQDQ CSCCSPTRTE PMQVALHCTN GSVVYHEVLN 2801 AMECKCSPRK CSK
	뉴클레오파이드 서열(서열번호: 1)
전장의 VWF	1 ATGATTCCCTG CCAGATTTCG CGGGGTGCTG CTTGCTCTGG CCCTCATTTT 51 GCCAGGGACC CTTTGTGCAG AAGGAACCTCG CGGCAGGTCA TCCACGGCCC 101 GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG 151 TACAGCTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA 201 ACGCTCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC 251 TCTCCGTGA TCTTGGGAA TTTTTGACA TCCATTGTT GTCAATGGT 301 ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG 351 GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCC 401 ATGGCTTGTG GCCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG 451 TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTGGCA ACTTTAACAT 501 CTTTGTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC 551 CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGCTGT 601 GAACGGGCAT CTCCCTCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT 651 GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT 701 TTGCCCCCTG CCACCCCTCTG GTGGACCCCG AGCCTTTGT GGCCCTGTGT 751 GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CTCGCCCTGC 801 CCTCCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAATG GTGCTGTACG 851 GCTGGACCGA CCACAGCGCG TGCAGCCAG TGTGCCCTGC TGGTATGGAG 901 TATAGGCAGT GTGTGCCCC TTGGCCAGG ACCTGCCAGA GCCTGCACAT 951 CAATGAAAATG TGTCAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG 1001 GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGTCCCTGC

[0108]

	1051      GTGCATTCCG GAAAGCGCTA CCCTCCCGC ACCTCCCTCT CTCGAGACTG 1101      CAACACCTGC ATTTGCCGAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT 1151      GTCAGGGGA GTGCCTTGTC ACTGGTCAAT CCCACTCAA GAGCTTTGAC 1201      AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGCCCCGGGA 1251      TTGCCAGGAC CACTCCCTCT CCATTGTCAAT TGAGACTGTC CAGTGTGCTG 1301      ATGACCGCGA CGCTGTGTC ACCCGCTCCG TCACCGTCCG GCTGCCTGGC 1351      CTGCACACA GCCTTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA 1401      TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC 1451      ATACAGTGAC GGCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG 1501      GACTGGGATG GCCGCCGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC 1551      CGGGAAGACC TGCGGCCGTG GTGGGAATTA CAATGCCAAC CAGGGCGACG 1601      ACTTCCTTAC CCCCTCTGGG CTGGCRGAGC CCCGGGTGGA GGACTTCGGG 1651      AACGCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG 1701      CGATCCCTGC GCCCTCAACC CGCCATGAC CAGGTTCTCC GAGGAGGCCTG 1751      GCGCGGTCTC GACGTCCCC ACATTCGAGG CCTGCCATCG TGCGGTCAAGC 1801      CGCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACCGTGTCT CCTGCTCGGA 1851      CGGCCCGAG TGCCTGTGCG GCGCCCTGGC CAGCTATGCC CGGCCCTGCG 1901      CGGGGAGAGG CGTGCCTGTC GCGTGGCGCG AGCCAGGGCG CTGTGAGCTG 1951      AACTGCCCGA AAGGCCAGGT GTACCTGCAG TGCAGGGACCC CCTGCAACCT 2001      GACCTGCCGC TCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCTGCC 2051      TGGAGGGCTG CTCTGCCCT CCAGGGCTCT ACATGGATGA GAGGGGGGAC 2101      TGCGTGCCTCA AGGCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA 2151      GCCAGAAAGAC ATCTTCTAG ACCATCACAC CATGTGCTAC TGTTGAGGATG 2201      GCTTCATGCA CTGTACCATG AGTGGAGTCC CCGGAAGCTT GCTGCCTGAC 2251      GCTGTCCCTCA GCAGTCCCCT GTCTCATCGC AGCAAAAGGA GCCTATCCTG 2301      TCGCCCCCCC ATGGTCAAGC TGGTGTGTCC CGCTGACAAC CTGCGGGCTG 2351      AAGGGCTCGA GTGTACCAAA ACGTGCAGA ACTATGACCT GGAGTGCATG 2401      AGCATGGGCT GTGTCTCTGG CTGCCCTG CCCCCGGGCA TGGCCGGCA 2451      TGAGAACAGA TGTTGCGGCC TGGAAAGGTG
--	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

[0109]

	TCCCTGCTTC CATCAGGGCA 2501 AGGAGTATGC CCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC 2551 TGTGGGGAC GGAAGTGGAA CTGCACAGAC CATGTGTGTG ATGCCACGTG 2601 CTCCACGATC GGCATGGCCC ACTACCTCAC CTTCGACGGG CTCAAATACC 2651 TGTTCCCCGG GGAGTGCCAG TACGTTCTGG TGCAGGATTA CTGCGGCAGT 2701 AACCTGGGA CCTTCGGAT CCTAGTGGGG ATAAAGGGAT GCAGCCACCC 2751 CTCAGTAAAA TGCAAGAAC GGGTCACCAT CCTGGTGGAG GGAGGAGAGA 2801 TTGAGCTGTT TGACGGGAG GTGAATGTGA AGAGGGCCAT GAAGGATGAG 2851 ACTCACTTG AGGTGGTGGA GTCTGGCCGG TACATCATTC TGCTGCTGGG 2901 CAAAGCCCTC TCCGTGGTCT GGGACCGCCA CCTGACCATC TCCGTGGTCC 2951 TGAAGCAGAC ATACCAGGAG AAAGTGTGTG GCCTGTGTGG GAATTGTGAT 3001 GGCATCCAGA ACAATGACCT CACCAGCAGC AACCTCCAAG TGGAGGAAGA 3051 CCCTGTGGAC TTTGGAACT CCTGGAAAGT GAGCTGGAG TGTGCTGACA 3101 CCAGAAAAGT GCCTCTGGAC TCATCCCCTG CCACCTGCCA TAACAACATC 3151 ATGAAGCAGA CGATGGTGGA TTCCCTCTGT AGAATCCTTA CCAGTGACGT 3201 CTTCCAGGAC TGCAACAAGC TGGTGGACCC CGAGCCATAT CTGGATGTCT 3251 GCATTTACGA CACCTGCTCC TGTGAGTCCA TTGGGACTG CGCCTGCTTC 3301 TCGGACACCA TTGCTGCCTA TGCCCACGTG TGTGCCAGC ATGGCAAGGT 3351 GGTGACCTGG AGGACGGCCA CATTGTGCC CCAGACCTGC GAGGAGAGA 3401 ATCTCCGGGA GAACGGGTAT GAGTGTGAGT GGCGCTATAA CAGCTGTGCA 3451 CCTGCCTGTC AAGTCACGTG TCAGCACCC GAGCCACTGG CCTGCCCTGT 3501 GCAGTGTGTG GAGGGCTGCC ATGCCACTG CCCTCCAGGG AAAATCCTGG 3551 ATGAGCTTT GCAGACCTGC GTTGACCCCTG AAGACTGTCC AGTGTGTGAG 3601 GTGGCTGCC GGCGTTTGTC CTCAGGAAAG AAAGTCACCT TGAATCCAG 3651 TGACCCCTGAG CACTGCCAGA TTTGCCACTG TGATGTTGTC AACCTCACCT 3701 GTGAAGCCTG CCAGGAGCCG GGAGGCCTGG TGGTGCCTCC CACAGATGCC 3751 CCGGTGAGCC CCACCACTCT GTATGTGGAG GACATCTCGG AACCGCCGTT 3801 GCACGATTC TACTGCAGCA GGCTACTGG CCTGGTCTTC CTGCTGGATG 3851 GCTCCTCCAG GCTGTCCGAG GCTGAGTTG AAGTGTGAA GGCCTTTGTG
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

[0110]

	3901 GTGGACATGA TGGAGCGGCT GCGCATCTCC CAGAAGTGGG TCCGCCTGGC 3951 CGTGGTGGAG TACCAAGACG GCTCCCACGC CTACATCGGG CTCAAGGACC 4001 GGAAGCGAAC GTCAGAGCTG CGGCCTGATTG CCAGGCCAGGT GAAGTATGCC 4051 GGCAGCCAGG TGGCCTCCAC CAGCGAGGTC TGAAATACA CACTGTTCCA 4101 AATCTTCAGC AAGATCGACC GCCCTGAAGC CTCCCGCATC GCCCTGCTCC 4151 TGATGGCCAG CCAGGAGCCC CAACGGATGT CCCGGAACCT TGTCCGCTAC 4201 GTCCAGGGCC TGAAGAAGAA GAAGGGTCATT GTGATCCCGG TGGGCATTGG 4251 GCCCCATGCC AACCTCAAGC AGATCCGCCT CATCGAGAAG CAGGCCCCCTG 4301 AGAACAAAGGC CTTCGTGCTG AGCAGTGTGG ATGAGCTGGA GCAGCAAAGG 4351 GACGAGATCG TTAGCTACCT CTGTGACCTT GCCCTGAAAG CCCCTCCTCC 4401 TACTCTGCC CCGACATGG CACAAGTCAC TGTGGCCCCG GGCTCTTG 4451 GGGTTTCGAC CCTGGGGCCC AAGAGGAAC CCATGGTCT GGATGTGGCG 4501 TTCGTCTGG AAGGATCGGA CAAAATTGGT GAAGCCGACT TCAACAGGAG 4551 CAAGGAGTTC ATGGAGGGAGG TGATTCAAGCG GATGGATGTG GGCCAGGACA 4601 GCATCCACGT CACGGTGCTG CAGTA ACATGGTGAC CGTGGAGTAC 4651 CCCTTCAGCG AGGCACAGTC CAAAGGGGAC ATCCCTGAGC GGCTGCCAGA 4701 GATCCGCTAC CAGGGCGCA ACAGGACCAA CACTGGGCTG GCCCTGCGGT 4751 ACCTCTCTGA CCACAGCTTC TTGGTCAGCC AGGGTGACCG GGAGCAGGCG 4801 CCCAACCTGG TCTACATGGT CACCGAAAT CCTGCTCTG ATGAGATCAA 4851 GAGGCTGCC GGAGACATCC AGGTGGTGCC CATTGGAGTG GGCCCTAATG 4901 CCAACGTGCA GGAGCTGGAG AGGATTGGCT GGCCCAATGC CCCTATCCTC 4951 ATCCAGGACT TTGAGACGCT CCCCCGAGAG GCTCCCTGACC TGTTGCTGCA 5001 GAGGTGCTGC TCCGGAGAGG GGCTGCAGAT CCCCACCTC TCCCTGCA 5051 CTGACTGCAG CCAGCCCCCTG GACGTGATCC TTCTCCCTGGA TGGCTCCTCC 5101 AGTTTCCAG CTTCTTATTT TGATGAAATG AAGAGTTTCG CCAAGGCTTT 5151 CATTCAAAA GCCAATATAG GGCTCGTCT CACTCAGGTG TCAGTGTG 5201 AGTATGGAAG CATCACCAC ATTGACGTG CATGGAACGT GGTCCCGGAG 5251 AAAGCCCCATT TGCTGAGCCT TGTGGACGTC ATGCAGCGGG AGGGAGGGCC 5301 CAGCCAATAC GGGGATGCCT TGGGCTTTGC
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

[0111]

	TGTGCGATAC TTGACTTCAG 5351 AAATGCATGG TGCCAGGCCG GGAGCCTCAA AGGCAGGTGGT CATCCTGTC 5401 ACGGACGCT CTGTGGATTG AGTGGATGCA GCAGCTGATG CGGCCAGGTC 5451 CAACAGACTG ACAGTGGTCC CTATTGGAAT TGGAGATCGC TACGATGCAG 5501 CCCAGCTACG GATCTGGCA GGCCCAGCAG GCGACTCCAA CGTGGTGAAG 5551 CTCCAGCGAA TCGAAGACCT CCCTACCATG GTCACCTTGG GCAATTCCIT 5601 CCTCCACAAA CTGTGCTCTG GATTGTTAG GATTTCATG GATGAGGATG 5651 GGAATGAGAA GAGGCCGGG GACGTCTGGA CCTTGCCAGA CCAGTGCCAC 5701 ACCGTGACTT GCCAGCCAGA TGGCCAGACC TTGCTGAAGA CTCATCGGGT 5751 CAACTGTGAC CGGGGCTGA GGCCTTCGTG CCCTAACAGC CAGTCCCCGT 5801 TTAAAGTGGA AGAGACCTGT GGCTGCCGCT GGACCTGCCC CTGYGTGTGC 5851 ACAGGCACCT CCACTCGGCA CATCGTGACC TTTGATGGGC AGAATTCAA 5901 GCTGACTGGC AGCTGTTCTT ATGTCCTATT TCAAAACAAG GAGCAGGAC 5951 TGGAGGTGAT TCTCCATAAT GGTGCCCTGCA GCCCTGGAGC AAGGCAGGGC 6001 TGCATGAAT CCATCGAGGT GAAGCACAGT GCCCTCTCCG TCGAGSTGCA 6051 CAGTGACATG GAGGTGACGG TGAATGGGAG ACTTGGTCTCT GTTCCTTACG 6101 TGGGTGGGAA CATGGAAGTC AACGTTTATG TGGCCATCAT CCATGAGGTC 6151 AGATTCAATC ACCTTGGTCA CATCTTCACA TTCACTCCAC AAAACAATGA 6201 GTTCCAACCTG CAGCTCAGCC CCAAGACTTT TGCTTCAAAG ACGTATGGTC 6251 TGTGTGGGAT CTGTGATGAG AACGGAGCCA ATGACTTCAT GCTGAGGGAT 6301 GGCACAGTCA CCACAGACTG GAAAACACTT GTTCAGGAAT CGACTGTGCA 6351 GCGGCCAGGG CAGACGTGCC AGCCCATCCT GGAGGAGCAG TGTCTTGTCC 6401 CCGACAGTCC CCACTGCCAG GTCCTCCTCT TACCACTGTT TGCTGAATGC 6451 CACAAGGTCC TGGCTCCAGC CACATTCTAT GCCATCTGCC AGCAGGACAG 6501 TTGCCACCA GAGCAAGTGT GTGAGGTGAT CGCCCTTTAT CCCACCTCT 6551 GTCGGACCAA CGGGGCTGTC GTTGACTGGA GGACACCTGA TTTCTGTGCT 6601 ATGTCATGCC CACCATCTCT GGTCTACAAC CACTGTGAGC ATGGCTGTCC 6651 CGGGCACTGT GATGGAACAG TGAGCTCCTG TGGGGACCAT CCCTCCGAAG 6701 GCTGTTTCTG CCCTCCAGAT AAAGTCATGT TGGAAAGGCAG CTGTGTCCCT
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

[0112]

	6751 GAAGAGGCCT GCACTCAGTG CATTGGTGAG GATGGAGTCC AGCACCAAGTT 6801 CCTGGAAGCC TGGGTCCCGG ACCACCAGCC CTGTCAGATC TGCACATGCC 6851 TCAGCGGGCG GAAGGTCAAC TGCACAACGC AGCCCTGCC CACGGCCAAA 6901 GCTCCCACGT GTGGCCTGTG TGAAGTAGCC CGCCTCCGCC AGAATGCAGA 6951 CCAGTGTGC CCCGAGTATG AGTGTGTGTG TGACCCAGTG AGCTGTGACC 7001 TGCCCCCAGT GCCTCACTGT GAACGTGGCC TCCAGCCCCAC ACTGACCAAC 7051 CCTGGCGAGT GCAGACCCAA CTTCACCTGC GCCTGCAGGA AGGAGGAGTG 7101 CAAAAGAGTG TCCCCACCCCT CCTGCCCCCC GCACCGTTTG CCCACCCCTC 7151 GGAAGACCCA GTGCTGTGAT GAGTATGAGT GTGCTTGCAA CTGTGTCAAC 7201 TCCACAGTGA GCTGTCCCCT TGGGTACTTG GCCTCAACCG CCACCAATGTA 7251 CTGTTGGCTGT ACCACACACCA CCTGCCTTCC CGACAAGGTG TGTGTCCACC 7301 GAAGCACCAT CTACCCGTG GGCCAGTTCT GGGAGGAGGG CTGCGATGTG 7351 TGCACCTGCA CGCACATGGA GGATGCCGTG ATGGGCCTCC CGGTGGCCCA 7401 GTGCTCCAG AAGCCCTGTG AGGACAGCTG TCGGTGGGG ITCACTTACG 7451 TTCTGCATGA AGGCGAGTGC TGTGGAAGGT GCCTGCCATC TGCCTGTGAG 7501 GTGGTGACTG GCTCACCGCG GGGGGACTCC CAGTCTTCCCT GGAAGAGTGT 7551 CGGCTCCAG TGGGCCTCCC CGGAGAACCC CTGCCTCATC AATGAGTGTG 7601 TCCGAGTGAA GGAGGAGGTC TTTATACAAC AAAGGAACGT CTCCCTGCC 7651 CAGCTGGAGG TCCCTGTCTG CCCCTCGGGC TTTCAGCTGA GCTGTAAGAC 7701 CTCAGCGTGC TGCCCAAGCT GTGCGCTGTGA GCGCATGGAG GCCTGCATGC 7751 TCAATGCCAC TGTGATATTGGG CCCGGGAAGA CTGTTGATGAT CGATGTGTG 7801 ACGACCTGCC GCTGCACTGGT GCAGGTGGGG GTCATCTCTG GATTCAAGCT 7851 GGAGTGCGAGG AAGACCACCT GCAACCCCTG CCCCCTGGGT TACAAGGAAG 7901 AAAATAAACAC AGGTGAATGT TGTGGGAGAT GTTTGCCTAC GGCTTGACCC 7951 ATTCAAGCTAA GAGGAGGACA GATCATGACA CTGAAAGCGTG ATGAGACGCT 8001 CCAGGATGGC TGTGATACTC ACTTCTGCAA GGTCAATGAG AGAGGGAGGT 8051 ACTTCTGGGA GAAGAGGGTC ACAGGCTGCC CACCCCTTGA TGAACACAAG 8101 TGTCTTGTG AGGGAGGTAA AATTATGAAA ATTCCAGGCA CCTGCTGTGA 8151 CACATGTGAG GAGCCTGAGT GCAACGACAT
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

[0113]

	CACTGCCAGG CTGCAGTATG 8201 TCAAGGTGGG AAGCTGTAAG TCTGAAGTAG AGGTGGATAT CCACTACTGC 8251 CAGGGCAAAT GTGCCAGCAA AGCCATGTAC TCCATTGACA TCAACGATGT 8301 GCAGGACCAAG TGCTCTGTCT GCTCTCCGAC ACGGACGGAG CCCATGCAGG 8351 TGGCCCTGCA CTGCACCAAT GGCTCTGTG TGTACCATGA GGTTCTCAAT 8401 GCCATGGAGT GCAAATGCTC CCCCAGGAAG TGCAGCAAGT GA
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

[0114]

[0115] 본원에 사용된 것으로서 VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함할 수 있으며, 여기서 VWF 단백질은 FVIII에 결합하여 FVIII에 대한 내인성 VWF(전장의 VWF)의 결합을 억제한다. D' 도메인 및 D3 도메인을 포함하는 VWF 단백질은 A1 도메인, A2 도메인, A3 도메인, D1 도메인, D2 도메인, D4 도메인, B1 도메인, B2 도메인, B3 도메인, C1 도메인, C2 도메인, CK 도메인, 이의 하나 이상의 단편, 또는 이의 어떠한 조합으로부터 선택된 VWF 도메인을 추가로 포함할 수 있다. 하나의 구현예에서, VWF 단백질은 (1) VWF의 D' 및 D3 도메인 또는 이의 단편; (2) VWF의 D1, D', 및 D3 도메인 또는 이의 단편; (3) VWF의 D2, D', 및 D3 도메인 또는 이의 단편; (4)

VWF의 D1, D2, D', 및 D3 도메인 또는 이의 단편; 또는 (5) VWF의 D1, D2, D', D3, 또는 A1 도메인 또는 이의 단편을 포함하거나, 이로 필수적으로 이루어지거나, 이로 이루어진다. 본원에 기술된 VWF 단백질은 VWF 정화 수용체 결합 부위를 함유하지 않는다. 본 발명의 VWF 단백질은 VWF 단백질에 연결되거나 융합된 어떠한 다른 서열도 포함할 수 있다. 예를 들면, 본원에 기술된 VWF 단백질은 시그널 웹타이드를 추가로 포함할 수 있다.

[0116] 하나의 구현예에서, VWF 단백질은 FVIII 단백질에 결합하거나 이와 연합된다. FVIII 단백질에 결합되거나 연합됨으로써, 본 발명의 VWF 단백질은 프로테아제 절단 및 FVIII 활성화로부터 FVIII을 보호하고, FVIII의 중쇄 및 경쇄를 안정화시키고, 스캐진저 수용체에 의한 FVIII의 정화를 방지한다. 다른 구현예에서, VWF 단백질은 FVIII 단백질에 결합하거나 이와 연합하거나 인지질 또는 활성화 단백질 C에 대한 FVIII 단백질의 결합을 차단하거나 방지한다. FVIII 단백질과 내인성의 전장의 VWF의 결합을 방지하거나 억제함으로써, 본 발명의 VWF 단백질은 내인성 VWF 정화 수용체에 의한 FVIII의 정화를 감소시킴으로써 FVIII 단백질이 반감기를 연장시킨다. 따라서, FVIII 단백질의 반감기 연장은 VWF 정화 수용체 결합 부위를 결여하는 VWF 단백질과 FVIII 단백질이 연합 및 이에 의한 VWF 정화 수용체 결합 부위를 함유하는 내인성 VWF으로부터 FVIII 단백질의 차폐 및/또는 보호에 기인한다. VWF 단백질에 결합하거나 이에 의해 보호된 FVIII 단백질은 FVIII 단백질의 재순환을 허용할 수 있다. 전장의 VWF 분자내 VWF 정화 경로 수용체 결합 부위를 제거함으로써, 본 발명의 FVIII/VWF 이종이량체는 VWF 정화 경로로부터 차폐되어, FVIII 반감기를 추가로 연장시킨다.

[0117] 하나의 구현예에서, 본 발명의 VWF 단백질은 VWF의 D' 도메인 및 VWF의 D3 도메인을 포함하고, 여기서 D' 도메인은 서열 번호 2의 764 내지 866번 아미노산에 대해 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하고, 여기서 VWF 단백질은 FVIII에 대한 내인성 VWF의 결합을 방지한다. 다른 구현예에서, VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함하며, 여기서 D3 도메인은 서열 번호 2의 867 내지 1240번 아미노산에 대해 적어도 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하고, 여기서 VWF 단백질은 FVIII에 대한 VWF 결합을 방지한다. 일부 구현예에서, VWF 단백질은 VWF의 D' 도메인 및 D3 도메인을 포함하며, 이들은 서열 번호 2의 764 내지 1240번 아미노산에 대해 적어도 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하고, 여기서 VWF 단백질은 FVIII에 대한 내인성 VWF의 결합을 방지한다. 다른 구현예에서, VWF 단백질은 서열 번호 2의 23 내지 1240번 아미노산에 대해 적어도 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 D1, D2, D', 및 D3 도메인을 포함하거나, 이로 필수적으로 이루어지거나 이루어지고, 여기서 VWF 단백질은 FVIII에 대한 VWF 결합을 방지한다. 여전히 다른 구현예에서, VWF 단백질은 이에 연결된 시그널 웹타이드를 추가로 포함한다.

[0118] 일부 구현예에서, 본 발명의 VWF 단백질은 (1) D'D3 도메인, D1D'D3 도메인, D2D'D3 도메인, 또는 D1D2D'D3 도메인 및 (2) 약 10개 이하의 아미노산(예를 들면, 서열 번호 2의 764 내지 1240번 아미노산 내지 서열 번호 2의 764 내지 1250번 아미노산으로부터의 어떠한 서열), 약 15개 이상의 아미노산(예를 들면, 서열 번호 2의 764 내지 1240번 아미노산 내지 서열 번호 2의 764 내지 1255번 아미노산으로부터의 어떠한 서열), 약 20개 이하의 아미노산(예를 들면, 서열 번호 2의 764 내지 1240번 아미노산 내지 서열 번호 2의 764 내지 1260번 아미노산으로부터의 어떠한 서열), 약 25개 아미노산(예를 들면, 서열 번호 2의 764 내지 1240번 아미노산 내지 서열 번호 2의 764 내지 1265번 아미노산으로부터의 어떠한 서열), 또는 약 30개 아미노산(예를 들면, 서열 번호 2의 764 내지 1240번 아미노산 내지 서열 번호 2의 764 내지 1260번 아미노산으로부터의 어떠한 서열)로 필수적으로 이루어지거나 이로 이루어진다. 특수한 구현예에서, D' 도메인 및 D3 도메인을 포함하거나 이로 필수적으로 이루어진 VWF 단백질은 서열 번호 2의 764 내지 1274번 아미노산도 전장의 성숙한 VWF도 아니다. 일부 구현예에서, D1D2 도메인은 D'D3 도메인과 트랜스 배향(in trans)으로 발현된다. 일부 구현예에서, D1D2 도메인은 D'D3 도메인과 시스 배향(in cis)으로 발현된다.

[0119] 다른 구현예에서, D1D2 도메인에 연결된 D'D3 도메인을 포함하는 VWF 단백질은 세포내 프로세싱 부위, 예를 들면, (PACE(푸린) 또는 PC5에 의한 프로세싱 부위)를 추가로 포함함으로써, 발현시 D'D3 도메인으로부터 D1D2 도메인의 절단을 허용한다. 세포내 프로세싱 부위의 비-제한적 예는 본원의 어딘가에 개재되어 있다.

[0120] 여전히 다른 구현예에서, VWF 단백질은 D' 도메인 및 D3 도메인을 포함하지만, (1) 서열 번호 2의 1241 내지 2813번 아미노산, (2) 서열 번호 2의 1270 내지 2813번 아미노산, (3) 서열 번호 2의 1271 내지 2813번 아미노산, (4) 서열 번호 2의 1272 내지 2813번 아미노산, (5) 서열 번호 2의 1273 내지 2813번 아미노산, (6) 서열 번호 2의 1274 내지 2813번 아미노산, 또는 이의 어떠한 조합으로부터 선택된 아미노산 서열을 포함하지 않는다.

[0121] 여전히 다른 구현예에서, 본 발명의 VWF 단백질은 D' 도메인, D3 도메인, A1 도메인을 포함하거나, 이로 필수적

으로 이루어지거나, 이로 이루어지며, 여기서 아미노산 서열은 서열 번호 2의 764 내지 1479번 아미노산에 대해 적어도 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하고, 여기서 VWF 단백질은 FVIII에 대한 내인성 VWF의 결합을 방지한다. 특수한 구현예에서, VWF 단백질은 서열 번호 2의 764 내지 1274번 아미노산이 아니다.

[0122] 일부 구현예에서, 본 발명의 VWF 단백질은 D' 도메인 및 D3 도메인을 포함하나, (1) A1 도메인, (2) A2 도메인, (3) A3 도메인, (4) D4 도메인, (5) B1 도메인, (6) B2 도메인, (7) B3 도메인, (8) C1 도메인, (9) C2 도메인, (10) CK 도메인, (11) CK 도메인 및 C2 도메인, (12) CK 도메인, C2 도메인, 및 C1 도메인, (13) CK 도메인, C2 도메인, C1 도메인, B3 도메인, (14) CK 도메인, C2 도메인, C1 도메인, B3 도메인, B2 도메인, (15) CK 도메인, C2 도메인, C1 도메인, B3 도메인, B2 도메인, B1 도메인, (16) CK 도메인, C2 도메인, C1 도메인, B3 도메인, B2 도메인, B1 도메인, D4 도메인, 및 A3 도메인, (17) CK 도메인, C2 도메인, C1 도메인, B3 도메인, B2 도메인, B1 도메인, D4 도메인, A3 도메인, 및 A2 도메인, (19) CK 도메인, C2 도메인, C1 도메인, B3 도메인, B2 도메인, B1 도메인, D4 도메인, A3 도메인, A2 도메인, 및 A1 도메인, 또는 (20) 이의 어떠한 조합으로부터 선택된 적어도 하나의 VWF 도메인을 포함하지 않는다.

[0123] 여전히 다른 구현예에서, VWF 단백질은 D'D3 도메인 및 하나 이상의 도메인 또는 모듈(module)을 포함한다. 이러한 도메인 또는 모듈의 예는 문헌[참조: Zhour 등, Blood published online April 6, 2012: DOI 10.1182/blood-2012-01-405134]에 개재된 도메인 및 모듈을 포함하나, 이에 한정되지 않는다. 예를 들면, VWF 단백질은 D'D3 도메인 및 A1 도메인, A2 도메인, A3 도메인, D4N 모듈, WWD4 모듈, C8-4 모듈, TIL-4 모듈, C1 모듈, C2 모듈, C3 모듈, C4 모듈, C5 모듈, C6 모듈, 또는 이의 어떠한 조합으로부터 선택된 하나 이상의 도메인 또는 모듈을 포함할 수 있다.

[0124] 특정의 구현예에서, 본 발명의 VWF 단백질은 다량체, 예를 들면, 이량체, 삼량체, 사량체, 오량체, 육량체, 칠량체, 또는 보다 높은 차수의 다량체를 형성한다. 다른 구현예에서, VWF 단백질은 단지 하나의 VWF 단백질을 지닌 단량체이다. 일부 구현예에서, 본 발명의 VWF 단백질은 하나 이상의 아미노산 치환, 결실, 첨가, 또는 변형을 가질 수 있다. 하나의 구현예에서, VWF 단백질은 아미노산 치환, 결실, 첨가 또는 변형을 포함함으로써, VWF 단백질은 이황화물 결합을 형성할 수 없거나 이량체 또는 다량체를 형성할 수 없다. 다른 구현예에서, 아미노산 치환은 D' 도메인 및 D3 도메인내에 있다. 특수한 구현예에서, 본 발명의 VWF 단백질은 서열 번호 2의 1099번 잔기, 1142번 잔기, 또는 1099번 및 1142번 잔기 둘 다에 상응하는 잔기에서 적어도 하나의 아미노산 치환을 함유한다. 적어도 하나의 아미노산 치환은 야생형 VWF에서 천연적으로 발생하지 않는 어떠한 아미노산일 수 있다. 예를 들어, 아미노산 치환은 시스테인외의 어떠한 아미노산, 예를 들면, 이소루이신, 알라닌, 루이신, 아스파라긴, 라이신, 아스프르트산, 메티오닌, 페닐알라닌, 글루탐산, 트레오닌, 글루타민, 트립토판, 글리신, 발린, 프롤린, 세린, 타이로신, 아르기닌, 또는 히스티딘일 수 있다. 다른 예에서, 아미노산 치환은 다량체를 형성하는 것으로부터 VWF 단백질을 방지하거나 억제하는 하나 이상의 아미노산을 가질 수 있다.

[0125] 일부 구현예에서, VWF 단백질은 VWF의 D'D3 도메인에 상응하는 336번 잔기(서열 번호 2의 1099번 잔기)에서 시스테인으로부터 알라닌으로의 아미노산 치환, 및 VWF의 D'D3 도메인에 상응하는 379번 잔기(서열 번호 2의 1142번 잔기)에서 시스테인으로부터 알라닌으로의 아미노산 치환을 포함한다.

[0126] 특정의 구현예에서, 본원에 유용한 VWF 단백질은 이의 FVIII와의 상호작용을 증진시키기 위해, 예를 들면, FVIII에 대한 친화성을 증진시키기 위해 추가로 변형될 수 있다. 비-제한적인 예로서, VWF 단백질은 서열 번호 2의 764번 아미노산에 상응하는 잔기에서 세린 잔기 및 서열 번호 2의 773번 아미노산에 상응하는 잔기에서 라이신 잔기를 포함한다. 잔기 764 및/또는 773은 FVIII에 대한 VWF 단백질의 결합 친화성에 기여할 수 있다. 다른 구현예에서, 본 발명에 유용한 VWF 단백질은 다른 변형을 지닐 수 있는데, 예를 들어, 단백질은 폐질화되거나, 글리코실화되거나, 헤실화되거나 폴리시알릴화될 수 있다.

### II.C.3. 이종 잔기

[0128] VWF 링커를 통해 VWF 단백질에 또는 FVIII 링커를 통해 FVIII 단백질에 융합될 수 있는 이종잔기는 이종 폴리펩타이드 또는 이종 비-폴리펩타이드 잔기일 수 있다. 특정의 구현예에서, 이종 잔기는 당해 분야에 공지된 반감기가 연장된 문자이고 폴리펩타이드, 비-폴리펩타이드 잔기, 또는 이를 둘 다의 조합을 포함한다. 이종 폴리펩타이드 잔기는 FVIII 단백질, 면역글로불린 불변 영역 또는 이의 일부, 알부민 또는 이의 단편, 알부민 결합 잔기, 트랜스페린 또는 이의 단편, PAS 서열, HAP 서열, 이의 유도체 또는 변이체, 사람 용모성 고나도트로핀의  $\beta$  소단위의 C-말단 웹타이드(CTP), 또는 이의 어떠한 조합을 포함할 수 있다. 일부 구현예에서, 비-폴리펩타이

드 결합 잔기는 폴리에틸렌 글리콜(PEG), 폴리시알산, 하이드록시에틸 전분(HES), 이의 유도체, 또는 이의 어떠한 조합을 포함한다. 특정의 구현예에서, 각각 동일하거나 상이한 문자일 수 있는, 1개, 2개, 3개 이상의 이종 잔기가 존재할 수 있다.

#### [0129] II.C.3.a 면역글로불린 불변 영역 또는 이의 부위

면역글로불린 불변 영역은 CH(불변 중쇄) 도메인(CH1, CH2 등)으로 나타낸 도메인으로 이루어진다. 동형, (즉, IgG, IgM, IgA IgD, 또는 IgE)에 따라서, 불변 영역은 3개 또는 4개의 CH 도메인으로 구성될 수 있다. 일부 동형(예를 들면 IgG) 불변 영역은 또한 힌지 영역을 함유한다. [참조: Janeway 등 2001, *Immunobiology*, Garland Publishing, N.Y., N.Y.]

본 발명의 키메라 단백질을 생산하기 위한 면역글로불린 불변 영역 또는 이의 일부는 다수의 상이한 공급원으로부터 수득될 수 있다. 일부 구현예에서, 면역글로불린 불변 영역 또는 이의 일부는 사람 면역글로불린으로부터 기원한다. 그러나, 면역글로불린 불변 영역 또는 이의 일부는 예를 들면, 설치류(예를 들면, 마우스, 랙트, 토끼, 기니아 퍼그) 또는 비-사람 영장류(예를 들면 침팬지, 마카크(macaque)) 종을 포함하는 다른 포유동물 종의 면역글로불린으로부터 기원할 수 있다. 더욱이, 면역글로불린 불변 영역 또는 이의 일부는 IgM, IgG, IgD, IgA and IgE를 포함하는 어떠한 면역글로불린 부류, 및 IgG1, IgG2, IgG3 및 IgG4를 포함하는 어떠한 면역글로불린 동형으로부터 기원할 수 있다. 하나의 구현예에서, 사람 동형 IgG1이 사용된다.

다양한 면역글로불린 불변 영역 유전자 서열(예를 들면 사람 불변 영역 유전자 서열)은 공공으로 접근가능한 기탁물의 형태로 이용가능하다. 불변 영역 도메인 서열은 특수한 효과기 기능을 가지거나(또는 특수한 효과기 기능을 결여하거나) 또는 면역원성을 감소시키기 위한 특수한 변형을 가지고도록 선택될 수 있다. 항체 및 항체를 암호화하는 유전자의 많은 서열은 발표되어 있으며 적합한 Ig 불변 영역 서열(예를 들면, 힌지, CH2, 및/또는 CH3 서열, 또는 이의 일부)은 당해 분야에 인식된 기술을 사용하여 이들 서열로부터 기원할 수 있다. 앞서의 방법 중 어느 것을 사용하여 수득된 유전 물질은 이후에 변경하거나 합성하여 본 발명의 폴리펩타이드를 수득할 수 있다. 본 발명의 영역은 불변 영역 DNA 서열의 대립형질, 변이체 및 돌연변이를 포함하는 것으로 추가로 인식될 것이다.

면역글로불린 고정 영역 또는 이의 일부의 서열은, 예를 들면, 목적한 도메인을 증폭시키기 위해 선택된 폴리머라제 쇄 반응 및 프라이머를 사용하여 클로닝될 수 있다. 항체로부터의 면역글로불린 불변 영역 또는 이의 일부의 서열을 클로닝하기 위하여, mRNA를 하이브리도마, 비장, 또는 림프 세포로부터 분리하여, DNA내로 역 전사시키고, 항체 유전자를 PCR에 의해 증폭시킬 수 있다. PCR 증폭 방법은 미국 특허 제 4,683,195호; 제4,683,202호; 제4,800,159호; 제4,965,188호; 및 문헌[참조: 예를 들면, "PCR Protocols: A Guide to Methods and Applications" Innis 등 eds., Academic Press, San Diego, CA (1990); Ho 등 1989. Gene 77:51; Horton 등 1993. Methods Enzymol. 217:270]에 상세히 기술되어 있다.

본원에 사용된 면역글로불린 불변 영역은 모든 도메인 및 힌지 영역 또는 이의 일부를 포함할 수 있다. 하나의 구현예에서, 면역글로불린 불변 영역 또는 이의 일부는 CH2 도메인, CH3 도메인, 및 힌지영역, 즉, Fc 영역 또는 FcRn 결합 파트너를 포함한다.

본원에 사용된 것으로서, 용어 "Fc 영역"은 천연의 면역글로불린의 Fc 영역에 상응하는 폴리펩타이드의 일부, 즉, 이의 2개의 중쇄의 각각의 Fc 도메인의 이량체성 연합에 의해 형성된 것으로서 정의된다. 천연의 Fc 영역은 다른 Fc 영역과의 동종이량체를 형성한다.

하나의 구현예에서, "Fc 영역"은 파파인 절단 부위의 바로 상부의 힌지 영역(즉, 114번인 중쇄 불변 영역의 제1 잔기로부터 취하여, IgG내 216번 잔기)에서 개시하여 항체의 C-말단에서 종결되는 단일의 면역글로불린 중쇄의 일부를 말한다. 따라서, 완전한 Fc 도메인은 적어도 힌지 도메인, CH2 도메인, 및 CH3 도메인을 포함한다.

면역글로불린 동형에 따라서, 면역글로불린 불변 영역의 Fc 영역은 CH2, CH3, 및 CH4 도메인, 및 힌지 영역을 포함할 수 있다. 면역글로불린의 Fc 영역을 포함하는 키메라 단백질은 증가된 안전성, 증가된 혈청 반감기(참조: Capon 등, 1989, *Nature* 337:525) 및 신생아 Fc 수용체(FcRn)와 같은 Fc 수용체에 대한 결합을 포함하는 키메라 단백질에 수개의 바람직한 특성을 부여한다(참조: 미국 특허 제 6,086,875호, 제6,485,726호, 제6,030,613호; 제WO 03/077834호; 제US2003-0235536A1호, 이의 전문은 본원에 참조로 포함된다).

면역글로불린 불변 영역 또는 이의 일부는 FcRn 결합 파트너일 수 있다. FcRn는 성인 내피 조직에서 활성이며 장, 폐 기도, 비강 표면, 질 표면, 결장 및 직장 표면의 내강(lumen)에서 발현된다(참조: 미국 특허 제

6,485,726호). FcRn 결합 파트너는 FcRn에 결합하는 면역글로불린의 일부이다.

[0139] FcRn 수용체는 사람을 포함하는 수개의 포유동물 종으로부터 분리된다. 사람 FcRn, 원숭이 FcRn, 랫 FcRn, 및 마우스 FcRn의 서열은 공지되어 있다(참조: Story 등 1994, J. Exp. Med. 180:2377). FcRn 수용체는 비교적 낮은 pH에서 IgG(그러나 IgA, IgM, IgD, 및 IgE와 같은 다른 면역글로불린은 아님)에 결합하여, 내강에서 장막 방향으로 IgG를 경세포적으로 활성적으로 수송한 후 사이질 유액에서 발견되는 비교적 높은 pH에서 IgG를 방출한다. 이는 성인 내피 조직(미국 특허 제 6,485,726호, 제6,030,613호, 제6,086,875호; 제WO 03/077834호; 제US2003-0235536A1호), 예를 들면, 폐 및 장 내피(참조: Israel 등 1997, Immunology 92:69) 신장 근위곡 관형 내피(renal tubular epithelium)(참조: Kobayashi 등 2002, Am. J. Physiol. Renal Physiol. 282:F358) 및 비강 내피, 질 표면, 및 담도계 표면에서 발현된다.

[0140] 본 발명에 유용한 FcRn 결합 파트너는 전체 IgG, IgG의 Fc 단편을 포함하는 FcRn 수용체, 및 FcRn 수용체의 완전한 결합 영역을 포함하는 다른 단편에 의해 특이적으로 결합될 수 있는 분자를 포함한다. FcRn 수용체에 결합하는 IgG의 Fc 부위의 영역은 X-선 결정학을 기본으로 기술되어 있다(참조: Burmeister 등 1994, Nature 372:379). Fc와 FcRn의 주요 접촉 부위는 CH2와 CH3 도메인의 연결부 근처이다. Fc-FcRn은 단일의 Ig 중쇄내에 모두 존재한다. FcRn 결합 파트너는 전체 IgG, IgG의 Fc 단편, 및 FcRn의 완전한 결합 영역을 포함하는 IgG의 다른 단편을 포함한다. 주요 접촉 부위는 CH2 도메인의 248, 250 내지 257, 272, 285, 288, 290, 291, 308 내지 311번 및 314 아미노산 잔기, 및 CH3 도메인의 385 내지 387, 428, 및 433 내지 436번 아미노산 잔기를 포함한다. 면역글로불린 또는 면역글로불린 단편, 또는 영역에 대해 이루어진 참조는 모두 문헌[참조: Kabat 등 1991, Sequences of Proteins of Immunological Interest, U.S. Department of Public Health, Bethesda, Md]를 기본으로 한다.

[0141] FcRn에 결합된 Fc 영역 또는 FcRn 결합 파트너는 FcRn에 의한 상피 장벽을 가로질러 효과적으로 왕복하여 바람직한 치료학적 분자를 전신계적으로 투여하기 위한 비-침입성 수단을 제공한다. 또한, Fc 영역 또는 FcRn 결합 파트너를 포함하는 융합 단백질은 FcRn을 발현하는 세포에 의해 세포내이입(endocytosis)된다. 그러나, 분해에 대해 표시하는 것 대신에, 이들 융합 단백질은 순환기내로 다시 재순환됨으로써 이들 단백질의 생체내 반감기를 증가시킨다. 특정의 구현예에서, 면역글로불린 불변 영역의 일부는 이황화물 결합 및 다른 비-특이적인 상호작용을 통해 Fc 영역 또는 다른 FcRn 결합 파트너와 전형적으로 연합하여 이량체 및 보다 높은 차수의 다량체를 형성하는 Fc 영역 또는 다른 FcRn 결합 파트너이다.

[0142] FcRn 결합 파트너 영역은 Fc 영역의 FcRn 수용체에 의해 후속적인 활성적 수송으로 FcRn 수용체에 의해 특이적으로 결합될 수 있는 분자 또는 이의 일부이다. '구체적으로 결합된'은 생리학적 조건하에서 비교적 안정한 복합체를 형성하는 2개의 분자를 말한다. 구체적인 결합은 일반적으로 중간 내지 고 능력의 낮은 친화성을 갖는 비특이적인 결합으로부터 구별되는 것으로서 고 친화성 및 낮은 내지 중간의 능력으로 특징화된다. 전형적으로, 결합은, 친화성 상수 KA가  $10^6 M^{-1}$  이상, 또는  $10^8 M^{-1}$  이상인 경우 특이적인 것으로 고려된다. 경우에 따라, 비-특이적인 결합은 결합 조건을 변화시킴으로써 특이적인 결합에 실질적으로 영향을 미치지 않고 감소될 수 있다. 분자의 농도, 용액의 이온 강도, 온도, 결합에 허여된 시간, 차단제(예를 들면 혈청 알부민, 우유 카제인), 등의 농도와 같은 적절한 결합 조건은 통사의 기술을 사용하여 기술자에 의해 최적화될 수 있다.

[0143] 미리아드 돌연변이체(Myriad mutant), 단편, 변이체, 및 유도체는 예를 들면, PCT 공보 제 WO 2011/069164 A2호, 제WO 2012/006623 A2호, 제WO 2012/006635 A2호, 또는 제WO 2012/006633 A2호에 기술되어 있으며, 이들 모두는, 이의 전문이 본원에 참조로 포함된다.

#### II.C.3.b. 알부민 또는 이의 단편, 또는 변이체

[0145] 특정의 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 단백질에 연결된 이종 잔기는 알부민 또는 이의 기능성 단편이다. 일부 구현예에서, VWF 단백질에 융합된 알부민은 FVIII 단백질에 융합된 알부민과 공유결합적으로 연합된다.

[0146] 이의 전장 형태에서 609개 아미노산의 단백질인, 사람 혈청 알부민(HSA, 또는 HA)은 혈청의 삼투압의 유의적인 비율에 관여하며 또한 내인성 및 외인성 리간드의 담체로서 기능한다. 본원에 사용된 것으로서 용어 "알부민"은 전장의 알부민 또는 이의 기능성 단편, 변이체, 유도체, 또는 유사체를 포함한다. 알부민 또는 이의 단편 또는 변이체의 예는 미국 특허 공보 제 2008/0194481A1호, 제2008/0004206 A1호, 제2008/0161243 A1호, 제2008/0261877 A1호, 또는 제2008/0153751 A1호 또는 PCT 출원 공보 제 2008/033413 A2호, 제2009/058322 A1호, 또는 제2007/021494 A2호에 개재되어 있으며, 이는, 이의 전문이 참조로 본원에 포함된다.

[0147]

### II.C.3.c. 알부민 결합 잔기

[0148]

특정의 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 단백질에 연결된 이종 잔기는 알부민 결합 잔기이며, 이는 알부민 결합 펩타이드, 세균 알부민 결합 도메인, 알부민-결합 항체 단편, 또는 이의 어떠한 조합도 포함한다. 예를 들면, 알부민 결합 단백질은 세균 알부민 결합 단백질, 도메인 항체를 포함하는 항체 또는 항체 단편일 수 있다(참조: 미국 특허 제6,696,245호). 예를 들어, 알부민 결합 단백질은 스트렙토코쿠스 단백질 G 중의 하나와 같은 세균 알부민 결합도메인일 수 있다[참조: Konig, T. and Skerra, A. (1998) *J. Immunol. Methods* 218, 73-83]. 접합 파트너로서 사용될 수 있는 알부민 결합 펩타이드의 다른 예는 예를 들면, Cys-Xaa<sub>1</sub>-Xaa<sub>2</sub>-Xaa<sub>3</sub>-Xaa<sub>4</sub>-Cys 컨센스 서열을 갖는 것들이며, 여기서 Xaa<sub>1</sub>는 Asp, Asn, Ser, Thr, 또는 Trp이고; Xaa<sub>2</sub>는 Asn, Gln이며, H는 Ile, Leu, 또는 Lys이고; Xaa<sub>3</sub>은 Ala, Asp, Phe, Trp, 또는 Tyr이며; Xaa<sub>4</sub>는 Asp, Gly, Leu, Phe, Ser, 또는 Thr이며, 이는 미국 특허원 제2003/0069395호 또는 덴니스(Dennis) 등의 문헌[참조: Dennis 등 (2002) *J. Biol. Chem.* 277, 35035-35043]에 기술된 바와 같다.

[0149]

### II.C.3.d. PAS 서열

[0150]

다른 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 단백질에 연결된 이종 잔기는 PAS 서열이다. 하나의 구현예에서, 키메라 분자는 VWF 링커를 통해 PAS 서열에 융합된 본원에 기술된 VWF 단백질을 포함한다. 다른 구현예에서, 본 발명의 키메라 분자는 VWF 링커를 통해 PAS 서열에 융합된 VWF 단백질을 포함하는 제1 쇄 및 FVIII 단백질과 추가의 임의의 PAS 서열을 포함하는 제2의 쇄를 포함하며, 여기서 PAS 서열은 FVIII 단백질에서 VWF 결합 부위를 차폐하거나 보호함으로써, FVIII 단백질과 내인성 VWF의 상호작용을 억제하거나 방지한다. 2개의 PAS 서열은 서로 공유결합으로 연합될 수 있다.

[0151]

본원에 사용된 것으로서 PAS 서열은 주로 알라닌 및 세린 잔기를 포함하거나 주로 알라닌, 세린, 및 프롤린 잔기를 포함하는 아미노산 서열을 의미하고, 당해 아미노산 서열은 생리학적 조건하에서 무작위적인 코일 구조를 형성한다. 따라서, PAS 서열은 빌딩 블록, 아미노산 중합체, 또는 키메라 단백질내 이종 잔기의 일부로서 사용될 수 있는 알라닌, 세린, 및 프롤린으로 필수적으로 이루어지거나 이루어진 서열 카세트이다. 여전히, 숙련가들은, 알라닌, 세린, 및 프롤린 이외의 잔기가 PAS 서열내 약간의 성분으로서 가해지는 경우 무작위적인 코일 구조를 형성할 수 있음을 인지하고 있다. 본원에 사용된 것으로서 용어 "약간의 성분"은, 알라닌, 세린, 및 프롤린 이외의 아미노산이 PAS 서열에서 특정의 정도로, 예를 들면, 약 12% 이하, 즉, PAS 서열의 100개 아미노산 중 약 12개의 아미노산, 약 10% 이하, 즉, PAS 서열의 100개 아미노산 중 약 10개의 아미노산, 약 9% 이하, 즉, 100개 아미노산 중 약 9개의 아미노산, 약 8% 이하, 즉, 100개 아미노산 중 약 8개의 아미노산, 약 6%, 즉, 100개 아미노산 중 약 6개의 아미노산, 약 5%, 즉, 100개 아미노산 중 약 5개의 아미노산, 약 4%, 즉, 100개의 아미노산 중 약 5개의 아미노산, 약 3%, 즉, 100개의 아미노산 중 약 3개의 아미노산, 약 2%, 즉, 100개 아미노산 중 약 2개의 아미노산, 약 1%, 즉, 100개의 아미노산 중 약 1개의 아미노산으로 첨가될 수 있음을 의미한다. 알라닌, 세린 및 프롤린과는 상이한 아미노산은 Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Tyr, 및 Val로 이루어진 그룹으로부터 선택될 수 있다.

[0152]

생리학적 조건하에서, PAS 서열 길이는 무작위적인 코일 구조를 형성함으로써 VWF 인자 또는 응고 활성의 단백질에 대해 생체내 및/또는 시험관내 안정성을 중재할 수 있다. 무작위적인 코일 도메인은 자체로서 안정한 구조 또는 기능을 채택하지 않으므로, 이것이 융합된 VWF 단백질 또는 FVIII 단백질에 의해 중재된 생물학적 활성은 필수적으로 보존된다. 다른 구현예에서, 무작위적인 코일 도메인을 형성하는 PAS 서열은 특히 혈액 혈장, 면역 원성, 등전점/정전 행위, 세포 표면 수용체 또는 내재화와 관련하여 생물학적으로 불활성이지만, 여전히 생분해성이며, 이는 PEG와 같은 합성 중합체보다 명확한 장점을 제공한다.

[0153]

무작위적인 코일 구조를 형성하는 PAS 서열의 비-제한적 예는 ASPAAPAPASPAAPAPSAPA(서열 번호 32), AAPASPAPAAAPSAPAPAAPS(서열 번호 33), APSSPSPSAPSSPSPASPSS(서열 번호 34), APSSPSPSAPSSPSPASPS(서열 번호 35), SSPSAPSPSSPASPSSPA(서열 번호 36), AASPAAPSAPPAAASPAAPSAPPA(서열 번호 37) 및 ASAAAPAAASAAASAPSAAA(서열 번호 38) 또는 이의 어떠한 조합으로 이루어진 그룹으로부터 선택된 아미노산 서열을 포함한다. PAS 서열의 추가의 예는 예를 들면, 미국 특허 공보 제2010/0292130 A1호 및 PCT 출원 공보 제WO 2008/155134 A1호에 공지되어 있다.

[0154]

### II.C.3.e. HAP 서열

[0155]

특정의 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 단백질에 연결된 이종

잔기는 글리신이 풍부한 아미노산 중합체(HAP)이다. HAP 서열은 글리신의 반복된 서열을 포함할 수 있으며, 이는, 길이가 적어도 50개 아미노산, 적어도 100개 아미노산, 120개 아미노산, 140개 아미노산, 160개 아미노산, 180개 아미노산, 200개 아미노산, 250개 아미노산, 300개 아미노산, 350개 아미노산, 400개 아미노산, 450개 아미노산, 또는 500개 아미노산이다. 하나의 구현예에서, HAP 서열은 HAP 서열에 융합되거나 연결된 잔기의 반감기를 연장시킬 수 있다. HAP 서열의 비-제한적인 예는  $(\text{Gly})_n$ ,  $(\text{Gly}_4\text{Ser})_n$  또는  $S(\text{Gly}_4\text{Ser})_n$ 을 포함하나, 이에 한정되지 않으며, 여기서,  $n$ 은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20이다. 하나의 구현예에서,  $n$ 은 20, 21, 22, 23, 24, 25, 26, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 또는 40이다. 다른 구현예에서,  $n$ 은 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200이다. [참조: 예를 들면, Schlapschy M 등, Protein Eng. Design Selection, 20: 273-284 (2007)].

#### [0156] II.C.3.f. 트랜스페린 또는 이의 단편

특정의 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 단백질에 연결된 이종 잔기는 트랜스페린 또는 이의 단편이다. 어떠한 트랜스페린도 사용하여 본 발명의 키메라 분자를 제조할 수 있다. 예로서, 야생형 사람 Tf(Tf)는 대략 75Kda(글리코실화에 대해 고려되지 않음)이고, 2개이 주요 도메인, N(약 330개 아미노산) 및 C(약 340개 아미노산)을 지닌 679개 아미노산 단백질이며, 이는 유전자 중복으로부터 기원하는 것으로 여겨진다. [참조: 진뱅크 수탁 번호 제NM001063호, 제XM002793호, 제M12530호, 제XM039845호, 제XM 039847호 및 제S95936호([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)), 이들 모두는, 이의 전문이 본원에 참조로 포함된다]. 트랜스페린은 2개의 도메인, N 도메인 및 C 도메인을 포함한다. N 도메인은 2개의 소도메인, N1 도메인 및 N2 도메인을 포함하며, C 도메인은 2개의 소 도메인, C1 도메인 및 C2 도메인을 포함한다.

하나의 구현예에서, 키메라 분자의 트랜스페린 부위는 트랜스페린 스플라이스 변이체를 포함한다. 하나의 예에서, 트랜스페린 스플라이스 변이체는 사람 트랜스페린의 스플라이스 변이체, 예를 들면, 진뱅크 수탁 번호 제AAA61140호일 수 있다. 다른 구현예에서, 키메라 분자의 트랜스페린 부위는 트랜스 페린 서열의 하나 이상의 도메인, 예를 들면, N 도메인, C 도메인, N1 도메인, N2 도메인, C1 도메인, C2 도메인 또는 이의 어떠한 조합을 포함한다.

#### [0159] II.C.3.g. 중합체, 예를 들면, 폴리에틸렌 글리콜(PEG)

다른 구현예에서, VWF 링커를 통해 VWF 단백질에 부착되거나 FVIII 링커를 통해 FVIII 단백질에 부착된 이종 잔기는 당해 분야에 공지된 가용성 중합체이며, 폴리에틸렌 글리콜, 에틸렌 글리콜/프로필렌 글리콜 공중합체, 카복시메틸셀룰로즈, 텍스트란, 또는 폴리비닐 알코올을 포함하나, 이에 한정되지 않는다. 가용성 중합체와 같은 이종 잔기는 키메라 분자내에의 어떠한 위치에도 부착될 수 있다.

특정의 구현예에서, 키메라 분자는 VWF 잔기를 통해 이종 잔기(예를 들면, Fc 영역)에 융합된 VWF 단백질이며, 여기서 VWF 단백질은 PEG에 추가로 연결된다. 다른 구현예에서, 키메라 분자는 VWF 링커를 통해 Fc 영역에 융합된 VWF 단백질 및 FVIII 단백질을 포함하며, 이는 서로 연합되어 있고, 여기서 FVIII 단백질은 PEG에 연결되어 있다.

또한, 폴리펩타이드의 증가된 가용성, 안전성 및 순환 시간, 또는 감소된 면역원성과 같이 추가의 장점을 제공할 수 있는 본 발명의 키메라 분자의 화학적으로 변형된 유도체가 본 발명에 의해 제공된다(참조: 미국 특허 제4,179,337호). 변형을 위한 화학적 잔기는 폴리에틸렌 글리콜, 에틸렌 글리콜/프로필렌 글리콜 공중합체, 카복시메틸셀룰로즈, 텍스트란, 또는 폴리비닐 알코올을 포함하나, 이에 한정되지 않는 수용성 중합체로부터 선택될 수 있다. 키메라 분자는 분자내 무작위한 위치에서 또는 N- 또는 C-말단에서, 또는 분자내 예정된 위치에서 변형될 수 있으며 1개, 2개, 3개 이상의 부착된 화학적 잔기를 포함할 수 있다.

중합체는 어떠한 분자량일 수 있으며, 측쇄되거나 측쇄되지 않을 수 있다. 하나의 구현예에서, 폴리에틸렌 글리콜의 경우, 분자량은 용이한 취급 및 제조를 위해 약 1kDa 내지 약 100kDa이다. 다른 크기가 목적한 프로파일(예를 들면, 요구되는 지속된 방출 기간, 생물학적 활성에 있어서 어떠한 것의 효과, 취급의 용이성, 항원성의 정도 또는 결여 및 단백질 또는 동족체에 대한 폴리에틸렌 글리콜의 다른 공지된 효과)에 따라 사용될 수 있다. 예를 들어, 폴리에틸렌 글리콜은, 평균 분자량이 약 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10,000, 10,500, 11,000, 11,500, 12,000, 12,500, 13,000, 13,500, 14,000, 14,500, 15,000, 15,500, 16,000, 16,500, 17,000, 17,500, 18,000, 18,500, 19,000, 19,500, 20,000, 25,000, 30,000, 35,000, 40,000, 45,000, 50,000, 55,000,

60,000, 65,000, 70,000, 75,000, 80,000, 85,000, 90,000, 95,000, 또는 100,000kDa일 수 있다.

[0164] 일부 구현예에서, 폴리에틸렌 글리콜은 측쇄된 구조를 가질 수 있다. 측쇄된 폴리에틸렌 글리콜은 예를 들면, 미국 특허 제5,643,575호; Morpurgo 등, *Appl. Biochem. Biotechnol.* 56:59-72 (1996); Vorobjev 등, *Nucleosides Nucleotides* 18:2745-2750 (1999); and Caliceti 등, *Bioconjug. Chem.* 10:638-646 (1999)에 기술되어 있으며, 이들 각각은, 이의 전문이 본원에 참조로 포함되어 있다.

[0165] 각각의 키메라 분자에 부착된 폴리에틸렌 글리콜 잔기의 수(즉, 치환 정도)는 또한 변할 수 있다. 예를 들어, 본 발명의 폐길화된 단백질은 평균 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20개 이상의 폴리에틸렌 글리콜 분자에 연결될 수 있다. 유사하게, 단백질 분자당 1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, 8-10, 9-11, 10-12, 11-13, 12-14, 13-15, 14-16, 15-17, 16-18, 17-19, 또는 18-20개의 폴리에틸렌 글리콜 잔기와 같은 범위내의 평균 치환도. 치환도를 측정하는 방법은, 예를 들면, 문헌[참조: Delgado 등, *Crit. Rev. Thera. Drug Carrier Sys.* 9:249-304 (1992)]에 논의되어 있다.

[0166] 다른 구현예에서, 본 발명에 사용된 FVIII 단백질은 하나 이상의 중합체에 접합된다. 중합체는 수용성일 수 있고 인자 VIII에 또는 인자 VIII에 접합된 다른 잔기에 공유결합적으로 또는 비-공유결합적으로 부착될 수 있다. 중합체의 비-제한적 예는 폴리(알킬렌 옥사이드), 폴리(비닐 피롤리돈), 폴리(비닐 알코올), 폴리옥사졸린, 또는 폴리(아크로일모르폴린)일 수 있다. 추가 유형의 중합체-접합된 FVIII은 미국 특허 제7,199,223호에 논의되어 있다.

#### II.C.3.h. 하이드록시에틸 전분(HES)

[0168] 특정의 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 단백질에 연결된 이종 잔기는 중합체, 예를 들면, 하이드록시에틸 전분(HES) 또는 이의 유도체이다.

[0169] 하이드록시에틸 전분(HES)은 천연적으로 존재하는 아밀로펙틴이고 체내에서 알파-아밀라제에 의해 분해된다. HES는 탄수화물 중합체 아밀로펙틴의 치환된 유도체이며, 이는 옥수수 전분 속에 95중량% 이하의 농도로 존재한다. HES는 유리한 생물학적 특성을 나타내며 임상에서 혈액 용적 대체제로서 및 혈액 회석 치료요법에서 사용된다[참조: Sommermeyer 등, *Krankenhauspharmazie*, 8(8), 271-278 (1987); and Weidler 등, *Arzneim.-Forschung/Drug Res.*, 41, 494-498 (1991)].

[0170] 아밀로펙틴은 글루코즈 잔기를 함유하며, 여기서 주체 알파-1,4-글리코시드 결합이 존재하며 측쇄부위에서 알파-1,6-글리코시드 결합이 발견된다. 당해 분자의 물리-화학적 특성은 글리코시드 결합의 유형에 의해 주로 측정된다. 닉(nick)이 있는 알파-1,4-글리코시드 결합으로 인하여, 회전 당 약 6개의 글루코즈-단량체를 지닌 나선 구조가 생산된다. 중합체의 물리-화학적 및 생화학적 특성은 치환을 통해 변형될 수 있다. 하이드록시에틸 그룹의 도입은 알칼리성 하이드록시에틸화를 통해 달성될 수 있다. 반응 조건을 적합하게 함으로써, 하이드록시에틸화에 대하여 비치환 글루코즈 단량체에서 각각의 하이드록실 그룹의 상이한 반응성을 이용할 수 있다. 이러한 사실로 인하여, 숙련가는 치환 양식에 제한된 정도로 영향을 미칠 수 있다.

[0171] HES는 분자량 분포 및 치환도에 의해 주로 특징화된다. DS로서 나타낸 치환도는 물 치환에 관한 것이며, 기술자에게 공지되어 있다. [참조: 상기 인용된 바와 같은 Sommermeyer 등, *Krankenhauspharmazie*, 8(8), 271-278 (1987), 특히 p. 273].

[0172] 하나의 구현예에서, 하이드록시에틸 전분은, 평균 분자량(중량 평균)이 1 내지 300kD, 2 내지 200kD, 3 내지 100kD, 또는 4 내지 70kD이다. 하이드록시에틸 전분은, 0.1 내지 3, 바람직하게는 0.1 내지 2, 보다 바람직하게는, 0.1 내지 0.9, 바람직하게는 0.1 내지 0.8의 몰 치환도, 및 하이드록시에틸 그룹과 관련하여 2 내지 20의 범위의 C2:C6 치환 사이의 비를 추가로 나타낼 수 있다. 평균 분자량이 약 130kD인 HES의 비-제한적 예는, 평균 치환도가 0.2 내지 0.8, 예를 들면, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 또는 0.8, 바람직하게는 0.4 내지 0.7, 예를 들면, 0.4, 0.5, 0.6, 또는 0.7인 HES이다. 구체적인 구현예에서, 평균 분자량이 약 130kD인 HES는 업자(Fresenius)로부터의 Voluven<sup>®</sup>이다. Voluven<sup>®</sup>은 인공 콜로이드이며, 예를 들면, 저혈량증의 치료요법 및 예방을 위한 치료학적 처방에서 사용된 용적 대체를 위해 사용된다. Voluven<sup>®</sup>의 특징은, 평균 분자량이 130,000+/-20,000D이고, 몰 치환도가 0.4이며 C2:C6 비가 약 9:1이다. 다른 구현예에서, 하이드록시에틸 전분의 평균 분자량의 범위는, 예를 들면, 4 내지 70kD 또는 10 내지 70kD 또는 12 내지 70kD 또는 18 내지 70kD 또는 50 내지 70kD 또는 4 내지 50kD 또는 10 내지 50kD 또는 12 내지 50kD 또는 18 내지 50kD 또는 4 내지 18kD 또는 10 내지 18kD 또는 12 내지 18kD 또는 4 내지 12kD 또는 10 내지 12kD 또는 4 내지 10kD이다. 여전히 다른 구현예에

서, 사용된 하이드록시에틸 전분의 평균 분자량은 4kD 이상 및 70kD 이하, 예를 들면, 약 10kD, 또는 9 내지 10kD 또는 10 내지 11kD 또는 9 내지 11kD의 범위, 또는 약 12kD, 또는 11 내지 12kD의 범위) 또는 12 내지 13kD 또는 11 내지 13kD, 또는 약 18kD, 또는 17 내지 18kD 또는 18 내지 19kD 또는 17 내지 19kD의 범위, 또는 약 30kD, 또는 29 내지 30, 30 내지 31kD, 또는 약 50kD의 범위, 또는 49 내지 50kD 또는 50 내지 51kD 또는 49 내지 51kD의 범위이다.

[0173] 특정의 구현예에서, 이종 잔기는, 평균 분자량이 상이하고/하거나 치환도가 상이하고/하거나 C2: C6 치환비가 상이한 하이드록시에틸 전분의 혼합물일 수 있다. 따라서, 평균 분자량이 상이하고 치환도가 상이하며 C2:C6 치환 비가 상이하거나, 평균 분자량이 상이하고 치환도가 상이하며 C2:C6 치환 비가 동일하거나 거의 동일하거나, 평균 분자량의 동일하거나 거의 동일하고 치환도가 상이하며 C2:C6 치환비가 상이하거나, 평균 분자량이 상이하고 치환도가 동일하거나 거의 동일하며 C2:C6 치환비가 동일하거나 거의 동일하거나, 평균 분자량이 동일하거나 거의 동일하고 치환비가 동일하거나 거의 동일하며 C2:C6 치환비가 상이하거나, 평균 분자량이 거의 동일하고 치환도가 거의 동일하며 C2:C6 치환비가 거의 동일한 하이드록시에틸 전분이 혼합물이 사용될 수 있다.

#### II.C.3.i. 폴리시알산(PSA)

[0175] 특정의 구현예에서, VWF 링커를 통해 VWF 단백질에 연결되거나 FVIII 링커를 통해 FVIII 링커에 연결된 된 비-폴리펩타이드 이종 잔기는 종합체, 예를 들면, 폴리시알산(PSA) 또는 이의 유도체이다. 폴리시알산(PSA)은 특정의 세균 균주에 의해 및 포유동물내에서 특정의 세포내에서 생산된 시알산의 천연적으로 존재하는 비축쇄된 중합체이다. [참조: Roth J., 등 (1993) in *Polysialic Acid: From Microbes to Man*, eds. Roth J., Rutishauser U., Troy F. A. (Birkhäuser Verlag, Basel, Switzerland), pp 335-348]. 이들은 n=약 80 이상의 시알산 잔기로부터 n=2까지의 다양한 중합화도에서 제한된 산 가수분해에 의해 또는 뉴라미니다제를 사용한 분해에 의해, 또는 중합체의 천연의, 세균적으로 기원한 형태의 분획화에 의해 생산될 수 있다. 상이한 폴리시알산의 조성은 변하여 단독중합체성 형태 즉 이. 콜라이(*E. coli*) 균주 K1 및 신경 세포 부착 분자 (N-CAM)의 배아 형태로 또한 발견된 그룹-B 뇌척수막염 균의 캡슐형 다당류를 포함하는 알파-2,8-연결된 폴리시알산이 존재하도록 한다. 혜테로중합체성 형태는 또한 이. 콜라이 균주 K92의 교호하는 알파-2,8 알파-2,9 폴리시알산 및 엔. 메닌기티디스(*N. meningitidis*)의 그룹 C 다당류와 같이 존재한다. 시알산은 또한 엔. 메닝기티디스(*N. meningitidis*)의 그룹 W135 또는 그룹 Y와 같은 시알산 이외의 단량체와의 교호하는 공중합체에서 발견될 수 있다. 포유동물에서 폴리시알산에 대한 수용체가 알려져 있지 않지만, 폴리시알산은 병리학적 세균에 의한 면역 및 보체 시스템의 회피 및 태아 발달 동안 미성숙 뉴우런의 신경교 접착성의 조절을 포함하는 중요한 생물학적 기능을 갖는다(여기서, 중합체는 항-부착성 기능을 갖는다)[참조: Cho and Troy, P.N.A.S., USA, 91 (1994) 11427-11431]. 이. 콜라이 균주 K1의 알파-2,8-연결된 폴리시알산은 또한 '콜로민산'으로 알려져 있으며 본 발명을 예시하는데 사용된다(다양한 길이로). 폴리시알산을 폴리펩타이드에 접착시키거나 접합시키는 다양한 방법은 기술되어 있다(예를 들면, 이의 전문이 본원에 참조로 포함된, 미국 특허 제5,846,951호; 제WO-A-0187922호, 및 제US 2007/0191597 A1호 참조).

#### II.C.4. XTEN 서열

[0177] 본원에 사용된 것으로서 "XTEN 서열"은 주로 소 친수성 아미노산으로 구성된, 비-천연적으로 존재하는, 실질적으로 비-반복성 서열을 지닌 연장된 길이의 폴리펩타이드를 말하며, 이의 서열은 생리학적 조건하에서 낮은 정도 또는 2차 또는 3차 구조를 갖는다. 키메라 단백질 파트너, XTEN은 담체로서 제공될 수 있으므로, 본 발명의 VWF 단백질 또는 FVIII 단백질에 연결되어 키메라 단백질을 생성하는 경우, 특정의 바람직한 약력학적, 물리화학적 및 약제학적 특성을 부여한다. 이러한 바람직한 특성은 향상된 약력학적 매개변수 및 용해도 특성을 포함하나, 이에 한정되지 않는다. 본원에 사용된 것으로서, "XTEN"은 일본쇄 항체 또는 경쇄 또는 중쇄의 Fc 단편과 같은 항체 또는 항체 단편을 구체적으로 배제한다.

[0178] 일부 구현예에서, 본 발명의 XTEN 서열은 약 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, 또는 2000개 이상의 아미노산 잔기를 갖는 웹타이드 또는 폴리펩타이드이다. 특정의 구현예에서, XTEN은 약 20 내지 약 3000 개 이상의 아미노산 잔기, 30 내지 약 2500개 이상의 아미노산 잔기, 40 내지 약 2000개 이상의 아미노산 잔기, 50 내지 약 1500개 이상의 아미노산 잔기, 60 내지 약 1000개 이상의 아미노산 잔기, 70 내지 약 900개 이상의 아미노산 잔기, 80 내지 약 800개 이상의 아미노산 잔기, 90 내지 약 700개 이상의 아미노산 잔기, 100 내지 약 600개 이상의 아미노산 잔기, 110 내지 약 500개 이상의 아미노산 잔기, 120 내지 약 400개 이상의 아미노산 잔

기를 지닌 웹타이드 또는 폴리웹타이드이다.

[0179] 본 발명의 XTEN 서열은 9 내지 14개 아미노산 잔기 또는 서열 모티프에 대해 적어도 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 동일한 아미노산 서열의 하나 이상의 서열 모티프를 포함할 수 있으며, 여기서 당해 모티프는 글리신(G), 알라닌(A), 세린(S), 트레오닌(T), 글루타메이트(E) 및 프롤린(P)으로 이루어진 그룹으로부터 선택된 4 내지 6개 유형의 아미노산을 포함하거나, 이로 필수적으로 이루어지거나, 이로 이루어진다. (참조: US 2010-0239554 A1).

[0180] 일부 구현예에서, XTEN은 비-오버랩핑된 서열 모티프를 포함하여 계열 서열을 생성하며 여기서 서열의 약 80%, 또는 적어도 약 85%, 또는 적어도 약 90%, 또는 약 91%, 또는 약 92%, 또는 약 93%, 또는 약 94%, 또는 약 95%, 또는 약 96%, 또는 약 97%, 또는 약 98%, 또는 약 99% 또는 약 100%가 표 2A로부터 선택된 단일의 모티프 계열로부터 선택된 비-오버랩핑 서열의 다중 단위로 이루어진다. 본원에 사용된 것으로서, "계열"은, XTEN이 표 2A로부터의 단일의 모티프 카테고리로부터 단지선택된 모티프; 즉, AD, AE, AF, AG, AM, AQ, BC, 또는 BD XTEN을 가지며, 계열 모티프로부터가 아닌 XTEN내 어떠한 다른 아미노산도 암호화된 뉴클레오타이드에 의한 제한 부위의 혼입, 절단 서열의 혼입을 허용하는 것과 같은, 요구되는 특성을 달성하기 위해, 또는 FVIII 또는 VWF에 대해 보다 우수한 연결을 달성하기 위해 선택된다. XTEN 계열의 일부 구현예에서, XTEN 서열은 AD 모티프 계열, 또는 AE 모티프 계열, 또는 AF 모티프 계열, 또는 AG 모티프 계열, 또는 AM 모티프 계열, 또는 AQ 모티프 계열, 또는 BC 모티프 계열, 또는 BD 모티프 계열의 비-오버래핑 서열 모티프의 다중 단위를 포함하며, 수득되는 XTEN은 위에서 기술한 상동성 범위를 나타낸다. 다른 구현예에서, XTEN은 표 2A의 2개 이상의 모티프 계열로부터의 모티프 서열의 다중 단위를 포함한다. 이들 서열은 하기에 보다 완전하게 기술된, 총 전하, 친수성, 2차 구조의 결여, 또는 모티프의 아미노산 조성에 의해 부여된 반복성의 결여와 같은 특성을 포함하는 바람직한 물리적/화학적 특징을 달성하기 위해 선택될 수 있다. 당해 단락에 기술된 본원의 위의 구현예에서, XTEN내로 포함된 모티프는 본원에 기술된 방법을 사용하여 선택되고 조립됨으로써 아미노산 잔기가 약 36 내지 약 3000개인 XTEN을 달성할 수 있다.

[0181]

표 2A. 12개 아미노산의 XTEN 서열 모티프 및 모티프 계열

모티프 계열	모티프 서열
AD	GESPGGSSGSES (서열 번호 49)
AD	GSEGSSGPGESS (서열 번호 50)
AD	GSSESGSSEGGP (서열 번호 51)
AD	GSGGEPSSEGSS (서열 번호 52)
AE, AM	GSPAGSPTSTEE (서열 번호 53)
AE, AM, AQ	GSEPATSGSETP (서열 번호 54)
AE, AM, AQ	GTSESATPESGP (서열 번호 55)
AE, AM, AQ	GTSTEPSEGSAP (서열 번호 56)
AF, AM	GSTSESPSGTAP (서열 번호 57)
AF, AM	GTSTPESGSASP (서열 번호 58)
AF, AM	GTSPSGESSTAP (SEQ ID NO: 59)
AF, AM	GSTSSTAESP GP (서열 번호 60)
AG, AM	GTPGSGTASSSP (서열 번호 61)
AG, AM	GSSTPSGATGSP (서열 번호 62)
AG, AM	GSSPSASTGTGP (SEQ ID NO: 63)
AG, AM	GASPGTSSTGSP (SEQ ID NO: 64)
AQ	GEPAGSPTSTSE (서열 번호 65)
AQ	GTGEPSSTPASE (서열 번호 66)
AQ	GSGPSTESAPTE (서열 번호 67)
AQ	GSETPSGPSETA (서열 번호 68)
AQ	GPSETSTSEPGA (서열 번호 69)
AQ	GSPSEPTEGTSA (서열 번호 70)
BC	GSGASEPTSTEP (서열 번호 71)
BC	GSEPATSGTEPS (서열 번호 72)
BC	GTSEPSTSEPGA (서열 번호 73)
BC	GTSTEPSEPGSA (서열 번호 74)
BD	GSTAGSETSTEA (서열 번호 75)
BD	GSETATSGSETA (서열 번호 76)
BD	GTSESATSESGA (서열 번호 77)
BD	GTSTEASEGSAS (서열 번호 78)

- 다양한 순열로 함께 사용되는 경우, 개개의 모티프 서열이

“계열 서열”을 생성하는 것에 주목한다.

[0182]

[0183]

XTEN은 FVIII 또는 VWF 또는 키메라 분자의 어떠한 다른 성분내로의 삽입 또는 이에 대한 연결을 위한 다양한 길이를 가질 수 있다. 하나의 구현예에서, XTEN 서열(들)의 길이는 융합 단백질내에서 달성될 특성 또는 기능을 기본으로 하여 선택된다. 의도된 특성 또는 기능에 따라서, XTEN은 담체로 제공될 수 있는 중간 길이의 서열 또는 보다 긴 서열일 수 있다. 특정의 구현예에서, XTEN은 약 6 내지 약 99개의 아미노산 잔기의 짧은 분절, 약 100 내지 약 399개의 아미노산 잔기의 중간 길이, 및 약 400 내지 약 1000개 및 약 3000개 이하의 아미노산 잔기의 보다 긴 길이를 포함한다. 따라서, FVIII 또는 VWF내로 삽입되거나 연결된 XTEN은, 길이가 약 6, 약 12, 약 36, 약 40, 약 42, 약 72, 약 96, 약 144, 약 288, 약 400, 약 500, 약 576, 약 600, 약 700, 약 800, 약 864, 약 900, 약 1000, 약 1500, 약 2000, 약 2500, 또는 약 3000개 이하인 아미노산 잔기 길이를 가질 수 있다. 다른 구현예에서, XTEN 서열은, 길이가 약 6 내지 약 50, 약 50 내지 약 100, 약 100 내지 150, 약 150 내지 250, 약 250 내지 400, 약 400 내지 약 500, 약 500 내지 약 900, 약 900 내지 1500, 약 1500 내지 2000, 또는 약 2000 내지 약 3000개인 아미노산 잔기이다. FVIII 또는 VWF내로 삽입된 XTEN의 정밀한 길이는 FVIII 또는 VWF의 활성에 역으로 영향을 미치지 않고 변할 수 있다. 하나의 구현예에서, 본원에 사용된 하나 이상의 XTEN은 36개의 아미노산, 42개의 아미노산, 72개의 아미노산, 144개의 아미노산, 288개의 아미노산, 576개의 아미노산, 또는 864개의 아미노산 길이를 가지며 XTEN 계열의 서열; 즉, AD, AE, AF, AG, AM, AQ, BC 또는 BD 중 하나 이상으로부터 선택될 수 있다.

[0184]

일부 구현예에서, 본 발명에 사용된 XTEN 서열은 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, 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, 및 AG2004로 이루어진 그룹으로부터 선택된 서열에 대해 적어도 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하다. (참조: US 2010-0239554 A1).

[0185] 하나의 구현예에서, XTEN 서열은 AE42, AE864, AE576, AE288, AE144, AG864, AG576, AG288, AG144, 및 이의 어떠한 조합으로 이루어진 그룹으로부터 선택된 아미노산 서열에 대해 적어도 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하다. 다른 구현예에서, XTEN 서열은 AE42, AE864, AE576, AE288, AE144, AG864, AG576, AG288, AG144, 및 이의 어떠한 조합으로 이루어진 그룹으로부터 선택된다. 구체적인 구현예에서, XTEN 서열은 AE288이다. 본 발명의 특정의 XTEN 서열에 대한 아미노산 서열은 표 2b에 나타나 있다.

#### 표 2B. XTEN 서열

XTEN	아미노산 서열
AE42 서열 번호 39	GAPGSPAGSPTSTEETSESATPESGPSEATSGSETPASS
AE72 서열 번호 40	GAP TSESATPESG PGSEPATSGS ETPGTSESAT PESGPGPSEPA TSGSETPGTS ESATPESGP GP TSTEPSEGSA PGASS
AE144 서열 번호 41	GSEPATSGSETPGTSESATPESGPGPSEPATSGSETPGSPAGSPTSTEETSTEPSEG SAPGPSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESA PESGPGPSEPATSGSETPGTSTEPSEGSAP
AG144 서열 번호 42	GTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSST GSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSA STGTGPGTPGSGTASSSPGSSTPSGATGSP
AE288 서열 번호 43	GTSESATPESGPGPSEPATSGSETPGTSESATPESGPGPSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGPAGSPTSTEETSESATPESGPGPSEPATSGSETPGTSESATPES GPGPAGSPTSTEETGPAGSPTSTEETSTEPSEGSAPGTSESATPESGPGPSESATPE SGPCTSESATPESGPGPSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEETSTEPSE GSAPGTSTEPSEGSAPGPSEPATSGSETPGTSESATPESGPGPSTEPSEGSAP
AG288 서열 번호 44	PGASPGTSSTGSPGASP GTGSPGTSSTGSPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSATGSPGSSPSASTG TGPGSSPSASTGTGPGASPGTSS TGSPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSS TGSPGASPGTSS TGSPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSS TGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSP
AE576 서열 번호 45	GSPAGSPTSTEETSESATPESGPGTSTEPSEGSAPGPAGSPTSTEETSTEPSEGSA PGTSTEPSEGSAPGTSESATPESGPGPSEPATSGSETPGSEPATSGSETPGSPAGSPTST EEGTSESATPESGPGPSTEPSEGSAPGTSTEPSEGSAPGPAGSPTSTEETSTEPSEG SAPGTSTEPSEGSAPGTSESATPESGPGPSTEPSEGSAPGTSESATPESGPGPSEPATSG SETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGPSEATPESGPGPAGSP TSTEETSESATPESGPGPSEPATSGSETPGTSESATPESGPGPSTEPSEGSAPGTSTEP SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGPAG SPTSTEETSTEPSEGSAPGTSESATPESGPGPSEPATSGSETPGTSESATPESGPGPSE ATSGSETPGTSESATPESGPGPSTEPSEGSAPGTSESATPESGPGPAGSPTSTEETGP

[0187]

사용된 XTEN 성분이 글리신(G), 알라닌(A), 세린(S), 트레오닌(T), 글루타메이트(E) 및 프롤린(P)으로부터 선택된 4, 5 또는 6개 유형의 아미노산으로 이루어진 이의 아미노산의 100% 미만이거나, 표2A로부터의 서열 모티프 또는 표 2B의 XTEN 서열로 이루어진 서열의 100% 미만인 이를 구현예에서, XTEN의 다른 아미노산 잔기는 어떠한 다른 14개의 천연의 L-아미노산으로부터 선택되지만 바람직하게는 친수성 아미노산으로부터 선택됨으로써 XTEN 서열은 적어도 약 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 적어도 약 99%의 친수성 아미노산을 함유한다. 글리신(G), 알라닌(A), 세린(S), 트레오닌(T), 글루타메이트(E) 및 프롤린(P)가 아닌 XTEN 아미노산은 XTEN 서열 전체에 산재되거나, 서열 모티프 내 또는 사이에 위치하거나 하나 이상의 짧은 신장의 XTEN 서열내에 놓축되어 예를 들면, XTEN과 다른 성분, 예를 들면, VWF 단백질 사이에 링커를 생성한다. XTEN 성분이 글리신(G), 알라닌(A), 세린(S), 트레오닌(T), 글루타메이트(E) 및 프롤린 이외의 아미노산을 포함하는 경우에, 본원에 개재된 방법에 의해 측정된 것으로서, 약 2% 미만 또는 약 1% 미만의 아미노산이 소수성 잔기이어서 생성되는 서열이 일반적으로 2차 구조를 결실하는, 예를 들면, 2% 이상의 알파 나선 또는 2%의 베타 쉬이트를 가지지 않도록 하는 것이 바람직하다. XTEN의 구조내에서 거의 선호되지 않는 소수성 잔기는 트립토판, 페닐알라닌, 타이로신, 루이신, 이소루이신, 발린, 및 메티오닌을 포함한다. 또한, XTEN 서열을 설계하여 다음의 아미노산을 5% 미만, 4% 미만 또는 3% 미만 또는 2% 미만으로 함유하도록 할 수 있다: 시스테인(이황화물 형성 및 산화를 방지하기 위하여), 메티오닌(산화를 방지하기 위하여), 아스파라기 및 글루타민(데스아미드화를 방지하기 위하여).

여). 따라서, 일부 구현예에서, 글리신(G), 알라닌(A), 세린(S), 트레오닌(T), 글루타메이트(E) 및 프롤린(P) 이외에 다른 아미노산을 포함하는 XTEN은 초우-파스만 알고리즘에 의해 측정된 것으로서 알파-나선 및 베타-쉬이트에 기여하는 5% 미만의 잔기를 갖는 서열을 가지며 GOR 알고리즘에 의해 측정된 것으로서 적어도 90%, 또는 적어도 약 95% 이상의 무작위적인 코일 형성을 갖는다.

[0190] 추가의 구현예에서, 본 발명에 사용된 XTEN 서열은 물리 또는 화학적 특성, 예를 들면, 본 발명의 키메라 단백질의 약력학에 영향을 미친다. 본 발명에 사용된 XTEN 서열은 다음의 유리한 특성들 중 하나 이상을 나타낼 수 있다: 구조적글곡성, 향상된 수용성, 고도의 프로테아제 내성, 저 면역원성, 포유동물 수용체에 대한 낮은 결합, 또는 증가된 유체역학적(또는 스트로크(Stoke)) 반경. 구체적인 구현예에서, 본 발명에서 FVIII 단백질에 연결된 XTEN 서열은 보다 긴 말단 반감기 또는 증가된 곡선하 영역(AUC)과 같은 약력학적 특성을 증가시킴으로써, 본원에 기술된 키메라 단백질은 야생형 FVIII과 비교하여 증가된 기간 동안 생체내에서 머무른다. 추가의 구현예에서, 본 발명에 사용된 XTEN 서열은 보다 긴 말단 반감기 또는 증가된 곡선하 영역(AUC)과 같은 약력학적 특성을 증가시킴으로써, FVIII 단백질은 야생형 FVIII과 비교하여 증가된 기간 동안 생체내에서 머무른다.

[0191] 다양한 방법 및 검정을 사용하여 XTEN 서열을 포함하는 단백질의 물리적/화학적 특성을 측정할수 있다. 이러한 방법은 분석적 원심분리, EPR, HPLC-이온 교환, HPLC-크기 배제, HPLC-역상, 광 산란 (light scattering), 모세관 전기영동, 원편광 이색성, 차등 주사 열량계, 형광성, HPLC-이온 교환, HPLC-크기 배출, IR, NMR, 라マン 분광법(Raman spectroscopy), 굴절률 측정, 및 UV/가시적 분광학을 포함하나, 이에 한정되지 않는다. 추가의 방법은 문헌[참조: Amau 등, *Prot Expr and Purif* 48, 1-13 (2006)]에 개재되어 있다.

[0192] XTEN 서열의 추가의 예는 본 발명에 따라 사용될 수 있는 미국 특허 공보 제 2010/0239554 A1호, 제 2010/0323956 A1호, 제2011/0046060 A1호, 제2011/0046061 A1호, 제2011/0077199 A1호, 또는 제2011/0172146 A1호, 또는 국제 특허 공보 제WO 2010091122 A1호, 제WO 2010144502 A2호, 제WO 2010144508 A1호, 제WO 2011028228 A1호, 제WO 2011028229 A1호, 제WO 2011028344 A2호, 또는 제WO2013123457 A1호, 또는 국제 특허원 제 PCT/US2013/049989호에 개재되어 있다.

### II.C.5. FVIII 단백질

[0194] 본원에 사용된 것으로서 "FVIII 단백질"은 달리 정의하지 않는 한, 응고에 있어서 이의 정상적인 역할에 있어서의 기능성 FVIII 폴리펩타이드를 의미한다. 용어 FVIII 단백질은 응고 경로에서 전장의 야생형 인자 VIII의 기능을 보유하는 기능성 단편, 변이체, 유사체, 또는 이의 유도체를 포함한다. "FVIII 단백질"은 FVIII 폴리펩타이드(또는 단백질) 또는 FVIII와 상호교환적으로 사용된다. FVIII 기능의 예는 응고를 활성화시키는 능력, 인자 IX에 대한 보조 인자로서 작용하는 능력, 또는  $\text{Ca}^{2+}$  및 인지질의 존재하에서 인자 IX와 함께 테나제 복합체를 형성한 후 인자 X를 활성화된 형태 Xa로 전환시키는 능력을 포함하나, 이에 한정되지 않는다. FVIII 단백질은 사람, 돼지, 양, 랙트, 또는 쥐 FVIII 단백질일 수 있다. 또한, 사람과 다른 종으로부터의 FVIII 사이의 비교는 기능에 요구되는 것으로 여겨지는 보존된 잔기를 확인하여 왔다(참조: Cameron 등, *Thromb. Haemost.* 79:317-22 (1998); US 6,251,632).

[0195] 다수의 시험이 응고 시스템: 활성화된 부분 트롬보플라스틴 시간(aPTT) 시험, 색원체 검정, ROTEM 검정, 프로트롬빈 시간(PT) 시험(또한 INR을 측정하는데 사용됨), 피브리노겐 시험(흔히 클라우쓰 시험(Clauss method)에 의해), 혈소판 수, 혈소판 기능 시험(흔히 PFA-100에 의해), TCT, 출혈 시간, 혼합 시험(환자의 혈장이 정상 혈장과 혼합되는 경우 비정상성이 교정되는지의 여부), 응고 인자 검정, 항인지질 항체, D-이량체, 유전 시험(예를 들면, 인자 V 라이덴 프로트롬빈 돌연변이(Leiden, prothrombin mutation) G20210A)), 희석된 러셀 사독 시간(dilute Russell's viper venom time: dRVVT), 다양한 혈소판 기능 시험(miscellaneous platelet function test), 혈액 응고 기능항진 (TEG 또는 소노클롯(Sonoclot)), 트롬보엘라스토메트리(thromboelastometry : TEM<sup>®</sup>, 예를 들면, ROTEM<sup>®</sup>), 또는 유글로불린 분해 시험(ELT)을 평가하는데 이용가능하다.

[0196] aPTT 시험은 "고유의"(또한 접촉 활성화 경로로 언급됨) 및 일반적인 응고 경로 둘 다의 효능을 측정하는 수행지시인자이다. 당해 시험은 시판되는 재조합체 응고 인자, 예를 들면, FVIII 또는 FIX의 응고 활성을 측정하기 위해 일반적으로 사용된다. 이는 프로트롬빈 시간(PT)과 함께 사용되며, 이는 고유의 경로를 측정한다.

[0197] ROTEM 분석은 지혈의 전체 역학: 응고 시간, 응괴 정보, 응괴 안전성 및 분해에 대한 정보를 제공한다. 트롬보엘라스토메트리(thromboelastometry)에서의 상이한 매개변수는 혈소판 응고 시스템, 혈소판 기능, 섬유소용해(fibrinolysis), 또는 이를 상호작용에 영향을 미치는 많은 인자의 활성에 의존적이다. 당해 검정은 2차 지혈의

완전한 고찰을 제공할 수 있다.

[0198] FVIII 폴리펩타이드 및 폴리뉴클레오타이드 서열은 많은 기능적 단편, 돌연변이체 및 변형된 버전에서와 같이 공지되어 있다. 사람 FVIII 서열(전장)의 예는 서열 번호 16 또는 18에서의 소서열(subsequence)로서 나타나 있다.

[0199] 표 3 전장의 FVIII(밀줄친 FVIII 시그널 웨პ타이드; FVIII 중쇄는 이중 밑줄쳐져 있다; B 도메인은 이탈릭체이다; 및 FVIII 경쇄는 평문이다)

[0200] 시그널 웨პ타이드: (서열 번호 15)

[0201] MOIELSTCFFLCLLRFCFS

[0202] 성숙 인자 VIII(서열 번호 16)<sup>\*</sup>

ATTRYYLGAVELSWDYMCSPLGELPVDARFPPRVPKSFPNTSVVYKKTLFVEFTDHLFNIAKPRRPWMGLL  
 GPTIQAEVYDTVTITLKNMASHPVSLHAGVSYWKASEGAEYDDNTSQRKEDDKVFPGGSHTYVWQVLKEN  
 GPMASDPCLCTYSYLSHVLDVKDLNSGLIGALLVCREGSLAKEKTOTLHKFILLFAVDEGKSWHSETKNSL  
 MCDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTPEVHSIFLEGHTFLVRNHRQASLET  
 SPITFLTAOTLIMDLGOFLLFCISSHCHDGMEAIVKVDSCPEEPQLRMKNNEFAEDYDDDLTDSEMDVVR  
 DDDNSPSFIRSVAKHPKTWVHYIAAEEDWDYAPLVIAPDORSYKSQYLNNGPDRIGRKYKKVRFMAYIP  
 DTFKTRALCHESGILGPILYGEVGDTLLTIFKNAASRPYNTYPHGIDTVRPLYSRRLPKGVKHLKDFP  
 PGEIFKYKWTVTVEDGPTKSPRCLTRYYSSFVNMERDLASGLIGPLICYKESVDORGNOIMSDKRNVILLE  
 SVFDENRSWYLTEENIQRFLPNPAVGQLEDPEFOASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDF  
 LSVFFSGYTFKHKMVYEDTTLTFFSGETVFMSENPGLWILGCHNSDFRNNGMTALLKVSSCDKNTGDYYP  
 DSYEDIISAYLLSKNNALEPRSFQSNSRHPSTRQKFNATTIPENDIEKTDPWFAHRTPMPKIQNVSSSDLLM  
 LLRQSFTPHGLSLSLDLQEAKYETFSDDPSPGAIDSNNLSEMTHFRPQLHHSGDMVFTPESGLQLRLNEKLG  
 TTAATELKLDKFKVSSTSNNLISTPDSNLAAGTDNTSSLGPPSMVHYDSQLDTTLFGKKSSPLTESGGPL  
 SLSEENNDSKLLESGLMNSQESSWGKNSVSTTESGRFLKGKRAHGCPALLTKDNALFKVSISSLKTNKTSNSA  
 TRRKTHIDGPSLNSPSVWQNILESDFEKVKTPLIHDRLMDKNATALRLNHNMSNKTTSSSKNMEMVQOK  
 KEGPIPPDAQNPDMSFFKMLFLPESARWIQRTHGKNSLNSGQGSPKQLVSLGPEKSVEGQNLSEKNKVVV  
 GKGEFTKDVGLKEMVFPSRNLFLTNLDNLHENNTNHQEKKTQEETIEKKETLIQENVVLQPQIHVTGKFM  
 KNLFLLSTRQNVEGSYDGAYAPVLQDFRSRNDSTNRTRKHTAHFSKGEEENLEGLGNQTKQIVEKYACTTR  
 ISPNTSQQNFVTQRSKRALQKFRLPLEETELEKRIIVDDTSTQWSKNMKHLPSTLTQIDYNEKEKGAITQS  
 PLSDCLTRSHSIQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDGSVQESSHFLQGAKK  
 NNLSLAITLEMTGDQREVGSLGTSATNSTVYKKVENTVLPKPDLPKTSGKVELLPKVHIYQKDLFPTETSN  
 GSPGHLDLVEGSSLQGTEGAIKWNEANRPGKVPFLRVATESAKTPSKLLDPLAWDNHYGTQIPKEEWKSQE  
 KSPEKTAFKKKDTILSLNACESNHAAIAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKRHQREITRTTLQ  
 SDQEEIDYDDTISVEMKKEDFIDYDENQSPRSFSQKKTQHFIAAVERLWDYGMSSSPHVLRNRAQSGSVP  
 QFKKVVQEFDTGSFTQPLYRGELNEHLGLGPYIRAEVEDNIMVTFRNQASRPSFYSSLISYEEDQRQGA  
 EPRKNFVVKPNETKTYFWKVQHMAPTKDEFDCAKAWAYFSDVDLEKDVHSGLIGPLLCVHTNTLNPAHGRQVT  
 VQEFAFFTIFDETKSWYFTENMERNCAPCNIQMEDPTFKENYRFAINGYIMDTLPGLVMAQDQRIRWYL  
 LSMGSNEIHSIHFGHVFTVKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLV  
 YSNKCTPLGMASGHIRDFQITASQYQGWAPKLARLHYSGSINAWSTKEPFWSIKVVDLLAPMIIHGIKTQG  
 AROKFSSLYISQFIIMYSLDGKQTYRGNSTGTLMMVFFGNVDSSGIKHNIFNQPIARYIRLHPTHYSIRS  
 TLRMELMGCDLNNSCMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHQGRSNAWRPQVNNPKEWLQV  
 DFQKTMKVTVQGVKSLLTMSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNNQDSFTPVVNSLDPLLTR  
 YLRIHPQSWVHQIALRMEVLGCEAQDLY

[0203]

[0204] 표 4. 전장의 FVIII를 암호화하는 뉴클레오타이드 서열(서열 번호 17)<sup>\*</sup>

661	ATG	CAAATAGAGC	TCTCCACCTG			
721	CTTCCTTCTG	TGCCCTTTGC	GATTCTGCTT	TAGTGCACC	AGAACGATACT	ACCTGGGTGC
781	AGTGGACTG	TCATGGACT	ATATGCAAAAG	TGATCTCGGT	GAGCTGCCTG	TGGACGCAAG
841	ATTTCCCTC	AGAGTCACCA	AATCTTTCC	ATTCAACACC	TCAGTCGTGT	ACAAAAAGAC
901	TCTGTTGTA	GAATTACCGG	ATCACCTTT	CAACATCGCT	AAGCCAAGGC	CACCTGGAT
961	GGGTCTGCTA	GGTCTTACCA	TCCAGGCTGA	GGTTTATGAT	ACAGTGGTCA	TTACACTTAA
1021	GAACATGGCT	TCCCACATCTG	TCAGTCTTCA	TGCTGTTGGT	GTATCCTACT	GGAAAGCTTC
1081	TGAGGGAGCT	GAATATGATG	ATCAGACCAAG	TCAAAGGGAG	AAAGAACGATG	ATAAAAGTCTT
1141	CCCTGGTGGA	AGCCATACAT	ATGTCCTGGCA	GGTCCTGAAA	GAGAATGGTC	CAATGGCCTC
1201	TGACCCACTG	TGCCCTTACCT	ACTCATATCT	TTCTCATGTG	GACCTGGTAA	AAGACTTGAA
1261	TTCAGGCCCTC	ATGGGAGCCC	TACTAGTATG	TAGAGAACGGG	AGTCTGGCCA	AGGAAAAGAC
1321	ACAGACCTTG	CACAAATTTA	TACTACTTTT	TGCTGTATTT	GATGAAGGGGA	AAAGTTGGCA
1381	CTCAGAAACA	AAAAGACTCCT	TGATGCAGGA	TAGGGATGCT	GCATCTGCTC	GGGCCTGGCC
1441	TAAAATGCAC	ACAGTCAATG	GTTATGTAAA	CAGGTCTCTG	CCAGGTCTGA	TTGGATGCCA
1501	CAGGAAATCA	GTCTATTGGC	ATGTGATTGG	AATGGGCACC	ACTCCTGAAG	TGCACTCAAT
1561	ATTCCCTCGAA	GGTCACACAT	TTCTTGTGAG	GAACCATCGC	CAGGCGTCCT	TGGAAATCTC

[0205]

1621 GCCAATAACT TTCCCTTACTG CTCAAACACT CTTGATGGAC CTTGGACAGT TTCTACTGTT  
 1681 TTGTCAATAC TCTTCCCACC AACATGATGG CATGGAAGCT TATGTCAAAG TAGACAGCTG  
 1741 TCCAGAGGAA CCCCAACTAC GAATGAAAAA TAATGAAGAA CGCGAAGACT ATGATGATGA  
 1801 TCTTACTGAT TCTGAAATGG ATGTGGTCAG GTTGATGAT GACAACCTCTC CTTCCCTTAT  
 1861 CCAAATTGCG TCAGTTGCCA AGAACATCC TAAAACCTGG GTACATTACA TTGCTGCTGA  
 1921 AGAGGAGGAC TGGGACTATG CTCCCTTAGT CCTCGCCCC GATGACAGAA GTTATAAAAG  
 1981 TCAATATTG AACAATGGC CTCAGGGAT TGGTAGGAAG TACAAAAAAG TCCGATTTAT  
 2041 GGCATACACA GATGAAACCT TTAAGACTCG TGAAGCTATT CAGCATGAAT CAGGAATCTT  
 2101 GGGACCTTTA CTTTATGGGG AAGTTGGAGA CACACTGTT ATTATATTAA AGAATCAAGC  
 2161 AAGCAGACCA TATAACATCT ACCCTCACCG AATCACTGAT GTCCGTCCTT TGTATTCAAG  
 2221 GAGATTACCA AAAGGTGAA AACATTGAA GGATTTCCA ATTCTGCCAG GAGAAATATT  
 2281 CAAATATAAA TGGACAGTGA CTGTTAGAAGA TGGGCAACT AAATCAGATC CTCGGTGCCT  
 2341 GACCCGCTAT TACTCTAGTT TCGTTAATAT GGAGAGAGAT CTAGCTTCAG GACTCATTGG  
 2401 CCCTCTCTC ATCTGCTACA AAGAATCTGT AGATCAAAGA GGAACACAGA TAATGTCAGA  
 2461 CAAGAGGAAT GTCATCTGT TTTCTGTATT TGATGAGAAC CGAAGCTGGT ACCTCACAGA  
 2521 GAATATACAA CGCTTTCTCC CCAATCCAGC TGGAGTCAG CTTGAGGATC CAGAGTTCCA  
 2581 AGCCTCCAAC ATCATGCACA GCATCAATGG CTATGTTTT GATACTTGC AGTTGTCAGT  
 2641 TTGTTGCAT GAGGTGGCAT ACTGGTACAT TCTAACGATT GGAGCACAGA CTGACTTCCT  
 2701 TTCTGTCTTC TTCTCTGGAT ATACCTTCAA ACACAAAATG GTCTATGAAG ACACACTCAC  
 2761 CCTATTCCA TTCTCAAGGAG AACTGTCCTT CATGTCGATG GAAAACCCAG GTCTATGGAT  
 2821 TCTGGGTGCA CACAACCTCAG ACTTTCGAA CAGAGGCATG ACCGCTTAC TGAAGGTTTC  
 2881 TAGTTGTGAC AAGAACACTG GTGATTATTA CGAGGACAGT TATGAAGATA TTTCAGCATA  
 2941 CTTGCTGAGT AAAAACATG CCATTGAATC AAGAACGCTTC TCCCGAGAATT CAAGACACCC  
 3001 TAGCACTAGG CAAAAGCAAT TTAATGCCAC CACAATTCCA GAAAATGACA TAGAGAAGAC  
 3061 TGACCCCTGG TTTGCACACA GAACACCTAT GCCTAAAATA CAAAATGTCT CCTCTAGTGA  
 3121 TTTGTTGATG CTCTTGCAC AGAGTCCTAC TCCACATGGG CTATCCTTAT CTGATCTCCA  
 3181 AGAACCCAAA TATGAGACTT TTTCTGTGATG TCCATCACCT GGAGCAATAG ACAGTAATAA  
 3241 CAGCTGTCT GAAAATGACAC ACTTCAGGCC ACAGCTCCAT CACAGTGGGG ACATGGTATT  
 3301 TACCCCTGAG TCAGGCCCTC AATTAAGATT AAATGAGAAA CTGGGGACAA CTGCAGCAAC  
 3361 AGAGTTGAAG AACTTGATT TCAAAGTTTCA TACTACATCA AATAATCTGA TTTCACAAAT  
 3421 TCCATCAGAC AATTTCGAG CAGGTTACTGA TAATACAAAGT CCCTTACGAC CCCCAAGTAT  
 3481 GCCAGTTCAT TATGATAGTC AATTAGATAC CACTCTATTT GGCAAAAGT CATCTCCCT  
 3541 TACTGAGTCT GGTGGACCTC TGAGCTTGAG TGAAGAAAAT AATGATTCAA AGTTGTTAGA  
 3601 ATCAGGTTA ATGAATAGCC AAGAAAGTTC ATGGGGAAA AATGTATCGT CAACAGAGAG  
 3661 TGGTAGGTTA TTTAAAGGGA AAAGAGCTCA TGGACCTGCT TTGTTGACTA AAGATAATGC  
 3721 CTTATTCAAA GTTAGCATCT CTTTGTAAA GACAAACAAA ACTTCCAATA ATTCAAGCAAC  
 3781 TAATAGAAAG ACTCACATTG ATGGCCCATG ATTATTAATT GAGAATAGTC CATCAGTCTG  
 3841 GCAAAATATA TTAGAAAGTG ACACTGAGTT TAAAAAAAGTG ACACCTTGA TTCAATGACAG  
 3901 AATGCTTATG GACAAAATG CTACAGCTT GAGGCTAAAT CATATGTCAA ATAAAATCAC  
 3961 TTCAATCAAA AACATGGAAA TGGTCCAACA GAAAAAAAGAG GGGCCCATTC CACCAGATGC  
 4021 ACAAAATCCA GATATGTCGT TCTTTAAAGAT GCTATTCTTG CCAGAACATCAG CAAGGTGGAT  
 4081 ACAAAAGGACT CATGGAAAGA ACTCTCTGAA CTCTGGCAA GGGCCCAAGTC CAAAGCAATT  
 4141 AGTATCCTTA GGACCAAGAAA AATCTGTGGA AGGTCAAGAAT TTCTTGCTG AGAAAAAACAA  
 4201 AGTGGTAGTGA GGAAAGGGTG AATTTCACAA GGACGTAGGA CTCAAAGAGA TGGTTTTCC  
 4261 AAGCAGCAGA AACCTATTTC TTACTAACTT GGATAATTAA CATGAAAATA ATACACACAA  
 4321 TCAAGAAAAA AAAATTCAAGG AAGAAATAGA AAAGAAGGAA ACATTAATCC AAGAGAATGT  
 4381 AGTTTGCCCT CAGATACATA CAGTGAETGG CACTAAGAAT TTCATGAAGA ACCTTTCTT  
 4441 ACTGAGCACT AGGCAAAATG TAGAAGGTT ATATGACGGG GCATATGCTC CAGTACTTCA  
 4501 AGATTTAGG TCATTAATG ATTCAACAAA TAGAACAAAG AAACACACAG CTCATTTCTC  
 4561 AAAAAAAAGGG GAGGAAGAAA ACTTGGAGG CTTGGGAAAT CAAACCAAGC AAATTGTTAGA  
 4621 GAAATATGCA TGCAACACAA GGATATCTCC TAATACAAGC CAGCAGAATT TTGTCACGCC  
 4681 ACGTAGTAAG AGAGCTTTGA AACAAATTCC ACTCCCACTA GAAGAAACAG AACTTGAAA  
 4741 AAGGATAATT GTGGATGACA CCTCAACCCA GTGGTCCAAA AACATGAAAC ATTTGACCCC  
 4801 GAGCACCCCTC ACACAGATAG ACTACAAATGA GAAGGAGAAA GGGGCCATTA CTCAGTCTCC  
 4861 CTTATCAGAT TGCCCTTACGA GGAGTCATAG CATCCCTCAA GCAAATAGAT CTCCATTACC  
 4921 CATTGCAAAG GTATCATCAT TTCCATCTAT TAGACCTATA TATCTGACCA GGGTCCTATT  
 4981 CCAAGACAAC TCTTCTCATC TTCCAGCAGC ATCTTATAGA AAGAAAGATT CTGGGGTCCA

5041 AGAAAGCACT CATTCTTAC AAGGAGCCAA AAAAATAAC CTTCTTTAG CCATTCTAAC  
 5101 CTTGGAGATG ACTGGTGATC AAAGAGAGGT TGGCTCCCTG GGGACAAGTG CCACAAATTC  
 5161 AGTCACATAC AAGAAAGTT AGAACACTGT TCTCCCGAAA CCAGACTTG CCAAACATC  
 5221 TGGCAAAGTT GAATTGCTTC CAAAAGTTCA CATTATTCAG AAGGACCTAT TCCCCTACGGA  
 5281 AACTAGCAAT GGGTCTCCTG GCCATCTGGA TCTCGTGGAA GGGAGCCTTC TTCAAGGGAAC  
 5341 AGAGGGAGCG ATTAAGTGGAA ATGAAGCAAA CAGACCTGGA AAAGTTCCCT TTCTGAGAGT  
 5401 AGCAACAGAA AGCTCTGCAA AGACTCCCTC CAAGCTATTG GATCCTCTTG CTTGGATAA  
 5461 CCACTATGGT ACTCAGATAC CAAAAGAAGA GTGGAATTC CAAAGAGAAGT CACCAAGAAAA  
 5521 AACAGCTTT AAGAAAAGG ATACCATTTC GTCCTGAAC GCTTGTGAAA GCAATCATGC  
 5581 AATAGCAGCA ATAAATGAGG GACAAAATAA GCCCAGAATAA GAAGTCACCT GGGCAAGCA  
 5641 AGTAGGACT GAAAGGCTGT GCTCTCAAA CCCACCAAGTC TTGAAACGCC ATCAACGGGA  
 5701 AATAACTCGT ACTACTCTTC AGTCAGATCA AGAGGAAATT GACTATGATG ATACCATATC  
 5761 AGTTGAAATG AAGAAGGAAAG ATTTTGACAT TTATGATGAG GATGAAAATC AGAGCCCCCG  
 5821 CAGCTTCAA AAGAAAACAC GACACTATT TATTGCTGCA GTGGAGAGGC TCTGGGATTA  
 5881 TGGGATGAGT AGCTCCCCAC ATGTTCTAAG AAACAGGGCT CAGAGTGGCA GTGTCCTCA  
 5941 GTTCAAGAAA GTTGTTCAGGAGAATTTC TGATGGCTCC TTTACTCAGC CCTTATACCG  
 6001 TGGAGAACTA ATGAACATT TGGGACTCCT GGGGCCATAT ATAAGAGCAG AAGTTGAAGA  
 6061 TAATATCATG GTAACCTTC GAAATCAGGC CTCTCGTCCC TATTCTTCT ATTCTAGCCT  
 6121 TATTCTTAT GAGGAAGATC AGAGGCAAGG AGCAGAACCT AGAAAAAAACT TTGTCAAGCC  
 6181 TAATGAAACC AAAACTTACT TTTGGAAAGT GCAACATCAT ATGGCACCCA CTAAGATGTA  
 6241 GTTGTACTGC AAAGCCTGGG CTTATTCTC TGATGTTGAC CTGGAAAAAG ATGTGCACTC  
 6301 AGGCCTGATT GGACCCCTTC TGGTCTGCCA CACTAACACA CTGAACCTTG CTCATGGGAG  
 6361 ACAAGTGACA GTACAGGAAT TTGCTCTGTT TTCACCACAT TTTGATGAGA CCAAAAGCTG  
 6421 GTACTTCACT GAAAATATGG AAAGAAACTG CAGGGCTCCC TGCAATATCC AGATGGAAGA  
 6481 TCCCACTTTT AAAGAGAATT ATCGCTTCCA TGCAATCAAT GGCTACATAA TGGATACACT  
 6541 ACCTGGCTTA GTAATGGCTC AGGATCAAAG GATTGATGG TATCTGCTCA GCATGGGCAG  
 6601 CAATGAAAAC ATCCATTCTA TTCATTTCAAG TGGACATGTG TTCACTGTAC GAAAAAAAGA  
 6661 GGAGTATAAA ATGGCACTGT ACAATCTCTA TCCAGGTGTT TTTGAGACAG TGGAAATGTT  
 6721 ACCATCCAAA GCTGGAATTG GGGGGGTGGA ATGCCATTG TGCGAGCATC TACATGCTGG  
 6781 GATGAGCACA CTTTTCTG TGATGACAA TAAGTGTCACT CTCCTCCCTGG GAATGGCTTC  
 6841 TGGACACATT AGAGATTTTC AGATTACAGC TTCAGGACAA TATGGACAGT GGGCCCCAAA  
 6901 GCTGGCCAGA CTTCATTATT CGGGATCAAT CAATGCCCTGG AGCACCAAGG AGCCCTTTTC  
 6961 TTGGATCAAG GTGGATCTGT TGGCACCAAT GATTATTCAC GGCATCAAGA CCCAGGGTGC  
 7021 CCGTCAGAAG TTCTCCAGCC TCTACATCTC TCAGTTTATC ATCATGTATA GTCTTGATGG  
 7081 GAAGAAGTGG CAGACTTATC GAGGAATTG CACTGGAAACC TTAATGGTCT TCTTGGCAA  
 7141 TGTGGATTCA TCTGGATAA AACACAATAT TTTTAACCCCT CCAATTATTG CTCGATACAT  
 7201 CCGTTTGAC CCAACTCATT ATAGCATTG CAGCACTCTT CGCATGGAGT TGATGGCTG  
 7261 TGATTTAAAT AGTTGCAGCA TGCCATTGGG AATGGAGAGT AAAGCAATAT CAGATGCACA  
 7321 GATTACTGCT TCATCCTACT TTACCAATT GTTTGCACCC TTGTCCTCCTT CAAAAGCTCG  
 7381 ACTTCACCTC CAAGGGAGGA GTAATGCCCTG GAGACCTCG GTGATAATC CAAAAGAGTG  
 7441 GCTGCAAGTG GACTTCCAGA AGAACATGAA AGTCACAGGA GTAACTACTC AGGGAGTAAA  
 7501 ATCTCTGCTT ACCAGCATGT ATGTGAAGGA GTTCCCTCATC TCCAGCAGTC AAGATGGCCA  
 7561 TCAGTGGACT CTCTTTTTTC AGAATGGCAA ACTAAAGGTT TTTCAGGGAA ATCAAGACTC  
 7621 CTTCACACCT GTGGTGAACCTCTAGACCC ACCGTTACTG ACTCGCTACC TTCAATTCA  
 7681 CCCCCAGAGT TGGGTGCACC AGATTGCCCT GAGGATGGAG GTTCTGGGCT GCGAGGCACA  
 7741 GGACCTCTAC

[0207]

[0208]

\* 밀줄친 핵산은 시그널 웨타이드를 암호화한다.

[0209]

FVIII 폴리펩타이드는 전장의 FVIII, N-말단에서 Met가 빠진 완전한 길이의 FVIII, 성숙한 FVIII(시그널 서열 빠짐), N-말단에서 추가의 Met를 지닌 성숙한 FVIII, 및/또는 B 도메인의 완전한 또는 부분 결실이 있는 FVIII를 포함한다. 특정의 구현예에서, FVIII 변이체는, 부분적인 또는 완전한 결실에 상관없이, B 도메인 결실을 포함한다.

[0210]

사람 FVIII 유전자는 포유동물 세포내에서 분리되어 발현된다[참조: Toole, J. J., 등, *Nature* 312:342-347 (1984); Gitschier, J., 등, *Nature* 312:326-330 (1984); Wood, W. I., 등, *Nature* 312:330-337 (1984); Vehar, G. A., 등, *Nature* 312:337-342 (1984); WO 87/04187; WO 88/08035; WO 88/03558; 및 특허 제4,757,006호]. FVIII 아미노산 서열은 미국 특허 제4,965,199호에 나타낸 바와 같이 cDNA로부터 유추되었다. 또한, 부분적으로 또는 완전한 B-도메인이 결실된 FVIII는 미국 특허 제4,994,371호 및 제4,868,112호에 나타난다. 일부 구현예에서, 사람 FVIII B-도메인은 미국 특허 제5,004,803호에 나타낸 바와 같이 사람 인자 V B-도메인으로 대체된다. 사람 인자 VIII를 암호화하는 cDNA 서열 및 아미노산 서열은 미국 특허 제7,211,559호의 서열 번호 1 및 2에 각각 나타나 있다.

[0211]

돼지 FVIII 서열은 문헌[참조: Toole, J. J., 등, *Proc. Natl. Acad. Sci. USA* 83:5939-5942 (1986)]에 발표되어 있다. 또한, 돼지 cDNA 라이브러리로부터의 FVIII 서열의 PCR 증폭으로부터 수득된 완전한 돼지 cDNA는 문헌[참조: Healey, J. F., 등, *Blood* 88:4209-4214 (1996)]에 보고되어 있다. 모든 도메인, 모든 소단위, 및 특이적인 아미노산 서열을 지닌 하이브리드 사람/돼지 FVIII은 미국 특허 제5,364,771호(Lollar 및 Runge), 및 제WO 93/20093호에 개재되어 있다. 보다 최근에, 돼지 FVIII의 A1 및 A2 도메인의 뉴클레오타이드 및 상응하는 아미

노산 서열 및 상응하는 사람 도메인이 치환된 돼지 A1 및/또는 A2 도메인과의 키메라 FVIII는 제WO 94/11503호에 보고되었다. 미국 특허 제5,859,204호(Lollar, J. S.)는 또한 돼지 cDNA 및 유추된 아미노산 서열을 개재하고 있다. 미국 특허 제6,458,563호는 B-도메인-결실된 돼지 FVIII을 개재하고 있다.

[0212] 미국 특허 제5,859,204호(Lollar, J. S.)는 감소된 항원성 및 감소된 면역반응성을 지닌 FVIII의 기능성 돌연변이체를 보고하고 있다. 미국 특허 제6,376,463호(Lollar, J. S.)는 또한 감소된 면역반응성을 지닌 FVIII의 돌연변이체를 보고하고 있다. 미국 특허원 공보 제2005/0100990호(Saenko 등)는 FVIII의 A2 도메인내 기능성 돌연변이를 보고하고 있다.

[0213] 하나의 구현예에서, FVIII 단백질(또는 키메라 단백질의 FVIII 부위)은 서열 번호 18의 1 내지 1438번 아미노산의 FVIII 아미노산 서열 또는 서열 번호 16의 1 내지 2332번 아미노산(시그날 웨타이드 없음)에 대해 적어도 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하며, 여기서 FVIII는 활성 활성을 가지는데, 예를 들면, 보조인자로서 인자 IX를 활성화시켜 인자 X를 활성화된 인자 X로 전환시킨다. FVIII(또는 키메라 단백질의 FVIII 부위)는 서열 번호 18의 1 내지 1438번 아미노산의 FVIII 아미노산 서열 또는 서열 번호 16의 1 내지 2332번 아미노산(시그날 웨타이드 없음)과 동일할 수 있다. FVIII 단백질은 또한 시그날 웨타이드를 포함할 수 있다.

[0214] 본원에 사용된 것으로서 "FVIII의 B-도메인"은 내부 아미노산 서열 동질성 및 단백질분해적 절단 부위, 예를 들면, 전장의 사람 FVIII의 잔기 Ser741 내지 Arg1648번에 의해 정의된 당해 분야에 공지된 B-도메인과 동일하다. 다른 사람 FVIII 도메인은 다음의 아미노산 잔기에 의해 정의된다: A1, 잔기 Ala1 내지 Arg372; A2, 잔기 Ser373 내지 Arg740; A3, 잔기 Ser1690 내지 Asn2019; C1, 잔기 Lys2020 내지 Asn2172; C2, 잔기 Ser2173 내지 Tyr2332. A3-C1-C2 서열은 잔기 Ser1690 내지 Tyr2332를 포함한다. 나머지 서열, 잔기 Glu1649 내지 Arg1689는 일반적으로 a3 산성 영역으로 언급된다. 돼지, 마우스 및 개 FVIII에 대한 B-도메인을 포함하는, 모든 도메인에 대한 경계부의 위치는 당해 분야에 또한 공지되어 있다. 하나의 구현예에서, FVIII의 B 도메인은 결실되어 있다 ("B-도메인-결실된 인자 VIII" 또는 "BDD FVIII"). BDD FVIII의 예는 REFACTO<sup>®</sup>(제조업체 BDD FVIII)이며, 이는 표 5의 서열의 인자 VIII 부위와 동일한 서열을 갖는다. (BDD FVIII 중쇄는 이중 밀줄쳐져 있고; B 도메인은 이탈릭체이며; BDD FVIII 경쇄는 평서문이다).

표 5. BDD FVIII(서열 번호 18)

```

ATRRYYLGAVELSWDYMOSDLGELPVDFPPRVPKSFPNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLI
GPTIQAEVYDTVVITLKNMASHPVSIHAVGVSYWKASEGAEYDDQTSQREKEKEDDKVFPGGSHTYVWQVLKEN
GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTCTLHKFILLFAVFDEGKSWHSETKNSL
MDRDAASARAWPKMHTVNGYVNRSILPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRJASLEI
SPITFLTAYTLLMDLGQFLLFCHISYHDGMEAAYVKVDSCPEEPPLRMKNN  
EEAEDYDDDLDSEMDVVRF
DDDNSPSFIIRSVAKKHPKTWVHYIAAEEDWYAPLVLAPLDDRSYKSCYLNNGPRIGRKYKKVRFMAYY
DETFKTREAQHESGILGPLIYGEVGDTILLIIFKNAQSRPYNIYPHGIDVRPLYSRIPKGVKHLKDFPI
PGEIFKYKWTVTVEDGPTKSDDPRCITRYSSFVNMERDILASGLIGPLILYCYESVDIRGNOIMSDKRNVILF
SVDENPWYLTENIRFLPNPAGVOLEDPEFOQASNIMHSINGYVFDSILSVCLHEVAYWYILSIGAONTDF
LSVFFSGYTFKHMVYEDTITLFFFSGETVFMSMENPGLWILGCHNSDFRNRGMTALKVSSCDKNTGDYYE
DSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTLQSDQEEIYDDDTISVEMKKEDFDIYDEDENQ
SPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSPQFKVVVFQEFTDGSFTQPLYRGELNEHLGL
LGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEF
DCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIDETKSWYFTIMRNCRAP
CNIQMEPTFKENYRFHAINGYIMDTLPGLVMQAQDQRIRWYLLSMGSNEIHSIHFSGHVFTVRKKEEYKMA
LYNLYPGVFETEMLPSKAGIWRVECLIGEHLHAGMSTLFVYSNKCQTPLGMASGHIRDFQITASGQYGQW
APKLARLHYSGSINAWSTKEPFSWIKVDLLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGN
STGTLMVFFGNVDSSGIKHNINFPPIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQI

```

TASSYFTNM<sup>F</sup>ATWS<sup>S</sup>PSKARL<sup>H</sup>LQGR<sup>S</sup>NAWR<sup>P</sup>QVNNP<sup>K</sup>E<sup>W</sup>LQVDFQ<sup>K</sup>TMKV<sup>T</sup>GV<sup>V</sup>T<sup>T</sup>QGV<sup>K</sup>SL<sup>T</sup>SMYV<sup>K</sup><sup>E</sup>FL<sup>I</sup>
SSSQDGHQWT<sup>L</sup>FFFQNGKV<sup>K</sup>V<sup>F</sup>QGNQDS<sup>F</sup>TPVV<sup>N</sup>SLD<sup>P</sup>PLL<sup>T</sup>RY<sup>L</sup>R<sup>I</sup>HPQ<sup>S</sup>WV<sup>H</sup>Q<sup>I</sup>ALRM<sup>E</sup>VLGCEAQDLY

[0218] 표 6. BDD FVIII를 암호화하는 뉴클레오타이드 서열(서열 번호 19)<sup>\*</sup>

661 A TGCAAATAGA GCTCTCCACC TGCTTCTTC  
 721 TGTGCCTTT GCGATTCTGC TTTAGTGCCA CCAGAACATA CTACCTGGGT GCAGTGGAAC  
 781 TGTATGGGA CTATATGCAA AGTGATCTG GTGAGCTGCC TGTGGACGCA AGATTTCCTC  
 841 CTAGAGTGC AAAATCTTT CCATTCAACA CCTCAGTCGT GTACAAAAG ACTCTGTTT  
 901 TAGAATTACAC GGATCACCTT TTCAACATCG CTAAGCCAAG GCCACCCCTGG ATGGGTCTGC  
 961 TAGGTCCTAC CATCCAGGCT GAGGTTATG ATACAGTGGT CATTACACTT AAGAACATGG  
 1021 CTTCCCATCC TGTCACTT CATGCTGTT GTGTATCCTA CTGGAAAGCT TCTGAGGGAG  
 1081 CTGAATATGA TGATCAGACC AGTCAAAGGG AGAAAAGAAGA TGATAAAAGTC TTCCCTGGTG  
 1141 GAAGCCATAC ATATGTCCTGG CAGGTCCCTGA AAGAGAATGG TCCAATGGCC TCTGACCCAC  
 1201 TGTGCCTTAC CTACTCATAT CTTTCTCATG TGGACCTGGT AAAAGACTTG AATTCAAGGCC  
 1261 TCATTGGAGC CCTACTAGTA TGTAGAGAAG GGAGTCTGGC CAAGGAAAAG ACACAGACCT  
 1321 TGCACAAATT TATACTACTT TTTGCTGTAT TTGATGAAGG GAAAAGTTGG CACTCAGAAA  
 1381 CAAAGAACTC CTTGATGCAG GATAGGGATG CTGCATCTGC TCGGGCCTGG CCTAAAATGC  
 1441 ACACAGTCAA TGTTTATGTA AACAGGTCTC TGCCAGGTCT GATTGGATGC CACAGGAAAT  
 1501 CAGTCTATTG GCATGTGATT GGAATGGGCA CCACTCCTGA AGTGC CTCA ATATTCCCTG  
 1561 AAGGTACAC ATTTCTTGTG AGGAACCATC GCCAGGGCTC CTGGAAATC TCGCCAATAA  
 1621 CTTTCCCTTAC TGCTCAAACA CTCTTGATGG ACCTTGGACA GTTCTACTG TTTTGTCTATA  
 1681 TCTCTTCCCA CCAACATGAT GGCATGGAAG CTTATGTCAA AGTAGACAGC TGTCCAGAGG  
 1741 AACCCCAACT ACGAATGAAA ATAATGAAG AAGCGGAAGA CTATGATGAT GATCTTACTG  
 1801 ATTCTGAAAT GGATGTGGTC AGGTTTGATG ATGACAACTC TCCTTCCTT ATCCAAATTC  
 1861 GCTCAGTTGC CAAGAACAT CCTAAAACCTT GGGTACATTA CATTGCTGCT GAAGAGGAGG  
 1921 ACTGGGACTA TGCTCCCTTA GTCCTCGCCC CCGATGACAG AAGTTATAAA AGTCAATATT  
 1981 TGAACAATGG CCCTCAGCGG ATTGGTAGGA AGTACAAAAA AGTCCGATTT ATGGCATA  
 2041 CAGATGAAAC CTTTAAGACT CGTGAAGCTA TTCAGCATGA ATCAGGAATC TTGGGACCTT  
 2101 TACTTTATGG GGAAGTTGGA GACACACTGT TGATTATATT TAAGAATCAA GCAAGCAGAC  
 2161 CATATAACAT CTACCCCTAC CGAACATCACTG ATGTCCTGTCC TTTGTATTCA AGGAGATTAC  
 2221 CAAAAGGTGT AAAACATTTC AAGGATTTTTC CAATTCTGCC AGGAGAAATA TTCAAATATA  
 2281 AATGGACAGT GACTGTAGAA GATGGGCCAA CTAATCAGA TCCTCGGTGC CTGACCCGCT  
 2341 ATTACTCTAG TTTCGTTAAT ATGGAGAGAG ATCTAGCTTC AGGACTCATT GGGCCTCTCC  
 2401 TCATCTGCTA CAAAGAACTC GTAGATCAAAG GAGGAACCA GATAATGTCA GACAAGAGGA  
 2461 ATGTCATCCT GTTTCTGTG TTTGATGAGA ACCGAAGCTG GTACCTCACA GAGAATATAC  
 2521 AACGCTTTCT CCCCAATCCA GCTGGAGTGC AGCTTGAGGA TCCAGAGTTC CAAGCCTCCA  
 2581 ACATCATGCA CAGCATCAAT GGGTATGTT TTGATAGTT GCAGTTGTCA GTTTGTTG  
 2641 ATGAGGTGGC ATACTGGTAC ATTCTAAGCA TTGGAGCACA GACTGACTTC CTTTCTGTCT  
 2701 TCTTCTCTGG ATATACCTTC AAACACAAAAA TGGTCTATGA AGACACACTC ACCCTATTCC  
 2761 CATTCTCAGG AGAAAATGTC TTCATGTCGA TGAAAACCC AGGTCTATGG ATTCTGGGGT  
 2821 GCCACAACTC AGACTTTCGG AACAGAGGCA TGACCCCTT ACTGAAGGTT TCTAGTTGTG  
 2881 ACAAGAACAC TGTTGATTAT TACGAGGACA GTTATGAAGA TATTCAGCA TACTTGCTGA  
 2941 GTAAAAACAA TGCCATTGAA CCAAGAACGT TCTCTCAAA CCCACCAGTC TTGAAACGCC  
 3001 ATCAACGGGA AATAACTCGT ACTACTCTTC AGTCAGATCA AGAGGAAATT GACTATGATG  
 3061 ATACCATATC AGTTGAAATG AAGAAGGAAG ATTITGACAT TTATGATGAG GATGAAAATC  
 3121 AGAGCCCCCG CAGCTTCAA AAGAAAACAC GACACTATT TATTGCTGCA GTGGAGAGGC  
 3181 TCTGGGATTA TGGGATGAGT AGCTCCCCAC ATGTTCTAAG AAACAGGGCT CAGAGTGGCA  
 3241 GTGTCCCTCA GTTCAAGAAA GTTGTCTTCC AGGAATTAC TGATGGCTCC TTTACTCAGC  
 3301 CCTTATACCG TGGAGAACTA AATGAACATT TGGGACTCCT GGGGCCATAT ATAAGAGCAG  
 3361 AAGTTGAAGA TAATATCATG GTAACCTTCA GAAATCAGGC CTCTCGTCCC TATTCTTCT  
 3421 ATTCTAGCCT TATTTCTTAT GAGGAAGATC AGAGGCAAGG AGCAGAACCT AGAAAAAAACT  
 3481 TTGTCAAGCC TAATGAAACC AAAACTTAAT TTTGGAAAGT GCAACATCAT ATGGCACCCA  
 3541 CTAAGAGATGA GTTGTCTGC AAAGCCTGGG CTTTATTCCT TGATGTTGAC CTGGAAAAG  
 3601 ATGTGCACTC AGGCCCTGATT GGACCCCTTC TGGTCTGCCA CACTAACACA CTGAACCCCTG  
 3661 CTCATGGGAG ACAAGTACA GTACAGGAAT TTGCTCTGTT TTTCACCATC TTTGATGAGA  
 3721 CCAAAAGCTG GTACTTCACT GAAAATATGG AAAGAAAATG CAGGGCTCCC TGCAATATCC

[0219]

3781 AGATGGAAGA TCCCACCTTT AAAGAGAATT ATCGCTTCCA TGCAATCAA GGCTACATAA  
 3841 TGGATACACT ACCTGGCTTA GTAATGGCTC AGGATCAAAG GATTCGATGG TATCTGCTCA  
 3901 GCATGGGAG CAATGAAAAC ATCCATTCTA TTCAATTTCAG TGGAACATGTG TTCACTGTAC  
 3961 GAAAAAAAGA GGAGTATAAA ATGGCACTGT ACAATCTCTA TCCAGGTGTT TTTGAGACAG  
 4021 TGGAAATGTT ACCATCCAAA GCTGGAATTG GGCGGGTGGGA ATGCCTTATT GGCGAGCATC  
 4081 TACATGCTGG GATGAGCACA CTTTTCTGG TGTACAGCAA TAAGTGTCAAG ACTCCCCTGG  
 4141 GAATGGCTTC TGGACACATT AGAGATTTC AGATTACAGC TTCAGGACAA TATGGACAGT  
 4201 GGGCCCCAAA GCTGGCCAGA CTTCATTATT CGGATCAAT CAATGCCCTGG AGCACCAAGG  
 4261 AGCCCTTTTC TTGGATCAAG GTGGATCTGT TGGCACCAAT GATTATTCAC GGCATCAAGA  
 4321 CCCAGGGTGC CCGTCAGAAG TTCTCCAGCC TCTACATCTC TCAGTTTATC ATCATGTATA  
 4381 GTCTTGATGG GAAGAAGTGG CAGACTTAC GAGGAATTTC CACTGGAACC TTAATGGTCT  
 4441 TCTTGGCAA TGTGGATTCA TCTGGGATAA AACACAAATAT TTTTAACCTT CCAATTATTG  
 4501 CTCGATACAT CGGTTTGAC CCAACTCATT ATAGCATTCTG CAGCACTCTT CGCATGGAGT  
 4561 TGATGGCTG TGATTTAAAT AGTTGCAGCA TGCCATTGGG AATGGAGAGT AAAGCAATAT  
 4621 CAGATGCAACA GATTACTGCT TCATCCTACT TTACCAAAATAT GTTTGCCACC TGGTCTCCTT  
 4681 CAAAAGCTCG ACTTCACCTC CAAGGGAGGA GTAATGCCCTG GAGACCTCAG GTGAATAATC  
 4741 CAAAAGAGTG GCTGCAAGTG GACTTCCAGA AGACAATGAA AGTCACAGGA GTAACTACTC  
 4801 AGGGAGTAAAT ATCTCTGCTT ACCAGCATGT ATGTGAAGGA GTTCCCTCATC TCCAGCAGTC  
 4861 AAGATGGCCA TCAGTGGACT CTCTTTTTC AGAATGGCAA AGTAAAGTT TTTCAGGGAA  
 4921 ATCAAGACTC CTTCACACCT GTGGTGAACCT CTCTAGACCC ACCGTTACTG ACTCGCTACC  
 4981 TTGGAATTCA CCCCCAGAGT TGGGTGCACC AGATTGCCCT GAGGATGGAG GTTCTGGGCT  
 5041 GCGAGGCACA GGACCTCTAC

[0220]

\* 밀줄친 핵산은 시그널 펩타이드를 암호화한다.

[0221]

"B-도메인 기원한 FVIII"은 완전하거나 부분적인 결실을 지닐 수 있으며 이는 미국 특허 제 6,316,226호, 제 6,346,513호, 제7,041,635호, 제5,789,203호, 제6,060,447호, 제5,595,886호, 제6,228,620호, 제5,972,885호, 제6,048,720호, 제5,543,502호, 제5,610,278호, 제5,171,844호, 제5,112,950호, 제4,868,112호, 및 제 6,458,563호에 개재되어 있다. 일부 구현예에서, 본 발명의 B-도메인-결실된 FVIII 서열은 미국 특허 제 6,316,226호(또는 미국 특허 제6,346,513호)의 컬럼 4, 4행 내지 컬럼 5, 28행 및 실시예 1 내지 5에 개재되어 있다. 다른 구현예에서, B-도메인 결실된 인자 VIII은 S743/Q1638 B-도메인 결실된 인자 VIII(SQ BDD FVIII) (예를 들면, 744번 아미노산 내지 1637번 아미노산으로부터의 결실을 지닌 인자 VIII, 예를 들면, 서열 번호 16의 1 내지 743번 아미노산 및 1638번 내지 2332번 아미노산을 지닌 인자 VIII, 즉, 서열 번호 18)이다. 일부 구현예에서, 본 발명의 B-도메인-결실된 FVIII은 미국 특허 제5,789,203호(또는 제US 6,060,447호, 제US 5,595,886호, 및 제US 6,228,620호)의 컬럼 2, 26 내지 51행 및 실시예 5 내지 8에 개재된 결실을 갖는다. 일부 구현예에서, B-도메인 결실된 인자 VIII은 미국 특허 제5,972,885호의 컬럼 1, 25행 내지 컬럼 컬럼 2, 40행; 미국 특허 제6,048,720호의 컬럼 6, 1 내지 22행 및 실시예 1; 미국 특허 제5,543,502호의 컬럼 2, 17 내지 46행; 미국 특허 제5,171,844호의 컬럼 4, 22행 내지 컬럼 5, 36행; 미국 특허 제5,112,950호의 컬럼 2, 55 내지 68행, 도 2, 및 실시예 1; 미국 특허 제4,868,112호의 컬럼 2, 2행 내지 컬럼 19, 21행 및 표 2; 미국 특허 제7,041,635호의 컬럼 2, 1행 내지 컬럼 3, 19행, 컬럼 3, 40행 내지 컬럼 4, 67행, 컬럼 7, 43행 내지 컬럼 8, 26행, 및 컬럼 11, 5행 내지 컬럼 13, 39행; 또는 미국 특허 제6,458,563호의 컬럼 4, 25 내지 53행에 기술된 결실을 갖는다.

[0223]

일부 구현예에서, B-도메인 결실된 FVIII은 대부분의 B 도메인의 결실을 갖지만, 제WO 91/09122호에 개재된 바와 같이, 주요 해독 생성물의 2개의 폴리펩타이드 쇄내로의 생체내 단백질분해 프로세싱에 필수적인 B 도메인의 아미노-말단 서열을 여전히 함유한다. 일부 구현예에서, B-도메인-결실된 FVIII은 747 내지 1638번 아미노산의 결실, 즉, B 도메인의 실질적으로 완전한 결실로 작제된다. [참조: Hoeben R.C., 등 *J. Biol. Chem.* 265 (13): 7318-7323 (1990)]. B-도메인-결실된 인자 VIII은 FVIII의 771 내지 1666번 아미노산 또는 868 내지 1562번 아미노산의 결실을 함유할 수 있다. [참조: Meulien P., 등 *Protein Eng.* 2(4): 301-6 (1988)]. 본 발명의 일부 인 추가의 B 도메인 결실은 982 내지 1562번 또는 760 내지 1639번 아미노산의 결실[참조: Toole 등, *Proc. Natl. Acad. Sci. U.S.A.* (1986) 83, 5939-5942)], 797 내지 1562번 아미노산의 결실[참조: Eaton, 등 *Biochemistry* (1986) 25:8343-8347)], 741 내지 1646번 아미노산의 결실[참조: Kaufman (PCT 공개된 특허원 제 WO 87/04187호)], 747 내지 1560번 아미노산[참조: Sarver, 등, *DNA* (1987) 6:553-564)], 741 내지 1648번 아미노산[참조: Pasek (PCT 출원 제88/00831호)], 또는 816 내지 1598번 또는 741 내지 1648번[참조: Lagner (Behring Inst. Mitt. (1988) No 82:16-25, EP 295597)]의 결실을 포함한다. 다른 구현예에서, BDD FVIII은 하나 이상의 N-연결된 글리코시드화 부위를 보유하는 B-도메인의 단편, 예를 들면, 전장의 FVIII 서열의 아미노산 서열에 상응하는 757, 784, 828, 900, 963, 또는 임의로 943번 잔기를 함유하는 FVIII 폴리펩타이드를 포함한다. B-도메인 단편의 예는 문헌[참조: Miao, H.Z., 등, *Blood* 103(a): 3412-3419 (2004), Kasuda, A,

등, *J. Thromb. Haemost.* 6: 1352-1359 (2008), 및 Pipe, S.W., 등, *J. Thromb. Haemost.* 9: 2235-2242 (2011)]에 개재된 바와 같은 B-도메인의 226번 아미노산 또는 163번 아미노산 (즉, B 도메인의 첫번째 226개 아미노산 또는 163개 아미노산은 보유되어 있다)을 포함한다. 일부 구현예에서, 부분 B-도메인을 지닌 FVIII은 FVIII198이다. FVIII198은 일본쇄 FVIIIFc 분자-226N6을 함유하는 부분 B-도메인이다. 226은 FVIII B-도메인의 N-말단 226개 아미노산을 나타내고, N6은 B-도메인에서 N-글리코시드화 부위를 나타낸다. 여전히 다른 구현예에서, BDD FVIII은 309번 잔기에서 점 돌연변이(Phe로부터 Ser으로)를 추가로 포함함으로써 BDD FVIII 단백질의 발현을 증진시킨다. 참조: Miao, H.Z., 등, *Blood* 103(a): 3412-3419 (2004). 여전히 다른 구현예에서, BDD FVIII은 B-도메인의 부위를 함유하지만, 하나 이상의 푸린 절단부위 (예를 들면, Arg1313 및 Arg 1648)를 함유하지 않는 FVIII 폴리펩타이드를 포함한다. [참조: Pipe, S.W., 등, *J. Thromb. Haemost.* 9: 2235-2242 (2011)]. 앞서의 결실 각각은 어떠한 FVIII 서열에서도 이루어질 수 있다.

[0224]

본 발명에 유용한 FVIII 단백질은 FVIII 응고 활성에 영향을 미치지 않는, 하나 이상의 이종 서열 또는 이에 화학적 또는 물리적 변형을 지닌 FVIII을 포함할 수 있다. 이러한 이종 서열 또는 화학적 또는 물리적 변형은 FVIII 단백질의 C-말단 또는 N-말단에 융합되거나 FVIII 단백질의 2개의 아미노산 잔기 사이에 삽입될 수 있다. FVIII 단백질내 이러한 삽입은 FVIII 응고 활성 또는 FVIII 기능에 영향을 미치지 않는다. 하나의 구현예에서, 삽입은 FVIII 단백질의 약력학적 특성(예를 들면, 반감기)를 개선시킨다. 다른 구현예에서, 삽입은 2, 3, 4, 5, 또는 6개 이상의 부위일 수 있다.

[0225]

하나의 구현예에서, FVIII은 1648번 아미노산(전장의 인자 VIII 또는 서열 번호 16), 754번 아미노산 754(S743/Q1638 B-도메인 결실된 인자 VIII 또는 서열 번호 16)에서 아르기닌, 또는 상응하는 아르기닌 잔기(다른 변이체에서) 이후에 우측으로 절단됨으로써, 중쇄 및 경쇄를 생성한다. 다른 구현예에서, FVIII은 금속 이온-매개된 비-공유결합성 결합에 의해 연결되거나 연합된, 중쇄 및 경쇄를 포함한다.

[0226]

다른 구현예에서, FVIII은 1648번 아미노산(전장의 FVIII 또는 서열 번호 16), 754번 아미노산 754(S743/Q1638 B-도메인 결실된 인자 VIII 또는 서열 번호 18)에서 아르기닌, 또는 상응하는 아르기닌 잔기(다른 변이체에서) 이후에 우측으로 절단되지 않은 일본쇄 FVIII이다. 일본쇄 FVIII은 하나 이상의 아미노산 치환을 포함할 수 있다. 하나의 구현예에서, 아미노산 치환은 1648번 잔기, 1645번 잔기, 또는 전장의 성숙한 인자 VIII 폴리펩타이드(서열 번호 16) 둘 다 또는 754번 잔기, 751번 잔기, 또는 SQ BDD 인자 VIII(서열 번호 18) 둘 다에 상응하는 잔기에 있다. 아미노산 치환은 아르기닌 이외의 어떠한 아미노산, 예를 들면, 이소루이신, 루이신, 라이신, 메티오닌, 페닐알라닌, 트레오닌, 트립토판, 발린, 알라닌, 아스파라긴, 아스파르트산, 시스테인, 글루탐산, 글루타민, 글리신, 프롤린, 셀레노시스테인, 세린, 타이로신, 히스티딘, 오르니틴, 피롤리신, 또는 타우린일 수 있다.

[0227]

FVIII은 트롬빈으로 절단된 후 FVIIIa로서 활성화되어, 활성화된 인자 IX(FIXa)에 대한 보조인자로서 추가로 제공될 수 있다. 그리고 활성화된 FIX는 활성화된 FVIII와 함께 Xase 복합체를 형성하고 인자 X를 활성화된 인자 X(FXa)로 전환시킨다. 활성화를 위해, FVIII는 372, 740, 및 1689번(B-도메인 결실된 FVIII 서열에서 372, 740, 및 795번 아미노산에 상응) 아미노산에서 3개의 아르기닌 잔기 이후에 트롬빈에 의해 절단되며, 당해 절단은 50kDa A1, 43kDa A2, 및 73kDa A3-C1-C2 쇄를 갖는 FVIIIa를 생성한다. 하나의 구현예에서, 본 발명에 유용한 FVIII 단백질은 비-활성 FVIII이다. 다른 구현예에서, FVIII 단백질은 활성화된 FVIII이다.

[0228]

VWF 단백질에 연결되거나 연합된 FVIII 폴리펩타이드는 서열 번호 16 또는 서열 번호 18에 대해 적어도 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하며, 여기서 당해 서열은 FVIII 응고 활성을 지니는데, 예를 들면, 보조인자로서 인자 IX를 활성화시켜 인자 X를 활성화된 인자 X(FXa)로 전환시킨다.

[0229]

일부 구현예에서, FVIII 단백질은 FVIII 단백질의 C-말단 또는 N-말단에 융합된 또는 FVIII 단백질내 2개의 인접한 아미노산 사이에 삽입된 하나 이상의 이종 잔기를 추가로 포함한다. 다른 구현예에서, 이종 잔기는 적어도 약 50개 아미노산, 적어도 약 100개 아미노산, 적어도 약 150개 아미노산, 적어도 약 200개 아미노산, 적어도 약 250개 아미노산, 적어도 약 300개 아미노산, 적어도 약 350개 아미노산, 적어도 약 400개 아미노산, 적어도 약 450개 아미노산, 적어도 약 500개 아미노산, 적어도 약 550개 아미노산, 적어도 약 600개 아미노산, 적어도 약 650개 아미노산, 적어도 약 700개 아미노산, 적어도 약 750개 아미노산, 적어도 약 800개 아미노산, 적어도 약 850개 아미노산, 적어도 약 900개 아미노산, 적어도 약 950개 아미노산, 또는 적어도 약 1000개 아미노산의 아미노산 서열을 포함한다. 일부 구현예에서, 키메라 단백질의 반감기는 야생형 FVIII보다 적어도 약 1.5배, 적어도 약 2배, 적어도 약 2.5배, 적어도 약 3배, 적어도 약 4배, 적어도 약 5배, 적어도 약 6배, 적어도 약 7배, 적어도 약 8배, 적어도 약 9배, 적어도 약 10배, 적어도 약 11배, 적어도 약 12배 더 길게 연장된다.

[0230] 다른 예시적인 FVIII 변이체는 또한 2013년 1월 17일자로 발표된 미국 공보 제US2013/0017997호, 2013년 8월 22일자로 발표된 국제 공보 제WO 2013/122617호, 또는 2014년 1월 16일자로 발표된 국제 공보 제WO 2014/011819호, 또는 국제 공보 제WO2013123457 A1호, 또는 국제 출원 제PCT/US2013/049989호에 개재되어 있다.

### III. 폴리뉴클레오타이드, 벡터, 숙주 세포, 및 제조 방법

[0231] 또한 본원에서는 본원에 기술된 키메라 분자를 암호화하는 폴리뉴클레오타이드가 제공된다. VWF 단백질이 VWF 링커를 통해 FVIII 단백질 및 이종 잔기 및 단일 폴리펩타이드 쇄로서 키메라 단백질내 XTEN 서열에 연결되는 경우, 본 발명은 단일의 폴리펩타이드 쇄를 암호화하는 단일의 폴리뉴클레오타이드에 관한 것이다. 키메라 단백질이 제1 및 제2의 폴리펩타이드 쇄들을 포함하고, 상기 제1 폴리펩타이드 쇄는 VWF 링커를 통해 VWF 단백질, XTEN 서열, 및 제1의 이종 잔기(예를 들면, 제1의 Fc 영역)을 포함하며, 상기 제2의 폴리펩타이드 쇄는 FVIII 단백질 및 제2의 이종 잔기(예를 들면 제2의 Fc 영역)를 포함하는 경우, 폴리뉴클레오타이드는 제1의 뉴클레오타이드 영역 및 제2의 뉴클레오타이드 영역을 포함할 수 있다. 하나의 측면에서, 제1의 뉴클레오타이드 영역 및 제2의 뉴클레오타이드 영역은 동일한 폴리뉴클레오타이드 상에 있다. 다른 측면에서, 제1의 뉴클레오타이드 영역 및 제2의 뉴클레오타이드 영역은 2개의 상이한 폴리뉴클레오타이드(예를 들면, 상이한 벡터) 상에 있다. 특정의 구현예에서, 본 발명은 제1의 뉴클레오타이드 쇄 및 제2의 뉴클레오타이드 쇄를 포함하는 폴리뉴클레오타이드의 세트에 관한 것이며, 여기서 상기 제1의 뉴클레오타이드 쇄는 VWF 단백질, XTEN 서열, VWF 링커, 및 키메라 단백질의 이종 잔기를 암호화하고 제2의 뉴클레오타이드 쇄는 FVIII 단백질 및 제2의 이종 잔기를 암호화한다. 일부 구현예에서, 본 발명은 제1의 뉴클레오타이드 쇄 및 제2의 뉴클레오타이드 쇄의 세트에 관한 것이며, 여기서, 상기 제1의 뉴클레오타이드 쇄는 VWF 단백질, 및 키메라 단백질의 이종 잔기를 암호화하고 제2의 뉴클레오타이드 쇄는 FVIII 링커를 통해 제2의 이종 잔기에 융합된 FVIII 단백질을 암호화하고, 여기서 적어도 하나의 XTEN 서열은 키메라 단백질에 융합된다. 다른 구현예에서, 본 발명은 제1의 뉴클레오타이드 쇄 및 제2의 뉴클레오타이드 쇄를 포함하는 폴리뉴클레오타이드의 세트에 관한 것이며, 여기서 제1의 뉴클레오타이드 쇄는 VWF 단백질, VWF 링커, 및 키메라 단백질이 이종 잔기를 암호화하고, 제2의 뉴클레오타이드 쇄는 FVIII 단백질, FVIII 링커, 및 제2의 이종 잔기를 암호화하며, 여기서 적어도 하나의 XTEN 서열은 키메라 단백질에 융합된다.

[0233] 다른 구현예에서, 폴리뉴클레오타이드의 세트는 단백질 전환효소(protein convertase)를 암호화하는 추가의 뉴클레오타이드 쇄(예를 들면, 키메라 폴리펩타이드가 단일의 폴리뉴클레오타이드 쇄에 의해 암호화된 경우 제2의 뉴클레오타이드 쇄 또는 키메라 단백질이 제2의 폴리뉴클레오타이드 쇄에 의해 암호화된 경우 제3의 뉴클레오타이드 쇄)를 추가로 포함한다. 단백질 전환효소는 전구단백질 전환효소 서브티리신/켁신(kexing) 제5형(PCSK5 또는 PC5), 전구단백질 전환효소 서브틸리신/켁신 제7형(PCSK7 또는 PC5), 효모 Kex 2, 전구단백질 전환효소 서브틸리신/켁신 제3형(PACE 또는 PCSK3), 또는 2개 이상의 이의 조합으로부터 선택될 수 있다. 일부 구현예에서, 단백질 전환효소는 PACE, PC5, 또는 PC7이다. 구체적인 구현예에서, 단백질 전환효소는 PC5 또는 PC7이다. 본원에 참조로 포함된 국제 출원 제PCT/US2011/043568호를 참조한다. 다른 구현예에서, 단백질 전환효소는 PACE/푸린이다.

[0234] 특정의 구현예에서, 본 발명은 VWF 링커를 통해 제1의 이종 잔기에 융합된 VWF의 D' 도메인 및 D3 도메인을 포함하는 VWF 단백질을 암호화하는 제1의 뉴클레오타이드 서열, FVIII 단백질 및 제2의 이종 잔기를 암호화하는 제2의 뉴클레오타이드 서열, 및 VWF의 D1 도메인 및 D2 도메인을 암호화하는 제3의 뉴클레오타이드 서열을 포함하는 폴리뉴클레오타이드의 세트를 포함하며 여기서 XTEN 서열은 제1쇄 또는 제2 쇄 내에 존재한다. 당해 구현예에서, D1 도메인 및 D2 도메인은 별도로 발현되어(VWF 단백질의 D'D3 도메인에 연결되지 않는다) 적절한 이황화물 결합 형성 및 D'D3 도메인의 폴딩이 이루어지도록 한다. D1D2 도메인 발현은 시스 또는 트랜스일 수 있다.

[0235] 본원에 사용된 것으로서, 발현 벡터는 삽입된 암호화 서열의 전사 및 해독에 필요한 요소를 함유하는 어떠한 핵산 작제물도 말하거나, RNA 바이러스 벡터의 경우에, 적절한 숙주 세포 내로 도입되는 경우, 복제 및 해독에 필수적인 성분을 말한다. 발현 벡터는 플라스미드, 파아지미드, 바이러스, 및 이의 유도체를 포함할 수 있다.

[0236] 본 발명의 발현 벡터는 키메라 분자를 암호화하는 폴리뉴클레오타이드를 포함할 것이다.

[0237] 하나의 구현예에서, 키메라 분자에 대한 암호화 서열은 발현 조절 서열에 작동적으로 연결되어 있다. 본원에 사용된 것으로서, 2개의 핵산 서열은, 각각의 성분 핵산 서열이 이의 기능성을 보유하도록 허용하는 방식으로 공유 연결되는 경우 작동적으로 연결되어 있다. 암호화 서열 및 유전자 발현 조절 서열은, 이들이 유전자 발현 조절 서열의 영향 또는 제어하에 암호화 서열의 발현 또는 전사 및/또는 해독이 일어나도록 하는 방식으로 공유결합으로 연결되어 있는 경우 작동적으로 연결된 것으로 일컬어진다. 2개의 DNA 서열은, 5' 유전자 발현 서열내

프로모터의 도입이 암호화 서열의 전사를 생성하고 2개의 서열 사이의 연결의 특성이 (1) 골격-쉬프트(frame-shift) 돌연변이의 도입을 생성하지 않거나, (2) 프로모터 영역의 능력을 방해하여 암호화 서열의 전사를 지시하지 않거나, (3) 상응하는 RNA 전사체가 단백질로 해독되도록 하는 능력을 방해하지 않는 경우, 작동적으로 연결된 것으로 일컬어진다. 따라서, 유전자 발현 서열은, 당해 유전자 발현 서열이 이러한 암호화 핵산 서열의 전사에 영향을 미쳐 수득되는 전사체가 바람직한 단백질 또는 폴리펩타이드로 해독되는 경우 암호화 핵산 서열에 작동적으로 연결될 수 있다.

[0238]

본원에 사용된 것으로서, 유전자 발현 조절 서열은, 이것이 작동적으로 연결된 암호화 핵산의 효율적인 전사 및 해독을 촉진하는, 프로모터 서열 또는 프로모터-인핸서 조합과 같은 어떠한 조절 뉴클레오타이드 서열이다. 유전자 발현 조절 서열은, 예를 들면, 구성적 또는 유도성 프로모터와 같은 포유동물 또는 바이러스 프로모터일 수 있다. 구성적 포유동물 프로모터는 다음의 유전자에 대한 프로모터를 포함하나, 이에 한정되지 않는다: 하이폭산틴 포스포리보실 트랜스퍼라제(HPRT), 아데노신 테아미나제, 피루베이트 키나제, 베타-액틴 프로모터, 및 다른 구성적 프로모터. 진핵 세포내에서 구성적으로 기능하는 예시적인 바이러스 프로모터는, 예를 들면, 사이토메갈로바이러스(CMV), 시미안 바이러스(예를 들면, SV40), 파필로마 바이러스, 아데노바이러스, 사람 면역결핍성 바이러스(HIV), 로우스 사코마 바이러스(Rous sarcoma virus), 사이토메갈로바이러스, 몰로니 백혈병 바이러스의 긴 말단 반복체(LTR), 및 다른 레트로바이러스, 및 헤르페스 단성 바이러스의 티미딘 키나제 프로모터를 포함한다. 다른 구성적 프로모터는 당해 분야의 통상의 기술자에게 공지되어 있다. 본 발명의 유전자 발현서열로서 유용한 프로모터는 또한 유도성 프로모터를 포함한다. 유도성 프로모터는 유도제의 존재하에서 발현된다. 예를 들면, 메탈로티오네인 프로모터는 특정의 금속 이온의 존재하에서 전사 및 해독을 촉진하도록 유도된다. 다른 유도성 프로모터는 당해 분야의 통상의 기술자에게 공지되어 있다.

[0239]

일반적으로, 유전자 발현 조절 서열은 필요할 경우, 전사 및 해독 각각의 개시에 포함된 5' 비-전사 및 5' 비-해독 서열, 예를 들면, TATA 박스, 캡핑 서열, CAAT 서열 등을 포함할 것이다. 특히, 이러한 5' 비-전사 서열은 작동적으로 결합된 암호화 핵산의 전사 조절을 위한 프로모터 서열을 포함하는 프로모터 영역을 포함할 것이다. 유전자 발현 서열은 바람직하게는 인핸서 서열 또는 상부 활성인자 서열을 임의로 포함한다.

[0240]

바이러스 벡터는 다음의 바이러스로부터의 핵산 서열을 포함하나, 이에 한정되지 않는다: 레트로바이러스, 예를 들면, 몰로니 쥐 백혈병 바이러스, 하비 쥐 육종 바이러스(Harvey murine sarcoma virus), 쥐 유방 종양 바이러스, 및 로우스 육종 바이러스; 아데노-관련된 바이러스; SV40-형 바이러스; 폴리오마바이러스; 엡슈타인-바르 바이러스(Epstein-Barr virus); 파필로마 바이러스(papilloma virus); 헤르페스 바이러스(herpes virus); 박시니아 바이러스(vaccinia virus); 폴리오 바이러스(polio virus); 및 레트로바이러스와 같은 RNA 바이러스. 당해 분야에 잘 공지된 다른 바이러스도 용이하게 사용할 수 있다. 특정의 바이러스 벡터는 비-세포변성 진핵세포 바이러스를 기본으로 하며, 여기서 비-필수 유전자는 목적한 유전자로 대체된다. 비-세포변성 바이러스는 레트로바이러스를 포함하며, 이의 생명 주기는 게놈성 바이러스 RNA의 DNA로의 역 전사와 숙주세포 DNA내로 후속적인 전구바이러스 통합을 포함한다. 레트로바이러스는 사람 유전자 치료요법 시도용으로 입증되어 왔다. 복제-결핍성인(즉, 바람직한 단백질의 합성을 지시할 수 있지만, 감염성 입자의 제조는 불가능한) 레트로바이러스가 가장 유용하다. 이러한 유전적으로 변경된 레트로바이러스 발현 벡터는 생체내에서 유전자의 고-효율성 형질변환에 대해 일반적인 활용을 갖는다. 복제-결핍성 레트로바이러스를 생산하기 위한 표준 프로토콜(외인성 유전 물질의 플라스미드 내로의 혼입, 패키징 세포주의 플라스미드를 사용한 형질감염, 재조합체 레트로바이러스의 패키징 세포주에 의한 생산, 조직 배양 배지로부터 바이러스 입자의 수집, 및 바이러스 입자를 사용한 표적 세포의 감염)은 문헌[참조: Kriegler, M., Gene Transfer and Expression, A Laboratory Manual, W.H. Freeman Co., New York (1990) and Murry, E. J., Methods in Molecular Biology, Vol. 7, Humana Press, Inc., Clifton, N.J. (1991)]에 제공된다.

[0241]

하나의 구현예에서, 바이러스는 아데노-관련 바이러스, 이-본쇄(double-stranded) DNA 바이러스이다. 아데노-관련 바이러스는 복제-결핍성으로 가공될 수 있으며 광범위한 세포형 및 종을 감염시킬수 있다. 이는 또한 열 및 지질 용매 안전성; 다양한 계통의 세포내에서 고 형질도입 빈도; 조혈 세포의 포함; 및 수퍼감염(superinfection) 억제의 결여와 같은 장점을 지님으로써 다중 계열의 형질도입을 허용한다. 전하는 바에 따르면, 아데노-관련 바이러스는 부위-특이적인 방식으로 사람 세포 DNA내로 통합됨으로써, 삽입 돌연변이유발의 가능성 및 레트로바이러스 감염의 삽입된 유전자 발현 특징의 가변성을 최소화시킬 수 있다. 또한, 야생형 아데노-관련 바이러스 감염은 선택적인 압력의 부재하에서 100회 이상의 계대배양을 위한 조직 배양물 속에서 수반되면, 이는, 아데노바이러스-관련 바이러스 게놈 통합이 비교적 안정적인 현상임을 내포한다. 아데노-관련 바이러스는 또한 염색체외 양식으로 기능할 수 있다.

- [0242] 다른 벡터는 플라스미드 벡터를 포함한다. 플라스미드 벡터는 당해 분야에 집중적으로 기술되어 있으며, 당해 분야의 숙련가에게 잘-공지되어 있다[참조: 예를 들면, Sambrook 등, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989]. 지난 수년내에, 플라스미드 벡터는 숙주 세포 내에서 복제하고 숙주 세포 내로 통합하는 것이 불가능하므로 생체내에서 유전자를 세포로 전달하는데 특히 유리한 것으로 밝혀졌다. 그러나, 숙주 세포와 혼화성인 프로모터를 갖는 이를 플라스미드는 플라스미드내에서 작동적으로 암호화된 유전자로부터 웨타이드를 발현할 수 있다. 상업적인 공급업자로부터 이용 가능한 일부 일반적으로 사용된 플라스미드는 pBR322, pUC18, pUC19, 다양한 pcDNA 플라스미드, pRC/CMV, 다양한 pCMV 플라스미드, pSV40, 및 pBlueScript를 포함한다. 특수한 플라스미드의 추가의 예는 pcDNA3.1, 카탈로그 번호 V79020; pcDNA3.1/하이드로, 카탈로그 번호 V87020; pcDNA4/myc-His, 카탈로그 번호 V86320; 및 pBudCE4.1, 카탈로그 번호 V53220를 포함하며, 모두는 Invitrogen(캘리포니아주 칼스바드 소재)으로부터 시판된다. 다른 플라스미드는 당해 분야의 통상의 기술자에게 잘 공지되어 있다. 또한, 플라스미드는 DNA의 특수 단편을 제거하고/하거나 첨가하는 표준분자 생물학 기술을 사용하여 통상적으로 설계할 수 있다.
- [0243] 본 발명의 단백질을 생산하는데 사용될 수 있는 하나의 곤충 발현 시스템에서, 오토그라파칼라포니카 (*Autographa californica*) 핵 다한증 바이러스(polyhidrosis virus)(AcNPV)가 외부 유전자를 발현하기 위한 벡터로서 사용된다. 바이러스는 스포도프테라 프루기페르다(*Spodoptera frugiperda*) 세포에서 성장한다. 암호화 서열은 바이러스의 비-필수 영역(예를 들면, 폴리헤드론 유전자)내로 클로닝되어 ACNPV 프로모터(예를 들면, 폴리헤드론 프로모터)의 조절 하에 위치할 수 있다. 암호화 서열의 성공적인 삽입은 폴리헤드론 유전자의 불활성화 및 폐쇄되지 않은 재조합체바이러스(즉, 폴리헤드론 유전자에 의해 암호화된 단백질성 외피를 결여한 바이러스)의 생산을 초래할 것이다. 이후에, 이를 재조합체 바이러스를 사용하여 삽입된 유전자가 발현된 스포도프테라 프루기페르다(*Spodoptera frugiperda*)를 세포내에서 감염시킨다. (참조: 예를 들면, Smith 등 (1983) *J Virol* 46:584; 미국 특허 제 4,215,051호). 이러한 발현의 추가의 예는 문헌[참조: Ausubel 등, eds. (1989) Current Protocols in Molecular Biology, Vol. 2, Greene Publish. Assoc. & Wiley Interscience]에서 찾을 수 있다.
- [0244] 본 발명의 단백질을 발현시키는데 사용될 수 있는 다른 시스템은 글루타민 신테타제 유전자 발현시스템이며, 이는 "GS 발현 시스템"(제조원: 영국 버크셔 소재의 Lonza Biologics PLC)으로 언급된다. 당해 발현 시스템은 미국 특허 제5,981,216호에 상세히 기술되어 있다.
- [0245] 포유동물 숙주 세포에서, 다수의 바이러스계 발현 시스템을 이용할 수 있다. 아데노바이러스가 발현 벡터로서 사용되는 경우에, 암호화 서열은 아데노바이러스 전사/해독 조절 복합체, 예를 들면, 레이트 프로모터(late promoter) 및 3부분(tripartite) 리더 서열에 연결될 수 있다. 당해 키메라 유전자는 이후에 시험관내 또는 생체내 재조합체 의해 아데노바이러스 게놈내에서 삽입될 수 있다. 바이러스 게놈의 비-필수적인 영역(예를 들면, 영역 E1 또는 E3)내에서의 삽입은 감염된 숙주내에서 생존하고 웨타이드를 발현할 수 있는 재조합체 바이러스를 생성할 것이다. [참조: 예를 들면, Logan & Shenk (1984) *Proc Natl Acad Sci USA* 81:3655]. 대안적으로, 박시니아 7.5 K 프로모터를 사용할 수 있다. [참조: 예를 들면, Mackett 등 (1982) *Proc Natl Acad Sci USA* 79:7415; Mackett 등 (1984) *J Virol* 49:857; Panicali 등 (1982) *Proc Natl Acad Sci USA* 79:4927].
- [0246] 생산 효율을 증가시키기 위하여, 폴리뉴클레오타이드를 효소 절단 부위에 의해 분리된 본 발명의 단백질의 다중 단위를 암호화하도록 설계할 수 있다. 수득되는 폴리웨타이드는 폴리웨타이드 단위를 회수하기 위하여 절단될 수 있다(예를 들면, 적절한 효소를 사용한 처리에 의함). 이는 단일 프로모터에 의해 구동된 폴리웨타이드의 수율을 증가시킬 수 있다. 적절한 바이러스 발현 시스템내에서 사용되는 경우, mRNA에 의해 암호화된 각각의 폴리웨타이드의 해독은 전사체내에서; 예를 들면, 내부 리보소움 도입 부위, IRES에 의해 내부적으로 지시된다. 따라서, 폴리시스템론성 작제물은 궁극적으로 다수의 개개 폴리웨타이드의 해독을 지시하는, 단일의 거대한 폴리시스템론성 mRNA의 전사를 지시한다. 이러한 시도는 다중단백질의 생산 및 효소적 프로세싱을 제거하고 단일의 프로모터에 의해 구동된 폴리웨타이드의 수율을 유의적으로 증가시킬 수 있다.
- [0247] 형질전환에 사용된 벡터는 일반적으로 선택가능한 마커를 함유함으로써 형질전환체를 확인할 수 있을 것이다. 세균 시스템에서, 이는 암피실린 또는 가나마이신과 같은 항생제 내성 유전자를 포함할 수 있다. 배양된 포유동물 세포내에서 사용하기 위한 선택가능한 마커는 네오마이신, 하이드로마이신, 및 메토트렉세이트와 같은 약물에 대해 내성을 부여하는 유전자를 포함한다. 선택가능한 마커는 중폭가능하고 선택가능한 마커일 수 있다. 하나의 중폭가능하고 선택가능한 마커는 디하이드로플레이트 리덕타제(DHFR) 유전자이다. [참조: Simonsen C C 등 (1983) *Proc Natl Acad Sci USA* 80:2495-9]. 선택가능한 마커는 문헌[참조: Thilly (1986) Mammalian Cell Technology, Butterworth Publishers, Stoneham, Mass.]에서 고찰가능하며, 선택가능한 마커의 선택은 당해 분

야의 통상의 기술 수준 내에 있다.

- [0248] 선택가능한 마커는, 목적 유전자와 동시에 별도의 플라스미드 상에서 세포내로 도입될 수 있거나, 이들은 동일한 플라스미드에서 도입될 수 있다. 동일한 플라스미드에서, 선택가능한 마커 및 목적 유전자는 상이한 프로모터 또는 동일한 프로모터의 조절하에 있을 수 있으며, 상이한 프로모터 배열은 디스트론성 메시지를 생산한다. 이러한 유형의 작제물은 당해 분야에 공지되어 있다(예를 들면, 미국 특허 제 4,713,339호).
- [0249] 발현 벡터는 재조합적으로 생산된 단백질의 용이한 정제를 허용하는 태그(tag)를 암호화할 수 있다. 예는 벡터 pUR278(참조: Ruther 등 (1983) *EMBO J* 2:1791)를 포함하나, 이에 한정되지 않으며, 여기서 발현될 단백질에 대한 암호화 서열은 lac z 암호화 영역과 프레임내에서 벡터내로 연결됨으로써 태그화된 융합 단백질이 생산된다; pGEX 벡터를 사용하여 글루타티온 S-트랜스퍼라제(GST) 태그를 지닌 본 발명의 단백질을 발현할 수 있다. 이를 단백질은 일반적으로 가용성이며 글루타티온-아가로즈 비드에 대한 흡수에 이어 유리 글루타티온의 존재하면서 용출에 의해 세포로부터 용이하게 정제될 수 있다. 벡터는 정제 후 태그의 용이한 제거를 위해 절단 부위(트롬빈 또는 인자 Xa 프로테아제 또는 PreScission Protease<sup>TM</sup>(제조원: 뉴저지주 피팩 소재의 Pharmacia)를 포함한다.
- [0250] 발현 벡터 또는 이후의 벡터는 적합한 표적 세포내로 형질감염되거나 공-형질감염되며, 이들은 폴리펩타이드를 발현할 것이다. 당해 분야에 공지된 형질감염 기술은 인산칼슘 침전(참조: Wigler 등 (1978) *Cell* 14:725), 전기천공(electroporation)(참조: Neumann 등 (1982) *EMBO J* 1:841), 및 리포좀-계 시약을 포함하나, 이에 한정되지 않는다. 다양한 숙주-발현 벡터 시스템을 이용하여 원핵 및 진핵 세포 둘 다를 포함하는 본원에 기술된 단백질을 발현할 수 있다. 이들은 적절한 암호화 서열을 함유하는 재조합체 박테리오파아지 DNA 또는 플라스미드 DNA 발현 벡터로 형질감염된 세균(예를 들면, 이. 콜라이); 적절한 암호화 서열을 함유하는 재조합체 효모 또는 진균 발현 벡터로 형질전환된 효모 또는 섬유상 진균; 적절한 암호화 서열을 함유하는 재조합체 바이러스 발현 벡터(예를 들면, 바콜로바이러스)로 감염된 곤충 세포 시스템; 재조합체 바이러스 발현 벡터(예를 들면, 칼리플라워 모자이크 바이러스(cauliflower mosaic virus) 또는 토마토 모자이크 바이러스 또는 담배 모자이크 바이러스)로 감염되거나 적절한 암호화 서열을 함유하는 재조합체 플라스미드 발현 벡터(예를 들면, Ti 플라스미드)로 형질전환된 식물 세포 시스템; 또는 포유동물 세포(예를 들면, HEK 293, CHO, Cos, HeLa, HKB11, 및 BHK 세포)를 포함하는 동물 세포 시스템을 포함하나, 이에 한정되지 않는다.
- [0251] 하나의 구현예에서, 숙주 세포는 진핵 세포이다. 본원에 사용된 것으로서, 진핵 세포는 명확한 핵을 갖는 어떠한 동물 또는 식물 세포를 말한다. 동물의 진핵 세포는 척추동물, 예를 들면, 포유동물의 세포 및 무척추동물의 세포, 예를 들면, 곤충을 포함한다. 식물의 진핵 세포는 제한없이, 효모 세포를 구체적으로 포함할 수 있다. 진핵 세포는 원핵 세포, 예를 들면, 세균과 구별된다.
- [0252] 특정의 구현예에서, 진핵 세포는 포유동물 세포이다. 포유동물 세포는 포유동물로부터 기원한 특정 세포이다. 포유동물 세포는 구체적으로 포유동물 세포주를 포함하나, 이에 한정되지 않는다. 하나의 구현예에서, 포유동물 세포는 사람 세포이다. 다른 구현예에서, 포유동물 세포는 HEK 293 세포이며, 이는 사람 배아 신장 세포주이다. HEK 293 세포는 버지니아주 마나사스 소재의 아메리칸 타입 컬ച류 컬렉션(American Type Culture Collection)으로부터 CRL-1533로서 및 Invitrogen(캘리포니아주 칼스바드 소재)으로부터 293-H 세포(카탈로그 번호 11631-017) 또는 293-F 세포(카탈로그 번호 11625-019)로서 이용가능하다. 일부 구현예에서, 포유동물 세포는 PER.C6<sup>®</sup> 세포이며, 이는 망막으로부터 기원한 사람세포주이다. PER.C6<sup>®</sup> 세포는 Crucell(네덜란드 라이덴 소재)로부터 이용가능하다. 다른 구현예에서, 포유동물 세포는 차이니즈 햄스터 난소(CHO) 세포이다. CHO 세포는 버지니아주 마나사스 소재의 아메리칸 타입 컬ച류 컬렉션으로부터 이용가능하다 (예를 들면, CHO-K1; CCL-61). 여전히 다른 구현예에서, 포유동물 세포는 아기 햄스터 신장(BHK) 세포이다. BHK 세포는 버지니아주 마나사스 소재의 아메리칸 타입 컬ച류 컬렉션으로부터 이용가능하다 (예를 들면, CRL-1632). 일부 구현예에서, 포유동물 세포는 HKB11 세포이며, 이는 HEK293 세포 및 사람 B 세포주의 하이브리드 세포이다 [참조: Mei 등, *Mol. Biotechnol.* 34(2): 165-78 (2006)].
- [0253] 하나의 구현예에서, 본 발명의 VWF 단백질, VWF 링커, 이종 잔기(smoiety) 또는 키메라 단백질을 암호화하는 플라스미드는 선택가능한 마커, 예를 들면, 제오신 내성을 추가로 포함하며, 키메라 단백질의 생산을 위해 HEK 293 세포내로 형질감염된다.
- [0254] 여전히 다른 구현예에서, 형질감염된 세포는 안정하게 형질감염된다. 이를 세포는 당해 분야의 숙련가에게 공지된 통상의 기술을 사용하여, 안정한 세포로서 선택되고 유지될 수 있다.

- [0255] 단백질의 DNA 작제물을 함유하는 숙주 세포는 적절한 성장 배지 속에서 성장한다. 본원에 사용된 바와 같이, 용어 "적절한 성장 배지"는 세포의 성장에 요구되는 영양소를 함유하는 배지를 의미한다. 세포 성장에 요구되는 영양소는 탄소원, 질소원, 필수 아미노산, 비타민, 무기질, 및 성장인자를 포함할 수 있다. 임의로, 배지는 하나 이상의 선택 인자를 함유할 수 있다. 임의로, 배지는 소 송아지 혈청 또는 태아 송아지 혈청(FCS)을 함유할 수 있다. 하나의 구현예에서, 배지는 IgG를 실질적으로 함유하지 않는다. 성장 배지는 일반적으로 예를 들면, DNA 작제물 상에서 선택가능한 마커에 의해 보충되거나 DNA 작제물과 함께 동시-형질감염된 필수 영양소 속에서 약물 선택 또는 결핍증에 의해 DNA 작제물을 함유하는 세포에 대해 일반적으로 선택될 것이다. 배양된 포유동물 세포는 일반적으로 혈청-함유 또는 혈청-비함유 배지(예를 들면, MEM, DMEM, DMEM/F12) 속에서 성장한다. 하나의 구현예에서, 배지는 CD293(제조원: 캘리포니아주 칼스바드 소재의 Invitrogen)이다. 다른 구현예에서, 배지는 CD17(캘리포니아주 칼스바드 소재의 Invitrogen)이다. 사용된 특수 세포주에 적절한 배지의 선택은 당해 분야의 통상의 기술자의 수준 내에 있다.
- [0256] 본원에 기술된 바와 같은 키메라 분자의 2개의 폴리펩타이드 쇄를 동시-발현시키기 위하여, 숙주세포는 쇄 둘다의 발현을 허용하는 조건하에서 배양한다. 본원에 사용된 것으로서, 배양은 살아있는 세포를 시험관내에서 적어도 정의된 시간 동안 유지하는 것을 말한다. 유지는 살아있는 세포의 집단에 있어서의 증가일 수 있으나, 포함하는 것을 필요로 하진 않는다. 예를 들면, 배양물 속에서 유지된 세포는 집단 속에서 정적이지만, 바람직한 생성물, 예를들면, 재조합체 단백질 또는 재조합체 융합 단백질을 생산할 수 있다. 진핵 세포를 배양하는데 적합한 조건은 당해 분야에 잘 공지되어 있으며 배양 배지, 배지 보충물, 온도, pH, 산소 포화 등의 적절한 선택을 포함한다. 시판 목적을 위해서, 배양은 진탕기 플라스크, 롤러 병(roller bottle), 중공 섬유 생물반응기(hollow fiber bioreactor), 교반-탱크 생물반응기, 에어리프트 생물반응기(airlift bioreactor), 웨이브 생물반응기(Wave bioreactor), 및 기타를 포함하는 확장 규모의 시스템의 어떠한 다양한 유형의 사용을 포함할 수 있다.
- [0257] 세포 배양 조건은 또한 키메라 분자내 제1 쇄 및 제2 쇄의 연합을 허용하도록 선택된다. 키메라 분자의 발현을 허용하는 조건은 비타민 K의 공급원의 존재를 포함할 수 있다. 예를 들면, 하나의 구현예에서, 안정하게 형질감염된 HEK 293 세포는 4mM 글루타민이 보충된 CD293 배지(제조원: 캘리포니아주 칼스바드 소재의 Invitrogen) 또는 OptiCHO 배지(제조원: 캘리포니아주 칼스바드 소재의 Invitrogen) 속에서 배양된다.
- [0258] 하나의 측면에서, 본 발명은 a) 숙주 세포를 키메라 분자를 암호화하는 폴리뉴클레오타이드로 형질감염시키는 단계 및 b) 숙주 세포를 배양 배지 속에서 키메라 분자를 발현시키는데 적합한 조건 하에서 배양시키는 단계(여기서 키메라 분자가 발현된다)를 포함하는 키메라 분자의 발현, 제조, 또는 생산 방법에 관한 것이다.
- [0259] 추가의 구현예에서, 키메라 분자를 함유하는 단백질 생성물은 배지내로 분비된다. 배지는 세포로부터 분리되어, 농축되고, 여과된 후, 2개 또는 3개의 친화성 컬럼, 예를 들면, 단백질 A 컬럼 및 1개 또는 2개의 이온 교환 컬럼 위에 통과한다.
- [0260] 특정의 측면에서, 본 발명은 본원에 기술된 방법에 의해 생산된 키메라 폴리펩타이드에 관한 것이다.
- [0261] 시험관내 생산은 본 발명의 목적한 변경된 폴리펩타이드를 다량으로 제공하기 위한 규모-확장을 허용한다. 조직 배양 조건하에서 포유동물 세포 배양을 위한 기술은 당해 분야에 공지되어 있으며 예를 들면 에어리프트 반응기 속에서 또는 연속 교반기 반응기 속에서 균질한 혼탁액 배양물, 또는 예를 들면 동공 섬유, 미세캡슐 속에, 아가로즈 미세비드 또는 세라믹 카트리지 위에 고정시키거나 인트랩(entrap)시키는 것을 포함한다. 필요한 경우 및/또는 바람직하게는, 폴리펩타이드의 용액은 통상의 크로마토그래피 방법, 예를 들면, 이온-교환 크로마토그래피, 소수성 상호작용 크로마토그래피(HIC, DEAE-셀룰로즈 위에서의 크로마토그래피 또는 친화성 크로마토그래피)에 의해 정제될 수 있다.
- [0262] 본 발명은 또한 제1의 이종 잔기에 융합된 VWF 단백질 및 XTEN 서열 및 제2의 이종 잔기에 융합된 FVIII 단백질을 포함하는 키메라 FVIII 단백질의 FVIII 활성을 증진시키는 방법을 포함하며, 당해 방법은 VWF 링커를 VWF 단백질과 제1이 이종 잔기 사이에 삽입시킴을 포함하고, 여기서, VWF 링커는 (i) 인자 VIII(FVIII)로부터의 a2 영역; (ii) FVIII로부터의 a1 영역; (iii) FVIII로부터의 a3 영역; (iv) X-V-P-R(여기서 X는 지방족 아미노산이다)(서열 번호 3) 및 PAR1 액소사이트 상호작용 모티프; 또는 (v) 이의 어떠한 조합으로부터 선택된 폴리펩타이드를 포함한다. 일부 구현예에서, FVIII 활성은 aPTT 검정 또는 ROTEM 검정으로 측정한다.
- [0263] IV. 약제학적 조성물
- [0264] 본 발명의 키메라 분자를 함유하는 조성물은 적합한 약제학적으로 허용되는 담체를 함유할 수 있다. 예를 들면,

이들은 작용 부위로 직접 전달하도록 설계된 제제로 활성 화합물의 프로세싱을 촉진하는 부형제 및/또는 보조제를 함유할 수 있다.

[0265] 약제학적 조성물은 거환 주사(bolus injection)에 의한 비경구 투여(즉, 정맥내, 피하, 또는 근육내)용으로 제형화될 수 있다. 주사용 제형은 단위 용량형, 예를 들면, 첨가된 방부제가 들어있는 앰플 또는 다중투여량 용기로 제공될 수 있다. 조성물은 혼탁제, 액체, 또는 오일 또는 수성 비히클중 유제와 같은 형태를 취할 수 있으며 혼탁화제, 안정화제 및/또는 분산제와 같은 제형 제제를 함유한다. 대안적으로, 활성 성분은 적합한 비히클, 예를 들면, 발열질이 유리된 물과 함께 구성하기 위한 분말 형태일 수 있다.

[0266] 비경구 투여용으로 적합한 제형은 또한 수용성 형태, 예를 들면, 수용성 염 속의 활성 화합물의 수용액을 포함한다. 또한, 적절한 오일 주사 혼탁액으로서의 활성 화합물의 혼탁액을 투여할 수 있다. 적합한 친지성 용매 또는 비히클은 지방 오일, 예를 들면, 참깨 오일, 또는 합성 지방산 에스테르, 예를 들면, 올레아이트 또는 트리글리세라이드를 포함한다. 수성 주사 혼탁액은 예를 들면, 나트륨 카복시메틸 셀룰로즈, 소르비톨 및 텍스트란을 포함하는 혼탁액의 점도를 증가시키는 물질을 함유할 수 있다. 임의로, 혼탁액은 또한 안정화제를 함유할 수 있다. 리포좀은 또한 세포 또는 사이질 공간내로 본 발명의 분자를 봉입(encapsulate)하기 위해 사용될 수 있다. 예시적인 약제학적으로 허용되는 담체는 생리학적으로 혼용성인 용매, 분산 매질, 피복제, 항세균 및 항진균제, 등장성 및 흡수 지연제, 물, 염수, 인산염 완충된 염수, 텍스트로즈, 글리세롤, 에탄올 등이다. 일부 구현예에서, 조성물은 등장성 제제, 예를 들면, 슈가, 만니톨, 소르비톨과 같은 다가알코올, 또는 염화나트륨을 포함한다. 다른 구현예에서, 화합물은 습윤제와 같은 약제학적으로 허용되는 물질 또는 활성 성분의 반감기 또는 효능을 향상시키는 습윤제 또는 유화제, 방부제 또는 완충제와 같은 미량의 보조 물질을 포함한다.

[0267] 본 발명의 조성물은 예를 들면, 액체(예를 들면, 주사가능하고 주입가능한 액제), 분산제, 혼탁제, 반고체 및 고체 용량형을 포함하는 다양한 형태일 수 있다. 바람직한 형태는 투여 방식 및 치료학적 적용에 의존한다.

[0268] 조성물은 액제, 미세 유제, 혼탁제, 리포좀, 또는 고 약물 농도에 적합한 다른 주문된 구조로서 제형화될 수 있다. 멸균 주사가능한 액제는 요구량의 활성 성분을 필요한 경우 상기 나열한 성분 중 하나 또는 조합이 들어 있는 적절한 용매 속에 포함시킨 후, 여과된 멸균에 의해 제조할 수 있다. 일반적으로, 분산제는 활성 성분을 염기성 분산 매질 및 상기 나열된 것들로부터의 요구되는 다른 성분들을 함유하는 멸균 비히클 속에 혼입시켜 제조한다. 멸균 주사가능한 액제를 제조하기 위한 멸균 분말의 경우에, 바람직한 제조 방법은 활성 성분 및 앞서 멸균-여과된 용액으로부터의 추가의 바람직한 성분의 분말을 수득하는 진공 건조 및 동결-건조이다. 용액의 적절한 유동성은, 예를 들면, 레시틴과 같은 피복제를 사용하고, 분산제의 경우에 요구되는 입자 크기를 유지하며 표면활성제를 사용함으로써 유지할 수 있다. 주사가능한 조성물의 연장된 흡수는 조성물 속에 흡수를 지연시키는 제제, 예를 들면, 모노스테아레이트 염 및 젤라틴을 포함시킴으로써 가져올 수 있다.

[0269] 활성 성분은, 방출이 조절된 제형 또는 장치를 사용하여 제형화할 수 있다. 이러한 제형 및 장치의 예는 이식체, 경피 패취, 및 미세캡슐화된 전달 시스템을 포함한다. 생분해가능하고, 생적합한 중합체, 예를 들면, 에틸렌 비닐 아세테이트, 다가무수물, 폴리글리콜산, 폴리오르토에스테르, 및 폴리락트산을 사용할 수 있다. 이러한 제형 및 장치의 제조 방법은 당해 분야에 공지되어 있다(참조: 예를 들면, Sustained and Controlled Release Drug Delivery Systems, J. R. Robinson, ed., Marcel Dekker, Inc., New York, 1978).

[0270] 주사가능한 데포트(depot) 제형은 폴리락타이드-폴리글리콜라이드와 같은 생분해가능한 중합체속에 약물의 미세 봉입된 매트릭스를 형성함으로써 제조할 수 있다. 중합체에 대한 약물의 비, 및 사용된 중합체의 특성에 따라서, 약물 방출 속도가 조절될 수 있다. 다른 예시적인 생분해가능한 중합체는 포리오르토에스테르 및 다가무수물이다. 데포트 주사가능한 제형은 또한 약물을 리포좀 또는 미세봉입체 속에 인트랩시켜 제조할 수 있다.

[0271] 보충적인 활성 화합물을 조성물 내로 혼입할 수 있다. 하나의 구현예에서, 본 발명의 키메라 분자는 다른 응고 인자, 또는 이의 변이체, 단편, 유사체 또는 유도체와 함께 제형화된다. 예를 들면, 응고 인자는 인자 V, 인자 VII, 인자 VIII, 인자 IX, 인자 X, 인자 XI, 인자 XII, 인자 XIII, 프로트롬빈, 피브린, 폰 빌데브란트 인자 또는 재조합체 가용성 조직 인자(rsTF) 또는 앞서의 어떠한 것의 활성화된 형태를 포함하나, 이에 한정되지 않는다. 지혈제의 응고 인자는 또한 항-섬유소용해성 약물, 예를 들면, 앱실론-아미노-카프로산, 트란엑삼산을 포함할 수 있다.

[0272] 약물 요법은 최적의 바람직한 반응을 제공하도록 조절할 수 있다. 예를 들면, 단일의 거환이 투여될 수 있으며, 수개의 분할된 투여량을 시간에 걸쳐 투여할 수 있거나, 투여량을 치료학적 상태의 지혈에 의해 나타낸 바와 같이 비례적으로 감소시키거나 증가시킬 수 있다. 용량의 투여 및 균일성의 용이함을 위하여 비경구 조성물을 용

량 단위로 제형화하는 것이 유리하다. [참조: 예를 들면, Remington's Pharmaceutical Sciences (Mack Pub. Co., Easton, Pa. 1980)].

[0273] 활성 화합물 외에, 액체 용량형은 물, 에틸 알코올, 에틸 카보네이트, 에틸 아세테이트, 벤질 알코올, 벤질 벤조에이트, 프로필렌 글리콜, 1,3-부틸렌 글리콜, 디메틸포름아미드, 오일, 글리세롤, 테트라하이드로푸르푸릴 알코올, 폴리에틸렌 글리콜, 및 소르비탄의 지방산 에스테르와 같은 불활성 성분을 함유할 수 있다.

[0274] 적합한 약제학적 담체의 비-제한적 예는 또한 문헌[참조: Remington's Pharmaceutical Sciences by E. W. Martin.]에 기술되어 있다. 부형제의 일부 예는 전분, 글루코즈, 락토즈, 슈크로즈, 젤라틴, 맥아, 쌀, 밀가루, 백악, 실리카겔, 스테아르산나트륨, 글리세롤 모노스테아레이트, 활석, 염화나트륨, 탈지 분유, 글리세롤, 프로필렌, 글리콜, 물, 에탄올 등을 포함한다. 조성물은 또한 pH 완충 시약, 및 습윤제 또는 유화제를 함유할 수 있다.

[0275] 경구 투여용으로, 약제학적 조성물은 통상의 수단으로 제조된 정제 또는 캡슐제의 형태를 취할 수 있다. 조성물은 또한 액체, 예를 들면, 시럽제 또는 혼탁제로서 제조될 수 있다. 액체는 혼탁화제(예를 들면, 소르비톨 시럽, 셀룰로즈 유도체 또는 수소화된 식용 지방), 유화제(레시틴 또는 아카시아), 비-수성 비히클(예를 들면, 아몬드 오일, 오일성 에스테르, 에틸 알코올, 또는 분획화된 야채 오일), 및 방부제(예를 들면, 메틸 또는 프로필-p-하이드록시벤조에이트 또는 소르브산)을 포함할 수 있다. 제제는 또한 풍미제, 착색제 및 감미제를 포함할 수 있다. 대안적으로, 조성물은 물 또는 다른 적합한 비히클과 함께 구성하기 위한 무수 제품으로서 제공될 수 있다.

[0276] 볼내 투여용으로, 조성물은 통상의 프로토콜에 따라서 정제 또는 로젠지제의 형태를 취할 수 있다.

[0277] 흡입에 의한 투여를 위해, 본 발명에 따라 사용하기 위한 화합물은 부형제의 존재 또는 부재하에 분무된 에어로졸 또는 가압 팩 또는 네뷸라이저(nebulizer)로부터의 에어로졸 분무의 형태로, 임의로 추진제, 예를 들면, 디클로로디플루오로메탄, 디클로로테트라플루오로메탄, 이산화탄소 또는 다른 적합한 가스와 함께 편리하게 전달된다. 가압된 에어로졸의 경우에 용량 단위는 계량된 양을 전달하기 위해 밸브를 제공함으로써 측정할 수 있다. 예를 들면, 흡입기 또는 취입기에서 사용하기 위한 젤라틴의 캡슐 및 카트리지의 화합물의 분말 혼합물 및 락토즈 또는 전분과 같은 적합한 분말 기재를 함유하여 제형화될 수 있다.

[0278] 약제학적 조성물은 또한 예를 들면, 코코아 버터 또는 다른 글리세라이드와 같은 통상의좌제 기재를 함유하는, 좌제 또는 보류 관장제(retention enema)로서 직장 투여용으로 제형화될 수 있다.

## V. 유전자 치료요법

[0280] 본 발명의 키메라 분자는 생체내에서 포유동물 속에서, 예를 들면, 사람 환자속에서, 출혈 응고 질환, 관절염증, 근육 출혈, 구강 출혈, 근육내로의 출혈, 구강 출혈, 외상, 두부 외상, 위장 출혈, 두개내 출혈, 복부내 출혈, 흉내 출혈, 골절, 중추 신경계 출혈, 인두후강내 출혈, 복막후 공간내 출혈, 또는 장요근집(iliopsoas sheath)내 출혈로부터 선택된 출혈 질병 또는 질환의 치료가 치료학적으로 유리할 수 있는 유전자 치료요법 시도를 사용하여 생산할 수 있다. 하나의 구현예에서, 출혈 질병 또는 질환은 혈우병이다. 다른 구현예에서, 출혈 질병 또는 질환은 혈우병 A이다. 이는 적합한 발현 조절 서열에 작동적으로 연결된 적합한 키메라 분자를 암호화하는 핵산의 투여를 포함한다. 특정의 구현예에서, 이들 서열은 바이러스 벡터내로 혼입된다. 이러한 유전자 치료요법을 위해 적합한 바이러스 벡터는 아데노바이러스 벡터, 렌티바이러스벡터, 바클로바이러스 벡터, 엡슈타인 바르 바이러스 벡터, 파포바바이러스 벡터, 박시니아 바이러스 벡터, 헤르페스 단성 바이러스 벡터, 및 아데노 관련 바이러스(AAV) 벡터를 포함한다. 바이러스 벡터는 복제-결핍성 바이러스 벡터일 수 있다. 다른 구현예에서, 아데노바이러스 벡터는 이의 E1 유전자 또는 E3 유전자에 결실을 갖는다. 아데노바이러스 벡터를 사용하는 경우, 포유동물은 선택가능한 마커 유전자를 암호화하는 핵산에 노출되지 않을 수 있다. 다른 구현예에서, 서열은 당해 분야의 숙련가에게 공지된 비-바이러스 벡터내로 혼입될 수 있다.

## VI. 키메라 단백질을 사용하는 방법

[0282] 본 발명은 또한 본 발명의 키메라 분자를 사용하여 이를 필요로 하는 피검자에서 출혈 사건의 빈도 또는 정도를 감소시키는 방법을 제공한다. 예시적인 방법은 이를 필요로 하는 피검자에게 치료학적 유효량의 본 발명의 키메라 분자를 투여함을 포함한다. 다른 측면에서, 본 발명은 본 발명의 키메라 분자를 사용하여 이를 필요로 하는 피검자에서 출혈사건의 발생을 방지하는 방법을 포함한다. 다른 측면에서, 본 발명의 재조합체 단백질을 암호화하는 DNA를 포함하는 조성물은 이를 필요로 하는 피검자에게 투여될 수 있다. 본 발명의 특정의 측면에서, 본 발명의 키메라 분자를 발현하는 세포는 이를 필요로 하는 피검자에게 투여될 수 있다. 본 발명의 특정의 측면에

서, 약제학적 조성물은 (i) 키메라 분자, (ii) 키메라 분자를 암호화하는 분리된 핵산, (iii) 키메라 분자를 암호화하는 핵산을 포함하는 벡터, (iv) 키메라 분자를 암호화하는 분리된 핵산 및/또는 키메라 분자를 암호화하는 핵산을 포함하는 벡터를 포함하는 세포, 또는 (v) 이의 조합을 포함하며, 약제학적 조성물은 허용되는 부형제 또는 담체를 추가로 포함한다.

[0283] 출혈 사건은 출혈 응고 질환으로부터 유발되거나 기원할 수 있다. 혈액 응고 질환은 또한 응고 장애로 언급될 수 있다. 하나의 예에서, 현재의 기재내용의 약제학적 조성물로 치료될 수 있는 혈액 응고 질환은 혈우병 또는 폰 빌레블란드 질병(vWD)이다. 다른 예에서, 본 발명의 약제학적 조성물로 치료될 수 있는 혈액 응고 질환은 혈우병 A이다.

[0284] 일부 구현예에서, 출혈 상태와 관련된 출혈 유형은 관절혈증, 근육 출혈, 구강 출혈, 출혈, 근육내로의 출혈, 구강 출혈, 외상, 두부 외상, 위장 출혈, 두개내 출혈, 복부내 출혈, 흉내 출혈, 골절, 중추 신경계 출혈, 인두 후강내 출혈, 복막후 공간내 출혈, 또는 장요근집내 출혈, 또는 이의 어떠한 조합으로부터 선택된다.

[0285] 다른 구현예에서, 출혈 상태로 고생하는 피검자는 예를 들면, 외과적 예방 또는 수술 전후 관리를 포함하는 수술 치료가 요구된다. 하나의 예에서, 수술은 소수술 및 대수술로부터 선택된다. 예시적인 수술 과정은 발치술, 편도선 절제술, 서혜부 헤르니아 절제술, 윤활막 절제술, 개두술, 골접합술, 외상 수술, 두개골내부 수술, 복부내 수술, 흉내 수술, 관절 대체술(예를 들면, 슬관절 치환술, 고관절 대치술 등), 심장 수술, 및 제왕절개 수술을 포함한다.

[0286] 다른 예에서, 피검자는 인자 IX와 병존적으로 치료된다. 본 발명의 화합물은 FIXa를 활성화할 수 있으므로, 이를 사용하여 피검자에게 FIXa를 투여하기 전에 FIXa 폴리펩타이드를 예비-활성화시킬 수 있다.

[0287] 본 발명의 방법은 예방학적 치료 또는 주문형 치료(on-demand treatment)가 요구되는 피검자에서 실시될 수 있다.

[0288] 본 발명의 키메라 분자를 포함하는 약제학적 조성물은 예를 들면, 국소(예를 들면, 경피 또는 안구), 경구, 복내, 비강, 질내, 직장 또는 비경구 투여를 포함하는 어떠한 적절한 투여 방식을 위해 제형화될 수 있다.

[0289] 본원에 사용된 것으로서 용어 비경구는 피하, 피내, 정맥내(예를 들면, 정맥내), 근육내, 척수, 두개내, 척추강내, 안내, 눈주위, 안와내, 활막내 및 복강내 주사, 및 어떠한 유사한 주사 또는 주입 기술을 포함한다. 조성물은 예를 들면, 혼탁제, 유제, 지속된 방출 제형, 크림제, 젤제 또는 산체일 수 있다. 조성물은 통상의 결합제 및, 트리글리세라이드와 같은 담체와 함께 좌제로서 제형화될 수 있다.

[0290] 본 발명이 본원에 상세히 기술되어 있으므로, 본 발명을 단지 나열할 목적이며 본 발명을 한정하는 것으로 의도되지 않는, 본원에 포함된, 다음의 실시예를 참조로 본 발명이 보다 명확하게 이해될 것이다. 본원에 언급된 모든 특허 및 공보는 참조로 표현하여 포함된다.

[0291] 실시예

[0292] 실시예 전체에서, 달리 기술하지 않는 한 다음의 물질 및 방법이 사용되었다.

[0293] 물질 및 방법

[0294] 일반적으로, 본 발명의 실시는, 달리 나타내지 않는 한, 화학, 생물리학, 분자생물학, 재조합체 DNA 기술, 면역학(특히, 예를 들면, 항체 기술), 및 전기영동에서 표준 기술을 사용한다. [참조: 예를 들면, Sambrook, Fritsch and 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, Ed., Irl Pr (1996); Antibodies: A Laboratory Manual, Harlow 등, CS.H.L. Press, Pub. (1999); and Current Protocols in Molecular Biology, eds. Ausubel 등, John Wiley & Sons (1992)].

[0295] 실시예 1. 다양한 VWF 작제물의 트롬빈-매개된 D'D3 방출의 평가

[0296] 당해 실시예는 도 2에 언급한 다양한 VWF 작제물의 37°C에서 트롬빈-매개된 D'D3 방출의 역학을 평가한다. 바이오코어(Biocore) 실험을 VWF의 D'D3 도메인과 Fc 사이에 상이한 트롬빈 결단가능한 링커를 함유하는 VWF-Fc 작제물을 사용하여 수행하였다. 궁극적인 목표는 본원에 기술된 바와 같이 FVIII-VWF 이종이량체에 대한 VWF-Fc 트롬빈 분해로부터 수집된 정보를 적용하는 것이다. 모든 VWF-D'D3 작제물을 칩 위에서 수행하여 100 내지 700 RU 범위의 단백질의 포획 밀도를 달성하였다. VWF 작제물을 칩 위에서 포획한 후, 5U/ml의 트롬빈을 표면 위에

서 5분 동안 주사하였다. Fc는 침에 결합되어 남은 반면, 절단가능한 작제물내 D'D3는 방출된다. 속도(RU/s) 대포획 밀도(RU)를 도 3 및 4에 플롯팅하였다. 분해 속도는 포획 밀도와 비례하는 반면 각각의 작제물에 대해 트롬빈 절단에 대한 민감성의 척도를 제공하였다.

[0297] 도 3은, 예측된 것으로서 VWF-052(이는 링커 영역내에서 트롬빈 절단 부위를 가지지 않는다)가트롬빈에 의해 절단되지 않음을 나타낸다. VWF-039(PAR1 부위를 지닌 LVPR)는 FVIII 절단 속도와 비교가능하다(데이타는 나타내지 않음). 따라서, VWF-039는 Fc로부터 완전한 D'D3 방출에 대한 벤치 마크로서 제공되었다. VWF-039와 관련하여 다양한 VWF-Fc 농도의 기울기의 비율을 사용하여 트롬빈 절단의 효능을측정하였다. VWF-039(PAR1 부위를 갖는 LVPR)는 트롬빈으로 VWF-031(LVPR)보다 대략 70 내지 80배 빠르게 절단된다. VWF-51(ALRPRVV)은 VWF-031(LVPR)보다 1.8배 신속하게 절단된다. LVPR 부위를 따라 288 XTEN을 함유하는 VWF-034는 VWF-031과 비교하여 보다 느린 절단을 나타내었다.

[0298] VWF-Fc 작제물은 또한 링커 영역내에서 FVIII 단백질의 상이한 산성 영역(a1, a2 및 a3)을 도입함으로써 제조하였다. D'D3과 Fc 영역 사이에 a2 영역을 함유하는 VWF-055는 VWF-039 작제물과 유사한 트롬빈절단을 나타내었다. 도 4에 나타낸 바와 같이, VWF-054(a1 영역) 및 VWF-056(a3 영역)은 ~ 5-배 감소된 트롬빈절단을 나타내었다.

[0299] 도 5는 상이한 VWF 작제물에 대한 트롬빈 절단 곡선의 기울기 값을 나타낸다. 이들 결과로부터 FVIII의 산성 영역 2(a2)는 보다 높은 효율적인 트롬빈 절단 부위인 것으로 여겨지며 본원에 기술된 바와 같이 FVIII-VWF 이종이량체내에 혼입되었다.

[0300] 실시예 2. HemA 환자 전혈 ROTEM 검정을 사용한 FVIII/VWFD'D3 이종이량체의 지혈 효능의 평가

[0301] 상이한 트롬빈 절단가능한 링커를 함유하는 FVIII/VWFD'D3 이종이량체를 혈우병에서 이들의 효능에 대해 HemA 공여자 전혈 ROTEM(회전 트롬보엘라스토메트리(rotational thromboelastometry)) 검정에서 평가하였다. 전혈 시료를 심각한 혈우병 A 출혈 질환을 지닌 공여자로부터 항-응고제로서 시트르산나트륨과함께 수집하였다. 출혈 시료 수집 40분 후, 상이한 트롬빈 절단가능한 링커 - FVIII155/VWF031 (48aa, LVPR 부위), FVIII155/VWF039 (26aa, LVPR+PAR1 부위), FVIII155/VWF055 (34aa, FVIII로부터의 a2) 를 함유하는 FVIII/VWFD'D3 이종이량체 변이체를 전혈 시료 속으로 희석시켜 FVIII 색원체 검정에 의해 측정된 것으로서 정상의 100%, 30%, 10%, 및 3%에서의 최종 농도로 희석시켰다. FVIII/VWFD'D3 이종이량체의 첨가 직후, ROTEM 반응을 CaCl<sub>2</sub>를 첨가하여 개시하였다. 응고 시간(시험의 개시로부터 2mm 진폭에 이르는 시간)을 장치로 기록하고 시료중 FVIII 농도에 대해 플롯팅하였다(도 6). 보다 강력한 FVIII/VWFD'D3 이종이량체는 보다 신속한 응고 과정을 유도하며, 따라서 덜 강력한 FVIII/VWFD'D3 이종이량체와 비교하여 보다 짧은 응고 시간을 생성하는 것으로 가설화하였다. 도 6에 나타낸 바와 같이, FVIII/VWF039 이종이량체가 첨가된 시료는 시험한 모든 농도에서 보다 짧은 응고 시간을 가졌으며, FVIII/VWF031 이종이량체가 첨가된 시료는 모든 농도에서 가장 긴 응고 시간을 가졌다. FVIII155/VWF055 이종이량체가 첨가된 시료에 대한 응고 시간은 중간이다. 따라서, 지혈 효능의 순위는 FVIII155/VWF039 > FVIII155/VWF055 > FVIII155/VWF031이다. 3개 분자들 사이에 유일한 차이는 VWF 단백질과 Fc 영역 사이의 트롬빈 절단가능한 링커이므로, 결과는, LVPR 부위 및 PAR1 엑소사이트 상호작용 모티프 및 FVIII의 a2 영역을 함유하는 링커가 LVPR 부위 단독보다 더 잘 작동함을 나타낸다.

[0302] 실시예 3. FVIII/VWF 이종이량체의 활성의 평가

[0303] FVIII-XTEN/VWF 이종이량체 작제물을 HEK293F 세포내에서 3개의 플라스미드: 제2 발현 FVIII-XTEN-Fc, 제2 발현 VWF-XTEN-Fc 및 제3 발현 PACE를 사용하여 형질감염시켰다. 폴리에틸렌이민 (PEI) 표준 프로토콜을 형질감염에 사용하였고 형질감염 5일 후, 조직 배양 배지를 수거하였다. FVIII-VWF 이종이량체의 다양한 조합을 배지로부터 정제하였다. 정제된 단백질의 활성을 표준 프로토콜을 사용하여 색원체성(2 단계) 및 aPTT(1 단계) 응고 검정 둘 다에서 시험하였다. FVIII와 Fc 사이 또는 D'D3와 Fc 사이에 FVIII의 산성 영역 2 (a2)의 도입(도 7a 및 도 7에서 나타낸 바와 같음)은 도 7c에서 나타낸 바와 같이, FVIII-VWF 이종이량체의 aPTT 활성을 개선시켰다. 예를 들면, FVIII169/VWF059 이종이량체는 D'D3-Fc 링커 영역에서 a2 트롬빈 절단 부위를 가지며 도 7c에 나타낸 바와 같이 D'D3Fc 링커내 LVPR 트롬빈 부위를 함유하는, FVIII169/VWF057보다 더 우수한 aPTT 활성을 갖는다.

[0304] 유사하게, FVIII와 Fc 사이의 a2 영역의 혼입은 표 7B에 나타낸 바와 같이, FVIII286/VWF059와 FVIII286/VWF062 사이의 aPTT 비에 대한 개원된 색원체성에 의해 입증되는 바와 같이 이종이량체의 1단계 응고 활성을 증가시켰다.

표 7A

서열 번호	작제물	FVIII & Fc (aa) 사이의 링커 길이	링커내 트롬빈 부위
1	FVIII169	-	없음
2	FVIII286	32	FVIII-a2
서열 번호	작제물	D'D3 와 Fc(aa) 사이의 링커 길이	링커내 트롬빈 부위
1	VWF057	144AE XTEN+ 35+ LVPR	LVPR
2	VWF059	144AE XTEN+ 32	FVIII-a2
3	VWF062	144AE XTEN	없음

표 7b

작제물	색원체성/aPTT 비
FVIII169/VWF057	2.51
FVIII169/VWF059	1.67
FVIII169/VWF062	2.7
FVIII286/VWF059	0.69
FVIII286/VWF062	0.83

표 7c

작제물	크로모(Chromo) 특이 활성 (IU/pmol)	aPTT 특이 활성 (IU/pmol)
FVIII169/VWF057	1.60	0.65
FVIII169/VWF059	1.60	0.90
FVIII169/VWF062	0.87	0.32
FVIII286/VWF059	1.35	1.96
FVIII286/VWF062	1.08	1.33

[0305]

[0306]

실시예 4: HemA 마우스 꼬리 클립 출혈 모델에서 FVIII-XTEN-Fc/D'D3-XTEN-Fc 이종이량체의 급성 효능

[0307]

상이한 트롬빈 절단가능한 링커를 함유하는 이종이량체의 급성 효능을 HemA 마우스 꼬리 클립출혈 모델을 사용하여 평가하였다.

[0308]

8-12주령의 수컷 HemA 마우스를 5개의 처리 그룹으로 무작위처리하고, SQ BDD-FVIII, rFVIII169/VWF034, rFVIII169/VWF057, rFVIII169/VWF059 또는 비히를 용액의 단일 정맥내 투여로 각각 처리하였다. FVIII(50 내지 100%의 정상 FVIII 혈장 수준을 재구성하기 위하여)의 사건 치료를 모사하기 위하여, 선택된 FVIII 치료 투여량은 FVIII aPTT 활성에 의해 측정된 것으로서 75 IU/kg 이다. 당해 투여량 수준에서, 모든 시험하는 FVIII 변이체는 투여 후 5분째에 정상의 쥐 혈장 FVIII 활성의 약 70%를 재구성할 것이다.

[0309]

꼬리 클립 과정은 다음과 같이 수행하였다. 요약하면, 마우스를 50mg/kg의 케타민/0.5mg/kg의 텍스메데토미딘 칵테일로 꼬리 손상 전에 마취시키고 37°C의 가열 패드에 두어 체온을 유지시켰다. 마우스의 꼬리를 이후에 37°C의 염수 속에 10분 동안 침지시켜 측 정맥을 팽창시켰다. 정맥 회석 후, FVIII 변이체 또는 비히를 용액을 꼬리 정맥을 통해 주사하고 꼬리의 5mm 먼 곳을 투여 5분 후 직선 칼날 #11 메스(scalpel)를 사용하여 절단하였다. 출혈되는 혈액을 13ml의 37°C 염수내로 30분 동안 수집하고 혈액 손실 용적을 혈액수집 튜브의 중량 변화로 측정하였다: 혈액 손실 용적 = (수집 튜브 말기 중량 - 초기 중량 + 0.10) ml. 통계적 분석을 t 시험 [콜모고로브-그미로브 시험(Kolmogorov-Smirnov test)] 및 일월변량분석(one way ANOVA)[크루스칼-발리스 시험(KRUSKAL-Wallis test), 시험 후: 둔스 다중 비교 시험(Dunns multiple comparison test)]를 사용하여 수행하였다.

[0310]

연구에서 각각의 개별 동물로부터의 혈액 손실 용적을 도 8에 플롯팅하였다. 혈액 손실 용적에 있어서 유의적인 감소가 비히를 처리된 동물과 비교하여 모든 FVIII 치료 그룹에서 관찰되었다( $p < 0.05$ , 표 8). 유사한 혈액 손실 감소를 BDD-FVIII 처리와 비교하여 모든 이종이량체 처리 그룹으로부터 관찰되었으며( $p > 0.5$ , 표 8), 이는,

이종이량체 분자가 요구된 처리에서 SQ BDD-FVIII와 같이 강력하게 효율적일 수 있음을 제안한다.

[0311] 표 8: 콜모고로브-스미르노프 시험(Kolmogorov-Smirnov test)의 P 값

	FVIII169/VWF034	FVIII169/VWF057	FVIII169/VWF059
BDD-FVIII	0.7591	0.9883	0.5176
비[하]클	0.0006	0.0006	0.0266

[0312] [0313] pSYN VWF057 뉴클레오파이드 서열(링커내 LVPR 트롭빈 부위를 지닌 VWF D'D3-Fc)(서열 번호 79)

```

1 ATGATTCTTG CCAGATTTGC CGGGGTGCTG CTTGCTCTGG CCCTCATTTT
51 GCCAGGGACC CTTTGTGCAAG AAGGAACACTCG CGGCAGGTCA TCCACGGCCC
101 GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG
151 TACAGCTTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA
201 ACGCTCCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC
251 TCTCCGTGTA TCTTGGGGAA TTTTTTGACA TCCATTGTT TGCAATGGT
301 ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG
351 GCTGTATCTA GAAAATGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCCT
401 ATGGCTTTGT GGCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCCTGCTG
451 TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTTGCA ACTTTAACAT
501 CTTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC
551 CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT
601 GAACGGGCAT CTCCCTCCCAG CAGCTCATGC AACATCTCCT CTGGGAAAT
651 GCAGAAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT
701 TTGCCCGCTG CCACCCCTCTG GTGGACCCCG AGCCTTTGT GGCCCTGTGT
751 GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CCTGCCCTGC
801 CCTCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAAATG GTGCTGTACG
851 GCTGGACCGA CCACAGCGCG TGCAGCCCCAG TGTGCCCTGC TGCTATGGAG
901 TATAGGCAGT GTGTGTCCTCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT
951 CAATGAAATG TGTCAAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG
1001 GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGCTCCCTGC
1051 GTGCATTCCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG
1101 CAACACCTGC ATTGCGCAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT
1151 GTCCAGGGGA GTGCCCTTGTC ACTGGTCAAT CCCACTTCAA GAGCTTGAC
1201 AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGG
1251 TTGCCAGGAC CACTCCCTCT CCATTGTCAAT TGAGACTGTC CAGTGTGCTG
1301 ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGTCCG GCTGCCCTGG

```

[0314]

1351 CTGCACAAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA  
 1401 TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC  
 1451 ATACAGTGAC GGCCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG  
 1501 GACTGGGATG CCCGGGGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC  
 1551 CGGGAAGAAC TCGGGGCTGT GTGGGAATTAA CAATGGCAAC CAGGGCGACG  
 1601 ACTTCCTTAC CCCCTCTGGG CTGGCGGAGC CCCGGGTGGA GGACTTCGGG  
 1651 AACGCCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG  
 1701 CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGGTTCTCC GAGGAGGGGT  
 1751 GCGCGGTCCCT GACGTCCCCC ACATTCGAGG CCTGCCATCG TGCGTCAGC  
 1801 CCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACGTGTGCT CCTGCTCGGA  
 1851 CGGCCGCGAG TGCGTGTGCG GCGCCCTGGC CAGCTATGCC GCGCCCTGCG  
 1901 CGGGGAGAGG CGTGCAGCGTC GCGTGGCGCG AGCCAGGCCG CTGTGAGCTG  
 1951 AACTGCCCCA AAGGCCAGGT GTACCTGCGAG TGCGGGACCC CCTGCAACCT  
 2001 GACCTGCCGC TCTCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCTGCC  
 2051 TGGAGGGCTG CTTCTGCCCT CCAGGGCTCT ACATGGATGA GAGGGGGGAC  
 2101 TGCGTGCCTCA AGGCCCAAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA  
 2151 GCCAGAAAGAC ATCTTCTCAG ACCATCACAC CATGTGCTAC TGTGAGGATG  
 2201 GCTTCATGCA CTGTACCATG AGTGGAGTCC CGGGAAGCTT GCTGCCCTGAC  
 2251 GCTGTCCCTCA GCAGTCCCCC GTCTCATCGC AGCAAAAGGA GCCTATCCTG  
 2301 TCGGGCCCCC ATGTCAGC TGGCTGTCC CGCTGACAAC CTGCCGGCTG  
 2351 AAGGGCTCGA GTGTACCAAAC ACCTGCCAGA ACTATGACCT GGAGTGCATG  
 2401 AGCATGGGCT GTGTCTCTGG CTGCTCTGC CCCCGGGCA TGTCCTGGCA  
 2451 TGAGAACAGA TGCGTGGGCC TGGAAAGGTG TCCCTGCTTC CATCAGGGCA  
 2501 AGGAGTATGC CCCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC  
 2551 TGTCGGGACC CGAAGTGGAA CTGCACAGAC CATGTGTTG ATGCCACGTG  
 2601 CTCCACGATC GGCGATGGCCC ACTACCTCAC CTTCGACGGG CTCAAATACC  
 2651 TGTTCCCCGG GGAGTGCAG TACGTTCTGG TGCAGGATTAA CTGCCGGCAGT  
 2701 AACCTGGGA CCTTCGGAT CCTAGTGGGG AATAAGGGAT GCAGGCCACCC  
 2751 CTCAGTAAA TCGAAGAAC GGGTACCAT CCTGGTGGAG GGAGGAGAGA  
 2801 TTGAGCTGTT TGACGGGGAG GTGAATGTGA AGAGGGCCAT GAAGGATGAG  
 2851 ACTCACTTTG AGGTGGTGGA GTCTGGCCCG TACATCATTG TGCTGCTGG  
 2901 CAAAGCCCTC CCTGTGGTCT GGGACCGCCA CCTGAGCATC TCCGTGGTCC  
 2951 TGAAGCAGAC ATACCAAGGAG AAAGTGTGTC GCCTGTGTC GAATTGAT  
 3001 GGCATCCAGA ACAATGACCT CACCAAGCAG AACCTCCAAG TGGAGGAAGA  
 3051 CCCCTGTGGAC TTTGGGAACT CCTGGAAAGT GAGCTCGCAG TGTGCTGACA  
 3101 CCAGAAAAGT GCCTCTGGAC TCATCCCCCTG CCACCTGCCA TAACAACATC  
 3151 ATGAAGCAGA CGATGGTGGA TTCCCTCTGT AGAATCCTTA CCAGTGACGT  
 3201 CTTCCAGGAC TCCAACAAGC TGGTGACCC CGAGCCATAT CTGGATGTCT  
 3251 GCATTACGA CACCTGCTCC TGTGAGTCCA TTGGGACTG CGCCGCATTC  
 3301 TGCGACACCA TTGCTGCCCTA TGCCCCACGTG TGTGCCAGC ATGGCAAGGT  
 3351 GGTGACCTGG AGGACGGCCA CATTGTGCC CCAAGAGCTGC GAGGAGAGGA  
 3401 ATCTCCGGGA GAAACGGGTAT GAGGCTGAGT GGCCTATATAA CAGCTGTGCA  
 3451 CCTGCCGTGTC AAGTCACGTG TCAGCACCT GAGCCACTGG CCTGCCCTGT  
 3501 GCAGTGTGTG GAGGGCTGCC ATGCCACTG CCCTCCAGGG AAAATCCTGG  
 3551 ATGAGCTTTT GCAGACCTGC GTTGACCCCTG AAGACTGTCC AGTGTGTGAG  
 3601 GTGGCTGGCC GGGCTTTTGC CTCAGGAAAG AAAGTCACCT TGAATCCCA  
 3651 TGACCCCTGAG CACTGCCAGA TTGCCCCACTG TGATGTTGTC AACCTCACCT  
 3701 GTGAGGCTG CCAAGGAGCCG ATATCGGGCG CGCCACATC AGAGAGGCC  
 3751 ACCCTGAAA GTGGTCCCGG GAGCGAGCCA GCCACATCTG GGTCGGAAAC  
 3801 GCCAGGCACA AGTGAGTCTG CAACTCCCGA GTCCGGACCT GGCTCCGAGC  
 3851 CTGCCACTAG CGGCTCCCGAG ACTCCGGAA CTTCCGAGAG CGCTACACCA  
 3901 GAAAGCGGAC CGGAAACCAG TACCGAACCT AGCGAGGGCT CTGCTCCGG  
 3951 CAGCCCAGCC GGCTCTCCTA CATCCACGGA GGAGGGCACT TCCGAATCCG  
 4001 CCACCCCGGA GTCAAGGGCA GGATCTGAAC CCGCTACCTC AGGCAGTGAG  
 4051 ACGCAGGAA CGAGCGAGTC CGCTACACCG GAGAGTGGG CAGGGAGGCC  
 4101 TGCTGGATCT CCTACGTCCA CTGAGGAAGG GTCAACCGCG GGCTCGCCCA  
 4151 CCAGCACTGA AGAAGGTGCC TCGAGGGCG GTGGAGGATC CGGTGGCGG

4201 GGATCCGGTG CGGGGGGATC CGGTGGCGGG GGATCCGGTG GCAGGGGGATC  
 4251 CGGTGGCGGG GGATCCCTGG TCCCCCGGG CAGCGGAGGC GACAAAACCTC  
 4301 ACACATGCC ACCGTGCCA GCTCCAGAAC TCCTGGCGG ACCGTCACTC  
 4351 TTCCCTCTCC CCCAAAAACC CAAGGACACC CTCATGATCT CCCGGACCCC  
 4401 TGAGGTCAAC TGCGTGGTGG TGGACGTGAG CCACGAAGAC CCTGAGGTCA  
 4451 AGTTCAACTG GTACGTGGAC GGCCTGGAG TGCATAATGC CAAGACAAAG  
 4501 CGCGGGGAGG AGCAAGTACAA CAGCACGTG CGTGTGGTCA CGTCCTCAC  
 4551 CGTCTGAC CAGGACTGGC TGAATGGCAA GGAGTACAAG TGCAAGGTCT  
 4601 CCAACAAAGC CCTCCCAGCC CCCATCGAGA AAACCATCTC CAAAGCCAAA  
 4651 GGGCAGCCCC GAGAACCCACA GGTGTACACC CTGCCCCCAT CCCGGGATGA  
 4701 GCTGACCAAG AACCAAGGTCA GCCTGACCTG CCTGGTCAA GGCTTCTATC  
 4751 CCAGCGACAT CGCCGTGGAG TGGGAGAGCA ATGGGAGGCC GGAGAACAAAC  
 4801 TACAAGACCA CGCCTCCCGT GTTGGACTCC GACGGCTCCT TCTTCCCTCA  
 4851 CAGCAAGCTC ACCGTGGACA AGAGCAGGTG GCAGCAGGGG AACGTCTTCT  
 4901 CATGCTCCGT GATGCATGAG GCTCTGCACA ACCACTACAC CGAGAACAGC  
 4951 CTCTCCCTGT CTCCGGTAA ATGA

[0315]

[0316]

[0317]

pSYN VWF057 단백질 서열(링커내에 LVPR 트롭빈 부위를 지닌 VWF D'D3-Fc): 짙은 밀줄은 링커 영역을 함유하는

트롭빈 절단가능한 LVPR을 나타낸다(서열 번호 80)

1 MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNNTFDGSM  
 51 YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVLGE FFDIHLFVNG  
 101 TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL  
 151 SDRYFNKTCG LCGNPNIFAE DDFMTQEGL TSDPYDFANS WALSSGEQWC  
 201 ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC  
 251 EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME  
 301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC  
 351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD  
 401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG  
 451 LHNSLVKLKH GAGVAMDQD IQPLPLLKGDL RIQHTVTASV RLSYGEDLQM  
 501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG  
 551 NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS  
 601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGRVVAWREPGRC  
 651 NCPKGQVYLQ CGTPCNLTCT SLSYPDEECN EACLEGCFCP PGLYMDERGD  
 701 CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD  
 751 AVLSSPLSHR SKRSLSRPP MVKLVCPADN LRAEGLCTK TCQNYDLECM  
 801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV  
 851 CRDRKWNCND GVCATCSTI GMAHYLTFDG LKYLFPGECC YVLVQDYCGS  
 901 NPGTFRILVG NKGSCHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE  
 951 THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD  
 1001 GIQNNDLTTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI  
 1051 MKQTMVDSSC RILTSDFVQD CNKLVDPPEPY LDVCIYDTCS CESIGDCAAF  
 1101 CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY EAEWRYNSCA  
 1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE  
 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP ISGAPTSESA  
 1251 TPESGPGSEP ATSGSETPGT SESATPESGP GSEPATSGSE TPGTSESATP  
 1301 ESGPGTSTEP SEGSAPGSPA GSPTSTEET SESATPESGP GSEPATSGSE  
 1351 TPGTSESATP ESGPGSPAGS PTSTEEGSPA GSPTSTEEGA SSGGGGSGGG  
 1401 SSGGGGSGGG GSGGGGSILVP RGSGGDKTHT CPPCPAPELL  
 1451 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDEPEVKF NWYVDGVEVH  
 1501 NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT  
 1551 ISKAKGQPRE PQVYTLPPSR DELTKNQVSL TCLVKGFYPS DIAVEWESNG  
 1601 QOPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH  
 1651 YTQKSLSLSP GK\*

[0318]

pSYN VWF059 뉴클레오파이드 서열(링커내에 산성 영역 2(a2) 트롭빈 부위를 지닌 VWF D'D3-Fc)(서열 번호 81)

1 ATGATTCTG CCGAGATTG CGGGGTGC TG CTTGCTCTGG CCCTCATTT  
 51 GCCAGGGACC CTTTGTGCAG AAGGAACCTCG CGGCAGGTCA TCCACGGCCC  
 101 GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTTGA TGGGAGCATG  
 151 TACAGCTTG CGGGATACTG CAGTTACCTC CTGGCAGGGG GCTGCCAGAA  
 201 ACGCTCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC  
 251 TCTCGTGTGA TCTTGGGAA TTTTTGACA TCCATTGTT TGTCAATGGT  
 301 ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG  
 351 GCTGTATCTA GAAACTGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGC  
 401 ATGGCTTTGT GGGCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG  
 451 TCAGACAGAT ACTTCAACAA GACCTGCGGG CTGTGTGGCA ACTTTAACAT  
 501 CTTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC  
 551 CTTATGACTT TGCCAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT  
 601 GAACGGGCAT CTCCTCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT  
 651 GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT  
 701 TTGCCCGCTG CCACCCCTTG GTGGACCCCG AGCCTTTGT GGCCCTGTGT  
 751 GAGAAGACTT TGTGTGAGTG TGCTGGGGGG CTGGAGTGCG CCTGCCCTG  
 801 CCTCCTGGAG TACGCCCGGA CCTGTGCCCA GGAGGGAAATG GTGCTGTACG  
 851 GCTGGACCGA CCACAGCGCG TGCAGGCCAG TGTGCCCTGC TGGTATGGAG  
 901 TATAGGCAGT GTGTGTCCCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT  
 951 CAATGAAATG TGTCAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG  
 1001 GACAGCTCT GGATGAAGGC CTCTGCGTGG AGAGCACCGA GTGCTCCCTG  
 1051 GTGATTCTCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG  
 1101 CAACACCTGC ATTGCGCAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT  
 1151 GTCCAGGGGA GTGCCCTGTG ACTGGTCAAT CCCACTTCAA GAGCTTTGAC  
 1201 AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGG  
 1251 TTGCCAGGAC CACTCCCTCT CCATTGTCA TGAGACTGTC CAGTGTGCTG  
 1301 ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGTCCG GCTGCCCTGGC  
 1351 CTGCACAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA  
 1401 TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC  
 1451 ATACAGTGCAG GGCCTCCGTG CGCCTCAGCT ACAGGGGAGGA CCTGCAGATG  
 1501 GACTGGGATG GCGCGGGGAG GCTGCTGGTG AAGCTGTCCC CCGTCTATGC  
 1551 CGGGAAGACC TGCAGGCTGT GTGGGAATTAA CAATGGCAAC CAGGGCGACG  
 1601 ACTTCCCTTAC CCCCTCTGGG CTGGCGGAGC CCCGGGTGGA GGACTTCGGG  
 1651 AACGCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG  
 1701 CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGTTCTCC GAGGAGGCGT  
 1751 GCGCGGTCTT GACGTCCCCC ACATTGAGG CCTGCCATCG TGCGTCAGC  
 1801 CCGCTGCCCT ACCTGCGGAA CTGCCGCTAC GACGTGTGCT CCTGCTCGGA  
 1851 CGGGCGCAG TGCCCTGTGCG GCGCCCTGGC CAGCTATGCC GCGGCCTGCG  
 1901 CGGGGAGAGG CGTGCAGCGT GCGTGGCGCG AGCCAGGCCG CTGTGAGCTG  
 1951 AACTGCCCGA AAGGCCAGGT GTACCTGCAG TGCGGGACCC CCTGCAACCT  
 2001 GACCTGCCGC TCTCTCTCTT ACCGGATGA GGAATGCAAT GAGGCCTGCC  
 2051 TGGAGGGCTG CTTCTGCCCT CCAGGGCTCT ACATGGATGA GAGGGGGGAC  
 2101 TGCGTGCCCA AGGCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA  
 2151 GCCAGAAAGAC ATCTTCTCA ACCATCACAC CATGTGCTAC TGTGAGGATG  
 2201 GCTTCATGCA CTGTACCATG AGTGGAGTCC CGCGAAGCTT GCTGCCCTGAC  
 2251 GCTGTCTCTCA GCAGTCCCCCT GTCTCATCGC AGCAAAAGGA GCCTATCCCTG  
 2301 TCGGCCCCCCC ATGGTCAAGC TGGTGTGTCC CGCTGACAAC CTGCGGGCTG  
 2351 AAGGGCTCGA GTGTACCAAAC ACAGTGCCAGA ACTATGACCT GGAGTGCATG  
 2401 AGCATGGGCT GTGTCTCTGG CTGCCCTGCG CCCCCGGGCA TGGTCCGGCA  
 2451 TGAGAACAGA TGTTGGCCC TGGAAAGGTG TCCCTGCTTC CATCAGGGCA  
 2501 AGGAGTATGC CCCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC  
 2551 TGTCGGGACC GGAAGTGGAA CTGCACAGAC CATGTGTGTG ATGCCACGTG  
 2601 CTCCACGATC GGCATGGCCC ACTACCTCAC CTTCGACGGG CTCAAATACC  
 2651 TGTTCCCCGG GGAGTGCAG TACGTTCTGG TGCAGGATTA CTGCGGCAGT  
 2701 AACCCCTGGGA CCTTTCGGAT CCTAGTGGGG AATAAGGGAT GCAGCCACCC

[0320]

2751 CTCAGTGAAA TGCAAGAAC GGGTCACCCT CCTGGTGGAG GGAGGAGAGA  
 2801 TTGAGCTTT TGACGGGAG GTGAATGTGA AGAGGCCAT GAAGGATGAG  
 2851 ACTCACTTGC AGGTGGTGA GTCTGCCGG TACATCATTC TGCTGCTGG  
 2901 CAAAGCCCTC TCCGTGGTCT GGGACCGCCA CCTGAGCATE TCCGTGGTCC  
 2951 TGAAGCAGAC ATACCAGGAG AAAGTGTGTG GCCTGTGTG GAATTGGAT  
 3001 GGCATCCAGA ACAATGACT CACCAGCAG AACCTCCAAG TGGAGGAAGA  
 3051 CCCGTGGAC TTGGAAACT CCTGAAAGT GAGCTCGAG TGTGCTGACA  
 3101 CCAGAAAAGT GCCTCTGGAC TCATCCCCTG CCACCTGCCA TAACAACATC  
 3151 ATGAAGCAGA CGATGGTGA TTCCTCTGT AGAATCTTA CCAGTGACGT  
 3201 CTTCCAGGAC TGCAACAAGC TGGTGACCC CGAGCCATAT CTGGATGTCT  
 3251 GCATTTACGA CACCTGCTCC TGTGAGTCCA TTGGGACTG CGCCGCATT  
 3301 TGGCACACCA TTGCTGCCA TGCCCCACGT TGTCGCCAGC ATGGCAAGGT  
 3351 GGTGACCTGG AGGACGGCCA CATTGTGCC CCAAGAGCTGC GAGGAGAGGA  
 3401 ATCTCCGGGA GAACGGGTAT GAGGCTGAGT GGGCCTATAA CAGCTGTGCA  
 3451 CCTGCCTGTC AAGTCACGTG TCAGCACCC GAGCCACTGG CCTGCCCTGT  
 3501 GCAGTGTGTG GAGGGCTGCC ATGCCCACTG CCCTCCAGGG AAAATCCTGG  
 3551 ATGAGCTTT GCAGACCTGC GTTGACCCCTG AAGACTGTCC AGTGTGTGAG  
 3601 GTGGCTGGCC GGCGTTTGC CTCAGGAAAG AAAGTCACCT TGAATCCAG  
 3651 TGACCCCTGAG CACTGCCAGA TTTGCCACTG TGATGTTGTC AACCTCACCT  
 3701 GTGAAGCCTG CCAGGAGCCG ATATCGGGCG CGCCAACATC AGAGAGGCC  
 3751 ACCCCTGAAA GTGGTCCCGG GAGCGAGCCA GCCACATCTG GGTGGAAAC  
 3801 GCCAGGCACA AGTGAGTCTG CAACTCCCAG GTCGGGACCT GGCTCCGAGC  
 3851 CTGCCACTAG CGGCTCCAGG ACTCCGGAA CTTCGGAGAG CGCTACACCA  
 3901 GAAAGCAGGAC CGCGAACCCAG TACCGAACCT AGCGAGGGCT CTGCTCCGG  
 3951 CAGCCCCAGCC GGCTCTCTA CATCCACGGA GGAGGGCACT TCCGAATCCG  
 4001 CCACCCCGGA GTCAGGGCCA GGATCTGAAC CCGCTACCTC AGGCAGTGAG  
 4051 ACGCCAGGAA CGAGCGAGTC CGCTACACCG GAGAGTGGGC CAGGGAGCCC  
 4101 TGCTGGATCT CCTACGTCCA CTGAGGAAGG GTCAACCAGCG GGCTCGCCA  
 4151 CCAGCACTGA AGAAGGTGCC TCGATATCTG ACAAGAACAC TGGTGATTAT  
 4201 TAGGAGGACA GTTATGAAGA TATTCAGCA TACTTGCTGA GTAAAAACAA  
 4251 TGCCATTGAA CCAAGAAGCT TCTCTGACAA AACTCACACA TGCCCACCGT  
 4301 GCCCAGCTCC AGAACTCCTG GGCGGACCGT CAGTCTTCCT CTTCCCCCA  
 4351 AAACCCAAGG ACACCCCTCAT GATCTCCCGG ACCCTGAGG TCACATGCGT  
 4401 GGTGGTGGAC GTGAGGCCAG AAGACCTGA GGTCAAGTTT AACTGGTACG  
 4451 TGGACGGCGT GGAGGTGCAT AATGCCAAGA CAAAGCCGCG GGAGGAGCAG  
 4501 TACAACAGCA CGTACCGTGT GGTCAAGCGT CTCACCGTCC TGCACCAGGA  
 4551 CTGGCTGAAT GGCAAGGAGT ACAAGTGCAA GGTCTCCAAC AAAGCCCTCC  
 4601 CAGCCCCCAT CGAGAAAACC ATCTCCAAAG CCAAAGGGCA GCCCCGAGAA  
 4651 CCACAGGTGT ACACCCCTGG CCCATCCCAG GATGAGCTGA CCAAGAACCA  
 4701 GGTCAAGCGT ACCTGCCTGG TCAAAGGCTT CTATCCCAGC GACATGCCG  
 4751 TGGAGTGGGA GAGCAATGGG CAGCCGGAGA ACAACTACAA GACCACGCC  
 4801 CCCGTGTTGG ACTCCGACGG CTCTCTCTC CTCTACAGCA AGCTCACCGT  
 4851 GGACAAGAGC AGGTGGCAGC AGGGGAACGT CTTCTCATGC TCCGTGATGC  
 4901 ATGAGGCTCT GCACAACAC TACACCGAGA AGAGCCTCTC CCTGTCTCCG  
 4951 GGTAATGA

[0321]

pSYN WVF059 단백질 서열(링커내에 LVPR 트롭린 부위를 지닌 WVF D'D3-Fc)-굵은 밑줄부위는 a2 영역(SEQ\_ID NO: 82)을 나타낸다.

1 MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNTFDGSM  
 51 YSFAGYCSYL LAGGCQKRSF SIIGDFQNGK RVSLSVYLGE FFDIHLFVNG  
 101 TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL  
 151 SDRYFNKTCG LCGFNIFAE DDFMTQEGLT TSDPYDFANS WALSSGEQWC  
 201 ERASPPSSSC NISSGEMQKG LWEQCQLLKS TSVFARCHPL VDPEPFVALC  
 251 EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME  
 301 YRQCVPSCAR TCQSLHINEM CQERCVDGCS CPEGQELDEG LCVESTECPC  
 351 VHSGKRYPSEG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD  
 401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRDAVC TRSVTVRLPG

[0323]

451 LHNSLVKLKH GAGVAMDQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM  
 501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG  
 551 NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS  
 601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCEL  
 651 NCPKGQVYLQ CGTPCNLTCA SLSYDEECN EACLEGCFCP PGLYMDERGD  
 701 CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD  
 751 AVLSSPLSHR SKRSLSCRPP MVKLVCPADN LRAEGLECKT TCQNYDLECM  
 801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV  
 851 CRDRKWNCND HVCDATCSTI GMAHYLTFDG LKYLFPGEQ YVLVQDYCGS  
 901 NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE  
 951 THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD  
 1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI  
 1051 MKQTMVDSSC RILTSDFQD CNKLVDPEPY LDVCIYDTCS CESIGDCAAF  
 1101 CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY EAEWRYNSCA  
 1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KIILDELLQTC VDPEDCPVCE  
 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP ISGAPTSESA  
 1251 TPESPGPSEP ATSGSETPGT SESATPESGP GSEPATSGSE TPGTSESATP  
 1301 ESGPGTSTEP SEGSAPGSPA GSPTSTEEGT SESATPESGP GSEPATSGSE  
 1351 TPGTSESATP ESGPGSPAGS PTSTEEGSPA GSPTSTEEGA SISDKNTGDY  
 1401 YEDESYEDISA YLLSKNNAIE PRSFSDKTHT CPPCPAPELL GGPSVFLFPP  
 1451 KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ  
 1501 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQP  
 1551 PQVYTLPPSR DELTKNQVSL TCLVKGFYPS DIAWEWESNG QPENNYKTTP  
 1601 PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP  
 1651 GK\*

[0324]

pSYN VWF062 뉴클레오티드 서열(링커내 트롭빈 부위가 없는 VWF D'D3-Fc)(서열 번호 83)

1 ATGATTCCCTG CCAGATTTCG CGGGGTGCTG CTTGCTCTGG CCCTCATTGG  
 51 GCCAGGGACC CTTTGTGCAAG AAGGAACCTCG CGGCAGGTCA TCCACGGCCC  
 101 GATGCAGCCT TTTCGGAAGT GACTTCGTCA ACACCTTGA TGGGAGCATG  
 151 TACAGCTTTG CGGGATACTG CAGTTACCTC CTGGCAGGG GCTGCCAGAA  
 201 ACGCTCCCTTC TCGATTATTG GGGACTTCCA GAATGGCAAG AGAGTGAGCC  
 251 TCTCCGTGTA TCTTGGGAA TTTTTTGACA TCCATTGTT TGTCATGGT  
 301 ACCGTGACAC AGGGGGACCA AAGAGTCTCC ATGCCCTATG CCTCCAAAGG  
 351 GCTGTATCTA GAAAATGAGG CTGGGTACTA CAAGCTGTCC GGTGAGGCCT  
 401 ATGGCTTTGT GGCCAGGATC GATGGCAGCG GCAACTTTCA AGTCCTGCTG  
 451 TCAGACAGAT ACTTCAACAA GACCTGGCGG CTGTGTGGCA ACTTTAACAT  
 501 CTTGCTGAA GATGACTTTA TGACCCAAGA AGGGACCTTG ACCTCGGACC  
 551 CTTATGACTT TGCGAAACTCA TGGGCTCTGA GCAGTGGAGA ACAGTGGTGT  
 601 GAACGGGCAT CTCCCTCCAG CAGCTCATGC AACATCTCCT CTGGGGAAAT  
 651 GCAGAAGGGC CTGTGGGAGC AGTGCCAGCT TCTGAAGAGC ACCTCGGTGT  
 701 TTGCCCGCTG CCACCCCTTG GTGGACCCCG AGCCTTTGT GGCCCTGTGT  
 751 GAGAAGACTT TGTGTGAGTG TGCTGGGGG CTGGAGTGCG CCTGCCCTGC  
 801 CCTCCCTGGAG TACGCCCGGA CCTGTGCCA GGAGGAATG GTGCTGTACG  
 851 GCTGGACCGA CCACAGCGCG TGCAGCCAG TGTGCCCTGC TGGTATGGAG  
 901 TATAGGCAGT GTGTGTCCCC TTGCGCCAGG ACCTGCCAGA GCCTGCACAT  
 951 CAATGAAATG TGTCAAGGAGC GATGCGTGGA TGGCTGCAGC TGCCCTGAGG  
 1001 GACAGCTCCT GGATGAAGGC CTCTGCGTGG AGACCAACCGA GTGCTCCCTGC  
 1051 GTGCAATTCCG GAAAGCGCTA CCCTCCCGGC ACCTCCCTCT CTCGAGACTG  
 1101 CAACACCTGC ATTGCGCAA ACAGCCAGTG GATCTGCAGC AATGAAGAAT  
 1151 GTCCAGGGGA GTGCCCTGTC ACTGGTCAAT CCCACTCAA GAGCTTTGAC  
 1201 AACAGATACT TCACCTTCAG TGGGATCTGC CAGTACCTGC TGGCCCGGG  
 1251 TTGCCAGGAC CACTCCTTCT CCATTGTCAT TGAGACTGTC CAGTGTGCTG  
 1301 ATGACCGCGA CGCTGTGTGC ACCCGCTCCG TCACCGCTCCG GCTGCCTGGC  
 1351 CTGCACAAACA GCCTTGTGAA ACTGAAGCAT GGGGCAGGAG TTGCCATGGA

[0326]

1401 TGGCCAGGAC ATCCAGCTCC CCCTCCTGAA AGGTGACCTC CGCATCCAGC  
 1451 ATACAGTGAC GGCCTCCGTG CGCCTCAGCT ACGGGGAGGA CCTGCAGATG  
 1501 GACTGGGATG GCGCGGGGAG GCTGCTGTC AAGCTGTCCC CCGCTATGC  
 1551 CGGAAAGACC TGCGGCCTGT GTGGGAATTAA CAATGGCAAC CAGGGCGACG  
 1601 ACTTCTTAC CCCCTCTGGG CTGGCGGAGC CCCGGGTGGA GGACTTCGGG  
 1651 AACGCCCTGGA AGCTGCACGG GGACTGCCAG GACCTGCAGA AGCAGCACAG  
 1701 CGATCCCTGC GCCCTCAACC CGCGCATGAC CAGGTTCTCC GAGGAGGCCT  
 1751 GCGCGGTCCCT GACGTCCCCC ACATTCGAGG CCTGCCATCG TGCGTCAGC  
 1801 CCGCTGCCCT ACCTCGGAA CTGCGCTAC GACGTGTGCT CCTGCTCGA  
 1851 CGGCCGCGAG TGCCTGTGCG GCGCCCTGTC CAGCTATGCC GCGGCCCTGCG  
 1901 CGGGAGAGG CGTGCACGCT GCCTGGCCG AGCCAGGCCG CTGTGAGCTG  
 1951 AACTGGCCGA AAAGCTCAGGT GTACCTGCAC TGCGGCACCC CCTGCAACCT  
 2001 GACCTGCCGC TCTCTCTT ACCCGGATGA GGAATGCAAT GAGGCCCTGCC  
 2051 TGGAGGGCTG CTTCTGCCCT CCAGGGCTCT ACATGGATGA GAGGGGGGAC  
 2101 TGCCTGCCCA AGGCCAGTG CCCCTGTTAC TATGACGGTG AGATCTTCCA  
 2151 GCCAGAAGAC ATCTTCTCAG ACCATCACAC CATGTGCTAC TGTGAGGATG  
 2201 GCTTCATGCA CTGTACCATG AGTGGAGTCC CGGGAAGCTT GCTGCCCTGAC  
 2251 GCTTCCTCA CGAGTCCCCC GTCTCATCGC AGCAAAAGGA GCCTATCCTG  
 2301 TCGGCCCCCCC ATGGTCAAGC TGGTGTGTC CGCTGACAAC CTGCGGGCTG  
 2351 AAGGGCTCGA GTTACCAAA ACGTGCCAGA ACTATGACCT GGAGTGATG  
 2401 AGCATGGGCT GTTCTCTGG CTGCCTCTGC CCCCCGGGCA TGCGCCGGCA  
 2451 TGAGAACAGA TGTGTGGCC TGGAAAGTC TCCCTGCTTC CATCAGGGCA  
 2501 AGGAGTATGC CCCCTGGAGAA ACAGTGAAGA TTGGCTGCAA CACTTGTGTC  
 2551 TGTGGGACG GGAAGTGGAA CTGCACAGAC CATGTGTGTC ATGCCACGTG  
 2601 CTCACGATC GGATGGCCC ACTACCTCAC CTTGCACGGG CTCAAATACC  
 2651 TGTTCCCCGG GGAGTGGCAG TACGTTCTGG TGCGAGGATTA CTGCGGGCAGT  
 2701 AACCTGGGA CCTTCGGAT CCTAGTGGGG AATAAGGGAT GCAGCCACCC  
 2751 CTCAGTGAAA TGCAAGAAC GGGTCACCAT CCTGGTGGAG GGAGGAGAGA  
 2801 TTGAGCTGTT TGACGGGGAG GTGAATGTGA AGAGGCCAT GAAGGATGAG  
 2851 ACTCACTTTG AGGTGGTGGA GTCTGGCCCG TACATCATTC TGCTGCTGGG  
 2901 CAAAGCCCTC TCCGTGGTCT GGGACCGCA CCTGAGCATC TCCGTGGTCC  
 2951 TGAAGCAGAC ATACCAAGGAG AAAGTGTGTC GCCTGTGTGG GAATTTTGAT  
 3001 GGCATCCAGA AAATGACCT CACCAGCAG AACCTCCAAG TGGAGGAAGA  
 3051 CCCTGTGGAC TTTGGGAACT CCTGGAAAGT GAGCTCGCAG TGTGCTGACA  
 3101 CCAGGAAAGT GCCTCTGGAC TCATCCCCCTG CCACCTGCCA TAACAACATC  
 3151 ATGAAGCAGA CGATGGTGGA TTCCCTCTGT AGAACCTTA CCAGTGACGT  
 3201 CTTCCAGGAC TGCAACAAGC TGGTGGACCC CGAGCCATAT CTGGATGTCT  
 3251 GCATTACGA CACCTGCTCC TGTGAGTCCA TTGGGACTG CGCCGCATTC  
 3301 TGCACACCA TTGCTGCTA TGCCCACGT TGTCGCCAGC ATGGCAAGGT  
 3351 GGTGACCTGG AGGACGGCCA CATTGTCGCC CCAGAGCTGC GAGGAGAGGA  
 3401 ATCTCCGGGA GAACGGGTAT GAGGCTGAGT GCGCTATAA CAGCTGTGCA  
 3451 CCTGCCCTGTC AAGTCACGTC TCAGCACCT GAGCCACTGG CCTGCCCTGT  
 3501 GCAGTGTGTG GAGGGCTGCC ATGCCACTG CCCTCCAGGG AAAATCCTGG  
 3551 ATGAGCTTTT GCAGACCTGC GTTGACCTG AAGACTGTCC AGTGTGTGAG  
 3601 GTGGCTGGCC GCGTTTGC CTCAGGAAAG AAAGTCACCT TGAATCCCAG  
 3651 TGACCCCTGAG CACTGCCAGA TTTGCCACTG TGATGTTGTC AACCTCACCT  
 3701 GTGAAGCCTG CCAGGAGCCG ATATCGGCG CGCCAACATC AGAGAGGCC  
 3751 ACCCCTGAAA GTGGTCCCG GAGCGAGCCA GCCACATCTG GTTCGGAAAC  
 3801 GCCAGGCACA AGTGAGCTG CAACTCCCGA GTCCGGACCT GGCTCCGAGC  
 3851 CTGCCACTAG CGGCTCCCGAG ACTCCGGAA CTTCCGAGAG CGCTACACCA  
 3901 GAAAGCGGAC CGGAAACCG TACCGAACCT AGCGAGGGCT CTGCTCGGG  
 3951 CAGCCCAGCC GGCTCTCTA CATCCACCGA GGAGGGCACT TCCGAATCCG  
 4001 CCACCCCGGA GTCAAGGCCA GGATCTGAAC CCGCTACCTC AGGCAGTGAG  
 4051 ACGCCAGGAA CGAGCGAGTC CGCTACACCG GAGAGTGGC CAGGGAGCCC  
 4101 TGCTGGATCT CCTACGTCCA CTGAGGAAGG GTCACCAAGCG GGTCTGCCA  
 4151 CCAGCACTGA AGAAAGGTGCC TCGAGCGACA AAAACTCACAC ATGCCACCG  
 4201 TGCCCACTGC CAGAACTCCT GGCGGACCG TCAGTCTTCC TCTTCCCCC

4251 AAAACCCAAG GACACCCCTCA TGATCTCCCG GACCCCTGAG GTCACATGCG  
 4301 TGGTGGTGGAA CGTGAGGCCAC GAAGACCTCG AGGTCAAGTT CAACTGGTAC  
 4351 GTGGACGGCG TGGAGGTGCA TAATGCCAAG ACAAAAGCCGC GGGAGGAGCA  
 4401 GTACAAACAGC ACGTACCGTG TGGTCAGCGT CCTCACCGTC CTGACCCAGG  
 4451 ACTGGCTGAA TGGCAAGGAG TACAAGTGCA AGGTCTCCAA CAAAGCCCTC  
 4501 CCAGCCCCCA TCGAGAAAAC CATCTCCAA GCGAACGGC AGCCCCGAGA  
 4551 ACCACAGGTG TACACCCCTGC CCCCCATCCCC GGATGAGCTG ACCAAGAAC  
 4601 AGGTCACTGC GACCTGCCTG GTCAAAGGCT TCTATCCCAG CGACATGCC  
 4651 GTGGAGTGGG AGAGCAATGG GCAGCCGGAG ACAACTACA AGACCAAGGCC  
 4701 TCCCGTGTG GACTCCGACG GCTCCTTCTT CCTCTACAGC AAGCTCACCG  
 4751 TGGACAAGAG CAGGTGGCAG CAGGGGAACG TCTTCTCATG CTCCGTGATG  
 4801 CATGAGGCTC TGCACAACCA CTACACGAG AAGAGCCTCT CCCTGTCTCC  
 4851 GGGTAAATGA

[0327]

[0328]

[0329]

pSYN VWF062 단백질 서열(링커내에 트로빈 부위가 없는 VWF D'D3-Fc)(서열 번호 84)

1 MIPARFAGVL LALALILPGT LCAEGTRGRS STARCSLFGS DFVNTFDGSM  
 51 YSFAGYCSYL LAGGCQKRSF SIIIGDFQNGK RVSLSVYLGE FFDIHLFVNG  
 101 TVTQGDQRVS MPYASKGLYL ETEAGYYKLS GEAYGFVARI DGSGNFQVLL  
 151 SDRYFNKTCG LCGFNIFAE DDFMTQEGLT TSDPYDFANS WALSSGEQWC  
 201 ERASPSSSC NISSGEMQKG LWEQCQLKS TSVFARCHPL VDPEPFVALC  
 251 EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME  
 301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC  
 351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD  
 401 NRYFTFSGIC QYLLARDQD HSFSIVIETV QCADDRAVC TRSVTVRLPG  
 451 LHNSLVKLKH GAGVAMDQGD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM  
 501 DWGDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG  
 551 NAWKLHGDCQ DLQKQHSDPC ALNPMTMTRFS EEACAVLTSP TFEACRRAVS  
 601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVRV AWREPGRCSEL  
 651 NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGFCFP PGLYMDERGD  
 701 CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD  
 751 AVLSSPLSHR SKRSLSCRPP MVKLVCPADN LRAEGLECTK TCQNYDLECM  
 801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV  
 851 CRDRKWNCND HVCDATCSTI GMAHYLTFDG LKYLFPGECCQ YVLVQDYCGS  
 901 NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE  
 951 THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD  
 1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI  
 1051 MKQTMVDSSC RILTSDFVQD CNKLVDPEPY LDVCIYDTCS CESIGDCAAF  
 1101 CDTIAAYAHV CAQHGKVVTW RTATLCPQSC EERNLRENGY EAEWRYNSCA  
 1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE  
 1201 VAGRRFASGK KVTLNPSDPE HCQICHCDVV NLTCEACQEP ISGAPTSESA  
 1251 TPESGPSEP ATSGSETPGT SESATPESGP GSEPATSGSE TPGTSESATP  
 1301 ESGPGTSTEP SEGSAPGSPA GSPTSTEGET SESATPESGP GSEPATSGSE  
 1351 TPGTSESATP ESGPGSPAGS PTSTEEGSPA GSPTSTEEGA SSDKTHTCPP  
 1401 CPAPELLGGP SVFLFPPKPK DTLMISRTP E VTCVVVDVSH EDPEVKFNWY  
 1451 VDGVEVHNAAK TKPREEQYNS TYRVVSVLTV LHQDWLNGKE YKCKVSNKAL  
 1501 PAPIEKTIISK AKGQPREPQV YTLPPSRDEL TKNQVSLTCL VKGFYPSDIA  
 1551 VEWESNGQPE NNYKTTPPV р DSDGSFFLYS KLTVDKSRWQ QGNVFSCSVM  
 1601 HEALHNHYTQ KSLSLSPGK\*

[0330]

pSYN FVIII 286 뉴클레오파이드 서열(FVIII와 Fc 사이에 a2 영역이 있는 FVIII-Fc) (서열 번호 85)

[0331] ATGCAAATAG AGCTCTCCAC CTGCTTCTTT CTGTGCCCTTT TGCATTCTG

[0332]

51 CTTTAGTGGC ACCAGAAGAT ACTACCTGGG TGCAGTGGAA CTGTCATGGG  
 101 ACTATATGCA AAGTGATCTC GGTGAGCTGC CTGTGGACGC AAGATTTCT  
 151 CCTAGAGTGC CAAAATCTTT TCCATTCAAC ACCTCAGTGC TGATACAAAAA  
 201 GACTCTGTT GTAGAATTCA CGGATCACCT TTCAACACATC GCTAACGCAA  
 251 GGCCACCCCTG GATGGGCTCG CTAGGTCCTA CCATCCAGGC TGAGGTTTAT  
 301 GATACAGTGG TCATTACACT TAAGAACATG GCTTCCCACAT CTGTCAGTCT  
 351 TCATGCTGTT GGTGATTCCT ACTGAAAGC TTCTGAGGGA GCTGAATATG  
 401 ATGATCAGAC CAGTCAAAGG GAGAAAGAAG ATGATAAAAGT CTTCCTGGT  
 451 GGAAGCCATA CATATGCTCG GCAGGTCCTG AAAGAGAATG GTCCAATGGC  
 501 CTCTGACCCA CTGTGCCCTA CCTACTCATA TCTTCTCAT GTGGACCTGG  
 551 TAAAAGACTT GAATTCAAGC CTCACTGGAG CCCTACTAGT ATGTTAGAGAA  
 601 GGGAGTCTGG CCAAGGAAAA GACACAGACC TTGACACAAAT TTATACTACT  
 651 TTTGCTGTA TTTGATGAAAG GGGAAAGTTG GCACCTCAGAA ACAAAAGAAT  
 701 CCTTGATGCA GGATAGGGAT GCTGCATCTG CTCGGGCCTG GCCTAAAATG  
 751 CACACAGTCA ATGGTTATGT AAACAGGTCT CTGCCAGGTC TGATTGGATG  
 801 CCACAGGAAA TCAGTCTATT GGCAATGTGAT TGGAATGGGC ACCACTCCTG  
 851 AAGTGCACCT AATATTCTC GAAGGTCAAA CATTCTTGT GAGGAACCAT  
 901 CGCCAGGCTA GCTTGGAAAT CTCGCAATA ACTTCCCTTA CTGCTCAAAC  
 951 ACTCTTGATG GACCTTGAC AGTTTCTACT GTTTTGTCT ATCTCTTCCC  
 1001 ACCAACATGAA TGGCATGGAA GCTTATGTC AAGTAGACAG CTGTCCAGAG  
 1051 GAACCCCCAAC TACGAATGAA AAATAATGAA GAAGCGGAAG ACTATGATGA  
 1101 TGATCTTACT GATTCTGAAA TGGATGTGGT CAGGTTGTGAT GATGACAAC  
 1151 CTCCTTCCTT TATCCAAATT CGCTCAGTTG CCAAGAAGCA TCCTAAAC  
 1201 TGGGTACATT ACATTCTGTC TGAAGAGGAG GACTGGACT ATGCTCCCTT  
 1251 AGTCCTCGCC CCCGATGACA GAAGGTATAA AAGTCATAT TTGAAACAATG  
 1301 GCCCTCAGCG CATTGGTAGG AAGTACAAAAA AAGTCGGATT TATGCGCATAC  
 1351 ACAGATGAAA CTTTTAAGAC TCGTGAAGCT ATTCAAGCATG AATCAGGAAT  
 1401 CTTGGGACCT TTACTTTATG GGGAAAGTTG AGACACACTG TTGATTATAT  
 1451 TTAAGAATCA AGCAACAGGA CCATATAACA TCTACCTCTA CGGAATCACT  
 1501 GATGTCCTGC CTTTGTTTTC AAGGAGATTAA CCAAAAGGTG TAAACATTT  
 1551 GAAGGATTTC CCAATTCTGC CAGGAGAAAT ATTCAAATAT AAATGGACAG  
 1601 TGACTGTAGA AGATGGGCA ACTAAATCAG ATCCTCGGTG CCTGACCCGC  
 1651 TATTACTCTA GTTGTGTTA TATGGAGAGA GATCTAGCTT CAGGACTCAT  
 1701 TGCCCTCTCCTCCTA CAAAGAATTC TGTTAGATCAA AGAGGAAACC  
 1751 AGATAATGTC AGACAAGAGG AATGTCATCC TGTTTCTGT ATTGATGAG  
 1801 AACCGAAGCT GGTACCTCAC AGAGAATATA CAACGCTTTC TCCCAATCC  
 1851 AGCTGGAGTG CAGCTTGAGG ATCCAGAGT CCAAGCCTCC AACATCATGC  
 1901 ACAGCATCAA TGGCTATGTT TTTGATAGTT TGCAGTTGTC AGTTTGTGTC  
 1951 CATGAGGTTG CATACTGGTA CATTCTGAGC ATTGGAGAC AGACTGACTT  
 2001 CCTTTCTGTC TTCTTCTCTG GATATAACCTT CAAACACAAA ATGGTCTATG  
 2051 AAGACACACT CACCCATTTC CCATTCTCAG GAGAAACTGT CTTCATGTCG  
 2101 ATGGAAAACCCTCAGGCTATG GATTCTGGGG TGCCACAACT CAGACTTTCG  
 2151 GAAACAGAGG ATGACCCCT TACTGAAGGT TTCTAGTTG GACAAGAAC  
 2201 CTGGTGATTA TTACGAGGAC AGTTATGAAAG ATATTCAGC ATACTGCTG  
 2251 AGTAAAACAA ATGCCATTGA ACCAAGAAGC TTCTCTCAAAC AGGGCGGCC  
 2301 AGGTACCTCA GAGTCTGCTA CCCCCGAGTC AGGGCCAGGA TCAGAGGCCAG  
 2351 CCACCTCCGG GTCTGAGACA CCCGGACTT CCGAGAGTGC CACCCCTGAG  
 2401 TCCGGACCCG GGTCCGAGCC CGGCACTTCC GGCTCCGAAA CTCCGGCAC  
 2451 AAGCGAGAGC GCTACCCAG AGTCAGGACC AGGAACATCT ACAGAGCCCT  
 2501 CTGAAGGCTC CGCTCAGGG TCCCCAGCCG GCAGTCCCCAC TAGCACCGAG  
 2551 GAGGGAAACCT CTGAAAGCGC CACACCCGAA TCAGGGCCAG GGTCTGAGCC  
 2601 TGCTTACCAAGC GGCAGGGAGA CACCAAGGCAC CTCTGAGTCC GCCACACCAAG  
 2651 AGTCCGGACC CGGATCTCCC GCTGGGAGCC CCACCTCCAC TGAGGAGGG  
 2701 TCTCTGCTG GCTCTCCAAC ATCTACTGAG GAAGGTACCT CAACCGAGCC  
 2751 ATCCGAGGGA TCAGCTCCCG GCACCTCAGA GTCCGGAAACC CGGGAGTCTG  
 2801 GACCCGGAAC TTCCGAAAGT GCCACACCAAG AGTCCGGTCC CGGGACTTCA  
 2851 GAATCAGCAA CACCCGAGTC CGGCCCTGGG TCTGAACCCG CCACAAGTGG

[0333]

2901 TAGTGAGACA CCAGGATCA AACCTGCTAC CTCAGGGTCA GAGACACCCG  
 2951 GATCTCCGGC AGGCTCACCA ACCTCCACTG AGGAGGGCAC CACACAGAA  
 3001 CCAARGCAGG GCTCCGCACC CGGAACAAGC ACTGAACCCA GTGAGGGTTTC  
 3051 AGCACCCGGC TCTGAGCCGG CCACAAAGTGG CAGTGAGACA CCCGGCACTT  
 3101 CAGAGAGTGC CACCCCCGAG AGTGGCCAG GCACATGAC CGAGCCCTCT  
 3151 GAAGGCAGTG CGCCAGCCTC GAGCCACCA GTCTTGAAAC GCATCAAGC  
 3201 TGAAATAACT CGTACTACTC TTCAAGTCAGA TCAAGAGGAA ATCGATTATG  
 3251 ATGATACCAT ATCAGTTGAA ATGAAGAAGG AAGATTTGA CATTATGAT  
 3301 GAGGATGAAA ATCAGAGCCC CGCGACGTTT CAAAAGAAAA CACGACACTA  
 3351 TTTTATGCT GCAGTGGAGA GGCTCTGGGA TTATGGATG AGTAGCTCCC  
 3401 CACATTTCT AAGAAACAGG GCTCAGAGTG GCAGTGTCCC TCAGTTCAAG  
 3451 AAAGTTGTT TCCAGGAATT TACTGATGCC TCCTTTACTC AGCCCTTATA  
 3501 CCGTGGAGAA CTAATGAAC ATTTGGACT CCTGGGCCA TATATAAGAG  
 3551 CAGAAGTTGA AGATAATATC ATGGTAACCT TCAGAAATCA GGCCCTCTCGT  
 3601 CCCTATTCCT CTCATTCTAG CCTTATTTCT TATGAGGAAG ATCAGAGGCA  
 3651 AGGAGCAGAA CCTAGAAAAAA ACTTTGTCAA GCCTAAATGAA ACCAAAACCTT  
 3701 ACTTTGGAA AGTGCACAT CATATGGCAC CCACAAAGA TGAGTTGAC  
 3751 TGCAAGCCT GGGCTTATT CTCTGATGTT GACCTGGAAA AAAGATGTGCA  
 3801 CTCAGGACCCCTT ATGGGACCCCTT CTCTGGTCTG CCACACTAAC ACAGTGAACC  
 3851 CTGCTCATGG GAGACAAAGTG ACAGTACAGG AATTGCTCT GTTTTCAACC  
 3901 ATCTTTGATG AGACCAAAAG CTGGTACTTC ACTGAAAATA TGAAAGAAA  
 3951 CTGAGGGCT CCCTGCAATA TCCAGATGGA AGATCCCCT TTTAAAGAGA  
 4001 ATTATGCTT CTCAGGACAT AATGGCTACA TAATGGATAC ACTACCTGGC  
 4051 TTAGTAATGG CTCAGGATCA AAGGATTGCA TGGTATCTGC TCAGCATGG  
 4101 CAGCAATGAA AACATCCATT CTATTCAATT CAGTGGACAT GTGTTCACTG  
 4151 TACGAAAAAA AGAGGAGTAT AAAATGGCAC TGACAAATCT CTATCCAGGT  
 4201 GTTTTGAGA CAGTGGAAAT GTTACCATCC AAAGCTGGAA TTGGCAGGGT  
 4251 GGAATGCCCTT ATTGGCGAGC ATCTACATGC TGGGATGAGC ACACCTTTTC  
 4301 TGGGTACAG CAATAAGTGT CAGACTCCCC TGGGAATGGC TTCTGGACAC  
 4351 ATTAGGAGTT TTCAGATTAC AGCTTCAGGA CAATATGGAC AGTGGGCC  
 4401 AAAGCTGGCC AGACTTCATT ATTCCGGATC AATCAATGCC TGAGCAGCA  
 4451 AGGAGCCCTT TTCTGGATC AAGGTGGATC TGTTGGCACC AATGATTATT  
 4501 CACGGCATCA AGACCCAGGG TGCCCGTCAG AAGTCTCCA GCCTCTACAT  
 4551 CTCTCAGTTT ATCATCATGT ATAGTCTTGA TGGGAAGAAG TGGCAGACTT  
 4601 ATCAGGAAA TTCCACTGGA ACCTTAATGG TCTTCTTGG CAATGTGGAT  
 4651 TCATCTGGGA TAAACACAA TATTCTTAAAC CCTCTCAATTAA TTGCTCGATA  
 4701 CATCCGTTT CACCCAACTC ATTATAGCAT TCGCAGCACT CTTCGCGATGG  
 4751 AGTTGATGGG CTGTGATTTA AATAGTTGCA GCATGCCATT GGAATGGAG  
 4801 AGTAAAGCAA TATCAGATGC ACAGATTACT GCTTCATCCT ACTTTACCAA  
 4851 TATGTTGCC ACCTGGTCTC TTTCAGAACG TCGACTTCAC CTCCAAGGG  
 4901 GGAGTAATGC CTGGAGACCT CAGGTGAATA ATCAXXAGA GTGGCTGCA  
 4951 GTGGACTCTC AGAAAGACAT GAAAGTCACA GGAGTAACCA CTCAAGGGAGT  
 5001 AAAATCTCTG CTTACAGCA TGATGTGAA GGAGTTCCCTC ATCTCCAGCA  
 5051 GTCAAGATGG CCATCAGTGG ACTCTCTTTT TTCAAGATGG CAAAGTAAAG  
 5101 GTTTTCAGG GAAATCAAGA CTCCCTCACA CCTGTGGTGA ACTCTCTAGA  
 5151 CCCACCGTTA CTGACTCGCT ACCTTCGAAT TCACCCCCAG AGTGGGGTGC  
 5201 ACCAGAGTGC CCTGAGGATG GAGGTTCTGG GCTCGAGGG ACAGGACCTC  
 5251 TACGACAAAG ACACTGGTGA TTATTACCGAG GACAGTTATG AAGATATTTC  
 5301 AGCATACTTG CTGAGTAAAA ACAATGCCAT TGAACCAAGA AGCTTCTCTG  
 5351 ACAAAACTCA CACATGCCCA CGGTGCCAG CTCAGAACT CCTGGGCC  
 5401 CCGTCAGTCT TCCCTTCTCC CCCCCAACCC AAGGACACCC TCATGATCTC  
 5451 CGGACCCCT GAGGTCACT GCGGGTGGT GGACCTGAGC CACGAAGACC  
 5501 CTGAGGTCAGA GTTCAACTGG TACGTGGAGC GCGTGGAGGT GCATAATGCC  
 5551 AAGACAAAGC CGCGGGAGGA GCAGTACAAC AGCACAGTAC GTGTGGTCAG  
 5601 CGTCCTCACC GTCCCTGCACC AGGACTGGCT GAATGGCAAG GAGTACAAGT  
 5651 GCAAGGCTC CAACAAAGCC CTCCCAAGCC CCAATCGAGAA AACCATCTCC  
 5701 AAAGCCAAAG GGCAGCCCCG AGAACACAG GTGTACACCC TGCCCCCATC

5751 CGGGGATGAG CTGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG  
 5801 GCTTCTATCC CAGCGACATC GCCGTGGAGT GGGAGAGCAA TGGCAGCCG  
 5851 GAGAACAACT ACAAGACACCG GCCTCCCGTG TTGGACTCCG ACAGCTCCTT  
 5901 CTTCCCTAC AGCAAGCTCA CGGTGGACAA GAGCAGGTGG CAGCAGGGGA  
 5951 ACGCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCAACTACACG  
 6001 CAGAAGAGCC TCTCCCTGTC TCCGGTAAAG TGA

[0334]

[0335]

[0336]

pSYN\_FVIII\_286 단백질 서열(FVIII와 Fc 사이에 추가의 a2 영역을 지닌 FVIII-Fc; 굵은 글씨 및 밑줄로 나타냄)(서열 번호 86)

1 ATRYYLGAV ELSWDYMQSD LGELPVDA RF PPRVPKSFPF NTSVVYKKTL  
 51 FVEFTDHLFN IAKPRPPWMG LLGPTIQA EV YDTVVITLKN MASHPVL HA  
 101 VGVS YWKASE GAEYDDQTSQ REKEDDKVFP GGSHTYVWQV LKENGPMASD  
 151 PLCLTYSYLS HVDLVKDLNS GLIGALLVCR EGSLAKEKTQ TLHKFILLFA  
 201 VFDEGKSWHS ETKNSLMQDR DAASARAWPK MHTVNGYVNR SLPGLIGCHR  
 251 KSVYWHVIGM GTTPEVHSIF LEGHTFLVRN HROASLEISP ITFLTAQTL  
 301 MDLGQFLFC HISSHQHDGM EAYVKVDFSCP EEPQLRMKNN EEAEDYDDDL  
 351 TDSEMDVVRF DDDNSPSFIQ IRSVAKKH PK TWVHYIAAEE EDWDYAPLVL  
 401 APDDRSYKSQ YLNNGPQRIG RKYKKVRFMA YTDETFKTR AIQHESGILG  
 451 PLLYGEVGDT LIIIFKNQAS RPYN IYPHGI TDVRPLYSRR LPKGVKHLKD  
 501 FPILPGEIFK YKWTVTVEDG PTKSDPRCLT RYSSFVNME RDLASGLIGP  
 551 LLIC YKESVD QRGNQIMS DK RNVILFSVFD ENR SWYL TEN IQRFLPNPAG  
 601 VQLEDPEFQA SNIMHSINGY VFDSLQLSVC LHEVAYWYIL SIGAQTDFLS  
 651 VFFSGYTFKH KMVYEDTLTL FPFSGETVFM SMENPGLWIL GCHNSDFRMR  
 701 GMTALLKVSS CDKNTGDYY DSYEDISAYL LSKNNAIEPR SFSQNGAPGT  
 751 SESATPESGP GSEPA TSGSE TPGTSESATP ESGPGSEPAT SGSETPGTSE  
 801 SATPESGP GT STEPSEGSAP GSPAGSPTST EEGTSESATP ESGPGSEPAT  
 851 SGSETPGTSE SATPESGP GPGS PAGSPTSTEE GSPAGSPTST EEGTSTEPSE  
 901 GSAPGTSESA TPESGP GT SATPESGP GSEPATSGSE  
 951 TPGSEPATSG SETPGSPAGS PTSTEETGTST EPSEGSAPGT STEPSEGSAP  
 1001 GSEPATSGSE TPGTSESATP ESGPGT STEP SEGSAPASSP PVLKRHQAEI  
 1051 TRTTLQSQE EIDYDDTISV EMKKEDFDIY DEDENQSPRS FQKKTRHYFI  
 1101 AAVERLWDYG MSSSPHVLRN RAQSGSPQF KKVFQEFTD GSFTQPLYRG  
 1151 ELNEHLGLLG PYIRAEVEDN IMVTFRNQAS RPYSFYSSL SYEEDQRQGA  
 1201 EPRKNFVKPN ETKTYFWXVQ HHMAPTKDEF DCKAWAYFSD VDLEKDVHSG  
 1251 LIGPLLVCHT NTLNPAHGRQ VTVQEFALFF TIFDET KSWY FTENMERNCR  
 1301 APCNIQMEDP TFKENYRFHA INGYIMDTLP GLVMAQDQRI RWYLLSMGSN  
 1351 ENIHSIHFG HVFTVRKKEE YKMALYNLYP GVFETVEMLP SKAGIWRVEC  
 1401 LIGEHLHAGM STLFLVYSNK CQTPLGMASG HIRDFQITAS GQYQGWAPKL  
 1451 ARLHYSGSIN AWSTKEPFSW IKV DLLAPMI IHGIKTQGAR QKFSSLYISQ  
 1501 FIIMYSLDGK KWQTYRGNST GTLMVFFGNV DSSGIKHNI NPPIIARYIR  
 1551 LHPHTHYSIRS TLRMELMGCD LNSCSMPLGM ESKAISDAQI TASSYFTNMF  
 1601 ATWSPSKARL HLQGRSNAWR PQVNNPKEWL QVDFQKTMKV TVGTTQGVKS  
 1651 LLTSMYVKEF LISSSDQDHQ WTLFFQNGKV KVFGQNQDSF TPVVNSLDPP  
 1701 LLTRYLRIHP QSWVHQIALR MEVLGCEAQD LYDKNTGDYY EDSYEDISAY  
 1751 LLSKNNATEP RSFESDKTHTC PPCPAPELLG GPSVFLFPPK PKDTLMISRT  
 1801 PEVTCVVVDV SHEDPEVFKN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL  
 1851 TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRD  
 1901 ELTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFLL  
 1951 YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K\*

[0337]

## FVIII 169 뉴클레오파이드 서열(서열 번호 87)

1 ATGCA AATAG AGCTC TCCAC CTGCT TCTTT CTGTG CCTTT TGCGA TTCTG  
 51 CTTTA GTGCC ACCAG AAGAT ACTAC CTGGG TGCAG TGGAA CTGTC ATGGG  
 101 ACTAT ATGCA AAGTG ATCTC GGTGA GCTGC CTGTG GACGC AAGAT TTCCT

[0339]

151 CCTAG AGTGC CAAAA TCTTT TCCAT TCAAC ACCTC AGTCG TGTAC AAAAA  
 201 GACTC TGTTT GTAGA ATTCA CGGAT CACCT TTTCA ACATC GCTAA GCCAA  
 251 GGCAA CCCTG GATGG GTCTG CTAGG TCCCTA CCATC CAGGC TGAGG TTTAT  
 301 GATAC AGTGG TCATT ACACT TAAGA ACATG GCTTC CCATC CTGTC AGTCT  
 351 TCATG CTGTT GGGTG ATCCT ACTGG AAAGC TTCTG AGGGA GCTGA ATATG  
 401 ATGAT CAGAC CAGTC AAAGG GAGAA AGAAG ATGAT AAAGT CTTCC CTGGT  
 451 GGAAG CCATA CATAT GTCTG GCAGG TCCCTG AAAGA GAATG GTCCA ATGGC  
 501 CTCTG ACCCA CTGTG CCTTA CCTAC TCATA TCTTT CTCAT GTGCC CCTGG  
 551 TAAAA GACTT GAATT CAGGC CTCTAT TGAGG CCCTA CTAGT ATGTA GAGAA  
 601 GGGAG TCTGG CCAAG GAAAA GACAC AGACC TTGCA CAAAT TTATA CTACT  
 651 TTTTG CTGTA TTTGA TGAAG GGAAA AGTTG GCACT CAGAA ACAAA GAACT  
 701 CCTTG ATGCA GGATA GGGAT GCTGC ATCTG CTCGG GCCTG GCCTA AAATG  
 751 CACAC AGTCA ATGGT TATGT AAACA GGTCT CTGCC AGGTC TGATT GGATG  
 801 CCACA GGAAA TCAGT CTATT GGCGT GTGAT TGAA TGGGC ACCAC TCCCTG  
 851 AAGTG CACTC AATAT TCCTC GAAGG TCACA CATTG CTGTT GAGGA ACCAT  
 901 CGCCA GGCTA GCTTG GAAAT CTCGC CAATA ACTTT CCTTA CTGCT CAAAC  
 951 ACTCT TGATG GACCT TGGAC AGTTT CTACT GTTTT GTCAT ATCTC TTCCC  
 1001 ACCAA CATGA TGGCA GCTTA TGTCA AAGTA GACAG CTGTC CAGAG  
 1051 GAACC CCAAC TACCA ATGAA AAATA ATGAA GAAGC GGAAG ACTAT GATGA  
 1101 TGATC TTACT GATTC TGAAA TGGAT GTGGT CAGGT TTGAT GATGA CAACT  
 1151 CTCCT TCCCT TATCC AAATT CGCTC AGTTG CCAAG AAGCA TCCCTA AAACCT  
 1201 TGGGT ACATT ACATT GCTGC TGAAG AGGAG GACTG GGACT ATGCT CCCTT  
 1251 AGTCC TCGCC CCCGA TGACA GAAGT TATAA AAGTC AATAT TTGAA CAATG  
 1301 GCCCT CAGCG GATTG GTAGG AAGTA CAAAA AAGTC CGATT TATGG CATACT  
 1351 ACAGA TGAAA CCTTT AAGAC TCGTG AAGCT ATTCA GCATG AATCA GGAAT  
 1401 CTTGG GACCT TTACT TTATG GGGAA GTTGG AGACA CACTG TTGAT TATAT  
 1451 TTAAG AATCA AGCAA GCAGA CCATA TAACA TCTAC CCTCA CGGAA TCAC  
 1501 GATGT CCGTC CTTTG TATTG AAGGA GATTA CCAAA AGGTG TAAAA CATT  
 1551 GAAGG ATTAT CCAAT TCTGC CAGGA GAAAT ATTCA AATAT AAATG GACAG  
 1601 TGACT GTAGA AGATG GGCCA ACTAA ATC..G ATCCT CGGTG CCTGA CCCGC  
 1651 TATTA CTCTA GTTTC GTTAA TATGG AGAGA GATCT AGCTT CAGGA CTCAT  
 1701 TGGCC CTCTC CTCTA CTGCT ACAA AATTC GATTC TGTAG ATCAA AGAGG AAACC  
 1751 AGATA ATGTC AGACA AGAGG ATATG CATCC TGTGTT TCTGT ATTG ATGAG  
 1801 AACCG AAGCT GGTAC CTCAC AGAGA ATATA CAAAG CTTTC TCCCC AATCC  
 1851 AGCTG GAGTG CAGCT TGAGG ATCCA GAGTT CCAAG CCTCC AACAT CATGC  
 1901 ACAGC ATCAA TGGCT ATGTT TTTGA TAGTT TGCAG TTGTC AGTTT GTTTG  
 1951 CATGA GGTGG CATACT TGGTA CATTG TAAGC ATTGG AGCAC AGACT GACTT  
 2001 CCTTT CTGTC TTCTT CTCTG GATAT ACCTT CAAAC ACAAA ATGGT CTATG  
 2051 AAGAC ACACT CACCC TATTC CCATT CTCAG GAGAA ACTGT CTTCA TGTG  
 2101 ATGGA AAACC CAGGT CTATG GATTC TGGGG TGCCA CAACT CAGAC TTTGG  
 2151 GAACA GAGGC ATGAC CGCCT TACTG AAGGT TTCTA GTTGT GACAA GAACA  
 2201 CTGGT GATTA TTACG AGGAC AGTTA TGAAG ATATT TCAGC AT..CT TGCTG  
 2251 AGTAA AAACA ATGCC ATTGA ACCAA GAAGC TTCTC TCAAA ACGGC GCGCC  
 2301 AGGTA CCTCA GAGTC TGCTA CCCCC GAGTC AGGGC CAGGA TCAGA GCCAG  
 2351 CCACC TCCGG GTCTG AGACA CCCGG GACTT CCGAG AGTGC CACCC CTGAG  
 2401 TCCGG ACCCG GGTCC GAGCC CGCCA CTTCC GGCTC CGAAA CTCCC GGCAC  
 2451 AAGCG AGAGC GCTAC CCCAG AGTCA GGACC AGGAA CATCT ACAGA GCCCT  
 2501 CTGAA GGCTC CGCTC CAGGG TCCCC AGCCG GCAGT CCCAC TAGCA CCGAG  
 2551 GAGGG AACCT CTGAA AGCGC CACAC CCGAA TCAGG GCCAG GGTCT GAGCC  
 2601 TGCTA CCAGC GGCAG CGAGA CACCA GGCAC CTCTG AGTCC GCCAC ACCAG  
 2651 AGTCC GGACC CGGAT CTCCC GCTGG GAGCC CCACC TCCAC TGAGG AGGGA  
 2701 TCTCC TGCTG GCTCT CCAAC ATCTA CTGAG GAAGG TACCT CAACC GAGCC  
 2751 ATCCG AGGGG TCAGC TCCCG GCACC TCAGA GTCGG CAACC CCGGA GTCTG  
 2801 GACCC GGAAC TTCCG AAAGT GCCAC ACCAG AGTCC GGTCC CGGGG CTTCA  
 2851 GAATC AGCAA CACCC GAGTC CGGCC CTGGG TCTGA ACCCG CCACA AGTGG  
 2901 TAGTG AGACA CCAGG ATCAG AACCT GCTAC CTCAG GGTCA GAGAC ACCCG  
 2951 GATCT CGGGC AGGCT CACCA ACCTC CACTG AGGAG GGCAC CAGCA CAGAA

[0340]

3001 CCAAG CGAGG GCTCC GCACC CGGAA CAAGC ACTGA ACCCA GTGAG GGTTC  
 3051 AGCAC CCGGC TCTGA GCCGG CCACA AGTGG CAGTG AGACA CCCGG CACTT  
 3101 CAGAG AGTGC CACCC CCGAG AGTGG CCCAG GCACT AGTAC CGAGC CCTCT  
 3151 GAAGG CAGTG CGCCA GCCTC GAGCC CACCA GTCTT GAAAC GCCAT CAAGC  
 3201 TGAAA TAACT CGTAC TACTC TTICAG TCAGA TCAAG AGGAA ATCGA TTATG  
 3251 ATGAT ACCAT ATCAT TTGAA ATGAA GAAGG AAGAT TTTGA CATTT ATGAT  
 3301 GAGGA TGAAA ATCAG AGCCC CGGCA GCTTT CAAAAA GAAAAA CACGA CACTA  
 3351 TTTTA TTGCT GCAGT GGAGA GGCTC TGGGA TTATG GGATG AGTAG CTCCC  
 3401 CACAT GTTCT AAGAA ACAGG GCTCA GAGTG GCAGT GTCCC TCAGT TCAAG  
 3451 AAAGT TGTCTT TCCAG GAATT TACTG ATGGC TCCTT TACTC AGCCC TTATA  
 3501 CCGTG GAGAA CTAAA TGAAC ATTIG GGACT CCTGG GGCCA TATAT AAGAG  
 3551 CAGAA GTTGA AGATA ATATC ATGGT AACTT TCAGA AATCA GGCCT CTCGT  
 3601 CCCTA TTCTT TCTAT TCTAG CCTTA TTTCT TATGA GGAAG ATCAG AGGCA  
 3651 AGGAG CAGAA CCTAG AAAAAA ACTTT GTCAA GCCTA ATGAA ACCAA AACCT  
 3701 ACTTT TGGAA AGTGC AACAT CATAT GGAC CCACT AAAAGA TGAGT TTGAC  
 3751 TGCCTA AGCCT GGGCT TATTT CTCTG ATGTT GACCT GGAAA AAGAT GTGCA  
 3801 CTCAG GCCTG ATTGG ACCCC TTCTG GTCTG GCCTA CTAAC ACACT GAACC  
 3851 CTGCT CATGG GAGAC AAGTG ACAGT ACAGG AATTG GCTCT GTTTT TCACC  
 3901 ATCTT TGATG AGACC AAAAG CTGGT ACTTC ACTGA AAATA TGAA AGAAA  
 3951 CTGCA GGGCT CCCCTG CAATA TCCAG ATGGA AGATC CCACT TTTAA AGAGA  
 4001 ATTAT CGCTT CCTATG CAATC AATGG CTACA TAATG GATAC ACTAC CTGGC  
 4051 TTAGT AATGG CTCAG GATCA AAGGA TTGGA TGGTA TCTGC TCAGC ATGGG  
 4101 CAGCA ATGAA AACAT CCATT CTATT CATT CAGTG GACAT GTGTT CACTG  
 4151 TACGA AAAAAA AGAGG AGTAT AAAAT GGAC TGTAC AATCT CTATC CAGGT  
 4201 GTTTT TGAGA CAGTG GAAAT GTTAC CATCC AAAGC TGGAA TTTGG CGGGT  
 4251 GGAAT GCCTT ATTGG CGAGC ATCTA CATGC TGGGA TGAGC ACACT TTTTC  
 4301 TGGTG TACAG CAATA AGTGT CAGAC TCCCC TGGGA ATGGC TTCTG GACAC  
 4351 ATTAG AGATT TTCAG ATTAC AGCTT CAGGA CAATA TGGAC AGTGG GCCCC  
 4401 AAAGC TGGCC AGACT TCATT ATTCC GGATC AATCA ATGCC TGGAG CACCA  
 4451 AGGAG CCCTT TTCTT GGATC AAGGT GGATC TGTTC GCACC AATGA TTATT  
 4501 CACGG CATCA AGACC CAGGG TGCCC GTCAG AAGTT CTCCA GCCTC TACAT  
 4551 CTCTC AGTTT ATCAT CATGT ATAGT CTTGA TGGGA AGAAG TGCCA GACTT  
 4601 ATCGA GGAAA TTCCA CTGGA ACCTT AATGG TCTTC TTTGG CAATG TGGAT  
 4651 TCATC TGGGA TAAAAA CACAA TATTG TTAAC CCTCC AATTG TTGCT CGATA  
 4701 CATCC GTTTG CACCC AACTC ATTAT AGCAT TCGCA GCACT CTTCG CATGG  
 4751 AGTTG ATGGG CTGTG ATTAA AATAG TTGCA GCATG CCATT GGGAA TGGAG  
 4801 AGTAA AGCAA TATCA GATGC ACAGA TTACT GCTTC ATCCT ACTTT ACCAA  
 4851 TATGT TTGCC ACCTG GTCTC CTCTCA AAAGC TCGAC TTCACT CCCTA AGGGA  
 4901 GGAGT AATGC CTGGA GACCT CAGGT GAATA ATCCA AAAGA GTGGC TGCA  
 4951 GTGGA CTTCC AGAAC ACAAT GAAAG TCACA GGAGT AACTA CTCAG GGAGT  
 5001 AAAAT CTCTG CTTAC CAGCA TGAT GTTAT GTGAA GGAGT TCCTC ATCTC CAGCA  
 5051 GTCAA GATGG CCATC AGTGG ACTCT CTTTT TTCAAG AATGG CAAAG TAAAG  
 5101 GTTTT TCAGG GAAAT CAAGA CTCTC TCACA CCTGT GGTGA ACTCT CTAGA  
 5151 CCCAC CGTTA CTGAC TCGCT ACCTT CGAAT TCACC CCCAG AGTTG GGTGC  
 5201 ACCAG ATTGC CCTGAA GGATG GAGGT TCTGG GCTGC GAGGC ACAGG ACCTC  
 5251 TACGA CAAAAA CTCAC ACATG CCCAC CGTGC CCAGC TCCAG AACTC CTGGG  
 5301 CGGAC CGTCA GTCCT CCTCT TCCCC CAAAA ACCCA AGGAC ACCCT CATGA  
 5351 TCTCC CGGAC CCCTG AGGTC ACATG CGTGG TGGTG GACGT GAGCC ACGAA  
 5401 GACCC TGAGG TCAAG TTCAA CTGGT ACGTG GACCG CGTGG AGGTG CATAA  
 5451 TGCCA AGACA AAGCC CGGGG AGGAG CAGTA CAACA GCACG TACCG TGTGG  
 5501 TCAGC GTCTT CACCG CCTGT CACCA GGACT GGCTG AATGG CAAGG AGTAC  
 5551 AAGTG CAAGG TCTCC AACAA AGCCC TCCCA GCCCC CATCG AGAAA ACCAT  
 5601 CTCCA AAGCC AAAGG GCAGC CCCGA GAACC ACAGG TGATC ACCCT GCCCC  
 5651 CATCC CGGGA TGAGC TGACC AAGAA CCAGG TCAGC CTGAC CTGCC TGGTC  
 5701 AAAGG CTTCT ATCCC AGCGA CATCG CCGTG GAGTG GGAGA GCAAT GGGCA  
 5751 GCCGG AGAAC AACTA CAAGA CCACG CCTCC CGTGT TGGAC TCCGA CGGCT  
 5801 CCTTC TTCCT CTACA GCAAG CTCAC CGTGG ACAAG AGCAG GTGGC AGCAG

[0341]

[0342]

5851 GGGAA CGTCT TCTCA TGCTC CGTGA TGCAAT GAGGC TCTGC ACAAC CACTA  
 5901 CACGC AGAAC AGCCT CTCCC TGCTC CCCGG TAAAT GA

[0344]

FVIII 169 단백질 서열(서열 번호 88)

1 MQIELSTCFF LCLLRFCFSA TRRYYLGAVE LSWDYMQSGL GELPVDARFP  
 51 PRVPKSFPFN TSVVYKKTLF VEFTDHFLFNI 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 TFLTAQTLMM DLGQFLLFCH ISSHQHDGME AYVKVDSCPE  
 351 EPQLRKMNN EAEDYDDDLT DSEMDVVRFD DDNSPSFIQI RSVAKKHPKT  
 401 WVHYIAEEE DWWDYAPLVLA PDDRSYKSQY LNNGPQRIGR KYKKVRFMAY  
 451 TDETFKTREA IQHESGILGP LLYGEVGDTL LIIFKNQASR PYNIYPHGIT  
 501 DVRPLYSRRL PKGVKHLKDF PILPGEIFKY KWTVTVEDGP TKSDPRLTR  
 551 YYSSFVNMER DLASGLIGPL LICYKESVDQ RGNQIMSDKR NVILFSVFDE  
 601 NRSWYLTERI QRFLPNPAGV QLEDPEFQAS NIMHSINGYV FDSDLQLSVCL  
 651 HEVAYWYIIS IGAQTDFLSV FFSGYTFKHK MVYEDTLLTF PFSGETVFMMS  
 701 MENPGLWILG CHNSDFRNRRG MTALLKVSSC DKNTGDDYYED SYEDISAYLL  
 751 SKNNAIEPRS FSQNGAPGTS ESATPESGPG SEPATSGSET PGTSESATPE  
 801 SGPGSEPAT S GSETPGTSES ATPESGPGS TEPSEGSAPG SPAGSPTSTE  
 851 EGTSESATPE SGPGSEPAT S 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 IDYDDDTISVE MKKEDFDIYD  
 1101 EDENQSPRSF QKKTRHYFIA AVERLWDYGM SSSPHVLRNR AQSGSVPQFK  
 1151 KVVFQEFTDG SFTQPLYRGE LNEHLLGP YIRAEVEDNI MVTFRNQASR  
 1201 PYSFYSSLIS YEEDQRQGAE PRKNFVKPNE TKTYFWKVQH HMAPTKDEF  
 1251 CKAWAYFSDV DLEKDVHSGL IGPLLVCHTN TLNPAGHRQV TVQEFLALFFT  
 1301 IFDETKSWSYF TENMERNCR A PCNIQMEDPT FKENYRFHAI NGYIMDTLPG  
 1351 LVMAQDQRIR WYLLSMGSNE NIHSIHFSGH VFTVRKKEEY KMALYNLYPG  
 1401 VFETVEMLPS KAGIWRVECL IGEHLHAGMS TLFVYSNKC QTPLGMASGH  
 1451 IRDFQITASG QYGQWAPKLA RLHYSGSINA WSTKEPFSWI KVDLLAPMII  
 1501 HGIKTQGARQ KFSSLYISQF IIMYSLDGKK WQTYRGNSTG TLMVFFGNVD  
 1551 SSGIKHNIFN PPIIARYIRL HPTHYSIRST LRMELMGCDL NSCSMPLGME  
 1601 SKAISDAQIT ASSYFTNMFA TWSPSKARLH LQGRSNAWRP QVNNPKEWLQ  
 1651 VDFQKTMKV T GTTQGVKSL LTSMYVKEFL ISSSQDGHQW TLFFQNGKVK  
 1701 VFQGNQDSFT PVVNSLDPP LTRYLRIHPQ SWVHQIALRM EVLGCEAQDL  
 1751 YDKTHTCPPC PAPELLGGPS VFLFPKKPKD TLMISRT EV TCVVVDVSHE  
 1801 DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL HQDWLNGKEY  
 1851 KCKVSNKALP APIEKTIKA KGQPREPQVY TLPPSRDELT KNQVSLTCLV  
 1901 KGFYPSDIAV EWESNGQOPEN NYKTTPPVLD SDGSFFLYSK LTVDKSRWQQ  
 1951 GNVFSCSVMH EALHNHYTQK SLSLSPGK\*

[0345]

WVF034 뉴클레오파이드 서열(서열 번호 91)

1 ATGAT TCCTG CCAGA TTTGC CGGGG TGCTG CTTGC TCTGG CCCTC ATTTT  
 51 GCCAG GGACC CTTTG TGCAG AAGGA ACTCG CGGCA GGTCA TCCAC GGCCC  
 101 GATGC AGCCT TTTCG GAACT GACTT CGTCA ACACC TTTGA TGGGA GCATG  
 151 TACAG CTTTG CGGGA TACTG CAGTT ACCTC CTGGC AGGGG GCTGC CAGAA  
 201 ACGCT CCTTC TCGAT TATTG GGGAC TTCCA GAATG GCAAG AGAGT GAGCC  
 251 TCTCC GTGTA TCTTG GGGAA TTTTT TGACA TCCAT TTGTT TGTCA ATGGT  
 301 ACCGT GACAC AGGGG GACCA AAGAG TCTCC ATGCC CTATG CCTCC AAAGG  
 351 GCTGT ATCTA GAAAC TGAGG CTGGG TACTA CAAGC TGTCC GGTGA GGCCT  
 401 ATGGC TTTGT GGCCA GGATC GATGG CAGCG GCAAC TTTCA AGTCC TGCTG

[0347]

451 TCAGA CAGAT ACTTC AACAA GACCT GCGGG CTGTG TGGCA ACTTT AACAT  
 501 CTTTG CTGAA GATGA CTTTA TGACC CAAGA AGGGA CCTTG ACCTC GGACC  
 551 CTTAT GACTT TGCCA ACTCA TGGGC TCTGA GCAGT GGAGA ACAGT GTGTT  
 601 GAACG GGCAT CTCCCT CCCAG CAGCT CATGC AACAT CTCCT CTGGG GAAAT  
 651 GCAGA AGGGC CTGTG GGAGC AGTGC CAGCT TCTGA AGAGC ACCTC GGTGT  
 701 TTGCC CGCTG CCACCA CTCTG GTGGA CCCCC AGCCT TTTGT GGCCC TGTGT  
 751 GAGAA GACTT TGTGT GAGTG TGCTG GGGGG CTGGA GTGCG CCTGC CCTGC  
 801 CCTCC TGGAG TACGC CCGGA CCTGT GCCCA GGAGG GAATG GTGCT GTACG  
 851 GCTGG ACCGA CCACA GCGCG TGCAG CCCAG TGTGC CCTGC TGGTA TGAG  
 901 TATAG GCAGT GTGTG TCCCC TTGCG CCAGG ACCTG CCAGA GCCTG CACAT  
 951 CAATG AAATG TGTCA GGAGC GATGC GTGGA TGGCT GCAGC TGCCC TGAGG  
 1001 GACAG CTCCT GGATG AAGGC CTCTG CGTGG AGAGC ACCGA GTGTC CCTGC  
 1051 GTGCA TTCCG GAAAG CGCTA CCCTC CCGGC ACCTC CCTCT CTCGA GACTG  
 1101 CAACA CCTGC ATTTG CCGAA ACAGC CAGTG GATCT GCAGC AATGA AGAAT  
 1151 GTCCA GGGGA GTGCC TTGTG ACTGG TCAAT CCCAC TTCAA GAGCT TTGAC  
 1201 AACAG ATACT TCACC TTCAG TGGGA TCTGC CAGTA CCTGC TGCCC CGGG  
 1251 TTGCC AGGAC CACTC CTTC CATT GTCAT TGAGA CTGTC CAGTG TGCTG  
 1301 ATGAC CGCGA CGCTG TGTGC ACCCG CTCCG TCACC GTCCG GTCG CTGGC  
 1351 CTGCA CAACA GCCTT GTGAA ACTGA AGCAT GGGGC AGGAG TTGCC ATGGA  
 1401 TGGCC AGGAC ATCCA GCTCC CCCCT CTGAA AGGTG ACCTC CGCAT CCAGC  
 1451 ATACA GTGAC GGCCT CCGTG CGCCT CAGCT ACAGG GAGGA CCTGC AGATG  
 1501 GACTG GGATG GCGCG GGGAG GCTGC TGGTG AAGCT GTCCC CGCTC TATGC  
 1551 CGGGG AGACC TGCGG CCTGT GTGGG ATTAA CAATG GCAAC CAGGG CGACG  
 1601 ACTTC CTTAC CCCCT CTGGG CTGGC GGAGC CCCGG GTGGA GGACT TCAGG  
 1651 AACGC CTGGA AGCTG CACGG GGACT GCCAG GACCT GCAGA AGCAG CACAG  
 1701 CGATC CCTGC GCCCT CAACC CGCGC ATGAC CAGGT TCTCC GAGGA GGGGT  
 1751 CGCGG GTCCT GACGT CCCCC ACATT CGAGG CCTGC CATCG TGCCC TCAGC  
 1801 CCGCT GCCCT ACCTG CGGAA CTGCC GCTAC GACGT GTGCT CCTGC TCGGA  
 1851 CGGGC GCGAG TGCCCT GTGCG CGGCC CTGGC CAGCT ATGCC CGGGC CTGG  
 1901 CGGGG AGAGG CGTGC GCGTC CGGTG CGCGG AGCCA GGCGG CTGTG AGCTG  
 1951 AACTG CCCGA AAGGC CAGGT GTACC TGCAG TGCGG GACCC CCTGC AACCT  
 2001 GACCT GCGCG TCTCT CTCTT ACCCG GATGA GGAAT GCAAT GAGGC CTGCC  
 2051 TGGAG GGCTG CTCTC GCCCC CCAGG GCCT ACATG GATGA GAGGG GGGAC  
 2101 TGCGT GCCCC AGGCC CAGTG CCCCT GTTAC TATGA CGGTG AGATC TTCAA  
 2151 GCCAG AAGAC ATCTT CTCAG ACCATCACAC CATGT GCTAC TGTGA GGATG  
 2201 GCTTC ATGCA CTGTA CCATG AGTGG AGTCC CCGGA AGCTT GTCG CTGAC  
 2251 GCTGT CCTCA GCACT CCCCT GTCTC ATCGC AGCAA AAGGA GCCTA TCCGT  
 2301 TCGGC CCCCC ATGGT CAAGC TGGTG TGTC CGCTG ACAAC CTGCG GGCTG  
 2351 AAGGG CTCGA GTGTA CCAAA ACGTG CCAGA ACTAT GACCT GGAGT GCATG  
 2401 AGCAT GGGCT GTGTC TCTGG CTGCC TCTGC CCCCC GGGCA TGGTC CGGCA  
 2451 TGAGA ACAGA TGTGT GGGCC TGGAA AGGTG TCCCT GCTTC CATCA GGGCA  
 2501 AGGAG TATGC CCGTG GAGAA ACAGT GAAGA TTGGC TGCAA CACCT GTGTC  
 2551 TGTCG GGACC GGAAG TGGAA CTGCA CAGAC CATGT GTGTT ATGCC ACGTG  
 2601 CTCCA CGATC GGCGAT GGGCC ACTAC CTAC CCTAC CTTCG ACAGG CTCAA ATACC  
 2651 TGTTC CCCCC GGAGT GGGAG TACGT TCTGG TGCG GATTA CTGCG GCAGT  
 2701 AACCC TGGGA CCTTT CGGAT CCTAG TGGGG AATAA GGGAT GCAGC CACCC  
 2751 CTCAG TGAAA TGCAA GAAAC GGGTC ACCAT CCTGG TGGAG GGAGG AGAGA  
 2801 TTGAG CTGTT TGACG GGGAG GTGAA TGTGA AGAGG CCCAT GAAGG ATGAG  
 2851 ACTCA CTTTC AGGTG GTGGA GTCTG GCCGG TACAT CATTC TGCTG CTGG  
 2901 CAAAG CCCTC TCCGT GGTCT GGGAC CGCCA CCTGA GCATC TCCGT GGTCC  
 2951 TGAAG CAGAC ATACC AGGAG AAGAT GTGTC GCGCTG TGTGG GAATT TTGAT  
 3001 GGCGT CCAGA AACAT GACCT CACCA GCAGC AACCT CCAAG TGGAG GAAGA  
 3051 CCCTG TGGAC TTTGG GAACT CCTGG AAGAT GAGCT CGCAG TGTGC TGACA  
 3101 CCAGA AAAGT GCCTC TGGAC TCATC CCCTG CCACC TGCCA TAACA ACATC  
 3151 ATGAA GCAGA CGATG GTGGA TTCTT CCTGT AGAAT CCTTA CCAGT GACGT  
 3201 CTTCC AGGAC TGCAA CAAGC TGGTG GACCC CGAGC CATAT CTGGA TGTCT  
 3251 GCATT TACGA CACCT GCTCC TGTGA GTCCA TTGGG GACTG CGCCG CATTC

3301 TGCAG CACCA TTGCT GCCTA TGCCC ACGTG TGTGC CCAGC ATGGC AAGGT  
 3351 GGTGA CCTGG AGGAC GGCA CATTG TGCCC CCAGA GCTGC GAGGA GAGGA  
 3401 ATCTC CGGGA GAACG GGTAT GAGGC TGAGT GGCAG TATAA CAGCT GTGCA  
 3451 CCTGC CTGTC AAGTC ACGTG TCAGC ACCCT GAGCC ACTGG CCTGC CCTGT  
 3501 GCAGT GTGTG GAGGG CTGCC ATGCC CACTG CCCTC CAGGG AAAAT CCTGG  
 3551 ATGAG CTTTT GCAGA CCTGC GTTGA CCCTG AAGAC TGTC AGTGT GTGAG  
 3601 GTGCC TGGCC GGCCT TTTGC CTCAG GAAAG AAAGT CACCT TGAAT CCCAG  
 3651 TGACC CTGAG CACTG CCAGA TTTGC CACTG TGATG TTGTC AACCT CACCT  
 3701 GTGAA GCCTG CCAGG AGCCG ATATC GGTTA CCTCA GAGTC TGCTA CCCCC  
 3751 GAGTC AGGGC CAGGA TCAGA GCCAG CCACC TCCGG GTCTG AGACA CCCGG  
 3801 GACTT CCGAG AGTGC CACCC CTGAG TCCGG ACCCG GGTC GAGCC CGCCA  
 3851 CTTCC GGCTC CGAAA CTCCC GGCAC AAGCG AGAGC GCTAC CCCAG AGTCA  
 3901 GGACG AGGAA CATCT ACAGA GCCCT CTGAA GGCTC CGCTC CAGGG CCCCC  
 3951 AGCCG GCAGT CCCAC TAGCA CCGAG GAGGG AACCT CTGAA AGCGC CACAC  
 4001 CCGAA TCAGG CCGG GGTCT GAGCC TGCTA CCAGC GGCAG CGAGA CACCA  
 4051 GGCAC CTCTG AGTCC GGCAC ACCAG AGTCC GGACC CGGAT CTCCC GCTGG  
 4101 GAGCC CCACC TCCAC TGAGG AGGGG TCTCC TGCTG GCTCT CCAAC ATCTA  
 4151 CTGAG GAAGG TACCT CAACC GAGCC ATCCG AGGGG TCAGC TCCCG GCACC  
 4201 TCAGA GTCGG CAACC CCGGA GTCTG GACCC GGAAC TTCCG AAAGT GCCAC  
 4251 ACCAG AGTCC GTGCC CGGG A CCTCA GAATC AGCAA CACCC GAGTC CGGCC  
 4301 CTGGG TCTGA ACCCG CCACA AGTGG TAGTG AGACA CCAGG ATCAG AACCT  
 4351 GCTAC CTCAG GTGCA GAGAC ACCCG GATCT CCGGC AGGCT CACCA ACCTC  
 4401 CACTG AGGAG GGCAC CAGCA CAGAA CCAAG CGAGG GCTCC GCACC CGGAA  
 4451 CAAGC ACTGA ACCCA GTGAG GGTTC AGCAC CCGGC TCTGA GCCGG CCACA  
 4501 AGTGG CAGTG AGACA CCCGG CACTT CAGAG AGTGC CACCC CCGAG AGTGG  
 4551 CCCAG GCACT AGTAC CGAGC CCTCT GAAGG CAGTG CGCCA GATTG TGGCG  
 4601 GTGGA GGTTC CGGTG CGGGG GGATC CGGTG CGGGG GGATC CGGTC CGGGG  
 4651 GGATC CGGTG CGGGG GGATC CCTGG TCCCC CGGGG CAGCG GAGGC GACAA  
 4701 AACTC ACACA TGCCC ACCGT GCCCA GCTCC AGAAC TCCTG GGCGG ACCGT  
 4751 CAGTC TTCCT CTTCC CCCCC AAACC CAAGG ACACC CTCAT GATCT CCCGG  
 4801 ACCCC TGAGG TCACA TGCGT GGTGG TGGA GGTGAG CCACG AAGAC CCTGA  
 4851 GGTCA AGTTC AACTG GTTAC TGGA GGCCT GGAGG TGCAT AATGC CAAGA  
 4901 CAAAG CCGCG GGAGG AGCAC TACAA CAGCA CGTAC CGTGT GGTCA GCGTC  
 4951 CTCAC CGTCC TGCAC CAGGA CTGGC TGAAAT GGCAA GGAGT ACAAG TGCAA  
 5001 GGTCT CCAAC AAAGC CCTCC CAGCC CCCAT CGAGA AAACC ATCTC CAAAG  
 5051 CCAAA GGGCA GCCCC GAGAA CCACA GGTGT ACACC CTGCC CCCAT CCCGG  
 5101 GATGA GCTGA CCAAG AACCA GGTCA GGCCTG ACCTG CCTGG TCAAA GGCTT  
 5151 CTATC CCAGC GACAT CGCCG TGGAG TGGGA GAGCA ATGGG CAGCC GGAGA  
 5201 ACAAC TACAA GACCA CGCC CCGCT GTTGG ACTCC GACGG CTCCCT TCTTC  
 5251 CTCTA CAGCA AGCTC ACCGT GGACA AGAGC AGTGT GCAGC AGGGG AACGT  
 5301 CTTCT CATGC TCCGT GATGC ATGAG GCTCT GCACA ACCAC TACAC GCAGA  
 5351 AGAGC CTCTC CCTGT CTCCG GGTAA ATGA

[0349]

VWF034 단백질 서열(서열 번호 92)

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 LWEQCQLLK TSFVARCHPL VDPEPFVALC  
 251 EKTLCECAGG LECACPALLE YARTCAQEGM VLYGWTDHSA CSPVCPAGME  
 301 YRQCVSPCAR TCQSLHINEM CQERCVDGCS CPEGQLLDEG LCVESTECPC  
 351 VHSGKRYPPG TSLSRDCNTC ICRNSQWICS NEECPGECLV TGQSHFKSFD  
 401 NRYFTFSGIC QYLLARDCQD HSFSIVIETV QCADDRAVC TRSVTVRLPG  
 451 LHNSLVKLKH GAGVAMDQD IQLPLLKGDL RIQHTVTASV RLSYGEDLQM

[0351]

501 DWDGRGRLLV KLSPVYAGKT CGLCGNYNGN QGDDFLTPSG LAEPRVEDFG  
 551 NAWKLHGDCQ DLQKQHSDPC ALNPRMTRFS EEACAVLTSP TFEACHRAVS  
 601 PLPYLRNCRY DVCSCSDGRE CLCGALASYA AACAGRGVV AWREPGRCSEL  
 651 NCPKGQVYLQ CGTPCNLTCR SLSYPDEECN EACLEGCFCP PGLYMDERGD  
 701 CVPKAQCPCY YDGEIFQPED IFSDHHTMCY CEDGFMHCTM SGVPGSLLPD  
 751 AVLSSPLSHR SKRSLSCRPP MVKLVCPADN LRAEGLECTK TCQNYDLECM  
 801 SMGCVSGCLC PPGMVRHENR CVALERCPCF HQGKEYAPGE TVKIGCNTCV  
 851 CRDRKWNTD HVCDATCSTI GMAHYLTFDG LKYLFPGECQ YVLVQDYCGS  
 901 NPGTFRILVG NKGCSHPSVK CKKRVTILVE GGEIELFDGE VNVKRPMKDE  
 951 THFEVVESGR YIILLLGKAL SVVWDRHLSI SVVLKQTYQE KVCGLCGNFD  
 1001 GIQNNDLTSS NLQVEEDPVD FGNSWKVSSQ CADTRKVPLD SSPATCHNNI  
 1051 MKQTMVDS SC RILTSDFQD CNKLVDPEPY LDVCIYDTCS CESIGDCAAF  
 1101 CDTIAAYAHV CAQHGKVVTW RTATLCQSC EERNLRENGY EAEWRYNSCA  
 1151 PACQVTCQHP EPLACPVQCV EGCHAHCPPG KILDELLQTC VDPEDCPVCE  
 1201 VAGRFFASGK 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 TCVVVVDVSHE  
 1601 DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL HQDWLNGKEY  
 1651 KCKVSNKALP APIEKTIKA KGQPREPQVY TLPPSRDELT KNQVSLTCLV  
 1701 KGFYPSDI AV EWESNGQOPEN NYKTTPPVLD SDGSFFLYSK LIVDKSRWQQ  
 1751 GNFSCSVMH EALHNHYTQK SLSLSPGK\*

[0352]

[0353] 구체적인 구현예의 앞서의 기술은 본 발명의 일반적인 특성을 충분히 나타내므로, 다른 것은 당해 분야의 기술 내 지식을 적용함으로써, 본 발명의 일반적인 개념으로부터 벗어남이 없이, 과도한 실험 없이도, 구체적인 구현 예와 같은 다양한 적용에 대해 용이하게 개질시키고/시키거나 조정할 수 있다. 따라서, 이러한 조정 및 변형은 본원에 나타낸 교시 및 안내를 기본으로 하여, 개재된 구현예의 등가물의 의미 및 범위 내에 있는 것으로 의도된다. 본원의 어법 또는 학술용어는 설명의 목적이며 제한하는 것이 아니므로, 본 명세서의 학술용어 또는 어법은 교시 및 안내의 측면에서 숙련가에 의해 해석되어야 함이 이해되어야 한다.

[0354]

본 발명의 다른 실시예는 본원에서 개시된 본 발명의 명세서와 실시의 고려로부터 기술 분야에서의 당업자에게 명백할 것이다. 명세서 및 실시에는 예시로서만 고려되며, 본 발명의 실제 영역 및 취지는 다음의 특허청구범위에 의해 나타나는 것으로 의도된다.

[0355]

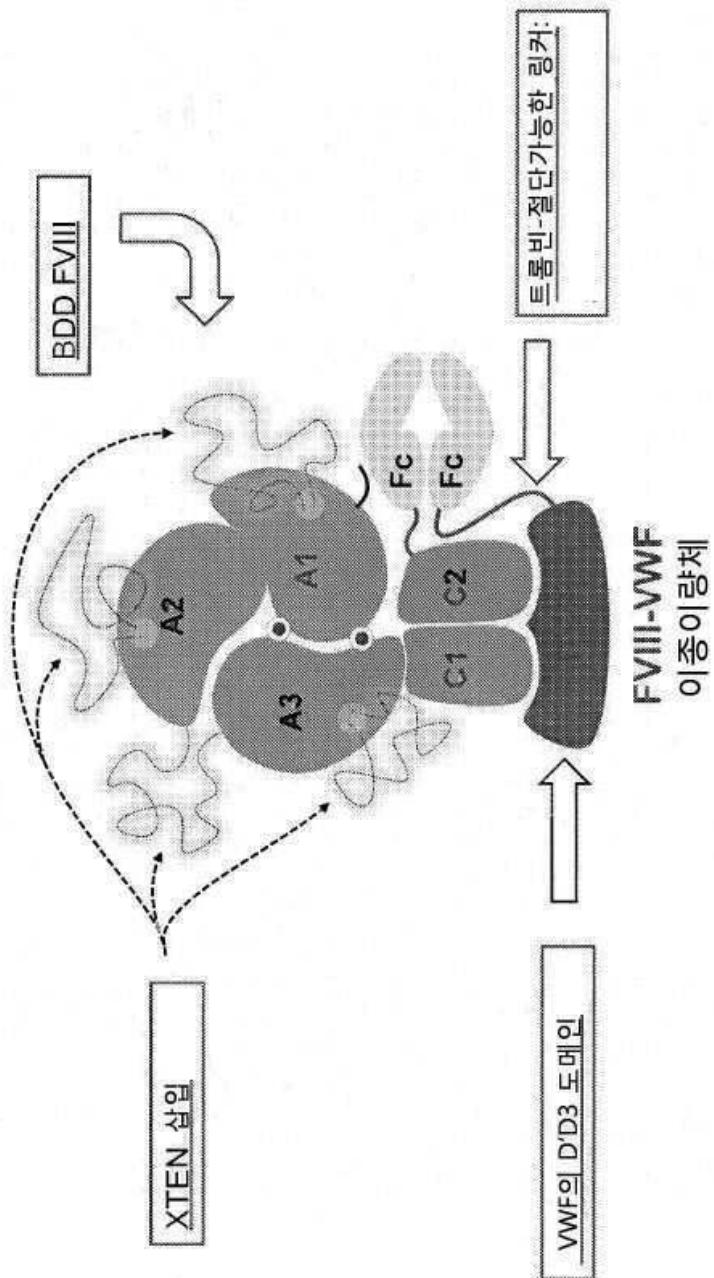
본원에 인용된 모든 특허 및 공보는, 이의 전문이 본원에 참조로 포함된다.

[0356]

본 출원은 2013년 6월 28일자로 출원된 미국 가특허원 제61/840,872호의 우선권의 이익을 청구한다. 상기 출원의 내용은, 이의 전문이 참조로 포함된다.

도면

도면1

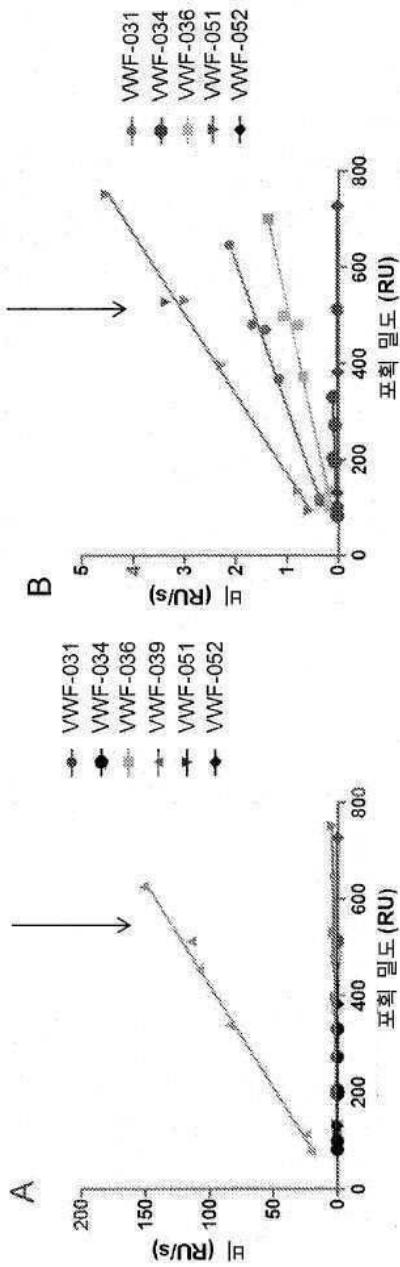


## 도면2

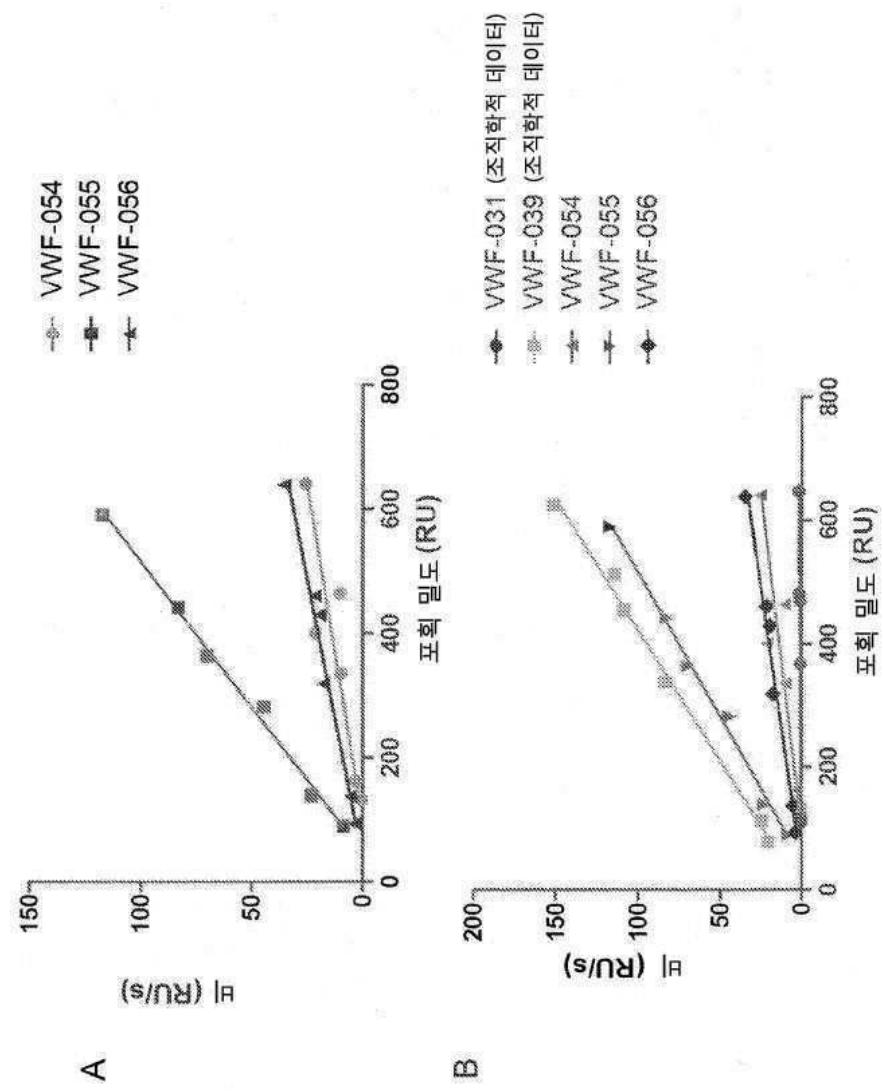
작제물	VWF-D'D3와 Fc 사이의 링커의 유형 및 길이
VWF-031	48aa (LVPR 부위)
VWF-034	288 XTEN+ 35aa (LVPR 부위)
VWF-035	73aa (LVPR 부위)
VWF-036	98aa (LVPR 부위)
VWF-039	26aa (LVPR+PAR1 부위)
VWF-051	54aa (ALRPRVV 부위)
VWF-052	48aa (트롬번 부위 번호)
VWF-054	40aa (FVIII으로부터의 a1 거의 없음)
VWF-055	34aa (FVIII으로부터의 a2 거의 없음)
VWF-056	46aa (FVIII으로부터의 a3 거의 없음)

## 도면3

절단률은 출발하는 포획 수준에 비례한다.  
기울기는 각각의 작제물에 대한 트롬빈 절단에 대한 민감성의 척도를 제공한다.



도면4



## 도면5

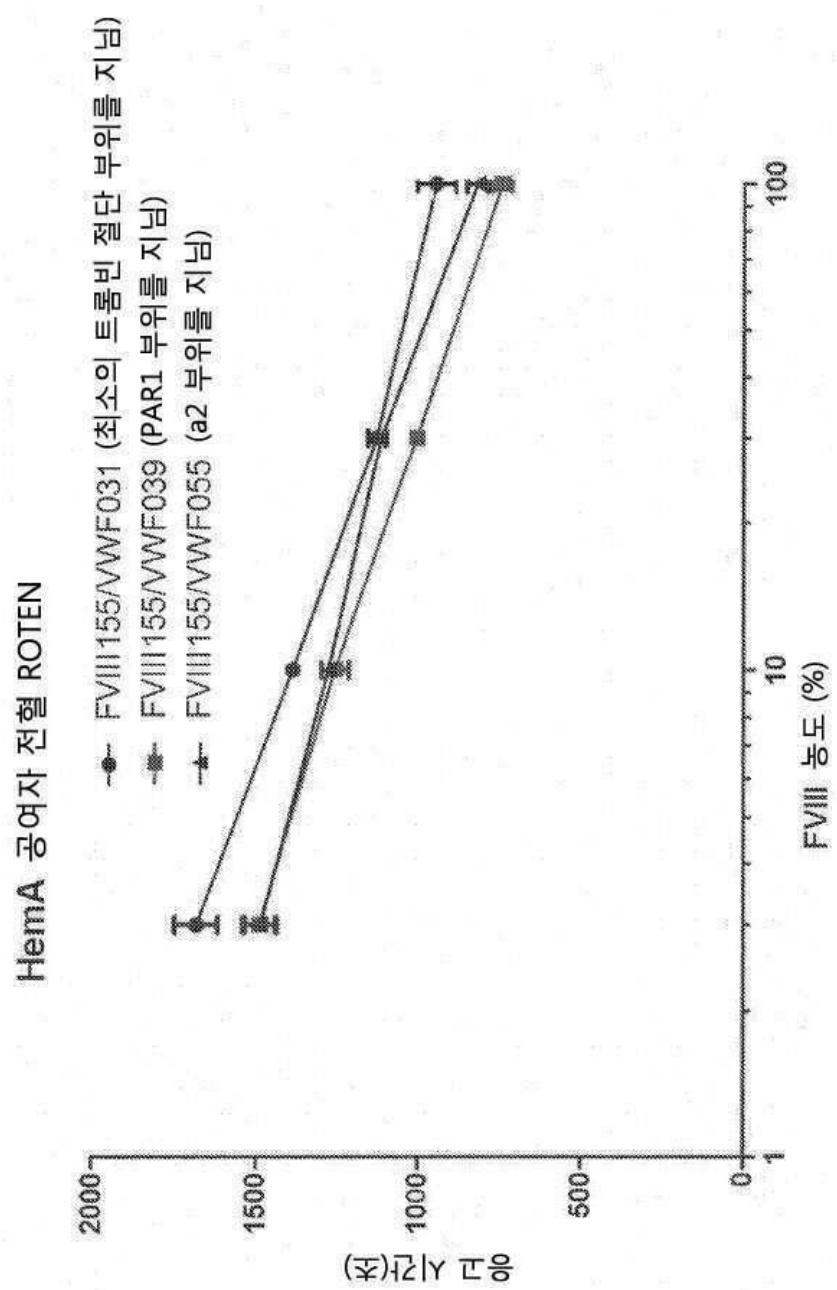
작제물	기울기 (°)
VWF-031	$3.32 \times 10^{-3} \pm 0.16 \times 10^{-3}$
VWF-034	$0.31 \times 10^{-3} \pm 0.085 \times 10^{-3}$
VWF-036	$2.06 \times 10^{-3} \pm 0.18 \times 10^{-3}$
VWF-039	$235.4 \times 10^{-3} \pm 9.1 \times 10^{-3}$
VWF-051	$6.00 \times 10^{-3} \pm 0.27 \times 10^{-3}$
VWF-052	$0.02 \times 10^{-3} \pm 0.006 \times 10^{-3}$
VWF-054	$46.35 \times 10^{-3} \pm 11.53 \times 10^{-3}$
VWF-055	$213.5 \times 10^{-3} \pm 8.5 \times 10^{-3}$
VWF-056	$55.6 \times 10^{-3} \pm 4.5 \times 10^{-3}$

기울기 VWF-039 / 기울기 VWF-031 = 71

기울기 VWF-055 / 기울기 VWF-031 = 65

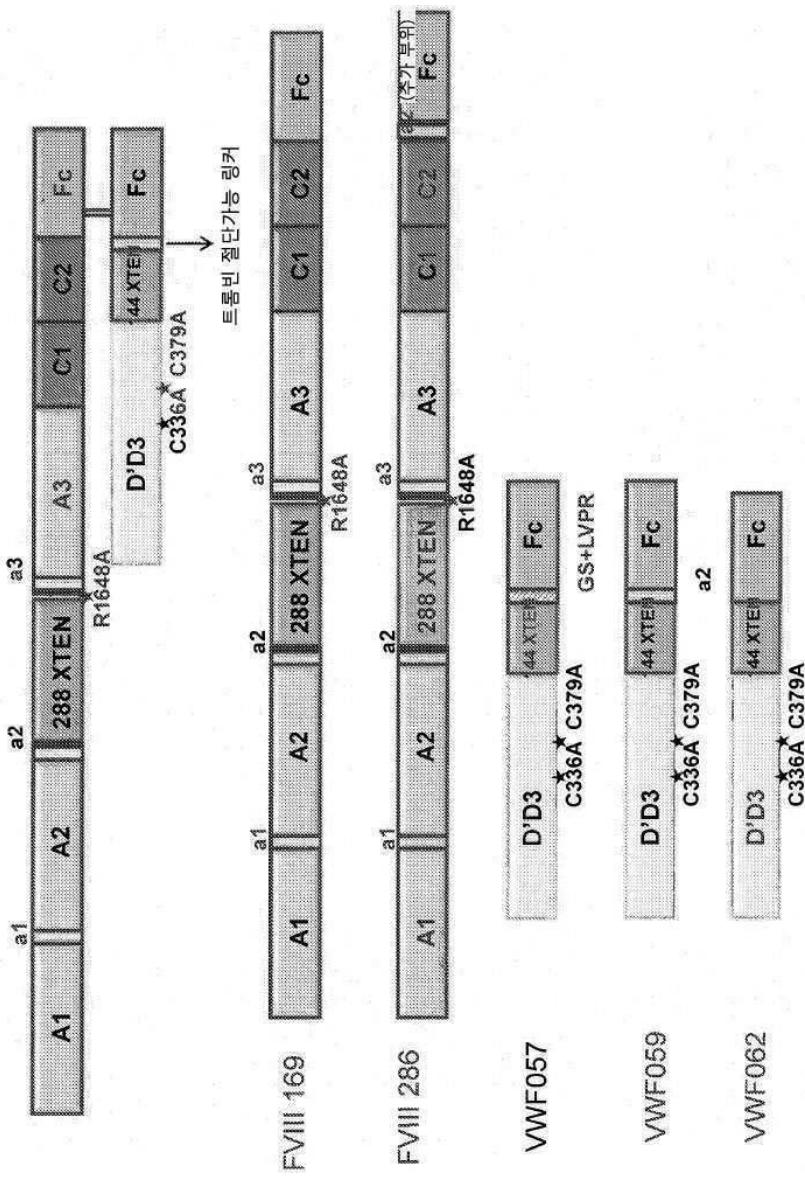
기울기 VWF-051 / 기울기 VWF-031 = 1.8

## 도면6

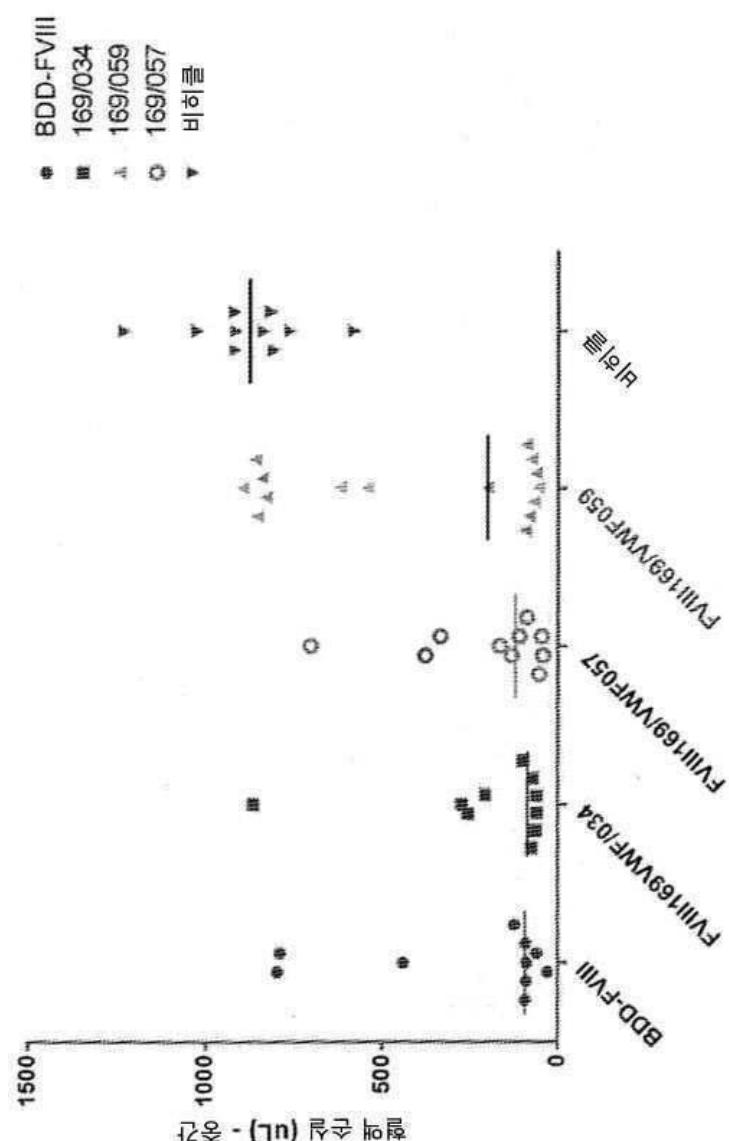


## 도면7

대표적인 FVIII-VWF 이종이형체의 막대 그래프



## 도면8



## 서열 목록

## SEQUENCE LISTING

&lt;110&gt; Biogen IDEC MA INC.

&lt;120&gt; THROMBIN CLEAVABLE LINKER WITH XTEN AND ITS USES THEREOF

&lt;130&gt; 2159.420PC01

&lt;140&gt; To be assigned

&lt;141&gt; 2014-06-27

&lt;150&gt; 61/840,872

&lt;151&gt; 2013-06-28

&lt;160&gt; 92

&lt;170&gt; PatentIn version 3.5

<210> 1  
<211> 8442  
<212> DNA  
<213> Homo sapiens  
<400> 1

atgattcctg ccagattgc cgggtgtcg ctgtctgg ccctcattt gccagggacc	60
ctttgtcag aaggaactcg cgccaggta tccacggccc gatgcagcct tttcgaaagt	120
gacttcgtca acaccttga tggagcatg tacagcttg cggatactg cagtacctc	180
ctggcagggg gctgccagaa acgctcctc tcgattattg gggacttcca gaatggcaag	240
agagttagcc tctccgtta tcttggaa tttttgaca tccattgtt tgtcaatgg	300
accgtgacac agggggacca aagagtctcc atgcctatg cctccaaagg gctgtatcta	360
gaaactgagg ctgggtacta caagctgtcc ggtgaggcct atggcttgc ggccaggatc	420
gatggcagcg gcaacttca agtcctgctg tcagacagat acttcaacaa gacctgcggg	480
ctgtgtggca acttaacat ctttgcgtaa gatgacttta tgacccaaga agggaccttgc	540
acctcggacc ctatgactt tgccaaactca tggctctga gcagtggaga acagtggtgt	600
gaacgggcat ctctccag cagctcatgc aacatctcct ctggggaaat gcagaaggc	660
ctgtgggagc agtgcagct tctgaagagc acctcggtgt ttgcccgtc ccacctctg	720
gtggaccccg agcctttgt gcccctgtgt gagaagactt tgtgtgagtg tgctgggggg	780
ctggagtgcg cctgcctgc ctcctggag tacgcccggc cctgtgccca ggagggaatg	840
tgctgtacg gctggaccga ccacagcgcg tgcagccag tgcctgc tggtatggag	900
tataggcagt gtgtgtcccc ttgcgcagg acctgcaga gcctgcacat caatgaaatg	960
tgtcaggagc gatgcgtgga tggctgcagc tgccctgagg gacagctctt ggtatggc	1020
ctctgcgtgg agaggaccga gtgtccctgc gtgcattccg gaaagcgcta ccctccggc	1080
acctccctct ctcgagactg caacacctgc atttgcgaa acagccagt gatctgcagc	1140
aatgaagaat gtccagggga gtgcctgtc actggtaat cccacttcaa gagcttgac	1200
aacagatact tcacccctcg tggatctgc cagttacctgc tggccggga ttgcaggac	1260
cactcctct ccattgtcat tgagactgtc cagttgtgtc atgaccgcga cgctgtgtgc	1320
acccgcctcg tcaccgtccg gtcgcctggc ctgcacaaca gccttgaa actgaagcat	1380
ggggcaggag ttgcgtgga tggccaggac atccagctcc ccctcctgaa aggtgaccc	1440
cgcatccagc atacagtgac ggcctccgtc cgcctcagct acggggagga cctgcagatg	1500
gactggatg gcccgggag gtcgtggc aagctgtccc ccgtctatgc cggaaagacc	1560

tgcggcctgt gtgggaatta caatggcaac cagggcgacg acttccttac cccctctggg	1620
ctggcrgagc cccgggtgga ggacttcggg aacgcctgga agctgcacgg ggactgccag	1680
gacctgcaga agcagcacag cgatccctgc gccctcaacc cgcgcatgac cagttctcc	1740
gaggaggcgt gcgcgtcct gacgtcccc acatcgagg ccigccatcg tgccgtcagc	1800
ccgctgcctt acctgcgaa ctgccgtac gacgtgtgct cctgctcgga cggccgcgag	1860
tgccctgtcg gcgcctggc cagctatgcc gcggcctgcf cggggagagg cgtgcgcgtc	1920
gcgtggcgcg acccaggccg ctgtgagctg aactgcccga aaggccaggt gtacctgcag	1980
tgcgggaccc cctgcaacct gacctgccgc tctctcttt acccgatga ggaatgcaat	2040
gaggccgtgcc tggaggcgtt cttctgcccc ccaggctct acatggatga gagggggac	2100
tgcgtgcca aggcccagtg cccctgttac tatgacggtg agatcttcca gccagaagac	2160
atcttctcag accatcacac catgtgtac tacatgtgtt gttcatgca ctgtaccatg	2220
agtggagtcc ccgaaagctt gctgcctgac gctgtcctca gcagtcccgt gtctcatcg	2280
agcaaaaagga gcctatcctg tcggccccc atggtaagc tgggtgttcc cgctgacaac	2340
ctgcggctg aagggtcgaa gtgtaccaa acgtgccaga actatgaccc ggagtgcatt	2400
agcatgggct gtgtctctgg ctgcctctgc ccccccggca tggccggca tgagaacaga	2460
tgtgtggccc tggaaagggtg tccctgttcc catcaggcga aggagtatgc ccctggagaa	2520
acagtgaaga ttggctgcaa cacttgttgc tgtcgggacc ggaagtggaa ctgcacagac	2580
catgtgtgt atgcacgtg ctccacgtac ggcatggccc actacccatc cttcgacggg	2640
ctcaaatacc tttccccgg ggagtgccag tacgttctgg tgcaggatta ctgcggcagt	2700
aaccctggga ctttcggat cctagtgggg aataaggat gcagccaccc ctcagtgaaa	2760
tgcaagaaac gggcacat cctgggtggag ggaggagaga ttgagctgtt tgacggggag	2820
gtgaatgtga agaggccat gaaggatgag actcaacttg aggtggtgga gtctggccgg	2880
tacatcatcc tgcgtctggg caaagccctc tccgtggctt gggaccgcca cctgagcatc	2940
tccgtggtcc tgaaggcagac ataccaggag aaagtgtgtt gcctgtgtgg gaattttgt	3000
ggcatccaga acaatgaccc caccagcgc aacccccaag tggaggaaga ccctgtggac	3060
tttggaaact cctggaaagt gagctcgacat tggctgacat ccagaaaagt gcctctggac	3120
tcatccccctg ccacctgcca taacaacatc atgaagcaga cgtgggttca ttccctctgt	3180
agaatccta ccagtgcacgt cttccaggac tgcaacaagc tgggtggaccc cgagccatat	3240
ctggatgtct gcattacga cacctgttcc tggatgttca ttggggactg cgcctgttcc	3300
tgcgacacca ttgtgccta tgcccacgtg tggccacgc atggcaaggt ggtgacctgg	3360
aggacggcca cattgtgccc ccagagctgc gaggagagga atctccggga gaacgggtat	3420

gagtgtgagt ggcgctataa cagctgtca cctgcctgtc aagtcacgtg tcagcaccc	3480
gagccactgg cctgcctgt gcagtgtgt gagggctgcc atgcccactg ccctccaggg	3540
aaaatcctgg atgagcttt gcagacctgc gttgaccctg aagactgtcc agtgtgtgag	3600
gtggctggcc ggcgtttgc ctcaggaaag aaagtacct tgaatccag tgaccctgag	3660
cactgccaga ttgccactg tcatgttgc aacacctcac gtgaaggctg ccaggagccg	3720
ggaggcctgg tgggcctcc cacagatgcc ccggtgagcc ccaccactct gtatgtggag	3780
gacatctcg aaccgccgtt gcacgattt tactgcagca ggctactgga cctggcttc	3840
ctgctggatg gctccctccag gctgtccag gctgagttt aagtgcgtaa ggccttgc	3900
gtggacatga tggagccgt ggcacatctcc cagaagtggg tccgcgtggc cgtggtgag	3960
taccacgacg gctcccacgc ctacatcggt ctcaaggacc ggaagcgacc gtcagagctg	4020
cgccgcattt ccagccaggt gaagtatgcg ggcagccagg tggcctccac cagcgaggtc	4080
ttgaaataaca cactgttcca aatcttcagc aagatcgacc gccctgaagc ctcccgatc	4140
gccctgtcc tcatggccag ccaggagccc caacggatgt cccggaactt tgtccgtac	4200
gtccaggggcc tgaagaagaa gaaggtcatt gtatcccgg tggcattgg gccccatgcc	4260
aacctcaagc agatccgcct catcgagaag caggccccgt agaacaaggc cttcgtgctg	4320
agcagtgtgg atgagctgga gcagcaaagg gacgagatcg ttagctaccc ctgtgaccc	4380
ccccctigaag cccctccctcc tactctgccc cccgacatgg cacaagtac tgtggcccg	4440
gggcctttgg ggtttcgac cctggggccc aagaggaact ccatggttct ggatgtggcg	4500
ttcgtcttgg aaggatcgga caaaaatttgtt gaagccact tcaacaggag caaggagttc	4560
atggaggagg tgattcagcg gatggatgtg ggccaggaca gcatccacgt cacggtgctg	4620
cagtaactct acatggtgac cgtggagtac cccttcagcg aggcacagtc caaagggac	4680
atcctgcagc gggtgcgaga gatccgtac cagggcggca acaggaccaa cactggctg	4740
gccctgcggggtt accctctga ccacagcttc ttggtcagcc agggtgaccg ggagcaggcg	4800
cccaacactgg tctacatggt caccggaaat cctgcctctg atgagatcaa gaggctgcct	4860
ggagacatcc aggtggtgcc cattggagtg ggccctaattt ccaacgtgca ggagctggag	4920
aggattggctt ggccaaatgc ccctatccctc atccaggact ttgagacgct ccccccggag	4980
gctccctgacc tggtgctgca gaggtgctgc tccggagagg ggctgcagat ccccaccc	5040
tccctgcac ctgactgcag ccagccccgt gacgtatcc ttctccctgg tggctccctcc	5100
agtttccag ctcttattt tcatgaaatg aagagttcg ccaaggctt cattcaaaa	5160

gccaatata ggcctcg tca cactcagg tgc tcagtgc agt atggaag catcaccacc	5220
attgacgtgc catggAACgt ggtcccggag aaagcccatt tgctgagcct tgtggacgtc	5280
atgcagcggg agggaggccc cagccaaatc gggatgcct tggcttgc tgtgcgatac	5340
ttgacttcag aaatgcatgg tgccaggccg ggagcctcaa aggccgttgt catcctggtc	5400
acggacgtct ctgtggattc agtggatgca gcagctgatg ccgcaggc caacagatg	5460
acagtgtcc ctattgaat tggagatgca tacgatgcag ccagctacg gatctggca	5520
ggcccagcag gcgactccaa cgtggtaag ctccagcga tcgaagacct ccctaccatg	5580
gtcaccttgg gcaattcctt cctccacaaa ctgtgctctg gatttggtag gatttgcatg	5640
gatgaggatg ggaatgagaa gagccccggg gacgtctgga cttgccaga ccagtgcac	5700
accgtgactt gccagccaga tggccagacc ttgtgaaga gtcatcggt caactgtgac	5760
cggggctga ggccttcgtg ccctaacagc cagtcctg ttaaagtggaa agagacctgt	5820
ggctgcccgtt ggacctgccc ctgygtgtgc acaggcact ccactcgca catcgacc	5880
tttgcgtggc agaattcaa gctgactggc agctgttctt atgccttatt tcaaaaacaag	5940
gagcaggacc tggaggtgat tctccataat ggtgcctgca gcctggagc aaggcaggc	6000
tgcataaat ccatcgaggt gaagcacagt gcctctccg tcgagstgca cagtgacatg	6060
gaggtgacgg tgaatggag actggctctt gttccttacg tgggtggaa catggaagtc	6120
aacgtttatg gtgcattcat gcatgagtc agatcaatc accttggta catcttaca	6180
ttcactccac aaaacaatga gttccaaactg cagtcagcc ccaagacttt tgcttcaaag	6240
acgtatggtc tggatggat ctgtgtgatg aacggagcca atgacttcat gctgaggat	6300
ggcacagtca ccacagactg gaaaacactt gttcaggaat ggactgtgca gcggccagg	6360
cagacgtgcc agcccatctt ggaggagcag tggatgtcc ccgacagctc ccactgccc	6420
gtcctcctt taccactgtt tgctgaatgc cacaaggccc tggctccagc cacatttat	6480
gccatctgcc agcaggacag ttgccaccag gagcaagtgt gtgaggatgat cccttatt	6540
gcccacccctt gtggaccaa cggggctgc gttgactgga ggacacctga ttctgtgt	6600
atgtcatgcc caccatctt ggtctacaac cactgtgagc atggctgtcc ccggcactgt	6660
gatggcaacg tggatgtcc tggggaccat ccctccgaag gctgttctg ccctccagat	6720
aaagtcatgt tggaaaggcag ctgtgtccct gaagaggcct gcactcagtg cattggtag	6780
gatggatgcc agcaccatgtt cctggaaagcc tgggtccccc accaccagcc ctgtcagatc	6840
tgcacatgcc tcagcggcgcg gaaggtcaac tgcacaacgc agccctgccc cacggccaaa	6900

gctcccacgt gtggcctgtg tgaagtagcc cgccctcgcc agaatgcaga ccagtgtgc	6960
cccagatgt agtgtgtgt tgacccagtg agctgtgacc tgccccagt gcctactgt	7020
gaacgtggcc tccagccac actgaccaac cctggcgagt gcagacccaa ctacacctgc	7080
gcctgcagga aggaggagtg caaaaagagtg tccccccct ccggccccc gcaccgttg	7140
cccacccttc ggaagaccca gtgctgtat gagtatgagt gtgcctgcaa ctgtgtcaac	7200
tccacagtga gctgtccct tgggtacttg gcctcaaccg ccaccaatga ctgtggctgt	7260
accacaacca cctgccttcc cgacaagggtg tgtgtccacc gaagcaccat ctaccctgtg	7320

ggccagttct gggaggaggg ctgcgtatgt tgccacgtgca ccgacatggaa ggatgccgtg	7380
atgggcctcc gctggccca gtgctcccg aagccctgtg aggacagctg tcggtcggc	7440
ttcacttacg ttctgcataa aggccgtgc tgtgaaaggt gcctgccatc tgccgtgag	7500
gtggtgactg gctcaccgcg gggggactcc cagtttccct ggaagaggtgt cggctccag	7560
tggcctccc cggagaaccc ctgcctcatc aatgagtgtg tccgagtgaa ggaggaggc	7620
tttataacaac aaaggaacgt ctccgtcccc cagctggagg tccctgtctg cccctcgcc	7680
ttttagctga gctgtaaagac ctcaagcgtgc tgcccaagct gtgcgtgtga ggcgtatggag	7740

gcctgcataatgc tcaatggcac tgtcattggg cccggaaaga ctgtgtatgat cgtatgtgtgc	7800
acgacccgtgcc gctgcataatgtt gcagggtgggg gtcataatctg gattcaagct ggagtgcagg	7860
aagaccacccat gcaacccctg cccctgggt tacaaggaag aaaataaac acgtgtatgt	7920
tgtggagat gtttgctac ggcttgcacc attcagctaa gaggaggaca gatcatgaca	7980
ctgaagcgtg atgagacgct ccaggatggc tgtgataactc acttctgcaa ggtcaatgag	8040
agaggagagt acttctggaa gaagagggtc acaggctgcc cacccttga tgaacacaag	8100
tgtcttgcgtg agggaggtaa aattatgaaa attccaggca cctgcgtgtga cacatgtgag	8160

gagcctgagt gcaacgcacat cactgccagg ctgcgtatgt tcaagggtggg aagctgtaa	8220
tctgaagtag aggtggatccactactgc caggcaaat gtgccagcaa agccatgtac	8280
tccattgaca tcaacgtgt gcaggaccag tgctctgtc gctctccgac acggacggag	8340
cccatgcagg tggccctgca ctgcaccaat ggctctgttgc tgtaccatga gggttctcaat	8400
gccatggagt gcaaattgtc ccccaggaag tgcagcaagt ga	8442

&lt;210&gt; 2

&lt;211&gt; 2813

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (2016)..(2016)

<223> where Xaa can be any amino acid other than cysteine

<400> 2

Met Ile Pro Ala Arg Phe Ala Gly Val Leu Leu Ala Leu Ala 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	Glu	Phe	Val
Ala	Leu	Cys	Glu
Lys	Thr	Leu	Cys
Glu			
245	250	255	
Cys	Ala	Gly	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
Ser	Asp	Phe	Asp
385	390	395	400
Asn	Arg	Tyr	Phe
Thr	Phe	Ser	Gly
Ile	Cys	Gln	Tyr
Leu	Leu	Leu	Ala
			Arg
405	410	415	
Asp	Cys	Gln	Asp
Asp	His	Ser	Phe
Ser	Ile	Val	Ile
Glu		Glu	Thr
Thr	Val	Gln	Cys
420	425	430	
Ala	Asp	Asp	Arg
Asp	Ala	Val	Cys
Thr	Arg	Ser	Val
Val	Thr	Val	Arg
			Leu
435	440	445	
Pro	Gly	Leu	His
Asn	Ser	Leu	Val
Lys	Leu	Lys	His
Gly	Ala	Gly	Val
450	455	460	
Ala	Met	Asp	Gly
Gln	Asp	Ile	Gln
Leu	Pro	Leu	Leu
Lys	Gly	Asp	Leu
465	470	475	480

Arg Ile Gln His Thr Val Thr Ala Ser Val Arg Leu Ser Tyr Gly Glu

485 490 495

Asp Leu Gln Met Asp Trp Asp Gly Arg Gly Arg Leu Leu Val Lys Leu

500 505 510

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

515 520 525

Gly Asn Gln Gly Asp Asp Phe Leu Thr Pro Ser Gly Leu Ala Glu Pro

530 535 540

Arg Val Glu Asp Phe Gly Asn Ala Trp Lys Leu His Gly Asp Cys Gln

545 550 555 560

Asp Leu Gln Lys Gln His Ser Asp Pro Cys Ala Leu Asn Pro Arg Met

565 570 575

Thr Arg Phe Ser Glu Glu Ala Cys Ala Val Leu Thr Ser Pro Thr Phe

580 585 590

Glu Ala Cys His Arg Ala Val Ser Pro Leu Pro Tyr Leu Arg Asn Cys

595 600 605

Arg Tyr Asp Val Cys Ser Cys Ser Asp Gly Arg Glu Cys Leu Cys Gly

610 615 620

Ala Leu Ala Ser Tyr Ala Ala Ala Cys Ala Gly Arg Gly Val Arg Val

625 630 635 640

Ala Trp Arg Glu Pro Gly Arg Cys Glu Leu Asn Cys Pro Lys Gly Gln

645 650 655

Val Tyr Leu Gln Cys Gly Thr Pro Cys Asn Leu Thr Cys Arg Ser Leu

660 665 670

Ser Tyr Pro Asp Glu Glu Cys Asn Glu Ala Cys Leu Glu Gly Cys Phe

675 680 685

Cys Pro Pro Gly Leu Tyr Met Asp Glu Arg Gly Asp Cys Val Pro Lys

690 695 700

Ala Gln Cys Pro Cys Tyr Tyr Asp Gly Glu Ile Phe Gln Pro Glu Asp

705 710 715 720

Ile Phe Ser Asp His His Thr Met Cys Tyr Cys Glu Asp Gly Phe Met

725	730	735
His Cys Thr Met Ser Gly Val Pro Gly Ser Leu Leu Pro Asp Ala Val		
740	745	750
Leu Ser Ser Pro Leu Ser His Arg Ser Lys Arg Ser Leu Ser Cys Arg		
755	760	765
Pro Pro Met Val Lys Leu Val Cys Pro Ala Asp Asn Leu Arg Ala Glu		
770	775	780
Gly Leu Glu Cys Thr Lys Thr Cys Gln Asn Tyr Asp Leu Glu Cys Met		
785	790	795
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 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 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> 4		
<212> PRT		

<213> Artificial Sequence

<220><223>

> Thrombin cleavage site

<220><221> MISC\_FEATURE

<222> (1)..(1)

<223> where X is an aliphatic amino acid

<400> 3

Xaa Val Pro Arg

1

<210> 4

<211> 34

<212> PRT

<213> Artificial Sequence

<220><223> a2 region

<400> 4

Ile Ser Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp

1	5	10	15
---	---	----	----

Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser

20	25	30
----	----	----

Phe Ser

<210> 5

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> a1 region

<400> 5

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

1	5	10	15
---	---	----	----

Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro

20	25	30
----	----	----

Ser Phe Ile Gln Ile Arg Ser Val

35	40
----	----

<210> 6

&lt;211&gt; 46

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; a3 region

&lt;400&gt; 6

Ile Ser Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile

1 5 10 15

Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp

20 25 30

Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln

35 40 45

&lt;210&gt; 7

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; PAR1 exosite interaction motif

&lt;400&gt; 7

Ser Phe Leu Leu Arg Asn

1 5

&lt;210&gt; 8

&lt;211&gt; 4

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; PAR1 exosite interaction motif

&lt;400&gt; 8

Pro Asn Asp Lys

1

&lt;210&gt; 9

&lt;211&gt; 5

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; PAR1 exosite interaction motif

&lt;400&gt; 9

Pro Asn Asp Lys Tyr

1 5

<210> 10

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> PAR1 exosite interaction motif

<400> 10

Pro Asn Asp Lys Tyr Glu

1 5

<210> 11

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> PAR1 exosite interaction motif

<400> 11

Pro Asn Asp Lys Tyr Glu Pro

1 5

<210> 12

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> PAR1 exosite interaction motif

<400> 12

Pro Asn Asp Lys Tyr Glu Pro Phe

1 5

<210> 13

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> PAR1 exosite interaction motif

<400>

> 13

Pro Asn Asp Lys Tyr Glu Pro Phe Trp

1 5

<210> 14

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> PAR1 exosite interaction motif

<400> 14

Pro Asn Asp Lys Tyr Glu Pro Phe Trp Glu

1 5 10

<210> 15

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Signal Peptide

<400> 15

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

1 5 10 15

Cys Phe Ser

<210> 16

<211> 2332

<212> PRT

<213> Homo sapiens

<400> 16

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 Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp

340 345 350

Ser Glu Met Asp Val Val Arg Phe 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 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
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
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
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 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> 17

<211>

> 7053

<212> DNA

<213> Homo sapiens

<400> 17

atgcaaata	60
agctctccac	
ctgcttcttt	
ctgtgcctt	
tgcgattctg	
cttttagtgcc	
accagaagat	120
actacctggg	
tgcagtggaa	
ctgtcatggg	
actatatgca	
aagtgtatc	
ctgtggacgc	
aagatttcct	180
cctagagtgc	
caaaatctt	
tccattcaac	
acctcagtcg	240
tgtacaaaaa	
gactctgttt	
gtagaattca	
cggatcacct	
tttcaacatc	
gctaaggcca	300
ggccaccctg	
gatgggtctg	
ctaggtccta	
ccatccaggc	
tgaggtttat	
gatacagtgg	360
tcattacact	
taagaacatg	
gcttcccata	
ctgtcagtct	
tcatgcttt	

ggtgtatcct	420
actggaaagc	
ttctgaggga	
gctgaatatg	
atgatcagac	
cagtcaaagg	
gagaaaagaag	480
atgataaagt	
cttccctgg	
ggaagccata	
catatgtctg	
gcaggcctg	
aaagagaatg	540
gtccaatggc	
ctctgaccca	
ctgtgcctt	
cctactcata	
tctttctcat	
gtggacacctg	600
taaaagactt	
gaattcaggc	
ctcattggag	
ccctactagt	
atgtagagaa	
gggagtcgtgg	660
ccaaggaaaa	
gacacagacc	
ttgcacaaat	
ttatactact	
ttttgctgta	



atacaaaaatg tctcctctag tgatttgg atgccttc gacagagtcc tactccat	2460
gggctatcct tatctgatct ccaagaagcc aaatatgaga cttttctga tgatccatca	2520
cctggagcaa tagacagtaa taacagcccg tctgaaatga cacacttcg gccacagctc	2580
catcacagt gggacatggt atttaccct gagtcaggcc tccaattaag attaaatgag	2640
aaactgggaa caactgcagc aacagagttg aagaaacttg attcaaagt ttctgtaca	2700
tcaaaataatc tgattcaac aattccatca gacaatttg cagcaggtac tgataataca	2760
agttccttag gaccccaag tatgccagtt cattatgata gtcaatttga taccactcta	2820
tttggcaaaa agtcatctcc cttactgag tctggggac ctctgagctt gagtgaagaa	2880
aataatgatt caaagttgtt agaatcaggt ttaatgaata gccaagaaag ttcattgggaa	2940
aaaaatgtat cgtcaacaga gagtggtagg ttattnaaag gggaaagagc tcatggacct	3000
gcttttgta ctaaagataa tgccttattc aaagtttagca tcttttgtt aaagacaac	3060
aaaacttcca ataattcagc aactaataga aagactcaca ttgatggccc atcattatta	3120
attgagaata gtcctcatgt ctggcaaaat atattagaaa gtgacactga gttttaaaaaa	3180
gtgacacacctt tgattcatga cagaatgctt atggaaaaaa atgctacagc tttgaggcta	3240
aatcatatgt caaataaaac tacttcatca aaaaacatgg aaatggtcca acagaaaaaa	3300
gagggccccca ttccaccaga tgcacaaaat ccagatatgt cgttcttaa gatgttattc	3360
ttgcagaat cagcaaggtg gatacaaagg actcatggaa agaactctt gaactctgg	3420
caaggccccca gtccaaagca attagtatcc tttaggaccag aaaaactgt ggaaggtcag	3480
aattcttgt ctgagaaaaaa caaagtggta gtaggaaagg gtgaatttac aaaggacgta	3540
ggactcaaag agatggttt tccaaggcagc agaaacctat ttcttactaa cttggataat	3600
ttacatgaaa ataatacaca caatcaagaa aaaaaatc aggaagaaat agaaaagaag	3660
gaaacattaa tccaagagaa tgtagtttg cctcagatac atacagtgc tggcactaag	3720
aatttcatga agaacctttt cttactgagc actaggcaaa atgtagaagg ttcattatgac	3780
ggggcatatg ctccagactact tcaagattt aggtcattaa atgattcaac aaatagaaca	3840
aagaaacaca cagtcattt ctcaaaaaaaaaa ggggaggaag aaaacttggaa aggcttggga	3900
aatcaaacca agcaaattgt agagaaatat gcatgcacca caaggataatc tcctaataca	3960
agccagcaga attttgtcac gcaacgtagt aagagagctt taaaacaatt cagactccca	4020
ctagaagaaa cagaacttga aaaaaggata attgtggatg acacctcaac ccagtggcc	4080
aaaaacatga aacatttgac cccgagcacc ctcacacaga tagactacaa tgagaaggag	4140

aaagggcca ttactcagtc tccttatca gattgccta cgaggagtca tagcatccct	4200
caagcaaata gatctccatt acccattgca aaggatcat cattccatc tattagacct	4260
atatatctga ccagggctt attccaagac aactttctc atttccagc agcatttat	4320
agaaagaag atictgggt ccaagaaagc agtcatttct tacaaggagc caaaaaaat	4380
aacccctt tagccattt aacccctggag atgactggc atcaagaga gttggctcc	4440
ctggggacaa gtgccacaaa ttcaagtaca tacaagaaag ttgagaacac tgcccccgg	4500
aaaccagact tgcccaaaac atctggcaaa gttgaattgc ttccaaaagt tcacattat	4560
cagaaggacc tattccctac ggaaacttagc aatgggtctc ctggccatct ggatctcg	4620
gaagggagcc ttcttcaggc aacagaggga gcgatatagt ggaatgaagc aaacagacct	4680
ggaaaagtcc ccttctgag agtagcaaca gaaagctcg caaagactcc ctccaaagct	4740
ttggatccctc ttgttggga taaccactat ggtactcaga taccaaaaga agagtggaaa	4800
tcccaagaga agtcaccaga aaaaacagct tttaagaaaa aggataccat ttgtccctg	4860
aacgcttgtg aaagcaatca tgcaatagca gcaataatg agggacaaaa taagccccaa	4920
atagaagtca cctggcaaa gcaaggtagg actgaaaggc tgtgctctca aaaccacca	4980
gtctgaaac gccatcaacg ggaaataact cgtactactc ttcaagtca tcaagaggaa	5040
attgactatg atgataccat atcagttgaa atgaagaagg aagatttgat cattatgat	5100
gaggatgaaa atcagagccc ccgcagctt caaaagaaaa cacgacacta ttttattgct	5160
gcagtggaga ggctctggga ttatggatg agtagctcc cacatgtct aagaaacagg	5220
gctcagatgt gcagtgtccc tcagttcaag aaagtgttt tccaggaatt tactgatggc	5280
tccttactc agcccttata ccgtggagaa ctaaatgaac atttggact cctggggcca	5340
tatataagag cagaagttga agataatatc atggtaactt tcagaaatca ggcctctcg	5400
ccctattcct tctattcttag ctttatttct tatgaggaag atcagaggca aggaggcagaa	5460
cctagaaaaa accttgcata gcctaataatgaa accaaaaactt acctttggaa agtgcacat	5520
catatggcac ccactaaaga ttagtttgc tgcaaaagctt gggcttattt ctctgatgtt	5580
gacctggaaa aagatgtgca ctcaggcctg attggacccc ttctggctcg ccacactaac	5640
acactgaacc ctgctcatgg gagacaagtg acagtcagg aatttgctct gtttttacc	5700
atctttgatg agaccaaaaag ctggtaactt actgaaaata tggaaagaaaa ctgcaggcgt	5760
ccctgcaata tccagatgga agatcccact tttaaagaga attatcgctt ccatgcaatc	5820
aatggctaca taatggatac actacctggc ttagtaatgg ctcaggatca aaggattcga	5880
tggtatctgc tcagcatggg cagcaatgaa aacatccatt ctattcattt cagtgacat	5940
gtgttcactg tacgaaaaaa agaggagttt aaaaatggcac tgtacaatct ctatccaggt	6000

gttttgaga cagtggaaat gttaccatcc aaagctggaa ttggcggtt ggaatgcctt	6060
attggcgagc atctacatgc tggatgagc acacttttc tgggtacag caataagtgt	6120
cagactcccc tggaaatggc ttctggacac attagagatt ttcagattac agcttcagga	6180
caatatgac agtggccccc aaagctggcc agacttcatt attccggatc aatcaatgcc	6240

tggagcacca aggagccctt ttcttggatc aagggtggatc tggcacc aatgattttt	6300
cacggcatca agacccaggg tgccgtcag aagtctcca gcctctacat ctctcagttt	6360
atcatcatgt atagtcttga tggagaagaag tggcagactt atcgaggaaa ttccactgg	6420
acctaatagg tcttcttgg caatgtggat tcatactggtaaaaacacaa tatttttaac	6480
cctccaatta ttgtcgata catccgttgc caccaactc attatagcat tcgcagcact	6540
cttcgcatgg agttgtatggg ctgtgattt aatagttgca gcatgccatt ggaaatggag	6600
agtaaagcaa tatcagatgc acagattact gttcatctt actttaccaa tatgtttgcc	6660

acctggctc cttaaaaagc tcgacttcac ctccaaggga ggagtaatgc ctggagacct	6720
caggtgaata atccaaaaga gtggctgcaa gtggacttcc agaagacaat gaaagtacaa	6780
ggagtaacta ctcaaggagt aaaatctctg cttaccagca tgtatgtgaa ggagttcctc	6840
atctccagca gtcaagatgg ccatcagtgg actctttttt ttcagaatgg caaagtaaag	6900
gttttcagg gaaatcaaga ctccatcaca cctgtggta actctctaga cccaccgtta	6960
ctgactcgct accttcgaat tcaccccaag agttgggtgc accagattgc cctgaggatg	7020
gagggtctgg gctgcgaggc acaggaccc tac	7053

&lt;210&gt; 18

&lt;211&gt; 1438

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; B-domain-deleted FVIII

&lt;400&gt; 18

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                    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 Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp  
 340                    345                    350  
 Ser Glu Met Asp Val Val Arg Phe 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 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 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  
 965 970 975

Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr  
 980 985 990

Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro  
 995 1000 1005

Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg  
 1010 1015 1020

Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu  
 1025 1030 1035

Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met

1040	1045	1050
Gly Ser Asn Glu Asn Ile His	Ser Ile His Phe Ser	Gly His Val
1055	1060	1065
Phe Thr Val Arg Lys Lys Glu	Glu Tyr Lys Met Ala	Leu Tyr Asn
1070	1075	1080
Leu Tyr Pro Gly Val Phe Glu	Thr Val Glu Met Leu	Pro Ser Lys
1085	1090	1095
Ala Gly Ile Trp Arg Val Glu	Cys Leu Ile Gly Glu	His Leu His
1100	1105	1110
Ala Gly Met Ser Thr Leu Phe	Leu Val Tyr Ser Asn	Lys Cys Gln
1115	1120	1125
Thr Pro Leu Gly Met Ala Ser	Gly His Ile Arg Asp	Phe Gln Ile
1130	1135	1140
Thr Ala Ser Gly Gln Tyr Gly	Gln Trp Ala Pro Lys	Leu Ala Arg
1145	1150	1155
Leu His Tyr Ser Gly Ser Ile	Asn Ala Trp Ser Thr	Lys Glu Pro
1160	1165	1170
Phe Ser Trp Ile Lys Val Asp	Leu Leu Ala Pro Met	Ile Ile His
1175	1180	1185
Gly Ile Lys Thr Gln Gly Ala	Arg Gln Lys Phe Ser	Ser Leu Tyr
1190	1195	1200
Ile Ser Gln Phe Ile Ile Met	Tyr Ser Leu Asp Gly	Lys Lys Trp
1205	1210	1215
Gln Thr Tyr Arg Gly Asn Ser	Thr Gly Thr Leu Met	Val Phe Phe
1220	1225	1230
Gly Asn Val Asp Ser Ser Gly	Ile Lys His Asn Ile	Phe Asn Pro
1235	1240	1245
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> 19

<211> 4371

<212> DNA

<213> Artificial Sequence

<220><223> B-domain-deleted FVIII

<400> 19

atgcaaata	agctctccac	ctgtttttt	ctgtgcctt	tgcattctg	cttttagtgcc	60
accagaagat	actacctggg	tgcagtggaa	ctgtcatggg	actatatgca	aagtgatctc	120
ggtagactgc	ctgtggacgc	aagatttcct	cctagagtgc	caaaatctt	tccattcaac	180

acctcagtcg tgtacaaaaa gactctgtt tagaaattca cgatcacct tttcaacatc	240
gctaagccaa gccaccctg gatgggtctg ctagtccta ccatccaggc tgaggttat	300
gatacagtgg tcattacact taagaacatg gcttccatc ctgcgtct tcgtgttt	360
ggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtcaaagg	420
gagaaagaag atgataaagt cttccctgg ggaaggcata catatgtctg gcaggctcg	480
aaagagaatg gtccaatggc ctctgaccca ctgtgccta cctactcata tctttctcat	540
gtggacctgg taaaagactt gaattcaggc ctcattggag ccctactagt atgttagagaa	600
gggaggtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta	660
tttgcataag ggaaaagttt gcactcagaa acaaagaact cttgtatgca ggataggat	720
gctgcacatcg ctccggcctg gcctaaaatg cacacagtca atggttatgt aaacaggct	780
ctgccaggctc tgattggatg ccacaggaaa tcagtctatt ggcattgtat tggaatggc	840
accactcctg aagtgcactc aatattcctc gaaggtcaca catttctgt gaggaaccat	900
cgcaggcgt cttggaaat ctgcataa acttcctta ctgcataaac actcttgatg	960
gaccttgac agtttctact gttttgtcat atctttccc accaacatga tggcatggaa	1020
gcttatgtca aagtagacag ctgtccagag gaaccccaac tacgaatgaa aaataatgaa	1080
gaagcggaaag actatgatga tgatcttact gattctgaaa tggatgtggc caggttgc	1140
gatgacaact ctccctcatt tatccaaatt cgctcgttg ccaagaagca tcctaaaact	1200
tgggtacatt acattgctgc tgaagaggag gactggact atgctccctt agtccctgcc	1260
cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg	1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa ccttaagac tcgtgaagct	1380
attcagcatg aatcaggaat cttggaccc ttactttatg gggaaaggtag agacacactg	1440
ttgattataat ttaagaatca agcaaggaga ccatataaca tctaccctca cggaatcact	1500
gatgtccgtc cttgttattc aaggagatta ccaaagggtg taaaacattt gaaggattt	1560
ccaaattctgc caggagaaat attcaaataat aaatggacag tgactgtaga agatggcca	1620
actaaatcag atccctcggtg cctgacccgc tattactcta gttcgttaa tatggagaga	1680
gatctagctt caggactcat tggccctctc ctcattgtct acaaagaatc tgttagatcaa	1740
agaggaaacc agataatgtc agacaagagg aatgtcatcc tttttctgt atttgatgag	1800
aaccgaagct ggtacctcagc agagaatata caacgtttc tccccatcc agctggagtg	1860
cagcttgagg atccagagtt ccaaggctcc aacatcatgc acagcatcaa tggctatgtt	1920

tttgatagtt tgca gttgtc agtttggg catgagggtg catactggta cattctaagc	1980
attggagcac agactgactt cctttctgtc ttcttcctg gataacacctt caaacacaaa	2040
atggcttatg aagacacact caccctattc ccattctcg gagaaactgt cttcatgtcg	2100
atggaaaacc caggcttatg gattctgggg tgccacaact cagactttcg gaacagaggc	2160
atgaccgcct tactgaaggt ttcttagttgt gacaagaaca ctgggtgatta ttacgaggac	2220
agttatgaag atattcagc atacttgcg agtaaaaaca atgccattga accaagaagc	2280
ttctctcaaa acccaccagt ctgaaacgc catcaacggg aaataactcg tactacttt	2340
cagtcagatc aagagggaaat tgactatgt gataccatat cagttgaaat gaagaaggaa	2400
gatttigaca tttatgtga ggatgaaaat cagagcccc gcagcttca aaagaaaaca	2460
cgacactatt ttattgctgc agtggagagg ctctggatt atggatgag tagtccccca	2520
catgttctaa gaaacagggc tcagagtggc agtgtccctc agttcaagaa agttgtttc	2580
caggaattta ctgatggctc cttaactcg cccttatacc gtggagaact aaatgaacat	2640
ttggactcc tggggccata tataagagca gaagttgaag ataataatcat ggtactttc	2700
agaaatcagg cctctcgcc ctattccttc tattctagcc ttatttctta tgaggaagat	2760
cagaggcaag gaggcagaacc tagaaaaaac tttgtcaagc ctaatgaaac caaaacttac	2820
ttttggaaag tcaacatca tatggcaccc actaaagatg agtttgactg caaagcctgg	2880
gcttatttct ctgatgttga cctggaaaaa gatgtgcact cagggctgtat tggaccctt	2940
ctggctgcc acactaacac actgaaccct gtcatggg gacaagtgc agtacaggaa	3000
tttgctctgt tttcaccat ctgttgcgatg accaaaagct ggtacttcac tggaaatatg	3060
gaaagaaaact gcagggctcc ctgcaatatc cagatggaaat atcccacttt taaagagaat	3120
tatcgcttcc atgcaatcaa tggctacata atggatacac tacctggctt agtaatggct	3180
caggatcaa ggattcgatg gtatctgctc agcatggca gcaatggaaa catccattct	3240
attcatttca gtggacatgt gttactgttca cgaaaaaaaaaggatataa aatggcactg	3300
tacaatctt atccagggtt ttttggatca gttggaaatgt taccatccaa agctggaaatt	3360
tggcgggtgg aatgcctt tggcgagcat ctacatgttggatggacacttttctg	3420
gtgtacagca ataagtgtca gactccctg ggaatggctt ctggacacat tagagat	3480
cagattacag cticaggaca atatggacag tggggccaa agctggccag acttcattat	3540
tccggatcaa tcaatgcctg gagcaccaag gagcccttt ctggatcaa ggtggatctg	3600
ttggcaccaa tgattattca cggcatcaag acccagggtg cccgtcagaa gttctccagc	3660
ctctacatct cttagttt catcatgtat agtcttgatg ggaagaagtgc gcaacttat	3720
cgagggaaattt ccactggaaac cttaatggtc ttctttggca atgtggattc atctggata	3780

aaacacaata ttttaaccc tccaattatt gctcgataca tccgttgca cccaactcat	3840
tatagcattc gcagcactt tcgcattggag ttgatggct gtgattaaa tagtgtcagc	3900
atgccattgg gaatggagag taaagcaata tcagatgcac agattactgc ttcatcctac	3960
tttaccaata tggtgccac ctggtctct tcaaaagctc gactcacct ccaaggagg	4020

agtaatgcct ggagacctca ggtgaataat ccaaagagt ggctgcaagt ggacttcag	4080
aagacaatga aagtacagg agtaactact cagggagtaa aatctctgct taccagcatg	4140
tatgtgaagg agttcctcat ctccagcgt caagatggcc atcagtggac tctttttt	4200
cagaatggca aagtaaaggt tttcaggga aatcaagact cttcacacc tgtggtaac	4260
tctctagacc caccgttact gactcgctac ctgcatttc acccccagag ttgggtcac	4320
cagattgccc tgaggatgga ggttctggc tgcgaggcac aggacctcta c	4371

&lt;210&gt; 20

&lt;211&gt; 11

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; PAR1 exosite interaction motif

&lt;400&gt; 20

Pro Asn Asp Lys Tyr Glu Pro Phe Trp Glu Asp

1 5 10

&lt;210&gt; 21

&lt;211&gt; 12

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; PAR1 exosite interaction motif

&lt;400&gt; 21

Pro Asn Asp Lys Tyr Glu Pro Phe Trp Glu Asp Glu

1 5 10

&lt;210&gt; 22

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; PAR1 exosite interaction motif

&lt;400&gt; 22

Pro Asn Asp Lys Tyr Glu Pro Phe Trp Glu Asp Glu Glu

1 5 10

<210> 23

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> PAR1 exosite interaction motif

<400> 23

Pro Asn Asp Lys Tyr Glu Pro Phe Trp Glu Asp Glu Glu Ser

1 5 10

<210> 24

<211> 26

<212> PRT

<213> Artificial Sequence

<220><223> VWF linker

<400> 24

Gly Gly Leu Val Pro Arg Ser Phe Leu Leu Arg Asn Pro Asn Asp Lys

1 5 10 15

Tyr Glu Pro Phe Trp Glu Asp Glu Glu Ser

20 25

<210> 25

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Thrombin cleavage site

<400> 25

Leu Val Pro Arg

1

<210> 26

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Thrombin cleavage site

&lt;400&gt; 26

Ala Leu Arg Pro Arg Val Val

1 5

&lt;210&gt; 27

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; FXIa cleavage site

&lt;400

&gt; 27

Thr Gln Ser Phe Asn Asp Phe Thr Arg

1 5

&lt;210&gt; 28

&lt;211&gt; 10

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; FXIa cleavage site

&lt;400&gt; 28

Ser Val Ser Gln Thr Ser Lys Leu Thr Arg

1 5 10

&lt;210&gt; 29

&lt;211&gt; 10

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Thrombin cleavage site

&lt;400&gt; 29

Asp Phe Leu Ala Glu Gly Gly Val Arg

1 5 10

&lt;210&gt; 30

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; Thrombin cleavage site

&lt;400&gt; 30

Thr Thr Lys Ile Lys Pro Arg

1 5

<210> 31

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Thrombin cleavage site

<400> 31

Leu Val Pro Arg Gly

1 5

<210> 32

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 32

Ala Ser Pro Ala Ala Pro Ala Pro Ala Ser Pro Ala Ala Pro Ala Pro

1 5 10 15

Ser Ala Pro Ala

20

<210> 33

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 33

Ala Ala Pro Ala Ser Pro Ala Pro Ala Ala Pro Ser Ala Pro Ala Pro

1 5 10 15

Ala Ala Pro Ser

20

<210> 34

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 34

Ala Pro Ser Ser Pro Ser Pro Ala Pro Ser Ser Pro Ser Pro Ala

1 5 10 15

Ser Pro Ser Ser

20

<210> 35

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 35

Ala Pro Ser Ser Pro Ser Pro Ala Pro Ser Ser Pro Ser Pro Ala

1 5 10 15

Ser Pro Ser

<210> 36

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 36

Ser Ser Pro Ser Ala Pro Ser Pro Ser Ser Pro Ala Ser Pro Ser Pro

1 5 10 15

Ser Ser Pro Ala

20

<210> 37

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 37

Ala Ala Ser Pro Ala Ala Pro Ser Ala Pro Pro Ala Ala Ser Pro

1 5 10 15

Ala Ala Pro Ser Ala Pro Pro Ala

20

<210> 38

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> PAS sequence

<400> 38

Ala Ser Ala Ala Ala Pro Ala Ala Ala Ser Ala Ala Ala Ser Ala Pro

1 5 10 15

Ser Ala Ala Ala

20

<210> 39

<211> 42

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AE42

<400> 39

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> 40

<211> 78

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AE72

<400> 40

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> 41

<211> 143

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AE144

<400> 41

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

&lt;210&gt; 42

&lt;211&gt; 144

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; XTEN AG144

&lt;400&gt; 42

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

&lt;210&gt; 43

&lt;211&gt; 288

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; XTEN AE288

&lt;400&gt; 43

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> 44

<211> 288

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AG288

<400> 44

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 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> 45

<211> 576

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AE576

<400> 45

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
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
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
Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro		
565	570	575
<210> 46		
<211> 576		
<212> PRT		

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; X TEN AG576

&lt;400&gt; 46

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                    320  
 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                    400

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                  480

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> 47

<211> 864

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AE864

<400> 47

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
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
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
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
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
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
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> 48

<211> 864

<212> PRT

<213> Artificial Sequence

<220><223> XTEN AG864

<400> 48

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
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
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  
 705 710 715 720

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  
 785 790 795 800

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> 49

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AD

<400> 49

Gly Glu Ser Pro Gly Gly Ser Ser Gly Ser Glu Ser

1 5 10

<210> 50

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AD

<400> 50

Gly Ser Glu Gly Ser Ser Gly Pro Gly Glu Ser Ser

1 5 10

<210> 51

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AD

<400> 51

Gly Ser Ser Glu Ser Gly Ser Ser Glu Gly Gly Pro

1 5 10

<210> 52

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AD

<400> 52

Gly Ser Gly Gly Glu Pro Ser Glu Ser Gly Ser Ser

1 5 10

<210> 53

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AE, AM

<400> 53

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

1 5 10

<210> 54

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AE, AM, AQ

<400> 54

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

1	5	10
---	---	----

<210> 55

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AE, AM, AQ

<400> 55

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

1	5	10
---	---	----

<210> 56

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

><223> XTEN Motif Family AE, AM, AQ

<400> 56

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

1	5	10
---	---	----

<210> 57

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AF, AM

<400> 57

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

1	5	10
---	---	----

<210> 58

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AF, AM

<400> 58

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

1 5 10

<210> 59

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> X TEN Motif Family AF, AM

<400> 59

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

1 5 10

<210> 60

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> X TEN Motif Family AF, AM

<400> 60

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

1 5 10

<210> 61

<211> 12

<212>

> PRT

<213> Artificial Sequence

<220><223> X TEN Motif Family AG, AM

<400> 61

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

1 5 10

<210> 62

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> X TEN Motif Family AG, AM

<400> 62

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

1 5 10

<210> 63

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AG, AM

<400> 63

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

1 5 10

<210> 64

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AG, AM

<400> 64

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

1 5 10

<210> 65

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AQ

<400> 65

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

1 5 10

<210> 66

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family AQ

<400> 66

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

1                   5                   10  
<210> 67  
<211> 12  
<212> PRT  
<213> Artificial Sequence  
<220><223> XTEN Motif Family AQ  
<400> 67  
Gly Ser Gly Pro Ser Thr Glu Ser Ala Pro Thr Glu

1                   5                   10  
<210> 68  
<211> 12  
<212> PRT  
<213> Artificial Sequence  
<220><223> XTEN Motif Family AQ

<400> 68  
Gly Ser Glu Thr Pro Ser Gly Pro Ser Glu Thr Ala  
1                   5                   10  
<210> 69  
<211> 12  
<212> PRT  
<213> Artificial Sequence  
<220><223> XTEN Motif Family AQ  
<400> 69

Gly Pro Ser Glu Thr Ser Thr Ser Glu Pro Gly Ala  
1                   5                   10  
<210> 70  
<211> 12  
<212> PRT  
<213> Artificial Sequence  
<220><223> XTEN Motif Family AQ  
<400> 70

Gly Ser Pro Ser Glu Pro Thr Glu Gly Thr Ser Ala  
1                   5                   10

<210> 71

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family BC

<400> 71

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

1	5	10
---	---	----

<210> 72

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family BC

<400> 72

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

1	5	10
---	---	----

<210> 73

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family BC

<400> 73

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

1	5	10
---	---	----

<210> 74

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> XTEN Motif Family BC

<400> 74

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

1	5	10
---	---	----

<210> 75

<211> 12

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; XTEN Motif Family BD

&lt;400&gt; 75

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

1 5 10

&lt;210&gt; 76

&lt;211&gt; 12

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; XTEN Motif Family BD

&lt;400&gt; 76

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

1 5 10

&lt;210&gt; 77

&lt;211&gt; 12

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; XTEN Motif Family BD

&lt;400&gt; 77

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

1 5 10

&lt;210&gt; 78

&lt;211&gt; 12

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; XTEN Motif Family BD

&lt;400&gt; 78

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

1 5 10

&lt;210&gt; 79

&lt;211&gt; 4974

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; VWF057 (VWF D'D3-Fc with LVPR thrombin site in the linker)

&lt;400&gt; 79

atgattcccg ccagatttgc cggggtgctg ctgtctgg ccctcatttt gccagggacc	60
ctttgtgcag aaggaactcg cggcaggta tccacggccc gatgcagcct tttcgaaagt	120
gacttcgtca acaccttga tgggagcatg tacagtttgc gggatactg cagttacctc	180
ctggcagggg gctgccagaa acgctccttc tcgattattt gggacttcca gaatggcaag	240
agagtgagcc tctccgtgta tcttgggaa tttttgaca tccatttgt tgtcaatgg	300
accgtgacac agggggacca aagagtctcc atgccatatg cttccaaagg gctgtatcta	360
gaaactgagg ctgggtacta caagctgtcc ggtgaggcct atggcttgc ggcaggatc	420
gatggcagcg gcaacttca agtctgtctc tcagacagat acttcaacaa gacctgggg	480
ctgtgtggca acttaacat cttgtctgaa gatgacttta tgacccaaga agggaccttgc	540
acctcggacc ctatgactt tgccaaactca tggctctga gcagtggaga acagtgggt	600
gaacgggcat ctcccccag cagctcatgc aacatctcct ctgggaaat gcagaaggc	660
ctgtgggagc agtgcagct tctgaagagc acctcggtgt ttgcccgtg ccaccctctg	720
gtggaccccg agcctttgt ggccctgtgt gagaagactt tgtgtgagtg tgctgggggg	780
ctggagtgcg cctgcctgc ctcctggag tacgccccga cctgtgcacca ggaggaaatg	840
gtgcgtacg gctggaccga ccacagcgcg tgcagcccgatgtgcctgc tggtatggag	900
tataggcagt gtgtgtcccc ttgcgcagg accgtccaga gcctgcacat caatgaaatg	960
tgtcaggagc gatgcgtgaa tggctgcgc tgccctgagg gacagctcct ggtatggc	1020
ctctgcgtgg agagcaccga gtgtccctgc gtgcattccg gaaagcgcta ccctccggc	1080
acctccctct ctgcagactg caacacctgc atttgcgaa acagccagt gatctgcagc	1140
aatgaagaat gtccaggaga gtgccttgctc actggtaat cccacttcaa gagctttgac	1200
aacagatact tcacccatcg tggatctgc cagttacctgc tggccggga ttgcaggac	1260
cactcctct ccattgtcat tgagactgtc cagttgtgtc atgaccgcga cgctgtgtgc	1320
acccgcctcg tcaccgtccg gtcgcctggc ctgcacaaca gccttgaa actgaagcat	1380
ggggcaggag ttgcatgaa tggccaggac atccagctcc ccctctgaa aggtgaccc	1440
cgcattccagc atacagtgac ggcctccgtc cgcctcagct acggggagga cctgcagatg	1500
gactgggatg gcccgggag gtcgtggtg aagctgtccc cgcgttatgc cggaaagacc	1560
tgcgcctgt gtggaaatta caatggcaac caggcgacg acttccttac ccctctggg	1620
ctggcggagc cccgggtgaa ggacttcggg aacgcctgaa agctgcacgg ggactgccag	1680

gacctgcaga agcagcacag cgatccctgc gccctaacc cgcgcatgac caggtttcc	1740
gaggaggcgt gcgcggctt gacgtcccc acattcgagg cctgccatcg tgccgtcagc	1800
ccgcgtgcct acctgcggaa ctgcccgtac gacgtgtgct cctgctcgga cggccgcgag	1860
tgccctgtcg ggcgcctggc cagctatgcc gcggccgtcg cggggagagg cgtgcgcgtc	1920
gcgtggcgcg agccaggccg ctgtgagctg aactgcccga aaggccaggt gtacctgcag	1980
tgcgggaccc cctgcaacct gacgtccgc tctctctt accggatga ggaatgaat	2040
gaggcctgcc tggaggcgtg cttctcccc ccaggctct acatggatga gaggggggac	2100
tgcgtgcca aggcccagtg cccctgttac tatgacggtg agatcttcca gccagaagac	2160
atcttctcag accatcacac catgtgttac tgtgaggatg gttcatgca ctgtaccatg	2220
agtggagtcc ccggaagctt gtcgcctgac gctgtctca gcagtccct gtctcatcgc	2280
agcaaaggaa gcctatcctg tggcccccc atggtaaagc tggtgtgtcc cgctgacaac	2340
ctgcgggctg aagggctcga gtgtacaaa acgtgcaga actatgaccc ggagtgcatt	2400
agcatggct gtgtctctgg ctgcctctgc ccccccggca tggccggca tgagaacaga	2460
tgtgtggccc tggaaaggtg tccctgttc catcaggca aggagtatgc ccctggagaa	2520
acagtgaaga ttggctgcaa cacttgttgc tgcgggacc ggaagtggaa ctgcacagac	2580
catgtgtgt atgcacgtg ctccacgtc ggcattggccc actacccac cttcgacggg	2640
ctcaaataacc tttccccgg ggagtgcac tacgttctgg tgcaggatta ctgcggcagt	2700
aaccctggaa ctttcggat cctagtgggg aataaggat gcagccaccc ctcaatgaaa	2760
tgcaagaaac gggcacat cctgggtggag ggaggagaga ttgagctgtt tgacggggag	2820
gtgaatgtga agaggccat gaaggatgag actcaattt aggtggtgga gtctggccgg	2880
tacatcattc tgctgtggg caaagccctc tccgtggctt gggaccgcca cctgagcatc	2940
tccgtggcc tgaaggcagac ataccaggag aaagtgtgtg gcctgtgtgg gaatttttat	3000
ggcatccaga acaatgaccc caccaggcgc aaccccaag tggaggaaga ccctgtggac	3060
tttggaaact cctggaaagt gagctgcag tgtgctgaca ccagaaaat gcctctggac	3120
tcatccccctg ccacctgcca taacaacatc atgaaggcaga cgtggtgga ttccctctgt	3180
agaatcctta ccagtgcacgt cttccaggac tgcaacaagc tggtgaccc cgagccatat	3240
ctggatgtct gcattacga cacctgtcc tgtgagtcca ttggggactg cgccgcattc	3300
tgcgacacca ttgctgccta tgcccacgtg tgtgcccagc atggcaaggt ggtgacctgg	3360
aggacggcca catttgtgccc ccagagctgc gaggagagga atctccggga gaacgggtat	3420
gaggctgagt ggcgcataa cagctgtgca cctgcgttc aagtcacgtg tcagcacct	3480
gagccactgg cctgcccgtt gcagtgtgtg gagggtgcg atgcccactg ccctccagg	3540

aaaatcctgg atgagcttt gcagacctgc gttgaccctg aagactgtcc agtgtgtgag 3600

gtggctggcc ggcgtttgc ctcaggaaag aaagtacct tgaatcccag tgaccctgag 3660  
 cactgccaga ttgccactg tcatgttgc aacacctgt gtgaaggctg ccaggagccg 3720  
 atatcggcg cgccaacatc agagagcgcc acccctgaaa gtggtcccg gagcgagcca 3780  
 gccacatctg ggtcgaaac gccaggcaca agtgagtctg caactcccga gtccggacct 3840  
 ggctccgagc ctgccactag cggctccgag actccggaa ctccgagag cgctacacca 3900  
 gaaagcggac ccggaaccag taccgaacct agcgaggct ctgtccggg cagcccagcc 3960  
 ggctctcta catccacgga ggagggcact tccgaatccg ccaccccgga gtcagggcca 4020

ggatctgaac ccgtacaccc aggcagttag acgccaggaa cgagcgagtc cgctacacccg 4080  
 gagagtggc cagggagccc tgctggatct cctacgtcca ctgaggaagg gtcaccagcg 4140  
 ggctcgccca ccagcaactga agaaggtgcc tcgagcggcg gtggaggatc cggtggcggg 4200  
 ggatccggtg cgccccggatc cggtggcggg ggatccggtg cgccccggatc cggtggcggg 4260  
 ggatccctgg tccccgggg cagcggaggc gacaaaactc acacatgcc accgtgccca 4320  
 gctccagaac tcctggcgg accgtcagtc ttccctttcc cccaaaaacc caaggacacc 4380  
 ctcatgatct cccggacccc tgaggtcaca tgcgtggtag tggacgtgag ccacgaagac 4440

cctgaggtca agttaactg gtacgtggac ggcgtggagg tgcataatgc caagacaaag 4500  
 ccgcgggagg agcagtacaa cagcacgtac cgtgtggtca gcgtccctac cgtccctgac 4560  
 caggactggc tgaatggcaa ggagtacaag tgcaaggtct ccaacaaagc cctccagcc 4620  
 cccatcgaga aaaccatctc caaagccaaa gggcagcccc gagaaccaca ggtgtacacc 4680  
 ctgccccat cccggatga gctgaccaag aaccaggatca gcctgacccctg cctggtaaa 4740  
 ggcttctatc ccagcgacat cgccgtggag tggagagca atggcagcc ggagaacaac 4800  
 tacaagacca cgcctccgt gttggactcc gacggctcc tttccctcta cagcaagctc 4860

accgtggaca agagcaggtg gcagcagggg aacgtttct catgctccgt gatgtcatgag 4920  
 gctctgcaca accactacac gcagaagagc ctctccctgt ctccggtaa atga 4974

<210> 80

<211> 1662

<212> PRT

<213> Artificial Sequence

<220><223> VWF057 (VWF D'D3-Fc with LVPR thrombin site linker)

<400> 80

Met Ile Pro Ala Arg Phe Ala Gly Val Leu Leu Ala Leu Ala 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		
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

785 790 795 800

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 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 Ala Pro Thr Ser Glu  
 1235 1240 1245  
 Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser  
 1250 1255 1260  
 Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser  
 1265 1270 1275  
 Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
 1280 1285 1290

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr  
 1295 1300 1305  
 Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro  
 1310 1315 1320  
 Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser  
 1325 1330 1335  
 Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
 1340 1345 1350

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala  
 1355 1360 1365  
 Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro  
 1370 1375 1380  
 Thr Ser Thr Glu Glu Gly Ala Ser Ser Gly Gly Gly Ser Gly  
 1385 1390 1395  
 Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
 1400 1405 1410

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Leu  
 1415 1420 1425  
 Val Pro Arg Gly Ser Gly Gly Asp Lys Thr His Thr Cys Pro Pro  
 1430 1435 1440  
 Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe  
 1445 1450 1455  
 Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu

1460	1465	1470
Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val		
1475	1480	1485
Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys		
1490	1495	1500
Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val		
1505	1510	1515
Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu		
1520	1525	1530
Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu		
1535	1540	1545
Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val		
1550	1555	1560
Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val		
1565	1570	1575
Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala		
1580	1585	1590
Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr		
1595	1600	1605
Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser		
1610	1615	1620
Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe		
1625	1630	1635
Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln		
1640	1645	1650
Lys Ser Leu Ser Leu Ser Pro Gly Lys		
1655	1660	
<210> 81		
<211> 4959		
<212> DNA		
<213> Artificial Sequence		

<220><223> VWF059 (VWF D'D3-Fc with acidic region thrombin site in the linker)

<400> 81

atgattcccg ccagatttgc cggggtgctg ctgtctgg ccctcatttt gccagggacc	60
ctttgtgcag aaggaactcg cggcaggta tccacggccc gatgcagcct tttcgaaagt	120
gacttcgtca acaccttga tgggagcatg tacagtttgc cgggatactg cagttacctc	180
ctggcagggg gctgccagaa acgctccttc tcgattattt gggacttcca gaatggcaag	240
agagtgagcc tctccgtgta tcttgggaa tttttgaca tccatttgt tgtcaatgg	300
accgtgacac agggggacca aagagtctcc atgcctatg ctcacaaagg gctgtatcta	360
gaaactgagg ctgggtacta caagctgtcc ggtgaggcct atggcttgc ggccaggatc	420
gatggcagcg gcaacttca agtctgtgt tcagacagat acttcaacaa gacctgggg	480
ctgtgtggca acttaacat cttgtgaa gatgacttta tgacccaaga agggacatttgc	540
acctcggacc ctatgactt tgccaaactca tggcctctga gcagtggaga acagtgggt	600
gaacgggcat ctcccccag cagctcatgc aacatctcct ctgggaaat gcagaaggc	660
ctgtgggagc agtgcacgt tctgaagagc acctcggtgt ttgcccgtg ccaccctctg	720
gtggaccccg agcctttgt ggccctgtgt gagaagactt tgtgtgagtg tgctgggggg	780
ctggagtgcg cctgcctgc cctcctggag tacgccccga cctgtgcacca ggaggaaatg	840
gtgcgtacg gctggaccga ccacagcgcg tgcagcccgatgtgcctgc tggtatggag	900
tataggcagt gtgtgtcccc ttgcgcagg acctgcaga gcctgcacat caatgaaatg	960
tgtcaggagc gatgcgtgaa tggctgcgc tgccctgagg gacagctcct ggtatggc	1020
ctctgcgtgg agagcaccga gtgtccctgc gtgcattccg gaaagcgcta ccctccggc	1080
acctccctct ctgcgacactg caacacactgc atttgcgaa acagccagt gatctgcac	1140
aatgaagaat gtccaggaaa gtgccttgac actggtaat cccacttcaa gagctttgac	1200
aacagatact tcacccatcg tggatctgc cagttacactgc tggccggaa ttgcaggac	1260
cactcctct ccattgtcat tgagactgtc cagttgtgtc atgaccgcga cgctgtgtgc	1320
acccgcctcg tcaccgtccg gtcgcctggc ctgcacaaca gccttgaa actgaagcat	1380
ggggcaggag ttgcatgaa tggccaggac atccagctcc ccctctgaa aggtgaccc	1440
cgcattccagc atacagtgac ggcctccgtc cgcctcagct acggggagga cctgcagatg	1500
gactgggatg gcccgggag gtcgtggatg aagctgtccc cctgttatgc cggaaagacc	1560
tgcgcctgt gtggaaatta caatggcaac caggcgacg acttccttac ccctctggg	1620
ctggcggagc cccgggtgaa ggacttcggg aacgcctgaa agctgcacgg ggactgccag	1680

gacctgcaga agcagcacag cgatccctgc gccctaacc cgcgcatgac caggtttcc	1740
gaggaggcgt gcgcggctt gacgtcccc acattcgagg cctgccatcg tgccgtcagc	1800
ccgcgtgcct acctgcggaa ctgcccgtac gacgtgtgct cctgctcgga cggccgcgag	1860
tgccctgtcg ggcgcctggc cagctatgcc gcggccgtcg cggggagagg cgtgcgcgtc	1920
gcgtggcgcg agccaggccg ctgtgagctg aactgcccga aaggccaggt gtacctgcag	1980
tgcgggaccc cctgcaacct gacctgccgc tctctctt accggatga ggaatgaat	2040
gaggcctgcc tggagggctg cttctcccc ccagggctt acatggatga gagggggac	2100
tgcgtgcca aggcccagtg cccctgttac tatgacggtg agatcttcca gccagaagac	2160
atcttctcag accatcacac catgtgttac tgtgaggatg gttcatgca ctgtaccatg	2220
agtggagtcc ccggaagctt gtcgcctgac gctgtctca gcagtccct gtctcatcg	2280
agcaaaaagga gcctatcctg tggcccccc atggtaaagc tggtgtgtcc cgctgacaac	2340
ctgcgggctg aagggctcga gtgtaccaa acgtgccaga actatgaccc ggagtgcatt	2400
agcatggct gtgtctctgg ctgcctctgc ccccccggca tggccggca tgagaacaga	2460
tgtgtggccc tggaaaggtg tccctgttcc catcaggcca aggagtatgc ccctggagaa	2520
acagtgaaga ttggctgcaa cacttgttgc tgcgggacc ggaagtggaa ctgcacagac	2580
catgtgtgt atgcacgtg ctccacgttcc ggcattggccc actacccatc cttcgacggg	2640
ctcaaataacc tttccccgg ggagtgcac tacgttctgg tgcaggatta ctgcggcagt	2700
aaccctggga ctttcggat cctagtgggg aataagggat gcagccaccc ctcaatggaa	2760
tgcaagaaac gggcacat cctgggtggag ggaggagaga ttgagctgtt tgacggggag	2820
gtgaatgtga agaggccat gaaggatgag actcaatttgg aggtggtgaa gtctggccgg	2880
tacatcattt tgcgtctggg caaagccctc tccgtggctt gggaccgcca cctgagcatc	2940
tccgtggcc tgaaggcagac ataccaggag aaagtgtgtg gcctgtgtgg gaattttgt	3000
ggcatccaga acaatgaccc caccaggcgc aaccccaag tggaggaaga ccctgtggac	3060
tttggaaact cctggaaagt gagctgcagc tgtgctgaca ccagaaaat gcctctggac	3120
tcatccccctg ccacctgcca taacaacatc atgaagcaga cgtgggttgc ttccctctgt	3180
agaatcctta ccagtgcacgt cttccaggac tgcaacaagc tggtgaccc cgagccatat	3240
ctggatgtct gcattacga cacctgttcc tgtgagtcca ttggggactg cgccgcattc	3300
tgcgacacca ttgctgccta tgcccacgtg tgtgcccacgc atggcaaggt ggtgacccgtt	3360
aggacggcca catttgtgccc ccagagctgc gaggagagga atctccggga gaacgggtat	3420
gaggctgagt ggcgcataa cagctgtgc cctgcgttc aagtcacgtg tcagcaccc	3480
gagccactgg cctgcccgtt gcagtgtgtg gagggtgcg atgcccactg ccctccagg	3540

aaaatcctgg atgagcttt gcagacctgc gttgaccctg aagactgtcc agtgtgtgag 3600

gtggctggcc ggcgtttgc ctcaggaaag aaagtacct tgaatcccag tgaccctgag 3660  
 cactgccaga ttgccactg tcatgttgc aacacctgt gtgaaggctg ccaggagccg 3720  
 atatcggcg cgccaacatc agagagcgcc acccctgaaa gtggtcccg gagcgagcca 3780  
 gccacatctg ggtcgaaac gccaggcaca agtgagtctg caactcccga gtccggacct 3840  
 ggctccgagc ctgccactag cggctccgag actccggaa ctccgagag cgctacacca 3900  
 gaaagcggac ccggaaccag taccgaacct agcgaggct ctgtccggg cagcccagcc 3960  
 ggctctcta catccacgga ggagggcact tccgaatccg ccaccccgga gtcagggcca 4020

ggatctgaac ccgttaccc aggcagttag acgccaggaa cgagcgagtc cgctacaccc 4080  
 gagagtggc cagggagccc tgctggatct cctacgtcca ctgaggaagg gtcaccagcg 4140  
 ggctcgccca ccagcaactga agaaggtgcc tcgatatctg acaagaacac tggtgattat 4200  
 tacgaggaca gttatgaaga tatttcagca tacttgctga gtaaaaacaa tgccattgaa 4260  
 ccaagaagct tctctgacaa aactcacaca tgcccaccgt gcccagctcc agaactctg 4320  
 ggcggaccgt cagtcttctt cttccccca aaacccaagg acaccctcat gatctccgg 4380  
 acccctgagg tcacatgcgt ggtggtgac gtgagccacg aagaccctga ggtcaagtcc 4440

aactggtagc tggacggcgt ggaggtgcat aatgccaaga caaagccgca ggaggaggcag 4500  
 tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat 4560  
 ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagccccat cgagaaaaacc 4620  
 atctccaaag ccaaggggca gccccgagaa ccacaggtgt acaccctgcc cccatccgg 4680  
 gatgagctga ccaagaacca ggtcagcgtc acctgcctgg tcaaaggctt ctatcccage 4740  
 gacatcgccg tggagtgggaa gagcaatggg cagccggaga acaactacaa gaccacgcct 4800  
 cccgttgttgg actccgacgg ctccttcttc ctctacagca agtcaccgt ggacaagagc 4860

aggtggcagc aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac 4920  
 tacacgcaga agagcctctc cctgtctcg ggttaatgaa 4959

<210> 82

<211> 1652

<212> PRT

<213> Artificial Sequence

<220><223> VWF059 (VWF D'D3-Fc with LVPR thrombin site in the linker)

<400> 82

Met Ile Pro Ala Arg Phe Ala Gly Val Leu Leu Ala Leu Ala 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		
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

785 790 795 800

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 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 Ala Pro Thr Ser Glu  
 1235 1240 1245  
 Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser  
 1250 1255 1260  
 Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser  
 1265 1270 1275  
 Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
 1280 1285 1290

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr  
 1295 1300 1305  
 Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro  
 1310 1315 1320  
 Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser  
 1325 1330 1335  
 Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
 1340 1345 1350

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala  
 1355 1360 1365  
 Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro  
 1370 1375 1380  
 Thr Ser Thr Glu Glu Gly Ala Ser Ile Ser Asp Lys Asn Thr Gly  
 1385 1390 1395  
 Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu  
 1400 1405 1410

Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Asp Lys Thr  
 1415 1420 1425  
 His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro  
 1430 1435 1440  
 Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile  
 1445 1450 1455  
 Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His

1460	1465	1470
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu		
1475	1480	1485
Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser		
1490	1495	1500
Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp		
1505	1510	1515
Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu		
1520	1525	1530
Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro		
1535	1540	1545
Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu		
1550	1555	1560
Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr		
1565	1570	1575
Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu		
1580	1585	1590
Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser		
1595	1600	1605
Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln		
1610	1615	1620
Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His		
1625	1630	1635
Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		
1640	1645	1650

&lt;210&gt; 83

&lt;211&gt; 4860

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; VWF062 (VWF D'D3-Fc with no thrombin site in the linker)

&lt;400&gt; 83

atgattcctg ccagattgc cgggtgctg cttgtctgg ccctcattt gccagggacc	60
ctttgtcgag aaggaactcg cggcaggta tccacggcc gatgcagcct tttcgaaagt	120
gacttcgtca acaccttga tggagcatg tacagtttg cggtatactg cagttaccc	180
ctggcaggaa gcigccagaa acgctccctc tcgattatg gggacttcca gaatggcaag	240
agagtgagcc tctccgtgta tcttgggaa tttttgaca tccatttgt tgtcaatggt	300
accgtgacac agggggacca aagagtctcc atgccatatg cctccaaagg gctgtatcta	360
gaaactgagg ctgggtacta caagctgtcc ggtgaggcct atggctttgt ggccaggatc	420
gatggcagcg gcaacttca agtcctgctg tcagacagat acttcaacaa gacctgcggg	480
ctgtgtggca acttaacat ctttgcitaa gatgacttta tgacccaaga agggacccgt	540
acctcggacc ctatgactt tgccaaactca tggctctga gcagtggaga acagtgggt	600
gaacgggcat ctctcccag cagtcatgc aacatctct ctggggaaat gcagaaggc	660
ctgtgggagc agtgcagct tctgaagagc acctcggtgt ttgcccgtg ccaccctctg	720
gtggaccccg agcffffgtt gcccctgtgt gagaagactt tgtgtgagtg tgctgggggg	780
ctggagtgcg cctgccctgc cctcctggag tacgccccga cctgtgccca ggagggaaatg	840
gtgctgtacg gctggaccga ccacagcgcg tgcagcccgag tgtgcctgc tggatggag	900
tataggcagt gtgtgtcccc ttgcgccagg acctgccaga gcctgcacat caatgaaatg	960
tgtcaggagc gatgcgtgga tggctgcagc tgccctgagg gacagctctt ggtatgaaggc	1020
ctctgcgtgg agagcaccga gtgtccctgc gtgcattccg gaaagcgcta ccctccggc	1080
acctccctct ctgcagactg caacacctgc atttgcgaa acagccagtg gatctgcagc	1140
aatgaagaat gtccagggga gtgccttgc actggtaat cccacttcaa gagcttgac	1200
aacagatact tcacccatcg tggatctgc cagttacctgc tggccggga ttgcaggac	1260
cactcctct ccattgtcat tgagactgtc cagttgtctg atgaccgcga cgctgtgtgc	1320
acccgcgtccg tcaccgtccg gctgcctggc ctgcacaaaca gccttgtgaa actgaagcat	1380
ggggcaggag ttgccccatggaa tggccaggac atccagctcc ccctcctgaa aggtgaccc	1440
cgcacatccagc atacagtgcg ggcctccgtg cgcctcagct acggggaggaa cctgcagat	1500
gactggatg gcccgggag gctgctgggtg aagctgtccc ccgtctatgc cggaaagacc	1560
tgcggcctgt gtggaaatta caatggcaac cagggcgacg acttccttac cccctctggg	1620
ctggcggagc cccgggtgga ggacttcggg aacgcctgga agctgcacgg ggactgccag	1680
gacctgcaga agcagcacag cgatccctgc gcccctcaacc cgccatgac cagttctcc	1740
gaggaggcgt ggcgggtcct gacgtcccc acattcaggc cctgcacatgc tgccgtcagc	1800
ccgctgcctt acctgcggaa ctgcccgtac gacgtgtgct cctgctcgaa cggccgcgag	1860

tgccctgtcg gcgcctggc cagctatgcc gcggcctgca cggggagagg cgtgcgcgtc	1920
gcgtggcgca agccaggccg ctgtgagctg aactgccga aaggccaggt gtacctgcag	1980
tgcgggaccc cctgcaacct gacctgccgc tctctcttt acccgatga ggaatgcaat	2040
gaggcctgcc tggagggctg cttctcccc ccagggctct acatggatga gagggggcac	2100
tcgtgccca aggcacagtg ccccttac tatgacggtg agatcttcca gccagaagac	2160
atcttctcag accatcacac catgtgtac tggaggatg gttcatgca ctgtaccatg	2220
agtggagtcc ccggaaagctt gtcgcctgac gctgtcctca gcagtcccgt gtctcatcg	2280
agcaaaaagga gcctatcctg tcggcccccc atggtaagc tggtgtgtcc cgctgacaac	2340
ctgccccctg aaggctcgaa gtgtaccaa acgtgccaga actatgacact ggagtgcata	2400
agcatggcgt gtgtctctgg ctgcctctgc ccccccggca tggccggca tgagaacaga	2460
tgtgtggccc tggaaaggtg tccctgctt catcaggca aggagtatgc ccctggagaa	2520
acagtgaaga ttggctgcaa cacttgcgtc tgcgggacc ggaagtggaa ctgcacagac	2580
catgtgtgtg atgccacgtg ctccacgatc ggcattggccc actacccac cttcgacggg	2640
ctcaaaatacc tggccccgg ggagtgcag tacgttctgg tgcaggatta ctgcggcagt	2700
aaccctggga ctttcggat cctagtgggg aataagggat gcagccaccc ctcagtgaaa	2760
tgcaagaaac gggtaccat cctggtgag ggaggagaga ttgagctgtt tgacggggag	2820
gtgaatgtga agaggcccat gaaggatgag actcaatttg aggtggtgga gtctggccgg	2880
tacatcatc tgcgtgtgg caaagccctc tccgtggct gggaccgcca cctgagcatc	2940
tccgtggcc tgaacgcac ataccaggaa aaagtgtgtc gctgtgtgg gaattttgat	3000
ggcatccaga acaatgaccc caccagcagc aaccccaag tggaggaaga ccctgtggac	3060
tttggaaact cttggaaagt gagctcgac tgcgtgaca ccagaaaagt gcctctggac	3120
tcatccccctg ccacctgcca taacaacatc atgaagcaga cgtgggtgg ttcctctgt	3180
agaatcccta ccagtgcgt ctccaggac tgcaacaagc tggtgaccc cgagccatat	3240
ctggatgtct gcattacga cacctgtcc tgcgtgtcca ttggggactg cgccgcattc	3300
tgcgacacca ttgtgccta tgcccacgtg tgcgtggcgc atggcaaggt ggtgacctgg	3360
aggacggcca cattgtgcc ccaagactgc gaggagagga atctccggga gaacgggtat	3420
gaggctgagt ggcgtataa cagctgtca cctgcgtgc aagtacgtg tcagcaccc	3480
gagccactgg ctgcctgt gcagtgtgtc gagggtggc atgcccactg ccctccagg	3540
aaaatcctgg atgagctttt gcagacccgt gttgaccctg aagactgtcc agtgtgtgag	3600

gtggctggcc ggcgtttgc ctcaggaaag aaagtacct tgaatcccag tgacccttag 3660

cactgccaga ttgccactg tcatgttgc aacacctt gtgaaggctg ccaggagccg 3720  
 atatcggcgcc cgccaacatc agagagcgcc accccctgaaa gtggtcccgg gagcgagcca 3780  
 gccacatctg ggtcgaaac gccaggcaca agtgagtctg caactcccga gtccggacct 3840  
 ggctccgagc ctgccactag cggctccgag actccggaa cttccgagag cgctacacca 3900  
 gaaagcggac ccggaaccag taccgaacct agcgaggct ctgctccgg cagcccagcc 3960  
 ggctctcta catccacgga ggagggcact tccgaatccg ccacccggaa gtcagggcca 4020  
 ggatctgaac ccgttaccc aggcagttag acgccaggaa cgagcggatc cgctacaccc 4080

gagagtgggc cagggagccc tgctggatct cctacgtcca ctgaggaagg gtcaccagcg 4140  
 ggctcgccca ccagcactga agaagggtgcc tcgagcgaca aaactcacac atgcccaccc 4200  
 tgcccagctc cagaactcct gggcggaccg tcagtttcc tttcccccc aaaacccaag 4260  
 gacaccctca tcatctcccg gacccttagt gtcacatgct tgggtggta cgtgagccac 4320  
 gaagaccctg aggtcaagtt caactggtagt gtggacggcg tggaggtgca taatgccaag 4380  
 acaaagccgc gggaggagca gtacaacagc acgtaccgtg tggcagcgt cttcaccgtc 4440  
 ctgcaccagg actggctgaa tggcaaggag tacaagtgca aggtctccaa caaagccctc 4500

ccagcccca tcgagaaaac catctccaa gccaaaggc agcccccaga accacaggtg 4560  
 tacaccctgc cccatcccg ggatgagctg accaagaacc aggtcagcct gacccctgc 4620  
 gtcaaaggct tctatcccg cgacatcgcc gtggagtgg agagcaatgg gcagccggag 4680  
 aacaactaca agaccacgccc tccgtgtt gactccgacg gtccttctt cctctacagc 4740  
 aagctcaccg tggacaagag caggtggcag cagggaaacg tcttctcatg ctccgtatg 4800  
 catgaggctc tgccacaacca ctacacgtag aagaccctt ccctgtctt gggtaatga 4860

&lt;210&gt; 84

&lt;211&gt; 1619

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; VWF062 (VWF D'D3-Fc with no thrombin site in the linker)

&lt;400&gt; 84

Met Ile Pro Ala Arg Phe Ala Gly Val Leu Leu Ala Leu 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
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
Leu Lys Tyr Leu Phe Pro Gly Glu Cys Gln Tyr Val Leu Val Gln Asp		880
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 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
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 Ala Pro Thr Ser Glu  
 1235 1240 1245  
 Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser

1250	1255	1260
Gly Ser	Glu Thr Pro Gly Thr	Ser Glu Ser Ala Thr
1265	1270	1275
Gly Pro	Gly Ser Glu Pro Ala	Thr Ser Gly Ser Glu
1280	1285	1290
Thr Ser	Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly
1295	1300	1305
Glu Pro	Ser Glu Gly Ser Ala	Pro Gly Ser Pro Ala
		Gly Ser Pro

1310	1315	1320
Thr Ser	Thr Glu Glu Gly Thr	Ser Glu Ser Ala Thr
1325	1330	1335
Gly Pro	Gly Ser Glu Pro Ala	Thr Ser Gly Ser Glu
1340	1345	1350
Thr Ser	Glu Ser Ala Thr Pro	Glu Ser Gly Pro Gly
1355	1360	1365
Gly Ser	Pro Thr Ser Thr Glu	Glu Gly Ser Pro Ala
		Gly Ser Pro

1370	1375	1380
Thr Ser	Thr Glu Glu Gly Ala	Ser Ser Asp Lys Thr
1385	1390	1395
Pro Pro	Cys Pro Ala Pro Glu	Leu Leu Gly Gly Pro
1400	1405	1410
Leu Phe	Pro Pro Lys Pro Lys	Asp Thr Leu Met Ile
1415	1420	1425
Pro Glu	Val Thr Cys Val Val	Val Asp Val Ser His
		Glu Asp Pro

1430	1435	1440
Glu Val	Lys Phe Asn Trp Tyr	Val Asp Gly Val Glu
1445	1450	1455
Ala Lys	Thr Lys Pro Arg Glu	Glu Gln Tyr Asn Ser
1460	1465	1470
Val Val	Ser Val Leu Thr Val	Leu His Gln Asp Trp
1475	1480	1485
		Leu Asn Gly

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

1490	1495	1500
Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro		
1505	1510	1515
Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn		
1520	1525	1530
Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp		
1535	1540	1545
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr		

1550	1555	1560
Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu		
1565	1570	1575
Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn		
1580	1585	1590
Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr		
1595	1600	1605
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		
1610	1615	

<210> 85

<211> 6033

<212> DNA

<213> Artificial Sequence

<220><223> FVIII 286 (FVIII-Fc with additional a2 region in between FVIII  
and Fc)

<400> 85

atgcaaatag agctctccac ctgtttctt ctgtgcctt tgcgattctg ctttagtgcc	60
accagaagat actacactggg tgcagtggaa ctgtcatggg actatatgca aagtgatctc	120
ggtagctgc ctgtggacgc aagatttcct cctagagtgc caaaatctt tccattcaac	180
acctcagtcg tgtacaaaaa gactctgttt gttagattca cggatcacct tttcaacatc	240
gctaagccaa ggccaccctg gatgggtctg ctaggtccta ccatccaggc tgaggttat	300

gatacagtgg tcattacact taagaacatg gttccccatc ctgtcagtct tcatgctgtt	360
-------------------------------------------------------------------	-----

ggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtcaaagg	420
gagaaaagaag atgataaagt ctccctggt ggaaggcata catatgtctg gcaggtcctg	480
aaagagaatg gtccaatggc ctctgaccca ctgtgccta cctactcata tctttctcat	540
gtggaccctgg taaaagactt gaattcaggc ctcattggag ccctactagt atgttagagaa	600
gggaggtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta	660
tttgcataatgg gaaaaagttt gcactcagaa acaaagaact cttgtatgca ggataggat	720
gctgcacatcg ctcggccctg gcctaaaatg cacacagtca atggttatgt aaacaggct	780
ctgccaggctc tgattggatg ccacaggaaa tcagtctatt ggcacatgtat tggaatggc	840
accacccctg aagtgcactc aatattccctc gaaggcaca catttttgtt gaggaccat	900
cggcaggctc gcttggaaat ctgcacata actttcctta ctgcacaaac actttgtatg	960
gaccttggac agtttctact gttttgtat atctttccc accaacatga tggcatggaa	1020
gcttatgtca aagtagacag ctgtccagag gaaccccaac tacgaatgaa aaataatgaa	1080
gaagcggaaag actatgtatgatgatgatgatgatgatgatgatgatgatgatgatgatgatgat	1140
gatgacaact ctcccccatt tatccaaatt cgctcagttt ccaagaagca tcctaaaact	1200
tgggtacatt acattgtgc tgaagaggag gactgggact atgctccctt agtcctcgcc	1260
cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg	1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa ccttaagac tcgtgaagct	1380
attcagcatg aatcaggaat cttggaccc ttactttatg gggaaatgg agacacactg	1440
ttgattatataa ttaagaatca agcaagcaga ccatataaca tctaccctca cgaaatcact	1500
gatgtccgtc ctgtgttattc aaggagatta ccaaaggatg taaaacattt gaaggatTTT	1560
ccaaattctgc caggagaaat attcaaatat aaatggacag tgactgttaga agatggccca	1620
actaaatcag atccctcggtg cctgacccgc tattactcta gtttcgttaa tatggagaga	1680
gatctagctt caggactcat tggccctctc ctcatctgtc acaaagaatc tgttagatcaa	1740
agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag	1800
aaccgaagct ggtacctcac agagaatata caacgcttcc tccccaaatcc agctggagtg	1860
cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatgtt	1920
tttgatagtt tgcagttgtc agtttgggg catgaggtgg catactggta cattctaagc	1980
attggagcac agactgactt cttttctgtc ttcttctctg gatatacctt caaacacaaaa	2040
atggctatg aagacacact cacccttattc ccattctcag gagaaactgt cttcatgtcg	2100
atggaaaacc caggctatg gattctgggg tgccacaact cagacttgc gaacagaggc	2160
atgaccgcct tactgaaggt ttcttagttt gacaagaaca ctgggtgatta ttacgaggac	2220

agtatgttgc	atatttcaggc	atacttgctg	agtaaaaaca	atgccattga	accaagaagc	2280
ttctctcaaa	acggcgcc	aggtaaccta	gagtctgcta	cccccgagtc	agggccagga	2340
ttagagccag	ccaccccg	gtctgagaca	ccggggactt	ccgagagtgc	cacccttag	2400
tccggacccg	ggtccgagcc	cgccacttcc	ggctccgaaa	ctcccgac	aagcgagagc	2460
gctaccccg	agtctcaggacc	aggaacatct	acagagccct	ctgaaggctc	cgctccagg	2520
tccccagccg	gcagtcccac	tagcaccgag	gagggAACCT	ctgaaAGCgc	cacacccgaa	2580
ttagggccag	ggtctgagcc	tgctaccagc	ggcagcgaga	caccaggcac	ctctgagtc	2640
gccacaccag	agtccggacc	cggtatccc	gctgggagcc	ccacccac	tgaggagg	2700
tctctgtgt	gctctccaac	atctactgag	gaaggtac	caaccgagcc	atccgagg	2760
ttagctcccg	gcacccataga	gtcgcaacc	ccggagtc	gaccggaaac	ttccgaaagt	2820
gccacaccag	agtccggtcc	cggtatcca	aatcagcaa	cacccgagtc	cggccctgg	2880
tctgaacccg	ccacaagtgg	tagttagaca	ccaggatcg	aacctgtac	ctcagggtca	2940
gagacacccg	gatctccggc	aggctcacca	acccactg	aggaggcac	cagcacagaa	3000
ccaagcgagg	gctccgcacc	cggaacaagc	actgaaccca	gtgagggttc	agcacccggc	3060
tctgagccgg	ccacaagtgg	cagtgagaca	ccggcactt	cagagagtgc	caccccgag	3120
agtggcccg	gcacttagtac	cgagccctt	gaaggcagt	cgccagc	ctc	3180
gtcttggaaac	gccatcaagc	tgaataact	cgtactactc	ttcagt	caga tcaagg	3240
atcgattatg	atgataccat	atcagtggaa	atgaagaagg	aagat	tttga catttat	3300
gaggatgaaa	atcagagccc	ccgcagctt	caaaagaaaa	cacgacacta	tttattgt	3360
gcagtgagaa	ggctctggaa	ttatggatg	atagctccc	cacatgttct	aagaacagg	3420
gctcagatg	gcagtgtccc	tcagtcaag	aaagttgttt	tccaggaatt	tactgatggc	3480
tcctttactc	agccctata	ccgtggagaa	ctaaatgaac	atttggact	cctggggcca	3540
tatataagag	cagaaggta	agataatatc	atggtaactt	tcagaaatca	ggcctctcg	3600
ccctattct	tctattct	ccttatttct	tatgaggaag	atcagaggca	aggagg	3660
cctagaaaaaa	actttgtcaa	gcctaattgaa	acccaaaactt	actttggaa	agtgcacat	3720
catatggcac	ccactaaaga	tgagttgac	tgcaaagct	gggttattt	ctctgatgtt	3780
gaccttggaaa	aagatgtgca	ctcaggcctg	attggacccc	ttctggctg	ccacactaac	3840
acactgaacc	ctgctcatgg	gagacaagtg	acagttacagg	aatttgc	tct	3900
atctttgtatg	agaccaaaag	ctggacttc	actgaaaata	tggaaagaaa	ctgcagg	3960
ccctgcaata	tccagatgga	agatcccact	tttaaagaga	attatcg	tttccatgc	4020

aatggctaca taatggatac actacctggc ttagtaatgg ctcaggatca aaggattcga	4080
tggtatctgc tcagcatggg cagcaatgaa aacatccatt ctattcattt cagtggacat	4140
gtgttcactg tacaaaaaaa agaggagtat aaaatggcac tgtacaatct ctatccaggt	4200
gttttgaga cagtggaaat gttaccatcc aaagctggaa ttggcggtt ggaatgcctt	4260
attggcgagc atctacatgc tggtatggc acacttttc tggtgtacag caataagtgt	4320
cagactcccc tggaaatggc ttctggacac attagagatt tttagattt acgtttcagga	4380
caatatggac agtggggccc aaagctggcc agacttcatt attccggatc aatcaatgcc	4440
tggagcacca aggagccctt ttcttggatc aagggtggatc tggtggcacc aatgattatt	4500
cacggcatca agacccaggg tgcccgtag aagttctcca gcctctacat ctctcagttt	4560
atcatcatgt atagtcttga tggaaagaag tggcagactt atcgaggaaa ttccactgga	4620
accttaatgg tcttcttgg caatgtggat tcacatggtaaaaacacaa tatttttaac	4680
cctccaatta ttgctcgata catccgttg cacccaaactc attatagcat tcgcagcact	4740
cttcgcatgg agtttatggg ctgtgattt aatagttgca gcatgccatt gggaaatggag	4800
agtaaagcaa tatcagatgc acagattact gcttcatcct actttaccaa tatgtttgcc	4860
acctggtctc cttaaaaagc tcgacttcac ctccaaggga ggagtaatgc ctggagacct	4920
caggtgaata atccaaaaga gtggctgcaa gtggacttcc agaagacaat gaaagtaca	4980
ggagtaacta ctcaaggagt aaaatctctg cttaccagca tgtatgtgaa ggaggttcctc	5040
atctccagca gtcaagatgg ccatcagtgg actctttt ttcagaatgg caaagtaag	5100
gttttcagg gaaatcaaga ctccttcaca cctgtggta actctctaga cccaccgtta	5160
ctgactcgct accttcgaat tcaccccccag agttgggtgc accagattgc cctgaggatg	5220
gagggtctgg gctgcgagggc acaggaccc tacgacaaga acactggta ttattacgag	5280
gacagttatg aagatatttc agcatacttg ctgagtaaaa acaatgccat tgaaccaaga	5340
agcttctctg acaaaaactca cacatgccca ccgtgccca ctcagaact cctggcgga	5400
ccgtcagtct tcctttccc cccaaaaccc aaggacaccc tcatgatctc ccggaccct	5460
gaggtcacat gcgtgggtgtt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg	5520
tacgtggacg gcgtggaggt gcataatgcc aagacaagc cgccggagga gcagtacaac	5580
agcacgtacc gtgtggtcag cgtcctcacc gtccctgacc aggactggct gaatggcaag	5640
gagtacaagt gcaaggcttc caacaaagcc ctcccaagcc ccatcgagaa aaccatctcc	5700
aaagccaaag ggcagccccc agaaccacag gtgtacaccc tgccccatc ccggatgag	5760

ctgaccaaga accaggtcag cctgacctgc ctggtaaaag gtttctatcc cagcgacatc 5820  
 gccgtggagt gggagagcaa tggcagccg gagaacaact acaagaccac gcctccgtg 5880  
 ttggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcagggtgg 5940  
 cagcagggga acgtcttctc atgctccgtg atgcattgagg ctctgcacaa ccactacacg 6000  
 cagaagagcc tctccctgtc tccggtaaa tga 6033

&lt;210&gt; 86

&lt;211&gt; 1991

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; FVIII 286 (FVIII-Fc with additional a2 region in between FVIII

and Fc)

&lt;400&gt; 86

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 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 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
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 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 Ala 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 Asn Thr Gly Asp Tyr Tyr Glu Asp Ser  
 1730 1735 1740

Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile  
 1745 1750 1755

Glu Pro Arg Ser Phe Ser Asp Lys Thr His Thr Cys Pro Pro Cys  
 1760 1765 1770

Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro  
 1775 1780 1785

Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val  
 1790 1795 1800

Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys  
 1805 1810 1815

Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr  
 1820 1825 1830

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser

1835                    1840                    1845  
 Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr  
 1850                    1855                    1860

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys  
 1865                    1870                    1875  
 Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr  
 1880                    1885                    1890  
 Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser  
 1895                    1900                    1905  
 Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val  
 1910                    1915                    1920

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr  
 1925                    1930                    1935  
 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys  
 1940                    1945                    1950  
 Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser  
 1955                    1960                    1965  
 Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys  
 1970                    1975                    1980

Ser Leu Ser Leu Ser Pro Gly Lys

1985                    1990

<210> 87

<211> 5937

<212> DNA

<213> Artificial Sequence

<220><223> FVIII 169

<400> 87

atgcaaata	g	actctccac	ctgtttt	ctgtgcctt	tgcgattctg	c	tttagtgcc	60	
accagaagat	a	ctacacctgg	tgca	gtggaa	ctgtcatgg	a	ctatatgca	aagtgtatc	120
ggtagactgc	c	ctgtggacgc	aagattcct	cctagagtgc	caaaatctt	tccattcaac			180
acctcagtcg	t	tgtacaaaaa	gactctgttt	gtagaattca	cggatcac	tttcaacatc			240

gctaagccaa gcccaccctg gatgggtctg ctagtccta ccatccaggc tgaggttat	300
gatacagtgg tcattacact taagaacatg gcttccatc ctgtcagtct tcatgctgtt	360
ggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtc当地 gagaaagaag atgataaagt cttccctgg ggaagccata catatgtctg gcaggcctg aaagagaatg gtccaatggc ctctgacca ctgtgccta cctactcata tcttcctcat gtggacctgg taaaagactt gaattcaggc ctcattggag ccctactagt atgtagagaa gggaggtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact tttgctgta tttgatgaag gaaaaagtgg cactcagaa acaaagaact cttgatgca ggataggat	420 480 540 600 660 720
gctgcacatcg ctccggcctg gcctaaaatg cacacagtca atggttatgt aaacaggct ctgcccaggc tgattggatg ccacaggaaa tcagtcattt ggcacatgtat tggaatggc accactcctg aagtgcactc aatattcctc gaaggtcaca catttctgt gaggaaccat cgccaggcta gcttggaaat ctgcacataa actttccta ctgtcaaac actttgatg gaccttggac agtttctact gtttgcattt atctttccc accaacatga tggcatgaa gcttatgtca aagttagacag ctgtccagag gaaccccaac tacgaatgaa aaataatgaa gaagcggaaag actatgatga tgcatttact gattctgaaa tggatgtgg caggttgc	780 840 900 960 1020 1080 1140
gatgacaact ctccctt tatccaaatt cgctcagttg ccaagaagca tcctaaaact tgggtacatt acattgcgc tgaagaggag gactggact atgctccctt agtccctgcc cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg aagtacaaaa aagtccgatt tatggcatac acagatgaaa ccttaagac tcgtgaagct attcagcatg aatcaggaat ctgggaccc ttactttatg gggaaatgg agacacactg ttgattataat ttaagaatca agcaaggaga ccatataaca tctaccctca cggaatact gatgtccgtc ctgttattt aaggagatta ccaaaggatg taaaacattt gaaggattt	1200 1260 1320 1380 1440 1500 1560
ccaattctgc caggagaaat attcaaatat aaatggacag tgactgtaga agatggcca actaaatcag atcctcggtg cctgacccgc tattactcta gttcgtaa tatggagaga gatctagctt caggactcat tggccctctc ctcattgtct acaaagaatc tgttagatcaa agaggaaacc agataatgtc agacaaggagg aatgtcatcc tgtttctgt atttgatgag aaccgaagct ggtacctcac agagaatata caacgcttcc tccccatcc agctggatg cagcttggg atccagagtt ccaaggctcc aacatcatgc acagcatcaa tggctatgtt tttgatagtt tgcagttgtc agttgttg catgaggatgg catactggta cattctaagc	1620 1680 1740 1800 1860 1920 1980

attggagcac agactgactt cctttctgtc ttcttctcg gatataccctt caaacacaaa	2040
atggcttatg aagacacact cacccatttc ccattctcg gagaaactgt ctcatgtcg	2100
atggaaaacc caggcttatg gattctgggg tgccacaact cagacttgc gaacagaggc	2160
atgaccgcct tactgaaggt ttcttagtgtt gacaagaaca ctggtgattt ttacgaggac	2220
agttatgaag atattcagc atacttgctg agtaaaaaca atgccattga accaagaagc	2280
ttctctcaaa acggcgcgcc aggtacctca gagtctgcta ccccgagtc agggccagga	2340
tcagagccag ccaccccggtctgagaca cccggactt ccgagagtgc caccctgag	2400
tccggacccg ggtccgagcc cgccacttcc ggctccgaaa ctccggcac aagcgagac	2460
gctaccccg agtcaggacc aggaacatct acagagccct ctgaaggctc cgctccaggg	2520
tccccagccg gcagtcccac tagcaccgag gaggAACCT ctgaaagcgc cacacccgaa	2580
tcagggccag ggtctgagcc tgctaccagc ggcaggaga caccaggcac ctctgagtcc	2640
gccacaccag agtccggacc cggatctccc gctgggagcc ccacccac tgaggaggga	2700
tctctgtctg gctctccaac atctactgag gaaggtacctt caaccgagcc atccgaggga	2760
tcagctcccg gcacccaga gtcggcaacc ccggactctg gacccggAAC ttccgaaagt	2820
gccacaccag agtccggtcc cgggacttca gaatcagcaa cacccgagtc cggccctggg	2880
tctgaacccg ccacaagtgg tagtgagaca ccaggatcg aacctgtac ctcagggtca	2940
gagacacccg gatctccggc aggctccacca acctccactg aggagggcac cagcacagaa	3000
ccaagcgagg gctccgcacc cggaaacaagc actgaaccca gtgagggttc agcacccggc	3060
tctgagccgg ccacaagtgg cagttagaca cccggactt cagagagtgc caccccgag	3120
agtggcccg gcactagtac cgagccctt gaaggcagtg cgccagcctc gagcccaacca	3180
gtcttgaac gccatcaagc tgaataact cgtactactc ttcatgtcaga tcaagaggaa	3240
atcgattatg atgataccat atcagttgaa atgaagaagg aagattttga catttatgt	3300
gaggatgaaa atcagagccc ccgcgcgttt caaaaagaaaa cacgacacta ttttattgt	3360
gcagtgaggaa ggctctggaa ttatggatg agtagtccc cacatgtct aagaaacagg	3420
gctcagatgt gcagtgtccc tcagttcaag aaagttgtt tccaggaaatt tactgtatgc	3480
tcctttactc agcccttata ccgtggagaa ctaaatgaac atttggact cctggggcca	3540
tatataagag cagaagttga agataatatc atggtaactt tcagaaatca ggcctctgt	3600
ccctatttct tctattctat ctttatttct tatgaggaag atcagaggca aggaggcagaa	3660
cctagaaaaa actttgtcaa gcctaattgaa accaaaaactt actttggaa agtgcacat	3720
catatggcac ccactaaaga ttagttgac tgcaaaaggctt gggcttattt ctctgatgtt	3780
gacctggaaa aagatgtgca ctcaggcctg attggacccc ttctggctg ccacactaac	3840

acactgaacc ctgctcatgg gagacaagtg acagtagcagg aatttgctct gttttcacc	3900
atctttagt agaccaaaag ctggtaacttc actgaaaata tggaaagaaaa ctgcaggct	3960
ccctgcaata tccagatgga agatcccact tttaaagaga attatcgctt ccatgcaatc	4020
aatggctaca taatggatac actacctggc ttagtaatgg ctcaggatca aaggattcga	4080
tggtatctgc tcagcatggg cagcaatgaa aacatccatt ctattcattt cagtggacat	4140
gtgttcactg tacgaaaaaa agaggagtat aaaatggcac tgtacaatct ctatccaggt	4200
gttttgaga cagtgaaat gttaccatcc aaagctggaa ttggcgggt ggaatgcctt	4260
attggcgagc atctacatgc tggatgagc acacttttc tgggtacag caataagtgt	4320
cagactcccc tggaaatggc ttctggacac attagagatt ttcaagattt acgttcagga	4380
caatatggac agtggggcccc aaagctggcc agacttcatt attccggatc aatcaatgcc	4440
tggagcacca aggagccctt ttcttgatc aaggtggatc tggtggcacc aatgattatt	4500
cacggcatca agacccagggg tgcccgtcag aagttctcca gcctctacat ctctcagttt	4560
atcatcatgt atagtcttga tggaaagaag tggcagactt atcgaggaaa ttccactgga	4620
accttaatgg tcttcttgg caatgtggat tcattggtaaaaaacacaa tatttttaac	4680
cctccaatta ttgctcgata catccgtttg caccaactc attatagcat tcgcagcact	4740
cttcgcatttgg agttgatggg ctgtgatTTT aatagttgca gcatgccatt gggatggag	4800
agtaaagcaa tatcagatgc acagattact gcttcatcct actttaccaa tatgtttgcc	4860
acctggctc ctcaaaagc tcgacttcac ctccaaggga ggagtaatgc ctggagacct	4920
caggtaata atccaaaaga gtggctgcaa gtggacttcc agaagacaat gaaagtacaca	4980
ggagtaacta ctcagggagt aaaatctctg cttaccagca tgtatgtgaa ggagttcctc	5040
atctccagca gtcaagatgg ccatcagtgg actctttttt ttcagaatgg caaagtaaag	5100
gttttcagg gaaatcaaga ctccttcaca cctgtggta actctctaga cccaccgtta	5160
ctgactcgct accttcgaat tcaccccccag agttgggtgc accagattgc cctgaggatg	5220
gaggTTCTGG gctcgagggc acaggaccc tacgacaaaaa ctcacacatg cccaccgtgc	5280
ccagctccag aactcctggg cgaccgtca gtcttcctt tcccccaaa acccaaggac	5340
accctcatga tctccggac ccctgagggtc acatgcgtgg tgggtggacgt gagccacgaa	5400
gaccctgagg tcaagttcaa ctggtaacgtg gacggcgtgg aggtgcataa tgccaagaca	5460
aagccgcggg aggagcagta caacagcagc taccgtgtgg tcagcgtctt caccgtcctg	5520
caccaggact ggctgaatgg caaggagtac aagtgcagg tctccaacaa agccctccca	5580
gcccccatcg agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac	5640
accctgcccc catccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggc	5700

aaaggcttct atcccagcga catcgccgtg gagtgggaga gcaatggca gccggagaac 5760

aactacaaga ccacgcctcc cgtgttggac tccgacggct ctttcttctt ctacagcaag 5820

ctcacccgtgg acaagagcag gtggcagcag gggAACGTCT tctcatgctc cgttatgcat 5880

gaggctctgc acaaccacta cacgcagaag agcctctccc tgtctccggg taaatga 5937

<210> 88

<211> 1978

<212> PRT

<213> Artificial Sequence

<220><223> FVIII 169

<400> 88

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
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
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
Trp Val His Tyr Ile Ala Ala 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
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
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
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 Ala Pro Cys Asn		
1310	1315	1320
Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His		
1325	1330	1335
Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met		
1340	1345	1350
Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser		
1355	1360	1365
Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr		
1370	1375	1380

Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr

1385 1390 1395

Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly

1400 1405 1410

Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly

1415 1420 1425

Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro

1430 1435 1440

Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala

1445 1450 1455

Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His

1460 1465 1470

Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser

1475 1480 1485

Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile

1490 1495 1500

Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser

1505 1510 1515

Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr

1520 1525 1530

Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn

1535 1540 1545

Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile

1550 1555 1560

Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg

1565 1570 1575

Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys

1580 1585 1590

Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln

1595 1600 1605

Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser

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

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

1850 1855 1860

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln

1865 1870 1875

Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln

1880 1885 1890

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

1895 1900 1905

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys

1910 1915 1920

Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr

1925 1930 1935

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val

1940 1945 1950

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr

1955 1960 1965

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys

1970 1975

<210> 89

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> Gly/Ser linker

<400> 89

Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser

1 5 10 15

<210> 90

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Gly/Ser linker

<400> 90

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser

20

<210> 91

<211> 5379

<212> DNA

<213> Artificial Sequence

<220><223> VWF034

<400> 91

atgattcctg ccagatttgc cgggtgtcg ctgtctgg ccctcattt gccagggacc	60
ctttgtgcag aaggaactcg cggcaggta tccacggccc gatgcagcct tttcggaagt	120

gacttcgtca acaccttga tgggagcatg tacagcttg cggtatactg cagttacctc	180
ctggcagggg gctgccagaa acgctcttc tcgattattg gggacttcca gaatggcaag	240
agagtgagcc tctccgtgta tcttgggaa tttttgaca tccatttgtt tgtcaatgg	300
accgtgacac agggggacca aagagtctcc atgccctatg cctccaaagg gctgtatcta	360
gaaactgagg ctgggtacta caagctgtcc ggtgaggcct atggcttgc ggccaggatc	420
gatggcagcg gcaacttca agtccgtcg tcagacagat acttcaacaa gacctgcggg	480
ctgtgtggca actttaacat ctttgcgaa gatgacttta tgacccaaga agggacattg	540

acctcggacc ctatgactt tgccaactca tgggtctga gcagtggaga acagtgggt	600
gaacgggcat ctcccccag cagctcatgc aacatctctt ctggggaaat gcagaaggc	660
ctgtgggagc agtgcagct tctgaagagc acctcggtt ttgcccgtc ccaccctcg	720
gtggaccccg agcctttgt gcccctgtgt gagaagactt tgtgtgagtg tgctgggggg	780
ctggagtgcg cctgccctgc ctcctggag tacgcccga cctgtgccca ggaggaaatg	840
gtgctgtacg gctggaccga ccacagcgcg tgcagccag tgcctgc tggtatggag	900
tataggcagt gtgtgtcccc ttgcgccagg acctgccaga gcctgcacat caatgaaatg	960

tgtcaggagc gatgcgtgga tggctgcagc tgccctgagg gacagctctt ggtatgaaggc	1020
ctctgcgtgg agagcaccga gtgtccctgc gtgcattccg gaaagcgcta ccctccggc	1080
acccctctt ctcgagactg caacacctgc atttgcgaa acagccagt gatctgcagc	1140
aatgaagaat gtccaggaga gtgcctgtc actggtaat cccactcaa gagcttgc	1200
aacagatact tcacccatg tggatctgc cagttacgtc tggccggga ttgcaggac	1260
cactcctctt ccattgtcat tgagactgtc cagttacgtc atgaccgcga cgctgtgtgc	1320

acccgctccg tcaccgtccg gtcgcctggc ctgcacaaca gccttgtgaa actgaagcat	1380
ggggcaggag ttgccatgga tggccaggac atccagctcc ccctccctgaa aggtgaccc	1440
cgcattccagc atacagtgac ggcctccgtcg cgcctcagct acggggagga cctgcagatg	1500
gactgggatg gccgcgggag gctgctggtg aagctgtccc cgcgttatgc cgggaagacc	1560
tgcggcctgt gtgggaatta caatggcaac caggcgcacg acttccttac cccctctgg	1620
ctggcggagc cccgggtgga ggacttcggg aacgcctgga agctgcacgg ggactgccag	1680
gacctgcaga agcagcacag cgatccctgc gcccctaacc cgcgcatacg cagttctcc	1740
gaggaggcgt gcgcggtcct gacgtcccc acattcgagg cctgcacatcg tgccgtcagc	1800
ccgctgcctt acctgcgaa ctggcgctac gacgtgtgct cctgctcgga cggccgcgag	1860
tgcctgtcg gcgcctggc cagctatgcc gcggcctgca cggggagagg cgtgcgcgtc	1920
gcgtggcgcg agccaggccg ctgtgagctg aactgcccga aaggccaggt gtacctgcag	1980
tgcgggacc cctgcaacct gacctgcccgc tctctcttt accggatga ggaatgaat	2040
gaggccgtgcc tggaggctg cttctgtcccc ccaggctct acatggatga gagggggac	2100
tgcgtgcaca aggcccagtg cccctgttac tatgacggtg agatcttcca gccagaagac	2160
atcttctcag accatcacac catgtgctac tgtgaggatg gttcatgca ctgtaccatg	2220
agtggagtcc ccgaaagctt gtcgcctgac gctgtcctca gcagtccct gtctcatcgc	2280
agcaaaaagga gcctatccctg tcggccccc atggtaaagc tgggtgtcc cgctgacaac	2340
ctgcggcgt aagggtcgat gtgtacccaa acgtgccaga actatgaccc ggagtgcata	2400
agcatgggct gtgtctctgg ctgcctctgc ccccccggca tggtccggca tgagaacaga	2460
tgtgtggccc tggaaaggatg tccctgcttc catcaggggca aggatgtatgc ccctggagaa	2520
acagtgaaga ttggctgcaa cacttgtgtc tgcgggacc ggaagtggaa ctgcacagac	2580
catgtgtgtg atgccacgtg ctccacgatc ggcacggccc actacccatc cttcgacggg	2640
ctcaaataacc tttccccgg ggagtgcac tacgttctgg tgcaggatta ctgcggcagt	2700
aacctggga ctttcggat cctagtgggg aataaggat gcaagccaccc ctcagtgaaa	2760
tgcaagaaac gggtcacccat cctgggtggag ggaggagaga ttgagctgtt tgacggggag	2820
gtgaatgtga agaggcccat gaaggatgag actcaacttg aggtggtgaa gtctggccgg	2880
tacatcattc tgcgtctggg caaagccctc tccgtggct gggaccgcca cctgacatc	2940
tccgtggcc tgaaggcagac ataccaggag aaagtgtgtc gcctgtgtgg gaattttgat	3000
ggcatccaga acaatgaccc caccaggcgc aacccatccaa tggaggaaga ccctgtggac	3060

tttggaaact cctggaaagt gagctcgag tgtgctgaca ccagaaaagt gcctctggac	3120
tcatccctg ccacctgcca taacaacatc atgaaggaga cgatggtgga ttcctcctgt	3180
agaatccta ccagtgacgt cttccaggac tgcaacaagc tggggaccc cgagccatat	3240
ctggaitct gcattiacga cacctgcicc tgtgagtcca tggggactg cgccgcattc	3300
tgcgacacca ttgctgccta tgcccacgtg tgtgccagc atggcaaggt ggtgacctgg	3360
aggacggcca cattgtgccc ccagagctgc gaggagagga atctccggga gaacgggtat	3420
gaggctgagt ggcgctataa cagctgtgca cctgcctgtc aagtacgtg tcagcacct	3480
gagccactgg cctgcccgt gcagtgtgtg gagggctgcc atgcccactg ccctccaggg	3540
aaaatctgg atgagcttt gcagacctgc gttgaccctg aagactgtcc agtgtgtgag	3600
gtggctggcc ggcgtttgc ctcaggaaag aaagtacacct tgaatcccag tgaccctgag	3660
cactgccaga ttgccactg tcatgttgta aacctcacct gtgaaggctg ccaggagccg	3720
atatcgggta cctcagagtc tgctaccccc gagtcagggc caggatcaga gccagccacc	3780
tccgggtctg agacacccgg gacttccgag agtgcaccc ctgagttccg accccgggtcc	3840
gagcccgcca ctccggctc cgaactccc ggcacaagcg agagcgtac cccagagtca	3900
ggaccaggaa catctacaga gccctctgaa ggctccgctc cagggcccc agccggcagt	3960
cccaactagca ccgaggaggg aacctctgaa agcgcacac ccgaatcagg gccagggtct	4020
gagcctgcta ccagcggcag cgagacacca ggcacctctg agtccggccac accagagtcc	4080
ggacccggat ctcccgtgg gagccccacc tccactgagg agggatctcc tgctggctct	4140
ccaacatcta ctgaggaagg tacctaacc gagccatccg agggatcagc tccggcacc	4200
ttagactcg caacccggaa gtctggaccc ggaacttccg aaagtgccac accagagtcc	4260
gttccggga cttcagaatc agcaacaccc gagtcggcc ctgggtctga acccgccaca	4320
agtggtagtg agacaccagg atcagaacct gctacctcag ggtcagagac accccggatct	4380
ccggcaggct caccaacctc cactgaggag ggcaccagca cagaaccaag cgagggctcc	4440
gcacccggaa caagcactga acccagttag ggtcagcac ccggctctga gcccggcaca	4500
agtggcagtg agacacccgg cacttcagag agtgcaccc ccgagagtg cccaggcact	4560
agtaccgagc cctctgaagg cagtgcgcca gattctggcg gtggaggttc cggtggcggg	4620
ggatccggtg ggggggatc cggtggcggg ggatccggtg ggggggatc cctggcccc	4680
cggggcagcg gaggcgacaa aactcacaca tgcccaccgt gcccagctcc agaactcctg	4740
ggcgaccgt cagtcttct cttccccca aaacccaagg acaccctcat gatctccgg	4800
acccctgagg tcacatgcgt ggtggtgac gtgagccacg aagaccctga ggtcaagttc	4860
aactggtagc tggacggcgt ggaggtgcat aatgccaaga caaagccgca ggaggagcag	4920

tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctaat	4980
ggcaaggagt acaaagtgcaa ggtctccaac aaagccctcc cagccccat cgagaaaacc	5040
atctccaaag ccaaaggcga gccccgagaa ccacaggtgt acaccctgcc cccatcccgg	5100
gatgagctga ccaagaacca ggtcagcctg acctgtctgg tcaaaggctt ctatcccagc	5160

gacatcgccg tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct	5220
cccggtttgg actccgacgg ctcttcttc ctctacagca agctcaccgt ggacaagagc	5280
aggtggcagc aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac	5340
tacacgcaga agagcctctc cctgtctccg ggttaatga	5379

&lt;210&gt; 92

&lt;211&gt; 1778

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; VWF034

&lt;400&gt; 92

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
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
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                    400

Asn Arg Tyr Phe Thr Phe Ser Gly Ile Cys Gln Tyr Leu Leu Ala Arg  
 405                    410                    415

Asp Cys Gln Asp His Ser Phe Ser Ile Val Ile Glu Thr Val Gln Cys  
 420                    425                    430

Ala Asp Asp Arg Asp Ala Val Cys Thr Arg Ser Val Thr Val Arg Leu  
 435                    440                    445

Pro Gly Leu His Asn Ser Leu Val Lys Leu Lys His Gly Ala Gly Val  
 450                    455                    460

Ala Met Asp Gly Gln Asp Ile Gln Leu Pro Leu Leu Lys Gly Asp Leu  
 465                    470                    475                    480

Arg Ile Gln His Thr Val Thr Ala Ser Val Arg Leu Ser Tyr Gly Glu  
 485                    490                    495

Asp Leu Gln Met Asp Trp Asp Gly Arg Gly Arg Leu Leu Val Lys Leu  
 500                    505                    510

Ser Pro Val Tyr Ala Gly Lys Thr Cys Gly Leu Cys Gly Asn Tyr Asn  
 515                    520                    525

Gly Asn Gln Gly Asp Asp Phe Leu Thr Pro Ser Gly Leu Ala Glu Pro  
 530                    535                    540

Arg Val Glu Asp Phe Gly Asn Ala Trp Lys Leu His Gly Asp Cys Gln  
 545                    550                    555                    560

Asp Leu Gln Lys Gln His Ser Asp Pro Cys Ala Leu Asn Pro Arg Met  
 565                    570                    575

Thr Arg Phe Ser Glu Glu Ala Cys Ala Val Leu Thr Ser Pro Thr Phe  
 580                    585                    590

Glu Ala Cys His Arg Ala Val Ser Pro Leu Pro Tyr Leu Arg Asn Cys  
 595                    600                    605

Arg Tyr Asp Val Cys Ser Cys Ser Asp Gly Arg Glu Cys Leu Cys Gly  
 610                    615                    620

Ala Leu Ala Ser Tyr Ala Ala Ala Cys Ala Gly Arg Gly Val Arg Val

625	630	635	640
Ala Trp Arg Glu Pro Gly Arg Cys Glu Leu Asn Cys Pro Lys Gly Gln			
645	650	655	
Val Tyr Leu Gln Cys Gly Thr Pro Cys Asn Leu Thr Cys Arg Ser Leu			
660	665	670	
Ser Tyr Pro Asp Glu Glu Cys Asn Glu Ala Cys Leu Glu Gly Cys Phe			
675	680	685	
Cys Pro Pro Gly Leu Tyr Met Asp Glu Arg Gly Asp Cys Val Pro Lys			
690	695	700	
Ala Gln Cys Pro Cys Tyr Tyr Asp Gly Glu Ile Phe Gln Pro Glu Asp			
705	710	715	720
Ile Phe Ser Asp His His Thr Met Cys Tyr Cys Glu Asp Gly Phe Met			
725	730	735	
His Cys Thr Met Ser Gly Val Pro Gly Ser Leu Leu Pro Asp Ala Val			
740	745	750	
Leu Ser Ser Pro Leu Ser His Arg Ser Lys Arg Ser Leu Ser Cys Arg			
755	760	765	
Pro Pro Met Val Lys Leu Val Cys Pro Ala Asp Asn Leu Arg Ala Glu			
770	775	780	
Gly Leu Glu Cys Thr Lys Thr Cys Gln Asn Tyr Asp Leu Glu Cys Met			
785	790	795	800
Ser Met Gly Cys Val Ser Gly Cys Leu Cys Pro Pro Gly Met Val Arg			
805	810	815	
His Glu Asn Arg Cys Val Ala Leu Glu Arg Cys Pro Cys Phe His Gln			
820	825	830	
Gly Lys Glu Tyr Ala Pro Gly Glu Thr Val Lys Ile Gly Cys Asn Thr			
835	840	845	
Cys Val Cys Arg Asp Arg Lys Trp Asn Cys Thr Asp His Val Cys Asp			
850	855	860	
Ala Thr Cys Ser Thr Ile Gly Met Ala His Tyr Leu Thr Phe Asp Gly			
865	870	875	880

Leu Lys Tyr Leu Phe Pro Gly Glu Cys Gln Tyr Val Leu Val Gln Asp  
 885                    890                    895  
 Tyr Cys Gly Ser Asn Pro Gly Thr Phe Arg Ile Leu Val Gly Asn Lys  
 900                    905                    910

Gly Cys Ser His Pro Ser Val Lys Cys Lys Lys Arg Val Thr Ile Leu  
 915                    920                    925  
 Val Glu Gly 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 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 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		