(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년07월06일
(11) 등록번호 10-2551733
(24) 등록일자 2023년06월30일

(51) 국제특허분류(Int. Cl.)
A61K 38/48 (2006.01) *A61K 48/00* (2006.01)
A61P 7/04 (2006.01) *C07K 14/745* (2006.01)
C12N 15/62 (2006.01) *C12N 15/86* (2006.01)
(52) CPC특허분류
A61K 38/4846 (2013.01)
A61K 48/0058 (2013.01)
(21) 출원번호 10-2022-7012028(분할)
(22) 출원일자(국제) 2018년05월22일
심사청구일자 2022년04월14일
(85) 번역문제출일자 2022년04월11일
(65) 공개번호 10-2022-0050238
(43) 공개일자 2022년04월22일
(62) 원출원 특허 10-2019-7037828
원출원일자(국제) 2018년05월22일
심사청구일자 2021년04월02일
(86) 국제출원번호 PCT/US2018/033866
(87) 국제공개번호 WO 2018/217731
국제공개일자 2018년11월29일
(30) 우선권주장
62/509,616 2017년05월22일 미국(US)

(56) 선행기술조사문헌
WO2006036502 A2
WO2014064277 A1
WO2016146757 A1
WO2016210170 A1

(73) 특허권자
다케다 야쿠힌 고교 가부시키가이샤
일본 오사까시 쥬오구 도쇼마찌 4-1-1
(72) 발명자
홀링 프란지스카
오스트레일리아 2230 가엔세른돌프 바우른펠트가쎄 3/9
랑글레 요하네스
오스트리아 1020 비엔나 스로츠버그스트라베 2/23
(뒷면에 계속)
(74) 대리인
양영준, 박보현

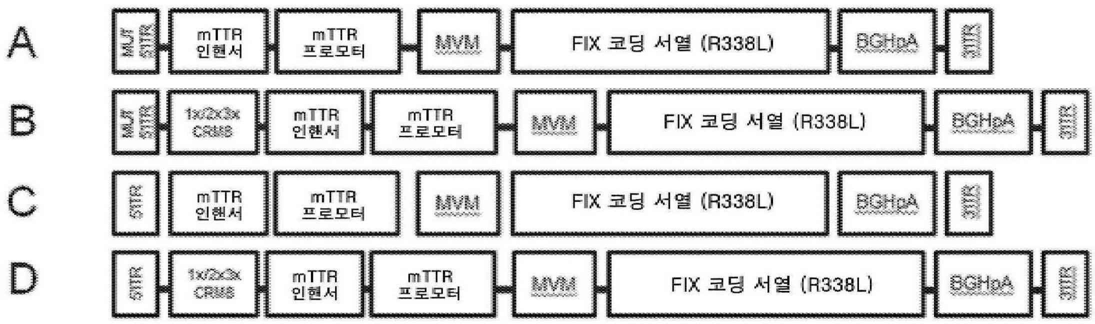
전체 청구항 수 : 총 24 항 심사관 : 최홍석

(54) 발명의 명칭 **B형 혈우병 유전자 요법을 위한 증가된 발현을 가지는 재조합 FIX 변이체를 암호화하는 바이러스 벡터**

(57) 요 약

본 명세서는 다른 측면들 중에서도, 포유류 세포에서의 발현을 위해 인자 IX 변이체를 암호화하는 코돈-변경된 폴리뉴클레오티드를 제공한다. 일부 실시 양태에서, 본 개시는 또한 B형 혈우병을 치료하기 위한 포유류 유전자 요법 벡터 및 방법을 제공한다.

대 표 도



(52) CPC특허분류
A61K 48/0066 (2013.01)
A61P 7/04 (2018.01)
C07K 14/745 (2013.01)
C12N 15/62 (2013.01)
C12N 15/86 (2013.01)
C12N 2750/14143 (2013.01)
C12N 2800/22 (2013.01)
C12N 2830/008 (2013.01)
C12N 2830/42 (2013.01)

(72) 발명자
포크너 팔코 군터
오스트리아 2304 오스/도나우 노이지들자일러 76아
로텐슈타이너 한스피터
오스트리아 1020 비엔나 헤이드가쎄 10/17
쉐이프린거 프리드리히
오스트리아 1090 비엔나 미헬베움가쎄 4/17

명 세 서

청구범위

청구항 1

B형 혈우병의 치료를 위한, 인자 IX 단백질을 암호화하는 인자 IX 폴리뉴클레오티드를 포함하는 핵산 조성물을 포함하는 제약 조성물이며, 상기 인자 IX 폴리뉴클레오티드는 서열 번호 17의 핵산 서열과 적어도 99% 동일한 핵산 서열을 포함하는 것인 제약 조성물.

청구항 2

제1항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 10개 이하의 CpG 디뉴클레오티드를 갖는 것인, 제약 조성물.

청구항 3

제1항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 3개 이하의 CpG 디뉴클레오티드를 갖는 것인, 제약 조성물.

청구항 4

제1항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 서열 번호 17의 핵산 서열을 포함하는 것인, 제약 조성물.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 서열 번호 2의 잔기 384에 대응하는 아미노산 위치에서 류신을 암호화하는 것인, 제약 조성물.

청구항 6

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드에 의해 암호화된 상기 인자 IX 단백질이 서열 번호 12와 비교하여 1 내지 10 개의 아미노산 치환을 갖는 것인, 제약 조성물.

청구항 7

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드에 의해 암호화된 상기 인자 IX 단백질이 서열 번호 12의 아미노산 서열을 갖는 것인, 제약 조성물.

청구항 8

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 프리-프로-리더 펩티드를 암호화하는 프리-프로-리더 폴리뉴클레오티드를 추가로 포함하고, 상기 프리-프로-리더 펩티드가 서열 번호 36의 아미노산 서열을 포함하는 것인, 제약 조성물.

청구항 9

제8항에 있어서, 상기 프리-프로-리더 폴리뉴클레오티드는 서열 번호 23의 핵산 서열을 갖는 것인, 제약 조성물.

청구항 10

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 서열 번호 9의 핵산 서열과 적어도 99% 동일한 핵산 서열을 갖는 것인, 제약 조성물.

청구항 11

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드는 서열 번호 9의 핵산 서열을 갖는 것인, 제약 조성물.

청구항 12

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드에 작동 가능하게 연결된 간-특이적 프로모터 요소를 추가로 포함하는, 제약 조성물.

청구항 13

제12항에 있어서, 상기 간-특이적 프로모터 요소가 프로모터 폴리뉴클레오티드의 하나의 카피를 포함하고, 상기 프로모터 폴리뉴클레오티드는 서열 번호 39와 적어도 95% 동일한 핵산 서열을 포함하는 것인, 제약 조성물.

청구항 14

제12항에 있어서, 상기 간-특이적 프로모터 요소가 3개 카피의 프로모터 폴리뉴클레오티드를 포함하고, 상기 프로모터 폴리뉴클레오티드는 서열 번호 39와 적어도 95% 동일한 핵산 서열을 포함하는 것인, 제약 조성물.

청구항 15

제13항에 있어서, 상기 프로모터 폴리뉴클레오티드는 서열 번호 39의 핵산 서열을 포함하는 것인, 제약 조성물.

청구항 16

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 인자 IX 폴리뉴클레오티드에 작동 가능하게 연결된 인트론을 추가로 포함하는, 제약 조성물.

청구항 17

제16항에 있어서, 상기 인트론은 서열 번호 53과 적어도 95% 동일한 핵산 서열을 포함하는 MVM 인트론 폴리뉴클레오티드를 포함하는 것인, 제약 조성물.

청구항 18

제17항에 있어서, 상기 MVM 인트론 폴리뉴클레오티드는 서열 번호 53의 핵산 서열을 포함하는 것인, 제약 조성물.

청구항 19

제16항에 있어서, 상기 인트론은 인자 IX 폴리펩티드를 암호화하는 뉴클레오티드 서열의 프로모터 요소와 번역 개시 부위 사이에 위치하는 것인, 제약 조성물.

청구항 20

제1항에 있어서, 서열 번호 40의 상기 핵산 서열을 포함하는, 제약 조성물.

청구항 21

제1항 내지 제3항 중 어느 한 항에 있어서, 포유류 유전자 요법 벡터를 포함하는, 제약 조성물.

청구항 22

제21항에 있어서, 상기 포유류 유전자 요법 벡터가 아데노-관련 바이러스 (AAV) 벡터인, 제약 조성물.

청구항 23

제22항에 있어서, 상기 아데노-관련 바이러스 벡터가 혈청형 8 아데노-관련 바이러스 (AAV-8) 벡터인, 제약 조성물.

청구항 24

제21항에 있어서, 상기 포유류 유전자 요법 벡터가 상기 인자 IX 단백질을 암호화하는 단일 가닥 폴리뉴클레오티드를 포함하는 것인, 제약 조성물.

발명의 설명

기 술 분 야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 미국 가출원 제62/509,616호 (2017년 5월 22일에 출원됨)에 대해 이익을 주장하며, 이는 모든 목적을 위해 그 전체가 본원에 명백히 참조로 포함된다.

[0003] 컴팩트 디스크 상에 제출된 "서열 목록", 표 또는 컴퓨터 프로그램, 목록 부록에 대한 참조

[0004] 본 명세서는 2018년 5월 21일에 작성되었고, 크기가 73 킬로바이트인 008073_5117_WO_Sequence_Listing.txt로 명명된, 함께 제출된 서열 목록 텍스트 사본을 참조로 포함한다.

배 경 기 술

[0005] 혈액 응고는 응고 캐스케이드로 지칭되는, 상호 의존적 생화학적 반응의 복잡하고 역동적인 생물학적 경로를 통해 진행된다. 응고 인자 VIII (FVIII) 및 인자 IX (FIX)가 캐스케이드의 핵심 구성 요소이다. 인자 VIII은 출혈 부위로 모집되어, 활성화된 인자 IX 및 인자 X (FX)와 Xase 복합체를 형성한다. Xase 복합체는 FX를 활성화시키고, 그 결과 프로트롬빈이 트롬빈으로 활성화되고, 그런 다음 응고 캐스케이드에서 안정적인 응괴를 생성하기 위해 다른 성분을 활성화시킨다 (Saenko *et al.*, *Trends Cardiovasc.Med.*, 9:185-192 (1999); Lenting *et al.*, *Blood*, 92:3983-3996 (1998)에서 리뷰됨).

[0006] B형 혈우병은 인자 IX 활성의 결핍을 특징으로 하는 선천적 X-연관 출혈 장애이다. 일반적으로, 감소된 인자 VIII/인자 IX 활성은 응고 캐스케이드에서 양성 피드백 루프를 억제한다. 이는 불완전한 응고를 일으켜, 증가된 지속 시간, 넓은 타박상, 자발적인 구강 및 비강 출혈, 관절 강직 및 만성 통증, 심한 경우에는 내부 출혈 및 빈혈이 있는 출혈 에피소드로 나타난다. (Zhang *et al.*, *Clinic.Rev. Allerg.Immunol.*, 37:114-124 (2009)).

[0007] 통상적으로, B형 혈우병은 B형 혈우병을 가진 개체에 대해 인자 IX 단백질 (예를 들어, 혈장 유래 또는 재조합 생성 인자 IX)을 투여하는 것으로 이루어진 인자 IX 대체 요법에 의해 치료된다. 인자 IX는 급성 출혈 에피소드에 대한 반응으로 출혈 에피소드의 빈도를 방지 또는 감소시키기 위해 예방적으로, 및/또는 수술 중 출혈을 관리하기 위해 수술주기적으로 투여된다. 그러나, 인자 IX 대체 요법의 여러 바람직하지 않은 특성이 있다.

[0008] 첫째, 인자 IX 대체 요법은 B형 혈우병을 치료 또는 관리하는 데 사용되지만 근본적인 인자 IX 결핍을 치료하지는 않는다. 이 때문에, B 형 혈우병 환자는 평생 동안 인자 IX 대체 요법이 필요하다. 지속적인 치료는 비싸고, 개인에게 엄격한 순응을 요구하며, 소량의 예방 용량이 누락되는 것으로 중증 B 형 혈우병 환자에게 심각한 결과를 초래할 수 있다.

[0009] 둘째, 통상적인 인자 IX 제품은 생체 내에서 약 24 시간의 비교적 짧은 반감기를 가지기 때문에 예방 인자 IX 대체 요법은 매주 2 내지 3 회 투여를 필요로 한다. 이로 인해 개인은 평생 동안 순응을 유지해야 하는 부담을 준다. 3 세대 "지속성" 인자 IX 약물은 투여 빈도를 감소시킬 수 있지만, 이러한 약물을 사용한 예방 인자 FIX 대체 요법은 여전히 월별, 주별 또는 더 빈번한 투여를 영원히 필요로 한다. 예를 들어, Nonacog 베타 페골 [페길화된 재조합 인자 IX] (Novo Nordisk, 미국 및 EP 규제 승인 보류)을 통한 예방 치료는 여전히 매주 투여를 필요로 한다 (Collins P.W., et al., Blood, 124(26):3880-86 (2014)). 더욱이, 화학적으로 변형된 생물학적 제제 (예를 들어, 페길화된 폴리펩티드)의 장기 효과는 아직 완전히 이해되지 않았다.

[0010] 셋째로, 중증 B형 혈우병 환자의 인자 IX 대체 요법의 최대 5%가 항-인자 IX 억제제 항체를 형성하여 요법을 비효율적으로 만든다 (Osooli and Berntorp, J. Intern.Med., 277(1):1-15 (2015)). 항-인자 VIII 억제 항체를 발달시킨 A 형 혈우병 환자를 치료하기 위해 사용될 수 있는 인자 VIII 우회 요법과는 달리, B형 혈우병의 치료를 위한 인자 IX 우회 요법은 존재하지 않는다.

[0011] 넷째로, 인자 IX 대체 요법은 환자의 체중에 따라, 용량 당 약 $1,000 내지 약 $3,000의 범위로, 비싸다 (Hemophilia Federation of America 온라인 자료). 따라서, 매주 2 회 투약하는 경우, 인자 IX 대체 요법은 연간 최대 $300,000의 비용이 들 수 있다.

[0012] 유전자 요법은 개체에 대해 인자 IX 활성의 1 회 용량을 제공하기 보다, 기능성 인자 IX 활성의 근본적인 저발현 (예를 들어, 미스센스 또는 넌센스 돌연변이로 인함)을 치료할 것이기 때문에 B형 혈우병의 치료에 큰 가능성을 제시한다. 인자 IX 대체 요법과 비교하여, 인자 IX를 제공하는 메커니즘의 차이로 인해, 인자 IX 유전자 요법 벡터의 단일 투여는 개체에게 더 길지 않더라도, 수년 동안 충분한 수준의 인자 IX를 제공할 수 있다. 이

는 치료 비용을 줄이고 지속적인 환자 순응의 필요성을 제거한다.

[0013] B형 혈우병의 인자 IX 유전자 요법 치료에 대한 개념 증명은 제시되었다. 예를 들어, Manno C.S., et al., Nat Med., 12(3):342-47 (2006)을 참조. 그러나, 치료학적 유효량의 인자 IX가 충분한 기간 동안 발현될 수 있는지에 대한 의문이 남아있다. 예를 들어, Giangrande, Semin Thromb Hemost.42(5):513-17 (2016)을 참조.

[0014] 코돈-최적화된 인자 IX를 구축하기 위한 여러 시도가 있었다. 예를 들어, WO 2006/036502는 ApoE HCR-1 인핸서 및 알파-1 안티 트립신 (AAT) 프로모터를 갖는 코돈-최적화된 인자 IX AAV 유전자 요법 벡터를 개시한다. 유사하게, WO 2014/064277 및 WO 2016/146757은 간-특이적 SERPIN 조절 요소의 하나 이상의 카피를 포함하는 코돈-최적화된 인자 VIII 및 인자 IX AAV 유전자 요법 벡터를 개시한다. 마지막으로, WO 2016/210170은 ApoE HCR-1 인핸서 및 알파-1 안티 트립신 (AAT) 프로모터를 갖는 코돈-최적화된 인자 IX AAV 유전자 요법 벡터를 개시한다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0015] 따라서, 개선된 인자 IX 유전자 요법 작제물이 필요하다. 예를 들어, 유전자 요법 벡터에 보다 효율적으로 포장되고, 이를 통해 전달되는 인자 IX를 암호화하는 합성, 코돈-변경된 핵산이 필요하다. 인자 IX를 보다 효율적으로 발현하는 합성, 코돈-변경된 핵산이 필요하다. 야생형 인자 IX와 비교하여 개선된 접힘 특성, 발현 세포로부터의 분비 개선 및/또는 증가된 활성을 가지는 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 핵산이 또한 필요하다. 이러한 인자 IX 암호화, 코돈-변경된 핵산은 인자 IX 결핍 (예를 들어, B형 혈우병)의 개선된 치료를 허용한다. 인자 IX 결핍 (예를 들어, B형 혈우병)의 치료와 관련된 상기 결함 및 다른 문제는 인자 IX 단백질을 암호화하는 개시된 코돈-변경된 핵산에 의해 감소되거나 제거된다.

[0016] 한 측면에서, 인자 IX 및 인자 IX 변이체를 암호화하는 핵산 조성물 (예를 들어, 코돈-변경된 폴리뉴클레오티드)이 기술된다. 핵산 조성물은 본원에 기재된 바와 같이, 인자 IX를 암호화하는 CS02, CS03, CS04, CS05 및 CS06 서열에 대해 높은 서열 동일성을 갖는 폴리뉴클레오티드를 포함한다. 본원에 기재된 핵산 조성물은 야생형 인자 IX 암호화 서열에 비해 증가된 인자 IX 발현을 제공한다. 핵산 조성물은 또한 AAV-기반 유전자 요법 비리온의 생산 증가를 허용한다. 일부 실시 양태에서, 본원에 기재된 핵산 조성물은 인자 IX를 암호화하는 야생형 서열과 비교하여 감소된 GC 함량을 갖거나 더 적은 CpG 디뉴클레오티드를 포함한다.

[0017] 일부 실시 양태에서, 핵산 조성물은 CS02-FL-NA (서열 번호 5), CS02-MP-NA (서열 번호 13), CS03-FL-NA (서열 번호 6), CS03-MP -NA (서열 번호 14), CS04-FL-NA (서열 번호 7), CS04-MP-NA (서열 번호 15), CS05-FL-NA (서열 번호 8), CS05-MP-NA (서열 번호 16), CS06-FL-NA (서열 번호 9) 및 CS06-MP-NA (서열 번호 17)로부터 선택된 개시된 서열에 대해 적어도 95%의 서열 동일성 (예를 들어, 적어도 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성)을 갖는 뉴클레오티드 서열을 가지는 인자 IX를 암호화하는 폴리뉴클레오티드를 포함한다.

[0018] 일부 실시 양태에서, 핵산 조성물은 인자IX 경쇄를 암호화하는 개시된 서열 (예를 들어, CS02-LC-NA (서열 번호 42), CS03-LC-NA (서열 번호 44), CS04-LC-NA (서열 번호46) 에 대한 CS05-LC-NA (서열 번호 48), 또는 CS06-LC-NA (서열 번호 50) 및 인자 IX 중쇄를 암호화하는 개시된 서열 (예를 들어, CS02-HC-NA (서열 번호 41), CS03-HC-NA (서열 번호 43), CS04-HC-NA (서열 번호 45), CS05-HC-NA (서열 번호 47) 또는 CS06-HC-NA (서열 번호 49))에 대해 적어도 95%의 서열 동일성 (예를 들어, 적어도 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성)을 갖는 뉴클레오티드 서열을 가지는 인자 IX를 암호화하는 폴리뉴클레오티드를 포함한다.

[0019] 일부 실시 양태에서, 핵산 조성물은 경쇄, 중쇄 및 경쇄의 C-말단을 중쇄의 N-말단에 연결하는 폴리펩티드 링커 (예를 들어, 활성화 펩티드)를 갖는 인자 IX 폴리펩티드를 암호화하는 폴리뉴클레오티드를 포함한다. 인자 IX 폴리펩티드의 경쇄는 CS02-LC-NA (서열 번호 42), CS03-LC-NA (서열 번호 44), CS04-LC-NA (서열 번호 46), CS05-LC-NA (서열 번호 48) 및 CS06-LC-NA (서열 번호 50) 중 하나에 대해 높은 서열 동일성을 가지는 제1 뉴클레오티드 서열에 의해 암호화된다. 인자 IX 폴리펩티드의 중쇄는 CS02-HC-NA (서열 번호41), CS03-HC-NA (서열 번호43), CS04-HC-NA (서열 번호 45), CS05-HC-NA (서열 번호 47) 및 CS06-HC-NA (서열 번호 49) 중 하나에 대해 높은 서열 동일성을 가지는 제 2 뉴클레오티드 서열에 의해 암호화된다. 폴리펩티드 링커는 단백질분해효

소 절단 부위 (예를 들어, 2 개의 인자 XIa 절단 부위)를 포함한다.

[0020] 상기 기술된 폴리뉴클레오티드의 일부 실시 양태에서, 폴리펩티드 링커는 야생형 인자 IX 활성화 펩티드 FIX-AP-AA (서열 번호 56; FIX-FL-AA (서열 번호 2)의 아미노산 192-226)에 대해 높은 서열 동일성을 가지는 아미노산 서열을 가진다. 일부 실시 양태에서, 폴리펩티드 링커는 CS02-AP-NA (서열 번호57), CS03-AP-NA (서열 번호 58), CS04- AP-NA (서열 번호 59), CS05-AP-NA (서열 번호 60) 및 CS06-AP-NA (서열 번호 61) 중 하나에 대해 높은 서열 동일성을 가지는 제 3 핵산 서열에 의해 암호화된다.

[0021] 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 예를 들어, 코딩된 인자 IX 단백질이 신호 펩티드 및 프로-펩티드를 포함하는 프리-프로-인자 IX 폴리펩티드를 암호화한다. 일부 실시 양태에서, 신호 펩티드, 프로-펩티드, 또는 신호 펩티드 및 프로-펩티드 둘 다 코돈-변경된 서열에 의해 암호화된다. 일부 실시 양태에서, 신호 펩티드, 프로-펩티드, 또는 신호 펩티드 및 프로-펩티드 둘 모두는 야생형 서열에 의해 암호화되는 반면, 성숙한 인자 IX 단일 사슬 폴리펩티드 (예를 들어, FIX-MP-AA (서열 번호 10); FIX-FL-AA (서열 번호 2)의 아미노산 47-461)를 암호화하는 핵산의 일부는 코돈 변경된다.

[0022] 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 예를 들어 야생형 인자 IX 아미노산 서열 (예를 들어, FIX-FL-AA (서열 번호2) 또는 FIX-MP-AA (서열 번호10))에 대해 하나 이상의 아미노산 치환을 갖는, 인자 IX 변이체 폴리펩티드를 암호화한다. 일부 실시 양태에서, 인자 IX 변이체는 야생형 인자 IX와 비교하여, 증가된 인자 IX 활성을 갖는 과활성 인자 IX 변이체이다. 특정 실시 양태에서, 암호화된 인자 IX 폴리펩티드는 인자 IX 프리 프로 폴리펩티드 서열 FIX-FL-AA (서열 번호 2)에 비해 '파두아' R384L 아미노산 치환, 성숙 인자 IX 단일 사슬 서열 FIX-MP-AA (서열 번호 10)에 비해 R338L 아미노산 치환을 가진다.

[0023] 한 측면에서, B형 혈우병을 치료하는 방법이 설명된다. 상기 방법은 본원에 기술된 (예를 들어, CS02, CS03, CS04, CS05 또는 CS06 인자 IX 코딩 서열에 대해 서열 동일성이 높은 폴리뉴클레오티드) 핵산 조성물 (예를 들어, 코돈-변경된 인자 IX 폴리뉴클레오티드 작제물)을 이를 필요로 하는 환자에게 투여하는 것을 포함한다. 일부 실시 양태에서, 인자 IX 폴리뉴클레오티드 작제물은 본원에 기재된 바와 같은 포유류 유전자 요법 벡터이다. 특정 실시 양태에서, 인자 IX 폴리뉴클레오티드 작제물은 아데노 관련 바이러스 (AAV) 벡터이다. 일부 실시 양태에서, 유전자 요법 벡터는 간 특이적 조절 제어 요소의 하나 이상의 카피 (예를 들어, CRM8 조절 제어 요소의 1 내지 3개의 카피)를 포함한다.

[0024] 한 측면에서, 아데노 관련 바이러스 (AAV) 입자를 생성하는 방법이 기술된다. 상기 방법은 본원에 기재된 (예를 들어, CS02, CS03, CS04, CS05, 또는 CS06 인자 IX 코딩 서열과 높은 서열 동일성을 갖는 폴리뉴클레오티드) 핵산 조성물 (예를 들어, 코돈-변경된 인자 IX 폴리뉴클레오티드 작제물)을 포유류 숙주 세포에 도입하는 것을 포함하며, 여기서 폴리뉴클레오티드는 포유류 숙주 세포에서의 복제에 적격이다.

도면의 간단한 설명

[0025] 도 1은 일부 구현에 따른 예시적인 인자 IX 유전자 요법 작제물을 예시한다. 자기 상보적 (A, B) 및 단일 가닥 (C, D) 벡터에 대한 서열 요소는 간 특이적 시스-조절 모듈 (CRM8) 없이 (A, C) 및 모듈과 함께 (B, D) 도시되어 있다.

도 2는 수탁 번호 CCDS14666.1 ("FIX-FL-NA")에 대한 야생형 인자 IX 코딩 서열 (서열 번호 1)을 나타낸다.

도 3은 인간에서 발현된 2개의 야생형 인자 IX 프리-프로-폴리펩티드 동형의 아미노산 서열을 나타낸다. 도 3a는 UniProt 수탁 번호 P00740 및 NCBI 수탁 번호 NP_000124.1 ("FIX-FL-AA")에 상응하는 제1의 더 긴 인자 IX 프리-프로-폴리펩티드 동형 (서열 번호 2)에 대한 야생형 아미노산 서열을 나타낸다. 도 3b는 NCBI 수탁 번호 NP_001300842.1 ("FIX2-FL-AA")에 상응하는 더 짧은 제 2 인자 IX 프리-프로-폴리펩티드 동형 (서열 번호 3)에 대한 야생형 아미노산 서열을 나타낸다.

도 4는 파두아 (R384L) 인자 IX 아미노산 서열 (서열 번호 4; "FIXp-FL-AA")을 나타낸다.

도 5는 일부 구현에 따라, R384L 아미노산 치환 (CS02-FL-NA)을 갖는 인자 IX 변이체를 암호화하는 CS02 코돈-변경된 뉴클레오티드 서열 (서열 번호 5)을 나타낸다.

도 6은 일부 구현에 따라, R384L 아미노산 치환 (CS03-FL-NA)을 갖는 인자 IX 변이체를 암호화하는 CS03 코돈-변경된 뉴클레오티드 서열 (서열 번호 6)을 나타낸다.

도 7은 일부 구현에 따라, R384L 아미노산 치환 (CS04-FL-NA)을 갖는 인자 IX 변이체를 암호화하는 CS04 코돈-

변경된 뉴클레오티드 서열 (서열 번호 7)을 나타낸다.

도 8은 일부 구현에 따라, R384L 아미노산 치환 (CS05-FL-NA)을 갖는 인자 IX 변이체를 암호화하는 CS05 코돈-변경된 뉴클레오티드 서열 (서열 번호 8)을 나타낸다.

도 9는 일부 구현에 따라, R384L 아미노산 치환 (CS06-FL-NA)을 갖는 인자 IX 변이체를 암호화하는 CS06 코돈-변경된 뉴클레오티드 서열 (서열 번호 9)을 나타낸다.

도 10은 2x10E11 vg/체중 kg의 용량으로, CRM8 간-특이적 시스-조절 제어 요소의 0, 1, 2 또는 3 개의 카피를 갖는 CS02 유전자 요법 작제물이 주사된 야생형 마우스에서의 FIX 항원 수준을 도시한다.

도 11은 인간에서 발현된 2개의 단일 사슬, 야생형 인자 IX 단백질 동형 (예를 들어, 신호 및 프로펩티드가 없음)의 아미노산 서열을 나타낸다. 도 11a는 UniProt 수탁 번호 P00740 및 NCBI 수탁 번호 NP_000124.1 ("FIX-MA-AA")에 상응하는 제1의 더 긴 인자 IX 프리-프로-폴리펩티드 동형 (서열 번호 10)에 대한 야생형 아미노산 서열을 나타낸다. 도 11b는 NCBI 수탁 번호 NP_001300842.1 ("FIX2-MA-AA")에 상응하는 더 짧은 제 2 인자 IX 프리-프로-폴리펩티드 동형 (서열 번호 11)에 대한 야생형 아미노산 서열을 나타낸다.

도 12는 단일 사슬 인자 IX (R338L) "파두아" 아미노산 서열 (서열 번호 12; "FIXp-MP-AA")을 나타낸다.

도 13은 일부 구현에 따라, R338L 아미노산 치환 (CS02-MP-NA)을 갖는 단일 사슬 인자 IX 변이체를 암호화하는 CS02 코돈-변경된 뉴클레오티드 서열 (서열 번호 13)을 나타낸다.

도 14는 일부 구현에 따라, R338L 아미노산 치환 (CS03-MP-NA)을 갖는 단일 사슬 인자 IX 변이체를 암호화하는 CS03 코돈-변경된 뉴클레오티드 서열 (서열 번호 14)을 나타낸다.

도 15는 일부 구현에 따라, R338L 아미노산 치환 (CS04-MP-NA)을 갖는 단일 사슬 인자 IX 변이체를 암호화하는 CS04 코돈-변경된 뉴클레오티드 서열 (서열 번호 15)을 나타낸다.

도 16은 일부 구현에 따라, R338L 아미노산 치환 (CS05-MP-NA)을 갖는 단일 사슬 인자 IX 변이체를 암호화하는 CS05 코돈-변경된 뉴클레오티드 서열 (서열 번호 16)을 나타낸다.

도 17은 일부 구현에 따라, R338L 아미노산 치환 (CS06-MP-NA)을 갖는 단일 사슬 인자 IX 변이체를 암호화하는 CS06 코돈-변경된 뉴클레오티드 서열 (서열 번호 17)을 나타낸다.

도 18은 일부 구현에 따라, 본원에 기재된 다수의 작제물의 프리-프로-펩티드 (PPP)를 암호화하는 핵산 서열 (NA)을 나타낸다.

도 19는 일부 구현에 따라, 본원에 기재된 다수의 작제물에 대한 신호 펩티드 (SP)를 암호화하는 핵산 서열 (NA)을 나타낸다.

도 20은 일부 구현에 따라, 본원에 기재된 다수의 작제물에 대해 프로-펩티드 (PP)를 암호화하는 핵산 서열 (NA)을 나타낸다.

도 21은 FIX 프리-프로-펩티드 (PPP)의 아미노산 서열 (AA)을 나타낸다.

도 22는 FIX 신호 펩티드 (SP)의 아미노산 서열 (AA)을 나타낸다.

도 23은 FIX 프로-펩티드 (PP)의 아미노산 서열 (AA)을 나타낸다.

도 24는 CRM8 서열 (서열 번호 39)의 핵산 서열을 나타낸다.

도 25a 및 b는 CS06-CRM8.3-ssV 작제물 (서열 번호 40)의 핵산 서열을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0026] I. 서론

[0027] AAV 기반 유전자 요법은 혈우병 치료에 큰 가능성을 가지고 있다. B형 혈우병의 경우, 적어도 일부 환자에서 약 10%의 FIX 수준이 1 년 이상 동안 유지될 수 있다는 점에서 최초의 임상 데이터가 고무적이다. 예를 들어, 최초의 인간 임상 시험에서, AVV-FIX 작제물의 간동맥 도관 삽입이 생체 내에서 인자 IX의 일시적인 발현을 초래함을 입증하였다. Kay M. et al., Nat Genet.24(3):257-61 (2000). 그러나, 형질 도입은 AAV-유래 캡시드 항원에 대한 면역계의 적당한 활성화를 초래하였다. Manno C.S. et al., Nat Med. 12(3):342-47 (2006) 및 Mingozzi

F. et al., Nat Med. 13(4):419-22 (2007).

[0028] 비-바이러스성 벡터는 비-항원성 담체 (예를 들어, 불활성 중합체, 지질 또는 나노 입자)와 관련된 네이키드 DNA 또는 DNA의 전달에 기초하기 때문에 면역원성이 적을 수 있다. 그러나, 비-바이러스성 벡터의 세포 형질 감염 속도는 바이러스 전달 벡터에 대한 것보다 낮다. 또한, 비-바이러스성 벡터로부터의 장기 발현은 작제물의 대규모 생산에 사용되는 박테리아 서열의 존재에 의해 방해받게 된다.

[0029] 그러나, 이러한 도전은 단순히 고용량의 유전자 요법 작제물을 투여함으로써 해결될 수 없다. 현재의 지식에 따르면, AAV-기반 유전자 요법 벡터의 벡터 용량은 $2x10^{12}$ vg/체중 kg 초과로 증가해서는 안된다. 이는 그러한 고용량에서 T 세포 면역 반응이 유발되어, 형질도입된 세포를 파괴하고 결과적으로 이식유전자 발현이 감소되거나 심지어 제거되기 때문이다. 따라서, FIX 유전자 요법을 B형 혈우병 환자에게 실행 가능한 치료 옵션으로 만들려면 FIX의 발현을 개선하기 위한 전략이 필요하다.

[0030] 따라서, 개선된 인자 IX 발현 및 활성을 지지하는 개선된 인자 IX 폴리펩티드 작제물은 두 치료적 접근법 모두를 향상시킬 것이다. 예를 들어, 바이러스 전달 방법은 작제물의 초기 용량을 감소시킴으로써 대상체의 면역계의 자극을 감소시킴으로써 개선될 것이다. 네이키드 DNA의 투여에 의존하는 방법은 치료용 폴리뉴클레오티드의 더 적은 카피로 더 큰 인자 IX 활성을 지원함으로써 개선될 것이다.

[0031] 본 개시는 인자 IX 유전자 요법과 관련된 이들 및 다른 문제를 해결하는 코돈-변경된 인자 IX 변이체 코딩 서열의 발견에 관한 것이다. 예를 들어, 본원에 개시된 폴리뉴클레오티드는 포유류 숙주에서 현저하게 개선된 인자 IX 발현 및 활성을 제공한다. 일부 구현에서, 이들 이점은 코돈-변경된 CS02, CS03, CS04, CS05 및 CS06 작제물에 대해 높은 서열 동일성을 갖는 인자 IX-암호화 폴리뉴클레오티드를 사용함으로써 실현된다. 일부 실시 양태에서, 이들 서열은 하기에 보다 상세히 기술되는 바와 같이 야생형 작제물과 비교하여 상당히 적은 CpG 디뉴클레오티드를 포함한다.

[0032] 유리하게는, 본원에 기재된 CS02, CS03, CS04, CS05 및 CS06 코돈-변경된 인자 IX 서열은 동등한 야생형 서열과 비교하여 생체 내에서 우수한 인자 IX 발현을 제공한다. 예를 들어, 실시예 1은 CS02, CS03, CS04, CS05 또는 CS06 코돈-변경된 인자 IX (R384L) 코딩 서열을 보유하는 자기 상보적 AAV 벡터가 야생형 인자 IX 코딩 서열을 보유하는 자기 상보적 AAV 벡터에 비해, 생체 내 인자 IX 활성의 20 배 내지 40 배 증가를 제공함을 보여준다. 유사하게, 야생형 인자 IX (R384L) 코딩 서열을 보유하는 자기 상보적 AAV 벡터에 비해 인자 IX 발현에서 2 배 내지 4 배 증가가 관찰된다 (표 2).

[0033] 유리하게는, CS02, CS03, CS04, CS05 및 CS06 코돈 변경된 서열로부터 생성된 개선된 인자 IX 활성은 인자 IX 코딩 서열의 업스트림에 간-특이적 조절 요소의 하나 이상의 카피를 도입함으로써 추가로 향상될 수 있다. 예를 들어, 실시예 2 및 3에서 입증된 바와 같이, 자기 상보적 AAV 인자 IX 벡터에 하나 이상의 간-특이적 CRM8 조절 제어 요소의 포함은 마우스 모델에서 인자 IX 발현을 2배 내지 3 배 및 인간 간세포에서 2배 내지 13배 추가로 증가시킨다 (각각 표 3 및 4). 마찬가지로, 단일-가닥 AAV 인자 IX 벡터에 간-특이적 CRM8 조절 제어 요소의 하나 이상의 카피를 포함시키는 것은 생체 내에서 2 배의 인자 IX 활성을 증가시키고 (마우스 모델; 표 5) 인간 간세포에서 최대 26 배 증가시켰다 (표 6).

[0034] 놀랍게도, 간-특이적 CRM8 조절 제어 요소가 결여된 코돈-변경된 인자 IX 폴리펩티드를 암호화하는 자기 상보적인 AAV 벡터가 유사한 단일 가닥 AAV 벡터보다 인자 IX 발현에서 더 큰 증가를 제공한 반면 (CS06-CRM.0-scV에 의해 제공되는 FIX 활성의 6.2배 증가를 표 6의 CS06-CRM.0-ssV (서열 번호40)에 의해 제공되는 인자 IX 활성 3.9 배 증가와 비교함), 간-특이적 CRM8 조절 제어 요소의 다중-카피를 함유하는 단일 가닥 AAV 인자 IX 벡터는 유사한 자기 상보적 AAV 벡터를 상당히 능가하였다 (표 4의 CS02-CRM8.0-scV에 비해 CS02-CRM8.3-scV에 의해 제공되는 인자 IX 활성의 12.8 배 증가를 표 6의 CS06-CRM8.0-scV에 비해 CS06-CRM8.3-ssV (서열 번호 40)에 의해 제공되는 인자 IX 활성의 16.8 배 증가와 비교함).

[0035] II. 정의

[0036] 본원에 사용된 바와 같이, 다음의 용어는 달리 명시되지 않는 한 그에 부여된 의미를 갖는다.

[0037] 본원에 사용된 바와 같이, 용어 "인자 IX" 및 "FIX" ("IX"는 "9"를 의미하는 로마 숫자를 지칭함)는 상호교환적으로 사용되며, 인자 IX 활성을 갖는 임의의 단백질 (예를 들어, 활성 FIX, 종종 "FIXa"로 지칭됨) 또는 IX 인자 활성, 특히 예를 들어, 유럽 약전 9.0 (그 내용이 본원에 참조로 포함됨)의 제 2.7.11장에 기재된 바와 같은 1 단계 인자 IX 응고 검정을 사용하여 측정된 바와 같은, 인자 VIII의 존재 하의 인자 X 절단 활성을 갖는 단백

질의 단백질 전구체 (예를 들어, 프로-단백질 또는 프리-프로-단백질, 종종 pFIX 및 ppFIX로 지칭됨)를 지칭한다.

[0038] 인자 IX는 신호 펩티드, 프로펩티드, 경쇄, 활성화 펩티드 및 중쇄 (종종 IX 프리-프로-폴리펩티드라고 지칭됨)를 포함하는 비활성, 단일 사슬 폴리펩티드로 번역된다. 인자 IX 프리-프로-펩티드는 번역 후 공정을 거쳐 활성 인자 IX 단백질 (예를 들어, FIXa)을 형성한다. 이 공정은 아직 비활성 상태인, 인자 IX 중쇄 및 인자 IX 경쇄를 함유하는 단일 사슬 성숙 인자 IX 폴리펩티드를 형성하기 위해, 신호 펩티드의 제거 (예를 들어, 절단에 의함), 이어서 프로펩티드의 제거 (예를 들어, 절단에 의함)를 포함한다. 성숙한 인자 IX 폴리펩티드는 인자 IX 경쇄와 인자 IX 중쇄 사이에서 활성화 펩티드를 삭제하기 위해 추가로 절단되어, 활성 인자 IX 단백질 (예를 들어, FIXa)을 형성한다. 인자 IX 경쇄 및 인자 IX 경쇄는 이황화 결합을 통해 연결된 상태로 유지된다.

[0039] 예를 들어, 야생형 인간 인자 IX 프리-프로-단백질은 먼저 암호화된 신호 펩티드 (FIX-FL-AA(서열 번호 2)의 아미노산 1-28)를 방출하도록 절단되어, 첫 번째 단일 사슬 프로-단백질을 형성한다. 그런 다음, 이 단일 사슬 프로-펩티드는 절단되어 프로펩티드 (FIX-FL-AA (서열 번호 2)의 아미노산 29-46)를 방출하여 제2 단일 사슬 프로단백질 (예를 들어, FIX-MP-AA (서열 번호 10), "MP" 명칭은 "성숙한 단백질"을 의미함)을 형성한다. 이어서, 제2 단일 사슬 프로-단백질은 인자 XIa에 의해 FIX 경쇄와 FIX 중쇄 사이에서 2회 절단되어 활성화 펩티드 (FIX-FL-AA (서열 번호2)의 아미노산 192-226)를 방출한다. 이는 이황화 결합을 통해 연결된 개별 경쇄 및 중쇄로 이루어진 활성 인자 IXa 단백질을 형성한다. 인자 IX의 구조, 기능 및 활성화에 대한 추가 정보는 예를 들어 Brandstetter H. et al. P.N.A.S.USA, 92(21):9796-800 (1995), Hopfner KP et al., Structure, 7(8):989-96 (1999), 및 Gailani D. et al., Thromb Res., 133 Suppl 1:S48-51 (2014)을 참조하며, 이 내용은 모든 목적을 위해 그 전체가 본원에 참조로 포함된다.

[0040] 본원에 기술된 바와 같이, 이러한 활성 인자 IXa 단백질은 하나 이상의 변이체를 포함할 수 있으며, 일부 실시양태에서 R338L 변이체가 특히 사용된다. 이는 본원에서 "FIXp-MP-NA"로 지칭되는 핵산 서열을 가지는 "FIXp-MP-AA"(서열 번호 12)로 지칭되고; "FIXp"는 최종 단백질에 파두아 R338L 변이체의 포함을 의미한다. 본원에 예시된 코돈-최적화된 서열 CS02-CS06은 R338L 변이체를 포함하는, FIXp 단백질을 암호화한다는 점을 유의해야 한다. 따라서 FIX의 정의에 구체적으로 포함된 것은 FIXp이다.

[0041] 본원에 사용된 용어 "인자 IX 폴리펩티드" 및 "FIX 폴리펩티드"는 예를 들어, 유럽 약전 9.0의 제 2.7.11 장에 기재된 1 단계 인자 IX 응고 검정을 사용하여 측정된 바와 같은, 특정 조건 하에서 인자 IX 세린 단백질분해효소 활성을 갖는 폴리펩티드를 지칭한다. 인자 IX 폴리펩티드는 상기 기재된 번역 후 가공에 의해 활성화될 때, 인자 IX 세린 단백질분해효소 활성뿐만 아니라 활성 인자 IX 단백질 자체를 갖는 인자 IX 단백질이 되는 단일-사슬 전구체 폴리펩티드 (인자 IX 프리-프로-폴리펩티드, 인자 IX 프로-펩티드 및 성숙된 단일-사슬 인자 IX 폴리펩티드를 포함)를 포함한다. 인자 IX 폴리펩티드의 정의에는 구체적으로 R338L 변이체를 포함하는 인자 IX 폴리펩티드가 포함된다. 예시적인 실시 양태에서, 인간 인자 IX 폴리펩티드는 경쇄 및 중쇄, FIX-MP-AA (서열 번호 10, 도 11a에 나타냄)를 포함하는 야생형 인간 인자 IX 폴리펩티드의 부분 또는 경쇄 및 중쇄, FIXp-MP-AA (서열 번호 12, 도12에 도시됨)를 포함하는 파두아 인간 인자 IX 폴리펩티드의 일부에 대해 높은 서열 동일성 (예를 들어, 적어도 85%, 90%, 95%, 99% 또는 그 이상)을 갖는 아미노산 서열을 포함하는 폴리펩티드를 지칭한다.

[0042] 본원에 사용된 바와 같이, 용어 "인자 IX 경쇄" 또는 간단히 "경쇄"는 인자 IX의 Gla 모듈, EGF-유사 1 및 EGF-유사 2 도메인을 함유하는 인자 IX 단일 사슬 폴리펩티드의 N-말단 부분으로부터 유래된 활성화된 인자 IXa 단백질의 폴리펩티드를 지칭한다. 예시적인 실시 양태에서, 인간 프리-프로-인자 IX 폴리펩티드 (FIX-FL-AA (서열 번호 2))의 아미노산 47-191은 인자 IX 경쇄를 구성한다. 본원에 사용된 바와 같이, 야생형 인자 IX 경쇄의 아미노산 서열은 FIX-LC-AA (서열 번호 62)로 지칭된다.

[0043] 본원에 사용된 바와 같이, 용어 "인자 IX 중쇄" 또는 간단히 "중쇄"는 인자 IX의 펩티드분해효소 S1 도메인을 함유하는 인자 IX 단일 사슬 폴리펩티드의 C-말단 부분으로부터 유래된 활성화된 인자 IXa 단백질의 폴리펩티드를 지칭한다. 예시적인 실시 양태에서, 인간 프리-프로-인자 IX 폴리펩티드 (FIX-FL-AA (서열 번호 2))의 아미노산 227-461은 인자 IX 중쇄를 구성한다. 본원에 사용된 바와 같이, 야생형 인자 IX 중쇄의 아미노산 서열은 R338L 변이체가 포함될 때 FIX-HC-AA (서열 번호 63) 및 FIXp-HC-AA (서열 번호 64)로 지칭된다.

[0044] 일반적으로, 인자 IX 경쇄 및 중쇄는 예를 들어 활성화 펩티드와 함께, 단일 폴리펩티드 사슬로 표현된다. 그러나, 일부 실시 양태에서, 인자 IX 경쇄 및 인자 VIII 중쇄는 별개의 폴리펩티드 사슬 (예를 들어, 공동-발현)로서 발현되고, 재구성되어 인자 IX 단백질 (예를 들어, 생체 내 또는 시험관 내)을 형성한다. 일반적으로, 본 발

명의 목적상, 2 개의 사슬이 개별적으로 발현되더라도, 이들은 상이한 발현 벡터가 아닌 동일한 발현 벡터 (예를 들어, 바이러스 게놈) 상에 존재한다.

[0045] 본원에 사용된 바와 같이, 용어 "인자 IX 활성화 펩티드" 또는 간단히 "활성화 펩티드"는 인자 IXa 단백질의 활성화시 인자 IX 단일-사슬 폴리펩티드로부터 절제된 펩티드를 지칭한다. 예시적인 실시 양태에서, 인간 프리-프로-인자 IX 폴리펩티드 (FIX-FL-AA (서열 번호 2))의 아미노산 192-226은 인자 IX 활성화 펩티드를 구성한다. 본원에 사용된 바와 같이, 야생형 인자 IX 활성화 펩티드의 아미노산 서열은 FIX-AP-AA (서열 번호 56)로 지칭된다.

[0046] 본원에 사용된 용어 "인자 IX 신호 펩티드" 또는 간단히 "신호 펩티드"는 신호 펩티드분해효소에 의해 인자 IX 프리-프로-폴리펩티드의 N-말단으로부터 절제된 펩티드를 지칭한다. 신호 펩티드는 새로 번역된 인자 IX 프리-프로-단백질을 소포체로 보낸다. 예시적인 실시 양태에서, 인간 프리-프로-인자 IX 폴리펩티드 (FIX-FL-AA (서열 번호 2))의 아미노산 1-28은 인자 IX 신호 펩티드를 구성한다. 본원에 사용된 바와 같이, 야생형 인자 IX 신호 펩티드의 아미노산 서열은 FIX-SP-AA (서열 번호 37)로 지칭된다. 본 발명의 다수의 신호 펩티드를 도 19 내지 22에 나타내고 있다.

[0047] 본원에 사용된 바와 같이, 용어 "인자 IX 프로-펩티드" 또는 간단히 "프로-펩티드"는 푸린에 의해 인자 IX 프로-폴리펩티드의 N-말단으로부터 절제된 (예를 들어, 신호 펩티드의 절단 후) 펩티드를 지칭한다. 프로-펩티드는 카르복실라제를 인접한 Gla 모듈로 모집하여 글루타민 잔기의 카르복실화를 촉진하는 γ- 카르복실화 인식 부위를 포함한다예시적인 실시 양태에서, 인간 프리-프로-인자 IX 폴리펩티드 (FIX-FL-AA (서열 번호 2))의 아미노산 29-46은 인자 IX 프로펩티드를 구성한다. 본원에 사용된 바와 같이, 야생형 인자 IX 프로-펩티드의 아미노산 서열은 FIX-PP-AA (서열 번호 38)로 지칭된다.

[0048] 본원에 사용된 바와 같이, 용어 "인자 IX 프리-프로-펩티드" 또는 간단히 "프리-프로-펩티드"는 인자 IX 신호 펩티드 및 프로 폴리펩티드의 응집체를 지칭한다. 예시적인 실시 양태에서, 인간 프리 프로 인자 IX 폴리펩티드 (FIX-FL-AA (서열 번호 2))의 아미노산 1-46은 인자 IX 프리-프로-펩티드를 구성한다. 본원에 사용된 바와 같이, 야생형 인자 IX 프리 프로펩티드의 아미노산 서열은 도 18에 나타낸 핵산 서열과 함께 FIX-PPP-AA (서열 번호 36)로 지칭되며, (R338L 변이체가 사용될 때 상응하는 FIXp-PPP-AA 및 FIXp-PPP-NA와 함께) FIX-PPP-NA (서열 번호 18)로 지칭된다.

[0049] 본원에서 달리 명시되지 않는 한, 인자 IX 아미노산의 넘버링은 전장, 야생형 인간 인자 IX 프리 프로 폴리펩티드 서열 (FIX-FL-AA)에서 상응하는 아미노산을 지칭하며, 도 3a의 서열 번호 2로 제시된다. 이와 같이, 본원에 개시된 인자 IX 폴리펩티드에서 아미노산 치환을 지칭할 때, 언급된 아미노산 수는 전장, 야생형 인자 IX 프리-프로-폴리펩티드 서열의 유사한 (예를 들어, 구조적으로 또는 기능적으로 동등 함) 및/또는 상동인 (예를 들어, 1차 아미노산 서열에 진화적으로 보존됨) 아미노산을 지칭한다. 예를 들어, R384L 아미노산 치환은 전장, 야생형 인간 인자 IX 프리-프로-펩티드 서열(FIX-FL-AA (서열 번호 2)의 384 위치에서 R에서 L 로의 치환, 성숙된 야생형 인자 IX 단일 사슬 폴리펩티드 (FIX-MP-AA (서열 번호 10))의 338 위치에서 R에서 L 로의 치환, 전장, 야생형 인간 인자 IX 프리-프로-펩티드 동형 2 서열 (FIX2-FL-AA (서열 번호 3)의 346 위치에서 R에서 L 로의 치환, 성숙된 야생형 인간 인자 IX 프리-프로-펩티드 동형 2 서열(FIX2-FL-AA (서열 번호3)의 300 위치에서 R에서 L 로의 치환, 및 야생형 인간 인자 IX 중쇄 서열 (FIX-HC-AA (서열 번호 63)의 158 위치에서 R에서 L로의 치환을 지칭한다. 따라서, 이들 명명법 모두는 상이한 인자 IX 작제물에서 동일한 "파두아" 아미노산 치환을 기술한다.

[0050] 본원에 기재된 바와 같이, 인자 IX 아미노산 넘버링 시스템은 인자 IX 프리-프로-펩티드 (예를 들어, 신호 펩티드 및 프로-펩티드를 포함하는 전장 야생형 인간 인자 IX 서열의 아미노산 1-46)에 의존하는지 여부가 포함된다. 프리-프로-펩티드가 포함되는 경우, 넘버링은 "프리-프로-펩티드 포함" 또는 "PPI"로 지칭된다. 프리-프로-펩티드가 포함되지 않은 경우, 넘버링은 "프리-프로-펩티드 배제" 또는 "PPE"로 지칭된다. 예를 들어, R384L은 PPE 넘버링에서 R338L과 동일한 아미노산 치환에 대한 PPI 넘버링이다. 유사하게, 인자 IX 아미노산 넘버링은 또한 인자 IX 동형에 의존적이다. 예를 들어, R384L은 동형 2 넘버링에서 R346L과 동일한 아미노산 치환에 대한 동형 1 넘버링이다. 달리 지시되지 않는 한, 모든 아미노산 넘버링은 전장, 야생형 인간 인자 IX 동형 1 서열 (FIX-FL-AA)에서 상응하는 아미노산을 나타내며, 도 3a의 서열 번호 2로 제시된다. 이 넘버링은 R384L "파두아" 돌연변이를 제외하고 동일한 아미노산 서열을 갖는 FIXp-FL-AA (서열 번호 4)에 대해 동일하다.

[0051] 야생형 인자 IX 폴리펩티드의 비제한적 예시는 인간 프리-프로-인자 IX (예를 들어, GenBank 수탁 번호 NP_000124.1 (FIX-FL-AA (서열 번호 2)) 및 NP_001300842.1 (FIX2-FL-AA (서열 번호3), 상응하는 신호 펩티드

(프리-프로-단백질의 아미노산 1-28) 및/또는 프로펩티드(프리-프로-단백질의 아미노산 29-46)가 없는 단일 사슬 인자 IX, 및 이의 천연 변이체; 돼지 프리-프로-인자 IX (예를 들어, UniProt 수탁 번호 P00741), 신호 펩티드가 없는 상응하는 단일 사슬 인자 IX 및 그의 천연 변이체; 마우스 프리-프로-인자 IX (예를 들어, UniProt 수탁 번호 P16294), 신호 펩티드가 없는 상응하는 단일 사슬 인자 IX 및 그의 천연 변이체; 랫트 프리-프로-인자 IX (예를 들어, UniProt 수탁 번호 P16296), 신호 펩티드가 없는 상응하는 단일 사슬 인자 IX 및 그의 천연 변이체; 및 다른 포유류 인자 VIII 동족체 (예를 들어, 침팬지, 유인원, 햄스터, 기니피그 등)를 포함한다.

[0052] 본원에 사용된 바와 같이, 인자 IX 폴리펩티드는 인자 VIII의 존재 하에서 인자 X 절단 활성을 갖는 천연 변이체 및 인공 작제물을 포함한다. 본 개시에 사용된 바와 같이, 인자 IX는 일부 기본 인자 IX 절단 활성 (예를 들어, 유럽 약전 9.0 (본원에서 제 2.7.11장의 인간 응고 인자 IX 검정에 대한 교시에 대한 참고로 포함됨)의 제 2.7.11 장에 따른 1 단계 응고 검정에서 분석된 바와 같은 상응하는 야생형 활성의 적어도 5%, 10%, 25%, 50%, 75%, 또는 그 이상의 활성)을 유지하는 임의의 천연 변이체, 대안 서열, 동형 또는 돌연변이 단백질을 포함한다. 인간 집단에서 발견되는 인자 IX 아미노산 변이 (FIX-FL-AA (서열 번호2)에 대한)의 예는 I17N, L20S, C28R, C28Y, V30I, R43L, R43Q, R43W, K45N, R46S, R46T, N48I, S49P, L52S, E53A, E54D, E54G, F55C, G58A, G58E, G58R, E66V, E67K, F71S, E73K, E73V, R75Q, E79D, T84R, Y91C, D93G, Q96P, C97S, P101R, C102R, C102R, G106D, G106S, C108S, D110N, I112S, N113K, Y115C, C119F, C119R, E124K, G125E, G125R, G125V, C134Y, I136T, N138H, G139D, G139S, C155F, G160E, Q167H, S169C, C170F, C178R, C178W, R191C, R191H, R226G, R226Q, R226W, V227D, V227F, V228F, V228L, Q241H, Q241K, C252S, C252Y, G253E, G253R, A265T, C268W, A279T, N283D, E291V, R294G, R294Q, V296M, H302R, N306S, I316F, L318R, L321Q, N328K, N328Y, P333H, P333T, T342K, T342M, I344L, G351D, W356C, G357E, G357R, K362E, G363W, A366D, R379G, R379Q, C382Y, L392F, L383I, R384L, K387E, I390F, M394K, F395I, F395L, C396F, C396S, A397P, R404T, C407R, C407S, D410H, S411G, S411I, G412E, G413R, P414T, V419E, F424V, T426P, S430T, W431G, W431R, G432S, E433A, G433K, C435Y, A436V, G442E, G442R, I443T, R449Q, R449W, Y450C, W453R및 I454T를 포함하지만, 이에 제한되지는 않는다. 하기에 더 상세히 논의된 바와 같이, 이 넘버링은 야생형 인간 FIX에 대한 것이다. 인간 집단에서 확인된 다른 아미노산 변이는 공지되어 있으며, 예를 들어, 국립 생물 공학 정보 센터 (National Center for Biotechnology Information, "NCBI")의 변이 뷰어를 사용하여, 수탁 번호 GCF_000001405.25를 찾을 수 있다. 인자 VIII 단백질은 또한 번역 후 변형을 함유하는 폴리펩티드를 포함한다.

[0053] 본 개시에서 특히 사용되는 것은 이른바 "파두아" 돌연변이, 성숙 단일 가닥 인자 IX 단백질 (R338L)의 위치 338, 인자 IX 프로-폴리펩티드 (R384L)의 위치 384에서 아르기닌에서 류신으로의 변화를 포함하는 FIX 단백질이다. 이 돌연변이는 FIX 단백질에 과기능 활성을 부여한다. 예를 들어, "파두아" 단백질 (예를 들어, R338L 돌연변이를 함유하는 인자 IX)은 생체 내 야생형 인자 IX보다 5 배 내지 10 배 더 활성인 것으로 나타났다. 미국특허 제6,531,298호; Simioni P. et al., N Engl J Med. 361(17):1671-75 (2009), 그 전문은 본원에 참조로 포함된다. 따라서, 본 개시는 때때로 본원에서 "FIXp" 또는 "pFIX"로 지칭되는 파두아-FIX 단백질을 암호화하는 아미노산 및 핵산 작제물을 제공한다.

[0054] 본원에 사용된 바와 같이, 용어 "인자 IX 폴리뉴클레오티드" 및 "FIX 폴리뉴클레오티드"는 예를 들어, 유럽 약전 9.0의 제2.7.11장에 기재된 1 단계 인자 IX 응고 검정을 사용하여 측정된 바와 같은, 특정 조건 하에서, 인자 IX 세린 단백질분해효소 활성을 갖는 인자 IX 폴리펩티드를 암호화하는 폴리뉴클레오티드를 지칭한다. 인자 IX 폴리뉴클레오티드는 인자 IX 프리-프로-폴리펩티드, 인자 IX 프로-펩티드 및 성숙된 단일-사슬 인자 IX 폴리펩티드를 포함하는 인자 IX 단일-사슬 전구체 폴리펩티드를 암호화하는 폴리뉴클레오티드를 포함하며, 이는 상기 기술된 번역 후 가공에 의해 활성화될 때, 인자 IX 세린 단백질분해효소 활성을 갖는 활성 인자 IX 단백질이 된다. 인자 IX 폴리뉴클레오티드의 정의에는 R338L 변이체를 포함하는 인자 IX 폴리펩티드를 암호화하는 폴리뉴클레오티드가 구체적으로 포함된다. 예시적인 실시 양태에서, 인간 인자 IX 폴리뉴클레오티드는 경쇄 및 중쇄, FIX-MP-AA (서열 번호 10, 도 11a에 나타냄)를 포함하는 야생형 인간 인자 IX 폴리펩티드의 일부 또는 경쇄 및 중쇄, FIXp-MP-AA (서열 번호 12, 도 12에 나타냄)를 포함하는 파두아 인간 인자 IX 폴리펩티드의 일부에 대해 서열 동일성이 높은 (예를 들어, 적어도 85%, 90%, 95%, 99% 또는 그 이상) 아미노산 서열을 포함하는 폴리펩티드를 암호화하는 폴리뉴클레오티드를 지칭한다.

[0055] 본원에 기재된 바와 같이, 인자 IX 폴리뉴클레오티드는 조절 요소, 예를 들어 프로모터, 인핸서, 종결자, 폴리아데닐화 서열 및 인트론 뿐만 아니라 바이러스 패키징 요소, 예를 들어 역 말단 반복 서열 ("ITR") 및/또는 비바이러스성 숙주 세포에서 폴리뉴클레오티드의 증식을 지지하는 다른 요소, 예를 들어 박테리아, 효모 또는 포유류 숙주 세포에서의 폴리뉴클레오티드의 복제를 지지하는 레플리콘을 포함할 수 있다.

[0056] 본 개시에서 코돈-변경된 인자 IX 폴리뉴클레오티드가 특히 사용된다. 본원에 기재된 바와 같이, 코돈-변경된 FIX 폴리뉴클레오티드는 고유-코딩된 인자 IX 작제물 (예를 들어, 야생형 인간 코돈을 사용하여 동일한 인자 IX 아미노산 서열을 암호화하는 폴리뉴클레오티드)에 의해 제공된 인자 IX 발현의 수준과 비교하여 생체 내 이식유전자 인자 IX의 증가된 발현을 제공한다. 본원에 사용된 바와 같이, 용어 "증가된 발현"은 고유-코딩된 인자 IX 작제물을 투여한 동물의 혈액의 이식유전자 인자 IX 단백질의 수준과 비교하여, 인자 IX를 암호화하는 코돈-변경된 폴리뉴클레오티드를 투여받은 동물의 혈액에서 이식유전자 인자 IX 단백질의 증가된 수준을 지칭한다. 단백질의 발현 증가는 인자 IX 활성의 증가로 이어진다; 따라서, 증가된 발현은 증가된 활성을 유발한다.

[0057] 일부 실시 양태에서, 증가된 발현은 고유-코딩된 인자 IX 폴리뉴클레오티드가 투여된 동물의 혈액에서 이식유전자 인자 IX 폴리펩티드의 수준과 비교하여, 코돈-변경된 인자 IX 폴리뉴클레오티드가 투여된 동물의 혈액에서 적어도 25% 이상의 이식유전자 인자 IX 폴리펩티드를 지칭한다. 본 개시의 목적을 위해, 증가된 발현은 근본적인 아미노산 치환, 예를 들어 "파두아" 돌연변이에 의해 야기된 과다활성보다는 코돈 서열의 변경에 의해 생성된 효과를 지칭한다. 즉, "파두아" 인자 IX 폴리뉴클레오티드를 암호화하는 코돈-최적화된 서열로부터 수득된 발현 수준을 고유-코딩된 "파두아" 단백질로부터 수득된 발현 수준과 비교한다. 일부 실시 양태에서, 증가된 발현은 고유 코딩된 인자 IX 폴리뉴클레오티드가 투여된 동물의 혈액에서 이식유전자 인자 IX 폴리펩티드 수준과 비교하여, 코돈-변경된 인자 IX 폴리뉴클레오티드를 투여한 동물의 혈액에서 적어도 50% 이상, 적어도 75% 이상, 적어도 100% 이상, 적어도 3 배 이상, 적어도 4 배 이상, 적어도 5 배 이상, 적어도 6 배 이상, 적어도 7 배 이상, 적어도 8 배 이상, 적어도 9 배 이상, 적어도 10 배 이상, 적어도 15 배 이상, 적어도 20 배 이상, 적어도 25 배 이상, 적어도 30 배 이상, 적어도 40 배 이상, 적어도 50 배 이상, 적어도 60 배 이상, 적어도 70 배 이상, 적어도 80 배 이상, 적어도 90 배 이상, 적어도 100 배 이상, 적어도 125 배 이상, 적어도 150 배 이상, 적어도 175 배 이상, 적어도 200 배 이상, 적어도 225 배 이상, 또는 적어도 250 배 이상의 이식유전자 인자 IX 폴리펩티드를 지칭한다. 동물의 혈액에서 인자 IX 폴리펩티드 수준은 예를 들어 인자 IX 폴리펩티드에 특이적인 ELISA 검정을 사용하여 측정될 수 있다.

[0058] 본원의 "인자 IX 활성" 또는 "인자 IX 세린 단백질분해효소 활성"은, 인자 VIIIa 보조 인자의 존재 하에서, 예를 들어 야생형 인자IX에서 Arg194-Ile195 펩티드 결합의 가수 분해를 통해 인자 X 폴리펩티드를 절단하여, 인자X를 인자 Xa로 활성화하는 능력을 의미한다. 활성 수준은 당 업계에 공지된 임의의 인자 IX 활성을 사용하여 측정될 수 있으며; 적합한 검정이 본원에 요약되어 있으며; 인자 IX 활성을 측정하기 위한 예시적인 검정은 본원에 제공된 실시예에서 사용된 유럽 약전 9.0의 제2.7.11장에 기재된 1 단계 인자 IX 응고 검정이다. 일부 실시 양태에서, FIX 활성이 결여된 인간 혈장이 1 단계 응고 검정에서 인자 IX특이성을 측정하기 위해 대조군으로 사용된다.

[0059] 특정 인자 IX 변이체는 *생체 내* 야생형 인자 IX변이체와 비교하여 특정 활성이 향상되었기 때문에, 예를 들어, 인간 '파두아' 변이체는 고유-코딩된 유형 인간 인자 IX보다 인자IX 세린 단백질분해효소 활성이 5 배 내지 10 배 더 크고, 일부 실시 양태에서, 인자 IX 폴리뉴클레오티드 조성물의 치료 잠재력은 예를 들어 증가된 인자 IX 발현 대신에 또는 이에 추가하는, 인자 IX 폴리뉴클레오티드를 투여한 동물의 혈액에서 인자 IX 활성의 증가에 의해 평가된다. 일부 실시 양태에서, 본원에 사용된 바와 같이, 증가된 인자 IX 활성은 고유-코딩된 인자 IX 폴리뉴클레오티드의 투여 전 동물의 혈액에서 기준선 인자 IX 활성에 비한, 고유-코딩된 인자 IX 폴리뉴클레오티드를 투여한 동물의 혈액에서 인자 IX 활성의 증가와 비교하여, 코돈-변경된 인자 IX 폴리뉴클레오티드의 투여 전 동물의 혈액에서 기준선 인자 IX 활성에 비해, 코돈-변경된 인자 IX 폴리뉴클레오티드가 투여된 동물의 혈액에서 인자 IX 활성이 더 크게 증가함을 지칭한다. 일부 실시 양태에서, 증가된 인자 IX 활성은 고유-코딩된 인자 IX 폴리뉴클레오티드의 투여 전 동물의 혈액에서 기준선 인자 IX 활성에 비한, 고유-코딩된 인자 IX 폴리뉴클레오티드를 투여한 동물의 혈액에서 인자 IX 활성의 증가와 비교하여, 코돈-변경된 인자 IX 폴리뉴클레오티드의 투여 전 동물의 혈액에서 기준선 인자 IX 활성에 비해, 코돈-변경된 인자 IX 폴리뉴클레오티드가 투여된 동물의 혈액에서 인자 IX 활성의 적어도25% 이상 증가를 지칭한다. 일부 실시 양태에서, 증가된 인자 IX 활성은 고유 코딩된 인자 IX 폴리뉴클레오티드의 투여 전 동물의 혈액에서 기준선 인자 IX 활성에 비한, 고유 코딩된 인자 IX 폴리뉴클레오티드를 투여한 동물의 혈액에서 인자 IX 활성의 증가와 비교하여, 코돈-변경된 인자 IX 폴리뉴클레오티드의 투여 전 동물의 혈액에서 기준선 인자 IX 활성에 비해, 코돈-변경된 인자 IX 폴리뉴클레오티드가 투여된 동물의 혈액에서 인자 IX 활성이 적어도 50% 이상, 적어도 75% 이상, 적어도 100% 이상, 적어도 3 배 이상, 적어도 4 배 이상, 적어도 5 배 이상, 적어도 6 배 이상, 적어도 7 배 이상, 적어도 8 배 이상, 적어도 9 배 이상, 적어도 10 배 이상, 적어도 15 배 이상, 적어도 20 배 이상, 적어도 25 배 이상, 적어도 30 배 이상, 적어도 40 배 이상, 적어도 50 배 이상, 적어도 60 배 이상, 적어도 70 배 이상, 적어도 80 배 이상, 적

어도 90 이상, 적어도 100 배 이상, 적어도 125 배 이상, 적어도 150 배 이상, 적어도 175 배 이상, 적어도 200 배 이상, 적어도 225 배 이상, 또는 적어도 250 배 이상 증가를 지칭한다. 본원에 기술된 바와 같이, 유럽 약전 9.0의 제 2.7.11 장에 기재된 1 단계 인자 IX 응고 검정을 사용하여 활성을 측정한다.

[0060] 본원에 사용된 바와 같이, 용어 "혈우병"은 감소된 혈액 응고 또는 응고를 특징으로 하는 질환 상태 그룹을 지칭한다. 혈우병은A형, B형 또는 C형 혈우병 또는 세 가지 질환 유형의 종합을 의미할 수 있다. A형 혈우병 (혈우병 A)은 인자 VIII (FVIII) 활성의 감소 또는 손실에 의해 발생하며 혈우병 아형 중 가장 두드러진다. B 형 혈우병 (혈우병 B)은 인자 IX (FIX)의 응고 기능의 손실 또는 감소로 인해 발생한다. C 형 혈우병 (혈우병 C)은 인자 XI (FXI)의 응고 활동의 손실 또는 감소의 결과이다. 혈우병 A와 B는 X-연관 질환이며, 혈우병 C는 상염색체 연관이다. 혈우병에 대한 기존의 치료에는 Bebulin®-VH를 포함하는 FVIII, FIX, 및 FXI 뿐만 아니라 FEIBA-VH, 데스모프레신 및 혈장 주입과 같은 응고 인자의 예방 및 주문형 투여를 모두 포함한다.

[0061] 본원에 사용된 바와 같이, 용어 "인자 IX 유전자 요법" 또는 "FIX 유전자 요법"은 인자 IX 결핍 (예를 들어 B형 혈우병)과 관련된 하나 이상의 증상 (예를 들어, 임상 인자)의 재발을 완화, 감소 또는 예방하기 위해 환자에게 인자 IX를 암호화하는 핵산을 제공하는 임의의 치료적 접근법을 포함한다. 상기 용어는 인자 IX 결핍 (예를 들어, 혈우병 B)이 있는 개인의 건강을 유지 또는 개선하기 위해, 임의의 변형된 형태의 인자 IX (예를 들어, 인자 VIII R384L 변이체)를 포함하는 인자 IX 분자를 암호화하는 핵산을 포함하는 임의의 화합물, 약물, 절차 또는 요법을 투여하는 것을 포함한다. 당업자는 예를 들어 본 개시에 따라 얻어진 결과에 기초하여 FIX 유전자 요법의 과정 또는 FIX 유전자 요법 치료제의 용량이 변경될 수 있음을 이해할 것이다.

[0062] 용어 "치료적 유효량 또는 용량" 또는 "치료적으로 충분한 양 또는 용량" 또는 "유효하거나 충분한 양 또는 용량"은 투여되는 치료 효과를 생성하는 용량을 지칭한다. 예를 들어, 혈우병 치료에 유용한 치료적 유효량의 약물은 혈우병과 관련된 하나 이상의 증상을 예방하거나 완화시킬 수 있는 양일 수 있다.

[0063] 일부 실시 양태에서, 치료적으로 유효한 치료는 대상체에서 출혈 사고의 빈도 및/또는 심각성을 감소시킨다.

[0064] 본원에 사용된 바와 같이, 용어 "유전자"는 폴리펩티드 사슬 (예를 들어, 코딩 영역)를 코딩하는 DNA 분자의 세그먼트를 지칭한다. 일부 실시 양태에서, 유전자는 폴리펩티드 사슬의 생성에 관여하는 코딩 영역 (예를 들어, 프로모터, 인핸서, 폴리아데닐화 서열, 5'-비번역 영역, 3'-비번역 영역 또는 인트론과 같은 조절 요소)을 선행, 추적 및/또는 개재하는 영역에 의해 위치된다.

[0065] 본원에 사용된 바와 같이, 용어 "조절 요소"는 세포에서 코딩 서열의 발현을 제공하는 프로모터, 인핸서, 종결자, 폴리아데닐화 서열, 인트론 등과 같은 뉴클레오티드 서열을 지칭한다.

[0066] 본원에 사용된 바와 같이, 용어 "프로모터 요소"는 코딩 서열의 발현을 제어하는 것을 돕는 뉴클레오티드 서열을 지칭한다. 일반적으로, 프로모터 요소는 유전자의 번역 시작 부위의 5'에 위치한다. 그러나, 특정 실시 양태에서, 프로모터 요소는 인트론 서열, 또는 코딩 서열의 3' 내에 위치할 수 있다. 일부 실시 양태에서, 유전자 요법 벡터에 유용한 프로모터는 표적 단백질의 고유 유전자 (예를 들어, 인자 VIII 프로모터)로부터 유래된다. 일부 실시 양태에서, 유전자 요법 벡터에 유용한 프로모터는 표적 유기체의 특정 세포 또는 조직에서의 발현에 특이적이다 (예를 들어, 간-특이적 프로모터). 또 다른 실시 양태에서, 복수의 잘 특성화된 프로모터 요소 중 하나가 본원에 기재된 유전자 요법 벡터에 사용된다. 잘 특성화된 프로모터 요소의 비제한적 예는 CMV 초기 프로모터, β-액틴 프로모터 및 메틸 CpG 결합 단백질 2 (MeCP2) 프로모터를 포함한다. 일부 실시 양태에서, 프로모터는 구성적 프로모터이며, 이는 표적 단백질의 실질적으로 일정한 발현을 유도한다. 다른 실시 양태에서, 프로모터는 유도성 프로모터이며, 이는 특정 자극 (예를 들어, 특정 치료 또는 작용제에 대한 노출)에 반응하여 표적 단백질의 발현을 유도한다. AAV-매개 유전자 요법을 위한 프로모터 설계에 대한 검토는 Gray et al. (Human Gene Therapy 22:1143-53 (2011), 그 내용이 모든 목적을 위해 그 전체가 참고로 명백하게 포함됨)을 참조한다.

[0067] 본원에 사용된 바와 같이, "CRM8" 요소는 작동 가능하게 연결된 유전자, 예를 들어 서열 번호 39에 대해 높은 서열 동일성을 가지는 간 특이적 방식으로, 인자 IX 폴리펩티드를 암호화하는 서열의 발현을 향상시키는 SERPINA1 유전자 (NCBI 수탁 번호 NM_000295.4)로부터 유래된 시스-작용 조절 모듈을 지칭한다. 본원에 사용된 바와 같이, CRM8 요소는 일부 실시 양태에서 인자 IX 폴리뉴클레오티드 내의 하나 이상의 카피, 예를 들어 1, 2, 3 개 이상의 카피에 포함되는 조절 요소의 단일 카피를 지칭한다. CRM8과 같은 CRM 요소에 대한 자세한 정보는 Chuah MK et al., Mol Ther., 22 (9):1605-13 (2014)를 참조.

[0068] 본원에 사용된 바와 같이, "MVM 인트론"은 서열 번호 53에 대해 높은 서열 동일성을 갖는 마우스의 미세 바이러

스로부터 유래된 인트론 서열을 지칭한다. MVM 인트론 자체에 대한 자세한 정보는 Haut and Pintel, J Virol.72(3):1834-43 (1998)을 참조하고, AAV 유전자 요법 벡터에서 MVM 인트론의 사용은 Wu Z et al., Mol Ther., 16(2):280-9 (2008)을 참조하며, 둘 다 본원에 참조로 포함된다.

[0069] 본원에 사용된 바와 같이, 용어 "작동 가능하게 연결된"은 제 2 뉴클레오티드 서열이 하나 이상의 제 1 참조 뉴클레오티드 서열과 관련된 특성 (예를 들어, 전사 속도)에 영향을 줄 수 있게 하는 제 1 참조 뉴클레오티드 서열 (예를 들어, 유전자)과 제 2 뉴클레오티드 서열 (예를 들어, 조절 제어 요소) 사이의 관계를 지칭한다. 본 개시의 맥락에서, 조절 요소가 인자 IX 이식유전자의 전사에 영향 (예를 들어, 촉진 또는 조직 선택적 영향)을 가하는 유전자 요법 벡터 내에 위치될 때 조절 요소는 인자 IX 이식유전자에 작동 가능하게 연결된다.

[0070] 본원에 사용된 바와 같이, 용어 "벡터"는 인자 IX 핵산을 숙주 세포 내로 전달하는데 사용되는 임의의 핵산 작제물을 지칭한다. 일부 실시 양태에서, 벡터는 핵산 작제물을 복제하는 기능을 하는 레플리콘을 포함한다. 유전자 요법에 유용한 벡터의 비제한적인 예는 플라스미드, 파지, 코스미드, 인공 염색체 및 바이러스를 포함하며, 이는 생체 내 자율 복제 단위로서 기능한다. 일부 실시 양태에서, 벡터는 인자 IX 핵산을 숙주 세포 내로 도입하기 위한 바이러스 벡터이다. 유전자 요법에 유용한 많은 변형된 진핵 바이러스가 당 업계에 공지되어 있다. 예를 들어, 아데노-관련 바이러스 (AAV)는 인간이 바이러스의 천연 숙주이고, 천연 바이러스는 어떠한 질환에도 기여하지 않으며, 바이러스는 가벼운 면역 반응을 나타내지 않기 때문에 인간 유전자 요법에 사용하기에 특히 적합하다.

[0071] 본원에 사용된 바와 같이, 용어 "인자 IX 바이러스 벡터"는 인자 IX 폴리펩티드를 암호화하는 인자 IX 폴리뉴클레오티드를 포함하는 재조합 바이러스를 지칭하며, 이는 적합한 동물 숙주 (예를 들어, 인간) 내로 도입될 때 인자 IX 폴리펩티드의 발현에 충분하다. 인자 IX 바이러스 벡터의 정의에 구체적으로 인자 IX 폴리펩티드를 암호화하는, 코돈-변경된 인자 IX 폴리뉴클레오티드가 바이러스의 게놈에 삽입된 재조합 바이러스가 포함된다. 또한 인자 IX 바이러스 벡터의 정의에 구체적으로 인자 IX 폴리펩티드를 암호화하는, 바이러스의 천연 게놈이 인자 IX 폴리뉴클레오티드로 대체된 재조합 바이러스가 포함된다. 인자 IX 바이러스 벡터의 정의에 인자 IX의 "파두아" 변이체를 암호화하는 인자 IX 폴리뉴클레오티드를 포함하는 재조합 바이러스가 포함된다.

[0072] 본원에 사용된 바와 같이, 용어 "인자 IX 바이러스 입자"는 인자 IX 폴리펩티드를 암호화하는 인자 IX 폴리뉴클레오티드를 캡슐화하는 바이러스 입자를 지칭하며, 이는 적합한 동물 숙주 (예를 들어, 인간)에 도입될 때 인자 IX 폴리펩티드의 발현에 특이적이다. 인자 IX 바이러스 입자의 정의에 구체적으로 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 인자 IX 폴리뉴클레오티드가 삽입된 게놈을 캡슐화하는 재조합 바이러스 입자가 포함된다. 또한 인자 IX 바이러스 입자의 정의에 구체적으로 바이러스의 고유 게놈을 대체하는 인자 IX 폴리펩티드를 암호화하는 인자 IX 폴리뉴클레오티드를 캡슐화하는 재조합 바이러스 입자가 포함된다. 인자 IX 바이러스 입자의 정의 내에 인자 IX 폴리뉴클레오티드를 캡슐화하는 재조합 바이러스 입자가 포함되는데, 이는 인자 IX의 "파두아" 변이체를 암호화한다.

[0073] 본원에서 "AAV" 또는 "아데노-관련 바이러스"는 바이러스의 파르보바이러스(Parvoviridae) 속 내의 데펜도파르보바이러스(Dependoparvovirus)를 의미한다. 본원에서 사용된 바와 같이, AAV는 자연 발생 AAV 캡 유전자에 의해 암호화된 캡시드 단백질을 사용하여 캡시드로 패키징된 재조합 인자 IX 폴리뉴클레오티드로부터 유래된 재조합 바이러스, 또는 비 천연 캡시드 캡 유전자에 의해 암호화된 캡시드 단백질을 사용하여 캡시드로 패키징된 재조합 인자 IX 폴리뉴클레오티드로부터 유래된 재조합 바이러스로 인자 IX 폴리뉴클레오티드가 삽입된 자연 발생 "야생형" AAV 게놈으로부터 유래된 바이러스를 지칭할 수 있다. AAV의 정의에 인자 IX 폴리뉴클레오티드를 캡슐화하는 AAV 유형 1 (AAV1), AAV 유형 2 (AAV2), AAV 유형 3 (AAV3), AAV 유형 4 (AAV4), AAV 유형 5 (AAV5), AAV 유형 6 (AAV6), AAV 유형 7 (AAV7), AAV 유형 8 (AAV8) 및 AAV 유형 9 (AAV9) 바이러스 및 인자 IX 폴리뉴클레오티드를 캡슐화하는 하나 이상의 변이체 AAV 캡시드 단백질에 의해 형성된 바이러스가 포함된다.

[0074] 본원에서 "AAV8", "AAV-8" 또는 "AAV 혈청형 8"은 인자 IX 폴리뉴클레오티드를 캡슐화하는 AAV8 캡시드 바이러스 단백질에 의해 형성된 바이러스를 의미한다.

[0075] 본원에 사용된 바와 같이, 용어 "CpG"는 단일 가닥의 DNA를 따르는 시토신-구아닌 디뉴클레오티드를 지칭하고, "p"는 둘 사이의 포스페이트 연결을 나타낸다.

[0076] 본원에 사용된 바와 같이, 용어 "CpG 아일랜드"는 통계적으로 상승된 밀도의 CpG 디뉴클레오티드를 가지는 폴리뉴클레오티드 내의 영역을 지칭한다. 본원에 사용된 바와 같이, 폴리뉴클레오티드 (예를 들어, 코돈-변경된 인자 IX 단백질을 암호화하는 폴리뉴클레오티드)의 영역은 하기 경우에 CpG 아일랜드이며, 200염기 쌍 윈도우에

걸쳐:(i) 이 영역은 50% 이상의 GC 함량을 가지며, (ii) 예상되는 CpG 디뉴클레오티드 당 관찰된 CpG 디뉴클레오티드의 비율이 하기 관계에 의해 정의된 바와 같이, 적어도0.6인 경우임:

[0077]

$$\frac{N[CpG]*N[\text{윈도우의 길이}]}{N[C]*N[G]} \geq 0.6.$$

[0078] CpG 아일랜드를 식별하는 방법에 대한 추가적 정보는 Gardiner-Garden M. et al., J Mol Biol., 196(2):261-82 (1987)을 *참조하며*, 이 내용은 모든 목적을 위해 그 전체가 본원에 참조로 포함된다.

[0079] 본원에 사용된 바와 같이, 용어 "핵산"은 단일 가닥 또는 이중 가닥 형태의 데옥시리보뉴클레오티드 또는 리보뉴클레오티드 및 이의 중합체, 및 이의 상보체를 지칭한다. 상기 용어는 공지된 뉴클레오티드 유사체 또는 합성, 자연 발생 및 비천연 발생인, 참조 핵산과 유사한 결합 특성을 가지며, 참조 뉴클레오티드와 유사한 방식으로 대사되는, 공지된 뉴클레오티드 유사체 또는 변형된 백본 잔기 또는 연결을 함유하는 핵산을 포함한다. 이러한 유사체의 예는 포스포로티오에이트, 포스포르아미데이트, 메틸포스포네이트, 키랄-메틸 포스포네이트, 2-O-메틸 리보뉴클레오티드 및 펩티드-핵산 (PNA)을 포함하지만, 이에 제한되지 않는다. 그러나, 환자에서 유전자요법에 사용하기 위한 본원에서 특히 유용한 실시 양태는 포스포디에스테르 결합을 사용한다.

[0080] 본원에서 "핵산 조성물"은 인자 IX 폴리뉴클레오티드를 암호화하는 인자 IX 폴리뉴클레오티드를 포함하는 분자의 임의의 분자 또는 제제를 의미한다. 핵산 조성물의 정의 내에 인자 IX 폴리뉴클레오티드, 인자 IX 폴리뉴클레오티드의 수용액, 인자 IX 폴리뉴클레오티드를 캡슐화하는 바이러스 입자 및 인자 IX 폴리뉴클레오티드를 캡슐화하는 바이러스 입자의 수성 제제가 포함된다. 본원에 개시된 바와 같은 핵산 조성물은 FIX 폴리펩티드를 암호화하는 코돈-변경된 FIX 유전자를 포함한다.

[0081] 용어 "아미노산"은 자연 발생 아미노산과 유사한 방식으로 기능하는 아미노산 유사체 및 아미노산 모방 체를 포함하는 자연 발생 및 비천연 아미노산을 지칭한다. 자연 발생 아미노산은 유전자 코드에 의해 암호화된 것뿐만 아니라 나중에 변형되는 아미노산, 예를 들어 히드록시프롤린, y-카르복시글루타메이트 및 O-포스포세린을 포함한다. 자연 발생 아미노산은 예를 들어 D- 및 L-아미노산을 포함할 수 있다. 아미노산 서열에 관해, 당업자는 암호화된 서열에서 단일 아미노산 또는 소량의 아미노산을 변경, 첨가 또는 결실시키는 핵산 또는 펩티드 서열에 대한 개별 치환, 결실 또는 첨가는 변형이 화학적으로 유사한 아미노산으로 아미노산의 치환을 초래하는 "보존적으로 변형된 변이체"임을 인식할 것이다. 기능적으로 유사한 아미노산을 제공하는 보존적 치환 표는 당업계에 잘 알려져 있다. 이러한 보존 적으로 변형된 변이체는 본 개시의 다형성 변이체, 종간 상동체 및 대립 유전자에 추가하고 이를 배제하지 않는다.

[0082] 기능적으로 유사한 아미노산을 제공하는 보존적 아미노산 치환은 당업계에 잘 알려져 있다. 특정 아미노산, 예를 들어 촉매적, 구조적 또는 입체적으로 중요한 아미노산의 기능성에 따라, 아미노산의 상이한 그룹화는 서로에 대한 보존적 치환으로 간주될 수 있다. 표 1은 아미노산의 전하 및 극성, 아미노산의 소수성, 아미노산의 표면 노출/구조적 성질, 및 아미노산의 2 차 구조 경향에 기초하여 보존적 치환으로 간주되는 아미노산의 그룹화를 제공한다.

[0083] 표 1. 단백질에서 잔기의 관능기에 기초한 보존적 아미노산 치환의 그룹화.

중요한 특징	보존적 그룹화
전하/극성	1. H, R 및 K 2. D 및 E 3. C, T, S, G, N, Q 및 Y 4. A, P, M, L, I, V, F 및 W
소수성	1. D, E, N, Q, R 및 K 2. C, S, T, P, G, H 및 Y 3. A, M, I, L, V, F 및 W
구조/표면 노출	1. D, E, N, Q, H, R 및 K 2. C, S, T, P, A, G, W 및 Y 3. M, I, L, V 및 F
이차 구조 경향	1. A, E, Q, H, K, M, L 및 R 2. C, T, I, V, F, Y 및 W 3. S, G, P, D 및 N
진화적 보존	1. D 및 E 2. H, K 및 R 3. N 및 Q 4. S 및 T 5. L, I 및 V 6. F, Y 및 W 7. A 및 G 8. M 및 C

[0084]

[0085] 2개 이상의 핵산 또는 펩티드 서열과 관련하여, 용어 "동일한" 또는 백분율 "동일성"은 하기 기술된 기본 매개변수가 있는 BLAST 또는 BLAST 2.0 서열 비교 알고리즘을 사용하여 측정한 경우 또는 수동 정렬 및 육안 검사에 의해 측정된 바와 같이 동일하거나 동일 (즉, 비교 창 또는 지정된 영역에 대한 최대 대응을 위해 비교 및 정렬된 경우 특정된 영역에 대해 약 60% 동일성, 바람직하게는 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 더 높은 동일성)한 특정된 백분율의 아미노산 잔기 또는 뉴클레오티드인 2개 이상의 서열 또는 하위 서열을 지칭한다.

[0086] 당업계에 공지된 바와 같이, 다수의 상이한 프로그램을 사용하여 단백질 (또는 하기 논의되는 바와 같은 핵산)이 공지된 서열에 대해 서열 동일성 또는 유사성을 갖는지 여부를 확인할 수 있다. 서열 동일성 및/또는 유사성은 하기를 포함하지만 이에 제한되지 않는 표준 기술을 사용하여 측정된다: Smith & Waterman, Adv.Appl.Math., 2:482 (1981)의 국소 서열 동일성 알고리즘, Needleman & Wunsch, J. Mol. Biol., 48:443 (1970) 의 서열 동일성 정렬 알고리즘, Pearson & Lipman, Proc. Natl. Acad. Sci. U.S.A., 85:2444 (1988)의 유사성 검색 방법, 이들 알고리즘 (위스콘신 유전학 소프트웨어 패키지의 GAP, BESTFIT, FASTA, 및 TFASTA, Genetics Computer Group, 575 Science Drive, Madison, WI)의 컴퓨터화된 구현, Devereux et al., Nucl.Acid Res., 12:387-395 (1984) 에 의해 기술된, 바람직하게는 기본 설정을 사용하거나 검사에 의한 Best Fit 서열 프로그램.바람직하게, 백분율 동일성은 하기 매개변수에 기초하여 FastDB에 의해 계산된다:불일치 페널티 1; 간격 페널티 1; 간격 크기 페널티 0.33; 및 연결 패널티 30, "Current Methods in Sequence Comparison and Analysis," Macromolecule Sequencing and Synthesis, Selected Methods and Applications, pp 127-149 (1988), Alan R. Liss, Inc, 이는 참조로 포함됨.

[0087] 유용한 알고리즘의 예는 PILEUP이다. PILEUP은 점진적 쌍 정렬을 사용하여 관련 서열 그룹에서 다중 서열 정렬을 만든다. 또한 정렬을 만드는 데 사용되는 클러스터링 관계를 보여주는 트리를 플롯할 수 있다. PILEUP은 Feng & Doolittle, J. Mol. Evol.35:351-360 (1987)의 점진적 정렬 방법을 단순화를 사용하고; 이 방법은 Higgins & Sharp CABIOS 5:151-153 (1989)에 의해 기술된 것과 유사하며, 둘 모두 참조로 포함된다. 기본 간격

가중치 3.00, 기본 간격 길이 가중치 0.10 및 가중치 끝 간격을 포함하는 유용한 PILEUP 매개변수.

[0088] 유용한 알고리즘의 또 다른 예는 하기에 설명된 BLAST 알고리즘이다. Altschul et al., J. Mol. Biol. 215, 403-410, (1990); Altschul et al., Nucleic Acids Res.25:3389-3402 (1997); 및 Karlin et al., Proc. Natl. Acad. Sci. U.S.A. 90:5873-5787 (1993), 둘 모두 참조로 포함된다. 특히 유용한 BLAST 프로그램은 Altschul et al., Methods in Enzymology, 266:460-480 (1996); http://blast.wustl/edu/blast/ README.html]로부터 획득한 WU-BLAST-2 프로그램이다. WU-BLAST-2는 여러 검색 매개변수를 사용하며, 대부분 검색 매개변수가 기본값으로 설정되어 있다. 조정 가능한 매개변수는 다음 값으로 설정된다. 오버랩 스팬 =1, 오버랩 비율 = 0.125, 워드 역치 (T) = 11. HSP S 및 HSP S2 매개변수는 동적 값이며 특정 서열의 구성 및 관심있는 서열을 검색할 특정 데이터베이스의 구성에 의존하는 프로그램 자체에 의해 설정되며: 그러나 감도를 높이기 위해 값을 조정할 수 있다.

[0089] 추가 유용한 알고리즘은 Altschul et al., Nucl.Acids Res., 25:3389-3402(참조로 포함)에 의해 보고된, 간극형(gapped) BLAST이다. 간극형 BLAST는 BLOSUM-62 대체 점수를 사용하며: 9로 설정된 역치 T 매개변수; 비간극형 확장을 유발하는 투 히트 방법; 10+k 비용의 k의 간격 길이를 충전하고; Xu는 16으로, Xg는 데이터베이스 검색 단계에서 40으로, 알고리즘의 출력 단계에서 67로 설정한다. 간극 정렬은 약 22 비트에 해당하는 점수에 의해 유발된다.

[0090] % 아미노산 서열 동일성 값은 일치하는 동일한 잔기의 수를 정렬된 영역에서 "더 긴" 서열의 총 잔기 수로 나눔으로써 측정된다. "더 긴" 서열은 정렬된 영역에서 가장 실제의 잔기를 갖는 서열이다 (배열 점수를 최대화하기 위해 WU-Blast-2에 의해 도입된 간격은 무시됨). 유사한 방식으로, 확인된 폴리펩티드의 코딩 서열에 대한 "백분율 (%) 핵산 서열 동일성"은 세포주기 단백질의 코딩 서열의 뉴클레오티드 잔기와 동일한 후보 서열의 뉴클레오티드 잔기의 백분율로 정의된다. 선호되는 방법은 WU-BLAST-2의 BLASTN 모듈을 기본 매개변수로 활용하고 오버랩 스팬 및 오버랩 비율을 각각 1 및 0.125로 설정한다.

[0091] 정렬은 정렬될 서열에서 간격의 도입을 포함할 수 있다. 또한, 도 3a의 야생형 인자 IX 서열 (서열 번호 2)에 의해 암호화된 단백질보다 더 많거나 적은 아미노산을 함유하는 서열의 경우, 한 실시 양태에서, 서열 동일성의 백분율은 아미노산 또는 뉴클레오티드의 총 수와 관련하여 동일한 아미노산 또는 뉴클레오티드의 수에 기초하여 결정될 수 있다. 따라서, 예를 들어, 하기 논의되는 바와 같이, 도 3a (서열 번호 2)에 나타낸 것보다 짧은 서열의 서열 동일성은 한 실시 양태에서 더 짧은 서열의 뉴클레오티드 수를 사용하여 결정될 것이다. 동일성 백분율 계산에서 상대적 가중치는 삽입, 결실, 치환 등과 같은 서열 변이의 다양한 발현에 할당되지 않는다.

[0092] 한 실시 양태에서, 동일성만이 양(+1)으로 스코어링되고, 간격을 포함하는 모든 형태의 서열 변이는 "0"의 값이 할당되며, 이는 서열 유사성 계산을 위해 하기 기술되는 바와 같이 가중 스케일 또는 매개변수에 대한 필요성을 제거한다. 서열 동일성 백분율은 예를 들어, 일치하는 동일한 잔기의 수를 정렬된 영역에서 "더 짧은" 서열의 총 잔기 수로 나누고 100을 곱함으로써 계산될 수 있다. "더 긴" 서열은 정렬된 영역에서 가장 실제의 잔기를 갖는 서열이다.

[0093] 용어 "대립유전자 변이체"는 특정 유전자좌에서 유전자의 다형성 형태, 뿐만 아니라 유전자의 mRNA 전 사체로부터 유래된 cDNA 및 이들에 의해 코딩된 폴리펩티드를 지칭한다. 용어 "바람직한 포유류 코돈"은 하기 목록으로부터 선택된 포유류 세포에서 발현되는 단백질에서 가장 빈번하게 사용되는 아미노산을 암호화하는 코돈 세트 중 코돈의 서브 세트를 지칭한다: Gly (GGC, GGG); Glu (GAG); Asp (GAC); Val (GTG, GTC); Ala (GCC, GCT); Ser (AGC, TCC); Lys (AAG); Asn (AAC); Met (ATG); Ile (ATC); Thr (ACC); Trp (TGG); Cys (TGC); Tyr (TAT, TAC); Leu (CTG); Phe (TTC); Arg (CGC, AGG, AGA); Gln (CAG); His (CAC); 및 Pro (CCC).

[0094] 본원에 사용된 바와 같이, 용어 "코돈-변경된"은 폴리펩티드(예를 들어, 인자 IX 단백질)를 암호화하는 폴리뉴클레오티드 서열을 지칭하며, 폴리펩티드를 암호화하는 천연 폴리뉴클레오티드의 적어도 하나의 코돈은 폴리뉴클레오티드 서열의 특성을 개선시키기 위해 변경되었다. 일부 실시 양태에서, 개선된 특성은 폴리펩티드를 암호화하는 mRNA의 전사 증가, mRNA의 안정성 증가 (예를 들어, 개선된 mRNA 반감기), 폴리펩티드의 번역 증가 및/또는 벡터 내의 폴리뉴클레오티드의 증가된 패키징을 촉진한다. 개선된 특성을 달성하기 위해 사용될 수 있는 변형의 비제한적 예는 특정 아미노산에 대한 코돈의 사용 및/또는 분포 변경, 전체 및/또는 국소 GC 함량 조정, AT-풍부 서열 제거, 반복 서열 요소 제거, 전체 및/또는 국소 CpG 디뉴클레오티드 함량 조정, 암호화 조절 요소 (예를 들어, TATA 박스 및 CCAAT 박스 요소) 제거, 인트론/엑손 스플라이스 부위 제거, 조절 서열 개선 (예를 들어, 코작 공통 서열 도입) 및 전사된 mRNA에서 2차 구조 (예를 들어, 스템-루프)를 형성할 수 있는 서열 요소 제거를 포함한다.

[0095] 본원에서 논의된 바와 같이, 본원의 개시 내용의 구성 성분을 지칭하는 다양한 명명법이 존재한다. "CS-번호" (예를 들어, "CS02", "CS03", "CS04", "CS05", "CS06" 등)은 FIX 폴리펩티드를 암호화하는 코돈 변경된 폴리뉴클레오티드 및/또는 변이체를 포함하는, 암호화된 폴리펩티드를 지칭한다. 예를 들어, CS02-FL은 전장(Full Length) 코돈 변경된 CS02 폴리뉴클레오티드 서열 또는 CS02 폴리뉴클레오티드 서열에 의해 암호화된 아미노산 서열 (때때로 본원에서 아미노산 서열(Amino Acid)에 대한 "CS02-FL-AA" 및 핵산(Nucleic Acid) 서열에 대한 "CS02-FL-NA" (서열 번호 5)로 지칭됨)을 지칭한다. 유사하게, "CS02-LC"는 FIX 폴리펩티드의 경쇄를 암호화하는 코돈 변경된 핵산 서열 ("CS02-LC-NA" (서열 번호 42)) 또는 CS02 폴리뉴클레오티드 서열에 의해 암호화된 FIX 경쇄의 아미노산 서열 (때로는 "CS02-LC-AA"로서 본원에 지칭됨)을 지칭한다. 마찬가지로, CS02-HC, CS02-HC-AA 및 CS02-HC-NA (서열 번호 41)는 FIX 중쇄에 대해 동일하다. 관련 기술 분야의 당업자에 의해 인식될 바와 같이, 단지 코돈-변경된 (예를 들어, 파두아 인자 IX변이체와 비교하여 추가의 아미노산 치환을 함유하지 않는) CS02, CS03, CS04, CS05, CS06 등과 같은 작제물에 대해, 아미노산 서열은 코돈 최적화에 의해 변경되지 않기 때문에 아미노산 서열은 동일할 것이다. 따라서, 본 개시의 서열 작제물은 CS02-FL-NA (서열 번호 5), CS02-FL-AA, CS02-LC-NA (서열 번호 42), CS02-LC-AA, CS02-HC-AA, CS02-HC-NA (서열 번호 41), CS03-FL-NA (서열 번호 6), CS03-FL-AA, CS03-LC-NA (서열 번호 44), CS03-LC-AA, CS03-HC-AA, CS03-HC-NA (서열 번호 43), CS04-FL-NA (서열 번호 7), CS04-FL-AA, CS04-LC-NA (서열 번호 46), CS04-LC-AA, CS04-HC-AA, CS04-HC-NA, CS05-FL-NA (서열 번호 8), CS05-FL-AA, CS05-LC-NA (서열 번호 48), CS05-LC-AA, CS05-HC-AA, CS05-HC-NA (서열 번호 47), CS06-FL-NA (서열 번호 9), CS06-FL-AA, CS06-LC-NA (서열 번호 50), CS06-LC-AA, CS06-HC-AA 및 CS06-HC-NA (서열 번호 49)를 포함하나, 이에 제한되지는 않는다. 본원의 모든 "CS" 작제물은 FIXp 아미노산 서열을 암호화하거나 함유하지만, CS 작제물의 정의 내에 포함되는 것은 인간 야생형 FIX 아미노산 서열을 암호화하거나 함유하는 것임을 유의하여야 한다.

[0096] 본원에 사용된 바와 같이, 용어 "간-특이적 발현"은 다른 조직과 비교하여 간 조직에서 특정 유전자 (예를 들어, 코돈-변경된, 이식유전자 인자 IX 유전자)의 우선적 또는 우세한 생체 내 발현을 지칭한다. 일부 실시 양태에서, 간-특이적 발현은 특정 유전자의 모든 발현의 적어도 50%가 대상체의 간 조직 내에서 발생한다는 것을 의미한다. 다른 실시 양태에서, 간 특이적 발현은 모든 발현의 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 100%가 특정 유전자는 대상체의 간 조직 내에서 발생한다는 것을 의미한다. 따라서, 간-특이적 조절 요소는 간 조직에서 유전자의 간-특이적 발현을 유도하는 조절 요소이다.

[0097] 본원에 사용된 바와 같이, 용어 X "미만" 및 X% "미만"은 X 값을 제외하는 0 내지 X, 예를 들어 X%를 제외하는, 0% 내지 X%의 범위를 지칭한다. 본원에 사용된 바와 같이, 상기 용어는 0 또는 0%에서 시작하지만 X 또는 X%를 포함하지 않는 범위와 상호교환적으로 사용된다.

[0098] 본원에 사용된 바와 같이, 용어 X "이하" 또는 X% "이하"는 X 값을 포함하는 0 내지 X의 범위, 예를 들어 X%를 포함하는 0% 내지 X%의 범위를 의미한다. 본원에 사용된 바와 같이, 상기 용어는 0 또는 0%에서 시작하여 X 또는 X%를 포함하는 범위와 상호교환적으로 사용된다.

[0099] 본원에 사용된 바와 같이, 용어 X "보다 큰" 또는 X% "보다 큰"은 값 X를 제외하는 X의 상한, 예를 들어 X%를 제외하는, X% 내지 100%의 범위를 지칭한다. 본원에 사용된 바와 같이, 상기 용어는 백분율의 맥락에서 X 또는 X%를 포함하지 않으며 시작하고 상한이 100%인 범위와 상호교환적으로 사용된다.

[0100] 본원에 사용된 바와 같이, 용어 "적어도" X 또는 "적어도" X%는 X 값을 포함하여 X 내지 상한의 범위, 예를 들어 X%를 포함하여, 예를 들어 X% 내지 100%의 범위를 지칭한다. 본원에 사용된 바와 같이, 상기 용어는 백분율의 맥락에서 X 또는 X%로 시작하고 이를 포함하는 상한이 100%인 범위와 상호교환적으로 사용된다.

[0101] 본원에서 사용된 용어 "'X'와 'Y' 사이", "'X'%와 'Y'% 사이", "'X' 내지 'Y'", "'X'% 내지 'Y'%"는 X 및 Y 값을 포함하여 X 내지 Y의 범위, 예를 들어 X% 및 Y%를 포함하여 X% 내지 Y%의 범위를 지칭한다. 본원에 사용된 바와 같이, 상기 용어는 X 또는 X%에서 시작하여 Y 또는 Y%를 포함하는 범위와 상호교환적으로 사용된다.

[0102] III. 코돈-변경된 인자 IX 폴리뉴클레오티드

[0103] 일부 실시 양태에서, 본 개시는 인자 IX 또는 인자 IX 변이체를 암호화하는 코돈 변경된 핵산 조성물 (특정 실시 양태에서 FIXp 발견 사용)을 제공한다. 이들 코돈-변경된 폴리뉴클레오티드는 AAV-기반 유전자 요법 작제물에서 투여될 때 인자 IX의 현저하게 개선된 발현을 제공한다. 코돈-변경된 폴리뉴클레오티드는 또한 통상적으로 코돈-최적화된 작제물과 비교하여 개선된 AAV-비리온 패키징을 입증한다. 실시예 1에서 입증된 바와 같이, 출원인은 과활성 R338L 아미노산 치환 (성숙한 단일 사슬 인자 IX 폴리펩티드 서열에 기초하고; 인자 IX 프리-프로-

단백질 서열에 기초한 R384L)을 갖는 인자 IXp 폴리펩티드를 암호화하는 여러 코돈-변경된 폴리뉴클레오티드 (예를 들어, CS02-FL-NA, CS03-FL-NA, CS04-FL-NA, CS05-FL-NA 및 CS06-FL-NA (각각 서열 번호 5 내지 9))의 발견을 통해 이들 이점을 달성하였다. 실시예 2 및 3에서 입증된 바와 같이, 인자 IX 분자를 암호화하는 유전자 요법 벡터에 하나 이상의 간-특이적 조절 제어 요소 (예를 들어, CRM8)의 통합은 인자 IX 및 인자 IX 활성의 생체 내 및 시험관 내 발현을 추가로 증가시켰다.

[0104] 야생형 인자 IX는 28 개의 아미노산 신호 펩티드 (FIX-SP-AA (서열 번호 37)) 및 18개의 아미노산 프로-펩티드 (FIX-PP-AA (서열 번호 38))로 암호화되고, 이는 인자 IXa의 활성화 전에 암호화된 폴리펩티드로부터 절단된다. 당업자에 의해 이해되는 바와 같이, 신호 펩티드 및/또는 프로-펩티드는 다른 유전자 또는 다른 유기체로부터의 신호 펩티드 및/또는 프로-펩티드에 의해 성숙되고, 대체될 수 있고, 또는 신호 및 프로-펩티드가 세포 가공에 의해 제거된 후 성숙한 폴리펩티드의 서열에 영향을 주지 않으면서 완전히 제거될 수 있다.

[0105] 따라서, 일부 실시 양태에서, 본원에 제공된 코돈-변경된 폴리뉴클레오티드 (예를 들어, 핵산 조성물)는 성숙한 인자 IX 단일 사슬 폴리펩티드, 즉 인자 IX 경쇄, 활성화 펩티드 및 중쇄 (예를 들어, 야생형 인자 IX 유전자에 의해 암호화된 전장 폴리펩티드의 아미노산 47-461; FIX-FL-AA (서열 번호 2))를 암호화하는 CS02-FL-NA, CS03-FL-NA, CS04-FL-NA, CS05-FL-NA 또는 CS06-FL-NA (각각 서열 번호 5 내지 9)와 서열 동일성이 높은 뉴클레오티드 서열을 갖는다.

[0106] 또한, 당업계에 공지된 바와 같이, 인간 야생형 인자 IX는 단백질의 활성화시 단일 사슬 인자 IX 폴리펩티드로부터 절제된 인자 IX 경쇄와 중쇄 사이에 위치된 34 개의 아미노산 활성화 펩티드를 갖는다. 활성화 펩티드가 활성 인자 IX 폴리펩티드로부터 제거되기 때문에, 펩티드 자체는 궁극적 인자 IX 활성에 필수적이다. 따라서, 본원에 개시된 코돈-변경된 폴리뉴클레오티드에 의해 암호화된 인자 IX 폴리펩티드가 인간 야생형 활성화 펩티드 서열 (FIX-AP-AA (서열 번호 56))과 높은 서열 동일성을 가질 필요는 없다. 그러나, 암호화된 활성화 펩티드는 인자 IX 폴리펩티드의 활성화시 절제 가능해야 한다. 예를 들어, 일부 실시 양태에서, 암호화된 활성화 펩티드는 생체 내에서 인간 인자 IX에 의해 인식되고 절단될 수 있는 그의 N-및 C-말단에 인자 XI 절단 부위를 포함해야 한다.

[0107] 따라서, 일부 실시 양태에서, 본원에 제공된 코돈-변경된 폴리뉴클레오티드 (예를 들어, 핵산 조성물)는 인간 야생형 FIX 경쇄 서열 (FIX-LC-AA (서열 번호 62) 및 인간 야생형 FIX 중쇄 서열 (FIX-HC-AA (서열 번호 63))에 대해 높은 서열 동일성을 가지는 단일 사슬 인자 IX 폴리펩티드를 암호화하고, 추가로 2 개의 인자 XI 절단 부위를 갖는 경쇄의 C-말단을 중쇄의 N-말단(예를 들어, 활성화 펩티드)에 결합시키는 폴리펩티드 링커를 추가로 암호화한다.

[0108] 일부 실시 양태에서, 코돈-변경된 폴리뉴클레오티드에 의해 암호화된 인자 IX 경쇄 및 중쇄는 각각 FIXp 중쇄를 포함하는 인간 인자 IX 경쇄 및 중쇄이다. 다른 실시 양태에서, 코돈-변경된 폴리뉴클레오티드에 의해 암호화된 인자 IX 경쇄 및 중쇄는 다른 포유류 (예를 들어, 돼지 인자 IX)로부터의 중쇄 및 경쇄 서열이다. 또 다른 실시 양태에서, 인자 IX 경쇄 및 중쇄는 키메라 경쇄 및 중쇄 (예를 들어, 인간 및 제 2 포유류 서열의 조합)이다. 또 다른 실시 양태에서, 인자 IX 경쇄 및 중쇄는 다른 포유류로부터의 경쇄 및 중쇄의 인간화된 버전, 예를 들어, 인간 잔기가 선택된 위치에서 치환되어 인간에게 투여될 때 생성되는 펩티드의 면역원성을 감소시키는 다른 포유류의 경쇄 및 중쇄 서열이다.

[0109] 인간 유전자의 GC 함량은 25% 미만에서 90% 초과로 광범위하게 다양하다. 그러나, 일반적으로 더 높은 GC 함량을 갖는 인간 유전자가 더 높은 수준으로 발현된다. 예를 들어, Kudla et al. (PLoS Biol., 4(6):80 (2006))은 유전자의 GC 함량을 증가시키면 전사를 증가시키고 mRNA 전사체의 더 높은 정상 상태 수준에 영향을 미침으로써, 암호화된 폴리펩티드의 발현을 증가시킨다는 것을 입증한다. 일반적으로, 코돈-최적화된 유전자 작제물의 바람직한 GC 함량은 60% 이상인 것으로 생각된다. 예를 들어, scAAV8.FIXR338L 유전자 요법 벡터에서 인자 IX 유전자는 야생형 코딩 서열의 GC 함량을 41% GC에서 61% GC로 증가시키기 위해 유전자 최적화 소프트웨어 (Geneart)를 사용하여, 구체적으로 코돈 변경되었다. Wu Z. et al., Mol Ther 16:280-89 (2008) 및 Monahan PE et al., Hum Gene Ther., 26(2):69-81 (2015)를 참조. 그러나, 천연 AAV 게놈은 약 56%의 GC 함량을 갖는다.

[0110] 따라서, 일부 실시 양태에서, 본원에 제공된 코돈-변경된 폴리뉴클레오티드 (예를 들어, 핵산 조성물)는 천연 AAV 비리온 (예를 들어, 약 56% GC)의 GC 함량과 보다 밀접하게 일치하는 CG 함량을 가지며, 이는 포유류 세포에서의 발현에 대해 통상적으로 코돈-최적화된 폴리뉴클레오티드의 바람직한 CG 함량 (예를 들어, 60% 이상의 GC) 보다 낮다. 예를 들어, CS02-FL-NA (서열 번호 5)는 약 54%의 GC 함량을 가지며, CS03-FL-NA (서열 번호 6)는 약 55%의 GC 함량을 가지며, CS04-FL-NA (서열 번호 7)는 약 54.5%의 GC 함량을 가지며, CS05-FL-NA (서

열 번호 8)는 약 56.6%의 GC 함량을 가지고, CS06-FL-NA (서열 번호 9)는 약 55%의 GC 함량을 갖는다. 이들 작제물은 더 높은 GC 함량을 갖는 유사하게 코돈-변경된 서열과 비교하여 개선된 비리온 패키징을 제공해야 한다.

[0111] 따라서, 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드 (예를 들어, CS02-CS06 인자 IX 코딩 서열 중 하나와 높은 서열 동일성을 갖는 폴리뉴클레오티드)의 전체 GC 함량은 60% 미만이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 59% 미만이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 58% 미만이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 57% 미만이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56% 이하이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55% 이하이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54% 이하이다.

[0112] 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 53% 내지 59%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54% 내지 59%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55% 내지 59%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56% 내지 59%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 53% 내지 58%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54% 내지 58%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55% 내지 58%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56% 내지 58%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 53% 내지 57%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54% 내지 57%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55% 내지 57%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56% 내지 57%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 53% 내지 56%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54% 내지 56%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55% 내지 56%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 53% 내지 55%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54% 내지 55%이다.

[0113] 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54±0.5%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54±0.4%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54±0.3%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54±0.2%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54±0.1%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 54%이다.

[0114] 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55±0.5%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55±0.4%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55±0.3%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55±0.2%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55±0.1%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 55%이다.

[0115] 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56±0.5%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56±0.4%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56±0.3%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오

티드의 전체 GC 함량은 56±0.2%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56±0.1%이다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 전체 GC 함량은 56%이다.

[0116] 이들 CpG 디뉴클레오티드 (즉, 시토신 뉴클레오티드에 이어지는 구아닌 뉴클레오티드)는 생체 내에서 톨-유사 수용체를 통해 면역 반응을 유도하는 것으로 이론화되었다. 일부 증거는 CpG-고갈된 AAV 벡터가 특정 상황 하에서 마우스에서 면역 검출을 회피한다는 것을 시사한다 (Faust et al., J. Clin.Invest. 2013; 123, 2994-3001). 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열 번호 1))은 20 개의 CpG 디뉴클레오티드를 함유한다.

[0117] 따라서, 일부 실시 양태에서, 본원에 제공된 핵산 조성물 (예를 들어, 코돈-변경된 폴리뉴클레오티드)은 인자 IX 코딩 서열에서 CpG 디뉴클레오티드의 수를 감소시키기 위해 코돈-변경된다. 예를 들어, CS02-FL-NA (서열 번호 5)는 CpG 디뉴클레오티드가 없고, CS03-FL-NA (서열 번호 6)는 CpG 디뉴클레오티드가 없으며, CS04-FL-NA (서열 번호 7)는 CpG 디뉴클레오티드가 없고, CS05-FL-NA (서열 번호 8)는 11 개의 CpG 디뉴클레오티드를 가지고, CS06-FL-NA (서열 번호 9)는 3 개의 CpG 디뉴클레오티드를 갖는다. 이들 작제물은 야생형 인자 IX 코딩 서열 및 더 많은 수의 CpG 디뉴클레오티드를 갖는 유사하게 코돈-변경된 서열보다 낮은 면역 원성 반응을 불법화해야 한다.

[0118] 따라서, 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열 (예를 들어, CS02-CS06 인자 IX 코딩 서열 중 하나와 높은 서열 동일성을 갖는 폴리뉴클레오티드)은 20개 미만의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 15개 미만의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 12 개 미만의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 10개 미만의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 5개 미만의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 3개 미만의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0119] 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 15개 이상의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다. 일부 실시 양태에서, 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드의 서열은 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1개 이하를 가지거나 CpG 디뉴클레오티드를 갖지 않는다.

[0120] A. 인자 IX 아미노산 치환

[0121] 본원에 기재된 인자 IX 작제물의 AAV-벡터 기반 발현의 효율을 추가로 증가시키기 위해, 일부 구현에 따라, 분비를 개선하고, 특이적 활성을 증가시키고/시키거나 인자 IX의 안정성을 향상시키는 것으로 알려진 아미노산 치환을 추가로 포함한다. FIX 활성의 혈장 수준을 증가시키는 다수의 잠재적인 인자 IX 변이체는 당업계에 공지되어 있다. 이들 변이체는 인자 IX 촉매 활성 (예를 들어, 과활성 돌연변이체)을 증가시키고, 안티트롬빈 III 및/또는 헤파린에 대한 내성을 증가시키고, 혈청 반감기를 증가시키고, 번역 후 변형의 패턴을 변경시키는 아미노산 치환을 포함한다.

[0122] 예를 들어, 잔기 R338 (PPE)의 돌연변이는 인자 IX의 응고 활성을 증가시킬 수 있다. 검토를 위해, 미국특허 제 6,531,298 호를 참조 (이의 내용은 모든 목적을 위해 그 전체가 참조로 포함됨). 미국특허 제 6,531,298호에 개시된 바와 같이, 이 위치에서 아르기닌에서 류신으로의 아미노산 치환은 인자 IX의 활성을 증가시킨다. 이는 *생체 내에서* 나중에 확인되었고, 여기서 R338L (PPE) 돌연변이는 생체 내에서 인자 IX 활성을 5 배에서 10 배로 증가시킨다. 검토를 위해, Simioni P. et al., N Engl J Med. 361(17):1671-75 (2009)를 참조 (그 전체가 본원에 참조로 포함됨). 따라서, 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 아르기닌 384 (PPI); 잔기 338 (PPE)에서 아미노산 치환을 갖는 인자 IX 폴리펩티드를 암호화한다. 특정 실시 양태에서, 아미

노산 치환은 R384L (PPI)이다. 다른 실시 양태에서, 잔기 384 (PPI)/338 (PPE)에서 류신 이외의 잔기로의 아미노산 치환.예를 들어, R384A (PPI) 아미노산 치환이 마우스에서 2 내지 6 배 더 높은 활성을 제공하는 것으로 보고되었다. Schuettrumpf J et al., Blood, 105(6):2316-23 (2005) (이의 내용은 모든 목적을 위해 그 전체가 참조로 명확히 포함됨).

[0123] 유사하게, 잔기 Y305, K311, S365 및 Y391의 돌연변이는 합성 기질에서 증가된 인자 IX 활성을 유발한다. 특히, K311M 및 K311T 단일 돌연변이는 합성 절단 기질에서 2.8 배 및 6.7 배 증가된 활성을 초래하였다. Sichler K. et al., J Biol Chem.278(6):4121-26 (2003) (다른 잔기 넘버링 사용). 또한, Y305F/K311T/Y391T 삼중 돌연변이체는 합성 기질 동일저자에서 7000 배 증가된 활성을 초래하였다. 따라서, 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 티로신 305 (PPI), 리신 311 (PPI) 및 티로신 391 (PPI) 중 하나 이상에서 아미노산 치환을 갖는 인자 IX 폴리펩티드를 암호화한다. 특정 실시 양태에서, 아미노산 치환은 K311M (PPI)이다. 특정 실시 양태에서, 아미노산 치환은 K311T (PPI)이다. 다른 특정 실시 양태에서, 아미노산 치환은 Y305F/K311T/Y391T (PPI)이다.

[0124] 개선된 특성을 제공하는 다른 아미노산 치환은 당업계에 공지되어 있고, 기재된 코돈-변경된 인자 IX 폴리뉴클레오티드에 포함될 수 있다. 예를 들어 미국특허 제 8,778,870 호를 참조 (이의 내용은 모든 목적을 위해 그 전체가 본원에 명백히 참조로 포함됨).

[0125] B. 인자 IX 단백질을 코딩하는 코돈-변형된 폴리뉴클레오티드

[0126] *CS02 코돈 변형된 폴리뉴클레오티드*

[0127] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드를 인코딩하는 인자 IX 폴리뉴클레오티드(예들 들어, 코돈-최적화된 폴리뉴클레오티드)를 포함하되, 여기서 인자 IX 폴리뉴클레오티드는 CS02-FL-NA(서열번호 5)와 서열 동일성이 높은 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0128] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 95% 동일성을 갖는다 . 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 97% 동일성을 갖는다 . 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-FL-NA(서열번호 5)이다.

[0129] 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0130] 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50%

내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0131] 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0132] 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열)은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-FL-NA(서열번호 5)에 대한 서열 동일성이 높은 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0133] 일부 구현예에서, 코딩된 인자 IX 폴리펩티드, 예들 들어, CS02-FL-NA(서열번호 5)에 대한 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형 인자 IX 프리-프로-단백질 서열 FIX-FL-AA (서열번호 2) 및/또는 파두아 (hFIX(R384L)) 프리-프로-단백질 서열 FIXp-FL-AA (서열번호 4)과 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0134] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.5% 동일성을 갖는

다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)이다.

[0135] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)이다.

[0136] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드(예들 들어, 세린 프로테아제 활성을 갖는)를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하되, 여기서 인자 IX 폴리뉴클레오티드는 CS02-MP-NA에 대한 높은 서열 동일성을 갖는 뉴클레오티드 서열(서열번호 13)을 포함한다. 일부 구현예에서, CS02-MP-NA에 대한 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열 (서열번호 13)은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA에 대한 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열 (서열번호 13)은 야생형 인자 IX 코딩 서열와 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다 (FIX-FL-NA (서열번호 1)).

[0137] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 95% 동일성을 갖는다 (서열번호 13). 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 96% 동일성을 갖는다 (서열번호 13). 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 97% 동일성을 갖는다 (서열번호 13). 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 98% 동일성을 갖는다 (서열번호 13). 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 99% 동일성을 갖는다 (서열번호 13). 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 99.5% 동일성을 갖는다 (서열번호 13). 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA와 적어도 99.9% 동일성을 갖는다 (서열번호 13). 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS02-MP-NA이다 (서열번호 13).

[0138] 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열(서열번호 13)은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열(서열번호 13)은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열(서열번호 13)은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열(서열번호 13)은 54% 미만의 GC 함량을 갖는다.

[0139] 일부 구현예에서, CS02-MP-NA 와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열(서열번호 13)은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 50% 내지 58%의 GC 함량을 갖는

다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열 (서열번호 13)은 50% 내지 54%의 GC 함량을 갖는다.

[0140] 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0141] 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0142] 일부 구현예에서, CS02-MP-NA(서열번호 13)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-SP-AA(서열번호 37)의 아미노산 서열을 갖는 인자 IX 신호 펩티드를 인코딩하는 인자 IX 신호 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS03-SP-NA (서열번호 26)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0143] 일부 구현예에서, CS02-MP-NA(서열번호 13)에 대해 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PP-AA(서열번호 38)의 아미노산 서열을 갖는 인자 IX 프로-펩티드를 인코딩하는 인자 IX 프로-펩티드 폴리뉴클

레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0144] 일부 구현예에서, CS02-MP-NA(서열번호 13)에 대해 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PPP-AA(서열번호 36)의 아미노산 서열을 갖는 인자 IX 프리-프로-펩티드를 인코딩하는 인자 IX 프리-프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0145] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS02-FL-NA(서열번호 5)에 대해 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA(서열번호 12)과 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0146] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0147] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9%

동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0148] 일 구현예에서, 본원에 제공된 코돈-변형된 폴리뉴클레오티드는 경쇄, 중쇄, 및 경쇄의 C-말단을 중쇄의 N-말단에 연결하는 폴리펩티드 링커를 포함하는 단일-쇄 인자 IX 폴리펩티드를 인코딩한다. 인자 IX 폴리펩티드의 경쇄는 인자 IX 경쇄를 인코딩하는 CS02-FL-NA (서열번호 5)의 일부인, CS02-LC-NA(서열번호 42)와 높은 서열 동일성을 갖는 제1 뉴클레오티드 서열에 의해 인코딩된다. 인자 IX 폴리펩티드의 경쇄는 인자 IX 중쇄를 인코딩하는 CS02-FL-NA (서열번호 5)의 일부인, CS02-HC-NA (서열번호 41)와 높은 서열 동일성을 갖는 제2 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 생체 내에서 성숙을 가능하게 하는 인자 XI 절단 부위를 포함한다(예를 들어, 전구체 단일 쇄 인자 IX 폴리펩티드의 발현 후).

[0149] 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)와 적어도 95% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)와 적어도 96% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)와 적어도 97% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)와 적어도 98% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41) 적어도 99% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)와 적어도 99.5% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)와 적어도 99.9% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS02-LC-NA 및 CS02-HC-NA (서열번호 42 및 41)이다.

[0150] 일부 구현예에서, 인자 IX 작제물의 폴리펩티드 링커는 CS02-AP-NA (서열번호 57)와 높은 서열 동일성을 갖는 제3 뉴클레오티드 서열에 의해 인코딩되며, 이는 야생형 인자 IX 활성화 폴리펩티드를 인코딩하는 코돈-변형된 서열, 예들 들어, FIX-FL-AA (서열번호 2)의 아미노산 192-226이다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 80% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 90% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 95% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 96% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 97% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 98% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)와 적어도 99% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS02-AP-NA (서열번호 57)이다.

[0151] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드는 또한 신호 펩티드 (예들 들어, 인자 IX 신호 펩티드) 및/또는 프로-펩티드 (예들 들어, 인자 IX 프로-펩티드)를 포함한다. 일부 구현예에서, 신호 펩티드는 야생형 인자 IX 신호 펩티드 (FIX-SP-AA (서열번호 37))이다. 일부 구현예에서, 신호 펩티드는 CS02-SP-NA (서열번호 25)와 높은 서열 동일성(예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 프로-펩티드는 야생형 인자 IX 프로-펩티드 (FIX-PP-AA (서열번호 38))이다. 일부 구현예에서, 프로-펩티드 펩티드는 CS02-PP-NA (서열번호 31)와 높은 서열 동일성(예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다.

[0152] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS02-LC-NA (서열번호 42) 및 CS02-HC-NA (서열번호 41)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0153] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97%

동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0154] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예를 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0155] 일부 구현예에서, 도 1을 참조하면, 구조 A의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-MP-NA (서열번호 13)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-FL-NA (서열번호 5)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0156] 일부 구현예에서, 도 1을 참조하면, 구조 B의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-MP-NA (서열번호 13)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-FL-NA (서열번호 5)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0157] 일부 구현예에서, 도 1을 참조하면, 구조 C의 폴리뉴클레오티드 (예를 들어, 단일-가닥 폴리뉴클레오티드)를 포

함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-MP-NA (서열번호 13)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-FL-NA (서열번호 5)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0158] 일부 구현예에서, 도 1을 참조하면, 구조 D의 폴리뉴클레오티드 (예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-MP-NA (서열번호 13)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS02-FL-NA (서열번호 5)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0159] *CS03 코돈 변형된 폴리뉴클레오티드*

[0160] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0161] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-FL-NA (서열번호 6)이다.

[0162] 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동

일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0163] 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0164] 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0165] 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1 개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-FL-NA (서열번호 6)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0166] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS03-FL-NA (서열번호 6)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형 인자 IX 프리-프로-단백질 서열 FIX-FL-AA (서열번호 2) 및/또는 파두아 (hFIX(R384L)) 프리-프로-단백질 서열 FIXp-FL-AA (서열번호 4)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0167] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)이다.

[0168] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)이다.

[0169] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드 (예들 들어, 세린 프로테아제 활성을 갖는)를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS03-MP-NA (서열번호 14)와 높은 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0170] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14) 와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14) 와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14) 와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14) 와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14) 와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14) 와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS03-MP-NA (서열번호 14)이다.

[0171] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된

폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0172] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0173] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0174] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0175] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-SP-AA (서열번호 37)의 아미노산 서열을 갖는 인자 IX 신호 펩티드를 암호화하는 인자 IX 신호 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오

티드는 CS03-SP-NA (서열번호 26)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0176] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PP-AA(서열번호 38)의 아미노산 서열을 갖는 인자 IX 프로-펩티드를 인코딩하는 인자 IX 프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0177] 일부 구현예에서, CS03-MP-NA (서열번호 14)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PPP-AA의 아미노산 서열을 갖는 인자 IX 프리-프로-펩티드 (서열번호 36)를 인코딩하는 인자 IX 프리-프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0178] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS03-FL-NA (서열번호 6)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0179] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0180] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA

(서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0181] 일 구현예에서, 본원에 제공된 코돈-변형된 폴리뉴클레오티드는 경쇄, 중쇄, 및 경쇄의 C-말단을 중쇄의 N-말단에 연결하는 폴리펩티드 링커를 포함하는 단일-쇄 인자 IX 폴리펩티드를 인코딩한다. 인자 IX 폴리펩티드의 경쇄는 인자 IX 경쇄를 인코딩하는 CS03-FL-NA (서열번호 6)의 일부인 CS03-LC-NA (서열번호 44)와 높은 서열 동일성을 갖는 제1 뉴클레오티드 서열에 의해 인코딩된다. 인자 IX 폴리펩티드의 중쇄는 인자 IX 중쇄를 인코딩하는 CS03-FL-NA (서열번호 6)의 일부인 CS03-HC-NA (서열번호 43)와 높은 서열 동일성을 갖는 제2 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 생체 내에서 성숙을 가능하게 하는 인자 XI 절단 부위를 포함한다(예를 들어, 전구체 단일 쇄 인자 IX 폴리펩티드의 발현 후).

[0182] 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 95% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 96% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 97% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 98% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 99% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 99.5% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 적어도 99.9% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)이다.

[0183] 일부 구현예에서, 인자 IX 작제물의 폴리펩티드 링커는 CS03-AP-NA (서열번호 58)와 높은 서열 동일성을 갖는 제3 뉴클레오티드 서열에 의해 인코딩되며, 이는 야생형 인자 IX 활성화 폴리펩티드, 예들 들어, FIX-FL-AA (서열번호 2)의 아미노산 192-226을 인코딩하는 코돈-변형된 서열이다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 80% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 90% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 95% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 96% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 97% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 98% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)와 적어도 99% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS03-AP-NA (서열번호 58)이다.

[0184] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드는 또한 신호 펩티드 (예들 들어, 인자 IX 신호 펩티드) 및/또는 프로-펩티드 (예들 들어, 인자 IX 프로-펩티드)를 포함한다. 일부 구현예에서, 신호 펩티드는 야생형 인자 IX 신호 펩티드 (FIX-SP-AA (서열번호 37))이다. 일부 구현예에서, 신호 펩티드는 CS03-SP-NA (서열번호 26)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 프로-펩티드는 야생형 인자 IX 프로-펩티드 (FIX-PP-AA (서열번호 38))이다. 일부 구현예에서, 프로-펩티드 펩티드는 CS03-PP-NA (서열번호 32)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다.

[0185] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS03-LC-NA (서열번호 44) 및 CS03-HC-NA (서열번호 43)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되

는 능력을 유지해야 한다 (예를 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0186] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0187] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예를 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0188] 일부 구현예에서, 도 1을 참조하면, 구조 A의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-MP-NA (서열번호 14)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-FL-NA (서열번호 6)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0189] 일부 구현예에서, 도 1을 참조하면, 구조 B의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-MP-NA (서열번호 14)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한

FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-FL-NA (서열번호 6)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0190] 일부 구현예에서, 도 1을 참조하면, 구조 C의 폴리뉴클레오티드 (예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물을 제공하며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-MP-NA (서열번호 14)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-FL-NA (서열번호 6)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0191] 일부 구현예에서, 도 1을 참조하면, 구조 D의 폴리뉴클레오티드(예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-MP-NA (서열번호 14)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-FL-NA (서열번호 6)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0192] *CS04 코돈 변형된 폴리뉴클레오티드*

[0193] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 가지는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 가지는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))과 비교하여 감소된 GC 함량을 가진다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 가지는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 수의 CpG 디뉴클레오티드를 가진다.

[0194] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-FL-NA (서열번호 7)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은

CS04-FL-NA (서열번호 7)이다.

[0195] 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0196] 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0197] 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0198] 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성

을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-FL-NA (서열번호 7)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0199] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS04-FL-NA (서열번호 7) 와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형 인자 IX 프리-프로-단백질 서열 FIX-FL-AA (서열번호 2) 및/또는 파두아 (hFIX(R384L)) 프리-프로-단백질 서열 FIXp-FL-AA (서열번호 4)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0200] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)이다.

[0201] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)이다.

[0202] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드 (예들 들어, 세린 프로테아제 활성을 갖음)를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0203] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-

MP-NA (서열번호 15)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS04-MP-NA (서열번호 15)이다.

[0204] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0205] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 50% 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0206] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0207] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드

의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0208] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-SP-AA (서열번호 37)의 아미노산 서열을 갖는 인자 IX 신호 펩티드를 인코딩하는 인자 IX 신호 폴리뉴클레오티드를 추가로 포함한다일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS03-SP-NA (서열번호 26)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0209] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PP-AA(서열번호 38)의 아미노산 서열을 갖는 인자 IX 프로-펩티드를 인코딩하는 인자 IX 프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0210] 일부 구현예에서, CS04-MP-NA (서열번호 15)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PPP-AA(서열번호 36)의 아미노산 서열을 갖는 인자 IX 프리-프로-펩티드를 인코딩하는 인자 IX 프리-프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0211] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS04-FL-NA (서열번호 7) 와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0212] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)

와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0213] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0214] 일 구현예에서, 본원에 제공된 코돈-변형된 폴리뉴클레오티드는 경쇄, 중쇄, 및 경쇄의 C-말단을 중쇄의 N-말단에 연결하는 폴리펩티드 링커를 포함하는 단일-쇄 인자 IX 폴리펩티드를 인코딩한다. 인자 IX 폴리펩티드의 경쇄는 인자 IX 경쇄를 인코딩하는 CS04-FL-NA (서열번호 7)의 일부인, CS04-LC-NA (서열번호 46)와 높은 서열 동일성을 갖는 제1 뉴클레오티드 서열에 의해 인코딩된다. 인자 IX 폴리펩티드의 중쇄는 인자 IX 중쇄를 인코딩하는 CS04-FL-NA (서열번호 7)의 일부인 CS04-HC-NA (서열번호 45)와 높은 서열 동일성을 갖는 제2 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 생체 내에서 성숙을 가능하게 하는 인자 XI 절단 부위를 포함한다(예를 들어, 전구체 단일 쇄 인자 IX 폴리펩티드의 발현 후).

[0215] 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 95% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 96% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 97% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 98% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 99% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 99.5% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)와 적어도 99.9% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS04-LC-NA 및 CS04-HC-NA (서열번호 46 및 45)이다.

[0216] 일부 구현예에서, 인자 IX 작제물의 폴리펩티드 링커는 야생형 인자 IX 활성화 폴리펩티드, 예들 들어, FIX-FL-AA (서열번호 2)의 아미노산 192-226을 인코딩하는 코돈-변형된 서열인, CS04-AP-NA (서열번호 59)와 높은 서열 동일성을 갖는 제3 뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)와 적어도 80% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)와 적어도 90% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 5 9)와 적어도 95% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)와 적어도 96% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)와 적어도 97% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)와 적어도 98% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)와 적어도 99% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS04-AP-NA (서열번호 59)이다.

[0217] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드는 또한 신호 펩티드 (예들 들어, 인자 IX 신호 펩티드) 및/또는 프로-펩티드 (예들 들어, 인자 IX 프로-펩티드)를 포함한다. 일부 구현예에서, 신호 펩티드는 야생형 인자 IX 신호 펩티드 (FIX-SP-AA (서열번호 37))이다. 일부 구현예에서, 신호 펩티드는 CS04-SP-NA (서열번호 27)와 높

은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드에 의해 인코딩된다. 일부 구현예에서, 프로-펩티드는 야생형 인자 IX 프로-펩티드 (FIX-PP-AA (서열번호 38))이다. 일부 구현예에서, 프로-펩티드 펩티드는 CS04-PP-NA (서열번호 33)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다.

[0218] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS04-LC-NA (서열번호 46) 및 CS04-HC-NA (서열번호 45)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0219] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0220] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0221] 일부 구현예에서, 도 1을 참조하면, 구조 A의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-MP-NA (서열번호 15)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 갖는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-FL-NA (서열번호 7)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인

자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0222] 일부 구현예에서, 도 1을 참조하면, 구조 B의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-MP-NA (서열번호 15)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-FL-NA (서열번호 7)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0223] 일부 구현예에서, 도 1을 참조하면, 구조 C의 폴리뉴클레오티드 (예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-MP-NA (서열번호 15)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-FL-NA (서열번호 7)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0224] 일부 구현예에서, 도 1을 참조하면, 구조 D의 폴리뉴클레오티드(예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-MP-NA (서열번호 15)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS04-FL-NA (서열번호 7)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0225] *CS05 코돈 변형된 폴리뉴클레오티드*

[0226] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0227] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 96% 동일성을 갖는다. .특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-FL-NA (서열번호 8)이다.

[0228] 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0229] 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0230] 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0231] 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서

열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-FL-NA (서열번호 8)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0232] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS05-FL-NA (서열번호 8)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형 인자 IX 프리-프로-단백질 서열 FIX-FL-AA (서열번호 2) 및/또는 파두아 (hFIX(R384L)) 프리-프로-단백질 서열 FIXp-FL-AA (서열번호 4)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0233] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)이다.

[0234] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)이다.

[0235] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드 (예들 들어, 세린 프로테아제 활성을 갖는)를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 수의 CpG 디뉴클레오티드를

갖는다.

[0236] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16) 와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16) 와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16) 와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16) 와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16) 와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16) 와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS05-MP-NA (서열번호 16)이다.

[0237] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0238] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0239] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0240] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는

다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0241] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-SP-AA (서열번호 37)의 아미노산 서열을 갖는 인자 IX 신호 펩티드를 인코딩하는 인자 IX 신호 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS03-SP-NA (서열번호 26)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0242] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PP-AA(서열번호 38)의 아미노산 서열을 갖는 인자 IX 프로-펩티드를 인코딩하는 인자 IX 프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0243] 일부 구현예에서, CS05-MP-NA (서열번호 16)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PPP-AA(서열번호 36)의 아미노산 서열을 갖는 인자 IX 프리-프로-펩티드를 인코딩하는 인자 IX 프리-프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0244] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예를 들어, CS05-FL-NA (서열번호 8)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예를 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0245] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을

갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0246] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0247] 일 구현예에서, 본원에 제공된 코돈-변형된 폴리뉴클레오티드는 경쇄, 중쇄, 및 경쇄의 C-말단을 중쇄의 N-말단에 연결하는 폴리펩티드 링커를 포함하는 단일-쇄 인자 IX 폴리펩티드를 인코딩한다. 인자 IX 폴리펩티드의 경쇄는 인자 IX 경쇄를 인코딩하는 CS05-FL-NA (서열번호 8)의 일부인 CS05-LC-NA (서열번호 48)와 높은 서열 동일성을 갖는 제1 뉴클레오티드 서열에 의해 인코딩된다. 인자 IX 폴리펩티드의 중쇄는 인자 IX 중쇄를 인코딩하는 CS05-FL-NA (서열번호 8)의 일부인 CS05-HC-NA (서열번호 47)와 높은 서열 동일성을 갖는 제2 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 생체 내에서 성숙을 가능하게 하는 인자 XI 절단 부위를 포함한다 (예를 들어, 전구체 단일 쇄 인자 IX 폴리펩티드의 발현 후).

[0248] 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47)와 적어도 95% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47) 와 적어도 96% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47), 와 적어도 97% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47) 와 적어도 98% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47) 와 적어도 99% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47), 와 적어도 99.5% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47)와 적어도 99.9% 서열 동일성을 가진다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS05-LC-NA 및 CS05-HC-NA (서열번호 48 및 47)이다.

[0249] 일부 구현예에서, 인자 IX 작제물의 폴리펩티드 링커는 야생형 인자 IX 활성화 폴리펩티드, 예들 들어, FIX-FL-AA (서열번호 2)의 아미노산 192-226을 인코딩하는 코돈-변형된 서열인, CS05-AP-NA (서열번호 60)와 높은 서열 동일성을 갖는 제3 뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60)와 적어도 80% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60) 와 적어도 90% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호

60). 와 적어도 95% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60). 와 적어도 96% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60) 와 적어도 97% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60) 와 적어도 98% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60) 와 적어도 99% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS05-AP-NA (서열번호 60)이다.

[0250] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드는 또한 신호 펩티드 (예들 들어, 인자 IX 신호 펩티드) 및/또는 프로-펩티드 (예들 들어, 인자 IX 프로-펩티드)를 포함한다. 일부 구현예에서, 신호 펩티드는 야생형 인자 IX 신호 펩티드 (FIX-SP-AA (서열번호 37))이다. 일부 구현예에서, 신호 펩티드는 CS05-SP-NA (서열번호 28)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 프로-펩티드는 야생형 인자 IX 프로-펩티드 (FIX-PP-AA (서열번호 38))이다. 일부 구현예에서, 프로-펩티드 펩티드는 CS05-PP-NA (서열번호 34)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다.

[0251] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS05-LC-NA (서열번호 48) 및 CS05-HC-NA (서열번호 47)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0252] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0253] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0254] 일부 구현예에서, 도 1을 참조하면, 구조 A의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS03-MP-NA (서열번호 14)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열

번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-FL-NA (서열번호 8)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0255] 일부 구현예에서, 도 1을 참조하면, 구조 B의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-MP-NA (서열번호 16)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-FL-NA (서열번호 8)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0256] 일부 구현예에서, 도 1을 참조하면, 구조 C의 폴리뉴클레오티드 (예를 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물을 포함하며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-MP-NA (서열번호 16)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-FL-NA (서열번호 8)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0257] 일부 구현예에서, 도 1을 참조하면, 구조 D의 폴리뉴클레오티드(예를 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-MP-NA (서열번호 16)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS05-FL-NA (서열번호 8)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0258] *CS06 코돈 변형된 폴리뉴클레오티드*

[0259] 일 구현예에서, 본원에 제공된 핵산 조성물 단일-쇄 인자 IX 폴리펩티드를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0260] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-FL-NA (서열번호 9)이다.

[0261] 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0262] 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0263] 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0264] 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은

15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-FL-NA (서열번호 9)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0265] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS06-FL-NA (서열번호 9)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형 인자 IX 프리-프로-단백질 서열 FIX-FL-AA (서열번호 2) 및/또는 파두아 (hFIX(R384L)) 프리-프로-단백질 서열 FIXp-FL-AA (서열번호 4)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0266] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-FL-AA (서열번호 2)이다.

[0267] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-FL-AA (서열번호 4)이다.

[0268] 일 구현예에서, 본원에 제공된 핵산 조성물은 단일-쇄 인자 IX 폴리펩티드 (예들 들어, 세린 프로테아제 활성을 갖음)를 인코딩하는 인자 IX 폴리뉴클레오티드 (예들 들어, 코돈-변형된 폴리뉴클레오티드)를 포함하며, 여기서 인자 IX 폴리뉴클레오티드는 CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 포함한다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드의 뉴클레오티드 서열은 야생형 인자 IX 코딩 서열 (FIX-FL-NA (서열번호 1))와 비교하여 감소된 수의 CpG 디뉴클레오티드를 갖는다.

[0269] 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 95% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 96% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 97% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 98% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 99% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 99.5% 동일성을 갖는다. 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)와 적어도 99.9% 동일성을 갖는다. 또 다른 특정 구현예에서, 코돈-변형된 폴리뉴클레오티드의 서열은 CS06-MP-NA (서열번호 17)이다.

[0270] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 60% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 59% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 58% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 57% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 56% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 55% 미만의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 54% 미만의 GC 함량을 갖는다.

[0271] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 60%의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 59%의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 58%의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 57%의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 56%의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 55%의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 50% 내지 54%의 GC 함량을 갖는다.

[0272] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±1.0의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.8의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.6의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.5의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.4의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.3의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.2의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%±0.1의 GC 함량을 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 53.8%의 GC 함량을 갖는다.

[0273] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 15개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 12개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 10개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 9개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 8개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 7개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 6개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 5개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 4개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 3개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 2개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 1개 이하의 CpG 디뉴클레오티드를 갖는다. 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 코돈-변형된 폴리뉴클레오티드의 서열은 CpG 디뉴클레오티드를 갖지 않는다.

[0274] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-SP-AA (서열번호 37)의 아미노산 서열을 갖는 인자 IX 신호 펩티드를 인코딩하는 인자 IX 신호 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS03-SP-NA (서열번호 26)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 신호 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0275] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PP-AA (서열번호 38)의 아미노산 서열을 갖는 인자 IX 프로-펩티드를 인코딩하는 인자 IX 프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프로-펩티드 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다.

[0276] 일부 구현예에서, CS06-MP-NA (서열번호 17)와 높은 서열 동일성을 갖는 인자 IX 폴리뉴클레오티드는 FIX-PPP-AA (서열번호 36)의 아미노산 서열을 갖는 인자 IX 프리-프로-펩티드를 인코딩하는 인자 IX 프리-프로-펩티드 폴리뉴클레오티드를 추가로 포함한다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 핵산 서열을 갖는다. 일부 구현예에서, 인자 IX 프리-프로-펩티드 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는

100% 동일한 핵산 서열을 갖는다.

[0277] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS06-FL-NA (서열번호 9)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0278] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0279] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0280] 일 구현예에서, 본원에 제공된 코돈-변형된 폴리뉴클레오티드는 경쇄, 중쇄, 및 경쇄의 C-말단을 중쇄의 N-말단에 연결하는 폴리펩티드 링커를 포함하는 단일-쇄 인자 IX 폴리펩티드를 인코딩한다. 인자 IX 폴리펩티드의 경쇄는 인자 IX 경쇄를 인코딩하는 CS06-FL-NA (서열번호 9)의 일부인 CS06-LC-NA (서열번호 50)와 높은 서열 동일성을 갖는 제1 뉴클레오티드 서열에 의해 인코딩된다. 인자 IX 폴리펩티드의 중쇄는 인자 IX 중쇄를 인코딩하는 CS06-FL-NA (서열번호 9)의 일부인 CS06-HC-NA (서열번호 49)와 높은 서열 동일성을 갖는 제2 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 생체 내에서 성숙을 가능하게 하는 인자 XI 절단 부위를 포함한다 (예를 들어, 전구체 단일 쇄 인자 IX 폴리펩티드의 발현 후).

[0281] 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 95% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 96% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 97% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 98% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 99% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 99.5% 서열 동일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)와 적어도 99.9% 서열 동

일성을 갖는다. 일부 구현예에서, 제1 및 제2 뉴클레오티드 서열은 각각 CS06-LC-NA 및 CS06-HC-NA (서열번호 50 및 49)이다.

[0282] 일부 구현예에서, 인자 IX 작제물의 폴리펩티드 링커는 야생형 인자 IX 활성화 폴리펩티드, 예들 들어, FIX-FL-AA (서열번호 2)의 아미노산 192-226을 인코딩하는 코돈-변형된 서열인 CS06-AP-NA (서열번호 61)와 높은 서열 동일성을 갖는 제3 뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61)와 적어도 80% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61) 와 적어도 90% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61) 와 적어도 95% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61) 와 적어도 96% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61) 와 적어도 97% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61) 와 적어도 98% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61) 와 적어도 99% 동일성을 갖는다. 일부 구현예에서, 제3 뉴클레오티드 서열은 CS06-AP-NA (서열번호 61)이다.

[0283] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드는 또한 신호 펩티드 (예들 들어, 인자 IX 신호 펩티드) 및/또는 프로-펩티드 (예들 들어, 인자 IX 프로-펩티드)를 포함한다. 일부 구현예에서, 신호 펩티드는 야생형 인자 IX 신호 펩티드 (FIX-SP-AA (서열번호 37))이다. 일부 구현예에서, 신호 펩티드는 CS06-SP-NA (서열번호 29)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다. 일부 구현예에서, 프로-펩티드는 야생형 인자 IX 프로-펩티드 (FIX-PP-AA (서열번호 38))이다. 일부 구현예에서, 프로-펩티드 펩티드는 CS06-PP-NA (서열번호 35)와 높은 서열 동일성 (예들 들어, 적어도 95%, 96%, 97%, 98%, 또는 99%)을 갖는 코돈-변형된 폴리뉴클레오티드 서열에 의해 인코딩된다.

[0284] 일부 구현예에서, 인코딩된 인자 IX 폴리펩티드, 예들 들어, CS06-LC-NA (서열번호 50) 및 CS06-HC-NA (서열번호 49)와 높은 서열 상동성을 갖는 폴리뉴클레오티드에 의해 인코딩된 폴리펩티드는, 야생형, 성숙 인자 IX 단일-쇄 폴리펩티드 서열 FIX-MP-AA (서열번호 10) 및/또는 성숙 파두아 (hFIX(R384L)) 단일-쇄 서열 FIXp-MP-AA (서열번호 12)와 높은 서열 동일성을 갖는다. 인코딩된 인자 IX 폴리펩티드는 기능 인자 IXa 단백질로 활성화되는 능력을 유지해야 한다 (예들 들어, 임의의 신호 펩티드 및 프로-펩티드의 제거, 및 활성화 폴리펩티드의 절단에 의해).

[0285] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 85% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 90% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 95% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 96% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 97% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 98% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.5% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)와 적어도 99.9% 동일성을 갖는다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIX-MP-AA (서열번호 10)이다.

[0286] 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 85% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384 (예들 들어, 성숙 인자 IX 단일-쇄 폴리펩티드 FIXp-MP-AA (서열번호 12)의 위치 338)에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 90% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 95% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 96% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 97% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 98% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.5% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신

을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)와 적어도 99.9% 동일성을 가지고 프리-프로-폴리펩티드의 위치 384에서 류신을 포함한다. 일 구현예에서, 인코딩된 인자 IX 폴리펩티드의 서열은 FIXp-MP-AA (서열번호 12)이다.

[0287] 일부 구현예에서, 도 1을 참조하면, 구조 A의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-MP-NA (서열번호 17)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-FL-NA (서열번호 9)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0288] 일부 구현예에서, 도 1을 참조하면, 구조 B의 자가-상보 폴리뉴클레오티드를 포함하는 핵산 조성물이 제공되며 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-MP-NA (서열번호 17)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-FL-NA (서열번호 9)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0289] 일부 구현예에서, 도 1을 참조하면, 구조 C의 폴리뉴클레오티드 (예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물을 포함하며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-MP-NA (서열번호 17)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는, 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-FL-NA (서열번호 9)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0290] 일부 구현예에서, 도 1을 참조하면, 구조 D의 폴리뉴클레오티드(예들 들어, 단일-가닥 폴리뉴클레오티드)를 포함하는 핵산 조성물이 제공되며, 여기서 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-MP-NA (서열번호 17)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는 성숙 인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 또한 FIX-SP-NA (서열번호 24), CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), 및 CS06-SP-NA (서열번호 29) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 신호 펩티드를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의

FIX 코딩 서열 부분은 또한 FIX-PP-NA (서열번호 30), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), 및 CS06-PP-NA (서열번호 35) 중 하나와 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 가지는 인자 IX 프로-펩티드 (임의로 전술한 인자 IX 신호 펩티드에 대한 핵산 서열과 조합하여)를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오티드의 FIX 코딩 서열 부분은 CS06-FL-NA (서열번호 9)와 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 동일성을 가지는 프리-프로-인자 IX 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.

[0291] C. 코돈-변형된 인자 IX 신호 및 프로-펩티드

[0292] 일 측면에서, 개시내용은 인자 IX 신호 펩티드, 인자 IX 프로-펩티드, 및 둘다 (예들 들어, 인자 IX 프리-프로-펩티드)를 인코딩하는 코돈-변형된 폴리뉴클레오티드를 제공한다. 이들 코돈-변형된 폴리뉴클레오티드는 인자 IX 발현을 개선하고 예들 들어, 인자 IX 단일-쇄 폴리펩티드 (예들 들어, 인자 IX 경쇄, 활성화 펩티드, 및 중쇄)를 인코딩하는 코돈-변형되거나 달리 폴리뉴클레오티드의 업스트림에 위치될 수 있다. 일반적으로, 인코딩된 펩티드는 야생형 인자 IX 신호 펩티드 (예들 들어, FIX-SP-AA (서열번호 37)), 프로-펩티드 (예들 들어, FIX-PP-AA (서열번호 38), 및 프리-프로-펩티드 (FIX-PPP-AA (서열번호 36))이다.

[0293] 특정 구현예에서, 인자 IX 신호 펩티드, 프로-펩티드, 및 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 높은 동일성을 갖는 서열 (예들 들어, CS02-SP-NA (서열번호 25), CS03-SP-NA (서열번호 26), CS04-SP-NA (서열번호 27), CS05-SP-NA (서열번호 28), CS06-SP-NA (서열번호 29), CS02-PP-NA (서열번호 31), CS03-PP-NA (서열번호 32), CS04-PP-NA (서열번호 33), CS05-PP-NA (서열번호 34), CS06-PP-NA (서열번호 35), CS02-PPP-NA (서열번호 19), CS03-PPP-NA (서열번호 20), CS04-PPP-NA (서열번호 21), CS05-PPP-NA (서열번호 22), 및 CS06-PPP-NA (서열번호 23) 중 하나와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%)을 갖는다.

[0294] *CS02 신호 및 프로-펩티드*

[0295] 일 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS02-SP-NA (서열번호 25)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0296] 일 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS02-PP-NA (서열번호 31)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0297] 일 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS02-PPP-NA (서열번호 19)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0298] *CS03 신호 및 프로-펩티드*

[0299] 일 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS03-SP-NA (서열번호 26)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS03-SP-NA (서열번호 26)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0300] 일 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS03-PP-NA (서열번호 32)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0301] 일 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS03-PPP-NA (서열번호 20)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0302] *CS04 신호 및 프로-펩티드*

[0303] 일 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴

클레오티드는 CS04-SP-NA (서열번호 27)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0304] 일 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS04-PP-NA (서열번호 33)을 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0305] 일 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS04-PPP-NA (서열번호 21)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0306] *CS05 신호 및 프로-펩티드*

[0307] 일 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS05-SP-NA (서열번호 28)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0308] 일 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS05-PP-NA (서열번호 34)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0309] 일 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS05-PPP-NA (서열번호 22)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0310] *CS06 신호 및 프로-펩티드*

[0311] 일 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 신호 펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS06-SP-NA (서열번호 29)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0312] 일 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS06-PP-NA (서열번호 35)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0313] 일 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 95% 서열 동일성을 갖는다. 다른 구현예에서, 인자 IX 프리-프로-펩티드를 인코딩하는 코돈-변형된 폴리뉴클레오티드는 CS06-PPP-NA (서열번호 23)와 적어도 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는다.

[0314] IV. 인자 IX 발현 벡터

[0315] 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 발현 벡터에 통합된다. 당업자에게 이해될 바와 같이, 많은 형태의 벡터가 본원에 개시된 코돈-변경된 인자 IX 폴리뉴클레오티드 서열을 사용하여 인자 IX 유전자 요법을 발효시키는데 사용될 수 있다. 발현 벡터의 비제한적인 예는 바이러스 벡터 (예를 들어, 유전자 요법에 적합한 벡터), 플라스미드 벡터, 박테리오파지 벡터, 코스미드, 파지미드, 인공 염색체 등을 포함한다.

[0316] 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 바이러스 유전자 요법 벡터에 통합된다. 바이러스 벡터의 비제한적 예는 하기를 포함한다: 레트로바이러스, 예를 들어, 몰로니 뮤린 백혈병 바이러스 (MMLV), 하비 뮤린 육종 바이러스, 뮤린 유선 종양 바이러스 및 라우스 육종 바이러스; 아데노 바이러스, 아데노 관련 바이러스; SV40 형 바이러스; 폴리오마바이러스; 엡스타인-바(Epstein-Barr) 바이러스; 유두종 바이러스; 헤르페스 바이러스; 백시니아 바이러스; 및 소아마비 바이러스.

[0317] 생체 내에서, 인자 IX는 주로 간에서 합성된다. 이와 같이, 간세포는 인자 IX 유전자 요법 작제물에 적합한 숙주 세포로서 표적화되어 왔다. 레트로바이러스 벡터 (예를 들어, Axelrod et al., 1990; Kay et al., 1992; Van den Driessche et al., 1999 및 Xu et al., 2003, 2005를 참조, 이의 개시는 그 전체가 모든 목적을 위해 참조로 명백히 포함됨), 렌티바이러스 (예를 들어, Ward et al., 2011, Brown et al., 2007 및 Matrai et al., 2011을 참조, 이의 개시는 그 전체가 모든 목적을 위해 참조로 명백히 포함됨), 아데노 관련 바이러스 (AAV)

(예를 들어, Herzog et al., 1999, 이의 개시는 그 전체가 모든 목적을 위해 참조로 명백히 포함됨), 아데노바이러스 벡터 (예를 들어, Brown et al., 2004 및 Ehrhardt & Kay, 2002를 참조, 이의 개시는 그 전체가 모든 목적을 위해 참조로 명백히 포함됨)를 포함하는 여러 종류의 바이러스 벡터가 유전자 요법 작제물의 간-표적 전달에 적합한 것으로 나타났다.

[0318] 일부 실시 양태에서, 유전자 요법 벡터는 레트로바이러스, 특히 복제-결핍 레트로바이러스이다. 복제-결핍 레트로바이러스의 생산을 위한 프로토콜은 당업계에 공지되어 있다. 검토를 위해 Kriegler, M., Gene Transfer and Expression, A Laboratory Manual, W.H.Freeman Co., New York (1990) 및 Murry, E.J., Methods in Molecular Biology, Vol. 7, Humana Press, Inc., Cliffton, N.J. (1991) 를 참조.

[0319] 한 실시 양태에서, 유전자 요법 벡터는 아데노-관련 바이러스 (AAV) 기반 유전자 요법 벡터이다. AAV 시스템은 이전에 기술되었고 일반적으로 당업계에 잘 알려져 있다 (Kelleher and Vos, Biotechniques, 17(6):1110-17 (1994); Cotten et al., P.N.A.S.U.S.A., 89(13):6094-98 (1992); Curiel, Nat Immun, 13(2-3):141-64 (1994); Muzyczka, Curr Top Microbiol Immunol, 158:97-129 (1992); 및 Asokan A, et al., Mol. Ther., 20(4):699-708 (2012), 각각은 모든 목적을 위해 그 전체가 본원에 참조로 포함됨). rAAV 벡터의 생성 및 사용에 관한 세부 사항은 예를 들어 미국특허 제 5,139,941 호 및 제 4,797,368호에 기술되어 있고, 각각은 모든 목적을 위해 그 전체가 본원에 참조로 포함된다. 특정 실시 양태에서, AAV 벡터는 AAV-8 벡터이다.

[0320] 간-특이적 인자 IX 발현을 위한 예시적인 AAV 전달 벡터는 WO 2009/130208에 기술되어 있으며, 이의 내용은 모든 목적을 위해 그 전체가 본원에 참조로 분명하게 포함된다. 벡터는 인간 인자 IX를 암호화하는 단일 가닥 AAV 벡터이며, 인자 cDNA를 구동하는 TTR Serp 조절 서열을 포함한다. 벡터는 또한 인간 인자 IX 유전자의 인트론 I 및 폴리아데닐화 신호를 포함한다.

[0321] 일부 실시 양태에서, 본원에 기술된 코돈-변경된 폴리뉴클레오티드는 레트로바이러스 발현 벡터에 통합된다. 이들 시스템은 전술되었으며, 당업계에 일반적으로 잘 알려져 있다 (Mann *et al.*, *Cell*, 33:153-159, 1983; Nicolas and Rubinstein, In:Vectors:A survey of molecular cloning vectors and their uses, Rodriguez and Denhardt, eds., Stoneham:Butterworth, pp. 494-513, 1988; Temin, In:Gene Transfer, Kucherlapati (ed.),New York:Plenum Press, pp. 149-188, 1986). 특정 실시 양태에서, 레트로바이러스 벡터는 렌티 바이러스 벡터이다 (예를 들어, Naldini et al., Science, 272(5259):263-267, 1996; Zufferey et al., Nat Biotechnol, 15(9):871-875, 1997; Blomer et al., J Virol., 71(9):6641-6649, 1997; 미국특허 제 6,013,516 호 및 제5,994,136호를 참조).

[0322] 일부 실시 양태에서, 본원에 기재된 코돈-변경된 폴리뉴클레오티드는 비바이러스 방법에 의해 대상체에 투여될 수 있다. 예를 들어, 네이키드 DNA는 전기 천공, 초음파천공, 입자 충격 또는 유체역학 전달에 의해 세포로 투여될 수 있다. DNA는 또한 중합체, 예를 들어 리포좀, 폴리솜, 폴리플레제(polypleses), 덴드리머와 함께 캡슐화되거나 커플링되어 복합체로서 대상체에 투여될 수 있다. 마찬가지로, DNA는 무기 나노 입자, 예를 들어 금, 실리카, 산화철 또는 인산 칼슘 입자에 커플링되거나 생체 내 세포로 전달하기 위해 세포-침투 펩티드에 부착될 수 있다.

[0323] 코돈-변경된 인자 IX 코딩 폴리뉴클레오티드는 또한 인공 염색체 발현 (ACE)과 같은 인공 염색체에 통합될 수 있다 (예를 들어, Lindenbaum et al., Nucleic Acids Res., 32(21):e172 (2004) 및 포유류 인공 염색체 (MAC)를 참조). 검토를 위해, 예를 들어, Perez-Luz and Diaz-Nido, J Biomed Biotechnol.2010:Article ID 642804 (2010)을 참조.

[0324] 진핵 생물 및 원핵 생물 발현 벡터를 포함하는, 세포 배양에서 코돈-변경된 폴리펩티드로부터 인자 IX 폴리펩티드의 발현을 위해 매우 다양한 벡터가 사용될 수 있다. 특정 실시 양태에서, 세포 배양에서 인자 IX 폴리펩티드를 발현시키는 데 사용하기 위해 플라스미드 벡터가 고려된다. 일반적으로, 숙주 세포와 호환되는 종으로부터 유래된 레플리콘 및 제어 서열을 함유하는 플라스미드 벡터가 이들 숙주와 관련하여 사용된다. 벡터는 복제 부위 뿐만 아니라 형질전환된 세포에서 표현형 선택을 제공할 수 있는 마킹 서열을 보유할 수 있다. 플라스미드는 하나 이상의 대조군 서열, 예를 들어 프로모터에 작동 가능하게 연결된 인자 IX 폴리펩티드를 암호화하는 코돈-변경된 폴리뉴클레오티드를 포함할 것이다.

[0325] 원핵 생물 발현을 위한 벡터의 비제한적인 예는 pRSET, pET, pBAD 등과 같은 플라스미드를 포함하고, 원핵 생물 발현 벡터에 사용되는 프로모터는 lac, trc, trp, recA, araBAD 등을 포함한다. 진핵 생물 발현을 위한 벡터의 예는 하기를 포함한다: (i) 효모에서의 발현을 위해, AOX1, GAP, GAL1, AUG1 등과 같은 프로모터를 사용하는

pAO, pPIC, pYES, pMET와 같은 벡터; (ii) 곤충 세포에서의 발현을 위해, PH, p10, MT, Ac5, OpIE2, gp64, polh 등과 같은 프로모터를 사용하는, pMT, pAc5, pIB, pMIB, pBAC 등과 같은 벡터, 및 (iii) 포유류 세포에서의 발현을 위해, pSVL, pCMV, pRc/RSV, pcDNA3, pBPV 등과 같은 벡터, 및 CMV, SV40, EF-1, UbC, RSV, ADV, BPV 및 β-액틴과 같은 프로모터를 사용하는 백시니아 바이러스, 아데노 관련 바이러스, 헤르페스 바이러스, 레트로바이러스 등과 같은 바이러스 시스템으로부터 유래된 벡터.

[0326] 일부 실시 양태에서, 본 개시 내용은 본원에 기재된 바와 같은 코돈-변경된 인자 IX 폴리뉴클레오티드, 벡터의 5' 및 3' 말단의 내부 말단 반복 (ITR) 서열, 인자 IX 폴리뉴클레오티드에 작동 가능하게 연결된 하나 이상의 프로모터 및/또는 인핸서 서열, 및 인자 IX 폴리뉴클레오티드 서열의 3' 말단에 이은 폴리아데닐화 신호를 포함하는 AAV 유전자 요법 벡터를 제공한다. 일부 실시 양태에서, 하나 이상의 프로모터 및/또는 인핸서 서열은 간-특이적 조절 제어 요소의 하나 이상의 카피를 포함한다.

[0327] 도 1은 일부 구현에 따른, 인자 IX 유전자 요법 벡터에 대한 몇몇 예시적인 구조를 도시한다. 도 1a는 돌연변이된 5' ITR, 절단된 TTR 인핸서/프로모터 서열, MVM 바이러스 인트론 서열, 코돈-변경된 인자 IX 코딩 서열, 폴리아데닐화 서열, 및 3'-ITR을 갖는 자기 상보적 AAV 벡터를 도시한다. 도 1b는 도 1a와 유사하지만 하나 이상의 (예를 들어, 하나, 둘, 셋 또는 그 이상의) 간-특이적 조절 제어 요소를 추가로 포함하는 인자 IX 폴리펩티드를 암호화하는 자기 상보적 AAV 벡터를 도시한다. 도 1c는 5'-ITR이 돌연변이되지 않고 자기 상보성을 방지하는 것만 제외하고 도 1a와 동일한 요소를 갖는 단일 가닥 벡터를 도시한다. 도 1d는 도 1a와 유사하지만 하나 이상의 (예를 들어, 하나, 둘, 셋 또는 그 이상의) 간-특이적 조절 제어 요소를 추가로 포함하는 인자 IX 폴리펩티드를 암호화하는 단일 가닥 AAV 벡터를 도시한다. 도 1에서 R384L '파두아' 아미노산 치환을 포함하는 인자 IX 단백질과 관련하여 도시되어 있지만, 일부 실시 양태에서, 도 1에 묘사된 바와 같은 일반적인 구조 (예를 들어, 구조 A, B, C 또는 D)를 갖는 인자 IX 뉴클레오티드 작제물은 R384L '파두아' 아미노산 치환을 포함하지 않는 인자 IX 단백질을 암호화한다.

[0328] 도 25는 AAV 인자 IX 유전자 치료 벡터 CS06-CRM8.3-ssV (서열 번호 40)의 뉴클레오티드 서열을 나타내며, 이는 도 1d에 도시된 유전자 치료 벡터 구조를 나타낸다. CS06-CRM8.3-ssV (서열 번호 40)의 뉴클레오티드 1-145는 AAV2 5'-ITR 서열 (서열 번호 51)이다. 5'-ITR 서열 다음에 뉴클레오티드 165-236, 238-309 및 311-382에서 간-특이적 CRM8 조절 제어 요소 CRM8 (서열 번호 39)의 3 개의 카피가 이어진다. CRM8 서열 다음에는 뉴클레오티드 383-712에서 절단된 TTR 인핸서/프로모터 서열 (서열 번호 52)이 존재한다. 다음으로, 벡터는 뉴클레오티드 724-800에서 마우스 미세 바이러스 (MVM) 인트론 (서열 번호 53)을 포함한다. 벡터의 뉴클레오티드 814-2199는 CS06 코돈-변경된 인자 IX (R384L) 코딩 서열 (CS06-FL-NA (서열 번호 9))이다. 뉴클레오티드 2208-2441에서 인자 IX 폴리뉴클레오티드 서열 다음에 BGH 폴리아데닐화 신호가 이어지고, 마지막으로 뉴클레오티드 2458-2602에서 AAV2 3'-ITR 서열 (서열 번호 55)이 이어진다.

[0329] 일부 실시 양태에서, 본 개시는 서열 번호 40의 뉴클레오티드 1-2602와 적어도 95%의 동일성을 갖는 서열을 포함하는 인자 IX 폴리뉴클레오티드를 제공한다. 일부 실시 양태에서, 본 개시는 서열 번호 40의 뉴클레오티드 1-2602와 적어도 99%의 동일성을 갖는 서열을 포함하는 인자 IX 폴리뉴클레오티드를 제공한다. 일부 실시 양태에서, 본 개시는 서열 번호 40의 뉴클레오티드 1-2602와 적어도99.5%의 동일성을 갖는 서열을 포함하는 인자 IX 폴리뉴클레오티드를 제공한다. 일부 실시 양태에서, 본 개시는 서열 번호 40의 뉴클레오티드 1-2602의 서열을 포함하는 인자 IX 폴리뉴클레오티드를 제공한다.

[0330] AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 및 AAV9를 포함하는 몇몇 AAV 혈청형이 특징화 되었다. 일반적으로, 임의의 AAV 혈청형이 본원에 기재된 인자 IX 유전자 요법 작제물에 사용될 수 있다. 그러나, 혈청형은 상이한 친화성을 가지며, 예를 들어 이들은 상이한 조직을 우선적으로 감염시킨다. 한 실시 양태에서, 인자 IX는 주로 간에서 생성되기 때문에, 개시된 유전자 요법 작제물에 대한 AAV 혈청형은 적어도 혈청형 AAV7, AAV8 및 AAV9에서 발견되는 간 친화성에 기초하여 선택된다. 따라서, 한 실시 양태에서, 인자 IX 유전자 요법 작제물은 AAV7 혈청형 벡터이다. 다른 실시 형태에서, 인자 IX 유전자 요법 작제물은 AAV8 혈청형 벡터이다. 또 다른 실시 형태에서, 인자 IX 유전자 요법 작제물은 AAV9 혈청형 벡터이다.

[0331] 본원에 기재된 인자 IX 유전자 요법 작제물은 단일 가닥 (예를 들어, 도 1c 및 1d에 도시된 바와 같은 ssAAV 벡터) 또는 자기 상보적 (예를 들어, 도 1a 및 1b에 도시된 scAAV 벡터) 일 수 있다. 연구와 이론은 자기 상보적 AAV 벡터가 더 나은 이식유전자 발현을 촉진해야 한다고 제시하지만, 번역 전에 제2 가닥 합성에 대한 요구를 우회함으로써, 단일-가닥 AAV 벡터가 실시예 5에 보고된 바와 같이, 유사한 자기 상보적 벡터보다 더 나은 인자 IX 발현을 촉진하는 것이 확인되었다.

[0332] *프로모터 및 인핸서*

[0333] 본원에 기재된 인자 IX 유전자 요법 작제물은 일반적으로 생체 내에서 유전자 발현을 유도하는 하나 이상의 프로모터 및/또는 인핸서 요소, 예를 들어 조절 요소를 포함한다. 일부 실시 양태에서, 프로모터 또는 인핸서 요소는 조직 의존적 방식으로, 예를 들어 특정 조직에서 두드러지게 발현을 유도한다. 인자 IX는 주로 간에서 합성되기 때문에, 일부 실시 양태에서, 본원에 기술된 유전자 요법 벡터는 간-특이적 조절 요소를 포함하며, 이는 간 세포로의 유전자 요법 벡터의 발현을 실질적으로 제한한다.

[0334] 일반적으로, 간 특이적 조절 요소는 간에서 독점적으로 발현되는 것으로 알려진 임의의 유전자로부터 유래될 수 있다. WO 2009/130208은 세르핀 펩티드분해효소 억제제, α-항트립신 (SERPINA1; GeneID 5265)으로 알려진, 클레이드 A 멤버 1, 아포지질단백질 C-I (APOC1; GeneID 341), 아포지질단백질 C-IV (APOC4; GeneID 346), 아포지질단백질 H (APOH; GeneID 350); 트랜스티레트린 (TTR; GeneID 7276), 알부민 (ALB; GeneID 213), 알돌라제 B (ALDOB; GeneID 229), 시토크롬 P450, 패밀리 2, 서브패밀리 E, 폴리펩티드 1 (CYP2E1; GeneID 1571), 피브리노겐 알파 사슬 (FGA; GeneID 2243), 트랜스페린 (TF; GeneID 7018), 합토글로빈 관련 단백질 (HPR; GeneID 3250)을 포함하는 간 특이적 방식으로 발현된 몇몇 유전자를 확인한다. 일부 실시 양태에서, 본원에 기재된 인자 IX 유전자 요법 작제물은 하나 이상의 이들 단백질의 게놈 유전자좌로부터 유래된 간-특이적 조절 요소를 포함한다. 이러한 요소의 몇몇 예시는 WO 2009/130208에 기술되어 있으며, 이 내용은 그 전체가 모든 목적을 위해, 본원에 참조로 분명하게 포함된다.

[0335] 간-특이적 조절 요소의 한 예는 일반적으로 "TTRe" 또는 "TTREnh"로 지칭되는, 트랜스티레트린 (TTR) 유전자로부터 유래한다. Hsieh J.L., et al., Cancer Sci., 100(3):537-45 (2009), 이 내용은 그 전체가 모든 목적을 위해, 본원에 참조로 분명하게 포함된다. 일부 실시 양태에서, 본원에 기재된 인자 IX 유전자 요법 작제물은 절단된 TTR 인핸서 및 프로모터 요소를 포함한다. 이들 요소의 예는 도 25와 같이 제공되는 CS06-CRM8.3-ssV의 뉴클레오티드 383-712 (서열 번호 40)에서 제공된다. 일부 실시 양태에서, 절단된 TTR 인핸서 및 프로모터 요소는 CS06-CRM8.3-ssV의 뉴클레오티드 383-712 (서열 번호 40)와 적어도85%의 서열 동일성을 갖는다. 다른 실시 양태에서, 절단된 TTR 인핸서 및 프로모터 요소는 CS06-CRM83-ssV (서열 번호 40)의 뉴클레오티드 383-712에 대해 적어도 90%, 95%, 96%, 97%, 98%, 99%, 99.5% 또는 100%의 서열 동일성을 갖는다.

[0336] 간-특이적 조절 요소의 다른 예는 PCT 공개 번호 WO 2016/146757에 기술된 바와 같이 SERPINA1 유전자로부터 유래하며, 이의 내용은 그 전체가 모든 목적을 위해 본원에 참고로 분명하게 포함된다. 이러한 요소의 예는 CS06-CRM8.3-ssV(서열 번호 40)의 뉴클레오티드 165-236에 제공된 CRM8 조절 제어 요소 (서열 번호 39)이다. 일부 실시 양태에서, SERPINA1-유도 조절 제어 요소는 CRM8 (서열 번호 39)에 대해 적어도 85% 서열 동일성을 갖는다. 다른 실시 양태에서, 절단된 SERPINA1-유도 조절 제어 요소는 CRM8(서열 번호 39)에 대해 적어도 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는다.

[0337] 일부 실시 양태에서, 인자 IX 유전자 요법 작제물은 도 1b 및 1d에 도시된 작제물에 의해 예시된 바와 같이, 하나 이상의 SERPINA1-유도 조절 제어 요소를 포함한다. 한 실시 양태에서, 작제물은 하나의 SERPINA1-유래 조절 제어 요소 (예를 들어, CRM8)를 포함한다. 다른 실시 양태에서, 작제물은 2 개의 SERPINA1-유도된 조절 제어 요소 (예를 들어, CRM8)를 포함한다. 다른 실시 양태에서, 작제물은 3 개의 SERPINA1-유도 조절 제어 요소 (예를 들어, CRM8)를 포함한다. 또 다른 실시 양태에서, 작제물은 4, 5, 6개 이상의 SERPINA1-유도 조절 제어 요소 (예를 들어, CRM8)를 포함한다.

[0338] 한 실시 양태에서, 인자 IX 유전자 요법 작제물은 도 1b, 1d 및 25에 예시된 바와 같이, 하나 이상의 SERPINA1 유래 조절 제어 요소 (예를 들어, CRM8) 및 절단된 TTR 인핸서 및 프로모터 요소를 포함한다.

[0339] *인트론*

[0340] 일부 실시 양태에서, 본원에 기재된 인자 IX 유전자 요법 작제물은 인자 IX 유전자의 발현을 증가시키기 위해 인트론, 예를 들어 바이러스 유래 인트론을 포함한다. 유전자 요법 작제물의 발현에 적합한 인트론은 당 업계에 공지되어 있다. 전형적으로, 인트론은 도 1 및 도 25에 도시된 인자 IX 작제물에서 예시된 바와 같이, 이식유전자 코딩 서열의 5'에 위치된다. 그러나, 일부 실시 양태에서, 인트론은 이식유전자 코딩 서열, 예를 들어 천연 인자 IX 인트론 접합부 또는 그 밖의 다른 위치, 또는 이식유전자 코딩 서열의 3'에 위치될 수 있다. 본원에 기술된 인자 IX 유전자 요법 작제물에 사용될 수 있는 인트론의 비제한적 예는 마우스 미세 바이러스(MVM) 인트론, 베타-글로빈 인트론 (betaIVS-11), 인자 IX (FIX) 인트론 A, 시미안 바이러스 40 (SV40) 스몰 T 인트론, 및 베타-액틴 인트론으로부터 유래된 인트론을 포함한다.

[0341] 한 실시 양태에서, 본원에 기술된 인자 IX 유전자 요법 작제물은 예를 들어, 도 1에 도시되고 도 25의 CS06-CRM8.3-ssV (서열 번호 40)의 뉴클레오티드 724-800에서 MVM 인트론 (서열 번호 53)에 의해 예시된 바와 같은 MVM-유래 인트론을 포함한다. 일부 실시 양태에서, 본원에 기재된 유전자 요법 작제물에 사용된 인트론은 MVM (서열 번호 53)과 적어도 85%의 서열 동일성을 갖는다. 다른 실시 양태에서, 본원에 기재된 유전자 요법 작제물에 사용된 인트론은 MVM(서열 번호 53)에 대해 적어도 90%, 95%, 96%, 97%, 98%, 99% 또는 100%의 서열 동일성을 갖는다.

[0342] *폴리아데닐화 신호*

[0343] 일부 실시 양태에서, 본원에 기재된 인자 IX 유전자 요법 작제물은 예를 들어 도 1에 도시된 바와 같은 폴리아데닐화 신호를 포함한다. 폴리아데닐화 신호는 인자 IX 이식유전자로부터 생성된 mRNA 전사체의 3' 말단에서 폴리-A 테일의 합성을 지시한다. 따라서, 폴리아데닐화 신호는 인자 IX 코딩 서열의 3'에 위치된다. 본원에 기재된 인자 IX 유전자 요법 작제물에 사용될 수 있는 폴리아데닐화 신호의 비제한적 예는 시미 안 바이러스 40 (SV40) 후기 유전자, 소 성장 호르몬 (BGH) 폴리아데닐화 신호, 및 미니멀 토끼 β-글로빈 (mRBG) 유전자로부터 유래된 폴리아데닐화 신호를 포함한다.

[0344] 한 실시 양태에서, 본원에 기술된 인자 IX 유전자 요법 작제물은, 예를 들어, 도 1에 도시되고 도 25의 CS06-CRM8.3-ssV (서열 번호 40)의 뉴클레오티드 2208-2441에서 BGHpA 신호 (서열 번호 54)에 의해 예시된 바와 같이, 소 성장 호르몬 (BGH) 폴리아데닐화 신호로부터 유래된 폴리아데닐화 신호를 포함한다. 일부 실시 양태에서, 본원에 기재된 유전자 요법 작제물에 사용된 폴리아데닐화 신호는 BGHpA 신호 (서열 번호 54)와 적어도 85%의 서열 동일성을 갖는다. 다른 실시 양태에서, 본원에 기재된 유전자 요법 작제물에 사용된 폴리아데닐화 신호는 BGHpA (서열 번호 54) 신호에 대해 적어도 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는다.

[0345] V. 방법

[0346] *생산*

[0347] 본원에 기재된 코돈-변경된 인자 IX 폴리뉴클레오티드 및 바이러스 벡터 (예를 들어, 핵산 조성물)는 핵산 증폭 및 벡터 생산을 위한 통상적인 방법에 따라 생산된다. 재조합 AAV 벡터의 대규모 생산을 위해 2 개의 주요 플랫폼이 개발되었다. 첫 번째 플랫폼은 포유류 세포에서의 복제를 기반으로 하는 반면, 두 번째 플랫폼은 무척추동물 세포에서의 복제를 기반으로 한다. 검토를 위해, Kotin RM, Hum.Mol. Genet., 20(R1):R2-6 (2011)을 참조, 이의 내용은 모든 목적을 위해 그 전체가 본원에 참조로 분명하게 포함된다.

[0348] 따라서, 본 개시 내용은 아데노 관련 바이러스 (AAV) 입자를 생산하는 방법을 제공한다. 일부 실시 양태에서, 방법은 본원에 기재된 바와 같이 CS02, CS03, CS04, CS05 또는 CS06 서열 중 하나에 대해 높은 뉴클레오티드 서열 동일성 (예를 들어, 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%)을 갖는 코돈-변경된 인자 IX 폴리뉴클레오티드 작제물을 폴리뉴클레오티드 작제물이 숙주 세포에서 복제하기에 적합한 숙주 세포 내로 도입하는 것을 포함한다.

[0349] 일부 실시 양태에서, 숙주 세포는 포유류 숙주 세포, 예를 들어 HEK, CHO 또는 BHK 세포이다. 특정 실시 양태에서, 숙주 세포는 HEK 293 세포이다. 일부 실시 양태에서, 숙주 세포는 무척추 동물 세포, 예를 들어 곤충 세포이다. 특정 실시 양태에서, 숙주 세포는 SF9 세포이다.

[0350] *제형*

[0351] 출혈 장애의 치료에 사용하기 위한 조성물이 본원에서 제공된다. 이러한 조성물은 본원에 기술된 바와 같이 치료 유효량의 코돈-변경된 인자 IX 폴리뉴클레오티드, 예를 들어, 인자 IX를 암호화하는 코돈-변경된 폴리뉴클레오티드를 포함하는 AAV 유전자 요법 벡터를 함유한다. 치료적으로 유효한 양의 코돈-변경된 FIX 폴리뉴클레오티드 (예를 들어, 코돈-변경된 인자 IX 코딩 서열을 포함하는 AAV 유전자 요법 벡터)는 전신, 경피 또는 국소 투여를 위해 적합한 약제학적 담체 또는 비히클과 혼합된다. 본원에 개시된 코돈-변경된 인자 IX 폴리뉴클레오티드의 최종 제형은 당업자의 능력 내에 있을 것이다.

[0352] *복용량*

[0353] 본 발명의 핵산 조성물은 이를 필요로 하는 환자에게 투여된다. 투여되는 치료 유전자 치료제의 양 또는 용량은 특정 코돈-변경된 FIX 폴리뉴클레오티드 작제물, 사용된 전달 벡터, 질환의 중증도 및 대상체의 일반적인 특성과 같은 인자에 의존한다. 정확한 투여량은 치료의 목적에 의존할 것이며, 공지된 기술을 사용하여 당업자에 의

해 확인될 수 있을 것이다 (예를 들어, Lieberman, Pharmaceutical Dosage Forms (vols. 1-3, 1992); Lloyd, Art, Science and Technology of Pharmaceutical Compounding (1999); Pickar, Dosage Calculations (1999); 및 Remington:*The Science and Practice of Pharmacy*, 20th Edition, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins를 참조). 특정 대상체의 치료를 위한 특정 투여량 및 투여 요법을 결정하는 것은 숙련된 의사의 능력 내에 있다.

[0354] 일부 실시 양태에서, 코돈-변경된 인자 IX 폴리뉴클레오티드를 갖는 유전자 요법 벡터 (예를 들어, AAV 유전자 요법 벡터 입자)는 치료적 유효량으로 이를 필요로 하는 대상체 (예를 들어, 경증, 중등, 또는 중증 B형 혈우병)에 대해 정맥 투여된다. 일부 실시 양태에서, 치료 유효 용량은 대상체의 킬로그램 체중 당 약 2x10E11 내지 2x10E14 벡터 게놈이다. 특정 실시 양태에서, 치료 유효 용량은 대상체의 킬로그램 체중 당 약 2x10E12 내지 2x10E13 벡터 게놈이다. 일부 구현 양태에서, 대상체는 대상체의 킬로그램 체중 당 2x10E11, 3x10E11, 4x10E11, 5x10E11, 6x10E11, 7x10E11, 8x10E11, 9x10E11, 1x10E12, 2x10E12, 3x10E12, 4x10E12, 5x10E12, 6x10E12, 7x10E12, 8x10E12, 9x10E12, 1x10E13, 2x10E13, 3x10E13, 4x10E13, 5x10E13, 6x10E13, 7x10E13, 8x10E13, 9x10E13, 1x10E14 또는 2x10E14 벡터 게놈이 투여된다.

[0355] 따라서, 본 개시는 인자 IX 결핍 (예를 들어, B형 혈우병)을 치료하는 방법을 제공한다. 일부 실시 양태에서, 방법은 본원에 기재된 바와 같이 CS02, CS03, CS04, CS05 또는 CS06 서열 중 하나에 대해 높은 뉴클레오티드 서열 동일성 (예를 들어, 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%)을 갖는 코돈-변경된 인자 IX 폴리뉴클레오티드 작제물을 이를 필요로 하는 환자에게 투여하는 것을 포함한다. 일부 실시 양태에서, 코돈-변경된 인자 폴리뉴클레오티드는 코돈-변경된 인자 IX 프리-프로-폴리펩티드 코딩 서열에 대해 높은 서열 동일성을 가지며, 예를 들어 CS02-FL-NA (서열 번호 5), CS03-FL-NA (서열 번호 6), CS04-FL-NA (서열 번호7), CS05-FL-NA (서열 번호8) 또는 CS06-FL-NA (서열 번호 9) 중 하나에 대한 높은 서열 동일성을 가진다. 일부 실시 양태에서, 코돈-변경된 인자 폴리뉴클레오티드는 코돈-변경된 성숙한 인자 IX 단일 사슬 폴리펩티드 코딩 서열에 대해 높은 서열 동일성을 가지며, 예를 들어 CS02-MP-NA (서열 번호13), CS03-MP-NA (서열 번호 14), CS04-MP-NA (서열 번호15), CS05-MP-NA (서열 번호16) 또는 CS06-MP-NA (서열 번호 17) 중 하나에 대한 높은 서열 동일성을 가진다.

[0356] 일부 실시 양태에서, 치료는 이를 필요로 하는 환자에게 본원에 기재된 바와 같이 CS02, CS03, CS04, CS05 또는 CS06 서열 중 하나에 대해 높은 뉴클레오티드 서열 동일성(예를 들어, 적어도95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%)을 갖는 코돈-변경된 인자 IX 폴리뉴클레오티드 작제물을 포함하는 유전자 요법 벡터를 투여하는 것을 포함한다. 한 실시 양태에서, 유전자 요법 벡터는 포유류 유전자 요법 벡터이다. 특정 실시 양태에서, 포유류 유전자 요법 벡터는 바이러스 벡터, 예를 들어 렌티바이러스, 레트로바이러스, 아데노 바이러스 또는 아데노 관련 바이러스 벡터이다.

[0357] 한 실시 양태에서, 유전자 요법 벡터는 코돈-변경된 인자 IX 코딩 서열을 암호화하는 바이러스 벡터를 보유하는 아데노-관련 바이러스 (AAV) 입자이다. 일반적으로, 바이러스 벡터는 각각의 말단에서 역 말단 반복 (ITR), 하나 이상의 발현 조절 요소, 코돈-변경된 인자 IX 코딩 서열 및 폴리-A 신호 서열을 포함한다. 특정 실시 양태에서, 유전자 요법 벡터는 간 특이적 조절 제어 요소 (예를 들어, CRM8 요소의 하나 이상의 카피)를 포함한다.

[0358] *생산*

[0359] 본원에 기재된 코돈-변경된 인자 IX 폴리뉴클레오티드 및 바이러스 벡터 (예를 들어, 핵산 조성물)는 핵산 증폭 및 벡터 생산을 위한 통상적인 방법에 따라 생산된다. 재조합 AAV 벡터의 대규모 생산을 위해 2 개의 주요 플랫폼이 개발되었다. 첫 번째 플랫폼은 포유류 세포에서의 복제를 기반으로 하는 반면, 두 번째 플랫폼은 무척추 동물 세포에서의 복제를 기반으로 한다. 검토를 위해, Kotin R.M., Hum.Mol. Genet., 20(R1):R2-6 (2011)을 참조, 이의 내용은 모든 목적을 위해 그 전체가 본원에 참조로 분명하게 포함된다.

[0360] 따라서, 본 개시 내용은 아데노 관련 바이러스 (AAV) 입자를 생산하는 방법을 제공한다. 일부 실시 양태에서, 방법은 본원에 기재된 바와 같이 CS02, CS03, CS04, CS05 또는 CS06 서열 중 하나에 대해 높은 뉴클레오티드 서열 동일성 (예를 들어, 적어도 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%)을 갖는 코돈-변경된 인자 IX 폴리뉴클레오티드 작제물을 폴리뉴클레오티드 작제물이 숙주 세포에서 복제하기에 적합한 숙주 세포 내로 도입하는 것을 포함한다.

[0361] 일부 실시 양태에서, 숙주 세포는 포유류 숙주 세포, 예를 들어 HEK, CHO 또는 BHK 세포이다. 특정 실시 양태에서, 숙주 세포는 HEK 293 세포이다. 일부 실시 양태에서, 숙주 세포는 무척추 동물 세포, 예를 들어 곤충 세포

이다. 특정 실시 양태에서, 숙주 세포는 SF9 세포이다.

[0362] *치료*

[0363] 일부 실시 양태에서, 본원에 기재된 핵산 조성물 (예를 들어, 코돈-변경된 폴리뉴클레오티드)은 공지된 투여 방법에 따라 이를 필요로 하는 대상체에 투여된다. 유전자 요법 벡터를 투여하는 방법은 당업계에 잘 알려져 있다. 이는 정맥 내 투여, 근육 내 주사, 간질 주사 및 간내 투여 (예를 들어, 간내 동맥 또는 정맥)를 포함하지만, 이에 제한되지는 않는다. 예를 들어, Chuah MK et al., Hum Gene Ther., 23(6):557-65 (2012); Chuah MK et al., J Thromb Haemost., 10(8):1566-69 (2012); Chuah MK et al., J Thromb Haemost.11 Suppl 1:99-110 (2013); VandenDriessche et al., Hum Gene Ther. 23(1):4-6 (2012); High KA, Blood, 120(23):4482-87 (2012); Matrai et al., Mol Ther., 18(3):477-90 (2010); 및 Matrai et al., Curr Opin Hematol., 17(5):387-92 (2010)를 참조하며, 검토를 위해 이들 각각이 본원에 참조로 포함된다.

[0364] *치료 효능 평가*

[0365] B형 혈우병 치료의 치료 효능은 예를 들어 치료되는 대상체로부터 혈액의 인자 IX-의존적 응고 잠재력을 측정함으로써 평가될 수 있다. 응고 잠재력을 평가하기 위한 측정법은 시험관내 활성화된 부분 트롬보플라스틴 시간 검정 (APPT), 인자 IX 발색 활성 검정, 혈액 응고 시간, 및 인자 IX 항원 수준 (예를 들어, 인자 IX-특이적 ELISA 사용)을 제한없이 포함한다. 치료 용량은 환자에서 야생형 수준의 FIX를 초래할 필요는 없으며; 오히려, 의미있거나 측정 가능한 방식으로 증상을 감소시키기에 충분한 발현은 본 발명의 목적을 위한 치료제로 간주된다는 것을 유의하여야 한다.

[0366] 국립 혈우병 재단 (National Hemophilia Foundation)에 따르면, 대상체는 그들의 혈장이 정상적인 인간 혈장의 인자 IX 활성의 6% 내지 49%를 함유할 때 경증 B형 혈우병을 갖는 것으로 분류된다. 경증 B형 혈우병 환자는 일반적으로 심각한 부상, 외상 또는 수술 후에만 출혈을 경험한다. 많은 경우, 경증 혈우병은 부상, 수술 또는 치아 추출로 인해 장기간 출혈이 발생할 때까지 진단되지 않는다. 첫 번째 에피소드는 성인이 될 때까지 발생하지 않을 수 있다. 경증 혈우병이 있는 여성은 종종 월경 곤란, 심각한 월경 기간을 경험할 수 있고, 출산 후 대출혈이 될 수 있다.

[0367] 국립 혈우병 재단에 따르면, 대상체는 그들의 혈장이 정상적인 인간 혈장의 인자 IX 활성의 1% 내지 5%를 함유할 때 중등 B형 혈우병을 갖는 것으로 분류된다. B 형 혈우병 환자는 부상 후 출혈이 발생하는 경향이 있다. 명백한 원인없이 발생하는 출혈을 자발적인 출혈 에피소드라고 한다.

[0368] 국립 혈우병 재단에 따르면, 대상체는 그들의 혈장이 정상적인 인간 혈장의 인자 IX 활동의 1% 미만을 함유할 때 중증 B형 혈우병을 갖는 것으로 분류된다. 중증 B형 혈우병을 앓고 있는 대상체는 손상 후 출혈을 경험하고 종종 관절 및 근육에 자발적인 출혈 에피소드를 가질 수 있다.

[0369] 일부 실시 양태에서, 정상적인 인간 혈장은 mL 당 1 IU의 인자 IX 활성을 함유하는 것으로 정의된다. 따라서, 일부 실시 양태에서, 경증 B형 혈우병으로 분류된 대상체로부터의 혈장은 mL 당 0.06 내지 0.49 IU의 인자 IX 활성을 함유한다. 일부 실시 양태에서, 중등 B형 혈우병으로 분류된 대상체로부터의 혈장은 mL 당 0.01 내지 0.05 IU의 인자 IX 활성을 함유한다. 일부 실시 양태에서, 중증 B형 혈우병으로 분류된 대상체로부터의 혈장은 mL 당 0.01 내지 0.05 IU의 인자 IX 활성을 함유한다.

[0370] 따라서, 일부 실시 양태에서, B형 혈우병 요법은 대상체의 혈액/혈장에서 인자 IX 활성의 평균 수준을 증가시킬 때 치료적으로 효과적이다. 일부 실시 양태에서, 치료적으로 유효한 치료는 대상체의 혈액/혈장에서 인자 IX 활성의 평균 수준을 적어도 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 또는 그 이상 증가시킨다. 특정 실시 양태에서, 치료적으로 유효한 혈우병 요법은 대상체의 혈액/혈장에서의 평균 인자 IX 활성을 적어도5% 증가시킨다. 다른 특정 실시 양태에서, 치료적으로 유효한 혈우병 치료는 대상의 혈액/혈장에서 평균 인자 IX 활성을 적어도 10% 증가시킨다. 다른 특정 실시 양태에서, 치료적으로 유효한 혈우병 치료는 대상체의 혈액/혈장에서의 평균 인자 IX 활성을 적어도 15% 증가시킨다. 다른 특정 실시 양태에서, 치료적으로 유효한 혈우병 치료는 대상체의 혈액/혈장에서 평균 인자 IX 활성을 적어도 20% 증가시킨다. 다른 특정 실시 양태에서, 치료적으로 유효한 혈우병 치료는 대상체의 혈액/혈장에서 평균 인자 IX 활성을 적어도 25% 증가시킨다. 다른 특정 실시 양태에서, 치료적으로 유효한 혈우병 치료는 대상체의 혈액/혈장에서 평균 인자 IX 활성을 적어도 30% 증가시킨다.

[0371] 일부 실시 양태에서, 치료적으로 유효한 치료는 대상체가 덜 심각한 형태의 B형 혈우병을 갖는 것으로 분류되도록 대상체 혈액에서 인자 IX 활성의 평균 수준을 증가시킨다. 예를 들어, 한 실시 양태에서, 중증 B형 혈우병으

로 원래 분류된 대상체는 치료적으로 효과적인 치료를 받은 후 중등 B형 혈우병 또는 경증 B형 혈우병으로 재분류된다. 또 다른 실시 양태에서, 원래 중등 B형 혈우병으로 분류된 대상체는 치료적으로 효과적인 치료를 받은 후 경증 B형 혈우병으로 재분류된다.

[0372] VI. 실시예

[0373] *실시예 1 - 코돈-변경된 인자 IX 발현 서열은 FIX 발현 수준을 향상시킨다*

[0374] 생체 내 이종성 인자 IX의 개선된 발현을 제공하는 유전자 요법 작제물을 생성하기 위해, R384L 아미노산 치환(FIXp-FL-AA 서열 번호4)을 갖는 전장 인자 IX 프리프로단백질을 암호화하는 자기 상보적 AAV8-기반 벡터의 패널을 구성하였다. 각각의 작제물의 인자 IX 코딩 서열은 여러 단계를 통해 인간에서의 발현을 개선시키기 위해 변경되었다. 각각의 인자 IX 코딩 서열은 바람직한/바람직하지 않은 서열 모티프를 설명하고 바람직한 인간 코돈에 대한 코돈-사용을 왜곡시키도록 설계된 알고리즘에 따라 변형되었다. 표 2에 보고된 바와 같이, 이 첫 번째 단계를 위해 여러 알고리즘을 사용하였다. 표 2에 보고된 알고리즘을 적용하여 발생된 중간 코돈 변경된 서열을, 그 다음 CpG 디뉴클레오티드를 감소 또는 제거하고, 최종 GC 함량을 조정하고, 바람직한 코돈 쌍을 허용하도록 조정하고, 바람직하지 않은 코돈 쌍을 피하도록 조정하는, 최종 코돈 사용을 조정하도록 수동적으로 추가로 변형하였다. 이러한 고려 사항에 대한 추가 정보는 예를 들어 Fath S. et al., PLoS.One., 6, e17596 (2011); Haas J. et al., Curr.Biol., 6, 315-324 (1996); Tats A., BMC Genomics.9:463 (2008), Grote A. et al., Nucleic Acids Research, 33 (Web Server issue), W526-W531 (2005), Mirsafian H. et al., Scientific World Journal., 639682 (2014), 및 Pechmann S. et al., Nat Struct Mol Biol. 20(2):237-43 (2013)을 참조, 이의 내용은 그 전체가, 모든 목적을 위해, 특히 코돈 변경 고려 사항의 교시를 위해 본원에 참조로 포함된다.

[0375] 각각은 동일한 FIX (R384L) 단백질 (FIXp-FL-AA) (서열 번호 4)에 대해 암호화된 각각의 생성된 코돈-변경된 코딩 서열 (예를 들어, 각각 도 5 내지 9에 도시된, CS02, CS03, CS04, CS05 및 CS06)을 생성하였다. CS02, CS03 및 CS04 작제물은 CpG 모티프를 함유하지 않는 반면, CS05 및 CS06은 각각 11 및 3 개의 CpG를 함유한다.

[0376] 대조군으로서 사용하기 위해, R384L 파두아 아미노산 치환을 갖거나 갖지 않는 야생형 FIX 코딩 서열을 포함하는 벡터 작제물을 또한 생성하였다. WH01 작제물은 R384L 파두아 돌연변이 없이 야생형 FIX 프리프로단백질을 암호화하고, 20 개의 CpG 디뉴클레오티드를 포함한다. WH02 작제물은 R384L 파두아 돌연변이를 갖는 야생형 FIX 프리프로단백질을 암호화하고, 19 개의 CpG 디뉴클레오티드를 포함한다.

[0377] WH01 및 WH02 작제물은 각각 코딩 서열에 20 및 19 개의 CpG를 포함한다. 대조적으로, CS02, CS03 및 CS04 작제물은 CpG 모티프를 함유하지 않는 반면, CS05 및 CS06 작제물은 각각 11 및 3 개의 CpG를 함유한다.

[0378] 도 1a에 나타낸 바와 같이, 코돈-변경된 인자 IX 코딩 서열을 마우스 트랜스티레트린 인핸서/프로모터 (서열 번호 52), 마우스 미세 바이러스 ("MVM"), 인트론 (서열 번호 53), R384L "파두아" 아미노산 치환을 포함하는 코돈-변경된 FIX 작제물을 함유하는 아데노-관련 바이러스 ("AAV") 이식유전자 카세트에 삽입하였다 (미국특허 제6,531,298호; Simione et al., NEJM 361:1671-75 (2009); R384L 돌연변이는 일반적으로 신호 및 프로펩티드가 결여된 인간 단일 사슬 FIX 단백질에서 야생형 아르기닌의 위치를 지칭하는 R338L 돌연변이로 보고되고, 소 성장 호르몬 폴리A 요소 (서열 번호 54)가 뒤따른다). 유전자 카세트는 AAV2 역 말단 반복 ("ITR")에 의해 측면에 위치한다 (서열 번호 51 및 55). 좌측 ITR 반복은 벡터의 자기 상보적 (sc) 표현형을 초래하는 말단 연결기 분해 부위에서의 돌연변이를 포함한다. 기본 벡터 디자인은 Wu et al., Mol. Ther. 16:280-89 (2008) 및 PCT 공개 번호 WO 2014/064277 A1에 기재되어 있으며, 그 내용은 모든 목적을 위해 그 전체가 본원에 참조로 포함된다.

[0379] CS 및 WH 인자 IX AAV 작제물을 B6/129P2-F9tm1Dws FIX 녹아웃 마우스 (Lin et al., Blood, 90:3962-66 (1997)에 기재됨)에 투여하였으며, 그 내용은 모든 목적을 위해 그 전체가 본원에 참고로 포함된다. AAV 벡터 희석물을 개별 동물 체중 (4x10E11 벡터 게놈/킬로그램 (vg/kg) 체중)을 기준으로 측면 꼬리 정맥을 통해 동물 (그룹 당 4 내지 8 마리 동물)에 주사하였다. 유리 모세관을 사용하여 공지된 절차에 따라 투여 후 역 궤도 천공에 의해 정해진 시간 간격으로 혈액 샘플을 수집하였다. 이어서, 혈액을 시트르산 나트륨 항응고제로 미리 채워진 튜브로 옮기고 표준 절차에 의해 혈장을 얻고 -20 ℃에서 동결시켰다.

[0380] 표 2에 보고된 바와 같이, 다양한 인자 IX 작제물의 발현을 측정하고, 마우스에 벡터를 꼬리 정맥 주사한 후 작제물의 효능을 판단하기 위해 FIX 녹아웃 마우스에서 14 일째 혈장 FIX 수준을 사용하였다. 14 일까지 녹아웃 마우스 모델의 발현 수준이 최대 FIX 발현에 거의 도달하였다. 표 2에 나타낸 바와 같이, WH02 FIX (R384L) 대조군 작제물은 4x10E11 벡터 게놈/킬로그램 (vg/kg) 체중의 투여 후 14 일에 FIX 1.03 단위로 발현되었다. 이

발현 수준을 코돈-변경된 인자 IX 작제물의 배수-농축을 측정하기 위한 기준으로서 사용하였다. 표 2에 보고된 바와 같이, CS 코돈-변경된 작제물은 야생형 폴리뉴클레오티드 서열에 의해 암호화된 WH02 대조군 작제물과 비교하여 약 2 배 내지 4 배 증가된 발현을 제공하였다. 특히, CS06 코돈-변경된 작제물은 WH02 대조군 작제물보다 4.2 배 더 큰 인자 IX 활성을 제공하였고, WH01 (야생형 인자 IX) 대조군 작제물보다 21.6 배 더 큰 인자 IX 활성을 제공하였다.

[0381] 표 2.야생형 코돈 서열 (WH01-wtFIX; WH02-FIX(R384L) 및 코돈-변경된 서열 (CS02-CS06)을 가지는 작제물로부터의 인자 IX의 발현.

[0382]

AAV 작제물	벡터 게놈의 변형	CpG 의 수	14 일 발현 수준 [Ｕ인간 FIX]	WH01 과 비교한 배수 발현	WH02 와 비교한 배수 발현
WH01	R338L (파두아) 돌연변이가 없는 인간 FIX 야생형 서열 (GeneBank NM000133.3)	20	0.20	1.0	0.19
WH02	R338L (파두아) 돌연변이를 갖는 인간 FIX 야생형 서열	19	1.03	5.2	1.0
CS02	R338L 돌연변이를 갖는 인간 FIX 서열: 인간 혈청 알부민 코돈에 대해 추가로 최적화된 Geneart 기본 알고리즘.	0	2.12	10.6	2.1
CS03	R338L 돌연변이를 갖는 인간 FIX 서열: 가장 자주 사용되는 인간 코돈에 최적화된 Geneart 기본 알고리즘 (Haas et al., 1996.Curr Biol. 6, 315-324).	0	1.98	9.9	1.9
CS04	R338L 돌연변이를 갖는 인간 FIX 서열: 간 코돈 사용에 대해 추가로 최적화된 Geneart 기본 알고리즘은 Uhlen et al., 2015 Science 347, 6220 에 기술됨.	0	2.77	13.9	2.7
CS05	R338L 돌연변이를 갖는 인간 FIX 서열: CpG 를 줄이기 위해 수정된 JCAT 알고리즘: (Grote et al., 2005.Nucleic Acids Res 1,33).	11	3.93	19.7	3.8
CS06	R338L 돌연변이를 갖는 인간 FIX 서열: 가장 많이 사용되는 인간 코돈에 대해 더욱 최적화된 Geneart 기본 알고리즘 (Haas et al., 1996.Curr Biol. 6, 315-324).	3	4.32	21.6	4.2

[0383] *실시예 2 -- 간-특이적 CRM8 요소는 마우스에서 FIX의 발현을 향상시킨다*

[0384] 코돈-변경된 작제물로부터의 인자 IX 발현 및 활성을 추가로 증가시키기 위해, Nair et al., Blood 123: 3195-99 (2014)에 보고된 바와 같이, 간 특이적 시스-조절 제어 요소 (CRM8 (서열 번호39))의 1 내지 3 개의 카피를 유전자 카세트에 통합하여, 도 1b에 도시된 작제물을 생성하였다. 0 개 (CS02-CRM8.0-V), 1개 (CS02-CRM8.1-V), 2 개 (CS02-CRM8.2-V) 또는 3 개 (CS02-CRM8.3-V)의 CRM8 대조군 요소를 가지는 CS02 코돈-변경된 FIX 코딩 서열을 보유하는 AAV 벡터를 꼬리 정맥 경로에 의해 야생형 마우스에 주사하였다. 이어서, 마우스 혈장에서 인간 FIX 항원을 인간 FIX-특이적 ELISA 검정으로 시간에 걸쳐 측정하였다.

[0385] 표 3에 보고된 바와 같이, CRM8 조절 요소의 사용은 감염 21일 후 CRM8 요소가 결여된 대조군 작제물로부터의 발현과 비교하여 생체 내 인자 IX 발현을 약 2 배 및 4 배 증가시켰다. 예를 들어, 단일 CRM8 요소를 포함하는 CS02-CRM8.1-V 벡터는 CS02-CRM8.0-V 대조군 벡터에서의 FIX 발현의 두 배를 제공했다. CRM8 요소의 다중 카피 포함으로 이러한 발현이 더욱 개선되었다. 예를 들어, 대조군 벡터에 대해 CRM8 요소의 2 개의 카피를 포함하는 벡터는 3 배 발현을 제공하고 CRM8 요소의 3 개의 카피를 포함하는 벡터는 3.4 배의 발현을 제공하였다.

[0386] 표 3. CRM8 조절 제어 요소의 0 내지 3개 카피를 가지는 코돈-변경된 AAV 벡터를 주사한 야생형 마우스의 혈장

에서 인자 IX 발현 수준.

[0387]

#	AAV 작제물	CRM8 요소의 수	FIX (ng/ml) 4 일	FIX (ng/ml) 11 일	FIX (ng/ml) 21 일	배수 증가 21 일
1	CS02-CRM8.0-scV	0	65.8	133.4	239.2	1.0
2	CS02-CRM8.1-scV	1	120.7	250.7	442.8	1.9
3	CS02-CRM8.2-scV	2	152.9	417.3	713.8	3.0
4	CS02-CRM8.3-scV	3	130.9	432.6	800.9	3.4

[0388] *실시예 3 - 간-특이적 CRM8 요소는 인간 간 세포에서 FIX의 발현을 향상시킨다*

[0389] 실시예 2에 기술된 바와 같이, CRM8 간-특이적 조절 제어 요소의 0 내지3개의 카피를 함유하는 CS02 인자 IX 유전자 요법 작제물을 인간 간 세포주 HepG2로 수행된 시험관 내 생체 효능 검정에 의해 추가로 시험하였다. 간략하게, 실시예 2에 기재된 바와 같이, HepG2 세포를 CS02-CRM8-V AAV 벡터 중 하나로 감염시키고, 감염 3 일 후 발색 기질 검정에 의해 FIX 활성을 측정하였다. 실시예 2에 보고된 결과와 일치하게, CRM8 조절 제어 요소를 함유하는 모든 벡터는 표 4에 보고된 바와 같이 더 높은 FIX 발현을 제공하였다. 놀랍게도, 다중 CRM8 요소를 사용하는 효과는 마우스 모델보다 인간 HepG2 세포에서 훨씬 더 두드러졌다. 예를 들어, 대조군 벡터에 대해 CRM8 요소의 2 개의 카피를 포함하는 벡터는 6.7 배의 발현을 제공하였고, CRM8 요소의 3 개의 카피를 포함하는 벡터는 12.8 배의 발현을 제공하였다. 이는 CRM8 조절 제어 요소가 이들 벡터에서 FIX 발현에 미치는 긍정적 효과를 확인시켜 준다.

[0390] 표 4.CRM8 조절 제어 요소의 0 내지 3개 카피를 갖는 코돈-변경된 AAV 벡터가 주입된 인간 간 HepG2 세포에서의 인자 IX 발현 수준.

[0391]

AAV 작제물	CRM8 요소의 수	FIX 활성 [생체역가 단위]	배수 증가
CS02-CRM8.0-scV	0	0.35	1
CS02-CRM8.1-scV	1	0.82	2.3
CS02-CRM8.2-scV	2	2.36	6.7
CS02-CRM8.3-scV	3	4.48	12.8

[0392] *실시예 4 - 단일 가닥 FIX AAV8 벡터는 동등한 자기 상보적 벡터와 유사한 생체 내 발현을 제공한다*

[0393] 일부 경우에, 자기 상보적 (sc) AAV 벡터는 유사한 단일 가닥 (ss) AAV 벡터보다 이식유전자 카세트를 보다 효율적으로 발현한다. 이것은 아마도 세포핵에서 자기 상보적 벡터 게놈의 코팅을 제거한 후 더 빠른 이중 가닥 형성으로 인한 것이다. 검토를 위해, McCarty D., Mol. Ther., (16):1648-56 (2008)를 참조, 그 내용은 모든 목적을 위해 그 전체가 본원에 참조로 포함된다.

[0394] 최근 연구는 EGFP 벡터를 사용하여 이 효과를 확인했다. Bell et al., Hum.Gene Ther. Methods, (27):228-37 (2016). 그러나, 이 연구는 이 효과가 전이 유전자 및 용량 의존적이라는 것을 보여주었다. 예를 들어, 자기 상보적 AAV8 벡터에서의 인간 오르니틴 트랜스카르바밀라제 (hOTC) 유전자 카세트는 상응하는 단일 가닥 벡터와 비교하여 마우스의 간에서 저용량에서 더 나은 발현을 나타냈다. 그러나, 이 효과는 고용량에서 입증될 수 없었으며, 적어도 비-분비 유전자에서 연구된 효과는 이식유전자 및 용량 의존적임을 시사한다. 동일저자

[0395] 단일-가닥 및 자기 상보적 설계의 맥락에서 개시된 코돈-변경된 FIX 유전자 작제물의 특성을 탐색하기 위해, CS06 코돈-변경된 FIX (R338L) 유전자 및 2 개의 온전한 ITR을 보유하는 단일-가닥 작제물을 도 1d 및 1c에 각각 도시된 바와 같이, CRM8 규제 제어 요소와 함께 또는 없이 구성하였다. 단일 가닥 (ss) 벡터를 HEK293 세포 시스템에서 생성하고, 인자 IX 발현을 실시예 1 내지3에 보고된 자기 상보적 작제물의 발현과 비교하였다.

[0396] 먼저, 자기 상보적 (sc) 및 단일 가닥 (ss) CS06-CRM8.0-V 작제물을 상기 기술된 바와 같이 B6/129P2-F9tm1Dws FIX 녹아웃 마우스에 주사 후 생체 내에서 시험하였다. 놀랍게도, 표 5에 보고된 바와 같이, 자기 상보적 (sc) 및 단일 가닥 (ss) CS06 벡터 작제물은 FIX 활성의 매우 유사한 혈장 수준을 보여, ss 벡터와 비교하여 sc 벡터의 보고된 이점을 시사하며, 본원에 기재된 코돈-변경된 인자 IX 작제물을 보유하지 않는다. 발현은 이식유전자 작제물, 전사체의 안정성, 작제물(들)에 사용된 프로모터, 시간 및 용량을 포함하는 많은 매개변수에 의존한다. 표 5에 나타낸 바와 같이, 출혈을 교정하고 FIX ko 마우스에서 장기간 발현을 얻기 위해 선택된 조건하에서, 상응하는 sc 및 ss 벡터는 실질적으로 유사한 발현 수준을 제공하였다.

[0397] FIX 발현에 대한 간-특이적 CRM8 조절 제어 요소의 효과는 또한 단일 가닥 벡터 배경에서 조사하였다. 표 5에 보고된 바와 같이, B6/129P2-F9tm1Dws FIX 녹아웃 마우스에서 단일 가닥 벡터에 하나의 CRM8 요소를 포함시키면 FIX 발현이 개선되었다. 3개의 CRM8 요소의 포함은 단일 가닥 CS06 작제물로부터의 FIX 발현을 CRM8 요소가 없는 자기 상보적 CS06 대조군보다 약간 2 배 이상 높은 수준으로 추가로 개선시켰다. 야생형 WH02 작제물과 비교하여, 단일 가닥 CS06 벡터는 3 개의 CRM8 조절 제어 요소와 쌍을 이룰 때 최대 7배 더 큰 발현을 제공하였다.

[0398] 표 5.다양한 단일 가닥 (ss) 및 자기 상보적 (sc) AAV 인자 IX 벡터를 주입한 FIX 녹아웃 마우스에서의 인자 IX 발현 수준.

AAV 작제물	CRM8 요소의 수	발현 수준¹ (7 일)	발현 수준 (14 일)	발현 수준 (28 일)	배수 증가 대 CS06 (7 일 / 14 일 / 28 일)	배수 증가 대 WH02 (7 일 / 14 일 / 28 일)
CS06-CRM8.0-ssV	0	1.38	2.73	2.99	0.7 / 0.9 / 0.9	1.5 / 2.7 /2.7
CS06-CRM8.1-ssV	1	1.92	3.57	3.47	1.0 / 1.1 / 1.0	2.1 / 3.5 / 3.2
CS06-CRM8.3-ssV (서열번호:40)	3	4.43	6.65	7.78	2.3 / 2.1 / 2.2	4.9 / 6.5 / 7.1
CS06-CRM8.0-scV	0	1.89	3.17	3.50	1.0 / 1.0/ 1.0	2.1 / 3.1 / 3.2
WH02-CRM8.0-scV	0	0.90	1.03	1.10	0.5 / 0.3 / 0.3	1.0

¹국제 단위 (평균 7 내지 8 마우스)에서의 FIX 활성; d, 일;

[0399]

[0400] *실시예 5 - 단일 가닥 FIX AAV8 벡터는 동등한 자기 상보적 벡터보다 인간 간 세포에서 더 나은 FIX 발현을 제공한다*

[0401] 이어서, 실시예 4에 기재된 단일 가닥 CS06 벡터로부터의 인자 IX 발현을 인간 HepG2 세포에서 조사하고, 유사한 자기 상보적 벡터 작제물과 비교하였다. 실시예 4에 보고된 생체 내 결과와 일치하는, CRM8 요소가 없는 단일 가닥 CS06 벡터는 HepG2 세포에서 비교 자기 상보적 벡터보다 약간 낮은 수준에서 FIX 발현을 제공하였다. 그러나, 단일 CRM8 요소의 포함은 표 6에 보고된 바와 같이, 단일 가닥 CS06 벡터로부터의 FIX 발현을 자기 상보적 CS06 벡터로부터의 발현보다 2.6 배 높은 수준으로 증가시켰다.

[0402] 그러나 가장 놀랍게도, 단일 가닥 CS06 벡터에 3 개의 CRM8 요소를 포함시키는 것은 FIX 발현을 자기 상보적 CS06 벡터로부터의 발현보다 16.8 배 높은 수준으로 증가시켰다. 증가된 FIX 발현은 WH02 대조군 벡터로부터의 FIX 발현보다 100 배 이상 더 크다. 요약하면, 3 개의 CRM8 요소를 함유하는 단일 가닥 CS06 벡터는 생체 내 및 시험관 내 생체역가 검정 모두에서 가장 높은 발현 수준을 제공한다.

[0403] 표 6. 인간 간 세포에서 단일 가닥 (ss) 및 자기 상보적 벡터로부터 인자 IX 발현 수준.

AAV 작제물	CRM8 요소의 수	FIX 활성 [생체역가 단위]	배수 증가 대 CS06-CRM8.0-scV	배수 증가 대 WH02-CRM8.0-scV
CS06-CRM8.0-ssV	0	0.24	0.6	3.9
CS06-CRM8.1-ssV	1	0.95	2.6	15.9
CS06-CRM8.3-ssV (서열 번호 40)	3	6.20	16.8	103.3
CS06-CRM8.0-scV	0	0.37	1.0	6.2
WH02-CRM8.0-scV	0	0.06	0.2	1.0

[0404]

[0405] 실시예 1 내지 5에 대한 재료 및 방법

[0406] 동물 실험. FIX 녹아웃 모델에서의 실험을 위해, FIX ko 마우스 균주 B6/129P2-F9tm1Dws (Lin et al., 1997.Blood 90 :3962-6에 의해 개발됨)를 사용하였다. 야생형 마우스 모델에서, 4 내지5 주령 수컷 C57BL6-J B16 마우스를 사용하였다. 두 종 모두 시판 육종가로부터 획득하였다. AAV 벡터 희석물을 개별 동물 체중을 기준으로 측면 꼬리 정맥을 통해 동물 (그룹당 4 내지8 마리 동물)에 주사하였다. 유리 모세관을 사용하여 공지된 절차에 따라 투약 후 역 궤도 천공에 의해 정해진 시간 간격으로 혈액 샘플링을 수행하였다. 이어서, 혈액을 시트르산 나트륨 항응고제로 미리 채워진 튜브로 옮기고 표준 절차에 의해 혈장을 얻고 -20℃에서 동결시켰다.

[0407] FIX 발색 기질 검정을 포함하는 HepG2 세포에서의 시험관 내 생체역가 검정.유전자 치료 벡터 제조를 위한 시험관 내 생체 효능 검정은 인간 간 세포주 HepG2 (ATCC HB-8065)에서 수행하였다. 히드록시우레아로 처리한 후, 세포를 AAV8FIX 벡터로 감염시키고 대략 96 시간 동안 배양하였다. 배양 시간 동안, FIX는 발현되어 세포 상청액으로 방출되었고 FIX-활성은 발색 종점 측정 (Rossix AB, Sweden)에 의해 측정하였다. 각각의 검정 실행은 700 내지 7000 범위의 MOI를 사용하여 정제된 AAV-FIX 벡터 물질의 표준 곡선을 포함한다. MOI 3270에서 표준의 FIX 활동은 1의 생체 역가 단위 (BPU)로 설정한다.

[0408] 마우스 혈장에서의 인간 FIX 정량. 녹아웃 마우스 혈장 FIX 응고 검정에서 인간 FIX를 정량하기 위해 표준 FIX 응고 분석을 사용하여 수행하였다. 야생형 마우스의 혈장에서 인간 FIX 항원을 정량하기 위해 인간 FIX를 특이적으로 검출하는 시판 ELISA 키트 (ASSERACHROM IX:AG (cat. nr. 00943 Stago BNL)를 사용하였다.

[0409] *실시예 6 - CRM8 요소의 포함에 의한 개선된 전사 효능*

[0410] 증가된 전사 효능으로 인한 CRM8-함유 벡터의 개선된 생체 효능의 결과를 표시하기 위해, 인간 간 세포주 (HepG2) 및 마우스 간 세포 (FIX 녹아웃 마우스)를 0개, 1개 또는 3개 CRM8 요소를 함유하는 단일 가닥 CS06 벡터로 형질도입시켰다. FIX mRNA 및 DNA 수준을 표준화된 FIX mRNA 및 DNA 수준 사이의 비로서 측정하고 제시하였다.

[0411] 시험관 내 모델에서, 1개의 CRM8 요소 (CS06-CRM8.1-ssV) 또는 3 개의 CRM8 요소 (CS06-CRM8.3-ssV (서열 번호 40))의 포함은 CRM8 요소 (CS06-CRM8.0-ssV)가 결여된 벡터로 각각 형질도입된 세포에서 보다 형질도입된 인간 간 세포에서 인간 FIX mRNA 수준이 각각, 5 배 및 23 배 더 높았다 (표 6). 유사하게, 생체 내 모델에서, 1 개 또는 3 개의 CRM8 요소를 함유하는 벡터로부터 마우스 간에서의 FIX 발현은 각각 CRM8 요소가 없는 벡터로부터의 FIX 발현보다 2.0 배 및 2.8 배 높았다 (표 6). 두 모델은 CRM8 요소(들)가 FIX 구성체의 전사 활성을 개선시키는 데 유리한 효과를 제공한다는 것을 지지한다.

[0412] 표 7. 인간 간 세포주 또는 마우스 간의 AAV8-FIX 형질 도입 후 FIX mRNA 수준.

[0413]

	시험관 내 간 세포주 (HepG2)		생체 내 마우스 간 조직	
AAV 작제물	정규화된 FIX 비율: mRNA/DNA	배수 증가 대 CS06-CRM8.0-ssV	정규화된 FIX 비율: mRNA/DNA	배수 증가 대 CS06-CRM8.0-ssV
CS06-CRM8.0-ssV	0.025	1	0.38	1
CS06-CRM8.1-ssV	0.14	5.4	0.78	2.0
CS06-CRM8.3-ssV	0.58	23.4	1.09	2.8

[0414] 실시예 6에 대한 방법

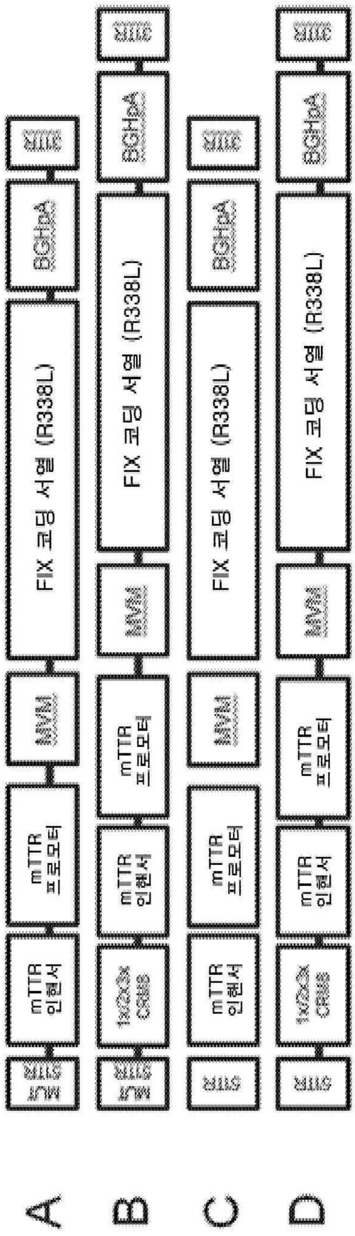
[0415] RNA 및 DNA 추출을 포함하는 정량적 실시간 중합 효소 연쇄 반응.표준 절차에 의해 냉동 간 (동물 실험 참조) 또는 HepG2 세포 (HepG2 세포의 시험관 내 생물역가 검정 참조)로부터 게놈 DNA 및 총 RNA를 추출하였다. 생체 내 실험의 분석을 위해, 14 일째에 각 그룹의 평균 FIX 활성 (평균 ± SD 내)에 근접한 처리군 당 3 마리의 동물의 서브 세트를 선택하였다. cDNA는 매뉴얼 (DNeasy Blood & Tissue Kit, Qiagen, Germany; RNeasy mini kit, Qiagen)에 따라 올리고 (dT20) 프라이머, SuperScript III 역전사 효소 (RT) 및 DNA 분해요소 처리된 총 RNA를 사용하여 합성하였다.

[0416] gDNA 및 cDNA 샘플 둘 모두에서 FIX-이식유전자 카피 수를 FIX 엑손 6의 96 bp 서열을 증폭시켜, 형광 기반 정량적 실시간 중합효소 연쇄 반응 (qPCR)에 의해 측정하였다. 마우스 β-액틴은 내인성 대조군으로서 작용하였고, 시판 TaqMan 검정을 사용하여 정량하였다. qPCR 데이터 분석은 표준 곡선의 선형 회귀 변수를 기반으로 반응 당 FIX 또는 β-액틴 카피를 계산하는 특정 장치의 소프트웨어를 사용하여 수행하였다. 또한, 결과를 1 ㎍의 RNA 또는 DNA로 정규화하고 mRNA:DNA 비를 계산하였다.

[0417] 본 명세서에 기술된 실시예 및 실시 양태는 단지 예시를 위한 것이며, 그에 대한 다양한 변경 또는 수정은 당업자에게 제안될 것이며 본 출원의 사상 및 범위 및 첨부된 청구항의 범위 내에 포함되는 것으로 이해된다. 본원에 인용된 모든 간행물, 특허 및 특허 출원은 모든 목적을 위해 그 전체가 본원에 참조로 포함된다.

도면

도면1



도면2

FIX-FL-NA

```
ATGCAGCGCGTGAACATGATCATGGCAGAATCACCAGGCCTCATCACCATCTGCCTTTTAGGATATC
TACTCAGTGCTGAATGTACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAA
GAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAA
AAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACTGAATTTTGGAAGC
AGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACAT
TAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTAACATGT
AACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCT
GTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGG
AAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTAT
GTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACT
TCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGG
TAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACTGT
GTTGAAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGAACATACAG
AGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAACTACAATGCAGCTATTAATAAGTACAA
CCATGACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGC
ATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAA
GAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGC
CACATGTCTTCGATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGA
GGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCT
TAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGT
ATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAA    (서열번호: 1)
```

도면3a

FIX-FL-AA

```
MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKRYNSG
KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVD
YVNSTEAETI LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIRII
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLRSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRY
VNWIKEKTKL T   (서열번호: 2)
```

도면3b

FIX2-FL-AA

```
mqrvnmimae spgliticll gyllsaectv fldhenanki lnrpkrynsg kleefvqgnl
erecmeekcs feearevfen terttefwkq yvdvtcnikn grceqfckns adnkvvcsct
egyrlaenqk scepavpfpc grvsvsqtsk ltraetvfpd vdyvnsteae tildnitqst
qsfndftrvv ggedakpgqf pwqvvlngkv dafcggsivn ekwivtaahc vetgvkitvv
agehnieete hteqkrnvir iiphhnynaa inkynhdial leldeplvln syvtpiciad
keytniflkf gsgyvsgwgr vfhkgrsalv lqylrvplvd ratclrstkf tiynnmfcag
fheggrdscq gdsggphvte vegtsfltgi iswgeecamk gkygiytkvs ryvnwikekt
klt   (서열번호: 3)
```

도면4

FIXp-FL-AA

```
MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKRYNSG
KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVD
YVNSTEAETI LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIRII
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLLSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRY
VNWIKEKTKL T
```
(서열번호: 4)

도면5

CS02-FL-NA

```
atgcagagggtgaacatgatcatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtcagcagagtgcacagtgttcctggaccatgagaatgccaacaagatcctgaacaggccc
aagagatacaactcaggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccagagaggtgtttgagaacacagagaggaccacagagttctgg
aagcagtatgtggatggagaccagtgtgagagcaacccttgcctgaatggaggcagctgcaaggat
gacatcaacagctatgagtgctggtgccctttttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcaggtgtgagcagttctgcaagaactcagctgacaacaaagtggtg
tgtagctgcacagagggctacagactggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagagtgtcagtgtcccagaccagcaagctgaccagagctgagacagtgttccctgat
gtggactatgtgaatagcacagaggctgagaccatcctggacaacatcacccagagcacccagtcc
ttcaatgacttcaccagagttgtgggaggagaggatgccaagcctggccagttcccctggcaggtg
gtgctgaatggcaaagtggatgccttctgtggaggcagcattgtgaatgagaagtggattgtgaca
gctgcccactgtgtggagacaggagtgaagatcacagtggtggctggagaacacaatattgaggag
acagagcacacagagcagaagaggaatgtcatcaggattatcccccaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcctctggtgctgaatagctat
gtgacccccatctgcattgctgacaaggagtacaccaacatcttcctgaagtttggctcaggctat
gtgtcaggctggggcagagtgttccacaagggcagatcagccctggtgctgcagtacctgagagtg
cccctggtggacagagccacctgcctgttgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggaggcagagacagctgccagggagactcaggaggaccccatgtgacagaa
gtggagggcaccagcttcctgacaggcatcatcagctggggagaggagtgtgccatgaagggcaag
tatggcatctacaccaaagtgagcagatatgtgaactggatcaaggagaaaaccaagctgacctga
```
(서열번호: 5)

도면6

CS03-FL-NA

atgcagagggtgaacatgatcatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcactgtgttcctggaccatgagaatgccaacaagatcctgaacaggccc
aagagatacaactctggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccagggaagtgtttgagaacactgagaggaccactgagttctgg
aagcagtatgtggatggggaccagtgtgagagcaacccttgcctgaatggggcagctgcaaggat
gacatcaacagctatgagtgctggtgccctttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcaggtgtgagcagttctgcaagaactctgctgacaacaaagtggtg
tgtagctgcactgagggctacagactggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagagtgtctgtgtcccagaccagcaagctgaccagagctgagactgtgttccctgat
gtggactatgtgaatagcactgaggctgagaccatcctggacaacatcacccagagcacccagtcc
ttcaatgacttcaccagagtggtggggggaggatgccaagcctggccagttcccctggcaggtg
gtgctgaatggcaaagtggatgccttctgtggggcagcattgtgaatgagaagtggattgtgact
gctgcccactgtgtggagactggggtgaagatcactgtggtggctggggaacacaatattgaggag
actgagcacactgagcagaagaggaatgtcatcaggattatcccccaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcctctggtgctgaatagctat
gtgacccccatctgcattgctgacaaggagtacaccaacatcttcctgaagtttggctctggctat
gtgtctggctggggcagagtgttccacaagggcaggtctgccctggtgctgcagtacctgagagtg
cccctggtggacagagccacctgcctgctgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggggcagagacagctgccaggggactctggggccccatgtgactgaa
gtggaGggcaccagcttcctgactggcatcatcagctgggggaggagtgtgccatgaaggcaag
tatggcatctacaccaaagtgagcaggtatgtgaactggatcaaggagaaaaccaagctgacctga
(서열번호: 6)

도면7

CS04-FL-NA

atgcagagggtgaacatgattatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcacagtgttcctggaccatgagaatgccaacaagatcctgaacaggccc
aagagatacaactctggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccagggaggtgtttgagaacacagagaggaccacagagttctgg
aagcagtatgtggatggtgaccagtgtgagagcaacccttgcctgaatggaggcagctgcaaggat
gacatcaacagctatgagtgctggtgccctttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcaggtgtgagcagttctgcaagaactctgctgacaacaaggtggtg
tgtagctgcacagagggctacagactggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagagtgtctgtgtcccagaccagcaagctgaccagagctgagacagtgttccctgat
gtggactatgtgaacagcacagaggctgagaccatcctggacaacatcacccagagcacccagtcc
ttcaatgacttcaccagagtggtgggaggagaggatgccaagcctggccagttcccctggcaggtg
gtgctgaatggcaaggtggatgccttctgtggaggcagcattgtgaatgagaagtggattgtgaca
gctgcccactgtgtggagacaggagtgaagatcacagtggtggctggagagcacaacattgaggag
acagagcacacagagcagaagaggaatgtgatcaggatcatccctcaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcctctggtgctgaacagctat
gtgacccctatctgcattgctgacaaggagtacaccaacatcttcctgaagtttggctctggctat
gtgtctggctggggcagagtgttccacaagggcaggtctgccctggtgctgcagtacctgagagtg
cccctggtggacagagccacctgcctgttgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggaggcagagacagctgccagggtgactctggaggaccccatgtgacagag
gtggaGggcaccagcttcctgacaggcatcatcagctggggagaggagtgtgccatgaaggcaag
tatggcatctacaccaaagtgagcagatatgtgaactggatcaaggagaagaccaagctgacctga
(서열번호: 7)

도면8

CS05-FL-NA

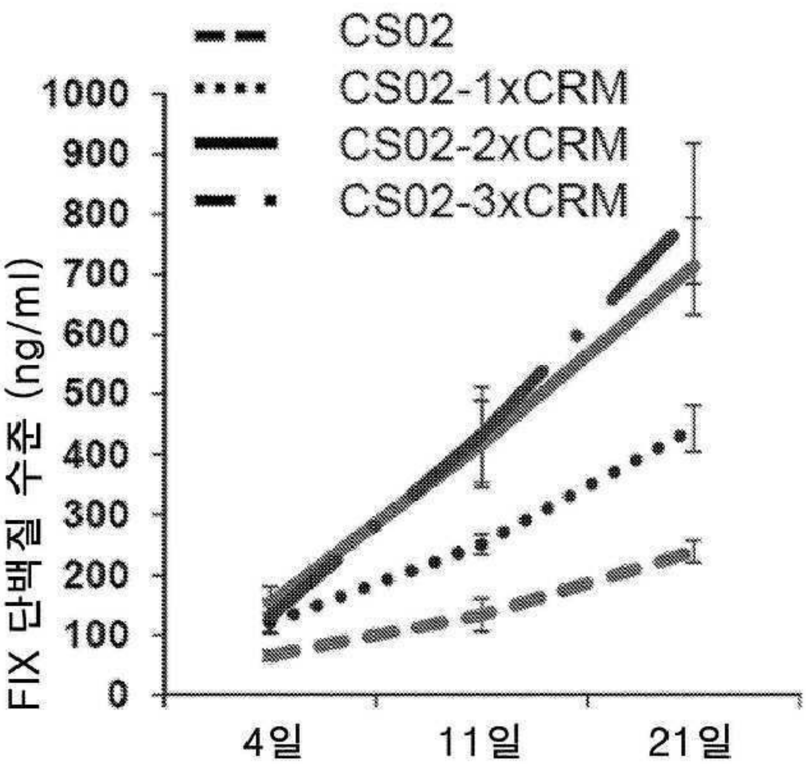
atgcagagggtgaacatgattatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcactgtgttcctggaccatgagaatgccaacaagatcctgaaccgcccc
aagcgctacaactctggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccagggaggtgtttgagaacactgagcgcaccactgagttctgg
aagcagtatgtggatggggaccagtgtgagagcaacccctgcctgaatgggggagctgcaaggat
gacatcaacagctatgagtgctggtgcccctttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggccgctgtgagcagttctgcaagaactctgctgacaacaaggtggtg
tgctcttgcactgagggctaccgcctggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagggtgtctgtgagccagaccagcaagctgaccagggctgagactgtgttccctgac
gtggactatgtgaacagcactgaggctgagaccatcctggacaacatcacccagagcacccagagc
ttcaatgacttcaccagggtggtgggaggagaggatgccaagcctggccagttcccctggcaggtg
gtgctgaatggcaaggtggatgccttctgtggaggcagcattgtgaatgagaagtggattgtgacc
gctgcccactgtgtggagactggagtgaagatcactgtggtggctggggagcacaacattgaggag
acagagcacacagagcagaagcgcaatgtgatcaggatcatccccccaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcccctggtgctgaacagctac
gtgacccccatctgcattgcagacaaggagtacaccaacatcttcctgaagtttggctctggctat
gtgtctggctggggcagggtgttccacaagggcaggtctgccctggtgctgcagtacctgagggtg
cccctggtggacagggccacctgcctgctgagcaccaagttcaccatctacaacaacatgttctgc
gctggcttccatgagggaggaagggacagctgccagggagactctggaggccccccatgtgacagag
gtggagggcaccagcttcctgacaggcatcatcagctgggggggaggagtgtgccatgaagggcaag
tatggcatctacaccaaagtgtcccgctatgtgaactggatcaaggagaagaccaagctgacctga
(서열번호: 8)

도면9

CS06-FL-NA

atgcagagggtcaacatgatcatggctgagtcccctggcctcatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcactgtcttcctggaccatgagaatgccaacaagatcctcaacaggccc
aagagatacaactctggcaaactggaggagtttgtccagggcaacctggagagggagtgcatggag
gagaagtgctcctttgaggaggccagggaggtctttgagaacactgagcgcaccactgagttctgg
aaacagtatgtggatggggaccagtgtgagtccaacccctgcctgaatggggcagctgcaaggat
gacatcaacagctatgagtgctggtgcccctttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcagatgtgagcagttctgcaagaactctgctgacaacaaggtggtg
tgctcctgcactgagggctaccgcctggctgagaaccagaagagctgtgagcctgctgtgccattc
ccatgtggcagagtctctgtgagccagaccagcaagctcaccagggctgagactgtgttccctgat
gtggactatgtgaacagcactgaggctgaaaccatcctggacaacatcacccagagcacccagagc
ttcaatgacttcaccagagtggtgggaggagaggatgccaagcctggccagttcccctggcaagtg
gtgctcaatggcaaggtggatgccttctgtgggggctccattgtgaatgagaagtggattgtcact
gctgcccactgtgtggagactggggtcaagatcactgtggtggctggggagcacaacattgaggag
actgagcacactgagcagaagcgcaatgtgatcaggatcatccccccaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcccctggtcctcaacagctat
gtgacccccatctgcattgctgacaaggagtacaccaacatcttcctcaagtttggctctggctat
gtctctggctggggcagagtgttccacaaaggcaggtctgccctggtgctccagtacctgagagtg
cccctggtggacagggccacctgcctcttgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggaggaagagacagctgccagggggactctggaggaccccatgtcactgag
gtggagggcacctccttcctcactggcatcatctcctggggagaggagtgtgccatgaaaggcaaa
tatggcatctacaccaaagtctccagatatgtcaactggatcaaggagaagaccaagctgacctga
(서열번호: 9)

도면10



도면11a

FIX-MP-AA

YNSG KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVD
YVNSTEAETI LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIRII
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLRSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRY
VNWIKEKTKL T (서열번호: 10)

도면11b

FIX2-MP-AA

ynsg kleefvqgnl erecmeekcs feearevfen terttefwkq yvdvtcnikn
grceqfckns adnkvvcsct egyrlaenqk scepavpfpc grvsvsqtsk ltraetvfpd
vdyvnsteae tildnitqst qsfndftrvv ggedakpgqf pwqvvlngkv dafcggsivn
ekwivtaahc vetgvkitvv agehnieete hteqkrnvir iiphhnynaa inkynhdial
leldeplvln syvtpiciad keytniflkf gsgyvsgwgr vfhkgrsalv lqylrvplvd
ratclrstkf tiynnmfcag fheggrdscq gdsggphvte vegtsfltgi iswgeecamk
gkygiytkvs ryvnwikekt klt (서열번호: 11)

도면12

FIXp-MP-AA

YNSG KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVD
YVNSTEAETI LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIRII
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLLSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRY
VNWIKEKTKL T (서열번호: 12)

도면13

CS02-MP-NA

tacaactcag gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag
gagaagtgca gctttgagga ggccagagag gtgtttgaga acacagagag gaccacagag
ttctggaagc agtatgtgga tggagaccag tgtgagagca acccttgcct gaatggaggc
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag
aactcagctg acaacaaagt ggtgtgtagc tgcacagagg gctacagact ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtcagt gtcccagacc
agcaagctga ccagagctga gacagtgttc cctgatgtgg actatgtgaa tagcacagag
gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga
gttgtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc
aaagtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagaacacaa tattgaggag
acagagcaca cagagcagaa gaggaatgtc atcaggatta tcccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg
aagtttggct caggctatgt gtcaggctgg ggcagagtgt tccacaaggg cagatcagcc
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc
tgccagggag actcaggagg accccatgtg acagaagtgg agggcaccag cttcctgaca
ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgagcagat atgtgaactg gatcaaggag aaaaccaagc tgacctga
(서열번호: 13)

도면14

CS03-MP-NA

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag
gagaagtgca gctttgagga ggccagggaa gtgtttgaga acactgagag gaccactgag
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccttgcct gaatgggggc
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag
aactctgctg acaacaaagt ggtgtgtagc tgcactgagg gctacagact ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc cctgtggca gagtgtctgt gtcccagacc
agcaagctga ccagagctga gactgtgttc cctgatgtgg actatgtgaa tagcactgag
gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga
gtggtggggg gggaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc
aaagtggatg ccttctgtgg gggcagcatt gtgaatgaga agtggattgt gactgctgcc
cactgtgtgg agactggggt gaagatcact gtggtggctg gggaacacaa tattgaggag
actgagcaca ctgagcagaa gaggaatgtc atcaggatta tcccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg
aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gctgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggggg cagagacagc
tgccaggggg actctggggg cccccatgtg actgaagtgg agggcaccag cttcctgact
ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgagcaggt atgtgaactg gatcaaggag aaaaccaagc tgacctga

(서열번호: 14)

도면15

CS04-MP-NA

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag
gagaagtgca gctttgagga ggccagggag gtgtttgaga acacagagag gaccacagag
ttctggaagc agtatgtgga tggtgaccag tgtgagagca acccttgcct gaatggaggc
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag
aactctgctg acaacaaggt ggtgtgtagc tgcacagagg gctacagact ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc cctgtggca gagtgtctgt gtcccagacc
agcaagctga ccagagctga gacagtgttc cctgatgtgg actatgtgaa cagcacagag
gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc
aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagagcacaa cattgaggag
acagagcaca cagagcagaa gaggaatgtg atcaggatca tccctcacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg
ctgaacagct atgtgacccc tatctgcatt gctgacaagg agtacaccaa catcttcctg
aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc
tgccagggtg actctggagg accccatgtg acagaggtgg agggcaccag cttcctgaca
ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgagcagat atgtgaactg gatcaaggag aagaccaagc tgacctga

(서열번호: 15)

도면16

CS05-MP-NA

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag
gagaagtgca gctttgagga ggccagggag gtgtttgaga acactgagcg caccactgag
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccctgcct gaatgggggg
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggcc gctgtgagca gttctgcaag
aactctgctg acaacaaggt ggtgtgctct tgcactgagg gctaccgcct ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc cctgtggca gggtgtctgt gagccagacc
agcaagctga ccagggctga gactgtgttc cctgacgtgg actatgtgaa cagcactgag
gctgagacca tcctggacaa catcacccag agcacccaga gcttcaatga cttcaccagg
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc
aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gaccgctgcc
cactgtgtgg agactggagt gaagatcact gtggtggctg gggagcacaa cattgaggag
acagagcaca cagagcagaa gcgcaatgtg atcaggatca tcccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcccctggtg
ctgaacagct acgtgacccc catctgcatt gcagacaagg agtacaccaa catcttcctg
aagtttggct ctggctatgt gtctggctgg ggcagggtgt tccacaaggg caggtctgcc
ctggtgctgc agtacctgag ggtgcccctg gtggacaggg ccacctgcct gctgagcacc
aagttcacca tctacaacaa catgttctgc gctggcttcc atgagggagg aagggacagc
tgccagggag actctggagg cccccatgtg acagaggtgg agggcaccag cttcctgaca
ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgtcccgct atgtgaactg gatcaaggag aagaccaagc tgacctga
(서열번호: 16)

도면17

CS06-MP-NA

tacaactctg gcaaactgga ggagtttgtc cagggcaacc tggagaggga gtgcatggag
gagaagtgct cctttgagga ggccagggag gtctttgaga acactgagcg caccactgag
ttctggaaac agtatgtgga tggggaccag tgtgagtcca acccctgcct gaatggggc
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca gatgtgagca gttctgcaag
aactctgctg acaacaaggt ggtgtgctcc tgcactgagg gctaccgcct ggctgagaac
cagaagagct gtgagcctgc tgtgccattc ccatgtggca gagtctctgt gagccagacc
agcaagctca ccagggctga gactgtgttc cctgatgtgg actatgtgaa cagcactgag
gctgaaacca tcctggacaa catcacccag agcacccaga gcttcaatga cttcaccaga
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaagtggt gctcaatggc
aaggtggatg ccttctgtgg gggctccatt gtgaatgaga agtggattgt cactgctgcc
cactgtgtgg agactggggt caagatcact gtggtggctg gggagcacaa cattgaggag
actgagcaca ctgagcagaa gcgcaatgtg atcaggatca tcccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcccctggtc
ctcaacagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctc
aagtttggct ctggctatgt ctctggctgg ggcagagtgt tccacaaagg caggtctgcc
ctggtgctcc agtacctgag agtgcccctg gtggacaggg ccacctgcct cttgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg aagagacagc
tgccaggggg actctggagg accccatgtc actgaggtgg agggcacctc cttcctcact
ggcatcatct cctggggaga ggagtgtgcc atgaaaggca aatatggcat ctacaccaaa
gtctccagat atgtcaactg gatcaaggag aagaccaagc tgacctga
(서열번호: 17)

도면18

FIX-PPP-NA

atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta
ggatatctac tcagtgctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt
ctgaatcggc caaagagg (서열번호: 18)

CS02-PPP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtcagcaga gtgcacagtg ttcctggacc atgagaatgc caacaagatc
ctgaacaggc ccaagaga (서열번호: 19)

CS03-PPP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc
ctgaacaggc ccaagaga (서열번호: 20)

CS04-PPP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgcacagtg ttcctggacc atgagaatgc caacaagatc
ctgaacaggc ccaagaga (서열번호: 21)

CS05-PPP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc
ctgaaccgcc ccaagcgc (서열번호: 22)

CS06-PPP-NA

atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgcactgtc ttcctggacc atgagaatgc caacaagatc
ctcaacaggc ccaagaga (서열번호: 23)

도면19

FIX-SP-NA

atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgccttta
ggatatctac tcagtgctga atgt (서열번호: 24)

CS02-SP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtcagcaga gtgc (서열번호: 25)

CS03-SP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (서열번호: 26)

CS04-SP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (서열번호: 27)

CS05-SP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (서열번호: 28)

CS06-SP-NA

atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (서열번호: 29)

도면20

FIX-PP-NA

acagtttttc ttgatcatga aaacgccaac aaaattctga atcggccaaa gagg
(서열번호: 30)

CS02-PP-NA

acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga
(서열번호: 31)

CS03-PP-NA

actgtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga
(서열번호: 32)

CS04-PP-NA

acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga
(서열번호: 33)

CS05-PP-NA

actgtgttcc tggaccatga gaatgccaac aagatcctga accgccccaa gcgc
(서열번호: 34)

CS06-PP-NA

actgtcttcc tggaccatga gaatgccaac aagatcctca acaggcccaa gaga
(서열번호: 35)

도면21

FIX-PPP-AA

MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKR (서열번호: 36)

도면22

FIX-SP-AA

MQRVNMIMAE SPGLITICLL GYLLSAEC (서열번호: 37)

도면23

FIX-PP-AA

TV FLDHENANKI LNRPKR (서열번호: 38)

도면24

CRM8

ggggaggctgctggtgaatattaaccaaggtcaccccagttatcggaggagcaaacaggggctaag
tccac (서열번호: 39)

도면25a

CS06-CRM8.3-ssV

```
ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcc
cgggctttgcccgggcggcctcagtgagcgagcgagcgcgcagagagggagtggccaactccatcac
taggggttcctgagtttaaacttcgtcgacggggaggctgctggtgaatattaaccaaggtcaccc
cagttatcggaggagcaaacaggggctaagtccaccggggaggctgctggtgaatattaaccaagg
tcaccccagttatcggaggagcaaacaggggctaagtccaccggggaggctgctggtgaatattaa
ccaaggtcaccccagttatcggaggagcaaacaggggctaagtccaccgagggcactgggaggatgt
tgagtaagatggaaaactactgatgacccttgcagagacagagtattaggacatgtttgaacagggg
ccgggcgatcagcaggtagctctagaggatccccgtctgtctgcacatttcgtagagcgagtgttcc
gatactctaatctccctaggcaaggttcatatttgtgtaggttacttattctccttttgttgactaa
gtcaataatcagaatcagcaggtttggagtcagcttggcagggatcagcagcctgggttggaaggag
ggggtataaaagccccttcaccaggagaagccgtcacacagactaggcgcgccctaaggtaagttgg
cgccgtttaagggatggttggttggtggggtattaatgtttaattacctttttacaggcctgaaga
tctgccaccatgcagagggtcaacatgatcatggctgagtcccctggcctcatcaccatctgcctgc
tgggctacctgctgtctgctgagtgcactgtcttcctggaccatgagaatgccaacaagatcctcaa
caggcccaagagatacaactctggcaaactggaggagtttgtccagggcaacctggagagggagtgc
atggaggagaagtgctcctttgaggaggccagggaggtctttgagaacactgagcgcaccactgagt
tctggaaacagtatgtggatggggaccagtgtgagtccaacccctgcctgaatgggggcagctgcaa
ggatgacatcaacagctatgagtgctggtgcccctttggctttgagggcaagaactgtgagctggat
gtgacctgcaacatcaagaatggcagatgtgagcagttctgcaagaactctgctgacaacaaggtgg
tgtgctcctgcactgagggctaccgcctggctgagaaccagaagagctgtgagcctgctgtgccatt
cccatgtggcagagtctctgtgagccagaccagcaagctcaccagggctgagactgtgttccctgat
gtggactatgtgaacagcactgaggctgaaaccatcctggacaacatcacccagagcacccagagct
tcaatgacttcaccagagtggtgggaggagaggatgccaagcctggccagttcccctggcaagtggt
gctcaatggcaaggtggatgccttctgtgggggctccattgtgaatgagaagtggattgtcactgct
gcccactgtgtggagactggggtcaagatcactgtggtggctggggagcacaacattgaggagactg
agcacactgagcagaagcgcaatgtgatcaggatcatcccccaccacaactacaatgctgccatcaa
caagtacaaccatgacattgccctgctggagctggatgagcccctggtcctcaacagctatgtgacc
cccatctgcattgctgacaaggagtacaccaacatcttcctcaagtttggctctggctatgtctctg
gctggggcagagtgttccacaaaggcaggtctgccctggtgctccagtacctgagagtgcccctggt
ggacagggccacctgcctcttgagcaccaagttcaccatctacaacaacatgttctgtgctggcttc
catgagggaggaagagacagctgccaggggggactctggaggaccccatgtcactgaggtggagggca
cctccttcctcactggcatcatctcctggggagaggagtgtgccatgaaaggcaaatatggcatcta
caccaaagtctccagatatgtcaactggatcaaggagaagaccaagctgacctgatgagcatgccta
gagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgt
gccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcg
cattgtcTgagtaggtgtcattctattctggggggtggggtggggcaggacagcaagggggaggatt
gggaagacaatagcaggcatgctggggaattaattaagctcgcgaaggaacccctagtgatggagtt
```

(계속)

도면25b

(계속)

ggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccg
ggctttgcccgggcggcctcagtgagcgagcgagcgcgcagagagggagtggccaagacgatttaaa
tgacaagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattcc
acacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcaca
ttaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaa
tcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactc
gctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatcc
acagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgta
aaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacg
ctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctcc
ctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaa
gcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagct
gggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag
tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcga
ggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagt
atttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggc
aaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaag
gatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgtta
agggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagt
tttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgagg
cacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataac
tacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccg
gctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactt
tatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatag
tttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttca
ttcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggtta
gctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggc
agcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactca
accaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggata
ataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaact
ctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttca
gcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagg
gaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcattta
tcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggggtt
ccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacct
ataaaaataggcgtatcacgaggccctttcgtctcgcgcgtttcggtgatgacggtgaaaacctctg
acacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgt
cagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattg
tactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgtaaggagaaaataccgcatcag
gcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcTtcgctatta
cgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagt
cacgacgttgtaaaacgacggccagtgaattcctcgagatttaaatgacg

(서열번호: 40)

서 열 목 록

SEQUENCE LISTING

<110> Baxalta Incorporated

Baxalta GmbH

Horling, Franziska

Lengler, Johannes

Falkner, Falko-Guenther

Rottensteiner, Hanspeter

Scheiflinger, Friedrich

<120> VIRAL VECTORS ENCODING RECOMBINANT FIX WITH INCREASED EXPRESSION FOR GENE THERAPY OF HEMOPHILIA B

<130> 008073-5117-WO

<150> US 62/509,616

<151> 2017-05-22

<160> 64

<170> PatentIn version 3.5

<210> 1

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> FIX-FL-NA

<400> 1

```
atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta      60
ggatatctac tcagtgctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt     120
ctgaatcggc caaagaggta taattcaggt aaattggaag agtttgttca agggaacctt     180
gagagagaat gtatggaaga aaagtgtagt tttgaagaag cacgagaagt tttgaaaac      240
actgaaagaa caactgaatt ttggaagcag tatgttgatg gagatcagtg tgagtccaat     300
ccatgtttaa atggcggcag ttgcaaggat gacattaatt cctatgaatg ttggtgtccc     360
tttggatttg aaggaaagaa ctgtgaatta gatgtaacat gtaacattaa gaatggcaga     420

tgcgagcagt tttgtaaaaa tagtgctgat aacaaggtgg tttgctcctg tactgaggga     480
tatcgacttg cagaaaacca gaagtcctgt gaaccagcag tgccatttcc atgtggaaga     540
gtttctgttt cacaaacttc taagctcacc cgtgctgaga ctgtttttcc tgatgtggac     600
tatgtaaatt ctactgaagc tgaaaccatt ttggataaca tcactcaaag cacccaatca     660
tttaatgact tcactcgggt tgttggtgga gaagatgcca aaccaggtca attcccttgg     720
caggttgttt tgaatggtaa agttgatgca ttctgtggag gctctatcgt taatgaaaaa     780
tggattgtaa ctgctgccca ctgtgttgaa actggtgtta aaattacagt tgtcgcaggt     840

gaacataata ttgaggagac agaacataca gagcaaaagc gaaatgtgat tcgaattatt     900
cctcaccaca actacaatgc agctattaat aagtacaacc atgacattgc ccttctggaa     960
ctggacgaac ccttagtgct aaacagctac gttacaccta tttgcattgc tgacaaggaa    1020
tacacgaaca tcttcctcaa atttggatct ggctatgtaa gtggctgggg aagagtcttc    1080
```

cacaaaggga gatcagcttt agttcttcag taccttagag ttccacttgt tgaccgagcc 1140

acatgtcttc gatctacaaa gttcaccatc tataacaaca tgttctgtgc tggcttccat 1200

gaaggaggta gagattcatg tcaaggagat agtgggggac cccatgttac tgaagtggaa 1260

gggaccagtt tcttaactgg aattattagc tggggtgaag agtgtgcaat gaaaggcaaa 1320

tatggaatat ataccaaggt atcccggtat gtcaactgga ttaaggaaaa aacaaagctc 1380

acttaa 1386

<210> 2

<211> 461

<212> PRT

<213> Artificial Sequence

<220><223> FIX-FL-AA

<400> 2

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu
20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg Tyr Asn
35 40 45

Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg Glu Cys
50 55 60

Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe Glu Asn
65 70 75 80

Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly Asp Gln
85 90 95

Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp Asp Ile
100 105 110

Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys Asn Cys
115 120 125

Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe
130 135 140

Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly

145 150 155 160

Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val Pro Phe

165 170 175

Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala

180 185 190

Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala Glu

195 200 205

Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp Phe

210 215 220

Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp

225 230 235 240

Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile

245 250 255

Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly

260 265 270

Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu

275 280 285

His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn

290 295 300

Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu

305 310 315 320

Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile

325 330 335

Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr

340 345 350

Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val

355 360 365

Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg

370 375 380

Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His

385 390 395 400

Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val

405 410 415

Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly

420 425 430

Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser

435 440 445

Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

450 455 460

<210> 3

<211> 423

<212> PRT

<213> Artificial Sequence

<220>

<223> FIX2-FL-AA

<400> 3

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr

1 5 10 15

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

20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg Tyr Asn

35 40 45

Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg Glu Cys

50 55 60

Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe Glu Asn

65 70 75 80

Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Val Thr Cys

85 90 95

Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe Cys Lys Asn Ser Ala Asp

100 105 110

Asn Lys Val Val Cys Ser Cys Thr Glu Gly Tyr Arg Leu Ala Glu Asn

115 120 125

Gln Lys Ser Cys Glu Pro Ala Val Pro Phe Pro Cys Gly Arg Val Ser

130 135 140

Val Ser Gln Thr Ser Lys Leu Thr Arg Ala Glu Thr Val Phe Pro Asp

145 150 155 160

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

165 170 175

Thr Gln Ser Thr Gln Ser Phe Asn Asp Phe Thr Arg Val Val Gly Gly

180 185 190

Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp Gln Val Val Leu Asn Gly

195 200 205

Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn Glu Lys Trp Ile

210 215 220

Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys Ile Thr Val Val

225 230 235 240

Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr Glu Gln Lys Arg

245 250 255

Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn Ala Ala Ile Asn

260 265 270

Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp Glu Pro Leu Val

275 280 285

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

290 295 300

Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser Gly Trp Gly Arg

305 310 315 320

Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln Tyr Leu Arg Val

325 330 335

Pro Leu Val Asp Arg Ala Thr Cys Leu Arg Ser Thr Lys Phe Thr Ile

340 345 350

Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly Gly Arg Asp Ser

355 360 365

Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu Val Glu Gly Thr

370 375 380

Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu Cys Ala Met Lys

385 390 395 400

Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr Val Asn Trp Ile

405 410 415

Lys Glu Lys Thr Lys Leu Thr

420

<210> 4

<211> 461

<212> PRT

<213> Artificial Sequence

<220><223> FIXp-FL-AA

<400> 4

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr

1 5 10 15

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

20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg Tyr Asn

35 40 45

Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg Glu Cys

50 55 60

Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe Glu Asn

65 70 75 80

Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly Asp Gln

85 90 95

Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp Asp Ile

100 105 110

Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys Asn Cys

115 120 125

Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe

130 135 140

Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly

145 150 155 160

Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val Pro Phe
165 170 175

Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala
180 185 190

Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala Glu
195 200 205

Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp Phe
210 215 220

Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp
225 230 235 240

Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile
245 250 255

Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly
260 265 270

Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu
275 280 285

His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn
290 295 300

Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu
305 310 315 320

Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile
325 330 335

Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr
340 345 350

Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val
355 360 365

Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Leu
370 375 380

Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His
385 390 395 400

Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val

405 410 415

Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly

420 425 430

Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser

435 440 445

Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

450 455 460

<210>

5

<211> 1385

<212> DNA

<213> Artificial Sequence

<220><223> CS02-FL-NA

<400> 5

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtcagcaga gtgcacagtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagagata caactcaggc aagctggagg agtttgtgca gggcaacctg 180

gagagggagt gcatggagga gaagtgcagc tttgaggagg ccagagaggt gtttgagaac 240

acagagagga ccacagagtt ctggaagcag tatgtggatg gagaccagtg tgagagcaac 300

ccttgcctga atggaggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgccct 360

tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcagg 420

tgtgagcagt tctgcaagaa ctcagctgac aacaaagtgg tgtgtagctg cacagagggc 480

tacagactgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcaga 540

gtgtcagtgt cccagaccag caagctgacc agagctgaga cagtgttccc tgatgtggac 600

tatgtgaata gcacagaggc tgagaccatc ctggacaaca tcacccagag cacccagtcc 660

ttcaatgact tcaccagagt tgtgggagga gaggatgcca agcctggcca gttcccctgg 720

caggtggtgc tgaatggcaa agtggatgcc ttctgtggag gcagcattgt gaatgagaag 780

tggattgtga cagctgccca ctgtgtggag acaggagtga agatcacagt ggtggctgga 840

gaacacaata ttgaggagac agagcacaca gagcagaaga ggaatgtcat caggattatc 900

ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag 960

ctggatgagc ctctggtgct gaatagctat gtgacccccca tctgcattgc tgacaaggag 1020

tacaccaaca tcttcctgaa gtttggctca ggctatgtgt caggctgggg cagagtgttc 1080

cacaagggca gatcagccct ggtgctgcag tacctgagag tgcccctggt ggacagagcc 1140 acctgcctgt tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggcttccat 1200

gagggaggca gagacagctg ccagggagac tcaggaggac cccatgtgac agaagtggag 1260

ggcaccagct tcctgacagg catcatcagc tggggagagg agtgtgccat gaagggcaag 1320

tatggcatct acaccaaagt gagcagatat gtgaactgga tcaaggagaa aaccaagctg 1380

acctg 1385

<210> 6

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> CS03-FL-NA

<400> 6

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagagata caactctggc aagctggagg agtttgtgca gggcaacctg 180

gagagggagt gcatggagga gaagtgcagc tttgaggagg ccagggaagt gtttgagaac 240

actgagagga ccactgagtt ctggaagcag tatgtggatg gggaccagtg tgagagcaac 300

ccttgcctga atgggggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgccct 360

tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcagg 420

tgtgagcagt tctgcaagaa ctctgctgac aacaaagtgg tgtgtagctg cactgagggc 480

tacagactgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcaga 540

gtgtctgtgt cccagaccag caagctgacc agagctgaga ctgtgttccc tgatgtggac 600

tatgtgaata gcactgaggc tgagaccatc ctggacaaca tcacccagag cacccagtcc 660

ttcaatgact tcaccagagt ggtggggggg gaggatgcca agcctggcca gttcccctgg 720

caggtggtgc tgaatggcaa agtggatgcc ttctgtgggg gcagcattgt gaatgagaag 780

tggattgtga ctgctgccca ctgtgtggag actggggtga agatcactgt ggtggctggg 840

gaacacaata ttgaggagac tgagcacact gagcagaaga ggaatgtcat caggattatc 900

cccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag 960

ctggatgagc ctctggtgct gaatagctat gtgacccccca tctgcattgc tgacaaggag 1020

tacaccaaca tcttcctgaa gtttggctct ggctatgtgt ctggctgggg cagagtgttc 1080

cacaagggca ggtctgccct ggtgctgcag tacctgagag tgcccctggt ggacagagcc 1140

acctgcctgc tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggcttccat 1200 gagggggca gagacagctg ccagggggac tctgggggcc cccatgtgac tgaagtggag 1260

ggcaccagct tcctgactgg catcatcagc tgggggagg agtgtgccat gaagggcaag 1320

tatggcatct acaccaaagt gagcaggtat gtgaactgga tcaaggagaa aaccaagctg 1380

acctga 1386

<210> 7

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> CS04-FL-NA

<400> 7

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcacagtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagagata caactctggc aagctggagg agtttgtgca gggcaacctg 180

gagagggagt gcatggagga gaagtgcagc tttgaggagg ccagggaggt gtttgagaac 240

acagagagga ccacagagtt ctggaagcag tatgtggatg gtgaccagtg tgagagcaac 300

ccttgcctga atggaggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgccct 360

tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcagg 420

tgtgagcagt tctgcaagaa ctctgctgac aacaaggtgg tgtgtagctg cacagagggc 480

tacagactgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcaga 540

gtgtctgtgt cccagaccag caagctgacc agagctgaga cagtgttccc tgatgtggac 600

tatgtgaaca gcacagaggc tgagaccatc ctggacaaca tcacccagag cacccagtcc 660

ttcaatgact tcaccagagt ggtgggagga gaggatgcca agcctggcca gttcccctgg 720

caggtggtgc tgaatggcaa ggtggatgcc ttctgtggag gcagcattgt gaatgagaag 780

tggattgtga cagctgccca ctgtgtggag acaggagtga agatcacagt ggtggctgga 840

gagcacaaca ttgaggagac agagcacaca gagcagaaga ggaatgtgat caggatcatc 900

cctcaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag 960

ctggatgagc ctctggtgct gaacagctat gtgaccccta tctgcattgc tgacaaggag 1020

tacaccaaca tcttcctgaa gtttggctct ggctatgtgt ctggctgggg cagagtgttc 1080

cacaagggca ggtctgccct ggtgctgcag tacctgagag tgcccctggt ggacagagcc 1140

acctgcctgt tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggcttccat 1200

gagggaggca gagacagctg ccagggtgac tctggaggac cccatgtgac agaggtggag 1260 ggcaccagct tcctgacagg catcatcagc tggggagagg agtgtgccat gaagggcaag 1320

tatggcatct acaccaaagt gagcagatat gtgaactgga tcaaggagaa gaccaagctg 1380

acctga 1386

<210> 8

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> CS05-FL-NA

<400> 8

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc 120

ctgaaccgcc ccaagcgcta caactctggc aagctggagg agtttgtgca gggcaacctg 180

gagagggagt gcatggagga gaagtgcagc tttgaggagg ccagggaggt gtttgagaac 240

actgagcgca ccactgagtt ctggaagcag tatgtggatg gggaccagtg tgagagcaac 300

ccctgcctga atggggggag ctgcaaggat gacatcaaca gctatgagtg ctggtgcccc 360

tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggccgc 420

tgtgagcagt tctgcaagaa ctctgctgac aacaaggtgg tgtgctcttg cactgagggc 480

taccgcctgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcagg 540

gtgtctgtga gccagaccag caagctgacc agggctgaga ctgtgttccc tgacgtggac 600

tatgtgaaca gcactgaggc tgagaccatc ctggacaaca tcacccagag cacccagagc 660

ttcaatgact tcaccagggt ggtgggagga gaggatgcca agcctggcca gttcccctgg 720

caggtggtgc tgaatggcaa ggtggatgcc ttctgtggag gcagcattgt gaatgagaag 780

tggattgtga ccgctgccca ctgtgtggag actggagtga agatcactgt ggtggctggg 840

gagcacaaca ttgaggagac agagcacaca gagcagaagc gcaatgtgat caggatcatc 900

ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag 960

ctggatgagc ccctggtgct gaacagctac gtgaccccca tctgcattgc agacaaggag 1020

tacaccaaca tcttcctgaa gtttggctct ggctatgtgt ctggctgggg cagggtgttc 1080

cacaagggca ggtctgccct ggtgctgcag tacctgaggg tgcccctggt ggacagggcc 1140

acctgcctgc tgagcaccaa gttcaccatc tacaacaaca tgttctgcgc tggcttccat 1200

gagggaggaa gggacagctg ccagggagac tctggaggcc cccatgtgac agaggtggag 1260

ggcaccagct tcctgacagg catcatcagc tggggggagg agtgtgccat gaagggcaag 1320 tatggcatct acaccaaagt gtcccgctat gtgaactgga tcaaggagaa gaccaagctg 1380

acctga 1386

<210> 9

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> CS06-FL-NA

<400> 9

atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcactgtc ttcctggacc atgagaatgc caacaagatc 120

ctcaacaggc ccaagagata caactctggc aaactggagg agtttgtcca gggcaacctg 180

gagagggagt gcatggagga gaagtgctcc tttgaggagg ccagggaggt ctttgagaac 240

actgagcgca ccactgagtt ctggaaacag tatgtggatg ggaccagtg tgagtccaac 300

ccctgcctga atgggggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgcccc 360

tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcaga 420

tgtgagcagt tctgcaagaa ctctgctgac aacaaggtgg tgtgctcctg cactgagggc 480

taccgcctgg ctgagaacca gaagagctgt gagcctgctg tgccattccc atgtggcaga 540

gtctctgtga gccagaccag caagctcacc agggctgaga ctgtgttccc tgatgtggac 600

tatgtgaaca gcactgaggc tgaaaccatc ctggacaaca tcacccagag cacccagagc 660

ttcaatgact tcaccagagt ggtgggagga gaggatgcca agcctggcca gttcccctgg 720

caagtggtgc tcaatggcaa ggtggatgcc ttctgtgggg gctccattgt gaatgagaag 780

tggattgtca ctgctgccca ctgtgtggag actggggtca agatcactgt ggtggctggg 840

gagcacaaca ttgaggagac tgagcacact gagcagaagc gcaatgtgat caggatcatc 900

ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag 960

ctggatgagc cctggtcct caacagctat gtgacccca tctgcattgc tgacaaggag 1020

tacaccaaca tcttcctcaa gtttggctct ggctatgtct ctggctgggg cagagtgttc 1080

cacaaaggca ggtctgccct ggtgctccag tacctgagag tgcccctggt ggacagggcc 1140

acctgcctct tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggcttccat 1200

gagggaggaa gagacagctg ccagggggac tctggaggac cccatgtcac tgaggtggag 1260

ggcacctcct tcctcactgg catcatctcc tggggagagg agtgtgccat gaaaggcaaa 1320

tatggcatct acaccaaagt ctccagatat gtcaactgga tcaaggagaa gaccaagctg 1380

acctga 1386

<210> 10

<211> 415

<212> PRT

<213> Artificial Sequence

<220><223> FIX-MP-AA

<400> 10

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg
1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe
20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly
35 40 45

Asp Gln Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp
50 55 60

Asp Ile Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys
65 70 75 80

Asn Cys Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu
85 90 95

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

Glu Gly Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val
115 120 125

Pro Phe Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr
130 135 140

Arg Ala Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu
145 150 155 160

Ala Glu Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn
165 170 175

Asp Phe Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe
180 185 190

Pro Trp Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly

195 200 205

Ser Ile Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu

210 215 220

Thr Gly Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu

225 230 235 240

Thr Glu His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His

245 250 255

His Asn Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu

260 265 270

Leu Glu Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile

275 280 285

Cys Ile Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser

290 295 300

Gly Tyr Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala

305 310 315 320

Leu Val Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys

325 330 335

Leu Arg Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly

340 345 350

Phe His Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro

355 360 365

His Val Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser

370 375 380

Trp Gly Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys

385 390 395 400

Val Ser Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

405 410 415

<210> 11

<211> 377

<212> PRT

<213> Artificial Sequence

<220><223> FIX2-MP-AA

<400> 11

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg
1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe
20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Val
35 40 45

Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe Cys Lys Asn Ser
50 55 60

Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly Tyr Arg Leu Ala
65 70 75 80

Glu Asn Gln Lys Ser Cys Glu Pro Ala Val Pro Phe Pro Cys Gly Arg
85 90 95

Val Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala Glu Thr Val Phe
100 105 110

Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala Glu Thr Ile Leu Asp
115 120 125

Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp Phe Thr Arg Val Val
130 135 140

Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp Gln Val Val Leu
145 150 155 160

Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn Glu Lys
165 170 175

Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys Ile Thr
180 185 190

Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr Glu Gln
195 200 205

Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn Ala Ala
210 215 220

Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp Glu Pro
225 230 235 240

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

245 250 255

Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser Gly Trp

260 265 270

Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln Tyr Leu

275 280 285

Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg Ser Thr Lys Phe

290 295 300

Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly Gly Arg

305 310 315 320

Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu Val Glu

325 330 335

Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu Cys Ala

340 345 350

Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr Val Asn

355 360 365

Trp Ile Lys Glu Lys Thr Lys Leu Thr

370 375

<210> 12

<211> 415

<212> PRT

<213> Artificial Sequence

<220><223> FIXp-MP-AA

<400> 12

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg

1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe

20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly

35 40 45

Asp Gln Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp

50 55 60

Asp Ile Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys
65 70 75 80

Asn Cys Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu
85 90 95

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

Glu Gly Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val
115 120 125

Pro Phe Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr
130 135 140

Arg Ala Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu
145 150 155 160

Ala Glu Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn
165 170 175

Asp Phe Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe
180 185 190

Pro Trp Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly
195 200 205

Ser Ile Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu
210 215 220

Thr Gly Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu
225 230 235 240

Thr Glu His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His
245 250 255

His Asn Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu
260 265 270

Leu Glu Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile
275 280 285

Cys Ile Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser
290 295 300

Gly Tyr Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala

305 310 315 320

Leu Val Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys

325 330 335

Leu Leu Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly

340 345 350

Phe His Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro

355 360 365

His Val Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser

370 375 380

Trp Gly Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys

385 390 395 400

Val Ser Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

405 410 415

<210> 13

<211> 1248

<212> DNA

<213> Artificial Sequence

<220><223> CS02-MP-NA

<400> 13

tacaactcag gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagagag gtgtttgaga acacagagag gaccacagag 120

ttctggaagc agtatgtgga tggagaccag tgtgagagca acccttgcct gaatggaggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300

aactcagctg acaacaaagt ggtgtgtagc tgcacagagg gctacagact ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtcagt gtcccagacc 420

agcaagctga ccagagctga gacagtgttc cctgatgtgg actatgtgaa tagcacagag 480

gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga 540

gttgtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 600

aaagtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc 660

cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagaacacaa tattgaggag 720 acagagcaca cagagcagaa gaggaatgtc atcaggatta tcccccacca caactacaat 780

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 840

ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg 900

aagtttggct caggctatgt gtcaggctgg ggcagagtgt tccacaaggg cagatcagcc 960

ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc 1020

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc 1080

tgccagggag actcaggagg accccatgtg acagaagtgg agggcaccag cttcctgaca 1140

ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 1200

gtgagcagat atgtgaactg gatcaaggag aaaaccaagc tgacctga 1248

<210> 14

<211> 1248

<212> DNA

<213> Artificial Sequence

<220><223> CS03-MP-NA

<400> 14

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagggaa gtgtttgaga acactgagag gaccactgag 120

ttctggaagc agtatgtgga tggggaccag tgtgagagca acccttgcct gaatgggggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300

aactctgctg acaacaaagt ggtgtgtagc tgcactgagg gctacagact ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc 420

agcaagctga ccagagctga gactgtgttc cctgatgtgg actatgtgaa tagcactgag 480

gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga 540

gtggtggggg gggaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 600

aaagtggatg ccttctgtgg gggcagcatt gtgaatgaga agtggattgt gactgctgcc 660

cactgtgtgg agactggggt gaagatcact gtggtggctg gggaacacaa tattgaggag 720

actgagcaca ctgagcagaa gaggaatgtc atcaggatta tcccccacca caactacaat 780

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 840

ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg 900

aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc 960 ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gctgagcacc 1020

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggggg cagagacagc 1080

tgccaggggg actctggggg cccccatgtg actgaagtgg agggcaccag cttcctgact 1140

ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 1200

gtgagcaggt atgtgaactg gatcaaggag aaaaccaagc tgacctga 1248

<210> 15

<211> 1250

<212> DNA

<213> Artificial Sequence

<220><223> CS04-MP-NA

<400> 15

cstacaactc tggcaagctg gaggagtttg tgcagggcaa cctggagagg gagtgcatgg 60

aggagaagtg cagctttgag gaggccaggg aggtgtttga gaacacagag aggaccacag 120

agttctggaa gcagtatgtg gatggtgacc agtgtgagag caacccttgc ctgaatggag 180

gcagctgcaa ggatgacatc aacagctatg agtgctggtg cccttttggc tttgagggca 240

agaactgtga gctggatgtg acctgcaaca tcaagaatgg caggtgtgag cagttctgca 300

agaactctgc tgacaacaag gtggtgtgta gctgcacaga gggctacaga ctggctgaga 360

accagaagag ctgtgagcct gctgtgccct tcccctgtgg cagagtgtct gtgtcccaga 420

ccagcaagct gaccagagct gagacagtgt tccctgatgt ggactatgtg aacagcacag 480

aggctgagac catcctggac aacatcaccc agagcaccca gtccttcaat gacttcacca 540

gagtggtggg aggagaggat gccaagcctg gccagttccc ctggcaggtg gtgctgaatg 600

gcaaggtgga tgccttctgt ggaggcagca ttgtgaatga gaagtggatt gtgacagctg 660

cccactgtgt ggagacagga gtgaagatca cagtggtggc tggagagcac aacattgagg 720

agacagagca cacagagcag aagaggaatg tgatcaggat catccctcac cacaactaca 780

atgctgccat caacaagtac aaccatgaca ttgccctgct ggagctggat gagcctctgg 840

tgctgaacag ctatgtgacc cctatctgca ttgctgacaa ggagtacacc aacatcttcc 900

tgaagtttgg ctctggctat gtgtctggct ggggcagagt gttccacaag ggcaggtctg 960

ccctggtgct gcagtacctg agagtgcccc tggtggacag agccacctgc ctgttgagca 1020

ccaagttcac catctacaac aacatgttct gtgctggctt ccatgaggga ggcagagaca 1080

gctgccaggg tgactctgga ggaccccatg tgacagaggt ggagggcacc agcttcctga 1140

caggcatcat cagctgggga gaggagtgtg ccatgaaggg caagtatggc atctacacca 1200 aagtgagcag atatgtgaac tggatcaagg agaagaccaa gctgacctga 1250

<210> 16

<211> 1248

<212> DNA

<213> Artificial Sequence

<220><223> CS05-MP-NA

<400> 16

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagggag gtgtttgaga acactgagcg caccactgag 120

ttctggaagc agtatgtgga tggggaccag tgtgagagca acccctgcct gaatgggggg 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggcc gctgtgagca gttctgcaag 300

aactctgctg acaacaaggt ggtgtgctct tgcactgagg gctaccgcct ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc cctgtggca gggtgtctgt gagccagacc 420

agcaagctga ccagggctga gactgtgttc cctgacgtgg actatgtgaa cagcactgag 480

gctgagacca tcctggacaa catcacccag agcacccaga gcttcaatga cttcaccagg 540

gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 600

aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gaccgctgcc 660

cactgtgtgg agactggagt gaagatcact gtggtggctg gggagcacaa cattgaggag 720

acagagcaca cagagcagaa gcgcaatgtg atcaggatca tcccccacca caactacaat 780

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gccccctggtg 840

ctgaacagct acgtgacccc catctgcatt gcagacaagg agtacaccaa catcttcctg 900

aagtttggct ctggctatgt gtctggctgg ggcagggtgt tccacaaggg caggtctgcc 960

ctggtgctgc agtacctgag ggtgcccctg gtggacaggg ccacctgcct gctgagcacc 1020

aagttcacca tctacaacaa catgttctgc gctggcttcc atgagggagg aagggacagc 1080

tgccagggag actctggagg cccccatgtg acagaggtgg agggcaccag cttcctgaca 1140

ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 1200

gtgtcccgct atgtgaactg gatcaaggag aagaccaagc tgacctga 1248

<210> 17

<211> 1248

<212> DNA

<213> Artificial Sequence

<220><223> CS06-MP-NA

<400> 17

tacaactctg gcaaactgga ggagtttgtc cagggcaacc tggagaggga gtgcatggag 60

gagaagtgct cctttgagga ggccagggag gtctttgaga acactgagcg caccactgag 120

ttctggaaac agtatgtgga tggggaccag tgtgagtcca acccctgcct gaatgggggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca gatgtgagca gttctgcaag 300

aactctgctg acaacaaggt ggtgtgctcc tgcactgagg gctaccgcct ggctgagaac 360

cagaagagct gtgagcctgc tgtgccattc ccatgtggca gagtctctgt gagccagacc 420

agcaagctca ccagggctga gactgtgttc cctgatgtgg actatgtgaa cagcactgag 480

gctgaaacca tcctggacaa catcacccag agcacccaga gcttcaatga cttcaccaga 540

gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaagtggt gctcaatggc 600

aaggtggatg ccttctgtgg gggctccatt gtgaatgaga agtggattgt cactgctgcc 660

cactgtgtgg agactggggt caagatcact gtggtggctg gggagcacaa cattgaggag 720

actgagcaca ctgagcagaa gcgcaatgtg atcaggatca tcccccacca caactacaat 780

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcccctggtc 840

ctcaacagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctc 900

aagtttggct ctggctatgt ctctggctgg ggcagagtgt tccacaaagg caggtctgcc 960

ctggtgctcc agtacctgag agtgcccctg gtggacaggg ccacctgcct cttgagcacc 1020

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg aagagacagc 1080

tgccaggggg actctggagg accccatgtc actgaggtgg agggcacctc cttcctcact 1140

ggcatcatct cctggggaga ggagtgtgcc atgaaaggca aatatggcat ctacaccaaa 1200

gtctccagat atgtcaactg gatcaaggag aagaccaagc tgacctga 1248

<210> 18

<211> 138

<212> DNA

<213> Artificial Sequence

<220><223> FIX-PPP-NA

<400> 18

atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta 60 ggatatctac tcagtgctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt 120

ctgaatcggc caaagagg 138

<210> 19

<211> 138

<212> DNA

<213> Artificial Sequence

<220><223> CS02-PPP-NA

<400> 19

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtcagcaga gtgcacagtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagaga 138

<210> 20

<211> 138

<212> DNA

<213> Artificial Sequence

<220><223> CS03-PPP-NA

<400> 20

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagaga 138

<210> 21

<211> 138

<212> DNA

<213> Artificial Sequence

<220><223> CS04-PPP-NA

<400> 21

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcacagtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagaga 138

<210> 22

<211> 138

<212> DNA

<213> Artificial Sequence

<220><223> CS05-PPP-NA

<400> 22

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc 120

ctgaaccgcc ccaagcgc 138

<210> 23

<211> 138

<212> DNA

<213> Artificial Sequence

<220><223> CS06-PPP-NA

<400> 23

atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgcactgtc ttcctggacc atgagaatgc caacaagatc 120

ctcaacaggc ccaagaga 138

<210> 24

<211> 84

<212> DNA

<213> Artificial Sequence

<220><223> FIX-SP-NA

<400> 24

atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta 60

ggatatctac tcagtgctga atgt 84

<210> 25

<211> 84

<212> DNA

<213> Artificial Sequence

<220><223> CS02-SP-NA

<400> 25

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtcagcaga gtgc 84

<210> 26

<211> 84

<212> DNA

<213> Artificial Sequence

<220><223> CS03-SP-NA

<400> 26

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 27

<211> 84

<212> DNA

<213> Artificial Sequence

<220><223> CS04-SP-NA

<400> 27

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 28

<211> 84

<212> DNA

<213> Artificial Sequence

<220><223> CS05-SP-NA

<400> 28

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 29

<211> 84

<212> DNA

<213> Artificial Sequence

<220><223> CS06--SP-NA

<400> 29

atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 30

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> FIX-PP-NA

<400> 30

acagtttttc ttgatcatga aaacgccaac aaaattctga atcggccaaa gagg 54

<210> 31

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> CS02-PP-NA

<400> 31

acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga 54

<210> 32

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> CS03-PP-NA

<400> 32

actgtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga 54

<210> 33

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223>

CS04-PP-NA

<400> 33

acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga 54

<210> 34

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> CS05-PP-NA

<400> 34

actgtgttcc tggaccatga gaatgccaac aagatcctga accgccccaa gcgc 54

<210> 35

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> CS06-PP-NA

<400> 35

actgtcttcc tggaccatga gaatgccaac aagatcctca acaggcccaa gaga 54

<210> 36

<211> 46

<212> PRT

<213> Artificial Sequence

<220><223> FIX-PPP-AA

<400> 36

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu
20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg
35 40 45

<210> 37

<211> 28

<212> PRT

<213> Artificial Sequence

<220><223> FIX-SP-AA

<400

> 37

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys
20 25

<210> 38

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> FIX-PP-AA

<400> 38

Thr Val Phe Leu Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro

1   5   10   15

Lys Arg

<210> 39

<211

> 72

<212> DNA

<213> Artificial Sequence

<220><223> CRM8

<400> 39

gggggaggct gctggtgaat attaaccaag gtcaccccag ttatcggagg agcaaacagg   60

ggctaagtcc ac   72

<210> 40

<211> 5276

<212> DNA

<213> Artificial Sequence

<220><223> CS06-CRM8.3-ssV

<400> 40

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc   60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg   120

gccaactcca tcactagggg ttcctgagtt taaacttcgt cgacggggga ggctgctggt   180

gaatattaac caaggtcacc ccagttatcg gaggagcaaa caggggctaa gtccaccggg   240

ggaggctgct ggtgaatatt aaccaaggtc accccagtta tcggaggagc aaacaggggc   300

taagtccacc gggggaggct gctggtgaat attaaccaag gtcaccccag ttatcggagg   360

agcaaacagg ggctaagtcc accgagggca ctgggaggat gttgagtaag atggaaaact   420

actgatgacc cttgcagaga cagagtatta ggacatgttt gaacaggggc cgggcgatca   480

gcaggtagct ctagaggatc cccgtctgtc tgcacatttc gtagagcgag tgttccgata   540

ctctaatctc cctaggcaag gttcatattt gtgtaggtta cttattctcc ttttgttgac   600

taagtcaata atcagaatca gcaggtttgg agtcagcttg gcagggatca gcagcctggg   660

ttggaaggag gggtataaa agccccttca ccaggagaag ccgtcacaca gactaggcgc   720

gccctaaggt aagttggcgc cgtttaaggg atggttggtt ggtggggtat taatgtttaa   780

ttaccttttt tacaggcctg aagatctgcc accatgcaga gggtcaacat gatcatggct   840

gagtcccctg gcctcatcac catctgcctg ctgggctacc tgctgtctgc tgagtgcact 900
gtcttcctgg accatgagaa tgccaacaag atcctcaaca ggcccaagag atacaactct 960
ggcaaactgg aggagtttgt ccagggcaac ctggagaggg agtgcatgga ggagaagtgc 1020

tcctttgagg aggccaggga ggtctttgag aacactgagc gcaccactga gttctggaaa 1080
cagtatgtgg atggggacca gtgtgagtcc aacccctgcc tgaatggggg cagctgcaag 1140
gatgacatca acagctatga gtgctggtgc ccctttggct ttgagggcaa gaactgtgag 1200
ctggatgtga cctgcaacat caagaatggc agatgtgagc agttctgcaa gaactctgct 1260
gacaacaagg tggtgtgctc ctgcactgag ggctaccgcc tggctgagaa ccagaagagc 1320
tgtgagcctg ctgtgccatt cccatgtggc agagtctctg tgagccagac cagcaagctc 1380
accagggctg agactgtgtt ccctgatgtg gactatgtga acagcactga ggctgaaacc 1440

atcctggaca acatcaccca gagcacccag agcttcaatg acttcaccag agtggtggga 1500
ggagaggatg ccaagcctgg ccagttcccc tggcaagtgg tgctcaatgg caaggtggat 1560
gccttctgtg gggctccat tgtgaatgag aagtggattg tcactgctgc ccactgtgtg 1620
gagactgggg tcaagatcac tgtggtggct ggggagcaca acattgagga gactgagcac 1680
actgagcaga agcgcaatgt gatcaggatc atcccccacc acaactacaa tgctgccatc 1740
aacaagtaca accatgacat tgccctgctg gagctggatg agcccctggt cctcaacagc 1800
tatgtgaccc ccatctgcat tgctgacaag gagtacacca acatcttcct caagtttggc 1860

tctggctatg tctctggctg gggcagagtg ttccacaaag gcaggtctgc cctggtgctc 1920
cagtacctga gagtgcccct ggtggacagg gccacctgcc tcttgagcac caagttcacc 1980
atctacaaca acatgttctg tgctggcttc catgagggag gaagagacag ctgccagggg 2040
gactctggag gaccccatgt cactgaggtg gagggcacct ccttcctcac tggcatcatc 2100
tcctggggag aggagtgtgc catgaaaggc aaatatggca tctacaccaa agtctccaga 2160
tatgtcaact ggatcaagga gaagaccaag ctgacctgat gagcatgcct agagctcgct 2220
gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc tcccccgtgc 2280

cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg 2340
catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca 2400
agggggagga ttgggaagac aatagcaggc atgctgggga attaattaag ctcgcgaagg 2460
aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg 2520
ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag 2580
cgcgcagaga gggagtggcc aagacgattt aaatgacaag cttggcgtaa tcatggtcat 2640

agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa 2700

gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc 2760

gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc 2820

aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact 2880

cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac 2940

ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 3000

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg 3060

acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 3120

gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 3180

ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 3240

gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 3300

cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 3360

taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 3420

atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa 3480

cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 3540

cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 3600

ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 3660

ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 3720

tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 3780

aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 3840

tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg 3900

gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag 3960

atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt 4020

tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag 4080

ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt 4140

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca 4200

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg 4260

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 4320

ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta 4380 tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca 4440

gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 4500

taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 4560

cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa 4620

agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 4680

gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 4740

ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 4800

ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtctcg 4860

cgcgtttcgg tgatgacggt gaaaacctct gacacatgca gctcccggag acggtcacag 4920

cttgtctgta agcggatgcc gggagcagac aagcccgtca gggcgcgtca gcgggtgttg 4980

gcgggtgtcg gggctggctt aactatgcgg catcagagca gattgtactg agagtgcacc 5040

atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc aggcgccatt 5100

cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct tcgctattac 5160

gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt 5220

cccagtcacg acgttgtaaa acgacggcca gtgaattcct cgagatttaa atgacg 5276

<210> 41

<211> 702

<212> DNA

<213> Artificial Sequence

<220><223> CS02-HC-NA

<400> 41

gttgtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 60

aaagtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc 120

cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagaacacaa tattgaggag 180

acagagcaca cagagcagaa gaggaatgtc atcaggatta tcccccacca caactacaat 240

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 300

ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg 360

aagtttggct caggctatgt gtcaggctgg ggcagagtgt tccacaaggg cagatcagcc 420

ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc 480

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc 540

tgccagggag actcaggagg accccatgtg acagaagtgg agggcaccag cttcctgaca 600

ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 660 gtgagcagat atgtgaactg gatcaaggag aaaaccaagc tg 702

<210> 42

<211> 435

<212> DNA

<213> Artificial Sequence

<220><223> CS02-LC-NA

<400> 42

tacaactcag gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagagag gtgtttgaga acacagagag gaccacagag 120

ttctggaagc agtatgtgga tggagaccag tgtgagagca acccttgcct gaatggaggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300

aactcagctg acaacaaagt ggtgtgtagc tgcacagagg gctacagact ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtcagt gtcccagacc 420

agcaagctga ccaga 435

<210> 43

<211> 705

<212> DNA

<213> Artificial Sequence

<220><223> CS03-HC-NA

<400> 43

gtggtggggg gggaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 60

aaagtggatg ccttctgtgg gggcagcatt gtgaatgaga agtggattgt gactgctgcc 120

cactgtgtgg agactggggt gaagatcact gtggtggctg gggaacacaa tattgaggag 180

actgagcaca ctgagcagaa gaggaatgtc atcaggatta tcccccacca caactacaat 240

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 300

ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg 360

aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc 420

ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gctgagcacc 480

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggggg cagagacagc 540

tgccaggggg actctggggg cccccatgtg actgaagtgg agggcaccag cttcctgact 600 ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 660

gtgagcaggt atgtgaactg gatcaaggag aaaaccaagc tgacc 705

<210> 44

<211> 435

<212> DNA

<213> Artificial Sequence

<220><223> CS03-LC-NA

<400> 44

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagggaa gtgtttgaga acactgagag gaccactgag 120

ttctggaagc agtatgtgga tggggaccag tgtgagagca acccttgcct gaatgggggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300

aactctgctg acaacaaagt ggtgtgtagc tgcactgagg gctacagact ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc 420

agcaagctga ccaga 435

<210> 45

<211> 705

<212> DNA

<213> Artificial Sequence

<220><223> CS04-HC-NA

<400> 45

gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 60

aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc 120

cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagagcacaa cattgaggag 180

acagagcaca cagagcagaa gaggaatgtg atcaggatca tccctcacca caactacaat 240

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 300

ctgaacagct atgtgacccc tatctgcatt gctgacaagg agtacaccaa catcttcctg 360

aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc 420

ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc 480

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc 540 tgccagggtg actctggagg accccatgtg acagaggtgg agggcaccag cttcctgaca 600

ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 660

gtgagcagat atgtgaactg gatcaaggag aagaccaagc tgacc 705

<210> 46

<211> 435

<212> DNA

<213> Artificial Sequence

<220><223> CS04-LC-NA

<400> 46

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagggag gtgtttgaga acacagagag gaccacagag 120

ttctggaagc agtatgtgga tggtgaccag tgtgagagca acccttgcct gaatggaggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300

aactctgctg acaacaaggt ggtgtgtagc tgcacagagg gctacagact ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc 420

agcaagctga ccaga 435

<210> 47

<211> 705

<212> DNA

<213> Artificial Sequence

<220><223> CS05-HC-NA

<400> 47

gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 60

aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gaccgctgcc 120

cactgtgtgg agactggagt gaagatcact gtggtggctg gggagcacaa cattgaggag 180

acagagcaca cagagcagaa gcgcaatgtg atcaggatca tcccccacca caactacaat 240

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcccctggtg 300

ctgaacagct acgtgacccc catctgcatt gcagacaagg agtacaccaa catcttcctg 360

aagtttggct ctggctatgt gtctggctgg ggcagggtgt tccacaaggg caggtctgcc 420

ctggtgctgc agtacctgag ggtgcccctg gtggacaggg ccacctgcct gctgagcacc 480 aagttcacca tctacaacaa catgttctgc gctggcttcc atgagggagg aagggacagc 540

tgccagggag actctggagg cccccatgtg acagaggtgg agggcaccag cttcctgaca 600

ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa 660

gtgtcccgct atgtgaactg gatcaaggag aagaccaagc tgacc 705

<210> 48

<211> 435

<212> DNA

<213> Artificial Sequence

<220><223> CS05-LC-NA

<400> 48

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60

gagaagtgca gctttgagga ggccagggag gtgtttgaga acactgagcg caccactgag 120

ttctggaagc agtatgtgga tggggaccag tgtgagagca acccctgcct gaatgggggg 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggcc gctgtgagca gttctgcaag 300

aactctgctg acaacaaggt ggtgtgctct tgcactgagg gctaccgcct ggctgagaac 360

cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gggtgtctgt gagccagacc 420

agcaagctga ccagg 435

<210> 49

<211> 705

<212> DNA

<213> Artificial Sequence

<220><223> CS06-HC-NA

<400> 49

gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaagtggt gctcaatggc 60

aaggtggatg ccttctgtgg gggctccatt gtgaatgaga agtggattgt cactgctgcc 120

cactgtgtgg agactggggt caagatcact gtggtggctg gggagcacaa cattgaggag 180

actgagcaca ctgagcagaa gcgcaatgtg atcaggatca tcccccacca caactacaat 240

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcccctggtc 300

ctcaacagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctc 360

aagtttggct ctggctatgt ctctggctgg ggcagagtgt tccacaaagg caggtctgcc 420 ctggtgctcc agtacctgag agtgcccctg gtggacaggg ccacctgcct cttgagcacc 480

aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg aagagacagc 540

tgccaggggg actctggagg accccatgtc actgaggtgg agggcacctc cttcctcact 600

ggcatcatct cctggggaga ggagtgtgcc atgaaaggca aatatggcat ctacaccaaa 660

gtctccagat atgtcaactg gatcaaggag aagaccaagc tgacc 705

<210> 50

<211> 435

<212> DNA

<213> Artificial Sequence

<220><223> CS06-LC-NA

<400> 50

tacaactctg gcaaactgga ggagtttgtc cagggcaacc tggagaggga gtgcatggag 60

gagaagtgct cctttgagga ggccagggag gtctttgaga acactgagcg caccactgag 120

ttctggaaac agtatgtgga tggggaccag tgtgagtcca acccctgcct gaatgggggc 180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag 240

aactgtgagc tggatgtgac ctgcaacatc aagaatggca gatgtgagca gttctgcaag 300

aactctgctg acaacaaggt ggtgtgctcc tgcactgagg gctaccgcct ggctgagaac 360

cagaagagct gtgagcctgc tgtgccattc ccatgtggca gagtctctgt gagccagacc 420

agcaagctca ccagg 435

<210> 51

<211> 145

<212> DNA

<213> Artificial Sequence

<220><223> AAV2_5'-ITR

<400> 51

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcct 145

<210> 52

<211> 330

<212> DNA

<213> Artificial Sequence

<220><223> truncated_TTR_enhancer/promoter

<400> 52

cgagggcact gggaggatgt tgagtaagat ggaaaactac tgatgaccct tgcagagaca 60

gagtattagg acatgtttga acaggggccg ggcgatcagc aggtagctct agaggatccc 120

cgtctgtctg cacatttcgt agagcgagtg ttccgatact ctaatctccc taggcaaggt 180

tcatatttgt gtaggttact tattctcctt ttgttgacta agtcaataat cagaatcagc 240

aggtttggag tcagcttggc agggatcagc agcctgggtt ggaaggaggg ggtataaaag 300

ccccttcacc aggagaagcc gtcacacaga 330

<210> 53

<211> 77

<212> DNA

<213> Artificial Sequence

<220><223> MVM_intron

<400> 53

ctaaggtaag ttggcgccgt ttaagggatg gttggttggt ggggtattaa tgtttaatta 60

cctttttac aggcctg 77

<210> 54

<211> 234

<212> DNA

<213> Artificial Sequence

<220><223> BGH_poly-adenylation_signal

<400> 54

cctagagctc gctgatcagc ctcgactgtg ccttctagtt gccagccatc tgttgtttgc 60

ccctcccccg tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa 120

aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg gggtggggtg 180

gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggaa 234

<210> 55

<211> 145

<212> DNA

<213> Artificial Sequence

<220><223> AAV2 3'-ITR

<400> 55

aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaa 145

<210> 56

<211> 35

<212> PRT

<213> Artificial Sequence

<220><223> FIX-AP-AA

<400> 56

Ala Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala

1 5 10 15

Glu Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp

20 25 30

Phe Thr Arg

35

<210> 57

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> CS02-AP-NA

<400> 57

gctgagacag tgttccctga tgtggactat gtgaatagca cagaggctga gaccatcctg 60

gacaacatca cccagagcac ccagtccttc aatgacttca ccaga 105

<210> 58

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> CS03-AP-NA

<400> 58

gctgagactg tgttccctga tgtggactat gtgaatagca ctgaggctga gaccatcctg 60

gacaacatca cccagagcac ccagtccttc aatgacttca ccaga 105

<210> 59

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> CS04-AP-NA

<400> 59

gctgagacag tgttccctga tgtggactat gtgaacagca cagaggctga gaccatcctg 60

gacaacatca cccagagcac ccagtccttc aatgacttca ccaga 105

<210> 60

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> CS05-AP-NA

<400> 60

gctgagactg tgttccctga cgtggactat gtgaacagca ctgaggctga gaccatcctg 60

gacaacatca cccagagcac ccagagcttc aatgacttca ccagg 105

<210> 61

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> CS06-AP-NA

<400> 61

gctgagactg tgttccctga tgtggactat gtgaacagca ctgaggctga aaccatcctg 60

gacaacatca cccagagcac ccagagcttc aatgacttca ccaga 105

<210> 62

<211> 145

<212> PRT

<213> Artificial Sequence

<220><223> FIX-LC-AA

<400> 62

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg

1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe

20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly

35 40 45

Asp Gln Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp

50 55 60

Asp Ile Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys

65 70 75 80

Asn Cys Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu

85 90 95

Gln Phe Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr

100 105 110

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

115 120 125

Pro Phe Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr

130 135 140

Arg

145

<210> 63

<211> 235

<212> PRT

<213> Artificial Sequence

<220><223> FIX-HC-AA

<400> 63

Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp Gln Val

1 5 10 15

Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn

20 25 30

Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys

35 40 45

Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr

50 55 60

Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn

65 70 75 80

Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp

85 90 95

Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile Ala Asp

100 105 110

Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser

115 120 125

Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln

130 135 140

Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg Ser Thr

145 150 155 160

Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly

165 170 175

Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu

180 185 190

Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu

195 200 205

Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr

210 215 220

Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

225 230 235

<210> 64

<211> 235

<212> PRT

<213> Artificial Sequence

<220><223> FIXp-HC-AA

<400> 64

Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp Gln Val

1 5 10 15

Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn

20 25 30

Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys

35 40 45

Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr

50 55 60

Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn

65 70 75 80

Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp

85 90 95

Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile Ala Asp

100 105 110

Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser

115 120 125

Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln

130 135 140

Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Leu Ser Thr

145 150 155 160

Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly

165 170 175

Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu

180 185 190

Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu

195 200 205

Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr

210 215 220

Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

225 230 235