(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년05월15일
(11) 등록번호 10-2808368
(24) 등록일자 2025년05월12일

(51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) *A61P 25/08* (2006.01)
A61P 25/16 (2006.01) *A61P 25/28* (2006.01)
C07K 14/47 (2006.01) *C07K 14/705* (2006.01)
C12N 15/86 (2006.01)
(52) CPC특허분류
C12N 15/1138 (2013.01)
A61P 25/08 (2018.01)
(21) 출원번호 10-2021-7043048
(22) 출원일자(국제) 2020년05월29일
심사청구일자 2022년11월14일
(85) 번역문제출일자 2021년12월29일
(65) 공개번호 10-2022-0066225
(43) 공개일자 2022년05월24일
(86) 국제출원번호 PCT/US2020/035431
(87) 국제공개번호 WO 2020/243651
국제공개일자 2020년12월03일
(30) 우선권주장
62/854,238 2019년05월29일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
JP2018088888 A
WO2018187363 A1

(73) 특허권자
엔코디드 테라퓨틱스, 인크.
미국 94080 캘리포니아주 사우스 샌프란시스코 오이스터 포인트 블러바드 341
(72) 발명자
태글리아텔라, 스테파니
미국 94080 캘리포니아주 사우스 샌프란시스코 오이스터 포인트 블러바드 341
태넌하우스, 앤
미국 94080 캘리포니아주 사우스 샌프란시스코 오이스터 포인트 블러바드 341
(뒷면에 계속)
(74) 대리인
양영준, 이상남

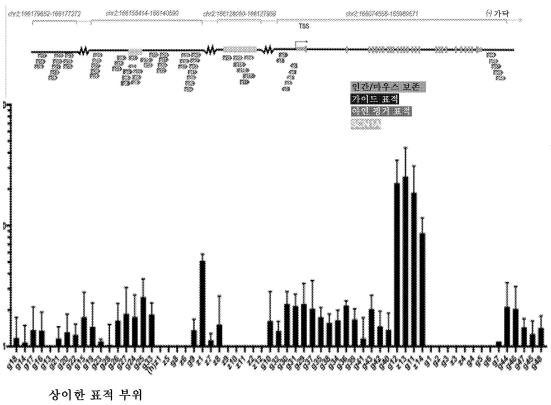
전체 청구항 수 : 총 27 항 심사관 : 정이든

(54) 발명의 명칭 **선택적 유전자 조절을 위한 조성물 및 방법**

(57) 요 약

SCN1a의 선택적 상향조절을 위한 조작된 전사 인자 및 질환 및 장애, 예컨대 드라베 증후군을 치료하기 위한 그의 용도가 본원에 제공된다. 또한, 마이크로RNA 결합 부위 및 파르브알부민 뉴런에서의 선택적 발현을 위한 그의 용도가 제공된다.

대 표 도



(52) CPC특허분류
A61P 25/16 (2018.01)
A61P 25/28 (2018.01)
C07K 14/4705 (2013.01)
C07K 14/705 (2013.01)
C12N 15/86 (2013.01)
C12N 2310/20 (2025.05)
C12N 2740/16043 (2013.01)
C12N 2750/14143 (2013.01)
C12N 2830/008 (2013.01)

(72) 발명자
라마무르티, 카르틱
미국 94080 캘리포니아주 사우스 샌프란시스코 오이스터 포인트 블러바드 341
영, 앤드류
미국 94080 캘리포니아주 사우스 샌프란시스코 오이스터 포인트 블러바드 341
오버코플러, 데이비드
미국 94080 캘리포니아주 사우스 샌프란시스코 오이스터 포인트 블러바드 341

(30) 우선권주장
62/857,727 2019년06월05일 미국(US)
63/008,569 2020년04월10일 미국(US)

명 세 서

청구범위

청구항 1

조작된 전사 인자를 코딩하는 핵산 서열을 포함하는 폴리뉴클레오티드이며, 여기서 조작된 전사 인자는 서열식별번호: 127의 아미노산 서열을 포함하는 것인 폴리뉴클레오티드.

청구항 2

제1항에 있어서, 핵산 서열이, 조작된 전사 인자를 코딩하는 핵산 서열에 작동가능하게 연결됨으로써 발현 카세트를 형성하는 조절 요소를 추가로 포함하는 것인 폴리뉴클레오티드.

청구항 3

제2항에 있어서, 조절 요소가 GAD2 프로모터, 인간 시냅신 프로모터, CBA 프로모터, CMV 프로모터, minCMV 프로모터, TATA 박스, 슈퍼 코어 프로모터 또는 EF1α 프로모터, 또는 그의 조합을 포함하는 것인 폴리뉴클레오티드.

청구항 4

제2항에 있어서, 조절 요소가 파르브알부민 (PV) 뉴런 선택적 조절 요소인 폴리뉴클레오티드.

청구항 5

제2항에 있어서, 조절 요소가 서열식별번호: 1-4 중 어느 하나를 포함하는 것인 폴리뉴클레오티드.

청구항 6

제5항에 있어서, 조절 요소가 서열식별번호: 2를 포함하는 것인 폴리뉴클레오티드.

청구항 7

제2항에 있어서, 발현 카세트가 흥분성 뉴런에서 조작된 전사 인자의 발현을 억제하는 요소를 추가로 포함하며, 여기서 흥분성 뉴런에서 조작된 전사 인자의 발현을 억제하는 요소는 서열식별번호: 7을 포함하는 것인 폴리뉴클레오티드.

청구항 8

제7항에 있어서, 서열식별번호: 71의 핵산 서열을 포함하는 폴리뉴클레오티드.

청구항 9

제2항 내지 제8항 중 어느 한 항의 폴리뉴클레오티드를 포함하는 발현 벡터.

청구항 10

제9항에 있어서, 바이러스 벡터인 발현 벡터.

청구항 11

제10항에 있어서, 바이러스 벡터가 아데노-연관 바이러스 (AAV)인 발현 벡터.

청구항 12

제11항에 있어서, AAV가 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, rh10, 조류 AAV, 소 AAV, 개 AAV, 말 AAV, 영장류 AAV, 비-영장류 AAV, 양 AAV, scAAV, scAAV1, scAAV2, scAAV5, scAAV8, scAAV9, 및 그의 임의의 것의 하이브리드로 이루어진 군으로부터 선택되는 혈청형을 갖는 것인 발현 벡터.

청구항 13

제11항에 있어서, 발현 카세트가 5' AAV 역전된 말단 반복 (ITR) 서열 및 3' AAV ITR 서열을 추가로 포함하는 것인 발현 벡터.

청구항 14

제13항에 있어서, 5' AAV ITR 서열 및 3' AAV ITR 서열이 각각, 독립적으로, AAV1, AAV2, AAV5, AAV8, 또는 AAV9로부터 유래된 ITR 서열인 발현 벡터.

청구항 15

제14항에 있어서, 발현 카세트가 서열식별번호: 71을 포함하는 것인 발현 벡터.

청구항 16

제15항에 있어서, 5' AAV 역전된 말단 반복 (ITR) 서열 및 3' AAV ITR 서열을 갖는 AAV를 포함하며, 여기서 AAV는 AAV9이고, 5' AAV ITR 서열 및 3' AAV ITR 서열은 AAV2로부터 유래된 것인 발현 벡터.

청구항 17

제2항 내지 제8항 중 어느 한 항의 폴리뉴클레오티드 또는 상기 폴리뉴클레오티드를 포함하는 발현 벡터를 포함하는 세포.

청구항 18

제17항에 있어서, 293 세포, A549 세포, 또는 HeLa 세포인 세포.

청구항 19

아데노-연관 바이러스 (AAV) 비리온을 제조하는 방법이며,

a) 재조합 AAV 비리온을 생산하기 위한 조건 하에 제17항에 따른 세포를 배양하는 단계;

b) 세포 배양물을 수거하는 단계; 및

c) 세포에 의해 생산된 재조합 AAV 비리온을 정제하는 단계

를 포함하는 방법.

청구항 20

SCN1A 반수체기능부전과 연관된 간질 또는 드라베 증후군을 치료하기 위한, 제9항의 발현 벡터 및 제약상 허용되는 담체를 포함하는 제약 조성물.

청구항 21

제20항에 있어서, 발현 벡터가 바이러스 벡터인 제약 조성물.

청구항 22

제21항에 있어서, 바이러스 벡터가 아데노-연관 바이러스 (AAV)인 제약 조성물.

청구항 23

제22항에 있어서, AAV가 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, rh10, 조류 AAV, 소 AAV, 개 AAV, 말 AAV, 영장류 AAV, 비-영장류 AAV, 양 AAV, scAAV, scAAV1, scAAV2, scAAV5, scAAV8, scAAV9, 및 그의 임의의 것의 하이브리드로 이루어진 군으로부터 선택되는 혈청형을 갖는 것인 제약 조성물.

청구항 24

제22항에 있어서, 발현 카세트가 5' AAV 역전된 말단 반복 (ITR) 서열 및 3' AAV ITR 서열을 추가로 포함하는

것인 제약 조성물.

청구항 25

제24항에 있어서, 5' AAV ITR 서열 및 3' AAV ITR 서열이 각각, 독립적으로, AAV1, AAV2, AAV5, AAV8, 또는 AAV9로부터 유래된 ITR 서열인 제약 조성물.

청구항 26

제25항에 있어서, 발현 카세트가 서열식별번호: 71을 포함하는 것인 제약 조성물.

청구항 27

제26항에 있어서, 발현 벡터가 5' AAV 역전된 말단 반복 (ITR) 서열 및 3' AAV ITR 서열을 갖는 AAV를 포함하며, 여기서 AAV는 AAV9이고, 5' AAV ITR 서열 및 3' AAV ITR 서열은 AAV2로부터 유래된 것인 제약 조성물.

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

청구항 111

삭제

청구항 112

삭제

청구항 113

삭제

청구항 114

삭제

청구항 115

삭제

청구항 116

삭제

청구항 117

삭제

청구항 118

삭제

청구항 119

삭제

청구항 120

삭제

청구항 121

삭제

청구항 122

삭제

청구항 123

삭제

청구항 124

삭제

청구항 125

삭제

청구항 126

삭제

청구항 127

삭제

청구항 128

삭제

청구항 129

삭제

청구항 130

삭제

청구항 131

삭제

청구항 132

삭제

청구항 133

삭제

청구항 134

삭제

청구항 135

삭제

청구항 136

삭제

청구항 137

삭제

청구항 138

삭제

청구항 139

삭제

청구항 140

삭제

청구항 141

삭제

청구항 142

삭제

청구항 143

삭제

청구항 144

삭제

청구항 145

삭제

청구항 146

삭제

청구항 147

삭제

청구항 148

삭제

청구항 149

삭제

청구항 150

삭제

청구항 151

삭제

청구항 152

삭제

청구항 153

삭제

청구항 154

삭제

청구항 155

삭제

청구항 156

삭제

청구항 157

삭제

청구항 158

삭제

청구항 159

삭제

청구항 160

삭제

청구항 161

삭제

청구항 162

삭제

청구항 163

삭제

청구항 164

삭제

청구항 165

삭제

청구항 166

삭제

청구항 167

삭제

청구항 168

삭제

청구항 169

삭제

청구항 170

삭제

청구항 171

삭제

청구항 172

삭제

청구항 173

삭제

청구항 174

삭제

청구항 175

삭제

청구항 176

삭제

청구항 177

삭제

청구항 178

삭제

청구항 179

삭제

청구항 180

삭제

청구항 181

삭제

청구항 182

삭제

청구항 183

삭제

청구항 184

삭제

청구항 185

삭제

청구항 186

삭제

청구항 187

삭제

청구항 188

삭제

청구항 189

삭제

청구항 190

삭제

청구항 191

삭제

청구항 192

삭제

청구항 193

삭제

청구항 194

삭제

청구항 195

삭제

청구항 196

삭제

청구항 197

삭제

청구항 198

삭제

청구항 199

삭제

청구항 200

삭제

청구항 201

삭제

청구항 202

삭제

청구항 203

삭제

청구항 204

삭제

청구항 205

삭제

청구항 206

삭제

발명의 설명

기 술 분 야

[0001] 상호 참조

[0002] 본 출원은 2019년 5월 29일에 출원된 미국 특허 가출원 번호 62/854,238; 2019년 6월 5일에 출원된 미국 특허 가출원 번호 62/857,727; 및 2020년 4월 10일에 출원된 미국 특허 가출원 번호 63/008,569를 우선권 주장하며, 이들 각각은 그 전문이 본원에 참조로 포함된다.

[0003] 서열 목록

[0004] 본 출원은 ASCII 포맷으로 전자 제출된 서열 목록을 함유하며, 이는 그 전문이 본원에 참조로 포함된다. 2020년 5월 28일에 생성된 상기 ASCII 카피는 파일명이 46482-724_601_SL.txt이고, 크기가 418,483 바이트이다.

배 경 기 술

[0005] 광범위한 인간 질환은 유전자의 비정상적 발현과 연관된다. 일부 경우에, 유전자 내의 유전자 돌연변이는 유전자가 이상조절되거나, 하향조절되거나 또는 전혀 발현되지 않도록 하여, 반수체기능부전을 초래한다. 일부 경우에, 유전자 내의 유전자 돌연변이는 유전자가 상향조절되도록 하여, 유전자의 과다발현을 초래한다. 유전 장애 또는 질환을 치료하는데 많은 도전과제가 존재한다. 하나의 접근법은 환자 세포 내로의 핵산의 치료적 전달을 수반하는 유전자 요법이다. 그러나, 유전자 요법과 연관된 다양한 도전과제, 예컨대 유전자 요법에 의해 도출된 원치않는 면역 반응, 오프-타겟 효과, 유전자 요법 비히클 (예를 들어, 바이러스)의 클로닝 능력에 대한 한계, 보다 긴 기간에 걸친 치료 효과의 지속 등은 여전히 해결되지 않고 있다. 중추 신경계 (CNS)는 유전자 및/또는 단백질 발현의 근본적인 손상을 다루는 요법의 개발을 위한 많은 고유한 도전과제를 제기한다. CNS 질환/장애의 증상의 관리를 돕는 약물이 있지만, 많은 CNS 질환/장애, 예를 들어 드라베 증후군은 특이적 치료법 또는 치유법이 없다. 따라서, 질환 또는 장애의 영향을 역전시키는 것을 돕기 위해 임의의 내인성 유전자의 발현을 조정할 수 있는 신규 조성물 및 방법, 특히 감소된 면역원성, 감소된 오프-타겟 효과, 표적 유전자에 대한 증가된 특이성 및/또는 증가된 치료 효능을 갖는 요법에 대한 필요가 존재한다.

발명의 내용

[0006] 한 측면에서, 본 출원은 세포에서 SCN1A 유전자의 발현을 증가시키는 비-자연 발생 전사 인자를 코딩하는 서열

을 포함하는 발현 카세트이며, 여기서 비-자연 발생 전사 인자는 적어도 2개의 전사 활성화 도메인 (TAD)에 작동가능하게 연결된 DNA 결합 도메인 (DBD)을 포함하고 이들은 하기 방식: TAD1-TAD2-DBD, DBD-TAD3-TAD4 또는 TAD1-TAD2-DBD-TAD3-TAD4로 연결된 것인 발현 카세트를 제공한다. 특정 실시양태에서, TAD1, TAD2, TAD3 및 TAD4는 하기: VP16, VP64, Viper, CITED2, CITED4, CREB3 또는 그의 기능적 단편으로부터 독립적으로 선택된다. 특정 실시양태에서, TAD1 및 TAD2는 동일한 TAD이다. 특정 실시양태에서, TAD1 및 TAD2는 CITED2 또는 그의 기능적 단편이다. 특정 실시양태에서, TAD1 및 TAD2는 CITED4 또는 그의 기능적 단편이다. 특정 실시양태에서, TAD3 및 TAD4는 동일한 TAD이다. 특정 실시양태에서, TAD3 및 TAD4는 CITED2 또는 그의 기능적 단편이다. 특정 실시양태에서, TAD3 및 TAD4는 CITED4 또는 그의 기능적 단편이다. 특정 실시양태에서, TAD1, TAD2, TAD3 및 TAD4는 동일한 TAD이다. 특정 실시양태에서, TAD1, TAD2, TAD3 및 TAD4는 CITED2 또는 그의 기능적 단편이다. 특정 실시양태에서, TAD1, TAD2, TAD3 및 TAD4는 CITED4 또는 그의 기능적 단편이다.

[0007] 특정 실시양태에서, 적어도 2개의 TAD 도메인 사이에 링커가 없다.

[0008] 특정 실시양태에서, 적어도 2개의 TAD 도메인 사이에 링커가 있다. 특정 실시양태에서, 링커는 GGSGGGSG (서열식별번호: 177) 또는 GGSGGGSGGGSGGGSG (서열식별번호: 178)를 포함하거나 그로 이루어진다.

[0009] 특정 실시양태에서, DBD는 18-27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다.

[0010] 특정 실시양태에서, DBD는 그의 가장 가까운 인간 대응물에 대해 적어도 80% 서열 동일성을 포함한다. 특정 실시양태에서, DBD는 그의 가장 가까운 인간 대응물에 대해 적어도 90% 서열 동일성을 포함한다. 특정 실시양태에서, DBD 및 적어도 2개의 TAD는 각각 그의 가장 가까운 인간 대응물에 대해 적어도 80% 서열 동일성을 포함한다. 특정 실시양태에서, DBD 및 적어도 2개의 TAD는 각각 그의 가장 가까운 인간 대응물에 대해 적어도 90% 서열 동일성을 포함한다.

[0011] 특정 실시양태에서, DBD는 가이드 RNA 및 뉴클레아제 불활성화된 Cas 단백질을 포함한다. 특정 실시양태에서, 뉴클레아제 불활성화된 Cas 단백질은 뉴클레아제 불활성화된 Cas9이다.

[0012] 특정 실시양태에서, DBD는 아연 핑거 도메인을 포함한다. 특정 실시양태에서, DBD는 6 내지 9개의 아연 핑거 도메인을 포함한다. 특정 실시양태에서, DBD는 6개의 아연 핑거를 포함한다. 특정 실시양태에서, DBD는 18개의 뉴클레오티드를 갖는 게놈 영역에 결합한다. 특정 실시양태에서, DBD는 9개의 아연 핑거를 포함한다. 특정 실시양태에서, DBD는 27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다.

[0013] 특정 실시양태에서, DBD는 서열식별번호: 148-151 중 어느 것에 대해 적어도 95% 서열 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DBD는 서열식별번호: 148-151 중 어느 하나를 갖는 서열을 포함한다.

[0014] 특정 실시양태에서, DBD는 인간 EGR1 또는 인간 EGR3으로부터 유래된다.

[0015] 특정 실시양태에서, DBD는 서열식별번호: 77-98 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DBD는 서열식별번호: 77-98을 포함한다.

[0016] 특정 실시양태에서, DBD는 서열식별번호: 92에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DBD는 서열식별번호: 92를 포함한다.

[0017] 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 130 또는 131에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 130 또는 131을 포함한다.

[0018] 특정 실시양태에서, 발현 카세트는 서열식별번호: 72 또는 73 중 어느 하나에 대해 적어도 90% 동일성을 갖는 뉴클레오티드 서열을 포함한다. 특정 실시양태에서, 발현 카세트는 서열식별번호: 72 또는 73 중 어느 하나의 뉴클레오티드 서열을 포함한다.

[0019] 특정 실시양태에서, 발현 카세트는 다른 세포 유형에서보다 PV 뉴런에서 더 높은 수준으로 전사 인자의 발현을 구동하는 조절 요소를 추가로 포함한다. 특정 실시양태에서, 조절 요소는 서열식별번호: 1-4 중 어느 하나를 포함한다. 특정 실시양태에서, 조절 요소는 서열식별번호: 2 또는 3을 포함한다.

[0020] 특정 실시양태에서, 발현 카세트는 PV 선택적 마이크로RNA 결합 부위를 추가로 포함한다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15 중 어느 하나에 대해 적어도 90% 동일성을 포함한다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15 중 어느 하나를 포함한다.

[0021] 특정 실시양태에서, 발현 카세트는 바이러스 벡터의 일부이다. 특정 실시양태에서, 바이러스 벡터는 AAV 바이러스이다. 특정 실시양태에서, AAV 바이러스는 AAV9 바이러스 또는 scAAV9 바이러스이다. 특정 실시양태에서, 바이러스 벡터는 렌티바이러스이다.

[0022] 또 다른 측면에서, 본 출원은 세포에서 SCN1A 유전자의 발현을 증가시키는 비-자연 발생 전사 인자를 코딩하는 서열을 포함하는 발현 카세트이며, 여기서 비-자연 발생 전사 인자는 전사 활성화 도메인에 작동가능하게 연결된 DNA 결합 도메인을 포함하고, 여기서 DNA 결합 도메인은 서열 LEPGEKP - [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH - TGKKTS (서열식별번호: 147)를 포함하는 아연 핑거 단백질이고, DNA 결합 도메인과 전사 활성화 도메인 사이에 HA 태그 (서열식별번호: 303)가 없는 것인 발현 카세트를 제공한다. 특정 실시양태에서, 전사 활성화 도메인은 VP16, VPR 또는 VP64 서열 또는 그의 기능적 단편을 포함한다. 특정 실시양태에서, 전사 활성화 도메인은 VP64를 포함한다.

[0023] 특정 실시양태에서, DNA 결합 도메인은 18-27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다. 특정 실시양태에서, DNA 결합 도메인은 n = 6 내지 9인 서열식별번호: 147을 포함하는 아연 핑거 도메인이다. 특정 실시양태에서, DNA 결합 도메인은 n = 6인 서열식별번호: 147을 포함하는 아연 핑거 도메인이다. 특정 실시양태에서, DNA 결합 도메인은 18개의 뉴클레오티드를 갖는 게놈 영역에 결합한다. 특정 실시양태에서, DNA 결합 도메인은 n = 9인 서열식별번호: 147을 포함하는 아연 핑거 도메인이다. 특정 실시양태에서, DNA 결합 도메인은 27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다.

[0024] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 148-151 중 어느 것에 대해 적어도 95% 서열 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 148-151 중 어느 하나를 갖는 서열을 포함한다.

[0025] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 77-91 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 77-91 중 어느 하나를 포함한다.

[0026] 특정 실시양태에서, 발현 카세트는 다른 세포 유형에서보다 PV 뉴런에서 더 높은 수준으로 전사 인자의 발현을 구동하는 조절 요소를 추가로 포함한다. 특정 실시양태에서, 조절 요소는 서열식별번호: 1-4 중 어느 하나를 포함한다. 특정 실시양태에서, 조절 요소는 서열식별번호: 2 또는 3을 포함한다.

[0027] 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 127에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 127을 포함한다.

[0028] 특정 실시양태에서, 발현 카세트는 서열식별번호: 93 또는 71 중 어느 하나에 대해 적어도 90% 동일성을 갖는 뉴클레오티드 서열을 포함한다. 특정 실시양태에서, 발현 카세트는 서열식별번호: 93 또는 71 중 어느 하나의 뉴클레오티드 서열을 포함한다.

[0029] 특정 실시양태에서, 발현 카세트는 PV 선택적 마이크로RNA 결합 부위를 추가로 포함한다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15 중 어느 하나에 대해 적어도 90% 동일성을 포함한다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15 중 어느 하나를 포함한다.

[0030] 특정 실시양태에서, 발현 카세트는 바이러스 벡터의 일부이다. 특정 실시양태에서, 바이러스 벡터는 AAV 바이러스이다. 특정 실시양태에서, AAV 바이러스는 AAV9 바이러스 또는 scAAV9 바이러스이다. 특정 실시양태에서, 바이러스 벡터는 렌티바이러스이다.

[0031] 또 다른 측면에서, 본 출원은 서열식별번호: 14 또는 15에 대해 적어도 80% 서열 동일성을 갖는 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드이며, 여기서 마이크로RNA 결합 부위는 흥분성 뉴런에서 트랜스진의 발현을 감소시키는 것인 폴리뉴클레오티드를 제공한다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 서열식별번호: 14를 포함한다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 서열식별번호: 15를 포함한다. 또 다른 측면에서, 본 출원은 PV 선택적 마이크로RNA 결합 부위 및 프로모터 및/또는 인핸서를 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 프로모터 및/또는 인핸서는 다른 세포 유형에서보다 파르브알부민 (PV) 뉴런에서 더 높은 수준으로 트랜스진의 발현을 구동하는 PV 선택적 조절 요소이다. 특정 실시양태에서, PV 선택적 조절 요소는 트랜스진에 작동가능하게 연결된다.

[0032] 또 다른 측면에서, 본 출원은 트랜스진에 작동가능하게 연결된 조절 요소 및 적어도 1개의 마이크로RNA 결합 부위를 포함하는 발현 카세트이며, 여기서 조절 요소는 다른 세포 유형에서보다 파르브알부민 (PV) 뉴런에서 더

높은 수준으로 트랜스진의 발현을 구동하고, 여기서 마이크로RNA 결합 부위는 흥분성 뉴런에서 트랜스진의 발현을 감소시키는 것인 발현 카세트를 제공한다. 특정 실시양태에서, 발현 카세트는 서열식별번호: 67을 포함하지 않는다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR128 (서열식별번호: 9)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR221 (서열식별번호: 11)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR222 (서열식별번호: 13)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR128 (서열식별번호: 9)에 대한 적어도 1개의 결합 부위 및 MIR221 (서열식별번호: 11)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR128 (서열식별번호: 9)에 대한 적어도 1개의 결합 부위, MIR221 (서열식별번호: 11)에 대한 적어도 1개의 결합 부위 및 MIR222 (서열식별번호: 13)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15를 포함한다.

[0033] 특정 실시양태에서, 트랜스진은 세포에서 SCN1A 유전자의 발현을 증가시키는 비-자연 발생 전사 인자를 포함하는 폴리펩티드를 코딩한다. 특정 실시양태에서, 전사 인자는 18-27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다. 특정 실시양태에서, 전사 인자는 DNA 결합 도메인을 포함한다. 특정 실시양태에서, 전사 인자는 DNA 결합 도메인 및 전사 활성화 도메인을 포함한다.

[0034] 특정 실시양태에서, DNA 결합 도메인은 그의 가장 가까운 인간 대응물에 대해 적어도 80% 서열 동일성을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 그의 가장 가까운 인간 대응물에 대해 적어도 90% 서열 동일성을 포함한다. 특정 실시양태에서, DNA 결합 도메인 및 전사 활성화 도메인은 둘 다 그의 가장 가까운 인간 대응물에 대해 적어도 80% 서열 동일성을 포함한다. 특정 실시양태에서, DNA 결합 도메인 및 전사 활성화 도메인은 둘 다 그의 가장 가까운 인간 대응물에 대해 적어도 90% 서열 동일성을 포함한다.

[0035] 특정 실시양태에서, DNA 결합 도메인은 가이드 RNA 및 뉴클레아제 불활성화된 Cas 단백질을 포함한다. 특정 실시양태에서, 뉴클레아제 불활성화된 Cas 단백질은 뉴클레아제 불활성화된 Cas9이다.

[0036] 특정 실시양태에서, DNA 결합 도메인은 아연 핑거 도메인을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 6 내지 9개의 아연 핑거 도메인을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 6개의 아연 핑거를 포함한다. 특정 실시양태에서, DNA 결합 도메인은 18개의 뉴클레오티드를 갖는 게놈 영역에 결합한다. 특정 실시양태에서, DNA 결합 도메인은 9개의 아연 핑거를 포함한다. 특정 실시양태에서, DNA 결합 도메인은 27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다.

[0037] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 148-151 중 어느 것에 대해 적어도 95% 서열 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 148-151 중 어느 하나를 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 92-98 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 92-98 중 어느 하나를 포함한다.

[0038] 특정 실시양태에서, DNA 결합 도메인은 서열 LEPGEKP - [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH - TGKKTS (서열식별번호: 147)를 포함하는 아연 핑거 단백질이다.

[0039] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 77-91 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 77-91 중 어느 하나를 포함한다.

[0040] 특정 실시양태에서, DNA 결합 도메인은 인간 EGR1 또는 인간 EGR3으로부터 유래된다.

[0041] 특정 실시양태에서, 전사 활성화 도메인은 VP16, VPR, VP64, CITED2, CITED4 또는 CREB3 서열 또는 그의 기능적 단편을 포함한다. 특정 실시양태에서, 전사 활성화 도메인은 인간 CITED2, CITED4 또는 CREB3 서열 또는 그의 기능적 단편을 포함한다.

[0042] 특정 실시양태에서, 조절 요소는 서열식별번호: 1-4 중 어느 하나를 갖는 서열을 포함한다. 특정 실시양태에서, 조절 요소는 서열식별번호: 2 또는 3을 갖는 서열을 포함한다.

[0043] 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 105, 106 및 127-129 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 105, 106 및 127-129 중 어느 하나를 포함한다.

[0044] 특정 실시양태에서, 트랜스진은 서열식별번호: 71, 74, 75, 76 또는 184 중 어느 하나에 대해 적어도 90% 동일성을 갖는 뉴클레오티드 서열을 포함한다. 특정 실시양태에서, 트랜스진은 서열식별번호: 71, 74, 75, 76 또는 184 중 어느 하나를 포함한다.

[0045] 특정 실시양태에서, 발현 카세트는 바이러스 벡터의 일부이다. 특정 실시양태에서, 바이러스 벡터는 AAV 바이러스이다. 특정 실시양태에서, AAV 바이러스는 AAV9 바이러스 또는 scAAV9 바이러스이다. 특정 실시양태에서, 바이러스 벡터는 렌티바이러스이다.

[0046] 또 다른 측면에서, 본 출원은 영장류에게 트랜스진 및 적어도 1개의 마이크로RNA 결합 부위를 포함하는 바이러스 벡터를 투여하는 것을 포함하며, 여기서 마이크로RNA 결합 부위는 흥분성 뉴런에서 트랜스진의 발현을 감소시키는 것인, 영장류의 파르브알부민 (PV) 뉴런에서의 트랜스진의 선택적 발현을 위한 방법을 제공한다.

[0047] 특정 실시양태에서, 바이러스 벡터는 트랜스진에 작동가능하게 연결된 조절 요소를 추가로 포함하며, 여기서 조절 요소는 다른 세포 유형에서보다 파르브알부민 (PV) 뉴런에서 더 높은 수준으로 트랜스진의 발현을 구동한다.

[0048] 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR128 (서열식별번호: 9)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR221 (서열식별번호: 11)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR222 (서열식별번호: 13)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR128 (서열식별번호: 9)에 대한 적어도 1개의 결합 부위 및 MIR221 (서열식별번호: 11)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 MIR128 (서열식별번호: 9)에 대한 적어도 1개의 결합 부위, MIR221 (서열식별번호: 11)에 대한 적어도 1개의 결합 부위 및 MIR222 (서열식별번호: 13)에 대한 적어도 1개의 결합 부위를 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, 마이크로RNA 결합 부위는 서열식별번호: 7, 14 또는 15를 포함한다.

[0049] 특정 실시양태에서, 트랜스진은 세포에서 SCN1A 유전자의 발현을 증가시키는 비-자연 발생 전사 인자를 코딩하는 서열을 포함한다.

[0050] 특정 실시양태에서, 전사 인자는 18-27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다.

[0051] 특정 실시양태에서, 전사 인자는 DNA 결합 도메인을 포함한다.

[0052] 특정 실시양태에서, 전사 인자는 DNA 결합 도메인 및 전사 활성화 도메인을 포함한다.

[0053] 특정 실시양태에서, DNA 결합 도메인은 그의 가장 가까운 인간 대응물에 대해 적어도 80% 서열 동일성을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 그의 가장 가까운 인간 대응물에 대해 적어도 90% 서열 동일성을 포함한다. 특정 실시양태에서, DNA 결합 도메인 및 전사 활성화 도메인은 둘 다 그의 가장 가까운 인간 대응물에 대해 적어도 80% 서열 동일성을 포함한다. 특정 실시양태에서, DNA 결합 도메인 및 전사 활성화 도메인은 둘 다 그의 가장 가까운 인간 대응물에 대해 적어도 90% 서열 동일성을 포함한다.

[0054] 특정 실시양태에서, DNA 결합 도메인은 가이드 RNA 및 뉴클레아제 불활성화된 Cas 단백질을 포함한다. 특정 실시양태에서, 뉴클레아제 불활성화된 Cas 단백질은 뉴클레아제 불활성화된 Cas9이다.

[0055] 특정 실시양태에서, DNA 결합 도메인은 아연 핑거 도메인을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 6 내지 9개의 아연 핑거 도메인을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 6개의 아연 핑거를 포함한다. 특정 실시양태에서, DNA 결합 도메인은 18개의 뉴클레오티드를 갖는 게놈 영역에 결합한다. 특정 실시양태에서, DNA 결합 도메인은 9개의 아연 핑거를 포함한다. 특정 실시양태에서, DNA 결합 도메인은 27개의 뉴클레오티드를 갖는 게놈 영역에 결합한다.

[0056] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 148-151 중 어느 것에 대해 적어도 95% 서열 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 148-151 중 어느 하나를 갖는 서열을 포함한다.

[0057] 특정 실시양태에서, DNA 결합 도메인은 인간 EGR1 또는 인간 EGR3으로부터 유래된다.

[0058] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 92-98 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 92-98 중 어느 하나를 포함한다.

[0059] 특정 실시양태에서, DNA 결합 도메인은 서열 LEPGEKP - [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH - TGKKTS (서열식별번호: 147)를 포함하는 아연 핑거 단백질이다.

[0060] 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 77-91 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, DNA 결합 도메인은 서열식별번호: 77-91 중 어느 하나를 포함한다.

[0061] 특정 실시양태에서, 전사 활성화 도메인은 VP16, VPR, VP64, CITED2, CITED4 또는 CREB3 서열 또는 그의 기능적 단편을 포함한다. 특정 실시양태에서, 전사 활성화 도메인은 인간 CITED2, CITED4 또는 CREB3 서열 또는 그의 기능적 단편을 포함한다.

[0062] 특정 실시양태에서, 조절 요소는 서열식별번호: 1-4 중 어느 하나를 갖는 서열을 포함한다. 특정 실시양태에서, 조절 요소는 서열식별번호: 2 또는 3을 갖는 서열을 포함한다.

[0063] 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 105, 106 및 127-129 중 어느 하나에 대해 적어도 90% 동일성을 갖는 서열을 포함한다. 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 105, 106 및 127-129 중 어느 하나를 포함한다.

[0064] 특정 실시양태에서, 트랜스진은 서열식별번호: 71, 74, 75, 76 또는 184 중 어느 하나에 대해 적어도 90% 동일성을 갖는 뉴클레오티드 서열을 포함한다. 특정 실시양태에서, 트랜스진은 서열식별번호: 71, 74, 75, 76 또는 184 중 어느 하나를 포함한다.

[0065] 특정 실시양태에서, 바이러스 벡터는 AAV 바이러스이다. 특정 실시양태에서, AAV 바이러스는 AAV9 바이러스 또는 scAAV9 바이러스이다. 특정 실시양태에서, 바이러스 벡터는 렌티바이러스이다.

[0066] 특정 실시양태에서, 영장류는 인간이다. 특정 실시양태에서, 영장류는 비-인간 영장류이다. 특정 실시양태에서, 비-인간 영장류는 구세계 원숭이, 오랑우탄, 고릴라, 침팬지, 마모셋, 게잡이 마카크, 레서스 마카크 또는 돼지-꼬리 마카크이다.

[0067] 또 다른 측면에서, 본 출원은 세포에서 SCN1A 유전자의 발현을 증가시키는 비-자연 발생 전사 인자를 코딩하는 서열을 포함하는 발현 카세트이며, 여기서 비-자연 발생 전사 인자는 서열식별번호: 128 또는 129에 대해 적어도 90% 동일성을 갖는 서열을 포함하는 것인 발현 카세트를 제공한다. 특정 실시양태에서, 비-자연 발생 전사 인자는 서열식별번호: 128 또는 129를 포함한다.

[0068] 또 다른 측면에서, 본 출원은 본원에 제공된 발현 카세트 중 어느 것을 투여함으로써 세포에서 SCN1A의 발현을 증가시키는 방법을 제공한다. 특정 실시양태에서, 세포는 뉴런 세포이다. 특정 실시양태에서, 뉴런 세포는 단극, 양극, 다극 또는 가단극 뉴런으로 이루어진 군으로부터 선택된다. 특정 실시양태에서, 세포는 GABA성(GABAergic) 뉴런이다. 특정 실시양태에서, 세포는 PV 뉴런이다. 특정 실시양태에서, 세포는 비-뉴런 세포이다. 특정 실시양태에서, 세포는 신경교 세포이다. 특정 실시양태에서, 신경교 세포는 성상세포, 핍지교세포, 상의 세포, 슈반 세포 및 위성 세포로 이루어진 군으로부터 선택된다. 특정 실시양태에서, 세포는 대상체 내의 세포이다. 특정 실시양태에서, 대상체는 포유동물이다. 특정 실시양태에서, 대상체는 인간이다. 특정 실시양태에서, SCN1A의 발현을 증가시키는 것은 질환, 장애 또는 증상을 치료한다. 특정 실시양태에서, 장애는 중추 신경계 장애이다. 특정 실시양태에서, 장애는 SCN1A 반수체기능부전과 연관된 간질이다. 특정 실시양태에서, 반수체기능부전은 대상체가 SCN1A 유전자의 기능 상실 돌연변이에 대해 이형접합인 결과이다. 특정 실시양태에서, 장애는 SCN1A 유전자 내의 삽입, 결실 또는 치환과 연관된 간질이다. 특정 실시양태에서, 장애는 SCN1A 유전자 내의 점 돌연변이와 연관된 간질이다. 특정 실시양태에서, 장애는 드라베 증후군이다. 특정 실시양태에서, 중추 신경계 장애의 증상은 뉴런 과다활성이다. 특정 실시양태에서, 중추 신경계 장애를 치료하는 것은 뉴런 과다활성을 감소시키는 것을 포함한다. 특정 실시양태에서, 중추 신경계 장애의 증상은 발작이다. 특정 실시양태에서, 중추 신경계 장애를 치료하는 것은 발작의 빈도를 감소시키는 것을 포함한다. 특정 실시양태에서, 중추 신경계 장애를 치료하는 것은 발작의 중증도를 감소시키는 것을 포함한다.

[0069] 또 다른 측면에서, 본 출원은 본원에 제공된 발현 카세트 중 어느 하나를 투여함으로써 CNS에서 SCN1A의 발현을 증가시키는 방법을 제공한다. 특정 실시양태에서, 발현 카세트는 일측 뇌실내 (ICV) 투여를 통해 투여된다. 특정 실시양태에서, 발현 카세트는 양측 뇌실내 (ICV) 투여를 통해 투여된다. 특정 실시양태에서, SCN1A의 증가된 발현은 뇌에서 발생한다. 특정 실시양태에서, SCN1A의 증가된 발현은 전두 피질, 두정 피질, 측두 피질, 해마, 수질 및/또는 후두 피질에서 발생한다. 특정 실시양태에서, SCN1A의 증가된 발현은 척추에서 발생한다. 특정 실시양태에서, SCN1A의 증가된 발현은 척수 및/또는 후근 신경절에서 발생한다.

[0070] 참조로 포함됨

[0071] 본 명세서에서 언급된 모든 간행물, 특허 및 특허 출원은 각각의 개별 간행물, 특허 또는 특허 출원이 참조로 포함되는 것으로 구체적으로 및 개별적으로 나타내어진 것과 동일한 정도로 본원에 참조로 포함된다.

도면의 간단한 설명

[0072] 본 발명의 신규 특색은 첨부된 청구범위에 구체적으로 제시된다. 본 발명의 특색 및 이점의 보다 나은 이해는 본 발명의 원리가 이용되는 예시적인 경우들을 제시하는 하기 상세한 설명 및 첨부된 도면을 참조하여 얻어질 것이다:

도 1은 염색체 2 상의 다양한 영역 (GRCh38.p12를 참조함)에 결합하는 조작된 전사 인자를 사용한 내인성 SCN1A의 상향조절을 예시한다. 데이터는 대조군 (EGFP-KASH) 조건 대비 SCN1A 발현의 배수 변화로서 제시된다.

도 2a, 도 2b 및 도 2c는 SCN1A-특이적 전사 활성인자를 사용한 HEK293 세포에서의 내인성 SCN1A의 상대 발현을 예시한다 (표 1 참조). 데이터는 대조군 조건 대비 배수 변화로서 제시되고, Log_{10} 스케일로 제시된다.

도 3a는 SCN1A-특이적 전사 활성인자 (구축물 30)를 사용한 GABA 뉴런에서의 내인성 SCN1A의 상대 발현을 예시한다. 데이터는 대조군 조건 (CBA-EGFP) 대비 배수 변화로서 제시된다.

도 3b는 SCN1A-특이적 전사 활성인자 (구축물 25 및 16)를 사용한 GABA 뉴런에서의 내인성 SCN1A의 상대 발현을 예시한다. 데이터는 대조군 조건 (CBA-EGFP) 대비 배수 변화로서 Log_{10}으로 제시된다.

도 4는 내인성 SCN1A 및 SCN1A 특이적 전사 인자 (구축물 30)에 의해 구동되는 40개의 가장 가까운 이웃 유전자의 상대 발현을 예시한다. 데이터는 대조군 조건 (CBA-EGFP-KASH) 대비 배수 변화로서 Log_{10}으로 제시된다.

도 5a 및 도 5b는 eGFP를 발현한 대조군 발현 카세트와 비교하여 생체내 SCN1A-특이적 전사 활성인자의 발현을 예시한다. 도 5a는 대조군 eGFP 또는 SCN1A 전사 활성인자를 포함하는 구축물 4가 주사된 마우스에서의 SCN1A 유전자의 상대 발현을 예시한다. 도 5b는 백분율 평균 eGFP의 면에서 SCN1A 발현의 변화를 예시한다. 이들 실험은 구축물 4에 의한 전사 활성화가 SCN1A 발현의 약 20-30% 상향조절을 유발하였다는 것을 나타낸다.

도 6a, 도 6b, 도 6c, 도 6d, 도 6e, 도 6f 및 도 6g는 대조군과 비교하여 다양한 SCN1A 특이적 전사 인자를 사용하였을 때 드라베 증후군의 $Scn1a^{tm1Kea}$ 녹아웃 마우스 모델에서의 고체온 발작에 대한 효과를 예시한다. P1 Scn1a +/- 마우스 (이형접합; HET)에게 AAV9-EGFP 또는 SCN1A 특이적 전사 인자 (구축물 31-34, 42 및 43 중 하나)를 발현하는 AAV9 벡터를 주입하였다. P26-P28에서 주입된 마우스에 대해 고체온 유발 발작 검정을 실행하고, 이들이 긴장성-간대성 발작을 경험한 내부 온도를 기록하였다. 도 6d는 DBD와 TAD 사이에 위치한 HA 태그를 함유하는 구축물 32와 HA 태그를 함유하지 않는 구축물 34 사이의 직접적 비교를 보여준다. 도 6e는 코딩 영역과 폴리A 테일 사이에 위치한 m1 마이크로RNA 결합 부위를 함유하는 구축물 31과 m1 마이크로RNA 결합 부위를 함유하지 않는 구축물 32 사이의 직접적 비교를 보여준다. 도 6h는 구축물 31을 사용하였을 때 드라베 증후군의 $Scn1a^{RX}$ 돌연변이 마우스 모델에서의 고체온 발작에 대한 효과를 예시한다 (PBS 주사된 대조군과 비교).

도 7a, 도 7b, 도 7c 및 도 7d는 다양한 조건 하의 드라베 증후군의 $Scn1a^{tm1Kea}$ 녹아웃 마우스 모델에서의 생존을 예시한다. 도 7a는 생존 검정에서 야생형 (PBS WT)과 Scn1a +/- 마우스 (PBS HET) 사이의 비교를 예시한다. P1 Scn1a +/- (N=53) 및 Scn1a +/+ (N=54) 마우스에게 PBS를 주입하였다. 마우스의 홈 케이지에서 매일 마우스를 관찰하였고, 임의의 사망의 경우에는 그 날짜를 기록하였다. Scn1a +/- 동물과 Scn1a +/+ 동물 사이에 생존에 있어서 유의한 차이가 있었다 ($P<0.0001$). 도 7b-d는 대조군과 비교하여 다양한 SCN1A 특이적 전사 인자로 처리된 마우스에 대한 드라베 증후군의 마우스 모델에서의 생존에 대한 효과를 예시한다. P1 Scn1a +/- 마우스에게 PBS 또는 SCN1A 특이적 전사 인자 (구축물 31 또는 33)를 발현하는 AAV9 벡터를 주입하였다. 마우스의 홈 케이지에서 매일 마우스를 관찰하였고, 임의의 사망의 경우에는 그 날짜를 기록하였다. 도 7d는 코딩 영역과 폴리A 테일 사이에 위치한 m1 마이크로RNA 결합 부위를 함유하는 구축물 31과 m1 마이크로RNA 결합 부위를 함유하지 않는 구축물 33 사이의 직접적 비교를 보여준다. 도 7e는 구축물 31을 사용한 드라베 증후군의 $Scn1a^{RX}$ 돌연변이 마우스 모델에서의 생존을 예시한다 (PBS 주사된 대조군과 비교).

도 8은 2마리의 시노몰구스 마카크에게 $1.2x10^{12}$ gc/동물로 투여된 SCN1A 특이적 전사 인자 (구축물 33)를 코딩하는 AAV9 벡터의 실질내 전달 후 상이한 뇌 조직에서의 상대 Scn1A mRNA 발현을 예시하며, 이는 2마리의 비처

리 대조군 동물에 대해 정규화된 것이다. 주사 28일 후에 모든 동물을 희생시키고, 택맨 PCR에 의해 조직 샘플에서 Scn1A mRNA를 정량화하였다. 데이터는 뇌로부터의 상이한 조직 절편에서의 표적 mRNA의 정규화된 발현으로서 보고된다. 상이한 세트의 Scn1a 유전자 유래 프라이머/프로브로도 유사한 결과가 기록되었다.

도 9a-f는 EF1a 프로모터, RE 2 프로모터 (서열식별번호: 2), 또는 EGFP 코딩 영역과 폴리A 부위 사이에 위치한 m1 마이크로RNA 결합 부위 (서열식별번호: 7)와 함께 RE 2 프로모터 (서열식별번호: 2)의 제어 하에 EGFP 트랜스진을 포함하는 AAV9 벡터로 처리한 후 마모셋 해마 치상회 영역에서의 EGFP의 발현 패턴을 보여준다. 해마의 치상회 영역의 대표적인 영역이 각각의 벡터 처리에 대해 제시된다. 상단 열은 DAPI로 염색된 세포 핵을 보여주고, 하단 열은 항-GFP 항체로 염색된 GFP 양성 영역을 보여준다. 도 9a (EF1a 처리)에서, 해마 CA4 문 영역은 황색으로 윤곽이 그려져 있으며, 화살표는 치상 세포 과립 세포체 층 (DG)을 가리킨다. 도 9b 및 도 9c는 동일한 영역을 중심으로 한 것이다. 흥분성 개재뉴런과 억제성 개재뉴런의 혼합물인 CA4 영역이 RE 2 + m1 조건에서 유일한 유의한 발현 영역이었기 때문에 강조된다. EF1a 및 RE 2가 트랜스진 발현을 구동하였을 때, GFP 발현은 더 넓어졌고, 해마의 다른 영역을 포함하였다. DG 세포 층은 주로 흥분성 뉴런을 함유하는 것으로 생각된다. EF1a 및 RE 2에 의해 구동된 GFP 발현은 DG 세포 층에서 가시적이지만 (도 9d 및 도 9e), RE 2 + m1 처리된 동물에서는 존재하지 않는다 (도 9f) (백색 화살촉).

도 10a-l은 EF1a 프로모터, RE 2 프로모터 (서열식별번호: 2), 또는 EGFP 코딩 영역과 폴리A 부위 사이에 위치한 m1 마이크로RNA 결합 부위 (서열식별번호: 7)와 함께 RE 2 프로모터 (서열식별번호: 2)의 제어 하에 EGFP 트랜스진을 포함하는 AAV9 벡터로 처리한 후 마모셋 해마 치상회 영역에서의 EGFP 발현 패턴이 RE 2 및 RE 2 + m1로 처리된 동물에서 파르브알부민 (PV) 양성 세포에 주로 국재화된다는 것을 보여준다. 해마의 치상회 영역의 대표적인 영역이 각각의 벡터 처리에 대해 제시된다. 상단 열은 GFP 양성 영역을 보여주고, 다음 하단 열은 PV에 대한 억제성 개재뉴런 마커로 염색된 동일한 영역을 보여준다. 도 10a-f의 박스 내 영역은 도 10g-l에서 더 높은 배율로 제시된다. RE 2 및 RE 2 + m1에 의해 구동된 GFP 발현은 주로 억제성 개재뉴런 마커 PV와 공동-국재화되는 반면 (도 10h 및 10k, 10i 및 10l 백색 화살촉), EG-EF1a에서는 GFP 발현이 PV 양성 세포에 용이하게 국재화되지 않는다 (도 10g 및 10j 백색 화살촉). 추가로, GFP 양성 세포는 EF1a 처리된 동물에서의 덜 뚜렷한 세포체 형태 (도 10g 황색 화살촉)와 비교하여 RE 2 및 RE 2+ m1 처리된 동물에서 피라미드 세포체를 갖는 고도로 분지된 세포의 뚜렷한 개재뉴런 형태 (도 10h 및 10i 황색 화살촉)를 갖는다.

도 11은 일측 뇌실내 (ICV) 투여를 통해 4.8E+13 또는 8E+13 vg/동물로 투여되는 AAV9-RE^{GABA}-eTF^{SCN1A}로 처리된 동물 (실시예 10 및 실시예 11)에 대한 전두 피질 (FC), 문측 두정 피질 (문측 PC), 측두 피질 (TC), 미측 두정 피질 (미측 PC), 해마 (Hip), 수질 (Med) 및 후두 피질 (OC) 조직 샘플에서의 VG/이배체 게놈을 보여준다. 각각의 데이터 포인트는 조직 샘플에 대한 VG/이배체 게놈을 나타내고, 수평 막대는 각각의 동물에서의 모든 조직 샘플에 대한 평균 VG/이배체 게놈을 나타낸다.

도 12는 일측 뇌실내 (ICV) 투여를 통해 4.8E+13 또는 8E+13 vg/동물로 투여되는 AAV9-RE^{GABA}-eTF^{SCN1A}로 처리된 동물 (실시예 10 및 실시예 11)에 대한 전두 피질 (FC), 문측 두정 피질 (문측 PC), 측두 피질 (TC), 미측 두정 피질 (미측 PC), 해마 (Hip), 수질 (Med) 및 후두 피질 (OC) 조직 샘플에서의 전사체/μg RNA를 보여준다. 각각의 데이터 포인트는 조직 샘플에 대한 VG/이배체 게놈을 나타내고, 수평 막대는 각각의 동물에서의 모든 조직 샘플에 대한 평균 VG/이배체 게놈을 나타낸다. ARFGAP2에 대한 평균 전사체는 1.85E+6/μg RNA였고, 파선 상부 경계선으로 표시된다. 검출 한계는 파선 하부 경계선으로 표시된다.

도 13은 뇌 외부의 말초 조직 샘플에서의 벡터 생체분포 (VG/이배체 게놈) 및 트랜스진 발현 (전사체/μg RNA)을 보여준다. 제시된 말초 조직 샘플은 척수 C2/L4 (SC C2/L4), 후근 신경절 C2/L4 (DRG C2/L4), 간, 비장, 심장, 신장, 폐, 췌장 및 고환/난소이다. 영장류 뇌에서의 평균 VCN (벡터 생체분포) 및 전사체 (트랜스진 발현)는 파선으로 표시된다.

발명을 실시하기 위한 구체적인 내용

[0073] 비-자연 발생이고 게놈 표적 부위에 결합하며 관심 내인성 유전자의 발현을 조정하도록 설계된 조작된 전사 인자 또는 eTF가 본원에 제공된다. 이러한 eTF는 관심 유전자의 발현 (RNA 및/또는 단백질 발현)을 상향조절 또는 하향조절하도록 설계될 수 있다. 또한, 바이러스 벡터 내로 혼입될 수 있고 파르브알부민 (PV) 뉴런에서 트랜스진의 선택적 발현을 제공할 수 있는 마이크로RNA 결합 부위가 본원에 제공된다.

[0074] 한 측면에서, 본 출원은 나트륨 전압 게이팅 채널 알파 서브유닛 1 (SCN1A) 유전자의 발현을 상향조절할 수 있

고 그의 상응하는 단백질 산물 Nav1.1의 발현을 증가시킬 수 있는 eTF; 및 Nav1.1의 결핍과 연관된 질환 또는 장애, 예컨대, 예를 들어 드라베 증후군을 치료하기 위한 그의 사용 방법을 제공한다.

[0075] 또 다른 측면에서, 본 출원은 흥분성 뉴런에서 마이크로RNA 결합 부위를 함유하는 mRNA의 발현을 감소시킴으로써 GABA성 또는 파르브알부민 (PV) 뉴런에서의 유전자의 선택적 발현을 유도하는 마이크로RNA 결합 부위 및 PV 뉴런에서의 관심 유전자의 선택적 발현을 위한 그의 사용 방법을 제공한다.

[0076] 정의

[0077] 본원에 사용된 단수 형태는 문맥이 달리 명확하게 나타내지 않는 한 복수 형태를 또한 포함하는 것으로 의도된다. 추가로, 용어 "포함한", "포함한다", "갖는", "갖는다", "가진" 또는 그의 변형어가 상세한 설명 및/또는 청구범위에 사용되는 경우에, 이러한 용어는 용어 "포함하는"과 유사한 방식으로 포괄적인 것으로 의도된다.

[0078] 용어 "약" 또는 "대략"은 관련 기술분야의 통상의 기술자에 의해 결정시 특정한 값에 대한 허용 오차 범위 내를 의미하며, 이는 부분적으로 값이 측정 또는 결정되는 방법, 즉 측정 시스템의 한계에 좌우될 것이다. 예를 들어, "약"은 관련 기술분야의 실시에 따라 1 또는 1 초과의 표준 편차 내에 있음을 의미할 수 있다. 대안적으로, "약"은 주어진 값의 20% 이하, 15% 이하, 10% 이하, 5% 이하 또는 1% 이하의 범위를 의미할 수 있다.

[0079] 용어 "결정하는", "측정하는", "평가하는", "사정하는", "검정하는", "분석하는" 및 그의 문법적 등가물은 임의의 형태의 측정을 지칭하는 것으로 본원에서 상호교환가능하게 사용될 수 있고, 요소가 존재하는지 아닌지를 결정하는 것 (예를 들어, 검출)을 포함한다. 이들 용어는 정량적 및/또는 정성적 결정 둘 다를 포함할 수 있다. 평가는 상대적이거나 절대적일 수 있다.

[0080] 용어 "발현"은 핵산 서열 또는 폴리뉴클레오티드가 DNA 주형으로부터 (예컨대 mRNA 또는 다른 RNA 전사체로) 전사되는 과정 및/또는 전사된 mRNA가 후속적으로 펩티드, 폴리펩티드 또는 단백질로 번역되는 과정을 지칭한다. 전사체 및 코딩된 폴리펩티드는 집합적으로 "유전자 산물"로 지칭될 수 있다. 폴리뉴클레오티드가 게놈 DNA로부터 유래된 경우에, 발현은 진핵 세포에서의 mRNA의 스플라이싱을 포함할 수 있다.

[0081] 본원에 사용된 "작동가능하게 연결된", "작동가능한 연결", "작동적으로 연결된" 또는 그의 문법적 등가물은 유전자 요소, 예를 들어 프로모터, 인핸서, 폴리아데닐화 서열 등의 병렬배치를 지칭하며, 여기서 요소는 이들이 예상된 방식으로 작동하도록 허용하는 관계에 있다. 예를 들어, 프로모터 및/또는 인핸서 서열을 포함할 수 있는 조절 요소가 코딩 서열의 전사를 개시하는 것을 돕는 경우에 조절 요소가 코딩 영역에 작동적으로 연결된다. 이러한 기능적 관계가 유지되는 한, 조절 요소와 코딩 영역 사이에 개재 잔기가 있을 수 있다.

[0082] 본원에 사용된 "벡터"는 폴리뉴클레오티드를 포함하거나 그와 회합하고 세포로의 폴리뉴클레오티드의 전달을 매개하는데 사용될 수 있는 거대분자 또는 거대분자의 회합을 지칭한다. 벡터의 예는 플라스미드, 바이러스 벡터, 리포솜 및 다른 유전자 전달 비히클을 포함한다. 벡터는 일반적으로 표적에서 유전자의 발현을 용이하게 하기 위해 유전자에 작동적으로 연결된 유전 요소, 예를 들어 조절 요소를 포함한다.

[0083] 본원에 사용된 "발현 카세트" 및 "핵산 카세트"는 함께 발현되거나 발현을 위해 작동가능하게 연결된 핵산 서열 또는 요소의 조합을 지칭하는 것으로 상호교환가능하게 사용된다. 일부 경우에, 발현 카세트는 조절 요소 및 발현을 위해 작동가능하게 연결된 유전자 또는 유전자들의 조합을 지칭한다.

[0084] 용어 "AAV"는 아데노-연관 바이러스의 약어이고, 바이러스 자체 또는 그의 유도체를 지칭하는 것으로 사용될 수 있다. 상기 용어는 달리 요구되는 경우를 제외하고는 모든 혈청형, 하위유형 및 자연 발생 및 재조합 형태 둘 다를 포괄한다. 약어 "rAAV"는 재조합 아데노-연관 바이러스를 지칭하며, 재조합 AAV 벡터 (또는 "rAAV 벡터")로도 지칭된다. 용어 "AAV"는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, rh10 및 그의 하이브리드, 조류 AAV, 소 AAV, 개 AAV, 말 AAV, 영장류 AAV, 비-영장류 AAV 및 양 AAV를 포함한다. AAV의 다양한 혈청형의 게놈 서열, 뿐만 아니라 천연 말단 반복부 (TR), Rep 단백질 및 캡시드 서브유닛의 서열은 관련 기술분야에 공지되어 있다. 이러한 서열은 문헌 또는 공공 데이터베이스, 예컨대 진뱅크 (GenBank)에서 찾아볼 수 있다. 본원에 사용된 "rAAV 벡터"는 AAV 기원이 아닌 폴리뉴클레오티드 서열 (즉, AAV에 대해 이종인 폴리뉴클레오티드), 전형적으로 세포의 유전자 형질전환을 위한 관심 서열을 포함하는 AAV 벡터를 지칭한다. 일반적으로, 이종 폴리뉴클레오티드는 적어도 1개, 일반적으로 2개의 AAV 역전된 말단 반복 서열 (ITR)이 플랭킹되어 있다. rAAV 벡터는 단일-가닥 (ssAAV) 또는 자기-상보적 (scAAV)일 수 있다. "AAV 바이러스" 또는 "AAV 바이러스 입자"는 적어도 1종의 AAV 캡시드 단백질 및 캡시드화 폴리뉴클레오티드 rAAV 벡터로 구성된 바이러스 입자를 지칭한다. 입자가 이종 폴리뉴클레오티드 (즉, 야생형 AAV 게놈 이외의 폴리뉴클

레오티드, 예컨대 포유동물 세포로 전달될 트랜스진)를 포함하는 경우, 이는 전형적으로 "rAAV 벡터 입자" 또는 간단히 "rAAV 입자"로 지칭된다. 따라서, rAAV 입자의 생산은 반드시 rAAV 벡터의 생산을 포함하는데, 이는 이러한 벡터가 rAAV 입자 내에 함유되기 때문이다.

[0085] 본원에 사용된 용어 "치료하다", "치료", "요법" 등은 질환 또는 장애의 진행의 완화, 지연 또는 둔화, 예방, 약화, 효과 또는 증상의 감소, 발병의 방지, 발병의 억제 또는 호전을 지칭한다. 본 개시내용의 방법은 임의의 포유동물에 사용될 수 있다. 예시적인 포유동물은 래트, 고양이, 개, 말, 소, 양, 돼지 및 보다 바람직하게는 인간을 포함하나 이에 제한되지는 않는다. 치료 이익은 치료될 기저 장애의 근절 또는 호전을 포함한다. 또한, 치료 이익은 대상체가 여전히 기저 장애를 앓을 수 있음에도 불구하고 대상체에서 개선이 관찰되도록 기저 장애와 연관된 생리학적 증상 중 1종 이상의 근절 또는 호전에 의해 달성된다. 일부 경우에, 예방적 이익을 위해, 특정한 질환이 발생할 위험이 있는 대상체에게 또는 질환이 진단되지 않았을 수 있더라도 이러한 질환의 생리학적 증상 중 1종 이상을 보고하는 대상체에게 치료제가 투여될 수 있다. 본 개시내용의 방법은 임의의 포유동물에 사용될 수 있다. 일부 경우에, 치료는 증상의 감소 또는 중지 (예를 들어, 발작의 빈도, 지속기간 및/또는 중증도의 감소)을 유발할 수 있다. 예방적 효과는 질환 또는 상태의 출현의 지연 또는 제거, 질환 또는 상태의 증상의 발병의 지연 또는 제거, 질환 또는 상태의 진행의 둔화, 정지 또는 역전 또는 그의 임의의 조합을 포함한다.

[0086] 용어 "유효량" 또는 "치료 유효량"은 하기 정의된 바와 같은 질환 치료를 포함하나 이에 제한되지는 않는 의도된 적용에 영향을 미치기에 충분한 본원에 기재된 조성물의 양을 지칭한다. 치료 유효량은 의도된 치료 용도 (생체내), 또는 치료될 대상체 및 질환 상태, 예를 들어 대상체의 체중 및 연령, 질환 상태의 중증도, 투여 방식 등에 따라 달라질 수 있으며, 이는 관련 기술분야의 통상의 기술자에 의해 용이하게 결정될 수 있다. 상기 용어는 또한 표적 세포에서 특정한 반응을 유도할 용량에 적용된다. 구체적 용량은 선택된 특정한 조성물, 따를 투여 요법, 다른 화합물과 조합되어 투여되는지 여부, 투여 시기, 투여될 조직 및 수행될 물리적 전달 시스템에 따라 달라질 것이다.

[0087] 뉴클레오티드 또는 펩티드 서열의 "단편"은 참조 또는 "전장" 서열보다 짧은 서열을 지칭한다.

[0088] 분자의 "변이체"는 이러한 서열의 대립유전자 변이, 즉 전체 분자 또는 그의 단편과 구조적 및 생물학적 활성이 실질적으로 유사한 서열을 지칭한다.

[0089] DNA 또는 단백질 서열의 "기능적 단편"은 전장 DNA 또는 단백질 서열의 생물학적 활성과 실질적으로 유사한 생물학적 활성 (기능적 또는 구조적)을 보유하는 단편을 지칭한다. DNA 서열의 생물학적 활성은 전장 서열에 기인하는 것으로 알려진 방식으로 발현에 영향을 미치는 그의 능력일 수 있다.

[0090] 용어 "대상체" 및 "개체"는 척추동물, 바람직하게는 포유동물, 보다 바람직하게는 인간을 지칭하는 것으로 본원에서 상호교환가능하게 사용된다. 본원에 기재된 방법은 질환 또는 상태의 동물 모델에서의 인간 치료, 수의학적 적용 및/또는 전임상 연구에 유용할 수 있다.

[0091] 용어 "생체내"는 대상체의 신체에서 일어나는 사건을 지칭한다.

[0092] 용어 "시험관내"는 대상체의 신체 외부에서 일어나는 사건을 지칭한다. 예를 들어, 시험관내 검정은 대상체 외부에서 실행되는 임의의 검정을 포괄한다. 시험관내 검정은 살아있거나 죽은 세포가 사용되는 세포-기반 검정을 포괄한다. 시험관내 검정은 또한 어떤 무손상 세포도 사용되지 않는 무세포 검정을 포괄한다.

[0093] 일반적으로, 상호교환가능하게 사용될 수 있는 "서열 동일성" 또는 "서열 상동성"은 각각 2개의 폴리뉴클레오티드 또는 폴리펩티드 서열의 정확한 뉴클레오티드-대-뉴클레오티드 또는 아미노산-대-아미노산 상응성을 지칭한다. 전형적으로, 서열 동일성을 결정하기 위한 기술은 2개의 뉴클레오티드 또는 아미노산 서열을 비교하는 것 및 그의 퍼센트 동일성을 결정하는 것을 포함한다. 서열 비교는, 예컨대 동일성을 평가하기 위한 목적으로, 니들만-분쉬(Needleman-Wunsch) 알고리즘 (예를 들어, www.ebi.ac.uk/Tools/psa/emboss_needle/에서 이용가능한 EMBOSS 니들 정렬기 (임의로 디폴트 설정을 가짐) 참조), BLAST 알고리즘 (예를 들어, blast.ncbi.nlm.nih.gov/Blast.cgi에서 이용가능한 BLAST 정렬 도구 (임의로 디폴트 설정을 가짐) 참조), 및 스미스-워터만(Smith-Waterman) 알고리즘 (예를 들어, www.ebi.ac.uk/Tools/psa/emboss_water/에서 이용가능한 EMBOSS 워터 정렬기 (임의로 디폴트 설정을 가짐) 참조)을 포함하나 이에 제한되지는 않는 임의의 적합한 정렬 알고리즘에 의해 수행될 수 있다. 최적 정렬은 디폴트 파라미터를 포함한 선택된 알고리즘의 임의의 적합한 파라미터를 사용하여 평가될 수 있다. 2개의 서열 사이의 "퍼센트 상동성"으로도 지칭되는 "퍼센트 동일성"은 2개의 최적으로 정렬된 서열 사이의 정확한 매치의 수를 참조 서열의 길이로 나누고 100을 곱하여 계산될 수 있

다. 퍼센트 동일성은 또한, 예를 들어 국립 보건원으로부터 입수가능한 버전 2.2.9를 포함한 진보된 BLAST 컴퓨터 프로그램을 사용하여 서열 정보를 비교함으로써 결정될 수 있다. BLAST 프로그램은 문헌 [Karlin and Altschul, Proc. Natl. Acad. Sci. USA 87:2264-2268 (1990)]의 정렬 방법에 기초하고, 문헌 [Altschul, et al., J. Mol. Biol. 215:403-410 (1990); Karlin and Altschul, Proc. Natl. Acad. Sci. USA 90:5873-5877 (1993); 및 Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)]에서 논의된 바와 같다. 간략하게, BLAST 프로그램은 동일한 정렬된 기호 (즉, 뉴클레오티드 또는 아미노산)의 수를 2개의 서열 중 더 짧은 것에서의 기호의 총수로 나눈 것으로 동일성을 정의한다. 상기 프로그램은 비교될 서열의 전체 길이에 걸쳐 퍼센트 동일성을 결정하는데 사용될 수 있다. 디폴트 파라미터는 짧은 질의 서열을 사용한, 예를 들어 blastp 프로그램을 사용한 검색을 최적화하기 위해 제공된다. 상기 프로그램은 또한 문헌 [Wootton and Federhen, Computers and Chemistry 17:149-163 (1993)]의 SEG 프로그램에 의해 결정시 질의 서열의 절편을 차폐하기 위한 SEG 필터의 사용을 가능하게 한다. 높은 서열 동일성은 일반적으로 대략 80% 내지 100%의 서열 동일성 범위 및 그 사이의 정수 값을 포함한다.

[0094] 단백질과 관련하여 본원에 사용된 "조작된"은 비-자연 발생 단백질, 예컨대 비제한적으로 자연 발생 단백질로부터 유래된 단백질 또는 자연 발생 단백질이 특정 특성을 갖도록 변형 또는 재프로그램화된 경우를 지칭한다.

[0095] 본원에 사용된 "합성" 및 "인공"은 자연 발생 인간 단백질에 대해 낮은 서열 동일성 (예를 들어, 50% 미만의 서열 동일성)을 갖는 단백질 또는 그의 도메인을 지칭하는 것으로 상호교환가능하게 사용된다. 예를 들어, VPR 및 VP64 도메인은 합성 전사활성화 도메인이다.

[0096] 본원에 사용된 "조작된 전사 인자" 또는 "eTF"는 특이적 표적 결합 부위에 결합하고/거나 변형 또는 대체된 전사 이펙터 도메인을 포함하도록 변형 또는 재프로그램화된 비-자연 발생 DNA 결합 단백질 또는 비-자연 발생 전사 조정제를 지칭한다.

[0097] 본원에 사용된 "DNA 결합 도메인"은 DNA 결합 단백질의 일부로서 1개 이상의 DNA 결합 모티프, 예컨대 아연 핑거 또는 염기성 헬릭스-루프-헬릭스 (bHLH) 모티프를 개별적으로 또는 집합적으로 지칭하는 것으로 사용될 수 있다.

[0098] 용어 "전사 활성화 도메인", "전사의 활성화 도메인", "전사활성화 도메인", "트랜스-활성화 도메인" 및 "TAD"는 본원에서 상호교환가능하게 사용되고, DNA 결합 도메인과 함께 전사 기구 (예를 들어, 일반적 전사 인자 및/또는 RNA 폴리머라제)에 직접 또는 보조-활성화제로 공지된 다른 단백질을 통해 접촉함으로써 프로모터로부터의 전사를 활성화시킬 수 있는 단백질의 도메인을 지칭한다.

[0099] 용어 "전사 억제인자 도메인", "전사의 억제인자 도메인" 및 "TRD"는 본원에서 상호교환가능하게 사용되고, DNA 결합 도메인과 함께 전사 기구 (예를 들어, 일반적 전사 인자 및/또는 RNA 폴리머라제)에 직접 또는 보조-억제인자로 공지된 다른 단백질을 통해 접촉함으로써 프로모터로부터의 전사를 억제할 수 있는 단백질의 도메인을 지칭한다.

[0100] 용어 "GRCh38.p12"는 2017년 12월 21일자로 진뱅크 어셈블리 수탁 번호 GCA_000001405.27을 갖는 게놈 참조 컨소시엄 휴먼 빌드(Genome Reference Consortium Human Build) 38 패치 릴리스 12 (GRCh38.p12)를 지칭한다.

[0101] 달리 나타내지 않는 한, 본원에 사용된 모든 용어는 관련 기술분야의 통상의 기술자에게 동일한 의미를 갖고, 본 발명의 실시는 관련 기술분야의 통상의 기술자의 지식 내에 있는 분자 생물학, 미생물학 및 재조합 DNA 기술의 통상적인 기술을 사용할 것이다.

[0102] SCN1A를 상향조절하는 조작된 전사 인자 (eTF)

[0103] 한 측면에서, 본 출원은 나트륨 전압 게이팅 채널 알파 서브유닛 1 (SCN1A) 유전자의 발현을 상향조절할 수 있고 그의 상응하는 단백질 산물 Nav1.1의 발현을 증가시킬 수 있는 eTF를 제공한다. SCN1A 유전자는 나트륨 채널을 조립하는데 사용되는 서브유닛을 코딩하는 유전자의 패밀리에 속한다. 양으로 하전된 나트륨 이온을 세포 내로 수송하는 이들 채널은 전기 신호를 생성하고 전송하는 세포의 능력에서 중요한 역할을 한다. SCN1A 유전자는 Nav1.1로 불리는 나트륨 채널의 한 부분 (알파 서브유닛)을 코딩한다. 이들 채널은 주로 뇌에서 발견되며, 여기서 이들은 세포 내로의 나트륨 이온의 유동을 제어한다. Nav1.1 채널은 하나의 신경 세포 (또는 뉴런)로부터 또 다른 것으로 신호를 전송하는데 관여한다. SCN1A 유전자 내의 여러 돌연변이는 다양한 중증도의 발작 장애의 스펙트럼인 열성 발작 플러스 동반 유전적 간질 (GEFS+)을 유발하는 것으로 밝혀졌다. 이들 상태는 영아기에서 시작하여 통상적으로 5세에 멈추는 단순 열성 (열-연관) 발작 및 열성 발작 플러스 (FS+)를 포함한다. FS+는 열성 발작 및 소아기를 지나 계속되는 열과 관련되지 않은 발작 (무열성 발작)을 포함한 다른

유형의 발작을 수반한다. GEFS+ 스펙트럼은 또한 보다 오래 지속되고 제어하기 어려울 수 있는 보다 심각한 발작을 유발하는 다른 상태, 예컨대 드라베 증후군 (영아기의 중증 근간대성 간질 또는 SMEI로도 공지됨)을 포함한다. 이들 재발성 발작 (간질)은 시간 경과에 따라 악화될 수 있고, 종종 뇌 기능의 저하를 동반한다. 많은 다른 돌연변이는 가족에서 진행되는 편두통 형태인 가족성 편마비성 편두통과 연관되었고, 적어도 1개의 돌연변이는 특정 항발작 의약의 유효성과 연관되었다. 따라서, SCN1A의 발현을 증가시키는 본원에 제공된 eTF는 Nav1.1 채널에서의 돌연변이와 연관된 다양한 질환 또는 장애를 치료하기 위해 사용될 수 있다.

[0104] 전사 인자 (TF)는 게놈 내의 특이적 서열에 결합하고 유전자의 발현을 제어하는 단백질이다. SCN1A를 상향조절하는 본원에 제공된 조작된 전사 인자 또는 eTF는 DNA 결합 도메인 (DBD), 및 전사 조정인자, 예를 들어 전사 활성화 도메인 (TAD) 또는 전사 억제인자 도메인 (TRD)인 적어도 1개의 도메인을 포함하는 비-자연 발생 단백질이다. 한 실시양태에서, SCN1A를 상향조절하는 eTF는 DBD 및 TAD (예를 들어, TAD-DBD 또는 DBD-TAD)를 포함할 수 있으며, 여기서 DBD 및 TAD는 동일한 단백질로부터 또는 상이한 단백질로부터 유래될 수 있다. 또 다른 실시양태에서, SCN1A를 상향조절하는 eTF는 DBD 및 2개의 TAD를 포함할 수 있으며, 여기서 DBD 및 TAD는 동일한 단백질로부터 유래되거나, DBD는 제1 단백질로부터 유래되고 TAD 둘 다는 제2 단백질로부터 유래되거나, DBD 및 1개의 TAD는 제1 단백질로부터 유래되고 제2 TAD는 제2 단백질로부터 유래되거나, 또는 DBD는 제1 단백질로부터 유래되고 1개의 TAD는 제2 단백질로부터 유래되고 제2 TAD는 제3 단백질로부터 유래된다 (예를 들어, TAD1-DBD-TAD1, TAD1-DBD-TAD2, TAD1-TAD1-DBD, TAD1-TAD2-DBD, DBD-TAD1-TAD1 또는 DBD-TAD1-TAD2). 또 다른 실시양태에서, SCN1A를 상향조절하는 eTF는 DBD 및 3개의 TAD를 포함할 수 있으며, 여기서 DBD 및 TAD는 동일한 단백질로부터 유래되거나, DBD는 제1 단백질로부터 유래되고 TAD는 1종 이상의 상이한 단백질로부터 유래되거나, 또는 DBD 및 모든 TAD는 모두 상이한 단백질로부터 유래되고, 예를 들어 TAD_X-TAD_X-TAD_X-DBD, TAD_X-TAD_X-DBD-TAD_X, TAD_X-DBD-TAD_X-TAD_X 또는 DBD-TAD_X-TAD_X-TAD_X이며, 여기서 각각의 X는 독립적으로 선택되고 다른 TAD 중 하나 또는 모두와 동일하거나 상이할 수 있다. 예는, 예를 들어 TAD1-TAD1-DBD-TAD1, TAD1-TAD1-DBD-TAD2, TAD1-TAD2-DBD-TAD1, TAD1-TAD2-DBD-TAD2, TAD1-TAD2-DBD-TAD3, TAD1-DBD-TAD1-TAD1, TAD1- DBD-TAD2-TAD2, TAD1-DBD-TAD1-TAD2, TAD2-DBD-TAD1-TAD2, TAD1-DBD-TAD2-TAD3, TAD1-TAD1-TAD1-DBD, TAD1-TAD2-TAD2-DBD, TAD1-TAD2-TAD2-DBD, TAD1-TAD2-TAD3-DBD, DBD-TAD1-TAD1-TAD1, DBD-TAD1-TAD1-TAD2, DBD-TAD1-TAD2-TAD2 또는 DBD-TAD1-TAD2-TAD3 등을 포함한다. 또 다른 실시양태에서, SCN1A를 상향조절하는 eTF는 DBD 및 4종의 TAD를 포함할 수 있으며, 여기서 DBD 및 TAD는 동일한 단백질로부터 유래되거나, DBD는 제1 단백질로부터 유래되고 TAD는 1종 이상의 상이한 단백질로부터 유래되거나, 또는 DBD 및 모든 TAD는 모두 상이한 단백질로부터 유래되고, 예를 들어 TAD_X-TAD_X-TAD_X-TAD_X-DBD, TAD_X-TAD_X-TAD_X-DBD-TAD_X, TAD_X-TAD_X-DBD-TAD_X-TAD_X, TAD_X- DBD-TAD_X-TAD_X-TAD_X 또는 DBD-TAD_X-TAD_X-TAD_X-TAD_X이며, 여기서 각각의 X는 독립적으로 선택되고 다른 TAD 중 하나 또는 모두와 동일하거나 상이할 수 있다. 예는, 예를 들어 TAD1-TAD1-DBD-TAD1-TAD1, TAD1-TAD1-DBD-TAD2-TAD2, TAD1-TAD2-DBD-TAD1-TAD2, TAD1-TAD2-DBD-TAD2-TAD1, TAD1-TAD2-DBD-TAD1-TAD3, TAD1-TAD3-DBD-TAD1-TAD2, TAD1-TAD2-DBD-TAD3-TAD4, TAD1-TAD1-TAD1-DBD-TAD2, TAD1-TAD2-TAD3-DBD-TAD4, TAD1-DBD-TAD1-TAD1-TAD2, TAD1-DBD-TAD2-TAD3-TAD4, TAD1-DBD-TAD1-TAD2-TAD3, TAD2-DBD-TAD1-TAD2-TAD3, TAD1-DBD-TAD2-TAD3-TAD4, TAD1-TAD1-TAD1-TAD1-DBD, TAD1-TAD2-TAD2-TAD3-DBD, TAD1-TAD2-TAD3-TAD4-DBD, DBD-TAD1-TAD1-TAD1-TAD1, DBD-TAD1-TAD1-TAD2-TAD2, DBD-TAD1-TAD2-TAD3-TAD4 또는 DBD-TAD1-TAD2-TAD3-TAD3 등을 포함한다. 한 실시양태에서, SCN1A를 상향조절하는 eTF는 DBD 및 DBD의 동일한 말단 (예를 들어, N-말단 또는 C-말단)에 위치한 2개의 TAD를 포함하며, 여기서 DBD는 제1 단백질로부터 유래되고 TAD 둘 다는 제2 단백질로부터 유래되거나, 또는 DBD는 제1 단백질로부터 유래되고 1개의 TAD는 제2 단백질로부터 유래되고 제2 TAD는 제3 단백질로부터 유래된다 (예를 들어, TAD1-TAD1-DBD, TAD1-TAD2-DBD, DBD-TAD1-TAD1 또는 DBD-TAD1-TAD2). 특정 실시양태에서, DBD는 다중 단백질로부터의 도메인을 함유하는 합성 구축물일 수 있다.

[0105] 특정 실시양태에서, DBD 및 TAD 및/또는 2개의 TAD는, 예를 들어 개재 아미노산 서열 없이 직접 접합될 수 있거나, DBD 및 TAD 및/또는 2개의 TAD는 펩티드 링커를 사용하여 접합될 수 있거나, 또는 그의 조합일 수 있다. 특정 실시양태에서, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 60, 70, 75, 80, 90 또는 100개의 아미노산 또는 1-5, 1-10, 1-20, 1-30, 1-40, 1-50, 1-75, 1-100, 5-10, 5-20, 5-30, 5-40, 5-50, 5-75, 5-100, 10-20, 10-30, 10-40, 10-50, 10-75, 10-100, 20-30, 20-40, 20-50, 20-75 또는 20-100개의 아미노산을 갖는 링커를 통해 DBD가 TAD에 접합되고/거나 1개의 TAD가 제2 TAD에 접합된다. 일부 경우에, DBD 및 TAD 및/또는 2개의 TAD는 이러한 도메인이 유래된 자연 발생 단백질에서 발견되는 자연 발생 개재 잔기를 통해 접합된다. 다른 실시양태에서, DBD 및 TAD 및/또는 2개의 TAD는 합성 또는 외인성 링커 서열을 통해 접합된다. 적합한 링커는 가요성, 절단가능, 비-절단가능, 친수성 및/또는 소수성일

수 있다. 특정 실시양태에서, DBD 및 TAD 및/또는 2개의 TAD는 복수의 글리신 및/또는 세린 잔기를 포함하는 링커를 통해 함께 융합될 수 있다. 글리신/세린 펩티드 링커의 예는 [GS]n, [GGGS]n (서열식별번호: 179), [GGGGS]n (서열식별번호: 180), [GGSG]n (서열식별번호: 181)을 포함하며, 여기서 n은 1 이상의 정수이다. 특정 실시양태에서, DBD 및 TAD 및/또는 2개의 TAD를 접합시키는데 유용한 링커는 GGSGGGSG (서열식별번호: 177)이다. 특정 실시양태에서, DBD 및 TAD 및/또는 2개의 TAD를 접합시키는데 유용한 링커는 GGSGGGSGGGSGGGSG (서열식별번호: 178)이다. 특정 실시양태에서, DBD가 2개의 TAD에 접합되는 경우에, 제1 및 제2 TAD는 동일하거나 상이한 링커로 DBD에 접합될 수 있거나, 또는 하나의 TAD는 링커로 DBD에 접합될 수 있고 다른 TAD는 DBD에 직접 (예를 들어, 개재 링커 서열 없이) 접합될 수 있거나, 또는 TAD 둘 다는 DBD에 직접 (예를 들어, 개재 링커 서열 없이) 접합될 수 있다. 특정 실시양태에서, DBD가 동일한 말단 (예를 들어, N-말단 또는 C-말단) 상의 2개의 TAD에 접합되는 경우에, 2개의 TAD를 연결하는 링커는 TAD를 DBD에 연결하는 링커와 동일하거나 상이할 수 있거나, 또는 TAD는 링커로 서로 접합될 수 있지만 TAD는 DBD에 직접 (예를 들어, 개재 링커 서열 없이) 접합되거나, 또는 TAD는 서로 직접 (예를 들어, 개재 링커 서열 없이) 접합될 수 있지만 TAD는 링커로 DBD에 접합된다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 DBD와 1개 이상의 TAD 사이에 위치한 1개 이상의 HA 태그(들) (예를 들어, 서열식별번호: 171)를 포함하지 않는다.

[0106] SCN1A를 상향조절하는 본원에 제공된 eTF는 자연 발생 전사 인자와 상이한 특성을 갖는다. 특정 실시양태에서, SCN1A를 상향조절하는 eTF는 DBD가 유래된 자연 발생 단백질과 비교하여 상이한 표적 부위에 결합하도록 변형된 자연 발생 단백질로부터 유래된 DBD를 포함하고, 이러한 변형된 DBD를 포함하는 eTF는 DBD가 유래된 자연 발생 단백질 (예를 들어, SCN1A 이외의 유전자)과 비교하여 상이한 유전자 (예를 들어, SCN1A)로부터의 발현을 조정한다. 다른 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 변형된 TAD를 포함하는 eTF가 TAD가 유래된 자연 발생 단백질 (예를 들어, SCN1A 이외의 유전자)과 비교하여 상이한 유전자 (예를 들어, SCN1A)로부터의 발현을 조정하도록 변형되고/거나 이러한 변형된 TAD를 포함하는 eTF가 TAD가 유래된 자연 발생 단백질과 비교하여 SCN1A의 발현을 상이하게 조정 (예를 들어, 상향조절 대 하향조절)하도록 변형된 자연 발생 단백질로부터 유래된 TAD를 포함한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 자연 발생 단백질로부터 유래된 DBD 및 자연 발생 단백질 (동일하거나 상이한 단백질)로부터 유래된 TAD를 포함하며, 여기서 DBD 및 TAD는 둘 다 변형되었다. 이러한 실시양태에서, DBD는 그것이 유래된 자연 발생 단백질과 비교하여 상이한 표적 부위에 결합할 수 있고/거나, 이러한 변형된 DBD 및 TAD를 포함하는 eTF는 도메인이 유래된 자연 발생 단백질 (예를 들어, SCN1A 이외의 유전자(들))과 비교하여 상이한 유전자 (예를 들어, SCN1A)로부터의 발현을 조정하고/거나, 이러한 변형된 DBD 및 TAD를 포함하는 eTF는 DBD 및 TAD 도메인이 유래된 자연 발생 단백질과 비교하여 SCN1A의 발현을 상이하게 조정한다 (예를 들어, 상향조절 대 하향조절함).

[0107] DNA 결합 도메인 (DBD)

[0108] SCN1A를 상향조절하는 본원에 제공된 eTF는 관심 표적 부위 (예를 들어, 본원에 제공된 eTF에 의해 결합되는 경우에 SCN1A의 상향조절을 유발하는 표적 부위)에 결합하는 임의의 적합한 DBD를 포함할 수 있다. 특정 실시양태에서, DBD는 합성적으로 설계된 DBD일 수 있다. 다른 실시양태에서, DBD는 자연 발생 단백질로부터 유래될 수 있다. DBD 패밀리는 염기성 헬릭스-루프-헬릭스 (bHLH) (예를 들어, c-Myc), 염기성-류신 지퍼 (예를 들어, C/EBP), 헬릭스-턴-헬릭스 (예를 들어, Oct-1) 및 아연 핑거 (예를 들어, EGR1 또는 EGR3)를 포함한다. 이들 패밀리는 광범위한 DNA 결합 특이성 및 유전자 표적을 나타낸다. 본원에서 고려된 바와 같이, 공지된 인간 전사 인자 단백질 중 어느 하나는 특이적 표적 부위를 인식하도록 DBD를 조작 및/또는 재프로그램화하여 내인성 SCN1A 유전자의 발현을 조정하기 위한 단백질 플랫폼으로서의 역할을 할 수 있다. 예시적인 실시양태에서, 본원에 제공된 DBD는 아연 핑거 도메인, TALEN 결합 도메인 또는 gRNA/Cas 복합체를 포함한다.

[0109] 본원에 제공된 DBD는 SCN1A의 상향조절을 유발하는 임의의 표적 부위를 인식하도록 설계될 수 있다. 예시적인 실시양태에서, DBD는 게놈 위치를 인식하고 eTF에 의해 결합되는 경우에 내인성 SCN1A 유전자의 발현을 상향조절하도록 설계된다. 본원에 제공된 eTF에 의해 결합되는 경우에 내인성 SCN1A 유전자의 발현을 조정할 수 있는 결합 부위는 SCN1A 유전자 발현의 조정을 유발하는 곳이라면 게놈 내의 어디에나 위치할 수 있다. 다양한 실시양태에서, 결합 부위는 SCN1A와 상이한 염색체 상에, SCN1A와 동일한 염색체 상에, SCN1A 유전자의 전사 개시 부위 (TSS)의 상류에, SCN1A 유전자의 TSS의 하류에, SCN1A 유전자의 TSS에 대해 근위에, SCN1A 유전자에 대해 원위에, SCN1A 유전자의 코딩 영역 내에, SCN1A 유전자의 인트론 내에, SCN1A 유전자의 폴리A 테일의 하류에, SCN1A 유전자를 조절하는 프로모터 서열 내에 또는 SCN1A 유전자를 조절하는 인핸서 서열 내에 위치할 수 있다.

[0110] DBD는, 예를 들어 오프-타겟 결합이 최소이거나 없이 DBD에 의한 표적 결합 부위 서열의 특이적 인식을 제공하는 한, 임의의 길이의 표적 결합 부위에 결합하도록 설계될 수 있다. 특정 실시양태에서, 표적 결합 부위는

eTF에 의해 결합되는 경우에 다른 모든 유전자와 비교하여 적어도 2배, 5배, 10배, 20배, 50배, 75배, 100배, 150배, 200배, 250배, 500배 또는 그 초과인 수준으로 SCN1A의 발현을 조정할 수 있다. 특정 실시양태에서, 표적 결합 부위는 eTF에 의해 결합되는 경우에 40개의 가장 가까운 이웃 유전자 (예를 들어, SCN1A의 코딩 서열의 상류 또는 하류에 있는, 염색체 상에 가장 가깝게 위치한 40개의 유전자)와 비교하여 적어도 2배, 5배, 10배, 20배, 50배, 75배, 100배, 150배, 200배, 250배, 500배 또는 그 초과인 수준으로 SCN1A의 발현을 조정할 수 있다. 특정 실시양태에서, 표적 결합 부위는 적어도 5 bp, 10 bp, 15 bp, 20 bp, 25 bp, 30 bp, 35 bp, 40 bp, 45 bp 또는 50 bp 또는 그 초과일 수 있다. 결합 부위의 특정 길이는 eTF 내의 DBD 유형에 의해 정보가 제공될 것이다. 일반적으로, 결합 부위의 길이가 길수록, 유전자 발현의 결합 및 조정에 대한 특이성이 더 크다 (예를 들어, 결합 부위가 길수록 오프-타겟 효과가 더 적음). 특정 실시양태에서, 보다 긴 표적 결합 부위를 인식하는 DBD를 갖는 eTF는 보다 짧은 표적 부위에 결합하는 DBD를 갖는 eTF에서 관찰된 오프-타겟 효과에 비해 비-특이적 결합과 연관된 오프-타겟 효과 (예컨대, 예를 들어 오프-타겟 유전자 또는 SCN1A 이외의 유전자의 발현의 조정)가 더 적다. 일부 경우에, 오프-타겟 결합의 감소는 더 짧은 표적 결합 부위를 인식하는 DBD를 갖는 대등한 eTF와 비교하여 적어도 1.2, 1.3, 1.4, 1.5, 2, 3, 4, 5, 6, 7, 8, 9 또는 10배 더 낮다.

[0111] 특정 실시양태에서, 본원에 제공된 DBD는 증가된 결합 친화도를 갖도록 변형되어, 그가 보다 긴 기간 동안 표적 결합 부위에 결합함으로써, DBD에 접합된 TAD가 보다 많은 전사 인자를 동원하고/거나 이러한 전사 인자를 보다 긴 기간 동안 동원하여 내인성 SCN1A 유전자의 발현 수준에 대해 보다 큰 효과를 발휘하도록 할 수 있다. 특정 실시양태에서, DBD는 목적하는 표적 부위에 대한 그의 특이적 결합 (또는 온-타겟 결합)을 증가시키도록 변형될 수 있고/거나 그의 비-특이적 또는 오프-타겟 결합을 감소시키도록 변형될 수 있다.

[0112] 다양한 실시양태에서, DBD 또는 eTF와 표적 결합 부위 사이의 결합은 다양한 방법을 이용하여 결정될 수 있다. 특정 실시양태에서, DBD 또는 eTF와 표적 결합 부위 사이의 특이적 결합은 이동성 변화 검정, DNase 보호 검정, 또는 단백질-DNA 결합을 검정하기 위한 관련 기술분야에 공지된 임의의 다른 시험관내 방법을 사용하여 결정될 수 있다. 다른 실시양태에서, eTF와 표적 결합 부위 사이의 특이적 결합은 기능적 검정을 사용하여, 예를 들어 표적 결합 부위가 eTF에 의해 결합되는 경우에 유전자 (예를 들어, SCN1A)의 발현 (RNA 또는 단백질)을 측정함으로써 결정될 수 있다. 예를 들어, 표적 결합 부위는 세포에 함유되거나 세포의 게놈 내로 통합된 벡터 상의 리포터 유전자 (예컨대, 예를 들어 eGFP) 또는 SCN1A 유전자의 상류에 위치할 수 있으며, 여기서 세포는 eTF를 발현한다. 대안적으로, eTF를 발현하는 벡터는 SCN1A 유전자를 자연 함유하는 세포 유형 내로 도입될 수 있다. 대조군 (예를 들어, eTF가 없거나 또는 상이한 표적 부위를 인식하는 eTF)과 비교하여 eTF의 존재 하에 리포터 유전자 (또는 SCN1A)의 발현 수준이 더 큰 것은 eTF의 DBD가 표적 부위에 결합한다는 것을 나타낸다. 적합한 시험관내 (예를 들어, 비-세포 기반) 전사 및 번역 시스템이 또한 유사한 방식으로 사용될 수 있다. 특정 실시양태에서, 표적 부위에 결합하는 eTF는 대조군 (예를 들어, eTF가 없거나 상이한 표적 부위를 인식하는 eTF)과 비교하여 리포터 유전자 또는 SCN1A의 적어도 2배, 3배, 5배, 10배, 15배, 20배, 30배, 50배, 75배, 100배, 150배 또는 그 초과의 발현을 가질 수 있다.

[0113] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 적어도 9bp, 12bp, 15bp, 18bp, 21bp, 24bp, 27bp, 30bp, 33bp 또는 36bp 크기; 9bp, 12bp, 15bp, 18bp, 21bp, 24bp, 27bp 또는 30bp 초과; 또는 9-33bp, 9-30bp, 9-27bp, 9-24bp, 9-21bp, 9-18bp, 9-15bp, 9-12bp, 12-33bp, 12-30bp, 12-27bp, 12-24bp, 12-21bp, 12-18bp, 12-15bp, 15-33bp, 15-30bp, 15-27bp, 15-24bp, 15-21bp, 15-18bp, 18-33bp, 18-30bp, 18-27bp, 18-24bp, 18-21bp, 21-33bp, 21-30bp, 21-27bp, 21-24bp, 24-33bp, 24-30bp, 24-27bp, 27-33bp, 27-30bp 또는 30-33bp인 표적 결합 부위를 인식한다. 예시적 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 18-27bp, 18bp 또는 27bp인 표적 결합 부위를 인식한다.

[0114] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 염색체 2 상에 위치하는 표적 결합 부위를 인식한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 염색체 2 상에 위치하는 SCN1A의 TSS의 상류 또는 하류 110 kb, 100 kb, 90 kb, 80 kb, 70 kb, 60 kb, 50 kb, 40 kb, 30 kb, 20 kb, 10 kb, 5 kb, 4 kb, 3 kb, 2 kb 또는 1 kb 내의 표적 결합 부위를 인식한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 염색체 2 상에 위치하는 SCN1A의 TSS의 상류 110 kb 내의 표적 결합 부위를 인식한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 염색체 2 상에 위치하는 SCN1A의 TSS의 하류 110 kb 내의 표적 결합 부위를 인식한다. 예시적 실시양태에서, 이러한 표적 결합 부위는 18-27bp, 18bp 또는 27bp이다.

[0115] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 염색체 2 상에 위치하는 위치 166179652-165989571, 위치 166128050-166127958, 위치 166155414-166140590, 위치 166179652-1661777272 또는 위치 1659990246-165989592 (모두 GRCh38.p12를 참조함) 내의 표적 결합 부위를 인식한다. 예시적 실시양태에서,

이러한 표적 결합 부위는 18-27bp, 18bp 또는 27bp이다.

[0116] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 염색체 2 상의 166178880, 166177369, 166177362, 166177299, 166155393, 166155264, 166149373, 166149176, 166149165, 166149118, 166148953, 166148565, 166142396, 166142391, 166142344, 166142239, 166141162, 166140928, 166140590, 165990076, 165989684, 165989571, 166155255, 166155099, 166148843, 166148361, 166142219, 166141090, 165990246, 165990193, 166149168, 166127991, 166128002, 166128037 또는 166128025 (모두 GRCh38.p12를 참조함)로부터 선택된 위치와 중첩되며, (iii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 1.2배 증가를 가져올 수 있는 표적 결합 부위를 인식한다.

[0117] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 18, 25, 30, 31 또는 35-66 중 어느 것을 포함하거나 그로 이루어진 표적 부위에 결합하고, (ii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 1.2배 증가를 가져올 수 있다.

[0118] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 염색체 2 상의 166155255, 166155099, 166148843, 166148361, 166142219, 166141090, 165990246, 165990193, 166149168, 166127991, 166128002, 166128037 또는 166128025 (모두 GRCh38.p12를 참조함)로부터 선택된 위치와 중첩되며, (iii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 2배 증가를 가져올 수 있는 표적 결합 부위를 인식한다.

[0119] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 18, 30, 31, 37, 38, 45, 47, 48, 49, 55, 61, 62 또는 64 중 어느 것을 포함하거나 그로 이루어진 표적 부위에 결합하고, (ii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 2배 증가를 가져올 수 있다.

[0120] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 염색체 2 상의 166149168, 166127991, 166128002, 166128037 또는 166128025 (모두 GRCh38.p12를 참조함)로부터 선택된 위치와 중첩되며, (iii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 5배 증가를 가져올 수 있는 표적 결합 부위를 인식한다.

[0121] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 18, 30, 31, 37 또는 38 중 어느 것을 포함하거나 그로 이루어진 표적 부위에 결합하고, (ii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 5배 증가를 가져올 수 있다.

[0122] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 염색체 2 상의 166128002, 166128037 또는 166128025 (모두 GRCh38.p12를 참조함)로부터 선택된 위치와 중첩되며, (iii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 15배 증가를 가져올 수 있는 표적 결합 부위를 인식한다.

[0123] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 30, 37 또는 38 중 어느 것을 포함하거나 그로 이루어진 표적 부위에 결합하고, (ii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 15배 증가를 가져올 수 있다.

[0124] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 염색체 2 상의 166128037 또는 166128025 (모두 GRCh38.p12를 참조함)로부터 선택된 위치와 중첩되며, (iii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 20배 증가를 가져올 수 있는 표적 결합 부위를 인식한다.

[0125] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 30 또는 38 중 어느 것을 포함하거나 그로 이루어진 표적 부위에 결합하고, (ii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 20배 증가를 가져올 수 있다.

[0126] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 염색체 2 상의 위치 166128025에서의 위치와 중첩되며, (iii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 25배 증가를 가져올 수 있는 표적 결합 부위를 인식한다.

[0127] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 30을 포함하거나 그로 이루어진 표적 부위에 결합하고, (ii) 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 25배 증가

를 가져올 수 있다.

[0128] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 서열식별번호: 18, 25, 30, 31 또는 35-66 중 어느 하나의 서열을 갖는 게놈 위치의 적어도 1 kb, 750 bp, 500 bp, 400 bp, 300 bp, 200 bp, 100 bp 또는 50 bp 내에 있는 게놈 영역에 결합하는 표적 결합 부위를 인식한다. 특정 실시양태에서, 표적 결합 부위는 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 1.2배, 2배, 5배, 15배, 20배 또는 25배 증가를 가져올 수 있다.

[0129] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 18-27bp, 18bp 또는 27bp이고, (ii) 서열식별번호: 18, 25, 30, 31 또는 35-66 중 어느 하나의 서열을 갖는 게놈 위치와 적어도 부분적으로 중첩되는 게놈 영역에 결합하는 표적 결합 부위를 인식한다. 특정 실시양태에서, 표적 결합 부위는 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 1.2배, 2배, 5배, 15배, 20배 또는 25배 증가를 가져올 수 있다.

[0130] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 하기 서열: 서열식별번호: 18, 25, 30, 31 또는 35-66 중 어느 하나를 갖는 표적 결합 부위를 인식한다. 특정 실시양태에서, 표적 결합 부위는 본원에 개시된 eTF에 의해 결합되는 경우에 SCN1A의 발현의 적어도 1.2배, 2배, 5배, 15배, 20배 또는 25배 증가를 가져올 수 있다.

[0131] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 대조군 (예를 들어, eTF가 없거나 또는 표적 부위를 인식하지 않는 eTF)과 비교하여 세포 또는 생체내에서 SCN1A 발현 (SCN1A RNA 및/또는 Nav1.1 단백질)의 적어도 1.5배, 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 15배, 20배, 25배, 50배, 100배 또는 그 초과 또는 적어도 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과의 상향조절을 유발한다. 다양한 실시양태에서, SCN1A 발현의 상향조절은 PCR 방법, 웨스턴 블롯 또는 면역검정을 사용하여 검출될 수 있다.

[0132] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 전사 활성화 검정에서 대조군에 비해 SCN1A 발현을 적어도 1.2배, 1.3배, 1.4배, 1.5배, 1.6배, 1.7배, 1.8배, 1.9배, 2배, 3배, 4배, 5배, 8배, 10배, 12배, 15배, 18배, 20배, 25배, 30배, 40배, 50배, 75배, 100배 또는 그 초과만큼 또는 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 증가시킬 수 있는 표적 부위에 결합한다. 예시적인 SCN1A 전사 활성화 검정이 본원의 실시예 3에 제공된다. 간략하게, HEK293을 eTF 또는 대조군 eGFP 리포터 구축물을 보유하는 플라스미드로 형질감염시킨다. 형질감염 48시간 후에, 세포를 수집하고, RNA를 단리하고, 역전사시키고, 생성된 cDNA 샘플을 qPCR (예를 들어, 서열식별번호: 185 및 186을 갖는 프라이머를 사용함)에 의해 분석하여 내인성 SCN1A 전사체의 수준을 정량화한다. GAPDH는 SCN1A 발현의 상대 수준을 결정하기 위한 참조 유전자로서 사용될 수 있다.

[0133] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 최소의 오프-타겟 효과, 예를 들어 비-특이적 결합과 연관된 오프-타겟 효과를 가지며, 예컨대, 예를 들어 오프-타겟 유전자 또는 SCN1A 이외의 유전자의 발현을 조정한다. 한 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 대조군과 비교하여 1개 이상의 오프-타겟 유전자에 대한 eTF에 의해 생산된 발현보다 적어도 5배, 10배, 15배, 20배, 25배, 30배, 40배 또는 50배 더 크게, 대조군과 비교하여 SCN1A를 특이적으로 상향조절한다. 예시적 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 대조군에 비해 eTF에 의해 생산된 40개의 가장 가까운 이웃 유전자 (예를 들어, 염색체 2 상의 SCN1A의 코딩 서열에 위치한 40개의 가장 가까운 유전자), 예를 들어 PLA2R1, ITGB6, RBMS1, TANK, PSMD14, TBR1, SLC4A10, DPP4, FAP, IFIH1, GCA, FIGN, GRB14, COBLL1, SLC38A11, SCN3A, SCN2A, CSRNP3, GALNT3, TTC21B, SCN9A, SCN7A, B3GALT1, STK39, CERS6, NOSTRIN, SPC25, ABCB11, DHRS9, BBS5, KLHL41, FASTKD1, PPIG, CCDC173, PHOSPHO2, KLHL23, SSB, METTL5, UBR3 및 MYO3B의 전사보다 적어도 15배 더 크게, 대조군과 비교하여 SCN1A 유전자로부터의 전사를 특이적으로 상향조절한다 (표 14 참조). 다양한 실시양태에서, SCN1A 유전자로부터의 전사의 상향조절은 PCR 방법을 사용하여 검출될 수 있다.

[0134] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 드라베 증후군의 $Scn1a^{tm1kea}$ 마우스 모델에서의 고체온 발작 (HTS) 검정에서 발작의 빈도를 감소시킬 수 있다. 특정 실시양태에서, 본원에 개시된 eTF는 HTS 검정에서 42.6℃에서의 발작 빈도를 대조군 (예를 들어, PBS 처리되거나 또는 eGFP를 코딩하는 서열을 포함하는 AAV 벡터로 처리됨)과 비교하여 적어도 1.2배, 1.3배, 1.4배, 1.5배, 1.6배, 1.7배, 1.8배, 1.9배, 2.0배 또는 그 초과만큼 또는 적어도 20%, 30% 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100% 또는 그 초과만큼 감소시킬 수 있다. 특정 실시양태에서, 본원에 개시된 eTF는 HTS 검정이 실행된 마우스의 적어도 60%, 62%, 65%, 70%, 75%, 76%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또

는 99%가 42.6℃에서의 발작이 없도록 상기 검정에서 42.6℃에서의 발작의 빈도를 감소시킬 수 있다. 예시적인 HTS 검정은 본원의 실시예 6에 기재되어 있다. 간략하게, 암컷 C57Bl/6J 마우스와 교배된 수컷 Scn1a +/- 마우스로부터 생산된 한배새끼들에게 본원에 제공된 바와 같은 SCN1A를 상향조절하는 eTF를 코딩하는 AAV9 벡터 또는 eGFP를 코딩하는 대조군 벡터를 P1에서 양측 ICV를 통해 투여할 수 있다. 마우스에게 ~1.0E10-5.0E12 gc/마우스를 투여할 수 있다. HTS 검정은 혼합된 129Stac X C57BL/6 배경의 P26-P28 SCN1A 이형접합 마우스 및 SCN1A 야생형 마우스에서 자세 상실이 동반된 제1 긴장성-간대성 발작의 발병까지 또는 43℃의 체온에 도달할 때까지 마우스의 체온을 2분마다 ~0.5℃만큼 증가시킴으로써 (제어 조건 하에 체온 모니터링을 수반함) 수행된다. 마우스는 실험의 전체 과정에 걸쳐 자세 상실이 동반된 발작이 검출되지 않는 경우에 발작이 없는 것으로 간주된다.

[0135] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 SCN1A에 대해 이형접합인 마우스, 예를 들어 $Scn1a^{tm1kea}$ 마우스 계통의 생존을 증가시킬 수 있다. 특정 실시양태에서, 본원에 개시된 eTF는 P100에서 SCN1A 이형접합 마우스의 생존율을 대조군 (예를 들어, PBS 처리되거나 또는 eGFP를 코딩하는 서열을 포함하는 AAV 벡터로 처리됨)과 비교하여 적어도 1.2배, 1.3배, 1.4배, 1.5배, 1.6배, 1.7배, 1.8배, 1.9배, 2.0배 또는 그 초과만큼 또는 적어도 20%, 30% 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100% 또는 그 초과만큼 증가시킬 수 있다. 특정 실시양태에서, 본원에 개시된 eTF는 검정이 실행된 마우스의 적어도 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%가 P100에서 여전히 살아있도록 P100에서 SCN1A 이형접합 마우스의 생존율을 증가시킬 수 있다. 예시적인 생존 검정이 본원의 실시예 7에 기재되어 있다. 간략하게, 암컷 C57Bl/6J 마우스와 교배된 수컷 Scn1a +/- 마우스로부터 생산된 한배새끼들에게 P1에서 양측 ICV를 통해 AAV9 벡터를 투여할 수 있다. 마우스에게 ~1.0E10-5.0E12 gc/마우스를 투여할 수 있다. P100까지 생존한 마우스의 수를 결정한다.

[0136] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 아연 핑거 단백질로부터의 것이거나, 아연 핑거 단백질로부터 유래되거나 또는 뉴클레아제 불활성화된 아연 핑거 단백질인 DBD를 포함할 수 있다. 아연 핑거는 폴드를 안정화시키기 위해 1개 이상의 아연 이온 (Zn^{2+})의 배위를 특징으로 하는 작은 단백질 구조적 모티프이다. 아연 핑거 (Znf) 도메인은 DNA 표적 부위와 탠덤 접촉하는 다수의 핑거-유사 돌출부를 함유하는 비교적 작은 단백질 모티프이다. 아연 핑거 모티프의 모듈 성질은 DNA 서열의 다수의 조합이 고도의 친화도 및 특이성으로 결합되는 것을 가능하게 하고, 따라서 특이적 DNA 서열에 표적화하고 결합할 수 있는 단백질의 조작에 이상적으로 적합하다. 많은 조작된 아연 핑거 어레이는 뮤린 전사 인자 Zif268의 아연 핑거 도메인에 기초한다. Zif268은 9 bp 서열에 고친화도로 집합적으로 결합하는 3개의 개별 아연 핑거 모티프를 갖는다. 매우 다양한 아연 핑거 단백질이 확인되었고, 본원에 추가로 기재된 바와 같이 구조에 기초하여 상이한 유형으로 특징화된다. 임의의 이러한 아연 핑거 단백질은 본원에 기재된 DBD와 관련하여 유용하다.

[0137] 아연 핑거 단백질을 설계하는 다양한 방법이 이용가능하다. 예를 들어, 관심 표적 DNA 서열에 결합하도록 아연 핑거 단백질을 설계하는 방법이 기재되어 있다 (예를 들어, 문헌 [Liu Q, et al., Design of polydactyl zinc-finger proteins for unique addressing within complex genomes, Proc Natl Acad Sci USA. 94 (11): 5525-30 (1997); Wright DA et al., Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly, Nat Protoc . Nat Protoc. 2006;1(3):1637-52; 및 CA Gersbach and T Gaj, Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies, Am Chem Soc 47: 2309-2318 (2014)] 참조). 추가로, 관심 DNA 표적 서열에 결합하도록 아연 핑거 단백질을 설계하기 위한 다양한 웹 기반 도구가 공개적으로 입수가능하다 (예를 들어, 월드 와이드 웹 omictools.com/zfns-category에서 이용가능한 OmicX로부터의 아연 핑거 뉴클레아제 설계 소프트웨어 툴 및 게놈 조작 데이터 분석 웹사이트; 및 월드 와이드 웹 scripps.edu/barbas/zfdesign/zfdesignhome.php에서 이용가능한 Scripps로부터의 아연 핑거 툴 설계 웹사이트 참조). 추가로, 관심 DNA 표적 서열에 결합하도록 아연 핑거 단백질을 설계하기 위한 다양한 상업적으로 입수가능한 서비스가 이용가능하다 (예를 들어, 크리에이티브 바이오랩스(Creative Biolabs)에 의해 제공된 상업적으로 입수가능한 서비스 또는 키트 (월드 와이드 웹 creative-biolabs.com/Design-and-Synthesis-of-Artificial-Zinc-Finger-Proteins.html), 애드진(Addgene)으로부터 이용가능한 아연 핑거 컨소시엄 모듈 어셈블리 키트 (월드 와이드 웹 addgene.org/kits/zfc-modular-assembly/), 또는 시그마 알드리치(Sigma Aldrich)로부터의 CompoZr 커스텀 ZFN 서비스 (월드 와이드 웹 sigmaaldrich.com/life-science/zinc-finger-nuclease-technology/custom-zfn.html) 참조).

[0138] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 1개 이상의 아연 핑거를 포함하는 DBD를 포함하

거나 또는 아연 핑거 단백질의 DBD로부터 유래된다. 일부 경우에, DBD는 다수의 아연 핑거를 포함하며, 여기서 각각의 아연 핑거는 그의 N-말단 또는 C-말단 또는 둘 다에서 아미노산 링커를 통해 또 다른 아연 핑거 또는 또 다른 도메인에 연결된다. 일부 경우에, 본원에 제공된 DBD는 표 9에 기재된 아연 핑거 유형 중 1개 이상으로부터의 1개 이상의 아연 핑거를 포함한다. 일부 경우에, 본원에 제공된 DBD는 복수의 아연 핑거 구조 또는 모티프, 또는 서열식별번호: 152-167 중 1개 이상을 갖는 복수의 아연 핑거 또는 그의 임의의 조합을 포함한다. 특정 실시양태에서, DBD는 X-[ZF-X]n 및/또는 [X-ZF]n-X를 포함하며, 여기서 ZF는 표 9에 열거된 모티프 중 어느 하나 (예를 들어, 서열식별번호: 136-146 중 어느 하나)를 갖는 아연 핑거 도메인이고, X는 1-50개의 아미노산을 포함하는 아미노산 링커이고, n은 1-15의 정수, 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15이고, 각각의 ZF는 독립적으로 DBD 내의 다른 ZF 서열과 동일한 서열 또는 상이한 서열을 가질 수 있고, 각각의 링커 X는 독립적으로 DBD 내의 다른 X 서열과 동일한 서열 또는 상이한 서열을 가질 수 있다. 각각의 아연 핑거는 그의 C-말단, N-말단 또는 둘 다에서 또 다른 서열, 아연 핑거 또는 도메인에 연결될 수 있다. DBD에서, 각각의 링커 X는 서열, 길이 및/또는 성질 (예를 들어, 가요성 또는 전하)이 동일할 수 있거나 또는 서열, 길이 및/또는 성질이 상이할 수 있다. 일부 경우에, 2개 이상의 링커는 동일할 수 있는 반면, 다른 링커는 상이하다. 예시적 실시양태에서, 링커는 표 9에 제공된 1개 이상의 자연 발생 아연 핑거 단백질에서 발견되는 아연 핑거를 연결하는 서열로부터 수득되거나 유래될 수 있다. 다른 실시양태에서, 적합한 링커 서열은, 예를 들어 5개 이상의 아미노산 길이의 링커를 포함한다. 또한, 6개 이상의 아미노산 길이의 예시적인 링커 서열에 대해서는 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조하며, 이들 각각은 그 전문이 본원에 포함된다. 본원에 제공된 DBD 단백질은 단백질의 개별 아연 핑거들 사이에 적합한 링커의 임의의 조합을 포함할 수 있다. 본원에 기재된 DBD 단백질은 단백질의 개별 아연 핑거들 사이에 적합한 링커의 임의의 조합을 포함할 수 있다.

[0139] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 1개 이상의 전형적 아연 핑거를 포함하는 DBD를 포함한다. 전형적 C2H2 아연-핑거는 아연 이온에 의해 배위되는, 하나의 쇄 내의 2개의 시스테인 및 또 다른 쇄 내의 2개의 히스티딘 잔기를 갖는다. 전형적 아연-핑거 도메인은 2개의 β-시트 및 1개의 α-헬릭스를 가지며, 여기서 α-헬릭스는 DNA 분자와 상호작용하고, 표적 부위에 대한 DBD 결합의 기초를 형성하고, "인식 헬릭스"로 지칭될 수 있다. 예시적 실시양태에서, 아연 핑거의 인식 헬릭스는 위치 -1, 2, 3 또는 6에서 적어도 1개의 아미노산 치환을 포함함으로써 아연 핑거 도메인의 결합 특이성을 변화시킨다. 다른 실시양태에서, 본원에 제공된 DBD는 1개 이상의 비-전형적 아연-핑거, 예를 들어 C2-H2, C2-CH 및 C2-C2를 포함한다.

[0140] 또 다른 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 하기 구조를 갖는 아연 핑거 모티프를 포함하는 DBD를 포함한다: LEPGEKP - [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH - TGKKTS (서열식별번호: 147), 여기서 n은 1-15의 정수, 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15이고, 각각의 X는 독립적으로 3 bp의 표적 서열에 결합할 수 있는 인식 서열 (예를 들어, 인식 헬릭스)이다. 예시적 실시양태에서, n은 3, 6 또는 9이다. 특히 바람직한 실시양태에서, n은 6이다. 다양한 실시양태에서, 각각의 X는 독립적으로 DBD 내의 다른 X 서열과 비교하여 동일한 아미노산 서열 또는 상이한 아미노산 서열을 가질 수 있다. 예시적 실시양태에서, 각각의 X는 월드 와이드 웹 scripps.edu/barbas/zfdesign/zfdesignhome.php에 위치한 Scripps로부터의 아연 핑거 설계 툴을 사용하여 3 bp의 관심 표적 결합 부위와 상호작용하도록 설계된 7개의 아미노산을 포함하는 서열이다.

[0141] DBD 내의 각각의 아연 핑거가 3 bp를 인식하기 때문에, DBD에 포함된 아연 핑거의 수는 DBD에 의해 인식되는 결합 부위의 길이에 대한 정보를 제공하며, 예를 들어 1개의 아연 핑거를 수반한 DBD는 3 bp를 갖는 표적 결합 부위를 인식할 것이고, 2개의 아연 핑거를 수반한 DBD는 6 bp를 갖는 표적 결합 부위를 인식할 것이고, 3개의 아연 핑거를 수반한 DBD는 9 bp를 갖는 표적 결합 부위를 인식할 것이고, 4개의 아연 핑거를 수반한 DBD는 12 bp를 갖는 표적 결합 부위를 인식할 것이고, 5개의 아연 핑거를 수반한 DBD는 15 bp를 갖는 표적 결합 부위를 인식할 것이고, 6개의 아연 핑거를 수반한 DBD는 18 bp를 갖는 표적 결합 부위를 인식할 것이고, 9개의 아연 핑거를 수반한 DBD는 27 bp를 갖는 표적 결합 부위를 인식할 것이라는 점 등이다. 일반적으로, 보다 긴 표적 결합 부위를 인식하는 DBD는 보다 큰 결합 특이성 (예를 들어, 보다 적은 오프-타겟 결합 또는 비-특이적 결합)을 나타낼 것이다.

[0142] 다른 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 DBD의 결합 특이성을 변화시키도록 (예를 들어, DBD에 의해 인식되는 표적 부위를 변화시키도록) 아연 핑거 도메인의 인식 헬릭스 중 1개 이상에 1개 이상의 아미노산 치환을 생성함으로써 자연 발생 아연 핑거 단백질로부터 유래된 DBD를 포함한다. 본원에 제공된 DBD는 임의의 자연 발생 아연 핑거 단백질로부터 유래될 수 있다. 다양한 실시양태에서, 이러한 DBD는 임의의 종, 예

를 들어 마우스, 래트, 인간 등의 아연 핑거 단백질로부터 유래될 수 있다. 예시적 실시양태에서, 본원에 제공된 DBD는 인간 아연 핑거 단백질로부터 유래된다. 특정 실시양태에서, 본원에 제공된 DBD는 표 9에 열거된 자연 발생 단백질로부터 유래된다. 예시적 실시양태에서, 본원에 제공된 DBD 단백질은 인간 EGR 아연 핑거 단백질, 예를 들어 EGR1, EGR2, EGR3 또는 EGR4로부터 유래된다.

[0143] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 자연 발생 단백질의 DBD 내에서 1개 이상의 아연 핑거를 반복시켜 DBD 단백질 내의 아연 핑거 도메인의 수를 증가시키도록 DBD를 변형시킴으로써 자연 발생 단백질로부터 유래된 DBD를 포함한다. 특정 실시양태에서, 이러한 변형은 자연 발생 단백질의 DBD 내의 이중반복, 삼중반복, 사중반복 또는 추가 다중반복의 아연 핑거를 포함한다. 일부 경우에, 인간 단백질의 DBD로부터의 1개의 아연 핑거가 다중반복되며, 예를 들어 동일한 아연 핑거 모티프의 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12개 또는 그 초과의 카피가 eTF의 DBD에서 반복된다. 일부 경우에, 자연 발생 단백질의 DBD로부터의 아연 핑거의 세트가 다중반복되며, 예를 들어 자연 발생 단백질의 DBD로부터의 3개 아연 핑거의 세트가 이중반복되어 6개의 아연 핑거를 수반한 DBD를 갖는 eTF가 생성되거나, 삼중반복되어 9개의 아연 핑거를 수반한 eTF의 DBD가 생성되거나, 또는 사중반복되어 12개의 아연 핑거를 수반한 eTF의 DBD가 생성되는 등이다. 일부 경우에, 자연 발생 단백질의 DBD로부터의 아연 핑거의 세트가 부분적으로 반복되어 더 많은 수의 아연 핑거를 갖는 eTF의 DBD가 형성되며, 예를 들어 eTF의 DBD는 4개의 아연 핑거를 포함하고, 여기서 이러한 아연 핑거는 eTF의 DBD 내의 총 4개의 아연 핑거에 대해 자연 발생 단백질로부터의 제1 아연 핑거의 1개 카피, 제2 아연 핑거의 1개 카피 및 제3 아연 핑거의 2개 카피를 나타낸다. 이어서, 이러한 DBD는 DBD의 결합 특이성을 변화시키도록 (예를 들어, DBD에 의해 인식되는 표적 부위를 변화시키도록) 아연 핑거 도메인의 인식 헬릭스 중 1개 이상에 1개 이상의 아미노산 치환을 만듦으로써 추가로 변형된다. 예시적 실시양태에서, DBD는 자연 발생 인간 단백질, 예컨대 인간 EGR 아연 핑거 단백질, 예를 들어 EGR1, EGR2, EGR3 또는 EGR4로부터 유래된다.

[0144] 인간 EGR1 및 EGR3은 3개 핑거의 C2H2 아연 핑거 DBD를 특징으로 한다. 아연 핑거에 대한 일반적 결합 규칙은 모든 3개의 핑거가 각각의 아연 핑거의 알파 헬릭스에서 동일한 아미노산을 사용하여 유사한 기하구조를 갖는 그의 동족 DNA 서열과 상호작용함으로써 표적 결합 부위 서열의 특이성 또는 인식을 결정하는 것을 제공한다. 이러한 결합 규칙은 목적하는 표적 결합 부위를 인식하는 DBD를 조작하도록 EGR1 또는 EGR3의 DBD를 변형시키는 것을 가능하게 한다. 일부 경우에, EGR1 또는 EGR3의 아연 핑거 모티프 내의 7-아미노산 DNA 인식 헬릭스는 공개된 아연 핑거 설계 규칙에 따라 변형된다. 특정 실시양태에서, EGR1 또는 EGR3의 3개 핑거의 DBD 내의 각각의 아연 핑거는, 예를 들어 1개 이상의 인식 헬릭스의 서열을 변경시킴으로써 및/또는 DBD 내의 아연 핑거의 수를 증가시킴으로써 변형된다. 특정 실시양태에서, EGR1 또는 EGR3은 목적하는 표적 부위에서 적어도 9, 12, 15, 18, 21, 24, 27, 30, 33, 36개 또는 그 초과의 염기 쌍의 표적 결합 부위를 인식하도록 재프로그램화된다. 특정 실시양태에서, ERG1 또는 EGR3으로부터 유래된 이러한 DBD는 적어도 3, 4, 5, 6, 7, 8, 9, 10, 11, 12개 또는 그 초과의 아연 핑거를 포함한다. 예시적 실시양태에서, DBD 내의 아연 핑거 중 1개 이상은 인식 헬릭스의 위치 -1, 2, 3 또는 6에서 적어도 1개의 아미노산 치환을 포함한다.

[0145] 다양한 실시양태에서, EGR1 또는 EGR3으로부터 유래된 DBD를 포함하는 SCN1A를 상향조절하는 eTF는 자연 발생 EGR1 또는 EGR3의 결합 특이성과 상이한 DNA 결합 특이성을 가지며, 예를 들어 DBD는 비변형된 EGR1 또는 EGR3에 의해 인식되는 결합 부위의 서열과 상이한 서열: (GCG(T/G)GGGCG) (서열식별번호: 182)을 갖는 표적 결합 부위를 인식한다.

[0146] 다른 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 gRNA/Cas 복합체인 DBD를 포함한다. CRISPR (클러스터링된 규칙적 간격의 짧은 회문식 반복부)/Cas9는 부위-특이적 게놈 표적화를 가능하게 하는 게놈 편집 도구이다. 유형 II CRISPR/Cas 시스템은 비코딩 RNA를 사용하여 Cas9 뉴클레아제가 부위-특이적 DNA 절단을 유도하도록 안내하는 원핵 적응 면역 반응 시스템이다. CRISPR/Cas9 시스템은 포유동물 세포에서 게놈 편집을 매개하는 간단한 RNA-프로그램가능한 방법을 생성하기 위해 이용되어 왔다. 단일 가이드 RNA (sgRNA)는 Cas9 뉴클레아제를 특이적 게놈 위치로 지시하기 위해 생성될 수 있고, 이어서 그 위치에 gRNA/Cas9 복합체가 결합한다. gRNA는 다양한 방법 및 도구를 사용하여 관심 표적 부위에 결합하도록 설계될 수 있다. 예를 들어, 관심 표적 DNA 서열에 결합하도록 gRNA를 설계하는 방법은 문헌 [Aach, et al. Flexible algorithm for identifying specific Cas9 targets in genomes. BioRxiv, Cold Spring Harbor Labs. doi: http://dx.doi.org/10.1101/005074 (2014); Bae, et al. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics. 30(10):1473-1475 (2014); Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotech 34, 184-191 (2016); Gratz, et al. Highly specific and efficient

CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 196(4):961-971 (2014); Heigwer, et al. E-CRISP: fast CRISPR target site identification. Nat Methods. 11(2):122-123 (2014); Ma, et al. A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes. Biomed Res Int. doi:http://doi.org/10.1155/2013/270805 (2013); Montague, et al. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 42(W1):W401-W407 (2014); Liu, et al. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. Bioinformatics. 31(22):3676-3678 (2015); Ran, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 520(7546):186-191 (2015); Wu, et al. Target specificity of the CRISPR-Cas9 system. Quant Biol. 2(2):59-70 (2015); Xiao, et al. CasOT: a genome-wide Cas9/gRNA off-target searching tool. Bioinformatics. 30(8):1180-1182 (2014); Zetsche, et al. Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas System. Cell. 163(3):759-771 (2015)]에 기재되어 있다. 추가로, 관심 DNA 표적 서열에 결합하도록 gRNA를 설계하기 위한 다양한 웹 기반 도구가 공개적으로 입수가능하다 (예를 들어, 월드 와이드 웹 atum.bio/eCommerce/cas9/input?multipleContacts=false의 AUTM으로부터 입수가능한 CRISPR gRNA 설계 도구; 월드 와이드 웹 portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design-crisprai의 브로드 인스티튜트(Broad Institute)로부터 입수가능한 CRISPRa/i gRNA 설계 도구; 월드 와이드 웹 e-crisp.org/E-CRISP/의 DKFZ 저먼 캔서 리서치 센터(DKFZ German Cancer Research Center)로부터 입수가능한 E-CRISP 설계 도구; 및 월드 와이드 웹 design.synthego.com/#/의 신테고(Synthego)로부터 입수가능한 녹아웃 가이드 설계 도구 참조). 추가로, 관심 DNA 표적 서열에 결합하도록 gRNA를 설계하기 위한 다양한 상업적으로 입수가능한 서비스가 이용가능하며, 예를 들어 IDT (월드 와이드 웹 idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE), 써모피셔(ThermoFisher) (월드 와이드 웹 thermofisher.com/order/custom-oligo/crispr) 및 진스크립트(GenScript) (월드 와이드 웹 genscript.com/gRNA-design-tool.html)에 의해 제공된 상업적으로 입수가능한 서비스를 참조한다.

[0147] 예시적 실시양태에서, gRNA/Cas 복합체인 DBD는 뉴클레아제 비활성화된 Cas 단백질 또는 dCas, 예컨대, 예를 들어 dCas9, 예컨대 뉴클레아제 비활성화된 스타필로코쿠스 아우레우스(*Staphylococcus aureus*) (dSaCas9) 또는 뉴클레아제 비활성화된 스트렙토코쿠스 피오게네스(*Streptococcus pyogenes*) Cas9 (dSpCas9)를 포함한다. gRNA는 gRNA/Cas 복합체를 목적하는 표적 부위로 표적화하는 표적화 영역 및 Cas 단백질과의 상호작용을 용이하게 하는 스캐폴드 영역을 포함하는 서열로서 제공된다. 임의의 적합한 gRNA 스캐폴드가 본원에 제공된 gRNA와 함께 사용될 수 있다. 예시적 실시양태에서, gRNA는 단일 gRNA 또는 sgRNA이고, 하기 스캐폴드 서열을 포함한다: 5'-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA-3' (서열식별번호: 183). 가이드 RNA의 표적화 영역은 스캐폴드 서열의 5' 말단에 부착되어 완전한 sgRNA를 형성한다. 특정 실시양태에서, gRNA 및 dCas 단백질은 동일한 발현 카세트로부터 발현될 수 있다. 특정 실시양태에서, U6 프로모터가 gRNA를 발현시키는데 사용된다. 다른 실시양태에서, gRNA는 dCas-TAD 단백질을 안정하게 발현하도록 조작된 세포에서, 예를 들어 dCas를 게놈 내로 안정하게 통합시킴으로써 또는 염색체외에서 안정하게 유지되는 플라스미드 상에서 발현될 수 있다.

[0148] 다른 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 TALEN으로부터의 것이거나, TALEN으로부터 유래되거나 또는 뉴클레아제 불활성화된 TALEN인 DBD를 포함할 수 있다. 전사 활성인자-유사 이펙터 뉴클레아제 (TALEN)는 DNA의 특이적 서열을 절단하도록 조작될 수 있는 뉴클레아제 도메인 및 DBD를 함유하는 제한 효소이다. TALEN은 TAL 이펙터 DNA 결합 도메인을 DNA 절단 도메인 (예를 들어, 뉴클레아제)에 접합시킴으로써 생성된다. 전사 활성인자-유사 이펙터 (TALE)는 목적하는 표적 DNA 서열에 결합함으로써 뉴클레아제 도메인을 특이적 위치로 지시하도록 조작될 수 있다.

[0149] TAL 이펙터는 크산토모나스(*Xanthomonas*) 박테리아로부터의 박테리아 단백질이다. DNA 결합 도메인은 분기 제12 및 제13 아미노산을 갖는 반복된 고도로 보존된 33-34개 아미노산 서열을 함유한다. 반복 가변 이잔기 (RVD)로 지칭되는 이들 2개의 위치는 고도로 가변적이고, 특이적 뉴클레오티드 인식과 강한 상관관계를 나타낸다. 아미노산 서열과 DNA 인식 사이의 이러한 직접적인 관계는 적절한 RVD를 함유하는 반복 절편들의 조합을 선택함으로써 목적하는 서열을 특이적으로 표적화하는 DBD의 조작을 가능하게 한다.

[0150] TALE를 설계하는 다양한 방법이 이용가능하다. 예를 들어, 관심 표적 DNA 서열에 결합하도록 TALE를 설계하는 방법은 문헌 [T. Cermak et al., Nucleic Acids Research. 39 (12): e82 (2011); F. Zhang F et al., Nature Biotechnology. 29 (2): 149-53 (2011); R. Morbitzer et al., Nucleic Acids Research. 39 (13): 5790-9 (2011); T. Li et al., Nucleic Acids Research. 39 (14): 6315-25 (2011); R. Geissler et al., PLOS One.

6(5): e19509 (2011); 및 E. Weber et al., PLOS One. 6 (5): e19722 (2011)]에 기재되어 있다. 추가로, 관심 DNA 표적 서열에 결합하도록 TALE를 설계하기 위한 다양한 웹 기반 도구가 공개적으로 입수가능하며, 예를 들어 월드 와이드 웹 e-talen.org/E-TALEN/TAL에서 이용가능한 E-Talen 및 월드 와이드 웹 tale-nt.cac.cornell.edu/node/add/single-tale에서 이용가능한 이펙터 뉴클레오티드 타겟터 2.0 도구를 참조한다. 추가로, 관심 DNA 표적 서열에 결합하도록 TALE를 설계하기 위한 다양한 상업적으로 입수가능한 서비스가 이용가능하며, 예를 들어 OmicX (월드 와이드 웹 omictools.com/), 애드진 (월드 와이드 웹 addgene.org/talen/guide/) 또는 써모피셔 (월드 와이드 웹 thermofisher.com/us/en/home/life-science/genome-editing/geneart-tals/tal-design-tool.html)에 의해 제공된 상업적으로 입수가능한 서비스를 참조한다. 추가로, TALE의 어셈블리에 적합한 올리고뉴클레오티드를 설계하기 위해 공개적으로 입수가능한 소프트웨어 프로그램 (DNAWorks)을 사용할 수 있다 (예를 들어, 문헌 [D. Hoover D Methods in Molecular Biology. 852: 215-23 (2012)] 참조).

[0151] 전사 조정 도메인

[0152] SCN1A를 상향조절하는 본원에 제공된 eTF는 전사를 조정할 수 있거나 (예를 들어, RNA 폴리머라제 II, CBP/p300, CREB 또는 KRAB) 또는 eTF가 DBD (예를 들어, 아연 핑거 DBD, gRNA/Cas DBD 또는 TALE DBD)를 통해 표적 부위에 결합되는 경우에 관심 유전자로부터의 유전자 발현 수준을 조정할 수 있는 1종 이상의 단백질 인자를 동원할 수 있는 임의의 적합한 도메인을 포함할 수 있다. 특정 실시양태에서, 이러한 도메인은 관심 유전자의 전사 또는 유전자 발현의 수준을 증가시키는 단백질 인자를 동원하며, 이는 전사 활성화 도메인 (TAD)이다. 다른 실시양태에서, 이러한 도메인은 관심 유전자로부터의 전사 또는 유전자 발현의 수준을 감소시키는 단백질 인자를 동원하며, 이는 전사 억제인자 도메인 (TRD)이다. 특정 실시양태에서, 전사 조정 도메인 (TAD 또는 TRD)은 합성적으로 설계된 도메인일 수 있다. 다른 실시양태에서, 전사 조정 도메인 (TAD 또는 TRD)은 자연 발생 단백질, 예를 들어 전사 인자, 전사 보조-활성인자, 전사 보조-억제인자 또는 사일렌서 단백질로부터 유래될 수 있다. 다양한 실시양태에서, 전사 조정 도메인 (TAD 또는 TRD)은 임의의 종, 예를 들어 마우스, 래트, 원숭이, 바이러스 또는 인간의 단백질로부터 유래될 수 있다.

[0153] 하나의 예시적 실시양태에서, SNC1A를 상향조절하는 본원에 제공된 eTF에 사용하기에 적합한 TAD는 바이러스 단백질로부터 유래된다. 바이러스 단백질로부터 유래된 예시적인 TAD는, 예를 들어 VP64 (서열식별번호: 133), VPR (서열식별번호: 132), VP16, VP128, p65, p300의 TAD 도메인 또는 그의 임의의 기능적 단편 또는 변이체 또는 그에 대해 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열을 포함한다.

[0154] 또 다른 예시적 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF에 사용하기에 적합한 TAD는 인간 단백질로부터 유래된다. 인간 단백질로부터 유래된 예시적인 TAD는, 예를 들어 CBP/p300-상호작용 전사활성인자 2 (CITED2) (서열식별번호: 134), CBP/p300-상호작용 전사활성인자 4 (CITED4) (서열식별번호: 135), EGR1 (서열식별번호: 176), CREB3 (서열식별번호: 224) 또는 EGR3 (서열식별번호: 175)의 TAD 도메인 또는 그의 임의의 기능적 단편 또는 변이체 또는 그에 대해 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열을 포함한다.

[0155] 특정 실시양태에서, SCN1A를 상향조절하는 eTF는 전사 활성화 도메인 또는 TAD와 접합되는 아연 핑거 DBD를 포함한다. 다양한 실시양태에서, 아연 핑거 DBD는 바이러스 단백질, 예컨대 VP64 또는 VPR로부터의 TAD, 또는 인간 단백질, 예컨대 CITED2, CITED4 또는 CREB3으로부터의 TAD에 접합될 수 있다. 특정 실시양태에서, 인간 단백질, 예를 들어 EGR1 또는 EGR3으로부터 유래된 아연 핑거 DBD는 인간 단백질, 예를 들어 CITED2, CITED4 또는 CREB3으로부터 유래된 TAD에 접합된다. 특정 실시양태에서, 인간 단백질, 예를 들어 EGR1 또는 EGR3으로부터 유래된 아연 핑거 DBD는 VP64 또는 VPR TAD에 접합된다. 특정 실시양태에서, 합성 아연 핑거 DBD 또는 인간 단백질, 예를 들어 EGR1 또는 EGR3에 대해 75% 미만의 서열 동일성을 갖는 아연 핑거 DBD는 인간 단백질, 예를 들어 CITED2, CITED4 또는 CREB3으로부터 유래된 TAD에 접합된다. 특정 실시양태에서, 합성 아연 핑거 DBD 또는 인간 단백질, 예를 들어 EGR1 또는 EGR3에 대해 75% 미만의 서열 동일성을 갖는 아연 핑거 DBD는 VP64 또는 VPR TAD에 접합된다.

[0156] 특정 실시양태에서, dCas 단백질은 TAD에 접합된다. 다양한 실시양태에서, dCas9는 바이러스 단백질, 예컨대 VP64 또는 VPR로부터의 TAD, 또는 인간 단백질, 예컨대 CITED2, CITED4 또는 CREB3으로부터의 TAD에 접합될 수 있다. 예시적 실시양태에서, dCas9는 VP64 또는 VPR TAD에 접합된다.

[0157] 특정 실시양태에서, TALE 단백질은 TAD에 접합된다. 다양한 실시양태에서, TALE는 바이러스 단백질, 예컨대

VP64 또는 VPR로부터의 TAD, 또는 인간 단백질, 예컨대 CITED2, CITED4 또는 CREB3으로부터의 TAD에 접합될 수 있다. 예시적 실시양태에서, TALE는 VP64 또는 VPR TAD에 접합된다.

[0158] SCN1A를 상향조절하고 인간 단백질에 고도로 상동인 eTF

[0159] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 1종 이상의 인간 단백질 (하기에 추가로 기재되는 바와 같음)에 대해 높은 퍼센트 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다. 특정 실시양태에서, 이러한 eTF는 인간 EGR1 또는 EGR3으로부터 유래된 DBD 및 인간 EGR1, EGR3, CITED2, CITED4 또는 CREB3으로부터 유래된 TAD를 포함할 수 있다. 이러한 eTF는 대상체에게 투여되는 경우에 면역원성을 거의 내지 전혀 갖지 않거나 또는 인간 단백질 서열에 대해 더 낮은 퍼센트 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 갖는다.

[0160] 특정 실시양태에서, SNC1A를 상향조절하는 본원에 제공된 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는다. SCN1A를 상향조절하는 본원에 제공된 eTF가 동일한 단백질로부터 유래된 DBD 및 TAD를 포함하는 경우에, 인간 단백질에 대한 퍼센트 동일성은 그것이 유래된 인간 단백질 (예를 들어, EGR1 또는 EGR3)과 매칭되는 eTF 내의 아미노산 잔기의 총수를, eTF 내의 아미노산 잔기의 총수로 나눔으로써 결정될 수 있다. SCN1A를 상향조절하는 것으로 제공된 eTF가 하나의 인간 단백질로부터의 DBD 및 상이한 인간 단백질로부터 유래된 TAD를 포함하는 경우에, 인간에 대한 퍼센트 동일성은 각각의 도메인의 인간에 대한 퍼센트 동일성을 별개로 계산하고 이 둘을 함께 합산함으로써 결정될 수 있으며, 예를 들어 (i) 그것이 유래된 인간 단백질 (예를 들어, EGR1 또는 EGR3)과 매칭되는 DBD 내의 아미노산 잔기의 총수를, eTF 내의 아미노산 잔기의 총수로 나누어 계산하고; (ii) 그것이 유래된 인간 단백질 (예를 들어, CITED2, CITED4 또는 CREB3)과 매칭되는 TAD 내의 아미노산 잔기의 총수를, eTF 내의 아미노산 잔기의 총수로 나누어 계산하고; (iii) (i)과 (ii)의 총수를 합산함으로써 결정될 수 있다. 이러한 한 실시양태에서, 도메인은 하기와 같이 나누어진다: 제1 도메인은 eTF의 N-말단으로부터 제2 도메인에 대한 코딩 서열의 시작까지이고, 제2 도메인은 제2 도메인에 대한 코딩 서열의 시작으로부터 eTF의 C-말단까지이다 (예를 들어, 구성 NLS-DBD-링커-NLS-TAD를 갖는 eTF의 경우, 제1 도메인은 NLS-DBD-링커일 것이고, 제2 도메인은 NLS-TAD일 것임). SNC1A를 상향조절하는 본원에 제공된 eTF가 1종의 인간 단백질로부터의 DBD 및 1종 이상의 상이한 인간 단백질로부터 유래된 2종의 TAD를 포함하는 경우에, 인간에 대한 퍼센트 동일성은 각각의 도메인의 인간에 대한 퍼센트 동일성을 별개로 계산하고 3개 모두를 함께 합산함으로써 결정될 수 있으며, 예를 들어 (i) 그것이 유래된 인간 단백질 (예를 들어, EGR1 또는 EGR3)과 매칭되는 DBD 내의 아미노산 잔기의 총수를, eTF 내의 아미노산 잔기의 총수로 나누어 계산하고; (ii) 그것이 유래된 인간 단백질 (예를 들어, CITED2, CITED4 또는 CREB3)과 매칭되는 제1 TAD 내의 아미노산 잔기의 총수를, eTF 내의 아미노산 잔기의 총수로 나누어 계산하고; (iii) 그것이 유래된 인간 단백질 (예를 들어, CITED2, CITED4 또는 CREB3)과 매칭되는 제2 TAD 내의 아미노산 잔기의 총수를, eTF 내의 아미노산 잔기의 총수로 나누어 계산하고; (iv) (i), (ii) 및 (iii)의 총수를 합산함으로써 결정될 수 있다. 이러한 한 실시양태에서, 도메인은 하기와 같이 나누어진다: 제1 도메인은 eTF의 N-말단으로부터 제2 도메인에 대한 코딩 서열의 시작까지이고, 제2 도메인은 제2 도메인에 대한 코딩 서열의 시작으로부터 제3 도메인에 대한 코딩 서열의 시작까지이고, 제3 도메인은 제3 도메인에 대한 코딩 서열의 시작으로부터 eTF의 C-말단까지이다 (예를 들어, 구성 NLS-TAD1-링커-NLS-DBD-링커-NLS-TAD2를 갖는 eTF의 경우, 제1 도메인은 NLS-TAD1-링커일 것이고, 제2 도메인은 NLS-DBD-링커일 것이고, 제3 도메인은 NLS-TAD2일 것임). 본 섹션에 기재된 바와 같은 1종 이상의 인간 단백질에 대한 퍼센트 동일성은 NCBI 웹사이트로부터의 월드 와이드 웹 blast.ncbi.nlm.nih.gov/에서 입수가능한, NCBI로부터 입수가능한 표준 단백질 BLAST 도구 (예를 들어, 디폴트 파라미터를 갖는 blastp (단백질 -> 단백질) 알고리즘을 사용하는 blastp 스위트 정렬 도구)를 사용하여 수득된 퍼센트 동일성 결과를 사용하여 결정될 수 있다.

[0161] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 자연 발생 인간 단백질에 대한 고도의 서열 동일성으로 인해 인간 대상체에서 유해한 면역 반응을 거의 유발하지 않거나, 최소로 유발하거나 또는 전혀 유발하지 않는 이점을 갖는다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 제공된 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 퍼센트 동일성을 포함하는 eTF, 예를 들어 1종 이상의 인간 단백질에 대해 50%, 55%, 65% 또는 70% 미만의 서열 동일성을 포함하는 eTF로 관찰된 면역원성과 비교하여 감소된 면역원성, 예를 들어 면역

원성의 적어도 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 10, 15, 20, 25, 30, 35, 40, 45 또는 50배 또는 그 초과의 감소를 유발한다. 일부 경우에, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다. 낮은 또는 최소 면역원성을 갖는 유전자 요법은 개선된 환자 내성, 치료 효과를 달성하는데 필요한 감소된 투여량, 1회 투여 후 연장된 치료 효과, 필요에 따라 다수회 또는 다중 용량으로 투여되는 능력, 투여당 보다 긴 기간에 걸친 지속적인 치료 효능, 증가된 안전성 및/또는 유전자 요법의 증가된 유효성을 포함한 여러 이점을 갖는다.

[0162] 특정 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 1종 이상의 자연 발생 인간 단백질로부터 유래된 DBD 및 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD를 포함하는 임의의 자연 발생 인간 단백질로부터 유래된 DBD를 포함할 수 있다. 예시적 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 자연 발생 아연 핑거 단백질, 예컨대, 예를 들어 표 1에 열거된 구축물 5-27, 36-41 또는 44-53 중 어느 하나로부터 유래된 DBD를 포함한다. 특정 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 인간 EGR 단백질, 예컨대 EGR1, EGR2, EGR3 또는 EGR4로부터 유래된 DBD를 포함한다. 예시적 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 인간 EGR1 또는 EGR3으로부터 유래된 DBD를 포함한다. 다양한 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 DBD의 1개 이상의 아연 핑거 도메인에서 최소 아미노산 변화 (예를 들어, 1, 2, 3, 4, 5, 6, 7 또는 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 2-3, 2-4, 2-5, 2-6, 2-7, 3-4, 3-5, 36 또는 3-7개 아미노산 변화)가 이루어져 DBD의 결합 특이성을 변경시켜 관심 표적 결합 부위를 인식하는, 인간 아연 핑거 단백질로부터 유래된 DBD를 포함한다. 이러한 서열 변형은 바람직하게는 DBD의 아연 핑거 도메인의 인식 헬릭스에서 이루어지는 반면, 나머지 인간 아연 핑거 DBD 또는 단백질 (TAD 포함)은 자연 발생 인간 단백질에 대해 가능한 한 많은 서열 동일성을 보존하도록 비변형된 채로 남아있다.

[0163] 특정 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 인간 단백질로부터 유래된 DBD에 접합된 인간 단백질로부터 유래된 1개 이상의 전사 조정 도메인 (예를 들어, TAD)을 포함한다. 다양한 실시양태에서, 전사 조정 도메인은 전사를 조정할 수 있거나 (예를 들어, RNA 폴리머라제 II, 보조-활성인자 단백질 또는 보조-억제인자 단백질) 또는 eTF가 DBD를 통해 표적 부위에 결합되는 경우에 관심 유전자로부터의 유전자 발현 수준을 조정할 수 있는 1종 이상의 단백질 인자를 동원할 수 있는 도메인을 갖는 임의의 자연 발생 인간 단백질로부터 유래될 수 있다. 예시적 실시양태에서, TAD는 인간 EGR 단백질, 예를 들어 인간 EGR1, EGR2, EGR3 또는 EGR4 또는 인간 인용 단백질, 예를 들어 인간 CITED2 또는 CITED4 단백질로부터 유래된다. 예시적 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 인간 EGR1 또는 EGR3 단백질로부터의 TAD를 포함한다. 또 다른 예시적 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 인간 CITED2 또는 CITED4 단백질로부터의 TAD를 포함한다.

[0164] 한 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 인간 DBD (hDBD) 및 인간 TAD (hTAD) (예를 들어, hTAD-hDBD 또는 hDBD-hTAD)를 포함할 수 있으며, 여기서 hDBD 및 hTAD는 동일한 인간 단백질로부터 또는 인간의 상이한 단백질로부터 유래될 수 있다. 또 다른 실시양태에서, 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 hDBD 및 2개의 hTAD를 포함할 수 있으며, 여기서 hDBD 및 hTAD는 동일한 인간 단백질로부터 유래되거나, hDBD는 제1 인간 단백질로부터 유래되고 hTAD 둘 다는 제2 인간 단백질로부터 유래되거나, hDBD 및 하나의 hTAD는 제1 인간 단백질로부터 유래되고 제2 hTAD는 제2 인간 단백질로부터 유래되거나, 또는 hDBD는 제1 인간 단백질로부터 유래되고, 하나의 hTAD는 제2 인간 단백질로부터 유래되고 제2 hTAD는 제3 인간 단백질로부터 유래된다 (예를 들어, hTAD1-hDBD-hTAD1, hTAD1-hDBD-hTAD2, hTAD1-hTAD1-hDBD, hTAD1-hTAD2-hDBD, hDBD-hTAD1-hTAD1 또는 hDBD-hTAD1-hTAD2).

[0165] 예시적 실시양태에서, 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 하기 구성 중 어느 하나를 포함한다: (i) 둘 다 인간 EGR1로부터 유래된 hDBD 및 hTAD; (ii) 둘 다 인간 EGR3으로부터 유래된 hDBD 및 hTAD; (iii) 인간 EGR1로부터 유래된 hDBD 및 CITED2로부터 유래된 hTAD (예를 들어, hEGR1 DBD-hCITED2 TAD 또는 hCITED2 TAD-hEGR1 DBD); (iv) 인간 EGR1로부터 유래된 hDBD 및 인간 CITED4로부터 유래된 hTAD (예를 들어, hEGR1 DBD-hCITED4 TAD 또는 hCITED4 TAD-hEGR1 DBD); (v) 인간 EGR3으로부터 유래된 hDBD 및 CITED2로부터 유래된 hTAD (예를 들어, hEGR3 DBD-hCITED2 TAD 또는 hCITED2 TAD-hEGR3 DBD);

(vi) 인간 EGR3으로부터 유래된 hDBD 및 인간 CITED4로부터 유래된 hTAD (예를 들어, hEGR3 DBD-hCITED4 TAD 또는 hCITED4 TAD-hEGR3 DBD); (vii) 인간 EGR1로부터 유래된 hDBD 및 CITED2로부터 유래된 2개의 hTAD (예를 들어, hCITED2 TAD-hEGR1 DBD-hCITED2 TAD, hCITED2 TAD-hCITED2 TAD-hEGR1 DBD 또는 hEGR1 DBD-hCITED2 TAD-hCITED2 TAD); (viii) 인간 EGR1로부터 유래된 hDBD 및 인간 CITED4로부터 유래된 2개의 hTAD (예를 들어, hCITED4 TAD-hEGR1 DBD-hCITED4 TAD, hCITED4 TAD-hCITED4 TAD-hEGR1 DBD 또는 hEGR1 DBD-hCITED4 TAD-hCITED4 TAD); (ix) 인간 EGR3으로부터 유래된 hDBD 및 인간 CITED2로부터 유래된 2개의 hTAD (예를 들어, hCITED2 TAD-hEGR3 DBD-hCITED2 TAD, hCITED2 TAD-hCITED2 TAD-hEGR3 DBD 또는 hEGR3 DBD-hCITED2 TAD-hCITED2 TAD); (x) 인간 EGR3으로부터 유래된 hDBD 및 인간 CITED4로부터 유래된 2개의 hTAD (예를 들어, hCITED4 TAD-hEGR3 DBD-hCITED4 TAD, hCITED4 TAD-hCITED4 TAD-hEGR3 DBD 또는 hEGR3 DBD-hCITED4 TAD-hCITED4 TAD); (xi) 인간 EGR1로부터 유래된 hDBD, 인간 CITED2로부터 유래된 제1 hTAD, 인간 CITED4로부터 유래된 제2 hTAD (예를 들어, hCITED2 TAD-hEGR1 DBD-hCITED4 TAD, hCITED4 TAD-hEGR1 DBD-hCITED2 TAD, hCITED2 TAD-hCITED4 TAD-hEGR1 DBD, hCITED4 TAD-hCITED2 TAD-hEGR1 DBD, hEGR1 DBD-hCITED4 TAD-hCITED2 TAD 또는 hEGR1 DBD-hCITED2 TAD-hCITED4 TAD); 또는 (xii) 인간 EGR3으로부터 유래된 hDBD, 인간 CITED2로부터 유래된 제1 hTAD, 인간 CITED4로부터 유래된 제2 hTAD (예를 들어, hCITED2 TAD-hEGR3 DBD-hCITED4 TAD, hCITED4 TAD-hEGR3 DBD-hCITED2 TAD, hCITED2 TAD-hCITED4 TAD-hEGR3 DBD, hCITED4 TAD-hCITED2 TAD-hEGR3 DBD, hEGR3 DBD-hCITED4 TAD-hCITED2 TAD 또는 hEGR3 DBD-hCITED2 TAD-hCITED4 TAD).

[0166] 특정 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 (i) 서열식별번호: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는 서열; (ii) 서열식별번호: 92-98 중 어느 하나를 포함하는 서열; (iii) (i) 또는 (ii)의 서열 중 어느 것에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iv) (i), (ii) 또는 (iii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다. 예시적 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다.

[0167] 특정 실시양태에서, SCN1A를 상향조절하고 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF는 DBD 및 TAD 도메인 이외에도 1개 이상의 아미노산 서열 또는 도메인, 예컨대 핵 국재화 신호 또는 링커 등을 추가로 포함할 수 있다. 추가로, 1종 이상의 인간 단백질에 대해 높은 퍼센트 서열 동일성을 갖는 본원에 제공된 eTF를 코딩하는 폴리뉴클레오티드는 eTF에 대한 코딩 서열 이외에도 1개 이상의 핵산 서열, 예컨대 프로모터, 인핸서, 폴리A 테일 등을 추가로 포함할 수 있다. 이러한 실시양태에서, 추가의 아미노산 서열 및/또는 핵산 서열 중 1개 이상은 바람직하게는 인간 서열로부터 유래된 인간 서열이거나 또는 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는다.

[0168] 예시적인 SCN1A eTF

[0169] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 서열식별번호: 152-167 중 어느 하나를 포함하는 인식 헬릭스를 포함하는 1개 이상의 아연 핑거 도메인을 갖는 DNA 결합 도메인을 포함한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 또는 12개의 아연 핑거 도메인을 갖는 DNA 결합 도메인을 포함하며, 여기서 각각의 아연 핑거 도메인은 독립적으로 서열식별번호: 152-167 중 어느 하나를 포함하는 인식 헬릭스를 포함한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 6개의 아연 핑거 도메인을 갖는 DNA 결합 도메인을 포함하며, 여기서 각각의 아연 핑거 도메인은 독립적으로 서열식별번호: 152-167 중 어느 하나를 포함하는 인식 헬릭스를 포함한다. 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 9개의 아연 핑거 도메인을 갖는 DNA 결합 도메인을 포함하며, 여기서 각각의 아연 핑거 도메인은 독립적으로 서열식별번호: 152-167 중 어느 하나를 포함하는 인식 헬릭스를 포함한다. 예시적 실시양태에서, 이러한 eTF는 서열식별번호: 147을 갖는 DNA 결합 도메인을 포함하며, 여기서 각각의 X는 독립적으로 서열식별번호: 152-167 중 어느 하나로부터 선택되고, n은 6 또는 9이다.

[0170] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) RSDNLVR x REDNLHT x RSDELVR x QSGNLTE x TSGHLVR x QNSTLTE (서열식별번호: 148) (여기서 x는 1-50개 아미노산의 링커일 수 있음)를 포함하는 서열, (ii) 서열식별번호: 148에 대해 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열, 또는 (ii) (i) 또는 (ii)의 기능적 단편 중 어느 하나를 갖는 DNA 결합 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 VP64, VPR, CITED2, CITED4 또는 CREB3으로부터 선택된 1개 이상의 TAD를 추가로 포함한다. 한 실시양태에서, 이러한 eTF는 DBD의 C-말단에 접합된 VPR TAD 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED2 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단 또는 C-말단에 접합된 2개의 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 서열식별번호: 18을 갖는 표적 부위에 결합할 수 있고, SCN1A의 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0171] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT (서열식별번호: 149) (여기서 x는 1-50개 아미노산의 링커일 수 있음)를 포함하는 서열, (ii) 서열식별번호: 149에 대해 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열, 또는 (ii) (i) 또는 (ii)의 기능적 단편 중 어느 하나를 갖는 DNA 결합 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 VP64, VPR, CITED2, CITED4 또는 CREB3으로부터 선택된 1개 이상의 TAD를 추가로 포함한다. 한 실시양태에서, 이러한 eTF는 DBD의 C-말단에 접합된 VPR TAD 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED2 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단 또는 C-말단에 접합된 2개의 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 서열식별번호: 30을 갖는 표적 부위에 결합할 수 있고, SCN1A의 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0172] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT (서열식별번호: 151)를 포함하는 서열, (ii) 서열식별번호: 151에 대해 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열, 또는 (ii) (i) 또는 (ii)의 기능적 단편 중 어느 하나를 갖는 DNA 결합 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 VP64, VPR, CITED2, CITED4 또는 CREB3으로부터 선택된 1개 이상의 TAD를 추가로 포함한다. 한 실시양태에서, 이러한 eTF는 DBD의 C-말단에 접합된 VPR TAD 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED2 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단 또는 C-말단에 접합된 2개의 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 서열식별번호: 32를 갖는 표적 부위에 결합할 수 있고, SCN1A의 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0173] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR (서열식별번호: 150)을 포함하는 서열, (ii) 서열식별번호: 150에 대해 적어도 89%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 서열, 또는 (ii) (i) 또는 (ii)의 기능적 단편 중 어느 하나를 갖는 DNA 결합 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 VP64, VPR, CITED2, CITED4 또는 CREB3으로부터 선택된 1개 이상의 TAD를 추가로 포함한다. 한 실시양태에서, 이러한 eTF는 DBD의 C-말단에 접합된 VPR TAD 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED2 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 DBD의 N-말단 또는 C-말단에 접합된 2개의 CITED4 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 서열식별번호: 31을 갖는 표적 부위에 결합할 수 있고, SCN1A의 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배,

50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0174] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 99-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는 서열; (ii) 서열식별번호: 77-98 중 어느 하나를 포함하는 서열; (iii) (i) 또는 (ii)의 서열 중 어느 것에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iv) (i), (ii) 또는 (iii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0175] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 99-102 또는 125-127 중 어느 하나를 포함하는 서열; (ii) 서열식별번호: 77-91 중 어느 하나를 포함하는 서열; (iii) (i) 또는 (ii)의 서열 중 어느 것에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iv) (i), (ii) 또는 (iii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0176] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는 서열; (ii) 서열식별번호: 92-98 중 어느 하나를 포함하는 서열; (iii) (i) 또는 (ii)의 서열 중 어느 것에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iv) (i), (ii) 또는 (iii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다. 예시적 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다.

[0177] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 127을 포함하는 서열; (ii) 서열식별번호: 127에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 서열식별번호: 77을 포함하고, 서열식별번호: 18을 갖는 표적 부위에 결합한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다.

[0178] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 128을 포함하는 서열; (ii) 서열식별번호: 128에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 서열식별번호: 92를 포함하고, 서열식별번호: 18을 갖는 표적 부위에 결합한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다. 예시적 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해

더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다.

[0179] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 129를 포함하는 서열; (ii) 서열식별번호: 129에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 서열식별번호: 92를 포함하고, 서열식별번호: 18을 갖는 표적 부위에 결합한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다. 예시적 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다.

[0180] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 130을 포함하는 서열; (ii) 서열식별번호: 130에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 서열식별번호: 92를 포함하고, 서열식별번호: 18을 갖는 표적 부위에 결합한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다. 예시적 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다.

[0181] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 (i) 서열식별번호: 131을 포함하는 서열; (ii) 서열식별번호: 131에 대해 적어도 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 포함하는 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 것의 기능적 단편 또는 변이체 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 eTF는 서열식별번호: 92를 포함하고, 서열식별번호: 18을 갖는 표적 부위에 결합한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있거나 또는 대조군과 비교하여 적어도 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% 또는 500% 또는 그 초과만큼 상향조절할 수 있다. 예시적 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 전체 서열 동일성을 갖는다. 특정 실시양태에서, 이러한 eTF는 1종 이상의 인간 단백질에 대해 더 낮은 전체 퍼센트 서열 동일성을 갖는 eTF와 비교하여 감소된 면역원성을 나타낸다. 다양한 실시양태에서, 면역원성의 감소는 엘리스팟 검정, 면역검정 또는 인 실리코 방법을 사용하여 측정될 수 있다.

[0182] 특정 실시양태에서, SCN1A를 상향조절하는 본원에 개시된 eTF는 gRNA/Cas 복합체를 포함하는 DBD를 포함하며, 여기서 gRNA는 서열식별번호: 35-66 중 어느 하나를 포함하는 표적화 서열을 포함한다. gRNA의 표적 서열은 서열: 5'-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA-3' (서열식별번호: 183)를 갖는 스캐폴드 서열의 5' 말단에 부착된다. 예시적 실시양태에서, Cas 단백질은 뉴클레아제 비활성화된 Cas9 단백질이다. 특정 실시양태에서, 이러한 eTF는 Cas 단백질에 접합된 1개 이상의 TAD를 추가로 포함하며, 여기서 TAD는 VP64, VPR, CITED2, CITED4 또는 CREB3으로부터 선택된다. 한 실시양태에서, 이러한 eTF는 Cas 단백질의 C-말단에 접합된 VPR TAD 도메인을 포함한다. 특정 실시양태에서, 이러한 eTF는 Cas 단백질의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED2 TAD를 포함한다. 특정 실시양태에서, 이러한 eTF는 Cas 단백질의 N-말단, C-말단 또는 N-말단 및 C-말단에 접합된 CITED4 TAD를 포함한다. 예시적 실시양태에서, 이러한 eTF는 SCN1A 발현을 대조군과 비교하여 적어도 2배, 5배, 10배, 15배, 20배, 25배, 30배, 40배, 50배 또는 그 초과만큼 상향조절할 수 있다.

[0183] 폴리뉴클레오티드

[0184] 또 다른 측면에서, 본 출원은 본원에 개시된 SNC1A를 상향조절하는 eTF 중 어느 것을 코딩하는 폴리뉴클레오티드를 제공한다. 또 다른 측면에서, 본 출원은 PV 선택적 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 특정 실시양태에서, 본 출원은 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 특정 실시양태에서, 본 출원은 본원에 개시된 바와 같은 SCN1A를 상향조절하는 eTF를 코딩하는 서열 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 특정 실시양태에서, 본 출원은 SCN1A를 상향조절하는 eTF를 코딩하는 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소 및 PV 선택적 조절 요소를 제공한다.

[0185] SCN1A를 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드

[0186] 특정 실시양태에서, 본 출원은 하기: (i) 서열식별번호: 77-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는, SCN1A를 상향조절하는 eTF 또는 그의 변이체 또는 기능적 단편을 코딩하는 핵산 서열; (ii) 서열식별번호: 77-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 갖는, SCN1A를 상향조절하는 eTF의 기능적 단편을 코딩하는 핵산; 또는 (iii) 서열식별번호: 77-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 갖는, SCN1A를 상향조절하는 eTF 또는 그의 변이체 또는 기능적 단편에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는, SCN1A를 상향조절하는 eTF를 코딩하는 핵산 중 어느 하나를 포함하는 폴리뉴클레오티드를 제공한다.

[0187] 특정 실시양태에서, 본 출원은 하기: (i) 서열식별번호: 92-98 중 어느 하나를 포함하는 DBD 또는 그의 변이체 또는 기능적 단편을 코딩하는 핵산 서열; (ii) 서열식별번호: 92-98 중 어느 하나를 갖는 DBD의 기능적 단편을 코딩하는 핵산; 또는 (iii) 서열식별번호: 92-98 중 어느 하나를 갖는 DBD 또는 그의 변이체 또는 기능적 단편에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 DBD를 코딩하는 핵산 중 어느 하나를 포함하며, 여기서 DBD는 서열식별번호: 92-98 중 어느 하나에 의해 결합된 표적 부위에 결합할 수 있는 것인 폴리뉴클레오티드를 제공한다.

[0188] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는 eTF를 코딩하는 핵산 서열; (ii) 서열식별번호: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 갖는 eTF의 기능적 단편을 코딩하는 핵산; 또는 (iii) 서열식별번호: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 갖는 eTF에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 eTF를 코딩하는 핵산 중 어느 하나를 포함하고, 여기서 eTF는 SCN1A를 상향조절할 수 있다.

[0189] 특정 실시양태에서, 본 출원은 본원에 개시된 eTF에 의해 결합되는 경우에 내인성 SCN1A를 상향조절할 수 있는 게놈 표적 부위에 결합하는 DBD를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 77-98 중 어느 하나를 포함하는 DBD를 코딩하는 핵산 서열; (ii) 서열식별번호: 77-98 중 어느 하나를 갖는 DBD의 기능적 단편을 코딩하는 핵산; 또는 (iii) 서열식별번호: 77-98 중 어느 하나를 갖는 DBD에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 eTF를 코딩하는 핵산 중 어느 하나를 포함하고, 여기서 DBD는 서열식별번호: 77-98 중 어느 하나에 의해 결합되는 표적 부위에 결합할 수 있다.

[0190] 특정 실시양태에서, 본 출원은 본원에 개시된 eTF에 의해 결합되는 경우에 내인성 SCN1A를 상향조절할 수 있는 게놈 표적 부위에 결합하는 DBD를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 148-151 중 어느 하나를 포함하는 DBD를 코딩하는 핵산 서열; (ii) 서열식별번호: 148-151 중 어느 하나를 갖는 DBD의 기능적 단편을 코딩하는 핵산; 또는 (iii) 서열식별번호: 148-151 중 어느 하나를 갖는 DBD에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 eTF를 코딩하는 핵산 중 어느 하나를 포함하고, 여기서 DBD는 서열식별번호: 92-98 중 어느 하나에 의해 결합된 표적 부위에 결합할 수 있다.

[0191] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 70-76 또는 184 중 어느 것을 갖는 핵산 서열; (ii) (i)의

서열 중 어느 하나의 기능적 단편을 갖는 핵산; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함하고, 여기서 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 eTF를 코딩한다.

[0192] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 70을 갖는 핵산 서열; (ii) 서열식별번호: 70의 기능적 단편을 갖는 핵산; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 127을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0193] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 71을 갖는 핵산 서열; (ii) 서열식별번호: 71의 기능적 단편을 갖는 핵산; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 127을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0194] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 72를 갖는 핵산 서열; (ii) 서열식별번호: 72의 기능적 단편을 갖는 핵산; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 130을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0195] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 73을 갖는 핵산 서열; (ii) 서열식별번호: 73의 기능적 단편을 갖는 핵산 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 131을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0196] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 74를 갖는 핵산 서열; (ii) 서열식별번호: 74의 기능적 단편을 갖는 핵산 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 127을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0197] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 75를 갖는 핵산 서열; (ii) 서열식별번호: 75의 기능적 단편을 갖는 핵산 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 127을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0198] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며, 여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 76을 갖는 핵산 서열; (ii) 서열식별번호: 76의 기능적 단편을 갖는 핵산 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 106을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0199] 특정 실시양태에서, 본 출원은 내인성 SCN1A를 조절할 수 있는 eTF를 코딩하는 폴리뉴클레오티드를 제공하며,

여기서 폴리뉴클레오티드는 하기: (i) 서열식별번호: 184를 갖는 핵산 서열; (ii) 서열식별번호: 184의 기능적 단편을 갖는 핵산 서열; 또는 (iii) (i) 또는 (ii)의 서열 중 어느 하나에 대해 적어도 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과의 서열 동일성을 갖는 핵산 중 어느 하나를 포함한다. 예시적 실시양태에서, 이러한 폴리뉴클레오티드는 SCN1A를 상향조절할 수 있는 서열식별번호: 106을 갖는 eTF 또는 그의 기능적 단편 또는 변이체를 코딩한다.

[0200] PV 뉴런에서의 선택적 발현을 위한 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드

[0201] 또 다른 측면에서, 본 출원은 파르브알부민 (PV) 뉴런에서 관심 유전자의 선택적 발현을 유도하는 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 마이크로RNA 또는 miRNA는 mRNA 분자 내의 상보적 인식 부위에 혼성화함으로써 전사후 유전자 발현을 조절하고, mRNA 전사체의 분해를 촉진함으로써 또는 mRNA에 의해 코딩되는 단백질의 번역을 억제함으로써 유전자 발현의 억제를 유도하는 작은 비-코딩 RNA (~20개 뉴클레오티드)이다. 본원에 제공된 마이크로RNA 결합 부위는 흥분성 뉴런에서 관심 유전자의 발현을 억제함으로써 PV 뉴런에서 관심 유전자의 선택적 발현을 촉진한다 (예를 들어, PV 선택적 마이크로RNA 결합 부위). 특정 실시양태에서, 흥분성 뉴런은 STAC, Slc17a7, Car12, Syt17, ITPKA, Col6a1, CamKII, Sv2b, INHBA 및/또는 DKK3 중 1종 이상을 발현하는 뉴런이다. 예시적 실시양태에서, 흥분성 뉴런은 CamKII를 발현하는 뉴런이다.

[0202] 특정 실시양태에서, 본 출원은 PV 선택적 발현을 촉진하는, 예를 들어 흥분성 뉴런에서 마이크로RNA 결합 부위를 포함하는 mRNA의 분해를 촉진하는 1개 이상의 마이크로RNA에 대한 1개 이상의 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. PV 선택적 발현을 촉진하는 예시적인 마이크로RNA는, 예를 들어 miR-128, miR-221 및 miR-222를 포함한다. 특정 실시양태에서, 본 출원은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 초과의 PV 선택적 마이크로RNA 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 초과의 miR-128 결합 부위 (서열식별번호: 9)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 초과의 miR-221 결합 부위 (서열식별번호: 11)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 초과의 miR-222 결합 부위 (서열식별번호: 13)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 1개의 miR-128 결합 부위, 적어도 1개의 miR-221 결합 부위 및 적어도 1개의 miR-222 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 1개의 miR-128 결합 부위 및 적어도 1개의 miR-222 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 1개의 miR-221 결합 부위 및 적어도 1개의 miR-222 결합 부위를 포함하는 폴리뉴클레오티드를 제공한다. 예시적 실시양태에서, 본 출원은 적어도 1개의 miR-128 결합 부위 (서열식별번호: 9) 및 적어도 1개의 miR-221 결합 부위 (서열식별번호: 11)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 2개의 miR-128 결합 부위 (서열식별번호: 9) 및 적어도 2개의 miR-221 결합 부위 (서열식별번호: 11)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 3개의 miR-128 결합 부위 (서열식별번호: 9) 및 적어도 3개의 miR-221 결합 부위 (서열식별번호: 11)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 4개의 miR-128 결합 부위 (서열식별번호: 9) 및 적어도 4개의 miR-221 결합 부위 (서열식별번호: 11)를 포함하는 폴리뉴클레오티드를 제공한다. 한 실시양태에서, 본 출원은 적어도 5개의 miR-128 결합 부위 (서열식별번호: 9) 및 적어도 5개의 miR-221 결합 부위 (서열식별번호: 11)를 포함하는 폴리뉴클레오티드를 제공한다. 이러한 실시양태에서, 결합 부위는 임의의 순서로 배열될 수 있다. 예를 들어, 2개의 miR-128 결합 부위 및 2개의 miR-221 결합 부위를 함유하는 구축물의 경우, 결합 부위는 하기 구성 중 어느 것으로 배열될 수 있다: miR-128 - miR-128 - miR-221 - miR221, miR-128 - miR-221 - miR-128 - miR-221, miR-128 - miR221 - miR221 - miR-128, miR-221 - miR128 - miR221 - miR128, miR-221 - miR128 - miR128 - miR221, or miR221 - miR221- miR128 - miR128. 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 4개의 miR-128 결합 부위 (서열식별번호: 9)에 이어서 4개의 miR-221 결합 부위 (서열식별번호: 11)를 갖는 서열, 예를 들어 miR-128 - miR-128 - miR128 - miR-128 - miR221 - miR221 - miR-221 - miR221을 포함한다. 또 다른 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 1개의 miR-221 서열 (서열식별번호: 11), 1개의 miR-222 서열 (서열식별번호: 13) 및 1개의 miR-128 결합 부위 (서열식별번호: 9)를 갖는 서열, 예를 들어 miR-221 - miR222 - miR128을 포함한다. 또 다른 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 하기 순서로 배열된 2개의 miR-221 서열 (서열식별번호: 11), 2개의 miR-222 서열 (서열식별번호: 13) 및 2개의 miR-128 결합 부위 (서열식별번호: 9)를 갖는 서열을 포함한다: miR-221 - miR222 - miR128 - miR-221 - miR222 - miR128.

[0203] 1개 초과의 마이크로RNA 결합 부위를 갖는 폴리뉴클레오티드에서, 결합 부위들은 폴리뉴클레오티드 서열에서 서

로 바로 인접할 수 있거나 (예를 들어, 결합 부위들 사이에 링커 또는 개재 서열이 없음), 또는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개 또는 그 초과의 뉴클레오티드 또는 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3 또는 1-2개의 뉴클레오티드만큼 서로 분리되어 있을 수 있다. 예시적 실시양태에서, 마이크로RNA 결합 부위는 약 5개의 뉴클레오티드만큼 또는 5개의 뉴클레오티드만큼 분리되어 있다. 예시적 실시양태에서, 마이크로RNA 결합 부위들을 분리하는 서열 (뿐만 아니라 마이크로RNA 결합 부위들 사이에 형성된 접합부 또는 마이크로RNA 결합 부위와 마이크로RNA 결합 부위들을 분리하는 서열 사이에 형성된 접합부)은 임의의 다른 마이크로RNA 또는 임의의 다른 뉴런 마이크로RNA에 상보적이지 않다.

[0204] 특정 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 7에 대해 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성을 갖는 마이크로RNA 결합 부위를 포함한다. 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 7을 포함하는 마이크로RNA 결합 부위를 포함한다.

[0205] 특정 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 14에 대해 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성을 갖는 마이크로RNA 결합 부위를 포함한다. 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 14를 포함하는 마이크로RNA 결합 부위를 포함한다.

[0206] 특정 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 15에 대해 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일성을 갖는 마이크로RNA 결합 부위를 포함한다. 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 15를 포함하는 마이크로RNA 결합 부위를 포함한다.

[0207] 특정 실시양태에서, 본원에 제공된 마이크로RNA 결합 부위는 mRNA 전사체의 3' 비번역 영역 내에, 예를 들어 번역 종결 코돈 (즉, TAA, TGA 또는 TAG) 다음에 및 폴리A 테일 앞에 위치한다. 마이크로RNA 결합 부위는 번역 종결 코돈에 바로 인접하여 위치할 수 있거나 또는 번역 종결 코돈으로부터 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개 또는 그 초과의 뉴클레오티드 또는 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3 또는 1-2개의 뉴클레오티드만큼 분리되어 있을 수 있고/거나, 폴리A 테일에 인접하여 위치할 수 있거나 또는 폴리A 테일로부터 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개 또는 그 초과의 뉴클레오티드 또는 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3 또는 1-2개의 뉴클레오티드만큼 분리되어 있을 수 있다.

[0208] 특정 실시양태에서, 본원에 제공된 마이크로RNA 결합 부위는 오프-타겟 세포 유형과 비교하여 PV 세포에서 선택적 유전자 발현을 일으킨다. 일부 경우에, 오프-타겟 세포 유형은 흥분성 뉴런, 비-PV CNS 세포-유형 및 비-뉴런 CNS 세포 유형을 포함하나 이에 제한되지는 않는다. 특정 실시양태에서, PV 선택적 마이크로RNA 결합 부위는 적어도 1, 2, 3, 4, 5종 또는 그 초과의 비-PV CNS 세포 유형에 비해 PV 뉴런에서 선택적 유전자 발현을 일으킨다. 일부 경우에, 비-PV CNS 세포는 흥분성 뉴런, 도파민성 뉴런, 성상세포, 소교세포, 운동 뉴런, 혈관 세포 또는 비-GABA성 뉴런 (예를 들어, GAD2, GAD1, NKX2.1, DLX1, DLX5, SST 및 VIP 중 1종 이상을 발현하지 않는 세포), 비-PV 뉴런 (예를 들어, 파르브알부민을 발현하지 않는 GABA성 뉴런), 또는 다른 CNS 세포 (예를 들어, PV, GAD2, GAD1, NKX2.1, DLX1, DLX5, SST 및 VIP 중 어떤 것도 발현하지 않는 CNS 세포 유형)이다. 예시적 실시양태에서, 본원에 제공된 PV 선택적 마이크로RNA 결합 부위는 흥분성 뉴런 (예를 들어, STAC, Slc17a7, Car12, Syt17, ITPKA, Col6a1, CamKII, Sv2b, INHBA 및/또는 DKK3 중 1종 이상을 발현하는 뉴런)에서의 발현을 감소시킴으로써 이러한 흥분성 뉴런과 비교하여 PV 뉴런에서 유전자 발현의 선택성 증가를 일으킨다. 일부 경우에, 세포 유형은 상이한 세포 마커, 형태, 표현형, 유전자형, 기능 및/또는 세포 유형을 분류하기 위한 임의의 다른 수단을 가짐으로써 구별된다.

[0209] PV 선택적 마이크로RNA 결합 부위에 의해 유도된 발현의 선택성은 다수의 방식으로 측정될 수 있다. 한 실시양태에서, 비-PV 세포 대비 PV 세포에서의 유전자 발현의 선택성은 PV 선택적 마이크로RNA 결합 부위를 함유하는 유전자로부터 검출가능한 수준의 전사체를 발현하는 PV 세포의 수를 유전자를 발현하는 세포의 총수와 비교함으로써 (예를 들어, 유전자를 발현하는 총 세포 (PV + 비-PV 세포) 대비 PV 세포의 비) 측정될 수 있다. 예를 들어, PV 뉴런에 대한 선택성은 형광 단백질 (예를 들어, eGFP)을 코딩하는 유전자 및 유전자 발현을 측정하기 위한 PV 선택적 마이크로RNA 결합 부위 및 제2 형광 표지 (예를 들어, 적색 형광 단백질)에 연결된 PV 세포를 확인하는 항체 (예를 들어, PV 뉴런과 특이적으로 상호작용하는 항-PV 항체)를 포함하는 발현 카세트를 사용한 면역조직화학 기반 공동국재화 검정을 사용하여 결정될 수 있다. PV 세포에서의 발현의 선택성은 PV 및 eGFP를

둘 다 발현하는 세포 (예를 들어, PV 세포)의 수를 eGFP를 발현하는 세포 (예를 들어, PV 세포 및 비-PV 세포)의 총수로 나누고 100을 곱하여 백분율로 전환시킴으로써 계산될 수 있다. 또 다른 예에서, PV 뉴런에 대한 선택성은 형광 단백질 (예를 들어, eGFP)을 코딩하는 유전자 및 유전자 발현을 측정하기 위한 PV 선택적 마이크로RNA 결합 부위 및 제2 형광 표지 (예를 들어, 적색 형광 단백질)에 연결된 PV 세포를 확인하는 제1 항체 (예를 들어, PV 뉴런과 특이적으로 상호작용하는 항-PV 항체) 및 흥분성 세포를 확인하는 제2 항체 (예를 들어, 흥분성 뉴런과 특이적으로 상호작용하는 항-CamKII 항체)를 포함하는 발현 카세트를 사용한 면역조직화학 기반 공동국재화 검정을 사용하여 결정될 수 있다. PV 세포에서의 발현의 선택성은 PV 및 eGFP를 둘 다 발현하는 세포 (예를 들어, PV 세포)의 수를 eGFP + PV 및 eGFP + CamKII를 발현하는 세포 (예를 들어, PV 세포 및 흥분성 세포)의 수로 나누고 100을 곱하여 백분율로 전환시킴으로써 계산될 수 있다. 트랜스진을 발현하는 PV 세포의 백분율이 높을수록, 마이크로RNA 결합 부위가 PV 세포에 대해 더 선택적이다. 특정 실시양태에서, 본원에 제공된 PV 선택적 마이크로RNA 결합 부위는 PV 세포에서의 발현에 대해 고도로 선택적일 수 있다. 예를 들어, 본원에 제공된 PV 선택적 마이크로RNA 결합 부위는 PV 뉴런에 대해 약 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 약 99% 초과의 선택성 (예를 들어, PV 뉴런/총 세포 x 100 또는 PV 뉴런/PV + 흥분성 뉴런 x 100)을 나타낼 수 있다.

[0210] 일부 경우에, 본원에 제공된 PV 선택적 마이크로RNA 결합 부위는 짧다. 일부 경우에, PV 선택적 마이크로RNA 결합 부위의 크기는 벡터, 예를 들어 바이러스 벡터 또는 rAAV의 클로닝 능력에 적합하여, 트랜스진, 프로모터 (및 임의적인 인핸서) 및 마이크로RNA 결합 부위의 합한 크기가 벡터의 클로닝 능력을 초과하지 않는다. 일부 경우에, PV 선택적 마이크로RNA 결합 부위는 약 500 bp, 400 bp, 300 bp, 250 bp, 225 bp, 215 bp, 210 bp, 200 bp, 150 bp, 140 bp, 135 bp, 130 bp, 125 bp, 120 bp, 115 bp, 110 bp, 100 bp, 90 bp, 80 bp, 75 bp, 70 bp, 65 bp, 60 bp 또는 50 bp 이하의 길이를 갖는다. 일부 경우에, PV 선택적 마이크로RNA 결합 부위는 약 50-500 bp, 50-400 bp, 50-300 bp, 50-250 bp, 50-200 bp, 50-100 bp, 50-75 bp, 50-70 bp, 100-500 bp, 100-400 bp, 100-300 bp, 100-250 bp, 100-200 bp, 100-150 bp, 100-140 bp, 100-135 bp, 200-500 bp, 200-400 bp, 200-300 bp 또는 200-250 bp이다.

[0211] 예시적 실시양태에서, 1개 이상의 PV 선택적 마이크로RNA 결합 부위를 포함하는 본원에 제공된 폴리뉴클레오티드는 서열식별번호: 67을 포함하지 않는다.

[0212] 발현 카세트

[0213] 또 다른 측면에서, 본 출원은 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드) 및 1개 이상의 조절 요소를 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드) 및 PV 선택적 프로모터를 포함하는 발현 카세트를 제공한다.

[0214] 특정 실시양태에서, 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드)는 eTF를 코딩하는 서열에 추가로 1개 이상의 조절 요소를 포함하는 발현 카세트의 일부이다. 예시적 실시양태에서, 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드)는 세포에서 트랜스진 (예를 들어, SCN1A를 선택적으로 상향조절하는 eTF)의 발현을 구동할 수 있도록 트랜스진 서열의 상류에 위치한 프로모터를 포함하는 발현 카세트의 일부이다.

[0215] 특정 실시양태에서, 본원에 개시된 발현 카세트는 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드) 및 세포에서 트랜스진 (예를 들어, SCN1A를 선택적으로 상향조절하는 eTF)의 발현을 구동할 수 있도록 트랜스진을 코딩하는 서열의 상류에 위치한 구성적 프로모터를 포함한다. 구성적 프로모터의 예는 GAD2 프로모터, 인간 시냅신 프로모터, CBA 프로모터, CMV 프로모터, minCMV 프로모터, TATA 박스, 슈퍼 코어 프로모터 또는 EF1α 프로모터 또는 그의 조합을 포함한다.

[0216] 특정 실시양태에서, 본원에 개시된 발현 카세트는 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드) 및 세포에서 트랜스진 (예를 들어, SCN1A를 선택적으로 상향조절하는 eTF)의 발현을 구동할 수 있는 짧은 프로모터를 포함한다. 특정 실시양태에서, 본원에 기재된 핵산 분자에 따라 사용하기에 적합한 짧은 프로모터는 500 bp, 450 bp, 400 bp, 350 bp, 300 bp, 250 bp, 225 bp, 200 bp, 175 bp, 150 bp, 145 bp, 140 bp, 135

bp, 130 bp, 125 bp, 120 bp, 115 bp, 110 bp, 105 bp, 100 bp, 95 bp, 90 bp, 85 bp, 80 bp 또는 75 bp 미만, 또는 약 80-300 bp, 80-275 bp, 80-250 bp, 80-200 bp, 80-150 bp, 80-125 bp, 80-120 bp, 80-115 bp, 80-110 bp, 80-105 bp, 80-100 bp, 85-300 bp, 85-275 bp, 85-250 bp, 85-200 bp, 85-150 bp, 85-125 bp, 85-120 bp, 85-115 bp, 85-110 bp, 85-105 bp, 85-100 bp, 90-300 bp, 90-275 bp, 90-250 bp, 90-200 bp, 90-150 bp, 90-125 bp, 90-120 bp, 90-115 bp, 90-110 bp, 90-105 bp, 90-100 bp, 95-300 bp, 95-275 bp, 95-250 bp, 95-200 bp, 95-150 bp, 95-125 bp, 95-120 bp, 95-115 bp, 95-110 bp, 95-105 bp, 95-100 bp, 100-300 bp, 100-275 bp, 100-250 bp, 100-200 bp, 100-150 bp, 100-125 bp, 100-120 bp, 100-115 bp, 100-110 bp 또는 100-105 bp를 포함한다. 예시적 실시양태에서, 본원에 기재된 발현 카세트에 따라 사용하기에 적합한 짧은 프로모터는 약 100-120 bp, 약 117 bp 또는 약 100 bp를 포함한다.

[0217] 특정 실시양태에서, 본원에 개시된 발현 카세트는, 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF 중 어느 하나를 코딩하고 본원에 개시된 바와 같은 마이크로RNA 결합 부위를 임의로 함유하는 폴리뉴클레오티드에 작동가능하게 연결된 (i) 서열식별번호: 1; (ii) 그의 변이체 또는 기능적 단편; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열 중 어느 하나를 포함하거나 그로 이루어진 짧은 프로모터를 포함한다. 짧은 프로모터 서열의 다른 예는 PCT 공개 번호 WO 2018/213786에서 찾아볼 수 있다.

[0218] 특정 실시양태에서, 본원에 개시된 발현 카세트는 본원에 제공된 폴리뉴클레오티드 (예를 들어, SCN1A를 상향조절하고/거나 PV 선택적 마이크로RNA 결합 부위를 함유하는 eTF를 코딩하는 서열을 포함하는 폴리뉴클레오티드) 및 관심 세포에서 트랜스진 (예를 들어, SCN1A를 선택적으로 상향조절하는 eTF)의 발현을 선택적으로 구동할 수 있도록 이러한 트랜스진을 코딩하는 서열의 상류에 위치한 세포 유형 선택적 프로모터를 포함한다. 특정 실시양태에서, 세포 유형 선택적 프로모터는 임의의 관심 세포 유형, 예컨대, 예를 들어 심장 세포, 간 세포, 근육 세포, 골 세포, 뉴런 또는 그의 하위 집단에 대해 선택적일 수 있다 (예를 들어, 그에서의 발현을 선택적으로 구동할 수 있음). 예시적 실시양태에서, 본원에 개시된 발현 카세트는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드, 및 PV 세포에서 eTF의 발현을 선택적으로 구동할 수 있도록 eTF를 코딩하는 서열의 상류에 위치한 PV 선택적 조절 요소 (예를 들어, 프로모터, 인핸서 및/또는 프로모터 및 인핸서), 및 임의로 PV 선택적 마이크로RNA 결합 부위를 포함한다. PV 선택적 조절 요소는 PV 뉴런에서의 유전자 발현을 특이적으로 조정하는 조절 요소를 지칭한다. 특정 실시양태에서, PV 선택적 조절 요소는 1종 이상의 다른 CNS 세포 유형에 비해 PV 뉴런에서 발현을 증진시킨다. 특정 실시양태에서, PV 선택적 조절 요소는 오프-타겟 세포 유형에서 전사 및/또는 번역 프로세스를 억제한다.

[0219] 특정 실시양태에서, 본원에 제공된 PV 선택적 조절 요소는 오프-타겟 세포 유형과 비교하여 PV 세포에서 선택적 유전자 발현을 일으킨다. 일부 경우에, 오프-타겟 세포 유형은 흥분성 뉴런, 비-PV CNS 세포-유형 및 비-뉴런 CNS 세포 유형을 포함하나 이에 제한되지는 않는다. 특정 실시양태에서, PV 선택적 조절 요소는 적어도 1, 2, 3, 4, 5종 또는 그 초과의 비-PV CNS 세포 유형에 비해 PV 뉴런에서 선택적 유전자 발현을 일으킨다. 일부 경우에, 비-PV CNS 세포는 흥분성 뉴런, 도파민성 뉴런, 성상세포, 소교세포, 운동 뉴런, 혈관 세포 또는 비-GABA성 뉴런 (예를 들어, GAD2, GAD1, NKX2.1, DLX1, DLX5, SST 및 VIP 중 1종 이상을 발현하지 않는 세포), 비-PV 뉴런 (예를 들어, 파르브알부민을 발현하지 않는 GABA성 뉴런), 또는 다른 CNS 세포 (예를 들어, PV, GAD2, GAD1, NKX2.1, DLX1, DLX5, SST 및 VIP 중 어떤 것도 발현하지 않는 CNS 세포 유형)이다. 일부 경우에, 본원에 제공된 PV 선택적 조절 요소는 비-PV GABA성 세포와 비교하여 PV 뉴런에서 유전자 발현의 선택성 증가를 일으킨다. 일부 경우에, 세포 유형은 상이한 세포 마커, 형태, 표현형, 유전자형, 기능 및/또는 세포 유형을 분류하기 위한 임의의 다른 수단을 가짐으로써 구별된다.

[0220] PV 선택적 조절 요소에 의해 구동되는 발현의 선택성은 다수의 방식으로 측정될 수 있다. 한 실시양태에서, 비-PV 세포 대비 PV 세포에서의 유전자 발현의 선택성은 PV 선택적 조절 요소에 작동가능하게 연결된 유전자로부터 검출가능한 수준의 전사체를 발현하는 PV 세포의 수를 유전자를 발현하는 세포의 총수와 비교함으로써 (예를 들어, 유전자를 발현하는 총 세포 (PV + 비-PV 세포) 대비 PV 세포의 비) 측정될 수 있다. 예를 들어, PV 뉴런에 대한 선택성은 유전자 발현을 측정하기 위한 PV 선택적 조절 요소에 작동가능하게 연결된 형광 단백질 (예를 들어, eGFP)을 코딩하는 유전자 및 제2 형광 표지 (예를 들어, 적색 형광 단백질)에 연결된 PV 세포를 확인하는 항체 (예를 들어, PV 뉴런과 특이적으로 상호작용하는 항-PV 항체)를 사용한 면역조직화학 기반 공동국재화 검정을 사용하여 결정될 수 있다. PV 세포에서의 발현의 선택성은 PV 및 eGFP를 둘 다 발현하는 세포 (예를 들어, PV 세포)의 수를 eGFP를 발현하는 세포 (예를 들어, PV 세포 및 비-PV 세포)의 총수로 나누고 100을 곱하여 백분율로 전환시킴으로써 계산될 수 있다. 트랜스진을 발현하는 PV 세포의 백분율이 높을수록, 조절 요소가

PV 세포에 대해 더 선택적이다. 특정 실시양태에서, 본원에 제공된 PV 선택적 조절 요소는 PV 세포에서의 발현에 대해 고도로 선택적일 수 있다. 예를 들어, 본원에 제공된 PV 선택적 조절 요소는 PV 뉴런에 대해 약 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 약 99% 초과의 선택성 (예를 들어, PV 뉴런/총 세포 x 100)을 나타낼 수 있다.

[0221] 일부 경우에, 본원에 제공된 PV 선택적 조절 요소는 짧다. 일부 경우에, PV 선택적 조절 요소의 크기는 벡터, 예를 들어 바이러스 벡터 또는 rAAV의 클로닝 능력에 적합하여, 트랜스진 및 1개 이상의 PV 선택적 조절 요소의 합한 크기가 벡터의 클로닝 능력을 초과하지 않는다. 일부 경우에, PV 선택적 조절 요소는 약 2050bp, 2000bp, 1900bp, 1800bp, 1700bp, 1600bp, 1500bp, 1400bp, 1300bp, 1200bp, 1100bp, 1000bp, 900bp, 800bp, 700bp, 600bp, 500bp, 400bp, 300bp, 200bp 또는 100bp 이하의 길이를 갖는다. 일부 경우에, PV 선택적 조절 요소는 약 500-600bp, 500-700bp, 500-800bp, 500-900bp, 500-1000bp, 500-1500bp, 500-2000bp 또는 500-2050bp이다.

[0222] 특정 실시양태에서, 본원에 제공된 PV 선택적 조절 요소는 (i) 서열식별번호: 2-4; (ii) 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열 중 어느 하나를 포함하거나 그로 이루어진다. 일부 경우에, 조절 요소는 서열식별번호: 2-4 중 어느 하나를 포함한다. PV 선택적 조절 요소의 다른 예는 PCT 공개 번호 WO 2018/187363에서 찾아볼 수 있다.

[0223] 예시적 실시양태에서, 본 출원은 PV 선택적 조절 요소의 제어 하에 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 핵산 서열을 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 서열식별번호: 2-4 중 어느 하나를 갖는 PV 선택적 조절 요소의 제어 하에 하기 서열: 서열식별번호: 77-98 중 어느 하나를 갖는 DBD를 포함하는, SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 핵산 서열을 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 서열식별번호: 2-4 중 어느 하나를 갖는 PV 선택적 조절 요소의 제어 하에 하기 서열: 서열식별번호: 99-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는, SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 핵산 서열을 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 서열식별번호: 2-4 중 어느 하나를 갖는 PV 선택적 조절 요소의 제어 하에 하기 서열: 서열식별번호: 67-73 중 어느 하나를 포함하는 핵산 서열을 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 서열식별번호: 2 중 어느 하나를 갖는 PV 선택적 조절 요소의 제어 하에 하기 서열: 서열식별번호: 148-151 중 어느 하나를 갖는 DBD를 포함하는, SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 핵산 서열을 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 서열식별번호: 2 중 어느 하나를 갖는 PV 선택적 조절 요소의 제어 하에 하기 서열: 서열식별번호: 99-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나를 포함하는, SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 핵산 서열을 포함하는 발현 카세트를 제공한다. 특정 실시양태에서, 본 출원은 서열식별번호: 2 중 어느 하나를 갖는 PV 선택적 조절 요소의 제어 하에 하기 서열: 서열식별번호: 67-76 또는 316 중 어느 하나를 포함하는 핵산 서열을 포함하는 발현 카세트를 제공한다.

[0224] 특정 실시양태에서, 본 출원은 PV 선택적 마이크로RNA 결합 부위 및 프로모터 및/또는 인핸서 서열을 포함하는 발현 카세트를 제공한다. 본원에 기재된 임의의 프로모터가 발현 카세트에 포함될 수 있다. 예시적 실시양태에서, 본원에 제공된 발현 카세트는 PV 선택적 마이크로RNA 결합 부위 및 PV 선택적 조절 요소를 포함한다. 특정 실시양태에서, 본원에 제공된 발현 카세트는 PV 선택적 마이크로RNA 결합 부위 및 (i) 서열식별번호: 2-4 중 어느 하나; (ii) 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소를 포함한다. 특정 실시양태에서, 본원에 제공된 발현 카세트는 (1) (i) 서열식별번호: 7, 14 또는 15 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위, 및 (2) (i) 서열식별번호: 2-4 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소를 포함한다. 특정 실시양태에서, 본원에 제공된 발현 카세트는 (1) (i) 서열식별번호: 7; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위, 및 (2) (i) 서열식별번

호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소를 포함한다. 한 실시양태에서, 본원에 제공된 발현 카세트는 서열식별번호: 7을 포함하는 마이크로RNA 결합 부위 및 서열식별번호: 2를 포함하는 PV 선택적 조절 요소를 포함한다. 특정 실시양태에서, 본원에 제공된 폴리뉴클레오티드는 (1) (i) 서열식별번호: 14; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위, 및 (2) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소를 포함한다. 한 실시양태에서, 본원에 제공된 발현 카세트는 서열식별번호: 15를 포함하는 마이크로RNA 결합 부위 및 서열식별번호: 2를 포함하는 PV 선택적 조절 요소를 포함한다. 특정 실시양태에서, 본원에 제공된 발현 카세트는 (1) (i) 서열식별번호: 15; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위, 및 (2) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소를 포함한다. 한 실시양태에서, 본원에 제공된 발현 카세트는 서열식별번호: 15를 포함하는 마이크로RNA 결합 부위 및 서열식별번호: 2를 포함하는 PV 선택적 조절 요소를 포함한다.

[0225] 특정 실시양태에서, 본원에 제공된 발현 카세트는 PV 선택적 마이크로RNA 결합 부위 및 본원에 제공된 바와 같은 SCN1A의 발현을 상향조절하는 eTF를 코딩하는 서열을 포함한다. 예시적 실시양태에서, 본 출원은 PV 선택적 조절 요소, 본원에 제공된 바와 같은 SCN1A의 발현을 상향조절하는 eTF 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2-4 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 77-131, 205, 207, 209, 213, 217, 219, 221 또는 223 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 7, 14 또는 15 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 또 다른 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 77 또는 127 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 7; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 또 다른 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 77 또는 127 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%

서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 14; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 또 다른 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 77 또는 127 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 15; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 또 다른 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 92 또는 106 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 7; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 또 다른 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 92 또는 106 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 14; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다. 또 다른 예시적 실시양태에서, 본 출원은 (1) (i) 서열식별번호: 2; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 조절 요소, (2) (i) 서열식별번호: 92 또는 106 중 어느 하나; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는, SCN1A를 상향조절하는 eTF를 코딩하는 서열, 및 (3) (i) 서열식별번호: 15; (ii) (i)의 그의 변이체, 기능적 단편 또는 조합; 또는 (iii) (i) 또는 (ii) 중 어느 하나에 대해 적어도 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 핵산 서열을 포함하는 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트를 제공한다.

[0226] 특정 실시양태에서, PV 선택적 조절 요소 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 본원에 제공된 발현 카세트는 5 kb, 4.9 kb, 4.8 kb, 4.7 kb, 4.6 kb, 4.5 kb, 4.4 kb, 4.3 kb, 4.2 kb, 4.1 kb, 4.0 kb, 3.9 kb, 3.8 kb, 3.7 kb, 3.6 kb, 3.5 kb, 3.4 kb, 3.3 kb, 3.2 kb, 3.1 kb, 3.0 kb, 2.9 kb, 2.8 kb, 2.7 kb, 2.6 kb, 2.5 kb, 2.4 kb, 2.3 kb, 2.2 kb, 2.1 kb, 2.0 kb, 1.9 kb, 1.8 kb, 1.7 kb, 1.6 kb 또는 1.5 kb 미만, 또는 약 1.5-5 kb, 1.5-4.7 kb, 1.5-4.5 kb, 1.5-4.0 kb, 1.5-3.5 kb, 1.5-3.0 kb, 1.5-2.5 kb, 1.5-2.0 kb 크기이다.

[0227] 특정 실시양태에서, 본원에 제공된 발현 카세트는 프로모터에 추가로 1개 이상의 추가의 조절 요소, 예컨대, 예를 들어 전사 개시 또는 종결과 연관된 서열, 인핸서 서열 및 효율적인 RNA 프로세싱 신호를 포함할 수 있다. 예시적인 조절 요소는, 예를 들어 인트론, 인핸서, UTR, 안정성 요소, WPRE 서열, 코작 컨센서스 서열, 번역후 반응 요소, 마이크로RNA 결합 부위 또는 폴리아데닐화 (폴리A) 서열 또는 그의 조합을 포함한다. 조절 요소는 유전자 발현의 전사기, 전사후기 또는 번역기에서 유전자 발현을 조정하는 기능을 할 수 있다. RNA 수준에서, 조절은 번역 (예를 들어, 번역을 위해 mRNA를 안정화시키는 안정성 요소), RNA 절단, RNA 스플라이싱 및/또는 전사 종결의 수준에서 발생할 수 있다. 다양한 실시양태에서, 조절 요소는 관심 세포 유형에서 유전자 발현 선택성을 증가시키고/거나, RNA 전사체가 생산되는 속도를 증가시키고/거나, 생산된 RNA의 안정성을 증가시키고/거나, RNA 전사체로부터의 단백질 합성 속도를 증가시키는 전사 인자를 코딩 영역으로 동원할 수 있다. 예시적 실시양태에서, 본원에 제공된 발현 카세트는 본원에 제공된 바와 같은 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함한다.

[0228] 특정 실시양태에서, 본원에 기재된 발현 카세트는 폴리A 서열을 추가로 포함한다. 적합한 폴리A 서열은, 예를 들어 길이가 약 75 bp인 인공 폴리A (PA75) (예를 들어, WO 2018/126116 참조), 소 성장 호르몬 폴리A, SV40 초기 폴리A 신호, SV40 후기 폴리A 신호, 토끼 베타 글로빈 폴리A, HSV 티미딘 키나제 폴리A, 프로타민 유전자 폴리A, 아데노바이러스 5 EIb 폴리A, 성장 호르몬 폴리A 또는 PBGD 폴리A를 포함한다. 예시적 실시양태에서, 본원에 제공된 발현 카세트에 사용하기에 적합한 폴리A 서열은 hGH 폴리A (서열식별번호: 17) 또는 합성 폴리A (서열식별번호: 16)이다. 전형적으로, 폴리A 서열은 본원에 기재된 발현 카세트에서 eTF를 코딩하는 폴리뉴클레오티드의 하류에 위치한다.

[0229] 특정 실시양태에서, 본원에 제공된 발현 카세트는 1개 이상의 핵 국재화 신호 (NLS)를 코딩하는 1개 이상의 핵산 서열을 추가로 포함한다. 세포 핵 내로 부착되는 단백질의 유입을 용이하게 하는 임의의 NLS 펩티드가 사용될 수 있다. NLS의 예는, 예를 들어 SV40 대형 T-항원 NLS, 뉴클레오플라스민 NLS, EGL-13 NLS, c-Myc NLS 및 TUS-단백질 NLS를 포함한다. 예를 들어, 문헌 [C. Dingwall et al., J. Cell Biol. 107: 841-9 (1988); J.P. Makkerh et al., Curr Biol. 6: 1025-7 (1996); 및 M. Ray et al., Bioconjug. Chem. 26: 1004-7 (2015)]을 참조한다. NLS는 eTF 단백질 서열 상의 어디에나 위치할 수 있지만, 바람직한 실시양태에서 eTF의 N-말단 또는 eTF의 도메인에 접합된다. 예시적 실시양태에서, 본원에 제공된 핵산 카세트는 eTF의 N-말단에 융합된 NLS를 수반한 eTF를 코딩한다. 다른 실시양태에서, 본원에 제공된 핵산 카세트는 eTF의 N-말단에 융합된 제1 NLS 및 eTF의 DBD와 TAD 도메인 사이에 위치한 제2 NLS를 수반한 eTF를 코딩한다.

[0230] 발현 벡터

[0231] 특정 실시양태에서, 본원에 기재된 발현 카세트는 발현 벡터 내로 혼입될 수 있다. 발현 벡터는 형질감염 또는 형질도입을 통해 발현 카세트를 표적 세포로 전달하는데 사용될 수 있다. 벡터는 발현 카세트 또는 트랜스진을 숙주 세포의 게놈 내로 통합시키는 벡터의 능력을 지칭하는 통합 또는 비-통합 벡터일 수 있다. 발현 벡터의 예는 (a) 비-바이러스 벡터, 예컨대 선형 올리고뉴클레오티드 및 원형 플라스미드를 포함하는 핵산 벡터; 인공 염색체, 예컨대 인간 인공 염색체 (HAC), 효모 인공 염색체 (YAC) 및 박테리아 인공 염색체 (BAC 또는 PAC); 에피솜 벡터; 트랜스포존 (예를 들어, 피기백); 및 (b) 바이러스 벡터, 예컨대 레트로바이러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터 및 아데노-연관 바이러스 벡터를 포함하나 이에 제한되지는 않는다.

[0232] 발현 벡터는 선형 올리고뉴클레오티드 또는 원형 플라스미드일 수 있고, 물리적 및 화학적 방법을 포함한 다양한 형질감염 방법을 통해 세포로 전달될 수 있다. 물리적 방법은 일반적으로 유전 물질의 세포내 전달을 용이하게 하는데 있어서 세포 막 장벽을 상쇄시키는 물리적 힘을 사용한 전달 방법을 지칭한다. 물리적 방법의 예는 바늘, 탄도 DNA, 전기천공, 초음파천공, 포토포레이션, 마그네토펙션 및 히드로포레이션의 사용을 포함한다. 화학적 방법은 일반적으로 화학적 담체가 핵산 분자를 세포로 전달하는 방법을 지칭하고, 무기 입자, 지질-기반 벡터, 중합체-기반 벡터 및 펩티드-기반 벡터를 포함할 수 있다.

[0233] 일부 실시양태에서, 발현 벡터는 양이온성 지질 (예를 들어, 양이온성 리포솜)을 사용하여 표적 세포에 투여된다. 다양한 유형의 지질, 예컨대, 예를 들어 지질 나노 에멀젼 (예를 들어, 하나의 불혼화성 액체가 또 다른 액체 중에 분산된, 유화제에 의해 안정화된 분산액임) 또는 고체 지질 나노입자가 유전자 전달을 위해 조사되었다.

[0234] 일부 실시양태에서, 발현 벡터는 펩티드 기반 전달 비히클을 사용하여 표적 세포에 투여된다. 펩티드 기반 전달 비히클은 전달될 유전 물질을 보호하고, 특이적 세포 수용체를 표적화하고, 엔도솜 막을 파괴하고, 유전 물질을 핵 내로 전달하는 이점을 가질 수 있다. 일부 실시양태에서, 발현 벡터는 중합체 기반 전달 비히클을 사

용하여 표적 세포에 투여된다. 중합체 기반 전달 비히클은 천연 단백질, 펩티드 및/또는 폴리사카라이드 또는 합성 중합체를 포함할 수 있다. 한 실시양태에서, 중합체 기반 전달 비히클은 폴리에틸렌이민 (PEI)을 포함한다. PEI는 DNA를 양으로 하전된 입자로 축합시킬 수 있으며, 이는 음이온성 세포 표면 잔기에 결합하고 세포내이입을 통해 세포 내로 도입된다. 다른 실시양태에서, 중합체 기반 전달 비히클은 폴리-L-리신 (PLL), 폴리(DL-락트산) (PLA), 폴리 (DL-락티드-코-글리코시드) (PLGA), 폴리오르니틴, 폴리아르기닌, 히스톤, 프로타민, 덴드리머, 키토산, 덱스트란의 합성 아미노 유도체 및/또는 양이온성 아크릴 중합체를 포함할 수 있다. 특정 실시양태에서, 중합체 기반 전달 비히클은 중합체, 예컨대, 예를 들어 PEG 및 PLL의 혼합물을 포함할 수 있다.

[0235] 특정 실시양태에서, 발현 벡터는 유전자 요법에 적합한 바이러스 벡터일 수 있다. 바이러스 유전자 요법 벡터 또는 유전자 전달 벡터의 바람직한 특징은 높은 역가로 재현가능하고 안정하게 증식 및 정제되는 능력; 표적화된 전달을 매개하는 능력 (예를 들어, 다른 곳에 광범위한 벡터 전파 없이 관심 조직 또는 기관에 특이적으로 트랜스진을 전달하는 능력); 및 유해 부작용을 유도하지 않으면서 유전자 전달 및 트랜스진 발현을 매개하는 능력을 포함할 수 있다.

[0236] 여러 유형의 바이러스, 예를 들어 아데노-연관 바이러스로 지칭되는 비-병원성 파르보바이러스가, 바이러스 감염 경로를 이용하지만 복제 및 독성으로 이어질 수 있는 바이러스 유전자의 후속 발현을 피함으로써 유전자 요법의 목적을 위해 조작되었다. 이러한 바이러스 벡터는, 바이러스 게놈으로부터 코딩 영역의 전부 또는 일부를 결실시키지만, 벡터 게놈을 바이러스 캡시드 내로 패키징하는 것 또는 벡터 핵산 (예를 들어, DNA)을 숙주 염색질 내로 통합시키는 것과 같은 기능에 필요할 수 있는 서열 (예를 들어, 말단 반복 서열)을 무손상으로 남김으로써 수득될 수 있다.

[0237] 다양한 실시양태에서, 적합한 바이러스 벡터는 레트로바이러스 (예를 들어, A-유형, B-유형, C-유형 및 D-유형 바이러스), 아데노바이러스, 파르보바이러스 (예를 들어 아데노-연관 바이러스 또는 AAV), 코로나바이러스, 음성 가닥 RNA 바이러스, 예컨대 오르토믹소바이러스 (예를 들어, 인플루엔자 바이러스), 랍도바이러스 (예를 들어, 광견병 및 수포성 구내염 바이러스), 파라믹소바이러스 (예를 들어 홍역 및 센다이), 양성 가닥 RNA 바이러스, 예컨대 피코르나바이러스 및 알파바이러스, 및 이중-가닥 DNA 바이러스, 예컨대 아데노바이러스, 헤르페스바이러스 (예를 들어, 단순 포진 바이러스 유형 1 및 2, 엡스타인-바르 바이러스, 시토메갈로바이러스) 및 폭스바이러스 (예를 들어, 백시니아, 계두 및 카나리폭스)를 포함한다. 레트로바이러스의 예는 조류 백혈증-육종 바이러스, 인간 T-림프향성 바이러스 유형 1 (HTLV-1), 소 백혈병 바이러스 (BLV), 렌티바이러스 및 스푸마바이러스를 포함한다. 다른 바이러스는, 예를 들어 노워크 바이러스, 토가바이러스, 플라비바이러스, 레오바이러스, 파포바바이러스, 헤파드나바이러스 및 간염 바이러스를 포함한다. 바이러스 벡터는 숙주 게놈 내로 통합되는 그의 능력에 따라 2개의 군 - 통합 및 비-통합으로 분류될 수 있다. 온코레트로바이러스 및 렌티바이러스는 숙주 세포 염색질 내로 통합될 수 있는 한편, 아데노바이러스, 아데노-연관 바이러스 및 포진 바이러스는 염색체외 에피솜으로서 세포 핵에서 우세하게 지속된다.

[0238] 특정 실시양태에서, 적합한 바이러스 벡터는 레트로바이러스 벡터이다. 레트로바이러스는 레트로비리다에 과의 바이러스를 지칭한다. 레트로바이러스의 예는 온코레트로바이러스, 예컨대 뮤린 백혈병 바이러스 (MLV), 및 렌티바이러스, 예컨대 인간 면역결핍 바이러스 1 (HIV-1)을 포함한다. 레트로바이러스 게놈은 단일-가닥 (ss) RNA이고, 시스 또는 트랜스로 제공될 수 있는 다양한 유전자를 포함한다. 예를 들어, 레트로바이러스 게놈은 유전자 발현, 역전사 및 숙주 염색체 내로의 통합을 위한 요소와 함께, 시스-작용 서열, 예컨대 2개의 긴 말단 반복부 (LTR)를 함유할 수 있다. 다른 성분은 새롭게 형성된 비리온으로의 특이적 RNA 패키징을 위한 패키징 신호 (psi 또는 ψ), 및 역전사 동안 양성 가닥 DNA 합성의 개시 부위인 폴리퓨린 트랙 (PPT)을 포함한다. 추가로, 레트로바이러스 게놈은 gag, pol 및 env 유전자를 포함할 수 있다. gag 유전자는 구조 단백질을 코딩하고, pol 유전자는 ssRNA를 동반하며 바이러스 RNA의 DNA로의 역전사를 수행하는 효소를 코딩하고, env 유전자는 바이러스 외피를 코딩한다. 일반적으로, gag, pol 및 env는 바이러스 복제 및 패키징을 위해 트랜스로 제공된다.

[0239] 특정 실시양태에서, 본원에 제공된 레트로바이러스 벡터는 렌티바이러스 벡터일 수 있다. 렌티바이러스의 적어도 5종의 혈청군 또는 혈청형이 인식된다. 상이한 혈청형의 바이러스는 특정 세포 유형 및/또는 숙주를 차등적으로 감염시킬 수 있다. 렌티바이러스는, 예를 들어 영장류 레트로바이러스 및 비-영장류 레트로바이러스를 포함한다. 영장류 레트로바이러스는 HIV 및 원숭이 면역결핍 바이러스 (SIV)를 포함한다. 비-영장류 레트로바이러스는 고양이 면역결핍 바이러스 (FIV), 소 면역결핍 바이러스 (BIV), 염소 관절염-뇌염 바이러스 (CAEV), 말 감염성 빈혈 바이러스 (EIAV) 및 비스나바이러스를 포함한다. 렌티바이러스 또는 렌티벡터는 정지 세포를 형질도입할 수 있다. 온코레트로바이러스 벡터와 같이, 렌티벡터의 설계는 시스- 및 트랜스-작용 서열의 분리에 기

초할 수 있다.

[0240] 특정 실시양태에서, 본 출원은 최적화된 치료 레트로바이러스 벡터에 의한 전달을 위해 설계된 발현 벡터를 제공한다. 레트로바이러스 벡터는 좌측 (5') LTR; 바이러스의 패키징 및/또는 핵 유입을 보조하는 서열; 프로모터; 임의로 1개 이상의 추가의 조절 요소 (예컨대, 예를 들어, 인핸서 또는 폴리A 서열); 임의로 렌티바이러스 역 반응 요소 (RRE); eTF를 코딩하는 서열에 작동가능하게 연결된 PV 선택적 조절 요소를 포함하는 구축물; 임의로 인슐레이터; 및 우측 (3') 레트로바이러스 LTR을 포함하는 렌티바이러스일 수 있다.

[0241] 예시적 실시양태에서, 본원에 제공된 바이러스 벡터는 아데노-연관 바이러스 (AAV)이다. AAV는 인간 및 일부 다른 영장류 종을 감염시키는 작은 복제-결손 비-외피보유 동물 바이러스이다. AAV는 인간 질환을 유발하지 않는 것으로 알려져 있고, 경도 면역 반응을 유도한다. AAV 벡터는 또한 숙주 세포 게놈 내로의 통합 없이 분열 세포 및 정지 세포 둘 다를 감염시킬 수 있다.

[0242] AAV 게놈은 ~4.7kb 길이인 선형 단일 가닥 DNA로 이루어진다. 게놈은 약 145bp 길이인 역전된 말단 반복부 (ITR) 서열이 플랭킹되어 있는 2개의 오픈 리딩 프레임 (ORF)으로 이루어진다. ITR은 5' 말단의 뉴클레오티드 서열 (5' ITR) 및 회문식 서열을 함유하는 3' 말단에 위치하는 뉴클레오티드 서열 (3' ITR)로 이루어진다. ITR은 제2 가닥 합성을 위한 DNA 복제의 개시 동안 프라이머로서 기능하는 T-형상 헤어핀 구조를 상보적 염기 쌍형성에 의해 형성하도록 폴딩됨으로써 시스로 기능한다. 2개의 오픈 리딩 프레임은 비리온의 복제 및 패키징에 수반되는 rep 및 cap 유전자를 코딩한다. 예시적 실시양태에서, 본원에 제공된 AAV 벡터는 rep 또는 cap 유전자를 함유하지 않는다. 이러한 유전자는 하기에 추가로 기재된 바와 같이 비리온을 생산하기 위해 트랜스로 제공될 수 있다.

[0243] 특정 실시양태에서, AAV 벡터는 스터퍼 핵산을 포함할 수 있다. 일부 실시양태에서, 스터퍼 핵산은 녹색 형광 단백질 또는 항생제 내성 유전자, 예컨대 카나마이신 또는 암피실린을 코딩할 수 있다. 특정 실시양태에서, 스터퍼 핵산은 (예를 들어, 5'와 3' ITR 서열들 사이에 위치하는 eTF 트랜스진 서열 및 조절 서열과 비교하여) ITR 서열의 외부에 위치할 수 있다.

[0244] AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12 및 AAV13을 포함한 AAV의 다양한 혈청형이 존재한다. 이들 혈청형은 그의 향성 또는 그가 감염시키는 세포의 유형이 상이하다. AAV는 다수의 혈청형 (예를 들어, 위형)으로부터의 게놈 및 캡시드를 포함할 수 있다. 예를 들어, AAV는 혈청형 5 또는 혈청형 9로부터의 캡시드에 패키징된 혈청형 2 (예를 들어, ITR)의 게놈을 포함할 수 있다. 유사형은 형질도입 효율을 개선시킬 뿐만 아니라 향성을 변경시킬 수 있다.

[0245] 일부 경우에, 혈액 뇌 장벽을 가로지르거나 CNS의 세포를 감염시킬 수 있는 AAV 혈청형이 바람직하다. 일부 경우에, AAV9 또는 그의 변이체는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소를 포함하는 본 개시내용의 발현 카세트를 전달하기 위해 사용된다. 일부 경우에, AAV9 또는 그의 변이체는 PV 선택적 마이크로RNA 결합 부위를 포함하는 본 개시내용의 발현 카세트를 전달하기 위해 사용된다. 일부 경우에, AAV9 또는 그의 변이체는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 본 개시내용의 발현 카세트를 전달하기 위해 사용된다.

[0246] 예시적 실시양태에서, 본 출원은 AAV에 의한 전달을 위해 설계된 발현 벡터를 제공한다. AAV는 임의의 혈청형, 예를 들어 AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-DJ 또는 키메라, 하이브리드 또는 변이체 AAV일 수 있다. AAV는 또한 자기-상보적 AAV (scAAV)일 수 있다. 특정 실시양태에서, AAV에 의한 전달을 위해 설계된 발현 벡터는 5' ITR 및 3' ITR을 포함한다. 특정 실시양태에서, AAV에 의한 전달을 위해 설계된 발현 벡터는 5' ITR, 프로모터, eTF를 코딩하는 트랜스진 및 3' ITR을 포함한다. 특정 실시양태에서, AAV에 의한 전달을 위해 설계된 발현 벡터는 5' ITR, 인핸서, 프로모터, eTF를 코딩하는 트랜스진, 폴리A 서열 및 3' ITR을 포함한다.

[0247] 숙주 세포

[0248] 또 다른 측면에서, 본 발명은 본원에 개시된 바와 같은 발현 카세트 또는 발현 벡터를 포함하는 숙주 세포에 관한 것이다. 숙주 세포는 박테리아 세포, 효모 세포, 곤충 세포 또는 포유동물 세포일 수 있다. 예시적 실시양태에서, 숙주 세포는 관심 바이러스에 의해 감염되기 쉽고 시험관내 배양에 적용가능한 임의의 세포주를 지칭한다.

[0249] 특정 실시양태에서, 본원에 제공된 숙주 세포는 생체외 유전자 요법 목적을 위해 사용될 수 있다. 이러한 실시

양태에서, 세포는 PV 선택적 마이크로RNA 결합 부위 및/또는 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열을 포함하는 핵산 분자 또는 발현 벡터로 형질감염되고, 후속적으로 환자 또는 대상체 내로 이식된다. 이식된 세포는 자가, 동종 또는 이종 기원을 가질 수 있다. 임상 사용을 위해, 세포 단리는 일반적으로 우수 제조 관리기준 (GMP) 조건 하에 수행될 것이다. 이식 전에, 세포 품질 및 미생물 또는 다른 오염물의 부재가 전형적으로 검사되고, 예컨대 방사선 및/또는 면역억제 처리를 사용한 사전조건화가 수행될 수 있다. 추가로, 숙주 세포가 세포 증식 및/또는 분화를 자극하는 성장 인자와 함께 이식될 수 있다.

[0250] 특정 실시양태에서, 숙주 세포는 생체외 유전자 요법에 사용될 수 있다. 바람직하게는, 상기 세포는 진핵 세포, 예컨대 포유동물 세포이고, 이들은 인간, 비-인간 영장류, 예컨대 유인원; 침팬지; 원숭이 및 오랑우탄, 개 및 고양이를 비롯한 가정 동물, 뿐만 아니라 말, 소, 돼지, 양 및 염소와 같은 가축, 또는 마우스, 래트, 기니 피그, 토끼, 햄스터 등을 비제한적으로 포함하는 다른 포유동물 종을 포함하나 이에 제한되지는 않는다. 관련 기술분야의 통상의 기술자는 이식될 환자 또는 대상체에 따라 보다 적절한 세포를 선택할 것이다.

[0251] 특정 실시양태에서, 본원에 제공된 숙주 세포는 자기-재생 및 다능성 특성을 갖는 세포, 예컨대 줄기 세포 또는 유도된 만능 줄기 세포일 수 있다. 줄기 세포는 바람직하게는 중간엽 줄기 세포이다. 중간엽 줄기 세포 (MSC)는 골모세포, 연골세포, 지방세포 또는 근세포 중 적어도 1종으로 분화될 수 있고, 임의의 유형의 조직으로부터 단리될 수 있다. 일반적으로, MSC는 골수, 지방 조직, 제대 또는 말초 혈액으로부터 단리될 것이다. 그의 수득 방법은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 유도된 만능 줄기 세포 (iPS 세포 또는 iPSC로도 공지됨)는 성체 세포로부터 직접적으로 생성될 수 있는 만능 줄기 세포의 한 유형이다. 야마나카 (Yamanaka) 등은 Oct3/4, Sox2, Klf4 및 c-Myc 유전자를 마우스 및 인간 섬유모세포 내로 전달하고 세포가 유전자를 발현하도록 강제함으로써 iPS 세포를 유도하였다 (WO 2007/069666). 후속적으로 톰슨(Thomson) 등은 Klf4 및 c-Myc 대신 Nanog 및 Lin28을 사용하여 인간 iPS 세포를 생산하였다 (WO 2008/118820).

[0252] 예시적 실시양태에서, 본원에 제공된 숙주 세포는 패키징 세포이다. 상기 세포는 부착 세포 또는 현탁 세포일 수 있다. 패키징 세포 및 헬퍼 벡터 또는 바이러스 또는 DNA 구축물(들)은 바이러스 벡터의 완전한 복제 및 패키징에 요구되는 모든 누락 기능을 트랜스로 함께 제공한다.

[0253] 바람직하게는, 상기 패키징 세포는 진핵 세포, 예컨대 원숭이, 인간, 개 및 설치류 세포를 포함한 포유동물 세포이다. 인간 세포의 예는 PER.C6 세포 (WO01/38362), MRC-5 (ATCC CCL-171), WI-38 (ATCC CCL-75), HEK-293 세포 (ATCC CRL-1573), HeLa 세포 (ATCC CCL2) 및 태아 레서스 폐 세포 (ATCC CL-160)이다. 비-인간 영장류 세포의 예는 베로 세포 (ATCC CCL81), COS-1 세포 (ATCC CRL-1650) 또는 COS-7 세포 (ATCC CRL-1651)이다. 개 세포의 예는 MDCK 세포 (ATCC CCL-34)이다. 설치류 세포의 예는 햄스터 세포, 예컨대 BHK21-F, HKCC 세포 또는 CHO 세포이다.

[0254] 포유동물 공급원에 대한 대안으로서, 본 발명에 사용하기 위한 세포주는 조류 공급원, 예컨대 닭, 오리, 거위, 메추라기 또는 꿩으로부터 유래될 수 있다. 조류 세포주의 예는 조류 배아 줄기 세포 (WO01/85938 및 WO03/076601), 불멸화 오리 망막 세포 (WO2005/042728), 및 조류 배아 줄기 세포 유래 세포, 예컨대 닭 세포 (WO2006/108846) 또는 오리 세포, 예컨대 EB66 세포주 (WO2008/129058 및 WO2008/142124)를 포함한다.

[0255] 또 다른 실시양태에서, 상기 숙주 세포는 곤충 세포, 예컨대 SF9 세포 (ATCC CRL-1711), Sf21 세포 (IPLB-Sf21), MG1 세포 (BTI-TN-MG1) 또는 하이 파이브(High Five)™ 세포 (BTI-TN-5B1-4)이다.

[0256] 특정 실시양태에서, 본 발명의 재조합 AAV 벡터/게놈 (예를 들어, PV 선택적 마이크로RNA 결합 부위 및/또는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열을 포함함)을 포함하는 본원에 제공된 숙주 세포는 1개 이상의 추가의 핵산 구축물, 예컨대, 예를 들어 (i) rep 및 cap 유전자를 코딩하지만 ITR 서열을 보유하지 않는 핵산 구축물 (예를 들어, AAV 헬퍼 플라스미드); 및/또는 (ii) AAV 복제에 필요한 아데노바이러스 기능을 제공하는 핵산 구축물 (예를 들어, 플라스미드)을 추가로 포함할 수 있다. 예시적 실시양태에서, 본원에 제공된 숙주 세포는 i) PV 선택적 마이크로RNA 결합 부위 및/또는 본원에 제공된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 (즉, 재조합 AAV 게놈)을 포함하는 발현 벡터; ii) ITR 서열을 보유하지 않는, AAV rep 및 cap 유전자를 코딩하는 핵산 구축물; 및 iii) 아데노바이러스 헬퍼 유전자를 포함하는 핵산 구축물 (하기에 추가로 기재된 바와 같음)을 포함한다.

[0257] 특정 실시양태에서, rep, cap 및 아데노바이러스 헬퍼 유전자는 단일 플라스미드 상에서 조합될 수 있다 (Blouin V et al. J Gene Med. 2004; 6(suppl): S223-S228; Grimm D. et al. Hum. Gene Ther. 2003; 7: 839-850). 따라서, 또 다른 예시적 실시양태에서, 본원에 제공된 숙주 세포는 i) PV 선택적 마이크로RNA 결합 부위

및/또는 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 (즉, 재조합 AAV 게놈)을 포함하는 발현 벡터; 및 ii) ITR 서열을 보유하지 않고 아데노바이러스 헬퍼 유전자를 추가로 포함하는, AAV rep 및 cap 유전자를 코딩하는 플라스미드를 포함한다.

[0258] 또 다른 실시양태에서, 본원에 제공된 숙주 세포는 a) PV 선택적 마이크로RNA 결합 부위 및/또는 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 (즉, 재조합 AAV 게놈)을 포함하는 발현 벡터; b) ITR 서열을 보유하지 않는, AAV rep 및 cap 유전자를 코딩하는 플라스미드; 및 c) 아데노바이러스 헬퍼 유전자 E2a, E4 및 VA RNA를 포함하는 플라스미드를 포함하며; 여기서 공동-형질감염은 아데노바이러스 E1 유전자를 구성적으로 발현하고 교차보완하는 세포, 바람직하게는 포유동물 세포, 예컨대 HEK-293 세포 (ATCC CRL-1573)에서 수행된다.

[0259] 특정 실시양태에서, AAV 벡터의 대규모 생산에 적합한 숙주 세포는 재조합 바큘로바이러스의 조합으로 감염될 수 있는 곤충 세포이다 (Urabe et al. Hum. Gene Ther. 2002; 13: 1935-1943). 예를 들어, SF9 세포는 AAV rep, AAV cap 및 패키징될 AAV 벡터를 각각 발현하는 3개의 바큘로바이러스 벡터로 공동-감염될 수 있다. 재조합 바큘로바이러스 벡터는 바이러스 복제 및/또는 패키징에 요구되는 바이러스 헬퍼 유전자 기능을 제공할 것이다.

[0260] 본 발명에 따른 유전자 요법을 위한 비리온의 구축 및 생산에 대한 추가의 지침은 문헌 [Viral Vectors for Gene Therapy, Methods and Protocols. Series: Methods in Molecular Biology, Vol. 737. Merten and Al-Rubeai (Eds.); 2011 Humana Press (Springer); Gene Therapy. M. Giacca. 2010 Springer-Verlag; Heilbronn R. and Weger S. Viral Vectors for Gene Transfer: Current Status of Gene Therapeutics. In: Drug Delivery, Handbook of Experimental Pharmacology 197; M. Schafer-Korting (Ed.). 2010 Springer-Verlag; pp. 143-170; Adeno-Associated Virus: Methods and Protocols. R. O. Snyder and P. Moulllier (Eds). 2011 Humana Press (Springer); Bunning H. et al. Recent developments in adeno-associated virus technology. J. Gene Med. 2008; 10:717-733; 및 Adenovirus: Methods and Protocols. M. Chillon and A. Bosch (Eds.); Third. Edition. 2014 Humana Press (Springer)]에서 찾아볼 수 있다.

[0261] 비리온 및 비리온의 제조 방법

[0262] 특정 실시양태에서, 본 출원은 PV 선택적 마이크로RNA 결합 부위 및/또는 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열을 포함하는 바이러스 벡터를 포함하는 바이러스 입자를 제공한다. 용어 "바이러스 입자" 및 "비리온"은 본원에서 상호교환가능하게 사용되고, 캡시드 내에 패키징된 바이러스 게놈 (예를 들어, 바이러스 발현 벡터)을 포함하는 감염성 및 전형적으로 복제-결손 바이러스 입자에 관한 것이며, 경우에 따라, 예를 들어 레트로바이러스의 경우, 캡시드를 둘러싸는 지질 외피일 수 있다. "캡시드"는 바이러스 게놈이 패키징되는 구조를 지칭한다. 캡시드는 단백질로 제조된 여러 올리고머 구조 서브유닛으로 이루어진다. 예를 들어, AAV는 3종의 캡시드 단백질: VP1, VP2 및 VP3의 상호작용에 의해 형성된 이십면체 캡시드를 갖는다. 한 실시양태에서, 본원에 제공된 비리온은 PV 선택적 조절 요소 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 AAV 벡터를 패키징함으로써 수득된 재조합 AAV 비리온 또는 rAAV 비리온이다. 또 다른 실시양태에서, 본원에 제공된 비리온은 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열에 작동가능하게 연결된 PV 선택적 조절 요소를 포함하는 AAV 벡터를 단백질 쉘에 패키징함으로써 수득된 재조합 AAV 비리온 또는 rAAV 비리온이다. 또 다른 실시양태에서, 본원에 제공된 비리온은 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열에 작동가능하게 연결된 PV 선택적 조절 요소 및 PV 선택적 마이크로RNA 결합 부위를 포함하는 AAV 벡터를 단백질 쉘에 패키징함으로써 수득된 재조합 AAV 비리온 또는 rAAV 비리온이다.

[0263] 특정 실시양태에서, 본원에 제공된 재조합 AAV 비리온은 특정한 AAV 혈청형으로부터 유래된 AAV 게놈을 동일한 특정한 혈청형의 AAV에 상응하는 천연 Cap 단백질에 의해 형성된 바이러스 입자 내에 캡시드화함으로써 제조될 수 있다. 다른 실시양태에서, 본원에 제공된 AAV 바이러스 입자는 상이한 혈청형으로부터의 단백질 내로 패키징된 소정의 AAV 혈청형의 ITR(들)을 포함하는 바이러스 벡터를 포함한다. 예를 들어, 문헌 [Bunning H et al. J Gene Med 2008; 10: 717-733]을 참조한다. 예를 들어, 소정의 AAV 혈청형으로부터의 ITR을 갖는 바이러스 벡터는 a) 동일하거나 상이한 AAV 혈청형으로부터 유래된 캡시드 단백질로 구성된 바이러스 입자 (예를 들어, AAV2 ITR 및 AAV9 캡시드 단백질; AAV2 ITR 및 AAV8 캡시드 단백질 등); b) 상이한 AAV 혈청형 또는 돌연변이체로부터의 캡시드 단백질의 혼합물로 구성된 모자이크 바이러스 입자 (예를 들어, AAV1 및 AAV9 캡시드 단백질을 갖는 AAV2 ITR); c) 상이한 AAV 혈청형 또는 변이체 사이의 도메인 스와핑에 의해 말단절단된 캡시드 단백질

로 구성된 키메라 바이러스 입자 (예를 들어, AAV9 도메인을 갖는 AAV8 캡시드 단백질을 갖는 AAV2 ITR); 또는 d) 표적 세포 특이적 수용체와의 엄격한 상호작용을 가능하게 하는, 선택적 결합 도메인을 디스플레이하도록 조작된 표적화된 바이러스 입자 (예를 들어, 펩티드 리간드의 삽입에 의해 유전자 말단절단된 AAV9 캡시드 단백질을 갖는 AAV5 ITR; 또는 캡시드 표면에의 펩티드 리간드의 커플링에 의해 비-유전자 변형된 AAV9 캡시드 단백질) 내로 패키징될 수 있다.

[0264] 통상의 기술자는 본원에 제공된 AAV 비리온이 임의의 AAV 혈청형의 캡시드 단백질을 포함할 수 있다는 것을 인지할 것이다. 한 실시양태에서, 바이러스 입자는 AAV1, AAV2, AAV5, AAV8 및 AAV9로 이루어진 군으로부터 선택된 AAV 혈청형으로부터의 캡시드 단백질을 포함하며, 이는 CNS로의 전달에 보다 적합하다 (M. Hocquemiller et al., Hum Gene Ther 27(7): 478-496 (2016)). 특정한 실시양태에서, 바이러스 입자는 발현 카세트의 5'ITR 및 3'ITR 서열이 AAV2 혈청형의 것이고 캡시드 단백질이 AAV9 혈청형의 것인 본 발명의 발현 카세트를 포함한다.

[0265] 형질감염, 안정한 세포주 생산 및 아데노바이러스-AAV 하이브리드, 헤르페스바이러스-AAV 하이브리드 (Conway, J E et al., (1997) J. Virology 71(11):8780-8789) 및 바큘로바이러스-AAV 하이브리드를 포함하는 감염성 하이브리드 바이러스 생산 시스템을 비롯한, rAAV 비리온의 생산을 위한 수많은 방법이 관련 기술분야에 공지되어 있다. rAAV 바이러스 입자의 생산을 위한 rAAV 생산 배양물은 모두 다음을 요구한다: 1) 바큘로바이러스 생산 시스템의 경우에, 예를 들어 인간-유래 세포주, 예컨대 HeLa, A549 또는 293 세포, 또는 곤충-유래 세포주, 예컨대 SF-9를 포함한 적합한 숙주 세포; 2) 야생형 또는 돌연변이 아데노바이러스 (예컨대 온도 감수성 아데노바이러스), 포진 바이러스, 바큘로바이러스, 또는 헬퍼 기능을 제공하는 플라스미드 구축물에 의해 제공되는 적합한 헬퍼 바이러스 기능; 3) AAV rep 및 cap 유전자 및 유전자 산물; 4) AAV ITR 서열이 플랭킹되어 있는 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함); 및 5) rAAV 생산을 지지하는 적합한 배지 및 배지 성분을 요구한다.

[0266] 다양한 실시양태에서, 본원에 기재된 숙주 세포는 하기 3종의 성분을 포함한다: (1) rep 유전자 및 cap 유전자, (2) 헬퍼 기능을 제공하는 유전자, 및 (3) ITR이 플랭킹되어 있는 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함). AAV rep 유전자, AAV cap 유전자 및 헬퍼 기능을 제공하는 유전자는, 상기 유전자를 벡터, 예컨대, 예를 들어 플라스미드 내로 혼입시키고, 상기 벡터를 숙주 세포 내로 도입함으로써 세포 내로 도입될 수 있다. rep, cap 및 헬퍼 기능 유전자는 동일한 플라스미드 내로 또는 상이한 플라스미드 내로 혼입될 수 있다. 바람직한 실시양태에서, AAV rep 및 cap 유전자는 하나의 플라스미드 내로 혼입되고, 헬퍼 기능을 제공하는 유전자는 또 다른 플라스미드 내로 혼입된다. 비리온 생산을 위한 숙주 세포의 생성을 위한 다양한 플라스미드 (예를 들어, AAV rep 및 cap 유전자, 헬퍼 기능 또는 트랜스진을 포함함)는 관련 기술분야에 널리 공지된 임의의 적합한 방법을 사용하여 세포 내로 도입될 수 있다. 형질감염 방법의 예는 인산칼슘과의 공동-침전, DEAE-덱스트란, 폴리브렌, 전기천공, 미세주사, 리포솜-매개 융합, 리포펙션, 레트로바이러스 감염 및 바이오리스틱 형질감염을 포함하나 이에 제한되지는 않는다. 특정 실시양태에서, rep 및 cap 유전자, 헬퍼 기능 및 ITR이 플랭킹되어 있는 트랜스진을 제공하는 플라스미드는 세포 내로 동시에 도입될 수 있다. 또 다른 실시양태에서, rep 및 cap 유전자 및 헬퍼 기능을 제공하는 플라스미드는 트랜스진을 포함하는 플라스미드의 도입 전 또는 후에 세포에 도입될 수 있다. 예시적 실시양태에서, 세포는 다음 3종의 플라스미드로 동시에 형질감염된다 (예를 들어, 삼중 형질감염 방법): (1) ITR이 플랭킹되어 있는 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함)을 포함하는 플라스미드, (2) AAV rep 및 cap 유전자를 포함하는 플라스미드, 및 (3) 헬퍼 기능을 제공하는 유전자를 포함하는 플라스미드. 예시적인 숙주 세포는 293, A549 또는 HeLa 세포일 수 있다.

[0267] 다른 실시양태에서, (1) AAV rep 및 cap 유전자, (2) 헬퍼 기능을 제공하는 유전자, 및 (3) ITR이 플랭킹되어 있는 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함) 중 1종 이상은 에피솜으로 및/또는 패키징 세포의 게놈 내로 통합되어 패키징 세포에 의해 보유될 수 있다. 한 실시양태에서, 숙주 세포는 AAV rep 및 cap 유전자 및 헬퍼 기능이 숙주 세포에서 안정하게 유지되고 ITR이 플랭킹되어 있는 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함)을 함유하는 플라스미드로 숙주 세포가 일시적으로 형질감염된 패키징 세포일 수 있다. 또 다른 실시양태에서, 숙주 세포는 AAV rep 및 cap 유전자가 숙

주 세포에서 안정하게 유지되고 ITR이 플랭킹된 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함)을 함유하는 플라스미드 및 헬퍼 기능을 함유하는 플라스미드로 숙주 세포가 일시적으로 형질감염된 패키징 세포이다. 또 다른 실시양태에서, 숙주 세포는 헬퍼 기능이 숙주 세포에서 안정적으로 유지되고 ITR이 플랭킹되어 있는 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함)을 함유하는 플라스미드 및 rep 및 cap 유전자를 함유하는 플라스미드로 숙주 세포가 일시적으로 형질감염된 패키징 세포일 수 있다. 또 다른 실시양태에서, 숙주 세포는 rep 및 cap 유전자, 헬퍼 기능 및 ITR이 플랭킹되어 있는 트랜스진 서열 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함)로 안정적으로 형질감염된 생산자 세포주일 수 있다. 예시적인 패키징 및 생산자 세포는 293, A549 또는 HeLa 세포로부터 유래될 수 있다.

[0268] 또 다른 실시양태에서, 생산자 세포주는 Rep 및 Cap 단백질을 제공하는 바큘로바이러스 발현 벡터로 감염된 곤충 세포주 (전형적으로 Sf9 세포)이다. 이 시스템은 아데노바이러스 헬퍼 유전자를 요구하지 않는다 (Ayuso E, et al., Curr. Gene Ther. 2010, 10:423-436).

[0269] 본원에 사용된 용어 "cap 단백질"은 천연 AAV Cap 단백질 (예를 들어 VP1, VP2, VP3)의 적어도 1종의 기능적 활성을 갖는 폴리펩티드를 지칭한다. cap 단백질의 기능적 활성의 예는 캡시드의 형성을 유도하고, 단일-가닥 DNA의 축적을 용이하게 하고, 캡시드 내로의 AAV DNA 패키징 (즉, 캡시드화)을 용이하게 하고, 세포 수용체에 결합하고, 숙주 세포 내로의 비리온의 진입을 용이하게 하는 능력을 포함한다. 원칙적으로, 임의의 Cap 단백질이 본 발명의 문맥에 사용될 수 있다.

[0270] Cap 단백질은 숙주 향성, 세포, 조직 또는 기관 특이성, 수용체 용법, 감염 효율 및 AAV 바이러스의 면역원성에 영향을 미치는 것으로 보고되었다. 따라서, rAAV에 사용하기 위한 AAV 캡은, 예를 들어 대상체의 종 (예를 들어 인간 또는 비-인간), 대상체의 면역학적 상태, 장기 또는 단기 치료에 대한 대상체의 적합성, 또는 특정한 치료 용도 (예를 들어 특정한 질환 또는 장애의 치료 또는 특정한 세포, 조직 또는 기관으로의 전달)를 고려하여 선택될 수 있다. 특정 실시양태에서, cap 단백질은 AAV1, AAV2, AAV5, AAV8 및 AAV9 혈청형으로 이루어진 군의 AAV로부터 유래된다. 예시적 실시양태에서, cap 단백질은 AAV9로부터 유래된다.

[0271] 일부 실시양태에서, 본 발명의 방법에 사용하기 위한 AAV Cap은 상기 언급된 AAV cap 중 1종 또는 그의 코딩 핵산의 돌연변이유발 (즉, 삽입, 결실 또는 치환)에 의해 생성될 수 있다. 일부 실시양태에서, AAV cap은 상기 언급된 AAV cap 중 1종 이상에 대해 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98% 또는 99% 또는 그 초과로 유사하다.

[0272] 일부 실시양태에서, AAV cap은 상기 언급된 AAV cap 중 2, 3, 4종 또는 그 초과로부터의 도메인을 포함하는 키메라이다. 일부 실시양태에서, AAV cap은 2 또는 3종의 상이한 AAV 또는 재조합 AAV로부터 기원하는 VP1, VP2 및 VP3 단량체의 모자이크이다. 일부 실시양태에서, rAAV 조성물은 상기 언급된 cap 중 1종 초과를 포함한다.

[0273] 일부 실시양태에서, rAAV 비리온에 사용하기 위한 AAV cap은 이종 서열 또는 다른 변형을 함유하도록 조작된다. 예를 들어, 선택적 표적화 또는 면역 회피를 부여하는 펩티드 또는 단백질 서열은 cap 단백질로 조작될 수 있다. 대안적으로 또는 추가로, cap은 rAAV의 표면이 폴리에틸렌 글리콜화 (즉, PEG화)되도록 화학적으로 변형될 수 있으며, 이는 면역 회피를 용이하게 할 수 있다. cap 단백질은 또한 (예를 들어, 그의 천연 수용체 결합을 제거하도록 또는 면역원성 에피토프를 차폐하도록) 돌연변이유발될 수 있다.

[0274] 본원에 사용된 용어 "rep 단백질"은 천연 AAV rep 단백질 (예를 들어 rep 40, 52, 68, 78)의 적어도 1종의 기능적 활성을 갖는 폴리펩티드를 지칭한다. rep 단백질의 기능적 활성의 예는 DNA 헬리카제 활성뿐만 아니라 AAV DNA 복제 기점의 인식, 결합 및 닉킹을 통한 DNA의 복제를 용이하게 하는 것을 포함한, 단백질의 생리학적 기능과 연관된 임의의 활성을 포함한다. 추가의 기능은 AAV (또는 다른 이종) 프로모터로부터의 전사의 조정 및 숙주 염색체 내로의 AAV DNA의 부위-특이적 통합을 포함한다. 특정한 실시양태에서, AAV rep 유전자는 혈청형 AAV1, AAV2, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 또는 AAVrh10; 보다 바람직하게는 AAV1, AAV2, AAV5, AAV8 및 AAV9로 이루어진 군으로부터 선택된 AAV 혈청형으로부터의 것일 수 있다.

[0275] 일부 실시양태에서, 본 발명의 방법에 사용하기 위한 AAV rep 단백질은 상기 언급된 AAV rep 중 1종 또는 그의 코딩 핵산의 돌연변이유발 (즉, 삽입, 결실 또는 치환)에 의해 생성될 수 있다. 일부 실시양태에서, AAV rep는

상기 언급된 AAV rep 중 1종 이상에 대해 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98% 또는 99% 또는 그 초과로 유사하다.

[0276] 본원에 사용된 표현 "헬퍼 기능" 또는 "헬퍼 유전자"는 AAV가 복제에 의존하는 바이러스 단백질을 지칭한다. 헬퍼 기능은 AAV 유전자 전사의 활성화, 단계 특이적 AAV mRNA 스플라이싱, AAV DNA 복제, cap 발현 산물의 합성 및 AAV 캡시드 어셈블리에 수반되는 단백질을 비제한적으로 포함하는, AAV 복제에 요구되는 단백질을 포함한다. 바이러스-기반 보조 기능은 임의의 공지된 헬퍼 바이러스, 예컨대 아데노바이러스, 헤르페스바이러스 (단순 포진 바이러스 유형-1 이외의 것) 및 백시니아 바이러스로부터 유래될 수 있다. 헬퍼 기능은 아데노바이러스 E1, E2a, VA 및 E4 또는 헤르페스바이러스 UL5, ULB, UL52 및 UL29 및 헤르페스바이러스 폴리머라제를 비제한적으로 포함한다. 바람직한 실시양태에서, AAV가 복제에 의존하는 단백질은 아데노바이러스로부터 유래된다.

[0277] 일부 실시양태에서, AAV가 본 발명의 방법에서의 사용을 위해 복제에 의존하는 바이러스 단백질은 상기 언급된 바이러스 단백질 중 1종 또는 그의 코딩 핵산의 돌연변이유발 (즉, 삽입, 결실 또는 치환)에 의해 생성될 수 있다. 일부 실시양태에서, 바이러스 단백질은 상기 언급된 바이러스 단백질 중 1종 이상에 대해 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98% 또는 99% 또는 그 초과로 유사하다.

[0278] AAV가 복제에 의존하는 cap 단백질, rep 단백질 및 바이러스 단백질의 기능을 검정하는 방법은 관련 기술분야에 널리 공지되어 있다.

[0279] 관심 트랜스진 (예를 들어, PV 선택적 마이크로RNA 결합 부위, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및/또는 PV 선택적 프로모터 중 1종 이상을 포함함)을 발현하기 위한 숙주 세포는 AAV 비리온의 어셈블리에 적절한 조건 하에 성장될 수 있다. 특정 실시양태에서, 숙주 세포는 AAV 비리온의 어셈블리 및 배지 내로의 비리온의 방출을 촉진하기 위해 적합한 기간 동안 성장된다. 일반적으로, 세포는 약 24시간, 약 36시간, 약 48시간, 약 72시간, 약 4일, 약 5일, 약 6일, 약 7일, 약 8일, 약 9일 또는 최대 약 10일 동안 성장될 수 있다. 약 10일 후 (또는 배양 조건 및 사용된 특정한 숙주 세포에 따라 더 빨리), 생산 수준은 일반적으로 유의하게 감소한다. 일반적으로, 배양 시간은 바이러스 생산 시점으로부터 측정된다. 예를 들어, AAV의 경우에, 바이러스 생산은 일반적으로 본원에 기재된 바와 같은 적절한 숙주 세포에 헬퍼 바이러스 기능을 공급할 때 시작된다. 일반적으로, 세포는 헬퍼 바이러스 감염 약 48 내지 약 100시간, 바람직하게는 약 48 내지 약 96시간, 바람직하게는 약 72 내지 약 96시간, 바람직하게는 약 68 내지 약 72시간 후에 (또는 바이러스 생산이 시작된 후에) 수거된다.

[0280] rAAV 생산 배양물은 이용되는 특정한 숙주 세포에 적합한 다양한 조건 하에 (넓은 온도 범위에 걸쳐, 다양한 시간 길이 등 동안) 성장될 수 있다. rAAV 생산 배양물은 적합한 부착-의존성 용기, 예컨대, 예를 들어 롤러 병, 중공 섬유 필터, 마이크로캐리어 및 충전층 또는 유동층 생물반응기에서 배양될 수 있는 부착-의존성 배양물을 포함한다. rAAV 벡터 생산 배양물은 또한, 예를 들어 스피너 플라스크, 교반 탱크 생물반응기 및 일회용 시스템, 예컨대 웨이브 백 시스템을 포함한 다양한 방식으로 배양될 수 있는 현탁액-적합화 숙주 세포, 예컨대 HeLa, 293 및 SF-9 세포를 포함할 수 있다.

[0281] 관련 기술분야에 공지된 적합한 배지가 rAAV 비리온의 생산에 사용될 수 있다. 이들 배지는 변형 이글 배지 (MEM), 둘베코 변형 이글 배지 (DMEM)를 포함한, 하이클론 래보러토리즈(Hyclone Laboratories) 및 JRH에 의해 생산된 배지를 비제한적으로 포함하며, 이들 각각은 그 전체가 본원에 참조로 포함된다. 특정 실시양태에서, rAAV 생산 배양 배지는 0.5%-20% (v/v 또는 w/v) 수준의 혈청 또는 혈청-유래 재조합 단백질로 보충될 수 있다. 대안적으로, rAAV 벡터는 동물-유래 산물이 없는 배지로도 지칭될 수 있는 무혈청 조건에서 생산될 수 있다.

[0282] 숙주 세포를 배양하여 AAV 비리온을 생산한 후, 이어서 생성된 비리온을 수거하고 정제할 수 있다. 특정 실시양태에서, AAV 비리온은 (1) 숙주 세포의 용해에 의해 생산 배양물의 숙주 세포로부터 수득될 수 있고/거나 (2) 형질감염 후 일정 기간, 바람직하게는 72시간 후 상기 세포의 배양 배지로부터 수득될 수 있다. rAAV 비리온은 세포가 무손상 세포로부터 배지 내로 rAAV 비리온의 방출을 유발하는 조건 하에 배양되는 한, 생산 배양물로부터 소비된 배지로부터 수거될 수 있다 (예를 들어, 미국 특허 번호 6,566,118 참조). 세포를 용해시키는 적합한 방법은 또한 관련 기술분야에 공지되어 있고, 예를 들어 다수의 동결/해동 사이클, 초음파처리, 마이크로유동화 및 화학물질, 예컨대 세제 및/또는 프로테아제로의 처리를 포함한다.

[0283] 수거 후, rAAV 비리온을 정제할 수 있다. 본원에 사용된 용어 "정제된"은 rAAV 비리온이 자연 발생하거나 초기에 제조되는 경우에 또한 존재할 수 있는 다른 성분 중 적어도 일부가 결여된 rAAV 비리온의 제제를 포함한다. 따라서, 예를 들어 정제된 rAAV 비리온은 공급원 혼합물, 예컨대 배양 용해물 또는 생산 배양물 상청액으로부터

그를 풍부화하기 위한 단리 기술을 사용하여 제조될 수 있다. 풍부화는 다양한 방식으로, 예컨대, 예를 들어 용액 중에 존재하는 DNase-저항성 입자 (DRP) 또는 게놈 카피 (gc)의 비율에 의해 또는 감염성에 의해 측정될 수 있거나, 또는 공급원 혼합물 중에 존재하는 제2의 잠재적 간섭 물질, 예컨대 오염물, 예컨대 생산 배양 오염물을 포함한 오염물, 또는 헬퍼 바이러스, 배지 성분 등을 포함한 공정중 오염물과 관련하여 측정될 수 있다.

[0284] 특정 실시양태에서, rAAV 생산 배양 수거물은 숙주 세포 파편을 제거하기 위해 정화될 수 있다. 일부 실시양태에서, 생산 배양 수거물은 다양한 표준 기술, 예컨대 원심분리 또는 0.2 μm 이상의 세공 크기의 필터 (예를 들어, 셀룰로스 아세테이트 필터 또는 일련의 심층 필터)를 통한 여과를 사용하여 정화될 수 있다.

[0285] 특정 실시양태에서, rAAV 생산 배양 수거물은 벤조나제(Benzonase)™로 추가로 처리되어 생산 배양물에 존재하는 임의의 고분자량 DNA를 소화시킨다. 일부 실시양태에서, 벤조나제™ 소화는 표준 조건 하에, 예를 들어 최종 농도 1-2.5 단위/ml의 벤조나제™로 주위 온도 내지 37℃ 범위의 온도에서 30분 내지 수시간의 기간 동안 수행된다.

[0286] 특정 실시양태에서, rAAV 비리온은 하기 정제 단계 중 1개 이상을 사용하여 단리 또는 정제될 수 있다: 평형 원심분리; 관통형 음이온 교환 여과; rAAV 입자를 농축시키기 위한 접선 흐름 여과 (TFF); 아파타이트 크로마토그래피에 의한 rAAV 포획; 헬퍼 바이러스의 열 불활성화; 소수성 상호작용 크로마토그래피에 의한 rAAV 포획; 크기 배제 크로마토그래피 (SEC)에 의한 완충제 교환; 나노여과; 및 음이온 교환 크로마토그래피, 양이온 교환 크로마토그래피 또는 친화성 크로마토그래피에 의한 rAAV 포획. 이들 단계는 단독으로, 다양한 조합으로, 또는 상이한 순서로 사용될 수 있다. rAAV 입자를 정제하는 방법은, 예를 들어 문헌 [Xiao et al., (1998) Journal of Virology 72:2224-2232]; 미국 특허 번호 6,989,264 및 8,137,948; 및 WO 2010/148143에서 발견된다.

[0287] 특정 실시양태에서, 정제된 AAV 비리온은 PBS에 대해 투석되고, 여과되고, -80℃에서 저장될 수 있다. 바이러스 게놈의 역가는 표준 곡선으로서 선형화된 플라스미드 DNA를 사용하여 정량적 PCR에 의해 결정될 수 있다 (예를 들어, 문헌 [Lock M, et al., Hum. Gene Ther. 2010; 21:1273-1285] 참조).

[0288] 제약 조성물

[0289] 특정 실시양태에서, 본 출원은 PV 선택적 마이크로RNA 결합 부위 및/또는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및 제약상 허용되는 담체를 포함하는 조성물을 제공한다. 다른 실시양태에서, 본 출원은 PV 선택적 마이크로RNA 결합 부위 및/또는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 서열 및 제약상 허용되는 담체를 포함하는 비리온을 제공한다. 예시적 실시양태에서, 이러한 조성물은 유전자 요법 적용에 적합하다. 제약 조성물은 바람직하게는 멸균되고, 제조 및 저장 조건 하에 안정하다. 멸균 용액은, 예를 들어 멸균 여과 막을 통한 여과에 의해 달성될 수 있다.

[0290] 제약 조성물 중 허용되는 담체 및 부형제는 바람직하게는 사용되는 투여량 및 농도에서 수용자에게 비독성이다. 허용되는 담체 및 부형제는 완충제, 예컨대 포스페이트, 시트레이트, HEPES 및 TAE, 항산화제, 예컨대 아스코르브산 및 메티오닌, 보존제, 예컨대 헥사메토늄 클로라이드, 옥타데실디메틸벤질 암모늄 클로라이드, 레조르시놀 및 벤즈알코늄 클로라이드, 단백질, 예컨대 인간 혈청 알부민, 젤라틴, 덱스트란 및 이뮤노글로불린, 친수성 중합체, 예컨대 폴리비닐피롤리돈, 아미노산, 예컨대 글리신, 글루타민, 히스티딘 및 리신 및 탄수화물, 예컨대 글루코스, 만노스, 수크로스 및 소르비톨을 포함할 수 있다. 본 개시내용의 제약 조성물은 주사가능한 제제의 형태로 비경구로 투여될 수 있다. 주사용 제약 조성물은 멸균 용액 또는 임의의 제약상 허용되는 액체를 비히클로서 사용하여 제제화될 수 있다. 제약상 허용되는 비히클은 멸균수 및 생리 염수를 포함하나 이에 제한되지는 않는다.

[0291] 본 개시내용의 제약 조성물은 마이크로캡슐, 예컨대 히드록실메틸셀룰로스 또는 젤라틴-마이크로캡슐 및 폴리메틸메타크릴레이트 마이크로캡슐로 제조될 수 있다. 본 개시내용의 제약 조성물은 또한 다른 약물 전달 시스템, 예컨대 리포솜, 알부민 마이크로구체, 마이크로에멀젼, 나노입자 및 나노캡슐로 제조될 수 있다. 유전자 요법을 위한 제약 조성물은 허용되는 희석제일 수 있거나 또는 유전자 전달 비히클이 포매되어 있는 느린 방출 매트릭스를 포함할 수 있다.

[0292] 본원에 제공된 제약 조성물은 비경구 투여, 피하 투여, 정맥내 투여, 근육내 투여, 동맥내 투여, 실질내 투여, 척수강내 투여, 대수조내 투여, 뇌실내 투여 또는 복강내 투여를 위해 제제화될 수 있다. 제약 조성물은 또한 비강, 분무, 경구, 에어로졸, 직장 또는 질 투여를 위해 제제화되거나 또는 그를 통해 투여될 수 있다. 한 실시양태에서, 본원에 제공된 제약 조성물은 CNS 또는 뇌 척수액 (CSF)에, 즉 실질내 주사, 척수강내 주사, 대수조내 주사 또는 뇌실내 주사에 의해 투여된다. 조직 표적은 특이적, 예를 들어 CNS일 수 있거나 또는 이는 여

러 조직, 예를 들어 근육 및 CNS 조직의 조합일 수 있다. 예시적인 조직 또는 다른 표적은 간, 골격근, 심장 근육, 지방 침착물, 신장, 폐, 혈관 내피, 상피, 조혈 세포, CNS 및/또는 CSF를 포함할 수 있다. 바람직한 실시양태에서, PV 선택적 마이크로RNA 결합 부위 및/또는 SCN1A를 선택적으로 상향조절하는 eTF를 포함하는 본원에 제공된 제약 조성물은 CNS 또는 CSF 주사에, 즉 실질내 주사, 척수강내 주사, 대수조내 주사 또는 뇌실내 주사에 의해 투여된다. 이들 방법 중 1개 이상은 본 개시내용의 제약 조성물을 투여하는데 사용될 수 있다.

[0293] 특정 실시양태에서, 본원에 제공된 제약 조성물은 "유효량" 또는 "치료 유효량"을 포함한다. 본원에 사용된 이러한 양은 목적하는 치료 결과를 달성하는데 필요한 투여량에서 및 필요한 기간 동안, 예컨대 SCN1A 발현의 수준을 증가시키고/거나 발작의 빈도 및/또는 지속기간을 감소시키는데 효과적인 양을 지칭한다.

[0294] 본 개시내용의 제약 조성물의 투여량은 투여 경로, 치료될 질환 및 대상체의 신체적 특징 (예를 들어, 연령, 체중, 전반적 건강)을 포함한 인자에 좌우된다. 투여량은 최적의 치료 반응을 제공하도록 조정될 수 있다. 전형적으로, 투여량은 유의한 독성을 유도하지 않으면서 질환을 효과적으로 치료하는 양일 수 있다. 한 실시양태에서, 본원에 제공된 AAV 벡터는 SCN1A 결핍 (예를 들어, 드라베 증후군 포함)의 치료를 위해 $5x10^{11}$ 내지 $1x10^{14}$ gc/kg (환자 체중 킬로그램당 게놈 카피 (gc/kg))의 범위 내의 양 또는 용량으로 환자에게 투여될 수 있다. 보다 특정한 실시양태에서, AAV 벡터는 약 $5x10^{11}$ gc/kg 내지 약 $3x10^{13}$ gc/kg 또는 약 $1x10^{12}$ 내지 약 $1x10^{14}$ gc/kg 또는 약 $1x10^{12}$ 내지 약 $1x10^{13}$ gc/kg 또는 약 $5x10^{11}$ gc/kg, $1x10^{12}$ gc/kg, $1.5x10^{12}$ gc/kg, $2.0x10^{12}$ gc/kg, $2.5x10^{12}$ gc/kg, $3x10^{12}$ gc/kg, $3.5x10^{12}$ gc/kg, $4x10^{12}$ gc/kg, $4.5x10^{12}$ gc/kg, $5x10^{12}$ gc/kg, $5.5x10^{12}$ gc/kg, $6x10^{12}$ gc/kg, $6.5x10^{12}$ gc/kg, $7x10^{12}$ gc/kg, $7.5x10^{12}$ gc/kg, $8x10^{12}$ gc/kg, $8.5x10^{12}$ gc/kg, $9x10^{12}$ gc/kg 또는 $9.5x10^{12}$ gc/kg의 범위 내에 포함되는 양으로 투여된다. gc/kg는, 예를 들어 qPCR 또는 디지털 액적 PCR (ddPCR)에 의해 결정될 수 있다 (예를 들어, 문헌 [M. Lock et al., Hum Gene Ther Methods. 2014 Apr;25(2): 115-25] 참조). 또 다른 실시양태에서, 본원에 제공된 AAV 벡터는 SCN1A 결핍 (예를 들어, 드라베 증후군 포함)의 치료를 위해 $1x10^{9}$ 내지 $1x10^{11}$ iu/kg (벡터의 감염 단위 (iu)/대상체 또는 환자 체중 (kg))의 범위 내의 양 또는 용량으로 환자에게 투여될 수 있다. 특정 실시양태에서, 제약 조성물은 필요에 따라 단위 용량으로 형성될 수 있다. 이러한 단일 투여 단위는 약 $1x10^{9}$ gc 내지 약 $1x10^{15}$ gc를 함유할 수 있다.

[0295] 본 개시내용의 제약 조성물은 그를 필요로 하는 대상체에게, 예를 들어 1회 이상 (예를 들어, 1-10회 이상) 매일, 매주, 매월, 2년마다, 매년 또는 의학적으로 필요한만큼 투여될 수 있다. 예시적 실시양태에서, 단일 투여로 충분하다. 한 실시양태에서, PV 선택적 마이크로RNA 결합 부위 및/또는 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 발현 카세트를 포함하는 제약 조성물은 인간 대상체에 사용하기에 적합하고, 실질내 주사, 척수강내 주사, 대수조내 주사 또는 뇌실내 주사에 의해 투여된다. 한 실시양태에서, 제약 조성물은 볼루스 주사에 의해 말초 정맥을 통해 전달된다. 다른 실시양태에서, 제약 조성물은 약 10분 (±5분), 약 20분 (±5분), 약 30분 (±5분), 약 60분 (±5분) 또는 약 90분 (±10분)에 걸친 주입에 의해 말초 정맥을 통해 전달된다.

[0296] 또 다른 측면에서, 본 출원은 1개 이상의 용기 내에 본원에 기재된 바와 같은 핵산 분자, 벡터, 숙주 세포, 비리온 또는 제약 조성물을 포함하는 키트를 추가로 제공한다. 키트는 환자에게 키트 내에 함유된 핵산 분자, 벡터, 숙주 세포 또는 비리온을 투여하는 방법을 기재하는 지침서 또는 패키징 물질을 포함할 수 있다. 키트의 용기는 임의의 적합한 물질, 예를 들어 유리, 플라스틱, 금속 등 및 임의의 적합한 크기, 모양 또는 형상일 수 있다. 특정 실시양태에서, 키트는 적합한 액체 또는 용액 형태로 핵산 분자, 벡터, 숙주 세포, 비리온 또는 제약 조성물을 함유하는 1개 이상의 앰풀 또는 주사기를 포함할 수 있다.

[0297] 치료 방법

[0298] 한 측면에서, 본 출원은 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 사용하는 방법을 제공한다. 특정 실시양태에서, 본 출원은 세포에서 SCN1A의 발현을 상향조절하기 위해 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 방법을 제공한다. 다양한 실시양태에서, 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF는 시험관내, 생체내 또는 생체외 세포에서 SCN1A의 발현을 조정하는데 사용될 수 있다.

[0299] 특정 실시양태에서, 본 출원은 SCN1A와 연관된 질환 또는 장애의 치료를 필요로 하는 대상체에게 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여함으로써 상기 질환 또는 장애를 치료하는 방법을 제공한다. 특정 실시양태에

서, 장애는 중추 신경계 장애이다. 예시적 실시양태에서, 질환 또는 장애는 SCN1A의 반수체기능부전과 연관된다. 특정 실시양태에서, 장애는 SCN1A 반수체기능부전과 연관된 간질이다. 특정 실시양태에서, 반수체기능부전은 대상체가 SCN1A 유전자의 기능 상실 돌연변이에 대해 이형접합인 결과이다. 특정 실시양태에서, 장애는 SCN1A 유전자 내의 삽입, 결실 또는 치환과 연관된 간질이다. 특정 실시양태에서, 장애는 SCN1A 유전자 내의 점 돌연변이와 연관된 간질이다. 특정 실시양태에서, 질환 또는 장애를 치료하는 방법은 SCN1A의 과소발현이 교정되거나, 건강한 개체의 수준 이내가 되거나 또는 표준 의료 관리에 의해 정의된 바와 같은 정상 범위 이내가 되도록 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것을 포함한다. 특정 실시양태에서, 본원에 개시된 방법은 SCN1A의 비정상적 발현을 유발하는 1개 이상의 돌연변이를 포함하는 내인성 SCN1A와 연관된 질환 또는 장애를 치료하는데 사용된다.

[0300] 특정 실시양태에서, 본 출원은 질환 또는 장애와 연관된 증상의 호전을 필요로 하는 대상체에게 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여함으로써 질환 또는 장애와 연관된 증상을 호전시키는 방법을 제공한다.

[0301] 예시적 실시양태에서, 본 출원은 SCN1A에서의 돌연변이 (예를 들어, 점 돌연변이, 치환, 결실, 역위 등), Nav1.1의 결핍 및/또는 Nav1.1의 감소된 활성과 연관된 질환, 장애 또는 증상의 치료를 필요로 하는 대상체에게 SCN1A 유전자 또는 그의 단백질 산물 Nav1.1의 발현을 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여함으로써 상기 질환, 장애 또는 증상을 치료하는 방법을 제공한다. 전압-게이팅 나트륨 이온 채널은 가로무늬근 및 뉴런 조직에서의 활동 전위의 생성 및 전파에 중요하다. 전압-게이팅 나트륨 이온 채널은 큰 중심 세공-형성 글리코실화 알파 서브유닛 및 2개의 보다 작은 보조 베타 서브유닛으로 이루어진 이형체 복합체이다. SCN1A 유전자에 의해 코딩되는 큰 알파 서브유닛 Nav1.1 서브유닛은 다양한 질환 또는 장애, 예컨대 드라베 증후군과 관련된다. Nav1.1은 뉴런에서 발현되고, SCN1B 유전자에 의해 발현된 Navβ1을 포함한 다양한 베타 서브유닛과 어셈블링될 수 있다.

[0302] 특정 실시양태에서, 본 출원은 내인성 SCN1A 유전자의 발현을 선택적으로 상향조절하는 eTF를 사용하여 SCN1A에서의 돌연변이 (예를 들어, 결실, 삽입, 역위, 점 돌연변이 (예를 들어, 넌센스 돌연변이, 미스센스 돌연변이) 등) 또는 Nav1.1의 감소된 활성과 연관된 질환을 치료하는 방법을 제공한다. SCN1A 돌연변이와 연관된 질환 및 장애는 드라베 증후군, 오타하라 증후군, 간질, 조기 영아 간질성 뇌병증 6 (EIEE6), 가족성 열성 발작 3A (FEB3A), 전신 긴장성-간대성 발작 동반 난치성 소아기 간질 (ICEGTC), 편두통, 가족성 편마비 3 (FHM3), 파나이이오토풀로스 증후군, 가족성 심방 세동 13 (ATFB13), 열성 발작 플러스 동반 전신 간질 유형 1 (gefs+ 유형 1), 브루가다 증후군, 비특이적 심장 전도 결손, 열성 발작 플러스 동반 전신 간질, 양성 가족성 영아 발작, 조기 영아 간질성 뇌병증11 (EIEE11), 양성 가족성 영아 간질, 신경변성, 타우병증 및 알츠하이머병을 포함하나 이에 제한되지는 않는다. 일부 경우에, 신경계 상태는 드라베 증후군이다. SCN1A에서의 돌연변이 또는 이상은 또한 발작 장애, 간질, 자폐증, 가족성 편마비성 편두통 유형 3 (FHM3), 열성 발작 플러스 동반 유전적 간질 (GEFS+), 및 특정 항발작 의약의 유효성과 연관되었다. 예를 들어, SCN1A에서의 ICS5N+5G>A 돌연변이는 항발작 약물 페니토인 및 카르바마제핀의 최대 안전 양 (용량)과 연관된다.

[0303] 특정 실시양태에서, 본 출원은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여함으로써 드라베 증후군을 갖거나 드라베 증후군이 발생할 위험이 있는 대상체를 치료하는 방법을 제공한다. 드라베 증후군은 생후 1년 내에 장기간의 열성 및 비-열성 발작을 특징으로 한다. 이 질환은 근간대성 및 부분 발작, 정신운동 지연 및 운동실조와 같은 다른 발작 유형으로 진행된다. 이는 인지 장애, 행동 장애 및 운동 결핍을 특징으로 한다. 행동 결핍은 종종 과잉행동 및 충동성 및 보다 드문 경우에는 자폐-유사 행동을 포함한다. 드라베 증후군은 또한 졸림 및 불면증을 포함한 수면 장애와 연관된다. 많은 환자에서, 드라베 증후군은 비-기능적 단백질의 생산을 유도하는 유전자 돌연변이에 의해 유발된다. 유전적 원인과 연관된 장애를 치료하는데 많은 난제가 존재한다. 따라서, 대부분의 기존 치료는 발작 및 다른 증상의 예방적 의학적 관리를 위한 것이었다.

[0304] 환자의 70-90%에서, 드라베 증후군은 SCN1A 유전자 내의 넌센스 돌연변이가 조기 정지 코돈 및 이에 따른 비-기능적 단백질을 생성함으로써 유발된다. 전형적으로, 나트륨 채널 세공의 S5 또는 S6 절편에서의 미스센스 돌연변이는 채널 기능의 상실 및 드라베 증후군의 발생을 유발한다. SCN1A 돌연변이, 예를 들어 넌센스 돌연변이, 미스센스 돌연변이, 결실, 삽입, 역위 등의 이형접합 유전은 모두 결손 나트륨 채널을 발생시키는데 필요한 것이고; 드라베 증후군을 갖는 환자는 여전히 1개의 정상 유전자 카피를 가질 것이다. 따라서, 질환은 반수체기능부전 중 하나를 특징으로 하고, 이에 따라 SCN1A의 기능적 카피의 발현을 증가시키는 것은 Nav1.1의 정상 생

산 수준을 회복시킬 수 있다.

[0305] 드라베 증후군과 연관된 증상은 발작, 기억 결함, 발달 지연, 불량한 근긴장도 및/또는 인지 문제를 포함한다. 본원에 기재된 발현 카세트, 발현 벡터 또는 바이러스 입자를 사용한 치료는 1종 이상의 증상의 개선, 예컨대 발작의 수, 지속기간 및/또는 강도의 감소를 유발할 수 있다. 드라베 증후군이 발생할 위험이 있는 대상체에게 본원에 기재된 바와 같은 유전자 요법을 투여하는 것은 드라베의 1종 이상의 증상의 발달을 예방하거나 또는 그의 진행을 늦출 수 있다.

[0306] 특정 실시양태에서, 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 사용한 치료는 발작 지속기간 및/또는 빈도, 예를 들어 드라베 증후군과 연관된 발작을, 비치료 대조군과 비교하여 또는 치료 전 수준과 비교하여 적어도 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 그 초과만큼 감소시킨다.

[0307] 일부 알츠하이머 환자에서, 뉴런의 흥분성에 영향을 미칠 수 있는 많은 펩티드 및 프로테아제를 수반하는 아밀로이드 β (Aβ)의 생산은 PV 뉴런에서 발작 및 Nav1.1 나트륨 채널의 하향조절을 유발한다. 또 다른 실시양태에서, 본 출원은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 본원에 기재된 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여함으로써 알츠하이머병을 앓고 있는 대상체를 치료하는 방법을 제공한다. 알츠하이머병과 연관된 증상은 단기 기억 상실, 인지 곤란, 발작, 및 언어 곤란, 실행 기능 곤란, 지각 곤란 (실인증) 및 운동 실행 곤란 (실행증)을 포함한다. SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 사용한 치료는 1종 이상의 알츠하이머병 증상의 개선, 예컨대 기억 상실의 진행의 감소 또는 1종 이상의 증상의 예방을 가져올 수 있다. 일부 경우에, 치료는 높은 감마 파워 뇌 활성의 교정을 가져올 수 있다. 치료는 치료 부재와 비교하여 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70% 또는 그 초과만큼 발작 빈도 및/또는 발작 중증도의 감소 또는 높은 감마 파워 활성의 감소를 유발할 수 있다. 일부 경우에, 치료는 인지 기능의 개선을 가져올 수 있다. 학습 및/또는 기억은 본원에 개시된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 사용한 치료 부재 또는 치료 전과 비교하여 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% 또는 100% 초과만큼 개선될 수 있다.

[0308] 일부 경우에, SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 사용한 치료는 높은 감마 파워 활성 (예를 들어, 알츠하이머병과 연관된 높은 감마 파워 활성)을, 비치료 대조군과 비교하여 또는 치료 전 수준과 비교하여 적어도 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% 또는 95%만큼 감소시킨다.

[0309] 파킨슨증은 느림 (운동완만), 강직 (강성), 진전 및 불균형 (자세 불안정)을 포함한, 파킨슨병 (PD)에서 발견되는 징후 및 증상의 집합을 지칭한다. 일부 경우에, 파킨슨병이 발생할 위험이 있거나 또는 파킨슨병을 앓고 있는 대상체에게 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것은 그의 1종 이상의 증상의 발생을 예방할 수 있거나 또는 파킨슨병의 진행을 치료 부재와 비교하여 적어도 10%, 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80% 또는 적어도 90%만큼 늦출 수 있다.

[0310] 특정 실시양태에서, 본 출원은 질환이 발생할 위험이 있는 대상체를 치료하는데 사용될 수 있는 방법을 제공한다. 대상체는 질환, 예를 들어 신경계 질환 또는 간질, 발작 및/또는 뇌병증과 연관된 질환에 걸리기 쉬운 것으로 알려져 있을 수 있다. 대상체는 유전적 사건으로 인해 또는 공지된 위험 인자로 인해 질환에 걸리기 쉬울 수 있다. 예를 들어, 대상체는 간질과 연관된 SCN1A에서의 돌연변이 (예컨대, 예를 들어 드라베 증후군)를 보유할 수 있다. SCN1A 유전자 내의 삽입, 결실, 역위, 전위 또는 치환 (예를 들어, 넌센스 돌연변이 및/또는 미스센스 돌연변이를 포함한 점 돌연변이) 중 어느 1개 이상을 포함한, (발현 수준을 감소시키거나, 단백질 기능을 손상시키거나 또는 둘 다의 조합에 의해) 그의 활성을 감소시키는 SCN1A 유전자 내의 임의의 돌연변이로 인해 대상체는 질환에 걸리기 쉬울 수 있다. 일부 경우에, 대상체는 대상체의 연령으로 인해 알츠하이머병과 같은 질환에 걸리기 쉬울 수 있다. 일부 경우에, 대상체는 불충분한 양의 SCN1A 단백질을 가질 수 있고, SCN1A와 연관된 질환을 치료하는 것은 본원에 기재된 바와 같은 내인성 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하

는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것을 수반한다.

[0311] 특정 실시양태에서, 본원에 제공된 내인성 SCN1A를 선택적으로 상향조절하는 eTF를 코딩하는 폴리뉴클레오티드를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 사용한 치료는 드라베 또는 다른 SCN1A 연관 질환 또는 장애와 연관된 증상, 예를 들어 SCN1A 반수체기능부전과 연관된 간질의 감소 또는 중지를 유발할 수 있다. 예를 들어, 치료는 학습, 기억, 인지 기능 및/또는 운동 기능을 개선시키고/거나; 발작의 빈도 및/또는 지속기간을 감소시키고/거나; 온도 감수성을 감소시킬 수 있다 (또는 발작을 촉발하기 위한 온도 임계치를 증가시킬 수 있음).

[0312] 또 다른 측면에서, 본 출원은 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것에 의한 PV 뉴런에서의 트랜스진의 선택적 발현을 위한 방법을 제공한다. 특정 실시양태에서, 본 출원은 트랜스진 및 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것에 의한 영장류의 PV 뉴런에서의 트랜스진의 선택적 발현을 위한 방법을 제공한다. 특정 실시양태에서, 본 출원은 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소 및 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것에 의한 PV 뉴런에서의 트랜스진의 선택적 발현을 위한 방법을 제공한다. 예시적 실시양태에서, 트랜스진은 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF 중 어느 것을 코딩하는 서열을 포함한다.

[0313] 특정 실시양태에서, 본 출원은 대상체에게 트랜스진 및 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것을 포함하는 유전자 요법을 위한 방법을 제공한다. 특정 실시양태에서, 본 출원은 대상체에게 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소 및 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것을 포함하는 유전자 요법을 위한 방법을 제공한다. 예시적 실시양태에서, 트랜스진은 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF 중 어느 것을 코딩하는 서열을 포함한다.

[0314] 특정 실시양태에서, 본 출원은 발현 카세트, 발현 벡터 또는 트랜스진 및 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 바이러스 입자를 투여하는 것을 포함하는, 질환 또는 장애를 치료하는 방법을 제공한다. 특정 실시양태에서, 본 출원은 대상체에게 트랜스진에 작동가능하게 연결된 PV 선택적 조절 요소 및 적어도 1개의 PV 선택적 마이크로RNA 결합 부위를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 투여하는 것을 포함하는, 질환 또는 장애를 치료하는 방법을 제공한다. 예시적 실시양태에서, 트랜스진은 본원에 기재된 바와 같은 SCN1A를 선택적으로 상향조절하는 eTF 중 어느 것을 코딩하는 서열을 포함한다. 특정 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 PV 선택적 조절 요소를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자는 PV 뉴런이 연루된 질환 또는 장애를 치료하는데 사용될 수 있다. 특정 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 PV 선택적 조절 요소를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자는 뉴런 상태를 치료하는데 사용된다. 치료에 적절한 뉴런 질환 또는 장애는 드라베 증후군, 알츠하이머병, 파킨슨병, 헌팅톤병, 근위축성 측삭 경화증 (ALS), 척수성 근육 위축 (SMA), 간질, 신경변성 장애, 운동 장애, 운동 장애, 기분 장애, 운동 뉴런 질환, 진행성 근육 위축 (PMA), 진행성 연수 마비, 가성구 마비, 원발성 측삭 경화증, AIDS의 신경계 결과, 발달 장애, 다발성 경화증, 신경발생 장애, 졸중, 척수 손상, 외상성 뇌 손상, 타우병증, 뉴런 과소흥분성 및/또는 발작을 포함하나 이에 제한되지는 않는다. 일부 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 PV 선택적 조절 요소를 포함하는 바이러스 벡터, 바이러스 입자 또는 제약 조성물은 정신 장애 (예를 들어, 정신분열증, 강박 장애, 중독, 우울증, 불안, 정신병); 자폐증 스펙트럼 장애 (예를 들어, 유약 X 증후군, 레트 증후군); 간질 (예를 들어, 드라베 증후군, 만성 외상성 뇌병증, 열성 발작 플러스 동반 전신 간질 (GEFS+), 간질성 뇌병증, 측두엽 간질, 초점성 간질, 결절성 경화증, SCN1A 반수체기능부전과 연관된 간질); 및/또는 신경변성 (예를 들어, 알츠하이머병, 파킨슨병)을 치료하는데 사용된다. 기능장애성 PV 뉴런과 연관된 질환, 예컨대 SCN1A 또는 Nav1.1에서의 기능 상실 돌연변이로 인한 것은 드라베 증후군, 오타하라 증후군, 간질, 조기 영아 간질성 뇌병증 6 (EIEE6), 가족성 열성 발작 3A (FEB3A), 전신 긴장성-간대성 발작 동반 난치성 소아기 간질 (ICEGTC), 편두통, 가족성 편마비 3 (FHM3), 파나이이오토풀로스 증후군, 가족성 심방 세동 13 (ATFB13), 열성 발작 플러스 동반 전신 간질 유형 1 (gefs+ 유형 1), 브루가다 증후군, 비특이적 심장 전도 결손, 열성 발작 플러스 동반 전신 간질, 양성 가족성 영아 발작, 조기 영아 간질성 뇌병증11 (EIEE11), 양성 가족성 영아 간질, 신경변성, 타우병증 및 알츠하이머병을 포함한다.

[0315] 특정 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 본원에 기재된 PV 선택적 조절 요

소를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자를 사용한 치료는 뉴런 질환 또는 장애와 연관된 증상을 개선시킨다. 예를 들어, 파킨슨병 환자는 치료에 대한 양성 반응을 나타내는 개선된 운동 기능에 대해 대증적으로 모니터링될 수 있다. 뉴런 장애가 발생할 위험이 있는 대상체에게 본원에 기재된 바와 같은 방법을 사용한 요법을 투여하는 것은 1종 이상의 증상의 발생을 예방하거나 또는 그의 진행을 늦출 수 있다.

[0316] 특정 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 PV 선택적 조절 요소를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자는 뉴런 질환, 예를 들어 SCN1A 반수체기능부전과 연관된 간질, 예컨대, 예를 들어 드라베 증후군으로 진단된 대상체를 치료하는데 사용될 수 있다. 다양한 실시양태에서, 본원에 개시된 임의의 뉴런 질환 또는 장애는 기지의 유전적 사건 (예를 들어, 관련 기술분야에 공지된 임의의 SCN1A 돌연변이)에 의해 유발되거나 또는 미지의 원인을 갖는다.

[0317] 특정 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 본원에 제공되는 PV 선택적 조절 요소를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자는 질환 또는 장애가 발생할 위험이 있는 대상체를 치료하는데 사용될 수 있다. 일부 실시양태에서, 대상체는 질환, 예를 들어 뉴런 질환 (예를 들어 SCN1A 반수체기능부전과 연관된 간질, 예컨대, 예를 들어 드라베 증후군)에 걸리기 쉬운 것으로 알려져 있을 수 있다. 일부 실시양태에서, 대상체는 유전적 사건으로 인해 또는 공지된 위험 인자로 인해 질환에 걸리기 쉬울 수 있다. 예를 들어, 대상체는 간질 또는 드라베 증후군과 연관된 SCN1A에서의 돌연변이, 예를 들어 삽입, 결실, 역위, 전위 또는 치환 (예를 들어, 넌센스 돌연변이 및/또는 미스센스 돌연변이를 포함한 점 돌연변이)을 보유할 수 있다.

[0318] 특정 실시양태에서, 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 본원에 제공된 PV 선택적 조절 요소를 포함하는 발현 카세트, 발현 벡터 또는 바이러스 입자는 질환 또는 장애와 연관된 1종 이상의 증상을 감소시키는데 사용될 수 있다. 예를 들어, 드라베 증후군과 연관된 증상은 발작, 기억 결함, 발달 지연, 불량한 근긴장도 및/또는 인지 문제를 포함한다. 본원에 제공된 트랜스진 및 PV 선택적 마이크로RNA 결합 부위 및 임의로 PV 선택적 조절 요소를 포함하는 바이러스 벡터, 바이러스 입자 또는 제약 조성물을 사용한 치료는 1종 이상의 증상의 개선, 예컨대 발작의 수, 지속기간 및/또는 강도의 감소를 유발할 수 있다.

[0319] 특정 실시양태에서, 본원에 기재된 방법은 PV 뉴런에서 트랜스진의 발현을 증가시키기 위해, 유전자 요법을 위해 또는 영장류에서 질환 또는 장애를 치료하기 위해 사용된다. 특정 실시양태에서, 영장류는 인간이다. 특정 실시양태에서, 영장류는 비-인간 영장류이다. 특정 실시양태에서, 비-인간 영장류는 구세계 원숭이, 오랑우탄, 고릴라, 침팬지, 게잡이 마카크, 레서스 마카크 또는 돼지-꼬리 마카크이다.

[0320] 용어 "대상체" 및 "개체"는 척추동물, 바람직하게는 포유동물, 보다 바람직하게는 인간을 지칭하는 것으로 본원에서 상호교환가능하게 사용된다. 본원에 기재된 방법은 질환 또는 상태의 동물 모델에서의 인간 치료, 수의학적 적용 및/또는 전임상 연구에 유용할 수 있다. 다양한 실시양태에서, 본원에 기재된 방법에 따라 치료될 수 있는 대상체는 포유동물, 예컨대, 예를 들어 마우스, 래트, 햄스터, 기니 피그, 저빌, 소, 양, 돼지, 염소, 당나귀, 말, 개, 고양이, 라마, 원숭이 (예를 들어, 구세계 원숭이, 마모셋 또는 마카크, 예컨대 레서스 마카크, 돼지-꼬리 마카크 또는 게잡이 마카크 (즉, 시노몰구스 원숭이)), 유인원 (예를 들어, 오랑우탄, 고릴라 또는 침팬지) 또는 인간이다. 예시적 실시양태에서, 대상체는 인간이다.

[0321] 하기 표는 본원에 개시된 서열을 제공한다.

[0322] 표 1. 예시적인 조작된 전사 인자가 본원에 개시된다. 조절 요소 (RE)의 서열은 하기 표 2 및 8에 개시된다. RE에 대해, m1이 나타나 있는 경우, 이는 m1 마이크로RNA 결합 부위 (서열식별번호: 7, 표 8)가 코딩 영역과 폴리A 테일 사이에 포함된다는 것을 의미한다. DNA 결합 도메인 (DBD)에 대한 서열은 하기 표 3에 개시된다. DBD의 경우, 조작된 아연 핑거 (eZF)는 구축물이 서열식별번호: 147에 제시된 식 (표 10)을 갖는다는 것을 나타내고; EGR1은 DBD가 야생형 인간 EGR1 (서열식별번호: 176; 표 12)로부터 유래된다는 것을 나타내며; EGR3은 DBD가 야생형 인간 EGR3 (서열식별번호: 175, 표 12)으로부터 유래된다는 것을 나타낸다. 표적 부위에 대한 서열 (예를 들어, DBD에 의해 결합된 서열)은 하기 표 4에 제공된다. 전사 활성화 도메인 (TAD)에 대한 서열은 하기 표 5에 개시된다. TAD의 경우, (c)는 TAD가 DBD의 c-말단에 위치한다는 것을 나타내고, (n)은 TAD가 DBD의 n-말단에 위치한다는 것을 나타내고, (n/c)는 TAD가 DBD의 n-말단 및 c-말단 둘 다에 위치한다는 것을 의미하고, 2x CITED4 (n)는 CITED4 TAD의 2 카피가 DBD의 n-말단에 위치한다는 것을 나타낸다. 전장 조작된 전사 인자 (DBD + TAD)에 대한 서열은 하기 표 6에 제공된다.

[0323]

구축물	RE	DBD	표적 부위	TAD (위치)	SEQ ID NO (DBD)	SEQ ID NO (DBD + TAD)
1	RE 1	eZF	Z13	VPR (c)	89	99
2	RE 1	eZF	Z1	VPR (c)	77	100
3	RE 1	eZF	Z13	VP64 (c)	89	101
4	RE 1	eZF	Z1	VP64 (c)	77	102
5	RE 1	EGR1	Z13	CITED4 (n/c)	93	103
6	RE 1	EGR1	Z13	CITED4 (n)	93	104
7	RE 1	EGR1	Z1	CITED4 (n/c)	92	105
8	RE 1	EGR1	Z1	CITED4 (n)	92	106
9	RE 1	EGR1	Z1	CITED4 (c)	92	107
10	RE 1	EGR1	Z1	CITED2 (c)	92	108
11	RE 1	EGR1	Z1	CITED2 (n)	92	109
12	RE 1	EGR3	Z1	CITED4 (n/c)	96	110
13	RE 1	EGR3	Z1	CITED4 (c)	96	111
14	RE 1	EGR3	Z1	CITED2 (c)	96	112
15	RE 1	EGR3	Z1	CITED2 (n)	96	113
16	CBA	EGR3	Z15	N/A	98	114
17	RE 1	EGR1	Z13	N/A	93	115
18	RE 1	EGR1	Z15	N/A	94	116
19	RE 1	EGR1	Z13	N/A	93	117
20	RE 1	EGR1	Z13	N/A	93	115
21	RE 1	EGR1	Z1	N/A	92	118
22	CBA	EGR3	Z13	N/A	97	119
23	RE 1	EGR1	Z17	N/A	95	120
24	RE 1	EGR1	Z13	N/A	93	121
25	CBA	EGR3	Z1	N/A	96	122
26	RE 1	EGR1	Z1	N/A	92	123
27	RE 1	EGR1	Z1	N/A	92	124
28	RE 1	eZF	Z8	VP64 (c)	84	125
29	RE 1	eZF	Z14	VP64 (c)	90	126
30	CBA	eZF	Z13	VPR (c)	89	99
31	RE 2 (m1)	eZF	Z1	VP64 (c)	77	102
32	RE 2	eZF	Z1	VP64 (c)	77	102
33	RE 2	eZF	Z1	VPR (c)	77	100
34	RE 2	eZF	Z1	VP64 (c)	77	127
35	RE 2 (m1)	eZF	Z1	VP64 (c)	77	127
36	RE 2	EGR1	Z1	CITED4 (n)	92	128
37	RE 2 (m1)	EGR1	Z1	CITED4 (n)	92	106
38	RE 2 (m1)	EGR1	Z1	CITED4 (n)	92	129
39	RE 2 (m1)	EGR1	Z1	CITED4 (n/c)	92	105
40	RE 2 (m1)	EGR1	Z1	2x CITED4 (n)	92	130
41	RE 2 (m1)	EGR1	Z1	2x CITED4 (n)	92	131
42	RE 2 (m2)	eZF	Z1	VP64 (c)	77	102
43	RE 2 (m3)	eZF	Z1	VP64 (c)	77	102
44	RE 2 (m2)	EGR1	Z1	CITED4 (n)	92	106

[0324]

구축물	RE	DBD	표적 부위	TAD (위치)	SEQ ID NO (DBD)	SEQ ID NO (DBD + TAD)
45	RE 2 (m3)	EGR1	Z1	CITED4 (n)	92	106
46	RE 1	EGR1	Z1	CITED4 (n)	92	205
47	RE 1	EGR1	Z1	2x CITED4 (n)	92	207
48	RE 1	EGR1	Z1	2x CITED4 (n)	92	209
49	CBA	eZF	Z1	CREB3 (n)	77	213
50	CBA	EGR1	Z1	CREB3 (n)	92 (C-말단 Lys 부재)	217
51	CBA	EGR1	Z13	CREB3 (n)	93 (C-말단 Lys 부재)	219
52	CBA	eZF	Z1	CREB3 (n); (c)에 TM 도메인 부재	77	221
53	CBA	EGR1	Z1	CREB3 (n); (c)에 TM 도메인 부재	92 (C-말단 Lys 부재)	223

[0325] 표 2. 본원에 개시된 다양한 조절 요소 (RE)에 대한 핵산 서열.

RE	서열	SEQ ID NO
RE 1	GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAG CGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTG ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTGGTACCGTGTGTATGCTCAGGGGCTGGGAAAG GAGGGGAGGGAGCTCCGGCTCAG	1
RE 2	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaacatatttt gaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaagggcgatcaacact atggtttgaaaaggggaaattgtagagcacagatgtgttcgtgtggcagtgtgctgtctctagc aatactcagagaagagagagaacaatgaaattctgattggccccagtgtgagcccagatgagg ttcagctgccaactttctctttcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttt tttttgagacagagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcac tgcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagctggaattac aggagtggcccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcac ccaggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagtgctggg attataggcgtcagccactatgcccaacccgaccaacctttttaaaataaatatttaaaaaattgg tatttcacatatatactagtatttacatttatccacacaaaacggacgggcctccgctgaaccagtg aggccccagacgtgcgcataaataaccctgcgtgctgcaccacctggggagaggggagg accacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttctctgccggtgg cactgggtagctgtggccaggtgtggtactttgatggggcccagggctggagctcaaggaag cgtcgcagggtcacagatctgggggaaccccggggaaaagcactgaggcaaaaccgccgc tcgtctcctacaatatatgggaggggggaggttgagtacgttctggattactcataagacctttttttt ttccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagaggcggtc aggcacctgcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctaccccggagccgt gcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcgagagggaactagcgaga acgaggaagcagctggaggtgacgccgggcagattacgcctgtcagggccgagccgagcg gatcgctgggcgctgtgcagaggaaaggcgggagtgcccggctcgctgtcgcagagccga ggtgggtaagctagcgaccacctggacttcccagcgcccaaccgtggcttttcagccaggtcc tctcctcccgcggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttccaggggc cgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggctagagtgcaaggtg actgtggttcttctctggccaagtccgagggagaacgtaaagatatgggccttttccccctctca ccttgtctcaccaaagtccctagtccccggagcagttagcctctttctttccagggaattagccag acacaacaacgggaaccagacaccgaaccagacatgcccgccccgtgcgccctccccgctc gctgcctttcctccctcttgtctctccagagccggatcttcaaggggagcctccgtgccccggc tgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgc agcctgtttctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccga gctg	2

[0326]

[0327]

RE	서열	SEQ ID NO
RE 3	TCAACAGGGGGACACTTGGGAAAGAAGGATGGGGACAG AGCCGAGAGGACTGTTACACATTAGAGAAACATCAGTGA CTGTGCCAGCTTTGGGGTAGACTGCACAAAAGCCCTGAG GCAGCACAGGCAGGATCCAGTCTGCTGGTCCCAGGAAGC TAACCGTCTCAGACAGAGCACAAAGCACCGAGACATGTG CCACAAGGCTTGTGTAGAGAGGTCAGAGGACAGCGTACA GGTCCCAGAGATCAAACTCAACCTCACCAGGCTTGGCAG CAAGCCTTTACCAACCCACCCCCACCCCACCCACCCTGCA CGCGCCCCTCTCCCCTCCCCATGGTCTCCCATGGCTATCT CACTTGGCCCTAAAATGTTTAAGGATGACACTGGCTGCTG AGTGGAAATGAGACAGCAGAAGTCAACAGTAGATTTTAG GAAAGCCAGAGAAAAAGGCTTGTGCTGTTTTTAGAAAGC CAAGGGACAAGCTAAGATAGGGCCCAAGTAATGCTAGTA TTTACATTTATCCACACAAAACGGACGGGCCTCCGCTGAA CCAGTGAGGCCCCAGACGTGCGCATAAATAACCCCTGCG TGCTGCACCACCTGGGGAGAGGGGGAGGACCACGGTAAA TGGAGCGAGCGCATAGCAAAAGGGACGCGGGGTCCTTTT CTCTGCCGGTGGCACTGGGTAGCTGTGGCCAGGTGTGGT ACTTTGATGGGGCCCAGGGCTGGAGCTCAAGGAAGCGTC GCAGGGTCACAGATCTGGGGGAACCCCGGGGAAAAGCA CTGAGGCAAAACCGCCGCTCGTCTCCTACAATATATGGG AGGGGGAGGTTGAGTACGTTCTGGATTACTCATAAGACC TTTTTTTTTTCCTTCCGGGCGCAAAACCGTGAGCTGGATTT ATAATCGCCCTATAAAGCTCCAGAGGCGGTCAGGCACCT GCAGAGGAGCCCCGCCGCTCCGCCGACTAGCTGCCCCCG CGAGCAACGGCCTCGTGATTTCCCCGCCGATCCGGTCCCC GCCTCCCCACTCTGCCCCCGCCTACCCCGGAGCCGTGCAG CCGCCTCTCCGAATCTCTCTCTTCTCCTGGCGCTCGCGTGC GAGAGGGAACTAGCGAGAACGAGGAAGCAGCTGGAGGT GACGCCGGGCAGATTACGCCTGTCAGGGCCGAGCCGAGC GGATCGCTGGGCGCTGTGCAGAGGAAAGGCGGGAGTGCC CGGCTCGCTGTCGCAGAGCCGAGGTGGGTAAGCTAGCGA CCACCTGGACTTCCCAGCGCCCAACCGTGGCTTTTCAGCC AGGTCCTCTCCTCCCGCGGCTTCTCAACCAACCCCATCCC AGCGCCGGCCACCCAACCTCCCGAAATGAGTGCTTCCTG CCCCAGCAGCCGAAGGCGCTACTAGGAACGGTAACCTGT TACTTTTCCAGGGGCCGTAGTCGACCCGCTGCCCGAGTTG CTGTGCGACTGCGCGCGCGGGGCTAGAGTGCAAGGTGAC TGTGGTTCTTCTCTGGCCAAGTCCGAGGGAGAACGTAAA GATATGGGCCTTTTTCCCCCTCTCACCTTGTCTCACCAAA GTCCCTAGTCCCCGGAGCAGTTAGCCTCTTTCTTTCCAGG GAATTAGCCAGACACAACAACGGGAACCAGACACCGAA	3

[0328]

RE	서열	SEQ ID NO
	CCAGACATGCCCGCCCCGTGCGCCCTCCCCGCTCGCTGCC TTTCCTCCCTCTTGTCTCTCCAGAGCCGGATCTTCAAGGG GAGCCTCCGTGCCCCCGGCTGCTCAGTCCCTCCGGTGTGC AGGACCCCGGAAGTCCTCCCCGCACAGCTCTCGCTTCTCT TTGCAGCCTGTTTCTGCGCCGGACCAGTCGAGGACTCTGG ACAGTAGAGGCCCCGGGACGACCGAGCTG	
RE 4	GCCCTCTAGGCCACCTGACCAGGTCCCCTCAGTCCCCCCC TTCCCACACTCCCACACTCAGCCCCCCTCCCCCCCCCCCG ACCCCTGCAGGATTATCCTGTCTGTGTTCCTGACTCAGCC TGGGAGCCACCTGGGCAGCAGGGGCCAAGGGTGTCCTAG AAGGGACCTGGAGTCCACGCTGGGCCAAGCCTGCCCTTT CTCCCTCTGTCTTCCGTCCCTGCTTGCGGTTCTGCTGAATG TGGTTATTTCTCTGGCTCCTTTTACAGAGAATGCTGCTGCT AATTTTATGTGGAGCTCTGAGGCAGTGTAATTGGAAGCC AGACACCCTGTCAGCAGTGGGCTCCCGTCCTGAGCTGCC ATGCTTCCTGCTCTCCTCCCGTCCCGGCTCCTCATTTCATG CAGCCACCTGTCCCAGGGAGAGAGGAGTCACCCAGGCCC CTCAGTCCGCCCCTTAAATAAGAAAGCCTCCGTTGCTCGG CACACATACCAAGCAGCCGCTGGTGCAATCT	4
CBA (CMV 인핸서 + 닭 베타 액틴 프로모터)	CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAG TACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGA CGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA GTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGggtcgaggtgagccccacgtt ctgcttcactctccccatctccccccctccccaccccaattttgtatttatttatttttttaattattttg tgcagcgatgggggcgggggggggggggcgcgcgccaggcggggcggggcggggcg aggggcggggcggggcgaggcggagaggtgcggcggcagccaatcagagcggcgcgct ccgaaagtttccttttatggcgaggcggcggcggcggccctataaaaagcgaagcgcgc ggcgggcgggagtcgctgcgttgccttcgccccgtgccccgctccgcgccgcctcgcgccg cccgccccggctctgactgaccgcgttactcccacaggtgagcgggcgggacggcccttctc ctccgggctgtaattagcgcttggtttaatgacggctcgtttcttttctgtggctgcgtgaaagcctt aaagggctccgggagggcccttttgtgcgggggggagcggctcgggggggtgcgtgcgtgtgt gtgtgcgtggggagcgccgcgtgcggcccgcgctgcccggcggctgtgagcgctgcgggc gcggcgcggggctttgtgcgctccgcgtgtgcgcgaggggagcgcggccgggggcggtgc cccgcggtgcgggggggctgcgagggggaacaaaggctgcgtgcggggtgtgtgcgtgggg gggtgagcaggggggtgtgggcgcggcggtcgggctgtaaccccccccctgcaccccccctccc cgagttgctgagcacggcccggcttcgggtgcggggctccgtgcggggcgtggcgcggggg ctcgccgtgccgggcgggggggtggcggcaggtgggggtgccgggcggggcggggccgcc ctcgggccggggagggctcgggggagggggcgcggcggccccggagcgccggcggctgt	5

RE	서열	SEQ ID NO
	cgaggcgcggcgagccgcagccattgccttttatggtaatcgtgcgagagggcgcagggact tcctttgtcccaaatctggcggagccgaaatctgggaggcgccgccgcaccccctctagcgggg cgcgggcgaagcggtgcggcgccggcaggaaggaaatgggcggggagggccttcgtgcg tcgccgcgccgccgtcccctctccatctccagcctcggggctgccgcaggggacggctgc cttcgggggggacggggcagggcggggttcggcttctggcgtgtgaccggcggctctagag cctctgctaaccatgttcatgccttcttcttttttcctacagctcctgggcaacgtgctggttgttgtgc tgtctcatcattttggcaaagaatt	
EF1알파	GAGTAATTCATACAAAAGGACTCGCCCCTGCCTTGGGGA ATCCCAGGGACCGTCGTTAAACTCCCACTAACGTAGAAC CCAGAGATCGCTGCGTTCCCGCCCCCTCACCCGCCCGCTC TCGTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTC CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGT CCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGG TGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGA TGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGA GAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTT TTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCG TGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGG CCCTTGCGTGCCTTGAATTACTTCCACGCCCCTGGCTGCA GTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGG TGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCTTCG CCTCGTGCTTGAGTTGAGGCCTGGCTTGGGCGCTGGGGCC GCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGC TGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGA CCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAA ATGCGGGCCAAGATCTGCACACTGGTATTTCGGTTTTTGG GGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCAC ATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAG AATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCT GGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTGG GCGGCAAGGCTGGCCCGGTCGGCACCAGTTGCGTGAGCG GAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCA AAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGA GTCACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGC CGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCC AGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGT CTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCC CCACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTG GCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGT TTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCA AAGTTTTTTTCTTCCATTTCAGGTGTCGTGA	6

[0329]

[0330] 표 3. 본원에 제공된 예시적인 DNA 결합 도메인 (DBD)에 대한 아미노산 서열. DBD의 경우, 조작된 아연 핑거 (eZF)는 구축물이 서열식별번호: 147에 제시된 식 (표 10)을 갖는다는 것을 나타내고; EGR1은 DBD가 야생형 인간 EGR1 (서열식별번호: 176; 표 12)로부터 유래된다는 것을 나타내며; EGR3은 DBD가 야생형 인간 EGR3 (서열식별번호: 175, 표 12)으로부터 유래된다는 것을 나타낸다. 표적 부위는 DBD에 의해 결합된 서열이고, 하기 표 4에 제공된다.

[0331]

DBD/표적 부위	서열	SEQ ID NO
eZF/z1	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS	77
eZF/z2	LEPGEKPYKCPECGKSFSTKNSLTEHQRTHTGEKPYKCPECGKSFSRADNLTEHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSTKNSLTEHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGKKTS	78
eZF/z3	LEPGEKPYKCPECGKSFSQAGHLASHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSTSGNLTEHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSQKSSLIAHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGKKTS	79
eZF/z4	LEPGEKPYKCPECGKSFSTTGNLTVHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSTSGNLTEHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSQRANLRAHQRTHTGKKTS	80
eZF/z5	LEPGEKPYKCPECGKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSTTGALTEHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGKKTS	81
eZF/z6	LEPGEKPYKCPECGKSFSRKDNLKNHQRTHTGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGEKPYKCPECGKSFSRKDNLKNHQRTHTGKKTS	82
eZF/z7	LEPGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS	83
eZF/z8	LEPGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTHTGKKTS	84
eZF/z9	LEPGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGKKTS	85
eZF/z10	LEPGEKPYKCPECGKSFSDCRDLARHQRTHTGEKPYKCPEC	86

[0332]

DBD/표적 부위	서열	SEQ ID NO
	GKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSRNDALTE HQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYK CPECGKSFSDPGNLVRHQRTHTGEKPYKCPECGKSFSQRA HLERHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEK PYKCPECGKSFSHRTTLTNHQRTHTGKKTS	
eZF/z11	LEPGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPEC GKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDPGNLVRH QRTHTGEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKC PECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSHRTTL TNHQRTHTGKKTS	87
eZF/z12	LEPGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPEC GKSFSDPGHLVRHQRTHTGEKPYKCPECGKSFSTSGELVRH QRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCP ECGKSFSSKKALTEHQRTHTGEKPYKCPECGKSFSQLAHLR AHQRTHTGEKPYKCPECGKSFSRSDHLTNHQRTHTGKKTS	88
eZF/z13	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPEC GKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFSREDNLHTH QRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCP ECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSREDNLH THQRTHTGKKTS	89
eZF/z14	LEPGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPEC GKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSQSGDLRR HQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKC PECGKSFSTSGNLVRHQRTHTGEKPYKCPECGKSFSRSDNL VRHQRTHTGKKTS	90
eZF/z15	LEPGEKPYKCPECGKSFSRRDELNVHQRTHTGEKPYKCPEC GKSFSRSDHLTNHQRTHTGEKPYKCPECGKSFSRSDDLVRH QRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKC PECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFSREDNL HTHQRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGEKPY KCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRED NLHTHQRTHTGKKTS	91
EGR1/z1	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SREDNLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHL RQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEH TKIHLRQKDK	92
EGR1/z13	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHL RQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRIC MRNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNLHTH TKIHLRQKDK	93
EGR1/z15	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRN FSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKIH LRQKDRPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHT HTKIHLRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKP FQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFARED NLHTHTKIHLRQKD	94

DBD/표적 부위	서열	SEQ ID NO
EGr1/z17	RPYACPVESCDRRFSDPGALVRHIRIHTGQKPFQCRICMRN FSRSDNLVRHIRTHTGEKPFACDICGRKFAQSGDLRRHTKI HLRQKDRPYACPVESCDRRFSTHLDLIRHIRIHTGQKPFQCR ICMRNFSTSGNLVRHIRTHTGEKPFACDICGRKFARSDNLV RHTKIHLRQKDRPYACPVESCDRRFSQSGHLTEHIRIHTGQ KPFQCRICMRNFSERSHLREHIRTHTGEKPFACDICGRKFAQ AGHLASHTKIHLRQKD	95
EGR3/z1	RPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRS FSREDNLHTHIRTHTGEKPFACEFCGRKFARSDELVRHAKI HLKQKEHACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRI CMRSFSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTLTE HAKIHLKQKEK	96
EGR3/z13	RPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRS FSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLHTHAKI HLKQKEHACPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRI CMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDNLHT HAKIHLKQKEK	97
EGR3/z15	RPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRS FSRSDHLTNHIRTHTGEKPFACEFCGRKFARSDDLVRHAKI HLKQKEHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRI CMRSFSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLHT HAKIHLKQKEHACPAEGCDRRFSTSHSLTEHLRIHTGHKPF QCRICMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDN LHTHAKIHLKQKEK	98

[0333]

[0334] 표 4. 본원에 개시된 DNA 결합 도메인에 의해 결합된 예시적인 표적 부위에 대한 표적 부위 서열 및 염색체 위치.

Chr 2 시작 위치	표적 부위 서열	표적 부위	표적 부위 서열에 대한 SEQ ID NO
166149168	CTAGGTCAAGTGTAGGAG	z1	18
166149158	ACTTGACCTAGACAGCCT	z2	19
166073978	TGAATAACTCATTAGTGA	z3	20
166073933	AAAGTACATTAGGCTAAT	z4	21
166149199	CCAGCACTGGTGCTTCGT	z5	22
166149176	AAGGCTGTCTAGGTCAAG	z6	23
166149168	CTAGGTCAAGTGTAGGAGACACAC	z7	24
166149165	GGTCAAGTGTAGGAGACA	z8	25
166149162	CAAGTGTAGGAGACACAC	z9	26
166149160	AGTGTAGGAGACACACTGCTGGCC	z10	27
166149160	AGTGTAGGAGACACACTG	z11	28
166149155	AGGAGACACACTGCTGGCCTG	z12	29
166128025	TAGGTACCATAGAGTGAG	z13	30
166127991	GAGGATACTGCAGAGGTC	z14	31
166127999	TAGGTACCATAGAGTGAGGCGAGGATG	z15	32
166127991	ATAGAGTGAGGCGAGGATGAAGCCGAG	z16	33
166127974	TGAAGCCGAGAGGATACTGCAGAGGTC	z17	34

[0335]

[0336] 표 5. 본원에 개시된 예시적인 전사 활성화 도메인 (TAD)에 대한 아미노산 서열.

TAD	서열	SEQ ID NO
VPR	DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDMLINSRSSGSPKKKRKVGSQYLPDTDDRHRIEEKRK RTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPAPQPY PFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPA PAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLS EALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQ GIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLP NGLLSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGS AISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVH EPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALR EMADTVIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTEDL NLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDTSLF	132
VP64	DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDML	133
CITED2	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFPS PHHHQQQQPQHAFNALMGEHIHYGAGNMNATSGIRHAMGPG TVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKL NNQYFNHHPYPHNHYMPDLHPAAGHQMNGTNQHFRDCNPKH SGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNSGGGSGSGNMPAS VAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELWL GQNEFDFMTDFVCKQQPSRVSC	134
CITED4	ADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSG LRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAH LQPVATPYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGG MDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGS APPAGSVSC	135
CREB3	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWALPLSE VPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHDHTYSLPRET VSMDLESESCRKEGTQMTPQHMEELAEQEIARLVLTDEEKSLL EKEGLILPETLPLTKTEEQILKRVR	224

[0337]

[0338] 표 6. 본원에 개시된 예시적인 조작된 전사 인자 (DBD + TAD)에 대한 아미노산 서열.

구축물	서열	SEQ ID NO
1	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNH QRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPEC GKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRED NLHTHQRTHTGKKTSKRPAATKKAGQAKKKKGSY PYDVPDYALEEASGSGRADALDDFDLDMLGSDALD DFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN SRSSGSPKKKRKVGSQYLPDTDDRHRIEEKRKRTYE TFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPA PQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAP PQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLG NSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPDEETSQAVKALREMADTVIPQKEEAAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTP ELNEILDTFLNDECLLHAMHISTGLSIFDTSLF	99
2	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEEASGSGRADALDDFDLDMLGSDALD DFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN SRSSGSPKKKRKVGSQYLPDTDDRHRIEEKRKRTYE TFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPA PQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAP PQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLG NSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPDEETSQAVKALREMADTVIPQKEEAAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTP ELNEILDTFLNDECLLHAMHISTGLSIFDTSLF	100
3	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNH QRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPEC GKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRED NLHTHQRTHTGKKTSKRPAATKKAGQAKKKKGSY PYDVPDYALEDALDDFDLDMLGSDALDDFDLDML	101

[0339]

[0340]

구축물	서열	SEQ ID NO
	GSDALDDFDLDMLGSDALDDFDLDML	
4	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDML	102
5	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFARE DNLHTHTKIHLRQKDRPYACPVESCDRRFSTSHSLT EHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGE KPFACDICGRKFAREDNLHTHTKIHLRQKDKLEMA DHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAG PGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQ PFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGP PGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLE LELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS C	103
6	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFARE DNLHTHTKIHLRQKDRPYACPVESCDRRFSTSHSLT EHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGE KPFACDICGRKFAREDNLHTHTKIHLRQKDK	104
7	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDKLEMAD HLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGP GLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQP	105

[0341]

구축물	서열	SEQ ID NO
	FPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLEL ELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	
8	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
9	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCD RRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNL HTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPF QCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKF AQNSTLTEHTKIHLRQKDKLEMADHLMLAEGYRLV QRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRGAP LGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIA HLQPVATPYPGRAAAPPNAPGGPPGPQPAPSAAAPP PPAHALGGMDAELIDEEALTSLELELGLHRVRELPE LFLGQSEFDCFSDLGSAPPAGSVSC	107
10	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCD RRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNL HTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPF QCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKF AQNSTLTEHTKIHLRQKDKLEMSGLEMADHMMAM NHGRFPDGTNGLHHHPAHRMGMGQFPSPHHHQQQ QPQHAFNALMGEHIHYGAGNMNATSGIRHAMGPG TVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLP ASMQLQKLNNQYFNHHPYPHNHYMPDLHPAAGHQ MNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSG SSSGGGAGSSNSGGGSGSGNMPASVAHVPAAMLPP NVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQN EFDFMTDFVCKQQPSRVSC	108
11	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRM GMGQFPSPHHHQQQQPQHAFNALMGEHIHYGAGN MNATSGVRHAMGPGTVNGGHPPSALAPAARFNNS QFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPH NHYMPDLHPAAGHQMNGTNQHFRDCNPKHSGGSS TPGGSGGSSTPGGSGSSSGGGAGSSNSGGGSGSGNM PASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGL DRIKELPELWLGQNEFDFMTDFVCKQQPSRVSCQSQ LIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSR SDNLVRHIRIHTGQKPFQCRICMRNFSREDNLHTHIR THTGEKPFACDICGRKFARSDELVRHTKIHLRQKDR	109

[0342]

구축물	서열	SEQ ID NO
	PYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNS TLTEHTKIHLRQKDK	
12	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGRPHACPAEGCDRRFSRSDNLVRH LRIHTGHKPFQCRICMRSFSREDNLHTHIRTHTGEKP FACEFCGRKFARSDELVRHAKIHLKQKEHACPAEG CDRRFSQSGNLTEHLRIHTGHKPFQCRICMRSFSTSG HLVRHIRTHTGEKPFACEFCGRKFAQNSTLTEHAKI HLKQKEKLEMADHLMLAEGYRLVQRPPSAAAAHG PHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGA LAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPG RAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMD AELIDEEALTSLELELGLHRVRELPELFLGQSEFDCF SDLGSAPPAGSVSC	110
13	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQC RICMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFAR SDELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTE HLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEK PFACEFCGRKFAQNSTLTEHAKIHLKQKEKLEMAD HLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGP GLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQP FPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLEL ELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	111
14	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQC RICMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFAR SDELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTE HLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEK PFACEFCGRKFAQNSTLTEHAKIHLKQKEKKAEKG GAPSASSAPPVSLAPVVTTCALEMSGLEMADHMMA MNHGRFPDGTNGLHHHPAHRMGMGQFPSPHHHQQ QQPQHAFNALMGEHIHYGAGNMNATSGIRHAMGP GTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSL PASMQLQKLNNQYFNHHPYPHNHYMPDLHPAAGH QMNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGS GSSSGGGAGSSNSGGGSGSGNMPASVAHVPAAMLP PNVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQN EFDFMTDFVCKQQPSRVSC	112
15	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRM GMGQFPSPHHHQQQQPQHAFNALMGEHIHYGAGN MNATSGVRHAMGPGTVNGGHPPSALAPAARFNNS QFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPH NHYMPDLHPAAGHQMNGTNQHFRDCNPKHSGGSS TPGGSGGSSTPGGSGSSSGGGAGSSNSGGGSGSGNM PASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGL	113

[0343]

구축물	서열	SEQ ID NO
	DRIKELPELWLGQNEFDFMTDFVCKQQPSRVSCRPH ACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICM RSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSDEL VRHAKIHLKQKEHACPAEGCDRRFSQSGNLTEHLRI HTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPFA CEFCGRKFAQNSTLTEHAKIHLKQKEKKAEKGGAP SASSAPPVSLAPVVTTCA	
16	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLF SGSSDSVVHYNQMATENVMDIGLTNEKPNPELSYS GSFQPAPGNKTVTYLGKFAFDSPSNWCQDNIISLMS AGILGVPPASGALSTQTSTASMVQPPQGDVEAMYP ALPPYSNCGDLYSEPVSFHDPQGNPGLAYSPQDYQS AKPALDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQG MDPIRVNPPPITPLETIKAFKDKQIHPGFGSLPQPPLT LKPIRPRKYPNRPSKTPLHERPHACPAEGCDRRFSRR DELNVHLRIHTGHKPFQCRICMRSFSRSDHLTNHIRT HTGEKPFACEFCGRKFARSDDLVRHAKIHLKQKEH ACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICM RSFSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNL HTHAKIHLKQKEHACPAEGCDRRFSTSHSLTEHLRI HTGHKPFQCRICMRSFSQSSSLVRHIRTHTGEKPFAC EFCGRKFAREDNLHTHAKIHLKQKEKKAEKGGAPS ASSAPPVSLAPVVTTCA	114
17	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRK FAREDNLHTHTKIHLRQKDRPYACPVESCDRRFSTS HSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTH TGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKA DKSVVASSATSSLSSYPSPVATSYPSPVTTSYPSPATT SYPSPVPTSFSSPGSSTYPSPVHSGFPSPSVATTYSSV PPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTI EIC	115
18	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK	116

[0344]

구축물	서열	SEQ ID NO
	TPPHERPYACPVESCDRRFSRRDELNVHIRIHTGQKP FQCRICMRNFSRSDHLTNHIRTHTGEKPFACDICGRK FARSDDLVRHTKIHLRQKDRPYACPVESCDRRFSRS DNLVRHIRIHTGQKPFQCRICMRNFSHRTTLTNHIRT HTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDRP YACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICM RNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNL HTHTKIHLRQKDKKADKSVVASSATSSLSSYPSPVA TSYPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYPSPV HSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSA STGLSDMTATFSPRTIEIC	
19	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRK FAREDNLHTHIRTHTGEKPFACDICGRKFSTSHSLTE HIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEK PFACDICGRKFAREDNLHTHTKIHLRQKDKKADKS VVASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYP SPVPTSFSSPGSSTYPSPVHSGFPSPSVATTYSSVPPA FPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	117
20	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRK FAREDNLHTHTKIHLRQKDRPYACPVESCDRRFSTS HSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTH TGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKA DKSVVASSATSSLSSYPSPVATSYPSPVTTSYPSPATT SYPSPVPTSFSSPGSSTYPSPVHSGFPSPSVATTYSSV PPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTI EIC	115
21	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI	118

[0345]

구축물	서열	SEQ ID NO
	SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGR KFARSDELVRHTKIHLRQKDRPYACPVESCDRRFSQ SGNLTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIR THTGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDK KADKSVVASSATSSLSSYPSPVATSYPSPVTTSYPSP ATTSYPSPVPTSFSSPGSSTYPSPVHSGFPSPSVATTY SSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFS PRTIEIC	
22	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLF SGSSDSVVHYNQMATENVMDIGLTNEKPNPELSYS GSFQPAPGNKTVTYLGKFAFDSPSNWCQDNIISLMS AGILGVPPASGALSTQTSTASMVQPPQGDVEAMYP ALPPYSNCGDLYSEPVSFHDPQGNPGLAYSPQDYQS AKPALDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQG MDPIRVNPPPITPLETIKAFKDKQIHPGFGSLPQPPLT LKPIRPRKYPNRPSKTPLHERPHACPAEGCDRRFSRS DNLVRHLRIHTGHKPFQCRICMRSFSHRTTLTNHIRT HTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKEH ACPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRICMR SFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDNLH THAKIHLKQKEKKAEKGGAPSASSAPPVSLAPVVTT CA	119
23	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSDPGALVRHIRIHTGQKP FQCRICMRNFSRSDNLVRHIRTHTGEKPFACDICGR KFAQSGDLRRHTKIHLRQKDRPYACPVESCDRRFST HLDLIRHIRIHTGQKPFQCRICMRNFSTSGNLVRHIR THTGEKPFACDICGRKFARSDNLVRHTKIHLRQKDR PYACPVESCDRRFSQSGHLTEHIRIHTGQKPFQCRIC MRNFSERSHLREHIRTHTGEKPFACDICGRKFAQAG HLASHTKIHLRQKDKKADKSVVASSATSSLSSYPSP VATSYPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYP SPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNS FSASTGLSDMTATFSPRTIEIC	120
24	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL	121

[0346]

구축물	서열	SEQ ID NO
	EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLTRHIRIHTGQKP FQCRICMRNFSHSTTLTNHIRTHTGEKPFACDICGRK FARSDNRKTHIRTHTGEKPFACDICGRKFSTSHSLTE HIRIHTGQKPFQCRICMRNFSQSSSLTRHIRTHTGEKP FACDICGRKFARSDNRKTHTKIHLRQKDKKADKSV VASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYPS PVPTSFSSPGSSTYPSPVHSGFPSPSVATTYSSVPPAF PAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	
25	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLF SGSSDSVVHYNQMATENVMDIGLTNEKPNPELSYS GSFQPAPGNKTVTYLGKFAFDSPSNWCQDNIISLMS AGILGVPPASGALSTQTSTASMVQPPQGDVEAMYP ALPPYSNCGDLYSEPVSFHDPQGNPGLAYSPQDYQS AKPALDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQG MDPIRVNPPPITPLETIKAFKDKQIHPGFGSLPQPPLT LKPIRPRKYPNRPSKTPLHERPHACPAEGCDRRFSRS DNLVRHLRIHTGHKPFQCRICMRSFSREDNLHTHIRT HTGEKPFACEFCGRKFARSDELVRHAKIHLKQKEH ACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRICM RSFSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTL TEHAKIHLKQKEKKAEKGGAPSASSAPPVSLAPVVT TCA	122
26	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLTRHIRIHTGQKP FQCRICMRNFSRSDNLTTHIRTHTGEKPFACDICGRK FARSDERKRHIRTHTGEKPFACDICGRKFSQSGNLTE HIRIHTGQKPFQCRICMRNFSTSGHLTRHIRTHTGEK PFACDICGRKFAQSSTRKEHTKIHLRQKDKKADKSV VASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYPS PVPTSFSSPGSSTYPSPVHSGFPSPSVATTYSSVPPAF PAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	123
27	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI	124

[0347]

구축물	서열	SEQ ID NO
	SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQGDLGLGTPDQKPFQGLESRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGR KFARSDELVRHIRTHTGEKPFACDICGRKFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDKKADKS VVASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYP SPVPTSFSSPGSSTYPSPVHSGFPSPSVATTYSSVPPA FPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	
28	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSS PADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRH QRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPEC GKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTSG HLVRHQRTHTGKKTSKRPAATKKAGQAKKKKGSY PYDVPDYALEDALDDFDLDMLGSDALDDFDLDML GSDALDDFDLDMLGSDALDDFDLDML	125
29	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS DPGALVRHQRTHTGEKPYKCPECGKSFSRSDNLVR HQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTG EKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPE CGKSFSTSGNLVRHQRTHTGEKPYKCPECGKSFSRS DNLVRHQRTHTGKKTSKRPAATKKAGQAKKKKGS YPYDVPDYALEDALDDFDLDMLGSDALDDFDLDM LGSDALDDFDLDMLGSDALDDFDLDML	126
30	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNH QRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPEC GKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRED NLHTHQRTHTGKKTSKRPAATKKAGQAKKKKGSY PYDVPDYALEEASGSGRADALDDFDLDMLGSDALD DFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN SRSSGSPKKKRKVGSQYLPDTDDRHRIEEKRKRTYE TFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPA PQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAP PQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLG NSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPDEETSQAVKALREMADTVIPQKEEAAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTP	99

[0348]

구축물	서열	SEQ ID NO
	ELNEILDTFLNDECLLHAMHISTGLSIFDTSLF	
31	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDML	102
32	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDML	102
33	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEEASGSGRADALDDFDLDMLGSDALD DFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLIN SRSSGSPKKKRKVGSQYLPDTDDRHRIEEKRKRTYE TFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPA PQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAP PQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLG NSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPDEETSQAVKALREMADTVIPQKEEAAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTP ELNEILDTFLNDECLLHAMHISTGLSIFDTSLF	100
34	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSD ALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	127
35	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC	127

[0349]

구축물	서열	SEQ ID NO
	GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSD ALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	
36	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSRED NLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIH LRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQK PFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGR KFAQNSTLTEHTKIHLRQKDK	128
37	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
38	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGGGSGGGSGQSQLIKPSRMRKYPN RPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHT GQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFACDI CGRKFARSDELVRHTKIHLRQKDRPYACPVESCDRR FSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGHLVR HIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQK DK	129
39	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDKLEMAD HLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGP	105

[0350]

구축물	서열	SEQ ID NO
	GLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQP FPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLEL ELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	
40	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCADHLMLAEGYRLVQRPPSAAAAHGPHALRTL PPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFG PPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPN APGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEA LTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPP AGSVSCQSQLIKPSRMRKYPNRPSKTPPHERPYACP VESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFS REDNLHTHIRTHTGEKPFACDICGRKFARSDELVRH TKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRIHT GQKPFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDI CGRKFAQNSTLTEHTKIHLRQKDK	130
41	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGADHLMLAEGYRLVQRPPSAAAAH GPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPG ALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYP GRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGGM DAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRKY PNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFA CDICGRKFARSDELVRHTKIHLRQKDRPYACPVESC DRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDK	131
42	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDML	102
43	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHT HQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNS	102

[0351]

구축물	서열	SEQ ID NO
	TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDML	
44	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
45	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
46	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGGGSGGGSGQSQLIKPSRMRKYPN RPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHT GQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFACDI CGRKFARSDELVRHTKIHLRQKDRPYACPVESCDRR FSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGHLVR HIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQK DK	205
47	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCADHLMLAEGYRLVQRPPSAAAAHGPHALRTL PPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFG PPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPN APGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEA LTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPP AGSVSCQSQLIKPSRMRKYPNRPSKTPPHERPYACP VESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFS REDNLHTHIRTHTGEKPFACDICGRKFARSDELVRH	207

[0352]

구축물	서열	SEQ ID NO
	TKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRIHT GQKPFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDI CGRKFAQNSTLTEHTKIHLRQKDK	
48	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGADHLMLAEGYRLVQRPPSAAAAH GPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPG ALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYP GRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGGM DAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRKY PNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFA CDICGRKFARSDELVRHTKIHLRQKDRPYACPVESC DRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDK	209
49	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRLEPGEKPYKCPECGKSFSRSDNLVR HQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTG EKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPE CGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQ RTHTGKKTSVYVGGLESRVLKYTAQNMELQNKVQ LLEEQNLSLLDQLRKLQAMVIEISNKTSSSSTCILVLL VSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALPS EDPYQLELPALQSEVPKDSTHQWLDGSDCVLQAPG NTSCLLHYMPQAPSAEPPLEWPFPDLFSEPLCRGPIL PLQANLTRKGGWLPTGSPSVILQDRYSG	213
50	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRRPYACPVESCDRRFSRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFA CDICGRKFARSDELVRHTKIHLRQKDRPYACPVESC DRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDVYVGGLESRVLKYTAQNMELQNKVQLLEEQ NLSLLDQLRKLQAMVIEISNKTSSSSTCILVLLVSFCL LLVPAMYSSDTRGSLPAEHGVLSRQLRALPSEDPYQ LELPALQSEVPKDSTHQWLDGSDCVLQAPGNTSCL LHYMPQAPSAEPPLEWPFPDLFSEPLCRGPILPLQAN LTRKGGWLPTGSPSVILQDRYSG	217

[0353]

구축물	서열	SEQ ID NO
51	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRRPYACPVESCDRRFSRSDNLVRHIRI HTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFA CDICGRKFAREDNLHTHTKIHLRQKDRPYACPVESC DRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSL VRHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHL RQKDVYVGGLESRVLKYTAQNMELQNKVQLLEEQ NLSLLDQLRKLQAMVIEISNKTSSSSTCILVLLVSFCL LLVPAMYSSDTRGSLPAEHGVLSRQLRALPSEDPYQ LELPALQSEVPKDSTHQWLDGSDCVLQAPGNTSCL LHYMPQAPSAEPPLEWPFPDLFSEPLCRGPILPLQAN LTRKGGWLPTGSPSVILQDRYSG	219
52	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRLEPGEKPYKCPECGKSFSRSDNLVR HQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTG EKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPE CGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQ RTHTGKKTSVYVGGLESRVLKYTAQNMELQNKVQ LLEEQNLSLLDQLRKLQAMVIEIS	221
53	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRRPYACPVESCDRRFSRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFA CDICGRKFARSDELVRHTKIHLRQKDRPYACPVESC DRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDVYVGGLESRVLKYTAQNMELQNKVQLLEEQ NLSLLDQLRKLQAMVIEIS	223

[0354] 표 7. 본원에 개시된 예시적인 조작된 전사 인자를 코딩하는 핵산 서열.

구축물	서열	SEQ ID NO
31 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataacccctgcgtgctgcaccacctggggagaggggagg gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgcccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA	67

[0355]

구축물	서열	SEQ ID NO
	CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTCGATCTCGA TATGTTATAAACTAGTaaagagaccggttcactgtgacagtaaaa gagaccggttcactgtgagaatgaaagagaccggttcactgtgatcggaaaaga gaccggttcactgtgagcggccttgaaacccagcagacaatgtagctcagtaga aacccagcagacaatgtagctgaatggaaacccagcagacaatgtagcttcgg agaaacccagcagacaatgtagctAAGCTTGGGTGGCATCCC TGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGG AAGTTGCCACTCCAGTGCCCACCAGCCTTGTCCT AATAAAATTAAGTTGCATCATTTTGTCTGACTAG GTGTCCTTCTATAATATTATGGGGTGGAGGGGGG TGGTATGGAGCAAGGGGCAAGTTGGGAAGACAA CCTGTAGGGCCTGCGGGGTCTATTGGGAACCAA GCTGGAGTGCAGTGGCACAATCTTGGCTCACTGC AATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTG CCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCAT GCATGACCAGGCTCAGCTAATTTTTGTTTTTTTGG TAGAGACGGGGTTTCACCATATTGGCCAGGCTGG TCTCCAACTCCTAATCTCAGGTGATCTACCCACC TTGGCCTCCCAAATTGCTGGGATTACAGGCGTGA ACCACTGCTCCCTTCCCTGTCCTT	
32 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataaccccctgcgtgctgcaccacctggggagaggggggag gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg	68

[0356]

[0357]

구축물	서열	SEQ ID NO
	gaggttgagtacgttctggattactcataagacctttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaacccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttcccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgccccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTCGATCTCGA TATGTTATAAACTAGTAAGCTTGGGTGGCATCCC TGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGG AAGTTGCCACTCCAGTGCCCACCAGCCTTGTCCT AATAAAATTAAGTTGCATCATTTTGTCTGACTAG GTGTCCTTCTATAATATTATGGGGTGGAGGGGGG TGGTATGGAGCAAGGGGCAAGTTGGGAAGACAA	

구축물	서열	SEQ ID NO
	CCTGTAGGGCCTGCGGGGTCTATTGGGAACCAA GCTGGAGTGCAGTGGCACAATCTTGGCTCACTGC AATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTG CCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCAT GCATGACCAGGCTCAGCTAATTTTTGTTTTTTTGG TAGAGACGGGGTTTCACCATATTGGCCAGGCTGG TCTCCAACTCCTAATCTCAGGTGATCTACCCACC TTGGCCTCCCAAATTGCTGGGATTACAGGCGTGA ACCACTGCTCCCTTCCCTGTCCTT	
33 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggccccca gacgtgcgcataaataaccccctgcgtgctgcaccacctggggagaggggggag gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgccccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC	69

[0358]

[0359]

구축물	서열	SEQ ID NO
	AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGAGGCC AGCGGTTCCGGACGGGCTGACGCATTGGACGAT TTTGATCTGGATATGCTGGGAAGTGACGCCCTCG ATGATTTTGACCTTGACATGCTTGGTTCGGATGC CCTTGATGACTTTGACCTCGACATGCTCGGCAGT GACGCCCTTGATGATTTCGACCTGGACATGCTGA TTAACTCTAGAAGTTCCGGATCTCCGAAAAAGAA ACGCAAAGTTGGTAGCCAGTACCTGCCCGACAC CGACGACCGGCACCGGATCGAGGAAAAGCGGAA GCGGACCTACGAGACATTCAAGAGCATCATGAA GAAGTCCCCCTTCAGCGGCCCCACCGACCCTAGA CCTCCACCTAGAAGAATCGCCGTGCCCAGCAGAT CCAGCGCCAGCGTGCCAAAACCTGCCCCCCAGC CTTACCCCTTCACCAGCAGCCTGAGCACCATCAA CTACGACGAGTTCCCTACCATGGTGTTCCCCAGC GGCCAGATCTCTCAGGCCTCTGCTCTGGCTCCAG CCCCTCCTCAGGTGCTGCCTCAGGCTCCTGCTCC TGCACCAGCTCCAGCCATGGTGTCTGCACTGGCT CAGGCACCAGCACCCGTGCCTGTGCTGGCTCCTG GACCTCCACAGGCTGTGGCTCCACCAGCCCCTAA ACCTACACAGGCCGGCGAGGGCACACTGTCTGA AGCTCTGCTGCAGCTGCAGTTCGACGACGAGGAT CTGGGAGCCCTGCTGGGAAACAGCACCGATCCT GCCGTGTTCACCGACCTGGCCAGCGTGGACAAC AGCGAGTTCCAGCAGCTGCTGAACCAGGGCATC CCTGTGGCCCCTCACACCACCGAGCCCATGCTGA TGGAATACCCCGAGGCCATCACCCGGCTCGTGAC AGGCGCTCAGAGGCCTCCTGATCCAGCTCCTGCC CCTCTGGGAGCACCAGGCCTGCCTAATGGACTGC TGTCTGGCGACGAGGACTTCAGCTCTATCGCCGA TATGGATTTCTCAGCCTTGCTGGGCTCTGGCAGC GGCAGCCGGGATTCCAGGGAAGGGATGTTTTTG CCGAAGCCTGAGGCCGGCTCCGCTATTAGTGACG TGTTTGAGGGCCGCGAGGTGTGCCAGCCAAAAC	

구축물	서열	SEQ ID NO
	GAATCCGGCCATTTCATCCTCCAGGAAGTCCATG GGCCAACCGCCCACTCCCCGCCAGCCTCGCACCA ACACCAACCGGTCCAGTACATGAGCCAGTCGGG TCACTGACCCCGGCACCAGTCCCTCAGCCACTGG ATCCAGCGCCCGCAGTGACTCCCGAGGCCAGTC ACCTGTTGGAGGATCCCGATGAAGAGACGAGCC AGGCTGTCAAAGCCCTTCGGGAGATGGCCGATA CTGTGATTCCCCAGAAGGAAGAGGCTGCAATCT GTGGCCAAATGGACCTTTCCCATCCGCCCCCAAG GGGCCATCTGGATGAGCTGACAACCACACTTGA GTCCATGACCGAGGATCTGAACCTGGACTCACCC CTGACCCCGGAATTGAACGAGATTCTGGATACCT TCCTGAACGACGAGTGCCTCTTGCATGCCATGCA TATCAGCACAGGACTGTCCATCTTCGACACATCT CTGTTTTAAACTAGTaataaaagatctttattttcattagatctgtgtgt tggttttttgtgtg	
34 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataaccccctgcgtgctgcaccacctggggagaggggggag gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt	70

[0360]

[0361]

구축물	서열	SEQ ID NO
	cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgcccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctcccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCGACGC GCTGGACGATTTCGATCTCGACATGCTGGGTTCT GATGCCCTCGATGACTTTGACCTGGATATGTTGG GAAGCGACGCATTGGATGACTTTGATCTGGACAT GCTCGGCTCCGATGCTCTGGACGATTTCGATCTC GATATGTTATAAAAGCTTGGGTGGCATCCCTGTG ACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGT TGCCACTCCAGTGCCCACCAGCCTTGTCCTAATA AAATTAAGTTGCATCATTTTGTCTGACTAGGTGT CCTTCTATAATATTATGGGGTGGAGGGGGGTGGT ATGGAGCAAGGGGCAAGTTGGGAAGACAACCTG TAGGGCCTGCGGGGTCTATTGGGAACCAAGCTG GAGTGCAGTGGCACAATCTTGGCTCACTGCAATC TCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTC AGCCTCCCGAGTTGTTGGGATTCCAGGCATGCAT GACCAGGCTCAGCTAATTTTTGTTTTTTTGGTAG AGACGGGGTTTCACCATATTGGCCAGGCTGGTCT CCAACTCCTAATCTCAGGTGATCTACCCACCTTG GCCTCCCAAATTGCTGGGATTACAGGCGTGAACC ACTGCTCCCTTCCCTGTCCTT	
26 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc	71

[0362]

구축물	서열	SEQ ID NO
	tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataacccctgcgtgctgcaccacctggggagaggggagg gaccacggtaaatggagcgagcgcatagcaaaagggacgcgggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgcccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCGACGC GCTGGACGATTTCGATCTCGACATGCTGGGTTCT	

[0363]

구축물	서열	SEQ ID NO
	GATGCCCTCGATGACTTTGACCTGGATATGTTGG GAAGCGACGCATTGGATGACTTTGATCTGGACAT GCTCGGCTCCGATGCTCTGGACGATTTCGATCTC GATATGTTATAAaaagagaccggttcactgtgacagtaaaagagac cggttcactgtgagaatgaaagagaccggttcactgtgatcggaaaagagaccg gttcactgtgagcggccttgaaacccagcagacaatgtagctcagtagaaaccc agcagacaatgtagctgaatggaaacccagcagacaatgtagcttcggagaaa cccagcagacaatgtagctAAGCTTGGGTGGCATCCCTGTG ACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGT TGCCACTCCAGTGCCCACCAGCCTTGTCCTAATA AAATTAAGTTGCATCATTTTGTCTGACTAGGTGT CCTTCTATAATATTATGGGGTGGAGGGGGGTGGT ATGGAGCAAGGGGCAAGTTGGGAAGACAACCTG TAGGGCCTGCGGGGTCTATTGGGAACCAAGCTG GAGTGCAGTGGCACAATCTTGGCTCACTGCAATC TCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTC AGCCTCCCGAGTTGTTGGGATTCCAGGCATGCAT GACCAGGCTCAGCTAATTTTTGTTTTTTTGGTAG AGACGGGGTTTCACCATATTGGCCAGGCTGGTCT CCAACTCCTAATCTCAGGTGATCTACCCACCTTG GCCTCCCAAATTGCTGGGATTACAGGCGTGAACC ACTGCTCCCTTCCCTGTCCTT	
40 (코딩)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCTCCATCTGCCGCTGC CGCCCACGGCCCCCACGCCCTGAGAACACTGCCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGCGCCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCCAGCcGCTGGCATCGCCCACCTGCAG CCTGTGGCCACCCCTTACCCCGGAAGAGCCGCCG CCCCTCCAAACGCCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAACTGGGCCTGCACAGAGTGCGGGAACTGC CTGAGCTGTTCCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCCTCCTGCCGGC TCTGTGTCCTGCgccgaccacctgatgctcgccgagggctaccgcct ggtgcagaggccgccgtccgccgccgccgcccatggccctcatgcgctccgg actctgccgccgtacgcgggcccgggcctggacagtgggctgaggccgcgg ggggctccgctggggccgccgccgccccgccaacccggggccctggcgtac ggggccttcgggccgccgtcctccttccagccctttccggccgtgcctccgccg gccgcgggcatcgcgcacctgcagcctgtggcgacgccgtaccccggccgc gcCgccgccccccaacgctccgggaggccccccgggcccgcagccggc cccaagcgccgcagccccgccgccgcccgcgcacgccctgggcggcatgga cgccgaactcatcgacgaggaggcgctgacgtcgctggagctggagctgggg ctgcaccgcgtgcgcgagctgcccgagctgttcctgggccagagcgagttcga ctgcttctcggacttggggtccgcgccgcccgccggctccgtgagctgccagtc ccagctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagcaag	72

[0364]

구축물	서열	SEQ ID NO
	acgccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgccgc ttctccCGCAGCGACAACCTGGTGAGAcacatccgcatccac acaggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCG AGAGGATAACTTGCACACTcacatccgcacccacacaggcg aaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGA TGAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgc ccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGG GAATCTGACTGAGcacatccgcatccacacaggccagaagcccttc cagtgccgcatctgcatgAGAaacttcagcACAAGTGGACATCT GGTACGCcacatccgcacccacacaggcgaaaagcccttcgcctgcgac atctgtggaagaaagtttgccCAGAATAGTACCCTGACCGAA cataccaagatccacttgcggcagaaggacaag	
41 (코딩)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCTCCATCTGCCGCTGC CGCCCACGGCCCCCACGCCCTGAGAACACTGCCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGCGCCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCCAGCTGCTGGCATCGCCCACCTGCAG CCTGTGGCCACCCCTTACCCCGGAAGAGCCGCCG CCCCTCCAAACGCCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAACTGGGCCTGCACAGAGTGCGGGAACTGC CTGAGCTGTTCCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCCTCCTGCCGGC TCTGTGTCCTGCGGCGGCAGCGGCGGCGGAAGC GGCgccgaccacctgatgctcgccgagggctaccgcctggtgcagaggcc gccgtccgccgccgccgcccatggccctcatgcgctccggactctgccgccgt acgcgggcccgggcctggacagtgggctgaggccgcgggggggctccgctgg ggccgccgccgcccccgccaacccggggccctggcgtacggggccttcgggc cgccgtcctccttccagccctttccggccgtgcctccgccggccgcgggcatcg cgcacctgcagcctgtggcgacgccgtaccccggccgcgcggccgcgcccc ccaacgctccgggaggccccccgggcccgcagccggccccaagcgccgca gccccgccgccgcccgcgcacgccctgggcggcatggacgccgaactcatc gacgaggaggcgctgacgtcgctggagctggagctggggctgcaccgcgtgc gcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctcggactt ggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggtggtt ctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccgg cccagcaagacgccccccacgaacgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCGCAGCGACAACCTGGTGAGAcacatc cgcatccacacaggccagaagcccttccagtgccgcatctgcatgAGAaact tcagcCGAGAGGATAACTTGCACACTcacatccgcacccac acaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGG AGCGATGAACTTGTCCGAcataccaagatccacttgcggcaga aggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAA TCAGGGAATCTGACTGAGcacatccgcatccacacaggccag aagcccttccagtgccgcatctgcatgAGAaacttcagcACAAGTGG	73

구축물	서열	SEQ ID NO
	ACATCTGGTACGCcacatccgcacccacacaggcgaaaagcccttc gcctgcgacatctgtggaagaaagtttgccCAGAATAGTACCCTG ACCGAAcataccaagatccacttgcggcagaaggacaag	
42 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataaccctgcgtgctgcaccacctggggagaggggggag gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgccccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC	74

[0365]

[0366]

구축물	서열	SEQ ID NO
	ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTCGATCTCGA TATGTTATAAACTAGTGAAACCCAGCAGACAAT GTAGCTAGACCCAGTAGCCAGATGTAGCTAAAG AGACCGGTTCACTGTGAAAGCTTGGGTGGCATCC CTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTG GAAGTTGCCACTCCAGTGCCCACCAGCCTTGTCC TAATAAAATTAAGTTGCATCATTTTGTCTGACTA GGTGTCCTTCTATAATATTATGGGGTGGAGGGGG GTGGTATGGAGCAAGGGGCAAGTTGGGAAGACA ACCTGTAGGGCCTGCGGGGTCTATTGGGAACCA AGCTGGAGTGCAGTGGCACAATCTTGGCTCACTG CAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCT GCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCA TGCATGACCAGGCTCAGCTAATTTTTGTTTTTTTG GTAGAGACGGGGTTTCACCATATTGGCCAGGCTG GTCTCCAACTCCTAATCTCAGGTGATCTACCCAC CTTGGCCTCCCAAATTGCTGGGATTACAGGCGTG AACCACTGCTCCCTTCCCTGTCCTT	
43 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataaccccctgcgtgctgcaccacctggggagaggggggag gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg	75

[0367]

구축물	서열	SEQ ID NO
	gaggttgagtacgttctggattactcataagacctttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggctttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgcccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTCAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTCGATCTCGA TATGTTATAAACTAGTGAAACCCAGCAGACAAT GTAGCTAGACCCAGTAGCCAGATGTAGCTAAAG AGACCGGTTCACTGTGAGAAACCCAGCAGACAA TGTAGCTAGACCCAGTAGCCAGATGTAGCTAAA GAGACCGGTTCACTGTGAAAGCTTGGGTGGCATC CCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCT	

구축물	서열	SEQ ID NO
	GGAAGTTGCCACTCCAGTGCCCACCAGCCTTGTC CTAATAAAATTAAGTTGCATCATTTTGTCTGACT AGGTGTCCTTCTATAATATTATGGGGTGGAGGGG GGTGGTATGGAGCAAGGGGCAAGTTGGGAAGAC AACCTGTAGGGCCTGCGGGGTCTATTGGGAACC AAGCTGGAGTGCAGTGGCACAATCTTGGCTCACT GCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCC TGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGC ATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTT GGTAGAGACGGGGTTTCACCATATTGGCCAGGCT GGTCTCCAACTCCTAATCTCAGGTGATCTACCCA CCTTGGCCTCCCAAATTGCTGGGATTACAGGCGT GAACCACTGCTCCCTTCCCTGTCCTT	
44 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataaccctgcgtgctgcaccacctggggagaggggagg gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgccccggg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga	76

[0368]

구축물	서열	SEQ ID NO
	ggccccgggacgaccgagctgGAATTCGCCACCATGGCCgc cgaccacctgatgctcgccgagggctaccgcctggtgcagaggccgccgtcc gccgccgccgcccatggccctcatgcgctccggactctgccgccgtacgcgggg cccgggcctggacagtgggctgaggccgcgggggggctccgctggggccgcc gccgccccgccaacccggggccctggcgtacggggccttcgggccgccgtc ctccttccagccctttccggccgtgcctccgccggccgcgggcatcgcgcacct gcagcctgtggcgacgccgtaccccggccgcgcggccgcgccccccaacgc tccgggaggcccccccgggcccgcagccggccccaagcgccgcagccccgc cgccgcccgcgcacgccctgggcggcatggacgccgaactcatcgacgagg aggcgctgacgtcgctggagctggagctggggctgcaccgcgtgcgcgagct gcccgagctgttcctgggccagagcgagttcgactgcttctcggacttggggtc cgcgccgcccgccggctccgtgagctgcggtggttctggtggtggttctggtca gtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagca agacgccccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgcc gcttctccCGCAGCGACAACCTGGTGAGAcacatccgcatcc acacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcC GAGAGGATAACTTGCACACTcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCG ATGAACTTGTCCGAcataccaagatccacttgcggcagaaggacc gcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAG GGAATCTGACTGAGcacatccgcatccacacaggccagaagccct tccagtgccgcatctgcatgAGAaacttcagcACAAGTGGACATC TGGTACGCcacatccgcacccacacaggcgaaaagcccttcgcctgcg acatctgtggaagaaagtttgccCAGAATAGTACCCTGACCG AAcataccaagatccacttgcggcagaaggacaagtaaCTCGAGGA AACCCAGCAGACAATGTAGCTAGACCCAGTAGC CAGATGTAGCTAAAGAGACCGGTTCACTGTGAA AGCTTGGGTGGCATCCCTGTGACCCCTCCCCAGT GCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTG CCCACCAGCCTTGTCCTAATAAAATTAAGTTGCA TCATTTTGTCTGACTAGGTGTCCTTCTATAATATT ATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGG CAAGTTGGGAAGACAACCTGTAGGGCCTGCGGG GTCTATTGGGAACCAAGCTGGAGTGCAGTGGCA CAATCTTGGCTCACTGCAATCTCCGCCTCCTGGG TTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTT GTTGGGATTCCAGGCATGCATGACCAGGCTCAGC TAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCAC CATATTGGCCAGGCTGGTCTCCAACTCCTAATCT CAGGTGATCTACCCACCTTGGCCTCCCAAATTGC TGGGATTACAGGCGTGAACCACTGCTCCCTTCCC TGTCCTT	
45 (RE + 코딩)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatg aaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgt gttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctt tcacatcttatgaaagtcatttaagcacaactaacttttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagc	184

[0369]

[0370]

구축물	서열	SEQ ID NO
	tggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacg ggggtttcaccatatcacccaggctggtctcgaactcctggcctcaagtgatcca cctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaac ccgaccaacctttttttaaaataaatatttaaaaaattggtatttcacatatatactagta tttacatttatccacacaaaacggacgggcctccgctgaaccagtgaggcccca gacgtgcgcataaataacccctgcgtgctgcaccacctggggagaggggggag gaccacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaaaccgccgctcgtctcctacaatatatgggagggg gaggttgagtacgttctggattactcataagacctttttttttccttccgggcgcaaa accgtgagctggatttataatcgccctataaagctccagaggcggtcaggcacct gcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctacccc ggagccgtgcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcga gagggaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagaggaaa ggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcgac cacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttt tccaggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcg gggctagagtgcaaggtgactgtggttcttctctggccaagtccgagggagaac gtaaagatatgggccttttttccccctctcaccttgtctcaccaaagtccctagtccc cggagcagttagcctctttctttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgtgcgccctccccgctcgctgccttt cctccctcttgtctctccagagccggatcttcaaggggagcctccgtgcccccgg ctgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctc gcttctctttgcagcctgtttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCgc cgaccacctgatgctcgccgagggctaccgcctggtgcagaggccgccgtcc gccgccgccgcccatggccctcatgcgctccggactctgccgccgtacgcggg cccgggcctggacagtgggctgaggccgcgggggggctccgctggggccgcc gccgccccgccaacccggggccctggcgtacggggccttcgggccgccgtc ctccttccagccctttccggccgtgcctccgccggccgcgggcatcgcgcacct gcagcctgtggcgacgccgtaccccggccgcgcggccgcgccccccaacgc tccgggaggccccccgggcccgcagccggccccaagcgccgcagccccgc cgccgcccgcgcacgccctgggcggcatggacgccgaactcatcgacgagg aggcgctgacgtcgctggagctggagctggggctgcaccgcgtgcgcgagct gcccgagctgttcctgggccagagcgagttcgactgcttctcggacttggggtc cgcgccgcccgccggctccgtgagctgcggtggttctggtggtggttctggtca gtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagca agacgccccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgcc gcttctccCGCAGCGACAACCTGGTGAGAcacatccgcatcc acacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcC GAGAGGATAACTTGCACACTcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCG ATGAACTTGTCCGAcataccaagatccacttgcggcagaaggacc gcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAG GGAATCTGACTGAGcacatccgcatccacacaggccagaagccct	

[0371]

구축물	서열	SEQ ID NO
	tccagtgccgcatctgcatgAGAaacttcagcACAAGTGGACATC TGGTACGCcacatccgcacccacacaggcgaaaagcccttcgcctgcg acatctgtggaagaaagtttgccCAGAATAGTACCCTGACCG AAcataccaagatccacttgcggcagaaggacaagtaaCTCGAGGA AACCCAGCAGACAATGTAGCTAGACCCAGTAGC CAGATGTAGCTAAAGAGACCGGTTCACTGTGAG AAACCCAGCAGACAATGTAGCTAGACCCAGTAG CCAGATGTAGCTAAAGAGACCGGTTCACTGTGA AAGCTTGGGTGGCATCCCTGTGACCCCTCCCCAG TGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGT GCCCACCAGCCTTGTCCTAATAAAATTAAGTTGC ATCATTTTGTCTGACTAGGTGTCCTTCTATAATAT TATGGGGTGGAGGGGGGTGGTATGGAGCAAGGG GCAAGTTGGGAAGACAACCTGTAGGGCCTGCGG GGTCTATTGGGAACCAAGCTGGAGTGCAGTGGC ACAATCTTGGCTCACTGCAATCTCCGCCTCCTGG GTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGT TGTTGGGATTCCAGGCATGCATGACCAGGCTCAG CTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCA CCATATTGGCCAGGCTGGTCTCCAACTCCTAATC TCAGGTGATCTACCCACCTTGGCCTCCCAAATTG CTGGGATTACAGGCGTGAACCACTGCTCCCTTCC CTGTCCTT	
8 (코딩)	ATGGCCgccgaccacctgatgctcgccgagggctaccgcctggtgcaga ggccgccgtccgccgccgccgcccatggccctcatgcgctccggactctgccg ccgtacgcgggcccgggcctggacagtgggctgaggccgcgggggctccg ctggggccgccgccgccccgccaacccggggccctggcgtacggggccttc gggccgccgtcctccttccagccctttccggccgtgcctccgccggccgcgggg catcgcgcacctgcagcctgtggcgacgccgtaccccggccgcgcggccgcg ccccccaacgctccgggaggccccccgggcccgcagccggccccaagcgcc gcagccccgccgccgcccgcgcacgccctgggcggcatggacgccgaactc atcgacgaggaggcgctgacgtcgctggagctggagctggggctgcaccgcg tgcgcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctcgg acttggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggtg gttctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaacc ggcccagcaagacgccccccccacgaacgcccttacgcttgcccagtggagtcc tgtgatcgccgcttctccCGCAGCGACAACCTGGTGAGAcac atccgcatccacacaggccagaagcccttccagtgccgcatctgcatgAGAa acttcagcCGAGAGGATAACTTGCACACTcacatccgcacc cacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCG GAGCGATGAACTTGTCCGAcataccaagatccacttgcggca gaaggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCA ATCAGGGAATCTGACTGAGcacatccgcatccacacaggcca gaagcccttccagtgccgcatctgcatgAGAaacttcagcACAAGTG GACATCTGGTACGCcacatccgcacccacacaggcgaaaagccct tcgcctgcgacatctgtggaagaaagtttgccCAGAATAGTACCCT GACCGAAcataccaagatccacttgcggcagaaggacaag	203
46 (코딩)	ATGGCCgccgaccacctgatgctcgccgagggctaccgcctggtgcaga ggccgccgtccgccgccgccgcccatggccctcatgcgctccggactctgccg ccgtacgcgggcccgggcctggacagtgggctgaggccgcgggggctccg	204

구축물	서열	SEQ ID NO
	ctggggccgccgccgccccgccaacccggggccctggcgtacggggccttc gggccgccgtcctccttccagccctttccggccgtgcctccgccggccgcggg catcgcgcacctgcagcctgtggcgacgccgtaccccggccgcgcggccgcg ccccccaacgctccgggaggccccccgggcccgcagccggccccaagcgcc gcagccccgccgccgcccgcgcacgccctgggcggcatggacgccgaactc atcgacgaggaggcgctgacgtcgctggagctggagctggggctgcaccgcg tgcgcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctcgg acttggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggtg gttctggtGGTGGCAGCGGGGGAGGTTCTGGTcagtccca gctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagcaagacg ccccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgccgcttct ccCGCAGCGACAACCTGGTGAGAcacatccgcatccacaca ggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCGAG AGGATAACTTGCACACTcacatccgcacccacacaggcgaaaa gcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGATGA ACTTGTCCGAcataccaagatccacttgcggcagaaggaccgccctta cgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGGGAA TCTGACTGAGcacatccgcatccacacaggccagaagcccttccagtg ccgcatctgcatgAGAaacttcagcACAAGTGGACATCTGGT ACGCcacatccgcacccacacaggcgaaaagcccttcgcctgcgacatctg tggaagaaagtttgccCAGAATAGTACCCTGACCGAAcatac caagatccacttgcggcagaaggacaag	
47 (코딩)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCTCCATCTGCCGCTGC CGCCCACGGCCCCCACGCCCTGAGAACACTGCCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGCGCCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCCAGCcGCTGGCATCGCCCACCTGCAG CCTGTGGCCACCCCTTACCCCGGAAGAGCCGCCG CCCCTCCAAACGCCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAACTGGGCCTGCACAGAGTGCGGGAACTGC CTGAGCTGTTCCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCCTCCTGCCGGC TCTGTGTCCTGCgccgaccacctgatgctcgccgagggctaccgcct ggtgcagaggccgccgtccgccgccgccgcccatggccctcatgcgctccgg actctgccgccgtacgcgggcccgggcctggacagtgggctgaggccgcgg ggggctccgctggggccgccgccgccccgccaacccggggccctggcgtac ggggccttcgggccgccgtcctccttccagccctttccggccgtgcctccgccg gccgcgggcatcgcgcacctgcagcctgtggcgacgccgtaccccggccgc gcCgccgcgccccccaacgctccgggaggccccccgggcccgcagccggc cccaagcgccgcagccccgccgccgcccgcgcacgccctgggcggcatgga cgccgaactcatcgacgaggaggcgctgacgtcgctggagctggagctgggg ctgcaccgcgtgcgcgagctgcccgagctgttcctgggccagagcgagttcga ctgcttctcggacttggggtccgcgccgcccgccggctccgtgagctgccagtc ccagctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagcaag	206

[0372]

구축물	서열	SEQ ID NO
	acgccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgccgc ttctccCGCAGCGACAACCTGGTGAGAcacatccgcatccac acaggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCG AGAGGATAACTTGCACACTcacatccgcacccacacaggcg aaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGA TGAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgc ccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGG GAATCTGACTGAGcacatccgcatccacacaggccagaagcccttc cagtgccgcatctgcatgAGAaacttcagcACAAGTGGACATCT GGTACGCcacatccgcacccacacaggcgaaaagcccttcgcctgcgac atctgtggaagaaagtttgccCAGAATAGTACCCTGACCGAA cataccaagatccacttgcggcagaaggacaag	
48 (코딩)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCTCCATCTGCCGCTGC CGCCCACGGCCCCCACGCCCTGAGAACACTGCCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGCGCCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCCAGCTGCTGGCATCGCCCACCTGCAG CCTGTGGCCACCCCTTACCCCGGAAGAGCCGCCG CCCCTCCAAACGCCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAACTGGGCCTGCACAGAGTGCGGGAACTGC CTGAGCTGTTCCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCCTCCTGCCGGC TCTGTGTCCTGCGGCGGCAGCGGCGGCGGAAGC GGCgccgaccacctgatgctcgccgagggctaccgcctggtgcagaggcc gccgtccgccgccgccgcccatggccctcatgcgctccggactctgccgccgt acgcgggcccgggcctggacagtgggctgaggccgcgggggggctccgctgg ggccgccgccgccccgccaacccggggccctggcgtacggggccttcgggc cgccgtcctccttccagccctttccggccgtgcctccgccggccgcgggcatcg cgcacctgcagcctgtggcgacgccgtaccccggccgcgcggccgcgcccc ccaacgctccgggaggccccccgggcccgcagccggccccaagcgccgca gccccgccgccgcccgcgcacgccctgggcggcatggacgccgaactcatc gacgaggaggcgctgacgtcgctggagctggagctggggctgcaccgcgtgc gcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctcggactt ggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggtggtt ctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccgg cccagcaagacgccccccacgaacgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCGCAGCGACAACCTGGTGAGAcacatc cgcatccacacaggccagaagcccttccagtgccgcatctgcatgAGAaact tcagcCGAGAGGATAACTTGCACACTcacatccgcacccac acaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGG AGCGATGAACTTGTCCGAcataccaagatccacttgcggcaga aggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAA TCAGGGAATCTGACTGAGcacatccgcatccacacaggccag aagcccttccagtgccgcatctgcatgAGAaacttcagcACAAGTGG	208

[0373]

[0374]

구축물	서열	SEQ ID NO
	ACATCTGGTACGCcacatccgcacccacacaggcgaaaagcccttc gcctgcgacatctgtggaagaaagtttgccCAGAATAGTACCCTG ACCGAAcataccaagatccacttgcggcagaaggacaag	
49 (코딩)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagag gaaagtggagatttggggacggcacccgatgaggccgtgagggccccactgg actgggcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtcccccagcgtcgttgaacattctcagctcctccaaccc tgccttgtccaccatgaccacacctactccctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgcctgagacacttcctctcactaa gacagaggaacaaattctgaaacgtgtgcggCTCGAACCAGGTGA AAAACCTTACAAATGTCCTGAATGTGGGAAATC ATTCAGTCGCAGCGACAACCTGGTGAGACATCA ACGCACCCATACAGGAGAAAAACCTTATAAATG TCCAGAATGTGGAAAGTCCTTCTCACGAGAGGAT AACTTGCACACTCATCAACGAACACATACTGGTG AAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCA ACGAACCCATACAGGCGAGAAGCCTTACAAATG TCCCGAGTGTGGCAAGAGCTTCTCACAATCAGGG AATCTGACTGAGCATCAACGAACTCATACCGGG GAAAAACCTTACAAGTGTCCAGAGTGTGGGAAG AGCTTTTCCACAAGTGGACATCTGGTACGCCACC AGAGGACACATACAGGGGAGAAGCCCTACAAAT GCCCCGAATGCGGTAAAAGTTTCTCTCAGAATAG TACCCTGACCGAACACCAGCGAACACACACTGG GAAAAAAACGAGTgtgtaCgttgggggtttagagagcCgggtctt gaaatacacagcccagaatatggagcttcagaacaaagtacagcttctggagga acagaatttgtcccttctagatcaactgaggaaactccaggccatggtgattgaga tCtcaaacaaaaccagcagcagcagcacctgcatcttggtcctGctagtctcctt ctgcctcctccttgtacctgctatgtactcctctgacacaaggggagcctgccag ctgagcatggagtgttgtcccgccagcttcgtgccctccccagtgaggaccctta ccagctggagctgcctgccctgcagtcagaagtgccgaaagacagcacacacc agtggttggacggctcagactgtgtactccaggcccctggcaacacttcctgcct gctgcattacatgcctcaggctcccagtgcagagcctcccctggagtggccCtt ccctgacctcttctcagagcctctctgccgaggtcccatcctcccccctgcaggca aatctcacaaggaagggaggatggcttcctactggtagcccctctgtcattttgca ggacagatactcaggc	212
50 (코딩)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagag gaaagtggagatttggggacggcacccgatgaggccgtgagggccccactgg actgggcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtcccccagcgtcgttgaacattctcagctcctccaaccc tgccttgtccaccatgaccacacctactccctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgcctgagacacttcctctcactaa gacagaggaacaaattctgaaacgtgtgcggcgcccttacgcttgcccagtgga gtcctgtgatcgccgcttctccCGCAGCGACAACCTGGTGAG AcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcatgA	216

[0375]

구축물	서열	SEQ ID NO
	GAaacttcagcCGAGAGGATAACTTGCACACTcacatccg cacccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgc cCGGAGCGATGAACTTGTCCGAcataccaagatccacttgcg gcagaaggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctcc CAATCAGGGAATCTGACTGAGcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcACAAGT GGACATCTGGTACGCcacatccgcacccacacaggcgaaaagcc cttcgcctgcgacatctgtggaagaaagtttgccCAGAATAGTACCC TGACCGAAcataccaagatccacttgcggcagaaggacgtgtaCgttg ggggtttagagagcCgggtcttgaaatacacagcccagaatatggagcttcag aacaaagtacagcttctggaggaacagaatttgtcccttctagatcaactgagga aactccaggccatggtgattgagatCtcaaacaaaaccagcagcagcagcacc tgcatcttggtcctGctagtctccttctgcctcctccttgtacctgctatgtactcctc tgacacaaggggagcctgccagctgagcatggagtgttgtcccgccagcttc gtgccctccccagtgaggacccttaccagctggagctgcctgccctgcagtcag aagtgccgaaagacagcacacaccagtggttggacggctcagactgtgtactcc aggcccctggcaacacttcctgcctgctgcattacatgcctcaggctcccagtgc agagcctccctggagtggccCttccctgacctcttctcagagcctctctgccga ggtcccatcctcccccctgcaggcaaatctcacaaggaagggaggatggcttcct actggtagcccctctgtcattttgcaggacagatactcaggc	
51 (코딩)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagag gaaagtggagatttggggacggcacccgatgaggccgtgagggccccactgg actgggcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtcccccagcgtcgttgaacattctcagctcctccaaccc tgccttgtccaccatgaccacacctactccctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgcctgagacacttcctctcactaa gacagaggaacaaattctgaaacgtgtgcggcgcccttacgcttgcccagtgga gtcctgtgatcgccgcttctccCGCTCAGACAACCTCGTTCGA cacatccgcatccacacaggccagaagcccttccagtgccgcatctgcatgAG Aaacttcagc CACCGGACTACACTCACGAACcacatccgca cccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgcc AGAGAAGACAATCTCCATACTcataccaagatccacttgcg gcagaaggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctcc ACCAGCCATTCTCTCACTGAAcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCAGTCT AGCTCACTGGTGAGGcacatccgcacccacacaggcgaaaagc ccttcgcctgcgacatctgtggaagaaagtttgccAGGGAGGATAAC CTGCATACGcataccaagatccacttgcggcagaaggacgtgtaCgtt gggggtttagagagcCgggtcttgaaatacacagcccagaatatggagcttca gaacaaagtacagcttctggaggaacagaatttgtcccttctagatcaactgagg aaactccaggccatggtgattgagatCtcaaacaaaaccagcagcagcagcac ctgcatcttggtcctGctagtctccttctgcctcctccttgtacctgctatgtactcct ctgacacaaggggagcctgccagctgagcatggagtgttgtcccgccagctt cgtgccctccccagtgaggacccttaccagctggagctgcctgccctgcagtca gaagtgccgaaagacagcacacaccagtggttggacggctcagactgtgtact ccaggcccctggcaacacttcctgcctgctgcattacatgcctcaggctcccagt gcagagcctccctggagtggccCttccctgacctcttctcagagcctctctgcc gaggtcccatcctcccctgcaggcaaatctcacaaggaagggaggatggcttc	218

구축물	서열	SEQ ID NO
	ctactggtagcccctctgtcattttgcaggacagatactcaggc	
52 (코딩)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagag gaaagtggagatttggggacggcacccgatgaggccgtgagggccccactgg actgggcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtcccccagcgtcgttgaacattctcagctcctccaacccc tgccttgtccaccatgaccacacctactccctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgcctgagacacttcctctcactaa gacagaggaacaaattctgaaacgtgtgcggCTCGAACCAGGTGA AAAACCTTACAAATGTCCTGAATGTGGGAAATC ATTCAGTCGCAGCGACAACCTGGTGAGACATCA ACGCACCCATACAGGAGAAAAACCTTATAAATG TCCAGAATGTGGAAAGTCCTTCTCACGAGAGGAT AACTTGCACACTCATCAACGAACACATACTGGTG AAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCA ACGAACCCATACAGGCGAGAAGCCTTACAAATG TCCCGAGTGTGGCAAGAGCTTCTCACAATCAGGG AATCTGACTGAGCATCAACGAACTCATACCGGG GAAAAACCTTACAAGTGTCCAGAGTGTGGGAAG AGCTTTTCCACAAGTGGACATCTGGTACGCCACC AGAGGACACATACAGGGGAGAAGCCCTACAAAT GCCCCGAATGCGGTAAAAGTTTCTCTCAGAATAG TACCCTGACCGAACACCAGCGAACACACACTGG GAAAAAAACGAGTgtgtaCgttgggggtttagagagcCgggtctt gaaatacacagcccagaatatggagcttcagaacaaagtacagcttctggagga acagaatttgtcccttctagatcaactgaggaaactccaggccatggtgattgaga tatca	220
53 (코딩)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagag gaaagtggagatttggggacggcacccgatgaggccgtgagggccccactgg actgggcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtcccccagcgtcgttgaacattctcagctcctccaacccc tgccttgtccaccatgaccacacctactccctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgcctgagacacttcctctcactaa gacagaggaacaaattctgaaacgtgtgcggcgcccttacgcttgcccagtgga gtcctgtgatcgccgcttctccCGCAGCGACAACCTGGTGAG AcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcatgA GAaacttcagcCGAGAGGATAACTTGCACACTcacatccg cacccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgc cCGGAGCGATGAACTTGTCCGAcataccaagatccacttgcg gcagaaggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctcc CAATCAGGGAATCTGACTGAGcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcACAAGT GGACATCTGGTACGCcacatccgcacccacacaggcgaaaagcc cttcgcctgcgacatctgtggaagaaagtttgccCAGAATAGTACCC TGACCGAAcataccaagatccacttgcggcagaaggacgtgtaCgttg gggtttagagagcCgggtcttgaaatacacagcccagaatatggagcttcag	222

[0376]

구축물	서열	SEQ ID NO
	aacaaagtacagcttctggaggaacagaatttgtcccttctagatcaactgagga aactccaggccatggtgattgagatCtca	

[0377]

[0378] 표 8. 예시적인 마이크로RNA 및 마이크로RNA 결합 부위에 대한 핵산 서열.

설명	서열	SEQ ID NO
M1 결합 부위	aaagagaccggttcactgtgacagtaaaagagaccggttcactgtgagaatgaaagag accggttcactgtgatcggaaaagagaccggttcactgtgagcggccttgaaacccagc agacaatgtagctcagtagaaacccagcagacaatgtagctgaatggaaacccagcag acaatgtagcttcggagaaacccagcagacaatgtagct	7
miR128 서열	UCACAGUGAACCGGUCUCUUU	8
miR128 결합 부위	AAAGAGACCGGTTCACTGTGA	9
miR221 서열	AGCUACAUUGUCUGCUGGGUUUC	10
miR221 결합 부위	GAAACCCAGCAGACAATGTAGCT	11
miR222 서열	AGCUACAUCUGGCUACUGGGUCU	12
miR222 결합 부위	AGACCCAGTAGCCAGATGTAGCT	13
M2 결합 부위	GAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCC AGATGTAGCTAAAGAGACCGGTTCACTGTGA	14
M3 결합 부위	GAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCC AGATGTAGCTAAAGAGACCGGTTCACTGTGAGAAAC CCAGCAGACAATGTAGCTAGACCCAGTAGCCAGATG TAGCTAAAGAGACCGGTTCACTGTGA	15

[0379]

[0380] 표 9. eTF를 생성하기 위한 상이한 유형의 아연 핑거 구조 및 예시적인 아연 핑거 단백질.

ZF 유형 명칭	SEQ ID NO	ZF 구조 (여기서 각각의 x는 독립적으로 임의의 잔기일 수 있음)	본원에 개시된 eTF 또는 eTF의 DNA 결합 도메인에 대한 단백질 플랫폼으로서의 역할을 할 수 있는 예시적인 단백질
아연 핑거 C2H2-유형 (ZNF)	136	C-x-C-x-H-x-H	KLF4, KLF5, EGR3, ZFP637, SLUG, ZNF750, ZNF281, ZBP89, GLIS1, GLIS3
고리 핑거 단백질 (RNF)	137	C-x-C-x-C-x-H-xxx-C-x-C-x-C-x-C	MDM2, BRCA1, ZNF179
PHD 핑거 단백질 (PHF)	138	C-x-C-x-C-x-C-xxx-H-x-C-x-C-x-C	KDM2A, PHF1, ING1
LIM 도메인 함유	139	C-x-C-x-H-x-C-x-C-x-C-x-C-x-(C,H,D)	ZNF185, LIMK1, PXN
핵 호르몬 수용체 (NR)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	VDR, ESR1, NR4A1
아연 핑거 CCCH-유형 (ZC3H)	141	C-x-C-x-C-x-H	RC3H1, HELZ, MBNL1, ZFP36, ZFP36L1
아연 핑거 FYVE-유형 (ZFYVE)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	EEA1, HGS, PIKFYVE
아연 핑거 CCHC-유형 (ZCCHC)	142	C-x-C-x-H-x-C	CNBP, SF1, LIN28A
아연 핑거 DHHC-유형 (ZDHHC)	143	C-x-C-x-H-x-C-xxx-C-x-C-x-H-x-C	ZDHHC2, ZDHHC8, ZDHHC9
아연 핑거 MYND-유형 (ZMYND)	144	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	PDCD2, RUNX1T1, SMYD2, SMYD1
아연 핑거 RANBP2-유형 (ZRANB)	145	C-x-C-x-C-x-C	YAF2, SHARPIN, EWSR1
아연 핑거 ZZ-유형 (ZZZ)	145	C-x-C-x-C-x-C	HERC2, NBR1, CREBBP
아연 핑거 C2HC-유형 (ZC2HC)	142	C-x-C-x-H-x-C	IKBKG, L3MBTL1, ZNF746
GATA 아연-핑거 도메인 함유 (GATAD)	145	C-x-C-x-C-x-C	GATA4, GATA6, MTA1
ZF 부류 호메오박스 및 위유전자	136	C-x-C-x-H-x-H	ADNP, ZEB1, ZHX1

[0381]

ZF 유형 명칭	SEQ ID NO	ZF 구조 (여기서 각각의 x는 독립적으로 임의의 잔기일 수 있음)	본원에 개시된 eTF 또는 eTF의 DNA 결합 도메인에 대한 단백질 플랫폼으로서의 역할을 할 수 있는 예시적인 단백질
THAP 도메인 함유 (THAP)	141	C-x-C-x-C-x-H	THAP1, THAP4, THAP11
아연 핑거 CXXC-유형 (CXXC)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	CXXC1, CXXC5, MBD1, DNMT1
아연 핑거 SWIM-유형 (ZSWIM)	141	C-x-C-x-C-x-H	MAP3K1, ZSWIM5, ZSWIM6
아연 핑거 AN1-유형 (ZFAND)	146	C-x-C-x-C-x-C-xxx-C-x-H-x-H-x-C	ZFAND3, ZFAND6, IGHMBP2
아연 핑거 3CxxC-유형 (Z3CXXC)	142	C-x-C-x-H-x-C	ZAR1, RTP1, RTP4
아연 핑거 CW-유형 (ZCW)	145	C-x-C-x-C-x-C	MORC1, ZCWPW1, KDM1B
아연 핑거 GRF-유형 (ZGRF)	145	C-x-C-x-C-x-C	TTF2, NEIL3, TOP3A
아연 핑거 MIZ-유형 (ZMIZ)	142	C-x-C-x-H-x-C	PIAS1, PIAS3, PIAS4
아연 핑거 BED-유형 (ZBED)	136	C-x-C-x-H-x-H	ZBED1, ZBED4, ZBED6
아연 핑거 HIT-유형 (ZNHIT)	144	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	ZNHIT3, DDX59, INO80B
아연 핑거 MYM-유형 (ZMYM)	145	C-x-C-x-C-x-C	ZMYM2, ZMYM3, ZMYM4
아연 핑거 마트린-유형 (ZMAT)	136	C-x-C-x-H-x-H	ZNF638, ZMAT1, ZMAT3, ZMAT5
아연 핑거 C2H2C-유형	136	C-x-C-x-H-x-H	MYT1, MYT1L, ST18
아연 핑거 DBF-유형 (ZDBF)	136	C-x-C-x-H-x-H	DBF4, DBF4B, ZDBF2
아연 핑거 PARP-유형	142	C-x-C-x-H-x-C	LIG3, PARP1

[0382]

[0383] 표 10. 예시적인 아연 핑거 DNA 결합 도메인에 대한 아미노산 서열.

DBD/표적 부위	서열	SEQ ID NO
eZF	LEPGEKP - [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH - TGKKTS, 여기서 n은 1-15의 정수이고, 각각의 X는 3 bp의 표적 서열에 결합할 수 있는 인식 서열임	147
Z1 표적 부위	RSDNLVR x REDNLHT x RSDELVR x QSGNLTE x TSGHLVR x QNSTLTE, 여기서 각각의 x는 1-50개의 아미노산 잔기를 포함하는 링커임	148
Z13 표적 부위	RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT, 여기서 각각의 x는 1-50개의 아미노산 잔기를 포함하는 링커임	149
Z14 표적 부위	DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR, 여기서 각각의 x는 1-50개의 아미노산 잔기를 포함하는 링커임	150
Z15 표적 부위	RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT, 여기서 각각의 x는 1-50개의 아미노산 잔기를 포함하는 링커임	151

[0384]

[0385] 표 11. 본원에 개시된 예시적인 아연 핑거 인식 서열에 대한 아미노산 서열.

[0386]

서열	SEQ ID NO
RSDNLVR	152
REDNLHT	153
RSDELVR	154
QSGNLTE	155
TSGHLVR	156
QNSTLTE	157
DPGALVR	158
HRTTLTN	159
QSGDLRR	160
TSHSLTE	161
THLDLIR	162
QSSSLVR	163
TSGNLVR	164
RRDELNV	165
RSDDLVR	166
RSDHLTN	167

[0387] 표 12. 본원에 개시된 다른 뉴클레오티드 및 아미노산 서열.

설명	서열	SEQ ID NO
EGR1 NLS 도메인	LIKPSRMRKYPNRPSK	168
SV40 NLS	PKKKRKV	169
뉴클레오플라스민 NLS	KRPAATKKAGQAKKKK	170
HA 태그	YPYDVPDYA	171
spA (합성 폴리A)	AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGG TTTTTTGTGTGCGGACCGCACGTG	16
hGH (인간 성장 호르몬 폴리A)	GGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCT GGCCCTGGAAGTTGCCACTCCAGTGCCCACCAGCCTT GTCCTAATAAAATTAAGTTGCATCATTTTGTCTGACTA GGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTG GTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGT AGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTG CAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCC TGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTT GTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAA TTTTTGTTTTTTTGGTAGAGACGGGGTTTCACCATATT GGCCAGGCTGGTCTCCAACTCCTAATCTCAGGTGATCT ACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGC GTGAACCACTGCTCCCTTCCCTGTCCTT	17
SCN1A 단백질	MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPD KKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLED LDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKI AIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKNVEY TFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITF AYVTEFVDLGNVSALRTFRVLRALKTISVIPGLKTIVGAL IQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQ WPPTNASLEEHSIEKNITVNYNGTLINETVFEFDWKSYIQ DSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRN PNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQLTLRAA GKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLE EAEQKEAEFQQMIEQLKKQQEAAQQAATATASEHSREP SAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGG EEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSSP HQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFAD DEHSTFEDNESRRDSLFVPRRHGERRNSNLSQTSRSSRM LAVFPANGKMHSTVDCNGVVSLVGGPSVPTSPVGQLLP EVIIDKPATDDNGTTTETEMRKRRSSSFHVSMDFLEDPSQ RQRAMSIASILTNTVEELEESRQKCPPCWYKFSNIFLIWD CSPYWLKVKHVVNLVVMDPFVDLAITICIVLNTLFMAM EHYPMTDHFNNVLTVGNLVFTGIFTAEMFLKIIAMDPYY YFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRV FKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIFAVV GMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVF	172

[0388]

[0389]

설명	서열	SEQ ID NO
	RVLCGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNL VVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRM HKGVAYVKRKIYEFIQQSFIRKQKILDEIKPLDDLNNKKD SCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDE SDYMSFINNPSLTVTVPIAVGESDFENLNTEDFSSESDLEE SKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACF TEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNW FETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVFT YIFILEMLLKWVAYGYQTYFTNAWCWLDFLIVDVSLVS LTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVV NALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYHCI NTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNF DNVGFGYLSLLQVATFKGWMDIMYAAVDSRNVELQPK YEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFG GQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQG MVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTT ILSRINLVFIVLFTGECVLKLISLRHYYFTIGWNIFDFVVVI LSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLIKGAK GIRTLLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNFAY VKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPIL NSKPPDCDPNKVNPGSSVKGDCGNPSVGIFFFVSYIIISFL VVVNMYIAVILENFSVATEESAEPLSEDDFEMFYEVWEK FDPDATQFMEFEKLSQFAAALEPPLNLPQPNKLQLIAMD LPMVSGDRIHCLDILFAFTKRVLGESGEMDALRIQMEER FMASNPSKVSYQPITTTLKRKQEEVSAVIIQRAYRRHLLK RTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEK TDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK	
dCAS 단백질	KRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANV ENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTD HSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGV HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLER LKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQ SFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLM GHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDEN EKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKG YRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIA KILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTH NLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQ KEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIE LAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKEN AKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEV DHIIPRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSS DSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRF SVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKV KSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANAD FIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEY KEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYS TRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLL MYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYL	173

설명	서열	SEQ ID NO
	TKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNK VVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYE VNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRV IGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQIIKKG	
dCAS9-VP64 구축물	MAPKKKRKVGIHGVPAAKRNYILGLAIGITSVGYGIIDYE TRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH RIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRN SKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEA KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGS PFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLY NALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPT LKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDI TARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQ EEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIF NRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIK VINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNR QTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLE AIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEE ASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRIS KTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMN LLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNK GYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFE EKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHR VDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKD NDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDE KNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLN AHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFV TVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIAS FYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYL ENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKK HPQIIKKGKRPAATKKAGQAKKKKGSYPYDVPDYALED ALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDML	174
WT EGR3 단백질 (인간)	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLFSGSS DSVVHYNQMATENVMDIGLTNEKPNPELSYSGSFQPAP GNKTVTYLGKFAFDSPSNWCQDNIISLMSAGILGVPPAS GALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLYS EPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMIPDY NLYHHPNDMGSIPEHKPFQGMDPIRVNPPPITPLETIKAF KDKQIHPGFGSLPQPPLTLKPIRPRKYPNRPSKTPLHERPH ACPAEGCDRRFSRSDELTRHLRIHTGHKPFQCRICMRSFS RSDHLTTHIRTHTGEKPFACEFCGRKFARSDERKRHAKI HLKQKEKKAEKGGAPSASSAPPVSLAPVVTTCA	175
WT EGR1 단백질 (인간)	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEM MLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGGGGG SNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNNEKVL VETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPLFSL	176

[0390]

[0391]

설명	서열	SEQ ID NO
	VSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSN DSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSAGTALQYPPP AYPAAKGGFQVPMIPDYLFPQQQGDLGLGTPDQKPFQG LESRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLI KPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRSDEL TRHIRIHTGQKPFQCRICMRNFSRSDHLTTHIRTHTGEKP FACDICGRKFARSDERKRHTKIHLRQKDKKADKSVVASS ATSSLSSYPSPVATSYPSPVTTSYPSPATTSYPSPVPTSFSS PGSSTYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSA VTNSFSASTGLSDMTATFSPRTIEIC	
링커	GGSGGGSG	177
링커	GGSGGGSGGGSGGGSG	178
링커	[GGGS]n	179
링커	[GGGGS]n	180
링커	[GGSG]n	181
WT EGR1 또는 EGR3에 대한 인식 부위	GCG(T/G)GGGCG	182
sgRNA 스캐폴드	GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAG GCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAG A	183
인간 CREB3 코딩 서열	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaaagt ggagatttggggacggcacccgatgaggccgtgagggccccactggactgggcgctg ccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctccctgctgagtc ccccagcgtcgttgaacattctcagctcctccaacccctgccttgtccaccatgaccacac ctactccctcccacgggaaactgtctctatggatctagagagtgagagctgtagaaaaga ggggacccagatgactccacagcatatggaggagctggcagagcaggagattgctagg ctagtactgacagatgaggagaagagtctattggagaaggaggggcttattctgcctgag acacttcctctcactaagacagaggaacaaattctgaaacgtgtgcggaggaagattcga aataaaagatctgctcaagagagccgcaggaaaaagaaggtgtaCgttgggggtttaga gagcCgggtcttgaaatacacagcccagaatatggagcttcagaacaaagtacagcttct ggaggaacagaatttgtcccttctagatcaactgaggaaactccaggccatggtgattgag atCtcaaacaaaaccagcagcagcagcacctgcatcttggtcctGctagtctccttctgc ctcctccttgtacctgctatgtactcctctgacacaaggggagcctgccagctgagcatg gagtgttgtcccgccagcttcgtgccctccccagtgaggacccttaccagctggagctgc ctgccctgcagtcagaagtgccgaaagacagcacacaccagtggttggacggctcaga ctgtgtactccaggcccctggcaacacttcctgcctgctgcattacatgcctcaggctccca gtgcagagcctcccctggagtggccCttccctgacctcttctcagagcctctctgccgag gtcccatcctcccccctgcaggcaaatctcacaaggaagggaggatggcttcctactggta gcccctctgtcattttgcaggcagatactcaggc	210
인간 CREB3 AA 서열	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWA LPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHD HTYSLPRETVSMDLESESCRKEGTQMTPQHMEELAEQEI ARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRVRRKI RNKRSAQESRRKKKVYVGGLESRVLKYTAQNMELQNK VQLLEEQNLSLLDQLRKLQAMVIEISNKTSSSSTCILVLL VSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALPSEDP YQLELPALQSEVPKDSTHQWLDGSDCVLQAPGNTSCLL	211

설명	서열	SEQ ID NO
	HYMPQAPSAEPPLEWPFPDLFSEPLCRGPILPLQANLTRK GGWLPTGSPSVILQDRYSG	
인간 CREB3 C-말단 도메인 (막횡단 영역 존재)	VYVGGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQL RKLQAMVIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSD TRGSLPAEHGVLSRQLRALPSEDPYQLELPALQSEVPKD STHQWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPLEW PFPDLFSEPLCRGPILPLQANLTRKGGWLPTGSPSVILQD RYSG	225
인간 CREB3 C-말단 도메인 (막횡단 영역 부재)	VYVGGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQL RKLQAMVIEIS	226
CREB3-TRE 코딩 서열	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaaagt ggagatttggggacggcacccgatgaggccgtgagggccccactggactgggcgctg ccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctccctgctgagtc ccccagcgtcgttgaacattctcagctcctccaaccccctgccttgtccaccatgaccacac ctactccctcccacgggaaactgtctctatggatctagagagtgagagctgtagaaaaga ggggacccagatgactccacagcatatggaggagctggcagagcaggagattgctagg ctagtactgacagatgaggagaagagtctattggagaaggaggggcttattctgcctgag acacttcctctcactaagacagaggaacaaattctgaaacgtgtgcggCTTGAGCC CGGAGAGAAGCCGTACAAGTGCCCTGAGTGCGGCAA GTCTTTTAGCAGAAGAGACGAACTTAATGTCCACCAG CGAACGCATACTGGTGAAAAGCCCTATAAATGTCCTG AATGTGGGAAATCATTCTCCAGCCGCAGAACCTGTAG GGCTCACCAGCGAACACACACCGGCGAAAAACCATA CAAATGTCCAGAATGCGGGAAATCCTTTTCTCAGTCAT CCAACTTGGTGAGACATCAACGCACGCACACTGGAGA AAAGCCTTACAAATGCCCGGAATGTGGAAAGTCTTTT TCCCAATTGGCCCATTTGCGAGCCCATCAGAGGACTC ACACGGGCGAGAAACCTTACAAATGCCCGGAATGCGG GAAATCTTTTTCAACGAGTGGCAACCTCGTAAGACAC CAAAGAACGCATACAGGCGAAAAGCCATATAAGTGTC CTGAGTGTGGTAAATCATTCTCACACAGGACCACCCT GACAAATCACCAGCGCACGCACACCGGCAAGAAGAC AAGCgtgtaCgttgggggtttagagagcCgggtcttgaaatacacagcccagaatat ggagcttcagaacaaagtacagcttctggaggaacagaatttgtcccttctagatcaactg aggaaactccaggccatggtgattgagatCtcaaacaaaaccagcagcagcagcacct gcatcttggtcctGctagtctccttctgcctcctccttgtacctgctatgtactcctctgacac aaggggagcctgccagctgagcatggagtgttgtcccgccagcttcgtgccctcccca gtgaggacccttaccagctggagctgcctgccctgcagtcagaagtgccgaaagacagc acacaccagtggttggacggctcagactgtgtactccaggcccctggcaacacttcctgc ctgctgcattacatgcctcaggctcccagtgcagagcctccccctggagtggccCttccct gacctcttctcagagcctctctgccgaggtcccatcctccccctgcaggcaaatctcacaa ggaagggaggatggcttcctactggtagcccctctgtcattttgcaggacagatactcagg c	214
CREB3-TRE	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWA LPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHD HTYSLPRETVSMDLESESCRKEGTQMTPQHMEELAEQEI ARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRVRLEP	215

[0392]

설명	서열	SEQ ID NO
	GEKPYKCPECGKSFSRRDELNVHQRTHTGEKPYKCPEC GKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLV RHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEK PYKCPECGKSFSTSGNLVRHQRTHTGEKPYKCPECGKSF SHRTTLTNHQRTHTGKKTSVYVGGLESRVLKYTAQNME LQNKVQLLEEQNLSLLDQLRKLQAMVIEISNKTSSSSTCI LVLLVSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALP SEDPYQLELPALQSEVPKDSTHQWLDGSDCVLQAPGNTS CLLHYMPQAPSAEPPLEWPFPDLFSEPLCRGPILPLQANL TRKGGWLPTGSPSVILQDRYSG	

[0393]

[0394] 실시예

[0395] 하기 실시예는 본 개시내용의 일부 측면을 추가로 기재하기 위해 포함되며, 본 발명의 범주를 제한하는데 사용되지 않아야 한다.

[0396] 실시예 1

[0397] SCN1A 특이적 전사 활성인자를 사용하여 SCN1A를 상향조절할 수 있는 표적 영역의 확인

[0398] 내인성 SCN1A 발현을 상향조절할 수 있는 게놈의 영역을 확인하기 위해, 상기 표 4 및 13에 제시된 바와 같은 게놈의 다양한 영역을 표적화하는 다양한 조작된 전사 인자 (아연 핑거 뉴클레아제 또는 gRNA/daCas9 구축물)를 설계하였다. gRNA/daCas9 구축물의 경우, gRNA는 상보적 게놈 가닥을 표적화하도록 설계되었기 때문에 gRNA는 표적 영역과 동일한 서열을 가졌다. dCas9 단백질은 dCAS9-VP64 구축물 (서열식별번호: 174)이었다.

[0399] HEK293 세포를 표준 방법에 따라 배양하고, 6-웰 플레이트의 웰당 조작된 전사 인자 또는 대조군 구축물을 보유하는 플라스미드 3ug으로 형질감염시켰다 (퓨진 HD, 프로메가(Promega)). 세포를 하기 표 13에 나타낸 구축물을 발현하는 플라스미드로 형질감염시켰다. 형질감염 48시간 후에, 세포를 수집하고, RNA를 단리하고 (퀴아젠(Qiagen) RNeasy 미니 키트), DNase 처리하였다. 올리고DT 프라이머 (슈퍼스크립트 IV, 인비트로젠(Invitrogen))를 사용하여 RNA (3ug)를 역전사시켰다. 퓨전 폴리머라제 (뉴 잉글랜드 바이오랩스(New England Biolabs)) 및 SYBR 그린 I을 사용하여 qPCR에 의해 cDNA 샘플을 분석하였다: (98℃에서 30초, 40x[98℃에서 10초, 66℃에서 15초, 72℃에서 15초]). SCN1A에 대한 프라이머 (5'-TGTCTCGGCATTGAGAACATTC-3' (서열식별번호: 185); 5'-ATTGGTGGGAGGCCATTGTAT-3' (서열식별번호: 186))를 사용하여 내인성 SCN1A 전사체의 수준을 정량화하고, 참조 유전자 (5'-ACCACAGTCCATGCCATCAC'-3' (서열식별번호: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (서열식별번호: 188))로서 GAPDH를 사용한 델타-델타 Ct 방법에 의해 SCN1A 발현의 상대 수준을 결정하였다. 데이터는 대조군 조건 대비 배수 변화로서 제시된다.

[0400] 결과는 도 1 및 하기 표 13에 대조군 조건 (예를 들어, EGFP-KASH 리포터 구축물) 대비 SCN1A 전사의 배수 변화로서 제시된다. 표 13은 대조군 조건 대비 전사의 적어도 1.5배 증가를 가져온 구축물에 대한 값을 보고한다.

[0401] 표 13. 상이한 게놈 표적 부위 및 상응하는 eTF가 전사에 미치는 효과. CON은 실험에 사용된 아연 핑거 구축물을 나타낸다 (표 1 참조). gRNA 구축물의 경우, gRNA가 상보적 DNA 가닥을 표적화하도록 설계되었기 때문에 표적 부위와 gRNA 서열은 동일하다.

CON	Chr 2 시작 위치	표적 서열	도 1로부터의 표적 부위	표적 부위에 대한 SEQ ID NO	평균	T검정
4	166149168	ctaggtcaagtgtaggag	z1	18	5.12090848	0.0096628
28	166149165	GGTCAAGTGTAGGAGACA	z8	25	1.52068773	0.62403349
3	166128025	taggtaccatagagtgag	z13	30	25.4730028	0.14942042
29	166127991	gaggatactgcagaggtc	z14	31	8.6766618	0.16432794
	166149176	aaggctgtctaggtcaagtgt	g9	35	1.36378425	0.18753821
	166149118	tgttcctccagattaacactt	g10	36	1.63040825	0.46710683
	166128002	gatgaagccgagaggatactg	g11	37	18.7579211	0.13148732
	166128037	gctgatttgtattaggtacca	g12	38	22.4892633	0.09291316
	166177299	AGAAAGCTGATACAGATACAA	g15	39	1.7542842	0.34519408
	166178880	ggtacgggcaaagatttcttg	g17	40	1.36801947	0.48762102
	166177299	AGAAAGCTGATACAGATACAA	g19	41	1.45636874	0.44464045
	166177369	ACACAATGAGCCACCTACAAG	g20	42	1.31187425	0.42605224
	166177362	GTGGCTCATTGTGTGTGTGCC	g22	43	1.25217773	0.26572657
	166155264	CATATCCCTGCAGGTTCAGAA	g24	44	1.75688991	0.28984533
	166155099	agagagagagagagagagaga	g25	45	2.05701745	0.42409102
	166155393	TTCTCAGTTTTGAAATTAAAA	g26	46	1.64498972	0.21705582
	166155255	TGGATTCTCTTCTGAACCTGC	g27	47	2.27026665	0.43195546
	166148361	TGCTGAGGCAGGACACAGTGT	g29	48	2.4290169	0.30364553
	166148843	ATCATCTGTAACCATCAAGGA	g30	49	2.58328006	0.0748197
	166148565	TCCTGCCTACTTAGTTTCAAG	g31	50	1.97097781	0.25980859
	166148953	ATTACAGTTCTGTCAGCATGC	g32	51	1.34500323	0.32186367
	166149373	TGGTCTCATTCTTTTTGTGGG	g33	52	1.71471378	0.32104302
	166142239	CGATATTTTCATGGATTCCTT	g34	53	1.7735976	0.21954265
	166142391	CTGACACTTACTTTGTCTAAA	g35	54	1.95513108	0.02069095
	166142219	AAAACTGGAACCGCATTCCCA	g36	55	2.08698135	0.0454403
	166142396	ACAAAGTAAGTGTCAGTGTGG	g37	56	1.30739959	0.72725347
	166142344	ATAATAGTTGTGTCTTTATAA	g38	57	1.55783618	0.29846459
	166141162	TGTACAAGCAGGGCTGCAAAG	g39	58	1.4663605	0.02946062
	166140590	GTTAACAAATACACTAAACAC	g40	59	1.37399196	0.33638238
	166140928	ttcaacaagctcccaagaagt	g42	60	1.46929899	0.24465271
	166141090	ATGTTCAAGGTGCAGAAGGAA	g43	61	2.04547409	0.09880194

[0402]

CON	Chr 2 시작 위치	표적 서열	도 1로부터의 표적 부위	표적 부위에 대한 SEQ ID NO	평균	T검정
	165990246	TGTTTGCTCAAACGTGCACCA	g44	62	2.13402102	0.25583999
	165989684	AAATAAGACATGAAAACAAGA	g45	63	1.27016182	0.32368695
	165990193	AAATATGTACCACAAGAAATG	g46	64	2.29522738	0.41829497
	165990076	TATCTGGTTTCTCTCACTGCT	g47	65	1.44542116	0.0947106
	165989571	ATTGCAAAGCATAATTTGGAT	g48	66	1.42246971	0.18117243

[0403]

[0404] 실시예 2

[0405] SCN1A 특이적 전사 인자를 사용한 HEK293 세포에서의 내인성 SCN1A의 상향조절

[0406] HEK293 세포를 표준 방법에 따라 배양하고, 6-웰 플레이트에 플레이팅하였다. 각각의 웰 내의 세포를 단일 조작된 전사 인자 구축물, WT 인간 CREB3 (서열식별번호: 211) 또는 EGFP 대조군 구축물을 보유하는 플라스미드 3ug으로 형질감염시켰다 (퓨진HD, 프로메가). 시험된 조작된 전사 인자 구축물은 구축물 1-27 및 46-53 (표 1) 및 CREB3-TRE (서열식별번호: 215; bZIP DNA 결합 도메인이 TET 프로모터-표적화 합성 ZF 도메인으로 대체된

CREB3)를 발현하는 플라스미드 (각각 개별적으로 시험됨)를 포함하였다. 형질감염 48시간 후에, 세포를 수집하고, RNA를 단리하고 (퀴아젠 RNeasy 미니 키트), DNase 처리하였다. 올리고DT 프라이머 (슈퍼스크립트 IV, 인비트로젠)를 사용하여 RNA (3ug)를 역전사시켰다. 퓨전 폴리머라제 (뉴 잉글랜드 바이오랩스) 및 SYBR 그린 I을 사용하여 qPCR에 의해 cDNA 샘플을 분석하였다: (98℃에서 30초, 40x[98℃에서 10초, 66℃에서 15초, 72℃에서 15초]). SCN1A에 대한 프라이머 (5'-TGTCTCGGCATTGAGAACATTC-3' (서열식별번호: 185); 5'-ATTGGTGGGAGGCCATTGTAT-3' (서열식별번호: 186))를 사용하여 내인성 SCN1A 전사체의 수준을 정량화하고, 참조 유전자 (5'-ACCACAGTCCATGCCATCAC'3' (서열식별번호: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (서열식별번호: 188))로서 GAPDH를 사용한 델타-델타 Ct 방법에 의해 SCN1A 발현의 상대 수준을 결정하였다. 데이터는 대조군 조건 대비 배수 변화로서 제시된다 (도 2a, 도 2b 및 도 2c 참조). 대조군 구축물은 RE 1 (서열식별번호: 1)의 제어 하에 발현된 EGFP로 이루어졌다. 조작된 전사 인자의 전달은 EGFP 조건에 비해 내인성 SCN1A 전사체에서 다양한 정도의 상향조절을 유도하였다.

[0407] 실시예 3

[0408] SCN1A 특이적 전사 인자를 사용한 GABA 뉴런에서의 내인성 SCN1A의 상향조절

[0409] iCell GABA 뉴런 (셀룰러 다이나믹스(Cellular Dynamics))을 6-웰 플레이트에 플레이팅하고 (~1E6개 세포/웰), 제조업체의 권장 프로토콜에 따라 유지하였다. 플레이팅 72시간 후에, 편재성 프로모터 (CBA 프로모터)의 제어 하에 EGFP 또는 도 3a의 활성인자 구축물 30 또는 도 3b의 구축물 25 또는 구축물 16을 발현하는 재조합 AAV (혈청형 AAV-DJ)를 대략 2E11 게놈 카피/웰로 배양 배지에 첨가하였다. 감염 후 1주 (도 3a) 또는 2주 (도 3b)에, RNA를 배양된 세포로부터 단리하고 (퀴아젠 RNeasy 미니 키트), DNase 처리하였다. 회수된 RNA를 올리고DT 프라이머 (슈퍼스크립트 IV, 인비트로젠)를 사용하여 역전사시켰다. 퓨전 폴리머라제 (뉴 잉글랜드 바이오랩스) 및 SYBR 그린 I을 사용하여 qPCR에 의해 cDNA 샘플을 분석하였다: (98℃에서 30초, 40x[98℃에서 10초, 66℃에서 15초, 72℃에서 15초]). SCN1A에 대한 프라이머 (5'-TGTCTCGGCATTGAGAACATTC-3' (서열식별번호: 185); 5'-ATTGGTGGGAGGCCATTGTAT-3' (서열식별번호: 186))를 사용하여 내인성 SCN1A 전사체의 수준을 정량화하고, 참조 유전자 (5'-ACCACAGTCCATGCCATCAC'3' (서열식별번호: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (서열식별번호: 188))로서 GAPDH를 사용한 델타-델타 Ct 방법에 의해 SCN1A 발현의 상대 수준을 결정하였다. 데이터는 대조군 조건 대비 배수 변화로서 제시된다 (도 3a 및 도 3b 참조). 조작된 전사 인자의 AAV-구동 발현은 배양된 iPS-유래 GABA 뉴런에서 내인성 SCN1A 전사체의 유의한 상향조절을 가져왔다.

[0410] 실시예 4

[0411] SCN1A 특이적 전사 인자를 사용한 GABA 뉴런에서의 내인성 SCN1A의 특이적 상향조절

[0412] iCell GABA 뉴런 (셀룰러 다이나믹스)을 6-웰 플레이트에 플레이팅하고 (~1E6개 세포/웰), 제조업체의 권장 프로토콜에 따라 유지하였다. 플레이팅 72시간 후에, CBA 프로모터의 제어 하에, CBA 프로모터에 의해 구동되는 VPR TAD에 융합된 아연 핑거 DBD를 포함하는, EGFP 또는 활성인자 (구축물 30)를 발현하는 재조합 AAV (혈청형 AAV-DJ)를 대략 2E11 게놈 카피/웰로 배양 배지에 첨가하였다.

[0413] 감염 1주 후에, RNA를 배양된 세포로부터 단리하고 (퀴아젠 RNeasy 미니 키트), DNase 처리하였다. 회수된 RNA로부터 TruSeq 가닥 mRNA 라이브러리 키트 (일루미나(Illumina))를 사용하여 RNAseq 라이브러리를 제조하고, 일루미나 NextSeq (2 x 75 사이클의 쌍형성된 말단 서열분석) 상에서 서열분석하였다. 서열분석 판독물을 인간 게놈 (RNASTAR)에 대해 정렬하고, DESeq2로 차등 발현 분석을 수행하였다. 데이터는 대조군 (AAVDJ-CBA-EGFP) 샘플 대비 배수 변화로서 제시된다 (도 4 참조). 결과는 표 14에 제시되고, 도 4는 대조군 조건 대비 배수 변화로서 제시된 내인성 SCN1A 및 40개의 가장 가까운 이웃 유전자 전사체의 상대 발현을 예시한다. 표 1에 기재된 바와 같은 구축물 30은 검사된 다른 유전자와 비교하여 SCN1A 유전자 또는 Nav1.1 단백질의 발현을 특이적으로 증가시킬 수 있었다. 이는 구축물 30의 전사 활성인자에 의해 인식되는 표적 부위가 SCN1A 유전자에 대해 특이적이어서, GABA 뉴런에서의 SCN1A 유전자 발현의 증가를 유발하였다는 것을 나타냈다.

[0414] 표 14. SCN1A 특이적 전사 인자 (구축물 30)로 처리된 GABA 뉴런에서의 내인성 SCN1A 및 40개의 가장 가까운 이웃 유전자의 전사에 대한 효과.

유전자 명칭	Chr 2 시작	Chr 2 종료	Chr 가닥	대조군 대비 배수 변화
PLA2R1	160788518	160919121	-	0.16367458
ITGB6	160956176	161128399	-	0.20679884
RBMS1	161128661	161350305	-	1.63514667
탱크	161993418	162092732	+	0.90946407
PSMD14	162164548	162268228	+	0.92699237
TBR1	162272604	162282381	+	0.53199642
SLC4A10	162280842	162841792	+	1.89407328
DPP4	162848750	162931052	-	2.82345284
FAP	163027193	163101661	-	2.26977379
IFIH1	163123588	163175213	-	1.46146481
GCA	163175349	163228105	+	2.58702426
FIGN	164449905	164592522	-	0.46785861
GRB14	165349321	165478358	-	0.5631965
COBLL1	165510133	165700189	-	0.43199257
SLC38A11	165752695	165812035	-	4.06730119
SCN3A	165944031	166060577	-	1.0807866
SCN2A	166095911	166248818	+	1.24475196
CSRNP3	166326156	166545917	+	0.82971233
GALNT3	166604100	166651192	-	0.33804418
TTC21B	166713984	166810353	-	1.58661143
SCN1A	166845669	166984523	-	62.9552975
SCN9A	167051694	167232503	-	1.71659087
SCN7A	167260082	167350757	-	0.29331967
B3GALT1	168675181	168730551	+	0.64436013
STK39	168810529	169104651	-	1.19821739
CERS6	169312371	169631644	+	0.86828378
NOSTRIN	169643048	169722024	+	1.82142718
SPC25	169690641	169769881	-	0.86880697
ABCB11	169779447	169887832	-	3.1441368
DHRS9	169921298	169952677	+	1.10381777

[0415]

유전자 명칭	Chr 2 시작	Chr 2 종료	Chr 가닥	대조군 대비 배수 변화
BBS5	170335687	170382432	+	0.65476347
KLHL41	170366211	170382772	+	0.87373377
FASTKD1	170386258	170430385	-	1.02786927
PPIG	170440849	170497916	+	1.09866236
CCDC173	170501934	170550943	-	0.67290779
PHOSPHO2	170550974	170558218	+	0.91339152
KLHL23	170550997	170633499	+	0.73926347
SSB	170648442	170668574	+	1.00631994
METTL5	170666590	170681441	-	1.21271497
UBR3	170683967	170940641	+	1.21350908
MYO3B	171034654	171511681	+	0.52839217

[0416]

[0417] 실시예 5

[0418] 생체내 발현 카세트로부터의 SCN1A의 발현

[0419] 생체내 SCN1A의 전사 활성인자의 발현을 시험하기 위해, 재조합 AAV9 벡터를 벡터 바이오랩스(Vector Biolabs)(펜실베니아주 말번)에 의해 생성하였다. 수컷 C57Bl/6 마우스 (군당 N=5, 7-8주령)에게 총 4개의 주사 부위에 대해, 배측 해마 (AP -2.0 mm, 측면 ±1.5, DV -1.4 mm, 경막으로부터) 및 복측 해마 (AP -3.1 mm, 측면 ±2.8, DV -3.8 mm, 경막으로부터) 내 양측으로 1.5 ul의 정제된 AAV 벡터를 주입하였다. AAV를 각각의 주사 후에 4m 휴지 기간으로 0.3 ul/분의 속도로 전달하였다. 처리 4주 후에, 마우스를 안락사시키고, 해마 조직을 해부하였다. 각각의 군에 대해, 대부분의 동물 (N=4)에서 좌측 및 우측 해마 조직 둘 다로부터의 조직을 수집하고 균질화를 위해 풀링하였으며, 예외로 1마리 동물에서는 좌측 해마만을 수집하고 균질화하였다. RNA를 균질물로부터 단리하고 (퀴아젠 RNeasy 미니 키트), DNase 처리하였다. 올리고DT 프라이머 (슈퍼스크립트 IV, 인비트로젠)를 사용하여 RNA (3μg)를 역전사시켰다. 퓨전 폴리머라제 (뉴 잉글랜드 바이오랩스) 및 SYBR 그린 I을 사용하여 qPCR에 의해 cDNA 샘플을 마우스 SCN1A의 발현에 대해 분석하였다: 98℃에서 30초, 40x[98℃에서 10초, 64℃에서 15초, 72℃에서 15초]. 마우스 SCN1A에 대한 프라이머 (5'-CAAAAAAGCCACAAAAGCCT-3' (서열식별번호: 189); 5'-TTAGCTCCGCAAGAAACATC-3' (서열식별번호: 190))를 사용하여 내인성 SCN1A 전사체의 수준을 정량화하고, 참조 유전자 (5'-ACCACAGTCCATGCCATCAC'3' (서열식별번호: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (서열식별번호: 188))로서 GAPDH를 사용한 델타-델타 Ct 방법에 의해 생체내 SCN1A 발현의 상대 수준을 결정하였다.

[0420] 도 5a 및 도 5b는 각각 AAV9 구축물이 주사된 5마리 동물의 평균 결과를 예시한다. eGFP 대조군 구축물은 eGFP 리포터 트랜스진을 포함하였다. 구축물 4 (표 1 참조)는 상기 표 1에 기재된 바와 같은 서열식별번호: 18을 포함하는 표적 서열을 인식하는 전사 활성인자를 포함하였다. 도 5a는 생체내 SCN1A의 상대 발현을 예시한다. 도 5b는 생체내 SCN1A 발현의 변화를 평균 eGFP 발현의 백분율로서 예시한다. 이들 결과는 발현 카세트 A의 SCN1A 전사 활성인자가 생체내에서 SCN1A 발현의 대략 20%-30% 상향조절을 유발하였다는 것을 나타냈다.

[0421] 이러한 발현 카세트는 드라베 증후군, 간질, 발작, 알츠하이머병, 파킨슨병 및/또는 SCN1A의 결핍 및/또는 손상된 활성과 연관된 임의의 다른 질환 또는 상태를 치료하기 위해 인간에서 사용되도록 적합화될 수 있다.

[0422] 실시예 6

[0423] 드라베 증후군의 마우스 모델에서의 고체온 발작 (HTS) 검정

[0424] A. 이형접합 Scn1a 녹아웃 마우스 모델

[0425] $Scn1a^{tm1Kea}$ 마우스 계통에서 발현 카세트를 사용한 드라베 증후군 및/또는 그의 증상의 치료를 시험하였다. 이러한 마우스 계통은 드라베 증후군에 대해 확립된 마우스 모델이다. $Scn1a^{tm1Kea}$ 마우스 계통은 CRE 레콤비나제를 요구하지 않는다. $Scn1a^{tm1Kea}$ 마우스 (잭슨 래보러토리(Jackson Laboratory)로부터 입수가능함; 문헌 [Hawkins et al., Scientific Reports, vol. 7: 15327 (2017)]에 기재됨)는 SCN1A의 제1 코딩 엑손의 결실을 포함한다. SCN1A 녹아웃 대립유전자에 대해 동형접합인 마우스는 진전, 운동실조, 발작을 특징으로 하고, 출생 후 제16일에 사망한다. C57BL/6 배경의 이형접합 마우스는 자발적 발작이 발생하고, 수주 내에 큰 백분율로 사망한다. 이러한 마우스 계통은 간질 및 드라베 증후군 치료의 안전성 및 효능을 연구하는데 사용될 수 있다. 추가의 정보에 대해서는 문헌 [Miller et al., Genes Brain Behav. 2014 Feb;13(2):163-72]을 참조한다.

[0426] $Scn1a^{tm1Kea}$ 마우스 계통에서 전사 활성인자의 효능을 시험하기 위해, 암컷 C57Bl/6J 사육동물과 +/- 교배된 수컷 Scn1a로부터 생산된 한배새끼에게 P1에서 양측 ICV를 통해 AAV 벡터를 투여하였다. 마우스에게 구축물 31-34 (표 1)를 투여하였다. 마우스를 P18에서의 이유시까지 그의 어미에 의해 방해받지 않게 둔 다음, 고체온 발작 (HTS) 검정을 개시할 때인 P26-P28까지 다시 방해받지 않게 두었다. 투여된 P1 마우스의 개별 한배새끼를 P18에서 이유시키고, 매일 사망을 관찰하였다. 혼합된 129Stac X C57BL/6 배경의 P26-P28 이형접합 (HET) 및 WT Scn1a 마우스에서 고체온 발작 유도를 수행하였다. 검정 전에 마우스는 가열 램프 (HL-1)와 직렬로 연결된 온도 제어 모듈 (TCAT 2DF, 피지템프(Physitemp))에 삽입 및 연결된 윤활 직장 온도 프로브 (Ret-4)를 가졌다. 이어서, 마우스를 큰 유리 비커에 넣고, 잠시 환경과 평형을 이루도록 하였다. 이후, 자세 상실이 동반된 제1 긴장성-간대성 발작의 발병까지 또는 43℃에 도달할 때까지 체온을 2분마다 ~0.5℃만큼 증가시켰다. 마우스가 자세 상실이 동반된 발작을 경험한 경우에, 실험을 종료하고, 마우스의 내부 체온을 기록하였다. 실험의 전체

과정에 걸쳐 자세 상실이 동반된 발작이 검출되지 않은 경우에, 그 마우스는 발작이 없는 것으로 간주하고, 검정을 종결하였다. P1에서 마우스로부터 조직 샘플을 수득하고, 실시간 PCR을 사용하여 실험 과정 동안 마우스의 유전자형결정을 수행하였다. 검정을 완료한 후 유전자형결정을 맹검해제하였으며, HET 또는 WT로서의 마우스의 상태는 수득된 데이터와 상관관계가 있었다. 데이터를 카플란-마이어 생존 곡선에 플롯팅하고, 유의성을 만텔-콕스 검정에 의해 결정하였다. 결과는 표 15 및 표 16 및 도 6a-e에 제시된다.

[0427] 표 15. 실시예 6에서 사용된 조건의 요약.

[0428]

구축물	투여량 (gc/마우스)
구축물 31 (도 6a & 6e)	6.0E+10
구축물 32 (도 6b, 6d, 6e)	3.1E+11
구축물 33 (도 6c)	5.8E+10
구축물 34 (도 6d)	4.9E+13

[0429] 표 16. 고체온 발작 검정의 결과의 요약.

[0430]

eTF 구축물 (도)	# 대조군 동물 (PBS 처리됨)	# 처리군 동물	42.6 ℃에서의 % 무발작	P 값
EGFP 리포터	16	N/A	44%	
구축물 31 (도 6a & 6e)	16	18	95	P<0.0001
구축물 32 (도 6b, 6d, 6e)	16	21	76	P<0.05
구축물 33 (도 6c)	16	14	93	P<0.001
구축물 34 (도 6d)	16	12	83	P<0.05

[0431] 상기 기재된 바와 같은 $Scn1a^{tm1Kea}$ 마우스에서 추가의 실험을 수행하여, HTS 검정에서 구축물 42 및 43을 발작에 대한 그의 효과에 대해 시험하였다. 이들 실험에서, 구축물 42를 P1에서 양측 ICV를 통해 $9x10^{10}$ gc/마우스로 투여하고, 구축물 43을 P1 또는 P5에서 양측 ICV를 통해 $6x10^{10}$ gc/마우스로 투여하였다. 결과는 도 6f (구축물 42) 및 도 6g (구축물 43)에 제시된다. 구축물 둘 다는 EGFP 대조군과 비교하여 발작의 유의한 감소를 나타냈다 (구축물 42 및 43 둘 다에 대해 $P < 0.0001$).

[0432] B. 이형접합 $Scn1a^{RX}$ 돌연변이 마우스 모델

[0433] $Scn1a^{RX}$ 마우스 계통에서 본 개시내용의 발현 카세트를 사용한 드라베 증후군 및/또는 그의 증상의 치료를 시험하였다. 이러한 마우스 계통은 드라베 증후군에 대해 확립된 마우스 모델이다. $Scn1a^{RX}$ 마우스 계통은 CRE 레콤비나제를 요구하지 않는다. $Scn1a^{RX}$ 마우스 (잭슨 래보러토리로부터 입수가능함; 문헌 [Ogiwara et al., J. Neuroscience, vol. 27: 5903-5914 (2007)]에 기재됨)는 SCN1A 유전자의 엑손 21에 기능 상실 단일 염기 넌센스 돌연변이 (CgG에서 TgA로; R1407X)를 포함한다. C57BL/6 배경의 이형접합 마우스는 자발적 발작이 발생하고, 수주 내에 큰 백분율로 사망한다.

[0434] $Scn1a^{RX}$ 마우스 계통에서 전사 활성인자의 효능을 시험하기 위해, $Scn1a^{RX/+}$ 정자와 CD-1 어미 내로 이식된 배아를 갖는 암컷 C57Bl/6J 난자의 수컷 및 IVF 교배로부터 생산된 한배새끼에게 P1에서 양측 ICV를 통해 AAV 벡터 (구축물 31)를 $5.1x10^{10}$ 게놈 카피 (gc)/마우스로 투여하거나 또는 PBS 대조군을 투여하였다. 마우스를 P18에서의 이유시까지 그의 어미에 의해 방해받지 않게 둔 다음, HTS 검정을 개시할 때인 P26-P28까지 다시 방해받지 않게 두었다. 투여된 P1 마우스의 개별 한배새끼를 P18에서 이유시키고, 매일 사망을 관찰하였다. C57BL/6 배경의 P26-P28 이형접합 (HET) 및 WT Scn1a 마우스에서 고체온 발작 유도를 수행하였다. 검정 전에 마우스는 가열 램프 (HL-1)와 직렬로 연결된 온도 제어 모듈 (TCAT 2DF, 피지템프(Physitemp))에 삽입 및 연결된 윤활 직장 온도 프로브 (Ret-4)를 가졌다. 이어서, 마우스를 큰 유리 비커에 넣고, 잠시 환경과 평형을 이루도록 하였다. 이후, 자세 상실이 동반된 제1 긴장성-간대성 발작의 발병까지 또는 43℃에 도달할 때까지 체온을 2분마다 ~0.5℃만큼 증가시켰다. 마우스가 자세 상실이 동반된 발작을 경험한 경우에, 실험을 종료하고, 마우스의 내부 체온을 기록하였다. 실험의 전체 과정에 걸쳐 자세 상실이 동반된 발작이 검출되지 않은 경우에, 그 마우스는 발작이 없는 것으로 간주하고, 검정을 종결하였다. P1에서 마우스로부터 조직 샘플을 수득하고, 실시간 PCR을 사

용하여 실험 과정 동안 마우스의 유전자형결정을 수행하였다. 검정을 완료한 후 유전자형결정을 맹검해제하였으며, HET 또는 WT로서의 마우스의 상태는 수득된 데이터와 상관관계가 있었다. 시험된 WT Scna1 마우스 중 어느 것도 발작을 경험하지 않았다. 구축물 31 처리된 (n=13) 및 PBS 대조군 처리된 (n=14) HET 마우스에 대한 데이터를 카플란-마이어 생존 곡선에 플롯팅하고, 유의성을 만텔-콕스 검정에 의해 결정하였다. 도 6h에 제시된 바와 같이, 구축물 31 처리된 HET 마우스는 PBS 대조군 처리된 HET 마우스에 비해 고체온 발작 유도의 유의한 감소를 나타냈다 ($P < 0.01$).

[0435] 실시예 7

[0436] 드라베 증후군의 마우스 모델에서의 생존 검정

[0437] A. 이형접합 Scn1a 녹아웃 마우스 모델

[0438] $Scn1a^{tm1Kea}$ 마우스 계통에서 전사 활성인자의 효능을 시험하기 위해, 암컷 C57Bl/6J 사육동물과 +/- 교배된 수컷 Scn1a로부터 생산된 한배새끼에게 P1에서 양측 ICV를 통해 AAV 벡터를 투여하였다. 마우스를 이유시까지 그의 어미에 의해 방해받지 않게 두었다. Scn1a +/- 마우스의 건강 상태의 관찰을 P18에서의 이유 후 매일 수행하였다. 어떤 원인으로든 그의 홈 케이지에서 죽은 채로 발견된 마우스는 그 날짜를 기록하였다. 데이터를 카플란-마이어 생존 곡선에 플롯팅하고, 유의성을 만텔-콕스 검정에 의해 결정하였다.

[0439] 결과는 표 17 및 도 7a-d에 제시된다.

[0440] 표 17. 생존 검정에 대한 조건 및 결과의 요약.

[0441]

SEQ ID	투여량 (gc/마우스)	# 대조군 동물 (PBS 처리됨)	# 처리군 동물	P100 (*P83)에서의 % 생존	P 값
PBS	N/A	53	N/A	49%	
구축물 33 (도 7c & 7d)	5.8E+10	53	29	76%	P<0.05
구축물 31 (도 7b & 7d)	6.0E+10	53	34	97%	P<0.0001

[0442] B. 이형접합 $Scn1a^{RX}$ 돌연변이 마우스 모델

[0443] $Scn1a^{RX}$ 마우스 계통에서 전사 활성인자의 효능을 시험하기 위해, $Scn1a^{RX/+}$ 정자와 CD-1 어미 내로 이식된 배아를 갖는 암컷 C57Bl/6J 난자의 수컷 및 IVF 교배로부터 생산된 한배새끼에게 P1에서 양측 ICV를 통해 AAV 벡터 (구축물 31)를 $5.1x10^{10}$ 게놈 카피 (gc)/마우스로 투여하거나 또는 PBS 대조군을 투여하였다. 마우스를 이유시까지 그의 어미에 의해 방해받지 않게 두었다. $Scn1a^{RX/+}$ 마우스의 건강 상태의 관찰을 P18에서의 이유 후 매일 수행하였다. 어떤 원인으로든 그의 홈 케이지에서 죽은 채로 발견된 마우스는 그 날짜를 기록하였다. 구축물 31 처리군 (n=27) 및 대조군 처리군 (n=18)에 대한 데이터를 카플란-마이어 생존 곡선에 플롯팅하고, 유의성을 만텔-콕스 검정에 의해 결정하였다.

[0444] 도 7e에 제시된 바와 같이, 구축물 31 처리된 $Scn1a^{RX/+}$ 마우스는 PBS 대조군 처리된 $Scn1a^{RX/+}$ 마우스에 비해 증가된 생존을 나타냈다 ($P < 0.0001$).

[0445] 실시예 8

[0446] SCN1A 특이적 전사 인자를 코딩하는 AAV로의 처리 후 비-인간 영장류에서의 SCN1A 전사 수준

[0447] 연구는 2 내지 3세의 수컷 시노몰구스 마카크 (마카카 파시쿨라리스(macaca fascicularis))를 사용하였다. 동물을 연구에 등록하기 전 세포-기반 중화 항체 검정에 의해 AAV9에 대한 교차-반응성 항체에 대해 사전스크리닝하였다. SCN1A 특이적 전사 인자 (구축물 33)를 발현하는 AAV9 또는 대조군을 PBS 중에 희석하고, 1.2E12 gc/동물로 실질내로 주사하였다. 각각의 반구에서 3개의 상이한 정위 좌표, 동물당 6개의 주사 부위를 주사에 대해 확인하였다. 부위당 10 ul 부피를 주사하였다. 우측 반구에서의 주사는 좌측에서의 주사와 대칭이었다. 2마리의 비처리 동물을 대조군으로서 사용하였다.

[0448] Scn1A mRNA 발현을 평가하기 위해, 역전사에 이어서 qPCR 방법을 수행하였다. 투여 28일 후의 부검에서, 대조군 및 처리군 동물로부터의 다양한 뇌 영역으로부터의 조직 절편 (전두 피질, 두정 피질, 측두 피질, 후두 피질, 해마, 수질, 소뇌; 각각 200 mg)을 RNAlater에 수집한 다음, 동결시켰다. 간략하게, 30 mg의 조직을 해부하고, RNA를 추출하고 (퀴아젠 Rneasy 지질 조직 미니 키트, 카탈로그 # 1023539 사용), 역전사에 의해 cDNA로 전환시키고 (어플라이드 바이오시스템즈(Applied Biosystems) 고용량 cDNA 역전사 키트, 카탈로그 # 4368814 사용), Scn1A 및 하우스키핑 유전자 GAPDH (어플라이드 바이오시스템즈, 카탈로그 # Rh02621745-gI FAM)에 대한 프라이머/프로브 세트를 사용하여 qPCR을 수행하였다.

[0449] SCN1A에 대한 프라이머/프로브 세트가 하기에 제공된다.

[0450] 표 18. 실시예 8에 사용된 프라이머 서열.

[0451]

유전자	SEQ ID NO	서열 (5'-3')	설명
Scn1A	191	CCATGGAACTGGCTCGATTTCAC	F-프라이머
	192	ATTGGTGGGAGGCCACTGTAT	R-프라이머
	193	AGGCCTGAAAACCATTGTGGGAGCCCT	프로브 (FAM)

[0452] 비교 Ct (ΔCt) 방법을 사용하여 상대 정량화 (RQ)에 의해 각각의 시험 샘플에서의 Scn1A의 유전자 발현을 결정하였다. 이 방법은 표적 유전자와 하우스키핑 유전자 사이의 Ct 차이 (ΔCt)를 측정한 다음, 처리 샘플의 ΔCt 값을 대조군 샘플과 비교한다.

[0453] ΔCt = 표적 유전자의 Ct 평균 - 하우스키핑 유전자의 Ct 평균

[0454] ΔΔCt = 처리 샘플의 ΔCt - 대조군 샘플의 ΔCt

[0455] 상대 발현 (처리 샘플) = $2^{-\Delta\Delta Ct}$

[0456] 데이터는 뇌로부터의 상이한 조직 절편에서의 표적 mRNA의 정규화된 발현으로서 보고된다 (도 8 참조). 도 8에 예시된 바와 같이, 실질내 주사 부위에 근접한 뇌 내의 부위가 가장 높은 수준의 SCN1A 전사체 발현을 나타냈다.

[0457] 실시예 9

[0458] PV 선택적 프로모터 및 마이크로RNA 결합 부위를 갖는 AAV로의 처리 후 비-인간 영장류의 PV 뉴런에서의 선택적 트랜스진 발현

[0459] 연구는 연구 등록 전 AAV9에 대한 교차-반응성 항체에 대해 사전스크리닝한 6마리의 마모셋 (칼리트릭스 자쿠스 (Callithrix jacchus))을 사용하였다. 2마리의 원숭이를 EF1알파 프로모터의 제어 하에 EGFP 트랜스진을 함유하는 AAV9로 처리하고, 2마리의 원숭이를 RE 2 (서열식별번호: 2)의 제어 하에 EGFP 트랜스진을 함유하는 AAV9로 처리하고, 2마리의 원숭이를 RE 2 (서열식별번호: 2)의 제어 하에 EGFP 트랜스진을 함유하고 또한 EGFP의 코딩 영역과 폴리A 테일 사이에 위치한 마이크로RNA 결합 부위 (서열식별번호: 7)를 함유하는 AAV9로 처리하였다. AAV9 벡터를 PBS 중에 희석하고, 동물을 동물당 총 6개의 주사 부위에 대해 각각의 반구의 해마/내후각 피질 내로의 3회 뇌내 주사 (각각 2 uL)로 처리하였다. EF1알파-EGFP를 함유하는 AAV9 벡터로 처리된 2마리의 동물은 각각 5.8E+11 gc/동물의 총 용량을 받았고, RE 2-EGFP를 함유하는 AAV9 벡터로 처리된 2마리의 동물은 각각 3.0E+11 gc/동물의 총 용량을 받았으며, RE 2 + m1-EGFP를 함유하는 AAV9 벡터로 처리된 2마리의 동물은 각각 2.3E+11 gc/동물의 총 용량을 받았다.

[0460] 면역조직화학을 사용하여 파르브알부민 (PV) 선택적 발현을 평가하였다. 투여 28일 후의 부검에서, 뇌의 다양한 영역으로부터의 조직 절편을 수집하였다. 부유 마모셋 뇌 절편 (35um)을 4% 파라포름알데히드 중에 고정시키고, 완충제 (PBS, 3% BSA, 3% 당나귀 혈청, 0.3% 트리톤-X 100, 0.2% 트윈-20)로 차단한 다음, 항-GFP (압캠(Abcam) ab290)에 이어서 알렉사-488에 접합된 2차 항체 (써모 A21206)로 염색하였다. 이어서, 항-PV 항체 (스완트(Swant)), 및 알렉사-647에 접합된 2차 항체 (써모 A31571) 및 4',6-디아미디노-2-페닐인돌 (DAPI)이 이어졌다. 절편을 탑재하고, 퍼킨엘머 벡트라3을 사용하여 영상화하였다.

[0461] 결과는 도 9a-f 및 10a-l에 제시된다.

[0462] 실시예 10

[0463] eTF^{SCN1A} 생체분포

[0464] 본 연구의 목적은 일측 뇌실내 (ICV) 주사를 통해 4.8E+13의 용량으로 투여되었을 때 어린 시노몰구스 마카크 원숭이의 중추 신경계 (CNS)에서의 eTF^{SCN1A}의 생체분포를 비교하는 것이었다. 각각의 동물에게 GABA 선택적 조절 요소의 제어 하에 eTF^{SCN1A}를 코딩하는 발현 카세트 (RE^{GABA}-eTF^{SCN1A})를 함유하는 AAV9를 주사하였다. AAV9 입자를 PBS + 0.001% 플루론산 중에 제제화하고, 4.8E+13 또는 8E+13 vg/동물의 용량으로 투여하였다. 2 ml 부피의 제제화된 바이러스 입자를 각각의 동물에게 투여하였다. 연구 설계는 표 19에 제시된다.

[0465] 24개월령 시노몰구스 마카크 원숭이를 표 19에 나타낸 바와 같이 군분류하였다. 연구의 개시 전에, 상기 기재된 NAb 역가 검정을 사용하여 동물로부터의 혈액 샘플을 AAV9에 대한 중화 항체 역가의 수준에 대해 시험하였다. 항체에 대해 낮은 또는 음성 결과를 갖는 동물을 연구용으로 선택하였다. 표준 외과적 절차를 사용하여 ICV 주사를 통해 샘플을 투여하였다. 해동된 투여 물질을 잠시 습윤 얼음 상에 저장하고, 투여 직전에 실온으로 가온하였다. 동물을 마취시키고, 수술을 위해 준비하고, MRI 호환 정위 프레임 (Kopf)에 탑재하였다. 기준선 MRI를 수행하여 표적 좌표를 확립하였다. 절개부를 만들고, 표적 위치 위의 두개골을 통해 단일 구멍을 뚫었다. 36" 마이크로-보어 연장 세트에 부착된 3 mL BD 시린지를 샘플과 함께 준비하고, 주입 펌프에 넣었다. 연장 라인을 프라이밍하였다. 경막을 개방하고, 투여 바늘을 구멍으로부터 13.0 내지 18.1 mm의 깊이로 전진시켰다. 조영 매체 주사 및 형광투시법을 사용하여 우측뇌실 내로의 척추 바늘의 배치를 확인하였다. 3.0" 22g 퀸케 BD 척추 후버 포인트 바늘에 조영제를 충전하여 배치를 결정한 후 프라이밍된 연장 라인 및 시린지를 부착하였다. 펌프 설정은 19 내지 20분 동안 0.1 mL/분이었다. 투여 후 완충제를 수동으로 밀어내어 연장 라인을 제거하였다. 주입 완료 후 1 내지 2분 동안 바늘을 제자리에 유지시킨 다음, 바늘을 회수하였다. 비히클 및 시험 물품을 제1일에 1회 투여하고, 대상체를 27- 또는 29-일 회복 기간 동안 유지하였다.

[0466] 표 19. 생체분포 연구 설계

군	성별	ID	용량 (VG/동물)
1군 (완충제 대조군)	M	21001	
	F	11501	
2군 (RE^{GABA}-eTF^{SCN1A})	M	2001	4.8E+13
	F	2501	4.8E+13
	M	3001	4.8E+13
	M	3002	8E+13

[0467]

[0468] 투여 후, 연구 지속기간에 걸쳐 동물을 일상적으로 모니터링하고, 혈액 샘플을 주기적으로 회수하였다. eTF^{SCN1A} 투여는 어떤 예상밖의 사망, 임상 소견 또는 육안 관찰과도 연관되지 않았다. AAV9-RE^{GABA}-eTF^{SCN1A} 처리된 동물은 제28일 ± 2일에 예정된 부검까지 생존하였다. 매일 또는 매주 신체 검사 동안 임상 또는 행동 징후, 체온 증가 또는 체중 감소가 관찰되지 않았다. AAV9-RE^{GABA}-eTF^{SCN1A} 처리된 동물에서 간 트랜스아미나제 (ALT 및 AST)의 일시적 상승이 관찰되었지만, 면역조정 없이 연구의 종료시까지 완전히 해소되었고, 혈청 빌리루빈 또는 알칼리성 포스파타제의 동반 증가는 나타나지 않았다. 다른 측정된 임상 화학 종점은 주목할 만하지 않았다. 현미경 관찰은 간 조직병리학 연구에서 보고되지 않았다. CSF 백혈구는 말단 수집물에서 처리전 값에 비해 상승하였지만, 대조군과 AAV9-RE^{GABA}-eTF^{SCN1A} 처리된 동물 사이에 대등하였다. AAV9-RE^{GABA}-eTF^{SCN1A} 연관된 뇌척수액 세포증가증은 관찰되지 않았다. 모든 동물에 걸친 비-뉴런 조직의 거시적-관찰 및 상세한 미세-조직병리학 검사는 주목할 만하지 않았다. 조직은 주요 말초 기관 (즉, 심장, 폐, 비장, 간 및 생식선)을 포함하였다. 뉴런 조직의 거시적-관찰 및 상세한 미세-조직병리학은 어떤 주목할 만한 소견도 나타내지 않았다. 조직은 뇌, 척수 및 연관된 후근 신경절 (경부, 흉부 및 요추 영역으로부터)을 포함하였다. 연구는 전문 신경병리학 장소에 있는 병리학자를 포함한 3명의 독립적인 병리학자에 의해 수행하였다.

[0469] 항-AAV9 캡시드 중화 항체가 투여 4주 후에 관찰되었기 때문에, AAV9의 ICV 투여는 투여후 혈청에서 면역 반응을 방지하지 않았다. 그러나, CSF에서의 중화 항-AAV9 항체 수준은 변화하지 않고 투여전 수준과 대등하게 유지되었다 (표 20).

[0470] 표 20. AAV9 혈청 NAb 역가

대상체 수	AAV9 혈청 NAb 역가			AAV9 CSF NAb 역가	
	스크린-전	주사시	주사 4-주 후	주사시	주사 4-주 후
21001	1:5	< 1:5	< 1:5	< 1:5	< 1:5
11501	< 1:5	< 1:5	< 1:5	< 1:5	< 1:5
2001	< 1:5	< 1:5	1:405	< 1:5	1:5
2501	< 1:5	< 1:5	1:135	< 1:5	1:5
3001	< 1:5	< 1:5	1:1215	< 1:5	< 1:5
3002	< 1:5	< 1:5	1:135	< 1:5	< 1:5

[0471]

[0472] 예정된 부검 동안 주요 기관 (심실, 간엽, 폐심장엽, 신장, 비장, 췌장 및 자궁경부 림프절)으로부터 투여 27-29일 후 샘플을 수집하였다. 8 밀리미터 펀치를 통해 펀칭물을 수집하고, 하기 논의된 바와 같이 추가로 가공하였다.

[0473] 실시예 11

[0474] 뇌에서의 eTF^{SCN1A}의 생체분포

[0475] ddPCR을 사용하여 뇌에서의 eTF^{SCN1A} 생체분포를 측정하였다. 시노몰구스 마카크 뇌 조직의 다양한 영역으로부터의 샘플 (FC: 전두 피질; PC: 두정 피질; TC: 측두 피질; Hip: 해마; Med: 수질; OC: 후두 피질)을 벡터 카피수에 대해 측정하여, 일측 ICV에 의해 AAV9로 투여되었을 때 GABA 선택적 조절 요소의 제어 하의 eTF^{SCN1A} (RE^{GABA}-eTF^{SCN1A})의 생체분포를 평가하였다. 조직 DNA를 DNeasy 혈액 & 조직 키트 (퀴아젠)로 단리하였다. DNA 양을 결정하고, UV 분광광도계를 사용하여 정규화하였다. 20 나노그램의 조직 DNA를, 프로브용 ddPCR 슈퍼 믹스 (dUTP 없음) (바이오-라드(Bio-Rad)) 및 eTF^{SCN1A} 서열의 영역에 대하여 지시된 택맨 프라이머 및 프로브와 함께 20 마이크로리터 반응물에 첨가하였다. 액적을 생성하고, 자동화 액적 발생기 및 써모 사이클러 (바이오-라드)를 사용하여 주형을 증폭시켰다. PCR 단계 후에, 플레이트에 로딩하고, QX2000 액적 판독기에 의해 판독하여 조직 내의 벡터 카피수를 결정하였다. 원숭이 알부민 (MfAlb) 유전자는 게놈 DNA 함량을 정규화하기 위한 내부 대조군으로서의 역할을 하였으며, 이를 동일한 반응에서 증폭시켰다. eTF^{SCN1A} 및 MfAlb에 대한 프라이머 및 프로브는 표 21에 제시된다.

[0476] 표 21. eTF^{SCN1A} 및 MfAlb에 대한 프라이머 및 프로브

프라이머 / 프로브 명칭		서열 (5'-3')
eTF^{SCN1A}	eTF^{SCN1A} 정방향 프라이머	GAATGTGGGAAATCATTCAGTCGC (SEQ ID NO: 194)
	eTF^{SCN1A} 역방향 프라이머	GCAAGTTATCCTCTCGTGAGAAGG (SEQ ID NO: 195)
	eTF^{SCN1A} 프로브	GCGACAACCTGGTGAGACATCAACGCACC (SEQ ID NO: 196)
Mf알부민	MfAlb 정방향 프라이머	GCTGTTATCTCTTGTGGGCTGT (SEQ ID NO: 197)
	MfAlb 역방향 프라이머	AAACTCATGGGAGCTGCCGGTT (SEQ ID NO: 198)
	MfAlb 프로브	CCACACAAATCTCTCCCTGGCATTG (SEQ ID NO: 199)

[0477]

[0478] eTF^{SCN1A}은 동물당 4.8E+13 바이러스 게놈으로 투여되었을 때 뇌 전반에 걸쳐 광범위하게 분포되었고, 평균 1.3-3.5 VG/이배체 게놈이었다 (도 11). 추가로, 동물당 4.8E+13 바이러스 게놈으로 투여된 RE^{GABA}-eTF^{SCN1A}의 뇌 전체에 걸친 유전자 전달을, 다양한 용량으로 ICV를 통해 투여된 eGFP의 뇌 전체에 걸친 유전자 전달과 비교하였을 때, 용량 증가에 따라 VG/이배체 게놈의 증가가 관찰되었다. 이는 뇌에서의 유전자 전달이 ICV를 통해 AAV9로 투여되었을 때 용량-의존적 방식으로 발생하였다는 것을 나타냈다.

[0479] 실시예 12

[0480] 뇌에서의 eTF^{SCN1A} 전사

[0481] GABA 선택적 조절 요소 RE^{GABA} (RE^{GABA}-eTF^{SCN1A})의 제어 하의 eTF^{SCN1A}의 전사를, ddPCR-기반 유전자 발현 검정을 사용하여 eTF^{SCN1A} mRNA를 측정함으로써 평가하였다. 뇌 조직에 대해 RNeasy 플러스 미니 키트 (퀴아젠) 또는 RNeasy 지질 조직 미니 키트 (퀴아젠)로 조직 RNA를 단리하였다. RNA 양을 결정하고, UV 분광광도계를 사용하여 정규화하고, 바이오애널라이저 RNA 칩을 사용하여 RNA 품질 (RIN)을 검사하였다. 1 마이크로그램의 조직 RNA를 DNase 처리, 및 ezDNase™ 효소 키트 (써모 피셔)를 수반한 슈퍼스크립트 VILO cDNA 합성 키트에 의한 cDNA 합성에 사용하였다. 50 마이크로그램의 RNA를 cDNA로 전환시켰다. cDNA를 프로브용 ddPCR 슈퍼 믹스 (dUTP 없음) (바이오-라드) 및 eTF^{SCN1A} 서열의 영역에 대하여 지시된 택맨 프라이머 및 프로브와 함께 20 마이크로리터 반응물에 첨가하였다 (표 22). 액적을 생성하고, 자동화 액적 발생기 및 써모 사이클러 (바이오-라드)를 사용하여 주형을 증폭시켰다. PCR 증폭 후에, 플레이트에 로딩하고, QX2000 액적 판독기에 의해 판독하여 조직에서의 유전자 발현 수준을 제공하였다. 원숭이 유전자 ARFGAP2 (MfARFGAP2) (써모 피셔 사이언티픽)는 유전자 발현 수준을 정규화하기 위한 내인성 대조군으로서의 역할을 하였으며, 이를 동일한 반응에서 증폭시켰다. ARFGAP2에 대한 평균 전사체는 1.85E+6/ug RNA였다 (도 12, 상부 경계선). 검출 한계는 하부 경계선에 의해 표시된다.

[0482] eTF^{SCN1A} mRNA가 모든 동물에서 뇌 전반에 걸쳐 관찰되었으며, 이는 GABA-선택적 프로모터 RE^{GABA}가 모든 AAV9-RE^{GABA}-eTF^{SCN1A} 처리된 마카크의 뇌 조직에서 전사상 활성이었다는 것을 나타낸다 (도 12). FC: 전두 피질; PC: 두정 피질; TC: 측두 피질; Hip: 해마; Med: 수질; OC: 후두 피질.

[0483] 표 22. eTF^{SCN1A} 서열의 영역에 대하여 지시된 택맨 프라이머 및 프로브

프라이머 / 프로브 명칭	설명	서열 (5'-3')
eTF^{SCN1A}	eTF^{SCN1A} 정방향 프라이머	GAATGTGGGAAATCATTCAGTCGC (SEQ ID NO: 200)
	eTF^{SCN1A} 역방향 프라이머	GCAAGTTATCCTCTCGTGAGAAGG (SEQ ID NO: 201)
	eTF^{SCN1A} 프로브	GCGACAACCTGGTGAGACATCAACGCACC (SEQ ID NO: 202)
Mf*ARFGAP2*	정방향, 역방향 프라이머, 프로브	써모 피셔 (Cat#: 4448491)

[0484]

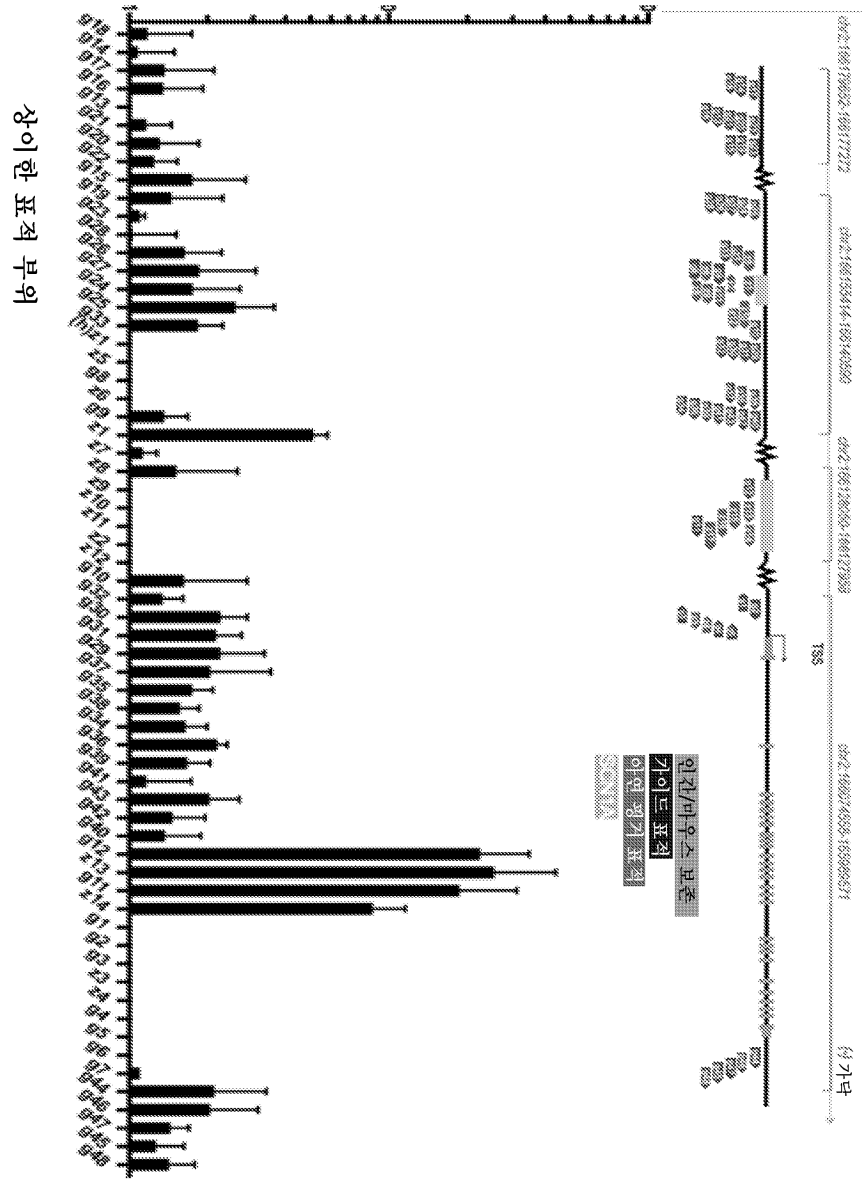
[0485] 실시예 13

[0486] 말초 조직에서의 eTF^{SCN1A} 생체분포 및 전사

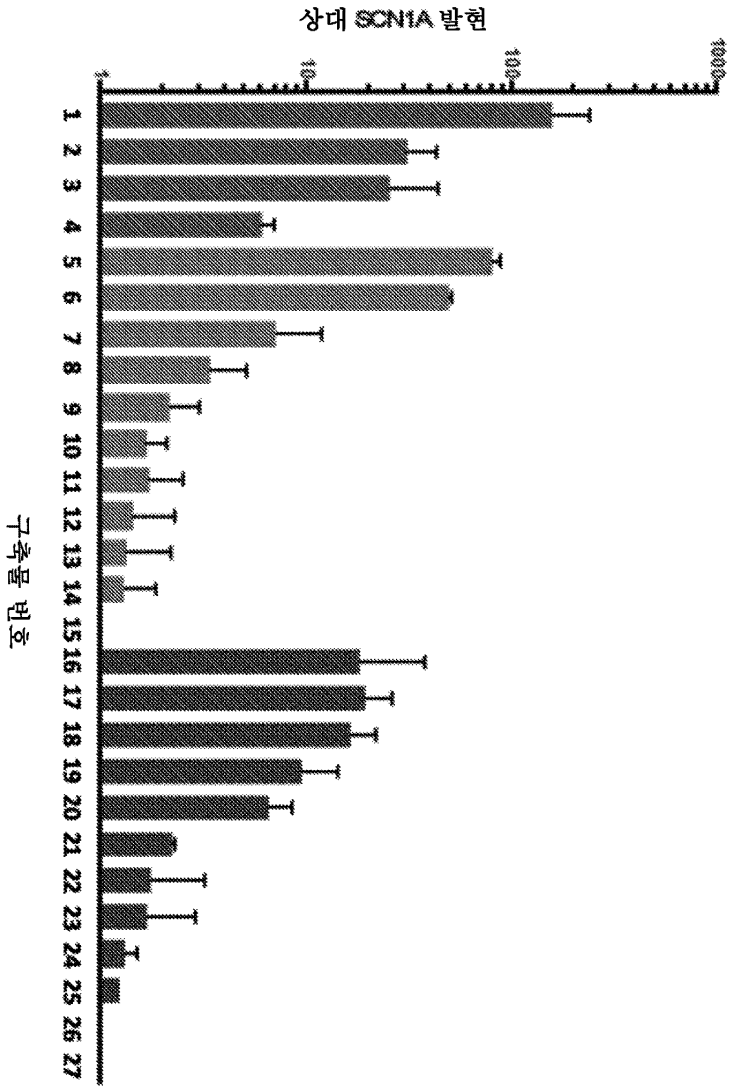
[0487] 추가로 다양한 기관에서 벡터 카피수를 측정하여, 일측 ICV에 의해 AAV9로 투여되었을 때 신체 전반에 걸친 조직에서의 RE^{GABA}-eTF^{SCN1A}의 형질도입을 평가하였다. 또한 ddPCR에 의해 eTF^{SCN1A}의 전사체 수준을 측정하여, 일측 ICV에 의해 AAV9로 투여되었을 때 신체 전반에 걸친 조직에서 GABA-선택적 조절 요소 RE^{GABA}의 제어 하의 eTF^{SCN1A} 전사 활성을 평가하였다. 두 방법은 일반적으로 상기 기재된 바와 같이 수행하였다. 척수 (SC) 및 후근 신경절 (DRG)에서의 RE^{GABA}-eTF^{SCN1A} 형질도입 및 eTF^{SCN1A}의 전사는 뇌에서 관찰된 수준과 대등하였다. 간을 제외하고, RE^{GABA}-eTF^{SCN1A} 형질도입은 뇌 외부의 말초 조직에서 보다 낮았다 (도 13). 간에서의 RE^{GABA}-eTF^{SCN1A}의 형질도입은 뇌에서보다 더 높았다. eTF^{SCN1A}의 전사는 심장, 폐 및 생식선을 포함한 말초 조직에서 검출되지 않았다. 그러나, 간에서의 eTF^{SCN1A} 전사체 수준은 뇌에서 측정된 eTF^{SCN1A}의 수준과 대등하였다. 추가로, 간에서의 eTF^{SCN1A} 전사는 존재하는 벡터 카피 수에 대해 정규화될 때 극도로 낮다 (뇌에서의 eTF^{SCN1A}의 전사와 비교하여 대략 1000배 더 낮음). 종합하면, 이는 GABA-선택적 조절 요소 RE^{GABA}의 제어 하의 eTF^{SCN1A}의 전사가 CNS로 제한된다는 것을 입증하였다.

도면

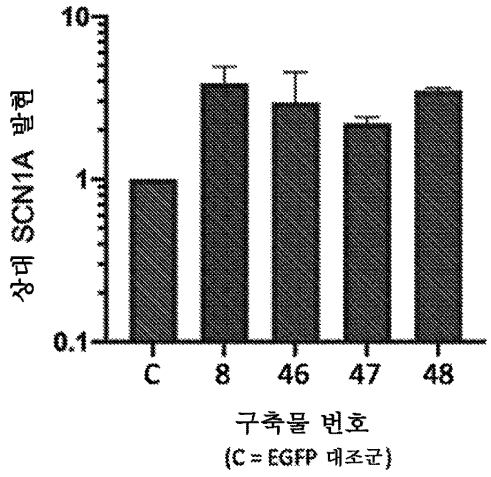
도면1



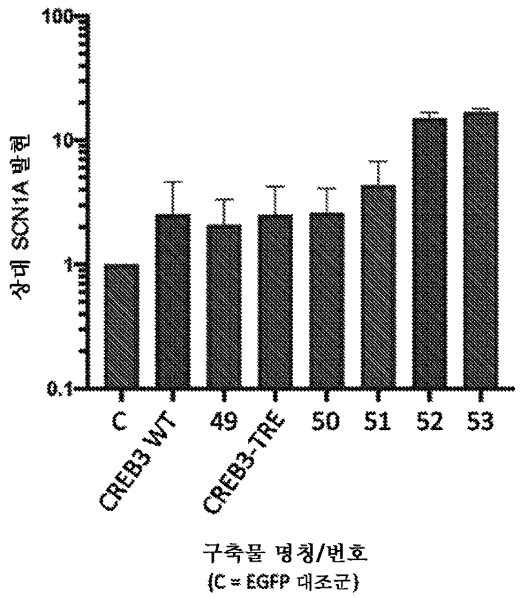
도면2a


상대 SCN1A 발현
1
10
100
1000
화합물 번호
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

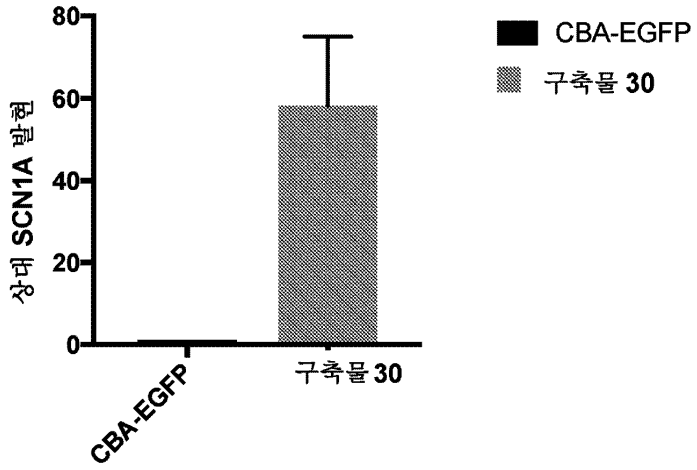
도면2b


10
1
0.1
상대 SCN1A 발현
C
8
46
47
48
구축물 번호
(C = EGFP 대조군)

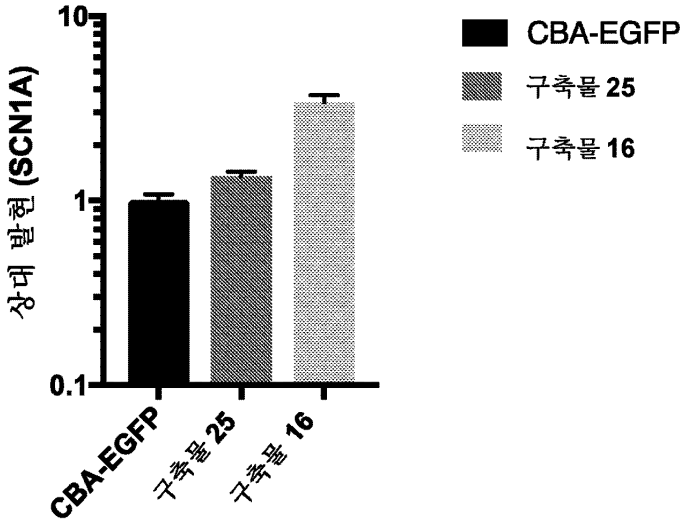
도면2c


100
10
1
0.1
상대 SCN1A 발현
C
CREB3 WT
49
CREB3-TRE
50
51
52
53
구축물 명칭/번호
(C = EGFP 대조군)

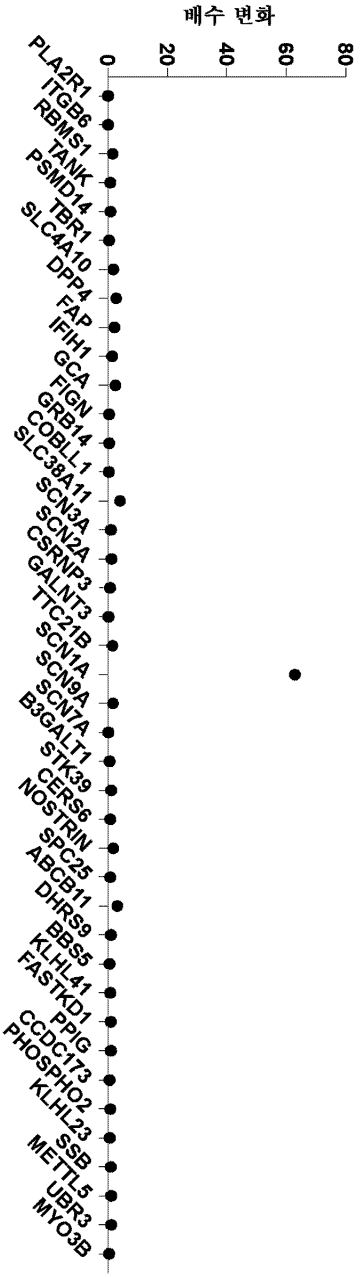
도면3a


80
60
40
20
0
상대 SCN1A 발현
CBA-EGFP
구축물 30
CBA-EGFP
구축물 30

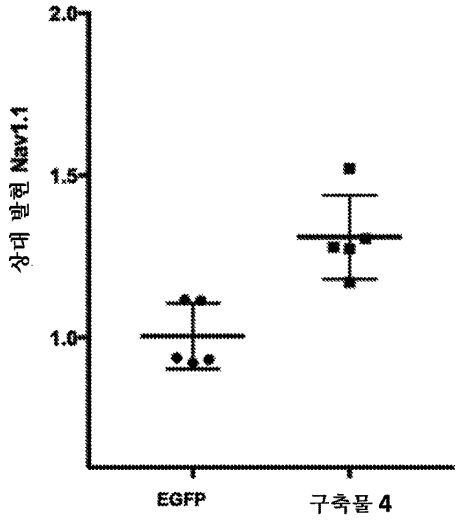
도면3b



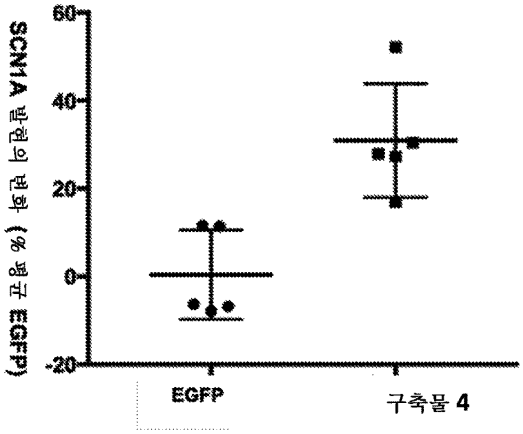
도면4


배수 변화
0
20
40
60
80
PLA2R1
ITGB6
RBMS1
TANK
PSMD14
TBR1
SLC4A10
DPP4
FAP
IFIH1
GCA
FIGN
GRB14
COBLL1
SLC38A11
SCN3A
SCN2A
CSRNP3
GALNT3
TTC21B
SCN1A
SCN9A
SCN7A
B3GALT1
STK39
CERS6
NOSTRIN
SPC25
ABCB11
DHRS9
BBS5
KLHL41
FASTKD1
PPIG
CCDC173
PHOSPHO2
KLHL23
SSB
METTL5
UBR3
MYO3B

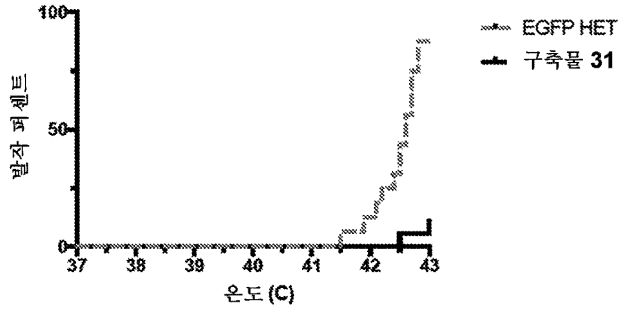
도면5a



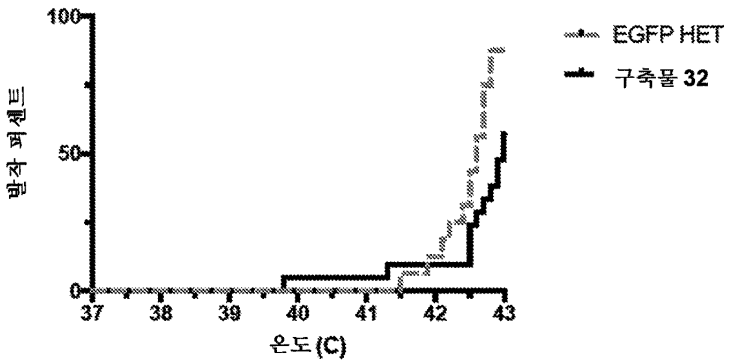
도면5b



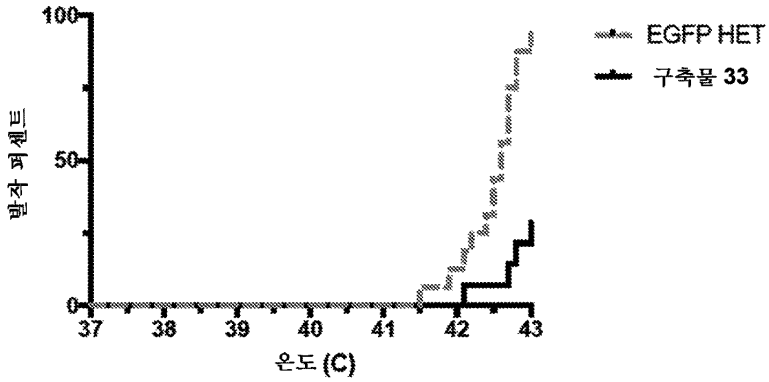
도면6a



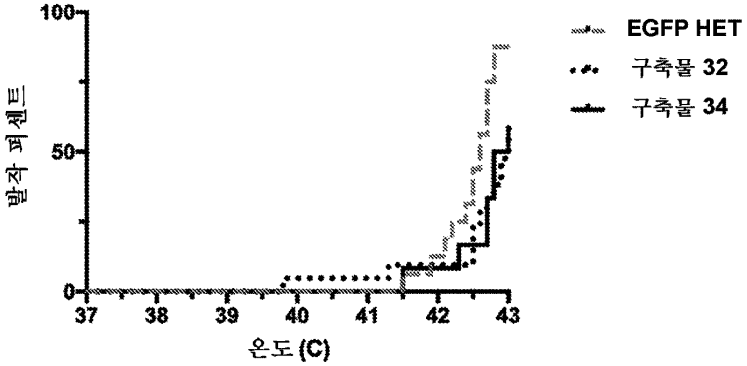
도면6b



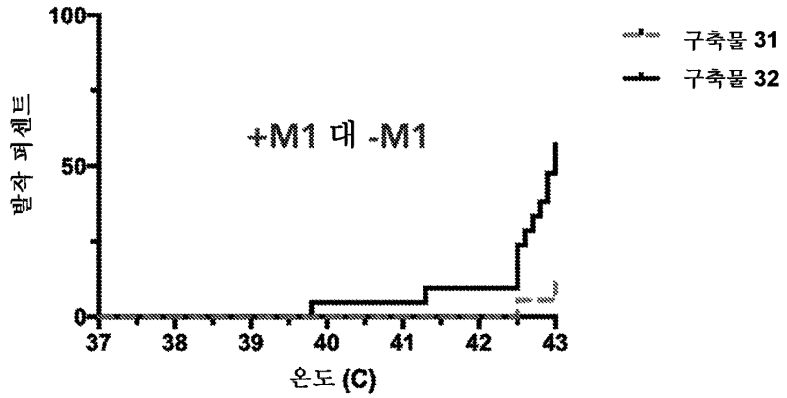
도면6c



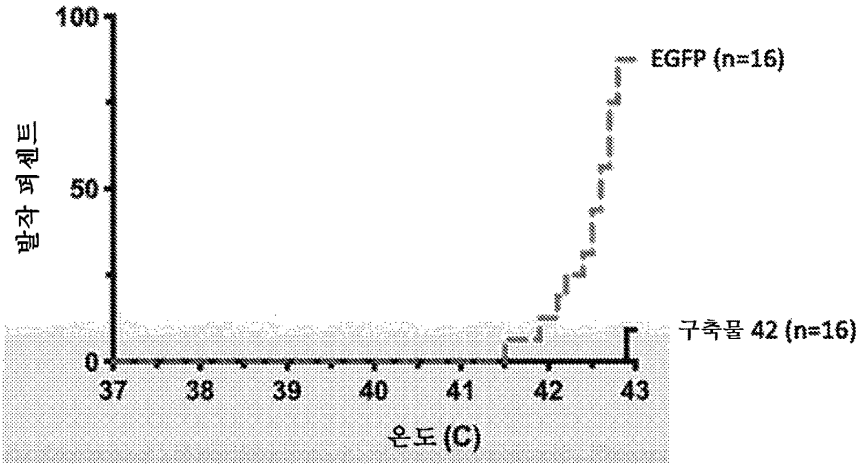
도면6d



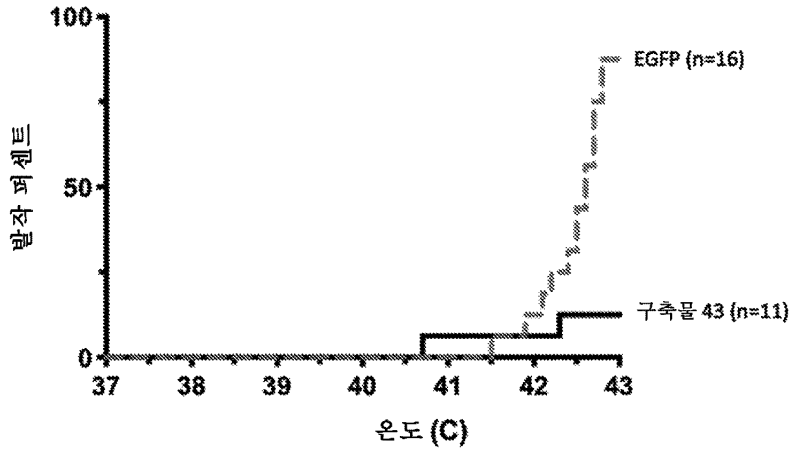
도면6e



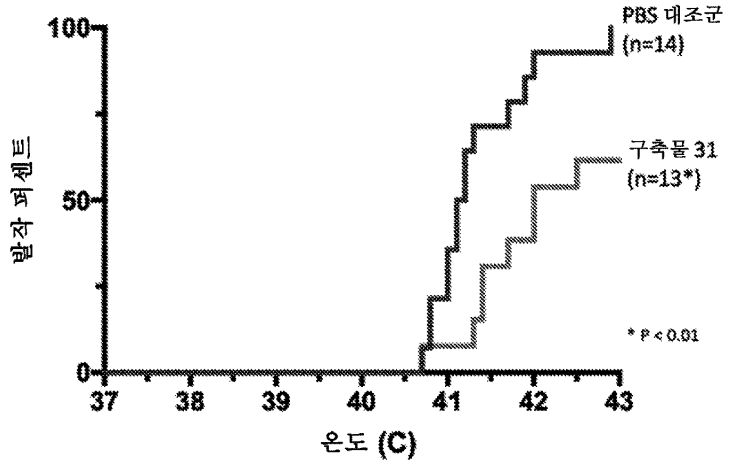
도면6f



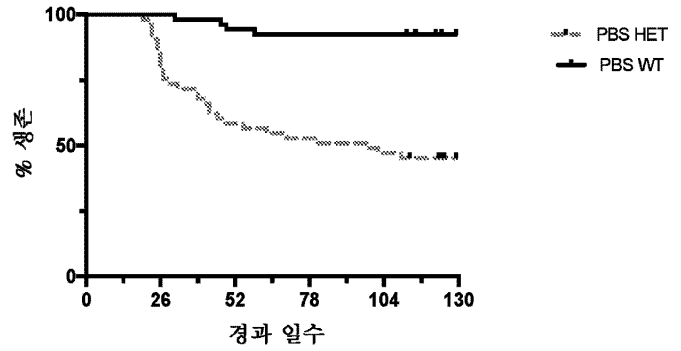
도면6g



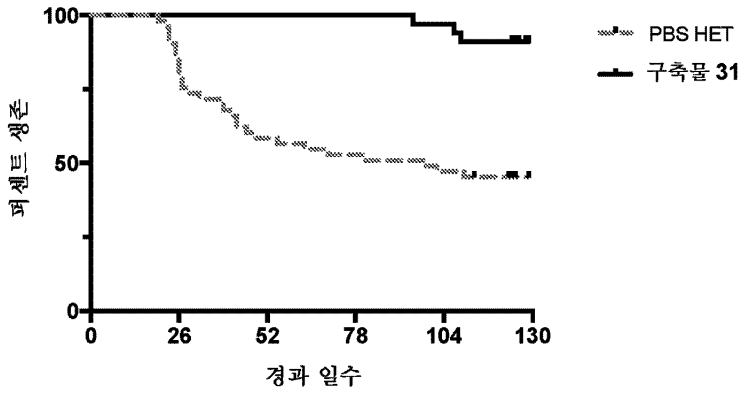
도면6h



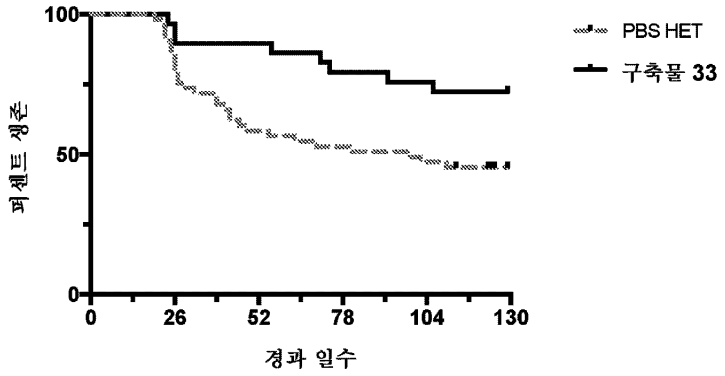
도면7a



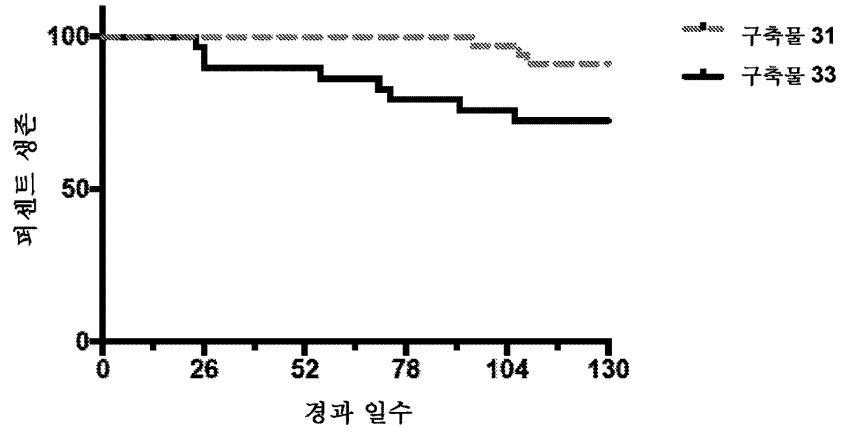
도면7b



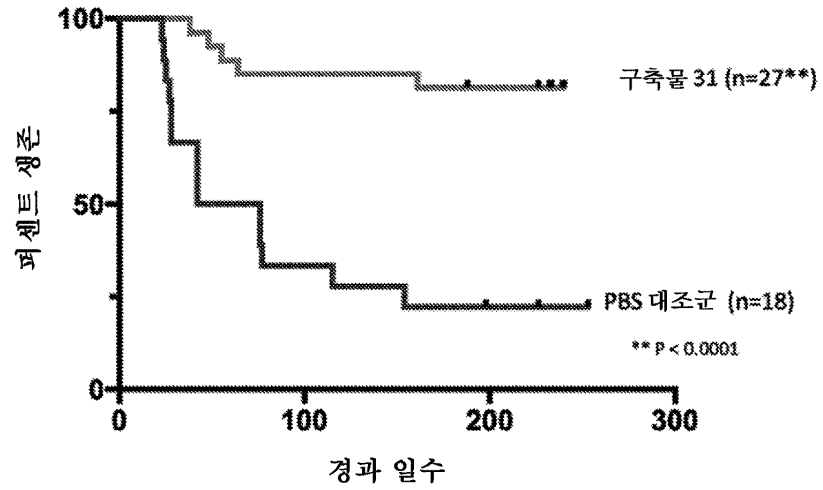
도면7c



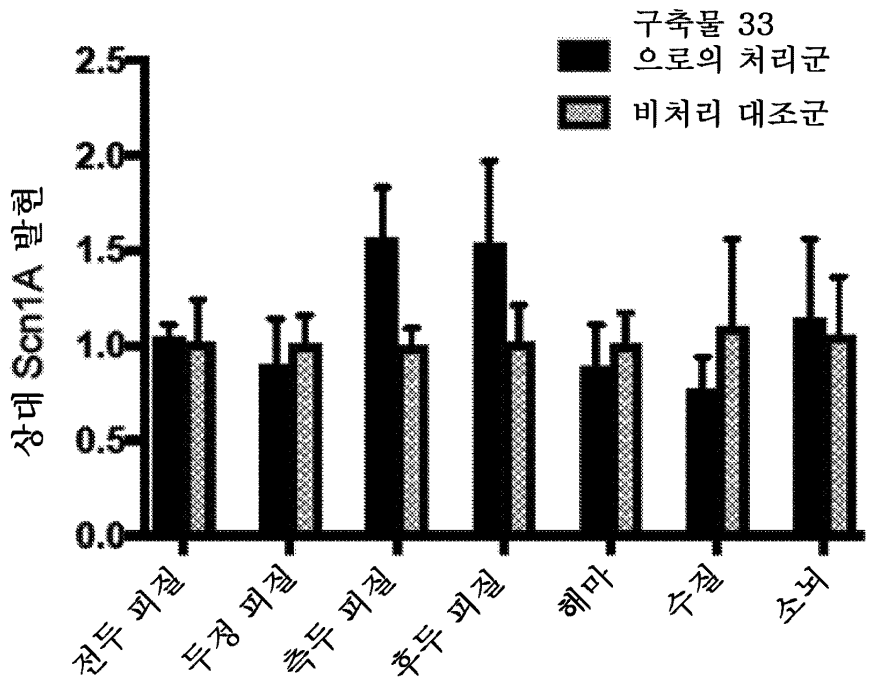
도면7d



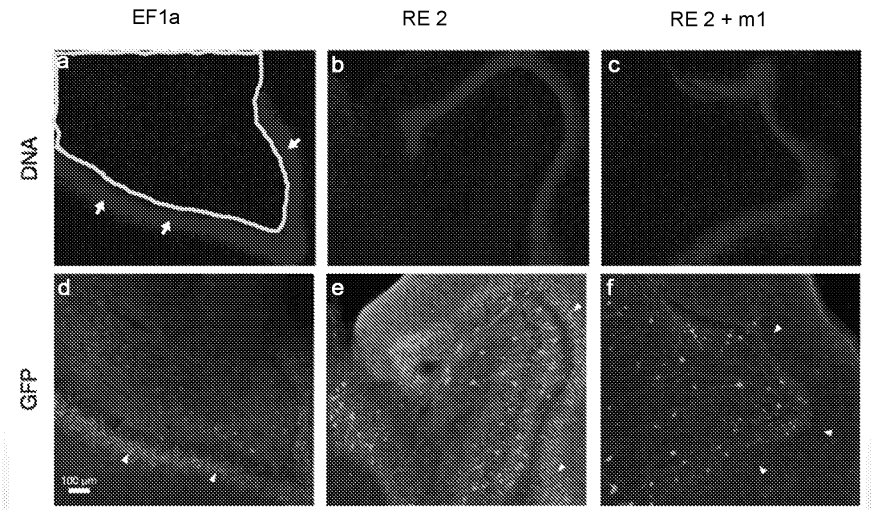
도면7e



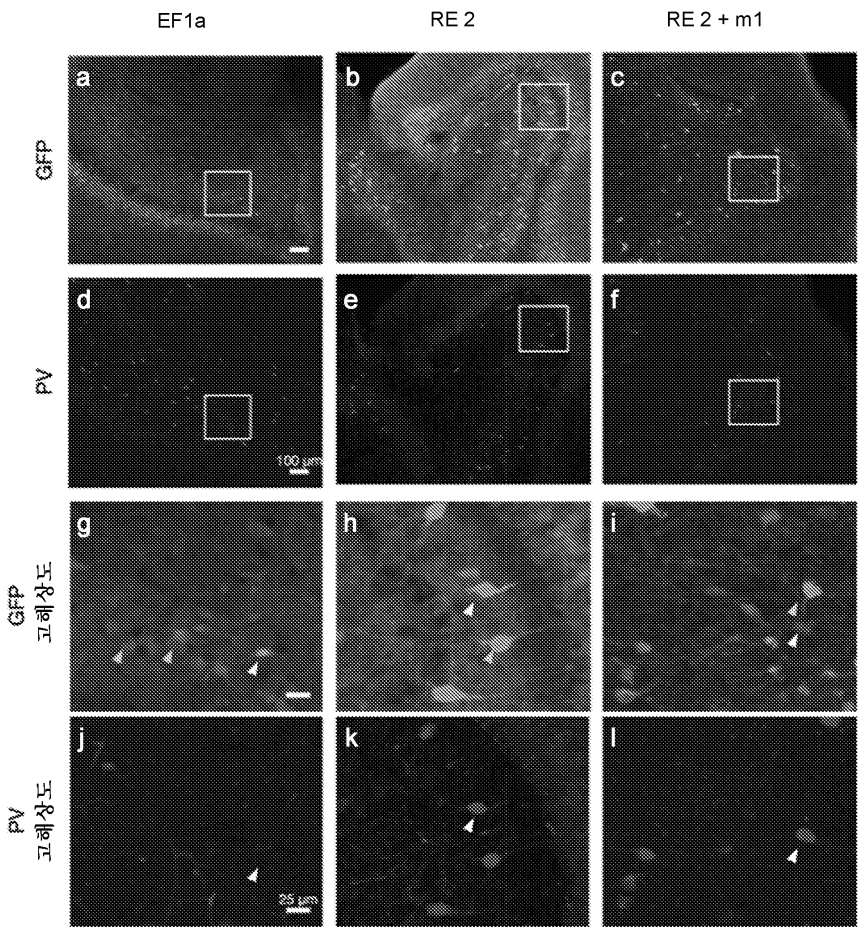
도면8


구축물 33
으로의 처리군
비처리 대조군
상대 Scn1A 발현
2.5
2.0
1.5
1.0
0.5
0.0
전두 피질
두정 피질
측두 피질
후두 피질
해마
수질
소뇌

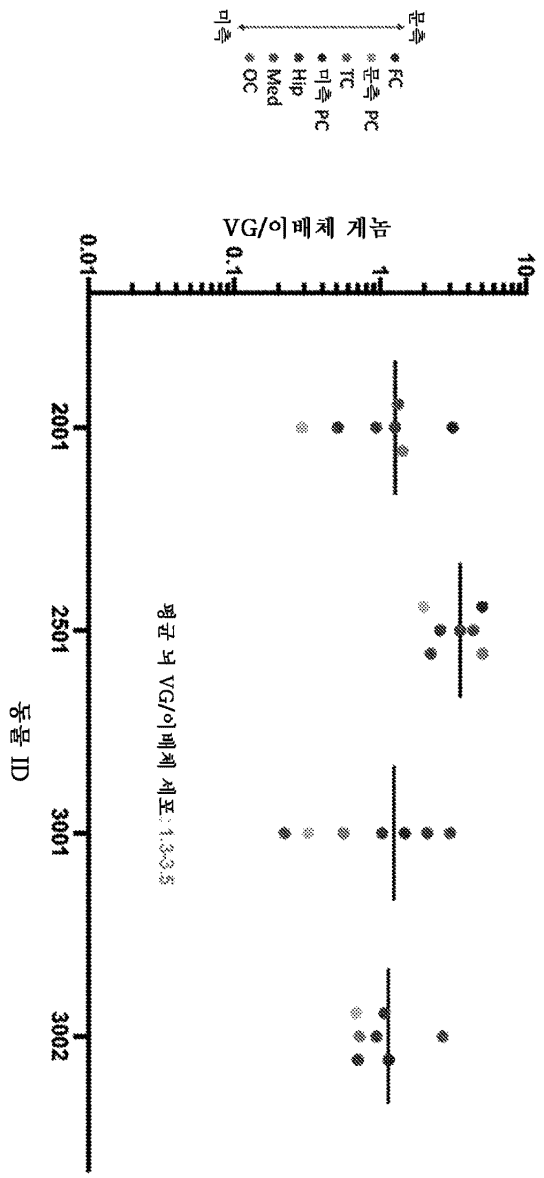
도면9


EF1a
RE 2
RE 2 + m1
DNA
GFP
a
b
c
d
e
f

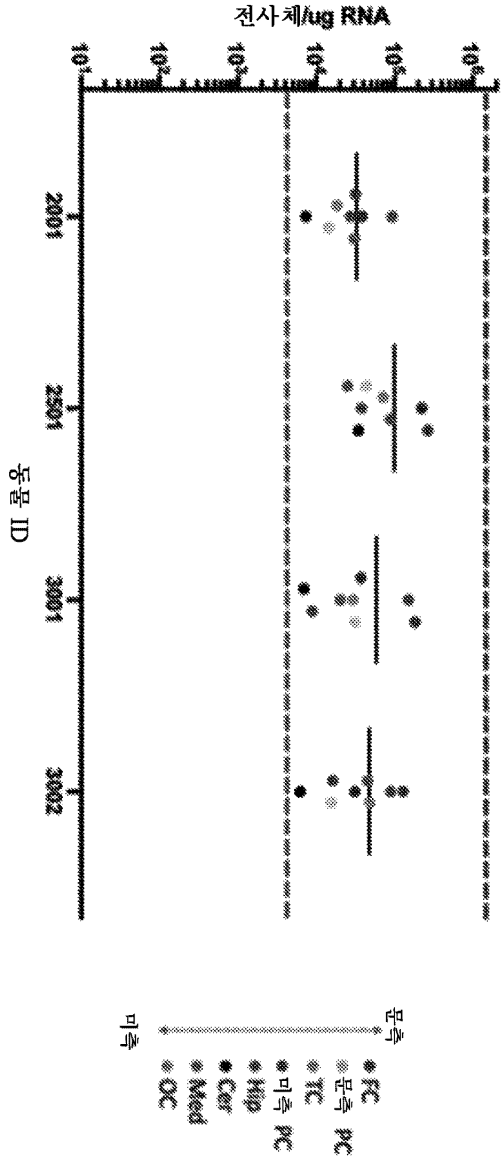
도면10


EF1a
RE 2
RE 2 + m1
GFP
PV
GFP 고해상도
PV 고해상도
a
b
c
d
e
f
g
h
i
j
k
l
100 μm
25 μm

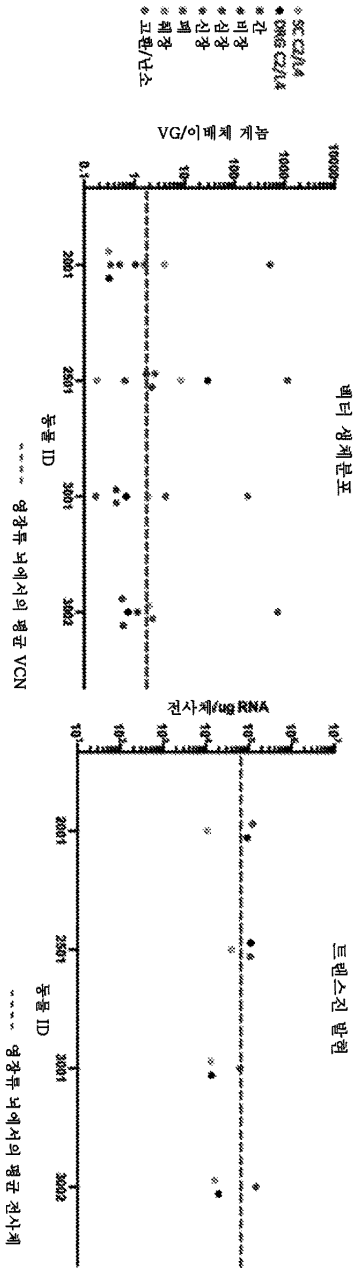
도면11


문측
미측
FC
문측 PC
TC
미측 PC
Hip
Med
OC
VG/이베체 계놈
0.01
0.1
1
10
2001
2501
3001
3002
동물 ID
평균 뇌 VG/이베체 세포: 1.3-3.5

도면12


전사체/ug RNA
동물 ID
2001
1002
1001
2002
많음
미늑
PC
공늑 PC
미늑 PC
HB
Cor
Mod
OC

도면13



서 열 목 록

SEQUENCE LISTING

<110> ENCODED THERAPEUTICS, INC.

<120> COMPOSITIONS AND METHODS FOR SELECTIVE GENE REGULATION

<130> 46482-724.601

<140><141><150> 63/008,569

<151> 2020-04-10

<150> 62/857,727

<151> 2019-06-05

<150> 62/854,238

<151> 2019-05-29

<160> 226

<170> PatentIn version 3.5

<210> 1

<211> 259

<212> DNA

<213> Artificial Sequence

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

<400> 1

gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt 60

ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac 120

tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg 180

tgggaggtct atataagcag agctggtacc gtgtgtatgc tcaggggctg ggaaaggagg 240

ggagggagct ccggctcag 259

<210> 2

<211> 2051

<212> DNA

<213> Artificial Sequence

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

<400> 2

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct g 2051

<210> 3

<211> 1878

<212> DNA

<213> Artificial Sequence

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

<400> 3

tcaacagggg gacacttggg aaagaaggat ggggacagag ccgagaggac tgttacacat 60

tagagaaaca tcagtgactg tgccagcttt ggggtagact gcacaaaagc cctgaggcag 120

cacaggcagg atccagtctg ctggtcccag gaagctaacc gtctcagaca gagcacaaag 180

caccgagaca tgtgccacaa ggcttgtgta gagaggtcag aggacagcgt acaggtccca 240

gagatcaaac tcaacctcac caggcttggc agcaagcctt taccaaccca cccccacccc 300

acccaccctg cacgcgcccc tctcccctcc ccatggtctc ccatggctat ctcacttggc 360

cctaaaatgt ttaaggatga cactggctgc tgagtggaaa tgagacagca gaagtcaaca 420

gtagatttta ggaaagccag agaaaaaggc ttgtgctgtt tttagaaagc caagggacaa 480

gctaagatag ggcccaagta atgctagtat ttacatttat ccacacaaaa cggacgggcc 540

tccgctgaac cagtgaggcc ccagacgtgc gcataaataa cccctgcgtg ctgcaccacc 600

tggggagagg gggaggacca cggtaaatgg agcgagcgca tagcaaaagg gacgcggggt 660

ccttttctct gccggtggca ctgggtagct gtggccaggt gtggtacttt gatggggccc 720

agggctggag ctcaaggaag cgtcgcaggg tcacagatct gggggaaccc cggggaaaag 780

cactgaggca aaaccgccgc tcgtctccta caatatatgg gagggggagg ttgagtacgt 840

tctggattac tcataagacc ttttttttt ccttccgggc gcaaaaccgt gagctggatt 900

tataatcgcc ctataaagct ccagaggcgg tcaggcacct gcagaggagc cccgccgctc 960

cgccgactag ctgcccccgc gagcaacggc ctcgtgattt ccccgccgat ccggtccccg 1020

cctccccact ctgcccccgc ctaccccgga gccgtgcagc cgcctctccg aatctctctc 1080

ttctcctggc gctcgcgtgc gagagggaac tagcgagaac gaggaagcag ctggaggtga 1140

cgccgggcag attacgcctg tcagggccga gccgagcgga tcgctgggcg ctgtgcagag 1200

gaaaggcggg agtgcccggc tcgctgtcgc agagccgagg tgggtaagct agcgaccacc 1260

tggacttccc agcgcccaac cgtggctttt cagccaggtc ctctcctccc gcggcttctc 1320

aaccaacccc atcccagcgc cggccaccca acctcccgaa atgagtgctt cctgccccag 1380

cagccgaagg cgctactagg aacggtaacc tgttactttt ccaggggccg tagtcgaccc 1440

gctgcccgag ttgctgtgcg actgcgcgcg cggggctaga gtgcaaggtg actgtggttc 1500

ttctctggcc aagtccgagg gagaacgtaa agatatgggc ctttttcccc ctctcacctt 1560

gtctcaccaa agtccctagt ccccggagca gttagcctct ttctttccag ggaattagcc 1620

agacacaaca acgggaacca gacaccgaac cagacatgcc cgccccgtgc gccctccccg 1680

ctcgctgcct ttcctccctc ttgtctctcc agagccggat cttcaagggg agcctccgtg 1740

cccccggctg ctcagtccct ccggtgtgca ggaccccgga agtcctcccc gcacagctct 1800

cgcttctctt tgcagcctgt ttctgcgccg gaccagtcga ggactctgga cagtagaggc 1860

cccgggacga ccgagctg 1878

<210> 4

<211> 509

<212> DNA

<213> Artificial Sequence

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

<400> 4

gccctctagg ccacctgacc aggtcccctc agtccccccc ttcccacact cccacactca 60

gccccctcc ccccccccg acccctgcag gattatcctg tctgtgttcc tgactcagcc 120

tgggagccac ctgggcagca ggggccaagg gtgtcctaga agggacctgg agtccacgct 180

gggccaagcc tgccctttct ccctctgtct tccgtccctg cttgcggttc tgctgaatgt 240

ggttatttct ctggctcctt ttacagagaa tgctgctgct aattttatgt ggagctctga 300

ggcagtgtaa ttggaagcca gacaccctgt cagcagtggg ctcccgtcct gagctgccat 360

gcttcctgct ctcctcccgt cccggctcct catttcatgc agccacctgt cccagggaga 420

gaggagtcac ccaggcccct cagtccgccc cttaaataag aaagcctccg ttgctcggca 480

cacataccaa gcagccgctg gtgcaatct 509

<210> 5

<211> 1644

<212> DNA

<213> Artificial Sequence

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

<400> 5

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300

catgggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 360

ccccaatttt gtatttattt atttttttaat tattttgtgc agcgatgggg gcgggggggg 420

gggggggcgcg cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga 480

ggtgcggcgg cagccaatca gagcggcgcg ctccgaaagt ttccttttat ggcgaggcgg 540

cggcggcggc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcct 600

tcgccccgtg ccccgctccg cgccgcctcg cgccgcccgc cccggctctg actgaccgcg 660

ttactcccac aggtgagcgg gcgggacggc ccttctcctc cgggctgtaa ttagcgcttg 720

gtttaatgac ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag 780

ggccctttgt gcgggggggа gcggctcggg gggtgcgtgc gtgtgtgtgt gcgtggggag 840

cgccgcgtgc ggcccgcgct gcccggcggc tgtgagcgct gcgggcgcgg cgcggggctt 900

tgtgcgctcc gcgtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg 960

gggctgcgag gggaacaaag gctgcgtgcg gggtgtgtgc gtgggggggt gagcaggggg 1020

tgtgggcgcg gcggtcgggc tgtaaccccc ccctgcaccc ccctccccga gttgctgagc 1080

acggcccggc ttcgggtgcg gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg 1140

gcgggggggtg gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg 1200

gctcgggggа ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg 1260

cagccattgc cttttatggt aatcgtgcga gagggcgcag ggacttcctt tgtcccaaat 1320

ctggcggagc cgaaatctgg gaggcgccgc cgcacccccт ctagcgggcg cgggcgaagc 1380

ggtgcggcgc cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc 1440

gtccccttct ccatctccag cctcggggct gccgcagggg gacggctgcc ttcgggggggg 1500

acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa 1560

ccatgttcat gccttcttct ttttcctaca gctcctgggc aacgtgctgg ttgttgtgct 1620

gtctcatcat tttggcaaag aatt 1644

<210> 6

<211> 1335

<212> DNA

<213> Artificial Sequence

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

<400> 6

gagtaattca tacaaaagga ctcgcccctg ccttggggaa tcccagggac cgtcgttaaa 60

ctcccactaa cgtagaaccc agagatcgct gcgttcccgc cccctcaccc gcccgctctc 120

gtcatcactg aggtggagaa gagcatgcgt gaggctccgg tgcccgtcag tgggcagagc 180

gcacatcgcc cacagtcccc gagaagttgg ggggaggggt cggcaattga accggtgcct 240

agagaaggtg gcgcggggta aactgggaaa gtgatgtcgt gtactggctc cgccttttc 300

ccgagggtgg gggagaaccg tatataagtg cagtagtcgc cgtgaacgtt ctttttcgca 360

acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg cctggcctct 420

ttacgggtta tggcccttgc gtgccttgaa ttacttccac gcccctggct gcagtacgtg 480

attcttgatc ccgagcttcg ggttggaagt gggtgggaga gttcgaggcc ttgcgcttaa 540

ggagcccctt cgcctcgtgc ttgagttgag gcctggcttg ggcgctgggg ccgccgcgtg 600

cgaatctggt ggcaccttcg cgcctgtctc gctgctttcg ataagtctct agccatttaa 660

aatttttgat gacctgctgc gacgcttttt ttctggcaag atagtcttgt aaatgcgggc 720

caagatctgc acactggtat ttcggttttt ggggccgcgg gcggcgacgg ggcccgtgcg 780

tcccagcgca catgttcggc gaggcggggc ctgcgagcgc ggccaccgag aatcggacgg 840

gggtagtctc aagctggccg gcctgctctg gtgcctggcc tcgcgccgcc gtgtatcgcc 900

ccgccctggg cggcaaggct ggcccggtcg gcaccagttg cgtgagcgga aagatggccg 960

cttcccggcc ctgctgcagg gagctcaaaa tggaggacgc ggcgctcggg agagcgggcg 1020

ggtgagtcac ccacacaaag gaaaagggcc tttccgtcct cagccgtcgc ttcatgtgac 1080

tccacggagt accgggcgcc gtccaggcac ctcgattagt tctcgagctt ttggagtacg 1140

tcgtctttag gttgggggga ggggttttat gcgatggagt ttccccacac tgagtgggtg 1200

gagactgaag ttaggccagc ttggcacttg atgtaattct ccttggaatt tgccctttt 1260

gagtttggat cttggttcat tctcaagcct cagacagtgg ttcaaagttt ttttcttcca 1320

tttcaggtgt cgtga 1335

<210> 7

<211> 214

<212> DNA

<213> Artificial Sequence

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

<400> 7

aaagagaccg gttcactgtg acagtaaaag agaccggttc actgtgagaa tgaaagagac 60

cggttcactg tgatcggaaa agagaccggt tcactgtgag cggccttgaa acccagcaga 120

caatgtagct cagtagaaac ccagcagaca atgtagctga atggaaaccc agcagacaat 180

gtagcttcgg agaaacccag cagacaatgt agct 214

<210> 8

<211> 21

<212> RNA

<213> Homo sapiens

<400> 8

ucacagugaa ccggucucuu u 21

<210> 9

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 9

aaagagaccg gttcactgtg a 21

<210> 10

<211> 23

<212> RNA

<213> Homo sapiens

<400> 10

agcuacauug ucugcugggu uuc 23

<210> 11

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 11

gaaacccagc agacaatgta gct 23

<210> 12

<211> 23

<212> RNA

<213> Homo sapiens

<400> 12

agcuacaucu ggcuacuggg ucu 23

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 13

agacccagta gccagatgta gct 23

<210> 14

<211> 67

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 14

gaaacccagc agacaatgta gctagaccca gtagccagat gtagctaaag agaccggttc 60

actgtga 67

<210> 15

<211> 134

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 15

gaaacccagc agacaatgta gctagaccca gtagccagat gtagctaaag agaccggttc 60

actgtgagaa acccagcaga caatgtagct agacccagta gccagatgta gctaaagaga 120

ccggttcact gtga 134

<210> 16

<211> 62

<212> DNA

<213> Artificial Sequence

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

<400> 16

aataaaagat ctttattttc attagatctg tgtgttggtt ttttgtgtgc ggaccgcacg 60

tg 62

<210> 17

<211> 477

<212> DNA

<213> Artificial Sequence

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

<400> 17

gggtggcatc cctgtgaccc ctccccagtg cctctcctgg ccctggaagt tgccactcca 60

gtgcccacca gccttgtcct aataaaatta agttgcatca ttttgtctga ctaggtgtcc 120

ttctataata ttatggggtg gaggggggtg gtatggagca aggggcaagt tgggaagaca 180

acctgtaggg cctgcggggt ctattgggaa ccaagctgga gtgcagtggc acaatcttgg 240

ctcactgcaa tctccgcctc ctgggttcaa gcgattctcc tgcctcagcc tcccgagttg 300

ttgggattcc aggcatgcat gaccaggctc agctaatttt tgtttttttg gtagagacgg 360

ggtttcacca tattggccag gctggtctcc aactcctaat ctcaggtgat ctacccacct 420

tggcctccca aattgctggg attacaggcg tgaaccactg ctcccttccc tgtcctt 477

<210> 18

<211> 18

<212> DNA

<213> Homo sapiens

<400> 18

ctaggtcaag tgtaggag 18

<210> 19

<211> 18

<212> DNA

<213> Homo sapiens

<400> 19

acttgaccta gacagcct 18

<210> 20

<211> 18

<212> DNA

<213> Homo sapiens

<400> 20

tgaataactc attagtga 18

<210> 21

<211> 18

<212> DNA

<213> Homo sapiens

<400> 21

aaagtacatt aggctaat 18

<210> 22

<211> 18

<212> DNA

<213> Homo sapiens

<400> 22

ccagcactgg tgcttcgt 18

<210> 23

<211> 18

<212> DNA

<213> Homo sapiens

<400> 23

aaggctgtct aggtcaag 18

<210> 24

<211> 24

<212> DNA

<213> Homo sapiens

<400> 24

ctaggtcaag tgtaggagac acac 24

<210> 25

<211> 18

<212> DNA

<213> Homo sapiens

<400> 25

ggtcaagtgt aggagaca 18

<210> 26

<211> 18

<212> DNA

<213> Homo sapiens

<400> 26

caagtgtagg agacacac 18

<210> 27

<211> 24

<212> DNA

<213> Homo sapiens

<400> 27

agtgtaggag acacactgct ggcc 24

<210> 28

<211> 18

<212> DNA

<213> Homo sapiens

<400> 28

agtgtaggag acacactg 18

<210> 29

<211> 21

<212> DNA

<213> Homo sapiens

<400> 29

aggagacaca ctgctggcct g 21

<210> 30

<211> 18

<212> DNA

<213> Homo sapiens

<400> 30

taggtaccat agagtgag 18

<210> 31

<211> 18

<212> DNA

<213> Homo sapiens

<400> 31

gaggatactg cagaggtc 18

<210> 32

<211> 27

<212> DNA

<213> Homo sapiens

<400> 32

taggtaccat agagtgaggc gaggatg 27

<210> 33

<211> 27

<212> DNA

<213> Homo sapiens

<400> 33

atagagtgag gcgaggatga agccgag 27

<210> 34

<211> 27

<212> DNA

<213> Homo sapiens

<400> 34

tgaagccgag aggatactgc agaggtc 27

<210> 35

<211> 21

<212> DNA

<213> Homo sapiens

<400> 35

aaggctgtct aggtcaagtg t 21

<210> 36

<211> 21

<212> DNA

<213> Homo sapiens

<400> 36

tgttcctcca gattaacact t 21

<210> 37

<211> 21

<212> DNA

<213> Homo sapiens

<400> 37

gatgaagccg agaggatact g 21

<210> 38

<211> 21

<212> DNA

<213> Homo sapiens

<400> 38

gctgatttgt attaggtacc a 21

<210> 39

<211> 21

<212> DNA

<213> Homo sapiens

<400> 39

agaaagctga tacagataca a 21

<210> 40

<211> 21

<212> DNA

<213> Homo sapiens

<400> 40

ggtacgggca aagatttctt g 21

<210> 41

<211> 21

<212> DNA

<213> Homo sapiens

<400> 41

agaaagctga tacagataca a 21

<210> 42

<211> 21

<212> DNA

<213> Homo sapiens

<400> 42

acacaatgag ccacctacaa g 21

<210> 43

<211> 21

<212> DNA

<213> Homo sapiens

<400> 43

gtggctcatt gtgtgtgtgc c 21

<210> 44

<211> 21

<212> DNA

<213> Homo sapiens

<400> 44

catatccctg caggttcaga a 21

<210> 45

<211> 21

<212> DNA

<213> Homo sapiens

<400> 45

agagagagag agagagagag a 21

<210> 46

<211> 21

<212> DNA

<213> Homo sapiens

<400> 46

ttctcagttt tgaaattaaa a 21

<210> 47

<211> 21

<212> DNA

<213> Homo sapiens

<400> 47

tggattctct tctgaacctg c 21

<210> 48

<211> 21

<212> DNA

<213> Homo sapiens

<400> 48

tgctgaggca ggacacagtg t 21

<210> 49

<211> 21

<212> DNA

<213> Homo sapiens

<400> 49

atcatctgta accatcaagg a 21

<210> 50

<211> 21

<212> DNA

<213> Homo sapiens

<400> 50

tcctgcctac ttagtttcaa g 21

<210> 51

<211> 21

<212> DNA

<213> Homo sapiens

<400> 51

attacagttc tgtcagcatg c 21

<210> 52

<211> 21

<212> DNA

<213> Homo sapiens

<400> 52

tggtctcatt ctttttgtgg g 21

<210> 53

<211> 21

<212> DNA

<213> Homo sapiens

<400> 53

cgatattttc atggattcct t 21

<210> 54

<211> 21

<212> DNA

<213> Homo sapiens

<400> 54

ctgacactta ctttgtctaa a 21

<210> 55

<211> 21

<212> DNA

<213> Homo sapiens

<400> 55

aaaactggaa ccgcattccc a 21

<210> 56

<211> 21

<212> DNA

<213> Homo sapiens

<400> 56

acaaagtaag tgtcagtgtg g 21

<210> 57

<211> 21

<212> DNA

<213> Homo sapiens

<400> 57

ataatagttg tgtctttata a 21

<210> 58

<211> 21

<212> DNA

<213> Homo sapiens

<400> 58

tgtacaagca gggctgcaaa g 21

<210> 59

<211> 21

<212> DNA

<213> Homo sapiens

<400> 59

gttaacaaat acactaaaca c 21

<210> 60

<211> 21

<212> DNA

<213> Homo sapiens

<400> 60

ttcaacaagc tcccaagaag t 21

<210> 61

<211> 21

<212> DNA

<213> Homo sapiens

<400> 61

atgttcaagg tgcagaagga a 21

<210> 62

<211> 21

<212> DNA

<213> Homo sapiens

<400> 62

tgtttgctca aacgtgcacc a 21

<210> 63

<211> 21

<212> DNA

<213> Homo sapiens

<400> 63

aaataagaca tgaaaacaag a 21

<210> 64

<211> 21

<212> DNA

<213> Homo sapiens

<400> 64

aaatatgtac cacaagaaat g 21

<210> 65

<211> 21

<212> DNA

<213> Homo sapiens

<400> 65

tatctggttt ctctcactgc t 21

<210> 66

<211> 21

<212> DNA

<213> Homo sapiens

<400> 66

attgcaaagc ataatttgga t 21

<210> 67

<211> 3585

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 67

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

aggggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtaa agagaccggt tcactgtgac agtaaaagag accggttcac tgtgagaatg 2940

aaagagaccg gttcactgtg atcggaaaag agaccggttc actgtgagcg gccttgaaac 3000

ccagcagaca atgtagctca gtagaaaccc agcagacaat gtagctgaat ggaaacccag 3060

cagacaatgt agcttcggag aaacccagca gacaatgtag ctaagcttgg gtggcatccc 3120

tgtgacccct ccccagtgcc tctcctggcc ctggaagttg ccactccagt gcccaccagc 3180

cttgtcctaa taaaattaag ttgcatcatt ttgtctgact aggtgtcctt ctataatatt 3240

atggggtgga ggggggtggt atggagcaag gggcaagttg ggaagacaac ctgtagggcc 3300

tgcggggtct attgggaacc aagctggagt gcagtggcac aatcttggct cactgcaatc 3360

tccgcctcct gggttcaagc gattctcctg cctcagcctc ccgagttgtt gggattccag 3420

gcatgcatga ccaggctcag ctaatttttg ttttttggt agagacgggg tttcaccata 3480

ttggccaggc tggtctccaa ctcctaatct caggtgatct acccaccttg gcctcccaaa 3540

ttgctgggat tacaggcgtg aaccactgct cccttccctg tcctt 3585

<210> 68

<211> 3371

<212> DNA

<213> Artificial Sequence

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

<400> 68

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtaa gcttgggtgg catccctgtg acccctcccc agtgcctctc ctggccctgg 2940

aagttgccac tccagtgccc accagccttg tcctaataaa attaagttgc atcattttgt 3000

ctgactaggt gtccttctat aatattatgg ggtggagggg ggtggtatgg agcaaggggc 3060

aagttgggaa gacaacctgt agggcctgcg gggtctattg ggaaccaagc tggagtgcag 3120

tggcacaatc ttggctcact gcaatctccg cctcctgggt tcaagcgatt ctcctgcctc 3180

agcctcccga gttgttggga ttccaggcat gcatgaccag gctcagctaa tttttgtttt 3240

tttggtagag acggggtttc accatattgg ccaggctggt ctccaactcc taatctcagg 3300

tgatctaccc accttggcct cccaaattgc tgggattaca ggcgtgaacc actgctccct 3360

tccctgtcct t 3371

<210> 69

<211> 4380

<212> DNA

<213> Artificial Sequence

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

<400> 69

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggaaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360
ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420
caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480
cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540
ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660
gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780
agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840
tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080
gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200
actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260
ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320
cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380
gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440
cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500
cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620
gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680
gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740
caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800
acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920
ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg aggccagcgg ttccggacgg gctgacgcat 2760

tggacgattt tgatctggat atgctgggaa gtgacgccct cgatgatttt gaccttgaca 2820

tgcttggttc ggatgccctt gatgactttg acctcgacat gctcggcagt gacgcccttg 2880

atgatttcga cctggacatg ctgattaact ctagaagttc cggatctccg aaaaagaaac 2940

gcaaagttgg tagccagtac ctgcccgaca ccgacgaccg gcaccggatc gaggaaaagc 3000

ggaagcggac ctacgagaca ttcaagagca tcatgaagaa gtccccccttc agcggcccca 3060

ccgaccctag acctccacct agaagaatcg ccgtgcccag cagatccagc gccagcgtgc 3120

caaaacctgc cccccagcct tacccccttca ccagcagcct gagcaccatc aactacgacg 3180

agttccctac catggtgttc cccagcggcc agatctctca ggcctctgct ctggctccag 3240

cccctcctca ggtgctgcct caggctcctg ctcctgcacc agctccagcc atggtgtctg 3300

cactggctca ggcaccagca cccgtgcctg tgctggctcc tggacctcca caggctgtgg 3360

ctccaccagc ccctaaacct acacaggccg gcgagggcac actgtctgaa gctctgctgc 3420

agctgcagtt cgacgacgag gatctgggag ccctgctggg aaacagcacc gatcctgccg 3480

tgttcaccga cctggccagc gtggacaaca gcgagttcca gcagctgctg aaccagggca 3540

tccctgtggc ccctcacacc accgagccca tgctgatgga ataccccgag gccatcaccc 3600

ggctcgtgac aggcgctcag aggcctcctg atccagctcc tgcccctctg ggagcaccag 3660

gcctgcctaa tggactgctg tctggcgacg aggacttcag ctctatcgcc gatatggatt 3720

tctcagcctt gctgggctct ggcagcggca gccgggattc cagggaaggg atgtttttgc 3780

cgaagcctga ggccggctcc gctattagtg acgtgtttga gggccgcgag gtgtgccagc 3840

caaaacgaat ccggccattt catcctccag gaagtccatg ggccaaccgc ccactccccg 3900

ccagcctcgc accaacacca accggtccag tacatgagcc agtcgggtca ctgaccccgg 3960

caccagtccc tcagccactg gatccagcgc ccgcagtgac tcccgaggcc agtcacctgt 4020

tggaggatcc cgatgaagag acgagccagg ctgtcaaagc ccttcgggag atggccgata 4080

ctgtgattcc ccagaaggaa gaggctgcaa tctgtggcca aatggacctt tcccatccgc 4140

ccccaagggg ccatctggat gagctgacaa ccacacttga gtccatgacc gaggatctga 4200

acctggactc acccctgacc ccggaattga acgagattct ggataccttc ctgaacgacg 4260

agtgcctctt gcatgccatg catatcagca caggactgtc catcttcgac acatctctgt 4320

tttaaactag taataaaaga tctttatttt cattagatct gtgtgttggt tttttgtgtg 4380

<210> 70

<211> 3332

<212> DNA

<213> Artificial Sequence

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

<400> 70

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

aggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttтc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgacg 2700

cgctggacga tttcgatctc gacatgctgg gttctgatgc cctcgatgac tttgacctgg 2760

atatgttggg aagcgacgca ttggatgact ttgatctgga catgctcggc tccgatgctc 2820

tggacgattt cgatctcgat atgttataaa agcttgggtg gcatccctgt gacccctccc 2880

cagtgcctct cctggccctg gaagttgcca ctccagtgcc caccagcctt gtcctaataa 2940

aattaagttg catcattttg tctgactagg tgtccttcta taatattatg gggtggaggg 3000

gggtggtatg gagcaagggg caagttggga agacaacctg tagggcctgc ggggtctatt 3060

gggaaccaag ctggagtgca gtggcacaat cttggctcac tgcaatctcc gcctcctggg 3120

ttcaagcgat tctcctgcct cagcctcccg agttgttggg attccaggca tgcatgacca 3180

ggctcagcta atttttgttt ttttggtaga gacggggttt caccatattg gccaggctgg 3240

tctccaactc ctaatctcag gtgatctacc caccttggcc tcccaaattg ctgggattac 3300

aggcgtgaac cactgctccc ttccctgtcc tt 3332

<210> 71

<211> 3546

<212> DNA

<213> Artificial Sequence

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

<400> 71

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttтc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgacg 2700

cgctggacga tttcgatctc gacatgctgg gttctgatgc cctcgatgac tttgacctgg 2760

atatgttggg aagcgacgca ttggatgact ttgatctgga catgctcggc tccgatgctc 2820

tggacgattt cgatctcgat atgttataaa aagagaccgg ttcactgtga cagtaaaaga 2880

gaccggttca ctgtgagaat gaaagagacc ggttcactgt gatcggaaaa gagaccggtt 2940

cactgtgagc ggccttgaaa cccagcagac aatgtagctc agtagaaacc cagcagacaa 3000

tgtagctgaa tggaaaccca gcagacaatg tagcttcgga gaaacccagc agacaatgta 3060

gctaagcttg ggtggcatcc ctgtgacccc tccccagtgc ctctcctggc cctggaagtt 3120

gccactccag tgcccaccag ccttgtccta ataaaattaa gttgcatcat tttgtctgac 3180

taggtgtcct tctataatat tatggggtgg aggggggtgg tatggagcaa ggggcaagtt 3240

gggaagacaa cctgtagggc ctgcggggtc tattgggaac caagctggag tgcagtggca 3300

caatcttggc tcactgcaat ctccgcctcc tgggttcaag cgattctcct gcctcagcct 3360

cccgagttgt tgggattcca ggcatgcatg accaggctca gctaattttt gtttttttgg 3420

tagagacggg gtttcaccat attggccagg ctggtctcca actcctaatc tcaggtgatc 3480

tacccacctt ggcctcccaa attgctggga ttacaggcgt gaaccactgc tcccttccct 3540

gtcctt 3546

<210> 72

<211> 1707

<212> DNA

<213> Artificial Sequence

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

<400> 72

atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ccgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcgccga ccacctgatg ctcgccgagg gctaccgcct ggtgcagagg 600

ccgccgtccg ccgccgccgc ccatggccct catgcgctcc ggactctgcc gccgtacgcg 660

ggcccgggcc tggacagtgg gctgaggccg cgggggggctc cgctggggcc gccgccgccc 720

cgccaacccg gggccctggc gtacggggcc ttcgggccgc cgtcctcctt ccagcccttt 780

ccggccgtgc ctccgccggc cgcgggcatc gcgcacctgc agcctgtggc gacgccgtac 840

cccggccgcg ccgccgcgcc ccccaacgct ccgggaggcc ccccgggccc gcagccggcc 900

ccaagcgccg cagccccgcc gccgcccgcg cacgccctgg gcggcatgga cgccgaactc 960

atcgacgagg aggcgctgac gtcgctggag ctggagctgg ggctgcaccg cgtgcgcgag 1020

ctgcccgagc tgttcctggg ccagagcgag ttcgactgct tctcggactt ggggtccgcg 1080

ccgcccgccg gctccgtgag ctgccagtcc cagctcatca aacccagccg catgcgcaag 1140

taccccaacc ggcccagcaa gacgcccccc cacgaacgcc cttacgcttg cccagtggag 1200

tcctgtgatc gccgcttctc ccgcagcgac aacctggtga gacacatccg catccacaca 1260

ggccagaagc ccttccagtg ccgcatctgc atgagaaact tcagccgaga ggataacttg 1320

cacactcaca tccgcaccca cacaggcgaa aagcccttcg cctgcgacat ctgtggaaga 1380

aagtttgccc ggagcgatga acttgtccga cataccaaga tccacttgcg gcagaaggac 1440

cgcccttacg cttgcccagt ggagtcctgt gatcgccgct tctcccaatc agggaatctg 1500

actgagcaca tccgcatcca cacaggccag aagcccttcc agtgccgcat ctgcatgaga 1560

aacttcagca caagtggaca tctggtacgc cacatccgca cccacacagg cgaaaagccc 1620

ttcgcctgcg acatctgtgg aagaaagttt gcccagaata gtaccctgac cgaacatacc 1680

aagatccact tgcggcagaa ggacaag 1707

<210> 73

<211> 1755

<212> DNA

<213> Artificial Sequence

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

<400> 73

atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ctgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcggcgg cagcggcggc ggaagcggcg ccgaccacct gatgctcgcc 600

gagggctacc gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg 660

ctccggactc tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg 720

gctccgctgg ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg 780

ccgccgtcct ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac 840

ctgcagcctg tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga 900

ggcccccccgg gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc 960

ctgggcggca tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag 1020

ctggggctgc accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac 1080

tgcttctcgg acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt 1140

ggtggttctg gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg 1200

cccagcaaga cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc 1260

cgcttctccc gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc 1320

ttccagtgcc gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc 1380

cgcacccaca caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg 1440

agcgatgaac ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct 1500

tgcccagtgg agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc 1560

cgcatccaca caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca 1620

agtggacatc tggtacgcca catccgcacc cacacaggcg aaaagcccct cgcctgcgac 1680

atctgtggaa gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg 1740

cggcagaagg acaag 1755

<210> 74

<211> 3438

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 74

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtga aacccagcag acaatgtagc tagacccagt agccagatgt agctaaagag 2940

accggttcac tgtgaaagct tgggtggcat ccctgtgacc cctccccagt gcctctcctg 3000

gccctggaag ttgccactcc agtgcccacc agccttgtcc taataaaatt aagttgcatc 3060

attttgtctg actaggtgtc cttctataat attatggggt ggaggggggt ggtatggagc 3120

aaggggcaag ttgggaagac aacctgtagg gcctgcgggg tctattggga accaagctgg 3180

agtgcagtgg cacaatcttg gctcactgca atctccgcct cctgggttca agcgattctc 3240

ctgcctcagc ctcccgagtt gttgggattc caggcatgca tgaccaggct cagctaattt 3300

ttgtttttttt ggtagagacg gggtttcacc atattggcca ggctggtctc caactcctaa 3360

tctcaggtga tctacccacc ttggcctccc aaattgctgg gattacaggc gtgaaccact 3420

gctcccttcc ctgtcctt 3438

<210> 75

<211> 3505

<212> DNA

<213> Artificial Sequence

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

<400> 75

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtga aacccagcag acaatgtagc tagacccagt agccagatgt agctaaagag 2940

accggttcac tgtgagaaac ccagcagaca atgtagctag acccagtagc cagatgtagc 3000

taaagagacc ggttcactgt gaaagcttgg gtggcatccc tgtgacccct ccccagtgcc 3060

tctcctggcc ctggaagttg ccactccagt gcccaccagc cttgtcctaa taaaattaag 3120

ttgcatcatt ttgtctgact aggtgtcctt ctataatatt atggggtgga ggggggtggt 3180

atggagcaag gggcaagttg ggaagacaac ctgtagggcc tgcggggtct attgggaacc 3240

aagctggagt gcagtggcac aatcttggct cactgcaatc tccgcctcct gggttcaagc 3300

gattctcctg cctcagcctc ccgagttgtt gggattccag gcatgcatga ccaggctcag 3360

ctaatttttg ttttttggt agagacgggg tttcaccata ttggccaggc tggtctccaa 3420

ctcctaatct caggtgatct acccaccttg gcctcccaaa ttgctgggat tacaggcgtg 3480

aaccactgct cccttccctg tcctt 3505

<210> 76

<211> 3804

<212> DNA

<213> Artificial Sequence

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

<400> 76

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttтc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccg ccgaccacct gatgctcgcc gagggctacc 2100

gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg ctccggactc 2160

tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg gctccgctgg 2220

ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg ccgccgtcct 2280

ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac ctgcagcctg 2340

tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga ggcccccggg 2400

gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc ctgggcggca 2460

tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag ctggggctgc 2520

accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac tgcttctcgg 2580

acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt ggtggttctg 2640

gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg cccagcaaga 2700

cgcccccccа cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc cgcttctccc 2760

gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc ttccagtgcc 2820

gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc cgcacccaca 2880

caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg agcgatgaac 2940

ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct tgcccagtgg 3000

agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc cgcatccaca 3060

caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca agtggacatc 3120

tggtacgcca catccgcacc cacacaggcg aaaagccctt cgcctgcgac atctgtggaa 3180

gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg cggcagaagg 3240
acaagtaact cgaggaaacc cagcagacaa tgtagctaga cccagtagcc agatgtagct 3300
aaagagaccg gttcactgtg aaagcttggg tggcatccct gtgacccctc cccagtgcct 3360
ctcctggccc tggaagttgc cactccagtg cccaccagcc ttgtcctaat aaaattaagt 3420
tgcatcattt tgtctgacta ggtgtccttc tataatatta tggggtggag gggggtggta 3480
tggagcaagg ggcaagttgg gaagacaacc tgtagggcct gcggggtcta ttgggaacca 3540
agctggagtg cagtggcaca atcttggctc actgcaatct ccgcctcctg ggttcaagcg 3600

attctcctgc ctcagcctcc cgagttgttg ggattccagg catgcatgac caggctcagc 3660
taatttttgt ttttttggta gagacggggt ttcaccatat tggccaggct ggtctccaac 3720
tcctaatctc aggtgatcta cccaccttgg cctcccaaat tgctgggatt acaggcgtga 3780
accactgctc ccttccctgt cctt 3804

<210> 77
<211> 176
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 77

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15
Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30
Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
35 40 45
Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60
Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val

65 70 75 80
Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg
100 105 110
Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125
Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr Gly

130 135 140
Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn
145 150 155 160
Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 78

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 78

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15
Phe Ser Thr Lys Asn Ser Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30
Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ala
35 40 45
Asp Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60
Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Leu Ala His Leu Arg

65 70 75 80
Ala His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95
Cys Gly Lys Ser Phe Ser Thr Lys Asn Ser Leu Thr Glu His Gln Arg
100 105 110
Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

115 120 125

Phe Ser Gln Ala Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr His

145 150 155 160

Leu Asp Leu Ile Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser

165 170 175

<210> 79

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 79

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Gln Ala Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly

20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu

35 40 45

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly Asn Leu Thr

65 70 75 80

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg

100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

115 120 125

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

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ala

145 150 155 160

Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly Lys Lys Thr Ser

165 170 175

<210> 80

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 80

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Thr Thr Gly Asn Leu Thr Val His Gln Arg Thr His Thr Gly

20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser

35 40 45

Gly Glu Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His

65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

85 90 95

Cys Gly Lys Ser Phe Ser Thr Ser Gly Asn Leu Thr Glu His Gln Arg

100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg

145 150 155 160

Ala Asn Leu Arg Ala His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 81

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 81

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Ser Arg Arg Thr Cys Arg Ala His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Thr
35 40 45

Gly Ala Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val

65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Gly Asp Leu Arg Arg His Gln Arg Thr His Thr Gly

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser

145 150 155 160

His Ser Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 82

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 82

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Arg Lys Asp Asn Leu Lys Asn His Gln Arg Thr His Thr Gly

20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro

35 40 45

Gly Ala Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His

65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

85 90 95

Cys Gly Lys Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg

100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

115 120 125

Phe Ser Thr Ser Gly Glu Leu Val Arg His Gln Arg Thr His Thr Gly

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Lys

145 150 155 160

Asp Asn Leu Lys Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser

165 170 175

<210> 83

<211> 232

<212> PRT

<213> Artificial Sequence

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

<400> 83

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

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

20 25 30

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

35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

50 55 60

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

65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

85 90 95

Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg

100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

115 120 125

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

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser

145 150 155 160

Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His Leu Val

180 185 190

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

195 200 205

Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu His Gln Arg

210 215 220

Thr His Thr Gly Lys Lys Thr Ser
225 230

<210> 84

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 84

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
145 150 155 160

Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 85

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 85

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Ser Lys Lys Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser
145 150 155 160

Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 86

<211> 232

<212> PRT

<213> Artificial Sequence

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

<400> 86

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Asp Cys Arg Asp Leu Ala Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn
35 40 45

Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn Asp Ala Leu Thr
65 70 75 80

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Asp Pro Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg
145 150 155 160

Ala His Leu Glu Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu Val
180 185 190

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg
210 215 220

Thr His Thr Gly Lys Lys Thr Ser

225 230

<210> 87

<211> 176

<212> PRT

<213> Artificial Sequence

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

<400> 87

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly

20 25 30

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

35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro Gly Asn Leu Val

65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

85 90 95

Cys Gly Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His Gln Arg

100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly

130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg

145 150 155 160

Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser

165 170 175

<210> 88

<211> 204

<212> PRT

<213> Artificial Sequence

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

<400> 88

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15
Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly

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

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

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

<210> 89

<211> 176

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 89

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg
35 40 45

Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
145 150 155 160

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 90

<211> 176

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 90

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu Arg
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
145 150 155 160

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 91

<211> 260

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 91

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Arg Asp Glu Leu Asn Val His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp His Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asp Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
145 150 155 160

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr
180 185 190

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg
210 215 220

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
225 230 235 240

Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly
245 250 255

Lys Lys Thr Ser
260

<210> 92

<211> 177

<212> PRT

<213> Artificial Sequence

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

<400> 92

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
1 5 10 15

Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu
35 40 45

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
50 55 60

Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr
65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
85 90 95

Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile
100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
115 120 125

Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr
130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln
145 150 155 160

Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys Asp

165 170 175

Lys

<210> 93

<211> 177

<212> PRT

<213> Artificial Sequence

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

<400> 93

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg

1 5 10 15

Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro

20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu

35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp

50 55 60

Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr

65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu

85 90 95

Ser Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Ile

100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg

115 120 125

Asn Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr

130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg

145 150 155 160

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

165 170 175

Lys

<210> 94

<211> 264

<212> PRT

<213> Artificial Sequence

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

<400> 94

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg

1 5 10 15

Arg Asp Glu Leu Asn Val His Ile Arg Ile His Thr Gly Gln Lys Pro

20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu

35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp

50 55 60

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

65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu

85 90 95

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile

100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg

115 120 125

Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr

130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg

145 150 155 160

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

165 170 175

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Thr

180 185 190

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

195 200 205

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser Leu

210 215 220

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp

225 230 235 240

Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr

245 250 255

Lys Ile His Leu Arg Gln Lys Asp

260

<210> 95

<211> 264

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 95

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Asp

1 5 10 15

Pro Gly Ala Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro

20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu

35 40 45

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp

50 55 60

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

65 70 75 80

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

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

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Glu Arg Ser His Leu
210 215 220
Arg Glu His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
225 230 235 240
Ile Cys Gly Arg Lys Phe Ala Gln Ala Gly His Leu Ala Ser His Thr
245 250 255
Lys Ile His Leu Arg Gln Lys Asp
260

<210> 96

<211> 175

<212> PRT

<213> Artificial Sequence

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

<400> 96

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

1 5 10 15

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro

20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn Leu

35 40 45

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu

50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ala

65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys

85 90 95

Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg Ile

100 105 110

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe

115 120 125

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

130 135 140

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

145 150 155 160

Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys

165 170 175

<210> 97

<211> 175

<212> PRT

<213> Artificial Sequence

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

<400> 97

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

1 5 10 15

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro

20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser His Arg Thr Thr Leu
35 40 45
Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu

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

115 120 125
Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu
130 135 140
Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
145 150 155 160
Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
165 170 175
<210> 98
<211> 261
<212> PRT
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 98
Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
1 5 10 15
Arg Asp Glu Leu Asn Val His Leu Arg Ile His Thr Gly His Lys Pro
20 25 30
Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu

50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Ala

65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys

85 90 95

Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile

100 105 110

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe

115 120 125

Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly Glu

130 135 140

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp

145 150 155 160

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala

165 170 175

Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu

180 185 190

Thr Glu His Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg

195 200 205

Ile Cys Met Arg Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile

210 215 220

Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg

225 230 235 240

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

245 250 255

Lys Gln Lys Glu Lys

260

<210> 99

<211> 753

<212> PRT

<213> Artificial Sequence

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

<400> 99

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
50 55 60

Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
85 90 95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
165 170 175

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Glu Ala

210 215 220

Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met
225 230 235 240
Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
245 250 255
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
260 265 270
Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Ser Ser Gly Ser
275 280 285

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

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

Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val
420 425 430
Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser
435 440 445
Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu
450 455 460

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

Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala
485 490 495

Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr
500 505 510

Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro
515 520 525

Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp
530 535 540

Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly
545 550 555 560

Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu
565 570 575

Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln
580 585 590

Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn
595 600 605

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His
610 615 620

Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu Asp
625 630 635 640

Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro
645 650 655

Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
660 665 670

Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp
675 680 685

Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr
690 695 700

Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro

705 710 715 720

Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu

725 730 735

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

740 745 750

Phe

<210> 100

<211> 753

<212> PRT

<213> Artificial Sequence

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

<400> 100

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala

1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr

35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg

50 55 60

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro

65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu

85 90 95

Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

100 105 110

Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln

115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

130 135 140

Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

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

Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met
225 230 235 240
Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
245 250 255
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
260 265 270
Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Ser Ser Gly Ser
275 280 285

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

Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr
355 360 365
Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile
370 375 380
Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln

385 390 395 400

Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln

405 410 415

Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val

420 425 430

Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser

435 440 445

Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu

450 455 460

Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val

465 470 475 480

Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala

485 490 495

Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr

500 505 510

Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro

515 520 525

Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp

530 535 540

Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly

545 550 555 560

Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu

565 570 575

Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln

580 585 590

Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn

595 600 605

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His

610 615 620

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

625 630 635 640

Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro
645 650 655
Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
660 665 670

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

His Ala Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu
740 745 750
Phe

<210> 101
<211> 272
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 101
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15
Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45
Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
50 55 60
Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro

65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu

85 90 95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln

115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

130 135 140

Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr

145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg

165 170 175

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

180 185 190

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

195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala

210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp

225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu

245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu

260 265 270

<210> 102

<211> 272

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 102

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15
Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30
Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45
Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

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

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

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205
Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220
Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240
Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu

245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 103

<211> 580

<212> PRT

<213> Artificial Sequence

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

<400> 103

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser

165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser

180 185 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn

195 200 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val

210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His

225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met

245 250 255

Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His

260 265 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

275 280 285

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

290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser

305 310 315 320

Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys

325 330 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser

340 345 350

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

355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His

370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His

385 390 395 400

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

405 410 415

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

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

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
545 550 555 560
Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
565 570 575
Ser Val Ser Cys
580

<210> 104
<211> 394
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<400> 104
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln

1 5 10 15
Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr

20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg

35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala

50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val

65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro

85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro

100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His

115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr

130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu

145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser

165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser

180 185 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn

195 200 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val

210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His

225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met

245 250 255

Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His

260 265 270

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

325 330 335
Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser
340 345 350
Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365
Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His
370 375 380
Thr Lys Ile His Leu Arg Gln Lys Asp Lys
385 390

<210> 105
<211> 580
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 105
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15
Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30
Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60
Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val

65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro

85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro

100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His

115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr

130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu

145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser

165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser

180 185 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn

195 200 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val

210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His

225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met

245 250 255

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His

260 265 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

275 280 285

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys

290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser

305 310 315 320

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
325 330 335
Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
340 345 350
Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365

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

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

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

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

565 570 575

Ser Val Ser Cys

580

<210> 106

<211> 394

<212> PRT

<213> Artificial Sequence

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

<400> 106

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr

20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro

85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro

100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu

145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser

165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190
Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
195 200 205
Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val

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

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

340 345 350
Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365
Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
370 375 380
Thr Lys Ile His Leu Arg Gln Lys Asp Lys
385 390
<210> 107
<211> 388
<212> PRT
<213> Artificial Sequence

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

polypeptide

<400> 107

Met Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
1 5 10 15

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
20 25 30

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
35 40 45

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
50 55 60

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
65 70 75 80

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
85 90 95

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
100 105 110

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
115 120 125

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
130 135 140

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
145 150 155 160

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
165 170 175

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
180 185 190

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His
195 200 205

Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
210 215 220

Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
225 230 235 240
Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
245 250 255

Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
260 265 270
Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
275 280 285
Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
290 295 300
Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
305 310 315 320

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

Ser Val Ser Cys
385

<210> 108
<211> 479
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 108
Met Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
1 5 10 15
Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val

20 25 30

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His

35 40 45

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met

50 55 60

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His

65 70 75 80

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

85 90 95

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys

100 105 110

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser

115 120 125

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys

130 135 140

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His

145 150 155 160

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

165 170 175

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

180 185 190

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ser Gly Leu

195 200 205

Glu Met Ala Asp His Met Met Ala Met Asn His Gly Arg Phe Pro Asp

210 215 220

Gly Thr Asn Gly Leu His His His Pro Ala His Arg Met Gly Met Gly

225 230 235 240

Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln Pro Gln His Ala

245 250 255

Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly Ala Gly Asn Met

260 265 270

Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro Gly Thr Val Asn

275 280 285

Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala Arg Phe Asn Asn

290 295 300

Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly Gly Ser Leu Pro

305 310 315 320

Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr Phe Asn His His

325 330 335

Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His Pro Ala Ala Gly

340 345 350

His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp Cys Asn Pro Lys

355 360 365

His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly Gly Ser Ser Thr

370 375 380

Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala Gly Ser Ser Asn

385 390 395 400

Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala Ser Val Ala His

405 410 415

Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp Thr Asp Phe Ile

420 425 430

Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met Gly Leu Asp Arg

435 440 445

Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn Glu Phe Asp Phe

450 455 460

Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg Val Ser Cys

465 470 475

<210> 109

<211> 476

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 109

Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln
35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly

50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Val Arg His Ala Met Gly Pro
65 70 75 80

Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
85 90 95

Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
100 105 110

Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr

115 120 125

Phe Asn His His Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His
130 135 140

Pro Ala Ala Gly His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp
145 150 155 160

Cys Asn Pro Lys His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly
165 170 175

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

180 185 190

Gly Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala
195 200 205

Ser Val Ala His Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp
210 215 220

Thr Asp Phe Ile Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met
225 230 235 240

Gly Leu Asp Arg Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn

245 250 255

Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg

260 265 270

Val Ser Cys Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr

275 280 285

Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys

290 295 300

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val

305 310 315 320

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile

325 330 335

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg

340 345 350

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys

355 360 365

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg

370 375 380

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg

385 390 395 400

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly

405 410 415

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser

420 425 430

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

435 440 445

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr

450 455 460

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys

465 470 475

<210> 110

<211> 554

<212> PRT

<213> Artificial Sequence

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

<400> 110

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln

1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr

20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg

35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala

50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val

65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro

85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro

100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His

115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr

130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu

145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser

165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser

180 185 190

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

195 200 205

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys

210 215 220

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn

225 230 235 240

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys

245 250 255

Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His

260 265 270

Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly

275 280 285

Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg

290 295 300

Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser

305 310 315 320

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

325 330 335

Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn

340 345 350

Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys

355 360 365

Leu Glu Met Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val

370 375 380

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

385 390 395 400

Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro

405 410 415

Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu

420 425 430

Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala

435 440 445

Val Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr

450 455 460

Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro
465 470 475 480

Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala
485 490 495

His Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu
500 505 510

Thr Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro

515 520 525

Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly
530 535 540

Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
545 550

<210> 111

<211> 362

<212> PRT

<213> Artificial Sequence

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

<400> 111

Met Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
1 5 10 15

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
20 25 30

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
35 40 45

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
50 55 60

Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
65 70 75 80

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

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

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

Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro
275 280 285
Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala
290 295 300
His Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu
305 310 315 320
Thr Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro

325 330 335

Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly
340 345 350
Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
355 360

<210> 112

<211> 479

<212> PRT

<213> Artificial Sequence

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

<400> 112

Met Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
1 5 10 15
Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys

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

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

145 150 155 160

Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys

165 170 175

Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val

180 185 190

Ser Leu Ala Pro Val Val Thr Thr Cys Ala Leu Glu Met Ser Gly Leu

195 200 205

Glu Met Ala Asp His Met Met Ala Met Asn His Gly Arg Phe Pro Asp

210 215 220

Gly Thr Asn Gly Leu His His His Pro Ala His Arg Met Gly Met Gly

225 230 235 240

Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln Pro Gln His Ala

245 250 255

Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly Ala Gly Asn Met

260 265 270

Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro Gly Thr Val Asn

275 280 285

Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala Arg Phe Asn Asn

290 295 300

Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly Gly Ser Leu Pro

305 310 315 320

Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr Phe Asn His His

325 330 335

Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His Pro Ala Ala Gly

340 345 350

His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp Cys Asn Pro Lys

355 360 365

His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly Gly Ser Ser Thr

370 375 380

Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala Gly Ser Ser Asn

385 390 395 400

Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala Ser Val Ala His

405 410 415

Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp Thr Asp Phe Ile

420 425 430

Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met Gly Leu Asp Arg

435 440 445

Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn Glu Phe Asp Phe

450 455 460

Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg Val Ser Cys

465 470 475

<210> 113

<211> 476

<212> PRT

<213> Artificial Sequence

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

<400> 113

Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly

1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg

20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln

35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly

50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Val Arg His Ala Met Gly Pro

65 70 75 80

Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala

85 90 95

Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly

100 105 110

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

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

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

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

355 360 365

Glu Gly Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His

370 375 380

Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met

385 390 395 400

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

405 410 415

Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala

420 425 430

Gln Asn Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys

435 440 445

Glu Lys Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro

450 455 460

Pro Val Ser Leu Ala Pro Val Val Thr Thr Cys Ala

465 470 475

<210> 114

<211> 559

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 114

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

1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala

20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln

35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro

50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn

65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
85 90 95
Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
100 105 110
Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
115 120 125

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

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

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

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu

325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Ala

340 345 350

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys

355 360 365

Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile

370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe

385 390 395 400

Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly Glu

405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp

420 425 430

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala

435 440 445

Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu

450 455 460

Thr Glu His Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg

465 470 475 480

Ile Cys Met Arg Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile

485 490 495

Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg

500 505 510

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

515 520 525

Lys Gln Lys Glu Lys Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser

530 535 540

Ser Ala Pro Pro Val Ser Leu Ala Pro Val Val Thr Thr Cys Ala

545 550 555

<210> 115

<211> 631

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 115

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala

210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr
370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys
405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
420 425 430

Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg
435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
450 455 460

Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly
465 470 475 480
Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu
485 490 495
Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys
500 505 510

Lys Ala Asp Lys Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser
515 520 525
Ser Tyr Pro Ser Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr
530 535 540
Ser Tyr Pro Ser Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr
545 550 555 560
Ser Phe Ser Ser Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser
565 570 575

Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro
580 585 590
Ala Phe Pro Ala Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn
595 600 605
Ser Phe Ser Ala Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser
610 615 620
Pro Arg Thr Ile Glu Ile Cys
625 630

<210> 116
<211> 719
<212> PRT
<213> Artificial Sequence

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

<400> 116
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15
Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr

20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe

35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser

65 70 75 80

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

85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn

100 105 110

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

115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn

130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly

145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser

165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys

180 185 190

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

195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala

210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr

225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu

245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys

260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro

275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu

290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg

305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg

325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Arg

340 345 350

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

355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu Thr

370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile

385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Thr Lys

405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser

420 425 430

Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile Arg

435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn

450 455 460

Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly

465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu

485 490 495

Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Arg

500 505 510

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Thr Ser

515 520 525

His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro Phe

530 535 540

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser Leu Val

545 550 555 560

Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile

565 570 575

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys

580 585 590

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala

595 600 605

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr

610 615 620

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr

625 630 635 640

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser

645 650 655

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala

660 665 670

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser

675 680 685

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu

690 695 700

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys

705 710 715

<210> 117

<211> 627

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 117

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15
Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr

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

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

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

210 215 220
Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240
Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu

245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys

260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro

275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu

290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg

305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg

325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser

340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe

355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr

370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile

385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ile Arg

405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys

420 425 430

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

435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser

450 455 460

Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe

465 470 475 480

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

485 490 495

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
500 505 510
Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
515 520 525
Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser

530 535 540
Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560
Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575
Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590
Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala

595 600 605
Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620
Glu Ile Cys
625

<210> 118
<211> 631
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 118
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30
Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45
Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320
Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

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

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

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

Ser Tyr Pro Ser Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr
530 535 540
Ser Tyr Pro Ser Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr

545 550 555 560

Ser Phe Ser Ser Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser

565 570 575

Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro

580 585 590

Ala Phe Pro Ala Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn

595 600 605

Ser Phe Ser Ala Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser

610 615 620

Pro Arg Thr Ile Glu Ile Cys

625 630

<210> 119

<211> 473

<212> PRT

<213> Artificial Sequence

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

<400> 119

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

1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala

20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln

35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro

50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn

65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn

85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly

100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
115 120 125
Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu

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

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

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

325 330 335
Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala
340 345 350
Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys

355 360 365

Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Leu Arg Ile

370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe

385 390 395 400

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

405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp

420 425 430

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu Lys Lys

435 440 445

Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val Ser

450 455 460

Leu Ala Pro Val Val Thr Thr Cys Ala

465 470

<210> 120

<211> 719

<212> PRT

<213> Artificial Sequence

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

<400> 120

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

1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr

20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe

35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser

65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

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

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

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

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300
Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320
Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg

325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Asp Pro

340 345 350

Gly Ala Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe

355 360 365

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

370 375 380

Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile

385 390 395 400

Cys Gly Arg Lys Phe Ala Gln Ser Gly Asp Leu Arg Arg His Thr Lys

405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser

420 425 430

Cys Asp Arg Arg Phe Ser Thr His Leu Asp Leu Ile Arg His Ile Arg

435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn

450 455 460

Phe Ser Thr Ser Gly Asn Leu Val Arg His Ile Arg Thr His Thr Gly

465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser

485 490 495

Asp Asn Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp Arg

500 505 510

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser

515 520 525

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

530 535 540

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Glu Arg Ser His Leu Arg

545 550 555 560

Glu His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile

565 570 575

Cys Gly Arg Lys Phe Ala Gln Ala Gly His Leu Ala Ser His Thr Lys
580 585 590
Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
595 600 605

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
610 615 620
Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
625 630 635 640
Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
645 650 655
Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
660 665 670

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser
675 680 685
Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
690 695 700
Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
705 710 715
<210> 121
<211> 627
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

polypeptide
<400> 121
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15
Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30
Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45
Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

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

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

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

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270
Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285
Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

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

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400
Cys Gly Arg Lys Phe Ala Arg Ser Asp Asn Arg Lys Thr His Ile Arg
405 410 415
Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
420 425 430
Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
450 455 460
Ser Ser Leu Thr Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
465 470 475 480
Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Asn Arg Lys
485 490 495
Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
515 520 525
Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540
Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser

545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser

565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala

580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala

595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile

610 615 620

Glu Ile Cys

625

<210> 122

<211> 473

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 122

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

1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala

20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln

35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro

50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn

65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn

85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly

100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
115 120 125

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

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

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

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

355 360 365

Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg Ile

370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe

385 390 395 400

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

405 410 415

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

420 425 430

Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys Lys

435 440 445

Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val Ser

450 455 460

Leu Ala Pro Val Val Thr Thr Cys Ala

465 470

<210> 123

<211> 627

<212> PRT

<213> Artificial Sequence

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

<400> 123

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

1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr

20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe

35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser

65 70 75 80

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

85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn

100 105 110

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

115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn

130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly

145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser

165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys

180 185 190

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

195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala

210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr

225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu

245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys

260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro

275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu

290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg

305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg

325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser

340 345 350

Asp Asn Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe

355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu Thr

370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile

385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Ile Arg

405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys

420 425 430

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly

435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser

450 455 460

Gly His Leu Thr Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe

465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Ser Ser Thr Arg Lys

485 490 495

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys

500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser

515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser

530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser

545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser

565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590
Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala

595 600 605
Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620
Glu Ile Cys
625

<210> 124
<211> 627
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 124
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

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

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

130 135 140

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

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

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

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

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

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

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

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540
Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560
Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575
Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605
Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620
Glu Ile Cys

625

<210> 125

<211> 272

<212> PRT

<213> Artificial Sequence

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

<400> 125

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala

1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

20 25 30

Ser Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr

35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg

50 55 60

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro

65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu

85 90 95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

100 105 110

Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln

115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

130 135 140

Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr

145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr

165 170 175

Ser Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr

180 185 190

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

195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala

210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp

225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu

245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu

260 265 270

<210> 126

<211> 272

<212> PRT

<213> Artificial Sequence

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

<400> 126

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala

1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

20 25 30

Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr His Thr

35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg

50 55 60

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro

65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu

85 90 95

Arg Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

100 105 110

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

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

225 230 235 240
Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
245 250 255
Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 127

<211> 261

<212> PRT

<213> Artificial Sequence

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

<400> 127

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala

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

35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg

50 55 60

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro

65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu

85 90 95

Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro

100 105 110

Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln

115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

130 135 140

Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr

145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln

165 170 175

Asn Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr

180 185 190

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

195 200 205

Lys Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser

210 215 220

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu

225 230 235 240

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe

245 250 255

Asp Leu Asp Met Leu

260

<210

> 128

<211> 386

<212> PRT

<213> Artificial Sequence

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

<400> 128

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

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

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
195 200 205

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser

210 215 220

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys
225 230 235 240

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
245 250 255

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
260 265 270

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
275 280 285

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
290 295 300

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
305 310 315 320

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
325 330 335

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
340 345 350

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
355 360 365

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
370 375 380

Asp Lys
385

<210> 129

<211> 402

<212> PRT

<213> Artificial Sequence

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

<400> 129

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

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

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

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

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

260 265 270

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys

275 280 285

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His

290 295 300

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val

305 310 315 320

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His

325 330 335

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met

340 345 350

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His

355 360 365

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

370 375 380

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys

385 390 395 400

Asp Lys

<210> 130

<211> 569

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 130

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln

1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr

20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg

35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

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

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

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

Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala Gly Ile Ala His
260 265 270
Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro
275 280 285
Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala

290 295 300

Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met Asp Ala Glu Leu
305 310 315 320

Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu Leu Gly Leu His
325 330 335

Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp
340 345 350

Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
355 360 365

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
370 375 380

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
385 390 395 400

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
405 410 415

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
420 425 430

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
435 440 445

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
450 455 460

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
465 470 475 480

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
485 490 495

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
500 505 510

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
515 520 525

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
530 535 540

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
545 550 555 560
Lys Ile His Leu Arg Gln Lys Asp Lys
565

<210> 131

<211> 585

<212> PRT

<213

> Artificial Sequence

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

<400> 131

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

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

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125
Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140
Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160
Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser

165 170 175

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

Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr
245 250 255
Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro
260 265 270
Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr
275 280 285
Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly
290 295 300

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

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

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
420 425 430

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

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

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
565 570 575
Lys Ile His Leu Arg Gln Lys Asp Lys
580 585

<210> 132
<211> 523
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 132
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe

20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp

35 40 45

Met Leu Ile Asn Ser Arg Ser Ser Gly Ser Pro Lys Lys Lys Arg Lys

50 55 60

Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu

65 70 75 80

Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys

85 90 95

Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile

100 105 110

Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln

115 120 125

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

130 135 140

Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu

145 150 155 160

Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro

165 170 175

Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro

180 185 190

Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys

195 200 205

Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu

210 215 220

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

225 230 235 240

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

245 250 255

Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro

260 265 270

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

275 280 285

Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu

290 295 300

Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp

305 310 315 320

Met Asp Phe Ser Ala Leu Leu Gly Ser Gly Ser Gly Ser Arg Asp Ser

325 330 335

Arg Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser

340 345 350

Asp Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro

355 360 365

Phe His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser

370 375 380

Leu Ala Pro Thr Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu

385 390 395 400

Thr Pro Ala Pro Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr

405 410 415

Pro Glu Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln

420 425 430

Ala Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys

435 440 445

Glu Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro

450 455 460

Arg Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu

465 470 475 480

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

485 490 495

Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser

500 505 510

Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe

515 520

<210> 133

<211> 50

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 133

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu

1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe

20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp

35 40 45

Met Leu

50

<210> 134

<211> 275

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 134

Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly

1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg

20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln

35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly

50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro
65 70 75 80
Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
85 90 95
Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
100 105 110
Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr

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

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

245 250 255
Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg
260 265 270
Val Ser Cys
275
<210> 135
<211> 183
<212> PRT
<213> Artificial Sequence

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

<400> 135

Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro
1 5 10 15

Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro
20 25 30

Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala
35 40 45

Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly
50 55 60

Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro
65 70 75 80

Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro
85 90 95

Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro
100 105 110

Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu
115 120 125

Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu
130 135 140

Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe
145 150 155 160

Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro
165 170 175

Pro Ala Gly Ser Val Ser Cys
180

<210> 136

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<400> 136

Cys Xaa Cys Xaa His Xaa His

1 5

<210> 137

<211> 17

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<400> 137

Cys Xaa Cys Xaa Cys Xaa His Xaa Xaa Xaa Cys Xaa Cys Xaa Cys Xaa

1 5 10 15

Cys

<210> 138

<211> 17

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<400> 138

Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa His Xaa Cys Xaa Cys Xaa

1 5 10 15

Cys

<210> 139

<211> 15

<212> PRT

<213> Artificial Sequence

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

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(8)

<223> Any amino acid

<220><221> MOD_RES

<222> (10)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (15)..(15)

<223> Cys, His, or Asp

<400> 139

Cys Xaa Cys Xaa His Xaa Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa

1 5 10 15

<210> 140

<211> 17

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<400> 140

Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa Cys Xaa Cys Xaa Cys Xaa

1 5 10 15

Cys

<210> 141

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<400> 141

Cys Xaa Cys Xaa Cys Xaa His

1 5

<210> 142

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<400> 142

Cys Xaa Cys Xaa His Xaa Cys

1 5

<210> 143

<211> 17

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223

> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<400> 143

Cys Xaa Cys Xaa His Xaa Cys Xaa Xaa Xaa Cys Xaa Cys Xaa His Xaa

1 5 10 15

Cys

<210> 144

<211> 17

<

212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<400> 144

Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa Cys Xaa Cys Xaa His Xaa

1 5 10 15

Cys

<210> 145

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<400> 145

Cys Xaa Cys Xaa Cys Xaa Cys

1 5

<210> 146

<211> 17

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (2)..(2)

<223> Any amino acid

<220><221> MOD_RES

<222> (4)..(4)

<223> Any amino acid

<220><221> MOD_RES

<222> (6)..(6)

<223> Any amino acid

<220><221> MOD_RES

<222> (8)..(10)

<223> Any amino acid

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><

221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<400> 146

Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa Cys Xaa His Xaa His Xaa

1 5 10 15

Cys

<210> 147

<211> 344

<212> PRT

<213> Artificial Sequence

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

polypeptide

<220><221> MISC_FEATURE

<222> (8)..(322)

<223> This region may encompass 1-15 "Tyr Lys Cys Pro Glu Cys

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

Pro" repeating units

<220><221> NON_CONS

<222> (18)..(19)

<220><221> NON_CONS

<222> (39)..(40)

<220><221> NON_CONS

<222> (60)..(61)

<220><221> NON_CONS

<222> (81)..(82)

<220><221> NON_CONS

<222> (102)..(103)

<220><221> NON_CONS

<222> (123)..(124)

<220><221> NON_CONS

<222> (144)..(145)

<220><221> NON_CONS

<222> (165)..(166)

<220><221> NON_CONS

<222> (186)..(187)

<220><221> NON_CONS

<222> (207)..(208)

<220><221> NON_CONS

<222> (228)..(229)

<220><221> NON_CONS

<222> (249)..(250)

<220><221> NON_CONS

<222> (270)..(271)

<220><221> NON_CONS

<222> (291)..(292)

<220><221> NON_CONS

<222> (312)..(313)

<220><221> NON_CONS

<222> (333)..(334)

<220>

<223> See specification as filed for detailed description of substitutions and preferred embodiments

<400> 147

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
20 25 30

Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys
35 40 45

Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr
50 55 60

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
65 70 75 80

Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro
100 105 110

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His
115 120 125

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser
130 135 140

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
145 150 155 160

Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr
180 185 190

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
195 200 205

Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly

210 215 220

Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
225 230 235 240
Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly
245 250 255
Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln
260 265 270

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
275 280 285
Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys
290 295 300
Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu
305 310 315 320
Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg
325 330 335

Thr His Thr Gly Lys Lys Thr Ser
340

<210> 148
<211> 292
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<220><221> MOD_RES
<222> (8)..(57)
<223> Any amino acid
<220><221> MISC_FEATURE
<222> (8)..(57)
<223> This region may encompass 1-50 residues
<220><221> MOD_RES
<222> (65)..(114)
<223> Any amino acid
<220><221> MISC_FEATURE

<222> (65)..(114)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (122)..(171)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (122)..(171)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (179)..(228)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (179)..(228)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (236)..(285)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (236)..(285)

<223> This region may encompass 1-50 residues

<400> 148

Arg Ser Asp Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp Asn Leu His Thr

50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

100 105 110

Xaa Xaa Arg Ser Asp Glu Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa

115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln Ser Gly Asn Leu

165 170 175

Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

210 215 220

Xaa Xaa Xaa Xaa Thr Ser Gly His Leu Val Arg Xaa Xaa Xaa Xaa Xaa

225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln Asn Ser

275 280 285

Thr Leu Thr Glu

290

<210> 149

<211> 292

<212> PRT

<213> Artificial Sequence

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

polypeptide

<220><221> MOD_RES

<222> (8)..(57)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (8)..(57)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (65)..(114)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (65)..(114)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (122)..(171)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (122)..(171)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (179)..(228)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (179)..(228)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (236)..(285)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (236)..(285)

<223> This region may encompass 1-50 residues

<400> 149

Arg Ser Asp Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Arg Thr Thr Leu Thr Asn

50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

100 105 110

Xaa Xaa Arg Glu Asp Asn Leu His Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa

115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser His Ser Leu

165 170 175

Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

210 215 220

Xaa Xaa Xaa Xaa Gln Ser Ser Ser Leu Val Arg Xaa Xaa Xaa Xaa Xaa

225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp

275 280 285

Asn Leu His Thr

290

<210> 150

<211> 292

<212> PRT

<213> Artificial Sequence

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

polypeptide

<220><221> MOD_RES

<222> (8)..(57)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (8)..(57)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (65)..(114)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (65)..(114)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (122)..(171)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (122)..(171)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (179)..(228)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (179)..(228)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (236)..(285)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (236)..(285)

<223> This region may encompass 1-50 residues

<400> 150

Asp Pro Gly Ala Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp Asn Leu Val Arg
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Gln Ser Gly Asp Leu Arg Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr His Leu Asp Leu
165 170 175

Ile Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220

Xaa Xaa Xaa Xaa Thr Ser Gly Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa
225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp
275 280 285

Asn Leu Val Arg
290

<210> 151

<211> 463

<212> PRT

<213> Artificial Sequence

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

<220><221> MOD_RES

<222> (8)..(57)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (8)..(57)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (65)..(114)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (65)..(114)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (122)..(171)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (122)..(171)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (179)..(228)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (179)..(228)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (236)..(285)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (236)..(285)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (293)..(342)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (293)..(342)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (350)..(399)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (350)..(399)

<223> This region may encompass 1-50 residues

<220><221> MOD_RES

<222> (407)..(456)

<223> Any amino acid

<220><221> MISC_FEATURE

<222> (407)..(456)

<223> This region may encompass 1-50 residues

<400> 151

Arg Arg Asp Glu Leu Asn Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp His Leu Thr Asn
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Arg Ser Asp Asp Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp Asn Leu
165 170 175

Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220
Xaa Xaa Xaa Xaa His Arg Thr Thr Leu Thr Asn Xaa Xaa Xaa Xaa Xaa
225 230 235 240
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245 250 255
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp
275 280 285
Asn Leu His Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
290 295 300
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
305 310 315 320
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
325 330 335

Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser His Ser Leu Thr Glu Xaa Xaa Xaa
340 345 350
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
355 360 365
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
370 375 380
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln
385 390 395 400

Ser Ser Ser Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
405 410 415
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
420 425 430
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
435 440 445
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp Asn Leu His Thr

450 455 460

<210> 152

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 152

Arg Ser Asp Asn Leu Val Arg

1 5

<210> 153

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 153

Arg Glu Asp Asn Leu His Thr

1 5

<210> 154

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 154

Arg Ser Asp Glu Leu Val Arg

1 5

<210> 155

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 155

Gln Ser Gly Asn Leu Thr Glu

1 5

<210> 156

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 156

Thr Ser Gly His Leu Val Arg

1 5

<210

> 157

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 157

Gln Asn Ser Thr Leu Thr Glu

1 5

<210> 158

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 158

Asp Pro Gly Ala Leu Val Arg

1 5

<210> 159

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 159

His Arg Thr Thr Leu Thr Asn

1 5

<210> 160

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 160

Gln Ser Gly Asp Leu Arg Arg

1 5

<210> 161

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 161

Thr Ser His Ser Leu Thr Glu

1 5

<210

> 162

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 162

Thr His Leu Asp Leu Ile Arg

1 5

<210> 163

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 163

Gln Ser Ser Ser Leu Val Arg

1 5

<210> 164

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 164

Thr Ser Gly Asn Leu Val Arg

1 5

<210> 165

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 165

Arg Arg Asp Glu Leu Asn Val

1 5

<210> 166

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 166

Arg Ser Asp Asp Leu Val Arg

1 5

<210

> 167

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 167

Arg Ser Asp His Leu Thr Asn

1 5

<210> 168

<211> 16

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 168

Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys

1 5 10 15

<210> 169

<211> 7

<212

> PRT

<213> Artificial Sequence

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

peptide

<400> 169

Pro Lys Lys Lys Arg Lys Val

1 5

<210> 170

<211> 16

<212> PRT

<213> Artificial Sequence

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

<400> 170

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 171

<211> 9

<212> PRT

<213> Artificial Sequence

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

<400> 171

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5

<210> 172

<211> 2009

<212> PRT

<213> Homo sapiens

<400> 172

Met Glu Gln Thr Val Leu Val Pro Pro Gly Pro Asp Ser Phe Asn Phe
1 5 10 15
Phe Thr Arg Glu Ser Leu Ala Ala Ile Glu Arg Arg Ile Ala Glu Glu
20 25 30
Lys Ala Lys Asn Pro Lys Pro Asp Lys Lys Asp Asp Asp Glu Asn Gly

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

100 105 110

Pro Phe Asn Pro Leu Arg Lys Ile Ala Ile Lys Ile Leu Val His Ser

115 120 125

Leu Phe Ser Met Leu Ile Met Cys Thr Ile Leu Thr Asn Cys Val Phe

130 135 140

Met Thr Met Ser Asn Pro Pro Asp Trp Thr Lys Asn Val Glu Tyr Thr

145 150 155 160

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

165 170 175

Gly Phe Cys Leu Glu Asp Phe Thr Phe Leu Arg Asp Pro Trp Asn Trp

180 185 190

Leu Asp Phe Thr Val Ile Thr Phe Ala Tyr Val Thr Glu Phe Val Asp

195 200 205

Leu Gly Asn Val Ser Ala Leu Arg Thr Phe Arg Val Leu Arg Ala Leu

210 215 220

Lys Thr Ile Ser Val Ile Pro Gly Leu Lys Thr Ile Val Gly Ala Leu

225 230 235 240

Ile Gln Ser Val Lys Lys Leu Ser Asp Val Met Ile Leu Thr Val Phe

245 250 255

Cys Leu Ser Val Phe Ala Leu Ile Gly Leu Gln Leu Phe Met Gly Asn

260 265 270

Leu Arg Asn Lys Cys Ile Gln Trp Pro Pro Thr Asn Ala Ser Leu Glu

275 280 285

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

290 295 300

Ile Asn Glu Thr Val Phe Glu Phe Asp Trp Lys Ser Tyr Ile Gln Asp

305 310 315 320

Ser Arg Tyr His Tyr Phe Leu Glu Gly Phe Leu Asp Ala Leu Leu Cys

325 330 335

Gly Asn Ser Ser Asp Ala Gly Gln Cys Pro Glu Gly Tyr Met Cys Val

340 345 350

Lys Ala Gly Arg Asn Pro Asn Tyr Gly Tyr Thr Ser Phe Asp Thr Phe

355 360 365

Ser Trp Ala Phe Leu Ser Leu Phe Arg Leu Met Thr Gln Asp Phe Trp

370 375 380

Glu Asn Leu Tyr Gln Leu Thr Leu Arg Ala Ala Gly Lys Thr Tyr Met

385 390 395 400

Ile Phe Phe Val Leu Val Ile Phe Leu Gly Ser Phe Tyr Leu Ile Asn

405 410 415

Leu Ile Leu Ala Val Val Ala Met Ala Tyr Glu Glu Gln Asn Gln Ala

420 425 430

Thr Leu Glu Glu Ala Glu Gln Lys Glu Ala Glu Phe Gln Gln Met Ile

435 440 445

Glu Gln Leu Lys Lys Gln Gln Glu Ala Ala Gln Gln Ala Ala Thr Ala

450 455 460

Thr Ala Ser Glu His Ser Arg Glu Pro Ser Ala Ala Gly Arg Leu Ser

465 470 475 480

Asp Ser Ser Ser Glu Ala Ser Lys Leu Ser Ser Lys Ser Ala Lys Glu

485 490 495

Arg Arg Asn Arg Arg Lys Lys Arg Lys Gln Lys Glu Gln Ser Gly Gly

500 505 510

Glu Glu Lys Asp Glu Asp Glu Phe Gln Lys Ser Glu Ser Glu Asp Ser

515 520 525

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

530 535 540

Tyr Glu Lys Arg Tyr Ser Ser Pro His Gln Ser Leu Leu Ser Ile Arg

545 550 555 560

Gly Ser Leu Phe Ser Pro Arg Arg Asn Ser Arg Thr Ser Leu Phe Ser

565 570 575

Phe Arg Gly Arg Ala Lys Asp Val Gly Ser Glu Asn Asp Phe Ala Asp

580 585 590

Asp Glu His Ser Thr Phe Glu Asp Asn Glu Ser Arg Arg Asp Ser Leu

595 600 605

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

610 615 620

Thr Ser Arg Ser Ser Arg Met Leu Ala Val Phe Pro Ala Asn Gly Lys

625 630 635 640

Met His Ser Thr Val Asp Cys Asn Gly Val Val Ser Leu Val Gly Gly

645 650 655

Pro Ser Val Pro Thr Ser Pro Val Gly Gln Leu Leu Pro Glu Val Ile

660 665 670

Ile Asp Lys Pro Ala Thr Asp Asp Asn Gly Thr Thr Thr Glu Thr Glu

675 680 685

Met Arg Lys Arg Arg Ser Ser Ser Phe His Val Ser Met Asp Phe Leu

690 695 700

Glu Asp Pro Ser Gln Arg Gln Arg Ala Met Ser Ile Ala Ser Ile Leu

705 710 715 720

Thr Asn Thr Val Glu Glu Leu Glu Glu Ser Arg Gln Lys Cys Pro Pro

725 730 735

Cys Trp Tyr Lys Phe Ser Asn Ile Phe Leu Ile Trp Asp Cys Ser Pro

740 745 750

Tyr Trp Leu Lys Val Lys His Val Val Asn Leu Val Val Met Asp Pro

755 760 765

Phe Val Asp Leu Ala Ile Thr Ile Cys Ile Val Leu Asn Thr Leu Phe

770 775 780

Met Ala Met Glu His Tyr Pro Met Thr Asp His Phe Asn Asn Val Leu

785 790 795 800

Thr Val Gly Asn Leu Val Phe Thr Gly Ile Phe Thr Ala Glu Met Phe

805 810 815

Leu Lys Ile Ile Ala Met Asp Pro Tyr Tyr Tyr Phe Gln Glu Gly Trp

820 825 830

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

835 840 845

```
Leu Ala Asn Val Glu Gly Leu Ser Val Leu Arg Ser Phe Arg Leu Leu
    850                 855                 860
Arg Val Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Met Leu Ile

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

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

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

    1055                1060                1065
Ile Gly  Lys Asp Leu Asp Tyr  Leu Lys Asp Val Asn  Gly Thr Thr
    1070                1075                1080
Ser Gly  Ile Gly Thr Gly Ser  Ser Val Glu Lys Tyr  Ile Ile Asp
```

1085 1090 1095

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

1100 1105 1110

Thr Val Pro Ile Ala Val Gly Glu Ser Asp Phe Glu Asn Leu Asn

1115 1120 1125

Thr Glu Asp Phe Ser Ser Glu Ser Asp Leu Glu Glu Ser Lys Glu

1130 1135 1140

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

1145 1150 1155

Ile Gly Ala Pro Val Glu Glu Gln Pro Val Val Glu Pro Glu Glu

1160 1165 1170

Thr Leu Glu Pro Glu Ala Cys Phe Thr Glu Gly Cys Val Gln Arg

1175 1180 1185

Phe Lys Cys Cys Gln Ile Asn Val Glu Glu Gly Arg Gly Lys Gln

1190 1195 1200

Trp Trp Asn Leu Arg Arg Thr Cys Phe Arg Ile Val Glu His Asn

1205 1210 1215

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

1220 1225 1230

Ala Leu Ala Phe Glu Asp Ile Tyr Ile Asp Gln Arg Lys Thr Ile

1235 1240 1245

Lys Thr Met Leu Glu Tyr Ala Asp Lys Val Phe Thr Tyr Ile Phe

1250 1255 1260

Ile Leu Glu Met Leu Leu Lys Trp Val Ala Tyr Gly Tyr Gln Thr

1265 1270 1275

Tyr Phe Thr Asn Ala Trp Cys Trp Leu Asp Phe Leu Ile Val Asp

1280 1285 1290

Val Ser Leu Val Ser Leu Thr Ala Asn Ala Leu Gly Tyr Ser Glu

1295 1300 1305

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

1310 1315 1320

Leu Arg Ala Leu Ser Arg Phe Glu Gly Met Arg Val Val Val Asn

1325 1330 1335

Ala Leu Leu Gly Ala Ile Pro Ser Ile Met Asn Val Leu Leu Val

1340 1345 1350

Cys Leu Ile Phe Trp Leu Ile Phe Ser Ile Met Gly Val Asn Leu

1355 1360 1365

Phe Ala Gly Lys Phe Tyr His Cys Ile Asn Thr Thr Thr Gly Asp

1370 1375 1380

Arg Phe Asp Ile Glu Asp Val Asn Asn His Thr Asp Cys Leu Lys

1385 1390 1395

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

1400 1405 1410

Asn Phe Asp Asn Val Gly Phe Gly Tyr Leu Ser Leu Leu Gln Val

1415 1420 1425

Ala Thr Phe Lys Gly Trp Met Asp Ile Met Tyr Ala Ala Val Asp

1430 1435 1440

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

1445 1450 1455

Met Tyr Leu Tyr Phe Val Ile Phe Ile Ile Phe Gly Ser Phe Phe

1460 1465 1470

Thr Leu Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln

1475 1480 1485

Gln Lys Lys Lys Phe Gly Gly Gln Asp Ile Phe Met Thr Glu Glu

1490 1495 1500

Gln Lys Lys Tyr Tyr Asn Ala Met Lys Lys Leu Gly Ser Lys Lys

1505 1510 1515

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

1520 1525 1530

Val Phe Asp Phe Val Thr Arg Gln Val Phe Asp Ile Ser Ile Met

1535 1540 1545

Ile Leu Ile Cys Leu Asn Met Val Thr Met Met Val Glu Thr Asp

1550 1555 1560

Asp Gln Ser Glu Tyr Val Thr Thr Ile Leu Ser Arg Ile Asn Leu

1565 1570 1575

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

1580 1585 1590

Ser Leu Arg His Tyr Tyr Phe Thr Ile Gly Trp Asn Ile Phe Asp

1595 1600 1605

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

1610 1615 1620

Leu Ile Glu Lys Tyr Phe Val Ser Pro Thr Leu Phe Arg Val Ile

1625 1630 1635

Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Ile Lys Gly Ala

1640 1645 1650

Lys Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser Leu Pro

1655 1660 1665

Ala Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe Ile

1670 1675 1680

Tyr Ala Ile Phe Gly Met Ser Asn Phe Ala Tyr Val Lys Arg Glu

1685 1690 1695

Val Gly Ile Asp Asp Met Phe Asn Phe Glu Thr Phe Gly Asn Ser

1700 1705 1710

Met Ile Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Gly

1715 1720 1725

Leu Leu Ala Pro Ile Leu Asn Ser Lys Pro Pro Asp Cys Asp Pro

1730 1735 1740

Asn Lys Val Asn Pro Gly Ser Ser Val Lys Gly Asp Cys Gly Asn

1745 1750 1755

Pro Ser Val Gly Ile Phe Phe Phe Val Ser Tyr Ile Ile Ile Ser

1760 1765 1770

Phe Leu Val Val Val Asn Met Tyr Ile Ala Val Ile Leu Glu Asn

1775 1780 1785

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

1790 1795 1800

Asp Phe Glu Met Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Asp

1805 1810 1815

Ala Thr Gln Phe Met Glu Phe Glu Lys Leu Ser Gln Phe Ala Ala

1820 1825 1830

Ala Leu Glu Pro Pro Leu Asn Leu Pro Gln Pro Asn Lys Leu Gln

1835 1840 1845

Leu Ile Ala Met Asp Leu Pro Met Val Ser Gly Asp Arg Ile His

1850 1855 1860

Cys Leu Asp Ile Leu Phe Ala Phe Thr Lys Arg Val Leu Gly Glu

1865 1870 1875

Ser Gly Glu Met Asp Ala Leu Arg Ile Gln Met Glu Glu Arg Phe

1880 1885 1890

Met Ala Ser Asn Pro Ser Lys Val Ser Tyr Gln Pro Ile Thr Thr

1895 1900 1905

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

1910 1915 1920

Arg Ala Tyr Arg Arg His Leu Leu Lys Arg Thr Val Lys Gln Ala

1925 1930 1935

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

1940 1945 1950

Leu Ile Lys Glu Asp Met Ile Ile Asp Arg Ile Asn Glu Asn Ser

1955 1960 1965

Ile Thr Glu Lys Thr Asp Leu Thr Met Ser Thr Ala Ala Cys Pro

1970 1975 1980

Pro Ser Tyr Asp Arg Val Thr Lys Pro Ile Val Glu Lys His Glu

1985 1990 1995

Gln Glu Gly Lys Asp Glu Lys Ala Lys Gly Lys

2000 2005

<210> 173

<211> 1052

<212> PRT

<213> Homo sapiens

<400> 173

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

1 5 10 15

Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly Val

20 25 30

Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg Ser

35 40 45

Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile Gln

50 55 60

Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His Ser

65 70 75 80

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

85 90 95

Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu Ala

100 105 110

Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr Gly

115 120 125

Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala Leu

130 135 140

Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys Asp

145 150 155 160

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

165 170 175

Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln Leu

180 185 190

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

195 200 205

Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys Asp

210 215 220

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

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

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

385 390 395 400
Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala Ile
405 410 415
Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln Gln
420 425 430
Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro Val
435 440 445
Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile Ile

450 455 460
Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg Glu

465 470 475 480

Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys Arg

485 490 495

Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr Gly

500 505 510

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

515 520 525

Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu Asp

530 535 540

Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro Arg

545 550 555 560

Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys Gln

565 570 575

Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu Ser

580 585 590

Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile Leu

595 600 605

Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu Tyr

610 615 620

Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp Phe

625 630 635 640

Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu Met

645 650 655

Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys Val

660 665 670

Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp Lys

675 680 685

Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp Ala

690 695 700

Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys Leu

705 710 715 720

Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys Gln
725 730 735
Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu Ile
740 745 750
Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr
755 760 765
Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn

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

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

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

965 970 975

Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu

980 985 990

Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn

995 1000 1005

Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr

1010 1015 1020

Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr

1025 1030 1035

Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly

1040 1045 1050

<210> 174

<211> 1148

<212> PRT

<213> Artificial Sequence

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

<400> 174

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala

1 5 10 15

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

20 25 30

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly

35 40 45

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg

50 55 60

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile

65 70 75 80

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His

85 90 95

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

100 105 110

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
115 120 125
Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
130 135 140
Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala

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

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

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

340 345 350
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala

355 360 365

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu

370 375 380

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser

385 390 395 400

Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile

405 410 415

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala

420 425 430

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln

435 440 445

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

450 455 460

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile

465 470 475 480

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

485 490 495

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys

500 505 510

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr

515 520 525

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

530 535 540

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu

545 550 555 560

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro

565 570 575

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

580 585 590

Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu

595 600 605

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
610 615 620
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
625 630 635 640
Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
645 650 655
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu

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

725 730 735
Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
740 745 750
Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
755 760 765
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
770 775 780
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile

785 790 795 800
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
805 810 815
Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
820 825 830
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
835 840 845
Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly

850 855 860

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr

865 870 875 880

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile

885 890 895

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

900 905 910

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr

915 920 925

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val

930 935 940

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser

945 950 955 960

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala

965 970 975

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

980 985 990

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile

995 1000 1005

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn

1010 1015 1020

Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser

1025 1030 1035

Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn

1040 1045 1050

Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys

1055 1060 1065

Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys

1070 1075 1080

Lys Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu

1085 1090 1095

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala

1100 1105 1110

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp

1115 1120 1125

Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe

1130 1135 1140

Asp Leu Asp Met Leu

1145

<210> 175

<211> 387

<212> PRT

<213> Homo sapiens

<400> 175

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

1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala

20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln

35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro

50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn

65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn

85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly

100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser

115 120 125

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

130 135 140

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

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

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

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Ala
340 345 350
Lys Ile His Leu Lys Gln Lys Glu Lys Lys Ala Glu Lys Gly Gly Ala
355 360 365
Pro Ser Ala Ser Ser Ala Pro Pro Val Ser Leu Ala Pro Val Val Thr
370 375 380
Thr Cys Ala
385

<210> 176

<211> 543

<212> PRT

<213> Homo sapiens

<400> 176

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

1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr

20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe

35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser

65 70 75 80

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

85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn

100 105 110

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

115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn

130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly

145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser

165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys

180 185 190

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

195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220
Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240
Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255
Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys

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

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

385 390 395 400
Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Thr Lys
405 410 415
Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
420 425 430
Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
435 440 445
Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr

450 455 460

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser

465 470 475 480

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala

485 490 495

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser

500 505 510

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu

515 520 525

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys

530 535 540

<210> 177

<211> 8

<212> PRT

<213> Artificial Sequence

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

<400> 177

Gly Gly Ser Gly Gly Gly Ser Gly

1 5

<210> 178

<211> 16

<212> PRT

<213> Artificial Sequence

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

<400> 178

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

1 5 10 15

<210> 179

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide
<220>
<223> See specification as filed for detailed description of
substitutions and preferred embodiments
<400> 179
Gly Gly Gly Ser
1
<210> 180
<211> 5
<212> PRT
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide
<220>
<223> See specification as filed for detailed description of
substitutions and preferred embodiments
<400> 180
Gly Gly Gly Gly Ser
1 5
<210> 181
<211> 4
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
peptide
<220>
<223> See specification as filed for detailed description of
substitutions and preferred embodiments
<400
> 181
Gly Gly Ser Gly
1
<210> 182

<211> 9

<212> DNA

<213> Homo sapiens

<400> 182

gcgkgggcg 9

<210> 183

<211> 76

<212> DNA

<213> Artificial Sequence

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

<400> 183

gttttagtac tctggaaaca gaatctacta aaacaaggca aaatgccgtg tttatctcgt 60

caacttgttg gcgaga 76

<210> 184

<211> 3871

<212> DNA

<213> Artificial Sequence

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

<400> 184

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccg ccgaccacct gatgctcgcc gagggctacc 2100

gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg ctccggactc 2160

tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg gctccgctgg 2220

ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg ccgccgtcct 2280

ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac ctgcagcctg 2340

tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga ggcccccccgg 2400

gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc ctgggcggca 2460

tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag ctggggctgc 2520

accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac tgcttctcgg 2580

acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt ggtggttctg 2640

gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg cccagcaaga 2700

cgcccccccа cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc cgcttctccc 2760

gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc ttccagtgcc 2820

gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc cgcacccaca 2880

caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg agcgatgaac 2940

ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct tgcccagtgg 3000

agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc cgcatccaca 3060

caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca agtggacatc 3120

tggtacgcca catccgcacc cacacaggcg aaaagccctt cgcctgcgac atctgtggaa 3180

gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg cggcagaagg 3240

acaagtaact cgaggaaacc cagcagacaa tgtagctaga cccagtagcc agatgtagct 3300

aaagagaccg gttcactgtg agaaacccag cagacaatgt agctagaccc agtagccaga 3360

tgtagctaaa gagaccggtt cactgtgaaa gcttgggtgg catccctgtg acccctcccc 3420

agtgcctctc ctggccctgg aagttgccac tccagtgccc accagccttg tcctaataaa 3480

attaagttgc atcattttgt ctgactaggt gtccttctat aatattatgg ggtggagggg 3540

ggtggtatgg agcaaggggc aagttgggaa gacaacctgt agggcctgcg gggtctattg 3600

ggaaccaagc tggagtgcag tggcacaatc ttggctcact gcaatctccg cctcctgggt 3660

tcaagcgatt ctcctgcctc agcctcccga gttgttggga ttccaggcat gcatgaccag 3720

gctcagctaa tttttgtttt tttggtagag acggggtttc accatattgg ccaggctggt 3780

ctccaactcc taatctcagg tgatctaccc accttggcct cccaaattgc tgggattaca 3840

ggcgtgaacc actgctccct tccctgtcct t 3871

<210> 185

<211> 22

<212> DNA

<213> Artificial Sequence

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

<400> 185

tgtctcggca ttgagaacat tc 22

<210> 186

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 186

attggtggga ggccattgta t 21

<210> 187

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 187

accacagtcc atgccatcac 20

<210> 188

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 188

tccaccaccc tgttgctgta 20

<210> 189

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 189

caaaaaagcc acaaaagcct 20

<210> 190

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 190

ttagctccgc aagaaacatc 20

<210> 191

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 191

ccatggaact ggctcgattt cac 23

<210> 192

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 192

attggtggga ggccactgta t 21

<210> 193

<211> 27

<212> DNA

<213> Artificial Sequence

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

<400> 193

aggcctgaaa accattgtgg gagccct 27

<210> 194

<211> 24

<212> DNA

<213> Artificial Sequence

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

primer

<400> 194

gaatgtggga aatcattcag tcgc 24

<210> 195

<211> 24

<212> DNA

<213> Artificial Sequence

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

primer

<400> 195

gcaagttatc ctctcgtgag aagg 24

<210> 196

<211> 29

<212> DNA

<213> Artificial Sequence

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

probe

<400> 196

gcgacaacct ggtgagacat caacgcacc 29

<210> 197

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 197

gctgttatct cttgtgggct gt 22

<210> 198

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 198

aaactcatgg gagctgccgg tt 22

<210> 199

<211> 25

<212> DNA

<213> Artificial Sequence

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

probe

<400> 199

ccacacaaat ctctccctgg cattg 25

<210> 200

<211> 24

<212> DNA

<213> Artificial Sequence

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

primer

<400> 200

gaatgtggga aatcattcag tcgc 24

<210> 201

<211> 24

<212> DNA

<213> Artificial Sequence

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

primer

<400> 201

gcaagttatc ctctcgtgag aagg 24

<210> 202

<211> 29

<212> DNA

<213> Artificial Sequence

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

<400> 202

gcgacaacct ggtgagacat caacgcacc 29

<210> 203

<211> 1182

<212> DNA

<213> Artificial Sequence

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

<400> 203

atggccgccg accacctgat gctcgccgag ggctaccgcc tggtgcagag gccgccgtcc 60

gccgccgccg cccatggccc tcatgcgctc cggactctgc cgccgtacgc gggccgggc 120

ctggacagtg ggctgaggcc gcgggggct ccgctggggc cgccgccgcc ccgccaaccc 180

ggggccctgg cgtacggggc cttcgggccg ccgtcctcct tccagccctt tccggccgtg 240

cctccgccgg ccgcgggcat cgcgcacctg cagcctgtgg cgacgccgta ccccggccgc 300

gcggccgcgc cccccaacgc tccgggaggc cccccgggcc cgcagccggc cccaagcgcc 360

gcagccccgc cgccgcccgc gcacgccctg ggcggcatgg acgccgaact catcgacgag 420

gaggcgctga cgtcgctgga gctggagctg gggctgcacc gcgtgcgcga gctgcccgag 480

ctgttcctgg gccagagcga gttcgactgc ttctcggact tggggtccgc gccgcccgcc 540

ggctccgtga gctgcggtgg ttctggtggt ggttctggtc agtcccagct catcaaaccc 600

agccgcatgc gcaagtaccc caaccggccc agcaagacgc ccccccacga acgcccttac 660

gcttgcccag tggagtcctg tgatcgccgc ttctcccgca gcgacaacct ggtgagacac 720

atccgcatcc acacaggcca gaagcccttc cagtgccgca tctgcatgag aaacttcagc 780

cgagaggata acttgcacac tcacatccgc acccacacag gcgaaaagcc cttcgcctgc 840

gacatctgtg gaagaaagtt tgcccggagc gatgaacttg tccgacatac caagatccac 900

ttgcggcaga aggaccgccc ttacgcttgc ccagtggagt cctgtgatcg ccgcttctcc 960

caatcaggga atctgactga gcacatccgc atccacacag gccagaagcc cttccagtgc 1020

cgcatctgca tgagaaactt cagcacaagt ggacatctgg tacgccacat ccgcacccac 1080

acaggcgaaa agcccttcgc ctgcgacatc tgtggaagaa agtttgccca gaatagtacc 1140

ctgaccgaac ataccaagat ccacttgcgg cagaaggaca ag 1182

<210> 204

<211> 1206

<212> DNA

<213> Artificial Sequence

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

<400> 204

atggccgccg accacctgat gctcgccgag ggctaccgcc tggtgcagag gccgccgtcc 60

gccgccgccg cccatggccc tcatgcgctc cggactctgc cgccgtacgc gggcccgggc 120

ctggacagtg ggctgaggcc gcgggggget ccgctggggc cgccgccgcc ccgccaaccc 180

ggggccctgg cgtacggggc cttcgggccg ccgtcctcct tccagccctt tccggccgtg 240

cctccgccgg ccgcgggcat cgcgcacctg cagcctgtgg cgacgccgta ccccggccgc 300

gcggccgcgc cccccaacgc tccgggaggc cccccgggcc cgcagccggc cccaagcgcc 360

gcagccccgc cgccgcccgc gcacgccctg ggcggcatgg acgccgaact catcgacgag 420

gaggcgctga cgtcgctgga gctggagctg gggctgcacc gcgtgcgcga gctgcccgag 480

ctgttcctgg gccagagcga gttcgactgc ttctcggact tggggtccgc gccgcccgcc 540

ggctccgtga gctgcggtgg ttctggtggt ggttctggtg gtggcagcgg gggaggttct 600

ggtcagtccc agctcatcaa acccagccgc atgcgcaagt accccaaccg gcccagcaag 660

acgccccccc acgaacgccc ttacgcttgc ccagtggagt cctgtgatcg ccgcttctcc 720

cgcagcgaca acctggtgag acacatccgc atccacacag gccagaagcc cttccagtgc 780

cgcatctgca tgagaaactt cagccgagag gataacttgc acactcacat ccgcacccac 840

acaggcgaaa agcccttcgc ctgcgacatc tgtggaagaa agtttgcccg gagcgatgaa 900

cttgtccgac ataccaagat ccacttgcgg cagaaggacc gcccttacgc ttgcccagtg 960

gagtcctgtg atcgccgctt ctcccaatca gggaatctga ctgagcacat ccgcatccac 1020

acaggccaga agcccttcca gtgccgcatc tgcatgagaa acttcagcac aagtggacat 1080

ctggtacgcc acatccgcac ccacacaggc gaaaagccct tcgcctgcga catctgtgga 1140

agaaagtttg cccagaatag taccctgacc gaacatacca agatccactt gcggcagaag 1200

gacaag 1206

<210> 205

<211> 402

<212> PRT

<213> Artificial Sequence

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

<400> 205

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gln Ser Gln Leu Ile Lys Pro
195 200 205

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His

210 215 220

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

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

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

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
385 390 395 400
Asp Lys

<210> 206
<211> 1707
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 206

atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ccgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcgccga ccacctgatg ctcgccgagg gctaccgcct ggtgcagagg 600

ccgccgtccg ccgccgccgc ccatggccct catgcgctcc ggactctgcc gccgtacgcg 660

ggcccgggcc tggacagtgg gctgaggccg cgggggggctc cgctggggcc gccgccgccc 720

cgccaacccg gggccctggc gtacggggcc ttcgggccgc cgtcctcctt ccagcccttt 780

ccggccgtgc ctccgccggc cgcgggcatc gcgcacctgc agcctgtggc gacgccgtac 840

cccggccgcg ccgccgcgcc ccccaacgct ccgggaggcc ccccgggccc gcagccggcc 900

ccaagcgccg cagccccgcc gccgcccgcg cacgccctgg gcggcatgga cgccgaactc 960

atcgacgagg aggcgctgac gtcgctggag ctggagctgg ggctgcaccg cgtgcgcgag 1020

ctgcccgagc tgttcctggg ccagagcgag ttcgactgct tctcggactt ggggtccgcg 1080

ccgcccgccg gctccgtgag ctgccagtcc cagctcatca aacccagccg catgcgcaag 1140

taccccaacc ggcccagcaa gacgcccccc cacgaacgcc cttacgcttg cccagtggag 1200

tcctgtgatc gccgcttctc ccgcagcgac aacctggtga gacacatccg catccacaca 1260

ggccagaagc ccttccagtg ccgcatctgc atgagaaact tcagccgaga ggataacttg 1320

cacactcaca tccgcaccca cacaggcgaa aagcccttcg cctgcgacat ctgtggaaga 1380

aagtttgccc ggagcgatga acttgtccga cataccaaga tccacttgcg gcagaaggac 1440

cgcccttacg cttgcccagt ggagtcctgt gatcgccgct tctcccaatc agggaatctg 1500

actgagcaca tccgcatcca cacaggccag aagcccttcc agtgccgcat ctgcatgaga 1560

aacttcagca caagtggaca tctggtacgc cacatccgca cccacacagg cgaaaagccc 1620

ttcgcctgcg acatctgtgg aagaaagttt gcccagaata gtaccctgac cgaacatacc 1680

aagatccact tgcggcagaa ggacaag 1707

<210> 207

<211> 569

<212> PRT

<213> Artificial Sequence

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

<400> 207

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Ala Asp His Leu Met Leu Ala
180 185 190

Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala Ala Ala Ala His
195 200 205

Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu
210 215 220
Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro
225 230 235 240
Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser
245 250 255

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

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

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
385 390 395 400
Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
405 410 415
Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
420 425 430
Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
435 440 445

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg

450 455 460

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp

465 470 475 480

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln

485 490 495

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro

500 505 510

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu

515 520 525

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp

530 535 540

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

545 550 555 560

Lys Ile His Leu Arg Gln Lys Asp Lys

565

<210> 208

<211> 1755

<212> DNA

<

213> Artificial Sequence

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

polynucleotide

<400> 208

atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ctgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcggcgg cagcggcggc ggaagcggcg ccgaccacct gatgctcgcc 600

gagggctacc gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg 660

ctccggactc tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg 720

gctccgctgg ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg 780

ccgccgtcct ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac 840

ctgcagcctg tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga 900

ggcccccccgg gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc 960

ctgggcggca tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag 1020

ctggggctgc accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac 1080

tgcttctcgg acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt 1140

ggtggttctg gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg 1200

cccagcaaga cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc 1260

cgcttctccc gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc 1320

ttccagtgcc gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc 1380

cgcacccaca caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg 1440

agcgatgaac ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct 1500

tgcccagtgg agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc 1560

cgcatccaca caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca 1620

agtggacatc tggtacgcca catccgcacc cacacaggcg aaaagccctt cgcctgcgac 1680

atctgtggaa gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg 1740

cggcagaagg acaag 1755

<210> 209

<211> 585

<212> PRT

<213> Artificial Sequence

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

<400> 209

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45
Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60
Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

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

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

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

Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr

275 280 285

Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly

290 295 300

Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala

305 310 315 320

Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser

325 330 335

Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu

340 345 350

Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala

355 360 365

Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser Gly

370 375 380

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg

385 390 395 400

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu

405 410 415

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile

420 425 430

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg

435 440 445

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr

450 455 460

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg

465 470 475 480

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp

485 490 495

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln

500 505 510

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro

515 520 525

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
530 535 540

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
545 550 555 560

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
565 570 575

Lys Ile His Leu Arg Gln Lys Asp Lys
580 585

<210> 210

<211> 1113

<212> DNA

<213

> Homo sapiens

<400> 210

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggaggaa gattcgaaat 480

aaaagatctg ctcaagagag ccgcaggaaa aagaaggtgt acgttggggg tttagagagc 540

cgggtcttga aatacacagc ccagaatatg gagcttcaga acaaagtaca gcttctggag 600

gaacagaatt tgtcccttct agatcaactg aggaaactcc aggccatggt gattgagatc 660

tcaaacaaaa ccagcagcag cagcacctgc atcttggtcc tgctagtctc cttctgcctc 720

ctccttgtac ctgctatgta ctcctctgac acaaggggga gcctgccagc tgagcatgga 780

gtgttgtccc gccagcttcg tgccctcccc agtgaggacc cttaccagct ggagctgcct 840

gccctgcagt cagaagtgcc gaaagacagc acacaccagt ggttggacgg ctcagactgt 900

gtactccagg cccctggcaa cacttcctgc ctgctgcatt acatgcctca ggctcccagt 960

gcagagcctc ccctggagtg gcccttccct gacctcttct cagagcctct ctgccgaggt 1020

cccatcctcc ccctgcaggc aaatctcaca aggaagggag gatggcttcc tactggtagc 1080

ccctctgtca ttttgcagga cagatactca ggc 1113

<210> 211

<211> 371

<212> PRT

<213> Homo sapiens

<400> 211

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu

1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg

20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp

35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu

50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr

65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser

85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu

100 105 110

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

115 120 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr

130 135 140

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

145 150 155 160

Lys Arg Ser Ala Gln Glu Ser Arg Arg Lys Lys Lys Val Tyr Val Gly

165 170 175

Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu

180 185 190

Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp

195 200 205

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

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

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

<210> 212

<211> 1590

<212> DNA

<213> Artificial Sequence

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

<400

> 212

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggctcga accaggtgaa 480

aaaccttaca aatgtcctga atgtgggaaa tcattcagtc gcagcgacaa cctggtgaga 540

catcaacgca cccatacagg agaaaaacct tataaatgtc cagaatgtgg aaagtccttc 600

tcacgagagg ataacttgca cactcatcaa cgaacacata ctggtgaaaa accatacaag 660

tgtcccgaat gtggtaaaag ttttagccgg agcgatgaac ttgtccgaca ccaacgaacc 720

catacaggcg agaagcctta caaatgtccc gagtgtggca agagcttctc acaatcaggg 780

aatctgactg agcatcaacg aactcatacc ggggaaaaac cttacaagtg tccagagtgt 840

gggaagagct tttccacaag tggacatctg gtacgccacc agaggacaca tacaggggag 900

aagccctaca aatgccccga atgcggtaaa agtttctctc agaatagtac cctgaccgaa 960

caccagcgaa cacacactgg gaaaaaaacg agtgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 213

<211> 530

<212> PRT

<213> Artificial Sequence

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

<400> 213

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

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

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

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

245 250 255

Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu

260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly

275 280 285

His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys

290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu

305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly

325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln

340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln

355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser

370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu

385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala

405 410 415

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

420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp

435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro

450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala

465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu

485 490 495

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

500 505 510

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

515 520 525

Ser Gly

530

<210> 214

<211> 1590

<212> DNA

<213> Artificial Sequence

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

<400> 214

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcttga gcccggagag 480

aagccgtaca agtgccctga gtgcggcaag tcttttagca gaagagacga acttaatgtc 540

caccagcgaa cgcatactgg tgaaaagccc tataaatgtc ctgaatgtgg gaaatcattc 600

tccagccgca gaacctgtag ggctcaccag cgaacacaca ccggcgaaaa accatacaaa 660

tgtccagaat gcgggaaatc cttttctcag tcatccaact tggtgagaca tcaacgcacg 720

cacactggag aaaagcctta caaatgcccg gaatgtggaa agtctttttc ccaattggcc 780

catttgcgag cccatcagag gactcacacg ggcgagaaac cttacaaatg cccggaatgc 840

gggaaatctt tttcaacgag tggcaacctc gtaagacacc aaagaacgca tacaggcgaa 900

aagccatata agtgtcctga gtgtggtaaa tcattctcac acaggaccac cctgacaaat 960

caccagcgca cgcacaccgg caagaagaca agcgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 215

<211> 530

<212> PRT

<213> Artificial Sequence

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

<400> 215

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu

1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg

20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp

35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu

50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr

65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser

85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu

100 105 110

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

115 120 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr

130 135 140

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

195 200 205

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

260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
275 280 285
Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
290 295 300
Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn
305 310 315 320
His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly

325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350
Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365
Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser

370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu

385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala

405 410 415

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

420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp

435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro

450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala

465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu

485 490 495

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

500 505 510

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

515 520 525

Ser Gly

530

<210> 216

<211> 1590

<212> DNA

<213> Artificial Sequence

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

<400> 216

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480

ccagtggagt cctgtgatcg ccgcttctcc cgcagcgaca acctggtgag acacatccgc 540

atccacacag gccagaagcc cttccagtgc cgcatctgca tgagaaactt cagccgagag 600

gataacttgc acactcacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc 660

tgtggaagaa agtttgcccg gagcgatgaa cttgtccgac ataccaagat ccacttgcgg 720

cagaaggacc gcccttacgc ttgcccagtg gagtcctgtg atcgccgctt ctcccaatca 780

gggaatctga ctgagcacat ccgcatccac acaggccaga agcccttcca gtgccgcatc 840

tgcatgagaa acttcagcac aagtggacat ctggtacgcc acatccgcac ccacacaggc 900

gaaaagccct tcgcctgcga catctgtgga agaaagtttg cccagaatag taccctgacc 960

gaacatacca agatccactt gcggcagaag gacgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 217

<211> 530

<212> PRT

<213> Artificial Sequence

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

<400> 217

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15
Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

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

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

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

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
225 230 235 240
Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg

245 250 255

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly

260 265 270

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser

275 280 285

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

290 295 300

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr

305 310 315 320

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly

325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln

340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln

355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser

370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu

385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala

405 410 415

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

420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp

435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro

450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala

465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu

485 490 495

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

500 505 510

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

515 520 525

Ser Gly

530

<210> 218

<211> 1590

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 218

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480

ccagtggagt cctgtgatcg ccgcttctcc cgctcagaca acctcgttcg acacatccgc 540

atccacacag gccagaagcc cttccagtgc cgcatctgca tgagaaactt cagccaccgg 600

actacactca cgaaccacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc 660

tgtggaagaa agtttgccag agaagacaat ctccatactc ataccaagat ccacttgcgg 720

cagaaggacc gcccttacgc ttgcccagtg gagtcctgtg atcgccgctt ctccaccagc 780

cattctctca ctgaacacat ccgcatccac acaggccaga agcccttcca gtgccgcatc 840

tgcatgagaa acttcagcca gtctagctca ctggtgaggc acatccgcac ccacacaggc 900

gaaaagccct tcgcctgcga catctgtgga agaaagtttg ccagggagga taacctgcat 960

acgcatacca agatccactt gcggcagaag gacgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 219

<211> 530

<212> PRT

<213> Artificial Sequence

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

<400> 219

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
100 105 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
115 120 125

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

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

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

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

370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu

385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala

405 410 415

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

420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp

435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro

450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala

465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu

485 490 495

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

500 505 510

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

515 520 525

Ser Gly

530

<210> 220

<211> 1140

<212> DNA

<213> Artificial Sequence

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

<400> 220

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggctcga accaggtgaa 480

aaaccttaca aatgtcctga atgtgggaaa tcattcagtc gcagcgacaa cctggtgaga 540

catcaacgca cccatacagg agaaaaacct tataaatgtc cagaatgtgg aaagtccttc 600

tcacgagagg ataacttgca cactcatcaa cgaacacata ctggtgaaaa accatacaag 660

tgtcccgaat gtggtaaaag ttttagccgg agcgatgaac ttgtccgaca ccaacgaacc 720

catacaggcg agaagcctta caaatgtccc gagtgtggca agagcttctc acaatcaggg 780

aatctgactg agcatcaacg aactcatacc ggggaaaaac cttacaagtg tccagagtgt 840

gggaagagct tttccacaag tggacatctg gtacgccacc agaggacaca tacaggggag 900

aagccctaca aatgccccga atgcggtaaa agtttctctc agaatagtac cctgaccgaa 960

caccagcgaa cacacactgg gaaaaaaacg agtgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatatca 1140

<210> 221

<211> 380

<212> PRT

<213> Artificial Sequence

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

<400> 221

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr

65 70 75 80

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

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

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
210 215 220
Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr
225 230 235 240
His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
245 250 255
Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu
260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
275 280 285
His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
290 295 300
Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu
305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350
Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365
Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser
370 375 380

<210> 222

<211> 1140

<212> DNA

<213> Artificial Sequence

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

<400> 222

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60
ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120
ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180
ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240
tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300
gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360
gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480
ccagtggagt cctgtgatcg ccgcttctcc cgcagcgaca acctggtgag acacatccgc 540
atccacacag gccagaagcc cttccagtgc cgcatctgca tgagaaactt cagccgagag 600
gataacttgc acactcacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc 660
tgtggaagaa agtttgcccg gagcgatgaa cttgtccgac ataccaagat ccacttgcgg 720
cagaaggacc gcccttacgc ttgcccagtg gagtcctgtg atcgccgctt ctcccaatca 780
gggaatctga ctgagcacat ccgcatccac acaggccaga agcccttcca gtgccgcatc 840

tgcatgagaa acttcagcac aagtggacat ctggtacgcc acatccgcac ccacacaggc 900
gaaaagccct tcgcctgcga catctgtgga agaaagtttg cccagaatag taccctgacc 960

gaacatacca agatccactt gcggcagaag gacgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

<210> 223

<211> 380

<212> PRT

<213> Artificial Sequence

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

<400> 223

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
100 105 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
115 120 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys
145 150 155 160

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
165 170 175

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
180 185 190

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg
195 200 205

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
210 215 220

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
225 230 235 240

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
245 250 255

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
260 265 270

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
275 280 285

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
290 295 300

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
305 310 315 320

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser
370 375 380

<210> 224

<211

> 155

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 224

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15
Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30
Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

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

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
115 120 125
Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
130 135 140
Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg
145 150 155

<210> 225

<211> 199

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 225

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

20 25 30

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

35 40 45

Ser Asn Lys Thr Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val

50 55 60

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

65 70 75 80

Gly Ser Leu Pro Ala Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala

85 90 95

Leu Pro Ser Glu Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser

100 105 110

Glu Val Pro Lys Asp Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys

115 120 125

Val Leu Gln Ala Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro

130 135 140

Gln Ala Pro Ser Ala Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu

145 150 155 160

Phe Ser Glu Pro Leu Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn

165 170 175

Leu Thr Arg Lys Gly Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile

180 185 190

Leu Gln Asp Arg Tyr Ser Gly

195

<210> 226

<211> 49

<212> PRT

<213> Artificial Sequence

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

<400> 226

Val Tyr Val Gly Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln

1 5 10 15

Asn Met Glu Leu Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu

20 25 30

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

35 40 45

Ser