



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0132392
(43) 공개일자 2014년11월17일

(51) 국제특허분류(Int. Cl.)
C12N 1/21 (2006.01) C12N 9/10 (2006.01)
C12N 9/24 (2006.01) C12P 7/18 (2006.01)
C12P 7/20 (2006.01) C12P 7/42 (2006.01)
C07K 14/245 (2006.01)
(21) 출원번호 10-2014-7027717
(22) 출원일자(국제) 2013년03월05일
심사청구일자 없음
(85) 번역문제출일자 2014년10월01일
(86) 국제출원번호 PCT/US2013/028970
(87) 국제공개번호 WO 2013/134174
국제공개일자 2013년09월12일
(30) 우선권주장
13/412,193 2012년03월05일 미국(US)

(71) 출원인
이 아이 듀폰 디 네모아 앤드 캄파니
미합중국 델라웨어주 (우편번호 19898) 월밍톤시
마아캣트 스트리트 1007
(72) 발명자
뢰블링-자스, 크리스틴
미국 19805 델라웨어주 월밍톤 라티머 드라이브 8
툼, 장-프랑수아
미국 19803 델라웨어주 월밍톤 하버힐 로드 627
(뒷면에 계속)
(74) 대리인
양영준, 심미성

전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **신규한 수크로스 수송체를 포함하는 재조합 박테리아**

(57) 요약

수크로스를 대사할 수 있는 재조합 박테리아가 기재된다. 재조합 박테리아는 수크로스 수송체 활성을 갖는 폴리펩티드를 암호화하는 신규한 뉴클레오티드 서열 및 수크로스 가수분해효소 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그들의 게놈 내에 또는 적어도 하나의 재조합 구축물 상에 포함한다. 이들 뉴클레오티드 서열은 동일하거나 상이한 프로모터에 각각 작동가능하게 연결된다. 수크로스를 대사하여 글리세롤 및/또는 글리세롤-유래 생성물, 예를 들어 1,3-프로판다이올 및 3-하이드록시프로피온산을 생성할 수 있는 재조합 박테리아도 기재된다.

(72) 발명자

반 디크, 티나, 케이.

미국 19803 델라웨어주 윌밍톤 캠브리지 드라이브
119

요우, 쟁

미국 19703 델라웨어주 클레이몬트 스킵 잭 레인
309

특허청구의 범위

청구항 1

(a) 수크로스 수송체 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열로서, 상기 폴리펩티드가 클러스탈(Clustal) W 정렬 방법에 기초하여, 서열 번호 68, 서열 번호 70, 서열 번호 72, 서열 번호 74, 서열 번호 76, 서열 번호 78, 서열 번호 80, 서열 번호 82, 서열 번호 84, 서열 번호 86, 서열 번호 88, 서열 번호 90, 서열 번호 92 및 서열 번호 94로 이루어진 군으로부터 선택되는 아미노산 서열과 비교시, 적어도 95% 서열 동일성을 갖는 뉴클레오티드 서열; 및

(b) 수크로스 가수분해효소(hydrolase) 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그의 게놈 내에 또는 적어도 하나의 재조합 구축물(construct) 상에 포함하는 재조합 박테리아로서,

(a) 및 (b)가 동일하거나 상이한 프로모터에 각각 작동가능하게 연결되고, 추가로 상기 재조합 박테리아가 수크로스를 대사할 수 있는 재조합 박테리아.

청구항 2

제1항에 있어서, 수크로스 가수분해효소 활성을 갖는 폴리펩티드가 EC 3.2.1.26 또는 EC 2.4.1.7로 분류되는 박테리아.

청구항 3

제 1항에 있어서, 프록토키나제 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그의 게놈 내에 또는 적어도 하나의 재조합 구축물 상에 추가로 포함하는 재조합 박테리아.

청구항 4

제3항에 있어서, 프록토키나제 활성을 갖는 폴리펩티드가 EC 2.7.1.4, EC 2.7.1.3 또는 EC 2.7.1.1로 분류되는 박테리아.

청구항 5

제1항에 있어서, 상기 박테리아가 하기의 속으로 이루어진 군으로부터 선택되는 재조합 박테리아:

에스케리키아(*Escherichia*), 클레브시엘라(*Klebsiella*), 시트로박터(*Citrobacter*) 및 아에로박터(*Aerobacter*).

청구항 6

제5항에 있어서, 상기 박테리아가 에스케리키아 콜라이(*Escherichia coli*)인 재조합 박테리아.

청구항 7

제 1항에 있어서, 재조합 박테리아가 1,3-프로판다이올, 글리세롤 및/또는 3-하이드록시프로피온산을 생성하는 재조합 박테리아.

청구항 8

a) 수크로스의 존재 하에서 제7항의 재조합 박테리아를 배양하는 단계; 및

b) 선택적으로, 생성된 글리세롤, 1,3-프로판다이올 및/또는 3-하이드록시프로피온산을 회수하는 단계를 포함하는 수크로스로부터 글리세롤, 1,3-프로판다이올 및/또는 3-하이드록시프로피온산의 제조 방법.

명세서

기술 분야

본 발명은 미생물학 및 분자 생물학 분야에 관한 것이다. 보다 구체적으로, 신규한 수크로스 수송체를 포함하는 재조합 박테리아, 및 수크로스로부터 글리세롤 및 글리세롤 유도 생성물과 같은 생성물을 생성하기 위한 이러한 재조합 박테리아의 이용 방법이 제공된다.

[0001]

배경 기술

- [0002] 많은 상업적으로 유용한 미생물은 그들의 주요 탄수화물원으로서 글루코스를 이용한다. 그러나, 글루코스의 높은 가격으로 상업적으로 바람직한 생성물의 생성을 위해 개발된 미생물에 의한 글루코스의 이용은 단점이 된다. 미생물 생성 시스템을 위한 탄수화물원으로서 수크로스의 이용 및 수크로스와 기타 당류 함유 혼합 공급원료(feedstock)의 이용은, 이들 재료가 일반적으로 더욱 낮은 가격으로 쉽게 입수가가능하기 때문에 상업적으로 더욱 바람직할 것이다.
- [0003] 생성 미생물은, 그것이 혼합 공급원료 내에 존재하는 임의의 수크로스를 이용할 수 있을 경우에 더욱 효율적으로 기능할 수 있다. 그러므로, 생성 미생물이 주요 탄소원으로서 수크로스를 효율적으로 이용하는 능력을 갖지 않는 경우, 그것은 효율적으로 작동할 수 없다. 예를 들어, 전형적으로 박테리아 세포는 당을 우선적으로 이용하는 것으로 나타나며, 글루코스를 가장 선호한다. 당의 혼합물을 함유하는 인공 배지에서, 전형적으로 다른 당에 앞서 글루코스가 전부 대사된다. 더욱이, 다수의 박테리아는 수크로스를 이용하는 능력이 결여되어 있다. 예를 들어, 에스케리키아 콜라이(*Escherichia coli*; *E. coli*) 균주의 50% 미만이 수크로스를 이용하는 능력을 가지고 있다. 따라서, 생성 미생물이 탄수화물원으로서 수크로스를 이용할 수 없는 경우, 미생물을 조작하여 그것이 수크로스를 이용할 수 있도록 하는 것이 바람직하다.
- [0004] 수크로스 이용 유전자를 통합함으로써 수크로스를 이용하도록 조작된 재조합 박테리아가 보고된 바 있다. 예를 들어, 리브시츠(Livshits) 등(미국 특허 제6,960,455호)은 수크로스 이용을 위한 대사 경로를 암호화하는(encoding) 유전자를 포함하는 에스케리키아 콜라이 균주를 이용한 아미노산의 생성을 기재하고 있다. 부가적으로, 문헌[Olson et al., (Appl. Microbiol. Biotechnol. 74:1031-1040, 2007)]에는 수크로스 분해를 담당하는 유전자를 지닌 에스케리키아 콜라이 균주가 기술되어 있으며, 이는 탄소원으로서 수크로스를 사용하여 L-타이로신 또는 L-페닐알라닌을 생성한다. 엘리엇(Eliot) 등(미국 특허 출원 공개 제2011/0136190 A1호)은 수크로스로부터 글리세롤 및 글리세롤 유도 생성물을 생성하는 재조합 박테리아를 기술하고 있다.
- [0005] 그러나, 새로운 수크로스 이용 유전자를 이용해서 수크로스를 이용하도록 조작되고, 개선된 수크로스 이용 능력을 갖는 박테리아 균주가 요구되고 있다. 추가적으로, 수크로스를 탄소원으로서 이용하여 글리세롤과 글리세롤 유도 생성물을 생성할 수 있는 박테리아 균주가 요구되고 있다.

발명의 내용

- [0006] 발명의 요약
- [0007] 일 실시형태는
- [0008] (a) 수크로스 수송체 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열로서, 상기 폴리펩티드가 클러스탈(Clustal) W 정렬 방법에 기초하여, 서열 번호 68, 서열 번호 70, 서열 번호 72, 서열 번호 74, 서열 번호 76, 서열 번호 78, 서열 번호 80, 서열 번호 82, 서열 번호 84, 서열 번호 86, 서열 번호 88, 서열 번호 90, 서열 번호 92 및 서열 번호 94로 이루어진 군으로부터 선택되는 아미노산 서열과 비교시, 적어도 95% 서열 동일성을 갖는 뉴클레오티드 서열; 및
- [0009] (b) 수크로스 가수분해효소(hydrolase) 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그의 게놈 내에 또는 적어도 하나의 재조합 구축물 상에 포함하는 재조합 박테리아를 제공하고,
- [0010] (a) 및 (b)는 동일하거나 상이한 프로모터에 각각 작동가능하게 연결되고(operably linked), 추가로 재조합 박테리아는 수크로스를 대사할 수 있다.
- [0011] 일 실시형태에서, 재조합 박테리아는 1,3-프로판다이올, 글리세롤 및/또는 3-하이드록시프로피온산을 생성한다.
- [0012] 다른 실시형태는 수크로스로부터 글리세롤, 1,3-프로판다이올 및/또는 3-하이드록시프로피온산을 제조하는 방법을 제공하며, 본 방법은
- [0013] a) 본 명세서에 개시된, 1,3-프로판다이올, 글리세롤 및/또는 3-하이드록시프로피온산을 생성하는 재조합 박테리아를 수크로스의 존재 하에 배양하는 단계; 및
- [0014] b) 선택적으로, 생성된 글리세롤, 1,3-프로판다이올 및/또는 3-하이드록시프로피온산을 회수하는 단계를 포함한다.

[0015] 간단한 서열 설명

[0016] 하기 서열들은 37 C.F.R. 1.821 1.825("뉴클레오티드 서열 및/또는 아미노산 서열 개시를 포함하는 특허 출원에 대한 요건 - 서열 규칙(Requirements for Patent Applicants Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules)")에 따르며, 세계 지적재산권 기구(WIPO) 표준 ST.25(2009) 및 EPO와 PCT(규칙 5.2 및 49.5(a bis)), 및 행정 지침(Administrative Instructions)의 섹션 208 및 부록 C의 서열 목록 요건에 따른다. 뉴클레오티드 및 아미노산 서열 데이터에 사용된 기호 및 체제는 37 C.F.R. § 1.822에 제시된 규정을 준수한다.

표 1

유전자의 요약 및 단백질 서열 번호

유전자	암호화 서열 서열 번호	암호화된 단백질 서열 번호
사카로마이세스 세레비시아(<i>Saccharomyces cerevisiae</i>) 유래의 GPD1	1	2
사카로마이세스 세레비시아 유래의 GPD2	3	4
사카로마이세스 세레비시아 유래의 GPP1	5	6
사카로마이세스 세레비시아 유래의 GPP2	7	8
클레브시엘라 뉴모니아(<i>Klebsiella pneumoniae</i>) 유래의 dhaB1	9	10
클레브시엘라 뉴모니아 유래의 dhaB2	11	12
클레브시엘라 뉴모니아 유래의 dhaB3	13	14
에스케리키아 콜라이 유래의 aldB	15	16
에스케리키아 콜라이 유래의 aldA	17	18
에스케리키아 콜라이 유래의 aldH	19	20
에스케리키아 콜라이 유래의 galP	21	22
에스케리키아 콜라이 EC3132 유래의 cscB	23	24
에스케리키아 콜라이 ATCC®13281 유래의 cscB	25	26
에스케리키아 콜라이 EC3132 유래의 cscA	27	28
에스케리키아 콜라이 ATCC13281 유래의 cscA	29	30
비피도박테리움 락티스(<i>Bifidobacterium lactis</i>) 균주 DSM 10140 ^T 유래의 bfrA	31	32
사카로마이세스 세레비시아 유래의 SUC2	33	34
코리네박테리움 글루타미쿰(<i>Corynebacterium glutamicum</i>) 유래의 scrB	35	36
류코노스톡 메센테로이데스(<i>Leuconostoc mesenteroides</i>) DSM 20193 유래의 수크로스 포스포릴라제 유전자	37	38
비피도박테리움 아돌레센티스(<i>Bifidobacterium adolescentis</i>) DSM 20083 유래의 sucP	39	40
아그로박테리움 투메파시엔스(<i>Agrobacterium tumefaciens</i>) 유래의 scrK	41	42
스트렙토코커스 뮤탄스(<i>Streptococcus mutans</i>) 유래의 scrK	43	44
에스케리키아 콜라이 유래의 scrK	45	46
클레브시엘라 뉴모니아 유래의 scrK	47	48
에스케리키아 콜라이 유래의 cscK	49	50
엔테로코커스 파에칼리스(<i>Enterococcus faecalis</i>) 유래의 cscK	51	52
사카로마이세스 세레비시아 유래의 HXK1	53	54
사카로마이세스 세레비시아 유래의 HXK2	55	56
클레브시엘라 뉴모니아 유래의 dhaT	57	58
클레브시엘라 뉴모니아 유래의 dhaX	59	60
시트로박터(<i>Citrobacter</i>) 종 유래의 scrT1	67	68
엔테로코커스 파에시움(<i>Enterococcus faecium</i>) 유래의 scrT3	69	70
코리네박테리움 글루쿠로노라이티쿰(<i>Corynebacterium glucuronolyticum</i>) 유래의 scrT4	71	72
비피도박테리움 애니말리스 하위종 락티스(<i>Bifidobacterium animalis subsp. lactis</i>) 유래의 scrT5	73	74

[0017]

비피도박테리움 갈리쿰(<i>Bifidobacterium gallicum</i>) 유래의 scrT6	75	76
비피도박테리움 롱쿰(<i>Bifidobacterium longum</i>) 유래의 scrT7	77	78
비피도박테리움 아틀레센티스 유래의 scrT8	79	80
비피도박테리움 롱쿰 유래의 scrT9	81	82
미트수오켈라 무타시나(<i>Mitsuokella multacida</i>) 유래의 scrT12	83	84
락토바실러스 안트리(<i>Lactobacillus antri</i>) 유래의 scrT13	85	86
락토바실러스 루미니스(<i>Lactobacillus ruminis</i>) 유래의 scrT14	87	88
예르시니아 프레데리크세나이(<i>Yersinia frederiksenii</i>) 유래의 scrT21	89	90
세라티아 프로테아마쿨란스(<i>Serratia proteamaculans</i>) 유래의 scrT25	91	92
에스케리키아 콜라이 유래의 scrT26	93	94
바이셀라 파라메센테로이데스(<i>Weissella paramesenteroides</i>) 유래의 scrT2	95	96
비피도박테리움 슈도칸테눌라툼(<i>Bifidobacterium pseudocatenulatum</i>) 유래의 scrT10	97	98
비피도박테리움 칸테눌라툼(<i>Bifidobacterium catenulatum</i>) 유래의 scrT11	99	100
클로스트리디움 써모셀룸(<i>Clostridium thermocellum</i>) 유래의 scrT15	101	102
그라눌리박터 베테스덴시스(<i>Granulibacter bethesdensis</i>) 유래의 scrT16	103	104
크로노박터 투리센시스(<i>Cronobacter turicensis</i>) 유래의 scrT17	105	106
엔테로박터(<i>Enterobacter</i>) 종 유래의 scrT18	107	108
시트로박터 코세리(<i>Citrobacter koseri</i>) 유래의 scrT19	109	110
바실러스 메가테리움(<i>Bacillus megaterium</i>) 유래의 scrT20	111	112
프로비덴시아 루스티지아니이(<i>Providencia rustigianii</i>) 유래의 scrT22	113	114
프로비덴시아 알칼리파시엔스(<i>Providencia alcalifaciens</i>) 유래의 scrT23	115	116
세라티아 오도리페라(<i>Serratia odorifera</i>) 유래의 scrT24	117	118
스트렙토마이세스 알부스(<i>Streptomyces albus</i>) 유래의 scrT27	119	120
아조박터 비넬란드ii(<i>Azotobacter vinelandii</i>) 유래의 scrT28	121	122
오에노코커스 오에니(<i>Oenococcus oeni</i>) 유래의 scrT29	123	124
락토바실러스 브레비스(<i>Lactobacillus brevis</i>) 유래의 scrT30	125	126
바이셀라 파라메센테로이데스 유래의 scrT31	127	128
만헤이미아 석시니시프로듀센스(<i>Mannheimia succiniciproducens</i>) 유래의 scrT32	129	130
아그레가티박터 아프로필러스(<i>Aggregatibacter aphrophilus</i>) 유래의 scrT33	131	132

[0018]

[0019] 서열 번호 61은 에스케리키아 콜라이 ATCC®13281 유래의 cscAKB 유전자 클러스터의 뉴클레오티드 서열이다.

[0020] 서열 번호 62는 플라스미드 pSYC0101의 뉴클레오티드 서열이다.

[0021] 서열 번호 63은 플라스미드 pSYC0103의 뉴클레오티드 서열이다.

[0022] 서열 번호 64는 플라스미드 pSYC0106의 뉴클레오티드 서열이다.

[0023] 서열 번호 65는 플라스미드 pSYC0109의 뉴클레오티드 서열이다.

[0024] 서열 번호 66은 플라스미드 pSYC0400/AGRO의 뉴클레오티드 서열이다.

[0025] 서열 번호 133은 플라스미드 pBHR-cscBKA의 뉴클레오티드 서열이다.

[0026] 서열 번호 134 내지 139 및 174 내지 177은 본 명세서의 실시예에 사용되는 프라이머의 뉴클레오티드 서열이다.

[0027] 서열 번호 140은 실시예 1 내지 34에 기재된 프로모터/MCS/이중 터미네이터 인서트의 뉴클레오티드 서열이다.

[0028] 서열 번호 141 내지 173은 본 명세서에 개시된 수크로스 수송체의 에스케리키아 콜라이에서의 발현을 위해 코돈 최적화된 핵산 서열이다.

발명을 실시하기 위한 구체적인 내용

- [0029] 본 명세서에 제시된 각각의 참고문헌의 개시내용은 그 전체 내용이 본 명세서에 참고로 포함된다.
- [0030] 본 명세서 및 첨부된 특허청구범위에 사용되는 바와 같이, 문맥이 명백하게 달리 지시하지 않는 한, 단수 형태 ("a", "an", 및 "the")는 복수 언급(plural reference)을 포함한다. 따라서, 예를 들어, "세포"에 대한 언급은 하나 이상의 세포 및 당업자에게 공지된 그의 균등물 등을 포함한다.
- [0031] 본 개시내용과 관련하여, 다수의 용어 및 약어가 사용된다. 하기의 정의가 제공된다.
- [0032] "오픈 리딩 프레임(open reading frame)"은 "ORF"로 약칭된다.
- [0033] "중합효소 연쇄 반응"은 "PCR"로 약칭된다.
- [0034] "아메리칸 타입 컬처 콜렉션(American Type Culture Collection)"은 "ATCC"로 약칭된다.
- [0035] 용어 "재조합 글리세롤-생성 박테리아"는 글리세롤 및/또는 글리세롤 유도 생성물을 생성할 수 있도록 유전자 조작된 박테리아를 의미한다.
- [0036] 용어 "수크로스 수송체 활성을 갖는 폴리펩티드"는 미생물 세포 내로의 수크로스 수송을 매개할 수 있는 폴리펩티드를 지칭한다.
- [0037] 용어 "프룩토키나제 활성을 갖는 폴리펩티드"는 D-프룩토스 + ATP에서 프룩토스-포스페이트 + ADP로의 전환을 촉매작용하는 능력을 갖는 폴리펩티드를 지칭한다. 전형적인 프룩토키나제는 EC 2.7.1.4이다. 프룩토스를 인산화하는 일부 능력을 갖는 효소는, 이 활성이 그들의 주된 활성인지 여부를 불문하고, 프룩토키나제로 지칭될 수 있다. 프룩토키나제 및 프룩토키나제 활성을 갖는 단백질을 암호화하는 유전자를 위해 사용되는 약어는 예를 들어, "Frk", "scrK", "cscK", "FK" 및 "KHK"를 포함한다. 프룩토키나제는 아그로박테리움 투메파시엔스 및 스트레토코커스 뮤탄스에서 scrK 유전자에 의해 암호화되며; 특정 에스케리키아 콜라이 균주에서 cscK 유전자에 의해 암호화된다.
- [0038] 용어 "수크로스 가수분해효소 활성을 갖는 폴리펩티드"는 수크로스의 가수분해를 촉매작용시켜, 글루코스 및 프룩토스를 생성하는 능력을 갖는 폴리펩티드를 지칭한다. 이러한 폴리펩티드는 종종 "인버타제(invertase)" 또는 " β -프룩토포라노시다제"로 지칭된다.
- [0039] 용어 "글리세롤 유도체" 및 "글리세롤 유도 생성물"은 본 명세서에서 상호교환가능하게 사용되며, 글리세롤로부터, 또는 글리세롤을 포함하는 경로에서 합성되는 화합물을 지칭한다. 이러한 생성물의 예는 3-하이드록시프로피온산, 메틸글리옥살, 1,2-프로판다이올 및 1,3-프로판다이올을 포함한다.
- [0040] 용어 "미생물 생성물"은 미생물에 의해 생성되는 생성물, 즉, 미생물이 기질을 대사시킨 결과물을 지칭한다. 미생물에 의해 생성물이 천연적으로 생성될 수 있거나, 생성물을 생성하도록 미생물을 유전자 조작할 수 있다. 용어 "포스포에놀피루베이트-당 포스포트랜스퍼라제 시스템", "PTS 시스템" 및 "PTS"는 본 명세서에서 상호교환가능하게 사용되며, 포스포에놀피루베이트-의존성 당 흡수 시스템을 말한다.
- [0041] 용어 "포스포캐리어(phosphocarrier) 단백질 HPr" 및 "PtsH"는 에스케리키아 콜라이에서 ptsH에 의해 암호화되는 포스포캐리어 단백질을 지칭한다. 용어 "포스포에놀피루베이트-단백질 포스포트랜스퍼라제" 및 "PtsI"는 에스케리키아 콜라이에서 ptsI에 의해 암호화되는 포스포트랜스퍼라제, EC 2.7.3.9를 지칭한다. 용어 "글루코스-특이적 IIA 컴포넌트" 및 "Crr"은 에스케리키아 콜라이에서 crr에 의해 암호화되는 EC 2.7.1.69로 표기되는 효소를 지칭한다. PtsH, PtsI, 및 Crr은 PTS 시스템을 포함한다.
- [0042] 용어 "PTS 마이너스"는, 그의 고유 상태에서 PTS 시스템을 함유하지 않는 미생물, 또는 PTS 유전자의 불활성화를 통해 PTS 시스템이 불활성화된 미생물을 지칭한다.
- [0043] 용어 "글리세롤-3-포스페이트 탈수소효소" 및 "G3PDH"는, 다이하이드록시아세톤 포스페이트(DHAP)에서 글리세롤 3-포스페이트(G3P)로의 전환을 촉매작용시키는 효소 활성을 담당하는 폴리펩티드를 지칭한다. 생체내 G3PDH는 NAD- 또는 NADP-의존성일 수 있다. 보조인자 특이적 글리세롤-3-포스페이트 탈수소효소를 구체적으로 지칭할 경우, 용어 "NAD-의존성 글리세롤-3-포스페이트 탈수소효소" 및 "NADP-의존성 글리세롤-3-포스페이트 탈수소효소"가 사용될 것이다. 일반적으로 NAD-의존성 및 NADP-의존성 글리세롤-3-포스페이트 탈수소효소가 NAD 및 NADP를 상호교환가능하게 사용할 수 있는 것이 사실이므로(예를 들어 gpsA에 의해 암호화되는 효소에 의해), 용어 NAD-의존성 및 NADP-의존성 글리세롤-3-포스페이트 탈수소효소는 상호교환가능하게 사용될 것이다. NAD-의존성 효소(EC 1.1.1.8)는, 예를 들어 본 명세서에서 DAR1로도 지칭되는, GPD1(서열 번호 1에 기재된 암호화 서열; 서열 번호 2에 기재된 암호화된 단백질 서열), 또는 GPD2(서열 번호 3에 기재된 암호화 서열; 서열 번호 4

에 기재된 암호화된 단백질 서열), 또는 GPD3을 포함하는 몇몇 유전자에 의하여 암호화된다. NADP-의존성 효소 (EC 1.1.1.94)는, 예를 들어 gpsA에 의해 암호화된다.

[0044] 용어 "글리세롤 3-포스파타제", "sn-글리세롤 3-포스파타제", "d,l-글리세롤 포스파타제" 및 "G3P 포스파타제"는 글리세롤 3-포스페이트 및 물에서 글리세롤 및 무기 포스페이트로의 전환을 촉매작용시킬 수 있는 효소 활성을 갖는 폴리펩티드를 지칭한다. G3P 포스파타제는 예를 들어 GPP1(서열 번호 5에 기재된 암호화 서열; 서열 번호 6에 기재된 암호화된 단백질 서열) 또는 GPP2(서열 번호 7에 기재된 암호화 서열; 서열 번호 8에 기재된 암호화된 단백질 서열)에 의하여 암호화된다.

[0045] 용어 "글리세롤 탈수효소" 또는 "탈수효소"는 글리세롤 분자에서 생성물인 3-하이드록시프로피온알데히드(3-HPA)로의 전환을 촉매작용시킬 수 있는 효소 활성을 갖는 폴리펩티드를 지칭한다.

[0046] 본 발명의 목적을 위한, 탈수효소는, 글리세롤 탈수효소(E.C. 4.2.1.30) 및 다이올 탈수효소(E.C. 4.2.1.28)를 포함하며, 이들은 각각 글리세롤 및 1,2-프로판다이올을 바람직한 기질로 갖는다. 탈수효소를 위한 유전자는, 특히 클레브시엘라 뉴모니아, 시트로박터 프레운다이(*Citrobacter freundii*), 클로스트리디움 파스퇴리아눔(*Clostridium pasteurianum*), 살모넬라 타이피무리움(*Salmonella typhimurium*), 클레브시엘라 옥시토카(*Klebsiella oxytoca*), 및 락토바실러스 레우테리(*Lactobacillus reuteri*)에서 확인되었다. 각 경우에서, 탈수효소는 3개의 서브유닛으로 구성된다: 큰 서브유닛 또는 "α" 서브유닛, 중간 서브유닛 또는 "β" 서브유닛 및 작은 서브유닛 또는 "γ" 서브유닛. 유전자는 또한, 예를 들어 문헌[Daniel et al. (FEMS Microbiol. Rev. 22, 553 (1999))] 및 문헌[Toraya and Mori (J. Biol. Chem. 274, 3372 (1999))]에 기재되어 있다. 글리세롤 탈수효소의 큰 서브유닛 또는 "α"(알파) 서브유닛을 암호화하는 유전자는 dhaB1(서열 번호 9에 기재된 암호화 서열, 서열 번호 10에 기재된 암호화된 단백질 서열), gldA 및 dhaB를 포함하고; 중간 서브유닛 또는 "β"(베타) 서브유닛을 암호화하는 유전자는 dhaB2(서열 번호 11에 기재된 암호화 서열, 서열 번호 12에 기재된 암호화된 단백질 서열), gldB 및 dhaC를 포함하고; 작은 서브유닛 또는 "γ"(감마) 서브유닛을 암호화하는 유전자는 dhaB3(서열 번호 13에 기재된 암호화 서열, 서열 번호 14에 기재된 암호화된 단백질 서열), gldC 및 dhaE를 포함한다. 다이올 탈수효소의 큰 서브유닛 또는 "α" 서브유닛을 암호화하는 기타 유전자는 pduC 및 pduA를 포함하고; 중간 서브유닛 또는 "β" 서브유닛을 암호화하는 기타 유전자는 pduD 및 pduB를 포함하고; 작은 서브유닛 또는 "γ" 서브유닛을 암호화하는 기타 유전자는 pduE 및 pduC를 포함한다.

[0047] 글리세롤 및 다이올 탈수효소는 글리세롤 및 일부 다른 기질에 의해 메커니즘-기반의 자멸적 불활성화(suicide inactivation)를 겪는다(문헌[Daniel et al., FEMS Microbiol. Rev. 22, 553 (1999)]). 용어 "탈수효소 재활성화 인자"는 탈수효소 활성의 재활성화를 담당하는 단백질을 지칭한다. 용어 "탈수효소 재활성화 활성", "탈수효소 활성의 재활성화" 및 "탈수효소 활성의 재생"은 상호교환가능하게 사용되며, 반응을 촉매작용시킬 수 없는 탈수효소를 반응을 촉매작용시킬 수 있는 것으로 전환시키는 현상 또는 탈수효소의 불활성화를 억제하는 현상 또는 생체내 탈수효소의 가용 반감기를 연장하는 현상을 지칭한다. 2가지 단백질이 탈수효소 재활성화 인자로서 관여되는 것으로 확인되었다(예를 들어, 미국 특허 제6,013,494호 및 그에 기재된 참고문헌; 상기 문헌[Daniel et al.]; 문헌[Toraya and Mori, J. Biol. Chem. 274, 3372 (1999)]; 및 문헌[Tobimatsu et al., J. Bacteriol. 181, 4110 (1999)] 참조). 단백질 중 하나를 암호화하는 유전자는, 예를 들어, orfZ, dhaB4, gdrA, pduG 및 ddrA를 포함한다. 2가지 단백질 중 제2 단백질을 암호화하는 유전자는, 예를 들어, orfX, orf2b, gdrB, pduH 및 ddrB를 포함한다.

[0048] 용어 "1,3-프로판다이올 산화환원효소", "1,3-프로판다이올 탈수소효소" 및 "DhaT"는 본 명세서에서 상호교환가능하게 사용되며, 3-HPA 및 1,3-프로판다이올의 상호전환을 촉매작용시킬 수 있는 효소 활성을 갖는 폴리펩티드(들)를 지칭하되, 단 이러한 활성을 암호화하는 유전자(들)는 그의 천연(즉, 야생형) 환경에서 탈수효소에 물리적으로 또는 전사에 의해 연결된 것으로 보이며; 예를 들어, 유전자는 dha 레귤론 내에서 발견되며, 이는 클레브시엘라 뉴모니아 유래의 dhaT의 경우와 같다. 1,3-프로판다이올 산화환원효소를 암호화하는 유전자는 클레브시엘라 뉴모니아, 시트로박터 프레운다이 및 클로스트리디움 파스퇴리아눔 유래의 dhaT를 포함하나, 이에 한정되지 않는다. 이들 유전자 각각은 III형 알코올 탈수소효소의 패밀리에 속하는 폴리펩티드를 암호화하며, 이는 보존된 철-결합 모티프를 나타내고, 3-HPA 및 1,3-프로판다이올의 NAD⁺/NADH 결합된 상호전환에 대한 선호를 갖는다(문헌[Johnson and Lin, J. Bacteriol. 169, 2050 (1987)]; 문헌[Daniel et al., J. Bacteriol. 177, 2151 (1995)]; 및 문헌[Leurs et al., FEMS Microbiol. Lett. 154, 337 (1997)]). 유사한 물리적 특성을 가진 효소들이 락토바실러스 브레비스 및 락토바실러스 부크네리(*Lactobacillus buchneri*)로부터 분리된 바 있다(문헌[Veiga da Dunha and Foster, Appl. Environ. Microbiol. 58, 2005 (1992)]).

- [0049] 용어 "dha 레귤론"은, 탈수효소 활성, 재활성화 활성, 및 1,3-프로판다이올 산화환원효소를 포함하나 이들에 한정되지 않는 다양한 생물학적 활성을 갖는 폴리펩티드를 암호화하는 관련 폴리뉴클레오티드 또는 오픈 리딩 프레임의 세트를 지칭한다. 전형적으로 dha 레귤론은, 미국 특허 제7,371,558호에 기술된 바와 같이 오픈 리딩 프레임 dhaR, orfY, dhaT, orfX, orfW, dhaB1, dhaB2, dhaB3, 및 orfZ를 포함한다.
- [0050] 용어 "알데히드 탈수소효소" 및 "Ald"는 알데히드에서 카르복실산으로의 전환을 촉매작용시키는 폴리펩티드를 지칭한다. 알데히드 탈수소효소는 NAD, NADP, FAD, 또는 PQQ와 같은 산화환원 보조인자를 사용할 수 있다. 전형적인 알데히드 탈수소효소로, EC 1.2.1.3(NAD-의존성); EC 1.2.1.4(NADP-의존성); EC 1.2.99.3(PQQ-의존성); 또는 EC 1.2.99.7(FAD-의존성)이 있다. NADP-의존성 알데히드 탈수소효소의 일례로, 에스케리키아 콜라이 유전자 aldB(서열 번호 15에 기재된 암호화 서열)에 의해 암호화된 AldB(서열 번호 16)가 있다. NAD-의존성 알데히드 탈수소효소는 예에는 에스케리키아 콜라이 유전자 aldA(서열 번호 17에 기재된 암호화 서열)에 의해 암호화된 AldA(서열 번호 18); 및 에스케리키아 콜라이 유전자 aldH(서열 번호 19에 기재된 암호화 서열)에 의해 암호화된 AldH(서열 번호 20)가 포함된다.
- [0051] 용어 "글루코키나제" 및 "Glc"는 본 명세서에서 상호교환가능하게 사용되며, D-글루코스 + ATP에서 글루코스 6-포스페이트 + ADP로의 전환을 촉매작용시키는 단백질을 지칭한다. 전형적인 글루코키나제는 EC 2.7.1.2이다. 글루코키나제는 에스케리키아 콜라이에서 glk에 의해 암호화된다.
- [0052] 용어 "포스포에놀피루베이트 카르복실라제" 및 "Ppc"는 본 명세서에서 상호교환가능하게 사용되며, 포스포에놀피루베이트 + H₂O + CO₂에서 포스페이트 + 옥살로아세트산으로의 전환을 촉매작용시키는 단백질을 지칭한다. 전형적인 포스포에놀피루베이트 카르복실라제는 EC 4.1.1.31이다. 포스포에놀피루베이트 카르복실라제는 에스케리키아 콜라이에서 ppc에 의해 암호화된다.
- [0053] 용어 "글리세르알데히드-3-포스페이트 탈수소효소" 및 "GapA"는 본 명세서에서 상호교환가능하게 사용되며, 글리세르알데히드 3-포스페이트 + 포스페이트 + NAD⁺에서 3-포스포-D-글리세로일-포스페이트 + NADH + H⁺로의 전환을 촉매작용시킬 수 있는 효소 활성을 갖는 단백질을 지칭한다. 전형적인 글리세르알데히드-3-포스페이트 탈수소효소는 EC 1.2.1.12이다. 글리세르알데히드-3-포스페이트 탈수소효소는 에스케리키아 콜라이에서 gapA에 의해 암호화된다.
- [0054] 용어 "유기 호흡 제어 단백질" 및 "ArcA"는 본 명세서에서 상호교환가능하게 사용되며, 총체적 조절 단백질(global regulatory protein)을 지칭한다. 유기 호흡 제어 단백질은 에스케리키아 콜라이에서 arcA에 의해 암호화된다.
- [0055] 용어 "메틸글리옥살 신타제" 및 "MgsA"는 본 명세서에서 상호교환가능하게 사용되며, 다이하이드록시아세톤 포스페이트에서 메틸글리옥살 + 포스페이트로의 전환을 촉매작용시킬 수 있는 효소 활성을 갖는 단백질을 지칭한다. 전형적인 메틸글리옥살 신타제는 EC 4.2.3.3이다. 메틸글리옥살 신타제는 에스케리키아 콜라이에서 mgsA에 의해 암호화된다.
- [0056] 용어 "포스포글루코네이트 탈수효소" 및 "Edd"는 본 명세서에서 상호교환가능하게 사용되며, 6-포스포-글루코네이트에서 2-케토-3-데옥시-6-포스포-글루코네이트 + H₂O로의 전환을 촉매작용시킬 수 있는 효소 활성을 갖는 단백질을 지칭한다. 전형적인 포스포글루코네이트 탈수효소는 EC 4.2.1.12이다. 포스포글루코네이트 탈수효소는 에스케리키아 콜라이에서 edd에 의해 암호화된다.
- [0057] 용어 "YciK"는 yciK에 의해 암호화되는 추정 효소를 지칭하며, 이는 에스케리키아 콜라이에서 코브(I)알라민 아데노실트랜스퍼라제를 암호화하는 유전자, btuR에 번역에 의해 연결된다.
- [0058] 용어 "코브(I)알라민 아데노실트랜스퍼라제"는 데옥시아데노실 부분을 ATP로부터 환원된 코리노이드로 전달할 수 있는 효소를 지칭한다. 전형적인 코브(I)알라민 아데노실트랜스퍼라제는 EC 2.5.1.17이다. 코브(I)알라민 아데노실트랜스퍼라제는 에스케리키아 콜라이에서 "btuR", 살모넬라 타이피무리움에서 "cobA" 및 슈도모나스 데니트리피칸스(*Pseudomonas denitrificans*)에서 "cob0" 유전자에 의해 암호화된다.
- [0059] 용어 "갈락토스-양성자 공동수송체" 및 "GalP"는 본 명세서에서 상호교환가능하게 사용되며, 당 및 양성자를 주변 세포질로부터 세포질로 수송할 수 있는 효소 활성을 갖는 단백질을 지칭한다. D-글루코스는 GalP에 있어서 선호되는 기질이다. 갈락토스-양성자 공동수송체는 에스케리키아 콜라이에서 galP에 의해 암호화된다(서열 번호 21에 기재된 암호화 서열, 서열 번호 22에 기재된 암호화된 단백질 서열).

- [0060] 용어 "비-특이적 촉매 활성"은 3-HPA 및 1,3-프로판다이올의 상호 전환을 촉매작용시킬 수 있는 효소 활성을 갖는 폴리펩티드(들)를 지칭하며, 특이적으로 1,3-프로판다이올 산화환원효소(들)를 배제한다. 전형적으로 이들 효소는 알코올 탈수소효소이다. 이러한 효소는, FAD 또는 FMN과 같은 플라빈을 포함하나, 이에 한정되지 않는, NAD⁺/NADH 외의 보조인자를 이용할 수 있다. 비-특이적 알코올 탈수소효소를 위한 유전자(yqhD)는, 예를 들어, 에스케리키아 콜라이 K-12 균주 내에서 내인적으로 암호화되고 기능적으로 발현되는 것으로 관찰된다.
- [0061] 용어 "1.6 장(long) GI 프로모터", "1.20 단(short)/장 GI 프로모터" 및 "1.5 장 GI 프로모터"는, 미국 특허 제7,132,527호에 기재된 바와 같이 스트렙토마이세스 리비단스(*Streptomyces lividans*) 글루코스 아이소머라제 유전자 유래의 프로모터를 함유하는 폴리뉴클레오티드 또는 단편을 지칭한다. 이들 프로모터 단편은 야생형 스트렙토마이세스 리비단스 글루코스 아이소머라제 유전자 프로모터에 비하여 그들의 활성을 감소시키는 돌연변이를 포함한다.
- [0062] 용어 "기능" 및 "효소 기능"은 본 명세서에서 상호교환가능하게 사용되며, 반응에 의해 그 자신은 소모되지 않으면서 특정 화학 반응이 일어나는 속도를 변경시키는 효소의 촉매 활성을 지칭한다. 적절한 조건 하에 생성물 또는 기질의 생성이 이루어질 수 있는 평형 중의 반응에 이러한 활성이 적용될 수 있다는 것이 이해된다.
- [0063] 용어 "폴리펩티드" 및 "단백질"은 본 명세서에 상호교환가능하게 사용된다.
- [0064] 용어 "탄소 기질" 및 "탄소원"은 본 명세서에서 상호교환가능하게 사용되며, 본 명세서에서 개시된 재조합 박테리아에 의하여 대사될 수 있는 탄소원, 특히 수크로스를 포함하는 탄소원을 지칭한다. 탄소원은 다른 당당류, 이당류, 올리고당류, 또는 다당류를 더 포함할 수 있다.
- [0065] 용어 "숙주 세포" 및 "숙주 박테리아"는 본 명세서에서 상호교환가능하게 사용되며, 외래 또는 이종 유전자를 수용할 수 있고 이들 유전자를 발현시켜 활성 유전자 생성물을 생성할 수 있는 박테리아를 지칭한다.
- [0066] 본 명세서에 사용되는 바와 같이, 용어 "생성 미생물"은, 재조합인 것들을 포함하나 이에 한정되지 않으며, 1,3-프로판다이올, 글리세롤, 3-하이드록시프로피온산, 다중불포화 지방산 등과 같은 특정 생성물을 제조하기 위해 사용되는 미생물을 지칭한다.
- [0067] 본 명세서에 사용되는 바와 같이, "핵산"은 폴리뉴클레오티드를 의미하며 데옥시리보뉴클레오티드 또는 리보뉴클레오티드 염기의 단일 또는 이중-가닥 폴리머를 포함한다. 핵산은 또한 단편 및 변형된 뉴클레오티드를 포함할 수 있다. 따라서, 용어 "폴리뉴클레오티드", "핵산 서열", "뉴클레오티드 서열" 또는 "핵산 단편"은 본 명세서에서 상호교환가능하게 사용되며, 합성된 비-천연 뉴클레오티드 염기 또는 변형된 뉴클레오티드 염기를 임의로 함유하는 단일- 또는 이중-가닥 RNA 또는 DNA의 폴리머를 지칭한다. 뉴클레오티드(일반적으로 그의 5'-모노포스페이트 형태로 발견됨)는 다음과 같은, 그들의 단일 문자 표기로 언급된다: 아데닐레이트 또는 데옥시아데닐레이트(각각 RNA 또는 DNA에 대한 것)는 "A", 시티딜레이트 또는 데옥시시티딜레이트는 "C", 구아닐레이트 또는 데옥시구아닐레이트는 "G", 유리딜레이트는 "U", 데옥시티미딜레이트는 "T", 퓨린(A 또는 G)은 "R", 피리미딘(C 또는 T)은 "Y", G 또는 T는 "K", A 또는 C 또는 T는 "H", "이노신"은 "I", 및 임의의 뉴클레오티드는 "N".
- [0068] 폴리뉴클레오티드는 합성된, 비천연 뉴클레오티드 염기 또는 변형된 뉴클레오티드 염기를 임의로 함유하는 단일- 또는 이중-가닥인 RNA 또는 DNA의 폴리머일 수 있다. DNA의 폴리머 형태인 폴리뉴클레오티드는 cDNA, 게놈 DNA, 합성 DNA, 또는 그의 혼합물의 하나 이상의 세그먼트로 이루어질 수 있다.
- [0069] "유전자"는 특정 단백질을 발현시키는 핵산 단편을 지칭하며, 이는 암호화 영역만을 지칭하거나 암호화 서열 앞에 있는 조절 서열(5' 비암호화 서열) 및 암호화 서열 뒤에 있는 조절 서열(3' 비암호화 서열)을 포함할 수 있다. "고유 유전자"는 그 자체의 조절 서열과 함께 천연에서 발견되는 유전자를 말한다. "키메라 유전자"는 천연에서 함께 발견되지 않는 조절 및 암호화 서열을 포함하는, 고유 유전자가 아닌 임의의 유전자를 말한다. 따라서, 키메라 유전자는 상이한 공급원으로부터 유래된 조절 서열과 암호화 서열, 또는 동일한 공급원으로부터 유래되었지만, 천연에서 발견되는 것과 상이한 방식으로 배열된 조절 서열과 암호화 서열을 포함할 수 있다. "내인성 유전자"는 유기체의 게놈에서 그의 천연의 위치에 있는 고유 유전자를 말한다. "외래" 유전자는 유전자 전달에 의해 숙주 유기체 내로 도입된 유전자를 지칭한다. 외래 유전자는 비고유 유기체에 삽입된 유전자, 고유 숙주 내의 새로운 위치에 도입된 유전자, 또는 키메라 유전자를 포함할 수 있다.
- [0070] 용어 "고유 뉴클레오티드 서열"은 숙주 미생물에서 통상 관찰되는 뉴클레오티드 서열을 지칭한다.
- [0071] 용어 "비-고유 뉴클레오티드 서열"은 숙주 미생물에서 통상 관찰되지 않는 뉴클레오티드 서열을 지칭한다.

- [0072] 용어 "고유 폴리펩티드"는 숙주 미생물에서 통상 관찰되는 폴리펩티드를 지칭한다.
- [0073] 용어 "비-고유 폴리펩티드"는 숙주 미생물에서 통상 관찰되지 않는 폴리펩티드를 지칭한다.
- [0074] 용어 "암호화하는" 및 "암호화"는 본 명세서에서 상호교환가능하게 사용되며, 유전자가 전사 및 번역의 메커니즘을 통해 아미노산 서열을 생성하는 과정을 지칭한다.
- [0075] 용어 "암호화 서열"은 특정 아미노산 서열을 암호화하는 뉴클레오티드 서열을 지칭한다.
- [0076] "적절한 조절 서열"은 암호화 서열의 업스트림(5' 비암호화 서열), 이의 내부 또는 이의 다운스트림(3' 비암호화 서열)에 위치하고, 전사, RNA 프로세싱 또는 안정성, 또는 관련된 암호화 서열의 번역에 영향을 미치는 뉴클레오티드 서열을 말한다. 조절 서열은 프로모터, 인핸서(enhancer), 사일런서(silencer), 5' 비번역 리더(untranslated leader) 서열(예를 들어, 전사 시작 부위와 번역 개시 코돈의 사이), 인트론, 폴리아데닐화 인식 서열, RNA 프로세싱 부위, 이펙터(effector) 결합 부위 및 스텝-루프 구조를 포함할 수 있다.
- [0077] 용어 "발현 카세트"는 선택된 유전자의 암호화 서열 및 선택된 유전자 생성물의 발현을 위해 필요한 암호화 서열의 앞에 있는 조절 서열(5' 비암호화 서열) 및 뒤에 있는 조절 서열(3' 비암호화 서열)을 포함하는 DNA의 단편을 지칭한다. 따라서, 발현 카세트는 전형적으로 1) 프로모터 서열; 2) 암호화 서열(즉, ORF) 및 3) 진핵생물에서 대개 폴리아데닐화 부위를 포함하는 3' 비번역 영역(예를 들어, 터미네이터)으로 구성된다. 발현 카세트(들)는 통상적으로 벡터 내에 포함되어, 클로닝 및 형질전환을 용이하게 한다. 각각의 숙주에 대해 올바른 조절 서열이 사용되는 한, 박테리아, 효모, 및 진균류를 포함하는 상이한 유기체를 상이한 발현 카세트로 형질전환시킬 수 있다.
- [0078] "형질전환"은, 숙주 유기체 내로 핵산 분자를 전달하여 유전적으로 안정한 유전성을 유발하는 것을 지칭한다. 핵산 분자는, 예를 들어, 자체적으로 복제되는 플라스미드일 수도 있고, 그것이 숙주 유기체의 게놈 내로 통합될 수도 있다. 핵산 단편으로 형질전환된 숙주 유기체는 "재조합" 또는 "형질전환" 유기체 또는 "형질전환체"로 지칭된다. "안정한 형질전환"은 핵산 단편을 핵 및 세포소기관 게놈 둘 모두를 포함하는 숙주 유기체의 게놈 내로 전달하여 유전적으로 안정한 유전성을 야기하는 것을 말한다. 대조적으로, "일시적 형질전환"은, 숙주 유기체의 핵, 또는 DNA-함유 소기관 내로 핵산 단편을 전달하여, 통합 또는 안정한 유전성이 없는 유전자 발현을 유발하는 것을 지칭한다.
- [0079] "코돈 축퇴성(degeneracy)"은 암호화된 폴리펩티드의 아미노산 서열에 영향을 주지 않으면서 뉴클레오티드 서열의 변이를 허용하는 유전 암호의 성질을 지칭한다. 주어진 아미노산을 특정하기 위한 뉴클레오티드 코돈의 사용에 있어서 특정 숙주 세포에서 나타나는 "코돈-편향(codon-bias)"이 당업자에게 주지되어 있다. 그러므로, 숙주 세포 내의 발현을 개선하기 위한 유전자를 합성할 경우에는, 그의 코돈 사용 빈도가 숙주 세포가 선호하는 코돈의 사용 빈도에 근접하도록 유전자를 고안하는 것이 바람직하다.
- [0080] 용어 "기능적으로 동등한 하위단편(subfragment)" 및 "기능적 동등 하위단편"은 본 명세서에서 상호교환가능하게 사용된다. 이들 용어는 분리된 핵산 단편의 일부 또는 하위서열을 지칭하며, 여기에는 단편 또는 하위단편이 활성 효소를 암호화하는지 여부에 무관하게 유전자 발현을 변경하거나 소정의 표현형을 생성하는 능력이 보유된다. 키메라 유전자는, 그것이 활성 효소를 암호화하는지 여부에 무관하게, 프로모터 서열에 대해 센스 또는 안티센스 배향으로, 그의 핵산 단편 또는 하위단편을 연결함으로써 억제에 사용하기 위해 고안될 수 있다.
- [0081] 용어 "보존된 도메인" 또는 "모티프"는 진화적으로 관련된 단백질의 정렬된 서열을 따라 특정 위치에서 보존된 아미노산의 세트를 의미한다. 다른 위치의 아미노산은 상동성 단백질 간에 달라질 수 있는 반면에, 특정 위치에서 고도로 보존된 아미노산은 단백질의 구조, 안정성 또는 활성에 있어서 필수적인 아미노산을 나타낸다.
- [0082] 용어 "실질적으로 유사한" 및 "실질적으로 상응하는"은 본 명세서에서 상호교환가능하게 사용된다. 그들은 하나 이상의 뉴클레오티드 염기의 변화가 유전자 발현을 매개하거나 소정의 표현형을 생성하는 핵산 단편의 능력에 영향을 미치지 않는 핵산 단편을 지칭한다. 이들 용어는 또한, 본 발명의 핵산 단편의 변형, 예를 들어 초기의 비변형된 단편에 비하여 얻어진 핵산 단편의 기능적 특성을 실질적으로 변경하지 않는 하나 이상의 뉴클레오티드의 결실 또는 삽입을 지칭한다. 그러므로, 당업자가 인식할 바와 같이, 본 발명은 특정 예시적 서열을 초과하여 포괄한다는 것이 이해된다. 더욱이, 본 발명이 포괄하는 실질적으로 유사한 핵산 서열이 본 명세서에 예시된 서열, 또는 본 명세서에 개시되고 본 명세서에 개시된 핵산 서열 중 임의의 것과 기능적으로 동등한 뉴클레오티드 서열의 임의의 부분에 혼성화되는 그들의 능력(중등도의 엄격 조건 하에, 예를 들어, 0.5X SSC(표준 소듐 시트레이트), 0.1% SDS(소듐 도데실 설페이트), 60°C)에 의해서도 정의된다는 것을, 당업자는 인식한다.

엄격성 조건을 조정하여, 관계가 먼 유기체로부터의 상동성 서열과 같이 중등으로 유사한 단편 내지 밀접하게 관련된 유기체로부터의 기능성 효소를 중복시키는 유전자와 같이 고도로 유사한 단편을 스크리닝할 수 있다. 혼성화후 세척은 엄격성 조건을 결정한다.

[0083] 용어 "선택적으로 혼성화한다"는, 엄격 혼성화 조건 하에 핵산 서열이 비-표적 핵산 서열에 대한 그의 혼성화보다 검출가능하게 더 큰 정도로(예를 들어, 백그라운드에 비해 적어도 2배) 특정된 핵산 표적 서열에 혼성화됨에 대한 언급, 및 비-표적 핵산의 실질적인 배제에 대한 언급을 포함한다. 선택적으로 혼성화되는 서열은, 뉴클레오티드 서열 중 하나의 상보체가 다른 뉴클레오티드 서열에 대해 전형적으로 적어도 약 80%의 서열 동일성, 또는 90%의 서열 동일성, 100%(즉, 완전히 상보적임) 이하의 서열 동일성을 갖는 2개의 뉴클레오티드 서열이다.

[0084] 용어 "엄격 조건" 또는 "엄격 혼성화 조건"은, 프로브가 그의 표적 서열에 선택적으로 혼성화될 조건에 대한 언급을 포함한다. 전형적으로 프로브는, 검출하고자 하는 핵산 서열에 상보적인 단일 가닥 핵산 서열이다. 프로브는 검출하고자 하는 핵산 서열에 "혼성화가능"하다. 일반적으로, 프로브는 약 1000개 미만의 뉴클레오티드 길이이며, 임의로 500개 미만의 뉴클레오티드 길이이다.

[0085] 혼성화 방법은 잘 정립되어 있다. 전형적으로, 프로브와 샘플은 핵산 혼성화를 허용할 조건 하에서 혼합된다. 이는 적절한 농도 및 온도 조건 하에 무기 또는 유기염의 존재 하에 프로브 및 시료를 접촉시킴을 포함한다. 임의로, 카오토로픽제(chaotropic agent)가 첨가될 수 있다. 핵산 혼성화는 다양한 검정 체제에 적응시킬 수 있다. 가장 적절한 것 중의 하나는 샌드위치 검정(sandwich assay) 체제이다. 샌드위치-형 검정의 주요 구성 요소는 고체 지지체이다. 고체 지지체에는, 서열의 일부에 대해 상보적이고 표지되지 않은 고정화된 핵산 프로브가 흡착되거나 공유적으로 결합된다.

[0086] 엄격 조건은 서열-의존성이며, 상이한 상황에서 상이할 것이다. 혼성화 및/또는 세척 조건의 엄격성을 제어함으로써, 프로브에 100% 상보적인 표적 서열을 확인할 수 있다(상동 프로빙(homologous probing)). 대안적으로, 더 낮은 정도의 유사성이 검출되도록 서열의 일부 불일치를 허용하도록 엄격성 조건을 조정할 수 있다(이종 프로빙(heterologous probing)).

[0087] 전형적으로 엄격 조건은, 염 농도가 pH 7.0 내지 8.3에서 약 1.5 M 미만의 Na 이온, 전형적으로는 약 0.01 내지 1.0 M의 Na 이온 농도(또는 다른 염)이며, 온도가 짧은 프로브(예를 들어, 10 내지 50개 뉴클레오티드)에 있어서는 적어도 약 30℃이고 긴 프로브(예를 들어, 50개 초과 뉴클레오티드)에 있어서는 적어도 약 60℃인 것들일 것이다. 엄격 조건은 또한, 포름아미드와 같은 불안정화제의 첨가로 달성할 수 있다. 예시적인 낮은 엄격성 조건은, 37℃에서 30 내지 35% 포름아미드, 1 M NaCl, 1% SDS(소듐 도데실 설페이트)의 완충 용액을 이용하는 혼성화, 및 50 내지 55℃에서 1X 내지 2X SSC(20X SSC = 3.0 M NaCl/0.3 M 트라이소듐 시트레이트) 중의 세척을 포함한다. 예시적인 중등의 엄격성 조건은, 37℃에서 40 내지 45% 포름아미드, 1 M NaCl, 1% SDS 중의 혼성화, 및 55 내지 60℃에서 0.5X 내지 1X SSC 중의 세척을 포함한다. 예시적인 높은 엄격성 조건은, 37℃에서 50% 포름아미드, 1 M NaCl, 1% SDS 중의 혼성화, 및 60 내지 65℃에서 0.1X SSC 중의 세척을 포함한다.

[0088] 특이성은 전형적으로 혼성화후 세척의 함수이며, 결정적 인자는 최종 세척 용액의 이온 세기 및 온도이다. DNA-DNA 혼성물의 경우, 열 용점(T_m)은 문헌[Meinkoth et al., Anal. Biochem. 138:267-284 (1984)]의 식으로부터 추산될 수 있다: $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\text{GC}\%) - 0.61 (\text{form}\%) - 500/L$ (식 중, M은 1가 양이온의 몰농도이고, GC%는 DNA 중 구아노신 및 시토신 뉴클레오티드의 백분율이고, form%는 혼성화 용액 중 포름아미드의 백분율이고, L은 혼성물의 염기쌍 길이이다). T_m 은 50%의 상보적인 표적 서열이 완전하게 일치하는 프로브에 혼성화되는 온도이다(정의된 이온 세기 및 pH 하에). T_m 은 각 1%의 불일치에 대하여 약 1℃ 감소되고; 이에 따라, T_m , 혼성화 및/또는 세척 조건은 원하는 동일성의 서열에 혼성화하도록 조절될 수 있다. 예를 들어, 90% 초과 동일성을 가진 서열을 탐색한다면, T_m 을 10℃ 감소시킬 수 있다. 일반적으로, 엄격 조건은 정의된 이온 세기 및 pH에서 특정 서열 및 그의 상보체에 대한 T_m 보다 약 5℃ 낮게 선택된다. 그러나, 매우 엄격한 조건은 T_m 보다 1, 2, 3 또는 4℃ 더 낮은 조건에서의 혼성화 및/또는 세척을 이용할 수 있고; 중등으로 엄격한 조건은 T_m 보다 6, 7, 8, 9 또는 10℃ 더 낮은 조건에서의 혼성화 및/또는 세척을 이용할 수 있고; 낮은 엄격성 조건은 T_m 보다 11, 12, 13, 14, 15 또는 20℃ 더 낮은 조건에서의 혼성화 및/또는 세척을 이용할 수 있다. 상기 등식, 혼성화 및 세척 조성물, 및 목적하는 T_m 을 사용하여, 혼성화 및/또는 세척 용액의 엄격성에서의 변화가 본질적으로 기재된다는 것을, 당업자는 이해할 것이다. 목적하는 불일치의 정도가 45℃(수용액)

또는 32℃(포름아미드 용액) 미만의 T_m 을 초래한다면, 더 높은 온도를 사용할 수 있도록 SSC 농도를 증가시키는 것이 바람직하다. 핵산 혼성화에 대한 광범위한 지침은 문헌[Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993)]; 및 문헌[Current Protocols in Molecular Biology, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995)]에서 찾을 수 있다. 혼성화 및/또는 세척 조건은 적어도 10, 30, 60, 90, 120, 또는 240 분 동안 적용될 수 있다.

[0089] 핵산 또는 폴리펩티드 서열과 관련하여 "서열 동일성" 또는 "동일성"은, 특정 비교창에 걸쳐 최대 상응도로 정렬되는 경우에 2개 서열 중의 동일한 핵산 염기 또는 아미노산 잔기를 지칭한다.

[0090] 따라서, "서열 동일성의 백분율"은, 비교창에 걸쳐 최적으로 정렬된 2개의 서열을 비교함으로써 결정된 값을 지칭하며, 여기서 비교창 내의 폴리뉴클레오티드 또는 폴리펩티드 서열의 부분은 2개 서열의 최적 정렬을 위한 기준 서열(부가 또는 결실을 포함하지 않음)과 비교하여 부가 또는 결실(즉, 갭)을 포함할 수 있다. 양자 모두의 서열 내에서 동일한 핵산 염기 또는 아미노산 잔기가 나타나는 위치의 개수를 결정하여 일치하는 위치의 개수를 산출하고, 일치하는 위치의 개수를 비교창 내의 위치의 총 개수로 나누고, 그 결과에 100을 곱하여 서열 동일성의 백분율을 산출함으로써 백분율을 계산한다. 서열 동일성 백분율의 유용한 예는 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 또는 95%, 또는 50% 내지 100%의 임의의 정수 백분율을 포함하나, 이에 한정되지 않는다. 이들 동일성은 본 명세서에 기술된 임의의 프로그램을 사용하여 결정할 수 있다.

[0091] 서열 정렬 및 동일성 또는 유사성 백분율 계산은, 레이저진 바이오인포매틱스 컴퓨팅 스위트(LASERGENE bioinformatics computing suite)(미국 위스콘신주 매디슨 소재의 디엔에이스타 인코포레이티드(DNASTAR Inc.))의 맥얼라인(MegAlign)TM 프로그램을 포함하나 이에 한정되지 않는, 상동성 서열을 검출하도록 고안된 다양한 비교 방법을 사용하여 결정할 수 있다. 본 출원과 관련하여, 서열 분석 소프트웨어를 분석에 이용하는 경우에, 달리 특정하지 않는 한, 분석 결과가 인용된 프로그램의 "디폴트 값"에 기초할 것임을 이해할 것이다. 본 명세서에서 사용된 "디폴트 값"은 최초 초기화할 때에 소프트웨어에 원래 로딩된 임의의 수치 또는 파라미터 세트를 의미할 것이다.

[0092] "클러스탈 V 정렬 방법"은 클러스탈 V(문헌[Higgins and Sharp, CABIOS. 5:151-153 (1989)]; 문헌[Higgins, D.G. et al., Comput. Appl. Biosci. 8:189-191 (1992)]에 기술됨)로 표기되고 레이저진 바이오인포매틱스 컴퓨팅 스위트(미국 위스콘신주 매디슨 소재의 디엔에이스타 인코포레이티드)의 맥얼라인TM 프로그램에서 관찰되는 정렬 방법에 상응한다. 다중 정렬의 경우, 디폴트 값은 갭 페널티(GAP PENALTY)=10 및 갭 길이 페널티(GAP LENGTH PENALTY)=10에 상응한다. 클러스탈 V 방법을 사용하는 단백질 서열의 쌍 정렬 및 동일성 백분율의 계산을 위한 디폴트 파라미터는 케이터플(KTUPLE)=1, 갭 페널티=3, 윈도우(WINDOW)=5 및 다이아고널스 세이브드(DIAGONALS SAVED)=5이다. 핵산의 경우, 이들 파라미터는 케이터플=2, 갭 페널티=5, 윈도우=4 및 다이아고널스 세이브드=4이다. 클러스탈 V 프로그램을 사용하여 서열들을 정렬한 후에, 동일한 프로그램 내의 "서열 거리" 표를 보면 "동일성 백분율"을 얻을 수 있다.

[0093] "클러스탈 W 정렬 방법"은 클러스탈 W(상기 문헌[Higgins and Sharp]; 상기 문헌[Higgins, D.G. et al.]에 기술됨)로 표기되고 레이저진 바이오인포매틱스 컴퓨팅 스위트(미국 위스콘신주 매디슨 소재의 디엔에이스타 인코포레이티드)의 맥얼라인TM v6.1 프로그램에서 관찰되는 정렬 방법에 상응한다. 다중 정렬을 위한 디폴트 파라미터는 갭 페널티=10, 갭 길이 페널티=0.2, 지연 이탈 서열(Delay Divergen Seqs)(%)=30, DNA 전이 중량(DNA Transition Weight)=0.5, 단백질 중량 매트릭스(Protein Weight Matrix)=고넷 시리즈(Gonnet Series), DNA 중량 매트릭스(DNA Weight Matrix)=IUB에 상응한다. 클러스탈 W 프로그램을 사용하여 서열들을 정렬한 후에, 동일한 프로그램 내의 "서열 거리" 표를 보면 "동일성 백분율"을 얻을 수 있다.

[0094] "BLASTN 정렬 방법"은 국립 생물공학 정보센터(NCBI: National Center for Biotechnology Information)에 의해 제공되는, 디폴트 파라미터를 사용하여 뉴클레오티드 서열을 비교하는 알고리즘이다.

[0095] 동일하거나 유사한 기능 또는 활성을 가진 폴리펩티드를 다른 종으로부터 확인함에 있어서 여러 수준의 서열 동일성이 유용하다는 것이 당업자에게 주지되어 있다. 동일성 백분율의 유용한 예는 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 또는 95%, 또는 50% 내지 100%의 임의의 정수 백분율을 포함하나, 이에 한정되지 않는다. 실제로, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 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%와 같은 50% 내지 100%의 임의의 정수의 아미노산 동일

성이 본 발명을 설명하는데 유용할 수 있다. 또한, 이러한 분리된 뉴클레오티드 단편의 임의의 전장 또는 부분 상보체도 관심대상이다.

[0096] 따라서, 본 발명은 본 명세서에 개시된 특정 예시적 뉴클레오티드 서열을 초과하여 포괄한다. 예를 들어, 유전 암호의 축퇴성을 반영하는 유전자 서열의 변경이 고려된다. 또한, 주어진 위치에 화학적으로 동등한 아미노산의 생성을 유발하면서 암호화된 단백질의 기능적 특성에 영향을 미치지 않는 유전자의 변경이 흔하다는 것은 당업계에 주지되어 있다. 본 명세서의 논의에 있어서 치환은 하기의 5개 그룹 중 하나 내에서의 교환으로 정의된다:

[0097] 1. 비극성 또는 약간 극성인 작은 지방족 잔기: Ala, Ser, Thr(Pro, Gly);

[0098] 2. 음으로 하전된 극성 잔기 및 그들의 아미드: Asp, Asn, Glu, Gln;

[0099] 3. 양으로 하전된 극성 잔기: His, Arg, Lys;

[0100] 4. 비극성인 큰 지방족 잔기: Met, Leu, Ile, Val(Cys); 및

[0101] 5. 큰 방향족 잔기: Phe, Tyr, Trp.

[0102] 따라서, 소수성 아미노산인 아미노산 알라닌에 대한 코돈은 소수성이 더 낮은 다른 잔기(예를 들어, 글리신) 또는 소수성이 더 높은 잔기(예를 들어, 발린, 류신 또는 아이소류신)를 암호화하는 코돈으로 치환될 수 있다. 유사하게, 음으로 하전된 하나의 잔기로 다른 하나를 치환하거나(예를 들어, 글루탐산을 아스파르트산으로) 양으로 하전된 하나의 잔기로 다른 하나를 치환하는(예를 들어, 아르기닌을 라이신으로) 변화도 기능적으로 동등한 생성물을 생성시키는 것으로 예상될 수 있다. 많은 경우에, 단백질 분자의 N-말단 및 C-말단 부분의 변경을 유발하는 뉴클레오티드 변화도 단백질의 활성을 변경하지 않을 것으로 예상된다.

[0103] 각각의 제안된 변형은 당업계의 통상적 기술로 충분히 가능하며, 암호화된 생성물의 생물학적 활성의 보존을 결정하는 것도 그러하다. 또한, 본 발명에 의해 포괄되는 실질적으로 유사한 서열이 상기 정의된 바와 같은 엄격 조건 하에 혼성화되는 그들의 능력에 의해서도 정의된다는 것을, 당업자는 인식한다.

[0104] 본 발명의 실질적으로 유사한 바람직한 핵산 단편은, 뉴클레오티드 서열이 본 명세서에 보고된 핵산 단편의 뉴클레오티드 서열과 적어도 70% 동일한 핵산 단편이다. 더욱 바람직한 핵산 단편은 본 명세서에 보고된 핵산 단편의 뉴클레오티드 서열과 적어도 90% 동일하다. 본 명세서에 보고된 핵산 단편의 뉴클레오티드 서열과 적어도 95% 동일한 핵산 단편이 가장 바람직하다.

[0105] 아미노산 또는 뉴클레오티드 서열의 "상당한 부분"은, 당업자에 의한 서열의 수동 평가에 의해, 또는 BLAST(기본 국소 정렬 검색 도구(Basic Local Alignment Search Tool); 문헌[Altschul, S. F., et al., J. Mol. Biol., 215:403-410(1993)])와 같은 알고리즘을 사용하는 컴퓨터-자동화 서열 비교 및 확인에 의해, 폴리펩티드 또는 유전자를 추정적으로 확인하기에 충분한 폴리펩티드의 아미노산 서열 또는 유전자의 뉴클레오티드 서열을 포함하는 부분이다. 일반적으로, 폴리펩티드 또는 핵산 서열을 공지의 단백질 또는 유전자와 상동인 것으로 추정적으로 확인하기 위해서는 10개 이상 연속되는 아미노산 또는 30개 이상의 뉴클레오티드의 서열이 필요하다. 더욱이, 뉴클레오티드 서열에 있어서는, 서열-의존적인 유전자 확인 방법(예를 들어, 서던 혼성화(Southern hybridization)) 및 분리 방법(예를 들어, 박테리아 콜로니 또는 박테리오파지 플라크의 체자리 혼성화(in situ hybridization))에 20 내지 30개의 연속된 뉴클레오티드를 포함하는 유전자 특이적 올리고뉴클레오티드 프로브가 사용될 수 있다. 또한, 12 내지 15개 염기의 짧은 올리고뉴클레오티드를 PCR에서 증폭 프라이머로 사용하여, 프라이머를 포함하는 특정 핵산 단편을 얻을 수 있다. 따라서, 뉴클레오티드 서열의 "상당한 부분"은, 서열을 포함하는 핵산 단편을 특이적으로 확인 및/또는 분리하기에 충분한 서열을 포함한다. 본 명세서에서는, 특정 단백질을 암호화하는 완전한 아미노산 및 뉴클레오티드 서열을 교시한다. 본 명세서에 보고된 서열의 이익을 얻어, 당업자는 이제 개시된 서열의 전체 또는 실질적인 부분을 당업계에 공지된 목적으로 사용할 수 있게 되었다.

[0106] 용어 "상보성"은, 역평행 배향으로 정렬될 때 왓슨-크릭(Watson-Crick) 염기-대합이 가능한 2개 서열의 뉴클레오티드 염기 간의 관계를 기술한다. 예를 들어, DNA에 있어서, 아데노신은 티민과 염기-대합이 가능하고 시토신은 구아닌과 염기-대합이 가능하다. 따라서 본 발명은, 첨부된 서열 목록 및 명세서에 보고된 완전한 서열뿐 아니라 실질적으로 유사한 핵산 서열에 상보적인 분리된 핵산 분자를 이용할 수 있다.

[0107] 용어 "분리된"은, 폴리펩티드 또는 뉴클레오티드 서열과 천연적으로 회합되어 있는 적어도 하나의 구성요소로부

터 분리된 폴리펩티드 또는 뉴클레오티드 서열을 지칭한다.

- [0108] "프로모터"는 암호화 서열 또는 기능성 RNA의 발현을 제어할 수 있는 DNA 서열을 지칭한다. 프로모터 서열은 근위 업스트림 요소 및 보다 원위의 업스트림 요소로 이루어지며, 후자의 요소는 흔히 인헨서로 지칭된다. 따라서, "인헨서"는 프로모터 활성을 자극할 수 있는 DNA 서열이며, 프로모터의 고유 요소이거나 프로모터의 조직-특이성 또는 수준을 향상시키도록 삽입된 이중 요소일 수 있다. 프로모터는 고유 유전자로부터 그 전체가 유래될 수도 있거나, 천연에서 발견되는 상이한 프로모터로부터 유래된 상이한 요소로 구성될 수도 있거나, 심지어 합성 DNA 세그먼트를 포함할 수도 있다. 상이한 프로모터가 상이한 조직 또는 세포 유형에서, 또는 상이한 발생 단계에서, 또는 상이한 환경적 조건에 반응하여, 유전자의 발현을 유도할 수 있음을 당업자는 이해할 것이다. 대부분의 경우에 조절 서열의 정확한 경계는 완벽하게 정의되지 않으므로, 일부 변이된 DNA 단편이 동일한 프로모터 활성을 가질 수 있음이 추가로 인정된다. 대부분의 세포 유형에서 대부분의 경우에 유전자가 발현되도록 하는 프로모터는 통상적으로 "구성성 프로모터"라고 지칭된다.
- [0109] "3' 비암호화 서열", "전사 터미네이터" 및 "종결 서열"은 본 명세서에서 상호교환가능하게 사용되며, 암호화 서열의 다운스트림에 위치하는 DNA 서열, 예를 들어 폴리아데닐화 인식 서열 및 mRNA 프로세싱 또는 유전자 발현에 영향을 미칠 수 있는 조절 신호를 암호화하는 다른 서열을 지칭한다. 폴리아데닐화 신호는 통상적으로 mRNA 전구체의 3' 말단에 폴리아데닐산 트랙트(tract)를 부가하는 데에 영향을 미치는 것을 특징으로 한다.
- [0110] 용어 "작동가능하게 연결된"은, 하나의 기능이 다른 하나에 의해 영향을 받도록 단일 핵산 단편상에서 핵산 서열들이 연계됨을 지칭한다. 예를 들어, 프로모터가 암호화 서열의 발현에 영향을 미칠 수 있는 경우에(즉, 암호화 서열은 프로모터의 전사 제어 하에 있음), 프로모터는 그 암호화 서열에 작동가능하게 연결된다. 암호화 서열은 센스 또는 안티센스 배향으로 조절 서열에 작동가능하게 연결될 수 있다. 다른 예에서, 본 발명의 상보적 RNA 영역은 직접 또는 간접적으로 표적 mRNA에 대해 5', 또는 표적 mRNA에 대해 3', 또는 표적 mRNA 내부에 작동가능하게 연결될 수 있거나, 제1 상보적 영역은 표적 mRNA에 대해 5'이고 그의 상보체는 3'이다.
- [0111] 본 명세서에 사용되는 표준 재조합 DNA 및 분자 클로닝 기술은 당업계에 주지되어 있으며, 문헌[Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989)]에 충분히 기술되어 있다. 형질전환 방법은 당업자에게 주지되어 있으며, 하기에 기술된다.
- [0112] "PCR" 또는 "중합효소 연쇄 반응"은 다량의 특정 DNA 세그먼트의 합성을 위한 기술이며 일련의 반복 사이클로 구성된다(미국 코네티컷주 노르워크 소재의 퍼킨 엘머 세터스 인스트루먼트(Perkin Elmer Cetus Instruments)). 전형적으로, 이중-가닥 DNA를 열변성시키고, 표적 세그먼트의 3' 경계에 상보적인 2개의 프라이머를 저온에서 어닐링시킨 후, 중간 온도에서 연장한다. 이들 3개의 연이은 단계의 1 세트를 "사이클"이라고 지칭한다.
- [0113] "플라스미드" 또는 "벡터"는, 흔히 세포의 중추 대사작용의 일부가 아닌 유전자를 지니는, 통상적으로 원형 이중-가닥 DNA 단편 형태의 염색체의 요소이다. 이러한 요소는, 세포 내로 발현 카세트(들)를 도입할 수 있는 독특한 구축물 내로 많은 뉴클레오티드 서열이 결합되거나 재조합된, 임의의 공급원으로부터 유래된 단일- 또는 이중-가닥 DNA 또는 RNA의 선형 또는 원형의 자체적으로 복제되는 서열, 게놈 통합 서열, 파지 또는 뉴클레오티드 서열일 수 있다.
- [0114] 용어 "유전자 변경된"은 유전공학, 형질전환 및/또는 돌연변이에 의해 유전 물질을 변화시키는 과정을 지칭한다.
- [0115] 용어 "재조합"은, 예를 들어 유전자 조작 기술에 의한 분리된 핵산 세그먼트의 조작에 의하거나 화학적 합성에 의해, 분리되어 있는 2개의 다른 서열 세그먼트의 인공 조합을 지칭한다. "재조합"은 또한, 이중 핵산의 도입에 의해 변형된 세포 또는 벡터, 또는 그렇게 변형된 세포로부터 유래하는 세포에 대한 언급을 포함하나, 의도적인 인위적 개입 없이 발생하는 것들과 같이 천연적으로 발생하는 사건(예를 들어, 자발적 돌연변이, 천연적 형질전환, 천연적 형질도입, 천연적 전위)에 의한 세포 또는 벡터의 변경은 포괄하지 않는다.
- [0116] 용어 "재조합 구축물", "발현 구축물", "키메라 구축물", "구축물" 및 "재조합 DNA 구축물"은 본 명세서에서 상호교환가능하게 사용된다. 재조합 구축물은 핵산 단편, 예를 들어, 천연에서는 함께 발견되지 않는 조절 및 암호화 서열의 인공 조합을 포함한다. 예를 들어, 재조합 구축물은 상이한 공급원으로부터 유래된 조절 서열 및 암호화 서열, 또는 동일한 공급원으로부터 유래되지만, 천연에서 발견되는 것과는 상이한 방식으로 배열된 조절 서열 및 암호화 서열을 포함할 수 있다. 이러한 구축물은 단독으로 사용될 수도 있고, 벡터와 함께 사용될 수도 있다. 벡터를 사용한다면, 당업자에게 주지된 바와 같이 벡터의 선택은 숙주 세포를 형질전환시키기 위해

사용될 방법에 따라 달라진다. 예를 들어, 플라스미드 벡터를 사용할 수 있다. 본 발명의 임의의 분리된 핵산 단편을 포함하는 숙주 세포를 성공적으로 형질전환시키고 선택하고 증식시키기 위하여 벡터 상에 존재해야 하는 유전 요소는 당업자에게 주지되어 있다. 또한, 숙주자는 상이한 독립적인 형질전환 사건이 발현 수준 및 패턴을 상이하게 하는 결과를 일으킬 수 있음을 인지할 것이며(문헌[Jones et al., EMBO J. 4:2411-2418 (1985)]; 문헌[De Almeida et al., Mol. Gen. Genetics 218:78-86 (1989)]), 이에 따라 목적하는 발현 수준 및 패턴을 나타내는 라인을 얻기 위해서는 그러한 다중의 사건을 스크리닝할 필요가 있다는 것을, 당업자는 또한 인식할 것이다. 이러한 스크리닝은 특히 DNA의 서던 분석, mRNA 발현의 노던 분석, 단백질 발현의 면역 블로팅 분석, 또는 표현형 분석에 의해 이루어질 수 있다.

[0117] 본 명세서에 사용되는 바와 같이, 용어 "발현"은 기능성 최종 생성물(예를 들어, mRNA 또는 단백질[전구체 또는 성숙체])의 생성을 지칭한다.

[0118] 용어 "도입된"은 핵산(예를 들어, 발현 구축물) 또는 단백질을 세포 내에 제공함을 의미한다. "도입된"은, 진핵 또는 원핵 세포 내로의 핵산의 혼입(여기서 핵산은 세포의 게놈 내로 혼입될 수 있음)에 대한 언급을 포함하며, 핵산 또는 단백질을 세포에 일시적으로 제공하는 것에 대한 언급을 포함한다. "도입된"은 유성 교배(sexually crossing)와 더불어 안정하거나 일시적인 형질전환 방법에 대한 언급을 포함한다. 따라서, 핵산 단편(예를 들어, 재조합 구축물/발현 구축물)을 세포 내로 삽입하는 것과 관련하여, "도입된"은 "트랜스펙션(transfection)" 또는 "형질전환" 또는 "형질도입"을 의미하며, 진핵 또는 원핵 세포 내로의 핵산 단편의 혼입(여기서 핵산 단편은 세포의 게놈(예를 들어, 염색체, 플라스미드, 플라스티드 또는 미토콘드리아 DNA) 내로 혼입되거나, 자체적 레플리콘(autonomous replicon)으로 전환되거나, 일시적으로 발현될 수 있음(예를 들어, 트랜스펙션된 mRNA))에 대한 언급을 포함한다.

[0119] 용어 "상동성"은, 유사한 촉매 기능을 가진, 공통의 진화적 기원의 단백질 또는 폴리펩티드를 지칭한다. 본 발명은 재조합 기술을 통해 상동성 단백질을 생성하는 박테리아를 포함할 수 있다.

[0120] 다양한 공급원 유래의 신규한 수크로스 수송체 유전자를 사용하여, 수크로스를 사용하도록 조작된 재조합 박테리아가 본 명세서에 개시된다. 구체적으로, 본 명세서에서 개시된 재조합 박테리아는 수크로스 수송체 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열로서, 상기 폴리펩티드가 클러스탈 W 정렬 방법에 기초하여, 서열 번호 68, 서열 번호 70, 서열 번호 72, 서열 번호 74, 서열 번호 76, 서열 번호 78, 서열 번호 80, 서열 번호 82, 서열 번호 84, 서열 번호 86, 서열 번호 88, 서열 번호 90, 서열 번호 92 및 서열 번호 94로 이루어진 군으로부터 선택되는 아미노산 서열과 비교시, 적어도 95% 서열 동일성을 갖는 뉴클레오티드 서열; 및 수크로스 가수분해효소 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그의 게놈 내에 또는 적어도 하나의 재조합 구축물 상에 포함한다. 상기에 기재된 수크로스 수송체 폴리펩티드를 암호화하는 뉴클레오티드 서열, 및 수크로스 가수분해효소 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 포함하는 재조합 박테리아는 당업계에 공지된 방법을 이용하여 게놈 내에 또는 적어도 하나의 재조합 구축물 상에, 적절한 숙주 박테리아 내로 상기 뉴클레오티드 서열을 도입함으로써 구축될 수 있으며, 이는 하기에 기재된 바와 같다. 일부 실시형태에서, 재조합 박테리아는 수크로스를 대사하여 글리세롤 및/또는 글리세롤 유도 생성물을 생성할 수 있다.

[0121] 본 명세서에서 개시된 재조합 박테리아의 구축에서의 이용에 적절한 숙주 박테리아는 하기 속의 유기체를 포함하지만, 이에 한정되지 않는다: 에스케리키아, 스트렙토코커스, 아그로박테리움, 바실러스, 코리네박테리움, 락토바실러스, 클로스트리디움, 글루코노박터(*Gluconobacter*), 시트로박터, 엔테로박터, 클레브시엘라, 아에로박터(*Aerobacter*), 메틸로박터(*Methylobacter*), 살모넬라, 스트렙토마이세스 및 슈도모나스.

[0122] 일부 실시형태에서, 숙주 박테리아는 하기의 속으로부터 선택된다: 에스케리키아, 클레브시엘라, 시트로박터 및 아에로박터.

[0123] 일부 실시형태에서, 숙주 박테리아는 에스케리키아 콜라이이다. 일부 실시형태에서, 숙주 박테리아는 PTS 마이너스이다. 이들 실시형태에서, 숙주 박테리아는 그의 고유 상태에서 PTS 마이너스이거나, 하기와 같이 PTS 유전자의 불활성화를 통해 PTS 마이너스가 되게 할 수 있다.

[0124] 생성 미생물에서는, 당의 수송과 수송되는 당의 인산화를 위한 포스포에놀피루베이트(PEP)의 사용을 연결하지 않는 것이 간혹 바람직하다. 용어 "하향 조절된"은 야생형 단백질(들)의 활성에 비해, 활성 단백질(들)의 활성에서의 감소 또는 제거를 말한다. PTS는, 이러한 유형의 수송에서 요구되는 단백질을 암호화하는 하나 이상의 내인성 유전자의 발현을 하향 조절시킴으로써 비활성화될 수 있다("PTS 마이너스" 유기체를 야기함). 하향 조절은 하나 이상의 이들 유전자가 "파괴(disruption)"되는 경우 전형적으로 일어나며, 이는 유전자의 일부 내에

서 삽입, 결실 또는 표적화된 돌연변이를 지칭하며, 이는 해당 유전자가 게놈에서 제거되고, 단백질이 번역되지 않도록 완전한 유전자 낙아웃(knockout)을 초래하거나 또는 단백질이 삽입, 결실, 아미노산 치환 또는 다른 표적화된 돌연변이를 갖도록 단백질이 번역되도록 한다. 단백질 내의 파괴의 위치는, 예를 들어 단백질의 N-말단 부분 내에, 또는 단백질의 C-말단 부분 내에 존재할 수 있다. 파괴된 단백질은 파괴되지 않은 단백질에 비해 손상된 활성을 가질 것이며, 비-기능성일 수 있다. 단백질 발현을 낮추거나 결여되게 하는 하향 조절은 또한, 조절 서열, 전사 및, 번역 인자 및/또는 신호 전달 경로의 조작을 통해, 또는 센스, 안티센스 또는 RNAi 기술, 또는 당업자에게 공지되어 있는 유사 메커니즘의 사용에 의해 유발할 수 있다.

[0125] 수크로스 수송체 폴리펩티드는 미생물 세포 내로의 수크로스의 수송을 매개할 수 있는 폴리펩티드이다. 문헌 [Jahreis et al. (J. Bacteriol. 184:5307-5316, 2002)]에 기재된 바와 같이, 수크로스 수송체 폴리펩티드, 예를 들어, cscB 유전자(서열 번호 25에 기재된 암호화 서열)에 의해 암호화된 에스케리키아 콜라이 ATCC®13281 유래의 CscB 폴리펩티드(서열 번호 26에 기재됨); 및 cscB 유전자(서열 번호 23에 기재된 암호화 서열)에 의해 암호화된 에스케리키아 콜라이 야생형 균주 EC3132 유래의 CscB 폴리펩티드(서열 번호 24에 기재됨)가 해당 분야에 공지되어 있다. 본 명세서에 개시된 수크로스 수송체 폴리펩티드는 표 1에 나타난 바와 같이 다양한 공급 원으로부터 분리되며, 서열 번호 68(서열 번호 67에 기재된 암호화 서열), 서열 번호 70(서열 번호 69에 기재된 암호화 서열), 서열 번호 72(서열 번호 71에 기재된 암호화 서열), 서열 번호 74(서열 번호 73에 기재된 암호화 서열), 서열 번호 76(서열 번호 75에 기재된 암호화 서열), 서열 번호 78(서열 번호 77에 기재된 암호화 서열), 서열 번호 80(서열 번호 79에 기재된 암호화 서열), 서열 번호 82(서열 번호 81에 기재된 암호화 서열), 서열 번호 84(서열 번호 83에 기재된 암호화 서열), 서열 번호 86(서열 번호 85에 기재된 암호화 서열), 서열 번호 88(서열 번호 87에 기재된 암호화 서열), 서열 번호 90(서열 번호 89에 기재된 암호화 서열), 서열 번호 92(서열 번호 91에 기재된 암호화 서열) 및 서열 번호 94(서열 번호 93에 기재된 암호화 서열)에 기재된 아미노산 서열을 갖는다. 그러나, 이들 폴리펩티드에 대해 입증된 활성이 이전에 공지되어 있지 않다.

[0126] 일부 실시형태에서, 수크로스 수송체 활성을 갖는 폴리펩티드는 클러스탈 W 정렬 방법에 기초하여, 서열 번호 68, 서열 번호 70, 서열 번호 72, 서열 번호 74, 서열 번호 76, 서열 번호 78, 서열 번호 80, 서열 번호 82, 서열 번호 84, 서열 번호 86, 서열 번호 88, 서열 번호 90, 서열 번호 92 또는 서열 번호 94에 기재된 바와 같은 아미노산 서열과 비교시, 적어도 95% 서열 동일성을 갖는다.

[0127] 일부 실시형태에서, 수크로스 수송체 활성을 갖는 폴리펩티드는 서열 번호 68, 서열 번호 70, 서열 번호 72, 서열 번호 74, 서열 번호 76, 서열 번호 78, 서열 번호 80, 서열 번호 82, 서열 번호 84, 서열 번호 86, 서열 번호 88, 서열 번호 90, 서열 번호 92 또는 서열 번호 94에 기재된 아미노산 서열을 갖는다.

[0128] 일부 실시형태에서, 수크로스 수송체 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열은 BLASTN 정렬 방법에 기초하여, 서열 번호 67, 서열 번호 69, 서열 번호 71, 서열 번호 73, 서열 번호 75, 서열 번호 77, 서열 번호 79, 서열 번호 81, 서열 번호 83, 서열 번호 85, 서열 번호 87, 서열 번호 89, 서열 번호 91 또는 서열 번호 93에 기재된 바와 같은 뉴클레오티드 서열과 비교시, 적어도 95% 서열 동일성을 갖는다.

[0129] 또한, 본 명세서에 개시된 재조합 박테리아는 수크로스 가수분해효소 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그들의 게놈 내에 또는 적어도 하나의 재조합 구축물 상에 포함한다. 수크로스 가수분해효소 활성을 갖는 폴리펩티드는, 수크로스를 가수분해하여 프룩토스 및 글루코스를 생성하는 촉매 작용을 하는 능력을 갖는다. 수크로스 가수분해효소 활성을 갖는 폴리펩티드는 공지되어 있으며, cscA 유전자(서열 번호 27에 기재된 암호화 서열)에 의해 암호화되는, 에스케리키아 콜라이 야생형 균주 EC3132 유래의 CscA(서열 번호 28에 기재됨), cscA 유전자(서열 번호 29에 기재된 암호화 서열)에 의해 암호화되는 에스케리키아 콜라이 ATCC®13281 유래의 CscA(서열 번호 30에 기재됨); bfrA 유전자(서열 번호 31에 기재된 암호화 서열)에 의해 암호화되는, 비피도박테리움 락티스 균주 DSM 10140^T 유래의 BfrA(서열 번호 32에 기재됨); SUC2 유전자(서열 번호 33에 기재된 암호화 서열)에 의해 암호화되는, 사카로마이세스 세레비시아 유래의 Suc2p(서열 번호 34에 기재됨); scrB 유전자(서열 번호 35에 기재된 암호화 서열)에 의해 암호화되는, 코리네박테리움 글루타미쿰 유래의 ScrB(서열 번호 36에 기재됨); 서열 번호 37에 기재된 암호화 유전자의 암호화 서열, 류코노스톡 메센테로이데스 DSM 20193 유래의 수크로스 포스포릴라제(서열 번호 38에 기재됨); 및 sucP 유전자(서열 번호 39에 기재된 암호화 서열)에 의해 암호화되는, 비피도박테리움 아돌레센티스 DSM 20083 유래의 수크로스 포스포릴라제(서열 번호 40에 기재됨)를 포함하지만, 이들에 한정되지 않는다.

[0130] 일부 실시형태에서, 수크로스 가수분해효소 활성을 갖는 폴리펩티드는 EC 3.2.1.26 또는 EC 2.4.1.7로 분류된다.

[0131] 일부 실시형태에서, 수크로스 가수분해효소 활성을 갖는 폴리펩티드는, 클러스탈 W 정렬 방법에 기초하여, 서열 번호 28, 서열 번호 30, 서열 번호 32, 서열 번호 34, 서열 번호 36, 서열 번호 38 또는 서열 번호 40에 기재된 아미노산 서열에 대해 적어도 95%의 서열 동일성을 갖는다.

[0132] 일부 실시형태에서, 수크로스 가수분해효소 활성을 갖는 폴리펩티드는 서열 번호 30에 기재된 아미노산 서열에 실질적으로 상응한다.

[0133] 본 명세서에 개시된 재조합 박테리아는 그 박테리아가 수크로스 가수분해에 의하여 생성된 프룩토스를 이용할 수 있게 하는 프룩토키나제 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 그들의 게놈 내에 또는 적어도 하나의 재조합 구축물 상에 추가로 포함할 수 있다. 프룩토키나제 활성을 갖는 폴리펩티드는 프룩토키나제(EC 2.7.1.4로 표기) 및 프룩토스 인산화 활성을 갖는 각종 헥소스 키나제(EC 2.7.1.3 및 EC 2.7.1.1)를 포함한다. 프룩토스 인산화 활성은 헥소키나제 및 케토헥소키나제에 의해 나타날 수 있다. 본 명세서에 개시된 재조합 박테리아를 구축하기 위해 사용될 수 있는, 다양한 미생물 유래의 폴리펩티드를 암호화하는 대표적인 유전자들은 표 2에 열거되어 있다. 프룩토스를 인산화할 수 있는 단백질과 실질적으로 유사한 단백질(예를 들어, 표 2에 열거된 유전자에 의해 암호화되는 것)도 또한 사용할 수 있다는 것을, 당업자는 알 것이다.

표 2

프룩토키나제 활성을 가진 효소를 암호화하는 서열

공급처	유전자 명칭	EC 번호	뉴클레오티드 단백질 서열 번호	서열 번호
아그로박테리움 투메파시엔스	<i>scrK</i> (프룩토키나제)	2.7.1.4	41	42
			43	44
스트렙토코커스 뮤탄스 에스케리키아 콜라이	<i>scrK</i> (프룩토키나제)	2.7.1.4	45	46
			47	48
클레브시엘라 뉴모니아 에스케리키아 콜라이	<i>scrK</i> (프룩토키나제)	2.7.1.4	49	50
			51	52
엔테로코커스 파에칼리스 사카로마이세스 세레비시아	<i>cscK</i> (프룩토키나제)	2.7.1.4	53	54
			55	56
사카로마이세스 세레비시아	<i>HXK1</i> (헥소키나제)	2.7.1.1		
	<i>HXK2</i> (헥소키나제)	2.7.1.1		

[0134]

[0135] 일부 실시형태에서, 프룩토키나제 활성을 갖는 폴리펩티드는 EC 2.7.1.4, EC 2.7.1.3 또는 EC 2.7.1.1로 분류된다.

[0136] 일부 실시형태에서, 프룩토키나제 활성을 갖는 폴리펩티드는, 클러스탈 W 정렬 방법에 기초하여, 서열 번호 42, 서열 번호 44, 서열 번호 46, 서열 번호 48, 서열 번호 50, 서열 번호 52, 서열 번호 54 또는 서열 번호 56에 기재된 아미노산 서열에 대하여 적어도 95%의 서열 동일성을 갖는다.

[0137] 일부 실시형태에서, 프룩토키나제 활성을 갖는 폴리펩티드는 서열 번호 50에 기재된 아미노산 서열을 갖는다.

[0138] 수크로스 수송체 활성을 갖는 폴리펩티드 및 수크로스 가수분해효소 활성을 갖는 폴리펩티드를 암호화하는 유전자의 암호화 서열은 동일하거나 상이한 미생물 종으로부터의 상동성 폴리펩티드를 암호화하는 뉴클레오티드 서열을 분리하는데 사용될 수 있다. 예를 들어, BLASTN과 같은 서열 분석 소프트웨어를 사용하여 공개 입수가능한 핵산 서열 데이터베이스를 검색함으로써, 상기 유전자들의 상동체들을 확인할 수 있다. 부가적으로, 서열-의존성 프로토콜을 사용하는 상동성 유전자의 분리는 당업계에 공지되어 있다. 서열-의존성 프로토콜의 예들로는, 이에 제한되지는 않지만, 핵산 혼성화 방법, DNA 및 RNA 증폭 방법이 포함되며, 이는 핵산 증폭 기술(예를 들어, 중합효소 연쇄반응(PCR)(몰리스(Mullis) 등의 미국 특허 제4,683,202호); 리가제 연쇄 반응(LCR)(문헌 [Tabor, S. et al., Proc., Acad. Sci. USA 82, 1074, 1985]); 또는 가닥 대체 증폭(SDA: strand displacement amplification)(문헌[Walker, et al., Proc. Natl. Acad. Sci. U.S.A., 89: 392, (1992)])의 다양한 사용에 의해 예시된다. 예를 들어, 상기 기재된 폴리펩티드를 암호화하는 뉴클레오티드 서열을 상동체의 확인을 위한 혼성화 프로브로 채용할 수 있다.

[0139] 다른 공급원으로부터 분리된 이들 폴리펩티드를 암호화하는 유전자 또한 본 명세서에 개시된 재조합 박테리아에

사용할 수 있다는 것을, 당업자는 인식할 것이다. 부가적으로, 코돈 축퇴성으로 인해 암호화된 폴리펩티드의 아미노산 서열에 영향을 미치지 않으면서 폴리펩티드를 암호화하는 뉴클레오티드 서열에 변이를 만들 수 있으며, 실질적으로 유사한 단백질을 생성하는 그 아미노산 치환, 결실 또는 부가는 암호화된 단백질에 포함될 수 있다.

[0140] 수크로스 수송체 활성을 갖는 폴리펩티드 및 수크로스 가수분해효소 활성을 갖는 폴리펩티드를 암호화하는 뉴클레오티드 서열은 원하는 서열에 결합되도록 설계된 프라이머를 이용하는 PCR(예를 들어 미국 특허 제4,683,202호 참조)을 이용하여, 분리될 수 있다. 축퇴 프라이머(degenerate primer) 또는 이중 프로브 혼성화의 사용에 의한 것과 같은 다른 유전자 분리 방법이 당업자에게 주지되어 있다. 뉴클레오티드 서열은 또한, 화학적으로 합성하거나 DNA2.0 인코포레이티드(DNA2.0 Inc.)(미국 캘리포니아주 멘로 파크 소재)와 같은 판매원로부터 구매할 수 있다. 뉴클레오티드 서열은 원하는 숙주 세포에서의 발현을 위해 코돈 최적화될 수 있다.

[0141] 폴리펩티드들의 발현은 당업자에게 공지된 많은 방법들 중 하나를 이용하여 초래될 수 있다. 예를 들어, 상기에 기재된 폴리펩티드들을 암호화하는 뉴클레오티드 서열들은 하나 이상의 다중카피 플라스미드 상에서, 또는 암호화 서열들의 하나 이상의 카피를 숙주 게놈 내에 통합시킴으로써 박테리아 내로 도입될 수 있다. 상기 폴리펩티드들을 암호화하는 뉴클레오티드 서열들은 숙주 박테리아 내에 별도로(예를 들어, 별도의 플라스미드들 상에서) 또는 임의의 조합으로(예를 들어, 단일 플라스미드 상에서) 도입될 수 있다.

[0142] 플라스미드(들) 상에 또는 게놈 내에 있는 도입된 암호화 영역은 적어도 하나의 고도로 활성인 프로모터로부터 발현될 수 있다. 통합된 암호화 영역은 그 자신의 프로모터를 갖는 키메라 유전자의 일부로서 도입될 수도 있고, 게놈에 내인성인 고도로 활성인 프로모터에 인접하여, 또는 고도로 발현되는 오페론 내에, 그것이 통합될 수도 있다. 적절한 프로모터는, CYC1, HIS3, GAL1, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO 및 lac, ara, tet, trp, lP_L , lP_R , T7, tac 및 trc(에스케리키아 콜라이에서의 발현에 유용함)와, amy, apr, npr 프로모터 및 다양한 파지 프로모터(바실러스에서의 발현에 유용함)를 포함하지만, 이들에 한정되지 않는다. 프로모터는, 파이네(Payne) 등(미국 특허 제7,132,527호)에 의해 기재된, 스트렙토마이세스 리비단스 글루코스 아이소머라제 프로모터 또는 그의 변이체일 수도 있다.

[0143] 일부 실시형태에서, 본 명세서에 개시된 재조합 박테리아는 글리세롤을 생성할 수 있다. 탄수화물 또는 당을 사용하여 글리세롤을 제조하기 위한 생물학적 방법은 효모 및 일부 박테리아, 다른 진균류, 및 조류에서 공지되어 있다. 박테리아 및 효모 양자 모두는, 해당 작용의 프룩토스-1,6-비스포스페이트 경로를 통해 글루코스 또는 다른 탄수화물을 전환함으로써 글리세롤을 생성한다. 본 명세서에 개시된 글리세롤 생성 방법에서는, 천연적으로 글리세롤을 생성하는 숙주 박테리아를 사용할 수 있다. 또한, 글리세롤 및 글리세롤 유도체의 생성을 위해 박테리아를 조작할 수 있다. 미국 특허 제7,005,291호에 기술된 바와 같이, 다양한 기질로부터 글리세롤을 생성하는 능력은 효소 활성 글리세롤-3-포스페이트 탈수소효소(G3PDH) 및/또는 글리세롤-3-포스파타제의 발현을 통해 제공될 수 있다. 숙주 박테리아에서 효소 활성을 발현시키기 위해 사용될 수 있는 이들 단백질을 암호화하는 유전자는 미국 특허 제7,005,291호에 기술되어 있다. 글리세롤-3-포스페이트 탈수소효소 활성을 갖는 폴리펩티드를 암호화하는 유전자의 적절한 예는 사카로마이세스 세레비시아 유래의 GPD1(서열 번호 1에 기재된 암호화 서열, 서열 번호 2에 기재된 암호화된 단백질 서열) 및 사카로마이세스 세레비시아 유래의 GPD2(서열 번호 3에 기재된 암호화 서열, 서열 번호 4에 기재된 암호화된 단백질 서열)를 포함하나, 이에 한정되지 않는다. 글리세롤-3-포스파타제 활성을 갖는 폴리펩티드를 암호화하는 유전자의 적절한 예는 사카로마이세스 세레비시아 유래의 GPP1(서열 번호 5에 기재된 암호화 서열, 서열 번호 6에 기재된 암호화된 단백질 서열) 및 사카로마이세스 세레비시아 유래의 GPP2(서열 번호 7에 기재된 암호화 서열, 서열 번호 8에 기재된 암호화된 단백질 서열)를 포함하나, 이에 한정되지 않는다.

[0144] 표적 내인성 유전자의 발현을 감소시킴으로써 글리세롤의 증가된 생성을 달성할 수 있다. 미국 특허 제7,005,291호에 기술된 바와 같이, 글리세롤 키나제 및 글리세롤 탈수소효소 활성을 암호화하는 내인성 유전자의 하향 조절은 글리세롤 생성을 더욱 증진시킨다. 미국 특허 제7,371,558호에 기술된 바와 같이, 글리세르알데히드 3-포스페이트 탈수소효소를 암호화하는 내인성 유전자의 발현을 감소시킴으로써 글리세롤로의 탄소의 증가된 채널링(channeling)이 이루어질 수 있다. 당업계에 공지된 임의의 방법, 예를 들어, PTS 시스템의 유전자의 하향 조절을 위한 상기 방법을 사용함으로써 하향 조절이 이루어질 수 있다.

[0145] 글리세롤은 유용한 생성물의 미생물 생성을 위한 기질을 제공한다. 이러한 생성물, 즉, 글리세롤 유도체의 예는 3-하이드록시프로피온산, 메틸글리옥살, 1,2-프로판다이올, 및 1,3-프로판다이올을 포함하나, 이에 한정되지 않는다.

- [0146] 일부 실시형태에서, 본 명세서에 개시된 재조합 박테리아는 1,3-프로판다이올을 생성할 수 있다. 글리세롤 유도체 1,3-프로판다이올은 폴리에스테르 섬유의 생성 및 폴리우레탄 및 환형 화합물의 제조에 있어서 잠재적 유용성을 갖는 모노머이다. 미국 특허 제5,686,276호에 기재된 바와 같이, 1,3-프로판다이올은 글리세롤 또는 다이하이드록시아세톤 이외의 탄소 기질의 생물전환에 의해 단일 미생물에 의해 생성될 수 있다. 이 생물전환에서는, 상기 기재된 바와 같이 탄소 기질로부터 글리세롤이 생성된다. 숙주 박테리아에 의해 암호화될 수 있거나 재조합에 의해 숙주 내로 도입될 수 있는 탈수효소에 의해 글리세롤이 중간체 3-하이드록시프로피온알데히드로 전환된다. 탈수효소는 글리세롤 탈수효소(E.C. 4.2.1.30), 다이올 탈수효소(E.C. 4.2.1.28) 또는 이러한 전환을 촉매작용시킬 수 있는 임의의 다른 효소일 수 있다. 글리세롤 탈수효소의 " α "(알파), " β "(베타), 및 " γ "(감마) 서브유닛을 암호화하는 유전자의 적절한 예는, 각각 클레브시엘라 뉴모니아 유래의 dhaB1(서열 번호 9에 기재된 암호화 서열), dhaB2(서열 번호 11에 기재된 암호화 서열) 및 dhaB3(서열 번호 13에 기재된 암호화 서열)을 포함하나, 이들에 한정되지 않는다. 3-하이드록시프로피온알데히드의 1,3-프로판다이올로의 추가의 전환은 1,3-프로판다이올 탈수소효소(E.C. 1.1.1.202) 또는 다른 알코올 탈수소효소에 의하여 촉매될 수 있다. 1,3-프로판다이올 탈수소효소를 암호화하는 유전자의 적절한 예는 클레브시엘라 뉴모니아 유래의 dhaT(서열 번호 57에 기재된 암호화 서열, 서열 번호 58에 기재된 암호화된 단백질 서열)이다.
- [0147] 박테리아를 재조합 조작하여 글리세롤 및 글리세롤 유도체 1,3-프로판다이올의 더욱 효율적인 생성을 제공할 수 있다. 예를 들어, 미국 특허 제7,005,291호는 글리세롤 키나제 및 글리세롤 탈수소효소의 내인성 활성 중 하나 또는 양자 모두를 파괴하면서, 글리세롤-3-포스페이트 탈수소효소 및 글리세롤-3-포스페이트 포스파타제 중 하나 또는 양자 모두의 외인성 활성을 발현시킴에서 유래하는 이점을 가진, 글리세롤 및 1,3-프로판다이올의 생성을 위한 형질전환된 미생물 및 그 생성 방법을 개시한다.
- [0148] 미국특허 제6,013,494호는, 외인성의 글리세롤-3-포스페이트 탈수소효소, 글리세롤-3-포스페이트 포스파타제, 탈수효소 및 1,3-프로판다이올 산화환원효소(예를 들어, dhaT)를 포함하는 단일의 미생물을 이용한 1,3-프로판다이올의 생성 방법을 기재한다. 미국 특허 제6,136,576호는, 탈수효소 및 단백질 X(이후 탈수효소 재활성화 인자 펩티드로서 확인됨)를 더 포함하는 재조합 미생물을 포함하는 1,3-프로판다이올의 생성 방법을 개시한다.
- [0149] 미국 특허 제6,514,733호는, 3-하이드록시프로피온알데히드를 1,3-프로판다이올로 전환시키는 비-특이적 촉매 활성에 의하여(dhaT에 의하여 암호화된 1,3-프로판다이올 산화환원효소와 구분됨) 역가(titer)의 유의한 증가(리터당 생성물 g)가 수득되는 방법의 개선을 기재한다. 부가적으로, 미국 특허 제7,132,527호는 1,3-프로판다이올의 생성에 유용한 벡터 및 플라스미드를 개시한다.
- [0150] 1,3-프로판다이올의 증가된 생성은 숙주 미생물에 대한 추가의 변형에 의하여 달성될 수 있으며, 이러한 변형은 일부 표적 유전자의 발현의 하향 조절 및 다른 표적 유전자의 발현의 상향 조절을 포함하며, 이는 미국 특허 제7,371,558호에 기재된 바와 같다. PTS 마이너스 숙주에서 탄소원으로 글루코스를 이용함에 있어서, 글루코키나제 활성의 발현이 증가될 수 있다.
- [0151] 증가된 발현 또는 상향 조절된 발현이 1,3-프로판다이올 생성을 증가시키는 추가의 유전자는 하기의 것들을 암호화하는 유전자를 포함한다:
- [0152] • 전형적으로 EC 4.1.1.31로 특성화되는 포스포에놀피루베이트 카르복실라제
 - [0153] • 전형적으로 EC 2.5.1.17로 특성화되는 코브(1)알라민 아데노실트랜스퍼라제
 - [0154] • 3-HPA 및 1,3-프로판다이올의 상호 전환을 촉매작용시키기에 충분하며, 특이적으로 1,3-프로판다이올 산화환원효소(들)를 배제하는 비-특이적 촉매 활성(전형적으로 이들 효소는 알코올 탈수소효소임)
 - [0155] 감소된 발현 또는 하향 조절된 발현이 1,3-프로판다이올 생성을 증가시키는 유전자는 하기의 것들을 암호화하는 유전자를 포함한다:
 - [0156] • 호기성 호흡 제어 단백질
 - [0157] • 메틸글리옥살 신타제
 - [0158] • 아세테이트 키나제
 - [0159] • 포스포트랜스아세틸라제

- [0160] • 알데히드 탈수소효소 A
- [0161] • 알데히드 탈수소효소 B
- [0162] • 트라이오스포스페이트 아이소머라제
- [0163] • 포스포글루코네이트 탈수효소
- [0164] 일부 실시형태에서, 본 명세서에 개시된 재조합 박테리아는 3-하이드록시프로피온산을 생성할 수 있다. 3-하이드록시프로피온산은 특수 합성에 있어서 유용성을 가지며, 화학 산업의 공지 기술에 의해 상업적으로 중요한 중간체로 전환될 수 있다(예를 들어, 탈수에 의해 아크릴산, 산화에 의해 말론산, 알코올과의 에스테르화 반응에 의해 에스테르, 그리고 환원에 의해 1,3-프로판다이올). 3-하이드록시프로피온산은 본 출원과 공개류 중인 그리고 공히 소유된 미국 특허 출원 공개 제2011/0144377 A1호에 기술된 바와 같이, 단일 미생물에 의해 발효성 탄소원으로부터 생물학적으로 생성될 수 있다. 1,3-프로판다이올의 생성에 관해 상기 기술된 바와 같이, 한가지 대표적인 생합성 경로에서는 탄소 기질이 3-하이드록시프로피온알데히드로 전환된다. 3-하이드록시프로피온알데히드는 알데히드 탈수소효소에 의해 3-하이드록시프로피온산으로 전환된다. 알데히드 탈수소효소의 적절한 예들로, 이에 제한되지는 않지만, 에스케리키아 콜라이 유전자 aldB(서열 번호 15에 기재된 암호화 서열)에 의하여 암호화되는 AldB(서열 번호 16); 에스케리키아 콜라이 유전자 aldA(서열 번호 17에 기재된 암호화 서열)에 의하여 암호화되는 AldA(서열 번호 18); 및 에스케리키아 콜라이 유전자 aldH(서열 번호 19에 기재된 암호화 서열)에 의하여 암호화되는 AldH(서열 번호 20)가 포함된다.
- [0165] 재조합 박테리아에 의해 1,3-프로판다이올 생성을 개선하기 위한 상기 변형 중 다수는 또한, 3-하이드록시프로피온산 생성을 개선하게 만들 수 있다. 예를 들어, 글리세롤 키나제의 제거는 G3P 포스파타제의 작용에 의해 G3P로부터 형성되는 글리세롤이 ATP를 소모하면서 G3P로 재전환되는 것을 방지한다. 또한, 글리세롤 탈수소효소(예를 들어, gldA)의 제거는, NAD-의존성 글리세롤-3-포스페이트 탈수소효소의 작용에 의해 DHAP로부터 형성되는 글리세롤이 다이하이드록시아세톤으로 전환되는 것을 방지한다. 돌연변이는, 효소 활성을 손상시키거나 개선하도록 구조 유전자에 지향될 수도 있고, 효소 활성의 발현 수준을 조정하도록 프로모터 영역 및 리보솜 결합 부위를 포함하는 조절 유전자에 지향될 수도 있다.
- [0166] 상향 조절 또는 하향 조절은 당업자에게 공지된 다양한 방법에 의해 달성될 수 있다. 유전자의 상향 조절 또는 하향 조절은, 그 유전자에 의해 암호화되는 단백질로부터 유래하는 세포 내에 존재하는 활성 수준이, 예를 들어 상응하는(또는 변경되지 않은) 야생형 유전자에 의해 암호화되는 단백질의 활성에 의한 대조군 활성 수준에 비교하여 변경됨을 지칭한다는 것이 주지되어 있다.
- [0167] 효소 경로에 관련되는 특정 유전자를 상향 조절하여 그들의 암호화된 기능(들)의 활성을 증가시킬 수 있다. 예를 들어, 선택된 유전자의 추가의 카피를 pBR322와 같은 다중카피 플라스미드 상에 숙주 세포 내로 도입할 수 있다. 이러한 유전자는 또한, 그들의 암호화된 기능의 증가된 활성을 유발하는 적절한 조절 서열과 함께 염색체 내로 통합될 수 있다. 비-고유 프로모터 또는 변경된 고유 프로모터의 제어 하에 있도록 표적 유전자를 변형시킬 수 있다. 돌연변이, 결실, 및/또는 치환에 의해 내인성 프로모터를 생체내 변경할 수 있다.
- [0168] 대안적으로, 소정의 유전자의 발현을 주어진 활성 수준에 대해 감소시키거나 제거하는 것이 유용할 수 있다. 유전자를 하향 조절(파괴)하는 방법은 당업자에게 공지되어 있다.
- [0169] 하향 조절은 암호화 영역 및/또는 조절(프로모터) 영역의 결실, 삽입, 또는 변경에 의해 발생할 수 있다. 무작위 돌연변이 후의 스크리닝 또는 선택에 의해, 또는 유전자 서열이 공지된 경우에는 당업자에게 공지된 분자생물학 방법에 의한 직접 개입에 의해, 특정 하향 조절을 얻을 수 있다. 하향 조절이 이루어지게 하는 특히 유용하나 배타적이지 않은 방법은 프로모터 강도를 변경하는 것이다.
- [0170] 또한, 유전자 발현의 하향 조절을 사용하여 관심의 대상인 단백질의 발현을 방지하거나 비-기능성인 단백질의 발현을 유발할 수 있다. 이는 예를 들어, 1) 암호화 영역 및/또는 조절(프로모터) 영역의 결실, 2) 암호화 영역 및/또는 조절(프로모터) 영역 내로의 외인성 핵산 서열의 삽입, 및 3) 암호화 영역 및/또는 조절(프로모터) 영역의 변경(예를 들어, DNA 염기쌍을 변화시킴으로써)에 의해 이루어질 수 있다. 무작위 돌연변이 후의 스크리닝 또는 선택에 의해 특정 파괴를 얻을 수도 있거나, 유전자 서열이 공지된 경우에는 당업자에게 공지된 분자생물학 방법을 사용하는 직접 개입에 의해 특정 파괴를 얻을 수도 있다. 특히 유용한 방법은 유의적인 양의 암호화 영역 및/또는 조절(프로모터) 영역의 결실이다.

- [0171] 재조합 단백질 발현을 변경하는 방법은 당업자에게 공지되어 있으며, 문헌[Baneyx, Curr. Opin. Biotechnol. (1999) 10:411]; 문헌[Ross, et al., J. Bacteriol. (1998) 180:5375]; 문헌[deHaseth, et al., J. Bacteriol. (1998) 180:3019]; 문헌[Smolke and Keasling, Biotechnol. Bioeng. (2002) 80:762]; 문헌[Swartz, Curr. Opin. Biotech. (2001) 12:195]; 및 문헌[Ma, et al., J. Bacteriol. (2002) 184:5733]에 부분적으로 개시되어 있다.
- [0172] 글리세롤 및 글리세롤 유도체를 포함하는 미생물 생성물의 생성에 있어서 수크로스를 대사시키기 위한 유전자 발현에 필요한 변화를 포함하는 재조합 박테리아는 상기와 같이, 당업계에서 주지된 기술을 사용하여 구축될 수 있다.
- [0173] 본 명세서에 개시된 재조합 박테리아의 구축은, 적절한 숙주 미생물 내에서의 글리세롤 및 그의 유도체의 생성에 수크로스를 이용하는 능력을 부여하는 암호화 영역의 클로닝, 형질전환 및 발현에 적절한 다양한 벡터 및 형질전환 및 발현 카세트를 사용하여 이루어질 수 있다. 적절한 벡터는 채용되는 박테리아와 상용성인 것들이다. 적절한 벡터는, 예를 들어 박테리아, 바이러스(예를 들어, 박테리오파지 T7 또는 M-13 유래의 파지), 코스미드, 효모 또는 식물로부터 유래될 수 있다. 이러한 벡터를 얻고 사용하기 위한 프로토콜은 당업자에게 공지되어 있다(상기 문헌[Sambrook et al.]).
- [0174] 목적하는 숙주 박테리아에서 본 발명을 위한 암호화 영역의 발현을 추진하기에 유용한 개시 제어 영역, 또는 프로모터는 다수이며 당업자에게 잘 알려져 있다. 발현을 추진할 수 있는 사실상 임의의 프로모터가 본 명세서에 사용하기에 적합하다. 예를 들어, 상기 열거된 프로모터 중 임의의 것을 사용할 수 있다.
- [0175] 종결 제어 영역은 또한 바람직한 숙주의 다양한 고유의 유전자로부터 유래될 수 있다. 선택적으로, 종결 부위는 불필요할 수도 있지만; 포함되는 경우가 가장 바람직하다.
- [0176] 본 폴리펩티드의 효과적인 발현을 위하여, 폴리펩티드를 암호화하는 뉴클레오티드 서열을 선택된 발현 제어 영역에 개시 코돈을 통해 작동가능하게 연결하여, 발현이 적절한 메신저(messenger) RNA의 형성을 유발하도록 한다.
- [0177] 미국 특허 제7,371,558호에 기재된 벡터 pSYCO101, pSYCO103, pSYCO106, 및 pSYCO109, 및 미국 특허 제7,524,660호에 기재된 pSYCO400/AGRO가 특히 유용하다. 이들 벡터의 필수 요소는 클레브시엘라 뉴모니아 및 사카로마이세스 세레비시아로부터 분리된 dha 레귤론으로부터 유래된다. 각각의 벡터는 오픈 리딩 프레임, dhaB1, dhaB2, dhaB3, dhaX(서열 번호 59에 기재된 암호화 서열; 서열 번호 60에 기재된 암호화된 폴리펩티드 서열), orfX, DAR1 및 GPP2(3개의 별도 오페론으로 배열됨)를 함유한다. pSYCO101, pSYCO103, pSYCO106, pSYCO109 및 pSYCO400/AGRO의 뉴클레오티드 서열은 각각 서열 번호 62, 서열 번호 63, 서열 번호 64, 서열 번호 65 및 서열 번호 66에 기재되어 있다. 벡터들 간의 차이는 하기 차트에 상술하였다[접두사 "p-"는 프로모터를 나타내며; 각각의 "(" 내에 포함된 오픈 리딩 프레임은 오페론 구성요소를 나타낸다]:
- [0178] pSYCO101(서열 번호 62):
- [0179] 다른 2개의 경로 오페론에 비교하여 반대 배향인 p-trc(Dar1_GPP2),
- [0180] p-1.6 장 GI(dhaB1_dhaB2_dhaB3_dhaX), 및
- [0181] p-1.6 장 GI(orfY_orfX_orfW).
- [0182] pSYCO103(서열 번호 63):
- [0183] 다른 2개의 경로 오페론에 비교하여 동일 배향인 p-trc(Dar1_GPP2),
- [0184] p-1.5 장 GI(dhaB1_dhaB2_dhaB3_dhaX), 및
- [0185] p-1.5 장 GI(orfY_orfX_orfW).
- [0186] pSYCO106(서열 번호 64):
- [0187] 다른 2개의 경로 오페론에 비교하여 동일 배향인 p-trc(Dar1_GPP2),
- [0188] p-1.6 장 GI(dhaB1_dhaB2_dhaB3_dhaX), 및
- [0189] p-1.6 장 GI(orfY_orfX_orfW).

- [0190] pSYCO109(서열 번호 65):
- [0191] 다른 2개의 경로 오페론에 비교하여 동일 배향인 p-trc(Dar1_GPP2),
- [0192] p-1.6 장 GI(dhaB1_dhaB2_dhaB3_dhaX), 및
- [0193] p-1.6 장 GI(orfY_orfX).
- [0194] pSYCO400/AGRO(서열 번호 66):
- [0195] 다른 2개의 경로 오페론에 비교하여 동일 배향인 p-trc(Dar1_GPP2),
- [0196] p-1.6 장 GI(dhaB1_dhaB2_dhaB3_dhaX), 및
- [0197] p-1.6 장 GI(orfY_orfX).
- [0198] 경로 오페론에 비교하여 반대 배향의 p-1.20 단/장 GI(scrK).
- [0199] 일단 적절한 발현 카세트가 구축되면, 그들을 사용하여 적절한 숙주 박테리아를 형질전환한다. 암호화 영역을 함유하는 카세트의 숙주 박테리아 내로의 도입은, 형질전환에 의한 것(예를 들어, 칼슘-투과화 세포, 또는 전기 천공을 사용함) 또는 재조합 파지 바이러스를 사용하는 트랜스펙션에 의한 것(상기 문헌[Sambrook et al.])과 같은 공지의 절차에 의해 이루어질 수 있다. 발현 카세트는 숙주 세포 내에서 안정한 플라스미드 상에 유지될 수 있다. 또한, 당업자에게 주지된 벡터 및 방법을 사용하는 상동 또는 무작위 재조합을 통해 발현 카세트를 숙주 박테리아의 게놈 내에 통합할 수 있다. 부위-특이적 재조합 시스템 또한 발현 카세트의 게놈 통합에 사용할 수 있다.
- [0200] 예시된 세포에 부가하여, 글리세롤 및/또는 그의 유도체를 포함하는 미생물 생성물의 생성을 증진하도록 특이적으로 고안된 단일 또는 다중 돌연변이를 갖는 세포 또한 사용할 수 있다. 정상적으로 탄소 공급 원료를 비-생성 경로에 전용하거나, 유의적인 이화산물 억제제를 나타내는 세포는, 이들 표현형 결함을 회피하도록 돌연변이될 수 있다.
- [0201] 돌연변이체를 생성시키는 방법은 통상적이며 당업계에 주지되어 있다. 일부 방법의 요약이 미국 특허 제 7,371,558호에 제공되어 있다. 방사선 또는 화학 약제를 사용하여 돌연변이체를 생성시키는 특정 방법이 당업계에 잘 입증되어 있다. 예를 들어, 문헌[Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA.] 또는 문헌[Deshpande, Mukund V., Appl. Biochem. Biotechnol. 36, 227 (1992)]을 참조한다.
- [0202] 돌연변이유발이 발생한 후에, 목적하는 표현형을 갖는 돌연변이체를 다양한 방법으로 선택할 수 있다. 무작위 스크리닝이 가장 통상적이며, 여기서 돌연변이된 세포는 목적하는 생성물 또는 중간체를 생성하는 능력에 대하여 선택된다. 대안적으로, 돌연변이된 개체군을 저항성 콜로니만 발생할 수 있는 선택적 배지 상에서 성장시킴으로써 돌연변이체의 선택적 분리를 수행할 수 있다. 돌연변이체 선택 방법은 고도로 개발되어 산업 미생물학 분야에 주지되어 있다. 예를 들어, 상기 문헌[Brock]; 문헌[DeMancilha et al., Food Chem. 14, 313 (1984)]을 참조한다.
- [0203] 본 발명에서 발효 배지는 수크로스를 탄소 기질로 포함한다. 글루코스 및 프룩토스와 같은 다른 탄소 기질 또한 존재할 수 있다.
- [0204] 적절한 발효 배지는, 탄소 기질에 부가하여, 예를 들어, 글리세롤 및 그의 유도체, 예를 들어 1,3-프로판다이올의 생성을 위해 필요한 효소 경로의 촉진 및 배양물의 성장에 적절한, 당업자에게 공지된 적절한 미네랄, 염, 보조인자, 완충제 및 다른 구성성분을 함유한다. 1,3-프로판다이올의 생성에 있어서, Co(II) 염 및/또는 비타민 B₁₂ 또는 그의 전구체가 특히 주목된다.
- [0205] 아데노실-코발아민(조효소 B₁₂)은 탈수효소 활성화에 있어서 중요한 보조인자이다. 원핵생물에서는 조효소 B₁₂의 합성이 발견되며, 이들 중 일부는 화합물의 드 노보(de novo) 합성이 가능하며, 예를 들어 에스케리키아 블라타타에(*Escherichia blattae*), 클레브시엘라 종, 시트로박터 종, 및 클로스트리디움 종이 있는 반면, 다른 것들은 부분 반응을 수행할 수 있다. 예를 들어, 에스케리키아 콜라이는 코린 고리 구조를 제작할 수 없지만, 코빈아미드를 코리노이드로 전환시키는 촉매 작용을 할 수 있으며 5'-데옥시아데노실기를 도입할 수 있다. 따라서, 비타민 B₁₂와 같은 조효소 B₁₂ 전구체를 에스케리키아 콜라이 발효에 공급할 필요가 있다는 것이 당업계에 공지되

어 있다. 비타민 B₁₂는 에스케리키아 콜라이 발효에 일정한 속도로 연속적으로 첨가하거나, 세포 매스의 생성과 일치하도록 단계화할 수 있거나, 또는 단일 또는 다중 볼루스 첨가로 첨가할 수도 있다.

[0206] 본 명세서에 기재된 형질전환된 에스케리키아 콜라이에 비타민 B₁₂가 첨가되지만, 드 노보 비타민 B₁₂ 생합성이 가능한 다른 박테리아 또한 적절한 생성 세포일 것이며, 이들 박테리아에는 비타민 B₁₂의 첨가가 불필요할 것으로 예상된다.

[0207] 전형적으로 박테리아 세포는 25 내지 40℃에서 수크로스를 함유하는 적절한 배지 내에서 성장한다. 본 명세서에 사용하기에 적절한 성장 배지의 예는, 루리아 베르타니(LB: Luria Bertani) 브로쓰, 사보로드 텍스트로스(SD: Sabouraud Dextrose) 브로쓰 또는 효모 배지(YM: Yeast medium) 브로쓰와 같은 상업적으로 제조된 통상적인 배지이다. 그밖에 정의되거나 합성된 성장 배지도 사용할 수 있으며, 특정 박테리아의 성장에 적절한 배지는 미생물학 또는 발효과학 분야의 당업자에게 공지되어 있을 것이다. 직접적으로 또는 간접적으로 이화산물 억제제를 조정하는 것으로 공지된 약제, 예를 들어 사이클릭 아데노신 2',3'-모노포스페이트의 사용 또한 반응 배지 내에 혼입될 수 있다. 마찬가지로, 1,3-프로판다이올 생성의 증진을 유발하는 효소 활성을 조정하는 것으로 공지된 약제(예를 들어, 메틸 비올로겐)의 사용을, 1,3-프로판다이올 생성 균주를 이용하는 유전자 조작과 함께, 또는 그 대안으로 사용할 수 있다.

[0208] 발효에 적절한 pH 범위는 pH 5.0 내지 pH 9.0이며, 여기서 pH 6.0 내지 pH 8.0이 초기 조건으로서 전형적이다.

[0209] 재조합 박테리아의 요구 조건에 따라, 호기성, 무산소, 또는 혐기성 조건 하에 반응을 수행할 수 있다. 유가식 발효는 한정되거나 과량인 탄소 공급물, 예를 들어 탄소 기질을 사용하여 수행될 수 있다.

[0210] 회분식 발효는 통상적으로 사용되는 방법이다. 고전적인 회분식 발효는, 배지의 조성이 발효 시작시에 설정되고 발효 중에 인공적으로 변경되지 않는 폐쇄 시스템이다. 따라서, 목적하는 박테리아를 발효의 시작시에 배지에 접종하며, 시스템에 아무것도 첨가하지 않고서 발효가 일어나게 한다. 그러나, 전형적으로 "회분식" 발효는 탄소원의 첨가와 관련하여 회분식이며, pH 및 산소 농도와 같은, 인자를 조절하려는 시도가 흔히 이루어진다. 회분식 시스템에서, 시스템의 대사산물 및 바이오매스 조성은 발효가 중단되는 시점까지 끊임없이 변화한다. 회분식 배양 내에서, 세포는 정적 유도기(static lag phase)를 통하여 고 성장 대수기(log phase)로, 그리고 최종적으로 성장 속도가 감소되거나 정지되는 정체기(stationary phase)로 완화된다. 미처리될 경우, 정체기의 세포는 결국 사멸할 것이다. 대수기의 세포는 일반적으로 최종 생성물 또는 중간체의 대량 생성에 책임이 있다.

[0211] 표준 회분식 시스템의 한 변형은 유가식(Fed-Batch) 시스템이다. 유가식 발효 방법 또한 본 명세서에 사용하기에 적합하며, 이는 발효가 진행됨에 따라 기질을 증분식으로 첨가한다는 점을 제외하고는 전형적인 회분식 시스템을 포함한다. 유가식 시스템은 이화산물 억제제가 세포의 대사작용을 저해하는 경향이 있을 때 그리고 배지 중에 한정된 양의 기질이 있는 것이 바람직한 경우 유용하다. 유가식 시스템 중 실제 기질 농도의 측정은 어려우며, 따라서 pH, 용존 산소 및 CO₂와 같은 폐가스의 분압과 같은 측정가능한 인자의 변화를 기반으로 하여 계산된다. 회분식 및 유가식 발효는 통상적이며 당업계에 주지되어 있고, 그 예는 상기 문헌[Brock]에서 발견할 수 있다.

[0212] 연속식 발효는 정의된 발효 배지가 연속적으로 생물 반응기에 첨가되고, 동일한 양의 조정된 배지가 프로세싱을 위하여 동시에 제거되는 개방 시스템이다. 연속 발효는 일반적으로, 세포가 주로 대수기 성장에 있는 경우 일정한 고밀도로 배양물을 유지시킨다.

[0213] 연속식 발효는 세포 성장 또는 최종 생성물 농도에 영향을 주는 하나의 인자 또는 임의의 개수의 인자의 조정을 허용한다. 예를 들어, 한 가지 방법은 제한 영양소, 예를 들어 탄소원 또는 질소의 수준을 고정된 비율로 유지하고 모든 다른 파라미터가 조정되게 할 것이다. 다른 시스템에서는, 배지의 탁도로 측정되는 세포 농도를 일정하게 유지하는 가운데, 성장에 영향을 주는 다수의 인자를 연속적으로 변경할 수 있다. 연속식 시스템은 안정 상태 성장 조건을 유지하도록 노력하므로, 배출되는 배지로 인한 세포 손실은 발효에서 세포 성장 속도에 대하여 균형이 맞추어져야 한다. 연속식 발효 방법에 있어서 영양소 및 성장 인자를 조정하는 방법뿐만 아니라, 생성물 형성 속도를 최대화하는 기술도 산업 미생물학 분야에서 주지되어 있으며, 다양한 방법이 상기 문헌[Brock]에 상술되어 있다.

[0214] 본 발명은 회분식, 유가식 또는 연속식 방법을 사용하여 실시할 수 있으며 임의의 알려진 발효 모드가 적합할 것으로 예상된다. 부가적으로, 세포를 기재 상에 전세포 촉매로서 고정시키고 글리세롤 및 글리세롤 유도체,

예를 들어 1,3-프로판다이올의 생성을 위한 발효 조건으로 처리할 수 있을 것으로 예상된다.

[0215] 일부 실시형태에는, 수크로스로부터 글리세롤, 1,3-프로판다이올, 및/또는 3-하이드록시프로피온산을 제조하는 방법이 제공된다. 본 방법은, 상기와 같이 수크로스의 존재 하에 재조합 박테리아를 배양하는 단계, 및 임의로 생성된 글리세롤, 1,3-프로판다이올, 및/또는 3-하이드록시프로피온산을 회수하는 단계를 포함한다. 당업계에 공지된 방법을 사용하여 생성물을 회수할 수 있다. 예를 들어, 원심분리, 여과, 경사분리(decantation) 등에 의해 발효 배지로부터 고체를 제거할 수 있다. 이어서, 증류, 액-액 추출 또는 막-기반의 분리와 같은 방법을 사용하여, 상기와 같이 고체를 제거하기 위해 처리된 발효 배지로부터 생성물을 분리할 수 있다.

[0216] 실시예

[0217] 본 발명은 하기 실시예에서 추가로 정의된다. 이들 실시예는 본 발명의 바람직한 실시형태를 나타내면서도 단지 예로써만 주어지는 것으로 이해해야 할 것이다. 상기 토의 및 이들 실시예로부터, 당업자는 본 발명의 필수적인 특징을 확인할 수 있고, 그의 취지 및 범주를 벗어나지 않으면서, 본 발명의 다양한 변화 및 개질을 하여서 본 발명이 다양한 용도와 조건에 적합하도록 할 수 있다.

[0218] 일반 방법

[0219] 본 실시예에 사용되는 표준 재조합 DNA 및 분자 클로닝 기술은 본 기술 분야에 잘 알려져 있으며, 문헌[Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989) (Maniatis)], 및 문헌[T. J. Silhavy, M. L. Bannan, and L. W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984)], 및 문헌[Ausubel, F. M. et al., Current Protocols in Molecular Biology, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987)]에 기술되어 있다.

[0220] 약어의 의미는 다음과 같다: "sec"는 초를 의미하며, "min"은 분을 의미하고, "h"는 시간을 의미하며, "nm"은 나노미터를 의미하고, "μL"은 마이크로리터를 의미하며, "mL"은 밀리리터를 의미하고, "L"은 리터를 의미하며, "mM"은 밀리몰을 의미하고, "M"은 몰을 의미하며, "g"는 그램을 의미하고, "μg"는 마이크로그램을 의미하며, "bp"는 염기쌍을 의미하고, "kbp"는 킬로염기쌍을 의미하며, "rpm"은 분당 회전수를 의미하고, "g"는 중력 상수를 의미하며, "ATCC"는 미국 버지니아주 머내서스 소재의 아메리칸 타입 컬처 콜렉션을 의미하고, "OD"는 광학 밀도를 의미하며, "HPLC"는 고성능 액체 크로마토그래피를 의미한다.

[0221] 배지 및 배양 조건:

[0222] 박테리아 배양물의 유지 및 성장을 위해 적절한 재료 및 방법은 당업계에 주지되어 있다. 하기 실시예에서의 이용에 적절한 기술은 문헌[Manual of Methods for General Bacteriology (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994)] 또는 문헌[Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989)]에 정리된 바와 같이 찾을 수 있다. 박테리아 세포의 유지 및 성장에 관해 기술된 모든 시약, 제한 효소 및 재료는 달리 특정되지 않는 한, 알드리치 케미컬즈(미국 위스콘신주 밀워키 소재), 비디 디아그노스틱 시스템즈(BD Diagnostic Systems)(미국 메릴랜드주 스파크스 소재), 라이프 테크놀로지즈(미국 메릴랜드주 록빌 소재), 뉴 잉글랜드 바이오랩스(New England Biolabs)(미국 매사추세츠주 비벌리 소재) 또는 시그마 케미컬 컴퍼니(Sigma Chemical Company)(미국 미주리주 세인트루이스 소재)로부터 입수할 수 있다.

[0223] LB(루리아 베르타니) 배지는 배지 1리터당 박토-트립톤(Bacto-tryptone)(10 g), 박토-효모 추출물(5 g) 및 NaCl(10 g)을 포함한다. LB 저염 배지는 배지 1리터당 박토-트립톤(10 g), 박토-효모 추출물(5 g) 및 NaCl(5 g)을 포함한다. 보충물은 하기 실시예에 기재된 것과 같이 첨가되었다. 모든 첨가물은 배지에 첨가되기 전에 미리 멸균되었다.

[0224] 분자 생물학적 기술:

[0225] 제한 효소 분해, 라이게이션, 형질전환 및 아가로스 겔 전기영동 방법을 문헌[Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press (1989)]에 기재된 것과 같이 수행하였다. 중합효소 연쇄 반응(PCR) 기술은 문헌[White, B., PCR Protocols: Current Methods and Applications, Volume 15 (1993), Humana Press Inc., New York, NY]에서 찾을 수 있다.

- [0226] 실시예 1 내지 34
- [0227] 추정된 수크로스 수송체 유전자를 함유하는 재조합 에스케리키아 콜라이 균주
- [0228] 이들 실시예의 목적은 다양한 추정된 수크로스 수송체 유전자를 함유하는 재조합 에스케리키아 콜라이 균주를 구축하려는 것이었다. 먼저, 수크로스 수송체가 결여되거나 수크로스 인버타제 및 프룩토키나제 유전자를 함유하는 본 명세서에서 PD03513으로 지칭되는 에스케리키아 콜라이 균주(즉, cscA+K+B⁻ (kanR) 균주)를 구축하였다. 이어서, 다양한 추정된 수크로스 수송체 유전자를 에스케리키아 콜라이 균주 PD03513로 도입하였다.
- [0229] 에스케리키아 콜라이 균주 PD03513의 구축
- [0230] 에스케리키아 콜라이 균주 PD03513을, 에스케리키아 콜라이 ATCC®13281 유래의 야생형 cscAKB 유전자 클러스터가 에스케리키아 콜라이 균주 FM5(ATCC® 번호 53911) 내의 yihP 유전자에서 통합된 에스케리키아 콜라이 균주(본 명세서에서 PD03085로 지칭)로부터 구축하였다. cscAKB 유전자 클러스터(서열 번호 61)를 람다 레드(Lambda Red) 방법에 의해 에스케리키아 콜라이 균주 FM5(ATCC® 번호 53911) 내의 yihP 위치에 통합시켰다. cscAKB 유전자 클러스터를, yihP 유전자의 플랜킹 서열을 포함하는 yihP cscA 프라이머(서열 번호 134) 및 yihP cscB 프라이머(서열 번호 135)를 사용하여, 미국 특허 출원 공개 제2011/0136190 A1호의 실시예 1에 설명한 바와 같이 구축한 플라스미드 pBHR-cscBKA(서열 번호 133)로부터 증폭시켰다. PstI 분해에 의해 선형화한 플라스미드 pBHR-cscBKA를 PCR 주형으로 사용하였다. 고충실도 PfuUltra® II 퓨전(Fusion) HS DNA 중합효소(스트라타진(Stratagene); 미국 캘리포니아주 라 졸라 소재)를 PCR 반응에서 사용하였다. PCR을 하기 사이클링 조건들을 이용하여 수행하였다: 95℃, 2분; 95℃, 30초, 60℃, 30초 및 72℃, 4분의 35 사이클; 및 그 후 72℃, 7분. 생성된 PCR 생성물을 4℃에서 보관하였다. PCR 생성물을 퀴아퀵(QIAquick) PCR 정제 키트(미국 캘리포니아주 발렌시아 소재의 퀴아젠(Qiagen))를 사용하여 정제하였다. 정제된 PCR 생성물을 람다 레드 재조합 절차(문헌 [Datsenko, K.A. and Wanner, B.L., 2000, Proc. Natl. Acad. Sci. USA 97, 6640-6645])에 따라, 람다 재조합 효소를 암호화하는 pKD46 플라스미드(레드 재조합효소 플라스미드, 진뱅크 수탁 번호 AY048746)를 함유하는 에스케리키아 콜라이 균주 FM5로 전기천공시켰다. 형질전환 혼합물을 10 g/L의 수크로스 및 100 µg/mL의 스펙티노마이신을 함유하는 MOPS 최소 플레이트 상에 플레이팅하였다. MOPS 최소 플레이트는 1XMOPS 완충제(미국 캘리포니아주 홀리스트 소재의 테크노바(Technova)), 1.32 mM의 KH₂PO₄(테크노바), 50 µg/L의 우라실 및 1.5 g/L의 박토(Bacto) 아가를 함유하였다. 플레이트를 37℃에서 2 내지 3일 동안 인큐베이션하였다. 최소 수크로스 플레이트 상에 성장시킨 콜로니를 피킹하여, 에스케리키아 콜라이 균주 PD03085를 제공하였다.
- [0231] 그 후, PD03085 내의 상기 클러스터 내의 cscB 유전자를 카나마이신 내성 카세트에 의한 이의 대체에 의해 부분적으로 결실시켰다. 카나마이신 내성 카세트를 pKD4 주형 플라스미드로부터 증폭시켰으며(문헌 [Datsenko and Wanner, Proc. Natl. Acad. Sci. USA 97:6640-6645, 2000]), 이는 cscB61 up kan 프라이머(서열 번호 136) 및 cscB353 down kan 프라이머(서열 번호 137)를 사용하였다. 고충실도 PfuUltra® II 퓨전(Fusion) HS DNA 중합효소(스트라타진(Stratagene); 미국 캘리포니아주 라 졸라 소재)를 PCR 반응에서 사용하였다. PCR을 하기 사이클링 조건들을 이용하여 수행하였다: 95℃, 2분; 95℃, 20초, 60℃, 20초 및 72℃, 1.5분의 30 사이클; 및 그 후 72℃, 3분. 생성된 PCR 생성물을 4℃에서 보관하였다. PCR 생성물을 퀴아퀵 PCR 정제 키트(퀴아젠)를 사용하여 정제하였다. 정제된 PCR 생성물을 람다 레드 재조합 절차에 따라 람다 재조합효소를 암호화하는 pKD46 플라스미드를 포함하는 PD03085 균주 내로 전기천공시켰다. 형질전환 혼합물을 25 µg/mL의 카나마이신을 포함하는 LB 플레이트 상에 플레이팅하였다. 카나마이신 내성 콜로니를 MOPS+10 g/L의 수크로스의 플레이트 상에서 점검하여 이들이 수크로스 상에서 성장할 수 없음을 확실하게 하였다. CscB의 잔기 61과 잔기 353 사이에 카나마이신 내성 카세트가 삽입된 것을 cscB 5' 프라이머(서열 번호 138) 및 cscB 3' 프라이머(서열 번호 139)를 사용한 PCR에 의해 확인하였다. 얻어진 FM5 yihP:cscA+K+B⁻(Δ61-353, kanR) 균주를 PD03513으로 표기하였다.
- [0232] 추정된 수크로스 수송체 유전자를 함유하는 에스케리키아 콜라이 균주의 구축
- [0233] 추정된 수크로스 수송체 유전자를 함유하는 에스케리키아 콜라이 균주를 하기 기재되는 벡터 pDMWP3 및 pDMWP4로부터 구축된 벡터 pDMWP12를 사용하여 구축하였다.
- [0234] 벡터 pDMWP3을 인테그레이티드 디엔에이 테크놀로지스, 인코포레이티드(Integrated DNA Technologies, Inc.) (미국 아이오와주 코랄빌 소재)로부터 얻었다. pDMWP3 벡터를, 인테그레이티드 디엔에이 테크놀로지스, 인코포레이티드에 의해 합성된 프로모터/MCS/이중 터미네이터 영역(서열 번호 140에 기재된 서열)을 pIDT-SMART 벡터(인테그레이티드 디엔에이 테크놀로지스, 인코포레이티드)로 클로닝시킴으로써 구축하였다.
- [0235] 벡터 pDMWP4를 플라스미드 pBR322로부터 구축하였다. TetR 유전자의 5' 말단 상의 scaI 부위 및 kpnI 부위 및

TetR 유전자의 3' 말단에서의 추가의 kpn1 부위를 플라스미드 pBR322로 도입시켰다. 추가로, kpn1 부위를 AmpR 유전자의 중간으로부터 제거하였다. 모든 제한 부위를 제조처의 프로토콜에 따라, 스트라타진의 퀵체인지(QuikChange) 위치-지정 돌연변이유발 키트(미국 캘리포니아주 라 줄라 소재의 스트라타진)를 사용하여 부가하거나 제거하였다.

[0236] 벡터 pDMWP3을 EcoR1 및 Kpn1으로 분해하고, 얻어진 438 bp 단편을, EcoR1 및 Kpn1으로 분해된 벡터 pDMWP4로 클로닝하여, 본 명세서에서 pBR*P1.20으로 지칭되는 벡터 pDMWP12를 제공하였다.

[0237] 다양한 추정된 수크로스 수송체 유전자를 에스케리키아 콜라이에서의 발현을 위해 코돈 최적화시키고, 코돈 최적화된 서열(서열 번호 141 내지 173에 기재)을 진스크립트 유에스에이 인코포레이티드(GenScript USA Inc.)(미국 뉴저지주 피스카타웨이 소재)에 의해 합성하였다. 합성 유전자를 HindIII 및 XmaI의 제한 부위에서 벡터 pDMWP12로 서브클로닝하였다. 벡터 내의 수송체 유전자의 존재를 서열 분석에 의해 확인하였다. 얻어진 벡터를 사용하여, PD03513의 일렉트로-컴피턴트(electro-competent) 세포를 형질전환시켰다. 콜로니를 LB-앰피실린 아가 플레이트로부터 선택하였다. 얻어진 균주를 PD050XX 시리즈로 표기하였다. 수크로스 수송체 유전체 인서트가 없는 대조군 플라스미드 pDMWP12도 또한 PD03513으로 형질전환시키고, 얻어진 균주를 PD03247로 표기하였다.

[0238] 수크로스 함유 배지에서의 다양한 수크로스 수송체를 함유하는 에스케리키아 콜라이 균주의 성장률을 바이오스크린(Bioscreen) C 성장 챔버(핀란드 헬싱키 소재의 바이오스크린(Bioscreen))를 사용하여 결정하였다. 각각의 균주를 16시간 동안 250 rpm에서 진탕시키면서, 33℃에서 100 mg/L의 카르베니실린을 함유하는 2 mL의 LB 배지에서 성장시켰다. 밤샘 배양물을, 바이오스크린 허니콤(honeycomb) 플레이트에서, 1XMOPS 완충제(미국 캘리포니아주 홀리스터 소재의 테크노바), 1.32 mM KH₂PO₄(테크노바)를 함유하고, 2, 5 또는 10 g/L의 수크로스 및 100 mg/L의 카르베니실린이 보충된 MOPS 최소 배지 150 μL에 1:100 희석하였다. 3회의 반복 실험을 각 시료에 대해 수행하였다. 성장 배지 블랭크(blank) 웰도 또한 포함시켰다. 허니콤 플레이트를 제조업자의 지시에 따라 바이오스크린 C 기기 내에 넣었다. 플레이트를 일정하게 진탕시키면서 33℃에서 인큐베이션시키고, OD를 600 nm에서 15분마다 기록하였다. 플라스미드 pDMWP12를 함유하는 에스케리키아 콜라이 균주 PD03247을 동일한 방식으로 성장시켜, 대조군으로 삼았다. 본 명세서에서 $\mu_{\max}$ 로 칭해지는 최대 성장률을 하기 절차를 이용하여 추정하였다. 먼저, 블랭크가 아닌 웰의 OD 값들로부터 블랭크 웰에서의 평균 OD 값들을 차감함으로써 백그라운드를 제거하였다. 그 후, 8개의 데이터점을 비선형 회귀를 이용하여 지수 곡선에 피팅(fitting)시킴으로써 8개의 데이터점(2시간의 성장을 커버함)으로 이루어진 슬라이딩 윈도우(sliding window)를 이용하여 성장률 파라미터를 추정하였다. 각 슬라이딩 윈도우에서, 피팅이 우수한 경우에만(즉, $R^2 > 0.95$), 추정된 성장률을 기록하였다. 모든 기록된 성장률로부터의 가장 큰 값을 $\mu_{\max}$ 로 취하였다. 3회의 반복 실험의 평균 $\mu_{\max}$ 및 표준 편차가 표 3에 제공되어 있다.

[0239] 표 3의 데이터로부터 알 수 있는 바와 같이, 실시예 2 내지 15의 추정된 수크로스 수송체 유전자를 함유하는 에스케리키아 콜라이 균주는 수크로스 수송체 유전자를 함유하지 않는 균주(비교예 1)에 비하여 수크로스 상에서 상당히 증가된 성장을 나타내었으며, 이에 의해, 이들 추정된 수송체의 기능을 입증하였다. 비교예 16 내지 33의 수크로스 수송체는 수크로스 상에서 성장을 거의 나타내지 않거나, 성장을 나타내지 않았으며, 이는 이들 추정된 수송체가 효율적인 수크로스 수송체로서 기능하지 않았음을 나타낸다.

표 3

수크로스에서 성장한 에스케리키아 콜라이 균주의 성장률

실시에	균주	수송체 유전자 (코돈 최적화된 CDS)	수크로스 상의 성장률(μ_{max})		
			2 g/L	5 g/L	10 g/L
비교예 1	PDO3247	없음	0.023 ± 0.0038	0.029 ± 0.0028	0.038 ± 0.001
2	PDO5001	<i>scrT1</i> (서열 번호 141)	0.411 ± 0.01	0.411 ± 0.002	0.399 ± 0.001
3	PDO5003	<i>scrT3</i> (서열 번호 142)	0.321 ± 0.015	0.390 ± 0.005	0.385 ± 0.005
4	PDO5004	<i>scrT4</i> (서열 번호 143)	0.239 ± 0.022	0.390 ± 0.001	0.390 ± 0.001
5	PDO5005	<i>scrT5</i> (서열 번호 144)	0.065 ± 0.015	0.127 ± 0.0002	0.176 ± 0.001
6	PDO5006	<i>scrT6</i> (서열 번호 145)	0.353 ± 0.0001	0.385 ± 0.006	0.382 ± 0.005
7	PDO5007	<i>scrT7</i> (서열 번호 146)	0.426 ± 0.002	0.403 ± 0.003	0.399 ± 0.005
8	PDO5008	<i>scrT8</i> (서열 번호 147)	0.072 ± 0.002	0.154 ± 0.0006	0.222 ± 0.001
9	PDO5009	<i>scrT9</i> (서열 번호 148)	0.251 ± 0.003	0.411 ± 0.0007	0.415 ± 0.003
10	PDO5012	<i>scrT12</i> (서열 번호 149)	0.028 ± 0.001	0.050 ± 0.003	0.103 ± 0.006
11	PDO5013	<i>scrT13</i> (서열 번호 150)	0.356 ± 0.002	0.359 ± 0.002	0.409 ± 0.005
12	PDO5014	<i>scrT14</i> (서열 번호 151)	0.238 ± 0.003	0.391 ± 0.013	0.387 ± 0.004
13	PDO5021	<i>scrT21</i> (서열 번호 152)	0.06 ± 0.0006	0.155 ± 0.002	0.271 ± 0.008
14	PDO5025	<i>scrT25</i> (서열 번호 153)	0.355 ± 0.0246	0.392 ± 0.0004	0.383 ± 0.004
15	PDO5026	<i>scrT26</i> (서열 번호 154)	0.034 ± 0.004	0.262 ± 0.001	0.287 ± 0.005
비교예 16	PDO5017	<i>scrT17</i> (서열 번호 155)	0.024 ± 0.003	0.032 ± 0.003	0.048 ± 0.002
비교예 17	PDO5018	<i>scrT18</i> (서열 번호 156)	0.047 ± 0.02	0.08 ± 0.01	0.086 ± 0.002
비교예 18	PDO5019	<i>scrT19</i> (서열 번호 157)	0.034 ± 0.01	0.06 ± 0.006	0.068 ± 0.008
비교예 19	PDO5020	<i>scrT20</i> (서열 번호 158)	0.028 ± 0.003	0.047 ± 0.01	0.052 ± 0.006
비교예 20	PDO5022	<i>scrT22</i> (서열 번호 159)	0.042 ± 0.003	0.044 ± 0.001	0.062 ± 0.003
비교예 21	PDO5023	<i>scrT23</i> (서열 번호 160)	0.029 ± 0.003	0.05 ± 0.001	0.075 ± 0.002
비교예 22	PDO5002	<i>scrT2</i> (서열 번호 161)	성장 부재	성장 부재	성장 부재
비교예 23	PDO5010	<i>scrT10</i> (서열 번호 162)	성장 부재	성장 부재	성장 부재
비교예 24	PDO5011	<i>scrT11</i> (서열 번호 163)	성장 부재	성장 부재	성장 부재
비교예 25	PDO5015	<i>scrT15</i> (서열 번호 164)	성장 부재	성장 부재	성장 부재
비교예 26	PDO5016	<i>scrT16</i> (서열 번호 165)	성장 부재	성장 부재	성장 부재
비교예 27	PDO5024	<i>scrT24</i> (서열 번호 166)	성장 부재	성장 부재	성장 부재
비교예 28	PDO5027	<i>scrT27</i> (서열 번호 167)	성장 부재	성장 부재	성장 부재
비교예 29	PDO5028	<i>scrT28</i> (서열 번호 168)	성장 부재	성장 부재	성장 부재
비교예 30	PDO5029	<i>scrT29</i> (서열 번호 169)	성장 부재	성장 부재	성장 부재
비교예 31	PDO5030	<i>scrT30</i> (서열 번호 170)	성장 부재	성장 부재	성장 부재
비교예 32	PDO5031	<i>scrT31</i> (서열 번호 171)	성장 부재	성장 부재	성장 부재
비교예 33	PDO5032	<i>scrT32</i> (서열 번호 172)	성장 부재	성장 부재	성장 부재
비교예 34	PDO5033	<i>scrT33</i> (서열 번호 173)	성장 부재	성장 부재	성장 부재

[0240]

[0241] 실시예 35 내지 52

[0242] 수크로스 수송체 유전자를 함유하는 재조합 에스케리키아 콜라이 균주

[0243] 이들 실시예의 목적은 본 명세서에 개시된 수크로스 수송체를 함유하며, 수크로스로부터 글리세롤 및 1,3-프로판다이올(PDO)을 생성하는 능력을 갖는 재조합 에스케리키아 콜라이 균주를 구축하는 것이었다.

[0244] 수크로스 수송체 유전자를 함유하는 에스케리키아 콜라이 균주를 PDO 생성 균주 TTab pSYCO400/AGRO로부터 구축하였다. 에스케리키아 콜라이 균주 TTab pSYCO400/AGRO, PTS 마이너스 균주를 하기과 같이 구축하였다. 미국 특허 제7,371,558호(실시예 17)에 기술된 균주 TT aldA로부터 aldB 유전자를 결실시킴으로써 균주 TTab를 생성시켰다. 간략하게는, 먼저 에스케리키아 콜라이 균주 MG1655에서 aldB의 1.5 kbp의 암호화 영역을 pKD3 플라스미드의 FRT-CmR-FRT 카세트(문헌[Datsenko and Wanner, Proc. Natl. Acad. Sci. USA 97:6640-6645, 2000])로 대체함으로써 aldB 결실을 일으켰다. pKD3을 주형으로 사용하여 프라이머 쌍 서열 번호 174 및 서열 번호 175로 대체 카세트를 증폭하였다. 프라이머 서열 번호 174는 aldB의 5'-말단에 대한 80 bp의 상동성, 및 pKD3에 대한 20 bp의 상동성을 포함한다. 프라이머 서열 번호 175는 aldB의 3' 말단에 대한 80 bp의 상동성, 및 pKD3에 대한 20 bp의 상동성을 포함한다. PCR 생성물을 겔-정제하고, MG1655/pKD46 컴피턴트 세포(미국 특허 제7,371,558호)로 전기천공시켰다. 재조합 균주를 12.5 mg/L의 클로람페니콜을 이용하여 LB(투리아 베르타니) 플레이트 상에서 선택하였다. 프라이머 쌍 서열 번호 176 및 서열 번호 177을 사용하여, PCR에 의해 aldB 유전자의 결실을 확인하였다. 야생형 균주는 1.5 kbp PCR 생성물을 제공한 반면에, 재조합 균주는 특징적인 1.1 kbp PCR 생성물을 제공하였다. P1 용해물을 제조하고, 이를 사용하여, 돌연변이를 TT aldA 균주로 옮겨, TT aldAΔaldB::Cm 균주를 형성하였다. 돌연변이가 존재함을 확인하기 위하여, 프라이머 쌍 서열 번호 176 및 서열 번호 177을 이용하는 계층 PCR에 의해 클로람페니콜-내성 클론을 확인하였다. FLP 재조합효소(상기 문헌[Datsenko and Wanner])를 이용하여 클로람페니콜 내성 마커를 제거하여, TTab를 생성하였다. 이어서, 균주 TTab를 미국

특허 제7,524,660호(실시예 4)에 기재된, pSYCO400/AGRO(서열 번호 66에 기재됨)로 형질전환시키고, 균주 TTab pSYCO400/AGRO를 생성하였다.

- [0245] 인용된 참고문헌에 기재된 바와 같이, 균주 TTab는 하기 변형을 포함한 에스케리키아 콜라이 균주 FM5(ATCC® 번호 53911)의 유도체이다:
- [0246] *glpK*, *gldA*, *ptsHI*, *crr*, *edd*, *arcA*, *mgsA*, *qor*, *ackA*, *pta*, *aldA* 및 *aldB* 유전자의 결실;
- [0247] *galP*, *glk*, *btuR*, *ppc*, 및 *yqhD* 유전자의 상향 조절; 및
- [0248] *gapA* 유전자의 하향 조절.
- [0249] 플라스미드 pSYCO400/AGRO는 글리세롤 생성 경로를 암호화하는 유전자(DAR1 및 GPP2) 및 글리세롤 탈수효소 및 관련 재활성화 인자를 암호화하는 유전자(*dhaB123*, *dhaX*, *orfX*, *orfY*), 및 프록토키나제를 암호화하는 유전자(*scrK*)를 포함한다.
- [0250] 균주 TTab pSYCO400/AGRO를 P1 형질도입을 위한 수용자로서 사용하였다. 공여자 균주는 실시예 1 내지 34에 기재된 바와 같이 구축된 PDO3513이었으며, 성장에 대한 선택은 25 µg/mL 카나마이신이 있는 LB 플레이트에서 이루어졌다. 카나마이신 및 스펙티노마이신에 대해 내성인 콜로니를 정제하고, PDO2737 [TTab/pSYCO400AGRO *yihP::cscKB Δ* (61-353)KanR&A]로 표기하였다.
- [0251] 플라스미드 pDMWP12 내의 다양한 추정된 수크로스 수송체를 표 4에 나타낸 바와 같이 전기천공법에 의해 균주 PDO2737로 형질전환시키고, 콜로니를 100 µg/mL 카르베닐린 및 100 µg/mL 스펙티노마이신을 함유하는 LB 저염 아가 플레이트 상에서 선택하였다. 각 수송체 유전자에 대하여 하나의 콜로니를 선택하여, 하기 실시예 53에 기재된 바와 같이 탄소원으로 수크로스를 사용하여 PDO 생성을 시험하였다.

표 4

수크로스 수송체를 함유하는 재조합 PDO 생성 에스케리키아 콜라이 균주

실시예	균주	수크로스 수송체
비교예 35	PDO5134	없음
36	PDO5101	<i>scrT1</i> (서열 번호 141)
37	PDO5103	<i>scrT3</i> (서열 번호 142)
38	PDO5104	<i>scrT4</i> (서열 번호 143)
39	PDO5105	<i>scrT5</i> (서열 번호 144)
40	PDO5106	<i>scrT6</i> (서열 번호 145)
41	PDO5107	<i>scrT7</i> (서열 번호 146)
42	PDO5108	<i>scrT8</i> (서열 번호 147)
43	PDO5109	<i>scrT9</i> (서열 번호 148)
44	PDO5112	<i>scrT12</i> (서열 번호 149)
45	PDO5113	<i>scrT13</i> (서열 번호 150)
46	PDO5114	<i>scrT14</i> (서열 번호 151)
비교예 47	PDO5119	<i>scrT19</i> (서열 번호 157)
48	PDO5121	<i>scrT21</i> (서열 번호 152)
비교예 49	PDO5122	<i>scrT22</i> (서열 번호 159)
50	PDO5123	<i>scrT23</i> (서열 번호 160)
51	PDO5125	<i>scrT25</i> (서열 번호 153)
52	PDO5126	<i>scrT26</i> (서열 번호 154)

- [0252]
- [0253] 실시예 53
- [0254] 수크로스 수송체 유전자를 함유하는 재조합 박테리아 균주에 의한 수크로스로부터의 글리세롤 및 1,3-프로판다이올의 생성
- [0255] 실시예 35 내지 52에 기재된 재조합 에스케리키아 콜라이 균주를 33°C에서 하룻밤 카르베닐린(100 µg/mL) 및 스펙티노마이신(100 µg/mL)을 함유하는 LB 저염 배지 1 mL에서 성장시켰다. 각 배양물의 200 µL 시료를 사용하여 10 g/L의 수크로스, 100 µg/mL의 스펙티노마이신 및 100 ng/mL의 비타민 B12가 있는 MOPS 최소 배지 12.5 mL에 접종하였다. 세포를 250 rpm에서 45시간(표 5의 균주) 또는 50시간(표 6의 균주) 동안 진탕시키면서 33°C에서 성장시켰다. 수크로스 수송체 유전자가 없는 대조군 균주 PDO5134를 각 배치에서 시험 균주와 함께 성장시켰다. 이어서, 배양물을 원심분리하고, 상청액을 0.22 µm 스피(Spin)-X 원심분리 튜브 필터(미국 뉴욕주 코닝 소재의 코닝 인코포레이티드(Corning Inc.))에 첨가하고, 10,000 g에서 1분 동안 원심분리하였다. 여액을 별도의 워터스(Waters) TCM 가열 챔버에서 85 °C로 가열한, 아미넥스(Aminex) HPX-87C HPLC 탄수화물 분석 컬

럼(미국 캘리포니아주 허큘리스 소재의 바이오-라드 레보라토리즈(Bio-Rad Laboratories), 카탈로그 번호 125-0095)을 갖춘 워터스 얼라이언스(Waters Alliance) 2690 HPLC 시스템(미국 매사추세츠주 밀포드 소재의 워터스 코퍼레이션(Waters Corp.))을 이용하여 HPLC에 의해 분석하였다. 바이오-라드 카보(carbo)-C 마이크로-가드(micro-guard) 컬럼(바이오-라드, 카탈로그 번호 125-0128)을 분석 컬럼 전에 사용하였다. 이동상은 0.05 mM CaO(시그마, #208159), 0.5 mM MES(시그마, #M3671), 0.05 mM HNO₃ (미국 뉴저지주 김스타운 소재의 이엠티 케미칼즈(EMD Chemicals), 카탈로그 번호 NX0409), pH 5.3으로 구성되었다. 유량은 0.5 mL/분이었다. 기지 농도로 용해된 상업적으로 구매한 순수 화합물로부터 얻은 표준 곡선을 기준으로 굴절률 검출에 의해 용출된 화합물을 정량화하였다.

표 5

수크로스로부터 PDO 및 글리세롤의 생성(45 시간의 발효)

균주	수송체 유전자	PDO(g/L)	글리세롤 (g/L)	잔여 수크로스 (g/L)
PDO5134	없음	0.01	0.33	10.82
PDO5101	<i>scrT1</i>	0.70	1.46	7.53
PDO5103	<i>scrT3</i>	0.83	0.77	7.83
PDO5104	<i>scrT4</i>	5.43	0.39	0.00
PDO5106	<i>scrT6</i>	3.68	1.32	0.00
PDO5107	<i>scrT7</i>	2.89	3.76	0.00
PDO5109	<i>scrT9</i>	3.48	3.10	0.00
PDO5113	<i>scrT13</i>	4.35	1.88	0.00
PDO5114	<i>scrT14</i>	4.71	1.34	0.00

[0256]

[0257]

표 5 및 6의 데이터로부터 알 수 있는 바와 같이, 수크로스 수송체 *scrT1*, *scrT3*, *scrT4*, *scrT5*, *scrT6*, *scrT7*, *scrT8*, *scrT9*, *scrT13*, *scrT14*, *scrT21*, *scrT23*, *scrT25* 및 *scrT26*을 함유하는 에스케리키아 콜라이 균주는 수크로스를 사용하고 PDO 및 글리세롤을 생성할 수 있었다. *scrT19*(비교예 47) 및 *scrT22*(비교예 49)를 함유하는 에스케리키아 콜라이 균주는 상당한 양으로 이들 생성물을 생성하기 위해 수크로스를 사용하지 않았다. *scrT12*가 수크로스 수송체인 것으로 나타났지만(실시예 10), *scrT12*를 함유하는 에스케리키아 콜라이 균주(PDO5112, 실시예 44)는 수크로스로부터 소량의 PDO 및 글리세롤을 생성하였다.

표 6

수크로스로부터의 PDO 및 글리세롤의 생성(50 시간의 발효)

균주	수송체 유전자	PDO(g/L)	글리세롤 (g/L)	잔여 수크로스 (g/L)
PDO5134	없음	0.07	0.01	10.96
PDO5105	<i>scrT5</i>	1.39	0.16	8.18
PDO5108	<i>scrT8</i>	1.14	0.23	8.68
PDO5112	<i>scrT12</i>	0.26	0.04	10.39
PDO5119	<i>scrT19</i>	0.12	0.03	10.73
PDO5121	<i>scrT21</i>	1.41	0.11	8.20
PDO5122	<i>scrT22</i>	0.18	0.01	10.41
PDO5123	<i>scrT23</i>	0.30	0.01	8.77
PDO5125	<i>scrT25</i>	4.05	0.25	3.09
PDO5126	<i>scrT26</i>	0.62	0.04	9.66

[0258]

서열 목록

SEQUENCE LISTING

<110> E.I. du Pont de Nemours and Co.

You, Zheng

<120> Recombinant Bacteria Having the Ability to Metabolize Sucrose

<130> CL5510 USNP

<160> 177

<170> PatentIn version 3.5

<210> 1

<211> 1176

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 1

atgtctgctg ctgctgatag attaaactta acttccggcc acttgaatgc tggtagaaag	60
agaagttcct cttctgtttc tttgaaggct gccgaaaagc ctttcaaggt tactgtgatt	120
ggatctggta actgggggtac tactattgcc aagggtggtt ccgaaaattg taagggatac	180
ccagaagttt tcgctccaat agtacaaatg tgggtgttcg aagaagagat caatggtgaa	240
aaattgactg aaatcataaa tactagacat caaaacgtga aatacttgcc tggcatcact	300
ctaccgaca atttggttgc taatccagac ttgattgatt cagtcaagga tgtcgacatc	360
atcgttttca acattccaca tcaatttttg ccccgatatc gtagccaatt gaaaggatcat	420
gttgattcac acgtcagagc tatctcctgt cttaaagggtt ttgaagttgg tgctaaaggt	480
gtccaattgc tatcctctta catcactgag gaactaggta ttcaatgtgg tgcctctatc	540
ggtgctaaca ttgccaccga agtcgctcaa gaacactggt ctgaaacaac agttgcttac	600
cacattccaa aggatttcag aggcgagggc aaggacgtcg accataaggt tctaaaggcc	660
ttgttccaca gaccttactt ccacgttagt gtcacgaag atgttgctgg tatctccatc	720
tgtgggtgctt tgaagaacgt tgttgcctta ggttgtggtt tcgtcgaagg tctaggctgg	780
ggtaacaacg cttctgctgc catccaaaga gtcggtttgg gtgagatcat cagattcggg	840
caaagtgttt tcccagaatc tagagaagaa acatactacc aagagtctgc tgggtgttgc	900
gatttgatca ccacctgcgc tgggtggtaga aacgtcaagg ttgctaggct aatggctact	960
tctggttaagg acgcctggga atgtgaaaag gagttgttga atggccaatc cgctcaaggt	1020
ttaattacct gcaaagaagt tcacgaatgg ttggaaacat gtggctctgt cgaagacttc	1080
ccattatttg aagccgtata ccaaactgtt tacaacaact acccaatgaa gaacctgccg	1140
gacatgattg aagaattaga tctacatgaa gattag	1176

<210> 2

<211> 391

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Ser Ala Ala Ala Asp Arg Leu Asn Leu Thr Ser Gly His Leu Asn
1 5 10 15
Ala Gly Arg Lys Arg Ser Ser Ser Ser Val Ser Leu Lys Ala Ala Glu
20 25 30
Lys Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr Thr
35 40 45
Ile Ala Lys Val Val Ala Glu Asn Cys Lys Gly Tyr Pro Glu Val Phe
50 55 60
Ala Pro Ile Val Gln Met Trp Val Phe Glu Glu Glu Ile Asn Gly Glu
65 70 75 80
Lys Leu Thr Glu Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr Leu
85 90 95
Pro Gly Ile Thr Leu Pro Asp Asn Leu Val Ala Asn Pro Asp Leu Ile
100 105 110
Asp Ser Val Lys Asp Val Asp Ile Ile Val Phe Asn Ile Pro His Gln
115 120 125
Phe Leu Pro Arg Ile Cys Ser Gln Leu Lys Gly His Val Asp Ser His
130 135 140
Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Val Gly Ala Lys Gly
145 150 155 160
Val Gln Leu Leu Ser Ser Tyr Ile Thr Glu Glu Leu Gly Ile Gln Cys
165 170 175
Gly Ala Leu Ser Gly Ala Asn Ile Ala Thr Glu Val Ala Gln Glu His
180 185 190
Trp Ser Glu Thr Thr Val Ala Tyr His Ile Pro Lys Asp Phe Arg Gly
195 200 205
Glu Gly Lys Asp Val Asp His Lys Val Leu Lys Ala Leu Phe His Arg
210 215 220
Pro Tyr Phe His Val Ser Val Ile Glu Asp Val Ala Gly Ile Ser Ile
225 230 235 240
Cys Gly Ala Leu Lys Asn Val Val Ala Leu Gly Cys Gly Phe Val Glu

245 250 255
 Gly Leu Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Val Gly
 260 265 270
 Leu Gly Glu Ile Ile Arg Phe Gly Gln Met Phe Phe Pro Glu Ser Arg
 275 280 285

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

Thr Cys Gly Ser Val Glu Asp Phe Pro Leu Phe Glu Ala Val Tyr Gln
 355 360 365
 Ile Val Tyr Asn Asn Tyr Pro Met Lys Asn Leu Pro Asp Met Ile Glu
 370 375 380
 Glu Leu Asp Leu His Glu Asp
 385 390

<210> 3

<211> 1323

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 3

atgcttgctg tcagaagatt aacaagatac acatttcctta agcgaacgca tccggtgtta 60
 tatactcgtc gtgcatataa aattttgcct tcaagatcta ctttcctaag aagatcatta 120

ttacaaacac aactgcactc aaagatgact gtcatacta atatcaaaca gcacaaacac 180
 tgtcatgagg accatcctat cagaagatcg gactctgccg tgtcaattgt acatttgaaa 240
 cgtgcgccct tcaaggttac agtgattggt tctggttaact gggggaccac catcgccaaa 300
 gtcatcgcg aaaacacaga attgcattcc catatcttcg agccagaggt gagaatgtgg 360
 gtttttgatg aaaagatcgg cgacgaaaat ctgacggata tcataaatac aagacaccag 420
 aacgttaaata atctacccaa tattgacctg ccccataatc tagtggccga tctgatctt 480

ttacactcca tcaaggggtgc tgacatcctt gttttcaaca tccctcatca atttttacca 540

aacatagtca aacaattgca aggccacgtg gccctcatg taagggccat ctctgtctta 600

aaagggttcg agttgggctc caagggtgtg caattgctat cctcctatgt tactgatgag 660

ttaggaatcc aatgtggcgc actatctggt gcaaacttgg caccggaagt ggccaaggag 720

cattggtccg aaaccaccgt ggcttaccaa ctaccaaagg attatcaagg tgatggcaag 780

gatgtagatc ataagatttt gaaattgctg ttccacagac cttacttcca cgtcaatgtc 840

atcgatgatg ttgctgggat atccattgcc ggtgccttga agaacgtcgt ggcacttgca 900

tgtggtttcg tagaaggatg gggatggggt aacaatgcct ccgcagccat tcaaaggctg 960

ggtttaggtg aaattatcaa gttcggtaga atgtttttcc cagaatccaa agtcgagacc 1020

tactatcaag aatccgctgg tgttcagat ctgatcacca cctgctcagg cggtagaaac 1080

gtcaaggttg ccacatacat ggccaagacc ggtaagtcag ccttgggaagc agaaaaggaa 1140

ttgcttaacg gtcaatccgc ccaagggata atcacatgca gagaagttca cgagtggcta 1200

caaacatgtg agttgaccca agaattccca ttattcgagg cagtctacca gatagtctac 1260

aacaacgtcc gcatggaaga cctaccggag atgattgaag agctagacat cgatgacgaa 1320

tag 1323

<210> 4

<211> 440

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 4

Met Leu Ala Val Arg Arg Leu Thr Arg Tyr Thr Phe Leu Lys Arg Thr

1 5 10 15

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

20 25 30

Ser Thr Phe Leu Arg Arg Ser Leu Leu Gln Thr Gln Leu His Ser Lys

35 40 45

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp

50 55 60

His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys

65 70 75 80

Arg Ala Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr

85 90 95
 Thr Ile Ala Lys Val Ile Ala Glu Asn Thr Glu Leu His Ser His Ile
 100 105 110
 Phe Glu Pro Glu Val Arg Met Trp Val Phe Asp Glu Lys Ile Gly Asp

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

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

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

 305 310 315 320
 Gly Leu Gly Glu Ile Ile Lys Phe Gly Arg Met Phe Phe Pro Glu Ser
 325 330 335

Lys Val Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile
340 345 350

Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala
355 360 365

Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly

370 375 380

Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu
385 390 395 400

Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Leu Phe Glu Ala Val Tyr
405 410 415

Gln Ile Val Tyr Asn Asn Val Arg Met Glu Asp Leu Pro Glu Met Ile
420 425 430

Glu Glu Leu Asp Ile Asp Asp Glu
435 440

<210> 5

<211> 816

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 5

atgaaacggt tcaatgtttt aaaatataac agaacaacaa aagcaaatat acaaaccatc 60
gcaatgcctt tgaccacaaa acctttatct ttgaaaatca acgccgctct attcgatgtt 120
gacggtacca tcatcatctc tcaaccagcc attgctgctt tctggagaga tticggtaaa 180
gacaagcctt acttcgatgc cgaacacggt attcacatct ctacacggttg gagaacttac 240
gatgccattg ccaagttcgc tccagacttt gctgatgaag aatacgttaa caagctagaa 300
ggtgaaatcc cagaaaagta cggatgaacac tccatcgaag ttccaggtgc tgtcaagttg 360

tgtaatgctt tgaacgcctt gccaaaggaa aaatgggctg tcgccacctc tggtagccgt 420
gacatggcca agaatgggtt cgacattttg aagatcaaga gaccagaata cttcatcacc 480
gccaatgatg tcaagcaagg taagcctcac ccagaacat acttaaaggg tagaaacggt 540
ttgggtttcc caattaatga acaagacca tccaaatcta aggttggtgt ctttgaagac 600
gcaccagctg gtattgtcgc tggtaaggct gctggctgta aaatcgttgg tattgctacc 660
actttcgatt tggacttctt gaaggaaaag ggttgtgaca tcattgtcaa gaaccacgaa 720

tctatcagag tcggtgaata caacgtgaa accgatgaag tcgaattgat ctttgatgac 780

tacttatacg ctaaggatga cttgttgaaa tggtaa 816

<210> 6

<211> 271

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 6

Met Lys Arg Phe Asn Val Leu Lys Tyr Ile Arg Thr Thr Lys Ala Asn

1 5 10 15

Ile Gln Thr Ile Ala Met Pro Leu Thr Thr Lys Pro Leu Ser Leu Lys

20 25 30

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

35 40 45

Pro Ala Ile Ala Ala Phe Trp Arg Asp Phe Gly Lys Asp Lys Pro Tyr

50 55 60

Phe Asp Ala Glu His Val Ile His Ile Ser His Gly Trp Arg Thr Tyr

65 70 75 80

Asp Ala Ile Ala Lys Phe Ala Pro Asp Phe Ala Asp Glu Glu Tyr Val

85 90 95

Asn Lys Leu Glu Gly Glu Ile Pro Glu Lys Tyr Gly Glu His Ser Ile

100 105 110

Glu Val Pro Gly Ala Val Lys Leu Cys Asn Ala Leu Asn Ala Leu Pro

115 120 125

Lys Glu Lys Trp Ala Val Ala Thr Ser Gly Thr Arg Asp Met Ala Lys

130 135 140

Lys Trp Phe Asp Ile Leu Lys Ile Lys Arg Pro Glu Tyr Phe Ile Thr

145 150 155 160

Ala Asn Asp Val Lys Gln Gly Lys Pro His Pro Glu Pro Tyr Leu Lys

165 170 175

Gly Arg Asn Gly Leu Gly Phe Pro Ile Asn Glu Gln Asp Pro Ser Lys

180 185 190

Ser Lys Val Val Val Phe Glu Asp Ala Pro Ala Gly Ile Ala Ala Gly
 195 200 205
 Lys Ala Ala Gly Cys Lys Ile Val Gly Ile Ala Thr Thr Phe Asp Leu
 210 215 220
 Asp Phe Leu Lys Glu Lys Gly Cys Asp Ile Ile Val Lys Asn His Glu
 225 230 235 240

Ser Ile Arg Val Gly Glu Tyr Asn Ala Glu Thr Asp Glu Val Glu Leu
 245 250 255
 Ile Phe Asp Asp Tyr Leu Tyr Ala Lys Asp Asp Leu Leu Lys Trp
 260 265 270

<210> 7

<211> 753

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 7

atgggattga ctactaaacc tctatctttg aaagttaacg ccgctttgtt cgacgtcgac	60
ggtagcatta tcatctctca accagccatt gctgcattct ggagggattt cggtaaggac	120
aaaccttatt tcgatgctga acacgttatc caagtctcgc atggttggag aacgtttgat	180
gccattgcta agttcgctcc agactttgcc aatgaagagt atgttaacaa attagaagct	240
gaaattccgg tcaagtaacg tgaaaaatcc attgaagtcc caggtgcagt taagctgtgc	300
aacgctttga acgctctacc aaaagagaaa tgggctgtgg caacttccgg taccctgat	360
atggcacaaa aatggttcga gcatctggga atcaggagac caaagtactt cattaccgct	420
aatgatgtca aacagggtaa gcctcatcca gaaccatata tgaagggcag gaatggctta	480
ggatatccga tcaatgagca agacccttcc aaatctaagg tagtagtatt tgaagacgct	540
ccagcaggta ttgccgccgg aaaagccgcc ggttgtaaga tcattggtat tgccactact	600
ttcgacttgg acttcctaaa ggaaaaaggc tgtgacatca ttgtcaaaaa ccacgaatcc	660
atcagagttg gcggctacaa tgccgaaaca gacgaagttg aattcatttt tgacgactac	720
ttatatgcta aggacgatct gttgaaatgg taa	753

<210> 8

<211> 250

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 8

Met Gly Leu Thr Thr Lys Pro Leu Ser Leu Lys Val Asn Ala Ala Leu
1 5 10 15

Phe Asp Val Asp Gly Thr Ile Ile Ile Ser Gln Pro Ala Ile Ala Ala
20 25 30

Phe Trp Arg Asp Phe Gly Lys Asp Lys Pro Tyr Phe Asp Ala Glu His
35 40 45

Val Ile Gln Val Ser His Gly Trp Arg Thr Phe Asp Ala Ile Ala Lys
50 55 60

Phe Ala Pro Asp Phe Ala Asn Glu Glu Tyr Val Asn Lys Leu Glu Ala
65 70 75 80

Glu Ile Pro Val Lys Tyr Gly Glu Lys Ser Ile Glu Val Pro Gly Ala
85 90 95

Val Lys Leu Cys Asn Ala Leu Asn Ala Leu Pro Lys Glu Lys Trp Ala
100 105 110

Val Ala Thr Ser Gly Thr Arg Asp Met Ala Gln Lys Trp Phe Glu His
115 120 125

Leu Gly Ile Arg Arg Pro Lys Tyr Phe Ile Thr Ala Asn Asp Val Lys
130 135 140

Gln Gly Lys Pro His Pro Glu Pro Tyr Leu Lys Gly Arg Asn Gly Leu
145 150 155 160

Gly Tyr Pro Ile Asn Glu Gln Asp Pro Ser Lys Ser Lys Val Val Val
165 170 175

Phe Glu Asp Ala Pro Ala Gly Ile Ala Ala Gly Lys Ala Ala Gly Cys
180 185 190

Lys Ile Ile Gly Ile Ala Thr Thr Phe Asp Leu Asp Phe Leu Lys Glu
195 200 205

Lys Gly Cys Asp Ile Ile Val Lys Asn His Glu Ser Ile Arg Val Gly
210 215 220

Gly Tyr Asn Ala Glu Thr Asp Glu Val Glu Phe Ile Phe Asp Asp Tyr
225 230 235 240

Leu Tyr Ala Lys Asp Asp Leu Leu Lys Trp

245 250

<210> 9

<211> 1668

<212> DNA

<213> *Klebsiella pneumoniae*

<220><221> CDS

<222> (1)..(1668)

<400> 9

atg aaa aga tca aaa cga ttt gca gta ctg gcc cag cgc ccc gtc aat 48

Met Lys Arg Ser Lys Arg Phe Ala Val Leu Ala Gln Arg Pro Val Asn

1 5 10 15

cag gac ggg ctg att ggc gag tgg cct gaa gag ggg ctg atc gcc atg 96

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met

20 25 30

gac agc ccc ttt gac ccg gtc tct tca gta aaa gtg gac aac ggt ctg 144

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

35 40 45

atc gtc gaa ctg gac ggc aaa cgc cgg gac cag ttt gac atg atc gac 192

Ile Val Glu Leu Asp Gly Lys Arg Arg Asp Gln Phe Asp Met Ile Asp

50 55 60

cga ttt atc gcc gat tac gcg atc aac gtt gag cgc aca gag cag gca 240

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Glu Arg Thr Glu Gln Ala

65 70 75 80

atg cgc ctg gag gcg gtg gaa ata gcc cgt atg ctg gtg gat att cac 288

Met Arg Leu Glu Ala Val Glu Ile Ala Arg Met Leu Val Asp Ile His

85 90 95

gtc agc cgg gag gag atc att gcc atc act acc gcc atc acg ccg gcc 336

Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala

100 105 110

aaa gcg gtc gag gtg atg gcg cag atg aac gtg gtg gag atg atg atg 384

Lys Ala Val Glu Val Met Ala Gln Met Asn Val Val Glu Met Met Met

115	120	125	
gcg ctg cag aag atg cgt gcc cgc cgg acc ccc tcc aac cag tgc cac			432
Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His			
130	135	140	
gtc acc aat ctc aaa gat aat ccg gtg cag att gcc gct gac gcc gcc			480
Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala			
145	150	155	160
gag gcc ggg atc cgc ggc ttc tca gaa cag gag acc acg gtc ggt atc			528
Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile			
165	170	175	
gcg cgc tac gcg ccg ttt aac gcc ctg gcg ctg ttg gtc ggt tcg cag			576
Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln			
180	185	190	
tgc ggc cgc ccc ggc gtg ttg acg cag tgc tcg gtg gaa gag gcc acc			624
Cys Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr			
195	200	205	
gag ctg gag ctg ggc atg cgt ggc tta acc agc tac gcc gag acg gtg			672
Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val			
210	215	220	
tcg gtc tac ggc acc gaa gcg gta ttt acc gac ggc gat gat acg ccg			720
Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro			
225	230	235	240
tgg tca aag gcg ttc ctc gcc tcg gcc tac gcc tcc cgc ggg ttg aaa			768
Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys			
245	250	255	
atg cgc tac acc tcc ggc acc gga tcc gaa gcg ctg atg ggc tat tcg			816
Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser			
260	265	270	
gag agc aag tcg atg ctc tac ctc gaa tcg cgc tgc atc ttc att act			864

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr	
275 280 285	
aaa ggc gcc ggg gtt cag gga ctg caa aac ggc gcg gtg agc tgt atc	912
Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile	
290 295 300	
ggc atg acc ggc gct gtg ccg tcg ggc att cgg gcg gtg ctg gcg gaa	960
Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu	
305 310 315 320	
aac ctg atc gcc tct atg ctc gac ctc gaa gtg gcg tcc gcc aac gac	1008
Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp	
325 330 335	
cag act ttc tcc cac tcg gat att cgc cgc acc gcg cgc acc ctg atg	1056
Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met	
340 345 350	
cag atg ctg ccg ggc acc gac ttt att ttc tcc ggc tac agc gcg gtg	1104
Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val	
355 360 365	
ccg aac tac gac aac atg ttc gcc ggc tcg aac ttc gat gcg gaa gat	1152
Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp	
370 375 380	
ttt gat gat tac aac atc ctg cag cgt gac ctg atg gtt gac ggc ggc	1200
Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly	
385 390 395 400	
ctg cgt ccg gtg acc gag gcg gaa acc att gcc att cgc cag aaa gcg	1248
Leu Arg Pro Val Thr Glu Ala Glu Thr Ile Ala Ile Arg Gln Lys Ala	
405 410 415	
gcg cgg gcg atc cag gcg gtt ttc cgc gag ctg ggg ctg ccg cca atc	1296
Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Pro Ile	
420 425 430	
gcc gac gag gag gtg gag gcc gcc acc tac gcg cac ggc agc aac gag	1344
Ala Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Asn Glu	

435	440	445	
atg ccg ccg cgt aac gtg gtg gag gat ctg agt gcg gtg gaa gag atg			1392
Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met			
450	455	460	
atg aag cgc aac atc acc ggc ctc gat att gtc ggc gcg ctg agc cgc			1440
Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Ser Arg			
465	470	475	480
agc ggc ttt gag gat atc gcc agc aat att ctc aat atg ctg cgc cag			1488
Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Arg Gln			
485	490	495	
cgg gtc acc ggc gat tac ctg cag acc tcg gcc att ctc gat cgg cag			1536
Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln			
500	505	510	
ttc gag gtg gtg agt gcg gtc aac gac atc aat gac tat cag ggg ccg			1584
Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro			
515	520	525	
ggc acc ggc tat cgc atc tct gcc gaa cgc tgg gcg gag atc aaa aat			1632
Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn			
530	535	540	
att ccg ggc gtg gtt cag ccc gac acc att gaa taa			1668
Ile Pro Gly Val Val Gln Pro Asp Thr Ile Glu			
545	550	555	
<210> 10			
<211> 555			
<212> PRT			
<213> Klebsiella pneumoniae			
<400> 10			
Met Lys Arg Ser Lys Arg Phe Ala Val Leu Ala Gln Arg Pro Val Asn			
1	5	10	15
Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met			
20	25	30	

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

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

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

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

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr

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

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

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

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

Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540
Ile Pro Gly Val Val Gln Pro Asp Thr Ile Glu
545 550 555
<210> 11
<211> 585
<212> DNA
<213> *Klebsiella pneumoniae*
<220><221> CDS
<222> (1)..(585)
<400> 11
gtg caa cag aca acc caa att cag ccc tct ttt acc ctg aaa acc cgc 48
Val Gln Gln Thr Thr Gln Ile Gln Pro Ser Phe Thr Leu Lys Thr Arg
1 5 10 15
gag ggc ggg gta gct tct gcc gat gaa cgc gcc gat gaa gtg gtg atc 96
Glu Gly Gly Val Ala Ser Ala Asp Glu Arg Ala Asp Glu Val Val Ile
20 25 30
ggc gtc ggc cct gcc ttc gat aaa cac cag cat cac act ctg atc gat 144
Gly Val Gly Pro Ala Phe Asp Lys His Gln His His Thr Leu Ile Asp
35 40 45
atg ccc cat ggc gcg atc ctc aaa gag ctg att gcc ggg gtg gaa gaa 192
Met Pro His Gly Ala Ile Leu Lys Glu Leu Ile Ala Gly Val Glu Glu
50 55 60
gag ggg ctt cac gcc cgg gtg gtg cgc att ctg cgc acg tcc gac gtc 240
Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80
tcc ttt atg gcc tgg gat gcg gcc aac ctg agc ggc tcg ggg atc ggc 288
Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95
atc ggt atc cag tcg aag ggg acc acg gtc atc cat cag cgc gat ctg 336
Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

ctg ccg ctc agc aac ctg gag ctg ttc tcc cag gcg ccg ctg ctg acg 384
 Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
 115 120 125

ctg gag acc tac cgg cag att ggc aaa aac gct gcg cgc tat gcg cgc 432
 Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
 130 135 140

aaa gag tca cct tcg ccg gtg ccg gtg gtg aac gat cag atg gtg cgg 480
 Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
 145 150 155 160
 ccg aaa ttt atg gcc aaa gcc gcg cta ttt cat atc aaa gag acc aaa 528

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
 165 170 175

cat gtg gtg cag gac gcc gag ccc gtc acc ctg cac atc gac tta gta 576
 His Val Val Gln Asp Ala Glu Pro Val Thr Leu His Ile Asp Leu Val
 180 185 190

agg gag tga 585
 Arg Glu

<210> 12

<211> 194

<212> PRT

<213> Klebsiella pneumoniae

<400> 12

Val Gln Gln Thr Thr Gln Ile Gln Pro Ser Phe Thr Leu Lys Thr Arg
 1 5 10 15

Glu Gly Gly Val Ala Ser Ala Asp Glu Arg Ala Asp Glu Val Val Ile
 20 25 30

Gly Val Gly Pro Ala Phe Asp Lys His Gln His His Thr Leu Ile Asp
 35 40 45

Met Pro His Gly Ala Ile Leu Lys Glu Leu Ile Ala Gly Val Glu Glu
 50 55 60

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

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

His Val Val Gln Asp Ala Glu Pro Val Thr Leu His Ile Asp Leu Val
180 185 190
Arg Glu

<210> 13

<211> 426

<212> DNA

<213> Klebsiella pneumoniae

<220><221> CDS

<222> (1)..(426)

<400> 13

atg agc gag aaa acc atg cgc gtg cag gat tat ccg tta gcc acc cgc 48

Met Ser Glu Lys Thr Met Arg Val Gln Asp Tyr Pro Leu Ala Thr Arg

1 5 10 15

tgc ccg gag cat atc ctg acg cct acc ggc aaa cca ttg acc gat att 96

Cys Pro Glu His Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile

20 25 30

acc ctc gag aag gtg ctc tct ggc gag gtg ggc ccg cag gat gtg cgg 144

Thr Leu Glu Lys Val Leu Ser Gly Glu Val Gly Pro Gln Asp Val Arg
35 40 45
atc tcc cgc cag acc ctt gag tac cag gcg cag att gcc gag cag atg 192
Ile Ser Arg Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
50 55 60
cag cgc cat gcg gtg gcg cgc aat ttc cgc cgc gcg gcg gag ctt atc 240
Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80
gcc att cct gac gag cgc att ctg gct atc tat aac gcg ctg cgc ccg 288
Ala Ile Pro Asp Glu Arg Ile Leu Ala Ile Tyr Asn Ala Leu Arg Pro
85 90 95
ttc cgc tcc tcg cag gcg gag ctg ctg gcg atc gcc gac gag ctg gag 336
Phe Arg Ser Ser Gln Ala Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu
100 105 110
cac acc tgg cat gcg aca gtg aat gcc gcc ttt gtc cgg gag tcg gcg 384
His Thr Trp His Ala Thr Val Asn Ala Ala Phe Val Arg Glu Ser Ala
115 120 125
gaa gtg tat cag cag cgg cat aag ctg cgt aaa gga agc taa 426

Glu Val Tyr Gln Gln Arg His Lys Leu Arg Lys Gly Ser
130 135 140
<210> 14
<211> 141
<212> PRT
<213> Klebsiella pneumoniae
<400> 14
Met Ser Glu Lys Thr Met Arg Val Gln Asp Tyr Pro Leu Ala Thr Arg
1 5 10 15
Cys Pro Glu His Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
20 25 30
Thr Leu Glu Lys Val Leu Ser Gly Glu Val Gly Pro Gln Asp Val Arg

35 40 45

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

100 105 110
 His Thr Trp His Ala Thr Val Asn Ala Ala Phe Val Arg Glu Ser Ala
 115 120 125
 Glu Val Tyr Gln Gln Arg His Lys Leu Arg Lys Gly Ser
 130 135 140

<210> 15
 <211> 1539
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(1539)
 <400> 15

atg acc aat aat ccc cct tca gca cag att aag ccc ggc gag tat ggt 48
 Met Thr Asn Asn Pro Pro Ser Ala Gln Ile Lys Pro Gly Glu Tyr Gly

1 5 10 15
 ttc ccc ctc aag tta aaa gcc cgc tat gac aac ttt att ggc ggc gaa 96
 Phe Pro Leu Lys Leu Lys Ala Arg Tyr Asp Asn Phe Ile Gly Gly Glu
 20 25 30
 tgg gta gcc cct gcc gac ggc gag tat tac cag aat ctg acg ccg gtg 144
 Trp Val Ala Pro Ala Asp Gly Glu Tyr Tyr Gln Asn Leu Thr Pro Val
 35 40 45

acc ggg cag ctg ctg tgc gaa gtg gcg tct tcg ggc aaa cga gac atc 192
 Thr Gly Gln Leu Leu Cys Glu Val Ala Ser Ser Gly Lys Arg Asp Ile
 50 55 60
 gat ctg gcg ctg gat gct gcg cac aaa gtg aaa gat aaa tgg gcg cac 240

Ala Gly Gly Val Ile Gly Glu Tyr Leu Ala Thr Ser Lys Arg Ile Ala

225	230	235	240	
aaa gtg gcg ttt acc ggc tca acg gaa gtg ggc caa caa att atg caa	768			
Lys Val Ala Phe Thr Gly Ser Thr Glu Val Gly Gln Gln Ile Met Gln				
245	250	255		
tac gca acg caa aac att att ccg gtg acg ctg gag ttg ggc ggt aag	816			
Tyr Ala Thr Gln Asn Ile Ile Pro Val Thr Leu Glu Leu Gly Gly Lys				
260	265	270		
tcg cca aat atc ttc ttt gct gat gtg atg gat gaa gaa gat gcc ttt	864			
Ser Pro Asn Ile Phe Phe Ala Asp Val Met Asp Glu Glu Asp Ala Phe				
275	280	285		
ttc gat aaa gcg ctg gaa ggc ttt gca ctg ttt gcc ttt aac cag ggc	912			
Phe Asp Lys Ala Leu Glu Gly Phe Ala Leu Phe Ala Phe Asn Gln Gly				
290	295	300		
gaa gtt tgc acc tgt ccg agt cgt gct tta gtg cag gaa tct atc tac	960			
Glu Val Cys Thr Cys Pro Ser Arg Ala Leu Val Gln Glu Ser Ile Tyr				
305	310	315	320	
gaa cgc ttt atg gaa cgc gcc atc cgc cgt gtc gaa agc att cgt agc	1008			
Glu Arg Phe Met Glu Arg Ala Ile Arg Arg Val Glu Ser Ile Arg Ser				
325	330	335		
ggt aac ccg ctc gac agc gtg acg caa atg ggc gcg cag gtt tct cac	1056			
Gly Asn Pro Leu Asp Ser Val Thr Gln Met Gly Ala Gln Val Ser His				
340	345	350		
ggg caa ctg gaa acc atc ctc aac tac att gat atc ggt aaa aaa gag	1104			
Gly Gln Leu Glu Thr Ile Leu Asn Tyr Ile Asp Ile Gly Lys Lys Glu				
355	360	365		
ggc gct gac gtg ctc aca ggc ggg cgg cgc aag ctg ctg gaa ggt gaa	1152			
Gly Ala Asp Val Leu Thr Gly Gly Arg Arg Lys Leu Leu Glu Gly Glu				
370	375	380		

ctg aaa gac ggc tac tac ctc gaa ccg acg att ctg ttt ggt cag aac 1200
 Leu Lys Asp Gly Tyr Tyr Leu Glu Pro Thr Ile Leu Phe Gly Gln Asn
 385 390 395 400
 aat atg cgg gtg ttc cag gag gag att ttt ggc ccg gtg ctg gcg gtg 1248
 Asn Met Arg Val Phe Gln Glu Glu Ile Phe Gly Pro Val Leu Ala Val
 405 410 415
 acc acc ttc aaa acg atg gaa gaa gcg ctg gag ctg gcg aac gat acg 1296
 Thr Thr Phe Lys Thr Met Glu Glu Ala Leu Glu Leu Ala Asn Asp Thr
 420 425 430
 caa tat ggc ctg ggc gcg ggc gtc tgg agc cgc aac ggt aat ctg gcc 1344
 Gln Tyr Gly Leu Gly Ala Gly Val Trp Ser Arg Asn Gly Asn Leu Ala
 435 440 445
 tat aag atg ggg cgc ggc ata cag gct ggg cgc gtg tgg acc aac tgt 1392
 Tyr Lys Met Gly Arg Gly Ile Gln Ala Gly Arg Val Trp Thr Asn Cys
 450 455 460
 tat cac gct tac ccg gca cat gcg gcg ttt ggt ggc tac aaa caa tca 1440
 Tyr His Ala Tyr Pro Ala His Ala Ala Phe Gly Gly Tyr Lys Gln Ser
 465 470 475 480
 ggt atc ggt cgc gaa acc cac aag atg atg ctg gag cat tac cag caa 1488
 Gly Ile Gly Arg Glu Thr His Lys Met Met Leu Glu His Tyr Gln Gln
 485 490 495
 acc aag tgc ctg ctg gtg agc tac tcg gat aaa ccg ttg ggg ctg ttc 1536
 Thr Lys Cys Leu Leu Val Ser Tyr Ser Asp Lys Pro Leu Gly Leu Phe
 500 505 510
 tga 1539
 <210> 16
 <211> 512
 <212> PRT
 <213> Escherichia coli
 <400> 16
 Met Thr Asn Asn Pro Pro Ser Ala Gln Ile Lys Pro Gly Glu Tyr Gly

1 5 10 15

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

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

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

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

Tyr Ala Thr Gln Asn Ile Ile Pro Val Thr Leu Glu Leu Gly Gly Lys
 260 265 270

Ser Pro Asn Ile Phe Phe Ala Asp Val Met Asp Glu Glu Asp Ala Phe
 275 280 285

Phe Asp Lys Ala Leu Glu Gly Phe Ala Leu Phe Ala Phe Asn Gln Gly
 290 295 300

Glu Val Cys Thr Cys Pro Ser Arg Ala Leu Val Gln Glu Ser Ile Tyr
 305 310 315 320

Glu Arg Phe Met Glu Arg Ala Ile Arg Arg Val Glu Ser Ile Arg Ser
 325 330 335

Gly Asn Pro Leu Asp Ser Val Thr Gln Met Gly Ala Gln Val Ser His
 340 345 350

Gly Gln Leu Glu Thr Ile Leu Asn Tyr Ile Asp Ile Gly Lys Lys Glu
 355 360 365

Gly Ala Asp Val Leu Thr Gly Gly Arg Arg Lys Leu Leu Glu Gly Glu
 370 375 380

Leu Lys Asp Gly Tyr Tyr Leu Glu Pro Thr Ile Leu Phe Gly Gln Asn
 385 390 395 400

Asn Met Arg Val Phe Gln Glu Glu Ile Phe Gly Pro Val Leu Ala Val
 405 410 415

Thr Thr Phe Lys Thr Met Glu Glu Ala Leu Glu Leu Ala Asn Asp Thr
 420 425 430

Gln Tyr Gly Leu Gly Ala Gly Val Trp Ser Arg Asn Gly Asn Leu Ala
 435 440 445

Tyr Lys Met Gly Arg Gly Ile Gln Ala Gly Arg Val Trp Thr Asn Cys
 450 455 460

Tyr His Ala Tyr Pro Ala His Ala Ala Phe Gly Gly Tyr Lys Gln Ser
 465 470 475 480

Gly Ile Gly Arg Glu Thr His Lys Met Met Leu Glu His Tyr Gln Gln
 485 490 495

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

500	505	510	
<210> 17			
<211> 1440			
<212> DNA			
<213> Escherichia coli			
<220><221> CDS			
<222> (1)..(1440)			
<400> 17			
atg tca gta ccc gtt caa cat cct atg tat atc gat gga cag ttt gtt			48
Met Ser Val Pro Val Gln His Pro Met Tyr Ile Asp Gly Gln Phe Val			
1 5 10 15			
acc tgg cgt gga gac gca tgg att gat gtg gta aac cct gct aca gag			96
Thr Trp Arg Gly Asp Ala Trp Ile Asp Val Val Asn Pro Ala Thr Glu			
20 25 30			
gct gtc att tcc cgc ata ccc gat ggt cag gcc gag gat gcc cgt aag			144
Ala Val Ile Ser Arg Ile Pro Asp Gly Gln Ala Glu Asp Ala Arg Lys			
35 40 45			
gca atc gat gca gca gaa cgt gca caa cca gaa tgg gaa gcg ttg cct			192
Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro Glu Trp Glu Ala Leu Pro			
50 55 60			
gct att gaa cgc gcc agt tgg ttg cgc aaa atc tcc gcc ggg atc cgc			240
Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys Ile Ser Ala Gly Ile Arg			
65 70 75 80			
gaa cgc gcc agt gaa atc agt gcg ctg att gtt gaa gaa ggg ggc aag			288
Glu Arg Ala Ser Glu Ile Ser Ala Leu Ile Val Glu Glu Gly Gly Lys			
85 90 95			
atc cag cag ctg gct gaa gtc gaa gtg gct ttt act gcc gac tat atc			336
Ile Gln Gln Leu Ala Glu Val Glu Val Ala Phe Thr Ala Asp Tyr Ile			
100 105 110			
gat tac atg gcg gag tgg gca cgg cgt tac gag ggc gag att att caa			384
Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr Glu Gly Glu Ile Ile Gln			

115	120	125	
agc gat cgt cca gga gaa aat att ctt ttg ttt aaa cgt gcg ctt ggt			432
Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu Phe Lys Arg Ala Leu Gly			
130	135	140	
gtg act acc ggc att ctg ccg tgg aac ttc ccg ttc ttc ctc att gcc			480
Val Thr Thr Gly Ile Leu Pro Trp Asn Phe Pro Phe Phe Leu Ile Ala			
145	150	155	160
cgc aaa atg gct ccc gct ctt ttg acc ggt aat acc atc gtc att aaa			528
Arg Lys Met Ala Pro Ala Leu Leu Thr Gly Asn Thr Ile Val Ile Lys			
165	170	175	
cct agt gaa ttt acg cca aac aat gcg att gca ttc gcc aaa atc gtc			576
Pro Ser Glu Phe Thr Pro Asn Asn Ala Ile Ala Phe Ala Lys Ile Val			
180	185	190	
gat gaa ata ggc ctt ccg cgc ggc gtg ttt aac ctt gta ctg ggg cgt			624
Asp Glu Ile Gly Leu Pro Arg Gly Val Phe Asn Leu Val Leu Gly Arg			
195	200	205	
ggg gaa acc gtt ggg caa gaa ctg gcg ggt aac cca aag gtc gca atg			672
Gly Glu Thr Val Gly Gln Glu Leu Ala Gly Asn Pro Lys Val Ala Met			
210	215	220	
gtc agt atg aca ggc agc gtc tct gca ggt gag aag atc atg gcg act			720
Val Ser Met Thr Gly Ser Val Ser Ala Gly Glu Lys Ile Met Ala Thr			
225	230	235	240
gcg gcg aaa aac atc acc aaa gtg tgt ctg gaa ttg ggg ggt aaa gca			768
Ala Ala Lys Asn Ile Thr Lys Val Cys Leu Glu Leu Gly Gly Lys Ala			
245	250	255	
cca gct atc gta atg gac gat gcc gat ctt gaa ctg gca gtc aaa gcc			816
Pro Ala Ile Val Met Asp Asp Ala Asp Leu Glu Leu Ala Val Lys Ala			
260	265	270	
atc gtt gat tca cgc gtc att aat agt ggg caa gtg tgt aac tgt gca			864
Ile Val Asp Ser Arg Val Ile Asn Ser Gly Gln Val Cys Asn Cys Ala			
275	280	285	

gaa cgt gtt tat gta cag aaa ggc att tat gat cag ttc gtc aat cgg	912
Glu Arg Val Tyr Val Gln Lys Gly Ile Tyr Asp Gln Phe Val Asn Arg	
290 295 300	
ctg ggt gaa gcg atg cag gcg gtt caa ttt ggt aac ccc gct gaa cgc	960
Leu Gly Glu Ala Met Gln Ala Val Gln Phe Gly Asn Pro Ala Glu Arg	
305 310 315 320	
aac gac att gcg atg ggg ccg ttg att aac gcc gcg gcg ctg gaa agg	1008
Asn Asp Ile Ala Met Gly Pro Leu Ile Asn Ala Ala Ala Leu Glu Arg	
325 330 335	
gtc gag caa aaa gtg gcg cgc gca gta gaa gaa ggg gcg aga gtg gcg	1056
Val Glu Gln Lys Val Ala Arg Ala Val Glu Glu Gly Ala Arg Val Ala	
340 345 350	
ttc ggt ggc aaa gcg gta gag ggg aaa gga tat tat tat ccg ccg aca	1104
Phe Gly Gly Lys Ala Val Glu Gly Lys Gly Tyr Tyr Tyr Pro Pro Thr	
355 360 365	
ttg ctg ctg gat gtt cgc cag gaa atg tgc att atg cat gag gaa acc	1152
Leu Leu Leu Asp Val Arg Gln Glu Met Ser Ile Met His Glu Glu Thr	
370 375 380	
ttt ggc ccg gtg ctg cca gtt gtc gca ttt gac acg ctg gaa gat gct	1200
Phe Gly Pro Val Leu Pro Val Val Ala Phe Asp Thr Leu Glu Asp Ala	
385 390 395 400	
atc tca atg gct aat gac agt gat tac ggc ctg acc tca tca atc tat	1248
Ile Ser Met Ala Asn Asp Ser Asp Tyr Gly Leu Thr Ser Ser Ile Tyr	
405 410 415	
acc caa aat ctg aac gtc gcg atg aaa gcc att aaa ggg ctg aag ttt	1296
Thr Gln Asn Leu Asn Val Ala Met Lys Ala Ile Lys Gly Leu Lys Phe	
420 425 430	
ggg gaa act tac atc aac cgt gaa aac ttc gaa gct atg caa ggc ttc	1344
Gly Glu Thr Tyr Ile Asn Arg Glu Asn Phe Glu Ala Met Gln Gly Phe	

435 440 445

cac gcc gga tgg cgt aaa tcc ggt att ggc ggc gca gat ggt aaa cat 1392

His Ala Gly Trp Arg Lys Ser Gly Ile Gly Gly Ala Asp Gly Lys His

450 455 460

ggc ttg cat gaa tat ctg cag acc cag gtg gtt tat tta cag tct taa 1440

Gly Leu His Glu Tyr Leu Gln Thr Gln Val Val Tyr Leu Gln Ser

465 470 475

<210> 18

<211> 479

<212> PRT

<213> Escherichia coli

<400> 18

Met Ser Val Pro Val Gln His Pro Met Tyr Ile Asp Gly Gln Phe Val

1 5 10 15

Thr Trp Arg Gly Asp Ala Trp Ile Asp Val Val Asn Pro Ala Thr Glu

20 25 30

Ala Val Ile Ser Arg Ile Pro Asp Gly Gln Ala Glu Asp Ala Arg Lys

35 40 45

Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro Glu Trp Glu Ala Leu Pro

50 55 60

Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys Ile Ser Ala Gly Ile Arg

65 70 75 80

Glu Arg Ala Ser Glu Ile Ser Ala Leu Ile Val Glu Glu Gly Gly Lys

85 90 95

Ile Gln Gln Leu Ala Glu Val Glu Val Ala Phe Thr Ala Asp Tyr Ile

100 105 110

Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr Glu Gly Glu Ile Ile Gln

115 120 125

Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu Phe Lys Arg Ala Leu Gly

130 135 140

Val Thr Thr Gly Ile Leu Pro Trp Asn Phe Pro Phe Phe Leu Ile Ala

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

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

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

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

 Phe Gly Pro Val Leu Pro Val Val Ala Phe Asp Thr Leu Glu Asp Ala
 385 390 395 400

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

His Ala Gly Trp Arg Lys Ser Gly Ile Gly Gly Ala Asp Gly Lys His
 450 455 460
 Gly Leu His Glu Tyr Leu Gln Thr Gln Val Val Tyr Leu Gln Ser
 465 470 475

<210> 19

<211> 1488

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(1488)

<400> 19

atg aat ttt cat cat ctg gct tac tgg cag gat aaa gcg tta agt ctc 48
 Met Asn Phe His His Leu Ala Tyr Trp Gln Asp Lys Ala Leu Ser Leu

1 5 10 15
 gcc att gaa aac cgc tta ttt att aac ggt gaa tat act gct gcg gcg 96
 Ala Ile Glu Asn Arg Leu Phe Ile Asn Gly Glu Tyr Thr Ala Ala Ala

20 25 30
 gaa aat gaa acc ttt gaa acc gtt gat ccg gtc acc cag gca ccg ctg 144
 Glu Asn Glu Thr Phe Glu Thr Val Asp Pro Val Thr Gln Ala Pro Leu

35 40 45

gcg aaa att gcc cgc ggc aag agc gtc gat atc gac cgt gcg atg agc 192
 Ala Lys Ile Ala Arg Gly Lys Ser Val Asp Ile Asp Arg Ala Met Ser

50 55 60
 gca gca cgc ggc gta ttt gaa cgc ggc gac tgg tca ctc tct tct ccg 240
 Ala Ala Arg Gly Val Phe Glu Arg Gly Asp Trp Ser Leu Ser Ser Pro

65 70 75 80

gct aaa cgt aaa gcg gta ctg aat aaa ctc gcc gat tta atg gaa gcc 288

Ala Lys Arg Lys Ala Val Leu Asn Lys Leu Ala Asp Leu Met Glu Ala
85 90 95

cac gcc gaa gag ctg gca ctg ctg gaa act ctc gac acc ggc aaa ccg 336

His Ala Glu Glu Leu Ala Leu Leu Glu Thr Leu Asp Thr Gly Lys Pro
100 105 110

att cgt cac agt ctg cgt gat gat att ccc ggc gcg gcg cgc gcc att 384

Ile Arg His Ser Leu Arg Asp Asp Ile Pro Gly Ala Ala Arg Ala Ile
115 120 125

cgc tgg tac gcc gaa gcg atc gac aaa gtg tat ggc gaa gtg gcg acc 432

Arg Trp Tyr Ala Glu Ala Ile Asp Lys Val Tyr Gly Glu Val Ala Thr
130 135 140

acc agt agc cat gag ctg gcg atg atc gtg cgt gaa ccg gtc ggc gtg 480

Thr Ser Ser His Glu Leu Ala Met Ile Val Arg Glu Pro Val Gly Val
145 150 155 160

att gcc gcc atc gtg ccg tgg aac ttc ccg ctg ttg ctg act tgc tgg 528

Ile Ala Ala Ile Val Pro Trp Asn Phe Pro Leu Leu Leu Thr Cys Trp
165 170 175

aaa ctc ggc ccg gcg ctg gcg gcg gga aac agc gtg att cta aaa ccg 576

Lys Leu Gly Pro Ala Leu Ala Ala Gly Asn Ser Val Ile Leu Lys Pro
180 185 190

tct gaa aaa tca ccg ctc agt gcg att cgt ctc gcg ggg ctg gcg aaa 624

Ser Glu Lys Ser Pro Leu Ser Ala Ile Arg Leu Ala Gly Leu Ala Lys
195 200 205

gaa gca ggc ttg ccg gat ggt gtg ttg aac gtg gtg acg ggt ttt ggt 672

Glu Ala Gly Leu Pro Asp Gly Val Leu Asn Val Val Thr Gly Phe Gly
210 215 220

cat gaa gcc ggg cag gcg ctg tcg cgt cat aac gat atc gac gcc att 720

His Glu Ala Gly Gln Ala Leu Ser Arg His Asn Asp Ile Asp Ala Ile

225	230	235	240	
gcc ttt acc ggt tca acc cgt acc ggg aaa cag ctg ctg aaa gat gcg				768
Ala Phe Thr Gly Ser Thr Arg Thr Gly Lys Gln Leu Leu Lys Asp Ala				
	245	250	255	
ggc gac agc aac atg aaa cgc gtc tgg ctg gaa gcg ggc ggc aaa agc				816
Gly Asp Ser Asn Met Lys Arg Val Trp Leu Glu Ala Gly Gly Lys Ser				
	260	265	270	
gcc aac atc gtt ttc gct gac tgc ccg gat ttg caa cag gcg gca agc				864
Ala Asn Ile Val Phe Ala Asp Cys Pro Asp Leu Gln Gln Ala Ala Ser				
	275	280	285	
gcc acc gca gca ggc att ttc tac aac cag gga cag gtg tgc atc gcc				912
Ala Thr Ala Ala Gly Ile Phe Tyr Asn Gln Gly Gln Val Cys Ile Ala				
	290	295	300	
gga acg cgc ctg ttg ctg gaa gag agc atc gcc gat gaa ttc tta gcc				960
Gly Thr Arg Leu Leu Leu Glu Glu Ser Ile Ala Asp Glu Phe Leu Ala				
305	310	315	320	
ctg tta aaa cag cag gcg caa aac tgg cag ccg ggc cat cca ctt gat				1008
Leu Leu Lys Gln Gln Ala Gln Asn Trp Gln Pro Gly His Pro Leu Asp				
	325	330	335	
ccc gca acc acc atg ggc acc tta atc gac tgc gcc cac gcc gac tgc				1056
Pro Ala Thr Thr Met Gly Thr Leu Ile Asp Cys Ala His Ala Asp Ser				
	340	345	350	
gtc cat agc ttt att cgg gaa ggc gaa agc aaa ggg caa ctg ttg ttg				1104
Val His Ser Phe Ile Arg Glu Gly Glu Ser Lys Gly Gln Leu Leu Leu				
	355	360	365	
gat ggc cgt aac gcc ggg ctg gct gcc gcc atc ggc ccg acc atc ttt				1152
Asp Gly Arg Asn Ala Gly Leu Ala Ala Ala Ile Gly Pro Thr Ile Phe				
	370	375	380	
gtg gat gtg gac ccg aat gcg tcc tta agt cgc gaa gag att ttc ggt				1200
Val Asp Val Asp Pro Asn Ala Ser Leu Ser Arg Glu Glu Ile Phe Gly				
385	390	395	400	

ccg gtg ctg gtg gtc acg cgt ttc aca tca gaa gaa cag gcg cta cag 1248

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

405 410 415

ctt gcc aac gac agc cag tac ggc ctt ggc gcg gcg gta tgg acg cgc 1296

Leu Ala Asn Asp Ser Gln Tyr Gly Leu Gly Ala Ala Val Trp Thr Arg

420 425 430

gac ctc tcc cgc gcg cac cgc atg agc cga cgc ctg aaa gcc ggt tcc 1344

Asp Leu Ser Arg Ala His Arg Met Ser Arg Arg Leu Lys Ala Gly Ser

435 440 445

gtc ttc gtc aat aac tac aac gac ggc gat atg acc gtg ccg ttt ggc 1392

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

450 455 460

ggc tat aag cag agc ggc aac ggt cgc gac aaa tcc ctg cat gcc ctt 1440

Gly Tyr Lys Gln Ser Gly Asn Gly Arg Asp Lys Ser Leu His Ala Leu

465 470 475 480

gaa aaa ttc act gaa ctg aaa acc atc tgg ata agc ctg gag gcc tga 1488

Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ser Leu Glu Ala

485 490 495

<210> 20

<211> 495

<212> PRT

<213> Escherichia coli

<400> 20

Met Asn Phe His His Leu Ala Tyr Trp Gln Asp Lys Ala Leu Ser Leu

1 5 10 15

Ala Ile Glu Asn Arg Leu Phe Ile Asn Gly Glu Tyr Thr Ala Ala Ala

20 25 30

Glu Asn Glu Thr Phe Glu Thr Val Asp Pro Val Thr Gln Ala Pro Leu

35 40 45

Ala Lys Ile Ala Arg Gly Lys Ser Val Asp Ile Asp Arg Ala Met Ser

50 55 60

Ala	Ala	Arg	Gly	Val	Phe	Glu	Arg	Gly	Asp	Trp	Ser	Leu	Ser	Ser	Pro
65				70				75				80			
Ala	Lys	Arg	Lys	Ala	Val	Leu	Asn	Lys	Leu	Ala	Asp	Leu	Met	Glu	Ala
85				90				95							
His	Ala	Glu	Glu	Leu	Ala	Leu	Leu	Glu	Thr	Leu	Asp	Thr	Gly	Lys	Pro
100				105				110							
Ile	Arg	His	Ser	Leu	Arg	Asp	Asp	Ile	Pro	Gly	Ala	Ala	Arg	Ala	Ile
115				120				125							
Arg	Trp	Tyr	Ala	Glu	Ala	Ile	Asp	Lys	Val	Tyr	Gly	Glu	Val	Ala	Thr
130				135				140							
Thr	Ser	Ser	His	Glu	Leu	Ala	Met	Ile	Val	Arg	Glu	Pro	Val	Gly	Val
145				150				155				160			
Ile	Ala	Ala	Ile	Val	Pro	Trp	Asn	Phe	Pro	Leu	Leu	Leu	Thr	Cys	Trp
165				170				175							
Lys	Leu	Gly	Pro	Ala	Leu	Ala	Ala	Gly	Asn	Ser	Val	Ile	Leu	Lys	Pro
180				185				190							
Ser	Glu	Lys	Ser	Pro	Leu	Ser	Ala	Ile	Arg	Leu	Ala	Gly	Leu	Ala	Lys
195				200				205							
Glu	Ala	Gly	Leu	Pro	Asp	Gly	Val	Leu	Asn	Val	Val	Thr	Gly	Phe	Gly
210				215				220							
His	Glu	Ala	Gly	Gln	Ala	Leu	Ser	Arg	His	Asn	Asp	Ile	Asp	Ala	Ile
225				230				235				240			
Ala	Phe	Thr	Gly	Ser	Thr	Arg	Thr	Gly	Lys	Gln	Leu	Leu	Lys	Asp	Ala
245				250				255							
Gly	Asp	Ser	Asn	Met	Lys	Arg	Val	Trp	Leu	Glu	Ala	Gly	Gly	Lys	Ser
260				265				270							
Ala	Asn	Ile	Val	Phe	Ala	Asp	Cys	Pro	Asp	Leu	Gln	Gln	Ala	Ala	Ser
275				280				285							
Ala	Thr	Ala	Ala	Gly	Ile	Phe	Tyr	Asn	Gln	Gly	Gln	Val	Cys	Ile	Ala
290				295				300							
Gly	Thr	Arg	Leu	Leu	Leu	Glu	Glu	Ser	Ile	Ala	Asp	Glu	Phe	Leu	Ala

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

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

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

<210> 21

<211> 1395

<212> DNA

<213> Escherichia coli

<400> 21

atgcctgacg ctaaaaaaca ggggcggtca aacaaggcaa tgacgttttt cgtctgcttc	60
cttgccgctc tggcgggatt actctttggc ctggatatcg gtgtaattgc tggcgactg	120
ccgtttattg cagatgaatt ccagattact tcgcacacgc aagaatgggt cgtaagctcc	180
atgatgttcg gtgcggcagt cgggtcggtg ggcagcggct ggctctcctt taaactcggg	240

cgcaaaaaga gcctgatgat cggcgcaatt ttgtttgttg ccggttcgct gttctctgcg 300
gctgcgccaa acgttgaagt actgattctt tcccgcttc tactggggct ggcggtgggt 360
gtggcctctt ataccgcacc gctgtacctc tctgaaattg cgccggaaaa aattcgtggc 420

agtatgatct cgatgtatca gttgatgac actatcgga tcctcggcgc ttatctttct 480
gataccgctc tcagctacac cgggtcatgg cgtgggatgc tgggtgtgat tatcatccc 540
gcaattttgc tgctgattgg tgtcttcttc ctgccagaca gcccacgttg gtttccgcc 600
aaacgccgtt ttgttgatgc cgaacgcgtg ctgctacgcc tgcgtgacac cagcgcgga 660
gcgaacgcg aactggatga aatccgtgaa agtttgagg ttaaacagag tggctgggcg 720
ctgtttaaag agaacagcaa ctccgccgc gcggtgttc ttggcgtact gttgcaggta 780
atgcagcaat tcaccgggat gaacgtcatc atgtattacg cgccgaaaat ctgcgaactg 840

gcgggttata ccaacactac cgagcaaatg tgggggaccg tgattgtcgg cctgaccaac 900
gtacttgcca ctttatcgc aatcgccctt gttgaccgt ggggacgtaa accaacgcta 960
acgtgggctc tcctggatga ggctgctggc atgggcgtac tcggtacaat gatgcatac 1020
ggtattcact ctccgtcggc gcagtatttc gccatcgcca tgctgctgat gtttattgtc 1080
ggttttgcca tgagtgcggc tccgtgatt tgggtactgt gctccgaaat tcagccgctg 1140
aaaggccgcg attttggcat cacctgtcc actgccacca actggattgc caacatgatc 1200
gttggcgcaa cgttctgac catgtcaac acgctgggta acgccaacac cttctgggtg 1260

tatgcggctc tgaacgtact gtttatcctg ctgacattgt ggctggtacc ggaaacaaaa 1320
cacgtttcgc tggaacatat tgaacgtaat ctgatgaaag gtcgtaaact gcgcgaaata 1380
ggcgctcacg attaa 1395

<210> 22

<211> 464

<212> PRT

<213> Escherichia coli

<400> 22

Met Pro Asp Ala Lys Lys Gln Gly Arg Ser Asn Lys Ala Met Thr Phe

1 5 10 15

Phe Val Cys Phe Leu Ala Ala Leu Ala Gly Leu Leu Phe Gly Leu Asp

20 25 30

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

35 40 45

Ile Thr Ser His Thr Gln Glu Trp Val Val Ser Ser Met Met Phe Gly
 50 55 60
 Ala Ala Val Gly Ala Val Gly Ser Gly Trp Leu Ser Phe Lys Leu Gly
 65 70 75 80
 Arg Lys Lys Ser Leu Met Ile Gly Ala Ile Leu Phe Val Ala Gly Ser
 85 90 95
 Leu Phe Ser Ala Ala Ala Pro Asn Val Glu Val Leu Ile Leu Ser Arg
 100 105 110
 Val Leu Leu Gly Leu Ala Val Gly Val Ala Ser Tyr Thr Ala Pro Leu
 115 120 125
 Tyr Leu Ser Glu Ile Ala Pro Glu Lys Ile Arg Gly Ser Met Ile Ser
 130 135 140
 Met Tyr Gln Leu Met Ile Thr Ile Gly Ile Leu Gly Ala Tyr Leu Ser
 145 150 155 160
 Asp Thr Ala Phe Ser Tyr Thr Gly Ala Trp Arg Trp Met Leu Gly Val
 165 170 175
 Ile Ile Ile Pro Ala Ile Leu Leu Leu Ile Gly Val Phe Phe Leu Pro
 180 185 190
 Asp Ser Pro Arg Trp Phe Ala Ala Lys Arg Arg Phe Val Asp Ala Glu
 195 200 205
 Arg Val Leu Leu Arg Leu Arg Asp Thr Ser Ala Glu Ala Lys Arg Glu
 210 215 220
 Leu Asp Glu Ile Arg Glu Ser Leu Gln Val Lys Gln Ser Gly Trp Ala
 225 230 235 240
 Leu Phe Lys Glu Asn Ser Asn Phe Arg Arg Ala Val Phe Leu Gly Val
 245 250 255
 Leu Leu Gln Val Met Gln Gln Phe Thr Gly Met Asn Val Ile Met Tyr
 260 265 270
 Tyr Ala Pro Lys Ile Phe Glu Leu Ala Gly Tyr Thr Asn Thr Thr Glu
 275 280 285
 Gln Met Trp Gly Thr Val Ile Val Gly Leu Thr Asn Val Leu Ala Thr

290	295	300	
Phe Ile Ala Ile Gly Leu Val Asp Arg Trp Gly Arg Lys Pro Thr Leu			
305	310	315	320
Thr Leu Gly Phe Leu Val Met Ala Ala Gly Met Gly Val Leu Gly Thr			
325	330	335	
Met Met His Ile Gly Ile His Ser Pro Ser Ala Gln Tyr Phe Ala Ile			
340	345	350	
Ala Met Leu Leu Met Phe Ile Val Gly Phe Ala Met Ser Ala Gly Pro			
355	360	365	
Leu Ile Trp Val Leu Cys Ser Glu Ile Gln Pro Leu Lys Gly Arg Asp			
370	375	380	
Phe Gly Ile Thr Cys Ser Thr Ala Thr Asn Trp Ile Ala Asn Met Ile			
385	390	395	400
Val Gly Ala Thr Phe Leu Thr Met Leu Asn Thr Leu Gly Asn Ala Asn			
405	410	415	
Thr Phe Trp Val Tyr Ala Ala Leu Asn Val Leu Phe Ile Leu Leu Thr			
420	425	430	
Leu Trp Leu Val Pro Glu Thr Lys His Val Ser Leu Glu His Ile Glu			
435	440	445	
Arg Asn Leu Met Lys Gly Arg Lys Leu Arg Glu Ile Gly Ala His Asp			
450	455	460	
<210>	23		
<211>	1248		
<212>	DNA		
<213>	Escherichia coli		
<400>	23		
atggcactga atattccatt cagaaatgcg tactatcggt ttgcatccag ttactcattt	60		
ctctttttta tttcctggtc gctgtggtgg tcgttatcac ctatttggct gaaaggacat	120		
ctaggattaa cagggacgga attaggtaca ctttattcgg tcaaccagtt taccagcatt	180		
ctatttatga tgttctacgg catcggtcag gataaactcg gtctgaagaa accgctcatc	240		
tggtgtatga gtttcattct ggtcttgacc ggaccgttta tgatttacgt ttatgaaccg	300		
ttactgcaaa gcaatttttc tgtaggtcta attctggggg cgctcttttt tggcctgggg	360		

tatctggcgg gatgcggttt gcttgacagc ttcaccgaaa aaatggcgcg aaattttcat 420

ttcgaatatg gaacagcgcg cgcctgggga tcttttggct atgctattgg cgcgttcttt 480

gccggtatat ttttagtat cagtccecat atcaacttct ggttggcttc getatttggc 540

gctgtattta tgatgatcaa catgcgtttt aaagataagg atcaccagtg catagcggcg 600

gatgcgggag gggtaaaaaa agaggatttt atcgagttt tcaaggatcg aaacttctgg 660

gttttcgtca tatttattgt ggggacgtgg tctttctata acatttttga tcaacaactc 720

tttctgtct tttatgcagg tttattcgaa tcacacgatg taggaacgcg cctgtatggt 780

tatctcaact cattccaggt ggtactcgaa gcgctgtgca tggcgattat tcctttcttt 840

gtgaatcggg tagggccaaa aaatgcattt cttatcggtg ttgtgattat ggcgttgcgt 900

atcctttcct gcgcgttggt cgtaaaccct tggattattt cattagttaa gctgttacat 960

gccattgagg ttccacttgg tgtcatatcc gtcttcaaat acagcgtggc aaactttgat 1020

aagcgctgt cgtcgacgat ctttctgatt ggttttcaaa ttgccagttc gcttgggatt 1080

gtgctgcttt caacgccgac tgggatactc ttgaccacg caggctacca gacagttttc 1140

ttcgcaattt cgggtattgt ctgcctgatg ttgctatttg gcattttctt cctgagtaaa 1200

aaacgcgagc aaatagttat ggaaacgcct gtaccttcag caatatag 1248

<210> 24

<211> 415

<212> PRT

<213> Escherichia coli

<400> 24

Met Ala Leu Asn Ile Pro Phe Arg Asn Ala Tyr Tyr Arg Phe Ala Ser

1 5 10 15

Ser Tyr Ser Phe Leu Phe Phe Ile Ser Trp Ser Leu Trp Trp Ser Leu

20 25 30

Tyr Ala Ile Trp Leu Lys Gly His Leu Gly Leu Thr Gly Thr Glu Leu

35 40 45

Gly Thr Leu Tyr Ser Val Asn Gln Phe Thr Ser Ile Leu Phe Met Met

50 55 60

Phe Tyr Gly Ile Val Gln Asp Lys Leu Gly Leu Lys Lys Pro Leu Ile

65 70 75 80

Trp Cys Met Ser Phe Ile Leu Val Leu Thr Gly Pro Phe Met Ile Tyr

85	90	95	
Val Tyr Glu Pro Leu Leu Gln Ser Asn Phe Ser Val Gly Leu Ile Leu			
100	105	110	
Gly Ala Leu Phe Phe Gly Leu Gly Tyr Leu Ala Gly Cys Gly Leu Leu			
115	120	125	
Asp Ser Phe Thr Glu Lys Met Ala Arg Asn Phe His Phe Glu Tyr Gly			
130	135	140	
Thr Ala Arg Ala Trp Gly Ser Phe Gly Tyr Ala Ile Gly Ala Phe Phe			
145	150	155	160
Ala Gly Ile Phe Phe Ser Ile Ser Pro His Ile Asn Phe Trp Leu Val			
165	170	175	
Ser Leu Phe Gly Ala Val Phe Met Met Ile Asn Met Arg Phe Lys Asp			
180	185	190	
Lys Asp His Gln Cys Ile Ala Ala Asp Ala Gly Gly Val Lys Lys Glu			
195	200	205	
Asp Phe Ile Ala Val Phe Lys Asp Arg Asn Phe Trp Val Phe Val Ile			
210	215	220	
Phe Ile Val Gly Thr Trp Ser Phe Tyr Asn Ile Phe Asp Gln Gln Leu			
225	230	235	240
Phe Pro Val Phe Tyr Ala Gly Leu Phe Glu Ser His Asp Val Gly Thr			
245	250	255	
Arg Leu Tyr Gly Tyr Leu Asn Ser Phe Gln Val Val Leu Glu Ala Leu			
260	265	270	
Cys Met Ala Ile Ile Pro Phe Phe Val Asn Arg Val Gly Pro Lys Asn			
275	280	285	
Ala Leu Leu Ile Gly Val Val Ile Met Ala Leu Arg Ile Leu Ser Cys			
290	295	300	
Ala Leu Phe Val Asn Pro Trp Ile Ile Ser Leu Val Lys Leu Leu His			
305	310	315	320
Ala Ile Glu Val Pro Leu Cys Val Ile Ser Val Phe Lys Tyr Ser Val			
325	330	335	

Ala Asn Phe Asp Lys Arg Leu Ser Ser Thr Ile Phe Leu Ile Gly Phe
 340 345 350
 Gln Ile Ala Ser Ser Leu Gly Ile Val Leu Leu Ser Thr Pro Thr Gly
 355 360 365
 Ile Leu Phe Asp His Ala Gly Tyr Gln Thr Val Phe Phe Ala Ile Ser
 370 375 380

Gly Ile Val Cys Leu Met Leu Leu Phe Gly Ile Phe Phe Leu Ser Lys
 385 390 395 400
 Lys Arg Glu Gln Ile Val Met Glu Thr Pro Val Pro Ser Ala Ile
 405 410 415

<210> 25

<211> 1248

<212> DNA

<213> Escherichia coli

<400> 25

atggcactga atattccatt cagaaatgcg tactatcggt ttgcatccag ttactcattt	60
ctcttttttta tttcctggtc gctgtggtgg tcgttatatcg ctatttggct gaaaggacat	120
ctagggttga caggacgga attaggtaca ctttattcgg tcaaccagtt taccagcatt	180
ctatttatga tgttctacgg catcggttcag gataaactcg gtctgaagaa accgctcatc	240
tggtgtatga gtttcaccc tggcttgacc ggaccgttta tgatttacgt ttatgaaccg	300
ttactgcaaa gcaatttttc tgtaggtcta attctggggg cgctattttt tggcttgggg	360
tatctggcgg gatgcggttt gcttgatagc ttcaccgaaa aaatggcgcg aaattttcat	420
ttcgaatatg gaacagcgcg gccttgggga tcttttggct atgctattgg cgcgttcttt	480
gccggcatat tttttagtat cagtccecat atcaacttct ggttggcttc gctatttggc	540
gctgtattta tgatgatcaa catgcgtttt aaagataagg atcaccagtg cgtagcggca	600
gatgcgggag gggtaaaaaa agaggatttt atcgagttt tcaaggatcg aaacttctgg	660
gttttctgca tattttattt ggggacgtgg tctttctata acatttttga tcaacaactt	720
tttctgtct tttattcagg tttattcgaa tcacacgatg taggaacgcg cctgtatggt	780
tatctcaact cattccaggt ggtactcgaa gcgctgtgca tggcgattat tcctttcttt	840
gtgaatcggg tagggccaaa aaatgcatta cttatcggag ttgtgattat ggcgttgcgt	900
atcctttcct gcgcgtgttt cgtaacccc tggattattt cattagttaa gttgttacct	960

gccattgagg ttccactttg tgtcatatcc gtcttcaa at acagcgtggc aaactttgat 1020

aagcgctgt cgtcgacgat ctttctgatt ggttttcaaa ttgccagttc gcttgggatt 1080

gtgctgcttt caacgccgac tgggatactc ttgaccacg caggctacca gacagttttc 1140

ttcgcaattt cgggtattgt ctgcctgatg ttgctatttg gcattttctt cttgagtaaa 1200

aaacgcgagc aaatagttat ggaaacgcct gtaccttcag caatatag 1248

<210> 26

<211> 415

<212> PRT

<213> Escherichia coli

<400> 26

Met Ala Leu Asn Ile Pro Phe Arg Asn Ala Tyr Tyr Arg Phe Ala Ser

1 5 10 15

Ser Tyr Ser Phe Leu Phe Phe Ile Ser Trp Ser Leu Trp Trp Ser Leu

20 25 30

Tyr Ala Ile Trp Leu Lys Gly His Leu Gly Leu Thr Gly Thr Glu Leu

35 40 45

Gly Thr Leu Tyr Ser Val Asn Gln Phe Thr Ser Ile Leu Phe Met Met

50 55 60

Phe Tyr Gly Ile Val Gln Asp Lys Leu Gly Leu Lys Lys Pro Leu Ile

65 70 75 80

Trp Cys Met Ser Phe Ile Leu Val Leu Thr Gly Pro Phe Met Ile Tyr

85 90 95

Val Tyr Glu Pro Leu Leu Gln Ser Asn Phe Ser Val Gly Leu Ile Leu

100 105 110

Gly Ala Leu Phe Phe Gly Leu Gly Tyr Leu Ala Gly Cys Gly Leu Leu

115 120 125

Asp Ser Phe Thr Glu Lys Met Ala Arg Asn Phe His Phe Glu Tyr Gly

130 135 140

Thr Ala Arg Ala Trp Gly Ser Phe Gly Tyr Ala Ile Gly Ala Phe Phe

145 150 155 160

Ala Gly Ile Phe Phe Ser Ile Ser Pro His Ile Asn Phe Trp Leu Val

165	170	175
Ser Leu Phe Gly Ala Val Phe Met Met Ile Asn Met Arg Phe Lys Asp		
180	185	190
Lys Asp His Gln Cys Val Ala Ala Asp Ala Gly Gly Val Lys Lys Glu		
195	200	205
Asp Phe Ile Ala Val Phe Lys Asp Arg Asn Phe Trp Val Phe Val Ile		
210	215	220
Phe Ile Val Gly Thr Trp Ser Phe Tyr Asn Ile Phe Asp Gln Gln Leu		
225	230	235
Phe Pro Val Phe Tyr Ser Gly Leu Phe Glu Ser His Asp Val Gly Thr		
245	250	255
Arg Leu Tyr Gly Tyr Leu Asn Ser Phe Gln Val Val Leu Glu Ala Leu		
260	265	270
Cys Met Ala Ile Ile Pro Phe Phe Val Asn Arg Val Gly Pro Lys Asn		
275	280	285
Ala Leu Leu Ile Gly Val Val Ile Met Ala Leu Arg Ile Leu Ser Cys		
290	295	300
Ala Leu Phe Val Asn Pro Trp Ile Ile Ser Leu Val Lys Leu Leu His		
305	310	315
Ala Ile Glu Val Pro Leu Cys Val Ile Ser Val Phe Lys Tyr Ser Val		
325	330	335
Ala Asn Phe Asp Lys Arg Leu Ser Ser Thr Ile Phe Leu Ile Gly Phe		
340	345	350
Gln Ile Ala Ser Ser Leu Gly Ile Val Leu Leu Ser Thr Pro Thr Gly		
355	360	365
Ile Leu Phe Asp His Ala Gly Tyr Gln Thr Val Phe Phe Ala Ile Ser		
370	375	380
Gly Ile Val Cys Leu Met Leu Leu Phe Gly Ile Phe Phe Leu Ser Lys		
385	390	395
Lys Arg Glu Gln Ile Val Met Glu Thr Pro Val Pro Ser Ala Ile		
405	410	415

<210> 27

<211> 1434

<212> DNA

<213> Escherichia coli

<400> 27

atgacgcaat ctcgattgca tgcggcgcaa aacgcactag caaaacttca cgagcgccga	60
ggtaacactt tctatcccca ttttcacctc gcgcctcctg ccgggtggat gaacgatcca	120
aacggcctga tctgggttaa cgatcgttat cacgcgtttt atcaacatca cccgatgagc	180
gaacactggg ggccaatgca ctggggacat gccaccagcg acgatatgat ccactggcag	240
catgagccta ttgcgctagc gccaggagac gagaatgaca aagacgggtg tttttcaggt	300
agtgctgtcg atgacaatgg tgtcctctca cttatctaca ccggacacgt ctggctcgat	360
gggtgcaggta atgacgatgc aattcgcgaa gtacaatgtc tggctaccag tcgggatggt	420
attcatttcg agaaacaggg tgtgatcctc actccaccag aaggcatcat gcacttccgc	480
gatactaaag tgtggcgtga agccgacaca tgggtggatgg tagtcggggc gaaagaccca	540
ggcaacacgg ggcagatcct gctttatcgc ggcagttcat tgcgtgaatg gactttcgat	600
cgcgtactgg cccacgtga tgcgggtgaa agctatatgt gggaatgtcc ggactttttc	660
agccttggcg atcagcatta tctgatgttt tccccgcagg gaatgaatgc cgagggatac	720
agttatcgaa atcgctttca aagtggcgta ataccggaa tgtggtcgcc aggacgactt	780
tttgcaaat ccgggcattt tactgaactt gataacgggc atgactttta tgcaccacaa	840
agctttgtag cgaaggatgg tccggctatt gttatcggct ggatggatat gtgggaatcg	900
ccaatgccct caaaacgtga aggtctggca ggctgcatga cgctggcgcg cgagctatca	960
gagagcaatg gcaaactcct acaacgcccg gtacacgaag ctgagtcgtt acgccagcag	1020
catcaatcta tctctccccg cacaatcagc aataaatatg ttttgcagga aaacgcgcaa	1080
gcagttgaga ttcagttgca gtgggcgctg aagaacagtg atgccgaaca ttacggatta	1140
cagctcggcg ctggaatgcg gctgtatatt gataaccaat ctgagcgact tgttttgtgg	1200
cggtattacc cacacgagaa tttagatggc taccgtagta ttccctccc gcagggtgac	1260
atgctcgccc taaggatatt tatcgataca tcatccgtgg aagtatttat taacgacggg	1320
gaggcggtag tgagtagccg aatatatccg cagccagaag aacgggaact gtcgctctat	1380
gcctcccacg gagtggctgt gctgcaacat ggagcactct ggcaactggg ttaa	1434

<210> 28

<211> 477

<212> PRT

<213> Escherichia coli

<400> 28

Met Thr Gln Ser Arg Leu His Ala Ala Gln Asn Ala Leu Ala Lys Leu

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

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

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

195 200 205
Gly Glu Ser Tyr Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asp
210 215 220
Gln His Tyr Leu Met Phe Ser Pro Gln Gly Met Asn Ala Glu Gly Tyr

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

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

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

 385 390 395 400
 Arg Tyr Tyr Pro His Glu Asn Leu Asp Gly Tyr Arg Ser Ile Pro Leu
 405 410 415
 Pro Gln Gly Asp Met Leu Ala Leu Arg Ile Phe Ile Asp Thr Ser Ser
 420 425 430
 Val Glu Val Phe Ile Asn Asp Gly Glu Ala Val Met Ser Ser Arg Ile
 435 440 445
 Tyr Pro Gln Pro Glu Glu Arg Glu Leu Ser Leu Tyr Ala Ser His Gly

 450 455 460
 Val Ala Val Leu Gln His Gly Ala Leu Trp Gln Leu Gly
 465 470 475

<210> 29

<211> 1434

<212> DNA

<213> Escherichia coli

<400> 29

atgacgcaat ctcgattgca tgcggcgcaa aacgccctag caaaacttca tgagcaccgg 60
ggtaacactt tctatcccca ttttcacctc gcgcctcctg ccgggtggat gaacgatcca 120
aacggcctga tctgggttaa cgatcgttat cagcggtttt atcaacatca tccgatgagc 180
gaacactggg ggccaatgca ctggggacat gccaccagcg acgatatgat ccactggcag 240

catgagccta ttgcgctagc gccaggagac gataatgaca aagacgggtg tttttcaggt 300
agtgtctgtc atgacaatgg tgtcctctca ctatctaca ccggacacgt ctggctcgat 360
gggtgcaggta atgacgatgc aattcgcgaa gtacaatgtc tggctaccag tcgggatggg 420
attcatttcg agaaacaggg tgtgatcctc actccaccag aaggaatcat gcacttccgc 480
gatactaaag tgtggcgtga agccgacaca tgggtggatgg tagtcggggc gaaagatcca 540
ggcaacacgg ggcagatcct gctttatcgc ggagttcgt tgcgtgaatg gaccttcgat 600
cgcgtactgg ccacgcctga tgcgggtgaa agctatatgt gggaatgtcc ggactttttc 660

agccttggcg atcagcatta tctgatgttt tccccgcagg gaatgaatgc cgagggatac 720
agttaccgaa atcgctttca aagtggcgta ataccggaa tgtgtcgcc aggacgactt 780
tttgcaaat ccgggcattt tactgaactt gataacgggc atgactttta tgcaccacaa 840
agcttttttag cgaaggatgg tcggcgatatt gttatcggct ggatggatat gtgggaatcg 900
ccaatgccct caaaacgtga aggatgggca ggctgcatga cgctggcgcg cgagctatca 960
gagagcaatg gcaaacttct acaacgccg gtacacgaag ctgagtcgtt acgccagcag 1020
catcaatctg tctctccccg cacaatcagc aataaatatg ttttcagga aaacgcgcaa 1080

gcagttgaga ttcagttgca gtgggcgctg aagaacagtg atgccgaaca ttacggatta 1140
cagctcgga ctggaatcg gctgtatatt gataaccaat ctgagcgact tgttttggg 1200
cggattacc cacacgagaa tttagacggc taccgtagta tccccctccc gcagcgtgac 1260
acgtcgccc taaggatatt tatcgataca tcatccgtgg aagtatttat taacgacggg 1320
gaagcgggta tgagtagtcg aatctatccg cagccagaag aacgggaact gtcgctttat 1380
gcctcccacg gagtggctgt gctgcaacat ggagcactct ggctactggg ttaa 1434

<210> 30

<211> 477

<212> PRT

<213> Escherichia coli

<400> 30

Met Thr Gln Ser Arg Leu His Ala Ala Gln Asn Ala Leu Ala Lys Leu

1 5 10 15

His Glu His Arg Gly Asn Thr Phe Tyr Pro His Phe His Leu Ala Pro

20 25 30

Pro Ala Gly Trp Met Asn Asp Pro Asn Gly Leu Ile Trp Phe Asn Asp

35 40 45

Arg Tyr His Ala Phe Tyr Gln His His Pro Met Ser Glu His Trp Gly

50 55 60

Pro Met His Trp Gly His Ala Thr Ser Asp Asp Met Ile His Trp Gln

65 70 75 80

His Glu Pro Ile Ala Leu Ala Pro Gly Asp Asp Asn Asp Lys Asp Gly

85 90 95

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

100 105 110

Tyr Thr Gly His Val Trp Leu Asp Gly Ala Gly Asn Asp Asp Ala Ile

115 120 125

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

130 135 140

Lys Gln Gly Val Ile Leu Thr Pro Pro Glu Gly Ile Met His Phe Arg

145 150 155 160

Asp Pro Lys Val Trp Arg Glu Ala Asp Thr Trp Trp Met Val Val Gly

165 170 175

Ala Lys Asp Pro Gly Asn Thr Gly Gln Ile Leu Leu Tyr Arg Gly Ser

180 185 190

Ser Leu Arg Glu Trp Thr Phe Asp Arg Val Leu Ala His Ala Asp Ala

195 200 205

Gly Glu Ser Tyr Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asp

210 215 220

Gln His Tyr Leu Met Phe Ser Pro Gln Gly Met Asn Ala Glu Gly Tyr

225 230 235 240
 Ser Tyr Arg Asn Arg Phe Gln Ser Gly Val Ile Pro Gly Met Trp Ser
 245 250 255

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

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

 Gly Met Arg Leu Tyr Ile Asp Asn Gln Ser Glu Arg Leu Val Leu Trp
 385 390 395 400
 Arg Tyr Tyr Pro His Glu Asn Leu Asp Gly Tyr Arg Ser Ile Pro Leu
 405 410 415
 Pro Gln Arg Asp Thr Leu Ala Leu Arg Ile Phe Ile Asp Thr Ser Ser
 420 425 430
 Val Glu Val Phe Ile Asn Asp Gly Glu Ala Val Met Ser Ser Arg Ile
 435 440 445

 Tyr Pro Gln Pro Glu Glu Arg Glu Leu Ser Leu Tyr Ala Ser His Gly
 450 455 460
 Val Ala Val Leu Gln His Gly Ala Leu Trp Leu Leu Gly
 465 470 475

<210> 31

<211> 1599

<212> DNA

<213> Bifidobacterium lactis

<400> 31

atggcaaccc ttcccacaa tattcccgcc aacggcattc tgacccccga cccggcgctc	60
gaccctgtgc tcacgccgat ctccggacat gccgagcagc tgtcactcgc cgaagcaggc	120
gtgtcggcac tggaaccac ccgcaacgac cgttgttacc cgaagttcca cattgcctcc	180
aatggcgggt ggatcaacga cccgaacggc ctgtgccgct acaacggacg ctggcacgtg	240
ttctaccagc tgcatcccca cggcacacag tggggcccga tgcatgggg ccacgtctcc	300
tccgacaaca tggtcgactg gcaccgcgaa cccatcgctt tcgcgccaag cctcgaacag	360
gaacgccacg gtgtgttctc cggttccgcc gtgattggcg acgacggcaa gccgtggatt	420
ttctacaccg gccaccgctg ggccaacggc aaggacaaca ccggaggcga ctggcaggtg	480
cagatgctcg ccaagccgaa gcacgacgaa ctgaagacct tcacgaagga gggcatgac	540
atcgactgcc ccaccgacga ggtggaccac cacttccgcg acccgaaggt gtggaagacc	600
ggtgacacct ggtatatgac cttcgggtgc tcgtcgaagg agcatcgtgg ccagatgtgg	660
ctgtacacgt cgagcgacat ggtgcactgg agcttcgacg ggggtgctgtt cgagcatccg	720
gatccgaacg tgttcatgct tgaatgcccc gatttcttcc cgatccgca tgcgcggggc	780
aacgagaaat gggtcacggt cttctccgcg atgggtgcca agccaaatgg cttcatgaac	840
cgcaacgtga acaatgccgg ctacatggtg ggcacatgga agccaggcga gagcttcaag	900
ccggagaccg agttccgct gtgggacgaa ggccataact tctatgcacc acagtcgttc	960
aacaccgaag ggcgccagat catgtacggc tggatgagcc cgttcgtcgc ccccatcccc	1020
atggaggagg acggctggtg cggcaacctc accctcccc gcgagatcac gctgggcgat	1080
gacggtgacc tggtcaccgc ccccaccatc gaaatggagg ggctgcgcga gaataccata	1140
ggcttcgact cgtcgcacct tggtagaac cagacctcca cgatcctcga cgatgacggc	1200
ggcgcccttg aaatcgagat gagactcgat ctgaacaaaa ccaccgccga acgcgccgga	1260
ctgcatgtgc atgccacaag gcacggccac tacacggcaa tcgtattcga cgcgcagatc	1320
ggcggcgctg tcatcgaccg gcagaacgtg gcgaacggag acaaaggcta ccgggtggcc	1380
aagctcagcg acaccgagct cgcagccgat acgcttgact tgcgcgtgtt catcgaccgc	1440
ggatgcgtcg aggtctacgt cgacggcggc aagcatgcga tgagctcgta ctcgttccct	1500
ggcgaatggc cagcgcctgt cgaactcgtg agcgaatccg gcaccacga catcgacacc	1560

ctcaccatgc actcgctcaa gtccatcgga ctcgagtga

1599

<210> 32

<211> 532

<212> PRT

<213> Bifidobacterium lactis

<400> 32

Met Ala Thr Leu Pro Thr Asn Ile Pro Ala Asn Gly Ile Leu Thr Pro

1 5 10 15

Asp Pro Ala Leu Asp Pro Val Leu Thr Pro Ile Ser Asp His Ala Glu

20 25 30

Gln Leu Ser Leu Ala Glu Ala Gly Val Ser Ala Leu Glu Thr Thr Arg

35 40 45

Asn Asp Arg Trp Tyr Pro Lys Phe His Ile Ala Ser Asn Gly Gly Trp

50 55 60

Ile Asn Asp Pro Asn Gly Leu Cys Arg Tyr Asn Gly Arg Trp His Val

65 70 75 80

Phe Tyr Gln Leu His Pro His Gly Thr Gln Trp Gly Pro Met His Trp

85 90 95

Gly His Val Ser Ser Asp Asn Met Val Asp Trp His Arg Glu Pro Ile

100 105 110

Ala Phe Ala Pro Ser Leu Glu Gln Glu Arg His Gly Val Phe Ser Gly

115 120 125

Ser Ala Val Ile Gly Asp Asp Gly Lys Pro Trp Ile Phe Tyr Thr Gly

130 135 140

His Arg Trp Ala Asn Gly Lys Asp Asn Thr Gly Gly Asp Trp Gln Val

145 150 155 160

Gln Met Leu Ala Lys Pro Asn Asp Asp Glu Leu Lys Thr Phe Thr Lys

165 170 175

Glu Gly Met Ile Ile Asp Cys Pro Thr Asp Glu Val Asp His His Phe

180 185 190

Arg Asp Pro Lys Val Trp Lys Thr Gly Asp Thr Trp Tyr Met Thr Phe

195 200 205

Gly Val Ser Ser Lys Glu His Arg Gly Gln Met Trp Leu Tyr Thr Ser
210 215 220

Ser Asp Met Val His Trp Ser Phe Asp Arg Val Leu Phe Glu His Pro
225 230 235 240

Asp Pro Asn Val Phe Met Leu Glu Cys Pro Asp Phe Phe Pro Ile Arg
245 250 255

Asp Ala Arg Gly Asn Glu Lys Trp Val Ile Gly Phe Ser Ala Met Gly
260 265 270

Ala Lys Pro Asn Gly Phe Met Asn Arg Asn Val Asn Asn Ala Gly Tyr
275 280 285

Met Val Gly Thr Trp Lys Pro Gly Glu Ser Phe Lys Pro Glu Thr Glu
290 295 300

Phe Arg Leu Trp Asp Glu Gly His Asn Phe Tyr Ala Pro Gln Ser Phe
305 310 315 320

Asn Thr Glu Gly Arg Gln Ile Met Tyr Gly Trp Met Ser Pro Phe Val
325 330 335

Ala Pro Ile Pro Met Glu Glu Asp Gly Trp Cys Gly Asn Leu Thr Leu
340 345 350

Pro Arg Glu Ile Thr Leu Gly Asp Asp Gly Asp Leu Val Thr Ala Pro
355 360 365

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

Leu Asp Leu Gly Thr Asn Gln Thr Ser Thr Ile Leu Asp Asp Asp Gly
385 390 395 400

Gly Ala Leu Glu Ile Glu Met Arg Leu Asp Leu Asn Lys Thr Thr Ala
405 410 415

Glu Arg Ala Gly Leu His Val His Ala Thr Ser Asp Gly His Tyr Thr
420 425 430

Ala Ile Val Phe Asp Ala Gln Ile Gly Gly Val Val Ile Asp Arg Gln
435 440 445

Asn Val Ala Asn Gly Asp Lys Gly Tyr Arg Val Ala Lys Leu Ser Asp

450 455 460
 Thr Glu Leu Ala Ala Asp Thr Leu Asp Leu Arg Val Phe Ile Asp Arg
 465 470 475 480

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

<210> 33

<211> 1599

<212> DNA

<213> *Saccharomyces cerevisiae*

<400

> 33

atgctttttgc aagcttttctt ttctctttttg gctgggttttg cagccaaaat atctgcatca 60
 atgacaaaacg aaactagcga tagacctttg gtccacttca cacccaacaa gggctggatg 120
 aatgacccaa atgggtttgtg gtacgatgaa aaagatgcc aatggcatct gtactttcaa 180
 tacaacccaa atgacaccgt atggggtacg ccattgtttt ggggccatgc tacttccgat 240
 gatttgacta attgggaaga tcaacccatt gctatcgctc ccaagcgtaa cgattcagg 300
 gcttttctctg gctccatggt ggttgattac aacaacacga gtgggttttt caatgatact 360
 attgatccaa gacaaagatg cgttgcgatt tggacttata acactcctga aagtgaagag 420

caatacatta gctattctct tgatggtggt tacactttta ctgaatacca aaagaaccct 480
 gttttagctg ccaactccac tcaattcaga gatccaaagg tgttctggta tgaaccttct 540
 caaaaatgga ttatgacggc tgccaaatca caagactaca aaattgaaat ttactcctct 600
 gatgacttga agtcctggaa gctagaatct gcatttgcca atgaaggttt cttaggctac 660
 caatacgaat gtccagggtt gattgaagtc ccaactgagc aagatccttc caaatcttat 720
 tgggtcatgt ttatttctat caaccagggt gcacctgctg gcggttcctt caaccaatat 780
 ttgtttggat ccttcaatgg tactcatttt gaagcgtttg acaatcaatc tagagtggta 840

gattttggta aggactacta tgccttgcaa actttcttca acactgaccc aacctacggt 900

tcagcattag gtattgcctg ggcttcaaac tgggagtaca gtgcctttgt cccaactaac 960
ccatggagat catccatgtc ttgtgtccgc aagttttctt tgaacactga atatcaagct 1020
aatccagaga ctgaattgat caatttgaaa gccgaaccaa tattgaacat tagtaatgct 1080
ggtccttggc ctgcctttgc tactaacaca actctaacta aggccaattc ttacaatgct 1140
gatttgagca actcgactgg taccctagag tttagattgg tttagctgt taacaccaca 1200
caaaccatat ccaaatccgt ctttgccgac ttatcacttt ggttcaaggg tttagaagat 1260

cctgaagaat atttgagaat gggttttgaa gtcagtgtct cttccttctt ttggaccgt 1320
ggtaactcta aggtcaagtt tgtcaaggag aacctatatt tcacaaacag aatgtctgtc 1380
aacaaccaac cattcaagtc tgagaacgac ctaagttact ataaagtgt cggcctactg 1440
gatcaaaaca tcttgaatt gtacttcaac gatggagatg tggtttctac aaatacctac 1500
ttcatgacca ccggtaacgc tctaggatct gtgaacatga ccactggtgt cgataatttg 1560
ttctacattg acaagttcca agtaagggaa gtaaaatag 1599

<210> 34

<211> 532

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 34

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys

1 5 10 15

Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His

20 25 30

Phe Thr Pro Asn Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr

35 40 45

Asp Glu Lys Asp Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn Pro Asn

50 55 60

Asp Thr Val Trp Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp

65 70 75 80

Asp Leu Thr Asn Trp Glu Asp Gln Pro Ile Ala Ile Ala Pro Lys Arg

85 90 95

Asn Asp Ser Gly Ala Phe Ser Gly Ser Met Val Val Asp Tyr Asn Asn

100 105 110

Thr Ser Gly Phe Phe Asn Asp Thr Ile Asp Pro Arg Gln Arg Cys Val

115	120	125
Ala Ile Trp Thr Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser		
130	135	140
Tyr Ser Leu Asp Gly Gly Tyr Thr Phe Thr Glu Tyr Gln Lys Asn Pro		
145	150	155
Val Leu Ala Ala Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp		
165	170	175
Tyr Glu Pro Ser Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp		
180	185	190
Tyr Lys Ile Glu Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu		
195	200	205
Glu Ser Ala Phe Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys		
210	215	220
Pro Gly Leu Ile Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr		
225	230	235
Trp Val Met Phe Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser		
245	250	255
Phe Asn Gln Tyr Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala		
260	265	270
Phe Asp Asn Gln Ser Arg Val Val Asp Phe Gly Lys Asp Tyr Tyr Ala		
275	280	285
Leu Gln Thr Phe Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly		
290	295	300
Ile Ala Trp Ala Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn		
305	310	315
Pro Trp Arg Ser Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr		
325	330	335
Glu Tyr Gln Ala Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu		
340	345	350
Pro Ile Leu Asn Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr		
355	360	365

Asn Thr Thr Leu Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn
370 375 380

Ser Thr Gly Thr Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr
385 390 395 400

Gln Thr Ile Ser Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys
405 410 415

Gly Leu Glu Asp Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser
420 425 430

Ala Ser Ser Phe Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val
435 440 445

Lys Glu Asn Pro Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro
450 455 460

Phe Lys Ser Glu Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu
465 470 475 480

Asp Gln Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser
485 490 495

Thr Asn Thr Tyr Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn
500 505 510

Met Thr Thr Gly Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val
515 520 525

Arg Glu Val Lys
530

<210> 35

<211> 1302

<212> DNA

<213> Corynebacterium glutamicum

<400> 35

gtgtgtgggg ctatgcacac agaactttcc agtttgcgcc ctgcgtacca tgtgactcct	60
ccgcagggca ggctcaatga tcccaacgga atgtacgtcg atggcgatac cctccacgtc	120
tactaccagc acgatccagg ttcccccttc gcaccaaagc gcaccggctg ggctcacacc	180
accacgccgt tgaccggacc gcagcgattg cagtggacgc acctgcccga cgctctttac	240

ccggaatgcat cciatgacct ggatggatgc tattccggtg gagccgtatt tactgacggc 300
 acacttaaac ttttctacac cggcaacctt aaaattgacg gcaagcgccg cgccacccaa 360
 aaacctgctg aagtcgagga cccaactggg ctgatgggag gcattcatcg ccgttcgcct 420
 aaaaatccgc ttatcgacgg acccgccagc ggtttcacac ccattaccg cgatcccatg 480
 atcagccctg atggtgatgg ttggaaaatg gttcttgggg cccaacgcga aaacctcacc 540
 ggtgcagcgg ttctataccg ctgcacagat ctgaaaact gggaattctc cggtgaaatc 600
 acctttgacc tcagtgatgc acaacctggt tctgctcctg atctcgttcc cggtggtctac 660

atgtgggaat gccccaaact ttttacgctt cgcatgaag aaactggcga agatctcgac 720
 gtgctgattt tctgtccaca aggattggac cgaatccacg atgaggttac tcactacgca 780
 agctctgacc agtcgggata tgctgtcggc aagcttgaag gaacgacctt ccgcgtcttg 840
 cgaggattca gcgagctgga tttcggccat gaattctacg caccgcaggt tgcagtaaac 900
 ggttctgatg cctggctcgt gggctggatg gggctgcccc cgagggatga tcaccaaca 960
 gttgcacggg aaggatgggt gcactgcctg actgtgcccc gcaagcttca tttgcgcaac 1020
 cagcgatct atcaagagct tcttctccca gagggggagt caggggtaat cagatctgta 1080

ttaggttctg aacctgtccg agtagacatc cgaggcaata tttccctcga gtgggatggt 1140
 gtccgtttgt ctgtggatcg tgggtggtgat cgtcgcgtag ctgaggtaaa acctggcgaa 1200
 ttagtgatcg cggacgataa tacagccatt gagataactg cagtgatgg acaggtttca 1260
 ttcgctttcc gggctttcaa aggtgacact attgagagat aa 1302

<210> 36

<211> 433

<212> PRT

<213> Corynebacterium glutamicum

<400> 36

Met Cys Gly Ala Met His Thr Glu Leu Ser Ser Leu Arg Pro Ala Tyr

1 5 10 15

His Val Thr Pro Pro Gln Gly Arg Leu Asn Asp Pro Asn Gly Met Tyr

20 25 30

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

35 40 45

Pro Phe Ala Pro Lys Arg Thr Gly Trp Ala His Thr Thr Thr Pro Leu

50 55 60

Thr Gly Pro Gln Arg Leu Gln Trp Thr His Leu Pro Asp Ala Leu Tyr

65					70					75					80				
Pro	Asp	Ala	Ser	Tyr	Asp	Leu	Asp	Gly	Cys	Tyr	Ser	Gly	Gly	Ala	Val				
85					90					95									
Phe	Thr	Asp	Gly	Thr	Leu	Lys	Leu	Phe	Tyr	Thr	Gly	Asn	Leu	Lys	Ile				
100					105					110									
Asp	Gly	Lys	Arg	Arg	Ala	Thr	Gln	Asn	Leu	Val	Glu	Val	Glu	Asp	Pro				
115					120					125									
Thr	Gly	Leu	Met	Gly	Gly	Ile	His	Arg	Arg	Ser	Pro	Lys	Asn	Pro	Leu				
130					135					140									
Ile	Asp	Gly	Pro	Ala	Ser	Gly	Phe	Thr	Pro	His	Tyr	Arg	Asp	Pro	Met				
145					150					155					160				
Ile	Ser	Pro	Asp	Gly	Asp	Gly	Trp	Lys	Met	Val	Leu	Gly	Ala	Gln	Arg				
165					170					175									
Glu	Asn	Leu	Thr	Gly	Ala	Ala	Val	Leu	Tyr	Arg	Ser	Thr	Asp	Leu	Glu				
180					185					190									
Asn	Trp	Glu	Phe	Ser	Gly	Glu	Ile	Thr	Phe	Asp	Leu	Ser	Asp	Ala	Gln				
195					200					205									
Pro	Gly	Ser	Ala	Pro	Asp	Leu	Val	Pro	Gly	Gly	Tyr	Met	Trp	Glu	Cys				
210					215					220									
Pro	Asn	Leu	Phe	Thr	Leu	Arg	Asp	Glu	Glu	Thr	Gly	Glu	Asp	Leu	Asp				
225					230					235					240				
Val	Leu	Ile	Phe	Cys	Pro	Gln	Gly	Leu	Asp	Arg	Ile	His	Asp	Glu	Val				
245					250					255									
Thr	His	Tyr	Ala	Ser	Ser	Asp	Gln	Cys	Gly	Tyr	Val	Val	Gly	Lys	Leu				
260					265					270									
Glu	Gly	Thr	Thr	Phe	Arg	Val	Leu	Arg	Gly	Phe	Ser	Glu	Leu	Asp	Phe				
275					280					285									
Gly	His	Glu	Phe	Tyr	Ala	Pro	Gln	Val	Ala	Val	Asn	Gly	Ser	Asp	Ala				
290					295					300									
Trp	Leu	Val	Gly	Trp	Met	Gly	Leu	Pro	Ala	Gln	Asp	Asp	His	Pro	Thr				
305					310					315					320				

Val Ala Arg Glu Gly Trp Val His Cys Leu Thr Val Pro Arg Lys Leu
325 330 335

His Leu Arg Asn His Ala Ile Tyr Gln Glu Leu Leu Leu Pro Glu Gly
340 345 350

Glu Ser Gly Val Ile Arg Ser Val Leu Gly Ser Glu Pro Val Arg Val
355 360 365

Asp Ile Arg Gly Asn Ile Ser Leu Glu Trp Asp Gly Val Arg Leu Ser
370 375 380

Val Asp Arg Gly Gly Asp Arg Arg Val Ala Glu Val Lys Pro Gly Glu
385 390 395 400

Leu Val Ile Ala Asp Asp Asn Thr Ala Ile Glu Ile Thr Ala Gly Asp
405 410 415

Gly Gln Val Ser Phe Ala Phe Arg Ala Phe Lys Gly Asp Thr Ile Glu
420 425 430

Arg

<210> 37

<211> 1473

<212> DNA

<213> Leuconostoc mesenteroides

<400> 37

atggaaattc aaaacaaagc aatgttgatc acttatgctg attcgttggg caaaaactta 60

aaagatgttc atcaagtctt gaaagaagat attggagatg cgattggtgg ggttcatttg 120

ttgcctttct tcccttcaac aggtgatcgc ggttttgcgc cagccgatta tactcgtgtt 180

gatgccgcat ttggtgattg ggcagatgtc gaagcattgg gtgaagaata ctatttgatg 240

tttgacttca tgattaacca tatttctcgt gaatcagtga tgtatcaaga ttttaagaag 300

aatcatgacg attcaaagta taaagatttc tttattcgtt gggaaaagt ctgggcaaag 360

gccggcgaaa accgtccaac acaagccgat gttgacttaa tttacaagcg taaagataag 420

gcaccaacgc aagaaatcac ttttgatgat ggcaaacag aaaacttgtg gaatactttt 480

ggtgaagaac aaattgacat tgatgttaat tcagccattg ccaaggaatt tattaagaca 540

acccttgaag acatggtaaa acatggtgct aacttgattc gtttggatgc ctttgcgtat 600

gcagttaaaa aagttgacac aaatgacttc ttcgttgagc cagaaatctg ggacactttg 660
aatgaagtac gtgaaatttt gacaccatta aaggctgaaa ttttaccaga aattcatgaa 720
cattactcaa tccttaaaaa gatcaatgat catggttact tcacctatga ctttgcatta 780
ccaatgacaa cgctttacac attgtattca ggtaagacaa atcaattggc aaagtggttg 840
aagatgtcac caatgaagca attcacaaca ttggacacgc atgatggtat tgggtgcgtt 900
gatgcccgtg atattctaac tgatgatgaa attgactacg cttctgaaca actttacaag 960

gttggcgcga atgtcaaaaa gacatattca tctgcttcat acaacaacct tgatatttac 1020
caaatlaact caacttatta ttcagcattg ggaaatgatg atgcagcata cttgttgagt 1080
cgtgtcttcc aagtctttgc gcctggaatt ccacaaattt attacgttgg tttgttggca 1140
ggtgaaaacg atatcgcgct tttggagtca actaaagaag gtcgtaatat taaccgtcat 1200
tactatacgc gtgaagaagt taagtcagaa gttaaagcag cagttgttgc taacttattg 1260
aagctattgt catggcgtaa tgaaagccct gcatttgatt tggtctggctc aatcacagtt 1320
gacacgcaa ctgatacaac aattgtggtg acacgtcaag atgaaaatgg tcaaaacaaa 1380

gctgtattaa cagccgatgc ggccaacaaa acttttgaaa tcgttgagaa tgggtcaaact 1440
gttatgagca gtgataattt gactcagaac taa 1473

<210> 38

<211> 490

<212> PRT

<213> *Leuconostoc mesenteroides*

<400> 38

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

1 5 10 15

Gly Lys Asn Leu Lys Asp Val His Gln Val Leu Lys Glu Asp Ile Gly

20 25 30

Asp Ala Ile Gly Gly Val His Leu Leu Pro Phe Phe Pro Ser Thr Gly

35 40 45

Asp Arg Gly Phe Ala Pro Ala Asp Tyr Thr Arg Val Asp Ala Ala Phe

50 55 60

Gly Asp Trp Ala Asp Val Glu Ala Leu Gly Glu Glu Tyr Tyr Leu Met

65 70 75 80

Phe Asp Phe Met Ile Asn His Ile Ser Arg Glu Ser Val Met Tyr Gln

85 90 95

Asp Phe Lys Lys Asn His Asp Asp Ser Lys Tyr Lys Asp Phe Phe Ile

100 105 110

Arg Trp Glu Lys Phe Trp Ala Lys Ala Gly Glu Asn Arg Pro Thr Gln

115 120 125

Ala Asp Val Asp Leu Ile Tyr Lys Arg Lys Asp Lys Ala Pro Thr Gln

130 135 140

Glu Ile Thr Phe Asp Asp Gly Thr Thr Glu Asn Leu Trp Asn Thr Phe

145 150 155 160

Gly Glu Glu Gln Ile Asp Ile Asp Val Asn Ser Ala Ile Ala Lys Glu

165 170 175

Phe Ile Lys Thr Thr Leu Glu Asp Met Val Lys His Gly Ala Asn Leu

180 185 190

Ile Arg Leu Asp Ala Phe Ala Tyr Ala Val Lys Lys Val Asp Thr Asn

195 200 205

Asp Phe Phe Val Glu Pro Glu Ile Trp Asp Thr Leu Asn Glu Val Arg

210 215 220

Glu Ile Leu Thr Pro Leu Lys Ala Glu Ile Leu Pro Glu Ile His Glu

225 230 235 240

His Tyr Ser Ile Pro Lys Lys Ile Asn Asp His Gly Tyr Phe Thr Tyr

245 250 255

Asp Phe Ala Leu Pro Met Thr Thr Leu Tyr Thr Leu Tyr Ser Gly Lys

260 265 270

Thr Asn Gln Leu Ala Lys Trp Leu Lys Met Ser Pro Met Lys Gln Phe

275 280 285

Thr Thr Leu Asp Thr His Asp Gly Ile Gly Val Val Asp Ala Arg Asp

290 295 300

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

305 310 315 320

Val Gly Ala Asn Val Lys Lys Thr Tyr Ser Ser Ala Ser Tyr Asn Asn

325 330 335

Leu Asp Ile Tyr Gln Ile Asn Ser Thr Tyr Tyr Ser Ala Leu Gly Asn

340 345 350
Asp Asp Ala Ala Tyr Leu Leu Ser Arg Val Phe Gln Val Phe Ala Pro

355 360 365
Gly Ile Pro Gln Ile Tyr Tyr Val Gly Leu Leu Ala Gly Glu Asn Asp

370 375 380
Ile Ala Leu Leu Glu Ser Thr Lys Glu Gly Arg Asn Ile Asn Arg His

385 390 395 400
Tyr Tyr Thr Arg Glu Glu Val Lys Ser Glu Val Lys Arg Pro Val Val

405 410 415
Ala Asn Leu Leu Lys Leu Leu Ser Trp Arg Asn Glu Ser Pro Ala Phe

420 425 430
Asp Leu Ala Gly Ser Ile Thr Val Asp Thr Pro Thr Asp Thr Thr Ile

435 440 445
Val Val Thr Arg Gln Asp Glu Asn Gly Gln Asn Lys Ala Val Leu Thr

450 455 460
Ala Asp Ala Ala Asn Lys Thr Phe Glu Ile Val Glu Asn Gly Gln Thr

465 470 475 480
Val Met Ser Ser Asp Asn Leu Thr Gln Asn

485 490

<210> 39

<211> 1515

<212> DNA

<213> Bifidobacterium adolescentis

<400> 39

atgaaaaaca aggtgcagct catcacttac gccgaccgcc ttggcgacgg caccatcaag 60

tcgatgaccg acattctgcg caccgccttc gacggcgtgt acgacggcgt tcacatcctg 120

ccgtttctca ccccgcttca cggcgccgac gcaggcttcg acccgatcga ccaccaacag 180

gtcgacgaac gtctcggcag ctgggacgac gtcgccgaac tctccaagac ccacaacatc 240

atggtcgacg ccatcgtaaa ccacatgagt tgggaatcca agcagttcca ggacgtgctg 300

gccaaggggc aggagtccga atactatccg atgttcctca ccatgagctc cgtgttcccg 360

aacggcgcca ccgaagagga cctggccggc atctaccgtc cgcgtccggg cctgccgttc 420

accactaca agttcgccgg caagacccgc ctctgtgtggg tcagcttcac cccgcagcag 480
 gttgacatcg acaccgattc cgacaagggt tgggaatacc tcatgtcgat ttctgaccag 540
 atggccgcct ctacgtcag ctacatccgc ctgcagccg tcggctatgg cgccaaggaa 600
 gccggcacca gctgcttcat gaccccgaa accttcaagc tgatctccg tctgctgag 660
 gaaggcgtca agcgcgggtt ggaaatctc atcgaagtgc actcctacta caagaagcag 720
 gtgaaatcg catccaaggt ggaccgcgtc tacgacttcg cctgcctcc gctgctgctg 780

cacgcgtga gcaccggcca cgtcgagccc gtcgccact ggaccgacat acgccgaac 840
 aacgccgtca cgtgctcga tacgcagcag ggcatcggcg tgatcgacat cggtccgac 900
 cagctcgacc gctcgtcaa gggctctgtg ccggtatgag acgtggacaa cctcgtcaac 960
 accatccacg ccaacaccca cggcgaatcc caggcagcca ctggcgccgc cgcatccaat 1020
 ctgcacctt accaggtcaa cagcacctac tattcgccgc tcgggtgcaa cgaccagcac 1080
 tacatcgccg cccgcgcggt gcagttcttc ctgccggcg tgccgcaagt ctactacgtc 1140
 ggcgcgctcg cggcaagaa cgacatggag ctgctgcgta agacgaataa cggccgcgac 1200

atcaatgcc attactactc caccgcggaa atcgacgaga acctcaagcg tccggtcgtc 1260
 aaggccctga acgcgtcgc caagttccgc aacgagctcg acgcgttcga cggcacgttc 1320
 tcgtacacca ccgatgacga cagtcctatc agcttcacct ggcgcgcgga aaccagccag 1380
 gccacgtga cgttcgagcc gaagcgcggt ctgggtgtgg acaacactac gccggtcgcc 1440
 atgttggaat gggaggattc cgcgggagac caccgttcgg atgatctgat cgccaatccg 1500
 cctgtcgtcg cctga 1515

<210> 40

<211> 504

<212> PRT

<213> Bifidobacterium adolescentis

<400> 40

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

1 5 10 15

Gly Thr Ile Lys Ser Met Thr Asp Ile Leu Arg Thr Arg Phe Asp Gly

20 25 30

Val Tyr Asp Gly Val His Ile Leu Pro Phe Phe Thr Pro Phe Asp Gly

35 40 45

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

50 55 60
 Leu Gly Ser Trp Asp Asp Val Ala Glu Leu Ser Lys Thr His Asn Ile
 65 70 75 80
 Met Val Asp Ala Ile Val Asn His Met Ser Trp Glu Ser Lys Gln Phe
 85 90 95
 Gln Asp Val Leu Ala Lys Gly Glu Glu Ser Glu Tyr Tyr Pro Met Phe
 100 105 110
 Leu Thr Met Ser Ser Val Phe Pro Asn Gly Ala Thr Glu Glu Asp Leu
 115 120 125

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

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

 Pro Leu Leu Leu His Ala Leu Ser Thr Gly His Val Glu Pro Val Ala
 260 265 270
 His Trp Thr Asp Ile Arg Pro Asn Asn Ala Val Thr Val Leu Asp Thr
 275 280 285
 His Asp Gly Ile Gly Val Ile Asp Ile Gly Ser Asp Gln Leu Asp Arg
 290 295 300

Ser Leu Lys Gly Leu Val Pro Asp Glu Asp Val Asp Asn Leu Val Asn
305 310 315 320

Thr Ile His Ala Asn Thr His Gly Glu Ser Gln Ala Ala Thr Gly Ala
325 330 335

Ala Ala Ser Asn Leu Asp Leu Tyr Gln Val Asn Ser Thr Tyr Tyr Ser
340 345 350

Ala Leu Gly Cys Asn Asp Gln His Tyr Ile Ala Ala Arg Ala Val Gln
355 360 365

Phe Phe Leu Pro Gly Val Pro Gln Val Tyr Tyr Val Gly Ala Leu Ala
370 375 380

Gly Lys Asn Asp Met Glu Leu Leu Arg Lys Thr Asn Asn Gly Arg Asp
385 390 395 400

Ile Asn Arg His Tyr Tyr Ser Thr Ala Glu Ile Asp Glu Asn Leu Lys
405 410 415

Arg Pro Val Val Lys Ala Leu Asn Ala Leu Ala Lys Phe Arg Asn Glu
420 425 430

Leu Asp Ala Phe Asp Gly Thr Phe Ser Tyr Thr Thr Asp Asp Asp Thr
435 440 445

Ser Ile Ser Phe Thr Trp Arg Gly Glu Thr Ser Gln Ala Thr Leu Thr
450 455 460

Phe Glu Pro Lys Arg Gly Leu Gly Val Asp Asn Thr Thr Pro Val Ala
465 470 475 480

Met Leu Glu Trp Glu Asp Ser Ala Gly Asp His Arg Ser Asp Asp Leu
485 490 495

Ile Ala Asn Pro Pro Val Val Ala
500

<210> 41

<211> 927

<212> DNA

<213> Agrobacterium tumefaciens

<400> 41

atgatcctgt gttgtgggtga agccctgata gacatgctgc cccggcagac gacgctgggt 60
gaggcgggct ttgcccctta cgcaggcgga gcggtcttca acacggcaat tgcgctgggg 120
cgtcttggcg tcccttcagc cttttttacc ggtctttccg acgacatgat gggcgatata 180
ctgcgggaga ccctgcgggc cagcaagggtg gatttcagct attgcgccac cctgtcgcgc 240
cccaccacca ttgcgttcgt taagctgggt gatggccatg cgacctacgc tttttacgac 300
gagaacaccg ccggccggat gatcaccgag gccgaacttc cggccttggg agcggattgc 360
gaagcgctgc atttcggcgc catcagccit attcccgaac cctgcggcag cacctatgag 420

gcgctgatga cgcgcgagca tgagaccgc gtcactcgc tcgatccgaa cattcgtccc 480
ggcttcatcc agaacaagca gtgcacatg gcccgcatcc gccgatggc ggcatgtct 540
gacatcgtca agttctcgga tgaggacctg gcgtgggtcg gtctggaagg cgacaggac 600
acgttggccc gccactgggt gcaccacggt gcaaaactcg tcgttgtcac ccgtggcgcc 660
aagggtgccg tgggttacag cgcaatctc aagggtgaag tggcctccga gcgcgtcgaa 720
gtggtcgata cgttcggcgc cggcgatacg ttcgatccg gcattcttgc ttcgctgaaa 780
atgcagggcc tgctgaccaa agcgcagggt gcttcgctga gcgaagagca gatcagaaaa 840

gctttggcgc ttggcgcgaa agccgctgcg gtcactgtct cgcgggctgg cgcaaatccg 900
cctttcgcgc atgaaatcgg tttgtga 927

<210> 42

<211> 308

<212> PRT

<213> Agrobacterium tumefaciens

<400> 42

Met Ile Leu Cys Cys Gly Glu Ala Leu Ile Asp Met Leu Pro Arg Gln

1 5 10 15

Thr Thr Leu Gly Glu Ala Gly Phe Ala Pro Tyr Ala Gly Gly Ala Val

20 25 30

Phe Asn Thr Ala Ile Ala Leu Gly Arg Leu Gly Val Pro Ser Ala Phe

35 40 45

Phe Thr Gly Leu Ser Asp Asp Met Met Gly Asp Ile Leu Arg Glu Thr

50 55 60

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

65 70 75 80

Pro Thr Thr Ile Ala Phe Val Lys Leu Val Asp Gly His Ala Thr Tyr

85 90 95
Ala Phe Tyr Asp Glu Asn Thr Ala Gly Arg Met Ile Thr Glu Ala Glu

100 105 110
Leu Pro Ala Leu Gly Ala Asp Cys Glu Ala Leu His Phe Gly Ala Ile

115 120 125
Ser Leu Ile Pro Glu Pro Cys Gly Ser Thr Tyr Glu Ala Leu Met Thr

130 135 140
Arg Glu His Glu Thr Arg Val Ile Ser Leu Asp Pro Asn Ile Arg Pro

145 150 155 160
Gly Phe Ile Gln Asn Lys Gln Ser His Met Ala Arg Ile Arg Arg Met

165 170 175
Ala Ala Met Ser Asp Ile Val Lys Phe Ser Asp Glu Asp Leu Ala Trp

180 185 190
Phe Gly Leu Glu Gly Asp Glu Asp Thr Leu Ala Arg His Trp Leu His

195 200 205
His Gly Ala Lys Leu Val Val Val Thr Arg Gly Ala Lys Gly Ala Val

210 215 220
Gly Tyr Ser Ala Asn Leu Lys Val Glu Val Ala Ser Glu Arg Val Glu

225 230 235 240
Val Val Asp Thr Val Gly Ala Gly Asp Thr Phe Asp Ala Gly Ile Leu

245 250 255
Ala Ser Leu Lys Met Gln Gly Leu Leu Thr Lys Ala Gln Val Ala Ser

260 265 270
Leu Ser Glu Glu Gln Ile Arg Lys Ala Leu Ala Leu Gly Ala Lys Ala

275 280 285
Ala Ala Val Thr Val Ser Arg Ala Gly Ala Asn Pro Pro Phe Ala His

290 295 300
Glu Ile Gly Leu

305

<210> 43

<211> 1404

<212> DNA

<213> Streptococcus mutans

<400> 43

cagctgatta tgcgtcagtt gaaaccctcg cttcttcagg aactgttgct gtaggtgata 60
gcttacttga agttaaaaa taagaaatat tatcagaaag accgtaaggt ctttttgact 120
gcttaaaaga ttcagtaaca atagtattaa agccttttgg ctaactaata ctigaaattt 180
agcaaattat gatataatgt taagtagtcc ttaagggtag attaagggtt tcaaatcca 240
aaaattgatt tggtaagtta agtaaaatat aagaggttta ttatgtctaa attatatggc 300

agcatcgaag ctggcggaac aaaatttgtc tgtgctgtag gtgatgaaaa ttttcaaatt 360
ttagaaaaag ttcagttccc aacaacaaca ccttatgaaa caatagaaaa aacagttgct 420
ttcttataaa aatttgaagc tgatttagcc agtgttgcca ttggttcttt tggccctatt 480
gatattgattc aaaattcaga cacttatggt tacattactt caacacaaaa gccaaactgg 540
gctaactgtt attttgcgg ctttaatttct aaagatttta aaattccatt ttactttacg 600
acagatgtta attcttctgc ttatggggaa acaattgtc gttcaaatgt taaaagtctg 660
gtttattata ctattggaac aggcatggga gcaggggcta ttcaaatgg cgaattcatt 720

ggcggtatgg gacatacgga agctggacac gtttacatgg ctccgcatcc caatgatgtt 780
catcatgggt ttgtaggcac ctgtcctttc cataaaggct gtttagaagg acttgcagcg 840
ggtcctagct tagaggctcg tactggtatt cgtgggtgagt taattgagca aaactcagaa 900
gtttgggata ttcaggcata ctacattgct caggcggtta ttcaagcgac tgtcctttat 960
cgtccgcaag tcattgtatt tggcggaggc gttatggcac aagaacatat gctcaatcgg 1020
gttcgtgaaa aatttacttc acttttgaat gactatcttc cagttccaga tgttaaagat 1080
tatattgtga caccagctgt tgcagaaaat ggttcagcaa cattgggaaa tctcgttta 1140

gctaaaaaga tagcagcgcg ttaattaaaa atgaattgga agattaaagc accttctaatt 1200
attcaatatt aaactgttag aatttacgtg aacgaaattt tcattttatg aggataatga 1260
agtgaatata attactcttg atttcctctg aaactagata gtggtatatt gaaaaacaga 1320
aaggagaaca ctatggaagg acctttgttt ttacaatcac aaatgcataa aaaaatctgg 1380
ggcggcaatc ggctcagaaa agaa 1404

<210> 44

<211> 293

<212> PRT

<213> Streptococcus mutans

<400> 44

Met Ser Lys Leu Tyr Gly Ser Ile Glu Ala Gly Gly Thr Lys Phe Val

1 5 10 15

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

20 25 30

Pro Thr Thr Thr Pro Tyr Glu Thr Ile Glu Lys Thr Val Ala Phe Phe

35 40 45

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

50 55 60

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

65 70 75 80

Thr Pro Lys Pro Asn Trp Ala Asn Val Asp Phe Val Gly Leu Ile Ser

85 90 95

Lys Asp Phe Lys Ile Pro Phe Tyr Phe Thr Thr Asp Val Asn Ser Ser

100 105 110

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

115 120 125

Tyr Thr Ile Gly Thr Gly Ile Gly Ala Gly Ala Ile Gln Asn Gly Glu

130 135 140

Phe Ile Gly Gly Met Gly His Thr Glu Ala Gly His Val Tyr Met Ala

145 150 155 160

Pro His Pro Asn Asp Val His His Gly Phe Val Gly Thr Cys Pro Phe

165 170 175

His Lys Gly Cys Leu Glu Gly Leu Ala Ala Gly Pro Ser Leu Glu Ala

180 185 190

Arg Thr Gly Ile Arg Gly Glu Leu Ile Glu Gln Asn Ser Glu Val Trp

195 200 205

Asp Ile Gln Ala Tyr Tyr Ile Ala Gln Ala Ala Ile Gln Ala Thr Val

210 215 220

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

225 230 235 240

Glu His Met Leu Asn Arg Val Arg Glu Lys Phe Thr Ser Leu Leu Asn

245 250 255
Asp Tyr Leu Pro Val Pro Asp Val Lys Asp Tyr Ile Val Thr Pro Ala

260 265 270
Val Ala Glu Asn Gly Ser Ala Thr Leu Gly Asn Leu Ala Leu Ala Lys

275 280 285
Lys Ile Ala Ala Arg

290

<210> 45

<211> 915

<212> DNA

<213> Escherichia coli

<400> 45

atgtcagcca aagtatgggt tttaggggat gcggtcgtag atctcttgcc agaatcagac 60

gggcgcctac tgccttgctc tggcggcgcg ccagctaacg ttgcggtggg aatcgccaga 120

ttaggcggaa caagtgggtt tataggtcgg gtgggggatg atccttttgg tgcgttaatg 180

caaagaacgc tgctaactga gggagtcgat atcacgtatc tgaagcaaga tgaatggcac 240

cggacatcca cggctgcttg cgatctgaac gatcaagggg aacgttcatt tacgtttatg 300

gtccgcccc gtgccgatct ttttttagag acgacagact tgccctgctg gcgacatggc 360

gaatggttac atctctgttc aattgcgttg tctgccgagc cttcgcgtac cagcgcattt 420

actgcatga cggcgcgacg gcatgccgga ggttttgtca gcttcgatcc taatattcgt 480

gaagatctat ggcaagacga gcatttgctc cgcttggtt tgcggcaggc gctacaactg 540

gcggatgtcg tcaagctctc ggaagaagaa tggcgactta tcagtggaaa aacacagaac 600

gatcaggata tatgcgcctt ggcaaaagag tatgagatcg ccatgctgtt ggtgactaaa 660

ggtgcagaag ggggtggtgt ctgttatcga ggacaagttc accattttgc tggaatgtct 720

gtgaattgtg tcgatatcac gggggcggga gatgcgttcg ttgccgggtt actcacaggt 780

ctgtcctcta cgggattatc tacagatgag agagaaatgc gacgaattat cgatctcgct 840

caacgttcgc gagcgttgc agtaacggcg aaaggggcaa tgacagcgct gccatgtcga 900

caagaactgg aatag 915

<210> 46

<211> 304

<212> PRT

<213> Escherichia coli

<400> 46

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

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

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

Leu Ile Ser Gly Lys Thr Gln Asn Asp Gln Asp Ile Cys Ala Leu Ala
195 200 205
Lys Glu Tyr Glu Ile Ala Met Leu Leu Val Thr Lys Gly Ala Glu Gly
210 215 220
Val Val Val Cys Tyr Arg Gly Gln Val His His Phe Ala Gly Met Ser

225 230 235 240
Val Asn Cys Val Asp Ser Thr Gly Ala Gly Asp Ala Phe Val Ala Gly
245 250 255

Leu Leu Thr Gly Leu Ser Ser Thr Gly Leu Ser Thr Asp Glu Arg Glu
260 265 270
Met Arg Arg Ile Ile Asp Leu Ala Gln Arg Cys Gly Ala Leu Ala Val
275 280 285
Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Cys Arg Gln Glu Leu Glu
290 295 300
<210> 47
<211> 924
<212> DNA
<213> Klebsiella pneumoniae
<400> 47
atgaatggaa aaatctgggt actcggcgat gcggtcgtcg atctcctgcc cgatggagag 60

ggccgcctgc tgcaatgccc cggcggcgcg ccggccaacg tggcggtcgg cgtggcgcg 120
ctcggcggtg acagcgggtt tateggcgcg gtcggcgacg atcccttcgg cgttttatg 180
cgtcacacc tggcgaggga gcaagtggat gtgaactata tgcgcctcga tgcggcgag 240
cgcacctcca cgttggtggt cgatctcgat agccacgggg agcgcacctt tacctttatg 300
gtccgtccga gcgcggacct gtctcttcag cccgaggatc tcccgcggtt tgccgcgggt 360
cagtggctgc acgtctgctc catcgtcttc agcgcggagc cgagccgcag cagcacattc 420
gcggcgatgg aggcgataaa gcgcgccggg ggctatgtca gcttcgacct caatatccgc 480

agcgacctgt ggcaggatcc gcaggacctt cgcgactgtc tcgaccgggc gctggccctc 540
gccgacgcca taaaactttc ggaagaggag ctggcggtta tcagcggcag cgacgacatc 600
gtcagcggca ccgccgggt gaacgccgcg ttccagccga cgctactgct ggtgacccag 660
ggtaaagcgg gggtcaggc cgccctgcgc gggcagggtta gccacttccc tgcccggcg 720
gtggtggcgg tcgataccac cggcgccggc gatgcctttg tcgccgggct actcgccggc 780
ctcgccgcc acggtatccc ggacaacctc gcagccctgg ctcccgacct cgcgtggcg 840
caaacctgcg gcgccctggc caccaccgcc aaaggcgcca tgaccgcct gcctacagg 900

gacgatcttc agcgtctgct gtga 924
<210> 48

<211> 307

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 48

Met Asn Gly Lys Ile Trp Val Leu Gly Asp Ala Val Val Asp Leu Leu

1 5 10 15

Pro Asp Gly Glu Gly Arg Leu Leu Gln Cys Pro Gly Gly Ala Pro Ala

20 25 30

Asn Val Ala Val Gly Val Ala Arg Leu Gly Gly Asp Ser Gly Phe Ile

35 40 45

Gly Arg Val Gly Asp Asp Pro Phe Gly Arg Phe Met Arg His Thr Leu

50 55 60

Ala Gln Glu Gln Val Asp Val Asn Tyr Met Arg Leu Asp Ala Ala Gln

65 70 75 80

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

85 90 95

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

100 105 110

Asp Leu Pro Pro Phe Ala Ala Gly Gln Trp Leu His Val Cys Ser Ile

115 120 125

Ala Leu Ser Ala Glu Pro Ser Arg Ser Thr Thr Phe Ala Ala Met Glu

130 135 140

Ala Ile Lys Arg Ala Gly Gly Tyr Val Ser Phe Asp Pro Asn Ile Arg

145 150 155 160

Ser Asp Leu Trp Gln Asp Pro Gln Asp Leu Arg Asp Cys Leu Asp Arg

165 170 175

Ala Leu Ala Leu Ala Asp Ala Ile Lys Leu Ser Glu Glu Glu Leu Ala

180 185 190

Phe Ile Ser Gly Ser Asp Asp Ile Val Ser Gly Thr Ala Arg Leu Asn

195 200 205

Ala Arg Phe Gln Pro Thr Leu Leu Leu Val Thr Gln Gly Lys Ala Gly

210 215 220

Val Gln Ala Ala Leu Arg Gly Gln Val Ser His Phe Pro Ala Arg Pro
225 230 235 240

Val Val Ala Val Asp Thr Thr Gly Ala Gly Asp Ala Phe Val Ala Gly
245 250 255

Leu Leu Ala Gly Leu Ala Ala His Gly Ile Pro Asp Asn Leu Ala Ala
260 265 270

Leu Ala Pro Asp Leu Ala Leu Ala Gln Thr Cys Gly Ala Leu Ala Thr
275 280 285

Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Tyr Arg Asp Asp Leu Gln
290 295 300

Arg Ser Leu

305

<210> 49

<211> 915

<212> DNA

<213> Escherichia coli

<400> 49

atgtcagcca aagtatgggt ttaggggat gcggtcgtag atctcttgcc agaatcagac 60
gggcgcctac tgccttgcc tggcggcgcg ccagctaacg ttgcggtggg aatcgccaga 120
ttaggcggaa caagtgggtt tataggtcgg gtgggggatg atccttttgg tgcgttaatg 180
caaagaacgc tgctaactga gggagtcgat atcacgtatc tgaagcaaga tgaatggcac 240
cggacatcca cggtgcttgt cgatctgaac gatcaagggg aacgttcatt tacgtttatg 300
gtccgccccca gtgccgatct ttttttagag acgacagact tgcctgctg gcgacatggc 360

gaatggttac atctctgttc aattgcgttg tctgccgagc cttcgcgtac cagcgcattt 420
actgcgatga cggcgatccg gcatgccgga ggttttgtca gcttcgatcc taatattcgt 480
gaagatctat ggcaagacga gcatttgctc cgcttgtgtt tgcggcaggc gctacaactg 540
gcggatgtcg tcaagctctc ggaagaagaa tggcgactta tcagtggaaa aacacagaac 600
gatcaggata tatgcgcctt ggcaaaagag tatgagatcg ccatgctgtt ggtgactaaa 660
ggtgcagaag ggggtgggtg ctgttatcga ggacaagttc accattttgc tggaatgtct 720
gtgaattgtg tcgatagcac gggggcggga gatgcgttcg ttgccgggtt actcacaggt 780

ctgtcctcta cgggattatc tacagatgag agagaaatgc gacgaattat cgatctcgct 840

caacgttgcg gagcgcttgc agtaacggcg aaaggggcaa tgacagcgct gccatgtcga 900
caagaactgg aatag 915

<210> 50
<211> 304
<212> PRT
<213> Escherichia coli
<400> 50

Met Ser Ala Lys Val Trp Val Leu Gly Asp Ala Val Val Asp Leu Leu
1 5 10 15
Pro Glu Ser Asp Gly Arg Leu Leu Pro Cys Pro Gly Gly Ala Pro Ala
20 25 30

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

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

Glu Asp Leu Trp Gln Asp Glu His Leu Leu Arg Leu Cys Leu Arg Gln
165 170 175
Ala Leu Gln Leu Ala Asp Val Val Lys Leu Ser Glu Glu Glu Trp Arg
180 185 190
Leu Ile Ser Gly Lys Thr Gln Asn Asp Gln Asp Ile Cys Ala Leu Ala

195 200 205
 Lys Glu Tyr Glu Ile Ala Met Leu Leu Val Thr Lys Gly Ala Glu Gly
 210 215 220
 Val Val Val Cys Tyr Arg Gly Gln Val His His Phe Ala Gly Met Ser
 225 230 235 240
 Val Asn Cys Val Asp Ser Thr Gly Ala Gly Asp Ala Phe Val Ala Gly
 245 250 255
 Leu Leu Thr Gly Leu Ser Ser Thr Gly Leu Ser Thr Asp Glu Arg Glu
 260 265 270
 Met Arg Arg Ile Ile Asp Leu Ala Gln Arg Cys Gly Ala Leu Ala Val
 275 280 285

Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Cys Arg Gln Glu Leu Glu
 290 295 300

<210> 51

<211> 879

<212> DNA

<213> Enterococcus faecalis

<400> 51

atgacagaaa aacttttagg aagtatcgaa gccggtggca caaaatttgt atgtggcggt 60
 gggacagatg atttgaccat cgtagaacgt gtcagttttc ccacaacaac cccagaagaa 120
 acaatgaaaa aagtaataga atttttccaa caatatacctt taaaagcgat tgggattggt 180
 tcatttggtc cgattgatat tcacgttgat tctcctacgt atggttatat cacttctaca 240

ccaaaattag ctggcgtaa ctttgacttg ttaggaacta tgaacaaca ttttgatgtg 300
 ccaatggcctt ggacaacgga tgtgaatgct gcggcatatg gtgagtatgt tgctggaaat 360
 gggcaacata catctagtig tgtatattat acaattggaa ctggtgttgg cgctggagcg 420
 attcaaaacg gtgagtttat tgaaggcttt agccaccag aaatggggca tgcgttagtt 480
 cgtcgtcatc ctgaagatac gtatgcagga aattgtcctt atcatggaga ttgtttagaa 540
 gggattgcag caggaccagc agttgaaggt cgttctggta aaaaaggaca tttattggaa 600
 gaggatcata aaacttggga attagaagct tattatttag cgcaagcggc gtacaatacg 660

actttattat tagcgccaga agtgatcatt ttaggtggcg gcgtcatgaa acaacgtcat 720
 ttgatgccga aagttcgtga aaaatttgct gaattagtca atggatatgt ggaaacaccg 780

cctttagaaa aatacttggg gacgcctctt ttagaagata atccaggaac aatcggttgc 840
 ttgaccttgg caaaaaaagc tttaatggct caaaaataa 879
 <210> 52
 <211> 292
 <212> PRT
 <213> Enterococcus faecalis
 <400> 52
 Met Thr Glu Lys Leu Leu Gly Ser Ile Glu Ala Gly Gly Thr Lys Phe
 1 5 10 15
 Val Cys Gly Val Gly Thr Asp Asp Leu Thr Ile Val Glu Arg Val Ser
 20 25 30
 Phe Pro Thr Thr Thr Pro Glu Glu Thr Met Lys Lys Val Ile Glu Phe
 35 40 45
 Phe Gln Gln Tyr Pro Leu Lys Ala Ile Gly Ile Gly Ser Phe Gly Pro
 50 55 60
 Ile Asp Ile His Val Asp Ser Pro Thr Tyr Gly Tyr Ile Thr Ser Thr
 65 70 75 80
 Pro Lys Leu Ala Trp Arg Asn Phe Asp Leu Leu Gly Thr Met Lys Gln
 85 90 95
 His Phe Asp Val Pro Met Ala Trp Thr Thr Asp Val Asn Ala Ala Ala
 100 105 110
 Tyr Gly Glu Tyr Val Ala Gly Asn Gly Gln His Thr Ser Ser Cys Val
 115 120 125
 Tyr Tyr Thr Ile Gly Thr Gly Val Gly Ala Gly Ala Ile Gln Asn Gly
 130 135 140
 Glu Phe Ile Glu Gly Phe Ser His Pro Glu Met Gly His Ala Leu Val
 145 150 155 160
 Arg Arg His Pro Glu Asp Thr Tyr Ala Gly Asn Cys Pro Tyr His Gly
 165 170 175
 Asp Cys Leu Glu Gly Ile Ala Ala Gly Pro Ala Val Glu Gly Arg Ser
 180 185 190
 Gly Lys Lys Gly His Leu Leu Glu Glu Asp His Lys Thr Trp Glu Leu

195	200	205	
Glu Ala Tyr Tyr Leu Ala Gln Ala Ala Tyr Asn Thr Thr Leu Leu Leu			
210	215	220	
Ala Pro Glu Val Ile Ile Leu Gly Gly Gly Val Met Lys Gln Arg His			
225	230	235	240
Leu Met Pro Lys Val Arg Glu Lys Phe Ala Glu Leu Val Asn Gly Tyr			
245	250	255	
Val Glu Thr Pro Pro Leu Glu Lys Tyr Leu Val Thr Pro Leu Leu Glu			
260	265	270	
Asp Asn Pro Gly Thr Ile Gly Cys Phe Ala Leu Ala Lys Lys Ala Leu			
275	280	285	
Met Ala Gln Lys			
290			
<210> 53			
<211> 1458			
<212> DNA			
<213> Saccharomyces cerevisiae			
<400> 53			
atggttcatt taggtccaaa gaaaccacag gctagaaagg gttccatggc tgatgtgccc	60		
aaggaattga tggatgaaat tcatcagttg gaagatatgt ttacagttga cagcgagacc	120		
ttgagaaagg ttgttaagca ctttatcgac gaattgaata aaggtttgac aaagaaggga	180		
ggtaacattc caatgattcc cggttgggtc atggaattcc caacaggtaa agaatctggt	240		
aactatttgg ccattgattt gggtggtact aacttaagag tcgtgttggt caagttgagc	300		
ggtaaccata cctttgacac cactcaatcc aagtataaac taccacatga catgagaacc	360		
actaagcacc aagaggagtt atggtccttt attgccgact ctttgaagga ctttatggtc	420		
gagcaagaat tgctaaacac caaggacacc ttaccattag gtttcacctt ctcgtaccca	480		
gcttcccaaa acaagattaa cgaaggtatt ttgcaaagat ggaccaaggg tttcgatatt	540		
ccaaatgtcg aaggccacga tgtctgcca ttgctacaaa acgaaatttc caagagagag	600		
ttgcctattg aaattgtagc attgattaat gatactgttg gtactttaat tgcctcatat	660		
tactactgacc cagagactaa gatgggtgtg attttcggta ctggtgtcaa cgggtgctttc	720		
tatgatgttg tttccgatat cgaaaagttg gagggcaaat tagcagacga tattccaagt	780		

aactctccaa tggctatcaa ttgtgaatat ggttccttcg ataatgaaca ttgtgtcttg 840
ccaagaacca agtacgatgt tgctgtcgac gaacaatctc caagacctgg tcaacaagct 900
tttgaaaaga tgacctccgg ttactacttg ggtgaattgt tgcgtctagt gttacttgaa 960
ttaaacgaga agggcttgat gttgaaggat caagatctaa gcaagttgaa acaaccatac 1020
atcatggata cctcctaccc agcaagaatc gaggatgata catttgaaaa cttggaagat 1080

actgatgaca tcttccaaaa ggactttggt gtcaagacca ctctgccaga acgtaagttg 1140
attagaagac ttigtgaatt gatcggtacc agagctgcta gattagctgt ttgtggtatt 1200
gccgctatctt gccaaaagag aggttacaag actgggcaca ttgccgctga cggttctgtc 1260
tataacaaat acccaggttt caaggaagcc gccgctaagg gtttgagaga tatctatgga 1320
tggactggtg acgcaagcaa agatccaatt acgattgttc cagctgagga tggttcaggt 1380
gcaggtgctg ctgttattgc tgcattgtcc gaaaaaagaa ttgccgaagg taagtctctt 1440
ggtatcattg gcgcttaa 1458

<210> 54

<211> 485

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 54

Met Val His Leu Gly Pro Lys Lys Pro Gln Ala Arg Lys Gly Ser Met

1 5 10 15

Ala Asp Val Pro Lys Glu Leu Met Asp Glu Ile His Gln Leu Glu Asp

20 25 30

Met Phe Thr Val Asp Ser Glu Thr Leu Arg Lys Val Val Lys His Phe

35 40 45

Ile Asp Glu Leu Asn Lys Gly Leu Thr Lys Lys Gly Gly Asn Ile Pro

50 55 60

Met Ile Pro Gly Trp Val Met Glu Phe Pro Thr Gly Lys Glu Ser Gly

65 70 75 80

Asn Tyr Leu Ala Ile Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu

85 90 95

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

100 105 110

Lys Leu Pro His Asp Met Arg Thr Thr Lys His Gln Glu Glu Leu Trp

115 120 125

Ser Phe Ile Ala Asp Ser Leu Lys Asp Phe Met Val Glu Gln Glu Leu

130 135 140

Leu Asn Thr Lys Asp Thr Leu Pro Leu Gly Phe Thr Phe Ser Tyr Pro

145 150 155 160

Ala Ser Gln Asn Lys Ile Asn Glu Gly Ile Leu Gln Arg Trp Thr Lys

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220

Glu Thr Lys Met Gly Val Ile Phe Gly Thr Gly Val Asn Gly Ala Phe

225 230 235 240

Tyr Asp Val Val Ser Asp Ile Glu Lys Leu Glu Gly Lys Leu Ala Asp

245 250 255

Asp Ile Pro Ser Asn Ser Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser

260 265 270

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

275 280 285

Val Asp Glu Gln Ser Pro Arg Pro Gly Gln Gln Ala Phe Glu Lys Met

290 295 300

Thr Ser Gly Tyr Tyr Leu Gly Glu Leu Leu Arg Leu Val Leu Leu Glu

305 310 315 320

Leu Asn Glu Lys Gly Leu Met Leu Lys Asp Gln Asp Leu Ser Lys Leu

325 330 335

Lys Gln Pro Tyr Ile Met Asp Thr Ser Tyr Pro Ala Arg Ile Glu Asp

340 345 350

Asp Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Ile Phe Gln Lys Asp

355 360 365
Phe Gly Val Lys Thr Thr Leu Pro Glu Arg Lys Leu Ile Arg Arg Leu

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

435 440 445
Pro Ile Thr Ile Val Pro Ala Glu Asp Gly Ser Gly Ala Gly Ala Ala
450 455 460
Val Ile Ala Ala Leu Ser Glu Lys Arg Ile Ala Glu Gly Lys Ser Leu
465 470 475 480
Gly Ile Ile Gly Ala
485

<210> 55

<211> 1461

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 55

atggttcatt taggtccaaa aaaaccacaa gccagaaagg gttccatggc cgatgtgcca 60

aaggaattga tgcaacaaat tgagaatttt gaaaaaattt tcaactgttc aactgaaact 120

ttacaagccg ttaccaagca cttcatttcc gaattggaaa agggtttgtc caagaagggt 180

ggtaacattc caatgattcc aggttgggtt atggatttcc caactggtaa ggaatccggt 240

gatttcttgg ccattgattt ggggtgtacc aacttgagag ttgtcttagt caagtgggc 300

ggtgaccgta ctttgacac cactcaatct aagtacagat taccagatgc tatgagaact 360

actcaaaatc cagacgaatt gtgggaattt attgccgact ctttgaaagc ttttattgat 420

gagcaattcc cacaaggtat ctctgagcca attccattgg gtttcacctt ttctttccca 480

gcttctcaaa acaaaatcaa tgaaggtatc ttgcaaagat ggactaaagg ttttgatatt 540

ccaaacattg aaaaccacga tgttgttcca atgttgcaaa agcaaatcac taagaggaat 600
atcccaattg aagtgttgc ttgtataaac gacactaccg gtacttttgt tgcttcttac 660
tacactgacc cagaaactaa gatgggtgtt atcttcggta ctggtgtcaa tgggtgcttac 720
tacgatgttt gticcgatat cgaaaagcta caaggaaaac tatctgatga cattccacca 780
tctgtctcaa tggccatcaa ctgtgaatac ggttccttcg ataatgaaca tgcgttttg 840
ccaagaacta aatacgatat caccattgat gaagaatctc caagaccagg ccaacaaacc 900

tttgaaaaaa tgccttctgg ttactactta ggtgaaatth tgcgtttggc cttgatggac 960
atgtacaaac aaggtttcat cttcaagaac caagacttgt ctaagttcga caagcctttc 1020
gtcatggaca ctctttacc agccagaatc gaggaagatc cattcgagaa cctagaagat 1080
accgatgact tgttccaaaa tgagttcggc atcaacacta ctgttcaaga acgtaaattg 1140
atcagacgtt tatctgaatt gattggtgct agagctgcta gattgtccgt ttgtggtatt 1200
gtgcttatct gtcaaaagag aggttacaag accggtcaca tcgctgcaga cggttccgtt 1260
tacaacagat acccaggttt caaagaaaag gctgccaatg ctttgaagga catttacggc 1320

tggactcaaa cctcactaga cgactacceca atcaagattg ttctgtctga agatggttcc 1380
ggtgctggtg ccgctgttat tgctgctttg gcccaaaaaa gaattgctga aggttaagtc 1440
gttggtatca tcggtgctta a 1461

<210> 56

<211> 486

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 56

Met Val His Leu Gly Pro Lys Lys Pro Gln Ala Arg Lys Gly Ser Met

1 5 10 15

Ala Asp Val Pro Lys Glu Leu Met Gln Gln Ile Glu Asn Phe Glu Lys

20 25 30

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

35 40 45

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

50 55 60

Met Ile Pro Gly Trp Val Met Asp Phe Pro Thr Gly Lys Glu Ser Gly

65 70 75 80

Asp Phe Leu Ala Ile Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu

85 90 95

Val Lys Leu Gly Gly Asp Arg Thr Phe Asp Thr Thr Gln Ser Lys Tyr

100 105 110

Arg Leu Pro Asp Ala Met Arg Thr Thr Gln Asn Pro Asp Glu Leu Trp

115 120 125

Glu Phe Ile Ala Asp Ser Leu Lys Ala Phe Ile Asp Glu Gln Phe Pro

130 135 140

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

145 150 155 160

Ala Ser Gln Asn Lys Ile Asn Glu Gly Ile Leu Gln Arg Trp Thr Lys

165 170 175

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

180 185 190

Gln Lys Gln Ile Thr Lys Arg Asn Ile Pro Ile Glu Val Val Ala Leu

195 200 205

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

210 215 220

Glu Thr Lys Met Gly Val Ile Phe Gly Thr Gly Val Asn Gly Ala Tyr

225 230 235 240

Tyr Asp Val Cys Ser Asp Ile Glu Lys Leu Gln Gly Lys Leu Ser Asp

245 250 255

Asp Ile Pro Pro Ser Ala Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser

260 265 270

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

275 280 285

Ile Asp Glu Glu Ser Pro Arg Pro Gly Gln Gln Thr Phe Glu Lys Met

290 295 300

Ser Ser Gly Tyr Tyr Leu Gly Glu Ile Leu Arg Leu Ala Leu Met Asp

305 310 315 320

Met Tyr Lys Gln Gly Phe Ile Phe Lys Asn Gln Asp Leu Ser Lys Phe

325 330 335
Asp Lys Pro Phe Val Met Asp Thr Ser Tyr Pro Ala Arg Ile Glu Glu

340 345 350
Asp Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Leu Phe Gln Asn Glu

355 360 365
Phe Gly Ile Asn Thr Thr Val Gln Glu Arg Lys Leu Ile Arg Arg Leu

370 375 380
Ser Glu Leu Ile Gly Ala Arg Ala Ala Arg Leu Ser Val Cys Gly Ile

385 390 395 400
Ala Ala Ile Cys Gln Lys Arg Gly Tyr Lys Thr Gly His Ile Ala Ala

405 410 415
Asp Gly Ser Val Tyr Asn Arg Tyr Pro Gly Phe Lys Glu Lys Ala Ala

420 425 430
Asn Ala Leu Lys Asp Ile Tyr Gly Trp Thr Gln Thr Ser Leu Asp Asp

435 440 445
Tyr Pro Ile Lys Ile Val Pro Ala Glu Asp Gly Ser Gly Ala Gly Ala

450 455 460
Ala Val Ile Ala Ala Leu Ala Gln Lys Arg Ile Ala Glu Gly Lys Ser

465 470 475 480
Val Gly Ile Ile Gly Ala

485

<210> 57

<211> 1164

<212> DNA

<213> *Klebsiella pneumoniae*

<400> 57

atgagctatc gtatgtttga ttatctggtg ccaaacttta acttttttgg ccccaacgcc	60
atttcgtag tcggcgaacg ctgccagctg ctggggggga aaaaagccct gctggtcacc	120
gacaaaggcc tcggggcaat taaagatggc gcagtggaca aaacctgca ttatctgcgg	180
gaggccggga tcgaggtggc gatctttgac ggcgtcgagc cgaacccgaa agacaccaac	240
gtgcgcgacg gcctcgccgt gtttcgccgc gaacagtgcg acatcatcgt caccgtgggc	300

ggcggcagcc cgcacgattg cggcaaaggc atcggcacgc cgcacacca tgaggcgat 360
ctgtaccagt atgccggaat cgagaccctg accaaccgcg tgccgcctat cgtcgcggtc 420
aataccaccg ccggcaccgc cagcgaggtc acccgccact gcgtcctgac caacaccgaa 480
accaaagtga agtttgtgat cgtcagctgg cgcaacctgc cgtcggtctc tatcaacgat 540
ccgctgctga tgatcggtaa accggccgcc ctgaccgcgg cgaccgggat ggatgcctg 600
acccacgcg tagaggccta tatctccaaa gacgctaacc cggtagacga cggcccgcc 660
atgcaggcga tccgcctcat cgcccgcaac ctgcgccagg ccgtggccct cggcagcaat 720

ctgcaggcgc gggaaaacat ggcctatgcc tctctgctgg ccgggatggc tttcaataac 780
gccaaacctg gctacgtgca cgccatggcg caccagctgg gcggcctgta cgacatgccg 840
cacggcgtgg ccaacgctgt cctgctgccg catgtggccc gctacaacct gatcgccaac 900
ccggagaaat tcgccgatat cgctgaactg atgggcgaaa atatcaccgg actgtccact 960
ctcgacgcgg cgaaaaaagc catgccgct atcacgcgtc tgtcgatgga tatcggtatt 1020
ccgcagcatc tgcgcgatct gggagtaaaa gaggccgact tcccctacat ggcggagatg 1080
gtctgaaag acggcaatgc gttctgaac ccgcgtaaag gcaacgagca ggagattgcc 1140

gcgattttcc gccaggcatt ctga 1164

<210> 58

<211> 387

<212> PRT

<213> Klebsiella pneumoniae

<400> 58

Met Ser Tyr Arg Met Phe Asp Tyr Leu Val Pro Asn Val Asn Phe Phe

1 5 10 15

Gly Pro Asn Ala Ile Ser Val Val Gly Glu Arg Cys Gln Leu Leu Gly

20 25 30

Gly Lys Lys Ala Leu Leu Val Thr Asp Lys Gly Leu Arg Ala Ile Lys

35 40 45

Asp Gly Ala Val Asp Lys Thr Leu His Tyr Leu Arg Glu Ala Gly Ile

50 55 60

Glu Val Ala Ile Phe Asp Gly Val Glu Pro Asn Pro Lys Asp Thr Asn

65 70 75 80

Val Arg Asp Gly Leu Ala Val Phe Arg Arg Glu Gln Cys Asp Ile Ile

85 90 95

Val Thr Val Gly Gly Gly Ser Pro His Asp Cys Gly Lys Gly Ile Gly
100 105 110

Ile Ala Ala Thr His Glu Gly Asp Leu Tyr Gln Tyr Ala Gly Ile Glu
115 120 125

Thr Leu Thr Asn Pro Leu Pro Pro Ile Val Ala Val Asn Thr Thr Ala
130 135 140

Gly Thr Ala Ser Glu Val Thr Arg His Cys Val Leu Thr Asn Thr Glu
145 150 155 160

Thr Lys Val Lys Phe Val Ile Val Ser Trp Arg Asn Leu Pro Ser Val
165 170 175

Ser Ile Asn Asp Pro Leu Leu Met Ile Gly Lys Pro Ala Ala Leu Thr
180 185 190

Ala Ala Thr Gly Met Asp Ala Leu Thr His Ala Val Glu Ala Tyr Ile
195 200 205

Ser Lys Asp Ala Asn Pro Val Thr Asp Ala Ala Ala Met Gln Ala Ile
210 215 220

Arg Leu Ile Ala Arg Asn Leu Arg Gln Ala Val Ala Leu Gly Ser Asn
225 230 235 240

Leu Gln Ala Arg Glu Asn Met Ala Tyr Ala Ser Leu Leu Ala Gly Met
245 250 255

Ala Phe Asn Asn Ala Asn Leu Gly Tyr Val His Ala Met Ala His Gln
260 265 270

Leu Gly Gly Leu Tyr Asp Met Pro His Gly Val Ala Asn Ala Val Leu
275 280 285

Leu Pro His Val Ala Arg Tyr Asn Leu Ile Ala Asn Pro Glu Lys Phe
290 295 300

Ala Asp Ile Ala Glu Leu Met Gly Glu Asn Ile Thr Gly Leu Ser Thr
305 310 315 320

Leu Asp Ala Ala Glu Lys Ala Ile Ala Ala Ile Thr Arg Leu Ser Met
325 330 335

Asp Ile Gly Ile Pro Gln His Leu Arg Asp Leu Gly Val Lys Glu Ala

340	345	350	
Asp Phe Pro Tyr Met Ala Glu Met Ala Leu Lys Asp Gly Asn Ala Phe			
355	360	365	
Ser Asn Pro Arg Lys Gly Asn Glu Gln Glu Ile Ala Ala Ile Phe Arg			
370	375	380	
Gln Ala Phe			
385			
<210> 59			
<211> 1824			
<212> DNA			
<213> <i>Klebsiella pneumoniae</i>			
<220><221> CDS			
<222> (1)..(1824)			
<400> 59			
atg ccg tta ata gcc ggg att gat atc ggc aac gcc acc acc gag gtg	48		
Met Pro Leu Ile Ala Gly Ile Asp Ile Gly Asn Ala Thr Thr Glu Val			
1 5 10 15			
gcg ctg gcg tcc gac tac ccg cag gcg agg gcg ttt gtt gcc agc ggg	96		
Ala Leu Ala Ser Asp Tyr Pro Gln Ala Arg Ala Phe Val Ala Ser Gly			
20 25 30			
atc gtc gcg acg acg ggc atg aaa ggg acg cgg gac aat atc gcc ggg	144		
Ile Val Ala Thr Thr Gly Met Lys Gly Thr Arg Asp Asn Ile Ala Gly			
35 40 45			
acc ctc gcc gcg ctg gag cag gcc ctg gcg aaa aca ccg tgg tcg atg	192		
Thr Leu Ala Ala Leu Glu Gln Ala Leu Ala Lys Thr Pro Trp Ser Met			
50 55 60			
agc gat gtc tct cgc atc tat ctt aac gaa gcc gcg ccg gtg att ggc	240		
Ser Asp Val Ser Arg Ile Tyr Leu Asn Glu Ala Ala Pro Val Ile Gly			
65 70 75 80			
gat gtg gcg atg gag acc atc acc gag acc att atc acc gaa tcg acc	288		
Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser Thr			

85	90	95	
atg atc ggt cat aac ccg cag acg ccg ggc ggg gtg ggc gtt ggc gtg			336
Met Ile Gly His Asn Pro Gln Thr Pro Gly Gly Val Gly Val Gly Val			
100	105	110	
ggg acg act atc gcc ctc ggg cgg ctg gcg acg ctg ccg gcg gcg cag			384
Gly Thr Thr Ile Ala Leu Gly Arg Leu Ala Thr Leu Pro Ala Ala Gln			
115	120	125	
tat gcc gag ggg tgg atc gta ctg att gac gac gcc gtc gat ttc ctt			432
Tyr Ala Glu Gly Trp Ile Val Leu Ile Asp Asp Ala Val Asp Phe Leu			
130	135	140	
gac gcc gtg tgg tgg ctc aat gag gcg ctc gac cgg ggg atc aac gtg			480
Asp Ala Val Trp Trp Leu Asn Glu Ala Leu Asp Arg Gly Ile Asn Val			
145	150	155	160
gtg gcg gcg atc ctc aaa aag gac gac ggc gtg ctg gtg aac aac cgc			528
Val Ala Ala Ile Leu Lys Lys Asp Asp Gly Val Leu Val Asn Asn Arg			
165	170	175	
ctg cgt aaa acc ctg ccg gtg gtg gat gaa gtg acg ctg ctg gag cag			576
Leu Arg Lys Thr Leu Pro Val Val Asp Glu Val Thr Leu Leu Glu Gln			
180	185	190	
gtc ccc gag ggg gta atg gcg gcg gtg gaa gtg gcc gcg ccg ggc cag			624
Val Pro Glu Gly Val Met Ala Ala Val Glu Val Ala Ala Pro Gly Gln			
195	200	205	
gtg gtg cgg atc ctg tcg aat ccc tac ggg atc gcc acc ttc ttc ggg			672
Val Val Arg Ile Leu Ser Asn Pro Tyr Gly Ile Ala Thr Phe Phe Gly			
210	215	220	
cta agc ccg gaa gag acc cag gcc atc gtc ccc atc gcc cgc gcc ctg			720
Leu Ser Pro Glu Glu Thr Gln Ala Ile Val Pro Ile Ala Arg Ala Leu			
225	230	235	240
att ggc aac cgt tcc gcg gtg gtg ctc aag acc ccg cag ggg gat gtg			768
Ile Gly Asn Arg Ser Ala Val Val Leu Lys Thr Pro Gln Gly Asp Val			
245	250	255	

cag tcg cgg gtg atc ccg gcg ggc aac ctc tac att agc ggc gaa aag	816
Gln Ser Arg Val Ile Pro Ala Gly Asn Leu Tyr Ile Ser Gly Glu Lys	
260 265 270	
cgc cgc gga gag gcc gat gtc gcc gag ggc gcg gaa gcc atc atg cag	864
Arg Arg Gly Glu Ala Asp Val Ala Glu Gly Ala Glu Ala Ile Met Gln	
275 280 285	
gcg atg agc gcc tgc gct ccg gta cgc gac atc cgc ggc gaa ccg ggc	912
Ala Met Ser Ala Cys Ala Pro Val Arg Asp Ile Arg Gly Glu Pro Gly	
290 295 300	
acc cac gcc ggc ggc atg ctt gag cgg gtg cgc aag gta atg gcg tcc	960
Thr His Ala Gly Gly Met Leu Glu Arg Val Arg Lys Val Met Ala Ser	
305 310 315 320	
ctg acc ggc cat gag atg agc gcg ata tac atc cag gat ctg ctg gcg	1008
Leu Thr Gly His Glu Met Ser Ala Ile Tyr Ile Gln Asp Leu Leu Ala	
325 330 335	
gtg gat acg ttt att ccg cgc aag gtg cag ggc ggg atg gcc ggc gag	1056
Val Asp Thr Phe Ile Pro Arg Lys Val Gln Gly Gly Met Ala Gly Glu	
340 345 350	
tgc gcc atg gag aat gcc gtc ggg atg gcg gcg atg gtg aaa gcg gat	1104
Cys Ala Met Glu Asn Ala Val Gly Met Ala Ala Met Val Lys Ala Asp	
355 360 365	
cgt ctg caa atg cag gtt atc gcc cgc gaa ctg agc gcc cga ctg cag	1152
Arg Leu Gln Met Gln Val Ile Ala Arg Glu Leu Ser Ala Arg Leu Gln	
370 375 380	
acc gag gtg gtg gtg ggc ggc gtg gag gcc aac atg gcc atc gcc ggg	1200
Thr Glu Val Val Val Gly Gly Val Glu Ala Asn Met Ala Ile Ala Gly	
385 390 395 400	
gcg tta acc act ccc ggc tgt gcg gcg ccg ctg gcg atc ctc gac ctc	1248
Ala Leu Thr Thr Pro Gly Cys Ala Ala Pro Leu Ala Ile Leu Asp Leu	
405 410 415	
ggc gcc ggc tcg acg gat gcg gcg atc gtc aac gcg gag ggg cag ata	1296

Gly Ala Gly Ser Thr Asp Ala Ala Ile Val Asn Ala Glu Gly Gln Ile

420 425 430

acg gcg gtc cat ctc gcc ggg gcg ggg aat atg gtc agc ctg ttg att 1344

Thr Ala Val His Leu Ala Gly Ala Gly Asn Met Val Ser Leu Leu Ile

435 440 445

aaa acc gag ctg ggc ctc gag gat ctt tcg ctg gcg gaa gcg ata aaa 1392

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

450 455 460

aaa tac ccg ctg gcc aaa gtg gaa agc ctg ttc agt att cgt cac gag 1440

Lys Tyr Pro Leu Ala Lys Val Glu Ser Leu Phe Ser Ile Arg His Glu

465 470 475 480

aat ggc gcg gtg gag ttc ttt cgg gaa gcc ctc agc ccg gcg gtg ttc 1488

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

485 490 495

gcc aaa gtg gtg tac atc aag gag ggc gaa ctg gtg ccg atc gat aac 1536

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

500 505 510

gcc agc ccg ctg gaa aaa att cgt ctc gtg cgc cgg cag gcg aaa gag 1584

Ala Ser Pro Leu Glu Lys Ile Arg Leu Val Arg Arg Gln Ala Lys Glu

515 520 525

aaa gtg ttt gtc acc aac tgc ctg cgc gcg ctg cgc cag gtc tca ccc 1632

Lys Val Phe Val Thr Asn Cys Leu Arg Ala Leu Arg Gln Val Ser Pro

530 535 540

ggc ggt tcc att cgc gat atc gcc ttt gtg gtg ctg gtg ggc ggc tca 1680

Gly Gly Ser Ile Arg Asp Ile Ala Phe Val Val Leu Val Gly Gly Ser

545 550 555 560

tcg ctg gac ttt gag atc ccg cag ctt atc acg gaa gcc ttg tcg cac 1728

Ser Leu Asp Phe Glu Ile Pro Gln Leu Ile Thr Glu Ala Leu Ser His

565 570 575

tat ggc gtg gtc gcc ggg cag ggc aat att cgg gga aca gaa ggg ccg 1776
Tyr Gly Val Val Ala Gly Gln Gly Asn Ile Arg Gly Thr Glu Gly Pro
580 585 590
cgc aat gcg gtc gcc acc ggg ctg cta ctg gcc ggt cag gcg aat taa 1824
Arg Asn Ala Val Ala Thr Gly Leu Leu Leu Ala Gly Gln Ala Asn
595 600 605
<210> 60
<211> 607
<212> PRT
<213> *Klebsiella pneumoniae*

<400> 60
Met Pro Leu Ile Ala Gly Ile Asp Ile Gly Asn Ala Thr Thr Glu Val
1 5 10 15
Ala Leu Ala Ser Asp Tyr Pro Gln Ala Arg Ala Phe Val Ala Ser Gly
20 25 30
Ile Val Ala Thr Thr Gly Met Lys Gly Thr Arg Asp Asn Ile Ala Gly
35 40 45
Thr Leu Ala Ala Leu Glu Gln Ala Leu Ala Lys Thr Pro Trp Ser Met
50 55 60
Ser Asp Val Ser Arg Ile Tyr Leu Asn Glu Ala Ala Pro Val Ile Gly
65 70 75 80
Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser Thr
85 90 95
Met Ile Gly His Asn Pro Gln Thr Pro Gly Gly Val Gly Val Gly Val
100 105 110
Gly Thr Thr Ile Ala Leu Gly Arg Leu Ala Thr Leu Pro Ala Ala Gln
115 120 125
Tyr Ala Glu Gly Trp Ile Val Leu Ile Asp Asp Ala Val Asp Phe Leu
130 135 140
Asp Ala Val Trp Trp Leu Asn Glu Ala Leu Asp Arg Gly Ile Asn Val
145 150 155 160
Val Ala Ala Ile Leu Lys Lys Asp Asp Gly Val Leu Val Asn Asn Arg

	165	170	175
Leu Arg Lys Thr	Leu Pro Val Val Asp Glu Val Thr	Leu Leu Glu Gln	
180	185	190	
Val Pro Glu Gly Val Met Ala Ala Val Glu Val Ala Ala Pro Gly Gln			
195	200	205	
Val Val Arg Ile Leu Ser Asn Pro Tyr Gly Ile Ala Thr Phe Phe Gly			
210	215	220	
Leu Ser Pro Glu Glu Thr Gln Ala Ile Val Pro Ile Ala Arg Ala Leu			
225	230	235	240
Ile Gly Asn Arg Ser Ala Val Val Leu Lys Thr Pro Gln Gly Asp Val			
245	250	255	
Gln Ser Arg Val Ile Pro Ala Gly Asn Leu Tyr Ile Ser Gly Glu Lys			
260	265	270	
Arg Arg Gly Glu Ala Asp Val Ala Glu Gly Ala Glu Ala Ile Met Gln			
275	280	285	
Ala Met Ser Ala Cys Ala Pro Val Arg Asp Ile Arg Gly Glu Pro Gly			
290	295	300	
Thr His Ala Gly Gly Met Leu Glu Arg Val Arg Lys Val Met Ala Ser			
305	310	315	320
Leu Thr Gly His Glu Met Ser Ala Ile Tyr Ile Gln Asp Leu Leu Ala			
325	330	335	
Val Asp Thr Phe Ile Pro Arg Lys Val Gln Gly Gly Met Ala Gly Glu			
340	345	350	
Cys Ala Met Glu Asn Ala Val Gly Met Ala Ala Met Val Lys Ala Asp			
355	360	365	
Arg Leu Gln Met Gln Val Ile Ala Arg Glu Leu Ser Ala Arg Leu Gln			
370	375	380	
Thr Glu Val Val Val Gly Gly Val Glu Ala Asn Met Ala Ile Ala Gly			
385	390	395	400
Ala Leu Thr Thr Pro Gly Cys Ala Ala Pro Leu Ala Ile Leu Asp Leu			
405	410	415	

Gly Ala Gly Ser Thr Asp Ala Ala Ile Val Asn Ala Glu Gly Gln Ile

420 425 430

Thr Ala Val His Leu Ala Gly Ala Gly Asn Met Val Ser Leu Leu Ile

435 440 445

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

450 455 460

Lys Tyr Pro Leu Ala Lys Val Glu Ser Leu Phe Ser Ile Arg His Glu

465 470 475 480

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

485 490 495

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

500 505 510

Ala Ser Pro Leu Glu Lys Ile Arg Leu Val Arg Arg Gln Ala Lys Glu

515 520 525

Lys Val Phe Val Thr Asn Cys Leu Arg Ala Leu Arg Gln Val Ser Pro

530 535 540

Gly Gly Ser Ile Arg Asp Ile Ala Phe Val Val Leu Val Gly Gly Ser

545 550 555 560

Ser Leu Asp Phe Glu Ile Pro Gln Leu Ile Thr Glu Ala Leu Ser His

565 570 575

Tyr Gly Val Val Ala Gly Gln Gly Asn Ile Arg Gly Thr Glu Gly Pro

580 585 590

Arg Asn Ala Val Ala Thr Gly Leu Leu Leu Ala Gly Gln Ala Asn

595 600 605

<210> 61

<211> 4146

<212> DNA

<213> Escherichia coli

<400> 61

ggatcccttg cccgctgttg atccgttggt ccacctgata ttatgttaac ccagtagcca 60

gagtgtctca tgttgcagca cagccactcc gtgggaggca taaagcgaca gttcccgttc 120

ttctggctgc ggatagattc gactactcat caccgcttcc ccgtcgttaa taaatacttc	180
cacggatgat gtatcgataa atatccttag ggcgagcgtg tcacgctgcg ggaggggaat	240
actacggttag ccgtctaaat tctcgtgtgg gtaataccgc cacaaaacaa gtcgctcaga	300
ttggttatca atatacagcc gcattccagt gccgagctgt aatccgtaat gttcggcatc	360
actgtttctc agcgcacct gcaactgaat ctcaactgct tgcgcgtttt cctgcaaaac	420
atatttattg ctgattgtgc ggggagagac agattgatgc tgctggcgta acgactcagc	480
ttcgtgtacc gggcgttgta gaagtttgcc attgctctct gatagctcgc gcgccagcgt	540
catgcagcct gcccatcctt cagttttga gggcattggc gattcccaca tatccatcca	600
gccgataaca atacgccgac catccttcgc taaaaagctt tgtggtgcat aaaagtcag	660
cccgttatca agttcagtaa aatgcccgga ttgtgcaaaa agtcgtcctg gcgaccacat	720
tccgggtatt acgccacttt gaaagcgatt tcggtaactg tatccctcgg cattcatcc	780
ctgcgggaa aacatcagat aatgctgac gccaaaggctg aaaaagtcgc gacattccca	840
catatagctt tcacccgcat cagcgtgggc cagtacgca tcgaaggctc attcacgcaa	900
cgaactgccg cgataaagca ggatctgcc cgtgttgctt ggatctttcg ccccactac	960
catccacat gtgtcggctt cagccacac tttaggatcg cggaagtgca tgattccttc	1020
tggtggagtg aggatcacac cctgtttctc gaaatgaata ccatcccac tggtagccag	1080
acattgtact tcgcaattg catcgtcatt acctgcacca tcgagccaga cgtgtccggt	1140
gtagataagt gagaggacac cattgtcatc gacagcacta cctgaaaaac acccgtcttt	1200
gtcattatcg tctcctggcg ctacgcgaat aggcctcatgc tgccagtgga tcatatcgtc	1260
gctgggtggca tgtccccagt gcattggccc ccagtgttcg ctcatcggt gatgttgata	1320
aaacgcgtga taacgatct taaaccagat caggccgttt ggatcgttca tccaccggc	1380
aggaggcgcg aggtgaaaat ggggtagaaa agtggttacc cggtgctcat gaagttttgc	1440
tagggcgttt tgcgccgcat gcaatcgaga ttgcgtcatt ttaatcatcc tggtaaagca	1500
aatttggtga attgttaacg ttaactttta taaaaataaa gtccttact ttcataaatg	1560
cgatgaatat cacaaatgtt aacgttaact atgacgtttt gtgatcgaat atgcatgttt	1620
tagtaaatcc atgacgattt tgcgaaaaag aggtttatca ctatgcgtaa ctcatgaa	1680
tttaaggga aaaaatgtca gccaaagtat gggttttagg ggatgcggtc gtagatctct	1740
tgccagaatc agacggggcg ctactgcctt gtcctggcgg cgcgccagct aacgttgagg	1800
tgggaatcg cagattagcg ggaacaagtg ggtttatagg tcgggtgggg gatgatcctt	1860

ttggtgcgtt aatgcaaaga acgctgctaa ctgagggagt cgatatcacg tatctgaagc	1920
aagatgaatg gcaccggaca tccacgggtgc ttgtcgatct gaacgatcaa ggggaacgtt	1980
catttacgtt tatggtccgc cccagtgccg atcttttttt agagacgaca gacttgcctt	2040
gctggcgaca tggcgaatgg ttacatctct gtccaattgc gttgtctgcc gaggccttcgc	2100
gtaccagcgc atttactcgc atgacggcga tccggcatgc cggaggtttt gtcagcttcg	2160
atcctaatat tcgtgaagat ctatggcaag acgagcattt gctccgcttg tgtttcgggc	2220
aggcgctaca actggcggat gtcgtcaagc tctcggaaga agaattggcga cttatcagtg	2280
gaaaaacaca gaacgatcag gatatatgcg ccttggcaaa agagtatgag atcgccatgc	2340
tgttggtgac taaagggtga gaaggggtgg tggctctgta tcgaggacaa gttcaccatt	2400
ttgctggaat gtctgtgaat tgtgtcgata gcacgggggc gggagatgcg ttcgttgccg	2460
ggttactcac aggtctgtcc tctacgggat tatctacaga tgagagagaa atgcgacgaa	2520
ttatcgatct cgctcaacgt tgcggagcgc ttgcagtaac ggcgaaagg gcaatgacag	2580
cgctgccatg tcgacaagaa ctggaatagi gagaagtaaa cggcgaagtc gctcttatct	2640
ctaaatagga cgtgaatttt ttaacgacag gcaggtaatt atggcactga atattccatt	2700
cagaaatgcg tactatcgtt ttgcatccag ttactcattt ctctttttta tttcctggtc	2760
gctgtggtgg tcgttatacg ctatttggct gaaaggacat ctagggttga caggacgga	2820
attaggtaca ctttattcgg tcaaccagtt taccagcatt ctatttatga tgttctacgg	2880
catcgttcag gataaacctg gtctgaagaa accgctcatc tgggtatga gtttcatcct	2940
ggtcttgacc ggaccgttta tgatttacgt ttatgaaccg ttactgcaa gcaatttttc	3000
tgtaggtcta attctggggg cgctattttt tggcttgggg tatctggcgg gatgcggttt	3060
gcttgatagc ttcaccgaaa aaatggcgcg aaattttcat ttcgaatatg gaacagcgcg	3120
cgcttgggga tcttttggct atgctattgg cgcgttcttt gccggcatat tttttagtat	3180
cagtcccat atcaacttct ggttggcttc gctatttggc gctgtattta tgatgatcaa	3240
catgcgtttt aaagataagg atcaccagtg cgtacggcga gatgcgggag gggtaaaaaa	3300
agaggatttt atcgagttt tcaaggatcg aaacttctgg gtttctgtca tatttatgt	3360
ggggacgtgg tctttctata acatttttga tcaacaactt tttcctgtct tttattcagg	3420
tttattcgaa tcacacgatg taggaacgcg cctgtatggt tatctcaact cattccaggt	3480
ggtactcgaa gcgctgtgca tggcgattat tcctttcttt gtgaatcggg tagggccaaa	3540
aaatgcatta cttatcggag ttgtgattat ggcgttgcgt atcctttcct gcgcgtgtt	3600
cgtaacccc tggattattt cattagtga gttgttacat gccattgagg ttccactttg	3660
tgtcatatcc gtcttcaaat acagcgtggc aaactttgat aagcgcctgt cgtcgacgat	3720

ctttctgatt ggttttcaaa ttgccagttc gcttgggatt gtgctgcttt caacgccgac	3780
tgggatactc ttgaccacg caggctacca gacagttttc ttcgcaattt cgggtattgt	3840
ctgcctgatg ttgctatttg gcatttttctt cttgagtaaa aaacgcgagc aaatagttaa	3900
ggaaacgcct gtaccttcag caatatagac gtaaactttt tccggttgtt gtcgatagct	3960
ctatatccct caaccggaag ataataatag taaaatgctt agccctgcta ataatgcct	4020
aatccaaacg cctcattcat gttctggtag agtcgctcaa atgtacttca gatgcgcggt	4080
tcgctgattt ccaggacatt gtcgtcattc agtgacctgt cccgtgtatc acggctcctgc	4140
gaattc	4146
<210> 62	
<211> 13669	
<212> DNA	
<213> Artificial Sequence	
<220><223> Plasmid	
<400> 62	
tagtaagcc ctgcctagat tttaatgcgg atgttgcat tacttcgcca actatigcga	60
taacaagaaa aagccagcct ttcatgatat atctccaat ttgttaggg cttattatgc	120
acgcttaaaa ataataaag cagacttgac ctgatatgtt ggctgtgagc aattatgtgc	180
ttagtgcac taacgcttga gtttaagccgc gccgcgaagc ggctgcggct tgaacgaatt	240
gttagacatt atttgccgac taccttggtg atctcgctt tcacgtagt gacaaattct	300
tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata agcctgtcta	360
gcttcaagta tgacgggctg atactgggcc ggcaggcgct ccattgccca gtcggcagcg	420
acatccttcg gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc	480
actacatttc gtcacatgcc agcccagtcg ggccggcgagt tccatagcgt taaggtttca	540
tttagcgct caaatagatc ctgttcagga accggatcaa agagtcttc cgcgcgtgga	600
cctaccaagg caacgctatg ttctcttgct ttgttcagca agatagccag atcaatgtcg	660
atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc	720
agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact	780
tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttccaa aaggtcgttg	840
atcaaagctc gccgcgttgt ttcatcaagc cttacggtca ccgtaaccag caaatcaata	900
tcactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac	960
gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt cgatacttcg	1020

gcatcaccg ctccctcat gatgtttaac ttgttttag ggcactgcc ctgctgcgta	1080
acatcgttgc tgcctcataa catcaaacat cgaccacg cgtaacgcgc ttgctgcttg	1140
gatgcccag gcatagactg taccctaaaa aaacagtcac aacaagccat gaaaaccgcc	1200
actgcgccgt taccaccgtt gcgttcggtc aaggttctgg accagttagc tgagcgcata	1260
cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc	1320
atccgtttcc acggtgtgcg tcaccgggca acctgggca gcagcgaagt cgaggcattt	1380
ctgtcctggc tggcgaacga gcgcaagggt tcggtctcca cgcacgtca ggcattggcg	1440
gccttctgtt tctctacgg caaggtgctg tgcacggatc tgccttggt tcaggagatc	1500
ggaagacctc ggccgtcgcg gcgcttgccg gtggtgctga ccccgatga agtggttcgc	1560
atcctcggtt ttctggaagg cgagcatcgt ttgttcgcc agcttctgta tggaaagggc	1620
atgcggatca gtgagggttt gcaactgcgg gtcaaggatc tggatttcga tcacggcacg	1680
atcatcgtgc gggagggcaa gggctccaag gatcgggcct tgatgttacc cgagagcttg	1740
gcaccagcc tgcgcgagca ggggaattaa tccccacggg tttgctgcc cgcaaagg	1800
ctgttctggt gttgctagtt tgttatcaga atcgagatc cggttcagc cggtttgccg	1860
gctgaaagcg ctatttcttc cagaattgcc atgattttt cccacggga ggcgtcactg	1920
gctcccggtt tgcggcagc ttgattcga taagcagcat cgcctgttc aggtgtcta	1980
tgtgtactg ttgagctga acaagtgtc tcaggtgttc aattcatgt tctagtgtt	2040
ttgtttact ggtttacct gttctattag gtgttatcgt ctgttcact gttacattgt	2100
cgatctgttc atggtgaaca gctttgaatg caccaaaaac tcgtaaaagc tctgatgtat	2160
ctatctttt tacaccgtt tcactgtgc atatggacag tttccctt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttactgata gatacaagag	2280
ccataagaac ctcagatcct tccgtattta gccagtatgt tctctagtgt ggttcgttgt	2340
tttgcgtga gccatgagaa cgaaccattg agatcact tactttgcat gtcactcaa	2400
aattttgcct caaaactggt gagctgaatt ttgcagtta aagcatcgt tagtggtttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatggtt ttggtatttt gtcaccattc	2520
atttttatct ggttgttctc aagttcgggt acgagatcca ttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccatt ggttaagcct tttaaactca	2700
tggtagttat tttcaagcat taacatgaac ttaattcat caaggctaat ctctatattt	2760
gccttgtgag tttcttttg tgttagttct ttaataacc actcataaat cctcatagag	2820
tatttgtttt caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880

aaaagataag gcaatatctc ttcactaaaa actaatctta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag cctttaacca aaggattcct gatttccaca	3000
gttctcgtca tcagctctct ggttgcttta gctaatacac cataagcatt ttcctactg	3060
atgttcatca tcigagcgta ttggtataa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tcgtggggtt gagtagtgcc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attggtctag gtgattttaa tcactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtctttt cctttgagtt gtgggtatct gtaaattctg ctagaccttt	3360
gctggaaaac ttgtaaattc tgctagacct tctgtaaatt ccgctagacc ttgtgtgtt	3420
ttttttgttt atattcaagt ggttataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcactacttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaacgtt gtttctcct ctacaaaaca gaccttaaaa ccctaaggc	3600
ttaagtagca cctcgcgaag ctccggcaaa tcgctgaata ttccttttgt ctccgaccat	3660
caggcacctg agtcgctgtc tttttcgtga cattcagttc gctgcgctca cggctctggc	3720
agtgaatggg ggtaaatggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccg tcacgggctt ctacggcggt tttatggcgg gtctgctatg	3840
tggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggcttaccg gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020
cgaccggagg cttttgactg ctaggcgatc tgtgctgttt gccacgggat gcagcaccag	4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcactggaac gagaagtcag	4140
gcgagccgtc acgcccttga caatgccaca tctgagcaa ataattcaac cactaaacaa	4200
atcaaccgag tttcccggag gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa	4260
caagttcaga caatcacctt ggccgcccgc cagcaaatgg cggcggcggt ggaaaaaaaa	4320
gccactgaga tcaacgtggc ggtggtgttt tccgtagtgt accgcggagg caacacgctg	4380
cttatccagc ggatggacga ggccttcgtc tccagctgcg atatttcctt gaataaagcc	4440
tggagcgctt gcagcctgaa gcaaggtacc catgaaatta cgtcagcgtt ccagccagga	4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatatt ttggcgcggc	4560
ctgccagtta tttttaatga gcaggtaatt ggccgctcgc gcgttagcgg cggtagcggtc	4620
gagcaggatc aattatttag ccagtgcgcc ctggattgtt tttccgcatt ataacctgaa	4680

gcgagaaggt atattatgag ctatcgatg ttccgccagg cattctgagt gttaacgagg 4740

ggaccgtcat gtcgctttca ccgccaggcg tacgcctgtt ttacgatccg cgcgggcacc 4800

atgccggcgc catcaatgag ctgtgctggg ggtcggagga gcagggggtc ccctgccaga 4860

ccataaccta tgacggaggc ggtgacgccg ctgcgctggg cgccctggcg gccagaagct 4920

cgcctctgcg ggtgggtatc gggctcagcg cgtccggcga gatagccctc actcatgccc 4980

agctgccggc ggacgcgccg ctggctaccg gacacgtcac cgatagcgac gatcaactgc 5040

gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact 5100

gaatgtatcg tatctatacc cgcaccgggg ataaaggcac caccgccctg tacggcggca 5160

gccgcatcga gaaagacat attcgcgtcg aggccctacgg caccgtcgat gaactgatat 5220

cccagctggg cgtctgttac gccacgacc gcgacgccgg gctgcgggaa agcctgcacc 5280

atattcagca gacgtgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga 5340

ccgccttag ccagacgac ggcaagagg agatcacgc cctggagcgg cttatcgacc 5400

gcaatatggc cgagagcggc ccgttaaac agttcgtgat cccggggagg aatctcgctt 5460

ctgccagct gcacgtggcg cgcaccagc cccgtcggct cgaacgcctg ctgacggcca 5520

tggaccgcgc gcatccgctg cgcgacgcgc tcaaagccta cagcaatcgc ctgtcggatg 5580

ccctgttctc catggcgca atcgaagaga ctaggcctga tgcttgcgt tgaactggcc 5640

tagcaaacac agaaaaagc ccgcacctga cagtgcgggc ttttttttc ctaggcgatc 5700

tgtgtgttt gccacggtat gcagaccag cgcgagatta tgggctcgca cgctcgactg 5760

tcggacgggg gcactggaac gagaagtcag gcgagccgtc acgcccttga caatgccaca 5820

tcctgagcaa ataattcaac cactaaacaa atcaaccgcg tttcccgag gtaaccaagc 5880

ttcaccttt gagccgatga acaatgaaaa gatcaaacg atttcagta ctggcccagc 5940

gccccgtcaa tcaggacggg ctgattggcg agtggcctga agaggggctg atcgccatgg 6000

acagccctt tgaccggtc tcttcagtaa aagtggacaa cggctctgac gtcgaaactgg 6060

acggcaaagc ccgggaccag tttagcatga tcgaccgatt tatcgccgat tacgcatca 6120

acgttgagcg cacagagcag gcaatgcgc tggaggcggg ggaatatgcc cgtatgctgg 6180

tggatattca cgtcagccgg gaggagatca ttgccatcac taccgccatc acgccggcca 6240

aagcgtcga ggtgatggcg cagatgaacg tggtaggat gatgatggcg ctgcagaaga 6300

tgctgcccc ccggacccc tccaaccagt gccacgtcac caatctcaa gataatccgg 6360

tgcagattgc cgtgacgcc gccaggccg ggtccgcgg cttctcagaa caggagacca 6420

cggtcggat cgcgcgtac gcgccgttta acgccctggc gctgttggc ggttcgcagt 6480
 gcggcgccc cggcgtgtg acgcagtgt cggtggaaga ggccaccgag ctggagctgg 6540
 gcatgcgtgg cttaccagc tacgccgaga cgggtgcgt ctaccgcacc gaagcggat 6600
 ttaccgacgg cgatgatac ccgtgggtcaa aggcgttcct cgccctggcc tacgcctccc 6660
 gcgggttgaa aatgcgtac acctccgca ccgcatccga agcgtgatg ggctattcgg 6720
 agagcaagtc gatgctctac ctgcaatgc gctgcatctt cattactaaa ggcccgggg 6780
 ttcagggact gcaaacggc gcggtgagct gtatcgcat gaccggcgct gtgccgtcgg 6840

 gcattcgggc ggtgctggcg gaaaacctga tcgcctctat gctcgacct gaagtggcgt 6900
 ccgccaacga ccagacttc tccactcgg atattcgccg caccgcgcg accctgatgc 6960
 agatgctgcc gggcaccgac tttattttct ccggctacag cgcggtgccg aactacgaca 7020
 acatgttcgc cgctcgaac ttcatgagg aagattttga tgattacaac atcctgcagc 7080
 gtgacctgat ggttgacggc ggcctgcgt cggtgaccga ggccgaaacc attgccattc 7140
 gccagaaagc ggccggggcg atccaggcgg tttccgcga gctggggctg ccgccaatcg 7200
 ccgacgagga ggtggaggcc gccacctacg cgcacggcag caacgagatg ccgccgcgta 7260

 acgtggtgga ggatctgagt gcggtggaag agatgatga gcgcaacatc accggcctcg 7320
 atattgtcgg cgcgtgagc cgcagcggct ttgaggatat cgccagcaat attctcaata 7380
 tgctgcgcca gggggtcacc ggcgattacc tgcagacctc ggccattctc gatcggcagt 7440
 tcgaggtggt gagtgcggc aacgacatca atgactatca ggggccgggc accggctatc 7500
 gcatctctgc cgaacgtgg gcggagatca aaaatatcc gggcgtggtt cagcccgaca 7560
 ccattgaata aggcggtatt cctgtgcaac agacaacca aattcagccc tcttttacc 7620
 tgaaaaccg cgagggcggg gtagcttctg ccgatgaacg cgccgatgaa gtggtgatcg 7680

 gcgtcggccc tgccttcgat aaacaccagc atcacactct gatcgatatg ccccatggcg 7740
 cgatcctcaa agagctgatt gccggggtgg aagaagagg gcttcacgcc cgggtggtgc 7800
 gcattctgcg cagctccgac gtctccttta tggcctggga tgcggccaac ctgagcggct 7860
 cggggatcgg catcggtatc cagtcgaagg ggaccacggt catccatcag cgcgatctgc 7920
 tgccgtcag caacctggag ctgttctccc aggcgccgct gctgacgtg gagacctacc 7980
 ggagattgg caaaaacgt gcgcgtatg cgcgcaaaga gtcaccttcg ccggtgccgg 8040
 tggatgaaga tcagatggtg cgccgaaat ttatggcaa agccgcgcta tttcatatca 8100

 aagagaccaa acatgtggtg caggacgccg agcccgctac cctgcacatc gacttagtaa 8160
 gggagtgacc atgagcgaga aaacctgcg cgtgcaggat tatccgttag ccaccgctg 8220
 cccggagcat atcctgacg ctaccggcaa accattgacc gatattacc tcgagaaggt 8280

gctctctggc gaggtgggcc cgcaggatgt gcggatctcc cgccagaccc ttgagtiacca	8340
ggcgcagatt gccgagcaga tgcagcgcca tgcggtggcg cgcaatttcc gccgcgcggc	8400
ggagcttata gccattcctg acgagcgcat tctggctatc tataacgcgc tgcgcccgtt	8460
ccgctcctcg caggcggagc tgctggcgat cgccgacgag ctggagcaca cctggcatgc	8520
gacagtgaat gccgcctttg tccgggagtc ggcggaagtg tatcagcagc ggcataagct	8580
gcgtaaagga agctaagcgg aggtcagcat gccgttaata gccgggattg atatcggcaa	8640
cgccaccacc gaggtggcgc tggcgtccga ctaccgcag gcgagggcgt ttgttgccag	8700
cgggatcgtc gcgacgacgg gcatgaaagg gacgcgggac aatatcgccg ggaccctcgc	8760
cgcgctggag caggccctgg cgaaaacacc gtggtcgatg agcgatgtct ctcgcatcta	8820
tcttaacgaa gccgcgccgg tgattggcga tgtggcgatg gagaccatca ccgagacat	8880
tatcaccgaa tcgacatga tcggtcataa ccgcagacg ccgggcgggg tgggcgttgg	8940
cgtggggacg aciatcgccc tcgggcggct ggcgacgctg ccggcggcgc agtatgccga	9000
ggggtggatc gtactgattg acgacgccgt cgatttcctt gacgccgtgt ggtggctcaa	9060
tgaggcgctc gaccggggga tcaactgggt ggcggcgac ctaaaaaagg acgacggcgt	9120
gctggtgaac aaccgcctgc gtaaaacct gccggtggtg gatgaagtga cgctgctgga	9180
gcaggtcccc gagggggtaa tggcggcggg ggaagtggcc gcgccgggcc aggtggtgcg	9240
gatacctgtc aatccctacg ggatcgccac ctctctcggg ctaagcccgg aagagaccca	9300
ggccatcgtc cccatcgccc gcgcctgat tggcaaccgt tccgcggtgg tgctcaagac	9360
cccgcagggg gatgtgcagt cgcggtgat ccggcgggc aacctctaca ttagcggcga	9420
aaagcgccgc ggagaggccg atgtcgccga ggcgcgga gccatcatgc aggcatgag	9480
cgctgcgct ccggtacgcg acatccgcgg cgaaccgggc acccagccg gcggcatgct	9540
tgagcgggtg cgcaaggtaa tggcgtccct gaccggccat gagatgagcg cgatatacat	9600
ccaggatctg ctggcgttgg atacgtttat tccgcgcaag gtgcaggcg ggatggccgg	9660
cgagtgcgcc atggagaatg ccgtcgggat ggcgcgcatg gtgaaagcgg atcgtctgca	9720
aatgcaggtt atcgcccgcg aactgagcgc ccgactgcag accgaggtgg tgggtggcgg	9780
cgtggaggcc aacatggcca tcgccggggc gttaccact cccggctgtg cggcgccgt	9840
ggcgatcctc gacctcgcg ccggctcgac ggatgcggcg atcgtcaac cgagggggca	9900
gataacggcg gtccatctcg ccggggcggg gaatatggtc agcctgttga ttaaaaccga	9960
gctgggcctc gaggatcttt cgctggcgga agcgataaaa aaataccgc tggccaaagt	10020
ggaaagcctg ttcagtattc gtcacgagaa tggcgcggtg gatttcttc gggaagccct	10080
cagcccgcg gtgttcgcca aagtggtga catcaaggag ggcaactgg tgccgatcga	10140

taacgccagc ccgtcgaaa aaattcgtct cgtgcgccgg caggcgaaag agaaagtgtt	10200
tgtcaccaac tgcctgcgcg cgctgcgcca ggtctcacc ggcggttcca ttcgcgatat	10260
cgcttttgtg gtgctgggtg gcggtcatc gctggacttt gagatccgc agcttatcac	10320
ggaagccttg tcgcactatg gcgtggtcgc cgggcagggc aatattcggg gaacagaagg	10380
gccgcgaat gcggtcgcca ccgggtgct actggccggt caggcgaatt aaacgggcgc	10440
tcgcgccagc ctctaggtac aaataaaaaa ggcacgtcag atgacgtgcc tttttcttg	10500
tctagagtac tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg	10560
ttttcccagt cacgacgttg taaaacgacg gccagtgaat tcgagctcgg taccgggggc	10620
ggccgcgcta gcgccgata cagctggagt ttgtagaaac gcaaaaaggc catccgtcag	10680
gatggccttc tgcttaattt gatgcctggc agtttatggc gggcgtcctg cccgccacc	10740
tccgggccgt tgcctcgcaa cgttcaaatc cgtcccggc ggatttgtcc tactcaggag	10800
agcgttcacc gacaacaac agataaaacg aaaggcccag tcttcgact gaccccttcg	10860
ttttatttga tgcctggcag ttccctactc tcgcatgggg agacccaca ctaccatcg	10920
cgctacggcg ttctacttct gaggctcgca tggggtcagg tgggaccacc gcgctactgc	10980
cgccaggcaa attctgtttt atcagaccgc ttctgcgttc tgatttaac tgtatcaggc	11040
tgaaaatctt ctctcatccg ccaaaacagc caagcttgca tgcctgcagc ccgggttacc	11100
atttcaacag atcgtcctta gcatataagt agtcgtcaaa aatgaattca acttcgtctg	11160
tttcggcatt gtagccgcca actctgatgg attcgtgggt ttgacaatg atgtcacagc	11220
ctttttcctt taggaagtc aagtcgaaag tagtggcaat accaatgac ttacaaccgg	11280
cggtttttcc ggcggaata cctgctggag cgtcttcaaa tactactacc ttagatttgg	11340
aagggtcttg ctcatgac ggatactcta agccattcct gcccttcaga tatggttctg	11400
gatgaggctt accctgtttg acatcattag cggtaatgaa gtactttggt ctctgattc	11460
ccagatgctc gaaccatttt tgtgccatat cacgggtacc ggaagttgcc acagccatt	11520
tctcttttgg tagagcgttc aaagcgttgc acagcttaac tgcacctggg acttcaatgg	11580
atttttcacc gtacttgacc ggaatttcag ctctaatctt gttacatac tcttcattgg	11640
caaagtctgg agcgaactta gcaatggcat caaacgttct ccaaccatgc gagacttga	11700
taacgtgttc agcatcgaaa taaggtttgt ccttacgaa atccctccag aatgcagcaa	11760
tggctggttg agagatgata atggtaccgt cgacgtcgaa caaagcgcg ttaactttca	11820
aagatagagg tttagtagtc aatcccataa ttctagtctg tttcctggat ccaataaatc	11880

taatcttcat gtagatctaa ttcttcaatc atgtccggca ggttcttcat tgggtagttg	11940
ttgtaaacga tttggtatag ggcttcaaataaat gggaagt cttcgacaga gccacatgtt	12000
tccaaccatt cgtgaacttc tttgcaggta attaaacctt gagcggattg gccattcaac	12060
aactcctttt cacattccca ggcttccctt ccagaagtag ccattagcct agcaaccttg	12120
acgtttctac caccagcgca ggtggtgatc aaatcagcaa caccagcaga ctcttggtag	12180
tatgtttctt ctctagattc tgggaaaaac atttgaccga atctgatgat ctcacccaaa	12240
ccgactcttt ggatggcagc agaagcgttg ttacccagc ctagaccttc gacgaaacca	12300
caacctaagg caacaacgtt cttcaaagca ccacagatgg agataccagc aacatcttcg	12360
atgacactaa cgtggaagta aggtctgtgg aacaaggcct ttagaacctt atggtcgacg	12420
tccttgccct cgctctgaa atcctttgga atgtggtgtaag caactgttgt ttcagaccag	12480
tgttcttgag cgacttcggt ggcaatgtta gcaccagata gagcaccaca ttgaatacct	12540
agttcctcag tgatgtaaga ggatagcaat tggacacctt tagcaccaac ttcaaaaccc	12600
tttagacagg agatagctct gacgtgtgaa tcaacatgac ctttcaattg gctacagata	12660
cggggcaaaa attgatgtgg aatgttgaaa acgatgatgt cgacatcctt gactgaatca	12720
atcaagtctg gattagcaac caaattgtcg ggtagagtga tgccaggcaa gtatttcacg	12780
ttttgatgtc tagtatttat gatttcagtc aatttttcac cattgatctc ttcttcgaac	12840
accacattt gtactattgg agcgaaaact tctgggtatc cttacaatt ttcggcaacc	12900
accttgcaa tagtagtacc ccagttacca gatccaatca cagtaacctt gaaaggcttt	12960
tcggcagcct tcaagaaac agaagaggaa cttctcttcc taccagcatt caagtggccg	13020
gaagttaagt ttaatctatc agcagcagca gccatggaat tgtcctcctt actagtcag	13080
gtctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacattata cgagccggat	13140
gattaattgt caacagctca ttccagaata tttgccagaa ccgttatgat gtcggcgcaa	13200
aaaacattat ccagaacggg agtgcgcctt gagcgacacg aattatgcag tgatttacga	13260
cctgcacagc cataccacag cttccgatgg ctgcctgacg ccagaagcat tgggtgcacgc	13320
tagccagtag atttaaatgg taccctctag tcaaggcctt aagtgagtcg tattacggac	13380
tggccgtcgt tttaaacgt cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc	13440
ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggccgc accgatcgcc	13500
cttccaaca gttgcgcagc ctgaatggcg aatggcgctt gatgcggat tttctcctta	13560
cgcatctgtg cgttatattca caccgatat ggtgcactct cagtacaatc tgctctgatg	13620
ccgatagtt aagccagccc cgacacccgc caacacccgc tgacgagct	13669

<211> 13543

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid

<400> 63

tagtaaagcc ctcgctagat tttaatgcgg atgttgcat tacttcgcca actattgcga	60
taacaagaaa aagccagcct ttcatgatat atctcccaat ttgtgtaggg cttattatgc	120
acgcttaaaa ataataaaag cagacttgac ctgatatgtt ggctgtgagc aattatgtgc	180
ttagtgcatc taacgcttga gttaagccgc gccgcgaagc ggctgcggct tgaacgaatt	240
gttagacatt atttgccgac taccttggtg atctcgctt tcacgtagt gacaaattct	300
tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata agcctgtcta	360
gcttcaagta tgacgggctg atactgggcc ggcaggcgct ccattgccca gtcggcagcg	420
acatccttcg gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc	480
actacatttc gtcctatgcc agcccagtcg ggccggcagc tccatagcgt taaggtttca	540
tttagcgctt caaatagatc ctgttcagga accggatcaa agagttcctc cgccgctgga	600
cctaccaagg caacgctatg ttctcttgct ttgtcagca agatagccag atcaatgtcg	660
atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc	720
agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatgggtgact	780
tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttccaa aaggctcgtt	840
atcaaagctc gccgcgttgt ttcatcaagc cttacggtca ccgtaaccag caaatcaata	900
tcactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac	960
gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt cgatacttcg	1020
gcgatcaccg ctccctcat gatgtttaac ttgttttag ggcgactgcc ctgctgcgta	1080
acatcgttgc tgctccataa catcaaacat cgaccacgga cgtaacgcgc ttgctgcttg	1140
gatgcccag gcatagactg taccccaaaa aaacagtcac aacaagccat gaaaaccgcc	1200
actgcgccgt taccaccgct gcgttcggtc aaggttctgg accagttgcg tgagcgcata	1260
cgctacttgc attacagctt acgaaccgaa caggttatg tccactgggt tcgtgccttc	1320
atccgtttcc acggtgtgct tcaccgggca accttgggca gcagcgaagt cgaggcattt	1380
ctgtcctggc tggcgaacga gcgcaagggt tcggtctcca cgcatcgtca ggcattggcg	1440
gccttgctgt tcttctacgg caaggtgctg tgcacggatc tgccttggt tcaggagatc	1500
ggaagacctc ggccgtcgcg gcgcttgccg gtggtgctga ccccgatga agtggttcgc	1560

atcctcgggtt ttctggaagg cgagcatcgt ttgttcgccc agcttctgta tggaaacgggc	1620
atgcggatca gtgaggggtt gcaactgcgg gtcaaggatc tggatttcga tcacggcacg	1680
atcatcgtgc gggagggcaa gggctccaag gatcgggcct tgatgttacc cgagagcttg	1740
gcacccagcc tgcgcgagca ggggaattaa ttcccacggg ttttgctgcc cgcaaacggg	1800
ctgttctggt gttgctagtt tgttatcaga atcgcagatc cggcttcagc cggtttgcg	1860
gctgaaagcg ctatttcttc cagaattgcc atgatttttt cccacaggga ggcgtcactg	1920
gctcccggtg tgcggcagc tttgattcga taagcagcat cgcctgttcc aggcgtgcta	1980
tgtgtgactg ttgagctgta acaagtgtc tcagggtgtc aatttcattg tctagttgct	2040
ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcattc gttacattgt	2100
cgatctgttc atggatgaac gctttgaatg caccaaaaac tcgtaaaagc tctgatgtat	2160
ctatcttttt tacaccgttt tcatctgtgc atatggacag ttttcccttt gatatgtaac	2220
ggatgaacagt tgttctactt ttgtttgtta gtcttgatgc ttactgata gatacaagag	2280
ccataagaac ctcagatcct tccgtattta gccagtatgt tctctagtgt gggttcgtgt	2340
ttttcggtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt ttgacagta aagcatcgtg tagtggtttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatgggtg ttgggtatttt gtcaccattc	2520
atttttatct gggtgttctc aagttcgggt acgagatcca tttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattgggttc aaaaccatt gggttaagcct tttaaaactca	2700
tggtagttaa tttaagcat taacatgaac ttaaattcat caaggctaata ctctatattt	2760
gccttgtagg ttttcttttg tgttagttct ttaataaacc actcataaat cctcatagag	2820
tatttggttt caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880
aaaagataag gcaatatctc ttactaaaa actaattcta attttctgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag ctttaacca aaggattcct gatttcaca	3000
gttctcgtca tcagctctct gggtgcttta gctaatacac cataagcatt ttcctactg	3060
atgttcacga tctgagcgta ttgggtataa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tcgtgggggt gagtagtgcc acacagcata aaattagctt gggttcacgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attggcttag gtgattttta tcactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtcccttt cctttgagtt gtgggtatct gtaaattctg ctagaccttt	3360
gctggaaaac ttglaaattc tgctagacct tctgtaaatt ccgctagacc tttgtgtgtt	3420

ttttttgttt atattcaagt ggttataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcactacttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaacgct gtttgcctct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca ccctcgcaag ctcgggcaaa tcgctgaata ttccttttgt ctccgaccat	3660
caggcacctg agtcgctgtc tttttcgtga cattcagttc gctgcgctca cggtcttgcc	3720
agtgaatggg ggtaaatggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccc tcacgggctt ctacgggctt tttatggcgg gtctgctatg	3840
tgggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggccttacc gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020
cgaccggagg cttttgactg ctaggcgatc tgtgtgttt gccacggtat gcagcaccag	4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcactggaac gagaagtcag	4140
gcgagccgtc acgcccttga ctatgccaca tcttgagcaa ataattcaac cactaaacaa	4200
atcaaccgag tttcccgagg gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa	4260
caagttcaga caatcacctt ggcccgccgc cagcaaatgg cggcgccggt ggaaaaaaaa	4320
gccactgaga tcaacgtggc ggtggtgttt tccgtagtgt acccgaggag caacacgtgt	4380
cttatccagc ggatggacga ggccttcgtc tccagctgag atatttcctt gaataaagcc	4440
tggagcgctt gcagcctgaa gcaaggtacc catgaaatta cgtcagcggc ccagccagga	4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatttt tggcggcggc	4560
ctgccagtta tttttaatga gcaggtaatt ggccgctgct gcgttagcgg cggtagcgtc	4620
gagcaggatc aattattagc ccagtgcgcc ctgattgttt tttccgattt ataacctgaa	4680
gcgagaaggt atattatgag ctatcgatg ttcgccagg cattctgagt gttaacgagg	4740
ggaccgtcat gtcgctttca ccgccaggcg tacgcctgtt ttacgatccg cgcgggcacc	4800
atgccggcgc catcaatgag ctgtgctggg ggctggagga gcagggggtc ccctgccaga	4860
ccataacctt tgacggaggc ggtgacgccg ctgcgctggg cgccctggcg gccagaagct	4920
cgccctgctg ggtgggtatc gggctcagcg cgtccggcga gatagccctc actcatgcc	4980
agctgccggc ggacgcgcgc ctggctaccg gacacgtcac cgatagcgac gatcaactgc	5040
gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact	5100
gaatgtatcg tatctatacc cgcaccgggg ataaaggcac caccgccctg tacggcggca	5160
gccgcatcga gaaagaccat attcgctgag aggcctacgg caccgtcgat gaactgatat	5220

cccagctggg cgctctgtac gccacgaccc gcgacgccgg gctgcgggaa agcctgcacc 5280

atattcagca gacgctgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga 5340

cccgcttag ccagacgac ggcaagagg agatcacgc cctggagcgg ctatcgacc 5400

gcaatatggc cgagagcggc ccgttaaac agttcgtgat cccggggagg aatctgcct 5460

ctgccagct gcacgtggcg cgcacccagt cccgtcggct cgaacgcctg ctgacggcca 5520

tggaccgcgc gcatccgctg cgcgacgcgc tcaaacgcta cagcaatcgc ctgtcggatg 5580

ccctgttctc catggcgca atcgaagaga ctaggcctga tgcttgcgct tgaactggcc 5640

tagcaaacac agaaaaaagc ccgcacctga cagtgcgggc ttttttttc ctaggcgatc 5700

tgctgtgttt gccacggtat gcagaccag cgcgagatta tgggctcgca cgctcgactg 5760

tcggacgggg gcaactgaac gagaagtcag gcgagccgtc acgcccttga ctatgccaca 5820

tcctgagcaa ataattcaac cactaaacaa atcaaccgcg tttcccgag gtaaccaagc 5880

ttcacctttt gagccgatga acaatgaaa gatcaaacg atttcagta ctggcccagc 5940

gccccgtcaa tcaggacggg ctgattggcg agtggcctga agaggggctg atcgccatgg 6000

acagcccctt tgacccggtc tcttcagtaa aagtggacaa cggctctgac gtcgaactgg 6060

acggcaaacg ccgggaccag tttagcatga tcgaccgatt tatcgccgat tacgcgatca 6120

acgttgagcg cacagagcag gcaatgcgcc tggaggcggg ggaaatagcc cgtatgctgg 6180

tggatattca cgtcagccgg gaggagatca ttgccatcac taccgccatc acgccggcca 6240

aagcggctga ggtgatggcg cagatgaacg tggtaggat gatgatggcg ctgcagaaga 6300

tgctgcccc ccggaccccc tccaaccagt gccacgtcac caatctcaa gataatccgg 6360

tgcatattgc cgctgacgcc gccgaggccg ggtaccgagg ctctcagaa caggagacca 6420

cggtcggtat cgcgcgtac gcgccgttta acgccctggc gctgttggtc ggttcgcagt 6480

gcggccgccc cggcgtgttg acgcagtgt cggtggaaga ggccaccgag ctggagctgg 6540

gcatgcgtgg cttaaccagc tacgccgaga cgggtgtcgg ctacggcacc gaagcggat 6600

ttaccgacgg cgatgatacg ccgtgggtcaa agcggttct cgcctcgcc tacgcctccc 6660

gcgggttgaa aatgcgtac acctccggca ccggatccga agcgtgatg ggctattcgg 6720

agagcaagtc gatgctctac ctgcaatgc gctgcatctt cattactaaa ggcgccgggg 6780

ttcagggact gcaaacggc gcggtagct gtatcgcat gaccggcgct gtgccgtcgg 6840

gcattcgggc ggtgctggcg gaaaacctga tcgcctctat gctcgacctc gaagtggcgt 6900

ccgccaacga ccagactttc tccactcgg atattcgcc caccgcgcgc acctgatgc 6960

agatgtctgcc gggcaccgac tttatcttct ccggttacag cgcggtgccg aactacgaca	7020
acatgttctgc cggctcgaac ttcgatgcgg aagattttga tgattacaac atcctgcagc	7080
gtgacctgat ggttgacggc ggcttgcgtc cggtgaccga ggcggaacc attgccattc	7140
gccagaaaagc ggcgcgggcg atccaggcgg ttttccgca gctggggctg ccgccaatcg	7200
ccgacgagga ggtggaggcc gccacctacg cgcacggcag caacgagatg ccgccgcgta	7260
acgtggtgga ggaatctgagt gcggtggaag agatgatgaa gcgcaacatc accggcctcg	7320
atattgtcgg cgcgctgagc cgcagcggct ttgaggatat cgccagcaat attctcaata	7380
tgctgcgcca gcgggtcacc ggcgattacc tgcagacctc ggccattctc gatcggcagt	7440
tcgaggtggt gagtgcggtc aacgacatca atgactatca ggggccgggc accggtatc	7500
gcatctctgc cgaacgctgg gcggagatca aaaatatcc gggcgtggtt cagcccgaca	7560
ccattgaata agcggttatt cctgtgcaac agacaacca aattcagccc tcttttacc	7620
tgaaaacccg cgagggcggg gtagcttctg ccgatgaacg cgccgatgaa gtggtgatcg	7680
gcgtcggccc tgccttcgat aaacaccagc atcacactct gatcgatatg ccccatggcg	7740
cgatcctcaa agagctgatt gccggggtgg aagaagaggg gcttcacgcc cgggtggtgc	7800
gcattctgcg cagtcggac gtctccttta tggcctggga tgcggccaac ctgagcggct	7860
cggggatcgg catcggtatc cagtccaagg ggaccacggc catccatcag cgcgatctgc	7920
tgccgtcag caacctggag ctgttctccc aggcgcgct gctgacgctg gagacctacc	7980
ggcagattgg caaaaacgt gcgcgtatg cgcgcaaaga gtcaccttcg ccggtgccgg	8040
tggtgaacga tcagatggtg cggccgaaat ttatggccaa agccgcgcta tttcatatca	8100
aagagaccaa acatgtggtg caggacggcg agcccgctac cctgcacatc gacttagtaa	8160
gggagtgacc atgagcgaga aaaccatgcg cgtgcaggat tatccgttag ccaccgctg	8220
cccggagcat atcctgacgc ctaccggcaa accattgacc gatattacc tcgagaaggt	8280
gtctctctggc gaggtgggcc cgcaggatgt gcggatctcc cgccagacc ttgagtacca	8340
ggcgcagatt gccgagcaga tgcagcgcca tgcggtggcg cgcaatttc gccgcgcggc	8400
ggagcttate gccattcctg acgagcgcat tctggctatc tataacgcgc tgcgccggtt	8460
ccgtctctcg caggcggagc tgctggcgat cgccgacgag ctggagcaca cctggcatgc	8520
gacagtgaat gccgcctttg tccgggagtc ggcggaagtg tatcagcagc ggcataagct	8580
gcgtaaagga agctaagcgg aggtcagcat gccgttaata gccgggattg atatcgga	8640
cgccaccacc gaggtggcgc tggcgtccga ctaccgcag gcgagggcgt ttgttgccag	8700
cgggatcgtc gcgacgacgg gcatgaaagg gacgcgggac aatatcgccg ggacctcgc	8760
cgcgctggag caggccctgg cgaaaacacc gtggtcgatg agcgatgtct ctgcgcatca	8820

tcttaacgaa gccgcgccgg tgattggcga tgtggcgatg gagaccatca ccgagaccat	8880
tatcaccgaa tcgacatga tcggtcataa cccgcagacg ccgggcgggg tggcggttgg	8940
cgtggggacg actatcgccc tcgggcggct ggcgacgctg ccggcgccgc agtatgccga	9000
ggggtggatc gtactgattg acgacgccgt cgatttcctt gacgccgtgt ggtggctcaa	9060
tgaggcgctc gaccggggga tcaactggtt ggccggcgatc ctcaaaaagg acgacggcgt	9120
gctggtgaac aaccgcctgc gtaaaacctt gccggtggtg gatgaagtga cgctgctgga	9180
gcaggtcccc gagggggtaa tggcggcggtt ggaagtggcc gcgccgggcc aggtggtgctg	9240
gatactgtcg aatccctacg ggatcgccac ctctctcggg ctaagcccgg aagagaccca	9300
ggccatcgct cccatcgccc gcgcctgat tggcaaccgt tccgcggtgg tgctcaagac	9360
cccgcagggg gatgtgcagt cgcggtgat cccggcgggc aacctctaca ttagcggcga	9420
aaagcgccgc ggagaggccg atgtcgccga gggcgcggaa gccatcatgc aggcgatgag	9480
cgcttgcgtt ccggtacgcg acatccgcgg cgaaccgggc acccacgccg gcggcatgct	9540
tgagcgggtg cgcaaggtaa tggcgtccct gaccggccat gagatgagcg cgatatacat	9600
ccaggatctg ctggcggtgg atacgtttat tccgcgcaag gtgcaggcgc ggatggccgg	9660
cgagtgcgcc atggagaatg ccgtcgggat ggccggcgatg gtgaaagcgg atcgtctgca	9720
aatgcaggtt atcgcccgcg aactgagcgc ccgactgcag accgaggtgg tggtagggcg	9780
cgtggaggcc aacatggcca tcgccggggc gtttaaccact cccggctgtg cggcgccgct	9840
ggcgatcctc gacctggcg cccggtcgac ggatgcggcg atcgtcaacg cggaggggca	9900
gataacggcg gtccatctcg ccggggcggg gaatatggtc agcctgttga ttaaaaccga	9960
gctgggcctc gaggatcttt cgctggcgga agcgataaaa aaataccgc tggccaaagt	10020
ggaaagcctg ttcagtattc gtcacagaa tggcgcggtg gatttcttc gggaagccct	10080
cagcccgcg gtgttcgcca aagtgggtga catcaaggag ggcgaaactg tccgatcga	10140
taacgccagc ccgttgaaa aaattcgtct cgtgcgccg caggcgaaag agaaagtgtt	10200
tgtcaccaac tgcctgcgcg cgctgcgcca ggtctcacc ggcggttcca ttcgcgatat	10260
cgcttttgtg gtgctgggtg gcggtcatc gctggacttt gagatccgc agcttatcac	10320
ggaagccttg tcgcactatg gcgtggtcgc cgggcagggc aatattcggg gaacagaagg	10380
gccgcgaat gcggtcgcca ccgggtgct actggccggt caggcgaatt aaacgggcgc	10440
tcgcgccagc ctctaggtaac aaataaaaaa ggcacgtcag atgacgtgcc tttttcttg	10500
tctagcgtgc accaatgctt ctggcgctcag gcagccatcg gaagctgtgg tatggctgtg	10560
caggctgtaa atcactgcat aattcgtgtc gctcaaggcg cactccggtt ctggataatg	10620
tttttgcgc cgacatcata acggttctgg caaatattct gaaatgagct gttgacaatt	10680

aatcatccgg ctctgataat gtgtggaatt gtgagcggat aacaatttca cacaggaaac	10740
agaccatgac tagtaaggag gacaattcca tggtctgtgc tgctgataga ttaaaacttaa	10800
cttccggcca cttgaatgct ggtagaaaga gaagttcctc ttctgtttct ttgaaggctg	10860
ccgaaaagcc tttaagggtt actgtgattg gatctggtaa ctgggggtact actattgcca	10920
aggtggttgc cgaaaattgt aagggaatac cagaagtttt cgctccaata gtacaaatgt	10980
gggtgttcga agaagagatc aatggtgaaa aattgactga aatcataaat actagacatc	11040
aaaacgtgaa atacttgctt ggcatcactc taccgacaa ttgtgttgct aatccagact	11100
tgattgattc agtcaaggat gtcgacatca tcgttttcaa cattccacat caatttttgc	11160
cccgtatctg tagccaattg aaaggtcatg ttgattcaca cgtcagagct atctcctgtc	11220
taaagggttt tgaagtgtgt gctaaagggt tccaattgct atcctcttac atcactgagg	11280
aactaggtat tcaatgttgt gctctatctg gtgctaacat tgccaccgaa gtcgctcaag	11340
aacactggtc tgaacaaca gttgcttacc acattccaaa ggatttcaga ggcgagggca	11400
aggacgtcga ccataagggt ctaaaggcct tgttcacag acctacttc cacgttagtg	11460
tcatcgaaga tttgtctgtt atctccatct gtggtgcttt gaagaacgtt gttgccttag	11520
gttgtgggtt cgtcgaagggt ctaggctggg gtaacaacgc ttctgctgcc atccaagag	11580
tcggtttggg tgagatcacc agattcggtc aaatgttttt ccagaatct agagaagaaa	11640
catactacca agagtctgct ggtgtgtgtg atttgatcac cacctgcgct ggtggtagaa	11700
acgtcaagggt tgctaggcta atggctactt ctggttaagga cgcctgggaa tgtgaaaagg	11760
agttgttgaa tggccaatcc gctcaagggt taattacctg caaagaagtt cacgaatggt	11820
tggaaacatg tggctctgtc gaagacttcc cattatttga agccgtatcc caaatcggtt	11880
acaacaacta ccaatgaag aacctgccgg acatgattga agaattagat ctacatgaag	11940
attagattta ttggatccag gaaacagact agaattatgg gattgactac taaacctcta	12000
tctttgaaag ttaacgccgc ttgtgtcgac gtcgacggta ccattatcat ctctcaacca	12060
gccattgctg cattctggag ggatttcggt aaggacaaac cttatttcga tgctgaacac	12120
gttatccaag tctcgcatgg ttggagaacg ttgatgcca ttgctaagtt cgctccagac	12180
tttgccaatg aagagtaigt taacaaatta gaagctgaaa ttccgggtcaa gtacggtgaa	12240
aatccattg aagtcacagg tgcagttaag ctgtgcaacg ctttgaacgc tctacaaaa	12300
gagaaatggg ctgtggcaac ttccggtacc cgtgatatgg cacaaaaatg gttcgagcat	12360
ctgggaatca ggagacaaaa gtacttcatt accgctaatt atgtcaaaca gggtaacct	12420

catccagaac catatctgaa gggcaggaat ggcttaggat atccgatcaa tgagcaagac 12480
 ccttcctaaat ctaaggtagt agtatttgaa gacgtccag caggtattgc cgccggaaaa 12540
 gccgccggtt gtaagatcat tggattgcc actactttcg acttggactt cctaaaggaa 12600
 aaaggctgtg acatcattgt caaaaaccac gaatccatca gagttggcgg ctacaatgcc 12660
 gaaacagacg aagtgaatt catttttgac gactacttat atgctaagga cgatctgttg 12720
 aaatggtaac ccgggctgca ggcatgcaag ctgggctgtt ttggcggatg agagaagatt 12780
 ttcagcctga tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttgcct 12840

ggcggcagta gcgcggtggt cccacctgac cccatgccga actcagaagt gaaacgccgt 12900
 agcgcgatg gtagtgtggg gtctcccat gcgagagtag ggaactgcca ggcatcaaat 12960
 aaaacgaaag gctcagtcga aagactgggc ctttcgtttt atctgttgtt tgtcggtgaa 13020
 cgctctctg agtaggacaa atccgccggg agcggatttg aacgttgcga agcaacggcc 13080
 cggagggtgg cgggcaggac gcccgcata aactgccagg catcaaatta agcagaaggc 13140
 catcctgacg gatggccttt ttgcgtttct acaaaactcca gctggatcgg gcgctagagt 13200
 atacatttaa atggtaccct ctagtcaagg ccttaagtga gtcgtattac ggactggccg 13260

tcgttttaca acgtcgtgac tgggaaaacc ctggcggtac ccaacttaat cgccttgacg 13320
 cacatcccc tttgccagc tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc 13380
 aacagttgcg cagcctgaat ggcgaaatggc gctgatgag gtattttctc cttacgcatc 13440
 tgtgcggtat ttacaccgc atatggtgca ctctcagtac aatctgctct gatgccgcat 13500
 agttaagcca gccccgacac ccgccaacac ccgctgacga gct 13543

<210> 64

<211> 13543

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid

<400> 64

tagtaaagcc ctgctagat tttaatgcgg atgttgcat tacttcgcca actattgcga 60

taacaagaaa aagccagcct ttcagatgat atctcccaat ttgtgtaggg cttattatgc 120
 acgcttaaaa ataataaaag cagacttgac ctgatagttt ggctgtgagc aattatgtgc 180
 ttagtgatc taacgcttga gttaagccgc gccgcgaagc ggctgcggct tgaacgaatt 240
 gttagacatt atttgccgac taccttggtg atctcgctt tcacgtagt gacaaattct 300
 tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata agcctgtcta 360
 gcttcaagta tgacgggctg atactgggcc ggaggcgct ccattgccca gtcggcagcg 420

acatccttcg gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc 480

actacatttc gctcatcgcc agcccagtcg ggcggcgagt tccatagcgt taaggtttca 540

tttagcgcct caaatagatc ctgttcagga accggatcaa agagtccctc cgccgctgga 600

cctaccaagg caacgctatg ttctcttgct ttgttcagca agatagccag atcaatgtcg 660

atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc 720

agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact 780

tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttcaa aaggtcgttg 840

atcaaagctc gccgcgttgt ttcatcaagc cttacggtca ccgtaaccag caaatcaata 900

tcactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac 960

gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt cgatacttcg 1020

gcgatcaccg ctccctcat gatgtttaac ttgttttag ggcgactgcc ctgctgcgta 1080

acatcgttgc tgctccataa catcaaacat cgaccacgg cgtaacgcgc ttgctgcttg 1140

gatgcccag gcatagactg taccctaaaa aaacagtcac aacaagccat gaaaaccgcc 1200

actgcgccgt taccaccgt gcgttcggtc aaggttctgg accagttgcg tgagcgcata 1260

cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc 1320

atccgtttcc acggtgtgcg taccgccgca accttgggca gcagcgaagt cgaggcattt 1380

ctgtctctggc tggcgaacga gcgcaagggt tcggtctcca cgcatcgtca ggcatitggcg 1440

gccttctgtg tcttctacgg caaggtgctg tgcacggatc tgccctggct tcaggagatc 1500

ggaagacctc ggccgtcgcg gcgcttgccg gtgggtgctga ccccgatga agtggttcgc 1560

atcctcggtt ttctggaagg cgagcatcgt ttgttcgcc agcttctgta tggaacgggc 1620

atgcggatca gtgagggttt gcaactgcgg gtcaaggatc tggatttcga tcacggcacg 1680

atcatcgtgc gggagggcaa gggctccaag gatcgggcct tgatgttacc cgagagcttg 1740

gcaccacgcc tgcgcgagca ggggaattaa ttcccacggg ttttctgcc cgcaaacggg 1800

ctgttctggt gttgctagtt gtttatcaga atcgcagatc cggcttcagc cggtttgccg 1860

gtgaaaagcg ctatttcttc cagaattgcc atgatttttt cccacggga ggcgtcactg 1920

gtcccggtg tgcggcagc ttgattcga taagcagcat cgcctgttcc aggtgtctta 1980

tgtgtgactg ttgagctgta acaagtgtc tcagggttcc aatttcattg tctagttgct 2040

ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcattc gttacattgt 2100

cgatctgttc atggtgaaca gctttgaatg caccataaac tcgtaaaagc tctgatgtat 2160

ctatcttttt tacaccgttt tcatctgtgc atatggacag ttttcccttt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttactgata gatacaagag	2280
ccataagaac ctcagatcct tccgtattta gccagtatgt tctctagtgt ggttcgttgt	2340
ttttgcgtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt ttgacagtta aagcatcgtg tagtgttttt	2460
cttagtcctg tatgtaggta ggaatctgat gtaatggttg ttggtatttt gtcaccattc	2520
atttttatct gggtgttctc aagttcgggt acgagatcca ttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccattt ggtaaagcct tttaaaactca	2700
tggtagttaa ttcaagcat taacatgaac ttaaattcat caaggctaata ctctatattt	2760
gccttgtgag ttttcttttg tgttagttct ttaataaacc actcataaat cctcatagag	2820
tatttgtttt caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880
aaaagataag gcaatatctc ttactataaa actaattcta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag ctttaacca aaggattcct gatttccaca	3000
gttctcgtca tcagctctct gggtgcttta gctaatacac cataagcatt ttcctactg	3060
atgttcatca tctgagcgta ttggtataaa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tcgtgggggt gagtagtgcc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attggtctag gtgattttaa tcactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtcccttt cctttgagtt gtgggtatct gtaaattctg ctagaccttt	3360
gctggaaaac ttgtaaattc tgctagacct tctgtaaatt ccgctagacc ttgtgtgtt	3420
ttttttgttt atattcaagt gggtataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcactacttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaagctt gtttgcctct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca cctcgcaag ctcgggcaaa tcgctgaata ttctttttgt ctccgacct	3660
caggcacctg agtcgctgtc ttttctgtga cattcagttc gctgcgctca cggctctggc	3720
agtgaatggg ggtaaatggc actacaggcg ctttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccg tcacgggctt ctcaggcggt tttatggcgg gtctgctatg	3840
tggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtagc aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggccttacc gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020

cgaccggagg cttttgactg ctaggcgatc tgtgctgttt gccacggtat gcagcaccag 4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcactggaac gagaagtcag 4140
gcgagccgtc acgcccttga caatgccaca tcctgagcaa ataattcaac cactaaacaa 4200
atcaaccgcg tttcccgagg gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa 4260

caagttcaga caatcacctt ggccgccgcc cagcaaatgg cggcggcggt ggaaaaaaaaa 4320
gccactgaga tcaacgtggc ggtggtgttt tccgtagtgt accgcggagg caacacgtctg 4380
cttatccagc ggatggacga ggcttcgtc tccagctgcg atatttcct gaataaagcc 4440
tggagcgct gcagcctgaa gcaaggtacc catgaaatta cgtcagcggt ccagccagga 4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatttt tggcggcggc 4560
ctgccagtta ttttaatga gcaggtaatt ggcccgctcg gcgttagcgg cggtacggtc 4620
gagcaggatc aattattagc ccagtgcgcc ctggattgtt tttccgatt ataacctgaa 4680

gcgagaaggt atattatgag ctatcgatg ttccgccagg cattctgagt gttaacgagg 4740
ggaccgtcat gtcgtttca ccgccaggcg tacgctgtt ttacgatccg cgcgggcacc 4800
atgccggcg catcaatgag ctgtgctggg ggtgggagga gcagggggtc ccctgccaga 4860
ccataaccta tgacggaggc ggtgacgccg ctgcgtggg cgccctggcg gccagaagct 4920
cgccctcgcg ggtgggtatc gggctcagcg cgtccggcga gatagccctc actcatgcc 4980
agctgccggc ggacgcgccg ctggctaccg gacacgtcac cgatagcgac gatcaactgc 5040
gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact 5100

gaatgtatcg tatctatacc cgcaccgggg ataaaggcac caccgccctg tacggcgga 5160
gccgcatcga gaaagacat attcgcgtcg aggcctacgg caccgtcgat gaactgatat 5220
cccagctggg cgtctgtac gccacgacc gcgacgccg gctgcgggaa agcctgcacc 5280
atattcagca gacgtgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga 5340
ccgccttag ccagacgatc ggcaagagg agatcaccg cctggagcgg ctatcgacc 5400
gcaatatggc cgagagcgcc ccgttaaaac agttcgtgat cccggggagg aatctgcct 5460
ctgccagct gcacgtggcg cgcaccagc cccgtcggct cgaacgcctg ctgacggcca 5520

tggaccgcg gcatccgtg cgcacgcgc taaacgcta cagcaatcgc ctgtcggatg 5580
ccctgttctc catggcgca atcgaagaga ctaggcctga tgcttgcgt tgaactggc 5640
tagcaaacac agaaaaagc ccgcacctga cagtgcgggc ttttttttc ctaggcgatc 5700
tgtgctgttt gccacggtat gcagcaccag cgcgagatta tgggctcgca cgctcgactg 5760
tcggacgggg gcactggaac gagaagtcag gcgagccgtc acgcccttga caatgccaca 5820
tcctgagcaa ataattcaac cactaaacaa atcaaccgcg tttcccgagg gtaaccaagc 5880

ttcacctttt gagccgatga acaatgaaaa gatcaaaacg atttcagta ctggcccagc	5940
gccccgtcaa tcaggacggg ctgattggcg agtggcctga agaggggctg atcgccatgg	6000
acagcccctt tgacccggtc tcttcagtaa aagtggacaa cggctctgac gtcgaactgg	6060
acggcaaacg ccgggaccag tttgacatga tcgaccgatt tatcgccgat tacgcgatca	6120
acgttgagcg cacagagcag gcaatgcgcc tggaggcggg ggaatatagcc cgtatgctgg	6180
tggatattca cgtcagccgg gaggagatca ttgccatcac taccgccatc acgccggcca	6240
aagcggctga ggtgatggcg cagatgaacg tggaggagat gatgatggcg ctgcagaaga	6300
tgctgcccc ccggaccccc tccaaccagt gccacgtcac caatctcaa gataatccgg	6360
tgcagattgc cgctgacgcc gccgaggccg ggatccgcgg cttctcagaa caggagacca	6420
cggtcggtat cgcgcgtac gcgccgttta acgccctggc gctgttggtc ggttcgcagt	6480
gcggccgccc cggcgtgttg acgcagtgtc cggagggaaga ggccaccgag ctggagctgg	6540
gcatgcgtgg cttaaccagc tacgccgaga cgggtgtcgg ctacggcacc gaagcggtat	6600
ttaccgacgg cgatgatacg ccgtgggtcaa aggcgttctt cgcctcggcc tacgcctccc	6660
gcgggttgaa aatgcgtac acctccggca ccggatccga agcgtgatg ggctattcgg	6720
agagcaagtc gatgctctac ctcgatcgc gctgcatctt cattactaaa ggccgcccggg	6780
ttcagggact gcaaaacggc gcggtgagct gtatcgcat gaccggcgt gtgccgtcgg	6840
gcattcgggc ggtgctggcg gaaaacctga tcgcctctat gctcgacctc gaagtggcgt	6900
ccgccaacga ccagacttcc tccactcgg atattcgccg caccgcgcgc accctgatgc	6960
agatgctgcc gggcaccgac tttattttct ccggctacag cgcggtgccg aactacgaca	7020
acatgttcgc cggctcgaac ttcatgctgg aagattttga tgattacaac atcctgcagc	7080
gtgacctgat ggttgacggc ggcctgcgtc cggtgaccga ggccgaaacc attgccattc	7140
gccagaaagc ggccgcccgc atccaggcgg ttttccgca gctggggctg ccgccaatcg	7200
ccgacgagga ggtggaggcc gccacctacg cgcacggcag caacgagatg ccgccgcgta	7260
acgtggtgga ggtatctgagt gcggtggaag agatgatgaa gcgcaacatc accggcctcg	7320
atattgtcgg cgcgctgagc cgcagcggct ttgaggatat cgccagcaat attctcaata	7380
tgctgcgcca gcgggtcacc ggcgattacc tcgagacctc ggccattctc gatcggcagt	7440
tcgaggtggt gagtgcggtc aacgacatca atgactatca ggggccgggc accggctatc	7500
gcatctctgc cgaacgttgg gcggagatca aaaatatcc gggcgtgggt cagcccagaca	7560
ccattgaata aggcggtatt cctgtgcaac agacaacca aattcagccc tcttttacc	7620

tgaaaacccg cgaggcgagg gtagcttctg ccgatgaacg cgccgatgaa gtggtgatcg	7680
gcgtcggccc tgccttcgat aaacaccagc atcacactct gatcgatatg ccccatggcg	7740
cgatcctcaa agagctgatt gccggggtgg aagaagaggg gcttcacgcc cgggtggtgc	7800
gcattctgcg cactgccgac gtctccttta tggcctggga tgcggccaac ctgagcggct	7860
cggggatcgg catcggtatc cagtcgaagg ggaccacggt catccatcag cgcgatctgc	7920
tgcgctcag caacctggag ctgttctccc aggcgcgct gctgacgctg gagacctacc	7980
ggcagattgg caaaaacgct gcgcgctaig cgcgcaaaga gtcaccttcg ccggtgccgg	8040
tggtgaacga tcagatggtg cgcccgaaat ttatggccaa agccgcgcta ttcatatca	8100
aagagaccaa acatgtggtg caggacgccg agcccgctac cctgcacatc gacttagtaa	8160
gggagtgcac atgagcgaga aaacctgctg cgtgcaggat tatccgttag ccaccgctg	8220
cccggagcat atcctgacgc ctaccggcaa accattgacc gatattaccc tcgagaaggt	8280
gtctcttgcc gaggtgggcc cgcaggatgt gcggatctcc cgccagaccc ttgagtacca	8340
ggcgcagatt gccgagcaga tgcagcgcca tgcggtggcg cgcaatttcc gccgcgcggc	8400
ggagcttacc gccattcctg acgagcgcat tctggctatc tataacgcgc tgcgcccggt	8460
ccgctcctcg caggcggagc tgctggcgat cgccgacgag ctggagcaca cctggcatgc	8520
gacagtgaat gccgcctttg tccgggagtc ggcggaagtg tatcagcagc ggcataagct	8580
gcgtaaagga agctaagcgg aggtcagcat gccgttaata gccgggattg atatcgga	8640
cgccaccacc gaggtggcgc tggcgctcga ctaccgcag cgcaggcgct ttgttgccag	8700
cgggatcgtc gcgacgacgg gcatgaaagg gacgcgggac aatatcgccg ggaccctgc	8760
cgcgctggag caggccctgg cgaaaacacc gtggtcgatg agcgatgtct ctgcatcta	8820
tcttaacgaa gccgcgccgg tgattggcga tgtggcgatg gagaccatca ccgagaccat	8880
tatcaccgaa tcgaccatga tcggtcataa ccgcgagacg ccggcggggg tggcggttg	8940
cgtggggacg actatcgccc tcggcgcgct ggcgacgctg ccggcgggcg agtatgccga	9000
ggggtggatc gtactgattg acgacgccgt cgatttcctt gacgccgtgt ggtggctcaa	9060
tgaggcgctc gaccggggga tcaacgtggt ggcgcgcatc ctcaaaaagg acgacggcgt	9120
gctggtgaac aaccgcctgc gtaaacctt gccggtggtg gatgaagtga cgctgctgga	9180
gcaggtcccc gagggggtaa tggcgcggt ggaagtggcc gcgccgggac aggtggtgcg	9240
gacctgtcgt aatccctacg ggatcgccac ctctctcggt ctaagcccg aagagacca	9300
ggccatcgtc cccatcgccc gcgccctgat tggcaaccgt tcccggtgg tgctcaagac	9360
cccgcagggg gatgtgcagt cgcggtgat cccggcgggc aacctctaca ttagcgcgga	9420
aaagcgccgc ggagaggccg atgtcgccga gggcgcgga gccatcatgc aggcgatgag	9480

cgcttgcgct ccggtacgcg acatccgagg cgaaccgggc acccacgccc gcggcatgct	9540
tgagcgggtg cgcaaggtaa tggcgctcct gaccggccat gagatgagcg cgatatacat	9600
ccaggatctg ctggcgggtg atacgtttat tccgcgcaag gtgcaggcg ggatggccgg	9660
cgagtgcgcc atggagaatg ccgtcgggat ggcggcgatg gtgaaagcgg atcgtctgca	9720
aatgcaggtt atcgcccgcg aactgagcgc ccgactgcag accgaggtgg tggtagggcg	9780
cgtggaggcc aacatggcca tcgccggggc gttaccact cccggctgtg cggcgccgct	9840
ggcgatcctc gacctggcg ccggctcgac ggatgcggcg atcgtcaacg cggaggggca	9900
gataacggcg gtccatctcg ccggggcggg gaatatgttc agcctgttga ttaaaaccga	9960
gctgggcctc gaggatcttt cgctggcgga agcgataaaa aaataccgc tggccaaagt	10020
ggaaagcctg ttcagtattc gtcacagaa tggcgcggtg gatttcttc gggaagccct	10080
cagcccgcg gtgttcgcca aagtgggtga catcaaggag ggcgaactgg tgccgatcga	10140
taacgccagc ccgctggaaa aaattcgtct cgtgcgccgg caggcgaaag agaaagtgtt	10200
tgtcaccaac tgcctgcgcg cgctgcgcca ggtctcacc cggggttcca ttcgcgatat	10260
cgcttttgtg gtgtgtgtg gcggctcatc gctggacttt gagatccgc agcttatcac	10320
ggaagcctt tgcactatg gcgtggtcgc cgggcagggc aatattcggg gaacagaagg	10380
gccgcgcaat gcggtcgcca ccgggctgct actggccggt caggcgaatt aaacgggcgc	10440
tcgcgccagc ctctaggtac aaataaaaaa ggcacgtcag atgacgtgcc tttttcttg	10500
tctagcgtgc accaatgctt ctggcgtcag gcagccatcg gaagctgtgg tatggctgtg	10560
caggtcgtaa atcactgcat aattcgtgtc gctcaaggcg cactcccgtt ctggataatg	10620
ttttttgcgc cgacatcata acggttctgg caaatattct gaaatgagct gttgacaatt	10680
aatcatccgg ctctataat gtgtggaatt gtgagcgat aacaatttca cacaggaaac	10740
agaccatgac tagtaaggag gacaattcca tggctgtgc tgcgataga ttaaaactaa	10800
cttccggcca ctigaatgct ggtagaaaga gaagttctc ttcgtttct ttgaaggctg	10860
ccgaaaagcc tttcaagggt actgtgattg gatctggtaa ctgggggtact actattgcca	10920
aggtggttgc cgaaaattgt aagggtatcc cagaagtgtt cgctccaata gtacaaatgt	10980
gggtgttcga agaagagatc aatggtgaaa aattgactga aatcataaat actagacatc	11040
aaaacgtgaa atacttgctt ggcatcactc tacccgacaa tttggttgct aatccagact	11100
tgattgattc agtcaaggat gtcgacatca tcgttttcaa cattccacat caatttttgc	11160
cccgtatctg tagccaattg aaaggtcatg ttgattcaca cgtcagagct atctcctgtc	11220
taaagggttt tgaagtgtgt gctaaagggt tccaattgct atcctcttac atcactgagg	11280
aactaggtat tcaatgttgt gctctatctg gtgctaacaat tgccaccgaa gtcgctcaag	11340

aacactggtc tgaacaaca gttgcttacc acattccaaa ggatttcaga ggcgagggca	11400
aggacgtcga ccataaggtt ctaaaggcct tgtccacag acctacttc cacgttagtg	11460
tcatcgaaga tgttgcctgt atctccatct gtgggtgttt gaagaacgtt gttgccttag	11520
gttgtggttt cgtcgaaggt ctaggtggg gtaacaacgc ttcgtctgcc atccaaagag	11580
tcggtttggg tgagatcatc agattcgttc aaatgttttt ccagaatct agagaagaaa	11640
catactacca agagtctgct ggtgttctg atttgatcac cacctgcgt ggtggtagaa	11700
acgtcaaggt tgctaggcta atggctactt ctggaagga cgcctgggaa tgtgaaaagg	11760
agttgttgaa tggccaatcc gctcaaggt taattacctg caaagaagtt cacgaatggt	11820
tggaacatg tggctctgtc gaagacttcc cattatttga agccgtatac caaatcgttt	11880
acaacaacta cccaatgaag aacctgccgg acatgattga agaattagat ctacatgaag	11940
attagattta ttggatccag gaaacagact agaattatgg gattgactac taaacctcta	12000
tctttgaaag ttaacgccgc tttgttcgac gtcgacggta ccattatcat ctctcaacca	12060
gccattgctg cattctggag ggatttcggt aaggacaaac cttatttcga tgctgaacac	12120
gttatccaag tctcgcattg ttggagaacg ttgatgccca ttgctaagtt cgctccagac	12180
tttgccaatg aagagtatgt taacaaatta gaagctgaaa ttccggtcaa gtacggtgaa	12240
aatccattg aagtcacagg tgcagttaag ctgtgcaacg ctttgaacgc tctacaaaa	12300
gagaaatggg ctgtggcaac ttccggtacc cgtgatatgg cacaaaaatg gttcgagcat	12360
ctgggaatca ggagacaaa gtacttcatt accgctaag atgtcaaaca gggtaacct	12420
catccagaac catatctgaa gggcaggaat ggcttaggat atccgatcaa tgagcaagac	12480
ccttccaaat ctaaggtagt agtatttgaa gacgtccag caggtattgc cgccggaaaa	12540
gccgccggtt gtaagatcat tggattgcc actactttcg acttggactt cctaaaggaa	12600
aaaggtgtg acatcattgt caaaaaccac gaatccatca gatttggcgg ctacaatgcc	12660
gaaacagacg aagtgaatt catttttgac gactacttat atgctaagga cgatctgttg	12720
aatggtaac cggggtgca ggcattgcaag ctgggtgtt ttggcggatg agagaagatt	12780
ttcagcctga tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttcct	12840
ggcggcagta gcgcggtgtt ccacactgac cccatgccga actcagaagt gaaacgccgt	12900
agcgcgatg gtagtgtggg gtctcccat gcgagagtag ggaactgcc ggcattcaat	12960
aaaacgaaag gtcagtcga aagactgggc ctttcgtttt atctgttgtt tgtcggtgaa	13020
cgctctctg agtaggacaa atccgccggg agcggatttg aacgttcga agcaacggcc	13080

cggagggtgg cgggcaggac gcccgcata aactgccagg catcaaatta agcagaaggc 13140
catcctgacg gatggccttt ttgcgtttct acaaaactcca gctggatcgg gcgctagagt 13200
atacatitaa atggtaccct ctagtcaagg ccttaagtga gtcgtattac ggactggccg 13260
tcgttttaca acgtcgtgac tgggaaaacc ctggcggttac ccaacttaat cgccttgacg 13320
cacatcccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgat cgccttcccc 13380
aacagtigcg cagcctgaat ggcgaaatggc gcctgatgcg gtattttctc cttacgcac 13440
tgtgcggtat ttcacaccgc atatggtgca ctctcagtac aatctgctct gatgccgcat 13500

agttaagcca gccccgacac ccgccaacac ccgctgacga gct 13543

<210> 65

<211> 13402

<212> DNA

<213> Artificial Sequence

<220><223> Plamid

<400> 65

tagtaaagcc ctgctagat tttaatgcgg atgttgcat tacttcgcca actatigcga 60
taacaagaaa aagccagcct ttcattgatat atctccaat ttgttaggg cttattatgc 120
acgtttaaaa ataataaaag cagactigac ctgatatgtt ggctgtgagc aattatgtgc 180
ttagtgcatc taacgcttga gttaagccgc gccgcgaagc ggctgcggct tgaacgaatt 240
gttagacatt atttgccgac taccttggtg atctgcctt tcacgtagtg gacaaattct 300

tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata agcctgtcta 360
gcttcaagta tgacgggctg atactgggcc ggcaggcgct ccattgccca gtcggcagcg 420
acatccttcg gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc 480
actacatttc gtcacatgcc agcccagtcg ggccggcgagt tccatagcgt taaggtttca 540
tttagcgct caaatagatc ctgttcagga accggatcaa agagtctctc cgcgcgtgga 600
cctaccaagg caacgctatg ttctcttgc tttgtcagca agatagccag atcaatgtcg 660
atcgtggctg gtcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc 720

agttcgcgt tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact 780
tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttccaa aaggtcgttg 840
atcaaagctc gccgcgttgt ttcattcaagc cttacggtca ccgtaaccag caaatcaata 900
tactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac 960
gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt cgatacttcg 1020
gcgatcaccg cttccctcat gatgtttaac ttgttttag ggcgactgcc ctgctgcgta 1080

acatcgttgc tgctccataa catcaaacat cgacccacgg cgtaacgcgc ttgctgcttg	1140
gatgcccagag gcatagactg taccceaaaa aaacagtcac aacaagccat gaaaaccgcc	1200
actgcgccgt taccaccgct gcgttcggtc aaggttcttg accagttgcg tgagcgcata	1260
cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc	1320
atccgtttcc acggtgtgcg tcacccggca accttgggca gcagcgaagt cgaggcattt	1380
ctgtcctggc tggcgaacga gcgcaagggt tcggtctcca cgcatcgta ggcattggcg	1440
gccttctgtg tcttctacgg caaggtgctg tgcacggatc tgccttggtc tcaggagatc	1500
ggaagacctc ggccgtcgcg gcgcttgccg gtggtgctga ccccgatga agtggttcgc	1560
atcctcgggt ttctggaagg cgagcatcgt ttgttcgccc agcttctgta tggaaagggc	1620
atgcggatca gtgagggttt gcaactgcgg gtcaaggatc tggatttcta tcacggcacg	1680
atcatcgtgc gggaggcgaa gggctccaag gatcgggcct tgatgttacc cgagagcttg	1740
gcaccagacc tgcgcgagca ggggaattaa ttcccacggg ttttctgcc cgcaaacggg	1800
ctgttctggt gttgctagtt tgttatcaga atcgagatc cggcttcagc cggtttgccg	1860
gctgaaagcg ctatttcttc cagaattgcc atgatttttt cccacggga ggcgtcactg	1920
gctcccggtg tctcggcagc ttgtattcga taagcagcat cgcctgttc aggcgttcta	1980
tgtgtactg ttgagctgta acaagtgtc tcagggtgtc aatttcattg tctagtgtg	2040
ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcattc gttacattgt	2100
cgatctgttc atggtgaaca gctttgaatg caccaaaaac tcgtaaaagc tctgatgtat	2160
ctatcttttt tacaccgttt tcatctgtgc atatggacag ttttcccttt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttactgata gatacaagag	2280
ccataagaac ctcatgctc tccgtattta gccagtatgt tctctagtgt ggttcgttgt	2340
ttttgcgtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt ttgacgtta aagcatcgtg tagtgttttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatggttg ttggtatttt gtcaccattc	2520
atttttatct ggttgttctc aagttcgggt acgagatcca tttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccattt ggttaagcct tttaaactca	2700
tggtagttaa ttcaagcat taacatgaac ttaaattcat caaggctaata ctctatattt	2760
gccttgtgag ttttcttttg tgttagttct ttaataaacc actcataaat cctcatagag	2820

tatttgTTTT caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880
aaaagataag gcaatatctc ttcactaaaa actaattcta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag cctttaacca aaggattcct gatttccaca	3000
gttctcgtca tcagctctct gggtgcttta gctaatacac cataagcatt ttccttactg	3060
atgttcatca tctgagcgta ttggttataa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tctgtggggtt gagtagtgcc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attgggtctag gtgattttta tctactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtccTTTT cctttgagtt gtgggtatct gtaaattctg ctagaccttt	3360
gctggaaaac ttgtaaattc tgctagacc cctgtaaatt ccgctagacc tttgtgtgtt	3420
ttttttgttt atattcaagt gggtataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tctactcttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaagct gtttgcctct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca ccttcgcaag ctcgggcaaa tcgctgaata ttctttttgt ctccgaccat	3660
caggcacctg agtcgctgtc tttttcgtga cattcagttc gctgcgctca cggtctctggc	3720
agtgaatggg ggtaaatggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccc tcacgggctt ctacgggctt tttatggcgg gtctgctatg	3840
tgggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggccttacc gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020
cgaccggagg cttttgactg ctaggcgatc tgtgctgttt gccacggtat gcagcaccag	4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcaactggaac gagaagtcag	4140
gcgagccgtc acgcccctga caatgccaca tcttgagcaa ataattcaac cactaaciaa	4200
atcaaccgcg tttcccgag gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa	4260
caagttcaga caatcacctt ggccgccgcc cagcaaatgg cggcggcggg ggaaaaaaaa	4320
gccactgaga tcaacgtggc ggtggtgttt tccgtagtgt accgcggagg caacacgctg	4380
cttatccagc ggatggacga ggcttctgtc tccagctgcg atatttcct gaataaagcc	4440
tggagcgcct gcagcctgaa gcaaggtagc catgaaatta cgtcagcggg ccagccagga	4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatatt tggcggcggc	4560
ctgccagtta tttttaatga gcaggttaatt ggccgccgtc gcgttagcgg cggtagcgtc	4620
gagcaggatc aattattagc ccagtgcgcc ctggattgtt tttccgcatt ataacctgaa	4680

gcgagaaggt atattatgag ctatcgatg ttcgccagg cattctgagt gttaacgagg 4740
ggaccgtcat gtcgctttca ccgccaggcg tacgcctgtt ttacgatccg cgcgggcacc 4800
atgccggcgc catcaatgag ctgtgctggg ggctggagga gcagggggtc ccctgccaga 4860
ccataaccta tgacggagggc ggtgacggcg ctgcgtggg cgccttggcg gccagaagct 4920

cgcccttgcg ggtgggtatc gggctcagcg cgtccggcga gatagccctc actcatgccc 4980
agctgccggc ggacgcggcg ctggctaccg gacacgtcac cgatagcgac gatcaactgc 5040
gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact 5100
gaatgtatcg tatctatacc cgcaccgggg ataaaggcac caccgccctg tacggcggca 5160
gccgcatcga gaaagacat attcgcgtcg aggcctacgg caccgtcgat gaactgatat 5220
cccagctggg cgtctgttac gccacgacc cgcacggcg gctgcgggaa agcctgcacc 5280
atattcagca gacgctgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga 5340

ccgccttag ccagacgac ggcaagagg agatcaccgc cctggagcgg cttatcgacc 5400
gcaatatggc cgagagcggc ccgttaaac agttcgtgat cccggggagg aatctcgctt 5460
ctgccagct gcacctgat gcttgcgtt gaactggcct agcaaacaca gaaaaagcc 5520
cgcacctgac agtgcgggt tttttttcc taggcgatct gtgctgttg ccacggtatg 5580
cagcaccagc gcgagattat gggctcgcac gctcgactgt cggacggggg cactggaacg 5640
agaagtcagg cgagccgtca cgcccttgac aatgccacat cctgagcaaa taattcaacc 5700
actaacaata tcaaccgctt ttcccggagg taaccaagct tcacctttt agccgatgaa 5760

caatgaaaag atcaaaacga ttgcatgtac tggcccagcg ccccgtaat caggacgggc 5820
tgattggcga gtggcctgaa gaggggctga tcgcatgga cagcccttt gaccgggtct 5880
cttcagtaaa agtggacaac ggtctgatcg tcgaactgga cggcaaacgc cgggaccagt 5940
ttgacatgat cgaccgattt atcgccgatt acgcatcaa cgttgagcgc acagagcagg 6000
caatgcgcct ggaggcggtg gaaatagccc gtatgctggt ggatattcac gtcagccggg 6060
aggagatcat tgccatcact accgccatca cgcgggcaa agcggtcgag gtgatggcgc 6120
agatgaacgt ggtggagatg atgatggcgc tgcagaagat gcgtgcccgc cggacccctt 6180

ccaaccagt ccacgtcacc aatctcaaag ataatccggt gcagattgcc gctgacgccg 6240
ccgagccgg gatccggcg tctcagaac aggagaccac ggtcgggtatc gcgcgtacg 6300
cgccgtttaa cgccctggcg ctgttggtcg gttcgcagt cggccgcccc ggcgtgttga 6360
cgcagtgtc ggtggaagag gccaccgagc tggagctggg catgcgtggc ttaaccagt 6420
acgccgagac ggtgtcggtc tacggcaccg aagcgggtatt taccgacggc gatgatacgc 6480
cgtggtcaaa ggctttctc gcctcggcct acgcctccc cgggttgaaa atgcgtaca 6540

cctccggcac cggatccgaa gcgctgatgg gctattcggg gagcaagtcg atgctctacc 6600

tcgaatcgcg ctgcatcttc attactaaag gcgccggggt tcagggactg caaaacggcg 6660

cggtagagctg tatcgcatg accggcgctg tgcgctcggg cattcgggcg gtgctggcgg 6720

aaaacctgat cgctctatg ctcgacctcg aagtggcgtc cgccaacgac cagactttct 6780

cccactcgga tattcgccg accgcgcgca cctgatgca gatgctgccg ggcaccgact 6840

ttattttctc cggtacagc gcggtgccga actacgaca catgttcgcc ggctcgaact 6900

tcgatgcgga agattttgat gattacaaca tcttcgagcg tgacctgat gtgacggcg 6960

gcctgcgtcc ggtgaccgag gcggaacca ttgccattcg ccagaaagcg gcgcgggcga 7020

tccaggcgggt tticcgcgag ctggggctgc cgccaatcg cgacgaggag gtggaggccg 7080

ccacctacgc gcacggcagc aacgagatgc cgcccgctaa cgtggtggag gatctgagt 7140

cgggtgaaga gatgatgaag cgcaacatca ccggcctcga tattgtcggc gcgctgagcc 7200

gcagcggctt tgaggatc gccagcaata ttctcaatat gctgcgccag cgggtcaccg 7260

gcgattacct gcagacctcg gccattctcg atcggcagtt cgaggtggtg agtgcggtca 7320

acgacatcaa tgactatcag gggccgggca ccggctatcg catctctgcc gaacgtggg 7380

cggagatcaa aaatattccg ggcgtgggtc agcccacac cattgaataa ggcggtattc 7440

ctgtgcaaca gacaacccaa attcagccct cttttaccct gaaaaccgc gagggcgggg 7500

tagctttctg cgatgaacgc gccgatgaag tggtagtcgg cgtcgccct gccttcgata 7560

aacaccagca tcacactctg atcgatatgc cccatggcg gatcctcaa gagctgattg 7620

ccgggggtgga agaagagggg cttcacgcc gggtgggtcg cattctgcgc acgtccgacg 7680

tctcctttat ggcttgggat gcggccaacc tgagcggtc ggggatcggc atcggtatcc 7740

agtcaaggg gaccacggtc atccatcagc gcgatctgct gccgctcagc aacctggagc 7800

tgttctccca ggcgcgctg ctgacgctgg agacctaccg gcagattggc aaaaacgctg 7860

cgcgctatgc gcgcaaagag tcaccttcgc cggtgccggg ggtgaacgat cagatggtgc 7920

ggccgaaatt tatggccaaa gccgcgctat ttcatatcaa agagaccaa catgtggtgc 7980

aggacccga gccgctcacc ctgcacatcg acttagtaag ggagtgacca tgagcgagaa 8040

aacctgcgc gtgcaggatt atccgttagc caccgctgc ccggagcata tctgacgcc 8100

taccggcaaa ccattgaccg atattacct cgagaagggt cttcttggcg aggtgggccc 8160

gcaggatgtg cggatctccc gccagaccct tgagtaccag gcgcagattg ccgagcagat 8220

gcagcgccat gcggtggcg gcaatttccg ccgcgcggcg gagcttatcg ccattcctga 8280

cgagcgcatc ctggctatct ataacgcgct gcgcccgttc cgctcctcgc aggcggagct	8340
gctggcgatc gccgacgagc tggagcacac ctggcatgcg acagtgaatg ccgcctttgt	8400
ccgggagtcg gcggaagtgt atcagcagcg gcataagctg cgtaaaggaa gctaagcggg	8460
ggtcagcatg ccgttaatat ccgggattga tatcggaac gccaccaccg aggtggcgct	8520
ggcgccgac taccgcagg cgaggcggtt tgttgccagc gggatcgtcg cgacacggg	8580
catgaaaggg acgcgggaca atatcgccgg gaccctcgcc gcgctggagc aggcctggc	8640
gaaaacaccg tggtcgatga gcgatgtctc tcgcatctat cttaacgaag ccgcgccgtt	8700
gattggcgat gtggcgatgg agaccatcac cgagaccatt atcaccgaat cgacatgat	8760
cggtcataac ccgcagacgc cgggcggggt gggcgttggc gtggggacga ctatcgccct	8820
cgggcggctg gcgacgctgc cggcgccgca gtatgccgag gggaggatcg tactgattga	8880
cgacgccgc gatttcttg acgcctgtg gtggctcaat gaggcgctcg accgggggat	8940
caacgtgggt gcggcgatcc tcaaaaagga cgacggcggt ctggtgaaca accgcctgcg	9000
taaaaccctg ccggtgggtg atgaagtac gctgctggag caggccccg agggggtaat	9060
ggcggcggtg gaagtggccg cgccgggcca ggtggtgcgg atcctgtcga atccctacgg	9120
gatcgccacc ttcttcgggc taagcccga agagaccag gccatcgcc ccatcgcccc	9180
cgccctgatt ggcaaccgtt ccgcgggtgt gctcaagacc ccgcaggggg atgtgcagtc	9240
gcgggtgatc ccggcgggca acctctacat tagcggcgaa aagcgccgcg gagaggccga	9300
tgtcgccgag ggcgcggaag ccatcatgca ggcatgagc gcctgcgctc cggtagcgca	9360
catccgggc gaaccgggca ccacgccgg cgcatgctt gagcgggtgc gcaagtaat	9420
ggcgccctg accggccatg agatgagcg gatatacatc caggatctgc tggcggtgga	9480
tacgtttatt ccgcgaagg tgcaggcggt gatggccgc gagtgcgcca tggagaatgc	9540
cgtcgggatg gcggcgatgg tgaaagcgga tcgtctgcaa atgcaggtta tcgcccgcga	9600
actgagcgcc cgactgcaga ccgaggtgtt ggtggcggc gtggaggcca acatggccat	9660
cgccggggcg ttaaccactc ccggctgtgc ggcccgctg gcgaccccg acctcgcgcc	9720
cggctcgacg gatgcggcga tcgtcaacgc ggaggggcag ataacggcgg tccatctcgc	9780
cggggcgggg aatatgttca gcctgttgat taaaaccgag ctgggcctcg aggatctttc	9840
gctggcgga gcgataaaaa aatacccgct ggccaaagt gaaagcctgt tcagtattcg	9900
tcacgagaat ggcgcggttg agttctttcg ggaagccctc agcccggcgg tgttcgcca	9960
agtgggtgac atcaaggagg gcgaactgt gccgatgat aacgccagcc cgctggaaaa	10020
aattcgtctc gtgcgccgc aggcgaaaga gaaagtgttt gtcaccaact gcctgcgcgc	10080
gctgcgccag gtctcaccgc gcggttccat tcgcatatc gcctttgtgg tgctgggtgg	10140

cggtcatcg ctggactttg agatcccga gcttatcacg gaagccttgt cgcactatgg	10200
cgtggtcgcc gggcaggga atattcgggg aacagaaggg ccgcgcaatg cggtcgccac	10260
cgggctgcta ctggccggtc aggcgaatta aacgggcgct cgcgccagcc tctaggtaca	10320
aataaaaaag gcacgtcaga tgacgtgcct tttttcttgt ctacgtgca ccaatgcttc	10380
tggcgtcagg cagccatcgg aagctgtggt atggctgtgc aggtcgtaaa tcaactgcata	10440
attcgtgtcg ctcaaggcgc actcccgttc tggataatgt tttttgcgcc gacatcataa	10500
cggttcttggc aaatattctg aaatgagctg ttgacaatta atcatccggc tcgtataatg	10560
tgtggaattg tgagcggata acaatttcac acaggaaaca gaccatgact agtaaggagg	10620
acaattccat ggctgtctgt gctgatagat taaacttaac ttccggccac ttgaatgctg	10680
gtagaaagag aagttcctct tctgtttctt tgaaggctgc cgaaaagcct ttcaaggtta	10740
ctgtgattgg atctggtaac tggggtacta ctattgccaa ggtggttgcc gaaaattgta	10800
agggatacc agaagttttc gctccaatag tacaatgtg ggtgttcgaa gaagagatca	10860
atggtgaaaa attgactgaa atcataaata ctagacatca aaacgtgaaa tacttgccctg	10920
gcacactct acccgacaat ttggttgcta atccagactt gattgattca gtcaaggatg	10980
tcgacatcat cgttttcaac attccacatc aatttttgcc ccgtatctgt agccaattga	11040
aaggatcatg tgattcacac gtcagagcta tctcctgtct aaagggtttt gaagtgggtg	11100
ctaaagggtg ccaattgcta tcctcttaca tcaactgagga actaggtatt caatgtgggtg	11160
ctctatctgg tgctaacatt gccaccgaag tcgctcaaga acactggtct gaaacaacag	11220
ttgcttacca cattccaaag gatttcagag gcgagggcaa ggacgtcgac cataaggttc	11280
taaaggcctt gtccacaga cttacttcc acgttagtgt catcgaagat gttgctggta	11340
tctccatctg tgggtgcttg aagaacgttg ttgccttagg ttgtggtttc gtcgaaggtc	11400
taggtcgggg taacaacgtt tctgtgtcca tccaaagagt cggtttgggt gagatcatca	11460
gattcggta aatgtttttc ccagaatcta gagaagaaac atactaccaa gactctgctg	11520
gtgttgctga ttgatcacc acctgcgtg gtggtagaaa cgtcaagggt gctaggctaa	11580
tggctacttc tggttaaggac gcctgggaat gtgaaaagga gttgttgaat ggccaatccg	11640
ctcaaggttt aattacctgc aaagaagttc acgaatggtt ggaacatgt ggctctgtcg	11700
aagacttccc attatttgaa gccgtatacc aaatcgttta caacaactac ccaatgaaga	11760
acctgccgga catgattgaa gaattagatc tacatgaaga ttagatttat tggatccagg	11820
aaacagacta gaattatggg attgactact aaacctctat ctttgaaagt taacgccgt	11880
ttgttcgacg tcgacggta cattatcatc tctcaaccag ccattgctgc attctggagg	11940
gatttcggta aggacaaacc ttatttcgat gctgaacacg ttatccaagt ctgcatggt	12000

tggagaacgt ttgatgcat tgctaagttc gctccagact ttgccaatga agagtatgtt	12060
aacaaattag aagctgaaat tccggtcaag tacggtgaaa aatccattga agtcccaggt	12120
gcagttaage tgtgcaacgc tttgaacgct ctacccaaag agaaatgggc tgtggcaact	12180
tccggtaccc gtgatatggc acaaaaatgg ttcgagcatc tgggaatcag gagaccaaaag	12240
tacttcatta ccgctaata tgtcaaacag ggtaagcctc atccagaacc atatctgaag	12300
ggcaggaatg gcttaggata tccgatcaat gagcaagacc cttccaaatc taaggtagta	12360
glatttgaag acgctccagc aggtattgcc gccggaaaag ccgccggttg taagatcatt	12420
ggtattgcca ctactttcga cttggacttc ctaaaggaaa aaggctgtga catcattgtc	12480
aaaaaccacg aatccatcag agttggcggc tacaatgccg aaacagacga agttgaattc	12540
atttttgacg actacttata tgctaaggac gatctgttga aatggtaacc cgggctgcag	12600
gcatgcaagc ttggctgttt tggcggatga gagaagatit tcagcctgat acagattaaa	12660
tcagaacgca gaagcggctc gataaaacag aatttgcctg gcggcagtag cgcggtggtc	12720
ccacctgacc ccatgccgaa ctcagaagtg aaacgccgta gcgccgatgg tagtgtgggg	12780
tctcccatg cgagagtagg gaactgccag gcatcaaata aaacgaaagg ctcagtcgaa	12840
agactgggcc ttctgtttta tctgttgttt gtcggtgaac gctctcctga gtaggacaaa	12900
tccgccggga gcggatttga acgttgcgaa gcaacggccc ggagggtggc gggcaggacg	12960
cccgccataa actgccaggc atcaaattaa gcagaaggcc atcctgacgg atggcctttt	13020
tgcgtttcta caaactccag ctggatcggg cgctagagta tacatttaaa tggtagcctc	13080
tagtcaaggc ctttaagttag tcgtattacg gactggccgt cgttttacia cgtcgtgact	13140
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct	13200
ggcgtaatag cgaagaggcc cgcaccgatc gcccttccca acagttgcgc agcctgaatg	13260
gcgaatggcg cctgatgcgg tattttctcc ttacgcatct gtgcggtatt tcacaccgca	13320
tatggtgcac tctcagtaca atctgctctg atgccgata gtttaagccag ccccgacacc	13380
cgccaacacc cgctgacgag ct	13402
<210> 66	
<211> 14443	
<212> DNA	
<213> Artificial Sequence	
<220><223> Plasmid	
<400> 66	
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgtagga	60

attaattccc ctgctcgcgc aggcctgggtg ccaagctctc gggtaacatc aaggcccgat	120
ccttgaggcc ctctttacag agatgaaaaa caaacccgca cgccaggcgg catcgcggtc	180
tcagagatat gtttacgtag atcgaagagc accggtgttt aaacgccctt gacgatgcca	240
catcctgagc aaataattca accactaaac aaatcaaccg cgtttcccgg aggtaacga	300
gctcatgac ctgtgtgtg gtgaagccct gatcgacatg ctgccccggc agacgacgt	360
gggtgaggcg ggctttgccc cttacgcagg cggagcggtc ttcaacacgg caattgcgt	420
ggggcgctct ggcgccctt cagccttttt taccggtctt tccgacgaca tgatggcgca	480
tatctgcgg gagaccctgc gggccagcaa ggtggatttc agctattgcg ccaccctgtc	540
gcgccccacc accattgctg tcgttaagct ggttgatggc catgcgacct acgcttttta	600
cgacgagaac accgccggcc ggatgatcac cgaggccgaa cttccggcct tgggagcgga	660
ttgcgaagcg ctgcatttcg gcgccatcag ccttattccc gaacctgcg gcagcaccta	720
tgaggcgctg atgacgcgcg agcatgagac ccgcgtcatc tcgctcgatc cgaacattcg	780
tcccggcttc atccagaaca agcagtcgca catggcccg c atccgccga tggcgcgat	840
gtctgacatc gtcaagttct cggatgagga cctggcgtgg ttcggtctgg aaggcgacga	900
ggacacgctt gcccgccact ggctgcacca cggtgcaaaa ctgctcgttg tcaccctgg	960
cgccaaggtt gccgtgggtt acagcgccaa tctcaagtg gaagtggcct ccgagcgct	1020
cgaagtggc gatacggctg gcgccggcga tacgttcgat gccggcattc ttgcttcgct	1080
gaaaatgcag ggctgtctga ccaaagcgca ggtggcttcg ctgagcgaag agcagatcag	1140
aaaagctttg gcgtttggcg cgaaagccgc tgcggtcact gtctcgcggg ctggcgcaaa	1200
tccgcctttc gcgcatgaaa tcggtttgtg attaatataa gcacgcagtc aaacaaaaaa	1260
cccgcgccat tgcgcgggtt tttttatgcc cgaaggcgcg ccagcacgca gtcaaacaaa	1320
aaaccgcgc cattgcgcgg gtttttttat gccgaacgg ccgaggtctt ccgatctct	1380
gaagccaggg cagatccgtg cacagcacct tgccgtagaa gaacagcaag gccgccaatg	1440
cctgacgatg cgtggagacc gaaaccttgc gctcgttcgc cagccaggac agaaatgcct	1500
cgacttcgct gctgcccaag gttgccgggt gacgcacacc gtggaaacgg atgaaggcac	1560
gaaccagtg gacataagcc tgttcggttc gtaagctgta atgcaagtag cgtatgcgt	1620
cacgcaactg gtccagaacc ttgaccgaac gcagcgggtg taacggcgca gtggcggttt	1680
tcatggcttg ttatgactgt tttttgggg tacagtctat gcctcgggca tccaagcagc	1740
aagcgcgtta cgccgtgggt cgatgtttga tgttatggag cagcaacgat gttacgcagc	1800
agggcagtcg ccctaaaaca aagttaaaca tcatgaggga agcggtagc gccgaagtat	1860

cgactcaact atcagaggta gttggcgtca tcgagcgcca tctcgaaccg acgttgctgg	1920
ccgtacattt gtacggctcc gcagtggatg gcggcctgaa gccacacagt gatattgatt	1980
tgctggttac ggtgaccgta aggcttgatg aaacaacgcg gcgagctttg atcaacgacc	2040
ttttggaaac ttcggcttcc cctggagaga gcgagattct ccgcgctgta gaagtcacca	2100
ttgttggtca cgacgacatc attccgtggc gttatccagc taagcgcgaa ctgcaatttg	2160
gagaatggca gcgaatgac attcttgacg gtatcttcga gccagccacg atcgacattg	2220
atctggctat ctigtgaca aaagcaagag aacatagcgt tgccttggtg ggtccagcgg	2280
cggaggaact ctttgatccg gttcctgaac aggatctatt tgaggcgcta aatgaaacct	2340
taacgctatg gaactcgccg cccgactggg ctggcgatga gcgaaatgta gtgcttacgt	2400
tgtccgcat ttgttacagc gcagtaaccg gcaaaatcgc gccgaaggat gtcgctgccg	2460
actgggcaat ggagcgctg ccggcccagt atcagcccgt catacttgaa gctagacagg	2520
cttatcttgg acaagaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg	2580
tccactacgt gaaaggcgag atcaccaagg tagtcggcaa ataatgtcta acaattcggt	2640
caagccgacg ccgcttcgcg gcgcggctta actcaagcgt tagatgcaact aagcacataa	2700
ttgctcacag ccaaactatc aggtcaagtc tgcttttatt attttaagc gtgcataata	2760
agccctacac aaattgggag atatatcatg aaaggctggc tttttcttgt tatcgcaata	2820
gttggcgaag taatcgcaac atccgcatta aaatctagcg agggctttac taagctcgtc	2880
agcgggtgtt ggcgggtgtc ggggctggct taactatgcg gcatcagagc agattgtact	2940
gagagtgcac catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat	3000
caggcgccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc	3060
ttcgctatta cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac	3120
gccagggttt tccagtcac gacgttgtaa aacgacggcc agtccgtaat acgactcact	3180
taaggccttg actagagggt accatttaaa tgtatactct agcgcccgat ccagctggag	3240
ttttagaaaa cgcaaaaagg ccatccgtca ggatggcctt ctgcttaatt tgatgcctgg	3300
cagtttatgg cgggcgtcct gcccgcacc ctccgggccg ttgcttcgca acgttcaaat	3360
ccgtccccg cggtattgtc ctactcagga gagcgttcac cgacaaaca cagataaaac	3420
gaaaggccca gtctttcgac tgagcctttc gttttatttg atgcctggca gttccctact	3480
ctcgcatggg gagacccac actaccatcg gcgctacggc gtttcacttc tgagtccggc	3540
atggggtcag gtgggaccac cgcgctactg ccgccaggca aattctgttt tatcagaccg	3600

cttctgcgtt ctgatttaat ctgtatcagg ctgaaaatct tctctcatcc gccaaaacag	3660
ccaagcttgc atgcctgcag cccgggttac catttcaaca gatcgtcctt agcatataag	3720
tagtcgtcaa aaatgaattc aacttcgtct gtttcggcat tgtagccgcc aactctgatg	3780
gattcgtggt ttttgacaat gatgtcacag cctttttcct ttaggaagtc caagtcgaaa	3840
gtagtggcaa taccaatgat cttacaaccg gcggcttttc cggcggcaat acctgctgga	3900
gcgtcttcaa atactactac cttagatttg gaagggtctt gctcattgat cggatacct	3960
aagccatttc tgccttcag atatggttct ggatgaggct taccctgttt gacatcatta	4020
gcggtaatga agtactttgg tctctgatt ccagatgct cgaaccattt ttgtgccata	4080
tcacgggtac cggaagtgc cacagcccat ttctcttttg gtagagcgtt caaagcgttg	4140
cacagcttaa ctgcacctgg gacttcaatg gatttttcac cgtacttgac cggaatttca	4200
gcttctaatt tgtaacata ctcttcattg gcaaagtctg gagcgaactt agcaatggca	4260
tcaaacgttc tccaacctg cgagacttgg ataacgtgtt cagcatcgaa ataaggtttg	4320
tccttaccga aatccctcca gaatgcagca atggctggtt gagagatgat aatggtaccg	4380
tcgacgtcga acaaagcggc gttaactttc aaagatagag gtttagtagt caatcccata	4440
attctagtct gtttctgga tccaataaat ctaatcttca tgtagatcta attcttcaat	4500
catgtccggc aggttcttca ttgggtagtt gttgtaaagc atttggata cggttcaaa	4560
taatgggaag tcttcgacag agccacatgt ttccaacct tcgtgaactt ctttgcaggt	4620
aattaaacct tgagcggatt ggccattcaa caactccttt tcacattccc aggcgtcctt	4680
accagaagta gccattagcc tagcaacctt gacgtttcta ccaccagcgc aggtggtgat	4740
caaatcagca acaccagcag actcttggta gtatgtttct tctctagatt ctgggaaaaa	4800
catttgaccg aatctgatga tctacccaa accgactctt tggatggcag cagaagcgtt	4860
gttaccag cctagacctt cgacgaaacc acaacctaag gcaacaacgt tcttcaaagc	4920
accacagatg gagataccag caacatcttc gatgacacta acgtggaagt aaggtctgtg	4980
gaacaaggcc tttagaacct tatggtcgac gtccttgccc tcgcctctga aatcctttgg	5040
aatgtggtaa gcaactgttg ttccagacca gtgttcttga gcgacttcgg tggcaatgtt	5100
agcaccagat agagcaccac attgaatacc tagttcttca gtgatgtaag aggatagcaa	5160
ttggacacct ttagcaccaa cttcaaaacc ctttagacag gagatagctc tgacgtgtga	5220
atcaacatga ctttcaatt ggctacagat acggggcaaa aattgatgtg gaatgttgaa	5280
aacgatgatg tcgacatcct tgactgaatc aatcaagtct ggattagcaa ccaaattgtc	5340
gggtagagt atgccaggca agtatttcac gttttgatgt ctagtattta tgatttcagt	5400
caatttttca ccattgatct cttcttcgaa caccacatt tgtactattg gagcgaaaac	5460

ttctgggtat cccttacaat ttctggcaac caccttggca atagtagtac cccagttacc 5520
 agatccaatc acagtaacct tgaaaggctt ttccggcagcc ttcaaagaaa cagaagagga 5580
 acttctcttt ctaccagcat tcaagtggcc ggaagttaag tttaatctat cagcagcagc 5640
 agccatggaa ttgtcctcct tactagtcac ggtctgtttc ctgtgtgaaa ttgttatccg 5700

 ctacacaattc cacacattat acgagccgga tgattaattg tcaacagctc atttcagaat 5760
 atttgccaga accgttatga tgtcggcgca aaaaacatta tccagaacgg gagtgcgcct 5820
 tgagcgacac gaattatgca gtgatttacg acctgcacag ccataccaca gcttccgatg 5880
 gctgcctgac gccagaagca ttggtgcacg ctagacaaga aaaaaggcac gtcatctgac 5940
 gtgccttttt tatttgtacc tagaggctgg cgcgagcgcc cgtttaattc gcctgaccgg 6000
 ccagtagcag cccggtggcg accgcattgc gcggcccttc tgttccccga atattgcctt 6060
 gcccggcgac cagcccatag tgcgacaagg ctcccgatgat aagctgcggg atctcaaagt 6120

 ccagcgatga gccgcccacc agcaccacaa aggcgatata gcgaatggaa ccgcccgggtg 6180
 agacctggcg cagcgcgcgc aggcagttgg tgacaaacac ttctctttc gcctgccggc 6240
 gcacgagacg aattttttcc agcgggctgg cgttatcgat cggcaccagt tcgcccctct 6300
 tgatgtacac cactttggcg aacaccgcgc ggctgagggc ttcccgaag aactccaccg 6360
 cgccattctc gtgacgaata ctgaacaggc ttccacttt ggccagcggg tattttttta 6420
 tcgcttccgc cagcgaaaga tcctcgaggc ccagctcggt tttaatcaac aggctgacca 6480
 tattccccgc cccggcgaga tggaccgcgc ttatctgccc ctcccgcttg acgatcgccg 6540

 catccgtcga gccggcgccg aggtcgagga tcgccagcgg cgcgcacag ccgggagtgg 6600
 ttaacgcccc ggcatggcc atgttggcct ccacgcgcgc caccaccacc tcggtctgca 6660
 gtcggcgctc cagttcgcgg gcgataacct gcatttgcag acgatccgct ttcaccatcg 6720
 ccgccatccc gacggcattc tccatggcgc actcgccggc catcccgcgc tgcacettgc 6780
 gcggaataaa cgtatccacc gccagcagat cctggatgta tatcgcgctc atctcatggc 6840
 cggtcaggga cgccattacc ttgcgcaccc gctcaagcat gccgccggcg tgggtgcccc 6900
 gttcgccgcg gatgtcgct accggagcgc aggcgctcat cgctgcatg atggcttccg 6960

 cgccctcggc gacatcgcc tctcccggc gcttttcgcc gctaatgtag aggttgcccc 7020
 ccgggatcac ccgcgactgc acatccccct gcgggggtctt gagcaccacc gcggaacggt 7080
 tgccaatcag ggcgcgggcg atggggacga tggcctgggt ctcttcggg cttagcccga 7140
 agaaggtggc gatcccgtag ggattcgaca ggatccgcac cacctggccc ggcgcggcca 7200
 ctccaccgc cgccattacc ccctcgggga cctgctccag cagcgtcact tcatccacca 7260
 ccggcagggt ttacgcagg cggttgttca ccagcacgcc gtcgtccttt ttgaggatcg 7320

ccgccaccac gttgatcccc cggtcgagcg cctcattgag ccaccacacg gcgtcaagga 7380

aatcgacggc gtcgtcaatc agtacgatcc acccctcggc atactgcgcc gccggcagcg 7440

tcgccagccg cccgagggcg atagtcgtcc ccacgccaac gccaccccg cccggcgtct 7500

gcgggttatg accgatcatg gtcgattcgg tgataatggt ctcggtgatg gtctccatcg 7560

ccacatcgcc aatcaccggc gcggcttcgt taagatagat gcgagagaca tcgtcatcg 7620

accacgggtg tttcgccagg gcctgctcca gcgcggcgag ggtcccgcg atattgtccc 7680

gcgtcccttt catgcccgtc gtcgcgacga tcccgtggc aacaaacgcc ctgcctgcg 7740

ggtagtcgga cgccagcgcc acctcggtgg tggcgttgcc gatatcaatc ccggctatta 7800

acggcatgct gacctccgt tagcttcctt tacgcagctt atgccgtgc tgatacactt 7860

ccgccgactc ccggacaaag gcggcattca ctgtcgcgt ccaggtgtgc tccagctcgt 7920

cggcgatcgc cagcagctcc gcctgcgagg agcggaacgg gcgcagcgcg ttatagatag 7980

ccagaatgcg ctgctcagga atggcgataa gctccgccg gcggcggaaa ttgcgcgcca 8040

ccgcatggcg ctgcatctgc tcggcaatct gcgcctggta ctcaagggtc tggcgggaga 8100

tccgcacatc ctgcgggccc acctcgccag agagcacctt ctcgagggtg atatcggtca 8160

atggtttgcc ggtaggcgtc aggatatgct ccgggcagcg ggtggctaac ggataatcct 8220

gcacgcgcat ggttttctcg ctcatggtca ctcccttact aagtcgatgt gcagggtgac 8280

gggctcggcg tctgcacca catgtttggt ctctttgata tgaaatagcg cggctttggc 8340

cataaatttc ggccgcacca tctgatcgtt caccaccggc accggcgaag gtgactcttt 8400

gcgcgcatag cgcgacggt ttttccaat ctgccgtag gtctccagcg tcagcagcgg 8460

gcctgggag aacagctcca ggttgctgag cggcagcaga tcgcgctgat ggatgaccgt 8520

ggtcccttc gactggatc cgatgccgat ccccgagccg ctcaggttgg ccgcatccca 8580

ggccataaag gagacgtcgg acgtgcgcag aatgcgcacc acccgggcgt gaagccctc 8640

ttcttcacc ccggcaatca gctctttgag gatcgcgcca tggggcatat cgatcagagt 8700

gtgatgctgg tgtttatcga aggcagggcc gacgccgatc accacttcat cggcgcgttc 8760

atcggcagaa gctacccgc cctcggggt tttcagggtg aaagagggt gaatttgggt 8820

tgtctgttc acaggaatac cgccttggtc aatggtgtcg ggctgaacca cccccgaat 8880

atttttgac tccgccagc gttcggcaga gatgcgatag ccggtgcccg gccctgata 8940

gtcattgatg tcgttgaccg cactcaccac ctgcaactgc cgatcgaaaa tggccgaggt 9000

ctgcaggtaa tcgccgtga cccgtggcg cagcatattg agaatttgc tggcgatac 9060

ctcaaagccg ctgcggctca gcgcgccgac aatatcgagg ccggtgatgt tgcgcttcat	9120
catctcttcc accgcactca gatcctccac cacgttacgc ggcgcatct cgttgctgcc	9180
gtgcgcgtag gtggcgccct ccacctctc gtcggcgatt ggcggcagcc ccagctcgcg	9240
gaaaaccgcc tggatcgccc gcgccgcttt ctggcgaaatg gcaatggttt cgccttcggt	9300
caccggacgc aggccgccgt caaccatcag gtcacgtgc aggatgttgt aatcatcaaa	9360
atcttccgca tcgaagtctg agccggcgaa catgttgtcg tagttcgga ccgctgtga	9420
gccggagaaa ataaagtcgg tgcccggcag catctgcatc aggggtgcgcg cgggtcggcg	9480
aatatccgag tgggagaaag tctggtcgtt ggcgacgcc acttcgaggt cgagcataga	9540
ggcgatcagg ttttccgcca gcaccgccc aatgccgcac ggacacagcg cggatcatgcc	9600
gatacagctc accgcgccgt ttgacgtcc ctgaacccg gcgcctttag taatgaagat	9660
gcagcgcat tcgaggtaga gcatcgactt gctctccgaa tagcccatca gcgttcgga	9720
tccggtgccg gaggtgtagc gcattttcaa cccgcgggag gcgtaggccg aggcgaggaa	9780
gcctttgac caccgctat catcgccgtc ggtaaatacc gcttcggtgc cgtagaccga	9840
caccgtctcg gcgtagctgg ttaagccacg catgccagc tccagctcgg tggcctcttc	9900
caccgagcac tgcgtcaaca cgccggggcg gccgcactgc gaaccgacca acagcgccag	9960
ggcgttaaac ggcgcgtagc gcgcgatacc gaccgtggtc tcctgttctg agaagccgcg	10020
gatcccgccc tcggcgccgt cagcggcaat ctgcaccgga ttatctttga gattggtgac	10080
gtggcactgg ttggaggggg tccggcgggc acgcatcttc tgcagcgcca tcatcatctc	10140
caccacgttc atctgcgcca tcacctgac cgttttggcc ggctgatgg cggtagtgt	10200
ggcaatgatc tcctcccgcc tgacgtgaat atccaccagc atacgggcta tttccaccgc	10260
ctccaggcgc attgcctgct ctgtgcgtc aacgttgatc gcgtaatcgg cgataaatcg	10320
gtcgatcatg tcaaactggt cccggcgctt gccgtccagt tcgacgatca gaccgtgtc	10380
cacttttact gaagagaccg ggtcaaagg gctgtccatg gcgatcagcc cctcttcagg	10440
ccactcgcca atcagcccgt cctgattgac ggggcgctgg gccagtactg caaatcgctt	10500
tgatcttttc attgttcatc ggctcaaaag gtgaagcttg gttacctcg ggaaacgcgg	10560
ttgatttgtt tagtggttga attatttgc caggatgtgg cattgtcaag ggcgtgacgg	10620
ctgcctgac ttctcgttcc agtgccccg tccgacagtc gagcgtgcga gcccataatc	10680
tcgcgtggt gctgcatacc gtggcaaaca gcacagatcg cctaggaaaa aaaaagcccc	10740
cactgtcagg tgcgggcttt tttctgtgtt tgctaggcca gttcaagcgc aagcatcagg	10800
gtgcagctgg gcagaggcga gattctctcc cgggatcacg aactgtttta acgggccgct	10860
ctcgccata ttgcggtcga taagccgtc cagggcggtg atctctctt cgccgatcgt	10920

ctggctcagg cgggtcaggc cccgcgcac gctggccagt tcagcccca gcacgaacag	10980
cgtctgctga atatggtgca ggctttcccg cagcccgcg tcgcggtcg tggcgtagca	11040
gacgcccagc tgggataca gttcatcgac ggtgccgtag gcctcgacgc gaatatggtc	11100
tttctcgatg cggctgccgc cgtacagggc ggtggtgcct ttatccccgg tgcgggtata	11160
gatacgtac attcagtttc tctcacttaa cggcaggact ttaaccagct gcccggcgtt	11220
ggcgccgagc gtacgcagtt gatcgctcgt atcggtgacg tgtccggtag ccagcggcgc	11280
gtccgccggc agctgggcat gaggtagggc tatctcgccg gacgcgtga gcccataacc	11340
caccgcagg ggcgagcttc tggccgccag ggcccccagc gcagcggtc caccgcctcc	11400
gtcatagggt atggtctggc aggggacccc ctgctcctcc agccccagc acagtcatt	11460
gatggcgccg gcatggtgcc cgcgcggtac gtaaacagg cgtacgcctg gcggtgaaag	11520
cgacatgacg gtcccctcgt taacactcag aatgcctggc ggaacatacg atagtcata	11580
atataccttc tcgcttcagg ttataatgcg gaaaaacaat ccaggcgca ctgggcta	11640
aattgacct gctcgacct accgccgta acgccgacgg cgccaattac ctgctcatta	11700
aaaataactg gcagccgcc gccaaaaata ataattcgt gttggttgg tagctgcaga	11760
ccgtacagag attgtcctg ctggaccgt gacgtaattt catgggtacc ttgcttcagg	11820
ctgcaggcgc tccaggcttt attcaggga atatcgacg tggagacga ggcctcgcc	11880
atccgtgga taagcagcgt gttgcctccg cgtcaacta cggaaaacac caccgccag	11940
ttgatctcag tggctttttt ttccaccgcc gccgccattt gctgggcggc ggccagggtg	12000
attgtctgaa ctgtttggct ctgtttcgc attctctccc gcaagcttg ttacctccg	12060
gaaacgcgtg tgatttgttt agtgggtgaa ttatttgctc aggatgtggc attgtcaagg	12120
gcgtgacggc tcgctgact tctcgttcca gtcccccg cgcacgtcg agcgtgcgag	12180
cccataatct cgcgtggtg ctgcataacc tggcaaacag cacagatcg ctagcagtea	12240
aaagcctccg gtcggaggct ttgactatt taaatgaatt cccgacagta agacgggtaa	12300
gcctgttgat gataccgtg ccttactggg tgcattagcc agtctgaatg acctgtcacg	12360
ggataatccg aagtgtcag actggaaaat cagagggcag gaactgctga acagcaaaaa	12420
gtcagatagc accacatagc agaccgcca taaaacgcc tgagaagccc gtgacgggt	12480
tttctgtat tatgggtagt ttcttgcat gaatccataa aaggcgctg tagtgccatt	12540
taccccatc cactgccaga gccgtgagcg cagcgaactg aatgtcacga aaaagacagc	12600
gactcagggt cctgatggc ggagacaaaa ggaatattca gcgatttgc cgagcttgcg	12660
agggtgctac ttaagccttt aggggtttta ggtctgttt gtagaggagc aaacagcgtt	12720
tgcgacatcc ttttgaata ctgcggaact gactaaagta gtgagttata cacagggtg	12780

ggatctattc tttttatctt tttttattct ttctttattc tataaattat aaccacttga 12840

atataaaca aaaaaacaca caaaggtcta gcggaattta cagagggtct agcagaattt 12900

acaagttttc cagcaaaggt ctagcagaat ttacagatac ccacaactca aaggaaaagg 12960

actagtaatt atcattgact agcccatctc aattggtata gtgattaaaa tcacctagac 13020

caattgagat gtatgtctga attagtgttt ttcaaagcaa atgaactagc gattagtgcg 13080

tatgacttaa cggagcatga aaccaagcta attttatgct gtgtggcact actcaacccc 13140

acgattgaaa accctacaag gaaagaacgg acggtatcgt tcacttataa ccaatacgt 13200

cagatgatga acatcagtag ggaaaatgct tatggtgtat tagctaaagc aaccagagag 13260

ctgatgacga gaactgtgga aatcaggaat cctttggtta aaggctttga gattttccag 13320

tggacaaact atgccaagtt ctcaagcgaa aaattagaat tagtttttag tgaagagata 13380

ttgccttate ttttccagtt aaaaaaattc ataaaatata atctggaaca tgttaagtct 13440

tttgaaaaca aatctctat gaggatttat gagtgggtat taaaagaact aacacaaaag 13500

aaaactcaca aggcaaatat agagattagc cttgatgaat ttaagttcat gttaatgctt 13560

gaaaataact accatgagtt taaaaggctt aaccaatggg ttttgaaacc aataagtaaa 13620

gatttaaaca cttacagcaa tatgaaattg gtggttgata agcgaggccg cccgactgat 13680

acgttgattt tccaagtga actagataga caaatggatc tcgtaaccga acttgagaac 13740

aaccagataa aaatgaatgg tgacaaaata ccaacaacca ttacatcaga ttcctaccta 13800

cataacggac taagaaaaac actacacgat gctttaactg caaaaattca gtcaccagt 13860

tttgaggcaa aatttttgag tgacatgcaa agtaagtatg atctcaatgg ttcgttctca 13920

tggtcacgc aaaaacaacg aaccacacta gagaacatac tggctaaata cggaaggatc 13980

tgaggttctt atggctcttg tatctatcag tgaagcatca agactaaca acaaaagtag 14040

aacaactgtt caccgttaca tatcaaaggg aaaactgtcc atatgcacag atgaaaacgg 14100

tgtaaaaaag atagatacat cagagctttt acgagttttt ggtgcattca aagctgttca 14160

ccatgaacag atcgacaatg taacagatga acagcatgta acacctaata gaacaggtga 14220

aaccagtaaa acaaagcaac tagaacatga aattgaacac ctgagacaac ttgttacagc 14280

tcaacagtca cacatagaca gcctgaaaca ggcgatgctg cttatcgaat caaagctgcc 14340

gacaacacgg gagccagtga cgctcccgt ggggaaaaaa tcatggcaat tctggaagaa 14400

atagcgcttt cagccggcaa accggctgaa gccggatctg cga 14443

<210> 67

<211> 1260

<212> DNA

<213> Citrobacter sp

<400> 67

atgaagataa atatgccgtt cagtaatgac aaatatcggt actcgccgg gtatctgctt 60
 tttttctttg ccgcctggtc gttgtgggtg tctttttatg cgatatggct aaaaaataaa 120
 cttggcctgt ccggaacgga gctgggaatg ctgtatgccg taaaccagtt ctttagcatg 180
 ctgtttatgc tggctctacgg ttttctgcag gataagctcg gcacccgtaa acaccttate 240
 tggctgatgg ggatagtcac cagctcagc ggcccgttcc tgatttatgt ttacgaaccg 300
 ctgctgacct ccaacttcaa acttggatg gcgctgggag ccattttctt tggccttggc 360
 tacctcgagg gttgtggctt ggtagaaagc ttcgtcgaaa aagtgagccg caaatcaac 420

tttgaattcg gcaccgccc cttgtgggga tcgcttggct acgccgcagg gacatttgtt 480
 ggcggtatct tcttcagcat caaccacac attaaacttct ggtgcgtatc ggtaatgggg 540
 gtgttattcc tgttgattaa cgtgtgttc aaaaccaact caccgcccc atcttctgta 600
 aaaacgcgtt ctctgaacc tgacgcgtg acccgaaagg attttctcac tatctttaa 660
 gatacgagc tctggttttt cgttatcttt gtcgtcggta cctggctggt ctatagcatc 720
 tacgatcagc agatgttccc ggtgttttac gccagcttat ttgacgatcc cgaactggca 780
 ccacgcgtat acggctacct caactcggtg caggtcttta tggaagccgt cggatatggc 840

ctggttccat tcttgattaa ccgcatcggg cctaaatccg cattgctgct gggtagcaca 900
 atcatggcct gtcgaatcct gggttcagca ctgttcaccg atatctatat tatctccttg 960
 attaaaatgc ttcatgcgtt ggaagtccca ctgtttgtta tttcagtgtt taaattcagc 1020
 gtagcgaatt ttgataaacg cctgtcatca acgatatatc tcattggctt caatatcgcc 1080
 agttccattg gcattatcgt gctgtcactg cctgtcggta agttgtttga taaagtgggc 1140
 tatcaggaaa tcttctgat tatggccagc attgtgataa taacactaat atttgctat 1200
 ttctcgttga gcaaaaagca tcatcagcag aagatgggaa atgaactggt gacagagtag 1260

<210> 68

<211> 419

<212> PRT

<213> Citrobacter sp

<400> 68

Met Lys Ile Asn Met Pro Phe Ser Asn Asp Lys Tyr Arg Tyr Ser Ser

1 5 10 15

Gly Tyr Leu Leu Phe Phe Phe Ala Ala Trp Ser Leu Trp Trp Ser Phe

20 25 30
 Tyr Ala Ile Trp Leu Lys Asn Lys Leu Gly Leu Ser Gly Thr Glu Leu
 35 40 45
 Gly Met Leu Tyr Ala Val Asn Gln Phe Phe Ser Met Leu Phe Met Leu
 50 55 60
 Val Tyr Gly Phe Leu Gln Asp Lys Leu Gly Thr Arg Lys His Leu Ile
 65 70 75 80
 Trp Leu Met Gly Ile Val Ile Thr Leu Ser Gly Pro Phe Leu Ile Tyr
 85 90 95
 Val Tyr Glu Pro Leu Leu Thr Ser Asn Phe Lys Leu Gly Met Ala Leu
 100 105 110
 Gly Ala Ile Phe Phe Gly Leu Gly Tyr Leu Ala Gly Cys Gly Leu Val
 115 120 125
 Glu Ser Phe Val Glu Lys Val Ser Arg Lys Phe Asn Phe Glu Phe Gly
 130 135 140
 Thr Ala Arg Leu Trp Gly Ser Leu Gly Tyr Ala Ala Gly Thr Phe Val
 145 150 155 160
 Gly Gly Ile Phe Phe Ser Ile Asn Pro His Ile Asn Phe Trp Cys Val
 165 170 175
 Ser Val Met Gly Val Leu Phe Leu Leu Ile Asn Val Leu Phe Lys Thr
 180 185 190
 Asn Ser Pro Ala Pro Ser Ser Val Lys Thr Arg Ser Pro Glu Pro Asp
 195 200 205
 Ala Leu Thr Arg Lys Asp Phe Leu Thr Ile Phe Lys Asp Thr Gln Phe
 210 215 220
 Trp Phe Phe Val Ile Phe Val Val Gly Thr Trp Ser Phe Tyr Ser Ile
 225 230 235 240
 Tyr Asp Gln Gln Met Phe Pro Val Phe Tyr Ala Ser Leu Phe Asp Asp
 245 250 255
 Pro Glu Leu Ala Pro Arg Val Tyr Gly Tyr Leu Asn Ser Val Gln Val
 260 265 270

Phe Met Glu Ala Val Gly Met Ala Leu Val Pro Phe Leu Ile Asn Arg
275 280 285

Ile Gly Pro Lys Ser Ala Leu Leu Leu Gly Gly Thr Ile Met Ala Cys
290 295 300

Arg Ile Leu Gly Ser Ala Leu Phe Thr Asp Ile Tyr Ile Ile Ser Leu

305 310 315 320
Ile Lys Met Leu His Ala Leu Glu Val Pro Leu Phe Val Ile Ser Val
325 330 335

Phe Lys Phe Ser Val Ala Asn Phe Asp Lys Arg Leu Ser Ser Thr Ile
340 345 350

Tyr Leu Ile Gly Phe Asn Ile Ala Ser Ser Ile Gly Ile Ile Val Leu
355 360 365

Ser Leu Pro Val Gly Lys Leu Phe Asp Lys Val Gly Tyr Gln Glu Ile

370 375 380
Phe Leu Ile Met Ala Ser Ile Val Ile Ile Thr Leu Ile Phe Gly Tyr
385 390 395 400

Phe Ser Leu Ser Lys Lys His His Gln Gln Lys Met Gly Asn Glu Leu
405 410 415

Val Thr Glu

<210> 69

<211> 1314

<212> DNA

<213> Enterococcus faecium

<400> 69

atgaaagggg atacaaatat atcgttggag gataaaaata tgtcaaaagt taacgtattt 60

aaaaatcaat cttatttaca aagttcagct acattattac tattttttgc ttcttggggt 120

gtttggtggt cattttttca actttggcta acatctgaat caaatggttt agggttatct 180

ggcagtgtctg taggaacagt attctcggca aattcgtag ttaccttaat ttgatgttt 240

atatttgaa cattacaaga taaattgtat attaaacgaa atttattaat ttttgcttct 300

gtattagcga cacttggtgg accatttttt atatggatat atgggccatt gctagataac 360

aattttaatt taggcattat tatgggagcg ctatttttgt cagctggata tttagcttct 420

gtaggagttt ttgaagctgt gtcagaaagg tttagtcgtt tatttggctt tgaatatgga 480

caagcaaggg cgtggggatc atttggttat gccttggttag cgcttttggc aggattttta 540

tttgtaaaaa atcctcattt aaacttttgg gcgggatcctt tctttggttc tttactattg 600

ttaaatttat tattttggaa ccctaaagt gaacgggaag caaatcaaaa ttttaatcaa 660

gaacaagctg aatcaaatag tattccttct ttaaaagaaa tgtttgatct aatgaaactg 720

cctcaattat ggacgataat catctttatt gtttttacat ggacatttta tacgggtattc 780

gatcaacaaa tgtttccggg attttatact ggtttgtttt caacatcagc taatggtgaa 840

aaaatatatg ggacattgaa tgctattcaa gtattttgtg aagcgtaat gatgggaatt 900

gttccaatca ttatgagaaa attaggggtt cgaaatactt tgttattagg tgtaaccatt 960

atgtgtgtac gaattggatt gtgcgggttt gcctcgacac cattatctgt ttcattgcata 1020

aaaatgttgc atgctttaga agtaccatta ttacattac caatgtttcg ctattttaca 1080

cttcattttg atacaaagct atcagcaacc ctctatatga taggatttca gatagctgct 1140

caaattgggc aagtgtttt atcaacacca ttgggaatat taagagacaa cgttggctat 1200

caaccaacat ttaaaattat ttctcttatt gtattactag caggcatata tgcattcttt 1260

attcttaaac aagatgatag agatgttcaa ggggatccat ttattcgagg ataa 1314

<210> 70

<211> 437

<212> PRT

<213> Enterococcus faecium

<400> 70

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

1 5 10 15

Val Asn Val Phe Lys Asn Gln Ser Tyr Leu Gln Ser Ser Ala Thr Leu

20 25 30

Leu Leu Phe Phe Ala Ser Trp Gly Val Trp Trp Ser Phe Phe Gln Leu

35 40 45

Trp Leu Thr Ser Glu Ser Asn Gly Leu Gly Leu Ser Gly Ser Ala Val

50 55 60

Gly Thr Val Phe Ser Ala Asn Ser Leu Val Thr Leu Ile Leu Met Phe

65 70 75 80

Ile Tyr Gly Thr Leu Gln Asp Lys Leu Tyr Ile Lys Arg Asn Leu Leu

85 90 95
 Ile Phe Ala Ser Val Leu Ala Thr Leu Val Gly Pro Phe Phe Ile Trp
 100 105 110
 Ile Tyr Gly Pro Leu Leu Asp Asn Asn Phe Asn Leu Gly Ile Ile Met

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

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

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

 305 310 315 320
 Met Cys Val Arg Ile Gly Leu Cys Gly Phe Ala Ser Thr Pro Leu Ser
 325 330 335

Val Ser Cys Ile Lys Met Leu His Ala Leu Glu Val Pro Leu Phe Thr
340 345 350

Leu Pro Met Phe Arg Tyr Phe Thr Leu His Phe Asp Thr Lys Leu Ser
355 360 365

Ala Thr Leu Tyr Met Ile Gly Phe Gln Ile Ala Ala Gln Ile Gly Gln

370 375 380

Val Ile Leu Ser Thr Pro Leu Gly Ile Leu Arg Asp Asn Val Gly Tyr
385 390 395 400

Gln Pro Thr Phe Lys Ile Ile Ser Leu Ile Val Leu Leu Ala Gly Ile
405 410 415

Tyr Ala Phe Phe Ile Leu Lys Gln Asp Asp Arg Asp Val Gln Gly Asp
420 425 430

Pro Phe Ile Arg Gly
435

<210

> 71

<211> 1299

<212> DNA

<213> Corynebacterium glucuronolyticum

<400> 71

atgtcaaaat tcatgcagca gctgaaaaac actgcgtatc agcagtcac agcgcaactc 60
ctgctcttct tcatgtcctg gggcatctgg tggctcttct tccagctttg gctctccagc 120
gaaacccgcg gtctcggctt caacggcagt gagatcggca ccatctactc ggtgaactcc 180
gccgtcacgc tcgtcctcat gctcgtctac ggaactgccc aagataagct tcgtactcgc 240
cgtaatttgg tgatcggat tgcagttcta atgagcttga ccggcccgtt cttcatgtgg 300
gtctactggc cactgctgca gagcgagtcg ctctatgtcc tcggtgttgg acttggcgca 360

atcttcatcg gtacggcttt tgtgggtca tgcccgtgt tcgaggcgct tgccgagcgc 420
atgtcccgaa aacacaactt cgaatatggc caggcccgcg cgtggggttc ctttggtac 480
gccatcgctg cactcctgc cggttcaac ttcacatca acccgcgat taacttctgg 540
atggcctcgg ccttcggcgt tctgttgctt ctcatcctcg ttttctggaa ggaaccggta 600
gcgcctcgta acgaaattgc agaggaggaa gtggaaaaca ccacacntag cgtcaaggaa 660
atggtgtctg ttctcaaagt gcccgcctc tgggtcgtca ttgtcctcgt gtcttcacg 720

tggacgttct acacggtctt cgaccagcag atgttccgc agttctacac ctcacttttt 780

agtgactccg ccaccggcga gcgaacctat ggctgtctca actccgtcca agtggttcgtc 840

gaggcggtga tgatgggaat cgtgcccac tacatgcgga aggtcggcgt gaagaacacc 900

ctcatgacgg gcttcgccgt catggcactg cgcacccatg gttgcgcggt cttcgcggac 960

ccagtcacca tctcctttgt caagatgttc cacgtctctg aggtaccact gtgcacctc 1020

cccatcttcc gctacttcac cctgcacttc cccacgaaga tctcggccac cttgtacatg 1080

gtcggcttcc agattgcctc gcaggtgggt aacgtcgtca tgtccccgat cctcggttcg 1140

ctgcgtgacc gcctcgggtt ccagccgacc ttctatgtca tctcgggaat cgtccttgc 1200

tccgctatct tcgcctggtt ggctctcaag ggcgataagg aacaagtgga gggcgatccc 1260

ttctaccgcg attcggaaact taaggagata caccaatga 1299

<210> 72

<211> 432

<212> PRT

<213> Corynebacterium glucuronolyticum

<400> 72

Met Ser Lys Phe Met Gln Gln Leu Lys Asn Thr Ala Tyr Gln Gln Ser

1 5 10 15

Ser Ala Gln Leu Leu Leu Phe Phe Met Ser Trp Gly Ile Trp Trp Ser

20 25 30

Phe Phe Gln Leu Trp Leu Ser Ser Glu Thr Arg Gly Leu Gly Phe Asn

35 40 45

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

50 55 60

Val Leu Met Leu Val Tyr Gly Thr Ala Gln Asp Lys Leu Arg Thr Arg

65 70 75 80

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

85 90 95

Phe Phe Met Trp Val Tyr Trp Pro Leu Leu Gln Ser Glu Ser Leu Tyr

100 105 110

Val Leu Gly Val Gly Leu Gly Ala Ile Phe Ile Gly Thr Ala Phe Val

115 120 125

Gly Ser Cys Pro Leu Phe Glu Ala Leu Ala Glu Arg Met Ser Arg Lys
130 135 140

His Asn Phe Glu Tyr Gly Gln Ala Arg Ala Trp Gly Ser Phe Gly Tyr
145 150 155 160

Ala Ile Val Ala Leu Leu Ala Gly Phe Asn Phe Thr Ile Asn Pro Ala

165 170 175

Ile Asn Phe Trp Met Ala Ser Ala Phe Gly Val Leu Leu Leu Ile
180 185 190

Leu Val Phe Trp Lys Glu Pro Val Ala Pro Arg Asn Glu Ile Ala Glu
195 200 205

Glu Glu Val Glu Asn Thr Thr Pro Ser Val Lys Glu Met Val Ser Val
210 215 220

Leu Lys Val Pro Ala Leu Trp Val Val Ile Val Leu Val Phe Phe Thr

225 230 235 240

Trp Thr Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro Gln Phe Tyr
245 250 255

Thr Ser Leu Phe Ser Asp Ser Ala Thr Gly Glu Arg Thr Tyr Gly Val
260 265 270

Leu Asn Ser Val Gln Val Phe Val Glu Ala Leu Met Met Gly Ile Val
275 280 285

Pro Ile Tyr Met Arg Lys Val Gly Val Lys Asn Thr Leu Met Thr Gly

290 295 300

Phe Ala Val Met Ala Leu Arg Ile Leu Gly Cys Ala Val Phe Ala Asp
305 310 315 320

Pro Val Thr Ile Ser Phe Val Lys Met Phe His Ala Leu Glu Val Pro
325 330 335

Leu Cys Ile Leu Pro Ile Phe Arg Tyr Phe Thr Leu His Phe Pro Thr
340 345 350

Lys Ile Ser Ala Thr Leu Tyr Met Val Gly Phe Gln Ile Ala Ser Gln

355 360 365

Val Gly Asn Val Val Met Ser Pro Ile Leu Gly Ser Leu Arg Asp Arg

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

420 425 430

<210> 73

<211> 1326

<212> DNA

<213> Bifidobacterium animalis

<400> 73

atggcaacaa ccacgaaggt gtggaggaac cctcctacc tgcaaagctc aaccggcatc	60
ttctgtttct tctgtccttg gggcatctgg tggctgttct tccagcgtg gctcaactcg	120
atgggactca acggcgcgga agtgggcacg atctattcga tcaactcgct ggccacgctc	180
atcctcatgt tcgggtacgg cctcatccag gacaatctcg gactcaagcg ccgtcttgtg	240
ctcgtcatct cgcgcatcgc cgcaactcgc ggacccttcg tgcagttcgt gtacgcgcgc	300
ctgatgagga cgaacatgat ggccgcccga ctcgtgggct ccgtcgttct ctccgcgggc	360
ttcatggcag gctgctcgtc catcgaggcc gtgaccgaac ggtacagccg ccgtttcaac	420
ttcgagtacg gccaatcccg cgcatggggt tccttcggct atgccattgt ggcgcttgtc	480
gccggcttcg tgttcaacat caaccgatg atcaacttct ggctcggctc cgcattcggc	540
gtgggcatgc tcatcgtgta cctcacctgg tatccggccg agcagcgca agcgtcaag	600
gaagccgcgc atccgaatgc cgcgccaact aaccgacca tcaaagacat gctcggcgtg	660
ctcaagatgc ccacgtgtg ggtgctcatc gtgttcatgc tgctcacaa cacgttctac	720
accgtattcg accagcagat gttecccacc tactacgcct cgctcttccc gaatgaggcc	780
accggcaacg ccgtctacgg cacgtcaac tcggtgcagg tgttctgcga atccgcatg	840
atgggcgtcg tgccgatcat catgcgcaag gtaggtgtgc gcaacgcgtt gctgctcgga	900
tccacggtga tgttccttcg catcgggctg tgcggcatct tccacgatcc ggtgtccatc	960
tcgatcgtca aaatgttcca cgccattgaa gttccgctgt tctgcctgcc ggcgttccgc	1020
tacttcacgc tccattcaa tccgaagctc tccgcgacgc tctacatggt cggttccag	1080
attgcctcac agatcggcca ggtcgtcttc tccacccgc tcggcatgct gcatgaccgc	1140

atgggcgacc gcacgacgtt cctgacgac tccgccatcg tgcttgctgc caccgtctac 1200
 ggattcttcg tgatcaagcg cgacgacgag caggtggatg gcgatccgtt catccgcgat 1260
 tcgaagaagc tgccgtcgct cgccaccgac gaggcgatcc tctccgcgga ttccgaggat 1320
 atgtaa 1326

<210> 74

<211> 441

<212> PRT

<213> Bifidobacterium animalis

<400> 74

Met Ala Thr Thr Thr Lys Val Trp Arg Asn Pro Ser Tyr Leu Gln Ser

1 5 10 15

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

20 25 30

Phe Phe Gln Arg Trp Leu Asn Ser Met Gly Leu Asn Gly Ala Glu Val

35 40 45

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

50 55 60

Gly Tyr Gly Leu Ile Gln Asp Asn Leu Gly Leu Lys Arg Arg Leu Val

65 70 75 80

Leu Val Ile Ser Ala Ile Ala Ala Leu Val Gly Pro Phe Val Gln Phe

85 90 95

Val Tyr Ala Pro Leu Met Arg Thr Asn Met Met Ala Ala Ala Leu Val

100 105 110

Gly Ser Val Val Leu Ser Ala Gly Phe Met Ala Gly Cys Ser Leu Ile

115 120 125

Glu Ala Val Thr Glu Arg Tyr Ser Arg Arg Phe Asn Phe Glu Tyr Gly

130 135 140

Gln Ser Arg Ala Trp Gly Ser Phe Gly Tyr Ala Ile Val Ala Leu Val

145 150 155 160

Ala Gly Phe Val Phe Asn Ile Asn Pro Met Ile Asn Phe Trp Leu Gly

165 170 175

Ser Ala Phe Gly Val Gly Met Leu Ile Val Tyr Leu Thr Trp Tyr Pro

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

Ile Leu Ser Ala Asp Ser Glu Asp Met

435	440	
<210>	75	
<211>	1398	
<212>	DNA	
<213>	Bifidobacterium gallicum	
<400>	75	
atggtgaata aaccgaagac cgcaaaaatc tggccaacc cgtcctat	60	
tttggcattt tctgtttctt ctgctcatgg ggcatctggg ggtccttctt ccagcgtg	120	
ctcaatacca ttggcctgaa cggcgcgga gtcggcaccg tctattccat caactcgtg	180	
gccacgtga tctcatgtt cggctacggc atcatccagg acaacctggg catcaagcg	240	
cgtctcgtgg tcgtcatcg caccatcgcg gcaactgatg gcccttctgt ccagttcgtg	300	
tacgcgccg tcatgcagac gaacatcatg gccgccgcc tgatcggctc cgtggtgctc	360	
tccgcgggt tcatgtcgg atgctcgtg attgaagcg ttaccgaac ctacagccg	420	
aagttcgggt tcgaatacgg ccagtcgcc gcatgggggt ccttcggcta cgccattgtg	480	
gccctgatcg ccggcattgt cttcaacatc aaccgatga tcaacttctg gctcggtcc	540	
gcattcggcg tgggcatgct catcgtgtac ctacgtggg acccgccga gcagcgcag	600	
gccctcaagg aagcgccga cccgaacgc gagaagtcca accgtcctt caaggacatg	660	
gtcaacgtgc tcaagatgcc gacgtgtgg gtgctcatca tcttcatgct gctgaccaac	720	
acgttttaca cggcttctga ccagcagatg ttcccgacct actacgctc gctgttccc	780	
agcattgaaa cgggcaaac ggtctacggc gtgctcaact ccatccagg cttctgcga	840	
tccgcgatga tggcgctgt cccgatcatc atgcgcaaga tcggcgtgca caacgcgtg	900	
ctgctggcg ccaccgtcat gttcttgcgc atcggcctgt gcggcatctt ccacgaccg	960	
gtagccatct ccatcgtcaa gatgttccac gccatcgaag ttccactgtt ctgcctgcc	1020	
gcgttccgct acttcacgt gcaattcaac ccgaagctct cggccacgt gtacatggtg	1080	
ggcttccaga tcgctcaca gatcggccag gttatcttct ccaccccgct gggcatgctg	1140	
cacgaccgt tcggcgacc caccacctt ctgtccatca gcggcatcgt gctgctggca	1200	
acgatctacg gcttcttct catcaagcg gacgacgag acgtggacgg cgatccgtt	1260	
ctgctgacc gcgaccgaa ggaaatgaa ctcatcgaag agaactgca gccagacgcc	1320	
gagctggaaa cgagcccggt aggcgtcgca gcacaggtgc gcgacaaccg cgcggtccag	1380	
ccggaatacg caagctga	1398	

<210> 76

<211> 465

<212> PRT

<213> Bifidobacterium gallicum

<400> 76

Met Val Asn Lys Pro Lys Thr Ala Lys Ile Trp Ser Asn Pro Ser Tyr

1 5 10 15

Leu Gln Ser Ser Phe Gly Ile Phe Leu Phe Phe Cys Ser Trp Gly Ile

20 25 30

Trp Trp Ser Phe Phe Gln Arg Trp Leu Asn Thr Ile Gly Leu Asn Gly

35 40 45

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

50 55 60

Leu Met Phe Gly Tyr Gly Ile Ile Gln Asp Asn Leu Gly Ile Lys Arg

65 70 75 80

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

85 90 95

Val Gln Phe Val Tyr Ala Pro Leu Met Gln Thr Asn Ile Met Ala Ala

100 105 110

Ala Leu Ile Gly Ser Val Val Leu Ser Ala Gly Phe Met Ser Gly Cys

115 120 125

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

130 135 140

Glu Tyr Gly Gln Ser Arg Ala Trp Gly Ser Phe Gly Tyr Ala Ile Val

145 150 155 160

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

165 170 175

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

180 185 190

Trp Tyr Pro Ala Glu Gln Arg Gln Ala Leu Lys Glu Ala Ala Asp Pro

195 200 205

Asn Ala Glu Lys Ser Asn Pro Ser Phe Lys Asp Met Val Asn Val Leu

210 215 220
 Lys Met Pro Thr Leu Trp Val Leu Ile Ile Phe Met Leu Leu Thr Asn
 225 230 235 240

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

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

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

 Glu Glu Asn Leu Gln Pro Asp Ala Glu Leu Glu Thr Ser Pro Val Gly
 435 440 445
 Val Ala Ala Gln Val Arg Asp Asn Arg Ala Val Gln Pro Glu Tyr Ala
 450 455 460

Ser

465

<210> 77

<211> 1338

<212> DNA

<213> Bifidobacterium longum

<400> 77

atggcaagtg caaccaagtc tgcattggaag aatccttctt atctgcagag ctctttcggc 60

atcttcatgt tcttctgctc ctggggcctc tgggtgtcct tcttcagcgc ctggctcctc 120

tcaggcgttg gattgaccaa tgctgaagtc ggccaccatct actccatcaa ctgcctggcc 180

accctggtea tcatgtttgt gtacggcgtg attcaggatc agctcgcat caagcgcaag 240

ctcgtcatcg tagtctcggc aatcgccgcc tgcgttggcc cattcgtcca attcgtttac 300

gccccgatga tctcgcgcgg tggcaccacc cgtcggatcg gcgcactcat cggtccatc 360

gttctgtctg ccggcttcat gtccggctgc tccctgttcg aggcctgcac cgaacgtac 420

tcccgtaaat tcggtttcga atatggccag tcccgtgctt ggggtctcctt cggttacgcc 480

atcgtggcgc tgtgcgcggc ctctctgttc aacatcaacc cgtgatcaa ctcttgggtc 540

ggctccgcat tggccctgg catgtctctc gtgtacgcct tctgggtccc ggccgagcag 600

aaggaagagc tcaagaagga aaccgacccg aacgcagccc ccaccaaccc gtccctcaag 660

gaaatggtcg ccgttctcaa gatgccgacc ctgtgggtgc tcatcgtctt catgtgtctg 720

accaacacct tctacaccgt gttcgatcag cagatgttcc cgacctacta cgccaacctc 780

ttccccactg aagaatcgg caacgccacc tacggcacc tgaacggttt ccaggcttctc 840

cttgagtccg caatgatggg cgtgggtccc atcatcatga agaagatcgg cgtgcgcaac 900

gtctgtctgc tcggcgtctc cgtgatgttc ctgcgcacgc gcttgtgcgg cgtgttccac 960

gaccgggtea ccatctccat cgtcaagctg ttccactcca tcgaagtgc gctgttctgc 1020

ctgccggcat tccgctactt cactctgcac ttgcacacca agctctctgc cacgtgtac 1080

atggtgggct tccagatcgc ttccaagtg ggtcaggtca tcttctcgac ccctctgggt 1140

gccttccacg acaagatggc tcagattctg ccgaacaacg acatgggac ccgcgtgacc 1200

ttctgggtca tctctgcat cgtgtgtgtc gcaatgattt acggcttctt cgtcatcaag 1260

catgatgatc aggaagtccg cggcgacccg ttctacaccg acaagcagct tcgccagatg 1320

gaagccgcca aggcctga 1338

<210> 78

<211> 445

<212> PRT

<213> Bifidobacterium longum

<400> 78

Met Ala Ser Ala Thr Lys Ser Ala Trp Lys Asn Pro Ser Tyr Leu Gln

1 5 10 15

Ser Ser Phe Gly Ile Phe Met Phe Phe Cys Ser Trp Gly Ile Trp Trp

20 25 30

Ser Phe Phe Gln Arg Trp Leu Ile Ser Gly Val Gly Leu Thr Asn Ala

35 40 45

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

50 55 60

Met Phe Val Tyr Gly Val Ile Gln Asp Gln Leu Gly Ile Lys Arg Lys

65 70 75 80

Leu Val Ile Val Val Ser Val Ile Ala Ala Cys Val Gly Pro Phe Val

85 90 95

Gln Phe Val Tyr Ala Pro Met Ile Leu Ala Gly Gly Thr Thr Arg Trp

100 105 110

Ile Gly Ala Leu Ile Gly Ser Ile Val Leu Ser Ala Gly Phe Met Ser

115 120 125

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

130 135 140

Gly Phe Glu Tyr Gly Gln Ser Arg Ala Trp Gly Ser Phe Gly Tyr Ala

145 150 155 160

Ile Val Ala Leu Cys Ala Gly Phe Leu Phe Asn Ile Asn Pro Leu Ile

165 170 175

Asn Phe Trp Val Gly Ser Ala Phe Gly Pro Gly Met Leu Leu Val Tyr

180 185 190

Ala Phe Trp Val Pro Ala Glu Gln Lys Glu Glu Leu Lys Lys Glu Thr

195 200 205

Asp Pro Asn Ala Ala Pro Thr Asn Pro Ser Leu Lys Glu Met Val Ala

210 215 220

Val Leu Lys Met Pro Thr Leu Trp Val Leu Ile Val Phe Met Leu Leu

225 230 235 240

Thr Asn Thr Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro Thr Tyr

245 250 255

Tyr Ala Asn Leu Phe Pro Thr Glu Glu Ile Gly Asn Ala Thr Tyr Gly

260 265 270

Thr Leu Asn Gly Phe Gln Val Phe Leu Glu Ser Ala Met Met Gly Val

275 280 285

Val Pro Ile Ile Met Lys Lys Ile Gly Val Arg Asn Ala Leu Leu Leu

290 295 300

Gly Ala Thr Val Met Phe Leu Arg Ile Gly Leu Cys Gly Val Phe His

305 310 315 320

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

325 330 335

Pro Leu Phe Cys Leu Pro Ala Phe Arg Tyr Phe Thr Leu His Phe Asp

340 345 350

Thr Lys Leu Ser Ala Thr Leu Tyr Met Val Gly Phe Gln Ile Ala Ser

355 360 365

Gln Val Gly Gln Val Ile Phe Ser Thr Pro Leu Gly Ala Phe His Asp

370 375 380

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

385 390 395 400

Phe Trp Val Ile Ser Ala Ile Val Leu Cys Ala Leu Ile Tyr Gly Phe

405 410 415

Phe Val Ile Lys His Asp Asp Gln Glu Val Gly Gly Asp Pro Phe Tyr

420 425 430

Thr Asp Lys Gln Leu Arg Gln Met Glu Ala Ala Lys Ala

435 440 445

<210> 79

<211> 1422

<212> DNA

<213> Bifidobacterium adolescentis

<400> 79

ctgaaatcag agcaggcgca agccaaaaca acatcggaag cgatcgctgc cgcgcggcag 60
cggcagcgcg aagagaaaaa gagaatcaaa atggcaagca aaacacgttc tgtatggaag 120
aatccttctt atctgcagag ctctctcggc attttcatgt tcttctgttc ctggggcatc 180
tggtggtcct tcttctcccg ctggctcact gaccgcaccc acggtctggg catgagctcc 240
gcggaacagg gccagatcta ctccatcaac tccttggcca ccttgggtcat catgttcgtt 300

tacggcacca ttcaggacca gctgggcatt aagcgtaagc tcgtgatctt catctctgcg 360
gtcgtgcat gcgttggccc gttcgtgcag ttcgtgtacc agccgatgct gaccgcccgc 420
ggcaccaccc gattcatcgg cgtgcttctc ggctccatcg tgctgtccgc aggttcatg 480
gccggtgctt cctgtttcga agccatcacc gaacgttact cccgtaagtt cggttctgaa 540
tacggccagt cccgcgcttg gggctccttc ggctacgtg tcgtggcact gtgcgcaggc 600
ttcctgttca acatcaaccc gctgctgaac ttctgggttg gttccatctg cggcctcagc 660
atgctgtgcg tctatgcttt ctgggttccg gccgagcaga aggaagaact caagaaggaa 720

gctgatccga acgcaactcc gaccaacccg tccttcaagg aaatggtctc cgtcctgaag 780
atgccgaccc tgtgggtgct catcgtcttc atgctgttca ccaacacctt ctacaccgtg 840
ttcgatcagc agatgttccc gaactactac gcctccctct tcccgaccac cgaaatcggc 900
aacgccacct acggcacctt gaactccttc caggtgttcc ttgagtccgc catgatgggc 960
gtcgtcccga tcatcatgaa gaagatcggc gtgcgtaact ccctgctgct cggcgccacc 1020
gtgatgttgc cccgtatcgg tctgtcggc gtgttccatg acccggtctc cgtctccatc 1080
gtcaagctgt tccactcat cgaggtaccg ctgttctgcc tgccggcggt ccgtacttc 1140

accctgcact tcgacacgaa gctgtctgcc accctgtaca tggttggttt ccagatcgtt 1200
tcccaggctg gccaggtgat tttctccacc ccgatgggtg ctctgcatga tgccatgggc 1260
gaccgtccga ctttcttcac catctctgcc atcgtgtttg cggtctgtgt ctacggcttc 1320
ttcgtcatca agaaggatga tcaggaagtc ggccggcgatc cgttctacac tgacaagcag 1380
ctcaaggcca tgaaggccgc tgatgcggaa gtgaaggcct ga 1422

<210> 80

<211> 473

<212> PRT

<213> Bifidobacterium adolescentis

<400> 80

Met Lys Ser Glu Gln Ala Gln Ala Lys Thr Thr Ser Glu Ala Ile Ala

1 5 10 15

Ala Ala Arg Gln Arg Gln Arg Glu Glu Lys Lys Arg Ile Lys Met Ala

20 25 30

Ser Lys Thr Arg Ser Val Trp Lys Asn Pro Ser Tyr Leu Gln Ser Ser

35 40 45

Phe Gly Ile Phe Met Phe Phe Cys Ser Trp Gly Ile Trp Trp Ser Phe

50 55 60

Phe Ser Arg Trp Leu Thr Asp Pro Thr His Gly Leu Gly Met Ser Ser

65 70 75 80

Ala Glu Gln Gly Gln Ile Tyr Ser Ile Asn Ser Leu Ala Thr Leu Val

85 90 95

Ile Met Phe Val Tyr Gly Thr Ile Gln Asp Gln Leu Gly Ile Lys Arg

100 105 110

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

115 120 125

Val Gln Phe Val Tyr Gln Pro Met Leu Thr Ala Gly Gly Thr Thr Arg

130 135 140

Phe Ile Gly Val Leu Leu Gly Ser Ile Val Leu Ser Ala Gly Phe Met

145 150 155 160

Ala Gly Cys Ser Leu Phe Glu Ala Ile Thr Glu Arg Tyr Ser Arg Lys

165 170 175

Phe Gly Phe Glu Tyr Gly Gln Ser Arg Ala Trp Gly Ser Phe Gly Tyr

180 185 190

Ala Val Val Ala Leu Cys Ala Gly Phe Leu Phe Asn Ile Asn Pro Leu

195 200 205

Leu Asn Phe Trp Val Gly Ser Ile Cys Gly Leu Ser Met Leu Cys Val

210 215 220

Tyr Ala Phe Trp Val Pro Ala Glu Gln Lys Glu Glu Leu Lys Lys Glu

225 230 235 240

Ala Asp Pro Asn Ala Thr Pro Thr Asn Pro Ser Phe Lys Glu Met Val

245 250 255
Ser Val Leu Lys Met Pro Thr Leu Trp Val Leu Ile Val Phe Met Leu

260 265 270
Phe Thr Asn Thr Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro Asn

275 280 285
Tyr Tyr Ala Ser Leu Phe Pro Thr Thr Glu Ile Gly Asn Ala Thr Tyr

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

305 310 315 320
Val Val Pro Ile Ile Met Lys Lys Ile Gly Val Arg Asn Ser Leu Leu

325 330 335
Leu Gly Ala Thr Val Met Phe Ala Arg Ile Gly Leu Cys Gly Val Phe

340 345 350
His Asp Pro Val Ser Val Ser Ile Val Lys Leu Phe His Ser Ile Glu

355 360 365
Val Pro Leu Phe Cys Leu Pro Ala Phe Arg Tyr Phe Thr Leu His Phe

370 375 380
Asp Thr Lys Leu Ser Ala Thr Leu Tyr Met Val Gly Phe Gln Ile Ala

385 390 395 400
Ser Gln Val Gly Gln Val Ile Phe Ser Thr Pro Met Gly Ala Leu His

405 410 415
Asp Ala Met Gly Asp Arg Pro Thr Phe Phe Thr Ile Ser Ala Ile Val

420 425 430
Phe Ala Ala Leu Val Tyr Gly Phe Phe Val Ile Lys Lys Asp Asp Gln

435 440 445
Glu Val Gly Gly Asp Pro Phe Tyr Thr Asp Lys Gln Leu Lys Ala Met

450 455 460
Lys Ala Ala Asp Ala Glu Val Lys Ala

465 470

<210> 81

<211> 1503

<212> DNA

<213> Bifidobacterium longum

<400> 81

ttgcgtgtcg aacaggtacg acacgccgat ttacgtcaat cgattgacgt aaatcgattg	60
acgtcgcata ataattacta cataacaact tctacaaagg cacggccgcc cgagcagaag	120
cgttacataa ccaataacca accaagtagt aatcaaagga tgattatggc aagtgaacc	180
aagtctgcat ggaagaatcc ttctatctg cagagctctt tcggcatctt catgttcttc	240
tgctcctggg gcatctggtg gtecttcttc cagcgctggc tcactcagg cgttgattg	300
accaatgctg aagtcggcac catctactcc atcaactcgc tggccaccct ggtcatcatg	360
tttgtgtacg gcgtgattca ggatcagctc ggcatcaagc gcaagctcgt catcgtagtc	420
tcggtaatcg ccgcctgcgt tggccattc gtccaattcg tttagccccc gatgatcctc	480
gccggtggca ccaccgctg gatcggcgca ctcatcggct ccatcgttct gtctgccggc	540
ttcatgtccg gctgtccct gttegaggcc gtcaccgaac gctactccg taaattcggg	600
ttcgaatatg gccagtcgcc tgcttggggc tccttcgggt acgccatcgt ggcgctgtgc	660
gccggcttcc tgttcaacat caaccgctg atcaacttct gggtcggctc cgcattcggc	720
cctggcatgc tctcgtgta cgcttcttgg gtcccggccg agcagaagga agagctcaag	780
aaggaaaccg acccgaacgc agccccacc aaccgctcc tcaaggaaat ggtcgccgtt	840
ctcaagatgc cgaccctgtg ggtgtctatc gtcttcatgc tegtaccaa caccttctac	900
accgtgttcg atcagcagat gtccccgacc tactacgcca acctcttccc cactgaagaa	960
atcggaacg ccacctacgg caccctgaac gggttccagg tcttcttga gtccgcaatg	1020
atgggcgtgg tcccgatcat catgaagaag atggcgctgc gcaacgctct gctgtcggc	1080
gtacacgtga tgttctcgcg catcggttg tggcgctgt tccacgacc ggtcaccatc	1140
tccatcgtca agctgttcca ctccatgaa gtgccgtgt tctgcctgcc ggcattccgc	1200
tacttactc tgcatttga caccaagctc tctgccacgc tgtacatggt gggcttccag	1260
atcgcttccc aagtgggtca ggtcatcttc tcgaccctc tgggtgcctt ccacgacaag	1320
atggctcaga ttctgccgaa caacgacatg ggatcccgcg tgaccttctg ggtcatctct	1380
gccatcgtgc tgtgcgcaat gatttacggc ttcttcgtca tcaagcatga tgatcaggaa	1440
gtcggcggcg accgttcta caccgacaag cagcttcgcc agatggaagc cgccaaggcc	1500
tga	1503

<210> 82

<211> 500

<212> PRT

<213> Bifidobacterium longum

<400> 82

Met Arg Val Glu Gln Val Arg His Ala Asp Leu Arg Gln Ser Ile Asp

1 5 10 15

Val Asn Arg Leu Thr Ser His Asn Asn Tyr Tyr Ile Thr Thr Ser Thr

20 25 30

Lys Ala Arg Pro Pro Glu Gln Lys Arg Tyr Ile Thr Asn Asn Gln Pro

35 40 45

Ser Ser Asn Gln Arg Met Ile Met Ala Ser Ala Thr Lys Ser Ala Trp

50 55 60

Lys Asn Pro Ser Tyr Leu Gln Ser Ser Phe Gly Ile Phe Met Phe Phe

65 70 75 80

Cys Ser Trp Gly Ile Trp Trp Ser Phe Phe Gln Arg Trp Leu Ile Ser

85 90 95

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

100 105 110

Ser Leu Ala Thr Leu Val Ile Met Phe Val Tyr Gly Val Ile Gln Asp

115 120 125

Gln Leu Gly Ile Lys Arg Lys Leu Val Ile Val Val Ser Val Ile Ala

130 135 140

Ala Cys Val Gly Pro Phe Val Gln Phe Val Tyr Ala Pro Met Ile Leu

145 150 155 160

Ala Gly Gly Thr Thr Arg Trp Ile Gly Ala Leu Ile Gly Ser Ile Val

165 170 175

Leu Ser Ala Gly Phe Met Ser Gly Cys Ser Leu Phe Glu Ala Val Thr

180 185 190

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

195 200 205

Trp Gly Ser Phe Gly Tyr Ala Ile Val Ala Leu Cys Ala Gly Phe Leu

210 215 220

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

225	230							235							240								
Pro Gly Met Leu Leu Val Tyr Ala Phe Trp Val Pro Ala Glu Gln Lys																							
245							250							255									
Glu Glu Leu Lys Lys Glu Thr Asp Pro Asn Ala Ala Pro Thr Asn Pro																							
260							265							270									
Ser Leu Lys Glu Met Val Ala Val Leu Lys Met Pro Thr Leu Trp Val																							
275							280							285									
Leu Ile Val Phe Met Leu Leu Thr Asn Thr Phe Tyr Thr Val Phe Asp																							
290							295							300									
Gln Gln Met Phe Pro Thr Tyr Tyr Ala Asn Leu Phe Pro Thr Glu Glu																							
305							310							315					320				
Ile Gly Asn Ala Thr Tyr Gly Thr Leu Asn Gly Phe Gln Val Phe Leu																							
325							330							335									
Glu Ser Ala Met Met Gly Val Val Pro Ile Ile Met Lys Lys Ile Gly																							
340							345							350									
Val Arg Asn Ala Leu Leu Leu Gly Ala Thr Val Met Phe Leu Arg Ile																							
355							360							365									
Gly Leu Cys Gly Val Phe His Asp Pro Val Thr Ile Ser Ile Val Lys																							
370							375							380									
Leu Phe His Ser Ile Glu Val Pro Leu Phe Cys Leu Pro Ala Phe Arg																							
385							390							395					400				
Tyr Phe Thr Leu His Phe Asp Thr Lys Leu Ser Ala Thr Leu Tyr Met																							
405							410							415									
Val Gly Phe Gln Ile Ala Ser Gln Val Gly Gln Val Ile Phe Ser Thr																							
420							425							430									
Pro Leu Gly Ala Phe His Asp Lys Met Ala Gln Ile Leu Pro Asn Asn																							
435							440							445									
Asp Met Gly Ser Arg Val Thr Phe Trp Val Ile Ser Ala Ile Val Leu																							
450							455							460									
Cys Ala Leu Ile Tyr Gly Phe Phe Val Ile Lys His Asp Asp Gln Glu																							
465							470							475					480				

Val Gly Gly Asp Pro Phe Tyr Thr Asp Lys Gln Leu Arg Gln Met Glu

485

490

495

Ala Ala Lys Ala

500

<210> 83

<211> 1266

<212> DNA

<213> Mitsukella multacida

<400> 83

atgggaaatc tcttgaaggc attttcgaat ccgttctaca ggacgagctc gcttgagatc	60
ctgctgttct tcgcgggctg gggcatctgg tggcgttct ttcagatctg gctgacgacg	120
aagcagggtc tcacgggcgc gcaggtcggc acgatttact ccttcggcag cgcggtcgcg	180
ctcgtcctga tggtcgtcta cggctccctg caggacaagc tcggcatgaa gaagacgatg	240
ctgaagtctc tcgccgtctg ccagatcttc gtcggcccggt tcttcacctg ggtctacgtg	300
ccgatgctcg ccgcgaactt ctacgtcggc gctgtcgtcg gtgccgtcta cctcgcggtg	360
gcgttctctg cggcctgccc tgtctttgag gcggtcacag agcgcctgag ccgccctac	420
tcctttgagt acggccaggc cagagcctgg ggctcgttcg gctatgccgt ggacgcgctc	480
tgcgcaggct tctctttcac gatgaacccg aacctgatct tctggacggg ctccgctgtg	540
gcggcgggtg agcttatcgt cttggtctcg atgacgccgg agaacgacgc ttcgcttacg	600
gcgcagtacg aggtcaaggc agagagcatc aaggagagca agacgccgtc gttcggcgag	660
atcgtcggcg tggtcaagct catcagggtc tggaagatga tcgtcttcgt catcatgagc	720
tggacgttct acaccgtctt tgaccagcag atgttcccgg agttcttcac gcgtttcttc	780
gcgacgccag aagcaggcca gcaggcttac ggctgtctca actccatcga agtcttcttc	840
gaattctcga tgatgggcct cgtgccgacg ctcatgcgcc gtatcggcgt tcgcaaggcc	900
atcctgctcg gctgcgccat catgatcgtc cgcacggcgg gctgcggcct cgtcacgaat	960
cctcttggcg tcgccgtcat caagctcttg cacgcaccgg aaacggcgct cttcatcttc	1020
gctgtcttcc gctacttcac gctgcacttt gacacgcgca tctcggcgac gctctacatg	1080
gtcggtttcc agatcgtgc acaggtcggc cagattatct tctcgacgcc gctcggcgcc	1140
ctgcatgaca gcatcggcta ccagagcact ttctcgtca tctccggcat cgtctgtgtg	1200
gccagcctct acgttttcgt catctcaag aaagacgacc agcaggtcga cggccagccg	1260
ctttga	1266

<210> 84

<211> 421

<212> PRT

<213> *Mitsuokella multacida*

<400> 84

Met Gly Asn Leu Leu Lys Ala Phe Ser Asn Pro Phe Tyr Arg Thr Ser

1 5 10 15

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

20 25 30

Phe Phe Gln Ile Trp Leu Thr Thr Lys Gln Gly Phe Thr Gly Ala Gln

35 40 45

Val Gly Thr Ile Tyr Ser Phe Gly Ser Ala Val Ala Leu Val Leu Met

50 55 60

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

65 70 75 80

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

85 90 95

Trp Val Tyr Val Pro Met Leu Ala Ala Asn Phe Tyr Val Gly Ala Val

100 105 110

Val Gly Ala Val Tyr Leu Ala Val Ala Phe Leu Ala Ala Cys Pro Val

115 120 125

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

130 135 140

Gly Gln Ala Arg Ala Trp Gly Ser Phe Gly Tyr Ala Val Ala Ala Leu

145 150 155 160

Cys Ala Gly Phe Leu Phe Thr Met Asn Pro Asn Leu Ile Phe Trp Thr

165 170 175

Gly Ser Ala Val Ala Ala Val Gln Leu Ile Val Leu Val Ser Met Thr

180 185 190

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

195 200 205

Ser Ile Lys Glu Ser Lys Thr Pro Ser Phe Gly Glu Ile Val Gly Val

210 215 220
Phe Lys Leu Ile Glu Val Trp Lys Met Ile Val Phe Val Ile Met Ser

225 230 235 240
Trp Thr Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro Glu Phe Phe

245 250 255
Thr Arg Phe Phe Ala Thr Pro Glu Ala Gly Gln Gln Ala Tyr Gly Val

260 265 270
Leu Asn Ser Ile Glu Val Phe Leu Glu Phe Leu Met Met Gly Leu Val

275 280 285
Pro Ile Leu Met Arg Arg Ile Gly Val Arg Lys Ala Ile Leu Leu Gly

290 295 300
Cys Ala Ile Met Ile Val Arg Ile Gly Gly Cys Gly Leu Val Thr Asn

305 310 315 320
Pro Leu Gly Val Ala Val Ile Lys Leu Leu His Ala Pro Glu Thr Ala

325 330 335
Leu Phe Ile Leu Ala Val Phe Arg Tyr Phe Thr Leu His Phe Asp Thr

340 345 350
Arg Ile Ser Ala Thr Leu Tyr Met Val Gly Phe Gln Ile Ala Ala Gln

355 360 365
Val Gly Gln Ile Ile Phe Ser Thr Pro Leu Gly Ala Leu His Asp Ser

370 375 380
Ile Gly Tyr Gln Ser Thr Phe Leu Val Ile Ser Gly Ile Val Cys Val

385 390 395 400
Ala Ser Leu Tyr Ala Phe Val Ile Leu Lys Lys Asp Asp Gln Gln Val

405 410 415
Asp Gly Gln Pro Leu

420

<210

> 85

<211> 1275

<212> DNA

<213> Lactobacillus antri

<400> 85

atgaaaaata gcaagttatc agcgtttaaa aacagctttt acctggagag ttcgcttagt 60
ctgctgctgt tcttcgccgc gtggggaatc tgggtgctgt tcttccaaat ctggctcacc 120
aatgacctcg gcttctctgg ggccaaggtc gggatgatct atactttcga ttcggcaatt 180
acgctggtct taatgttcat ctacgggtca gtgcaagaca agctcggcat taaacgccgg 240
ctgctgattg gggttacat cctggaaatg ctcttgggc ctttctttac ctggatttac 300
gcgccactgc tgcactctaa ctttatcttc ggcgcttct taggttcct ctacctctcc 360

tttgcctttc tggcggcgtc cccgaccttc gaggcctcg cagaacggat gagccggcgg 420
tacagctttg aataggtcg ggcccgggcc tgggggtcat ttggttacgc cgtttcggca 480
ttgtgtgccg gctacctctt caccatcagt ccctacatcg tcttttggct cagcagcggg 540
attagcttgc taaccttctt cctgctctgc tttggccgga ctaagagccc cacacaggtt 600
gcccgttacg agaataaggc cgaggaagaa cacgacgcgg ataagccgag tttcaaagag 660
atcatcagt tttcaagct caagcagttg tgggaattgg ttttcttcat tattttcagc 720
gggtcctttt acacggtctt tgaccagcag atgtttcccc agttctttac ccaatttttc 780

aagacggcgg cccagggaaa cacgacctac ggaatcctca attcgattga agtcttctc 840
gaagcaatta tgatggcgat tgttccttgg attatgaaga agatcggggg ccgcaagacc 900
ctcttgattg gggtcacat tatgttcttg cggatcggcc tctgcggcct ggtcgtcagc 960
ccggtcggga tctcgattgt gaagctcttt cacgccccgg aaacggccat ctttgcctg 1020
gcgatgttcc gctatttgac cctccacttt gacaccgggc tatcggcgac gatgtacatg 1080
gtggttgggc agattgccgg tcaaatcggc cagatcatcc tgtcgacgcc cctgggaatg 1140
ctccacgacc ggatcggcta ccgggcgacc ttcttggtta tttcgtgat tgtgatttgc 1200

gctgcggtat acgcattcgt cattttgcgc aaggataacc aggaggttga cggtaacca 1260
ctagaaaaca actaa 1275

<210> 86

<211> 424

<212> PRT

<213> Lactobacillus antri

<400> 86

Met Lys Asn Ser Lys Leu Ser Ala Phe Lys Asn Ser Phe Tyr Leu Glu

1 5 10 15

Ser Ser Leu Ser Leu Leu Leu Phe Phe Ala Ala Trp Gly Ile Trp Trp

20 25 30

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

35 40 45

Lys Val Gly Met Ile Tyr Thr Phe Asp Ser Ala Ile Thr Leu Val Leu

50 55 60

Met Phe Ile Tyr Gly Ser Val Gln Asp Lys Leu Gly Ile Lys Arg Arg

65 70 75 80

Leu Leu Ile Gly Val Thr Ile Leu Glu Met Leu Leu Gly Pro Phe Phe

85 90 95

Thr Trp Ile Tyr Ala Pro Leu Leu His Ser Asn Phe Ile Leu Gly Ala

100 105 110

Phe Leu Gly Ser Leu Tyr Leu Ser Phe Ala Phe Leu Ala Ala Ser Pro

115 120 125

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

130 135 140

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

145 150 155 160

Leu Cys Ala Gly Tyr Leu Phe Thr Ile Ser Pro Tyr Ile Val Phe Trp

165 170 175

Leu Ser Ser Gly Ile Ser Leu Leu Thr Phe Leu Leu Leu Cys Phe Gly

180 185 190

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

195 200 205

Glu Glu His Asp Ala Asp Lys Pro Ser Phe Lys Glu Ile Ile Ser Val

210 215 220

Phe Lys Leu Lys Gln Leu Trp Glu Leu Val Phe Phe Ile Ile Phe Ser

225 230 235 240

Gly Ser Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro Gln Phe Phe

245 250 255

Thr Gln Phe Phe Lys Thr Ala Ala Gln Gly Asn Thr Ala Tyr Gly Ile

260 265 270

Leu Asn Ser Ile Glu Val Phe Leu Glu Ala Ile Met Met Ala Ile Val

275 280 285
Pro Trp Ile Met Lys Lys Ile Gly Val Arg Lys Thr Leu Leu Ile Gly

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

355 360 365
Ile Gly Gln Ile Ile Leu Ser Thr Pro Leu Gly Met Leu His Asp Arg
370 375 380
Ile Gly Tyr Arg Ala Thr Phe Leu Val Ile Ser Leu Ile Val Ile Cys
385 390 395 400
Ala Ala Val Tyr Ala Phe Val Ile Leu Arg Lys Asp Asn Gln Glu Val
405 410 415
Asp Gly Gln Pro Leu Glu Asn Asn
420

<210> 87

<211> 1323

<212> DNA

<213> Lactobacillus ruminis

<400> 87

atgatgccga ttcttgacaa ttggaaagga attttattta tgaacgatat gaataaaagc 60
ggacggatgt cacaactgaa gaatccgttc ttacaagca atgcgacaaa tattctcatg 120
ttctttgctg gctggggcat ctggtggtca ttcttcaga tctggctgac aaccaagcag 180
gggttcaccg gagcccaggt tggcgagata tactccttca actcggcggt ctcactgatt 240
gccaaccttg ttacagcaa cattcaggac aggctcggcc tcaaacgcaa ccttttgatc 300
ttctgcgcct gcctgcaggt gttcctcggg cccttcttca cgttctctt cgtgccgatg 360

cttcatgcc aacctgaact cggcgctctg atcggttcat gctacctgac gcttgccat 420

ctttccgcat ccccgatgtt cgaggcactg acggaacgtg caagccgccg cttcaactat 480
cagtatgggt cagcgcgtgc ctggggctcg ttcggatatg ccgtatccgc cttgcttgca 540
ggattcgtct tcacaatcaa tccgtcgtcg ctgttctgga tcggctctgc catcgtctgtt 600
gtcctttctt tctgtctttt gttctggaac cctgtccgca acaaggagac ggttgccaga 660
tttgaaaatg aaatggtcag ggaacgtgag aactccaagc ctgggtcaag ggacttcctc 720
aatgtcttca aggttcgcag cctttgggaa atcgccatit tcttgtctt cagcgttaca 780

ttctacacga ttttcgatca gcagatgttt cctcagttct tcaactcagtt cttcaagacc 840
caggcaatgg gcgatcacat gtatgggac ctgaactcgg ttgaggtgtt cctcgaagca 900
ctcatgatgg gcctgggtcc gcttctcatg aagaagatcg gcgtccgccg cacgattctt 960
gtcggcgtga cgttcatgtt catcagaatc ggtggctgcg gtctgattac gaacctctt 1020
ggcgtttcaa tgatcaagct tctccatcg cctgaaacgg ccattttctg cgtcgtaatg 1080
ttccgttact acactctgca ctacgatccg cgagtatcag ccacgatcaa tatcgtaacg 1140
ggcattgcgg gttcgttcgg ccagatactt ctctcaacgc cgcttggact tctgcgtgac 1200

cacatcggtc atcagccgac cttctctggt atcgccggca tcgtattctg cgccggcacc 1260
tacggcttat tcatcattcg aagggatgat caggaagtaa acggagagag gctgtctgaa 1320
taa 1323

<210> 88

<211> 440

<212> PRT

<213> Lactobacillus ruminis

<400> 88

Met Met Pro Ile Ser Asp Asn Trp Lys Gly Ile Leu Phe Met Asn Asp

1 5 10 15

Met Asn Lys Ser Gly Arg Met Ser Gln Leu Lys Asn Pro Phe Phe Thr

20 25 30

Ser Asn Ala Thr Asn Ile Leu Met Phe Phe Ala Gly Trp Gly Ile Trp

35 40 45

Trp Ser Phe Phe Gln Ile Trp Leu Thr Thr Lys Gln Gly Phe Thr Gly

50 55 60

Ala Gln Val Gly Glu Ile Tyr Ser Phe Asn Ser Ala Phe Ser Leu Ile

65 70 75 80

Ala Asn Leu Val Tyr Ser Asn Ile Gln Asp Arg Leu Gly Leu Lys Arg

85

90

95

Asn Leu Leu Ile Phe Cys Ala Cys Leu Gln Val Phe Leu Gly Pro Phe

100

105

110

Phe Thr Phe Leu Phe Val Pro Met Leu His Ala Asn Leu Glu Leu Gly

115

120

125

Ala Leu Ile Gly Ser Cys Tyr Leu Thr Leu Ala Tyr Leu Ser Ala Ser

130

135

140

Pro Met Phe Glu Ala Leu Thr Glu Arg Ala Ser Arg Arg Phe Asn Tyr

145

150

155

160

Gln Tyr Gly Ser Ala Arg Ala Trp Gly Ser Phe Gly Tyr Ala Val Ser

165

170

175

Ala Leu Leu Ala Gly Phe Val Phe Thr Ile Asn Pro Ser Leu Leu Phe

180

185

190

Trp Ile Gly Ser Ala Ile Ala Val Val Leu Leu Leu Leu Leu Leu Phe

195

200

205

Trp Asn Pro Val Arg Asn Lys Glu Thr Val Ala Arg Phe Glu Asn Glu

210

215

220

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

225

230

235

240

Asn Val Phe Lys Val Arg Ser Leu Trp Glu Ile Ala Ile Phe Leu Val

245

250

255

Phe Ser Gly Thr Phe Tyr Thr Ile Phe Asp Gln Gln Met Phe Pro Gln

260

265

270

Phe Phe Thr Gln Phe Phe Lys Thr Gln Ala Met Gly Asp His Met Tyr

275

280

285

Gly Ile Leu Asn Ser Val Glu Val Phe Leu Glu Ala Leu Met Met Gly

290

295

300

Leu Val Pro Leu Leu Met Lys Lys Ile Gly Val Arg Arg Thr Ile Leu

305

310

315

320

Val Gly Val Thr Phe Met Phe Ile Arg Ile Gly Gly Cys Gly Leu Ile

325

330

335

Thr Asn Pro Leu Gly Val Ser Met Ile Lys Leu Leu His Ala Pro Glu
340 345 350

Thr Ala Ile Phe Cys Val Val Met Phe Arg Tyr Tyr Thr Leu His Tyr
355 360 365

Asp Pro Arg Val Ser Ala Thr Ile Asn Ile Val Thr Gly Ile Ala Gly
370 375 380

Ser Phe Gly Gln Ile Leu Leu Ser Thr Pro Leu Gly Leu Leu Arg Asp
385 390 395 400

His Ile Gly Tyr Gln Pro Thr Phe Leu Val Ile Ala Gly Ile Val Phe
405 410 415

Cys Ala Gly Ile Tyr Gly Leu Phe Ile Ile Arg Arg Asp Asp Gln Glu
420 425 430

Val Asn Gly Glu Arg Leu Ser Glu
435 440

<210> 89

<211> 1242

<212> DNA

<213> Yersinia frederiksenii

<400> 89

atgaaacatt ctgtccgtaa tcaatatctg atcttaagtg gcttattgtt tacgtttttc 60
tttacttggt catcggcatt ctctttattc tccatatggc tcaatcaata ttaggatta 120
aaaggtaccg aaacaggggc gactttttcc gccattgcct taacggcact ttgcgtcaa 180

ccgctttatg gcgtgataca agataagttg gggctaataa aacatctttt atgggccatt 240
ggtattttgc tgctgatcag tggccctttt ttattttatg tttatgcccc tttattgcgt 300
gtcaacatgc tggttggtgc cgttaccggt ggcttatata tggggatgac gttctttgcc 360
ggtattggtg cgcttgagtc ttataccgaa cgagttagcc gtattagtgg gtttgagttt 420
ggtaaagccc gtatgtgggg atcgctgggg tggcggggtg caaccttttt tgctggcatg 480
ttgtttaata ttaatcccaa cattaatttc tggatggcat cggcatcggc cgcgatattt 540
ttactgttgt tgtggcactt acatgaagtt aaaacagcgg ctatggggca gttggaatac 600

ggtaagaata gtgccctgac actgagtgat acattgtcac tgtttcgtat gccgcgtttc 660
tgggcgctgg tggatattgt caccggtgtg agcgtttata acgtctatga ccagcaattt 720

ccggtctatt tctcctctct atttactgac cgacgccacg gcaatgaaat gtacggcttt 780
 ctttaattcac tacaggtatt cctagaggct ggtgggtatgt tcctcgcgcc ttttctggtt 840
 aaccgtattg gcgcgaaaaa gggcttactg ctgagcggat taatcatggc aatgcgcata 900
 ttgggttcag gggtggcaca agatgcagtc accatctcat tgatgaagtt attacatgca 960
 gtggagttag ctatcttctc cattgcgatg tttaagtata tcgccgcaa tttcgaccgc 1020

cgtttgcag ccacgcttta tctggtggga tttcagttta ttaccaagt ctatgccagc 1080
 glattttcgc cggtggcagg taaaggctat gacctgatcg gggtcgtga tacctatctg 1140
 atcatgggag gcattgtcct cggattaaca gcaatttctt gttttatgct gcgcggcgag 1200
 tcgcgtacgg atgatacttc cctacaatta accactaagt ga 1242

<210> 90

<211> 413

<212> PRT

<213> *Yersinia frederiksenii*

<400> 90

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

1 5 10 15

Phe Thr Phe Phe Phe Thr Trp Ser Ser Ala Phe Ser Leu Phe Ser Ile

20 25 30

Trp Leu Asn Gln Tyr Val Gly Leu Lys Gly Thr Glu Thr Gly Ala Thr

35 40 45

Phe Ser Ala Ile Ala Leu Thr Ala Leu Cys Ala Gln Pro Leu Tyr Gly

50 55 60

Val Ile Gln Asp Lys Leu Gly Leu Lys Lys His Leu Leu Trp Ala Ile

65 70 75 80

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

85 90 95

Pro Leu Leu Arg Val Asn Met Leu Val Gly Ala Val Thr Gly Gly Leu

100 105 110

Tyr Met Gly Met Thr Phe Phe Ala Gly Ile Gly Ala Leu Glu Ser Tyr

115 120 125

Thr Glu Arg Val Ser Arg Ile Ser Gly Phe Glu Phe Gly Lys Ala Arg

130	135	140	
Met Trp Gly Ser Leu Gly Trp Ala Gly Ala Thr Phe Phe Ala Gly Met			
145	150	155	160
Leu Phe Asn Ile Asn Pro Asn Ile Asn Phe Trp Met Ala Ser Ala Ser			
	165	170	175
Ala Ala Ile Phe Leu Leu Leu Leu Trp His Leu His Glu Val Lys Thr			
	180	185	190
Ala Ala Met Gly Gln Leu Glu Tyr Gly Lys Asn Ser Ala Leu Thr Leu			
195	200	205	
Ser Asp Thr Leu Ser Leu Phe Arg Met Pro Arg Phe Trp Ala Leu Val			
210	215	220	
Val Phe Val Thr Gly Val Ser Val Tyr Asn Val Tyr Asp Gln Gln Phe			
225	230	235	240
Pro Val Tyr Phe Ser Ser Leu Phe Thr Asp Arg Arg His Gly Asn Glu			
	245	250	255
Met Tyr Gly Phe Leu Asn Ser Leu Gln Val Phe Leu Glu Ala Gly Gly			
	260	265	270
Met Phe Leu Ala Pro Phe Leu Val Asn Arg Ile Gly Ala Lys Lys Gly			
275	280	285	
Leu Leu Leu Ser Gly Leu Ile Met Ala Met Arg Ile Leu Gly Ser Gly			
290	295	300	
Leu Ala Gln Asp Ala Val Thr Ile Ser Leu Met Lys Leu Leu His Ala			
305	310	315	320
Val Glu Leu Pro Ile Leu Leu Ile Ala Met Phe Lys Tyr Ile Ala Ala			
	325	330	335
Asn Phe Asp Pro Arg Leu Ser Ala Thr Leu Tyr Leu Val Gly Phe Gln			
340	345	350	
Phe Ile Thr Gln Val Tyr Ala Ser Val Phe Ser Pro Leu Ala Gly Lys			
355	360	365	
Gly Tyr Asp Leu Ile Gly Phe Ala Asp Thr Tyr Leu Ile Met Gly Gly			
370	375	380	

Ile Val Leu Gly Leu Thr Ala Ile Ser Cys Phe Met Leu Arg Gly Glu
385 390 395 400

Ser Arg Thr Asp Asp Pro Ser Leu Gln Leu Thr Thr Lys
405 410

<210> 91

<211> 1260

<212> DNA

<213> *Serratia proteamaculans*

<400> 91

atgaaccgcg aaacaaaaa atattatgtg cttctcagcg gcctgttggt tttcttcttc 60
tttacctggt catccagctt ttcactgac tccatctggc tgaaccagaa aatcggcctg 120
aaagggactg aaaccgggct gatcttcgcg gcaatgtcga tcatggcggt gtgcgcccac 180
ccgctgtacg gctttattca ggacaaactt gggtctgcga agcacctgct gctgtttgtc 240
ggcgtgctgc tgttgctcac cggcccggtt tttatctatg tctacgcccc gctgctgcag 300

agcaaccttg tggtcggcgc actggtgggc ggctgtttg tcagcctggc gttcaatgcc 360
ggtattggcg cgtcggaaac ctataccgaa cgagtcagcc gcatcgctcg tttcgaattc 420
ggccggggcg gtatgtgggg gtcattgggc tgggccagcg ccaccttctt tgccggcttt 480
aactacaata tcgaccccaa tatcaacttc tggatcgctt cggcctcggc ggcagtgttt 540
ctgctgttgc tgtggcaagt gcgtgagctg aaaccaacg ccatggccgg tctggaatac 600
ggcaagccgg aaaacctgaa gctgcaggac gcattggccc tgcctgcgct gccgggggtt 660
tgggcgctgg tgggtgttgt gctgggcacc agcatctacg gcgtgtttga ccagcagttc 720

ccggtgtatt tcgcctcgca gttccccacc caggaagaag gcaaccgcat gtacggtttc 780
cttaattcgc tgcagggtgt tctggaggcc ggtggcatgt tcctggcccc gctgctggtt 840
aaccgcattg gcataaagca aagcctgttg ctggccagca gcgtgatgac gctgcgcatg 900
gtcggttccg gctttgccag cggcgccctg atgatttccg ccatgaaact getgcacgcc 960
glagaattgc caatcctgct ggtggcgatg ttcaagtaca tcaccaccg tttcgacagc 1020
cgctgtcct ccacgtgta cctggtgggc ttccagtta tcagccaaat tgcgccggt 1080
tttctggcac cgtggccgg ttatggttac gaccgcatcg gctttgccga cacctatttg 1140

ctgatgggtt gcgcggtggc cgggaccacg ctgatttctt gcttctgct gcgcggcgag 1200
accgtcgcca gtgcgctca atttcaatcc acgttaaaat caagtgagcc aaccaatga 1260

<210> 92

<211> 419

<212> PRT

<213> *Serratia proteamaculans*

<400> 92

Met Asn Arg Glu Thr Lys Lys Tyr Tyr Val Leu Leu Ser Gly Leu Leu

1 5 10 15

Phe Phe Phe Phe Phe Thr Trp Ser Ser Ser Phe Ser Leu Ile Ser Ile

20 25 30

Trp Leu Asn Gln Lys Ile Gly Leu Lys Gly Thr Glu Thr Gly Leu Ile

35 40 45

Phe Ala Ala Met Ser Ile Met Ala Leu Cys Ala Gln Pro Leu Tyr Gly

50 55 60

Phe Ile Gln Asp Lys Leu Gly Leu Arg Lys His Leu Leu Leu Phe Val

65 70 75 80

Gly Val Leu Leu Leu Leu Thr Gly Pro Phe Phe Ile Tyr Val Tyr Ala

85 90 95

Pro Leu Leu Gln Ser Asn Leu Val Val Gly Ala Leu Val Gly Gly Val

100 105 110

Phe Val Ser Leu Ala Phe Asn Ala Gly Ile Gly Ala Leu Glu Ser Tyr

115 120 125

Thr Glu Arg Val Ser Arg Ile Val Gly Phe Glu Phe Gly Arg Ala Arg

130 135 140

Met Trp Gly Ser Leu Gly Trp Ala Ser Ala Thr Phe Phe Ala Gly Phe

145 150 155 160

Asn Tyr Asn Ile Asp Pro Asn Ile Asn Phe Trp Ile Ala Ser Ala Ser

165 170 175

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

180 185 190

Asn Ala Met Ala Gly Leu Glu Tyr Gly Lys Pro Glu Asn Leu Lys Leu

195 200 205

Gln Asp Ala Leu Ala Leu Leu Arg Leu Pro Gly Phe Trp Ala Leu Val

210 215 220

Val Phe Val Leu Gly Thr Ser Ile Tyr Gly Val Phe Asp Gln Gln Phe

225 230 235 240

Pro Val Tyr Phe Ala Ser Gln Phe Pro Thr His Glu Glu Gly Asn Arg

245 250 255

Met Tyr Gly Phe Leu Asn Ser Leu Gln Val Phe Leu Glu Ala Gly Gly

260 265 270

Met Phe Leu Ala Pro Leu Leu Val Asn Arg Ile Gly Ile Lys Gln Ser

275 280 285

Leu Leu Leu Ala Ser Ser Val Met Ala Leu Arg Met Val Gly Ser Gly

290 295 300

Phe Ala Ser Gly Ala Leu Met Ile Ser Ala Met Lys Leu Leu His Ala

305 310 315 320

Val Glu Leu Pro Ile Leu Leu Val Ala Met Phe Lys Tyr Ile Thr Thr

325 330 335

Arg Phe Asp Ser Arg Leu Ser Ser Thr Leu Tyr Leu Val Gly Phe Gln

340 345 350

Phe Ile Ser Gln Ile Val Ala Gly Phe Leu Ala Pro Leu Ala Gly Tyr

355 360 365

Gly Tyr Asp Arg Ile Gly Phe Ala Asp Thr Tyr Leu Leu Met Gly Cys

370 375 380

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

385 390 395 400

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

405 410 415

Pro Thr Gln

<210> 93

<211> 1239

<212> DNA

<213> Escherichia coli

<400> 93

atgaaaaaac ggcctactcg aagttacatg ctgctcagcg ctctgctgtt ctttttcttt 60
gtgacctggt cctcatcaag ttcactgctc tcaatctggc ttcaccagga agtggggcta 120
aaagcatcgg aaaccggcat tttttttca gtattatccg tctccgcgct cttcgcgcag 180
gtctgttatg gcittattca ggaccgactt ggtctgcgca aacatttggt atggtttatc 240
accgcgttgt tgatcctctc cggcccggct tatctgcttt ttagttattt gctgagcgtt 300
aatattctgc tgggcagcgt attcgggggc ttatttatcg ggctgacgtt taatgggggt 360

atcggcggtt tggagtccta taccgagcgc gtgcgcgctc aaaglacctt tgagtttggg 420
cgggcacgca tgtgggggtc tctgggctgg gcagttgccg cgttttttgc cgggttactg 480
tttaatatca accctgacct taacttcttg gtggcttcat gctcagggtt aatcttcttc 540
tgccctctgg cccgattaaa ggtggccgcg ccggcaagca tggagaaact cgaaattggc 600
gctaaaaaag tttctctgga agacgccctg cgtctgctta ctctgccgcg ctcttgggca 660
ctgatattct tcgtggtcgg aacctgcatt tacggcgtat acgatcagca attcccgttc 720
tatttctcat cacagttccc gacattacgc gaagggaacg agatgtttgg ctatttaaac 780

tccttccagg tctttctcga ggccgcaggt atgttttgtg cgccgtggct ggttaatcgc 840
attggtgcta aaaatggctt gatattcgca ggaatgggtga tggcgctgcg catgattact 900
tcagggtcgg tggaaggccc cctgcttatt tccattacca aactgcttca cgcggtcgaa 960
ctgccaatat tgtagtcgc catatttaaa tacaacagtc tgaatttcga caaacgtctc 1020
tcctccacca tttatctggt gggatttgcc tgcaccagct ccgtcattgg taccgtattg 1080
tccccgctgg caggctttag ctatgagaga tttggcttcg cccaatccta tctgatcatg 1140
ggcatcatgg tggtcagcac cacgtttatt tccattttcc ttttgcgctc aactaaatcc 1200

tcactcgagc catcttttct gcagcaaaaa gctgtgtaa 1239

<210> 94

<211> 412

<212> PRT

<213> Escherichia coli

<400> 94

Met Lys Lys Arg Pro Thr Arg Ser Tyr Met Leu Leu Ser Ala Leu Leu

1 5 10 15

Phe Phe Phe Phe Val Thr Trp Ser Ser Ser Ser Ser Leu Leu Ser Ile

20 25 30

Trp Leu His Gln Glu Val Gly Leu Lys Ala Ser Glu Thr Gly Ile Ile

35 40 45
 Phe Ser Val Leu Ser Val Ser Ala Leu Phe Ala Gln Val Cys Tyr Gly
 50 55 60
 Phe Ile Gln Asp Arg Leu Gly Leu Arg Lys His Leu Leu Trp Phe Ile
 65 70 75 80
 Thr Ala Leu Leu Ile Leu Ser Gly Pro Ala Tyr Leu Leu Phe Ser Tyr
 85 90 95
 Leu Leu Ser Val Asn Ile Leu Leu Gly Ser Val Phe Gly Gly Leu Phe
 100 105 110

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

 Leu Ile Phe Phe Cys Leu Leu Ala Arg Leu Lys Val Ala Ala Pro Ala
 180 185 190
 Ser Met Glu Lys Leu Glu Ile Gly Ala Lys Lys Val Ser Leu Glu Asp
 195 200 205
 Ala Leu Arg Leu Leu Thr Leu Pro Arg Phe Trp Ala Leu Ile Phe Phe
 210 215 220
 Val Val Gly Thr Cys Ile Tyr Gly Val Tyr Asp Gln Gln Phe Pro Val
 225 230 235 240

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

Phe Ala Gly Met Val Met Ala Leu Arg Met Ile Thr Ser Gly Leu Val
290 295 300

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

Glu Arg Phe Gly Phe Ala Gln Ser Tyr Leu Ile Met Gly Ile Met Val
370 375 380
Phe Ser Thr Thr Phe Ile Ser Ile Phe Leu Leu Arg Ser Thr Lys Ser
385 390 395 400
Ser Ser Glu Pro Ser Phe Leu Gln Gln Lys Ala Val
405 410

<210> 95

<211> 1290

<212> DNA

<213> Weissella paramesenteroides

<400> 95

atggggggac gtcgtatgtc taaatatatg aaggctttta aaaatgagtc ttatcttcaa 60

agttcagtga caatgttatt gtacttcgca tcttggggta tttggtggtc tttctttcaa 120

ttatgggttaa ctgccaaaag cggtttgaat ttggatggta gtgccgttgg aacaatttat 180

tctgctaatt cattagtac cttagtatta atgttttttt atggcacttt acaagataaa 240

ttgggaatta aaaggcattt actaatattc agcgggtgtt gtgcggcatt agtaggtcca 300

ttcttttatt atgcttatga accaatgtta catactaact tcaccatggg ctgtgtgtgt 360

ggttcaattt ttctttcggc aggattttta gcaacgacag gtatttatga ggcactagtt 420

gagcgcttca gtcgcgtttt caaatttgaa tatggtcagg cacgcgatg gggatcattt 480

ggctatgcag tgggtgcact attggctggc catttatgtg tcattaaccc tgatttaaatt 540

ttctgggttg gatcaatatt tgggtgtgcta ttattgctaa atgtttgttt ttgggtgccca 600

aaagctgaac gtcttgaacg agtacgtagt gctcaagaaa ctgaaaaaac agtaccaagc 660
gtaagagaaa tgttgtcttt gttgaaatg cgtgatcttt gggtagttat tgtcttgatt 720
ttctttacat ggactttcta tactgtatgt gatcaacaaa tgtttccgag ctctcatgca 780
ggattatgtt catctgttgc acaagggtcaa cagatgtacg gtaatttgaa ttcgatccag 840
gtctttgttg aagcaattat gatgggtatt gtaccagtca ttatgaataa aattggtgtt 900

cgtaacacat tattgttagg aattgcaatt atggcaattc gaatcggatt atgtggcttt 960
attgataacc cagttgctat ttcgtttgtg aagatgttgc actcgtttga aacaccatta 1020
tttattctgt cgatttttag atattttaca cttcattttg atacaaaatt atcggcaaca 1080
ttgtatatga taggcttcca agttgccgca cagttaggtc aggttttctt ttcaacaccg 1140
ttgggtatgt tacgagataa ctcaggttat gcagtgacat tccatattat tacaattatc 1200
gttatcgctg cgggaatcta tgcattcttt gttattaaaa aagataatca agatgttaat 1260
ggcgagccat tagttagtaa agctgcttaa 1290

<210> 96

<211> 429

<212> PRT

<213> Weissella paramesenteroides

<400> 96

Met Gly Gly Arg Arg Met Ser Lys Tyr Met Lys Ala Phe Lys Asn Glu

1 5 10 15

Ser Tyr Leu Gln Ser Ser Val Thr Met Leu Leu Tyr Phe Ala Ser Trp

20 25 30

Gly Ile Trp Trp Ser Phe Phe Gln Leu Trp Leu Thr Ala Lys Ser Gly

35 40 45

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

50 55 60

Leu Val Thr Leu Val Leu Met Phe Phe Tyr Gly Thr Leu Gln Asp Lys

65 70 75 80

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

85 90 95

Leu Val Gly Pro Phe Phe Ile Tyr Ala Tyr Glu Pro Met Leu His Thr

100 105 110

Asn Phe Thr Met Gly Leu Val Val Gly Ser Ile Phe Leu Ser Ala Gly

115 120 125

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

130 135 140

Arg Val Phe Lys Phe Glu Tyr Gly Gln Ala Arg Ala Trp Gly Ser Phe

145 150 155 160

Gly Tyr Ala Val Val Ala Leu Leu Ala Gly His Leu Phe Val Ile Asn

165 170 175

Pro Asp Leu Asn Phe Trp Val Gly Ser Ile Phe Gly Val Leu Leu Leu

180 185 190

Leu Asn Val Cys Phe Trp Val Pro Lys Ala Glu Arg Leu Glu Arg Val

195 200 205

Arg Ser Ala Gln Glu Thr Glu Lys Thr Val Pro Ser Val Arg Glu Met

210 215 220

Leu Ser Leu Leu Lys Met Arg Asp Leu Trp Val Val Ile Val Leu Ile

225 230 235 240

Phe Phe Thr Trp Thr Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro

245 250 255

Ser Phe Tyr Ala Gly Leu Phe Ser Ser Val Ala Gln Gly Gln Gln Met

260 265 270

Tyr Gly Asn Leu Asn Ser Ile Gln Val Phe Val Glu Ala Ile Met Met

275 280 285

Gly Ile Val Pro Val Ile Met Asn Lys Ile Gly Val Arg Asn Thr Leu

290 295 300

Leu Leu Gly Ile Ala Ile Met Ala Ile Arg Ile Gly Leu Cys Gly Phe

305 310 315 320

Ile Asp Asn Pro Val Ala Ile Ser Phe Val Lys Met Leu His Ser Phe

325 330 335

Glu Thr Pro Leu Phe Ile Leu Ser Ile Phe Arg Tyr Phe Thr Leu His

340 345 350

Phe Asp Thr Lys Leu Ser Ala Thr Leu Tyr Met Ile Gly Phe Gln Val

355 360 365
Ala Ala Gln Leu Gly Gln Val Phe Leu Ser Thr Pro Leu Gly Met Leu

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

420 425
<210> 97
<211> 1185
<212> DNA
<213> Bifidobacterium pseudocatenulatum

<400> 97
atgacctccg ccgagcaggg ccagatctac tccatcaact ccctggcaac cctggtcac 60
atgttcgect acggcgccat ccaggaccaa ctgggcatca agcgtaaact cgtgatcttc 120
gtctccgcca tcgcagcgct cgctggcccg ttctgtcagt tcgtgtatgc cccgatgctg 180
acagccggcg gcaccaccg ttctcatcggc gtgctcatcg gctccatcgt gctgtccgca 240
ggtttcatgg ccggtgctc cctgttcgag gcactcaccg aacgctactc ccgcaagttc 300
ggcttcgaat acggccagtc ccgagcttgg ggctcattcg gctacgcat cgtggcgctg 360
tgtgcaggct tctgttcaa tatcaaccg ctgctgaact tctgggtcgg ctccatctgc 420

ggcctcggca tgctgtgcat ctacgtttc tgggttcgg ccgagcagaa ggaagagctc 480
aagaaggaag ctgatccgaa cgcggctccg accaaccgt cttcaagga aatgatttc 540
gttctgaaga tgccgacct gtgggtgctc atcgttttca tgctgttcac caacacctc 600
tacaccgtgt tgatecagca gatgttcccg aactactacg cttccctctt ctccaccact 660
gaaatcggca acgccaccta cggcacctg aactccttc aggtgttcct cgagtccgcc 720
atgatgggtg tcgttccgat catcatgaag aagatcggcg tgcgcaactc cctgctgctc 780
ggtgcaaccg tgatgttct gcgcacggc ctgtgcggcg tgttcacga tccggtcagc 840

gtctccatcg tcaagctgtt ccactccatc gaggttcgc tgttctgcct gccggcattc 900
cgctacttca cctgcactt cgacaccaag ctgtcggcaa cctgtacat ggtgggcttc 960
cagatcgctt ccagatcgg ccaggtgatc ttctccacc cgatgggcgc cctgcatgac 1020

gccatgggcg atcgcccgac cttcttcacc atctccggca tcgtgctggc agccctgac 1080
 tacggcttct tcgtgatcaa gaaggacgac caggaagtcg gcggtgatcc gttctacacc 1140
 gacaagcagc tcaaggccca ggccgctgca gaagcgaatg cctga 1185

<210> 98

<211> 394

<212> PRT

<213> Bifidobacterium pseudocatenulatum

<400> 98

Met Thr Ser Ala Glu Gln Gly Gln Ile Tyr Ser Ile Asn Ser Leu Ala

1 5 10 15

Thr Leu Val Ile Met Phe Ala Tyr Gly Ala Ile Gln Asp Gln Leu Gly

20 25 30

Ile Lys Arg Lys Leu Val Ile Phe Val Ser Ala Ile Ala Ala Leu Val

35 40 45

Gly Pro Phe Val Gln Phe Val Tyr Ala Pro Met Leu Thr Ala Gly Gly

50 55 60

Thr Thr Arg Phe Ile Gly Val Leu Ile Gly Ser Ile Val Leu Ser Ala

65 70 75 80

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

85 90 95

Ser Arg Lys Phe Gly Phe Glu Tyr Gly Gln Ser Arg Ala Trp Gly Ser

100 105 110

Phe Gly Tyr Ala Ile Val Ala Leu Cys Ala Gly Phe Leu Phe Asn Ile

115 120 125

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

130 135 140

Leu Cys Ile Tyr Ala Phe Trp Val Pro Ala Glu Gln Lys Glu Glu Leu

145 150 155 160

Lys Lys Glu Ala Asp Pro Asn Ala Ala Pro Thr Asn Pro Ser Phe Lys

165 170 175

Glu Met Ile Ser Val Leu Lys Met Pro Thr Leu Trp Val Leu Ile Val

	180	185	190
Phe Met Leu Phe Thr Asn Thr Phe Tyr Thr Val Phe Asp Gln Gln Met			
195	200	205	
Phe Pro Asn Tyr Tyr Ala Ser Leu Phe Ser Thr Thr Glu Ile Gly Asn			
210	215	220	
Ala Thr Tyr Gly Thr Leu Asn Ser Phe Gln Val Phe Leu Glu Ser Ala			
225	230	235	240
Met Met Gly Val Val Pro Ile Ile Met Lys Lys Ile Gly Val Arg Asn			
245	250	255	
Ser Leu Leu Leu Gly Ala Thr Val Met Phe Leu Arg Ile Gly Leu Cys			
260	265	270	
Gly Val Phe His Asp Pro Val Ser Val Ser Ile Val Lys Leu Phe His			
275	280	285	
Ser Ile Glu Val Pro Leu Phe Cys Leu Pro Ala Phe Arg Tyr Phe Thr			
290	295	300	
Leu His Phe Asp Thr Lys Leu Ser Ala Thr Leu Tyr Met Val Gly Phe			
305	310	315	320
Gln Ile Ala Ser Gln Ile Gly Gln Val Ile Phe Ser Thr Pro Met Gly			
325	330	335	
Ala Leu His Asp Ala Met Gly Asp Arg Pro Thr Phe Phe Thr Ile Ser			
340	345	350	
Gly Ile Val Leu Ala Ala Leu Ile Tyr Gly Phe Phe Val Ile Lys Lys			
355	360	365	
Asp Asp Gln Glu Val Gly Gly Asp Pro Phe Tyr Thr Asp Lys Gln Leu			
370	375	380	
Lys Ala Gln Ala Ala Ala Glu Ala Asn Ala			
385	390		
<210>	99		
<211>	1086		
<212>	DNA		
<213>	Bifidobacterium catenulatum		

<400> 99

gtgctgcgcg gcagtggcac caacaaaaag aaacagagaa acaacatgac aagcaacact	60
cgatccgcac ggaagaatcc ttccatcttg cagagtccct tcggcatctt catgtttctt	120
tgctcgtggg gcatctgggt gtccttcttc tcccgctggc tcaccgaccc gacccacggc	180
ctgggcatga cctccgcaga gcagggccag atctactcca ttaactcctt ggcaaccctg	240
gtcatcatgt tcgcctatgg tgccatccag gatcagctgg gcatcaagcg caagctcgtg	300
atcttcgttt ccgccatcgc agcgctcgtc ggcccgttcg tacagtctgt gtacgccccg	360
atgctgactg ccggcggcac caccgtttc atcggcgtgc tcacggctc catcgtgctc	420
tcctcaggct tcattggcgg ctgctccctg ttccaagcgc tcaccgaacg ctactcccgt	480
aagtctggct tcgaatacgg ccagtcctgc gcttggggct ccttcggcta cgccatcgtg	540
gcactgtgcg caggtttctt gttcaacatc aaccgcgtgc tgaacttctg ggtcggctcc	600
atctgcggcc ttggcatgct gtgcgtctac gctttctggg ttccggccaa gcaaaaggaa	660
gagctcaaga aagaagccga tccgaacgcg gctccaacca atccatcctt caaagagatg	720
atctccgttc tgaagatgcc gactctgtgg gtgctcatcg tcttcatgct gttcaccaac	780
acctttata ccgtgttcga ccagcagatg ttccctaact actacgcttc cctcttcccg	840
accaccgaaa tcggcaacgc cacctacggc accctgaact ccttcagggt gttcctcgag	900
tccgccatga tgggcgtcgt tccgattatc atgaagaaga tcggcgttcg taactccctg	960
ctgctcggcg caaccgtgat gttcctgcgt atcggcctgt gcggcgtgtt ccacgatcca	1020
gtcagtgtct ccacgtcaa gctgttccac tccatcgagg ttccgctgtt ctgcctaccg	1080
gcataa	1086

<210> 100

<211> 361

<212> PRT

<213> Bifidobacterium catenulatum

<400> 100

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

Phe Phe Ser Arg Trp Leu Thr Asp Pro Thr His Gly Leu Gly Met Thr

50

55

60

Ser Ala Glu Gln Gly Gln Ile Tyr Ser Ile Asn Ser Leu Ala Thr Leu

65

70

75

80

Val Ile Met Phe Ala Tyr Gly Ala Ile Gln Asp Gln Leu Gly Ile Lys

85

90

95

Arg Lys Leu Val Ile Phe Val Ser Ala Ile Ala Ala Leu Val Gly Pro

100

105

110

Phe Val Gln Phe Val Tyr Ala Pro Met Leu Thr Ala Gly Gly Thr Thr

115

120

125

Arg Phe Ile Gly Val Leu Ile Gly Ser Ile Val Leu Ser Ser Gly Phe

130

135

140

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

145

150

155

160

Lys Phe Gly Phe Glu Tyr Gly Gln Ser Arg Ala Trp Gly Ser Phe Gly

165

170

175

Tyr Ala Ile Val Ala Leu Cys Ala Gly Phe Leu Phe Asn Ile Asn Pro

180

185

190

Leu Leu Asn Phe Trp Val Gly Ser Ile Cys Gly Leu Gly Met Leu Cys

195

200

205

Val Tyr Ala Phe Trp Val Pro Ala Lys Gln Lys Glu Glu Leu Lys Lys

210

215

220

Glu Ala Asp Pro Asn Ala Ala Pro Thr Asn Pro Ser Phe Lys Glu Met

225

230

235

240

Ile Ser Val Leu Lys Met Pro Thr Leu Trp Val Leu Ile Val Phe Met

245

250

255

Leu Phe Thr Asn Thr Phe Tyr Thr Val Phe Asp Gln Gln Met Phe Pro

260

265

270

Asn Tyr Tyr Ala Ser Leu Phe Pro Thr Thr Glu Ile Gly Asn Ala Thr

275

280

285

Tyr Gly Thr Leu Asn Ser Phe Gln Val Phe Leu Glu Ser Ala Met Met

290	295	300	
Gly Val Val Pro Ile Ile Met Lys Lys Ile Gly Val Arg Asn Ser Leu			
305	310	315	320
Leu Leu Gly Ala Thr Val Met Phe Leu Arg Ile Gly Leu Cys Gly Val			
	325	330	335
Phe His Asp Pro Val Ser Val Ser Ile Val Lys Leu Phe His Ser Ile			
	340	345	350
Glu Val Pro Leu Phe Cys Leu Pro Ala			
	355	360	
<210> 101			
<211> 1176			
<212> DNA			
<213> Clostridium thermocellum			
<400> 101			
atgattttgg ataatcttaa acgttcagaa agatacccg tctcttttat tctgttttat			60
tccttgtttt atatgggcct tgcggtatgt ggctgtgtta tgctgtgtta ttggaaggg			120
ctgggctatg acaatacggg tataggaaca tttctttcaa tcagttcgtt tgcggcctg			180
ttgacacagc ccatttgggg tgcataagt gaccgggcaa aatccaaaa caatgtgctg			240
aaaatgttgg tgcttttcag cagcattgcc atttttatgt ttcgtctctc gggcaactat			300
tactatatat ttgcggtaat ggttgtttat gcctttttcc aaacgccc at tactccgata			360
ggtgatgcga ttacattgga gtatattact gacacaaaat ggaagtatgg cccgataagg			420
cttgcgggtg cattgggata tgcggtgatg gcatttatcg gaggggcatt gacaagaaaa			480
aatatcaacg ctattttctt tataatgcttt gtcataaggta ttatgtcttt gattacagta			540
tttagaatgc caacggtaaa aggacatcaa tcggacggaa acaagctttc cattttagaa			600
gttttcaaaa acagcgaact tgtgtgtgct atgggattta cacttggtat tcatactacc			660
atgggttttt ataatacttt ctttccgatt tactataaaa acatgggtgc tgacaacacc			720
attctgggat tggcgggtgt tatcggtcgc gcgagtgaat taatcttctt tgttttcggc			780
gacaggataa taaaacgttt gggaatcaag ttacgctgt tcggtgcagc ggttgttgca			840
gttgtacggt gggcaagttt gggattgatt aacaatattt ttgcagtgtc tgcactccaa			900
attctccatg gttttatatt cattgttttg gcctactcca tggcaacata tatcaataat			960
gagatgccac ctgaattgaa ggcctcagga cagacggtaa actccgtcat aggtttgggt			1020

atttcagga taattggaag tacaggcggc ggtgtgataa gtgatttaat cggaatcagg 1080
caggtattct ttttaaattc gggtattgtt cttgcttcaa ttgtcatttt tggcgcaata 1140
tttttggtaa gaagacaaaa aattacagga caatag 1176

<210> 102

<211> 391

<212> PRT

<213> Clostridium thermocellum

<400> 102

Met Ile Leu Asp Asn Leu Lys Arg Ser Glu Arg Tyr Pro Phe Ser Phe

1 5 10 15
Ile Leu Phe Tyr Ser Leu Phe Tyr Met Gly Leu Ala Val Phe Gly Val

20 25 30
Phe Met Pro Val Tyr Leu Glu Gly Leu Gly Tyr Asp Asn Thr Asp Ile

35 40 45
Gly Thr Phe Leu Ser Ile Ser Ser Phe Val Gly Leu Phe Ala Gln Pro

50 55 60
Ile Trp Gly Val Ile Ser Asp Arg Ala Lys Ser Lys Asn Asn Val Leu

65 70 75 80
Lys Met Leu Val Leu Phe Ser Ser Ile Ala Ile Phe Met Phe Arg Leu

85 90 95
Ser Gly Asn Tyr Tyr Tyr Ile Phe Ala Val Met Val Val Tyr Ala Phe

100 105 110
Phe Gln Thr Pro Ile Thr Pro Ile Gly Asp Ala Ile Thr Leu Glu Tyr

115 120 125
Ile Thr Asp Thr Lys Trp Lys Tyr Gly Pro Ile Arg Leu Ala Gly Ala

130 135 140
Leu Gly Tyr Ala Val Met Ala Phe Ile Gly Gly Ala Leu Thr Arg Lys

145 150 155 160
Asn Ile Asn Ala Ile Phe Phe Ile Cys Phe Val Ile Gly Ile Met Ser

165 170 175
Leu Ile Thr Val Phe Arg Met Pro Thr Val Lys Gly His Gln Ser Asp

180 185 190

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

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

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

325 330 335
 Ile Gly Leu Gly Ile Ser Arg Ile Ile Gly Ser Thr Gly Gly Gly Val
 340 345 350
 Ile Ser Asp Leu Ile Gly Ile Arg Gln Val Phe Phe Leu Asn Ser Val
 355 360 365
 Ile Val Leu Ala Ser Ile Val Ile Phe Gly Ala Ile Phe Leu Val Arg
 370 375 380
 Arg Gln Lys Ile Thr Gly Gln
 385 390

<210> 103

<211> 1311

<212> DNA

<213> Granulibacter bethesdensis

<400> 103

ttggttttgg atatggacgc aatgagcggg acgatatcgg ttagtccagt ctgtcagat 60
 ggtagcagccc cctctgacag tcgacaaaag cagggagaat ctgctcttct cgacaagttg 120
 aatttttttc ttgctgatgt gcgggatggt ctggggcctt accttgccgt gtacctgttg 180
 tcggcttcgg gcagggacgg gcgctgggac gagtccagtg tcgggctggt gctgaccatc 240
 agtgggtgtg tcgggctggt tgcgcaaacc cctgccggag cactggttga tcgcagtcga 300
 aacaaacccc gcctgctggc tgtcgcgac ttgctggtaa cgctcagcac ttgctgctg 360

ccctttatgt ccggtttgcc gctggtgacg ctgaccaat ctatggccgc cgtggcgggg 420
 gcgattttcg cccctgttat tgcggcgatg acgctgggtc tgggtgggaac ggatgggttt 480
 gccagcgcgga tcgggcgtaa cgagctcttc aatcatctcg ggaatgcagt ctcgcgcgca 540
 attgccgggc ttctggcgtg gcattacggc ccggtcgtgg tgttctggct gatgggccta 600
 ctggccattg ccggactggt aacggtgctg aggattgata accggcatat agacaatgat 660
 ctggccaacg ggggggaagc caatggaatg gatggggccg ctctgaagga tcaggcatcg 720
 ccaggcctat ggcggtattc ggcgatcat cctggcctga tgacctttgc cgtgctggtg 780

tttttgttcc atctggccaa tgcggccatg cttacctcgg tcagtcagct tctgatgcgg 840
 gaggtcggaa aggatatggc gacctctctg gcggcagcct gtattgttgc cgcccagatg 900
 gtgatggtgc cgggtggcgt gctggtgggg cggtttgtcg atcgatcgg caccgccccg 960
 atctttctgt tggcattcgg tattctggcg ctgcgtggga tgctgtatcc gctttcaggt 1020
 aatccatggt ggctattggg ggtgcaattg ctggatgggg tgggcgccgg gatattcggg 1080
 gccttgttcc cgggtggtgt agccgatctg acgaaggga cgggcgggtt caatatcagc 1140
 cagggcgccg tggcgacggc acagggcatt ggtgctgccc tgagtgccgg tgtggcgggg 1200

ctgattatcg tcaaagccgg ttattccgtg gcgtttttct cgctggctgg tgtggctgga 1260
 ctgggctttg tctgtacgc cctgctgatg cctgaaaccc ggcacctgtg a 1311

<210> 104

<211> 436

<212> PRT

<213> Granulibacter bethesdensis

<400> 104

Met Val Leu Asp Met Asp Ala Met Ser Gly Thr Ile Ala Val Ser Pro

1 5 10 15

Val Leu Ser Asp Gly Gln Pro Pro Ser Asp Ser Arg Gln Lys Gln Gly

20 25 30

Glu Ser Ala Leu Leu Asp Lys Leu Asn Phe Phe Leu Ala Asp Val Arg

35 40 45

Asp Gly Leu Gly Pro Tyr Leu Ala Val Tyr Leu Leu Ser Ala Ser Gly

50 55 60

Arg Asp Gly Arg Trp Asp Glu Ser Ser Val Gly Leu Val Leu Thr Ile

65 70 75 80

Ser Gly Val Val Gly Leu Val Ala Gln Thr Pro Ala Gly Ala Leu Val

85 90 95

Asp Arg Ser Arg Asn Lys Pro Arg Leu Leu Ala Val Ala Ile Leu Leu

100 105 110

Val Thr Leu Ser Thr Leu Leu Leu Pro Phe Met Ser Gly Leu Pro Leu

115 120 125

Val Thr Leu Thr Gln Ser Met Ala Ala Val Ala Gly Ala Ile Phe Ala

130 135 140

Pro Val Ile Ala Ala Met Thr Leu Gly Leu Val Gly Thr Asp Gly Phe

145 150 155 160

Ala Arg Arg Ile Gly Arg Asn Glu Ser Phe Asn His Leu Gly Asn Ala

165 170 175

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

180 185 190

Val Val Phe Trp Leu Met Gly Leu Leu Ala Ile Ala Gly Leu Val Thr

195 200 205

Val Leu Arg Ile Asp Asn Arg His Ile Asp Asn Asp Leu Ala Asn Gly

210 215 220

Gly Glu Ala Asn Gly Met Asp Gly Ala Ala Leu Lys Asp Gln Ala Ser

225 230 235 240

Pro Gly Leu Trp Arg Ile Leu Ala Asp His Pro Gly Leu Met Thr Phe

245 250 255

Ala Val Leu Val Phe Leu Phe His Leu Ala Asn Ala Ala Met Leu Thr

260 265 270

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

275 280 285
Ser Leu Ala Ala Ala Cys Ile Val Ala Ala Gln Met Val Met Val Pro

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

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

420 425 430
Thr Arg His Leu
435

<210> 105

<211> 1263

<212> DNA

<213> Cronobacter turicensis

<400> 105

atgtctgaag tcttacaatc cacacgtca gcgtatataa agtcagctt gtttgtcttt	60
ttatttactt tcacctggtc tgccggtttc ggtatgtatg ctatctggct tggatgataa	120
gcgggactga gtgcaattga tatcggtacg gtttttgccg tcaacggtgt gtttgccgtc	180
ctgggttaac ctgtatatgg ctttatcatg gacaaaaccg gcatgaaaaa gcatttgctc	240
tatttcgtct gccttatttc cgcactgatg gccccttttt atctgtgggtg ctacttgccc	300

ttactacaaa gtcactttct ggcaggcatg atcattggcg cgttatTTTT cagcttcggc 360
 tggtagcgcg gtgtagccgc tgaagagtct tatgtggatc gtttagccg cctgtataat 420
 atggagttcg gtcgtatcag aatgtgggcc gcactgggat gggcggttgc tgcatacgtt 480
 tctggctacg tttataacat ttccccaata attaacTTTA ccatcggcag cgtatctgca 540
 gtgtgtatgt tgctgggtct aatttcactt aaagttgatc atttcgggca ggaagaaaat 600
 agcgttctgt ctaaagataa aattgttatc gcggatgtgt ttcatctgtt cagaaatcca 660
 aaattctgga tgctcgtct gtatatgca ggcgtggcct ggatgatgtt tattgccgaa 720

cagcagtTct cccgttatTT cgtcactttc tttaacacca aagaacaggg taatgccatg 780
 tttggttata tgagcacgac ccagtcgtgt accgaattct ttggcatgat gctggttcct 840
 gcgctggTca accggattgg cgccaaacag gggatgctgc tcacaggcct gatcgtcagt 900
 ttgcgcctta tcatctcggg tctgacaacc gacctctca tcatctgtct tgtaaaaccg 960
 ctttacggga ttgaaatcgc gctgattctg gtctccgtct tcaagtacat tgctgagcac 1020
 ttcgacaaaa ggattaacgc gaccatgtat ctgcttggct atcaggccat ggtttatgtg 1080
 ggctccatcg tggttgcccc tcccgtggc tacgctatg acaccatagg ttttgaacat 1140

acgtatctga ttatgggcac catttacta ctgtttacga ttgtgtcggc tttcacactt 1200
 tctgattgtc aaaagcataa aaaagtcacc tccgctccc atacagatca agcattaaat 1260
 taa 1263

<210> 106

<211> 420

<212> PRT

<213> Cronobacter turicensis

<400> 106

Met Ser Glu Val Leu Gln Ser Thr Arg Ser Ala Tyr Ile Lys Leu Ser

1 5 10 15

Leu Phe Val Phe Leu Phe Thr Phe Thr Trp Ser Ala Gly Phe Gly Met

20 25 30

Tyr Ala Ile Trp Leu Gly Asp Lys Ala Gly Leu Ser Ala Ile Asp Ile

35 40 45

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

50 55 60

Val Tyr Gly Phe Ile Met Asp Lys Thr Gly Met Lys Lys His Leu Leu

65 70 75 80

Tyr Phe Val Cys Leu Ile Ser Ala Leu Met Ala Pro Phe Tyr Leu Trp

85 90 95

Cys Tyr Leu Pro Leu Leu Gln Ser His Phe Leu Ala Gly Met Ile Ile

100 105 110

Gly Ala Leu Phe Phe Ser Phe Gly Trp Tyr Ala Gly Val Ala Ala Glu

115 120 125

Glu Ser Tyr Val Asp Arg Phe Ser Arg Leu Tyr Asn Met Glu Phe Gly

130 135 140

Arg Ile Arg Met Trp Ala Ala Leu Gly Trp Ala Val Ala Ala Ser Phe

145 150 155 160

Ser Gly Tyr Val Tyr Asn Ile Ser Pro Ile Ile Asn Phe Thr Ile Gly

165 170 175

Ser Val Ser Ala Val Cys Met Leu Leu Val Leu Ile Ser Leu Lys Val

180 185 190

Asp His Phe Gly Gln Glu Glu Asn Ser Val Leu Ser Lys Asp Lys Ile

195 200 205

Val Ile Ala Asp Val Phe His Leu Phe Arg Asn Pro Lys Phe Trp Met

210 215 220

Phe Ala Leu Tyr Ile Ala Gly Val Ala Trp Met Met Phe Ile Ala Glu

225 230 235 240

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

245 250 255

Gly Asn Ala Met Phe Gly Tyr Met Ser Thr Thr Gln Ser Cys Thr Glu

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

Leu Tyr Gly Ile Glu Ile Ala Leu Ile Leu Val Ser Val Phe Lys Tyr

325	330	335	
Ile Ala Glu His Phe Asp Lys Arg Ile Asn Ala Thr Met Tyr Leu Leu			
340	345	350	
Gly Tyr Gln Ala Met Val Tyr Val Gly Ser Ile Val Val Ala Pro Pro			
355	360	365	
Ala Gly Tyr Ala Tyr Asp Thr Ile Gly Phe Glu His Thr Tyr Leu Ile			
370	375	380	
Met Gly Thr Ile Ser Leu Leu Phe Thr Ile Val Ser Ala Phe Thr Leu			
385	390	395	400
Ser Asp Cys Gln Lys His Lys Lys Val Thr Ser Ala Ser His Thr Asp			
405	410	415	
Gln Ala Leu Asn			
420			
<210> 107			
<211> 1269			
<212> DNA			
<213> Enterobacter sp			
<400> 107			
atggaaaata catcaacgaa taacagaacc gaatattaca aaatcagtat ttttatcttt	60		
ctctatttct ttacgtggtc ggcaagtatt ggtttgctgg cgatctggct cggacaaaag	120		
gccaatctca gcgggtcggg tatcgggacg gtatttgcgg ttaacggtat tttctccgtg	180		
attttaaaac cgatctacgg ctatattctc gataagatcg ggatgagcaa atacctgctc	240		
tattttgtgg tggcgaatgc ggccctgatg gcaccgttct ttatttatgt ctatcagccg	300		
ctgttaatgt ccaacacgat gctgggtatt attatcggcg cgctctatct aagctttgcc	360		
tggatgcgg gcgtggcagc gtgtgaatcc tattccgacg gtttcagtcg tctaaacggc	420		
atggagtttg gtcagattcg catgtggggt tcacttggct gggcgggtggc gtcgtcgttc	480		
tcgggcctgc tgtttaacct ctgcctcgcg tataatttta ttctgggtag cgtggcgta	540		
gtcatcatgc tcattgttct gctgaacctg aaagtgaaca ctaactccgt tcatgccagt	600		
gacgtactga ccaaagagaa aatcgccctc tcggatgttt acgcgctgct gcgcagccgc	660		
aagttttggg cttctgcct gtatgtcgcg ggcgtggcgt ggatgatgtt tatcgccgag	720		
cagcagttct cgcgctactt cgtcaccttc ttgacgata tccatcaggg caacgcggta	780		

ttcggctacc tcggcacggt tcagtcgggt atggaatttg tcatgtatat ggtgattccg 840
 ctgttcgtga actttatcgg tgccaagcgc ggattactca tcgtcgggtt gttcgtcggg 900
 gcgcgtctga ttatttccgg cctgtgcgat tcccacttat taatttccgt tctcaagcct 960
 ttatacggtc tggaaatctg ttgtctgctg gtgtcgggtat ttaaataatat cgccgagcat 1020
 ttcgataaac gtgtaaafgc caccatgtat ttattgggct atcaggcgat gctgtatgtc 1080
 ggaaatgtgg ttatctcgtc accggcggga tatatgtatg accgcatcgg atttgaacag 1140

acctatatca ttatgggcgc aacggcgcgt acctttacgc tgctctctgc ctttacgtta 1200
 tccgcctgcc aaagcaaagc gcgtggcaca tctgcaatct ctgtacgaga acagaatcca 1260
 tctcgataa 1269

<210> 108

<211> 422

<212> PRT

<213> Enterobacter sp

<400> 108

Met Glu Asn Thr Ser Thr Asn Asn Arg Thr Glu Tyr Tyr Lys Ile Ser

1 5 10 15

Ile Phe Ile Phe Leu Tyr Phe Phe Thr Trp Ser Ala Ser Ile Gly Leu

20 25 30

Leu Ala Ile Trp Leu Gly Gln Lys Ala Asn Leu Ser Gly Ser Val Ile

35 40 45

Gly Thr Val Phe Ala Val Asn Gly Ile Phe Ser Val Ile Leu Lys Pro

50 55 60

Ile Tyr Gly Tyr Ile Leu Asp Lys Ile Gly Met Ser Lys Tyr Leu Leu

65 70 75 80

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

85 90 95

Val Tyr Gln Pro Leu Leu Met Ser Asn Thr Met Leu Gly Ile Ile Ile

100 105 110

Gly Ala Leu Tyr Leu Ser Phe Ala Trp Tyr Ala Gly Val Ala Ala Cys

115 120 125

Glu Ser Tyr Ser Asp Arg Phe Ser Arg Leu Asn Gly Met Glu Phe Gly

130 135 140

Gln Ile Arg Met Trp Gly Ser Leu Gly Trp Ala Val Ala Ser Ser Phe
145 150 155 160

Ser Gly Leu Leu Phe Asn Leu Ser Pro Ala Tyr Asn Phe Ile Leu Gly
165 170 175

Ser Val Ala Ser Val Ile Met Leu Ile Val Leu Leu Asn Leu Lys Val
180 185 190

Asn Thr Asn Ser Val His Ala Ser Asp Val Leu Thr Lys Glu Lys Ile
195 200 205

Ala Pro Ser Asp Val Tyr Ala Leu Leu Arg Ser Arg Lys Phe Trp Ala
210 215 220

Phe Cys Leu Tyr Val Ala Gly Val Ala Trp Met Met Phe Ile Ala Glu
225 230 235 240

Gln Gln Phe Ser Arg Tyr Phe Val Thr Phe Phe Asp Asp Ile His Gln
245 250 255

Gly Asn Ala Val Phe Gly Tyr Leu Gly Thr Val Gln Ser Gly Met Glu
260 265 270

Phe Val Met Tyr Met Val Ile Pro Leu Phe Val Asn Phe Ile Gly Ala
275 280 285

Lys Arg Gly Leu Leu Ile Val Gly Leu Phe Val Gly Ala Arg Leu Ile
290 295 300

Ile Ser Gly Leu Cys Asp Ser His Leu Leu Ile Ser Val Leu Lys Pro
305 310 315 320

Leu Tyr Gly Leu Glu Ile Cys Leu Leu Leu Val Ser Val Phe Lys Tyr
325 330 335

Ile Ala Glu His Phe Asp Lys Arg Val Asn Ala Thr Met Tyr Leu Leu
340 345 350

Gly Tyr Gln Ala Met Leu Tyr Val Gly Asn Val Val Ile Ser Ser Pro
355 360 365

Ala Gly Tyr Met Tyr Asp Arg Ile Gly Phe Glu Gln Thr Tyr Ile Ile
370 375 380

Met Gly Ala Thr Ala Leu Thr Phe Thr Leu Leu Ser Ala Phe Thr Leu

385 390 395 400
 Ser Ala Cys Gln Ser Lys Trp Arg Gly Thr Ser Ala Ile Ser Val Ala
 405 410 415

Glu Gln Asn Pro Ser Arg

420

<210> 109

<211> 1260

<212> DNA

<213> *Citrobacter koseri*

<400> 109

atgaataaaa atgatatgtc tgaagaagac aggcgtaaaa tttatattac gctcagtc	60
tttatcctgg tctacttctt ctcatggaaa gcgacacttg atacttattc tttttggcta	120
tcagaaaaga taggccttga tgggtgtgcg attggtattg tttttgccgc aaacggcttt	180
tgtgcggtac taatcaaacc tgtatatggg tttcttaccg atcggctagg gttaaggaaa	240
gatctgttat tctttatctc gcttatctcg ataaccgttt tccccctttt cttttatata	300
tacaaaccac tattgcagtc tgcgttgtat gctgggtattg cgttgggcgg tatattttta	360
agtatgggat actacgcagg gtgtgcagcc gcagaatcat accttgatcg ctttggccgt	420
ttgtttgatac ttgagtttgg tcaaatccgc atgtggggcg cgattgggtc tgtattctct	480
gctgcaagta caggctatat ttttaatat aacccaatga tcaatttcgg gttaagttcc	540
gcagggtgcg tgataattct tattattgtt ttacgatga agatcgaagt caatgaagag	600
glaaaaaatc gcatcattgc taaagataaa gcgacaatat acgatattat agcattat	660
agcacacgca atttctggac gtttgtttta tatgtttgtg gtgttgtgtg gattat	720
gtcgcagaac aacaatatcc acgtttcttt atcacatttt atgccagtaa agaaatcggg	780
catcaaatgt ttggttatct tgggtctgcc agtctgggag tagagtttct gttcatgatg	840
attgcacctg gcattgtcaa ccgtctgggg ccaaaaaccg gcttgcgtgat tacaggcttt	900
gttgtggccg ctagatttat tggttctggt ttggttacaa gtgcaataac aattggaatt	960
atcaaattgt cgtgggggat agaaatgtca ttctgttgg tttcaatttt caaatatctg	1020
gaaattaact ttgataaaaa agtaaatggt tcaatgtatt tattaggcta tcaaagcgtt	1080
aactatatcg gaacagttat cttgtcgccg ttgagtgggt acttgtatga gcgattaggt	1140
ttttcacata catattttat gatgggatgt acagtattgt ttttactat tatatccgca	1200
tttcttcttc aaaataacaa acgaacagaa ccagtcattg cggagaatgc gcatgcttga	1260

<210> 110

<211> 577

<212> PRT

<213> *Citrobacter koseri*

<400> 110

Met Arg Met Leu Glu Arg Asn Met Leu Arg Glu Asn Ile Ile Asp Lys

1 5 10 15

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

20 25 30

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

35 40 45

Gly Lys Phe Ala Leu Ala Leu Ser Phe Leu Ile Glu Arg Phe Pro Ala

50 55 60

Arg Ala Asp Gln Trp Ser Gly Val Trp His Glu Leu Val Lys Ala Pro

65 70 75 80

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

85 90 95

His Cys Leu Gly Leu Leu Asn Thr Ile Phe Ser Asp Gln Gln Leu Gln

100 105 110

Arg Leu Thr Glu Lys Leu Arg Trp Gln Glu Leu Val Asp Glu Glu Ser

115 120 125

Trp Lys Leu Asn Pro Pro Phe Pro Thr Asn Tyr Tyr Gly Val Ala Phe

130 135 140

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

145 150 155 160

Ser Gln Lys Ile Leu Gln Gln Leu Val Glu His Tyr Arg Thr Tyr Ser

165 170 175

Gln His Gly Cys Val Asp Glu Thr Glu Gly Lys Gly Arg Phe Asp Arg

180 185 190

Tyr Ser Ile Leu Leu Ile Ala Glu Ile Cys Gln Arg His Ile Asp Thr

195 200 205

Gly Leu Glu Ile Cys Asp Trp Leu Lys Asp Ala Leu Arg Gln Val Ser

210	215	220	
Thr Leu Ile Leu Ala Met Leu Asn Asp Glu Gly Val Gly Phe Leu Trp			
225	230	235	240
Gly Arg Ser Ile Gly Ala Tyr Gly Asp Ser Ala Phe Asn Glu Ile Leu			
245	250	255	
Thr Ile Ser Met Arg Leu Gly Leu Leu Asp Arg Gln Glu Cys Ala Ala			
260	265	270	
Ala Ser Arg Phe Thr Leu Ala Cys Ser Glu Arg Phe Leu Asn Phe Trp			
275	280	285	
Tyr Asp Ser Asn Glu Gln Ser Val Asn Leu Trp Phe Tyr Gly Arg Gln			
290	295	300	
Thr Asp Ala Tyr Arg Ala Glu His Arg Leu Val Gly Glu Asn Ile Ser			
305	310	315	320
Leu Ser Cys Gln His Leu Cys Val Gln Arg Ala Trp Ala Asp Val Ser			
325	330	335	
Phe Asp Ala Thr Pro Phe Val Leu Pro Glu Gln Thr Leu Lys Phe Thr			
340	345	350	
Pro Phe Cys Asn Asp Lys Tyr Thr Arg Gly Leu Phe His Trp Tyr Asp			
355	360	365	
Gly Lys Arg Leu Phe Val Leu Pro Leu Ile Asn Gly Asp Lys His Tyr			
370	375	380	
Phe Ala Thr Ser Pro Tyr Phe Pro Val Pro Phe Ser Ala Gly Leu Ile			
385	390	395	400
Thr Gly Val Ala Gln Gly His Ala Pro Leu Trp Val Pro Gly Leu Val			
405	410	415	
Asp Ser Arg Gly Cys Ile Leu Arg Pro Leu Val Trp Phe Gly Asp Cys			
420	425	430	
Gly Tyr Gln Lys Thr Lys Asn Gly Trp Glu Ile Glu Ile Asn Tyr Ser			
435	440	445	
Ala Leu Asn Val Val Met Glu Asn Gly Val Leu Leu Ser Glu Pro Lys			
450	455	460	

Lys Asp Tyr Ser Cys Gln Cys Arg Thr Arg Tyr Phe Ile Glu Pro Ser
 465 470 475 480
 Thr Leu Thr Arg Val Asp Thr Phe Ser Phe Leu Lys His Ser Glu Met

485 490 495
 Tyr Leu Glu Leu Gln Cys Ala Val Phe Pro Glu Lys Met Arg Met Ile
 500 505 510

His Ser Ala Asp Ser Leu His Ile Asp Tyr Glu Ser Asn Gly Ile Gln
 515 520 525
 Ser Leu Glu Leu Asn Gly Phe Glu Asp Tyr Cys Val Glu Arg Thr Ala

530 535 540
 Leu Cys Ser Pro Tyr Gly Ala Leu Gly Gln Gln Ile Thr Gly Arg Arg

545 550 555 560
 Asp Val Ser Gly Ala Lys Asp Val Thr Val Ser Trp Gln Ile Cys Tyr
 565 570 575

Cys

<210> 111

<211> 1242

<212> DNA

<213> Bacillus megaterium

<400> 111

atgaaaagtt caaaaagttt atattggaag ctgagtgcac attttttctt cttctttttt 60
 acgtggctgt ccagctactc tttattttct atctggtag ggcaagagat taagtgaac 120
 ggatcagcaa caggcttaac cttttctgtg aacgctatct ttgctttgtg tatgcagcct 180

ttgtacggat atatatctga caggataggt ttaaagaaac acgttttatt ttttattagc 240
 tgtcttttag tgttttagg accattttat atttttgttt atggaccttt gcttcaatac 300
 aatgtgtga taggggctat tattggaggc ttatatntag gtagtagcgtt ttggcgggc 360
 atcggagcga ttgaaacata catagagaaa gtcagccgaa aatataaatt tgagtatgga 420
 aaatctagaa tgtaggggtc attaggctgg gctgcagcta cgttttttgc aggacaatta 480
 tttaatatta atcctcacat taatttttgg gtggcatctg tttcagcggg tatectagtc 540
 gctattatatt tttctgtaaa agtagaaatg agcagttatg aaatggaaaa agctgaatcc 600

gtcacgctaa gagatgtagg tagtttattt ttattgaagg aattttgggt ttttatgatt 660
 tatgttgttg gtgttacatg cgtatatggg gtatacgatc agcaatttcc tatctattat 720
 gcatctttgt ttcctactga atcgataggt aaccaagtat ttggctactt aaattcgttt 780
 caagtatatt tagaagctgg aatgatgttc gcagcacctg ttattgtaaa taagatcggt 840
 gcaaaaaata gttaatttt agcaggattt ctaatgggat ttcgtattat tgggtcaggg 900
 ctgttgttg gtccaatcgg aatatcttca atgaaactca ttcatgcatt agaacttctt 960
 attatgctca ttgccatctt taaatatttg gcagctaact ttgacacgcg cctatcttcc 1020

atcctttatt tggttggatt ccaatttgct tcgcaaatag gtgcttcagt tctttcaccc 1080
 atagtgggag gattatatga tagcgtaggt ttttccgta cctatttaat tatgggaggt 1140
 atggttttgg tatttaacgt tatttcgatg ttacactac taaactctaa aaaacataga 1200
 ttataagaa aagatgtaca agaaaacacg cagattatat ag 1242

<210> 112

<211> 413

<212> PRT

<213> Bacillus megaterium

<400> 112

Met Lys Ser Ser Lys Ser Leu Tyr Trp Lys Leu Ser Ala Tyr Phe Phe
 1 5 10 15

Phe Phe Phe Phe Thr Trp Ser Ser Ser Tyr Ser Leu Phe Ser Ile Trp
 20 25 30

Leu Gly Gln Glu Ile Lys Leu Asn Gly Ser Ala Thr Gly Leu Ile Phe
 35 40 45

Ser Val Asn Ala Ile Phe Ala Leu Cys Met Gln Pro Leu Tyr Gly Tyr
 50 55 60

Ile Ser Asp Arg Ile Gly Leu Lys Lys His Val Leu Phe Phe Ile Ser
 65 70 75 80

Cys Leu Leu Val Phe Val Gly Pro Phe Tyr Ile Phe Val Tyr Gly Pro
 85 90 95

Leu Leu Gln Tyr Asn Val Val Ile Gly Ala Ile Ile Gly Gly Leu Tyr
 100 105 110

Leu Gly Val Ala Phe Leu Ala Gly Ile Gly Ala Ile Glu Thr Tyr Ile
 115 120 125

Glu Lys Val Ser Arg Lys Tyr Lys Phe Glu Tyr Gly Lys Ser Arg Met
130 135 140

Trp Gly Ser Leu Gly Trp Ala Ala Ala Thr Phe Phe Ala Gly Gln Leu
145 150 155 160

Phe Asn Ile Asn Pro His Ile Asn Phe Trp Val Ala Ser Val Ser Ala
165 170 175

Val Ile Leu Val Ala Ile Ile Phe Ser Val Lys Val Glu Met Ser Ser
180 185 190

Tyr Glu Met Glu Lys Ala Glu Ser Val Thr Leu Arg Asp Val Gly Ser
195 200 205

Leu Phe Leu Leu Lys Glu Phe Trp Phe Phe Met Ile Tyr Val Val Gly
210 215 220

Val Thr Cys Val Tyr Gly Val Tyr Asp Gln Gln Phe Pro Ile Tyr Tyr
225 230 235 240

Ala Ser Leu Phe Pro Thr Glu Ser Ile Gly Asn Gln Val Phe Gly Tyr
245 250 255

Leu Asn Ser Phe Gln Val Phe Leu Glu Ala Gly Met Met Phe Ala Ala
260 265 270

Pro Phe Ile Val Asn Lys Ile Gly Ala Lys Asn Ser Leu Ile Leu Ala
275 280 285

Gly Phe Leu Met Gly Phe Arg Ile Ile Gly Ser Gly Leu Val Val Gly
290 295 300

Pro Ile Gly Ile Ser Ser Met Lys Leu Ile His Ala Leu Glu Leu Pro
305 310 315 320

Ile Met Leu Ile Ala Ile Phe Lys Tyr Leu Ala Ala Asn Phe Asp Thr
325 330 335

Arg Leu Ser Ser Ile Leu Tyr Leu Val Gly Phe Gln Phe Ala Ser Gln
340 345 350

Ile Gly Ala Ser Val Leu Ser Pro Ile Val Gly Gly Leu Tyr Asp Ser
355 360 365

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

370	375	380	
Phe Asn Val Ile Ser Met Phe Thr Leu Leu Asn Ser Lys Lys His Arg			
385	390	395	400
Phe Ile Arg Lys Asp Val Gln Glu Asn Thr Gln Ile Ile			
405	410		
<210>	113		
<211>	1263		
<212>	DNA		
<213>	Providencia rustigianii		
<400>	113		
atgggaaaga taataggagc aactcggcac gcttatttta tccttagcgg ctgctattc	60		
accttttttt ggacatggtc atctgcgttt tcgctggctt ctccttggt cagtcagaag	120		
gtgggcttaa aggggacgga tacgggaatt atcttctcgg tgatcagttt aaccgcattc	180		
tgtgcccagc cactttatgg atttatccaa gataagcttg gattacgcaa aaacttactt	240		
tggtacattg gtgtgatgct ggtgatcagt ggtccatggt ttatttttgt ttgtactcca	300		
ttgttacgtt ggaatatattt cctcggatca atcagtttag gcgtgtatgt cggggcaacg	360		
tttttcgctg ggataggggc tttagaatct tacactgagc gtgtaagtcg aattgcggga	420		
ttcgaatatg gtcgagcgcg catgtgggga tcattgggat gggctggagc gacattcttc	480		
gccggtatct tattcaatat taatccaaat tataatttcg cgatgggacg ggtatccgca	540		
ttaattttca tcattttatt atggagatta agtcatgttc gtcctcaagc gatgaatgag	600		
ctggaatttg gtaaaagtaa tgcactaaaa ctgtcagatg ccttgggggtt actcaaatg	660		
cgttcatattt ggtctctggt cgtttttgtg tgtggcgta gcgtttataa cgtatatgac	720		
caacaattcc caatctattt tgcctcacta ttagtgaac aaagtcattg taatgaaatg	780		
tttggttacc tgaattcact gcaagttttt ctagaagcag gcggcatgtt ttagcccca	840		
ttcttagtca accgaattgg cgccaagaat ggattactcc tcagtggatt tgtaatggcg	900		
ctacgtatct taggttcggg tttagcagat gacaccatta cgatttcttt catgaaatta	960		
cttcacgccg ttgaacttcc tatectactg atatctttat ttaaataat cactacagtt	1020		
ttcgataagc gtttatctgc cacgatttat ctcgtgggct tccagtttat tgcctcactg	1080		
tgcgccacgg tgctatcccc attagctggg atcagttatg accggattgg ctttgcggat	1140		
acctatatga ttttggcggg tattgtcttt acgttaacca ttatttcttg tgtgctgctg	1200		
aaatctgaaa acagctcaaa agttcacaat gtcagtattc agaaaaaac ggtagaacta	1260		

tga 1263

<210> 114

<211> 420

<212> PRT

<213> Providencia rustigianii

<400> 114

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

1 5 10 15

Gly Leu Leu Phe Thr Phe Phe Trp Thr Trp Ser Ser Ala Phe Ser Leu

20 25 30

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

35 40 45

Gly Ile Ile Phe Ser Val Ile Ser Leu Thr Ala Phe Cys Ala Gln Pro

50 55 60

Leu Tyr Gly Phe Ile Gln Asp Lys Leu Gly Leu Arg Lys Asn Leu Leu

65 70 75 80

Trp Tyr Ile Gly Val Met Leu Val Ile Ser Gly Pro Trp Phe Ile Phe

85 90 95

Val Cys Thr Pro Leu Leu Arg Trp Asn Ile Phe Leu Gly Ser Ile Ser

100 105 110

Leu Gly Val Tyr Val Gly Ala Thr Phe Phe Ala Gly Ile Gly Ala Leu

115 120 125

Glu Ser Tyr Thr Glu Arg Val Ser Arg Ile Ala Gly Phe Glu Tyr Gly

130 135 140

Arg Ala Arg Met Trp Gly Ser Leu Gly Trp Ala Gly Ala Thr Phe Phe

145 150 155 160

Ala Gly Ile Leu Phe Asn Ile Asn Pro Asn Tyr Asn Phe Ala Met Gly

165 170 175

Ser Val Ser Ala Leu Ile Phe Ile Ile Leu Leu Trp Arg Leu Ser Asp

180 185 190

Val Arg Pro Gln Ala Met Asn Glu Leu Glu Phe Gly Lys Ser Asn Ala

195 200 205

Leu Lys Leu Ser Asp Ala Leu Gly Leu Leu Lys Met Arg Ser Phe Trp

210

215

220

Ser Leu Val Val Phe Val Cys Gly Val Ser Val Tyr Asn Val Tyr Asp

225

230

235

240

Gln Gln Phe Pro Ile Tyr Phe Ala Ser Leu Phe Ser Glu Gln Ser His

245

250

255

Gly Asn Glu Met Phe Gly Tyr Leu Asn Ser Leu Gln Val Phe Leu Glu

260

265

270

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

275

280

285

Lys Asn Gly Leu Leu Leu Ser Gly Phe Val Met Ala Leu Arg Ile Leu

290

295

300

Gly Ser Gly Leu Ala Asp Asp Thr Ile Thr Ile Ser Phe Met Lys Leu

305

310

315

320

Leu His Ala Val Glu Leu Pro Ile Leu Leu Ile Ser Leu Phe Lys Tyr

325

330

335

Ile Thr Thr Val Phe Asp Lys Arg Leu Ser Ala Thr Ile Tyr Leu Val

340

345

350

Gly Phe Gln Phe Ile Ala Ser Leu Cys Ala Thr Val Leu Ser Pro Leu

355

360

365

Ala Gly Ile Ser Tyr Asp Arg Ile Gly Phe Ala Asp Thr Tyr Met Ile

370

375

380

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

385

390

395

400

Lys Ser Glu Asn Ser Ser Lys Val His Asn Val Ser Ile Gln Lys Lys

405

410

415

Thr Val Glu Leu

420

<210> 115

<211> 1266

<212> DNA

<213> Providencia alcalifaciens

<400> 115

atgggcaaaa aaacacgcgc taacatacat gcctatctca cattaagcgg cttattat	60
acattctttt ggacatggtc atccgtgttt tccttagtct ctctgtggct cagccaaaaa	120
attggcttaa aggttactga taccggcatt attttttcag tcattagtct aacggctttt	180
tgtgcacagc ctctgtatgg atttattcag gacagactcg gtttacgcaa aaatttatta	240
tggtttctgg gtgggttgtt attaatacagt ggaccttggg ttatttttgt ctgtactccg	300
ttattaaaat ggaacatttt tgcgggtgcc gcagcgttag ggatttatgt cggagcgacg	360
ttctttgctg gtatcgggtgc acttgaatcc tacactgaac gggttagccg catcgcgagc	420
tttgaatatg ggctgtcccg tatgtgggga tcgtctgggtt gggcgggcgc cactttcttt	480
gctgggctgt tatttaatat cgatcccaac ttaaatttcg ccatggggtc tgtcagcgca	540
ctgctgttta tcattctgtt atggcgttta cgtaacgttc gtcctcatc tatgaacgaa	600
cttgaattcg gtaaaactca atcgtaact gtctctgatg cgctgggact ttacgtatg	660
cgctcatttt ggtcttttagt aatatttgc tgtggtgtga gtgtttataa cgtctacgac	720
caacaatttc cagtctattt tgccctccctg tttgcgaatc aaaatgaagg caatgagatg	780
ttcggctact taaattcgtc tcaggtcttt ttagaggcag gcgggatgtt ttagcgcct	840
ttccttgtaa atcgatcgg tgctaaacgg ggtctattgt taagtgggtt gatcatggcg	900
ctacgtattt tgggttcagg acttgccgat gacgccatca ctatctcatt catgaagcta	960
ctgcatgccg ttgaacttcc tattctctta atttcattat ttaaatacat tacgtctgta	1020
tttgataaac gcttatccgc aaccatctat ttagtcggct tccagtttac ggcgtcactg	1080
tgtgcttcat tcttatcccc gttagcaggg tacggctatg acacattagg cttcgccgat	1140
acttatatga tctcgcgcgc catcgtcatt tccatgacga taatctctag cgtacttctg	1200
caatcagaaa acacatcaac cgtcagcaac cccatgacaa cgtcgaaaaa aacgggtgga	1260
ttatga	1266

<210> 116

<211> 421

<212> PRT

<213> Providencia alcalifaciens

<400> 116

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

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

Gly Leu Leu Phe Thr Phe Phe Trp Thr Trp Ser Ser Val Phe Ser Leu

20 25 30

Val Ser Leu Trp Leu Ser Gln Lys Ile Gly Leu Lys Gly Thr Asp Thr

35 40 45

Gly Ile Ile Phe Ser Val Ile Ser Leu Thr Ala Phe Cys Ala Gln Pro

50 55 60

Leu Tyr Gly Phe Ile Gln Asp Arg Leu Gly Leu Arg Lys Asn Leu Leu

65 70 75 80

Trp Phe Leu Gly Gly Leu Leu Leu Ile Ser Gly Pro Trp Phe Ile Phe

85 90 95

Val Cys Thr Pro Leu Leu Lys Trp Asn Ile Phe Ala Gly Ala Ala Ala

100 105 110

Leu Gly Ile Tyr Val Gly Ala Thr Phe Phe Ala Gly Ile Gly Ala Leu

115 120 125

Glu Ser Tyr Thr Glu Arg Val Ser Arg Ile Ala Ser Phe Glu Tyr Gly

130 135 140

Arg Ala Arg Met Trp Gly Ser Leu Gly Trp Ala Gly Ala Thr Phe Phe

145 150 155 160

Ala Gly Leu Leu Phe Asn Ile Asp Pro Asn Leu Asn Phe Ala Met Gly

165 170 175

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

180 185 190

Val Arg Pro His Thr Met Asn Glu Leu Glu Phe Gly Lys Thr Gln Ser

195 200 205

Leu Thr Val Ser Asp Ala Leu Gly Leu Leu Arg Met Arg Ser Phe Trp

210 215 220

Ser Leu Val Ile Phe Val Cys Gly Val Ser Val Tyr Asn Val Tyr Asp

225 230 235 240

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

245 250 255

Gly Asn Glu Met Phe Gly Tyr Leu Asn Ser Leu Gln Val Phe Leu Glu

260 265 270
Ala Gly Gly Met Phe Leu Ala Pro Phe Leu Val Asn Arg Ile Gly Ala

275 280 285
Lys Arg Gly Leu Leu Leu Ser Gly Leu Ile Met Ala Leu Arg Ile Leu

290 295 300
Gly Ser Gly Leu Ala Asp Asp Ala Ile Thr Ile Ser Phe Met Lys Leu

305 310 315 320
Leu His Ala Val Glu Leu Pro Ile Leu Leu Ile Ser Leu Phe Lys Tyr

325 330 335
Ile Thr Ser Val Phe Asp Lys Arg Leu Ser Ala Thr Ile Tyr Leu Val

340 345 350
Gly Phe Gln Phe Thr Ala Ser Leu Cys Ala Ser Phe Leu Ser Pro Leu

355 360 365
Ala Gly Tyr Gly Tyr Asp Thr Leu Gly Phe Ala Asp Thr Tyr Met Ile

370 375 380
Leu Ala Ala Ile Val Ile Ser Met Thr Ile Ile Ser Ser Val Leu Leu

385 390 395 400
Gln Ser Glu Asn Thr Ser Thr Val Ser Asn Pro Met Thr Thr Ser Lys

405 410 415
Lys Thr Val Glu Leu

420

<210> 117

<211> 1371

<212> DNA

<213> *Serratia odorifera*

<400> 117

atgttaacgt taacttttgt gattaacatc gcaatccctg atgttggtgg ctttatttgg	60
ccaaagggtta acgttaacaa taaggctgac atcatccatc agggcatctc acatgaaccg	120
cgaacaaaaa aatattatgt gcttctaagc ggctgtgtgt ttttcttctt ttttacctgg	180
tcggccagct tttcactgat ctccatctgg ctaaaccaga aaatcggcct gaaaggcacc	240
gaaaccggcc tgatcttctc ggccatgtcg atcatggcgc tgtgcgtgca gccgctgtac	300

gggtttgttc aggacaaact cgggttgcgt aagcatctgc tgcgtttcgt cggcgtgctg 360
 ttgttgctca cggggccgtt ctttatctat gtttacgccc cgctgctgca aaccaatctg 420
 gtggccggcg cacttggttg cggcctgttc gtcagcctgg cgttcaacgc cggcatcggc 480
 gcgctggaat cciataccga acgcgtcagc cgcacgtctg gcttcgagtt cggccgggcg 540
 cgcacgtggg gctcgtggg ctgggccagc gccaccttct ttgccggctt tatctataac 600
 atcaaccga acatcaactt ttggattgcc tcggcctcgg cggcggtatt cctgctgttg 660
 ctgtggcagg tgcgagagct gaaacccaac gccatggccg ggttgaata cggcaaaccg 720

gaaaacctga agttgcagga cgccctggcg ttgctgcgcc tgccggggtt ctgggcgctg 780
 gtggtgttcg tgcctggcac cagcatctac ggcgtcttcg accagcagtt cccggtgtac 840
 ttcgcttcgc agttcgcgac ccatgaggaa ggcaaccgca tgtacggctt ccttaactcg 900
 ctgcaagtgt ttctggagcg cggtagcatg ttcttgccgc cgctgctggt caaccgtatc 960
 ggtatcaagc aaagcctgct gctggccagc agcgtgatgg cgtttcgcat gttcggctct 1020
 ggtttcgcca atggcgcgct gatgatttcc gccatgaaac tgctgcatgc cgtggaactg 1080
 ccgattctgc tggtagccat gttcaatac atcaccaccc gcttcgacag ccgctgtcc 1140

tccacgtgt atctgggtggg cttccagttt atcagccaga tcgtcgccgg cattctggcg 1200
 ccgctggccg gtctcggtta cgaccgcac ggctttgccg acacctatit gctgatgggc 1260
 tgcgcgtcg ccggtactac gctggtttcc tgcctcttgc tgcgcggcga gactccgcc 1320
 cgcgcacctc aatttcaatc cagataaaa tcaagtgage caatccaatg a 1371

<210> 118

<211> 456

<212> PRT

<213> *Serratia odorifera*

<400> 118

Met Leu Thr Leu Thr Phe Val Ile Asn Ile Ala Ile Pro Asp Val Gly

1 5 10 15

Gly Phe Ile Trp Pro Lys Val Asn Val Asn Asn Lys Ala Asp Ile Ile

20 25 30

His Gln Gly Ile Ser His Glu Pro Arg Asn Lys Lys Tyr Tyr Val Leu

35 40 45

Leu Ser Gly Leu Leu Phe Phe Phe Phe Thr Trp Ser Ala Ser Phe

50 55 60

Ser Leu Ile Ser Ile Trp Leu Asn Gln Lys Ile Gly Leu Lys Gly Thr

65				70				75				80			
Glu	Thr	Gly	Leu	Ile	Phe	Ser	Ala	Met	Ser	Ile	Met	Ala	Leu	Cys	Val
				85				90				95			
Gln	Pro	Leu	Tyr	Gly	Phe	Val	Gln	Asp	Lys	Leu	Gly	Leu	Arg	Lys	His
				100				105				110			
Leu	Leu	Leu	Phe	Val	Gly	Val	Leu	Leu	Leu	Leu	Thr	Gly	Pro	Phe	Phe
				115				120				125			
Ile	Tyr	Val	Tyr	Ala	Pro	Leu	Leu	Gln	Thr	Asn	Leu	Val	Ala	Gly	Ala
130				135				140							
Leu	Val	Gly	Gly	Leu	Phe	Val	Ser	Leu	Ala	Phe	Asn	Ala	Gly	Ile	Gly
145				150				155				160			
Ala	Leu	Glu	Ser	Tyr	Thr	Glu	Arg	Val	Ser	Arg	Ile	Val	Gly	Phe	Glu
				165				170				175			
Phe	Gly	Arg	Ala	Arg	Met	Trp	Gly	Ser	Leu	Gly	Trp	Ala	Ser	Ala	Thr
180				185				190							
Phe	Phe	Ala	Gly	Phe	Ile	Tyr	Asn	Ile	Asn	Pro	Asn	Ile	Asn	Phe	Trp
195				200				205							
Ile	Ala	Ser	Ala	Ser	Ala	Ala	Val	Phe	Leu	Leu	Leu	Leu	Trp	Gln	Val
210				215				220							
Arg	Glu	Leu	Lys	Pro	Asn	Ala	Met	Ala	Gly	Leu	Glu	Tyr	Gly	Lys	Pro
225				230				235				240			
Glu	Asn	Leu	Lys	Leu	Gln	Asp	Ala	Leu	Ala	Leu	Leu	Arg	Leu	Pro	Gly
				245				250				255			
Phe	Trp	Ala	Leu	Val	Val	Phe	Val	Leu	Gly	Thr	Ser	Ile	Tyr	Gly	Val
260				265				270							
Phe	Asp	Gln	Gln	Phe	Pro	Val	Tyr	Phe	Ala	Ser	Gln	Phe	Ala	Thr	His
275				280				285							
Glu	Glu	Gly	Asn	Arg	Met	Tyr	Gly	Phe	Leu	Asn	Ser	Leu	Gln	Val	Phe
290				295				300							
Leu	Glu	Ala	Gly	Gly	Met	Phe	Leu	Ala	Pro	Leu	Leu	Val	Asn	Arg	Ile
305				310				315				320			

Gly Ile Lys Gln Ser Leu Leu Leu Ala Ser Ser Val Met Ala Phe Arg
325 330 335

Met Phe Gly Ser Gly Phe Ala Asn Gly Ala Leu Met Ile Ser Ala Met
340 345 350

Lys Leu Leu His Ala Val Glu Leu Pro Ile Leu Leu Val Ala Met Phe
355 360 365

Lys Tyr Ile Thr Thr Arg Phe Asp Ser Arg Leu Ser Ser Thr Leu Tyr
370 375 380

Leu Val Gly Phe Gln Phe Ile Ser Gln Ile Val Ala Gly Ile Leu Ala
385 390 395 400

Pro Leu Ala Gly Leu Gly Tyr Asp Arg Ile Gly Phe Ala Asp Thr Tyr
405 410 415

Leu Leu Met Gly Cys Ala Val Ala Gly Thr Thr Leu Val Ser Cys Phe
420 425 430

Leu Leu Arg Gly Glu Thr Pro Ala Arg Ala Pro Gln Phe Gln Ser Thr
435 440 445

Ile Lys Ser Ser Glu Pro Ile Gln
450 455

<210> 119

<211> 1314

<212> DNA

<213>

Streptomyces albus

<400> 119

atgaccgtcc agagccctgc ggctccgagc cccgccgcgg ccgtcccgtg gcgcggccgg	60
cgctgcctga acttcggcct gatcagcgcg gccctgttca tgttcttcgt gacctggtcg	120
ctctcctggt cactcttctc catctggctg acccaggaca tcggcctgag cccgggccgc	180
agttcgtcgg tgcacggcgc caactcgtc ggctgcctgg tcacgatgcc ggtctacggc	240
ttcctccagg accggctggg cctgcgcaag aacctgctgt actggatcgg cgccctgatg	300
ctgctggtcg gcccgtgta catctacgtc tacggcccgc tgctcaagtc gcacttcgcc	360
ctcggcctgg tggtcggctc ggccctacac gcgatggcct tcgcggctgc ggtggccacc	420
ctggagagct acgcggagcg gctgagccgg ttccacggat tcgagttcgg ccgggcccgg	480

atgttcggct cgctcggtcg ggccggcgcg accttcttgg ccggccgcct cttaacatc 540
gacccggacc tgaccttctg ggccgcctcc ggcacggcgg tggcttctgt cggactgctg 600
gtggcgatcc gggtagacca cggacggcgg gccgccgcg tcgaccacgc cggcgctcg 660
gtctccctag ccgactgggc cgcgtctctg cgtacccgg ctttctgggg gctgctgctc 720
ttcgtggtcg cggtagacgc gacgtacaac acctacgacg cgaatgtccc ctctacttc 780
tcctcgtctt tcgcgacgga cgccgacggc aaccggatgt acagcgatct caactccgtc 840

caggtcttcc tggaggcggg cgggatggcg ctggcgccgt tcctgggtcaa ccggctcggc 900
ccgaagaagt cctgctggt ctccggctgc gtcacggcga cccggatctt cctctccggc 960
atcgtcacgg acccggtcgc gatctcggcg gtgaagctcc tgcacgccgt cgaactgccc 1020
atcatgctga tcgcgatctt caagtacgtg aaccgccact tcgagccgcg tctctcctcc 1080
accatctacc tgatcggtt ccagatggcg acgcaggtcg gcgccgccgt catctcccc 1140
ttggcgggccc tgggttacga ccgcctcggc tacgccccga cgtacctggt gatgtcggcg 1200
ctggtggcca ccttcacct ggtctcggtc ttaccctgc gcgccgacgc caagggcacg 1260

cacgggctgc ccgcggtggg agccgccgag ggccgctccc cggcagcggg ctga 1314

<210> 120

<211> 437

<212> PRT

<213> Streptomyces albus

<400> 120

Met Thr Val Gln Ser Pro Ala Ala Pro Ser Pro Ala Ala Ala Val Pro

1 5 10 15

Val Arg Gly Arg Arg Arg Leu Asn Phe Gly Leu Ile Ser Ala Ala Leu

20 25 30

Phe Met Phe Phe Val Thr Trp Ser Leu Ser Trp Ser Leu Phe Ser Ile

35 40 45

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

50 55 60

Ile Gly Ala Asn Ser Leu Gly Cys Leu Val Thr Met Pro Val Tyr Gly

65 70 75 80

Phe Leu Gln Asp Arg Leu Gly Leu Arg Lys Asn Leu Leu Tyr Trp Ile

85 90 95

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

100	105	110	
Pro Leu Leu Lys Ser His Phe Ala Leu Gly Leu Val Val Gly Ser Ala			
115	120	125	
Tyr Leu Ala Met Ala Phe Ala Val Ala Val Ala Thr Leu Glu Ser Tyr			
130	135	140	
Ala Glu Arg Leu Ser Arg Phe His Gly Phe Glu Phe Gly Arg Ala Arg			
145	150	155	160
Met Phe Gly Ser Leu Gly Trp Ala Ala Ala Thr Phe Leu Ala Gly Arg			
165	170	175	
Leu Phe Asn Ile Asp Pro Asp Leu Thr Phe Trp Ala Ala Ser Gly Thr			
180	185	190	
Ala Val Val Phe Val Gly Leu Leu Val Ala Ile Arg Val Thr Asp Gly			
195	200	205	
Arg Arg Ala Ala Ala Val Asp His Ala Ala Ala Ser Val Ser Leu Ala			
210	215	220	
Asp Val Ala Ala Leu Leu Arg Tyr Pro Ala Phe Trp Gly Leu Leu Leu			
225	230	235	240
Phe Val Val Ala Val Thr Ala Thr Tyr Asn Thr Tyr Asp Ala Met Phe			
245	250	255	
Pro Ser Tyr Phe Ser Ser Leu Phe Ala Thr Asp Ala Asp Gly Asn Arg			
260	265	270	
Met Tyr Ser Asp Leu Asn Ser Val Gln Val Phe Leu Glu Ala Gly Gly			
275	280	285	
Met Ala Leu Ala Pro Phe Leu Val Asn Arg Leu Gly Pro Lys Lys Ser			
290	295	300	
Leu Leu Val Ser Gly Cys Val Met Ala Thr Arg Ile Phe Leu Ser Gly			
305	310	315	320
Ile Val Thr Asp Pro Val Ala Ile Ser Ala Val Lys Leu Leu His Ala			
325	330	335	
Val Glu Leu Pro Ile Met Leu Ile Ala Ile Phe Lys Tyr Val Asn Arg			
340	345	350	

His Phe Glu Pro Arg Leu Ser Ser Thr Ile Tyr Leu Ile Gly Phe Gln
355 360 365

Met Ala Thr Gln Val Gly Ala Ala Val Ile Ser Pro Leu Ala Gly Leu
370 375 380

Gly Tyr Asp Arg Leu Gly Tyr Ala Pro Thr Tyr Leu Val Met Ser Ala
385 390 395 400

Leu Val Ala Thr Phe Thr Leu Val Ser Val Phe Thr Leu Arg Ala Asp
405 410 415

Ala Lys Gly Thr His Gly Leu Pro Ala Val Gly Ala Ala Glu Gly Arg
420 425 430

Ser Pro Ala Ala Gly
435

<210> 121

<211> 1284

<212> DNA

<213> Azobactervinelandii

<400> 121

atgcatctcc cgtaaaccg cgaatactgg ctgatcagtg gcctgctgtt tttcttcttc 60
ttcgcttggc cgtcgagcta ttcgctgttc tcgatctggc tgcaccgggt catcggtctg 120
agcggtagcg aaaccggctt catcttcgcc gccaacgcca tcgccgcgct gttcatccag 180
cccttctacg gcgccctgca ggaccgcctg ggctgtccc gcaggctgct ggtctggatc 240
ggcgtgctgc tcgcctgcgc cgcgcccttc gcgatccacg tctacgccgg cctgctggcc 300
cacgacgtgg tgctcggcgc cctggtcggc gccgccttcc tggccttcgc catgctcgcc 360

ggcgtgggcg tcacgagtc ctataccgag cggtgtgc gccacgccgg cttcagttc 420
ggcaccacgc ggatgtgggg ctgctgggc tgggcggccg cgacgggctg ggccggcgtg 480
gtgttcgaca tcgaccgga catgccttc tacatgagca gcgccgccgg cctggcgctc 540
ctgctgatcc tctcccgtt ggacatggac gccttcagga accacgagac gcgggccggc 600
gagtcgccc cggtacgcac cgccgacatc ctgcaactgc tgcgcctgcc gcgcttctgg 660
gcgttcagcc tgttctgac cggggtcgcc ggggtctaca tgatctacga gcagcagttc 720
ccggtgtact tcgcctcctt ctccctacc ccggaagaag gcaccgcgc ctacggctac 780

ctgaactcct cgcaggtaact ggtcgaggcc ctgctcatgc tggtcgcgcc ctgggtggtc 840

gccccgaccg gcgccaata cggcctgctg ctggccggca ccatcatgtt cgtgcgcatac 900
ctcggctcgg ggctggccag cgatgcctgg accatcgccg cctgcaagat gctgcatgcc 960
atcgaggtgc cgatcctgtt ggctcggtg ttcaagtaca tctgcgtcaa cttcgatgcg 1020
cggtgtcgg ccacgtcta cctggtcagt ttccagttcg cccagcaact gaccgccaatg 1080
ctgttgctgc cgcccggtg ctatggctac gacaggtcgc gcttcgccag cgtctacctg 1140
ctgctggcca gtctggtcgg cgctgcctg ctgctgtcct ggctcgtgct gcgcgccgac 1200

ccgtccagca agaccacccg ggccggcgac ggcccccgg aactgccggc catcgcccag 1260
gcggccaacc attacgaacc ctga 1284

<210> 122

<211> 427

<212> PRT

<213> Azobacter vinelandii

<400> 122

Met His Leu Pro Leu Lys Arg Glu Tyr Trp Leu Ile Ser Gly Leu Leu

1 5 10 15

Phe Phe Phe Phe Phe Ala Trp Ser Ser Ser Tyr Ser Leu Phe Ser Ile

20 25 30

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

35 40 45

Phe Ala Ala Asn Ala Ile Ala Ala Leu Phe Ile Gln Pro Phe Tyr Gly

50 55 60

Ala Leu Gln Asp Arg Leu Gly Leu Ser Arg Arg Leu Leu Val Trp Ile

65 70 75 80

Gly Val Leu Leu Ala Cys Ala Ala Pro Phe Ala Ile His Val Tyr Ala

85 90 95

Gly Leu Leu Ala His Asp Val Val Leu Gly Ala Leu Val Gly Ala Ala

100 105 110

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

115 120 125

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

130 135 140

Met Trp Gly Ser Leu Gly Trp Ala Ala Ala Thr Gly Val Ala Gly Val

145 150 155 160
 Val Phe Asp Ile Asp Pro Asp Ile Ala Phe Tyr Met Ser Ser Ala Ala

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

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

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

 355 360 365
 Gly Tyr Asp Arg Leu Gly Phe Ala Ser Val Tyr Leu Leu Leu Ala Ser
 370 375 380
 Leu Val Gly Ala Cys Leu Leu Leu Ser Trp Ser Leu Leu Arg Ala Asp
 385 390 395 400

Pro Ser Ser Lys Thr Thr Arg Ala Ala Asp Gly Ala Pro Glu Leu Pro

405

410

415

Ala Ile Ala Gln Ala Ala Asn His Tyr Glu Pro

420

425

<210> 123

<211> 1278

<212> DNA

<213> Oenococcus oeni

<400> 123

ttgatcatga ttaaccaatc tgaaattaaa actgctaata agaaaaaaga tttctgggga 60

tttccattct ctactttac ctatTTTTTT atttggcaaa tcattaacgg ttatttaact 120

ctatggatgg aacagatagg acacctaaat ggtgcagaag ctggagtgtg attttctatg 180

atggctggat tttctctcat ttttcaacca tcatttggaa ttatttccga taaattaatt 240

tttaagaaaa atttgattct gacaatttca atcgctggaa tcttgattgg accatatTTT 300

cagtggcttt tcttaccgtt aatgaaaata aactcaacat tcgttgctat tattactggt 360

atTTTTTTaa gtttcatttt gaatgggtggc gtttagcgtaa tagaacaata tgttcaaaga 420

gcaagtctcg ctaaccattt tgaatcggga cattcaagaa tgggtggatc tttggcagga 480

atgtgtgcct cttttcttgg tgggcaacta tttttgtgga gtcctaattc tattttttgg 540

gcctgcacta tttcagcaat gatccttaca ggcttggtat tatttaatga caagatccat 600

atggaaaatg caaatcttgc cgggtggaact tccagtcgtt taaacctaaa acagatctct 660

tcaatattta aaataaagaa tttttggttc ctctcactat tttatattgg cacggcctct 720

atttacgata tctttgatca gcaatttatt atatttttta aatctTTTTT tgataccgca 780

tcacaaggaa cagctgccta tagttatatg acaactgcgc agattggtat cgaatttata 840

cttatgtttc caatgccttg gataattaat cgaattggag cgcgtaatgg cctaatagca 900

tatggaacaa ttttagctat tagaattatt ggtagtgcac tatctccaaa tcttttttgg 960

gttattttgc tacgcttgct ggctggatta gaaatgcctc ttatcctagt ttcaatcatg 1020

aaatatatat cgggtgcctt tgaccttcgc ctttatgcca ctatatatgc gctgtcatca 1080

aactttgcaa aacaaataac tgaattcttt ttagcaactg ctgcaggaaa aatgtatgat 1140

tcaatcggct ttaccatac ttatttaatt cttggagtca tcgtcgcaat ctttacgatt 1200

tttacggcac tttttcttaa aaaagaaaac cctgttcaag caggagaaag agacagagct 1260

ggaaacccca aagcttaa 1278

<210> 124

<211> 425

<212> PRT

<213> *Oenococcus oeni*

<400> 124

Met Ile Met Ile Asn Gln Ser Glu Ile Lys Thr Ala Asn Lys Lys Lys

1 5 10 15

Asp Phe Trp Gly Phe Pro Phe Ser His Phe Thr Tyr Phe Phe Ile Trp

20 25 30

Gln Ile Ile Asn Gly Tyr Leu Thr Leu Trp Met Glu Gln Ile Gly His

35 40 45

Leu Asn Gly Ala Glu Ala Gly Val Val Phe Ser Met Met Ala Gly Phe

50 55 60

Ser Leu Ile Phe Gln Pro Ser Phe Gly Ile Ile Ser Asp Lys Leu Ile

65 70 75 80

Phe Lys Lys Asn Leu Ile Leu Thr Ile Ser Ile Ala Gly Ile Leu Ile

85 90 95

Gly Pro Tyr Phe Gln Trp Leu Phe Leu Pro Leu Met Lys Ile Asn Ser

100 105 110

Thr Phe Val Ala Ile Ile Thr Gly Ile Phe Leu Ser Phe Ile Leu Asn

115 120 125

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

130 135 140

Asn His Phe Glu Tyr Gly His Ser Arg Met Gly Gly Ser Leu Ala Gly

145 150 155 160

Met Cys Ala Ser Phe Leu Gly Gly Gln Leu Phe Leu Trp Ser Pro Asn

165 170 175

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

180 185 190

Leu Leu Phe Asn Asp Lys Ile His Met Glu Asn Ala Asn Leu Ala Gly

195 200 205

Gly Thr Ser Ser Arg Leu Asn Leu Lys Gln Ile Ser Ser Ile Phe Lys

210 215 220
 Ile Lys Asn Phe Trp Phe Leu Ser Leu Phe Tyr Ile Gly Thr Ala Ser

 225 230 235 240
 Ile Tyr Asp Ile Phe Asp Gln Gln Phe Ile Ile Phe Phe Lys Ser Phe

 245 250 255
 Phe Asp Thr Ala Ser Gln Gly Thr Ala Ala Tyr Ser Tyr Met Thr Thr

 260 265 270
 Ala Gln Ile Gly Ile Glu Phe Ile Leu Met Phe Pro Met Pro Trp Ile

 275 280 285
 Ile Asn Arg Ile Gly Ala Arg Asn Gly Leu Ile Ala Tyr Gly Thr Ile

 290 295 300
 Leu Ala Ile Arg Ile Ile Gly Ser Ala Leu Ser Pro Asn Leu Phe Trp
 305 310 315 320
 Val Ile Leu Leu Arg Leu Leu Ala Gly Leu Glu Met Pro Leu Ile Leu

 325 330 335
 Val Ser Ile Met Lys Tyr Ile Ser Gly Ala Phe Asp Leu Arg Leu Tyr

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

 355 360 365
 Phe Phe Leu Ala Thr Ala Ala Gly Lys Met Tyr Asp Ser Ile Gly Phe

 370 375 380
 His His Thr Tyr Leu Ile Leu Gly Val Ile Val Ala Ile Phe Thr Ile
 385 390 395 400
 Phe Thr Ala Leu Phe Leu Lys Lys Glu Asn Pro Val Gln Ala Gly Glu

 405 410 415
 Arg Asp Arg Ala Gly Asn Pro Lys Ala

 420 425

<210> 125

<211> 1290

<212> DNA

<213> Lactobacillus brevis

<400> 125

atggatagtg caaaaactga atcgaacat ttctggggat ttccctatc tcaacttagc 60
tactttctca tttgggccgt tgttaacggt tatttaaccc tctggatgga acaagttgcc 120
cacttaaacg ggaccgaatc tggtgccgtc ttctcaatga tggcgggtat ttcttgate 180
ttccaaccta tctttggggt tgtttctgat aaattactct tcaagaaaaa cttaacttta 240
acgattgcga tcgtcgggat cttcatcggt ccttacttcc aatgggtctt catgcccatc 300
ttacacatta gcgccttctt agtagccatc gttactggga cttttctgag ttcatccta 360

aacggtgggg tgtccgtcat tgaacaatac gtacagcggg ctagtttgge taacggcttc 420
gaatatgccc actcccggtt cgggtggttc gtcgccggtg tcgttgcttc attagtggc 480
gggcgaatct tcctctggaa gcctaactcc atcttctggg catgtacat tgccgctatt 540
atcttaactt tcttattact cttctctgac aaggttaact tagataacgc tgccgctgct 600
ggtgacacgt ccgattcact tgatatgaag acggttctgt ctgtcttcaa acttaagaac 660
ctctgggttt tagctatttt ctacatgggc gcttctgcct tattcgatgt attgatcaa 720
cagtttatca tcttctttaa gactttcttt gataccgctg ctcaggggac cctggtttac 780

tcttacatga gtctgctca aacagctatc gaattttgtt taatgttccc aatgccatgg 840
attatcaaca aaattgggtc tcgtaatggt ttgttaattt acggtttcat tacctgcatc 900
cggatcttag gtctgcttt atctcctaac tggatctggg ttgttggttt ccgtttgtta 960
gctgggttgg aaatgccatt gctattaact tccatcatga aatacattgc tgggtgccttt 1020
gatattcggg tatacggcac gggttacgca ttggcgtaaa actttgctaa acaggtatcc 1080
gtcttcatct tcctcacagt tgctggtcgg atgtatgacg ttatcgggta tcaacacacg 1140
tacatcatca tggggatcgt ggtcttcttc attacctgt ttgccgtctt cttctgaag 1200

aaagaagatc cgattcaagc cggatgaagtc gaggatccta atgttaaagc taaagtggaa 1260
gcggcttctt caactgatag tcgtgaatga 1290

<210> 126

<211> 429

<212> PRT

<213> Lactobacillus brevis

<400> 126

Met Asp Ser Ala Lys Thr Glu Ser Lys His Phe Trp Gly Phe Pro Leu

1 5 10 15

Ser His Phe Ser Tyr Phe Phe Ile Trp Ala Val Val Asn Gly Tyr Leu

20 25 30

Thr Leu Trp Met Glu Gln Val Ala His Leu Asn Gly Thr Glu Ser Gly

35 40 45

Ala Val Phe Ser Met Met Ala Gly Ile Ser Leu Ile Phe Gln Pro Ile

50 55 60

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

65 70 75 80

Thr Ile Ala Ile Val Gly Ile Phe Ile Gly Pro Tyr Phe Gln Trp Val

85 90 95

Phe Met Pro Ile Leu His Ile Ser Ala Phe Leu Val Ala Ile Val Thr

100 105 110

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

115 120 125

Gln Tyr Val Gln Arg Ala Ser Leu Ala Asn Gly Phe Glu Tyr Ala His

130 135 140

Ser Arg Val Gly Gly Ser Val Ala Gly Val Val Ala Ser Leu Val Ala

145 150 155 160

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

165 170 175

Ile Ala Ala Ile Ile Leu Thr Phe Leu Leu Leu Phe Ser Asp Lys Val

180 185 190

Asn Leu Asp Asn Ala Ala Ala Ala Gly Asp Thr Ser Asp Ser Leu Asp

195 200 205

Met Lys Thr Val Leu Ser Val Phe Lys Leu Lys Asn Leu Trp Val Leu

210 215 220

Ala Ile Phe Tyr Met Gly Ala Ser Ala Leu Phe Asp Val Phe Asp Gln

225 230 235 240

Gln Phe Ile Ile Phe Phe Lys Thr Phe Phe Asp Thr Ala Ala Gln Gly

245 250 255

Thr Leu Val Tyr Ser Tyr Met Ser Ser Ala Gln Thr Ala Ile Glu Phe

260 265 270

Cys Leu Met Phe Pro Met Pro Trp Ile Ile Asn Lys Ile Gly Ser Arg

275 280 285
Asn Gly Leu Leu Ile Tyr Gly Phe Ile Thr Cys Ile Arg Ile Leu Gly

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

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

420 425
<210> 127
<211> 1245
<212> DNA
<213> Weissella paramesenteroides
<400> 127

atggatcaaa cacaaaatac acaaaaaaag cacttttggg gtttttttgg aactgatgtc 60
agttatttct ttatttggca aattgtacaa ggtttcttgg tactttgggtt aaaacaagaa 120
gccgatttga gtggtggcca agccggattt gttttttcat ttttggcatt tgcctcatta 180
ttttatcaac cagtatttgg tataatttcg gataaattgg tctttaaaaa gacgttgcta 240
ttatcaataa caattgctgc tatcctcatt ggcccathtt tccaatgggtt attcatacca 300

ttacttcatt ttagttcatt attaggggcg attattgggtg gcttcgcatt agcttatgta 360
cttaatgggtg gtgtggcagt tattgaatct tatgtggaac gcgcaagctt agccaatggc 420

tttgaatatg gtcatgcccc tatgggagga tcaattgggtg gtgcactggc atcgttcgtg 480
gcaggattgg tgtttgttaa aaatccattt ttggttttct ggatttgctc acttgctggt 540
attattttga cttgcttgggt aatcttttgg gatcggatta attttgctaa tgcagatgcg 600
gcagaaggag caagttcagc accattaaat gccaaaacta ttttatctat ttttaaaatt 660
aaaaatttct ggatattaag tattttcttt atgggaacat cagcaattta tgatgtgttt 720

gatcaacagt tcccgttatt ttatcaaaca ttcttcacga gtgctgcagc tgggtaccgta 780
gcctacagcc gcctactaac gacacaaatt gcgttagaag ccctattaat gattccaatg 840
ccttgataa ttaataaaat tgggtgctaag aatggattga ttgcttatgg tgtgcttact 900
tttctacgta ttacgttaag cgcaattgct cctaatttct ggttcttaac attgttctgt 960
ctacttgcat catttgaaat gccactcttc ttagtctcaa caatgaagta tattgccgga 1020
gcatttgatt taagattata cgcaaccatt tatgctttgg cgttcaattt tgctaaacag 1080
atatcgttat ttattttctc tacagttgcc ggtgatcttt atgatagcgt gggattccat 1140

acgacatatt tcactcttagc tgcattagtc gcttttgtaa caatcgttgc tgtatttatg 1200
ttgaagaaag aaaatccagc tcaggcgggt gaagtcgcgg agtga 1245

<210> 128

<211> 414

<212> PRT

<213> Weissella paramesenteroides

<400> 128

Met Asp Gln Thr Gln Asn Thr Gln Lys Lys His Phe Trp Ser Phe Phe

1 5 10 15

Gly Thr Asp Val Ser Tyr Phe Phe Ile Trp Gln Ile Val Gln Gly Phe

20 25 30

Leu Val Leu Trp Leu Lys Gln Glu Ala Asp Leu Ser Gly Gly Gln Ala

35 40 45

Gly Phe Val Phe Ser Phe Leu Ala Phe Ala Ser Leu Phe Tyr Gln Pro

50 55 60

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

65 70 75 80

Leu Ser Ile Thr Ile Ala Ala Ile Leu Ile Gly Pro Phe Phe Gln Trp

85 90 95

Leu Phe Ile Pro Leu Leu His Phe Ser Ser Leu Leu Gly Ala Ile Ile

100 105 110

Gly Gly Phe Ala Leu Ala Tyr Val Leu Asn Gly Gly Val Ala Val Ile

115 120 125

Glu Ser Tyr Val Glu Arg Ala Ser Leu Ala Asn Gly Phe Glu Tyr Gly

130 135 140

His Ala Arg Met Gly Gly Ser Ile Gly Gly Ala Leu Ala Ser Phe Val

145 150 155 160

Ala Gly Leu Val Phe Val Lys Asn Pro Phe Leu Val Phe Trp Ile Cys

165 170 175

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

180 185 190

Ile Asn Phe Ala Asn Ala Asp Ala Ala Glu Gly Ala Ser Ser Ala Pro

195 200 205

Leu Asn Ala Lys Thr Ile Leu Ser Ile Phe Lys Ile Lys Asn Phe Trp

210 215 220

Ile Leu Ser Ile Phe Phe Met Gly Thr Ser Ala Ile Tyr Asp Val Phe

225 230 235 240

Asp Gln Gln Phe Pro Val Phe Tyr Gln Thr Phe Phe Thr Ser Ala Ala

245 250 255

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

260 265 270

Glu Ala Leu Leu Met Ile Pro Met Pro Trp Ile Ile Asn Lys Ile Gly

275 280 285

Ala Lys Asn Gly Leu Ile Ala Tyr Gly Val Leu Thr Phe Leu Arg Ile

290 295 300

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

305 310 315 320

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

325 330 335

Tyr Ile Ala Gly Ala Phe Asp Leu Arg Leu Tyr Ala Thr Ile Tyr Ala

340 345 350
Leu Ala Phe Asn Phe Ala Lys Gln Ile Ser Leu Phe Ile Phe Ser Thr

355 360 365
Val Ala Gly Asp Leu Tyr Asp Ser Val Gly Phe His Thr Thr Tyr Phe

370 375 380
Ile Leu Ala Ala Leu Val Ala Leu Val Thr Ile Val Ala Val Phe Met

385 390 395 400
Leu Lys Lys Glu Asn Pro Ala Gln Ala Gly Glu Val Ala Glu

405 410

<210> 129

<211> 1278

<212> DNA

<213> Mannheimia succinicproduces

<400> 129

gtgcttatga caagccaaaa taagataaat gcggttcctt caaatcaaaa tttttatttg 60
aacaaccgta actattggat ctttagcggg tactttttcg tctatttctt tattatggcc 120
acatgttate ctttcttagg catttggttg ggagatatta acggtctgtc gggagaagat 180
cgaggcacag tatttgctat gatgtctttc ttgccccttt gctttcagcc tgtttttggc 240
tatgtgtcag ataaattggg attgaaaaaa catctcctct gggtattagg tataagcctt 300
ttaatctatg cacccttttt tatctataatt ttgacaccgt tgcctaaagt gaatgtttgg 360
ttaggttcct tgggtggcgg cgcctatac ggctttgttt tccaggcggg agcaccggct 420

tcggaagctt atattgagcg ggtcagccgt cgtagtaaat ttgaatacgg gcgggtcaga 480
atgttcggaa tgttcggctg ggctatttgc gcctccatag ccggagtcct ttacgtacc 540
aatcctaate tggctctttg gcttggttct atagcctctc taatcctttt acttttgata 600
gccctagcta agccggaaca aaccagcact gtgcaaatcg cggagaaatt aggggcta 660
aaaaatccgg ttaacctcag acaggcattc gctctcctca aactaccgaa attctgggcg 720
cttctggctt atgtaattgg aattgcctgt gtttacgaca tcttcgacca acaattcggt 780
aatTTTTTTT atacctTTTT tgaatcgac gaacaaggaa tcaaatgtt tggttatgta 840

accactgccg gtgaattgtt aaacgccctg attatgttct ttgttccct tatcattaac 900
cgtatcgggt ctaaaaatgc tctgttatt gcgggaacca ttatgagcgt gcggatcac 960
ggctcttctt atgcaatcga ggcctggcat gtggtagttt taaaaactct gcatagttt 1020

gaagttcctt ttacctagt aggactatit aaatacattg ctaatgtgtt tgaagtiacat 1080
 ttttctgcga ccatttatct ggtagcctgc cactttgcc aacagatcgg caatatgctg 1140
 gtatccccctt tggtaggcgc ttggtacgat acctatgggt tccaagatac ttacctaatt 1200
 ctaggctgca ttgccgccgg attcacctg ctttcgggtt ttacacttac cggcaaatcc 1260

ttatcgtctc aatcataa 1278

<210> 130

<211> 425

<212> PRT

<213> Mannheimia succiniciproducens

<400> 130

Met Leu Met Thr Ser Gln Asn Lys Ile Asn Ala Val Pro Ser Asn Gln

1 5 10 15

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

20 25 30

Phe Val Tyr Phe Phe Ile Met Ala Thr Cys Tyr Pro Phe Leu Gly Ile

35 40 45

Trp Leu Gly Asp Ile Asn Gly Leu Ser Gly Glu Asp Arg Gly Thr Val

50 55 60

Phe Ala Met Met Ser Phe Phe Ala Leu Cys Phe Gln Pro Val Phe Gly

65 70 75 80

Tyr Val Ser Asp Lys Leu Gly Leu Lys Lys His Leu Leu Trp Val Leu

85 90 95

Gly Ile Ser Leu Leu Ile Tyr Ala Pro Phe Phe Ile Tyr Ile Phe Ala

100 105 110

Pro Leu Leu Lys Val Asn Val Trp Leu Gly Ser Leu Val Gly Gly Ala

115 120 125

Tyr Ile Gly Phe Val Phe Gln Ala Gly Ala Pro Ala Ser Glu Ala Tyr

130 135 140

Ile Glu Arg Val Ser Arg Arg Ser Lys Phe Glu Tyr Gly Arg Val Arg

145 150 155 160

Met Phe Gly Met Phe Gly Trp Ala Ile Cys Ala Ser Ile Ala Gly Val

165										170										175									
Leu	Tyr	Ala	Thr	Asn	Pro	Asn	Leu	Val	Phe	Trp	Leu	Gly	Ser	Ile	Ala														
180					185					190																			
Ser	Leu	Ile	Leu	Leu	Leu	Leu	Ile	Ala	Leu	Ala	Lys	Pro	Glu	Gln	Thr														
195					200					205																			
Ser	Thr	Val	Gln	Ile	Ala	Glu	Lys	Leu	Gly	Ala	Asn	Lys	Asn	Pro	Val														
210					215					220																			
Asn	Leu	Arg	Gln	Ala	Phe	Ala	Leu	Leu	Lys	Leu	Pro	Lys	Phe	Trp	Ala														
225					230					235					240														
Leu	Leu	Ala	Tyr	Val	Met	Gly	Ile	Ala	Cys	Val	Tyr	Asp	Ile	Phe	Asp														
245					250					255																			
Gln	Gln	Phe	Gly	Asn	Phe	Phe	Asn	Thr	Phe	Phe	Glu	Ser	His	Glu	Gln														
260					265					270																			
Gly	Ile	Lys	Met	Phe	Gly	Tyr	Val	Thr	Thr	Ala	Gly	Glu	Leu	Leu	Asn														
275					280					285																			
Ala	Leu	Ile	Met	Phe	Phe	Val	Pro	Leu	Ile	Ile	Asn	Arg	Ile	Gly	Ala														
290					295					300																			
Lys	Asn	Ala	Leu	Leu	Ile	Ala	Gly	Thr	Ile	Met	Ser	Val	Arg	Ile	Ile														
305					310					315					320														
Gly	Ser	Ser	Tyr	Ala	Ile	Glu	Ala	Trp	His	Val	Val	Val	Leu	Lys	Thr														
325					330					335																			
Leu	His	Met	Phe	Glu	Val	Pro	Phe	Tyr	Leu	Val	Gly	Leu	Phe	Lys	Tyr														
340					345					350																			
Ile	Ala	Asn	Val	Phe	Glu	Val	His	Phe	Ser	Ala	Thr	Ile	Tyr	Leu	Val														
355					360					365																			
Ala	Cys	His	Phe	Ala	Lys	Gln	Ile	Gly	Asn	Met	Leu	Val	Ser	Pro	Leu														
370					375					380																			
Val	Gly	Ala	Trp	Tyr	Asp	Thr	Tyr	Gly	Phe	Gln	Asp	Thr	Tyr	Leu	Ile														
385					390					395					400														
Leu	Gly	Cys	Ile	Ala	Ala	Gly	Phe	Thr	Leu	Leu	Ser	Val	Phe	Thr	Leu														
405					410					415																			

Thr Gly Lys Ser Leu Ser Ser Gln Ser

420

425

<210> 131

<211> 1257

<212> DNA

<213> Aggregatibacter aphrophilus

<400> 131

atgagtaatc aaacttcac taaatcgcaa tacttaacca atagtaatta ttggattttt	60
agtgcgtatt tttttgcttt tttcttcatt atggcgacct gtcacccctt tttagggtatt	120
tggctcgggtg atattcacgg attaaaaggg gagaaaattg gctacgtctt ttcgttcatt	180
tctctttttg ctttactctt tcagcctatt ttaggatttt tatctgataa attaggtatt	240
cgaaaacatc tgctttggct acttgcgata ttgcttcttt tctatgcacc attctttatt	300
tatgtttttg ctccgttatt gaaaactaat ttatggttgg gggttattgc cgggtgtgca	360
tatatgggat ttgtgttcca agcaggcgca cctgcgtcag aagcctacat tgagcgtatt	420
agccgtttag atggatttga atatggcaga acacggctat tcgggatgtt aggttgggca	480
atctgtgcct ctattgcagg caatttataat agttctcaac cgaatgcggt tttctggctt	540
ggctccgcca ccgcccgtgt gttattgggt ctgattttcc ttgcgaaaac agatagcagt	600
aatacggctc aagtcgttga taaattaggt gtaaataaaa gtccaattac actaaaacaa	660
gccttaaagc tcttttctct acctcgtttc tgggcactat taacctatgt tgttggagtg	720
gcttgtgttt atgacatttt tgaccagcaa ttcggtaatt tcttcaatac gttttttgaa	780
tctaaagaac aaggaatgaa attcttttgg tttgttaacaa cgggaggaga gctattaaat	840
gcaacaatta tgtttttcgt gcctttattg attaatcgta ttggggcgaa gaatgcgtta	900
ttgattgcag gatcaattat gaggatttgc attatgggct catcttttgc aacagaagca	960
tggcacgtga ttgtattaaa aacattacat atgtttgaag tgccattcta ttgtgttggga	1020
gtatttaagt atattgctga tgtttttgaa gtccgttttt ctgcaacat ttacttagtc	1080
tcttgccact ttcaaagca gattggtaat atgattttat cgcctgccgt agggacttta	1140
tatgacatgt atggttttca atctacctat tttatcttag gttgtattgc tcttacattt	1200
acatcagttt ctgtatttac tttagtaaat accaagaaat tagtgaacaa cgcttaa	1257

<210> 132

<211> 417

<212> PRT

<213> Aggregatibacter aphrophilus

<400> 132

Met Ser Asn Gln Thr Ser Ser Lys Ser Gln Tyr Leu Thr Asn Ser Asn

1 5 10 15

Tyr Trp Ile Phe Ser Ala Tyr Phe Phe Ala Phe Phe Phe Ile Met Ala

20 25 30

Thr Cys His Pro Phe Leu Gly Ile Trp Leu Gly Asp Ile His Gly Leu

35 40 45

Lys Gly Glu Lys Ile Gly Tyr Val Phe Ser Phe Ile Ser Leu Phe Ala

50 55 60

Leu Leu Phe Gln Pro Ile Leu Gly Phe Leu Ser Asp Lys Leu Gly Ile

65 70 75 80

Arg Lys His Leu Leu Trp Leu Leu Ala Ile Leu Leu Leu Phe Tyr Ala

85 90 95

Pro Phe Phe Ile Tyr Val Phe Ala Pro Leu Leu Lys Thr Asn Leu Trp

100 105 110

Leu Gly Val Ile Ala Gly Gly Ala Tyr Met Gly Phe Val Phe Gln Ala

115 120 125

Gly Ala Pro Ala Ser Glu Ala Tyr Ile Glu Arg Ile Ser Arg Leu Asp

130 135 140

Gly Phe Glu Tyr Gly Arg Thr Arg Leu Phe Gly Met Leu Gly Trp Ala

145 150 155 160

Ile Cys Ala Ser Ile Ala Gly Asn Leu Tyr Ser Ser Gln Pro Asn Ala

165 170 175

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

180 185 190

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

195 200 205

Leu Gly Val Asn Lys Ser Pro Ile Thr Leu Lys Gln Ala Leu Lys Leu

210 215 220

Phe Ser Leu Pro Arg Phe Trp Ala Leu Leu Thr Tyr Val Val Gly Val

225 230 235 240

Ala Cys Val Tyr Asp Ile Phe Asp Gln Gln Phe Gly Asn Phe Phe Asn

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

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

370 375 380
 Gly Phe Gln Ser Thr Tyr Phe Ile Gly Cys Ile Ala Leu Thr Phe Thr
 385 390 395 400
 Ser Val Ser Val Phe Thr Leu Val Asn Thr Lys Lys Leu Val Asn Asn
 405 410 415

Ala

<210> 133

<211> 9317

<212> DNA

<213> Artificial Sequence

<220><223> Plasmid

<400> 133

tcgaggaatt cgcaggaccg tgatacacgg gacaggtcac tgaatgacga caatgtcctg

60

gaaatcagcg aaccgcgcatt ctgaagtaca tttagcgcgac tgtaccagaa catgaatgag 120
 gcgttttgat taggcgatta ttagcagggc taagcathtt actattatta ttttccggtt 180
 gagggatata gagctatcga caacaaccgg aaaaagtta cgtctatatt gctgaaggta 240
 caggcgtttc cataactatt tgctcgcgtt ttttactcaa gaagaaaatg ccaaatagca 300
 acatcaggca gacaataccc gaaattgcga agaaaactgt ctggtagcct gcgtggtcaa 360
 agagtatccc agtcggcggtt gaaagcagca caatcccaag cgaactggca atttgaaaac 420
 caatcagaaa gatcgtcgac gacaggcgct tatcaaagtt tgccacgctg ttttgaaga 480

 cggatatgac acaaagtga acctcaatgg catgtaacaa cttcactaat gaaataatcc 540
 aggggttaac gaacagcgcg caggaaagga tacgcaacgc cataatcaca actccgataa 600
 gtaatgcatt ttttggccct acccgattca caaagaaagg aataatcgcc atgcacagcg 660
 cttcagtagc cacctggaat gagttgagat aaccatacag gcgcgttcct acatcgtgtg 720
 attcgaataa acctgaataa aagacaggaa aaagttgttg atcaaaaatg ttatagaaag 780
 accacgtccc cacaataaat atgacgaaaa ccagaagtt tcgatccttg aaaactgcga 840
 taaaatcctc tttttttacc cctcccgcat ctgccgttac gcactggtga tccttatctt 900

 taaaacgcat gttgatcatc ataaatacag cgccaaatag cgagaccaac cagaagtga 960
 tatggggact gatactaaaa aatatgccgg caaagaacgc gccaatagca tagccaaaag 1020
 atccccaggc gcgcgctgtt ccatattcga aatgaaaatt tcgcgccatt ttttcggtga 1080
 agctatcaag caaacgcgat cccgccagat accccaagcc aaaaaatagc gccccagaa 1140
 ttagacctac agaaaaattg ctttgagta acggttcata aacgtaaatc ataaacggtc 1200
 cggatcaagc caggatgaaa ctcatacc agatgagcgg tttcttcaga ccgagtttat 1260
 cctgaacgat gccgtagaac atcataata gaatgctggt aaactggttg accgaataaa 1320

 gtgtacctaa ttccgtccct gtcaacccta gatgtccttt cagccaaata gcgataaacg 1380
 accaccagc cgaccaggaa ataaaaaga gaaatgagta actggatgca aaacgatagt 1440
 acgcatctct gaatggaata ttcagtcca taattacctg cctgtcgtta aaaaattcac 1500
 gtctatatta gagataagag cgacttcgcc gtttacttct cactattcca gttctgtcg 1560
 acatggcagc gctgtcattg cccctttcgc cgttactgca agcgtccgc aacgttgagc 1620
 gagatcgata attcgtcgca tttctctctc atctgtagat aatcccgtag aggacagacc 1680
 tgtgagtaac ccggcaacga acgcatctcc cgccccgtg ctatcgacac aattcacaga 1740

 cattccagca aaatggtgaa cttgtcctcg ataacagacc accaccctt ctgcaccttt 1800
 agtcaccaac agcatggcga tctcatactc ttttgccagg gcgcataat cctgatcgtt 1860
 ctgtgttttt ccaactgataa gtcgccattc ttttccgag agcttgacga catccgccag 1920

ttgtagcgcc tgccgcaaac acaagcggag caaatgctcg tcttgccata gatcttcacg	1980
aatattagga tcgaagctga caaaacctcc ggcatgccgg atcgccgtca tcgcagtaaa	2040
tgcgctggta cgcgaaaggct cggcagacaa cgcaattgaa cagagatgta accattcgcc	2100
atgtcgccag cagggcaagt ctgtcgtctc taaaaaaga tcggcactgg ggccggacat	2160
aaacgtaaat gaacgttccc cttgatcggt cagatcgaca agcaccgtgg atgtccggtg	2220
ccattcatct tgcctcagat acgtgataac gactccctca gttagcagcg ttctttgcat	2280
taacgcacca aaaggatcat cccccaccg acctataaac ccacttggtc cgctaactct	2340
ggcgattccc accgcaacgt tagctggcgc gccgccagga caaggcagta ggccgccgtc	2400
tgattctggc aagagatcta cgaccgcatc ccctaaaacc catactttgg ctgacatctt	2460
tttcccttaa attcatctga gttacgata gtgataaacc tctttttcgc aaaatcgta	2520
tggatttact aaaacatgca tattcgatca caaaacgtca tagttaacgt taacatttgt	2580
gataatcatc gcatttatga aagtaaggga ctttatcttt ataaaagta acgttaacaa	2640
ttcaccaaat ttgcttaacc aggatgatta aaatgacgca atctcgattg catgcggcgc	2700
aaaacgccct agcaaaactt catgagcacc ggggtaacac tttctatccc cattttcacc	2760
tcgcgcctcc tgccgggtgg atgaacgac caaacgccct gatctggttt aacgatcggt	2820
atcacgcgtt ttatcaacat catccgatga gcgaacactg ggggccaatg cactggggac	2880
atgccaccag cgacgatatg atccactggc agcatgagcc tattgcgcta gcgccaggag	2940
acgataatga caaagacggg tgtttttcag gtagtgctgt cgatgacaat ggtgtcctct	3000
cacttatcta caccggacac gtctggctcg atggtgcagg taatgacgat gcaattcgcg	3060
aagtacaatg tciggttacc agtcgggatg gtattcatct cgagaaacag ggtgtgatcc	3120
tcactccacc agaaggaatc atgcacttcc gcgacctaag agtggtggcg gaagccgaca	3180
catggtggat ggtagtggg gcgaaagatc caggcaacac ggggcagatc ctgctttatc	3240
gcggcagttc gttgcgtgaa tggaccttcg atcgcttact ggcccacgt gatgcgggtg	3300
aaagctatat gtgggaatgt ccggactttt tcagccttgg cgatcagcat tatctgatgt	3360
tttccccgca gggaatgaat gccgagggat acagttaccg aaatcgcttt caaagtggcg	3420
taatacccg aatgtggctg ccaggacgac tttttgcaca atccgggcat ttactgaac	3480
ttgataacgg gcatgacttt tatgcaccac aaagcttttt agcgaaggat ggtcggcgta	3540
ttgttatcgg ctggatggat atgtgggaat cgccaatgcc ctcaaacgt gaaggatggg	3600
caggctgcat gacgtggcg cgcgagctat cagagagcaa tggcaaacct ctacaacgcc	3660
cggtacagga agctgagtcg ttacgccagc agcatcaatc tgtctctccc cgcacaatca	3720
gcaataaata tgttttgcag gaaaacgcgc aagcagttga gattcagttg cagtgggcgc	3780

tgaagaacag tgatgccgaa cattacggat tacagctcgg cactggaatg cggctgtata	3840
ttgataacca atctgagcga cttgttttgi ggcggtatta cccacacgag aatttagacg	3900
gctacgtag tattccctc cgcagcgtg acacgctcgc cctaaggata tttatcgata	3960
catcatccgt ggaagtattt attaacgacg gggaagcggg gatgagtagt cgaatctatc	4020
cgcagccaga agaacgggaa ctgtcgcttt atgcctccca cggagtggct gtgctgcaac	4080
atggagcact ctggctactg gggttaacata atatcagggt gaacaacgga tcaacagcgg	4140
gcaagggatc cacgaagctt cccatgggtga cgtcaccggg aaaccagcaa tagacataag	4200
cggctattta acgacctgc cctgaaccga cgaccgggtc gaatttgctt tcgaatttct	4260
gccattcatc cgcttattat acttattcag gcgtagcacc aggcgtttta gggcaccaat	4320
aactgcctta aaaaaattac gccccgcct gccactcatc gcagtactgt tgtaattcat	4380
taagcattct gccgacatgg aagccatcac agacggcatg atgaacctga atcgccagcg	4440
gcatcagcac ctgtcgctt tgcgtataat atttgcccat ggtgaaaacg ggggcgaaga	4500
agttgtccat attggccacg tttaaatcaa aactggtgaa actcaccag ggattggctg	4560
agacgaaaa catattctca ataaacctt tagggaaata ggccaggttt tcaccgtaac	4620
acgccacatc ttgcgaatat atgtgtagaa actgccggaa atcgtcgtgg tattcactcc	4680
agagcgatga aaacgtttca gtttctcat ggaaaacggg gtaacaagg tgaacactat	4740
cccatatcac cagctcaccg tctttcattg ccatacggaa ttccggatga gcattcatca	4800
ggcgggcaag aatgtgaata aaggccgat aaaacttggt cttatttttc tttacggtct	4860
ttaaaaaggc cgtaatatcc agctgaacgg tctggttata ggtacattga gcaactgact	4920
gaaatgcctc aaaatgttct ttacgatgcc attgggatat atcaacgggt gtatatccag	4980
tgattttttt ctccatttta gcttccttag ctctgaaaa tctcgataac tcaaaaaata	5040
cgcccggtag tgatcttatt tcattatggt gaaagtigga acctcttacg tgccgatcaa	5100
cgctctattt tcgcaaaaag ttggcccagg gcttcccggg atcaacaggg acaccaggat	5160
ttattttatc tcggaagtga tcttccgtca caggatatta ttcggcgcaa agggcctcgt	5220
gatacgctta tttttatagg ttaatgtcat gataataatg gtttcttaga cgtcaggtag	5280
cacttttcgg ggaatgtgc gcgcccgt tctgtctggc gctgggcctg tttctggcgc	5340
tggacttccc gctgttccgt cagcagcttt tcgcccacgg ccttgatgat cgcggcgccc	5400
ttggcctgca tatccgatt caacggcccc agggcgctca gaacgggctt caggcgtcc	5460
cgaaggtctc gggccgtctc ttgggcttga tcggccttct tgcgcatctc acgcgtcct	5520

gcgggcgccct gtagggcagg ctcatacccc tgccgaaccg cttttgtcag cgggtcggcc 5580

acggcttccg gcgtctcaac gcgctttgag attcccagct tttcggccaa tccctgcggt 5640

gcataggcgc gtggctcgac cgcttgcggg ctgatggtga cgtggccac tggtgccgc 5700

tccaggccct cgtagaacgc ctgaatgcgc gtgtgacgtg ctttgcgtcc ctgatgccc 5760

cgttgcagcc ctagatcggc cacagcggcc gcaaactggg tctggtcgcg ggtcatctgc 5820

gctttgttgc cgtagaactc cttggccgac agcctgccgt cctgcgtcag cggcaccacg 5880

aacgcggtca tgtgcgggct ggtttcgtca cggtggtatgc tggccgtcac gatgcgatcc 5940

gccccgtact tgtccgccag ccacttgtgc gccttctcga agaacgccgc ctgctgttct 6000

tggctggccg acttccacca ttccgggctg gccgtcatga cgtactcgac cgccaacaca 6060

gcgtccttgc gccgcttctc tggcagcaac tcgcgcagtc ggcccatcgc ttcctcggtg 6120

ctgctggccg cccagtgcgc gttctctggc gtctctgtgg cgtcagcgtt gggcgtctcg 6180

cgctcgggt aggcgtgctt gagactggcc gccacgttgc ccattttcgc cagcttcttg 6240

catcgcatga tcgcgtatgc cgccatgcct gccctccct tttgggtgcc aaccggtcgc 6300

acggggggcag cgcaaggcgg tgccctccgc gggccactca atgcttgagt atactcacta 6360

gactttgctt cgcaaagtcg tgaccgccia cggcggctgc ggcgccctac gggttgctc 6420

tccgggcttc gccctgcgcg gtgcgtgcgc tcccttgcca gcccgtagat atgtggacga 6480

tggccgcgag cggccaccgg ctggctcgct tcgctcggcc cgtggacaac cctgctggac 6540

aagctgatgg acaggctgcg cctgcccacg agcttgacca cagggttgc ccaccggcta 6600

cccagccttc gaccacatac ccaccggctc caactgcgcg gcctgcggcc ttgccccatc 6660

aattttttta attttctctg gggaaaagcc tccggcctgc ggctgcgcg cttcgttgc 6720

cggttggaca ccaagtggaa ggcgggtcaa ggctcgcgca gcgaccgcgc agcggcttgg 6780

ccttgacgcg cctggaacga cccaagccta tgcgagtggg ggcagtcgaa ggcgaagccc 6840

gcccgcctgc cccccgagac ctgcaggggg gggggggcgc tgaggtctgc ctgctgaaga 6900

aggtgttgc gactcatacc aggcctgaat cgccccatca tccagccaga aagtgaggga 6960

gccacggttg atgagagctt tgtttaggt ggaccagttg gtgattttga acttttgctt 7020

tgccacggaa cggctcgcgt tgtcgggaag atgcgtgac tgatccttca actcagcaaa 7080

agttcgattt attcaacaaa gccgcgtcc cgtcaagtca gcgtaatgct ctgccagtgt 7140

tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat 7200

ttattcatac caggattatc aataccatat ttttgaataa gccgtttctg taatgaagga 7260

gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 7320

actcgtccaa catcaatata acctattaat ttcccctcgt caaaaataag gttatcaagt 7380

gagaaatcac catgagtgc gactgaatcc ggtgagaatg gcaaaagctt atgcatttct 7440
 ttccagactt gttaacagg ccagccatta cgctcgatc caaaatcact cgcatacaacc 7500
 aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc gctgttaaaa 7560
 ggacaattac aaacaggaat cgaatgcaac cggcgcagga aactgccag cgcatacaaca 7620

 atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt cccggggatc 7680
 gcagtgggta gtaacctgc atcatcagga gtacggataa aatgcttgat ggtcggaaga 7740
 ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg 7800
 ctacctttgc catgtttcag aaacaactct ggccgcacgg gcttccata caatcgatag 7860
 attgtcgac ctgattgcc gacattatcg cgagccatt tataccata taaatcagca 7920
 tccatgttgg aatttaacg cggcctcgag caagacgttt cccgttgaat atggctcata 7980
 acacccttg tattactgtt tatgtaagca gacagtttta ttgttcata tgatatattt 8040

 ttatcttggt caatgtaaca tcagagattt tgagacacaa cgtggcttcc cccccccc 8100
 ctgcaggtcc cgagcctcac ggccgaggt gcgggggttc caagggggca gcgccacctt 8160
 gggcaaggcc gaagcccggt cagtcgatca acaagcccg gagggggcac tttttgccgg 8220
 agggggagcc gcgccgaagg cgtgggggaa cccgcaggg gtgccttctt ttgggcacca 8280
 aagaactaga tataggcgga aatgcgaaag acttaaaaat caacaactta aaaaaggggg 8340
 gtacgcaaca gtcattggt gcacccccg caatagctca ttgcgtaggt taaagaaaat 8400
 ctgtaattga ctgccacttt tacgcaacgc ataattgttg tcgcgtgcc gaaaagtgc 8460

 agctgattgc gcatggtgcc gcaaccgtgc ggccacctac cgcattgaga taagcatggc 8520
 cacgcagtcc agagaaatcg gattcaagc caagaacaag cccggtcact ggttgcaaac 8580
 ggaacgcaaa gcgcatgagg cgtgggcccgt gcttattgct aggaacacca cggcggcaat 8640
 gctgctgcat cactcgtgg cgcagatggg ccaccagaac gccgtggtgg tcagccagaa 8700
 gacattttc aagctcatcg gacgttcttt gcggacggtc caatcgcag tcaaggactt 8760
 ggtggccgag cgttgatct cgtcgtgaa gctcaacgg cccggcaccg tgcggccta 8820
 cgtggtcaat gaccgctgg cgtggggcca gcccgcgac cagttgcgc tgcggtgtt 8880

 cagtccgcc gtggtggttg atcacgacga ccaggacgaa tcgtgttgg gcatggcga 8940
 cctgcgccg atcccgacc tgtatccgg cgagcagcaa ctaccgacc gcccggcga 9000
 ggagccgcc agccagccc gattccggg catggaacca gacctgccag cttgaccga 9060
 aacggaggaa tgggaacgg gcgggcagca gcgcctgcc atgccgatg agccgtgttt 9120
 tctggacgat ggagagcgt tggagccgc gacacgggt acgtgccgc gccggtagca 9180
 cttgggttgc gcagcaacc gtaagtgcgc tgtccagac tatcggtgt agccgctcg 9240

ccgccctata ccttgtctgc ctccccgcgt tgcgtcgcgg tgcattggagc cgggccacct 9300

cgacctgaat ggaagcc 9317

<210> 134

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 134

accattgtgg cgatgggttg cttctacagc ctgaacgaga ggatcccttg cccgctgttg 60

<210> 135

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 135

ttacgggctt ctatctcttc cacaatgcgg acatacatct gaattcgcag gaccgtgata 60

<210> 136

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 136

attaggtaca ctttattcgg tcaaccagtt taccagcatt cgtcttgagc gattgtgtag 60

<210> 137

<211> 60

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 137

tcggcggttga aagcagcaca atccaagcg aactggcaat tgaatatacct ccttagttcc 60

<210> 138

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 138

atggcactga atattccatt c 21

<210> 139

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 139

ctatatgtct gaaggtacag 20

<210> 140

<211> 447

<212> DNA

<213> Artificial Sequence

<220><223> Promoter/MCS/terminator insert

<400> 140

aggaattccc taggcgatct gtgctgtttg ccacggatg cagcaccagc gcgagattat 60

gggctcgac gctcgactgt cggacggggg cactggaacg agaagtcagg cgagccgtca 120

cgcccttgac gatgccacat cctgagcaaa taattcaacc actaaacaaa tcaaccgcgt 180

ttcccggagg taaccaagct tgcccggatc cgcattgcgcg gccgcgtcga ctctagttaa 240

aacccccggg tgatcgatag ctcttaatta agttgtttgc caatgtaatg ccgctgcacc 300

caggcatcaa ataaaacgaa aggctcagtc gaaagactgg gcctttcgtt ttatctgttg 360

tttgtcgggtg aacgctctct actagagtca cactggctca ccttcgggtg ggcctttctg 420

cgtttataca gctgtcggtg ccgccag 447

<210> 141

<211> 1260

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Citrobacter sp

<400> 141

atgaaaatca atatgccgtt ctccaatgac aaataccgtt atagttcggg ctacctgctg 60

ttctttcttcg ctgcgtgggc cctgtgggtg agttttctacg caatctggct gaaaaacaaa 120
ctgggcctgt ccggcaccga actgggcatg ctgtatgctg ttaatcagtt ttctccatg 180
ctgttcacgc tggctacggt cttttctgaa gataaactgg gcaccgcgtg acatctgatt 240
tggctgatgg gcattgtgat cagcgtgtca ggtccgttcc tgatctatgt ttacgaaccg 300
ctgctgacct cgaactttaa actgggcatg gcactgggtg ctattttctt tggctcgggt 360
tatctggcag gttgcggcct ggtggaatct tttgtggaaa aagttttctg taaattcaac 420
ttcgaatttg gcaccgcacg tctgtggggc tctctgggtt acgcggccgg tacgttcgtt 480

ggcgggtattt tctttagcat caaccgcac attaatTTTT ggtgtgtctc tgtgatgggc 540
gtcctgttcc tgctgatcaa cgtgtgtttt aaaaccaata gtccggcacc gagctctgtg 600
aaaaccgtt ccccggaacc ggatgctctg acgcgcaaag acttcctgac catctttaaa 660
gatacgagct tctggttttt cgttattttt gtggttggca cgtggagtgt ctattccatc 720
tacgaccagc aaatgttccc ggtgttttat gcgagcctgt ttgatgacct ggaactggcc 780
ccgcgtgttt atggttacct gaactctgtt caagtcttca tggaagcggg tggcatggcc 840
ctggtcccggt ttctgattaa tcgtatcggt ccgaaaagcg cactgctgct gggcggcacc 900

atcatggcat gccgcattct ggggtcagct ctgtttacgg atatctacat catctcgctg 960
atcaaaatgc tgcatgcgtt ggaagtcccg ctgttcgtca tttcagtgtt caaattttcg 1020
gtggccaact ttgacaaacg cctgagttcc accatttacc tgatcggtt taatatcgcg 1080
tcatcgattg gtattatcgt gctgagctctg ccggttggca aactgttcga taaagtgggt 1140
tatcaggaaa tttttctgat catggccagc atcgtcatta tcacctgat ttccggtac 1200
tttagcctgt ctaaaaaaca tcaccagcaa aaaatgggta acgaactggt gacggaataa 1260

<210> 142

<211> 1314

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Enterococcus faecium*

<400> 142

atgaaaggtg ataccaacat tagcctggaa gacaaaaata tgagcaaagt caacgtgttc 60
aaaaaccaat cctatctgca atcgtccgcc accctgctgc tgtttttcgc atcatggggc 120
gtttgggtgt cgtttttcca actgtggctg accagcgaat ctaatggcct gggctctgagt 180
ggctccgctg tcggtacggt gttttcagcg aactcgtcgg tcacctgat cctgatgttc 240
atttatggta cgctgcagga taaactgtac atcaaacgta atctgctgat ttttgcttct 300

gttctggcga ccctggtcgg cccgtttttc atctggattt atggtcgct gctggacaac 360

aattttaacc tgggcattat catgggtgct ctgttctga gcgccggcta cctggcatct 420

gtgggtgttt ttgaagcgg tagcgaacgt ttctctcgcc tgtttggctt cgaatatggt 480

caggcacgtg ctggggcag ctttggttac gcgctggttg cctgctggc aggccttctg 540

ttcgtcaaaa atccgcatct gaacttctgg gccggcagct tttcggttc cctgctgctg 600

ctgaatctgc tgttttgaa cccgaaagt gaacgcgaag ctaaccagaa tttcaaccag 660

gaacaagcgg aatcaaattc gatcccgagc ctgaaagaaa tgtttgatct gatgaaactg 720

ccgcaactgt ggaccattat catttttatt gtcttcacct ggacgtttta tacggtgttc 780

gaccagcaaa tgtttccggg ctctacacc ggtctgttta gcacgtctgc aaatggcgaa 840

aaaatctatg gcacctgaa cgccattcag gttttctgcg aagcactgat gatgggcatc 900

gtgccatca ttatgcgtaa actgggcgtt cgcaatagc tgcgtctggg tgtgaccatc 960

atgtgcgttc gtattggcct gtgtggtttt gccagcacc cgctgagtgt ctctgtatt 1020

aaaatgctgc atgcactgga agtgccgtg ttacgtgc cgatgtttcg ctatttcacc 1080

ctgcacttcg atacgaaact gagtccacc ctgtacatga tcggctttca gattgcggcc 1140

cagatcggtc aagtattct gtccaccccg ctgggcatcc tgcgcgataa cgtcggttat 1200

caaccgacgt ttaaaatcat tagtctgatt gtgctgctgg ctggtatcta cgcgtttttc 1260

attctgaaac aggatgaccg tgatgtgcaa ggcgaccgt ttattcgagg ttaa 1314

<210> 143

<211> 1299

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Corynebacterium glucuronolyticum

<400> 143

atgtcaaaat tcatgcaaca actgaaaaat acggcttacc aacaatcatc ggcacaactg 60

ctgctgttct ttatgtcgtg gggcatttgg tggagctttt tccaactgtg gctgagctct 120

gaaacccgtg gcctgggttt taatggttcc gaaattggca cgatctatag cgtgaactct 180

gcagtgacc tggttctgat gctggtttac ggcacggctc aggataaact gcgcaccgt 240

cgcaatctgg ttattggtat cgcggtccig atgagcctga cgggcccgtt tttcatgtgg 300

gigtattggc cgtgctgca aagtgaatcc ctgtacgtcc tgggtgtggg tctgggtgcg 360

atttttatcg gcaccgcat cgtgggcagt tgcccgtgt ttgaagcgt ggccgaactg 420

atgtcccga aacataactt tgaatatggt caggcacgtg cttggggtag tttcggtac 480

gcaattgtgg ctctgctggc gggctttaac ttcaccatta atccggcaat caacttttgg 540
atggcgagcg ccttcggtgt tctgctgctg ctgattctgg tgttttgaa agaaccggtt 600

gcaccgcgca atgaaatcgc tgaagaagaa gtggaaaaca ccacgccgtc agtgaaagaa 660
atggtttcgg tcctgaaagt gccggccctg tgggtgggta tcgtgctggt gtttttcacc 720
tggacgtttt atacggtttt cgatcagcaa atgtttccgc agttctacac cagcctgttt 780
tcagactcgg cgaccggtga acgtacgtat ggcgtcctga acagcgtcca agtgtttgtt 840
gaagccctga tgatgggcat tgttccgac tacatgcgta aagtcggtgt gaaaaacacc 900
ctgatgacgg gctttgccgt gatggcactg cgcattctgg gttgcgctgt gttcgcggat 960
ccggttacca tcagctttgt caaaatgttc catgcgctgg aagtgcgct gtgtattctg 1020

ccgatctttc gctatttcac cctgcacttt ccgacgaaaa tttcagcgac cctgtacatg 1080
gttggtttcc agatcgctc gcaagtcggc aatgtcgtga tgagtcgat tctgggttcc 1140
ctgcgtgac gcctgggctt tcagccgacc ttctatgtga ttagcggtat cgttctggtc 1200
tctgccatct ttgatggct ggctctgaaa ggtgacaaaag aacaggttga aggcgatccg 1260
ttctatcgtg actctgaact gaaagaaatt caccaataa 1299

<210> 144
<211> 1326
<212> DNA
<213> Artificial Sequence
<220><223> codon optimized CDS from Bifidobacterium animalis

<400> 144

atggcgacca cgaccaaagt ctggcgtaat ccgagttatc tgcaatcctc cacgggcatt 60
tttctgttct tctgttcttg gggcatctgg tggagctttt tccagcgttg gctgaactcc 120
atgggcctga atggtgctga agttggcacc atttatagta tcaactccct ggcgacgctg 180
attctgatgt ttggctacgg tctgatccag gataatctgg gcctgaaacg tcgcctggtg 240
ctggttatta gcgcaatcgc agccctgggtg ggtccgtttg tgcaattcgt ttatgcaccg 300
ctgatgcgca ccaacatgat ggcgcgcgcg ctggttggca gcgtgggttct gctgcaggc 360
tttatggctg gttgcagcct gattgaagcc gtcacggaac gttattctcg tcgctttaac 420

ttcgaatacg gtcagagtcg cgcgtggggc tccttcggtt acgctatcgt cgcgtgggtt 480
gcgggctttg ttttcaacat caaccgatg atcaacttct ggctgggttc tgccttcggc 540
gtgggtatgc tgattgttta tctgacctgg tatccggcgg aacaacgtga agcactgaaa 600
gaagccgcag atccgaacgc tgcgccgacc aatccgacga ttaaagacat gctgggcgtg 660

ctgaaaaatgc cgaccctgtg ggtttctgac gtctttatgc tgctgaccaa cacgttttat 720
accgttttcg atcagcaaat gtttccgacg tattacgcga gtctgttccc gaacgaagcg 780
accggcaatg ccgtttacgg tacgctgaac agcgtccagg tgttttgtga atctgctatg 840

atgggtgtcg tgccgattat catgcgtaaa gttggcgtcc gcaatgcact gctgctgggt 900
agcaccgtga tgtttctgcg tattggcctg tgcggtatct tccatgatcc ggtcagcatt 960
tctatcgtga aaatgtttca cgctattgaa gtcccgtgtg tctgtctgcc ggcgtttcgc 1020
tatttcaccc tgcattttaa tccgaaactg agtgccacgc tgtacatggg ggcgtttcag 1080
attgcatcac agatcggtca agttgtcttc tcgaccccg cgggcatgct gcacgatcgt 1140
atgggtgacc gcaccacgtt tctgaccatt tccgcatcg ttttgccgc aacggcttat 1200
ggctttttcg tgattaaacg tgatgacgaa caagtggatg gtgacccgtt catcccgat 1260

agcaaaaaac tgccgtcgct ggcaaccgac gaagctattc tgtcagcgga ttcggaagac 1320
atgtaa 1326

<210> 145
<211> 1398
<212> DNA
<213> Artificial Sequence
<220><223> codon optimized CDS from Bifidobacterium gallicum
<400> 145

atggtcaata aaccgaaaac cgctaaaatc tggagcaatc cgtcatacct gcaatcttcc 60
tttggcatct tctgtttctt ctgttcctgg ggcatctggg ggagcttttt ccagcgctgg 120
ctgaacacca ttggcctgaa tgggtcggaa gtcggcaccg tgtatagtat caactccctg 180

gccacgctga ttctgatgtt tggctacggg attatccaag ataatctggg catcaaactg 240
cgcttggtgg ttgtcattgc aaccatcgcg gccctgattg gtccgtttgt tcagttcgtc 300
tatgtccgc tgatgcaaac gaacatcatg gcggcggcac tgattggcag tgtggttctg 360
tccgcaggct ttatgtcagg ttgctcgctg atcgaagctc tgaccgaacg ctatagccgt 420
aaatttggct tcgaatacgg tcagagccgc gcgtggggct cttttggta cgccattgtc 480
gcactgattg ctggcatcgt gttcaacatc aaccgatga tcaacttctg gctgggttcc 540
gcgttcggcg tcggtatgct gattgtgtat ctgacctggg atccggcgga acagcgtcaa 600

gccctgaaag aagccgcaga tccgaacgcc gaaaaaagca atccgtcttt taaagacatg 660
gttaacgtcc tgaaaatgcc gacgtgtgg gttctgatta tcttcatgct gctgaccaat 720
acgttttata ccgtgttcga tcagcaaatg tttccgacgt attacgcctc actgttcccg 780

tcgatcgaaa ccggcaacac ggtgtacggt gttctgaata gtattcaggt gttttgcgaa 840
tccgcaatga tgggtgtcgt gccgattatc atcgcaaaaa tcggcggttcg taacgcactg 900
ctgctgggtg ctaccgtgat gtttctgcgc atcggcctgt gcggtatattt ccatgatccg 960
gttgcaatta gtatcgtaa aatgtttcac gctattgaag tgccgctgtt ctgtctgccg 1020

gcgtttcgtt atttcaccct gcattttaat ccgaaactgt ctgcgacgct gtacatggtt 1080
ggctttcaga ttgcctcaca gatcgggtcaa gtgattttct cgaccccgtt gggcatgctg 1140
cacgatcgct ttggtgaccg taccacgttc ctgagcattt ctggcatcgt gctgctggca 1200
acgatctatg gctttttcgt tattaacgt gatgacgaac atgtcgatgg cgacccgttt 1260
ctgcgtgatc gcgaccgtaa agaaatggaa ctgattgaag aaaacctgca gccggatgcc 1320
gaactggaaa ccagccccgtt ggggtgttgc gcacaagtgc gtgacaatcg tgcggttcaa 1380
ccggaatatg cctcttaa 1398

<210> 146

<211> 1338

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Bifidobacterium longum

<400> 146

atggcgctcag cgaccaaacc ggctgggaaa aaccgcctct atctgcaatc ctcatcggc 60
atcttcattgt tcttctgttc gtggggcatt tgggtggtcat tttccagcg ttggtgatc 120
tcgggcgttg gtctgacgaa cgccgaagtt ggcaccattt atagcatcaa ttctctggca 180
accctggtga ttatgtttgt gtacggcggt attcaggatc aactgggtat caaacgtaaa 240
ctggttattg tggtttagct catcgcggcc tgcgtgggtc cgtttgtcca gtctgtgtat 300

gcaccgatga ttctggcggg cggcaccacg cgttggatcg gtgctctgat tggttcaatc 360
gtgctgtcgg cgggctttat gagggttgc tcctgttcg aagctgttac cgaacgttat 420
tctcgcaaat ttggcttcga atacggtcag agccgcgcct ggggtctttt tggttatgca 480
attgtggctc tgtgtcggg ctttctgttc aacattaatc cgctgatcaa cttttgggtt 540
ggttcagcat tcggtccggg catgctgctg gtttacgctt tttgggtccc ggcggaacaa 600
aaagaagaac tgaaaaaaga aacggatccg aacgcagctc cgaccaatcc gtcgctgaaa 660
gaaatggttg cggtcctgaa aatgccgacg ctgtgggttc tgattgtctt tatgctgctg 720

accaacacgt ttataccgt gttcgaccag caaatgtttc cgacgtatta cgtaacctg 780
tttccgaccg aagaaatcgg caacgcgacc tacggcacgc tgaatggttt tcagggtttc 840

ctggaagcgc ccatgatggg tgcgtgccg attatcatga agaaaattgg cgttcgtaat	900
gccctgctgc tgggtgcaac ggtcatgttt ctgcgcatcg gcctgtgcgg tgtgttccat	960
gatccggtta ccattagtat cgtcaaaactg ttctactcca ttgaagtgcc gctgttctgt	1020
ctgccggcgt ttcgttatit caccctgcat tttagacaga aactgagcgc caccctgtac	1080
atggttggct tccagattgc aagccaagtg ggtcaagtta tcttttctac gccgctgggc	1140
gccttcacgc ataaaatggc acaaatcctg ccgaacaatg acatgggtag tcgtgtcacc	1200
ttttgggtga ttccgctat cgtgctgtgt gcgctgattt atggcttttt cgtcatcaaa	1260
catgatgacc aggaagtggg cggtgatccg ttctacaccg acaaaact gcgcaaatg	1320
gaagcggcca aagcgtaa	1338
<210> 147	
<211> 1422	
<212> DNA	
<213> Artificial Sequence	
<220><223> codon optimized CDS from Bifidobacterium adolescentis	
<400> 147	
atgaaatcgg aacaggctca ggctaaaacg acctcagaag caatcgcggc tgcccgtcaa	60
cgccaacgtg aagaaaaaaa acgcatcaaa atggcgtaaa aaaccggttc ggtgtggaaa	120
aaaccgagtt atctgcagag ctcttttggc attttcatgt tttctgtctc ctggggatc	180
tgggtggagct ttttctctcg ctggctgacc gatccgacgc atggcctggg tatgagttcc	240
gccgaacagg gccaaattta tagtatcaat tccctggcaa cgctgggtcat tatgtttgtg	300
tacggcacca ttcaggacca actgggtatc aaacgtaaac tggttatatt catcagcgca	360
gttgcggcct gtgtcgggcc gtttgtccag ttcgtgtatc aaccgatgct gaccgcgggc	420
ggcaccacgc gttttattgg tgcctgtctg ggtagcatcg tgctgtctgc cggctttatg	480
gcaggttgca gcctgttcga agctattacc gaacgttatt ctgcgaaatt tggcttcgaa	540
tacggtcaga gtcgtgcctg gggctccttt ggttatgcag tggttgctct gtgtgcgggc	600
tttctgttca acattaatcc gctgtgaac ttttgggtgg gctcaatctg cggctctgtc	660
atgctgtgtg tgtacgcctt ctgggttccg gcagaacaga aagaagaact gaaaaagaa	720
gtgatccga acgcgacccc gacgaatccg agctttaag aaatgggttc tgcctgaaa	780
atgccgacgc tgggggttct gattgtcttt atgctgttca ccaacacgtt ttatacctg	840
ttcgatcagc aaatgtttcc gaattattac gctagtctgt tcccgaccac ggaaatcggc	900
aacgcgacct acggtacgct gaatagtttt cagggtttcc tggaatccgc catgatgggc	960

gtcgtgccga ttatcatgaa gaaaattggc gtgcgtaatt ctctgctgct ggggtgctacc 1020
gttatgtttg cgcgcacatcgg cctgtgcggg gtcttccatg atccggtgtc agtttcgatt 1080
gttaaactgt ttcacagcat cgaagtcccg ctgttctgtc tgccggcggt tcgttatttc 1140
accctgcatt ttgacacgaa actgtcagcc accctgtaca tgggtgggctt tcagattgca 1200
tcacaagtgg gtcaagttat cttctcgacc ccgatgggtg cactgcacga tgcaatgggt 1260
gaccgcccga cttttttcac gattagcgcc atcgtctttg cagctctggt gtatggcttt 1320

ttcgttatca aaaaagatga ccaggaagtg ggcgggtgatc cgttttacac cgacaacaa 1380
ctgaaagcaa tgaaagcggc cgatgctgaa gttaaagcgt aa 1422

<210> 148
<211> 1503
<212> DNA
<213> Artificial Sequence
<220><223> codon optimized CDS from Bifidobacterium longum
<400> 148

atgcgtgtcg aacaagtccg ccatgctgat ctgcgccagt ccattgacgt gaaccgcctg 60
acctcacata acaactacta cattaccacc agcacgaaag ctctgccgcc ggaacagaaa 120
cgctatatta cgaacaatca gccgagctct aaccaacgta tgatcatggc aagtgtctacc 180

aaatccgcat ggaaaaatcc gtcataacctg cagagttcct ttggcatttt catgtttttc 240
tgcagttggg gtatctgggtg gtcttttttc caacgctggc tgatttcggg cgtgggtctg 300
acgaacgctg aagttaggcac catttattca atcaattcgc tggcgaccct ggtgatcatg 360
tttgtgtacg gcgttattca ggatcaactg ggtatcaaac gtaaactggt tattgttgtt 420
agtgtcatcg cggcctgtgt ggggtccgttt gtccagttcg tgtatgtcc gatgattctg 480
gcgggcggca ccacgcgttg gatcggtgcc ctgattggtt caatcgtgct gtcggcaggc 540
tttatgagcg gttgctctct gttcgaagcc gttaccgaac gttatagccg caaatttggc 600

ttcgaatacg gtcagagccg tgcattggggc tcttttgggt atgcgattgt ggccctgtgt 660
gcaggctttc tgttcaacat taatccgctg atcaactttt ggggttggtag cgccttcggt 720
ccgggcatgc tgcgtgttta tgctttttgg gtcccggcag aacagaaaaga agaactgaaa 780
aaagaaacgg atccgaacgc agctccgacc aatccgtctc tgaaagaaat ggttgcggtc 840
ctgaaaatgc cgacgctgtg ggttctgatt gtctttatgc tgctgaccaa cacgttttat 900
accgtgttcg accagcaaat gtttccgacg tattacgcca acctgtttcc gaccgaagaa 960
atcggcaacg caacctacgg cacgtgaat ggttttcagg ttttcctgga atccgcatg 1020

atgggtgtcg tgccgattat catgaagaaa attggcggttc gtaatgctct gctgctgggt 1080
 gcgacgggtca tgtttctgcg catcggcctg tgcggtgtgt tccatgatcc ggttaccatt 1140
 tcaatcgtea aactgtttca ctcgattgaa gtgccgctgt tctgtctgcc ggcgtttcgc 1200
 tatttcaccc tgcattttga cagaaactg tctgtaccc tgtacatggt tggcttcag 1260
 attgcgagcc aagtgggtca agttatcttt tctacccgc tgggcgcttt ccacgataaa 1320
 atggcgcaaa ttctgccgaa caatgacatg ggtagtcgtg tcaccttttg ggtgatttcc 1380
 gccatcgtgc tgtgcgcact gatttatggc tttttcgtca tcaaacatga tgaccaggaa 1440

gtgggcggtg atccgtttta caccgacaaa caactgcgcc aaatggaagc ggccaaagcg 1500
 taa 1503

<210> 149

<211> 1266

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Mitsuokella multacida

<400> 149

atgggtaatc tgctgaaagc gttctcgaat ccgttttata gcacctctc cctggaaatc 60
 ctgctgttct ttgctggttg gggatatttg tggagctttt tccagatttg gctgaccacg 120
 aaacagggct ttacgggtgc acaagtgggc accatctaca gctttggttc tgcagtcgct 180

ctggtgctga tgttcgttta tggcagcctg caggataaac tgggtatgaa gaaaaccatg 240
 ctgaaatttt tcgcggtttg ccaaattctg gtcggcccgt tttcacctg ggtttatgtg 300
 ccgatgctgg cggccaactt ttacgtgggc gcggtggttg gtgccgtgta tctggcggtt 360
 gcctttctgg cagcttgccc ggtcttcgaa gccgtgaccg aacgtctgag tcgtcgtat 420
 tcctttgaat acggtcaggc acgcgcttgg ggctccttcg gttacgtgtg tgcggccctg 480
 tgtgcgggct ttctgttcac gatgaaccg aatctgatct tttggaccg ttcagcggtt 540
 gcggcgggtg aactgatcgt tctggtctcg atgacgccg aaaatgatgc aagcctgacc 600

gctcaatatg aagtgaaagc cgaatcaatt aaagaatcga aaaccccgag ctttggcgaa 660
 attgtgggtg ttttcaaact gatcgaagtg tggaaaatga ttgtctttgt gatcatgtcg 720
 tggacctttt acacggtttt cgaccagcaa atgtttccgg aatttttcac ccgtttcttt 780
 gcgacgccgg aagccggcca gcaagcatai ggtgtgctga actccattga agtttttctg 840
 gaattttcga tgatgggcct ggtgccgatt ctgatgcgtc gcacggcgt tcgtaaagcc 900
 attctgctgg gttgcgcaat tatgatcgtc cgcacggcg gttgtggcct ggtgacgaat 960

ccgctgggtg ttgccgtcat taaactgctg catgctccgg aaaccgcgct gtttatcctg 1020

gctgtgtttc gttacttcac cctgcacttc gatacgcgca tttctgcaac cctgtatatg 1080

gttggttttc agatcgcggc ccaggtcggg caaattatct tcagtacgcc gctgggcgcg 1140

ctgcatgact ctattgggta ccagagtacc tttctgggta tttccggcat cgtgtgtgtt 1200

gcctcactgt atgcattcgt catcctgaaa aaagatgacc agcaagttga tggtcagccg 1260

ctgtaa 1266

<210> 150

<211> 1275

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Lactobacillus antri

<400> 150

atgaaaaact ctaaactgag cgcattcaaa aactcgttct atctggaaag tagcctgtcc 60

ctgctgctgt tcttcgcgcg ctgggggtatt tggaggagct tttccagat ctggctgacc 120

aacgatctgg gcttttagtg tgcaaaagtc ggcatgattt atacgttcga ctccgcgac 180

accctggttc tgatgtttat ttacggctca gtccaggata aactgggtat taaacgtcgc 240

ctgctgattg gcgttacgat cctggaaatg ctgctgggtc cgtttttcac ctggatctat 300

gcgccgtgc tgcatagtaa tttcattctg ggcgcatttc tgggttact gtacctgtcg 360

tttgctttcc tggcggccag cccgaccttt gaagcactgg ctgaacgtat gagccgtcgc 420

tattctttcg aatcgggtcg tgcgcgtgcg tggggctcat ttggttatgc agtgtcggca 480

ctgtgcgcag gctacctgtt tacgatttcg ccgtacatcg tgttttggt gagctctggc 540

attagcctgc tgaccttctt gctgctgtgt tttggtcgta ccaaaagccc gacgcaggtg 600

gctcgcgatg aaaacaaagc ggaagaagaa cacgatgccg acaaacggag cttcaaagaa 660

atcatctctg ttttcaaaact gaaacaactg tgggaactgg tgtttttcat tatctttagt 720

ggctccttct atacggtttt tgatcagcaa atgtttccgc agtttttcac ccaatttttc 780

aaaacggcag ctcagggcaa caccgcctac ggcattctga acagcatcga agtttttctg 840

gaagctatta tgatggcgat cgtcccgtag attatgaaga aaattggcgt gcgtaaaacg 900

ctgctgattg gtgttaccat catgtttctg cgcacatggcc tgtgcggtct ggtggtagc 960

ccggtgggca tttccatcgt taaactgttc catgccccgg aaaccgcaat tttgcactg 1020

gctatgtttc gttatctgac cctgcacttt gatacgcgcc tgtctgcgac catgtacatg 1080

gtcgtgggtc agattgccgg ccagatcggg caaattatcc tgagtacgcc gctgggcatg 1140

ctgcatgacc gtatcggtta tcgcgccacc ttcctgggtca tttccctgat tgtgatctgt	1200
gcggccggtt acgcatttgt cattctgcgc aaagataatc aggaagtgga cggtaaccg	1260
ctggaaaaca attaa	1275
<210> 151	
<211> 1323	
<212> DNA	
<213> Artificial Sequence	
<220><223> codon optimized CDS from Lactobacillus ruminis	
<400> 151	
atgatgccga ttagcgacaa ctggaaaggt atcctgttta tgaatgacat gaacaaaagc	60
ggtcgtatga gccaaactgaa aaaccggtt ttcacctcaa acgctacgaa tattctgatg	120
tttttcgcgg gctgggggat ttgggtggtcg tttttccaga tctggctgac cacgaaacag	180
ggctttaccg gtgcccgaagt ggggtgaaatt tatagtttta actccgcttt ctactgatt	240
gcgaacctgg ttacagcaa tatccaggat cgtctgggtc tgaaacgcaa tctgctgac	300
ttttgcgcgt gtctgcaagt cttctgggc cgtttttca ctttctgtt cgttccgatg	360
ctgcatgcta acctggaact ggggtcgctg attggtagct gctatctgac cctggcttac	420
ctgagcgctg ctccgatgtt cgaagccctg acggaacgtg catctcgctg ctttaactat	480
cagtacgggt ctgcgcgtgc gtggggcagt ttcggttatg cggtttccgc cctgctggca	540
ggttttgtct tcacatcaa tccgtcactg ctgttttga ttggctcggc aatcgctgtg	600
gttctgctgc tgctgctgct gttctggaac ccggtgcgta ataaagaaac cgttgcccgc	660
tttgaaaacg aaatggtgcg tgaacgcgaa aattcgaaac cgggtagccg tgattttctg	720
aatgtcttca aagtgcgcag tctgtgggaa attgcaatct tcttggtttt ctccgcacc	780
ttctacagca tcttcgatca gcaaatgttc ccgcagtttt tcaccaatt tttcaaacg	840
caggctatgg gtgacatat gtacggcatc ctgaatagcg ttgaagtctt tctggaagcg	900
ctgatgatgg gtctgggtcc gctgctgatg aagaaaattg gtgttcgtcg caccatcctg	960
gtgggcgtta cgtttatgtt cattcgtatc ggcggttgcg gtctgattac caaccgctg	1020
ggcgtgagta tgatcaaat gctgcatgcc ccggaaccg caatcttttg cgtcgtgatg	1080
ttccgttatt acacgctgca ctatgatccg cgcgtctctg ccaccattaa tatcgtgacg	1140
ggcattgcag gtagctttgg ccagatcctg ctgtctaccc cgctgggtct gctgcgtgac	1200
cacattggct atcaaccgac gtttctgggt attgccggtg tcgtcttctg tgcaggtatc	1260
tacggcctgt ttattatccg tcgcgatgac caggaagtga atggcgaacg cctgtccgaa	1320

taa 1323

<210> 152

<211> 1248

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Yersinia frederiksenii*

<400> 152

aagcttatga aacatagcgt ccgtaaccaa tacctgattc tgagcggctc gctgttcacc 60

ttctttcttta cctggctcgtc ggcatcttagc ctgttcagta tttggctgaa ccagtatgtt 120

ggcctgaaag gcaccgaaac ggggtgcgacc ttttccgcaa ttgctctgac ggcactgtgc 180

gcacagccgc tgtacgggtg catccaagat aaactgggcc tgaaaaaaca tctgctgtgg 240

gcgattggta tctgctgct gattagcggc ccgtttttca tctatgtcta cgcgccgctg 300

ctgcgtgtga atatgctggt tgggtccgtc accggcggtc tgtatatggg catgacgttt 360

ttcgcgggca ttggtgccct ggaatcatac accgaacgtg tgtcacgcat ctcggtttt 420

gaatttggca aagcccgcat gtggggctcg ctgggttggg caggltgcaac gtttttcgca 480

ggcatgctgt tcaacatcaa cccgaacatc aacttctgga tggcaagtgc ttccgcggcc 540

atctttctgc tgctgctgtg gcactcgcac gaagtgaata cgcagctat gggtaactg 600

gaatatggca aaaacagcgc actgacctg agtgatacgc tgtccctgtt ccgtatgccg 660

cgcttttggg ccttggtggt ttttgttacc ggtgtctctg tgtataatgt ttacgatcag 720

caatttccgg tctatttcag ctctctgttt acggaccgtc gccatggtaa cgaaatgtac 780

ggcttcctga atagtctgca ggtctttctg gaagcaggcg gtatgttcct ggctccgttt 840

ctggtgaacc gtattggtgc gaaaaaaggc ctgctgctgt ctggcctgat tatggtatg 900

cgcatcctgg gcagcggctc ggcacaagat gctgtgacca tctctctgat gaaactgctg 960

cacgcagttg aactgccgat tctgctgac gctatgttca aatatattgc ggccaatttt 1020

gaccgcgctc tgagcgcgac cctgtatctg gtgggttttc agttcattac gcaagtgtac 1080

gcgagcgttt tctctccgt ggccggcaaa ggttatgac tgatcggtt tgcggacacc 1140

tacctgatta tggcggtat cgttctgggt ctgacggcca tttctgttt tatgctgcgt 1200

ggcgaatcac gcaccgatga cccgtcgtg caactgacca cgaaataa 1248

<210> 153

<211> 1260

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Serratia proteamaculans*

<400> 153

```

atgaatcgtg aaaccaaaaa atactatgtc ctgctgtcgg gtcgtgtgtt cttcttcttc      60
tttacctggt cttcgtcgtt ctgctgtgatt tcaatctggc tgaaccagaa aattggcctg      120
aaaggcaccg aaacgggcct gatctttgcg gccatgtcga ttatggcgct gtgcgcccag      180
ccgctgtatg gctttatcca agataaacgt ggctctgcgt aacatctgct gctgttcgtt      240
ggcgtcctgc tgcgtctgac cggtcggttt ttcatttatg tgtacgcacc gctgctgcag      300

tcaaacctgg tggttggcgc tctggttggc ggtgtgtttg tttcgtggc attcaatgct      360
ggcatcggtg cgctggaaag ctataccgaa cgtgtctctc gcattgtggg ctttgaattt      420
ggctcgtcac gcatgtgggg ctcactgggt tgggcatcgg caacgttttt cgcgggtttc      480
aactacaata ttgaccgcaa catcaatttt tggattgcaa gcgctagcgc ggcgggtgtt      540
ctgctgctgc tgtggcaggt ccgtgaactg aaaccgaacg cgatggccgg cctggaatat      600
ggtaaacggg aaaatctgaa actgcaagat gcactggctc tgcctgcctc gccgggtttt      660
tgggcgctgg tcgtgttcgt gctgggcacc agcatctatg gtgtgtttga ccagcaattc      720

ccggtttact ttgcctctca gttcccgacg catgaagaag gcaaccgat gtacggcttt      780
ctgaacagcc tgcaagtctt cctggaagca ggcggtatgt tttcggctcc gctgctggtt      840
aatcgtattg gcatcaaaac gtctctgctg ctggcgagct ctgttatggc cctgcgcatg      900
gtcggcagtg gttttgcgtc cggtgccctg atgattagtg cgatgaaact gctgcacgcc      960
gtcgaactgc cgatcctgct ggttgcgatg ttcaaataata ttaccacgcg ttttgatagt      1020
cgctgagtt ccaccctgta cctggtgggc ttccagtca tttccaaat cgttgccggt      1080
tttctggcac cgttggtggt ctatggttac gatcgtatcg gctttgcaga cacctatctg      1140

ctgatgggtt gcgcagtggc aggcaccacg ctgattagct gttttctgct gcgcggtgaa      1200
acggttgcaa gtgctccgca gttccaatcc accctgaaat catcggaacc gacgcagtaa      1260

```

<210> 154

<211> 1239

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Escherichia coli*

<400> 154

```

atgaaaaaac gcccgaccgg ctcatatcatg ctgctgtccg cactgctgtt cttcttcttc      60
gtgacctgga gttcgtcctc ttctctgctg agcatttggc tgcatcagga agtgggcctg      120

```

aaagcgtctg aaaccggtat tatctttagt gtccgtgccg tgtcagcact gtttgctcag 180

gtttgctatg gcttcatcca agatcgtctg ggctcgcga aacacctgct gtggtttatt 240

accgcactgc tgatcctgtc gggcccggct tatctgtgt tctcttacct gctgagtgtg 300

aacattctgc tgggtagcgt ttttggcggc ctgttcattg gcccgacgtt taatggcggc 360

atcgggtgtc tggaaatcga taccgaacgt gtgcgcgcc agagcacgtt tgaatttggc 420

cgtgcacgta tgtggggctc tctgggttgg gcagtcgcaa cctttttcgc gggctctgtg 480

tttaacatta atccggacct gaacttctg gtgcctcgt gcagcggcct gatctttttc 540

tgtctgtcgg cactctgaa agtggcagca ccggcctcca tggaaaaact ggaaattggc 600

gcgaaaaaag tticactgga agatgctctg cgtctgtga ccccgccgcg cttttgggcg 660

ctgattttct ttgtggttgg cactgcatc tatggtgtct acgaccagca atttccggtg 720

tatttcagct ctacgtttcc gacctgcgc gaaggcaacg aatgtttgg ttacctgaat 780

agctttcaag ttttcttggg agcggcgggt atgttctgtg cgccgtggct ggtcaaccgt 840

attggcgcca aaaatggctc gatctttgcg ggcatggtca tggccctgcg catgattacc 900

tctggcctgg tggaaaggtc gctgtgatt agtatcacga aactgctgca tgcagttgaa 960

ctgccgattc tgcgtgtcgc tatcttcaa tacaactctc tgaacttcga taaacgtctg 1020

agttccacca ttacctggg gggctttgca tctacctcat cggttatcgg tacggctctg 1080

tcgccgtcgg caggcttcag ctatgaacgt ttgtgttcg ctacagagta cctgattatg 1140

ggcatcatgg tgttctccac cacttttatt tcaatcttcc tgcgtcgtc cacgaaaagc 1200

tctagtgaac cgtcatttct gcagcaaaaa gccgtttaa 1239

<210> 155

<211> 1263

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Cronobacter turicensis

<400> 155

atgagcgaag tcttgaatc aaccgttcc gcctacatca aactgagcct gttcgtcttt 60

ctgttcacct ttacctggag tcttggcttt ggcatgtatg cgatttggct gggcgataaa 120

gccggctcga gtgcaattga catcgccacg gtctttgctg tgaacggtgt tttcgcggtc 180

ctggtgaaac cgttttatgg cttcatcatg gataaaaccg gtatgaaaaa acatctgctg 240

tactttgtgt gctgatttag cgactgatg gctccgttct atctgtggtg ttacctgccg 300

ctgctgcagt cacactttct ggccggcatg attatcggtg cactgttttt ctcgttcggc 360

tggatgctgg gtgttgcggc cgaagaaagc tatgtcgatc gtttttctcg cctgtacaat 420
 atggaatttg gccgtatccg catgtgggca gctctgggtt gggcagtggc agccagcttt 480

 tctggetatg tttacaacat tagcccgatt atcaatttca ccatcggtag tgtctccgca 540
 gtgtgcatgc tgctggctct gatttctctg aaagtggacc attttggcca ggaagaaaac 600
 agcgtctgtg ctaaagataa aattgttatc gcggacgtct ttcacctgtt ccgtaatccg 660
 aaattttgga tgttcgcaat gtatatcgcg ggtgttgcct ggatgatgtt tattgctgaa 720
 cagcaattta gccgctattt cgtgaccttt ttcaacacga aagaacaggg caatgccatg 780
 ttcggttaca tgagtaccac gcaatcctgc acggaatfff tcggcatgat gctgggtgcc 840
 gccctggtta accgtatcgg cgcaaaacag ggtatgctgc tgaccggtct gattgttagt 900

 ctgcgcctga ttatctccgg cctgaccacg gatccgctga ttatctgtct ggtgaaaccg 960
 ctgtatggta ttgaaatcgc actgattctg gtttctgtct ttaaatacat cgctgaacat 1020
 ttcgacaaac gtattaatgc gaccatgtat ctgctgggct accaagccat ggtgtatgtt 1080
 ggtagtatcg tggttgcacc gccgctggc tatgcgtacg ataccattgg ttttgaacac 1140
 acgtacctga ttatgggcac catctccctg ctgtttacga ttgtctcagc cttcacctg 1200
 tcggattgtc agaaacataa aaaagtgacc tcagcttcgc acacggacca agcgtgaac 1260
 taa 1263

<210> 156

<211> 1269

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Enterobacter sp

<400> 156

atggaaaata cctcaacgaa caatcgcacc gaatactata aaatcagcat cttcatcttc 60
 ctgtacttct tcacctggc agcgtcaatt ggctgtctgg ctatctggct gggtcagaaa 120
 gcgaacctgt caggctcggg gatcggcacc gtgtttgcag ttaatggcat cttcagcgtt 180
 attctgaaac cgatctatgg ctacattctg gacaaaatcg gcatgtcgaa atatctgctg 240
 tactttgtgg ttgcaatgag cgctctgatg gcgccgtttt tcatttatgt ttaccagccg 300

 ctgctgatgt ctaacacat gctgggcatt atcattgggtg cctgtatct gagtttcgca 360
 tggtagccg gtgtggcagc atgcgaaagc tattctgac gtttttccg cctgaatggc 420
 atggaatttg gtcagatccg tatgtggggc tactgggtt gggcagttgc tagctctttt 480
 tctggcctgc tgttcaacct gagtccggcc tataatttta tttgggttag tctcgcatcc 540

gtgatcatgc tgattgttct gctgaacctg aaagttaaca ccaatagcgt ccatgcgtct 600
gatgtgctga cgaaagaaaa aatcgccccg tccgacgtct atgcactgct gcgttcacgc 660
aaattttggg ctttctgtct gtacgttgct ggcgtcgcgt ggatgatgtt tatcgcagaa 720

cagcaatfff cgcgctatff cgtgacctff ttcatgaca ttaccaggga caacgccgtt 780
tttggctatc tgggtacggt ccaaagcggg atggaatttg tgatgtacat ggttattccg 840
ctgtttgtta atttcacgg cgcaaaacgt ggtctgctga ttgtcggcct gtttgtgggt 900
gctcgcctga tcatttcagg cctgtgcgat tcgcatctgc tgattagcgt gctgaaaccg 960
ctgtatggtc tggaatctg tctgtgctg gttagtgtct ttaaatacat tgcgaacac 1020
ttcgacaaac gtgtcaacgc gaccatgtat ctgctgggct accaggccat gctgtatgtg 1080
ggtaatgtcg tgatcagttc cccggcgggt tatatgtacg atcgattgg ttttgaacaa 1140

acctacatca ttatgggcgc gacggccctg acctttacgc tgctgtctgc tttcacctg 1200
agtgcgtgcc agtccaaatg gcgtggtacg agcgccattt cggggcaga acaaaatccg 1260
agccgctaa 1269

<210> 157
<211> 1260
<212> DNA
<213> Artificial Sequence
<220><223> codon optimized CDS from *Citrobacter koseri*
<400> 157

atgaacaaaa atgactcggc ggaaaaaaat cgtcgcaaaa tctacatcac gctgtcgtg 60
tttattctgg ttacttctt ctctggaaa gcgacctgg atactactc cttttggctg 120

tcagaaaaaa ttggcctgga cgggtgcgcg attggcatcg tgtttgcggc caacggtttc 180
tgcgccgttc tgattaaacc ggtctatggc tttctgatcg atcgtctggg tctgcgcaaa 240
gacctgctgt ttttcattag tctgatttcc atcacctgt tccgttttt cttttatatac 300
tacaaaccgc tgcgtcagtc tgctctgtac gcgggcattg cctggggcgg tatctttctg 360
tcgatgggct attacgcagg ttgtgcggcg gcggaaagct atctggatcg tttcgccgc 420
ctgtttgacc tggaatttgg tcaaatcgt atgtggggcg cgatcggttc agtcttttcg 480
gccgcaagca ccggttacat cttaacatc aaccgatga tcaattcgg cctgagctct 540

gctgggtgcgc tgatcatcct gatcatcgtt ttaccatga aaatcgaagt gaacgaagaa 600
gttaaaaaac gtatcatcgc aaaagataaa gctacgatct acgacatcat cgccctgttt 660
agcaccgcga atttctggac gtttgttctg tacgtctgcg gcgtggtttg gattatcttc 720

gtggcagaac agcaatatcc gcgcttcttt atcaccttct acgcttctaa agaaatcggt 780
catcagatgt tcggctatct gggctgctgcg agtctgggcg tcgaatttct gtttatgatg 840
attgcgccgg gtatcgtgaa tcgtctgggc ccgaaaaccg gtctgctgat tacgggcttt 900
gtcgtggcgg cagccttcat cggctcgggt ctggttacca gcgcaattac gatcggcatt 960

atcaaaactga gctgggggat tgaaatgtct tttctgctgg ttagcatctt caaataacctg 1020
gaaatcaact tcgataaaaa agtcaacggc tccatgtatc tgctgggtta ccagtcagtg 1080
aactacattg gcaccgttat cctgtctccg ctgagtggct atctgtacga acgtctgggt 1140
ttttcccata cgtatttcat gatgggttgt accgtgctgt tctttacgat tatcagtgcg 1200
tttctgctgc aaaacaataa acgcaccgaa ccggttattg ccgaaaatgc ccacgcataa 1260

<210> 158

<211> 1242

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Bacillus megaterium

<400> 158

atgaaatcct caaaatcgct gtattggaaa ctgtcggtt acttcttctt cttcttcttt 60
acctggatcat cctcttattc gctgttcagc atttggctgg gccaggaaat caaactgaac 120
ggctcagcca ccggtctgat ttttagcgtg aatgctatct tcgcgctgtg catgcaaccg 180
ctgtatggct acatttcaga tcgtatcggt ctgaaaaaac atgttctgtt tttcatttcg 240
tgtctgctgg tgtttgtgg cccgttttat atcttctgtg acggtcgctt gctgcagtat 300
aacgtgggtta ttggcgtat tatcgccggt ctgtacctgg gtgttgcctt tctggcaggc 360
attggtgcaa tcgaaacgta catcgaaaaa gtcagtcgta aatacaaatt cgaatacggc 420

aaaagccgta tgtggggctc tctgggttgg gcggccgcaa ctttttctgc gggtaactg 480
tttaacatta atccgcacat caatttctgg gtggctagcg tttctgcggt cattctgggtg 540
gcgattatct ttagcgtcaa agtggaaatg agctcttatg aaatggaaaa agccgaaagt 600
gttacgtgc gtgatgtcgg ctccctgttt ctgctgaaag aattttggtt tttcatgac 660
tacgtcgtgg gcgttacgtg cgtctatggt gtgtacgacc agcaatttcc gatttattac 720
gcaagtctgt tcccagccga atccatcggc aaccaggtct ttggtatct gaacagcttt 780
caagtgttcc tggaagcggg catgatgttt gctgcgccgt tcattgtgaa caaaatcggt 840

gctaaaaata gcctgattct ggcgggcttt ctgatgggtt tccgcattat cggctctggt 900
ctggttgtcg gcccgattgg tatcagttcc atgaaactga ttcatgcctt ggaactgccg 960

atcatgctga ttgcaatctt taaatatctg gccgcaaact tcgatacccg tctgtcatcg 1020
attctgtacc tggttggctt tcagttcgcc tctcaaattg gtgcaagtgt gctgtccccg 1080
atcgttggcg gtctgtatga ctcagttggt ttttcgcgca cctacctgat tatgggcggt 1140
atggttcttg tctttaacgt gatcagcatg ttcacgctgc tgaacagcaa aaaacaccgt 1200
tttattcgca aagacgtgca ggaaaatacc caaattatct aa 1242

<210> 159

<211> 1263

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Providencia rustigianii

<400> 159

atgggcaaaa tcattggcgc gaccgcacac gcctacttta tcctgagcgg cctgctgttt 60
accttcttct ggacctggag ttcggcattt agcctgggtt ctctgtggct gtctcagaaa 120
gtcggcctga aaggcaccga tacgggtatt atcttcagtg ttatctccct gaccgcattt 180
tgcgctcagc cgctgtatgg ttctattcaa gacaaactgg gcctgcgtaa aaacctgctg 240
tggtacattg gtgtgatgct gggtatcagt ggcccgtggt ttattttcgt ttgtaccccg 300

ctgctgcgct ggaatatatt tctgggcagt atctccctgg gtgtctatgt gggcgcgacc 360
tttttcgccg gcatcggcgc actggaatcc tatacgaac gtgtgtcacg cattgctggt 420
tttgaatacg gtctgtcacg tatgtggggc tcgctgggtt gggcaggtgc aaccttttc 480
gcaggtattc tgttcaacat caaccgaac tacaatttcg ccatgggctc agtgtcggca 540
ctgatcttta ttatcctgct gtggcgtctg agcgtatgtc gccgcaggc aatgaacgaa 600
ctggaatttg gtaaatcgaa tgctctgaaa ctgagcgacg cgctgggcct gctgaaaatg 660
cgttcattct ggagcctggt ggtgtttgtg tgcggtgtta gcgtctataa cgtttacgat 720

cagcaatttc cgatttatatt cgccagcctg ttttctgaac agagtcattg taacgaaatg 780
tttgctatc tgaacagcct gcaagttttc ctggaagccg gcggtatgtt cctggcacccg 840
tttctggtca accgtatcgg tgctaaaaat ggctgtctgc tgtcaggctt tgtcatggct 900
ctgcgcattc tgggctcggg tctggcggat gacaccatta cgatcagctt tatgaaactg 960
ctgcacgcgg tggaactgcc gattctgctg atctctctgt tcaaatacat caccacggtt 1020
ttcgataaac gtctgtctgc caccatctac ctggtcggtt ttcagttcat tgctagtctg 1080
tgcgcaacgg tgctgtctcc gctggcaggt attagttagt atcgcatcgg ctttgccgac 1140

acctacatga ttctggcagg catcgtcttc acctgacga ttatctcctg tgtgtgctg 1200

aaatcagaaa acagctctaa agtccataat gtgagcattc agaagaaaac cgtggaactg 1260
 taa 1263
 <210> 160
 <211> 1266
 <212> DNA
 <213> Artificial Sequence
 <220><223> codon optimized CDS from Providencia alcalifaciens
 <400> 160
 atgggcaaaa aaacccgtgc taacattcat gcttatctga ccctgagcgg cctgctgttc 60
 accttcttct ggacgtggtc gagcgtgttt tctactgggt cgctgtggct gtctcagaaa 120
 attggcctga aaggcaccga tacgggtatt atcttcagcg ttatctctct gaccgctttt 180
 tgcgcgcagc cgctgtatgg ttccattcaa gaccgtctgg gcctgcgcaa aaacctgctg 240
 tggtttctgg gcggtctgct gctgattagc ggtccgtggt ttatcttctg gtgtaccccg 300
 ctgctgaaat ggaatatatt tgccggcgcg gccgcactgg gtatctacgt tggcgtctacc 360
 tttttcgcgg gcattgggtgc cctggaatcc tatacgaac gtgtcagtcg catcgcatcc 420
 tttgaatacg gtcgtgctcg tatgtggggc tctactgggt gggcaggtgc aaccttttc 480
 gccggtctgc tgtttaacat tgatccgaac ctgaatttcg caatgggcag tgtctccgct 540
 ctgctgttta ttatcctgct gtggcgtctg cgcaacgtgc gtccgcatc catgaatgaa 600
 ctggaatttg gtaaaaccca gtactgacg gtgtcggatg cactgggtct gctgcgtatg 660
 cgcagtttct gggtccctgt catctttgtg tgcggtgtgt ctgtttataa cgtttacgac 720
 cagcaatttc cggctctatt cgcaagtctg tttgctaacc agaagagagg taacgaaatg 780
 tttggctatc tgaacagcct gcaagtgttc ctggaagcgg gcggtatgtt cctggccccg 840
 tttctggtta atcgtattgg tgcaaaacgc ggctgctgc tgagcggcct gattatggct 900
 ctgcgtatcc tgggctccgg tctggccgat gacgcaatta ccatctcatt tatgaaactg 960
 ctgcacgcgg ttgaactgcc gattctgctg atcagcctgt tcaaatacat cacgtctgtt 1020
 ttcgataaac gcctgagtcg aaccatctac ctggtcgggt ttcagttcac ggcgagcctg 1080
 tgtgcctcat tctgtgcgcc gctggcgggc tatggttacg ataccctggg cttcgccgac 1140
 acgtatatga ttctggctgc gattgtgatc tcgatgacca ttatcagtc tgttctgctg 1200
 caaagcgaag acaccagcac ggtctctaata ccgatgacca cgagcaagaa aaccgtggaa 1260
 ctgtaa 1266
 <210> 161
 <211> 1290

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Weissella paramesenteroides

<400> 161

```
atgggtggtc gccgtatgtc aaaatatatg aaagccttca aaaacgaaag ctacctgcaa      60
tcctcgggtca cgatgctgct gtacttcgct agctggggca tttggtggag ctttttccaa    120
ctgtggctga ccgcgaaatc aggtctgaac ctggatggct cggccgtggg tacgatctat    180
agtgcgaatt ccttggttac cctggtcctg atgtttttct acggcacgct gcaagacaaa    240
ctgggtatta aacgtcatct gctgatcttt agcgggtgtt gcgcagccct ggtcgggccg    300
tttttcattt atgcgtacga accgatgctg cacaccaact ttacgatggg cctggtggtt    360
```

```
ggttcaattt ttctgtcggc tggcttctctg gcgaccacgg gtatctatga agccctggtc    420
gaacgtttta gccgcgtgtt taaattcgaa tatggccagg cgcgtgcctg gggtctcttt    480
ggttacgcag tcgtggctct gctggcgggt cacctgtttg tgattaacce ggatctgaat    540
ttttgggttg gcagtatctt cggtgtcctg ctgctgctga atgtctgttt ttgggtgccg    600
aaagccgaac gcctggaacg tgttcgcagc gcacaggaaa ccgaaaaaac ggtgccgagt    660
gttcgtgaaa tgctgtccct gctgaaaatg cgcgatctgt gggttgtcat tgtgctgac    720
tttttcacct ggacgtttta taccgttttc gaccagcaaa tgtttccgag ttctacgct    780
```

```
ggcctgttta gctctgtggc gcagggccag caaatgtatg gtaacctgaa cagcattcaa    840
gtgttcgttg aagccattat gatgggtatc gtcccgtga ttatgaacaa aatcggcgtt    900
cgtaataccc tgctgctggg tattgccatc atggcaattc gcatcggcct gtgcggtttt    960
attgataacc cgttggaatc cagcttcgtt aaaatgctgc attcttttga aacgccgctg    1020
ttcattctgt caatctttcg ttatttcacc ctgcactttg acacgaaact gtcggcgacc    1080
ctgtacatga ttggctttca ggtcgcagct caactgggtc aagtgttctt gagtaccccg    1140
ctgggcatgc tgcgcgataa ttccggttat gccgttacgt ttcacattat caccattatc    1200
```

```
gtcattgcgg cgggtatcta cgcatTTTTT gtgatcaaaa aagataacca ggacgtgaat    1260
ggtgaaccgc tggtttctaa agcagcttaa                                1290
```

<210> 162

<211> 1185

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Bifidobacterium pseudocatenulatum

<400> 162

atgacctcgg cggaacaagg ccaaattctac tcaatcaact cgctggctac cctgggttatt	60
atgttcgctt atgggtgctat ccaagaccaa ctgggcatca aacgtaaact ggttatcttt	120
gtcagcgcca ttgcggccct ggttgggtccg tttgtgcaat tcgtttatgc accgatgctg	180
accgcgggcg gcaccacgcg ttctattggt gttctgattg gttcaatcgt cctgtcggca	240
ggctttatgg ctggttgtag tctgttcgaa gcactgaccg aacgttattc ccgcaaattt	300
ggcttcgaat acggtcagag ccgtgcttgg ggctcttttg gttatgctat cgtggcgctg	360
tgtgccggct ttctgttcaa cattaatccg ctgctgaact tttgggtggg tagcatctgc	420
ggcctgggta tgctgtgtat ttacgcgttc tgggttccgg ccgaacaaaa agaagaactg	480
aaaaaagaag cggatccgaa cgcagctccg acgaatccga gttttaaaga aatgatctcc	540
gtcctgaaaa tgccgaccct gtgggtcctg attgtgttta tgctgttcac caacacgttt	600
tataccgtgt tcgatcagca aatgtttccg aattattacg cgagtcgttt ctccaccacg	660
gaaattggca acgccaccta cggtagctg aacagctttc aggttttcct ggaatcggcc	720
atgatggcgg tgggtccgat tatcatgaag aaaattggcg ttctgaatag cctgctgctg	780
ggtgcaaccg tcatgtttct gcgcattggc ctgtgcggtg tgttccatga tccggtcagc	840
gtgtctatcg tgaaactgtt tcaactcaatt gaagttccgc tgttctgtct gccggcgttt	900
cgttatttca cctgcattt tgacacgaaa ctgtcggcaa cctgtacat ggttggcttt	960
cagattgcta gccagatcgg tcaagtcatt ttctctacgc cgatgggtgc actgcacgat	1020
gcaatgggtg accgcccgcg ctttttcacg atttctggca tcgtgctggc ggccctgatc	1080
tatggctttt tcgtcatcaa aaaagatgac caggaagtgg gcggtgatcc gttttacacc	1140
gacaaacaac tgaaagccca agcggcggcg gaagccaacg cataa	1185

<210> 163

<211> 1086

<212> DNA

<213> Artificial Sequence

<220><223> codon oprimized CDS from Bifidobacterium catenulatum

<400> 163

atgtcgcgtg gctctggcac caataaaaaa aaacaacgta ataacatgac ctcaaacacc	60
cgctccgcct ggaaaaatcc gtcctacctg cagagctctt ttggcatttt catgtttttc	120
tgctcatggg gtatctgggtg gtcatttttc tcgcgttggc tgaccgatcc gacgcatggc	180
ctgggtatga cctcggccga acagggccaa atttatagca tcaactctct ggcaacgctg	240

gtgatcatgt ttgcgtacgg cgccattcag gaccaactgg gtatcaaacg taaactggtc 300
attttcgtga gtgccatcgc ggccctgggtt ggtccgtttg ttcagttcgt ctatgcaccg 360
atgctgaccg cgggcggcac caccggtttt attggtgtgc tgattggttc tatcgttctg 420
agttccggct ttatggccgg ttgcagtcig ttcgaagcac tgaccgaacg ttattcccgc 480
aaatttggct tcgaatacgg tcagagtcgc gcctggggct cctttgggta tgcaattgtt 540

gctctgtgtg cgggctttct gttcaacatc aatccgtgc tgaatttttg ggtcggtagc 600
atttcggcc tgggtatgct gtgtgtctac gctttctggg tgccggcgaa acaaaaagaa 660
gaactgaaaa aagaagcgga tccgaacgca gctccgacca atccgtcatt taaagaaatg 720
atctcgggtc tgaaaatgcc gacgtgtgg gtgctgattg tttttatgct gttcaccaac 780
acgttttata ccgttttcga ccagcaaatg tttccgaatt attacgttc cctgttcccg 840
accacggaaa ttggcaacgc gacctcgggt acgctgaata gctttcaggt gttcctggaa 900
tctgctatga tgggcgtggt tccgattatc atgaagaaaa ttggcgtccg taacagcctg 960

ctgctgggtg caaccgtgat gtttctgcgc attggcctgt gtggtgtttt ccacgatccg 1020
gtctctgtga gcatcgtgaa actgttccat agcattgaag tcccgtgtt ctgtctgcgc 1080
gcgtaa 1086

<210> 164
<211> 1176
<212> DNA
<213> Artificial Sequence
<220><223> codon optimized CDS from Clostridium thermocellum
<400> 164

atgattctgg acaacctgaa acgctccgaa cgtatccgt tctcttttat cctgttttac 60
tccctgtttt acatgggtct ggctgtgttt ggtgttttca tgccgggtcta cctggaaggt 120

ctgggctatg ataacaccga catcggcacg tttctgagca ttagctcttt tgtcggctctg 180
ttcgcacagc cgatctgggg cgtgattagc gatcgtgcca aatctaaaaa caatgtgctg 240
aaaatgctgg ttctgtttag ttccattgcg atctttatgt tccgcctgag cggtaattat 300
tactatattt tcgcggtgat ggtggtttac gcctttttcc aaaccccgat tacgccgac 360
ggtgatgcca tcacctgga atacattacc gacacgaaat ggaaatatgg tccgatccgt 420
ctggcagggt ctcgtgggta tgcggtgatg gcctttattg gcggtgcact gacccgcaaa 480
aacatcaacg ctatcttttt catttgcttt gttatcgga tcatgagcct gattaccgtc 540

ttccgtatgc cgacggtgaa aggtcatcag tcagacggca acaaactgtc gatcctggaa 600

gttttcaaaa attctgaact ggtcctgctg atgggtttta ccctgggtgat ccacaccacg 660
atgggcttct acaacacgtt ttccccgac tactacaaaa acatgggtgc ggataatacg 720
atcctgggtc tggcagtgtt tattggctca gcttcggaaa ttatctttct ggttttcggc 780
gaccgtatta tcaaacgcct ggggtattaaa ttaccctgt tcggtgcggc cgtcgtggca 840
gttgtccgtt gggcttcact gggctcgtac aacaatattt ttgcagtgc ggctctgcag 900
atcctgcatg gctttatttt catcgttctg gcgtactcga tggccaccta tattaacaat 960

gaaatgccgc cggaactgaa agcgagcggc cagacggcga actctgtgat cggctcgggc 1020
attagtcgta ttatcggctc caccggcggc ggcgttatca gtgatctgat tggcatccgc 1080
caagtctttt tctgaatag tgtgattgtt ctggcgccca ttgtgatctt tggcgccatc 1140
ttcctggctc gtcgccagaa aattaccggc caataa 1176

<210> 165

<211> 1311

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Granulibacter bethesdensis

<400> 165

atggttctgg atatggatgc gatgagcggc acgattgccg tgagcccggc cctgtcggat 60

ggtcagccgc cgtcggattc ccgtcagaaa cagggcgaat ccgcgctgct ggataaactg 120
aactttttcc tggcagatgt gcgtgacggc ctgggtccgt atctggcagt ttacctgctg 180
agtgcctccg gtcgtgatgg ccgctgggac gaaagctctg tgggcctggt tctgaccatt 240
agtgggtggt ttggcctggt cgcgcagacg ccggccgggtg cactggttga tcgttcccgc 300
aacaaccgc gctcgtggc tgttcgacg ctgctggcga ccctgagtac gctgctgctg 360
ccgtttatgt ccggtctgcc gctggtgacc ctgacgcaaa gcatggcggc cgttgcgggt 420
gcaattttcg ccccggtcat cgcagctatg accctgggtc tgggtggcac ggatggtttt 480

gcacgtcgca ttggccgcaa cgaatcattc aatcatctgg gtaacgcagt ttcggcggcc 540
atcgcaggtc tgctggcatg gcactatggt ccggtcgtgg tgttttggt gatgggcctg 600
ctggctattg cgggtctggt caccgtgctg cgtattgata atcgccatat cgataatgac 660
ctggcaaacg gcggtgaagc taacggcatg gatggtgcag ctctgaaaga ccaggcatca 720
ccgggcctgt ggcgcattct ggcagacat ccgggtctga tgacgtttgc cgttctggtc 780
tttctgttcc acctggcaaa tgcggccatg ctgacctcag tgtcgcaact gctgatgcgt 840
gaagttggta aagatatggc gacgtcgtg gcggcggcat gcattgtggc agcacaatg 900

gtgatggttc cggtcgcat gctggtcggc cgttttgtgg accgcattgg caccctccg 960
atctttctgc tggcattcgg cattctggca ctgcgtggta tgctgtacc gctgagcggc 1020
aatccgtggt ggctgctggg tgtgcaactg ctggatggcg ttggtgcagg catctttggc 1080
gtctgtttcc cggctgtggt tgcggacctg accaagggtta cgggccgttt taacattagc 1140
cagggtgccg tggcaaccgc tcaaggtatc ggtgctgcac tgtctgcggg cgtcgcaggt 1200
ctgattatcg tgaagcggg ctatagcgtt gcctttttct ctctggcagg tctcgtggt 1260
ctgggcttcg tgctgtacgc gctgctgatg ccggaaaccc gtcacctgta a 1311

<210> 166

<211> 1371

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Serratia odorifera*

<400> 166

atgctgacc tgaccttctg tatcaatc gctatcccg acgtggcgcg ttttatctgg 60
ccgaaagtga atgtgaaca caaagctgat attatccatc agggcatttc tcacgaaccg 120
cgtaacaaaa aatattacgt tctgctgagt ggtctgctgt tttcttttt ctttacctgg 180
agcgcgtctt ttagtctgat ttctatctgg ctgaatcaga aaatcggcct gaaagcacc 240
gaaacgggcc tgatttttag cgcatgagc atcatggccc tgtgcgtcca gccgctgtat 300

ggcttcgtgc aagacaaact gggctcgcgc aaacatctgc tgctgtttgt cgggtgtgctg 360
ctgctgctga ccggtccgtt ctttatttat gtgtacgcac cgctgctgca gacgaacctg 420
gtggcaggcg ctctggttgg cggctctgtc gtcagcctgg cgtttaatgc cggcattggt 480
gccctggaat cctataccga acgtgtttca cgcacgtcgc gcttcgaatt tggctgtgca 540
cgcatgtggg gcctctctggg ttgggcgagt gccacgttct ttgcgggttt catctacaac 600
atcaaccga acatcaattt ctggatcgca tccgcttcag cggccgtctt tctgctgctg 660
ctgtggcaag tgcgtgaact gaaaccgaac gcgatggccg gcctggaata tggtaaaccg 720

gaaaatctga aactgcaaga tgcactggct ctgctgcgtc tgccgggttt ctgggcgctg 780
gttgtgtttg ttctgggcac tcctatttat ggtgttttcg accagcaatt tccggtctac 840
ttcgcatcac agtttgctac gcatgaagaa ggcaaccgta tgtacggttt cctgaatagc 900
ctgcaagttt ttctggaagc gggcggatat ttcttgccac cgctgctggt gaaccgtatt 960
ggcatcaaac agtccctgct gctggcaagc tctgtgatgg ctttcgcat gtttggtctg 1020
ggttttgcaa atggtgctct gatgattagc gcaatgaaac tgctgcacgc tgtggaactg 1080

ccgattctgc tggttgccat gtttaaatat atcaccacgc gtttcgattc gcgcctgagt 1140

tccacctgt accctggttgg cttccagttt attagccaaa tcgtcgcagg tattctggca 1200

ccgctggcag gtctgggtta tgatcgcatc ggctttgcag acacctacct gctgatgggt 1260

tgcgcagtgg caggcaccac gctggtttct tgtttcctgc tgcgtggtga aaccccgga 1320

cgcgctccgc agtttcaaag tacgattaaa tcatcggaac cgatccagta a 1371

<210> 167

<211> 1314

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Streptomyces albus

<400> 167

atgacgggcc agtccccggc tgctccgagt ccggtgctg ctgtcccggc gcgtggctgt 60

cgccgcctga attttggctt gattagtgcg gccctgttca tgtttttcgt tacctggagc 120

ctgtcttggc gtctgttttc gatttggctg acgcaggata tcggcctgtc tccgggccgt 180

agctctctgg tgattggcgc gaactctctg ggttgcctgg tcaccatgcc ggtgtacggc 240

tttctgcaag accgtctggg tctgcgcaaa aatctgctgt attggattgg cgcctgatg 300

ctgctgggtg gcccggttta tatctacgtt tatggcccgc tgcgaaatc tcatttcgca 360

ctgggcctgg tggttggtag tgcttacctg gccatggcat ttgctgtggc ggttgccacc 420

ctggaatcct atgcggaacg tctgtcacgc ttccacggct ttgaatttgg tcgtgccgc 480

atgtttggct ctctgggttg ggcggcggca acgttctctg caggctgtct gtttaacatt 540

gatccggacc tgaccttctg ggcagcaagt ggcacggcgg tcgtgtttgt tggctctgtg 600

gtcgccatcc gtgtgaccga tggctgtcgc gctgcggccg ttgaccatgc ggcggcgagc 660

gtgagcctgg cagatgtcgc cgcactgtc cgtatccgg cgttctgggg tctgtctgtg 720

tttgtgtcgc cagtaccgc tacgtacaac acctatgatg cgatgtttcc gagctacttc 780

agttccctgt ttgcgacgga tgccgacggc aaccgcatgt attcggacct gaatagcgtc 840

caggtgttcc tggaagcagg cggatatggc ctggctccgt ttctgggtta tcgtctgggc 900

ccgaaaaaat ctctgtggt tagcggttgt gtcattggcca cccgcatttt tctgtccggt 960

atcgtgacgg atccggttgc aattagcgcg gtgaaactgc tgcatgcagt ggaactgccg 1020

atcatgctga ttgctatctt caaatcgtg aatcgtcact ttgaaccgcg cctgtcatcg 1080

accatttata tgatcggtt tcagatggcg acgcaagtgc gtgctgcggt gatcagcccc 1140

ctggcaggtc tgggttatga tcgtctgggc tacgcaccga cctatctggt gatgtcggcg 1200

ctggttgcca ccttcacgct ggtagcgctc ttaccctgc gtgcagacgc aaaaggcacg 1260
cacggtctgc cggccgtggg cgccgcagaa ggtcgctccc cggctgcggg ttaa 1314

<210> 168

<211> 1284

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Azotobacter vinelandii*

<400> 168

atgcacctgc cgtgaaacg cgaatactgg ctgatctcgg gtctgctgtt cttcttcttc 60
tttgcttggc cgtcatccta ctactgttc tcaatttggc tgcacgtgt tatcggcctg 120
tcgggaccgc aaacgggctt tattttcgcg gccaacgcca ttgcagctct gttcatccag 180
ccgttttatg gcgcactgca agaccgtctg ggtctgagcc gtgcctgct ggtctggatt 240
ggtgtgctgc tggcatgcgc agcaccgttt gcaatccatg tgtacgccgg tctgctggca 300

cacgatgtgg ttctgggcgc actggttggc gcagcttttc tggttttcgc gatgctggcg 360
ggcgtgggtg ttattgaatc atataccgaa cgtctgtcgc gccatgccgg ctttgaattt 420
ggcaccacgc gtatgtgggg ctccctgggt tgggcggccg caaccggtgt ggcaggtgtc 480
gtgttcgata ttgaccgga tatgccttt tacatgagct ctgctgcggg cctggcattc 540
ctgctgatcc tgagccgtct ggacatggat gcctttcgta accacgaaac ccgcgtggt 600
gaatctgcgc cggttcgac ggccgatatt ctgcaactgc tgcgtctgcc gcgcttttgg 660
gccttcagtc tgtttctgac cggcgtggca ggtgtttata tgatctacga acagcaattc 720

ccggtgtatt ttgcgagctt tttcccgacc ccggaagaag gcacgcgcgc ctatggttac 780
ctgaatagtt cccaagtctt ggtggaagct ctgctgatgc tggtcgcgcc gtgggttgtc 840
gcacgtaccg gtgcaaaata cggctctgtc ctggcgggca cgattatgtt tgtgcgcac 900
ctgggcagtg gtctggcttc cgacgcgtgg accattgccg catgcaaaat gctgcatgct 960
attgaagtgc cgatcctgct ggttcggctc ttcaaataca tctgtgttaa cttcgatgct 1020
cgtctgagcg cgacctgta cctggtctct tttcagttcg ccagcaact gacggcaatg 1080
ctgctgagcc cggctgttgg ctatggttac gatgcctgg gctttgcgag tgtttatctg 1140

ctgctggctt ccctggtcgg tgcgtgtctg ctgctgagct ggtctctgct gcgtgccgac 1200
ccgtcatcga aaaccacgc cgctgcggat ggtgcaccgg aactgccggc cattgcacag 1260
gccgcaaate actatgaacc gtaa 1284

<210> 169

<211> 1278

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Oenococcus oeni*

<400> 169

atgattatga tcaaccaatc cgaaattaaa accgccaaca aaaaaaaga ctctctggggc 60
ttcccgttct cgcacttcac ctacttcttt atctggcaga ttatcaacgg ttatctgacc 120

ctgtggatgg aacaaattgg ccacttgaat ggtgcggaag ccggcgtggt ttctcagtatg 180
atggccggtt ttccctgat ttccagccg tcatttggca tcactctgga taaactgac 240
ttcaagaaaa acctgatcct gaccattagt atcgaggta ttctgatcgg ccggtatttc 300
caatggctgt ttctgccgt gatgaaaatc aacagcacct ttgtcgcat taccagggc 360
atctttctga gtttcattct gaatggcggg gtctccgtga ttgaacagta tgttcaacgt 420
gcaagcctgg ctaaccattt tgaatcggg cacagccgca tgggcgggtc tctggcaggc 480
atgtgcgcta gcttctctgg cggtcaactg ttctgtgga gcccgaaacag catcttctgg 540

gcatgtacca ttctgtctat gatcctgacg ggtctgtctg tgtttaacga taaaattcac 600
atggaaaacg caaatctggc gggcggcacc agctctctgc tgaacctgaa acagatcagt 660
tccatcttca aaatcaaaaa ctctctgggtc ctgtcactgt tttatattgg cacggcgtcg 720
atctacgata tcttcgacca gcaattcatc atctttttca aatcattttt cgacaccgca 780
tcgcagggta cggcggccta ttctacatg accacggcac aaattggcat cgaatttatt 840
ctgatgtttc cgatgccgtg gattatcaac cgtatcgggtg cacgcaatgg cctgattgct 900
tatggcacca tcttggcgat tcgtattatc ggcagtgtc tgtccccgaa cctgttttgg 960

gtgattctgc tgcgcctgct ggccgggtctg gaaatgccgc tgattctggt tagcatcatg 1020
aaatacattt ctggcgctt cgatctgcgc ctgtatgcga ccatctacgc cctgtcatcg 1080
aactttgcta aacagattac cgaatttttc ctggcgaccg cggcgggcaa aatgtatgac 1140
tctatcggtt tccatcacac gtacctgatt ctgggtgtta tcgtcgccat ttttaccatc 1200
ttcacggcac tgtttctgaa aaaagaaaac ccggtgcagg caggtgaacg tgatcgcgcg 1260
ggcaatccga aagcctaa 1278

<210> 170

<211> 1290

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Lactobacillus brevis*

<400> 170

atggactccg caaaaaccga atcgaaacac ttctggggct tcccgtgtgc gcactttctcc 60
tactttctta tctgggccgt cgtgaatggc tatctgaccc tgtggatgga acaggttgcc 120
catctgaacg gcacggaaag tggcgccgtc ttttccatga tggcaggcat ttcactgac 180
tttcaaccga ttttcgggtg ggtttcggat aaactgctgt ttaagaaaaa cctgacctg 240
accattgcga tcgtgggcat tttcatcggc ccgtactttc agtgggtttt catgccgatt 300
ctgcatatca gtgcgtttct ggtggccatt gttaccggca cgtttctgag tttcattctg 360

aacggcgggt tctccgtgat cgaacagtat gtgcaacgtg cgagcctggc caatggtttt 420
gaatacgcac actcacgctg tggcggttcg gtcgcagggtg tcgtggctag cctggttgcg 480
ggtcgtattt ttctgtggaa accgaactct attttctggg cgtgcacat cgcggccatt 540
atcctgacgt ttctgtctgt gttctccgac aaagtgaacc tggataatgc ggccggcgcg 600
ggtgatacct cagactcgtg ggatatgaaa acggtcctga gtgtgtttaa actgaaaaat 660
ctgtgggttc tggccatttt ctatatgggc gcatccgtc tgtttgacgt tttcgtacg 720
caattcatca tctttttcaa aaccttttc gacacggcag ctcagggcac cctgggtgat 780

agctacatga gctctgcaca aacggctatt gaattttgcc tgatgttccc gatgccgtgg 840
attatcaaca aaatcggtc tcgtaatggt ctgctgattt atggctttat cacctgtatt 900
cgcacctcgt gtagcgtctg gtctccgaac tggatttggg ttgtcggctt ccgtctgctg 960
gcaggctcgt aaatgccgtg gctctgaccc tcaattatga aatatatcg gggtgccttt 1020
gacatccgcc tgtatgctac ggtgtacgca ctggcttcga atttcgcga acaagttagc 1080
gtctttatct tctctaccgt cgcgggccgt atgtacgatg tgatcggtta tcaacacacc 1140
tacattatca tgggcattgt ggtgtttttc atcacgtgt ttgccgtgtt tttcctgaaa 1200

aaagaagacc cgattcaggc cggatgaagt gaagatccga acgtgaaagc aaaagttgaa 1260
gcggccagtt ccaccgattc tcgcaataa 1290

<210> 171

<211> 1245

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from *Weissella paramesenteroides*

<400> 171

atggatcaga cgcagaatac gcagaaaaaa cacttttggc cgttcttcgg cacggatgtc 60
agctacttct tcatttggca gattgtgcag ggctttctgg tcctgtggct gaaacaggaa 120

gctgatctga gcggcgggtca agcggggttt gttttcagct ttctggcctt cgcattctctg 180

ttttaccagc cgggtgttcgg cattatcagc gacaaactgg ttttaagaa aaccctgctg 240

ctgtctatta cgatcgcggc cattctgac ggccggtttt tccaatggct gttcattccg 300

ctgctgcatt ttagctctct gctgggcgca attatcggcg gtttcgctct ggctgatgtc 360

ctgaacggcg gtgtggcggg tattgaaagt tacgtggaac gtgcctccct ggcaaatggc 420

tttgaatatg gtcacgctcg catgggcggg tcaatcggcg gtgcactggc atcggttcgtt 480

gcaggctctgg tctttgtgaa aaaccggttc ctggtctttt ggatttgcag tctggcgggc 540

attatcctga cctgtctggg gatcttcggg gatcgtatta actttgctaa tgcagacgcg 600

gcggaagggtg ctagttccgc accgctgaac gccaaaaacca tccgtctat cttcaaaatc 660

aaaaacttct ggatcctgag cattttcttt atgggtacgt ctgcaattta tgatgtcttt 720

gaccagcaat tcccggtgtt ttaccagacc tttttcacga gtgcggccgc aggcaccgtg 780

gcatactccc gtctgtgac caccgaaatc gccctggaag cactgctgat gattccgatg 840

ccgtggatta tcaacaaaat cggcgctaaa aatggtctga ttgcgtatgg tgttctgacc 900

ttctctcgta tcacgtgag cgccattgca ccgaacttct gggttctgac ctttgtgcgc 960

ctgctggcgt cattcgaaat gccgctgttt ctggtttcga ccatgaaata tatcgctggc 1020

gcgtttgatc tgcgtctgta tgccagatt tacgccctgg cattcaattt tgcaaaacag 1080

atctcactgt tcatttttag caccgttgcc ggcgatctgt atgactccgt cggtttccat 1140

accacgtact ttatcctggc tgcgtgggtt gcgctggtta cgattgttgc agtctttatg 1200

ctgaaaaaag aaaatccgc tcaggcgggt gaagtggcgg aataa 1245

<210> 172

<211> 1278

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Mannheimia succiniciproducens

<400> 172

atgctgatga cctcgcaaaa taaatcaac gccgtgccgt cgaacaaaa cttctacctg 60

aataaccgca actactggat cttctcaggt tatttctttg ttacttttt catcatggcg 120

acctgctatc cgtttctggg catttggctg ggtgatatca acggcctgtc aggtgaagac 180

cgtggcacgg ttttcgcaat gatgtcgttt ttgcgtctgt gttttcagcc ggttttcggc 240

tacgtcagcg ataaactggg tctgaaaaaa catctgctgt gggttctggg tatttctctg 300

ctgatctatg caccgttttt catttacatc ttgtctccgc tgctgaaagt gaatgtttgg 360

ctgggcagtc tggtagggcgg tgcataatatt ggctttgttt tccaggcagg tgctccggcg 420
tccgaagcct atatacgaacg tggtttcacgt cgctcgaaat ttgaatacgg tcgtgtccgc 480

atgtttggca tggtcggttg ggccatttgc gcaagcatcg ctggtgtcct gtatgcaacc 540
aaccgaatc tggtagtttg gctgggcagc attgcttctc tgatcctgct gctgctgatt 600
gccctggcaa aaccggaaca gacctctacg gtgcaaatcg cggaaaaact gggcgccaac 660
aaaaatccgg ttaacctgcg ccaggctttt gcgctgctga aactgccgaa attctgggcg 720
ctgctggcct atgttatggg tattgcgtgt gtctacgata tctttgacca gcaattcggc 780
aactttttca ataccttttt cgaaagccac gaacaaggca ttaaaatgtt tggttatgtc 840
accacggcgg gtgaactgct gaatgccctg attatgtttt tcgtgccgct gattatcaac 900

cgtagcggcg caaaaaatgc tctgctgatt gccggcacca tcatgagcgt gcgcattatc 960
ggcagctctt acgcgattga agcctggcat gtggttgtcc tgaaaacgct gcacatgttt 1020
gaagtcccg tctatctggt gggctctgtt aaatacatcg cgaacgtctt tgaagtgcatt 1080
ttctcagcca ccatattatc ggttgcatgc cacttcgcta aacagatcgg caatatgctg 1140
gtgtcgccgc tggtaggcgc gtggtatgat acctacgggt ttcaagacac gtacctgatt 1200
ctgggctgta tcgcggccgg ttttacgtcg ctgagtgtgt tcaccctgac gggcaaaagt 1260
ctgagttccc agtcctaa 1278

<210> 173

<211> 1257

<212> DNA

<213> Artificial Sequence

<220><223> codon optimized CDS from Aggregatibacter aphrophilus

<400> 173

atgagcaatc aaacctctc taaatcgaa tactgacca acagcaacta ctggatcttc 60
tcggcatact tcttcgcctt tttctttatt atggcaacct gccatccgtt tctgggcatt 120
tggtcgggtg atatccacgg cctgaaaggt gaaaaaatcg gttacgtgtt tagtttcac 180
tccctgtttg cgctgctgtt ccagccgatt ctgggccttc tgagcgacaa actgggtatc 240
cgtaaacatc tgctgtggct gctggcaatt ctgctgctgt tttatgctcc gtttttcac 300

tacgttttcg ctccgctgct gaaaaccaac ctgtggctgg gtgtgattgc cggcggtgca 360
tatatgggct ttgttttcca ggcgggtgcc cggcaagtg aagcctatat tgaacgtatc 420
tcccgcctgg atggctttga atacggctgt acccgctgt tcggcatgct gggttgggct 480
atttgcgcga gtatcgccgg taacctgtac agctctcaac cgaatgccgt cttttggctg 540

ggctctgcta cggcgggtgt tctgctgggt ctgattttcc tggcgaaaac cgatagtcc 600

aacacggccc aggtcgtgga caaactgggt gtgaataaat caccgatcac cctgaaacaa 660

gcgctgaaac tgttttcgct gccgcgtttc tgggcactgc tgacgtatgt tgcggtgtg 720

gcttgtgttt acgataatgt tgaccagcaa ttcggcaact tttcaaacac ctttttcgaa 780

agcaaagaac agggcatgaa atttttcgggt tatgttacca cgggcgggtga actgctgaat 840

gccacgatca tgtttttcgt cccgctgctg attaacgta tcggcgcaaa aaatgctctg 900

ctgattgcag gttaaatcat gtcgattcgc atcatgggct catcgtttgc gaccgaagcc 960

tggcatgtga ttgttctgaa aacgctgcac atgtttgaag ttccgttcta tctggtcggc 1020

gtgtttaaat acattgcgga tgtctttgaa gtgcgcttct ccgccacat ctatctggtg 1080

agctgccact tttctaaaca gattggtaac atgacacctga gcccggcagt tggcacctg 1140

tatgacatgt acggttttca atctacgtac ttcattctgg gctgtatcgc tctgaccttt 1200

acgagcggtt ctgtcttcac cctgggttaat acgaaaaaac tggtaacaa tgcgtaa 1257

<210> 174

<211> 100

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 174

cgtctaccct tgttatacct cacaccgcaa ggagacgac atgaccaata atcccccttc 60

agcacagatt aagccccggc gtgtaggctg gagctgcttc 100

<210> 175

<211> 100

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 175

gcatcaggca atgaataccc aatgcgacca gcttcttata tcagaacagc cccaacggtt 60

tatccgagta gctcaccagc catatgaata tcctccttag 100

<210> 176

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 176

atgaccaata atcccccttc ag

22

<210> 177

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 177

gcttcttata tcagaacagc c

21