



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년05월31일
(11) 등록번호 10-2404550
(24) 등록일자 2022년05월27일

- (51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) A61K 48/00 (2006.01)
C07K 14/755 (2006.01)
- (52) CPC특허분류
C12N 15/86 (2013.01)
A61K 48/0008 (2013.01)
- (21) 출원번호 10-2018-7015191
- (22) 출원일자(국제) 2016년11월11일
심사청구일자 2021년11월11일
- (85) 번역문제출일자 2018년05월29일
- (65) 공개번호 10-2018-0070700
- (43) 공개일자 2018년06월26일
- (86) 국제출원번호 PCT/US2016/061688
- (87) 국제공개번호 WO 2017/083764
국제공개일자 2017년05월18일
- (30) 우선권주장
62/255,323 2015년11월13일 미국(US)
- (56) 선행기술조사문헌
US20150071883 A1
US20130017997 A1

- (73) 특허권자
다케다 야쿠힌 교교 가부시키키가이샤
일본 오사카시 주오구 도쇼마찌 4-1-1
- (72) 발명자
팔크너 팔코-컨터
오스트리아 2304 오르트/도나우 노이시에틀자일 76아
호링 프란치스카
오스트리아 2230 겐세른도르프 바움펠트가쎄 3/9
(뒷면에 계속)
- (74) 대리인
양영준

전체 청구항 수 : 총 10 항

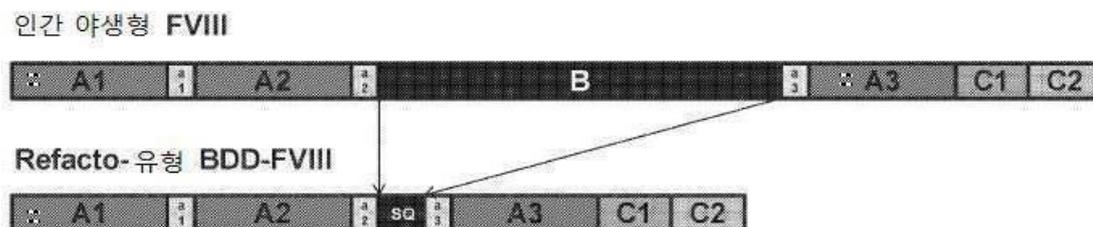
심사관 : 손영희

(54) 발명의 명칭 **A형 혈우병의 유전자 요법을 위한 증가된 발현을 갖는 재조합 FVIII 변이체를 인코딩하는 바이러스 벡터**

(57) 요약

본 발명은 다른 양상 중에서, 포유류 세포에서 발현을 위한 인자 VIII 변이체를 인코딩하는 코돈-변경된 폴리뉴클레오티드를 제공한다. 일부 구체예에서, 본 발명은 또한, A형 혈우병을 치료하기 위한 포유류 유전자 요법 벡터 및 방법을 제공한다.

대표도



(52) CPC특허분류

A61K 48/0058 (2013.01)

A61K 48/0066 (2013.01)

C07K 14/755 (2013.01)

C12N 2750/14142 (2013.01)

C12N 2750/14143 (2013.01)

C12N 2840/007 (2013.01)

(72) 발명자

렝글러 조하네스

오스트리아 1020 비엔나 슈로츠베르크스트라쎄
2/23

로텐스타이너 한스페터

오스트리아 1020 비엔나 하이트가쎄 10/17

샤이프링거 프리드리히

오스트리아 1090 비엔나 미셀바움가쎄 4/17

명세서

청구범위

청구항 1

서열 번호: 1의 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드에 있어서, 상기 폴리뉴클레오티드는 인자 VIII 폴리펩티드를 인코딩하는 것을 특징으로 하는 폴리뉴클레오티드.

청구항 2

청구항 1에 있어서, 인자 VIII 폴리펩티드를 인코딩하는 폴리뉴클레오티드에 작동가능하게 연결된 프로모터 요소를 더욱 포함하는 것을 특징으로 하는 폴리뉴클레오티드.

청구항 3

청구항 2에 있어서, 프로모터 요소는 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열의 상류에 간-특이적 프로모터 서열인 것을 특징으로 하는 폴리뉴클레오티드.

청구항 4

청구항 3에 있어서, 간-특이적 프로모터 서열 및 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열 사이에 배치된 인트론 서열을 더욱 포함하는 것을 특징으로 하는 폴리뉴클레오티드.

청구항 5

청구항 1 내지 4 중에서 어느 한 항에 따른 폴리뉴클레오티드를 포함하는 아데노 연관된 바이러스 (AAV) 벡터.

청구항 6

청구항 1 내지 4 중에서 어느 한 항에 따른 폴리뉴클레오티드를 포함하는 아데노 연관된 바이러스 (AAV) 입자.

청구항 7

청구항 1 내지 4 중에서 어느 한 항에 따른 폴리뉴클레오티드를 포함하는 아데노 연관된 바이러스 (AAV) 입자로 감염된 단리된 숙주 세포.

청구항 8

청구항 1 내지 4 중에서 어느 한 항에 따른 폴리뉴클레오티드를 포유류 숙주 세포 내로 도입하는 것을 포함하고, 상기 폴리뉴클레오티드는 포유류 숙주 세포에서 복제에 적격성인 것인, 아데노 연관된 바이러스 (AAV) 입자를 생산하기 위한 시험관내 방법.

청구항 9

청구항 6에 따른 아데노 연관된 바이러스 (AAV) 입자를 포함하는, A형 혈우병을 치료하기 위한 제약 조성물.

청구항 10

숙주 세포를 청구항 6에 따른 아데노 연관된 바이러스 (AAV) 입자와 접촉시키는 것을 포함하는, 숙주 세포를 형질도입하기 위한 시험관내 방법.

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

- 청구항 29
- 삭제
- 청구항 30
- 삭제
- 청구항 31
- 삭제
- 청구항 32
- 삭제
- 청구항 33
- 삭제
- 청구항 34
- 삭제
- 청구항 35
- 삭제
- 청구항 36
- 삭제
- 청구항 37
- 삭제
- 청구항 38
- 삭제
- 청구항 39
- 삭제
- 청구항 40
- 삭제
- 청구항 41
- 삭제
- 청구항 42
- 삭제
- 청구항 43
- 삭제
- 청구항 44
- 삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

발명의 설명

기술 분야

[0001] 관련된 출원에 대한 교차 참조

[0002] 본 출원은 2015년 11월 13일자 제출된 US 특허가출원 번호 62/255,323에 우선권을 주장하고, 이것은 전체적으로 본원에 참조로서 편입된다.

[0003] 서열 목록

[0004] 본 출원은 서열 목록을 내포하는데, 이것은 ASCII 형식으로 전자적으로 제출되었고 전체적으로 본원에 참조로서 편입된다. 2016년 11월 7일자에 창출된 상기 ASCII 사본은 008073_5115_W0_Sequence_Listing.txt로 명명되고 크기에서 183,311 바이트이다.

배경 기술

[0005] 발명의 배경

[0006] 혈액 응고는 응고 연쇄 반응으로서 지칭된, 상호의존성 생화학적 반응의 복합적이고 동적인 생물학적 경로를 통해 진행된다. 응고 인자 VIII (FVIII)는 이러한 연쇄 반응에서 핵심 성분이다. 인자 VIII는 출혈 부위로 모집되고, 그리고 활성화된 인자 IX (FIXa) 및 인자 X (FX)와 Xase 복합체를 형성한다. Xase 복합체는 FX를 활성화시키고, 이것은 차례로, 프로트롬빈을 트롬빈으로 활성화시키고, 이것은 이후, 응고 연쇄 반응에서 다른 성분을 활성화시켜 안정된 응괴를 산출한다 (Saenko *et al.*, *Trends Cardiovasc. Med.*, 9:185-192 (1999); Lenting *et al.*, *Blood*, 92:3983-3996 (1998)에서 리뷰됨).

[0007] A형 혈우병은 인자 VIII 활성화에서 결함에 의해 특징되는 선천성 X 연관된 출혈 장애이다. 축소된 인자 VIII 활성화는 응고 연쇄 반응에서 양성 피드백 루프를 저해한다. 이것은 불완전 응고를 유발하고, 이것은 증가된 지속 기간을 갖는 출혈 에피소드, 광범위한 타박상, 자발적 입과 코 출혈, 관절 경직 및 만성 통증, 그리고 아마도 심각한 사례에서 내출혈 및 빈혈로서 현성한다 (Zhang *et al.*, *Clinic. Rev. Allerg. Immunol.*, 37:114-124 (2009)).

[0008] 전통적으로, A형 혈우병은 인자 VIII 보상 요법에 의해 치료되는데, 이것은 인자 VIII 단백질 (가령, 혈장-유래된 또는 재조합적으로-생산된 인자 VIII)을 A형 혈우병을 앓는 개체에 투여하는 것으로 구성된다. 인자 VIII는 급성 출혈 에피소드에 대한 응답으로 및/또는 수술 동안 출혈을 관리하기 위해 수술전후에 출혈 에피소드의 빈도를 예방하거나 또는 감소시키기 위해 예방적으로 투여된다. 하지만, 인자 VIII 보상 요법의 여러 바람직하지 않은 특질이 있다.

[0009] 첫 번째, 인자 VIII 보상 요법은 A형 혈우병을 치료하거나 또는 관리하는데 이용되지만, 근원적인 인자 VIII 결함을 치유하지 못한다. 이것 때문에, A형 혈우병을 앓는 개체는 그들의 삶의 지속 기간 동안 인자 VIII 보상 요

법을 필요로 한다. 연속 치료는 값비싸고, 그리고 단지 몇 번의 예방적 용량의 결여도 중증도 A형 혈우병을 앓는 개체에게 심각한 결과를 초래할 수 있기 때문에, 개체가 엄격한 순응도를 유지하는 것을 필요로 한다.

[0010] 두 번째, 인자 VIII가 생체내에서 상대적으로 짧은 반감기를 갖기 때문에, 전통적인 예방적 인자 VIII 보상 요법은 2 일 또는 3 일 마다 투여를 필요로 한다. 이것은 개체에게 그들의 일생 동안 순응도를 유지해야 하는 부담을 준다. 비록 3세대 "지속성" 인자 VIII 약물이 투여 빈도를 감소시킬 수 있긴 하지만, 이들 약물로 예방적 인자 FVIII 보상 요법은 여전히 영구적으로 월 1회, 주 1회, 또는 더욱 빈번한 투여를 필요로 한다. 가령, ELOCTATE™ [항혈우병 인자 (재조합), Fc 융합 단백질]로 예방적 처치는 3 일 내지 5 일 마다 투여를 필요로 한다 (ELOCTATE™ Prescribing Information, Biogen Idec Inc., (2015)). 게다가, 화학적으로 변형된 생물체제 (가령, 폐길화된 폴리펩티드)의 장기간 효과는 아직 완전히 이해되지는 않고 있다.

[0011] 세 번째, 인자 VIII 보상 요법을 제공받는 모든 개체 중에서 15% 및 30% 사이가 항-인자 VIII 저해제 항체를 형성하고, 상기 요법을 비효율적으로 만든다. 인자 VIII 우회로 요법 (가령, 혈장-유래된 또는 재조합적으로-생산된 프로트롬빈 복합체 농축물의 투여)이 저해제 항체를 형성하는 개체에서 혈우병을 치료하는데 이용될 수 있다. 하지만, 인자 VIII 우회로 요법은 인자 VIII 보상 요법보다 덜 효과적이고 (Mannucci P.M., J Thromb Haemost., 1(7):1349-55 (2003)), 그리고 심혈관 합병증의 증가된 위험과 연관될 수 있다 (Luu and Ewenstein, Haemophilia, 10 Suppl. 2:10-16 (2004)).

[0012] 체세포 유전자 요법은 A형 혈우병의 치료에 대한 큰 희망을 여전히 주고 있는데, 그 이유는 이것이 인자 VIII 활성의 1회 용량을 개체에게 제공하기 보다는, 근원적인 과소발현 기능적 인자 VIII 활성 (가령, 미스센스 또는 넌센스 돌연변이로 인해)을 치료할 것이기 때문이다. 인자 VIII 보상 요법과 비교하여, 작용 기전에서 이러한 차이 때문에, 인자 VIII 유전자 요법 벡터의 1회 투여는 인자 VIII를 수년 동안 개체에 제공하고, 치료 비용을 감소시키고, 그리고 지속된 환자 순응도에 대한 요구를 제거할 수 있다.

[0013] 응고 인자 IX (FIX) 유전자 요법은 축소된 인자 IX 활성에 의해 특징되는 관련된 혈액 응고 장애인 B형 혈우병을 앓는 개체를 치료하는데 효과적으로 이용되었다 (Manno C.S., et al., Nat Med., 12(3):342-47 (2006)). 하지만, 인자 VIII 유전자 요법은 여러 독특한 과제를 제시한다. 가령, 전장, 야생형 인자 VIII 폴리펩티드 (2351개 아미노산; UniProt 수탁 번호 P00451)는 전장, 야생형 인자 IX 폴리펩티드 (461개 아미노산; UniProt 수탁 번호 P00740)보다 5배 크다. 따라서, 야생형 인자 VIII의 코딩 서열은 7053개 염기쌍인데, 이것은 전통적인 AAV 유전자 요법 벡터에서 포장되기에는 너무 크다. 게다가, 인자 VIII의 B-도메인 결실된 변이체의 보고된 재조합 발현 (BDD-FVIII)이 불량하였다. 따라서, 여러 그룹이 BDD-FVIII 구조체의 코돈 사용빈도를 변경하는 것을 시도하였지만, 성공은 제한적이었다.

발명의 내용

해결하려는 과제

[0014] **발명의 짧은 요약**

[0015] 따라서, 코딩 서열이 유전자 요법 벡터 내로 더욱 효율적으로 포장되고, 그리고 이들을 통해 전달되는 인자 VIII 변이체가 요구된다. 인자 VIII를 더욱 효율적으로 발현하는 합성, 코돈-변경된 핵산이 또한 필요하다. 이런 인자 VIII 변이체 및 코돈-변경된 핵산은 인자 VIII 결함 (가령, A형 혈우병)의 향상된 치료를 허용한다. 상기 결함 및 인자 VIII 결함 (가령, A형 혈우병)의 치료와 연관된 다른 문제는 개시된 코돈-변경된 인자 VIII 변이체에 의해 감소되거나 또는 제거된다.

[0016] 일부 구체예에 따라서, 본 발명은 인자 VIII 중쇄 (가령, CS01-HC-NA, CS04-HC-NA, 또는 CS23-HC-NA) 및 경쇄 (CS01-LC-NA, CS04-LC-NA, 또는 CS23-LC-NA)의 개시된 코돈-변경된 서열에 높은 서열 동일성을 갖는 인자 VIII 변이체를 인코딩하는 핵산을 제공한다. 일부 구체예에서, 이들 핵산은 인자 VIII 중쇄와 경쇄를 코딩하는 서열 사이에, 선천적 인자 VIII B-도메인을 대체하는 링커 서열 (가령, 퓨린 개열 부위를 포함하는 링커 서열)을 인코딩하는 서열을 더욱 포함한다.

[0017] 한 양상에서, 본 발명은 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공한다. 인자 VIII 폴리펩티드는 경쇄, 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 CS04-HC-NA (서열 번호: 3)와 최소한 95% 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 FVIII 폴리펩티드의 경쇄는 CS04-LC-NA (서열 번호: 4)와 최소한 95% 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 퓨린 개열 부위를

포함한다.

- [0018] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리펩티드 링커는 BDL004 (서열 번호: 6)와 최소한 95% 동일성을 갖는 세 번째 뉴클레오티드 서열에 의해 인코딩된다.
- [0019] 한 양상에서, 본 발명은 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공한다. 인자 VIII 폴리펩티드는 경쇄, 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 CS01-HC-NA (서열 번호: 24)와 최소한 95% 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 FVIII 폴리펩티드의 경쇄는 CS01-LC-NA (서열 번호: 25)와 최소한 95% 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 퓨린 개열 부위를 포함한다.
- [0020] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리펩티드 링커는 BDL001 (서열 번호: 5)과 최소한 95% 동일성을 갖는 세 번째 뉴클레오티드 서열에 의해 인코딩된다.
- [0021] 한 양상에서, 본 발명은 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공한다. 인자 VIII 폴리펩티드는 경쇄, 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 CS23-HC-NA (서열 번호: 22)와 최소한 95% 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 FVIII 폴리펩티드의 경쇄는 CS23-LC-NA (서열 번호: 23)와 최소한 95% 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 퓨린 개열 부위를 포함한다.
- [0022] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리펩티드 링커는 BDL023 (서열 번호: 7)과 최소한 95% 동일성을 갖는 세 번째 뉴클레오티드 서열에 의해 인코딩된다.
- [0023] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 개별 중쇄 서열 (가령, CS04-HC-NA (서열 번호: 3), CS01-HC-NA (서열 번호: 24), 또는 CS23-HC-NA (서열 번호: 22))와 최소한 96% 동일성을 갖고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 개별 경쇄 서열 (가령, CS04-LC-NA (서열 번호: 4), CS01-LC-NA (서열 번호: 25), 또는 CS23-LC-NA (서열 번호: 23))와 최소한 96% 동일성을 갖는다.
- [0024] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 개별 중쇄 서열 (가령, CS04-HC-NA (서열 번호: 3), CS01-HC-NA (서열 번호: 24), 또는 CS23-HC-NA (서열 번호: 22))와 최소한 97% 동일성을 갖고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 개별 경쇄 서열 (가령, CS04-LC-NA (서열 번호: 4), CS01-LC-NA (서열 번호: 25), 또는 CS23-LC-NA (서열 번호: 23))와 최소한 97% 동일성을 갖는다.
- [0025] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 개별 중쇄 서열 (가령, CS04-HC-NA (서열 번호: 3), CS01-HC-NA (서열 번호: 24), 또는 CS23-HC-NA (서열 번호: 22))와 최소한 98% 동일성을 갖고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 개별 경쇄 서열 (가령, CS04-LC-NA (서열 번호: 4), CS01-LC-NA (서열 번호: 25), 또는 CS23-LC-NA (서열 번호: 23))와 최소한 98% 동일성을 갖는다.
- [0026] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 개별 중쇄 서열 (가령, CS04-HC-NA (서열 번호: 3), CS01-HC-NA (서열 번호: 24), 또는 CS23-HC-NA (서열 번호: 22))와 최소한 99% 동일성을 갖고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 개별 경쇄 서열 (가령, CS04-LC-NA (서열 번호: 4), CS01-LC-NA (서열 번호: 25), 또는 CS23-LC-NA (서열 번호: 23))와 최소한 99% 동일성을 갖는다.
- [0027] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 개별 중쇄 서열 (가령, CS04-HC-NA (서열 번호: 3), CS01-HC-NA (서열 번호: 24), 또는 CS23-HC-NA (서열 번호: 22))와 최소한 99.5% 동일성을 갖고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 개별 경쇄 서열 (가령, CS04-LC-NA (서열 번호: 4), CS01-LC-NA (서열 번호: 25), 또는 CS23-LC-NA (서열 번호: 23))와 최소한 99.5% 동일성을 갖는다.
- [0028] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 개별 중쇄 서열 (가령, CS04-HC-NA (서열 번호: 3), CS01-HC-NA (서열 번호: 24), 또는 CS23-HC-

NA (서열 번호: 22))와 최소한 99.9% 동일성을 갖고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 개별 경쇄 서열 (가령, CS04-LC-NA (서열 번호: 4), CS01-LC-NA (서열 번호: 25), 또는 CS23-LC-NA (서열 번호: 23))와 최소한 99.9% 동일성을 갖는다.

- [0029] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 CS04-HC-NA (서열 번호: 3)이고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 CS04-LC-NA (서열 번호: 4)이다.
- [0030] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 CS01-HC-NA (서열 번호: 24)이고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 CS01-LC-NA (서열 번호: 25)이다.
- [0031] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인자 VIII 폴리펩티드의 중쇄를 인코딩하는 첫 번째 뉴클레오티드 서열은 CS23-HC-NA (서열 번호: 22)이고, 그리고 인자 FVIII 폴리펩티드의 경쇄를 인코딩하는 두 번째 뉴클레오티드 서열은 CS23-LC-NA (서열 번호: 23)이다.
- [0032] 한 양상에서, 본 발명은 CS04-FL-NA와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 인자 VIII 폴리펩티드를 인코딩한다.
- [0033] 한 양상에서, 본 발명은 CS01-FL-NA와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 인자 VIII 폴리펩티드를 인코딩한다.
- [0034] 한 양상에서, 본 발명은 CS23-FL-NA와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 인자 VIII 폴리펩티드를 인코딩한다.
- [0035] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-FL-NA (서열 번호: 1), CS01-FL-NA (서열 번호: 13), 또는 CS23-FL-NA (서열 번호: 20))와 최소한 96% 동일성을 갖는다.
- [0036] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-FL-NA (서열 번호: 1), CS01-FL-NA (서열 번호: 13), 또는 CS23-FL-NA (서열 번호: 20))와 최소한 97% 동일성을 갖는다.
- [0037] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-FL-NA (서열 번호: 1), CS01-FL-NA (서열 번호: 13), 또는 CS23-FL-NA (서열 번호: 20))와 최소한 98% 동일성을 갖는다.
- [0038] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-FL-NA (서열 번호: 1), CS01-FL-NA (서열 번호: 13), 또는 CS23-FL-NA (서열 번호: 20))와 최소한 99% 동일성을 갖는다.
- [0039] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-FL-NA (서열 번호: 1), CS01-FL-NA (서열 번호: 13), 또는 CS23-FL-NA (서열 번호: 20))와 최소한 99.5% 동일성을 갖는다.
- [0040] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-FL-NA (서열 번호: 1), CS01-FL-NA (서열 번호: 13), 또는 CS23-FL-NA (서열 번호: 20))와 최소한 99.9% 동일성을 갖는다.
- [0041] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)이다.
- [0042] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)이다.
- [0043] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)이다.
- [0044] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 95% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0045] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 96% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.

- [0046] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 97% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0047] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 98% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0048] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 99% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0049] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 99.5% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0050] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)와 최소한 99.9% 동일성을 갖는 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0051] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 CS04-FL-AA (서열 번호: 2)의 아미노산 서열을 포함하는 인자 VIII 폴리펩티드를 인코딩한다.
- [0052] 한 양상에서, 본 발명은 CS04-SC1-NA (서열 번호: 9)와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 단일 사슬 인자 VIII 폴리펩티드를 인코딩한다.
- [0053] 한 양상에서, 본 발명은 CS04-SC2-NA (서열 번호: 11)와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 단일 사슬 인자 VIII 폴리펩티드를 인코딩한다.
- [0054] 한 양상에서, 본 발명은 CS01-SC1-NA (서열 번호: 26)와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 단일 사슬 인자 VIII 폴리펩티드를 인코딩한다.
- [0055] 한 양상에서, 본 발명은 CS01-SC2-NA (서열 번호: 27)와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 단일 사슬 인자 VIII 폴리펩티드를 인코딩한다.
- [0056] 한 양상에서, 본 발명은 CS23-SC1-NA (서열 번호: 28)와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 단일 사슬 인자 VIII 폴리펩티드를 인코딩한다.
- [0057] 한 양상에서, 본 발명은 CS23-SC2-NA (서열 번호: 29)와 최소한 95% 동일성을 갖는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 제공하고, 여기서 상기 폴리뉴클레오티드는 단일 사슬 인자 VIII 폴리펩티드를 인코딩한다.
- [0058] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-SC1-NA (서열 번호: 9), CS04-SC2-NA (서열 번호: 11), CS01-SC1-NA (서열 번호: 26), CS01-SC2-NA (서열 번호: 27), CS23-SC1-NA (서열 번호: 28), 또는 CS23-SC2-NA (서열 번호: 29))와 최소한 96% 동일성을 갖는다.
- [0059] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-SC1-NA (서열 번호: 9), CS04-SC2-NA (서열 번호: 11), CS01-SC1-NA (서열 번호: 26), CS01-SC2-NA (서열 번호: 27), CS23-SC1-NA (서열 번호: 28), 또는 CS23-SC2-NA (서열 번호: 29))와 최소한 97% 동일성을 갖는다.
- [0060] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-SC1-NA (서열 번호: 9), CS04-SC2-NA (서열 번호: 11), CS01-SC1-NA (서열 번호: 26), CS01-SC2-NA (서열 번호: 27), CS23-SC1-NA (서열 번호: 28), 또는 CS23-SC2-NA (서열 번호: 29))와 최소한 98% 동일성을 갖는다.
- [0061] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-SC1-NA (서열 번호: 9), CS04-SC2-NA (서열 번호: 11), CS01-SC1-NA (서열 번호: 26), CS01-SC2-NA (서열 번호: 27), CS23-SC1-NA (서열 번호: 28), 또는 CS23-SC2-NA (서열 번호: 29))와 최소한 99% 동일성을 갖는다.
- [0062] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-SC1-NA (서열 번호: 9), CS04-SC2-NA (서열 번호: 11), CS01-SC1-NA (서열 번호: 26), CS01-SC2-NA (서열

번호: 27), CS23-SC1-NA (서열 번호: 28), 또는 CS23-SC2-NA (서열 번호: 29))와 최소한 99.5% 동일성을 갖는다.

- [0063] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 개별 전장 폴리뉴클레오티드 서열 (가령, CS04-SC1-NA (서열 번호: 9), CS04-SC2-NA (서열 번호: 11), CS01-SC1-NA (서열 번호: 26), CS01-SC2-NA (서열 번호: 27), CS23-SC1-NA (서열 번호: 28), 또는 CS23-SC2-NA (서열 번호: 29))와 최소한 99.9% 동일성을 갖는다.
- [0064] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)이다.
- [0065] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)이다.
- [0066] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)이다.
- [0067] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)이다.
- [0068] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)이다.
- [0069] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)이다.
- [0070] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 95% 동일성을 갖는다.
- [0071] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 96% 동일성을 갖는다.
- [0072] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 97% 동일성을 갖는다.
- [0073] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 98% 동일성을 갖는다.
- [0074] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 99% 동일성을 갖는다.
- [0075] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 99.5% 동일성을 갖는다.
- [0076] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택되는 서열과 최소한 99.5% 동일성을 갖는다.
- [0077] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01-SC2-NA, CS04-SC2-NA 및 CS23-SC2-NA로 구성된 군에서 선택된다.
- [0078] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인코딩된 인자 VIII 폴리펩티드는 2개 연속 아미노산 사이에

배치된 글리코실화 폴리펩티드를 포함한다.

- [0079] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 또한, 인자 VIII 폴리펩티드를 인코딩하는 폴리뉴클레오티드에 작동가능하게 연결된 프로모터 요소를 포함한다.
- [0080] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 또한, 인자 VIII 폴리펩티드를 인코딩하는 폴리뉴클레오티드에 작동가능하게 연결된 인핸서 요소를 포함한다.
- [0081] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 또한, 인자 VIII 폴리펩티드를 인코딩하는 폴리뉴클레오티드에 작동가능하게 연결된 폴리아데닐화 요소를 포함한다.
- [0082] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 폴리뉴클레오티드는 또한, 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열에 작동가능하게 연결된 인트론을 포함한다.
- [0083] 앞서 설명된 폴리뉴클레오티드의 한 구체예에서, 인트론은 프로모터 요소 및 인자 VIII 폴리펩티드를 인코딩하는 뉴클레오티드 서열의 번역 개시 부위 (가령, 첫 번째 코딩 ATG) 사이에 배치된다.
- [0084] 다른 양상에서, 본 발명은 앞서 설명된 바와 같은 폴리뉴클레오티드를 포함하는 포유류 유전자 요법 벡터를 제공한다.
- [0085] 앞서 설명된 포유류 유전자 요법 벡터의 한 구체예에서, 포유류 유전자 요법 벡터는 아데노 연관된 바이러스 (AAV) 벡터이다.
- [0086] 앞서 설명된 포유류 유전자 요법 벡터의 한 구체예에서, AAV 벡터는 AAV-8 벡터이다.
- [0087] 다른 양상에서, 본 발명은 A형 혈우병을 치료하기 위한 방법을 제공하고, 상기 방법은 앞서 설명된 바와 같은 포유류 유전자 요법 벡터를 치료가 필요한 환자에 투여하는 것을 포함한다.
- [0088] 다른 양상에서, 본 발명은 A형 혈우병을 치료하기 위한, 앞서 설명된 바와 같은 포유류 유전자 요법 벡터를 제공한다.
- [0089] 다른 양상에서, 본 발명은 A형 혈우병을 치료하기 위한 약제의 제조에서, 앞서 설명된 바와 같은 포유류 유전자 요법 벡터의 용도를 제공한다.

도면의 간단한 설명

[0090] **도면의 간단한 설명**

도면 1은 야생형 및 ReFacto-유형 인간 인자 VIII 단백질 구조체의 개략적 도해를 보여준다.

도면 2a 및 2b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CS04 코돈-변경된 뉴클레오티드 서열 (서열 번호: 1) (전장 코딩 서열의 경우 "CS04-FL-NA")을 보여준다.

도면 3은 일부 구체예에 따른 CS04 코돈-변경된 뉴클레오티드 서열에 의해 인코딩된 인자 VIII 변이체 아미노산 서열 (서열 번호: 2) (전장 아미노산 서열의 경우 "CS04-FL-AA")을 보여준다.

도면 4는 일부 구체예에 따른 인자 VIII 변이체의 중쇄를 인코딩하는 CS04 코돈-변경된 뉴클레오티드 서열 (서열 번호: 3) ("CS04-HC-NA")의 부분을 보여준다.

도면 5는 일부 구체예에 따른 인자 VIII 변이체의 경쇄를 인코딩하는 CS04 코돈-변경된 뉴클레오티드 서열 (서열 번호: 4) ("CS04-LC-NA")의 부분을 보여준다.

도면 6은 일부 구체예에 따라서 B-도메인 치환된 링커에 대한 예시적인 코딩 서열 (서열 번호 5-7)을 보여준다. BDL001 (서열 번호: 5), BDL004 (서열 번호: 6) 및 BDL023 (서열 번호: 7)은 B-도메인 치환된 링커를 각각 인코딩하는 CS01, CS04 및 CS23 코돈-변경된 뉴클레오티드 서열의 개별 부분이다.

도면 7a, 7b 및 7c는 일부 구체예에 따라서 CS04 코돈-변경된 뉴클레오티드 서열을 내포하는 AAV 벡터 서열 (서열 번호: 8) ("CS04-AV-NA")을 보여준다.

도면 8a 및 8b는 일부 구체예에 따른 단일 사슬 인자 VIII 변이체를 인코딩하는 CS04△(760-1667) (SPI; CS04△(741-1648), SPE) 코돈-변경된 뉴클레오티드 서열 (서열 번호: 9) ("CS04-SC1-NA")을 보여준다.

도면 9는 일부 구체예에 따라서 CS01△(760-1667) (SPI; CS01△(741-1648), SPE), CS04△(760-1667) (SPI;

CS04△(741-1648), SPE) 및 CS23△(760-1667) (SPI; CS23△(741-1648), SPE) 코돈-변경된 뉴클레오티드 서열에 의해 인코딩된 인자 VIII 변이체 아미노산 서열 (서열 번호: 10) (각각, "CS01-SC1-AA", "CS04-SC1-AA" 및 "CS23-SC1-AA")을 보여준다.

도면 10a 및 10b는 일부 구체예에 따른 단일 사슬 인자 VIII 변이체를 인코딩하는 CS04△(772-1667) (SPI; CS04△(753-1648), SPE) 코돈-변경된 뉴클레오티드 서열 (서열 번호: 11) ("CS04-SC2-NA")을 보여준다.

도면 11은 일부 구체예에 따른 CS01△(772-1667) (SPI; CS01△(753-1648), SPE), CS04△(772-1667) (SPI; CS04△(753-1648), SPE) 및 CS23△(772-1667) (SPI; CS23△(753-1648), SPE) 코돈-변경된 뉴클레오티드 서열에 의해 인코딩된 인자 VIII 변이체 아미노산 서열 (서열 번호: 12) (각각, "CS01-SC2-AA", "CS04-SC2-AA" 및 "CS23-SC2-AA")을 보여준다.

도면 12a 및 12b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CS01 코돈-변경된 뉴클레오티드 서열 (서열 번호: 13) ("CS01-FL-NA")을 보여준다.

도면 13a 및 13b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CS08 코돈-변경된 뉴클레오티드 서열 (서열 번호: 14) ("CS08-FL-NA")을 보여준다.

도면 14a 및 14b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CS10 코돈-변경된 뉴클레오티드 서열 (서열 번호: 15) ("CS10-FL-NA")을 보여준다.

도면 15a 및 15b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CS11 코돈-변경된 뉴클레오티드 서열 (서열 번호: 16) ("CS11-FL-NA")을 보여준다.

도면 16a 및 16b는 일부 구체예에 따른 CS40 야생형 ReFacto 코딩 서열 (서열 번호: 17) ("CS40-FL-NA")을 보여준다.

도면 17a 및 17b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CH25 코돈-변경된 뉴클레오티드 서열 (서열 번호: 18) ("CH25-FL-NA")을 보여준다.

도면 18은 일부 구체예에 따른 야생형 인간 인자 VIII 아미노산 서열 (서열 번호: 19) ("FVIII-FL-AA")을 보여준다.

도면 19는 AscI 및 NotI 제한 부위를 통해, 합성 Refacto-유형 BDD-FVIII DNA 서열을 벡터 중추 pCh-BB01 내로 삽입함으로써, pCS40, pCS01, pCS04, pCS08, pCS10, pCS11 및 pCh25 구조체를 클로닝하기 위한 반응식을 도해한다.

도면 20은 아가로스 겔 전기이동에 의해 분석될 때, AAV 벡터 유전체 제조물의 완전성을 보여준다. 레인 1, DNA 마커; 레인 2, vCS40; 레인 3, vCS01; 레인 4, vCS04. AAV 벡터는 모두 동일한-크기의 유전체를 갖고, 거의 5 kb에서 이동한다 (화살표, 오른쪽). 왼쪽 측면에서 척도는 킬로베이스 (kb)에서 DNA 단편의 크기를 지시한다.

도면 21은 PAGE 및 은 염색에 의한 AAV 벡터 제조물의 단백질 분석을 보여준다. 레인 1, 단백질 마커 (M); 레인 2, vCS40, 레인 3, vCS01; 및 레인 4, vCS04. 이들 구조체 모두 VP1, VP2 및 VP3으로 구성되는 동일한 AAV8 캡시드를 갖는다 (오른쪽 화살표). 왼쪽 측면에 척도는 킬로달톤 (kDa)에서 단백질 마커의 크기를 지시한다.

도면 22a 및 22b는 일부 구체예에 따른 인자 VIII 변이체를 인코딩하는 CS23 코돈-변경된 뉴클레오티드 서열 (서열 번호: 20) ("CS23-FL-NA")을 보여준다.

도면 23은 일부 구체예에 따른 CS23 코돈-변경된 뉴클레오티드 서열에 의해 인코딩된 인자 VIII 변이체 아미노산 서열 (서열 번호: 21) ("CS23-FL-AA")을 보여준다.

도면 24는 일부 구체예에 따른 인자 VIII 변이체의 중쇄를 인코딩하는 CS23 코돈-변경된 뉴클레오티드 서열 (서열 번호: 22) ("CS23-HC-NA")의 부분을 보여준다.

도면 25는 일부 구체예에 따른 인자 VIII 변이체의 경쇄를 인코딩하는 CS23 코돈-변경된 뉴클레오티드 서열 (서열 번호: 23) ("CS23-LC-NA")의 부분을 보여준다.

도면 26은 일부 구체예에 따른 인자 VIII 변이체의 중쇄를 인코딩하는 CS01 코돈-변경된 뉴클레오티드 서열 (서열 번호: 24) ("CS01-HC-NA")의 부분을 보여준다.

도면 27은 일부 구체예에 따른 인자 VIII 변이체의 경쇄를 인코딩하는 CS01 코돈-변경된 뉴클레오티드 서열 (서

열 번호: 25) ("CS01-LC-NA")의 부분을 보여준다.

도면 28a 및 28b는 일부 구체예에 따른 단일 사슬 인자 VIII 변이체를 인코딩하는 CS01△(760-1667) (SPI; CS01△(741-1648), SPE) 코돈-변경된 뉴클레오티드 서열 (서열 번호: 26) ("CS01-SC1-NA")을 보여준다.

도면 29a 및 29b는 일부 구체예에 따른 단일 사슬 인자 VIII 변이체를 인코딩하는 CS01△(772-1667) (SPI; CS01△(753-1648), SPE) 코돈-변경된 뉴클레오티드 서열 (서열 번호: 27) ("CS01-SC2-NA")을 보여준다.

도면 30a 및 30b는 일부 구체예에 따른 단일 사슬 인자 VIII 변이체를 인코딩하는 CS23△(760-1667) (SPI; CS23△(741-1648), SPE) 코돈-변경된 뉴클레오티드 서열 (서열 번호: 28) ("CS23-SC1-NA")을 보여준다.

도면 31a 및 31b는 일부 구체예에 따른 단일 사슬 인자 VIII 변이체를 인코딩하는 CS23△(772-1667) (SPI; CS23△(753-1648), SPE) 코돈-변경된 뉴클레오티드 서열 (서열 번호: 29) ("CS23-SC2-NA")을 보여준다.

발명을 실시하기 위한 구체적인 내용

[0091] 발명의 상세한 설명

[0092] I. 도입

[0093] AAV-기초된 유전자 요법은 혈우병 환자의 치료에 대한 큰 희망을 여전히 주고 있다. B형 혈우병의 경우에, 첫 번째 임상적 데이터는 약 10%의 FIX 수준이 1년 이상 동안 최소한 일부 환자에서 유지될 수 있다는 점에서 고무적이다. 하지만, A형 혈우병의 경우에, AAV 벡터로 5-10%의 치료적 발현 수준을 달성하는 것이 다양한 이유로 인해 과제로 남아있다. 첫 번째, 인자 VIII 코딩 서열은 전통적인 AAV-기초된 벡터에게는 너무 크다. 두 번째, 가공된 B-도메인 결실되거나 또는 절두된 인자 VIII 구조체는 심지어 코돈-최적화될 때에도 생체내에서 불량한 발현을 겪는다. 세 번째, 이들 B-도메인 결실되거나 또는 절두된 인자 VIII 변이체 구조체는 생체내에서 짧은 반감기를 갖고, 불량한 발현의 효과를 악화시킨다. 네 번째, 심지어 발현될 때에도, FVIII는 다른 응고 인자, 예를 들면, 인자 IX처럼, 세포로부터 효율적으로 분비되지 않는다.

[0094] 게다가, 이들 과제는 더욱 높은 용량의 유전자 요법 구조체를 단순히 투여함으로써 다뤄질 수 없다. 현재 지식에 따르면, AAV-기초된 유전자 요법 벡터의 벡터 용량은 2×10^{12} vg/kg 체중을 초과하여 증가되어야 한다. 이것은 이런 고용량에서 T 세포 면역 반응이 촉발되기 때문인데, 상기 반응은 형질도입된 세포를 파괴하고, 그리고 결과로서, 도입유전자 발현이 감소되거나 또는 심지어 제거된다. 이런 이유로, FVIII의 발현을 향상시키는 전략이 FVIII 유전자 요법을 A형 혈우병 환자에 대한 실행가능한 치료 옵션으로 만드는데 필요하다.

[0095] 본 발명은 인자 VIII 유전자 요법과 연관된 이런 저런 문제를 해결하는 코돈-변경된 인자 VIII 변이체 코딩 서열의 발견에 관계한다. 가령, 본원에서 개시된 폴리뉴클레오티드는 포유류 세포에서 현저하게 향상된 발현을 제공하고, 그리고 안정된 패키징 상호작용으로 인해 향상된 비리온 포장을 진시한다. 일부 실행에서, 이들 이점은 코돈 변경된 CS01, CS04 및 CS23 구조체에 높은 서열 동일성 (가령, CS01-HC, CS04-HC 및 CS23-HC 중쇄 코딩 서열 중에서 한 가지에 높은 서열 동일성, 그리고 CS01-LC, CS04-LC 및 CS23-LC 경쇄 코딩 서열 중에서 한 가지에 높은 서열 동일성)을 갖는 인자 VIII의 중쇄와 경쇄에 대한 코딩 서열을 이용함으로써 실현된다.

[0096] 일부 실행에서, 본원에서 설명된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 분자는 야생형 B-도메인을 절두시키거나, 결실시키거나, 또는 대체함으로써 단축되었다. 따라서, 이들 폴리뉴클레오티드는 더욱 큰 폴리펩티드, 예를 들면, 야생형 인자 VIII를 비효율적으로 발현하는 전통적인 유전자 요법 벡터를 통해 인자 VIII를 발현하는데 더욱 적합하다.

[0097] 유리하게는, CS01, CS04 및 CS23 코돈-변경된 인자 VIII 변이체 코딩 서열은 생체내에서 B-도메인 결실된 인자 VIII 구조체의 우수한 발현을 제공하는 것으로 본원에서 밝혀진다. 가령, 실시예 2 및 표 4에서, CS01 (서열 번호: 13), CS04 (서열 번호: 1) 및 CS23 (서열 번호: 20) 코딩 서열을 갖는 AAV-기초된 유전자 요법 벡터의 정맥 내 투여는 인자 VIII 녹아웃 생쥐에서, 야생형 폴리뉴클레오티드 서열 (서열 번호: 17)로 인코딩된 상응하는 CS40 구조체에 비하여, 인자 VIII 발현에서 18-배, 74-배 및 30-배 증가를 제공하는 것으로 증명된다 (표 4).

[0098] 게다가, CS01 및 CS04 코돈-변경된 인자 VIII 변이체 코딩 서열은 우수한 비리온 포장 및 바이러스 생산을 제공하는 것으로 본원에서 또한 밝혀진다. 가령, 실시예 1에서, CS01 및 CS04 구조체를 내포하는 AAV 벡터 구조체는 동일한 양의 세포 펠렛으로부터 단리될 때, 야생형 폴리뉴클레오티드 서열로 인코딩된 상응하는 CS40 구조체에 비하여, 5 내지 7-배 큰 바이러스 수율을 제공하는 것으로 증명된다.

[0099] II. 정의

[0100] 본원에서 이용된 바와 같이, 다음의 용어는 달리 명시되지 않으면, 그들에 생득된 의미를 갖는다.

[0101] 본원에서 이용된 바와 같이, 용어 "인자 VIII" 및 "FVIII"는 교체가능하게 이용되고, 그리고 인자 VIII 활성을 갖는 임의의 단백질 (가령, 활성 FVIII, FVIIIa로서 종종 지칭됨) 또는 인자 VIII 활성, 특히 인자 IXa 보조인자 활성을 갖는 단백질의 단백질 전구체 (가령, 프로-단백질 또는 프리-프로-단백질)를 지칭한다. 한 예시적인 구체예에서, 인자 VIII 폴리펩티드는 야생형 인자 VIII 폴리펩티드의 중쇄와 경쇄에 높은 서열 동일성 (가령, 최소한 70%, 75%, 80%, 85%, 90%, 95%, 99%, 또는 그 이상)을 갖는 서열을 갖는 폴리펩티드를 지칭한다. 일부 구체예에서, 인자 VIII 폴리펩티드의 B-도메인은 인자 VIII 폴리펩티드를 인코딩하는 폴리뉴클레오티드의 크기를 감소시키기 위해 결실되거나, 절두되거나, 또는 링커 폴리펩티드로 대체된다. 한 예시적인 구체예에서, CS04-FL-AA의 아미노산 20-1457은 인자 VIII 폴리펩티드를 구성한다.

[0102] 야생형 인자 VIII 폴리펩티드의 무제한적 실례는 인간 프리-프로-인자 VIII (가령, GenBank 수탁 번호 AAA52485, CAA25619, AAA58466, AAA52484, AAA52420, AAV85964, BAF82636, BAG36452, CAI41660, CAI41666, CAI41672, CAI43241, CA003404, EAW72645, AAH22513, AAH64380, AAH98389, AAI11968, AAI11970, 또는 AAB61261), 상응하는 프로-인자 VIII 및 이들의 자연 변이체; 돼지 프리-프로-인자 VIII (가령, UniProt 수탁 번호 F1RZ36 또는 K7GSZ5), 상응하는 프로-인자 VIII 및 이들의 자연 변이체; 생쥐 프리-프로-인자 VIII (가령, GenBank 수탁 번호 AAA37385, CAM15581, CAM26492, 또는 EDL29229), 상응하는 프로-인자 VIII 및 이들의 자연 변이체; 쥐 프리-프로-인자 VIII (가령, GenBank 수탁 번호 AAQ21580), 상응하는 프로-인자 VIII 및 이들의 자연 변이체; 쥐 프리-프로-인자 VIII; 그리고 다른 포유류 인자 VIII 동족체 (가령, 원숭이, 유인원, 햄스터, 기니 피그 등)를 포함한다.

[0103] 본원에서 이용된 바와 같이, 인자 VIII 폴리펩티드는 인자 IX 보조인자 활성을 갖는 자연 변이체 및 인공 구조체를 포함한다. 본 발명에서 이용된 바와 같이, 인자 VIII는 일부 기저 인자 IX 보조인자 활성 (가령, 상응하는 야생형 활성의 최소한 5%, 10%, 25%, 50%, 75%, 또는 그 이상)을 유지하는 임의의 자연 변이체, 대안적 서열, 동종형, 또는 돌연변이체 단백질을 포괄한다. 인간 개체군에서 발견된 인자 VIII 아미노산 변이 (FVIII-FL-AA (서열 번호: 19)에 비하여)의 실례는 제한 없이, S19R, R22T, Y24C, Y25C, L26P/R, E30V, W33G, Y35C/H, G41C, R48C/K, K67E/N, L69P, E72K, D75E/V/Y, P83R, G89D/V, G92A/V, A97P, E98K, V99D, D101G/H/V, V104D, K108T, M110V, A111T/V, H113R/Y, L117F/R, G121S, E129V, G130R, E132D, Y133C, D135G/Y, T137A/I, S138R, E141K, D145H, V147D, Y155H, V159A, N163K, G164D/V, P165S, C172W, S176P, S179P, V181E/M, K185T, D186G/N/Y, S189L, L191F, G193R, L195P, C198G, S202N/R, F214V, L217H, A219D/T, V220G, D222V, E223K, G224W, T252I, V253F, N254I, G255V, L261P, P262L, G263S, G266F, C267Y, W274C, H275L, G278R, G280D, E284K, V285G, E291G/K, T294I, F295L, V297A, N299I, R301C/H/L, A303E/P, I307S, S308L, F312S, T314A/I, A315V, G323E, L326P, L327P/V, C329F, I331V, M339T, E340K, V345A/L, C348R/S/Y, Y365C, R391C/H/P, S392L/P, A394S, W401G, I405F/S, E409G, W412G/R, K427I, L431F/S, R437P/W, I438F, G439D/S/V, Y442C, K444R, Y450D/N, T454I, F455C, G466E, P470L/R/T, G474E/R/V, E475K, G477V, D478N, T479R, F484C, A488G, R490G, Y492C/H, Y492H, I494T, P496R, G498R, R503H, G513S/V, I522Y, K529E, W532G, P540T, T541S, D544N, R546W, R550C/G/H, S553P, S554C/G, V556D, R560T, D561G/H/Y, I567T, P569R, S577F, V578A, D579A/H, N583S, Q584H/K/R, I585R/T, M586V, D588G/Y, L594Q, S596P, N601D/K, R602G, S603I/R, W604C, Y605H/S, N609I, R612C, N631K/S, M633I, S635N, N637D/I/S, Y639C, L644V, L650F, V653A/M, L659P, A663V, Q664P, F677L, M681I, V682F, Y683C/N, T686R, F698L, M699T/V, M701I, G705V, G710W, N713I, R717L/W, G720D/S, M721I/L, A723T, L725Q, V727F, E739K, Y742C, R795G, P947R, V1012L, E1057K, H1066Y, D1260E, K1289Q, Q1336K, N1460K, L1481P, A1610S, I1698T, Y1699C/F, E1701K, Q1705H, R1708C/H, T1714S, R1715G, A1720V, E1723K, D1727V, Y1728C, R1740G, K1751Q, F1762L, R1768H, G1769R, L1771P, L1775F/V, L1777P, G1779E/R, P1780L, I1782R, D1788H, M1791T, A1798P, S1799H, R1800C/G/H, P1801A, Y1802C, S1803Y, F1804S, L1808F, M1842I, P1844S, T1845P, E1848G, A1853T/V, S1858C, K1864E, D1865N/Y, H1867P/R, G1869D/V, G1872E, P1873R, L1875P, V1876L, C1877R/Y, L1882P, R1888I, E1894G, I1901F, E1904D/K, S1907C/R, W1908L, Y1909C, A1939T/V, N1941D/S, G1942A, M1945V, L1951F, R1960L/Q, L1963P, S1965I, M1966I/V, G1967D, S1968R, N1971T, H1973L, G1979V, H1980P/Y, F1982I, R1985Q, L1994P, Y1998C, G2000A, T2004R, M2007I, G2013R, W2015C, R2016P/W, E2018G, G2022D, G2028R, S2030N, V2035A, Y2036C, N2038S, 2040Y, G2045E/V, I2051S, I2056N, A2058P, W2065R, P2067L, A2070V, S2082N, S2088F, D2093G/Y, H2101D, T2105N, Q2106E/P/R, G2107S, R2109C, I2117F/S, Q2119R, F2120C/L, Y2124C, R2135P, S2138Y, T2141N, M2143V, F2145C, N2148S, N2157D,

P2162L, R2169C/H, P2172L/Q/R, T2173A/I, H2174D, R2178C/H/L, R2182C/H/P, M2183R/V, L2185S/W, S2192I, C2193G, P2196R, G2198V, E2200D, I2204T, I2209N, A2211P, A2220P, P2224L, R2228G/L/P/Q, L2229F, V2242M, W2248C/S, V2251A/E, M2257V, T2264A, Q2265R, F2279C/I, I2281T, D2286G, W2290L, G2304V, D2307A, P2319L/S, R2323C/G/H/L, R2326G/L/P/Q, Q2330P, W2332R, I2336F, R2339T, G2344C/D/S, 그리고 C2345S/Y를 포함한다. 인자 VIII 단백질은 또한, 번역후 변형을 내포하는 폴리펩티드를 포함한다.

[0104] 일반적으로, 인자 VIII를 인코딩하는 폴리뉴클레오티드는 활성 인자 VIII 단백질 (가령, FVIIIa)을 형성하는 번역후 처리를 겪는 비활성 단일 사슬 폴리펩티드 (가령, 프리-프로-단백질)을 인코딩한다. 가령, 도면 1을 참고하면, 야생형 인간 인자 VIII 프리-프로-단백질은 먼저 개열되어 인코딩된 신호 펩티드 (도시되지 않음)를 방출하고, 첫 번째 단일 사슬 프로-단백질 ("인간 야생형 FVIII로서 도시됨)을 형성한다. 프로-단백질은 이후, B 및 A3 도메인 사이에 개열되어, 인자 VIII 중쇄 (가령, A1 및 A2 도메인) 및 B-도메인을 포함하는 첫 번째 폴리펩티드, 그리고 인자 VIII 경쇄 (가령, A3, C1 및 C3 도메인 포함)를 포함하는 두 번째 폴리펩티드를 형성한다. 첫 번째 폴리펩티드는 B-도메인을 제거하고, 그리고 또한, 성숙 인자 VIIIa 단백질에서 인자 VIII 경쇄와 연관된 상태로 남아있는 A1 및 A2 도메인을 분리하기 위해 더욱 개열된다. 인자 VIII 성숙 과정의 리뷰를 위해, Graw et al., Nat Rev Genet., 6(6):488-501 (2005)을 참조하고, 이것의 내용은 모든 점에서 본원에 전체적으로 참조로서 편입된다.

[0105] 하지만, 일부 구체예에서, 인자 VIII 폴리펩티드는 단일 사슬 인자 VIII 폴리펩티드이다. 단일 사슬 인자 VIII 폴리펩티드는 자연 개열 부위를 제거하고, 그리고 임의선택적으로, 인자 VIII의 B-도메인을 제거하거나, 절두하거나, 또는 대체하도록 가공된다. 따라서, 이들은 개열 (임의선택적 신호 및/또는 리더 펩티드의 개열 제외)에 의해 성숙되지 않고, 그리고 단일 사슬로서 활성이다. 단일 사슬 인자 VIII 폴리펩티드의 무제한적 실례는 Zollner et al. (Thromb Res, 134(1):125-31 (2014)) 및 Donath et al. (Biochem J., 312(1):49-55 (1995))에서 설명되고, 이들의 개시는 모든 점에서 본원에 전체적으로 참조로서 편입된다.

[0106] 본원에서 이용된 바와 같이, 용어 "인자 VIII 중쇄" 또는 단순히 "중쇄"는 인자 VIII 폴리펩티드의 A1 및 A2 도메인의 응집체를 지칭한다. 한 예시적인 구체예에서, CS04-FL-AA (서열 번호: 2)의 아미노산 20-759는 인자 VIII 중쇄를 구성한다.

[0107] 본원에서 이용된 바와 같이, 용어 "인자 VIII 경쇄" 또는 단순히 "경쇄"는 인자 VIII 폴리펩티드의 A3, C1 및 C2 도메인의 응집체를 지칭한다. 한 예시적인 구체예에서, 아미노산 774-1457 CS04-FL-AA (서열 번호: 2)는 인자 VIII 경쇄를 구성한다. 일부 구체예에서, 인자 VIII 경쇄는 생체내에서 성숙 동안 방출되는 산성 a3 펩티드를 배제한다.

[0108] 일반적으로, 인자 VIII 중쇄와 경쇄는 예로서, 임의선택적 B-도메인 또는 B-도메인 치환된 링커와 함께, 단일 폴리펩티드 사슬로서 발현된다. 하지만, 일부 구체예에서, 인자 VIII 중쇄 및 인자 VIII 경쇄는 별개의 폴리펩티드 사슬로서 발현되고 (가령, 공동발현되고), 그리고 재구성되어 인자 VIII 단백질을 형성한다 (가령, 생체내에서 또는 시험관내에서).

[0109] 본원에서 이용된 바와 같이, 용어 "B-도메인 치환된 링커" 및 "인자 VIII 링커"는 교체가능하게 이용되고, 그리고 야생형 인자 VIII B-도메인의 절두된 이형 (가령, FVIII-FL-AA (서열 번호: 19)의 아미노산 760-1667) 또는 인자 VIII 폴리펩티드의 B-도메인을 대체하도록 가공된 펩티드를 지칭한다. 본원에서 이용된 바와 같이, 인자 VIII 링커는 일부 구체예에 따라서 인자 VIII 변이체 폴리펩티드에서 인자 VIII 중쇄의 C 말단 및 인자 VIII 경쇄의 N 말단 사이에 배치된다. B-도메인 치환된 링커의 무제한적 실례는 U.S. 특허 번호 4,868,112, 5,112,950, 5,171,844, 5,543,502, 5,595,886, 5,610,278, 5,789,203, 5,972,885, 6,048,720, 6,060,447, 6,114,148, 6,228,620, 6,316,226, 6,346,513, 6,458,563, 6,924,365, 7,041,635 및 7,943,374; U.S. 특허 출원 공개 번호 2013/024960, 2015/0071883 및 2015/0158930; 그리고 PCT 공개 번호 WO 2014/064277 및 WO 2014/127215에서 개시되고, 이들의 개시는 모든 점에서 본원에 전체적으로 참조로서 편입된다.

[0110] 본원에서 달리 특정되지 않으면, 인자 VIII 아미노산의 넘버링은 도면 18에서 서열 번호: 19로서 제공된, 전장, 야생형 인간 인자 VIII 서열에서 상응하는 아미노산 (FVIII-FL-AA)을 지칭한다. 따라서, 본원에서 개시된 인자 VIII 변이체 단백질에서 아미노산 치환을 지칭할 때, 언급된 아미노산 번호는 전장, 야생형 인자 VIII 서열에서 유사한 (가령, 구조적으로 또는 기능적으로 동등한) 및/또는 상동한 (가령, 일차 아미노산 서열에서 진화적으로 보존된) 아미노산을 지칭한다. 가령, T2105N 아미노산 치환은 전장, 야생형 인간 인자 VIII 서열 (FVIII-FL-AA; 서열 번호: 19)의 위치 2105에서 T에서 N으로 치환, 그리고 CS04에 의해 인코딩된 인자 VIII 변이체 단백질 (CS04-FL-AA; 서열 번호: 2)의 위치 1211에서 T에서 N으로 치환을 지칭한다.

- [0111] 본원에서 설명된 바와 같이, 인자 VIII 아미노산 넘버링 시스템은 인자 VIII 신호 펩티드 (가령, 전장, 야생형 인간 인자 VIII 서열의 아미노산 1-19)가 포함되는 지에 의존한다. 신호 펩티드가 포함되는 경우에, 넘버링은 "신호 펩티드 포함" 또는 "SPI"로서 지칭된다. 신호 펩티드가 포함되지 않는 경우에, 넘버링은 "신호 펩티드 배제" 또는 "SPE"로서 지칭된다. 가령, F328S는 SPE 넘버링에서 F309S와 동일한 아미노산에 대한 SPI 넘버링이다. 달리 지시되지 않으면, 모든 아미노산 넘버링은 도면 18에서 서열 번호: 19로서 제공된, 전장, 야생형 인간 인자 VIII 서열 (FVIII-FL-AA)에서 상응하는 아미노산을 지칭한다.
- [0112] 본원에서 설명된 바와 같이, 코돈-변경된 폴리뉴클레오티드는 선천적으로-코딩된 인자 VIII 구조체 (가령, 야생형 인간 코돈을 이용하여 동일한 인자 VIII 구조체를 인코딩하는 폴리뉴클레오티드)에 의해 제공된 인자 VIII 발현의 수준과 비교하여, 생체내에서 (가령, 유전자 요법 벡터의 일부로서 투여될 때) 유전자도입 인자 VIII의 증가된 발현을 제공한다. 본원에서 이용된 바와 같이, 용어 "증가된 발현"은 선천적으로-코딩된 인자 VIII 구조체가 투여된 동물의 혈액 내에 유전자도입 인자 VIII 활성의 수준과 비교하여, 코돈-변경된 폴리뉴클레오티드 인코딩 인자 VIII이 투여된 동물의 혈액 내에 유전자도입 인자 VIII 활성의 증가된 수준을 지칭한다. 활성 수준은 당해 분야에서 공지된 임의의 인자 VIII 활성을 이용하여 측정될 수 있다. 인자 VIII 활성을 결정하기 위한 예시적인 검정은 Technochrome FVIII 검정 (Technoclone, Vienna, Austria)이다.
- [0113] 일부 구체예에서, 증가된 발현은 선천적으로 코딩된 인자 VIII 폴리뉴클레오티드가 투여된 동물의 혈액 내에 유전자도입 인자 VIII 활성의 수준과 비교하여, 코돈-변경된 인자 VIII 폴리뉴클레오티드가 투여된 동물의 혈액 내에 최소한 25% 큰 유전자도입 인자 VIII 활성을 지칭한다. 일부 구체예에서, 증가된 발현은 선천적으로 코딩된 인자 VIII 폴리뉴클레오티드가 투여된 동물의 혈액 내에 유전자도입 인자 VIII 활성의 수준과 비교하여, 코돈-변경된 인자 VIII 폴리뉴클레오티드가 투여된 동물의 혈액 내에 최소한 50% 큰, 최소한 75% 큰, 최소한 100% 큰, 최소한 3-배 큰, 최소한 4-배 큰, 최소한 5-배 큰, 최소한 6-배 큰, 최소한 7-배 큰, 최소한 8-배 큰, 최소한 9-배 큰, 최소한 10-배 큰, 최소한 15-배 큰, 최소한 20-배 큰, 최소한 25-배 큰, 최소한 30-배 큰, 최소한 40-배 큰, 최소한 50-배 큰, 최소한 60-배 큰, 최소한 70-배 큰, 최소한 80-배 큰, 최소한 90-배 큰, 최소한 100-배 큰, 최소한 125-배 큰, 최소한 150-배 큰, 최소한 175-배 큰, 최소한 200-배 큰, 최소한 225-배 큰, 또는 최소한 250-배 큰 유전자도입 인자 VIII 활성을 지칭한다.
- [0114] 본원에서 설명된 바와 같이, 코돈-변경된 폴리뉴클레오티드는 선천적으로-코딩된 인자 VIII 구조체 (가령, 야생형 인간 코돈을 이용하여 동일한 인자 VIII 구조체를 인코딩하는 폴리뉴클레오티드)에 의해 제공된 벡터 생산의 수준과 비교하여 증가된 벡터 생산을 제공한다. 본원에서 이용된 바와 같이, 용어 "증가된 바이러스 생산"은 선천적으로-코딩된 인자 VIII 구조체로 접종된 세포 배양액에서 벡터 수율과 비교하여, 코돈-변경된 폴리뉴클레오티드 인코딩 인자 VIII로 접종된 세포 배양액에서 증가된 벡터 수율 (가령, 배양액 리터당 역가)을 지칭한다. 벡터 수율은 당해 분야에서 공지된 임의의 벡터 역가 검정을 이용하여 측정될 수 있다. 벡터 수율 (가령, AAV 벡터의)을 결정하기 위한 예시적인 검정은 AAV2 반전된 말단 반복을 표적으로 하는 qPCR이다 (Aurnhammer, Human Gene Therapy Methods: Part B 23:18-28 (2012)).
- [0115] 일부 구체예에서, 증가된 바이러스 생산은 동일한 유형의 배양액에서 선천적으로-코딩된 인자 VIII 구조체의 수율과 비교하여, 최소한 25% 큰 코돈-변경된 벡터 수율을 지칭한다. 일부 구체예에서, 증가된 벡터 생산은 동일한 유형의 배양액에서 선천적으로-코딩된 인자 VIII 구조체의 수율과 비교하여, 최소한 50% 큰, 최소한 75% 큰, 최소한 100% 큰, 최소한 3-배 큰, 최소한 4-배 큰, 최소한 5-배 큰, 최소한 6-배 큰, 최소한 7-배 큰, 최소한 8-배 큰, 최소한 9-배 큰, 최소한 10-배 큰, 최소한 15-배 큰, 또는 최소한 20-배 큰 코돈-변경된 벡터 수율을 지칭한다.
- [0116] 본원에서 이용된 바와 같이, 용어 "혈우병"은 감소된 혈액 응고 또는 응고에 의해 광범위하게 특징되는 일군의 질환 상태를 지칭한다. 혈우병은 유형 A, 유형 B, 또는 유형 C 혈우병, 또는 3개 질환 유형 모두의 복합을 지칭할 수 있다. 유형 A 혈우병 (A형 혈우병)은 인자 VIII (FVIII) 활성의 감소 또는 상실에 의해 유발되고, 그리고 혈우병 아형 중에서 가장 두드러진다. 유형 B 혈우병 (B형 혈우병)은 인자 IX (FIX) 응고 기능의 상실 또는 감소로부터 발생한다. 유형 C 혈우병 (C형 혈우병)은 인자 XI (FXI) 응고 활성에서 상실 또는 감소의 결과이다. A형과 B형 혈우병은 X 연관된 질환이고, 반면 C형 혈우병은 상염색체 연관된 질환이다. 혈우병에 대한 전통적인 치료는 응고 인자, 예를 들면, FVIII, Bebulin®-VH를 비롯한 FIX 및 FXI뿐만 아니라 FEIBA-VH, 데스모프레신, 그리고 혈장 주입의 예방적 및 주문형 투여 둘 모두를 포함한다.
- [0117] 본원에서 이용된 바와 같이, 용어 "FVIII 유전자 요법"은 혈우병과 연관된 하나 또는 그 이상의 증상 (가령, 임상적 인자)을 완화하거나, 축소하거나, 또는 이의 재발을 예방하기 위해, 인자 VIII를 인코딩하는 핵산을 환자

에게 제공하는 임의의 치료적 접근법을 포함한다. 상기 용어는 혈우병을 앓는 개체의 건강을 유지하거나 또는 향상시키기 위해, 인자 VIII의 임의의 변형된 형태 (가령, 인자 VIII 변이체)를 비롯하여, 인자 VIII 분자를 인코딩하는 핵산을 포함하는 임의의 화합물, 약물, 시술, 또는 섭생을 투여하는 것을 포괄한다. 당업자는 FVIII 요법의 코스 또는 FVIII 치료적 작용제의 용량이 예로서, 본 발명에 따라 획득된 결과에 근거하여 변화될 수 있다는 것을 인지할 것이다.

[0118] 본원에서 이용된 바와 같이, 용어 "우회로 요법"은 혈우병과 연관된 하나 또는 그 이상의 증상 (가령, 임상적 인자)을 완화하거나, 축소하거나, 또는 이의 재발을 예방하기 위해, 비-인자 VIII 지혈 작용제, 화합물 또는 응고 인자를 환자에게 제공하는 임의의 치료적 접근법을 포함한다. 비-인자 VIII 화합물 및 응고 인자는 인자 VIII 저해제 우회로 활성 (FEIBA), 재조합 활성화된 인자 VII (FVIIa), 프로트롬빈 복합체 농축물, 그리고 활성화된 프로트롬빈 복합체 농축물을 포함하지만 이들에 한정되지 않는다. 이들 비-인자 VIII 화합물 및 응고 인자는 재조합이거나 또는 혈장-유래될 수 있다. 당업자는 우회로 요법의 코스 또는 우회로 요법의 용량이 예로서, 본 발명에 따라 획득된 결과에 근거하여 변화될 수 있다는 것을 인지할 것이다.

[0119] 본원에서 이용된 바와 같이, 인자 VIII 분자를 인코딩하는 핵산 및 전통적인 A형 혈우병 치료적 작용제를 투여하는 것을 포함하는 "복합 요법"은 혈우병과 연관된 하나 또는 그 이상의 증상 (가령, 임상적 인자)을 완화하거나, 축소하거나, 또는 이의 재발을 예방하기 위해, 인자 VIII 분자를 인코딩하는 핵산 및 인자 VIII 분자 및/또는 비-인자 VIII 지혈 작용제 (가령, 우회로 치료적 작용제) 둘 모두를 환자에게 제공하는 임의의 치료적 접근법을 포함한다. 상기 용어는 인자 VIII의 임의의 변형된 형태를 비롯하여, 인자 VIII 분자를 인코딩하는 핵산을 포함하는 임의의 화합물, 약물, 시술, 또는 섭생을 투여하는 것을 포괄하는데, 이것은 혈우병을 앓는 개체의 건강을 유지하거나 또는 향상시키는데 유용하고 본원에서 설명된 치료적 작용제 중에서 한 가지를 포함한다.

[0120] 용어 "치료 효과량 또는 용량" 또는 "치료적으로 충분한 양 또는 용량" 또는 "효과적인 또는 충분한 양 또는 용량"은 이것이 투여되면 치료 효과를 산출하는 용량을 지칭한다. 가령, 혈우병을 치료하는데 유용한 약물의 치료 효과량은 혈우병과 연관된 하나 또는 그 이상의 증상을 예방하거나 또는 완화할 수 있는 양일 수 있다. 정확한 용량은 치료의 목적에 의존할 것이고, 그리고 공지된 기술을 이용하여 당업자에 의해 확인될 수 있을 것이다 (가령, Lieberman, *Pharmaceutical Dosage Forms* (vols. 13, 1992); Lloyd, *The 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 참조한다).

[0121] 본원에서 이용된 바와 같이, 용어 "유전자"는 폴리펩티드 사슬을 코딩하는 DNA 분자의 분절 (가령, 코딩 영역)을 지칭한다. 일부 구체예에서, 유전자는 폴리펩티드 사슬을 생산하는데 관련되는, 코딩 영역 직전에, 코딩 영역 이후에 및/또는 코딩 영역을 개재하는 영역 (가령, 조절 요소, 예를 들면, 프로모터, 인핸서, 폴리아데닐화 서열, 5'-비번역 영역, 3'-비번역 영역, 또는 인트론)에 의해 배치된다.

[0122] 본원에서 이용된 바와 같이, 용어 "조절 요소"는 세포에서 코딩 서열의 발현을 제공하는 뉴클레오티드 서열, 예를 들면, 프로모터, 인핸서, 종결인자, 폴리아데닐화 서열, 인트론 등을 지칭한다.

[0123] 본원에서 이용된 바와 같이, 용어 "프로모터 요소"는 코딩 서열의 발현을 제어하는데 보조하는 뉴클레오티드 서열을 지칭한다. 일반적으로, 프로모터 요소는 유전자의 번역 시작 부위의 5'에 위치된다. 하지만, 일정한 구체예에서, 프로모터 요소는 인트론 서열, 또는 코딩 서열의 3' 내에 위치될 수 있다. 일부 구체예에서, 유전자 요법 벡터에 유용한 프로모터는 표적 단백질의 선천적 유전자로부터 유래된다 (가령, 인자 VIII 프로모터). 일부 구체예에서, 유전자 요법 벡터에 유용한 프로모터는 표적 생물의 특정 세포 또는 조직에서 발현에 특이적이다 (가령, 간-특이적 프로모터). 또 다른 구체예에서, 복수의 충분히 특징된 프로모터 요소 중에서 한 가지가 본원에서 설명된 유전자 요법 벡터에서 이용된다. 충분히 특징된 프로모터 요소의 무제한적 실례는 CMV 초기 프로모터, β -액틴 프로모터 및 메틸 CpG 결합 단백질 2 (MeCP2) 프로모터를 포함한다. 일부 구체예에서, 프로모터는 구조성 프로모터인데, 이것은 표적 단백질의 실제적으로 일정한 발현을 주동한다. 다른 구체예에서, 프로모터는 유도성 프로모터인데, 이것은 특정 자극 (가령, 특정 치료 또는 작용제에 노출)에 대한 응답으로 표적 단백질의 발현을 주동한다. AAV-매개된 유전자 요법을 위한 프로모터 설계의 리뷰를 위해, Gray et al. (*Human Gene Therapy* 22:1143-53 (2011))을 참조하고, 이의 내용은 모든 점에서 전체적으로 명시적으로 참조로서 편입된다.

[0124] 본원에서 이용된 바와 같이, 용어 "벡터"는 핵산 (가령, 인자 VIII 유전자 요법 구조체를 인코딩하는)을 숙주 세포 내로 전달하는데 이용되는 임의의 운반체를 지칭한다. 일부 구체예에서, 벡터는 레플리콘을 포함하는데, 이것은 표적 핵산과 함께, 운반체를 복제하는 기능을 한다. 유전자 요법에 유용한 벡터의 무제한적 실례는 플라

스미드, 파지, 코스미드, 인공 염색체 및 바이러스를 포함하는데, 이들은 생체내에서 복제의 자율적인 단위로서 기능한다. 일부 구체예에서, 벡터는 표적 핵산 (가령, 인자 VIII 변이체를 인코딩하는 코돈-변경된 폴리뉴클레오티드)을 도입하기 위한 바이러스 운반체이다. 유전자 요법에 유용한 많은 변형된 진핵 바이러스는 당해 분야에서 공지된다. 가령, 아데노 연관된 바이러스 (AAVs)가 인간 유전자 요법에서 이용하기 특히 충분히 적합한데, 그 이유는 인간이 상기 바이러스에 대한 자연 숙주이고, 선천적 바이러스가 임의의 질환에 기여하는 것으로 알려져 있지 않고, 그리고 상기 바이러스가 경미한 면역 반응을 유발하기 때문이다.

[0125] 본원에서 이용된 바와 같이, 용어 "CpG 섬"은 CpG 디뉴클레오티드의 통계학적으로 상승된 밀도를 갖는, 폴리뉴클레오티드 내에 영역을 지칭한다. 본원에서 이용된 바와 같이, 폴리뉴클레오티드 (가령, 코돈-변경된 인자 VIII 단백질을 인코딩하는 폴리뉴클레오티드)의 영역은 200-염기쌍 윈도우 위에서: (i) 상기 영역이 50%보다 큰 GC 함량을 갖고, 그리고 (ii) 예상된 CpG 디뉴클레오티드마다 관찰된 CpG 디뉴클레오티드의 비율이 아래의 관계에 의해 규정된 바와 같이, 최소한 0.6이면, CpG 섬이다:

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

[0127] CpG 섬을 확인하기 위한 방법에 관한 추가 정보를 위해, Gardiner-Garden M. et al., J Mol Biol., 196(2):261-82 (1987)를 참조하고, 이것의 내용은 모든 점에서 명시적으로 본원에 전체적으로 참조로서 편입된다.

[0128] 본원에서 이용된 바와 같이, 용어 "핵산"은 데옥시리보뉴클레오티드 또는 리보뉴클레오티드 및 단일- 또는 이중가닥 형태에서 이들의 중합체, 그리고 이들의 보체를 지칭한다. 상기 용어는 공지된 뉴클레오티드 유사체 또는 변형된 중추 잔기 또는 연쇄를 내포하는 핵산을 포괄하는데, 이들은 합성, 자연발생 및 비자연발생이고, 참고 핵산과 유사한 결합 성질을 갖고, 그리고 참고 뉴클레오티드와 유사한 방식으로 물질대사된다. 이런 유사체의 실례는 제한 없이, 포스포로티오에이트, 포스포라미데이트, 메틸 포스포네이트, 키랄-메틸 포스포네이트, 2-O-메틸 리보뉴클레오티드, 그리고 펩티드-핵산 (PNAs)을 포함한다.

[0129] 용어 "아미노산"은 자연발생 및 비자연 아미노산뿐만 아니라 자연발생 아미노산과 유사한 방식으로 기능하는 아미노산 유사체 및 아미노산 모방체를 지칭한다. 자연발생 아미노산은 유전자 코드에 의해 인코딩된 아미노산뿐만 아니라 추후 변형되는 아미노산, 예를 들면, 히드록시프롤린, γ-카르복시글루타메이트 및 O-포스포세린이다. 자연발생 아미노산은 예로서, D- 및 L-아미노산을 포함할 수 있다. 본원에서 이용된 아미노산은 또한, 비자연 아미노산을 포함할 수 있다. 아미노산 유사체는 자연 발생 아미노산과 동일한 기본 화학 구조, 다시 말하면, 수소에 결합된 임의의 탄소, 카르복실 기, 아미노 기, 그리고 R 기를 갖는 화합물, 예를 들면, 호모세린, 노르류신, 메티오닌 술폭시드, 또는 메티오닌 메틸 술포늄을 지칭한다. 이런 유사체는 변형된 R 기 (가령, 노르류신) 또는 변형된 펩티드 중추를 갖지만, 자연 발생 아미노산과 동일한 기본 화학 구조를 유지한다. 아미노산 모방체는 아미노산의 일반적인 화학 구조와 상이한 구조를 갖지만, 자연발생 아미노산과 유사한 방식으로 기능하는 화학적 화합물을 지칭한다. 아미노산은 본원에서, 그들의 통상적으로 알려진 3 문자 기호에 의해 또는 IUPAC-IUB Biochemical Nomenclature Commission에 의해 권장되는 1-문자 기호에 의해 지칭될 수 있다. 뉴클레오티드는 유사하게, 그들의 통상적으로 인정되는 1-문자 코드에 의해 지칭될 수 있다.

[0130] 아미노산 서열에 관하여, 당업자는 인코딩된 서열 내에 단일 아미노산 또는 적은 비율의 아미노산을 변경, 부가 또는 결실하는 핵산 또는 펩티드 서열에 개별 치환, 결실 또는 부가가 이런 변경이 화학적으로 유사한 아미노산으로 아미노산의 치환을 유발하는 경우에, "보존성으로 변형된 변이체"라는 것을 인지할 것이다. 기능적으로 유사한 아미노산을 제공하는 보존성 치환 표는 당업자에게 알려져 있다. 이런 보존성으로 변형된 변이체는 본 발명의 다형성 변이체, 중간 동족체, 그리고 대립유전자에 더해지고 이들을 배제하지 않는다.

[0131] 기능적으로 유사한 아미노산을 제공하는 보존성 아미노산 치환은 당업자에게 알려져 있다. 특정 아미노산, 예를 들면, 촉매성, 구조적, 또는 입체적으로 중요한 아미노산의 기능성에 따라, 아미노산의 상이한 군화가 서로에 대해 보존성 치환인 것으로 고려될 수 있다. 표 1은 아미노산의 전하 및 극성, 아미노산의 소수성, 아미노산의 표면 노출/구조적 본성, 그리고 아미노산의 이차 구조 성향에 근거하여 보존성 치환인 것으로 고려되는 아미노산의 군화를 제공한다.

[0132] 표 1. 단백질 내에 잔기의 기능성에 근거된 보존성 아미노산 치환의 군화.

표 1

[0133]

중요한 특징	보존성 군화
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 성향	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

[0134]

2개 또는 그 이상의 핵산 또는 펩티드 서열의 맥락에서 용어 "동일한," 또는 "동일성" 퍼센트는 예로서, 아래에 설명된 디폴트 파라미터를 갖는 BLAST 또는 BLAST 2.0 서열 비교 알고리즘을 이용하여, 또는 수동 정렬과 시각적 검사에 의해 계측될 때 동일하거나 또는 동일한 아미노산 잔기 또는 뉴클레오타이드의 특정된 백분율 (즉, 비교 윈도우 또는 지정된 영역에 걸쳐 최고 상응을 위해 비교되고 정렬될 때, 특정된 영역에 걸쳐 약 60% 동일성, 바람직하게는 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 더욱 높은 동일성)을 갖는 2개 또는 그 이상의 서열 또는 하위서열을 지칭한다.

[0135]

당해 분야에서 공지된 바와 같이, 다수의 상이한 프로그램이 단백질 (또는 아래에 논의된 바와 같이 핵산)이 공지된 서열에 서열 동일성 또는 유사성을 갖는 지를 확인하는데 이용될 수 있다. 서열 동일성 및/또는 유사성은 Smith & Waterman, Adv. Appl. Math., 2:482 (1981)의 국부 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 방법, 이들 알고리즘의 컴퓨터화 실행 (Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, WI에서 GAP, BESTFIT, FASTA 및 TFASTA), Devereux et al., Nucl. Acid Res., 12:387-395 (1984)에 의해 설명된 최고 적합 서열 프로그램 (바람직하게는, 디폴트 세팅을 이용), 또는 검사를 포함하지만 이들에 한정되지 않는, 당해 분야에서 공지된 표준 기술을 이용하여 결정된다. 바람직하게는, 동일성 퍼센트는 다음의 파라미터에 근거하여 FastDB에 의해 계산된다: 1의 부정합 페널티; 1의 갭 페널티; 0.33의 갭 크기 페널티; 및 30의 연결 페널티, "Current Methods in Sequence Comparison and Analysis," Macromolecule Sequencing and Synthesis, Selected Methods and Applications, pp 127-149 (1988), Alan R. Liss, Inc, 이들 모두 참조로서 편입됨.

[0136]

유용한 알고리즘의 실례는 PILEUP이다. PILEUP는 진행성, 쌍별 정렬을 이용하여 일군의 관련된 서열로부터 복수 서열 정렬을 창출한다. 이것은 또한, 정렬을 창출하는데 이용된 군집화 관계를 보여주는 나무를 플롯팅할 수 있다. PILEUP는 Feng & Doolittle, J. Mol. Evol. 35:351-360 (1987)의 진행성 정렬 방법의 단순화를 이용하고; 상기 방법은 Higgins & Sharp CABIOS 5:151-153 (1989)에 의해 설명된 방법과 유사한데, 이들 둘 모두 참조로서 편입된다. 유용한 PILEUP 파라미터는 3.00의 디폴트 갭 가중, 0.10의 디폴트 갭 길이 가중, 그리고 가중된 단부 갭을 포함한다.

[0137]

유용한 알고리즘의 다른 실례는 참조로서 편입되는 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)에서 설명된 BLAST 알고리즘이다. 특히 유용한 BLAST 프로그램은 Altschul et al., Methods in Enzymology, 266:460-480 (1996); [http://blast.wustl.edu/blast/ README.html](http://blast.wustl.edu/blast/README.html)]으로부터 획득된 WU-BLAST-2 프로그램이다. WU-BLAST-2는 여러 검색 파라미터를 이용하는데, 이들 중에서 대부분은 디폴트 값에 세팅된다. 조정가능한 파라미터는 다음의 값으로 세팅된다: 중복 스캔 =1, 중복 분율 = 0.125, 단어 역치 (T) = 11. HSP S 및 HSP S2 파라미터는 동적 값이고, 그리고 특정 서열의 구성 및 관심되는 서열이 검색되는 특정 데이터베이스의 구성에 따라 프로그램 그 자체에 의해 확립된다; 하지만, 이들 값은 감수성을 증가시키기 위해 조정될 수 있다.

[0138] 추가 유용한 알고리즘은 참조로서 편입되는 Altschul et al., Nucl. Acids Res., 25:3389-3402에 의해 보고된 바와 같은 깡트 BLAST이다. 깡트 BLAST는 BLOSUM-62 치환 점수; 9로 세팅된 역치 T 파라미터; 깡이 없는 연장을 촉발하기 위한 2-히트 방법; 10+k의 비용으로 k의 충전 깡 길이; 16에 세팅된 Xu, 그리고 데이터베이스 검색 시기의 경우 40 및 알고리즘의 출력 시기의 경우 67에 세팅된 Xg를 이용한다. 깡트 정렬은 ~22 비트에 상응하는 점수에 의해 촉발된다.

[0139] % 아미노산 서열 동일성 값은 정렬된 영역 내에 "더욱 긴" 서열의 잔기의 총수에 의해 나뉘어진 정합 동일한 잔기의 숫자에 의해 결정된다. "더욱 긴" 서열은 정렬된 영역 내에 가장 많은 실제 잔기를 갖는 것이다 (정렬 점수를 최대화하기 위해 WU-Blast-2에 의해 도입된 깡은 무시된다). 유사한 방식으로, 확인된 폴리펩티드의 코딩 서열에 대하여 "핵산 서열 동일성 퍼센트 (%)"는 세포 주기 단백질의 코딩 서열 내에 뉴클레오티드 잔기와 동일한, 후보 서열 내에 뉴클레오티드 잔기의 백분율로서 규정된다. 바람직한 방법은 각각, 1 및 0.125에 세팅된 중복 스캔 및 중복 분율을 갖는 디폴트 파라미터에 세팅된 WU-BLAST-2의 BLASTN 모듈을 활용한다.

[0140] 정렬은 정렬되는 서열 내에 깡의 도입을 포함할 수 있다. 이에 더하여, 도면 2의 서열 (서열 번호:1)에 의해 인코딩된 단백질보다 더욱 많은 또는 더욱 적은 아미노산을 내포하는 서열의 경우에, 한 구체예에서, 서열 동일성의 백분율은 아미노산 또는 뉴클레오티드의 총수에 관계하여 동일한 아미노산 또는 뉴클레오티드의 숫자에 근거하여 결정될 것으로 이해된다. 따라서, 한 구체예에서, 예로서 도면 2에서 도시된 서열 (서열 번호:1)보다 짧은 서열의 서열 동일성은 아래에 논의된 바와 같이, 더욱 짧은 서열에서 뉴클레오티드의 숫자를 이용하여 결정될 것이다. 동일성 퍼센트 계산에서 상대적 가중은 서열 변이의 다양한 현성, 예를 들면, 삽입, 결실, 치환 등에 배정되지 않는다.

[0141] 한 구체예에서, 단지 동일성만 양으로 (+1) 채점되고, 그리고 깡을 포함하는 서열 변이의 모든 형태는 "0"의 값이 배정되는데, 이것은 서열 유사성 계산을 위해 아래에 설명된 바와 같은 가중된 척도 또는 파라미터에 대한 필요를 배제시킨다. 서열 동일성 퍼센트는 예로서, 정합 동일한 잔기의 숫자를 정렬된 영역 내에 "더욱 짧은" 서열의 잔기의 총수로 나누고 100을 곱함으로써 계산될 수 있다. "더욱 긴" 서열은 정렬된 영역 내에 가장 많은 실제 잔기를 갖는 것이다.

[0142] 용어 "대립형질 변이체"는 특정 유전자 좌위에서 유전자의 결정다형뿐만 아니라 이들 유전자의 mRNA 전사체로부터 유래된 cDNAs, 그리고 이들에 의해 인코딩된 폴리펩티드를 지칭한다. 용어 "바람직한 포유류 코돈"은 다음의 목록에서 선택되는, 포유류 세포에서 발현되는 단백질에서 가장 빈번하게 이용되는, 아미노산을 인코딩하는 코돈의 세트 중에서 코돈의 부분집합을 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).

[0143] 본원에서 이용된 바와 같이, 용어 "코돈-변경된"은 폴리펩티드 (가령, 인자 VIII 변이체 단백질)를 인코딩하는 선천적 폴리뉴클레오티드의 최소한 하나의 코돈이 폴리뉴클레오티드 서열의 성질을 향상시키기 위해 변화된, 상기 폴리펩티드를 인코딩하는 폴리뉴클레오티드 서열을 지칭한다. 일부 구체예에서, 향상된 성질은 폴리펩티드를 코딩하는 mRNA의 증가된 전사, mRNA의 증가된 안정성 (가령, 향상된 mRNA 반감기), 폴리펩티드의 증가된 번역 및/또는 벡터 내에 폴리뉴클레오티드의 증가된 포장을 증진한다. 향상된 성질을 달성하는데 이용될 수 있는 변경의 무제한적 실례는 특정 아미노산에 대한 코돈의 사용빈도 및/또는 분포를 변화시키고, 전역 및/또는 국부 GC 함량을 조정하고, AT-풍부한 서열을 제거하고, 반복된 서열 요소를 제거하고, 전역 및/또는 국부 CpG 디뉴클레오티드 함량을 조정하고, 잠적 조절 요소 (가령, TATA 상자 및 CCAAT 상자 요소)를 제거하고, 인트론/엑손 스플라이스 부위를 제거하고, 조절 서열을 향상시키고 (가령, 코자크 공통 서열의 도입), 그리고 전사된 mRNA 내에 이차 구조 (가령, 스템-루프)를 형성할 수 있는 서열 요소를 제거하는 것을 포함한다.

[0144] 본원에서 논의된 바와 같이, 본원에서 발명의 성분을 지칭하는 다양한 명명법이 있다. "CS-숫자" (가령, "CS04", "CS01", "CS23" 등)는 변이체를 비롯하여, FVIII 폴리펩티드를 인코딩하는 코돈 변경된 폴리뉴클레오티

드 및/또는 인코딩된 폴리펩티드를 지칭한다. 가령, CS01-FL은 CS01 폴리뉴클레오티드 서열에 의해 인코딩된 **전장 코돈** 변경된 CS01 폴리뉴클레오티드 서열 또는 아미노산 서열을 지칭한다 (때때로, **아미노산** 서열의 경우 "CS01-FL-AA" 및 **핵산** 서열의 경우 "CS01-FL-NA"로서 본원에서 지칭됨). 유사하게, "CS01-LC"는 FVIII 폴리펩티드의 경쇄를 인코딩하는 코돈 변경된 핵산 서열 ("CS01-LC-NA") 또는 CS01 폴리뉴클레오티드 서열에 의해 인코딩된 FVIII 경쇄의 아미노산 서열 (때때로, "CS01-LC-AA"로서 본원에서 또한 지칭됨)을 지칭한다. 유사하게, CS01-HC, CS01-HC-AA 및 CS01-HC-NA는 FVIII 중쇄에서와 동일하다. 당업자에 의해 인지되는 바와 같이, 단지 코돈-변경된 구조체, 예를 들면, CS01, CS04, CS23 등 (가령, 이들은 Refacto와 비교하여 추가 아미노산 치환을 내포하지 않는다)의 경우에, 아미노산 서열은 동일할 것인데, 그 이유는 아미노산 서열이 코돈 최적화에 의해 변경되지 않기 때문이다. 따라서, 본 발명의 서열 구조체는 CS01-FL-NA, CS01-FL-AA, CS01-LC-NA, CS01-LC-AA, CS01-HC-AA, CS01-HC-NA, CS04-FL-NA, CS04-FL-AA, CS04-LC-NA, CS04-LC-AA, CS04-HC-AA, CS04-HC-NA, CS23-FL-NA, CS23-FL-AA, CS23-LC-NA, CS23-LC-AA, CS23-HC-AA 및 CS23-HC-NA를 포함하지만 이들에 한정되지 않는다.

[0145] **III. 코돈-변경된 인자 VIII 변이체**

[0146] 일부 구체예에서, 본 발명은 인자 VIII 변이체를 인코딩하는 코돈-변경된 폴리뉴클레오티드를 제공한다. 이들 코돈-변경된 폴리뉴클레오티드는 AAV-기초된 유전자 요법 구조체에서 투여될 때, 인자 VIII의 현저하게 향상된 발현을 제공한다. 코돈-변경된 폴리뉴클레오티드는 또한, 전통적으로 코돈-최적화된 구조체와 비교하여, 향상된 AAV-비리온 포장을 보여준다. 실시예 2 및 표 4에서 증명된 바와 같이, 출원인은 인간 야생형 인자 VIII 중쇄와 경쇄, 그리고 생체내에서 활성 FVIIIa 단백질의 성숙을 조장하는 퓨린 개열 부위를 내포하는 짧은, 14개 아미노산, B-도메인 치환된 링커 ("SQ" 링커)를 갖는 인자 VIII 폴리펩티드를 인코딩하는 3개의 코돈-변경된 폴리뉴클레오티드 (CS01-FL-NA, CS04-FL-NA 및 CS23-FL-NA)의 발견을 통해 이들 이점을 달성하였다.

[0147] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 최소한, 인자 VIII 중쇄 및 인자 VIII 경쇄를 인코딩하는 CS01, CS04 또는 CS23 내에 서열 (각각, 서열 번호 13, 1 및 20)과 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 당해 분야에서 공지된 바와 같이, 인자 VIII의 B-도메인은 생체내 활성에 불필요하다. 따라서, 일부 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 인자 VIII B-도메인을 완전하게 결여한다. 일부 구체예에서, 선천적 인자 VIII B-도메인은 퓨린 개열 부위를 내포하는 짧은 아미노산 링커, 예를 들면, CS01, CS04 또는 CS23의 아미노산 760-773으로 구성되는 "SQ" 링커 (각각, 서열 번호 2, 2 및 21) 구조체로 대체된다. "SQ" 링커는 BDL004로서 또한 지칭된다 (아미노산 서열의 경우 -AA 및 뉴클레오티드 서열의 경우 -NA).

[0148] 한 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 중쇄와 경쇄는 각각, 인간 인자 VIII 중쇄와 경쇄이다. 다른 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 중쇄와 경쇄는 다른 포유동물로부터 중쇄와 경쇄 서열 (가령, 돼지 인자 VIII)이다. 또 다른 구체예에서, 인자 VIII 중쇄와 경쇄는 키메라 중쇄와 경쇄 (가령, 인간 및 두 번째 포유류 서열의 조합)이다. 또 다른 구체예에서, 인자 VIII 중쇄와 경쇄는 다른 포유동물로부터 중쇄와 경쇄의 인간화 버전, 예를 들면, 인간에 투여될 때 결과의 펩티드의 면역원성을 감소시키기 위해 선별된 위치에서 인간 잔기가 치환되는 다른 포유동물로부터 중쇄와 경쇄 서열이다.

[0149] 인간 유전자의 GC 함량은 25% 이하로부터 90% 이상까지 폭넓게 변한다. 하지만, 일반적으로, 더욱 높은 GC 함량을 갖는 인간 유전자는 더욱 높은 수준에서 발현된다. 가령, Kudla et al. (PLoS Biol., 4(6):80 (2006))은 유전자의 GC 함량을 증가시키는 것이 일차적으로, 전사를 증가시키고 mRNA 전사체의 더욱 높은 항정 상태 수준을 달성함으로써, 인코딩된 폴리펩티드의 발현을 증가시킨다는 것을 증명한다. 일반적으로, 코돈-최적화된 유전자 구조체의 원하는 GC 함량은 60%와 동등하거나 또는 이보다 크다. 하지만, 선천적 AAV 유전체는 대략 56%의 GC 함량을 갖는다.

[0150] 따라서, 일부 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 선천적 AAV 비리온의 GC 함량 (가령, 대략 56% GC)에 더욱 가깝게 정합하는 CG 함량을 갖는데, 이것은 포유류 세포에서 발현을 위해 전통적으로 코돈-최적화되는 폴리뉴클레오티드의 바람직한 CG 함량 (가령, 60% GC에서 또는 이를 초과)보다 낮다. 실시예 1에서 개시된 바와 같이, 약 56%의 GC 함량을 갖는 CS04-FL-NA (서열 번호: 1)은 더욱 높은 GC 함량을 갖는 유사하게 코돈-변경된 코딩 서열과 비교하여, 향상된 비리온 포장을 갖는다.

[0151] 따라서, 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 60%보다 적다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전

반적인 GC 함량은 59%보다 적다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 58%보다 적다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 57%보다 적다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56% 이내이다.

[0152] 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 54% 내지 59%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 55% 내지 59%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56% 내지 59%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 54% 내지 58%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 55% 내지 58%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56% 내지 58%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 54% 내지 57%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 55% 내지 57%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56% 내지 57%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 54% 내지 56%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 55% 내지 56%이다.

[0153] 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56 ± 0.5%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56 ± 0.4%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56 ± 0.3%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56 ± 0.2%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56 ± 0.1%이다. 일부 구체예에서, 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드의 전반적인 GC 함량은 56%이다.

[0154] **A. 인자 VIII B-도메인 치환된 링커**

[0155] 일부 구체예에서, FVIII 중쇄 및 경쇄 사이에 연쇄 (가령, 야생형 인자 VIII에서 B-도메인)는 더욱 변경된다. AAV 포장 능력의 크기 제약으로 인해, B-도메인 결실된, 절단된, 및 또는 링커 치환된 변이체는 FVIII 유전자 요법 구조체의 효력을 향상시킬 것이다. 가장 전통적으로 이용되는 B-도메인 치환된 링커는 SQ FVIII의 링커인데, 이것은 B 도메인의 단지 14개 아미노산만 링커 서열로서 유지한다. 돼지 VIII의 다른 변이체 (U.S. 특허 번호 6,458,563에서 설명된 "OBI-1")는 CHO 세포에서 충분히 발현되고, 그리고 24개 아미노산의 약간 더욱 긴 링커를 갖는다. 일부 구체예에서, 본원에서 설명된 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 구조체는 SQ-유형 B-도메인 링커 서열을 포함한다. 다른 구체예에서, 본원에서 설명된 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 구조체는 OBI-1-유형 B-도메인 링커 서열을 포함한다.

[0156] 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 폴리펩티드는 야생형 인간 인자 VIII B-도메인 (FVIII-FL-AA; 서열 번호: 19)의 아미노산 760-762/1657-1667을 포함하는 SQ-유형 B-도메인 링커 (SFSQNPVLRHQR; BDL-SQ-AA; 서열 번호: 30)를 포함한다 (Sandberg et al. *Thromb. Haemost.* 85:93 (2001)). 일부 구체예에서, SQ-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 하나의 아미노산 치환을 갖는다. 일부 구체예에서, SQ-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 2개의 아미노산 치환을 갖는다.

[0157] 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 폴리펩티드는 야생형 인간 인자 VIII B-도메인 (FVIII-FL-AA; 서열 번호: 19)의 아미노산 760/1582-1667을 포함하는 Greengene-유형 B-도메인 링커를 포함한다 (Oh et al., *Biotechnol. Prog.*, 17:1999 (2001)). 일부 구체예에서, Greengene-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 하나의 아미노산 치환을 갖는다. 일부 구체예에서, Greengene-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 2개의 아미노산 치환을 갖는다.

[0158] 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 폴리펩티드는 야생형 인간 인자 VIII B-도메인 (FVIII-FL-AA; 서열 번호: 19)의 아미노산 760-769/1657-1667을 포함하는 연장된 SQ-유형 B-도메인 링커를 포함한다 (Thim et al., *Haemophilia*, 16:349 (2010)). 일부 구체예에서, 연장된 SQ-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 하나의 아미노산 치환을 갖는다. 일부 구체예에서, 연장된 SQ-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 2개의 아미노산 치환을 갖는다.

- [0159] 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 폴리펩티드는 야생형 돼지 인자 VIII B-도메인으로부터 아미노산 SFAQNSRPPSASAPKPPVLRHQR (서열 번호: 31)을 포함하는 돼지 OBI-1-유형 B-도메인 링커를 포함한다 (Toschi et al., Curr. Opin. Mol. Ther. 12:517 (2010)). 일부 구체예에서, 돼지 OBI-1-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 하나의 아미노산 치환을 갖는다. 일부 구체예에서, 돼지 OBI-1-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 2개의 아미노산 치환을 갖는다.
- [0160] 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 폴리펩티드는 야생형 인간 인자 VIII B-도메인 (FVIII-FL-AA; 서열 번호: 19)의 아미노산 760-772/1655-1667을 포함하는 인간 OBI-1-유형 B-도메인 링커를 포함한다. 일부 구체예에서, 인간 OBI-1-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 하나의 아미노산 치환을 갖는다. 일부 구체예에서, 인간 OBI-1-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 2개의 아미노산 치환을 갖는다.
- [0161] 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 폴리펩티드는 야생형 돼지 인자 VIII B-도메인으로부터 아미노산 SFSQNSRHQAYRYRRG (서열 번호: 32)를 포함하는 O8-유형 B-도메인 링커를 포함한다 (Toschi et al., Curr. Opin. Mol. Ther. 12:517 (2010)). 일부 구체예에서, 돼지 OBI-1-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 하나의 아미노산 치환을 갖는다. 일부 구체예에서, 돼지 OBI-1-유형 B-도메인 링커는 상응하는 야생형 서열에 비하여, 2개의 아미노산 치환을 갖는다.
- [0162] 인자 VIII 구조체로부터 B-도메인의 제거는 아마도 B-도메인이 활성화 동안 제거되기 때문에, 활성화된 효소 (가령, FVIIIa)의 활성화에 영향을 주는 것으로 보이지 않는다. 하지만, 인자 VIII의 B-도메인은 예로서, N- 또는 O-연결된 글리코실화에 의해 번역후 변형되는 여러 잔기를 내포한다. 야생형 인자 VIII B-도메인의 인실리코 분석 (Prediction of N-glycosylation sites in human proteins, R. Gupta, E. Jung and S. Brunak, in preparation (2004))은 이들 부위 중에서 최소한 4개가 생체내에서 글리코실화되는 것으로 예측한다. B-도메인 내에 이들 변형은 생체내에서 인자 VIII의 번역후 조절 및/또는 반감기에 기여하는 것으로 생각된다.
- [0163] 인자 VIII B-도메인이 성숙 인자 VIIIa 단백질에서 부재하긴 하지만, 전구체 인자 VIII 분자의 B-도메인 내에 글리코실화는 활성화에 앞서 상기 단백질의 순환 반감기를 증가시킬 수 있다. 따라서, 일부 구체예에서, 본원에서 설명된 인코딩된 인자 VIII 구조체의 폴리펩티드 링커는 생체내에서 글리코실화를 허용하는, 하나 또는 그 이상의 글리코실화 서열을 포함한다. 일부 구체예에서, 폴리펩티드 링커는 최소한 하나의 공통 글리코실화 서열 (가령, N- 또는 O-연결된 글리코실화 공통 서열)을 포함한다. 일부 구체예에서, 폴리펩티드 링커는 최소한 2개의 공통 글리코실화 서열을 포함한다. 일부 구체예에서, 폴리펩티드 링커는 최소한 3개의 공통 글리코실화 서열을 포함한다. 일부 구체예에서, 폴리펩티드 링커는 최소한 4개의 공통 글리코실화 서열을 포함한다. 일부 구체예에서, 폴리펩티드 링커는 최소한 5개의 공통 글리코실화 서열을 포함한다. 일부 구체예에서, 폴리펩티드 링커는 최소한 6, 7, 8, 9, 10개, 또는 그 이상의 공통 글리코실화 서열을 포함한다.
- [0164] 일부 구체예에서, 폴리펩티드 링커는 최소한 하나의 N-연결된 글리코실화 서열 N-X-S/T를 내포하고, 여기서 X는 P, S 또는 T 이외에 임의의 아미노산이다. 일부 구체예에서, 폴리펩티드 링커는 최소한 2개의 N-연결된 글리코실화 서열 N-X-S/T를 내포하고, 여기서 X는 P, S 또는 T 이외에 임의의 아미노산이다. 일부 구체예에서, 폴리펩티드 링커는 최소한 3개의 N-연결된 글리코실화 서열 N-X-S/T를 내포하고, 여기서 X는 P, S 또는 T 이외에 임의의 아미노산이다. 일부 구체예에서, 폴리펩티드 링커는 최소한 4개의 N-연결된 글리코실화 서열 N-X-S/T를 내포하고, 여기서 X는 P, S 또는 T 이외에 임의의 아미노산이다. 일부 구체예에서, 폴리펩티드 링커는 최소한 5개의 N-연결된 글리코실화 서열 N-X-S/T를 내포하고, 여기서 X는 P, S 또는 T 이외에 임의의 아미노산이다. 일부 구체예에서, 폴리펩티드 링커는 최소한 6, 7, 8, 9, 10개, 또는 그 이상의 N-연결된 글리코실화 서열 N-X-S/T를 내포하고, 여기서 X는 P, S 또는 T 이외에 임의의 아미노산이다.
- [0165] **B. 개열가능한 링커를 갖는 인자 VIII 변이체를 인코딩하는 코돈-변경된 폴리뉴클레오티드**
- [0166] *CS04 코돈 변경된 폴리뉴클레오티드*
- [0167] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 생체내에서 개열가능한 링커를 갖는 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 인자 VIII 중쇄를 인코딩하는 CS04-FL-NA (서열 번호: 1)의 일부인 CS04-HC-NA (서열 번호: 3)와 높은 서열 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 VIII 폴리펩티드의 경쇄는 인자 VIII 경쇄를 인코딩하는 CS04-FL-NA (서열 번호: 1)의 일부인 CS04-LC-NA (서열 번호: 4)와 높은 서열 동

일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 퓨린 개열 부위를 포함하는데, 이것은 생체내에서 성숙을 허용한다 (가령, 전구체 폴리펩티드의 생체내 발현 또는 투여 후).

- [0168] 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 95% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 96% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 97% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 98% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 99% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 99.5% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 동일하다.
- [0169] 일부 구체예에서, 인자 VIII 구조체의 폴리펩티드 링커는 CS04-FL-AA (서열 번호: 2)의 아미노산 760-773에 상응하는 14개-아미노산 링커를 인코딩하는 BDLO04 (서열 번호: 6)와 높은 서열 동일성을 갖는 세 번째 뉴클레오티드 서열에 의해 인코딩된다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 동일하다.
- [0170] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS04-FL-NA (서열 번호: 1)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-FL-NA (서열 번호: 1)와 동일하다.
- [0171] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS04-FL-AA (서열 번호: 2)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-FL-AA (서열 번호: 2)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-FL-AA (서열 번호: 2)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-FL-AA (서열 번호: 2)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-FL-AA (서열 번호: 2)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-FL-AA (서열 번호: 2)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-FL-AA (서열 번호: 2)와 동일하다.
- [0172] CS01 코돈 변경된 폴리뉴클레오티드
- [0173] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 생체내에서 개열가능한 링커를 갖는 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 인자 VIII 중쇄를 인코딩하는 CS01-FL-NA (서열 번호: 13)의 일부인 CS01-HC-NA (서열 번호: 24)와 높은 서열 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 VIII 폴리펩티드의 경쇄는 인자 VIII 경쇄를 인코딩하는 CS01-FL-NA (서열 번호: 13)의 일부인 CS01-LC-NA (서열 번호: 25)와 높은 서열 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 퓨린 개열 부위를 포함하는데, 이것은 생체내에서 성숙을 허용한다 (가령, 전구체 폴리펩티드의 생체내 발현 또는 투여 후).
- [0174] 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 95% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 96% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째

제와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 97% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 98% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 99% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 99.5% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 99.9% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 동일하다.

[0175] 일부 구체예에서, 인자 VIII 구조체의 폴리펩티드 링커는 CS01-FL-AA (서열 번호: 2)의 아미노산 760-773에 상응하는 14개-아미노산 링커를 인코딩하는 BDLO04 (서열 번호: 6)와 높은 서열 동일성을 갖는 세 번째 뉴클레오티드 서열에 의해 인코딩된다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDLO04 (서열 번호: 6)와 동일하다.

[0176] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS01-FL-NA (서열 번호: 13)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-FL-NA (서열 번호: 13)와 동일하다.

[0177] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS01-FL-AA (서열 번호: 2)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-FL-AA (서열 번호: 2)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-FL-AA (서열 번호: 2)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-FL-AA (서열 번호: 2)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-FL-AA (서열 번호: 2)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-FL-AA (서열 번호: 2)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-FL-AA (서열 번호: 2)와 동일하다.

[0178] CS23 코돈 변경된 폴리뉴클레오티드

[0179] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 생체내에서 개열가능한 링커를 갖는 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 인자 VIII 중쇄를 인코딩하는 CS23-FL-NA (서열 번호: 20)의 일부인 CS23-HC-NA (서열 번호: 22)와 높은 서열 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 VIII 폴리펩티드의 경쇄는 인자 VIII 경쇄를 인코딩하는 CS23-FL-NA (서열 번호: 20)의 일부인 CS23-LC-NA (서열 번호: 23)와 높은 서열 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 폴리펩티드 링커는 퓨린 개열 부위를 포함하는데, 이것은 생체내에서 성숙을 허용한다 (가령, 전구체 폴리펩티드의 생체내 발현 또는 투여 후).

[0180] 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 95% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 96% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 97% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 98% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 99% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소

한 99.5% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 99.9% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 동일하다.

[0181] 일부 구체예에서, 인자 VIII 구조체의 폴리펩티드 링커는 CS23-FL-AA (서열 번호: 21)의 아미노산 760-773에 상응하는 14개-아미노산 링커를 인코딩하는 BDL004 (서열 번호: 6)와 높은 서열 동일성을 갖는 세 번째 뉴클레오티드 서열에 의해 인코딩된다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDL004 (서열 번호: 6)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDL004 (서열 번호: 6)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDL004 (서열 번호: 6)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDL004 (서열 번호: 6)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 세 번째 뉴클레오티드 서열은 BDL004 (서열 번호: 6)와 동일하다.

[0182] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS23-FL-NA (서열 번호: 20)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-FL-NA (서열 번호: 20)와 동일하다.

[0183] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS23-FL-AA (서열 번호: 21)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-FL-AA (서열 번호: 21)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-FL-AA (서열 번호: 21)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-FL-AA (서열 번호: 21)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-FL-AA (서열 번호: 21)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-FL-AA (서열 번호: 21)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-FL-AA (서열 번호: 21)와 동일하다.

[0184] **C. 단일 사슬 인자 VIII 단백질을 인코딩하는 코돈-변경된 폴리뉴클레오티드**

[0185] B-도메인의 C 말단 단부에서 위치한 퓨린 개열 부위가 제거되는 인자 VIII 구조체는 인자 VIII 분자의 정상적인 성숙이 일어날 수 없음에도 불구하고, 단일 사슬 폴리펩티드로서 활성을 유지한다 (Leyte et al. (1991)). 유사하게, 약독화된 퓨린 부위를 갖는 B-도메인 결실된 인자 VIII 구조체 (R1664H 아미노산 치환을 내포)는 야생형 퓨린 개열 부위를 갖는 상응하는 인자 VIII 구조체보다 생물학적으로 더욱 활성이다 (Siner et al. (2013)). 따라서, 일부 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 단일 사슬 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 단일 사슬 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 폴리펩티드 링커를 포함한다. 폴리펩티드 링커는 퓨린 개열 부위를 포함하지 않는다.

[0186] *단일 사슬 CS04 코돈 변경된 폴리뉴클레오티드*

[0187] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 단일 사슬 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 임의선택적 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 인자 VIII 중쇄를 인코딩하는 CS04-FL-NA (서열 번호: 1)의 일부인 CS04-HC-NA (서열 번호: 3)와 높은 서열 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 VIII 폴리펩티드의 경쇄는 인자 VIII 경쇄를 인코딩하는 CS04-FL-NA (서열 번호: 1)의 일부인 CS04-LC-NA (서열 번호: 4)와 높은 서열 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 임의선택적 폴리펩티드 링커는 퓨린 개열 부위를 포함하지 않는다.

[0188] 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 95% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 96% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두

번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 97% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 98% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 99% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 99.5% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 최소한 99.9% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS04-HC-NA 및 CS04-LC-NA (서열 번호 3 및 4)와 동일하다.

[0189] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS04-SC1-NA (서열 번호: 9)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC1-NA (서열 번호: 9)와 동일하다.

[0190] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS04-SC2-NA (서열 번호: 11)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS04-SC2-NA (서열 번호: 11)와 동일하다.

[0191] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 단일 사슬 인자 VIII 변이체는 CS04-SC1-AA (서열 번호: 10; 인간 인자 VIII Δ (760-1667) (SPI; *HsFVIII Δ* (741-1648), SPE))와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS04-SC1-AA (서열 번호: 10)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC1-AA (서열 번호: 10)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC1-AA (서열 번호: 10)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC1-AA (서열 번호: 10)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC1-AA (서열 번호: 10)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC1-AA (서열 번호: 10)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC1-AA (서열 번호: 10)와 동일하다.

[0192] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 단일 사슬 인자 VIII 변이체는 CS04-SC2-AA (서열 번호: 12; 인간 인자 VIII Δ (772-1667) (SPI; *HsFVIII Δ* (753-1648), SPE))와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS04-SC2-AA (서열 번호: 12)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC2-AA (서열 번호: 12)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC2-AA (서열 번호: 12)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC2-AA (서열 번호: 12)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC2-AA (서열 번호: 12)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC2-AA (서열 번호: 12)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS04-SC2-AA (서열 번호: 12)와 동일하다.

[0193] 단일 사슬 CS01 코돈 변경된 폴리뉴클레오티드

[0194] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 단일 사슬 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 임의선택적 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중

쇄는 인자 VIII 중쇄를 인코딩하는 CS01-FL-NA (서열 번호: 13)의 일부인 CS01-HC-NA (서열 번호: 24)와 높은 서열 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 VIII 폴리펩티드의 경쇄는 인자 VIII 경쇄를 인코딩하는 CS01-FL-NA (서열 번호: 13)의 일부인 CS01-LC-NA (서열 번호: 25)와 높은 서열 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 임의선택적 폴리펩티드 링커는 퓨린 개열 부위를 포함하지 않는다.

[0195] 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 95% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 96% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 97% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 98% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 99% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 99.5% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 최소한 99.9% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS01-HC-NA 및 CS01-LC-NA (서열 번호 24 및 25)와 동일하다.

[0196] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS01-SC1-NA (서열 번호: 26)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC1-NA (서열 번호: 26)와 동일하다.

[0197] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS01-SC2-NA (서열 번호: 27)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS01-SC2-NA (서열 번호: 27)와 동일하다.

[0198] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 단일 사슬 인자 VIII 변이체는 CS01-SC1-AA (서열 번호: 10; 인간 인자 VIII Δ (760-1667) (SPI; *HsFVIII Δ* (741-1648), SPE))와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS01-SC1-AA (서열 번호: 10)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC1-AA (서열 번호: 10)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC1-AA (서열 번호: 10)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC1-AA (서열 번호: 10)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC1-AA (서열 번호: 10)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC1-AA (서열 번호: 10)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC1-AA (서열 번호: 10)와 동일하다.

[0199] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 단일 사슬 인자 VIII 변이체는 CS01-SC2-AA (서열 번호: 12; 인간 인자 VIII Δ (772-1667) (SPI; *HsFVIII Δ* (753-1648), SPE))와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS01-SC2-AA (서열 번호: 12)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC2-AA (서열 번호: 12)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC2-AA (서열 번호: 12)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC2-AA (서열

번호: 12)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC2-AA (서열 번호: 12)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC2-AA (서열 번호: 12)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS01-SC2-AA (서열 번호: 12)와 동일하다.

[0200] 단일 사슬 CS23 코돈 변경된 폴리뉴클레오티드

[0201] 한 구체예에서, 본원에서 제공된 코돈-변경된 폴리뉴클레오티드는 단일 사슬 인자 VIII 변이체 폴리펩티드를 인코딩하는 뉴클레오티드 서열을 포함한다. 인자 VIII 폴리펩티드는 인자 VIII 경쇄, 인자 VIII 중쇄, 그리고 중쇄의 C 말단을 경쇄의 N 말단에 연결하는 임의선택적 폴리펩티드 링커를 포함한다. 인자 VIII 폴리펩티드의 중쇄는 인자 VIII 중쇄를 인코딩하는 CS23-FL-NA (서열 번호: 20)의 일부인 CS23-HC-NA (서열 번호: 22)와 높은 서열 동일성을 갖는 첫 번째 뉴클레오티드 서열에 의해 인코딩된다. 인자 VIII 폴리펩티드의 경쇄는 인자 VIII 경쇄를 인코딩하는 CS23-FL-NA (서열 번호: 20)의 일부인 CS23-LC-NA (서열 번호: 23)와 높은 서열 동일성을 갖는 두 번째 뉴클레오티드 서열에 의해 인코딩된다. 임의선택적 폴리펩티드 링커는 퓨린 개별 부위를 포함하지 않는다.

[0202] 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 95% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 96% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 97% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 98% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 99% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 99.5% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 최소한 99.9% 서열 동일성을 갖는다. 일부 구체예에서, 첫 번째와 두 번째 뉴클레오티드 서열은 각각, CS23-HC-NA 및 CS23-LC-NA (서열 번호 22 및 23)와 동일하다.

[0203] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS23-SC1-NA (서열 번호: 28)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC1-NA (서열 번호: 28)와 동일하다.

[0204] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드는 CS23-SC2-NA (서열 번호: 29)와 높은 서열 동일성을 갖는 뉴클레오티드 서열을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 95% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 96% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 뉴클레오티드 서열은 CS23-SC2-NA (서열 번호: 29)와 동일하다.

[0205] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 단일 사슬 인자 VIII 변이체는 CS23-SC1-AA (서열 번호: 10; 인간 인자 VIII Δ (760-1667) (SPI; HsFVIII Δ (741-1648), SPE))와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS23-SC1-AA (서열 번호: 10)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC1-AA (서열 번호: 10)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC1-AA (서열 번호: 10)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC1-AA (서열 번호: 10)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC1-AA (서열 번호: 10)와

최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC1-AA (서열 번호: 10)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC1-AA (서열 번호: 10)와 동일하다.

[0206] 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 단일 사슬 인자 VIII 변이체는 CS23-SC2-AA (서열 번호: 12; 인간 인자 VIII Δ (772-1667) (SPI; HsFVIII Δ (753-1648), SPE))와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 코돈-변경된 폴리뉴클레오티드에 의해 인코딩된 인자 VIII 변이체는 CS23-SC2-AA (서열 번호: 12)와 높은 서열 동일성을 갖는 아미노산 서열을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC2-AA (서열 번호: 12)와 최소한 97% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC2-AA (서열 번호: 12)와 최소한 98% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC2-AA (서열 번호: 12)와 최소한 99% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC2-AA (서열 번호: 12)와 최소한 99.5% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC2-AA (서열 번호: 12)와 최소한 99.9% 동일성을 갖는다. 일부 구체예에서, 아미노산 서열은 CS23-SC2-AA (서열 번호: 12)와 동일하다.

[0207] **D. 인자 VIII 발현 벡터**

[0208] 일부 구체예에서, 본원에서 설명된 코돈-변경된 폴리뉴클레오티드는 발현 벡터 내로 통합된다. 발현 벡터의 무제한적 실례는 바이러스 벡터 (가령, 유전자 요법에 적합한 벡터), 플라스미드 벡터, 박테리오파지 벡터, 코스미드, 파지미드, 인공 염색체 등을 포함한다.

[0209] 바이러스 벡터의 무제한적 실례는 다음을 포함한다: 레트로바이러스, 예를 들면, 모로니 뮤린 백혈병 바이러스 (MLLV), 하비 뮤린 육종 바이러스, 뮤린 유방 종양 바이러스 및 라우스 육종 바이러스; 아데노바이러스, 아데노 연관된 바이러스; SV40-유형 바이러스; 폴리오마바이러스; 엡스타인 바르 바이러스; 유두종 바이러스; 헤르페스 바이러스; 우두바이러스; 그리고 소아마비 바이러스.

[0210] 일부 구체예에서, 본원에서 설명된 코돈-변경된 폴리뉴클레오티드는 유전자 요법 벡터 내로 통합된다. 일부 구체예에서, 유전자 요법 벡터는 레트로바이러스, 그리고 특히, 복제-결합성 레트로바이러스이다. 복제-결합성 레트로바이러스의 생산을 위한 프로토콜은 당해 분야에서 공지된다. 리뷰를 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., Clifton, N.J. (1991)를 참조한다.

[0211] 한 구체예에서, 유전자 요법 벡터는 아데노 연관된 바이러스 (AAV) 기초된 유전자 요법 벡터이다. AAV 시스템은 앞서 설명되었고 전반적으로 당해 분야에서 널리 공지된다 (Kelleher and Vos, *Biotechniques*, 17(6):1110-17 (1994); Cotten et al., *Proc Natl Acad Sci USA*, 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), 이들은 각각 모든 점에서 본원에 전체적으로 참조로서 편입된다). rAAV 벡터의 산출 및 이용에 관련된 상세는 예로서, U.S. 특허 번호 5,139,941 및 4,797,368에서 설명되고, 이들은 각각 모든 점에서 본원에 전체적으로 참조로서 편입된다. 특정한 구체예에서, AAV 벡터는 AAV-8 벡터이다.

[0212] 일부 구체예에서, 본원에서 설명된 코돈-변경된 폴리뉴클레오티드는 레트로바이러스 발현 벡터 내로 통합된다. 이들 시스템은 앞서 설명되었고 전반적으로 당해 분야에서 널리 공지된다 (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). 특정한 구체예에서, 레트로바이러스 벡터는 렌티바이러스 벡터이다 (가령, 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; U.S. 특허 번호 6,013,516 및 5,994,136을 참조한다).

[0213] 진핵 및 원핵 발현 벡터를 비롯하여, 매우 다양한 벡터가 세포 배양 동안 코돈-변경된 폴리펩티드로부터 인자 VIII 폴리펩티드의 발현에 이용될 수 있다. 일정한 구체예에서, 플라스미드 벡터가 세포 배양 동안 인자 VIII 폴리펩티드를 발현하는데 이용이 예기된다. 일반적으로, 숙주 세포와 양립성인 종으로부터 유래되는 레플리콘 및 제어 서열을 내포하는 플라스미드 벡터가 이들 숙주와 관련하여 이용된다. 벡터는 복제 부위뿐만 아니라 형질전환된 세포에서 표현형 선별을 제공할 수 있는 표지 서열을 보유할 수 있다. 플라스미드는 하나 또는 그 이상의 제어 서열, 예를 들면, 프로모터에 작동가능하게 연결된 인자 VIII 폴리펩티드를 인코딩하는 코돈-변경된 폴리뉴클레오티드를 포함할 것이다.

[0214] 원핵 발현을 위한 벡터의 무제한적 실례는 플라스미드, 예를 들면, pRSET, pET, pBAD 등을 포함하는데, 여기서

원핵 발현 벡터에서 이용된 프로모터는 lac, trc, trp, recA, araBAD 등을 포함한다. 진핵 발현을 위한 벡터의 실례는 다음을 포함한다: (i) 효모에서 발현의 경우, AOX1, GAP, GAL1, AUG1 등과 같은 프로모터를 이용하는, pAO, pPIC, pYES, pMET와 같은 벡터; (ii) 곤충 세포에서 발현의 경우, PH, p10, MT, Ac5, OpIE2, gp64, polh 등과 같은 프로모터를 이용하는, pMT, pAc5, pIB, pMIB, pBAC 등과 같은 벡터, 그리고 (iii) 포유류 세포에서 발현의 경우, CMV, SV40, EF-1, UbC, RSV, ADV, BPV 및 β -액틴과 같은 프로모터를 이용하는, pSVL, pCMV, pRc/RSV, pcDNA3, pBPV 등과 같은 벡터, 그리고 바이러스 시스템, 예를 들면, 우두바이러스, 아데노 연관된 바이러스, 헤르페스 바이러스, 레트로바이러스 등으로부터 유래된 벡터.

IV. 실시예

실시예 1 - 코돈-변경된 인자 VIII 변이체 발현 서열의 작제

A형 혈우병의 유전자 요법에 효과적인 인자 VIII 코딩 서열을 창출하기 위해, 2개의 장애물이 극복되어야 했다. 첫 번째, 전통적인 유전자 요법 전달 벡터 (가령, AAV 비리온)의 유전체 크기 제한 때문에, 인코딩된 인자 VIII 폴리펩티드는 상당히 단축되어야 했다. 두 번째, (i) 전달 벡터 내에 포장 상호작용을 안정시키고, (ii) mRNA 중간을 안정시키고, 그리고 (iii) mRNA의 전사/번역의 견고성을 향상시키기 위해 코딩 서열이 변경되어야 했다.

첫 번째 목적을 달성하기 위해, 출원인은 본원에서 "FVIII-BDD-SQ"로서 지칭되는 B-도메인 결실된 인자 VIII 변이체 구조체로 시작하였다. 이러한 구조체에서, B-도메인은 "SQ" 서열로서 지칭되는 14개 아미노산 서열로 대체된다. 재조합 FVIII-BDD-SQ는 상품명 REFACTO® 하에 판매되고, 그리고 A형 혈우병의 관리에 효과적인 것으로 나타났다. 하지만, 인자 VIII 중쇄와 경쇄에 대한 인간 야생형 핵산 서열을 포함하는, FVIII-BDD-SQ에 대한 선천적 코딩 서열은 유전자 요법 벡터에서 무효하게 발현된다.

선천적 FVIII-BDD-SQ의 불량한 발현을 다루기 위해, Ward et al. (Blood, 117:798 (2011)) 및 McIntosh et al. (Blood, 121, 3335-3344 (2013))에서 설명된 바와 같이 변형된, Fath et al. (PLoS ONE, 6:e17596 (2011))에서 설명된 코돈 최적화 알고리즘이 첫 번째 중간 코딩 서열 CS04a를 창출하기 위해 FVIII-BDD-SQ 서열에 적용되었다. 하지만, 출원인은 변형된 알고리즘을 이용하여 창출된 CS04a 서열이 상기 서열을 더욱 변형함으로써 향상될 수 있다는 것을 인식하였다. 따라서, 출원인은 CpG 섬, 그리고 AT-풍부한 및 GC-풍부한 스트레치의 국부 과다표현을 방지하면서, CpG 디뉴클레오티드를 제외하고, 아르기닌에 대한 CGC 코돈을 제외하고, 류신 및 세린 코돈 분포를 변화시키고, 고도로 보존된 코돈 쌍을 제외하고, 그리고 잠적 TATA 상자, CCAAT 상자 및 스플라이스 부위 요소를 제거하였다.

첫 번째, 변형된 알고리즘은 CpG-디뉴클레오티드를 내포하는 코돈 (가령, 아르기닌 코돈)을 비-CpG-디뉴클레오티드 코돈으로 조직적으로 대체하고, 그리고 인접한 코돈에 의해 창출된 CpG-디뉴클레오티드를 제거하거나/방지한다. CpG 디뉴클레오티드의 이러한 엄격한 회피는 통상적으로, DNA 백신의 근육내 주사 후 TLR-유도된 면역성을 예방하기 위해 행워진다. 하지만, 이렇게 하는 것은 코돈 최적화 가능성을 제한한다. 가령, 변형된 알고리즘은 CGX 아르기닌 코돈의 완전한 세트의 이용을 배제한다. 이것은 인간 세포에서 발현을 위한 유전자의 코딩에서 특히 파괴적인데, 그 이유는 CGC가 고도로 발현된 인간 유전자에서 가장 빈번하게 이용되는 아르기닌 코돈이기 때문이다. 추가적으로, 인접한 코돈에 의한 CpGs의 창출을 방지하는 것은 최적화 가능성을 더욱 제한한다 (가령, 함께 이용될 수 있는 코돈 쌍의 숫자를 제한한다).

TLR-유도된 면역성이 간-지향된, AAV-기초된 유전자 요법과 연관된 문제점일 것으로 예상되지 않기 때문에, CpGs를 포함하는 코돈 및 CpGs를 창출하는 인접한 코돈은 우선적으로 인자 VIII 경쇄를 코딩하는 서열에서 중간 코딩 서열 CS04a 내로 (가령, FVIII-BDD-SQ 코딩 서열의 3' 단부에서) 제외되었다. 이것은 바람직한 인간 코돈, 특히 아르기닌에 대한 것들의 더욱 빈번한 이용을 허용하였다. 하지만, CpG 부위의 높은 빈도를 갖는 코딩 서열의 영역인 CpG 섬의 창출을 방지하기 위해 주의를 기울였다. 이것은 전사 시작 부위의 하류에 CpG 도메인이 유전자 발현의 높은 수준을 증진한다는 것을 암시하는, Krinner et al. (Nucleic Acids Res., 42(6):3551-64 (2014))의 교시에 반한다.

두 번째, 변형된 알고리즘은 일정한 코돈, 예를 들면, 류신의 경우 CTG, 발린의 경우 GTG 및 글루타민의 경우 CAG를 배타적으로 적용한다. 하지만, 이것은 예로서, Haas et al. (Current Biology, 6(3):315-24 (1996))에서 제안된 바와 같이, 균형화된 코돈 이용의 원리를 위반한다. 변형된 알고리즘에 의한 바람직한 코돈의 과용을 설명하기 위해, 코돈 변경에 적용된 다른 규칙 (가령, CpG 빈도 및 GC 함량)에 의해 허용되는 경우에 대체 류신 코돈이 제외되었다.

세 번째, 변형된 알고리즘은 일정한 기준 (가령, CG-디뉴클레오티드의 존재)이 부합될 때, 코돈 쌍이 자연에서

어떻게 보존되는 지에 대한 고려 없이 이들을 대체한다. 진화에 의해 보존되었을 수 있는 유익한 성질을 설명하기 위해, 상기 알고리즘에 의해 대체된 대부분의 보존된 코돈 쌍 및 대부분의 보존된 바람직한 코돈 쌍은 코돈 변경에 적용된 다른 규칙 (가령, CpG 빈도 및 GC 함량)에 의해 허용되는 경우에, 예로서 Tats et al. (BMC Genomics 9:463 (2008))에서 설명된 바와 같이 분석되고 조정되었다.

[0224] 네 번째, 중간 코딩 서열에서 이용된 세린 코돈이 또한, 추가공되었다. 구체적으로, AGC, TCC 및 TCT 세린 코돈이 인간 코돈 사용빈도에 대해 전반적으로 더욱 우수하게 정합하도록, 더욱 높은 빈도에서 변형된 코딩 서열 내로 도입되었다 (Haas et al., 위와 같음).

[0225] 다섯 번째, TATA 상자, CCAAT 상자 요소 및 인트론/엑손 스플라이스 부위가 선별검사되고, 그리고 변형된 코딩 서열로부터 제거되었다. 코딩 서열을 변형할 때, AT-풍부한 또는 GC 풍부한 스트레치의 국부 과다표현을 방지하기 위해 주의를 기울였다.

[0226] 최종적으로, 코딩 서열 내에서 코돈 사용빈도를 최적화하는 것에 더하여, 중간 코딩 서열 CS04a를 더욱 정밀화할 때, 근원적인 AAV 비리온의 구조적 요건이 고려되었다. AAV 벡터 (가령, AAV 비리온의 핵산 부분)는 그들의 캡시드 내로 단일 가닥 DNA 분자로서 포장된다 (리뷰를 위해, Daya and Berns, Clin. Microbiol Rev., 21(4):583-93 (2008)을 참조한다). 벡터의 GC 함량은 이런 이유로, 유전체의 포장, 그리고 따라서, 생산 동안 벡터 수율에 영향을 줄 가능성이 높다. 많은 알고리즘과 유사하게, 여기에서 이용된 변형된 알고리즘은 최소한 60%의 GC 함량을 갖는 최적화된 유전자 서열을 창출한다 (Fath et al., PLoS One, 6(3):e17596 (2011) (erratum in: PLoS One, (6)3 (2011)을 참조한다). 하지만, AAV8 캡시드 단백질은 약 56%의 더욱 낮은 GC 함량을 갖는 뉴클레오티드 서열에 의해 인코딩된다. 따라서, 선천적 AAV8 캡시드 단백질 코딩 서열을 더욱 우수하게 모의하기 위해, 중간 코딩 서열 CS04a의 GC 함량은 56%까지 감소되었다.

[0227] 도면 2에서 도시된 결과의 CS04 코딩 서열은 56%의 전체 GC 함량을 갖는다. 상기 서열의 CpG-디뉴클레오티드 함량은 중간이다. 하지만, CpG 디뉴클레오티드는 코딩 서열의 하류 부분, 예를 들면, 인자 VIII 경쇄를 코딩하는 부분 내에 지배적으로 존재한다. CS04 서열은 야생형 인자 VIII (Genbank 수탁 M14113)에서 상응하는 코딩 서열에 79.77% 뉴클레오티드 서열 동일성을 갖는다.

[0228] 비교 목적을 위해, 여러 다른 코돈-최적화된, ReFacto 구조체가 제조되었다. CS01은 CS04에 대해 행위된 바와 같이 Ward et al.,에 의해 변형된 바와 같은 Fath et al.,의 코돈-최적화 알고리즘을 적용함으로써 작제되었다. 하지만, CS04와 달리, CS01 구조체는 CpG 섬을 전혀 내포하지 않는다. CS08 ReFacto 구조체는 Radcliff P.M. et al., Gene Therapy, 15:289-97 (2008)에서 설명된 바와 같이 코돈 최적화되었는데, 이의 내용은 전체적으로 모든 점에서 본원에 명시적으로 참조로서 편입된다. CS10 코돈-최적화된 ReFacto 구조체는 Eurofins Genomics (Ebersberg, Germany)로부터 획득되었다. CS11 코돈-최적화된 ReFacto 구조체는 Integrated DNA Technologies, Inc. (Coralville, USA)로부터 획득되었다. CH25 코돈-최적화된 ReFacto 구조체는 ThermoFischer Scientific의 GeneArt services (Regensburg, Germany)로부터 획득되었다. CS40 ReFacto 구조체는 야생형 인자 VIII 코딩 서열로 구성된다. CS23을 작제하는데 이용된 알고리즘은 코돈-최적화를 위한 온라인 도구인 JCAT 도구 (www.jcat.de)에 근거된다 (Grote et al., 2005; Nucl. Acids Res. W526-31). 상기 서열은 알부민 상과의 코돈 사용빈도를 더욱 반영하기 위해 더욱 변형되었다 (Mirzafian et al. 2014: Sc. Word Journal 2014, ID 639682). 각각의 ReFacto 코딩 서열 사이에 공유된 서열 동일성은 아래의 표 2에서 제시된다.

[0229] 표 2 - 코돈-변경된 인자 VIII 구조체에 대한 동일성 퍼센트 매트릭스.

표 2

	CS01	CS04	CS08	CS10	CS11	CS40	CH25	CS23
CS01	100%							
CS04	93.0%	100%						
CS08	80.7%	82.2.%	100%					
CS10	79.1%	79.4%	78.4%	100%				
CS11	78.3%	78.3%	78.1%	77.5%	100%			
CS40	79.6%	79.8%	76.7%	77.6%	75.4%	100%		
CH25	81.3%	85.1%	85.0%	79.9%	79.4%	75.8%	100%	
CS23	84.3%	89.2%	85.1%	80.3%	79.9	76.5%	93.2%	100%

- [0231] 각 구조체의 플라스미드는 상이한 합성 DNA 단편을 동일한 벡터 중추 플라스미드 (pCh-BB01) 내로 클로닝함으로써 작제되었다. 측면에서 접하는 AscI 및 NotI 효소 제한 부위를 갖는 Refacto-유형 BDD-FVIII 단편의 DNA 합성은 ThermoFischer Scientific (Regensburg, Germany)에 의해 행하여졌다. 벡터 중추는 간-특이적 무린 트랜스 티레틴 유전자로부터 유래된 프로모터/ 인핸서 서열, 개별 Refacto-유형 BDD-FVIII의 삽입을 위한 AscI 및 NotI 효소 제한 부위, 그리고 합성 polyA 부위를 포괄하는 2개의 측면에서 접하는 AAV2-유래된 반전된 말단 반복 (ITRs)을 내포한다. AscI 및 NotI 부위를 통해, 제조된 벡터 중추 및 삽입물의 걸찰 후, 결과의 플라스미드는 밀리그램 규모에서 증폭되었다. 이들 구조체의 Refacto-유형 BDD-FVIII 서열은 직접적인 염기서열결정 (Microsynth, Balgach, Switzerland)에 의해 실증되었다. 이러한 클로닝은 pCS40, pCS01, pCS04, pCS08, pCS10, pCS11, pCh25 및 pCS23으로 명명된 7개의 상이한 플라스미드 구조체를 유발하였다 (도면 19). 이들 구조체는 동일한 벡터 중추를 갖고, 그리고 동일한 B-도메인 결실된 FVIII 단백질 (Refacto-유형 BDD-FVIII)을 인코딩하지만, 그들의 FVIII 코딩 서열에 대해 다르다.
- [0232] AAV8-기초된 벡터는 Grieger JC, et al. (Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector, Mol Ther., Oct 6. (2015) doi: 10.1038/mt.2015.187. [인쇄에 앞서 Epub])에서 설명된 바와 같이, 3가지 플라스미드 형질감염 방법에 의해 제조되었는데, 이의 내용은 전체적으로 모든 점에서 본원에 명시적으로 참조로서 편입된다. HEK293 현탁액 세포가 상응하는 FVIII 벡터 플라스미드, 보조 플라스미드 pXX6-80 (아데노바이러스 보조 유전자를 보유), 그리고 포장 플라스미드 pGSK2/8 (rep2 및 cap8 유전자를 기여)을 이용한 플라스미드 형질감염에 이용되었다. AAV8 구조체를 단리하기 위해, 1 리터 배양액의 세포 펠렛이 Grieger et al. (2015, 위와 같음)에서 설명된 바와 같이, 아이오딕사놀 구배를 이용하여 처리되었다. 이러한 절차는 vCS01, vCS04, vCS08, vCS10, vCS11 및 vCH25로 불리는 벡터 제조물을 유발하였다. 벡터는 AAV2 반전된 말단 반복을 표적으로 하는 보편적인 qPCR 절차를 이용한 qPCR에 의해 정량되었다 (Aurnhammer, Human Gene Therapy Methods: Part B 23:18-28 (2012)). AAV2 반전된 말단 반복을 보유하는 대조 벡터 플라스미드는 표준 곡선을 준비하는데 적합하였다. 결과의 vCS04 구조체는 도면 7a-7c에서 서열 번호: 8로서 표시된다.
- [0233] 벡터 유전체의 완전성은 AAV 아가로즈 겔 전기이동에 의해 분석되었다. 전기이동은 Fagone et al., Human Gene Therapy Methods 23:1-7 (2012)에서 설명된 바와 같이 수행되었다. 간단히 말하면, AAV 벡터 제조물은 0.5% SDS의 존재에서 75 °C에서 10 분 동안 배양되고, 그리고 이후, 실온으로 식혀졌다. 거의 1.5E10 벡터 유전체 (vg)가 1% 1xTAE 아가로오스 겔 상에서 라인마다 부하되고, 그리고 7 V/cm의 겔 길이에서 60 분 동안 전기영동되었다. 겔은 이후, 2x GelRed (Biotium Cat# 41003) 용액에서 염색되고 ChemiDocTMMP (Biorad)에 의해 영상화되었다. 도면 20에서 도시된 결과는 vCS01, vCS04 및 vCS40 바이러스 벡터가 5kb 범위에서 상이한 띠에 의해 지시된, 동일한-크기산정된 유전체를 갖는다는 것을 증명한다 (도면 20, 라인 2-4). 대략 5.2 kb의 벡터 크기에도 불구하고, 상기 유전체는 다소간 과대 유전체 (4.7 kb의 AAV 야생형 유전체에 비하여)의 정확한 포장을 확증하는 균질한 띠이다. 모든 다른 vCS 벡터 제조물은 동일한 유전체 크기를 보여준다 (데이터 제시되지 않음).
- [0234] 캡시드 단백질의 예상된 패턴을 확증하기 위해, SDS PAGE, 그 이후에 은 염색이 벡터 vCS01, vCS04 및 vCS40으로 수행되었다 (도면 21). 상기 도면에서 보여 지는 바와 같이, 하류 정제 절차는 VP1, VP2 및 VP3의 예상된 단백질 패턴을 전시하는 고도로 정제된 물질을 유발하였다 (도면 21, 라인 2-4). 동일한 패턴이 모든 다른 바이러스 제조물에서 목격되었다 (도시되지 않음). AAV 제조물의 SDS-PAGE 절차는 표준 절차에 따라 행하여졌다. 각 레인은 1E10 vg의 개별 바이러스 구조체를 내포하였고, 그리고 제조업체의 사용설명서에 따라서 4-12% Bis-Tris (NuPAGE® Novex, Life Technologies) 겔에서 분리되었다. 은 염색은 제조업체의 사용설명서에 따라서 SilverQuest™ 키트 (Novex, Life Technologies)로 수행되었다.
- [0235] 놀랍게도, AAV 벡터 vCS01 및 vCS04는 vCS40 야생형 코딩 구조체 및 다른 코돈-최적화된 구조체와 비교하여, AAV 바이러스 생산에서 더욱 높은 수율에 의해 계측된 더욱 높은 비리온 포장을 가졌다. 표 3에서 보여 지는 바와 같이, vCS01 및 vCS04 벡터는 vCS40보다 훨씬 우수하게 복제되고, AAV 역가에서 5-7 배 수율 증가를 제공하였다.
- [0236] 표 3 - 세포 펠렛으로부터 정제될 때, AAV 벡터 구조체 vCS01, vCS04 및 vCD40으로 획득된 세포 배양액 리터당 수율.

표 3

구조체	백터 농도 [vg/ml] x10E12	수율 [vg /리터] x10E12	배수적 증가 wt와 대비하여
vCS40	2.0	11.0	-
vCS01	9.2	51.4	4.7
vCS04 - 표본 1	17.6	79.2	7.2
vCS04 - 표본 2	15.9	58.8	5.4

[0237]

[0238]

[0239]

[0240]

[0241]

[0242]

[0243]

실시예 2 - 코돈-변경된 인자 VIII 변이체 발현 서열의 생체내 발현

코돈-변경된 인자 VIII 변이체 서열의 생물학적 효능을 시험하기 위해, 실시예 1에서 설명된 ReFacto-유형 FVIII 구조체가 인자 VIII을 결여하는 생쥐에게 투여되었다. 간단히 말하면, 이들 검정은 생쥐의 체중 킬로그램 당 4E12 백터 유전체 (vg)의 꼬리 정맥 주사에 의해 C57Bl/6 FVIII 녹아웃 (ko) 생쥐 (군마다 6-8마리 동물)에서 수행되었다. 혈액은 안와후방 천자에 의해 주사 후 14 일에 뽑혀졌고, 그리고 혈장이 표준 절차를 이용하여 준비되고 동결되었다. 14 일자에 발현 수준이 선택되었는데, 그 이유는 이러한 시점에서 저해성 항체의 영향이 최소이기 때문인데, 이들 항체는 추후 시점에 이러한 생쥐 모형의 일부 동물에서 목격된다. 생쥐 혈장에서 FVIII 활성은 제조업체 (Technoclon, Vienna, Austria)에 의해 제안된 바와 같이, 단지 약간만 변형되어 수행된 Technochrome FVIII 검정을 이용하여 결정되었다. 상기 검정을 위해, 혈장 표본은 적절하게 희석되고, 그리고 트롬빈, 활성화된 인자 IX (FIXa), 인지질, 인자 X 및 칼슘을 내포하는 검정 시약과 혼합되었다. 트롬빈에 의한 FVIII 활성화 이후에, FIXa, 인지질 및 칼슘과의 복합체가 형성된다. 이러한 복합체는 FX를 활성화된 FX (FXa)로 활성화시키고, 이것은 차례로, 발색 기질로부터 파라-니트로아닐리드 (pNA)를 개열한다. pNA 형성의 동역학은 405 nm에서 측정된다. 상기 비율은 표본 내에 FVIII 농도에 직접적으로 비례한다. FVIII 농도는 참고 곡선으로부터 판독되고, 그리고 결과는 IU FVIII/밀리리터에서 제공된다.

아래 표 4에서 제시된 결과는 상업적인 알고리즘을 이용하여 설계된 코돈-변경된 서열 (CS10, CS11 및 CH25)이 야생형 BDD-인자 VIII 구조체 (CS40)와 비교하여 BDD-인자 VIII에서 단지 근소한 증가 (3-4 배)만을 제공한다는 것을 증명한다. 유사하게, Radcliffe et al.에서 설명된 바와 같이 제조된 코돈-변경된 BDD-인자 VIII 구조체 (CS08)는 BDD-FVIII 발현에서 단지 3-4 배 증가만을 제공하였다. 이러한 결과는 Radcliff et al.에서 보고된 결과와 일치한다. 놀랍게도, CS01, CS04 및 CS23 구조체는 생체내 생물효능 검정에서 훨씬 높은 BDD-FVIII 발현 (각각, 18-, 74- 및 30-배 증가)을 제공하였다.

표 4 - 상이한 AAV 백터 구조체에 의해 유도된 FVIII-녹아웃 생쥐의 혈장에서 FVIII의 발현.

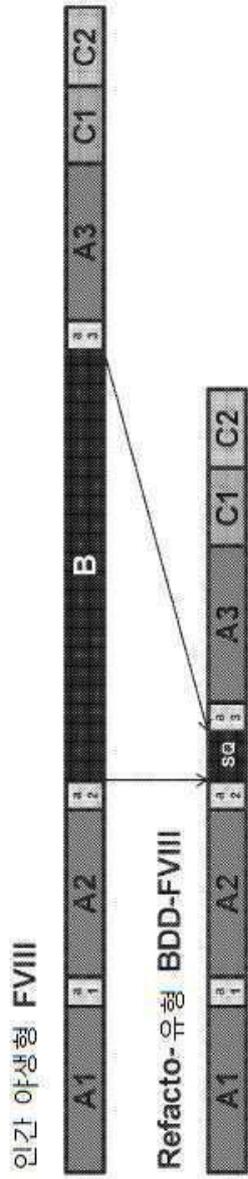
표 4

구조체	코돈 알고리즘	평균 FVIII 14 일자에 발현 [IU/ml]	표준 편차	생쥐의 숫자	배수적 증가 wt와 대비하여
vCS40	인간 야생형	0.03	0.03	12	-
vCS01	출원인'	0.55	0.28	22	18.3
vCS04	출원인'	2.21	1.20	55	73.7
vCS08	Radcliffe et al.	0.11	0.01	6	3.6
vCS10	Eurofins	0.09	0.01	7	3.0
vCS11	IDT	0.08	0.02	8	2.7
vCH25	GeneArt	0.13	0.12	18	4.3
vCS23	출원인'	0.91	0.32	5	30.3

본 명세서에 기재된 실시예와 구체에는 단지 예시적인 목적을 위한 것이고, 그리고 이에 비추어 다양한 개변은 당해 분야의 당업자에게 제시되고 본원의 기술적 사상과 이해범위 및 첨부된 청구항의 범위 내에 포함되는 것으로 이해된다. 본원에서 인용된 모든 간행물, 특허 및 특허 출원은 본원에 전체적으로 참조로서 편입된다.

도면

도면1



도면2a

CS04-FL-NA

atgcagattgagctgagcacctgcttcttctctgtgctgctgaggttctgcttctctgcccaccagga
gatactacctggggctgtggagcttcttgggactacatgcaagtctgacctgggggagctgctgt
ggatgccagggttccccaccagagtgcccaaatccttcccattcaaacacctctgtgggtctacaagaag
acctcttftgtggagttcactgaccacctgttcaacattgccaaccaggccacctggatgggac
tcttgggacccaccattcaggctgaggtgtatgacctgtgggtatcaacctcaagaacatggcctc
ccacctgtgagcctgcacatgctgtggggctcagctactggaaggcctctgagggggctgagtatgat
gaccagacctcccagagggagaaggaggatgacaaagtgttccctgggggagccacacctatgtgt
ggcaggctcctcaaggagaatggccccatggcctctgacctctgctgacctactcctacettct
tcatgtggacctggtcaaggacctcaactctggactgattggggcctgctggtgtgcagggagggc
tccctggccaaagagaagaccagacctgcacaagtctcattctctgtttgctgtctttgatgagg
gcaagagctggcactctgaaaccaagaactcctgatgcaggacagggatgctgctctgcccagggc
ctggccaagatgcacactgtgaaatggctatgtgaacaggagcctgctggactcattggctgccac
aggaatctgtctactggcatgtgatggcatggggacaacctgaggtgcactccatttctctgg
agggccacaccttctgttcaggaaaccacagaccaggccagcctggagatcagccccatcaccttct
cactgccagacctgctgatggacctcggacagttctctgtgttctgcccacatcagctcccaccag
catgatggcatggaggcctatgtcaagggtggacagctgacctgaggagccacagctcaggatgaaga
acaatgaggaggctgaggactatgatgatgacctgactgactctgagatggatgtggtccgctttga
tgatgacaacagcccatccttctcagatcaggctctgtggccaagaaacccccaaagacctgggtg
cactacattgctgctgaggaggaggactgggactatgccccactggtcctggcccctgatgacagga
gctacaagagccagctacctcaacaatggcccacagaggattggacgcaagtacaagaaagtccaggtt
catggctacactgatgaaacctcaagaccaggaggccatcagcatgagtctggcatcctggggc
ccactcctgtatggggaggtgggggacacctgctcatcatctcaagaaccaggcctccaggcct
acaacatctaccacatggcatcactgatgtcaggccccctgtacagccgaggctgccaaaggggt
gaaacacctcaaggacttccccattctgcttggggagatcttcaagtacaagtggactgtcactgtg
gaggatggaccaaccaaaatctgaccccagggtgctcaccagatactactccagctttgtgacatgg
agagggacctggcctctggcctgattggcccactgctcctctgctacaaggagtctgtggaccagag
gggaaaccagatcatgtctgacaagaggaatgtgattctgttctctgtctttgatgagaacaggagc
tggtaacctgactgagaacattcagcgttctctgcccaccctgctgggggtgcagctggaggacctg
agttccaggccagcaacatcatgcactccatcaatggctatgtgtttgacagcctccagctttctgt
ctgctgcatgagggtggcctactgggtacattcttctattggggcccagactgacttcttctgtc
tcttctctgtgctacaccttcaaacacaagatgggtgatgaggacacctgacctcttcccattct
ctggggagactgtgttcatgagcatggagaacctggcctgtggatcttgggatgccacaactctga
cttccgcaacaggggcatgactgcccctgctcaaggtctcctcctgtgacaagaacactggggactac
tatgaggacagctatgaggacatctctgctacctgctcagcaagaacaatgccattgagcccagga
gcttcagccagaatccacctgtcctgaaacgccaccagaggagatcaccaggaccacctccagtc
tgaccaggaggagattgactatgatgacaccttctgtggagatgaagaaaggactttgacatc
tatgacgaggacgagaaccagagcccaaggagcttcagagaagaagaccaggcactactcattgctg
ctgtggagcgcctgtgggactatggcatgagctccagccccatgtcctcaggaaacaggggccagtc
tggctctgtgcccaggttcaagaaagtggctcttccaagagttcactgatggcagcttccccagccc
ctgtacagaggggagctgaatgagcacctgggactcctgggcccatacatcagggtgaggtggagg
acaacatcatgggtgaccttccgcaaccaggcctccaggccctacagcttctacagctcctcctcag
ctatgagGaggaccagaggcagggggctgagccaagcaagaactttgtgaaaccaatgaaaccaag
acctactctggaaagtccagcaccacatggccccaccaaggatgagtttgactgcaaggcctggg

(계속)

도면2b

```

octacttctctgatgtggaacctggagaaggatgtgcaactctggcctgattggcccactcctgggtctg
ccacaccaacaccctgaacctgcccaggaaggcaagtgaactgtgcaaggagtttgcctcttctctc
accatctttgatgaaaccaagagctggtaacttcaactgagaacatggagcgcaactgcaggggcccat
gcaacattcagatggaggacccccaccttcaaagagaactaccgcttccatgccatcaatggctacat
catggacacacctggcctgggcttgcctatggcccagaccagaggatcaagggtgtaacctgctttctatg
ggctccaatgagaacatcactccatccacttctctgggcatgtcttcaactgtgcgcaagaaggagg
agtacaagatggcctgtacaacctetacctggggcttttgagaactgtggagatgctgcctccaa
agctggcatctggaggggtggagtgctcattggggagcactgcctgctggcatgagcaccctgttc
ctggctacagcaacaagtgcagacccccctgggaatggcctctggccaatcagggaacttccaga
tcaactgctctggccagatggccagtgggcccccagctggccaggtccactactctggatccat
caatgctggagcaccaggagccatcagctggatcaaaaggagcctgctggccccatgatcctc
catggcatcaagaccaggggggcaggcagaagtctccagcctgacatcagccagttcatcatca
tgtacagcctggatggcaagaaatggcagacctacagagcactccactggaacactcatggtctt
ctttggcaatgtggacagctctggcatcaagcacaacatcttcaaccccacatcatcggcagatac
atcaggtgcacccccaccactacagcatccgcagcaccctcaggatggagctgatgggctgtgacc
tgaactcctgcagcatgcccctgggcatggagagcaaggccatttctgatgccagatcactgcctc
cagctacttcaccaacatgtttgccacctggagcccaagcaaggccaggtgcacctccagggaagg
agcaatgctggagggcccagggtcaacaacccaaaggagtgctgcagggtggactccagaagacca
tgaaggctcaactggggtgaccaccccagggggtcaagagcctgctcaccagcatgtatgtgaaggagt
cctgatcagctccagccaggatggccaccagtgggaccctcttctccagaatggcaagggtcaagggt
ttccagggcaaccaggacagcttcacCctgtgggtgaacagcctggacccccctcctgaccagat
acctgaggttcaaccccagagctgggtccaccagattgccttgaggatggaggtcctgggatgtga
ggcccaggacctgtactga (서열 번호:1)
    
```

도면3

CS04-FL-AA

```

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDAR' FPPRVPKSFPFNTSVVYK
KTLFVBEFTDHLFNIAKPRPPWMLLGPITQAEVYDTVVI TLKNMASHPVSLHAVGVSYWKASEGAEY
DDQTSQREKEDDKVFPGGSHYVWQVLKENGFMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCRE
GSLAKEKTQTLHKFIFLFAVFEDEKGSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGC
HRKSVYWHVIGMGTTPPEVHSIFLEGHTFLVRNHRQASLEISPIITFLTAQTLMLDLGQFLLFCHTSSH
QHDGMEAYVKVDSCEPEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDONSPSFIQIRSVAKKHPKTW
VHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQIRGRKYKKVRFMAYTDETFKTREAIQHESGIL
GPLLYGEVGDITLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVT
VEDGPTKSDPRCLTRYYSFVNMERDLASGLIGPLLI CYKESVDQRGNQIMSDKRNVLFSVFDENR
SWYLTENIQRFLENPAGVOLEDPEFOASNIHMSINGYVFDLSQLSVCLHEVAYWYILSIGAQDFLS
VFFSGYTFKHKMVEYEDTLTFPFSGETVFMSENPGWLILGCHNSDFRNRGMTALLKVSSCDKNTGD
YYEDSYEDISAYLLSKNNAIEPRFSQNPVLRHQREITRTTLOSQDEEIDYDDTISVEMKKEDEFD
IYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSPHVLRNRAQSGSVQFKKVVFOEFTDGSFTQ
PLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPKKNFVKPNET
KTYFWKVQHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQFALE
FTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQIRWYLLS
MGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTL
FLVYSNKQOTPLGMASGHIRDFQITASGOYQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMI
IHGIKTQGARQKFSLSYLSQFIIMYSLDGKKWQTYRCNSTGTLMVFFGNVDSSGKHNIFNPPIIAR
YIRLHPHTYSIRSTLRMELMGCDLNSCSMPLGMEKAI SDAQITASSYFTNMFATWSPSKARLHLQG
RSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVK
VFQGNQDSFPPVNSLDPPLLTRYLRIPHQSWVHQIALRMEVLGCEAQDLY (서열 번호: 2)
    
```

도면4

CS04-HC-NA

```

                                gcc
accaggagat actacctggg ggetgtggag ctttcttggg actacatgca gtctgacctg
ggggagctgc ctgtggatgc caggttcccc cccagagtgc ccaaatecett cccattcaac
acctotgtgg tetacaagaa gacctotttt gtggagtcca ctgaccacct gttcaacatt
gccaaaacca ggecacctctg gatgggactc ctgggacceca ccattcagge tgagggtgat
gacctgtgg teatcacctt caagaacatg gcctcccacc ctgtgagcct gcctgctgtg
ggggteagct actggaaggc ctctgagggg gctgagtatg atgaccagac ctcccagagg
gagaaggagg atgacaagt gttccctggg ggcaaccaca cotatgtgtg gcaggtectc
aaggagaatg gccccatggc ctctgacceca ctctgcctga cctaectcta cctttctcat
gtggacctgg tcaaggacct caactctgga ctgattgggg cectgctggt gtgcaggggag
ggctccctgg ccaaagagaa gaccocagacc ctgcacaagt teattctcct gtttgctgtc
tttgatgagg gcaagagctg gcactctgaa accaagaact cectgatgca ggacagggat
gctgcctctg ccagggectg gcccaagatg cacactgtga atggctatgt gaacaggagc
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg
acaaccctct aggtgcactc cattiectctg gagggccaca ccttctctggt caggaaccac
agacaggcca gccctggagat cagccccatc accttctctc ctgcccagac cctgctgatg
gacctcggac agttctgtct gttctgccac atcagctccc accagcatga tggcatggag
gcctatgtca aggtggacag ctgcccctgag gagccacagc tcaggatgaa gaacaatgag
gaggctgagg actatgatga tgacctgaet gactctgaga tggatgtggt ccgctttgat
gatgacaaca gcccatcctt catteagatc aggtctgtgg ccaagaaaca ccccagacc
tggtgtcact acattgctgc tgaggaggag gactgggact atgccccact ggtcctggcc
cctgatgaca ggagctacaa gagccagtac ctcaacaatg gcccacagag gattggacgc
aagtacaaga aagtcaagtt catggcctac actgatgaaa ccttcaagac cagggaggcc
attcagcatg agtctggcat cctgggcccc ctctgtatg gggagggtgg ggacaccctg
ctcctcctct tcaagaacca ggctccagg cctacaaca totaccaca tggcatcact
gatgtcaggc cctgtacag ccgcaaggctg ccaaaggggg tgaaacacct caaggacttc
cccattctgc ctggggagat ctcaagtac aagtggactg tcactgtgga ggatggacca
accaaatctg accccaggtg cctcaccaga tactactcca gotttgtgaa catggagagg
gacctggcct ctggcctgat tggcccactg ctcatctgct acaaggagtc tgtggaccag
aggggaaacc agatcctatg tgacaagagg aatgtgatto tgttctctgt ctttgatgag
aacaggagct ggtacctgac tgagaacatt cagcgttcc tgcccacccc tgctgggtg
cagctggagg acctgagtt ccaggccagc aacatcatgc actccatcaa tgctatgtg
tttgacagcc tcagcttctc tgtctgcctg catgagggtg cctactggta cttctttct
attggggccc agactgactt cctttctgtc ttcttctctg gctacacctt caaacacaag
atggtgtatg aggacacct gapcctcttc ccattctctg gggagactgt gttcatgagc
atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttcct caacaggggc
atgaetgcc tgctcaaagt ctctcctgtg gacaagaaca ctggggacta ctatgaggac
agctatgagg acatctctgc ctacctctc agcaagaaca atgccattga gccccagg
(서열 번호: 3)

```

도면5

CS04-LC-NA

```

                                g agatcaaccag gaccacocctc
cagtctgacc aggaggagat tgactatgat gacaaccattt ctgtggagat gaagaaagag
gactttgaca tctatgacga ggacgagaac cagagoccaa ggagcttcca gaagaagacc
aggcactact tcattgctgc tgtggagcgc ctgtgggact atggcatgag ctccagcccc
catgtectca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggctctc
caagagtcca ctgatggcag cttcacccag cccctgtaca gaggggagct gaatgagcac
ctgggactcc tgggcccata catcagggct gaggtggagg acaacatcat ggtgacctc
cgcaaccagg cctccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac
cagaggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac
ttctggaaag tccagcacca catggccccc accaaggatg agtttgactg caaggcctgg
gocfactctc ctgatgtgga cctggagaag gatgtgcact ctggcctgat tggcccactc
ctggctctgc acaccaaac cctgaacct gcccatggaa ggcaagtgac tgtgcaggag
tttgccctct tcttcacat ctttgatgaa accaagagct ggtacttca tgaagaacatg
gagcgcacat gcagggcccc atgcaacatt cagatggagg accccacctt caaagagaac
taaccgttcc atgccatcaa tggctacatc atggacaccc tgcctgggct tgtcatggcc
caggaccaga ggatcaggtg gtaccctgtt tctatgggct ccaatgagaa cctcactcc
atccaactct ctggcctgt cttcactgtg cgcaagaagg aggagtaca gatggccctg
tacaacctct accctggggt ctttgagact gtggagatgc tgcctccaa agctggcatc
tggagggtgg agtgcctcat tggggagcac ctgcatgctg gcatgagcac cctgttctg
gtctacagca acaagtcca gacccccctg ggaatggct ctggccacat cagggacttc
cagatcactg cctctggcca gtatggccaq tggccccc aactggccag gctccactac
tctggatcca tcaatgctg gagcaccag gagccattca gctggatcaa agtggacctg
ctggccccc a tgatcatcca tggcatcaag acccaggggg ccaggcagaa gttctccagc
ctgtacatca gccagttcaf catcatgtac agcctggatg gcaagaaatg gcagacctac
agaggcaact ccactggaac actcatggtc ttctttggca atgtggacag ctctggcatc
aagcacaaca tcttcaacc cccaatcctc gccagataca tcaggctgca ccccaccac
tacagcatcc gcagcaccct caggatggag ctgatgggct gtgacctgaa ctctgcagc
atgccctgg gcattggagag caaggocatt tctgatgcc agatcactgc ctccagctac
ttcaccaca tgtttggccac ctggagocca agcaaggcca ggtgcacct ccagggaagg
agcaatgct ggaggcccca ggtcaacaac ccaaaggagt ggtgcagggt ggacttccag
aagaccatga aggtcactgg ggtgaccacc cagggggtca agagcctgct caccagcatg
tatgtgaagg agttcctgat cagctccagc caggatggcc accagtggac cctctctctc
cagaatggca aggtcaagg gtctcagggc aaccaggaca gcttcaccc tgtgggtgac
agcctggacc ccccctcct gaccagatac ctgaggatc accccagag ctgggtccac
cagattgccc tgaggatgga ggtcctggga tgtgaggecc aggaactgta c

```

(서열 번호: 4)

도면6

BDL001 - ago ttctctctcaga atccacctgt cctgaagaga caccagaga (서열 번호: 5)
BDL004 - ago ttccagccaga atccacctgt cctgaacac caccagagg (서열 번호: 6)
BDL023 - ago ttccagccaga acccccctgt gctgaagagg caccagagg (서열 번호: 7)

도면7a

CS04-AV-NA

```

1   togcgcggttt  cggatgatgac  ggtgaaaacc  tctgacacat  gcagctcccg  gagacggtca
61  cagcttgtct  gtaagcggat  gccgggagca  gacaagcccc  tcagggcgcg  tcagcgggtg
121  ttggcgggtg  tcggggctgg  cttaaactatg  cggcctcaga  gcagattgta  ctgagagtgc
181  accatattgcg  gtgtgaaata  ccgcacagat  gcgtaaggag  aaaataccgc  atcaggcgcc
241  attcgcatt  caggctgcgc  aactgttggg  aagggcgatc  ggtgcgggcc  tcttcgctat
301  tacgccagct  ggcgaaagg  ggatgtgctg  caaggcgatt  aagttgggtg  acgccagggg
361  tttccagtc  acgacgttgt  aaaaacgacg  ccagtgaatt  cctcgagatt  taaatgacgt
421  tggcactcc  ctctctgcgc  gctcctcgc  tcactgaggg  cgggggacca  aaggtcgccc
481  gacgcocgg  ctttgcccgg  gccgcctcag  tgagcgagcg  agcgcgcaga  gagggagtgg
541  ccaactccat  cactagggg  tcttgagttt  aaactctgct  gacgattcga  gcttgggctg
601  caggtcgagg  gcactgggag  gatgttgagt  aagatggaaa  actactgatg  acccttgca
661  agacagagta  ttaggacatg  ttgaaacagg  gccggggcga  tcagcaggta  gctctagagg
721  atcccctct  dtctgcacat  ttctgtaggc  gagtgttccg  atactcta  ctccctaggg
781  aaggttcata  tttgtgtagg  ttacttattc  tcttttgg  gactaagtca  ataatcagaa
841  tcageagggt  tggagtcagc  ttggcaggg  tcagcagcct  gggttggaag  gagggggtat
901  aaaagccct  tcaccaggag  aagccgtcac  acagactagg  cgcgcaccg  ccaccatgca
961  gattgagctg  agcaactgct  tcttctgtg  cctgctgagg  ttctgattct  ctgccaccag
1021  gagatactac  ctggggctg  tggagctttc  ttgggactac  atgcagtctg  acctggggga
1081  gctgcctgtg  gatgccaggt  tcccaccag  agtgcocaaa  tcttcccat  tcaaacctc
1141  tgtggtctac  aagaagacc  tctttgtgga  gttcactgac  cacctgttca  acattgcaa
1201  acccaggcca  ccttgatgg  gactcctggg  adccaccatt  caggctgagg  tgtatgacac
1261  tgtggtctac  accctcaaga  acatggcctc  ccaccctgtg  agcctgcatg  ctgtgggggt
1321  cagctactgg  aaggcctctg  aggggctga  gtatgatgac  cagacotccc  agaggagaa
1381  ggaggatgac  aaagtgttcc  ctggggcag  ccacacctat  gtgtgagag  tctcaagga
1441  gaatggccc  atggcctctg  acccactctg  cctgacctac  tctaccttt  ctcatgtgga
1501  cctggtcaag  gacctcaact  ctggactgat  tggggcctg  ctggtgtgca  gggagggctc
1561  cctggccaaa  gagaagacc  agaccctgca  caagtccatt  ctctgtttg  ctgtctttga
1621  tgaggcaag  agctggcact  ctgaaaccaa  gaactccctg  atgcaggaca  gggatgctgc
1681  ctctgccagg  gcctggccca  agatgcacac  tgtgaatggc  tatgtgaaca  ggagcctgcc
1741  tggactcatt  gctgcccaca  ggaatctgt  ctactgcat  gtgattggca  tggggacaac
1801  ccctgaggtg  cactccattt  tcttgaggg  ccacaccttc  ctggtcagga  accacagaca
1861  ggccagcctg  gagatcagcc  ccatacctt  cctcactgac  cagaccctgc  tgatggacct
1921  cggacagttc  ctgctgttct  gccacatcag  ctcccaccag  catgatggca  tggaggccta
1981  tgtcaaggtg  gacagctgcc  ctgaggagcc  acagctcagg  atgaagaaca  atgaggaggc
2041  tgaggactat  gatgatgacc  tgactgactc  tgagatggat  gtggtccgct  ttgatgatga
2101  caacagccca  tcttcattc  agatcaggtc  tgtggccaag  aaacacccca  agacctgggt
2161  gcactacatt  gctgctgagg  aggaggactg  ggactatgcc  ccactggtcc  tggccctga
2221  tgacaggagc  tacaagagcc  agtaoctcaa  caatggccca  cagaggattg  gacgcaagta
2281  caagaaagtc  aggttcattg  cctacactga  tgaaaccttc  aagaccaggg  aggcattca
2341  gcatgagttc  ggcattctgg  gccactctct  gtatggggag  gtgggggaca  cctgtctcat
2401  catcttcagg  aaccaggcct  ccaggcccta  caacatctac  ccacatggca  tcactgatgt
2461  caggccctg  tacagccgca  ggctgccaaa  gggggtgaaa  cacctcaagg  acttcccat

```

도면7b

2521 tctgcctggg gagatcttca agtacaagtg gactgtcact gtggaggatg gaccaaccaa
 2581 atctgacccc aggtgcctca ccagatacta ctccagcttt gtgaacatgg agagggacct
 2641 ggctcttggc ctgattggcc cactgctcat ctgctacaag gagtctgtgg accagagggg
 2701 aaaccagatc atgtctgaca agaggaatgt gattctgttc tctgtctttg atgagaacag
 2761 gagctggtac ctgactgaga acattcagcg cttctgtccc aacctgctg ggtgacagct
 2821 ggaggacctt gaggttccagg ccagcaacat catgcactcc atcaatggct atgtgtttga
 2881 cagcctccag ctttctgtct gcctgcctga ggtggcctac tggtagattc tttctattgg
 2941 ggcccagact gacttccctt ctgtctcttt ctctggctac accttcaaac acaagatggt
 3001 gtatgaggac acctgaccc tcttcccatt ctctggggag actgtgttca tgagcatgga
 3061 gaaccctggc ctgtggattc tgggatgcaa caactctgac ttcggcaaca gggccatgac
 3121 tgccttctc aaagtctctt cctgtgacaa gaacctggg gactactatg aggacagcta
 3181 tgaggacatc tctgctacc tgcctagcaa gaacaatgcc attgagccca ggagcttcag
 3241 ccagaatcca cctgtcctga aacgccacca gagggagatc accaggacca cctccagtc
 3301 tgaccaggag gagattgact atgatgacac ctttctgtg gagatgaaga aagaggactt
 3361 tgacatctat gacgaggacg agaaccagag cccaaggagc ttcagaaga agaccaggca
 3421 ctacttctatt gctctgtgg agcgcctgtg ggactatggc atgagctcca gcccctatgt
 3481 cctcaggaac agggcccagt ctgctctgt gccacagttc aagaaagtgg tcttccaaga
 3541 gttcactgat ggcagcttca cccagcccct gtacagaggg gactggaatg agcactggg
 3601 actcctgggc ccatacatca gggctgaggt ggaggacaad atcatggtga ccttccgcaa
 3661 ccaggcctcc agggcctaca gcttctacag ctccctcctc agctatgagg aggaccagag
 3721 gcagggggct gagccacgca agaactttgt gaaacccaat gaaaccaaga cctactctg
 3781 gaaagtccag caccacatgg cccccaccaa ggatgagttt gactgcaagg cctgggccta
 3841 cttctctgat gtggacctgg agaagatgt gcactctggc ctgattggcc cactcctggt
 3901 ctgccacacc aacaccctga acctgccc a tggaggcaa gtgaactgtc aggagttgc
 3961 cctcttcttc accatctttg atgaaaccaa gagctggtac ttcactgaga acatggagcg
 4021 caactgcagg gccccatgca acatcagat ggaggacct accttcaaag agaactaccg
 4081 cttccatgac atcaatggct acatcattga cacctgctt gggcttgtca tggcccagga
 4141 ccagaggatc aggtggtacc tgccttctat gggctccaat gagaacatc actccatcca
 4201 cttctctggg catgtcttca ctgtgccc aa gaaggaggag tacaagatgg cctgttaca
 4261 cctctaccct ggggtctttg agactgtgga gatgtgcc tccaaagctg gcctctgga
 4321 ggtggagtg ctcattgggg agcacctgca tgcctggcat agcacctgt tctgtgtca
 4381 cagcaacaag tgcagacc cctgggaat ggcctctggc cacatcagg acttccagat
 4441 cactgcctct ggccagtatg gccagtggc cccaagctg gccaggctcc actactctgg
 4501 atccatcaat gcctggagca ccaaggagcc attcagctgg atcaaagtgg acctgctggc
 4561 ccccatgatc atccatggca tcaagaccca gggggccagg cagaagttct ccagcctgta
 4621 catcagccag ttcacatca tgtacagcct ggtggcaag aatggcaga cctacagagg
 4681 caactcact ggaacactca tggctcttct tggcaatgtg gacagctctg gcataagca
 4741 caacatcttc aaccoccaa tcatcgccag atacatcagg ctgcacccca cccactacag
 4801 catcgcagc acctcagga tggagotgat gggctgtgac ctgaaactct gcagcatgcc
 4861 cctgggcatg gagagcaagg ccatctctga tgcctagatc actgctcca gctacttca
 4921 caacatggtt gccacctgga gcccaagcaa ggcaggctg cacctccagg gaaggagcaa
 4981 tgcctggagg cccagggtca acaacccaaa ggagtggctg caggtggact tccagaagac

도면7c

```

5041 catgaaggtc actgggggtga ccaccocaggg ggtcaagagc ctgctcacca goatgtatgt
5101 gaaggagbttc ctgabcagct ccagpcagga tggccaccag tggaccctct tcttccagaa
5161 tggcaagggtc aagggtgtcc agggcaacca ggacagcttc acccctgtgg tgaacagcct
5221 ggaccccccc ctccatgaeca gatacctgag gattcaccoc cagagctggg tccaccagat
5281 tgcocctgagg atggaggctc tgggatgtga ggcccaggac ctgtactgat gacgagcggc
5341 cgcctcttagt agcagtatog ataataaaaag atctttatit tcattagatc tgtgtgttgg
5401 ttttttgtgt gtttaattaag ctccggaagg aaccocctagt gatggagtty gccactccct
5461 ctctggcgcgc tcgctcgctc actgaggccg ggccgacaaa ggtcccccga cgcocgggct
5521 ttgcccgggc ggccctcagt agcagagcag cgcgcagaga gggagtggcc aagacgattt
5581 aatbgacaag ctggcgtaa tcatggtcat agctgtttcc tgbtgaaat tgttatccgc
5641 tcacaattcc acacaacata cgagccggaa gcataaagt taaagcctgg ggtgccta
5701 gagtgaagta actcacatta attgogttgc gctcactgoc cgtttccag tccggaaacc
5761 tgtogtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggf ttgctattg
5821 ggccctcttc cgttccctcg ctccactgact cgtcgcctc ggtcgttcgg ctgocggcag
5881 cgglatcagc tccctcaaag gcgtaatac ggttatccac agaatcaggg gataacgcag
5941 gaaagaacat gtgagcaaaa ggccagcaaa aggcacggaa ccgtaaaaag gcccggttg
6001 tggcgttttt ccataggctc cgcocccctg acgagcatca caaaaatcga cgtcaagtc
6061 agaggtggcg aaaccgcaca ggaatataaa gataccaggc gtttccccc tggagctccc
6121 tcttgocctc tctgttcccg accctgcgcg ttaccggata cctgtccgcc tttctccct
6181 cgggaagcgt ggccgtttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg
6241 ttcctccaa gctgggctac gtgcacgaac ccccggttca gcccgaccgc tgccttat
6301 ccgtaacla tcttcttag tccaaccccg taagacacga cttatcgcca ctggcagcag
6361 ccactggtaa caggattagc agagccaggt atgtaggcgg tgcacagag tctttgaagt
6421 ggtggcctaa ctaccgctac actagaagaa cagtatttgg tatctgcct ctgctgaagc
6481 cagttaccct cggaaaaaga gttggttagct cttgatccgg caaacaaacc accgctggta
6541 gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaaggga tctcaagaag
6601 atcctttgat cttttctacg ggtctgacg ctccagtga aaaaaactca cgttaaggga
6661 ttttggtcat gagattatca aaaaggatct tccactgat ccttttaaat taaaaatgaa
6721 gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgctaa
6781 tcagtgaagg acctatctca gcgatctgtc tatttctgtc atccatagtt gectgactcc
6841 ccgtcgtgta gataactaog atacgggagg gcttaccatc tggcccagc gctgcaatga
6901 taccgcgaga cccagctca ccgctccag atttatcagc aataaaccag ccagccggaa
6961 gggccgagcg cagaagtgtt cctgcaactt tatccgcctc catccagctc attaattgtt
7021 gccgggaagc tagagtaagt agttccgcag ttaatagttt gcgcaagtt gttgcoattg
7081 ctacaggeat cgtggtgtca cgtcgtcgt ttggtatggc ttcattcagc tccggtccc
7141 aacgatcaag gcgagttaca tgatcccca tgttggtgca aaaaagggtt agctccttcc
7201 gtcctccgat cgttgcaga agtaagttgg ccgagtggt atcactcatg gttatggcag
7261 cactgcataa ttctcttact gtcctgccat ccgtaagatg cttttctgtg actggtgagt
7321 actcaaccaa gtcattctga gaatagtgtg tgcggcgacc gagtgtctct tgcocggcgt
7381 caatacggga taataccgog ccacatagca gaactttaa agtgctcact attgaaaac
7441 gttcttcggg gcgaaaaact tcaaggatct taccgctgt gagatccagt tccatgtaac
7501 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag
7561 caaaaaacag aaggcaaat gccgcaaaa agggcaatag ggccacacgg aatgttgaa
7621 tactcatact ctctcctttt caatattatt gaagcattta tcagggttat tgtctcatga
7681 gcgatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc
7741 cccgaaaagt gccacctgac gttcaagaaa ccattattat catgacatta acctabaaaa
7801 ataggogtat cagcaggccc tttcgtc (서열 번호: 8)

```

도면8a

CS04Δ(760-1667) - CS04-SC1-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCGTGCGCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGATAC
TACCTGGGGGCTGTGGAGCTTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCCTGTGGATGCCAGG
TTCCCACCCAGAGTGCCTAAATCTTCCCATTCAACACCTCTGTGGTCTACAAGAAGACCTCTTTGTGGAG
TTCACTGACCCACCTGTCAACATGCGAAACCCAGGCCACCTGGATGGGACTCCTGGGACCCACCATTCAG
GCTGAGGTGTATGACACTGTGGTCACTACCCCTCAAGAACATGGCCCTCCACCCCTGTGAGCCTGCATGCTGTG
GGGGTCAGCTACTGGAAGGCCCTCTGAGGGGGCTGAGTATGATGACCCAGACCTCCAGAGGGAGAGGAGGAT
GACAAAGTGTCCCTGGGGGAGCCACACCTATGTGTGGCAGGTCTCAAGGAGAAATGGCCCATGGCCCTCT
GACCCACTCTGCCTGACCTACTCCTACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATT
GGGGCCCTGCTGGTGTGCAGGGAGGGCTCCCTGGCCAAAGAGAGAACCCAGACCTGCACAAGTTCAATCTCT
CTGTTTGTGTCTTTGATGAGGGCAAGAGCTGGCCTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGAT
GCTGCCCTCTGCCAGGGCCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTTGGCATGGGGCAACCCCTGAGGTGCACCTCCATT
TTCCCTGGAGGGCCACACCTTCCCTGGTCAGGAACACAGACAGGCCAGCCTGGAGATCAGCCCCATCACCTTC
CTCACTGCCAGACCTGCTGATGGACCTCGGACAGTCCCTGCTGTTCTGCCACATCAGCTCCACCAGCAT
GATGGCATGGAGGCCATGTCAAGGTGGACAGCTGCCCTGAGGAGCCACAGCTCAGGATGAAGAACAATGAG
GAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCCGCPTTGATGATGACAACAGC
CCATCCTTCATTGAGATCAAGTCTGTGGCCAAAGAAACACCCCAAGACCTGGGTGGACTACATTGCTGCTGAG
GAGGAGACTGGGACTATGCCCACTGGTCCCTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAAC
AATGGCCCAAGAGGATTGGACGCAAGTACAAGAAAGTCAGGTTCCATGGCCCTACACTGATGAAACCTTCAAG
ACCAGGGAGGCCATTGAGCATGAGTGTGGCATCCTGGGCCACTCCTGTATGGGAGGTGGGGGACACCCCTG
CTCATCATCTTCAAGAACCAGGCCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTGAGGCC
CTGTACAGCCCGAGGCTGCCAAAGGGGGTGAACACCTCAAGGACTTCCCCATTCTGCCCTGGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAATCTGACCCAGGTGCCCTCACCAGTACTAC
TCCAGCTTTGTGAACATGGAGAGGGGACCTGGCCCTCTGGCCCTGATTGGCCCACTGCTCATCTGCTACAAGGAG
TCTGTGGACCAGAGGGGAAACGAGATCATGTGTGACAAGAGGAATGTGATTTCTGTCTCTGTCTTTGATGAG
AACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCTGCCCAACCCCTGCTGGGGTGCAGCTGGAGGAC
CCTGAGTTCAGGCCAGGAAACATCATGCACTCCATCAATGGCTATGTGTTTGACAGCCTCCAGCTTTCTGTCT
TGCCTGCATGAGGTGGCCACTGCTGATATCTTTCTATGGGGCCAGACTGACTTCCTTTCTGTCTCTCTC
TCTGGCTACACCTTCAARCAAGATGGTGTATGAGGACACCTGACCCCTCTTCCCATCTCTGGGGAGACT
GTGTTTATGAGCATGAGAACCCCTGGCCCTGTGGATTCGGGATGCCACAACCTGACTTCCGCAACAGGGGC
ATGACTGCCCTGCTCAAAGTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACAGCTATGAGGAC
ATCTCTGCCACTCTGCTCAGCAAGAACAATGCCATTGAGCCAGGGAGATCACCAGGACCACCCCTCCAGTCT
GACCAGGAGGAGATTGACTATGATGACACCATTTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGAC
GAGGACGAGAACGAGGCCAAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGCGC
CTGTGGGACTATGGCATGAGCTCCAGCCCCATGTCTCAGGAAACAGGGCCAGTCTGGCTCTGTGCCACAG
TTCAAGAAAGTGGTCTTCCAAAGATTCAGTGTGGCAGCTTCAACCAGCCCTGTACAGAGGGGAGCTGAAT
GAGCACCTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGGACTTCCGCAAC
CAGGCCCTCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGGGCTGAG
CCAGCCAAAGACTTTGTGAAACCCAAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACACATGGCCCC

(계속)

도면8b

ACCAAGGATGAGTTTGAAGTCAAGGCGTGGGCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACCTCT
 GGCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCCTGCCCATGGAGGCAAGTACTGTG
 CAGGAGTTTGGCCCTCTCTTACCATOTTTGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGCGC
 AACTGCAGGCCCCATGCAACATTCAGATGGAGGACCCACCTTCAAAGAGAACTACCGCTTCCATGCCATC
 AATGGCTACATCATGGACACCTGCCCTGGGCTTGTCTATGGCCAGGACCAGAGGATCAGGTGGTACCTGGCTT
 TCTATGGGCTCCAATGAGAACATTCACCTCCATCCACTTCTCTGGGCATGTCTTCACTGTGGCAAGAAGGAG
 GAGTACAAGATGGCCCTGTACAACCTCTACCTGGGGTCTTTGAGACTGTGGAGATGTGCCCTCCAAAGCT
 GGCTACTGGAGGGTGGAGTGCCTCATTTGGGAGCACCTGCATGCTGGCATGAGCACCCCTGTCTCTGGCTTAC
 AGCAACAAGTGCCAGACCCCTGGGAATGGCTCTGGCCACATCAGGACTTCCAGATCACTGCCTCTGGC
 CAGTATGGCCAGTGGGCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGCACCAG
 GAGCCATTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAGGGGGCC
 AGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGGCAAGAAATGGCAG
 ACCTACAGAGGCAACTCCACTGGARACTCATGGCTCTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCAC
 AACATCTTCAACCCCCCAATCATCGCCAGATACATCAGGCTGCACCCACCCACTACAGCATCCGAGCACC
 CTCAGGATGGAGCTGATGGSTGTGACCTGAACTCCTGCAGCATGCCCTGGGCATGGAGAGCAAGGCCATT
 TCTGATGCCCAGATCACTGCCTCCAGCTACTTCCACCAACATGTTTGGCCCTGGAGCCCAAGCAGGCCAGG
 CTGCACCTCCAGGGAGGAGCAATGCCTGGAGGCCCCAGGTCACCAACCCAAAGGAGTGGCTGCAGGTGGAC
 TTCCAGAAGACCATGAAGSTCACTGGGGTGACCACCCAGGGGTCAGAGCCTGCTCACCAGCATGTATGTG
 AAGGAGTTCCTGATCAGCTCCAGCCAGGATGGCCACCAAGTGGACCCTCTTCTCCAGAAATGGCAAGGTCAA
 GTGTTCCAGGGCAACCAGGACAGCTTCAACCCCTGTGGTGAACAGCCTGGACCCCCCTCCTGACCAGATAC
 CTGAGGATTCACCCCCAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCTGGATGTGAGGCCACG
 GACCTGTACTGA (서열 번호: 9)

도면9

CS04A(760-1667) - CS04-SC1-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSPFPNTSVVYKK
 TLFVEFTDHLFNIAKPRPPWMGLLPTIQAEVYDTVVI TLKNMASHPVSLHAVGVSVWKASEGAEYD
 DQTSQREKEDDKVFPFGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVLDLVKDLNSGLIGALLVCREG
 SLAKEKTQLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCH
 RKSVYVHVIGMGTTPVHSLIFLEIGHTFLVRNHRQASLEISPIITFLTAQTLMLDLGQFLLFCHISSHQ
 HDGMEA YVKVDSCEPEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDNNSPFIQIRSVAKKHPKTVV
 HYIAAEEDWDYAPLV LAPDDRSYKSQYLNNQPQIRGRKYKVRFMAYTDET FKTREATQHESGILG
 PLYGVEVDLFLIIFKNQASRPYNIYPHGTIDVRPLYSRRLPKGVKHLKDFPILPGEIPKYKWTVTV
 EDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLLYCYKESVDQRGNQIMSDKRNVLFSVFDENRS
 WYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVPDSLQLSVCLHEVAYWYIISIGAQTDFLSV
 FFSGYTFKHKMVEDTLTLFPFSGETVFMSENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDY
 YEDSYEDISAYLLSKNNAIEPREITRITLQSDQEEIDYDDTISVEMKREDFDIYDEDENQSPRSFQK
 KTRHYFIAAVERLWDYGMSSSPHVLNRNAQSGSVPQFKKVVVFQEFDTGDSFTQPLYRGELNEHLGLLG
 PYIRAEVEDNIMVTRFNQASRPYSFYSSLISYEDQRQGAEPKRNFKVKNETKTYFWKVQHHMARTK
 DEFDCWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTEN
 MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRTIRWYLLSMGSNENIHSIHFSGH
 VFTVRKKEEYKMALYNLYPGVFETVEMLEPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMA
 SGHIRDFQITASGQYQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSS
 LYISQFIIMYSLDGKKWQTYRGNSTGILMVFFGNVDSGKHNIFNPPHARYIRLHPHYSIRSTL
 RMELMGCDLNSCSMPLGMSKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNAWRPQVNNPKEW
 LQVDFQKTMKVTGVTTCGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFEQGNQDSFTPVVNS
 LDPPLLTRYLRIHPQSVWHQIALRMEVLGCEAQDLY (서열 번호: 10)

도면10a

CS04Δ(772-1667) - CS04-SC2-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
 AGATACTACCTGGGGGCTGTGGAGCTTCTTGGGACTACATGCAGTCTGACCTGGGGGAGCTGCCCT
 GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAATCCTTCCCATTC AACACCTCTGTGGTCTACAAG
 AAGACCCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCTGGATG
 GGACTCCTGGGACCCACCATT CAGGCTGAGGTGTATGACACTGTGGTCATCACCCCAAGAACATG
 GCCTCCCACCCCTGTGAGCCTGCATGCTGTGGGGTCAGCTACTGGAAGGCCTCTGAGGGGCTGAG
 TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTCCCTGGGGGCAGCCACACC
 TATGTGTGGCAGGTCCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCCTGACCTACTCC
 TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
 AGGGAGGGCTCCCTGGCCAAAGAGAAGACCCAGACCCCTGCACAAGTTCATTCTCCTGTTTGCTGTC
 TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
 TCTGCCAGGGCTGGCCAAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCTGGACTC
 ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAGGTGCAC
 TCCATTTCTGGAGGGCCACACCTTCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
 CCCATCACCTTCCCTACTGCCAGACCCCTGCTGATGGACCTCGGACAGTTCTCTGTTCTGCGAC
 ATCAGCTCCCACAGCATGATGGCATGGAGGCCATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
 CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
 GATGTGGTCCGCTTTGATGATGACAACAGCCCATCCTTCATT CAGATCAGGTCTGTGGCCAAGAAA
 CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCACTGGTC
 CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCACAGAGGATTGGACGC
 AAGTACAAGAAAGTCAGGTT CATGGCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATT CAG
 CATGAGTCTGGCATCCTGGGCCACTCCTGTATGGGGAGGTGGGGGACACCCCTGCTCATCATCTTC
 AAGAACCAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGT CAGGCCCTGTAC
 AGCCGCAGGCTGCCAAAGGGGTTGAAACACCTCAAGGACTTCCCATCTGCTGGGGAGATCTTC
 AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCAGGTGCTCACCAGA
 TACTACTCCAGCTTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCACTGCTCATC
 TGCTACAAGGAGTCTGTGGACCAGAGGGGAAACCAGATCATGTCTGACAAGAGGAATGTGATTCTG
 TTCTCTGCTTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCTGCCCCAAC
 CCTGCTGGGGTGCAGCTGGAGGACCTGAGTTCAGGCCAGCAACATCATGCACTCCATCAATGGC
 TATGTGTTTGACAGCCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTTCTTCT
 ATTGGGGCCAGACTGACTTCCCTTCTGCTTCTTCTCTGGCTACACCTTCAAACACAAGATGGTG
 TATGAGGACACCCTGACCCTCTTCCCATCTCTGGGGAGACTGTGTT CATGAGCATGGAGAACCCT
 GGCTGTGGATTCTGGGATGCCACAACCTGACTTCGGCAACAGGGGCATGACTGCCCTGCTCAA
 GTCTCCTCTGTGACAAAGAACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTAC
 CTGCTCAGCAAGAACAATGCCATTGAGCCAGGAGCTTCAGCCAGAATTCAGACACCCCAAGCACC
 AGGGAGATCACAGGACCACCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGACACCATTTCT
 GTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAGAGCCCAAGGAGCTTC

(계속)

도면10b

CAGAAGAAGACCAGGCCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGGACTATGGCATGAGCTCC
AGCCCCCATGTCTCAGGAACAGGGCCAGTCTGGCTCTGTGCCACAGTTCAGAAAGTGGTCTTC
CAAGAGTTCAGTATGGCAGCTTCACCCAGCCCCGTACAGAGGGGAGCTGAATGAGCACCTGGGA
CTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTCCCGCAACCAGGCC
TCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGACAGGGGGCTGAG
CCACGCAAGAACTTTGTGAAACCCAAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCATG
GCCCCACCAAGGATGAGTTTGACTGCAAGGCCCTGGGCCACTTCTCTGATGTGGACCTGGAGAAG
GATGTGCACCTCTGGCCCTGATTGGCCACTCTCTGGTCTGCCACACCAACACCCCTGAACCTGCCCCAT
GGAAGGCAAGTACTGTGCAGGAGTTTGCCCTCTTCTTACCATCTTTGATGAAACCAAGAGCTGG
TACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAGGACCCCCACC
TTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCITGGGCTTGT
ATGGCCAGGACCAGAGGATCAGGTGGTACCTGCTTTCTATGGGCTCCAATGAGAACATTCAGCTCC
ATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCCCTGTACAAC
CTCTACCCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGGGTGGAG
TGCCCTCATTGGGGAGCACCTGCATGCTGGCATGAGCACCCCTGTTCCCTGGTCTACAGCAACAAGTGC
CAGACCCCCCTGGGAATGGCCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCCTGGCCAGTAT
GGCCAGTGGGCCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCCTGGAGCACCAG
GAGCGATTACAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAG
GGGGCCAGGCAGAAGTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGGC
AAGAAATGGCAGACCTACAGAGGCCAACTCCACTGGAACTACTCATGGTCTTCTTTGGCAATGTGGAC
AGCTCTGGCATCAAGCAACAACATCTCAACCCGCCAATCATCGCCAGATACATCAGGCTGCACCCC
ACCCACTACAGCATCCGCAGCACCCCTCAGGATGGAGCTGATGGGCTGTGACCTGAAGTCTCTGCAGC
ATGCCCTGGGCATGGAGAGCAAGGCCAATTTCTGATGCCAGATCACTGCCTCCAGCTACTTCAACC
AACATGTTTGCACCTGGAGCCCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAGGAGCAATGCCCTGG
AGGCCACAGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTCCAGAAGACCATGAAGGTCACT
GGGCTGACCCACCCAGGGGCTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTCTCTGATCAGC
TCCAGCCAGGATGGCCACCACTGGACCCCTCTTCTTCCAGAATGGCAAGGTCAAGGTGTTCCAGGGC
AACCAGGACAGCTTCAACCCCTGTGGTGAACAGCCTGGACCCCCCTCTGACCAGATACCTGAGG
ATTCAACCCCAAGAGCTGGGTCCACCAGATTGCCCTGAGGATGGAGGTCTGGGATGTGAGGCCAG
GACCTGTACTGA (서열 번호: 11)

도면11

CS04Δ(772-1667) - CS04-SC2-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFFPRVPKSFPFNTSVVYKKTLEVEF
TDHLFNIAPRPPMGLLGPITQAEVYDVTVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDK
VFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVLDLVKDLNSGLIGALIVCREGSLAKEKTQTLHKFILLFA
VFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTEVHSIFLEG
HTFLVRNHRQASLEISFITFLTAQTLMLDLGQFLLECHISSHQHDGMEAYVKVDSCEPEEPQLRMKNNEEAEDY
DDDLTDSEMDVVRFDNDNSPSPFIQIRSVAKKHPKTWVHYIAAEEEDMDYAPLVLAPDDRSYKSQYLNNGFQRI
GRKYKVRFMAYTDETFKTRERAIQHESGILGPLYGVEVGLTLLIIFKNQASRPYNTYPHGITDVRFLYSRRLP
KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLLICYESVDQRGNQ
IMSDKRNVILFVFDENRSWYLTENIQRFPLPNPAGVQLEDPEFQASNMHSINGYVFDLQLSVCLHEVAYWY
ILSIGAQTFDLSVFFSGYTFKHKMVEDTLTLFFPSGETVFMSENENPGLWILGCHNSDFRNRGMTALLKVVSSC
DKNTGDYEDSYEDI SAYLLSKNNAIEPRSPSQNSRHPSTREITRTTLQSDQEEIDYDDTISVEMKKEDEDIY
DEENQSPRSPQKTRHYFIAAVERLWDYGMSSSPHVLNRBAQSGSVPPFKKVVFQEFDTGDFTPQLYRGLN
EHLGGLGPIRAEVEDNIMVTFRNQASRPYSFYSSLI SYEEDQRQGAEPKRFVKNETKTYFWKVQHMAPT
KDFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAGRQVTVQEFALFPTIFDETKSWYFTENMERNC
RAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYK
MALYNLYPGVFETVEMLPSKAGIWRVECLIGBHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGG
WAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMI IHGIKTQGARQKFSLYISQFIIMYSLDGKKWQTYRGN
STGTLMVFFGNVDSSGKHNINPPIIARYIRLHPHTYSIRSTLRMELMGCDLNSCSMPLGMESKAI SDAQIT
ASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQTKMKVTGVTQGVKSLEPTSMYVKEFLISS
SQDGHOWTLFPQNGKVKVFEQGNQDSFTPVVNSLDEPLLTRYLRIHPQSWVHGIALRMEVLGCEAQLLY
(서열 번호: 12)

도면12a

CS01-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
 AGATACTACCTGGGGGCTGTGGAACCTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
 GTGGATGCCAGGTTCCCACCCAGAGTGCCTCAAGTCCCTCCCATTCAACACCTCTGTGGTCTACAAG
 AAGACACTCTTTGTGGAATCACTGACCACCTGTTC AACATTGCAAAACCCAGACCACCTGGATG
 GGACTCTGGGACCCACCATT CAGGCTGAGGTGTATGACACTGTGGTCATCACCCTCAAGAACATG
 GCATCCCACCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
 TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACC
 TATGTGTGGCAAGTCTCAAGGAGAATGGACCATGGCATCTGACCCACTCTGCCTGACATACTCC
 TACCTTTCTCATGTGGACTGGTCAAGGACCTCAACTCTGGACTGATGGGGACTGCTGGTGTGC
 AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGCTGTC
 TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
 TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCAC TGCTGGACTC
 ATTGGCTGCCACAGGAATCTGTCTACTGGCATGTGATGGCATGGGGACAACCCCTGAAGTGCAC
 TCCATTTTCTGGAGGGACACACCTTCTGGT CAGGAACCACAGACAAGCCTCTCTGGAGATCTCT
 CCCATCACCTTCTCACTGCACAGACACTGCTGATGGACCTGGACAGTTCCCTGCTGTCTGCCAC
 ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
 CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
 GATGTGGTCAGATTTGATGATGACAACCTCTCCATCCTTCATT CAGATCAGGTCTGTGGCAAAGAAA
 CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCCTGGTC
 CTGGCCCCGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGCCCCACAAAGAATTGGAAGA
 AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATT CAG
 CATGAGTCTGGCATTCTGGGACCCTCTGTATGGGGAGTGGGAGACACCCTGCTCATCATCTTC
 AAGAACCAGGCCCTCCAGGCCCTACAAGATCTACCCACATGGCATCACTGATGT CAGGCCCTGTAC
 AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCTGGAGAGATCTTC
 AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCAGSTGCCTCACCAGA
 TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCCTGCTCATC
 TGCTACARGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
 TTCTCTGTCTTTGATGAGAACAGATCATGTTACCTGACTGAGAACATTCAGAGATTCCTGCCAAC
 CCTGCTGGGGTGCAACTGGAAGACCCTGAGTTCAGGCAAGCAACATCATGCACTCCATCAATGCC
 TATGTGTTTGACTCTCTCCAGCTTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
 ATTTGGGGCACAACCTGACTTCTCTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
 TATGAGGACACCCCTGCACTCTTCCCATCTCTGGGGAACTGTGTT CATGAGCATGGAGAACCCT
 GGACTGTGGATTCTGGGATGCCACAACCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAA
 GTCTCCTCCTGTGACAAGAACACTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTCCCTAC
 CTGCTCAGCAAGAACAATGCCATTGAGCCAGAAGCTTCTCTCAGAAATCCACCTGTCTGGAAGAGA
 CACCAGAGAGAGATCACCAGGACAACCCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACC
 ATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCAAGA
 TCATTCCAGAAGAAGACAAGACACTACTTCATTGCTGCTGTGGAAAGACTGTGGGACTATGGCATG
 TCTTCCCTCCTCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG
 GTCTTCCAGGAGTTCATGATGGCTCATTCACCCAGCCCCTGTACAGAGGGGAAC TGAATGAGCAC

(계속)

도면12b

CTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTGAGAAAC
 CAGGCCCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGG
 GCTGAGCCAAGAAAGAACTTTGTGAAACCCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
 CACATGGCACCCACCAGGATGAGTTTACTGCAAGGCCCTGGGCATACTTCTCTGATGTGGACCTG
 GAGAAAGATGTGCACCTTGGCCTGATTGSCCCACTCCTGGTCTGCCACACCAACACCCTGAACCCT
 GCACATGGAAGGCAAGTGACTGTGCAGGAGTTTGCCCTCTTCTTCACCATCTTTGATGAAACCAAG
 TCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCCAGATGGAAGAC
 CCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGG
 CTTGTTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTTCATGGGATCCAATGAGAACATT
 CACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCCCTG
 TACAACCTCTACCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
 GTGGAATGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCCTGTTCTGGTCTACAGCAAC
 AAGTGCCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGGCTCTGGC
 CAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCA
 ACCAAGGAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTTCATGGCATCAAG
 ACACAGGGGGCAAGACAGAAATTCCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTG
 GATGGCAAGAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCCTCATGGTCTTCTTTGGCAAT
 GTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCTCCATCATGCCAGATACATCAGGCTG
 CACCCACCCCACTACTCAATCAGATCAACCCCTCAGGATGGAACATGATGGGATGTGACCTGAACCTC
 TGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTTCTGATGCCAGATCACTGCATCCTCTTAC
 TTCACCAACATGTTTGGCACCTGGTCACCATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGCAAT
 GCCTGGAGACCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAGACAATGAAA
 GTCACTGGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGATGTGAAGGAGTTCTG
 ATCTCTTCCCTCACAGGATGGCCACCAGTGGACACTCTTCTTCCAGAAATGGCAAAGTCAAGGTGTT
 CAGGGCAACCAGGACTCTTTCACACCTGTGGTGAACCTCACTGGACCCCCCTCTTGACAAGATAC
 CTGAGAATTCACCCCAAGTCTTGGGTCCACCAGATTGCCCTGAGAAATGGAAGTCTGGGATGTGAG
 GCACAAGACCTGTACTGA (서열 번호:13)

도면13a

CS08-FL-NA

ATGCAGATCGAACTGAGCACTTGCTTCTTCCTGTGTCTCCTGCGCTPTTGCTTCTCCGCCACAAGG
 AGATACATCTCGGTGCGGTGGAGCTCAGCTGGGACTACATGCAGAGCGACTTGGGTGAACTGCCT
 GTGGACGCCAGGTTTCCACCCCGCGTSCCCAAAGATTTCCCGTTCAACACCAGTGTGCTGTACAAG
 AAAACCCCTCTCGTGGAAATTCACCGACCACCTGTTCAACATCGCCAAACCGGCCCTCCCTGGATG
 GGGCTGCTCGGCCCGACGATCCAGGCTGAGGTCTATGACACGGTGGTGTATACCCCTCAAGAACATG
 GCTAGCCACCCGGTGTGAGCTGCACGCCGTGGGCGTGTCTATTGGAAAGCGTCCGAGGGTGC GGAG
 TACGATGACCAGACTTCACAGCGGGAGAAGGAAGACGACAAAGTGTCCCCGGGGGTCCCACACC
 TATGCTGGCAGGTCTGAAGGAGAAATGGTCCATGGCTCCGACCCATTGTGCCTCACCTACTCT
 TACCTAAGCCATGTGGATCTCGTCAAGGACCTGAACTCGGGGCTGATCGGGCGCCCTGCTCGTGTGC
 CGGGAGGGCTCACTGGCCAAGGAGAAGACCCAAACTCTGCACAAGTTCATCCTGCTGTTCCGCGTA
 TTCCAGCAGGGGAAGTCTGGCACFCGGAGACCAAGAACAGCCTGATGCAGGACCCGACGCGAGCC
 TCGGCCCGTGGTGGCCAAAGATGCACACCGTGAACGGCTACGTTAACAGGAGCCTACCCGGCCTG
 ATCGGCTGCCACCGCAAATCGGTCFACTGGCATGTGATCGGAATGGGCACAACGCCCGAGGTCCAC
 AGTATCTTCCCTCGAGGGCCACACTTTCCCTGGTCCGGAATCACCGCCAGGCCAGCCTGGAGATCAGC
 CCCATAACCTTTCTGACGGCGCAGACCTTACTCATGGATCTCGGCCAGTTCCTCCTGTTCTGCCAC
 ATTTCTGCCACCAGCACGATGGGATGGAAGCATATGTGAAAGTGGACTCCTGCCCGAGGAACCC
 CAGCTTAGGATGAAGAACAATGAGGAGGCCGAGGACTACGACGATGACCTTACCGATTGAGAAATG
 GACGTAGTACGCTTTGACGACGACAACCTCTCCATCCTTCATACAGATTGCTCCGTCCGCAAGAAG
 CACCCTAAGACTTGGGTGCACCTACATCGCGGCCGAGGAGGAGACTGGGATTATGCTCCCCTGGTG
 CTGGCCCCCGACGACCGCAGCTACAAGAGCCAGTACCTGAATAACGGGGCCCAGCGCATCGGCCGG
 AAGTACAAGAAAGTGGCGTTTATGGCTTACACGGACGAGACCTTCARGACCCGGGAGGCTATCCAG
 CATGAGAGCGGCATCTTGGGGCCCCCTCCTGTACGGGGAAGTTGGAGACACACTGCTGATCATCTTC
 AAGAACCAGGCGAGCAGGCCCTACAACATCTACCCCCACGGCATTACCGATGTCCGGCCGTTGTAC
 AGCCGACGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCGATCCTGCCGGCGGAGATCTTC
 AAGTACAAGTGGACTGTGACCGTGGAGGATGGGCCGACCAAGAGCGATCCGCGCTGCCTGACCCGT
 TACTACTCCAGCTTTGTCAATATGGAGCGCGACCTCGCTAGCGGCTTGATTGGCCCTCTGCTGATC
 TGCTACAAGGAGTCCGTGGACCAGAGGGGAATCAGATCATGAGTGACAAGAGGAACGTGATCCTG
 TTCTCCGTGTTGACGAAAACCGCAGCTGGTATCTCACCAGAAATATCCAGCGCTTCTGCCCAAC
 CCGGCCGGTGTGCAGCTGGAGGACCCCGAGTTTCAGGCCAGCAACATCATGCATTCTATCAACGGA
 TATGTGTTTATTCCCTGCAGCTCTCAGTGTGTCTGCACGAGGTCCGCTACTGGTATATCCTCAGC
 ATTGGGGCACAGACCGACTTCTGAGCGTGTCTTCTCCGGGTATACCTTCAAGCACAAGATGGTG
 TACGAGGATACCCTGACCCTGTTCCCTTTAGCGGCGAAACCGTGTATATGCTATGGAGAACCCC
 GGGCTCTGGATCCTTGGCTGCCATAACTCCGACTTCCGCAACCGCGGAATGACCCGCTCCTGAAA
 GTGTGAGTTGTGACAAGAACACCGCGACTATTACGAGGACAGTTACGAGGACATCTCTGCGTAC
 CTCCTTAGCAAGAATAACGCCATCGAGCCAAGATCCTTCAGCCAGAACCCCCAGTGTGAGAGG
 CATCAGCGGGAGATCACCCGACGACCTGCAGTCCGATCAGGAGGAGATTGATTACGACGACACG
 ATCAGTGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAAGATGAAAACAGTCCCCTCGG
 TCCTTCCAAAAGAAGACCCGGCCTACTTTCATCGCGCTGTGGAACGCCTGTGGGACTATGGAATG

(계속)

도면13b

TCTTCTAGCCCTCACGTTTTGAGGAACCGCGCCAGTCGGGCAGCSTGCCCCAGTTCAGAAAAGTG
 GTGTTCCAGGAGTTCACCGACGGCTCCTTCACCCAGCCACTTTACCGGGGCGAGCTCAATGAACAT
 CTGGGCCCTGCTGGGACCCCTACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACATTCCGGAA
 CAGGCCAGCAGACCATACAGTTTTCTACAGTTCACTCATCTCCTACGAGGAGACCAGCGCCAGGGG
 GCTGAACCCCGTAAGAACTTCGTGAAGCCAAACGAAACAAAGACCTACTTCTGGAAAGGTCCAGCAC
 CACATGGCACCTACCAGGACGAGTTCGATTGCAAGGCCTGGGGCTACTTCTCCGACGTGGACCTG
 GAGAAAAGATGTGCACAGCGGCCTGATTTGGCCCTCTGCTGGTGTGTGCACACGAACACACTCAACCCT
 GCACACGGGGCGGCGAGGTCACTGTGCGAGGAATTCGGCCCTGTTCTTTACCATCTTTGATGAGACGAAG
 TCCTGSTATTTCAACGAAAACATGGAGAGGAACGGCGCGCACCCCTGCAACATCCAGATGGAAGAT
 CCGACATTCAGGAGAACTACCGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCTGGC
 CTCGTGATGGCCCAAGACCAGCGTATCCGCTGGTATCTGCTGTGATGGGCTCCACGAGAACATC
 CATAGTATCCACTTCAGCGGGCATGTCTTCAGGTTGAGGAAAAAGGAGGAGTACAAGATGGCACGTG
 TACAACCTCTATCCCGGCGTGTTCGAGACCGTGGAGATGCTGCCCTCCAAGGCCGGCATCTGGAGA
 GTGGAATGCCTGATCGGGGAGCACCTCCACGCTGGGATGTCCACGCTGTTCCCTGTTTACAGCAAT
 AAGTGGCAGACCCCTCTGGGCATGGCGAGCGGCCACATCCGCGACTTCCAGATTCAGCCAGCGGC
 CAGTACCGTCACTGGGCTCCAAAGCTGGCCCGTCTGCACTACTCCGGATCCATCAACGCCTGGTCC
 ACCAAGGAACCGTTCCTGGATCAAAGTAGACCTGCTAGCCCCCATGATCATTCACGGCATCAAG
 ACACAAGGCGCCCGACAGAAGTTCGAGGCCTCTATATCTCCAGTTCATCATCATGTATAGCCTG
 GACGGAAAGAAGTGGCAGACTTACCGGGAAACTCGACAGGGACCCCTGATGSTATTTCTTCGGTAAC
 GTGGACAGCTCCGGAATCAAGCACAAACATCTTCAACCCACCCATTATCGCCCGCTACATCCGGCTG
 CACCCCACTCACTATAGCATTAGGTCCACCCTGCGAATGGAGCTCATGGGCTGTGACCTGAACAGC
 TGTAGCATGCCCCCTCGGCATGGAGTCTAAGGCGATCTCCGACGCACAGATAACGGCATCATCTTAC
 TTTACCAAGATGTTGCTACCTGGTCCCCCTCCAAGGCCCGACTCCACCTGCAAGGGAGATCCAAC
 GCCTGGCGGCACAGGTCAACAATCCCAAGGAGTGGCTGCAAGTGGACTTTCAGAAAACATGAAA
 GTACCCGGAGTGACCACACAGGGAGTGAAGTCTCTGCTGACCAGCATGTACGTGAAGGAGTTCCTC
 ATCTCCAGTTCGCAGGATGGCCACCAGTGGACGTTGTTCTTCCAAAACGGTAAAGTCAAAGTCTTC
 CAAGGGAACCAGGACAGCTTTACACCCGTCGTGAACCTCCCTGGACCCCCCGCTTCTCACTAGATAC
 CTCCGCATCCACCCTCAGAGCTGGGTGCACCAGATTGCCCTGCGCATGGAGGTTCTGGGSTGTGAA
 GCCCAGGACCTGTAATAA (서열 번호:14)

도면14a

CS10-FL-NA

ATGCAGATTGAGCTCTCCACCTGCTTCTTTCTCTGCCTTCTTCGCTTCTGCTTTTCTGCCACACGC
 AGGTACTATTTGGGAGCAGTGGAACTGAGCTGGGATTACATGCAGAGTGACCTTGGTGAACCTTCT
 GTGGACGCTCGTTTTCCACCTAGAGTTCCCAAGTCCTTCCCCTTCAACACCTCAGTGGTCTACAAG
 AAAACGCTGTTTTGTGGAGTTCACTGACCACCTCTTCAACATTGCCAAACCAAGACCCCTTGGATG
 GGATTGCTGGGACCCACAATAACAAGCAGAAGTCTACGACACGGTGGTGATTACCCTGAAGAACATG
 GCGTCCACACCCTGTTTTCACTTTCACGCTGTGGGGTCAGTTATTGGAAAGCCTCAGAGGGTGCGGAA
 TACGATGATCAAACCAGCCAGAGGGAGAAGGAAGATGACAAGGTCTTTCTGGGGGTAGCCATAACC
 TATGTTTTGGCAGGTGCTGAAAGAGAAATGGGCCCTATGGCCTCTGATCCCTTGTGCCCTCACATACTCT
 TACCTGAGTCACTGCGACCTGGTGAAGACCTGAATAGCGGTCTGATTGGTGCCTGCTTGTGTTGT
 AGAGAGGGGAGTTTTGGCCAAGGAGAAAACCTCAGACTCTCCACAAGTTTATCCTCCTGTTTGTGTG
 TTCCGACGAGGGCAAGTCTTGGCACCTGAAACAAAGAACTCCCTGATGCAGGACAGAGATGCTGCA
 TCTGCAAGGGCTTGGCCAAAAATGCACACAGTGAACGGCTATGTGAATCGATCACTGCCAGGACTG
 ATAGGCTGTCACTCGCAAGTCACTGTATTGGCAGTTATCGGGATGGGAAACAACCCAGAAGTGCAC
 AGCATCTTCTTGGGGCCACACTTCTGCTGGTTCGGAATCATAGACAGGCCAGCCTTGAGATCAGC
 CCAATCACCTTTCTGACTGCCCAAACTTGGCTGATGGATCTGGGACAGTTCTCCTGTTTTGTGTCAC
 ATCTCCTCCCACCAACATGACGGGATGGAGGCTTATGTGAAGTTCGATAGCTGTCCGGAGGAACCA
 CAACTGAGGATGAAGAACAACGAAGAGGCAGAGGACTATGACGACGATCTGACTGACAGTGAATG
 GACGTGGTTCCGGTTCGACGATGACAAATCTCCTTCATTTATCCAGATCCGTTCCTGGCCAAGAAG
 CACCCCAAGACTTGGGTTTCACTTACATGGCTGCTGAGGAGGAGGATTGGGACTACGCCCCCTTGGTG
 TTGGCCCCAGACGATCGCTCATACAAGAGCCAGTACCTTAACAATGGTCCACAAGGATCGGCCGG
 AAGTACAAGAAGGTTAGATTTATGGCTTATACCGACGAGACTTTTAAAACCTAGGGAAGCAATTCAG
 CATGAAGTGGCATTCTTGGACCCCTGCTGTATGGCGAGGTTGGCGACACCCTGCTGATTATCTTT
 AAGAACCAGGCAAGCCGGCCCTACAACATCTACCCGCACGGCATAACCGATGTACGACCCCTGTAC
 AGTCGCAGACTTCTTAAAGGGGTGAAACACCTGAAGGACTTCCCAATTCTGCCCGGGGAGATCTTC
 AAGTATAAATGGACCGTGACGGTTGAGGATGGTCCCACAAAGTCCGATCCGAGATGCCTTACCCGA
 TATTATTCCAGCTTCGTGAACATGGAAAGGGACCTGGCCAGCGGGCTGATTGGCCCACTGCTGATT
 TGTTACAAGGAGTCTGTGATCAAAGAGGAAACCAATAATGAGCGACAAACGTAAACGTATCCTG
 TTCAGCGTCTTTGATGAGAATAGAAGCTGGTACCTCACAGAAAATATTCAGCGGTTTCTGCTAAC
 CCCGAGCGCTCCAGCTGGAAGATCCCGAGTTCBAGCCCTCAAACATCATGCATAGCATCAACGGA
 TACGTATTCGATAGCCTGCAGCTGTCCTGCTCCATGAAGTGCCATATGGTACATCCTGAGT
 ATCGGGGCGCAGACCGACTTCCCTGAGCGTGTCTTTTCTGATACACGTTCAAACACAAAATGGTC
 TATGAAGATACCCCTGACTCTGTTCCATTCTCAGGAGAGACAGTCTTTATGATATGGAAAATCCT
 GGACTGTGGATCCTGGGCTGTCACAATTCTGATTTTCGGAACAGAGGCATGACAGCCCTGCTTAAA
 GTGAGCTCATGCGACAAGAACACCGGTGATTACTACGAAGATAGCTATGAGGACATCAGTGCCTAT
 TTGCTCTCCAAGAACAACGCTATCGAGCCACGGTCTTTCAGTCAAGATCCTCCCGTCTGAAGCGG
 CATCAGCGGAAATAACAGGCACAACCCCTTCAGTCAAGCAAGAGGAAATCGACTACGATGATACT
 ATCTCTGTGGAGATGAAGAAGGAGGATTTCCGACATTTACGACGAGGACGAGAATCAGTCCCAAGG
 AGCTTTCAGAAGAAAACAAGACACTATTTTCATTGCCCGCGTGGAGCGACTGTGGGACTACGGCATG

(계속)

도면14b

TCTAGCTCTCCGCATGTACTTAGAANTAGGGCAGAAAAGCGGATCCGTGCCTCAGTTAAGAAAGTT
 GTCTTTCAGGAGTTTACAGATGGCTCCTTCACCCAGCCCTTGATCGCGGGGAACTCAATGAACAC
 CTGGGCCTCCTGGGTCCCTTATATTAGGGCCGAAGTCGAGGACAATATCATGGTGACCTTTAGGAAC
 CAGGCATCTAGACCTTACTCTTTCTACTCCTCCCTGATATCCTATGAGGAGGACCAGCGGCAAGGC
 GGTGAGCCTCGGAAGAACTTTGTGAAGCCAAATGAAACAAAACATACTTTTGGAAAGTTTCAGCAC
 CACATGGCTCCCACGAAGGACGAATTTGACTGTAAAGCCTGGGCCTACTTCTCAGATGTAGATCTC
 GAGAAAGACGTGCACTCAGGGCTCATTTGGTCCCCTCCTGGTCTGTATATACTAATAACCCCTCAATCCA
 GCACACGGACGTGAGGTAACCGTCCAGGAATTTGCCCTGTTCTTTACCATTTTCGATGAGACTAAA
 TCCTGGTACTTTTACCAGAAAACATGGAGAGGAATTCAGAGCCCCATGCAACATCCAGATGGAGGAC
 CCTACCTTCAAAGAGAACTATCGCTTCCATGCCATTAACGGTTACATATATGGATACTCTCCCAGGA
 CTTGTGATGGCAGGATCAGCGGATAAGATGGTATCTGTTGAGCATGGGCTCCAACGAGAATATT
 CACAGCATCCATTTCTCCGGTCACTGTTTACAGTGAGAAAGAAAGAGTACAAGATGGCTCTG
 TATAATCTCTATCCAGCGGTATTCGAAACGGTGGAGATGTTGCCTAGCAAGGCCCGGCATTTGGCGA
 GTAGAATGCCCTTATCGGGGAACATCTGCATGCCGGAAATGAGCACGCTCTTCCCTGGTGTATAGTAAC
 AAGTGCCAGACTCCGCTGGGCATGGCATCTGGCCATATACGGGACTTTCAGATTACGGCTAGCGGG
 CAGTATGGGCAGTGGGCACCCAAACTTGGCGGACTGCACATTTACGGCTCTATCAATGCATGGTCC
 ACCAAGGAACCCCTTCTCTTGGATTAAAGTGGACCTTTTGGCGCCCATGATAATCCATGGGATCAAA
 ACCCAGGGCGCTCGTCAGAAATTTCTCATCACTCTACATCTCTCAGTTCATAATAATGTATTCACTG
 GATGGGAAGAAATGGCAGACTTACAGAGGAAACAGCACCGGGACGCTGATGGTGTCTTTGGCAAC
 GTGGACAGCAGCGGCATGAAACACAACATCTTCAATCCTCCCATTATTGCCCGTTATATTAGACTG
 CATCCCACTCACTACTCTATACGCAGCACACTTAGGATGGAGCTCATGGGATGCGACCTGAACAGT
 TGTAGTATGCCCTTGGGGATGGAGTCAAAAGCTATAAGCGACGCACAAATTACAGCTAGCTCTTAC
 TTTACGAATATGTTCCGCACGTGGAGCCCAAGCAAAGCCCGGCTGCATTTGCAGGGTCGGAGTAAT
 GCTTGGCGCCACAGGTGAATAACCTAAGGAATGGTTGCAAGTAGATTTCCAGAAAACATGAAG
 GTAACCGCGTCACTACACAGGGAGTCAAGTCCCTCTTGACCTCTATGTACGTCAGGAGTTCCTG
 ATTAGCAGCAGTCAGGATGGGCACCAATGGACACTGTTCTTCCAGAAATGGGAAAGTTAAAGTATTT
 CAGGGTAACCAGGACTCCTTTACACCTGTGGTGAATAGCCTCGACCCACCCCTGCTGACACGATAC
 CTCCGCATCCACCCCTCAGTCTTGGGTGCATCAAAATGCCCTGCCAATGGAGGTGTTGGGATGCCAA
 GCTCAGGACCTCTACTGA (서열 번호: 15)

도면15a

CS11-FL-NA

ATGCAGATCGAACTCTCTACTTGGCTTCTTCTCTGTGCCTTCTGAGSTTCTGCTTCTCTGCCACTCGC
 CGATATTACCTCGGGGCCGTGGAGTTGAGTTGGGACTACATGCAATCAGATCTGGGCGAACTCCCT
 GTGGATGCCCGATTCCCACCGCGCGTGGCCAAAGTCTTCCCATTTAAAFACCTTCTGTGGTGTACAAG
 AAGACATTGTTTGTGGAGTTTACCGATCACCTGTTCACATCGCCAAACCGGGGCCCCCATGGATG
 GGTCTGCTTGGGCCCACCATTCAAGCGGAGGTCTATGATACAGTGGTGATAACGCTTAAGAACATG
 GCGAGCCACCAGTGTCTCTGCATGCCGTTGGTGTATCATATTGGAAGGCCAGCGAAGGAGCGGAG
 TACGATGACCAGACCTCTCAGAGAGAGAAGGAAGACGATAAGGTTTTTCTTGGCGGAAGTCATACA
 TATGTATGGCAGGTCTGAAAGAGAATGGGCCGATGGCTTCTGACCCCTTTGTCTTACCTATAGT
 TATCTGAGCCACGTGGACCTGCTCAAGGACCTCAACAGTGGTCTGATTGGGGCTCTGCTTGTTTGT
 AGAGAGGGTAGCTTGGCTAAGGAGAAAACCCAAACATCCATAAGTTCATTTTGTCTGTTGCGGGT
 TTCGACGAGGAAAAGAGTTGGCACAGCGAAACAAAGAATTCACTGATGCAAGACAGGGACGCCGCT
 TCCGCAAGGGCTTGGCCTAAGATGCATACGGTGAATGGGTATGTGAACCGGAGCCTCCCGGGGCTG
 ATCGGCTGCCATCGCAAGTCTGTTTACTGGCACGTCAATGGAAATGGGGACAACGCCAGAGGTACAT
 AGTATATTTCTTGAAGGCCACACGTTCCCTCGTACGGAAACCACCGACAGGCTTCCCTGGAGATAAGC
 CGCATTACCTTCTGACCGCTCAGACTCTGCTGATGGACCTTGGCCAGTTTCTCTCTGTTCTGCCAT
 ATTAGCAGCCACCAGCACGCGGTATGGAAGCATACTGTGAAAGTCGATAGCTGTCTCTGAGGAGCCT
 CAGCTCAGAATGAAGAACAACGAGGAGGCCGAAGACTATGACGATGACCTTACAGATTCGGAGATG
 GACGTGGTGGCTTTGACGACGATAACAGTCCTAGTTTCATTCAAATCAGATCGGTAGCCAAAAAG
 CATCCAAAGACATGGGTGCATTACATTGCAGCCGAAGAGGAGGATTGGGATATGCGCCCTTGT
 CTGGCTCCAGATGACAGGAGCTATAAGTCCCACTACTTGAACAACGGGCCACAGCGAATCGGTAGA
 AAATATAAGGAGGTAAGATTCATGGCCTACACTGACGAAACATTTAAAACCCAGGGAAGCTATCCAA
 CACGAATCTGGAATTTCTCGGCCCTCTGCTCTACGGTGAAGTGGGGGACACCTTGTCTGATCATTTC
 AAAATCAGGCATCCAGGCCTTACAACATATACCCCATGGCATCACCGATGTCCGCCCCTGTAT
 TCCAGAAGACTCCCCAAGGGAGTGAAACATCTGAAAGATTTTCCATCCTGCCGGGCGAGATCTTT
 AAATACAAATGGACTGTGACTGTAGAGGACGGGCCCTACAAAATCAGACCCACGGTGCCTGACAAGG
 TATTACAGTAGCTTCGTCAACATGGAACGCGACCTCGCCAGCGGACTCATTTGGCCACTGTTGATC
 TGTTACAAAGAGTCAAGTGGATCAGAGGGGAAATCAGATCATGAGCGATAAGAGAAACGTTATCCTG
 TTTAGTGTCTTCGACGASAAACCGGTCTTGGTACCTTACTGAGAACATCCAGAGGTTCTTGGCGAAT
 CCGGCTGGCGTTCACTCGAGGACCCAGAGTTCAGGCCAGTAATATAATGCACTCAATCAACGGT
 TATGTGTTCCGATAGCCTGCAGCTGAGCGTCTGCCTCCACGAGGTAGCCTATTGGTACATATTGTCC
 ATCGGGGCTCAGACCGATTTCTGTCCGTGTTCTTTAGCGGGTATACCTTTAAACATAAAATGGTC
 TATGAAGACACCCTGACCCGTTCCTCCGGTGAAGTGTGTTTCAATGTCATGTCATGGGAGAACCA
 GGGCTGTGGATCCTGGGTTGTCACAATAGTGACTTTAGGAATCGGGGAATGACGGCACTGCTGAAG
 GTGAGTTCTTCGGATAAAAAATACAGGAGATTACTATGAGGATAGTTACGAGGATATCAGTGCCTAT
 CTGCTTTCAAAAACAACCGCAATTGAGCCCGGTCTTTCTCACAAAACCCCGGTGCTGAAGCGC
 CACCAGCGGAAAATTACCCGGACAACCTTGCAGTCCGACCAGGAGGAAATCGATTATGACGATACT
 ATCAGTGTAGAAATGAAAAAGGAGGATTTTGATATTTACGACGAAGACGAGAACCAGTCTCCGCGA

(계속)

도면15b

AGTTTTCAGAAGAAAACGCGACACTACTTTATAGCTGCCGTGGAACGACTCTGGGATTATGGCATG
 TCCTCCAGCCCTCATGTCTTAGGAATCGAGCGCAGAGTGGCTCTGTGCCTCAGTTCAAAAAGGTT
 GTGTTCCAGGAATTCACCGACGGCTCATTTACCCAGCCGCTGTACAGAGGCSAATCAACGAACAC
 CTTGGGCTGCTTGGGCCATATATTCGAGCAGAGGTGGAAGATAATATCATGGTAACCTTTAGAAAC
 CAGGCGTCAAGACCCTATTCCCTTCTACAGTTCTCTGATCAGCTACGAGGAGGACCAAAGACAGGGA
 GCTGAACCCAGGAAGAATTTGTGAAACCTAATGAGACCAAGACCTACTTCTGGAAGGTCCAGCAC
 CATATGGCCCAACTAAAGATGAATTCGATTGCAAGGCCCTGGGCTTATTTACGCGACGTGGATCTC
 GAAAAGGATGTGCACAGCGGGTTGATCGGACCGCTTTTGGTGTGCCACACAAATACCCTCAATCCT
 GCCCACGGGCGGCAGGTCACAGTTCAGAGTTTGCACCTCTTCTTTACAATATTTGACGAGACAAAG
 TCATGGTATTTTACAGAGAATATGGAGAGAAATTTGTCGCGCACCTTGCAACATTCAGATGGAGGAC
 CCCACATTTAAGGAGAATTACAGATTTTCATGCTATCAATGGGTACATTTATGGATACTCTGCCTGGT
 CTGGTCATGGCCCAGGATCAGCGCATAGGTGGTACTTGCTGAGCATGGGATCTAATGAGAATATA
 CACAGCATTCACTTCAGTGGCCACGTTTTTACTGTTAGAAAGAAGGAGGAGTACAAAATGGCGCTC
 TACAACCTTTACCCGGGTGTGTTTGAGACAGTGGAGATGCTGCCAAGCAAGGCAGGCATCTGGAGG
 GTTGAGTGTCTTATTGGGGAGCATCTGCATGCTGGAATGTCCACCCTCTTTCTTGTGTACAGCAAT
 AAGTGCCAGACACCCTTGGCATGGCCAGGGCCACATTAGGGACTTTCAGATAACTGCCAGTGGGA
 CAGTACGGCCAGTGGGCTCCCAAGCTTGCAAGACTCCACTACTCCGGAAGCATAAACGCATGGAGC
 ACCAAGGAACCCCTTCTCTTGGATTAAGGTGGACCTGCTGGCSCCAATGATCATTCACGGCATAAAA
 ACCCAAGGGGCACGACAGAAATTTTCATCTTTGTATATTAGTCAAGTTTATCATCATGTACAGCTTG
 GATGGAAGAAGTGGCAGACGTACAGGGCAATTCACAGGAACACTTATGGTGTTTTTTGGGAAT
 GTCGATTCCAGCGGGATCAAACATAACATCTTCAATCCTCCTATTATCGCCCGATATATCCGCCTG
 CACCCITACGCATTACTCCATCAGGTCCACATTTAGAGATGGAACGTATGGGGTGGCACCCTGAATAGT
 TGTAGTATGCCACTGGGCATGGAGTCTAAGCCATCAGCGATGCACAGATCACTGCCAGCTCTTAC
 TTCACCAACATGTTTGCAACTTGGTCCCCCTCTAAAGCTCGCCTGCATCTGCAGGGACGCTCAAAT
 GCATGGCGACCACAGGTGAACAATCCAAAAGAGTGGCTCCAGGTCGACTTTCAGAAAGACAATGAAG
 GTAACAGGAGTGACAACCCAGGGTGTAAAAAGCCTCCTTACGAGTATGTACGTTAAGGAGTTTCTG
 ATTTCTAGCTCCAGGACGGACACCAGTGGACTCTGTCTTCCAGAACGGCAAAGTGAAGGTATTT
 CAGGGAACCAGGATTCCTTTACCCCGGTAGTGAATAGCCTGGATCCACCCTTGTGACCCGCTAT
 CTGAGAATTCATCCACAATCCTGGGTGCATCAGATTGCCCTCCGGATGGAAGTGTCTGGCTGTGAA
 GCTCAGGATCTGTATTAG (서열 번호: 16)

도면16a

CS40-FL-NA

ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGGGATTCTGCTTTAGTGCCACCAGA
 AGATACTACCTGGGTGCAGTGGAACTGTCATGGGACTATATGCAAAGTGATCTCGGTGAGCTGCCCT
 GTGGACGCAAGATTTCTCCTAGAGTGCCAAAATCTTTCCATTCAACACCTCAGTCGTGTACAAA
 AAGACTCTGTTTGTAGAAATCACGGATCACCTTTTCAACATCGCTAAGCCAAGGCCACCCTGGATG
 GGTCTGCTAGGTCTTACCATCCAGSCTGAGGTTTATGATACAGTGGTCATTACACTTAAGAACATG
 GCTTCCCATCCTGTGCTGCTTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTGAA
 TATGATGATCAGACCAGTCAAAGGGAGAAAGAAGATGATAAAGTCTTCCCTGGTGAAGCCATACA
 TATGCTCTGGCAGSTCCTGAAAGAGAAATGGTCCAATGGCCTCTGACCCACTGTGCCCTTACCTACTCA
 TATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGT
 AGAGAAGGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAAAATTTATACTACTTTTTGCTGTA
 TTTGATGAAGGGAAAAGTTGGCAGTCAAGAAACAAGAACTCCTTGTATGCAGGATAGGGATGCTGCA
 TCTGCTCGGGCCTGGCCTAAAATGCACAGAGTCAATGTTTATGTAACAGGTCTCTGCCAGGTCTG
 ATTGGATGCCACAGSAAATCAGTCTATTTGGCATGTGATTGGAATGGGCACCACCTCTGAAGTGCAC
 TCAATATTCTCGAAGGTCAACACATTTCTTGTGAGGAACCATCGCCAGGCGTCTTGGAAAATCTCG
 CCAATAACTTTCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTTGTCTAT
 ATCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCC
 CAACTACGAATGAAAAATAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATG
 GATGTGCTCAGGTTTGTATGATGACAACCTCTCCTTCTTTATCCAAATTCGCTCAGTTGCCAAGAAG
 CATCTAAAACCTTGGGTACATTACATTGCTGCTGAAGAGGAGGACTGGGACTATGCTCCCTTAGTC
 CTGCCCCCGATGACAGAAGTTATAAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGTAGG
 AAGTACAAAAAAGTCCGATTTATGGCATAACACAGATGAAACCTTTAAGACTCGTGAAGCTATTCAG
 CATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTGATTATATTT
 AAGAATCAAGCAAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCCGTCCTTTGTAT
 TCAAGGAGATTACCAAAAAGGTGTA AAAACATTTGAAGGATTTTCCAAATCTGCCAGGAGAAAATATTC
 AAATATAAATGGACAGTACTGTAGAAAGATGGGCCAACTAAATCAGATCCTCGGTGCCCTGACCCGC
 TATTACTCTAGTTTTGTTAATATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCCTCATC
 TGCTACAAAGAATCTGTAGATCAAAGAGGAAAACCAGATAATGTCAGACAAGAGGAATGTCATCCTG
 TTTTCTGTATTTGATGAGAACCGAAGCTGGTACCCTCACAGAGAATATACAACGCTTTCTCCCCAAT
 CCAGCTGGAGTGCAGCTTGGAGATCCAGAGTTCCAAGCCTCCAACATPCATGCACAGCATCAATGGC
 TATGTTTTTGTATAGTTTGCAGTTGTGAGTTTGTGTTGTCATGAGGTGGCATACTGGTACATTCTAAGC
 ATTGGAGCACAGACTGACTTCTTCTGTCTTCTTCTCTGGATATACCTTCAAAGCACAAAATGGTC
 TATGAAGACACACTCACCTATTCCCATTTCTCAGGAGAAACTGTCTTCATGTCCGATGGAAAACCCA
 GGTCTATGGATTCTGGGGTGCCACAACCTCAGACTTTCGGAACAGAGGCATGACCCGCTTACTGAAG
 GTTCTAGTTGTGACAAGAACACTGGTGATTATTACGAGGACAGTTATGAAGATATTTGAGCATAAC
 TTGCTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCAGAAATCCACCAGTCTTGAACGC
 CATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATGACTATGATGATAC
 ATATCAGTTGAAATGAAGAAGGAAGATTTTGCATTTATGATGAGGATGAAAATCAGAGCCCCCGC
 AGCTTTCAAAAGAAAACACGACACTATTTTATGCTGCAGTGGAGAGGCTCTGGGATTATGGGATG

(계속)

도면16b

AGTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGTT
 GTTTTCCAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGCAGAACAAATGAACAT
 TTGGGACTCCCTGGGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACTTTCAGAAAT
 CAGGCCCTCGTCCCTATTCCCTTCTATTCTAGCCTTATTTCTTATGAGGABGATCAGAGGCAAGGA
 GCAGAACCTAGAAAAAATTTGTCAAGCCTAATGAAACCAAAAATTTACTTTTGGAAAGTGCAACAT
 CATATGGCACCCACTAAAGATGAGTTTACTGCAABAGCCTGGGCTTATTTCTCTGATGTTGACCTG
 GAAAAAGATGTGCACTCAGGCCTGATGGACCCCTTCTGGTCTGCCACACTAACACACTGAACCCF
 GCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCCACCATCTTTGATGAGACCAAA
 AGCTGGTACTTCACTGAAAATATGGAAAGAAAATGCAAGGGCTCCCTGCAATATCCAGATGGAAGAT
 CCCACTTTTAAAGAGAATTTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGC
 TTAGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATC
 CATTCTATTCATTTTCAGTGGACATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTG
 TACAATCTCTATCCAGTGTTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAAATTTGGCGG
 GTGGAAATGCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAAT
 AAGTGTGAGACTCCCTGGGAATGGCTTCTGGACACATAGAGATTTTCAGATTACAGCTTCAGGA
 CAATATGGACAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCGGGATCAATCAATGCCTGGAGC
 ACCAAGGAGCCCTTTTCTTGGATCAAGGTGGATCTGTTGGCACCAATGATTATTCACGGCATCAAG
 ACCCAGGGTGCCCGTCAGAAGTTCCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTT
 GATGGGAAGAAGTGGCAGACTTATCGAGGAAATCCACTGGAACCTTAATGGTCTTCTTTGGCAAT
 GTGGATTCATCTGGGATAAAACACAATATTTTTAACCCCTCCAATTATTGCTCGATACATCCGTTTG
 CACCCAACTCATTATAGCATTTCGAGCACTCTTCGCATGGAGTTSATGGGCTGTGATTTAAATAGT
 TGCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCCTAC
 TTTACCAATATGTTTGGCACCTGGTCTCCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAAT
 GCCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAA
 GTCACAGGAGTAACCTACTGAGGGAGTAAAATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTC
 ATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTTCAGAAATGGCAAAGTAAAGGTTTTT
 CAGGGAAATCAAGACTCCTTCACACCTGTGGTGAACCTCTCTAGACCCACCGTTACTGACTCGCTAC
 CTTCGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAG
 GCACAGGACCTCTACTGA (서열 번호: 17)

도면17a

CH25-FL-NA

ATGCAGATCGAGCTGTCCACATGCTTTTTTCTGTGCCTGCTGCGGTTCTGCTTCAGCGCCACCCGG
 CGGTACTACCTGGGGCCCGTGGAGCTGTCTCTGGGACTACATGCAGAGCGACCTGGGCGAGCTGCC
 GTGGACGCCCGGTTCCCCCCCAGAGTGCCTCAAGAGCTTCCCCTTCAACACCAGCGTGGTGTACAAG
 AAAACCTGTTCSTGGAGTTCACCGACACCTGTCAACATCGCCAGCCAGGCCCCCTGGATG
 GGCCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGTATCACCCTGAAGAACATG
 GCCAGCGACCCCGTGAAGCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCCTCGGAGGGCGCCGAG
 TACGACGACCAGACCAGCCAGCGGGAGAAAGAGGACGACAAAGTCTTTCTGGCGGCAGCCACACC
 TACGTGTGGCAGGTCTGAAAGAAAACCGCCCCATGGCTCCGACCCCTGTGCTGACCTACAGC
 TACCTEAGCCACSTGGACCTGTTGAAGGACCTGAACAGCGGGCTGATGGGGCCCTGCTGCTCTGC
 CGGGAGGGCAGCCTGGCCAAAAGAGAAAACCCAGACCTGCACAAGTTCATCTGCTGTTTCGCGGTG
 TTCGACGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACCCGGGACGCGGCC
 TCTGCCAGAGCCTGGCCCAAGATGCACACCGTGAACGGCTACGTGAACAGAAGCCTGCCCGGCTG
 ATTGGCTGCCACCGGAAGAGCTGTACTGGCAGTGTATCGGCATGGGCACACACCCGAGGTGCAC
 AGCATCTTTCTGSAAGGGCACACCTTTCTGGTGGGAACACCGGCAGGCCAGCCTGGAAATCAGC
 CCTATCACCTTCTGACCGCCAGACACTGCTGATGGACCTGGGCCAGTTCCTGCTGTTTTGCCAC
 ATCAGCTCTCACCAGCACGACGGCATGGAAGCCTACGTGAAGGTGGACTCCTGCCCGAGGAACCC
 CAGCTGCGGATGAAGAACACCGAGGAAGCCGAGGACTACGACGACGACCTGACCGACAGCGAGATG
 GACGTGGTGGCTTCGACGACGACAAAGCCCGAGCTTCATCCAGATCAGAAGCCTGGCCAAGAAG
 CACCCCAAGACCTGGGTGCACTACATCGCCCGGAGGAAGAGGACTGGGACTACGCCCCCTGGTG
 CTGGCCCCCGACGACAGAAGCTACAAGAGCCAGTACCTGAACAATGGCCCCAGCGGATCGGGCCG
 AAGTACAAGAAAGTGGGTTTCATGGCTACACCGACGAGACCTCAAGACCCGGGAGGCCATCCAG
 CACGAGAGCGGCATCTGGGCCCTGCTGTACGGGGAAGTGGGGACACACTGCTGATCATCTTC
 AAGAACCAGGCCAGCCCGCCTACAACATCTACCCCCACGGCATCACCGACGTGGGGCCCTGTAC
 AGCAGGCGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCTGCCCGGCGAGATCTTC
 AAGTACAAGTGGACCGTGACCGTGGAGGACGGCCCCACCAAGAGCGACCCAGATGCCTGACCCGG
 TACTACAGCAGCTTCGTGAACATGGAACGGGACCTGGCTCCGGCTGATCGGACCTCTGCTGATC
 TGCTACAAAGAAAGCGTGGACGAGCGGGCAACCAGATCATGAGCGACAAGCGGAACGTGATCCTG
 TTCAGCGTGTTCGATGAGAACCGGTCCTGGTATCTGACCGAGAATCCAGCGGTTTCTGCCCAAC
 CCTGCCGGGGTGCAGCTGGAAGATCCCGAGTTCAGGCCAGCAACATCATGCACCTCCATCAATGGC
 TACGTGTTGCACAGCTGCAGCTGTCCGTGTGTCTGCACGAGGTGGCTACTGGTACATCTGAGC
 ATCGGCCCCAGACCGACTTCTGAGCGTGTCTTCAGCGGCTACACCTTCAAGCACAAGATGGTG
 TACGAGGACACCCTGACCCTGTTCCCTTTCAGCGGCGAGACCGTTCATGAGCATGGAAAACCCC
 GGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCCGGAACCGGGCATGACCGCCCTGCTGAAG
 GTGTCCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTACGAGGATATCAGCGCCTAC
 CTGCTGTCCAAGAACAACGCCATCGAGCCAGAAAGCTTCAGCCAGAACCCCTGTGCTGAAGCGG
 CACCAGAGAGATCACCCGGACCACCTGCAGTCCGACCAGGAAGAGATCGATTACGACGACACC

(계속)

도면17b

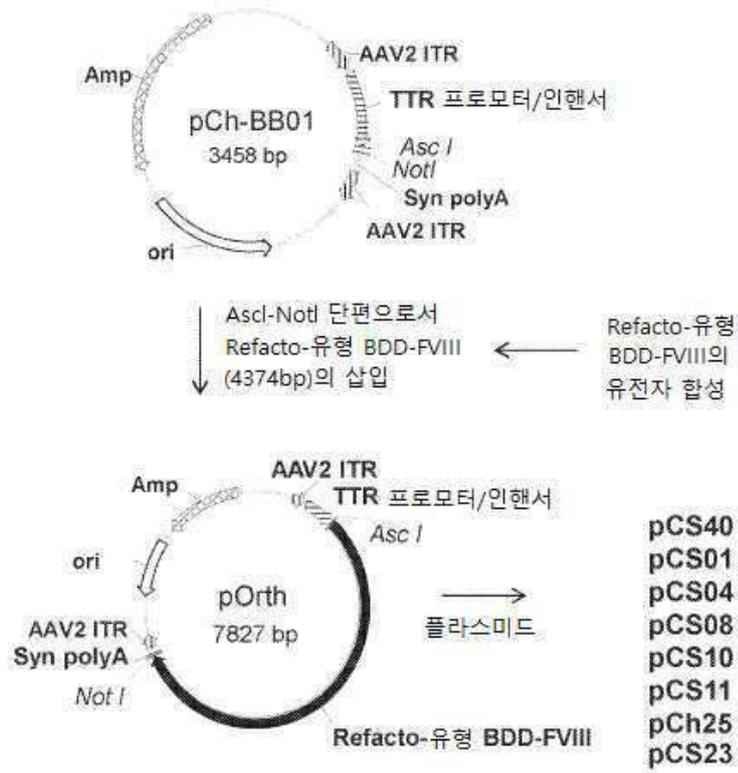
ATCAGCGTGGAGATGAAAAAGAAGATTTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCCGG
 TCCTTCCAGAAGAAAACCCGGCACTACTTTATCGCCGCGCTGGAGCGGCTGTGGGACTACGGCATG
 AGCAGCAGCCCCACGTGCTGCGGAACCGGGCCAGAGCGGCAGCGTGCCCCAGTTCAAGAAAGTG
 GTGTTCCAGGAATTCACCGACGGCAGCTTCACCCAGCCCTGTACCGGGGCGAGCTGAACGAGCAC
 CTGGGGCTGCTGGGGCCCTACATCAGGGCCGAAGTGGAGGACAACATCATGGTGACCTTCCGGAAT
 CAGGCCAGCAGACCCTACTCCTTCTACAGCAGCCTGATCAGCTACGAAGAGGACCAGCGGCAGGGC
 GCTGAACCCCGGAAGAACTTCGTGAAGCCCAATGAGACCAAGACCTACTTCTGGAAAGTGCAGCAC
 CACATGGCCCCCACCAGGACGAGTTCGACTGCAAGGCCTGGGCTACTTCAGCGACGTGGATCTG
 GAAAAGGACGTGCACTCTGGACTGATGGCCCTCTGCTGGTGTGCCACACCAACACCCTGAACCCC
 GCCCAGGCGGGCAGGTGACCGTGCAGGAATTCGCCCTGTTCTTCACCATCTTCGACGAGACCAAG
 TCCTGGTACTTCACCGAGAATATGGAACGGAACCTGCAGAGCCCCCTGCAACATCCAGATGGAAGAT
 CCTACCTTCAAAGAGAACTACCGGTTCCACGCCATCAACGGCTACATCATGGACACCCTGCCTGGC
 CTGGTGTATGGCCAGGACCAGAGGATCCGGTGGTATCTGCTGTCCATGGGCAGCAACGAGAATATC
 CACAGCATCCACTTCAGCGGCCACGTGTTCCACCGTGAAGAAAGAAAGAGTACAAGATGGCCCTG
 TACAACCTGTACCOCGGCGTGTTCGAGACCGTGGAGATGCTGCCAGCAAGGCGGCATCTGGCGG
 GTGGASTGTCTGATCGGCGAGCACCTGCATGCCGGGATGAGCACCTGTTTCTGGTGTACAGCAAC
 AAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCCGGGACTTCCAGATCACCGCCTCCGGC
 CAGTACGGCCAGTGGGCCCCCAAGCTGGCCCGGCTGCACTACAGCGGCAGCATCAACGCCTGGTCC
 ACCAAGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATATGATCATCCACGGCATTAA
 ACCCAGGGGGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTG
 GACGGCAAGAAGTGGCAGACCTACCGGGCAACAGCACCGGCACCCCTGATGGTGTCTTCGGCAAC
 GTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCATCATCGCCCGGTACATCCGGCTG
 CACCCACCCACTACAGCATCAGATCCACCCCTGGGGATGGAACCTGATGGGCTGCGACCTGAACCTC
 TGCAGCATGCCTCTGGGCATGAAAGCAAGGCCATCAGCGACGCCCAGATCACAGCCAGCAGCTAC
 TTCACCAACATGTTCCGCCACCTGGTCCCCCTCCAAGGCCAGGCTGCACCTGCAGGGCCGGTCCAAC
 GCCTGGCGGCCTCAGGTGAACAACCCCAAAGAATGGCTGCAGGTGGACTTTCAGAAAACCATGAAG
 GTGACCGGCGTGACCACCCAGGGCGTGAAGAAGCCTGCTGACCAGCATGTACGTGAAAGAGTTTCTG
 ATCAGCAGCAGCCAGGACGGCCACCACTGGACCCCTGTTCTTTCAAGACGGCAAGGTGAAAGTGTTC
 CAGGGCAACCAGGACTCCTTACCCCGTGGTGAACCTCCCTGGACCCCCCTGCTGACCCGCTAC
 CTGCGGATCCACCCCACTCTTGGGTGCACCAGATCCGCCCTGAGGATGGAAGTGTGGGATGTGAG
 GCCCAGGATCTGTACTGA (서열 번호: 18)

도면18

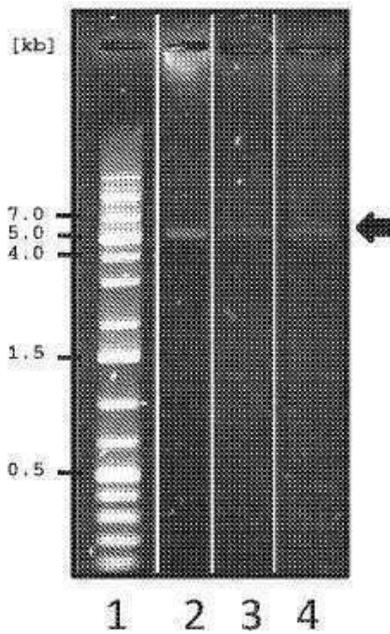
FVIII-FL-AA

mqieltstcfff	lcllrfcfssa	trryylgave	lswdymqsdl	gelpvdarfp	prvpksfpfn
tsvvykktlf	veftdhlfnl	akprppwmg	lgptiqaevy	dtvvitlknm	ashpvslhav
gvsywkaseg	aeyddqtsqr	ekeddkvfp	gshtyvwwvl	kengpmasdp	lcltysylsh
vdlvkdlnsq	ligallvcre	gslakektqt	lkkfillfav	fdegkswhse	tknslmqdrd
aasarawpkm	htvngyvnr	lpgligchrk	svywhvigmg	ttpevhsifl	eghtflvrnh
rqasleispi	tfltaqtllm	dlqqfllfch	isshqhdgmé	ayvkvdscpe	epqirmkne
eaedydddl	dsemdvvrfd	ddnspsfigi	rsvakkhpk	wwhyiaae	dwdyaplvl
pddrsyksqy	lmgpqrigr	kykkvrfmay	tdefktrrea	iqhesqilgp	llygevgdtl
liifknqasr	pyniypggit	dvrplysrri	pkgvkhkdf	pilpgeifky	kwvtvtedgp
tksdprcltr	yyssfvmer	dlasgligpl	licykesvdg	rgnqimsdkr	nvilfsvfde
nrswylteni	qrflprpagv	qledpefqas	nimhsingyv	fdslqlsvcl	hevaywyils
igaqtdflsv	ffsgytfkhh	myvedtltlf	pfsgetvims	menpglwiig	chnsdfrarg
mtallkvssc	dkntgdyyed	syedisayll	sknnaiexpr	fsqnsrhpst	rqkgfnatti
pendiektdp	wfahrtpmk	iqnvssddl	mlrqsptph	qlslsdlqea	kyetfsddps
pgaidssnsl	semthfrpql	hhsqdmvftp	esglqrlne	klgttaatel	kkldfkvsst
snnlistips	dnlaagtdnt	sslgppsmv	hydsqldttl	fgkksplte	sggplslsee
nndskillesg	lmsqesswg	knvsstesgr	lfkgkrahgp	alltkdnalf	kvsisilktn
ktsnnsatnr	kthidgpsll	ienspsvwgn	ilesdtefkk	vtplihdrml	mdknatalrl
nhmsnkttas	knmemvqqk	egpipdaqn	pdmsffkmlf	lpesarwiqr	thgknslnag
qgpspkqlvs	lgpeksvegg	nflseknkv	vgkgeftkdv	glkemvfps	rnlfltnldn
lhenathnqe	kkigeeiekk	etliqenvvl	pqihvtvgtk	nfmknlflls	trqnvegsyd
gayapvlqdf	rslnstart	kkhtahfskk	geeenleglg	nqtkqiveky	acttrispt
sqqnfvtrqs	kralkgfrlp	leetelekri	ivddstqws	knmkhltpst	ltqidyneke
kgaitqspis	dcltrshsip	qanrspipia	kvssfpirp	yltrvlfqd	nshlpaasy
rkkdsgvqes	shflggakkn	nlsalittle	mtgdqrevgs	lgtsatnsvt	ykkventvlp
kpdlpktsgk	vellpkvhiy	qkdlfptets	ngspghldlv	egsllqgteg	aikwneanrp
gkvpflrvat	essaktpskl	ldplawdnhy	gtqipkeewk	sqekspekta	fkkttilsl
nacesnhaia	aineggnkpe	ievtwakqgr	terlcsqnp	vlkrhqreit	rttiqsdqee
idyddtisve	mkkedfdiyd	edengsprsf	qkktthyfia	averlwdygm	sssphvlrnr
aqsgsvpqfk	kvvfgeftdg	sftqplyrge	lnehlgllgp	yiraavedni	mvtrfnqasr
pysfysslis	yeedqrqgae	prknfvkpne	tktyfwkvqh	hmaptkdefd	ckawayfsdy
dlekdvhsgl	igp1lvchn	tlnpahgrqv	tvqefalfft	ifdetkswyf	tenmerncra
pcniqmedpt	fkenyrfhai	ngyimdtlpg	lymaqdrir	wyllsmgsne	nihsihfagh
vftvrkkeey	kmalynlypg	vfetvemlps	kagiwrvecl	igehlhagms	tlfivysakc
qtplgmasgh	irdfqitasg	gyggwapkla	rlhysgsina	wstkepfswi	kvdlapmi
hgiktqgarq	kfsslyisqf	iimysldgkk	wqtyrgnstg	tlmvffgnvd	ssgikhnifn
ppliaryirl	hpthysirst	lrmelmgcdl	nscsmplgme	skaisdaqit	assyftmfma
twspkskarlh	lqgrsnawrp	gvnnpkewlq	vdqktmkvt	gvttqgvksl	ltsmyvkefl
issqdghqw	tlffqngkvk	vfggnqdsft	pvvnslpppl	ltrylrhpbq	swvhqialrm
evlgceaqdl	y (서열 번호: 19)				

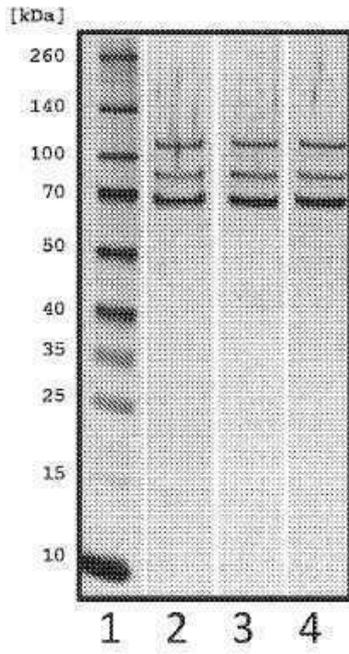
도면19



도면20



도면21



도면22a

CS23-FL-NA

atgcagattgagctgagcaacctgcttcttccctgtgacctgctgaggttctgcttctctgccaccagg
 agatactacotgggcgccgtggagctgagctgggactacatgcagctctgacctggcgagctgacct
 gtggacgccagggttcccccccagagtgcacaagagcttccccctcaacacctcagtggtgtacaag
 aagacctgttctgtggagttcaaccgaccacctgttcaacatcgccaagcccaggccccctggatg
 ggctgctgggccccaccatccaggccgaggtgtacgacacctgggtgatccctgaagaacatg
 gccagccaccccgctgagcctgcaccgctgggctgagctactggaaggcctctgagggcgccgag
 tatgacgaccagaccagccagaggagaaggaggacgacaagggtgttccccggcggcagccacacc
 tacgtgtggcagggtgctgaaggagaacggccccatggccagcgacctctgtgctgacctacagc
 tacctgagccacctggacctgggaaggacctgaactctggcctgatcggcgccctgctggtgtgc
 agggagggcagcctggccaaggagaagaccagacctgcacaagtctatcctgctgttccgctgt
 ttogatgagggcaagagctggcacagcgagaccagaacagcctgatagcaggacagggatgccgct
 tctgccagggcctggccccagatgcacacctggaacggctacgtgaacaggagcctgcccgccctg
 atcggctgccacaggaagtctgtgactggcacgtgatcggcatgggcaccacccccagggtgac
 agcatcttctggagggccacaccttctctgggtgaggaaccaaggcaggccagcctggagatcagc
 ccoctaccttctgacctgcccagacctgctgatggaacctgggcaagtctctgctgttctgccac
 atcagcagccaaccagcacagcggcatgggacctacgtgaagggtggacagctgccccgaggagccc
 cagctgaggatgaagaacaacgaggaggccgaggactatgatgatgacctgaccgactctgagatg
 gacgtggtgaggtttgatgatgacaacagccccagcttcatccagatcagggtctgtggccaagaag
 ccccccaagacctgggtgacctacatcgccgcccaggaggaggactgggactaccgccccctgggtg
 ctggcccccgacgacaggagctacaagagcccagctacctgaacaacggccccagaggatcggcagg
 aagtacaagaaggctcagattcatggcctacaccgaagagaccttcaagaccaggaggccatccag
 cagagctctggcatcctgggccccctgctgtacggcgagggtgggccaacctgctgatcatcttc
 aagaaccaggccagcaggccctacaacatctacccccacggcatcaccgatgtgaggccctgtac
 agcaggaggctgcccaaggcgtgaagcactgaaggacttccccatcctgccggcgagatcttc
 aagtacaagtggaacctgacctggaggatggccccaccaagtctgaccccagggtgctgaccagg
 tactacagcagcttcgtgaacatggagagggaacctggcctctggcctgatcggccctctgctgatc
 tgctacaaggagagcgtggaccagaggggcaaccagatcagctctgacaagaggaacgtgatcctg
 ttctctgtgtogatgagaacaggagctgggtatctgaccgagaacatccagagggttctgcccaac
 cccgcccggctgcagctggaggaccccaggttccaggccagcaacatcatgcacagcatcaacggc
 tacgtgttcgacagcctgcagctgtctgtgtgacctgcacgaggtggcctactggtacatctctgagc
 atcggcgcccagaccgacttctctgtctgtgttcttctctggetacaccttcaagcacaagatgggtg
 taegaggacacctgacctgttccccctcagcggcgagaccgtgttatgagcatggagaacccc
 ggctgtggatcctgggtgccacaacagcagcttcaggaacaggggatgaccgacctgctgaaa
 gtcagcagctgcgacaagaacaccggcgactactacgaggacagctacgaggacatcagcgcctac
 ctgctgagcaagaacaacgccatcgagcccaggagcttcagccagaacccccctgctgaagagg
 caccagaggagatcaccaggaccacctgcagagcgaccaggaggagatcgaactatgatgacacc

(계속)

도면22b

atcagcgtggagatgaagaaggaggacttcgacatctacgacgaggacgagaaccagagccccagg
 agcttccagaagaagaccaggaactacttcatcgccgctggagaggtgtgggactatggcatg
 agcagcagccccacgtgctgaggaacagggcccagagcggcagcgtgcccagttcaagaaggtg
 gtgttccaggagttcaccgacggcagcttcaaccagccccgtacagaggcagctgaacgagcac
 ctgggctgtgtggccccctacatcagggccgaggtggaggacaacatcatggtgaccttcaggaac
 cagggccagcagggccctacagcttctacagcagcctgatcagctacgaggaggaccagagggcagggc
 gccgagcccaggaagaacttcgtgaagcccacgagaccaagacctaacttctggaaggtgcagcac
 cacatggccccaccaggacaggttcgaetgcaaggcctgggctacttctctgatgtggacctg
 gagaaggacgtgcacagcggcctgatcggccccctgctggtgtgcccacccaacacctgaacccc
 gccccagggcagggcaggtgacgtgcaggagtteggcctgttcttccaccttctegacgagaccaag
 agctggtaacttcaecgagaacatggagaggaactgcagggccccctgcaacatccagatggaggac
 ccaccttcaaggagaactacaggttccaogccatcaacggctacatcatggacaacctgccccggc
 ctggtgatggcccaggaccagaggatcaggtggtatctgctgagcatgggcagcaacgagaacatc
 cacagcatccacttcagcggccaaggtgttaccggtgaggaagaaggaggagtacaagatggccctg
 tacaacctgtaccccggcgtgttccgagaccgtggagatgctgccagcaaggccggcatctggagg
 gtggagtgccctgatcggcagcaccctgcacccggcagtgagcaccctgttccctggtgtacagcaac
 aagtgccagacccccctgggcatggccagcggccacatcagggacttccagatcaccgcctctggc
 cagtacggccagtgggcccccaagctggccaggctgcactacagcggcagcatcaacgcctggagc
 accaaggagcccttcagctggatcaaggtggaacctgctggcccccatgatcatccacggcatcaag
 acccagggcgccagggcagaagttccagcagcctgtacatcagccagttcateatcatgtacagcctg
 gacggcaagaagtgccagacctacaggggcaacagcaccggcaccctgatggtgttcttcggcaac
 gtggacagcagcggcatcaagcacaacatcttcaaccccccatcategccaggtacatcaggetg
 caccccccactacagcabcaggagcaccctgggatggaactgatgggctgagacctgaacagc
 tgcagcatgccctgggcatggagagcaaggccatctctgacgccagatcaccgccagcagctac
 ttcaccaacatgttcgccacctggagccccagcaaggccagcctgcacctgcagggcagggagcaac
 gctggagggccccaggtgaacaacccccaaaggagtggtgcaggtggacttccagaagaccatgaag
 gtgaccggcgtgaccaccagggcgtgaagagcctgctgaccagcatgtacgtgaaggagttcctg
 atcagcagcagccagcagccaccagctggacctgttcttccagaaggcgaagtgaaagtggtc
 cagggcaaccagcagcagcttccCcccgtggtgaacagcctggaacccccctgctgaccaggtat
 ctgaggtaccacccccagagctgggtgcaccagatcgccctgagaatggaagtgctgggatgcgag
 gccaggacctgtactga (서열 번호: 20)

도면23

CS23-FL-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDAR' FPPRVPKSFPFNTSVVYK
 KTLFVEFTDHLFNIAKPRPPWMLLGPITQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEY
 DDQTSQREKEDDKVFPGGSHTYVWQVLKENGFMASDPLCLTYSLSHVDLVKDLNSGLIGALLVCRE
 GSLAKEKTQTLHKFILLFAVDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGC
 HRKSVYWHVIGMGTTPPEVHSIFLEGHTFLVRNHRQASLEISPIITFLTAQTLMDLGQFLFCHISSH
 QHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSMDVVRFDDDNSPFIQIRSVAKKHPKTW
 VHIIAAEEEDWDYAPLVLPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETEKTRAIQHESGIL
 GPLLYGEVGDPELLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIPKYKWTVT
 VEDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLLIICYKESVDQRGNQIMSDKRNVI LFSVFDENR
 SWYLTENIQRF LPNPAGVQLEDPEFOASNIMHSINGYVFDLSQLSVCLHEVAYWYILSIGAQTDFLS
 VFFSGYTFKHKMVYEDTLTLPFSGETVFMSENPLWILGCHNSDFRNRGMTALLKVSSCDKNTGD
 YEDSYEDISAYLLSKNNAIEPRSFSQNPVVKRHRQREITRTTLQSDQEEIDYDDTISVEMKKEDEFD
 IYDEDENQSPRSFQKTRHYFIAAVERLWDYGMSSSPHVLNRNRAQSGSVFQFKKVVQEFDTDGSFTQ
 PLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPKRNFKPNE'
 KTYFWKVQHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALF
 FTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRPHAINGYIMDTLPGLVMAQDQRI RWYLLS
 MGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFEVEMLPKAGIWRVECLIGEHLHAGMSTL
 FLVYSNKQTPPLGMASGHIRDFQITASGQYQWAPKLARLHYSGSINAWSTKEPFSWIKVDDLAPMI
 IHGIKTQGARQKFSLSYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIAR
 YIRLHPHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQ
 RSNAWRPQVNNPKWQVDFQKTMKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVK
 VFQGNQDSFTPVVNSLDPPLLTRYLRTHPQSWVHQIALRMEVLGCEAQDLY (서열 번호: 21)

도면24

CS23-HC-NA

```

accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gctctgacctg
ggcgagctgc ctgtggacgc caggttcccc ccagagtgcc ccaagagctt ccccttcaac
acctcagtgg tgtacaagaa gacctgttcc gtggagtcca ccgaccacct gttcaacatc
gccaaagcca ggccccctcg gatgggcctg ctgggccccca ccataccaggc cgaggtgtac
gacaccctgg tgatccctc gaagaacatg gccagccacc ccgtgagcct gcacgccgtg
ggcgtgagct actggaaggc ctctgagggc gccgagtatg accgaccagac cagccagagg
gagaaggagg acgacaaggt gttccccggc gccagccaca cctacgtgtg gcaggtgctg
aaggagaacg gccccatggc cagcgacccc ctgtgcctga cctacagcta cctgagccac
gtggacctgg tgaaggacct gaactctggc ctgateggcg cctgctggt gtgcaggag
ggcagcctgg ccaaggagaa gaccagacc ctgcacaagt tcactctget gttcgccgtg
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgatgca ggacagggat
gccgcctctg ccagggcctg gcccaagatg cacaccgtga acggtacgt gaacaggagc
ctgccccggc tgatcggtg ccacaggaag tctgtgtact ggcacgtgat cggcatgggc
accacccccg aggtgcacag catcttctctg gagggccaca ccttctggt gaggaaccac
aggcaggcca gccctggagat cagccccatc accttctga ccgcccagac cctgctgatg
gacctgggcc agttcctgct gttctgccc atcagcagcc accagcacga cggcatggag
gcctacgtga aggtggacag ctgccccgag gagccccagc tgaggatgaa gaacaaccag
gaggccgagg actatgatga tgacctgacc gactctgaga tggacgtggt gaggtttgat
gatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc
tggtgact acatcgccc cgaggaggag gactgggact acgccccct ggtgctggcc
cccgacgaca ggagctacaa gagccagtac ctgaacaacg gcccccagag gatcggcagg
aagtacaaga aggtcagatt catggcctac accgacgaga ccttcaagac cagggaggcc
atccagcacg agtctggcat cctgggcccc ctgctgtacg gcgaggtggg cgacaccctg
ctgatcatct tcaagaacca ggcagcagg cctacaaca tetacccca cggcatacc
gatgtgaggc ccctgtacag caggaggctg ccaaggggcg tgaagcacct gaaggactt
ccatcctgc ccggcgagat ctcaagtac aagtggaccg tgaccgtgga ggatggccc
accaagtctg accccaggtg cctgaccagg tactacagca gcttcgtgaa catggagagg
gacctggcct ctggcctgat cggccccctg ctgatctgct acaaggagag cgtggaccag
aggggcaacc agatcatgtc tgacaagagg aacgtgatcc tgttctctgt gttogatgag
aacaggagct ggtatctgac cgagaacatc cagaggttcc tgcccaccc cgccggcgtg
cagctggagg acccogagtt ccaggcccagc aacatcatgc adagcatcaa cggtacgtg
ttcgacagcc tgcagctgtc tgtgtgectg cacgaggtgg cctaotggtg catcctgagc
atggcgccc agaccgactt cctgtctgtg ttcttctctg gctacacctt caagcacaag
atggtgtacg aggacacct gacctgttc ccttcagcg gogagacct gttcatgagc
atggagaacc ccggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacagggc
atgaccgcc tgcgtaaaagt cagcagctgc gacaagaaca ccggcgacta ctacgaggac
agctacgagg acatcagcgc ctacctgctg agcaagaaca acgcatcga gccccagg
(서열 번호: 22)

```

도면25

CS23-LC-NA

```

                                g agatcaccag gaccaccctg
cagagcgacc aggaggagat cgactatgat gacaccatca gcgtggagat gaagaaggag
gacttcgaca tetacgacga ggacgagaac cagagcccca ggagcttcca gaagaagacc
aggcactact tcatcgccgc cgtggagagg ctgtgggact atggcatgag cagcagcccg
cacgtgctga ggaacagggc ccagagcggc agcgtgcccc agttcaagaa ggtggtgttc
cagaggttca ccgacggcag cttcaccag cccctgtaca gaggcgagct gaacgagcac
ctgggcctgc tgggcacctc catcagggcc gaggtggagg acaacatcat ggtgacctc
aggaaccagg ccagcaggcc ctacagcttc tacagcagcc tgatcagcta cgaggaggac
cagaggcagg gcgccgagcc caggaagaac ttcgtgaagc ccaacgagac caagacctac
ttctggaagg tgcagacca catggccccc accaaggacg agttcgactg caaggcctgg
gcctaactct ctgatgtgga cctggagaag gaodtgaca gcggcctgat cggccccctg
ctggtgtgcc acaccaacac cctgaacccc gccacggca ggcaggtgac cgtgcaggag
ttcgccctgt tcttcaecat cttcgaagag accaagagct ggtacttcac cqagaacatg
gagaggaact gcagggcccc ctgcaacatc cagatggagg accccacctt caaggagaac
tacaggttc acgcoatcaa cggctacatc atggacacc tgcocggcct ggtgatggc
caggaccaga ggatcaggtg gtatctgctg agcatgggca gcaacgaga catccacagc
atccacttca ggggccacgt gttcacctg aggaagaagg aggagtaca gatggccctg
tacaacctgt accccggcgt gttcgagacc gtggagatgc tgcccagcaa ggccggcacc
tgaggggtg agtgctgat cggcgagcac ctgcacggc gcatgagcac cctgttctg
gtgtacagca acaagtgcc gacccccctg ggcabggcca gcggccacat cagggacttc
cagatcaccg cctctggcca gtacggccag tgggccccca agctggccag gctgcactac
agcggcagca tcaacgcctg gacaccaag gagcccttca gctggatcaa ggtggacctg
ctggccccca tgatcatcca cggcatcaag acccagggcg ccaggcagaa gttcagcagc
ctgtacatca gccagttcat catcatgtac agcctggacg gcaagaagtg gcagacctac
aggggcaaca gcaccggcac cctgatggtg ttcttcggca acgtggacag cagcggcacc
aagcacaaca tcttcaacce ccccatcacc gccaggtaca tcaggctgca ccccaccac
tacagcatca ggagcaccct gcggatggaa ctgatgggct gcgacctgaa cagctgcagc
atgcccctgg gcatggagag caaggccatc totgacgccc agatcacccg cagcagctac
ttaccaaca tgttggccac ctggagcccc agcaaggcca ggotgcacct gcagggcagg
agcaacgcct ggagcccca ggtgaacaac cccaaggagt ggotgcaggt ggacttccag
aagaccatga aggtgaccgg cgtgaccacc cagggcgtga agagcctgct gaccagcatg
tacgtgaagg agttcctgat cagcagcagc caggacggcc accagtggac cctgttcttc
cagaacggca aagtgaaggt gttccagggc aaccaggaca gcttcacccc cgtggtgac
agcctggacc cccccctgct gaccaggtat ctgaggatcc acccccagag ctgggtgac
cagatcgccc tgagaatgga agtgctggga tgcgaggccc aggacctgta c

```

(서열 번호: 23)

도면26

CS01-HC-NA

```

                                gcc
accaggagat actaactggg ggctgtggaa ctttcttggg actacatgca gtctgaacctg
ggagagctgc ctgtggatgc caggttccca cccagagtgc ccaagtcctt cccattccaac
acctctgtgg tctacaagaa gacactcttt gtggaattca ctgaccacct gbtcaacatt
gcaaaaccca gaccaccctg gatgggactc ctgggaccea ccatteagge tgaggtgtat
gacactgtgg tcateacct caagaacatg goatcccacc ctgtgtctct gcatgctgtg
ggagtctcat actggaaggo ctotgaaggg gctgagtatg atgaccagac atcccagaga
gagaagagg atgacaaggt gtccctctgg ggatctcaā cctatgtgtg gcaagtctc
aaggagaatg gaaccatggc atctgaccca ctctgcctga catactctca cctttctcat
gtggaacctg tcaaggacct caactctgga ctgattgggg cactgctggt gtgcaggaa
ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcattctctt gtttctgtc
ttgatgagg gcaagtcttg gcaacttgaa acaagaact cctgatgca agacagggat
gctgctctg ccaggctatg gcccaagatg cacactgtga atggtatgt gaacagatca
ctgcctggac taattggctg ccapaggaa totgtctact ggcattgtgt tggcatggg
ccaaccctg aagtgcact cactttctct gagggacaca ccttctggt caggaaccac
agacaagcct ctctggagat ctctccatc accttctca ctgcacagac actgctgatg
gacctggac agtctctgtt gttctgcac atctcttccc accagcatga tggcatggaa
gcctatgtca aggtggactc atgcoctgag gaaccacagc tcaggatgaa gaacaatgag
gaggtgagg actatgatga tgacctgact gactctgaga tggatgtgt cagatttgat
gatgacaact ctccatctt ccttcagatc aggtctgtgg caaagaaaca cccaagaca
tgggtgcaet acattgctgc tgaggaagag gactgggact atgcaccact ggtctctggc
cctgatgaca ggagctacaa gtctcagtao ctcaacaatg gcccaaaaag aattggaaga
aagtacaaga aagtcaagatt catggctac actgatgaaa ccttcaagac aagagaagcc
attcagcatg agtctggcat tctgggacca ctctgtatg gggaaagtgg agacacctg
ctcctctct tcaagaacca ggcctccagg cctacaaca tetaccaca tggcatcact
gatgtcaggc cctgttacag caggagactg ccaaaaggg tgaaaacact caaggacttc
cccattctg ctggagagat ctccaagtac aagtggactg tcactgtgga ggatggacca
ācaaaagtctg accccaggtg cctcaccaga tactactct ctttgtgaa catggagaga
gacctggcat ctggactgat tggaccactg ctcatctgt acāaggagt tgtggaccag
agaggcaacc agatcātgtc tgacaagaga aatgtgatto tgttctctgt ctttgatgag
aacagatcat ggtacctgac tgagaacatt cagagattcc tgccaacc tgctggggtg
caactggaag acctgagtt ccaggcaagc aacatcātgc actccatcaa tggctatgtg
ttgactctc tccagcttcc tgtctgctg catgaggtgg cctactggta cattctttct
attggggcac aaactgaett cctttctgtc ttctctctg gatcacctt caagcacaag
atggtgtatg aggacacct gatctctctg ccattctctg gggaaactgt gtctatgagc
atggagaacc ctggaactgt gattctggga tgccaacact ctgacttcag aaacagggga
atgactgac tgotcaaggt ctctctctgt gacaagaaca ctgggacta ctatgaggac
tcttatgagg acatctctgc ctactctctc agcaagaaca atgccattga gcccaaga
( 서열 번호: 24 )

```

도면27

CS01-LC-NA

```

                                g agatcaccag gacaaccctc
cagtcctgacc aggaagagat tgactatgat gacaccattt ctgtggagat gaagaaggag
gactttgaca totatgatga ggacggagaac cagtctccaa gatcattcca gaagaagaca
agacactact tcattgctgc tgtggaaaga ctgtgggact atggcatgtc ttctctccc
catgtectca ggaacagggc acagtctggc totgtgccac agtcaagaa agtggctctc
caggagtcca ctgatggctc attcaccag ccectgtaca gaggggaact gaatgagcac
ctgggactcc tgggaccata catcaggctt gaggtggaag acaacatcat ggtgacattc
agaaccagg cctccaggcc ctacagcttc tactcttccc tcacagcta tgaggaagac
cagagacaag gggctgagcc aagaaagaac ttgtgaaac ccaatgaaac caagacctac
ttctggaaag tccagcacca catggcacc cccaaggatg agtttgactg caaggcctgg
gcatacttct ctgatgtgga cctggagaaa gatgtgact ctggcctgat tggcccactc
ctggctctgc acaccaacac cctgaaccct gcacatggaa ggcaagtgc tgtgcaggag
tttgccctct tcttcacct cttfgatgaa accaagtcac ggtactcac tgagaacatg
gagagaaact gcagagcacc atgcaacatt cagatggaag accccacctt caaggagaac
tacaggttcc atgccatcaa tggctacatc atggacacc tgcctgggct tgtcatggca
caggaccaga gaatcagatg gtacctgctt totatgggat ccaatgagaa cattcactcc
atccacttct ctgggcatgt cttcactgtg agaaagaagg aggaatacaa gatggccctg
tacaacctct accctggggg ctttgagact gtggagatgc tgcctccaa agctggcatc
tggagggtgg aatgcctcat tggggagcac ctgcatgctg gcatgtcaac cctgttctctg
gtctacagca acaagtgcc gacaccctg ggaatggcct ctggccacat cagggacttc
cagatcactg cctctggcca gtatggccag tgggcacca aactggccag gctccaetac
tetggctcca tcaatgcctg gtaaccaag gagecattct cttggatcaa ggtggacctg
ctggcaccca tgatcattca tggcatdaag acacagggg caagacagaa attctctct
ctgtacatct cacagtteat catcatgtac tctctggatg gcaagaagtg gcagacatac
agaggcaact ccactggcac cctcatggtc ttctttggca atgtggacag ctctggcatc
aagcacaaca tcttcaacce tccatcatt gccagataca tcaggctgca ccccaccac
tactcaatca gateaacctt caggatgga ctgatgggat gtgacctgaa ctctgctca
atgcccctgg gaatggagag caaggccatt tctgatgccc agatcactgc atcctcttac
ttcaaccaada tgtttgccac ctggtcacca tcaaaaagcca ggcctgcacct cdaggggaaga
agcaatgcct ggagacccca ggtcaacaac ccaaaggaat ggctgcaagt ggacttccag
aagacaatga aagtcaactg ggtgacaacc caggggtca agtctctgct cacctcaatg
tatgtgaagg agttctgat ctcttctca caggatggcc accagtgga actcttctc
cagaatggca aagtcaaggt gttccaggg aaccaggact ctttcaacc tgtggtgaac
tcaactggacc cccccctct gacaagatac ctgagaatto acccccagtc ttgggtccac
cagattgccc tgagaatgga agtcttggga tgtgaggcac aagacctgta c

```

(서열 번호: 25)

도면28a

CS01A(760-1667) - CS01-SC1-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTTCTGTGCCGTGCTGAGATTCTGCTTCTCTGCCACCAGGAGATAC
 TACCCTGGGGGCTGTGGAACTTTCTTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCCTGTGGATGCCAGG
 TTCCCACCCAGAGTGCCCAAGTCCCTTCCCATTC AACACCTCTGTGGTCTACAAGAGACACTCTTTGTGGAA
 TTCCTGACCCACCTGTTCAACATTTGCAAAAACCCAGACCACCTGGATGGGACTCCTGGGACCCACCATTGAG
 GCTGAGGTGTATGACACTGTGGTCATCACCTCAAGAACATGGCATCCCACCTGTGTCTCTGCATGCTGPG
 GGAGTCTCATACTGGAAAGCCCTGAAAGGGCTGAGTATGATGACCCAGACATCCCAGAGAGAGAAAGAGGAT
 GACAAAGGTGTTCCCTGGGGGACTCACACCTATGTGTGGCAAGTCTCAAGGAGAAATGGACCCATGGCATCT
 GACCCACTCTGCCTGACATACTCCTACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGAT
 GGGGCACTGCTGGTGTGCAGGGAAGGATCCCTGGCCAAAGGAGAAAACCCAGACACTGCACAAGTTCAATTC
 CTGTTTGTCTCTTTGATGAGGGCAAGTCTTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGAT
 GCTGCCTCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTC
 ATTTGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTTGGCATGGGGACAACCCCTGAAGTGCACCTCCAT
 TTCTGGAGGGACACACCTTCCCTGGTCAGGAACCAAGCAAGCCCTCTCTGGAGATCTCTCCCATCACTTC
 CTCCTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTCTGCCACATCTCTPCCCACAGCAT
 GATGCCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCCACAGCTCAGGATGAAGAACAATGAG
 GAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCAAGTTGATGATGACAACCTCT
 CCATCCTTCATTCAGATCAGGTCTGTGGCAAGAAACACCCCAAGACATGGGTGCACACTACATTTGCTGCTGAG
 GAAGAGGACTGGGACTATGCACCACTGGTCCCTGGCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAAC
 AATGGCCCAAAAGAAATTTGAAGAAAGTACAAGAAAGTCAGATTCATGGCTACACTGATGAAACCTTCAAG
 ACAAGAGAAGCCATTTCAGCATGAGTCTGGCATTCTGGACCACTCCTGTATGGGGAAAGTGGGAGACCCCTG
 CTCATCATCTTCAAGAACCAGGCCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTGAGGCC
 CTGTACAGCAGGAGACTGCCAAAAGGGTGAACACCTCAAGGACTTCCCATTTCTGCCTGGAGAGATCTTC
 AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCAGGTGCCTCACCAGATACTAC
 TCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTTGGACCACTGCTCATCTGCTACAAGGAG
 TCTGTGGACCAGAGAGGCAACCAGATGATGTCTGACAAGAGAAATGTGATTCGTCTCTCTGTCTTTGATGAG
 AACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCCAACCCCTGCTGGGGTGCAACTGGAAGAC
 CCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGCTATGTGTTGACTCTCTCCAGCTTTCTGTC
 TGCCCTGCATGAGGTGGCTACTGGTACATTCCTTCTATTTGGGGCACAACTGACTTCCTTTCTGTCTCTCTC
 TCTGGATACACCTTCAAGCACAAAGATGGTGTATGAGGACACCCTGACACTCTTCCCATTCCTTGGGGAACT
 GTGTTTCATGAGCATGGAGAACCCCTGGACTGTGGATTCGGGATGCCACAACCTGACTTCAGAAACAGGGGA
 ATGACTSCACTGCTCAAAGTCTCCTCCTGTGACAAGAACAACCTGGGGACTACTATGAGGACTCTTATGAGGAC
 ATCTCTGCCTACTGTCTCAGCAAGAACAATGCCATTTGAGCCAGAGAGATCACCAGGACAACCCCTCCAGTCT
 GACCAGGAAGAGATTGACTATGATGACACCTTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGAT
 GAGGACGAGAACAGTCTCCAAAGATCATCCAGAAGAAGACAAGACACTACTTCATTTGCTGCTGTGGAAAGA
 CTGTGGACTATGGCATGTCTTCCCTCCTCCCATGTCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAG
 TTCAAGAAAGTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCTGTACAGAGGGGAACTGAAT
 GAGCACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACACATCATGGTGCATTCAGAAAC
 CAGGCTCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGGCTGAG
 CCAAGAAAGAACTTTGTGAAACCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACACATGGCACCC

(계속)

도면28b

```

ACCAAGGATGAGTTTGACTGCAAGGCCCTGGGCATACTTCTCTGATGTGGACCTGGAGAAAGATGTGCACTCT
GGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACACCCCTGAACCTTGCACATGGAAGGCCAAGTACTGTG
CAGGAGTTTGGCCCTCTTCTTCCACATCTTTGATGAAACCAAGTCATGGTACTTCACTGAGAACATGGAGAGA
AACTGCAGAGCACCATGCAACATTCAGATGGAGACCCACCTTCAAGGAGAACTACAGGTTCATGCCATC
AATGGCTACATCATGGACACCCTGCCCTGGGCTTGTTCATGGCACAGGACCAGAGAATCAGATGGTACCTGCTT
TCTATGGGATCCAAATGAGAACATTCACCTCCATCCACTTCTCTGGGCATGCTTCACTGTGAGAAAGAAGGAG
GAATACAAGATGGCCCTGTACAACTCTACCCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCT
GGCATCTGGAGGGTGAATGCCCTCAATGGGGAGCACCTGCATGCTGGCATGTCAACCCCTGTTCCCTGGTCTAC
AGCAACAAGTGCAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCCTCTGGC
CAGTATGGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCAACCAAG
GAGCCATTCTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTGATGGCATCAAGACACAGGGGGCA
AGACAGAAATTCTCCTCTCTGTACATCTCAGAGTTCATCATCATGTAATCTCTGGATGGCAAGAAGTGGCAG
ACATACAGAGGCAACTCCACTGGCACCCCTCATGGTCTTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCAC
AACATCTTCAACCCCTGCCATCATTGCCAGATACATCAGGCTGCACCCACCCACTACTCAATCAGATCAACC
CTCAGGATGGAACATGATGGGATGTGACCTGAACTCCTGCTCAATGCCCCCTGGGAATGGAGAGCAAGGCCATT
TCTGATGCCAGATCACTGCATCCTCTTACTTCACCAACATGTTTGGCCCTGGTCAACCATCAAAAGCCAGG
CTGCACCTCCAGGGAGAAGCAATGCCCTGGAGACCCACAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGAC
TTCCASAGACAATGAAAGTCACTGGGCTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTG
AAGGAGTTCTTATCTTCTTCCACAGGATGGCCACCCAGTGGACACTCTTCTTCCAGAAATGGCAAGTCAAG
GTGTTCAGGGCAACCCAGGACTCTTTCACACCTGTGGTGAACCTCACTGGACCCCCCTCCTGACAGATAC
CTGAGAATTCACCCCCAGTCTTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCTGGGATGTGAGGCACAA
GACCTGTAATGA (서열 번호: 28)
    
```

도면29a

CS01A(772-1667) - CS01-SC2-NA

ATGCAGATGAGCTGTCCACCTGCTTCTTTCTGTGCCTGCTGAGATTCTGCTTCTCTGCCACCAGG
 AGATACTACCTGGGGGCTGTGGAACCTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
 GTGGATGCCAGGTTCCCACCCAGAGTGGCCAAAGTCTTCCCATTCAACACCTCTGTGGTCTACAAG
 AAGACACTCTTTGTGGAATTCACCTGACCACCTGTTCAACATTGCAAAAACCCAGACCACCCTGGATG
 GGACTCCTGGGACCCACCATTGAGGCTGAGGTGTATGACACTGTGGTCATCACCCCAAGAACATG
 GCATCCCACCCTGTGTCTCTGCATGCTGTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
 TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGGATCTCACACC
 TATGTGTGGCAAGTCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCC
 TACCTTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGG
 AGGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTTGTCTGTC
 TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
 TCTGCCAGGGCATGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
 ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGGACAACCCCTGAAGTGCAC
 CCCATCACCTTCCTCACCTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
 ATCTCTTCCCACCAGCATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
 CAGCTCAGGATGAAGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
 GATGTGGTCAGATTTGATGATGACAACCTCTCCATCCTTCATTGAGATCAGGTCTGTGGCAAAGAAA
 CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCCTGGTCT
 CTGGCCCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCAAAAGAATTGGAAGA
 AAGTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTGAG
 CATGAGTCTGGCAFTCTGGGACCACCTCCTGTATGGGGAAGTGGGAGACACCCTGCTCATCATCTTC
 AAGAACCAGGCCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTGAGGCCCTGTAC
 AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCATTCTGCTGGAGAGATCTTC
 AAGTACAAGTGGACTSTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAAGGTGCTCACCAGA
 TACTACTCCTCTTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCCTGCTCATC
 TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
 TTCTCTGTCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCTGCCAAC
 CCTGCTGGGGTGCACCTGGGAGACCTGAGTTCAGGCAAGCAACATCATGCACTCCATCAATGGC
 TATGTGTTTGACTCTCTCCAGCTTCTGTCTGCCTGCATGAGGTGGCCTACTGGTACATTCTTTCT
 ATTTGGGGCACAACCTGACTTCCCTTCTGTCTTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
 TATGAGGACACCCCTGACACTCTTCCCATTCTCTGGGGAACTGTGTTTATGAGCATGGAGAACCCCT
 GGACTGTGGATTCTGGGATGCCACAACCTGACTTCAGAAACAGGGGAATGACTGCCTGCTCAA
 GTCTCCTCCTGTGACAAGAACAACCTGGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCTAC
 CTGCTCAGCAAGAACAATGCCATTGAGCCAGAAAGCTTCTCTCAGAATTCAGACACCCAGCACC
 AGGGAGATCACCAAGGACACCCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACCAATTTCT
 GTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGACGAGAACCAGTCTCCAAGATCATTC

(계속)

도면29b

CAGAAGAAGACAAGACACTACTTCATGCTGCTGTGGAAAGACTGTGGGACTATGGCATGTCTCC
TCTCCCATGTCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAGAAAGTGGTCTTC
CAGGAGTTCAGTGTGGCTCATTACCCAGCCCTGTACAGAGGGAACTGAATGAGCACCTGGGA
CTCCTGGGACCATAACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAGAAACCAGGCC
TCCAGGCCCTACAGCTTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGGGCTGAG
CCAAGAAAGAACTTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCATG
GCACCCACCAAGGATGAGTTTGAAGGCTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTGGAGAAA
GATGTGCACTCTGGCCTGATTGGCCCACTCCTGGTCTGCCACACCAACAECCTGAAACCTGCACAT
GGAAGGCAAGTGAAGTGTGCAGGAGTTTGGCCCTCTTCTTCACCATCTTTGATGAAACCAAGTCAATGG
TACTTCACTGAGAACATGGAGAGAACTGCAGAGCACCATGCAACATTCAGATGGAAGACCCACCC
TTCAAGGAGAAGTACAGGTTCCATGGCATCAATGGCTACATCATGGACACCCCTGCCTGGGCTTGTG
ATGGCACAGGACCAGAGAATCAGATGGTACCTGCTTCTATGGGATCCAATGAGAACATTCACTCC
ATCCACTTCTCTGGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCCCTGTACAAC
CTCTACCCCTGGGGTCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGGGTGGAA
TGCCTCATTTGGGAGCACCTGCATGCTGGCATGTCACCCTGTTCTGGTCTACAGCAACAAGTGC
CAGACACCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCCTCTGGCCAGTAT
GGCCAGTGGGCACCCAAACTGGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCAACCAAG
GAGCCATTCTCTTGGATCAAGSTGGACCTGCTGGCACCCATGATCATTTCATGGCATCAAGACACAG
GGGCAAGACAGAAATTTCTCTCTGTACATCTCACAGTTCATCATCATGTACTCTCTGGATGGC
AAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCCTCATGCTCTTTGGCAATGTGGAC
AGCTCTGGCATCAAGCACAACATCTTCAACCCTCCCATCATTGGCCAGATACATCAGGCTGCACCCC
ACCCACTACTCAATCAGATCAACCCCTCAGGATGGAAGTGTGGGATGTGACCTGAACTCCTGCTCA
ATGCCCTTGGGAATGGAGAGCAAGGCCATTTCTGATGCCAGATCACTGCATCTCTTACTTCACC
AACATGTTTGCACCTGGTCAACATCAAAAGCCAGGCTGCACCTCCAGGGAAGAAGCAATGCCTGG
AGACCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACT
GGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCTGATCTCT
TCCTCACAGGATGGCCACCAAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTGTTCCAGGGC
AACCAGGACTCTTTCACACCTGTGGTGAAGTCACTGGACCCCCCTCTGACAAGATACCTGAGA
ATTCACCCCAAGTCTTGGGTCCACCAGATGCCCCTGAGAATGGAAGTCTGGGATGTGAGGCACAA
GACCTGTACTGA (서열 번호: 27)

도면30a

CS23A(760-1667) - CS23-SC1-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGATAC
 TACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCCTGTGGACGOCAGG
 TTCCCCCCCAGAGTGCCCCAAGAGCTTCCCCTTCAACAACCTCAGTGGTGTACAAGAAGACCCTGTTCTGTGGAG
 TTCACCGACCACCTGPTCAACATCGCCAAAGCCAGGCCCCCTGGATGGCCCTGCTGGGCCCCACCATCCAG
 GCGGAGGTGTACGACACCCGTGGTGATCACCTGAAGAACATGGCCAGCCACCCCGTGGAGCTGCACGCCCTG
 GGCTGAGCTACTGGAAGCCCTCTGAGGGCCGCGAGTATGACGAGCAGACCAGCCAGAGGGAGAAGGAGGAC
 GACAAGGTGTTCGCCGGCGGCAGCCACACCTACGTGTGGCAGGTGCTGAAGGAGAACGGGCCCATGGCCAGE
 GAGCCCTGTGCCTGACCTACAGCTACCTGAGCCACCTGGACCTGGTGAAGGACCTGAACCTGTGCCCTGATC
 GCGCCCTGCTGGTGTGACAGGAGGGCAGCCCTGGCCAAAGGAGAGACCCAGACCCTGCACAAGTTCATCCTG
 CTGTTCCGCGTGTTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGACAGCCCTGATGCAGGACAGGGAT
 GCCGCTCTGCCAGGCGCTGGCCAAAGATGCACACCGTGAACCGCTACGTGAACAGGAGCCTGCCCGCCCTG
 ATCGCTGCCACAGGAAGTCTGTGTAAGTGGCACGTGATCGGCATGGGCAACCACCCCGAGGTGCACAGCATC
 TTCTGGAGGGCCACACCTTCTGTTGAGGAACACAGGCAGGCCAGCCCTGGAGATCAGCCCCATCACCTTC
 CTGACCGCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCACATCAGCAGCCAGCCAGCAC
 GACGCCATGGAGCCCTACGTGAAGGTGGACAGCTGCCCCGAGGAGCCCCAGCTGAGGATGAAGAACRACGAG
 GAGGCCGAGGACTATGATGATGACCTGACCCACTCTGAGATGGAGGTGGTGGGTTTGTATGATGACRACAGC
 CCCAGCTTCATCCAGATCAGGTCTGTGGCCAAAGAGCACCOCRAAGACCTGGGTGCACTACATCGCCGCGGAG
 GAGGAGGACTGGGACTACGCCCCCTGTTGCTGGCCCGGACGACAGGAGCTACAAGAGCCAGTACTGAAAC
 AACGCCCCAGAGGATCGGCAGGAAGTACAAGAAGGTACAGATTCATGGCTACAGCCAGGAGACCTCAAG
 ACCAGGGAGGCCATCCAGCAAGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGGAGGTGGGCGACACCCCTG
 CTGATCATCTTCAAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCACGGCATCACCGATGTGAGGCC
 CTGTACAGCAGGAGGCTGCCAAAGGGCTGAAGCACCTGAAGSACTTCCCCATCTGCCCGGCGAGATCTT
 AAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCAGGTGCTGACCAGGTACTAC
 AGCAGCTTCTGTAACATGGAGAGGGACCTGGCTCTGGCTGATGGGCCCTGCTGATCTGCTACRAGGAG
 AGCCTGGACAGAGGGCCAAACAGATCATGTCTGACRAGAGGAACGTGATCCTGTCTCTGTGTTGATGAG
 AACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCCTGCCCAACCCGCGCGGCTGCAGCTGGAGGAC
 CCGGAGTTCAGGCGCAGCAACATCATGCACAGCATCAACGGCTAGCTGTTCGACAGCCTGCAGCTGTCTGTG
 TGCTGACAGAGGTGGCTACTGGTACATCCTGAGCATCGGCGCCAGACCGACTTCCTGTCTGTCTCTTCT
 TCTGGCTACACCTTCAAGCACAAAGATGTTGTACGAGGACACCTGACCTGTTCCCTTTCAGCGGCGGAGACC
 GTGTTCAATGAGCATGGAGAACCOCGGCTGTGGATCTGGGCTGCCACACAGCGACTTCAGGAACAGGGGC
 ATGACCGCCCTGCTGAAAGTCAAGAGCTGCGACAAGAACACCGCGACTACTACGAGGACAGCTACGAGGAC
 ATCAGCGCTACCTGCTGAGCAGAACACAGCCATCGAGCCAGGGAGATCACCGGACACCCCTGCAGAGC
 GACCAGGAGGAGATCGACTATGATGACACCATCAGCGTGGAGATGAAGAGGAGGACTTCGACATCTACGAC
 GAGGACGAGAACCCAGAGCCCGAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATCGCCGCGCTGGAGAG
 CTGTGGGATATGSCATGAGCAGCAGCCCCACCTGCTGAGGAACAGGGCCAGAGCGGCAGCTGCCCCAG
 TTCAAGAAGTGGTGTTCAGGAGTTGACCGAGCGCAGCTTCAACCCAGCCCTGTACAGAGGCGAGCTGAAC
 GAGCACCTGGCCCTGCTGGGCCCTACATCAGGGCCGAGGTGGAGGACACATCATGGTGCACCTTCAGGAAC
 CAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCGAGGGCGCCGAG

(계속)

도면30b

CCCAGGAAGAACTTCGTAAGCCCAACGAGACCAAGACCTACTTCTGGAAGGTGCAGCACACATGGCCCCG
 ACCAAGGACAGTTCGACTGC AAGGCCCTGGGCCCTACTTCTCTGATGTGGACCTGGAGAAGGACGTGCACAGC
 GGCTGATCGGCCCTGCTGGTGTGCCACACCAACACCCCTGAACCCCGCCACGGCAGGCAGGTGACCGTG
 CAGGAGTTCCGCTGTTCCTTCAACATCTTCGACGAGACCAAGAGCTGGTACTTCAACGAGAACATGGAGAGG
 AACTGCAGGGCCCCCTGCACATCCAGATGGAGGACCCACCTTCAAGGAGAACTACAGGTTCCAGGCCATC
 AACGGCTACATCATGGACACCCTGCCCGCCCTGGTGTGGCCAGGACCAAGAGGATCAGGTGGTATCTGCTG
 AGCATGGGCAGCAACGAGAACATCCACAGCATCCACTTCAGCGGCCACCTGTTCCACCGTGGGAAGAAGGAG
 GAGTACAAGATGGCCCTGTACAACCTGTACCCCGCGCTGTTGAGAGCCCTGGAGATGCTGCCAGCAAGGCC
 GGCATCTGGAGGGTGGAGTGCCTGATCGGCGAGCACCTGCACCGCCGGCATGAGCACCCCTGTTCTCTGGTGTAC
 AGCAACAAGTGCAGACCCOCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCCTCTGGC
 CAGTACGGCCAGTGGGCCCCAGCTGGCCAGGCTGC ACTACAGCGGCAGCATCAACGCCCTGGAGCACCAAG
 GAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCAGGCATCAAGACCCAGGGCCGC
 AGGCAGAAGTTCAAGAGCCTGTACATCAGCCAGTTCAATCATATGTACAGCCTGGACGGCAAGAAGTGGCAG
 ACCTACAGGGCAACAGCACCCGACCCCTGATGGTGTCTTCCGCCAAGCTGGACAGCAGCGGCATCAAGCAC
 AACATCTTCAACCCCCCATCATCGCCAGGTACATCAGGCTGCACCCACCCACTACAGCATCAGGAGCACCC
 TGTCCGATGGAACCTGATGGGCTGCGACCTGAACAGCTGCAGCATGCCCTGGGCATGGAGAGCAAGGCCATC
 TCTGACGCCAGATCACCGCCAGCAGCTACTTCACCAACATGTTGCCACCTGGAGCCCCAGCAAGGCCAGG
 CTGCCCTGCAGGGCAGGAGCAACGCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGAC
 TTCAGAGAAGCCATGAAGGTGACCGGCTGACGACCCAGGSCCTGAAGAGCCTGCTGACCCAGCATGTACGTG
 AAGGAGTTCCTGATCAGCAGCAGCCAGGACGGCCAGCAGTGGACCCCTGTTCTTCCAGAACGGCAAGTGAAG
 GTGTTCCAGGGCAACCAGGACAGCTTCAACCCCTGTTGGTGAACAGCCTGGACCCCCCTGCTGACAGGTAT
 CTGAGGATCCACCCCCAGAGCTGGGTGCACCAAGATCGCCCTGAGAATGGAGTCTGGGATGCGAGGCCCCAG
 GACCTGTACTGA (서열 번호: 28)

도면31a

CS23Δ(772-1667) - CS23-SC2-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
 AGATACTAECTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCT
 GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCAGTGGTGTACAAG
 AAGACCTGTTCGTGGAGTTCCACCGACCACCTGTTCAACATCGCCAAGCCCAGGCCCCCCCTGGATG
 GGCTGTCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGTATCACCTGAAGAACATG
 GCCAGCCACCCCGTGAGCCTGCACGCCGTGGGCGTGAGCTACTGGAAGGCTCTGAGGGCGCCGAG
 TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGCAGCCACACC
 TACGTGTGGCAGGTGCTGAAGGAGAACGCCCCCATGGCCAGCGACCCCCTGTGCTGACCTACAGC
 TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGGCCCTGCTGGTGTGC
 AGGGAGGGCAGCCTGGCCAAGGAGAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTCCGCGTG
 TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAGAACAGCCTGATGCAGGACAGGGATGCCGCC
 TCTGCCAGGGCCTGGCCCAAGATGCACACCCTGAACGGCTACGTGAACAGGAGCCTGCCCGGCTG
 ATCGGCTGCCAECAGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCGAGGTGCAC
 AGCATCTTCTGGAGGGCCACACCTTCTGGTGAAGAACACAGGCAGGCCAGCCTGGAGATCAGC
 CCCATCACCTTCTGACCGCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCAC
 ATCAGCAGCCACCAGCACGACCGCATGGAGGCCCTACGTGAAGGTGGACAGCTGCCCGAGGAGCCC
 CAGCTGAGGATGAAGAACAACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
 GACGTGGTGAAGTTTGTATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAG
 CACCCCAAGACCTGGGTGCACTACATCGCCCGCGAGGAGGAGGACTGGGACTACGCCCCCTGGTG
 CTGGCCCCGACGACAGGAGCTACAGAGCCAGTACCTGAACAACGGCCCCCAGAGGATCGGCAGG
 AAGTACAAGAAGGTCAGATTCATGGCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
 CACGAGTCTGGCATCCTGGGCCCCCTGCTGTACGGCGAGGTGGGCGACACCCTGCTGATCATCTTC
 AAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCAACGGCATCACCGATGTGAGGCCCTGTAC
 AGCAGGAGGCTGCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCTCCTGCCCGGCGAGATCTTC
 AAGTACAAGTGGACCGTGAACCGTGGAGGATGGCCCCAACAAGTCTGACCCAGGTGCTGACCAGG
 TACTACAGCAGCTTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCCCTGCTGATC
 TGCTACAAGGAGAGCGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAACGTGATCCTG
 TTCTCTGTGTTTCGATGAGAACAGGAGCTGGTATCTGACCAGAACATCCAGAGGTTCTGCCCAAC
 CCCCGCGGCGTGCAGCTGGAGGACCCCGAGTTCAGGCCAGCAACATCATGCACAGCATCAACGGC
 TACGTGTTGACAGCCTGCAGCTGTCTGTGTGCTGCACGAGGTGGCCTACTGTTACATCCTGAGC
 ATCGGGCCCCAGACCAGCTTCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
 TACGAGGACACCCTGACCCTGTTCCCCTTCAGCGGGCAGACCGTGTTCATGAGCATGGAGAACCCC
 GGCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGGGCATGACCGCCCTGCTGAAA
 GTCAGCAGCTGCGACAAGAACACCGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
 CTGCTGAGCAAGAACAACGCCATCGAGCCCAGGAGCTTCAGCCAGAACTCCAGACACCCAGCAC

(계속)

도면31b

```

AGGGAGATCACCAGGACCACCCCTGCAGAGCCAGGAGGAGATCGACTATGATGACACCATCAGC
GTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCAGGAGCTTC
CAGAAGAAGACCAGGCCTACTTTCATCGCCCGGTGGAGAGGCTGTGGGACTATGGCATGAGCAGC
AGCCCCCAGCTGCTGAGGAACAGGGCCCCAGAGCGGCAGCGTGCCTCAGTTCAAGAAGGTGGTGTTC
CAGGAGTTCACCCGACGGCAGCTTCACCCAGCCCCGTGTACAGAGGCGAGCTGAACGAGCACCTGGGC
CTGCTGGGCCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACCAGGGC
AGCAGGCCCTACAGCTTCTACAGCAGCTGATCAGCTACGAGGAGGACCAGAGGCAGGGCGCCGAG
CCCAGGAAGAACTTCGTGAAGCCCCAAGCAGACCAAGACCTACTTCTGGAGGTTGCAGCACCACATG
GCCCCCACCAGGACGAGTTCGACTGCAGGCCCTGGGCCCTACTTCTCTGATGTGGACCTGGAGAAG
GACGTGCACAGCGGCCCTGATCGGCCCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCCGCCAC
GGCAGGCAGGTGACCGTGCAGGAGTTCGCCCTGTTCTTCCACATCTTGGACGAGACCAAGAGCTGG
TACTTCCACCGAGAACATGGAGAGGAATGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCCCACC
TTCAAGGAGAACTACAGGTTCACAGCCATCAACGGCTACATCATGGACACCCTGCCCGGCCCTGGTG
ATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAACATCCACAGC
ATCCACTTCAGCGGCCAGGTGTTCCCGGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAA
CTGTACCCCGGGCTGTTCGAGACCGTGGAGATGCTGCCCAGCAAGGCCGGCATCTGGAGGGTGGAG
TGCCATGATCGCCGAGCACCTGCACGCCGGCATGAGCACCCCTGTTCTGGTGTACAGCAACAAGTGC
CAGACCCCTCGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCTCTGGCCAGTAC
GGCCAGTGGGCCCCCCAAGCTGGCCAGGCTGCACCTACAGCGGCAGCATCAACGCCTGGAGCACCAG
GAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAGACCCAG
GGCGCCAGGCAGAASTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGACGGC
AAGAAGTGGCAGACCTACAGGGGCAACAGCACCCGGCACCTGATGGTGTTCCTCGGCCAAGTGGAC
AGCAGCGGCATCAAGCACAAACATCTTCAACCCCCCATCATCGCCAGGTACATCAGGCTGCACCCC
ACCCACTACAGCATCAGGAGCACCCCTGCGGATGGAACGATGGGCTGCGACCTGAACAGCTGCAGC
ATGCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCAGATCACCGCCAGCAGCTACTTCACC
AACATGTTTCGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGCGAGGAGCAACGCCTGG
AGGCCCCAGSTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAGGTGACC
GGCGTGACCAACCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTCCTGATCAGC
AGCAGCCAGGACGGCCACCAGTGGACCCTGTTCTTCCAGAACGGCAAAGTGAAGGTGTTCCAGGGC
AACCAGGACAGCTTCACCCCGTGGTGAACAGCCTGGACCCCCCTGCTGACCAGGTATCTGAGG
ATCCACCCCCAGAGCTGGGTGCACCAGATCGCCCTGAGAATGGAAGTCTGGGATGCCAGGCCAG
GACCTGTACTGA (서열 번호: 29)

```

서열목록

SEQUENCE LISTING

<110> BAXALTA INCORPORATED
 BAXALTA GMBH

<120> VIRAL VECTORS ENCODING RECOMBINANT FVIII VARIANTS WITH INCREASED
 EXPRESSION FOR GENE THERAPY OF HEMOPHILIA A

<130> 008073-5115-WO

<140><141><150> 62/255,323

<151> 2015-11-13

<160> 32

<170> PatentIn version 3.5

<210> 1

<211> 4374

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 1

```

atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc      60

accaggagat actacctggg ggctgtggag ctttcttggg actacatgca gtctgacctg      120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaatcctt cccattcaac      180
acctctgtgg tctacaagaa gacctcttt gtggagtca ctgaccacct gttcaacatt      240
gccaaacca ggccacctg gatgggactc ctgggacca ccattcaggc tgagggtgat      300
gacctgtgg tcatcacct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg      360
ggggtcagct actggaaggc ctctgagggg gctgagtatg atgaccagac ctcccagagg      420
gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgtgtg gcaggtcctc      480

aaggagaatg gcccatggc ctctgacca ctctgcctga cctactccta ctttctcat      540
gtggacctgg tcaaggacct caactctgga ctgattgggg cctgctggt gtgcagggag      600
ggctccctgg ccaagagaa gaccagacc ctgcacaagt tcattctcct gtttgctgtc      660
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgatgca ggacagggat      720
gtgcctctg ccaggcctg gcccaagatg cacactgtga atggctatgt gaacaggagc      780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg      840
acaaccctg aggtgcactc ctttctctg gagggccaca ctttctggt caggaaccac      900

agacaggcca gcctggagat cagcccatc accttctca ctgccagac cctgctgatg      960
gacctggac agttctgct gttctgccac atcagctccc accagcatga tggcatggag      1020
gcctatgtca aggtggacag ctgccctgag gagccacagc tcaggatgaa gaacaatgag      1080
gaggctgagg actatgatga tgacctgact gactctgaga tggatgtggt ccgctttgat      1140
gatgacaaca gccatcctt cattcagatc aggtctgtgg ccaagaaaca cccaagacc      1200
tgggtgcact acattgctgc tgaggaggag gactgggact atgccccact ggtcctggcc      1260
cctgatgaca ggagctacaa gagccagtac ctcaacaatg gccacagag gattggacgc      1320

aagtacaaga aagtcaggtt catggcctac actgatgaaa cttcaagac caggagggcc      1380
attcagcatg agtctggcat cctgggcca ctctgtatg gggaggtggg ggacacctg      1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccaca tggcatcact      1500
gatgtcagge cctgtacag ccgaggctg ccaaaggggg tgaaacacct caaggacttc      1560
ccattctgc ctggggagat cttcaagtac aagtggactg tcaactgtga ggatggacca      1620
accaaatctg accccaggtg cctcaccaga tactactcca gctttgtgaa catggagagg      1680

```

gacctggcct ctggcctgat tggcccactg ctcatctgct acaaggagtc tgtggaccag 1740

aggggaacc agatcatgtc tgacaagagg aatgtgattc tgttctctgt ctttgatgag 1800

aacaggagct ggtacctgac tgagaacatt cagcgcttcc tgccaacce tgctggggtg 1860

cagctggagg acctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg 1920

tttgacagcc tccagcttcc tgtctgctg catgaggtgg cctactggta cattctttct 1980

attggggccc agactgactt cttttctgtc ttcttctctg gctacacctt caaacacaag 2040

atggtgtatg aggacacct gaccctcttc ccattctctg gggagactgt gttcatgagc 2100

atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc 2160

atgactgccc tgctcaaagt ctctcctgt gacaagaaca ctggggacta ctataggac 2220

agctatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gcccaggagc 2280

ttcagccaga atccacctgt cctgaaacgc caccagaggg agatcaccag gaccacctc 2340

cagtctgacc aggaggagat tgactatgat gacaccattt ctgtggagat gaagaaagag 2400

gactttgaca tctatgacga ggacgagaac cagagcccaa ggagcttcca gaagaagacc 2460

aggcactact teattgtctg tgtggagcgc ctgtgggact atggcatgag ctccagcccc 2520

catgtcctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggcttcc 2580

caagagtcca ctgatggcag cttcaaccag ccctgtaca gaggggagct gaatgagcac 2640

ctgggactcc tgggcccata catcagggct gaggtggagg acaacatcat ggtgaccttc 2700

cgcaaccagg cctccagccc ctacagcttc tacagctccc tcatcagcta tgaggaggac 2760

cagaggcagg ggctgagcc acgcaagaac ttgtgaaac ccaatgaaac caagacctac 2820

ttctggaaag tccagacca catggcccc accaaggatg agtttgactg caaggcctgg 2880

gcctacttct ctgatgtgga cctggagaag gatgtgact ctggcctgat tggcccactc 2940

ctggtctgcc acaccaaac cctgaacct gcccattgaa ggcaagtgac tgtgcaggag 3000

tttgcctct tctcaccat ctttgatgaa accaagagct ggtacttcc tgagaacatg 3060

gagcgcaact gcagggcccc atgcaacatt cagatggagg accccacctt caaagagaac 3120

taccgcttcc atgcatcaa tggctacatc atggacacct tgcctgggct tgtcatggcc 3180

caggaccaga ggalcaggtg gtacctgctt tctatgggct ccaatgagaa cattcactcc 3240

atccacttct ctggcatgt cttcactgtg cgcaagaagg aggagtaca gatggccctg 3300

tacaacctct acctgggggt ctttgagact gtggagatgc tgcctccaa agctggcatc 3360

tggaggtgg agtgcctcat tggggagcac ctgcatgctg gcatgagcac cctgttctg 3420

gcttacagca acaagtgcc aacccccctg ggaatggcct ctggccacat cagggacttc 3480
 cagatcactg cctctggcca gtatggccag tgggccccca agctggccag gctccactac 3540
 tctggatcca tcaatgcctg gagcaccaag gagccattca gctggatcaa agtggacctg 3600
 ctggccccca tgatcatcca tggcatcaag acccaggggg ccaggcagaa gttctccagc 3660
 ctgtacatca gccattcat catcatgtac agcctggatg gcaagaaatg gcagacctac 3720
 agaggcaact ccaactggaac actcatggtc ttctttggca atgtggacag ctctggcatc 3780
 aagcacaaca tcttcaacct cccaatcatc gccagatata tcaggctgca ccccaccac 3840

tacagcatcc gcagaccct caggatggag ctgatgggct gtgacctgaa ctctgcagc 3900
 atgccctgg gcatggagag caaggccatt tctgatgccc agatcactgc ctccagctac 3960
 ttcaccaaca tgittgccac ctggagccca agcaaggcca ggctgcacct ccaggaagg 4020
 agcaatgctt ggaggcccc ggtcaacaac ccaaaggagt ggctgcaggt ggacttcag 4080
 aagaccatga aggtcactgg ggtgaccacc caggggtca agagcctgct caccagcatg 4140
 tatgtgaagg agttctgat cagctccagc caggatggcc accagtggac cctcttcttc 4200
 cagaatggca aggtcaaggt gttccagggc aaccaggaca gcttcacccc tgttgtgaac 4260

agcctggacc cccccctct gaccagatac ctgaggattc acccccagag ctgggtccac 4320
 cagattgccc ttaggatgga ggtcctggga tgtgaggccc aggacctgta ctga 4374

<210> 2

<211> 1457

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 2

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

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

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335
 Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

 Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365
 Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380
 Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400
 Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

 Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430
 Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445
 Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460
 Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480

 Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495
 His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510
 Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525
 Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540

 Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560
 Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu

Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp
 820 825 830

Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln
 835 840 845

Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr
 850 855 860

Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His
 865 870 875 880

Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile
 885 890 895

Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser
 900 905 910

Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg
 915 920 925

Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val
 930 935 940

Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp
 945 950 955 960

Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu
 965 970 975

Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His
 980 985 990

Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe
 995 1000 1005

Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn
 1010 1015 1020

Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys
 1025 1030 1035

Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr
 1040 1045 1050

Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr

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

Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys
 1295 1300 1305

Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn
 1310 1315 1320

Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln
 1325 1330 1335

Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu
 1340 1345 1350

Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val
 1355 1360 1365

Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys
 1370 1375 1380

Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu
 1385 1390 1395

Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp
 1400 1405 1410

Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr
 1415 1420 1425

Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala
 1430 1435 1440

Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr
 1445 1450 1455

<210> 3

<211> 2220

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 3

gccaccagga gatactacct gggggctgtg gagctttctt gggactacat gcagtctgac 60
 ctgggggagc tgctgtgga tgccagggtc ccaccagag tgcccaaadc cttccattc 120
 aacacctctg tggctctaca gaagaccctc tttgtggagt tcactgacca cctgttcaac 180

attgccaaac ccaggccacc ctggatggga ctctgggac ccaccattca ggctgaggtg 240
 tatgacactg tggatcac cctcaagaac atggcctccc accctgtgag cctgcatgct 300
 gtgggggtca gctactggaa ggccctgtag ggggctgagt atgatgacca gacctcccag 360
 agggagaagg aggatgacaa agtgttcctt gggggcagcc acacctatgt gtggcaggtc 420

 ctcaaggaga atggcccat ggccctgac ccactctgcc tgacctactc ctacctttct 480
 catgtggacc tggcaagga cctcaactct ggactgattg gggccctgct ggtgtgcagg 540
 gagggctccc tggccaaaga gaagaccag accctgcaca agttcattct cctgtttgct 600
 gtctttgatg agggcaagag ctggcactct gaaaccaaga actccctgat gcaggacagg 660
 gatgctgcct ctgccagggc ctggcccaag atgcacactg tgaatggcta tgtgaacagg 720
 agcctgcctg gactcattgg ctgccacagg aaatctgtct actggcatgt gattggcatg 780
 gggacaacce ctgaggtgca ctccattttc ctggagggcc acaccttctt ggtcaggaac 840

 cacagacagg ccagcctgga gatcagcccc atcaccttcc tcaactgcca gacctgctg 900
 atggacctcg gacagttcct gctgttctgc cacatcagct cccaccagca tgatggcatg 960
 gaggcctatg tcaaggtgga cagctgacct gaggagccac agctcaggat gaagaacaat 1020
 gagggaggctg aggactatga tgatgacctg actgactctg agatggatgt ggtccgcttt 1080
 gatgatgaca acagcccatc cttcattcag atcaggtctg tggccaagaa acacccaag 1140
 acctgggtgc actacattgc tgctgaggag gaggactggg actatgcccc actggtcctg 1200
 gcccctgatg acaggagcta caagagccag tacctcaaca atggcccaca gaggattgga 1260

 cgcaagtaca agaaagttag gttcatggcc tacactgatg aaaccttcaa gaccagggag 1320
 gccattcagc atgagtctgg catcctgggc ccactcctgt atggggaggt gggggacacc 1380
 ctgctcatca tttcaagaa ccagccctcc aggcctaca acatctacc acatggcatc 1440
 actgatgta ggcctctgta cagccgagg ctgccaaagg gggatgaaaca cctcaaggac 1500
 ttccccattc tgctgggga gatcttcaag tacaagtga ctgtcactgt ggaggatgga 1560
 ccaaccaaat ctgaccagc gtgcctcacc agatactact ccagctttgt gaacatggag 1620
 agggacctgg cctctggcct gattggccca ctgctcatct gctacaagga gtctgtggac 1680

 cagaggggaa accagatcat gtctgacaag aggaatgtga ttctgttctc tgtctttgat 1740
 gagaacagga gctggtacct gactgagaac attcagcgt tctgccccaa cctgctggg 1800
 gtgcagctgg aggaccctga gttccaggcc agcaacatca tgcactccat caatggctat 1860
 gtgtttgaca gcctccagct ttctgtctgc ctgcatgagg tggcctactg gtacattctt 1920
 tctattgggg cccagactga ctctctttct gtctctttct ctggtctacac ctccaacac 1980
 aagatggtgt atgaggacac cctgacctc ttccattct ctggggagac tgtgttcatg 2040

agcatggaga accctggcct gtggattctg ggatgccaca actctgactt cgcgaacagg 2100

ggcatgactg cctctctcaa agtctcctcc tctgacaaga aacttgggga ctactatgag 2160

gacagctatg aggacatctc tgcctacctg ctgagcaaga acaatgccat tgagcccagg 2220

<210> 4

<211> 2052

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 4

gagatcacca ggaccaccct ccagtctgac caggaggaga ttgactatga tgaccaccatt 60

tctgtggaga tgaagaaaga ggactttgac atctatgacg aggacgagaa ccagagccca 120

aggagcttcc agaagaagac caggcactac ttcattgctg ctgtggagcg cctgtgggac 180

tatggcatga gctccagccc ccatgtctc aggaacaggg cccagtctgg ctctgtgcca 240

cagttcaaga aagtggcttt ccaagagttc actgatggca gcttcacca gccctgtac 300

agaggggagc tgaatgagca cctgggactc ctgggcccac acatcagggc tgaggtggag 360

gacaacatca tggtagacct cgcgaaccag gcctccagge cctacagctt ctacagctcc 420

ctcatcagct atgaggagga ccagaggcag ggggctgagc cacgcaagaa ctttgtgaaa 480

cccaatgaaa ccaagacctt cttctggaaa gtccagcacc acatggcccc caccaaggat 540

gagtttgact gcaaggcctg ggcctacttc tctgatgtgg acctggagaa ggatgtgcac 600

tctggcctga ttggcccact cctggtctgc cacaccaaca ccctgaacce tgcccatgga 660

aggcaagtga ctgtgcagga gtttgcctc ttcttcacca tctttgatga aaccaagagc 720

tggtagctca ctgagaacat ggagcgaac tgcagggcc catgcaacat tcagatggag 780

gacccacct tcaagagaa ctaccgctc catgcccata atggctacat catggacacc 840

ctgcctgggc ttgtcatggc ccaggaccag aggatcaggt ggtacctgct ttctatgggc 900

tccaatgaga acattcactc catccacttc tctgggcatg tcttactgt gcgaagaag 960

gaggagtaca agatggccct gtacaacctc tacctgggg tctttgagac tgtggagatg 1020

ctgcctcca aagctggcat ctggagggig gactgcctca ttggggagca cctgcatgct 1080

ggcatgagca cctgttctt ggtctacagc aacaagtcc agacccccct gggaatggcc 1140

tctggccaca tcagggactt ccagatcact gcctctggcc agtatggcca gtgggcccc 1200

aagctggcca ggctccacta ctctggatcc atcaatgcct ggagaccaa ggagccattc 1260

agctggatca aagtggacct gctggccccc atgatcatcc atggcatcaa gaccagggg 1320
 gccaggcaga agttctccag cctgtacatc agccagtca tcatcatgta cagcctggat 1380
 ggcaagaaat ggcagaccta cagaggcaac tccactggaa cactcatggt cttctttggc 1440

aatgtggaca gctctggcat caagcacaac atcttcaacc cccaatcat cgccagatac 1500
 atcaggctgc accccaccca ctacagcatc cgcagcacc tcaggatgga gctgatgggc 1560
 tgtgacctga actcctgcag catgcccttg ggcatggaga gcaaggccat ttctgatgcc 1620
 cagatcactg cctccagcta cttaccaac atgtttgcca cctggagccc aagcaaggcc 1680
 aggctgcacc tccaggggaag gagcaatgcc tggaggcccc aggtcaaca cccaaaggag 1740
 tggctgcagg tggacttcca gaagacatg aaggctactg gggtgaccac ccagggggtc 1800
 aagagcctgc tcaccagcat gtatgtgaag gatttcctga tcagctccag ccaggatggc 1860

caccagtgga cctcttctt ccagaatggc aaggtaagg tgttccagg caaccaggac 1920
 agcttcacc ctgtggtgaa cagcctggac cccccctcc tgaccagata cctgaggatt 1980
 cacccccaga gctgggtcca ccagattgcc ctgaggatgg aggtcctggg atgtgaggcc 2040
 caggacctgt ac 2052

<210> 5

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 5

agcttctctc agaatccacc tgtctgaag agacaccaga ga 42

<210> 6

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 6

agcttcagcc agaatccacc tgtctgaaa cgccaccaga gg 42

<210> 7

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 7

agcttcagcc agaaccccc cgtgctgaag aggcaccaga gg 42

<210> 8

<211> 7827

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 8

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatgatcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagcgcgc 240

attcgccatt caggctgctc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300

tacgccagct ggcgaaaggg ggatgtgctg caagcgcatt aagtgggta acgccagggt 360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cctcgagatt taaatgacgt 420

tggccactcc ctctctgctc gctcctcgc tctactgaggc cgggcgacca aaggtcgccc 480

gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgcaga gagggagtgg 540

ccaactccat cactaggggt tcttgagttt aaacttcgctc gacgattcga gcttgggctg 600

caggtcgagg gactgggag gatgttgagt aagatgaaa actactgatg acccttcgag 660

agacagagta ttaggacatg tttgaacagg ggccgggcca tcagcaggta gctctagagg 720

atccccgtct gtctgcacat ttcgtagagc gagtgttccg atactctaata ctccctaggc 780

aaggttcata tttgtgtagg ttacttattc tccttttgtt gactaagtca ataatacagaa 840

tcagcagggt tggagtgcgc ttggcaggga tcagcagcct gggttggaa gagggggtat 900

aaaagcccct tcaccaggag aagccgtcac acagactagg cgcgccaccg ccaccatgca 960

gattgagctg agcacctgct tcttctctgt cctgctgagg ttctgcttct ctgccaccag 1020

gagatactac ctgggggctg tggagctttc ttgggactac atgcagtctg acctggggga 1080

gctgcctgtg gatgccaggt tcccaccag agtgccc aaa tccttccat tcaacacctc 1140

tgtggictac aagaagacc tctttgtgga gttcactgac cacctgttca acattgccaa 1200

 acccaggcca cctggatgg gactcctggg acccaccatt caggctgagg tgtatgacac 1260
 tgtggtcate accctcaaga acatggcctc ccaccctgtg agcctgcatg ctgtgggggt 1320
 cagctactgg aaggcctctg agggggctga gtatgatgac cagacctccc agagggagaa 1380
 ggaggatgac aaagtgttcc ctgggggcag ccacacctat gtgtggcagg tcctcaagga 1440
 gaatggcccc atggcctctg acccactctg cctgacctac tcctaccttt ctcattgtgga 1500
 cctggccaag gacctcaact ctggactgat tggggccctg ctggtgtgca gggagggctc 1560
 cctggccaaa gagaagacc agacctgca caagttcatt ctctgtttg ctgtctttga 1620

 tgaggcaag agctggcact ctgaaaccaa gaactccctg atgcaggaca gggatgctgc 1680
 ctctgccagg gcctggccca agatgcacac tftgaatggc tatftgaaca ggacctgcc 1740
 tggactcatt ggetgccaca ggaaatctgt ctactggcat gtgattggca tggggacaac 1800
 ccctgaggtg cactccattt tcctggaggg ccacaccttc ctggtcagga accacagaca 1860
 ggccagcctg gagatcagcc ccatcacctt cctcactgcc cagacctgc tgatggacct 1920
 cggacagtte ctgtgttct gccacatcag ctcccaccag catgatggca tggaggccta 1980
 tgtcaaggtg gacagctgcc ctgaggagcc acagctcagg atgaagaaca atgaggaggc 2040

 tgaggactat gatgatgacc tgactgactc tgagatggat gtggtccgct ttgatgatga 2100
 caacagccca tccttcattc agatcaggtc tftggccaag aacacccca agacctgggt 2160
 gcactacatt gctgctgagg aggaggactg ggactatgcc ccactggctc tggcccctga 2220
 tgacaggagc tacaagagcc agtacctcaa caatggccca cagaggattg gacgcaagta 2280
 caagaaagtc aggttcattg cctacactga tgaaccttc aagaccaggg aggccattca 2340
 gcatgagtct ggcatcctgg gccactcct gtatggggag ftgggggaca ccctgctcat 2400
 catcttcaag aaccaggcct ccaggcccta caacatctac ccacatggca tcaactgatgt 2460

 caggcccctg tacagccgca ggctgcaaaa gggggtgaaa cacctcaagg acttcccat 2520
 tctgcctggg gagatcttca agtacaagtg gactgtcact gtggaggatg gaccaaccaa 2580
 atctgacccc aggtgcctca ccagatacta ctccagcttt gtgaacatgg agagggacct 2640
 ggctctggc ctgattggcc cactgctcat ctgctacaag gactctgtgg accagagggg 2700
 aaaccagatc atgtctgaca agaggaatgt gattctgttc tctgtctttg atgagaacag 2760
 gagctggtac ctgactgaga acattcagcg cttcctgccc aacctgctg gggatgacct 2820
 ggaggacctg gattccagg ccagcaacat catgactcc atcaatggct atgtgtttga 2880

cagcctccag ctttctgtct gcctgcatga ggtggcctac tggtagattc tttctattgg 2940
 ggcccagact gacttccttt ctgtcttctt ctctggctac accttcaaac acaagatggt 3000
 gtatgaggac accctgacct tcttccatt ctctggggag actgtgttca tgagcatgga 3060
 gaacctggc ctgtggattc tgggatgcca caactctgac ttccgcaaca ggggcatgac 3120
 tgccctgctc aaagtctctt cctgtgacaa gaacctggg gactactatg aggacagcta 3180
 tgaggacatc tctgcctacc tgctcagcaa gaacaatgcc attgagccca ggagcttcag 3240
 ccagaatcca cctgtcctga aacgccacca gagggagatc accaggacca ccctccagtc 3300

 tgaccaggag gagattgact atgatgacac ctttctgtg gagatgaaga aagaggactt 3360
 tgacatctat gacgaggacg agaaccagag cccaaggagc ttccagaaga agaccaggca 3420
 ctacttcatt gctgctgtgg agcgctgtg ggactatggc atgagctcca gccccatgt 3480
 cctcaggaac agggcccagt ctggtctgtt gccacagttc aagaaagtgg tcttccaaga 3540
 gttactgat ggcagcttca cccagcccct gtacagaggg gagctgaatg agcacctggg 3600
 actcctgggc ccatacatca gggctgaggt ggaggacaac atcatggtga ccttccgcaa 3660
 ccaggcctec aggccttaca gcttctacag ctccctcacc agctatgagg aggaccagag 3720

 gcagggggct gagccacgca agaactttgt gaaaccaat gaaaccaaga cctacttctg 3780
 gaaagtccag caccacatgg cccccacaa ggatgagttt gactgcaagg cctgggccta 3840
 cttctctgat gtggacctgg agaaggatgt gcaactctggc ctgattggcc cactcctggt 3900
 ctgccacacc aacaccctga acctgcccc tggaaaggcaa gtgactgtgc aggagtttgc 3960
 cctcttcttc accatctttg atgaaaccaa gagctggtac ttactgaga acatggagcg 4020
 caactgcagg gcccctgca acattcagat ggaggacccc accttcaaag agaactaccg 4080
 cttccatgcc atcaatggct acatcatgga caccctgctt gggcttgtca tggcccagga 4140

 ccagaggate aggtggtacc tgctttctat gggctccaat gagaacattc actccatcca 4200
 cttctctggg catgtcttca ctgtgcgcaa gaaggaggag tacaagatgg cctgtlaca 4260
 cctctacctt ggggtctttg agactgtgga gatgctgccc tccaaagtct gcactctggag 4320
 ggtggagtgc ctattgggg agcacctgca tgctggcatg agcacctgt tectggtcta 4380
 cagcaacaag tgccagacct ccctgggaat ggctctggc cacatcaggg acttccagat 4440
 cactgctctt ggccagtatg gccagtgggc ccccaagtct gccaggctcc actactctgg 4500
 atccatcaat gcctggagca ccaaggagcc attcagctgg atcaaagtgg acctgctggc 4560

 ccccatgate atccatggca tcaagacca gggggccagg cagaagtctt ccagcctgta 4620
 catcagccag ttcatcatca tgtacagcct ggatggcaag aatggcaga cctacagagg 4680
 caactccact ggaacctca tggcttctt tggcaatgtg gacagctctg gcatcaagca 4740

caacatcttc aacccccca tcatcgccag atacatcagg ctgcaccca cccactacag 4800
catccgcagc accctcagga tggagctgat gggctgtgac ctgaactcct gcagcatgcc 4860
cctgggcatg gagagcaagg ccattttctga tgcccagatc actgcctcca gctacttcac 4920
caacatgttt gccacctgga gcccaagcaa ggccaggctg cacctccagg gaaggagcaa 4980

tgcttggagg ccccaggtca acaacccaaa ggagtggctg caggaggact tccagaagac 5040
catgaaggte actggggtga ccaccagggt ggtaagagc ctgctcacca gcatgtatgt 5100
gaaggagttc ctgatcagct ccagccagga tggccaccag tggaccctct tcttcagaa 5160
tggcaaggte aaggtgttcc agggcaacca ggacagcttc acccctgtgg tgaacagcct 5220
ggaccccccc ctctgacca gataactgag gattcacccc cagagctggg tccaccagat 5280
tgcctgagg atggaggtcc tgggatgtga ggcccaggac ctgtactgat gacgagcggc 5340
cgctcttagt agcagtatcg ataataaaag atctttatct tcattagatc tgtgtgttgg 5400

ttttttgtgt gtttaattaag ctcgcaagg aaccctagt gatggagtg gccactccct 5460
ctctgcgcgc tcgctcgctc actgaggccg ggcgacaaa ggtcgcccga cgccccggct 5520
ttgccggggt ggctcagtg agcgagcgag cgcgagaga gggagtggcc aagacgattt 5580
aatgacaag cttagcgtaa tcatggtcat agctgtttcc tgtgtgaaat tttatccgc 5640
tcacaattcc acacaacata cgagccgga gcataaagt taaagcctgg ggtgcctaat 5700
gagtgagcta actcacatta attgcgttgc gctcactgcc cgtttccag tgggaaacc 5760
tgtcgtgcca gctgcattaa tgaatcgcc aacgcgagg gagaggcgt ttgcgtattg 5820

ggcgtcttc cgttctctg ctcactgact cgtgcgctc ggtcgttcgg ctgcggcgag 5880
cggatcagc tactcaaag gcggaatac ggttatccac agaatcagg gataacgcag 5940
gaaagaacat gtgagcaaaa ggccagcaaa agccagga cctgaaaaag gccgcgttgc 6000
tggcgttttt ccatagctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 6060
agagggtggc aaaccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 6120
tcgtgcgctc tctgttccg accctgccgc ttaccgata cctgtccgc tttctcctt 6180
cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg 6240

ttcgtccaa gctgggctgt gtgcacgaac cccccgtca gcccgaccgc tgcgccttat 6300
ccgtaacta tegtcttag tccaaccgg taagacacga cttatcgcca ctggcagcag 6360
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 6420
ggtggcctaa ctacggctac actagaagaa cagtatttgg tatctgcgct ctgctgaagc 6480
cagttacctt cgaaaaaaga gttggttagc cttagtcgg caaacaacc accgctggta 6540
gcggtggttt tttgtttgc aagcagcaga ttacgcgag aaaaaagga tctcaagaag 6600

atcctttgat cttttctacg gggctctgacg ctcagtggaa cgaaaactca cgtaaagga 6660

ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa 6720

gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa 6780

tcagtgaggc acctatctca gcgatctgic tatttcgttc atccatagtt gcctgactcc 6840

ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 6900

taccgcgaga cccacgctca ccggetccag atttatcagc aataaaccag ccageccgaa 6960

gggccgagcg cagaagtggc cctgcaactt tatccgctc catccagtct attaattgtt 7020

gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 7080

ctacaggcat cgigggtgca cgctcgtcgt ttggatggc ttcattcagc tccggttccc 7140

aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agctccttcg 7200

gtcctccgat cgttgtcaga agtaagttgg ccgcagtggt atcaactcatg gttatggcag 7260

cactgcataa ttctcttact gtcatgcat ccgtaagatg cttttctgtg actggtgagt 7320

actcaacca gtcattctga gaatagtga tgcggcgacc gagttgctct tgccccgct 7380

caatacggga taataccgcg ccacatagca gaactttaaa agtgcctcacc attggaaaac 7440

gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 7500

ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 7560

caaaaacagg aaggcaaaat gccgcaaaa agggaataag ggcgacacgg aatgttgaa 7620

tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 7680

gcggatacat atttgaatgt atttagaaaa ataacaat aggggttccg cgcacatttc 7740

ccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa 7800

ataggcgtat cacgaggccc tttcgtc 7827

<210> 9

<211> 4332

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 9

atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc 60

accaggagat actacctggg ggctgtggag ctttcttggg actacatgca gctgacctg 120

ggggagctgc ctgtggatgc caggttccca ccagagtgc ccaaatcctt cccattcaac 180

acctctgtgg tctacaagaa gacctctttt gtggagtca ctgaccacct gttcaacatt 240
 gccaaacca ggccacctg gatgggactc ctgggacca ccattcaggc tgaggtgat 300
 gacactgtgg tcatcacct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg 360

 ggggtcagct aciggaaggc ctctgagggg gctgagtatg atgaccagac ctcccagagg 420
 gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgtgtg gcaggtcctc 480
 aaggagaatg gccccatggc ctctgacceca ctctgcctga cctactccta cettttctcat 540
 gtggacctgg tcaaggacct caactctgga ctgattgggg ccctgctggt gtgcaggagg 600
 ggctccctgg ccaaagagaa gaccagacc ctgcacaagt tcattctcct gtttctgtc 660
 tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgatgca ggacagggat 720
 gctgcctctg ccagggcctg gcccaagatg cacactgtga atggctatgt gaacaggagc 780

 ctgcctggac teattgctg ccacaggaaa tetgtctact ggcatgtgat tggcatgggg 840
 acaaccctg aggtgcactc cttttctctg gagggccaca ccttctggt caggaaccac 900
 agacaggcca gcctggagat cagcccatc accttctca ctgcccagac cctgctgatg 960
 gacctcgac agttctctg gttctgceac atcagctccc accagcatga tggcatggag 1020
 gcctatgtca aggtggacag ctgacctgag gagccacagc tcaggatgaa gaacaatgag 1080
 gaggtctgagg actatgatga tgacctgact gactctgaga tggatgtggt ccgctttgat 1140
 gatgacaaca gcccatcctt cattcagatc aggtctgtgg ccaagaaaca cccaagacc 1200

 tgggtgcaact acattgctgc tgaggaggag gactgggact atgccccact ggtcctggcc 1260
 cctgatgaca ggagctacaa gagccagtac ctcaacaatg gccacagag gattggacgc 1320
 aagtacaaga aagtcaggtt catggcctac actgatgaaa ccttcaagac cagggaggcc 1380
 attcagcatg agtctggcat cctgggcca ctctgtatg gggaggtggg ggacaccctg 1440
 ctcatcatct tcaagaacca ggctccagg ccctacaaca tctaccaca tggcatcact 1500
 gatgtcaggc ccctgtacag ccgcaggctg ccaaaggggg tgaaacacct caaggacttc 1560
 cccattctgc ctggggagat cttcaagtac aagtggactg tcactgtgga ggatggacca 1620

 accaaatctg accccaggtg cctcaccaga tactactcca gctttgtgaa catggagagg 1680
 gacctggcct ctggcctgat tggccactg ctcatctgct acaaggagt cgtggaccag 1740
 aggggaaacc agatcatgct tgacaagagg aatgtgattc tgttctctgt ctttgatgag 1800
 aacaggagct ggiacctgac tgagaacatt cagcgcttcc tgccaaccc tgctggggtg 1860
 cagctggagg accctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg 1920
 tttgacagcc tccagcttcc tgtctgcctg catgaggtgg cctactggta cattcttct 1980

attggggccc agactgactt cctttctgtc ttcttctctg gctacacctt caaacacaag 2040

 atggtgtatg aggacacctt gacctcttc ccattctctg gggagactgt gttcatgagc 2100
 atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc 2160
 atgactgccc tgctcaaagt ctctctctgt gacaagaaca ctggggacta ctatgaggac 2220
 agctatgagg acatctctgc ctacctgtc agcaagaaca atgccattga gcccagggag 2280
 atcaccagga ccacctcca gtctgaccag gaggagattg actatgatga caccatttct 2340
 gtggagatga agaaagagga ctttgacatc tatgacgagg acgagaacca gagcccaagg 2400
 agcttccaga agaagaccag gcaactctc attgctgctg tggagcgctt gtgggactat 2460

 ggcatgagct ccagcccca tgtctcagg aacagggccc agtctggctc tgtgccacag 2520
 ttcaagaaag tggcttcca agagttcact gatggcagct tcaccagcc cctgtacaga 2580
 ggggagctga atgagcacct gggactcctg ggccataca tcagggctga ggtggaggac 2640
 aacatcatgg tgaccttccg caaccaggcc tccaggcctt acagcttcta cagctcctc 2700
 atcagctatg aggaggacca gaggcagggg gctgagccac gcaagaactt tgtgaaacce 2760
 aatgaaacca agacctactt ctggaaagtc cagcaccaca tggccccac caaggatgag 2820
 tttgactgca aggcctgggc ctacttctct gatgtggacc tggagaagga tgtgactct 2880

 ggctgattg gccactcct ggtctgccac accaacacce tgaacctgc ccatggaagg 2940
 caagtgactg tgcaggagt ttccctcttc ttaccatct ttgatgaaac caagagctgg 3000
 tacttactg agaacatgga gcgcaactgc agggcccat gcaacattca gatggaggac 3060
 cccacctta aagagaacta ccgcttccat gccatcaatg gctacatcat ggacacctg 3120
 cctgggcttg tcatggccca ggaccagagg atcaggtggt acctgcttc tatgggctcc 3180
 aatgagaaca ttactccat ccacttctct gggcatgtct tcactgtgcg caagaaggag 3240
 gagtacaaga tggccctgta caacctctac cctggggtct ttgagactgt ggagatgctg 3300

 ccctcaaag ctggcatctg gaggtggag tgctcattg gggagcacct gcatgctggc 3360
 atgagcacc tgttcttgg ctacagcaac aagtgccaga cccccctggg aatggcctct 3420
 ggccacatca gggacttcca gatcactgcc tctggccagt atggccagt ggcccccaag 3480
 ctggccaggc tccactactc tggatccatc aatgcctgga gcaccaagga gccattcagc 3540
 tggatcaaag tggacctgtt ggccccatg atcatccatg gcatcaagac ccagggggcc 3600
 aggcagaagt tctccagct gtacatcagc cagttcatca tcatgtacag cctggatggc 3660
 aagaaatggc agacctacag aggcaactcc actggaacac tcatggtctt ctttgcaat 3720

gtggacagct ctggcatcaa gcacaacatc ttcaaccccc caatcatcgc cagatacatc 3780
 aggctgcacc ccaccacta cagcatccgc agcacctca ggatggagct gatgggctgt 3840
 gacctgaact cctgcagcat gccctgggc atggagagca aggccatttc tgatgccag 3900
 atcactgcct ccagctactt caccaacatg ttgccacct ggagcccaag caaggccagg 3960
 ctgcacctcc aggaaggag caatgcctgg aggccccagg tcaacaacc aaaggagtgg 4020
 ctgcaggtgg acttccagaa gaccatgaag gtcaactggg tgaccacca ggggtcaag 4080
 agcctgctca ccagcatgta tgtgaaggag ttctgatca gctccagcca ggatggccac 4140

cagtgacc tcttctcca gaatggcaag gtcaaggtgt tccaggcaa ccaggacagc 4200
 ttcaccctg tgggaacag cctggacccc ccctcctga ccagatacct gaggattcac 4260
 ccccagagct gggccacca gattgccctg aggatggagg tcctgggatg tgaggcccag 4320
 gacctgtact ga 4332

<210> 10

<211> 1443

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 10

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

1 5 10 15
 Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser

20 25 30
 Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg

35 40 45
 Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val

50 55 60
 Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile

65 70 75 80
 Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln

85 90 95
 Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser

100 105 110

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

850 855 860
 Glu His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp
 865 870 875 880
 Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe
 885 890 895
 Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu

 900 905 910
 Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp
 915 920 925
 Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys
 930 935 940
 Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser
 945 950 955 960
 Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro

 965 970 975
 Ala His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr
 980 985 990
 Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg
 995 1000 1005
 Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe
 1010 1015 1020
 Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp

 1025 1030 1035
 Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp
 1040 1045 1050
 Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His
 1055 1060 1065
 Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys
 1070 1075 1080
 Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu

 1085 1090 1095

Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile
 1100 1105 1110

Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr
 1115 1120 1125

Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile
 1130 1135 1140

Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala
 1145 1150 1155

Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp
 1160 1165 1170

Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala
 1175 1180 1185

Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys
 1190 1195 1200

Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu
 1205 1210 1215

Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr
 1220 1225 1230

Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His
 1235 1240 1245

Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His
 1250 1255 1260

Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met
 1265 1270 1275

Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser
 1280 1285 1290

Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr
 1295 1300 1305

Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu
 1310 1315 1320

Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys

ggctccctgg ccaaagagaa gaccagacc ctgcacaagt tcattctcct gtttgctgtc 660
 ttgatgagg gcaagagctg gcactctgaa accaagaact ccctgatgca ggacagggat 720
 gctgcctctg ccagggcctg gccaagatg cacactgtga atggctatgt gaacaggagc 780

 ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg 840
 acaaccctg aggtgcactc cttttctctg gagggccaca ctttctggt caggaaccac 900
 agacaggcca gcctggagat cagecccatc accttctca ctgcccagac cctgctgatg 960
 gacctcggac agttcctgct gttctgccac atcagctccc accagcatga tggcatggag 1020
 gcctatgta aggtggacag ctgcccctgag gagccacagc tcaggatgaa gaacaatgag 1080
 gaggtgagg aciatgatga tgacctgact gactctgaga tggatgtggt ccgctttgat 1140
 gatgacaaca gcccatcctt cattcagatc aggtctgtgg ccaagaaaca cccaagacc 1200

 tgggtgact acattgtctg tgaggaggag gactgggact atgccccact ggtcctggcc 1260
 cctgatgaca ggagctacaa gagccagtac ctcaacaatg gccacagag gattggacgc 1320
 aagtacaaga aagtcaggtt catggcctac actgatgaaa cttcaagac cagggaggcc 1380
 attcagcatg agtctggcat cctgggceca ctctgtatg gggaggtggg ggacaccctg 1440
 ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccaca tggcatcact 1500
 gatgtcagge cctgtacag ccgcaggtg ccaaaggggg tgaaacacct caaggacttc 1560
 cccattctgc ctggggagat cttcaagtac aagtggactg tcaactgtga ggatggacca 1620

 accaaatctg accccaggtg cctcaccaga tactactcca gctttgtgaa catggagagg 1680
 gacctggcct ctggcctgat tggcccactg ctcatctgct acaaggagtc tgtggaccag 1740
 aggggaaacc agatcatgtc tgacaagagg aatgtgattc tgttctctgt ctttgatgag 1800
 aacaggagct ggtacctgac tgagaacatt cagcgttcc tgccaacce tgctggggtg 1860
 cagctggagg acctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg 1920
 ttgacagcc tcagcttcc tgtctgctg catgaggtgg cctactggta cattcttct 1980
 attggggccc agactgactt ctttctgctc ttcttctctg gctacacctt caaacacaag 2040

 atggtgtatg aggacacct gaccctcttc ccattctctg gggagactgt gttcatgagc 2100
 atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc 2160
 atgactgccc tgetcaaagt ctctcctgt gacaagaaca ctggggacta ctataggac 2220
 agctatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gccaggagc 2280
 ttcagccaga attccagaca cccagcacc agggagatca ccaggaccac cctccagtct 2340
 gaccaggagg agattgacta tgatgacacc atttctgtgg agatgaagaa agaggacttt 2400

gacatctatg acgaggacga gaaccagagc ccaaggagct tccagaagaa gaccaggcac 2460

tacttcattg ctgctgtgga gcgcctgtgg gactatggca tgagctccag ccccatgtc 2520
ctcaggaaca gggcccagtc tggctctgtg ccacagttca agaaagtggc cttccaagag 2580
ttcaactgatg gcagcttcac ccagcccctg tacagagggg agctgaatga gcacctggga 2640
ctcctgggcc catacatcag ggctgaggtg gaggacaaca tcatggtgac cttccgcaac 2700
caggcctcca ggcctacag cttctacagc tcctcatca gctatgagga ggaccagagg 2760
cagggggctg agccacgcaa gaactttgtg aaaccaatg aaaccaagac ctacttctgg 2820
aaagtccagc accacatggc cccaccaag gatgagttg actgcaaggc ctgggcctac 2880

ttctctgatg tggacctgga gaaggatgtg cactctggcc tgattggccc actcctggtc 2940
tgccacacca acacctgaa cctgcccac ggaaggcaag tgactgtgca ggagtttgcc 3000
ctcttcttca ccatctttga tgaaccaag agctggact tcactgagaa catggagcgc 3060
aactgcaggc ccccatgcaa cattcagatg gaggaccca cttcaaaga gaactaccgc 3120
ttccatgcca tcaatggcta catcatggac accctgctg ggcttgtcat ggcccaggac 3180
cagaggatca ggtggtacct gctttctatg gctccaatg agaacattca ctccatccac 3240
ttctctgggc atgtcttca tgtgcgaag aaggaggagt acaagatggc cctgtacaac 3300

ctctaccctg gggctcttga gactgtggag atgctgcct ccaaagctgg catctggagg 3360
gtggagtgcc tcattgggga gcacctgcat gctggcatga gcacctgtt cctggictac 3420
agcaacaagt gccagacccc cctgggaatg gcctctggcc acatcagga cttccagatc 3480
actgcctctg gccagtatgg ccagtgggcc cccaagctgg ccaggctcca ctactctgga 3540
tccatcaatg cctggagcac caaggagcca ttcagctgga tcaaagtgga cctgctggcc 3600
cccatgatca tccatggcat caagaccag ggggccaggc agaagtctc cagcctgtac 3660
atcagccagt tcatcatcat gtacagcctg gatggcaaga aatggcagac ctacagaggc 3720

aactccactg gaacactcat ggtcttctt ggcaatgtgg acagctctgg catcaagcac 3780
aacatcttca acccccaat catgccaga tacatcagc tgcacccac ccaactacagc 3840
atccgcagca cctcaggat ggagctgatg ggctgtgacc tgaactcctg cagcatgccc 3900
ctgggcatgg agagcaaggc ctttctgat gccagatca ctgcctccag ctacttacc 3960
aacatgtttg ccacctggag cccaagcaag gccaggctgc acctccaggg aaggagcaat 4020
gcctggaggc cccaggtcaa caaccxaaag gactggctgc aggtggactt ccagaagacc 4080
atgaaggtca ctggggtgac caccagggg gtcaagagcc tgctcaccag catgtatgtg 4140

aaggagttcc tgatcagctc cagccaggat ggccaccagt ggacctctt cttccagaat 4200
 ggcaaggtca aggtgttcca gggcaaccag gacagcttca cccctgtggt gaacagcctg 4260
 gaccccccc tctgaccag atacctgagg attcaccccc agagctgggt ccaccagatt 4320
 gcctgagga tggaggtcct gggatgtgag gcccaggacc tgtactga 4368

<210> 12

<211> 1455

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 12

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

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

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

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

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

Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn
 915 920 925
 Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His
 930 935 940
 His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr
 945 950 955 960
 Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly
 965 970 975
 Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg
 980 985 990
 Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu
 995 1000 1005
 Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg
 1010 1015 1020
 Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn
 1025 1030 1035
 Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro
 1040 1045 1050
 Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu
 1055 1060 1065
 Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly
 1070 1075 1080
 His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu
 1085 1090 1095
 Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro
 1100 1105 1110
 Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His
 1115 1120 1125
 Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys
 1130 1135 1140
 Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe

1145	1150	1155
Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu		
1160	1165	1170
Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys		
1175	1180	1185
Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile		
1190	1195	1200
Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser		
1205	1210	1215
Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys		
1220	1225	1230
Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val		
1235	1240	1245
Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe		
1250	1255	1260
Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His		
1265	1270	1275
Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp		
1280	1285	1290
Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile		
1295	1300	1305
Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe		
1310	1315	1320
Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg		
1325	1330	1335
Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu		
1340	1345	1350
Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr		
1355	1360	1365
Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe		
1370	1375	1380

Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe

1385 1390 1395

Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe

1400 1405 1410

Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr

1415 1420 1425

Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg

1430 1435 1440

Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr

1445 1450 1455

<210> 13

<211> 4374

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 13

```

atgcagattg agctgtccac ctgcttcttt ctgtgcctgc tgagattctg cttctctgcc      60
accaggagat actacctggg ggctgtggaa ctttcttggg actacatgca gtctgacctg      120
ggagagctgc ctgtggatgc caggttccca cccagagtgc ccaagtcctt cccattcaac      180
acctctgtgg tctacaagaa gacctcttt gtggaattca ctgaccacct gttcaacatt      240

gcaaaaccca gaccacctg gatgggactc ctgggacca ccattcaggc tgagggtgat      300
gacctgtgg tcatcacct caagaacatg gcatcccacc ctgtgtctct gcatgctgtg      360
ggagtctcat actggaaagc ctctgaaggg gctgagtatg atgaccagac atcccagaga      420
gagaaagagg atgacaaggt gtteccctggg ggatctcaca cctatgtgtg gcaagtcctc      480
aaggagaatg gacccatggc atctgacca cctgcctga catactccta ctttctcat      540
gtggacctgg tcaaggacct caactctgga ctgattgggg cactgctggt gtgcagggaa      600
ggatccctgg ccaaggagaa aaccagaca ctgcacaagt tcattctcct gtttgetgtc      660

tttgatgagg gcaagtcttg gcactctgaa acaaagaact ccctgatgca agacagggat      720
gtgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca      780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg      840

```

acaaccctg aagtgcactc cttttcctg gagggacaca ctttctggt caggaaccac 900
 agacaagcct ctctggagat ctctccatc accttctca ctgcacagac actgctgatg 960
 gacctggac agttctgct gttctgccac atctcttccc accagcatga tggcatggaa 1020
 gcctatgtca aggtggactc atgccctgag gaaccacagc tcaggatgaa gaacaatgag 1080

gaggctgagg actatgatga tgacctgact gactctgaga tggatgtggt cagatttgat 1140
 gatgacaact ctccatcctt cattcagatc aggtctgtgg caaagaaaca cccaagaca 1200
 tgggtgact acattgctgc tgaggaagag gactgggact atgcaccact ggtcctggcc 1260
 cctgatgaca ggagctaca gtctcagtac ctcaacaatg gcccaaaag aattggaaga 1320
 aagtacaaga aagttagatt catggcctac actgatgaaa cttcaagac aagagaagcc 1380
 attcagcatg agtctggcat tctgggacca ctctgtatg gggaagtggg agacaccctg 1440
 ctcatcatc tcaagaacca ggctccagg ccctacaaca tctaccaca tggcatcact 1500

gatgtcaggc ccctgtacag caggagactg caaaagggg tgaacacct caaggacttc 1560
 cccattctgc ctggagagat cttcaagtac aagtggactg tcactgtgga ggatggacca 1620
 acaaagtctg accccaggtg cctcaccaga tactactcct cttttgtgaa catggagaga 1680
 gacctggcat ctggactgat tggaccactg ctcatctgct acaaggagtc tgtggaccag 1740
 agaggcaacc agatcatgct tgacaagaga aatgtgattc tgttctctgt ctttgatgag 1800
 aacagatcat ggtacctgac tgagaacatt cagagattcc tgccaaccc tgctggggtg 1860
 caactggaag accctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgtg 1920

tttgactctc tccagcttc tgtctgctg catgaggtgg cctactggtc cattcttct 1980
 attggggcac aaactgactt ctttctgct ttcttctctg gatacacctt caagcacaag 2040
 atggtgatg aggacacct gacactctc ccattctctg gggaaactgt gttcatgagc 2100
 atggagaacc ctggactgtg gattctggga tgccacaact ctgacttcag aaacagggga 2160
 atgactgcac tgctcaaagt ctctctctg gacaagaaca ctggggacta ctatgaggac 2220
 tcttatgagg acatctctgc ctacctctc agcaagaaca atgccattga gcccagaagc 2280
 ttctctcaga atccacctg cctgaagaga caccagagag agatcaccag gacaaccctc 2340

cagtctgacc aggaagagat tgactatgat gaccacattt ctgtggagat gaagaaggag 2400
 gactttgaca tctatgatga ggacgagaac cagtctccaa gatcattcca gaagaagaca 2460
 agacactact tcattgctgc tgtgaaaga ctgtgggact atggcatgct ttcctctccc 2520
 catgtcctca ggaacagggc acagtctgct tctgtgccac agttcaagaa agtggctctc 2580
 caggagtcca ctgatggctc attcaccag ccctgtaca gaggggaact gaatgagcac 2640
 ctgggactcc tgggaccata catcagggtc gaggtggaag acaacatcat ggtgacattc 2700

agaaaccagg cctccaggcc ctacagcttc tactcttccc tcatcagcta tgaggaagac 2760

 cagagacaag gggctgagcc aagaaagaac tttgtgaac ccaatgaaac caagacctac 2820
 ttctggaaag tccagcacca catggcacc accaaggatg agtttgactg caaggcctgg 2880
 gcatacttct ctgatgtgga cctggagaaa gatgtgcact ctggcctgat tggccactc 2940
 ctggtctgcc acaccaaac cctgaacct gcacatggaa ggcaagtgac tgtgcaggag 3000
 tttgccctct tcttccat ctttgatgaa accaagtcac ggtacttcac tgagaacatg 3060
 gagagaaact gcagagcacc atgcaacatt cagatggaag accccacctt caaggagaac 3120
 tacaggttcc atgcatcaa tggctacatc atggacaccc tgcctgggct tgtcatggca 3180

 caggaccaga gaatcagatg gtacctgctt tctatgggat ccaatgagaa cattcactcc 3240
 atccacttct ctgggcatgt cttcactgtg agaaagaagg aggaatacaa gatggcctg 3300
 tacaacctct acctgggggt ctttgagact gtggagatgc tgcctccaa agctggcatc 3360
 tggagggtgg aatgccatc tggggagcac ctgcatgctg gcatgtcaac cctgttcctg 3420
 gtctacagca acaagtcca gacaccctg ggaatggcct ctggccacat cagggacttc 3480
 cagatcactg cctctggcca gtatggccag tgggcaccca aactggccag getccactac 3540
 tctggctcca tcaatgcatg gtcaaccaag gagccattct cttggatcaa ggtggacctg 3600

 ctggcaccca tgatcattca tggcatcaag acacaggggg caagacagaa atctctctct 3660
 ctgtacatct cacagttcat catcatgtac tctctggatg gcaagaagtg gcagacatac 3720
 agaggcaact cactggcac cctcatggtc ttctttggca atgtggacag ctctggcatc 3780
 aagcacaaca tcttcaacc tcccatcatt gccagataca tcaggctgca ccccaccac 3840
 tactcaatca gatcaacct caggatggaa ctgatgggat gtgacctgaa ctctgtctca 3900
 atgcccttg gaatggagag caagccatt tctgatgccc agatcactgc atcctcttac 3960
 ttcaccaaca tgtttgccac ctggtcacca tcaaaagcca ggctgcacct ccaggaaga 4020

 agcaatgcct ggagaccca ggtcaacaac ccaaaggaat ggctgcaagt ggacttccag 4080
 aagacaatga aagtcactgg ggtgacaacc caggggtca agtctctgct cacctcaatg 4140
 tatgtgaagg agttcctgat ctcttctca caggatggcc accagtggac actcttcttc 4200
 cagaatggca aagtcaaggt gttccagggc aaccaggact ctttcacacc tgtggtgaac 4260
 tcaactggacc cccccctct gacaagatac ctgagaattc acccccagtc ttgggtccac 4320
 cagattgccc tgagaatgga agtctggga tgtgaggcac aagacctgta ctga 4374

<210> 14

<211> 4374

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 14

```

atgcagatcg aactgagcac ttgcttcttc ctgtgtctcc tgcgcttttg cttctccgcc      60
acaaggagat actatctcgg tgccgtggag ctgagctggg actacatgca gagcgacttg      120
ggtgaactgc ctgtggacgc caggtttcca ccccgctgac ccaagagttt cccgttcaac      180
accagtgtgc tgtacaagaa aacctcttc gtggaattca ccgaccacct gttcaacatc      240
gccaaaccgc gccctccctg gatggggctg ctgggccga cgatccaggc tgaggctat      300
gacacggtgg tgattaccct caagaacatg gctagccacc cggtagcct gcacgccgtg      360

ggcgtgtcct attgaaaagc gtccgagggt gcggagtacg atgaccagac ttcacagcgg      420
gagaaggaag acgacaaagt gttccccggg ggttcccaca cctatgtctg gcaggctctg      480
aaggagaatg gtcctatggc ctccgacca ttgtgcctca cctactctta cctaagccat      540
gtggatctcg tcaaggacct gaactcgggg ctgatcggcg cctgctcgt gtgccgggag      600
ggctcactgg ccaaggagaa gacccaaact ctgcacaagt tcatcctgct gttcgcggtg      660
ttcgacgagg ggaagtctg gcaactccgag accaagaaca gcctgatgca ggaccgcgac      720
gcagcctcgg cccgtgcgtg gccaaagatg cacaccgtga acggctacgt taacaggagc      780

ctaccggccc tgatcggctg ccaccgcaa tgggtctact ggcatgtgat cggaatgggc      840
acaaccccc aggtccacag tatcttctc gagggccaca ctttctggt cgggaatcac      900
cgccaggcca gcctggagat cagccccata acctttctga cggcgcagac ctactcatg      960
gatctcggcc agttctctc gttctgccac atttctccc accagcacga tgggatggaa      1020
gcatatgtga aagtggactc ctgccccgag gaaccccagc ttaggatgaa gaacaatgag      1080
gaggccgagg actacgacga tgacctacc gattcagaaa tggacgtagt acgctttgac      1140
gacgacaact ctccatcct catacagatt cgctccgtcg ccaagaagca ccctaagact      1200

tgggtgcact acatcgcggc cgaggaggag gactgggatt atgctcccct ggtgctggcc      1260
cccgacgacc gcagctacaa gagccagtac ctgaataacg ggccccagcg catcgccgg      1320
aagtacaaga aagtgcggtt catggcttac acggacgaga cttcaagac cggggaggct      1380
atccagcatg agagcggcat cttggggccc ctctgtacg gcgaagtgg agacacactg      1440
ctgatcatct tcaagaacca ggcgagcagg ccctacaaca tctacccca cggcattacc      1500
gatgtccgce cgttgtacag ccgacggctg cccaagggcg tgaagcacct gaaggacttt      1560

```

ccgatcctgc cgggcgagat cttcaagtac aagtggactg tgaccgtgga ggatgggccg 1620

 accaagagcg atccgcgctg cctgaccctg tactactcca gctttgtcaa tatggagcgc 1680
 gacctcgcta gggccttgat tggcctctg ctgatctgct acaaggagtc cgtggaccag 1740
 agggggaatc agatcatgag tgacaagagg aacgtgatcc tgttctccgt gttcgacgaa 1800
 aaccgcagct ggtatctcac cgagaatc cagcgcttcc tgccaaccc ggccggtgtg 1860
 cagctggagg accccgagtt tcagccagc aacatcatgc attctatcaa cggatatgtg 1920
 tttgattccc tgcagctctc agtgtgtctg cagcaggtcg cctactggta taccctcagc 1980
 attggggcac agaccgactt cctgagcgtg ttcttctccg ggtatacctt caagcacaag 2040

 atggtgtacg aggatacctt gaccctgttc ccctttagcg gcgaaaccgt gtttatgtct 2100
 atggagaacc cgggctctg gatccttggc tgccataact ccgacttccg caaccgcgga 2160
 atgaccgcgc tctgaaagt gtcgagttgt gacaagaaca ccggcgacta ttacgaggac 2220
 agttacgagg acatctctgc gtacctcctt agcaagaata acgcatcga gccaagatcc 2280
 ttcagccaga acccccagt gctgaagagg catcagcggg agatcaccg cacgacctg 2340
 cagtcggate agaggagat tgattacgac gacacgatca gtgtggagat gaagaaggag 2400
 gacttcgaca tctacgacga agatgaaaac cagtccctc ggtccttcca aaagaagacc 2460

 cggcactact tcatcggcg tgtggaacgc ctgtgggact atggaatgtc ttctagccct 2520
 cacgttttga ggaaccgcgc ccagtcgggc agcgtgcccc agttcaagaa agtggigttc 2580
 caggagtcca ccgacggctc cttcaccag ccactttacc ggggcgagct caatgaacat 2640
 ctgggcctgc tgggacccta catcagggtc gaggtggagg acaacatcat ggtgacattc 2700
 cggaatcagg ccagcagacc atacagtctc tacagtccac tcatctccta cgaggaggac 2760
 cagcggcagg gggctgaacc ccgtaagaac ttcgtgaagc caaacgaaac aaagacctac 2820
 ttctggaagg tccagcaca catggcacct accaaggacg agttcgattg caaggcctgg 2880

 gcctacttct ccgacgtgga cctggagaaa gatgtgcaca gcggcctgat tggcctctg 2940
 ctggtgtgtc acacgaacac actcaacctt gcacacgggc ggcaggtcac tgtgcaggaa 3000
 ttcgcctgtt tctttaccat ctttgatgag acgaagtcct ggtatttcac cgaaaacatg 3060
 gagaggaact gccgcgacc ctgcaacatc cagatggaag atccgacatt caaggagaac 3120
 taccggttcc atgcatcaa tggctacatc atggacacce tgctggcct cgtgatggcc 3180
 caagaccagc gtatccgctg gtatctgctg tcgatgggct ccaacgagaa catccatagt 3240
 atccacttca gcgggcatgt cttcacggtg aggaaaaagg aggagtaca gatggcactg 3300

tacaacctct atccccggct gttcgagacc gtggagatgc tgcctccaa ggccggcatc 3360
 tggagagtgg aatccctgat cggcgagcac ctccacgtg ggatgtccac gctgttctc 3420
 gtttacagca ataagtgcc aaccctctg ggcatggcga gcggccacat ccgcgacttc 3480
 cagattacag ccagcggcca gtacggtcag tgggctccaa agctggccc tctgactac 3540
 tccggatcca tcaacgctg gtccaccaag gaaccgttct cctggatcaa agtagacctg 3600
 ctagecccca tgatcattca cggcatcaag acacaaggcg cccgacagaa gttctcgagc 3660
 ctctatatct ccagttcat catcatgtat agcctggacg gaaagaagtg gcagacttac 3720

cgcggaact cgacaggac cctgatggta ttcttcgta acgtggacag ctccggaatc 3780
 aagcacaaca tttcaacc acccattatc gcccgctaca tccgctgca cccactcac 3840
 tatagatta ggtccacct gcgaatggag ctcatggct gtgacctgaa cagctgtagc 3900
 atgccctcg gcatggagtc taaggcgatc tccgacgac agataacggc atcatctac 3960
 tttaccaaca tgctcgctac ctggcccc tccaaggccc gactccacct gcaagggaga 4020
 tccaacgct ggcggccaca ggtcaacaat cccaaggagt ggctgcaagt ggactttcag 4080
 aaaaatga aagtcaccgg agtgaccaca caggagtgta agtctctgct gaccagcatg 4140

tacgtgaagg agttctcat ctccagttcg caggatggcc accagtggac gttgttcttc 4200
 caaacggta aagtcaaagt cttccaagg aaccaggaca gctttacacc cgtcgtgaac 4260
 tccctggacc cccgcttct cactagatac ctccgcatcc accctcagag ctgggtgac 4320
 cagattgcc tgcgcatgga ggttctgggg tgtgaagccc aggacctgta ctaa 4374

<210> 15

<211> 4374

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 15

atgcagattg agctctccac ctgctcttt ctctgccttc ttcgcttctg cttttctgcc 60

acacgcaggt actatttggg agcagtggaa ctgagctggg attacatgca gagtgcctt 120

ggtgaacttc ctgtggacgc tcgttttcca cctagagttc ccaagtcctt cccttcaac 180

acctcagtgg tctacaagaa aacgctgttt gtggagtca ctgaccacct cttcaacatt 240

gccaaccaa gaccccttg gatgggattg ctgggacca caatacaagc agaagtctac 300

gacacggtgg tgattacct gaagaacatg gcgtcacacc ctgtttcact tcacgctgtt 360

ggggtcagtt attggaaagc ctgagagggt gcggaatagc atgatcaaac cagccagagg 420

gagaaggaag atgacaaggt ctttctggg gtagccata cctatgttg gcagggtctg 480

 aaagagaatg ggcctatggc ctctgatccc ttgtgcctca catactctta cctgagtcac 540
 gtcgacctgg tgaagacct gaatagcggc ctgattggg cactgcttgt ttgtagagag 600
 gggagtttgg ccaaggagaa aactcagact ctccacaagt ttatcctcct gtttgctgtg 660
 ttcgacgagg gcaagtcttg gcactctgaa acaagaact ccctgatgca ggacagagat 720
 gctgcatctg caaggccttg gccaaaaatg cacacagtga acggctatgt gaatcgatca 780
 ctgccaggac tgataggctg tcatcgcaag tcagtgtatt ggacgcttat cgggatggga 840
 acaactccag aagtgcacag catcttctct gagggccaca ctttctggt tcggaatcat 900

 agacaggcca gccttgagat cagcccaatc acctttctga ctgcccaaac cttgctgatg 960
 gatctgggac agttcctcct gttttgtcac atctcctccc accaacaatga cgggatggag 1020
 gcttatgtga aggtcgatag ctgtccggag gaaccacaac tgaggatgaa gaacaacgaa 1080
 gaggcagagg aciatgacga cgatctgact gacagtgaaa tggacgtggt tcggttcgac 1140
 gatgacaatt ctcttcatt tatccagatc cgttccgtgg ccaagaagca cccaagact 1200
 tgggttcatt acatcgctgc tgaggaggag gattgggact acgcgccctt ggtgttgccc 1260
 ccagacgatc gtcatacaa gagccagtac cttacaatg gtccacaag gatcggccgg 1320

 aagtacaaga aggttagatt tatggcttat accgacgaga cttttaaacc tagggaagca 1380
 attcagcatg aaagtggcat tcttggacc ctgctgtatg gcgaggttgg cgacaccctg 1440
 ctgattatct ttaagaacca ggcaagccgg ccctacaaca tctaccgca cggcataacc 1500
 gatgtacgac cctgtacag tcgcagactt cctaaagggg tgaaacacct gaaggacttc 1560
 ccaattctgc ccggggagat cttcaagtat aaatggaccg tgacggttga ggatgtccc 1620
 acaaagtccg atccgagatg ccttaccga tattattcca gcttcgtgaa catggaaggg 1680
 gacctggcca ggggctgat tggcccactg ctgatttgtt acaaggagtc tgtcgatcaa 1740

 agaggaaacc aaataatgag cgacaaacgt aacgtcatcc tgttcagcgt ctttgatgag 1800
 aatagaagct ggtacctcac agaaaaatatt cagcggtttc tgcctaacc cgcaggcgtc 1860
 cagctggaag atcccagatt ccaagcctca aacatcatgc atagcatcaa cggatacgtg 1920
 ttcgatagcc tgcagctgtc cgtctgtctc catgaagtgg catatttgta catcctgagt 1980
 atcggggcgc agaccgactt cctgagcgtg ttctttctg gatacacgtt caaacacaaa 2040
 atggtctatg aagataacct gactctgttt ccattctcag gagagacagt ctttatgagt 2100
 atggaaaatc ctggactgtg gatcctgggc tgcacaatt ctgattttcg gaacagagge 2160

atgacagccc tgcttaaagt gagtcatgc gacaagaaca ccggtgatta ctacgaagat 2220
 agctatgagg acatcagtc gtatttgtc tccaagaaca acgctatcga gccacggtct 2280
 ttcagtcaga atcctcccgt tctgaagcgg catcagcgcg aaataacacg cacaaccctt 2340
 cagtcagacc aagaggaat cgactacgat gatactatct ctgtggagat gaagaaggag 2400
 gatttcgaca ttacgacga ggacgagaat cagtcccaa ggagctttca gaagaaaaca 2460
 agacactatt tcattgccgc cgtggagcga ctgtgggact acggcatgtc tagctctccg 2520
 catgtactta gaaatagggc acaaagcggg tccgtgcctc agtttaagaa agttgtcttt 2580

 caggagtta cagatggctc cttcaccag cccttgtatc gcggggaact caatgaacac 2640
 ctgggcctcc tgggtcctta tattagggcc gaagtcgagg acaatatcat ggtgaccttt 2700
 aggaaccagg catctagacc ttactcttc tactcctccc tgatataccta tgaggaggac 2760
 cagcggcaag gcctgagcc tcggaagaac tttgtgaagc caaatgaaac caaacatac 2820
 ttttgaaag ttcagcacca catggctccc acgaaggacg aatttgactg taaagcctgg 2880
 gcctacttct cagatgtaga tctcgagaaa gacgtgcact cagggtcat tggccctc 2940
 ctggtctgtc atactaatac cctcaatcca gcacacggac gtcaggtaac cgtccaggaa 3000

 tttgccctgt tccttaccat tttcgatgag actaaatcct ggtactttac cgaacatg 3060
 gagaggaatt gcagagcccc atgcaacatc cagatggagg acctacctt caaagagaac 3120
 tatecttcc atgccattaa cggttacatt atggatactc tcccaggact tgtgatggca 3180
 caggatcagc ggataagatg gtatctgttg agcatgggct ccaacgagaa tattcacagc 3240
 atccatttct ccggtcacgt gtttacagtg agaaagaaag aagagtaca gatggctctg 3300
 tataatctct atccaggcgt attcgaaacg gtggagatgt tgcctagcaa ggccgacatt 3360
 tggcgagtag aatgccttat cggggaacat ctgcatgccg gaatgagcac gctcttctg 3420

 gtgtatagta acaagtgcca gactccgtg ggcatggcat ctggccatat acgggacttt 3480
 cagattacgg ctagcgggca gtatgggcag tggcaccca aacttgcgcg actgcactat 3540
 tcaggctcta tcaatgcatg gtccaccaag gaaccttct cttggattaa ggtggacctt 3600
 ttggcgccea tgataatcca tgggatcaaa acccagggcg ctcgtcagaa attctcatca 3660
 ctctacatct ctcagttcat aataatgtat tctctggatg ggaagaaatg gcagacttac 3720
 agaggaaca gcaccgggac gctgatggtg ttctttggca acgtggacag cagcggcatc 3780
 aaacacaaca tcttcaatcc tcccattatt gccggttata ttagactgca tccaactcac 3840

 tactctatac gcagcacact taggatggag ctcatgggat gcgacctgaa cagttgtagt 3900
 atgcccttgg ggatggagtc caaagctata agcagcgcac aaattacagc tagctcttac 3960
 tttacgaata tgctgccac gtggagccca agcaaagccc ggctgcattt gcagggtcgg 4020

agtaatgctt ggcgccaca ggtgaataac cctaaggaat ggttgcaagt agatttccag 4080
 aaaactatga aggtaaccgg cgtcactaca cagggagtca agtcctctt gacctctatg 4140
 tacgtcaagg agttcctgat tagcagcagt caggatgggc accaatggac actgttcttc 4200
 cagaatggga aagttaaagt atttcagggt aaccaggact cctttacacc tgtggtgaat 4260

 agcctcgacc caccctgct gacacgatac ctccgcatcc accctcagtc ttgggtgcat 4320
 caaattgccc tgcgaatgga ggtgttggga tgcgaagctc aggacctcta ctga 4374
 <210> 16
 <211> 4374
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide
 <400> 16
 atgcagatcg aactctctac ttgcttcttc ctgtgccttc tgaggttctg ctctctgcc 60
 actcgcgatg attacctcgg ggccgtggag ttgagttggg actacatgca atcagatctg 120
 ggccaactcc ctgtggatgc ccgattccca ccgcgcgtgc ccaagtcttt cccatttaat 180

 acttctgtgg tgiacaagaa gacattgttt gtggagtta ccgatcacct gttcaacatc 240
 gccaaaccgc ggccccatg gatgggtctg cttgggcca ccattcaagc ggaggtctat 300
 gatacagtgg tgataacgct taagaacatg gcgagccacc cagtgtctct gcattgccgtt 360
 ggtgtatcat attggaaggc cagcgaagga gcggagtacg atgaccagac ctctcagaga 420
 gagaaggaag acgataaggt ttttctggc ggaagtcata catatgtatg gcaggtcctg 480
 aaagagaatg ggccgatggc ttctgacccc ctttgtctta cctatagtta tctgagccac 540
 gtggacctgg tcaaggacct caacagtggg ctgattgggg ctctgcttgt ttgtagagag 600

 ggtagcttgg ctaaggagaa aacccaaaca ctccataagt tcattttgct gttcgcggtg 660
 ttcgacgagg gaaagagtg gcacagcga acaaagaatt cactgatgca agacagggac 720
 gccgcttccg caagggttg gcctaagatg catacggatg atgggtatgt gaaccggagc 780
 ctccggggc tgatcgggtg ccacgcaag tctgtttact ggcacgtcat tggaatgggg 840
 acaacccag aggtacatag tatatttctt gaaggccaca cgttctctgt acggaaccac 900
 cgacaggctt ccctggagat aagccccatt acctttctga ccgctcagac tctgctgatg 960
 gacctggcc agtttctct gttctgceat attagcagcc accagcacga cggatggaa 1020

 gcatacgtga aagtcgatag ctgtcctgag gacccctcagc tcagaatgaa gaacaacgag 1080

gaggccgaag aci atgacga tgacctaca gattccgaga tggacgtggt gcgctttgac 1140
 gacgataaca gtctagttt cattcaaatc agatccgtag ccaaaaagca tccaaagaca 1200
 tgggtgcatt acattgcagc cgaagaggag gattgggatt atgcgccct tgttctggct 1260
 ccagatgaca ggagctataa gtcccagtac ttgaacaacg ggccacagcg aatcggtaga 1320
 aaataaaga aggtaagatt catggcctac actgacgaaa catttaaac cagggaagct 1380
 atccaacacg aatctggaat tctcgccct ctgctctacg gtgaggtggg ggacacctg 1440

 ctgatcattt tcaaaaatca ggcatccagg cttacaaca tatacccca tggcatcacc 1500
 gatgtccgc cgtgtattc cagaagactc cccaaggag tgaacatct gaaagatttt 1560
 cccatctgc cggcgagat ctttaatac aaatggactg tgactgtaga ggacggcct 1620
 acaaatcag acccacggtg cctgacaagg tattacagta gcttcgtcaa catggaacgc 1680
 gacctgccca gcggactcat tggcccactg ttgatctgtt acaaagagtc agtggatcag 1740
 aggggaaatc agatcatgag cgataagaga aacgttatcc tgtttagtgt cttcgacgag 1800
 aaccggtctt ggtaccttac tgagaacatc cagaggttcc tgccgaatcc ggctggcgtt 1860

 cagctcgagg acccagagtt ccagccagt aatataatgc actcaatcaa cggttatgtg 1920
 ttcgatagcc tgcagctgag cgtctgcctc cagaggtag cctattgta catattgtcc 1980
 atcggggctc agaccgattt tctgtccgtg ttctttagcg ggtatacctt taaacataaa 2040
 atggtctatg aagacacct gacctgttc ccattctccg gtgagactgt gttcatgtcc 2100
 atggagaacc cagggtgtg gatcctgggg tgcacaata gtgactttag gaatcgggga 2160
 atgacggcac tgctgaaggt gaggttctgc gataaaaata caggagatta ctatgaggat 2220
 agttacgagg atatcagtc ctatctgctt tcaaaaaaca acgcaattga gccccgtct 2280

 ttctcaaaa accccccgt gctgaagcgc caccagcgcg aaattaccg gacaacctg 2340
 cagtcgacc aggaggaat cgattatgac gatactatca gtgtagaaat gaaaaaggag 2400
 gatthtgata tttacgacga agacgagaac cagtctccgc gaagttttca gaagaaaacg 2460
 cgacactact ttatagctgc cgtggaacga ctctgggatt atggcatgtc ctccagcct 2520
 catgtcctta ggaatcgagc gcagagtggc tctgtgctc agttcaaaaa ggttgtgttc 2580
 caggaattca ccgacggctc atttaccag ccgctgtaca gaggcgaact caacgaacac 2640
 cttgggtgc ttggccata tattcgagca gaggtggaag ataatatcat ggtaacctt 2700

 agaaaccagg cgicaagacc ctattccttc tacagttctc tgatcagcta cgaggaggac 2760
 caaagacagg gagctgaacc caggaagaac tttgtgaac ctaatgagac caagacctac 2820
 ttctggaagg tccagacca tatggcccca actaaagatg aattcgattg caaggcctgg 2880
 gcttatttca gcgacgtgga tctcgaag gatgtgcaca gcgggttgat cggaccgctt 2940

ttggtgtgcc acacaaatac cctcaatcct gccacgggc ggcaggtcac agttcaagag 3000
 ttgactct tctttacaat atttgacgag acaaagtcac ggtatcttac agagaatag 3060
 gagagaaatt gtcgcgacc ttgcaacatt cagatggagg accccacatt taaggagaat 3120

tacagatttc atgctatcaa tgggtacatt atggatactc tgcctggtct ggtcatggcc 3180
 caggatcagc gcataagggtg gtacttgctg agcatgggat ctaatgagaa tatacacagc 3240
 attcacttca gtggccacgt ttttactggt agaaagaagg aggagtacaa aatggcgctc 3300
 tacaaccttt acccgggtgt gtttgagaca gtggagatgc tgccaagcaa ggcaggcac 3360
 tggagggttg agtgtcttat tggggagcat ctgcatgctg gaatgtccac cctctttctt 3420
 gtgtacagca ataagtcca gacaccgctt ggcatggcca gcggccacat tagggacttt 3480
 cagataactg ccagtggaca gtacggccag tgggctcca agcttgcaag actccactac 3540

tccggaagca taaacgcatg gagcaccaag gaacccttct cttggattaa ggtggacctg 3600
 ctggcgccaa tgatcattca cggcataaaa acccaagggg cacgacagaa attttcatct 3660
 ttgtatatta gtcagtttat catcatgtac agcttggatg gaaagaagtg gcagacgtac 3720
 aggggcaatt ctacaggaac acttatggtg ttttttggga atgtcgatc cagcgggatc 3780
 aacataaca tcttcaatcc tcttattatc gccgatata tccgcctgca cctacgcat 3840
 tactccatca ggtccacatt gagaatggaa ctgatgggtg gcgacctgaa tagttgtagt 3900
 atgccactgg geatggagtc taaagccatc agcgatgcac agatcactgc cagctcttac 3960

ttaccaaca tgittgcaac ttggtcccc tctaaagctc gcctgcatct gcagggacgc 4020
 tcaaatgcat ggcgaccaca ggtgaacaat ccaaaagagt ggctccaggt cgactttcag 4080
 aagacaatga aggtaacagg agtgacaacc cagggtgtaa aaagcctcct tacgagtatg 4140
 tacgttaagg agtttctgat ttctagctcc caggacggac accagtggac tctgttcttc 4200
 cagaacggca aagtgaaggt atttcaggga aaccaggatt cttttacccc ggtagtgaat 4260
 agcctggatc caccgttctg gaccgctat ctgagaattc atccacaatc ctgggtgcat 4320
 cagattgccc tccggatgga agtgctcggc tgtgaagctc aggatctgta ttag 4374

<210> 17

<211> 4374

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 17

atgcaaatag agctctccac ctgcttcttt ctgtgccttt tgcgattctg ctttagtgcc 60
 accagaagat actacctggg tgcagtgga ctgtcatggg actatatgca aagtgatctc 120
 ggtgagctgc ctgtggacgc aagatttctt cctagagtgc caaaatcttt tccattcaac 180
 acctcagtcg tgiacaaaa gactctgttt gtagaattca cggatcacct tttcaacatc 240
 gctaagccaa ggccaccctg gatgggtctg ctaggctcta ccatccaggc tgaggtttat 300

 gatacagtgg tcattacact taagaacatg gcttcccatc ctgtcagtct tcatgctgtt 360
 ggtgtatcct actggaaagc ttctgaggga gctgaatatg atgatcagac cagtcaaagg 420
 gagaaagaag atgataaagt cttccctggg ggaagccata catatgtctg gcaggctctg 480
 aaagagaatg gtccaatggc ctctgacceca ctgtgcctta cctactcata tctttctcat 540
 gtggacctgg taaaagactt gaattcaggc ctcatggag ccctactagt atgtagagaa 600
 gggagtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttctgta 660
 tttgatgaag ggaaaagtg gcactcagaa acaaagaact cttgatgca ggatagggat 720

 gctgcatctg ctctggcctg gcctaaaatg cacacagtca atggttatgt aaacaggtct 780
 ctgccaggtc tgattggatg ccacaggaaa tcagtctatt ggcatgtgat tggaatgggc 840
 accactcctg aagtgcactc aatattcttc gaaggtcaca catttcttgt gaggaacat 900
 cgccaggcgt ccttggaat ctcgccaata actttcctta ctgctcaaac actcttgatg 960
 gaccttgac agtttctact gttttgctat atctcttccc accaacaatga tggcatggaa 1020
 gcttatgtca aagtagacag ctgtccagag gaacccaac tacgaatgaa aaataatgaa 1080
 gaagcggaag actatgatga tgatcttact gattctgaaa tggatgtggt caggtttgat 1140

 gatgacaact ctcttctctt tatccaaat cgctcagttg ccaagaagca tcctaaaact 1200
 tgggtacatt acattgtctg tgaagaggag gactgggact atgctccctt agtcctcgcc 1260
 cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gatttgtagg 1320
 aagtacaaaa aagtccgatt tatggcatac acagatgaaa cctttaagac tcgtgaagct 1380
 attcagcatg aatcaggaat cttgggacct ttactttatg gggaaagtgg agacacactg 1440
 ttgattatat ttaagaatca agcaagcaga ccatataaca tctaccctca cggaatcact 1500
 gatgtccgtc cttgtattc aaggagatta ccaaagggtg taaaacattt gaaggatttt 1560

 ccaattctgc caggagaaat attcaaatat aaatggacag tgactgtaga agatgggcca 1620
 actaaatcag atcctcggtg cctgacccgc tattactcta gtttcgttaa tatggagaga 1680
 gatctagctt caggactcat tggcctctc ctcatctgct acaaagaatc ttagatcaa 1740
 agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag 1800
 aaccgaagct ggiacctcac agagaatata caacgctttc tcccaatcc agctggagtg 1860

cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatgtt 1920
 tttgatagtt tgcagttgtc agtttgtttg catgaggtgg catactggta cattctaagc 1980

 attgagcac agactgactt cttttctgtc ttcttctctg gatatacctt caaacacaaa 2040
 atggctctatg aagacacact caccctattc ccattctcag gagaaactgt cttcatgtcg 2100
 atggaaaacc caggtctatg gattctgggg tgccacaact cagactttcg gaacagaggc 2160
 atgaccgcct tactgaaggt ttctagtgtg gacaagaaca ctggtgatta ttacgaggac 2220
 agttatgaag atatctcagc atacttctg agtaaaaaca atgccattga accaagaagc 2280
 ttctcccaga atccaccagt cttgaaacgc catcaacggg aaataactcg tactactctt 2340
 cagtcagatc aagaggaat tgactatgat gataccatat cagttgaaat gaagaaggaa 2400

 gattttgaca tttatgatga ggatgaaaat cagagccccc gcagctttca aaagaaaaca 2460
 cgacactatt ttatgtctgc agtggagagg ctctgggatt atgggatgag tagctcccca 2520
 catgttctaa gaaacagggc tcagagtggc agtgtcctc agttcaagaa agttgttttc 2580
 caggaattta ctgatggctc ctttactcag cccttatacc gtggagaact aaatgaacat 2640
 ttgggactcc tggggccata tataagagca gaagttgaag ataatatcat ggtaactttc 2700
 agaaatcagg ccctcgtcc ctattccttc tattctagcc ttatttctta tgaggaagat 2760
 cagaggcaag gagcagaacc tagaaaaaac ttgtcaagc ctaatgaaac caaaacttac 2820

 ttttgaaag tgcaacatca tatggcacc actaaagatg agtttgactg caaagcctgg 2880
 gcttatttct ctgatgttga cctggaaaaa gatgtgact caggcctgat tggaccctt 2940
 ctggtctgcc aactaacac actgaacct gctcatggga gacaagtac agtacaggaa 3000
 tttgctctgt tttcacat ctttgatgag accaaaagct ggtacttcac tgaaaatag 3060
 gaaagaaact gcagggctcc ctgcaatc cagatggaag atcccactt taaagagaat 3120
 tategettcc atgcaatcaa tggctacata atggatacac tacctggctt agtaatgget 3180
 caggatcaaa ggattcgatg gtatctgtc agcatgggca gcaatgaaa catccattct 3240

 attcatttca gtggacatgt gttcactgta cgaaaaaag aggagtataa aatggcactg 3300
 tacaatctct atccaggtgt ttttgagaca gtggaaatgt taccatcaa agctggaatt 3360
 tggcgggtgg aatgccttat tggcgagcat ctacatctg ggatgagcac acttttctg 3420
 gtgtacagca ataagtgtca gactccctg ggaatggctt ctggacacat tagagatttt 3480
 cagattacag cttcaggaca atatggacag tgggcccaca agctggccag acttcattat 3540
 tccgatcaa tcaatgcctg gagcaccaag gagccctttt cttgatcaa ggtggatctg 3600

ttggcaccaa tgattattca cggcatcaag acccaggggtg cccgtcagaa gttctccagc 3660

ctctacatct ctcagtttat catcatgtat agtcttgatg ggaagaagtg gcagacttat 3720

cgaggaatt ccaactggaac cttaatggtc ttctttggca atgtggattc atctgggata 3780

aaacacaata tttttaacct tccaattatt gctcgataca tccgtttgca cccaactcat 3840

tatagcattc gcagcactct tcgcatggag ttgatgggct gtgatttaa tagttgcagc 3900

atgccattgg gaatggagag taaagcaata tcagatgcac agattactgc ttcacacctac 3960

ttaccaata tgttgcccac ctggtctcct tcaaaagctc gacttcacct ccaagggagg 4020

agtaatgcct ggagacctca ggtgaataat ccaaaagagt ggctgcaagt ggacttccag 4080

aagacaatga aagtcacagg agtaactact cagggagtaa aatctctgct taccagcatg 4140

tatgtgaagg agttcctcat ctccagcagt caagatggcc atcagtggac tctctttttt 4200

cagaatggca aagtaaaggt ttttcagga aatcaagact cttcacacc tgtggtgaac 4260

tctctagacc caccgttact gactcgctac ctctgaattc acccccagag ttgggtgcac 4320

cagattgccc ttaggatgga ggttctgggc tgcgaggcac aggacctcta ctga 4374

<210> 18

<211> 4374

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 18

atgcagatcg agctgtccac atgctttttt ctgtgcctgc tgcggttctg cttcagcgcc 60

accggcggt actacctggg cgccgtggag ctgtcctggg actacatgca gagcgacctg 120

ggcgagctgc ccgtggacgc ccggttccc ccagagtgc ccaagagctt cccttcaac 180

accagcgtgg tgtacaagaa aacctgttc gtggagtca ccgaccacct gttcaacatc 240

gccaaagccc ggccccctg gatgggcctg ctgggcccga ccatccaggc cgagggttac 300

gacaccgtgg tgatcacct gaagaacatg gccagccacc ccgtgagcct gcacgccgtg 360

ggcgtgagct actggaaggc ctccgagggc gccgagtac acgaccagac cagccagcgg 420

gagaaagagg acgacaaagt ctttctggc gccagccaca cctacgtgtg gcaggctctg 480

aaagaaaacg gccccatggc ctccgaccc ctgtgcctga cctacagcta cctgagccac 540

gtggacctgg tgaaggacct gaacagcggg ctgattgggg ccttctgtgt ctgccgggag 600

ggcagcctgg ccaagagaa aaccagacc ctgcacaagt tcatcctgct gttcgccgtg 660

ttcgacgagg gcaagagctg gcacagcgag accaagaaca gcctgatgca ggaccgggac 720
 gccgcctctg ccagagcctg gcccaagatg cacaccgtga acggctacgt gaacagaagc 780
 ctgcccggcc tgattggctg ccaccggaag agcgtgtact ggacacgtgat cggcatgggc 840

 accacaccgg aggtgcacag catctttctg gaagggcaca ctttctggt gcggaaccac 900
 cggcaggcca gcctggaat cagccctatc accttctga ccgccagac actgctgatg 960
 gacctgggcc agttctgct gttttgccac atcagctctc accagcacga cggcatggaa 1020
 gcctacgtga aggtggactc ctgccccgag gaacccagc tgcgatgaa gaacaacgag 1080
 gaagccgagg actacgacga cgacctgacc gacagcgaga tggacgtggt gcggttcgac 1140
 gacgacaaca gccccagctt catccagatc agaagcgtgg ccaagaagca cccaagacc 1200
 tgggtgcact acatcgccgc cgaggaagag gactgggact acgccccct ggtgctggcc 1260

 cccgacgaca gaagctaaa gagccagtac ctgaacaatg gccccagcg gatcgccgg 1320
 aagtacaaga aagtgcggtt catggcctac accgacgaga cttcaagac cggggaggcc 1380
 atccagcacg agagcggcat cctgggcccc ctgctgtacg gcgaagtggg cgacacactg 1440
 ctgatcatct tcaagaacca ggccagccgg cctacaaca tctacccca cggcatcacc 1500
 gacgtcggc ccctgtacag caggcggctg cccaagggcg tgaagcacct gaaggacttc 1560
 cccatcctgc cggcgagat cttcaagtac aagtggaccg tgaccgtgga ggacggcccc 1620
 accaagagcg accccagatg cctgacccgg tactacagca gttcgtgaa catggaacgg 1680

 gacctggcct cgggctgat cggacctctg ctgatctgct acaaagaaag cgtggaccag 1740
 cggggcaacc agatcatgag cgacaagcgg aacgtgatcc tgttcagcgt gttcgtgag 1800
 aaccggtcct ggiatctgac cgagaacatc cagcggtttc tgccaaccc tgccggggtg 1860
 cagctggaag atcccagtt ccagccagc aacatcatgc actccatcaa tggtacgtg 1920
 ttcgacagcc tgcagctgtc cgtgtgtctg cacgaggtgg cctactgta catcctgagc 1980
 atcggcgccc agaccgactt cctgagcgtg ttcttcagcg gctacacctt caagcacaag 2040
 atggtgtacg aggacaccct gacctgttc ctttcagcg gcgagaccgt gttcatgagc 2100

 atgaaaaacc cggcctgtg gatcctgggc tgccacaaca gcgacttccg gaaccggggc 2160
 atgaccgcc tgctgaaggt gtccagctgc gacaagaaca ccggcgacta ctacaggac 2220
 agctacgagg atatcagcgc ctacctgtg tccaagaaca acgcatcga gccagaagc 2280
 ttcagccaga acccccctgt gctgaagcgg caccagagag agatcacccg gaccacctg 2340
 cagtccgacc aggaagagat cgattacgac gacacatca gcgtggagat gaaaaaagaa 2400
 gatttcgaca tctacgacga ggacgagaac cagagcccc ggtccttcca gaagaaaacc 2460

cggcactact ttatcgccgc cgtggagcgg ctgtgggact acggcatgag cagcagcccc 2520

 cacgtgctgc ggaaccgggc ccagagcggc agcgtgcccc agttcaagaa agtgggtttc 2580
 caggaattca ccgacggcag cttcacccag ccctgtacc ggggcgagct gaacgagcac 2640
 ctggggctgc tggggcccta catcagggcc gaagtggagg acaacatcat ggtgaccttc 2700
 cggaatcagg ccagcagacc ctactccttc tacagcagcc tgatcagcta cgaagaggac 2760
 cagcggcagg gcgctgaacc ccggaagaac ttcgtgaagc ccaatgagac caagacctac 2820
 ttctgaaaag tgcagcacca catggcccc accaaggacg agttcgactg caaggcctgg 2880
 gcctacttca gcgacgtgga tctggaaaag gacgtgcact ctggactgat tggccctctg 2940

 ctggtgtgcc acaccaacac cctgaacccc gccacggcc ggcaggtgac cgtgcaggaa 3000
 ttcgccctgt tcttaccat cttcgacgag accaagtctt ggtacttcac cgagaatatg 3060
 gaacggaact gcagagcccc ctgcaacatc cagatggaag atcctacctt caaagagaac 3120
 taccggttcc acgcatcaa cggctacatc atggacacc tgcctggcct ggtgatggcc 3180
 caggaccaga ggatccggtg gtatctgctg tccatgggca gcaacgagaa tatccacagc 3240
 atccacttca gggccacgt gttcacctg aggaagaaag aagagtacaa gatggccctg 3300
 tacaacctgt accccggcgt gttcgagacc gtggagatgc tgcccagcaa ggccggcatc 3360

 tggcgggtgg agtgtctgat cggcgagcac ctgcatgccg ggatgagcac cctgtttctg 3420
 gtgtacagca acaagtgcc gacccccctg ggcatggcca gcggccacat ccgggacttc 3480
 cagatcaccg cctccggcca gtacggccag tggccccca agctggcccg gctgcactac 3540
 agcggcagca tcaacgctg gtccacaaa gagcccttca gctggatcaa ggtggacctg 3600
 ctggccccta tgatcatcca cggcattaag acccagggcg ccaggcagaa gttcagcagc 3660
 ctgtacatca gccattcat catcatgtac agcctggacg gcaagaagt gcagacctac 3720
 cggggcaaca gcaccggcac cctgatggtg ttcttcggca acgtggacag cagcggcatc 3780

 aagcacaaca tcttcaacc cccatcatc gcccggtaca tccggctgca cccacccac 3840
 tacagcatca gatccacct gcggatggaa ctgatgggct gcgacctgaa ctctgcagc 3900
 atgcctctgg gcatggaag caaggccatc agcgacgccc agatcacagc cagcagctac 3960
 ttcaccaaca tgttgccac ctggtcccc tccaaggcca ggctgcacct gcagggccgg 4020
 tccaacgctt ggggccctca ggtgaacaac cccaaagaat ggctgcaggt ggactttcag 4080
 aaaacctga aggtgaccgg cgtgaccacc cagggcgtga aaagcctgct gaccagcatg 4140
 tacgtgaaag agttttctgat cagcagcagc caggacggcc accagtggac cctgttcttt 4200

cagaacggca aggtgaaagt gttccagggc aaccaggact ccttcacccc cgtgggtgaac 4260
 tccttggacc cccccctgct gacccgctac ctgcggatcc acccccagtc ttgggtgcac 4320
 cagatcgccc tgaggatgga agtgctggga tgtgaggccc aggatctgta ctga 4374

<210> 19

<211> 2351

<212> PRT

<213> Homo sapiens

<400> 19

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
 1 5 10 15
 Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser
 20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
 35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
 50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
 65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
 85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
 100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
 115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
 130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
 145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
 165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
 180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
 195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
 210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
 225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
 260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met

Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700
 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720
 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735

 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
 740 745 750
 Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro
 755 760 765
 Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp
 770 775 780
 Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys
 785 790 795 800

 Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser
 805 810 815
 Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr
 820 825 830
 Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn
 835 840 845
 Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly
 850 855 860

 Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu
 865 870 875 880
 Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys
 885 890 895
 Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn
 900 905 910
 Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met
 915 920 925

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

930 935 940
 Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu
 945 950 955 960
 Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu
 965 970 975
 Ser Ser Trp Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe
 980 985 990

 Lys Gly Lys Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala
 995 1000 1005
 Leu Phe Lys Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser
 1010 1015 1020
 Asn Asn Ser Ala Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser
 1025 1030 1035
 Leu Leu Ile Glu Asn Ser Pro Ser Val Trp Gln Asn Ile Leu Glu
 1040 1045 1050

 Ser Asp Thr Glu Phe Lys Lys Val Thr Pro Leu Ile His Asp Arg
 1055 1060 1065
 Met Leu Met Asp Lys Asn Ala Thr Ala Leu Arg Leu Asn His Met
 1070 1075 1080
 Ser Asn Lys Thr Thr Ser Ser Lys Asn Met Glu Met Val Gln Gln
 1085 1090 1095
 Lys Lys Glu Gly Pro Ile Pro Pro Asp Ala Gln Asn Pro Asp Met
 1100 1105 1110

 Ser Phe Phe Lys Met Leu Phe Leu Pro Glu Ser Ala Arg Trp Ile
 1115 1120 1125
 Gln Arg Thr His Gly Lys Asn Ser Leu Asn Ser Gly Gln Gly Pro
 1130 1135 1140
 Ser Pro Lys Gln Leu Val Ser Leu Gly Pro Glu Lys Ser Val Glu
 1145 1150 1155
 Gly Gln Asn Phe Leu Ser Glu Lys Asn Lys Val Val Val Gly Lys
 1160 1165 1170

Gly	Glu	Phe	Thr	Lys	Asp	Val	Gly	Leu	Lys	Glu	Met	Val	Phe	Pro
1175						1180					1185			
Ser	Ser	Arg	Asn	Leu	Phe	Leu	Thr	Asn	Leu	Asp	Asn	Leu	His	Glu
1190						1195					1200			
Asn	Asn	Thr	His	Asn	Gln	Glu	Lys	Lys	Ile	Gln	Glu	Glu	Ile	Glu
1205						1210					1215			
Lys	Lys	Glu	Thr	Leu	Ile	Gln	Glu	Asn	Val	Val	Leu	Pro	Gln	Ile
1220						1225					1230			
His	Thr	Val	Thr	Gly	Thr	Lys	Asn	Phe	Met	Lys	Asn	Leu	Phe	Leu
1235						1240					1245			
Leu	Ser	Thr	Arg	Gln	Asn	Val	Glu	Gly	Ser	Tyr	Asp	Gly	Ala	Tyr
1250						1255					1260			
Ala	Pro	Val	Leu	Gln	Asp	Phe	Arg	Ser	Leu	Asn	Asp	Ser	Thr	Asn
1265						1270					1275			
Arg	Thr	Lys	Lys	His	Thr	Ala	His	Phe	Ser	Lys	Lys	Gly	Glu	Glu
1280						1285					1290			
Glu	Asn	Leu	Glu	Gly	Leu	Gly	Asn	Gln	Thr	Lys	Gln	Ile	Val	Glu
1295						1300					1305			
Lys	Tyr	Ala	Cys	Thr	Thr	Arg	Ile	Ser	Pro	Asn	Thr	Ser	Gln	Gln
1310						1315					1320			
Asn	Phe	Val	Thr	Gln	Arg	Ser	Lys	Arg	Ala	Leu	Lys	Gln	Phe	Arg
1325						1330					1335			
Leu	Pro	Leu	Glu	Glu	Thr	Glu	Leu	Glu	Lys	Arg	Ile	Ile	Val	Asp
1340						1345					1350			
Asp	Thr	Ser	Thr	Gln	Trp	Ser	Lys	Asn	Met	Lys	His	Leu	Thr	Pro
1355						1360					1365			
Ser	Thr	Leu	Thr	Gln	Ile	Asp	Tyr	Asn	Glu	Lys	Glu	Lys	Gly	Ala
1370						1375					1380			
Ile	Thr	Gln	Ser	Pro	Leu	Ser	Asp	Cys	Leu	Thr	Arg	Ser	His	Ser
1385						1390					1395			
Ile	Pro	Gln	Ala	Asn	Arg	Ser	Pro	Leu	Pro	Ile	Ala	Lys	Val	Ser

1400	1405	1410
Ser Phe Pro Ser Ile Arg Pro Ile Tyr Leu Thr Arg Val Leu Phe		
1415	1420	1425
Gln Asp Asn Ser Ser His Leu Pro Ala Ala Ser Tyr Arg Lys Lys		
1430	1435	1440
Asp Ser Gly Val Gln Glu Ser Ser His Phe Leu Gln Gly Ala Lys		
1445	1450	1455
Lys Asn Asn Leu Ser Leu Ala Ile Leu Thr Leu Glu Met Thr Gly		
1460	1465	1470
Asp Gln Arg Glu Val Gly Ser Leu Gly Thr Ser Ala Thr Asn Ser		
1475	1480	1485
Val Thr Tyr Lys Lys Val Glu Asn Thr Val Leu Pro Lys Pro Asp		
1490	1495	1500
Leu Pro Lys Thr Ser Gly Lys Val Glu Leu Leu Pro Lys Val His		
1505	1510	1515
Ile Tyr Gln Lys Asp Leu Phe Pro Thr Glu Thr Ser Asn Gly Ser		
1520	1525	1530
Pro Gly His Leu Asp Leu Val Glu Gly Ser Leu Leu Gln Gly Thr		
1535	1540	1545
Glu Gly Ala Ile Lys Trp Asn Glu Ala Asn Arg Pro Gly Lys Val		
1550	1555	1560
Pro Phe Leu Arg Val Ala Thr Glu Ser Ser Ala Lys Thr Pro Ser		
1565	1570	1575
Lys Leu Leu Asp Pro Leu Ala Trp Asp Asn His Tyr Gly Thr Gln		
1580	1585	1590
Ile Pro Lys Glu Glu Trp Lys Ser Gln Glu Lys Ser Pro Glu Lys		
1595	1600	1605
Thr Ala Phe Lys Lys Lys Asp Thr Ile Leu Ser Leu Asn Ala Cys		
1610	1615	1620
Glu Ser Asn His Ala Ile Ala Ala Ile Asn Glu Gly Gln Asn Lys		
1625	1630	1635

Pro Glu Ile Glu Val Thr Trp Ala Lys Gln Gly Arg Thr Glu Arg
 1640 1645 1650

Leu Cys Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg Glu
 1655 1660 1665

Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr
 1670 1675 1680

Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile
 1685 1690 1695

Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys
 1700 1705 1710

Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr
 1715 1720 1725

Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser
 1730 1735 1740

Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr
 1745 1750 1755

Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu
 1760 1765 1770

His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp
 1775 1780 1785

Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser
 1790 1795 1800

Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly
 1805 1810 1815

Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr
 1820 1825 1830

Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu
 1835 1840 1845

Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu
 1850 1855 1860

Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His

1865	1870	1875
Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln		
1880	1885	1890
Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp		
1895	1900	1905
Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn		
1910	1915	1920
Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His		
1925	1930	1935
Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met		
1940	1945	1950
Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser		
1955	1960	1965
Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr		
1970	1975	1980
Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr		
1985	1990	1995
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly		
2000	2005	2010
Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly		
2015	2020	2025
Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro		
2030	2035	2040
Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala		
2045	2050	2055
Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His		
2060	2065	2070
Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser		
2075	2080	2085
Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile		
2090	2095	2100

Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser
 2105 2110 2115
 Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr
 2120 2125 2130

 Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn
 2135 2140 2145
 Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile
 2150 2155 2160
 Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg
 2165 2170 2175
 Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys
 2180 2185 2190

 Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln
 2195 2200 2205
 Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser
 2210 2215 2220
 Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp
 2225 2230 2235
 Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe
 2240 2245 2250

 Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys
 2255 2260 2265
 Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser
 2270 2275 2280
 Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys
 2285 2290 2295
 Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val
 2300 2305 2310

 Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His
 2315 2320 2325
 Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu

cccgacgaca ggagctacaa gagccagtac ctgaacaacg gccccagag gatcggcagg 1320

 aagtacaaga aggtcagatt catggcctac accgacgaga cttcaagac cagggaggcc 1380
 atccagcagc agtctggcat cctgggcccc ctgctgtacg gcgaggtggg cgacaccctg 1440
 ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctaccccca cggcatcacc 1500
 gatgtgagge ccctgtacag caggaggctg cccaagggcg tgaagcacct gaaggacttc 1560
 cccatcctgc cggcgagat cttcaagtac aagtggaccg tgaccgtgga ggatggcccc 1620
 accaagtctg accccaggctg cctgaccagg tactacagca gcttcgtgaa catggagagg 1680
 gacctggcct ctggcctgat cggccccctg ctgatctgct acaaggagag cgtggaccag 1740

 aggggcaacc agatcatgtc tgacaagagg aacgtgatcc tgttctctgt gttcgtgag 1800
 aacaggagct ggtatctgac cgagaacatc cagaggttcc tgccaaccc cgccggcgtg 1860
 cagctggagg accccgagtt ccaggccagc aacatcatgc acagcatcaa cggctacgtg 1920
 ttcgacagcc tgcagctgtc tgtgtgcctg cagaggtgg cctactggta catcctgagc 1980
 atcggcgccc agaccgactt cctgtctgtg ttcttctctg gctacacctt caagcacaag 2040
 atggtgtacg aggacacctt gacctgttc cccttcagcg gcgagaccgt gttcatgagc 2100
 atggagaacc ccggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc 2160

 atgaccgccc tgetgaaagt cagcagctgc gacaagaaca ccggcgacta ctacaggac 2220
 agctacgagg acatcagcgc ctacctgtg agcaagaaca acgcatcga gcccaggagc 2280
 ttcagccaga accccccgt gctgaagagg caccagaggg agatcaccag gaccaccctg 2340
 cagagcgacc aggaggagat cgactatgat gacacatca gcgtggagat gaagaaggag 2400
 gacttcgaca tctacgacga ggacgagaac cagagcccca ggagcttcca gaagaagacc 2460
 aggactact tcatcgccg cgtggagagg ctgtgggact atggcatgag cagcagcccc 2520
 cacgtgctga ggaacagggc ccagagcggc agcgtgcccc agttcaagaa gttggtgttc 2580

 caggagtcca ccgacggcag cttcaccag ccctgtaca gagcgagct gaacgagcac 2640
 ctgggcctgc tgggccccta catcagggcc gaggtggagg acaacatcat ggtgaccttc 2700
 aggaaccagg ccagcaggcc ctacagcttc tacagcagcc tgatcagcta cgaggaggac 2760
 cagaggcagg gcgccgagcc caggaagaac ttcgtgaagc ccaacgagac caagacctac 2820
 ttctggaagg tgcagacca catggcccc accaaggacg agttcgactg caaggcctgg 2880
 gcctacttct ctgatgtgga cctggagaag gacgtgcaca gcggcctgat cggccccctg 2940
 ctggtgtgcc acaccaacac cctgaacccc gccacggca ggcaggtgac cgtgcaggag 3000

ttcgccctgt tcttcacat cttcgacgag accaagagct ggtacttcac cgagaacatg 3060
 gagaggaact gcagggcccc ctgcaacatc cagatggagg accccacctt caaggagaac 3120
 tacaggttcc acgccatcaa cggtacatc atggacaccc tgcccggcct ggtgatggcc 3180
 caggaccaga ggatcaggtg gtatctgctg agcatgggca gcaacgagaa catccacagc 3240
 atccacttca gcggccactg gttcacctg aggaagaagg aggagtaca gatggccctg 3300
 tacaacctgt accccggcgt gttcgagacc gtggagatgc tgcccagcaa ggccggcatc 3360
 tggaggttg agtgcctgat cggcgagcac ctgcacgccg gcatgagcac cctgttctg 3420

gtgtacagca acaagtgcc aacccccctg ggcatggcca gcggccacat cagggacttc 3480
 cagatcaccg cctctggcca gtacggccag tgggccccca agctggccag gctgactac 3540
 agcggcagca tcaacgcctg gagcaccaag gagcccttca gctggatcaa ggtggacctg 3600
 ctggccccca tgatcatca cggcatcaag acccagggcg ccaggcagaa gttcagcagc 3660
 ctgtacatca gccagttcat catcatgtac agcctggacg gcaagaagtg gcagacctac 3720
 aggggcaaca gcaccggcac cctgatggtg ttcttcggca acgtggacag cagcggcatc 3780
 aagcacaaca tcttcaacc ccccatcatc gccaggfaca tcaggctgca ccccaccac 3840

tacagcatca ggagcacct gcggatgga ctgatgggct gcgacctgaa cagctgcagc 3900
 atgccctgg gcatggagag caagccatc tctgacgcc agatcaccgc cagcagctac 3960
 ttcaccaaca ttttcgcc ctggagcccc agcaaggcca ggctgcacct gcagggcagg 4020
 agcaacgctt ggaggcccc ggtgaacaac cccaaggagt ggctgcaggt ggacttccag 4080
 aagaccatga aggtgaccgg cgtgaccacc cagggcgtga agagcctgct gaccagcatg 4140
 tacgtgaagg agttcctgat cagcagcagc caggacggcc accagtggac cctgttcttc 4200
 cagaacggca aagtgaaggt gttccagggc aaccaggaca gttcaccac cgtggtgaac 4260

agcctggacc cccccctgct gaccaggtat ctgaggatcc acccccagag ctgggtgcac 4320
 cagatcgccc tgagaatgga agtgctggga tgcgaggccc aggacctgta ctga 4374

<210> 21

<211> 1457

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polypeptide

<400> 21

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

1 5 10 15

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

 Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
 35 40 45
 Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
 50 55 60
 Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
 65 70 75 80
 Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
 85 90 95

 Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
 100 105 110
 His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
 115 120 125
 Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
 130 135 140
 Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
 145 150 155 160

 Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
 165 170 175
 Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
 180 185 190
 Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
 195 200 205
 Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
 210 215 220

 Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
 225 230 235 240
 Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255
 Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val

260 265 270
 Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285

 Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300
 Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320
 Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335
 Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

 Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365
 Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380
 Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400
 Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

 Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430
 Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445
 Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460
 Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480

 Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495
 His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510

Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525

Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
 565 570 575

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
 580 585 590

Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605

Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620

Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640

Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
 645 650 655

Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
 660 665 670

Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
 675 680 685

Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700

Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720

Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735

Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
 740 745 750

Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu

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

Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn
 1010 1015 1020
 Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys
 1025 1030 1035
 Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr
 1040 1045 1050

 Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr
 1055 1060 1065
 Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe
 1070 1075 1080
 Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met
 1085 1090 1095
 Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met
 1100 1105 1110

 Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly
 1115 1120 1125
 Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser
 1130 1135 1140
 Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg
 1145 1150 1155
 Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro
 1160 1165 1170

 Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser
 1175 1180 1185
 Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro
 1190 1195 1200
 Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe
 1205 1210 1215
 Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp
 1220 1225 1230

 Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu

1235	1240	1245
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1250	1255	1260
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1265	1270	1275
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1280	1285	1290
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1295	1300	1305
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1310	1315	1320
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1325	1330	1335
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1340	1345	1350
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1355	1360	1365
Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys		
1370	1375	1380
Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu		
1385	1390	1395
Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp		
1400	1405	1410
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr		
1415	1420	1425
Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala		
1430	1435	1440
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr		
1445	1450	1455

<210> 22

<211> 2220

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 22

```

gccaccagga gatactacct gggcgccgtg gagctgagct gggactacat gcagtctgac      60
ctgggcgagc tgctgtgga cgccaggttc cccccagag tgccaagag cttccccttc      120
aacacctcag tgggtgataa gaagaccctg ttcgtggagt tcaccgacca cctgttcaac      180
atcgccaage ccaggcccc ctggatgggc ctgctgggcc ccaccatcca ggccgaggtg      240
tacgacaccg tggtgatcac cctgaagaac atggccagcc acccctgag cctgcacgcc      300
gtggcgctga gctactggaa ggcctctgag ggcgccgagt atgacgacca gaccagccag      360
agggagaagg aggacgacaa ggtgttcccc ggcggcagcc acacctacgt gtggcaggtg      420

ctgaaggaga acggcccat ggccagcgac ccctgtgcc tgacctacag ctacctgagc      480
cacgtggacc tggngaagga cctgaactct ggcctgatcg ggcctctgct ggtgtgcagg      540
gaggcgacc tggccaagga gaagaccag accctgcaca agttcatcct getgttccgc      600
gtgttcgatg agggcaagag ctggcacagc gagaccaaga acagcctgat gcaggacagg      660
gatgccgct ctgccagggc ctggccaag atgcacaccg tgaacggcta cgtgaacagg      720
agcctgcccg gectgatcgg ctgccacagg aagtctgtgt actggcacgt gatcgcatg      780
ggcaccacc cagaggtgca cagcatcttc ctggagggcc acaccttct ggtgaggaac      840

cacaggcagg ccagcctgga gatcagcccc atcaccttcc tgaccgcca gaccctgctg      900
atggacctgg gccagttcct gctgttctgc cacatcagca gccaccagca cgacggcatg      960
gaggcctacg tgaaggtgga cagctgcccc gaggagcccc agctgaggat gaagaacaac      1020
gaggaggccg aggactatga tgatgacctg accgactctg agatggacgt ggtgaggttt      1080
gatgatgaca acagccccag cttcatccag atcaggtctg tggccaagaa gcacccaag      1140
acctgggtgc actacatcgc cgccgaggag gaggactggg actacgcccc cctggtgctg      1200
gccccgacg acaggagcta caagagccag tacctgaaca acggccccca gaggatcggc      1260

aggaagtaca agaaggtcag attcatggcc tacaccgacg agaccttaa gaccaggag      1320
gccatccagc acgagtctgg catctgggc cccctgctgt acggcgaggt gggcgacacc      1380
ctgctgatca tctcaagaa ccaggccagc aggcctaca acatctacc ccacggcatc      1440
accgatgtga ggcccctgta cagcaggagg ctgccaagg gcgtgaagca cctgaaggac      1500
ttcccctcc tgcccggcga gatcttcaag tacaagtga ccgtgaccgt ggaggatggc      1560
cccaccaagt ctgaccagc gtgcctgacc aggtactaca gcagcttctg gaacatggag      1620

```

agggacctgg ccctcggcct gatcggcccc ctgctgatct gctacaagga gagcgtggac 1680

 cagaggggca accagatcat gtctgacaag aggaacgtga tcctgttctc tgtgttcgat 1740
 gagaacagga gctggtatct gaccgagaac atccagaggt tcctgcccaa ccccgcgggc 1800
 gtgcagctgg aggaccccga gttccaggcc agcaacatca tgcacagcat caacggctac 1860
 gtgttcgaca gcctgcagct gtctgtgtgc ctgcacgagg tggcctactg gtacatcctg 1920
 agcatcggcg ccagaccga ctctctgtct gtgttcttct ctggctacac cttcaagcac 1980
 aagatggtgt acgaggacac cctgacctg ttccccttca gcgcgagac cgtgttcattg 2040
 agcatggaga accccggcct gtggatcctg ggctgccaca acagcgactt caggaacagg 2100

 ggcatgaccg ccctgctgaa agtcagcagc tgcgacaaga acaccggcga ctactacgag 2160
 gacagctacg aggacatcag cgcctacctg ctgagcaaga acaacgccat cgagcccagg 2220

 <210> 23
 <211> 2052
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 polynucleotide
 <400> 23

 gagatcacca ggaccacct gcagagcgac caggaggaga tcgactatga tgacaccatc 60
 agcgtggaga tgaagaagga ggacttcgac atctacgacg aggacgagaa ccagagcccc 120
 aggagcttcc agaagaagac caggcactac ttcattgccc ccgtggagag gctgtgggac 180

 tatggcatga gcagcagccc ccacgtgctg aggaacaggg cccagagcgg cagcgtgccc 240
 cagttcaaga agtggtgttt ccaggagttc accgacggca gttcacca gecctgtac 300
 agaggcgagc tgaacgagca cctgggcctg ctgggccctt acatcagggc cgaggaggag 360
 gacaacatca tggtagcctt caggaaccag gccagcaggc cctacagctt ctacagcagc 420
 ctgatcagct acgaggagga ccagaggcag ggcgccgagc ccaggaagaa cttcgtgaag 480
 cccaacgaga ccaagaccta cttctggaag gtgcagcacc acatggcccc caccaaggac 540
 gatttcgact gcaaggcctg ggcctacttc tctgatgtgg acctggagaa ggacgtgcac 600

 agcggcctga tcggccccct gctggtgtgc cacaccaaca ccctgaacce cgcccacggc 660
 aggcaggtga ccgtgcagga gttcgcctg ttcttcacca tcttcgacga gaccaagagc 720
 tggtagctca ccgagaacat ggagaggaac tgcaggcccc cctgcaacat ccagatggag 780
 gacccacct tcaaggagaa ctacaggttc cagccatca acggctacat catggacacc 840

ctgccccggcc tggatgatggc ccaggaccag aggatcaggt ggtatctgct gagcatgggc 900
agcaacgaga acatccacag catccacttc agcggccacg tgttcaccgt gaggaagaag 960
gaggagtaca agatggccct gtacaacctg taccgccgcg tgttcgagac cgtggagatg 1020

ctgcccagca aggccggcat ctggagggtg gagtgcctga tcggcgagca cctgcacgcc 1080
ggcatgagca cctgtttctt ggtgtacagc aacaagtgcc agacccccct gggcatggcc 1140
agcggccaca tcagggactt ccagatcacc gcctctggcc agtacggcca gtgggcccc 1200
aagctggcca ggctgacta cagcggcagc atcaacgctt ggagcaccaa ggagcccttc 1260
agctggatca aggtggacct gctggcccc atgatcatcc acggcatcaa gaccagggc 1320
gccaggcaga agttcagcag cctgtacatc agccagtca tcatcatgta cagcctggac 1380
ggcaagaagt ggcagaccta caggggcaac agcaccggca ccctgatggt gttcttcggc 1440

aacgtggaca gcagcggcat caagcacaac atcttcaacc ccccatcat cgccaggtag 1500
atcaggctgc accccacca ctacagcatc aggagcacc tgcgatgga actgatgggc 1560
tgcgacctga acagctgcag catgccctg ggcattgaga gcaaggccat ctctgacgcc 1620
cagatcaccg ccagcagcta ctccaccaac atgttcgcca cctggagccc cagcaaggcc 1680
aggctgcacc tgcagggcag gagcaacgcc tggaggcccc aggtgaacaa cccaaggag 1740
tggctgcagg tggacttcca gaagaccatg aaggtgaccg gcgtgaccac ccaggcctg 1800
aagagcctgc tgaccagcat gtacgtgaag gatttctga tcagcagcag ccaggacggc 1860

caccagtgga cctgtttctt ccagaacggc aaagtgaagg tgttcaggg caaccaggac 1920
agcttcaccc ccgtggtgaa cagcctggac cccccctgc tgaccaggta tetgaggatc 1980
cacccccaga gctgggtgca ccagatcgcc ctgagaatgg aagtgctggg atgcgaggcc 2040
caggacctgt ac 2052

<210> 24
<211> 2220
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 24
gccaccagga gatactacct gggggctgtg gaactttctt gggactacat gcagtctgac 60
ctgggagagc tgctgtgga tgccaggctc ccaccagag tgeccaagtc ctcccatc 120
aacacctctg tggctctaca gaagacactc tttgtggaat tcaactgacca cctgttcaac 180

attgcaaac ccagaccacc ctggatggga ctctctgggac ccaccattca ggctgaggtg 240
 tatgacactg tggatcac cctcaagaac atggcatccc accctgtgtc tctgcatgct 300
 gtgggagtct catactggaa agcctctgaa ggggctgagt atgatgacca gacatcccag 360
 agagagaaag aggatgacaa ggtgttcctt gggggatctc acacctatgt gtggcaagtc 420
 ctcaaggaga atggacccat ggcatctgac ccactctgcc tgacatactc ctacctttct 480

 catgtggacc tggcaagga cctcaactct ggactgattg gggcactgct ggtgtgcagg 540
 gaaggatccc tggccaagga gaaaaccag aactgcaca agttcattct cctgtttgct 600
 gtctttgatg agggcaagtc ttggcactct gaaacaaaga actccctgat gcaagacagg 660
 gatgctgcct ctgccagggc atggcccaag atgcacactg tgaatggcta tgtgaacaga 720
 tcaactgctg gactcattgg ctgccacagg aaatctgtct actggcatgt gattggcatg 780
 gggacaacc ctgaagtga ctccattttc ctggaggggac acaccttctt ggtcaggaac 840
 cacagacaag cctctctgga gatctctccc atcaccttcc tcaactgcaca gacactgctg 900

 atggaccttg gacagttcct gctgttctgc cacatctctt cccaccagca tgatggcatg 960
 gaagcctatg tcaaggtgga ctcatgacct gaggaaccac agctcaggat gaagaacaat 1020
 gaggaggctg aggactatga tgatgacctg actgactctg agatggatgt ggtcagattt 1080
 gatgatgaca actctccatc ctccattcag atcaggctctg tggcaaagaa acacccaag 1140
 acatgggtgc actacattgc tgctgaggaa gaggactggg actatgcacc actggtcctg 1200
 gcccctgatg acaggagcta caagtctcag tacctcaaca atggcccaca aagaattgga 1260
 agaaagtaca agaaagtcag attcatggcc tacactgatg aaaccttcaa gacaagagaa 1320

 gccattcagc atgagtctgg cattctggga ccactcctgt atggggaagt gggagacacc 1380
 ctgctcatca tcttcaagaa ccaggcctcc aggcctaca acatctacc acatggcatc 1440
 actgatgta gcccctgta cagcaggaga ctgccaaaag gggtgaaaca cctcaaggac 1500
 ttccccattc tgcttgaga gatcttcaag tacaagtgga ctgtcactgt ggaggatgga 1560
 ccaacaaagt ctgaccacc gtgcctcacc agatactact cctcttttgt gaacatggag 1620
 agagacctgg catctggact gattggacca ctgctcatct gctacaagga gtctgtggac 1680
 cagagaggca accagatcat gtctgacaag agaaatgtga ttctgttctc tgtctttgat 1740

 gagaacagat catggtacct gactgagaac attcagagat tctgcccac cctgctggg 1800
 gtgcaactgg aagaccctga gttccaggca agcaacatca tgcaactccat caatggctat 1860
 gtgtttgact ctctccagct ttctgtctgc ctgcatgagg tggcctactg gtacattctt 1920
 tctattgggg cacaaaactg ctctctttct gtcttcttct ctggatacac ctccaagcac 1980
 aagatggtgt atgaggacac cctgacactc ttccattct ctggggaac tgtgttcatg 2040

agcatggaga accctggact gtggattctg ggatgccaca actctgactt cagaaacagg 2100
ggaatgactg cactgctcaa agtctcctcc tgtgacaaga aactgggga ctactatgag 2160

gactcttatg aggacatctc tgcctacctg ctgagcaaga acaatgcat tgagcccaga 2220
<210> 25
<211> 2052
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 25

gagatcacca ggacaacct ccagctgac caggaagaga ttgactatga tgacaccatt 60
tctgtggaga tgaagaagga ggactttgac atctatgatg aggacgagaa ccagtctcca 120
agatcattcc agaagaagac aagacactac tcatttctg ctgtgaaag actgtgggac 180
tatggcatgt cttctctcc ccattgctc aggaacaggg cacagtctgg ctctgtgcca 240

cagttcaaga aagtggctt ccaggagttc actgatggct cattcacca gccctgtac 300
agaggggaac tgaatgagca cctgggactc ctgggacat acatcaggc tgaggtggaa 360
gacaacatca tggtagacatt cagaaaccag gcctccagc cctacagct ctactctcc 420
ctcatcagct atgaggaaga ccagagaca ggggctgagc caagaaagaa ctttgtaaa 480
cccaatgaa ccaagacct cttctggaaa gtccagcacc acatggcacc caccaaggat 540
gagtttgact gcaaggcctg ggcatcttc tctgatgtgg acctggagaa agatgtgcac 600
tctggcctga ttggcccact cctggtctgc cacaccaaca ccctgaacc tgccatgga 660

aggcaagtga ctgtgcagga gtttccctc ttcttcacca tctttgatga aaccaagtea 720
tggtagtcca ctgagaacat ggagagaaac tgagagcac catgcaacat tcagatggaa 780
gacccacct tcaaggagaa ctacaggttc catgcatca atggctacat catggacacc 840
ctgcctgggc ttgtcatggc acaggaccag agaatcagat ggtacctgct ttctatggga 900
tccaatgaga acattcactc catccattc tctgggcatg tcttactgt gagaaagaag 960
gaggaataca agatggccct gtacaacctc tacctgggg tctttgagac tgtggagatg 1020
ctgcctcca aagctggcat ctggagggtg gaatgcctca ttggggagca cctgcatgct 1080

ggcatgtcaa cctgttctt ggtctacagc aacaagtcc agacaccct gggaatggcc 1140
tctggccaca tcagggactt ccagatcact gcctctggcc agtatggcca gtgggcacc 1200
aaactggcca ggctccacta ctctggctcc atcaatgcat ggtcaaccaa ggagccattc 1260

tcttggatca aggtggacct gctggcacc atgatcattc atggcatcaa gacacagggg 1320
gcaagacaga aattctcctc tctgtacatc tcacagtca tcatcatgta ctctctggat 1380
ggcaagaagt ggcagacata cagaggcaac tccactggca ccctcatggt ctctcttggc 1440
aatgtggaca gctctggcat caagcacaac atcttcaacc ctcccatcat tgccagatac 1500

atcaggctgc accccacca ctactcaatc agatcaacc tcaggatgga actgatggga 1560
tgtgacctga actcctgctc aatgcccttg ggaatggaga gcaaggccat ttctgatgcc 1620
cagatcactg catcctctta cttcaccaac atgtttgcca cctggtcacc atcaaaagcc 1680
aggctgcacc tccaggggaag aagcaatgcc tggagacccc aggtcaaca cccaaaggaa 1740
tggtgcaag tggacttcca gaagacaatg aaagtcaactg gggtgacaac ccagggggtc 1800
aagtctctgc tcacctcaat gtatgtgaag gagttcctga tctcttctc acaggatggc 1860
caccagtgga cactcttctt ccagaatggc aaagtcaagg tgttccaggg caaccaggac 1920

tctttcacac ctgtggtgaa ctactggac cccccctcc tgacaagata cctgagaatt 1980
cacccccagt ctgggtcca ccagattgcc ctgagaatgg aagtctggg atgtgaggca 2040
caagacctgt ac 2052

<210> 26
<211> 4332
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 26
atgcagattg agctgtccac ctgcttcttt ctgtgcctgc tgagattctg ctctctgcc 60
accaggagat actacctggg ggctgtggaa ctttcttggg actacatgca gctgacctg 120

ggagagctgc ctgtggatgc caggttcca ccagagtgc ccaagtcctt ccattcaac 180
acctctgtgg tctacaagaa gacctcttt gtggaattca ctgaccacct gttcaacatt 240
gcaaaacca gaccacctg gatgggactc ctgggacca ccattcagc tgagggtat 300
gacctgtgg tcatccctc caagaacatg gcatcccacc ctgtgtctct gcatgctgtg 360
ggagtctcat actgaaagc ctctgaaggg gctgagtatg atgaccagac atcccagaga 420
gagaaagagg atgacaagg gttccctggg ggatctcaca cctatgtgtg gcaagtcctc 480
aaggagaatg gaccatggc atctgacca ctctgcctga catactccta cttttctcat 540

gtggacctgg tcaaggacct caactctgga ctgattgggg cactgctggt gtgcaggaa 600

ggatccctgg ccaaggagaa aaccagaca ctgcacaagt tcattctcct gtttgctgtc 660
 ttgatgagg gcaagtcttg gcactctgaa acaaagaact ccctgatgca agacagggat 720
 gctgcctctg ccagggcatg gccaagatg cacactgtga atggctatgt gaacagatca 780
 ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg 840
 acaaccctg aagtgcactc cattttctg gagggacaca ctttctggt caggaaccac 900
 agacaagcct ctctggagat ctctcccatc accttctca ctgcacagac actgctgatg 960

 gaccttggac agttcctgct gttctgccac atctcttccc accagcatga tggcatggaa 1020
 gcctatgca aggtggactc atgccctgag gaaccacagc tcaggatgaa gaacaatgag 1080
 gaggtgagg aciataatga tgacctgact gactctgaga tggatgtggt cagatttgat 1140
 gatgacaact ctccatcctt cattcagatc aggtctgtgg caaagaaaca cccaagaca 1200
 tgggtgact acattgctgc tgaggaagag gactgggact atgcaccact ggtcctggcc 1260
 cctgatgaca ggagctacaa gtctcagtac ctcaacaatg gccacaaaag aattggaaga 1320
 aagtacaaga aagtcagatt catggcctac actgatgaaa cttcaagac aagagaagcc 1380

 attcagcatg agtctggcat tctgggacca ctctgtatg gggaagtggg agacaccctg 1440
 ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccaca tggcatcact 1500
 gatgtcagge cctgtacag caggagactg ccaaaagggg tgaaacacct caaggacttc 1560
 cccattctgc ctggagagat ctcaagtac aagtggactg tcaactgtga ggatggacca 1620
 acaaagtctg accccaggatg cctcaccaga tactactcct cttttgtgaa catggagaga 1680
 gacctggcat ctggactgat tggaccactg ctcatctgct acaaggagt cgtggaccag 1740
 agaggcaacc agatcatgct tgacaagaga aatgtgattc tgttctctgt ctttgatgag 1800

 aacagatcat ggtacctgac tgagaacatt cagagattcc tgccaacce tgcctgggtg 1860
 caactggaag acctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgtg 1920
 ttgactctc tcagcttcc tgtctgctg catgaggtgg cctactggta cattcttct 1980
 attggggcac aaactgactt ctttctgctc ttcttctctg gatacactt caagcacaag 2040
 atggtgatg aggacacct gacactctc ccattctctg gggaaactgt gttcatgagc 2100
 atggagaacc ctggactgtg gattctggga tgccacaact ctgacttcag aacagggga 2160
 atgactgcac tgcctaaagt ctctcctgt gacaagaaca ctggggacta ctataggac 2220

 tcttatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gccagagag 2280
 atcaccagga caacctcca gtctgaccag gaagagattg actatgatga caccattct 2340
 tgggatgata agaaggagga ctttgacatc tatgatgagg acgagaacca gtctccaaga 2400
 tcattccaga agaagacaag acactactc attgctgctg tggaaagact tggggactat 2460

ggcatgtctt ccctcccaca tgtctcagg aacagggcac agtctggctc tgtgccacag 2520
 ttcaagaaag tggctctcca ggagttcact gatggctcat tcaccagcc cctgtacaga 2580
 ggggaactga atgagcacct gggactcctg ggaccataca tcagggtga ggtggaagac 2640

 aacatcatgg tgacattcag aaaccaggcc tccaggcctt acagcttcta ctcttcctc 2700
 atcagctatg aggaagacca gagacaaggg gctgagccaa gaaagaactt tgtgaaacce 2760
 aatgaaacca agacctactt ctggaaagtc cagcaccaca tggcaccac caaggatgag 2820
 ttgactgca aggctgggc atacttctct gatgtggacc tggagaaaga tgtgactct 2880
 ggctgattg gccactcct ggtctgccac accaacaccc tgaacctgc acatggaagg 2940
 caagtactg tgcaggagt tgcctcttc ttcacatct ttgatgaaac caagtcatgg 3000
 tacttactg agaacatgga gagaaactgc agagcacat gcaacattca gatggaagac 3060

 cccacttca aggagaacta caggttccat gccatcaatg gctacatcat ggacacctg 3120
 cctgggcttg tcatggcaca ggaccagaga atcagatgg acctcttctc tatgggatcc 3180
 aatgagaaca ttactccat ccacttctct gggcatgtct tcaactgtgag aaagaaggag 3240
 gaatacaaga tggccctgta caacctctac cctggggtct ttgagactgt ggagatgctg 3300
 cctccaaag ctggcatctg gaggttgaa tgcctcattg gggagcacct gcatgctggc 3360
 atgtcaacce tgttcttgg ctacagcaac aagtgccaga caccctggg aatggcctct 3420
 ggccacatca gggacttcca gateactgcc tctggccagt atggccagt ggcaaccaaa 3480

 ctggccagge tccactactc tggctccatc aatgcatgg caaccaagga gccattctct 3540
 tggatcaagg tggacctgct ggacccatg atcattcatg gcatcaagac acagggggca 3600
 agacagaaat tctctctct gtacatctca cagttcatca tcatgtactc tctggatggc 3660
 aagaagtggc agacatacag aggcaactcc actggcacc tcatggtctt ctttggcaat 3720
 gtggacagct ctggcatcaa gcacaacatc tcaacctc ccatcattgc cagatacatc 3780
 aggctgcacc ccaccacta ctcaatcaga tcaacctca ggatggaact gatgggatgt 3840
 gacctgaact cctgctcaat gccctggga atggagagca aggccatttc tgatgccag 3900

 atcactgcat cccttactt caccaacatg tttgccacct ggtcaccatc aaaagccagg 3960
 ctgcacctc aggaagaag caatgcctgg agacccagg tcaacaacc aaaggaatgg 4020
 ctgcaagtgg acttccagaa gacaatgaaa gtcaactggg tgacaacca gggggtcaag 4080
 tctctgctca ccicaatgta tgtgaaggag ttctgatct cttctcaca ggatggccac 4140
 cagtggacac tcttctcca gaatggcaaa gtcaaggtgt tccagggcaa ccaggactct 4200
 ttcacacctg tggatgaact actggacccc ccctctga caagatact gagaattcac 4260

ccccagtctt ggggccacca gattgccctg agaatggaag tcctgggatg tgaggcacia 4320

gacctgtact ga 4332

<210> 27

<211> 4368

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 27

atgcagattg agctgtccac ctgcttcttt ctgtgcctgc tgagattctg cttctctgcc 60

accaggagat actacctggg ggctgtggaa ctttcttggg actacatgca gctgacctg 120

ggagagctgc ctgtggatgc caggttccca cccagagtgc ccaagtcctt cccattcaac 180

acctctgtgg tctacaagaa gacctcttt gtggaattca ctgaccacct gttcaacatt 240

gcaaaaccca gaccacctg gatgggactc ctgggacca ccattcaggc tgaggtgtat 300

gacctgtgg tcatcacct caagaacatg gcatcccacc ctgtgtctct gcatgctgtg 360

ggagtctcat actggaaagc ctctgaaggg gctgagtatg atgaccagac atcccagaga 420

gagaaagagg atgacaaggt gtccctggg ggatctcaca cctatgtgtg gcaagtcctc 480

aaggagaatg gaccatggc atctgacca ctctgcctga catactccta ctttctcat 540

gtggacctgg tcaaggacct caactctgga ctgattgggg cactgctggt gtgcagggaa 600

ggatccctgg ccaaggagaa aaccagaca ctgcacaagt tcattctcct gtttctgtc 660

tttgatgagg gcaagtcttg gcactctgaa acaaagaact ccctgatgca agacagggat 720

gctgcctctg ccaggcatg gcccaagatg cacactgtga atggctatgt gaacagatca 780

ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg 840

acaaccctg aagtgcactc cattttctg gagggacaca ctttctggt caggaaccac 900

agacaagcct ctctggagat ctctccatc accttctca ctgcacagac actgctgatg 960

gaccttggac agttctgct gttctgccac atctcttccc accagcatga tggcatggaa 1020

gcctatgta aggtggactc atgcctgag gaaccacagc tcaggatgaa gaacaatgag 1080

gaggctgagg aciatgatga tgacctgact gactctgaga tggatgtggt cagatttgat 1140

gatgacaact ctccatcctt cattcagatc aggtctgtgg caaagaaaca cccaagaca 1200

tgggtgact acattgtctg tgaggaagag gactgggact atgcaccact ggtcctggcc 1260

cctgatgaca ggagctacaa gtctcagtac ctcaacaatg gcccaaaag aattggaaga 1320

aagtacaaga aagtcagatt catggcctac actgatgaaa ctttcaagac aagagaagcc 1380
attcagcatg agtctggcat tctgggacca ctctgtatg gggaagtggg agacaccctg 1440
ctcatcatct tcaagaacca ggctccagg ccctacaaca tctaccaca tggcatcact 1500

gatgtcaggc ccctgtacag caggagactg caaaagggg tgaacacct caaggacttc 1560
cccatcttgc ctggagagat cttcaagtac aagtggactg tcaactgtga ggatggacca 1620
acaaagtctg accccagggtg cctcaccaga tactactcct cttttgtgaa catggagaga 1680
gacctggcat ctggactgat tggaccactg ctcatctgct acaaggagt c tgtggaccag 1740
agaggcaacc agatcatgtc tgacaagaga aatgtgattc tgttctctgt ctttgatgag 1800
aacagatcat ggtacctgac tgagaacatt cagagattcc tgccaaccc tgctggggtg 1860
caactggaag accctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgtg 1920

tttgactctc tccagcttcc tgtctgcctg catgaggtgg cctactggta cattctttct 1980
attggggcac aaactgactt ctttctgtc ttcttctctg gatacacctt caagcacaag 2040
atggtgtatg aggacacct gacactcttc ccattctctg gggaaactgt gttcatgagc 2100
atggagaacc ctggactgtg gattctggga tgccacaact ctgacttcag aaacagggga 2160
atgactgcac tgctcaaagt ctctcctgt gacaagaaca ctggggacta ctataggac 2220
tcttatgagg acatctctgc ctacctctc agcaagaaca atgccattga gccagaagc 2280
ttctctcaga attccagaca ccccagcacc agggagatca ccaggacaac cctccagtct 2340

gaccaggaag agattgacta tgatgacacc atttctgtgg agatgaagaa ggaggacttt 2400
gacatctatg atgaggacga gaaccagtct ccaagatcat tccagaagaa gacaagacac 2460
tacttcattg ctgctgtgga aagactgtgg gactatggca tgtcttctc tccccatgtc 2520
ctcaggaaca gggcacagtc tggctctgtg ccacagtca agaaagtggc cttccaggag 2580
ttcaactgatg gtcattcac ccageccctg tacagagggg aactgaatga gcacctggga 2640
ctcctgggac calacatcag ggctgaggtg gaagacaaca tcatggtgac attcagaaac 2700
caggcctcca ggcctacag cttctactct tcctcatca gctatgagga agaccagaga 2760

caaggggctg agccaagaaa gaactttgtg aaaccaatg aaaccaagac ctacttctgg 2820
aaagtccagc accacatggc accaccaag gatgagttg actgcaagc ctgggcatac 2880
ttctctgatg tggacctgga gaaagatgtg cactctggcc tgattggccc actcctggtc 2940
tgccacacca acacctgaa ccctgcacat ggaaggcaag tgactgtgca ggagtttgcc 3000
ctcttctca ccatcttga tgaaccaag tcatggtact tcactgagaa catggagaga 3060
aactgcagag caccatgcaa cattcagatg gaagacceca cttcaagga gaactacagg 3120

ttccatgcc a tcaatggcta catcatggac accctgcctg ggcttgtcat ggcacaggac 3180

cagagaatca gatggtacct gctttctatg ggatccaatg agaacattca ctccatccac 3240

ttctctgggc atgtcttcac tgtgagaaaag aaggaggaat acaagatggc cctgtacaac 3300

ctctaccctg gggctcttga gactgtggag atgctgcct ccaaagctgg catctggagg 3360

gtggaatgcc tcattgggga gcaacctgcat gctggcatgt caacctgtt cctggcttac 3420

agcaacaagt gccagacacc cctgggaatg gcctctggcc acatcagga cttccagatc 3480

actgcctctg gccagtatgg ccagtgggca cccaaactgg ccaggctcca ctactctggc 3540

tccatcaatg catggtcaac caaggagcca ttctcttga tcaagtgga cctgctggca 3600

cccatgatca ttcatggcat caagacacag ggggcaagac agaaattctc ctctctgtac 3660

atctcacagt tcatcatcat gtactctctg gatggcaaga agtggcagac atacagagge 3720

aactccactg gcacctcat ggtcttcttt ggcaatgtgg acagctctgg catcaagcac 3780

aacatcttca acctcccat cattgccaga tacatcagge tgcacccac cactactca 3840

atcagatcaa cctcaggat ggaactgatg ggatgtgacc tgaactctg ctcaatgccc 3900

ctgggaatgg agagcaagge cattctgat gccagatca ctgcatctc ttacttcacc 3960

aacatgtttg ccacctggtc accatcaaaa gccaggctgc acctccaggg aagaagcaat 4020

gcctggagac cccaggtcaa caaccxaaag gaatggctgc aagtggactt ccagaagaca 4080

atgaaagtca ctggggtgac aaccagggg gtcaagtctc tgctcacctc aatgtatgtg 4140

aaggagtcc tgatctctc ctcacaggat ggccaccagt ggacactctt cttccagaat 4200

ggcaaagtca aggtgttcca gggcaaccag gactcttca cacctgtggt gaactcactg 4260

gaccccccc tctgacaag atacctgaga attcacccc agtcttgggt ccaccagatt 4320

gcctgagaa tgaagtctt gggatgtgag gcacaagacc tgtactga 4368

<210> 28

<211> 4332

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 28

atgcagattg agctgagcac ctgcttctc ctgtgcctgc tgaggttctg cttctctgcc 60

accaggagat actacctggg gcgcgtggag ctgagctggg actacatgca gtctgacctg 120

ggcagactgc ctgtggacgc caggttcccc cccagagtgc ccaagagctt cccttcaac 180

acctcagtgg tgiacaagaa gaccctgttc gtggagtca ccgaccacct gttcaacatc 240
 gccaaagccca ggccccctg gatgggctg ctgggccccca ccatccaggc cgaggtgtac 300
 gacaccgtgg tgatcacct gaagaacatg gccagccacc ccgtgagcct gcacgccgtg 360

 ggcgtgagct aciikkaaggc ctctgagggc gccgagtatg acgaccagac cagccagagg 420
 gagaaggagg acgacaaggt gttccccggc ggccagccaca cctacgtgtg gcaggtgtg 480
 aaggagaacg gccccatggc cagcagcccc ctgtgcctga cctacagcta cctgagccac 540
 gtggacctgg tgaaggacct gaactctggc ctgatcggcg ccctgctggt gtgcaggagg 600
 ggccagcctgg ccaaggagaa gaccagacc ctgcacaagt tcatcctgct gttcgccgtg 660
 ttcgatgagg gcaagagctg gcacagcag accaagaaca gcctgatgca ggacagggat 720
 gccgcctctg ccagggcctg gcccaagatg cacaccgtga acggctacgt gaacaggagc 780

 ctgccccggc tgatcggtc ccacaggaag tetgtgtact ggcacgtgat cggcatgggc 840
 accacccccg aggtgcacag catcttctg gagggccaca ccttctggt gaggaaccac 900
 aggcaggcca gcctggagat cagccccatc accttctga ccgccagac cctgctgatg 960
 gacctgggc agttctgct gttctgccc atcagcagcc accagcacga cggcatggag 1020
 gcctactga aggtggacag ctgccccgag gagccccagc tgaggatgaa gaacaacgag 1080
 gaggccgagg actatgatga tgacctgacc gactctgaga tggacgtggt gaggtttgat 1140
 gatgacaaca gcccagctt catccagatc aggtctgtgg ccaagaagca cccaagacc 1200

 tgggtgcaact acatcgccg cgaggaggag gactgggact acgccccct ggtgctggcc 1260
 cccgacgaca ggagctaaa gagccagtac ctgaacaacg gccccagag gatcggcagg 1320
 aagtacaaga aggtcagatt catggcctac accgacgaga ccttcaagac caggagggcc 1380
 atccagcagc agtctggcat cctgggcccc ctgctgtacg gcgaggtggg cgacaccctg 1440
 ctgatcatct tcaagaacca ggccagcagg cctacaaca tctaccccc cggcatcacc 1500
 gatgtgaggc ccctgtacag caggaggctg cccaaggcg tgaagcacct gaaggacttc 1560
 cccatcctgc ccggcgagat ctcaagtac aagtggaccg tgaccgtgga ggatggcccc 1620

 accaagtctg accccaggtg cctgaccagg tactacagca gcttctgaa catggagagg 1680
 gacctggcct ctggcctgat cggccccctg ctgatctgct acaaggagag cgtggaccag 1740
 aggggcaacc agatcatgic tgacaagagg aacgtgatcc tgttctctgt gttcgatgag 1800
 aacaggagct ggiatctgac cgagaacatc cagaggttcc tgccaaccc cgccggcgtg 1860
 cagctggagg accccgagtt ccaggccagc aacatcatgc acagcatcaa cggctacgtg 1920
 ttcgacagcc tgcagctgic tgtgtgcctg cacgaggtgg cctactggtc catcctgagc 1980

atcggcgccc agaccgactt cctgtctgtg ttcttctctg gctacacctt caagcacaag 2040

 atggtgtacg aggacacctt gacctgttc cccttcagcg gcgagaccgt gttcatgagc 2100
 atggagaacc cggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc 2160
 atgaccgccc tgctgaaagt cagcagctgc gacaagaaca ccggcgacta ctacaggagc 2220
 agctacgagg acatcagcgc ctacctgtg agcaagaaca acgccatcga gcccgaggag 2280
 atcaccagga ccacctgca gagcgaccag gaggagatcg actatgatga caccatcagc 2340
 gtggagatga agaaggagga cttcgacatc tacgacgagg acgagaacca gagccccagg 2400
 agcttcaga agaagaccag gcaactctt atcgccgccc tggagaggct gtgggactat 2460

 ggcatgagca gcagcccca cgtgtgagg aacagggccc agagcggcag cgtgccccag 2520
 ttcaagaagg tggigtcca ggagttacc gacggcagct tcaccagcc cctgtacaga 2580
 ggcgagctga acgagacct gggcctgtg ggcccctaca tcagggccga gttggaggac 2640
 aacatcatgg tgaccttcag gaaccaggcc agcaggccct acagcttcta cagcagcctg 2700
 atcagctacg aggaggacca gaggcagggc gccgagccca ggaagaactt cgtgaagccc 2760
 aacgagacca agacctactt ctggaagggt cagcaccaca tggccccac caaggacgag 2820
 ttcgactgca aggcctgggc ctacttctct gatgtggacc tggagaagga cgtgcacagc 2880

 ggctgatcg gccccctgt ggtgtgccac accaacacc tgaaccccgc ccacggcagg 2940
 caggtgaccg tgcaggagtt cgccctgttc ttaccatct tcgacgagac caagagctgg 3000
 tacttcaccg agaacatgga gaggaactgc agggccccct gcaacatcca gatggaggac 3060
 cccaccttea aggagaacta caggttccac gccatcaacg gctacatcat ggacaccctg 3120
 cccggcctgg tgatggccca ggaccagagg atcaggtggt atctgctgag catgggcagc 3180
 aacgagaaca tccacagcat ccacttcagc ggccacgtgt tcaccgtgag gaagaaggag 3240
 gagtacaaga tggccctgta caacctgtac cccggcgtgt tcgagaccgt ggagatgctg 3300

 cccagcaagg ccggcatctg gaggtggag tgctgatcg gcgagcacct gcacgccggc 3360
 atgagcacc tgttctgtgt gtacagcaac aagtgccaga cccccctggg catggccagc 3420
 ggccacatca gggacttcca gatcaccgcc tctggccagt acggccagtg ggcccccaag 3480
 ctggccaggc tgactacag cggcagcatc aacgcctgga gcaccaagga gccttcagc 3540
 tggatcaagg tggacctgtt ggccccatg atcatccacg gcatcaagac ccagggcgcc 3600
 aggcagaagt tcagcagcct gtacatcagc cagttcatca tcatgtacag cctggacggc 3660
 aagaagtggc agacctacag gggcaacagc accggcacc tgatggtgtt cttcgcaac 3720

gtggacagca gcggcatcaa gcacaacatc ttcaaccccc ccatcatcgc caggtacatc 3780
aggctgcacc ccaccacta cagcatcagg agcacctgc ggatggaact gatgggctgc 3840
gacctgaaca gctgcagcat gccctgggc atggagagca aggccatctc tgacgccag 3900
atcacgccca gcagctactt caccaacatg ttcgccacct ggagccccag caaggccagg 3960
ctgcacctgc agggcaggag caacgcctgg aggccccagg tgaacaacc caaggagtgg 4020
ctgcaggtgg acttccagaa gacatgaag gtgaccggcg tgaccacca gggcgtgaag 4080
agcctgctga ccagcatgta cgtgaaggag ttctgatca gcagcagcca ggacggccac 4140

cagtggaacc tgttcttcca gaacggcaaa gtgaaggtgt tccagggcaa ccaggacagc 4200
ttcacccccg tggatgaacag cctggacccc ccctgctga ccaggtatct gaggatccac 4260
ccccagagct gggatgacaa gatcgccttg agaatggaag tgctgggatg cgaggcccag 4320
gacctgtact ga 4332

<210> 29
<211> 4368
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 29
atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc 60

accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgacctg 120
ggcagactgc ctgtggacgc caggttcccc ccagagtgc ccaagagctt cccttcaac 180
acctcagtgg tgtacaagaa gacctgttc gtggagtca ccgaccacct gttcaacatc 240
gccaagccca ggccccctg gatgggcctg ctgggccccca ccatccaggc cgagggttac 300
gacacctggg tgatcacct gaagaacatg gccagccacc ccgtgacct gcacgccctg 360
ggcgtgagct actggaaggc ctctgagggc gccgagtatg acgaccagac cagccagagg 420
gagaaggagg acgacaaggt gttccccggc ggcagccaca cctacgtgtg gcagggtctg 480

aaggagaacg gcccattggc cagcagcccc ctgtgcctga cctacagcta cctgagccac 540
gtggacctgg tgaaggacct gaactctggc ctgatcggcg ccctgctggt gtgcaggag 600
ggcagcctgg ccaaggagaa gaccagacc ctgcacaagt tcatcctgct gttcgccctg 660
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgatgca ggacagggat 720
gccgcctctg ccagggcctg gcccaagatg cacacctga acggctacgt gaacaggagc 780
ctgccccgcc tgatcggctg ccacaggaag tctgtgtact ggcacgtgat cggcatgggc 840

accacccccg aggtgcacag catcttcctg gagggccaca ccttcctggt gaggaaccac 900

aggcaggcca gcctggagat cagcccatc accttcctga ccgccagac cctgctgatg 960

gacctgggce agttcctgct gttctgceac atcagcagcc accagcacga cggcatggag 1020

gcctactgta aggtggacag ctgccccgag gagccccagc tgaggatgaa gaacaacgag 1080

gaggccgagg actatgatga tgacctgacc gactctgaga tggacgtggt gaggtttgat 1140

gatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca cccaagacc 1200

tgggtgcact acatcgccgc cgaggaggag gactgggact acgccccct ggtgctggcc 1260

cccgacgaca ggagctaaa gagccagtac ctgaacaacg gccccagag gatcggcagg 1320

aagtacaaga aggtcagatt catggcctac accgacgaga ccttcaagac cagggaggcc 1380

atccagcacg agtctggcat cctgggcccc ctgctgtacg gcgaggtggg cgacaccctg 1440

ctgatcatct tcaagaacca ggccagcagg cctacaaca tctacccca cggcatcacc 1500

gatgtgaggc ccctgtacag caggaggctg cccaaggcg tgaagcacct gaaggacttc 1560

cccatcctgc cggcgagat cttcaagtac aagtggaccg tgaccgtgga ggatggcccc 1620

accaagtctg accccaggtg cctgaccagg tactacagca gcttcgtgaa catggagagg 1680

gacctggcct ctggcctgat cggccccctg ctgatctgct acaaggagag cgtggaccag 1740

aggggcaacc agatcatgtc tgacaagagg aacgtgatcc tgttctctgt gttcgtatgag 1800

aacaggagct ggiatctgac cgagaacatc cagaggttcc tgccaacce cgccggcgtg 1860

cagctggagg accccagtt ccaggccagc aacatcatgc acagcatcaa cggctacgtg 1920

ttcgacagcc tgcagctgtc tgtgtgcctg cacgaggtgg cctactggtg catcctgagc 1980

atcggcgccc agaccgactt cctgtctgtg ttcttctctg gctacacctt caagcacaag 2040

atggtgtacg aggacacct gacctgttc ccttcagcg gcgagaccgt gttcatgagc 2100

atggagaacc cggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc 2160

atgaccgcc tgctgaaagt cagcagctgc gacaagaaca ccggcgacta ctacaggac 2220

agctacgagg acatcagcgc ctacctgtg agcaagaaca acgcatcga gccaggagc 2280

ttcagccaga actccagaca cccagcacc agggagatca ccaggaccac cctgcagagc 2340

gaccaggagg agatcgacta tgatgacacc atcagcgtgg agatgaagaa ggaggacttc 2400

gacatctacg acgaggacga gaaccagagc cccaggagct tccagaagaa gaccaggcac 2460

tacttcatcg ccgccgtgga gaggctgtgg gactatggca tgagcagcag cccccactg 2520

ctgaggaaca gggcccagag cggcagcgtg cccagttca agaaggtggt gttccaggag 2580

ttcaccgacg gcagcttcac ccagcccctg tacagaggcg agctgaacga gcacctgggc 2640
 ctgctgggcc cctacatcag ggccgagggtg gaggacaaca tcatggtgac cttcaggaac 2700
 caggccagca ggccttacag cttctacagc agcctgatca gctacgagga ggaccagagg 2760
 cagggcgccg agcccaggaa gaacttcgtg aagcccaacg agaccaagac ctacttctgg 2820
 aaggtgcagc accacatggc ccccaacaag gacgagttcg actgcaagc ctgggcctac 2880
 ttctctgatg tggacctgga gaaggactg cacagcggcc tgatcgcccc cctgctggtg 2940
 tgccacacca acacctgaa ccccgccac ggaggcagg tgaccgtgca ggagttcgcc 3000

ctgtttctca ccatcttcga cgagaccaag agctggtact tcaccgagaa catggagagg 3060
 aactgcaggg ccccctgcaa catccagatg gaggacceca cttcaagga gaactacagg 3120
 ttccacgcca tcaacggcta catcatggac acctgcccc gcctggtgat ggcccaggac 3180
 cagaggatca ggtggtatct gctgagcatg ggagcaacg agaacatcca cagcatccac 3240
 ttcagcggcc acgtgttcac cgtgaggaag aaggaggagt acaagatggc cctgtacaac 3300
 ctgtaccccc gcgtgttcga gaccgtggag atgctgcca gcaaggccgg catctggagg 3360
 gtggagtgcc tgatcggcga gcacctgcac gccggcatga gcacctgtt cctggtgtac 3420

agcaacaagt gccagacccc cctgggcatg gccagcggcc acatcagga cttccagatc 3480
 accgcctctg gccagtacgg ccagtgggcc cccaagctgg ccaggctgca ctacagcggc 3540
 agcatcaacg cctggagcac caaggagccc ttcagctgga tcaagtgga cctgctggcc 3600
 cccatgatca tccacggcat caagaccag gccgccaggc agaagttcag cagcctgtac 3660
 atcagccagt tcatcatcat gtacagcctg gacggcaaga agtggcagac ctacaggggc 3720
 aacagcaccg gcacctgat ggtgttcttc ggcaactgg acagcagcg catcaagcac 3780
 aacatctca accccccat catcgccagg tacatcaggc tgcacccac cactacagc 3840

atcaggagca cctgcggat ggaactgatg ggtgcgacc tgaacagctg cagcatgccc 3900
 ctgggcatgg agagcaaggc catctctgac gccagatca ccgccagcag ctacttcacc 3960
 aacatgttcg ccacctggag cccagcaag gccaggctgc acctgcaggg caggagcaac 4020
 gcctggagge cccaggtgaa caacccaag gagtggctgc aggtggactt ccagaagacc 4080
 atgaagtgga ccggcgtgac caccagggc gtgaagagcc tgctgaccag catgtactgt 4140
 aaggagtcc tgatcagcag cagccaggac ggccaccagt ggacctgtt cttcagaac 4200
 ggcaaagtga aggtgttcca gggcaaccag gacagcttca ccccgtgtt gaacagcctg 4260

gaccccccc tgctgaccag gtatctgagg atcccccc agagctgggt gcaccagatc 4320
 gccctgagaa tggaagtgtt gggatgagc gccagggacc tgtactga 4368

<210> 30

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 30

Ser Phe Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg

1 5 10

<210> 31

<211> 24

<212> PRT

<213> Sus sp.

<400> 31

Ser Phe Ala Gln Asn Ser Arg Pro Pro Ser Ala Ser Ala Pro Lys Pro

1 5 10 15

Pro Val Leu Arg Arg His Gln Arg

20

<210> 32

<211> 16

<212> PRT

<213> Sus sp.

<400> 32

Ser Phe Ser Gln Asn Ser Arg His Gln Ala Tyr Arg Tyr Arg Arg Gly

1 5 10 15