



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0160052
(43) 공개일자 2022년12월05일

(51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) C07K 14/47 (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/86 (2013.01)
C07K 14/47 (2013.01)
(21) 출원번호 10-2022-7037261
(22) 출원일자(국제) 2021년03월26일
심사청구일자 없음
(85) 번역문제출일자 2022년10월25일
(86) 국제출원번호 PCT/EP2021/058016
(87) 국제공개번호 WO 2021/191447
국제공개일자 2021년09월30일
(30) 우선권주장
20382236.6 2020년03월26일
유럽특허청(EPO)(EP)

(71) 출원인
스플라이스바이오 에스.엘.
스페인 08028 바르셀로나 10-12 칼레 발디리 레이
삭 파크 싸이언티픽 드 바르셀로나
(72) 발명자
프루토스 도밍구에즈 실비아
스페인 08028 바르셀로나 파크 싸이언티픽 드 바
르셀로나 칼레 발디리 레이삭 10-12 스플라이스바
이오 에스.엘.
카엘레즈 비달 제러드
스페인 08028 바르셀로나 파크 싸이언티픽 드 바
르셀로나 칼레 발디리 레이삭 10-12 스플라이스바
이오 에스.엘.
(뒷면에 계속)
(74) 대리인
유미특허법인

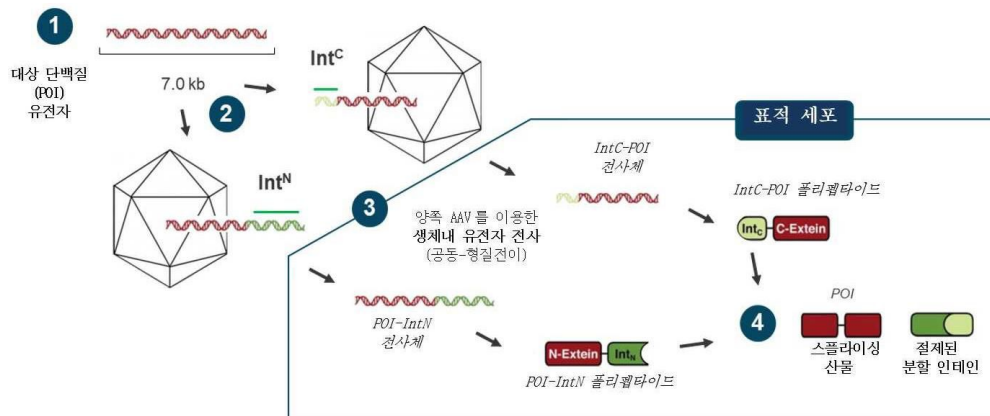
전체 청구항 수 : 총 30 항

(54) 발명의 명칭 분할 인테인 및 그 용도

(57) 요약

본 발명은 유전자 요법을 위한 거대 유전자를 재구성하기 위한 조작된 분할 인테인 및 분해 신호 (탈안정화 도메인, 데그론)를 가진 이의 조합물의 용도에 관한 것이다.

대표도



(52) CPC특허분류

C12N 9/22 (2013.01)
C07K 2319/60 (2013.01)
C07K 2319/92 (2013.01)
C12N 2750/14143 (2013.01)
C12N 2800/40 (2013.01)

(72) 발명자

오테로 빌바오 아나벨

스페인 08028 바르셀로나 파크 싸이언티픽 드 바르
셀로나 칼레 발디리 레이삭 10-12 스플라이스바이
오 에스.엘.

빌라 페렐로 미겔

스페인 08028 바르셀로나 파크 싸이언티픽 드 바르
셀로나 칼레 발디리 레이삭 10-12 스플라이스바이
오 에스.엘.

명세서

청구범위

청구항 1

조성물로서,

a. 분할 인테인 (split intein) N-단편을 포함하는 폴리펩타이드를 코딩하는 제1 폴리뉴클레오티드로서, 분할 인테인 N-단편이, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 서열번호 27의 CfaN, 서열번호 30의 CatN 및 서열번호 38의 Gp41N, 또는 서열번호 39의 ConN과 같은 이들의 임의의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는, 제1 폴리뉴클레오티드; 및

b. 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하는 제2 폴리뉴클레오티드로서, 분할 인테인 C-단편이, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 서열번호 28의 CfaC, 서열번호 31의 CatC 및 서열번호 104의 Gp41C, 또는 CfaCmut (서열번호 29) 또는 ConC (서열번호 105)와 같은 이들의 임의의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는, 제2 폴리뉴클레오티드

를 포함하고,

상기 조성물의 폴리뉴클레오티드 2종 모두 함께 하나의 제형으로 또는 분리하여 서로 다른 제형으로 포장될 수 있으며;

상기 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드는, 양쪽 단편이 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편에 연결되어 전체 단백질이 형성되는 방식으로, 각각 재구성할 단백질의 N-말단 단편 및 C-말단 단편을 코딩하고;

상기 재구성할 단백질이 25 KDa보다 크고; 및

상기 조성물이 다음과 같은 추가의 특징:

- 분할 인테인 N-단편이 펩타이드 결합을 통해 테그론과 직접 추가로 연결되고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결됨; 및/또는
- 분할 인테인 C-단편이 펩타이드 결합을 통해 테그론과 직접 추가로 연결되고, 테그론은 인테인 C-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 C-단편의 C-말단은 재구성할 단백질의 C-말단 단편에 펩타이드 결합을 통해 직접 연결됨

을 갖는 조성물.

청구항 2

제1항에 있어서,

상기 분할 인테인 N-단편은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된 서열번호 27의 CfaN이고, 상기 분할 인테인 C-단편은 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된 서열번호 28의 CfaC인, 조성물.

청구항 3

제1항에 있어서,

상기 분할 인테인 N-단편은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된 서열번호 27의 CfaN이고, 상기 분할 인테인 C-단편은 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된 서열번호 29의 CfaCmut 인, 조성물.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는, 조성물.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는, 조성물.

청구항 6

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는, 조성물.

청구항 7

제2항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는, 조성물.

청구항 8

제2항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는, 조성물.

청구항 9

제2항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는, 조성물.

청구항 10

제3항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는, 조성물.

청구항 11

제3항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는, 조성물.

청구항 12

제3항에 있어서,

상기 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론이 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는, 조성물.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서,

상기 조성물은

- o 상기 분할 인테인 N-단편이 펩타이드 결합을 통해 테그론과 직접 추가로 연결되고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- o 상기 분할 인테인 C-단편이 펩타이드 결합을 통해 테그론과 직접 추가로 연결되고, 테그론은 인테인 C-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 C-단편의 C-말단은 재구성할 단백질의 C-말단 단편에 펩타이드 결합을 통해 직접 연결되는 것을 특징으로 하는, 조성물.

청구항 14

제13항에 있어서

- a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-말단을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 28의 CfaC 또는 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 조성물.

청구항 15

제13항에 있어서,

- a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 28의 CfaC 또는 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 조성물.

청구항 16

제13항에 있어서,

- a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고, 추가적으로 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 28의 CfaC 또는 서열번호 29의 CfaCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 조성물.

청구항 17

제13항에 있어서,

- a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 38의 Gp41N 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 104의 Gp41C 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 조성물.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서,

상기 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드는, 폴리뉴클레오티드 2종 모두 각각의 단백질 복합체로 번역되어 본 발명의 방법에 따라 조합되었을 때, ABCA4 단백질의 N-말단 단편이 ABCA4 단백질의 C-말단 단편에 연결되어 전체 ABCA4 단백질이 생성되는 방식으로, ABCA4 단백질의 N-말단 단편 및 C-말단 단편을 각각 코딩하는, 조성물.

청구항 19

제18항에 있어서,

- a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 데그

론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;

상기 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1176 또는 1-1178번 위치를 코딩하고; 상기 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1177-2273 또는 1179-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1150-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1139번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1140-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1176번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1177-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1178번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1179-2273번 위치를 코딩하는, 조성물.

청구항 20

제18항에 있어서,

a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;

상기 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1176 또는 1-1178번 위치를 코딩하고; 상기 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1177-2273 또는 1179-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1150-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1139번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1140-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1176번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1177-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1178번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1179-2273번 위치를 코딩하는, 조성물.

청구항 21

제18항에 있어서,

a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 테그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 테그론은 인테인 N-단편과 테그론

사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;

상기 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1176 또는 1-1178번 위치를 코딩하고; 상기 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1177-2273 또는 1179-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1150-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1139번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1140-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1176번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1177-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1178번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1179-2273번 위치를 코딩하는, 조성물.

청구항 22

제18항에 있어서,

a. 제1 폴리뉴클레오티드가, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

b. 제2 폴리뉴클레오티드가, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로 펩타이드 링커를 통해 연결되고; 추가적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 연결된, 서열번호 28의 CfaC, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut, 또는 이들의 임의의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;

상기 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1149번 위치를 코딩하고; 상기 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1150-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1150-2273번 위치를 코딩하는, 조성물.

청구항 23

제18항에 있어서,

상기 제1 폴리뉴클레오티드는 제17항의 a에 따라 정의되고, 상기 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1095 또는 1-1185번 위치를 코딩하고; 상기 제2 폴리뉴클레오티드는 제17항의 b에 따라 정의되고, 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1096-2273 또는 1186-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1095번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1096-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1185번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1186-2273번 위치를 코딩하는, 조성물.

청구항 24

제1항 내지 제23항 중 어느 한 항에 있어서,

상기 폴리뉴클레오티드 2종 모두 적절한 숙주 세포에서 폴리뉴클레오티드의 증폭을 허용하는 벡터에 포함되는, 조성물.

청구항 25

제24항에 있어서,

상기 벡터가 아데노-부속 바이러스 (AAV)인, 조성물.

청구항 26

제25항에 있어서,

상기 벡터가 혈청형 1, 2, 3, 4, 5, 6, 7, 8 또는 9의 AAV인, 조성물.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서,

요법에 사용하기 위한 것인, 조성물.

청구항 28

하기 단계를 포함하는, 세포에서 대상 유전자를 발현시키는 방법:

(i) 세포를, 제1항 내지 제12항 중 어느 한 항에 정의된 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드와 접촉시키는 단계로서, 상기 폴리뉴클레오티드 하나 이상이 데그론에 펩타이드 결합으로 직접 연결된 분할 인테인 단편을 코딩하는, 단계;

(ii) 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드를 발현시켜, 제1 융합 단백질과 제2 융합 단백질을 생산하는 단계; 및

(iii) 제1 단백질과 제2 단백질을 접촉시켜, 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간산물이 생성되고, 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결되는 단계.

청구항 29

하기 단계를 포함하는, 세포에서 대상 유전자를 발현시키는 방법:

(i) 세포를, 제13항 내지 제23항 중 어느 한 항에 정의된 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드와 접촉시키는 단계로서, 상기 폴리뉴클레오티드 2종이 데그론과 펩타이드 결합으로 직접 연결된 분할 인테인 단편을 코딩하는, 단계;

(ii) 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드를 발현시켜, 제1 융합 단백질과 제2 융합 단백질을 생산하는 단계; 및

(iii) 제1 단백질과 제2 단백질을 접촉시켜, 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간산물이 생성되고, 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결되는 단계.

청구항 30

제28항 또는 제29항에 있어서,

상기 폴리뉴클레오티드 2종 모두 아데노 부속 바이러스 (AAV) 안에 포함된, 방법.

발명의 설명

기술 분야

본 발명은 생물공학 분야에 포함되며, 보다 상세하게는 분할 인테인 (split intein) 및 그 용도에 관한 것이다.

[0001]

배경 기술

- [0002] 유전자 요법에서 AAV (아데노-부속 바이러스) 안에 캡슐화하기 어려운 유전자는 2개의 단편으로 분할할 수 있는 것으로 알려져 있다. 각 단편에 분할 인테인 (split intein)을 융합하여, 자신의 AAV 내에서 전달하여야 한다. 표적 세포에 감염되면, 인테인들은 원하는 표적 단백질로 생성 가능할 것이다 (도 1 참조).
- [0003] 그럼에도 불구하고, 이러한 방식의 주요한 한가지 한계는 단백질 중간산물 (N-엑스테인-ItN 및 IntC-C-엑스테인)뿐 아니라 절단된 분할 인테인이 축적된다는 것이다. 이 문제를 해결하기 위해, 본 출원에서는, 일부 인테인을 단백질 중간산물의 축적을 상당히 감소시키는 특정한 분해 신호와 조합시킬 수 있는 것으로, 확인되었다. 본 발명은 이러한 인테인과 분해 신호를 조합하는데 집중하였다. 또한, 본 발명자들은 생체내에서 스플라이싱 수율을 증가시키는데 기여하는 일부 인테인-데그론 조합들도 동정하게 되었다.

발명의 내용

- [0004] 바람직한 제1 측면에서, 본 발명은 조성물에 관한 것으로, 조성물은
- [0005] a. 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하는 제1 폴리뉴클레오티드로서, 분할 인테인 N-단편이, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 서열번호 27의 CfaN, 서열번호 30의 CatN 및 서열번호 38의 Gp41N 또는 서열번호 39의 ConN과 같은 이들의 임의의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는, 제1 폴리뉴클레오티드; 및
- [0006] b. 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하는 제2 폴리뉴클레오티드로서, 분할 인테인 C-단편이, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 서열번호 28의 CfaC, 서열번호 31의 CatC 및 서열번호 104의 Gp41C, 또는 CfaCmut (서열번호 29) 또는 ConC (서열번호 105)와 같은 이들의 임의의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는, 제2 폴리뉴클레오티드를 포함하고,
- [0007] 조성물의 2종의 폴리뉴클레오티드는 함께 단일 제형으로 또는 분리하여 서로 다른 제형들로 포장될 수 있으며;
- [0008] 제1 및 제2 폴리뉴클레오티드는 각각, 양쪽 단편이 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편에 연결되어 전체 단백질을 형성되는 방식으로, 재구성할 단백질의 N-말단 단편 및 C-말단 단편을 코딩하며;
- [0009] 재구성할 단백질은 25 KDa 초과이고; 및
- [0010] 조성물은 추가적으로 하기와 같이 특정된다:
- [0011] - 분할 인테인 N-단편은 펩타이드 결합을 통해 데그론에 직접 추가로 연결되되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및/또는
- [0012] - 분할 인테인 C-단편은 펩타이드 결합을 통해 데그론에 직접 추가로 연결되되, 데그론은 인테인 C-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 C-단편의 C-말단은 재구성할 단백질의 C-말단 단편에 펩타이드 결합을 통해 직접 연결된다.
- [0013] 제1 측면에 대한 바람직한 구현예에서, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택된다. 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택된다. 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택된다.
- [0014] 제1 측면에 대한 다른 바람직한 구현예에서, 분할 인테인 N-단편은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되는 서열번호 27의 CfaN이고, 분할 인테인 C-단편은 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되는 서열번호 28의 CfaC이며; 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되고; 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 DD1, DD3, PEST, SopE, L2, L9, M4

또는 V12로 이루어진 목록으로부터 선택되고; 또 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택된다.

[0015] 제1 측면에 대한 다른 바람직한 구현예에서, 분할 인테인 N-단편은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되는 서열번호 27의 CfaN이고, 분할 인테인 C-단편은 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되는 서열번호 29의 CfaCmut이며; 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되고; 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되고; 더욱 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택된다.

[0016] 제1 측면에 대한 다른 바람직한 구현예에서, 분할 인테인 N-단편은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 서열번호 32의 NpuN이고, 분할 인테인 C-단편은 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut이며; 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7로 이루어진 목록으로부터 선택되고; 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택되고; 더욱 더 바람직하게는, 분할 인테인 N-단편 및/또는 분할 인테인 C-단편에 직접 연결되는 데그론은 SopE, L2, L9, M4 또는 V12로 이루어진 목록으로부터 선택된다.

[0017] 제1 측면에 대한 다른 바람직한 구현예에서, 조성물은 데그론 2개를 포함하는 것을 특징으로 하며, 특히 조성물은 하기를 특징으로 한다:

[0018] o 분할 인테인 N-단편은 펩타이드 결합을 통해 데그론에 직접 추가로 연결되며, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

[0019] o 분할 인테인 C-단편은 펩타이드 결합을 통해 데그론에 직접 추가로 연결되며, 데그론은 인테인 C-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 C-단편의 C-말단은 재구성할 단백질의 C-말단 단편에 펩타이드 결합을 통해 직접 연결된다.

[0020] 제1 측면에 대한 바람직한 구현예 또는 임의의 바람직한 구현예에서, 조성물은 하기를 포함하는 것을 특징으로 한다:

[0021] a. 제1 폴리뉴클레오티드로서, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제1 폴리뉴클레오티드; 및

[0022] b. 제2 폴리뉴클레오티드로서, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7으로 이루어진 군으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가로 연결된, 서열번호 28의 CfaC 또는 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제2 폴리뉴클레오티드.

- [0023] 제1 측면에 대한 바람직한 구현에 또는 임의의 바람직한 구현예에서, 조성물은 하기를 포함한다:
- [0024] a. 제1 폴리뉴클레오티드로서, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제1 폴리뉴클레오티드; 및
- [0025] b. 제2 폴리뉴클레오티드로서, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 28의 CfaC 또는 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제2 폴리뉴클레오티드.
- [0026] 제1 측면에 대한 다른 바람직한 구현예에서, 또는 이의 바람직한 구현예들 중 어느 하나에 있어서, 조성물은 하기를 포함한다:
- [0027] a. 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하는 제1 폴리뉴클레오티드로서, 데그론이 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단이 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제1 폴리뉴클레오티드; 및
- [0028] b. 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 28의 CfaC 또는 서열번호 29의 CfaCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하는 제2 폴리뉴클레오티드로서, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제2 폴리뉴클레오티드.
- [0029] 제1 측면에 대한 다른 바람직한 구현예에서, 또는 이의 바람직한 구현예들 중 어느 하나에 있어서, 조성물은 하기를 포함한다:
- [0030] a. 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 38의 Gp41N 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하는 제1 폴리뉴클레오티드로서, 데그론이 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단이 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제1 폴리뉴클레오티드; 및
- [0031] b. 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 104의 Gp41C 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하는 제2 폴리뉴클레오티드로서, 데그론이 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통

해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단이 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결된, 제2 폴리뉴클레오티드.

[0032] 제1 측면에 대한 다른 바람직한 구현예에서, 또는 이의 바람직한 구현예들 중 어느 하나에 있어서, 제1 및 제2 폴리뉴클레오티드는, 이들 2종의 폴리뉴클레오티드가 본 발명의 방법에 따라 이의 각각의 단백질 복합체로 번역되어 조합되었을 때, ABCA4 단백질의 N-말단 단편이 ABCA4 단백질의 C-말단 단편과 연결되어 전체 ABCA4 단백질을 형성되는 방식으로, ABCA4 단백질의 N-말단 단편 및 C-말단 단편을 각각 코딩한다.

[0033] 제1 및 제2 폴리뉴클레오티드가 ABCA4 단백질의 N-말단 단편 및 C-말단 단편을 각각 코딩하는 상기에 언급된 구현예와 관련하여, 본 발명은 하기 대안을 제공한다.

[0034] 제1 대안:

[0035] a. 제1 폴리뉴클레오티드는, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

[0036] b. 제2 폴리뉴클레오티드는, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;

[0037] 상기한 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1176 또는 1-1178번 위치를 코딩하고; 상기한 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1177-2273 또는 1179-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1150-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1139번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1140-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1176번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1177-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1178번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1179-2273번 위치를 코딩한다.

[0038] 제2 대안:

[0039] a. 제1 폴리뉴클레오티드는, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및

[0040] b. 제2 폴리뉴클레오티드는, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;

- [0041] 상기한 제1 폴리뉴클레오타이드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1176 또는 1-1178번 위치를 코딩하고; 상기한 제2 폴리뉴클레오타이드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1177-2273 또는 1179-2273번 위치를 코딩하며, 제1 폴리뉴클레오타이드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1150-2273번 위치를 코딩하고, 제1 폴리뉴클레오타이드가 1-1139번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1140-2273번 위치를 코딩하고, 제1 폴리뉴클레오타이드가 1-1176번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1177-2273번 위치를 코딩하고, 제1 폴리뉴클레오타이드가 1-1178번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1179-2273번 위치를 코딩한다.
- [0042] 제3 대안:
- [0043] a. 제1 폴리뉴클레오타이드는, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- [0044] b. 제2 폴리뉴클레오타이드는, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut 또는 이들의 임의의 기능적인 등가 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;
- [0045] 상기한 제1 폴리뉴클레오타이드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1176 또는 1-1178번 위치를 코딩하고; 상기한 제2 폴리뉴클레오타이드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1177-2273 또는 1179-2273번 위치를 코딩하며, 제1 폴리뉴클레오타이드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1150-2273번 위치를 코딩하고, 제1 폴리뉴클레오타이드가 1-1139번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1140-2273번 위치를 코딩하고, 제1 폴리뉴클레오타이드가 1-1176번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1177-2273번 위치를 코딩하고, 제1 폴리뉴클레오타이드가 1-1178번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1179-2273번 위치를 코딩한다.
- [0046] 제4 대안:
- [0047] a. 제1 폴리뉴클레오타이드는, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 27의 CfaN 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- [0048] b. 제2 폴리뉴클레오타이드는, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고; 선택적으로 DD1, DD3, PEST, SopE, L2, L9, M4 또는 V12으로 이루어진 목록으로부터 선택되는 데그론에 펩타이드 결합을 통해 직접 추가적으로 연결된, 서열번호 28의 CfaC, 서열번호 29의 CfaCmut 또는 서열번호 36의 NpuCmut, 또는 이들의 임의의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하되, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;
- [0049] 상기한 제1 폴리뉴클레오타이드는 ABCA4 단백질의 N-말단 단편의 1-1149번 위치를 코딩하고; 상기한 제2 폴리뉴클레오타이드는 ABCA4 단백질의 C-말단 단편의 1150-2273번 위치를 코딩하며, 제1 폴리뉴클레오타이드가 1-1149번 위치를 코딩하는 경우 제2 폴리뉴클레오타이드는 1150-2273번 위치를 코딩한다.
- [0050] 제5 대안:

- [0051] a. 제1 폴리뉴클레오티드는, 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고, 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7로 이루어진 목록으로부터 선택되는 테그론이 펩타이드 결합을 통해 직접 연결된, 서열번호 38의 Gp41N 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 N-단편을 포함하는 폴리펩타이드를 코딩하되, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 및
- [0052] b. 제2 폴리뉴클레오티드는, 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결되고, 추가적으로 CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7로 이루어진 목록으로부터 선택되는 테그론이 펩타이드 결합을 통해 직접 연결된, 서열번호 104의 Gp41C 또는 이의 기능적으로 동등한 변이체로 이루어진 목록으로부터 선택되는 분할 인테인 C-단편을 포함하는 폴리펩타이드를 코딩하되, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 N-단편의 C-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되며;
- [0053] 상기한 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1095 또는 1-1185번 위치를 코딩하고; 상기한 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1096-2273 또는 1186-2273번 위치를 코딩하며, 제1 폴리뉴클레오티드가 1-1095번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1096-2273번 위치를 코딩하고, 제1 폴리뉴클레오티드가 1-1185번 위치를 코딩하는 경우 제2 폴리뉴클레오티드는 1186-2273번 위치를 코딩한다.
- [0054] 제1 측면에 대한 다른 바람직한 구현예에서, 또는 이의 바람직한 구현예들 또는 대안들 중 어느 하나에 있어서, 폴리뉴클레오티드 2종 모두 적절한 숙주 세포에서 이러한 폴리뉴클레오티드의 증폭 또는 삽입을 허용하는 벡터 안에 포함된다. 바람직하게는, 이러한 벡터는 아데노 부속 바이러스 (AAV)이며, 더 바람직하게는 벡터는 혈청형 1, 2, 3, 4, 5, 6, 7, 8 또는 9의 AAV이다.
- [0055] 제1 측면에 대한 다른 추가적인 바람직한 구현예에서, 또는 이의 바람직한 구현예들 또는 대안들 중 어느 하나에 있어서, 본원에 기술된 조성물은 요법에 사용하기 위한 것이며, 특히 표 1에 표시된 임의 질환에 대해 사용하기 위한 것이다.
- [0056] 본 발명의 추가적인 측면은 하기를 포함하는 세포에서 대상 유전자를 시험관내 또는 생체내 발현하는 방법에 관한 것이다:
- [0057] (i) 세포를, 상기한 측면, 구현예 또는 대안 중 어느 하나에 정의된 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드와 접촉시키는 단계로서, 바람직하게는, 폴리뉴클레오티드 2종 모두 아데노 부속 바이러스 (AAV) 안에 포함되고, 하나 이상의 폴리뉴클레오티드가 펩타이드 결합을 통해 테그론과 직접 연결된 분할 인테인을 코딩하는 것인, 단계,
- [0058] (ii) 제1 및 제2 융합 단백질이 생산되도록 제1 및 제2 폴리뉴클레오티드의 발현을 허용하는 단계; 및
- [0059] (iii) 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간산물이 형성되고 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단과 공유 연결되도록, 제1 단백질과 제2 단백질 간에 접촉을 허용하는 단계.

도면의 간단한 설명

- [0060] **도 1.** 인테인 및 테그론과 인테론의 조합을 이용하여 유전자 요법을 위한 거대 단백질을 재구성하기 위한 전략에 대한 일반적인 도식. 상단 패널: (1) 대상 유전자는 신중하게 선택된 위치에서 재조합 방식으로 분할하고, 수득한 5' 말단에는 IntN이, 3'에는 IntC가, 단백질 발현시 단백질의 N-말단 단편이 이의 C-말단에 융합된 IntN과 함께 발현되고 단백질의 C-말단 단편이 이의 N-말단에 융합된 IntC와 함께 발현되게 하는 방식으로, 재조합 융합된다. (2) 각 구조체는 개별 AAV 안에 캡슐화된다. AAV 2종을 특정 세포에 공동-형질전이되도록 환자에 투여한다. (3) 공동-형질전이되면, AAV를 통해 전달된 DNA가 RNA로 전사되고, 단백질로 번역될 것이며, 단백질 수준에서 인테인은 단백질 트랜스-스플라이싱 반응을 통해 원하는 분할된 생성물로 재구성될 것이다 (4). 하단 패널: 인테인을 출발 물질의 축적을 방지하고 절단된 인테인을 제거하기 위해 분해 신호와 조합한다. 테그론은 N-

인테인의 C-말단 및/또는 C-인테인의 N-말단에 융합된다.

도 2. 상단 패널: H6 태그가 달린 전장 EGFP를 제조하는 EGFP^N-IntN-H6 및 IntC-EGFP^C-H6 구조체 간의 반응 도식. 하단 패널: Cfa 또는 Npu 인테인을 이용한, N-말단 단편, C-말단 단편으로 형질감염된 세포 또는 이들 2종의 단편으로 공동-형질감염된 세포에 대한 형광 현미경 사진.

도 3. EGFP-인테인 폴리뉴클레오티드로 형질전환된 HEK293 세포 용해물에 대한 웨스턴 블롯 분석. 좌측 패널: EGFP-인테인 플라스미드로 형질감염된 세포에 대한 웨스턴 블롯 분석. 세포에 전장 EGFP 플라스미드 또는 EGFP^N-IntN, IntC-EGFP^C 플라스미드를 형질감염시키거나, 또는 이들 2종의 플라스미드를 공동-형질감염시킨다. Cfa 또는 Npu 인테인을 함유한 플라스미드를 검사한다. 우측 패널: Npu 대비 스플라이싱된 산물의 배수 증가 정량. 생산 수율은 로딩 대조군으로서 β -튜불린을 이용한 밀도측정으로 결정한다. 그래프는, 세포에 전장 EGFP 플라스미드 (EGFP)를 형질감염시키거나 또는 EGFP^N-CfaN 및 CfaC-EGFP^C (Cfa)로, 또는 EGFP^N-NpuN 및 NpuC-EGFP^C (Npu)로 공동-형질감염시킬 경우, Npu에 대한 EGFP의 배수 증가를 도시한다.

도 4. HEK293 형질감염 세포를 유세포 측정으로 분석한 것이다. 좌측 패널: 여러가지 구조체 및 구조체들의 조합으로 형질감염한 세포에 대한 유세포 측정 데이터. 검정 곡선은 EGFP의 N-말단 단편 또는 C-말단 단편으로 형질감염된 대조군을 나타낸다. 청색 곡선은 전장 EGFP를 코딩하는 플라스미드로 형질감염된 세포 레플리카이트 4종이고, 녹색 곡선은 EGFP^N-CfaN 및 CfaC-EGFP^C (71번 위치에서 EGFP 분할)로 공동-형질감염된 세포이고, 노란색 곡선은 EGFP^N-NpuN 및 NpuC-EGFP^C (71번 위치에서 EGFP 분할)로 공동-형질감염된 세포이다. 중앙 패널: Cfa 인테인을 사용하여 전장 EGFP에 해당하는 신호를 90% 이상으로 회복할 수 있음을 보여주는, 각 샘플 세트에서 수득한 평균 형광 강도 (MFI). 우측 패널: 형질감염 효율이 이들 3세트 모두에서 비슷함을 보여주는, 각 샘플에서의 양성 세포 수를 나타낸 그래프.

도 5: Cfa 및 Npu 사이의 1150번 위치에서 ABCA4의 스플라이싱 수율 비교. 세포에 ABCA4^N1150-IntN 및 IntC-ABCA4^C1150을 공동-형질감염시키고, 세포용해 후 WB에 의해 분석하였다. ABCA4^N1150은 잔기 1-1149번 범위의 ABCA4의 N-말단 단편이고, ABCA4^C1150은 잔기 1150-2273번 범위의 ABCA4의 C-말단 단편에 해당한다. Cfa 또는 Npu를 가진 구조체를 검사하였다. 좌측 패널은 항-ABCA4 mAb를 이용한 블롯을 도시한다. 우측 패널은 웨스턴 블롯의 밀도측정 정량 결과이다.

도 6. Cfa 및 Npu 사이의 1140번 위치에서 ABCA4의 스플라이싱 수율 비교. 세포에 ABCA4^N1140-IntN 및 IntC-ABCA4^C1140을 공동-형질감염시키고, 세포용해 후 WB에 의해 분석하였다. ABCA4^N1140은 잔기 1-1139번 범위의 ABCA4의 N-말단 단편이고, ABCA4^C1140은 잔기 1140-2273번 범위의 ABCA4의 C-말단 단편에 해당한다. Cfa 또는 Npu를 가진 구조체를 검사하였다. 좌측 패널 및 중앙 패널은 항-FLAG 태그 및 항-ABCA4 mAb를 각각 이용한 블롯 결과를 도시한다. 우측 패널은 웨스턴 블롯의 밀도측정 정량 결과이다.

도 7: 1140, 1150, 1179 및 1188번 위치에서 ABCA4의 스플라이싱 수율 비교. 세포에 대응되는 ABCA4^N-CfaN 및 CfaC (또는 CfaCmut)-ABCA4^C 구조체를 공동-형질감염시키고, 세포용해 후 WB에 의해 분석하였다. 그래프는 재구성된 ABCA4 단백질의 WB에 의한 밀도측정 정량 결과 및 전장 단백질을 코딩하는 플라스미드로 형질감염 시 수득한 ABCA4 수준의 비교 결과를 도시한다.

도 8: 테그론을 함유한 PTS의 반응 도식. 대상 단백질의 N-말단 단편을 IntN, 링커, 이 경우 His6 태그, 및 테그론과 융합한다. 단백질의 C-말단은, N-말단이 테그론과 연결된, IntC와 융합한다. 스플라이싱 반응 결과, 이에 융합된 임의의 테그론이 제거되고 각 분해 신호를 가진 인테인이 적출된, 원하는 전장 단백질이 만들어진다.

도 9: 리포터 단백질 EGFP를 재구성하기 위한 인테인과 테그론의 조합. HEK293 세포에 지정된 구조체를 형질감염시키고, 세포 용해 후 항-His6 태그 mAb를 이용한 웨스턴 블롯으로 분석하였다. 그 결과, CfaN 또는 C 인테론과 분해 신호를 조합한 구조체의 사용시 출발 물질 (EGFP-CfaN-6H) 뿐만 아니라 인테인 (CfaN-6H)으로부터 발생하는 임의의 신호가 소실되는 것으로 확인된다.

도 10: 거대 단백질 ABCA4를 재구성하기 위한 인테인과 테그론의 조합. HEK293 세포에 ABCA4^N1150-CfaN 및 CfaC-ABCA4^C1150을 테그론을 첨가한 경우와 첨가하지 않은 경우로 형질감염시키고, 세포 용해 후 항-FLAG 태그

mAb를 이용한 웨스턴 블롯에 의해 분석하였다. 세포를 다른 2 시점 (24시간 및 48시간)에 수집하였다. 상단 패널: 웨스턴 블롯 분석. 하단 패널: 웨스턴 블롯에 대한 밀도측정 정량. 그 결과, Cfa N 또는 C 인테인을 분해 신호와 조합하는 구조체를 이용한 경우 출발 물질 (ABCA4^N-CfaN 및 CfaC-ABCA4^C)로부터 발생하는 신호가 사라진 것으로 확인된다. 흥미롭게도, 또한 Cfa 인테인과 분해 신호의 조합시 재구성된 전장 ABCA4의 수율이 증가되는 결과가 확인된다.

도 11: 인테인-데그론 조합물의 효과는 프로테오솜에 의해 매개된다. 세포에 ABCA4^N1150-CfaN 및 CfaC-ABCA4^C1150 구조체를 데그론을 첨가한 경우와 첨가하지 않은 경우로 공동-형질감염시키고; 24시간 후, 세포를 지정된 기간 동안 (30분, 3시간, 6시간 및 24시간) 프로테오솜 저해제 MG132의 존재 또는 부재 하에 인큐베이션 하였다. 지정된 시간 경과 후, 세포를 용해시키고, 항-FLAG mAb를 이용한 웨스턴 블롯에 의해 분석하였다. 프로테오솜 저해제 MG132의 추가시 단편에 해당하는 밴드의 강도 증가를 유발하며, 이는 이의 분해가 저해되었음을 의미한다.

도 12: 여러가지 인테인 또는 데그론과 인테인의 조합을 이용한 재구성 수율 비교. 좌측 패널: Cfa, Cfa-SopE 또는 Npu를 이용한 1150번 위치에서 ABCA4 분할의 PTS를 경유한 재구성에 대한 웨스턴 블롯. 세포에 ABCA4^N1150-CfaN 및 CfaC-ABCA4^C1150 (Cfa), ABCA4^N1150-CfaN-SopE 및 SopE-CfaC-ABCA4^C1150 (Cfa-SopE), 또는 ABCA4^N1150-NpuN 및 NpuC-ABCA4^C1150 (Npu)을 공동-형질감염시켰다. 48시간 후, 세포를 수집하여 세포 용해한 다음 WB에 의해 분석하였다. Cfa 구조체의 두플리케이트와 CfaSopE 및 Npu에 대한 트리플리케이트를 분석하였다. 우측 패널: 인테인 및 데그론으로 구성된 각 조합에서 수득된 생산물 및 남아있는 출발 물질에 대한 밀도 측정 정량 결과. 결과는, Cfa를 데그론 (SopE)와 조합하여 이용할 경우, 생산물의 재구성이 얼마나 더 높은지를 보여준다. 흥미롭게도, 데이터는 Npu를 이용한 재구성이 가장 낮다는 것을 보여주며, 또한 이는 미-반응 출발 물질이 가장 많이 잔류하는 것임을 보여준다.

도 13: 분해 신호와 함께 또는 분해 신호 없이 CfaCmut를 이용한 경우 1140번 위치에서 ABCA4의 재구성. 좌측 패널: 세포에 ABCA4^N1140-CfaN 및 CfaCmut-ABCA4^C1140, ABCA4^N1140-CfaN-SopE 및 SopE-CfaC-ABCA4^C1140 (Cfa-SopE), 또는 전장 ABCA4 (ABCA)를 공동-형질감염시켰다. 48시간 후, 세포를 수집하여 세포 용해한 다음 항-FLAG 태그 mAb를 이용한 WB에 의해 분석하였다. 우측 패널: 전장 ABCA4의 형질감염과 비교한, ABCA4 재구성 수준에 대한 밀도 측정 정량 결과.

도 14: 리포터 단백질 EGFP를 DD1 데그론을 이용해 재구성하기 위한 인테인과 데그론의 조합. HEK293 세포에 지정된 구조체를 형질감염시키고, 세포 용해한 후 항-His6 태그 mAb를 이용한 웨스턴 블롯에 의해 분석하였다. Cfa 인테인과 DD1 데그론의 조합시, 부적절한 출발 물질과 절제된 인테인이 소거된다.

도 15: 질량 분광 측정 분석. 세포에 ABCA4^N1150-CfaN-SopE 및 SopE-CfaC-ABCA4^C1150을 형질감염시키고, 48시간 후 수집하여 세포 용해하였다. 세포 용해물을 SDS-PAGE 겔에서 전개시켰으며, 약 300 KDa 밴드를 추출하여 단백질 분해 처리 (재료 및 방법 참조) 후 LC-MS/MS에 의해 분석하였다. 결과는 도에 요약 기술한다. 녹색 부분은 MSMS에 의해 위 발견을 (FDR) 1% 미만으로 동정된 펩타이드를 표시하고, 이다. 노란색 부분은 FDR 5% 미만으로 동정된 펩타이드를 표시한다. 적색 밑줄 친 부분은 분할 부위의 서열을 표시한다.

도 16: 오르소고날 컨센서스 인테인 (orthogonal consensus inteins)을 이용한 3-조각 절찰.

도 17: 오르소고날 컨센서스 인테인을 이용한 3-조각 절찰. 도 16에 도시된 N, M 및 C 구조체들을 HEK293에 각각 단독으로, 쌍으로, 또는 3종을 함께 형질감염시켰다. 세포를 용해한 후 N, M- 및 C-말단 단편 (좌측 패널)의 C-말단에 존재하는 3FT 태그에 대한 항체 또는 POIN-단편 (우측)에 대한 항체를 이용한 웨스턴 블롯으로 분석하였다.

도 18: A) 본 발명의 분할 인테인 N-단편의 발현 수준에 대한 데그론의 효과 분석. 데그론 함유 구조체 및 데그론 비-함유 구조체를 HEK293 세포에 형질감염시키고, 발현 수준을 웨스턴 블롯으로 분석하였다. N-인테인 단편에 데그론 부가시, 발현된 단백질의 검출 수준이 감소한다. B) 본 발명의 분할 인테인 C-단편의 발현 수준에 대한 데그론의 효과 분석. 데그론 함유 구조체 및 데그론 비-함유 구조체를 HEK293 세포에 형질감염시키고, 발현 수준을 웨스턴 블롯으로 분석하였다. C-인테인 단편에 데그론 부가시, 발현된 단백질의 검출 수준이 감소한다. C) N-말단 단편 또는 C-말단 단편으로 형질감염된 또는 이들 단편이 공동-형질감염된 세포에 대한 웨스턴 블롯 분석 결과를 도시한다. 데그론 비-함유 구조체 (데그론 없음) 또는 대표적인 데그론 함유 구조체 (SopE, L2 및

L9)의 결과들을 나타낸다. 우측 패널은 지정된 각 태그론 함유 구조체에서 수득된 PTS 생산물의 양을 태그론 비 함유 구조체에서 수득된 결과와 비교해 정량한 결과를 제시한다. 일부 태그론은 태그론 비-함유 구조체와 비교해 수율이 더 높음을 알 수 있다. 흥미롭게도, 태그론들 모두 작동하며, 이들 모두 함유한 경우 태그론 비-함유 구조체에서 수득되는 생산물의 적어도 25%가 재구성될 수 있다. 일부 태그론은 태그론 비-함유 구조체에서 수득되는 수준의 50% 이상으로 재구성할 수 있으며, 최종적으로 일부 태그론은 100% (또는 그 이상)의 재구성을 달성한다.

도 19. 여러가지 태그론과의 PTS 반응을 인테인을 검출할 수 있는 여러가지 유형의 겔을 이용하여 WB에 의해 분석하였다. 알 수 있는 바와 같이, 몇몇 태그론은 다른 것 (예를 들어, SpoE 및 PEST)과 비교해 절단되는 인테인의 양을 감소시키며, L2, L9, M2, M4 또는 DD3과 같은 다른 태그론은 인테인 밴드를 완전히 제거한다.

도 20. 여러가지 태그론 조합들 및 태그론 비-함유 구조체를 이용한 분할 위치 1150번에서 PTS 반응에 대한 웨스턴 블롯 분석 결과. 이 경우, N-단편에 위치한 태그론은 고정하고 (좌측 패널 L9 및 우측 패널 L2, N으로 표시), C-단편에는 여러가지 태그론 (L2, L9 M4, M2, V12, DD3 및 DD1, C로 표시)을 사용하였다.

도 21: N-말단 및 C-말단 인테인 ABCA4 단편에서 태그론 L9 함유 구조체 또는 비-함유한 구조체를 이용한 1140 및 11 1179번 위치에서의 PTS 반응들에 대한 WB 및 정량 결과.

도 22: gp41 인테인을 이용한 PTS 반응에 대한 웨스턴 블롯 분석. 좌측 패널: Cfa 인테인을 이용한 1150번 위치 및 gp41 인테인을 이용한 1185번 위치에서 ABCA4 분할의 PTS를 통한 재구성에 대한 웨스턴 블롯. 48시간 후, 세포를 수집하여 세포 용해한 다음 WB에 의해 분석하였다. 우측 패널: 각 인테인을 이용하여 수득한 생산물 및 전장 단백질에 대한 밀도 측정 정량 결과.

도 23: 생체내 망막 EGFP: (A) 전장 EGFP를 코딩하는 AAV8 또는 본 발명의 N (N, EGFPN-CfaN) 및 C-말단 단편 (C, CfaC-EGFPC)을 코딩하는 AAV 2종을 주사한 마우스에서의 안저 자동 형광 (Fundus Auto Fluorescence, FAF). (B) 식염수 또는 처리 망막에 대한 면역 조직화학 분석 (IHC). 결과는 단백질 스포라이싱 매개 EGFP 재구성을 의미하는, 광-수용체에서의 네이티브 EGFP 형광성을 나타낸다. 모든 구조체는 GRK1 프로모터의 통제 하에 배치하였으며, N+C 구조체를 서로 다른 용량으로 주사하였다. BSS (균형 멸균 식염수, Balanced Sterile Saline).

발명을 실시하기 위한 구체적인 내용

- [0061] 본 발명은 조작된 또는 자연 생성 분할 인테인의 이용 방법뿐 아니라 유전자 요법을 위한 거대 단백질을 재구성하기 위한 이와 분해 신호 (탈안정화 도메인, 태그론)의 조합에 관한 것이다.
- [0062] 본원에 기술된 방법은 거대 단백질을 시험관내, 생체외 및 생체내에서 재구성할 수 있다. 생체내 이용할 경우, 비-제한적으로, 중추 신경계 (CNS), 말초신경계 (PNS), 근육, 간, 눈, 췌장, 망막, 신장, 내이, 심장, 폐, 혈액, 비장, 피부 등의 임의의 원하는 조직에서 원하는 거대 표적 단백질을 재구성할 수 있을 것이다.
- [0063] 본원에서, 용어 "인테인"은 전구체 단백질로부터 인테인 서열을 잘라 측면 서열 (N- 및 C-엑스테인)과 펩타이드 결합으로 연결하는 단백질 스포라이싱 반응을 촉매할 수 있는, 자연-생성 또는 인공-구축된 폴리펩타이드 서열을 의미한다. 이는 전형적으로 아미노산 150-550개 크기이며, 또한 호밍 엔도뉴클레아제 도메인을 함유할 수도 있다. 공지된 인테인 목록은 <http://www.inteins.com> 및 <https://inteins.biocenter.helsinki.fi/index.php>에 공개되어 있다.
- [0064] 본원에서, 용어 "분할 인테인"은, N-말단 서열과 C-말단 서열이 비-공유적으로 재-연합 또는 재구성되어 트랜스-스포라이싱 반응에 기능하는 인테인이 될 수 있는 단편들로 분리되도록, N-말단 아미노산 서열과 C-말단 아미노산 서열이 펩타이드 결합으로 직접 연결되어 있지 않은, 임의의 인테인을 지칭한다.
- [0065] 용어 "펩타이드 결합"은 카르복시 성분으로서 지칭되는 한 분자의 카르복시 부분이 아미노 성분으로서 지칭되는 다른 분자의 아미노 부분과 반응하여, 한 분자를 해리하면서 분자 2종 간에 형성되는 화학적 공유 결합 -CO-NH-를 지칭한다. 예를 들어, 단백질 생성 (proteinogenic) L-아미노산은 물 분자가 해리되면서 연결되어 펩타이드 결합을 형성할 수 있다. 따라서, 단백질 및 펩타이드는 펩타이드 결합에 의해 함께 연결된 아미노산 잔기 채로 서 간주할 수 있다. 펩타이드 결합은 "아미드 결합" 또는 "아미드 연결"이다.
- [0066] 용어 "폴리펩타이드", "펩타이드" 또는 "단백질"은 본원에서 상호 호환적으로 아미노산들로 된 폴리머를 지칭한다.

- [0067] 용어 "아미노산"은 자연 생성 아미노산과 합성 아미노산뿐 아니라 자연 생성 아미노산과 비슷한 방식으로 기능하는 아미노산 유사체 및 아미노산 모방체를 지칭한다. 아울러, 용어 "아미노산"은 D-아미노산 및 L-아미노산(입체이성질체) 둘다 포괄한다.
- [0068] 용어 "천연 아미노산" 또는 "자연 생성 아미노산"은 자연 생성 아미노산 20종을 포함하며; 종종 생체내에서 번역 후 변형된 아미노산, 예를 들어, 하이드록시프롤린, 포스포세린 및 포스포트레오닌 등과; 비-제한적으로, 2-아미노아디프산, 하이드록시라이신, 이소데스모신, 노르-발린, 노르-루신 및 오르니틴 등의 기타 특이 아미노산을 포괄한다.
- [0069] 본원에서, 용어 "비-천연 아미노산" 또는 "합성 아미노산"은 천연 아미노산과 구조적으로 비슷하며 "α" (알파) 위치에서 아민기로 치환된 카르복시산 또는 이의 유도체를 지칭한다. 변형된 아미노산 또는 특이 아미노산에 대한 비-제한적인 예시적인 예로는 2-아미노아디프산, 3-아미노아디프산, β-알라닌, 2-아미노부티르산, 4-아미노부티르산, 6-아미노카프로익산, 2-아미노헵타노익산, 2-아미노이소부티르산, 3-아미노이소부티르산, 2-아미노피델산, 2,4-다이아미노부티르산, 데스모신, 2,2'-다이아미노피델산, 2,3-다이아미노프로피온산, N-에틸글리신, N-에틸아스파라긴, 하이드록시 라이신, 알리오 (alio) 하이드록시 라이신, 3-하이드록시프롤린, 4-하이드록시프롤린, 이소데스모신, 알로이소루신, N-메틸글리신, N-메틸이소루신, 6-N-메틸-라이신, N-메틸발린, 노르발린, 노르루신, 오르니틴, p-아세틸페닐알라닌, p-할로페닐알라닌, p-프로파르글리옥시페닐알라닌, p-아지도페닐알라닌, p-벤조일페닐알라닌 등을 포함한다. 이 군은 또한 "천연 아미노산"의 D-이성질체를 포함한다.
- [0070] 본원에서, 용어 "분할 인테인 N-단편" 또는 "N-말단 분할 인테인" 또는 "N-말단 인테인 단편" 또는 "N-말단 인테인 서열" (약칭 "IntN")은 트랜스-스플라이싱 반응에 기능하는, 즉, 기능성 분할 인테인 C-단편과 연합하여 숙주 단백질로부터 자신을 잘라낼 수 있으며, 엑스테인 또는 측면 서열과 펩타이드 결합 결합을 촉매하여 완전한 인테인을 형성할 수 있는, 또는 분할 인테인 C-단편과의 연합시 "N-말단 절단", 즉 엑스테인과 분할 인테인 N-단편의 N-말단 사이의 펩타이드 결합을 친핵성 공격하여 펩타이드 결합의 파괴를 유발하는, N-말단 아미노산 서열을 포함하는 임의의 인테인 서열을 지칭한다. 따라서, IntN은 트랜스-스플라이싱이 발생하면 스플라이싱에 의해 제거되는 서열을 또한 포함한다. IntN은 자연 생성 인테인 서열의 N-말단 부분의 변형 서열을 포함할 수 있다. 예를 들어, 이는 이러한 추가적인 및/또는 돌연변이된 잔기의 함유가 IntN을 트랜스-스플라이싱에 비-기능성으로 만들지 않는 한, 추가적인 아미노산 잔기 및/또는 돌연변이된 잔기를 포함할 수 있다. 바람직하게는, 추가적인 및/또는 돌연변이된 잔기의 함유는 IntN의 트랜스-스플라이싱 활성을 개선 또는 강화한다.
- [0071] 본원에서, 용어 "테그론"은 다른 폴리펩타이드에 재조합을 통해 융합되었을 때 이 단백질의 프로테오솜 분해 경로 또는 임의의 기타 세포성 분해 기전을 통한 분해를 가속화하는, 자연-생성 또는 인공적으로-구축된 폴리펩타이드 서열을 의미한다.
- [0072] 본원에서 상호 호환적으로 사용되는 바와 같이, 용어 "분할 인테인 C-단편", "C-말단 분할 인테인", "C-말단 인테인 단편" 및 "C-말단 인테인 서열" (약칭 "IntC")은, 트랜스-스플라이싱 반응에 기능하는 C-말단 아미노산 서열을 포함하는 임의의 인테인 서열을 지칭한다. 즉, 기능성 분할 인테인 N-단편과 연합하여, 숙주 단백질로부터 자신을 잘라내고, 엑스테인 또는 측면 서열과 펩타이드 결합을 통한 결합을 촉매할 수 있거나, 또는 분할 N-인테인과 연합시 "C-말단 절단"을 촉매하는, 즉 엑스테인과 분할 인테인 C-단편 사이의 펩타이드 결합을 친핵성 공격하여 펩타이드 결합을 파괴하는 것을 촉매하는, C-말단 아미노산 서열을 포함하는 임의의 인테인 서열을 지칭한다. 따라서, IntC는 트랜스-스플라이싱이 발생하였을 때 스플라이싱에 의해 제거되는 서열을 또한 포함한다. IntC는 자연 생성 인테인 서열의 C-말단 부분의 변형 서열을 포함할 수 있다. 예를 들어, 이는 이러한 추가적인 및/또는 돌연변이된 잔기의 함유가 IntC를 트랜스-스플라이싱에 비-기능성으로 만들지 않는 한, 추가적인 아미노산 잔기 및/또는 돌연변이된 잔기를 포함할 수 있다. 바람직하게는, 추가적인 및/또는 돌연변이된 잔기의 함유는 IntC의 트랜스-스플라이싱 활성을 개선 또는 강화한다.
- [0073] 본원에서, 본 발명의 맥락에서, "재구성할 단백질의 N-말단 단편"은 단백질의 N-말단 단편, 특히 25 KDa 초과, 50 KDa 초과 또는 100 KDa 초과 단백질 단편을 지칭함에 유념한다. 본원에서, 용어 "단백질의 N-말단 단편"은 따라서 단백질 (성숙한 또는 비-성숙 형태)의 N-말단을 함유한 가변적인 길이의 단편을 지칭한다. 구체적인 구현예에서, N-말단 단편은 전체 단백질 길이의 100% 미만, 90% 미만, 80% 미만, 70% 미만, 60% 미만, 50% 미만, 40% 미만, 30% 미만, 20% 미만, 10% 미만, 5% 미만의 길이에 해당하는 단편이다.
- [0074] 본원에서, 본 발명의 맥락에서, "재구성할 단백질의 C-말단 단편"은 단백질의 C-말단 단편, 특히 25 KDa 초과, 50 KDa 초과 또는 100 KDa 초과 단백질 단편을 지칭함에 유념한다. 본원에서, 용어 "단백질의 C-말단 단편"은 따라서 단백질의 C-말단을 함유한 가변적인 길이의 단편을 지칭한다. 구체적인 구현예에서, C-말단 단편은 전체

단백질 길이의 100% 미만, 90% 미만, 80% 미만, 70% 미만, 60% 미만, 50% 미만, 40% 미만, 30% 미만, 20% 미만, 10% 미만, 5% 미만의 길이에 해당하는 단편이다.

[0075] 본원에서, 용어 "폴리뉴클레오티드"는 포스포다이에스테르 결합 (또는 유사한 구조 변이체 또는 이의 합성 유사체)을 통해 연결된 뉴클레오티드 단위 (데옥시리보뉴클레오티드 또는 리보뉴클레오티드 또는 유사한 구조 변이체 또는 이의 합성 유사체) 여러개로 구성된 폴리머를 지칭한다. 용어 폴리뉴클레오티드는 이중 가닥 또는 단일 가닥 게놈 및 cDNA, RNA, 임의의 합성 및 유전자 조작된 폴리뉴클레오티드, 및 센스 및 안티-센스 폴리뉴클레오티드 (본 발명에서 센스 가닥만 기술된 경우에도)를 포괄한다. 이는 단일-가닥 및 이중-가닥 분자, DNA-DNA, DNA-RNA 및 RNA-RNA 하이브리드를 포괄한다.

[0076] **분할 인테인 N-단편**

[0077] 본 발명은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접 연결되는 분할 인테인 N-단편을 지칭한다. 이 구조체는 N-말단에서 C-말단으로 다음과 같은 일반적인 구조 형태를 가질 것이다:

[0078] (재구성할 단백질의 N-말단 단편) - InteinN.

[0079] 재구성할 단백질은 임의의 거대 유전자일 수 있으며, 이의 인지시 긍정적인 치료 효과를 발휘할 수 있다. 본 발명으로 재구성할 수 있는 단백질에 대한 비-제한적인 예는 ABCA4 단백질, 및 아래 표 1에 열거된 유전자들에 의해 코딩되는 단백질들이다. 이들 단백질의 재구성 목적은 이러한 유전자의 올바른 버전을 제공함으로써 이를 코딩하는 유전자 내 돌연변이와 관련있는 질환을 치료하고자 하는 것이다. 아울러, 재구성할 단백질은 질환을 치료하기 위한 단백질 또는 효소일 수 있으나, 반드시 돌연변이된 것을 치환하는 것은 아니다. 예를 들어, 재구성할 단백질은 유전자, 염기 또는 프라임 편집 (prime editing)을 위한 CRISPR/Cas9 시스템일 수 있다.

[0080] **표 1.**

표 1

[0081]

질환	단백질 (유전자)	Uniprot ID	코딩 영역 크기
어서 증후군 2C (USH2C)	부착 G-단백질 커플링된 수용체 V1 (ADGRV1)	Q8WXG9	18.9
황반 변성, 노인성, 1 (ARMD1)	헤미센틴-1 (HMCN1)	Q96RW7	18.9
열성 경련, 가족성, 4 (FEB4)	부착 G-단백질 커플링된 수용체 V1 (ADGRV1)	Q8WXG9	18.9
어서 증후군 2A (USH2A)	Usherin (USH2A)	O75445	15.6
망막 색소변성증 39 (RP39)	Usherin (USH2A)	O75445	15.6
알스톰 증후군	알스톰 증후군 단백질 1 (ALMS1)	Q8TCU4	12.5
뒤센 근위축증	디스트로핀 (DMD)	P11532	11.6
난청, 상염색체 열성, 3 (DFNB3)	비전형적인 미오신-XV (MYO15A)	Q9UKN7	10.8
어서 증후군 1D/F (USH1DF)	카드헤린-23 (CDH23)	Q9H251	10.1
어서 증후군 1D (USH1D)	카드헤린-23 (CDH23)	Q9H251	10.1
뇌하수체 선종 5, 여러 유형 (PITA5)	카드헤린-23 (CDH23)	Q9H251	10.1
난청, 상염색체 열성, 12 (DFNB12)	카드헤린-23 (CDH23)	Q9H251	10.1
망막 색소변성증 25 (RP25)	단백질 눈 셔트 상동체 (protein eyes shut homolog)(EYS)	Q5T1H1	9.4
근이영양증, 지대형 ar 23 (LGMDR23)	라미닌 서브타입 α-2 (LAMA2)	P24043	9.4
메로신-결합성 선천성 근이영양증 1A (MDC1A)	라미닌 서브타입 α-2 (LAMA2)	P24043	9.4
본 빌레브란트병 1 (VWD1)	폰 빌레브란트 인자 (VWF)	P04275	8.4
본 빌레브란트병 2 (VWD2)	폰 빌레브란트 인자 (VWF)	P04275	8.4
본 빌레브란트병 3 (VWD3)	폰 빌레브란트 인자 (VWF)	P04275	8.4
신생아 일과성 수포성 피부박리증 (TBDN)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
손발톱 장애, 무-증상성 선천성, 8 (NDNC8)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
각질화 결단을 동반한, 디스트로피성 수포성 표피박리증 (EBDSC)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
디스트로피성 수포성 표피박리증, Bart 타입 (B-DEB)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
소양성, 수포성 표피 박리증 (EBP)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
디스트로피성 수포성 표피 박리증, 전경골형 (PR-DEB)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1

디스트로피성 수포성 표피 박리증 ar (RDEB)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
디스트로피성 수포성 표피 박리증 ad (DDEB)	콜라겐 α-1(VII) 체인 (COL7A1)	Q02388	8.1
척수소뇌성 실조증 6 (SCA6)	전압-의존적인 P/Q형 칼슘 채널 서브타입 α-1A (CACNA1A)	000555	7.7
편두통, 가족성 반신불수, 1 (FHM1)	전압-의존적인 P/Q형 칼슘 채널 서브타입 α-1A (CACNA1A)	000555	7.7
간헐성 운동실조 2 (EA2)	전압-의존적인 P/Q형 칼슘 채널 서브타입 α-1A (CACNA1A)	000555	7.7
간질성 뇌병증, 초기 유아, 42 (EIEE42)	전압-의존적인 P/Q형 칼슘 채널 서브타입 α-1A (CACNA1A)	000555	7.7
시니어-로켄 증후군 6 (SLSN6)	290 kDa의 센트로솜 단백질 (CEP290)	015078	7.5
메켈 증후군 4 (MKS4)	290 kDa의 센트로솜 단백질 (CEP290)	015078	7.5
레베르 선천성 흑암시 10 (LCA10)	290 kDa의 센트로솜 단백질 (CEP290)	015078	7.5
주버트 증후군 5 (JBTS5)	290 kDa의 센트로솜 단백질 (CEP290)	015078	7.5
바멧-비들 증후군 14 (BBS14)	290 kDa의 센트로솜 단백질 (CEP290)	015078	7.5
잠복 황반 이영양증 (OCMD)	망막 색소변성증 1-유사 1 단백질 (RP1L1)	Q81WN7	7.2
고알도스테론혈증, 가족성, 4 (HALD4)	전압-의존적인 T형 칼슘 채널 서브타입 α-1H (CACNA1H)	095180	7.1
혈우병 A	응고 인자 FVIII	P00451	7.1
간질, 특발성 전신 6 (EIG6)	전압-의존적인 T형 칼슘 채널 서브타입 α-1H (CACNA1H)	095180	7.1
간질, 소아기 결신 6 (ECA6)	전압-의존적인 T형 칼슘 채널 서브타입 α-1H (CACNA1H)	095180	7.1
난청, ar, 28 (DFNB28)	TRIO 및 F-액틴-결합 단백질 (TRIOBP)	Q9H2D6	7.1
망막 색소변성증 13 (RP13)	Pre-mRNA-프로세싱-스플라이싱 인자 8 (PRPF8)	Q6P2Q9	7.0
스타가르트 질환 1 (STGD1)	망막-특이 인지질-수송 ATPase ABCA4 (ABCA4)	P78363	6.8
망막 색소변성증 19 (RP19=)	망막-특이 인지질-수송 ATPase ABCA4 (ABCA4)	P78363	6.8
황반 변성, 노화성, 2 (ARMD2)	망막-특이 인지질-수송 ATPase ABCA4 (ABCA4)	P78363	6.8
황색 반점 안저 (FFM)	망막-특이 인지질-수송 ATPase ABCA4 (ABCA4)	P78363	6.8
원뿔-막대세포 이상증 3 (CORD3)	망막-특이 인지질-수송 ATPase ABCA4 (ABCA4)	P78363	6.8
어서 증후군 1B (USH1B)	비전형적인 미오신-VIIa (MYO7A)	Q13402	6.7
난청 ar 2 (DFNB2)	비전형적인 미오신-VIIa (MYO7A)	Q13402	6.7
난청 ad 11 (DFNA11)	비전형적인 미오신-VIIa (MYO7A)	Q13402	6.7
망막 색소변성증 1 (RP1)	산소-조절되는 단백질 1 (RP1)	P56715	6.5
근이영양증, 지대형, ar 2 (LGMDR2)	Dysferlin (DYSF)	075923	6.5
미요시 근이영양증 1 (MMD1)	Dysferlin (DYSF)	075923	6.5
전경골근 발병 원위부 근병증 (DMAT)	Dysferlin (DYSF)	075923	6.5
난청, ad, 12 (DFNA12)	α-tectorin (TECTA)	075443	6.5
난청, ar, 21 (DFNB21)	α-tectorin (TECTA)	075443	6.5
망막 색소변성증 33 (RP33)	U5 소형 핵 리보핵단백질 200 kDa 헬리카제 (SNRNP200)	075643	6.4
진행성 가족성 심장 차단 1A (PFHB1A)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
간 QT 증후군 3 (LQT3)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
부르가다 증후군 1 (BRGDA1)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
동기능부전증후군 1 (SSS1)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
가족성 발작성 심실 원섬유화 1 (VF1)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
영아 돌연사 증후군 (SIDS)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
심방 정지 1 (ATRST1)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2

심근증, 확장 1E (CMD1E)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
심방세동, 가족성, 10 (ATFB10)	소듐 채널 단백질 타입 5 서브타입 α (SCN5A)	Q14524	6.2
발작, 양성 가족성 영아기, 3 (BFIS3)	소듐 채널 단백질 2형 서브타입 α (SCN2A)	Q99250	6.2
편두통, 가족성 반신불수, 3 (FHM3)	소듐 채널 단백질 1형 서브타입 α (SCN1A)	P35498	6.2
강직간대 발작을 동반한 난치성 소아 간질 (ICEGTC)	소듐 채널 단백질 1형 서브타입 α (SCN1A)	P35498	6.2
영아 연축	소듐 채널 단백질 2형 서브타입 α (SCN2A)	Q99250	6.2
열성 경련 플러스 2를 동반한 전신 간질 (GEFS+2)	소듐 채널 단백질 1형 서브타입 α (SCN1A)	P35498	6.2
열성 경련, 가족성, 3A (FEB3A)	소듐 채널 단백질 1형 서브타입 α (SCN1A)	P35498	6.2
간질성 뇌병증, 초기 영아, 6 (EIEE6)	소듐 채널 단백질 1형 서브타입 α (SCN1A)	P35498	6.2
간질성 뇌병증, 초기 영아, 62 (EIEE62)	소듐 채널 단백질 3형 서브타입 α (SCN3A)	Q9NY46	6.2
간질성 뇌병증, 초기 영아, 11 (EIEE11)	소듐 채널 단백질 2형 서브타입 α (SCN2A)	Q99250	6.2
간질, 다양한 초점의 가족성 국소 4 (FFEVF4)	소듐 채널 단백질 3형 서브타입 α (SCN3A)	Q9NY46	6.2
드라벡 증후군	소듐 채널 단백질 1형 서브타입 α (SCN1A)	P35498	6.2
난청, ar, 9 (DFNB9)	Otoferlin (OTOF)	Q9HC10	6.2
청각 신경병증, ar, 1 (AUNB1)	Otoferlin (OTOF)	Q9HC10	6.2
난청, ad, 4A (DFNA4A)	Myosin-14 (MYH14)	Q7Z406	6.2
말초 신경병증, 근병증, 애성 및 청각 상실 (PNMHH)	Myosin-14 (MYH14)	Q7Z406	6.2
난청, ar, 77 (DFNB77)	리폭시게나제 상동체 도메인-함유 단백질 1 (LOXHD1)	Q8IVV2	6.2
발작, 양성 가족성 영아, 5 (BFIS5)	소듐 채널 단백질 8형 서브타입 α (SCN8A)	Q9UQD0	6.1
근육간대경련, 가족성, 2 (MYOCL2)	소듐 채널 단백질 8형 서브타입 α (SCN8A)	Q9UQD0	6.1
간질성 뇌병증, 초기 영아, 13 (EIEE13)	소듐 채널 단백질 8형 서브타입 α (SCN8A)	Q9UQD0	6.1
소뇌 운동실조를 동반 또는 비-동반한 인지장애 (CIAT)	소듐 채널 단백질 8형 서브타입 α (SCN8A)	Q9UQD0	6.1
야맹증, 선천 전지성 2A (CSNB2A)	전압-의존적인 L형 칼슘 채널 서브타입 α-1F (CACNA1F)	O60840	6.0
원뿔-막대세포 이영양증, X-연결된 3 (CORDX3)	전압-의존적인 L형 칼슘 채널 서브타입 α-1F (CACNA1F)	O60840	6.0
알란트 섬 눈 질환 (AIED)	전압-의존적인 L형 칼슘 채널 서브타입 α-1F (CACNA1F)	O60840	6.0
어서 증후군 1F (USH1F)	프로토카드헤린-15 (PCDH15)	Q96QU1	5.9
어서 증후군 1D/F (USH1DF)	프로토카드헤린-15 (PCDH15)	Q96QU1	5.9
난청, ar, 23 (DFNB23)	프로토카드헤린-15 (PCDH15)	Q96QU1	5.9
난청, ar, 16 (DFNB16)	Stereocilin (STRC)	Q7RTU9	5.2
난청-불임 증후군 (DIS)	Stereocilin (STRC)	Q7RTU9	5.2
난청, ar, 30 (DFNB30)	미오신-IIa (MYO3A)	Q8NEV4	4.8
귀척추거대골단이형성증 ad (OSMEDA)	콜라겐 α-2(XI) 체인 (COL11A2)	P13942	4.7
귀척추거대골단이형성증 ar (OSMEDB)	콜라겐 α-2(XI) 체인 (COL11A2)	P13942	4.7
난청 ad 13 (DFNA13)	콜라겐 α-2(XI) 체인 (COL11A2)	P13942	4.7
난청 ar 53 (DFNB53)	콜라겐 α-2(XI) 체인 (COL11A2)	P13942	4.7
섬유연골형성 2 (FBCG2)	콜라겐 α-2(XI) 체인 (COL11A2)	P13942	4.7
낭포성 섬유증 (CF)	낭포성 섬유증 막횡단 전도 조절인자 (CFTR)	P13569	4.6
선천성 양측 정관 결손증 (CBAVD)	낭포성 섬유증 막횡단 전도 조절인자 (CFTR)	P13569	4.6

판코니 빈혈, 상보성군 A (FANCA)	판코니 빈혈 군 A 단백질 (FANCA)	015360	4.4
망막 색소변성증 12 (RP12)	단백질 크럼스 상동체 1 (CRB1)	P82279	4.2
색소 정맥염 맥락망막 위축증 (PPCRA)	단백질 크럼스 상동체 1 (CRB1)	P82279	4.2
레베르 선천성 흑내장 8 (LCA8)	단백질 크럼스 상동체 1 (CRB1)	P82279	4.2
난청 ad 22 (DFNA22)	비전형적인 미오신-VI (MYO6)	Q9UM54	4.1
난청 ar 37 (DFNB37)	비전형적인 미오신-VI (MYO6)	Q9UM54	4.1
비대성 심근증을 수반한 난청 ad 22 (DFNHCM)	비전형적인 미오신-VI (MYO6)	Q9UM54	4.1
난청, ar, 30 (DFNB30)	미오신-IIb (MYO3B)	Q8WXR4	4.0
레베르 선천성 흑내장 6 (LCA6)	X-연결된 망막 색소변성증 GTPase 조절인자-상호작용성 단백질 1 (RPGRIP1)	Q96KN7	4.0
원뿔-막대세포 이형양증 13 (CORD13)	X-연결된 망막 색소변성증 GTPase 조절인자-상호작용성 단백질 1 (RPGRIP1)	Q96KN7	4.0

[0082] LCA = 레베르 선천성 흑암시

[0083] RP = 망막 색소변성증

[0084] ad = 상염색체 우성

[0085] ar = 상염색체 열성

[0086] PR = 광수용체

[0087] RPE = 망막 색소 상피

[0088] ECM = 세포외 매트릭스

[0089] 따라서, 제1 측면은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접 연결된 분할 인테인 N-단편 (이하, "본 발명의 분할 인테인 N-단편")에 관한 것이며, 여기서 전술한 바와 같이, IntN은 직접 또는 링커를 이용해 N-말단 단편의 C-말단에 연결된다.

[0090] 본 발명의 제1 측면에 대한 바람직한 구현에는 테그론이 펩타이드 결합을 통해 직접 연결된 분할 인테인 N-단편 (이하, "본 발명의 분할 인테인 N-단편 테그론")에 관한 것이며, 테그론은 인테인 N-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접 연결된다.

[0091] 따라서, 본 발명의 분할 인테인 N-단편 테그론의 구성은 N-말단에서 C-말단으로 방향으로 (재구성할 단백질의 N-말단 부분) - (인테인-N 단편) - (테그론)일 것이다.

[0092] IntN과 테그론 사이에 링커가 도입되면, 본 발명의 분할 인테인 N-단편 테그론의 구조체의 구성은 (재구성할 단백질의 N-말단 부분) - (인테인-N 단편) - (링커) - (테그론)일 것이다.

[0093] 바람직하게는, 본 발명의 임의의 분할 인테인 N-단편의 인테인 N 또는 본 발명의 임의의 분할 인테인 N-단편 테그론의 인테인 N은 아래 표 2에 열거된 서열들 중 어느 하나로부터 선택할 수 있으며:

[0094] **표 2**

표 2

명칭	서열	서열번호
Cfa^N	CLSYDTEILTVEYGFLPIGKIVEERIECTVYTDKNGFVYTQPIAQWHNRGEQEVFEYCLEDGSI IRATKDHKFMITDQGMLPIDEIFERGLDLKQVDGLP	27
Cat^N	CLSGDTMIEILDDGGIIQKISMEDLYQRLA	30
Npu^N	CLSYETEILTVEYGLLPPIGKIVEKRIECTVYSVDNNGNIYTPVAQWHRGEQEVFEYCLEDGSL IRATKDHKFMITVDGQMLPIDEIFERELDLMRVDNLPN	32
Gp41N	CLDLKTQVQTPQGMKEISNIQVGLVLSTGYNEVLNVFPKSKKSYKITLEDGKEIICSEEHF PTQTGEMNISGGLKEGMCCLYVKE	38
ConN	CLSYDTEILTVEYGAVPIGKIVEENIECTVYSVDENGFTVYTQPIAQWHRGEQEVFEYCLEDGST IRATKDHKFMTEDEGMLPIDEIFEQGLDLKQVKGLPD	39

- [0096] 바람직하게는, 인테인 N은 서열번호 27의 CfaN 인테인 또는 이의 임의의 변이체, 예를 들어 서열번호 39의 ConN; 또는 서열번호 30의 CatN 인테인 또는 이의 임의의 변이체; 또는 서열번호 32의 NpuN 인테인 또는 이의 임의의 변이체; 또는 서열번호 38의 Gp41N 인테인 또는 이의 임의의 변이체이다.
- [0097] 본원에서, 용어 "변이체"는 특정 폴리펩타이드 서열과 실질적으로 유사한 폴리펩타이드 분자를 지칭한다. 본 발명에서, 이는 임의의 이의 폴리펩타이드 서열과 실질적으로 유사한 폴리펩타이드로서 Cfa, Npu, Cat 또는 Gp41 인테인의 C-인테인 또는 N-인테인의 변이체를 지칭하며, 예를 들어 Con 인테인은 본원에서 Cfa 변이체로서 이해된다. 따라서, 변이체는 변이체의 기원이 되는 폴리펩타이드와 구조 및 생물학적 활성이 유사할 수 있다. 즉, 변이체는 폴리펩타이드 서열의 돌연변이를 지칭할 수 있다. 용어 "돌연변이"는 기원이 되는 폴리펩타이드와 비교해 하나 이상의 아미노산이 부가, 결손, 치환되거나 또는 달리 화학적으로 변형된 폴리펩타이드 분자 서열을 지칭한다. 돌연변이는 기원이 되는 폴리펩타이드 분자와 실질적으로 동일한 특성을 보유할 수 있거나, 또는 대상 서열의 생물학적 활성이 결핍될 수 있다. 특정 구현예에서, 인테인 변이체는 비-촉매성 Cys 잔기, 즉 N-인테인의 첫번째 잔기가 세린 또는 알라닌으로 돌연변이된 돌연변이를 포함한다.
- [0098] 바람직하게는, 용어 "변이체"는 본 발명에서 일반적으로 이해되는 바와 같이 무차별성 (promiscuity)이 증가된 변이체를 지칭하며, 여기서 무차별성은 분할 부위에 바로 인접한 잔기의 동일성과는 독립적으로, 단백질 트랜스-스플라이싱 (PTS) 반응을 수행하는 인테인의 능력으로서 이해된다. 보다 구체적으로, 인테인 무차별성은 분할 부위, 즉 대상 단백질에 인테인이 삽입되는 부위에 바로 인접한 아미노산에 대한 동일성과는 독립적으로, PTS 반응을 수행하는 능력을 의미한다. 분할 부위를 아래에 대략적으로 도시한다:
- $$\begin{array}{ccccccc} -3 & -2 & -1 & & +1 & +2 & +3 \\ \text{X-X-X-인테인 N} & & & & \text{인테인 C-X-X-X} \end{array}$$
- [0099]
- [0100] 전형적으로, 인테인은 이들 위치에서 일부 아미노산에 대해 강한 선호성을 가진다. 예를 들어, Cat 인테인 (서열번호 30)은 +1번 위치에 Cys를, -1번 위치에 Glu를 선호한다 (Stevens, A. J., Sekar, G., Gramespacher, J. A., Cowburn, D., & Muir, T. W. (2018) An Atypical Mechanism of Split Intein Molecular Recognition and Folding. *Journal of the American Chemical Society*, 140(37), 11791-11799. <http://doi.org/10.1021/jacs.8b07334>). 더 많은 무차별 변이체는 심지어 이러한 바람직한 잔기가 없는 경우에도 단백질 트랜스-스플라이싱 반응을 양호한 수율로 수행할 수 있을 것이다.
- [0101] 특정 구현예에서, 인테인 변이체는, 인테인의 촉매성 및 제2 셸 촉진자 잔기 (shell accelerator residue)가 유지되고, 비-촉매성 잔기 또는 제2 셸 외부의 잔기만 돌연변이된, 돌연변이를 포함한다. 제2 셸 촉진자 잔기는, 인테인의 스플라이싱 활성의 제어하는데 중요한 역할을 수행하는, 인테인의 활성부에 인접한 잔기이다. 예를 들어, Cfa 인테인 및 DnaE 패밀리의 촉매성 및 촉진자 잔기는 잘 알려져 있으며, 당해 기술 분야에서 기술되어 있다 (Stevens, A. J., Brown, Z. Z., Shah, N. H., Sekar, G., Cowburn, D., & Muir, T. W. (2016). *Journal of the American Chemical Society*. <http://doi.org/10.1021/jacs.5b13528>). 마찬가지로, DnaE, GyrA, GyrB, DnaB, TerL, gp41, IMPDH 등의 복수의 다른 인테인 패밀리에서도 촉매 잔기가 조사 및 특정된 바 있다 (Shah, N. H., & Muir, T. W. (2014). Inteins: Nature's Gift to Protein Chemists. *Chemical Science* (Royal Society of Chemistry: 2010), 5(1), 446-461. <http://doi.org/10.1039/C3SC52951G>). 제2 셸 촉진자 잔기를 동정하기 위한 방법 역시 언급된 바 있으며 (Stevens et al. 2016 *Journal of the American Chemical Society*. <http://doi.org/10.1021/jacs.5b13528>), 이를 이용해 본 발명에서 이용하기에 적절한 특성을 가진 인테인 변이체 (N-인테인 및 C-인테인 둘다)를 구축할 수 있었다.
- [0102] 특정 구현예에서, 인테인 변이체는 인테인의 촉매 잔기가 유지되는 돌연변이를 포함하며, 변이체는 빠른 스플라이싱 속도 (반감기 5분 미만), 카오토로픽제 (chaotropic agent) 또는 고온의 존재 하에 높은 활성을 비롯한 CfaN, NpuN, CatN 또는 gp41N 인테인의 핵심적인 기능적인 특징을 유지한다. 보다 특정한 구현예에서, CfaN, CatN, NpuN 또는 gp41N의 변이체는 이들 임의의 서열들에 대한 기능적으로 동등한 변이체이다.
- [0103] 본원에서, 용어 "기능적으로 동등한 변이체"는 기능이 실질적으로 유지되거나 또는 개선되는 한, 아미노산 하나 이상의 변형, 삽입 및/또는 결손에 의해 서열로부터 파생된 모든 단백질을 의미하는 것으로 이해되며, 특히 분할 인테인 N-단편의 기능적으로 동등한 변이체의 경우 그 활성을 유지하는 것을 의미한다. 본원에서, 용어 "활성"은 분할 인테인 C-단편에 결합시 분할 인테인 N-단편이 단백질 트랜스-스플라이싱 반응을 수행할 수 있는 능력을 의미한다.

- [0104] CfaN, NpuN, CatN 또는 gp41N 인테인에 대한 기능적으로 동등한 변이체의 예는 아래에 나타낸다:
- [0105] - CfaN의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:
- [0106] Cys28 및/또는 Cys59가 Ser, Thr 또는 Ala으로 돌연변이된, CfaN 서열 (서열번호 27).
- [0107] 추가적인 아미노산, 예를 들어, 링커, 태그론 또는 검출용 태그가 Met1의 N-말단에 포함된, CfaN 서열 (서열번호 27).
- [0108] 유지되어야 하는 잔기: Cys1, Lys70, His72, Met75, Met81을 제외하고, 임의의 잔기가 물리화학적 특성이 비슷한 잔기로 돌연변이된, CfaN 서열 (서열번호 27).
- [0109] 유지되어야 하는 잔기: Cys1, Asp5, Phe15, Glu24, Thr32, Lys35, Phe38, Val39, Ile44, Asn49, Ile65, Thr69, Lys70, His72, Met75, Thr77, Met81, Gly91, Lys95, Gln96, Gly99를 제외하고는, 임의의 잔기가 물리화학적 특성이 비슷한 잔기로 돌연변이된, CfaN 서열 (서열번호 27).
- [0110] 잔기: Cys1, Lys70, His72, Met75, Met81은 유지되고, 돌연변이: Asp5Glu, Phe15Leu, Glu24Lys, Thr32Ser, Lys35Asn, Phe38Asn, Val39Ile, Ile44Val, Asn49Asp, Ile65Leu, Thr77Val, Gly91Glu, Lys95Met, Gln96Arg, Gly99Asn 중 어느 하나 또는 이의 임의 조합이 도입된, CfaN 서열 (서열번호 27).
- [0111] 잔기: Cys1, Lys70, His72, Met75, Met81은 유지되고, 돌연변이: Asp5Glu, Phe15Leu, Glu24Lys, Thr32Ser, Lys35Asn, Phe38Asn, Val39Ile, Ile44Val, Asn49Asp, Ile65Leu, Thr77Val, Gly91Glu, Lys95Met, Gln96Arg, Gly99Asn 중 1-6개 (그 범위에 1개 및 6개가 포함됨)가 도입된, CfaN 서열 (서열번호 27).
- [0112] 잔기: Cys1, Asp5, Glu24, Phe38, Val39, Ile44, Lys70, His72, Met75, Met81, Gly91, Lys95, Gln96, Gly99는 유지되고, 돌연변이: Phe15Ala, Thr32Ser, Lys35Glu, Asn49Asp, Ile65Thr, Thr77Glu 중 어느 하나 또는 임의 조합이 도입된, CfaN 서열 (서열번호 27).
- [0113] - NpuN의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:
- [0114] Cys28 및/또는 Cys59가 Ser, Thr 또는 Ala으로 돌연변이된, NpuN 서열 (서열번호 32).
- [0115] 유지되어야 하는 잔기: Cys1, Thr69, Lys70, His72, Met75, Met81을 제외한, NpuN의 임의의 잔기가 물리화학적 특성이 비슷한 잔기로 돌연변이된, NpuN 잔기 (서열번호 32).
- [0116] NpuN의 임의의 잔기가 CfaN 서열에서 대응되는 아미노산으로 돌연변이될 수 있는, NpuN 서열 (서열번호 32).
- [0117] - Gp41N의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:
- [0118] Cys59 및/또는 Cys83이 Ser, Thr 또는 Ala으로 돌연변이된, Gp41N 서열 (서열번호 38).
- [0119] 유지되어야 하는 잔기: Cys1, His63을 제외한, Gp41N의 임의의 잔기가 물리화학적 특성이 비슷한 잔기로 돌연변이될 수 있으며, 잔기 60이 Ser 또는 Thr인, Gp41N 서열 (서열번호 38).
- [0120] - CatN의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:
- [0121] 유지되어야 하는 Cys1을 제외한, CatN의 임의의 잔기가 물리화학적 특성이 비슷한 잔기로 돌연변이될 수 있는, CatN 서열 (서열번호 30).
- [0122] 구체적인 구현예에서, 인테인 N은 서열번호 27, 서열번호 30, 서열번호 32 또는 서열번호 38 각각에 대해 전체 서열에서 적어도 90%의 서열 동일성을 가진 서열번호 27, 서열번호 30, 서열번호 32 또는 서열번호 38의 아미노산 서열의 변이체를 포함하거나 또는 이로 구성된다. 특정 구현예에서, 서열번호 27, 서열번호 30, 서열번호 32 또는 서열번호 38의 인테인 N의 변이체는 서열번호 27, 서열번호 30, 서열번호 32 또는 서열번호 38 각각에 대해 전체 서열에서 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99%의 서열 동일성을 가진다.
- [0123] 2종 이상의 아미노산 서열 또는 뉴클레오타이드 서열의 맥락에서, 용어 "동일성", "동일한", "동일성 %" 또는 "서열 동일성"은, 서열 동일성의 일부로서 임의의 보존적인 아미노산 치환을 고려하지 않고, 최대한으로 일치되도록 비교 및 정렬 (필요에 따라 갭 도입)하였을 때, 명시된 동일한 아미노산 잔기 %를 가지거나 또는 동일한 2 이상의 서열 또는 부분 서열을 지칭한다. 동일성 %는 서열 비교 소프트웨어 또는 알고리즘을 이용해 또는 육안 검사에 의해 측정할 수 있다. 아미노산 서열의 정렬을 달성하기 위해 이용가능한 다양한 알고리즘 및 소프트웨어들이 당해 기술 분야에 공지되어 있다. 서열 정렬 알고리즘에 대한 이러한 비-제한적인 일 예는 N BLAST 및

XBLAST 프로그램 (Altschul et al., 1991, Nucleic Acids Res., 25:3389-402)에 통합된, Karlin et al., 1990, Proc. Natl. Acad. Sci., 87:2264-8에서 언급되고 Karlin et al., 1993, Proc. Natl. Acad. Sci., 90:5873-7에서 수정된 알고리즘이다. 특정 구현예에서, Gapped BLAST를 Altschul et al., 1997, Nucleic Acids Res. 25:3389-402에 기술된 바와 같이 이용할 수 있다. BLAST-2, WU-BLAST-2 (Altschul et al., 1996, Methods in Enzymology, 266:460-80), ALIGN, ALIGN-2 (Genentech, South San Francisco, California) 또는 Megalign (DNASTAR)은 서열 정렬에 이용할 수 있는 공개적으로 이용가능한 부가적인 소프트웨어 프로그램이다. 일부 대안적인 구현예에서, Needleman 및 Wunsch (J. Mol. Biol. 48:444-53 (1970))의 알고리즘을 통합한 GCG 소프트웨어 패키지의 GAP 프로그램을 이용해 아미노산 서열 2종 간의 동일성 %를 결정할 수 있다 (예를 들어, Blossum 62 매트릭스 또는 PAM250 매트릭스와, 갭 가중치 16, 14, 12, 10, 8, 6 또는 4 및 길이 가중치 1, 2, 3, 4, 5 적용). 대안적으로, 특정 구현예에서, 아미노산 서열들 간의 동일성 %는 Myers 및 Miller의 알고리즘 (CABIOS, 4:1 1-7 (1989))을 이용해 결정한다. 예를 들어, 동일성 %는 ALIGN 프로그램 (버전 2.0) 및 PAM120을 잔기 표, 갭 길이 패널티 12 및 갭 패널티 4를 적용해 이용함으로써 결정할 수 있다. 특정한 정렬 소프트웨어로 최대 정렬하기 위한 적절한 매개변수는 당해 기술 분야의 당업자에 의해 결정될 수 있다. 특정 구현예에서, 정렬 소프트웨어의 디폴트 매개변수를 이용한다. 특정 구현예에서, 제1 아미노산 서열의 제2 아미노산 서열에 대한 동일성 % "X"는 $100 \times (Y/Z)$ 로 계산하며, 여기서 Y는 (육안 검사 또는 특정한 서열 정렬 프로그램에 의해 정렬한) 제1 서열과 제2 서열의 정렬에서 동일한 매칭으로서 계산되는 아미노산 잔기의 개수이고, Z는 제2 서열에서 잔기의 총 수이다. 제2 서열이 제1 서열보다 더 길다면, 양쪽 서열 전체를 고려한 글로벌 정렬을 이용하며, 즉 각 서열에서 모든 문자 및 null을 정렬해야 한다. 이 경우, 상기와 동일한 식을 이용할 수 있으며, 단 Z 값은 제1 서열과 제2 서열이 중첩되는 영역의 길이이며, 이 영역은 제1 서열의 길이와 실질적으로 동일한 길이를 가진다.

[0124] 비-제한적인 예로서, 임의의 특정 폴리펩타이드가 참조 서열에 대해 특정한 % (예, 적어도 80%, 적어도 85%, 적어도 90%, 일부 구현예에서, 적어도 95%, 96%, 97%, 98% 또는 99%)의 서열 동일성을 가지는 지는, 특정 구현예에서, Bestfit 프로그램 (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 5371 1)을 이용해 확인할 수 있다. Bestfit은 Smith and Waterman, Advances in Applied Mathematics 2:482-9 (1981)의 국소 상동성 알고리즘을 이용해, 2 서열 간의 최상의 상동성 세그먼트를 찾는다. Bestfit 또는 특정 서열이 본 발명에 따른 참조 서열에 대해 예를 들어 95% 동일한지를 결정하기 위한 임의의 기타 서열 정렬 프로그램을 이용하는 경우, 동일성 %가 참조 아미노산 서열의 전장에 대해 계산하고 참조 서열의 뉴클레오타이드 총 수에 대해 최대 5% 상동성으로 갭을 허용하도록, 매개변수를 설정한다.

[0125] 분할 인테인 N-단편과 테그론 사이의 용어 "링커"는 인테인 N-단편의 C-말단을 테그론 서열의 N-말단과 연결하는 아미노산 서열을 지칭한다. 링커는 바람직하게는 아미노산 1-100개, 1-5개, 1-10개, 1-50개, 또는 1-25개로 이루어진 폴리펩타이드일 것이다. 링커는 Gly 풍부 펩타이드 또는 Gly-Ser 풍부 펩타이드일 수 있다. Gly-Ser 풍부 펩타이드에 대한 일 예는 서열 GGS이거나, 또는 일반식 (GGS)_n의 상기 링커의 폴리머일 수 있으며, 여기서 n은 1-10일 수 있다. 링커에 대한 일반식은 ((G)_nS)_y일 것이며, 여기서 n은 1-5이고, y는 1-10이다. 아울러, 예 피토프 태그, 예를 들어, 핵사히스티딘 (His6, 서열: GHHHHHHG (서열번호 84) 태그 또는 트리플 FLAG 태그 (3FT, 서열: DYKDHGDYKDHDIDYKDDDDK (서열번호 103))를 링커로서 이용할 수 있다.

[0126] 다른 특정 구현예에서, 펩타이드 결합으로 분할 인테인 N-단편에 직접 연결되는 테그론은 아래 표 3에 열거된 임의 서열로부터 선택된다:

[0127] **표 3**

표 3

테그론	서열	서열번호
CL1	ACKNWFSSLSHFVIHL	40
Deg1	MNKIPKDLLNPQITDEFKSSILDINKLFSICCNLPKLPESVTTEEEVELRDILGFLSRAN	41
PEST	RSSSPSDSDTSGFSSGSDHLSDLISLRLIS	42
DD1	RISFGSPPPMAXG; X=임의의 아미노산, 바람직하게는 Gly (G)	43
DD2	GSPPPMAXG; X=임의의 아미노산, 바람직하게는 Gly (G)	44
DD3 또는 VI5	TNGILKLGCG	45
M1	CSEIIPMSRSTPISTMG	46
M2	VSFANLNSLIVGILRFHW	47

SopE 또는 SopE100	MTKITLSPQNFRIQKQETTLKEKSTEKNSLAKSILAVKNHFIELRSKLSERFISHKNTESSATHFHRGSASEGRAVLTKNVKDFMLQTLNDIDIRGSA	48
SopE-1-78	MTKITLSPQNFRIQKQETTLKEKSTEKNSLAKSILAVKNHFIELRSKLSERFISHKNTESSATHFHRGSASEGRAVL	49
SopE-15-78	QETTLKEKSTEKNSLAKSILAVKNHFIELRSKLSERFISHKNTESSATHFHRGSASEGRAVL	50
SopE-15-50	QETTLKEKSTEKNSLAKSILAVKNHFIELRSKLS	51
L2 또는 CL2	SLISLPLPTRVKFSSLLIRIMKIITMTFPKKLRS	52
L6	FYYPIWFARVLLVHYQ	53
L9 또는 CL9	SNPFSSLFASLLIDSVSLKSNWDTSSSSCLISFFSSVMFSSTTRS	54
L10	CRQRFSCHLTASYPQSTVTPFLAFLRRDFFLRHNSSAD	55
L11	GAPHVVLFDLRLITNPLSHIQSVSLQITLIFCSLPSLILSKFLQV	56
L12 또는 CL12	NTPLFSKSFSTTCGVAKKTLLAQISSLFFLLSSNIAV	57
L15	PTVKNSPKIFCLSSPYLAFNLEYLSLRIFSTLSKCSNTLLTSL	58
L16	SNQLKRLWLWLEVRSFDRTLRRPWHLPS	59
M3	KSVTLERSPKFLNWFVSLFKVITTG	34
M4	YMSILRCASGKISIAAPPYIF	35
M5	AGESFNFVMKLLYKHPILPCLKTLLSIRSSCSPR	37
VI2	TNGSVLREFTLLELEVTRNTEL	85
DD4	RDSFGSPPPMAXG; X=임의의 아미노산, 바람직하게는 Gly (G)	80
DD5	GFWI	81
DD6	RFKG	82
DD7	KFYK	83

- [0129] 데그론에 대한 비-제한적인 기타 예로는 DHFR (데하이드로폴레이트 리덕타제), FKBP (FK506 결합 단백질), FRB (FKBP-라파마이신 결합 단백질) 및 PDE5 (포스포다이에스테라제 타입 5)와 관련한 폴리펩타이드들뿐 아니라 소포체 관련 분해 (ERAD) 경로의 분해 신호를 포함한다.
- [0130] 또 다른 바람직한 구현예에서, 데그론은 아미노산 75개 미만의 폴리펩타이드이며, 이는 본 발명의 인테인 N 단편에 연결되어 이의 분해를 유도한다.
- [0131] 또 다른 바람직한 구현예에서, 데그론은 본 발명의 분할 인테인 N-말단 단편 또는 분할 인테인 C-말단 단편에 융합되었을 때, 이 단편의 발현 수준을 10% 초과, 20% 초과, 30% 초과, 50% 초과, 75% 초과 또는 90% 초과로 감소시키는, 폴리펩타이드이다. 중요하게는, 본 발명의 단편에 데그론의 병합은 본 발명의 분할 인테인 N-말단 단편과 분할 인테인 C-말단 단편 간의 단백질 트랜스-스플라이싱으로부터 발생하는 재구성된 단백질의 수율에 현저한 영향을 미치지 않는다. 이러한 바람직한 구현예의 목적에서, 발현 수준은 도 18에 도시된 바와 같이 측정하며, 즉 데그론을 가진 단편의 발현 수준은 도 18에 나타난 조건과 유사한 조건에서 데그론이 없는 단편의 발현 수준과 비슷하며, 전자가 후자의 적어도 10% 또는 적어도 20% 또는 적어도 30% 또는 적어도 50% 또는 적어도 75% 또는 적어도 90%로 낮아야 한다.
- [0132] 또 다른 바람직한 구현예에서, 본 발명의 분할 인테인 N-단편 데그론의 N 인테인과 데그론의 바람직한 조합은 하기로 이루어진 목록으로부터 선택된다:
- [0133] - CfaN 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7. 바람직하게는, CfaN 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4; 또는
- [0134] - Gp41N 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7. 바람직하게는, Gp41N 또는 이의 기능적으로 동등한 변이체; 및 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4.
- [0135] 도 18A에서, 일부 대표적인 데그론이 분할 N-인테인 단편, 특히 CfaN 인테인에 융합되었을 때, 가지는 효과를 보여준다. 데그론 (SopE, L2, L9, M4, V12)의 부가는 데그론이 융합된 N-인테인 함유 단백질 융합물의 검출가능한 발현 수준을 현저하게 감소시킨다. 흥미롭게도, 실시예 및 도 18, 20 및 21에 나타난 바와 같이, 일부 바람직한 데그론 (SopE, L2, L9, M4 또는 V12)의 함유는 스플라이싱 산물의 생성 수율에는 부정적인 영향을 미치지 않는다. 이러한 결과는, 데그론 부가에 의해 달성되는 바와 같이, 단백질 트랜스-스플라이싱 반응의 출발 물질

수준 감소시 단백질 스플라이싱 수율 감소가 예상될 것이므로, 당해 기술 분야에서는 예상하지 못하였다.

[0136] 또 다른 바람직한 구현예에서, 분할 인테인 N-단편 및 테그론의 바람직한 조합은 다음으로 이루어진 목록으로부터 선택된다: CfaN 인테인; 및 테그론 DD1, DD2, DD3, SopE, L2, L9, M4, V12, DD4, DD5, DD6 또는 DD7 중 임의의 테그론.

[0137] 따라서, 본 발명자들은, CfaN 인테인이 전술한 임의의 테그론, 특히 DD1, DD2, DD3, SopE, L2, L9, M4, V12, DD4, DD5, DD6 또는 DD7과 조합시, N-말단 인테인 단편의 발현 수준 감소가 검출되는 것으로, 관찰하였다 (도 9, 18A). 중요하게도, 이는 단백질 트랜스-스플라이싱 반응의 수율에는 영향을 미치지 않았다 (도 9, 18C, 20 또는 21). 이런 결과는, 전형적으로 단백질 트랜스-스플라이싱 반응에 참여하는 인테인 단편 2개 중 하나의 수율, 발현 또는 안정성의 저하가 생성되는 스플라이싱 산물의 전체적인 양 감소로 이어지므로, 선행 기술에서 예상하지 못하였다.

[0138] 본 발명자들은 이러한 효과를 ABCA4와 같은 막 단백질 (도 18, 20 및 21), 뿐 아니라 세포질 용해성 단백질, 예를 들어 EGFP에서도 관찰하였으며, 이는 작용의 일반성을 예시해준다. 도 9에서, EGFPN-CfaN에 테그론 부가시 단백질에 해당하는 밴드가 완전히 없어지는 것을 관찰할 수 있었다. 그러나, (테그론 함유, 또는 테그론 비-함유) CfaC-EGFP 단편 부가시, 전장 EGFP 산물은 테그론 비-함유 경우와 동일한 수율로 생성되었다. 이 예는, 테그론 하나 또는 심지어 2중 (각 단편에 하나)의 부가시, 표적 단백질의 수율에 부정적인 영향이 없다는 것을, 보여준다.

[0139] 분할 인테인 N-단편을 포함하는 복합체

[0140] 다른 측면에서, 본 발명은 다음 복합체/복합체들 중 임의의 복합체에 관한 것이다: 본 발명의 분할 인테인 N-단편 또는 본 발명의 분할 인테인 N-단편 테그론, 이하, 본 발명의 제1 복합체/복합체들, 각 복합체는

[0141] (i) 재구성할 대상 단백질의 N-말단 단편, 및

[0142] (ii) 상기 단락 (본 발명의 분할 인테인 N-단편 테그론)에서 정의된 바와 같이, 테그론이 펩타이드 결합으로 직접, 선택적으로는 링커를 통해 연결된, 분할 인테인 N-단편 또는 분할 인테인 N-단편을 포함하고;

[0143] 복합체/복합체들은 선택적으로 (i)과 (ii) 사이에 링커를 포함하고,

[0144] 여기서,

[0145] - 대상 단백질의 N-말단 단편은 분할 인테인 N-단편의 N-말단에 아미드 연결을 통해 연결되거나, 또는

[0146] - 복합체/복합체들이 링커를 포함할 경우, 대상 단백질의 N-말단 단편은 아미드 연결에 의해 링커에 결합되거나 및/또는 링커는 아미드 연결에 의해 분할 인테인 N-단편의 N-말단에 결합된다.

[0147] 상기한 복합체들 중 어느 하나에 이용가능한 대상 단백질에 대한 예시적인 비-제한적인 예는 상기 표 1에 나타난 것들이다. 추가적인 대상 단백질은 항체, 항체 단편, 예를 들어, Fc 도메인, scFv, 나노바디, 이중 특이성 항체, 단백질, 바람직하게는 >25, >50, >100 KDa의 임의의 단백질들로부터 선택될 수 있다.

[0148] 이미 언급한 바와 같이, 선택적으로, 대상 단백질 및 분할 인테인 N-단편은, 대상 단백질과 N-인테인 사이에 링커가 위치하도록 링커를 통해 연결될 수 있다. 링커의 특성은 대상 단백질의 특성에 따라 결정될 것이다. 특정 구현예에서, 링커는 펩타이드이다. 특정 구현예에서, 링커는 아미노산 1개, 2개, 3개, 4개, 5개, 10개, 20개, 50개, 100개 또는 그 이상의 길이의 펩타이드이다. 바람직하게는, 링커의 N-말단은 대상 단백질의 C-말단에 연결되고, 링커의 C-말단은 펩타이드 결합으로 N-인테인의 N-말단에 연결된다.

[0149] 다른 특정 구현예에서, 복합체는 대상 단백질의 N-말단 단편과 분할 인테인 N-단편 사이에 링커를 포함하지 않는다. 이러한 특정 구현예에서, 대상 단백질은 분할 인테인 N-단편의 N-말단에 아미드 연결을 통해 연결된다.

[0150] 다른 특정 구현예에서, 복합체가 링커를 포함할 경우, 복합체는 융합 단백질이다. 용어 "융합 단백질"은 당해 기술 분야에 잘 알려져 있으며, 서로 다른 기원, 천연 및/또는 인공적인 서열 2 이상을 포함하는 인공적으로 고안된 단일한 폴리펩타이드 체인을 지칭한다. 융합 단백질은 정의에 따라 자연계에서 그 자체로 발견되지 않는다.

[0151] 분할 인테인 C-단편

[0152] 다른 측면에서, 본 발명은 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접 연결된 분할 인테인 C-단

편에 관한 것이다 (이하, "본 발명의 분할 인테인 C-단편").

[0153] 더 바람직하게는, 본 발명은 테그론에 펩타이드 결합으로 직접 연결된 분할 인테인 C-단편에 관한 것이며 (이하, "본 발명의 분할 인테인 C-단편 테그론"), 테그론은 인테인 C-단편과 테그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 C-단편의 C-말단은 재구성할 단백질의 C-말단 단편의 N-말단에 펩타이드 결합으로 직접 연결된다.

[0154] 따라서, 본 발명의 분할 인테인 C-단편 테그론의 구조체 구성은 N-말단에서 C-말단 방향으로 (테그론) - (인테인-C 단편) - (재구성할 단백질의 C-말단 부분)일 것이다.

[0155] 테그론과 인테인 C 사이에 링커가 존재할 경우, 본 발명의 분할 인테인 C-단편 테그론의 구조체 구성은 N-말단에서 C-말단 방향으로 (테그론) - (링커) - (인테인-C 단편) - (재구성할 단백질의 C-말단 부분)일 것이다.

[0156] 바람직하게는, 본 발명의 임의의 분할 인테인 C-단편의 인테인 C 또는 본 발명의 분할 인테인 C-단편 테그론의 인테인 C는 하기 표에 열거된 것들 중 임의의 것으로 선택되며:

[0157] 표 4

표 4

명칭	서열	서열번호
Cfa^C	MVKIIISRKSLGTQNVYDIGVEKDHNFLLKGLVASN	28
Cfa^{Cmut}	MVKIIISRKSLGTQNVYDIGVGEPHNFLLKGLVASN	29
Cat^C	MFKLNTKNIKVLTPSGFKSFGSIQKVYPFYHHIIFFDGEIKCSDNHSFGKDKIKASTIKVGDYLGKKVLYNEIVEEGIYLYDLLNVGEDNLYYTNGIVSHN	31
Npu^C	MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN	33
Npu^{Cmut}	MIKIATRKYLGKQNVYDIGVGEPHNFALKNGFIASN	36
Gp41C	MMLKKILKIEELDERELIDIEVSGNHLFYANDILTHN	104
ConC	MVKIIISRQSLGKQNVYDIGVEKDHNFLLANGLIASN	105

[0159] 바람직하게는, 인테인 C는 서열번호 28의 인테인 또는 CfaCmut (서열번호 29) 또는 ConC (서열번호 105)와 같은 임의의 변이체이거나; 또는 서열번호 31의 CatC 인테인 또는 임의의 변이체, 또는 NpuC 인테인 (서열번호 33) 또는 NpuCmut (서열번호 36)와 같은 임의의 변이체, 또는 서열번호 104의 GP41C 인테인 또는 임의의 변이체; 또는 상기에서 정의되는 바와 같이 무차별성이 증가된 임의의 변이체이다. 인테인 C의 변이체는 서열번호 28, 29, 31, 33, 36, 104 및 105를 포함하며, N-말단 메티오닌 잔기가 존재하지 않는다.

[0160] 특정 구현예에서, 무차별성이 증가한 CfaC, NpuC 또는 ConC 변이체들은 Stevens et al. Proc. Natl. Acad. Sci. 2017에 언급된 바와 같이 21, 22 및 23번 위치에 아미노산 Gly, Glu 및 Pro를 가진다.

[0161] 특정 구현예에서, CfaC, NpuC CatC 또는 gp41c 인테인의 기능적으로 동등한 변이체에 대한 예는 다음과 같이 설명된다:

[0162] - CfaC의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:

[0163] Met1이 제거된 CfaC (서열번호 28) 서열.

[0164] N-말단 Met1에 부가적인 아미노산, 예를 들어, 링커, 테그론 또는 검출용 태그를 가진, CfaC (서열번호 28) 서열.

[0165] 유지되어야 하는 잔기: Asp17, His24, Asn36을 제외한 CfaC의 임의 잔기가 물리화학적 특성이 비슷한 잔기로 치환된 CfaC (서열번호 28) 서열.

[0166] 유지되어야 하는 잔기: Asp17, Asp23, His24, Ser35, Asn36을 제외한 CfaC의 임의 잔기가 물리화학적 특성이 비슷한 잔기로 치환된 CfaC (서열번호 28) 서열.

[0167] 유지되어야 하는 잔기: Asp17, Asp23, His24, Ser35, Asn36을 제외한 CfaC의 임의 잔기가 물리화학적 특성이 비슷한 잔기로 치환된 CfaC (서열번호 28) 서열.

- [0168] 잔기: Asp17, Asp23, His24, Ser35, Asn36은 유지되고 아미노산: Val2, Ile5, Ser6, Ser9, Lys22, Leu27, Leu32, Val33 중 임의의 잔기는 돌연변이된, CfaC (서열번호 28) 서열.
- [0169] 잔기: Glu21Gly, Lys22Glu 및 Asp23Pro이 돌연변이된, CfaC (서열번호 28) 서열. 이 변이체는 CfaCmut 인테인 (서열번호 29)에 해당한다.
- [0170] 지정된 위치에 존재해야 하는 잔기: Asp17, Gly21, Glu22, Pro23, His24, Asn36을 제외한 임의의 잔기가 물리화학적 특성이 유사한 잔기로 돌연변이된, CfaC (서열번호 28) 서열.
- [0171] 지정된 위치에 잔기: Asp17, Gly21, Glu22, Pro23, His24, Ser35, Asn36이 존재하고 아미노산: Val2, Ile5, Ser9, Leu27, Leu32, Val33 중 어느 것이 돌연변이된, CfaC (서열번호 28) 서열.
- [0172] 돌연변이: Val2Ile, Ile5Ala, Ser6Thr, Ser9Tyr, Thr12Lys, Leu27Ala, Leu32Phe, Val33Ile 중 1-5개 (1 및 5가 범위에 포함됨)가 도입된, CfaC (서열번호 28) 또는 CfaCmut (서열번호 29) 서열.
- [0173] - NpuC의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:
- [0174] Met1이 제거된 NpuC (서열번호 33) 서열.
- [0175] Met1의 N-말단에 추가적인 아미노산, 예를 들어, 링커, 테그론 또는 검출용 태그가 포함된, NpuC (서열번호 33) 서열.
- [0176] 유지되어야 하는 잔기: Asp17, His24, Asn36을 제외한 NpuC의 임의의 잔기가 물리화학적 특성이 유사한 잔기로 돌연변이될 수 있는, NpuC (서열번호 33) 서열.
- [0177] 유지되어야 하는 잔기: Asp17, His24, Ser35, Asn36을 제외한 NpuC의 임의의 잔기가 물리화학적 특성이 유사한 잔기로 돌연변이될 수 있는, NpuC (서열번호 33) 서열.
- [0178] 잔기: Asp17, His24, Ser35, Asn36은 유지되고 아미노산: Val2, Ile5, Ser9, Leu27, Leu32, Val33 중 임의의 아미노산은 돌연변이된, NpuC (서열번호 33) 서열.
- [0179] 잔기: Asp17, His24, Ser35, Asn36은 유지되고 돌연변이: Ile2Val, Ala5Ile, Tyr9Ser, Ala27Leu, Phe32Leu, Ile33Val 중 임의의 돌연변이는 병합된, NpuC (서열번호 33) 서열.
- [0180] Npu의 무차별성을 높이기 위해 잔기: Glu21Gly, Arg22Glu, Asp23Pro가 돌연변이된, NpuC (서열번호 33) 서열.
- [0181] - Gp41C의 기능적으로 동등한 변이체는 하기로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다:
- [0182] Met1이 제거된 Gp41C (서열번호 104) 서열.
- [0183] Met1의 N-말단에 추가적인 아미노산, 예를 들어 링커, 테그론 또는 검출용 태그가 함유된, Gp41C (서열번호 104) 서열.
- [0184] 유지되어야 하는 잔기: His26, His36, Asn37을 제외한 Gp41C의 임의의 잔기가 물리화학적 특성이 유사한 잔기로 돌연변이될 수 있는, Gp41C (서열번호 104) 서열.
- [0185] 특정 구현예에서, 인테인 C는 서열번호 28, 서열번호 31, 서열번호 33 또는 서열번호 104 각각에 대해, 전체 서열에서, 적어도 90%의 서열 동일성을 가진 서열번호 28, 서열번호 31, 서열번호 33 또는 서열번호 104의 아미노산 서열의 변이체를 포함하거나 또는 이로 구성된다. 특정 구현예에서, 서열번호 28, 서열번호 31, 서열번호 33 또는 서열번호 104의 인테인 C의 변이체는 서열번호 28, 서열번호 31, 서열번호 33 또는 서열번호 104 각각에 대해 전체 서열에서 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99%의 서열 동일성을 가진다.
- [0186] 다른 특정 구현예에서, 분할 인테인 C-단편에 펩타이드 결합으로 직접 연결된 테그론은 표 3에 열거된 것들 중 임의의 것으로 선택된다. 본 섹션에서 이용가능한 대상 단백질에 대한 예시적인 비-제한적인 예는 상기 표 1에 나타낸 것이다.
- [0187] 또 다른 바람직한 구현예에서, 본 발명의 분할 인테인 C-단편 테그론에서 C 인테인과 테그론의 바람직한 조합은 하기로 이루어진 목록으로부터 선택된다:
- [0188] - CfaC 또는 이의 기능적으로 동등한 변이체, 예를 들어 CfaCmut; 및 하기 테그론들 중 임의의 것의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12,

L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7. 바람직하게는, CfaC 또는 이의 기능적으로 동등한 변이체, 예를 들어 CfaCmut; 및 하기 테그론들 중 임의의 것의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, DD4, DD5, DD6 또는 DD7; 또는

[0189] - Gp41C 또는 이의 기능적으로 동등한 변이체; 및 다음 테그론들 중 임의의 테그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7. 바람직하게는, Gp41C 또는 이의 기능적으로 동등한 변이체; 및 다음 테그론들 중 임의의 테그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4.

[0190] 도 18B에서, C-말단 구조체에 테그론의 융합 효과를 관찰할 수 있다. N-말단 테그론에 대해서 관찰되는 바와 같이, 바람직한 조합을 이용할 경우, C-말단 구조체에의 테그론 융합은 구조체의 발현 수준을 감소시키지만, 스플라이싱 수율에 대한 효과는 관찰되지 않는다. 다음의 테그론은 특히 CfaC와 잘 작동한다: DD1, DD2, DD3, SopE, L2, L9, M4, V12, DD4, DD5, DD6 또는 DD7. N-말단 단편에 대해 관찰되는 바와 같이, 본 발명의 C-말단 단편에 테그론의 병합시 (WB에 의해 검출되는 바와 같이) 발현을 감소시키지만, 흥미롭게도, PTS 수율, 즉 원하는 산물의 형성에는 부정적인 효과가 관찰되지 않았다. 이러한 결과는 EGFP 및 ABCA4에 대해서 본원에 나타낸 예들을 비롯하여, 수종의 단백질들에서 관찰되었다.

[0191] **분할 인테인 C-단편을 포함하는 복합체**

[0192] 다른 측면에서, 본 발명은 하기 복합체/복합체들 중 임의의 복합체에 관한 것이다: 본 발명의 분할 인테인 C-단편 또는 본 발명의 분할 인테인 C-단편 테그론, 이하 본 발명의 제2 복합체/복합체들, 여기서 이들 복합체 각각은

[0193] (i) 대상 단백질의 C-말단 단편, 및

[0194] (ii) 상기 섹션에서 정의된 바와 같이, 테그론에 펩타이드 결합으로 직접, 선택적으로 링커를 통해, 연결된, 분할 인테인 C-단편 또는 분할 인테인 C-단편 (본 발명의 분할 인테인 C-단편 테그론)을 포함하고,

[0195] 복합체/복합체들은 선택적으로 (i)과 (ii) 사이에 링커를 포함하고,

[0196] 여기서,

[0197] - 대상 단백질의 C-말단 단편은 분할 인테인 C-단편의 C-말단에 아미드 연결에 의해 연결되거나, 또는

[0198] - 복합체/복합체들이 링커를 포함할 경우, 대상 단백질의 C-말단 단편은 링커에 아미드 연결을 통해 결합되거나, 및/또는 링커는 분할 인테인 C-단편의 C-말단에 아미드 연결을 통해 결합된다.

[0199] 본 발명에서 이용가능한 대상 단백질에 대한 예시적인 비-제한적인 예는 상기 표 1에 열거된 것이다. 대상 단백질에 대한 추가적인 예는 항체, 항체 단편, 예를 들어, Fc 도메인, scFv, 나노바디, 이중 특이성 항체, 단백질, 바람직하게는 >25, >50, >100 KDa의 임의의 단백질들로부터 선택될 수 있다.

[0200] 용어 "대상 단백질" 및 "링커"는 본 발명의 제1 복합체와 관련하여 앞서 정의된 바 있다. 본 발명의 제1 복합체의 링커 및 대상 단백질에 대한 모든 구체적인 구현예들은 본 발명의 제2 복합체에도 모두 적용된다.

[0201] 특정 구현예에서, 복합체는 대상 화합물과 분할 인테인 C-단편 사이에 링커를 포함하지 않는다. 이러한 특정 구현예에서, 대상 화합물은 분할 인테인 C-단편의 C-말단에 아미드 연결에 의해 연결된다.

[0202] 특정 구현예에서, 복합체는 대상 화합물과 분할 인테인 C-단편 사이에 링커를 포함한다. 이러한 특정 구현예에서, 대상 화합물은 링커에 대상 화합물 및 링커의 화학적 성질에 따라 임의의 적절한 수단에 의해 결합될 수 있다. 이러한 특정 구현예에서, 링커는 분할 인테인 C-단편의 C-말단에 아미드 연결에 의해 결합된다. 다른 특정 구현예에서, 대상 화합물은 링커에 아미드 연결에 의해 연결되며, 이 경우 링커는 분할 인테인 C-단편의 C-말단에 임의의 적절한 수단에 의해 결합될 수 있다. 다른 특정 구현예에서, 대상 화합물은 링커에 아미드 연결에 의해 결합되고, 링커는 분할 인테인 C-단편의 C-말단에 아미드 연결에 의해 결합된다.

[0203] 특정 구현예에서, 복합체가 링커를 포함할 경우, 링커는 펩타이드 링커이다. 이러한 특정 구현예에서, 복합체는 융합 단백질이다.

[0204] **본 발명의 복합체를 포함하는 조성물**

[0205] 다른 측면에서, 본 발명은 본 발명의 제1 복합체/복합체들 및/또는 제2 복합체/복합체들을 포함하는, 이하 본

발명의 제1 조성물로 지칭되는, 조성물에 관한 것이다.

[0206] 용어 "조성물"은 명시된 구성성분을 함유한 제품뿐 아니라 명시된 양의 명시된 구성성분들의 조합으로부터 직접 또는 간접적으로 생성되는 임의의 제품을 포괄하는 것으로 의도된다. 조성물의 구성성분들은 함께 단일 제형으로 또는 서로 다른 제형으로 분리하여 포장될 수 있다. 따라서, 일 구현예에서, 본 발명의 제1 복합체는 본 발명의 제2 복합체와 단일 제형으로 함께 포장된다. 다른 구현예에서, 본 발명의 제1 복합체와 본 발명의 제2 복합체는 분리하여 포장된다.

[0207] 특히 바람직한 구현예에서, 제1 복합체 및 제2 복합체는, 이들 2가지 복합체가 본 발명의 방법에 따라 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편과 연결되어 전체 단백질이 생성되는 방식으로, 동일 단백질의 N-말단 단편 및 C-말단 단편을 각각 포함한다.

[0208] 본 발명의 폴리뉴클레오티드, 벡터 및 숙주 세포

[0209] 다른 측면에서, 본 발명은 본 발명의 제1 복합체 또는 제2 복합체를 코딩하는 폴리뉴클레오티드에 관한 것이다.

[0210] 바람직하게는, 바람직한 구현예에서, 본 발명은 하나의 폴리뉴클레오티드가 본 발명의 제1 복합체/복합체들을 코딩하고 (이하, 본 발명의 제1 폴리뉴클레오티드), 다른 폴리뉴클레오티드는 본 발명의 제2 복합체/복합체들을 코딩하며 (이하, 본 발명의 제2 폴리뉴클레오티드); 본 발명의 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드는, 폴리뉴클레오티드 2종이 모두 각각의 단백질 복합체로 번역되거나 및/또는 본 발명의 방법에 따라 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편에 연결되어 재구성할 전체 단백질이 생성되는 방식으로, 각각 동일한 단백질의 N-말단 단편 및 C-말단 단편을 코딩하는, 2종의 폴리뉴클레오티드에 관한 것이다. 본 발명의 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드가 각각 바람직하게는 동일한 인테인의 분할 인테인을 코딩하고, 즉 N-인테인 및 C-인테인이 동족 쌍 (cognate pair)임에 유념하는 것이 중요하다 (Shah et al. Journal of the American Chemical Society 2012 <https://doi.org/10.1021/ja303226x>). 동족 인테인 쌍에 대한 예는 CfaN/CfaC (또는 CfaCmut 등의 임의의 변이체), NpuN/NpuC (또는 임의의 변이체), CatN/CatC (또는 임의의 변이체), Gp41N/Gp41C (또는 임의의 변이체)를 포함한다. 그러나, 폴리뉴클레오티드 2종이 모두 동일한 인테인에 속하는 분할 인테인을 코딩하는지 유무와는 독립적으로, 단백질로 번역되면, 분할 인테인 각각의 N-말단 및 C-말단은 트랜스-스플라이싱 반응에 기능하는 인테인으로 비-공유적으로 재-연합하거나 또는 재구성할 수 있는 별도의 단편들이 되어야 한다.

[0211] 바람직한 구현예에서, 본 발명은, 폴리뉴클레오티드 하나는 본 발명의 분할 인테인 N-단편 테그론을 코딩하고 (이하, 본 발명의 테그론을 코딩하는 제1 폴리뉴클레오티드), 다른 폴리뉴클레오티드는 본 발명의 분할 인테인 C-단편 테그론을 코딩하며 (이하, 본 발명의 테그론을 코딩하는 제2 폴리뉴클레오티드); 본 발명의 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드는, 폴리뉴클레오티드 2종이 모두 각각의 단백질 복합체로 번역되어 본 발명의 방법에 따라 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편에 연결되어 전체 단백질이 생성되는 방식으로, 각각 동일한 단백질의 N-말단 단편 및 C-말단 단편을 코딩하는, 상기 언급된 2종의 폴리뉴클레오티드에 관한 것이다. 전술한 구현예에서와 마찬가지로, 본 발명의 테그론을 코딩하는 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드가 각각 바람직하게는 동일한 인테인의 분할 인테인을 코딩함에 유념하는 것이 중요하다. 적어도 상기에서 반영된 바와 같이, 본원에서 각각의 폴리뉴클레오티드는, 단백질로 번역되면, N-말단과 N-말단 서열이 트랜스-스플라이싱 반응에서 기능하는 인테인으로 비-공유적으로 재-연합하거나 또는 재구성할 수 있는 별도의 단편들이 되도록, 분할 인테인들을 코딩하여야 한다.

[0212] 또 다른 바람직한 구현예에서, 본 발명은 오르소고날 분할 인테인 쌍들을 이용한 3-조각 결찰 전략에 관한 것이다. 이러한 방법에서, 제1 폴리뉴클레오티드는 POI^N -CfaN을 코딩하고 (POI^N 은 대상 단백질의 N 단편으로서 이해됨); 제2 폴리뉴클레오티드는 CfaC- POI^M -CatN을 코딩하고 (이 경우에 POI^M 은 대상 단백질의 중간산물 단편임); 제3 폴리뉴클레오티드는 CatC- POI^C 를 코딩하는 (POI^C 은 대상 단백질의 C 단편으로서 이해됨), 폴리뉴클레오티드 3종이 존재한다. 대안적으로, Cfa 인테인 및 Cat 인테인의 위치는, 다음의 구조 구성을 가진 구조체를 생성하도록, 교체될 수 있다: POI^N -CatN, CatC- POI^M -CfaN 및 CfaC- POI^C . CfaN은 서열번호 27의 CfaN 인테인 단편 서열번호 27 뿐 아니라 이의 임의의 변이체이다. CatN은 서열번호 30의 CatN 인테인 단편 뿐 아니라 이의 임의의 변이체이다. CfaC는 서열번호 28의 CfaC 인테인 단편 뿐 아니라 서열번호 29 등의 이의 임의의 변이체이다. CatC는 서열번호 31의 CatC 인테인 단편 뿐 아니라 이의 임의의 변이체이다.

[0213] 본원에서, 폴리뉴클레오티드 2종을 참조하는 구현예들에서, 제1 폴리뉴클레오티드 (또는 구체적으로, 본 발명의

데그론을 코딩하는 제1 폴리뉴클레오타이드)가 본 발명의 임의의 분할 인테인 N-단편 또는 본 발명의 임의의 분할 인테인-N 데그론 단편일 수 있음에 유념한다. 구체적으로, 본 발명의 제1 폴리뉴클레오타이드 (또는 구체적으로, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오타이드)는 서열번호 1,3, 5, 8, 10, 12, 14, 17, 19, 21, 23, 24, 25, 66, 68, 70, 72, 74, 76 및 78로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다. 더 바람직하게는, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오타이드는 다음과 같은 구체적인 인테인과 데그론의 임의 조합을 코딩한다:

- [0214] - CfaN 또는 이의 기능적으로 동등한 변이체, 예를 들어 ConN; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7. 바람직하게는, CfaN 또는 이의 기능적으로 동등한 변이체, 예를 들어 ConN; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4; 또는
- [0215] - Gp41N 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7. 바람직하게는, Gp41N 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4.
- [0216] 본원에서, 폴리뉴클레오타이드 2종을 참조하는 구현예들에서, 제2 폴리뉴클레오타이드 (또는 구체적으로, 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오타이드)가 본 발명의 임의의 분할 인테인 C-단편 또는 본 발명의 임의의 분할 인테인-C 데그론 단편일 수 있음에 유념한다. 구체적으로, 본 발명의 제2 폴리뉴클레오타이드 (또는 구체적으로, 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오타이드)는 서열번호 2, 4, 6, 9, 11, 13, 15, 18, 20, 22, 26, 67, 69, 71, 73, 75, 77 및 79로 이루어진 목록으로부터 임의의 것으로 선택할 수 있다. 더 바람직하게는, 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오타이드는 다음과 같은 구체적인 인테인; 및 데그론의 임의 조합을 코딩한다:
- [0217] - CfaC 또는 이의 기능적으로 동등한 변이체, 예를 들어 CfaCmut 또는 ConC; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PESTt, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7. 바람직하게는, CfaC 또는 이의 기능적으로 동등한 변이체, 예를 들어 CfaCmut; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, DD4, DD5, DD6 또는 DD7; or
- [0218] - Gp41C 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7. 바람직하게는, Gp41C 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4.
- [0219] 바람직한 구현예에서, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오타이드 및 제2 폴리뉴클레오타이드는, 폴리뉴클레오타이드 2종이 모두 각각의 단백질 복합체로 번역되어 본 발명에 따라 조합되었을 때 ABCA4 단백질의 N-말단 단편이 ABCA4 단백질의 C-말단 단편과 연결되어 전체 ABCA4 단백질이 생성되는 방식으로, ABCA4 단백질의 각각 N-말단 단편 및 C-말단 단편을 코딩한다. 더 바람직하게는, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오타이드는 ABCA4 단백질의 N-말단 단편의 1-1149, 1-1139, 1-1178 또는 1-1187번 위치 및 다음과 같은 인테인과 데그론의 임의의 특정한 조합을 코딩하고:
- [0220] - CfaN 또는 이의 기능적으로 동등한 변이체, 예를 들어 ConN; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7. 바람직하게는, CfaN 또는 이의 기능적으로 동등한 변이체, 예를 들어 ConN; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4; 및
- [0221] 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오타이드는 ABCA4 단백질의 C-말단 단편의 1150-2273, 1140-2273, 1179-2273 또는 1188-2273번 위치 및 다음과 같은 인테인과 데그론의 임의의 특정한 조합을 코딩한다:
- [0222] - CfaC 또는 이의 기능적으로 동등한 변이체, 예를 들어 CfaCmut 또는 ConC; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 및 DD7. 바람직하게는, CfaC 또는 이의 기능적

으로 동등한 변이체, 예를 들어 CfaCmut; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, DD4, DD5, DD6 또는 DD7.

- [0223] ABCA4 단백질의 N-말단 단편이 N-말단 단편의 C-말단 단편에 연결되어 전체 N-말단 단편 단백질을 형성하므로, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오티드가 1-1149번 위치를 코딩하는 경우, 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1150-2273번 위치를 코딩하는 것이, 중요하다. 나머지 위치들도 마찬가지이다.
- [0224] 전술한 바람직한 구현예를 이용해, 출발 N-말단 단편 및 C-말단 단편의 수준을 현저하게 낮추면서, 표적 단백질, 이들 예에서 ABCA4의 효율적인 재구성이 달성될 수 있는 것으로, 관찰할 수 있다. 구조체들 중 하나 또는 둘다에 하나 이상의 데그론 함유시, N-말단 단편 및/또는 C-말단 단편의 가속화된 분해로 인해 발현 감소가 달성된다. 당해 기술에 기반하면, 이러한 출발 물질의 수준 감소시 스플라이스 산물이 감소될 것으로 예상된다. 그럼에도 불구하고, 본 발명자들은 반대되는 거동을 관찰하였으며, 전장 스플라이싱된 산물의 수율에는 영향이 없으며, 일부 경우에는 심지어 개선될 수 있다. 아울러, 본 발명자들은, 본원에 기술된 방법을 이용하여 단백질을 재구성함으로써, ATPase 활성 분석으로 결정된 바와 같이 기능성 단백질을 생성할 수 있는 것으로 관찰하였다. 또한, 본 발명자들은 마우스 망막에서 표적 단백질이 생체내 효과적으로 재구성됨을 관찰하였다.
- [0225] 다른 바람직한 구현예에서, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오티드는 ABCA4 단백질의 N-말단 단편의 1-1095 또는 1-1185번 위치 및 다음과 같은 인테인과 데그론의 임의의 특정한 조합을 코딩하고:
- [0226] - Gp41N 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7. 바람직하게는, Gp41N 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4; 및
- [0227] 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오티드는 ABCA4 단백질의 C-말단 단편의 1096-2273 또는 1186-2273번 위치 및 다음과 같은 인테인과 데그론의 임의의 특정한 조합을 코딩한다:
- [0228] - Gp41C 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, PEST, DD1, DD2, DD3, M1, M2, SopE, SopE-1-78, SopE-15-78, SopE-15-50, L2, L6, L9, L10, L11, L12, L15, L16, M3, M4, M5, V12, DD4, DD5, DD6 또는 DD7. 바람직하게는, Gp41C 또는 이의 기능적으로 동등한 변이체; 및 다음 데그론들 중 임의의 데그론의 조합: CL1, Deg1, DD1, DD2, DD3, SopE, L2, L9, M4, V12, 또는 DD4.
- [0229] 앞서 경우에서와 같이, ABCA4 단백질의 N-말단 단편이 ABCA4 단백질의 C-말단 단편에 연결되어 전체 ABCA4 단백질을 생성하여야 하므로, 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오티드가 1-1095번 위치를 코딩하는 경우, 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오티드는 따라서 ABCA4 단백질의 C-말단 단편의 1096-2273번 위치를 코딩함에 유념하는 것이 중요하다. 나머지 위치들도 마찬가지이다.
- [0230] 상기한 (본 발명의 제1 및 제2 폴리뉴클레오티드, 또는 본 발명의 데그론을 코딩하는 제1 및 제2 폴리뉴클레오티드에 대한) 서열을 코딩하는 뉴클레오티드 서열들의 추가적인 바람직한 조합들은 다음과 같은 쌍들 중 임의의 쌍으로부터 선택된다: 1 및 2; 3 및 4; 5 및 6; 66 및 67; 68 및 69; 70 및 71; 72 및 73; 74 및 75; 76 및 77; 및 78 및 79.
- [0231] 용어 "본 발명의 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드"는 특히 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드를 망라하고, 후술한 바로부터 폴리뉴클레오티드에 대한 전술한 대안들을 이미 모두 망라하는 용어로서 본 발명의 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드만을 지칭하는 것임에 유념한다.
- [0232] 바람직하게는, 본 발명은 본 발명의 제1 폴리뉴클레오티드 및/또는 제2 폴리뉴클레오티드 또는 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오티드 및/또는 제2 폴리뉴클레오티드 및/또는 분리하여 또는 연결하여 3-조각 결합 시스템의 POI-인테인 단편들 각각을 포함하는, 조성물, 이하 본 발명의 제2 조성물을 지칭한다. 이러한 구체적인 맥락에서, 용어 "조성물"은 명시된 구성성분들을 함유한 생산물을 망라하는 것으로 의도된다. 조성물의 구성성분들은 함께 단일 제형으로, 또는 분리하여 서로 다른 제형으로 포장될 수 있다. 따라서, 일 구현예에서, 본 발명의 제1 폴리뉴클레오티드는 본 발명의 제1 폴리뉴클레오티드와 단일한 제형으로 함께 포장된다. 다른 구현예에서, 본 발명의 제1 폴리뉴클레오티드 및 본 발명의 제2 폴리뉴클레오티드는 분리하여 포장된다. 더 바람직하게는, 본 발명의 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드는, 양쪽 복합체가 본 발명에 따른 방법으로 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편과 연결되어 재구성할 전체 단백질이 생성되는 방식

으로, 재구성할 N-말단 단편 및 C-말단 단편을 각각 코딩한다.

- [0233] 상기에 언급된 본 발명의 폴리뉴클레오티드들 모두 자체 단리되거나 또는 적절한 숙주 세포에서 이러한 폴리뉴클레오티드의 전달 및/또는 증폭을 허용하는 벡터의 일부를 구성하는 것으로 확인할 수 있다. 따라서, 다른 측면에서, 본 발명은 상기 언급된 바와 같은 전술한 본 발명의 임의의 폴리뉴클레오티드를 포함하는 벡터에 관한 것이다.
- [0234] 이러한 폴리뉴클레오티드를 삽입하기에 적합한 벡터는 원핵생물에서의 발현 벡터, 예를 들어, pUC18, pUC19, Bluescript 및 이의 유도체, mpl8, mpl9, pBR322, pMB9, ColE1, pCR1, RP4, 파지 및 pSA3 및 pAT28와 같은 "서들" 벡터; 효모 발현 벡터, 예를 들어, 2 마이크로 플라스미드 타입의 벡터, 통합 플라스미드, YEP 벡터, 센트로미어 플라스미드 등; 곤충 세포에서의 발현 벡터, 예를 들어 pAC 및 pVL 시리즈 벡터; 식물 발현 벡터, 예를 들어, pIBI, pEarleyGate, pAVA, pCambia, pGSA, pGWB, pMDC, pMY, pORE 시리즈 등; 및 진핵생물 세포에서의 발현 벡터, 예를 들어 임의의 상업적으로 이용가능한 벡로바이러스 시스템을 이용하여 곤충 세포를 형질감염시키는데 적합한 벡로바이러스로부터 유래한 벡터들이다. 진핵생물 세포에 대한 벡터는 바람직하게는 바이러스 벡터 (아데노바이러스, 아데노 부속 바이러스 (AAV), 레트로바이러스와 같은 아데노바이러스 관련 바이러스, 특히 렌티바이러스뿐만 아니라 비-바이러스 벡터, 예를 들어 pSilencer 4.1-CMV (Ambion), pcDNA3, pcDNA3.1/hyg, pHCV/Zeo, pCR3.1, pEFI/His, pIND/GS, pRc/HCMV2, pSV40/Zeo2, pTRACER-HCMV, pUB6/V5-His, pVAX1, pZeoSV2, pCI, pSVL, PKSV-10, pBPV-1, pML2d 및 pTDT1을 포함한다. 바람직하게는, 벡터는 아데노 부속 바이러스 (AAV)이다. 바람직하게는, 벡터는 혈청형 1, 2, 3, 4, 5, 6, 7, 8 또는 9의 AAV이다. 이량체성 또는 자기-상보성 AAV 벡터 (scAAV) 역시 이러한 폴리뉴클레오티드를 삽입하기 위해 이용할 수 있다.
- [0235] 따라서, 본 발명은 바람직하게는 유전자 요법을 위한 벡터로서 AAV의 개발에 관한 것이다. 바람직하게는, 이들 벡터는 벡터의 DNA에서 rep 및 cap를 제거함으로써 이의 삽입 능력 (integrative capacity)이 제거된 것이다. 원하는 유전자 (폴리뉴클레오티드)는 유전자의 전사를 지시하기 위한 프로모터와 함께, 단일 가닥 벡터 DNA가 숙주 세포 DNA 중합효소 복합체에 의해 이중 가닥 DNA로 변환된 후 핵에서 콘카타머 형성을 보조하는 역위 말단 반복체 (ITR)들 사이에 삽입된다. AAV-기반의 유전자 요법 벡터는 숙주 세포 핵에서 에피솜 콘카타머를 형성한다. 이러한 콘카타머는 비-분열 세포에서는 숙주 세포의 생명에는 영향을 주지 않고 유지된다. 분열성 세포에서는, 에피솜 DNA가 숙주 세포 DNA와 함께 복제되지 않으므로, AAV DNA가 세포 분열을 통해 상실된다. 숙주 계능으로의 AAV DNA의 무작위 삽입이 검출되기도 하지만, 매우 낮은 빈도로 발생한다.
- [0236] 원하는 유전자 (폴리뉴클레오티드)는 구조체를 발현시킬 의도한 숙주 세포에서 기능하는 조절 인자와 조합될 것이다. 당해 기술 분야의 당업자는 적절한 숙주 세포, 예를 들어 포유류 또는 인간 숙주 세포에서 사용하기 위한 조절 인자들을 선택할 수 있다. 조절 인자로는 예를 들어 프로모터, 전사 종결 서열, 번역 종결 서열, 인핸서, 신호 펩타이드 및 폴리아데닐화 인자를 포함한다. 본 발명의 폴리뉴클레오티드는 프로모터 서열과 작동가능하게 연결될 수 있다. 본 발명에서 사용이 고려되는 프로모터로는 비-제한적으로 네이티브 유전자 프로모터, 사이토메갈로바이러스 (CMV) 프로모터 (KF853603.1, bp 149-735), 키메라 CMV/닭 β -액틴 프로모터 (CBA) 및 CBA (smCBA) 프로모터의 말단 절단형 (US8298818 및 Light-Driven Cone Arrestin Translocation in Cones of Postnatal Guanylate Cyclase-1 Knockout Mouse Retina Treated with AAVGC 1), 로돕신 프로모터 (NG 009115, bp 4205-5010), 인터포토수용체 레티노이드 결합 단백질 (IRBP) 프로모터 (NG_029718.1, bp 4777-5011), 난황형 황반 이영양증 2 (VMD2) 프로모터 (NG 009033.1, bp 4870-5470), PR-특이 인간 G 단백질-커플링된 수용체 키나제 1 (hGRK1; AY327580.1 bp1793-2087 또는 bp 1793-1991)(Haire et al. 2006; 미국 특허 8,298,818), 근위 뮤라인 로돕신 프로모터 (MOPS)를 포함한다. 그러나, 당해 기술 분야에서 공지된 임의의 적절한 프로모터들을 이용할 수 있다. 특정 구현예에서, 프로모터는 CMV 또는 hGRK1 프로모터이다. 일 구현예에서, 프로모터는 한 가지 조직 또는 일군의 조직들에서 선택적으로 활성이지만 다른 조직에서는 활성이 낮거나 또는 비활성인, 조직-특이적인 프로모터이다. 일 구현예에서, 프로모터는 광수용체-특이적인 프로모터이다. 추가적인 구현예에서, 프로모터는 원추 세포-특이 및/또는 간상 세포-특이 프로모터이다.
- [0237] 바람직한 프로모터는 CMV, GRK1, CBA 및 IRBP 프로모터이다. 또한 바람직한 프로모터는 다양한 프로모터로부터 유래한 조절 인자들을 조합한 하이브리드 프로모터이다 (예, CMV 프로모터, CBA 프로모터 및 Sv40 키메라 인트론으로부터 유래한 인핸서를 조합한 키메라 CBA 프로모터, 본원에서 CBA 하이브리드 프로모터로 지칭됨).
- [0238] 또한, AAV는 면역원성이 매우 낮아, 중화 항체 생성으로 제한되지 않는 것으로 보이며, 명백하게 정의된 세포독성 반응을 유도하지 않는다. 이런 특징은 휴지기 세포를 감염시키는 능력과 더불어 인간 유전자 요법을 위한 벡터로서 아데노바이러스에 비해 우수성을 가진다.

- [0239] AAV 게놈, 트랜스크립토 및 프로테옴: AAV 게놈은 약 4.7 킬로베이스 길이의 (+) 또는 (-) 센스 단일 가닥의 데옥시리보핵산 (ssDNA)으로 구성된다. 게놈은 DNA 가닥의 양쪽 말단에 역위 말단 반복체 (ITR)와 2개의 오픈 리딩 프레임 (ORF): red 및 cap로 구성되어 있다. 전자는 AAV 생활 주기에 필요한 Rep 단백질을 코딩하는 중첩되는 유전자 4개로 구성되고, 후자는 캡시드 단백질: VP1, VP2 및 VP3의 중첩되는 뉴클레오티드 서열로 구성되며, 이는 서로 상호작용하여 대칭적인 20면체 캡시드를 형성한다.
- [0240] 역위 말단 반복체 (ITR) 서열은 각각 염기 145개로 구성된다. 이는 대칭성으로 인해 그렇게 지칭되었으며, AAV 게놈을 효율적으로 증폭하는데 필수적인 것으로 입증되어 있다. 이들 서열의 다른 특성은 헤어핀을 형성하는 능력인데, 이는 제2 DNA 가닥의 프리마제-독립적인 합성을 허용하는 소위 자기-프라이밍에 기여한다. ITR은 또한 AAV DNA를 숙주 세포 게놈 (인간 염색체 19번)에 삽입하여 이로부터 구조하는데, 또한 완전히 조립된 데옥시리보뉴클레아제-내성 AAV 입자의 구축과 더불어 AAV DNA를 효율적으로 캡시드화하는데 필수적인 것으로 알려져 있다.
- [0241] 유전자 요법과 관련하여, ITR은 치료학적 유전자 옆에 cis로 필요한 유일한 서열인 것으로 보이며, 구조 (cap) 및 패키징 (rep) 유전자를 trans로 전달할 수 있다. 이러한 추측 하에, 리포터 또는 치료학적 유전자를 보유한 재조합 AAV (rAAV) 벡터를 효율적으로 생성하기 위한 많은 방법들이 확립되었다. 그러나, ITR은 효과적인 복제 및 캡슐화를 위해 cis로 필요한 유일한 인자가 아닌 것으로도 발표되어 있다. 몇몇 연구 그룹들에서는 rep 유전자의 코딩 서열 내에 서열 지정된 cis-작용하는 Rep-의존적인 인자 (CARE)를 발견하였다. CARE는 cis로 존재할 경우 복제 및 캡슐화를 높이는 것으로 입증되었다.
- [0242] 2006년까지 AAV 혈청형 11종이 이미 언급되어 있다. 공지된 혈청형들 모두 복수의 다양한 조직 타입들의 세포를 감염시킬 수 있다. 요법에 이의 사용이 중요할 가능성이 있는 항성 (tropism) 범위를 바꾸기 위해 조직 특이성은 AAV 벡터의 캡시드 혈청형 및 슈도타이핑에 의해 결정된다. 본 발명에서, AAV 혈청형 1, 2, 3, 4, 5, 6, 7, 8 또는 9의 ITR이 바람직하다.
- [0243] 혈청형 2 혈청형 2 (AAV2)가 지금까지 가장 널리 조사되었다. AAV2는 골격근, 뉴런, 혈관 평활근 세포 및 간세포에 대한 천연적인 항성을 나타낸다.
- [0244] 또한, 벡터와의 접촉 후 벡터가 통합된 세포를 식별할 수 있는 리포터 또는 마커 유전자를 포함할 수 있다.
- [0245] 본 발명의 맥락에서 유용한 리포터 유전자는 lacZ, 루시페라제, 티미딘 키나제, GFP 등을 포함한다. 본 발명의 맥락에서 유용한 마커 유전자로는 예를 들어 아미노글리코시드 G418에 대해 내성을 부여하는 네오마이신 내성 유전자; 히그로마이신에 대해 내성을 부여하는 히그로마이신 포스포트랜스퍼라제 유전자; 오르니틴 데카르복실라제 (2-(다이플루오로메틸)-DL-오르니틴 (DFMO)에 대해 내성을 부여하는 ODC 유전자; 메토틱렉세이트에 대해 내성을 부여하는 다이하이드로폴레이트리덕타제 유전자; 푸로마이에신에 대해 내성을 부여하는 푸로마이에신-N-아세틸 트랜스퍼라제 유전자; 제옥신에 대해 내성을 부여하는 ble 유전자; 9-베타-D-크실로푸라노스 아데닌에 대해 내성을 부여하는 아데노신 데아미나제 유전자; N-(포스포아세틸)-L-아스파르테이트의 존재시 세포 증식을 허용하는 시토신 데아미나제 유전자; 아미노프테린 (아미노프테린)의 존재시 세포 증식을 허용하는 티미딘 키나제; 크산틴의 존재 및 구아닌의 부재시 세포 증식을 허용하는 크산틴-구아닌 포스포리보실트랜스퍼라제 유전자; 트립토판 대신 인돌 존재시 세포 증식을 허용하는 *E. coli*의 trpB 유전자; 세포가 히스티딘 대신 히스티디놀을 이용하는 것을 허용하는 *E. coli*의 hisD 유전자 등이 있다. 선별 유전자는 진핵생물 세포에서 유전자를 발현시키는데 적합한 프로모터 (예를 들어, CMV 또는 SV40 프로모터), 최적화된 번역 개시 부위 (예를 들어, 소위 코자크 법칙에 따른 부위 또는 IRES), 폴리아데닐화 부위, 예를 들어 SV40 폴리아데닐화 또는 포스포글리세레이트 키나제 부위, 인트론, 예를 들어, β -글로불린 유전자 인트론을 부가적으로 포함할 수 있는 플라스미드에 통합된다. 대안적으로, 동일 벡터에 리포터 유전자 및 마커 유전자 둘다의 조합을 동시에 이용하는 것이 가능하다.
- [0246] 한편, 당해 기술 분야의 당업자에게 공지된 바와 같이, 벡터의 선정은 이후에 도입될 숙주 세포에 좌우될 것이다. 예를 들어, 폴리뉴클레오티드가 도입되는 벡터는 또한 효모 인공 염색체 (YAC), 박테리아 인공 염색체 (BAC) 또는 PI-유래 인공 염색체 (PAC)일 수 있다. YAC, BAC 및 PAC의 특징들은 당해 기술 분야의 당업자에게 알려져 있다. 이러한 유형의 벡터에 대한 상세 정보는, 예를 들어 Giraldo and Montoliu (Giraldo, P. & Montoliu L., 2001 Size matters: use of YACs, BACs and PACs in transgenic animals, Transgenic Research 10(2): 83-110)에 제공되어 있다. 본 발명의 벡터는 당해 기술 분야의 당업자들에게 공지된 통례적인 방법에 의해 수득할 수 있다 (Sambrook J. et al., 2000 "Molecular cloning, a Laboratory Manual", 3rd ed., Cold Spring Harbor Laboratory Press, N.Y. Vol 1-3).

- [0247] 본 발명의 폴리뉴클레오티드는 숙주 세포에 생체내에서 네이키드 DNA 플라스미드로 도입할 수 있지만, 또한 비-제한적으로, 형질감염, 전기천공 (예, 경피 전기천공), 미세주입, 형질전이, 세포 융합, DEAE 텍스트란, 칼슘 포스페이트 석출, 유전자 총 사용 또는 DNA 벡터 트랜스포터 사용 등의 당해 기술 분야에 공지된 방법에 의해 벡터를 이용해 도입할 수 있다. 네이키드 DNA를 제형화하여 포유류 근육 조직에 투여하는 방법들 역시 알려져 있다. Feigner P, et al., US 5,580,859, 및 US 5,589,466을 참조한다. 양이온성 올리고펩타이드, DNA 결합 단백질 유래 펩타이드 또는 양이온성 폴리머와 같이, 생체내 핵산 형질감염을 용이하게 하기 위한 기타 분자들 역시 유용하다. Bazile D, et al., WO 1995021931, 및 Byk G, et al., WO 1996025508을 참조한다.
- [0248] 폴리뉴클레오티드를 숙주 세포에 도입하기 위해 이용할 수 있는 다른 잘 알려진 방법은 입자 총 (particle bombardment) (바이올리스틱 형질전환으로 알려져 있음)이다. 바이올리스틱 형질전환은 여러가지 방식들 중 한가지 방식으로 일반적으로 달성된다. 일반적인 한가지 방법은 세포에서 불활성 또는 생물학적 활성 입자를 추진시키는 것을 수반한다. Sanford J, et al., US 4,945,050, US 5,036,006 및 US 5,100,792를 참조한다.
- [0249] 대안적으로, 벡터는 리포펙션에 의해 생체내 도입할 수 있다. 양이온성 지질의 사용은 음으로 하전된 핵산의 캡슐화를 촉진할 수 있으며, 또한 음으로 하전된 세포 막과의 융합을 촉진할 수 있다. Feigner P, Ringold G, Science 1989; 337:387-388을 참조한다. 핵산 전달에 특히 유용한 지질 화합물 및 조성물들이 알려져 있다. Feigner P, et al., US 5,459,127, Behr J, et al., WO1995018863, 및 Byk G, WO1996017823을 참조한다.
- [0250] 마지막으로, 특히 바람직하게는, 벡터는 비-제한적으로 아데노바이러스 벡터, 아데노-부속 바이러스 (AAV) 벡터, 슈도타입 AAV 벡터, 헤르페스 바이러스 벡터, 레트로바이러스 벡터, 렌티바이러스 벡터, 바쿨로바이러스 벡터 등의 바이러스 전달 시스템에 의해 생체내 도입할 수 있다. 슈도타입 AAV 벡터는 한가지 AAV 혈청형의 게놈이 2번째 AAV 혈청형의 캡시드 안에 포함하는 것이며; 예를 들어, AAV2/8 벡터는 AAV8 캡시드와 AAV 2 게놈을 포함한다 (Auricchio et al. (2001) Hum. Mol. Genet. 10(26):3075-81). 이러한 벡터는 또한 키메라 벡터로 알려져 있다. 그외 전달 시스템에 대한 예로는 생체의 전달 시스템이 있으며, 이는 비-제한적으로 DNA 형질감염 방법, 예를 들어 전기천공, DNA 바이올리스틱, 지질-매개 형질감염, 컴팩트 DNA-매개 형질감염을 포함한다.
- [0251] AAV 벡터의 구축은 당해 기술 분야의 당업자에게 공지된 기법을 이용해 아래 절차로 수행할 수 있다. 아데노-부속 바이러스 벡터 구축 및 요법에서의 사용에 대한 이론 및 실무는 몇가지 과학 및 특허 간행물들에 예시되어 있다 (하기 문헌들은 원용에 의해 본원에 포함됨: Flotte TR. Adeno-associated virus-based gene therapy for inherited disorders. Pediatr Res. 2005 Dec;58(6):1143-7; Goncalves MA. Adeno-associated virus: from defective virus to effective vector, Virol J. 2005 May 6;2:43; Surace EM, Auricchio A. Adeno-associated viral vectors for retinal gene transfer. Prog Retin Eye Res. 2003 Nov;22(6):705-19; Mandel RJ, Manfredsson FP, Foust KD, Rising A, Reimsnider S, Nash K, Burger C. Recombinant adeno-associated viral vectors as therapeutic agents to treat neurological disorders. Mol Ther. 2006 Mar;13(3):463-83).
- [0252] AAV 벡터를 함유한 약학적 조성물에 적합한 투여 형태로는 비-제한적으로 주사가능한 용액 또는 현탁액, 눈 로션제 및 안 연고 등이 있다. 따라서, 다른 측면에서, 본 발명은 본 발명의 폴리뉴클레오티드 또는 벡터를 포함하는 숙주 세포에 관한 것이다. 세포는 당해 기술 분야의 당업자에게 공지된 통례적인 방법으로 획득할 수 있다 (예를 들어, 상기 인용된 Sambrook et al., 참조).
- [0253] 본원에서, 용어 "숙주 세포"는 본 발명에 따른 폴리뉴클레오티드 또는 벡터와 같은 본 발명의 핵산이 도입되어, 본 발명의 분할 인테인 N-단편 또는 분할 인테인 N-단편을 포함하는 융합 단백질을 발현할 수 있는 세포를 지칭한다. 용어 "숙주 세포" 및 "재조합 숙주 세포"는 본원에서 상호 호환적으로 사용된다. 이들 용어는 특정 대상 세포뿐 아니라 이러한 세포의 자손 또는 잠재적인 자손을 지칭하는 것으로 이해하여야 한다. 돌연변이 또는 환경적인 영향으로 인해 연속 세대에서 일정한 변형이 발생할 수 있으므로, 상기한 자손은 실제 부모 세포와 동일하지 않을 수 있지만, 여전히 본 발명의 용어 범위에 포함된다. 이 용어는 이중적인 DNA의 도입에 의해 변형될 수 있는 임의의 증식가능한 세포를 포괄한다. 바람직하게는, 숙주 세포는 본 발명의 폴리뉴클레오티드가 안정적으로 발현되고, 번역 후 수정되고, 적절한 세포내 영역으로 국지화되고, 적절한 전사 기전을 참여시킬 수 있는 것이다. 적절한 숙주 세포의 선정은 또한 검출 신호의 선정에 의해 영향을 받을 것이다. 예를 들어, 전술한 바와 같이, 리포터 구조체는 전사 조절 단백질에 반응하여 유전자 전사 저해 또는 활성화시 선별가능한 또는 스크리닝 가능한 형질을 제공할 수 있으며; 최적의 선별 또는 스크리닝을 달성하기 위해, 숙주 세포 표현형이 고려될 것이다. 본 발명의 숙주 세포는 원핵생물 세포 및 진핵생물 세포를 포함한다. 원핵생물은 그람 음성 또는 그람 양성 유기체, 예를 들어 *E. coli* 또는 바실러스를 포함한다. 원핵생물 세포도, 바람직하게는, 본 발명의 폴리뉴클레오티드 또는 벡터를 포함하는 전사 조절 서열을 증폭시키기 위해, 이용될 것으로 이해된다. 형질전환에

적합한 원핵생물 숙주 세포로는 예를 들어 *E. coli*, 바실러스 썩틸리스, 살모넬라 티피무리움 및 슈도모나스, 스트렙토마이세스 및 스타필로코커스 속에 속하는 다양한 기타 종 등이 있다. 진핵생물 세포로는 비-제한적으로 효모 세포, 식물 세포, 진균 세포, 곤충 세포 (예를 들어, 벡로로바이러스), 포유류 세포, 및 기생 유기체 세포, 예를 들어, 트리파노솜 (trypanosome) 등이 있다. 본원에서, 효모는 엄격한 분류학적 의미에서 효모, 즉 단세포 유기체뿐만 아니라 사상성 진균의 효모-유사 다세포 진균을 포함한다. 예시적인 종으로는 클루베리 락티스 (*Kluyverei lactis*), 시조사카로마이세스 폼베 (*Schizosaccharomyces pombe*) 및 우스틸라코 마이디스 (*Ustilago maydis*)가 있으며, 사카로마이세스 세레비시아 (*Saccharomyces cerevisiae*)가 바람직하다. 본 발명의 실시예에 이용할 수 있는 기타 효모는 뉴로스포라 크라사 (*Neurospora crassa*), 아스퍼질러스 나이거 (*Aspergillus niger*), 아스퍼질러스 니들란스 (*Aspergillus nidulans*), 피키아 파스토리스 (*Pichia pastoris*), 칸디다 트로피칼리스 (*Candida tropicalis*) 및 한센울라 폴리모르파 (*Hansenula polymorpha*)가 있다. 포유류 숙주 세포 배양 시스템은 COS 세포, L 세포, 3T3 세포, 중국 햄스터 난소 (CHO) 세포, 배아 줄기 세포와 같은 확립된 세포주들을 포함하며, BHK, HeK 또는 HeLa 세포가 바람직하다. 바람직하게는 진핵생물 세포가 재조합 유전자 발현에 이용된다.

[0254] 바람직한 구현예에서, 본 발명의 제2 조성물은 요법, 특히 표 1에 동정된 임의 질환에서 조성물에 대해 표기된 유전자 타입에 따라 이용하기 위한 것이다 (질환과 유전자 간의 연관성은 표 1에 명확하게 언급되어 있으며, 따라서 당업자라면 올바른 조합을 식별할 수 있음이 자명하다).

[0255] **세포에서 대상 단백질을 코딩하는 유전자의 발현 방법**

[0256] 다른 측면에서, 본 발명은 하기를 포함하는, 시험관내 또는 생체내에서 세포에서 대상 유전자를 발현시키는 방법, 이하 대상 유전자를 발현시키는 제1 방법에 관한 것이다:

[0257] (i) 세포를

[0258] (a) 본 발명의 제1 폴리뉴클레오타이드 또는 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오타이드, 및

[0259] (b) 본 발명의 제2 폴리뉴클레오타이드 또는 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오타이드와 접촉시키는 단계,

[0260] (ii) 제1 융합 단백질 및 제2 융합 단백질이 생산되도록 제1 폴리뉴클레오타이드 및 제2 폴리뉴클레오타이드를 발현시키는 단계; 및

[0261] (iii) 제1 단백질과 제2 단백질을 접촉시켜, 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간 산물이 생성되고, 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결되는 단계.

[0262] 특정 구현예에서, 제1 폴리뉴클레오타이드는 데그론을 코딩하는 제1 폴리뉴클레오타이드이고, 제2 폴리뉴클레오타이드는 데그론을 코딩하는 제2 폴리뉴클레오타이드이고, 바람직하게는, 대상 단백질은 > 25 KDa, > 50 KDa 또는 > 100 KDa이며, 따라서 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결되면 전체 단백질이 만들어진다.

[0263] 세포와 본 발명의 제1 폴리뉴클레오타이드 및/또는 제2 폴리뉴클레오타이드 중 임의의 것의 접촉은 시험관내 또는 생체내에서 대상 폴리뉴클레오타이드를 세포에 도입할 수 있는 임의의 적절한 수단, 예를 들어 형질감염, 전기천공, 미세주입법, 형질전이, 리포펙션, 세포 융합, DEAE 텍스트란, 칼슘 포스페이트 석출, 유전자총 사용 또는 DNA 벡터 트랜스포터 사용에 의해 이루어질 수 있다. 바람직하게는, 벡터는 아데노-부속 바이러스 (AAV)이다.

[0264] 바람직한 구현예에서, 본 발명은 하기를 포함하는 세포에서 대상 단백질을 코딩하는 유전자를 발현시키는 방법, 이하 대상 유전자를 발현시키기 위한 제1 방법에 관한 것으로:

[0265] (i) 세포를,

[0266] (a) 본 발명의 제1 폴리뉴클레오타이드 또는 본 발명의 데그론을 코딩하는 제1 폴리뉴클레오타이드를 포함하는 제1 AAV, 및

[0267] (b) 본 발명의 제2 폴리뉴클레오타이드 또는 본 발명의 데그론을 코딩하는 제2 폴리뉴클레오타이드를 포함하는 제2 AAV와 접촉시키는 단계,

[0268] (ii) 제1 폴리뉴클레오타이드와 제2 폴리뉴클레오타이드를 발현시켜, 제1 융합 단백질 및 제2 융합 단백질을 생산

하는 단계, 및

- [0269] (iii) 제1 융합 단백질과 제2 융합 단백질을 접촉시켜, 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간산물이 생성되고, 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결되는 단계,
- [0270] 바람직하게는, 제1 폴리뉴클레오티드는 데그론을 코딩하는 제1 폴리뉴클레오티드이고, 제2 폴리뉴클레오티드는 데그론을 코딩하는 제2 폴리뉴클레오티드이며, 더 바람직하게는, 대상 단백질은 > 25 KDa, > 50 KDa 또는 > 100 KDa이므로, 대상 제1 폴리펩타이드의 C-말단이 대상 제2 폴리펩타이드의 N-말단에 공유 연결시 전체 단백질이 생성된다.
- [0271] 이러한 의미에서, 실시예에 예시된 바와 같이, 데그론의 존재는 출발 물질 (즉, 본 발명의 분할 인테인 N- 및/또는 C-말단 단편 데그론)의 빠른 분해를 야기할 것이다. 선택되는 데그론은, 정상 상태에서, 출발 물질의 양이 데그론의 부재시 관찰되는 것보다 감소되는 방식으로, 단백질 스플라이싱과 호환되는 분해 카이네틱스를 가진다. 그러나, 스플라이싱 산물의 수준은 데그론 부재시 관찰되는 수준과 비교해 유지되거나 또는 증가된다.
- [0272] 본 발명의 대상 유전자를 발현시키기 위한 방법에서, 세포를 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드와 동시에 접촉시키거나, 또는 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드와 임의 순서로 순차적으로 접촉시키는 것이 고려되며, 즉, 세포를 먼저 제1 폴리뉴클레오티드와 접촉시킨 다음 제2 폴리뉴클레오티드와 접촉시키거나, 또는 먼저 제2 폴리뉴클레오티드와 접촉시킨 다음 제1 폴리뉴클레오티드와 접촉시킬 수 있다. 이러한 폴리뉴클레오티드를 코딩하는 벡터에도 동일하게 적용되어야 한다.
- [0273] 숙주 세포로서 앞서 정의된 임의의 세포를 이러한 방법에 이용할 수 있다.
- [0274] ***
- [0275] 본 발명은 하기 실시예를 들어 기술될 것이며, 실시예는 단순 예시로서 간주되며 본 발명의 범위를 제한하는 것으로 간주되어서는 안된다.
- [0276] **실시예**
- [0277] 아래 표에서, 본원에 이미 열거된 서열과 더불어 첨부물에 의해 형성되는 서열을 나타내기 위한 명명 방식을 확립한다. 이들 서열은 하기 표에서 N-말단에서 C-말단 방향으로, 이를 구성하는 서열번호들로 식별되는 서열을 나열함으로써 표시된다. 예를 들어, 서열번호 106의 1-1149번 아미노산들이 서열번호 27 (CfaN)에 직접 연결되고 그 다음으로 서열번호 103 (3FT)이 직접 연결된 서열의 경우 사용되는 명명 방식은 서열번호 106 (1번에서 1149번)-서열번호 27-서열번호 103일 것이다. 이러한 구체적인 예는 N-말단으로부터 시작해, 서열번호 106의 1번 아미노산에서부터 1149번까지, 그 다음으로 펩타이드 결합으로 직접 연결된 서열번호 27이, 그 다음으로 펩타이드 결합으로 또한 연결된 서열번호 103에 대응되는 폴리펩타이드가 바로 위치하는, 서열로서 해석하여야 한다. 그래서, 이 서열은 N-말단에서 C-말단 방향으로 하기 구성을 가질 것이다:
- [0278] [서열번호 106 (잔기 1번부터 1149번)] - [서열번호 27] - [서열번호 103]
- [0279] 괄호 "[]" 안 서열은 해당 서열의 모든 아미노산을 나타내고, "-"는 괄호 안 서열 2개를 연결하는 펩타이드 결합을 의미한다. 특정 서열의 잔기 x-y는 x위치에서 y위치까지의 양쪽 말단을 포함한 모든 아미노산 잔기를 의미한다. 예를 들어, 서열번호 106 (잔기 1-1149)은 양쪽 말단을 포함하여 서열번호 106의 1번 잔기부터 1149번 잔기를 포함하는 서열을 의미한다.
- [0280] 표 5. 실시예 전체에 사용되는 서열 목록:

표 5

구조체	*서열	서열번호
ABCA4	M ¹ GFVRQIQLLLWKNWTLRKRQKIRFVVELVWPLSLFLVLIWLRNANPLYS HHECHFPNKAMP SAGMLPWLQGIFCNVNNPCFQSPTPGESPGIVSNYNNNS ILARVYRDFQELLMNAPE SQHLGRIW TELHILSQFMDTLRTHPERIAGRIGRI RDILKDEETLTFLIKNIGLSD SVVYLLINSQVRPEQFAHGVPDLALKDIACS EALLERFIIFSQRRGAKTVRYALCSLSQGTLQWIEDTLYANVDFFKLFRVLP TLLD SRSQGINLRSWGGILSDMSPRIQEFIHRSMDQLLWVTRPLMQNGGP ETFTKLMGILSDLLCGYPEGGGSRVLSFNWYEDNNYKAFLGIDSTRKDPIY SYDRRTTSFCNALIQSLESNPLTKIAWRAAKPLLMGKILYTPDSPAARRILK NANSTFEELEHVRKLVKAWEEVGPQIWYFFDNSTQNMNIRDTLGNPTVK DFLNRQLGEEGITAEAILNFLYKGPRESQADDMANFDRDIFNITDRTLRL VNQYLECLVLDKFESYNDETQLTQRALSLLEENMFAGVVFDPDMPWTS SLPPHVKYKIRMDIDVVEKTNKIKDRYWD SGPRADPVEDFRYIWGGFAYLQ DMVEQGITRSQVQAEAPVGIYLQQMPYPCFVDDSFMIILNRCFPIFMVLAWI YSVSMTVKSVILEKELRLKETLKNQGVSNVWCTWFLDSFSIMSMSIFLL TIFIMHGRILHYSDPFILFLLLAFSTATIMLCFLLSTFFSKASLAAACSGVIY FTLYLPHILCFAWQDRMTAELKKAVSLLSPVAFGFGTEYLVRFEQGLGL QWSNIGNSPTEGDEFSFLLSMQMMLLDAAVYGLLAWYLDQVFPGDYGT LPWYFLLQESYWLGGEGCSTREERALEKTEPLTEETEDPEHPEGIHD SFFE	106

[0281]

	REHPGWVPGVCVKNLVKIFPCGRPAVDRLNITFYENQITAFLGHNAGK TTTSLILTGLLPPTSGTVLVGGRDIETSLDAVRQSLGMCQHNILFHHLTV AEHMLFYAQLKGKSQEEAQLEMEAMLED TGLHHKRNEEAQDLSGGMQRK LSVAIAFVGDAKVVLDEPTSGVDPYS ¹⁰⁹ RRSIWDLKLYRSGRTIIMSTHH MDEADLLGDRIAIQAQRLYC ¹¹⁴ SGTPLFLKNC ¹¹⁵ FGTGLYLTIVRKMKN QSQRKGSEGT C ¹¹⁷ SC ¹¹⁷ SSKGF S ¹¹⁸ TTC ¹¹⁸ PAHVDDLTPQVLDGDVNE LMDVVLHHVPEAKLVEICIGQELIFLLPNKNFKHRAVYSLFRELEETLADLG LSSFGISDTPLEEFLKVTEDSDSGPLFAGGAQQKRENVNPRHPCLGPREK AGQTPQDSNV CSPGAPAAHPEGQPPPEPECPGPQLNTGTQLVLQHVQALL VKRFQHTIRSHKDFLAQIVLPATFVFLALMLSIVIPPFGEYPALTLHPWIYG QQYTFFSMDEPGSEQFTVLADVLLNKPGFGNRCLKEGWLPEYPCGNSTP WKTPSVSPNITQLFQKQKW TQVNPSPCRCSTREKLTMLPECPEGAGGLP PPQRTQRSTEILQDLTDRNISDFLVKTYPALIRSSLKSKFWVNEQRYGGISIG GKLPVVPITGEALVGFLSDLGRIMNVSGGPITREASKEIPDFLKHLETEDNI KVWFNNKGWHALVSFLNVAHNAILRASLPKDRSPEEYGITVISQPLNLTKE QLSEITVLTTSDAVVAICVIFSMSPVPA SFVLYLIQERVNKS KHLQFISGV PTTYWVTNFLWDIMNYSV SAGLVVGIFIGFQKKAYTSPENLPALVALLLY GWAVIPMMYPASFLFDVPSTAYVALSCANLFIGINSSAITFILELPENNRTL LRFNAVLRKLLIVPHFCLGRGLIDLALSQAVTDVYARFGEEHSANPFHWD LIGKNLFAMVVEGVVYFLLTLLVQRHFFLSQWIAEPTKEPIVDEDDDAEE RQRITGGNKTDILRLHELTKIYPGTSSPAVDRLCVGVRPGE CFGLLGVN GA GKTTTFKMLTGDTT VTSGDATVAGK SILTNISEVHQNMGYCQFDAQDEL LTGREHLYLYARLRGVPABEIEK VANW SIKSLGLTVYADCLAGTYSGGNKR KLSTAIALIGCPPLVLLDEPTTGMDPQARRMLWNVIVSIIREGRAVVL TSHS MEECEALCTRLAIMVKGAFCRMGTIQHLKSKFGDGYIVTMKIKSPKDDLLP DLNPVEQFFQGNFP GSVQRRERHYNMLQFQVSSSSLARIFQLLLSHKDLLIE EYSVTQTTLQVFNFAKQQTESHDLPLHPRAAGASRQAQD ²²⁷	
3FT	DYKDHDGDYKDHDIDYKDDDDK	103

[0282]

구조체	*서열	서열번호
ABCA4-1150-Cfa ^N -3FT	서열번호 106 (1-1149)-서열번호 27-서열번호 103	1
ABCA4-1150-Cfa ^C -3FT	서열번호 28-서열번호 106 (1150-2273)-서열번호 103	2
ABCA4-1140-Cfa ^N -3FT	서열번호 106 (1-1139)-서열번호 27-서열번호 103	3
ABCA4-1140-Cfa ^{Cmut} -3FT	서열번호 29-서열번호 106 (1140-2273)-서열번호 103	4
ABCA4-1188-Cfa ^N -3FT	서열번호 106 (1-1187)-서열번호 27-서열번호 103	5
ABCA4-1188-Cfa ^{Cmut} -3FT	서열번호 29-서열번호 106 (1188-2273)-서열번호 103	6
ABCA4-3FT	서열번호 106-서열번호 103	7
ABCA4-1150-N ^{pu} ^N -3FT	서열번호 106 (1-1149)-서열번호 32-서열번호 103	8
ABCA4-1150-N ^{pu} ^C -3FT	서열번호 33-서열번호 106 (1150-2273)-서열번호 103	9

ABCA4-1140-N ^{pu} ^N -3FT	서열번호 106 (1-1139)-서열번호 32-서열번호 103	10
ABCA4-1140-N ^{pu} ^C -3FT	서열번호 33-서열번호 (1140-2273)-서열번호 103	11
ABCA4-1188-N ^{pu} ^N -3FT	서열번호 106 (1-1187)-서열번호 32- 서열번호 103	12
ABCA4-1188-N ^{pu} ^C -3FT	서열번호 33-서열번호 106 (1188-2273)-서열번호 103	13
EGFP-71- Cfa ^N	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD ATYGKLTCLKFICTTGKLPVPWPTLVTTLTLYGVQCLSY DTEILTVEYGFLLPIGKIVEERIECTVYTVDKNGFVYTQP IAQWHNRGEQEVFEYCLEDGSIIRATKDHKFMTTDGG MLPIDEIFERGLDLKQVDGLPGHHHHHHG	14
EGFP-71- Cfa ^C	MVKIISRKSLGTQNVYDIGVEKDHNFLKNGLVASNCF SRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT RAEVKFEGLTLVNRIELKGIDFKEDGNILGHKLEYNN SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQ QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHML LEFVTAAGITLGMDELYKGHHHHHHG	15
EGFP	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD ATYGKLTCLKFICTTGKLPVPWPTLVTTLTLYGVQCFSR YPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTR AEVKFEGLTLVNRIELKGIDFKEDGNILGHKLEYNN SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQ NTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHML EFVTAAGITLGMDELYKGHHHHHHG	16
EGFP-71- N ^{pu} ^N	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGD ATYGKLTCLKFICTTGKLPVPWPTLVTTLTLYGVQCLSY ETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQP VAQWHDNRGEQEVFEYCLEDGSLIRATKDHKFM TVD GQMLPIDEIFERELDLMRVDNLPNGHHHHHHG	17
EGFP-71- N ^{pu} ^C	MIKIATRKYLKQNVYDIGVERDHNFLKNGFIASNCF SRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT RAEVKFEGLTLVNRIELKGIDFKEDGNILGHKLEYNN SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQ QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHML LEFVTAAGITLGMDELYKGHHHHHHG	18
ABCA4-1150-Cfa ^N -3FT- SopE	서열번호 106(1-1149)-서열번호 27-서열번호 103- 서열번호 48	19
ABCA4-1150-Cfa ^C -3FT- SopE	서열번호 28-서열번호 106(1150-2273)-서열번호 103-서열번호 48	20
ABCA4-1140-Cfa ^N -3FT- SopE	서열번호 106(1-1139)-서열번호 27-서열번호 103- 서열번호 48	21
ABCA4-1140-Cfa ^{Cmut} -	서열번호 29-서열번호 106(1140-2273)-서열번호	22

3FT-SopE	103-서열번호 48	
EGFP-71-Cfa^N-SopE	서열번호 14-서열번호 48	23
EGFP-71-Cfa^C-SopE	MTKITLSPQNFRIQKQETTLKEKSTENSLAKSILAV KNHFIELRSKLSERFISHKNTESSATHFHRGSASEGRA VLTKNVVKDFMLQTLNDIDIRGSAVKIISRKSLGTQNV YDIGVEKDHNFLKNGLVASNCFSRYPDHMKQHDFFK SAMPEGYVQERTIFFKDDGNYKTRAEVKFEGLTVNR IELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNG IKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDN HYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDE LYKGHHHHHHG	24
EGFP-71-Cfa^N-DD1	서열번호 14-서열번호 43	25
EGFP-71-Cfa^C-DD1	MRISFGSPPPMAGGVKIIIRKSLGTQNVYDIGVEKDHN FLLKNGLVASNCFSRYPDHMKQHDFFKSAMPEGYVQ ERTIFFKDDGNYKTRAEVKFEGLTVNRIELKGIDFKED GNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRH NIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSAL SKDPNEKRDHMLLEFVTAAGITLGMDELYKGHHHH HHG	26

[0285]

구조체	*서열	서열번호
ABCA4 ^N -1150	서열번호 106 (1-1149)	60
ABCA4 ^C -1150	서열번호 106 (1150-2273)	61
ABCA4 ^N -1140	서열번호 106 (1-1139)	62
ABCA4 ^C -1140	서열번호 106 (1140-2273)	63
ABCA4 ^N -1188	서열번호 106 (1-1187)	64
ABCA4 ^C -1188	서열번호 106 (1188-2273)	65
ABCA4-1150-Cfa ^N	서열번호 60-서열번호 27	66
ABCA4-1150-Cfa ^C	서열번호 28-서열번호 61	67
ABCA4-1140-Cfa ^N	서열번호 62-서열번호 27	68
ABCA4-1140-Cfa ^{Cmut}	서열번호 29-서열번호 63	69
ABCA4-1188-Cfa ^N	서열번호 64-서열번호 27	70
ABCA4-1188-Cfa ^{Cmut}	서열번호 29-서열번호 65	71
ABCA4-1150-Cfa ^N -SopE	서열번호 66-서열번호 103-서열번호 48	72
ABCA4-1150-Cfa ^C -SopE	서열번호 48-서열번호 67-서열번호 103	73
ABCA4-1140-Cfa ^N -SopE	서열번호 68-서열번호 103-서열번호 48	74
ABCA4-1140-Cfa ^{Cmut} -SopE	서열번호 48-서열번호 69-서열번호 103	75
ABCA4-1150-Cfa ^N -DD1	서열번호 66-서열번호 43, X = Gly.	76
ABCA4-1150-Cfa ^C -DD1	서열번호 67-서열번호 43, X = Gly.	77
ABCA4-1140-Cfa ^N -DD1	서열번호 68-서열번호 43, X = Gly.	78
ABCA4-1140-Cfa ^{Cmut} -DD1	서열번호 69-서열번호 43, X = Gly.	79

[0286]

[0287]

재료 및 방법

[0288]

재료:

[0289]

올리고뉴클레오타이드는 Eurofins genomics에서 구입하였다. 합성 유전자는 GENEWIZ에서 구입하였다. 클로닝용 Pfu Ultra 융합 증합효소 및 모든 제한 효소는 ThermoFisher Scientific에서 구입하였다. 클로닝을 위한 고-검 피턴시 세포는 XL10-Gold 화학적 컴피턴트 *E. coli*에서 구축하였다. HEK293T 세포는 ATCC에서 입수하였다. DNA 정제 키트는 Qiagen에서 구입하였다. 모든 플라스미드는 MacroGen사에서 서열분석하였다. LB (Luria Bertani) 배지 및 모든 완충제 염은 ThermoFisher Scientific에서 구입하였다. 코마시 브릴리언트 블루, MG-132 프로테아좀 저해제, 페닐메탄 설펜일 플루오라이드, 요오도아세트아미드, NH₄CO₃, DTT, 포름산, 소 태아 혈청 및 대두 아소랙틴은 Sigma-Aldrich에서 구입하였다. 아세토니트릴 (ACN)은 Carlo-Erba에서 구입하였다. EDTA-비함유 완전 프로테아제 저해제는 Roche에서 구입하였다. 리포펙타민 2000 형질감염 시약, DMEM 고 글루코스 GlutaMAX 첨

가물, RPMI 1640 배지 GlutaMAX 첨가물, RIPA 세포용해 및 추출 완충제, BCA 단백질 분석 키트, MES-SDS 전개 완충제, 사전-염색된 단백질 래더 및 SDS-PAGE (비스-트리스 및 트리스-아세테이트 겔)는 ThermoFisher Scientific에서 구입하였다. 1차 항-6xHis 태그 마우스 단일클론 항체, 항-FLAG 태그 마우스 단일클론 항체, 항-ABCA4 토끼 다클론 항체 및 항-튜불린 토끼 다클론 항체는 Invitrogen에서 구입하였다. 2차 염소 항-마우스 IgG (H+L) 고 교차-흡착 항체 alexa fluor plus 680 및 염소 항-토끼 IgG (H+L) 2차 항체 dylight 800 4X PEG는 Invitrogen에서 구입하였다. 도데실 말토시드 (D310) 및 콜레스테릴 헤미숙시네이트 (CH210) 용액은 Anatrace에서 구입하였다. 트립신은 Promega에서 구입하였다.

[0290] 장치:

[0291] 전기분무 이온화 질량 분광측정 분석 (ESI-MS)은 LTQ-Orbitrap Velos (Thermo Scientific) 질량 분광기가 연결된 Acquity liquid chromatographer (Waters)에서 수행하였다. 겔 및 웨스턴-블롯은 LI-COR Odyssey Infrared Imager으로 촬영하였다. 세포 용해는 SFX550 Branson sonifier를 이용해 수행하였다. FACS 측정은 Gallios Beckman Coulter에서 수행하였다.

[0292] 재조합 DNA 클로닝

[0293] 구조체 ABCA4-1150-CfaN (서열번호 1) 및 ABCA4-1150-CfaC (서열번호 2)를 제조하기 위한 합성 유전자를 구입하여, pEGFP-N1 발현 벡터에 KpnI 및 NotI 제한 효소를 이용해 도입하였다. NpuN 및 NpuC에 대한 합성 유전자를 구입하고, ABCA4-1150-CfaN 및 ABCA4-1150-CfaC에 제한효소 프리 클로닝에 의해 도입해, 구조체 ABCA4-1150-NpuN (서열번호 8) 및 ABCA4-1150-NpuC (서열번호 9)를 수득하였다.

[0294] 구조체 ABCA4-3FT (서열번호 7), ABCA4-1140-CfaN (서열번호 3), ABCA4-1140-CfaCmut (서열번호 4), ABCA4-1188-CfaN (서열번호 5), ABCA4-1188-CfaCmut (서열번호 6) ABCA4-1179-CfaN (서열번호 93) 및 ABCA4-1179-CfaCmut (서열번호 94)는 제한 효소 프리 클로닝에 의해 준비하였다. CfaCmut에서 돌연변이를 Pfu Ultra II HF 중합효소를 이용한 역위 PCR을 이용해 도입하였다. ABCA4-1140-NpuN (서열번호 10), ABCA4-1140-NpuC (서열번호 11), ABCA4-1188-NpuN (서열번호 12) 및 ABCA4-1188-NpuC (서열번호 13)는 제한 효소 프리 클로닝에 의해 준비하였다.

[0295] 구조체 EGFP-71-CfaN (서열번호 14), EGFP-71-CfaC (서열번호 15), EGFP (서열번호 16), EGFP-71-NpuN (서열번호 17) 및 EGFP-71-NpuC (서열번호 18)는 제한 효소 프리 클로닝에 의해 준비하였다.

[0296] SopE에 대한 합성 유전자를 구입하고, ABCA4-1150-CfaN, ABCA4-1150-CfaC, ABCA4-1140-CfaN, ABCA4-1140-CfaC, EGFP-71-CfaN 및 EGFP-71-CfaC에 제한 효소 프리 클로닝에 의해 도입하여, 구조체 ABCA4-1150-CfaN-SopE (서열번호 19), ABCA4-1150-CfaC-SopE (서열번호 20), ABCA4-1140-CfaN-SopE (서열번호 21), ABCA4-1140-CfaC-SopE (서열번호 22), EGFP-71-CfaN-SopE (서열번호 23) 및 EGFP-71-CfaC-SopE (서열번호 24)를 구축하였다. DD1에 대한 유전자는 증폭 연장 PCR에 의해 제조하고, EGFP-71-CfaN 및 EGFP-71-CfaC에 제한 효소 프리 클로닝에 의해 도입하여, 구조체 EGFP-71-CfaN-DD1 (서열번호 25) 및 EGFP-71-CfaC-DD1 (서열번호 26)을 구축하였다. 모든 재조합 플라스미드의 동일성을 서열분석으로 검증하였으며, 대응되는 단백질 서열을 표 1에 기재한다. 태그론을 함유한 구조체는 분할 단백질 유전자, 분할 인테인 및 태그론 등의 여러가지 인자들을 발현 플라스미드에 클로닝함으로써 마찬가지로 구축하였다.

[0297] AAV 제조

[0298] CfaN 인테인과 함께 EGFP 또는 ABCA4 단편을 코딩하는 본 발명의 N-단편을 코딩하는 AAV를 표준적인 트리플 형질감염 전략으로 제조하였다. 야생형 AAV2 ITR을 가진 rAAV8 벡터는 플라스미드 pAAV-Rep2-Cap8, pHelper 및 AAV-전이유전자-플라스미드를 HEK293 세포에 폴리에틸렌이민 (PEI) 매개 공동-형질감염을 통해 제조하였다.

[0299] 세포 및 배지를 형질감염 후 48시간에 회수하였다. 0.1% 트리톤을 처리하여 세포로부터 바이러스를 해리시키고, 배양 배지로부터 AAV 입자를 PEG-석출에 의해 수득하였다. 세포용해 조산물의 정제 및 빈 캡시드의 제거는 Zolotukhin 및 동료의 방법에 따른 이오딕사놀 농도구배에 의해 수행하였다 [Zolotukhin S, Byrne BJ, Mason E, Zolotukhin I, Potter M, Chestnut K, Summerford C, Samulski, RJ, Mucyczka N. (1999) Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. Gene Ther; 6(6): 973-85].

[0300] 정제한 배치를 농축하고, Vivaspin 20 Centrifugal Concentrator 100.000 MWCO (Sartorius, Cat.No. VS 064 1)을 이용해, 정량적인 중합효소 연쇄 반응 (qPCR)에 의해 결정된 바에 따른 최종 농도 1×10^{13} vg/ml로 제형화하

었다. 사용한 제형 완충제는 BSS (Alcon)+Pluronic에 의해 특정된다.

- [0301] 바이러스 배치를 농축한 후 Acrodisc(R) 시린지 필터 (Acrodisc PP, PES, 0.2 μ m 1 cm²)를 사용해 여과하여 -80°C에서 보관하였다. ml 당 계승 카피수로서 바이러스 역가는 ITR 특이적인 PCR-프라이머를 이용한 qPCR에 의해 결정하였다. 또한, 캡시드 역가는 AAV8 ELISA (Progen)를 제조사의 지침에 따라 사용해 수득하였다.
- [0302] 최종 생산물에 존재하는, 1×10^{10} vg에 해당하는, 단백질을 단백질 가시화를 위해 10% SDS PAGE에서 Pierce Silver stain kit (Cat#24612)를 이용해 전개하였다. VP1, VP2 및 VP3 단백질이 올바른 화학량론 1:1:10으로 검출가능하였으며, 이는 AAV 조제물의 순도가 > 95%임을 의미한다.
- [0303] HEK293 세포에서 인테인 플라스미드의 형질감염:
- [0304] HEK293T 세포는 10% FBS 및 항생제가 첨가된 DMEM에서 37°C 및 6% CO₂ 분위기 하에 유지시켰다. 세포는 약 80% 컨플루언스에서 리포펙타민 2000을 이용해 공동-형질감염시켰으며, 6웰 플레이트 포맷에서 각 플라스미드 1.25 μ g을 사용하였다. 전장 유전자를 코딩하는 플라스미드를 이용한 실험에서는, 스크램블 플라스미드를 전장 플라스미드와 함께 공동-형질감염시켜, 인테인 플라스미드 2종을 이용하였을 때와 동일한 수준의 DNA 형질감염을 달성하였다. 형질감염 후 48시간에 세포를 회수하고, EGFP를 웨스턴 블롯으로 분석하였다.
- [0305] HEK293 세포에서 테그론-인테인 플라스미드 형질감염:
- [0306] 세포를 약 80% 컨플루언스에서 리포펙타민 2000과 각 플라스미드 1.25 μ g을 6웰 플레이트 포맷으로 이용해 공동-형질감염시켰다. 인테인-테그론 플라스미드를 형질감염시킨 세포를 형질감염 후 48시간에 회수하고, 웨스턴 블롯으로 분석하였다. 시간 경과 실험에서, 형질감염 후 24시간 및 48시간에 세포를 회수하고, 웨스턴 블롯으로 분석하였다.
- [0307] 프로테아좀 저해제 실험을 MG-132 저해제를 이용해 수행하였다. 세포를 공동-형질감염시키고, 형질감염 후 24시간에 DMSO 또는 MG-132 (DMSO에 용해)를 최종 농도 50 μ M로 첨가하였다. 30분, 3시간, 6시간 및 24시간에 세포를 회수하여, 웨스턴 블롯으로 분석하였다.
- [0308] 웨스턴 블롯 분석:
- [0309] EGFP 형질감염된 세포 (HEK293 및 WERI-RB1 세포)를 프로테아제 저해제 및 1 mM 페닐메틸설포닐이 첨가된 RIPA 완충제에서 세포용해 처리하였다. 세포용해 후, 샘플을 BCA 단백질 분석 키트를 사용해 정량하였다. 총 단백질 10 μ g을 함유한 샘플을 95°C에서 10분간 1X Laemmli 샘플 완충제에서 변성시켰다. 세포용해물을 12% 비스-트리스 SDS-PAGE 겔에서 40분간 165V에서 분리시켰다. 면역-블롯팅용 항체는 EGFP를 검출하기 위한 항-6xHis 태그 및 로딩 대조군으로서 항- β -튜불린이었다. 웨스턴 블롯에 의해 검출한 EGFP 밴드에 대한 정량화를 LI-COR Odyssey Infrared Imager를 이용해 수행하였다.
- [0310] ABCA4 형질감염 세포 (HEK293)를 프로테아제 저해제 및 1 mM 페닐메틸설포닐이 첨가된 PBS (1:1) 중의 도데실 말토시드 (D310) 및 콜레스테릴 헤미숙시네이트 (CH210) 용액에서 세포용해 처리하였다. 세포용해 후, ABCA4 샘플을 BCA 단백질 분석 키트로 정량하였다. 총 단백질 25 μ g을 함유한 샘플을 37°C에서 15분간 2.5 mg/ml 아슬렉틴이 함유된 1X Laemmli 샘플 완충제에서 변성시켰다. 세포용해물을 3-8% 트리스-아세이트 SDS-PAGE 겔에서 1.5시간 동안 150V에서 분리시켰다. 면역-블롯팅용 항체는 ABCA4 단백질을 검출하기 위한 항-FLAG 태그 또는 항-ABCA4 및 로딩 대조군으로서 항- β -튜불린이었다. 웨스턴 블롯에 의해 검출한 ABCA4 밴드에 대한 정량화를 LI-COR Odyssey Infrared Imager를 이용해 수행하였다.
- [0311] 형광-활성화된 세포 분류 측정:
- [0312] HEK293T 세포에 각 EGFP-인테인 플라스미드 2 μ g을 공동-형질감염시키고, 음성 대조군으로서 HEK293T 세포를 배양하였으나 형질감염시키진 않았다. 전장 유전자를 코딩하는 플라스미드를 이용한 실험에서는, 스크램블 플라스미드를 전장 플라스미드와 함께 공동-형질감염시켜, 인테인 플라스미드 2종을 이용하였을 때와 동일한 수준의 DNA 형질감염을 달성하였다. 형질감염 후 48시간에 세포를 유세포 측정기로 분석하였다. 형질감염 세포 및 비-형질감염 세포를 각각 FACS 분석 완충제 (PBS, 2% FSA, 2 mM EDTA)에 재현탁하였다. EGFP+ 세포의 %를 Beckman Coulter Gallios 유세포 측정기에서 Kaluza 유세포 측정 분석 소프트웨어를 사용하여 여러가지 형질감염 세포를 비-형질감염 세포와 비교함으로써 분석하였다. Dapi 염색을 이용해 사멸 세포의 수를 검출하였다.
- [0313] 질량 분광 측정 분석:

[0314] 겔 밴드를 암모늄 바이카보네이트 (50 mM NH_4HCO_3) 및 아세토니트릴 (ACN)로 행구었다. 샘플을 20 mM DTT를 60 분간 60℃에서 처리해 환원시키고, 55 mM 요오도아세트아미드를 25℃에서 30분간 암 조건에서 처리해 알킬화하였다. 그 후, 샘플을 2시간 동안 밤새 37℃에서 트립신과 함께 이중 절단 처리하였다 (서열 등급 변형 트립신). 마지막으로, 제조된 펩타이드 혼합물을 겔 매트릭스로부터 5% 포름산 (FA) / 50% ACN, 및 100 % ACN을 사용해 추출하고, SpeedVac 진공 시스템에서 건조시켰다. 수득한 펩타이드 혼합물을 C18 tip (PolyLC Inc.)을 제조사의 프로토콜에 따라 사용해 행구었다. 마지막으로, 행군 펩타이드 용액을 건조시켰다. 트립신 절단 혼합물을 1% FA 용액에 재현탁하고, 각 샘플에 대해 분액을 크로마토그래피 분리를 위해 주입하였다. 펩타이드를 Symmetry C18 트랩 컬럼 (5 μm 180 μm x 20mm; Waters)에 포획하고, C18 역상 모세관 컬럼 (ACQUITY UPLC M-Class Peptide BEH column; 130Å, 1.7 μm , 75 μm x 250 mm, Waters)을 사용해 분리하였다. 펩타이드 용출에 이용한 농도구배는 25분간 B 1 → 40%, 이후 5분간 40%에서 50%로 농도 구배 (A: 0.1% FA; B: 100% ACN, 0.1% FA)였으며, 유속은 250 nL/분이었다. 용출된 펩타이드는 이미터 니들 (PicoTip™, New Objective)에서 적용 전압 2000V으로 전기분무 이온화에 투입하였다. 펩타이드 질량 (m/z 300-1700)을 데이터 의존적인 방식으로 분석하였으며, 풀스캔 MS를 400m/z에서 분해능 60,000 FWHM으로 Orbitrap에서 획득하였다. 15번째까지 가장 풍부한 펩타이드 (최소 강도 500 카운트)를 각 MS 스캔에서 선택하여, CID (38% 정규화된 충돌 에너지) 및 충돌 기체로서 헬륨을 이용해 선형적인 이온 트랩으로 단편화하였다. 데이터베이스 검색을 Sequest HT 검색 엔진에서 Thermo Proteome Discover를 이용해 수행하였다.

[0315] 생체내 실험

[0316] 실험은 4주령의 야생형 동물, 129Sv 품종에서 수행하였다. 검사 샘플을 마우스에 망막하 주사하였다.

[0317] 망막하 주사를 동공 확장 후 5 ml 해밀턴 주사기로 33G 바늘을 사용해 수행하였다. 주사 후 4주차에 EGFP 발현 안저 평가를 디지털 이미징 시스템과 연결된 Canon UVI 망막 카메라를 사용해 수행하였다. 망막 구조 및 바깥핵 층 (ONL)의 정량화를 SD-OCT (빛간섭 단층 촬영)에 의해 Spectralis 이미징 시스템 (Heidelberg Engineering Inc.)을 이용해 평가하였다.

[0318] 이미지 촬영 후, 동물을 안락사시키고, 면역 조직화학 및 웨스턴 블롯으로 추가로 분석하기 위해 눈에서 탈핵 처리하였다. 6주령의 C57BL/6 마우스 내이에 백터를 주사하는 추가적인 실험을 수행하였다.

[0319] 결과 및 고찰

[0320] 최근 몇가지 새로운 인테인들은 컨센서스 설계에 기반하여 조작되었으며 (Stevens et al., 2016; Stevens, Sekar, Gramespacher, Cowburn, & Muir, 2018), 자연 생성 인테인보다 우수한 특성을 가지는 것으로 입증되었다. Cfa로 지칭되는 이들 인테인 중 하나는 DnaE 인테인 정렬로부터 컨센서스 설계에 의해 구축되었으며, Npu와 같은 DnaE 패밀리의 성능이 가장 우수한 인테인들 중 일부보다 우수한 특성을 가진 것으로 입증되었다. Cfa는 더 빠른 카이네틱스, 더 높은 발현 수준 및 고온 및 변성체의 고 농도와 같은 높은 용인 내지 극한의 조건을 가진 것으로 보고되었다. 흥미롭게도, 상동성 수준이 90% 이상인 Cfa 변이체는 비슷한 특성을 나타낸다. 동일한 컨센서스 설계 전략을 TerL-AceL 인테인 패밀리에 적용하여, Cat로 지칭되는 컨센서스 서열을 구축하였으며, 이 역시 TerL-AceL 패밀리의 나머지 구성원들보다 개선된 특성을 보인다.

[0321] 컨센서스 인테인이 유전자 요법 활용에서 단백질 트랜스-스플라이싱에 의한 단백질 재구성에 유용성을 제공하는지를 입증하기 위해, 직접 비교를 수행하였다. 비교는 세포에 정제된 플라스미드를 형질감염시켜 수행하였다.

[0322] 처음 실험은 리포터 유전자로서 EGFP를 이용해 수행하였다. 간략하게는, EGFP는 71번 위치에서, N-인테인에 재조합에 의해 융합된 N-말단 단편 (잔기 1-70, EGFPN)과 C-인테인에 융합된 C-말단 단편 (잔기 71-239, EGFPNC)으로 분할되었다. Cfa 컨센서스 서열의 재구성 효율을 Npu와 비교하기 위해, 다음과 같은 구조체 4종을 구축하였다: EGFPN-CfaN, EGFPN-NpuN, CfaC-EGFPNC 및 NpuC-EGFPNC. 배양한 HEK293 세포에 N-말단 단편과 C-말단 단편을 코딩하는 플라스미드를 동일 몰 함량으로 공동-형질감염시키고, 형광 현미경, 유세포 측정 분석 및 웨스턴 블롯팅에 의해 스플라이싱 효율을 모니터링하였다.

[0323] 본 결과는, 도 2, 3 및 4를 참조하여, 컨센서스 Cfa 인테인이 Npu 인테인과 비교해 EGFP 재구성을 2.5배 증가시키는 것으로 확인되었으며, 이는 컨센서스 설계에 의해 수득된 인테인이 공동-형질감염시 표적 단백질을 더 높은 수율로 제공함을 의미한다. Cfa가 시험관내에서 Npu보다 더 빨리 분할하는 것으로 알려져 있지만, 이의 N-말단 단편은 세포에 개별 형질전환 또는 형질감염시 더 잘 발현되었으며, 만일 이것이 양쪽 단편을 동일 세포에 공동-형질감염하였을 때 달성되는지에 대해서는 입증된 바 없다는 점을 중요하게 지적하여야 한다. 이들 실험에서, IntN 및 IntC 단편을 동일 세포에서 발현시켰을 때, Cfa가 기존에 알려진 매우-빠른 분할 인테인, 예를 들

어 Npu와 비교해 표적 단백질의 재구성 수율에 명백한 이점을 제공하는 것으로, 입증되었다. EGFP의 N-단편과 C-단편 및 N-인테인과 C-인테인을 코딩하는 유전자 역시 생체내 전달하기 위한 재조합 AAV에 통합하였다. 그 결과, Cfa 매개 단백질 스플라이싱으로 망막 및 내이 등의 여러 장기들에서 EGFP 재조합이 달성되는 것으로 확인되었다 (도 23).

- [0324] Cfa 컨센서스 서열의 이러한 긍정적인 특징이 다른 단백질에서도 관찰되는지를 검증하기 위해, ABCA4 단백질을 이용한 실험을 반복 실시하였다. ABCA4는 스타가르트 질환에서 돌연변이된 거대 단백질로서, 이의 재구성은 질환을 치료하기 위한 실행가능한 전략으로서 제안된 바 있다. AAV 유전자 요법에 기반한 몇가지 방식들은 현재 이를 재구성하기 위해 탐색 중에 있다. ABCA4는 이의 큰 크기로 인해, 단일 AAV로 캡슐화할 수 없으며, 그래서 ABCA4를 2개의 단편으로 전달해 표적 세포 안에서 재구성하는 다른 전략을 제안하였다. Cfa와 같은 조작된 컨센서스 서열이 자연 생성 인테인, 예를 들어 Npu보다 유익성을 제공하는지를 검증하기 위해, ABCA4를 1150번 위치에서 분할하였으며, N-인테인 및 C-인테인을 단백질의 N-말단 단편 및 C-말단 단편으로 각각 클로닝하였다.
- [0325] $ABCA4^N(1-1149)$ -IntN 및 $IntC-ABCA4^C(1150-2273)$ 를 Cfa 또는 Npu 인테인과 함께 발현 플라스미드에 클로닝한 다음 HEK293 세포에 공동-형질감염시켰다. 세포를 용해 처리하고, 웨스턴 블롯으로 단백질 재구성 수율을 결정하였다. 전장 ABCA4 역시 공동-형질감염시켜 대조군으로 이용하였다.
- [0326] EGFP에서 관찰된 바와 같이, Npu보다는 Cfa를 이용한 경우에 더 높은 재구성 수율이 관찰되었다 (도 5 참조).
- [0327] CfaC 단편에 특이적인 돌연변이를 가진 Cfa의 변이체, 즉 CfaCmut (Stevens et al., 2017) 역시 최근 Cfa 이점을 추가하는 것으로 보고되었다. 특히, CfaCmut 변이체는, 심지어 C-엑스테인의 +2 위치에 네이티브 Phe 잔기가 아니더라도 단백질 단편을 효율적으로 스플라이싱할 수 있는 것으로, 입증되었다. 상기 정의된 바와 같이, +2번 위치는 IntC의 마지막 아미노산을 기준으로 +2 위치를 지칭한다. ABCA4가 1150번 위치에서 분할된다면, C-엑스테인 +2 위치에 Phe가 남겨지며, 이는 Npu 뿐 아니라 Cfa에 대해서도 최적의 잔기이다.
- [0328] ABCA4가 1150번 위치에서 이를 분할되어 재구성할 수 있음에도 불구하고, 이는 치료학적 목적에서 가능한 최상의 위치가 아닐 수 있다. 1150 이외의 다른 위치에서 이를 분할함으로써 ABCA4를 재구성하는 능력을 조사하기 위해, 인테인에 대해 +2 위치에 Phe가 있어야하는 위치들로 한정하지 않도록 CfaCmut를 사용하였다.
- [0329] ABCA4 단백질이 분할 부위의 +2 위치에 Phe가 없는 부위에서 분할되는 수종의 ABCA4 구조체들을 클로닝하였다. 검사한 추가적인 부위 2곳은, +2 부위가 Ser인 1140번 위치와, +2 부위가 Pro인 1188번 위치였다. 구조체들을 전술한 바와 같이 클로닝하여 검사하였다. 구체적으로, $ABCA4(1-1139)$ -IntN, $IntC-ABCA4(1140-2273)$ 구조체를 이용해 1140번 위치에서 ABCA4 스플라이싱을 평가하고, $ABCA4(1-1187)$ -IntN 및 $IntC-ABCA4(1188-2273)$ 를 이용해 1188번 위치에서 ABCA4 분할을 평가하였다. 이들 부위는 ABCA4의 위상 구조에 기반하여, folded 도메인의 존재를 고려하여, 선정하였다. 부위는 ABCA4의 세포내 부위에서 잘-정의된 folded 도메인의 바깥쪽을 선택하였다. 여러 부위를 조사하였으며, 이들 모두에서 Cfa 및 Cfa 변이체가 Npu보다 재구성 수율이 높고, 분할 부위에 따라 효율이 다르다는 것을 관찰하였다 (도 5, 6 및 7). 또한, 본 발명자들은, 이러한 방식으로 1179 및 1177번 위치에서 ABCA4 분할의 재구성을 달성할 수 있음을, 입증하였다. 1140 및 1188에서 상기 정의된 구조체와 유사한 구조체를 클로닝하고, 이를 이용해 1179 및 1177번 위치에서의 스플라이싱 효율을 검사하였으며, 1179번 위치에서의 결과를 도 21에 나타낸다.
- [0330] 본 발명자들은, 또한, gp41과 같은 다른 효과적인 인테인을 이용해 분할 단편들로부터 ABCA4를 재구성할 수 있음을 입증하였다. 1185 및 1095번 위치에서 분할된 ABCA4 구조체들을 설계하고, N-말단 인테인 및 C-말단 인테인에 대한 대응되는 융합체를 콜로닝한 다음, 이의 스플라이싱 활성을 전술한 바와 같이 분석하였다. 분할 위치 1096 및 gp41 인테인에서 PTS 반응의 웨스턴 블롯 분석 결과를 도 22에 나타낸다.
- [0331] 아래 표 7은 단백질을 여러 부위 (1177, 1179, 1096 및 1185)에서 분할한 경우에 수득되는 ABCA4 단편들 및 대응되는 ABCA4-인테인N 또는 인테인C-ABCA4 융합체의 서열을 나타낸다.
- [0332] 재구성 수율 이외에, 유전자 요법에서 분할 인테인의 사용의 또 다른 주요 한계는 미-반응 출발 물질 및/또는 인테인의 존재이다. 부적절한 출발 물질의 축적을 방지하고 절단된 인테인 단편을 제거하기 위해, 도 8에 도시된 일반적인 구성으로 구조체에 태그론을 추가하였다.
- [0333] 본 발명자들은, ABCA4에 대한 유전자 요법 방식을 개발하기 위해, 다른 거대 유전자에서 돌연변이에 의해 유발되는 질환으로 전환될 뿐 아니라 질환을 치료하기 위한 단백질을 재구성하기 위해, 예를 들어 CRISPR/Cas9 시스템을 재구성하기 위해, 인테인과 조합 사용할 수 있는 수종의 태그론을 동정하였다.

- [0334] 적절한 테그론-인테인 조합을 식별하기 위해, 선택 테그론을 포함하는 EGFP-Int 구조체들을 클로닝하였다. 테그론은 N-인테인의 C-말단에, C-인테인의 N-말단에 클로닝하였다. 검출 목적으로 2개의 인자 사이에 링커로서 His6 태그를 포함시켰다. 원리 증명으로, SopE 탈안정화 도메인 (1-100)뿐만 아니라 펩타이드 서열 DD1으로 이루어진 테그론을 이용하였다 (표 3 참조).
- [0335] 그 결과, 테그론의 함유가 실제 출발 물질과 인테인을 임의의 검출가능한 양으로 제거하는 것으로 확인된다 (도 9, 도 14 참조). 중요하게도, 예상치 못하게도, 2개의 테그론의 동시 함유는 단백질 스플라이싱 반응에 어떠한 수율 감소 없이 임의의 미-반응 출발 물질을 제거할 수 있는 것으로, 입증되었다.
- [0336] 흥미롭게도, 테그론 전략을 거대 유전자 ABCA4에 적용할 경우, 비슷한 결과가 관찰되었지만, 스플라이싱된 ABCA4 생산물의 수준 역시 증가하였는데, 이는 컨센서스 인테인과 테그론 간의 예상치 못한 상승적인 작용을 의미한다 (도 10 참조).
- [0337] 프로테오솜의 저해제가 이러한 결과에 미치는 효과를 또한 조사하였다 (도 11). 본 발명자들은, 프로테오솜 저해제가 테그론의 효과를 감소시키는 것으로 확인하였으며, 즉 SopE 분해가 프로테오솜 매개로 이루어진다는 것이 검증되었다. 흥미롭게도, 본 발명자들은, 인테인과 테그론의 조합시, 출발 물질의 존재를 제거할 뿐 아니라 ABCA4의 재구성 수준을 유지 또는 심지어 증가시킬 수 있다는 것을, 관찰하였다 (도 10).
- [0338] 또한, 본 발명자들은 컨센서스 인테인 Cfa와 테그론의 조합 효과를 조사하고 이를 테그론 부재 또는 다른 인테인의 사용과 비교하기 위한 실험을 수행하였다. 흥미롭게도, 본 발명자들은 (도 12), Cfa와 테그론의 사용이 출발 물질의 수준을 현저하게 낮추고 또한 생산물의 양을 증가시키는 것을, 관찰하였다. 또한, 본 발명자들은, 이러한 효과가 ABCA4의 1140번 위치 (도 13 및 도 21) 및 1179번 위치 (도 21)에서도 관찰됨을 검증하였다. 중요하게도, 본 발명자들은, Cfa (및 Cfamut) 인테인과 일부 특징들을 공유하는 수종의 여러가지 테그론을 조합한 관찰된 효과들이 유지되는 것을 확인하였다 (실시예 3 참조).
- [0339] ABCA4에서 수득한 결과를 토대로, 이러한 방식을 적용할 수 있는 기타 질환, 유전자 및 단백질은 표 1에 열거된 것을 포함한다.
- [0340] **실시예 2. 오르소고날 분할 인테인 쌍들을 이용한 3-조각 결찰 전략**
- [0341] **재료 및 방법:**
- [0342] HEK293 세포에 3-조각 플라스미드 형질감염:
- [0343] 세포를 약 80% 컨플루언스에서 리포펙타민 2000과 각 플라스미드 1.25 μ g을 6웰 플레이트 포맷으로 이용해 공동-형질감염시켰다. 3-조각 플라스미드를 형질감염시킨 세포를 형질감염 후 48시간에 회수하고, 웨스턴 블롯으로 분석하였다. 대조군으로서, 세포에 플라스미드 2종 플라스미드 (POIN-CfaN+CfaC-POIM-CatN, CfaC-POIM-CatN+CatC-POIC 및 POIN-CfaN+ CatC-POIC)을 공동-형질감염시켰다.
- [0344] **웨스턴 블롯 분석**
- [0345] 3-조각 형질감염 세포 (HEK293)를 프로테아제 저해제 및 1 mM 페닐메틸설포닐이 첨가된 RIPA 완충제에서 세포용해 처리하였다. 세포용해 후, 샘플을 BCA 단백질 분석 키트를 사용해 정량하였다. 총 단백질 10 μ g을 함유한 샘플을 95°C에서 10분간 1X Laemmli 샘플 완충제에서 변성시켰다. 세포용해물을 12% 비스-트리스 SDS-PAGE 겔에서 40분간 165V에서 분리시켰다. 면역-블롯팅용 항체는 POI를 검출하기 위한 항-FLAG 태그 및 로딩 대조군으로서 항- β -튜불린이었다.
- [0346] **결과**
- [0347] 본 발명자들은 매우 빠른, 컨센서스 인테인을 테그론과 조합하는 것이 출발 물질과 절단된 인테인의 수준을 감소시키면서 거대 단백질을 재구성하기 위한 실행가능한 전략임을 입증하였다. 이 전략은, 본원에 기술된 것과 같은 적절한 전달 벡터와 조합할 경우, 유전자 치환 전략으로 거대 단백질을 재구성하고 아울러 생체내 치료학적 물질을 구축하기 위해 유전자 요법에서 이용할 수 있었다. 그러나, 심지어 더 큰 단백질의 경우에, 이 전략은 충분하지 않을 수 있다. 예를 들어, 코딩 영역이 7-8 kb보다 큰 단백질의 경우, 이의 코딩 유전자를 2개의 단편으로 분할하는 것은 AAV 벡터 안에 캡슐화될 수 있는 단편을 만들기엔 충분하지 않을 것이다. 이러한 단백질의 경우, 코딩 유전자는 3개의 조각으로 분할해야 할 수 있다. 분리된 3개의 단편을 인테인을 이용해 단백질로 조립하기 위해, 오르소고날 인테인 쌍들이 필요할 것이다. 최대 재구성 수율을 달성하기 위해, 고 효율의 오르소고날 인테인이 필요할 것이다. 본 발명자들은 Cfa를 Cat 인테인과 조합하기로 결정하였다. 이들 인테인 2종

은 컨센서스 설계를 통해 수득하였으며, 고 발현 수율, 열 안정성, 카오토로픽제에 대한 용인성 및 빠른 스플라이싱 카이네틱스 등의 일부 공통적인 특징을 공유한다. 본 발명자들은 도 16에 나타난 구조 구성에 따라 일련의 구조체들을 설계하였다: POIN-CfaN, CfaC-POIM-CatN 및 CatC-POIC. 대안적으로, Cfa 및 Cat 인테인의 위치를 교체하여, 다음과 같은 구조 구성을 가진 구조체를 구축하였다: POIN-CatN, CatC-POIM-CfaN 및 CfaC-POIC.

[0348] 본 발명자들은 오르소고날리티 (orthogonality)를 검사하였으며, 2종의 인테인이 실제 오르소고날이고, 서로, 즉, Cfa의 N-단편 또는 Cat 인테인은 서로 반응하지 않고 단지 각각의 동족 쌍과만 반응함을, 입증하였다. 단편 3개를 동일 물량으로 공동-형질감염시켰을 때, 검출된 주요 생산물은 단편 3개가 전장의 원하는 생산물로 조립됨으로써 생성되는 것이었다. 중요하게도, 본 발명자들은 미-반응 물질을 낮은 수준으로만 검출하였으며, 단편 2종의 반응으로 인해 생기는 생산물인 임의의 중간산물은 검출되지 않았다. 이러한 결과는, 오르소고날 컨센서스 인테인의 고 활성 쌍을 사용하는 것이 개개 단편 3개로부터 거대 단백질을 조립하기 위한 적절한 전략임을 보여준다.

[0349] 실시예 3.

[0350] 본 발명을 실시하기 위해, 수종의 데그론을 검사하였다. AAV 유전자 요법 활용에서 아미노산 75개 미만의 데그론이 제조되는 전이유전자의 크기를 최소화하는데 바람직할 것이므로, 크기에 기반하여 문헌들에서 데그론을 선정하였다. 일부 경우에, 데그론은 공지된 N-말단 데그론 또는 C-말단 데그론을 조합함으로써 설계하였다.

[0351] 데그론은 N-단편의 C-말단에, C-단편의 N-말단에 클로닝하였으며, 이들 단편 2개는 용해성 세포질 단백질에 대한 대표적인 예로서 전장 EGFP 단백질을 재구성하도록 의도하였다. 도 9에서, 결과 요약으로, 데그론 하나 또는 2종 (EGFP 단편 각각에 대해 하나) 사용시 스플라이싱 수율에 부정적인 영향 없이 부적절한 출발 물질 단편이 소거됨을 알 수 있었다. 비슷한 결과는 다른 데그론에서도 관찰되었으며 (도 14), 이는 본 방식의 일반성을 보여준다.

[0352] N-단편의 C-말단에, C-단편의 N-말단에 데그론을 클로닝하였으며, 이들 단편 2개는 대표적인 예로서 전장 ABCA4 단백질을 1140, 1150 및 1179 등의 여러가지 분할 위치에서 재구성하도록 의도하였다. 아래 표 6은 검사한 데그론을 크기, 서열 및 기원을 비롯하여 모두 나타내며, 또한 이의 유비퀴틴화 및 궁극적인 분해를 담당하는 리가제도 표시한다.

[0353] 표 6.

표 6

Name	bp		DD origin	Human ortholog
DD1	39	RISFGSPPPMAGG	H. sapiens	KLHDC2
L2	105	SLISLPLPTRVKFSSLLIRIMKIITMTFPKKLRS	S. cerevisiae	UBE2G2
L9	139	SNPFSSLFGLASLLIDSVSLKSNWDTSSSSCLISFFSSVMFSSTTRS	S. cerevisiae	UBE2G2
SopE50	105	QETTLKEKSTEKNSLAKSILAVKNHFIELRSKLS	S. typhimurium	UBE2D3
SopE70	189	QETTLKEKSTEKNSLAKSILAVKNHFIELRSKLSERFISHKNTESSATHFHRSASEGRAVL	S. typhimurium	UBE2D3
PEST	90	RSSSPSDSDTSFGSSGSDHLSDLISLRIS	Xenopus	KLHDC2
M2	57	VSAFNLNSLIVGILRFHW	S. cerevisiae	TEB4
M4	63	YMSILRCASGKISIAAPPYIF	S. cerevisiae	LTN1
V12	69	TNGSVLREFTLLELVVTRNTEL	S. cerevisiae	TEB4
DD3	33	TNGILKGCQG	S. cerevisiae	TEB4
DD4	39	RDSFGSPPPMAGG	H. sapiens	KLHDC2
DD5	12	GFWI	S. cerevisiae	KLHDC2
DD6	12	RFKG	S. cerevisiae	TEB4
DD7	12	KFYK	S. cerevisiae	TEB4
T1	63	SVLEENRPFQAQLSNVYFTIL	S. cerevisiae	TEB4
T2	63	KLKGLGKRKRREDLEIRFIL	S. cerevisiae	TEB4
T3	63	WKLLLVVGLVLVKHHDGAAH	S. cerevisiae	TEB4

[0354]

- [0355] 데그론이 출발 물질 및 절단된 인테인의 분해를 유도할 수 있는지를 검증하고 또한 단백질 트랜스-스플라이싱 반응의 수율에 대한 효과, 즉 전장 ABCA4 단백질을 재구성하는 능력을 검증하기 위해 데그론을 검사하였다. 세포에 전술한 여러가지 데그론을 함유한 N-말단 구조체 및 C-말단 구조체를 공동-형질감염시켰다. 세포를 용해 처리한 다음 웨스턴 블롯으로 분석하여 PTS 생산물의 존재를 데그론 함유 및 비-함유 상태에서 검출하였다. 도 18에서, 동일한 데그론을 N-단편 및 C-단편에 사용한 경우의 결과를 나타낸다. 좌측 패널의 겔 (도 18C)은 PTS 생산물 (ABCA4)과 아울러 각 반응에서의 출발 물질 (ABCA-IntN-DD 및 DD-IntC-ABCAC)을 보여준다. PTS는 데그론을 사용하지 않았을 경우의 결과에 해당하며, 다른 레인에서의 표지는 각 경우에 데그론을 사용한 것을 나타낸다. 우측 패널 (도 18C)은 각 경우에 PTS 생산물의 양에 대한 정량 결과를 나타낸다. 알 수 있는 바와 같이, 놀랍게도, 일부 데그론은 데그론 비-함유 구조체보다 더 높은 수율을 제공하였다. 흥미롭게도, 데그론들 모두 작동하며, 이들 모두 데그론 비-함유 경우에 수득되는 생산물의 적어도 25%를 재구성할 수 있었다. 일부 데그론은 데그론 비-함유시 수득되는 수준의 50% 초과로 재구성을 달성할 수 있으며, 최종적으로, 일부 데그론은 재구성 수율이 100% (또는 그 이상)이었다. 특히, 재구성 수율이 100% (또는 그 이상)인 데그론은 다음과 같다: DD1, V12, M4, L2, L9, DD3 및 SopE100.
- [0356] 절단된 인테인에 대한 데그론의 효과 역시 조사하였다 (도 19). PTS와 여러가지 데그론의 반응물을 높은 %의 아크릴아미드 겔을 이용한 WB (웨스턴 블롯)에 의해 분석하여 인테인을 검출하였다. 알 수 있는 바와 같이, 몇몇 데그론은 절단된 인테인의 양을 다른 것 (예를 들어, SpoE 및 PEST)과 비교해 감소시켰으며, L2 (서열번호 52), L9 (서열번호 54), M2 (서열번호 47), M4 (서열번호 35) 또는 DD3 (서열번호 45)와 같은 기타 데그론들은 인테인 밴드를 완전히 제거하였다.
- [0357] 아울러, 본 발명자들은 N-말단 또는 C-말단 위치에서 여러가지 데그론 이용 효과를 조사하기 위한 실험을 수행하였다. 예로서, 도 20에서, N-단편 위치의 데그론을 고정하고, C-단편 위치의 데그론을 바꾼 경우의, 웨스턴 블롯 분석 결과를 도시한다. 도 20에 나타난 결과는, 바람직한 데그론 일부를 조합할 경우, 결과에 대한 추가적인 개선이 달성될 수 있으며, 원하는 스플라이싱 생산물의 수율의 감소 없이 출발 물질이 추가적으로 감소됨을 보여준다. 이들 조합은 특정 활용에, 특히 임의의 출발 물질을 완전히 제거하는 것이 중요한 경우에 유용할 수 있다.
- [0358] 수득한 결과를 토대로 데그론을 3가지 범주로 분류할 수 있었다:
- [0359] * 데그론 부재시 달성되는 수율의 적어도 25%로, 표적 단백질의 재구성을 허용하는 데그론.
- [0360] * 데그론 부재시 달성되는 수율의 약 50% 이상으로, 표적 단백질의 재구성을 허용하는 데그론.
- [0361] * 데그론 부재시 달성되는 수율과 동일한 또는 우수한 수율로, 표적 단백질의 재구성을 허용하는 데그론.
- [0362] CfaN과 CfaC 또는 CfaCmut 인테인의 조합을 이용할 경우, ABCA4의 여러 부위가 본원에 기술된 본 발명을 이용하기에 적합한 것으로 입증되었다. 구체적으로, 뉴클레오티드 결합 도메인 1 (NBD1)의 C-말단과 7번째 막횡단 도메인 (TMD) 사이의, 1140, 1150, 1177, 1179 및 1188번 위치 등의 위치. ABCA4 재구성 수율은 1140, 1150 및 1179번 위치에서 가장 우수하였으며, 1188번 위치가 가장 낮은 효율로 작동하였지만, 여전히 비-컨센서스 Npu 인테인을 이용해 수득한 것보다는 더 높은 수준으로 ABCA4 재구성이 달성되었다.
- [0363] CfaN과 CfaC mut 인테인의 조합은 무차별성의 증가를 제공한다. 실제, 본 발명자들은, 본원에서, 최초로, 이러한 무차별성 증가가 ABCA4와 같은 막횡단 단백질에서 관찰되고, 높은 재구성 수율이 여러가지 스플라이싱 부위에서 CfaN-CfaCmutant 쌍을 이용하여 수득됨을, 입증하였다. 아울러, 출발 물질이 각각 10% 미만으로 관찰되었으며 (도 12에 기술된 바와 같이 추정), 이는 Npu 인테인과 ABCA4를 이용한 기존에 보고된 결과 (Auricchio et al. 2019)에서 출발 물질의 상당량이 반응하지 않고 남아있어, 이는 단백질-트랜스 스플라이싱이 이러한 막 단백질에 본질적으로 더 어려운 것으로 시사되었다는 사실에 기반하면 예상할 수 없었으므로, 이러한 스플라이싱 반응은 이들 인테인을 사용할 경우에 (CfaN 및 CfaC mut 인테인) 매우 효율적인 것으로 입증되었다.
- [0364] 본원에 기재된 결과들을 토대로, 본 발명자들은 다음과 같은 특징을 가진 임의의 부위가 컨센서스 인테인 CfaN 및 CfaCmutant 쌍을 이용할 경우에 효율적으로 작동할 것으로 결론내렸다: (1) 효소적 또는 막횡단 도메인의 바깥 부위, (2) (전술한 바와 같이) 분할 부위의 +1 위치 등의 위치는 Cys 또는 Ser과 같은 친핵 잔기에 해당하여야 하며, +2 위치는 Pro를 제외한 임의의 부위일 것이다.
- [0365] **표 7**

표 7

[0366]

구조체	^a 서열	서열번호
ABCA4 ^N -1177	서열번호 106 (1-1176)	87
ABCA4 ^C -1177	서열번호 106 (1177-2273)	88
ABCA4 ^N -1179	서열번호 106 (1-1178)	89
ABCA4 ^C -1179	서열번호 106 (1179-2273)	90
ABCA4-1177-Cfa ^N	서열번호 87-서열번호 27	91
ABCA4-1177-Cfa ^{Cmut}	서열번호 29-서열번호 88	92
ABCA4-1179-Cfa ^N	서열번호 89-서열번호 27	93
ABCA4-1179-Cfa ^{Cmut}	서열번호 29-서열번호 90	94
ABCA4 ^N -1096	서열번호 106 (1-1095)	95
ABCA4 ^C -1096	서열번호 106 (1096-2273)	96
ABCA4 ^N -1185	서열번호 106 (1-1184)	97
ABCA4 ^C -1185	서열번호 106 (1185-2273)	98
ABCA4-1096-gp41 ^N	서열번호 95-서열번호 38	99
ABCA4-1096-gp41 ^C	서열번호 104-서열번호 96	100
ABCA4-1185-gp41 ^N	서열번호 97-서열번호 38	101
ABCA4-1185-gp41 ^C	서열번호 104-서열번호 98	102

[0367] 항목

[0368]

1. 재구성할 단백질의 N-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 분할 인테인 N-단편을 포함하는 폴리뉴클레오티드를 코딩하는 제1 폴리뉴클레오티드, 및 재구성할 단백질의 C-말단 단편에 펩타이드 결합으로 직접, 선택적으로는 펩타이드 링커를 통해 연결된, 분할 인테인 C-단편을 포함하는 폴리뉴클레오티드를 코딩하는 제2 폴리뉴클레오티드를 포함하는 조성물로서,

[0369]

상기 조성물의 구성성분 2종이 단일 제형으로 함께, 또는 분리하여 서로 다른 제형으로 포장될 수 있으며;

[0370]

제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드는, 양쪽 단편이 조합되면 단백질의 N-말단 단편이 단백질의 C-말단 단편에 연결되어 전체 단백질을 형성되는 방식으로, 재구성할 단백질의 N-말단 단편 및 C-말단 단편을 각각 코딩하고;

[0371]

상기 각각의 폴리뉴클레오티드는, 단백질로 번역되면, N-말단 서열과 C-말단 서열이 트랜스-스플라이싱 반응에서 기능하는 인테인으로 비-공유적으로 재-연합 또는 재구성할 수 있는 별개의 단편들이 되도록, 분할 인테인을 코딩하여야 하며,

[0372]

상기 재구성할 단백질은 25 KDa보다 큰, 조성물.

[0373]

2. 제1 폴리뉴클레오티드는 데그론에 펩타이드 결합으로 직접 연결된 분할 인테인 N-단편을 코딩하고, 데그론은 인테인 N-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 C-말단을 통해 인테인 N-단편에 연결되고, 분할 인테인 N-단편의 N-말단은 재구성할 단백질의 N-말단 단편에 펩타이드 결합을 통해 직접 연결되고; 상기 제2 폴리뉴클레오티드는 데그론에 펩타이드 결합으로 직접 연결된 분할 인테인 C-단편을 코딩하고, 데그론은 인테인 C-단편과 데그론 사이에 링커를 이용해 또는 링커 없이 인테인의 N-말단을 통해 인테인 C-단편에 연결되고, 분할 인테인 C-단편의 C-말단은 재구성할 단백질의 C-말단 단편에 펩타이드 결합을 통해 직접 연결된, 항목 2에 따른 조성물.

[0374]

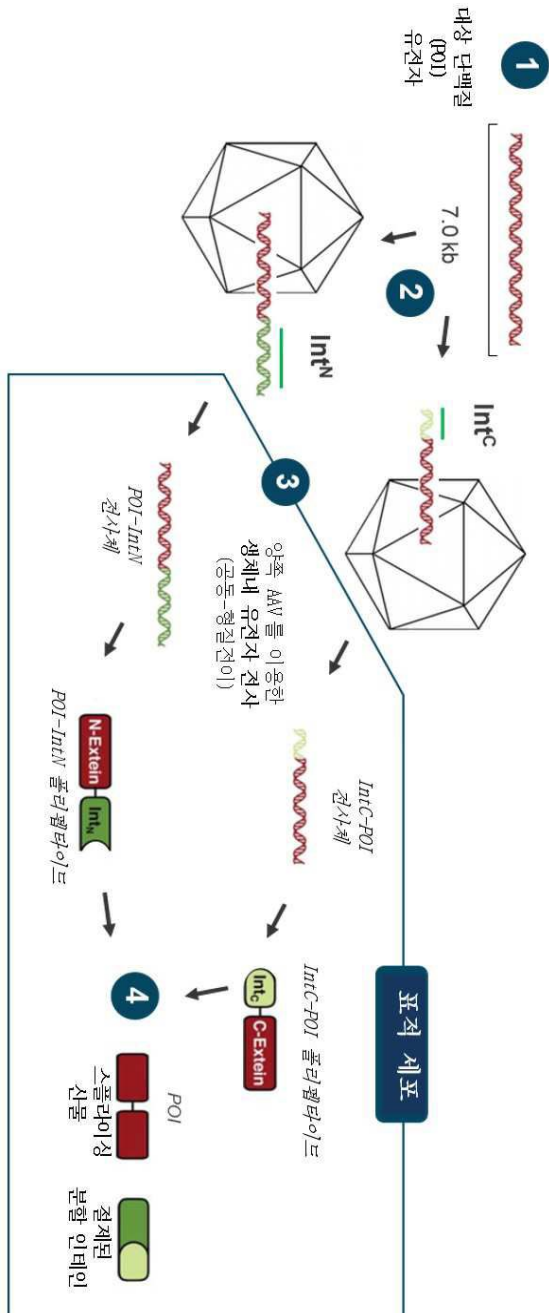
3. 제1 폴리뉴클레오티드가 서열번호 27의 CfaN 인테인 또는 이의 임의의 변이체를 코딩하고, 제2 폴리뉴클레오티드가 서열번호 28의 CfaC 인테인 또는 이의 임의의 변이체를 코딩하고, 변이체는 CfaCmut (서열번호 29)와 같

은 이들 임의의 서열에 대해 서열 동일성이 90% 이상인 서열번호 27 또는 서열번호 28의 분할 인테인 N-단편 또는 분할 인테인 C-단편으로서 이해되는, 항목 1 또는 2에 따른 조성물.

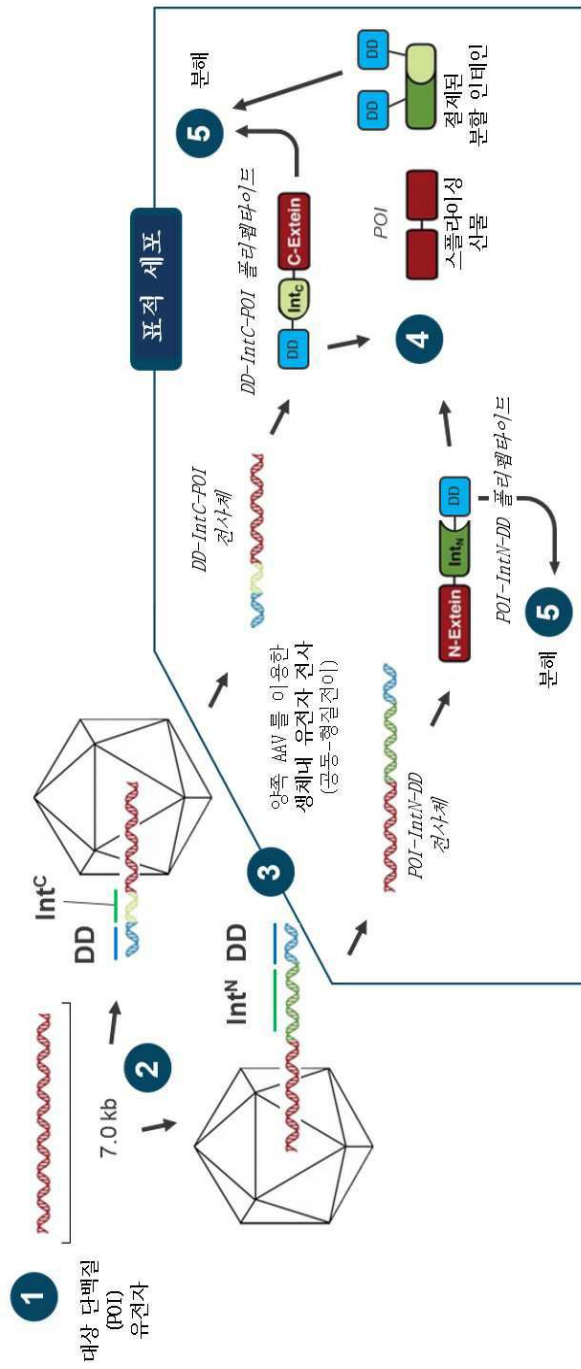
- [0375] 4. 테그론이 서열번호 40-59, 34, 35 및 37로 이루어진 목록으로부터 선택되는, 항목 2 또는 3에 따른 조성물.
- [0376] 5. 분할 인테인이 청구항 3에 따라 정의되고, 테그론이 서열번호 42, 43, 49, 50, 51, 54 및 34로 이루어진 목록으로부터 선택되는, 항목 2에 따른 조성물.
- [0377] 6. 폴리뉴클레오티드 2종 모두 적절한 숙주 세포에서 폴리뉴클레오티드의 증폭을 허용하는 벡터에 포함되는, 항목 1-5 중 어느 하나에 따른 조성물.
- [0378] 7. 아테노-부속 바이러스 (AAV)인, 항목 6에 따른 조성물.
- [0379] 8. 벡터가 혈청형 1, 2, 3, 4, 5, 6, 7, 8 또는 9의 AAV인, 항목 7에 따른 조성물.
- [0380] 9. 전체 단백질을 코딩하는 유전자가 표 1에 열거된 단백질들로 이루어진 목록으로부터 선택되는, 항목 1-8 중 어느 하나에 따른 조성물.
- [0381] 10. 요법에 사용하기 위한, 항목 1-9 중 어느 하나에 따른 조성물.
- [0382] 11. 하기 단계를 포함하는, 대상 유전자를 발현시키는 방법:
- [0383] (i) 세포를 하기와 접촉시키는 단계:
- [0384] (a) 항목 1에서 정의되는 제1 폴리뉴클레오티드, 및
- [0385] (b) 항목 1에서 정의되는 제2 폴리뉴클레오티드,
- [0386] (ii) 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드를 발현시켜, 제1 융합 단백질과 제2 융합 단백질을 생산하는 단계, 및
- [0387] (iii) 제1 단백질과 제2 단백질을 접촉시켜, 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간 산물을 생성하고, 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단을 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결하는 단계.
- [0388] 12. 하기 단계를 포함하는, 대상 유전자를 발현시키는 방법:
- [0389] (i) 세포를 하기와 접촉시키는 단계:
- [0390] (a) 항목 2에서 정의되는 제1 폴리뉴클레오티드, 및
- [0391] (b) 항목 2에서 정의되는 제2 폴리뉴클레오티드,
- [0392] (ii) 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드를 발현시켜, 제1 융합 단백질과 제2 융합 단백질을 생산하는 단계, 및
- [0393] (iii) 제1 단백질과 제2 단백질을 접촉시켜, 분할 인테인 N-단편이 분할 인테인 C-단편에 결합하여 인테인 중간 산물을 생성하고, 인테인 중간산물이 반응하여 대상 제1 폴리펩타이드의 C-말단을 대상 제2 폴리펩타이드의 N-말단에 공유적으로 연결하는 단계.
- [0394] 13. 제1 폴리뉴클레오티드는 서열번호 27의 CfaN 인테인 또는 이의 임의의 변이체를 코딩하고, 제2 폴리뉴클레오티드는 서열번호 28의 CfaC 인테인 또는 이의 임의의 변이체를 코딩하며, 상기 변이체는 CfaCmut (서열번호 29)와 같은 이들 임의의 서열에 대해 서열 동일성이 90% 이상인 서열번호 27 또는 서열번호 28의 분할 인테인 N-단편 또는 분할 인테인 C-단편으로서 이해되는, 항목 11 또는 12에 따른 조성물.
- [0395] 14. 분할 인테인이 청구항 3에 따라 정의되고, 테그론이 서열번호 42, 43, 49, 50, 51, 54 및 34로 이루어진 목록으로부터 선택되는, 항목 11 또는 12에 따른 조성물.
- [0396] 15. 폴리뉴클레오티드 2종 모두 아테노 부속 바이러스 (AAV) 안에 포함되는, 항목 11-14 중 어느 하나에 따른 조성물.

도면

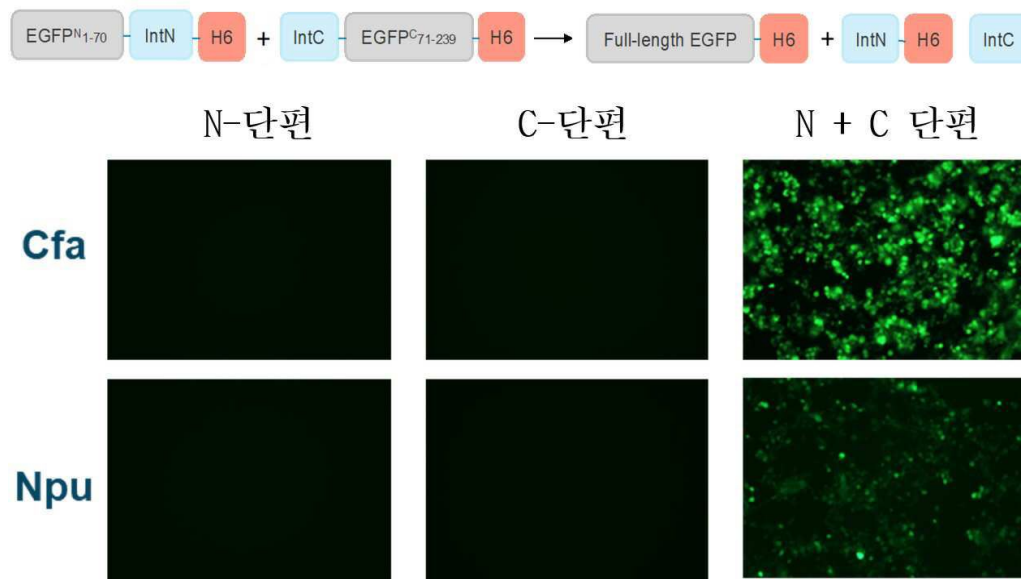
도면1



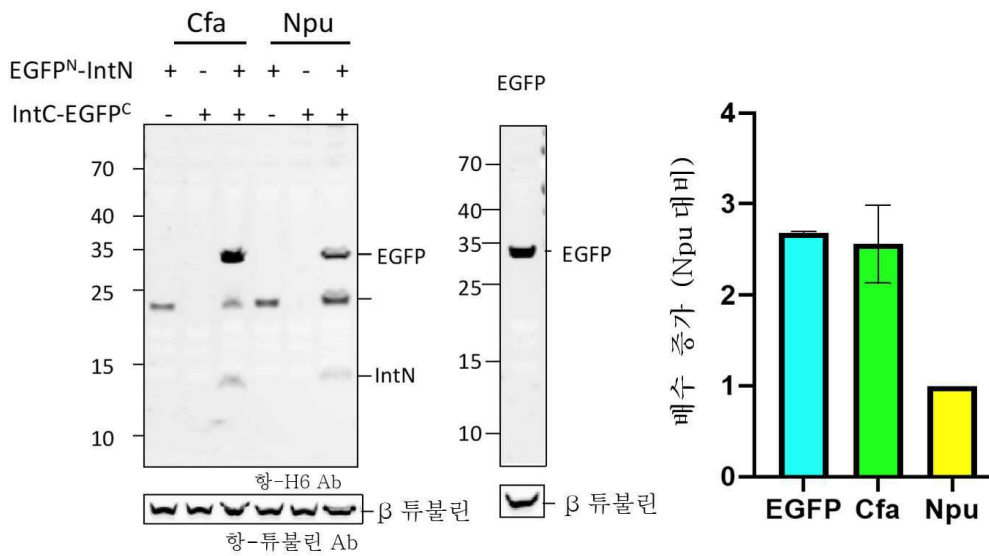
도면1a



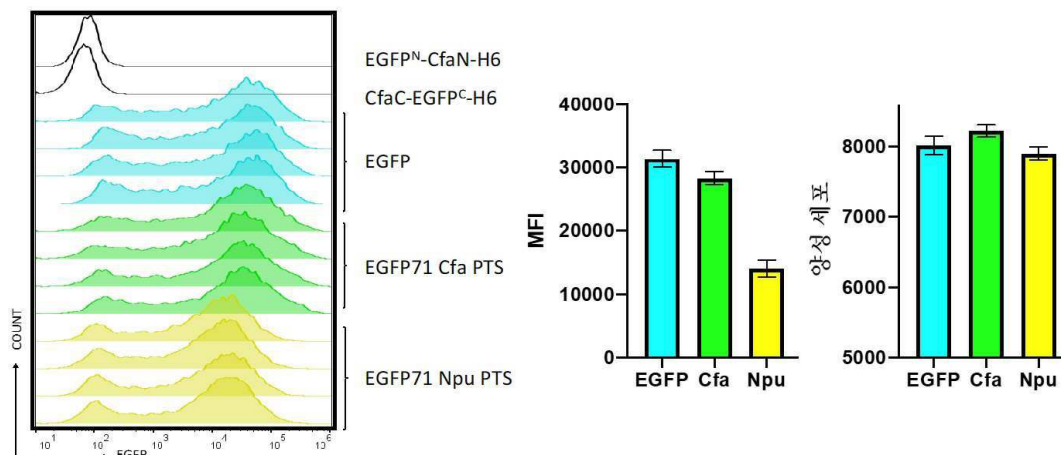
도면2



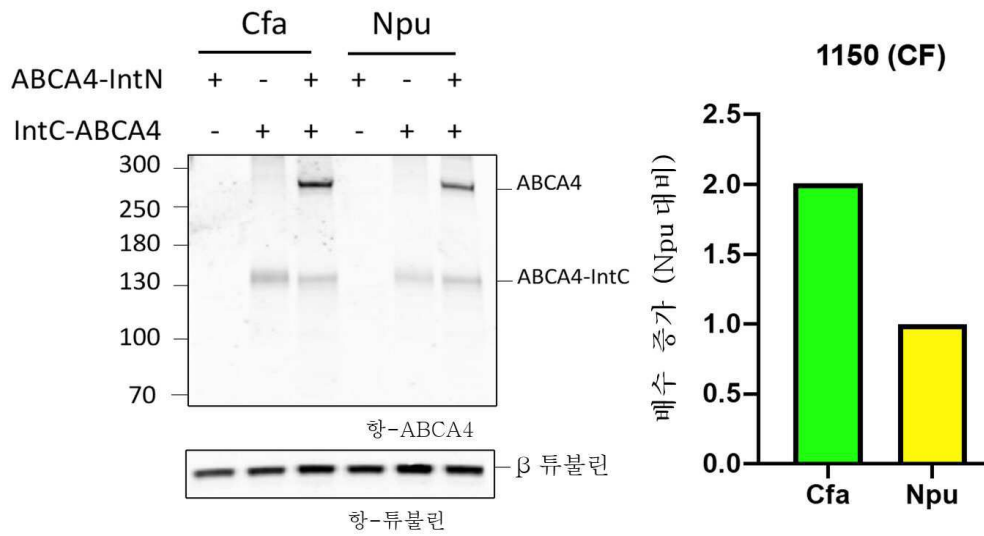
도면3



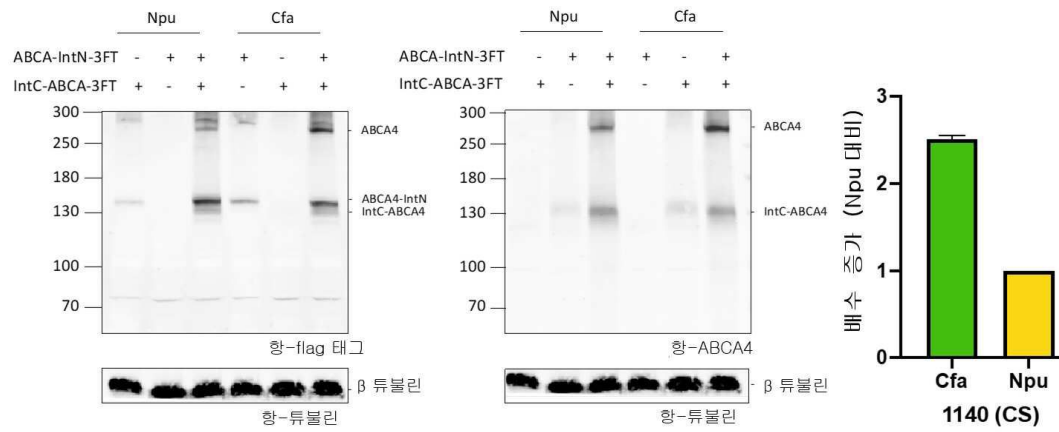
도면4



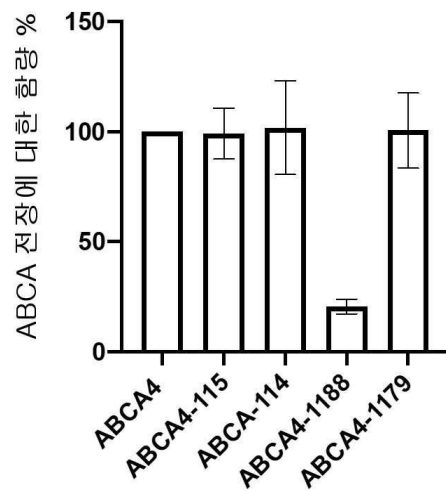
도면5



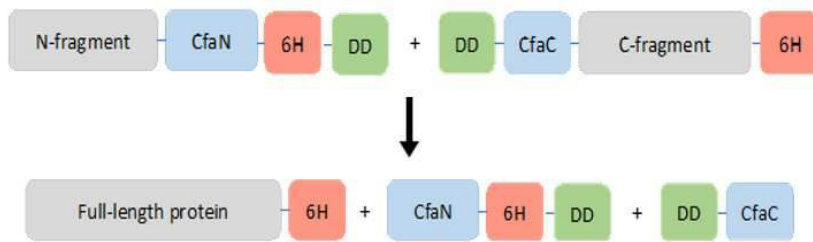
도면6



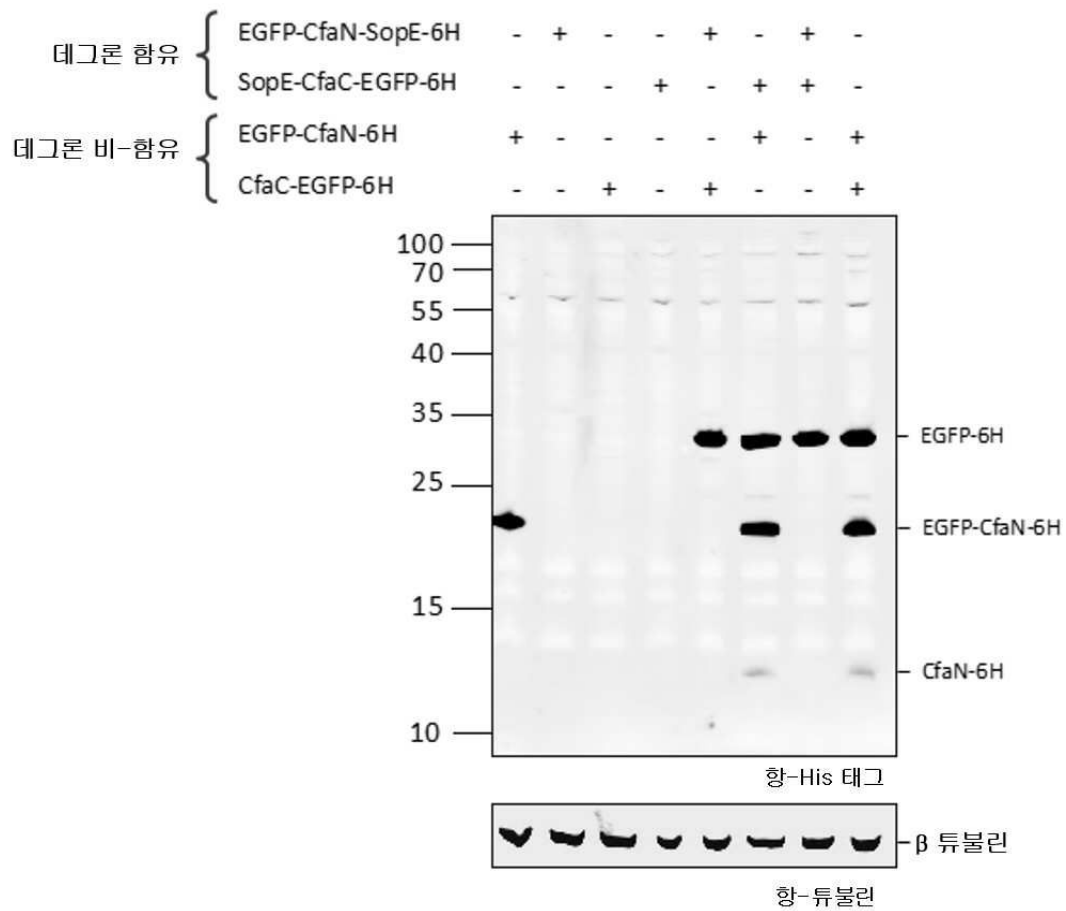
도면7



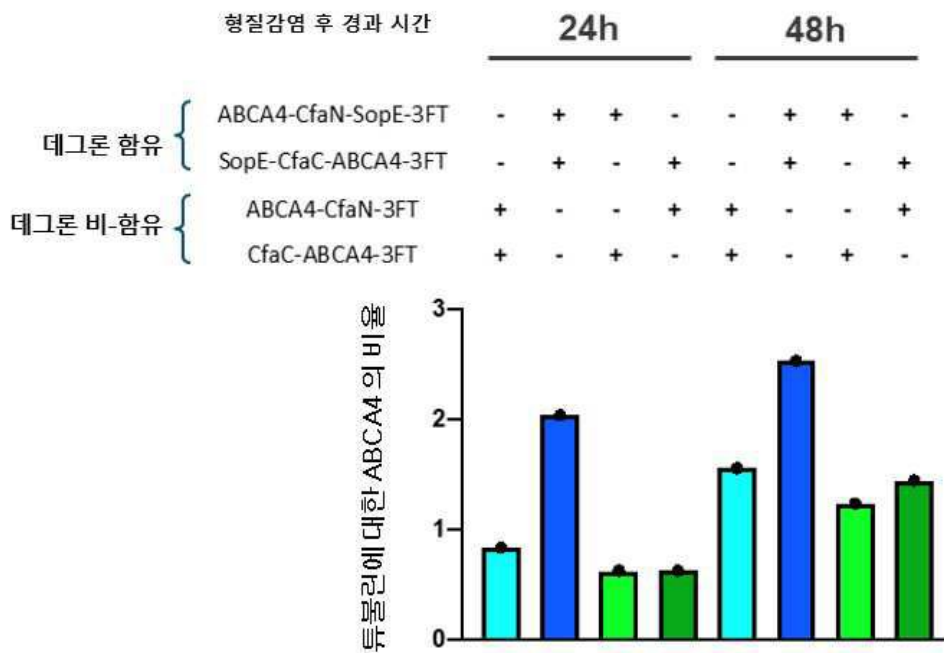
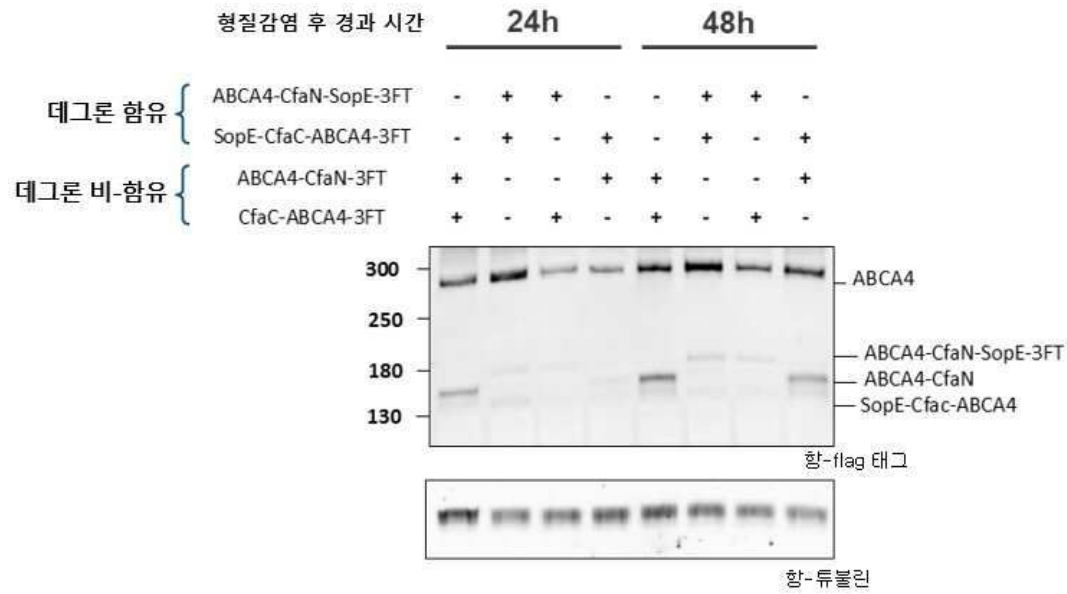
도면8



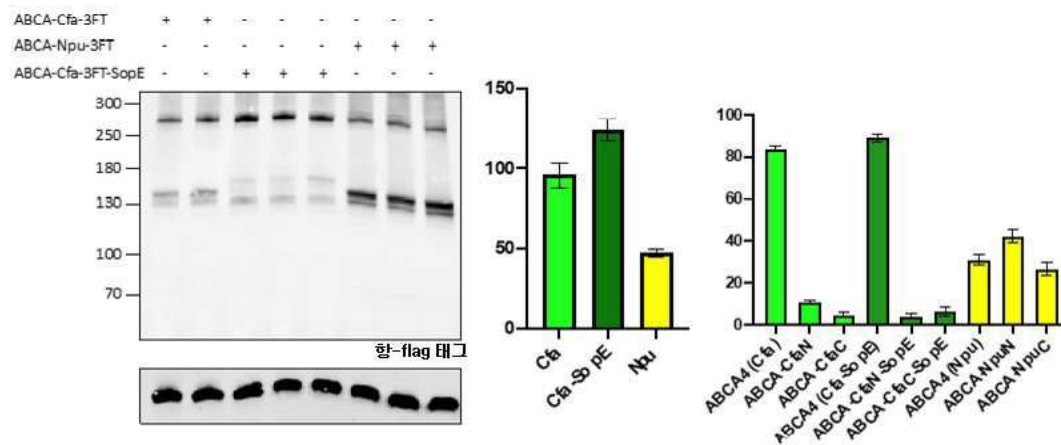
도면9



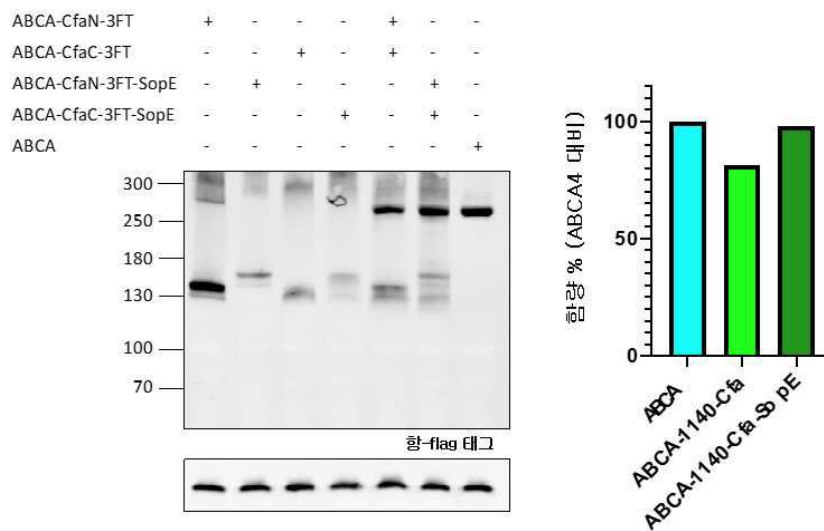
도면10



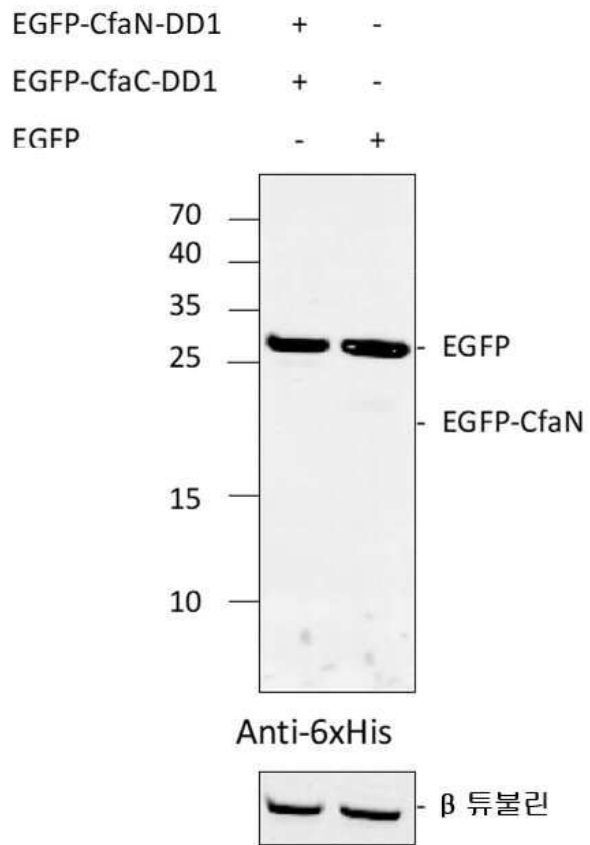
도면12



도면13



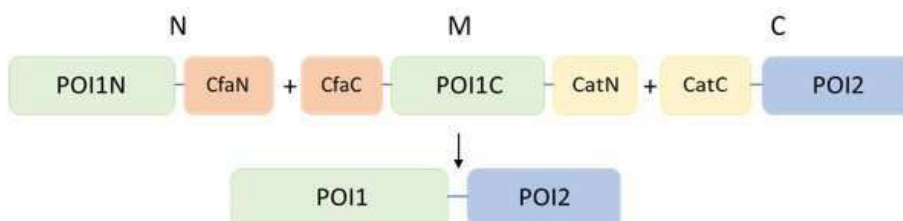
도면14



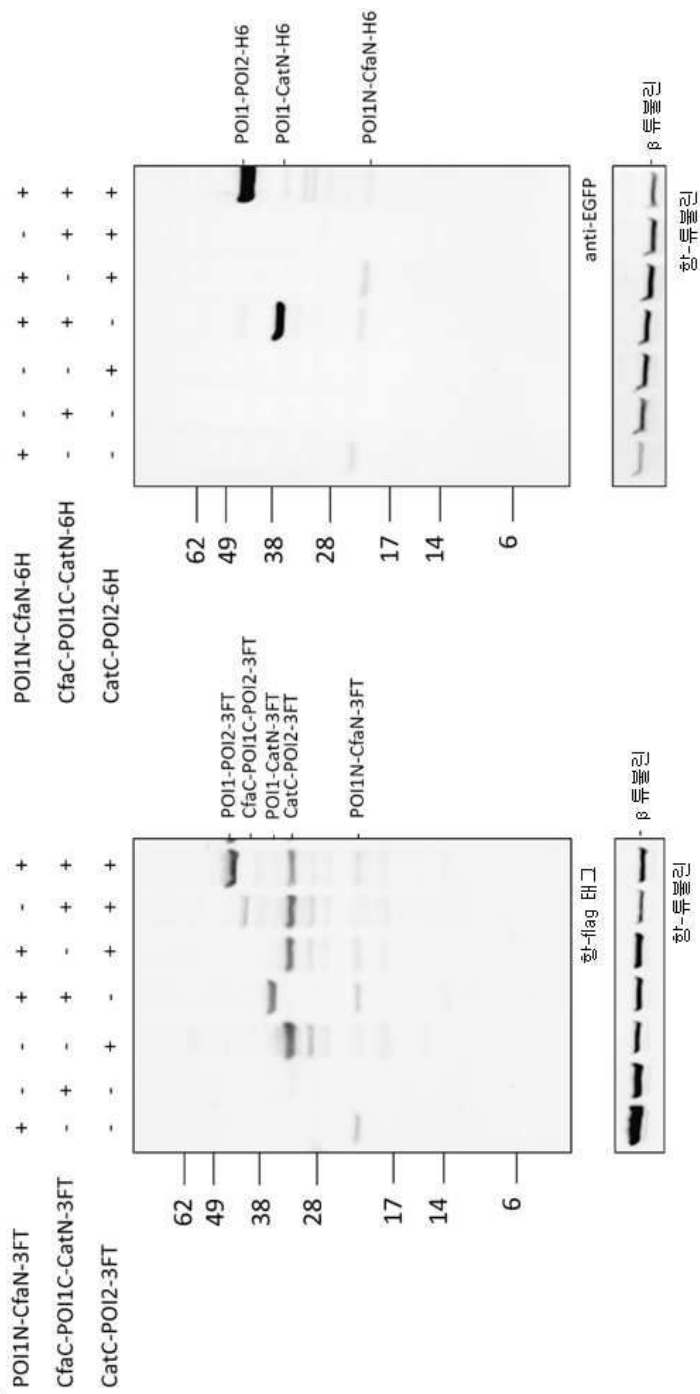
도면15

Sequence Comparison									
1	11	21	31	41	51	61	71	81	91
1	MGFVQYQULL	WQWTLIRK	QKRIFFVVLV	WELSLFVL	WLRWNPVS	RHECHFNKA	MFAGMLPL	QGIENWNP	CFASPTFEGS
101	ILARVYDTPQ	ELANAFESQ	HLGRINTEML	ILGQFMTLM	TUPERIAGS	IRINDIHOE	ETLFLFIRN	IGLDDGVVYL	LINQVTPQ
201	KDIACSEALL	EFFIIFQDPA	CAKTVYALC	SLSQCTIOMT	EDLIYANVF	KLDFVLPPL	LDERSQCIHL	DSWCCILSDM	SPRIQFIHR
301	RPLMONGCPE	TETKUMILS	DULCCYPEG	GSRLSENWY	EDNNYKAFIC	IDSTRKPTIY	SYDRRTSTC	NALIOSLESN	PLUKIAWRA
401	TPQSPAARAI	LKNANSTEE	LEHVRKLVEA	WEVGPQIYV	FFDNSTQDM	IDTULGNPTV	KDFLNGQLGE	EGITAEAIUN	FLXKGPRESQ
501	DIFENITRCL	LUNQYIECL	VLDKEFSYND	ETQUTORALS	LLEENFWAS	VVFEMVMT	SSFFVAVTK	INMEIDWEX	INKIKRYND
601	FRYTWGFEAY	IQDAVEGGIT	RSQVQAEAPV	GIYUQMPVP	CFYDDFMCI	LNRCFTPMAY	LAMIVGVCHT	VKGIVLEKEL	RLKETLKNQS
701	FLDSEFSMS	SIFILLITFM	HCZILVSDP	FILFLIFLAF	STATIMLCIL	LSIFFSASL	AAACSCVIIF	TLVIDPHILCF	AWQDPAEL
801	AFCEGTEYLV	RFEQIGLGLQ	WSNIGNSPTE	GDEFSFILSN	QMILDAAY	GLLAWVIDQV	FPDYSCTIPL	WYFILQESYV	LGSEGCSTRE
901	TEETEDPEEP	EGIEDSHFER	EHPGWPGVC	VKNIVK	EEPCGAPDRLN	ITEYENCITA	FLCHNGAGKT	TTLSEILTGUL	PPTSGVVLVG
1001	VKSLGKCTQ	WHILFHILTY	RESHLUTYAL	KGKQEEAQL	EEEMLEEDG	LHHERNEEAQ	DLSGGMRKL	SVALAFVGBA	KVILDEP'S
1101	WDLULRYRGS	RTIINGTHIM	DEADLQDRI	AIIAQGRLYC	GGPFLWRMC	FOTGLXITLV	DKMNIQQR	KSGEIGTCCG	CKKFGVTCFA
1201	LDCEVNEIUD	WVLEHVREAK	LVECIQOEIL	FLNNKNFTVH	DAVASJFEEL	HEFLADIOLS	SFCISDTDL	EIEFLKVTEDS	DSCPLFACCA
1301	HPCIGPRLA	QHTPCDSHC	SPAPAAHTE	QDEPESECP	QPOLNFTOL	VLDHVALIV	KRFQHTIRSH	KDFIAQVILP	AIFVFLAHL
1401	PALLHFW-Y	SQQYTFESMD	EPDSEQFTVL	ADVLNKPGE	ENRCLKEGML	PEYPCGNSTP	WKTPSYSPWI	TOLFQCLKMT	QVNPSPSCRC
1501	ECPEGAGGUP	PPQETQASTE	ILQOLTDRI	SDFLVATYFA	LIRSSKSNF	WYNEQRNGCI	SIGCKLPVVP	ITGEALVHEL	SDLGRHAWVS
1601	ELPOTLEHLE	TEDNKNWEN	NKGWIALVGF	LNVAIRNAIRL	ASLEKOSPE	EYGITVISOQ	UNLTKQLGE	ITVITTOVDA	VVAICVIFGM
1701	LIQEDYNKSK	HLQISICVSD	TIYVITNFWL	DIMAYSUSAC	LVGCIICIQ	KKATSEHML	DALVALLLIV	CHAVIDMAYD	ASPLEVDVST
1801	FIZINSALT	FILELFENNR	TILR	HAVILR	KLIVFPHFC	LGRGLIDLAL	SOAVTDVYAR	FGEHGANPF	HMDLIGNLF
1901	FLSQWIAEPT	KEPIVDEDDO	VAZERQRIIT	GSNTIDILSL	HELTKEVPT	SSPANDULCV	QVRPGSCFGL	LGVNGACKTI	IFKALTCDT
2001	KSILTNISEV	RQNRNGYCPQF	DAIDEJUTGR	EHLYVYARL	QVPAHILIV	ANWSIKSLGL	TVYADCLAGT	YSGGNKKULS	TATALIGCPP
2101	MDPOQRUMALW	WYIVGIIHES	RAYVLFQHM	HECEALOTRL	ALMWKATRC	MGTIQHLKGN	FQGGYIVTKEK	IKGHKDLILP	DLPVPEQFTQ
2201	SHVBLQFCV	SSSLABIFQ	LLLSHSDSL	IEEVSVTQTT	LDQVFNKAK	QOFESHULPL	HPSALCASQ	LQDUVYXHDG	DYXDHIDYK

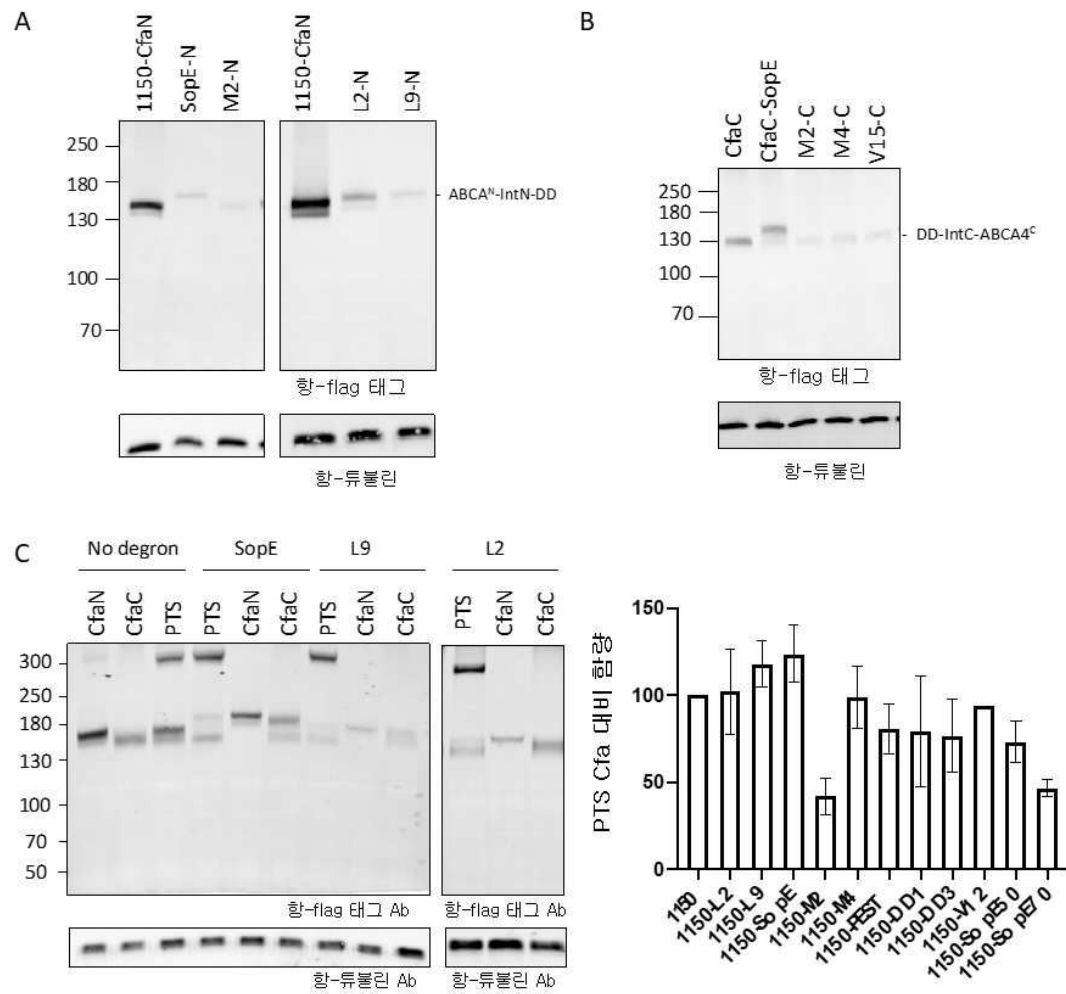
도면16



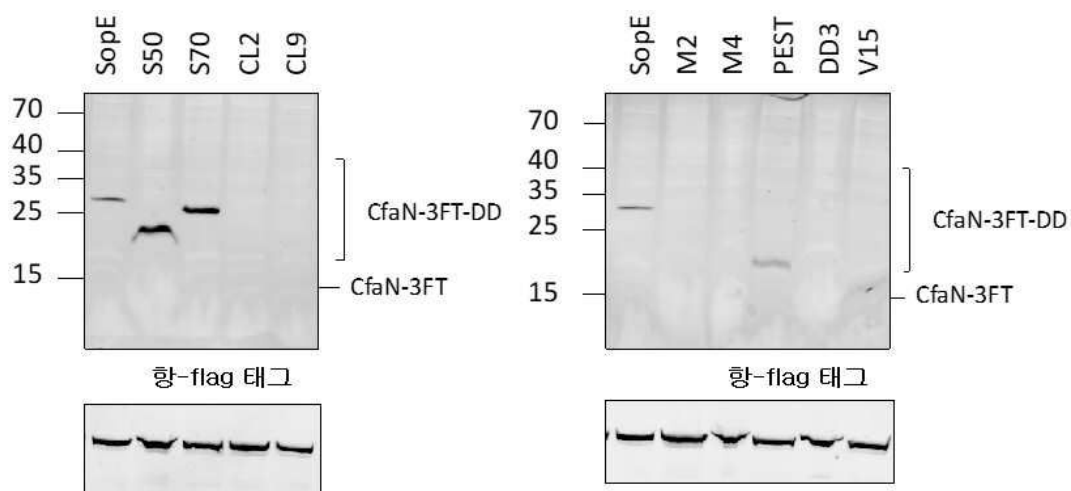
도면17



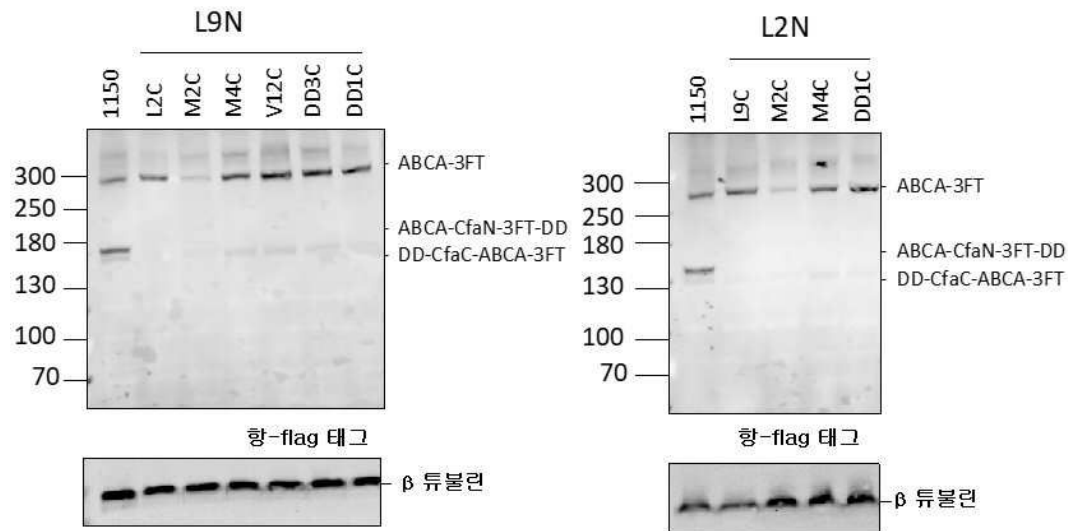
도면18



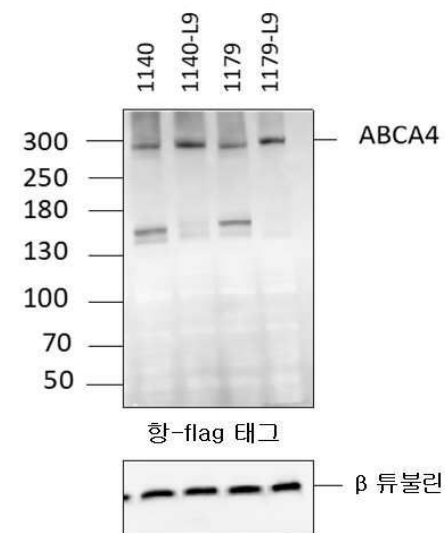
도면19



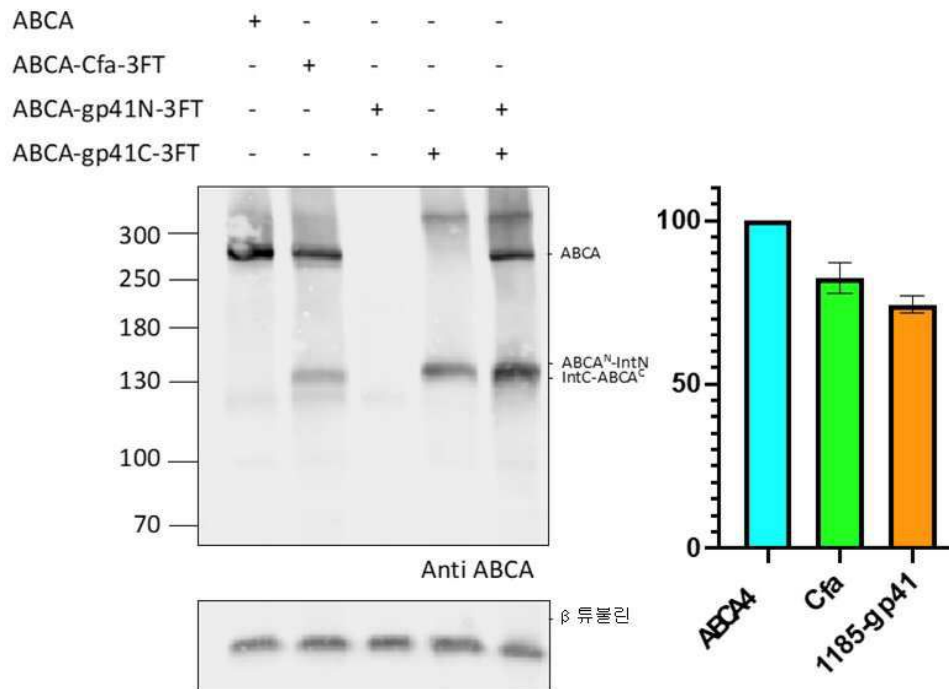
도면20



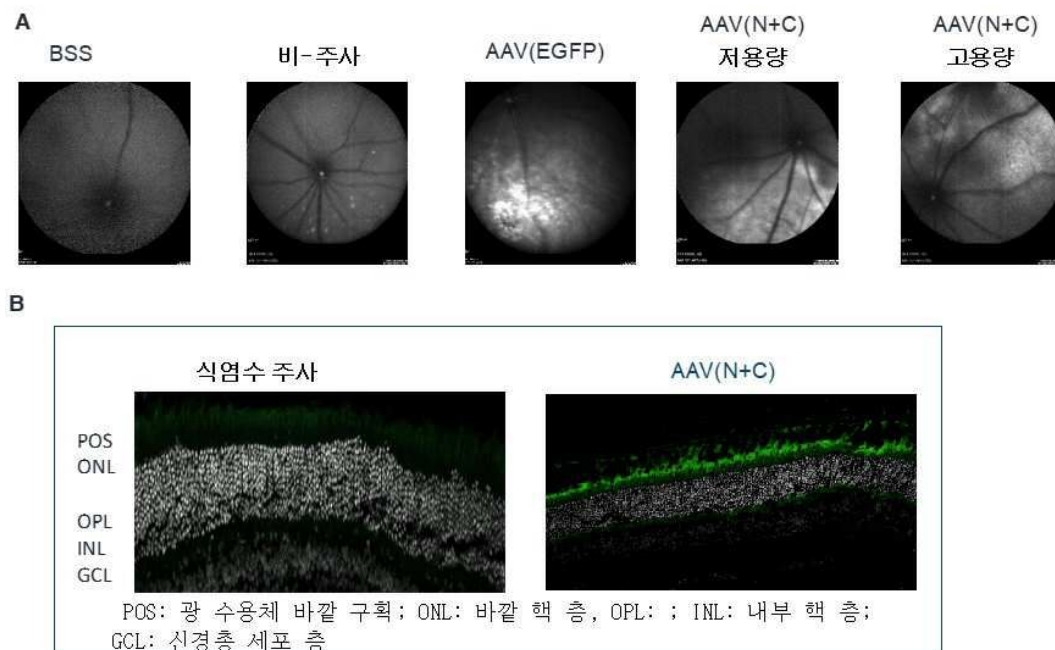
도면21



도면22



도면23



서열 목록

SEQUENCE LISTING

<110> SPLICEBIO S.L.

<120> Split inteins and their uses

<130> 905 992

<140> EPPCT/2021/058016

<141> 2020-03-26

<160> 108

<170> BiSSAP 1.3.6

<210> 1

<211> 1272

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaN-3FT

<400> 1

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
180 185 190

His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
195 200 205

Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
225 230 235 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
325 330 335

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

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu

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

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

 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605

 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915	920	925	
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro			
930	935	940	
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala			
945	950	955	960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu			
	965	970	975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg			
	980	985	990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys			
	995	1000	1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met			
	1010	1015	1020
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu			
1025	1030	1035	1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn			
	1045	1050	1055
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val			
	1060	1065	1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro			
	1075	1080	1085
Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu			
	1090	1095	1100
Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met			
1105	1110	1115	1120
Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly			
	1125	1130	1135
Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Leu Ser			
	1140	1145	1150
Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr Gly Phe Leu Pro Ile Gly			
	1155	1160	1165

Lys Ile Val Glu Glu Arg Ile Glu Cys Thr Val Tyr Thr Val Asp Lys
1170 1175 1180

Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala Gln Trp His Asn Arg Gly
1185 1190 1195 1200

Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu Asp Gly Ser Ile Ile Arg
1205 1210 1215

Ala Thr Lys Asp His Lys Phe Met Thr Thr Asp Gly Gln Met Leu Pro
1220 1225 1230

Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp Leu Lys Gln Val Asp Gly
1235 1240 1245

Leu Pro Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile
1250 1255 1260

Asp Tyr Lys Asp Asp Asp Asp Lys
1265 1270

<210> 2

<211> 1182

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaC-3FT

<400> 2

Met Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr
1 5 10 15

Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys Asn Gly Leu
20 25 30

Val Ala Ser Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg
35 40 45

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

Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp
65 70 75 80

Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met

85

90

95

Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile

100

105

110

Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg

115

120

125

Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu

130

135

140

Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe

145

150

155

160

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

165

170

175

Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly

180

185

190

Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser

195

200

205

Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro

210

215

220

Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln

225

230

235

240

His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser

245

250

255

His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe

260

265

270

Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro

275

280

285

Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe

290

295

300

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

305

310

315

320

Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp

325

330

335

Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser
 340 345 350

Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln
 355 360 365

Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr
 370 375 380

Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln
 385 390 395 400

Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn
 405 410 415

Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser
 420 425 430

Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser
 435 440 445

Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val
 450 455 460

Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro
 465 470 475 480

Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu
 485 490 495

Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His
 500 505 510

Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala
 515 520 525

Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile
 530 535 540

Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val
 545 550 555 560

Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser
 565 570 575

Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg

580 585 590
 Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr
 595 600 605

 Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val
 610 615 620
 Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala
 625 630 635 640
 Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu
 645 650 655
 Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe
 660 665 670

 Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe
 675 680 685
 Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu
 690 695 700
 Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu
 705 710 715 720
 Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala
 725 730 735

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

 Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile Ile
 805 810 815
 Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr Lys
 820 825 830

Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val Gly
835 840 845

Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly
850 855 860

Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser
865 870 875 880

Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu
885 890 895

Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu
900 905 910

Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly
915 920 925

Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser
930 935 940

Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly
945 950 955 960

Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro
965 970 975

Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala
980 985 990

Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg
995 1000 1005

Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys
1010 1015 1020

Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr
1025 1030 1035 1040

Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met
1045 1050 1055

Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val
1060 1065 1070

Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg

1075 1080 1085
 His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg
 1090 1095 1100
 Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu
 1105 1110 1115 1120

Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala
 1125 1130 1135
 Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala
 1140 1145 1150
 Gly Ala Ser Arg Gln Ala Gln Asp Asp Tyr Lys Asp His Asp Gly Asp
 1155 1160 1165
 Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys
 1170 1175 1180

<210> 3

<211> 1262

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaN-3FT

<400> 3

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
 1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95

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

340 345 350
Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365
Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys

370 375 380
Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400
Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415
Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430
Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met

435 440 445
Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460
Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480
Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495
Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

500 505 510
Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525
Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540
Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560
Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

565 570 575
Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro

625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser

690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile

755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp

820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

835 840 845
Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880
Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys

885 890 895
Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910
Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135

Arg Leu Tyr Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr

1140 1145 1150

Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr

1155 1160 1165

Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala

1170 1175 1180

Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu

1185 1190 1195 1200

Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr

1205 1210 1215

Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp

1220 1225 1230

Leu Lys Gln Val Asp Gly Leu Pro Asp Tyr Lys Asp His Asp Gly Asp

1235 1240 1245

Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys

1250 1255 1260

<210> 4

<211> 1192

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaCmut-3FT

<400> 4

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

1 5 10 15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30
 Val Ala Ser Asn Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe
 35 40 45
 Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln

 50 55 60
 Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly
 65 70 75 80
 Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln
 85 90 95
 Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His
 100 105 110
 Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe

 115 120 125
 Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe
 130 135 140
 Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly
 145 150 155 160
 Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp
 165 170 175
 Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu

 180 185 190
 Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly
 195 200 205
 Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala
 210 215 220
 His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln
 225 230 235 240
 Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu

 245 250 255
 Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala
 260 265 270

Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser
275 280 285

Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro
290 295 300

Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly
305 310 315 320

Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly
325 330 335

Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys
340 345 350

Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr
355 360 365

Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser
370 375 380

Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro
385 390 395 400

Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr
405 410 415

Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val
420 425 430

Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp
435 440 445

Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro
450 455 460

Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu
465 470 475 480

Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser
485 490 495

Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile
500 505 510

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

515 520 525
 Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg
 530 535 540
 Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu
 545 550 555 560
 Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp

 565 570 575
 Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala
 580 585 590
 Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His
 595 600 605
 Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn
 610 615 620
 Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val

 625 630 635 640
 Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn
 645 650 655
 Leu Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile
 660 665 670
 Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala
 675 680 685
 Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser

 690 695 700
 Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu
 705 710 715 720
 Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe
 725 730 735
 Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr
 740 745 750
 Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His

 755 760 765

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

1010 1015 1020
 His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met
 1025 1030 1035 1040
 Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser
 1045 1050 1055
 Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys
 1060 1065 1070
 Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly

1075 1080 1085
 Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln
 1090 1095 1100
 Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu
 1105 1110 1115 1120
 Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr
 1125 1130 1135
 Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser

1140 1145 1150
 His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala
 1155 1160 1165
 Gln Asp Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile
 1170 1175 1180
 Asp Tyr Lys Asp Asp Asp Asp Lys

1185 1190

<210> 5

<211> 1310

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1188-CfaN-3FT

<400> 5

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
 50 55 60
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125
 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala

 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser

260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr

515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu

530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr

545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

565 570 575

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

580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe

595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val

610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro

625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro

645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys

660 665 670

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

675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser

690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met

705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe

725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser

740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile

755 760 765
Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp

770 775 780
Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800
Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815
Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830
Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

835 840 845
Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880
Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895
Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

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

965 970 975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135

Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly

1140 1145 1150

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

1155 1160 1165

Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe

1170 1175 1180

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

1185 1190 1195 1200

Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr

1205 1210 1215

Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala

1220 1225 1230

Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu

1235 1240 1245

Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr

1250 1255 1260
 Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp
 1265 1270 1275 1280
 Leu Lys Gln Val Asp Gly Leu Pro Asp Tyr Lys Asp His Asp Gly Asp

1285 1290 1295
 Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys

1300 1305 1310

<210> 6

<211> 1144

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1188-Cfacmut-3FT

<400> 6

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

1 5 10 15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30

Val Ala Ser Asn Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln

35 40 45

Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His

50 55 60

Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe

65 70 75 80

Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe

85 90 95

Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly

100 105 110

Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp

115 120 125

Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu

130 135 140

Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly
145 150 155 160

Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala
165 170 175

His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln
180 185 190

Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu
195 200 205

Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala
210 215 220

Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser
225 230 235 240

Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro
245 250 255

Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly
260 265 270

Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly
275 280 285

Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys
290 295 300

Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr
305 310 315 320

Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser
325 330 335

Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro
340 345 350

Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr
355 360 365

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

Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp

385 390 395 400
 Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro
 405 410 415

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

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

 Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His
 545 550 555 560
 Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn
 565 570 575
 Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val
 580 585 590
 Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn
 595 600 605

 Leu Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala Val Ile
 610 615 620
 Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala
 625 630 635 640

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

885	890	895
Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr		
900	905	910
Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu		
915	920	925
Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp		
930	935	940
Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn		
945	950	955
Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser		
965	970	975
His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met		
980	985	990
Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser		
995	1000	1005
Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys		
1010	1015	1020
Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly		
1025	1030	1035
Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln		
1045	1050	1055
Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu		
1060	1065	1070
Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr		
1075	1080	1085
Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser		
1090	1095	1100
His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala		
1105	1110	1115
Gln Asp Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile		
1125	1130	1135

Asp Tyr Lys Asp Asp Asp Asp Lys

1140

<210> 7

<211> 2295

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-3FT

<400> 7

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180	185	190	
His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala			
195	200	205	
Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr			
210	215	220	
Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile			
225	230	235	240
Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val			
245	250	255	
Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser			
260	265	270	
Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile			
275	280	285	
His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met			
290	295	300	
Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser			
305	310	315	320
Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser			
325	330	335	
Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp			
340	345	350	
Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser			
355	360	365	
Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys			
370	375	380	
Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr			
385	390	395	400
Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser			
405	410	415	
Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu			
420	425	430	

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

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

 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605

 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670

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

675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720
 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735

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

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

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

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
1125 1130 1135

Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
1140 1145 1150

Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
1155 1160 1165

Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe

1170	1175	1180	
Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val			
1185	1190	1195	1200
Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val			
	1205	1210	1215
Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu			
	1220	1225	1230
Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg			
	1235	1240	1245
Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile			
	1250	1255	1260
Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser			
1265	1270	1275	1280
Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn			
	1285	1290	1295
Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln			
	1300	1305	1310
Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His			
	1315	1320	1325
Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu			
	1330	1335	1340
Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val			
1345	1350	1355	1360
Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala Gln			
	1365	1370	1375
Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile			
	1380	1385	1390
Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro Trp			
	1395	1400	1405
Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser			
1410	1415	1420	

Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe
1425 1430 1435 1440

Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly
1445 1450 1455

Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln
1460 1465 1470

Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys
1475 1480 1485

Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu
1490 1495 1500

Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu
1505 1510 1515 1520

Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val Lys
1525 1530 1535

Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val
1540 1545 1550

Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val
1555 1560 1565

Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly
1570 1575 1580

Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys
1585 1590 1595 1600

Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys
1605 1610 1615

Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu Asn
1620 1625 1630

Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser
1635 1640 1645

Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr
1650 1655 1660

Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala

1665 1670 1675 1680
 Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser
 1685 1690 1695

 Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu
 1700 1705 1710
 Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe
 1715 1720 1725
 Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly
 1730 1735 1740
 Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu
 1745 1750 1755 1760

 Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro
 1765 1770 1775
 Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr
 1780 1785 1790
 Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala
 1795 1800 1805
 Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg
 1810 1815 1820

 Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys
 1825 1830 1835 1840
 Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp
 1845 1850 1855
 Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp
 1860 1865 1870
 Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val
 1875 1880 1885

 Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln
 1890 1895 1900
 Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp
 1905 1910 1915 1920

Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp

1925

1930

1935

Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser

1940

1945

1950

Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe

1955

1960

1965

Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met

1970

1975

1980

Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly

1985

1990

1995

2000

Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr

2005

2010

2015

Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His

2020

2025

2030

Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu

2035

2040

2045

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

2050

2055

2060

Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser

2065

2070

2075

2080

Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu

2085

2090

2095

Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val

2100

2105

2110

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

2115

2120

2125

Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val

2130

2135

2140

Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys

2145

2150

2155

2160

Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp

2165 2170 2175
Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn

2180 2185 2190
Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe
2195 2200 2205

Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser
2210 2215 2220

His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr
2225 2230 2235 2240

Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His
2245 2250 2255

Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln
2260 2265 2270

Asp Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
2275 2280 2285

Tyr Lys Asp Asp Asp Asp Lys
2290 2295

<210> 8

<211> 1273

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-NpuN-3FT

<400> 8

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95

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

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

 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285

 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser

305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350

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

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

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

 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly

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

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

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
 930 935 940
 Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
 945 950 955 960
 Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
 965 970 975
 Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
 980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
 995 1000 1005
 Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
 1010 1015 1020
 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
1075 1080 1085
Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
1090 1095 1100
Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
1125 1130 1135
Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Leu Ser
1140 1145 1150
Tyr Glu Thr Glu Ile Leu Thr Val Glu Tyr Gly Leu Leu Pro Ile Gly
1155 1160 1165
Lys Ile Val Glu Lys Arg Ile Glu Cys Thr Val Tyr Ser Val Asp Asn
1170 1175 1180

Asn Gly Asn Ile Tyr Thr Gln Pro Val Ala Gln Trp His Asp Arg Gly
1185 1190 1195 1200
Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu Asp Gly Ser Leu Ile Arg
1205 1210 1215
Ala Thr Lys Asp His Lys Phe Met Thr Val Asp Gly Gln Met Leu Pro
1220 1225 1230
Ile Asp Glu Ile Phe Glu Arg Glu Leu Asp Leu Met Arg Val Asp Asn
1235 1240 1245

Leu Pro Asn Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp
1250 1255 1260
Ile Asp Tyr Lys Asp Asp Asp Asp Lys
1265 1270

<210> 9

<211> 1182

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-NpuC-3FT

<400> 9

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr

1 5 10 15

Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe

20 25 30

Ile Ala Ser Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg

35 40 45

Lys Met Lys Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys

50 55 60

Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp

65 70 75 80

Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met

85 90 95

Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile

100 105 110

Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg

115 120 125

Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu

130 135 140

Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe

145 150 155 160

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

165 170 175

Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly

180 185 190

Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser

195 200 205

Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro

210 215 220

Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln
 225 230 235 240
 His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser
 245 250 255
 His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe
 260 265 270
 Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro

 275 280 285
 Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe
 290 295 300
 Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val
 305 310 315 320
 Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp
 325 330 335
 Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser

 340 345 350
 Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln
 355 360 365
 Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr
 370 375 380
 Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln
 385 390 395 400
 Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn

 405 410 415
 Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser
 420 425 430
 Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser
 435 440 445
 Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val
 450 455 460
 Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro

465 470 475 480
 Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu
 485 490 495
 Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His
 500 505 510
 Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala
 515 520 525
 Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile

 530 535 540
 Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val
 545 550 555 560
 Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser
 565 570 575
 Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg
 580 585 590
 Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr

 595 600 605
 Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val
 610 615 620
 Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala
 625 630 635 640
 Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu
 645 650 655
 Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe

 660 665 670
 Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe
 675 680 685
 Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu
 690 695 700
 Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu
 705 710 715 720

Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala

725 730 735

Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His

740 745 750

Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala

755 760 765

Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val Gln

770 775 780

Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro

785 790 795 800

Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile Ile

805 810 815

Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr Lys

820 825 830

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

835 840 845

Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly

850 855 860

Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser

865 870 875 880

Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu

885 890 895

Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu

900 905 910

Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly

915 920 925

Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser

930 935 940

Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly

945 950 955 960

Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro

965 970 975
 Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala

 980 985 990
 Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg
 995 1000 1005
 Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys
 1010 1015 1020
 Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr
 1025 1030 1035 1040
 Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met

 1045 1050 1055
 Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val
 1060 1065 1070
 Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg
 1075 1080 1085
 His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg
 1090 1095 1100
 Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu

 1105 1110 1115 1120
 Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala
 1125 1130 1135
 Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala
 1140 1145 1150
 Gly Ala Ser Arg Gln Ala Gln Asp Asp Tyr Lys Asp His Asp Gly Asp
 1155 1160 1165
 Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys

 1170 1175 1180
 <210> 10
 <211> 1263
 <212> PRT
 <213> Artificial Sequence

<220>

<223> ABCA4-1140-NpuN-3FT

<400> 10

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile

225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300

 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365

 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430

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

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe

725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750

 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765
 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780
 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
 805 810 815

 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
 820 825 830
 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
 835 840 845
 Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
 850 855 860
 Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
 865 870 875 880

 Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
 885 890 895
 Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
 900 905 910
 Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
 915 920 925
 Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
 930 935 940

 Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
 945 950 955 960
 Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
 965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135

Arg Leu Tyr Cys Leu Ser Tyr Glu Thr Glu Ile Leu Thr Val Glu Tyr

1140 1145 1150

Gly Leu Leu Pro Ile Gly Lys Ile Val Glu Lys Arg Ile Glu Cys Thr

1155 1160 1165

Val Tyr Ser Val Asp Asn Asn Gly Asn Ile Tyr Thr Gln Pro Val Ala

1170 1175 1180

Gln Trp His Asp Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu

1185 1190 1195 1200

Asp Gly Ser Leu Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Val

1205 1210 1215

Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Glu Leu Asp

1220 1225 1230
 Leu Met Arg Val Asp Asn Leu Pro Asn Asp Tyr Lys Asp His Asp Gly
 1235 1240 1245
 Asp Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Lys
 1250 1255 1260

<210> 11

<211> 1192

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-NpuC-3FT

<400> 11

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr
 1 5 10 15
 Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe
 20 25 30
 Ile Ala Ser Asn Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe
 35 40 45
 Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln

50 55 60
 Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly
 65 70 75 80
 Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln
 85 90 95
 Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His
 100 105 110
 Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe

115 120 125
 Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe
 130 135 140
 Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly
 145 150 155 160

Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp
165 170 175
Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu
180 185 190
Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly
195 200 205
Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala
210 215 220
His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln
225 230 235 240
Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu
245 250 255
Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala
260 265 270
Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser
275 280 285
Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro
290 295 300
Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly
305 310 315 320
Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly
325 330 335
Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys
340 345 350
Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr
355 360 365
Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser
370 375 380
Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro
385 390 395 400
Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr

405 410 415
 Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val
 420 425 430
 Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp

 435 440 445
 Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro
 450 455 460
 Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu
 465 470 475 480
 Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser
 485 490 495
 Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile

 500 505 510
 Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu
 515 520 525
 Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg
 530 535 540
 Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu
 545 550 555 560
 Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp

 565 570 575
 Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala
 580 585 590
 Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His
 595 600 605
 Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn
 610 615 620
 Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val

 625 630 635 640
 Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn
 645 650 655

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

900 905 910
Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu

915 920 925
His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile

930 935 940
Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr

945 950 955 960
Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu

965 970 975
Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp

980 985 990
Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn

995 1000 1005
Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser

1010 1015 1020
His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met

1025 1030 1035 1040
Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser

1045 1050 1055
Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys

1060 1065 1070
Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly

1075 1080 1085
Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln

1090 1095 1100
Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu

1105 1110 1115 1120
Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr

1125 1130 1135
Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser

1140 1145 1150

His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala

1155

1160

1165

Gln Asp Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile

1170

1175

1180

Asp Tyr Lys Asp Asp Asp Asp Lys

1185

1190

<210> 12

<211> 1311

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1188-NpuN-3FT

<400> 12

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1

5

10

15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20

25

30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35

40

45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50

55

60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65

70

75

80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85

90

95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100

105

110

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

115

120

125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130

135

140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190

 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320

 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380

 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro

645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700

 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720
 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750
 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765

 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780
 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
 805 810 815
 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
 820 825 830

 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
 835 840 845
 Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
 850 855 860
 Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
 865 870 875 880
 Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
 885 890 895

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

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

 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055
 Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
 1060 1065 1070
 Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
 1075 1080 1085

 Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
 1090 1095 1100
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly

1140 1145 1150
 Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
 1155 1160 1165
 Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe
 1170 1175 1180
 Ser Thr Thr Cys Leu Ser Tyr Glu Thr Glu Ile Leu Thr Val Glu Tyr
 1185 1190 1195 1200
 Gly Leu Leu Pro Ile Gly Lys Ile Val Glu Lys Arg Ile Glu Cys Thr
 1205 1210 1215

 Val Tyr Ser Val Asp Asn Asn Gly Asn Ile Tyr Thr Gln Pro Val Ala
 1220 1225 1230
 Gln Trp His Asp Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu
 1235 1240 1245
 Asp Gly Ser Leu Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Val
 1250 1255 1260
 Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Glu Leu Asp
 1265 1270 1275 1280

 Leu Met Arg Val Asp Asn Leu Pro Asn Asp Tyr Lys Asp His Asp Gly
 1285 1290 1295
 Asp Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys
 1300 1305 1310

 <210> 13
 <211> 1144
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> ABCA4-1188-NpuC-3FT
 <400> 13
 Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr
 1 5 10 15
 Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe

20 25 30
 Ile Ala Ser Asn Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln
 35 40 45
 Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His
 50 55 60
 Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe
 65 70 75 80
 Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe

 85 90 95
 Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly
 100 105 110
 Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp
 115 120 125
 Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu
 130 135 140
 Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly

 145 150 155 160
 Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala
 165 170 175
 His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln
 180 185 190
 Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu
 195 200 205
 Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala

 210 215 220
 Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser
 225 230 235 240
 Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro
 245 250 255
 Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly
 260 265 270

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

275 280 285

Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys

290 295 300

Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr

305 310 315 320

Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser

325 330 335

Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro

340 345 350

Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr

355 360 365

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

370 375 380

Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp

385 390 395 400

Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro

405 410 415

Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu

420 425 430

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

435 440 445

Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile

450 455 460

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

465 470 475 480

Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg

485 490 495

Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu

500 505 510

Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp

515 520 525
Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala

530 535 540
Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His

545 550 555 560
Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn

565 570 575
Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val

580 585 590
Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn

595 600 605
Leu Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile

610 615 620
Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala

625 630 635 640
Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser

645 650 655
Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu

660 665 670
Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe

675 680 685
Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr

690 695 700
Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His

705 710 715 720
Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val

725 730 735
Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser

740 745 750
Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp

755 760 765

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

1010 1015 1020
 Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly
 1025 1030 1035 1040
 Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln

 1045 1050 1055
 Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu
 1060 1065 1070
 Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr
 1075 1080 1085
 Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser
 1090 1095 1100
 His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala

 1105 1110 1115 1120
 Gln Asp Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile
 1125 1130 1135
 Asp Tyr Lys Asp Asp Asp Asp Lys
 1140
 <210> 14
 <211> 179
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> EGFP-71- CfaN
 <400> 14
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly

 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60

Leu Thr Tyr Gly Val Gln Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr
 65 70 75 80
 Val Glu Tyr Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile

85 90 95
 Glu Cys Thr Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln
 100 105 110
 Pro Ile Ala Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr
 115 120 125
 Cys Leu Glu Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe
 130 135 140
 Met Thr Thr Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg

145 150 155 160
 Gly Leu Asp Leu Lys Gln Val Asp Gly Leu Pro Gly His His His His
 165 170 175
 His His Gly

<210> 15

<211> 213

<212> PRT

<213> Artificial Sequence

<220>

<223> EGFP-71- CfaC

<400> 15

Met Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr
 1 5 10 15
 Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30
 Val Ala Ser Asn Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His
 35 40 45
 Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr
 50 55 60
 Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys

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

85 90 95

Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr

100 105 110

Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile

115 120 125
Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln

130 135 140

Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val

Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys

Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr

180 185 190

Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Gly His His

His His His His Gly

210

<210

$$> 16$$

<211> 247

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> EGFP

<400> 16

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1 5 10 15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly

20 25 30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr

50

55

60

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

65

70

75

80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu

85

90

95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu

100

105

110

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

115

120

125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr

130

135

140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn

145

150

155

160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser

165

170

175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly

180

185

190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu

195

200

205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe

210

215

220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Gly

225

230

235

240

His His His His His His Gly

245

<210> 17

<211> 180

<212> PRT

<213> Artificial Sequence

<220>

<223> EGFP-71- NpuN

<400> 17

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

1 5 10 15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly

20 25 30

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

35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr

50 55 60

Leu Thr Tyr Gly Val Gln Cys Leu Ser Tyr Glu Thr Glu Ile Leu Thr

65 70 75 80

Val Glu Tyr Gly Leu Leu Pro Ile Gly Lys Ile Val Glu Lys Arg Ile

85 90 95

Glu Cys Thr Val Tyr Ser Val Asp Asn Asn Gly Asn Ile Tyr Thr Gln

100 105 110

Pro Val Ala Gln Trp His Asp Arg Gly Glu Gln Glu Val Phe Glu Tyr

115 120 125

Cys Leu Glu Asp Gly Ser Leu Ile Arg Ala Thr Lys Asp His Lys Phe

130 135 140

Met Thr Val Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg

145 150 155 160

Glu Leu Asp Leu Met Arg Val Asp Asn Leu Pro Asn Gly His His His

165 170 175

His His His Gly

180

<210> 18

<

211> 213

<212> PRT

<213> Artificial Sequence

<220>

<223> EGFP-71- NpuC

<400> 18

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr

1 5 10 15

Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe

20 25 30

Ile Ala Ser Asn Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His

35 40 45

Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr

50 55 60

Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys

65 70 75 80

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

85 90 95

Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr

100 105 110

Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile

115 120 125

Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln

130 135 140

Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val

145 150 155 160

Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys

165 170 175

Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr

180 185 190

Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Gly His His

195 200 205

His His His His Gly

210

<210> 19

<211> 1371

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaN-3FT-SopE

<400> 19

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile

 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp

 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser

 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn

465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn

485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr

515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu

530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr

545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

565 570 575

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

580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe

595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val

610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro

625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro

645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys

660 665 670

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

675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser

690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met

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

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu

965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135

Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Leu Ser

1140 1145 1150

Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr Gly Phe Leu Pro Ile Gly

1155 1160 1165

Lys Ile Val Glu Glu Arg Ile Glu Cys Thr Val Tyr Thr Val Asp Lys

1170 1175 1180

Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala Gln Trp His Asn Arg Gly

1185 1190 1195 1200

Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu Asp Gly Ser Ile Ile Arg

1205 1210 1215
Ala Thr Lys Asp His Lys Phe Met Thr Thr Asp Gly Gln Met Leu Pro
1220 1225 1230
Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp Leu Lys Gln Val Asp Gly

1235 1240 1245
Leu Pro Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile

1250 1255 1260
Asp Tyr Lys Asp Asp Asp Asp Lys Thr Lys Ile Thr Leu Ser Pro Gln
1265 1270 1275 1280
Asn Phe Arg Ile Gln Lys Gln Glu Thr Thr Leu Leu Lys Glu Lys Ser
1285 1290 1295
Thr Glu Lys Asn Ser Leu Ala Lys Ser Ile Leu Ala Val Lys Asn His

1300 1305 1310
Phe Ile Glu Leu Arg Ser Lys Leu Ser Glu Arg Phe Ile Ser His Lys

1315 1320 1325
Asn Thr Glu Ser Ser Ala Thr His Phe His Arg Gly Ser Ala Ser Glu

1330 1335 1340
Gly Arg Ala Val Leu Thr Asn Lys Val Val Lys Asp Phe Met Leu Gln
1345 1350 1355 1360

Thr Leu Asn Asp Ile Asp Ile Arg Gly Ser Ala

1365 1370

<210> 20

<211> 1281

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaC-3FT-SopE

<400> 20

Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln

1 5 10 15

Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala

20 25 30

Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys
 35 40 45

 Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
 50 55 60
 His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn
 65 70 75 80
 Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile
 85 90 95
 Arg Gly Ser Ala Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln
 100 105 110

 Asn Val Tyr Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys
 115 120 125
 Asn Gly Leu Val Ala Ser Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr
 130 135 140
 Leu Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu
 145 150 155 160
 Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala
 165 170 175

 His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn
 180 185 190
 Glu Leu Met Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val
 195 200 205
 Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe
 210 215 220
 Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu
 225 230 235 240

 Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu
 245 250 255
 Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe
 260 265 270
 Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro

275 280 285
 Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn
 290 295 300

 Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro
 305 310 315 320
 Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu
 325 330 335
 Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr
 340 345 350
 Ile Arg Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr
 355 360 365

 Phe Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly
 370 375 380
 Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr
 385 390 395 400
 Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu
 405 410 415
 Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys
 420 425 430

 Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys
 435 440 445
 Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys
 450 455 460
 Trp Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu
 465 470 475 480
 Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro
 485 490 495

 Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr
 500 505 510
 Asp Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile
 515 520 525

Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly
 530 535 540
 Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu
 545 550 555 560

Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser
 565 570 575
 Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu
 580 585 590
 Lys His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys
 595 600 605
 Gly Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile
 610 615 620

Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile
 625 630 635 640
 Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu
 645 650 655
 Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val
 660 665 670
 Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile
 675 680 685

Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val
 690 695 700
 Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn
 705 710 715 720
 Tyr Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln
 725 730 735
 Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu
 740 745 750

Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser
 755 760 765
 Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala

770 775 780
 Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu
 785 790 795 800
 Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg
 805 810 815

 Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile
 820 825 830
 Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly
 835 840 845
 Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn
 850 855 860
 Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu
 865 870 875 880

 Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr
 885 890 895
 Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln
 900 905 910
 Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu
 915 920 925
 Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu
 930 935 940

 Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn
 945 950 955 960
 Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr
 965 970 975
 Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn
 980 985 990
 Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala
 995 1000 1005

 Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg
 1010 1015 1020

Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser
 1025 1030 1035 1040
 Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr
 1045 1050 1055
 Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile
 1060 1065 1070

 Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp
 1075 1080 1085
 Pro Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg
 1090 1095 1100
 Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu
 1105 1110 1115 1120
 Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys
 1125 1130 1135

 Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile
 1140 1145 1150
 Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu
 1155 1160 1165
 Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln
 1170 1175 1180
 Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser
 1185 1190 1195 1200

 Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu
 1205 1210 1215
 Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val
 1220 1225 1230
 Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro
 1235 1240 1245
 Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp Asp Tyr Lys Asp His
 1250 1255 1260

 Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp

1265 1270 1275 1280

Lys

<210> 21

<211> 1361

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaN-3FT-SopE

<400> 21

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
180 185 190

His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
195 200 205

Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
225 230 235 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
325 330 335

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

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu

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

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

 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605

 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915	920	925	
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro			
930	935	940	
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala			
945	950	955	960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu			
	965	970	975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg			
	980	985	990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys			
	995	1000	1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met			
	1010	1015	1020
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu			
1025	1030	1035	1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn			
	1045	1050	1055
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val			
	1060	1065	1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro			
	1075	1080	1085
Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu			
	1090	1095	1100
Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met			
1105	1110	1115	1120
Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly			
	1125	1130	1135
Arg Leu Tyr Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr			
	1140	1145	1150
Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr			
	1155	1160	1165

Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala
1170 1175 1180

Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu
1185 1190 1195 1200

Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr
1205 1210 1215

Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp
1220 1225 1230

Leu Lys Gln Val Asp Gly Leu Pro Asp Tyr Lys Asp His Asp Gly Asp
1235 1240 1245

Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys Thr Lys
1250 1255 1260

Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln Glu Thr Thr
1265 1270 1275 1280

Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala Lys Ser Ile
1285 1290 1295

Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys Leu Ser Glu
1300 1305 1310

Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr His Phe His
1315 1320 1325

Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn Lys Val Val
1330 1335 1340

Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile Arg Gly Ser
1345 1350 1355 1360

Ala

<210> 22

<211> 1291

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaC-3FT-SopE

<400> 22

```

Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln
1           5           10           15
Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala
           20           25           30
Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys
           35           40           45
Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
           50           55           60

His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn
65           70           75           80
Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile
           85           90           95
Arg Gly Ser Ala Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln
           100          105          110
Asn Val Tyr Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys
           115          120          125

Asn Gly Leu Val Ala Ser Asn Cys Ser Gly Thr Pro Leu Phe Leu Lys
           130          135          140
Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys
145          150          155          160
Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser
           165          170          175
Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr
           180          185          190

Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val
           195          200          205
Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu
           210          215          220
Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala

```

225 230 235 240
 Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser
 245 250 255

 Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val
 260 265 270
 Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln
 275 280 285
 Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu
 290 295 300
 Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala
 305 310 315 320

 Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro
 325 330 335
 Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln
 340 345 350
 Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp
 355 360 365
 Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu
 370 375 380

 Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr
 385 390 395 400
 Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp
 405 410 415
 Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn
 420 425 430
 Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu
 435 440 445

 Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro
 450 455 460
 Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro
 465 470 475 480

Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro
485 490 495

Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln
500 505 510

Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp
515 520 525

Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser
530 535 540

Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly
545 550 555 560

Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu
565 570 575

Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg
580 585 590

Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu
595 600 605

Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val
610 615 620

Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro
625 630 635 640

Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro
645 650 655

Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr
660 665 670

Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe
675 680 685

Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys
690 695 700

Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp
705 710 715 720

Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly

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

Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala

980 985 990

Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln

995 1000 1005

Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr

1010 1015 1020

Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala

1025 1030 1035 1040

Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu

1045 1050 1055

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

1060 1065 1070

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

1075 1080 1085

Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met

1090 1095 1100

Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val

1105 1110 1115 1120

Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu

1125 1130 1135

Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His

1140 1145 1150

Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys

1155 1160 1165

Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe

1170 1175 1180

Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn

1185 1190 1195 1200

Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln

1205 1210 1215

Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val

1220 1225 1230
 Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln
 1235 1240 1245
 Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser
 1250 1255 1260
 Arg Gln Ala Gln Asp Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp
 1265 1270 1275 1280

His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys

1285 1290

<210> 23

<211> 278

<212> PRT

<213> Artificial Sequence

<220>

<223> EGFP-71-CfaN-SopE

<400> 23

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile

35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr

50 55 60

Leu Thr Tyr Gly Val Gln Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr

65 70 75 80

Val Glu Tyr Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile

85 90 95

Glu Cys Thr Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln

100 105 110

Pro Ile Ala Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr

115 120 125

Cys Leu Glu Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe
 130 135 140
 Met Thr Thr Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg
 145 150 155 160
 Gly Leu Asp Leu Lys Gln Val Asp Gly Leu Pro Gly His His His His
 165 170 175
 His His Gly Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln
 180 185 190
 Lys Gln Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser
 195 200 205
 Leu Ala Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg
 210 215 220
 Ser Lys Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser
 225 230 235 240
 Ala Thr His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu
 245 250 255
 Thr Asn Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile
 260 265 270
 Asp Ile Arg Gly Ser Ala
 275
 <210> 24
 <211> 312
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> EGFP-71-CfaC-SopE
 <400> 24
 Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln
 1 5 10 15
 Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala
 20 25 30
 Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys

35 40 45
 Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
 50 55 60
 His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn

 65 70 75 80
 Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile
 85 90 95
 Arg Gly Ser Ala Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln
 100 105 110
 Asn Val Tyr Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys
 115 120 125
 Asn Gly Leu Val Ala Ser Asn Cys Phe Ser Arg Tyr Pro Asp His Met

 130 135 140
 Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln
 145 150 155 160
 Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala
 165 170 175
 Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys
 180 185 190
 Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu

 195 200 205
 Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys
 210 215 220
 Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly
 225 230 235 240
 Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp
 245 250 255
 Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala

 260 265 270
 Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu
 275 280 285

Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys

290 295 300

Gly His His His His His His Gly

305 310

<210> 25

<211> 192

<212> PRT

<213> Artificial Sequence

<220>

<223> EGFP-71-CfaN-DD1

<400> 25

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

1 5 10 15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly

20 25 30

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

35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr

50 55 60

Leu Thr Tyr Gly Val Gln Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr

65 70 75 80

Val Glu Tyr Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile

85 90 95

Glu Cys Thr Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln

100 105 110

Pro Ile Ala Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr

115 120 125

Cys Leu Glu Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe

130 135 140

Met Thr Thr Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg

145 150 155 160

Gly Leu Asp Leu Lys Gln Val Asp Gly Leu Pro Gly His His His His

165 170 175
His His Gly Arg Ile Ser Phe Gly Ser Pro Pro Pro Met Ala Gly Gly
180 185 190

<210> 26

<211> 226

<212> PRT

<213> Artificial Sequence

<220>

<223> EGFP-71-CfaC-DD1

<400> 26

Met Arg Ile Ser Phe Gly Ser Pro Pro Pro Met Ala Gly Gly Val Lys
1 5 10 15
Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr Asp Ile Gly
20 25 30
Val Glu Lys Asp His Asn Phe Leu Leu Lys Asn Gly Leu Val Ala Ser
35 40 45
Asn Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe
50 55 60

Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe
65 70 75 80
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly
85 90 95
Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu
100 105 110
Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His
115 120 125

Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn
130 135 140
Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp
145 150 155 160
His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro
165 170 175

Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn
180 185 190

Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly
195 200 205

Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Gly His His His His His
210 215 220

His Gly

225

<210> 27

<211> 101

<212> PRT

<213> Artificial Sequence

<220>

<223> CfaN

<400> 27

Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr Gly Phe Leu
1 5 10 15

Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr Val Tyr Thr
20 25 30

Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala Gln Trp His
35 40 45

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

Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr Asp Gly Gln
65 70 75 80

Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp Leu Lys Gln
85 90 95

Val Asp Gly Leu Pro
100

<210> 28

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> CfaC

<400> 28

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

1 5 10 15

Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30

Val Ala Ser Asn

35

<210> 29

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> CfaCmut

<400> 29

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

1 5 10 15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30

Val Ala Ser Asn

35

<210> 30

<211> 30

<212> PRT

<213> Artificial Sequence

<220>

<223>

> CatN

<400> 30

Cys Leu Ser Gly Asp Thr Met Ile Glu Ile Leu Asp Asp Asp Gly Ile

1 5 10 15

Ile Gln Lys Ile Ser Met Glu Asp Leu Tyr Gln Arg Leu Ala

20 25 30

<210> 31

<211> 104

<212> PRT

<213> Artificial Sequence

<220>

<223> CatC

<400> 31

Met Phe Lys Leu Asn Thr Lys Asn Ile Lys Val Leu Thr Pro Ser Gly

1 5 10 15

Phe Lys Ser Phe Ser Gly Ile Gln Lys Val Tyr Lys Pro Phe Tyr His

20 25 30

His Ile Ile Phe Asp Asp Gly Ser Glu Ile Lys Cys Ser Asp Asn His

35 40 45

Ser Phe Gly Lys Asp Lys Ile Lys Ala Ser Thr Ile Lys Val Gly Asp

50 55 60

Tyr Leu Gln Gly Lys Lys Val Leu Tyr Asn Glu Ile Val Glu Glu Gly

65 70 75 80

Ile Tyr Leu Tyr Asp Leu Leu Asn Val Gly Glu Asp Asn Leu Tyr Tyr

85 90 95

Thr Asn Gly Ile Val Ser His Asn

100

<210> 32

<211> 102

<212> PRT

<213> Artificial Sequence

<220>

<223> NpuN

<400> 32

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

1 5 10 15

Pro Ile Gly Lys Ile Val Glu Lys Arg Ile Glu Cys Thr Val Tyr Ser

20 25 30

Val Asp Asn Asn Gly Asn Ile Tyr Thr Gln Pro Val Ala Gln Trp His

35

40

45

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

50

55

60

Leu Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Val Asp Gly Gln

65

70

75

80

Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Glu Leu Asp Leu Met Arg

85

90

95

Val Asp Asn Leu Pro Asn

100

<210> 33

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> NpuC

<400> 33

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr

1

5

10

15

Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe

20

25

30

Ile Ala Ser Asn

35

<210> 34

<211> 28

<212> PRT

<213> Artificial Sequence

<220>

<223> M3

<400> 34

Lys Ser Val Thr Leu Glu Ser Arg Ser Pro Lys Phe Leu Asn Trp Phe

1

5

10

15

Ser Val Phe Ser Leu Phe Lys Val Ile Thr Thr Gly

20

25

<210> 35

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> M4

<400> 35

Tyr Met Ser Ile Leu Arg Cys Ala Ser Gly Lys Ile Ser Ile Ala Ala

1

5

10

15

Pro Pro Tyr Ile Phe

20

<210> 36

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223>

> NpuCmut

<400> 36

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr

1

5

10

15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Ala Leu Lys Asn Gly Phe

20

25

30

Ile Ala Ser Asn

35

<210> 37

<211> 34

<212> PRT

<213> Artificial Sequence

<220>

<223> M5

<400> 37

Ala Gly Glu Ser Phe Asn Phe Met Val Lys Leu Leu Tyr Lys His Pro

1 5 10 15

Ile Leu Pro Cys Leu Lys Thr Leu Leu Ser Ile Arg Ser Ser Cys Ser

20 25 30

Pro Arg

<210> 38

<211> 88

<212> PRT

<213> Artificial Sequence

<220>

<223> GP41N

<400> 38

Cys Leu Asp Leu Lys Thr Gln Val Gln Thr Pro Gln Gly Met Lys Glu

1 5 10 15

Ile Ser Asn Ile Gln Val Gly Asp Leu Val Leu Ser Asn Thr Gly Tyr

20 25 30

Asn Glu Val Leu Asn Val Phe Pro Lys Ser Lys Lys Lys Ser Tyr Lys

35 40 45

Ile Thr Leu Glu Asp Gly Lys Glu Ile Ile Cys Ser Glu Glu His Leu

50 55 60

Phe Pro Thr Gln Thr Gly Glu Met Asn Ile Ser Gly Gly Leu Lys Glu

65 70 75 80

Gly Met Cys Leu Tyr Val Lys Glu

85

<210> 39

<211> 102

<212> PRT

<213> Artificial Sequence

<220>

<223> ConN

<400> 39

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

1 5 10 15
 Pro Ile Gly Lys Ile Val Glu Glu Asn Ile Glu Cys Thr Val Tyr Ser
 20 25 30
 Val Asp Glu Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala Gln Trp His
 35 40 45
 Asp Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu Asp Gly Ser
 50 55 60

Thr Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Glu Asp Gly Glu
 65 70 75 80
 Met Leu Pro Ile Asp Glu Ile Phe Glu Gln Gly Leu Asp Leu Lys Gln
 85 90 95
 Val Lys Gly Leu Pro Asp
 100

<210> 40

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> CL1

<400> 40

Ala Cys Lys Asn Trp Phe Ser Ser Leu Ser His Phe Val Ile His Leu
 1 5 10 15

<210> 41

<211> 62

<212> PRT

<213> Artificial Sequence

<220>

<223> Deg1

<400> 41

Met Asn Lys Ile Pro Ile Lys Asp Leu Leu Asn Pro Gln Ile Thr Asp
 1 5 10 15
 Glu Phe Lys Ser Ser Ile Leu Asp Ile Asn Lys Lys Leu Phe Ser Ile
 20 25 30

Cys Cys Asn Leu Pro Lys Leu Pro Glu Ser Val Thr Thr Glu Glu Glu

35

40

45

Val Glu Leu Arg Asp Ile Leu Gly Phe Leu Ser Arg Ala Asn

50

55

60

<210> 42

<211> 30

<212> PRT

<213> Artificial Sequence

<220>

<223> PEST

<400> 42

Arg Ser Ser Ser Pro Ser Asp Ser Asp Thr Ser Gly Phe Ser Ser Gly

1

5

10

15

Ser Asp His Leu Ser Asp Leu Ile Ser Ser Leu Arg Ile Ser

20

25

30

<210> 43

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> DD1

<220>

<221> VARIANT

<

<222> 12

<223> X= any aminoacid, preferably Gly

<400> 43

Arg Ile Ser Phe Gly Ser Pro Pro Pro Met Ala Xaa Gly

1

5

10

<210> 44

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> DD2

<220>

<221> VARIANT

<222> 8

<223> X= any aminoacid, preferably Gly

<400> 44

Gly Ser Pro Pro Pro Met Ala Xaa Gly

1 5

<210> 45

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> DD3

<400> 45

Thr Asn Gly Ile Leu Lys Leu Gly Cys Gln Gly

1 5 10

<210> 46

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> M1

<400> 46

Cys Ser Glu Ile Ile Pro Met Ser Arg Ser Thr Pro Ile Ser Thr Met

1 5 10 15

Gly

<210> 47

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> M2

<400> 47

Val Ser Phe Ala Phe Asn Leu Asn Ser Leu Ile Val Gly Ile Leu Arg

1 5 10 15

Phe His Trp

<210> 48

<211> 100

<212> PRT

<213> Artificial Sequence

<220>

<223> SopE

<400> 48

Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln

1 5 10 15

Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala

20 25 30

Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys

35 40 45

Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr

50 55 60

His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn

65 70 75 80

Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile

85 90 95

Arg Gly Ser Ala

100

<210> 49

<211> 78

<212> PRT

<213> Artificial Sequence

<220>

<223> SopE-1-78

<400> 49

Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln

1 5 10 15
 Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala
 20 25 30
 Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys
 35 40 45
 Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
 50 55 60
 His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu

65 70 75

<210> 50

<211> 63

<212> PRT

<213> Artificial Sequence

<220>

<223> SopE-15-78

<400> 50

Gln Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu
 1 5 10 15
 Ala Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser
 20 25 30
 Lys Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala
 35 40 45

Thr His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu
 50 55 60

<210> 51

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> SopE-15-50

<400> 51

Gln Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu
 1 5 10 15
 Ala Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser
 20 25 30
 Lys Leu Ser
 35

<210> 52

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> L2

<400> 52

Ser Leu Ile Ser Leu Pro Leu Pro Thr Arg Val Lys Phe Ser Ser Leu
 1 5 10 15
 Leu Leu Ile Arg Ile Met Lys Ile Ile Thr Met Thr Phe Pro Lys Lys
 20 25 30
 Leu Arg Ser
 35

<210> 53

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> L6

<400> 53

Phe Tyr Tyr Pro Ile Trp Phe Ala Arg Val Leu Leu Val His Tyr Gln

1 5 10 15

<210> 54

<211> 46

<212> PRT

<213> Artificial Sequence

<220>

<223> L9

<400> 54

Ser Asn Pro Phe Ser Ser Leu Phe Gly Ala Ser Leu Leu Ile Asp Ser
 1 5 10 15
 Val Ser Leu Lys Ser Asn Trp Asp Thr Ser Ser Ser Ser Cys Leu Ile
 20 25 30
 Ser Phe Phe Ser Ser Val Met Phe Ser Ser Thr Thr Arg Ser
 35 40 45

<210> 55

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> L10

<400> 55

Cys Arg Gln Arg Phe Ser Cys His Leu Thr Ala Ser Tyr Pro Gln Ser
 1 5 10 15
 Thr Val Thr Pro Phe Leu Ala Phe Leu Arg Arg Asp Phe Phe Phe Leu
 20 25 30
 Arg His Asn Ser Ser Ala Asp
 35

<210> 56

<211> 46

<212> PRT

<213> Artificial Sequence

<220>

<223> L11

<400> 56

Gly Ala Pro His Val Val Leu Phe Asp Phe Glu Leu Arg Ile Thr Asn
 1 5 10 15
 Pro Leu Ser His Ile Gln Ser Val Ser Leu Gln Ile Thr Leu Ile Phe
 20 25 30
 Cys Ser Leu Pro Ser Leu Ile Leu Ser Lys Phe Leu Gln Val

35 40 45

<210> 57

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> L12

<400> 57

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

1 5 10 15

Lys Lys Thr Leu Leu Leu Ala Gln Ile Ser Ser Leu Phe Phe Leu Leu

20 25 30

Leu Ser Ser Asn Ile Ala Val

35

<210> 58

<211> 45

<212> PRT

<213> Artificial Sequence

<220>

<223> L15

<400> 58

Pro Thr Val Lys Asn Ser Pro Lys Ile Phe Cys Leu Ser Ser Ser Pro

1 5 10 15

Tyr Leu Ala Phe Asn Leu Glu Tyr Leu Ser Leu Arg Ile Phe Ser Thr

20 25 30

Leu Ser Lys Cys Ser Asn Thr Leu Leu Thr Ser Leu Ser

35 40 45

<210> 59

<211> 30

<212> PRT

<213> Artificial Sequence

<220>

<223> L16

<400> 59

Ser Asn Gln Leu Lys Arg Leu Trp Leu Trp Leu Leu Glu Val Arg Ser
 1 5 10 15
 Phe Asp Arg Thr Leu Arg Arg Pro Trp Ile His Leu Pro Ser
 20 25 30

<210> 60

<211> 1149

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4N-1150

<400> 60

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
 1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160

Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser

405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met

 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

 500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro

 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

900 905 910
Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915 920 925
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

930 935 940
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu

965 970 975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

1045 1050 1055
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085
Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100
Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120
Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135
Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn

1140 1145

<210> 61

<211> 1124

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4C-1150

<400> 61

Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn

1 5 10 15

Ile Gln Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser

20 25 30

Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro

35 40 45

Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu

50 55 60

His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu

65 70 75 80

Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser

85 90 95

Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser

100 105 110

Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr

115 120 125

Glu Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys

130 135 140

Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys

145 150 155 160

Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro

165 170 175

Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly

180 185 190

Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala

195 200 205
 Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe
 210 215 220
 Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met
 225 230 235 240

 Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu
 245 250 255
 His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu
 260 265 270
 Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys
 275 280 285
 Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr
 290 295 300

 Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn
 305 310 315 320
 Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser
 325 330 335
 Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu
 340 345 350
 Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg
 355 360 365

 Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe
 370 375 380
 Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys
 385 390 395 400
 Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys
 405 410 415
 Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser
 420 425 430

 Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu
 435 440 445

Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp
450 455 460

Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val Ser
465 470 475 480

Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys
485 490 495

Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu
500 505 510

Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser
515 520 525

Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val
530 535 540

Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser
545 550 555 560

Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val
565 570 575

Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu
580 585 590

Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro
595 600 605

Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala
610 615 620

Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser
625 630 635 640

Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn
645 650 655

Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr
660 665 670

Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro
675 680 685

His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala

690 695 700
 Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro
 705 710 715 720
 Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu
 725 730 735
 Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe
 740 745 750

 Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu
 755 760 765
 Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn
 770 775 780
 Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly
 785 790 795 800
 Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly
 805 810 815

 Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr
 820 825 830
 Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr
 835 840 845
 Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn
 850 855 860
 Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly
 865 870 875 880

 Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu
 885 890 895
 Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr
 900 905 910
 Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg
 915 920 925
 Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu
 930 935 940

Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu
 945 950 955 960
 Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu
 965 970 975
 Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala
 980 985 990
 Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu
 995 1000 1005

Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser
 1010 1015 1020
 Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe
 1025 1030 1035 1040
 Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met
 1045 1050 1055
 Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu
 1060 1065 1070

Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr
 1075 1080 1085
 Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr
 1090 1095 1100
 Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg
 1105 1110 1115 1120
 Gln Ala Gln Asp

<210> 62

<211> 1139

<212> PRT

<213> Artificial Sequence

<220>

<223>

> ABCA4N-1140

<400> 62

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
 50 55 60

 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125

 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190

 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320

 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380

 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445

 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

500	505	510
Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr		
515	520	525
Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu		
530	535	540
Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr		
545	550	555
Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp		
565	570	575
Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly		
580	585	590
Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe		
595	600	605
Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val		
610	615	620
Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro		
625	630	635
Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro		
645	650	655
Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys		
660	665	670
Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn		
675	680	685
Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser		
690	695	700
Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met		
705	710	715
His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe		
725	730	735
Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser		
740	745	750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005
 Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
 1010 1015 1020

 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055
 Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
 1060 1065 1070
 Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
 1075 1080 1085

 Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
 1090 1095 1100
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135
 Arg Leu Tyr

<210> 63

<211> 1134

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4C-1140

<400> 63

Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly Thr Gly Leu
 1 5 10 15
 Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg Lys
 20 25 30
 Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr
 35 40 45

Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly
 50 55 60

Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu Ala
 65 70 75 80

Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn
 85 90 95

Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu
 100 105 110

Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr
 115 120 125

Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly
 130 135 140

Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn Pro
 145 150 155 160

Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln
 165 170 175

Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu Gly
 180 185 190

Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly
 195 200 205

Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg Phe
 210 215 220

Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu
 225 230 235 240

Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro
 245 250 255

Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly
 260 265 270

Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe
 275 280 285

Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg

290 295 300
 Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr
 305 310 315 320

 Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln
 325 330 335
 Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser
 340 345 350
 Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly
 355 360 365
 Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln
 370 375 380

 Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro
 385 390 395 400
 Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln
 405 410 415
 Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile
 420 425 430
 Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met
 435 440 445

 Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro
 450 455 460
 Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe
 465 470 475 480
 Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His
 485 490 495
 Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu
 500 505 510

 Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln
 515 520 525
 Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala
 530 535 540

Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu
 545 550 555 560
 Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile
 565 570 575

Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp
 580 585 590
 Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile
 595 600 605
 Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu
 610 615 620
 Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr
 625 630 635 640

Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr Val Ala Leu
 645 650 655
 Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe
 660 665 670
 Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala
 675 680 685
 Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg
 690 695 700

Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala
 705 710 715 720
 Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile
 725 730 735
 Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu
 740 745 750
 Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala
 755 760 765

Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu
 770 775 780
 Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg

785 790 795 800
 Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val
 805 810 815
 Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu
 820 825 830

 Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly
 835 840 845
 Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile
 850 855 860
 Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln
 865 870 875 880
 Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu
 885 890 895

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

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

 Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu
 1025 1030 1035 1040

Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly
1045 1050 1055
Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser
1060 1065 1070
Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp
1075 1080 1085

Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln
1090 1095 1100
Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro
1105 1110 1115 1120
Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp
1125 1130

<210> 64

<211> 1187

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4N-1188

<400> 64

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15
Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
20 25 30
Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
35 40 45
Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
50 55 60
Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80
Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
85 90 95
Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125
 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala

 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser

 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser

 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe

595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro

 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met

 705 710 715 720
 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750
 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765
 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp

 770 775 780
 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
 805 810 815
 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
 820 825 830
 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

 835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
 1140 1145 1150
 Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser

1155 1160 1165
 Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe
 1170 1175 1180

Ser Thr Thr

1185

<210> 65

<211> 1086

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4C-1188

<400> 65

Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly
 1 5 10 15
 Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu Ala

20 25 30
 Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn
 35 40 45
 Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu
 50 55 60
 Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr
 65 70 75 80
 Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly

85

90

95

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

340 345 350
 Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln
 355 360 365
 Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile
 370 375 380
 Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met
 385 390 395 400
 Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro

 405 410 415
 Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe
 420 425 430
 Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His
 435 440 445
 Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu
 450 455 460
 Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln

 465 470 475 480
 Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala
 485 490 495
 Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu
 500 505 510
 Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile
 515 520 525
 Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp

 530 535 540
 Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile
 545 550 555 560
 Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu
 565 570 575
 Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr
 580 585 590

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

595 600 605

Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe

610 615 620

Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala

625 630 635 640

Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg

645 650 655

Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala

660 665 670

Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile

675 680 685

Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu

690 695 700

Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala

705 710 715 720

Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu

725 730 735

Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg

740 745 750

Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val

755 760 765

Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu

770 775 780

Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly

785 790 795 800

Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile

805 810 815

Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln

820 825 830

Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu

835 840 845
Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala

850 855 860
Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu
865 870 875 880

Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile
885 890 895

Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr
900 905 910

Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser

915 920 925
Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu
930 935 940

Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala
945 950 955 960

Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp
965 970 975

Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu

980 985 990
Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly
995 1000 1005

Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser
1010 1015 1020

Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp
1025 1030 1035 1040

Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln

1045 1050 1055
Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro
1060 1065 1070

Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp

1075 1080 1085

<210> 66

<211> 1250

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaN

<400> 66

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270

 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335

 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400

 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser

690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720

 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750
 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765
 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780

 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
 805 810 815
 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
 820 825 830
 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
 835 840 845

 Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
 850 855 860
 Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
 865 870 875 880
 Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
 885 890 895
 Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
 900 905 910

 Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
 915 920 925
 Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
 930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
 945 950 955 960
 Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
 965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
 980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
 995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
 1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
 1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
 1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
 1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135

Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Leu Ser
 1140 1145 1150

Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr Gly Phe Leu Pro Ile Gly
 1155 1160 1165

Lys Ile Val Glu Glu Arg Ile Glu Cys Thr Val Tyr Thr Val Asp Lys
 1170 1175 1180

Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala Gln Trp His Asn Arg Gly

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

 Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp Leu Lys Gln Val Asp Gly
 1235 1240 1245
 Leu Pro
 1250
 <210> 67
 <211> 1160
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> ABCA4-1150-CfaC
 <400> 67
 Met Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr
 1 5 10 15
 Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys Asn Gly Leu
 20 25 30

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

 Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile
 100 105 110
 Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg
 115 120 125

Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu
130 135 140

Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe
145 150 155 160

Leu Lys Val Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly
165 170 175

Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly
180 185 190

Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser
195 200 205

Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro
210 215 220

Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln
225 230 235 240

His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser
245 250 255

His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe
260 265 270

Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro
275 280 285

Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe
290 295 300

Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val
305 310 315 320

Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp
325 330 335

Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser
340 345 350

Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln
355 360 365

Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr

370 375 380
 Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln
 385 390 395 400
 Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn
 405 410 415

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

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

 Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val
 545 550 555 560
 Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser
 565 570 575
 Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg
 580 585 590
 Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr
 595 600 605

 Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val
 610 615 620

Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala
625 630 635 640

Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu
645 650 655

Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe
660 665 670

Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe
675 680 685

Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu
690 695 700

Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu
705 710 715 720

Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala
725 730 735

Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His
740 745 750

Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala
755 760 765

Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val Gln
770 775 780

Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro
785 790 795 800

Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile Ile
805 810 815

Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr Lys
820 825 830

Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val Gly
835 840 845

Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly
850 855 860

Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser

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

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

 Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg
 995 1000 1005
 Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys
 1010 1015 1020
 Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr
 1025 1030 1035 1040
 Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met
 1045 1050 1055

 Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val
 1060 1065 1070
 Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg
 1075 1080 1085
 His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg
 1090 1095 1100
 Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu
 1105 1110 1115 1120

Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala

1125

1130

1135

Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala

1140

1145

1150

Gly Ala Ser Arg Gln Ala Gln Asp

1155

1160

<210> 68

<211> 1240

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaN

<400> 68

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1

5

10

15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20

25

30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35

40

45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50

55

60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65

70

75

80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85

90

95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100

105

110

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

115

120

125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130

135

140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205

 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270

 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335

 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro

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

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
1125 1130 1135

Arg Leu Tyr Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr

1140 1145 1150
Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr
1155 1160 1165

Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala
1170 1175 1180

Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu
1185 1190 1195 1200

Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr
1205 1210 1215

Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp
1220 1225 1230

Leu Lys Gln Val Asp Gly Leu Pro
1235 1240

<210> 69

<211> 1170

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaCmut

<400> 69

Met Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr
1 5 10 15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu
20 25 30

Val Ala Ser Asn Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe
35 40 45

Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln
50 55 60

Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly
65 70 75 80

Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln
85 90 95

Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His
100 105 110

Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe
115 120 125

Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe
130 135 140

Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly
145 150 155 160

Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp
165 170 175

Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu
180 185 190

Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly
195 200 205

Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala
210 215 220

His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln
225 230 235 240

Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu
245 250 255

Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala
260 265 270

Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser
275 280 285

Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro
290 295 300

Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly
305 310 315 320

Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly
325 330 335

Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys

340 345 350
 Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr
 355 360 365

 Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser
 370 375 380
 Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro
 385 390 395 400
 Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr
 405 410 415
 Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val
 420 425 430

 Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp
 435 440 445
 Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro
 450 455 460
 Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu
 465 470 475 480
 Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser
 485 490 495

 Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile
 500 505 510
 Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu
 515 520 525
 Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg
 530 535 540
 Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu
 545 550 555 560

 Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp
 565 570 575
 Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala
 580 585 590

Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His
595 600 605

Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn
610 615 620

Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val
625 630 635 640

Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn
645 650 655

Leu Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala Val Ile
660 665 670

Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala
675 680 685

Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser
690 695 700

Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu
705 710 715 720

Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe
725 730 735

Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr
740 745 750

Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His
755 760 765

Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val
770 775 780

Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser
785 790 795 800

Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp
805 810 815

Asp Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr
820 825 830

Asp Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser

835 840 845
Ser Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys

850 855 860
Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys
865 870 875 880

Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala
885 890 895

Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly
900 905 910

Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu
915 920 925

His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile
930 935 940

Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr
945 950 955 960

Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu
965 970 975

Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp
980 985 990

Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn
995 1000 1005

Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser
1010 1015 1020

His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met
1025 1030 1035 1040

Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser
1045 1050 1055

Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys
1060 1065 1070

Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly
1075 1080 1085

Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln

1090 1095 1100

Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu

1105 1110 1115 1120

Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr

1125 1130 1135

Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser

1140 1145 1150

His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala

1155 1160 1165

Gln Asp

1170

<210> 70

<211> 1288

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1188-CfaN

<400> 70

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125
 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140

 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205

 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270

 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335

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

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575

Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe

595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655

 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720

 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750
 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765
 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780

 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
 805 810 815
 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
 820 825 830
 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
 835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090	1095	1100	
Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met			
1105	1110	1115	1120
Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly			
	1125	1130	1135
Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly			
	1140	1145	1150
Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser			
1155	1160	1165	
Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe			
1170	1175	1180	
Ser Thr Thr Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr			
1185	1190	1195	1200
Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr			
	1205	1210	1215
Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala			
	1220	1225	1230
Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu			
1235	1240	1245	
Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr			
1250	1255	1260	
Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp			
1265	1270	1275	1280
Leu Lys Gln Val Asp Gly Leu Pro			
	1285		
<210> 71			
<211> 1122			
<212> PRT			
<213> Artificial Sequence			
<220>			
<223> ABCA4-1188-Cfacmut			

<400> 71

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

1 5 10 15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30

Val Ala Ser Asn Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln

35 40 45

Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His

50 55 60

Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe

65 70 75 80

Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe

85 90 95

Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly

100 105 110

Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp

115 120 125

Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu

130 135 140

Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly

145 150 155 160

Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala

165 170 175

His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln

180 185 190

Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu

195 200 205

Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala

210 215 220

Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser

225 230 235 240

Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro

245										250					255									
Trp	Ile	Tyr	Gly	Gln	Gln	Tyr	Thr	Phe	Phe	Ser	Met	Asp	Glu	Pro	Gly									
260										265					270									
Ser	Glu	Gln	Phe	Thr	Val	Leu	Ala	Asp	Val	Leu	Leu	Asn	Lys	Pro	Gly									
275										280					285									
Phe	Gly	Asn	Arg	Cys	Leu	Lys	Glu	Gly	Trp	Leu	Pro	Glu	Tyr	Pro	Cys									
290										295					300									
Gly	Asn	Ser	Thr	Pro	Trp	Lys	Thr	Pro	Ser	Val	Ser	Pro	Asn	Ile	Thr									
305										310					315					320				
Gln	Leu	Phe	Gln	Lys	Gln	Lys	Trp	Thr	Gln	Val	Asn	Pro	Ser	Pro	Ser									
325										330					335									
Cys	Arg	Cys	Ser	Thr	Arg	Glu	Lys	Leu	Thr	Met	Leu	Pro	Glu	Cys	Pro									
340										345					350									
Glu	Gly	Ala	Gly	Gly	Leu	Pro	Pro	Pro	Gln	Arg	Thr	Gln	Arg	Ser	Thr									
355										360					365									
Glu	Ile	Leu	Gln	Asp	Leu	Thr	Asp	Arg	Asn	Ile	Ser	Asp	Phe	Leu	Val									
370										375					380									
Lys	Thr	Tyr	Pro	Ala	Leu	Ile	Arg	Ser	Ser	Leu	Lys	Ser	Lys	Phe	Trp									
385										390					395					400				
Val	Asn	Glu	Gln	Arg	Tyr	Gly	Gly	Ile	Ser	Ile	Gly	Gly	Lys	Leu	Pro									
405										410					415									
Val	Val	Pro	Ile	Thr	Gly	Glu	Ala	Leu	Val	Gly	Phe	Leu	Ser	Asp	Leu									
420										425					430									
Gly	Arg	Ile	Met	Asn	Val	Ser	Gly	Gly	Pro	Ile	Thr	Arg	Glu	Ala	Ser									
435										440					445									
Lys	Glu	Ile	Pro	Asp	Phe	Leu	Lys	His	Leu	Glu	Thr	Glu	Asp	Asn	Ile									
450										455					460									
Lys	Val	Trp	Phe	Asn	Asn	Lys	Gly	Trp	His	Ala	Leu	Val	Ser	Phe	Leu									
465										470					475					480				
Asn	Val	Ala	His	Asn	Ala	Ile	Leu	Arg	Ala	Ser	Leu	Pro	Lys	Asp	Arg									
485										490					495									

Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu
500 505 510

Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp
515 520 525

Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala
530 535 540

Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His
545 550 555 560

Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn
565 570 575

Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val
580 585 590

Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn
595 600 605

Leu Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile
610 615 620

Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala
625 630 635 640

Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser
645 650 655

Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu
660 665 670

Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe
675 680 685

Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr
690 695 700

Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His
705 710 715 720

Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val
725 730 735

Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser

740 745 750
 Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp
 755 760 765

 Asp Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr
 770 775 780
 Asp Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser
 785 790 795 800
 Ser Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys
 805 810 815
 Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys
 820 825 830

 Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala
 835 840 845
 Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly
 850 855 860
 Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu
 865 870 875 880
 His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile
 885 890 895

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

 Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser
 965 970 975
 His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met
 980 985 990

Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser
995 1000 1005

Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys
1010 1015 1020

Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly
1025 1030 1035 1040

Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln
1045 1050 1055

Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu
1060 1065 1070

Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr
1075 1080 1085

Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser
1090 1095 1100

His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala
1105 1110 1115 1120

Gln Asp

<210> 72

<211> 1349

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaN-SopE

<400> 72

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80

 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125
 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140

 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205

 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270

 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335

Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr

545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590

 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655

 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720

 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750
 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765
 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780

 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

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

Lys Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala
1300 1305 1310
Thr His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr
1315 1320 1325
Asn Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp
1330 1335 1340
Ile Arg Gly Ser Ala
1345
<210> 73
<211> 1259
<212> PRT
<213> Artificial Sequence
<220>
<223> ABCA4-1150-CfaC-SopE
<400> 73
Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln
1 5 10 15
Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala
20 25 30
Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys
35 40 45
Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
50 55 60
His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn
65 70 75 80
Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile
85 90 95
Arg Gly Ser Ala Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln
100 105 110
Asn Val Tyr Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Lys
115 120 125
Asn Gly Leu Val Ala Ser Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr

130	135	140	
Leu Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu			
145	150	155	160
Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala			
165	170	175	
His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn			
180	185	190	
Glu Leu Met Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val			
195	200	205	
Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe			
210	215	220	
Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu			
225	230	235	240
Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu			
245	250	255	
Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe			
260	265	270	
Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro			
275	280	285	
Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn			
290	295	300	
Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro			
305	310	315	320
Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu			
325	330	335	
Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr			
340	345	350	
Ile Arg Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr			
355	360	365	
Phe Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly			
370	375	380	

Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr
 385 390 395 400
 Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu
 405 410 415
 Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys
 420 425 430
 Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys
 435 440 445

 Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys
 450 455 460
 Trp Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu
 465 470 475 480
 Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro
 485 490 495
 Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr
 500 505 510

 Asp Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile
 515 520 525
 Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly
 530 535 540
 Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu
 545 550 555 560
 Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser
 565 570 575

 Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu
 580 585 590
 Lys His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys
 595 600 605
 Gly Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile
 610 615 620
 Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile

625 630 635 640

 Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu

 645 650 655
 Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val

 660 665 670
 Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile

 675 680 685
 Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val

 690 695 700

 Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn
 705 710 715 720
 Tyr Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln

 725 730 735
 Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu

 740 745 750
 Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser

 755 760 765

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

 770 775 780
 Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu
 785 790 795 800
 Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg

 805 810 815
 Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile

 820 825 830

 Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly

 835 840 845
 Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn

 850 855 860
 Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu
 865 870 875 880

Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr
885 890 895

Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln
900 905 910

Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu
915 920 925

Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu
930 935 940

Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn
945 950 955 960

Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr
965 970 975

Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn
980 985 990

Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala
995 1000 1005

Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg
1010 1015 1020

Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser
1025 1030 1035 1040

Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr
1045 1050 1055

Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile
1060 1065 1070

Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp
1075 1080 1085

Pro Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg
1090 1095 1100

Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu
1105 1110 1115 1120

Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys

1125 1130 1135
Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile
1140 1145 1150

Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu
1155 1160 1165

Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln
1170 1175 1180

Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser
1185 1190 1195 1200

Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu
1205 1210 1215

Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val
1220 1225 1230

Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro
1235 1240 1245

Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp
1250 1255

<210> 74

<211> 1339

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaN-SopE

<400> 74

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15
Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30
Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45
Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile

225 230 235 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val

245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser

260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile

275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met

290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser

305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser

 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr

 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu

 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr

 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560

Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575
Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590
Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605
Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620
Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640
Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655
Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670
Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685
Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700
Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720
His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735
Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750
Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765
Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780
Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800
Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly

805 810 815
Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp

820 825 830
Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

835 840 845
Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp

850 855 860
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp

865 870 875 880
Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys

885 890 895
Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

900 905 910
Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915 920 925
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

930 935 940
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu

965 970 975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135

Arg Leu Tyr Cys Leu Ser Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr

1140 1145 1150

Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu Arg Ile Glu Cys Thr

1155 1160 1165

Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala

1170 1175 1180

Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu

1185 1190 1195 1200

Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His Lys Phe Met Thr Thr

1205 1210 1215

Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp

1220 1225 1230

Leu Lys Gln Val Asp Gly Leu Pro Thr Lys Ile Thr Leu Ser Pro Gln

1235 1240 1245

Asn Phe Arg Ile Gln Lys Gln Glu Thr Thr Leu Leu Lys Glu Lys Ser

1250 1255 1260

Thr Glu Lys Asn Ser Leu Ala Lys Ser Ile Leu Ala Val Lys Asn His

1265 1270 1275 1280

Phe Ile Glu Leu Arg Ser Lys Leu Ser Glu Arg Phe Ile Ser His Lys

1285 1290 1295

Asn Thr Glu Ser Ser Ala Thr His Phe His Arg Gly Ser Ala Ser Glu

1300 1305 1310
 Gly Arg Ala Val Leu Thr Asn Lys Val Val Lys Asp Phe Met Leu Gln
 1315 1320 1325
 Thr Leu Asn Asp Ile Asp Ile Arg Gly Ser Ala
 1330 1335
 <210> 75
 <211> 1269
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> ABCA4-1140-CfaCmut-SopE

 <400> 75
 Met Thr Lys Ile Thr Leu Ser Pro Gln Asn Phe Arg Ile Gln Lys Gln
 1 5 10 15
 Glu Thr Thr Leu Leu Lys Glu Lys Ser Thr Glu Lys Asn Ser Leu Ala
 20 25 30
 Lys Ser Ile Leu Ala Val Lys Asn His Phe Ile Glu Leu Arg Ser Lys
 35 40 45
 Leu Ser Glu Arg Phe Ile Ser His Lys Asn Thr Glu Ser Ser Ala Thr
 50 55 60

 His Phe His Arg Gly Ser Ala Ser Glu Gly Arg Ala Val Leu Thr Asn
 65 70 75 80
 Lys Val Val Lys Asp Phe Met Leu Gln Thr Leu Asn Asp Ile Asp Ile
 85 90 95
 Arg Gly Ser Ala Val Lys Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln
 100 105 110
 Asn Val Tyr Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys
 115 120 125

 Asn Gly Leu Val Ala Ser Asn Cys Ser Gly Thr Pro Leu Phe Leu Lys
 130 135 140
 Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys
 145 150 155 160

Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser
 165 170 175
 Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr
 180 185 190

 Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val
 195 200 205
 Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu
 210 215 220
 Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala
 225 230 235 240
 Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser
 245 250 255

 Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val
 260 265 270
 Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln
 275 280 285
 Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu
 290 295 300
 Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala
 305 310 315 320

 Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro
 325 330 335
 Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln
 340 345 350
 Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp
 355 360 365
 Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu
 370 375 380

 Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr
 385 390 395 400
 Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp

405 410 415
 Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn
 420 425 430
 Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu
 435 440 445

 Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro
 450 455 460
 Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro
 465 470 475 480
 Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro
 485 490 495
 Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln
 500 505 510

 Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp
 515 520 525
 Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser
 530 535 540
 Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly
 545 550 555 560
 Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu
 565 570 575

 Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg
 580 585 590
 Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu
 595 600 605
 Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val
 610 615 620
 Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro
 625 630 635 640

 Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro
 645 650 655

Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr
 660 665 670
 Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe
 675 680 685
 Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys
 690 695 700

 Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp
 705 710 715 720
 Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly
 725 730 735
 Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser
 740 745 750
 Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp
 755 760 765

 Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro
 770 775 780
 Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile
 785 790 795 800
 Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg
 805 810 815
 Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe
 820 825 830

 Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln
 835 840 845
 Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn
 850 855 860
 Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val
 865 870 875 880
 Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe
 885 890 895

 Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp

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

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

Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala
 1025 1030 1035 1040
 Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu
 1045 1050 1055
 Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys
 1060 1065 1070
 Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val
 1075 1080 1085

Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met
 1090 1095 1100
 Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val
 1105 1110 1115 1120
 Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu
 1125 1130 1135
 Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His
 1140 1145 1150

Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys

1155 1160 1165

Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe

1170 1175 1180

Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn

1185 1190 1195 1200

Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln

1205 1210 1215

Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val

1220 1225 1230

Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln

1235 1240 1245

Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser

1250 1255 1260

Arg Gln Ala Gln Asp

1265

<210> 76

<211> 1263

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1150-CfaN-DD1

<400> 76

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125

 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190

 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380

 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445

 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510

 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

565					570					575						
Val	Val	Glu	Lys	Thr	Asn	Lys	Ile	Lys	Asp	Arg	Tyr	Trp	Asp	Ser	Gly	
580					585					590						
Pro	Arg	Ala	Asp	Pro	Val	Glu	Asp	Phe	Arg	Tyr	Ile	Trp	Gly	Gly	Phe	
595					600					605						
Ala	Tyr	Leu	Gln	Asp	Met	Val	Glu	Gln	Gly	Ile	Thr	Arg	Ser	Gln	Val	
610					615					620						
Gln	Ala	Glu	Ala	Pro	Val	Gly	Ile	Tyr	Leu	Gln	Gln	Met	Pro	Tyr	Pro	
625					630					635					640	
Cys	Phe	Val	Asp	Asp	Ser	Phe	Met	Ile	Ile	Leu	Asn	Arg	Cys	Phe	Pro	
645					650					655						
Ile	Phe	Met	Val	Leu	Ala	Trp	Ile	Tyr	Ser	Val	Ser	Met	Thr	Val	Lys	
660					665					670						
Ser	Ile	Val	Leu	Glu	Lys	Glu	Leu	Arg	Leu	Lys	Glu	Thr	Leu	Lys	Asn	
675					680					685						
Gln	Gly	Val	Ser	Asn	Ala	Val	Ile	Trp	Cys	Thr	Trp	Phe	Leu	Asp	Ser	
690					695					700						
Phe	Ser	Ile	Met	Ser	Met	Ser	Ile	Phe	Leu	Leu	Thr	Ile	Phe	Ile	Met	
705					710					715					720	
His	Gly	Arg	Ile	Leu	His	Tyr	Ser	Asp	Pro	Phe	Ile	Leu	Phe	Leu	Phe	
725					730					735						
Leu	Leu	Ala	Phe	Ser	Thr	Ala	Thr	Ile	Met	Leu	Cys	Phe	Leu	Leu	Ser	
740					745					750						
Thr	Phe	Phe	Ser	Lys	Ala	Ser	Leu	Ala	Ala	Ala	Cys	Ser	Gly	Val	Ile	
755					760					765						
Tyr	Phe	Thr	Leu	Tyr	Leu	Pro	His	Ile	Leu	Cys	Phe	Ala	Trp	Gln	Asp	
770					775					780						
Arg	Met	Thr	Ala	Glu	Leu	Lys	Lys	Ala	Val	Ser	Leu	Leu	Ser	Pro	Val	
785					790					795					800	
Ala	Phe	Gly	Phe	Gly	Thr	Glu	Tyr	Leu	Val	Arg	Phe	Glu	Glu	Gln	Gly	
805					810					815						

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060	1065	1070	
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro			
1075	1080	1085	
Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu			
1090	1095	1100	
Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met			
1105	1110	1115	1120
Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly			
1125	1130	1135	
Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Leu Ser			
1140	1145	1150	
Tyr Asp Thr Glu Ile Leu Thr Val Glu Tyr Gly Phe Leu Pro Ile Gly			
1155	1160	1165	
Lys Ile Val Glu Glu Arg Ile Glu Cys Thr Val Tyr Thr Val Asp Lys			
1170	1175	1180	
Asn Gly Phe Val Tyr Thr Gln Pro Ile Ala Gln Trp His Asn Arg Gly			
1185	1190	1195	1200
Glu Gln Glu Val Phe Glu Tyr Cys Leu Glu Asp Gly Ser Ile Ile Arg			
1205	1210	1215	
Ala Thr Lys Asp His Lys Phe Met Thr Thr Asp Gly Gln Met Leu Pro			
1220	1225	1230	
Ile Asp Glu Ile Phe Glu Arg Gly Leu Asp Leu Lys Gln Val Asp Gly			
1235	1240	1245	
Leu Pro Arg Ile Ser Phe Gly Ser Pro Pro Pro Met Ala Gly Gly			
1250	1255	1260	
<210> 77			
<211> 1173			
<212> PRT			
<213> Artificial Sequence			
<220>			
<223> ABCA4-1150-CfaC-DD1			
<400> 77			

Met Arg Ile Ser Phe Gly Ser Pro Pro Pro Met Ala Gly Gly Val Lys

1 5 10 15

Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr Asp Ile Gly

20 25 30

Val Glu Lys Asp His Asn Phe Leu Leu Lys Asn Gly Leu Val Ala Ser

35 40 45

Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys

50 55 60

Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser

65 70 75 80

Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr

85 90 95

Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val

100 105 110

Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu

115 120 125

Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala

130 135 140

Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser

145 150 155 160

Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val

165 170 175

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

180 185 190

Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu

195 200 205

Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala

210 215 220

Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro

225 230 235 240

Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln

245 250 255
Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp

260 265 270
Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu

275 280 285
Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr

290 295 300
Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp

305 310 315 320
Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn

325 330 335
Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu

340 345 350
Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro

355 360 365
Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro

370 375 380
Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro

385 390 395 400
Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln

405 410 415
Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp

420 425 430
Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser

435 440 445
Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly

450 455 460
Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu

465 470 475 480
Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg

485 490 495

Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu
500 505 510
Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val
515 520 525
Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro
530 535 540
Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro
545 550 555 560
Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr
565 570 575
Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe
580 585 590
Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys
595 600 605
Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp
610 615 620
Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly
625 630 635 640
Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser
645 650 655
Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp
660 665 670
Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro
675 680 685
Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile
690 695 700
Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg
705 710 715 720
Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe
725 730 735
Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln

740 745 750
Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn
755 760 765
Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val

770 775 780
Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe
785 790 795 800
Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp
805 810 815
Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly
820 825 830
Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Pro

835 840 845
Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro
850 855 860
Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr
865 870 875 880
Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala
885 890 895
Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln

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

965 970 975
Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val
980 985 990

Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met

995 1000 1005

Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val

1010 1015 1020

Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu

1025 1030 1035 1040

Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His

1045 1050 1055

Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys

1060 1065 1070

Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe

1075 1080 1085

Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn

1090 1095 1100

Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln

1105 1110 1115 1120

Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val

1125 1130 1135

Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln

1140 1145 1150

Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser

1155 1160 1165

Arg Gln Ala Gln Asp

1170

<210> 78

<211> 1253

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaN-DD1

<400> 78

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

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

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

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

 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr

 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val

 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn

 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720
 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser

 740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

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

1250

<210

> 79

<211> 1183

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1140-CfaCmut-DD1

<400> 79

Met Arg Ile Ser Phe Gly Ser Pro Pro Pro Met Ala Gly Gly Val Lys

1 5 10 15

Ile Ile Ser Arg Lys Ser Leu Gly Thr Gln Asn Val Tyr Asp Ile Gly

20 25 30

Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu Val Ala Ser

35 40 45

Asn Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly Thr Gly

50 55 60

Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg

65 70 75 80

Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr

85 90 95

Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp

100 105 110

Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu

115 120 125

Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro

130 135 140

Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu

145 150 155 160

Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp

165 170 175

Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser

180 185 190
 Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn
 195 200 205
 Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro
 210 215 220
 Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu
 225 230 235 240
 Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr

 245 250 255
 Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg
 260 265 270
 Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala Gln Ile Val
 275 280 285
 Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile
 290 295 300
 Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr

 305 310 315 320
 Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln
 325 330 335
 Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn
 340 345 350
 Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser
 355 360 365
 Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe

 370 375 380
 Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys
 385 390 395 400
 Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala
 405 410 415
 Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu
 420 425 430

Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr

435 440 445

Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu

450 455 460

Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro

465 470 475 480

Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile

485 490 495

Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile

500 505 510

Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp

515 520 525

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

530 535 540

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

545 550 555 560

Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu

565 570 575

Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val

580 585 590

Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val

595 600 605

Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe

610 615 620

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

625 630 635 640

Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe

645 650 655

Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala

660 665 670

Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met

675 680 685
Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr Val Ala

690 695 700
Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr
705 710 715 720
Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn
725 730 735
Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly
740 745 750
Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr

755 760 765
Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu
770 775 780
Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe
785 790 795 800
Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile
805 810 815
Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala

820 825 830
Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu
835 840 845
Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala
850 855 860
Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu
865 870 875 880
Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr

885 890 895
Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser
900 905 910
Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro
915 920 925

Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr

930 935 940

Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val

945 950 955 960

Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys

965 970 975

Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala

980 985 990

Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr

995 1000 1005

Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val

1010 1015 1020

Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met

1025 1030 1035 1040

Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly

1045 1050 1055

Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly

1060 1065 1070

Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu

1075 1080 1085

Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro

1090 1095 1100

Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val

1105 1110 1115 1120

Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys

1125 1130 1135

Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp

1140 1145 1150

Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu

1155 1160 1165

Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp

1170 1175 1180

<210> 80

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> DD4

<220>

<221> VARIANT

<222> 12

<223> X= any aminoacid, preferably Gly

<400> 80

Arg Asp Ser Phe Gly Ser Pro Pro Pro Met Ala Xaa Gly

1 5 10

<210> 81

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> DD5

<400> 81

Gly Phe Trp Ile

1

<210> 82

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> DD6

<400> 82

Arg Phe Lys Gly

1

<210> 83

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> DD7

<400> 83

Lys Phe Tyr Lys

1

<210> 84

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Hexahistidine

<400> 84

Gly His His His His His His Gly

1 5

<210> 85

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> V12

<400> 85

Thr Asn Gly Ser Val Leu Arg Glu Phe Thr Leu Leu Glu Leu Glu Val

1 5 10 15

Val Thr Arg Asn Thr Glu Leu

20

<210> 86

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> T1

<400> 86

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

1 5 10 15

Tyr Phe Thr Ile Leu

20

<210> 87

<211> 1176

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4N-1177

<400> 87

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile

225 230 235 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val

245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser

260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile

275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met

290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser

305 310 315 320

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

325 330 335

Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp

340 345 350

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser

355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys

370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr

385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser

405 410 415
Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu

420 425 430
Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met

435 440 445
Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu

450 455 460
Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn

465 470 475 480
Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn

485 490 495
Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

500 505 510
Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr

515 520 525
Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu

530 535 540
Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr

545 550 555 560
Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

565 570 575
Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly

580 585 590
Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe

595 600 605
Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val

610 615 620
Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro

625 630 635 640
Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro

645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

900 905 910
 Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
 915 920 925
 Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

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

995 1000 1005
 Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
 1010 1015 1020
 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055
 Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070
 Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
 1075 1080 1085
 Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
 1090 1095 1100
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly

1125 1130 1135
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
 1140 1145 1150

Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
1155 1160 1165
Gln Arg Lys Gly Ser Glu Gly Thr
1170 1175
<210> 88
<211> 1097
<212> PRT
<213> Artificial Sequence
<220>
<223> ABCA4C-1177
<400> 88
Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val
1 5 10 15
Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu
20 25 30
Met Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys
35 40 45
Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His
50 55 60
Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp
65 70 75 80
Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile
85 90 95
Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly
100 105 110
Gly Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu
115 120 125
Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys
130 135 140
Ser Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu
145 150 155 160
Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu

165 170 175
 Gln His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg
 180 185 190
 Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val

 195 200 205
 Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr
 210 215 220
 Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe
 225 230 235 240
 Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp
 245 250 255
 Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly

 260 265 270
 Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro
 275 280 285
 Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr
 290 295 300
 Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu
 305 310 315 320
 Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro

 325 330 335
 Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg
 340 345 350
 Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser
 355 360 365
 Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile
 370 375 380
 Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu

 385 390 395 400
 Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly
 405 410 415

Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His
420 425 430

Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp
435 440 445

His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg
450 455 460

Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val
465 470 475 480

Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr
485 490 495

Val Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe
500 505 510

Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu
515 520 525

Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro
530 535 540

Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser
545 550 555 560

Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys
565 570 575

Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu
580 585 590

Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu
595 600 605

Phe Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu
610 615 620

Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe
625 630 635 640

Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu
645 650 655

Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu

660 665 670
 Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu
 675 680 685
 His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe
 690 695 700
 Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val

 705 710 715 720
 Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu
 725 730 735
 Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile
 740 745 750
 Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr
 755 760 765
 Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val

 770 775 780
 Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala
 785 790 795 800
 Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr
 805 810 815
 Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser
 820 825 830
 Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp

 835 840 845
 Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg
 850 855 860
 Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys
 865 870 875 880
 Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser
 885 890 895
 Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys

 900 905 910

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

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

1025 1030 1035 1040
 Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu
 1045 1050 1055
 Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe
 1060 1065 1070
 Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala
 1075 1080 1085
 Ala Gly Ala Ser Arg Gln Ala Gln Asp
 1090 1095

<210> 89

<211> 1178

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4N-1179

<400> 89

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

 50 55 60
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu

 115 120 125
 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala

 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val

 245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser

 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys

 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met

 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu

500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp

 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro

 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser

 690 695 700
 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720
 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile

755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp

770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val

785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly

805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp

820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp

850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp

865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys

885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu

965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005
 Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

 1010 1015 1020
 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055
 Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
 1060 1065 1070
 Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

 1075 1080 1085
 Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
 1090 1095 1100
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly

 1140 1145 1150
 Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
 1155 1160 1165
 Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser
 1170 1175
 <210> 90
 <211> 1095
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> ABCA4C-1179
 <400> 90
 Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His Val Asp Asp

1 5 10 15

 Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp
 20 25 30
 Val Val Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly
 35 40 45
 Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe Lys His Arg Ala
 50 55 60
 Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly
 65 70 75 80

 Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu
 85 90 95
 Lys Val Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala
 100 105 110
 Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro
 115 120 125
 Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro
 130 135 140

 Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu
 145 150 155 160
 Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln His
 165 170 175
 Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg Ser His
 180 185 190
 Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe Val Phe Leu
 195 200 205

 Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala
 210 215 220
 Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser
 225 230 235 240
 Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu Ala Asp Val Leu
 245 250 255

Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu
 260 265 270

Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr Pro Ser Val
 275 280 285

Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp Thr Gln Val
 290 295 300

Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys Leu Thr Met
 305 310 315 320

Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg
 325 330 335

Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile
 340 345 350

Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu
 355 360 365

Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile
 370 375 380

Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly
 385 390 395 400

Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile
 405 410 415

Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu
 420 425 430

Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala
 435 440 445

Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser
 450 455 460

Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser
 465 470 475 480

Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu
 485 490 495

Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser Met

500 505 510
 Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val
 515 520 525

 Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr
 530 535 540
 Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser
 545 550 555 560
 Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr
 565 570 575
 Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Tyr
 580 585 590

 Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp
 595 600 605
 Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile
 610 615 620
 Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn
 625 630 635 640
 Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile
 645 650 655

 Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu
 660 665 670
 Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser
 675 680 685
 Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met
 690 695 700
 Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg
 705 710 715 720

 His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile
 725 730 735
 Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg Ile Ile Thr
 740 745 750

Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu Thr Lys Ile
755 760 765

Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val Gly Val
770 775 780

Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys
785 790 795 800

Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly
805 810 815

Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val
820 825 830

His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu
835 840 845

Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val
850 855 860

Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu
865 870 875 880

Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly
885 890 895

Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro
900 905 910

Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg
915 920 925

Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala
930 935 940

Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr
945 950 955 960

Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile
965 970 975

Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys
980 985 990

Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu

995 1000 1005
 Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His
 1010 1015 1020
 Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile
 1025 1030 1035 1040

 Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr
 1045 1050 1055
 Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys
 1060 1065 1070
 Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly
 1075 1080 1085
 Ala Ser Arg Gln Ala Gln Asp
 1090 1095
 <210> 91
 <211> 1277
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> ABCA4-1177-CfaN
 <400> 91
 Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
 1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
 50 55 60

 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125

 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190

 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320

 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp

340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380

 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445

 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510

 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575

 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590

Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

835	840	845	
Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp			
850	855	860	
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp			
865	870	875	880
Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys			
885	890	895	
Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly			
900	905	910	
Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly			
915	920	925	
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro			
930	935	940	
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala			
945	950	955	960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu			
965	970	975	
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg			
980	985	990	
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys			
995	1000	1005	
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met			
1010	1015	1020	
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu			
1025	1030	1035	1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn			
1045	1050	1055	
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val			
1060	1065	1070	
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro			
1075	1080	1085	

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
1125 1130 1135

Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
1140 1145 1150

Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
1155 1160 1165

Gln Arg Lys Gly Ser Glu Gly Thr Cys Leu Ser Tyr Asp Thr Glu Ile
1170 1175 1180

Leu Thr Val Glu Tyr Gly Phe Leu Pro Ile Gly Lys Ile Val Glu Glu
1185 1190 1195 1200

Arg Ile Glu Cys Thr Val Tyr Thr Val Asp Lys Asn Gly Phe Val Tyr
1205 1210 1215

Thr Gln Pro Ile Ala Gln Trp His Asn Arg Gly Glu Gln Glu Val Phe
1220 1225 1230

Glu Tyr Cys Leu Glu Asp Gly Ser Ile Ile Arg Ala Thr Lys Asp His
1235 1240 1245

Lys Phe Met Thr Thr Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe
1250 1255 1260

Glu Arg Gly Leu Asp Leu Lys Gln Val Asp Gly Leu Pro
1265 1270 1275

<210>

92

<211> 1133

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1177-CfaCmut

<400> 92

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

1 5 10 15
 Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu
 20 25 30
 Val Ala Ser Asn Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys
 35 40 45
 Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp

 50 55 60
 Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu Ala Lys
 65 70 75 80
 Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys
 85 90 95
 Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu
 100 105 110
 Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro

 115 120 125
 Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly Pro
 130 135 140
 Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg
 145 150 155 160
 His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp
 165 170 175
 Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln

 180 185 190
 Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr
 195 200 205
 Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg Phe Gln
 210 215 220
 His Thr Ile Arg Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro
 225 230 235 240
 Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro

 245 250 255

Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln
 260 265 270
 Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr
 275 280 285
 Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys
 290 295 300
 Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro

 305 310 315 320
 Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys
 325 330 335
 Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr
 340 345 350
 Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly
 355 360 365
 Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp

 370 375 380
 Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala
 385 390 395 400
 Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg
 405 410 415
 Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr
 420 425 430
 Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn

 435 440 445
 Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp
 450 455 460
 Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn
 465 470 475 480
 Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn
 485 490 495
 Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr

500 505 510
 Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu
 515 520 525
 Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile
 530 535 540
 Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr
 545 550 555 560
 Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser

 565 570 575
 Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile
 580 585 590
 Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly
 595 600 605
 Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val
 610 615 620
 Ala Leu Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro

 625 630 635 640
 Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser
 645 650 655
 Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile
 660 665 670
 Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val
 675 680 685
 Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly

 690 695 700
 Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg
 705 710 715 720
 Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly
 725 730 735
 Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu
 740 745 750

Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu

755 760 765

Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu

770 775 780

Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu

785 790 795 800

His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp

805 810 815

Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly

820 825 830

Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp

835 840 845

Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu

850 855 860

Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe

865 870 875 880

Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr

885 890 895

Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn

900 905 910

Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala

915 920 925

Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala

930 935 940

Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly

945 950 955 960

Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile

965 970 975

Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu Glu

980 985 990

Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe

995 1000 1005
 Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly

 1010 1015 1020
 Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro
 1025 1030 1035 1040
 Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser

 1045 1050 1055
 Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser
 1060 1065 1070
 Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser

 1075 1080 1085
 Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val
 1090 1095 1100
 Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu
 1105 1110 1115 1120
 His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp

 1125 1130

 <210> 93
 <211> 1279
 <212> PRT
 <213> Artificial Sequence
 <220>

 <223> ABCA4-1179-CfaN
 <400> 93
 Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
 1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125

 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190

 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
325 330 335

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

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525

Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540

Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr

545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575

 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640

 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700

 Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
 705 710 715 720
 His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
 725 730 735
 Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
 740 745 750
 Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
 755 760 765

 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780
 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly

805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp

820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala

835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp

850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp

865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys

885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu

965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

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

<210> 94

<211> 1131

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1179-CfaCmut

<400> 94

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

1 5 10 15

Asp Ile Gly Val Gly Glu Pro His Asn Phe Leu Leu Lys Asn Gly Leu

20 25 30

Val Ala Ser Asn Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala

35 40 45

His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn

50 55 60

Glu Leu Met Asp Val Val Leu His His Val Pro Glu Ala Lys Leu Val

65 70 75 80

Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro Asn Lys Asn Phe

85 90 95

Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu

100 105 110

Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp Thr Pro Leu Glu

115 120 125

Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser Gly Pro Leu Phe

130 135 140

Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn Pro Arg His Pro

145 150 155 160

Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn

165 170 175

Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro

180 185 190

Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu

195 200 205

Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr

210 215 220
 Ile Arg Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr
 225 230 235 240
 Phe Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly

 245 250 255
 Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln Tyr
 260 265 270
 Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr Val Leu
 275 280 285
 Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg Cys Leu Lys
 290 295 300
 Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys

 305 310 315 320
 Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys
 325 330 335
 Trp Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu
 340 345 350
 Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro
 355 360 365
 Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr

 370 375 380
 Asp Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile
 385 390 395 400
 Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly
 405 410 415
 Gly Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu
 420 425 430
 Ala Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser

 435 440 445
 Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu
 450 455 460

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

705 710 715 720
 Glu Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn
 725 730 735
 Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu
 740 745 750
 Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr

 755 760 765
 Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln
 770 775 780
 Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu
 785 790 795 800
 Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu
 805 810 815
 Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn

 820 825 830
 Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr
 835 840 845
 Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn
 850 855 860
 Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala
 865 870 875 880
 Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg

 885 890 895
 Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser
 900 905 910
 Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr
 915 920 925
 Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile
 930 935 940
 Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp

 945 950 955 960

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

1010 1015 1020
 Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu
 1025 1030 1035 1040
 Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln
 1045 1050 1055
 Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser
 1060 1065 1070
 Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu

1075 1080 1085
 Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val
 1090 1095 1100
 Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro
 1105 1110 1115 1120
 Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp
 1125 1130

<210> 95

<211> 1095

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4N-1096

<400> 95

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
 1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
 50 55 60

 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
 65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
 115 120 125

 Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190

 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255

 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
325 330 335

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

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445

Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460

Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480

Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495

Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510

Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr

515	520	525
Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu		
530	535	540
Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr		
545	550	555
Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp		
565	570	575
Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly		
580	585	590
Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe		
595	600	605
Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val		
610	615	620
Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro		
625	630	635
		640
Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro		
645	650	655
Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys		
660	665	670
Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn		
675	680	685
Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser		
690	695	700
Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met		
705	710	715
His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe		
725	730	735
Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser		
740	745	750
Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile		
755	760	765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
 770 775 780
 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
 785 790 795 800
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
 805 810 815
 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
 820 825 830

 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
 835 840 845
 Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
 850 855 860
 Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
 865 870 875 880
 Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
 885 890 895

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

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

1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

 1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

 1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

 1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr

 1090 1095

<210> 96

<211> 1178

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4C-1096

<400> 96

Ser Arg Arg Ser Ile Trp Asp Leu Leu Leu Lys Tyr Arg Ser Gly Arg

1 5 10 15

Thr Ile Ile Met Ser Thr His His Met Asp Glu Ala Asp Leu Leu Gly

 20 25 30

Asp Arg Ile Ala Ile Ile Ala Gln Gly Arg Leu Tyr Cys Ser Gly Thr

 35 40 45

Pro Leu Phe Leu Lys Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr Leu

 50 55 60

Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg Lys Gly Ser Glu Gly

65 70 75 80

Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys Pro Ala His

 85 90 95

Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu

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

Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser Thr Pro Trp Lys Thr
355 360 365

Pro Ser Val Ser Pro Asn Ile Thr Gln Leu Phe Gln Lys Gln Lys Trp
370 375 380

Thr Gln Val Asn Pro Ser Pro Ser Cys Arg Cys Ser Thr Arg Glu Lys
385 390 395 400

Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro
405 410 415

Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp
420 425 430

Arg Asn Ile Ser Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg
435 440 445

Ser Ser Leu Lys Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly
450 455 460

Ile Ser Ile Gly Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala
465 470 475 480

Leu Val Gly Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly
485 490 495

Gly Pro Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys
500 505 510

His Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly
515 520 525

Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu
530 535 540

Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr
545 550 555 560

Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile
565 570 575

Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile
580 585 590

Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln

595 600 605
 Glu Arg Val Asn Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser
 610 615 620

 Pro Thr Thr Tyr Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr
 625 630 635 640
 Ser Val Ser Ala Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys
 645 650 655
 Lys Ala Tyr Thr Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu
 660 665 670
 Leu Leu Tyr Gly Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe
 675 680 685

 Leu Phe Asp Val Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn
 690 695 700
 Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu
 705 710 715 720
 Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys
 725 730 735
 Leu Leu Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp
 740 745 750

 Leu Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu
 755 760 765
 Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu
 770 775 780
 Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr Leu Leu
 785 790 795 800
 Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu Pro Thr Lys
 805 810 815

 Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu Glu Arg Gln Arg
 820 825 830
 Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu Arg Leu His Glu Leu
 835 840 845

Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys
850 855 860
Val Gly Val Arg Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly
865 870 875 880

Ala Gly Lys Thr Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val
885 890 895
Thr Ser Gly Asp Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile
900 905 910
Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile
915 920 925
Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu
930 935 940

Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile
945 950 955 960
Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr
965 970 975
Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly
980 985 990
Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro
995 1000 1005

Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu
1010 1015 1020
Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala
1025 1030 1035 1040
Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met
1045 1050 1055
Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val
1060 1065 1070

Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn
1075 1080 1085
Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg

1090 1095 1100
 Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu
 1105 1110 1115 1120
 Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile
 1125 1130 1135

Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn
 1140 1145 1150
 Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg
 1155 1160 1165
 Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp
 1170 1175

<210> 97

<211> 1184

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4N-1185

<400> 97

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
 20 25 30
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
 35 40 45
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
 50 55 60
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn
 85 90 95
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
 100 105 110

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

355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr

 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu

 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr

 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly

 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605

Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640

Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655

Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670

Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp

850 855 860
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880
Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895
Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

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

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

1025 1030 1035 1040
Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
1045 1050 1055
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
1060 1065 1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
1075 1080 1085
Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
 1140 1145 1150
 Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
 1155 1160 1165
 Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe
 1170 1175 1180
 <210> 98
 <211> 1089
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> ABCA4C-1185
 <400> 98
 Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val
 1 5 10 15
 Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val
 20 25 30
 Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu
 35 40 45
 Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg
 50 55 60
 Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile
 65 70 75 80
 Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser
 85 90 95
 Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn
 100 105 110
 Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln

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

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

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

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

 Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly
 545 550 555 560
 Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu
 565 570 575
 Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro
 580 585 590
 Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr
 595 600 605

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

610 615 620
 Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg
 625 630 635 640
 Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys
 645 650 655
 Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp
 660 665 670

 Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp
 675 680 685
 Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val
 690 695 700
 Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln
 705 710 715 720
 Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp
 725 730 735

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

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

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

 Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His
 930 935 940
 Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val
 945 950 955 960
 Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys
 965 970 975
 Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp
 980 985 990

 Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn
 995 1000 1005
 Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe
 1010 1015 1020
 Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser
 1025 1030 1035 1040
 His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr
 1045 1050 1055

 Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His
 1060 1065 1070
 Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln
 1075 1080 1085
 Asp

<210> 99

<211> 1183

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1096-gp41N

<400> 99

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270

 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335

 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
 370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400

 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu

450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495
 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525

 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590

 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655

 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670
 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685
 Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
 690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
 965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
 980 985 990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
 995 1000 1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
 1010 1015 1020
Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055
Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
 1060 1065 1070
Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
 1075 1080 1085
Thr Ser Gly Val Asp Pro Tyr Cys Leu Asp Leu Lys Thr Gln Val Gln
1090 1095 1100

Thr Pro Gln Gly Met Lys Glu Ile Ser Asn Ile Gln Val Gly Asp Leu
1105 1110 1115 1120
Val Leu Ser Asn Thr Gly Tyr Asn Glu Val Leu Asn Val Phe Pro Lys
 1125 1130 1135
Ser Lys Lys Lys Ser Tyr Lys Ile Thr Leu Glu Asp Gly Lys Glu Ile
 1140 1145 1150
Ile Cys Ser Glu Glu His Leu Phe Pro Thr Gln Thr Gly Glu Met Asn
1155 1160 1165

Ile Ser Gly Gly Leu Lys Glu Gly Met Cys Leu Tyr Val Lys Glu
1170 1175 1180

<210> 100

<211> 1215

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1096-gp41C

<400> 100

Met Met Leu Lys Lys Ile Leu Lys Ile Glu Glu Leu Asp Glu Arg Glu

1 5 10 15

Leu Ile Asp Ile Glu Val Ser Gly Asn His Leu Phe Tyr Ala Asn Asp

20 25 30

Ile Leu Thr His Asn Ser Arg Arg Ser Ile Trp Asp Leu Leu Leu Lys

35 40 45

Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met Asp Glu

50 55 60

Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly Arg Leu

65 70 75 80

Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly Thr Gly

85 90 95

Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg

100 105 110

Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr

115 120 125

Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp

130 135 140

Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu

145 150 155 160

Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu Pro

165 170 175

Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg Glu Leu

180 185 190

Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile Ser Asp

195 200 205

Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser Asp Ser

210 215 220
 Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn Val Asn

 225 230 235 240
 Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln Thr Pro
 245 250 255
 Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His Pro Glu
 260 265 270
 Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu Asn Thr
 275 280 285
 Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val Lys Arg

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

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

 420 425 430
 Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly Ala
 435 440 445
 Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Ile Leu
 450 455 460

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

705 710 715 720
 Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr Val Ala
 725 730 735
 Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile Thr

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

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

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

 930 935 940
 Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr Cys Pro
 945 950 955 960

Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His Leu Tyr

965 970 975

Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu Lys Val

980 985 990

Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala Asp Cys

995 1000 1005

Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala

1010 1015 1020

Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu Pro Thr

1025 1030 1035 1040

Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val

1045 1050 1055

Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met

1060 1065 1070

Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly

1075 1080 1085

Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly

1090 1095 1100

Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu

1105 1110 1115 1120

Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe Pro

1125 1130 1135

Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe Gln Val

1140 1145 1150

Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser His Lys

1155 1160 1165

Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu Asp

1170 1175 1180

Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His Asp Leu

1185 1190 1195 1200

Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln Asp

1205 1210 1215

<210> 101

<211> 1272

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1185-gp41N

<400> 101

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu

130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu

145 150 155 160

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

165 170 175

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

180 185 190

His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
195 200 205

Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
225 230 235 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
325 330 335

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

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415

Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430

Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met

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

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

 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560
 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605

 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620
 Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
 625 630 635 640
 Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
 645 650 655
 Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
 660 665 670

 Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
 675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700

Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720

His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735

Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765

Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780

Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800

Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830

Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845

Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860

Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

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

 Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
 995 1000 1005
 Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met
 1010 1015 1020
 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu
 1025 1030 1035 1040
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn
 1045 1050 1055

 Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val
 1060 1065 1070
 Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro
 1075 1080 1085
 Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu
 1090 1095 1100
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met
 1105 1110 1115 1120

 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
 1125 1130 1135
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
 1140 1145 1150
 Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
 1155 1160 1165
 Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe
 1170 1175 1180

Cys Leu Asp Leu Lys Thr Gln Val Gln Thr Pro Gln Gly Met Lys Glu
 1185 1190 1195 1200
 Ile Ser Asn Ile Gln Val Gly Asp Leu Val Leu Ser Asn Thr Gly Tyr
 1205 1210 1215
 Asn Glu Val Leu Asn Val Phe Pro Lys Ser Lys Lys Lys Ser Tyr Lys
 1220 1225 1230
 Ile Thr Leu Glu Asp Gly Lys Glu Ile Ile Cys Ser Glu Glu His Leu
 1235 1240 1245

Phe Pro Thr Gln Thr Gly Glu Met Asn Ile Ser Gly Gly Leu Lys Glu
 1250 1255 1260

Gly Met Cys Leu Tyr Val Lys Glu

1265 1270

<210> 102

<211> 1126

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4-1185-gp41C

<400> 102

Met Met Leu Lys Lys Ile Leu Lys Ile Glu Glu Leu Asp Glu Arg Glu

1 5 10 15

Leu Ile Asp Ile Glu Val Ser Gly Asn His Leu Phe Tyr Ala Asn Asp

20 25 30

Ile Leu Thr His Asn Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu

35 40 45

Thr Pro Glu Gln Val Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val

50 55 60

Val Leu His His Val Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln

65 70 75 80

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

85 90 95

Ala Ser Leu Phe Arg Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu

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

Pro Glu Cys Pro Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr
355 360 365

Gln Arg Ser Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser
370 375 380

Asp Phe Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys
385 390 395 400

Ser Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly
405 410 415

Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe
420 425 430

Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr
435 440 445

Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr
450 455 460

Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu
465 470 475 480

Val Ser Phe Leu Asn Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu
485 490 495

Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln
500 505 510

Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr
515 520 525

Thr Ser Val Asp Ala Val Val Ala Ile Cys Val Ile Phe Ser Met Ser
530 535 540

Phe Val Pro Ala Ser Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn
545 550 555 560

Lys Ser Lys His Leu Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr
565 570 575

Trp Val Thr Asn Phe Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala
580 585 590

Gly Leu Val Val Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr

595 600 605
 Ser Pro Glu Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly
 610 615 620
 Trp Ala Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val
 625 630 635 640
 Pro Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly
 645 650 655
 Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn
 660 665 670

 Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val
 675 680 685
 Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser
 690 695 700
 Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala
 705 710 715 720
 Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val
 725 730 735

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

 Pro Gly Thr Ser Ser Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg
 805 810 815
 Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr
 820 825 830
 Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp
 835 840 845

Ala Thr Val Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His
850 855 860

Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu
865 870 875 880

Thr Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro
885 890 895

Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly
900 905 910

Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn
915 920 925

Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu
930 935 940

Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg
945 950 955 960

Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val
965 970 975

Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg
980 985 990

Leu Ala Ile Met Val Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln
995 1000 1005

His Leu Lys Ser Lys Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile
1010 1015 1020

Lys Ser Pro Lys Asp Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln
1025 1030 1035 1040

Phe Phe Gln Gly Asn Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr
1045 1050 1055

Asn Met Leu Gln Phe Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe
1060 1065 1070

Gln Leu Leu Leu Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser
1075 1080 1085

Val Thr Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln

1090 1095 1100
 Gln Thr Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala
 1105 1110 1115 1120

Ser Arg Gln Ala Gln Asp
 1125

<210> 103

<211> 22

<212> PRT

<213> Artificial Sequence

<220>

<223> 3FT

<400> 103

Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr

1 5 10 15

Lys Asp Asp Asp Asp Lys

20

<210> 104

<211> 37

<212> PRT

<213> Artificial Sequence

<220>

<223> GP41C

<400> 104

Met Met Leu Lys Lys Ile Leu Lys Ile Glu Glu Leu Asp Glu Arg Glu

1 5 10 15

Leu Ile Asp Ile Glu Val Ser Gly Asn His Leu Phe Tyr Ala Asn Asp

20

25

30

Ile Leu Thr His Asn

35

<210> 105

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> ConC

<400> 105

Met Val Lys Ile Ile Ser Arg Gln Ser Leu Gly Lys Gln Asn Val Tyr

1 5 10 15

Asp Ile Gly Val Glu Lys Asp His Asn Phe Leu Leu Ala Asn Gly Leu

20 25 30

Ile Ala Ser Asn

35

<210> 106

<211> 2273

<212> PRT

<213> Artificial Sequence

<220>

<223> ABCA4

<400> 106

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr

1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro

20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu

35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala

50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro

65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu

100 105 110

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

115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
 130 135 140
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
 145 150 155 160
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
 165 170 175

 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
 180 185 190
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
 195 200 205
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
 210 215 220
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
 225 230 235 240

 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
 245 250 255
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
 260 265 270
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
 275 280 285
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
 290 295 300

 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
 305 310 315 320
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
 325 330 335
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
 340 345 350
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
 355 360 365

 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys

370 375 380
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
 385 390 395 400
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
 405 410 415
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
 420 425 430

 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
 435 440 445
 Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
 450 455 460
 Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
 465 470 475 480
 Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
 485 490 495

 Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
 500 505 510
 Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
 515 520 525
 Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
 530 535 540
 Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
 545 550 555 560

 Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
 565 570 575
 Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
 580 585 590
 Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
 595 600 605
 Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
 610 615 620

Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640
Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655
Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670
Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685

Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700
Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met
705 710 715 720
His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735
Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765
Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780
Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800
Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815

Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830
Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845
Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp

865 870 875 880

Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys

885 890 895

Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly

900 905 910

Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly

915 920 925

Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro

930 935 940

Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala

945 950 955 960

Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu

965 970 975

Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg

980 985 990

Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys

995 1000 1005

Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met

1010 1015 1020

Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu

1025 1030 1035 1040

Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn

1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val

1060 1065 1070

Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro

1075 1080 1085

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu

1090 1095 1100

Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met

1105 1110 1115 1120

Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly
1125 1130 1135

Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly
1140 1145 1150

Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser
1155 1160 1165

Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe
1170 1175 1180

Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val
1185 1190 1195 1200

Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val
1205 1210 1215

Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu
1220 1225 1230

Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg
1235 1240 1245

Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile
1250 1255 1260

Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser
1265 1270 1275 1280

Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn
1285 1290 1295

Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln
1300 1305 1310

Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His
1315 1320 1325

Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu
1330 1335 1340

Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val
1345 1350 1355 1360

Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala Gln

1365	1370	1375	
Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile			
1380	1385	1390	
Val Ile Pro Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro Trp			
1395	1400	1405	
Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser			
1410	1415	1420	
Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe			
1425	1430	1435	1440
Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly			
1445	1450	1455	
Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln			
1460	1465	1470	
Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys			
1475	1480	1485	
Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu			
1490	1495	1500	
Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu			
1505	1510	1515	1520
Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val Lys			
1525	1530	1535	
Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val			
1540	1545	1550	
Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val			
1555	1560	1565	
Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly			
1570	1575	1580	
Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys			
1585	1590	1595	1600
Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys			
1605	1610	1615	

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

1620 1625 1630

Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser

1635 1640 1645

Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr

1650 1655 1660

Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala

1665 1670 1675 1680

Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser

1685 1690 1695

Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu

1700 1705 1710

Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe

1715 1720 1725

Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly

1730 1735 1740

Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu

1745 1750 1755 1760

Pro Ala Leu Val Ala Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro

1765 1770 1775

Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr

1780 1785 1790

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

1795 1800 1805

Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg

1810 1815 1820

Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys

1825 1830 1835 1840

Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp

1845 1850 1855

Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp

1860	1865	1870	
Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val			
1875	1880	1885	
Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln			
1890	1895	1900	
Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp			
1905	1910	1915	1920
Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp			
1925	1930	1935	
Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser			
1940	1945	1950	
Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe			
1955	1960	1965	
Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met			
1970	1975	1980	
Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly			
1985	1990	1995	2000
Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr			
2005	2010	2015	
Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His			
2020	2025	2030	
Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu			
2035	2040	2045	
Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala			
2050	2055	2060	
Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser			
2065	2070	2075	2080
Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu			
2085	2090	2095	
Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val			
2100	2105	2110	

Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His
 2115 2120 2125
 Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val
 2130 2135 2140
 Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys
 2145 2150 2155 2160

Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp
 2165 2170 2175
 Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn
 2180 2185 2190
 Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe
 2195 2200 2205
 Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser
 2210 2215 2220

His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr
 2225 2230 2235 2240
 Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His
 2245 2250 2255
 Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln
 2260 2265 2270
 Asp

<210> 107

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> T2

<400> 107

Lys Leu Lys Gly Leu Gly Lys Arg Cys Lys Arg Arg Glu Asp Leu Glu

1 5 10 15
 Ile Arg Phe Ile Leu

20

<210> 108

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> T3

<400> 108

Trp Lys Leu Leu Leu Trp Val Gly Leu Val Leu Val Leu Lys His His

1 5 10 15

Asp Gly Ala Ala His

20