



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0075813
(43) 공개일자 2020년06월26일

(51) 국제특허분류(Int. Cl.)
C12N 15/70 (2006.01) C12N 15/77 (2006.01)
C12N 9/00 (2006.01) C12P 17/16 (2006.01)
C12P 17/18 (2006.01) C12P 7/40 (2006.01)
(52) CPC특허분류
C12N 15/70 (2013.01)
C12N 15/77 (2013.01)
(21) 출원번호 10-2020-7002518
(22) 출원일자(국제) 2018년07월12일
심사청구일자 없음
(85) 번역문제출일자 2020년01월23일
(86) 국제출원번호 PCT/EP2018/068989
(87) 국제공개번호 WO 2019/012058
국제공개일자 2019년01월17일
(30) 우선권주장
17181503.8 2017년07월14일
유럽특허청(EPO)(EP)

(71) 출원인
바이오신티아 앱스
덴마크 코펜하겐 2100 포어비에바이 3
(72) 발명자
히니 한스 예스퍼
덴마크 코펜하겐 2100 포어비에바이 3 바이오신티
아 앱스
벨리 애네 피오
덴마크 코펜하겐 2100 포어비에바이 3 바이오신티
아 앱스
밀링-피터슨 닐스
덴마크 코펜하겐 2100 포어비에바이 3 바이오신티
아 앱스
(74) 대리인
리엔목특허법인

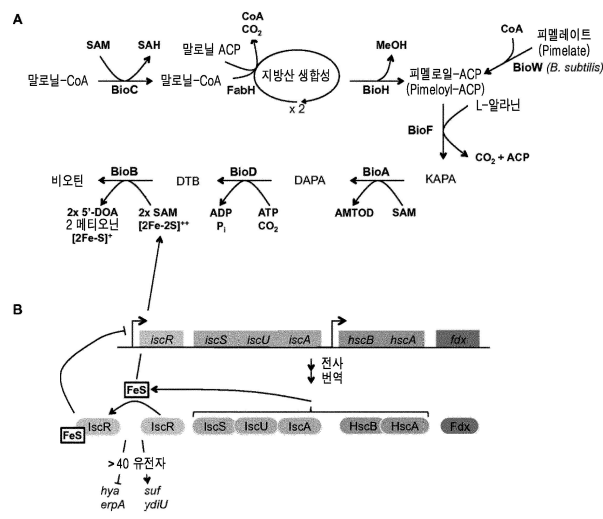
전체 청구항 수 : 총 18 항

(54) 발명의 명칭 개선된 철-황 클러스터 전달을 갖는 세포 공장

(57) 요약

본 발명은 돌연변이 철 황 클러스터 조절자 (IscR)를 코딩하는 변형된 유전자뿐만 아니라 비오틴, 리포산 또는 티아민의 생합성을 증가시키는 폴리펩티드를 코딩하는 하나 이상의 전이 유전자를 특징으로 하는, 개선된 철-황 클러스터 전달이 가능한 유전자 변형 박테리아 세포를 제공한다. 본 발명은 본 발명의 유전자 변형 박테리아를 사용하여 비오틴, 리포산 또는 티아민을 생산하는 방법뿐만 아니라; 비오틴, 리포산 또는 티아민 생산을 위한 유전자 변형 박테리아 세포의 사용을 제공한다.

대표도 - 도1



(52) CPC특허분류

C12N 9/00 (2013.01)

C12Y 201/01197 (2013.01)

C12Y 203/01047 (2013.01)

C12Y 206/01062 (2013.01)

C12Y 208/01006 (2013.01)

C12Y 208/01008 (2013.01)

C12Y 301/01085 (2013.01)

C12Y 401/99017 (2013.01)

C12Y 401/99019 (2013.01)

명세서

청구범위

청구항 1

하기를 포함하는, 비오틴 또는 리포산 또는 티아민의 증가된 생산을 위한 유전자 변형 박테리아:

a) 돌연변이 IscR 폴리펩티드를 코딩하는 유전자 변형 내인성 *iscR* 유전자로서,

상기 돌연변이 IscR 폴리펩티드의 아미노산 서열은 서열번호 2, 4, 6, 8, 10, 12 및 14로 이루어진 군으로부터 선택된 서열과 적어도 80% 아미노산 서열 상동성을 갖는 것이고, 상기 아미노산 서열은 i) L15X, C92X, C98X, C104X, 및 H107X로 이루어진 군으로부터 선택된 적어도 하나의 아미노산 치환을 갖는 것이고; 상기 X는 서열번호 2, 4, 6, 8, 10, 12 및 14에 상응하는 아미노산 잔기 이외의 임의의 아미노산인 것인, 유전자; 및

b) 적어도 하나의 전이 유전자로서,

ii) 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드;

iii) 리포산 신타아제 (EC2.8.1.8) 활성을 갖는 폴리펩티드;

iv) HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드; 및

v) 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 적어도 하나의 전이 유전자.

청구항 2

청구항 1에 있어서, 상기 돌연변이 IscR 폴리펩티드에서 상기 적어도 하나의 아미노산 치환은 다음으로 이루어진 군으로부터 선택되는 것인, 비오틴 또는 리포산 또는 티아민의 증가된 생산을 위한 유전자 변형 박테리아:

a. L15X, 상기 X는 F, Y, M 및 W 중 어느 하나임;

b. C92X, 상기 X는 Y, A, M, F 및 W 중 어느 하나임;

c. C98X, 상기 X는 A, V, I, L, F 및 W 중 어느 하나임;

d. C104X, 상기 X는 A V, I, L, F 및 W 중 어느 하나임; 및

e. H107X; 상기 X는 A, Y, V, I, 및 L 중 어느 하나임.

청구항 3

청구항 1 또는 2에 있어서, 상기 비오틴 신타아제 활성을 갖는 폴리펩티드 (EC 2.8.1.6)를 코딩하는 적어도 하나의 전이 유전자는 다음으로 이루어진 군으로부터 선택된 하나 이상의 폴리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함하고:

a. SAM (S-아데노실메티오닌)-의존적 메틸트랜스퍼라제 (EC 2.1.1.197) 활성을 갖는 폴리펩티드;

b. 7-케토-8-아미노펠라르곤산 (KAPA) 신타아제 (EC 2.3.1.47) 활성을 갖는 폴리펩티드;

c. 7,8-디아미노펠라르곤산 (DAPA) 신타아제 (EC:2.6.1.62 또는 EC2.6.1.105) 활성을 갖는 폴리펩티드; 및

d. 데티오비오틴 (DTB) 신타아제 (EC 6.3.3.3) 활성을 갖는 폴리펩티드;

e. 피멜로일-[아실-운반 단백질] 메틸 에스테르 에스테라제 (EC 3.1.1.85)를 갖는 폴리펩티드; 및

f. 6-카복시헥사노에이트-CoA 리가제 (EC 6.2.1.14) 활성을 갖는 폴리펩티드

상기 박테리아는 비오틴의 증가된 생산을 위한 것인 유전자 변형 박테리아.

청구항 4

청구항 1 또는 2에 있어서, 상기 적어도 하나의 전이 유전자는 리포산 신타아제 (EC 2.8.1.8) 활성을 갖는 폴리

펩티드를 코딩하고, 다음으로 이루어진 군으로부터 선택된 하나 이상의 폴리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함하는 것이고:

- a. 옥타노일트랜스퍼라제 (EC 2.3.1.181) 활성을 갖는 폴리펩티드;
- b. 피루브산 탈수소효소 (EC 2.3.1.12)의 디하이드로리포일라이신-잔기 아세틸트랜스퍼라제 성분을 포함하는 폴리펩티드; 및

- c. 리포에이트-단백질 리가아제 A (EC:6.3.1.20) 활성을 갖는 폴리펩티드;

상기 박테리아는 리포산의 증가된 생산을 위한 것인 유전자 변형 박테리아.

청구항 5

청구항 1 또는 2에 있어서, 상기 HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드를 코딩하는 적어도 하나의 전이 유전자 및/또는 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드를 코딩하는 적어도 하나의 전이 유전자는 다음으로 이루어진 군으로부터 선택된 하나 이상의 폴리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함하는 것이고:

- a. ThiS 아데닐트랜스퍼라제 (EC2.7.7.73) 활성을 갖는 ThiF 폴리펩티드;
- b. 티아민 포스페이트 신타아제 (EC 2.5.1.3) 활성을 갖는 ThiE 폴리펩티드;
- c. 티아졸 신타아제 활성 (E.C.2.8.1.10) 활성을 갖는 ThiG 폴리펩티드;
- d. 포스포하이드록시메틸피리미딘 키나아제 (EC 2.7.4.7) 활성을 갖는 ThiD 폴리펩티드;
- e. 글리신 옥시다아제 (EC 1.4.3.19) 활성을 갖는 ThiO 폴리펩티드;
- f. 황-운반 단백질 활성을 갖는 ThiS 폴리펩티드;
- g. 하이드록시에틸티아졸 키나아제 (EC2.7.1.50) 활성을 갖는 ThiM 폴리펩티드; 및
- h. 모노-포스페이트 포스파타아제 (E.C. 3.1.3.-) 활성을 갖는 폴리펩티드;

상기 박테리아는 티아민의 증가된 생산을 위한 것인 유전자 변형 박테리아.

청구항 6

청구항 5에 있어서, 상기 박테리아는 다음을 코딩하는 추가적인 전이 유전자를 포함하는 것인 유전자 변형 박테리아:

- a. HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 ThiC 폴리펩티드;
- b. 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 ThiH 폴리펩티드 또는 글리신 옥시다아제 (EC 1.4.3.19) 활성을 갖는 ThiO 폴리펩티드;
- c. ThiS 아데닐트랜스퍼라제 (EC 2.7.7.73) 활성을 갖는 ThiF 폴리펩티드;
- d. 티아민 포스페이트 신타아제 (EC 2.5.1.3) 활성을 갖는 ThiE 폴리펩티드;
- e. 티아졸 신타아제 (E.C.2.8.1.10) 활성을 갖는 ThiG 폴리펩티드;
- f. 포스포하이드록시메틸피리미딘 키나아제 (EC 2.7.4.7) 활성을 갖는 ThiD 폴리펩티드;
- g. 황-운반 단백질 활성을 갖는 ThiS 폴리펩티드; 및
- h. 티아민 모노-포스페이트 포스파타아제 (E.C. 3.1.3.-) 활성을 갖는 폴리펩티드.

청구항 7

청구항 1 내지 6 중 어느 한 항에 있어서, 상기 적어도 하나의 전이 유전자 및 상기 하나 이상의 추가적인 전이 유전자는 항상성 프로모터(constitutive promoter)에 작동 가능하게 연결된 것인 유전자 변형 박테리아.

청구항 8

청구항 1 내지 7 중 어느 한 항에 있어서, 상기 박테리아는 에서리키아 (*Escherichia*), 바실러스 (*Bacillus*), 브레비박테리움 (*Brevibacterium*), 버크홀데리아 (*Burkholderia*), 캄필로박터 (*Campylobacter*), 코리네박테리움 (*Corynebacterium*), 슈도모나스 (*Pseudomonas*), 셀라티아 (*Serratia*), 락토바실러스 (*Lactobacillus*), 락토코커스 (*Lactococcus*), 아시네토박터 (*Acinetobacter*), 슈도모나스 (*Pseudomonas*), 및 아세토박터 (*Acetobacter*)로 이루어진 군으로부터 선택된 박테리아의 속인 것인 유전자 변형 박테리아.

청구항 9

하기 단계를 포함하는 비오틴을 생산하는 방법:

- 청구항 1 내지 3, 7 및 8 중 어느 한 항에 따른 유전자 변형 박테리아를 배양물을 생산하기 위한 증식 배지 내로 도입하는 단계로서, 여기에서 상기 적어도 하나의 전이 유전자는 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드를 코딩하는 것인 단계;
- 상기 배양물을 배양하는 단계; 및
- 상기 배양에 의해 생산된 비오틴을 회수하고, 선택적으로 회수된 비오틴을 정제하는 단계.

청구항 10

리포산을 생산하는 방법으로서,

- 청구항 1, 2, 4, 7 및 8 중 어느 한 항에 따른 유전자 변형 박테리아를 배양물을 생산하기 위한 증식 배지 내로 도입하는 단계로서, 여기에서 상기 적어도 하나의 전이 유전자는 리포산 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드를 코딩하는 것인 단계;
- 상기 배양물을 배양하는 단계; 및
- 상기 배양에 의해 생산된 리포산을 회수하고, 선택적으로 회수된 리포산을 정제하는 단계를 포함하는 방법.

청구항 11

티아민을 생산하는 방법으로서,

- 청구항 1, 2, 및 5 내지 8 중 어느 한 항에 따른 유전자 변형 박테리아를 배양물을 생산하기 위한 증식 배지 내로 도입하는 단계로서, 여기에서 상기 적어도 하나의 전이 유전자는 HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드 및/또는 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드를 코딩하는 것인 단계;
- 상기 배양물을 배양하는 단계; 및
- 상기 배양에 의해 생산된 티아민을 회수하고, 선택적으로 회수된 티아민을 정제하는 단계를 포함하는 방법.

청구항 12

청구항 9 내지 11 중 어느 한 항에 있어서, 상기 증식 배지는 글루코스, 말토스, 갈락토스, 프럭토스, 수크로스, 아라비노스, 자일로스, 라피노스, 만노스, 및 락토스, 또는 이들의 임의의 조합으로부터 선택된 탄소원을 포함하는 것인, 비오틴, 리포산 및 티아민 중 어느 하나를 생산하는 방법.

청구항 13

유전자 변형 박테리아에서 비오틴, 리포산 또는 티아민 중 어느 하나의 생산을 증가시키기 위한 돌연변이 *iscR* 폴리펩티드를 코딩하는 유전적으로 변형된 유전자의 용도로서, 상기 박테리아는:

- 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드;
- 리포산 신타아제 (EC 2.8.1.8) 활성을 갖는 폴리펩티드;
- HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드; 및
- 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드로 이루어진 군으로부터 선택된 폴리펩티드를 코딩하는 적어도 하나의 전이유전자를 포함하고 발현하는 것이고;

상기 유전적으로 변형된 유전자는 돌연변이 *IscR* 폴리펩티드를 코딩하는 내인성 *iscR* 유전자이고, 여기에서 상

기 돌연변이 IscR 폴리펩티드의 아미노산 서열은 서열번호 2, 4, 6, 8, 10, 12 및 14와 적어도 80% 아미노산 서열 상동성을 갖는 것이고, 상기 아미노산 서열은: L15X, C92X, C98X, C104X, 및 H107X로 이루어진 군으로부터 선택된 적어도 하나의 아미노산 치환을 갖는 것이고; 상기 X는 서열번호 2, 4, 6, 8, 10, 12 및 14에 상응하는 아미노산 잔기 이외의 임의의 아미노산인 것인 용도.

청구항 14

청구항 13에 있어서, 상기 돌연변이 IscR 폴리펩티드에서 상기 적어도 하나의 아미노산 치환은 다음으로 이루어진 군으로부터 선택되는 것인, 유전자 변형 박테리아에서 비오틴, 리포산 또는 티아민 중 어느 하나의 생산을 증가시키기 위한 돌연변이 iscR 폴리펩티드를 코딩하는 유전적으로 변형된 유전자의 용도:

- a. L15X, 상기 X는 F, Y, M 및 W 중 어느 하나임;
- b. C92X, 상기 X는 Y, A, M, F 및 W 중 어느 하나임;
- c. C98X, 상기 X는 A, V, I, L, F 및 W 중 어느 하나임;
- d. C104X, 상기 X는 A, V, I, L, F 및 W 중 어느 하나임; 및
- e. H107X; 상기 X는 A, Y, V, I, 및 L 중 어느 하나임.

청구항 15

비오틴, 리포산 또는 티아민 중 어느 하나의 증가된 생산을 위한, 청구항 1 내지 8 중 어느 한 항에 따른 유전자 변형 박테리아의 용도.

청구항 16

청구항 1 내지 8 중 어느 한 항에 있어서, 상기 박테리아는 하기 군으로부터 선택된 하나 이상의 유전자를 더 포함하는 것인 유전자 변형 박테리아:

- a. 플라보독신/페레독신-NADP 환원 효소 (EC:1.18.1.2 및 EC 1.19.1.1)를 코딩하는 유전자;
- b. 피루브산-플라보독신/페레독신 산화 환원 효소 (EC 1.2.7)를 코딩하는 유전자;
- c. 플라보독신을 코딩하는 유전자;
- d. 페레독신을 코딩하는 유전자; 및
- e. 플라보독신 및 페레독신-NADP 환원 효소를 코딩하는 유전자;

여기서, 상기 하나 이상의 유전자는 상기 박테리아에서 상기 하나 이상의 유전자의 발현을 증가시킬 수 있는 비-천연(non-native) 프로모터에 작동 가능하게-연결되는 것이고, 상기 하나 이상의 유전자는 천연 유전자(native gene) 또는 전이 유전자일 수 있다.

청구항 17

비오틴, 리포산 또는 티아민 중 어느 하나의 증가된 생산을 위한, 청구항 16에 따른 유전자 변형 박테리아의 용도.

청구항 18

청구항 9 내지 12 중 어느 한 항에 있어서, , 상기 유전자 변형 박테리아는 다음의 군으로부터 선택된 하나 이상의 유전자를 더 포함하는 것인 비오틴, 리포산 및 티아민 중 어느 하나를 생산하는 방법:

- a. 플라보독신/페레독신-NADP 환원 효소 (EC:1.18.1.2 및 EC 1.19.1.1)를 코딩하는 유전자;
- b. 피루브산-플라보독신/페레독신 산화 환원 효소 (EC 1.2.7)를 코딩하는 유전자;
- c. 플라보독신을 코딩하는 유전자;
- d. 페레독신을 코딩하는 유전자; 및
- e. 플라보독신 및 페레독신-NADP 환원 효소;

상기 하나 이상의 유전자는 상기 박테리아에서 상기 하나 이상의 유전자의 발현을 증가시킬 수 있는 비-천연 프로모터에 작동 가능하게-연결된 것이고, 상기 하나 이상의 유전자는 천연 유전자 또는 전이 유전자일 수 있다.

발명의 설명

기술 분야

[0001] 본 발명은 돌연변이 철 황 클러스터 조절자 (IscR)를 코딩하는 변형된 유전자뿐만 아니라 비오틴, 리포산 또는 티아민의 생합성을 증가시키는 폴리펩티드를 코딩하는 하나 이상의 전이 유전자를 특징으로 하는, 개선된 철-황 클러스터 전달이 가능한 유전자 변형 박테리아 세포에 관한 것이다. 본 발명은 또한 본 발명의 유전자 변형 박테리아를 사용하여 비오틴, 리포산 또는 티아민을 생산하는 방법뿐만 아니라; 비오틴, 리포산 또는 티아민 생산을 위한 유전자 변형 박테리아 세포의 사용에 관한 것이다.

배경 기술

[0002] 비오틴 (비타민 B7 또는 비타민 H라고도 알려짐), 및 티아민 (비타민 B2이라고도 알려짐)은 사람에게 필수적인식이 비타민인데, 이는 다른 후생동물들과 마찬가지로, 그들은 비오틴 또는 티아민을 생산할 수 없기 때문이다. 리포산 (LA: Lipoic acid)은 황-함유, 비타민-유사 항산화제이고, 박테리아, 식물 및 동물에서 소량으로 합성된다. 세가지 모두 식이 보충제로 널리 사용된다. 이들 비타민 또는 비타민-유사 화합물들의 생산은 현재 화학적 합성에 의존하는데, 이는 비용이 많이 든다. 이들의 제조를 위한 생합성 방법은 현재 및 미래 요구를 충족시키기 위한 대안적이고, 보다 비용효율적인 방법을 제공할 것이다.

[0003] 비오틴은 모든 생물 형태에 존재하는 아세틸-CoA 카복실라제 (ACC: acetyl-CoA carboxylase)와 같은 특정 카르복실화 반응을 촉매시키는 효소의 필수적인 보조인자이고, 지방산 생합성의 중요한 구성 요소인 말로닐-CoA를 생산한다. 사실상, 비오틴은 지방산 생합성 경로와 관련된 선형 경로에 의해 합성된다. 대장균 (*E. coli*: *Escherichia coli*)에서 비오틴 생합성의 초기 기질은 지방산 합성의 개시 대사 산물이기도 한 말로닐-ACP이다. 지방산 사이클을 들어가기 전에, 말로닐-ACP는 SAM (S-아데노실메티오닌)-의존적 메틸트랜스퍼라제, BioC에 의해 마스킹되어, 말로닐-ACP 메틸 에스테르를 생성한다. 이어서, 2회의 지방산 사슬 신장은 피멜로일-에스테르-ACP 분자를 생성한다. 전용 에스테라제, BioH에 의한 피멜로일-에스테르-ACP의 O-메틸기의 가수분해는 이들 분자를 지방산 신장 사이클을 빠져나가게 한다. 이어서, 중간체(intermediate), 피멜로일-에스테르-ACP는 비오틴-특이적 경로를 통해 비오틴으로 전환된다 (도 1A). 이 경로에서, BioF는 피멜로일-ACP와 알라닌의 PLP-의존적 탈탄산 알돌 축합(decarboxylative aldol condensation)을 촉매하여 KAPA (8-아미노-7-옥소노나노에이트)를 생산한다. BioA (및 BioK)는 KATA의 PLP-의존적 아미노기 전이반응을 촉매하여 DAPA (7,8-디아미노펠라르코네이트)를 생산하고, 여기에서 공여자는 SAM이고; 부산물은 S-아데노실옥소메티오닌이다. BioD는 ATP-구동 카르복실화 및 DAPA의 고리 폐쇄를 촉매하여 데스티오비오틴 (DTB: desthiobiotin)에서 티오펜(thiophane) 고리를 형성한다. 비오틴 합성 경로의 최종 단계는 BioB (비오틴 신타아제)에 의해 2개의 탄화수소 사이에 황 다리 결합(sulfur bridge)의 도입과 관련되어 있으므로, 비오틴을 생산하기 위한 알려진 가장 복잡한 반응 중 하나이다. BioB는 이합체로 발견되는, S-아데노실-L-메티오닌 (SAM 또는 AdoMet) 라디칼 효소이고, 이의 활성 부위에 2개의 철-황 클러스터: $[2Fe-2S]^{2+}$ 및 $[4Fe-4S]^{2+}$ 를 포함한다. DTB에서 티오펜 고리를 생성하는데 필요한 황 원자는 BioB에서 $[2Fe-2S]^{2+}$ 클러스터로부터 조달된 것으로 여겨진다. 결과적으로, DTB 합성에 소비되는 BioB 이합체에서 철-황 클러스터는 촉매반응의 각 라운드 후에 재생산되는 것으로 생각된다.

[0004] 리포산 (LA: Lipoic acid)는 활성 산소종의 강력한 스캐빈저(scavenger)일 뿐만 아니라, 이에 따라 중요한 항산화제이고, 또한 α -케토산 탈수소효소에 대한 보조인자이다. LA는 지방산 대사과정에서 중간체로부터 새롭게 합성된다 (도 2). *E. coli*에서 LA 합성에 참여하는 3개의 효소는 Lp1A (리포산-단백질 리가아제), LipB (옥탄오일 단백질 ACP 운반 단백질: 단백질 트랜스퍼라제), 및 LipA (리포산 신타아제)이다. *lp1A* 유전자에 의해 코딩되는 Lp1A는 ATP-의존적인 방식으로 타겟 효소의 E2 서브유닛의 비리포일화된-아포-리포일 도메인 (unlipoylated-apo-lipoyl domain)에 외인성 옥탄산의 컨주게이션을 촉매할 수 있다. *lipB* 유전자에 의해 코딩되는 LipB는 ACP로부터의 옥타닐 잔기의 타겟 효소의 E2 서브유닛의 아포-리포일 도메인으로의 이동을 촉매할 수 있다. *AceF* 유전자는 피루브산 탈수소효소의 E2 서브유닛의 리포일 도메인을 코딩한다. *lipA* 유전자에 의해

코딩되는 LipA는 2개의 C-S 결합의 형성을 담당한다. LipA-유도 반응은 이의 기능을 수행하기 위해 철-황 클러스터 (4Fe-4S) 및 SAM (metK 유전자에 의해 생성됨)을 필요로 한다. 리포산은 주로 다수의 다중-효소 복합체에 서 단백질-결합된 리포아미드(lipoamide) 모이어티로서 세포에서 발견된다.

[0005] 티아민 생합성은 박테리아, 일부 원생동물, 식물, 및 진균에서 특성화되었다. 티아민의 티아졸 및 피리미딘 모이어티는 별도로 합성된다 (도 3). 피리미딘 모이어티, 4-아미노-5-하이드록시메틸-2-메틸피리미딘 포스페이트 (HMP-P)는 드 노보(de novo) 퓨린 생합성 경로에서 중간체인 5-아미노이미다졸 리보타이드 (AIR: aminoimidazole ribotide)로부터 유래된다. 그람-음성 박테리아에서, AIR의 HMP-P로의 전환은 서브유닛 당 1개의 [4Fe-4S] 클러스터에 결합하는 *thiC* 유전자 산물인 HMP-P 신타제에 의한 라디칼 S-아테노실-L-메티오닌 (SAM)-의존적 반응에서 촉매된다.

[0006] 이어서, HMP-P는 티아졸 유닛과의 커플링 전에 ThiD 키나아제에 의해 HMP-PP로 인산화된다. 티아졸 모이어티, 5-(2-하이드록시메틸)-4-메틸티아졸 포스페이트 (HET-P)는 L-티로신 및 1-데옥시-D-자일로스 포스페이트 (DXP) 및 시스테인으로부터 유래되고; 여기에서 황 원자는 L-시스테인으로부터 유래되는 것으로 예상된다. *thiH* 유전자에 의해 코딩되는 티로신 리아제(Tyrosine lyase)는 서브유닛 당 1개의 [4Fe-4S]에 결합하고, 티로신의 2-이미노아세테이트 및 4-크레솔로의 라디칼-매개 절단을 촉매작용한다. 티아졸 모이어티의 합성은 적어도 5개의 유전자 *thiF*, *this*, *thiG*, *thiH* 및 *thiI*의 발현을 필요로 한다.

[0007] 이후, 피리미딘 및 티아졸 모이어티는 *thiE*에 의해 코딩되는 티아민-포스페이트 신타아제 (EC 2.5.1.3)의 작용에 의해 TMP를 형성하도록 조합된다. 따라서, TMP는 모든 공지된 티아민 생합성 경로의 첫번째 산물이다. *대장균* 및 다른 장내 세균에서, TMP는 ATP의 존재 하에 *thiL*에 의해 코딩되는 티아민-포스페이트 키나아제 (EC 2.7.4.16)에 의해 보조인자 TPP로 인산화될 수 있다. 티아민 모노-포스페이트 포스파타제 (E.C 3.1.3.-)를 발현하는 전이 유전자를 포함하는 박테리아 균주들은 TMP를 티아민으로 전환시킬 수 있고, 이로써 티아민 생산을 증가시킬 수 있다.

[0008] 박테리아-기반의 세포 공장의 사용은 바이오, 리포산 및 티아민의 생합성 생산을 위한 잠재적 경로이다. 바이오-제품의 생산을 위한 세포 공장으로서 재조합 *대장균*의 이점은 다음과 같은 사실로 인해 널리 인식된다: (i) 이는 글루코스-염 배지에서 배양될 경우 및 최적의 환경 조건 하에서 약 20분의 배가 시간(doubling time)을 가진 비할데 없는 빠른 성장 속도를 갖는다, (ii) 높은 세포 밀도를 쉽게 달성한다; 여기에서 *대장균* 액체 배양물의 이론적 밀도 한계는 약 200 g 건조 세포 중량/l 또는 대략 1×10^{13} 생존 가능한(viable) 박테리아/mL로 추정된다. 또한, *대장균*은 이종 단백질의 발현에 다루기 쉬운(amenable) 유기체일 뿐만 아니라; *대장균*의 유전자 변형을 위한 많은 분자 도구 및 사용 가능한 프로토콜이 있다; 이들 모두는 원하는 바이오-제품의 높은-수준의 생산을 수득하기 위해 필수적일 수 있다.

[0009] *대장균*에서, 바이오틴 오페론 구조는 반대 가닥(bioO 유전자좌) 상의 중복 프로모터의 조절 하에 *bioA* 및 *bioBFC*로 나뉘지는데, *bioH*는 *대장균* 염색체의 다른 곳에 위치한다. 바이오틴 오페론의 발현은 바이오틴-결합 억제자(BirA: biotin-bound repressor)에 의해 하향-조절되고; 바이오틴-결합 억제자는 바이오틴 오페론에서 오퍼레이터(operator)에 결합한다. BirA은 또한 바이오틴을 세포의 카복실라제로 전달하는 바이오틴 리가아제로서 기능을 한다. 바이오틴 리가아제로부터 전사 억제자로의 BirA의 기능 전환은 각각의 세포 내 바이오틴 및 아포-카복실라제 풀에 의해 조절된다. *대장균*에서 바이오틴 오페론 (*bioA* 및 *bioBFC*)의 과-발현은 성장을 저해하는 것으로 보고되었다 (Ifuku, O. et al., 1995). 이 저해의 원인은 알려져 있지 않았고, 이는 바이오틴 합성을 증가시키는 장애물을 형성한다.

[0010] 일반적으로, 박테리아-기반의 세포 공장 (예를 들어, *대장균*)에서 바이오틴, 리포산 및 티아민의 생산을 용이하게 하기 위해 이러한 복잡한 생합성 경로들의 병목 현상을 규명할 필요가 있고, 박테리아-기반의 세포 공장은 이들 각각의 경로 효소들의 증가된 수준을 성장 및 생산하는 이들의 능력을 제한할수 있는 다양한 원인을 극복하도록 맞춤-제작된다.

발명의 내용

[0011] 발명의 요약

[0012] 일 양상에 따르면, 본 발명은 바이오틴, 리포산 또는 티아민 중 어느 하나의 증가된 생산을 위해 유전자 변형 박테리아를 제공하며; 상기 박테리아는:

[0013] · 돌연변이 IscR 폴리펩티드를 코딩하는 유전자로 변형 내인성 *iscR* 유전자로서, 상기 돌연변이 IscR 폴리펩티

드의 아미노산 서열은 서열번호 2, 4, 6, 8, 10, 12 및 14와 80% 서열 상동성을 갖는 것이고, 상기 아미노산 서열은 다음으로 이루어진 군으로부터 선택된 적어도 하나의 아미노산 치환을 갖는 것인:

- [0014] o L15X, C92X, C98X, C104X, and H 107X; 상기 X는 서열번호 2, 4, 6, 8, 10, 12 및 14에 상응하는 아미노산 잔기 이외의 임의의 아미노산인 것인 유전자, 및
- [0015] · 이들 중 선택된 폴리펩티드를 코딩하는 적어도 하나의 전이유전자:
- [0016] o 증가된 비오틴 생산을 위한 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드,
- [0017] o 증가된 리포산 생산을 위한 리포산 신타아제 (EC 2.8.1.8) 활성을 갖는 폴리펩티드,
- [0018] o 증가된 티아민 생산을 위한 HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드, 및
- [0019] o 증가된 티아민 생산을 위한 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드.
- [0020] 바람직하게, 상기 돌연변이 IscR 폴리펩티드 내 적어도 하나의 아미노산 치환은 다음으로 이루어진 군으로부터 선택된다:
- [0021] o L15X, 상기 X는 F, Y, M 및 W 중 어느 하나임;
- [0022] o C92X, 상기 X는 Y, A, V, I, G, L, M, F 및 W 중 어느 하나임;
- [0023] o C98X, 상기 X는 A, V, I, G, L, F 및 W 중 어느 하나임;
- [0024] o C104X, 상기 X는 A V, I, G, L, F 및 W 중 어느 하나임; 및
- [0025] o H 107X; 상기 X는 A, Y, M, F, W, V, I, G, 및 L 중 어느 하나임.
- [0026] 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드를 코딩하는 하나의 전이 유전자를 포함하는 비오틴의 증가된 생산을 위한 본 발명에 따른 유전자 변형 박테리아는 다음으로 이루어진 군으로부터 선택된 하나 이상의 폴리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함할 수 있다:
- [0027] o SAM (S-아데노실메티오닌)-의존적 메틸트랜스퍼라제 (BioC; EC 2.1.1.197) 활성을 갖는 폴리펩티드;
- [0028] o 7-케토-8-아미노펠라르곤산 (KAPA) 신타아제 (BioF; EC 2.3.1.47) 활성을 갖는 폴리펩티드;
- [0029] o 7,8-디아미노펠라르곤산 (DAPA) 신타아제 (BioA; EC:2.6.1.62) 또는 L-리신:8-아미노-7-옥소노나노에이트 아미노트랜스퍼라제(BioK; EC:2.6.1.105) 활성을 갖는 폴리펩티드;
- [0030] o 데티오비오틴 (dethiobiotin, DTB) 신타아제 (BioD; EC 6.3.3.3) 활성을 갖는 폴리펩티드, 및
- [0031] o 피멜로일-[아실-운반 단백질] 메틸 에스테르 에스테라제 (BioH; EC 3.1.1.85)를 갖는 폴리펩티드 또는 6-카복시헥사노에이트-CoA 리가제 (BioW; EC 6.2.1.14) 활성을 갖는 폴리펩티드.
- [0032] 바람직하게, 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드를 코딩하는 하나의 전이 유전자를 포함하는 비오틴의 증가된 생산을 위한 본 발명에 따른 유전자 변형 박테리아는 SAM (S-아데노실메티오닌)-의존적 메틸트랜스퍼라제 (BioC; EC 2.1.1.197) 활성; 7-케토-8-아미노펠라르곤산 (KAPA) 신타아제 (BioF; EC 2.3.1.47) 활성; 및 7,8-디아미노펠라르곤산 (DAPA) 신타아제 (BioA; EC:2.6.1.62) 활성을 갖는 폴리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함할 수 있다.
- [0033] 리포산 신타아제 (EC 2.8.1.8) 활성을 갖는 폴리펩티드를 코딩하는 하나의 전이 유전자를 포함하는 리포산의 증가된 생산을 위한 본 발명에 따른 유전자 변형 박테리아는 다음으로 이루어진 군으로부터 선택된 하나 이상의 폴리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함할 수 있다:
- [0034] o 옥타노일트랜스퍼라제 (EC 2.3.1.181) 활성을 갖는 폴리펩티드, 및
- [0035] o 피루브산 탈수소효소 (EC 2.3.1.12)의 디하이드로리포일라이신-잔기 아세틸트랜스퍼라제 성분을 포함하는 폴리펩티드, 및
- [0036] o 리포에이트-단백질 리가아제 A (EC:6.3.1.20) 활성을 갖는 폴리펩티드.
- [0037] HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 ThiC 폴리펩티드를 코딩하는 하나의 전이 유전자, 및/또는 티로신 리아제(EC 4.1.99.19) 활성을 갖는 ThiH 폴리펩티드를 코딩하는 하나의 전이 유전자를 포함하는 티아민의 증가된 생산을 위한 본 발명에 따른 유전자 변형 박테리아는 다음으로 이루어진 군으로부터 선택된 하나 이상의 폴

리펩티드를 코딩하는 추가적인 전이 유전자를 더 포함할 수 있다:

- [0038] o ThiS 아테닐트랜스퍼라제 (EC 2.7.7.73) 활성을 갖는 ThiF 폴리펩티드;
- [0039] o 티아민 포스페이트 신타아제 (EC 2.5.1.3) 활성을 갖는 ThiE 폴리펩티드;
- [0040] o 티아졸 신타아제 (E.C.2.8.1.10) 활성을 갖는 ThiG 폴리펩티드;
- [0041] o 포스포하이드록시메틸피리미딘 키나아제 (EC 2.7.4.7) 활성을 갖는 ThiD 폴리펩티드;
- [0042] o 황-운반 단백질 활성을 갖는 ThiS 폴리펩티드;
- [0043] o 모노-포스페이트 포스파타제 (E.C. 3.1.3.-) 활성을 갖는 폴리펩티드; 및
- [0044] o 글리신 옥시다아제 (EC 1.4.3.19) 활성을 갖는 ThiO 폴리펩티드; 및 선택적으로는, 하이드록시메틸티아졸 키나아제 (2.7.1.50) 활성을 갖는 ThiM 폴리펩티드를 코딩하는 추가적인 전이 유전자.
- [0045] 바람직하게, 상기 유전자 변형 박테리아는 다음의 폴리펩티드를 코딩하는 전이 유전자를 포함한다: ThiC (*thiC* 유전자에 의해 코딩됨); ThiD (*thiD* 유전자에 의해 코딩됨), ThiE (*thiE* 유전자에 의해 코딩됨), ThiF (*thiF* 유전자에 의해 코딩됨), sulfur-carrier protein (*thiS* 유전자에 의해 코딩됨), ThiG (*thiG* 유전자에 의해 코딩됨), TMP phosphatase (TMP 포스파타제 유전자에 의해 코딩됨); 및 ThiH (*thiH* 유전자에 의해 코딩됨) 또는 ThiO (*thiO* 유전자에 의해 코딩됨). 구현예에 따르면, 상기 세포는 효소 ThiM (*ThiM* 유전자에 의해 코딩됨)을 코딩하는 전이 유전자를 더 포함할 수 있다.
- [0046] 바람직하게 본 발명의 유전자 변형 박테리아에서 상기 적어도 하나의 전이 유전자 및 상기 하나 이상의 추가적인 전이 유전자는 항상성 프로모터(constitutive promoter)에 작동 가능하게 연결된 것이다 (상기 프로모터는 전이 유전자를 포함한 오페론에 작동 가능하게 연결된 것일 수 있음).
- [0047] 본 발명의 유전자 변형 박테리아는 바람직하게 에서리키아 (*Escherichia*), 바실러스 (*Bacillus*), 브레비박테리움 (*Brevibacterium*), 버크홀데리아 (*Burkholderia*), 캄필로박터 (*Campylobacter*), 코리네박테리움 (*Corynebacterium*), 슈도모나스 (*Pseudomonas*), 셀라티아 (*Serratia*), 락토바실러스 (*Lactobacillus*), 락토코커스 (*Lactococcus*), 아시네토박터 (*Acinetobacter*), 슈도모나스 (*Pseudomonas*), 및 아세토박터 (*Acetobacter*)로 이루어진 군으로부터 선택된 속의 중, 보다 바람직하게 에서리키아 또는 코리네박테리움의 종은, 예를 들어, *에서리키아 콜라이* (*Escherichia coli*) 또는 코리네박테리움 글루타미쿰 (*Corynebacterium glutamicum*)이다.
- [0048] 제2 구현예에 따르면, 본 발명은 비오틴을 생산하는 방법으로서:
- [0049] o 본 발명에 따른 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드를 코딩하는 전이 유전자를 포함하는 유전자 변형 박테리아를 배양물을 생산하기 위한 증식 배지 내로 도입하는 단계;
- [0050] o 상기 배양물을 배양하는 단계; 및
- [0051] o 상기 배양에 의해 생산된 비오틴을 회수하고, 선택적으로 회수된 비오틴을 정제하는 단계를 포함하는 방법을 제공한다.
- [0052] 제3 구현예에 따르면, 본 발명은 리포산을 생산하는 방법으로서:
- [0053] o 본 발명에 따른 리포산 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드를 코딩하는 전이 유전자를 포함하는 유전자 변형 박테리아를 배양물을 생산하기 위한 증식 배지 내로 도입하는 단계;
- [0054] o 상기 배양물을 배양하는 단계; 및
- [0055] o 상기 배양에 의해 생산된 리포산을 회수하고, 선택적으로 회수된 리포산을 정제하는 단계를 포함하는 방법을 제공한다.
- [0056] 제4 구현예에 따르면, 본 발명은 티아민을 생산하는 방법으로서:
- [0057] o 본 발명에 따른 HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드를 코딩하는 전이 유전자, 및/또는 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드를 코딩하는 전이 유전자를 포함하는 유전자 변형 박테리아를 배양물을 생산하기 위한 증식 배지 내로 도입하는 단계;
- [0058] o 상기 배양물을 배양하는 단계; 및

- [0059] o 상기 배양에 의해 생산된 티아민을 회수하고, 선택적으로 회수된 티아민을 정제하는 단계를 포함하는 방법을 제공한다.
- [0060] 바람직하게 비오틴, 리포산 및 티아민 중 어느 하나를 생산하는 방법에서 사용된 증식 배지는 배지는 글루코스, 말토스, 갈락토스, 프럭토스, 수크로스, 아라비노스, 자일로스, 라피노스, 만노스, 및 락토스, 또는 이들의 임의의 조합으로부터 선택된 탄소원을 포함한다.
- [0061] 제4 구현예에 따르면, 본 발명은 비오틴 신타아제를 코딩하는 전이 유전자를 발현하는 박테리아 세포에서 비오틴 생산을 증가시키기 위해 돌연변이 *iscR* 폴리펩티드를 코딩하는 유전적으로 변형된 유전자의 용도를 제공한다. 상기 돌연변이 *iscR* 폴리펩티드는 서열번호 2, 4, 6, 8, 10, 12 및 14와 적어도 80% 서열 상동성을 갖는 것이고; 상기 아미노산 서열은 L15X, Cys92X, Cys98X, Cys104X, 및 His107X로 이루어진 군으로부터 선택된 적어도 하나의 아미노산 치환을 갖는 것이고, 상기 X는 서열번호 2, 4, 6, 8, 10, 12 및 14에 상응하는 아미노산 잔기 이외의 임의의 아미노산이다.
- [0062] 제4 구현예에 따르면, 본 발명은 박테리아에서 비오틴, 리포산 또는 티아민 중 어느 하나의 생산을 증가시키기 위한 돌연변이 *iscR* 폴리펩티드를 코딩하는 유전적으로 변형된 유전자의 용도를 제공한다. 상기 박테리아는 이들로부터 선택된 폴리펩티드를 코딩하는 적어도 하나의 전이 유전자를 포함하고 발현한다:
- [0063] · 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드,
- [0064] · 리포산 신타아제 (EC 2.8.1.8) 활성을 갖는 폴리펩티드,
- [0065] · HMP-P 신타아제 (EC 4.1.99.17) 활성을 갖는 폴리펩티드, 및 티로신 리아제 (EC 4.1.99.19) 활성을 갖는 폴리펩티드.
- [0066] 상기 유전적으로 변형된 유전자는 돌연변이 *IscR* 폴리펩티드를 코딩하는 내인성 *iscR* 유전자이고, 상기 돌연변이 *IscR* 폴리펩티드의 아미노산 서열은 서열번호 2, 4, 6, 8, 10, 12 및 14와 적어도 80% 아미노산 서열 상동성을 갖는 것이고,
- [0067] 상기 아미노산 서열은: L15X, C92X, C98X, C104X, 및 H107X로 이루어진 군으로부터 선택된 적어도 하나의 아미노산 치환을 갖는 것이고; 상기 X는 서열번호 2, 4, 6, 8, 10, 12 및 14에 상응하는 아미노산 잔기 이외의 임의의 아미노산이다.
- [0068] 제5 구현예에 따르면, 본 발명은 비오틴, 리포산 또는 티아민의 증가된 생산을 위한 본 발명에 따른 유전자 변형 박테리아의 용도를 제공한다.
- [0069] 제6 구현예에 따르면, 본 발명은 비오틴, 리포산 또는 티아민 중 어느 하나의 증가된 생산을 위한 본 발명에 따른 유전자 변형 박테리아를 제공한다. 상기 박테리아는 전자 공여자 NADPH로부터 SAM-라디칼 이온-황 클러스터 효소의 $[4Fe-4S]^{2+}$ 클러스터로의 증가된 전자 이동을 매개할 수 있는 폴리펩티드; 예를 들어, 플라보독신/페레독신 환원 효소 및 플라보독신 환원 시스템 또는 피루브산-플라보독신/페레독신 산화 환원 효소 시스템을 코딩하는 하나 이상의 유전자를 더 포함한다.

도면의 간단한 설명

- [0070] **도 1 A)** 박테리아의 비오틴 경로 및 비오틴의 합성을 일으키는 각각의 효소적 단계의 중간체를 나타내는 그림 (cartoon). SAM: S-아데노실-L-메티오닌, SAH: S-아데노실-L-호모시스테인, CoA: 코엔자임 A, ACP: 아실기 운반 단백질, KAPA: 7-케토-8-아미노펠라르곤산, AMTOD: S-아데노실-2-옥소-4-티오메틸부틸레이트, DAPA: 7,8-디아미노펠라르곤산, DTB: 데스티오비오틴, 5'DOA: 5'-데옥시아데노신. B) *isc*-오페론 구조 및 *IscR*의 조절 메커니즘뿐만 아니라 Fe-S-클러스터 형성에서의 역할을 나타내는 그림. 상기 *isc* 오페론은 다음의 유전자들의 발현을 조절하는 *IscR*을 코딩하는 *iscR* 유전자를 포함한다: *iscS* (시스테인 디설파라제 (cysteine desulphurase)), *iscU* (스캐폴드), *iscA* (A-타입 단백질), *HscB* (DnaJ-유사 코-샤페론), *HscA* (DnaK-유사 샤페론), 및 *fdx* (페레독신). *IscR*은 또한 *hyaA*, *ydiU*, *erpA*, 및 *sufA* 유전자를 포함하는 > 40 유전자를 조절한다.
- 도 2** 박테리아의 리포산 경로 및 리포산화된 리포일 도메인 (리포산 합성)을 일으키는 각각의 효소적 단계의 중간체를 나타내는 그림. 상기 경로에서의 주요한 효소는 LipA (리포산 신타아제) 및 LipB (옥타노일 단백질 ACP 운반 단백질: 단백질 트랜스퍼라제)를 포함하고, 기질은 SAM: S-아데노실-L-메티오닌. LipA는 리포에이트-단백질 리가아제 A; EC: 6.3.1.20이다.

도 3 박테리아의 티아민 경로 및 티아민 (THI); 티아민 모노포스페이트 (TMP) 및 티아민 디포스페이트 (TPP)의 합성을 일으키는 각각의 효소적 단계의 중간체를 나타내는 그림. 중간체의 약어: 5-아미노이미다졸 리보뉴클레오타이드 (AIR: aminoimidazole ribonucleotide), 4-아미노-2-메틸-5-(포스포옥시메틸)피리미딘 (HMP-P: 4-amino-2-methyl-5-(phosphoxymethyl)pyrimidine), 4-아미노-2-메틸-5-(디포스포메틸)피리미딘 (HMP-PP: 4-amino-2-methyl-5-(diphosphomethyl) pyrimidine), 1-데옥시-D-자일룰로스 5-포스페이트 (DXP: 1-deoxy-D-xylulose 5-phosphate), 디하이드로글리신 (DHG: dehydroglycine), 4-메틸-5-(2-포스포옥시에틸)티아졸 (THZ-P: 4-methyl-5-(2-phosphoxylethyl)thiazole), 아데노신 트리포스페이트 (ATP: adenosine triphosphate), 아데노신 모노포스페이트 (AMP: adenosine monophosphate), S-아데노실-L-메티오닌 (SAM: S-adenosyl-L-methionine), 환원된 니코틴아마이드 아데닌 디뉴클레오타이드 포스페이트 (NADPH: reduced nicotinamide adenine dinucleotide phosphate), 니코틴아마이드 아데닌 디뉴클레오타이드 포스페이트 (NADP+: nicotinamide adenine dinucleotide phosphate), 환원된 페레독신 (Fdx red: reduced ferredoxin), 산화된 페레독신 (Fdx ox: oxidized ferredoxin).

도 4 IPTG 유도성 *bioB* 발현 플라스미드를 포함하는 $\Delta bioB$ 균주 (대장균 BW25113) (오른쪽 패널); 및 IPTG 유도성 프레임시프트된 *bioB* (조기 종결 코돈) 발현 플라스미드를 포함하는 표준 균주 (대장균 BW25113) (왼쪽 패널)의 시간 경과에 따라 측정된 세포 밀도 (OD_{600} 에서 측정됨)의 그래픽 표현 (Graphical presentation). 0.1 g/L DTB, 50 μ g/mL 카나마이신 및 0 (점), 0.01 (삼각형), 또는 0.1 (사각형) mM IPTG를 갖는 200 μ l mMOPS에서 성장된 4개의 생물학적 복제물에 대하여 OD_{620} 은 멀티스칸을 사용하여 측정되고, OD_{600} 으로 전환되었다. 각 지수 성장률 값은 인접한 상자에 표시된다.

도 5 275 rpm 진탕을 하며 37°C에서 20시간 동안의 인큐베이션 후 (하기 기재된 비오틴 정량화하는 방법과 같음), 0 또는 0.244 μ g 비오틴/mL까지 증가하는 농도의 비오틴이 보충된 40 μ g/mL 제오신을 갖는 150 μ l mMOPS 상에서 성장된 플라스미드 pBS451을 포함하는 대장균 BS1011의 배양물의 최종 세포 밀도 (OD_{600} 에서 측정됨)를 나타낸 산포도의 그래픽 표현. 수직의 회색 점선은 비오틴 바이오어세이를 위한 0.024 내지 0.24 μ g 비오틴/L의 최적의 농도 범위를 식별한다.

도 6 IPTG-유도성 *bioB* 발현 플라스미드 (pBS412)를 각각 포함하는 4개의 생물학적 복제물에서 아미노산 돌연변이를 갖는 돌연변이 *IscR*을 발현하는 3개의 상이한 *iscR* 돌연변이 균주 (BS1377, L15F), (BS1375, C92Y) 및 (BS1353, H107Y) 및 대장균 BW25113 *bioB* 균주 (BS1011, Ref) (균주에 대하여 표 1 참조)의 비오틴 생산을 나타내는 바 다이어그램. 균주들은 24시간 동안 37°C에서 275 rpm 진탕과 함께 0.1 g/l DTB 및 50 μ g/mL 카나마이신을 갖는 400 μ l mMOPS에서 성장되었다. 바는 평균 비오틴 생산 값 (높이) 및 IPTG 유도 수준 (회색 음영)을 예시하고, 검은색 점은 개별적인 복제물 배양물로부터의 비오틴 생산을 나타내고, 수평의 점선은 참조 야생형 균주로부터의 최대 비오틴 생산을 표시한다. IPTG가 없이 배양된 경우 검출 가능한 수준의 비오틴을 생산한 균주는 없었다.

도 7 돌연변이 *IscR*을 발현하는 *iscR* 돌연변이 균주 (BS1353, H107Y), 및 대장균 BW25113 $\Delta bioB$ 균주 (BS1011, Ref)의 세포 밀도 및 비오틴 생산의 그래픽 표현이고, 상기 각 균주는 IPTG-유도성 *bioB* 유전자 발현 플라스미드 (pBS412)를 포함한다. 상기 데이터는 25시간 동안 모니터링된 각각의 *iscR* H107Y 돌연변이 균주 (굵은 어두운 선) 및 표준 균주, 대장균 BW25113 $\Delta bioB$ 균주 (얇은 회색 점선)의 3개의 생물학적 복제물의 측정된 OD_{600} 평균, 및 25시간 동안 모니터링된 각각의 *iscR* H107Y 돌연변이 균주 (굵은 어두운 점) 및 표준 균주, 대장균 BW25113 $\Delta bioB$ 균주 (얇은 회색 점)에 의한 비오틴 생산을 나타낸다. 상기 균주들은 250 mL 배플 진탕 플라스크 (baffled shake flask)에서 0.1 g/l DTB, 0.01 (A) 또는 0.5 mM IPTG (B) 및 50 μ g/mL 카나마이신을 갖는 50 mL mMOPS에서 37°C에서 275 rpm 진탕과 함께 성장되었다. 성장물은 검정색 박스에 나타낸다.

도 8 동정된 돌연변이 균주의 *iscR* 유전자에서 뉴클레오타이드 및 아미노산 서열 돌연변이의 위치를 나타내기 위해 주석이 달린 *IscR* 코딩 서열을 나타낸 그림

도 9 막대(sticks, WT 아미노산은 회색, 돌연변이 아미노산은 검정색으로 표시됨)로 표시된 L15F 및 H107Y *iscR* 돌연변이를 갖는 *hya* DNA 결합 부위 (검정색)에 결합된 *IscR* 이합체 (회색)의 결정 구조 (PDB entry 4HF1)를 나타낸 그림; 및 돌연변이된 잔기를 강조한 확대 이미지.

도 10 IPTG-유도성 *bioB* 발현 플라스미드 및 *isc*-오페론 (*iscSUA-hscBA-fdx*, *iscR* 유전자가 결핍된 천연 대장균 *isc* 오페론 구조에 상응함) 또는 강한 리보솜 결합 부위 (RBS: strong ribosomal binding site)에 작동 가능하게 연결된 대장균 *suf*-오페론 (*sufABCDSE*)을 포함하는 플라스미드 및 미디엄 카피 넘버 플라스미드(p15A ori)

또는 대조군 플라스미드로부터의 T5 LacO 억제 프로모터를 포함하는 *대장균* 균주의 비오틴 생산을 나타내는 바 다이어그램. 상기 대조군 플라스미드는 *suf-* 또는 *isc-*오페론 대신에 IPTG-유도성 GFP를 코딩하는 유전자를 포함하였다. 각 균주의 생물학적 삼중물 (triplicates)은 100 µg/mL 암피실린 및 50 µg/mL 스펙티노마이신을 갖는 mMOPS에서 0.1 g/l (DTB)를 기질로서 제공하여 낮은 (0.01 mM IPTG) 및 높은 (0.1 mM IPTG) 유도 하에 배양되었다. 비오틴 생산이 성장 기반의 바이오어쎄이를 이용하여 평가된 후, 상기 균주들은 딥 웰 플레이트에서 24 시간 동안 37°C에서 275 rpm으로 성장되었다. 바는 평균 비오틴 성장 값 (높이) 및 IPTG 유도 수준 (회색 음영)을 예시하고, 검은색 점은 개별적인 복제물로부터의 비오틴 생산을 나타내고, X표(crosses)는 OD₆₀₀으로 측정된 각 균주의 종점(end-point, end) 세포 밀도를 나타낸다. 0.01 mM IPTG로 유도된 경우, 검출 가능한 수준의 비오틴을 생산한 균주는 없었다.

도 11 삼중물에서 수행된 4개의 상이한 샘플에서의 BioB 단백질 발현 수준 및 비오틴 생산의 상관관계의 그래픽 표현. 상기 균주는 pBS430 (*bioB* 프레임시프트 IPTG 유도성 플라스미드)을 갖는 BS1013 (*대장균* BW25113, 백그라운드 균주), pBS412 (*bioB* IPTG 유도성 플라스미드)를 갖는 BS1011 ($\Delta bioB$ 를 갖는 BS1013), pBS412를 갖는 BS1353 (*iscR* H107Y 돌연변이를 갖는 BS1011)이었다. 균주들은 0.1 g/l DTB 및 그래프에 나타낸 바와 같은 IPTG를 갖는 mMOPS에서 성장되었다.

도 12 IPTG-유도성 *bioB* 발현 플라스미드 및 다음 중 *iscR*의 게놈 변이체를 포함하는 *대장균* $\Delta bioB$ 균주의 비오틴 생산을 나타내는 바 다이어그램: 야생형 (*iscR* WT), 22번 위치에서 종결 코돈으로 돌연변이된 E22* 글루탐산을 코딩하는 녹-아웃 돌연변이 (*iscR* KO), 92번 위치에서 시스테인으로부터 티로신으로의 치환을 코딩하는 돌연변이 (*iscR* C92Y). 바는 IPTG 유도의 주어진 수준 (회색의 음영)에서의 비오틴 생산 평균 값 (높이)을 예시하고, 점은 개별적인 복제물로부터의 비오틴 생산을 나타낸다. 각 균주의 생물학적 삼중물은 100 µg/mL 암피실린을 갖는 mMOPS에서 0.1 g/l DTB를 기질로서 제공하여 없는 (0 mM IPTG), 낮은 (0.01 mM IPTG) 및 낮은 (0.1 mM IPTG) 유도 하에 배양되었다. 비오틴 생산이 성장 기반의 바이오어쎄이를 이용하여 평가된 후, 각 균주는 딥 웰 플레이트에서 24시간 동안 37°C에서 275 rpm으로 성장되었다.

도 13 비오틴-오페론 플라스미드 및 다음 중 *iscR*의 게놈 변이체를 포함하는 *대장균* $\Delta bioA$ $\Delta bioB$ CD 균주에 의한 비오틴 생산을 나타내는 바 다이어그램: 야생형 (*iscR* WT), 돌연변이 *iscR* H107Y (107번 위치에서 히스티딘으로부터 티로신으로의 치환을 코딩함), 92번 위치에서 시스테인으로부터 티로신으로의 치환을 코딩하는 돌연변이 (*iscR* C92Y). 각 균주의 생물학적 사중물 (quadruplicate)은 0.1 데스티오비오틴 (DTB)을 기질로서 제공하거나 제공하지 않고 10 µg/mL 테트라사이클린을 갖는 mMOPS에서 배양되었다. 비오틴 생산이 성장 기반의 바이오어쎄이를 이용하여 평가된 후, 상기 균주들은 딥 웰 플레이트에서 24시간 동안 37°C에서 275 rpm으로 성장되었다. 바는 비오틴 생산 평균 값 (높이) 및 DTB의 공급 여부 (회색의 음영)를 예시하고, 검은색 점은 개별적인 복제물로부터의 비오틴 생산을 나타낸다.

도 14 IPTG 유도성 *lipA* (pBS993, 표 4 참조)를 각각 포함하는 3개의 생물학적 복제물에서 아미노산 돌연변이를 갖는 돌연변이 *IscR*을 발현하는 2개의 상이한 *iscR* 돌연변이 균주 (BS1375 C92Y) 및 (BS1353, H107Y) 및 대조군 균주 (BS1011, *IscR* WT) (균주에 대하여 표 1 참조)의 리포산 생산 및 생산 (X표) 24시간 후의 종점(최종) OD₆₀₀을 나타내는 바 다이어그램. 균주들은 100 g/mL 암피실린, 0.1 mM 비오틴, 0.6 g/l 옥탄산 및 0.01 mM IPTG를 갖는 400 µl mMOPS에서 24시간 동안 37°C에서 275 rpm 진탕과 함께 성장되었다. 바는 리포산 생산 평균 값 (높이)을 예시하고, 검은색 점은 개별적인 복제물 배양물로부터의 리포산 생산을 나타낸다. 종점 OD₆₀₀이 동일하게 유지되더라도, 평균 리포산 생산은 1.79-배 증가하는 것을 볼 수 있다.

도 15 표준 균주 (*대장균* BW25113), $\Delta lipA$ (WT, 삼각형); 및 돌연변이 *iscR* (C92Y)를 갖는 $\Delta lipA$ 균주 (C92Y, 사각형)로서, 상기 균주 모두 IPTG 유도성 *lipA* 발현 플라스미드 (pBS1037)를 포함하는 균주의 시간 경과에 따라 측정된, 세포 밀도 (OD₆₂₀에서 측정됨)의 그래픽 표현. OD₆₂₀은 0.6 g/L 옥탄산, 100 g/mL 암피실린 및 0 내지 0.04 mM IPTG (증가에 따라 회색 음영의 어둠이 증가함)을 갖는 200 µl mMOPS에서 성장된 6개의 생물학적 균주 복제물에 대하여 멀티스칸을 이용하여 측정되었다. 각 생존율 (GR: growth rate)은 오른쪽에 나타난다.

도 16 전체 티아민 경로 유전자, *thiCEFSGHMD*를 발현하는 플라스미드 (pBS140)를 각각 포함하는 4개의 생물학적 복제물에서 아미노산 돌연변이를 갖는 돌연변이 *IscR*을 발현하는 2개의 상이한 *iscR* 돌연변이 균주 (BS2019, C92Y) 및 (BS2020, H107Y) 및 *대장균* BW25113 $\Delta thiP$, *thiL** 균주 (BS750, Ref) (균주에 대하여 표 5 참조)의 티아민 생산을 나타내는 바 다이어그램. 균주들은 50 µg/mL 카나마이신을 갖는 400 µl mMOPS에서 24시간 동안

37℃에서 275 rpm 진탕과 함께 성장되었다. 바는 종점 OD₆₀₀에 대하여 보정된 티오크롬 어쎬이 (티아민, TMP 및 TPP를 함유함)에 의해 측정된 바와 같은 상층액에서의 티아민 생산 평균 값 (높이)을 예시하고, 검은색 점은 개별적인 복제물 배양물로부터의 티아민 생산을 나타낸다. OD 정규화(normalized) 역가는 돌연변이 균주 (BS2019 및 BS2020)에서 표준 균주 (BS750)와 비교하여 1.43배 향상된 것을 확인할 수 있다.

도 17 IPTG-유도성 BioB 과발현 플라스미드 pBS679 단독으로 (BS1937) 또는 FldA-Fpr의 항시적 과발현을 갖는 pBS1112 (BS2185) 또는 GFP의 항시적 과발현을 갖는 pBS1054 (BS2707)를 더한 대장균 $\Delta bioABFCD$ *iscR* H107Y (107번 위치에서 히스티딘으로부터 티로신으로의 치환을 코딩함) 균주에 의한 비오틴 생산을 나타내는 바 다이어그램. 각 균주는 100 μ g/ml 암피실린, 기질로서 0.1 g/L 테스티오비오틴 (DTB), 및 0, 0.01, 0.025, 0.05, 0.075 또는 0.1 mM IPTG를 갖는 mMOPS에서 배양되었다. BS2185 및 BS2707을 위한 배지는 50 μ g/ml 카나마이신을 포함하는 것을 제외하고는 동일하였다. 비오틴 생산이 성장 기반의 바이오어쎬이를 이용하여 평가된 후, 상기 균주들은 덤 웰 플레이트에서 24시간 동안 37℃에서 275 rpm 진탕과 함께 성장되었다. 바는 각각의 균주에 의한 비오틴 생산 값 (높이)을 예시한다: BS1937 (검은색 바); BS2185 (회색 바); 및 BS2707 (체크무늬 회색).

도 18 IPTG-유도성 BioB 과발현 플라스미드 pBS679를 포함하는 대장균 $\Delta bioABFCD$ *iscR* H107Y (107번 위치에서 히스티딘으로부터 티로신으로의 치환을 코딩함) 균주에 의한 비오틴 생산을 나타내는 바 다이어그램. BS2185는 FldA-Fpr의 항시적 과발현을 갖는 pBS1112를 더 포함한다. BS1937은 100 μ g/ml 암피실린, 기질로서 0.1 g/L 테스티오비오틴 (DTB), 및 0.025 mM IPTG 유도를 갖는 mMOPS에서 배양되었다. BS2185를 위한 배지는 동일하나, 50 μ g/ml 카나마이신을 더 포함하였다. 비오틴 생산이 성장 기반의 바이오어쎬이를 이용하여 평가된 후, 상기 균주들은 덤 웰 플레이트에서 24시간 동안 37℃에서 275 rpm과 함께 성장되었다. 어두운 회색 바는 비오틴 생산 평균 값 (높이) (BS1937 n=6 및 BS2185 n=8)을 예시하고, 밝은 회색 바는 종점 OD₆₀₀을 예시한다. 검은색 점은 비오틴 생산 및 개별적인 복제물로부터의 종점 OD₆₀₀을 나타낸다.

발명을 실시하기 위한 구체적인 내용

정의:

아미노산 서열 상동성: 본원에서 사용된 용어 “서열 상동성”은 실질적으로 동일한 길이의 2개의 아미노산 서열 사이의 상동성 정도의 정량적인 치수를 나타낸다. 비교되는 2개의 서열은 갭(gaps)의 삽입 또는 대안적으로, 단백질 서열의 말단에서의 절단에 의해 가능한 최상의 핏(fit)을 제공하도록 정렬되어야 한다. 서열 상동성은 ((Nref- Ndif) 100)/(Nref)로 계산될 수 있고, 상기 Ndif는 정렬된 경우의 2개의 서열에서 비-동일한 잔기의 총 수이고, Nref는 서열 중 하나의 잔기의 수이다. 서열 상동성 계산은 바람직하게는 BLAST 프로그램, 예를 들어, BLASTP 프로그램 (Pearson W.R and D.J. Lipman (1988)) (www.ncbi.nlm.nih.gov/cgi-bin/BLAST)을 사용하여 자동화된다. 다중 서열 정렬은 <http://www2.ebi.ac.uk/clustalw/> 에서 이용 가능한 Thompson J., et al 1994에 의해 기술된 바와 같은 디폴트 파라미터를 갖는 서열 정렬 방법 ClustalW로 수행된다. 바람직하게, 폴리펩티드에서 하나 이상의 아미노산 잔기의 치환, 삽입, 첨가 또는 결실의 수는, 이의 비교 폴리펩티드와 비교하여, 즉, 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10 이하의 치환, 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10 이하의 삽입, 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10 이하의 첨가, 및 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10 이하의 결실로 한정된다. 바람직하게, 상기 치환은 보존적 아미노산 치환이다: 제1 군: 글리신, 알라닌, 발린, 류신, 이소류신; 제2 군: 세린, 시스테인, 셀레노시스테인, 트레오닌, 메티오닌; 제3 군: 프롤린; 제4 군: 페닐알라닌, 티로신, 트립토판; 제5 군: 아스파르테이트, 글루타메이트, 아스파라긴, 글루타민인 군의 구성원 내에서의 교환으로 제한됨.

아미노산 약어: 류신 (L), 시스테인 (C), 및 히스티딘 (H).

내인성 유전자: 숙주 박테리아와 기원이 동일한 박테리아 세포 계놈 내의 유전자 (즉 숙주 박테리아의 천연 유전자)이다. 내인성 유전자는 당업계에 알려진 도구를 사용하여 유전적으로 변형될 수 있으며, 이에 의해 유전적으로 변형된 내인성 유전자는 이것이 유래된 모체 내인성 유전자에 의해 코딩되는 폴리펩티드와 하나 이상의 위치에서 아미노산 서열이 상이한 돌연변이 폴리펩티드를 코딩한다.

계놈: 세포 또는 유기체에 존재하는 유전적 물질이다; 상기 세포 또는 유기체를 구축하고 유지하기 위해 필요한 모든 정보를 포함하는 계놈은 세포 또는 유기체 내에 존재하는 염색체(들) 및 플라스미드(들) 모두에서 유전 물질을 포함한다.

GFP: 녹색 형광 단백질.

- [0077] **gi 번호:** 유전자정보 검색번호(genInfo identifier)는 DDBJ/EMBL/GenBank로부터의 뉴클레오티드 서열, SWISS-PROT, PIR 및 다른 많은 것들로부터의 단백질 서열을 포함하여, Entrez로 처리된 모든 서열에 NCBI에 의해 할당되는 데이터베이스 근원과 관계 없이, 특정 서열을 식별하는 독특한 정수이다.
- [0078] **Isc 경로:** 철 황 클러스터 경로; *iscR* 유전자를 포함한 *isc* 오페론에 의해 코딩됨.
- [0079] **멀티스칸(Multiskan):** 필터-기반의 마이크로플레이트 광도계; 600 - 620nm을 포함한, 340 내지 850 nm 범위의 파장에서 96 또는 384-웰 플레이트 포맷으로부터 흡광도를 측정하기 위함. 플레이트를 최대 50℃의 선택된 온도에서 광도계에서 인큐베이트된다. 광도계는 Thermo Scientific에 의해 제공된다.
- [0080] **천연 유전자 (Native gene):** 숙주 박테리아와 동일한, 박테리아 세포 계놈 내에서 내인성 유전자.
- [0081] **비-천연 프로모터 (Non-native promoter):** 본 발명의 유전자 변형 박테리아와 관련하여, 상기 세포에서 유전자 또는 전이 유전자에 작동 가능하게-연결된 프로모터이며, 상기 프로모터는 자연에서 발견된 박테리아 세포에서 상기 유전자 또는 전이 유전자에 작동 가능하게-연결된 것으로 발견되지 않을 것임.
- [0082] **OD (Optical Density):** 광학 밀도
- [0083] **전이 유전자(Transgene):** 계놈 공학에 의해 박테리아의 계놈 내로 도입된 외인성 유전자이다. 본 발명의 내용에서, 상기 계놈은 염색체 및 에피솜 유전자 요소를 모두 포함한다.
- [0084] **발명의 상세한 설명**
- [0085] 비오틴, 리포산 및 티아민의 합성을 위한 생합성 경로의 공통적인 특징은 복잡한 라디칼-매개 분자 재배열을 촉매하기 위한 하나 이상의 SAM 또는 AdoMet 라디칼 효소에 대한 필요 조건이다. 비오틴 신타아제, 리포산 신타아제, HMP-P 신타아제, 및 티로신 리아제는 이들 경로에서 이들 필수 단계를 촉매하는 것으로 알려진 각각의 효소들이다. 비오틴 오페론의 과발현, 또는 심지어 BirA 억제자에 의한 피드백 조절에 영향을 받지 않는 돌연변이 비오틴 오페론의 사용에 의한, *대장균*에서 비오틴 생합성을 증가시키기 위한 초기 시도의 실패는 성장의 강력한 억제 때문이었다 (Ifuku, O. *et al.*, 1995). 관찰된 성장 억제에 대하여 증거-기반의 어떠한 설명도 없는 경우; 비오틴 신타아제 과-발현의 독성을 설명할 수 있는 세포 인자를 규명하기 위한 대안적인 접근법이 필요했다.
- [0086] 본 발명에 의해 제공되는, 이 문제에 대한 해결책은 박테리아 세포 공장 (예를 들어, *대장균*)의 세포들에서 비오틴 신타아제, 리포산 신타아제, HMP-P 신타아제 및 티로신 리아제의 발현을 증가시키기 위해 동일하게 적용 가능한 것으로 나타난다. 이 문제를 해결하기 위한 접근법은 불완전한 오류-정정(error-correcting) 폴리머라제에 의해 생성된 백그라운드 돌연변이의 축적으로 인해 진화된 계놈 다양성을 갖는 *대장균* 세포의 라이브러리를 생성하는 것이었다. 이러한 라이브러리의 세포를 IPTG-유도성 *bioB* 유전자 발현 카세트를 포함하는 플라스미드로 형질전환시켰다. 후보 돌연변이는 돌연변이 세포가 유래된 모체 *대장균* 균주에서 BioB 발현 독성을 유도하기에 충분한 농도에서 IPTG 존재 하에 성장할 수 있는 라이브러리 내의 세포들이었다.
- [0087] 선택된 BioB-발현 돌연변이 균주의 증식을 위한 유전적 기초는 전체 계놈 시퀀싱에 의해 확립되었다. 놀랍게도, 3개의 균주들은 천연 철 황 클러스터 조절 유전자 (*iscR*)에서 돌연변이를 갖는 것으로 밝혀졌다; 이것은 다면 발현성 전사 인자 (IscR) [서열번호 2]를 코딩한다. Fe-S 클러스터는 많은 단백질 및 필수적 효소들의 보조인자로서, 비오틴과 같은 S-함유 화합물의 합성을 위해 단독으로 필요할 뿐만 아니라, 산화환원- (redox-) 또는 철-관련 스트레스 조건에 대한 센서로서 이들에 다양한 생화학적 능력을 부여한다.
- [0088] IscR은 Fe-S 클러스터 단순-단백질(holo-protein), 또는 Fe-S 클러스터가 없는 아포단백질(apoprotein)로서 2가지 상태로 존재한다. IscR의 Fe-S 클러스터의 어셈블리는 *isc* 오페론에 의해 코딩되는 Isc 경로에 의해 촉매된다. *isc* 오페론은 먼저 조절자 (IscR)를 코딩하고, 그런 다음 시스테인 디설파라제(desulphurase) (IscS), 스캐폴드 (IscU), A-타입 단백질 (IscA), DnaJ-유사 코-샤페론(co-chaperone) (HscB), DnaK-유사 샤페론 (HscA) 및 페레독신 (Fdx)을 코딩한다. Isc 경로는 IscR 완전효소의 어셈블리를 위해 필수적일 뿐만 아니라, *대장균*에서 Fe-S 클러스터 생물발생(biogenesis)을 위한 주요한 경로이다 (도 1B).
- [0089] IscR의 2가지 형태 간의 비율은 [2Fe-2S] 클러스터의 세포 수준에 의해 결정되고, 이는 결국 철- 및 산소 수준을 포함한 여러 요인에 의해 영향을 받는다 (Py, B. & Barras, 2010). 철-풍부한 조건 하에, IscR은 주로 완전 효소로 존재하고, *isc* 오페론의 전사 억제자로 작용한다. 그러나, 철-적은 ([2Fe-2S] 클러스터의 낮은 수준) 조건 하에, IscR은 아포-단백질 상태로 전환되고, 이는 *isc* 오페론의 전사를 가능하게 한다. 이의 아포-단백질 상태에서, IscR은 산화 스트레스 하에 Fe-S 클러스터 생물발생을 촉매하는 *sufABCDSE* 오페론의 활성인자

(activator)로서 역할을 한다.

- [0090] 대장균에서 2개의 Fe-S-클러스터 어셈블리의 발현을 조절하는 것 외에도, IscR은 산화 스트레스 매커니즘 (예를 들어, *sodA*), 특이적 및 전면적(global) 조절자 (예를 들어, *yqjI* 및 *soxS*), 아미노산 생합성 (예를 들어, *argE*), 및 알려지지 않은 기능을 가진 다양한 유전자와 같은 작용의 다양한 메커니즘에 관여된 >40 개의 유전자들을 조절한다. IscR의 역할은 IscR 조절 환경이 호기성 및 혐기성 조건 사이에서 변화한다는 사실에 의해 더욱 복잡해진다 (Martin, and Imlay, 2012; Giel et al., 2006).
- [0091] IscR의 항상성 역할; 및 전면적 유전자 조절에서의 이의 역할의 관점에서; 이의 조절 특성의 임의의 변형의 결과는 예측할 수 없으며, 아마도 세포 대사에 대해 심오하다. 또한, 황 형성 (*suf*) 및 *isc* 경로 모두의 증가된 발현으로 인해, Fe-S 클러스터 생물발생이 증가되는 세포 조건은 축적된 Fe-S 클러스터가 펜톤 반응(fenton reactions)에 의해 퍼옥사이드 라디칼을 생성할 위험을 생성한다.
- [0092] 이러한 관점에서, 3개의 분리된 개별적인 돌연변이에 의해 입증된 바와 같이, IscR이 세포 BioB의 활성 및 독성에 대해 매우 중요한 것으로 발견되는 것은 매우 예상하지 못한 것이었다. 또한, IscR 단백질의 중요성은 예상치 못한 것이었고, 이는 Fe-S 클러스터를 합성하고 조립하는 증가된 능력을 제공하는 *isc* 오프론 또는 *suf* 오프론의 과-발현이 *bioB*를 과-발현하는 세포에서 비오틴 생산을 증가시키는 것으로 발견되지 않았기 때문이다 (실시예 1, 도 10 참조). 또한, *iscR* 유전자 녹-아웃에 의한 세포 *iscR* 조절의 제거는 세포에서 비오틴 생산을 증가시키는데 실패하였다 (실시예 1, 도 12 참조).
- [0093] 돌연변이 세포에서 BioB 발현의 독성을 제거하는 IscR 단백질에서의 3가지의 다른 돌연변이는 아미노산들의 단일 아미노산 치환, L15 [서열번호 16], C92 [서열번호 18] 및 H107 [서열번호 20]이었다 (도 8). 3개의 돌연변이 중 2개는 IscR의 잔기에 정확하게 상응하고, 이들 각각은 IscR 단순-단백질의 형성에 필수적인 것으로 알려져 있다. 대장균에서 관찰된 바와 같이, IscR은 특이한 Fe-S 클러스터 결합(ligation) 메커니즘을 가짐으로써, Fe-S 클러스터 결합에 필수적인 잔기는 H107뿐만 아니라, C92, C98, 및 C104이다. 이 비정형 결합은 다른 Fe-S 단백질에 비해 IscR의 완전효소 상태의 낮은 안정성을 부여할 수 있고, 이는 결국 낮은 Fe-S 조건 동안 아포-단백질 상태로의 전환을 설명한다 (Fleischhacker et al., 2012).
- [0094] 이론에 의해 구속되고자 하는 것은 아니나, 이는 본 발명의 돌연변이 *iscR* 유전자를 발현하는 반면에, 철-황 클러스터 함유 유전자들(비오틴 신타아제, 리포산 신타아제, HMP-P 신타아제 및 티로신 리아제)의 어셈블리를 심지어 이들의 과-발현 동안에 촉진하는 세포에서, Fe-S 클러스터 생물발생의 항상성 조절 및 세포 성장에 필요한 전면적인 유전자 조절이 유래없이 보존됨을 시사한다.
- [0095] 요약하면, 본 발명자들은 Fe-S 클러스터의 결합에 필요한 하나 이상의 아미노산 잔기의 부족을 특징으로 하는, 돌연변이 IscR 단백질을 코딩하는 돌연변이 *iscR* 유전자로서, 그 결과 발현된 돌연변이 IscR 단백질이 아포-단백질 형태로만 존재하도록 하는 *iscR* 유전자를 규명하였다. 박테리아에서 비오틴, 리포산 또는 티아민의 생산을 증가시키기 위해 노력으로, 효소를 함유하는 철-황 클러스터의 합성은 상당한 장애물을 구성하는 것으로 보여진다. 본 발명에 의해 제공된 바와 같이, 이 문제에 대한 해결책은 아포-단백질 상태로 존재하는 돌연변이 IscR 단백질을 코딩하는 유전자를 포함하는 세포 공장으로서 이들 효소의 과-발현에 의해 촉진된다. 본 발명의 다양한 구현에는 하기에 보다 자세히 기술된다.
- [0096] **I 비오틴의 생산을 위한 유전자 변형 박테리아 세포**
- [0097] 본 발명은 증가된 수준의 비오틴을 생산할 수 있는 유전자 변형 박테리아 세포를 제공한다. 상기 박테리아 세포는 야생형 IscR을 대체하여 돌연변이 IscR을 발현하고, 비오틴 신타아제 (EC 2.8.1.6을 갖는 비오틴 신타아제)를 코딩하는 전이 유전자를 포함하도록 유전적으로 변형된다. 선택적으로, 상기 유전자 변형 박테리아 세포는 비오틴 경로 (도 1A)의 추가적인 단계를 촉매하는 폴리펩티드를 코딩하는 하나 이상의 추가적인 전이 유전자를 더 포함할 수 있다. 비오틴 경로에서 단계들을 촉매하는 이들 폴리펩티드 수준의 증가는 박테리아 세포에서 비오틴 경로에서의 중간체 및 상기 경로의 최종 생성물 (비오틴) 모두의 합성을 향상시킨다.
- [0098] 유전자 변형 박테리아 세포에 의해 발현되는, 돌연변이 IscR 폴리펩티드는 폴리펩티드 백본 (아포-단백질)을 특징으로 하는 IscR 폴리펩티드 군의 야생형 구성원으로부터 유래된다. IscR 폴리펩티드 군의 야생형 구성원의 아미노산 서열은 서열번호 2, 4, 6, 8, 10, 12 및 14 중 어느 하나로부터 선택된 서열과 적어도 70, 75, 80, 85, 90, 95, 96, 98, 100% 아미노산 서열 상동성을 갖는다. 본 발명에 따른 돌연변이 IscR 폴리펩티드의 아미노산 서열은 적어도 하나의 아미노산 치환에 의해 유래된 상응하는 야생형 IscR 폴리펩티드의 아미노산 서열과 상이하다; 상기 치환은 L15X, C92X, C98X, C104X, 및 H107X로부터 선택된다; 상기 치환 아미노산인 X는 돌연변이가

유래된 야생형 IscR에서 상응하는 위치에서 발견되는 아미노산 이외의 임의의 아미노산이다.

[0099] 대안적인 구현예에서, 상기 돌연변이 IscR에서 아미노산 치환은 L15X로서, 상기 X는 L 이외의 임의의 아미노산이고, 보다 바람직하게 X는 페닐알라닌 (F), 티로신 (Y), 메티오닌 (M) 및 트립토판 (W)으로부터 선택되는 것인 L15X; C92X로서, 상기 X는 C 이외의 임의의 아미노산이고, 보다 바람직하게는 X는 티로신 (Y), 알라닌 (A), 메티오닌 (M), 페닐알라닌 (F) 및 트립토판 (W)으로부터 선택되는 것인 C92X; C98X로서, 상기 X는 C 이외의 임의의 아미노산이고, 보다 바람직하게는 X는 알라닌 (A), 발린 (V), 이소류신 (I), 류신 (L), 페닐알라닌 (F) 및 트립토판 (W)으로부터 선택되는 것인 C98X; Cys104X로서, 상기 X는 C 이외의 임의의 아미노산이고, 보다 바람직하게는 X는 알라닌 (A), 발린 (V), 이소류신 (I), 류신 (L), 페닐알라닌 (F) 및 트립토판 (W)으로부터 선택되는 것인 Cys104X; 및 His107X로서, 상기 X는 H 이외의 임의의 아미노산이고, 보다 바람직하게는 X는 알라닌 (A), 티로신 (Y), 발린 (V), 이소류신 (I), 및 류신 (L)으로부터 선택되는 것인 His107X로부터 선택되는 것이다. 예를 들어, 상기 돌연변이 IscR에서 아미노산 치환은 L15F, C92Y, C92A, C98A, Cys104A, H107Y, 및 H107A 중에서 선택되는 것일 수 있다.

[0100] 본 발명의 유전자 변형 박테리아 세포에 의해(야생형 IscR을 대체하여) 발현되는 돌연변이 IscR은 염색체 상 또는 자기-복제 플라스미드 상에, 박테리아 세포의 계놈에 위치한, 유전적으로 변형된 유전자에 의해 코딩된다. 상기 염색체에서 유전적으로 변형된 *iscR* 유전자는 천연 계놈에서 야생형 *iscR* 유전자와 동일한 위치에 있는 계놈에 위치할 수 있다. 천연 야생형 *iscR* 유전자는 결실되거나 유전적으로 변형된 *iscR* 유전자에 의해 직접적으로 치환되므로, 본 발명의 유전자 변형 박테리아 세포의 계놈은 천연 야생형 *iscR* 유전자가 결여된다. 유전적으로 변형된 *iscR* 유전자의 발현을 구동하는 프로모터는 상기 유전적으로 변형된 *iscR* 유전자가 유래되거나 대체된 야생형 *iscR* 유전자의 천연 프로모터일 수 있다. 대안적으로, 상기 프로모터는 이종의 항시성 또는 유도성 프로모터일 수 있다. 상기 프로모터가 이종의 항시성 프로모터인 경우, 적합한 프로모터는: *apFab* 패밀리를 포함하는 반면에, 적합한 유도성 프로모터는: *pBad* (아라비노스 유도성 [서열번호 233] 및 *LacI* [서열번호 234]를 포함한다. 적합한 종결자(terminator)는 [서열번호 235-237]를 포함하는 *apFAB* 종결자 패밀리의 구성원을 포함한다.

[0101] 본 발명에 따른 비오틴 신타아제 (EC 2.8.1.6) 활성을 갖는 폴리펩티드는 데스티오비오틴의 비오틴으로의 전환을 촉매하는 비오틴 신타아제 활성을 갖는 폴리펩티드이다. 비오틴 신타아제의 이 군의 구성원은 광범위한 속 (genera)에 속하는 박테리아에서 발견된 유전자에 의해 코딩된다. 비오틴 신타아제 활성을 갖는 폴리펩티드의 아미노산 서열은 이들 중 어느 하나로부터 선택된 서열과 적어도 70, 75, 80, 85, 90, 95, 96, 98, 100% 아미노산 서열 상동성을 갖는다: 서열번호 22 (기원: *에서리키아 콜라이(Escherichia coli)*); 서열번호 27 (기원: *칸디다투스 클로라시도박테리움 썬모필룸 B(Candidatus Chloracidobacterium thermophilum B)*); 서열번호 29 (기원: *스트렙토마이세스 리디커스(Streptomyces lydicus)*); 서열번호 31 (기원: *파라코커스 테니트리피칸스(Paracoccus denitrificans)*); 서열번호 33 (기원: *파라코커스 테니트리피칸스 PD1222*); 서열번호 35 (기원: *아그로박테리움 비티스(Agrobacterium vitis)*); 서열번호 37 (기원: *루에제리아 포메로이(Ruegeria pomeroyi)*); 서열번호 39 (기원: *아그로박테리움 파브룸(Agrobacterium fabrum)*); 서열번호 41 (기원: *시멕스 렉툴라리우스의 볼바키아속 내공생자(Wolbachia endosymbiont of Cimex lectularius)*); 서열번호 43 (기원: *스핀고모나스 파우치모빌리스(Sphingomonas paucimobilis)*); 서열번호 45 (기원: *애시디싸이오바실러스 페리보란스(Acidithiobacillus ferrivorans)*); 서열번호 47 (기원: *갈리오넬라 캡시페리포르만스(Gallionella capsiferiformans)*); 서열번호 49 (기원: *랄스토니아 유티로파(Ralstonia eutropha)*); 서열번호 51 (기원: *보르데텔라 파라퍼투스(Bordetella parapertussis)*); 서열번호 53 (기원: *푸실리모나스 종(Pusillimonas sp.)*); 서열번호 55 (기원: *케나르카이움 심비오숨 종(Cenarchaeum symbiosum sp.)*); 서열번호 57 (기원: *알리사이클로바실러스 아시도칼다리우스 종(Alicyclobacillus acidocaldarius sp.)*); 서열번호 59 (기원: *게오바실러스 썬모글루코시다시우스(Geobacillus thermoglucosidasius)*); 서열번호 61 (기원: *바실러스 서브틸리스(Bacillus subtilis)*); 서열번호 63 (기원: *리시나바실러스 스파이리쿠스(Lysinibacillus sphaericus)*); 서열번호 65 (기원: *메틸로코커스 캡슐라투스(Methylococcus capsulatus)*); 서열번호 67 (기원: *레클레르시아 아데카르복시라타(Leclercia adecarboxylata)*); 서열번호 69 (기원: *크로모할로박터 살렉시겐스(Chromohalobacter salexigens)*); 서열번호 71, 73, 75, 77, 79, 81, 83, 85, 87 (기원: *슈도모나스 종(Pseudomonas spp.)*).

[0102] 유전자 변형 박테리아 세포에서 추가적인 전이 유전자에 의해 코딩되고, 및 비오틴 경로의 중간체 및 생성물 모두의 합성을 증가시키는 역할을 활성을 갖는 폴리펩티드는 다음과 같다:

[0103] a) SAM (S-아데노실메티오닌)-의존적 메틸트랜스퍼라제 (**BioC**; EC 2.1.1.197) 활성을 갖는 폴리펩티드; 예를 들

어, 서열번호 89와 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;

[0104] b) 7-케토-8-아미노펠라르곤산 (KAPA) 신타아제 (**BioF**; EC 2.3.1.47) 활성을 갖는 폴리펩티드, 예를 들어, 서열번호 91과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;

[0105] c) 7,8-디아미노펠라르곤산 (DAPA) 신타아제 (**BioA**; EC: 2.6.1.62) 활성을 갖는 폴리펩티드, 예를 들어, 서열번호 93과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드; 또는 L-리신:8-아미노-7-옥소노나노에이트 아미노트랜스퍼라제 (**BioK**; EC: 2.6.1.105) 활성을 갖는 폴리펩티드, 예를 들어, 서열번호 97과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;

[0106] d) 데스티오비오틴 (desthiobiotin, DTB) 신타아제 (**BioD**; E.C 6.3.3.3) 활성을 갖는 폴리펩티드, 예를 들어, 서열번호 95와 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드; 및 선택적으로

[0107] f) 피멜로일-[아실-운반 단백질] 메틸 에스테르 에스테라제 (**BioH**; EC: 3.1.1.85) 활성을 갖는 폴리펩티드, 예를 들어, 서열번호 99와 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드; 또는

[0108] g) 6-카복시핵사노에이트-CoA 리가제 (**BioW**; EC 6.2.1.14) 활성을 갖는 폴리펩티드; 예를 들어, 서열번호 101과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;

[0109] 비오틴 경로의 추가적인 단계를 촉매하는 폴리펩티드를 코딩하는 하나 이상의 추가적인 전이 유전자와 함께, BioB를 코딩하는 전이 유전자는 박테리아 세포 염색체 내로 또는 자기-복제 플라스미드 상에 통합되는 유전자 변형 박테리아 세포의 게놈에 위치한다. BioB 및 비오틴 경로 효소들(BioABFCD 및 H 또는 W)에서 하나 이상의 효소를 코딩하는 전이 유전자는 하나 이상의 오페론 내의 게놈에 존재할 수 있다.

[0110] 하나 이상의 추가적인 전이 유전자와 함께 BioB를 코딩하는 전이 유전자의 발현을 유도하는 프로모터는 바람직하게는 이중의 항시성-프로모터 또는 유도성-프로모터일 수 있는 비-천연 프로모터이다. 상기 프로모터가 이중의 항시성 프로모터인 경우, 적합한 프로모터는 apFab 패밀리 [서열번호 230-232]를 포함하는 반면에, 적합한 유도성 프로모터는: pBad (아라비노스 유도성 [서열번호 233] 및 LacI [서열번호 234]를 포함한다. 적합한 종결자는 [서열번호 235-237]를 포함하는 apFAB 종결자 패밀리의 구성원을 포함한다. 선택된 프로모터 및 종결자는 BioB에 대한 코딩 서열; 및 BioC, BioD, BioA, BioF, 및 BioW 또는 BioH 폴리펩티드에 대한 하나 이상의 코딩 서열에 작동 가능하게 연결되는 것일 수 있고, 또한 선택된 Bio 폴리펩티드를 코딩하는 하나 이상의 오페론에 작동 가능하게 연결되는 것일 수 있다.

[0111] II 본 발명에 따른 유전자 변형 박테리아를 이용하여 비오틴을 생산하고 검출하는 방법

[0112] 비오틴은 비오틴의 생합성을 위해 적합한 탄소원을 포함하고; 배양에 의해 생산된 비오틴을 최종적으로 회수할 뿐만 아니라, 성장을 지지하기에 적합한 배양 배지 내로 세포를 도입함으로써, 본 발명의 유전적 변형된 박테리아 세포(예를 들어, 유전자 변형 대장균 세포)를 사용하여 비오틴은 생산되고 추출될 수 있다.

[0113] 비오틴 신타아제 (BioB)를 코딩하는 전이 유전자를 포함하는 본 발명의 유전자 변형 박테리아 세포는 공급된 탄소원이 데스티오비오틴(DTB)를 포함하는 경우, 비오틴을 증가된 수준으로 생산할 것이다. 본 발명의 유전자 변형 박테리아 세포는 BioA, BioF, BioC, BioD, 및 BioH 또는 BioW 각각을 코딩하는 전이 유전자를 추가적으로 포함하고, 이는 공급된 탄소원이 글루코스, 말토스, 갈락토스, 프럭토스, 수크로스, 아라비노스, 자일로스, 라피노스, 만노스, 및 락토스 중에서 선택되는 경우, 비오틴을 생산할 것이다 (실시예 1, 도 13).

[0114] 본 발명의 유전자 변형 박테리아 세포에 의해 생산된 세포 외 비오틴을 정량화하는 방법은 실시예 1.5에 기술된다. 상기 방법은 본 발명의 세포의 배양으로부터 유래된 세포 외 증식 배지로 보충된 비오틴-결핍 증식 배지 내에 플라스미드 pBS451을 포함하는 BS1011의 비오틴-결핍된(starved) 하룻밤 동안의 배양의 성장을 측정하는 것에 기초한, 바이오어쎄이 (bioassay)이다. 도 5에 나타난 바와 같이, 비오틴 기준(standards)의 알려진 농도 범위로 보충된 경우, 비오틴 바이오어쎄이 검정 곡선 (calibration curve)이 비오틴-결핍된 하룻밤 동안의 배양의 성장을 측정함으로써 작성된다.

[0115] III 리포산의 생산을 위한 유전자 변형 박테리아 세포

[0116] 본 발명은 리포산의 증가된 수준을 생산할 수 있는 유전자 변형 박테리아 세포를 제공한다. 본 발명에 따르면, 상기 박테리아 세포는 리포산 신타아제 (EC 2.8.1.8)를 코딩하는 전이 유전자를 포함할 뿐만 아니라, 야생형 IscR 대신에 돌연변이 IscR을 발현하도록 유전적으로 변형된다 (섹션 I 참조). LipA는 2개의 황 결합의 형성을 촉진함으로써 공유결합된 옥타노일-도메인을 리포일 도메인으로의 전환을 촉매한다. 선택적으로, 상기 유전자

변형 박테리아 세포는 리포산 합성 경로의 추가적인 단계를 촉매하는 폴리펩티드를 코딩하는 하나 이상의 추가적인 전이 유전자를 더 포함할 수 있고 (도 2), 보다 구체적으로는 예를 들어 *LipB* 유전자의 상기 코딩된 폴리펩티드 *LipB*; EC:2.3.1.181, 및 예를 들어 *aceF* 유전자의 상기 코딩된 폴리펩티드 E2; EC:2.3.1.12. 박테리아 세포에서 리포산 경로에서의 단계를 촉매하는 이들 폴리펩티드의 수준의 증가는 상기 경로의 중간체 및 최종 산물 모두의 합성을 증가시킨다. *LpIA*, 리포에이트-단백질 리가아제 A; EC:6.3.1.20를 코딩하는 추가적인 전이 유전자는 옥타노일 모이어티의 활성화된 리포일 도메인으로의 전달을 촉매함으로써, 옥탄산이 공급된 세포에서 리포산의 합성을 촉진하는 역할을 한다.

[0117] 리포산 신타아제는 광범위한 속에 속하는 광범위한 박테리아 및 진균에서 발견되는 유전자에 의해 코딩된다. 리포산 신타아제 활성을 갖는 폴리펩티드의 아미노산 서열은 다음 중 어느 하나로부터 선택된 서열과 적어도 70, 75, 80, 85, 90, 95, 96, 98, 100% 아미노산 서열 상동성을 갖는다: 서열번호 103 (기원: *에서리키아 콜라이*); 서열번호 105 (기원: *바실러스 서브틸리스*); 서열번호 107 (기원: *사카로미세스 세레비제*(*Saccharomyces cerevisiae*)); 서열번호 109 (기원: *슈도모나스 푸티다* (*Pseudomonas putida*); 서열번호 111 (기원: *박테로이데스 프라길리스* (*Bacteroides fragilis*)); 및 서열번호 113 (기원: *스트렙토마이세스 쉐리칼라* (*Streptomyces coelicolor*)).

[0118] 유전자 변형 박테리아 세포에서 추가적인 전이 유전자에 의해 코딩되고, 리포산 경로의 중간체 및 생성물 모두의 합성을 증가시키는 역할을 하는 폴리펩티드는 다음과 같다:

[0119] a) 옥타노일트랜스퍼라제 활성을 갖는 폴리펩티드 (ACP로부터 타겟 효소의 E2 서브유닛의 아포-리포일 도메인으로의 옥타닐 잔기의 이동을 위한; *LipB*; EC: 2.3.1.181, 예를 들어, 서열번호 115 (기원: *에서리키아 콜라이*) 또는 서열번호 117 (기원: *시겔라플렉스너리* (*Shigella flexneri*))과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;

[0120] b) 피루브산 탈수소효소 (E2; EC: 2.3.1.12)의 디하이드로리포일라이신-잔기 아세틸트랜스퍼라제 성분을 포함하는 폴리펩티드, 예를 들어 서열번호 119 (기원: *에서리키아 콜라이*), 또는 서열번호 121 (기원: *클렙시엘라 옥시토카* (*Klebsiella oxytoca*)) 또는 서열번호 239 (하이브리드 서열(hybrid sequence))과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드.

[0121] c) 리포에이트-단백질 리가아제 A (EC:6.3.1.20) 활성을 갖는 폴리펩티드, 예를 들어 서열번호 123 (기원: *에서리키아 콜라이*) 또는 서열번호 125 (기원: *클렙시엘라 옥시토카*)과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드.

[0122] 리포산 경로의 추가적인 단계를 촉매하는 폴리펩티드를 코딩하는 하나 이상의 추가적인 전이 유전자와 함께 리포산 신타아제를 코딩하는 전이 유전자는 박테리아 세포 염색체 내로 또는 자기-복제(self-replicating) 플라스미드 상에 통합되는 유전자 변형 박테리아 세포의 게놈에 위치한다. *LipA*를 코딩하는 전이 유전자 및 리포산 경로 효소를 코딩하는 하나 이상의 전이 유전자 (*IpB*, *IpIA*, 및 *AceF*)는 하나 이상의 오페론 내에 게놈에 존재할 수 있다.

[0123] *LipA*를 코딩하는 전이 유전자 및 하나 이상의 추가적인 전이 유전자의 발현을 유도하는 프로모터는 바람직하게는 이중의 항시성-프로모터 또는 유도성-프로모터일 수 있는 비-천연 프로모터이다. 상기 프로모터가 이중의 항시성 프로모터인 경우, 적합한 프로모터는 apFab 패밀리 [서열번호 230-232]를 포함하는 반면에, 적합한 유도성 프로모터는: pBad (아라비노스 유도성 [서열번호 233] 및 LacI [서열번호 234]를 포함한다. 적합한 종결자는 [서열번호 235-237]를 포함하는 apFAB 종결자 패밀리의 구성원을 포함한다. 선택된 프로모터 및 종결자는 개별적인 유전자 조절을 제공하기 위해 또는 오페론의 조절을 위해, 각각의 유전자에 작동 가능하게 연결되는 것일 수 있다.

[0124] IV 본 발명의 유전자 변형 박테리아를 이용하여 리포산을 생산하고 검출하는 방법

[0125] 실시예 2 및 도 14에 예시된 바와 같이, 리포산은 본 발명의 유전자 변형 박테리아 세포(예를 들어, 유전자 변형 *대장균* 세포)를 이용하여 적합한 배양 배지 내로 상기 세포를 도입하고; 세포에 의해 생산된 리포산을 최종적으로 회수함으로써 생산될 수 있다.

[0126] 리포산 신타아제 (*LipA*)를 코딩하는 전이 유전자를 포함하는 본 발명의 유전자 변형 박테리아 세포는 공급된 탄소원이 옥탄산 (OA)를 포함하는 경우 리포산을 생산할 것이다. 상기 세포는 적합한 탄소원, 예를 들어, 글루코스, 말토스, 갈락토스, 프럭토스, 수크로스, 아라비노스, 자일로스, 라피노스, 만노스, 및 락토스 중에서 선택

된 탄소원으로 공급될 경우, 리포산을 생산할 것이다.

- [0127] 본 발명의 유전자 변형 박테리아 세포에 의해 생산된 세포 외 리포산을 정량화하는 방법은 실시예 2에 기술된다. 상기 방법은 본 발명의 세포로부터 추출된 리포산으로 보충된 최소 배지 상에서 리포산-의존적 영양 요구성(auxotrophic) 대장균 균주의 성장을 측정하는 것에 기초한, 바이오어쎈이다.
- [0128] **V 티아민의 생산을 위한 유전자 변형 박테리아 세포**
- [0129] 본 발명은 티아민을 증가된 수준으로 생산할 수 있는 유전자 변형 박테리아 세포를 제공한다. 본 발명에 따르면, 상기 박테리아 세포는 야생형 IscR을 대체하여 돌연변이 IscR을 발현하고, *thiC*에 의해 코딩되는 HMP-P 신타아제라고도 불리는 포스포메틸피리미딘 신타아제 (EC 4.1.99.17); 또는 *thiH*에 의해 코딩되는 티로신 리아제 (2-이미노아세테이트 신타아제 (EC 4.1.99.19)라고도 불림)를 코딩하는 전이 유전자를 포함하도록 유전적으로 변형된다.
- [0130] 상기 유전자 변형 박테리아 세포는 티아민 합성 경로에서 추가적인 단계를 촉매하는 폴리펩티드를 코딩하는 하나 이상의 전이 유전자를 더 포함할 수 있다 (도 3). 박테리아 세포에서 티아민 경로에서의 단계를 촉매하는 이들 폴리펩티드의 수준의 증가는 상기 경로의 중간체 및 최종 산물 모두의 합성을 증가시킨다. 예를 들어, 상기 박테리아 세포는 다음을 코딩하는 하나 이상의 전이 유전자를 더 포함할 수 있다: ThiE 티아민 포스페이트 신타아제 (EC 2.5.1.3); [ThiS] 아데닐일트랜스퍼라제 (EC 2.7.7.73) (예를 들어, *thiF* 유전자에 의해 코딩됨); ThiG 티아졸 신타아제 (E.C.2.8.1.10); ThiS 황-운반 단백질; ThiD 포스포하이드록시메틸피리미딘 키나아제 (EC 2.7.4.7) 및 티아민 모노-포스페이트 포스파타제 (E.C. 3.1.3.-); ThiO 글리신 옥시다아제 (EC 1.4.3.19); 및 ThiM 하이드록시에틸티아졸 키나아제 (2.7.1.50).
- [0131] HMP-P 신타아제는 광범위한 속에 속하는 광범위한 박테리아 및 진균에서 발견되는 유전자에 의해 코딩된다. HMP-P 신타아제 활성을 갖는 폴리펩티드의 아미노산 서열은 다음 중 어느 하나로부터 선택된 서열과 적어도 70, 75, 80, 85, 90, 95, 96, 98, 100% 아미노산 서열 상동성을 갖는다: 서열번호 201 (기원: *에서리키아 콜라이*); 서열번호 203 (기원: *시네크코쿠스_이롱가투스(Synechococcus_elongatus)*); 서열번호 205 (기원: *코리네박테리움 글루타미쿰 (Corynebacterium glutamicum)*); 서열번호 207 (기원: *칸디다투스 바우마니아 시사델리니콜라 (Candidatus Baumannia cicadellinicola)*). 티로신 리아제는 2-이미노아세테이트 신타아제 (EC 4.1.99.19)라고도 불린다. HMP-P 신타아제 활성을 갖는 폴리펩티드의 아미노산 서열은 서열번호 217의 서열과 적어도 70, 75, 80, 85, 90, 95, 96, 98, 100% 아미노산 서열 상동성을 갖는다.
- [0132] 유전자 변형 박테리아 세포에서 하나 이상의 추가적인 전이 유전자에 의해 코딩되고, 티아민 경로의 중간체 및 생성물 모두의 합성을 증가시키는 역할을 하는 폴리펩티드는 다음과 같다:
- [0133] a) [ThiS] 아데닐일트랜스퍼라제 (EC 2.7.7.73) 활성을 갖는 폴리펩티드, 예를 들어 서열번호 211과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;
- [0134] b) 티아민 포스페이트 신타아제 (EC 2.5.1.3) 활성을 갖는 폴리펩티드, 예를 들어 서열번호 209와 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;
- [0135] c) 티아졸 신타아제 (E.C.2.8.1.10) 활성을 갖는 폴리펩티드, 예를 들어 서열번호 215와 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;
- [0136] d) 포스포하이드록시메틸피리미딘 키나아제 (EC 2.7.4.7) 활성을 갖는 폴리펩티드, 예를 들어 서열번호 225와 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;
- [0137] e) 글리신 옥시다아제 (EC 1.4.3.19) 활성을 갖는 폴리펩티드; 예를 들어 서열번호 219, 221, 및 223으로부터 선택된 서열과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;
- [0138] f) ThiS 황-운반 활성을 갖는 폴리펩티드 예를 들어 서열번호 213과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드;
- [0139] g) 티아민 모노-포스페이트 포스파타제 (E.C. 3.1.3.-) 활성을 갖는 폴리펩티드; 예를 들어 서열번호 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199 중 어느 하나로부터 선택된 서열과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드; 및
- [0140] h) ThiM 하이드록시에틸티아졸 키나아제 (2.7.1.50) 활성을 갖는 폴리펩티드, 예를 들어 서열번호 227과 80,

85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는 폴리펩티드.

- [0141] 바람직하게는, 상기 유전자 변형 박테리아 세포는 다음의 효소를 코딩하는 전이 유전자를 포함한다: *ThiC* (*thiC* 유전자에 의해 코딩됨); *ThiD* (*thiD* 유전자에 의해 코딩됨), *ThiE* (*thiE* 유전자에 의해 코딩됨), *ThiF* (*thiF* 유전자에 의해 코딩됨), 황-운반 단백질 (*thiS* 유전자에 의해 코딩됨), *ThiG* (*thiG* 유전자에 의해 코딩됨), TMP 포스파타제 (TMP 포스파타제 유전자에 의해 코딩됨); 및 *ThiH* (*thiH* 유전자에 의해 코딩됨) 또는 *ThiO* (*thiO* 유전자에 의해 코딩됨). 구현예에 따르면, 상기 세포는 효소 *ThiM* (*ThiM* 유전자에 의해 코딩됨)을 코딩하는 전이 유전자를 더 포함할 수 있다.
- [0142] 본 발명의 유전자 변형 박테리아에서 티아민 합성 수준은 티아민-포스페이트 키나아제를 코딩하는 내인성 *thiL* 유전자의 돌연변이에 의해 더 증가될 수 있다. 돌연변이 *thiL* 유전자는 서열번호 228의 뉴클레오티드 서열을 가지고, 모체 야생형 유전자와 비교하여 G133D 치환을 갖는 폴리펩티드 [서열번호 229]를 코딩하는 뉴클레오티드 133-135에서의 돌연변이 (GGT에서 GAC)를 갖는다.
- [0143] 티아민 경로에서 추가적인 단계를 촉매하는 폴리펩티드를 코딩하는 하나 이상의 추가적인 전이 유전자와 함께, *thiC*에 의해 코딩되는, HMP-P 신타아제 (EC 4.1.99.17); 또는 *thiH*에 의해 코딩되는 티로신 리아제 (EC 4.1.99.19)를 코딩하는 전이 유전자는 박테리아 세포 염색체 내로 또는 자기-복제 플라스미드 상에 통합되는 유전자 변형 박테리아 세포의 게놈에 위치한다. *thiC* 또는 *thiH* 전이 유전자 및 티아민 경로에서의 효소를 코딩하는 하나 이상의 전이 유전자는 하나 이상의 오펜론 내에 게놈에 위치할 수 있다.
- [0144] *thiC* 또는 *thiH* 전이 유전자 및 하나 이상의 추가적인 전이 유전자의 발현을 유도하는 프로모터는 바람직하게는 이중의 항시성-프로모터 또는 유도성-프로모터일 수 있는 비-천연 프로모터이다. 상기 프로모터가 이중의 항시성 프로모터인 경우, 적합한 프로모터는 apFab 패밀리 [서열번호 230-232]를 포함하는 반면에, 적합한 유도성 프로모터는: pBad (아라비노스 유도성 [서열번호 233] 및 LacI [서열번호 234]를 포함한다. 적합한 종결자는 [서열번호 235-237]를 포함하는 apFAB 종결자 패밀리의 구성원을 포함한다. 선택된 프로모터 및 종결자는 개별적인 유전자 조절을 제공하기 위해 또는 오펜론의 조절을 위해, 각각의 유전자에 작동 가능하게 연결되는 것일 수 있다.
- [0145] **VI 본 발명에 따른 유전자 변형 박테리아를 이용하여 티아민을 생산하고 검출하는 방법**
- [0146] 실시예 3 및 도 16에 예시된 바와 같이, 티아민, 티아민 모노포스페이트 (TMP) 및 티아민 디포스페이트 (TPP)는 본 발명의 유전자 변형 박테리아 세포 (예를 들어, 유전자 변형 *대장균* 세포)를 이용하여 적합한 배양 배지 내로 상기 세포를 도입하고; 최종적으로 티아민, 및 추가적으로는 상기 세포에 의해 생산된 TPP 및 TMP를 회수함으로써 생산될 수 있다.
- [0147] HMP-P 신타아제를 코딩하는 전이 유전자를 포함하는 본 발명의 유전자 변형 박테리아 세포는 공급된 탄소원이 글루코스, 말토스, 갈락토스, 프럭토스, 수크로스, 아라비노스, 자일로스, 라피노스, 만노스, 및 락토스 중에서 선택되는 경우, 티아민, TPP 및 TMP를 생산할 것이다.
- [0148] 본 발명의 유전자 변형 박테리아 세포에 의해 생산된 티아민을 정량화하는 방법은 실시예 3에 기술되고; 티아민 기준과 비교한 고압력 액체 크로마토그래피 (High Pressure Liquid Chromatography)의 사용을 포함할 수 있다.
- [0149] **VII 비오틴, 리포산 또는 티아민의 생산을 위한 유전자 변형 박테리아 세포를 설계하는 방법**
- [0150] 본 발명의 박테리아 세포에서 비오틴, 리포산 또는 티아민의 합성과 관련된 효소 활성을 갖는 하나 이상의 폴리펩티드를 코딩하는 하나 이상의 전이 유전자를 클로닝하고 도입하기에 적합한 통합 (Integration) 및 자기-복제 벡터는 통상의 기술자에게 상업적으로 이용 가능하고 알려져 있다 (예를 들어, Sambrook et al., Molecular Cloning : A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989 참조). 박테리아 세포는 이중 DNA의 세포로의 도입에 의해 유전적으로 조작된다. 본 발명의 박테리아 세포에서 비오틴, 리포산 또는 티아민 합성과 관련된 효소 활성을 갖는 하나 이상의 폴리펩티드를 코딩하는 유전자의 이중 발현은 실시예 1, 2 및 3에서 각각 입증된다.
- [0151] 본 발명의 박테리아 세포에서 비오틴, 리포산 또는 티아민 합성과 관련된 효소 활성을 갖는 하나 이상의 폴리펩티드를 코딩하는, 핵산 분자는 자기-복제 벡터에 의해 숙주 세포 내로 도입되거나, 당업계의 표준인 방법 및 기술을 사용하여 숙주 세포 게놈 내로 선택적으로 통합될 수 있다. 예를 들어, 핵산 분자는 화학적 변형 (chemical transformation) 및 전기천공법을 포함한 형질전환 (transformation), 형질도입 (transduction), 유전자총법 (particle bombardment), 등과 같은 표준 프로토콜에 의해 도입될 수 있다. 청구된 발명의 효소를 코

당하는 핵산 분자를 발현하는 것은 게놈 내로 핵산 분자를 통합함으로써 달성될 수 있다.

- [0152] 본 발명의 박테리아 세포에서 천연 내인성 *iscR* 유전자의 유전적 변형은 표준 재조합 방법을 적합한 모체 박테리아 세포에 적용함으로써, 내인성 *iscR* 유전자의 결실(녹아웃) 및 섹션 I에 기재된 바와 같은 돌연변이 IscR 폴리펩티드를 코딩하는 전이 유전자를 이용한 삽입/치환에 의해 수행될 수 있다 (Datsenko KA, et al. ; 2000).
- [0153] 비오틴, 리포산 또는 티아민의 생산을 위한, 본 발명에 따른 유전자 변형 박테리아 세포는 박테리아(bacterium)일 수 있고, 적합한 박테리아의 비-완전한(exhaustive) 리스트는 다음과 같이 제공된다: *에서리키아* (*Escherichia*), *브레비박테리움* (*Brevibacterium*), *버크홀데리아* (*Burkholderia*), *캠필로박터* (*Campylobacter*), *코리네박테리움* (*Corynebacterium*), *슈도모나스* (*Pseudomonas*), *셀라티아* (*Serratia*), *락토바실러스* (*Lactobacillus*), *락토코커스* (*Lactococcus*), *아세트박터* (*Acetobacter*), *아시네토박터* (*Acinetobacter*), *슈도모나스* (*Pseudomonas*) 등으로 이루어진 군으로부터 선택된 박테리아의 속에 속하는 종.
- [0154] 본 발명의 바람직한 박테리아 종은 *에서리키아 콜라이*, *슈도모나스 푸티다*, *셀라티아 마르센세스* (*Serratia marcescens*), 및 *코리네박테리움 글루타미쿰*이다.
- [0155] VIII 본 발명의 유전자 변형 박테리아 세포의 비오틴 생산 능력은 증가된 전자 전달에 의해 향상된다.
- [0156] 산화된 $[4Fe-4S]^{2+}$ 클러스터를 포함하는 SAM-라디칼 철-황 클러스터 효소, 예를 들어, BioB, ThiC 및 LipA는 $[4Fe-4S]^+$ 클러스터로 환원하기 위해 전자 이동이 필요하다. 환원된 $[4Fe-4S]^+$ 클러스터만이 촉매에 필요한 SAM-라디칼을 생성할 수 있다. 전자 공여자 NADPH로부터 $[4Fe-4S]^{2+}$ 로의 전자 이동은 플라보독신/페레독신 환원 효소 (Fpr) 및 플라보독신 (FldA) 환원 시스템에 의해 또는 피루브산-플라보독신/페레독신 산화환원 효소 시스템에 의해 매개될 수 있다.
- [0157] 추가적인 구현예에서, 비오틴, 리포산 또는 티아민을 생산할 수 있는 본 발명에 따른 유전자 변형 박테리아 세포는 다음의 군으로부터 선택된 하나 이상의 유전자를 더 포함한다: 플라보독신/페레독신-NADP 환원 효소 (EC:1.18.1.2 및 EC 1.19.1.1)를 코딩하는 유전자; 피루브산-플라보독신/페레독신 산화 환원 효소 (EC 1.2.7)를 코딩하는 유전자; 플라보독신을 코딩하는 유전자; 페레독신을 코딩하는 유전자; 플라보독신 및 페레독신-NADP 환원 효소를 코딩하는 유전자. 상기 하나 이상의 유전자에 작동 가능하게-연결된 프로모터는 상기 박테리아에서 상기 하나 이상의 유전자의 발현을 증가시킬 수 있다; 상기 하나 이상의 유전자는 천연 유전자 또는 전이 유전자일 수 있다. 바람직하게, 상기 작동 가능하게-연결된 프로모터는 본 발명의 유전자 변형 박테리아가 유래된 모체 박테리아 보다 더 높은 수준으로 상기 박테리아에서 상기 하나 이상의 유전자의 발현을 증가시킨다. 바람직하게, 본 발명의 유전자 변형 박테리아 세포는 플라보독신/페레독신-NADP 환원 효소 (EC:1.18.1.2 및 EC 1.19.1.1)를 코딩하는 유전자 및 플라보독신을 코딩하는 유전자; 또는 플라보독신 및 페레독신-NADP 환원 효소 모두에 대한 코딩 서열을 포함하는 단일 유전자를 포함한다. 추가적으로 상기 유전자 변형 박테리아 세포는 페레독신을 코딩하는 유전자를 더 포함할 수 있다.
- [0158] 본 발명의 유전자 변형 박테리아 세포에서 전자 이동 경로의 구성요소를 발현하는 유전자의 과발현은 이들의 SAM-라디칼 철-황 클러스터 효소의 세포 활성을 증가시킨다 (본 발명의 비오틴-생산 세포에 대한 실시예 4에 예시된 바와 같음).
- [0159] 바람직하게, 본 발명의 유전자 변형 박테리아 세포에서 천연 유전자 또는 전이 유전자에 의해 코딩되는 폴리펩티드는 플라보독신/페레독신 환원효소 (EC:1.18.1.2 및 EC 1.19.1.1) 활성을 가지고, 이는 다음 중 어느 하나로부터 선택된 서열과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는다: 서열번호 241 (기원: *대장균*으로부터 *fpr* 유전자); 서열번호 243 (기원: *바실러스 서브틸리스* 168로부터 *yumC* 유전자); 서열번호 245 (기원: *슈도모나스 푸티다* KT2440로부터 *fpr-I* 유전자); 서열번호 247 (기원: *스트렙토마이세스 베네주엘라* ATCC 10712 -로부터 SVEN_0113 유전자); 서열번호 249 (기원: *코리네박테리움 글루타미쿰* ATCC 13032로부터 *Cgl2384* 유전자), 및 서열번호 251 (기원: *스핑고박테리움* 종 (*Sphingobacterium* sp.) JB170으로부터 SJN15614.1 유전자).
- [0160] 바람직하게, 본 발명의 유전자 변형 박테리아 세포에서 천연 유전자 또는 전이 유전자에 의해 코딩되는 폴리펩티드는 피루브산-플라보독신/페레독신 산화 환원 효소 (EC 1.2.7) 활성을 가지고, 이는 다음 중 어느 하나로부터 선택된 서열과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는다: 서열번호 253 (기원: *대장균* K12 MG1655로부터 *YdbK* 유전자); 서열번호 255 (기원: *지오박터 설퍼레두스* (*Geobacter sulfurreducens*) AM-1으로부터 *por* 유전자); 서열번호 257 (기원: *스트렙토마이세스 프라텐시스* (*Streptomyces*

pratensis) ATCC 33331로부터 *Sfla_2592* 유전자; 서열번호 259 (기원: *프로피오니박테리움 프레우텐레이치트* (*Propionibacterium freudenreichii*) DSM 20271로부터 RM25_0186 유전자); 서열번호 261 (기원: *시네키크시스티스* 종 (*Synechocystis* sp.) PCC 6803으로부터 *nifJ* 유전자)

[0161] 바람직하게, 본 발명의 유전자 변형 박테리아 세포에서 천연 유전자 또는 전이 유전자에 의해 코딩되는 폴리펩티드는 플라보독신이고, 이는 다음 중 어느 하나로부터 선택된 서열과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는다: 서열번호 263 (기원: *대장균* K12 MG1655로부터 *fldA* 유전자); 서열번호 265 (기원: *대장균* K12 MG1655로부터 *fldB* 유전자); 서열번호 267 (기원: *바실러스 서브틸리스* 168로부터 *ykuN* 유전자); 서열번호 269 (기원: *시네키크시스티스* 종 PCC 6803로부터 *isiB* 유전자; 서열번호 271 (기원: *스트렙토마이세스 베네주엘라* ATCC 10712로부터 *wrbA* 유전자); 서열번호 273 (기원: *메타노코커스 아이올리쿠스* 난카이-3 (*Methanococcus aeolicus* Nankai-3)으로부터 PRK06242 유전자).

[0162] 바람직하게, 본 발명의 유전자 변형 박테리아 세포에서 천연 유전자 또는 전이 유전자에 의해 코딩되는 폴리펩티드는 페레독신이고, 이는 다음 중 어느 하나로부터 선택된 서열과 80, 85, 90, 95 또는 100% 서열 상동성을 갖는 아미노산 서열을 갖는다: 서열번호 275 (기원: *대장균*으로부터 *fdx* 유전자); 서열번호 277 (기원: *바실러스 서브틸리스* 168로부터 *fer* 유전자); 서열번호 279 (기원: *코리네박테리움 글루타미쿰* ATCC 13032로부터 *fdxB* 유전자); 서열번호 281 (기원: *시네키크시스티스* 종 PCC 6803로부터 *fdx* 유전자); 서열번호 283 (기원: *스트렙토마이세스 베네주엘라* ATCC 10712로부터 SVEN_7039 유전자); 서열번호 285 (기원: *메타노코커스 아이올리쿠스* 난카이-3 으로부터 *fdx* 유전자).

[0163] 유전자 발현을 증가시킬 수 있는 프로모터는 상기 박테리아에서 전자 이동 경로의 폴리펩티드를 코딩하는 천연 유전자 또는 전이 유전자에 작동 가능하게-연결된 경우 바람직하게 비-천연 프로모터이다. 상기 프로모터는 항시성 apFab309 프로모터 패밀리 [서열번호 230-232]의 구성원일 수 있다. 바람직하게 상기 비-천연 프로모터는 상기 천연 유전자 또는 전이 유전자에 작동 가능하게-연결된 경우 상기 유전자 변형 박테리아에서 이 박테리아가 유래된 모체 박테리아 보다 높은 수준까지 상기 코딩된 폴리펩티드의 발현을 증가시킨다. 상기 천연 유전자 또는 전이 유전자에 작동 가능하게-연결될 수 있는 적합한 종결자는 apFAB378 종결자 패밀리 [서열번호 235-237]를 포함한다.

[0164] 실시예

[0165] 실시예 1: 바이오틴 생산을 향상시킬 수 있는 유전자 변형 *대장균* 균주의 동정 및 특성화

[0166] 1 방법

[0167] 1. 1 : 실시예에서 사용된 하기 에세리키아 균주를 하기 열거한다.

표 1

균주

[0168]

명칭	설명
BS1013	하기 유전자형을 갖는 <i>E. coli</i> K-12 BW25113 모체 균주:rrnB3 Δ lacZ4787 hsdR514 Δ (araBAD)567 Δ (rhaBAD)568 rph-1
BS1011	<i>E. coli</i> K-12 BW25113로부터 유래된 Δ bioB ¹ (JW0758-1)
BS1353	<i>iscR</i> 유전자에 H107Y 돌연변이를 포함하는 BS1011 파생체(derivative)
BS1113	IPTG-유도성 BioB 발현을 제공하는 pBS412 플라스미드를 포함하는 BS1011 파생체
BS1375	<i>iscR</i> 유전자에 C92Y 돌연변이를 포함하는 BS1011 파생체
BS1377	<i>iscR</i> 유전자에 L15F 돌연변이를 포함하는 BS1011 파생체

[0169] ¹ Δ bioB 유전자의 결실 전 뉴클레오타이드 서열은 서열번호 21이었다.

[0170] 1.2 : 실시예에서 사용된 하기 플라스미드를 하기 나열한다.

표 2

[0171]

플라스미드

명칭	설명
pBS412	T5 lacO 억제 프로모터(repressed promoter)로부터의 BioB [서열번호 22] 과발현 플라스미드 (kanR, SC101)
pBS430	T5 lacO 억제 프로모터 [서열번호 25]로부터의 <i>bioB</i> ¹ (kanR, SC101)에서 초기 프레임 시프트 돌연변이를 갖는 pBS412
pBS451	항시적으로 발현되는 GFP [서열번호 287] (zeoR, p15A)
pBS281	미디엄 카피 넘버 플라스미드(medium copy number plasmid) (p15A ori)에서 복제된 IPTG 유도성 T5 프로모터로부터의 <i>E. coli</i> <i>isc</i> 오페론 (<i>iscSUA-hscBA-fdx</i>)
pBS282	미디엄 카피 넘버 플라스미드 (p15 ori)에서 복제된 IPTG 유도성 T5 프로모터로부터의 <i>E. coli</i> <i>suf</i> 오페론 (<i>sufABCDSE</i>)
pBS231	IPTG 유도성 T5 프로모터로부터 sfGFP 단백질을 코딩하는 유전자를 발현하는 미디엄 카피 넘버 플라스미드 (p15 ori)
pBS936	<i>bio</i> 오퍼레이터 위치에 "타입 9(type 9)" 돌연변이를 갖는 <i>E. coli</i> 유래의 천연 비오틴-오페론 (Ifuku et al., 1993)

[0172]

¹: *bioB* 프레임시프트(frameshift) 유전자의 뉴클레오타이드 서열은 서열번호 23을 갖는다.

[0173]

1.3 배지 및 첨가제:

[0174]

각각의 실시예에서 사용된 증식 배지(mMPOS)는 다음의 조성물을 갖는다: 1.32 mM K2HP04; 2 g/l D-글루코스; 0.0476 mg/l 판토텐산 칼슘; 0.0138 mg/l p-아미노벤조산; 0.0138 mg/l p- 하이드록시벤조산; 0.0154 mg/l 2,3-디하이드록시벤조산, 및 1x 변형된 MOPS 버퍼.

[0175]

10 x 변형된 MOPS는 0.4 M MOPS (3-(N-모르폴리노)프로판 설푼산 (3-(N-morpholino)propane sulfonic acid)); 0.04 M 트리신(Tricine); 0.1 mM FeSO47H20; 95 mM NH4Cl; 2.76 mM K2SO4; 5 μM CaCl22H2O; 5.25 mM MgCl2; 0.5 M NaCl; 및 미량영양소 스톡 용액(micronutrient stock solution)의 5000x 희석물을 포함한다.

[0176]

미량영양소 스톡 용액:

성분	화합식	FW	50 mL 당 그램
몰리브데넘산 암모늄 (ammonium molybdate)	(NH4)6Mo7O244H2O	1235.9	0.009
붕산	H3BO3	61.83	0.062
염화 코발트	CoCl2	237.9	0.018
황산동(cupric sulfate)	CuSO4	249.7	0.006
염화 망가니즈	MnCl2	197.9	0.040
황산아연	ZnSO4	287.5	0.007

[0177]

하기의 항생제 스톡(stocks)이 적용되었다: 1000x 희석물을 수득하기 위해 표시된 바와 같이 증식 배지에 첨가된; 암피실린 (amp, 100 mg/mL), 카나마이신 (kan, 50 mg/mL), 제오신(zeocin, zeo, 40 mg/mL).

[0179]

1.4 대장균 균주 라이브러리의 구축:

[0180]

진화된 게놈 다양성을 갖는 대장균 라이브러리는 세포를 kan으로 보충된 mMOPS 배지(mMOPS-kan)에서 하룻밤 동안 정치 배양하고, mMOPS-kan에서 생성된 배양물의 100x 희석물을 제조하고, 하룻밤 동안의 배양 및 희석의 연속적인 단계를 5회 반복함으로써 대장균 균주 BS1011의 세포로부터 유래된다. 이 과정은 불완전한 오류-정정 폴리머라제에 의해 생성된 백그라운드 돌연변이의 축적을 허용함으로써 유전적 다양성을 만든다. 배양 및 희석의 각 라운드 후에, 증가된 BioB 발현을 견디도록 적응된 세포의 진화를 검출하기 위해, 세포 배양물의 샘플은 IPTG를 갖는 mMOPS 플레이트 상에 플레이팅하였다(하기 참조). 이어서, 각 라이브러리의 세포는 BioB 과-발현 플라스미드, pBS412로 형질전환되었다.

[0181]

1.5 돌연변이 균주의 선별

[0182]

IPTG를 0, 0.0001, 0.001, 0.01, 0.1 및 1 mM의 농도로 포함하는 mMOPS (Ø=9츠)를 포함하는 일련의 1.5% 아가 플레이트 상에서 pBS412를 포함하는 BS1011의 mMOPS-kan에서 하룻밤 동안(o/n)의 배양으로부터 유래된 10⁴, 10⁵, 10⁶ 및 10⁷ 각각의 세포를 플레이팅함으로써, 선별 어췌이(selection assay)가 개발되었다. 이어서, 상기 플레이트는 37℃에서 최대 36시간 동안 인큐베이션되었고, 세포 성장은 간격을 두고 평가되었다. 이러한 조건들

하에, 0.1 mM IPTG으로 pBS412로부터의 BioB 발현의 유도는 최대 10^5 세포의 성장을 억제하는 것으로 발견된 반면에, 1 mM IPTG으로의 유도는 싱글 펩트리 디쉬 상에 플레이팅된 경우에 적어도 10^7 세포의 성장을 억제하였다. 10^5 세포의 세포 집단에 대한 1 mM IPTG로의 유도를 포함한 선택압(selection pressure)은 BioB 발현에 대한 더 높은 강건함을 갖는 균주를 식별하기 위해 최적의 건으로 발견되었고; 이에 따라 다음과 같이 실시되었다:

- [0183] 1) 섹션 1.4에 기재된 바와 같이, 각 라이브러리로부터 약 10^5 세포는 각 mMOPS-kan-1 mM IPTG 아가 플레이트 상에 플레이팅되었고, 최대 24시간 동안 37°C에서 인큐베이션되었다.
- [0184] 2) 단일 콜로니는 전-배양물(pre-cultures)을 생산하기 위해 mMOPS-kan 액체 배지에서 성장되었고, 이어서 전-배양물은 0.00, 0.01 또는 0.1 mM IPTG로 보충된 mMOPS-kan에서 수행된 비오틴 바이오어세이를 이용함으로써 이들의 비오틴 생산에 대해 평가되었다 (하기 섹션 1.6에 기재된 바와 같음). 각각의 전-배양물의 세포는 20% 글리세롤에서 글리세롤 스톱으로 보존되었다.
- [0185] 3) 1.5 mg 이상의 비오틴/l (세포 외 비오틴으로 검출됨)을 생산하는 콜로니는 mMOPS-kan 아가 플레이트 상에 재-배열되었고 (re-streaked), 최대 24시간 동안 37°C에서 인큐베이션되었고, 이어서 생물학적 복제물에서 비오틴 생산에 대하여 재-바이오어세이되었다 (re-bioassayed) (하기 섹션 1.6에 상세히 기재된 바와 같음).
- [0186] 4) 선별된 비오틴 과-발현 균주의 세포는 다음과 같이 평가되었다:
- [0187] a) 모체 균주 BS1011의 계능과 비교하여 선별된 균주의 세포의 계능에서의 유전적 돌연변이를 식별하기 위해 전체 계능 시퀀싱을 선별된 균주의 세포로부터 분리된 DNA 상에서 다음과 같이 수행하였다: 선별된 균주는 5-10 mL mMOPS-kan에서 성장되었고, 상기 세포는 이후 수확되었다; 계능 DNA는 Invitrogen Purelink 계능 DNA 추출 키트: (<https://www.thermofisher.com/order/catalog/product/K182001>)를 이용하여, 수확된 세포로부터 분리되었다; 상기 추출된 DNA는 전체 계능 시퀀싱에 적용되었다.
- [0188] b) pBS412 플라스미드의 선택된 균주의 세포를 양생(curing)하는 단계; 이어서 상기 양생된 균주의 세포를 pBS412 플라스미드로 재-형질전환시키는 단계; 및 최종적으로 상기 형질전환된 균주의 세포의 배양물의 비오틴 생산에 대한 바이오어세이를 다시하는 단계. pBS412 플라스미드의 세포를 양생하는 단계는 항생제를 포함하지 않고 1 mM IPTG를 갖는 풍부한(rich) 루리아 브로스(LB: Luria Broth) 배지에서 세포를 37°C에서 하룻밤 동안 증식시키고; 생성된 배양물의 세포를 LB 아가 플레이트 상에 배열(streaking out)시키고 37°C에서 하룻밤 동안 인큐베이션 함으로써 수행되었다. 아가 플레이트로부터 단일 콜로니는 50 μ l LB 배지에서 희석되었고, 5 μ l는 37°C에서 하룻밤 동안 인큐베이션된 LB 및 LB-amp 아가 플레이트 상에 점을 표시하기 위해 사용되었다. LB-amp 플레이트 상이 아닌, LB 플레이트 상에서 성장된 이들 단일 콜로니는 재-형질전환을 위해 양생된 균주로 사용되는 단일 콜로니를 수득하기 위해 LB 플레이트 상에 재-배열되었다.
- [0189] 생물학적 복제물에서 측정된 형질전환된 균주의 세포에 의한 비오틴 생산 (하기 섹션 1.6에 자세히 기재됨. 간 단히, 비오틴 생산은 0.00, 0.01 또는 0.1 mM IPTG로 보충된 mMOPS-kan에서 생물학적 복제물에 대해 재-평가되었다. 동시에, 각각의 형질전환된 균주의 세포의 성장률은 멀티스칸 FC에서 24시간의 기간 동안의 혐기성 성장을 위해 “빠른 진탕(fast shaking)” 으로 37°C에서 투명한 통기성 밀봉으로 밀봉된 미량 정량판(microtiter plate)에서 200 μ L mMOPS-kan 배지에서 측정되었다. 세포 성장은 30분 마다 OD620을 측정함으로써 모니터링되었다.
- [0190] 1.6 비오틴 생산을 정량화하기 위한 바이오어세이
- [0191] 전-배양물은 96 딥(deep)-웰 플레이트에서 400 μ L mMOPS-kan에서 선별된 단일 세포 콜로니로부터 각각 제조되었고, 16-18시간 동안 275 rpm에서 진탕하면서 37°C에서 인큐베이션되었다. 생산 배양물(production cultures)은 ~0.03의 초기 OD600을 제공하기에 충분한 전-배양물의 4 μ L을 갖는, 96 딥-웰 플레이트에서, 0.1 g/L 테스티오비오틴으로 보충된 400 μ L mMOPS-kan을 접종하고, 선택적으로 최종 농도가 최대 1 mM까지 IPTG를 포함함으로써 생산되었다. 이어서, 배양물은 24시간 동안 275 rpm 진탕으로 37°C에서 성장되었다. 배양물의 OD₆₀₀을 측정한 후 8분 동안 4000 G에서의 원심분리에 의해, 96 딥-웰 플레이트의 세포는 펠렛화(pelleted)되었다. 각 배양물 상층액으로부터의 상층액은 초순수(ultrapure water, Mili-Q)에서 0.05 nM 내지 0.50 nM 비오틴의 농도 범위까지 희석되었다. 동시에, 0.1 nM (0.024 μ g 비오틴/L) 내지 1 nM (0.24 μ g 비오틴/L)의 농도 범위에서 >5 비오틴 표준물(biotin standards)은 Milli-Q 물에서 제조되었다. 각 희석된 상층액의 15 μ L 및 각각의 비오

틴 표준은 미량정량판의 웰에 첨가되었다; 상기 각 웰은 플라스미드 pBS451을 포함하는 BS1011의 비오틴-결합된 하룻밤 동안의 배양물의 135 μ L를 포함하고, 상기 하룻밤 동안의 배양물은 제오신이 보충된 mMOPS에서 0.01의 초기 OD₆₂₀까지 희석되었다. 상기 플레이트는 통기성 밀봉으로 밀봉되었고, OD₆₀₀이 측정되기 전에 20시간 동안 275 rpm 진탕으로 37 °C에서 인큐베이션되었다. 비오틴 표준물의 범위를 사용하여, 이 바이오어쎈이로 수득된 비오틴 바이오어쎈이 검정 곡선은 도 5에 나타낸다.

[0192] 1.7 게놈 돌연변이의 동정

[0193] 모든 차세대 염기서열 분석(NGS: Next-Generation Sequencing) 데이터의 경우, CLC 게놈 워크벤치 버전 9.5.3 (CLC genomic workbench version 9.5.3, Qiagen에 의해 제공됨)이 모체 균주 게놈과 비교하여 선별된 균주의 세포의 게놈에서의 돌연변이(변이체 검출에 의한 단일 또는 일부 치환, 결실 또는 삽입 및 인델(InDels) 및 구조적 변이체에 의한 더 큰 삽입/결실)을 동정하기 위해 사용되었다. 80%의 컷-오프는 시퀀싱 과정에 의해 도입된 잘못된 뉴클레오타이드로부터의 게놈 돌연변이와 구별하기 위해, 돌연변이가 제공된 박테리아 균주의 세포로부터 분리된 DNA 분자 (게놈) 집단의 85% 이상에서 존재해야 한다는 것을 의미하는 "유의미한 돌연변이"로 정의하는데 사용되었다.

[0194] NCBI의 게놈 수탁 번호 CP009273은 참조 서열로서 사용되었고, 시퀀싱에 의해 서열이 입증된Keio Δ bioB 홀터 돌연변이 (scar mutation)를 고려하였다.

[0195] 1.8 *iscR* 돌연변이의 프로테오믹스(proteomics) 랜드스케이프(landscape)의 특성화

[0196] 1 mM IPTG 유도에서의 BS1353 + pBS412뿐만 아니라 0.025 mM IPTG 유도 수준에서의 BS1013 + pBS430, BS1011 + pBS412 및 BS1353 + pBS412의 단백질 함량은 LC-MS 및 효율적인 단백질 추출을 조합한 최근에 개발된 접근 방법에 의해 측정되었다 (Schmidt et al, 2015). 2.0% FDR의 펩티드 역치(threshold)에 따른 분석을 위해 식별되는 펩티드의 최소 개수로서, 3개의 펩타이드가 선택되었다. 단백질 발현에서 유의적인 변화는 스캐폴드 뷰어 4.7.5 (Scaffold Viewer 4.7.5)를 사용한 다중 테스트에 대한 Benjamini-Hochberg 보정을 통한 분산 분석 (ANOVA, Analysis of Variance)에 기초하여 0.5% 신뢰 구간으로 보고된다.

[0197] 0.5의 OD₆₀₀에 도달할 때까지, 약 10 세대 동안의 IPTG 유도로 mMOPS에서 성장되었다. 10⁸ 세포는 최대 속도에서 4°C에서의 원심분리에 의해 수확되었고; 아이스-콜드 PBS 버퍼에서 1회 세척되었고; 최대 속도에서 4°C에서의 원심분리에 의해 재-펠렛화되고 PBS 버퍼를 제거한 후 액체 질소에서 순간-동결 (snap-frozen)되었다.

[0198] 2. 결과

[0199] 2.1 BioB의 과발현은 유독하다

[0200] 저-카피 플라스미드로부터의 고유한 비오틴 신타아제 유전자 (*bioB*)를 보유하지만, IPTG-유동성 프레임시프트된 대장균 *bioB* 유전자(조기 종결 코돈으로 인해 비-기능성 비오틴 신타아제를 코딩함)를 발현하는 대장균은 IPTG가 있거나 없는 mMOPS-kan 배지에서 호기성으로 성장할 수 있다 이는 대장균 BW25113, BioB 프레임시프트 돌연변이의 지수 성장곡선을 나타낸 도 4 (왼쪽 패널)에 예시된다. 반대로, 대장균 녹-아웃 균주 Δ bioB에서 기능성 비오틴 신타아제 유전자 (*bioB*)의 과-발현은 성장에 유독하여, 유도기 (lag phase)에서 매우 유의적인 확장을 야기한다. 이는 저-카피 플라스미드 (Sc101 복제 기점) 상에서 IPTG-유도성 T5 프로모터로부터 대장균 *bioB* 유전자를 발현하는 대장균 녹-아웃 균주 Δ bioB의 성장을 나타낸 도 4 (오른쪽 패널)에 예시된다. 도 4에 나타낸 바와 같이, IPTG 수준의 증가에 반응한 *bioB* 발현의 증가의 유도는 (회색의 어둠) 유도기에 유의적으로 영향을 미치는 반면, 성장률은 약간 영향을 받는다 (검정색 박스).

[0201] 2.2 향상된 비오틴 생산 역가 (production titers)를 갖는 *iscR* 돌연변이 균주의 분리

[0202] 진화된 게놈 다양성을 갖는 대장균 라이브러리 (섹션 1.4 및 1.5 참조)는 *bioB* 유전자 발현에 대한 향상된 내성 및 증가된 비오틴 생산을 갖는 균주에 대해 스크리닝되었다. 선별된 균주의 전체 게놈 시퀀싱으로 3개의 특이한 돌연변이를 동정하였고, 이들 각각은 L15F, C92Y 및 H107Y 중에서 하나의 아미노산 치환을 갖는 *iscR* 폴리펩티드를 코딩하는 철-황 클러스터 조절자 (*iscR*) 유전자를 포함하고, 상기 코딩된 조절자의 아미노산 서열은 각각 서열번호 16, 18 및 20이다. 섹션 1.6 (및 도 5)에 기재된 바와 같이, 비오틴 생산 수준은 바이오어쎈이를 이용하여 측정되었다. 대장균 BW25113 Δ bioB 표준 균주 (reference strain)뿐만 아니라, 각각의 *iscR* 돌연변이 균주에 대한 비오틴 생산 역가는 2개의 상이한 IPTG 농도 수준의 부재 또는 존재에서 (증가된 IPTG는 더 진한 회색임) 0.1 g/l DTB로 보충된 mMOPS에서 성장된 4개의 생물학적 복제물 (검정색 점)에 대하여 도 6에 나타낸다.

표준 균주에서 비오틴 생산은 표준 균주의 성장에 유독한 IPTG 수준에 상응하는 0.01 mM 이상의 IPTG 수준에서 억제되었고 (도 4 참조), 반면에 *iscR* 돌연변이 균주는 0.01 - 0.1 mM IPTG에서 성장하였고 비오틴을 생산하였다. 3개의 *iscR* 돌연변이 균주 모두 0.01 mM의 IPTG 농도에서 표준 균주 (점선)에 비해 약 1.5배 더 많은 비오틴을 생산한다. *iscR* 돌연변이 균주는 0.1 mM의 IPTG 농도에서 표준 균주의 가장 높은 생산 역가 (~ 1.5 mg biotin/l)에 비해 최대 2배 더 많이 (~3.2 mg 비오틴/l) 생산한다.

[0203] 2.3 *IscR* H107Y 돌연변이 균주에서의 비오틴 생산 및 성장

[0204] *iscR* (H107Y) 돌연변이 균주의 성장 특성 (growth profile) 및 비오틴 생산 역가는 2개의 상이한 IPTG 유도 수준 (도 7A에서 0.01 mM) 및 0.05 mM (도 7B)에서 250 mL 진탕-플라스크 실험에서 0.1 g/l DTB로 보충된 50 mL mMOPS에서 특성화되었다. 낮은 IPTG 수준에서 (도 7A), *IscR* 돌연변이 균주 (어두운 회색) 및 표준 균주 (밝은 회색)는 ~ 1.1 mg 비오틴/l의 최종 역가로 성장 및 비오틴 생산 역가에 대하여 유사하였다. 그러나, 높은 IPTG 유도 수준에서 (도 7B에서 0.5 mM) 표준 균주 (밝은 회색)의 성장은 심각하게 억제되었던 반면에, *IscR* 돌연변이 균주는 낮은 IPTG 유도 수준에서와 동일한 성장 특성을 보여주었다. 또한, *IscR* 돌연변이 균주의 비오틴 생산 역가는 25시간의 성장 후에 최대 ~2.2 mg 비오틴/l까지, 약 2배로 증가되었다.

[0205] 2.4 *IscR* 돌연변이의 작용의 메커니즘

[0206] 도 6에 나타난 바와 같이, 향상된 비오틴 내성 표현형은 동정된 3개의 *IscR* 돌연변이 균주 모두에 대하여 명확하게 입증되었다. C92 돌연변이 (C92Y)의 비오틴 내성을 향상시키는 능력은 *IscR*의 [Fe-S] 클러스터 결합 특성에서 C92의 역할에 의한 것으로 제안된다. C92Y 돌연변이로 인한 [Fe-S] 클러스터 결합 특성의 손실은 *IscR*의 *Isc*-오페론 억제 반응을 불활성화시키는 것으로 제안된다. 동시에, C92Y 돌연변이 *IscR*에서 *IscR*의 프로모터 기능은 온전하게 유지되어, 다중 세포 과정에 필수적인 다른 경로를 활성화시키는데 그 기능을 유지하는 것으로 제안된다. *IscR*의 [Fe-S] 클러스터 결합 특성을 제공하는 유사한 필수적 역할은 H107에 기인하고; 여기에서 H107Y 돌연변이는 *대장균*에서 비오틴 내성을 유사하게 향상시킬 수 있다. *IscR*에서 L15F 또한 철-황 클러스터 결합을 방해하여, 철-황 클러스터 고갈을 부분적으로 극복하는 것으로 제안된다. 도 9는 DNA에 결합될 때 L15 및 H107의 위치를 나타내고 (hya, PDB 엔트리 4HF1), L15는 *IscR* 각각의 서브유닛의 내부에 위치함을 알 수 있다. 페닐알라닌은 류신보다 상당히 큰 아미노산이고, 단백질의 3차원 폴딩을 방해할 수 있다.

[0207] 2.5 *대장균* 균주 단독에서의 *isc*-오페론 또는 *suf* 오페론의 과발현은 비오틴 생산을 향상시키기에 충분하지 않다.

[0208] *isc*-오페론 (*iscSUA-hscBA-fdx*, *iscR* 유전자가 제외된 천연 *대장균* 오페론 구조와 상응함) 또는 *suf*-오페론 (*sufABCDSE*, 천연 *대장균* 오페론 구조와 상응함)의 *대장균*에서의 비오틴 생산에 대한 직접적인 효과를 측정하기 위해, 각각의 오페론은 강한 RBS 및 IPTG 유도성 T5 프로모터의 조절 하에 배치된 미디엄 카피 넘버 플라스미드 (p15A ori) 내로 클로닝하였다. *isc*- 또는 *suf*-오페론 대신에 슈퍼 폴더 녹색 형광 단백질 (sfGFP: super folder Green Fluorescent Protein)을 코딩하는 유전자를 포함하는 플라스미드는 대조군으로 적용되었다. 각각의 플라스미드는 IPTG-유도성 bioB 발현 플라스미드를 포함하는 *대장균* 균주의 세포 내로 형질전환되었다. IPTG-유도성 bioB 발현 플라스미드 뿐만 아니라, 1) IPTG-유도성 *isc*-오페론, 2) IPTG-유도성 *suf*-오페론 또는 3) IPTG-유도성 GFP (대조군): 중 하나를 포함하는 생물학적 3중 (triplicate) 콜로니는 100 µg/mL 암피실린 및 50 µg/mL 스펙티노마이신을 갖는 400 L mMOPS에서 낮은 (0.01 mM IPTG) 및 높은 (0.1 mM IPTG) 유도 하에서 배양한 후, 비오틴 생산에 대하여 분석되었다 (섹션 1.5에 기재된 바와 같음).

[0209] 그래프 (도 10)으로부터, 모든 균주에서 비오틴 생산은 IPTG-유도성임에도 불구하고; 검출 가능한 비오틴 생산 수준에 도달하기에 필요한 IPTG 농도는 도 6에 나타난 표준 및 돌연변이 *iscR* 균주와 비교하였을 때, 0.01 mM IPTG 내지 0.1 mM IPTG로 증가되었다. 또한, 비오틴 생산 역가는 *isc*- 또는 *suf*-오페론의 과발현에 의해 도 10의 sfGFP 균주와 비교하였을 때, 현저히 감소되었다. 또한, *isc*-오페론의 과발현은 *suf*-오페론의 과발현 보다 훨씬 더 비오틴 생산을 억제하였다. *iscR*에 싱글 포인트 돌연변이 (single point mutation)를 갖는 돌연변이 균주에서의 비오틴 생산의 관찰된 증가와 함께, 이들 균주에서 *isc*-오페론의 결과적인 억제-해제(de-repression)는 이들 균주에서의 개선된 비오틴 생산의 유일한/주요한 원인일 가능성은 낮다.

[0210] 2.6 BioB 단백질 함량은 비오틴 생합성과 관련이 있다

[0211] 야생형 및 돌연변이 백그라운드 균주에서 BioB 과발현의 분자적 효과를 조사하기 위해, 야생형 백그라운드 균주: BS1013 보유 (holding) pBS430; 비오틴 생산 플라스미드를 갖는 야생형 *iscR* 균주: BS1011 보유 pBS412; 및 bioB 생산 플라스미드를 갖는 돌연변이 *iscR* 균주: BS1353 보유 pBS412에 대하여 프로테오믹스 측정이 수행

되었다. 모든 균주는 0.1 g/L DTB 및 0.025 mM IPTG로 mMOPS에서 성장되었다. 후자 균주는 1 mM IPTG 유도에서 추가로 성장되었다. 세포는 프로테오믹스 분석을 위해 수확되었고, 나머지 세포 배양물은 다른 곳에 기재된 바이오어썰이를 이용하여 비오틴 생산이 측정되기 전에, 총 24시간 동안 인큐베이션이 유지되었다.

[0212] 그래프 (도 11)로부터, 측정된 비오틴 단백질 수준은 비오틴 생산과 밀접한 관련이 있다 (R^2 값 0.96). 선형 상관관계는 향상된 BioB 발현을 촉진시키는 것이 *IscR* 돌연변이 세포 공장에서 비오틴 생산을 향상시키는 핵심임을 나타낸다. 프로테오믹스 데이터의 ANOVA 분석은 추가적인 29 단백질에서 발현의 현저한 증가 (95% 신뢰 구간, p-값 0.00166)를 나타냈다. 이들 중에서 *isc*-오페론 (*IscA* 및 *IscS*) 및 *suf*-오페론 (*SufB* 및 *SufS*)의 구성원이 있다.

[0213] 2.7 *iscR* 녹아웃 돌연변이에서 비오틴 생산은 향상되지 않는다

[0214] *iscR* 유전자의 번역 녹아웃 (translational knockout)은 *iscR*에서의 22번 위치 상에서 글루탐산을 코딩하는 코돈 (E, GAA)을 종결 코돈 (*, TGA)으로 변환함으로써, MAGE에 의해, BW25113 $\Delta bioB$ 균주 내로 도입되었다. 상기 코돈의 성공적인 변환은 상기 영역의 PCR 증폭 및 이후의 생가 시퀀싱 (Sanger sequencing)에 의해 입증되었다. 야생형 *iscR*, *iscR* 녹아웃 (E22*), 및 돌연변이 *iscR* (C92Y)를 코딩하는 유전자를 갖는 균주는 IPTG-유도성 *bioB* 플라스미드 pBS412로 형질전환되었고, 상기 기재된 바와 같은 3개의 상이한 IPTG 유도 수준 (0, 0.01, 및 0.1 mM)에서 0.1 g/L DTB 및 50 μ g/L 카나마이신으로 보충된 mMOPS에서 성장된 생물학적 복제물 (n=3)에서 비오틴 생산에 대해 시험하였다.

[0215] IPTG 유도에 의해 *bioB* 발현을 유도한 경우 *iscR* 녹아웃 (*iscR* KO) 및 야생형 *iscR* (*iscR* WT) 사이에 비오틴 생산에 있어 유의적인 차이는 관찰되지 않았다. 이는 *iscR*을 녹아웃하는 것이 비오틴 생산을 개선시키지 않는다는 증거를 제공한다. *iscR* WT 및 *iscR* KO 균주 모두와 비교하여, *IscR* C92Y 치환을 코딩하는 돌연변이 *iscR*에서 비오틴 생산에 있어 유의적인 개선이 다시 관찰되었다.

[0216] 2.8 본 발명의 *iscR* 돌연변이 균주에서 드-노보 (De-novo) 비오틴 생산이 향상된다

[0217] *bioA* 유전자 및 전체 비오틴-오페론 ($\Delta bioB$ - $\Delta bioD$)이 결실되고, *iscR* WT, *iscR* H107Y 돌연변이 또는 *iscR* C92Y 유전자를 포함하는, BW25113 대장균 균주는 천연 대장균 *bioA* 및 *bioO* 오퍼레이터 상에 싱글 포인트 돌연변이 (타입 9 돌연변이, Ifuku et al., 1993)를 갖는 비오틴-오페론을 항시적으로 과발현하는 테트라사이클린 내성 플라스미드로 형질전환되었다. 상기 기재된 바와 같이 (도 13), 0.1 g/L DTB를 첨가하거나 첨가하지 않고 10 μ g/ml 테트라사이클린을 갖는 mMOPS (2 g 글루코스/L)에서의 생물학적 복제물 (n=4)에서 3개의 상이한 균주에 대하여 비오틴 생산이 평가되었다.

[0218] 기질, DTB가 증식 배지에 첨가된 경우, 모든 3개의 균주에서 비오틴 역가의 현저한 증가가 관찰되었고, 이는 DTB를 비오틴으로 전환하는, *bioB* 효소 반응 자체가 더이상 이들 균주에서 비오틴 생산에 장애물 (bottleneck)이 아님을 나타낸다 (도 13). 또한, *iscR* WT 균주와 비교하여 *iscR* 돌연변이 균주 모두에서 글루코스로부터 비오틴의 드-노보 생산의 현저한 증가가 관찰되었다. 이들 결과를 고려하면, 본 발명의 모든 돌연변이 *iscR* 균주는 직접적인 전구체, DTB, 및 글루코스 모두로부터 향상된 비오틴 생산을 지지하는 것으로 추론될 수 있다.

[0219] 실시예 2: 리포산 생산을 향상시킬 수 있는 유전자 변형 대장균 균주의 설계 및 특성화

[0220] 실시예에서 사용된 하기의 *에세리키아 콜라이* 균주는 하기에 나열된다.

표 3

[0221] 균주

명칭	설명
BS1912	<i>E. coli</i> K-12 BW25113로부터 유래된 $\Delta lipdA$
BS2114	<i>iscR</i> 에서 H107Y 돌연변이를 포함하는 BS1912 파생체

[0222] 실시예에서 사용된 하기의 플라스미드는 하기에 나열된다.

표 4

[0223]

플라스미드

명칭	설명
pBS993	AceF [서열번호 119] 발현의 추가적인 항시성 발현을 갖는 T5 lacO 억제 프로모터 [서열번호 234]로부터의 Lipid A [서열번호 103] 과발현 플라스미드 (kanR, SC101)
pBS1037	AceF [서열번호 119] 발현의 낮은 RBS 강도 및 SC101 대신에 p15A 복제 기점을 갖는 pBS993 파생체
pBS451	항시적으로 발현되는 GFP [서열번호 287] (zero, p15A)

[0224]

플라스미드 pBS1037 상에 클로닝된, LipA [서열번호 103]를 코딩하는 IPTG-유도성 전이 유전자는, 실시예 1에 기재된 바와 같이, 천연 *iscR* 유전자를 포함하는 *대장균* 숙주 균주 또는 천연 *iscR* 유전자가 C92Y 또는 H107Y 치환을 갖는 IscR 단백질을 코딩하는 돌연변이 *iscR* 유전자로 치환된 *대장균* 숙주 균주 내로 도입되었다; 상개 2개의 균주는 bioB 또는 *lipA*의 녹-아웃을 더 포함한다. 상기 균주들은 0.1 mM 비오틴 (Δ bioB 균주를 위함), *lipA* 발현의 유도를 위한 IPTG 및 기질로서 0.6 g/l 옥탄산으로 보충된 mMOPS 배지 (섹션 1.3에 기재된 바와 같은)에서, 24시간 동안 37°C에서, 상층액에서의 자유(free) 리포산의 측정 전에, 배양되었다. Δ *lipA* 균주 (BS1912 및 BS2114, 표 3)의 경우 37°C에서 24시간의 성장이 뒤따랐다.

[0225]

기재된 균주들의 배양된 세포에 의해 생산된, 리포산은 pBS451을 포함한 BS1912를 이용한 섹션 1.6에 기재된 것과 유사한 바이오어쎄이를 이용하여 상층액으로부터 측정되었다. 리포산의 정량화를 위한, 성장-기반의 바이오어쎄이는 리포산을 합성할 수 없는 영양요구성 *대장균* 단일 Δ *lipA* 돌연변이 균주를 이용하여 수행되었다 (Herbert and Guest, 1975) (pBS451을 포함하는 BS1912). 상기 상층액에서의 자유 리포산 농도는 리포산의 단독 공급원으로서 생산 균주로부터 회수된 리포산으로 보충되고, 탄소원으로서 50 nM na-숙시네이트를 갖는 최소 배지에서의 리포산 영양요구성 균주의 성장을 측정함으로써 결정되었다. 공지된 농도 범위의 리포산 표준으로 보충된 최소 배지 상에서 영양요구성 균주가 성장되는 경우와, 리포산 바이오어쎄이 검정 곡선은 병렬로 수행되었다.

[0226]

상기 시험은 IscR 단백질의 돌연변이 형태 (C92Y 또는 H107 치환을 갖는 IscR 단백질)를 코딩하는 유전자를 포함하는 *대장균* 균주에서의 *LipA* 유전자의 과-발현은, IscR 단백질의 천연 형태를 코딩하는 유전자를 포함하는 모체 *대장균* 균주에서의 *LipA* 유전자의 과발현과 비교하여, 더 안정적인 생산 및 리포산 역가의 80% 증가를 나타냄을 입증한다 (도 14). 따라서 *iscR* WT 균주 (pBS993을 갖는 BS1011)의 생산 역가에서의 표준 편차는 2.73인 반면에 *iscR* C92Y (pBS993을 갖는 BS1375)의 경우 1.42이고 *iscR* H107Y (pBS993을 갖는 BS1353)의 경우 0.11로 낮다 (균주 표준에 대한 표 1 참조). 개별적인 균주의 평균 생산 역가에 기초하여, 리포산 생산은 WT 균주와 비교하여 돌연변이 균주에서 1.79배 개선되었다 (도 14 참조).

[0227]

WT *iscR* 균주 (삼각형, pBS1037을 포함하는 BS1912) 및 *iscR* 돌연변이 균주 (사각형, pBS1047을 포함하는 BS2114) 둘다의 경우에서, LipA의 과발현은 IPTG에 의한 LipA의 증가된 유도에 반응하여 성장률이 감소하는 명확한 경향을 나타내었다 (더 어두운 음영을 넣은 회색, 도 15). 그러나, WT *iscR* 균주의 성장률은 0.01 mM 내지 0.03 mM 내지의 시험된 모든 IPTG 유도 수준에서 돌연변이 *iscR* 균주와 비교하여 보다 극심하게 감소되었다 (도 15 참조).

[0228]

실시예 3 티아민 생산을 향상시킬 수 있는 유전자 변형 *대장균* 균주의 설계 및 특성화

[0229]

실시예에서 사용된 하기의 *에세리키아 콜라이* 균주는 하기에 나열된다.

표 5

균주

[0230]

명칭	설명
BS750	코딩된 TMP 키나아제에서 G133D 치환을 일으키는 천연 <i>thiL</i> 유전자: 코돈 133에서 GGT로부터 GAC로의 점 돌연변이를 포함하는 BS1013 파생체, BW25113 Δ <i>thiP</i>
BS2019	<i>iscR</i> 에 C92Y 돌연변이를 포함하는 BS750 파생체
BS2020	<i>iscR</i> 에 H107Y 돌연변이를 포함하는 BS750 파생체

[0231]

실시예에서 사용된 하기의 플라스미드는 하기에 나열된다.

표 6

[0232]

명칭	설명
pBS140	<i>thiC</i> 오페론 (apFAB46 프로모터 [서열번호 147] 및 apFAB377 종결자 [서열번호 153]에 기능적으로 연결됨) 및 <i>thiM</i> 오페론 (apFAB71 프로모터 [서열번호 149] 및 apFAB378 종결자 [서열번호 152]에 기능적으로 연결됨)의 조합으로 구성됨; <i>대장균</i> 티아민 경로 유전자 <i>thiCEFSGHMD</i> 를 포함하는 벡터
pBS100	pBS140의 구성에 사용되기 위한 빈 벡터
pBS93	pFAB70 프로모터 [서열번호 148] 및 apFAB381 종결자 [서열번호 154]에 기능적으로 연결된 <i>대장균</i> 에서의 발현에 최적화된 애기장대 (<i>Arabidopsis thaliana</i>) AT5G32470.1 포스파타아제 코돈을 코딩하는 합성 유전자 (synthetic gene)를 포함하는 벡터
pBS209	<i>대장균</i> 으로부터 추가적인 <i>thiC</i> 를 갖는 pBS140에 기반한 플라스미드

[0233]

실시에 1에 기재된 바와 같이, 플라스미드 pBS140에 클로닝된 티아민 경로 유전자 *thiCEFSGHMD*는 천연 *iscR* 유전자를 포함하는 *대장균* 숙주 균주 (BS750) (표준 균주)뿐만 아니라 상기 천연 *iscR* 유전자가 C92Y 또는 H107Y 치환을 각각 갖는 *IscR*을 코딩하는 돌연변이 *iscR* 유전자에 의해 치환된 이 표준 균주 (BS2019 및 BS2020)의 파생체 내로 도입되었다. 상기 균주들은 딥 배양 플레이트의 개별적인 웰에서 24시간 동안 37°C에서 mMOPS 배지 (섹션 1.3에 기재된 바와 같음)에서 배양되었다.

[0234]

기재된 균주들의 배양된 세포들에 의해 생산된 세포 외 및 세포 내 티아민, TMP 및 TPP는 다음과 같이 회수되고 추출되었다: 각 배양물의 0.4 mL는 배양 플레이트에서 5분 동안 4000 x g에서 원심분리에 의해 4°C에서 수확되었다. 나머지 모든 단계는 얼음 상에서 수행되었다. 세포 외 TPP, TMP 및 티아민의 분석을 위해 상층액의 40 µg은 부드럽게 제거되었다. 남은 상층액을 따라낸 후, 배양 플레이트는 잔류 배지를 제거하기 위해 뒤집은 후 불텍싱되었다 (vortexed). 100 µL 아이스-콜드 HPLC 그레이드 메탄올 (HPLC grade methanol)을 배양 플레이트의 각 웰에 첨가되었고; 세포는 다시 불텍싱되었다. 얼음 상에서 최소 20분 동안 인큐베이션 후, 세포 잔해물은 5분 동안 4000 x g에서의 원심분리에 의해 펠렛화되었다. 상층액은 추가적인 분석을 위한 세포 내 추출물로서 사용되었다.

[0235]

형광 검출기를 이용하여 TPP, TMP 및 티아민을 검출하기 위해, 각 배양물에 의해 생산된 티아민 화합물은 강한 형광성인 티오크롬으로 유도체화되었다 (derivatized). 모든 단계는 실온에서 수행되었다. 세포 외 및 세포 내 추출물의 40 µL 부피는 4M 포타슘 아세테이트의 80 µL에 첨가되었고, 피펫팅 (pipetting)에 의해 혼합되었다. 새로 제조된 7M NaOH에서의 3.8 mM 포타슘 페리사이아나이드 (potassium ferricyanide) 40 µL가 첨가되고 혼합되었다. 새로 제조된 포화된 KH₂PO₄에서의 0.06% H₂O₂ 40 µL의 첨가에 의해 상기 반응은 억제되었다 (quenched). 상기 추출물은 6M HCl 47 µL의 첨가에 의해 중화되었고 이후 HPLC 또는 멀티스칸을 이용한 직접적인 형광 측정에 의해 분석되었다. 유도체화된 모든 화합물은 상기 분석된 추출물과 병렬하여 티오크롬으로 유도체화된 새로 제조된 TPP, TMP 및 티아민 표준의 형광 표준 곡선을 이용하여 정량화되었다.

[0236]

상기 시험은 *IscR* 단백질의 돌연변이 형태 (C92Y 또는 H107Y 치환을 갖는 *IscR* 단백질)를 코딩하는 유전자를 포함하는 숙주 *대장균* 균주 (BS2019 및 BS202)에서의 TMP 포스파타아제 유전자 (At5g32470)와 조합하여 *thiC* 유전자 및 *thiH* 유전자를 포함한 티아민 경로 유전자의 과-발현은, *IscR* 단백질의 천연 형태를 코딩하는 유전자를 포함하는 숙주 *대장균* 균주 (BS750)에서의 과-발현과 비교하여, 티아민, TMP 및 TPP, 특히 티아민의 생합성을 향상 시킴을 입증한다.

[0237]

보다 구체적으로, 상기 시험은 WT *iscR*을 갖는 균주 (BS750)와 pBS140을 사용하는 경우 *iscR* 돌연변이를 코딩하는 균주 (BS2020, H107Y 또는 BS2019, C92Y) 사이에 티아민 (티아민, TMP 및 TPP)의 OD-표준화된 세포 외 생산에 있어 1.43 배의 증가를 나타내었다 (도 16).

[0238]

실시에 4 비오틴을 생산할 수 있는 유전자 변형 *대장균* 균주의 생산성을 증가시키기 위한 플라보독신/페레독신 환원 효소 (Fpr)의 과발현 및 플라보독신 (F1dA) 환원 시스템

[0239]

1 방법

[0240]

1.1 : 실시예에서 사용된 하기의 *에세리키아 콜라이* 균주는 하기에 나열된다.

표 7

[0241]

균주

명칭	설명
BS1013	다음의 유전자형을 갖는 <i>E. coli</i> K-12 BW25113 모체 균주: <i>rrnB3</i> Δ (<i>lacZ</i> 4787 <i>hsdR</i> 514 Δ (<i>araBAD</i>)567 Δ (<i>rhaBAD</i>)568 <i>rph</i> -1
BS1011	<i>E. coli</i> K-12 BW25113로부터 유래된 Δ <i>bioB</i> (JW0758-1)
BS1353	<i>iscR</i> 에 H107Y 돌연변이를 포함하는 BS1011 파생체
BS1615	Δ <i>bioAFCD</i> 의 추가적인 결실을 갖는 BS1011 파생체
BS1937	IPTG-유도성 <i>BioB</i> 발현을 제공하는 pBS679 플라스미드를 포함하는 BS1615 파생체
BS2185	IPTG-유도성 <i>BioB</i> 발현을 제공하는 pBS679 플라스미드 및 <i>FldA</i> -Fpr 항시적 발현을 제공하는 pBS1112를 포함하는 BS1615 파생체
BS2707	IPTG-유도성 <i>BioB</i> 발현을 제공하는 pBS679 플라스미드 및 GFP 항시적 발현을 제공하는 pBS1054를 포함하는 BS1615 파생체

[0242]

실시예에서 사용된 하기의 플라스미드는 하기에 나열된다.

표 8

[0243]

플라스미드

명칭	설명
pBS679	T5 <i>lacO</i> 억제 프로모터 [서열번호 25]로부터의 <i>BioB</i> [서열번호 22] 과발현 플라스미드 (<i>ampR</i> , pSC101) 및
pBS1054	apFAB378 종결자 [서열번호 292]를 갖는 apFAB309 항시적 프로모터 [서열번호 291]로부터의 GFP [서열번호 276] 과발현 플라스미드 (<i>kanR</i> , pBR322)
pBS1112	apFAB306 항시적 프로모터 (apFAB306- <i>FldA</i> -Fpr 유전자-apFAB378 종결자 [서열번호 288])로부터의 <i>FldA</i> -Fpr 과발현 플라스미드 (<i>kanR</i> , pBR322)

[0244]

*BioB*를 코딩하는 IPTG-유도성 전이 유전자는 플라스미드 pBS679 상에 클로닝되었고; GFP를 코딩하는 항시적으로-조절되는 전이 유전자는 플라스미드 pBS1054에 클로닝되었고; *FldA*-렉을 코딩하는 합성 오페론 (synthetic operon)을 포함하는 항시적으로-조절되는 전이 유전자는 플라스미드 pBS1112 상에 클로닝되었다. pBS679는 실시예 1에 기재된 바와 같이, 천연 *iscR* 유전자가 H107Y 치환을 갖는 *IscR* 단백질을 코딩하는 돌연변이 *iscR* 유전자에 의해 치환되고, BS1937 균주를 생성하는 *bioAFCD* 유전자의 녹-아웃을 더 포함하는 대장균 숙주 균주 (BS1615) 내로 도입되었다. 상기 BS1937 균주는 BS2707 (대조군 균주) 및 BS2185 균주 각각을 생성하도록 플라스미드 pBS1054 또는 pBS1112로 더 형질전환되었다.

[0245]

상기 균주들은 적합한 항생제, *BioB*-매개 촉매를 위한 기질로서 0.1 g/l을 가지고, *BioB* 유전자의 발현을 유도하기 위한 0, 0.01, 0.025, 0.05, 0.075 또는 0.1 mM IPTG로 보충된 mMOPS 배지 (실시예 1.3에 기재된 바와 같음)에서 배양되었다. 상기 세포들은 딥 웰 배양 플레이트의 개별적인 웰에서 24시간 동안 37°C에서 인큐베이션되었다. 최종 OD가 예측되었고, 상층액은 원심분리에 의해 수확되었고, 실시예 1.6에 의해 기재된 바와 같이 비오틴 바이오어췌이에 의해 상층액으로부터 비오틴이 정량화되었다.

[0246]

균주 BS1937에 대한 도 12 및 도 17에 나타난 바와 같이, 유전적으로 변형된 내인성 *iscR* 유전자를 포함하는 대장균 세포에서 *BioB* 유전자 발현이 IPTG 농도의 증가에 의해 유도된 경우, 상기 세포는 상응하는 비오틴 생산에 있어 점진적인 증가를 나타낸다. 이들 유전자 변형 세포에서 비오틴 생산은 이의 모체 균주 BS1937 및 *FldA*-Fpr 대신에 GFP를 코딩하는 전이유전자를 발현하는 대조군 균주와 비교하여, *FldA*-Fpr을 코딩하는 전이 유전자의 공동-발현 (BS2185 균주)에 의해 더 향상된다.

[0247]

IscR 단백질의 돌연변이 형태 (H107Y 치환을 갖는 *IscR* 단백질)를 코딩하는 유전자, *BioB* (pBS679) 및 *FldA*-Fpr 유전자 (pBS1112)의 과발현을 위한 플라스미드를 포함하는 BS2185 균주의 비오틴 생산은 대조군 균주 BS1937 (*FldA*-Fpr 유전자의 과발현 없음)과 비교하여 2.12-배 향상된다 (도 18).

[0248]

참고문헌

[0249]

Datsenko KA, Wanner BL. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* ;97(12) : 6640-5.

[0250]

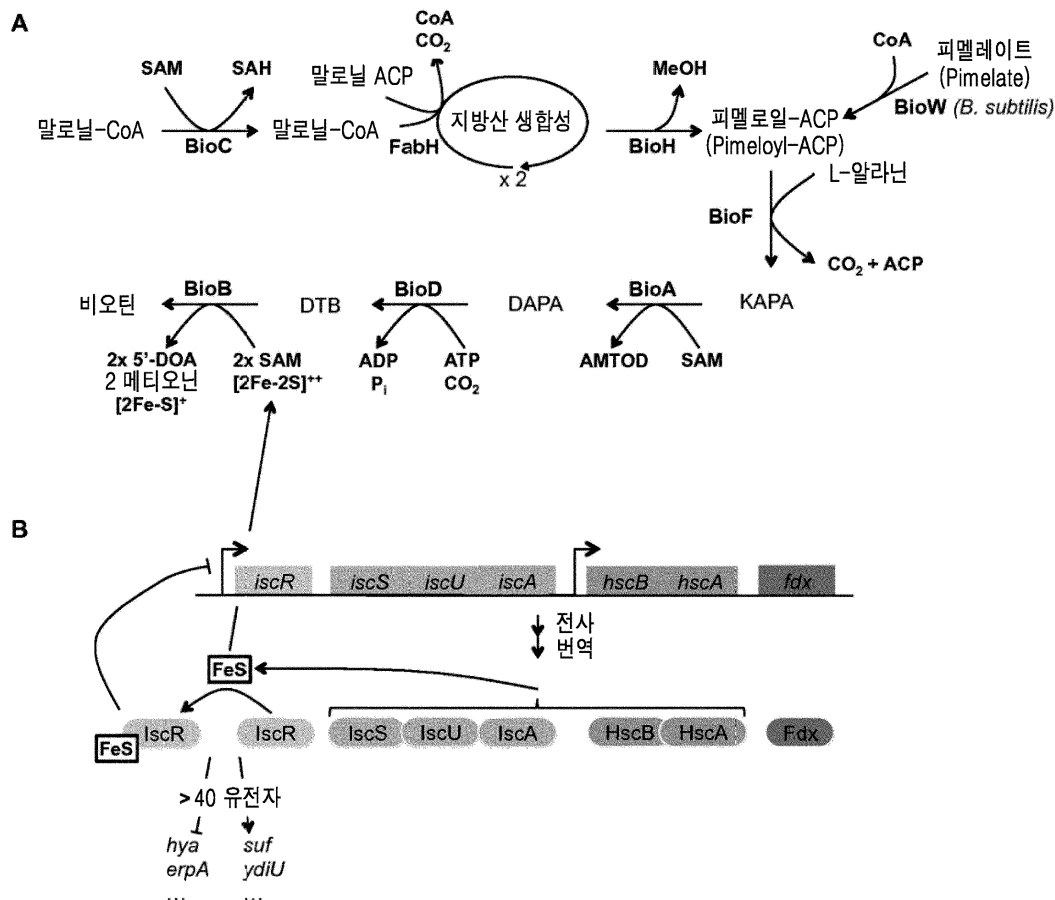
Fleischhacker, A. S. *et al.* (2012) Characterization of the [2Fe-2S] cluster of *Escherichia coli*

transcription factor IscR. Biochemistry 51 : 4453-4462.

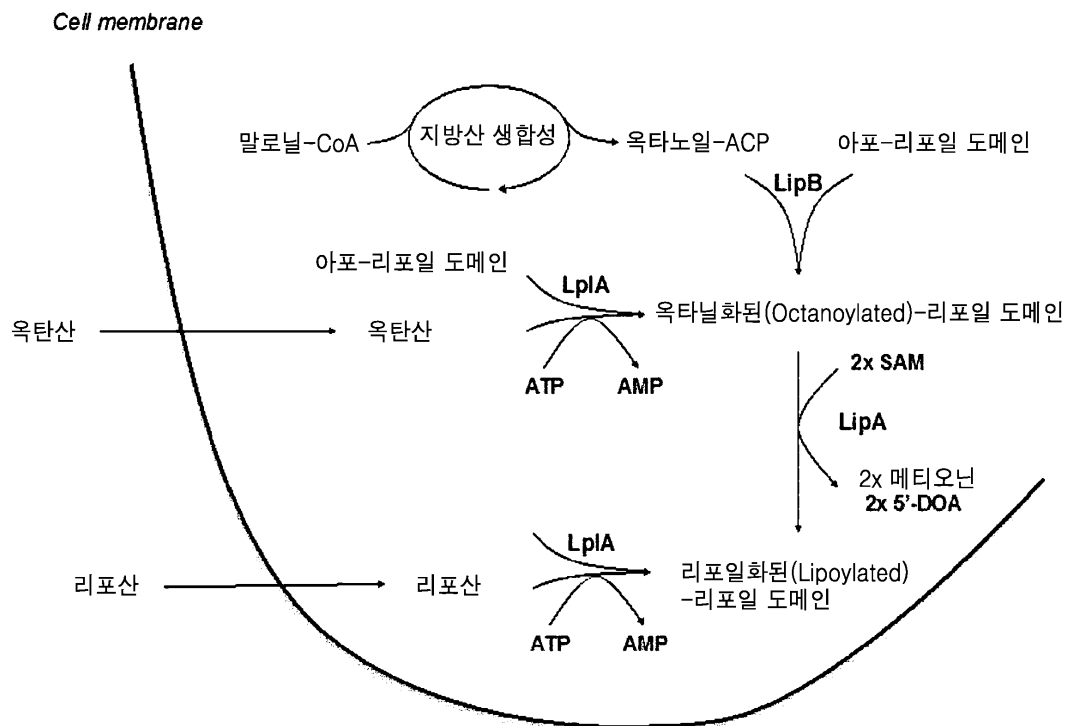
- [0251] Giel, J. L., Rodionov, D., Liu, M., Blattner, F. R. & Kiley, P. J. (2006) IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O₂-regulated genes in *Escherichia coli*. Mol. Microbiol. 60, 1058-1075.
- [0252] Herbert AA, Guest JR. (1975) Lipoic acid content of *Escherichia coli* and other microorganisms. Arch Microbiol. ; 106(3) :259-66. Epub 1975/12/31. pmid: 814874
- [0253] Ifuku, O. et al. Sequencing analysis of mutation points in the biotin operon of biotin-overproducing *Escherichia coli* mutants. Biosci Biotechnol Biochem 57, 760-765 (1993).
- [0254] Ifuku, O. et al., (1995) "Molecular analysis of growth inhibition caused by overexpression of the biotin operon in *Escherichia coli*." Bioscience, biotechnology, and biochemistry 59(2): 184-189.
- [0255] Martin, J. E. & Imlay, J. A. (2012) Replication during periods of iron starvation. 80, 319-334.
- [0256] Py, B. & Barras, F. (2010) Building Fe-S proteins: bacterial strategies. Nat. Rev. Microbiol. 8, 436-446.
- [0257] Schmitt, A, Kochanowski K., Vedelaar S., Ahrne E., Volkmer B., Callipo L., Knoop K., Bauer M., Aebbersold R., & Heinemann M., (2015) The quantitative and condition-dependent *Escherichia coli* proteome. Nature Biotechnology 2015; doi: 10.1038

도면

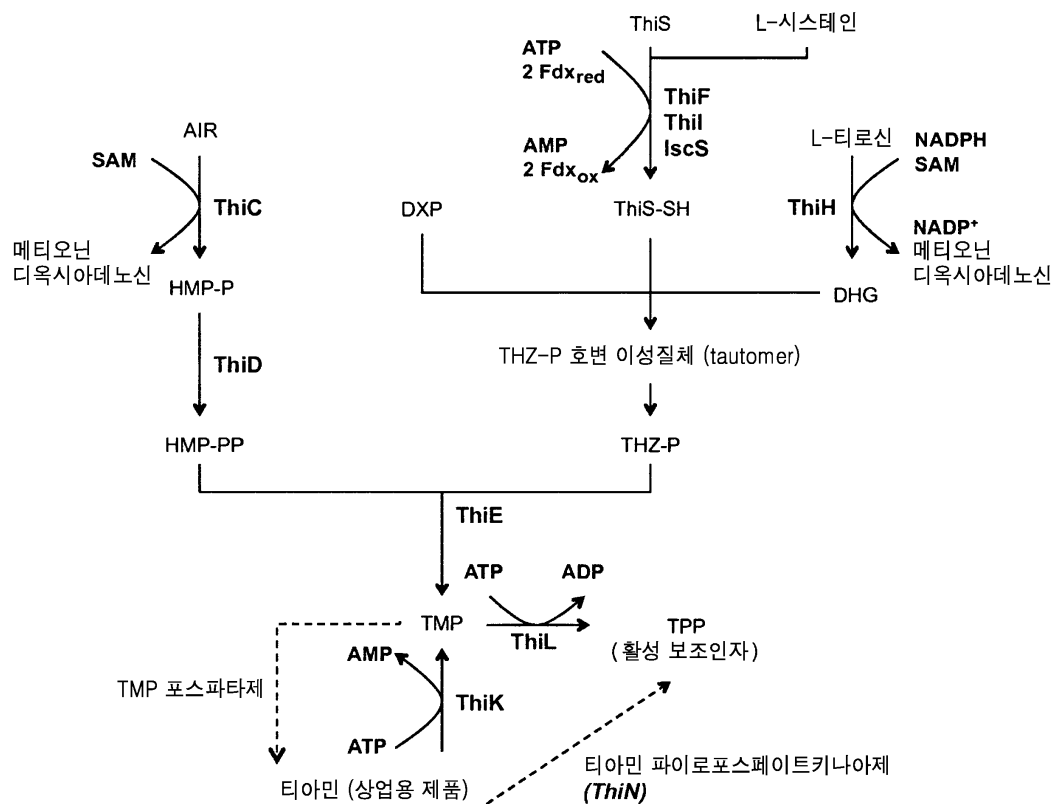
도면1



도면2

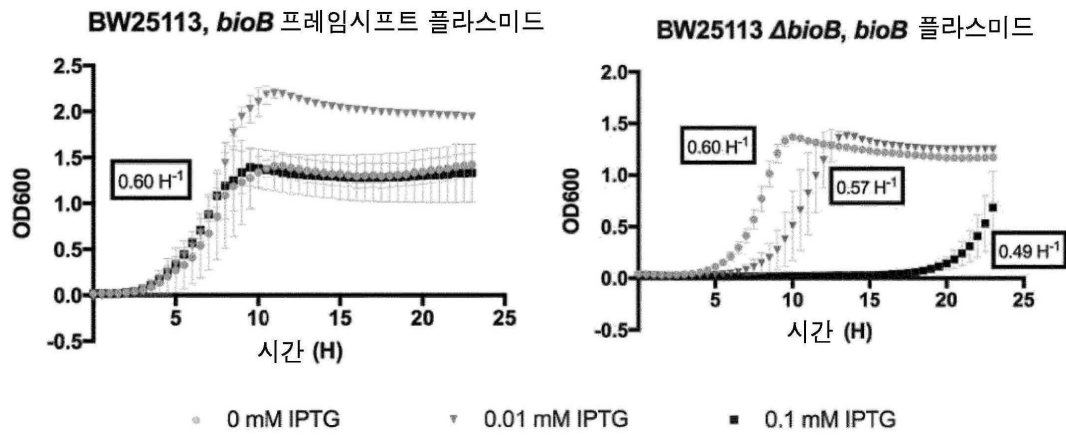


도면3



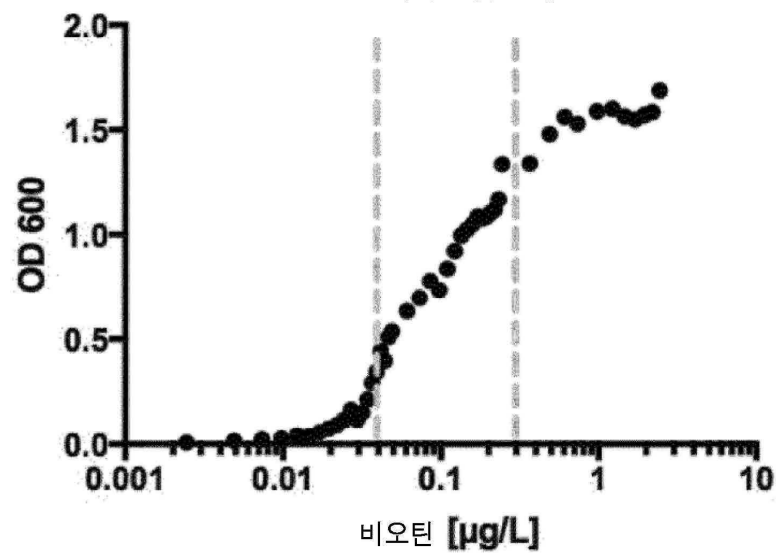
도면4

증식 프로파일

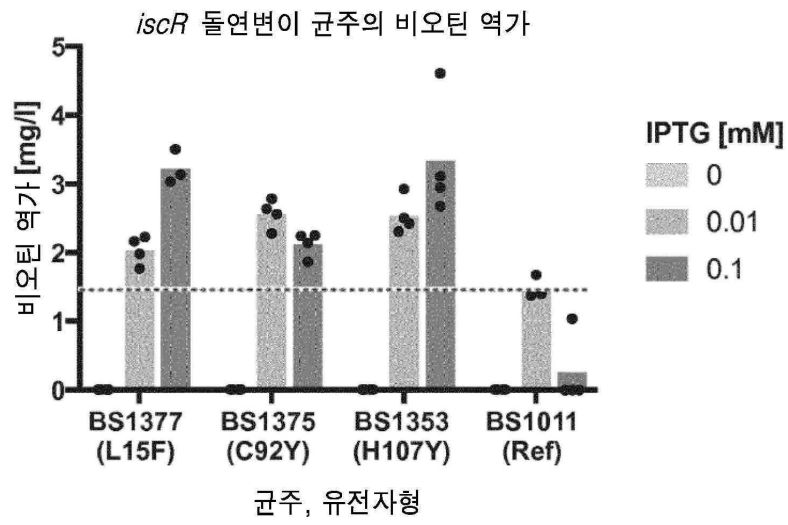


도면5

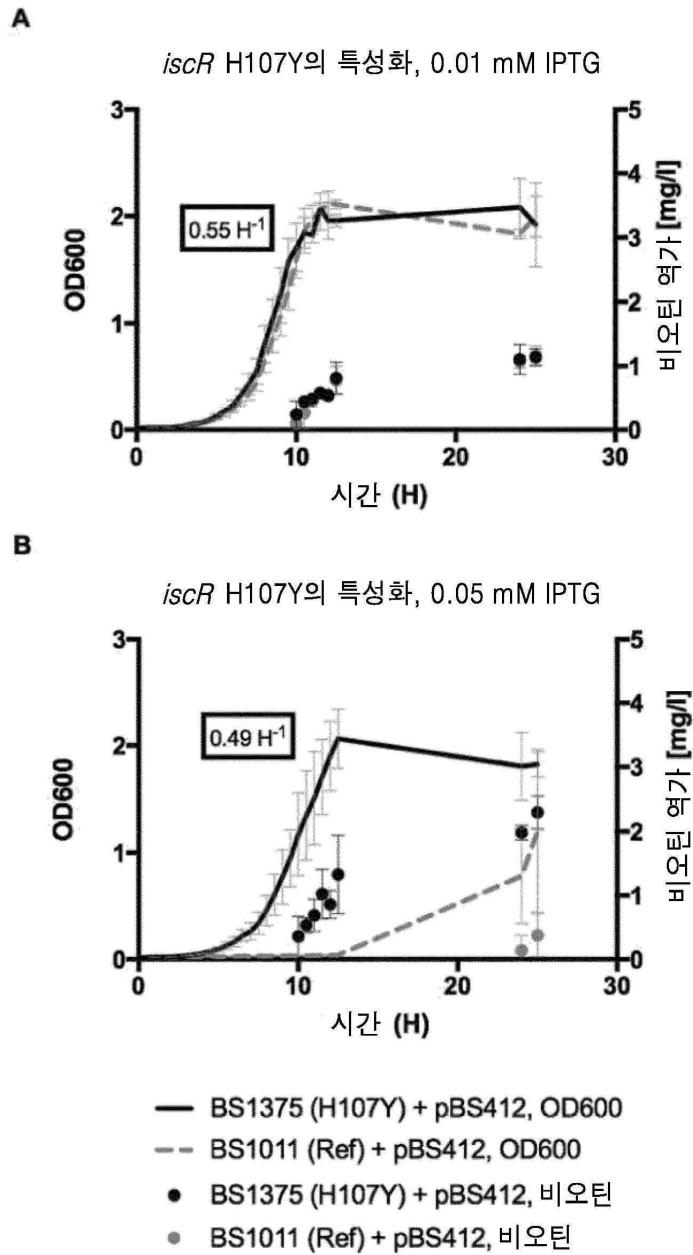
바이오어쎈이 검정 곡선



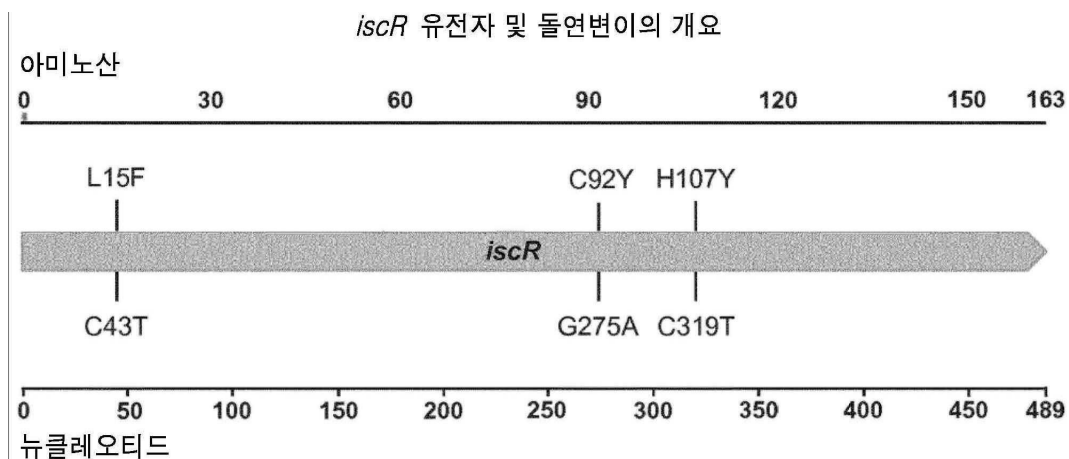
도면6



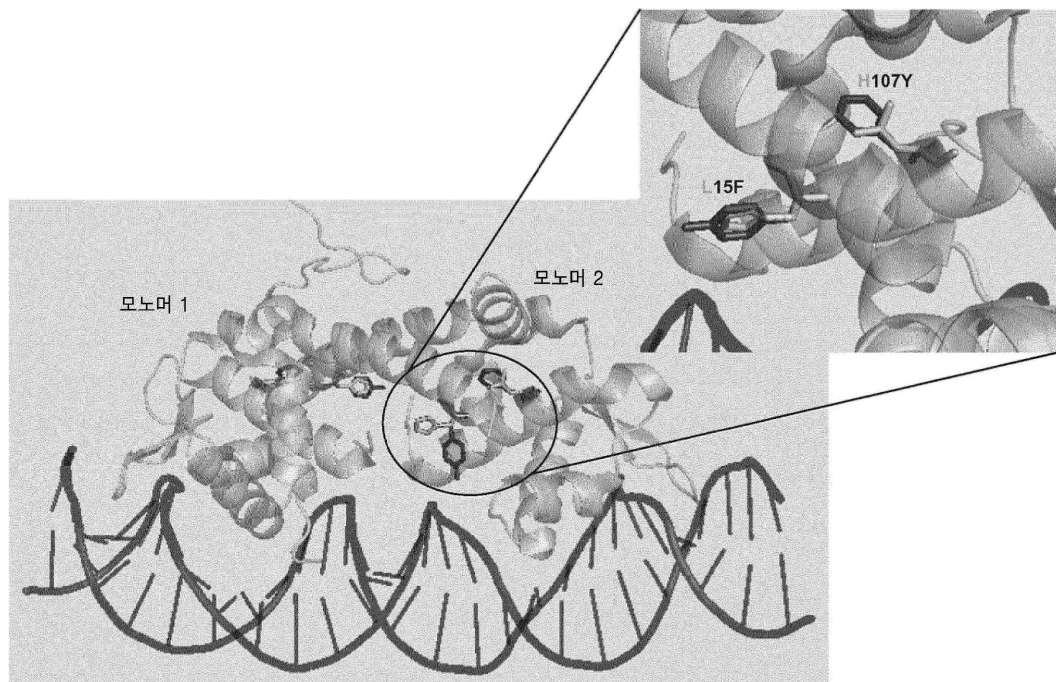
도면7



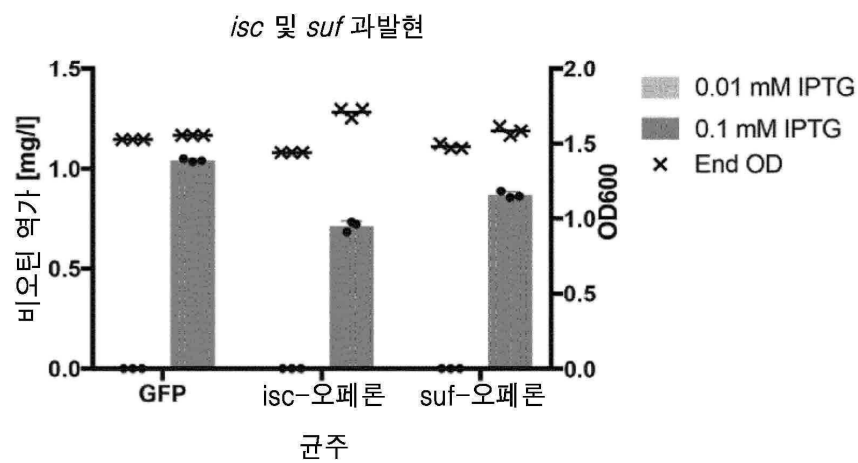
도면8



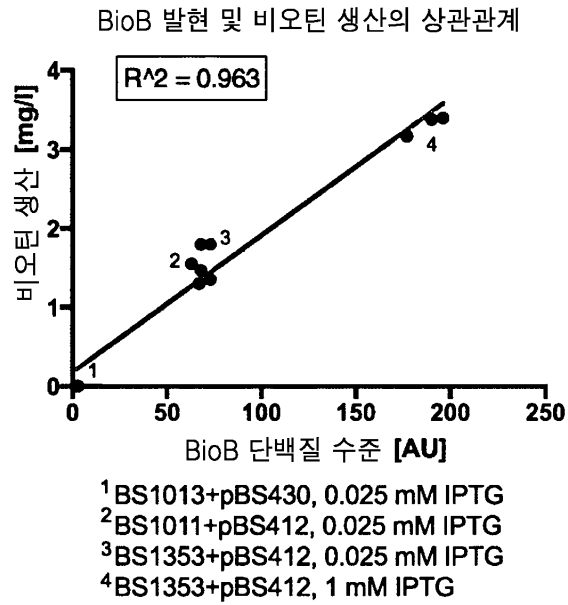
도면9



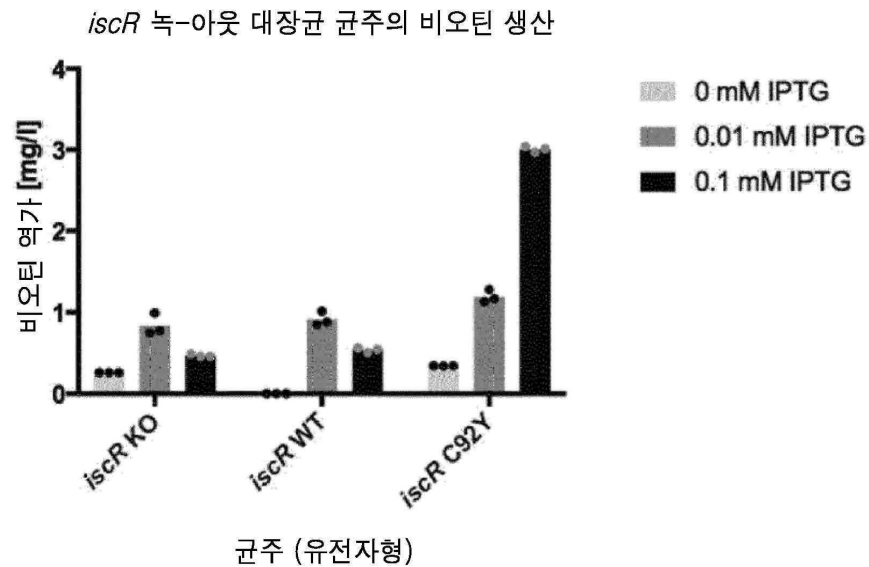
도면10



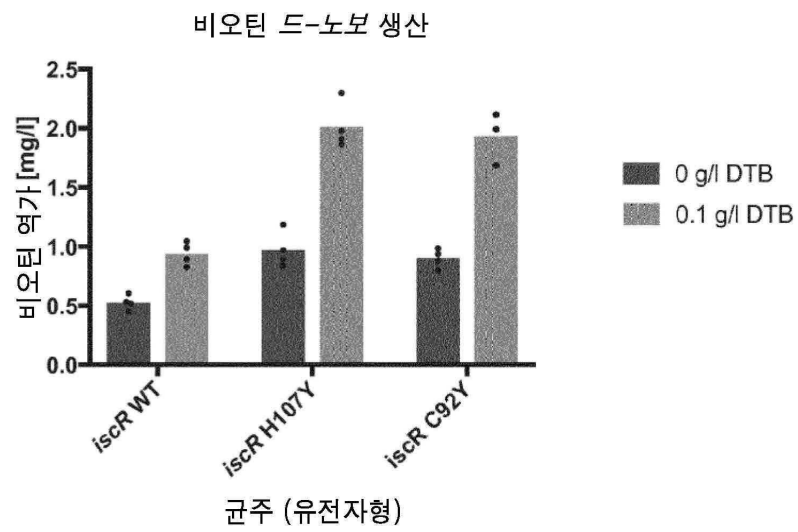
도면11



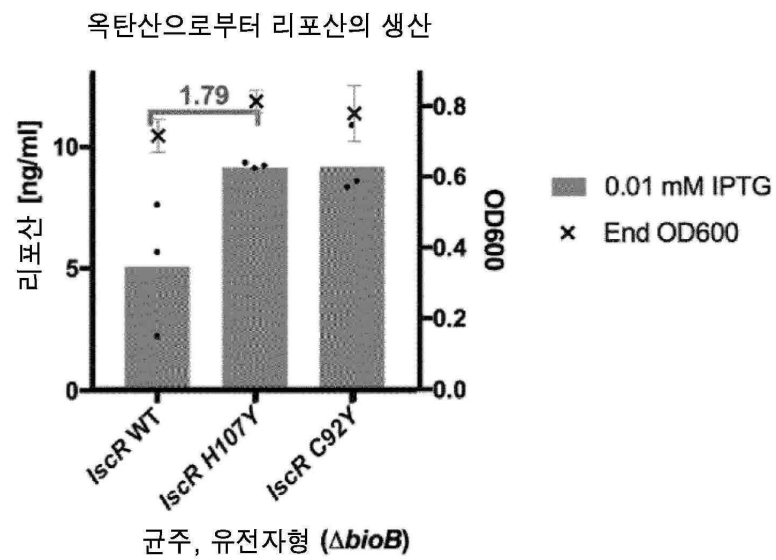
도면12



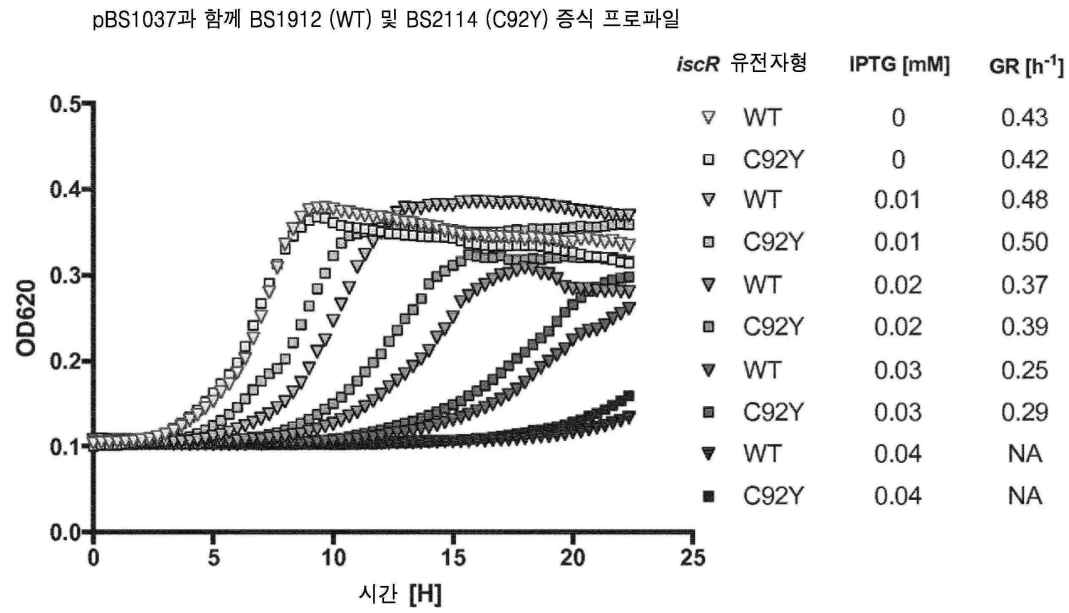
도면13



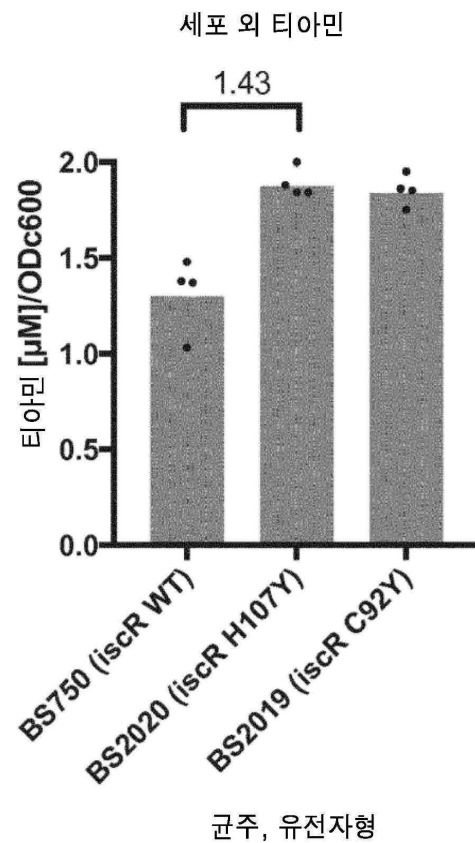
도면14



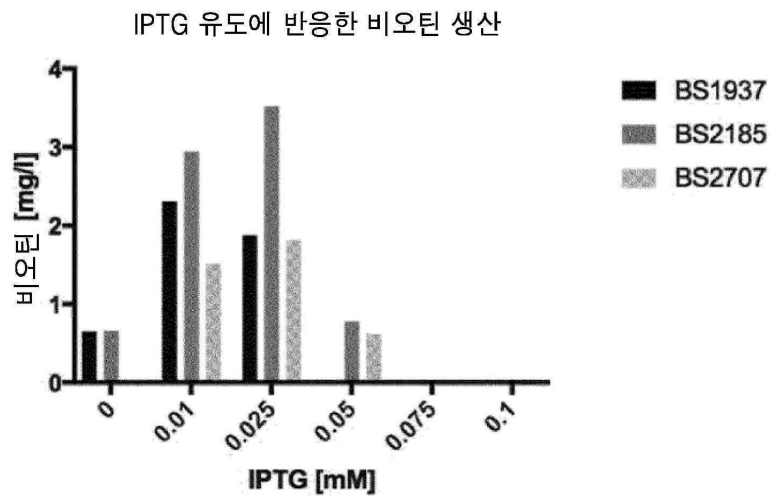
도면15



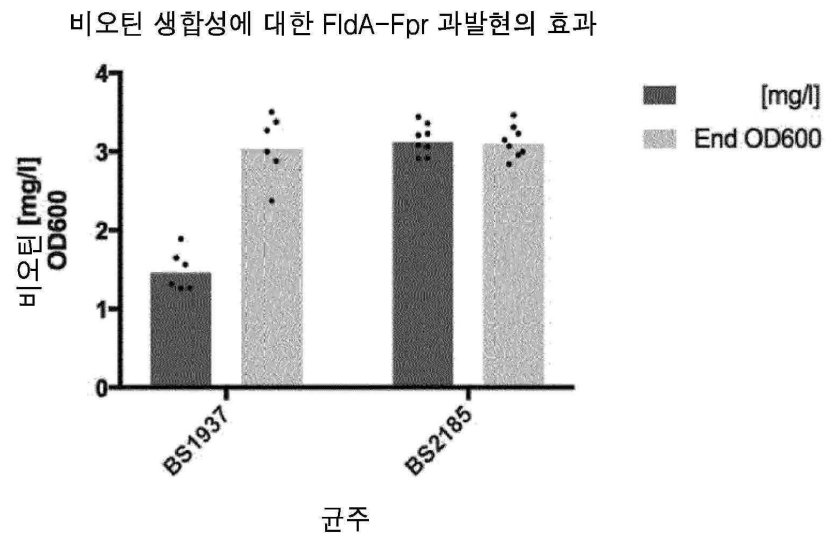
도면16



도면17



도면18



서열 목록

SEQUENCE LISTING

<110> Biosyntia ApS

<120> Cell factory having improved iron-sulfur cluster delivery

<130> P2296PC00

<150> EP17181503.8

<151> 2017-07-14

<160> 292

<170> PatentIn version 3.5

<210> 1

<211> 489

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(489)

<223> iscR WT gene encoding Iron Sulfur Cluster Regulator protein
(IscR)

<400> 1

atg aga ctg aca tct aaa ggg cgc tat gcc gtg acc gca atg ctt gac 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15
gtt gcg ctc aac tct gaa gcg ggc ccg gta ccg ttg gct gat att tcc 96

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30
gaa cgt cag gga att tcc ctt tct tat ctg gaa caa ctg ttt tcc cgt 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45
ctg cgt aaa aat ggt ctg gtt tcc agc gta cgt gga cca ggc ggt ggt 192

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60
tat ctg tta ggc aaa gat gcc agc agc atc gcc gtt ggc gaa gta att 240

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

65 70 75 80

agc gcc gtt gac gaa tct gta gat gcc acc cgt tgt cag ggt aaa ggc 288

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

85 90 95
ggc tgc cag ggc ggc gat aaa tgc ctg acc cac gcg ctg tgg cgt gat 336

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp

100 105 110
ttg agc gac cgt ctc acc ggt ttt ctc aac aac att act tta ggc gaa 384

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

115 120 125

ctg gtt aat aac cag gaa gtg ctg gat gtg tct ggt cgt cag cat act 432
 Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Thr
 130 135 140

cac gac gcg cca cgc acc cgc aca caa gac gcg atc gac gtt aag tta 480
 His Asp Ala Pro Arg Thr Arg Thr Gln Asp Ala Ile Asp Val Lys Leu

145 150 155 160

cgc gct taa 489
 Arg Ala

<210> 2
 <211> 162
 <212> PRT
 <213> Escherichia coli
 <400> 2

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp
 1 5 10 15

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser
 20 25 30

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg
 35 40 45

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly
 50 55 60

Tyr Leu Leu Gly Lys Asp Ala Ser Ser Ile Ala Val Gly Glu Val Ile
 65 70 75 80

Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Gly
 85 90 95

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp
 100 105 110

Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu
 115 120 125

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

130 135 140

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

145 150 155 160

Arg Ala

<210> 3

<211> 486

<212> DNA

<213> Shigella sonnei

<220><221> CDS

<222> (1)..(486)

<223> iscR gene encoding Iron Sulfur Cluster Regulator

<400> 3

atg cgc ctg acc agc aaa ggc cgc tat gcg gtg acc gcg atg ctg gat 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15

gtg gcg ctg aac agc gaa gcg ggc ccg gtg ccg ctg gcg gat att agc 96

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

gaa cgc cag ggc att agc ctg agc tat ctg gaa cag ctg ttt agc cgc 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

ctg cgc aaa aac ggc ctg gtg agc agc gtg cgc ggc ccg ggc ggc ggc 192

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

tat ctg ctg ggc aaa gat gcg agc agc att gcg gtg ggc gaa gtg att 240

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

65 70 75 80

agc gcg gtg gat gaa agc gtg gat gcg acc cgc tgc cag ggc aaa ggc 288

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

85 90 95

ggc tgc cag ggc ggc gat aaa tgc ctg acc cat gcg ctg tgg cgc gat 336
 Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp
 100 105 110
 ctg agc gat cgc ctg acc ggc ttt ctg aac aac att acc ctg ggc gaa 384
 Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu
 115 120 125
 ctg gtg aac aac cag gaa gtg ctg gat gtg agc ggc cgc cag cat acc 432
 Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Thr
 130 135 140
 cat gat gcg ccg cgc acc cgc att ctg gat gcg att gat gtg aaa ctg 480
 His Asp Ala Pro Arg Thr Arg Ile Leu Asp Ala Ile Asp Val Lys Leu
 145 150 155 160
 cgc gcg 486

Arg Ala

<210> 4

<211> 162

<212> PRT

<213> Shigella sonnei

<400> 4

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp
 1 5 10 15
 Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser
 20 25 30
 Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg
 35 40 45
 Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly
 50 55 60
 Tyr Leu Leu Gly Lys Asp Ala Ser Ser Ile Ala Val Gly Glu Val Ile
 65 70 75 80
 Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Gly
 85 90 95

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp

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

<210> 5

<211> 489

<212>

> DNA

<213> *Citrobacter pasteurii*

<220><221> CDS

<222> (1)..(489)

<223> iscR gene encoding Iron Sulfur Cluster Regulator

<400> 5

atg cgc ctg acc agc aaa ggc cgc tat gcg gtg acc gcg atg ctg gat 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15

gtg gcg ctg aac agc gaa acc ggc ccg gtg ccg ctg gcg gat att agc 96

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

20 25 30

gaa cgc cag ggc att agc ctg agc tat ctg gaa cag ctg ttt agc cgc 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

ctg cgc aaa aac ggc ctg gtg agc agc gtg cgc ggc ccg ggc ggc ggc 192

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

tat ctg ctg ggc aaa gat gcg agc agc att gcg gtg ggc gaa gtg att 240

Tyr Leu Leu Gly Lys Asp Ala Ser Ser Ile Ala Val Gly Glu Val Ile
65 70 75 80
agc gcg gtg gat gaa agc gtg gat gcg acc cgc tgc cag ggc aaa ggc 288
Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Gly
85 90 95
ggc tgc cag ggc ggc gat aaa tgc ctg acc cat gcg ctg tgg cgc gat 336

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp
100 105 110
ctg agc gat cgc ctg acc ggc ttt ctg aac aac att acc ctg ggc gaa 384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu
115 120 125
ctg gtg aac aac cag gaa gtg ctg gat gtg agc ggc cgc cag cat acc 432
Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Thr

130 135 140
cat gat gcg ccg cgc acc aac cgc gcg cag gat gcg att gat gtg aaa 480
His Asp Ala Pro Arg Thr Asn Arg Ala Gln Asp Ala Ile Asp Val Lys
145 150 155 160
ctg cgc gcg 489
Leu Arg Ala

<210> 6

<211> 163

<212> PRT

<213> Citrobacter pasteurii

<400> 6

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

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

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

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

<210> 7

<211> 489

<212> DNA

<213> *Enterobacter timonensis*

<220><221> CDS

<222> (1)..(489)

<223> iscR gene encoding Iron Sulfur Cluster Regulator

<400> 7

atg cgc ctg acc agc aaa ggc cgc tat gcg gtg acc gcg atg ctg gat 48
Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp
1 5 10 15
gtg cgc ctg aac agc gaa gcg ggc ccg gtg ccg ctg gcg gat att agc 96
Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser
20 25 30
gaa cgc cag ggc att agc ctg agc tat ctg gaa cag ctg ttt agc cgc 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg
35 40 45
ctg cgc aaa aac ggc ctg gtg agc agc gtg cgc ggc ccg ggc ggc ggc 192
Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly
50 55 60
tat ctg ctg ggc aaa gat gcg ggc agc att gcg gtg ggc gaa gtg att 240
Tyr Leu Leu Gly Lys Asp Ala Gly Ser Ile Ala Val Gly Glu Val Ile
65 70 75 80
agc gcg gtg gat gaa agc gtg gat gcg acc cgc tgc cag ggc aaa ggc 288
Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Gly
85 90 95
ggc tgc cag ggc ggc gat aaa tgc ctg acc cat gcg ctg tgg cgc gat 336
Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp
100 105 110
ctg agc gat cgc ctg acc ggc ttt ctg aac aac att acc ctg ggc gaa 384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu
115 120 125
ctg gtg aac aac cag gaa gtg ctg gat gtg agc ggc cgc cag cat agc 432
Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Ser
130 135 140
cat gat agc cag cgc aac acc cgc gcg cag gat gcg att gat gtg aaa 480
His Asp Ser Gln Arg Asn Thr Arg Ala Gln Asp Ala Ile Asp Val Lys
145 150 155 160
ctg cgc gcg 489
Leu Arg Ala

<210> 8

<211> 163

<212> PRT

<213> Enterobacter timonensis

<400> 8

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

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

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

130 135 140
His Asp Ser Gln Arg Asn Thr Arg Ala Gln Asp Ala Ile Asp Val Lys
145 150 155 160
Leu Arg Ala

<210> 9
<211> 489
<212> DNA
<213> *Pluralibacter gergoviae*
<220><221> CDS
<222> (1)..(489)
<223> iscR gene encoding Iron Sulfur Cluster Regulator
<400> 9

atg cgc ctg acc agc aaa ggc cgc tat gcg gtg acc gcg atg ctg gat

48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1	5	10	15	
gtg gcg ctg aac agc gaa agc ggc ccg gtg ccg ctg gcg gat att agc				96
Val Ala Leu Asn Ser Glu Ser Gly Pro Val Pro Leu Ala Asp Ile Ser				
	20	25	30	
gaa cgc cag ggc att agc ctg agc tat ctg gaa cag ctg ttt agc cgc				144
Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg				
	35	40	45	
ctg cgc aaa aac ggc ctg gtg agc agc gtg cgc ggc ccg ggc ggc ggc				
Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly				192
	50	55	60	
tat ctg ctg ggc aaa gat gcg ggc agc att gcg gtg ggc gaa gtg att				240
Tyr Leu Leu Gly Lys Asp Ala Gly Ser Ile Ala Val Gly Glu Val Ile				
65	70	75	80	
agc gcg gtg gat gaa agc gtg gat gcg acc cgc tgc cag ggc aaa gcg				288
Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Ala				
	85	90	95	
ggc tgc cag ggc ggc gat aaa tgc ctg acc cat gcg ctg tgg cgc gat				336
Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp				
	100	105	110	
ctg agc gat cgc ctg acc ggc ttt ctg aac aac att acc ctg ggc gaa				384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu				
	115	120	125	
ctg gtg aac aac cag gaa gtg ctg gat gtg agc gat cgc cag cat att				432
Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Asp Arg Gln His Ile				
	130	135	140	
cat gaa acc cag cgc agc acc cgc agc cag gat gcg att gat gtg aaa				480
His Glu Thr Gln Arg Ser Thr Arg Ser Gln Asp Ala Ile Asp Val Lys				
145	150	155	160	

ctg cgc gcg

489

Leu Arg Ala

<210> 10

<211> 163

<212> PRT

<213> *Pluralibacter gergoviae*

<400> 10

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15

Val Ala Leu Asn Ser Glu Ser Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

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

65 70 75 80

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

85 90 95

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp

100 105 110

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

115 120 125

Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Asp Arg Gln His Ile

130 135 140

His Glu Thr Gln Arg Ser Thr Arg Ser Gln Asp Ala Ile Asp Val Lys

145 150 155 160

Leu Arg Ala

<210> 11

<211> 486

<212> DNA

<213> Buttiauxella

<220><221> CDS

<222> (1)..(486)

<223> iscR gene encoding Iron Sulfur Cluster Regulator

<400> 11

atg cgc ctg acc agc aaa ggc cgc tat gcg gtg acc gcg atg ctg gat 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15

gtg gcg ctg aac agc gaa agc ggc ccg gtg ccg ctg gcg gat att agc 96

Val Ala Leu Asn Ser Glu Ser Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

gaa cgc cag ggc att agc ctg agc tat ctg gaa cag ctg ttt agc cgc 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

ctg cgc aaa aac ggc ctg gtg gcg agc gtg cgc ggc ccg ggc ggc ggc 192

Leu Arg Lys Asn Gly Leu Val Ala Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

tat ctg ctg ggc aaa gaa gcg agc gcg att gcg gtg ggc gaa gtg att 240

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

65 70 75 80

agc gcg gtg gat gaa agc gtg gat gcg acc cgc tgc gcg ggc aaa ggc 288

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

85 90 95

ggc tgc cag ggc ggc gat aaa tgc ctg acc cat gcg ctg tgg cgc gat 336

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp

100 105 110

ctg agc gat cgc ctg acc ggc ttt ctg aac aac att acc ctg ggc gaa 384

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

115 120 125

ctg gtg aac aac cag gaa gtg ctg gat gtg agc ggc cgc cag cat aac 432

Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Asn
 130 135 140
 gaa aac cat cgc agc acc cgc agc cag gat gcg att gat gtg aaa ctg 480
 Glu Asn His Arg Ser Thr Arg Ser Gln Asp Ala Ile Asp Val Lys Leu
 145 150 155 160
 cgc gcg 486
 Arg Ala

<210> 12
 <211> 162
 <212> PRT
 <213> Buttiiauxella
 <400> 12

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

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

Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu
 115 120 125
 Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Asn
 130 135 140
 Glu Asn His Arg Ser Thr Arg Ser Gln Asp Ala Ile Asp Val Lys Leu

145 150 155 160

Arg Ala

<210> 13

<211> 489

<212> DNA

<213> Kosakonia sacchari

<220><221> CDS

<222> (1)..(489)

<223> iscR gene encoding Iron Sulfur Cluster Regulator

<400> 13

atg cgc ctg acc agc aaa ggc cgc tat gcg gtg acc gcg atg ctg gat 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15

gtg gcg ctg aac agc gaa gcg ggc ccg gtg ccg ctg gcg gat att agc 96

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

gaa cgc cag ggc att agc ctg agc tat ctg gaa cag ctg ttt agc cgc 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

ctg cgc aaa aac ggc ctg gtg gcg agc gtg cgc ggc ccg ggc ggc ggc 192

Leu Arg Lys Asn Gly Leu Val Ala Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

tat ctg ctg ggc aaa gat gcg aac acc att gcg gtg ggc gaa gtg att 240

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

65 70 75 80

agc gcg gtg gat gaa agc gtg gat gcg acc cgc tgc cag ggc aaa agc 288

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

85 90 95

ggc tgc cag ggc ggc gat aaa tgc ctg acc cat gcg ctg tgg cgc gat 336

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp

100	105	110	
ctg agc gat cgc ctg acc ggc ttt ctg aac aac att acc ctg ggc gaa			384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu			
115	120	125	
ctg gtg aac aac cag gaa att ctg gat gtg agc gat cgc cag cat aac			432
Leu Val Asn Asn Gln Glu Ile Leu Asp Val Ser Asp Arg Gln His Asn			
130	135	140	
aac gaa agc cat cgc aac acc cgc ggc cag gat gcg att gat gtg aaa			480
Asn Glu Ser His Arg Asn Thr Arg Gly Gln Asp Ala Ile Asp Val Lys			
145	150	155	160
ctg cgc gcg			489
Leu Arg Ala			

<210> 14

<211> 163

<212> PRT

<213> Kosakonia sacchari

<400> 14

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

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

Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu
115 120 125

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

Asn Glu Ser His Arg Asn Thr Arg Gly Gln Asp Ala Ile Asp Val Lys
145 150 155 160

Leu Arg Ala

<210> 15

<211> 489

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(489)

<223> iscR mutant gene encoding Iron Sulfur Cluster Regulator protein
(IscR) with L15F substitution

<400>

15

atg aga ctg aca tct aaa ggg cgc tat gcc gtg acc gca atg ttt gac 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Phe Asp

1 5 10 15

gtt gcg ctc aac tct gaa gcg ggc ccg gta ccg ttg gct gat att tcc 96

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

gaa cgt cag gga att tcc ctt tct tat ctg gaa caa ctg ttt tcc cgt 144

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

ctg cgt aaa aat ggt ctg gtt tcc agc gta cgt gga cca ggc ggt ggt 192

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

tat ctg tta ggc aaa gat gcc agc agc atc gcc gtt ggc gaa gta att 240

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

65	70	75	80	
agc gcc gtt gac gaa tct gta gat gcc acc cgt tgt cag ggt aaa ggc				288
Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Gly				
	85	90	95	
ggc tgc cag ggc ggc gat aaa tgc ctg acc cac gcg ctg tgg cgt gat				336
Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp				
	100	105	110	
ttg agc gac cgt ctc acc ggt ttt ctc aac aac att act tta ggc gaa				384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu				
	115	120	125	
ctg gtt aat aac cag gaa gtg ctg gat gtg tct ggt cgt cag cat act				432
Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Thr				
	130	135	140	
cac gac gcg cca cgc acc cgc aca caa gac gcg atc gac gtt aag tta				480
His Asp Ala Pro Arg Thr Arg Thr Gln Asp Ala Ile Asp Val Lys Leu				
145	150	155	160	
cgc gct taa				489
Arg Ala				
<210> 16				
<211> 162				
<212> PRT				
<213> Escherichia coli				
<400> 16				
Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Phe Asp				
1	5	10	15	
Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser				
	20	25	30	
Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg				
	35	40	45	
Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly				
50	55	60		

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

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

130 135 140
 His Asp Ala Pro Arg Thr Arg Thr Gln Asp Ala Ile Asp Val Lys Leu
 145 150 155 160
 Arg Ala

<210> 17
 <211> 489
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(489)
 <223> iscR mutant gene encoding Iron Sulfur Cluster Regulator protein
 (IscR) with C92Y substitution
 <400> 17

atg aga ctg aca tct aaa ggg cgc tat gcc gtg acc gca atg ctt gac 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp
 1 5 10 15
 gtt gcg ctc aac tct gaa gcg ggc ccg gta ccg ttg gct gat att tcc 96
 Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser
 20 25 30
 gaa cgt cag gga att tcc ctt tct tat ctg gaa caa ctg ttt tcc cgt 144
 Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35	40	45	
ctg cgt aaa aat ggt ctg gtt tcc agc gta cgt gga cca ggc ggt ggt			192
Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly			
50	55	60	
tat ctg tta ggc aaa gat gcc agc agc atc gcc gtt ggc gaa gta att			240
Tyr Leu Leu Gly Lys Asp Ala Ser Ser Ile Ala Val Gly Glu Val Ile			
65	70	75	80
agc gcc gtt gac gaa tct gta gat gcc acc cgt tat cag ggt aaa ggc			288
Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Tyr Gln Gly Lys Gly			
85	90	95	
ggc tgc cag ggc ggc gat aaa tgc ctg acc cac gcg ctg tgg cgt gat			336
Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp			
100	105	110	
ttg agc gac cgt ctc acc ggt ttt ctc aac aac att act tta ggc gaa			384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu			
115	120	125	
ctg gtt aat aac cag gaa gtg ctg gat gtg tct ggt cgt cag cat act			432
Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Thr			
130	135	140	
cac gac gcg cca cgc acc cgc aca caa gac gcg atc gac gtt aag tta			480
His Asp Ala Pro Arg Thr Arg Thr Gln Asp Ala Ile Asp Val Lys Leu			
145	150	155	160
cgc gct taa			489
Arg Ala			
<210>	18		
<211>	162		
<212>	PRT		
<213>	Escherichia coli		
<400>	18		
Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp			

1 5 10 15

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

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

65 70 75 80

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

85 90 95

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr His Ala Leu Trp Arg Asp

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

Arg Ala

<210> 19

<211> 489

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(489)

<223> iscR mutant gene encoding Iron Sulfur Cluster Regulator protein
(IscR) with H107Y substitution

<400> 19

atg aga ctg aca tct aaa ggg cgc tat gcc gtg acc gca atg ctt gac 48

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1	5	10	15	
gtt gcg ctc aac tct gaa gcg ggc ccg gta ccg ttg gct gat att tcc				96
Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser				
	20	25	30	
gaa cgt cag gga att tcc ctt tct tat ctg gaa caa ctg ttt tcc cgt				144
Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg				
	35	40	45	
ctg cgt aaa aat ggt ctg gtt tcc agc gta cgt gga cca ggc ggt ggt				192
Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly				
	50	55	60	
tat ctg tta ggc aaa gat gcc agc agc atc gcc gtt ggc gaa gta att				240
Tyr Leu Leu Gly Lys Asp Ala Ser Ser Ile Ala Val Gly Glu Val Ile				
65	70	75	80	
agc gcc gtt gac gaa tct gta gat gcc acc cgt tgt cag ggt aaa ggc				288
Ser Ala Val Asp Glu Ser Val Asp Ala Thr Arg Cys Gln Gly Lys Gly				
	85	90	95	
ggc tgc cag ggc ggc gat aaa tgc ctg acc tac gcg ctg tgg cgt gat				336
Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr Tyr Ala Leu Trp Arg Asp				
	100	105	110	
ttg agc gac cgt ctc acc ggt ttt ctc aac aac att act tta ggc gaa				384
Leu Ser Asp Arg Leu Thr Gly Phe Leu Asn Asn Ile Thr Leu Gly Glu				
	115	120	125	
ctg gtt aat aac cag gaa gtg ctg gat gtg tct ggt cgt cag cat act				432
Leu Val Asn Asn Gln Glu Val Leu Asp Val Ser Gly Arg Gln His Thr				
	130	135	140	
cac gac gcg cca cgc acc cgc aca caa gac gcg atc gac gtt aag tta				480
His Asp Ala Pro Arg Thr Arg Thr Gln Asp Ala Ile Asp Val Lys Leu				
145	150	155	160	

cgc gct taa

489

Arg Ala

<210> 20

<211> 162

<212> PRT

<213> Escherichia coli

<400> 20

Met Arg Leu Thr Ser Lys Gly Arg Tyr Ala Val Thr Ala Met Leu Asp

1 5 10 15

Val Ala Leu Asn Ser Glu Ala Gly Pro Val Pro Leu Ala Asp Ile Ser

20 25 30

Glu Arg Gln Gly Ile Ser Leu Ser Tyr Leu Glu Gln Leu Phe Ser Arg

35 40 45

Leu Arg Lys Asn Gly Leu Val Ser Ser Val Arg Gly Pro Gly Gly Gly

50 55 60

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

65 70 75 80

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

85 90 95

Gly Cys Gln Gly Gly Asp Lys Cys Leu Thr Tyr Ala Leu Trp Arg Asp

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

Arg Ala

<210> 21

<211> 1041

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(1041)

<223> bioB gene encoding biotin synthase (EC 2.8.1.6)

<400> 21

atg gct cac cgc cca cgc tgg aca ttg tgc caa gtc aca gaa tta ttt 48

Met Ala His Arg Pro Arg Trp Thr Leu Ser Gln Val Thr Glu Leu Phe

1 5 10 15

gaa aaa ccg ttg ctg gat ctg ctg ttt gaa gcg cag cag gtg cat cgc 96

Glu Lys Pro Leu Leu Asp Leu Leu Phe Glu Ala Gln Gln Val His Arg

20 25 30

cag cat ttc gat cct cgt cag gtg cag gtc agc acg ttg ctg tgc att 144

Gln His Phe Asp Pro Arg Gln Val Gln Val Ser Thr Leu Leu Ser Ile

35 40 45

aag acc gga gct tgt ccg gaa gat tgc aaa tac tgc ccg caa agc tgc 192

Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys Tyr Cys Pro Gln Ser Ser

50 55 60

cgc tac aaa acc ggg ctg gaa gcc gag cgg ttg atg gaa gtt gaa cag 240

Arg Tyr Lys Thr Gly Leu Glu Ala Glu Arg Leu Met Glu Val Glu Gln

65 70 75 80

gtg ctg gag tgc gcg cgc aaa gcg aaa gcg gca gga tgc acg cgc ttc 288

Val Leu Glu Ser Ala Arg Lys Ala Lys Ala Ala Gly Ser Thr Arg Phe

85 90 95

tgt atg ggc gcg gcg tgg aag aat ccc cac gaa cgc gat atg ccg tac 336

Cys Met Gly Ala Ala Trp Lys Asn Pro His Glu Arg Asp Met Pro Tyr

100 105 110

ctg gaa caa atg gtg cag ggg gta aaa gcg atg ggg ctg gag gcg tgt 384

Leu Glu Gln Met Val Gln Gly Val Lys Ala Met Gly Leu Glu Ala Cys

115 120 125

atg acg ctg ggc acg ttg agt gaa tct cag gcg cag cgc ctc gcg aac 432

Met Thr Leu Gly Thr Leu Ser Glu Ser Gln Ala Gln Arg Leu Ala Asn

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

aaa ctg ctg acc acg ccg aat ccg gaa gaa gat aaa gac ctg caa ctg 912
 Lys Leu Leu Thr Thr Pro Asn Pro Glu Glu Asp Lys Asp Leu Gln Leu
 290 295 300
 ttc cgc aaa ctg ggg cta aat ccg cag caa act gcc gtg ctg gca ggg 960
 Phe Arg Lys Leu Gly Leu Asn Pro Gln Gln Thr Ala Val Leu Ala Gly
 305 310 315 320
 gat aac gaa caa cag caa cgt ctt gaa cag gcg ctg atg acc ccg gac 1008

Asp Asn Glu Gln Gln Gln Arg Leu Glu Gln Ala Leu Met Thr Pro Asp
 325 330 335
 acc gac gaa tat tac aac gcg gca gca tta tga 1041
 Thr Asp Glu Tyr Tyr Asn Ala Ala Ala Leu
 340 345

<210> 22

<211> 346

<212> PRT

<213> Escherichia coli

<400> 22

Met Ala His Arg Pro Arg Trp Thr Leu Ser Gln Val Thr Glu Leu Phe

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

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

115 120 125
Met Thr Leu Gly Thr Leu Ser Glu Ser Gln Ala Gln Arg Leu Ala Asn

130 135 140
Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu Phe
145 150 155 160

Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr Gln Glu Arg Leu Asp Thr
165 170 175

Leu Glu Lys Val Arg Asp Ala Gly Ile Lys Val Cys Ser Gly Gly Ile
180 185 190

Val Gly Leu Gly Glu Thr Val Lys Asp Arg Ala Gly Leu Leu Leu Gln

195 200 205
Leu Ala Asn Leu Pro Thr Pro Pro Glu Ser Val Pro Ile Asn Met Leu
210 215 220

Val Lys Val Lys Gly Thr Pro Leu Ala Asp Asn Asp Asp Val Asp Ala
225 230 235 240

Phe Asp Phe Ile Arg Thr Ile Ala Val Ala Arg Ile Met Met Pro Thr
245 250 255

Ser Tyr Val Arg Leu Ser Ala Gly Arg Glu Gln Met Asn Glu Gln Thr

260 265 270
Gln Ala Met Cys Phe Met Ala Gly Ala Asn Ser Ile Phe Tyr Gly Cys
275 280 285

Lys Leu Leu Thr Thr Pro Asn Pro Glu Glu Asp Lys Asp Leu Gln Leu
290 295 300

Phe Arg Lys Leu Gly Leu Asn Pro Gln Gln Thr Ala Val Leu Ala Gly
305 310 315 320

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

325 330 335
Thr Asp Glu Tyr Tyr Asn Ala Ala Ala Leu

340 345

<210> 23

<211> 1040

<212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(222)
 <223> Mutant bioB gene with frameshift encoding inactive truncated
 biotin synthase (EC 2.8.1.6)
 <400> 23
 atg gcc acc gcc cac gct gga cat tgt cgc aag tca cag aat tat ttg 48
 Met Ala Thr Ala His Ala Gly His Cys Arg Lys Ser Gln Asn Tyr Leu
 1 5 10 15
 aaa aac cgt tgc tgg atc tgc tgt ttg aag cgc agc agg tgc atc gcc 96
 Lys Asn Arg Cys Trp Ile Cys Cys Leu Lys Arg Ser Arg Cys Ile Ala
 20 25 30
 agc att tcg atc ctc gtc agg tgc agg tca gca cgt tgc tgt cga tta 144
 Ser Ile Ser Ile Leu Val Arg Cys Arg Ser Ala Arg Cys Cys Arg Leu
 35 40 45
 aga ccg gag ctt gtc cgg aag att gca aat act gcc cgc aaa gct cgc 192
 Arg Pro Glu Leu Val Arg Lys Ile Ala Asn Thr Ala Arg Lys Ala Arg
 50 55 60
 gct aca aaa ccg ggc tgg aag ccg agc ggt tgatggaagt tgaacaggtg 242
 Ala Thr Lys Pro Gly Trp Lys Pro Ser Gly
 65 70
 ctggagtcgg cgcgcaaagc gaaagcggca ggatcgacgc gcttctgtat gggcgcggcg 302
 tggaagaatc cccacgaacg cgatatgccg tacctggaac aaatggtgca gggggtaaaa 362
 gcgatggggc tggaggcgtg tatgacgtg ggcacgttga gtgaatctca ggcgcagcgc 422
 ctgcgaacg ccgggctgga ttactacaac cacaacctgg acacctcgcc ggagttttac 482
 ggcaatatca tcaccacacg cacttatcag gaacgcctcg atacgtgga aaaagtgcgc 542
 gatgccggga tcaaagtctg ttctggcggc attgtgggct taggcgaaac ggtaaaagat 602
 cgcgccggat taitgtgca actggcaaac ctgccgacgc cgccggaaag cgtgccaatc 662
 aacatgctgg tgaaggtgaa aggcacgccg cttgccgata acgatgatgt cgatgccttt 722
 gattttattc gcaccattgc ggtcgcgcgg atcatgatgc caacctctta cgtgcgcctt 782

tctgccggac gcgagcagat gaacgaacag actcaggcga tgtgctttat ggcaggcgca 842
aactcgattt tctacgggtg caaactgctg accacgccga atccggaaga agataaagac 902
ctgcaactgt tccgcaaact ggggctaaat cgcagcaaaa ctgccgtgct ggcaggggat 962
aacgaacaac agcaacgtct tgaacaggcg ctgatgaccc cggacaccga cgaatattac 1022
aacgcggcag cattatga 1040

<210> 24

<211> 74

<212> PRT

<213> Escherichia coli

<400> 24

Met Ala Thr Ala His Ala Gly His Cys Arg Lys Ser Gln Asn Tyr Leu

1 5 10 15

Lys Asn Arg Cys Trp Ile Cys Cys Leu Lys Arg Ser Arg Cys Ile Ala

20 25 30

Ser Ile Ser Ile Leu Val Arg Cys Arg Ser Ala Arg Cys Cys Arg Leu

35 40 45

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

50 55 60

Ala Thr Lys Pro Gly Trp Lys Pro Ser Gly

65 70

<210> 25

<211> 97

<212> DNA

<213> Escherichia coli

<220><221> promoter

<222> (1)..(97)

<223> T5 lacO repressed promoter

<400> 25

tcataaaaaa tttatttgct ttgtgagcgg ataacaatta taatagattc aaatcggagg 60

ttctctaact agtatctcta gagctaagga ggtaaat 97

<210> 26

<211> 1041

<212> DNA

<213> Candidatus Chloracidobacterium thermophilum B

<220><221> CDS

<222> (1)..(1041)

<223> bioB gene encoding biotin synthase from Candidatus

Chloracidobacterium thermophilum B

<400> 26

atg agc cag ccg ctg gtt cgt ttt gat tgg acc cgt gat gaa ctg cgt 48

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

1 5 10 15

gca ctg cat gat ctg ccg ctg ctg gaa ctg att cat cgt gca gca acc 96

Ala Leu His Asp Leu Pro Leu Leu Glu Leu Ile His Arg Ala Ala Thr

20 25 30

gtt cat cgt acc tgt cat gat ccg cag gaa gtt cag gtt tgt cgt ctg 144

Val His Arg Thr Cys His Asp Pro Gln Glu Val Gln Val Cys Arg Leu

35 40 45

att agc att aaa acc ggc ggt tgt ccg gaa gat tgt ggt tat tgt agc 192

Ile Ser Ile Lys Thr Gly Gly Cys Pro Glu Asp Cys Gly Tyr Cys Ser

50 55 60

cag agc gca cat tat gaa acc ggt att gca gca cag ccg ctg ctg gat 240

Gln Ser Ala His Tyr Glu Thr Gly Ile Ala Ala Gln Pro Leu Leu Asp

65 70 75 80

aaa gcc acc gtt gtt gca att gca gaa cgt gca aaa gca cat ggt gtt 288

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

85 90 95

agc cgt gtt tgt ctg ggt gca gca tgg cgt aat gtt cgt gat gat gca 336

Ser Arg Val Cys Leu Gly Ala Ala Trp Arg Asn Val Arg Asp Asp Ala

100 105 110

cag ttt gaa gca gtt ctg gat att gtt cgt agc gtt aat gca ctg ggt 384

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

115 120 125

att gaa gtt tgt tgt acc ctg ggt atg ctg acc gaa gcc cag gca cgt 432

Ile Glu Val Cys Cys Thr Leu Gly Met Leu Thr Glu Ala Gln Ala Arg	
130	135
140	
cgt ctg gaa gaa gca ggt ctg tat gca tat aat cat aat ctg gat acc	480
Arg Leu Glu Glu Ala Gly Leu Tyr Ala Tyr Asn His Asn Leu Asp Thr	
145	150
155	160
agc cgt gaa tat tat ggt cgt gtt gtt acc acc cgt acc tat gat gat	528
Ser Arg Glu Tyr Tyr Gly Arg Val Val Thr Thr Arg Thr Tyr Asp Asp	
165	170
175	
cgc ctg gaa acc ctg gca aat gtt cgc aaa acc ggt gtt acc ctg tgt	576
Arg Leu Glu Thr Leu Ala Asn Val Arg Lys Thr Gly Val Thr Leu Cys	
180	185
190	
acc ggt ggt att ctg ggt ctg ggt gaa agc acc gat gat cgt att ggt	624
Thr Gly Gly Ile Leu Gly Leu Gly Glu Ser Thr Asp Asp Arg Ile Gly	
195	200
205	
ctg ctg cat acc ctg gcc acc atg aat ccg cat ccg gaa agc gtt ccg	672
Leu Leu His Thr Leu Ala Thr Met Asn Pro His Pro Glu Ser Val Pro	
210	215
220	
att aat ctg ctg acc cgt gtt ccg ggt acc ccg atg gaa aat gaa gca	720
Ile Asn Leu Leu Thr Arg Val Pro Gly Thr Pro Met Glu Asn Glu Ala	
225	230
235	240
gaa gtt agc gtt tgg gaa acc ctg cgt gtt att gca acc gca cgt att	768
Glu Val Ser Val Trp Glu Thr Leu Arg Val Ile Ala Thr Ala Arg Ile	
245	250
255	
gca atg ccg cgt agt gtt att cgt ctg agc gca ggt cgt acc cag ctg	816
Ala Met Pro Arg Ser Val Ile Arg Leu Ser Ala Gly Arg Thr Gln Leu	
260	265
270	
agc gaa gaa gca cag gca ctg tgt ttt ctg gcc ggt gca aat agc att	864
Ser Glu Glu Ala Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile	
275	280
285	

ttt agc agc gat gca cgt atg atg ctg acc cgc gtt agc ccg acc aat 912

Phe Ser Ser Asp Ala Arg Met Met Leu Thr Arg Val Ser Pro Thr Asn

290 295 300

gat tat gat gaa gat gcc cag ctg ctg aat aaa ctg ggt ctg cat ccg 960

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

305 310 315 320

cgt gtt ccg ttt aaa gat gca ccg aat gca aaa acc gca ggt tgt gca 1008

Arg Val Pro Phe Lys Asp Ala Pro Asn Ala Lys Thr Ala Gly Cys Ala

325 330 335

agc gca gcc acc gca acc ctg cag gaa aaa taa 1041

Ser Ala Ala Thr Ala Thr Leu Gln Glu Lys

340 345

<210> 27

<211> 346

<212> PRT

<213> Candidatus Chloracidobacterium thermophilum B

<400> 27

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

1 5 10 15

Ala Leu His Asp Leu Pro Leu Leu Glu Leu Ile His Arg Ala Ala Thr

20 25 30

Val His Arg Thr Cys His Asp Pro Gln Glu Val Gln Val Cys Arg Leu

35 40 45

Ile Ser Ile Lys Thr Gly Gly Cys Pro Glu Asp Cys Gly Tyr Cys Ser

50 55 60

Gln Ser Ala His Tyr Glu Thr Gly Ile Ala Ala Gln Pro Leu Leu Asp

65 70 75 80

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

85 90 95

Ser Arg Val Cys Leu Gly Ala Ala Trp Arg Asn Val Arg Asp Asp Ala

100 105 110

Gln Phe Glu Ala Val Leu Asp Ile Val Arg Ser Val Asn Ala Leu Gly
115 120 125

Ile Glu Val Cys Cys Thr Leu Gly Met Leu Thr Glu Ala Gln Ala Arg
130 135 140

Arg Leu Glu Glu Ala Gly Leu Tyr Ala Tyr Asn His Asn Leu Asp Thr
145 150 155 160

Ser Arg Glu Tyr Tyr Gly Arg Val Val Thr Thr Arg Thr Tyr Asp Asp
165 170 175

Arg Leu Glu Thr Leu Ala Asn Val Arg Lys Thr Gly Val Thr Leu Cys
180 185 190

Thr Gly Gly Ile Leu Gly Leu Gly Glu Ser Thr Asp Asp Arg Ile Gly
195 200 205

Leu Leu His Thr Leu Ala Thr Met Asn Pro His Pro Glu Ser Val Pro
210 215 220

Ile Asn Leu Leu Thr Arg Val Pro Gly Thr Pro Met Glu Asn Glu Ala
225 230 235 240

Glu Val Ser Val Trp Glu Thr Leu Arg Val Ile Ala Thr Ala Arg Ile
245 250 255

Ala Met Pro Arg Ser Val Ile Arg Leu Ser Ala Gly Arg Thr Gln Leu
260 265 270

Ser Glu Glu Ala Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile
275 280 285

Phe Ser Ser Asp Ala Arg Met Met Leu Thr Arg Val Ser Pro Thr Asn
290 295 300

Asp Tyr Asp Glu Asp Ala Gln Leu Leu Asn Lys Leu Gly Leu His Pro
305 310 315 320

Arg Val Pro Phe Lys Asp Ala Pro Asn Ala Lys Thr Ala Gly Cys Ala
325 330 335

Ser Ala Ala Thr Ala Thr Leu Gln Glu Lys
340 345

<210> 28

<211> 1164

<212> DNA

<213> Streptomyces lydicus

<220><221> CDS

<222> (1)..(1164)

<223> bioB gene encoding biotin synthase from Streptomyces lydicus A02

<400> 28

atg ccg tat gtt cgt att aat gca atg gat ctg ctg aat acc ctg gtt 48

Met Pro Tyr Val Arg Ile Asn Ala Met Asp Leu Leu Asn Thr Leu Val

1 5 10 15

gat aaa ggt ctg cgt cgt gaa ctg ccg acc cgt gaa gaa gca ctg gca 96

Asp Lys Gly Leu Arg Arg Glu Leu Pro Thr Arg Glu Glu Ala Leu Ala

20 25 30

gtt ctg gca acc agc gat gat gaa ctg ctg gat gtt gtt gca gca gcg 144

Val Leu Ala Thr Ser Asp Asp Glu Leu Leu Asp Val Val Ala Ala Ala

35 40 45

ggg aaa gtt cgc cgt cag tgg ttt ggt cgt cgt gtt aaa ctg aat tat 192

Gly Lys Val Arg Arg Gln Trp Phe Gly Arg Arg Val Lys Leu Asn Tyr

50 55 60

ctg gtg aat ctg aaa agc ggt ctg tgt ccg gaa gat tgt agc tat tgt 240

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

65 70 75 80

agc cag cgt ctg ggt agc aaa gca gaa att ctg aaa tat acc tgg ctg 288

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

85 90 95

aaa ccg gat gat gca agt aaa gca gcc gca gca ggc gtt gcc ggt ggt 336

Lys Pro Asp Asp Ala Ser Lys Ala Ala Ala Ala Gly Val Ala Gly Gly

100 105 110

gca aaa cgt gtt tgt ctg gtt gca agc ggt cgt ggt ccg acc gat aaa 384

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

115 120 125

gat gtt gat cgt gtg agc gaa acc att agc gca att aaa gaa cag aat	432
Asp Val Asp Arg Val Ser Glu Thr Ile Ser Ala Ile Lys Glu Gln Asn	
130 135 140	
gaa ggt gtg gaa gtt tgt gca tgt ctg ggt ctg ctg agc gat ggt cag	480
Glu Gly Val Glu Val Cys Ala Cys Leu Gly Leu Leu Ser Asp Gly Gln	
145 150 155 160	
gca gat cgt ctg cgt gca gcc ggt gca gat gca tat aat cat aat ctg	528
Ala Asp Arg Leu Arg Ala Ala Gly Ala Asp Ala Tyr Asn His Asn Leu	
165 170 175	
aat acc agt gaa gca acc tat ggt gat att tgt acc acc cat gat ttt	576
Asn Thr Ser Glu Ala Thr Tyr Gly Asp Ile Cys Thr Thr His Asp Phe	
180 185 190	
agc gat cgt gtt agc acc gtt cag cag gca cag gca gca ggt atg agc	624
Ser Asp Arg Val Ser Thr Val Gln Gln Ala Gln Ala Ala Gly Met Ser	
195 200 205	
gca tgt agc ggt ctg att gca ggt atg ggt gaa agc gat gca gat ctg	672
Ala Cys Ser Gly Leu Ile Ala Gly Met Gly Glu Ser Asp Ala Asp Leu	
210 215 220	
gtg gat gtg gtt ttt gca ctg cgt gaa ctg gat ccg gat agc gtt ccg	720
Val Asp Val Val Phe Ala Leu Arg Glu Leu Asp Pro Asp Ser Val Pro	
225 230 235 240	
gtt aat ttt ctg att ccg ttt gaa ggt acc ccg ctg gca aaa gaa tgg	768
Val Asn Phe Leu Ile Pro Phe Glu Gly Thr Pro Leu Ala Lys Glu Trp	
245 250 255	
aat ctg acc ccg cag cgt gcc ctg cgt att ctg gca atg gtt cgt ttt	816
Asn Leu Thr Pro Gln Arg Ala Leu Arg Ile Leu Ala Met Val Arg Phe	
260 265 270	
gtt tgt ccg gat gtt gaa gtt cgt ctg gca ggc ggt cgt gaa gtt cat	864
Val Cys Pro Asp Val Glu Val Arg Leu Ala Gly Gly Arg Glu Val His	

275	280	285	
ctg cgt agc ctg cag ccg ctg gca ctg cat ctg gtt aat agc att ttt			912
Leu Arg Ser Leu Gln Pro Leu Ala Leu His Leu Val Asn Ser Ile Phe			
290	295	300	
ctg ggt gat tat ctg acc agc gaa ggt cag gcc ggt aaa gaa gat ctg			960
Leu Gly Asp Tyr Leu Thr Ser Glu Gly Gln Ala Gly Lys Glu Asp Leu			
305	310	315	320
gcc atg att gcc gat gca ggt ttt gaa gtg gaa ggt aca gat acc acc			1008
Ala Met Ile Ala Asp Ala Gly Phe Glu Val Glu Gly Thr Asp Thr Thr			
325	330	335	
acc ctg ccg gaa cat cgt aca gat gct gca gtt cag ccg gca ccg gaa			1056
Thr Leu Pro Glu His Arg Thr Asp Ala Ala Val Gln Pro Ala Pro Glu			
340	345	350	
ccg gca gca gat gca gca gtt cct gca cct cct agt gaa gaa aca cgt			1104
Pro Ala Ala Asp Ala Ala Val Pro Ala Pro Pro Ser Glu Glu Thr Arg			
355	360	365	
cgt gat ctg gtt agc gtt cgt cgt cgt ggt gca ggt acc gaa ctg ccg			1152
Arg Asp Leu Val Ser Val Arg Arg Arg Gly Ala Gly Thr Glu Leu Pro			
370	375	380	
ccg aat gca taa			1164
Pro Asn Ala			

385

<210> 29

<211> 387

<212> PRT

<213> Streptomyces lydicus

<400> 29

Met Pro Tyr Val Arg Ile Asn Ala Met Asp Leu Leu Asn Thr Leu Val

1 5 10 15

Asp Lys Gly Leu Arg Arg Glu Leu Pro Thr Arg Glu Glu Ala Leu Ala

20 25 30

Val Leu Ala Thr Ser Asp Asp Glu Leu Leu Asp Val Val Ala Ala Ala

35 40 45
 Gly Lys Val Arg Arg Gln Trp Phe Gly Arg Arg Val Lys Leu Asn Tyr
 50 55 60
 Leu Val Asn Leu Lys Ser Gly Leu Cys Pro Glu Asp Cys Ser Tyr Cys
 65 70 75 80
 Ser Gln Arg Leu Gly Ser Lys Ala Glu Ile Leu Lys Tyr Thr Trp Leu
 85 90 95
 Lys Pro Asp Asp Ala Ser Lys Ala Ala Ala Ala Gly Val Ala Gly Gly
 100 105 110

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

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

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

Leu Arg Ser Leu Gln Pro Leu Ala Leu His Leu Val Asn Ser Ile Phe
290 295 300

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

Arg Asp Leu Val Ser Val Arg Arg Arg Gly Ala Gly Thr Glu Leu Pro
370 375 380

Pro Asn Ala

385

<210> 30

<211> 975

<212> DNA

<213> Paracoccus denitrificans

<220><221> CDS

<222> (1)..(975)

<223> bioB gene encoding biotin synthase from Paracoccus denitrificans

PD1222

<400> 30

atg att cgt acc gat tgg aca atg gca gaa gca tgg gca att cat gca 48

Met Ile Arg Thr Asp Trp Thr Met Ala Glu Ala Trp Ala Ile His Ala

1 5 10 15

ctg ccg ttt gca gat ctg atg cat cgc gca cag acc ctg cat cgt gca 96

Leu Pro Phe Ala Asp Leu Met His Arg Ala Gln Thr Leu His Arg Ala

20 25 30

cat ttt gat ccg aat gca att gaa acc gca agc ctg ctg agc att aaa 144

His Phe Asp Pro Asn Ala Ile Glu Thr Ala Ser Leu Leu Ser Ile Lys

35	40	45	
acc ggt ggt tgt ccg gaa gat tgt ggt tat tgt agc cag agc gca cat			192
Thr Gly Gly Cys Pro Glu Asp Cys Gly Tyr Cys Ser Gln Ser Ala His			
50	55	60	
cat gat acc ggt gtt aaa gca acc aaa ctg atg ggt acc gaa gaa gtt			240
His Asp Thr Gly Val Lys Ala Thr Lys Leu Met Gly Thr Glu Glu Val			
65	70	75	80
ctg gca gca gca cgt cgt gca aaa gca agc ggt gca cag cgt ttt tgt			288
Leu Ala Ala Ala Arg Arg Ala Lys Ala Ser Gly Ala Gln Arg Phe Cys			
85	90	95	
atg ggt gca gca tgg cgt agc ccg aaa gat cgt gat atg gat aaa ctg			336
Met Gly Ala Ala Trp Arg Ser Pro Lys Asp Arg Asp Met Asp Lys Leu			
100	105	110	
tgt gat atg gtt cgt ggt gtt gca gaa ctg ggt ctg gaa acc tgt atg			384
Cys Asp Met Val Arg Gly Val Ala Glu Leu Gly Leu Glu Thr Cys Met			
115	120	125	
acc ctg ggt atg ctg agc ccg gaa cag gtt gca cgt ctg aaa gca gca			432
Thr Leu Gly Met Leu Ser Pro Glu Gln Val Ala Arg Leu Lys Ala Ala			
130	135	140	
ggt ctg gat ttt tat aat cat aat att gat acc agc ccg gaa tat tat			480
Gly Leu Asp Phe Tyr Asn His Asn Ile Asp Thr Ser Pro Glu Tyr Tyr			
145	150	155	160
gcc cag att gcc agc acc cgt aca atg gaa aat cgt ctg gat acc gtt			528
Ala Gln Ile Ala Ser Thr Arg Thr Met Glu Asn Arg Leu Asp Thr Val			
165	170	175	
gaa cag gtt cgt aaa ggt ggt att aaa gtt tgt tgt ggt ggt att ctg			576
Glu Gln Val Arg Lys Gly Gly Ile Lys Val Cys Cys Gly Gly Ile Leu			
180	185	190	
ggt atg ggt gaa gca gaa gaa gat cgt att gca atg ctg gtt acc ctg			624

Gly Met Gly Glu Ala Glu Glu Asp Arg Ile Ala Met Leu Val Thr Leu
195 200 205
gca acc ctg ccg gca cat ccg gat agc gtt ccg gtt aat ctg tgg aat 672
Ala Thr Leu Pro Ala His Pro Asp Ser Val Pro Val Asn Leu Trp Asn
210 215 220
gaa att gaa ggt gtt ccg gtt cag gca cgt gca cag gca gtt gat ccg 720
Glu Ile Glu Gly Val Pro Val Gln Ala Arg Ala Gln Ala Val Asp Pro
225 230 235 240
ttt gcc ctg gtt cgt att gtt gca ctg gca cgt att ctg atg ccg gca 768
Phe Ala Leu Val Arg Ile Val Ala Leu Ala Arg Ile Leu Met Pro Ala
245 250 255
agc gtt gtt cgt ctg agc gca ggt cgt acc ggt atg agc gat gaa ctg 816
Ser Val Val Arg Leu Ser Ala Gly Arg Thr Gly Met Ser Asp Glu Leu
260 265 270
cag gca ctg tgt ttt ctg gcg ggt gca aat agc att ttt gtt ggt gat 864
Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Val Gly Asp
275 280 285
cag ctg ctg acc acc ggt aat ccg gca gca tgg aaa gat cag gat ctg 912
Gln Leu Leu Thr Thr Gly Asn Pro Ala Ala Trp Lys Asp Gln Asp Leu
290 295 300
ctg agt cgt ctg ggt atg cat att gca ccg gca cag gcc cgt ccg cgt 960
Leu Ser Arg Leu Gly Met His Ile Ala Pro Ala Gln Ala Arg Pro Arg
305 310 315 320
gtt gca gca gaa taa 975
Val Ala Ala Glu

<210> 31

<211> 324

<212> PRT

<213> Paracoccus denitrificans

<400> 31

Met Ile Arg Thr Asp Trp Thr Met Ala Glu Ala Trp Ala Ile His Ala

1 5 10 15

Leu Pro Phe Ala Asp Leu Met His Arg Ala Gln Thr Leu His Arg Ala

20 25 30

His Phe Asp Pro Asn Ala Ile Glu Thr Ala Ser Leu Leu Ser Ile Lys

35 40 45

Thr Gly Gly Cys Pro Glu Asp Cys Gly Tyr Cys Ser Gln Ser Ala His

50 55 60

His Asp Thr Gly Val Lys Ala Thr Lys Leu Met Gly Thr Glu Glu Val

65 70 75 80

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

85 90 95

Met Gly Ala Ala Trp Arg Ser Pro Lys Asp Arg Asp Met Asp Lys Leu

100 105 110

Cys Asp Met Val Arg Gly Val Ala Glu Leu Gly Leu Glu Thr Cys Met

115 120 125

Thr Leu Gly Met Leu Ser Pro Glu Gln Val Ala Arg Leu Lys Ala Ala

130 135 140

Gly Leu Asp Phe Tyr Asn His Asn Ile Asp Thr Ser Pro Glu Tyr Tyr

145 150 155 160

Ala Gln Ile Ala Ser Thr Arg Thr Met Glu Asn Arg Leu Asp Thr Val

165 170 175

Glu Gln Val Arg Lys Gly Gly Ile Lys Val Cys Cys Gly Gly Ile Leu

180 185 190

Gly Met Gly Glu Ala Glu Glu Asp Arg Ile Ala Met Leu Val Thr Leu

195 200 205

Ala Thr Leu Pro Ala His Pro Asp Ser Val Pro Val Asn Leu Trp Asn

210 215 220

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

225 230 235 240

Phe Ala Leu Val Arg Ile Val Ala Leu Ala Arg Ile Leu Met Pro Ala

245 250 255
Ser Val Val Arg Leu Ser Ala Gly Arg Thr Gly Met Ser Asp Glu Leu

260 265 270
Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Val Gly Asp

275 280 285
Gln Leu Leu Thr Thr Gly Asn Pro Ala Ala Trp Lys Asp Gln Asp Leu

290 295 300
Leu Ser Arg Leu Gly Met His Ile Ala Pro Ala Gln Ala Arg Pro Arg

305 310 315 320
Val Ala Ala Glu

<210> 32

<211> 963

<212> DNA

<213> Paracoccus denitrificans

<220><221> CDS

<222> (1)..(963)

<223> bioB gene encoding biotin synthase from Paracoccus denitrificans
PD1222(2)

<400> 32

atg ccg cag cag agc cgt agc gca gca gaa att tat cat cag ccg ctg 48

Met Pro Gln Gln Ser Arg Ser Ala Ala Glu Ile Tyr His Gln Pro Leu

1 5 10 15

atg gat ctg ctg ttt cag gca cag acc gtt cat cgt gca cat ttt gat 96

Met Asp Leu Leu Phe Gln Ala Gln Thr Val His Arg Ala His Phe Asp

20 25 30

ccg aat gtt gtt cag tgt agc aaa ctg ctg agc att aaa acc ggt ggt 144

Pro Asn Val Val Gln Cys Ser Lys Leu Leu Ser Ile Lys Thr Gly Gly

35 40 45

tgt ccg gaa gat tgt gca tat tgt agc cag agc gca cgt aat ggt agc 192

Cys Pro Glu Asp Cys Ala Tyr Cys Ser Gln Ser Ala Arg Asn Gly Ser

50	55	60	
gaa ctg agc gca agc aaa ctg atg gaa gtt cag cgt gtt ctg gca gaa			240
Glu Leu Ser Ala Ser Lys Leu Met Glu Val Gln Arg Val Leu Ala Glu			
65	70	75	80
gca cgt cgt gca aaa gaa gcc ggt gca acc cgt tat tgt atg ggt gca			288
Ala Arg Arg Ala Lys Glu Ala Gly Ala Thr Arg Tyr Cys Met Gly Ala			
	85	90	95
gca tgg cgt agc ccg aaa gaa cgt gat atg ccg gca gtt ctg gca atg			336
Ala Trp Arg Ser Pro Lys Glu Arg Asp Met Pro Ala Val Leu Ala Met			
	100	105	110
att cgt ggt gtt aaa gca atg ggt atg gaa acc tgt atg acc ctg ggt			384
Ile Arg Gly Val Lys Ala Met Gly Met Glu Thr Cys Met Thr Leu Gly			
	115	120	125
atg ctg gat gca gat cag gca ctg cgt ctg aaa gat gcc ggt ctg gat			432
Met Leu Asp Ala Asp Gln Ala Leu Arg Leu Lys Asp Ala Gly Leu Asp			
	130	135	140
tat tat aac cat aat att gat acc agc gaa cgc tat tat agc gaa att			480
Tyr Tyr Asn His Asn Ile Asp Thr Ser Glu Arg Tyr Tyr Ser Glu Ile			
145	150	155	160
att acc acc cgc acc ttt cag gat cgc att gaa acc ctg gaa cgt gtt			528
Ile Thr Thr Arg Thr Phe Gln Asp Arg Ile Glu Thr Leu Glu Arg Val			
	165	170	175
cag gca gca ggt att aat gtt tgt gcc ggt ggt att gtt ggt atg ggt			576
Gln Ala Ala Gly Ile Asn Val Cys Ala Gly Gly Ile Val Gly Met Gly			
	180	185	190
gaa acc gca gaa gat cgt att agc atg ctg gaa acc ctg gca ggt ctg			624
Glu Thr Ala Glu Asp Arg Ile Ser Met Leu Glu Thr Leu Ala Gly Leu			
	195	200	205
gaa gtg ccg ccg cag agc gtt ccg att aat atg ctg atg ccg atg gca			672

Glu Val Pro Pro Gln Ser Val Pro Ile Asn Met Leu Met Pro Met Ala
210 215 220
ggt acc ccg ctg gca gat gtt ccg cgt ctg gat gca att gaa atg gtt 720
Gly Thr Pro Leu Ala Asp Val Pro Arg Leu Asp Ala Ile Glu Met Val
225 230 235 240
cgt acc att gca acc gca cgt att ctg atg ccg gca agc tat gtt cgt 768
Arg Thr Ile Ala Thr Ala Arg Ile Leu Met Pro Ala Ser Tyr Val Arg
245 250 255
ctg agc gca ggt cgt agc gaa atg agc gat gaa atg cag gca atg tgt 816
Leu Ser Ala Gly Arg Ser Glu Met Ser Asp Glu Met Gln Ala Met Cys
260 265 270
ttt ttt gca ggc gca aat agc att ttt gtt ggt gat acc ctg ctg acc 864
Phe Phe Ala Gly Ala Asn Ser Ile Phe Val Gly Asp Thr Leu Leu Thr
275 280 285
gca ggt aat ccg gat gaa gat aaa gat gca ctg ctg ttt gca aaa ctg 912
Ala Gly Asn Pro Asp Glu Asp Lys Asp Ala Leu Leu Phe Ala Lys Leu
290 295 300
ggt ctg cgt gca gaa gtt ccg gaa gca agc cag gaa ggt tgt gca gca 960
Gly Leu Arg Ala Glu Val Pro Glu Ala Ser Gln Glu Gly Cys Ala Ala
305 310 315 320
taa 963

<210> 33

<211> 320

<212> PRT

<213> Paracoccus denitrificans

<400> 33

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

Cys Pro Glu Asp Cys Ala Tyr Cys Ser Gln Ser Ala Arg Asn Gly Ser

50 55 60

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

65 70 75 80

Ala Arg Arg Ala Lys Glu Ala Gly Ala Thr Arg Tyr Cys Met Gly Ala

85 90 95

Ala Trp Arg Ser Pro Lys Glu Arg Asp Met Pro Ala Val Leu Ala Met

100 105 110

Ile Arg Gly Val Lys Ala Met Gly Met Glu Thr Cys Met Thr Leu Gly

115 120 125

Met Leu Asp Ala Asp Gln Ala Leu Arg Leu Lys Asp Ala Gly Leu Asp

130 135 140

Tyr Tyr Asn His Asn Ile Asp Thr Ser Glu Arg Tyr Tyr Ser Glu Ile

145 150 155 160

Ile Thr Thr Arg Thr Phe Gln Asp Arg Ile Glu Thr Leu Glu Arg Val

165 170 175

Gln Ala Ala Gly Ile Asn Val Cys Ala Gly Gly Ile Val Gly Met Gly

180 185 190

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

195 200 205

Glu Val Pro Pro Gln Ser Val Pro Ile Asn Met Leu Met Pro Met Ala

210 215 220

Gly Thr Pro Leu Ala Asp Val Pro Arg Leu Asp Ala Ile Glu Met Val

225 230 235 240

Arg Thr Ile Ala Thr Ala Arg Ile Leu Met Pro Ala Ser Tyr Val Arg

245 250 255

Leu Ser Ala Gly Arg Ser Glu Met Ser Asp Glu Met Gln Ala Met Cys

260 265 270

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

275 280 285

Ala Gly Asn Pro Asp Glu Asp Lys Asp Ala Leu Leu Phe Ala Lys Leu

290 295 300
 Gly Leu Arg Ala Glu Val Pro Glu Ala Ser Gln Glu Gly Cys Ala Ala

305 310 315 320
 <210> 34
 <211> 987
 <212> DNA
 <213> Agrobacterium vitis
 <220><221> CDS
 <222> (1)..(987)
 <223> bioB gene encoding biotin synthase from Agrobacterium vitis S4
 <400> 34

atg agc gaa gca gcc ggt gaa att cgt aat gat tgg agc gtt gaa gaa 48
 Met Ser Glu Ala Ala Gly Glu Ile Arg Asn Asp Trp Ser Val Glu Glu
 1 5 10 15
 att gtg acc ctg cat aat ctg ccg ctg ctg gaa ctg att ggt cat gca 96

Ile Val Thr Leu His Asn Leu Pro Leu Leu Glu Leu Ile Gly His Ala
 20 25 30
 aat gca gtt cat ggt cgt cat cat aat ccg aat gtt gtt cag aaa gca 144
 Asn Ala Val His Gly Arg His His Asn Pro Asn Val Val Gln Lys Ala
 35 40 45
 agc ctg ctg agc att aaa acc ggt ggt tgt ccg gaa gat tgt gca tat 192
 Ser Leu Leu Ser Ile Lys Thr Gly Gly Cys Pro Glu Asp Cys Ala Tyr

50 55 60
 tgt ccg cag agc gca cat cat cgt gaa gtt aaa ctg acc aaa gat cgt 240
 Cys Pro Gln Ser Ala His His Arg Glu Val Lys Leu Thr Lys Asp Arg
 65 70 75 80
 ctg atg cag ccg gaa acc gtt ctg gca ctg gca aaa cgt gca aaa gat 288
 Leu Met Gln Pro Glu Thr Val Leu Ala Leu Ala Lys Arg Ala Lys Asp
 85 90 95

gcc ggt gca gaa cgt ttt tgt atg ggt gca gca tgg cgt cag gtt cgt 336
 Ala Gly Ala Glu Arg Phe Cys Met Gly Ala Ala Trp Arg Gln Val Arg

100	105	110	
gat ggt aaa gaa ttt gat gca gtt ctg aca atg gtt cgt ggt gtg cgt			384
Asp Gly Lys Glu Phe Asp Ala Val Leu Thr Met Val Arg Gly Val Arg			
115	120	125	
gat ctg ggt atg gaa gca tgt gtt acc ctg ggt atg ctg gaa aaa cat			432
Asp Leu Gly Met Glu Ala Cys Val Thr Leu Gly Met Leu Glu Lys His			
130	135	140	
cag gcc gaa aaa ctg gca gaa gca ggt ctg acc gca tat aat cat aat			480
Gln Ala Glu Lys Leu Ala Glu Ala Gly Leu Thr Ala Tyr Asn His Asn			
145	150	155	160
ctg gat acc agc ccg gaa ttt tat ggc gaa att att acc acc cgt agc			528
Leu Asp Thr Ser Pro Glu Phe Tyr Gly Glu Ile Ile Thr Thr Arg Ser			
165	170	175	
tat gca gat cgt ctg gaa acc ctg agc att gtt cgt agc ttt ggt att			576
Tyr Ala Asp Arg Leu Glu Thr Leu Ser Ile Val Arg Ser Phe Gly Ile			
180	185	190	
gat ctg tgt tgt ggt ggt att att ggt atg ggt gaa acc att cgt gat			624
Asp Leu Cys Cys Gly Gly Ile Ile Gly Met Gly Glu Thr Ile Arg Asp			
195	200	205	
cgt gca agt atg ctg cag gtt ctg gca agc atg cgt ccg cat ccg gaa			672
Arg Ala Ser Met Leu Gln Val Leu Ala Ser Met Arg Pro His Pro Glu			
210	215	220	
agt gtt ccg att aat gca ctg gtt ccg gtt gaa ggt acc ccg ctg gca			720
Ser Val Pro Ile Asn Ala Leu Val Pro Val Glu Gly Thr Pro Leu Ala			
225	230	235	240
gca atg ccg cgt att gat ccg ctg gaa ctg gtt cgt atg gtt gca acc			768
Ala Met Pro Arg Ile Asp Pro Leu Glu Leu Val Arg Met Val Ala Thr			
245	250	255	
gca cgt att gtt atg ccg aaa agc acc gtt cgt ctg agc gca ggt cgt			816
Ala Arg Ile Val Met Pro Lys Ser Thr Val Arg Leu Ser Ala Gly Arg			
260	265	270	

agc acc ctg aat cgt gaa gca cag att ctg tgt ctg gtt agc ggt gca 864

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

275 280 285

aat agc gtt ttt tat ggt gat acc ctg ctg acc acc ccg aat gca ggt 912

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

290 295 300

att ggt gaa gat gaa gca ctg ttt gca gca att ggt gca ctg ccg cat 960

Ile Gly Glu Asp Glu Ala Leu Phe Ala Ala Ile Gly Ala Leu Pro His

305 310 315 320

gaa gca gca ccg ctg gcc gca gaa taa 987

Glu Ala Ala Pro Leu Ala Ala Glu

325

<210> 35

<211> 328

<212> PRT

<213> Agrobacterium vitis

<400> 35

Met Ser Glu Ala Ala Gly Glu Ile Arg Asn Asp Trp Ser Val Glu Glu

1 5 10 15

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

20 25 30

Asn Ala Val His Gly Arg His His Asn Pro Asn Val Val Gln Lys Ala

35 40 45

Ser Leu Leu Ser Ile Lys Thr Gly Gly Cys Pro Glu Asp Cys Ala Tyr

50 55 60

Cys Pro Gln Ser Ala His His Arg Glu Val Lys Leu Thr Lys Asp Arg

65 70 75 80

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

85 90 95

Ala Gly Ala Glu Arg Phe Cys Met Gly Ala Ala Trp Arg Gln Val Arg

100 105 110

Asp Gly Lys Glu Phe Asp Ala Val Leu Thr Met Val Arg Gly Val Arg
115 120 125

Asp Leu Gly Met Glu Ala Cys Val Thr Leu Gly Met Leu Glu Lys His
130 135 140

Gln Ala Glu Lys Leu Ala Glu Ala Gly Leu Thr Ala Tyr Asn His Asn
145 150 155 160

Leu Asp Thr Ser Pro Glu Phe Tyr Gly Glu Ile Ile Thr Thr Arg Ser
165 170 175

Tyr Ala Asp Arg Leu Glu Thr Leu Ser Ile Val Arg Ser Phe Gly Ile
180 185 190

Asp Leu Cys Cys Gly Gly Ile Ile Gly Met Gly Glu Thr Ile Arg Asp
195 200 205

Arg Ala Ser Met Leu Gln Val Leu Ala Ser Met Arg Pro His Pro Glu
210 215 220

Ser Val Pro Ile Asn Ala Leu Val Pro Val Glu Gly Thr Pro Leu Ala
225 230 235 240

Ala Met Pro Arg Ile Asp Pro Leu Glu Leu Val Arg Met Val Ala Thr
245 250 255

Ala Arg Ile Val Met Pro Lys Ser Thr Val Arg Leu Ser Ala Gly Arg
260 265 270

Ser Thr Leu Asn Arg Glu Ala Gln Ile Leu Cys Leu Val Ser Gly Ala
275 280 285

Asn Ser Val Phe Tyr Gly Asp Thr Leu Leu Thr Thr Pro Asn Ala Gly
290 295 300

Ile Gly Glu Asp Glu Ala Leu Phe Ala Ala Ile Gly Ala Leu Pro His
305 310 315 320

Glu Ala Ala Pro Leu Ala Ala Glu
325

<210> 36

<211> 957

<212> DNA

<213> Ruegeria pomeroyi

<220><221> CDS

<222> (1)..(957)

<223> bioB gene encoding biotin synthase from Ruegeria pomeroyi DSS-3

<400

> 36

atg gcc gaa gca att cgt agc gat tgg agc gtt gat gaa gtt gaa gca 48

Met Ala Glu Ala Ile Arg Ser Asp Trp Ser Val Asp Glu Val Glu Ala

1 5 10 15

ctg ctg cgt ctg ccg ctg ctg gat ctg gtt ggt cgt gca aat ggt gtt 96

Leu Leu Arg Leu Pro Leu Leu Asp Leu Val Gly Arg Ala Asn Gly Val

20 25 30

cat cgt gca cat cat gca ccg gat gat att cag aaa gca agc ctg ctg 144

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

35 40 45

agc att aaa acc ggt ggt tgt ccg gaa gat tgt gca tat tgc ccg cag 192

Ser Ile Lys Thr Gly Gly Cys Pro Glu Asp Cys Ala Tyr Cys Pro Gln

50 55 60

agc gca cat cat cgt gaa gtg gaa ctg acc cgt gaa aaa ctg atg aat 240

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

65 70 75 80

ccg gat cat gtt gtt agc ctg gca cgt cgt gcc cag cgt gcc ggt gcc 288

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

85 90 95

gaa cgt ttt tgt atg ggt gca gca tgg cgt cag gtt cgt gat ggt gca 336

Glu Arg Phe Cys Met Gly Ala Ala Trp Arg Gln Val Arg Asp Gly Ala

100 105 110

gaa ttt gat aat gtt ctg gca atg gtt cgt ggt gtt cgt gca ctg ggt 384

Glu Phe Asp Asn Val Leu Ala Met Val Arg Gly Val Arg Ala Leu Gly

115 120 125

atg gaa gca tgt gtt acc ctg ggt atg ctg cgt ccg cat cag gca cag 432

Met Glu Ala Cys Val Thr Leu Gly Met Leu Arg Pro His Gln Ala Gln

130 135 140

cgt ctg gca gaa gca ggt ctg acc gca tat aat cat aat ctg gat acc 480

 Arg Leu Ala Glu Ala Gly Leu Thr Ala Tyr Asn His Asn Leu Asp Thr
 145 150 155 160
 agc ccg gaa ttt tat ggt cag att att ggt acc cgt acc tat cag gat 528
 Ser Pro Glu Phe Tyr Gly Gln Ile Ile Gly Thr Arg Thr Tyr Gln Asp
 165 170 175
 cgt ctg gat acc ctg gca tat tgt cgt gat gca ggt att gaa ctg tgt 576
 Arg Leu Asp Thr Leu Ala Tyr Cys Arg Asp Ala Gly Ile Glu Leu Cys

 180 185 190
 tgt ggt ggt att att ggc atg ggt gaa agc ctg cgt gat cgt gca gca 624
 Cys Gly Gly Ile Ile Gly Met Gly Glu Ser Leu Arg Asp Arg Ala Ala
 195 200 205
 atg ctg cag gtt ctg gcc aat ttt gca ccg cat ccg gaa agc gtt ccg 672
 Met Leu Gln Val Leu Ala Asn Phe Ala Pro His Pro Glu Ser Val Pro
 210 215 220

 att aat gca ctg att ccg att gaa ggt acc ccg ctg gca cat cgt gaa 720
 Ile Asn Ala Leu Ile Pro Ile Glu Gly Thr Pro Leu Ala His Arg Glu
 225 230 235 240
 cgt gtt ggt att ttt gat ctg gtt cgt atg gtt gca acc gca cgt att 768
 Arg Val Gly Ile Phe Asp Leu Val Arg Met Val Ala Thr Ala Arg Ile
 245 250 255
 att atg ccg ctg acc cgt gtt cgt ctg agc gca ggt cgt agt gat ttt 816

 Ile Met Pro Leu Thr Arg Val Arg Leu Ser Ala Gly Arg Ser Asp Phe
 260 265 270
 agc gcc gca gaa cag gca ctg tgt ttt ctg gcg ggt gca aat agc gtt 864
 Ser Ala Ala Glu Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Val
 275 280 285
 ttt tat ggt gat gtt ctg ctg acc gca ccg aat gca ggt acc ggt gca 912
 Phe Tyr Gly Asp Val Leu Leu Thr Ala Pro Asn Ala Gly Thr Gly Ala

290 295 300
 gat gca gaa ctg ttt gca gca ctg ggt gca ctg gaa acc gca taa 957
 Asp Ala Glu Leu Phe Ala Ala Leu Gly Ala Leu Glu Thr Ala
 305 310 315
 <210> 37
 <211> 318
 <212> PRT
 <213> Ruegeria pomeroyi
 <400> 37
 Met Ala Glu Ala Ile Arg Ser Asp Trp Ser Val Asp Glu Val Glu Ala
 1 5 10 15

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

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

 Arg Leu Ala Glu Ala Gly Leu Thr Ala Tyr Asn His Asn Leu Asp Thr
 145 150 155 160
 Ser Pro Glu Phe Tyr Gly Gln Ile Ile Gly Thr Arg Thr Tyr Gln Asp
 165 170 175
 Arg Leu Asp Thr Leu Ala Tyr Cys Arg Asp Ala Gly Ile Glu Leu Cys

180 185 190
Cys Gly Gly Ile Ile Gly Met Gly Glu Ser Leu Arg Asp Arg Ala Ala
195 200 205

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

Ser Ala Ala Glu Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Val
275 280 285
Phe Tyr Gly Asp Val Leu Leu Thr Ala Pro Asn Ala Gly Thr Gly Ala
290 295 300
Asp Ala Glu Leu Phe Ala Ala Leu Gly Ala Leu Glu Thr Ala
305 310 315

<210> 38

<211> 1020

<212> DNA

<213> Agrobacterium fabrum

<220><221> CDS

<222> (1)..(1020)

<223> bioB gene encoding biotin synthase from Agrobacterium fabrum str.

C58

<400> 38

atg gat cag ctg gca acc cag att gat ggt aaa ccg gca agc att ccg 48
Met Asp Gln Leu Ala Thr Gln Ile Asp Gly Lys Pro Ala Ser Ile Pro
1 5 10 15
gca gtt gaa acc agc agc agc ctg gaa gaa gcc aaa att att tat aat 96
Ala Val Glu Thr Ser Ser Ser Leu Glu Glu Ala Lys Ile Ile Tyr Asn
20 25 30

ctg ccg ttt aat gat ctg ctg ttt cgc gcc cag cag gtt cat cgt tgt 144

Leu Pro Phe Asn Asp Leu Leu Phe Arg Ala Gln Gln Val His Arg Cys
35 40 45

cat ttt gat gcc aat gca att cag atg agc cgt ctg ctg agc att aaa 192

His Phe Asp Ala Asn Ala Ile Gln Met Ser Arg Leu Leu Ser Ile Lys
50 55 60

acc ggc ggt tgt ccg gaa gat tgt agc tat tgt agc cag agc gca cgt 240

Thr Gly Gly Cys Pro Glu Asp Cys Ser Tyr Cys Ser Gln Ser Ala Arg

65 70 75 80

aat ccg acc ggt ctg aaa gca agc aaa ctg atg gaa gtt gaa cgt gtt 288

Asn Pro Thr Gly Leu Lys Ala Ser Lys Leu Met Glu Val Glu Arg Val
85 90 95

ctg gca gaa gca cgt aaa gca aaa gaa ggt ggt gca acc cgt tat tgt 336

Leu Ala Glu Ala Arg Lys Ala Lys Glu Gly Gly Ala Thr Arg Tyr Cys
100 105 110

atg ggt gca gca tgg cgt aat ccg aaa gaa cgt gat atg gaa gca gtt 384

Met Gly Ala Ala Trp Arg Asn Pro Lys Glu Arg Asp Met Glu Ala Val
115 120 125

gtt gca atg gtt gaa ggt gtt aaa gca ctg gat atg gaa acc tgt atg 432

Val Ala Met Val Glu Gly Val Lys Ala Leu Asp Met Glu Thr Cys Met
130 135 140

acc ctg ggt atg ctg acc ccg gaa cag agc gaa cgt ctg gca gat gca 480

Thr Leu Gly Met Leu Thr Pro Glu Gln Ser Glu Arg Leu Ala Asp Ala
145 150 155 160

ggt ctg gat tat tat aat cat aat gtg gat acc agc gaa cgt ttt tat 528

Gly Leu Asp Tyr Tyr Asn His Asn Val Asp Thr Ser Glu Arg Phe Tyr
165 170 175

agc gaa att att acc acc cgt acc ttt gaa gat cgc ctg gaa acc ctg 576

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

180	185	190	
gcc aat gtt cgt gat gca ggt att aaa gtt tgt gca ggc ggt att ctg			624
Ala Asn Val Arg Asp Ala Gly Ile Lys Val Cys Ala Gly Gly Ile Leu			
195	200	205	
ggt atg ggt gaa acc gtg gaa gat cgt att agc atg ctg gtt acc ctg			672
Gly Met Gly Glu Thr Val Glu Asp Arg Ile Ser Met Leu Val Thr Leu			
210	215	220	
gca aat ctg ccg gtt ccg ccg gaa agc gtt ccg att aat atg ctg att			720
Ala Asn Leu Pro Val Pro Pro Glu Ser Val Pro Ile Asn Met Leu Ile			
225	230	235	240
ccg att ccg ggt agc aaa ctg gca aat gca gat ccg gtt gat ccg att			768
Pro Ile Pro Gly Ser Lys Leu Ala Asn Ala Asp Pro Val Asp Pro Ile			
245	250	255	
gat ttt gtt cgt acc att gca ctg gca cgt att ctg atg ccg cgt agc			816
Asp Phe Val Arg Thr Ile Ala Leu Ala Arg Ile Leu Met Pro Arg Ser			
260	265	270	
cat gtt cgt ctg agc gca ggt cgt acc gaa atg agc gat gaa acc cag			864
His Val Arg Leu Ser Ala Gly Arg Thr Glu Met Ser Asp Glu Thr Gln			
275	280	285	
gca ctg tgt ttt ctg gcc ggt gca aat agc att ttt att ggt gaa acc			912
Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Ile Gly Glu Thr			
290	295	300	
ctg ctg acc gca gat aat ccg ggt gaa gat cat gat acc gca ctg ttt			960
Leu Leu Thr Ala Asp Asn Pro Gly Glu Asp His Asp Thr Ala Leu Phe			
305	310	315	320
cgt cgt ctg ggt ctg aaa ccg atg gaa ctg cag agc agc gaa gcc ggt			1008
Arg Arg Leu Gly Leu Lys Pro Met Glu Leu Gln Ser Ser Glu Ala Gly			
325	330	335	
ggt tgt cgt taa			1020
Gly Cys Arg			

<210> 39

<211> 339

<212> PRT

<213> *Agrobacterium fabrum*

<400> 39

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

1 5 10 15

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

20 25 30

Leu Pro Phe Asn Asp Leu Leu Phe Arg Ala Gln Gln Val His Arg Cys

35 40 45

His Phe Asp Ala Asn Ala Ile Gln Met Ser Arg Leu Leu Ser Ile Lys

50 55 60

Thr Gly Gly Cys Pro Glu Asp Cys Ser Tyr Cys Ser Gln Ser Ala Arg

65 70 75 80

Asn Pro Thr Gly Leu Lys Ala Ser Lys Leu Met Glu Val Glu Arg Val

85 90 95

Leu Ala Glu Ala Arg Lys Ala Lys Glu Gly Gly Ala Thr Arg Tyr Cys

100 105 110

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

115 120 125

Val Ala Met Val Glu Gly Val Lys Ala Leu Asp Met Glu Thr Cys Met

130 135 140

Thr Leu Gly Met Leu Thr Pro Glu Gln Ser Glu Arg Leu Ala Asp Ala

145 150 155 160

Gly Leu Asp Tyr Tyr Asn His Asn Val Asp Thr Ser Glu Arg Phe Tyr

165 170 175

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

180 185 190

Ala Asn Val Arg Asp Ala Gly Ile Lys Val Cys Ala Gly Gly Ile Leu

195 200 205

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

210 215 220

Ala Asn Leu Pro Val Pro Pro Glu Ser Val Pro Ile Asn Met Leu Ile

225 230 235 240

Pro Ile Pro Gly Ser Lys Leu Ala Asn Ala Asp Pro Val Asp Pro Ile

245 250 255

Asp Phe Val Arg Thr Ile Ala Leu Ala Arg Ile Leu Met Pro Arg Ser

260 265 270

His Val Arg Leu Ser Ala Gly Arg Thr Glu Met Ser Asp Glu Thr Gln

275 280 285

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

290 295 300

Leu Leu Thr Ala Asp Asn Pro Gly Glu Asp His Asp Thr Ala Leu Phe

305 310 315 320

Arg Arg Leu Gly Leu Lys Pro Met Glu Leu Gln Ser Ser Glu Ala Gly

325 330 335

Gly Cys Arg

<210> 40

<211> 954

<212> DNA

<213> Wolbachia endosymbiont of Cimex lectularius

<220><221> CDS

<222> (1)..(954)

<223> bioB gene encoding biotin synthase from Wolbachia endosymbiont of
Cimex lectularius

<400> 40

atg acc gaa gaa tgg acc ttt gcc aaa gca gat cag att ttt aat ttt 48

Met Thr Glu Glu Trp Thr Phe Ala Lys Ala Asp Gln Ile Phe Asn Phe

1 5 10 15

cgc ttt ccg gaa ctg att tat att gca cag acc gaa cat cgc aaa cag 96

Pro Phe Pro Glu Leu Ile Tyr Ile Ala Gln Thr Glu His Arg Lys Gln

20	25	30	
ttt aat ccg agc gaa gtg cag att agc acc ctg ctg agc att aaa acc			144
Phe Asn Pro Ser Glu Val Gln Ile Ser Thr Leu Leu Ser Ile Lys Thr			
35	40	45	
ggt agc tgt ccg gaa aat tgt agc tat tgt ccg cag agc gca cat tat			192
Gly Ser Cys Pro Glu Asn Cys Ser Tyr Cys Pro Gln Ser Ala His Tyr			
50	55	60	
aat acc ggt ctg cag aaa aaa ccg ctg ctg gaa att gca gaa gtt att			240
Asn Thr Gly Leu Gln Lys Lys Pro Leu Leu Glu Ile Ala Glu Val Ile			
65	70	75	80
gaa gca gca aaa tgt gca aaa gaa gca ggt agc acc cgt ttt tgt atg			288
Glu Ala Ala Lys Cys Ala Lys Glu Ala Gly Ser Thr Arg Phe Cys Met			
85	90	95	
ggt gca gca tgg cgt ggt ccg cgt gat cag gat ctg aaa gtt gtt tgt			336
Gly Ala Ala Trp Arg Gly Pro Arg Asp Gln Asp Leu Lys Val Val Cys			
100	105	110	
gaa atg att cgt gaa gtt aaa aaa ctg ggt ctg gaa acc tgt gtt acc			384
Glu Met Ile Arg Glu Val Lys Lys Leu Gly Leu Glu Thr Cys Val Thr			
115	120	125	
ctg ggt ctg ctg aaa gat cat cag gcc aac atg ctg aaa gaa gcc ggt			432
Leu Gly Leu Leu Lys Asp His Gln Ala Asn Met Leu Lys Glu Ala Gly			
130	135	140	
ctg gat ttt tat aac cat aac atc gat acc agc gaa gaa tat tat aac			480
Leu Asp Phe Tyr Asn His Asn Ile Asp Thr Ser Glu Glu Tyr Tyr Asn			
145	150	155	160
aaa gtg atc acc acc cgt acc ttt cag gat cgt ctg gaa acc ctg gaa			528
Lys Val Ile Thr Thr Arg Thr Phe Gln Asp Arg Leu Glu Thr Leu Glu			
165	170	175	
tgt gtt cgt gca agc ggt att aaa gtt tgt tgt ggt ggt att ctg ggt			576
Cys Val Arg Ala Ser Gly Ile Lys Val Cys Cys Gly Gly Ile Leu Gly			
180	185	190	

atg ggt gaa acc aat gaa gat cgt att aaa atg ctg gtt ctg ctg gca 624
Met Gly Glu Thr Asn Glu Asp Arg Ile Lys Met Leu Val Leu Leu Ala
195 200 205

aat ctg aat gat ccg ccg gaa agc gtt ccg att aat acc ctg att aaa 672

Asn Leu Asn Asp Pro Pro Glu Ser Val Pro Ile Asn Thr Leu Ile Lys
210 215 220

att ccg ggt acc ccg ctg gaa aat gtt gca gat gtt gat ccg ttt gat 720
Ile Pro Gly Thr Pro Leu Glu Asn Val Ala Asp Val Asp Pro Phe Asp
225 230 235 240

ttt gtt cgt acc att gca att gcc cgt att att atg ccg aaa agc tat 768
Phe Val Arg Thr Ile Ala Ile Ala Arg Ile Ile Met Pro Lys Ser Tyr

245 250 255

att cgt ctg agc gca ggt cgt gaa aaa atg agc gat gaa ctg cag gca 816
Ile Arg Leu Ser Ala Gly Arg Glu Lys Met Ser Asp Glu Leu Gln Ala
260 265 270

ctg tgt ttt ctg gcc ggt gca aat agc att ttt tat ggt gaa aaa ctg 864
Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr Gly Glu Lys Leu
275 280 285

ctg acc gca cag aat ccg att ccg gaa cag gat aat cat ctg ttt cag 912
Leu Thr Ala Gln Asn Pro Ile Pro Glu Gln Asp Asn His Leu Phe Gln
290 295 300

cgt ctg ggc ctg cag aaa ctg gca ctg ctg cgt gaa aat taa 954
Arg Leu Gly Leu Gln Lys Leu Ala Leu Leu Arg Glu Asn
305 310 315

<210> 41
<211> 317
<212> PRT
<213> Wolbachia endosymbiont of Cimex lectularius

<400> 41
Met Thr Glu Glu Trp Thr Phe Ala Lys Ala Asp Gln Ile Phe Asn Phe
1 5 10 15

Pro Phe Pro Glu Leu Ile Tyr Ile Ala Gln Thr Glu His Arg Lys Gln
20 25 30

Phe Asn Pro Ser Glu Val Gln Ile Ser Thr Leu Leu Ser Ile Lys Thr
35 40 45

Gly Ser Cys Pro Glu Asn Cys Ser Tyr Cys Pro Gln Ser Ala His Tyr
50 55 60

Asn Thr Gly Leu Gln Lys Lys Pro Leu Leu Glu Ile Ala Glu Val Ile
65 70 75 80

Glu Ala Ala Lys Cys Ala Lys Glu Ala Gly Ser Thr Arg Phe Cys Met
85 90 95

Gly Ala Ala Trp Arg Gly Pro Arg Asp Gln Asp Leu Lys Val Val Cys
100 105 110

Glu Met Ile Arg Glu Val Lys Lys Leu Gly Leu Glu Thr Cys Val Thr
115 120 125

Leu Gly Leu Leu Lys Asp His Gln Ala Asn Met Leu Lys Glu Ala Gly
130 135 140

Leu Asp Phe Tyr Asn His Asn Ile Asp Thr Ser Glu Glu Tyr Tyr Asn
145 150 155 160

Lys Val Ile Thr Thr Arg Thr Phe Gln Asp Arg Leu Glu Thr Leu Glu
165 170 175

Cys Val Arg Ala Ser Gly Ile Lys Val Cys Cys Gly Gly Ile Leu Gly
180 185 190

Met Gly Glu Thr Asn Glu Asp Arg Ile Lys Met Leu Val Leu Leu Ala
195 200 205

Asn Leu Asn Asp Pro Pro Glu Ser Val Pro Ile Asn Thr Leu Ile Lys
210 215 220

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

Phe Val Arg Thr Ile Ala Ile Ala Arg Ile Ile Met Pro Lys Ser Tyr
245 250 255

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

260 265 270
 Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr Gly Glu Lys Leu
 275 280 285
 Leu Thr Ala Gln Asn Pro Ile Pro Glu Gln Asp Asn His Leu Phe Gln
 290 295 300
 Arg Leu Gly Leu Gln Lys Leu Ala Leu Leu Arg Glu Asn
 305 310 315
 <210>
 42
 <211> 1026
 <212> DNA
 <213> Sphingomonas paucimobilis
 <220><221> CDS
 <222> (1)..(1026)
 <223> bioB gene encoding biotin synthase from Sphingomonas paucimobilis
 NBRC 13935
 <400> 42
 atg acc acc acc ccg gca ctg agc agt gaa gca acc ccg cgt acc gat 48
 Met Thr Thr Thr Pro Ala Leu Ser Ser Glu Ala Thr Pro Arg Thr Asp
 1 5 10 15
 tgg acc cgt gca gaa att gca gca ctg ttt gat ctg ccg ttt acc gaa 96
 Trp Thr Arg Ala Glu Ile Ala Ala Leu Phe Asp Leu Pro Phe Thr Glu
 20 25 30
 ctg ctg ttt cgt gcc gca gaa gtt cat cgt gca cat cat gca gca gat 144
 Leu Leu Phe Arg Ala Ala Glu Val His Arg Ala His His Ala Ala Asp
 35 40 45
 cag gtt cag ctg agc acc ctg ctg agc att aaa acc ggt ggt tgt ccg 192
 Gln Val Gln Leu Ser Thr Leu Leu Ser Ile Lys Thr Gly Gly Cys Pro
 50 55 60
 gaa gat tgt ggt tat tgt agc cag agc acc cat gca gat acc ggt ctg 240
 Glu Asp Cys Gly Tyr Cys Ser Gln Ser Thr His Ala Asp Thr Gly Leu
 65 70 75 80
 aaa gca acc aaa ctg atg gat ccg cgt gca gtt ctg cag gca gca gca 288

Lys Ala Thr Lys Leu Met Asp Pro Arg Ala Val Leu Gln Ala Ala Ala
85 90 95

cag gca aaa gat cat ggt agc acc cgt ttt tgt atg ggt gca gca tgg 336
Gln Ala Lys Asp His Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp
100 105 110

cgt aat ccg aaa gat cgt gat atg ccg gca att gtt gaa atg gtg aaa 384
Arg Asn Pro Lys Asp Arg Asp Met Pro Ala Ile Val Glu Met Val Lys
115 120 125

ggt gtt cgt gca atg ggt atg gaa acc tgt atg acc ctg ggt atg ctg 432

Gly Val Arg Ala Met Gly Met Glu Thr Cys Met Thr Leu Gly Met Leu
130 135 140

acc gat gca cag gca cag acc ctg gca gaa gca ggt ctg gat tat tat 480
Thr Asp Ala Gln Ala Gln Thr Leu Ala Glu Ala Gly Leu Asp Tyr Tyr
145 150 155 160

aat cat aat att gat acc agc ccg gaa cgt tat ggt gat gtg att acc 528
Asn His Asn Ile Asp Thr Ser Pro Glu Arg Tyr Gly Asp Val Ile Thr
165 170 175

acc cgt agc ttt ggt gaa cgt ctg gaa acc ctg gaa cat gtt cgt gat 576
Thr Arg Ser Phe Gly Glu Arg Leu Glu Thr Leu Glu His Val Arg Asp
180 185 190

gca ggt att aat gtt tgt tgt ggt ggt att gtt ggt atg ggt gaa acc 624
Ala Gly Ile Asn Val Cys Cys Gly Gly Ile Val Gly Met Gly Glu Thr
195 200 205

cgt ggt gat cgt gtt ggt ttt att cat gca ctg gca acc ctg ccg gtt 672
Arg Gly Asp Arg Val Gly Phe Ile His Ala Leu Ala Thr Leu Pro Val
210 215 220

cat ccg ggt agt gtg ccg gtt aat gca ctg gtt ccg gtt aaa ggt acc 720
His Pro Gly Ser Val Pro Val Asn Ala Leu Val Pro Val Lys Gly Thr
225 230 235 240

gtt ctg ggt gat atg ctg gca gat acc ccg ctg gca aaa att gat gat 768

Val Leu Gly Asp Met Leu Ala Asp Thr Pro Leu Ala Lys Ile Asp Asp

245 250 255

att gaa ttt gtt cgt acc gtt gca gtt gca cgt att acc atg ccg cat 816

Ile Glu Phe Val Arg Thr Val Ala Val Ala Arg Ile Thr Met Pro His

260 265 270

agc atg gtt cgt ctg agc gca ggt cgt gaa agc atg agc gat gca acc 864

Ser Met Val Arg Leu Ser Ala Gly Arg Glu Ser Met Ser Asp Ala Thr

275 280 285

cag gca ctg tgt ttt ctg gcc ggt gca aat agc att ttt acc ggt gat 912

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

290 295 300

aaa ctg ctg acc gca ggt aat gca ggc gat gat aaa gat gca gcc ctg 960

Lys Leu Leu Thr Ala Gly Asn Ala Gly Asp Asp Lys Asp Ala Ala Leu

305 310 315 320

ttt gca cgt ctg ggt ctg acc ccg atg gca gca gaa tgt aaa gtt gaa 1008

Phe Ala Arg Leu Gly Leu Thr Pro Met Ala Ala Glu Cys Lys Val Glu

325 330 335

ctg gaa gca gca gaa taa 1026

Leu Glu Ala Ala Glu

340

<210> 43

<211> 341

<212> PRT

<213> Sphingomonas paucimobilis

<400> 43

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

1 5 10 15

Trp Thr Arg Ala Glu Ile Ala Ala Leu Phe Asp Leu Pro Phe Thr Glu

20 25 30

Leu Leu Phe Arg Ala Ala Glu Val His Arg Ala His His Ala Ala Asp

35 40 45
 Gln Val Gln Leu Ser Thr Leu Leu Ser Ile Lys Thr Gly Gly Cys Pro
 50 55 60

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

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

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

 Ile Glu Phe Val Arg Thr Val Ala Val Ala Arg Ile Thr Met Pro His
 260 265 270
 Ser Met Val Arg Leu Ser Ala Gly Arg Glu Ser Met Ser Asp Ala Thr
 275 280 285

Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Thr Gly Asp
290 295 300

Lys Leu Leu Thr Ala Gly Asn Ala Gly Asp Asp Lys Asp Ala Ala Leu
305 310 315 320

Phe Ala Arg Leu Gly Leu Thr Pro Met Ala Ala Glu Cys Lys Val Glu
325 330 335

Leu Glu Ala Ala Glu
340

<210> 44

<211> 951

<212> DNA

<213> Acidithiobacillus ferrivorans

<220><221> CDS

<222> (1)..(951)

<223> bioB gene encoding biotin synthase from Acidithiobacillus
ferrivorans SS3

<400> 44

atg aat acc acc gca ccg ccg cag acc ctg gat gca att ctg gaa att 48
Met Asn Thr Thr Ala Pro Pro Gln Thr Leu Asp Ala Ile Leu Glu Ile

1 5 10 15
tat gcc agc ccg ttt aat gat ctg att ttt gaa gca cag aaa gtg cat 96
Tyr Ala Ser Pro Phe Asn Asp Leu Ile Phe Glu Ala Gln Lys Val His

20 25 30
cgc ctg cat ttt gat ccg aat gcc att cag tgt agc acc ctg ctg agc 144
Arg Leu His Phe Asp Pro Asn Ala Ile Gln Cys Ser Thr Leu Leu Ser

35 40 45

att aaa acc ggt ggt tgt ccg gaa gat tgt ggt tat tgt agc cag agc 192
Ile Lys Thr Gly Gly Cys Pro Glu Asp Cys Gly Tyr Cys Ser Gln Ser

50 55 60
gca cat cat cag acc gca ctg aaa gca gaa gca ctg atg gat ctg gaa 240
Ala His His Gln Thr Ala Leu Lys Ala Glu Ala Leu Met Asp Leu Glu

65 70 75 80

cag gtt cgt gca gca gca cag gaa gca aaa gca aat ggt gca cag cgt	288
Gln Val Arg Ala Ala Ala Gln Glu Ala Lys Ala Asn Gly Ala Gln Arg	
85 90 95	
ctg tgt atg ggt gca gca tgg cgt agc ccg cat gat aaa gat att gaa	336
Leu Cys Met Gly Ala Ala Trp Arg Ser Pro His Asp Lys Asp Ile Glu	
100 105 110	
aaa gtt gca gca atg att ggt gtt gtt aaa gaa tat ggt ctg gaa agc	384
Lys Val Ala Ala Met Ile Gly Val Val Lys Glu Tyr Gly Leu Glu Ser	
115 120 125	
tgt gtt acc ctg ggt atg ctg aaa ccg ggt cag gca gaa cgt ctg cag	432
Cys Val Thr Leu Gly Met Leu Lys Pro Gly Gln Ala Glu Arg Leu Gln	
130 135 140	
aat gcg ggt ctg gat tat tat aat cat aat ctg gat acc agc ccg gaa	480
Asn Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu	
145 150 155 160	
ttt tat ggt gaa gtt att cat acc cgt agc tat cag gat cgt ctg gat	528
Phe Tyr Gly Glu Val Ile His Thr Arg Ser Tyr Gln Asp Arg Leu Asp	
165 170 175	
acc ctg gaa acc gtt cgt agc gca ggt att aaa att tgt agc ggt ggc	576
Thr Leu Glu Thr Val Arg Ser Ala Gly Ile Lys Ile Cys Ser Gly Gly	
180 185 190	
att ctg ggt atg ggt gaa agc cgt cgt gat cgt gca cgt atg ctg cag	624
Ile Leu Gly Met Gly Glu Ser Arg Arg Asp Arg Ala Arg Met Leu Gln	
195 200 205	
att ctg gca cat ctg ccg cag gca ccg gaa agc att ccg att aat gca	672
Ile Leu Ala His Leu Pro Gln Ala Pro Glu Ser Ile Pro Ile Asn Ala	
210 215 220	
ctg gtt ccg gtt ccg ggt acc ccg ctg gaa gca gca gaa ccg att gat	720
Leu Val Pro Val Pro Gly Thr Pro Leu Glu Ala Ala Glu Pro Ile Asp	

225 230 235 240
 ggt ttt gaa ttt gtt cgt acc att gca gtt gca cgt att ctg ttt ccg 768
 Gly Phe Glu Phe Val Arg Thr Ile Ala Val Ala Arg Ile Leu Phe Pro
 245 250 255
 aaa gca tat gtt cgt ctg agc gca ggt cgt ggt gca atg agc gat gaa 816
 Lys Ala Tyr Val Arg Leu Ser Ala Gly Arg Gly Ala Met Ser Asp Glu
 260 265 270

 ctg cag gca ctg gca ttt ctg gcc ggt gca aat agc att ttt ctg ggt 864
 Leu Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn Ser Ile Phe Leu Gly
 275 280 285
 gat cgt ctg ctg acc acc gat aat gca agc atg ggt cat gat cag agc 912
 Asp Arg Leu Leu Thr Thr Asp Asn Ala Ser Met Gly His Asp Gln Ser
 290 295 300
 ctg ttt agc cgt ctg ggt ctg cat cgt agc gaa gca taa 951

 Leu Phe Ser Arg Leu Gly Leu His Arg Ser Glu Ala
 305 310 315
 <210> 45
 <211> 316
 <212> PRT
 <213> Acidithiobacillus ferrivorans
 <400> 45
 Met Asn Thr Thr Ala Pro Pro Gln Thr Leu Asp Ala Ile Leu Glu Ile
 1 5 10 15
 Tyr Ala Ser Pro Phe Asn Asp Leu Ile Phe Glu Ala Gln Lys Val His
 20 25 30
 Arg Leu His Phe Asp Pro Asn Ala Ile Gln Cys Ser Thr Leu Leu Ser

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

85 90 95
Leu Cys Met Gly Ala Ala Trp Arg Ser Pro His Asp Lys Asp Ile Glu

100 105 110
Lys Val Ala Ala Met Ile Gly Val Val Lys Glu Tyr Gly Leu Glu Ser

115 120 125
Cys Val Thr Leu Gly Met Leu Lys Pro Gly Gln Ala Glu Arg Leu Gln

130 135 140
Asn Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu

145 150 155 160
Phe Tyr Gly Glu Val Ile His Thr Arg Ser Tyr Gln Asp Arg Leu Asp

165 170 175
Thr Leu Glu Thr Val Arg Ser Ala Gly Ile Lys Ile Cys Ser Gly Gly

180 185 190
Ile Leu Gly Met Gly Glu Ser Arg Arg Asp Arg Ala Arg Met Leu Gln

195 200 205
Ile Leu Ala His Leu Pro Gln Ala Pro Glu Ser Ile Pro Ile Asn Ala

210 215 220
Leu Val Pro Val Pro Gly Thr Pro Leu Glu Ala Ala Glu Pro Ile Asp

225 230 235 240
Gly Phe Glu Phe Val Arg Thr Ile Ala Val Ala Arg Ile Leu Phe Pro

245 250 255
Lys Ala Tyr Val Arg Leu Ser Ala Gly Arg Gly Ala Met Ser Asp Glu

260 265 270
Leu Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn Ser Ile Phe Leu Gly

275 280 285
Asp Arg Leu Leu Thr Thr Asp Asn Ala Ser Met Gly His Asp Gln Ser

290 295 300
Leu Phe Ser Arg Leu Gly Leu His Arg Ser Glu Ala

305 310 315

<210> 46

<211> 996

<212> DNA

<213> Gallionella capsiferriiformans

<220><221> CDS

<222> (1)..(996)

<223> bioB gene encoding biotin synthase from Gallionella
capsiferriiformans ES-2

<400> 46

atg aat acc cag acc att gcc ttt cat cat ccg gtt aaa cgt acc gca 48

Met Asn Thr Gln Thr Ile Ala Phe His His Pro Val Lys Arg Thr Ala

1 5 10 15
acc ccg gaa cgt tgg agc gtt gaa gca gtt gaa agc ctg ttt gca ctg 96

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

20 25 30
ccg ttt gcc gat ctg ctg tat cgt gca cag cag gtt cat cgt gaa cat 144

Pro Phe Ala Asp Leu Leu Tyr Arg Ala Gln Gln Val His Arg Glu His

35 40 45
ttt gat ccg aat cag gtt cag ctg agc acc ctg ctg agc att aaa acc 192

Phe Asp Pro Asn Gln Val Gln Leu Ser Thr Leu Leu Ser Ile Lys Thr

50 55 60
ggg ggt tgt agc gaa gat tgt ggt tat tgt ccg cag agc gca ttt cat 240

Gly Gly Cys Ser Glu Asp Cys Gly Tyr Cys Pro Gln Ser Ala Phe His

65 70 75 80
agc acc ggt gtt gaa gat cgt aaa atg ctg gca ctg gat gca gtt att 288

Ser Thr Gly Val Glu Asp Arg Lys Met Leu Ala Leu Asp Ala Val Ile

85 90 95
gaa gca gca aaa gca gca cag gca gca ggc gca gat cgt ttt tgt atg 336

Glu Ala Ala Lys Ala Ala Gln Ala Ala Gly Ala Asp Arg Phe Cys Met

100 105 110
ggg gca gca tgg cgt gaa ccg agc gaa gca gat atg ctg agc gtt gtt 384

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

115 120 125

gat atg gtt cag gca gtt cgt ggc ctg ggt atg gaa acc tgt gca acc	432
Asp Met Val Gln Ala Val Arg Gly Leu Gly Met Glu Thr Cys Ala Thr	
130 135 140	
ctg ggt atg ctg aat gat gca cag acc gaa cag ctg cgt gca gcc ggt	480
Leu Gly Met Leu Asn Asp Ala Gln Thr Glu Gln Leu Arg Ala Ala Gly	
145 150 155 160	
ctg gat tat tat aat cat aat ctg gat acc agc ccg gaa ttt tat ggc	528
Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu Phe Tyr Gly	
165 170 175	
gat att att agc acc cgt gat tat cag gat cgc ctg gat acc ctg gaa	576
Asp Ile Ile Ser Thr Arg Asp Tyr Gln Asp Arg Leu Asp Thr Leu Glu	
180 185 190	
cgc gtt cgt cgt gca ggt atg cat gtt tgt agc ggt ggt att gtt ggt	624
Arg Val Arg Arg Ala Gly Met His Val Cys Ser Gly Gly Ile Val Gly	
195 200 205	
atg ggt gaa agt ctg acc gaa cgt gca ggt ctg gtt gcc cag ctg gca	672
Met Gly Glu Ser Leu Thr Glu Arg Ala Gly Leu Val Ala Gln Leu Ala	
210 215 220	
aat ctg aat ccg tat ccg gaa agc gtt ccg att aat aat ctg gtt aaa	720
Asn Leu Asn Pro Tyr Pro Glu Ser Val Pro Ile Asn Asn Leu Val Lys	
225 230 235 240	
gtt gaa ggt acc ccg ctg gca gat gca gca gaa ctg gat ccg ctg gat	768
Val Glu Gly Thr Pro Leu Ala Asp Ala Ala Glu Leu Asp Pro Leu Asp	
245 250 255	
ttt gtt cgt acc att gca gtt gca cgt att acc atg ccg acc gca cgt	816
Phe Val Arg Thr Ile Ala Val Ala Arg Ile Thr Met Pro Thr Ala Arg	
260 265 270	
gtt cgt ctg agc gca ggt cgt cag gca atg agc gat gca att cag gca	864
Val Arg Leu Ser Ala Gly Arg Gln Ala Met Ser Asp Ala Ile Gln Ala	
275 280 285	
ctg tgt ttt ctg gcc ggt gca aat agc att ttt tat ggt gaa cag ctg	912

Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr Gly Glu Gln Leu
 290 295 300
 ctg acc acc ggt aat ccg gaa gtt gaa cgt gat cgt gca ctg atg gat 960

Leu Thr Thr Gly Asn Pro Glu Val Glu Arg Asp Arg Ala Leu Met Asp
 305 310 315 320
 aaa ctg ggt atg tat ccg ttt gca gat aaa cat taa 996

Lys Leu Gly Met Tyr Pro Phe Ala Asp Lys His
 325 330

<210> 47

<211> 331

<212> PRT

<213> Gallionella capsiferriiformans

<400> 47

Met Asn Thr Gln Thr Ile Ala Phe His His Pro Val Lys Arg Thr Ala

1 5 10 15
 Thr Pro Glu Arg Trp Ser Val Glu Ala Val Glu Ser Leu Phe Ala Leu
 20 25 30

Pro Phe Ala Asp Leu Leu Tyr Arg Ala Gln Gln Val His Arg Glu His
 35 40 45

Phe Asp Pro Asn Gln Val Gln Leu Ser Thr Leu Leu Ser Ile Lys Thr
 50 55 60

Gly Gly Cys Ser Glu Asp Cys Gly Tyr Cys Pro Gln Ser Ala Phe His

65 70 75 80
 Ser Thr Gly Val Glu Asp Arg Lys Met Leu Ala Leu Asp Ala Val Ile
 85 90 95

Glu Ala Ala Lys Ala Ala Gln Ala Ala Gly Ala Asp Arg Phe Cys Met
 100 105 110

Gly Ala Ala Trp Arg Glu Pro Ser Glu Ala Asp Met Leu Ser Val Val
 115 120 125

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

130 135 140

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

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

 260 265 270
 Val Arg Leu Ser Ala Gly Arg Gln Ala Met Ser Asp Ala Ile Gln Ala
 275 280 285
 Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr Gly Glu Gln Leu
 290 295 300
 Leu Thr Thr Gly Asn Pro Glu Val Glu Arg Asp Arg Ala Leu Met Asp
 305 310 315 320
 Lys Leu Gly Met Tyr Pro Phe Ala Asp Lys His
 325 330

<210> 48

<211> 1029

<212> DNA

<213> *Ralstonia eutropha*

<220><221> CDS

<222> (1)..(1029)

<223> bioB gene encoding biotin synthase from *Ralstonia eutropha* JMP134

<400> 48

atg aat cag gca gca cag acc gtt gca acc att agc gca gaa gcc ctg	48
Met Asn Gln Ala Ala Gln Thr Val Ala Thr Ile Ser Ala Glu Ala Leu	
1 5 10 15	
cgt cag acc gca cgt aat acc cat gca ctg ccg gaa gat gca cgt tgg	96
Arg Gln Thr Ala Arg Asn Thr His Ala Leu Pro Glu Asp Ala Arg Trp	
20 25 30	
cgt gtt gat gat gtt gca gca ctg ttt gcc ctg ccg ttt aat gat ctg	144
Arg Val Asp Asp Val Ala Ala Leu Phe Ala Leu Pro Phe Asn Asp Leu	
35 40 45	
ctg ttt cgt gca cag cag gtt cat cgt gaa aat ttt gat gca aat acc	192
Leu Phe Arg Ala Gln Gln Val His Arg Glu Asn Phe Asp Ala Asn Thr	
50 55 60	
gtt cag ctg agc acc ctg ctg agc att aaa acc ggt ggt tgt gaa gaa	240
Val Gln Leu Ser Thr Leu Leu Ser Ile Lys Thr Gly Gly Cys Glu Glu	
65 70 75 80	
gat tgt ggt tat tgt ccg cag agc gca cat cat gat gcg ggt gtt aaa	288
Asp Cys Gly Tyr Cys Pro Gln Ser Ala His His Asp Ala Gly Val Lys	
85 90 95	
gcc gaa aaa ctg atg gaa ctg gat gaa gtt ctg gaa gca gca cgt gca	336
Ala Glu Lys Leu Met Glu Leu Asp Glu Val Leu Glu Ala Ala Arg Ala	
100 105 110	
gca aaa gca aat ggt gca acc cgt ttt tgt atg ggt gca gca tgg cgt	384
Ala Lys Ala Asn Gly Ala Thr Arg Phe Cys Met Gly Ala Ala Trp Arg	
115 120 125	
agc ccg aaa gat cgt cat ctg gaa ccg gtt atg gat atg gtt cgt gaa	432
Ser Pro Lys Asp Arg His Leu Glu Pro Val Met Asp Met Val Arg Glu	
130 135 140	
gtt aaa gca atg ggt ctg gaa acc tgt gtt acc ctg ggt atg ctg aaa	480
Val Lys Ala Met Gly Leu Glu Thr Cys Val Thr Leu Gly Met Leu Lys	
145 150 155 160	
gca gaa cag gcc cag cag ctg aaa gat gcc ggt ctg gat tat tat aat	528

Ala Glu Gln Ala Gln Gln Leu Lys Asp Ala Gly Leu Asp Tyr Tyr Asn	
165	170
175	
cat aat ctg gat acc agc ccg gaa ttt tat ggc aaa att att acc acc	576
His Asn Leu Asp Thr Ser Pro Glu Phe Tyr Gly Lys Ile Ile Thr Thr	
180	185
190	
cgt acc tat cag gat cgc ctg gat acc att ggc cat gtt cgt gat gca	624
Arg Thr Tyr Gln Asp Arg Leu Asp Thr Ile Gly His Val Arg Asp Ala	
195	200
205	
ggg att aat gtt tgt tgt ggt ggt att gtt ggt atg ggt gaa agc cgt	672
Gly Ile Asn Val Cys Cys Gly Gly Ile Val Gly Met Gly Glu Ser Arg	
210	215
220	
gaa gcc cgt gca ggt ctg att gca cag ctg gca aat atg gat ccg tat	720
Glu Ala Arg Ala Gly Leu Ile Ala Gln Leu Ala Asn Met Asp Pro Tyr	
225	230
235	240
ccg gaa agc gtt ccg att aat aat ctg gtt cag gtt gaa ggt acc ccg	768
Pro Glu Ser Val Pro Ile Asn Asn Leu Val Gln Val Glu Gly Thr Pro	
245	250
255	
ctg gca ggt acc gaa gca ctg gat ccg ttt gaa ttt gtt cgt acc att	816
Leu Ala Gly Thr Glu Ala Leu Asp Pro Phe Glu Phe Val Arg Thr Ile	
260	265
270	
gca gtt gca cgt att acc atg ccg ggt gca atg gtt cgt ctg agc gca	864
Ala Val Ala Arg Ile Thr Met Pro Gly Ala Met Val Arg Leu Ser Ala	
275	280
285	
ggg cgt gaa gca atg gat gaa gca ctg cag gca ctg tgt ttt atg gcc	912
Gly Arg Glu Ala Met Asp Glu Ala Leu Gln Ala Leu Cys Phe Met Ala	
290	295
300	
ggg gca aat agc att ttt tat ggt gaa aaa ctg ctg acc acc ggt aat	960
Gly Ala Asn Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Gly Asn	
305	310
315	320

ccg cag gca gat cgt gat cgt gca ctg ctg gca cgt ctg gat att cgt 1008

Pro Gln Ala Asp Arg Asp Arg Ala Leu Leu Ala Arg Leu Asp Ile Arg

325 330 335

gca gaa ggt tat gca ggt taa 1029

Ala Glu Gly Tyr Ala Gly

340

<210> 49

<211> 342

<212> PRT

<213> Ralstonia eutropha

<400> 49

Met Asn Gln Ala Ala Gln Thr Val Ala Thr Ile Ser Ala Glu Ala Leu

1 5 10 15

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

20 25 30

Arg Val Asp Asp Val Ala Ala Leu Phe Ala Leu Pro Phe Asn Asp Leu

35 40 45

Leu Phe Arg Ala Gln Gln Val His Arg Glu Asn Phe Asp Ala Asn Thr

50 55 60

Val Gln Leu Ser Thr Leu Leu Ser Ile Lys Thr Gly Gly Cys Glu Glu

65 70 75 80

Asp Cys Gly Tyr Cys Pro Gln Ser Ala His His Asp Ala Gly Val Lys

85 90 95

Ala Glu Lys Leu Met Glu Leu Asp Glu Val Leu Glu Ala Ala Arg Ala

100 105 110

Ala Lys Ala Asn Gly Ala Thr Arg Phe Cys Met Gly Ala Ala Trp Arg

115 120 125

Ser Pro Lys Asp Arg His Leu Glu Pro Val Met Asp Met Val Arg Glu

130 135 140

Val Lys Ala Met Gly Leu Glu Thr Cys Val Thr Leu Gly Met Leu Lys

145 150 155 160

Ala Glu Gln Ala Gln Gln Leu Lys Asp Ala Gly Leu Asp Tyr Tyr Asn
165 170 175
His Asn Leu Asp Thr Ser Pro Glu Phe Tyr Gly Lys Ile Ile Thr Thr
180 185 190
Arg Thr Tyr Gln Asp Arg Leu Asp Thr Ile Gly His Val Arg Asp Ala
195 200 205
Gly Ile Asn Val Cys Cys Gly Gly Ile Val Gly Met Gly Glu Ser Arg
210 215 220
Glu Ala Arg Ala Gly Leu Ile Ala Gln Leu Ala Asn Met Asp Pro Tyr
225 230 235 240
Pro Glu Ser Val Pro Ile Asn Asn Leu Val Gln Val Glu Gly Thr Pro
245 250 255
Leu Ala Gly Thr Glu Ala Leu Asp Pro Phe Glu Phe Val Arg Thr Ile
260 265 270
Ala Val Ala Arg Ile Thr Met Pro Gly Ala Met Val Arg Leu Ser Ala
275 280 285
Gly Arg Glu Ala Met Asp Glu Ala Leu Gln Ala Leu Cys Phe Met Ala
290 295 300
Gly Ala Asn Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Gly Asn
305 310 315 320
Pro Gln Ala Asp Arg Asp Arg Ala Leu Leu Ala Arg Leu Asp Ile Arg
325 330 335
Ala Glu Gly Tyr Ala Gly
340
<210> 50
<211> 1008
<212> DNA
<213> Bordetella parapertussis
<220><221> CDS
<222> (1)..(1008)
<223> bioB gene encoding biotin synthase from Bordetella parapertussis
12822

<400> 50

atg cat acc gca tat att ccg gtt ccg aca ccg gtt cgt ccg ccg agt 48

Met His Thr Ala Tyr Ile Pro Val Pro Thr Pro Val Arg Pro Pro Ser

1 5 10 15

gca gaa cgt tgg ccg ctg gca gca gtt gca gaa ctg ttt gaa ctg ccg 96

Ala Glu Arg Trp Pro Leu Ala Ala Val Ala Glu Leu Phe Glu Leu Pro

20 25 30

ttt ctg gat ctg ctg cat cgt gca cag cag gtt cat cgt cag cat ttt 144

Phe Leu Asp Leu Leu His Arg Ala Gln Gln Val His Arg Gln His Phe

35 40 45

gat gca aat acc gtt cag ctg agc agc ctg ctg agc att aaa acc ggt 192

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

50 55 60

ggt tgt ccg gaa gat tgt gca tat tgt ccg cag agc gca cat tat gat 240

Gly Cys Pro Glu Asp Cys Ala Tyr Cys Pro Gln Ser Ala His Tyr Asp

65 70 75 80

acc ggt gtt gat gca gat aaa ctg atg ccg ctg gat gaa gtt gtt cgt 288

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

85 90 95

gca gcc cgt gca gca cag gca aat ggt gca cag cgt ttt tgt atg ggt 336

Ala Ala Arg Ala Ala Gln Ala Asn Gly Ala Gln Arg Phe Cys Met Gly

100 105 110

gca gca tgg cgt agc ccg aaa ccg cat cat ctg gaa gca gtg gca gaa 384

Ala Ala Trp Arg Ser Pro Lys Pro His His Leu Glu Ala Val Ala Glu

115 120 125

atg att ggt gca gtt aaa gca ctg ggt atg gaa acc tgt gtt acc ctg 432

Met Ile Gly Ala Val Lys Ala Leu Gly Met Glu Thr Cys Val Thr Leu

130 135 140

ggt atg ctg cgt gat ggt cag gca gaa cag ctg aaa gca gca ggc ctg 480

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

145 150 155 160

gat tat tat aat cat aat ctg gat acc gca ccg gaa ttt tat ggt aaa 528

Asp Tyr Tyr Asn His Asn Leu Asp Thr Ala Pro Glu Phe Tyr Gly Lys
165 170 175

att att agc acc cgt acc tat cag gat cgt ctg gat acc ctg cag cag 576

Ile Ile Ser Thr Arg Thr Tyr Gln Asp Arg Leu Asp Thr Leu Gln Gln
180 185 190

gtt cgt gaa gca ggt att aat gtt tgt tgt ggt ggt att gtt ggt atg 624

Val Arg Glu Ala Gly Ile Asn Val Cys Cys Gly Gly Ile Val Gly Met

195 200 205

ggt gaa agc cgt cgt gat cgt gca ggt ctg gtt gca cag ctg gca aat 672

Gly Glu Ser Arg Arg Asp Arg Ala Gly Leu Val Ala Gln Leu Ala Asn
210 215 220

atg gaa ccg tat ccg gaa agc gtt ccg att aat aat ctg gtt cag gtg 720

Met Glu Pro Tyr Pro Glu Ser Val Pro Ile Asn Asn Leu Val Gln Val
225 230 235 240

gaa ggt acc ccg ctg gcg ggt gca gaa acc ctg gat ccg ttt gaa ttt 768

Glu Gly Thr Pro Leu Ala Gly Ala Glu Thr Leu Asp Pro Phe Glu Phe
245 250 255

att cgt acc att gca gtt gca cgt att acc atg ccg ctg gcc aaa gtt 816

Ile Arg Thr Ile Ala Val Ala Arg Ile Thr Met Pro Leu Ala Lys Val
260 265 270

cgt ctg agc gca ggt cgt gaa acc atg agc gat agc gaa cag gca ctg 864

Arg Leu Ser Ala Gly Arg Glu Thr Met Ser Asp Ser Glu Gln Ala Leu
275 280 285

tgt ttt atg gcc ggt gca aat agc att ttt tat ggt gat gtt ctg ctg 912

Cys Phe Met Ala Gly Ala Asn Ser Ile Phe Tyr Gly Asp Val Leu Leu
290 295 300

aca acc ggt aat ccg cag gtt gaa gca gat cgt cgt ctg ctg cag cgt 960

Thr Thr Gly Asn Pro Gln Val Glu Ala Asp Arg Arg Leu Leu Gln Arg

305	310										315					320					
ctg ggt atg cgt gca gaa ggt ctg ccg tgt gca gcc ggt cag gca taa																					1008
Leu Gly Met Arg Ala Glu Gly Leu Pro Cys Ala Ala Gly Gln Ala																					
	325										330					335					
<210>	51																				
<211>	335																				
<212>	PRT																				
<213>	Bordetella parapertussis																				
<400>	51																				
Met His Thr Ala Tyr Ile Pro Val Pro Thr Pro Val Arg Pro Pro Ser																					
1	5										10					15					
Ala Glu Arg Trp Pro Leu Ala Ala Val Ala Glu Leu Phe Glu Leu Pro																					
	20										25					30					
Phe Leu Asp Leu Leu His Arg Ala Gln Gln Val His Arg Gln His Phe																					
	35										40					45					
Asp Ala Asn Thr Val Gln Leu Ser Ser Leu Leu Ser Ile Lys Thr Gly																					
	50										55					60					
Gly Cys Pro Glu Asp Cys Ala Tyr Cys Pro Gln Ser Ala His Tyr Asp																					
65	70										75					80					
Thr Gly Val Asp Ala Asp Lys Leu Met Pro Leu Asp Glu Val Val Arg																					
	85										90					95					
Ala Ala Arg Ala Ala Gln Ala Asn Gly Ala Gln Arg Phe Cys Met Gly																					
	100										105					110					
Ala Ala Trp Arg Ser Pro Lys Pro His His Leu Glu Ala Val Ala Glu																					
	115										120					125					
Met Ile Gly Ala Val Lys Ala Leu Gly Met Glu Thr Cys Val Thr Leu																					
	130										135					140					
Gly Met Leu Arg Asp Gly Gln Ala Glu Gln Leu Lys Ala Ala Gly Leu																					
145	150										155					160					
Asp Tyr Tyr Asn His Asn Leu Asp Thr Ala Pro Glu Phe Tyr Gly Lys																					
	165										170					175					
Ile Ile Ser Thr Arg Thr Tyr Gln Asp Arg Leu Asp Thr Leu Gln Gln																					

180 185 190
Val Arg Glu Ala Gly Ile Asn Val Cys Cys Gly Gly Ile Val Gly Met
195 200 205

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

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

<210> 52
<211> 1044
<212> DNA
<213> Pusillimonas sp. T7-7
<220><221> CDS
<222> (1)..(1044)
<223> bioB gene encoding biotin synthase from Pusillimonas sp. T7-7
<400> 52

atg gca gca atg aaa ccg gca att ccg agc cat acc ccg acc ccg gat 48
Met Ala Ala Met Lys Pro Ala Ile Pro Ser His Thr Pro Thr Pro Asp
1 5 10 15
cat gca ccg cag gca tgg ggt att gca cag att ctg cgt ctg tat gaa 96

His Ala Pro Gln Ala Trp Gly Ile Ala Gln Ile Leu Arg Leu Tyr Glu

20 25 30

ctg ccg ttt ctg gat ctg ctg cat cag gca cag gcc gtt cat cgt gca 144

Leu Pro Phe Leu Asp Leu Leu His Gln Ala Gln Ala Val His Arg Ala

35 40 45

cat cat cag ccg aat acc gtt cag ctg agc agc ctg ctg agc att aaa 192

His His Gln Pro Asn Thr Val Gln Leu Ser Ser Leu Leu Ser Ile Lys

50 55 60

acc ggt gca tgt ccg gaa gat tgt gca tat tgt ccg cag agc gca cgt 240

Thr Gly Ala Cys Pro Glu Asp Cys Ala Tyr Cys Pro Gln Ser Ala Arg

65 70 75 80

cat gat acc ggt ggt aaa cag gaa gca ctg atg ccg gtt gca gaa gtt 288

His Asp Thr Gly Gly Lys Gln Glu Ala Leu Met Pro Val Ala Glu Val

85 90 95

ctg gaa gca gca cgt aaa gca aaa gca aat ggt gca cag cgt ttt tgt 336

Leu Glu Ala Ala Arg Lys Ala Lys Ala Asn Gly Ala Gln Arg Phe Cys

100 105 110

atg ggt gca gca tgg cgt agc ccg acc gca cgt cag ctg gat agc gtt 384

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

115 120 125

gtt gaa atg gtt ggt gca gtt aaa gca ctg ggt ctg gaa acc tgt gtt 432

Val Glu Met Val Gly Ala Val Lys Ala Leu Gly Leu Glu Thr Cys Val

130 135 140

acc ctg ggt atg ctg aaa gaa ggt cag gca gaa cgt ctg cgt gat gcg 480

Thr Leu Gly Met Leu Lys Glu Gly Gln Ala Glu Arg Leu Arg Asp Ala

145 150 155 160

ggt ctg gat tat tat aat cat aat ctg gat acc agc ccg gaa ttt tat 528

Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu Phe Tyr

165 170 175

ggt aat att att acc acc cgt agc tat cag gat cgt ctg gat acc ctg 576
Gly Asn Ile Ile Thr Thr Arg Ser Tyr Gln Asp Arg Leu Asp Thr Leu
180 185 190

gaa cgt gtt cgt aat gcc ggt gtt cat gtt tgt tgt ggt ggt att gtt 624
Glu Arg Val Arg Asn Ala Gly Val His Val Cys Cys Gly Gly Ile Val
195 200 205

ggt ctg ggt gaa agc cgt aaa gaa cgt gca ggt ctg gtt gca cag ctg 672

Gly Leu Gly Glu Ser Arg Lys Glu Arg Ala Gly Leu Val Ala Gln Leu
210 215 220

gca aat ctg agc ccg tat ccg gaa agc gtt ccg gtt aat aat ctg gtt 720
Ala Asn Leu Ser Pro Tyr Pro Glu Ser Val Pro Val Asn Asn Leu Val
225 230 235 240

aaa gtt gca ggt acc ccg ctg gat gcc acc ccg gat att gat ccg ttt 768
Lys Val Ala Gly Thr Pro Leu Asp Ala Thr Pro Asp Ile Asp Pro Phe
245 250 255

gaa ttt gtt cgt acc att gca gtt gca cgt att acc atg ccg cgt gca 816
Glu Phe Val Arg Thr Ile Ala Val Ala Arg Ile Thr Met Pro Arg Ala
260 265 270

gtt gtt cgt ctg agc gca ggt cgt gaa gca atg agc gat gca att cag 864
Val Val Arg Leu Ser Ala Gly Arg Glu Ala Met Ser Asp Ala Ile Gln
275 280 285

gca ctg tgt ttt atg gcc ggt gca aat agc att ttt tat ggt gaa cag 912
Ala Leu Cys Phe Met Ala Gly Ala Asn Ser Ile Phe Tyr Gly Glu Gln
290 295 300

ctg ctg aca aca gca aat ccg cag ctg agt cag gat cag caa ctg ttt 960
Leu Leu Thr Thr Ala Asn Pro Gln Leu Ser Gln Asp Gln Gln Leu Phe
305 310 315 320

cag cgt ctg ggt ctg aca gca acc ccg gca gat ccg gca cgt ccg gca 1008

Gln Arg Leu Gly Leu Thr Ala Thr Pro Ala Asp Pro Ala Arg Pro Ala
325 330 335

cat ctg gaa cat cat cat gaa gca acc ctg gca taa 1044

His Leu Glu His His His Glu Ala Thr Leu Ala

340

345

<210> 53

<211> 347

<212> PRT

<213> Pusillimonas sp. T7-7

<400> 53

Met Ala Ala Met Lys Pro Ala Ile Pro Ser His Thr Pro Thr Pro Asp

1

5

10

15

His Ala Pro Gln Ala Trp Gly Ile Ala Gln Ile Leu Arg Leu Tyr Glu

20

25

30

Leu Pro Phe Leu Asp Leu Leu His Gln Ala Gln Ala Val His Arg Ala

35

40

45

His His Gln Pro Asn Thr Val Gln Leu Ser Ser Leu Leu Ser Ile Lys

50

55

60

Thr Gly Ala Cys Pro Glu Asp Cys Ala Tyr Cys Pro Gln Ser Ala Arg

65

70

75

80

His Asp Thr Gly Gly Lys Gln Glu Ala Leu Met Pro Val Ala Glu Val

85

90

95

Leu Glu Ala Ala Arg Lys Ala Lys Ala Asn Gly Ala Gln Arg Phe Cys

100

105

110

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

115

120

125

Val Glu Met Val Gly Ala Val Lys Ala Leu Gly Leu Glu Thr Cys Val

130

135

140

Thr Leu Gly Met Leu Lys Glu Gly Gln Ala Glu Arg Leu Arg Asp Ala

145

150

155

160

Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu Phe Tyr

165

170

175

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

180

185

190

Glu Arg Val Arg Asn Ala Gly Val His Val Cys Cys Gly Gly Ile Val

195 200 205

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

210 215 220

Ala Asn Leu Ser Pro Tyr Pro Glu Ser Val Pro Val Asn Asn Leu Val

225 230 235 240

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

245 250 255

Glu Phe Val Arg Thr Ile Ala Val Ala Arg Ile Thr Met Pro Arg Ala

260 265 270

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

275 280 285

Ala Leu Cys Phe Met Ala Gly Ala Asn Ser Ile Phe Tyr Gly Glu Gln

290 295 300

Leu Leu Thr Thr Ala Asn Pro Gln Leu Ser Gln Asp Gln Gln Leu Phe

305 310 315 320

Gln Arg Leu Gly Leu Thr Ala Thr Pro Ala Asp Pro Ala Arg Pro Ala

325 330 335

His Leu Glu His His His Glu Ala Thr Leu Ala

340 345

<210> 54

<211> 957

<212> DNA

<213> Cenarchaeum symbiosum A

<220><221> CDS

<222> (1)..(957)

<223> bioB gene encoding biotin synthase from Cenarchaeum symbiosum A

<400> 54

atg ggt att gca gaa tgt cgt gat aaa gtt ctg ggt ggt ggt gaa ctg 48

Met Gly Ile Ala Glu Cys Arg Asp Lys Val Leu Gly Gly Gly Glu Leu

1 5 10 15

acc aaa gat gaa gca cgt ggt ctg atg gaa gca gat gtt acc gaa ctg	96
Thr Lys Asp Glu Ala Arg Gly Leu Met Glu Ala Asp Val Thr Glu Leu	
20 25 30	
gcc gca gca gca gat gaa att acc cgt cgt ttt aat ggt gat ggt gtg	144
Ala Ala Ala Ala Asp Glu Ile Thr Arg Arg Phe Asn Gly Asp Gly Val	
35 40 45	
gat gtg gaa cag ctg aat aat att aaa cgt gat ggt tgc agc gaa gat	192
Asp Val Glu Gln Leu Asn Asn Ile Lys Arg Asp Gly Cys Ser Glu Asp	
50 55 60	
tgt acc ttt tgt ggt cag agc gcc ttt tat gat gca gat aaa gaa ccg	240
Cys Thr Phe Cys Gly Gln Ser Ala Phe Tyr Asp Ala Asp Lys Glu Pro	
65 70 75 80	
cat ccg ctg ccg gaa ccg gaa gaa gtt gtt cgt gca gcc ctg aaa gca	288
His Pro Leu Pro Glu Pro Glu Glu Val Val Arg Ala Ala Leu Lys Ala	
85 90 95	
aaa aaa gaa gaa gcc agc agc tat tgt ctg gtt gcc gca tgg cgt gaa	336
Lys Lys Glu Glu Ala Ser Ser Tyr Cys Leu Val Ala Ala Trp Arg Glu	
100 105 110	
ccg acc ccg gaa ggt ttt gaa aaa gtt tgt acc att att cag gaa att	384
Pro Thr Pro Glu Gly Phe Glu Lys Val Cys Thr Ile Ile Gln Glu Ile	
115 120 125	
aat acc cat gtt ggt att agc gtt gaa tgt agc ctg ggt ttt ctg acc	432
Asn Thr His Val Gly Ile Ser Val Glu Cys Ser Leu Gly Phe Leu Thr	
130 135 140	
cgt gaa cgt gca gca cgt ctg aaa ggt ctg ggt gtt aaa cgt tat aat	480
Arg Glu Arg Ala Ala Arg Leu Lys Gly Leu Gly Val Lys Arg Tyr Asn	
145 150 155 160	
cat aat ctg gaa acc gcc cgt agc aaa ttt ccg gaa att tgt agc acc	528
His Asn Leu Glu Thr Ala Arg Ser Lys Phe Pro Glu Ile Cys Ser Thr	
165 170 175	
cat acc tat gaa gat cgt ctg gat acc ctg gaa att gca cgt gaa gcc	576

His Thr Tyr Glu Asp Arg Leu Asp Thr Leu Glu Ile Ala Arg Glu Ala
180 185 190
ggg ctg gaa ctg tgt acc ggt ggt att att ggt atg ggt gaa agc cgt 624
Gly Leu Glu Leu Cys Thr Gly Gly Ile Ile Gly Met Gly Glu Ser Arg
195 200 205
ggg cag cgt att gaa ctg gca atg gaa ctg gca cgt att cgt ccg gaa 672
Gly Gln Arg Ile Glu Leu Ala Met Glu Leu Ala Arg Ile Arg Pro Glu
210 215 220
gaa gca acc gtt aat att ctg gtt ccg gtt cag ggt acc ccg atg gaa 720
Glu Ala Thr Val Asn Ile Leu Val Pro Val Gln Gly Thr Pro Met Glu
225 230 235 240
ctg cag gca ccg ctg ccg ccg ggt gaa gca gaa cgt ttt ttt gca ctg 768
Leu Gln Ala Pro Leu Pro Pro Gly Glu Ala Glu Arg Phe Phe Ala Leu
245 250 255
gtt cgt ttt ctg ctg ccg cgt agc gtt gtt aaa att agc ggt ggt cgt 816
Val Arg Phe Leu Leu Pro Arg Ser Val Val Lys Ile Ser Gly Gly Arg
260 265 270
gaa aaa gca ctg gat gat gat ggt cgt gca att ctg cgt ggt ggt gca 864
Glu Lys Ala Leu Asp Asp Asp Gly Arg Ala Ile Leu Arg Gly Gly Ala
275 280 285
aat ggt att att acc agc ggt tat ctg aca atg ggt ggt aat gat agc 912
Asn Gly Ile Ile Thr Ser Gly Tyr Leu Thr Met Gly Gly Asn Asp Ser
290 295 300
agc gca gat atg gaa atg att cgt gaa gca ggt ctg gaa gca taa 957
Ser Ala Asp Met Glu Met Ile Arg Glu Ala Gly Leu Glu Ala
305 310 315
<210> 55
<211> 318
<212> PRT
<213> Cenarchaeum symbiosum A
<400> 55

Met Gly Ile Ala Glu Cys Arg Asp Lys Val Leu Gly Gly Gly Glu Leu
1 5 10 15

Thr Lys Asp Glu Ala Arg Gly Leu Met Glu Ala Asp Val Thr Glu Leu
 20 25 30

Ala Ala Ala Ala Asp Glu Ile Thr Arg Arg Phe Asn Gly Asp Gly Val
 35 40 45

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

Cys Thr Phe Cys Gly Gln Ser Ala Phe Tyr Asp Ala Asp Lys Glu Pro
65 70 75 80

His Pro Leu Pro Glu Pro Glu Glu Val Val Arg Ala Ala Leu Lys Ala
 85 90 95

Lys Lys Glu Glu Ala Ser Ser Tyr Cys Leu Val Ala Ala Trp Arg Glu
 100 105 110

Pro Thr Pro Glu Gly Phe Glu Lys Val Cys Thr Ile Ile Gln Glu Ile
 115 120 125

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

Arg Glu Arg Ala Ala Arg Leu Lys Gly Leu Gly Val Lys Arg Tyr Asn
145 150 155 160

His Asn Leu Glu Thr Ala Arg Ser Lys Phe Pro Glu Ile Cys Ser Thr
 165 170 175

His Thr Tyr Glu Asp Arg Leu Asp Thr Leu Glu Ile Ala Arg Glu Ala
 180 185 190

Gly Leu Glu Leu Cys Thr Gly Gly Ile Ile Gly Met Gly Glu Ser Arg
 195 200 205

Gly Gln Arg Ile Glu Leu Ala Met Glu Leu Ala Arg Ile Arg Pro Glu
 210 215 220

Glu Ala Thr Val Asn Ile Leu Val Pro Val Gln Gly Thr Pro Met Glu
225 230 235 240

Leu Gln Ala Pro Leu Pro Pro Gly Glu Ala Glu Arg Phe Phe Ala Leu

245	250	255	
Val Arg Phe Leu Leu Pro Arg Ser Val Val Lys Ile Ser Gly Gly Arg			
260	265	270	
Glu Lys Ala Leu Asp Asp Asp Gly Arg Ala Ile Leu Arg Gly Gly Ala			
275	280	285	
Asn Gly Ile Ile Thr Ser Gly Tyr Leu Thr Met Gly Gly Asn Asp Ser			
290	295	300	
Ser Ala Asp Met Glu Met Ile Arg Glu Ala Gly Leu Glu Ala			
305	310	315	
<210> 56			
<211> 999			
<212> DNA			
<213> Alicyclobacillus acidocaldarius subsp. acidocaldarius DSM 446			
<220><221> CDS			
<222> (1)..(999)			
<223> bioB gene encoding biotin synthase from Alicyclobacillus			
acidocaldarius subsp. acidocaldarius DSM 446			
<400> 56			
atg atg aaa att gat tat cag acc aat tgg att gat ctg gcc cgt cgt	48		
Met Met Lys Ile Asp Tyr Gln Thr Asn Trp Ile Asp Leu Ala Arg Arg			
1	5	10	15
gtt ctg gat ggt cgt ggt gtt acc cgt gaa gaa gca ctg gat att ctg	96		
Val Leu Asp Gly Arg Gly Val Thr Arg Glu Glu Ala Leu Asp Ile Leu			
20	25	30	
cgt tct agc gat gat gaa ctg ctg gat ctg ctg gca gca gcc ttt ctg	144		
Arg Ser Ser Asp Asp Glu Leu Leu Asp Leu Leu Ala Ala Ala Phe Leu			
35	40	45	
att cgt cgc cgc tat ttt ggc aaa aaa gtg aaa ctg aat atg att att	192		
Ile Arg Arg Arg Tyr Phe Gly Lys Lys Val Lys Leu Asn Met Ile Ile			
50	55	60	
aat gcc aaa agc aaa atg tgc ccg gaa gat tgc gcc tat tgc agc cag	240		
Asn Ala Lys Ser Lys Met Cys Pro Glu Asp Cys Ala Tyr Cys Ser Gln			

65	70	75	80	
agc gcc att agc aaa gca ccg gtt agc aaa tat ccg ctg gtt agt aaa				288
Ser Ala Ile Ser Lys Ala Pro Val Ser Lys Tyr Pro Leu Val Ser Lys				
85	90	95		
gaa gaa att att gcc ggt gca cgt gaa gca gaa cgt cgt aaa gca ggt				336
Glu Glu Ile Ile Ala Gly Ala Arg Glu Ala Glu Arg Arg Lys Ala Gly				
100	105	110		
acc tat tgt att gtt att agc ggt cgt cgt ccg agc gat cgt gaa att				384
Thr Tyr Cys Ile Val Ile Ser Gly Arg Arg Pro Ser Asp Arg Glu Ile				
115	120	125		
gaa cgt att gca gaa gca gtt gaa gaa att cgt gca acc acc acc ctg				432
Glu Arg Ile Ala Glu Ala Val Glu Glu Ile Arg Ala Thr Thr Thr Leu				
130	135	140		
aaa att tgt tgt tgt ctg ggt ctg ctg acc ccg gca cag gca gat cgt				480
Lys Ile Cys Cys Cys Leu Gly Leu Leu Thr Pro Ala Gln Ala Asp Arg				
145	150	155	160	
ctg gca cgt gcg ggt gtt cat cgt tat aat cat aat ctg aat acc agc				528
Leu Ala Arg Ala Gly Val His Arg Tyr Asn His Asn Leu Asn Thr Ser				
165	170	175		
cgt gat cgt tat ggt gat att tgt acc acc cat acc tat gat gat cgc				576
Arg Asp Arg Tyr Gly Asp Ile Cys Thr Thr His Thr Tyr Asp Asp Arg				
180	185	190		
gtt cgt acc ctg gaa cat gtg aaa gaa gca ggt att agc ccg tgt agc				624
Val Arg Thr Leu Glu His Val Lys Glu Ala Gly Ile Ser Pro Cys Ser				
195	200	205		
ggt gtt att ttt ggt atg ggt gaa agc gat gaa gaa gcc gtt gat atg				672
Gly Val Ile Phe Gly Met Gly Glu Ser Asp Glu Glu Ala Val Asp Met				
210	215	220		
gcc ttt gcc ctg aaa gaa atg gat gca gat agc att ccg tgt aat ttt				720

Ala Phe Ala Leu Lys Glu Met Asp Ala Asp Ser Ile Pro Cys Asn Phe

225 230 235 240
ctg aat ccg att ccg ggt acc ccg ctg gaa ggt atg gaa acc ctg aat 768

Leu Asn Pro Ile Pro Gly Thr Pro Leu Glu Gly Met Glu Thr Leu Asn

245 250 255
ccg cgt cgt tgt ctg aaa ctg ctg tgt atg atg cgt ttt gtt aat ccg 816

Pro Arg Arg Cys Leu Lys Leu Leu Cys Met Met Arg Phe Val Asn Pro

260 265 270

agc aaa gaa att cgt att gcg ggt ggt cgt gaa cgt aat ctg cgt agc 864

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

275 280 285

ctg cag gtt ctg ggt ctg tat ccg gca aat agc att ttt gtt ggt gat 912

Leu Gln Val Leu Gly Leu Tyr Pro Ala Asn Ser Ile Phe Val Gly Asp

290 295 300

tat ctg acc acc ccg ggt cag gca ccg acc gaa gat tgg gca atg att 960

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

305 310 315 320

gaa gat ctg ggt ttt gaa att gaa gaa tgt gca ctg taa 999

Glu Asp Leu Gly Phe Glu Ile Glu Glu Cys Ala Leu

325 330

<210> 57

<211> 332

<212> PRT

<213> Alicyclobacillus acidocaldarius subsp. acidocaldarius DSM 446

<400> 57

Met Met Lys Ile Asp Tyr Gln Thr Asn Trp Ile Asp Leu Ala Arg Arg

1 5 10 15

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

20 25 30

Arg Ser Ser Asp Asp Glu Leu Leu Asp Leu Leu Ala Ala Ala Phe Leu

35 40 45

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

290 295 300
Tyr Leu Thr Thr Pro Gly Gln Ala Pro Thr Glu Asp Trp Ala Met Ile
305 310 315 320
Glu Asp Leu Gly Phe Glu Ile Glu Glu Cys Ala Leu

325 330

<210> 58
<211> 996
<212> DNA
<213> Geobacillus thermoglucosidasius C56-YS93
<220><221> CDS
<222> (1)..(996)
<223> bioB gene encoding biotin synthase from Geobacillus
thermoglucosidasius C56-YS93

<400> 58

atg att aat tgg ctg gcc ctg gca gat cgt gtt att gca ggt cat gaa 48
Met Ile Asn Trp Leu Ala Leu Ala Asp Arg Val Ile Ala Gly His Glu
1 5 10 15

ctg acc gat gaa gaa gca ctg gca att ctg gat tgt ccg gat gaa gaa 96
Leu Thr Asp Glu Glu Ala Leu Ala Ile Leu Asp Cys Pro Asp Glu Glu
20 25 30

ctg ctg ctg ctg atg cag ggt gcc tat aat att cgt cgc acc tat tat 144
Leu Leu Leu Leu Met Gln Gly Ala Tyr Asn Ile Arg Arg Thr Tyr Tyr
35 40 45

ggc aat aaa gtt aaa ctg aat atg att att aat gcc aaa agc ggt ctg 192

Gly Asn Lys Val Lys Leu Asn Met Ile Ile Asn Ala Lys Ser Gly Leu
50 55 60

tgc ccg gaa aat tgc ggc tat tgc gca cag agc gca gtt agc acc gca 240
Cys Pro Glu Asn Cys Gly Tyr Cys Ala Gln Ser Ala Val Ser Thr Ala
65 70 75 80

ccg gtt aaa acc tat aaa atg gtt gat aaa gaa acc ctg att cgt ggt 288
Pro Val Lys Thr Tyr Lys Met Val Asp Lys Glu Thr Leu Ile Arg Gly

85	90	95	
gca gaa gaa gca tat cgt atg cgt att ggt acc tat tgt att gtt gca			336
Ala Glu Glu Ala Tyr Arg Met Arg Ile Gly Thr Tyr Cys Ile Val Ala			
100	105	110	
agc ggt cgt ggt ccg agc gaa aaa gaa att gat acc gtt gtg agc gcc			384
Ser Gly Arg Gly Pro Ser Glu Lys Glu Ile Asp Thr Val Val Ser Ala			
115	120	125	
gtt aaa gaa att aaa gaa cgt ttt ggt ctg aaa att tgt gca tgt ctg			432
Val Lys Glu Ile Lys Glu Arg Phe Gly Leu Lys Ile Cys Ala Cys Leu			
130	135	140	
ggt att ctg aaa ccg gaa cag gca gca cgt ctg aaa gaa gcc ggt gtt			480
Gly Ile Leu Lys Pro Glu Gln Ala Ala Arg Leu Lys Glu Ala Gly Val			
145	150	155	160
gat cgc tat aat cat aat att aat acc agc aaa gaa cat cat ccg aat			528
Asp Arg Tyr Asn His Asn Ile Asn Thr Ser Lys Glu His His Pro Asn			
165	170	175	
att acc acc agc cat acc tat gat gat cgc gtg cgt acc gtt gaa acc			576
Ile Thr Thr Ser His Thr Tyr Asp Asp Arg Val Arg Thr Val Glu Thr			
180	185	190	
gtt aaa cag gca ggt att agc ccg tgt agc ggt gtt att att ggt atg			624
Val Lys Gln Ala Gly Ile Ser Pro Cys Ser Gly Val Ile Ile Gly Met			
195	200	205	
cgt gaa acc aaa cag gat gtt att aat atg gca cgt agt ctg cgc att			672
Arg Glu Thr Lys Gln Asp Val Ile Asn Met Ala Arg Ser Leu Arg Ile			
210	215	220	
ctg gat gca gat agc att ccg gtt aat ttt ctg cat gca att gat ggt			720
Leu Asp Ala Asp Ser Ile Pro Val Asn Phe Leu His Ala Ile Asp Gly			
225	230	235	240
acc ccg ctg gca ggt acc aat gaa ctg gat ccg cgt tat tgt ctg aaa			768
Thr Pro Leu Ala Gly Thr Asn Glu Leu Asp Pro Arg Tyr Cys Leu Lys			
245	250	255	

gtt ctg gca ctg ttt cgt tat atg aat ccg acc aaa gaa att cgt att 816

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

260 265 270

gcc ggt ggt cgt gaa gtt aat ctg cgt agc ctg cag ccg ctg ggt ctg 864

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

275 280 285

tat gca gca aat agc att ttt gtt ggt gat tat ctg acc acc gca ggt 912

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

290 295 300

cag gaa aaa agc gaa gat tat cgc atg ctg gaa gat ctg ggt ttt gaa 960

Gln Glu Lys Ser Glu Asp Tyr Arg Met Leu Glu Asp Leu Gly Phe Glu

305 310 315 320

att gat ttt gcc gaa gaa cag cag gtt gtt tgt taa 996

Ile Asp Phe Ala Glu Glu Gln Gln Val Val Cys

325 330

<210> 59

<211> 331

<212> PRT

<213> Geobacillus thermoglucosidasius C56-YS93

<400> 59

Met Ile Asn Trp Leu Ala Leu Ala Asp Arg Val Ile Ala Gly His Glu

1 5 10 15

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

20 25 30

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

35 40 45

Gly Asn Lys Val Lys Leu Asn Met Ile Ile Asn Ala Lys Ser Gly Leu

50 55 60

Cys Pro Glu Asn Cys Gly Tyr Cys Ala Gln Ser Ala Val Ser Thr Ala

65 70 75 80

Pro Val Lys Thr Tyr Lys Met Val Asp Lys Glu Thr Leu Ile Arg Gly

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

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

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

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

<210> 60

<211> 1008

<212> DNA

<213> *Bacillus subtilis* subsp. *subtilis* str. 168

<220><221> CDS

<222> (1)..(1008)

<223> bioB gene encoding biotin synthase from *Bacillus subtilis* subsp.
subtilis str. 168

<400> 60

atg aat cag tgg atg gaa ctg gca gat cgt gtt ctg gcc ggt gca gaa 48

Met Asn Gln Trp Met Glu Leu Ala Asp Arg Val Leu Ala Gly Ala Glu

1 5 10 15

gtt acc gat gaa gaa gca ctg agc att ctg cat tgc ccg gat gaa gat 96

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

20 25 30

atc ctg ctg ctg atg cat ggt gcc ttt cat att cgc aaa cat ttt tat 144

Ile Leu Leu Leu Met His Gly Ala Phe His Ile Arg Lys His Phe Tyr

35 40 45

ggc aaa aaa gtg aaa ctg aat atg att atg aac gcc aaa agc ggt ctg 192

Gly Lys Lys Val Lys Leu Asn Met Ile Met Asn Ala Lys Ser Gly Leu

50 55 60

tgt ccg gaa aat tgc ggt tat tgc agc cag agc gca att agc aaa gca 240

Cys Pro Glu Asn Cys Gly Tyr Cys Ser Gln Ser Ala Ile Ser Lys Ala

65 70 75 80

ccg att gaa agc tat cgt atg gtt aat aaa gaa acc ctg ctg gaa ggt 288

Pro Ile Glu Ser Tyr Arg Met Val Asn Lys Glu Thr Leu Leu Glu Gly

85 90 95

gca aaa cgt gca cat gat ctg aat att ggt acc tat tgt att gtt gca 336

Ala Lys Arg Ala His Asp Leu Asn Ile Gly Thr Tyr Cys Ile Val Ala

100 105 110

agc ggt cgt ggt ccg agc aat cgc gaa gtt gat cag gtt gtt gat gca 384

Ser Gly Arg Gly Pro Ser Asn Arg Glu Val Asp Gln Val Val Asp Ala

115	120	125	
ggt cag gaa att aaa gaa acc tat ggt ctg aaa att tgt gca tgt ctg			432
Val Gln Glu Ile Lys Glu Thr Tyr Gly Leu Lys Ile Cys Ala Cys Leu			
130	135	140	
ggt ctg ctg aaa ccg gaa cag gca aaa cgt ctg aaa gat gca ggc gtt			480
Gly Leu Leu Lys Pro Glu Gln Ala Lys Arg Leu Lys Asp Ala Gly Val			
145	150	155	160
gat cgt tat aat cat aat ctg aat acc agc cag cgt aac cat agc aat			528
Asp Arg Tyr Asn His Asn Leu Asn Thr Ser Gln Arg Asn His Ser Asn			
165	170	175	
att acc acc agc cat acc tat gat gat cgt gtg aat acc gtt gaa att			576
Ile Thr Thr Ser His Thr Tyr Asp Asp Arg Val Asn Thr Val Glu Ile			
180	185	190	
gcc aaa gaa agt ggt ctg agc ccg tgt agc ggt gca att att ggt atg			624
Ala Lys Glu Ser Gly Leu Ser Pro Cys Ser Gly Ala Ile Ile Gly Met			
195	200	205	
aaa gaa acc aaa cag gat gtt att gat att gcg aaa agc ctg aaa gca			672
Lys Glu Thr Lys Gln Asp Val Ile Asp Ile Ala Lys Ser Leu Lys Ala			
210	215	220	
ctg gat gca gat agc att ccg gtg aat ttt ctg cat gca att gat ggt			720
Leu Asp Ala Asp Ser Ile Pro Val Asn Phe Leu His Ala Ile Asp Gly			
225	230	235	240
acc ccg ctg gaa ggt gtt aat gaa ctg aat ccg ctg tat tgt ctg aaa			768
Thr Pro Leu Glu Gly Val Asn Glu Leu Asn Pro Leu Tyr Cys Leu Lys			
245	250	255	
ggt ctg gca ctg ttt cgt ttt att aat ccg agc aaa gaa att cgt att			816
Val Leu Ala Leu Phe Arg Phe Ile Asn Pro Ser Lys Glu Ile Arg Ile			
260	265	270	
agc ggt ggt cgt gaa gtt aat ctg cgt acc ctg cag ccg ctg ggt ctg			864

Ser Gly Gly Arg Glu Val Asn Leu Arg Thr Leu Gln Pro Leu Gly Leu

275 280 285
tat gca gca aat agc att ttt gtt ggt gat tat ctg acc acc gca ggt 912

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

290 295 300
cag gaa gaa acc gaa gat cat aaa atg ctg agc gat ctg ggt ttt gaa 960

Gln Glu Glu Thr Glu Asp His Lys Met Leu Ser Asp Leu Gly Phe Glu

305 310 315 320

gtt gaa agc gtt gaa gaa atg aaa gca agc ctg agc gca aaa agc taa 1008

Val Glu Ser Val Glu Glu Met Lys Ala Ser Leu Ser Ala Lys Ser

325 330 335

<210> 61

<211> 335

<212> PRT

<213> Bacillus subtilis subsp. subtilis str. 168

<400> 61

Met Asn Gln Trp Met Glu Leu Ala Asp Arg Val Leu Ala Gly Ala Glu

1 5 10 15

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

20 25 30

Ile Leu Leu Leu Met His Gly Ala Phe His Ile Arg Lys His Phe Tyr

35 40 45

Gly Lys Lys Val Lys Leu Asn Met Ile Met Asn Ala Lys Ser Gly Leu

50 55 60

Cys Pro Glu Asn Cys Gly Tyr Cys Ser Gln Ser Ala Ile Ser Lys Ala

65 70 75 80

Pro Ile Glu Ser Tyr Arg Met Val Asn Lys Glu Thr Leu Leu Glu Gly

85 90 95

Ala Lys Arg Ala His Asp Leu Asn Ile Gly Thr Tyr Cys Ile Val Ala

100 105 110

Ser Gly Arg Gly Pro Ser Asn Arg Glu Val Asp Gln Val Val Asp Ala

115 120 125
Val Gln Glu Ile Lys Glu Thr Tyr Gly Leu Lys Ile Cys Ala Cys Leu

130 135 140
Gly Leu Leu Lys Pro Glu Gln Ala Lys Arg Leu Lys Asp Ala Gly Val

145 150 155 160
Asp Arg Tyr Asn His Asn Leu Asn Thr Ser Gln Arg Asn His Ser Asn

165 170 175
Ile Thr Thr Ser His Thr Tyr Asp Asp Arg Val Asn Thr Val Glu Ile

180 185 190
Ala Lys Glu Ser Gly Leu Ser Pro Cys Ser Gly Ala Ile Ile Gly Met

195 200 205
Lys Glu Thr Lys Gln Asp Val Ile Asp Ile Ala Lys Ser Leu Lys Ala

210 215 220
Leu Asp Ala Asp Ser Ile Pro Val Asn Phe Leu His Ala Ile Asp Gly

225 230 235 240
Thr Pro Leu Glu Gly Val Asn Glu Leu Asn Pro Leu Tyr Cys Leu Lys

245 250 255
Val Leu Ala Leu Phe Arg Phe Ile Asn Pro Ser Lys Glu Ile Arg Ile

260 265 270
Ser Gly Gly Arg Glu Val Asn Leu Arg Thr Leu Gln Pro Leu Gly Leu

275 280 285
Tyr Ala Ala Asn Ser Ile Phe Val Gly Asp Tyr Leu Thr Thr Ala Gly

290 295 300
Gln Glu Glu Thr Glu Asp His Lys Met Leu Ser Asp Leu Gly Phe Glu

305 310 315 320
Val Glu Ser Val Glu Glu Met Lys Ala Ser Leu Ser Ala Lys Ser

325 330 335

<210> 62

<211> 996

<212> DNA

<213> Lysinibacillus sphaericus

<220><221> CDS

<222> (1)..(996)

<223> bioB gene encoding biotin synthase from Lysinibacillus sphaericus

<400> 62

atg aat ttt ctg cag gtt gcc cag gaa gtt att gat ggc aaa att att 48

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

1 5 10 15

agc aat gaa gaa gcc ctg gcc att ctg aat agc aaa gat gat gaa ctg 96

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

20 25 30

ctg cag ctg atg gat ggt gca ttt gcc att cgc cgc cat tat tat ggt 144

Leu Gln Leu Met Asp Gly Ala Phe Ala Ile Arg Arg His Tyr Tyr Gly

35 40 45

aaa aaa gtg aaa ctg aac atg att atg aac gcc aaa agc ggt tat tgc 192

Lys Lys Val Lys Leu Asn Met Ile Met Asn Ala Lys Ser Gly Tyr Cys

50 55 60

ccg gaa gat tgt ggt tat tgt agc cag agc agc aaa agc acc gca ccg 240

Pro Glu Asp Cys Gly Tyr Cys Ser Gln Ser Ser Lys Ser Thr Ala Pro

65 70 75 80

att gaa aaa tat ccg ttt att acc aaa gaa gaa att ctg gcc ggt gcc 288

Ile Glu Lys Tyr Pro Phe Ile Thr Lys Glu Glu Ile Leu Ala Gly Ala

85 90 95

aaa cgt gcc ttt gat aat aaa att ggt acc tat tgc att gtt gcc agc 336

Lys Arg Ala Phe Asp Asn Lys Ile Gly Thr Tyr Cys Ile Val Ala Ser

100 105 110

ggt cgt ggt ccg acc cgt aaa gat gtt aat gtt gtg agc gaa gca gtt 384

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

115 120 125

acc gaa att aaa gaa aaa tat ggc ctg aaa gtt tgt gca tgt ctg ggt 432

Thr Glu Ile Lys Glu Lys Tyr Gly Leu Lys Val Cys Ala Cys Leu Gly

130 135 140

ctg ctg aaa gaa gaa cag gca cag cag ctg aaa gaa gcc ggt gtt gat	480
Leu Leu Lys Glu Glu Gln Ala Gln Gln Leu Lys Glu Ala Gly Val Asp	
145 150 155 160	
cgc tat aat cat aat ctg aac acc agc gaa cgt cat cat agc ttt att	528
Arg Tyr Asn His Asn Leu Asn Thr Ser Glu Arg His His Ser Phe Ile	
165 170 175	
acc acc agc cat acc tat gaa gat cgt gtg aac acc gtg gaa att gtg	576
Thr Thr Ser His Thr Tyr Glu Asp Arg Val Asn Thr Val Glu Ile Val	
180 185 190	
aaa aaa cat ggt att agc ccg tgt agc ggt gca att att ggt atg aaa	624
Lys Lys His Gly Ile Ser Pro Cys Ser Gly Ala Ile Ile Gly Met Lys	
195 200 205	
gaa acc cgt gaa gat gtt gtt aat att gca cgt gca ctg cat cag ctg	672
Glu Thr Arg Glu Asp Val Val Asn Ile Ala Arg Ala Leu His Gln Leu	
210 215 220	
gat gca gat agc att ccg gtt aat ttt ctg aat gca att gat ggt acc	720
Asp Ala Asp Ser Ile Pro Val Asn Phe Leu Asn Ala Ile Asp Gly Thr	
225 230 235 240	
aaa ctg gaa ggt acc cgt gat ctg aat ccg cgt tat tgt ctg aaa gtg	768
Lys Leu Glu Gly Thr Arg Asp Leu Asn Pro Arg Tyr Cys Leu Lys Val	
245 250 255	
ctg gca ctg ttt cgt tat att aat ccg acc aaa gaa att cgt att agc	816
Leu Ala Leu Phe Arg Tyr Ile Asn Pro Thr Lys Glu Ile Arg Ile Ser	
260 265 270	
ggt ggt cgt gaa att aat ctg ggt agc ctg cag ccg ctg ggt ctg tat	864
Gly Gly Arg Glu Ile Asn Leu Gly Ser Leu Gln Pro Leu Gly Leu Tyr	
275 280 285	
gca gca aat agc att ttt gtt ggt gat tat ctg acc acc gca ggt cag	912
Ala Ala Asn Ser Ile Phe Val Gly Asp Tyr Leu Thr Thr Ala Gly Gln	
290 295 300	
gaa gcc aat agc gat tat cgt atg ctg gaa gat ctg ggt ttt gaa att	960

Glu Ala Asn Ser Asp Tyr Arg Met Leu Glu Asp Leu Gly Phe Glu Ile
305 310 315 320
gaa ctg acc cag aaa cag gaa gca gca ttt tgt taa 996
Glu Leu Thr Gln Lys Gln Glu Ala Ala Phe Cys
325 330
<210> 63
<211> 331
<212> PRT
<213> Lysinibacillus sphaericus
<400> 63
Met Asn Phe Leu Gln Val Ala Gln Glu Val Ile Asp Gly Lys Ile Ile
1 5 10 15
Ser Asn Glu Glu Ala Leu Ala Ile Leu Asn Ser Lys Asp Asp Glu Leu
20 25 30
Leu Gln Leu Met Asp Gly Ala Phe Ala Ile Arg Arg His Tyr Tyr Gly
35 40 45
Lys Lys Val Lys Leu Asn Met Ile Met Asn Ala Lys Ser Gly Tyr Cys
50 55 60
Pro Glu Asp Cys Gly Tyr Cys Ser Gln Ser Ser Lys Ser Thr Ala Pro
65 70 75 80
Ile Glu Lys Tyr Pro Phe Ile Thr Lys Glu Glu Ile Leu Ala Gly Ala
85 90 95
Lys Arg Ala Phe Asp Asn Lys Ile Gly Thr Tyr Cys Ile Val Ala Ser
100 105 110
Gly Arg Gly Pro Thr Arg Lys Asp Val Asn Val Val Ser Glu Ala Val
115 120 125
Thr Glu Ile Lys Glu Lys Tyr Gly Leu Lys Val Cys Ala Cys Leu Gly
130 135 140
Leu Leu Lys Glu Glu Gln Ala Gln Gln Leu Lys Glu Ala Gly Val Asp
145 150 155 160
Arg Tyr Asn His Asn Leu Asn Thr Ser Glu Arg His His Ser Phe Ile

165 170 175

Thr Thr Ser His Thr Tyr Glu Asp Arg Val Asn Thr Val Glu Ile Val

180 185 190

Lys Lys His Gly Ile Ser Pro Cys Ser Gly Ala Ile Ile Gly Met Lys

195 200 205

Glu Thr Arg Glu Asp Val Val Asn Ile Ala Arg Ala Leu His Gln Leu

210 215 220

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

225 230 235 240

Lys Leu Glu Gly Thr Arg Asp Leu Asn Pro Arg Tyr Cys Leu Lys Val

245 250 255

Leu Ala Leu Phe Arg Tyr Ile Asn Pro Thr Lys Glu Ile Arg Ile Ser

260 265 270

Gly Gly Arg Glu Ile Asn Leu Gly Ser Leu Gln Pro Leu Gly Leu Tyr

275 280 285

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

290 295 300

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

305 310 315 320

Glu Leu Thr Gln Lys Gln Glu Ala Ala Phe Cys

325 330

<210> 64

<211> 1002

<212> DNA

<213> Methylococcus capsulatus str. Bath

<220><221> CDS

<222> (1)..(1002)

<223> bioB gene encoding biotin synthase from Methylococcus capsulatus
str. Bath

<400> 64

atg cat gca gaa gtt gca gtt atg acc aat cag gaa cgt gca gaa gaa 48

Met	His	Ala	Glu	Val	Ala	Val	Met	Thr	Asn	Gln	Glu	Arg	Ala	Glu	Glu	
1				5					10					15		
ccg	gtt	ctg	cgt	cat	gat	tgg	acc	cag	ggc	gaa	gca	gaa	gca	ctg	ttt	96
Pro	Val	Leu	Arg	His	Asp	Trp	Thr	Gln	Gly	Glu	Ala	Glu	Ala	Leu	Phe	
				20				25						30		
gca	ctg	ccg	ttt	aat	gaa	ctg	ctg	ttt	cag	gca	cag	acc	att	cat	cgt	144
Ala	Leu	Pro	Phe	Asn	Glu	Leu	Leu	Phe	Gln	Ala	Gln	Thr	Ile	His	Arg	
				35				40					45			
cgt	cat	ttt	gat	ccg	aat	gaa	gtt	cag	gtt	agc	agc	ctg	ctg	agc	att	192
Arg	His	Phe	Asp	Pro	Asn	Glu	Val	Gln	Val	Ser	Ser	Leu	Leu	Ser	Ile	
				50				55					60			
aaa	acc	ggc	gca	tgt	agc	gaa	gat	tgt	gca	tat	tgt	ccg	cag	agc	gca	240
Lys	Thr	Gly	Ala	Cys	Ser	Glu	Asp	Cys	Ala	Tyr	Cys	Pro	Gln	Ser	Ala	
65						70					75				80	
cat	tat	gaa	acc	ggc	gtt	aaa	cgt	gaa	agc	ctg	atg	agc	ctg	gaa	gat	288
His	Tyr	Glu	Thr	Gly	Val	Lys	Arg	Glu	Ser	Leu	Met	Ser	Leu	Glu	Asp	
						85				90				95		
gtt	ctg	gaa	gca	gca	cag	cgt	gca	cgt	gaa	gaa	ggc	gca	acc	cgt	ttt	336
Val	Leu	Glu	Ala	Ala	Gln	Arg	Ala	Arg	Glu	Glu	Gly	Ala	Thr	Arg	Phe	
						100				105				110		
tgt	atg	ggc	gca	gca	tgg	cgt	agc	ccg	cgt	gat	ggc	gat	ctg	gaa	gca	384
Cys	Met	Gly	Ala	Ala	Trp	Arg	Ser	Pro	Arg	Asp	Gly	Asp	Leu	Glu	Ala	
						115				120				125		
att	gca	gca	atg	gtt	gaa	ggc	gtt	aaa	gca	ctg	ggc	atg	gaa	acc	tgt	432
Ile	Ala	Ala	Met	Val	Glu	Gly	Val	Lys	Ala	Leu	Gly	Met	Glu	Thr	Cys	
						130				135				140		
gtt	acc	gcc	ggc	atg	ctg	agt	gat	gaa	cag	gcc	cgt	cgt	ctg	aaa	gaa	480
Val	Thr	Ala	Gly	Met	Leu	Ser	Asp	Glu	Gln	Ala	Arg	Arg	Leu	Lys	Glu	
145						150				155					160	
gcc	ggc	ctg	gat	tat	tat	aat	cat	aat	ctg	gat	acc	agc	gaa	agc	tat	528
Ala	Gly	Leu	Asp	Tyr	Tyr	Asn	His	Asn	Leu	Asp	Thr	Ser	Glu	Ser	Tyr	

165	170	175	
tat ggc gaa att att acc acc cgc acc tat cag gat cgt ctg gat acc			576
Tyr Gly Glu Ile Ile Thr Thr Arg Thr Tyr Gln Asp Arg Leu Asp Thr			
180	185	190	
ctg cag cgt gtt cgt gat gca ggt atg cat gtt tgt tgt ggt ggt att			624
Leu Gln Arg Val Arg Asp Ala Gly Met His Val Cys Cys Gly Gly Ile			
195	200	205	
gtt ggt atg ggt gaa agc gca gca gat cgt gca ggt ctg ctg att ggt			672
Val Gly Met Gly Glu Ser Ala Ala Asp Arg Ala Gly Leu Leu Ile Gly			
210	215	220	
ctg gca aat ctg ccg cgt cat ccg gaa agc gtt ccg att aat ctg ctg			720
Leu Ala Asn Leu Pro Arg His Pro Glu Ser Val Pro Ile Asn Leu Leu			
225	230	235	240
gtt cgt gtt gaa ggt acc ccg ctg gca gat acc gca gca ctg gat ccg			768
Val Arg Val Glu Gly Thr Pro Leu Ala Asp Thr Ala Ala Leu Asp Pro			
245	250	255	
ttt gat ttt gtt cgt acc gtt gca gtt gca cgt att atg atg ccg gca			816
Phe Asp Phe Val Arg Thr Val Ala Val Ala Arg Ile Met Met Pro Ala			
260	265	270	
agc cgt gtt cgt ctg agc gca ggt cgt agc gat atg agc gat gaa atg			864
Ser Arg Val Arg Leu Ser Ala Gly Arg Ser Asp Met Ser Asp Glu Met			
275	280	285	
cag gca ctg tgt ttt ctg gcc ggt gca aat agc att ttt tat ggt gat			912
Gln Ala Leu Cys Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr Gly Asp			
290	295	300	
cgt ctg ctg acc acc gaa aat ccg cag gcc cag cgt gat cgc cgt ctg			960
Arg Leu Leu Thr Thr Glu Asn Pro Gln Ala Gln Arg Asp Arg Arg Leu			
305	310	315	320
ttt gcc cgt ctg ggt ctg cgt atg gca ggt ctg ggt tgt taa			1002
Phe Ala Arg Leu Gly Leu Arg Met Ala Gly Leu Gly Cys			
325	330		

<210> 65

<211> 333

<212> PRT

<213> *Methylococcus capsulatus* str. Bath

<400> 65

Met His Ala Glu Val Ala Val Met Thr Asn Gln Glu Arg Ala Glu Glu

1 5 10 15

Pro Val Leu Arg His Asp Trp Thr Gln Gly Glu Ala Glu Ala Leu Phe

20 25 30

Ala Leu Pro Phe Asn Glu Leu Leu Phe Gln Ala Gln Thr Ile His Arg

35 40 45

Arg His Phe Asp Pro Asn Glu Val Gln Val Ser Ser Leu Leu Ser Ile

50 55 60

Lys Thr Gly Ala Cys Ser Glu Asp Cys Ala Tyr Cys Pro Gln Ser Ala

65 70 75 80

His Tyr Glu Thr Gly Val Lys Arg Glu Ser Leu Met Ser Leu Glu Asp

85 90 95

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

100 105 110

Cys Met Gly Ala Ala Trp Arg Ser Pro Arg Asp Gly Asp Leu Glu Ala

115 120 125

Ile Ala Ala Met Val Glu Gly Val Lys Ala Leu Gly Met Glu Thr Cys

130 135 140

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

145 150 155 160

Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Glu Ser Tyr

165 170 175

Tyr Gly Glu Ile Ile Thr Thr Arg Thr Tyr Gln Asp Arg Leu Asp Thr

180 185 190

Leu Gln Arg Val Arg Asp Ala Gly Met His Val Cys Cys Gly Gly Ile

195 200 205

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

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

Phe Ala Arg Leu Gly Leu Arg Met Ala Gly Leu Gly Cys
 325 330

<210> 66

<211> 1041

<212> DNA

<213> Leclercia adecarboxylata

<220><221> CDS

<222> (1)..(1041)

<223> bioB gene encoding biotin synthase from Leclercia adecarboxylata

<400> 66

atg gca cat cag acc cgt tgg acc ctg agc cag gtt acc gca ctg ttt 48
 Met Ala His Gln Thr Arg Trp Thr Leu Ser Gln Val Thr Ala Leu Phe
 1 5 10 15

gaa aaa ccg ctg ctg gaa ctg ctg ttt gaa gca cag cag att cat cgt 96
 Glu Lys Pro Leu Leu Glu Leu Leu Phe Glu Ala Gln Gln Ile His Arg
 20 25 30

cag cat ttt gat ccg cag cag att cag gtt agc acc ctg ctg agc att 144
 Gln His Phe Asp Pro Gln Gln Ile Gln Val Ser Thr Leu Leu Ser Ile

35	40	45	
aaa acc ggt gca tgt ccg gaa gat tgt aaa tat tgt ccg cag agc gca			192
Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys Tyr Cys Pro Gln Ser Ala			
50	55	60	
cgt tat aaa acc ggt ctg gaa tca gaa cgt ctg atg gaa gtt gaa cag			240
Arg Tyr Lys Thr Gly Leu Glu Ser Glu Arg Leu Met Glu Val Glu Gln			
65	70	75	80
gtt ctg gaa agc gca cgt cag gca aaa aat gca ggt agc acc cgt ttt			288
Val Leu Glu Ser Ala Arg Gln Ala Lys Asn Ala Gly Ser Thr Arg Phe			
85	90	95	
tgt atg ggt gca gca tgg aaa aat ccg cat gaa cgt gat atg ccg tat			336
Cys Met Gly Ala Ala Trp Lys Asn Pro His Glu Arg Asp Met Pro Tyr			
100	105	110	
ctg gaa cag atg gtt cag ggt gtt aaa gca atg ggt ctg gaa gca tgt			384
Leu Glu Gln Met Val Gln Gly Val Lys Ala Met Gly Leu Glu Ala Cys			
115	120	125	
atg acc ctg ggt acc ctg gat gat acc cag gca cag cgt ctg gca agc			432
Met Thr Leu Gly Thr Leu Asp Asp Thr Gln Ala Gln Arg Leu Ala Ser			
130	135	140	
gca ggt ctg gat tat tat aat cat aat ctg gat acc agc ccg gaa ttt			480
Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu Phe			
145	150	155	160
tat ggc aat att att acc acc cgc acc tat cag gaa cgc ctg gat acc			528
Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr Gln Glu Arg Leu Asp Thr			
165	170	175	
ctg gat aaa gtt cgt gat gca ggt att aaa gtt tgt agc ggt ggt att			576
Leu Asp Lys Val Arg Asp Ala Gly Ile Lys Val Cys Ser Gly Gly Ile			
180	185	190	
gtt ggt ctg ggt gaa acc gtt acc gat cgt gca ggt ctg ctg ctg cag			624

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

195 200 205

ctg gca aat ctg ccg acc ccg ccg gaa agc gtt ccg att aat atg ctg 672

Leu Ala Asn Leu Pro Thr Pro Pro Glu Ser Val Pro Ile Asn Met Leu

210 215 220

gtt aaa gtt aaa ggt acc ccg ctg gcc gat aat gat gat gtt gat gca 720

Val Lys Val Lys Gly Thr Pro Leu Ala Asp Asn Asp Asp Val Asp Ala

225 230 235 240

ttt gat ttt att cgt acc att gca gtg gcc cgt gtg atg atg ccg acc 768

Phe Asp Phe Ile Arg Thr Ile Ala Val Ala Arg Val Met Met Pro Thr

245 250 255

agc ttt gtt cgt ctg agc gcc ggt cgt gaa cag atg aat gaa cag acc 816

Ser Phe Val Arg Leu Ser Ala Gly Arg Glu Gln Met Asn Glu Gln Thr

260 265 270

cag gcc atg tgt ttt atg gcc ggt gca aat agc att ttt tat ggt tgt 864

Gln Ala Met Cys Phe Met Ala Gly Ala Asn Ser Ile Phe Tyr Gly Cys

275 280 285

aaa ctg ctg acc acc ccg aat ccg gaa gaa gat aaa gat gtt cag ctg 912

Lys Leu Leu Thr Thr Pro Asn Pro Glu Glu Asp Lys Asp Val Gln Leu

290 295 300

ttt cgt aaa ctg ggt ctg aat ccg cag cag acc gca gtt ctg acc ggt 960

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

305 310 315 320

gat aat gaa cag cag cat cag ctg gaa cag cag ctg att aat gca gat 1008

Asp Asn Glu Gln Gln His Gln Leu Glu Gln Gln Leu Ile Asn Ala Asp

325 330 335

acc gat cag ttt tat aat gca gca acc gtt taa 1041

Thr Asp Gln Phe Tyr Asn Ala Ala Thr Val

340 345

<210> 67

<211> 346

<212> PRT

<213> Leclercia adecarboxylata

<400> 67

Met Ala His Gln Thr Arg Trp Thr Leu Ser Gln Val Thr Ala Leu Phe

1 5 10 15

Glu Lys Pro Leu Leu Glu Leu Leu Phe Glu Ala Gln Gln Ile His Arg

20 25 30

Gln His Phe Asp Pro Gln Gln Ile Gln Val Ser Thr Leu Leu Ser Ile

35 40 45

Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys Tyr Cys Pro Gln Ser Ala

50 55 60

Arg Tyr Lys Thr Gly Leu Glu Ser Glu Arg Leu Met Glu Val Glu Gln

65 70 75 80

Val Leu Glu Ser Ala Arg Gln Ala Lys Asn Ala Gly Ser Thr Arg Phe

85 90 95

Cys Met Gly Ala Ala Trp Lys Asn Pro His Glu Arg Asp Met Pro Tyr

100 105 110

Leu Glu Gln Met Val Gln Gly Val Lys Ala Met Gly Leu Glu Ala Cys

115 120 125

Met Thr Leu Gly Thr Leu Asp Asp Thr Gln Ala Gln Arg Leu Ala Ser

130 135 140

Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro Glu Phe

145 150 155 160

Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr Gln Glu Arg Leu Asp Thr

165 170 175

Leu Asp Lys Val Arg Asp Ala Gly Ile Lys Val Cys Ser Gly Gly Ile

180 185 190

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

195 200 205

Leu Ala Asn Leu Pro Thr Pro Pro Glu Ser Val Pro Ile Asn Met Leu

210 215 220
Val Lys Val Lys Gly Thr Pro Leu Ala Asp Asn Asp Asp Val Asp Ala
225 230 235 240
Phe Asp Phe Ile Arg Thr Ile Ala Val Ala Arg Val Met Met Pro Thr

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

305 310 315 320
Asp Asn Glu Gln Gln His Gln Leu Glu Gln Gln Leu Ile Asn Ala Asp
325 330 335
Thr Asp Gln Phe Tyr Asn Ala Ala Thr Val
340 345

<210> 68

<211> 1113

<212> DNA

<213> Chromohalobacter salexigens DSM 3043

<220><221> CDS

<222> (1)..(1113)

<223> bioB gene encoding biotin synthase from Chromohalobacter
salexigens DSM 3043

<400> 68

atg acc gca cag agc cgt gat ccg gca tgg acc gat gca agc ccg acc 48
Met Thr Ala Gln Ser Arg Asp Pro Ala Trp Thr Asp Ala Ser Pro Thr
1 5 10 15
ttt cag ccg acc atg cgt cat gat tgg agc ctg gaa gaa att gaa gca 96
Phe Gln Pro Thr Met Arg His Asp Trp Ser Leu Glu Glu Ile Glu Ala
20 25 30

ctg ttt gca ctg ccg ttt aat gat ctg ctg ttt cgt gca cag cag gtt 144

Leu Phe Ala Leu Pro Phe Asn Asp Leu Leu Phe Arg Ala Gln Gln Val
35 40 45

cat cgt gca cat ttt gat ccg aat gca gtt cag gtt agc acc ctg ctg 192

His Arg Ala His Phe Asp Pro Asn Ala Val Gln Val Ser Thr Leu Leu
50 55 60

agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa tat tgt ccg cag 240

Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys Tyr Cys Pro Gln

65 70 75 80

agc ggt cat tat aat acc ggc ctg ggt aaa gaa aaa ctg ctg gaa att 288

Ser Gly His Tyr Asn Thr Gly Leu Gly Lys Glu Lys Leu Leu Glu Ile
85 90 95

gaa aaa gtt gtt gaa cag gcc cgt gca gca aaa gca gca ggc gca agc 336

Glu Lys Val Val Glu Gln Ala Arg Ala Ala Lys Ala Ala Gly Ala Ser
100 105 110

cgt ttt tgt atg ggt gca gca tgg cgt agc ccg cgt gaa aaa gat ctg 384

Arg Phe Cys Met Gly Ala Ala Trp Arg Ser Pro Arg Glu Lys Asp Leu
115 120 125

cgt gtt gtt acc gaa atg gtt ggt cgt gtt aaa gca ctg ggc ctg gaa 432

Arg Val Val Thr Glu Met Val Gly Arg Val Lys Ala Leu Gly Leu Glu
130 135 140

acc tgt atg acc ctg ggt atg gtt gat gtt gat cag gca cgt cgt ctg 480

Thr Cys Met Thr Leu Gly Met Val Asp Val Asp Gln Ala Arg Arg Leu
145 150 155 160

gca gaa gcc ggt ctg gat tat tat aat cat aat ctg gat acc agc ccg 528

Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro
165 170 175

gat tat tat ggt gaa att att acc acc cgt acc tat gca gat cgt ctg 576

Asp Tyr Tyr Gly Glu Ile Ile Thr Thr Arg Thr Tyr Ala Asp Arg Leu

180	185	190	
gaa acc ctg gcc aat gtt cgt gaa gca ggt atg aaa gtt tgt agt ggt			624
Glu Thr Leu Ala Asn Val Arg Glu Ala Gly Met Lys Val Cys Ser Gly			
195	200	205	
ggt att ctg ggt atg ggt gaa gca ccg cgt gat cgc gca gca ctg ctg			672
Gly Ile Leu Gly Met Gly Glu Ala Pro Arg Asp Arg Ala Ala Leu Leu			
210	215	220	
cag cag ctg gtt cgt ctg gat ccg cat ccg gaa agc gtt ccg att aat			720
Gln Gln Leu Val Arg Leu Asp Pro His Pro Glu Ser Val Pro Ile Asn			
225	230	235	240
atg ctg gtt aaa gtt ccg ggt acc ccg atg gaa aat gtt gaa gat atg			768
Met Leu Val Lys Val Pro Gly Thr Pro Met Glu Asn Val Glu Asp Met			
245	250	255	
gat ccg ctg acc ttt att cgt gca att gca gtt gca cgt att ctg atg			816
Asp Pro Leu Thr Phe Ile Arg Ala Ile Ala Val Ala Arg Ile Leu Met			
260	265	270	
ccg aaa agc cat gtt cgt ctg agc gca ggt cgt gaa cag atg gat gaa			864
Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu Gln Met Asp Glu			
275	280	285	
agc acc cag gca ctg gca ttt ctg gcc ggt gca aat agc att ttt tat			912
Ser Thr Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr			
290	295	300	
ggt gat acc ctg ctg acc acc ggt aat ccg cag gtt gaa cgt gat cgt			960
Gly Asp Thr Leu Leu Thr Thr Gly Asn Pro Gln Val Glu Arg Asp Arg			
305	310	315	320
gcc ctg ttt gat aaa ctg ggt ctg cat ccg gaa ccg agc gat ccg cat			1008
Ala Leu Phe Asp Lys Leu Gly Leu His Pro Glu Pro Ser Asp Pro His			
325	330	335	
gca gat gat gcc cat cgt gat gat gaa cag gca gaa att gca ctg gca			1056
Ala Asp Asp Ala His Arg Asp Asp Glu Gln Ala Glu Ile Ala Leu Ala			
340	345	350	

cat gca att cag cgt cag cgt gat gat gca ctg ttt tat gat gca acc 1104
 His Ala Ile Gln Arg Gln Arg Asp Asp Ala Leu Phe Tyr Asp Ala Thr
 355 360 365
 cgt ggt taa 1113

Arg Gly

370

<210> 69

<211> 370

<212> PRT

<213> Chromohalobacter salexigens DSM 3043

<400> 69

Met Thr Ala Gln Ser Arg Asp Pro Ala Trp Thr Asp Ala Ser Pro Thr
 1 5 10 15
 Phe Gln Pro Thr Met Arg His Asp Trp Ser Leu Glu Glu Ile Glu Ala
 20 25 30

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

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

Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu Asp Thr Ser Pro
165 170 175

Asp Tyr Tyr Gly Glu Ile Ile Thr Thr Arg Thr Tyr Ala Asp Arg Leu
180 185 190

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

Gly Ile Leu Gly Met Gly Glu Ala Pro Arg Asp Arg Ala Ala Leu Leu
210 215 220

Gln Gln Leu Val Arg Leu Asp Pro His Pro Glu Ser Val Pro Ile Asn
225 230 235 240

Met Leu Val Lys Val Pro Gly Thr Pro Met Glu Asn Val Glu Asp Met
245 250 255

Asp Pro Leu Thr Phe Ile Arg Ala Ile Ala Val Ala Arg Ile Leu Met
260 265 270

Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu Gln Met Asp Glu
275 280 285

Ser Thr Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn Ser Ile Phe Tyr
290 295 300

Gly Asp Thr Leu Leu Thr Thr Gly Asn Pro Gln Val Glu Arg Asp Arg
305 310 315 320

Ala Leu Phe Asp Lys Leu Gly Leu His Pro Glu Pro Ser Asp Pro His
325 330 335

Ala Asp Asp Ala His Arg Asp Asp Glu Gln Ala Glu Ile Ala Leu Ala
340 345 350

His Ala Ile Gln Arg Gln Arg Asp Asp Ala Leu Phe Tyr Asp Ala Thr
355 360 365

Arg Gly
370

<210> 70

<211> 1059

<212> DNA

<213> Pseudomonas caeni

<220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from Pseudomonas caeni

<400> 70

atg acc acc agc ccg cat gca gat acc cgt cat gat tgg acc ctg gca 48

Met Thr Thr Ser Pro His Ala Asp Thr Arg His Asp Trp Thr Leu Ala

1 5 10 15

gaa gtt acc gca ctg ctg cag cag ccg ttt aat gat ctg att ttt cag 96

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

20 25 30

gca cag agc gtt cat cgt cag cat ttt aat gca aat cgt gtt cag gtt 144

Ala Gln Ser Val His Arg Gln His Phe Asn Ala Asn Arg Val Gln Val

35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa 192

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

50 55 60

tat tgt ccg cag agc ggt cat tat aat acc ggt ctg gat aaa gaa aaa 240

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Asp Lys Glu Lys

65 70 75 80

ctg atg gaa gtg cag aaa gtt ctg gat gaa gcc aaa cgt gcc aaa gaa 288

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

85 90 95

att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc 336

Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser

100 105 110

gca aaa gat ctg ccg tat gtt ctg gaa atg gtt aaa ggt gtt aaa gca 384

Ala Lys Asp Leu Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Ala

115 120 125

atg ggt atg gaa acc tgt atg acc ctg ggt aaa ctg gat gaa gca cag 432

Met Gly Met Glu Thr Cys Met Thr Leu Gly Lys Leu Asp Glu Ala Gln

130 135 140

acc aaa gcg ctg gca gat gcg ggt ctg gat tat tat aat cat aat ctg 480
Thr Lys Ala Leu Ala Asp Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160
gat acc agc ccg gaa ttt tat ggt aat att att acc acc cgt acc tat 528
Asp Thr Ser Pro Glu Phe Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr
165 170 175
gca gaa cgt ctg cag acc ctg agc tat gtt cgt gat gca ggt atg aaa 576
Ala Glu Arg Leu Gln Thr Leu Ser Tyr Val Arg Asp Ala Gly Met Lys
180 185 190
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc att gca gat cgt 624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Ile Ala Asp Arg
195 200 205
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc 672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220
gtt ccg att aat atg ctg gtt aaa gtt gaa ggt acc ccg ctg gaa aat 720
Val Pro Ile Asn Met Leu Val Lys Val Glu Gly Thr Pro Leu Glu Asn
225 230 235 240
gca gaa gat gtt gat ccg ttt gat ttt att cgt atg ctg gca gtt gca 768
Ala Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Met Leu Ala Val Ala
245 250 255
cgc att atg atg ccg aaa agc cat gtt cgt ctg agc gca ggt cgt gaa 816
Arg Ile Met Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270
cag atg aat gaa cag atg cag agc ctg gca ttt ctg gcc ggt gca aat 864
Gln Met Asn Glu Gln Met Gln Ser Leu Ala Phe Leu Ala Gly Ala Asn
275 280 285
agc att ttt tat ggt gaa aaa ctg ctg acc acc gca aat ccg cag gca 912
Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala
290 295 300
gat aaa gat atg cag ctg ttt gca cgt ctg ggt att aaa ccg gaa gca 960

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

305	310	315	320	
cgt gaa gaa tat gca gat gaa gtt cat cag gca gca att gaa cat gca	1008			
Arg Glu Glu Tyr Ala Asp Glu Val His Gln Ala Ala Ile Glu His Ala				
325	330	335		
att att gaa cag cgt gat gcc agc ctg ttt tat gat gca gca acc cat	1056			
Ile Ile Glu Gln Arg Asp Ala Ser Leu Phe Tyr Asp Ala Ala Thr His				
340	345	350		
taa				1059

<210> 71

<211> 352

<212> PRT

<213> Pseudomonas caeni

<400> 71

Met Thr Thr Ser Pro His Ala Asp Thr Arg His Asp Trp Thr Leu Ala

1	5	10	15

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

20	25	30

Ala Gln Ser Val His Arg Gln His Phe Asn Ala Asn Arg Val Gln Val

35	40	45

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

50	55	60

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Asp Lys Glu Lys

65	70	75	80

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

85	90	95

Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser

100	105	110

Ala Lys Asp Leu Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Ala

115	120	125

Met Gly Met Glu Thr Cys Met Thr Leu Gly Lys Leu Asp Glu Ala Gln

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

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

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

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

 <210> 72
 <211> 1059
 <212> DNA
 <213> Pseudomonas monteilii
 <220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from *Pseudomonas monteilii*

<400> 72

atg agc gca agc acc att gca acc acc cgt cat gat tgg acc ctg gcc	48
Met Ser Ala Ser Thr Ile Ala Thr Thr Arg His Asp Trp Thr Leu Ala	
1 5 10 15	
gaa gtt cgt gca ctg ttt cag cag ccg ttt aat gat ctg ctg ttt cag	96
Glu Val Arg Ala Leu Phe Gln Gln Pro Phe Asn Asp Leu Leu Phe Gln	
20 25 30	
gca cag agc gtt cat cgt gca cat ttt gat gca aat cgt gtt cag gtt	144

Ala Gln Ser Val His Arg Ala His Phe Asp Ala Asn Arg Val Gln Val	
35 40 45	
agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa	192
Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys	
50 55 60	
tat tgc ccg cag agc ggt cat tat aat acc ggt ctg gaa aaa cag aaa	240
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Gln Lys	

65 70 75 80	
ctg atg gaa gtt cag aaa gtt ctg gaa gaa gca gca cgt gcc aaa gca	288
Leu Met Glu Val Gln Lys Val Leu Glu Glu Ala Ala Arg Ala Lys Ala	
85 90 95	
att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc	336
Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser	
100 105 110	

gca aaa gat atg ccg tat gtt ctg cag atg gtt cag ggt gtt aaa gca	384
Ala Lys Asp Met Pro Tyr Val Leu Gln Met Val Gln Gly Val Lys Ala	
115 120 125	
atg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg gat cgc gaa cag	432
Met Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Arg Glu Gln	
130 135 140	

acc gca gca ctg gca gaa gca ggt ctg gat tat tat aat cat aat ctg 480

Thr Ala Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160

gat acc agc ccg gaa ttt tat ggt agc att att acc acc cgt acc tat 528

Asp Thr Ser Pro Glu Phe Tyr Gly Ser Ile Ile Thr Thr Arg Thr Tyr
165 170 175

gca gaa cgt ctg cag acc ctg gca tat gtt cgt gat gca ggt atg aaa 576

Ala Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ala Gly Met Lys

180 185 190

att tgt agc ggt ggt att ctg ggt atg ggt gaa agc ctg gat gat cgt 624

Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205

gca aat ctg ctg att cag ctg gcc aat ctg ccg gaa cat ccg gaa agc 672

Ala Asn Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220

gtt ccg att aat atg ctg gtt aaa gtt gcc ggt acc ccg ctg gca aat 720

Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Ala Asn
225 230 235 240

gaa gaa gat gtt gat ccg ttt gat ttt att cgt atg ctg gca gtt gca 768

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

cgt att ctg atg ccg cag agc cat gtt cgt ctg agc gca ggt cgt gaa 816

Arg Ile Leu Met Pro Gln Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270

cag atg aat gaa cag atg cag gca ctg gca ttt ctg gcc ggt gca aat 864

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn
275 280 285

agc att ttt tat ggt gaa aaa ctg ctg acc acc ggt aat ccg cag gca 912

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Gly Asn Pro Gln Ala

290	295	300	
gat cgt gat atg cag ctg ttt gca cgt ctg ggt att cag ccg gaa gcc			960
Asp Arg Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Gln Pro Glu Ala			
305	310	315	320
ggt gaa ggt cat gca gat gaa gtt cat cag gca gca att gaa cag gca			1008
Gly Glu Gly His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala			
	325	330	335
gtt att gaa cag cgt aat ggt gaa ctg ttt tat gat gca gtt agc gca			1056
Val Ile Glu Gln Arg Asn Gly Glu Leu Phe Tyr Asp Ala Val Ser Ala			
	340	345	350
taa			1059
<210> 73			
<211> 352			
<212> PRT			
<213> Pseudomonas monteilii			
<400> 73			
Met Ser Ala Ser Thr Ile Ala Thr Thr Arg His Asp Trp Thr Leu Ala			
1	5	10	15
Glu Val Arg Ala Leu Phe Gln Gln Pro Phe Asn Asp Leu Leu Phe Gln			
	20	25	30
Ala Gln Ser Val His Arg Ala His Phe Asp Ala Asn Arg Val Gln Val			
	35	40	45
Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys			
	50	55	60
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Gln Lys			
65	70	75	80
Leu Met Glu Val Gln Lys Val Leu Glu Glu Ala Ala Arg Ala Lys Ala			
	85	90	95
Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser			
	100	105	110
Ala Lys Asp Met Pro Tyr Val Leu Gln Met Val Gln Gly Val Lys Ala			
	115	120	125

Met Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Arg Glu Gln
130 135 140

Thr Ala Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160

Asp Thr Ser Pro Glu Phe Tyr Gly Ser Ile Ile Thr Thr Arg Thr Tyr
165 170 175

Ala Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ala Gly Met Lys
180 185 190

Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205

Ala Asn Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220

Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Ala Asn
225 230 235 240

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

Arg Ile Leu Met Pro Gln Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn
275 280 285

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Gly Asn Pro Gln Ala
290 295 300

Asp Arg Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Gln Pro Glu Ala
305 310 315 320

Gly Glu Gly His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
325 330 335

Val Ile Glu Gln Arg Asn Gly Glu Leu Phe Tyr Asp Ala Val Ser Ala
340 345 350

<210> 74

<211> 1059

<212> DNA

<213> Pseudomonas massiliensis CB1

<220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from Pseudomonas massiliensis
CB1

<400> 74

atg agc gca agc ctg aat agc ccg ctg cgt cat gat tgg acc ctg agc 48

Met Ser Ala Ser Leu Asn Ser Pro Leu Arg His Asp Trp Thr Leu Ser

1 5 10 15

gaa gtt aaa gcc ctg ttt acc cag ccg ttt aat gat ctg ctg ttt cat 96

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

20 25 30

gca atg agc gtt cat cgt gca cat ttt gat ccg aat cag gtt cag gtt 144

Ala Met Ser Val His Arg Ala His Phe Asp Pro Asn Gln Val Gln Val

35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa 192

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

50 55 60

tat tgt ccg cag agc ggt cat tat aat acc ggc ctg gaa aaa gaa aaa 240

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Glu Lys

65 70 75 80

ctg ctg gaa gtg cag aaa gtg att gaa gaa gca gca cgt gcc aaa gca 288

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

85 90 95

att ggt agc acc cgt ttt tgt atg ggt gca gca tgg aaa cat ccg agc 336

Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser

100 105 110

gca aaa gat atg ccg tat gtt ctg gaa atg gtt cgt ggt gtt aaa gca 384

Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Arg Gly Val Lys Ala

115 120 125

ctg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg gat cgt gat cag 432

Leu Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Arg Asp Gln
130 135 140
acc gtt gca ctg gca gaa gca ggt ctg gat tat tat aat cat aat ctg 480
Thr Val Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160
gat acc agc ccg gaa ttt tat ggc aat att att acc acc cgt acc tat 528
Asp Thr Ser Pro Glu Phe Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr
165 170 175
ggt gaa cgt ctg cag acc ctg gcc tat gtt cgt gat gca ggt atg aaa 576
Gly Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ala Gly Met Lys
180 185 190
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc ctg gat gat cgt 624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc 672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220
gtt ccg att aat atg ctg gtt aaa gtt gcc ggt acc ccg ctg gaa aat 720
Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Glu Asn
225 230 235 240
gca gaa gat gtt gat ccg ttt gat ttt att cgt atg ctg gca gtt gca 768
Ala Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Met Leu Ala Val Ala
245 250 255
cgt att ctg atg ccg cgt agc cat gtt cgt ctg agc gca ggt cgt gaa 816
Arg Ile Leu Met Pro Arg Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270
cag atg aat gaa cag atg cag gcc ctg gca ttt atg gcc ggt gca aat 864
Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn
275 280 285
agc att ttt tat ggt gaa aaa ctg ctg acc acc gca aat ccg cag gca 912
Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala

290 295 300
gat aaa gat atg cgt ctg ttt gca cgt ctg ggt att cgt ccg gaa gca 960
Asp Lys Asp Met Arg Leu Phe Ala Arg Leu Gly Ile Arg Pro Glu Ala

305 310 315 320
cgt gaa gaa cat gat gat gaa gtt cat cag gca gca att gaa cag gca 1008
Arg Glu Glu His Asp Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala

325 330 335
ctg gtt gaa cag cgt agc ggt gaa ctg ttt tat gat gca gca gca gtt 1056
Leu Val Glu Gln Arg Ser Gly Glu Leu Phe Tyr Asp Ala Ala Ala Val

340 345 350
taa 1059
<210> 75
<211> 352
<212> PRT
<213> Pseudomonas massiliensis CB1
<400> 75
Met Ser Ala Ser Leu Asn Ser Pro Leu Arg His Asp Trp Thr Leu Ser
1 5 10 15
Glu Val Lys Ala Leu Phe Thr Gln Pro Phe Asn Asp Leu Leu Phe His
20 25 30
Ala Met Ser Val His Arg Ala His Phe Asp Pro Asn Gln Val Gln Val
35 40 45
Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys
50 55 60
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Glu Lys
65 70 75 80
Leu Leu Glu Val Gln Lys Val Ile Glu Glu Ala Ala Arg Ala Lys Ala
85 90 95
Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser
100 105 110
Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Arg Gly Val Lys Ala

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

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

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

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

<210> 76

<211> 1059

<212> DNA

<213> Pseudomonas putida F1

<220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from Pseudomonas putida F1

<400> 76

atg agc gca agc acc acc gca acc acc cgt cat gat tgg agc ctg gca 48

Met Ser Ala Ser Thr Thr Ala Thr Thr Arg His Asp Trp Ser Leu Ala

1 5 10 15

gaa gtt aaa gcc ctg ttt cag cag ccg ttt aat gat ctg ctg ttt cag 96

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

20 25 30

gca cag acc gtt cat cgt gca cat ttt aat ccg aat cgt gtt cag gtt 144

Ala Gln Thr Val His Arg Ala His Phe Asn Pro Asn Arg Val Gln Val

35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgc ccg gaa gat tgt aaa 192

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

50 55 60

tat tgt ccg cag agc ggt cat tat aat acc ggt ctg gaa aaa cag aaa 240

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Gln Lys

65 70 75 80

ctg atg gaa gtt cag aaa gtt ctg gaa gaa gca gcc cgt gca aaa gca 288

Leu Met Glu Val Gln Lys Val Leu Glu Glu Ala Ala Arg Ala Lys Ala

85 90 95

att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc 336

Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser

100 105 110

gca aaa gat atg ccg tat gtt ctg gaa atg gtt aaa ggt gtt aaa gca 384

Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Ala

115 120 125

atg ggc ctg gaa acc tgt atg acc ctg ggt aaa ctg gat cag gaa cag 432

Met Gly Leu Glu Thr Cys Met Thr Leu Gly Lys Leu Asp Gln Glu Gln

130	135	140	
acc aaa gca ctg gca cat gcg ggt ctg gat tat tat aat cat aat ctg			480
Thr Lys Ala Leu Ala His Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu			
145	150	155	160
gat acc agc ccg gaa ttt tat ggc agc att att acc acc cgt acc tat			528
Asp Thr Ser Pro Glu Phe Tyr Gly Ser Ile Ile Thr Thr Arg Thr Tyr			
165	170	175	
agc gaa cgt ctg cag acc ctg gca tat gtt cgt gat gca ggt atg aaa			576
Ser Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ala Gly Met Lys			
180	185	190	
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc ctg gat gat cgt			624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg			
195	200	205	
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc			672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser			
210	215	220	
gtt ccg att aat atg ctg gtt aaa gtt gca ggt acc ccg ctg gcc gaa			720
Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Ala Glu			
225	230	235	240
gaa gaa gat gtt gat ccg ttt gat ttt att cgt atg ctg gca gtt gca			768
Glu Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Met Leu Ala Val Ala			
245	250	255	
cgt att ctg atg ccg aaa agc cat gtt cgt ctg agc gca ggt cgt gaa			816
Arg Ile Leu Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu			
260	265	270	
cag atg aat gaa cag atg cag gcg ctg gca ttt atg gcc ggt gca aat			864
Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn			
275	280	285	
agc att ttt tat ggt gaa aaa ctg ctg acc acc gcc aat ccg cag gca			912

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala

290	295	300	
gat aaa gat atg cag ctg ttt gca cgt ctg ggt att aaa ccg gaa gca			960
Asp Lys Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Lys Pro Glu Ala			
305	310	315	320
cgt gaa gaa cat gca gat gaa gtt cat cag gca gca att gaa cag gca			1008
Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala			
325	330	335	
ctg gtt gaa cag cgt agc agc gaa atg ttt tat gat gca gca acc gca			
Leu Val Glu Gln Arg Ser Ser Glu Met Phe Tyr Asp Ala Ala Thr Ala			
340	345	350	
taa			1059

<210> 77

<211> 352

<212> PRT

<213> Pseudomonas putida F1

<400> 77

Met	Ser	Ala	Ser	Thr	Thr	Ala	Thr	Thr	Arg	His	Asp	Trp	Ser	Leu	Ala
1			5						10					15	

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

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

Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Ala

115 120 125

Met Gly Leu Glu Thr Cys Met Thr Leu Gly Lys Leu Asp Gln Glu Gln

130 135 140

Thr Lys Ala Leu Ala His Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu

145 150 155 160

Asp Thr Ser Pro Glu Phe Tyr Gly Ser Ile Ile Thr Thr Arg Thr Tyr

165 170 175

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

180 185 190

Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg

195 200 205

Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser

210 215 220

Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Ala Glu

225 230 235 240

Glu Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Met Leu Ala Val Ala

245 250 255

Arg Ile Leu Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu

260 265 270

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn

275 280 285

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala

290 295 300

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

305 310 315 320

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

325 330 335

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

340 345 350

<210> 78

<211> 1059

<212> DNA

<213> Pseudomonas thermotolerans

<220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from Pseudomonas
thermotolerans

<400> 78

atg aat gca agc gtt gca gca gca att cgt cat gat tgg acc ctg gca 48

Met Asn Ala Ser Val Ala Ala Ala Ile Arg His Asp Trp Thr Leu Ala

1 5 10 15

gaa gtt aaa gcc ctg ttt gca ctg ccg ttt aat gat ctg ctg tat cag 96

Glu Val Lys Ala Leu Phe Ala Leu Pro Phe Asn Asp Leu Leu Tyr Gln

20 25 30

gca cag acc gtt cat cgt cag tat ttt gat gca aat cgt gtt cag gtt 144

Ala Gln Thr Val His Arg Gln Tyr Phe Asp Ala Asn Arg Val Gln Val

35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa 192

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

50 55 60

tat tgt ccg cag agc ggt cat tat aat acc ggc ctg gaa aaa cag aaa 240

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Gln Lys

65 70 75 80

ctg atg gaa gtt cag aaa gtt ctg cag gca gca gca gaa gca aaa gca 288

Leu Met Glu Val Gln Lys Val Leu Gln Ala Ala Ala Glu Ala Lys Ala

85 90 95

atg ggt agc acc cgt ttt tgt atg ggt gca gca tgg aaa cat ccg agc 336

Met Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser

100 105 110

gca aaa gat ttt ccg tat gtt ctg gaa atg gtt aaa ggt gtt aaa gca 384

Ala Lys Asp Phe Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Ala

115	120	125	
ctg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg tca cgt gaa cag			432
Leu Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Ser Arg Glu Gln			
130	135	140	
acc cag gca ctg gcg gaa gcc ggt ctg gat tat tat aat cat aat ctg			480
Thr Gln Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu			
145	150	155	160
gat acc agc ccg gaa ttt tat ggt cgt att att acc acc cgt acc tat			528
Asp Thr Ser Pro Glu Phe Tyr Gly Arg Ile Ile Thr Thr Arg Thr Tyr			
165	170	175	
gca gaa cgt ctg cag acc ctg gca tat gtt cgt gaa gca ggt atg aaa			576
Ala Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Glu Ala Gly Met Lys			
180	185	190	
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc ctg gat gat cgt			624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg			
195	200	205	
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc			672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser			
210	215	220	
gtt ccg att aat atg ctg gtt aaa gtt cag ggt acc ccg ctg gca gat			720
Val Pro Ile Asn Met Leu Val Lys Val Gln Gly Thr Pro Leu Ala Asp			
225	230	235	240
gca gaa gat gtt gat ccg ttt gat ttt att cgt acc ctg gcg gtt gca			768
Ala Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Thr Leu Ala Val Ala			
245	250	255	
cgt att atg atg ccg aaa agc cat gtt cgt ctg agc gca ggt cgc gaa			816
Arg Ile Met Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu			
260	265	270	
cag atg aat gaa cag atg cag gcg ctg gca ttt ctg gcc ggt gca aat			864

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn
275 280 285
agc att ttt tat ggt gaa aaa ctg ctg acc acc ggt aat ccg cag gca 912
Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Gly Asn Pro Gln Ala
290 295 300
gaa aaa gat ctg cag ctg ttt cgt cgt ctg ggt att cag ccg gaa gaa 960
Glu Lys Asp Leu Gln Leu Phe Arg Arg Leu Gly Ile Gln Pro Glu Glu
305 310 315 320
cgt gaa gaa cat gca gat gaa gtt cat cag gcc gca att gaa cag gcc 1008
Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
325 330 335
ctg gcc gaa cag cgt gat agc cag ctg ttt tat gat gca gca agc gca 1056
Leu Ala Glu Gln Arg Asp Ser Gln Leu Phe Tyr Asp Ala Ala Ser Ala
340 345 350
taa 1059
<210> 79
<211> 352
<212> PRT
<213> Pseudomonas thermotolerans
<400> 79
Met Asn Ala Ser Val Ala Ala Ala Ile Arg His Asp Trp Thr Leu Ala
1 5 10 15
Glu Val Lys Ala Leu Phe Ala Leu Pro Phe Asn Asp Leu Leu Tyr Gln
20 25 30
Ala Gln Thr Val His Arg Gln Tyr Phe Asp Ala Asn Arg Val Gln Val
35 40 45
Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys
50 55 60
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Gln Lys
65 70 75 80
Leu Met Glu Val Gln Lys Val Leu Gln Ala Ala Ala Glu Ala Lys Ala
85 90 95

Met Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser
100 105 110

Ala Lys Asp Phe Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Ala
115 120 125

Leu Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Ser Arg Glu Gln
130 135 140

Thr Gln Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160

Asp Thr Ser Pro Glu Phe Tyr Gly Arg Ile Ile Thr Thr Arg Thr Tyr
165 170 175

Ala Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Glu Ala Gly Met Lys
180 185 190

Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205

Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220

Val Pro Ile Asn Met Leu Val Lys Val Gln Gly Thr Pro Leu Ala Asp
225 230 235 240

Ala Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Thr Leu Ala Val Ala
245 250 255

Arg Ile Met Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Leu Ala Gly Ala Asn
275 280 285

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Gly Asn Pro Gln Ala
290 295 300

Glu Lys Asp Leu Gln Leu Phe Arg Arg Leu Gly Ile Gln Pro Glu Glu
305 310 315 320

Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
325 330 335

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

340 345 350

<210> 80

<211> 1059

<212> DNA

<213> pseudomonad ancestor

<220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from pseudomonad ancestor

<400> 80

atg agc gca agc acc aat agc ccg ctg cgt cat gat tgg acc ctg agc 48

Met Ser Ala Ser Thr Asn Ser Pro Leu Arg His Asp Trp Thr Leu Ser

1 5 10 15

gaa gtt aaa gca ctg ttt acc cag ccg ttt aat gat ctg ctg ttt cat 96

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

20 25 30

gca atg acc gtt cat cgt gca cat ttt gat ccg aat cag gtt cag gtt 144

Ala Met Thr Val His Arg Ala His Phe Asp Pro Asn Gln Val Gln Val

35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa 192

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

50 55 60

tat tgt ccg cag agc ggc cat tat aat acc ggc ctg gaa aaa gaa aaa 240

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Glu Lys

65 70 75 80

ctg atg gaa gtg cag aaa gtt att gaa gaa gca gca cgt gcc aaa gca 288

Leu Met Glu Val Gln Lys Val Ile Glu Glu Ala Ala Arg Ala Lys Ala

85 90 95

att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc 336

Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser

100 105 110

gca aaa gat atg ccg tat gtt ctg gaa atg gtt cgt ggt gtt aaa gca 384

Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Arg Gly Val Lys Ala
115 120 125
atg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg gat cag gat cag 432
Met Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Gln Asp Gln
130 135 140
acc gtt gca ctg gca gaa gca ggt ctg gat tat tat aat cat aat ctg 480

Thr Val Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160
gat acc agc ccg gaa ttt tat ggt aat att att acc acc cgt acc tat 528
Asp Thr Ser Pro Glu Phe Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr
165 170 175
ggt gaa cgt ctg cag acc ctg gcc tat gtt cgt gat gca ggt atg aaa 576
Gly Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ala Gly Met Lys

180 185 190
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc ctg gat gat cgt 624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc 672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220

gtt ccg att aat atg ctg gtt aaa gtt gca ggt acc ccg ctg gaa aat 720
Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Glu Asn
225 230 235 240
gca gaa gat gtt gat ccg ttt gat ttt att cgt atg ctg gca gtt gca 768
Ala Glu Asp Val Asp Pro Phe Asp Phe Ile Arg Met Leu Ala Val Ala
245 250 255
cgt att ctg atg ccg cgt agc cat gtt cgt ctg agc gca ggt cgt gaa 816

Arg Ile Leu Met Pro Arg Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270
cag atg aat gaa cag atg cag gcc ctg gca ttt atg gcc ggt gca aat 864
Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn

275 280 285
 agc att ttt tat ggt gaa aaa ctg ctg acc acc gca aat ccg cag gca 912
 Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala

290 295 300
 gat aaa gat atg cag ctg ttt gca cgt ctg ggt att cgt ccg gaa gca 960
 Asp Lys Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Arg Pro Glu Ala

305 310 315 320
 cgt gaa gaa cat gat gat gaa gtt cat cag gca gca att gaa cag gca 1008
 Arg Glu Glu His Asp Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala

325 330 335
 ctg gtt gaa cag cgt agc agc gaa atg ttt tat gat gca gca gca gtt 1056
 Leu Val Glu Gln Arg Ser Ser Glu Met Phe Tyr Asp Ala Ala Ala Val

340 345 350
 taa 1059

<210> 81

<211> 352

<212> PRT

<213> pseudomonad ancestor

<400> 81

Met Ser Ala Ser Thr Asn Ser Pro Leu Arg His Asp Trp Thr Leu Ser
 1 5 10 15

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

20 25 30
 Ala Met Thr Val His Arg Ala His Phe Asp Pro Asn Gln Val Gln Val

35 40 45
 Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys

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

65 70 75 80

Leu Met Glu Val Gln Lys Val Ile Glu Glu Ala Ala Arg Ala Lys Ala

85 90 95

Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser
100 105 110

Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Arg Gly Val Lys Ala
115 120 125

Met Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Gln Asp Gln
130 135 140

Thr Val Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160

Asp Thr Ser Pro Glu Phe Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr
165 170 175

Gly Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ala Gly Met Lys
180 185 190

Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205

Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220

Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Glu Asn
225 230 235 240

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

Arg Ile Leu Met Pro Arg Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn
275 280 285

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala
290 295 300

Asp Lys Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Arg Pro Glu Ala
305 310 315 320

Arg Glu Glu His Asp Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
325 330 335

Leu Val Glu Gln Arg Ser Ser Glu Met Phe Tyr Asp Ala Ala Ala Val

340 345 350

<210> 82

<211> 1059

<212> DNA

<213> Pseudomonas aeruginosa PA01

<220><221> CDS

<222> (1)..(1059)

<223> bioB gene encoding biotin synthase from Pseudomonas aeruginosa
PA01

<400> 82

atg agc gca acc gca agc gtt gca acc cgt cat gat tgg agc ctg gca 48
Met Ser Ala Thr Ala Ser Val Ala Thr Arg His Asp Trp Ser Leu Ala
1 5 10 15

gaa gtt cgt gca ctg ttt gaa cag ccg ttt aat gat ctg ctg ttt cag 96
Glu Val Arg Ala Leu Phe Glu Gln Pro Phe Asn Asp Leu Leu Phe Gln
20 25 30

gca cag acc gtt cat cgt gca cat ttt gat ccg aat cgt gtt cag gtt 144
Ala Gln Thr Val His Arg Ala His Phe Asp Pro Asn Arg Val Gln Val
35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa 192

Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys
50 55 60

tat tgt ccg cag agc ggt cat tat aat acc ggc ctg gat aaa gaa aaa 240
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Asp Lys Glu Lys
65 70 75 80

ctg atg gaa gtt cag aaa gtt ctg gaa gca gca gca gaa gca aaa gca 288
Leu Met Glu Val Gln Lys Val Leu Glu Ala Ala Ala Glu Ala Lys Ala

85 90 95

att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc 336
Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser
100 105 110

gca aaa gat atg ccg tat gtt ctg gaa atg gtt aaa ggt gtt aaa aaa 384

Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Lys	
115 120 125	
ctg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg acc cag gaa cag	432
Leu Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Thr Gln Glu Gln	
130 135 140	
acc cag gcg ctg gca gat gcc ggt ctg gat tat tat aat cat aat ctg	480
Thr Gln Ala Leu Ala Asp Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu	
145 150 155 160	
gat acc agc ccg gaa ttt tat ggt aat att att acc acc cgt acc tat	528
Asp Thr Ser Pro Glu Phe Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr	
165 170 175	
agc gaa cgt ctg cag acc ctg gcc tat gtt cgt gaa gca ggt atg aaa	576
Ser Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Glu Ala Gly Met Lys	
180 185 190	
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc gtt gat gat cgt	624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Val Asp Asp Arg	
195 200 205	
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc	672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser	
210 215 220	
gtt ccg att aat atg ctg gtt aaa gtt aaa ggt acc ccg ctg gcc gaa	720
Val Pro Ile Asn Met Leu Val Lys Val Lys Gly Thr Pro Leu Ala Glu	
225 230 235 240	
gaa aaa gat gtt gat ccg ttt gat ttt att cgt acc ctg gca gtt gcc	768
Glu Lys Asp Val Asp Pro Phe Asp Phe Ile Arg Thr Leu Ala Val Ala	
245 250 255	
cgt att atg atg ccg aaa agc cat gtt cgt ctg agc gca ggt cgt gaa	816
Arg Ile Met Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu	
260 265 270	

cag atg aat gaa cag atg cag gca ctg gca ttt atg gcc ggt gca aat 864

Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn
275 280 285

agc att ttt tat ggt gaa aaa ctg ctg acc acc aaa aat ccg cag gcc 912

Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Lys Asn Pro Gln Ala
290 295 300

gaa aaa gat atg cag ctg ttt gca cgt ctg ggt att aaa ccg gaa gaa 960

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

305 310 315 320

cgt gaa gaa cat gca gat gaa gtt cat cag gca gca att gaa cag gcc 1008

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

325 330 335

ctg gtt gaa cag cgt gaa agc aaa ctg ttt tat aat gca gca agc gca 1056

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

340 345 350

taa 1059

<210> 83

<211> 352

<212> PRT

<213> Pseudomonas aeruginosa PA01

<400> 83

Met Ser Ala Thr Ala Ser Val Ala Thr Arg His Asp Trp Ser Leu Ala

1 5 10 15

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

20 25 30

Ala Gln Thr Val His Arg Ala His Phe Asp Pro Asn Arg Val Gln Val

35 40 45

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

50 55 60

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Asp Lys Glu Lys

65 70 75 80

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

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

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

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

 Glu Lys Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Lys Pro Glu Glu
 305 310 315 320
 Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala

325	330	335	
Leu Val Glu Gln Arg Glu Ser Lys Leu Phe Tyr Asn Ala Ala Ser Ala			
340	345	350	
<210> 84			
<211> 1059			
<212> DNA			
<213> Pseudomonas balearica			
<220><221> CDS			
<222> (1)..(1059)			
<223> bioB gene encoding biotin synthase from Pseudomonas balearica			
<400> 84			
atg agc gca acc gca agc att gca acc cgt cat gat tgg agc ctg gcg	48		
Met Ser Ala Thr Ala Ser Ile Ala Thr Arg His Asp Trp Ser Leu Ala			
1 5 10 15			
gaa gtt aaa gca ctg ttt gaa cag ccg ttt aat gat ctg ctg ttt cag	96		
Glu Val Lys Ala Leu Phe Glu Gln Pro Phe Asn Asp Leu Leu Phe Gln			
20 25 30			
gca cag acc gtt cat cgt gca cat ttt gat ccg aat cgt gtt cag gtt	144		
Ala Gln Thr Val His Arg Ala His Phe Asp Pro Asn Arg Val Gln Val			
35 40 45			
agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgc aaa	192		
Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys			
50 55 60			
tat tgt ccg cag agc ggt cat tat aat acc ggt ctg gat aaa gaa aaa	240		
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Asp Lys Glu Lys			
65 70 75 80			
ctg atg gaa gtt cag aaa gtt ctg gaa gca gca gca gaa gcc aaa gca	288		
Leu Met Glu Val Gln Lys Val Leu Glu Ala Ala Ala Glu Ala Lys Ala			
85 90 95			
att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc	336		
Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser			

100	105	110	
gca aaa gat atg ccg tat gtt ctg gaa atg gtt aaa ggt gtt aaa aaa			384
Ala Lys Asp Met Pro Tyr Val Leu Glu Met Val Lys Gly Val Lys Lys			
115	120	125	
ctg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg gat cag gaa cag			432
Leu Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Gln Glu Gln			
130	135	140	
acc cag gcc ctg gca gaa gca ggc ctg gat tat tat aat cat aat ctg			480
Thr Gln Ala Leu Ala Glu Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu			
145	150	155	160
gat acc agc ccg gaa ttt tat ggt aat att att acc acc cgt acc tat			528
Asp Thr Ser Pro Glu Phe Tyr Gly Asn Ile Ile Thr Thr Arg Thr Tyr			
165	170	175	
agc gaa cgc ctg cag acc ctg gca tat gtt cgt gaa gca ggt atg aaa			576
Ser Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Glu Ala Gly Met Lys			
180	185	190	
att tgt agc ggt ggt att ctg ggt atg ggt gaa agc gtt gat gat cgt			624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Val Asp Asp Arg			
195	200	205	
gca ggt ctg ctg att cag ctg gca aat ctg ccg gaa cat ccg gaa agc			672
Ala Gly Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser			
210	215	220	
ggt ccg att aat atg ctg gtt aaa gtg aaa ggt acc ccg ctg gcc gaa			720
Val Pro Ile Asn Met Leu Val Lys Val Lys Gly Thr Pro Leu Ala Glu			
225	230	235	240
gaa aaa gat gtt gat ccg ttt gat ttt att cgt acc ctg gcc gtt gca			768
Glu Lys Asp Val Asp Pro Phe Asp Phe Ile Arg Thr Leu Ala Val Ala			
245	250	255	
cgt att atg atg ccg aaa agc cat gtt cgt ctg agc gca ggt cgt gaa			816

Arg Ile Met Met Pro Lys Ser His Val Arg Leu Ser Ala Gly Arg Glu
260 265 270
cag atg aat gaa cag atg cag gcg ctg gca ttt atg gcc ggt gca aat 864
Gln Met Asn Glu Gln Met Gln Ala Leu Ala Phe Met Ala Gly Ala Asn
275 280 285
agc att ttt tat ggt gaa aaa ctg ctg acc acc gca aat ccg cag gca 912
Ser Ile Phe Tyr Gly Glu Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala
290 295 300
gaa aaa gat atg cag ctg ttt gca cgt ctg ggt att aaa ccg gaa gaa 960
Glu Lys Asp Met Gln Leu Phe Ala Arg Leu Gly Ile Lys Pro Glu Glu
305 310 315 320
cgt gaa gaa cat gca gat gaa gtt cat cag gca gca att gaa cag gca 1008
Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
325 330 335
ctg gtt gaa cag cgt gat agc cag ctg ttt tat gat gca gca agc gca 1056
Leu Val Glu Gln Arg Asp Ser Gln Leu Phe Tyr Asp Ala Ala Ser Ala
340 345 350
taa 1059
<210> 85
<211> 352
<212> PRT
<213> *Pseudomonas balearica*
<400> 85
Met Ser Ala Thr Ala Ser Ile Ala Thr Arg His Asp Trp Ser Leu Ala
1 5 10 15
Glu Val Lys Ala Leu Phe Glu Gln Pro Phe Asn Asp Leu Leu Phe Gln
20 25 30
Ala Gln Thr Val His Arg Ala His Phe Asp Pro Asn Arg Val Gln Val
35 40 45
Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys
50 55 60
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Asp Lys Glu Lys

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

Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
325 330 335

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

<210> 86

<211> 1056

<212> DNA

<213> Pseudomonas fluorescens SBW25

<220><221> CDS

<222> (1)..(1056)

<223> bioB gene encoding biotin synthase from Pseudomonas fluorescens
SBW25

<400> 86

atg agc gca agc acc acc gcc acc ctg cgt cat gat tgg agc ctg gca 48
Met Ser Ala Ser Thr Thr Ala Thr Leu Arg His Asp Trp Ser Leu Ala
1 5 10 15

gaa gtt aaa gca ctg ttt gtt cag ccg ttt aat gat ctg ctg ttt cag 96
Glu Val Lys Ala Leu Phe Val Gln Pro Phe Asn Asp Leu Leu Phe Gln
20 25 30

gca cag acc gtt cat cgt gca cat ttt gat gca aat cgt gtt cag gtt 144
Ala Gln Thr Val His Arg Ala His Phe Asp Ala Asn Arg Val Gln Val
35 40 45

agc acc ctg ctg agc att aaa acc ggt gca tgt ccg gaa gat tgt aaa 192

Ser Thr Leu Leu Ser Ile Lys Thr Gly Ala Cys Pro Glu Asp Cys Lys
50 55 60

tat tgt ccg cag agc ggt cat tat aat acc ggc ctg gaa aaa gaa aaa 240
Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Glu Lys
65 70 75 80

ctg atg gaa gtt cag aaa gtt ctg gaa gaa gca gca cgt gca aaa gca 288
Leu Met Glu Val Gln Lys Val Leu Glu Glu Ala Ala Arg Ala Lys Ala

85

90

95

att ggt agc acc cgt ttt tgc atg ggt gca gca tgg aaa cat ccg agc 336
Ile Gly Ser Thr Arg Phe Cys Met Gly Ala Ala Trp Lys His Pro Ser
100 105 110

gca aaa gat atg ccg tat gtt ctg cag atg gtt aaa ggt gtt aaa gca 384
Ala Lys Asp Met Pro Tyr Val Leu Gln Met Val Lys Gly Val Lys Ala
115 120 125

atg ggt ctg gaa acc tgt atg acc ctg ggt cgt ctg gat cag gat cag 432
Met Gly Leu Glu Thr Cys Met Thr Leu Gly Arg Leu Asp Gln Asp Gln
130 135 140

acc gaa gca ctg gca cag gca ggt ctg gat tat tat aat cat aat ctg 480
Thr Glu Ala Leu Ala Gln Ala Gly Leu Asp Tyr Tyr Asn His Asn Leu
145 150 155 160

gat acc agc ccg gaa ttt tat ggt agc att att acc acc cgt acc tat 528

Asp Thr Ser Pro Glu Phe Tyr Gly Ser Ile Ile Thr Thr Arg Thr Tyr
165 170 175

ggt gaa cgt ctg cag acc ctg gca tat gtt cgt gat agc ggt atg aaa 576
Gly Glu Arg Leu Gln Thr Leu Ala Tyr Val Arg Asp Ser Gly Met Lys
180 185 190

att tgt agc ggt ggt att ctg ggt atg ggt gaa agc ctg gat gat cgt 624
Ile Cys Ser Gly Gly Ile Leu Gly Met Gly Glu Ser Leu Asp Asp Arg
195 200 205

gca aat ctg ctg att cag ctg gcc aat ctg ccg gaa cat ccg gaa agc 672
Ala Asn Leu Leu Ile Gln Leu Ala Asn Leu Pro Glu His Pro Glu Ser
210 215 220

gtt ccg att aat atg ctg gtt aaa gtt gca ggt acc ccg ctg gaa aat 720
Val Pro Ile Asn Met Leu Val Lys Val Ala Gly Thr Pro Leu Glu Asn
225 230 235 240

gcc gaa gat att gat ccg ttt gat ttt att cgt atg ctg gca gtt gca 768
Ala Glu Asp Ile Asp Pro Phe Asp Phe Ile Arg Met Leu Ala Val Ala
245 250 255

cgt att ctg atg ccg cgt agc cat gtt cgt ctg agc gca ggt cgt gaa 816

Arg Ile Leu Met Pro Arg Ser His Val Arg Leu Ser Ala Gly Arg Glu
 260 265 270
 gca atg aat gaa cag atg cag gca ctg gcc ttt ttt gcc ggt gca aat 864

Ala Met Asn Glu Gln Met Gln Ala Leu Ala Phe Phe Ala Gly Ala Asn
 275 280 285
 agc att ttt tat ggt gat aaa ctg ctg acc acc gca aat ccg cag gca 912
 Ser Ile Phe Tyr Gly Asp Lys Leu Leu Thr Thr Ala Asn Pro Gln Ala
 290 295 300
 gat aaa gat atg cag ctg ttt agc cgt ctg ggt att ctg ccg gaa gca 960
 Asp Lys Asp Met Gln Leu Phe Ser Arg Leu Gly Ile Leu Pro Glu Ala

305 310 315 320
 cgt gaa gaa cat gca gat gaa gtt cat cag gca gca att gaa cag gcc 1008
 Arg Glu Glu His Ala Asp Glu Val His Gln Ala Ala Ile Glu Gln Ala
 325 330 335
 ctg gtt gaa cag aaa agc agc gaa cag ttt tat aat gca gca gtt taa 1056
 Leu Val Glu Gln Lys Ser Ser Glu Gln Phe Tyr Asn Ala Ala Val
 340 345 350

<210> 87

<211> 351

<212> PRT

<213> Pseudomonas fluorescens SBW25

<400> 87

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

50 55 60

Tyr Cys Pro Gln Ser Gly His Tyr Asn Thr Gly Leu Glu Lys Glu Lys

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

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

325 330 335

Leu Val Glu Gln Lys Ser Ser Glu Gln Phe Tyr Asn Ala Ala Val

340 345 350

<210> 88

<211> 756

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(756)

<223> bioC gene encoding SAM (S-adenosylmethionine)-dependent
methyltransferase (BioC)

<400> 88

atg gca acg gtt aat aaa caa gcc att gca gcg gca ttt ggt cgg gca 48

Met Ala Thr Val Asn Lys Gln Ala Ile Ala Ala Ala Phe Gly Arg Ala

1 5 10 15

gcc gca cac tat gag caa cat gca gat cta cag cgc cag agt gct gac 96

Ala Ala His Tyr Glu Gln His Ala Asp Leu Gln Arg Gln Ser Ala Asp

20 25 30

gcc tta ctg gca atg ctt cca cag cgt aaa tac acc cac gta ctg gac 144

Ala Leu Leu Ala Met Leu Pro Gln Arg Lys Tyr Thr His Val Leu Asp

35 40 45

gcg ggt tgt gga cct ggc tgg atg agc cgc cac tgg cgg gaa cgt cac 192

Ala Gly Cys Gly Pro Gly Trp Met Ser Arg His Trp Arg Glu Arg His

50 55 60

gcg cag gtg acg gcc tta gat ctc tcg ccg cca atg ctt gtt cag gca 240

Ala Gln Val Thr Ala Leu Asp Leu Ser Pro Pro Met Leu Val Gln Ala

65 70 75 80

cgc cag aag gat gcc gca gac cat tat ctg gcg gga gat atc gaa tcc 288

Arg Gln Lys Asp Ala Ala Asp His Tyr Leu Ala Gly Asp Ile Glu Ser

85 90 95

ctg ccg tta gcg act gcg acg ttc gat ctt gca tgg agc aat ctc gca 336

Leu Pro Leu Ala Thr Ala Thr Phe Asp Leu Ala Trp Ser Asn Leu Ala
100 105 110

gtg cag tgg tgc ggt aat tta tcc acg gca ctc cgc gag ctg tat cgg 384
Val Gln Trp Cys Gly Asn Leu Ser Thr Ala Leu Arg Glu Leu Tyr Arg
115 120 125

gtg gtg cgc ccc aaa ggc gtg gtc gcg ttt acc acg ctg gtg cag gga 432
Val Val Arg Pro Lys Gly Val Val Ala Phe Thr Thr Leu Val Gln Gly
130 135 140

tcg tta ccc gaa ctg cat cag gcg tgg cag gcg gtg gac gag cgt ccg 480

Ser Leu Pro Glu Leu His Gln Ala Trp Gln Ala Val Asp Glu Arg Pro
145 150 155 160

cat gct aat cgc ttt tta ccg cca gat gaa atc gaa cag tcg ctg aac 528
His Ala Asn Arg Phe Leu Pro Pro Asp Glu Ile Glu Gln Ser Leu Asn
165 170 175

ggc gtg cat tat caa cat cat att cag ccc atc acg ctg tgg ttt gat 576
Gly Val His Tyr Gln His His Ile Gln Pro Ile Thr Leu Trp Phe Asp

180 185 190

gat gcg ctc agt gcc atg cgt tcg ctg aaa ggc atc ggt gcc acg cat 624
Asp Ala Leu Ser Ala Met Arg Ser Leu Lys Gly Ile Gly Ala Thr His
195 200 205

ctt cat gaa ggg cgc gac ccg cga ata tta acg cgt tcg cag ttg cag 672
Leu His Glu Gly Arg Asp Pro Arg Ile Leu Thr Arg Ser Gln Leu Gln
210 215 220

cga ttg caa ctg gcc tgg ccg caa cag cag ggg cga tat cct ctg acg 720
Arg Leu Gln Leu Ala Trp Pro Gln Gln Gln Gly Arg Tyr Pro Leu Thr
225 230 235 240

tat cat ctt ttt ttg gga gtg att gct cgt gag taa 756
Tyr His Leu Phe Leu Gly Val Ile Ala Arg Glu
245 250

<210> 89
<211> 251

<212> PRT

<213> Escherichia coli

<400> 89

Met Ala Thr Val Asn Lys Gln Ala Ile Ala Ala Ala Phe Gly Arg Ala

1 5 10 15

Ala Ala His Tyr Glu Gln His Ala Asp Leu Gln Arg Gln Ser Ala Asp

20 25 30

Ala Leu Leu Ala Met Leu Pro Gln Arg Lys Tyr Thr His Val Leu Asp

35 40 45

Ala Gly Cys Gly Pro Gly Trp Met Ser Arg His Trp Arg Glu Arg His

50 55 60

Ala Gln Val Thr Ala Leu Asp Leu Ser Pro Pro Met Leu Val Gln Ala

65 70 75 80

Arg Gln Lys Asp Ala Ala Asp His Tyr Leu Ala Gly Asp Ile Glu Ser

85 90 95

Leu Pro Leu Ala Thr Ala Thr Phe Asp Leu Ala Trp Ser Asn Leu Ala

100 105 110

Val Gln Trp Cys Gly Asn Leu Ser Thr Ala Leu Arg Glu Leu Tyr Arg

115 120 125

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

130 135 140

Ser Leu Pro Glu Leu His Gln Ala Trp Gln Ala Val Asp Glu Arg Pro

145 150 155 160

His Ala Asn Arg Phe Leu Pro Pro Asp Glu Ile Glu Gln Ser Leu Asn

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220

Arg Leu Gln Leu Ala Trp Pro Gln Gln Gln Gly Arg Tyr Pro Leu Thr
225 230 235 240

Tyr His Leu Phe Leu Gly Val Ile Ala Arg Glu
245 250

<210> 90

<211> 1155

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(1155)

<223> bioF gene encoding (Bio F) 7-keto-8-aminopelargonic acid (KAPA)
synthase

<400> 90

atg agc tgg cag gag aaa atc aac gcg gcg ctc gat gcg cgg cgt gct 48

Met Ser Trp Gln Glu Lys Ile Asn Ala Ala Leu Asp Ala Arg Arg Ala

1 5 10 15

gcc gat gcc ctg cgt cgc cgt tat ccg gtg gcg caa gga gcc gga cgc 96

Ala Asp Ala Leu Arg Arg Arg Tyr Pro Val Ala Gln Gly Ala Gly Arg

20 25 30

tgg ctg gtg gcg gat gat cgc cag tat ctg aac ttt tcc agt aac gat 144

Trp Leu Val Ala Asp Asp Arg Gln Tyr Leu Asn Phe Ser Ser Asn Asp

35 40 45

tat tta ggt tta agc cat cat ccg caa att atc cgt gcc tgg cag cag 192

Tyr Leu Gly Leu Ser His His Pro Gln Ile Ile Arg Ala Trp Gln Gln

50 55 60

ggg gcg gag caa ttt ggc atc ggt agc ggc ggc tcc ggt cac gtc agc 240

Gly Ala Glu Gln Phe Gly Ile Gly Ser Gly Gly Ser Gly His Val Ser

65 70 75 80

ggt tat agc gtg gtg cat cag gca ctg gaa gaa gag ctg gcc gag tgg 288

Gly Tyr Ser Val Val His Gln Ala Leu Glu Glu Glu Leu Ala Glu Trp

85 90 95

ctt ggc tat tcg cgg gca ctg ctg ttt atc tct ggt ttc gcc gct aat 336

Leu Gly Tyr Ser Arg Ala Leu Leu Phe Ile Ser Gly Phe Ala Ala Asn
100 105 110

cag gca gtt att gcc gcg atg atg gcg aaa gag gac cgt att gct gcc 384
Gln Ala Val Ile Ala Ala Met Met Ala Lys Glu Asp Arg Ile Ala Ala
115 120 125

gac cgg ctt agc cat gcc tca ttg ctg gaa gct gcc agt tta agc ccg 432
Asp Arg Leu Ser His Ala Ser Leu Leu Glu Ala Ala Ser Leu Ser Pro

130 135 140

tcg cag ctt cgc cgt ttt gct cat aac gat gtc act cat ttg gcg cga 480
Ser Gln Leu Arg Arg Phe Ala His Asn Asp Val Thr His Leu Ala Arg
145 150 155 160

ttg ctt gct tcc ccc tgt ccg ggg cag caa atg gtg gtg aca gaa ggc 528
Leu Leu Ala Ser Pro Cys Pro Gly Gln Gln Met Val Val Thr Glu Gly
165 170 175

gtg ttc agc atg gac ggc gat agt gcg cca ctg gcg gaa atc cag cag 576
Val Phe Ser Met Asp Gly Asp Ser Ala Pro Leu Ala Glu Ile Gln Gln
180 185 190

gta acg caa cag cac aat ggc tgg ttg atg gtc gat gat gcc cac ggc 624
Val Thr Gln Gln His Asn Gly Trp Leu Met Val Asp Asp Ala His Gly
195 200 205

acg ggc gtt atc ggg gag cag ggg cgc ggc agc tgc tgg ctg caa aag 672

Thr Gly Val Ile Gly Glu Gln Gly Arg Gly Ser Cys Trp Leu Gln Lys
210 215 220

gta aaa cca gaa ttg ctg gta gtg act ttt ggc aaa gga ttt ggc gtc 720
Val Lys Pro Glu Leu Leu Val Val Thr Phe Gly Lys Gly Phe Gly Val
225 230 235 240

agc ggg gca gcg gtg ctt tgc tcc agt acg gtg gcg gat tat ctg ctg 768
Ser Gly Ala Ala Val Leu Cys Ser Ser Thr Val Ala Asp Tyr Leu Leu

245	250	255	
caa ttc gcc cgc cac ctt atc tac agc acc agt atg ccg ccc gct cag			816
Gln Phe Ala Arg His Leu Ile Tyr Ser Thr Ser Met Pro Pro Ala Gln			
260	265	270	
gcg cag gca tta cgt gcg tcg ctg gcg gtc att cgc agt gat gag ggt			864
Ala Gln Ala Leu Arg Ala Ser Leu Ala Val Ile Arg Ser Asp Glu Gly			
275	280	285	
gat gca cgg cgc gaa aaa ctg gcg gca ctc att acg cgt ttt cgt gcc			912
Asp Ala Arg Arg Glu Lys Leu Ala Ala Leu Ile Thr Arg Phe Arg Ala			
290	295	300	
gga gta cag gat ttg ccg ttt acg ctt gct gat tca tgc agc gcc atc			960
Gly Val Gln Asp Leu Pro Phe Thr Leu Ala Asp Ser Cys Ser Ala Ile			
305	310	315	320
cag cca ttg att gtc ggt gat aac agc cgt gcg tta caa ctg gca gaa			1008
Gln Pro Leu Ile Val Gly Asp Asn Ser Arg Ala Leu Gln Leu Ala Glu			
325	330	335	
aaa ctg cgt cag caa ggc tgc tgg gtc acg gcg att cgc ccg cca acc			1056
Lys Leu Arg Gln Gln Gly Cys Trp Val Thr Ala Ile Arg Pro Pro Thr			
340	345	350	
gta ccc gct ggt act gcg cga ctg cgc tta acg cta acc gct gcg cat			1104
Val Pro Ala Gly Thr Ala Arg Leu Arg Leu Thr Leu Thr Ala Ala His			
355	360	365	
gaa atg cag gat atc gac cgt ctg ctg gag gtg ctg cat ggc aac ggt			1152
Glu Met Gln Asp Ile Asp Arg Leu Leu Glu Val Leu His Gly Asn Gly			
370	375	380	
taa			1155
<210>	91		
<211>	384		
<212>	PRT		
<213>	Escherichia coli		
<400>	91		

Met Ser Trp Gln Glu Lys Ile Asn Ala Ala Leu Asp Ala Arg Arg Ala

1 5 10 15
Ala Asp Ala Leu Arg Arg Arg Tyr Pro Val Ala Gln Gly Ala Gly Arg

20 25 30
Trp Leu Val Ala Asp Asp Arg Gln Tyr Leu Asn Phe Ser Ser Asn Asp

35 40 45
Tyr Leu Gly Leu Ser His His Pro Gln Ile Ile Arg Ala Trp Gln Gln

50 55 60
Gly Ala Glu Gln Phe Gly Ile Gly Ser Gly Gly Ser Gly His Val Ser

65 70 75 80
Gly Tyr Ser Val Val His Gln Ala Leu Glu Glu Glu Leu Ala Glu Trp

85 90 95
Leu Gly Tyr Ser Arg Ala Leu Leu Phe Ile Ser Gly Phe Ala Ala Asn

100 105 110
Gln Ala Val Ile Ala Ala Met Met Ala Lys Glu Asp Arg Ile Ala Ala

115 120 125
Asp Arg Leu Ser His Ala Ser Leu Leu Glu Ala Ala Ser Leu Ser Pro

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

145 150 155 160
Leu Leu Ala Ser Pro Cys Pro Gly Gln Gln Met Val Val Thr Glu Gly

165 170 175
Val Phe Ser Met Asp Gly Asp Ser Ala Pro Leu Ala Glu Ile Gln Gln

180 185 190
Val Thr Gln Gln His Asn Gly Trp Leu Met Val Asp Asp Ala His Gly

195 200 205
Thr Gly Val Ile Gly Glu Gln Gly Arg Gly Ser Cys Trp Leu Gln Lys

210 215 220
Val Lys Pro Glu Leu Leu Val Val Thr Phe Gly Lys Gly Phe Gly Val

225 230 235 240
Ser Gly Ala Ala Val Leu Cys Ser Ser Thr Val Ala Asp Tyr Leu Leu

245 250 255
Gln Phe Ala Arg His Leu Ile Tyr Ser Thr Ser Met Pro Pro Ala Gln

260 265 270
Ala Gln Ala Leu Arg Ala Ser Leu Ala Val Ile Arg Ser Asp Glu Gly

275 280 285
Asp Ala Arg Arg Glu Lys Leu Ala Ala Leu Ile Thr Arg Phe Arg Ala

290 295 300
Gly Val Gln Asp Leu Pro Phe Thr Leu Ala Asp Ser Cys Ser Ala Ile

305 310 315 320
Gln Pro Leu Ile Val Gly Asp Asn Ser Arg Ala Leu Gln Leu Ala Glu

325 330 335
Lys Leu Arg Gln Gln Gly Cys Trp Val Thr Ala Ile Arg Pro Pro Thr

340 345 350
Val Pro Ala Gly Thr Ala Arg Leu Arg Leu Thr Leu Thr Ala Ala His

355 360 365
Glu Met Gln Asp Ile Asp Arg Leu Leu Glu Val Leu His Gly Asn Gly

370 375 380

<210> 92

<211> 1290

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(1290)

<223> bioA gene encoding 7,8-Diaminopelargonic Acid (DAPA) Synthase
(BioA)

<400> 92

atg aca acg gac gat ctt gcc ttt gac caa cgc cat atc tgg cac cca 48
Met Thr Thr Asp Asp Leu Ala Phe Asp Gln Arg His Ile Trp His Pro

1 5 10 15

tac aca tcc atg acc tcc cct ctg ccg gtt tat ccg gtg gtg agc gcc 96
Tyr Thr Ser Met Thr Ser Pro Leu Pro Val Tyr Pro Val Val Ser Ala

20	25	30	
gaa ggt tgc gag ctg att ttg tct gac ggc aga cgc ctg gtt gac ggt			144
Glu Gly Cys Glu Leu Ile Leu Ser Asp Gly Arg Arg Leu Val Asp Gly			
35	40	45	
atg tcg tcc tgg tgg gcg gcg atc cac ggc tac aat cac ccg cag ctt			192
Met Ser Ser Trp Trp Ala Ala Ile His Gly Tyr Asn His Pro Gln Leu			
50	55	60	
aat gcg gcg atg aag tcg caa att gat gcc atg tcg cat gtg atg ttt			240
Asn Ala Ala Met Lys Ser Gln Ile Asp Ala Met Ser His Val Met Phe			
65	70	75	80
ggc ggt atc acc cat gcg cca gcc att gag ctg tgc cgc aaa ctg gtg			288
Gly Gly Ile Thr His Ala Pro Ala Ile Glu Leu Cys Arg Lys Leu Val			
85	90	95	
gcg atg acg ccg caa ccg ctg gag tgc gtt ttt ctc gcg gac tcc ggt			336
Ala Met Thr Pro Gln Pro Leu Glu Cys Val Phe Leu Ala Asp Ser Gly			
100	105	110	
tcc gta gcg gtg gaa gtg gcg atg aaa atg gcg ttg cag tac tgg caa			384
Ser Val Ala Val Glu Val Ala Met Lys Met Ala Leu Gln Tyr Trp Gln			
115	120	125	
gcc aaa ggc gaa gcg cgc cag cgt ttt ctg acc ttc cgc aat ggt tat			432
Ala Lys Gly Glu Ala Arg Gln Arg Phe Leu Thr Phe Arg Asn Gly Tyr			
130	135	140	
cat ggc gat acc ttt ggc gcg atg tcg gtg tgc gat ccg gat aac tca			480
His Gly Asp Thr Phe Gly Ala Met Ser Val Cys Asp Pro Asp Asn Ser			
145	150	155	160
atg cac agt ctg tgg aaa ggc tac ctg cca gaa aac ctg ttt gct ccc			528
Met His Ser Leu Trp Lys Gly Tyr Leu Pro Glu Asn Leu Phe Ala Pro			
165	170	175	
gcc ccg caa agc cgc atg gat ggc gaa tgg gat gag cgc gat atg gtg			576
Ala Pro Gln Ser Arg Met Asp Gly Glu Trp Asp Glu Arg Asp Met Val			
180	185	190	

ggc ttt gcc cgc ctg atg gcg gcg cat cgt cat gaa atc gcg gcg gtg 624
 Gly Phe Ala Arg Leu Met Ala Ala His Arg His Glu Ile Ala Ala Val
 195 200 205

atc att gag ccg att gtc cag ggc gca ggc ggg atg cgc atg tac cat 672

Ile Ile Glu Pro Ile Val Gln Gly Ala Gly Gly Met Arg Met Tyr His
 210 215 220

ccg gaa tgg tta aaa cga atc cgc aaa ata tgc gat cgc gaa ggt atc 720
 Pro Glu Trp Leu Lys Arg Ile Arg Lys Ile Cys Asp Arg Glu Gly Ile
 225 230 235 240

ttg ctg att gcc gac gag atc gcc act gga ttt ggt cgt acc ggg aaa 768
 Leu Leu Ile Ala Asp Glu Ile Ala Thr Gly Phe Gly Arg Thr Gly Lys
 245 250 255

ctg ttt gcc tgt gaa cat gca gaa atc gcg ccg gac att ttg tgc ctc 816
 Leu Phe Ala Cys Glu His Ala Glu Ile Ala Pro Asp Ile Leu Cys Leu
 260 265 270

ggg aaa gcc tta acc ggc ggc aca atg acc ctt tcc gcc aca ctc acc 864
 Gly Lys Ala Leu Thr Gly Gly Thr Met Thr Leu Ser Ala Thr Leu Thr
 275 280 285

acg cgc gag gtt gca gaa acc atc agt aac ggt gaa gcc ggt tgc ttt 912
 Thr Arg Glu Val Ala Glu Thr Ile Ser Asn Gly Glu Ala Gly Cys Phe
 290 295 300

atg cat ggg cca act ttt atg ggc aat ccg ctg gcc tgc gcg gca gca 960
 Met His Gly Pro Thr Phe Met Gly Asn Pro Leu Ala Cys Ala Ala Ala
 305 310 315 320

aac gcc agc ctg gcg att ctc gaa tct ggc gac tgg cag caa cag gtg 1008

Asn Ala Ser Leu Ala Ile Leu Glu Ser Gly Asp Trp Gln Gln Gln Val
 325 330 335

gcg gat att gaa gta cag ctg cgc gag caa ctt gcc ccc gcc cgt gat 1056
 Ala Asp Ile Glu Val Gln Leu Arg Glu Gln Leu Ala Pro Ala Arg Asp
 340 345 350

gcc gaa atg gtt gcc gat gtg cgc gta ctg ggg gcc att ggc gtg gtc 1104

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

355 360 365
gaa acc act cat ccg gtg aat atg gcg gcg ctg caa aaa ttc ttt gtc 1152
Glu Thr Thr His Pro Val Asn Met Ala Ala Leu Gln Lys Phe Phe Val

370 375 380
gaa cag ggt gtc tgg atc cgg cct ttt ggc aaa ctg att tac ctg atg 1200
Glu Gln Gly Val Trp Ile Arg Pro Phe Gly Lys Leu Ile Tyr Leu Met
385 390 395 400

ccg ccc tat att att ctc ccg caa cag ttg cag cgt ctg acc gca gcg 1248
Pro Pro Tyr Ile Ile Leu Pro Gln Gln Leu Gln Arg Leu Thr Ala Ala

405 410 415
gtt aac cgc gcg gta cag gat gaa aca ttt ttt tgc caa taa 1290
Val Asn Arg Ala Val Gln Asp Glu Thr Phe Phe Cys Gln

420 425
<210> 93
<211> 429
<212> PRT
<213> Escherichia coli

<400> 93

Met Thr Thr Asp Asp Leu Ala Phe Asp Gln Arg His Ile Trp His Pro
1 5 10 15

Tyr Thr Ser Met Thr Ser Pro Leu Pro Val Tyr Pro Val Val Ser Ala
20 25 30

Glu Gly Cys Glu Leu Ile Leu Ser Asp Gly Arg Arg Leu Val Asp Gly
35 40 45

Met Ser Ser Trp Trp Ala Ala Ile His Gly Tyr Asn His Pro Gln Leu
50 55 60

Asn Ala Ala Met Lys Ser Gln Ile Asp Ala Met Ser His Val Met Phe
65 70 75 80

Gly Gly Ile Thr His Ala Pro Ala Ile Glu Leu Cys Arg Lys Leu Val
85 90 95

Ala Met Thr Pro Gln Pro Leu Glu Cys Val Phe Leu Ala Asp Ser Gly
100 105 110

Ser Val Ala Val Glu Val Ala Met Lys Met Ala Leu Gln Tyr Trp Gln
115 120 125

Ala Lys Gly Glu Ala Arg Gln Arg Phe Leu Thr Phe Arg Asn Gly Tyr
130 135 140

His Gly Asp Thr Phe Gly Ala Met Ser Val Cys Asp Pro Asp Asn Ser
145 150 155 160

Met His Ser Leu Trp Lys Gly Tyr Leu Pro Glu Asn Leu Phe Ala Pro
165 170 175

Ala Pro Gln Ser Arg Met Asp Gly Glu Trp Asp Glu Arg Asp Met Val
180 185 190

Gly Phe Ala Arg Leu Met Ala Ala His Arg His Glu Ile Ala Ala Val
195 200 205

Ile Ile Glu Pro Ile Val Gln Gly Ala Gly Gly Met Arg Met Tyr His
210 215 220

Pro Glu Trp Leu Lys Arg Ile Arg Lys Ile Cys Asp Arg Glu Gly Ile
225 230 235 240

Leu Leu Ile Ala Asp Glu Ile Ala Thr Gly Phe Gly Arg Thr Gly Lys
245 250 255

Leu Phe Ala Cys Glu His Ala Glu Ile Ala Pro Asp Ile Leu Cys Leu
260 265 270

Gly Lys Ala Leu Thr Gly Gly Thr Met Thr Leu Ser Ala Thr Leu Thr
275 280 285

Thr Arg Glu Val Ala Glu Thr Ile Ser Asn Gly Glu Ala Gly Cys Phe
290 295 300

Met His Gly Pro Thr Phe Met Gly Asn Pro Leu Ala Cys Ala Ala Ala
305 310 315 320

Asn Ala Ser Leu Ala Ile Leu Glu Ser Gly Asp Trp Gln Gln Gln Val
325 330 335

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

340 345 350
Ala Glu Met Val Ala Asp Val Arg Val Leu Gly Ala Ile Gly Val Val

355 360 365
Glu Thr Thr His Pro Val Asn Met Ala Ala Leu Gln Lys Phe Phe Val

370 375 380

Glu Gln Gly Val Trp Ile Arg Pro Phe Gly Lys Leu Ile Tyr Leu Met

385 390 395 400
Pro Pro Tyr Ile Ile Leu Pro Gln Gln Leu Gln Arg Leu Thr Ala Ala

405 410 415
Val Asn Arg Ala Val Gln Asp Glu Thr Phe Phe Cys Gln

420 425

<210> 94

<211> 678

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(678)

<223> bioD gene encoding Dethiobiotin (DTB) Synthetase (BioD)

<400> 94

gtg agt aaa cgt tat ttt gtc acc gga acg gat acc gaa gtg ggg aaa 48

Val Ser Lys Arg Tyr Phe Val Thr Gly Thr Asp Thr Glu Val Gly Lys

1 5 10 15

act gtc gcc agt tgt gca ctt tta caa gcc gca aag gca gca ggc tac 96

Thr Val Ala Ser Cys Ala Leu Leu Gln Ala Ala Lys Ala Ala Gly Tyr

20 25 30

cgg acg gca ggt tat aaa ccg gtc gcc tct ggc agc gaa aag acc ccg 144

Arg Thr Ala Gly Tyr Lys Pro Val Ala Ser Gly Ser Glu Lys Thr Pro

35 40 45

gaa ggt tta cgc aat agc gac gcg ctg gcg tta cag cgc aac agc agc 192

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

50 55 60

ctg cag ctg gat tac gca aca gta aat cct tac acc ttc gca gaa ccc 240

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

65	70	75	80	
act tcg ccg cac atc atc agc gcg caa gag ggc aga ccg ata gaa tca				288
Thr Ser Pro His Ile Ile Ser Ala Gln Glu Gly Arg Pro Ile Glu Ser				
	85	90	95	
ttg gta atg agc gcc gga tta cgc gcg ctt gaa caa cag gct gac tgg				336
Leu Val Met Ser Ala Gly Leu Arg Ala Leu Glu Gln Gln Ala Asp Trp				
	100	105	110	
gtg tta gtg gaa ggt gct ggc ggc tgg ttt acg ccg ctt tct gac act				384
Val Leu Val Glu Gly Ala Gly Gly Trp Phe Thr Pro Leu Ser Asp Thr				
	115	120	125	
ttc act ttt gca gat tgg gta aca cag gaa caa ctg ccg gtg ata ctg				432
Phe Thr Phe Ala Asp Trp Val Thr Gln Glu Gln Leu Pro Val Ile Leu				
	130	135	140	
gta gtt ggt gtg aaa ctc ggc tgt att aat cac gcg atg ttg act gca				480
Val Val Gly Val Lys Leu Gly Cys Ile Asn His Ala Met Leu Thr Ala				
	145	150	155	160
cag gta ata caa cac gcc gga ctg act ctg gcg ggt tgg gtg gcg aac				528
Gln Val Ile Gln His Ala Gly Leu Thr Leu Ala Gly Trp Val Ala Asn				
	165	170	175	
gat gtt acg cct ccg gga aaa cgt cac gct gaa tat atg acc acg ctc				576
Asp Val Thr Pro Pro Gly Lys Arg His Ala Glu Tyr Met Thr Thr Leu				
	180	185	190	
acc cgc atg att ccc gcg ccg ctg ctg gga gag atc ccc tgg ctt gca				624
Thr Arg Met Ile Pro Ala Pro Leu Leu Gly Glu Ile Pro Trp Leu Ala				
	195	200	205	
gaa aat cca gaa aat gcg gca acc gga aag tac ata aac ctt gcc ttg				672
Glu Asn Pro Glu Asn Ala Ala Thr Gly Lys Tyr Ile Asn Leu Ala Leu				
	210	215	220	

ttg tag

678

Leu

225

<210> 95

<211> 225

<212> PRT

<213> Escherichia coli

<400> 95

Val Ser Lys Arg Tyr Phe Val Thr Gly Thr Asp Thr Glu Val Gly Lys

1 5 10 15

Thr Val Ala Ser Cys Ala Leu Leu Gln Ala Ala Lys Ala Ala Gly Tyr

20 25 30

Arg Thr Ala Gly Tyr Lys Pro Val Ala Ser Gly Ser Glu Lys Thr Pro

35 40 45

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

50 55 60

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

65 70 75 80

Thr Ser Pro His Ile Ile Ser Ala Gln Glu Gly Arg Pro Ile Glu Ser

85 90 95

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

100 105 110

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

115 120 125

Phe Thr Phe Ala Asp Trp Val Thr Gln Glu Gln Leu Pro Val Ile Leu

130 135 140

Val Val Gly Val Lys Leu Gly Cys Ile Asn His Ala Met Leu Thr Ala

145 150 155 160

Gln Val Ile Gln His Ala Gly Leu Thr Leu Ala Gly Trp Val Ala Asn

165 170 175

Asp Val Thr Pro Pro Gly Lys Arg His Ala Glu Tyr Met Thr Thr Leu

180 185 190

Thr Arg Met Ile Pro Ala Pro Leu Leu Gly Glu Ile Pro Trp Leu Ala
195 200 205

Glu Asn Pro Glu Asn Ala Ala Thr Gly Lys Tyr Ile Asn Leu Ala Leu
210 215 220

Leu

225

<210> 96

<211> 1347

<212> DNA

<213> Bacillus subtilis

<220><221> CDS

<222> (1)..(1347)

<223> bioK gene encoding L-lysine:8-amino-7-oxononanoate
aminotransferase (BioK)

<400> 96

atg act cac gat tta atc gaa aaa agc aaa aag cac ttg tgg ctg ccc 48

Met Thr His Asp Leu Ile Glu Lys Ser Lys Lys His Leu Trp Leu Pro

1 5 10 15

ttc act cag atg aaa gat tat gac gaa aac cct ttg atc att gaa agc 96

Phe Thr Gln Met Lys Asp Tyr Asp Glu Asn Pro Leu Ile Ile Glu Ser

20 25 30

ggc aca ggg att aaa gta aaa gat atc aat ggg aag gaa tat tac gac 144

Gly Thr Gly Ile Lys Val Lys Asp Ile Asn Gly Lys Glu Tyr Tyr Asp

35 40 45

ggc ttc agt tct gtt tgg ctg aac gtt cac ggg cac cgc aag aag gag 192

Gly Phe Ser Ser Val Trp Leu Asn Val His Gly His Arg Lys Lys Glu

50 55 60

ctt gat gac gca atc aag aag caa tta ggt aag att gcg cat agt acg 240

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

65 70 75 80

ctt tta gga atg acg aat gtc cct gca act caa tta gct gaa aca ctt 288

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

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

Cys Thr Thr Tyr Asp Val Leu Met Ile Val Asp Glu Val Ala Thr Gly	
245 250 255	
ttc ggt cgc acc ggg aaa atg ttt gca tgc gaa cat gaa aac gtg caa	816
Phe Gly Arg Thr Gly Lys Met Phe Ala Cys Glu His Glu Asn Val Gln	
260 265 270	
ccg gat ctg atg gcc gca ggc aag ggt atc acg ggc gga tac ctt ccg	864
Pro Asp Leu Met Ala Ala Gly Lys Gly Ile Thr Gly Gly Tyr Leu Pro	
275 280 285	
att gcg gtt act ttt gcc acc gaa gac att tat aag gca ttt tac gat	912
Ile Ala Val Thr Phe Ala Thr Glu Asp Ile Tyr Lys Ala Phe Tyr Asp	
290 295 300	
gat tat gaa aac ttg aag acc ttt ttt cat gga cac tct tac aca gga	960
Asp Tyr Glu Asn Leu Lys Thr Phe Phe His Gly His Ser Tyr Thr Gly	
305 310 315 320	
aat caa ctg ggt tgt gca gtc gca ctg gag aat ctg gca ctg ttt gaa	1008
Asn Gln Leu Gly Cys Ala Val Ala Leu Glu Asn Leu Ala Leu Phe Glu	
325 330 335	
agc gaa aac att gtt gag cag gtc gct gaa aaa tcg aag aaa tta cat	1056
Ser Glu Asn Ile Val Glu Gln Val Ala Glu Lys Ser Lys Lys Leu His	
340 345 350	
ttt tta tta caa gat tta cat gcc ttg cca cat gta ggc gac atc cgt	1104
Phe Leu Leu Gln Asp Leu His Ala Leu Pro His Val Gly Asp Ile Arg	
355 360 365	
caa tta gga ttc atg tgt ggt gcg gag tta gtc cgt agc aaa gaa aca	1152
Gln Leu Gly Phe Met Cys Gly Ala Glu Leu Val Arg Ser Lys Glu Thr	
370 375 380	
aaa gag ccc tat ccc gct gat cgt cgc atc ggt tac aaa gtc agt ctt	1200
Lys Glu Pro Tyr Pro Ala Asp Arg Arg Ile Gly Tyr Lys Val Ser Leu	
385 390 395 400	
aaa atg cgt gaa tta ggg atg ttg aca cgc ccg ttg gga gat gtt att	1248
Lys Met Arg Glu Leu Gly Met Leu Thr Arg Pro Leu Gly Asp Val Ile	

405 410 415
gca ttt ttg cct ccg tta gcg tct acc gcg gag gag ctg agt gag atg 1296
Ala Phe Leu Pro Pro Leu Ala Ser Thr Ala Glu Glu Leu Ser Glu Met
420 425 430
gta gca att atg aag caa gcc att cac gaa gtt act tcc ttg gaa gac 1344
Val Ala Ile Met Lys Gln Ala Ile His Glu Val Thr Ser Leu Glu Asp
435 440 445
tga 1347
<210> 97
<211> 448
<212> PRT
<213> Bacillus subtilis
<400> 97
Met Thr His Asp Leu Ile Glu Lys Ser Lys Lys His Leu Trp Leu Pro
1 5 10 15
Phe Thr Gln Met Lys Asp Tyr Asp Glu Asn Pro Leu Ile Ile Glu Ser
20 25 30
Gly Thr Gly Ile Lys Val Lys Asp Ile Asn Gly Lys Glu Tyr Tyr Asp
35 40 45
Gly Phe Ser Ser Val Trp Leu Asn Val His Gly His Arg Lys Lys Glu
50 55 60
Leu Asp Asp Ala Ile Lys Lys Gln Leu Gly Lys Ile Ala His Ser Thr
65 70 75 80
Leu Leu Gly Met Thr Asn Val Pro Ala Thr Gln Leu Ala Glu Thr Leu
85 90 95
Ile Asp Ile Ser Pro Lys Lys Leu Thr Arg Val Phe Tyr Ser Asp Ser
100 105 110
Gly Ala Glu Ala Met Glu Ile Ala Leu Lys Met Ala Phe Gln Tyr Trp
115 120 125
Lys Asn Ile Gly Lys Pro Glu Lys Gln Lys Phe Ile Ala Met Lys Asn
130 135 140

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

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

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

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

385	390	395	400
Lys Met Arg Glu Leu Gly Met Leu Thr Arg Pro Leu Gly Asp Val Ile			
405	410	415	
Ala Phe Leu Pro Pro Leu Ala Ser Thr Ala Glu Glu Leu Ser Glu Met			
420	425	430	
Val Ala Ile Met Lys Gln Ala Ile His Glu Val Thr Ser Leu Glu Asp			
435	440	445	
<210>	98		
<211>	771		
<212>	DNA		
<213>	Escherichia coli		
<220><221>	CDS		
<222>	(1)..(771)		
<223>	bioH gene encoding Pimeloyl-[acyl-carrier protein] methyl ester		
	esterase (BioH)		
<400>	98		
atg aat aac atc tgg tgg cag acc aaa ggt cag ggg aat gtt cat ctt			48
Met Asn Asn Ile Trp Trp Gln Thr Lys Gly Gln Gly Asn Val His Leu			
1	5	10	15
gtg ctg ctg cac gga tgg gga ctg aat gcc gaa gtg tgg cgt tgc att			96
Val Leu Leu His Gly Trp Gly Leu Asn Ala Glu Val Trp Arg Cys Ile			
20	25	30	
gac gag gaa ctt agc tgc cat ttt acg ctg cac ctt gtt gac ctg ccc			144
Asp Glu Glu Leu Ser Ser His Phe Thr Leu His Leu Val Asp Leu Pro			
35	40	45	
ggc ttc ggg cgt agc cgg gga ttt ggt gcg ctg tca ctt gct gat atg			192
Gly Phe Gly Arg Ser Arg Gly Phe Gly Ala Leu Ser Leu Ala Asp Met			
50	55	60	
gcc gaa gcc gtg ctg caa cag gca cct gat aaa gcc att tgg tta ggc			240
Ala Glu Ala Val Leu Gln Gln Ala Pro Asp Lys Ala Ile Trp Leu Gly			
65	70	75	80

tgg agt ctg ggc ggg ctg gtg gca agc cag att gcg tta acc cat ccc	288
Trp Ser Leu Gly Gly Leu Val Ala Ser Gln Ile Ala Leu Thr His Pro	
85 90 95	
gag cgt gtt cag gcg ctg gtc acc gtg gcg tcg tca cct tgt ttt agt	336
Glu Arg Val Gln Ala Leu Val Thr Val Ala Ser Ser Pro Cys Phe Ser	
100 105 110	
gct cgt gac gag tgg ccg ggg ata aaa ccg gac gtg ctg gcg gga ttt	384
Ala Arg Asp Glu Trp Pro Gly Ile Lys Pro Asp Val Leu Ala Gly Phe	
115 120 125	
cag cag caa ctc agt gat gat ttt cag cgt aca gtg gag cgg ttc ctg	432
Gln Gln Gln Leu Ser Asp Asp Phe Gln Arg Thr Val Glu Arg Phe Leu	
130 135 140	
gcg tta caa acc atg ggg act gaa acg gcg cgc cag gat gcg cgg gcg	480
Ala Leu Gln Thr Met Gly Thr Glu Thr Ala Arg Gln Asp Ala Arg Ala	
145 150 155 160	
ttg aag aaa acc gtt ctg gcg tta ccg atg ccg gag gtt gac gtg ctt	528
Leu Lys Lys Thr Val Leu Ala Leu Pro Met Pro Glu Val Asp Val Leu	
165 170 175	
aat ggc ggg ctg gaa atc ctg aaa acg gtc gat ctc cgt cag ccg ctg	576
Asn Gly Gly Leu Glu Ile Leu Lys Thr Val Asp Leu Arg Gln Pro Leu	
180 185 190	
caa aac gtg tcc atg ccg ttt ttg cga ttg tat ggc tat ctc gac ggt	624
Gln Asn Val Ser Met Pro Phe Leu Arg Leu Tyr Gly Tyr Leu Asp Gly	
195 200 205	
ctg gtg ccg cgc aaa gtg gtg ccg atg ctg gat aaa ctt tgg cct cac	672
Leu Val Pro Arg Lys Val Val Pro Met Leu Asp Lys Leu Trp Pro His	
210 215 220	
agc gaa tca tat atc ttc gcc aaa gcg gcc cat gcg cca ttt att tcg	720
Ser Glu Ser Tyr Ile Phe Ala Lys Ala Ala His Ala Pro Phe Ile Ser	
225 230 235 240	
cat ccg gcc gag ttt tgt cac ctg ctg gtg gcg ttg aag cag agg gtg	768

His Pro Ala Glu Phe Cys His Leu Leu Val Ala Leu Lys Gln Arg Val

245 250 255

tag

771

<210> 99

<211> 256

<212> PRT

<213> Escherichia coli

<400> 99

Met Asn Asn Ile Trp Trp Gln Thr Lys Gly Gln Gly Asn Val His Leu

1 5 10 15

Val Leu Leu His Gly Trp Gly Leu Asn Ala Glu Val Trp Arg Cys Ile

20 25 30

Asp Glu Glu Leu Ser Ser His Phe Thr Leu His Leu Val Asp Leu Pro

35 40 45

Gly Phe Gly Arg Ser Arg Gly Phe Gly Ala Leu Ser Leu Ala Asp Met

50 55 60

Ala Glu Ala Val Leu Gln Gln Ala Pro Asp Lys Ala Ile Trp Leu Gly

65 70 75 80

Trp Ser Leu Gly Gly Leu Val Ala Ser Gln Ile Ala Leu Thr His Pro

85 90 95

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

100 105 110

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

115 120 125

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

130 135 140

Ala Leu Gln Thr Met Gly Thr Glu Thr Ala Arg Gln Asp Ala Arg Ala

145 150 155 160

Leu Lys Lys Thr Val Leu Ala Leu Pro Met Pro Glu Val Asp Val Leu

165 170 175

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

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

245 250 255

<210> 100

<211> 777

<212> DNA

<213> Bacillus subtilis

<220><221> CDS

<222> (1)..(777)

<223> bioW gene encoding 6-carboxyhexanoate-CoA ligase (BioW)

<400> 100

atg caa gaa gag acg ttc tat tca gtg cgt atg cgc gct tca atg aat 48

Met Gln Glu Glu Thr Phe Tyr Ser Val Arg Met Arg Ala Ser Met Asn

1 5 10 15

ggc tcc cac gaa gat gga ggt aag cac atc tcc ggg ggt gag cgc ctt 96

Gly Ser His Glu Asp Gly Gly Lys His Ile Ser Gly Gly Glu Arg Leu

20 25 30

atc ccg ttc cac gag atg aaa cat acc gtc aac gct ttg ctt gag aag 144

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

35 40 45

ggc ctt tct cat tct cgt ggg aaa cct gat ttt atg caa att cag ttt 192

Gly Leu Ser His Ser Arg Gly Lys Pro Asp Phe Met Gln Ile Gln Phe

50 55 60

gaa gag gtt cac gag tca atc aag aca atc cag ccc tta cct gtg cac 240

Glu Glu Val His Glu Ser Ile Lys Thr Ile Gln Pro Leu Pro Val His

65 70 75 80

acc aac gaa gtt agc tgc ccc gaa gaa gga caa aaa ctt gca cgc ttg	288
Thr Asn Glu Val Ser Cys Pro Glu Glu Gly Gln Lys Leu Ala Arg Leu	
85 90 95	
tta ctg gag aaa gaa ggg gtg agc cgc gac gtt att gaa aag gct tac	336
Leu Leu Glu Lys Glu Gly Val Ser Arg Asp Val Ile Glu Lys Ala Tyr	
100 105 110	
gaa caa att ccc gag tgg tcg gat gtc cgt ggt gcc gta ttg ttt gat	384
Glu Gln Ile Pro Glu Trp Ser Asp Val Arg Gly Ala Val Leu Phe Asp	
115 120 125	
att cat acg ggc aag cgt atg gat cag acg aaa gaa aag ggg gtg cgc	432
Ile His Thr Gly Lys Arg Met Asp Gln Thr Lys Glu Lys Gly Val Arg	
130 135 140	
gtc tct cgt atg gac tgg ccc gac gct aac ttt gag aaa tgg gcc tta	480
Val Ser Arg Met Asp Trp Pro Asp Ala Asn Phe Glu Lys Trp Ala Leu	
145 150 155 160	
cac agc cac gtg cca gca cat tca cgc atc aag gag gca ctg gca ctt	528
His Ser His Val Pro Ala His Ser Arg Ile Lys Glu Ala Leu Ala Leu	
165 170 175	
gct agc aag gtg tcc cgt cac ccg gca gtc gtt gcc gaa ttg tgt tgg	576
Ala Ser Lys Val Ser Arg His Pro Ala Val Val Ala Glu Leu Cys Trp	
180 185 190	
agc gac gat cca gat tac atc acc gga tat gta gct ggt aaa aaa atg	624
Ser Asp Asp Pro Asp Tyr Ile Thr Gly Tyr Val Ala Gly Lys Lys Met	
195 200 205	
ggg tac caa cgc att acc gca atg aag gag tac ggg acc gag gag gga	672
Gly Tyr Gln Arg Ile Thr Ala Met Lys Glu Tyr Gly Thr Glu Glu Gly	
210 215 220	
tgt cgt gtc ttc ttc atc gac ggc tcg aac gat gtt aat act tac att	720
Cys Arg Val Phe Phe Ile Asp Gly Ser Asn Asp Val Asn Thr Tyr Ile	
225 230 235 240	

cac gac ttg gag aaa cag ccg atc ctg att gaa tgg gaa gaa gac cac 768

His Asp Leu Glu Lys Gln Pro Ile Leu Ile Glu Trp Glu Glu Asp His
245 250 255

gat agc tga 777

Asp Ser

<210> 101

<211> 258

<212> PRT

<213> Bacillus subtilis

<400> 101

Met Gln Glu Glu Thr Phe Tyr Ser Val Arg Met Arg Ala Ser Met Asn

1 5 10 15
Gly Ser His Glu Asp Gly Gly Lys His Ile Ser Gly Gly Glu Arg Leu

20 25 30
Ile Pro Phe His Glu Met Lys His Thr Val Asn Ala Leu Leu Glu Lys

35 40 45
Gly Leu Ser His Ser Arg Gly Lys Pro Asp Phe Met Gln Ile Gln Phe

50 55 60
Glu Glu Val His Glu Ser Ile Lys Thr Ile Gln Pro Leu Pro Val His

65 70 75 80
Thr Asn Glu Val Ser Cys Pro Glu Glu Gly Gln Lys Leu Ala Arg Leu

85 90 95
Leu Leu Glu Lys Glu Gly Val Ser Arg Asp Val Ile Glu Lys Ala Tyr

100 105 110
Glu Gln Ile Pro Glu Trp Ser Asp Val Arg Gly Ala Val Leu Phe Asp

115 120 125
Ile His Thr Gly Lys Arg Met Asp Gln Thr Lys Glu Lys Gly Val Arg

130 135 140
Val Ser Arg Met Asp Trp Pro Asp Ala Asn Phe Glu Lys Trp Ala Leu

145 150 155 160

His Ser His Val Pro Ala His Ser Arg Ile Lys Glu Ala Leu Ala Leu
165 170 175
Ala Ser Lys Val Ser Arg His Pro Ala Val Val Ala Glu Leu Cys Trp
180 185 190
Ser Asp Asp Pro Asp Tyr Ile Thr Gly Tyr Val Ala Gly Lys Lys Met
195 200 205
Gly Tyr Gln Arg Ile Thr Ala Met Lys Glu Tyr Gly Thr Glu Glu Gly
210 215 220
Cys Arg Val Phe Phe Ile Asp Gly Ser Asn Asp Val Asn Thr Tyr Ile
225 230 235 240
His Asp Leu Glu Lys Gln Pro Ile Leu Ile Glu Trp Glu Glu Asp His
245 250 255
Asp Ser

<210> 102

<211> 966

<

212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(966)

<223> lipA gene encoding lipoic acid synthase (LipA)

<400> 102

atg agt aaa ccc att gtg atg gaa cgc ggt gtt aaa tac cgc gat gcc 48

Met Ser Lys Pro Ile Val Met Glu Arg Gly Val Lys Tyr Arg Asp Ala

1 5 10 15

gat aag atg gcc ctt atc ccg gtt aaa aac gtg gca aca gag cgc gaa 96

Asp Lys Met Ala Leu Ile Pro Val Lys Asn Val Ala Thr Glu Arg Glu

20 25 30

gcc ctg ctg cgc aag ccg gaa tgg atg aaa atc aag ctt cca gcg gac 144

Ala Leu Leu Arg Lys Pro Glu Trp Met Lys Ile Lys Leu Pro Ala Asp

35 40 45

tct aca cgt atc cag ggc atc aaa gcc gca atg cgc aaa aat ggc ctg 192

Ser	Thr	Arg	Ile	Gln	Gly	Ile	Lys	Ala	Ala	Met	Arg	Lys	Asn	Gly	Leu		
50						55						60					
cat	tct	gtc	tgc	gag	gaa	gcc	tcc	tgc	cct	aac	ctg	gcg	gaa	tgc	ttc	240	
His	Ser	Val	Cys	Glu	Glu	Ala	Ser	Cys	Pro	Asn	Leu	Ala	Glu	Cys	Phe		
65			70			75			80								
aac	cac	ggc	aca	gca	acg	ttt	atg	atc	ctc	ggc	gct	att	tgt	acc	cgc	288	
Asn	His	Gly	Thr	Ala	Thr	Phe	Met	Ile	Leu	Gly	Ala	Ile	Cys	Thr	Arg		
			85			90			95								
cgt	tgt	ccg	ttc	tgt	gat	gtt	gcc	cac	ggg	cgc	ccg	gta	gct	cct	gat	336	
Arg	Cys	Pro	Phe	Cys	Asp	Val	Ala	His	Gly	Arg	Pro	Val	Ala	Pro	Asp		
			100			105			110								
gcc	aat	gaa	cca	gtg	aaa	ctg	gcg	cag	acc	att	gcc	gat	atg	gcg	ctg	384	
Ala	Asn	Glu	Pro	Val	Lys	Leu	Ala	Gln	Thr	Ile	Ala	Asp	Met	Ala	Leu		
115			120			125											
cgt	tat	gtg	gtt	atc	acc	tcc	gtt	gac	cgt	gat	gac	ctg	cgc	gat	ggc	432	
Arg	Tyr	Val	Val	Ile	Thr	Ser	Val	Asp	Arg	Asp	Asp	Leu	Arg	Asp	Gly		
130			135			140											
ggg	gcc	cag	cac	ttt	gcg	gat	tgc	att	act	gcc	att	cgg	gaa	aaa	agc	480	
Gly	Ala	Gln	His	Phe	Ala	Asp	Cys	Ile	Thr	Ala	Ile	Arg	Glu	Lys	Ser		
145			150			155			160								
ccg	caa	atc	aaa	att	gaa	act	ctg	gtg	ccg	gat	ttc	cgc	ggg	cgt	atg	528	
Pro	Gln	Ile	Lys	Ile	Glu	Thr	Leu	Val	Pro	Asp	Phe	Arg	Gly	Arg	Met		
			165			170			175								
gat	cgt	gct	ctg	gat	att	ctg	act	gca	acg	cca	cca	gat	gtg	ttc	aac	576	
Asp	Arg	Ala	Leu	Asp	Ile	Leu	Thr	Ala	Thr	Pro	Pro	Asp	Val	Phe	Asn		
180			185			190											
cat	aac	ctg	gaa	aac	gta	ccg	cgt	att	tac	cgt	cag	gta	cgg	cct	ggg	624	
His	Asn	Leu	Glu	Asn	Val	Pro	Arg	Ile	Tyr	Arg	Gln	Val	Arg	Pro	Gly		
195			200			205											

gca gat tac aac tgg tgc ctg aag ctg ctg gaa cgc ttt aaa gaa gcg 672

Ala Asp Tyr Asn Trp Ser Leu Lys Leu Leu Glu Arg Phe Lys Glu Ala

210 215 220

cat ccg gaa atc ccg acc aag tct ggt ctg atg gtg gga ctg ggt gaa 720

His Pro Glu Ile Pro Thr Lys Ser Gly Leu Met Val Gly Leu Gly Glu

225 230 235 240

acc aat gaa gaa att att gag gta atg cgc gac ctg cgc cgt cat ggt 768

Thr Asn Glu Glu Ile Ile Glu Val Met Arg Asp Leu Arg Arg His Gly

245 250 255

gtg acg atg tta acg ctg ggg caa tat ttg cag cca agc cgc cat cac 816

Val Thr Met Leu Thr Leu Gly Gln Tyr Leu Gln Pro Ser Arg His His

260 265 270

ctg ccg gtt caa cgt tac gtt agc ccg gat gag ttc gac gaa atg aaa 864

Leu Pro Val Gln Arg Tyr Val Ser Pro Asp Glu Phe Asp Glu Met Lys

275 280 285

gcc gaa gcg ctg gcg atg ggc ttt acc cat gct gca tgc ggt ccg ttt 912

Ala Glu Ala Leu Ala Met Gly Phe Thr His Ala Ala Cys Gly Pro Phe

290 295 300

gtc cgc tct tct tac cac gcc gat ttg cag gcg aaa ggg atg gaa gtt 960

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

305 310 315 320

aag taa 966

Lys

<210> 103

<211> 321

<212> PRT

<213> Escherichia coli

<400> 103

Met Ser Lys Pro Ile Val Met Glu Arg Gly Val Lys Tyr Arg Asp Ala

1 5 10 15

Asp Lys Met Ala Leu Ile Pro Val Lys Asn Val Ala Thr Glu Arg Glu
 20 25 30
 Ala Leu Leu Arg Lys Pro Glu Trp Met Lys Ile Lys Leu Pro Ala Asp
 35 40 45
 Ser Thr Arg Ile Gln Gly Ile Lys Ala Ala Met Arg Lys Asn Gly Leu
 50 55 60
 His Ser Val Cys Glu Glu Ala Ser Cys Pro Asn Leu Ala Glu Cys Phe
 65 70 75 80
 Asn His Gly Thr Ala Thr Phe Met Ile Leu Gly Ala Ile Cys Thr Arg
 85 90 95
 Arg Cys Pro Phe Cys Asp Val Ala His Gly Arg Pro Val Ala Pro Asp
 100 105 110
 Ala Asn Glu Pro Val Lys Leu Ala Gln Thr Ile Ala Asp Met Ala Leu
 115 120 125
 Arg Tyr Val Val Ile Thr Ser Val Asp Arg Asp Asp Leu Arg Asp Gly
 130 135 140
 Gly Ala Gln His Phe Ala Asp Cys Ile Thr Ala Ile Arg Glu Lys Ser
 145 150 155 160
 Pro Gln Ile Lys Ile Glu Thr Leu Val Pro Asp Phe Arg Gly Arg Met
 165 170 175
 Asp Arg Ala Leu Asp Ile Leu Thr Ala Thr Pro Pro Asp Val Phe Asn
 180 185 190
 His Asn Leu Glu Asn Val Pro Arg Ile Tyr Arg Gln Val Arg Pro Gly
 195 200 205
 Ala Asp Tyr Asn Trp Ser Leu Lys Leu Leu Glu Arg Phe Lys Glu Ala
 210 215 220
 His Pro Glu Ile Pro Thr Lys Ser Gly Leu Met Val Gly Leu Gly Glu
 225 230 235 240
 Thr Asn Glu Glu Ile Ile Glu Val Met Arg Asp Leu Arg Arg His Gly
 245 250 255
 Val Thr Met Leu Thr Leu Gly Gln Tyr Leu Gln Pro Ser Arg His His

260 265 270
 Leu Pro Val Gln Arg Tyr Val Ser Pro Asp Glu Phe Asp Glu Met Lys
 275 280 285
 Ala Glu Ala Leu Ala Met Gly Phe Thr His Ala Ala Cys Gly Pro Phe

290 295 300
 Val Arg Ser Ser Tyr His Ala Asp Leu Gln Ala Lys Gly Met Glu Val
 305 310 315 320
 Lys

<210> 104
 <211> 897
 <212> DNA
 <213> Bacillus subtilis subsp. Subtilis, str. 168
 <220><221> CDS
 <222> (1)..(897)
 <223> lipA gene encoding a lipoic acid synthase (LipA)
 <400> 104

atg gcg aag aag gat gaa cac ctg aga aag cca gaa tgg ctt aaa att 48
 Met Ala Lys Lys Asp Glu His Leu Arg Lys Pro Glu Trp Leu Lys Ile

1 5 10 15
 aag ctg aat aca aac gaa aac tac act ggc tta aaa aag tta atg cgt 96
 Lys Leu Asn Thr Asn Glu Asn Tyr Thr Gly Leu Lys Lys Leu Met Arg

20 25 30
 gag aat aac tta cat act gtc tgt gag gag gca aaa tgt cca aat ata 144
 Glu Asn Asn Leu His Thr Val Cys Glu Glu Ala Lys Cys Pro Asn Ile

35 40 45
 cac gaa tgc tgg gcc gtt cgg cgt acc gcg acg ttt atg ata ctg ggc 192
 His Glu Cys Trp Ala Val Arg Arg Thr Ala Thr Phe Met Ile Leu Gly

50 55 60
 tcc gtc tgc acg aga gca tgt cgt ttt tgc gcg gtt aaa acc ggc ctg 240
 Ser Val Cys Thr Arg Ala Cys Arg Phe Cys Ala Val Lys Thr Gly Leu

65 70 75 80

ccg act gag ctt gac ttg caa gag cca gag cgc gtg gct gat tca gtt 288

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

85 90 95

gcc ctt atg aac ctg aaa cac gcc gtt atc acg gcg gtc gcc cgt gac 336

Ala Leu Met Asn Leu Lys His Ala Val Ile Thr Ala Val Ala Arg Asp

100 105 110

gat caa aaa gat ggt gga gcg gga ata ttc gca gaa acg gta cgt gct 384

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

115 120 125

atc cgc cgg aag tct cca ttt acc acg att gaa gtg ctg ccg agc gat 432

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

130 135 140

atg gcc ggt aat tat gat aac ctt aag acc ttg atg gac aca cgg ccg 480

Met Gly Gly Asn Tyr Asp Asn Leu Lys Thr Leu Met Asp Thr Arg Pro

145 150 155 160

gat att ctg aat cac aac atc gag act gta cgg cgt tta aca cca aga 528

Asp Ile Leu Asn His Asn Ile Glu Thr Val Arg Arg Leu Thr Pro Arg

165 170 175

gtc aga gca cgc gct acc tat gat cgc agc ctt gaa ttc ttg cgt cgc 576

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

180 185 190

gcc aaa gag atg cag ccc gac ata cca aca aaa agt agt ata atg atc 624

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

195 200 205

ggc ttg gga gaa aca aaa gaa gaa atc atc gag gtc atg gat gac ctt 672

Gly Leu Gly Glu Thr Lys Glu Glu Ile Ile Glu Val Met Asp Asp Leu

210 215 220

ttg gca aac aac gtg gac ata atg gcc att ggg caa tac ttg caa cca 720

Leu Ala Asn Asn Val Asp Ile Met Ala Ile Gly Gln Tyr Leu Gln Pro

225 230 235 240
act aaa aag cac tta aaa gtt cag aaa tac tat cat cct gat gaa ttt 768
Thr Lys Lys His Leu Lys Val Gln Lys Tyr Tyr His Pro Asp Glu Phe
 245 250 255
gcc gag ttg aag gaa atc gcc atg cag aag ggg ttt tca cat tgc gag 816
Ala Glu Leu Lys Glu Ile Ala Met Gln Lys Gly Phe Ser His Cys Glu
 260 265 270
gcg ggt ccg ttg gtc cgt tca agt tac cac gcg gac gaa cag gtg aat 864
Ala Gly Pro Leu Val Arg Ser Ser Tyr His Ala Asp Glu Gln Val Asn
 275 280 285
gaa gcg tca aag aaa cgc caa gca caa gct taa 897
Glu Ala Ser Lys Lys Arg Gln Ala Gln Ala
 290 295
<210> 105
<211> 298
<212> PRT
<213> Bacillus subtilis subsp. Subtilis, str. 168

<400> 105
Met Ala Lys Lys Asp Glu His Leu Arg Lys Pro Glu Trp Leu Lys Ile
1 5 10 15
Lys Leu Asn Thr Asn Glu Asn Tyr Thr Gly Leu Lys Lys Leu Met Arg
 20 25 30
Glu Asn Asn Leu His Thr Val Cys Glu Glu Ala Lys Cys Pro Asn Ile
 35 40 45
His Glu Cys Trp Ala Val Arg Arg Thr Ala Thr Phe Met Ile Leu Gly
 50 55 60

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

Asp Gln Lys Asp Gly Gly Ala Gly Ile Phe Ala Glu Thr Val Arg Ala
 115 120 125

Ile Arg Arg Lys Ser Pro Phe Thr Thr Ile Glu Val Leu Pro Ser Asp
 130 135 140

Met Gly Gly Asn Tyr Asp Asn Leu Lys Thr Leu Met Asp Thr Arg Pro
 145 150 155 160

Asp Ile Leu Asn His Asn Ile Glu Thr Val Arg Arg Leu Thr Pro Arg
 165 170 175

Val Arg Ala Arg Ala Thr Tyr Asp Arg Ser Leu Glu Phe Leu Arg Arg
 180 185 190

Ala Lys Glu Met Gln Pro Asp Ile Pro Thr Lys Ser Ser Ile Met Ile
 195 200 205

Gly Leu Gly Glu Thr Lys Glu Glu Ile Ile Glu Val Met Asp Asp Leu
 210 215 220

Leu Ala Asn Asn Val Asp Ile Met Ala Ile Gly Gln Tyr Leu Gln Pro
 225 230 235 240

Thr Lys Lys His Leu Lys Val Gln Lys Tyr Tyr His Pro Asp Glu Phe
 245 250 255

Ala Glu Leu Lys Glu Ile Ala Met Gln Lys Gly Phe Ser His Cys Glu
 260 265 270

Ala Gly Pro Leu Val Arg Ser Ser Tyr His Ala Asp Glu Gln Val Asn
 275 280 285

Glu Ala Ser Lys Lys Arg Gln Ala Gln Ala
 290 295

<210> 106

<211> 1245

<212> DNA

<213> *Saccharomyces cerevisiae*, S288C

<220><221> CDS

<222> (1)..(1245)

<223> lipA gene encoding a lipoic acid synthase (LipA)

<400> 106

atg tac aga cgt agt gta ggg gtg ctt ttc gta ggg cgt aac act cgg 48

Met Tyr Arg Arg Ser Val Gly Val Leu Phe Val Gly Arg Asn Thr Arg

1 5 10 15

tgg atc agc agc acg atc cgg tgt ggc act agc gca acc cgc cct att 96

Trp Ile Ser Ser Thr Ile Arg Cys Gly Thr Ser Ala Thr Arg Pro Ile

20 25 30

cgt agt aac gcg ttg aac act gac tca gat aac gct agt gtg cgc gta 144

Arg Ser Asn Ala Leu Asn Thr Asp Ser Asp Asn Ala Ser Val Arg Val

35 40 45

ccc gta ggg aac agc acg gag gta gaa aat gcg acc tcg caa ctg aca 192

Pro Val Gly Asn Ser Thr Glu Val Glu Asn Ala Thr Ser Gln Leu Thr

50 55 60

ggt act tcg ggc aag aga cgg aag gga aat aga aag cgg ata acg gaa 240

Gly Thr Ser Gly Lys Arg Arg Lys Gly Asn Arg Lys Arg Ile Thr Glu

65 70 75 80

ttt aaa gat gcc tta aac ctg ggt ccc tcg ttc gcc gat ttc gta tca 288

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

85 90 95

ggt aag gct tcg aaa atg att ttg gat cct ttg gag aaa gcc cgc cag 336

Gly Lys Ala Ser Lys Met Ile Leu Asp Pro Leu Glu Lys Ala Arg Gln

100 105 110

aac aca gag gag gct aag aag tta cct aga tgg ttg aaa gtg cct att 384

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

115 120 125

cct aaa ggc acg aac tac cac aaa ttg aaa ggt gat gta aaa gaa tta 432

Pro Lys Gly Thr Asn Tyr His Lys Leu Lys Gly Asp Val Lys Glu Leu

130 135 140

gga ctt tcc act gtc tgt gag gaa gct cgc tgc cca aat atc gga gag 480

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

145 150 155 160

tgc tgg gga gga aag gac aaa tcc aag gct acg gcg acg ata atg ttg	528
Cys Trp Gly Gly Lys Asp Lys Ser Lys Ala Thr Ala Thr Ile Met Leu	
165 170 175	
ctt ggc gat acg tgc acg cgc ggt tgc cgt ttt tgc tct gta aag act	576
Leu Gly Asp Thr Cys Thr Arg Gly Cys Arg Phe Cys Ser Val Lys Thr	
180 185 190	
aac cgc aca ccc tcg aaa cct gat cct atg gag ccc gaa aat aca gca	624
Asn Arg Thr Pro Ser Lys Pro Asp Pro Met Glu Pro Glu Asn Thr Ala	
195 200 205	
gag gcg att aaa cgt tgg ggc ctg ggc tat gtt gta tta act aca gta	672
Glu Ala Ile Lys Arg Trp Gly Leu Gly Tyr Val Val Leu Thr Thr Val	
210 215 220	
gat cgt gat gat tta gtc gac ggt ggg gca aat cat ttg gcg gag act	720
Asp Arg Asp Asp Leu Val Asp Gly Gly Ala Asn His Leu Ala Glu Thr	
225 230 235 240	
gtc aga aag atc aag caa aaa gcg ccg aat act ctt gtg gag act ctt	768
Val Arg Lys Ile Lys Gln Lys Ala Pro Asn Thr Leu Val Glu Thr Leu	
245 250 255	
tcc ggt gac ttt aga ggg gac tta aag atg gtt gac atc atg gca caa	816
Ser Gly Asp Phe Arg Gly Asp Leu Lys Met Val Asp Ile Met Ala Gln	
260 265 270	
tgc ggc tta gat gtt tac gct cat aac tta gaa acc gtt gaa agt tta	864
Cys Gly Leu Asp Val Tyr Ala His Asn Leu Glu Thr Val Glu Ser Leu	
275 280 285	
act ccc cac gta aga gac cgg cgg gcc aca tac aga caa agt ctg agc	912
Thr Pro His Val Arg Asp Arg Arg Ala Thr Tyr Arg Gln Ser Leu Ser	
290 295 300	
gtt ttg gag cgt gcc aaa gcc acg gtg cca tca tta atc act aaa acc	960
Val Leu Glu Arg Ala Lys Ala Thr Val Pro Ser Leu Ile Thr Lys Thr	
305 310 315 320	
tct ata atg tta ggc tta ggg gaa acg gac gaa caa ata acg caa acc	1008

Ser Ile Met Leu Gly Leu Gly Glu Thr Asp Glu Gln Ile Thr Gln Thr
325 330 335

tta aaa gat tta cgg aac att caa tgc gac gtc gtg acc ttt ggt caa 1056
Leu Lys Asp Leu Arg Asn Ile Gln Cys Asp Val Val Thr Phe Gly Gln
340 345 350

tat atg cgg cca aca aag cgg cac atg aag gtg gtc gag tac gta aaa 1104
Tyr Met Arg Pro Thr Lys Arg His Met Lys Val Val Glu Tyr Val Lys
355 360 365

ccc gaa aaa ttt gat tac tgg aag gag cgg gct ctt gag atg ggt ttc 1152

Pro Glu Lys Phe Asp Tyr Trp Lys Glu Arg Ala Leu Glu Met Gly Phe
370 375 380

tta tat tgc gcc tct ggg cca ctt gta cgc tct agc tat aaa gcg ggt 1200
Leu Tyr Cys Ala Ser Gly Pro Leu Val Arg Ser Ser Tyr Lys Ala Gly
385 390 395 400

gag gca ttt atc gag aat gtt tta aaa aaa aga aac atg aag tga 1245
Glu Ala Phe Ile Glu Asn Val Leu Lys Lys Arg Asn Met Lys

405 410

<210> 107
<211> 414
<212> PRT
<213> Saccharomyces cerevisiae, S288C
<400> 107

Met Tyr Arg Arg Ser Val Gly Val Leu Phe Val Gly Arg Asn Thr Arg
1 5 10 15

Trp Ile Ser Ser Thr Ile Arg Cys Gly Thr Ser Ala Thr Arg Pro Ile
20 25 30

Arg Ser Asn Ala Leu Asn Thr Asp Ser Asp Asn Ala Ser Val Arg Val
35 40 45

Pro Val Gly Asn Ser Thr Glu Val Glu Asn Ala Thr Ser Gln Leu Thr
50 55 60

Gly Thr Ser Gly Lys Arg Arg Lys Gly Asn Arg Lys Arg Ile Thr Glu

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

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

Pro Glu Lys Phe Asp Tyr Trp Lys Glu Arg Ala Leu Glu Met Gly Phe
 370 375 380
 Leu Tyr Cys Ala Ser Gly Pro Leu Val Arg Ser Ser Tyr Lys Ala Gly
 385 390 395 400
 Glu Ala Phe Ile Glu Asn Val Leu Lys Lys Arg Asn Met Lys
 405 410

<210> 108

<211> 1017

<212> DNA

<213> Pseudomonas putida KT2440

<220><221> CDS

<222> (1)..(1017)

<223> lipA gene encoding a lipoic acid synthase (LipA)

<400> 108

atg acg acc gtt cag gaa gcc gtc ccg aat ttg atc ccc acg caa gat	48
Met Thr Thr Val Gln Glu Ala Val Pro Asn Leu Ile Pro Thr Gln Asp	
1 5 10 15	
gct acg ccg cgc ccg gct ccc aaa aag gtt gaa gct ggt gtg aag ttg	96
Ala Thr Pro Arg Pro Ala Pro Lys Lys Val Glu Ala Gly Val Lys Leu	
20 25 30	
cgc gga gcg gat aaa gtt gcc cgc atc cct gtg aaa att att ccg aca	144

Arg Gly Ala Asp Lys Val Ala Arg Ile Pro Val Lys Ile Ile Pro Thr	
35 40 45	
gat gag tta ccg aaa aaa cct gac tgg atc cgc gtc cgc atc cct gta	192
Asp Glu Leu Pro Lys Lys Pro Asp Trp Ile Arg Val Arg Ile Pro Val	
50 55 60	

tcg ccc gag gtt gat cgt atc aaa caa ttg ttg cgt aag cac aaa ttg	240
Ser Pro Glu Val Asp Arg Ile Lys Gln Leu Leu Arg Lys His Lys Leu	
65 70 75 80	
cat agc gtc tgc gag gag gcc tcc tgc cca aat ttg ggc gag tgc ttt	288
His Ser Val Cys Glu Glu Ala Ser Cys Pro Asn Leu Gly Glu Cys Phe	
85 90 95	
agc ggc ggc acg gct act ttc atg att atg ggc gac atc tgt aca cgt	336
Ser Gly Gly Thr Ala Thr Phe Met Ile Met Gly Asp Ile Cys Thr Arg	
100 105 110	
cgt tgt cct ttt tgc gac gtg gga cac gga cgc cca aag ccg ctg gat	384
Arg Cys Pro Phe Cys Asp Val Gly His Gly Arg Pro Lys Pro Leu Asp	
115 120 125	
ttg gat gag cca aaa aat ctt gca gtt gcg att gca gac ttg cgc tta	432
Leu Asp Glu Pro Lys Asn Leu Ala Val Ala Ile Ala Asp Leu Arg Leu	
130 135 140	
aag tac gtg gtt atc aca tcg gtt gat cgt gac gat tta cgt gac ggg	480
Lys Tyr Val Val Ile Thr Ser Val Asp Arg Asp Asp Leu Arg Asp Gly	
145 150 155 160	
ggc gct caa cat ttt gcc gac tgc atc cgt gag atc cgc gca ctg tcc	528
Gly Ala Gln His Phe Ala Asp Cys Ile Arg Glu Ile Arg Ala Leu Ser	
165 170 175	
ccg ggg gtg cag ctg gag act ttg gtt ccg gac tac cgt gga cgc atg	576
Pro Gly Val Gln Leu Glu Thr Leu Val Pro Asp Tyr Arg Gly Arg Met	
180 185 190	
gat gtt gca ttg gag att acc gcc cag gag cct cca gat gtg ttc aac	624
Asp Val Ala Leu Glu Ile Thr Ala Gln Glu Pro Pro Asp Val Phe Asn	
195 200 205	
cat aat ctt gag aca gtc cca cgc tta tat aag gct gct cgt ccg ggt	672
His Asn Leu Glu Thr Val Pro Arg Leu Tyr Lys Ala Ala Arg Pro Gly	

210	215	220	
tct gat tac gac tgg agt tta gac ttg ttg caa aaa ttt aag cag ctg			720
Ser Asp Tyr Asp Trp Ser Leu Asp Leu Leu Gln Lys Phe Lys Gln Leu			
225	230	235	240
gtt ccg cat gtg cca act aag tct gga ctg atg tta gga tta gga gaa			768
Val Pro His Val Pro Thr Lys Ser Gly Leu Met Leu Gly Leu Gly Glu			
	245	250	255
aca gat gag gaa gtc att gaa gta atg cat cgt atg cgt gag cat gat			816
Thr Asp Glu Glu Val Ile Glu Val Met His Arg Met Arg Glu His Asp			
	260	265	270
atc gac atg ttg act ctg gga cag tac ctt cag ccc tcg cgt tcg cat			864
Ile Asp Met Leu Thr Leu Gly Gln Tyr Leu Gln Pro Ser Arg Ser His			
	275	280	285
ctt cct gtt cag cgt ttt gtt cat ccc gat act ttt gct tgg ttc gcg			912
Leu Pro Val Gln Arg Phe Val His Pro Asp Thr Phe Ala Trp Phe Ala			
	290	295	300
gaa gag gga tac aag atg ggg ttc aaa aat gtc gct tct gga cca ttg			960
Glu Glu Gly Tyr Lys Met Gly Phe Lys Asn Val Ala Ser Gly Pro Leu			
305	310	315	320
gta cgc tcg tca tat cac gca gac cag cag gct cat gag gcc aaa att			1008
Val Arg Ser Ser Tyr His Ala Asp Gln Gln Ala His Glu Ala Lys Ile			
	325	330	335
aag ctt tga			1017
Lys Leu			

<210> 109

<211> 338

<212> PRT

<213> Pseudomonas putida KT2440

<400> 109

Met Thr Thr Val Gln Glu Ala Val Pro Asn Leu Ile Pro Thr Gln Asp

1 5 10 15
 Ala Thr Pro Arg Pro Ala Pro Lys Lys Val Glu Ala Gly Val Lys Leu

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

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

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

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

Thr Asp Glu Glu Val Ile Glu Val Met His Arg Met Arg Glu His Asp
 260 265 270
 Ile Asp Met Leu Thr Leu Gly Gln Tyr Leu Gln Pro Ser Arg Ser His
 275 280 285
 Leu Pro Val Gln Arg Phe Val His Pro Asp Thr Phe Ala Trp Phe Ala
 290 295 300
 Glu Glu Gly Tyr Lys Met Gly Phe Lys Asn Val Ala Ser Gly Pro Leu
 305 310 315 320
 Val Arg Ser Ser Tyr His Ala Asp Gln Gln Ala His Glu Ala Lys Ile
 325 330 335
 Lys Leu

<210> 110

<211> 867

<

212> DNA

<213> Bacteroides fragilis 638R

<220><221> CDS

<222> (1)..(867)

<223> lipA gene encoding a lipoic acid synthase (LipA)

<400> 110

atg ggg aac gac aag cgc gtt cgc aag cct gag tgg tta aaa att tct 48

Met Gly Asn Asp Lys Arg Val Arg Lys Pro Glu Trp Leu Lys Ile Ser

1 5 10 15

att ggt gca aat gag cgc tac acc gag act aaa cgt atc gtg gaa agc 96

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

20 25 30

cat tgt ctt cac acc atc tgc agt tct ggg cgc tgc ccg aat atg ggg 144

His Cys Leu His Thr Ile Cys Ser Ser Gly Arg Cys Pro Asn Met Gly

35 40 45

gaa tgt tgg ggg aaa ggg aca gca acc ttt atg atc gct ggt gac atc 192

Glu Cys Trp Gly Lys Gly Thr Ala Thr Phe Met Ile Ala Gly Asp Ile

50	55	60	
tgc act cgc tct tgc aag ttc tgt aat acc caa acc ggg cgc ccc tta			240
Cys Thr Arg Ser Cys Lys Phe Cys Asn Thr Gln Thr Gly Arg Pro Leu			
65	70	75	80
cct tta gac ccg gat gaa ccc acc cac gtt gcc gaa tct att gca tta			288
Pro Leu Asp Pro Asp Glu Pro Thr His Val Ala Glu Ser Ile Ala Leu			
	85	90	95
atg aag ctg tca cat gca gtc att aca agc gta gac cgt gac gac ctt			336
Met Lys Leu Ser His Ala Val Ile Thr Ser Val Asp Arg Asp Asp Leu			
	100	105	110
ccg gac tta gga gca gca cat tgg gct cag act atc cgc gag atc aag			384
Pro Asp Leu Gly Ala Ala His Trp Ala Gln Thr Ile Arg Glu Ile Lys			
	115	120	125
cgt ttg aat ccg gaa act acc aca gag gtt tta att cct gac ttt cag			432
Arg Leu Asn Pro Glu Thr Thr Thr Glu Val Leu Ile Pro Asp Phe Gln			
	130	135	140
gga cgt aag gaa ctt atc gac caa gtc att aag gcg tgt ccc gaa att			480
Gly Arg Lys Glu Leu Ile Asp Gln Val Ile Lys Ala Cys Pro Glu Ile			
145	150	155	160
att tca cat aac atg gaa acg gtc aaa cgc att tcg ccg cag gtt cgt			528
Ile Ser His Asn Met Glu Thr Val Lys Arg Ile Ser Pro Gln Val Arg			
	165	170	175
tct gca gcg aat tac cac act agt ctt gaa gtc att cgt cag att gct			576
Ser Ala Ala Asn Tyr His Thr Ser Leu Glu Val Ile Arg Gln Ile Ala			
	180	185	190
gaa agc ggg atc act gct aaa tcg ggc att atg gtt ggg ttg ggt gag			624
Glu Ser Gly Ile Thr Ala Lys Ser Gly Ile Met Val Gly Leu Gly Glu			
	195	200	205
act ccc gcc gaa gtc gaa gag ctt atg gac gac ttg atc tca gtc ggt			672

Thr Pro Ala Glu Val Glu Glu Leu Met Asp Asp Leu Ile Ser Val Gly
210 215 220
tgc aaa atc ctg acc atc ggt caa tat ctt caa cct aca cat aag cat 720
Cys Lys Ile Leu Thr Ile Gly Gln Tyr Leu Gln Pro Thr His Lys His
225 230 235 240
ttc ccg gtt gct gct tac att acc cca gaa cag ttc gcc gtc tat aag 768
Phe Pro Val Ala Ala Tyr Ile Thr Pro Glu Gln Phe Ala Val Tyr Lys
245 250 255
gag acg ggc ttg aag aaa ggt ttt gag cag gtg gag tca gcg ccc ctt 816
Glu Thr Gly Leu Lys Lys Gly Phe Glu Gln Val Glu Ser Ala Pro Leu
260 265 270
gtg cgc tct tct tat cac gca gaa aaa cac atc cgc ttt aat aac aag 864
Val Arg Ser Ser Tyr His Ala Glu Lys His Ile Arg Phe Asn Asn Lys
275 280 285
taa 867
<210> 111
<211> 288
<212> PRT
<213> Bacteroides fragilis 638R
<400> 111
Met Gly Asn Asp Lys Arg Val Arg Lys Pro Glu Trp Leu Lys Ile Ser
1 5 10 15
Ile Gly Ala Asn Glu Arg Tyr Thr Glu Thr Lys Arg Ile Val Glu Ser
20 25 30
His Cys Leu His Thr Ile Cys Ser Ser Gly Arg Cys Pro Asn Met Gly
35 40 45
Glu Cys Trp Gly Lys Gly Thr Ala Thr Phe Met Ile Ala Gly Asp Ile
50 55 60
Cys Thr Arg Ser Cys Lys Phe Cys Asn Thr Gln Thr Gly Arg Pro Leu
65 70 75 80
Pro Leu Asp Pro Asp Glu Pro Thr His Val Ala Glu Ser Ile Ala Leu
85 90 95

Met Lys Leu Ser His Ala Val Ile Thr Ser Val Asp Arg Asp Asp Leu
100 105 110

Pro Asp Leu Gly Ala Ala His Trp Ala Gln Thr Ile Arg Glu Ile Lys
115 120 125

Arg Leu Asn Pro Glu Thr Thr Thr Glu Val Leu Ile Pro Asp Phe Gln
130 135 140

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

Ile Ser His Asn Met Glu Thr Val Lys Arg Ile Ser Pro Gln Val Arg
165 170 175

Ser Ala Ala Asn Tyr His Thr Ser Leu Glu Val Ile Arg Gln Ile Ala
180 185 190

Glu Ser Gly Ile Thr Ala Lys Ser Gly Ile Met Val Gly Leu Gly Glu
195 200 205

Thr Pro Ala Glu Val Glu Glu Leu Met Asp Asp Leu Ile Ser Val Gly
210 215 220

Cys Lys Ile Leu Thr Ile Gly Gln Tyr Leu Gln Pro Thr His Lys His
225 230 235 240

Phe Pro Val Ala Ala Tyr Ile Thr Pro Glu Gln Phe Ala Val Tyr Lys
245 250 255

Glu Thr Gly Leu Lys Lys Gly Phe Glu Gln Val Glu Ser Ala Pro Leu
260 265 270

Val Arg Ser Ser Tyr His Ala Glu Lys His Ile Arg Phe Asn Asn Lys
275 280 285

<210> 112

<211> 954

<212> DNA

<213> Streptomyces coelicolor A3(2)

<220><221> CDS

<222> (1)..(954)

<223> lipA gene encoding a lipoic acid synthase (LipA)

<400> 112

atg agc gcg gta gct cct gac ggc cgc aag atg ctg cgt ttg gag gtt 48

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

1 5 10 15

cgt aac agc caa acg cct atc gaa cgc aag ccc gag tgg atc aag act 96

Arg Asn Ser Gln Thr Pro Ile Glu Arg Lys Pro Glu Trp Ile Lys Thr

20 25 30

cgt gct aag atg ggt ccc gaa tac acg aag atg cag aac ttg gtc aag 144

Arg Ala Lys Met Gly Pro Glu Tyr Thr Lys Met Gln Asn Leu Val Lys

35 40 45

tcg gag ggc tta cat acg gta tgc cag gaa gct ggg tgc ccc aac att 192

Ser Glu Gly Leu His Thr Val Cys Gln Glu Ala Gly Cys Pro Asn Ile

50 55 60

tat gag tgc tgg gag gat cgt gaa gca acg ttt ttg atc gga ggc gat 240

Tyr Glu Cys Trp Glu Asp Arg Glu Ala Thr Phe Leu Ile Gly Gly Asp

65 70 75 80

cag tgc act cgt cgc tgc gac ttt tgt caa atc gat aca gga aaa cct 288

Gln Cys Thr Arg Arg Cys Asp Phe Cys Gln Ile Asp Thr Gly Lys Pro

85 90 95

gaa gcc ctt gat cgt gat gag cca cgt cgt gta gga gaa tcg gtc gtc 336

Glu Ala Leu Asp Arg Asp Glu Pro Arg Arg Val Gly Glu Ser Val Val

100 105 110

aca atg gat ctt aac tat gct acc att act ggc gtt gca cgt gat gat 384

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

115 120 125

ctg ccc gat ggt gga gct tgg ctt tac gct gaa act gtt cgt cag atc 432

Leu Pro Asp Gly Gly Ala Trp Leu Tyr Ala Glu Thr Val Arg Gln Ile

130 135 140

cat gaa cag act gcg ggt cgc gaa gcc ggg cgt acc aaa gtt gaa ctt 480

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

145 150 155 160

ctt gca cct gat ttt aac gcg gtg cca gag ctg tta cgc gaa gtc ttt 528
 Leu Ala Pro Asp Phe Asn Ala Val Pro Glu Leu Leu Arg Glu Val Phe
 165 170 175
 gaa tcg cgc cct gaa gtc ttc gct cac aat gta gag aca gta cca cgc 576

 Glu Ser Arg Pro Glu Val Phe Ala His Asn Val Glu Thr Val Pro Arg
 180 185 190
 atc ttt aag cgt att cgt cca ggg ttt cgc tat gag cgc agt ctt aag 624
 Ile Phe Lys Arg Ile Arg Pro Gly Phe Arg Tyr Glu Arg Ser Leu Lys
 195 200 205
 gtt atc act gat gcg cgt gat ttt gga ttg gtc acc aaa tcg aac ctt 672
 Val Ile Thr Asp Ala Arg Asp Phe Gly Leu Val Thr Lys Ser Asn Leu

 210 215 220
 atc ttg gga atg ggg gaa aca cgc gag gaa att tcg gaa gcc tta aaa 720
 Ile Leu Gly Met Gly Glu Thr Arg Glu Glu Ile Ser Glu Ala Leu Lys
 225 230 235 240
 cag ctg cat gaa gcg ggt tgc gag ctt att aca atc acc cag tat ctt 768
 Gln Leu His Glu Ala Gly Cys Glu Leu Ile Thr Ile Thr Gln Tyr Leu
 245 250 255

 cgt cct agt gtc cgt cat cat ccg gtc gaa cgc tgg gta aaa ccc cag 816
 Arg Pro Ser Val Arg His His Pro Val Glu Arg Trp Val Lys Pro Gln
 260 265 270
 gaa ttt gtg gaa ctt aag gag gag gcg gag caa att ggc ttt tcg ggt 864
 Glu Phe Val Glu Leu Lys Glu Glu Ala Glu Gln Ile Gly Phe Ser Gly
 275 280 285
 gta atg tca ggc ccc ctt gtg cgt tcc tcg tac cgt gcg ggg cgc ttg 912

 Val Met Ser Gly Pro Leu Val Arg Ser Ser Tyr Arg Ala Gly Arg Leu
 290 295 300
 tac gga atg gct atg gaa cag cgc cgt tcc gca acc gtt tga 954
 Tyr Gly Met Ala Met Glu Gln Arg Arg Ser Ala Thr Val
 305 310 315
 <210> 113

<211> 317

<212> PRT

<213> Streptomyces coelicolor A3(2)

<400> 113

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

1 5 10 15

Arg Asn Ser Gln Thr Pro Ile Glu Arg Lys Pro Glu Trp Ile Lys Thr

20 25 30

Arg Ala Lys Met Gly Pro Glu Tyr Thr Lys Met Gln Asn Leu Val Lys

35 40 45

Ser Glu Gly Leu His Thr Val Cys Gln Glu Ala Gly Cys Pro Asn Ile

50 55 60

Tyr Glu Cys Trp Glu Asp Arg Glu Ala Thr Phe Leu Ile Gly Gly Asp

65 70 75 80

Gln Cys Thr Arg Arg Cys Asp Phe Cys Gln Ile Asp Thr Gly Lys Pro

85 90 95

Glu Ala Leu Asp Arg Asp Glu Pro Arg Arg Val Gly Glu Ser Val Val

100 105 110

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

115 120 125

Leu Pro Asp Gly Gly Ala Trp Leu Tyr Ala Glu Thr Val Arg Gln Ile

130 135 140

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

145 150 155 160

Leu Ala Pro Asp Phe Asn Ala Val Pro Glu Leu Leu Arg Glu Val Phe

165 170 175

Glu Ser Arg Pro Glu Val Phe Ala His Asn Val Glu Thr Val Pro Arg

180 185 190

Ile Phe Lys Arg Ile Arg Pro Gly Phe Arg Tyr Glu Arg Ser Leu Lys

195 200 205

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

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

260 265 270
 Glu Phe Val Glu Leu Lys Glu Glu Ala Glu Gln Ile Gly Phe Ser Gly
 275 280 285
 Val Met Ser Gly Pro Leu Val Arg Ser Ser Tyr Arg Ala Gly Arg Leu
 290 295 300
 Tyr Gly Met Ala Met Glu Gln Arg Arg Ser Ala Thr Val

305 310 315

<210> 114

<211> 642

<212> DNA

<213> Escherichia coli

<220><

221> CDS

<222> (1)..(642)

<223> lipB encoding octanoyltransferase (LipB)

<400> 114

atg tat cag gat aaa att ctt gtc cgc cag ctc ggt ctt cag cct tac 48

Met Tyr Gln Asp Lys Ile Leu Val Arg Gln Leu Gly Leu Gln Pro Tyr

1 5 10 15

gag cca atc tcc cag gct atg cat gaa ttc acc gat acc cgc gat gat 96

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

20 25 30

agt acc ctt gat gaa atc tgg ctg gtc gag cac tat ccg gta ttc acc 144

Ser Thr Leu Asp Glu Ile Trp Leu Val Glu His Tyr Pro Val Phe Thr

35 40 45

caa ggt cag gca gga aaa gcg gag cac att tta atg ccg ggt gat att 192

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

50	55	60	
ccg gtg atc cag agc gat cgc ggt ggg cag gtg act tat cac ggg ccg			240
Pro Val Ile Gln Ser Asp Arg Gly Gly Gln Val Thr Tyr His Gly Pro			
65	70	75	80
ggg caa cag gtg atg tat gtg ttg ctt aac ctg aaa cgc cgt aaa ctc			288
Gly Gln Gln Val Met Tyr Val Leu Leu Asn Leu Lys Arg Arg Lys Leu			
85	90	95	
ggt gtg cgt gaa ctg gtg acc ttg ctt gag caa aca gtg gtg aat acc			336
Gly Val Arg Glu Leu Val Thr Leu Leu Glu Gln Thr Val Val Asn Thr			
100	105	110	
ctg gct gaa ctg ggt ata gaa gcg cat cct cgg gct gac gcg cca ggt			384
Leu Ala Glu Leu Gly Ile Glu Ala His Pro Arg Ala Asp Ala Pro Gly			
115	120	125	
gtc tat gtt ggg gaa aag aaa att tgc tca ctg ggt tta cgt att cga			432
Val Tyr Val Gly Glu Lys Lys Ile Cys Ser Leu Gly Leu Arg Ile Arg			
130	135	140	
cgc ggt tgt tca ttc cac ggt ctg gca tta aac gtc aat atg gat ctt			480
Arg Gly Cys Ser Phe His Gly Leu Ala Leu Asn Val Asn Met Asp Leu			
145	150	155	160
tca cca ttt tta cgt att aat cct tgt ggg tat gcc gga atg gaa atg			528
Ser Pro Phe Leu Arg Ile Asn Pro Cys Gly Tyr Ala Gly Met Glu Met			
165	170	175	
gct aaa ata tca caa tgg aaa ccc gaa gcg acg act aat aat att gct			576
Ala Lys Ile Ser Gln Trp Lys Pro Glu Ala Thr Thr Asn Asn Ile Ala			
180	185	190	
cca cgt tta ctg gaa aat att tta gcg cta cta aac aat ccg gac ttc			624
Pro Arg Leu Leu Glu Asn Ile Leu Ala Leu Leu Asn Asn Pro Asp Phe			
195	200	205	
gaa tat att acc gct taa			642

Glu Tyr Ile Thr Ala

210

<210> 115

<211> 213

<212> PRT

<213> Escherichia coli

<400> 115

Met Tyr Gln Asp Lys Ile Leu Val Arg Gln Leu Gly Leu Gln Pro Tyr

1 5 10 15

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

20 25 30

Ser Thr Leu Asp Glu Ile Trp Leu Val Glu His Tyr Pro Val Phe Thr

35 40 45

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

50 55 60

Pro Val Ile Gln Ser Asp Arg Gly Gly Gln Val Thr Tyr His Gly Pro

65 70 75 80

Gly Gln Gln Val Met Tyr Val Leu Leu Asn Leu Lys Arg Arg Lys Leu

85 90 95

Gly Val Arg Glu Leu Val Thr Leu Leu Glu Gln Thr Val Val Asn Thr

100 105 110

Leu Ala Glu Leu Gly Ile Glu Ala His Pro Arg Ala Asp Ala Pro Gly

115 120 125

Val Tyr Val Gly Glu Lys Lys Ile Cys Ser Leu Gly Leu Arg Ile Arg

130 135 140

Arg Gly Cys Ser Phe His Gly Leu Ala Leu Asn Val Asn Met Asp Leu

145 150 155 160

Ser Pro Phe Leu Arg Ile Asn Pro Cys Gly Tyr Ala Gly Met Glu Met

165 170 175

Ala Lys Ile Ser Gln Trp Lys Pro Glu Ala Thr Thr Asn Asn Ile Ala

180 185 190

Pro Arg Leu Leu Glu Asn Ile Leu Ala Leu Leu Asn Asn Pro Asp Phe

195 200 205

Glu Tyr Ile Thr Ala

210

<210> 116

<211> 576

<212> DNA

<213> *Shigella flexneri*

<220><221> CDS

<222> (1)..(576)

<223> lipB gene encoding a octanoyltransferase (LipB)

<400> 116

atg cat gaa ttc acc gat acc cgc gat aat agt acc ctt gat gaa atc 48

Met His Glu Phe Thr Asp Thr Arg Asp Asn Ser Thr Leu Asp Glu Ile

1 5 10 15

tgg ctg gtc gag cac tat ccg gta ttc acc caa ggt cag gca gga aaa 96

Trp Leu Val Glu His Tyr Pro Val Phe Thr Gln Gly Gln Ala Gly Lys

20 25 30

gcg gag cac att tta atg ccg ggt gat att ccg gtg atc cag agc gat 144

Ala Glu His Ile Leu Met Pro Gly Asp Ile Pro Val Ile Gln Ser Asp

35 40 45

cgc ggt ggg cag gtg act tat cac ggg ccg gga caa cag gtg atg tat 192

Arg Gly Gly Gln Val Thr Tyr His Gly Pro Gly Gln Gln Val Met Tyr

50 55 60

gtg ttg ctt aac ctg aaa cgc cgt aaa ctc ggt gtg cgt gaa ctg gtg 240

Val Leu Leu Asn Leu Lys Arg Arg Lys Leu Gly Val Arg Glu Leu Val

65 70 75 80

acc ttg ctt gag caa aca gtg gtg aat acc ctg gct gaa ctg ggt ata 288

Thr Leu Leu Glu Gln Thr Val Val Asn Thr Leu Ala Glu Leu Gly Ile

85 90 95

gaa gcg cat cct cgg gct gac gcg cct ggt gtc tat gtc ggg gaa aag 336

Glu Ala His Pro Arg Ala Asp Ala Pro Gly Val Tyr Val Gly Glu Lys

100 105 110
aaa att tgc tca ctg ggt tta cga att cga cgc ggt tgt tca ttc cac 384

Lys Ile Cys Ser Leu Gly Leu Arg Ile Arg Arg Gly Cys Ser Phe His
115 120 125
ggt ctg gca tta aac gtc aat atg gat ctt tca cca ttt tta cgt att 432

Gly Leu Ala Leu Asn Val Asn Met Asp Leu Ser Pro Phe Leu Arg Ile
130 135 140
aat cct tgt ggg tat gcc gga atg gaa atg gct aaa ata tca caa tgg 480
Asn Pro Cys Gly Tyr Ala Gly Met Glu Met Ala Lys Ile Ser Gln Trp

145 150 155 160
aaa ccc gaa gcg acg act aat aat att gct cca cgt tta ctg gaa aat 528
Lys Pro Glu Ala Thr Thr Asn Asn Ile Ala Pro Arg Leu Leu Glu Asn

165 170 175
att tta gcg cta cta aac aat ccg gac ttc gaa tat att acc gct taa 576
Ile Leu Ala Leu Leu Asn Asn Pro Asp Phe Glu Tyr Ile Thr Ala

180 185 190

<210> 117
<211> 191
<212> PRT
<213> Shigella flexneri
<400> 117

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

50 55 60
Val Leu Leu Asn Leu Lys Arg Arg Lys Leu Gly Val Arg Glu Leu Val
65 70 75 80

Thr Leu Leu Glu Gln Thr Val Val Asn Thr Leu Ala Glu Leu Gly Ile
85 90 95
Glu Ala His Pro Arg Ala Asp Ala Pro Gly Val Tyr Val Gly Glu Lys
100 105 110
Lys Ile Cys Ser Leu Gly Leu Arg Ile Arg Arg Gly Cys Ser Phe His
115 120 125
Gly Leu Ala Leu Asn Val Asn Met Asp Leu Ser Pro Phe Leu Arg Ile
130 135 140
Asn Pro Cys Gly Tyr Ala Gly Met Glu Met Ala Lys Ile Ser Gln Trp
145 150 155 160
Lys Pro Glu Ala Thr Thr Asn Asn Ile Ala Pro Arg Leu Leu Glu Asn
165 170 175
Ile Leu Ala Leu Leu Asn Asn Pro Asp Phe Glu Tyr Ile Thr Ala

180 185 190
<210> 118
<211> 1893
<212> DNA
<213> Escherichia coli
<220><221> CDS
<222> (1)..(1893)
<223> aceF gene encoding dihydrolipoyllysine-residue acetyltransferase
component of pyruvate dehydrogenase (E2)

<400> 118
atg gct atc gaa atc aaa gta ccg gac atc ggg gct gat gaa gtt gaa 48
Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu
1 5 10 15
atc acc gag atc ctg gtc aaa gtg ggc gac aaa gtt gaa gcc gaa cag 96
Ile Thr Glu Ile Leu Val Lys Val Gly Asp Lys Val Glu Ala Glu Gln
20 25 30
tcg ctg atc acc gta gaa ggc gac aaa gcc tct atg gaa gtt ccg tct 144
Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ser
35 40 45

ccg cag gcg ggt atc gtt aaa gag atc aaa gtc tct gtt ggc gat aaa 192

Pro Gln Ala Gly Ile Val Lys Glu Ile Lys Val Ser Val Gly Asp Lys
50 55 60

acc cag acc ggc gca ctg att atg att ttc gat tcc gcc gac ggt gca 240
Thr Gln Thr Gly Ala Leu Ile Met Ile Phe Asp Ser Ala Asp Gly Ala
65 70 75 80

gca gac gct gca cct gct cag gca gaa gag aag aaa gaa gca gct ccg 288
Ala Asp Ala Ala Pro Ala Gln Ala Glu Glu Lys Lys Glu Ala Ala Pro

85 90 95

gca gca gca cca gcg gct gcg gcg gca aaa gac gtt aac gtt ccg gat 336
Ala Ala Ala Pro Ala Ala Ala Ala Lys Asp Val Asn Val Pro Asp

100 105 110

atc ggc agc gac gaa gtt gaa gtg acc gaa atc ctg gtg aaa gtt ggc 384
Ile Gly Ser Asp Glu Val Glu Val Thr Glu Ile Leu Val Lys Val Gly

115 120 125

gat aaa gtt gaa gct gaa cag tcg ctg atc acc gta gaa ggc gac aag 432
Asp Lys Val Glu Ala Glu Gln Ser Leu Ile Thr Val Glu Gly Asp Lys

130 135 140

gct tct atg gaa gtt ccg gct ccg ttt gct ggc acc gtg aaa gag atc 480
Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly Thr Val Lys Glu Ile

145 150 155 160

aaa gtg aac gtg ggt gac aaa gtg tct acc ggc tcg ctg att atg gtc 528

Lys Val Asn Val Gly Asp Lys Val Ser Thr Gly Ser Leu Ile Met Val
165 170 175

ttc gaa gtc gcg ggt gaa gca ggc gcg gca gct ccg gcc gct aaa cag 576
Phe Glu Val Ala Gly Glu Ala Gly Ala Ala Ala Pro Ala Ala Lys Gln

180 185 190

gaa gca gct ccg gca gcg gcc cct gca cca gcg gct ggc gtg aaa gaa 624
Glu Ala Ala Pro Ala Ala Ala Pro Ala Pro Ala Ala Gly Val Lys Glu

195	200	205	
gtt aac gtt ccg gat atc ggc ggt gac gaa gtt gaa gtg act gaa gtg			672
Val Asn Val Pro Asp Ile Gly Gly Asp Glu Val Glu Val Thr Glu Val			
210	215	220	
atg gtg aaa gtg ggc gac aaa gtt gcc gct gaa cag tca ctg atc acc			720
Met Val Lys Val Gly Asp Lys Val Ala Ala Glu Gln Ser Leu Ile Thr			
225	230	235	240
gta gaa ggc gac aaa gct tct atg gaa gtt ccg gcg ccg ttt gca ggc			768
Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly			
245	250	255	
gtc gtg aag gaa ctg aaa gtc aac gtt ggc gat aaa gtg aaa act ggc			816
Val Val Lys Glu Leu Lys Val Asn Val Gly Asp Lys Val Lys Thr Gly			
260	265	270	
tcg ctg att atg atc ttc gaa gtt gaa ggc gca gcg cct gcg gca gct			864
Ser Leu Ile Met Ile Phe Glu Val Glu Gly Ala Ala Pro Ala Ala Ala			
275	280	285	
cct gcg aaa cag gaa gcg gca gcg ccg gca ccg gca gca aaa gct gaa			912
Pro Ala Lys Gln Glu Ala Ala Ala Pro Ala Pro Ala Ala Lys Ala Glu			
290	295	300	
gcc ccg gca gca gca cca gct gcg aaa gcg gaa ggc aaa tct gaa ttt			960
Ala Pro Ala Ala Ala Pro Ala Ala Lys Ala Glu Gly Lys Ser Glu Phe			
305	310	315	320
gct gaa aac gac gct tat gtt cac gcg act ccg ctg atc cgc cgt ctg			1008
Ala Glu Asn Asp Ala Tyr Val His Ala Thr Pro Leu Ile Arg Arg Leu			
325	330	335	
gca cgc gag ttt ggt gtt aac ctt gcg aaa gtg aag ggc act ggc cgt			1056
Ala Arg Glu Phe Gly Val Asn Leu Ala Lys Val Lys Gly Thr Gly Arg			
340	345	350	
aaa ggt cgt atc ctg cgc gaa gac gtt cag gct tac gtg aaa gaa gct			1104
Lys Gly Arg Ile Leu Arg Glu Asp Val Gln Ala Tyr Val Lys Glu Ala			
355	360	365	

atc aaa cgt gca gaa gca gct ccg gca gcg act ggc ggt ggt atc cct	1152
Ile Lys Arg Ala Glu Ala Ala Pro Ala Ala Thr Gly Gly Gly Ile Pro	
370 375 380	
ggc atg ctg ccg tgg ccg aag gtg gac ttc agc aag ttt ggt gaa atc	1200
Gly Met Leu Pro Trp Pro Lys Val Asp Phe Ser Lys Phe Gly Glu Ile	
385 390 395 400	
gaa gaa gtg gaa ctg ggc cgc atc cag aaa atc tct ggt gcg aac ctg	1248
Glu Glu Val Glu Leu Gly Arg Ile Gln Lys Ile Ser Gly Ala Asn Leu	
405 410 415	
agc cgt aac tgg gta atg atc ccg cat gtt act cac ttc gac aaa acc	1296
Ser Arg Asn Trp Val Met Ile Pro His Val Thr His Phe Asp Lys Thr	
420 425 430	
gat atc acc gag ttg gaa gcg ttc cgt aaa cag cag aac gaa gaa gcg	1344
Asp Ile Thr Glu Leu Glu Ala Phe Arg Lys Gln Gln Asn Glu Glu Ala	
435 440 445	
gcg aaa cgt aag ctg gat gtg aag atc acc ccg gtt gtc ttc atc atg	1392
Ala Lys Arg Lys Leu Asp Val Lys Ile Thr Pro Val Val Phe Ile Met	
450 455 460	
aaa gcc gtt gct gca gct ctt gag cag atg cct cgc ttc aat agt tcg	1440
Lys Ala Val Ala Ala Ala Leu Glu Gln Met Pro Arg Phe Asn Ser Ser	
465 470 475 480	
ctg tcg gaa gac ggt cag cgt ctg acc ctg aag aaa tac atc aac atc	1488
Leu Ser Glu Asp Gly Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn Ile	
485 490 495	
ggt gtg gcg gtg gat acc ccg aac ggt ctg gtt gtt ccg gta ttc aaa	1536
Gly Val Ala Val Asp Thr Pro Asn Gly Leu Val Val Pro Val Phe Lys	
500 505 510	
gac gtc aac aag aaa ggc atc atc gag ctg tct cgc gag ctg atg act	1584
Asp Val Asn Lys Lys Gly Ile Ile Glu Leu Ser Arg Glu Leu Met Thr	
515 520 525	
att tct aag aaa gcg cgt gac ggt aag ctg act gcg ggc gaa atg cag	1632

Ile Ser Lys Lys Ala Arg Asp Gly Lys Leu Thr Ala Gly Glu Met Gln

530 535 540
ggc ggt tgc ttc acc atc tcc agc atc ggc ggc ctg ggt act acc cac 1680

Gly Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His

545 550 555 560
ttc gcg ccg att gtg aac gcg ccg gaa gtg gct atc ctc ggc gtt tcc 1728

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

565 570 575

aag tcc gcg atg gag ccg gtg tgg aat ggt aaa gag ttc gtg ccg cgt 1776

Lys Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg

580 585 590

ctg atg ctg ccg att tct ctc tcc ttc gac cac cgc gtg atc gac ggt 1824

Leu Met Leu Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly

595 600 605

gct gat ggt gcc cgt ttc att acc atc att aac aac acg ctg tct gac 1872

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

610 615 620

att cgc cgt ctg gtg atg taa 1893

Ile Arg Arg Leu Val Met

625 630

<210> 119

<211> 630

<212> PRT

<213> Escherichia coli

<400> 119

Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu

1 5 10 15

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

20 25 30

Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ser

35 40 45

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

290 295 300
 Ala Pro Ala Ala Ala Pro Ala Ala Lys Ala Glu Gly Lys Ser Glu Phe
 305 310 315 320
 Ala Glu Asn Asp Ala Tyr Val His Ala Thr Pro Leu Ile Arg Arg Leu

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

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

 450 455 460
 Lys Ala Val Ala Ala Ala Leu Glu Gln Met Pro Arg Phe Asn Ser Ser
 465 470 475 480
 Leu Ser Glu Asp Gly Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn Ile
 485 490 495
 Gly Val Ala Val Asp Thr Pro Asn Gly Leu Val Val Pro Val Phe Lys
 500 505 510
 Asp Val Asn Lys Lys Gly Ile Ile Glu Leu Ser Arg Glu Leu Met Thr

 515 520 525
 Ile Ser Lys Lys Ala Arg Asp Gly Lys Leu Thr Ala Gly Glu Met Gln
 530 535 540

Gly Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His
 545 550 555 560
 Phe Ala Pro Ile Val Asn Ala Pro Glu Val Ala Ile Leu Gly Val Ser
 565 570 575
 Lys Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg

580 585 590
 Leu Met Leu Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly
 595 600 605
 Ala Asp Gly Ala Arg Phe Ile Thr Ile Ile Asn Asn Thr Leu Ser Asp
 610 615 620
 Ile Arg Arg Leu Val Met
 625 630

<210> 120

<211> 1890

<212> DNA

<213> Klebsiella oxytoca

<220><221> CDS

<222> (1)..(1890)

<223> aceF gene encoding dihydrolipoyllysine-residue acetyltransferase

component of pyruvate dehydrogenase (E2)

<400> 120

atg gct atc gag atc aag gtg ccc gac atc ggc gct gac gaa gta gaa 48

Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu

1 5 10 15

att acc gag atc ctg gtt aaa gtc ggt gat aag gta gag gca gaa cag 96

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

20 25 30

agt tta atc act gta gaa ggc gat aag gca tcc atg gag gta cca tca 144

Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ser

35 40 45

ccg caa gct ggt gta gtt aaa gag atc aaa gtg agt gta ggc gac aaa 192

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

50	55	60	
act gaa acc ggt aag tta att atg atc ttc gat tca gcc gac ggg gca			240
Thr Glu Thr Gly Lys Leu Ile Met Ile Phe Asp Ser Ala Asp Gly Ala			
65	70	75	80
gcc gcc gct gct ccc gca cag gaa gag aaa aag gag gcg gca cct gcc			288
Ala Ala Ala Ala Pro Ala Gln Glu Glu Lys Lys Glu Ala Ala Pro Ala			
	85	90	95
gct gca gca cca gca gcc gct tct gca aaa gag gtg cat gta cct gac			336
Ala Ala Ala Pro Ala Ala Ala Ser Ala Lys Glu Val His Val Pro Asp			
100	105	110	
att gga ggc gac gaa gta gaa gta aca gag att atg gtc aag gtt ggc			384
Ile Gly Gly Asp Glu Val Glu Val Thr Glu Ile Met Val Lys Val Gly			
115	120	125	
gat aca atc gca gcg gaa caa agc tta att acg gta gaa ggc gat aaa			432
Asp Thr Ile Ala Ala Glu Gln Ser Leu Ile Thr Val Glu Gly Asp Lys			
130	135	140	
gca agc atg gaa gtt ccc gct ccc ttc gct ggg act gta aaa gaa atc			480
Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly Thr Val Lys Glu Ile			
145	150	155	160
aag att aac acc ggc gac aag gtt tcc acc ggc tca tta att atg atc			528
Lys Ile Asn Thr Gly Asp Lys Val Ser Thr Gly Ser Leu Ile Met Ile			
	165	170	175
ttc gaa gta gca gga gct gca cct gcg gca gcg cct gcg aaa gcg gag			576
Phe Glu Val Ala Gly Ala Ala Pro Ala Ala Ala Pro Ala Lys Ala Glu			
180	185	190	
gct gca cct gca gcg gcg gct ccc gcc gct agt ggt agt aaa gaa gtg			624
Ala Ala Pro Ala Ala Ala Ala Pro Ala Ala Ser Gly Ser Lys Glu Val			
195	200	205	
cac gtt ccg gac atc gga ggt gac gag gtc gaa gtc act gaa gtg atg			672

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

210	215	220	
gta aaa gca ggg gat aaa atc gca gcc gag cag agt tta att aca gtc			720
Val Lys Ala Gly Asp Lys Ile Ala Ala Glu Gln Ser Leu Ile Thr Val			
225	230	235	240
gaa ggc gat aag gcg tct atg gaa gtt cca gcg cca ttc gcc ggt aca			768
Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly Thr			
245	250	255	
gta aag gaa att aag atc agc act gga gat aaa gtc tca act ggt tca			816
Val Lys Glu Ile Lys Ile Ser Thr Gly Asp Lys Val Ser Thr Gly Ser			
260	265	270	
ttg atc atg gtc ttc gaa gtc gaa ggc gcc gca cct gcg gcg gca ccg			864
Leu Ile Met Val Phe Glu Val Glu Gly Ala Ala Pro Ala Ala Ala Pro			
275	280	285	
gca gcg gct gcc gct cca gca cca gct gct gcg ccc gca caa gct gca			912
Ala Ala Ala Ala Ala Pro Ala Pro Ala Ala Ala Pro Ala Gln Ala Ala			
290	295	300	
aaa cca gct gcc ccc gca gcg aag gcc gaa ggc aag agc gag ttc gca			960
Lys Pro Ala Ala Pro Ala Ala Lys Ala Glu Gly Lys Ser Glu Phe Ala			
305	310	315	320
gag aat gat gcg tac gta cat gcg aca cca ttg att cgc cgt ttg gca			1008
Glu Asn Asp Ala Tyr Val His Ala Thr Pro Leu Ile Arg Arg Leu Ala			
325	330	335	
cgc gaa ttc ggt gtg aac ctg gct aaa gtg aag gga acg ggg cgc aaa			1056
Arg Glu Phe Gly Val Asn Leu Ala Lys Val Lys Gly Thr Gly Arg Lys			
340	345	350	
ggc cgc att ttg cgt gag gac gtc cag gct tat gtt aaa gaa gca gtc			1104
Gly Arg Ile Leu Arg Glu Asp Val Gln Ala Tyr Val Lys Glu Ala Val			
355	360	365	

aag cgt gcc gaa gca gcg cca gcc gca acc ggc ggg ggg atc cca ggc	1152
Lys Arg Ala Glu Ala Ala Pro Ala Ala Thr Gly Gly Gly Ile Pro Gly	
370 375 380	
atg tta ccc tgg cca aag gta gac ttt tca aaa ttt ggg gag gta gaa	1200
Met Leu Pro Trp Pro Lys Val Asp Phe Ser Lys Phe Gly Glu Val Glu	
385 390 395 400	
gag gtt gag ttg gga cgc atc cag aag atc tcc ggt gca aat ttg tgc	1248
Glu Val Glu Leu Gly Arg Ile Gln Lys Ile Ser Gly Ala Asn Leu Ser	
405 410 415	
cgc aac tgg gtc atg att ccg cac gtc act cac ttt gac aaa acg gac	1296
Arg Asn Trp Val Met Ile Pro His Val Thr His Phe Asp Lys Thr Asp	
420 425 430	
att acc gat ttg gag gct ttt cgt aag caa caa aat gct gag gcg gag	1344
Ile Thr Asp Leu Glu Ala Phe Arg Lys Gln Gln Asn Ala Glu Ala Glu	
435 440 445	
aag cgt aaa ttg gac gtg aag ttc acc ccg gtg gtg ttc att atg aag	1392
Lys Arg Lys Leu Asp Val Lys Phe Thr Pro Val Val Phe Ile Met Lys	
450 455 460	
gca gtg gcc gca gca ctt gaa cag atg ccg cgt ttc aac tca tcc ctg	1440
Ala Val Ala Ala Ala Leu Glu Gln Met Pro Arg Phe Asn Ser Ser Leu	
465 470 475 480	
tca gag gat gct caa cgt ctt acc tta aag aag tac atc aac att ggt	1488
Ser Glu Asp Ala Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn Ile Gly	
485 490 495	
gtt gct gtg gac acg cca aat ggg ttg gta gtc ccc gtg ttt aaa gac	1536
Val Ala Val Asp Thr Pro Asn Gly Leu Val Val Pro Val Phe Lys Asp	
500 505 510	
gtt aat aag aag tcc att aca gag tta tgc cgc gaa tta act gtt atc	1584
Val Asn Lys Lys Ser Ile Thr Glu Leu Ser Arg Glu Leu Thr Val Ile	
515 520 525	
agc aag aaa gca cgc gat ggg aag ctg act gcc ggc gaa atg caa ggc	1632

Ser Lys Lys Ala Arg Asp Gly Lys Leu Thr Ala Gly Glu Met Gln Gly
530 535 540
gga tgt ttt acc atc tcg agt att ggg gga tta ggg acc aca cat ttt 1680
Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His Phe
545 550 555 560
gca ccc atc gtc aat gca cct gaa gta gct atc tta ggg gta tca aaa 1728
Ala Pro Ile Val Asn Ala Pro Glu Val Ala Ile Leu Gly Val Ser Lys
565 570 575
tca gcg atg gag ccg gtt tgg aat ggg aaa gag ttc gta ccc cgt ctg 1776
Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg Leu
580 585 590
atg atg cca att tca ctg tca ttc gac cat cgc gtc att gat ggc gcg 1824
Met Met Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly Ala
595 600 605
gat ggc gca cgc ttt atc aca att atc aat aat atg ctt tca gat att 1872
Asp Gly Ala Arg Phe Ile Thr Ile Ile Asn Asn Met Leu Ser Asp Ile
610 615 620
cgc cgt tta gtc atg taa 1890

Arg Arg Leu Val Met

625

<210> 121

<211> 629

<212> PRT

<213> Klebsiella oxytoca

<400> 121

Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu

1 5 10 15

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

20 25 30

Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ser

35

40

45

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

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

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

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

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

Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His Phe

545 550 555 560

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

565 570 575

Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg Leu

580 585 590

Met Met Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly Ala

595 600 605

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

610 615 620

Arg Arg Leu Val Met

625

<210> 122

<211> 1017

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(1017)

<223> lplA gene encoding lipoate-protein ligase A (LplA)

<400> 122

atg tcc aca tta cgc ctg ctc atc tct gac tct tac gac ccg tgg ttt 48

Met Ser Thr Leu Arg Leu Leu Ile Ser Asp Ser Tyr Asp Pro Trp Phe

1 5 10 15

aac ctg gcg gtg gaa gag tgt att ttt cgc caa atg ccc gcc acg cag 96

Asn Leu Ala Val Glu Glu Cys Ile Phe Arg Gln Met Pro Ala Thr Gln

20 25 30

cgc gtt ctg ttt ctc tgg cgc aat gcc gac acg gta gta att ggt cgc 144

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

35 40 45

gcg cag aac ccg tgg aaa gag tgt aat acc cgg cgg atg gaa gaa gat 192

Ala Gln Asn Pro Trp Lys Glu Cys Asn Thr Arg Arg Met Glu Glu Asp

50	55	60	
aac gtc cgc ctg gcg cgg cgc agt agc ggt ggc ggc gcg gtg ttc cac			240
Asn Val Arg Leu Ala Arg Arg Ser Ser Gly Gly Gly Ala Val Phe His			
65	70	75	80
gat ctc ggc aat acc tgc ttt acc ttt atg gct ggc aag ccg gag tac			288
Asp Leu Gly Asn Thr Cys Phe Thr Phe Met Ala Gly Lys Pro Glu Tyr			
	85	90	95
gat aaa act atc tcc acg tcg att gtg ctc aat gcg ctg aac gcg ctc			336
Asp Lys Thr Ile Ser Thr Ser Ile Val Leu Asn Ala Leu Asn Ala Leu			
100	105	110	
ggc gtc agc gcc gaa gcg tcc gga cgt aac gat ctg gtg gtg aaa acc			384
Gly Val Ser Ala Glu Ala Ser Gly Arg Asn Asp Leu Val Val Lys Thr			
115	120	125	
gtc gaa ggc gac cgc aaa gtc tca ggc tcg gcc tat cgc gaa acc aaa			432
Val Glu Gly Asp Arg Lys Val Ser Gly Ser Ala Tyr Arg Glu Thr Lys			
130	135	140	
gat cgc ggc ttc cac cac ggc acc ttg cta ctc aat gcc gac ctc agc			480
Asp Arg Gly Phe His His Gly Thr Leu Leu Leu Asn Ala Asp Leu Ser			
145	150	155	160
cgc ctg gca aac tat ctc aat ccg gat aaa aag aaa ctg gcg gcg aaa			528
Arg Leu Ala Asn Tyr Leu Asn Pro Asp Lys Lys Lys Leu Ala Ala Lys			
165	170	175	
ggc att acg tcg gta cgt tcc cgc gtg acc aac ctc acc gag ctg ttg			576
Gly Ile Thr Ser Val Arg Ser Arg Val Thr Asn Leu Thr Glu Leu Leu			
180	185	190	
ccg ggg atc acc cat gag cag gtt tgc gag gcc ata acc gag gcc ttt			624
Pro Gly Ile Thr His Glu Gln Val Cys Glu Ala Ile Thr Glu Ala Phe			
195	200	205	
ttc gcc cat tat ggc gag cgc gtg gaa gcg gaa atc atc tcc ccg aac			672
Phe Ala His Tyr Gly Glu Arg Val Glu Ala Glu Ile Ile Ser Pro Asn			
210	215	220	

aaa acg cca gac ttg cca aac ttc gcc gaa acc ttt gcc cgc cag agt 720
Lys Thr Pro Asp Leu Pro Asn Phe Ala Glu Thr Phe Ala Arg Gln Ser
225 230 235 240

agc tgg gaa tgg aac ttc ggt cag gct ccg gca ttc tcg cat ctg ctg 768
Ser Trp Glu Trp Asn Phe Gly Gln Ala Pro Ala Phe Ser His Leu Leu
245 250 255

gat gaa cgc ttt acc tgg ggc ggc gtg gaa ctg cat ttc gac gtt gaa 816
Asp Glu Arg Phe Thr Trp Gly Gly Val Glu Leu His Phe Asp Val Glu
260 265 270

aaa ggc cat atc acc cgc gcc cag gtg ttt acc gac agc ctc aac ccc 864

Lys Gly His Ile Thr Arg Ala Gln Val Phe Thr Asp Ser Leu Asn Pro
275 280 285

gcg ccg ctg gaa gcc ctc gcc gga cga ctg caa ggc tgc ctg tac cgc 912
Ala Pro Leu Glu Ala Leu Ala Gly Arg Leu Gln Gly Cys Leu Tyr Arg
290 295 300

gca gat atg ctg caa cag gag tgc gaa gcg ctg ttg gtt gac ttc ccg 960
Ala Asp Met Leu Gln Gln Glu Cys Glu Ala Leu Leu Val Asp Phe Pro

305 310 315 320
gaa cag gaa aaa gag cta cgg gag tta tcg gca tgg atg gcg ggg gct 1008
Glu Gln Glu Lys Glu Leu Arg Glu Leu Ser Ala Trp Met Ala Gly Ala
325 330 335

gta agg tag 1017
Val Arg

<210> 123

<211> 338

<212> PRT

<213> Escherichia coli

<400> 123

Met Ser Thr Leu Arg Leu Leu Ile Ser Asp Ser Tyr Asp Pro Trp Phe
1 5 10 15

Asn Leu Ala Val Glu Glu Cys Ile Phe Arg Gln Met Pro Ala Thr Gln
 20 25 30
 Arg Val Leu Phe Leu Trp Arg Asn Ala Asp Thr Val Val Ile Gly Arg
 35 40 45
 Ala Gln Asn Pro Trp Lys Glu Cys Asn Thr Arg Arg Met Glu Glu Asp
 50 55 60
 Asn Val Arg Leu Ala Arg Arg Ser Ser Gly Gly Gly Ala Val Phe His
 65 70 75 80
 Asp Leu Gly Asn Thr Cys Phe Thr Phe Met Ala Gly Lys Pro Glu Tyr
 85 90 95
 Asp Lys Thr Ile Ser Thr Ser Ile Val Leu Asn Ala Leu Asn Ala Leu
 100 105 110
 Gly Val Ser Ala Glu Ala Ser Gly Arg Asn Asp Leu Val Val Lys Thr
 115 120 125
 Val Glu Gly Asp Arg Lys Val Ser Gly Ser Ala Tyr Arg Glu Thr Lys
 130 135 140
 Asp Arg Gly Phe His His Gly Thr Leu Leu Leu Asn Ala Asp Leu Ser
 145 150 155 160
 Arg Leu Ala Asn Tyr Leu Asn Pro Asp Lys Lys Lys Leu Ala Ala Lys
 165 170 175
 Gly Ile Thr Ser Val Arg Ser Arg Val Thr Asn Leu Thr Glu Leu Leu
 180 185 190
 Pro Gly Ile Thr His Glu Gln Val Cys Glu Ala Ile Thr Glu Ala Phe
 195 200 205
 Phe Ala His Tyr Gly Glu Arg Val Glu Ala Glu Ile Ile Ser Pro Asn
 210 215 220
 Lys Thr Pro Asp Leu Pro Asn Phe Ala Glu Thr Phe Ala Arg Gln Ser
 225 230 235 240
 Ser Trp Glu Trp Asn Phe Gly Gln Ala Pro Ala Phe Ser His Leu Leu
 245 250 255
 Asp Glu Arg Phe Thr Trp Gly Gly Val Glu Leu His Phe Asp Val Glu

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

 305 310 315 320
 Glu Gln Glu Lys Glu Leu Arg Glu Leu Ser Ala Trp Met Ala Gly Ala
 325 330 335
 Val Arg

<210> 124

<211> 1017

<212> DNA

<213> Klebsiella oxytoca

<220><221> CDS

<222> (1)..(1017)

<223> lplA gene encoding lipoate-protein ligase A (LplA)

<400> 124

atg tca acc ttg cgt ttg ctg tta tcc gac agt tat gac cca tgg ttt 48
 Met Ser Thr Leu Arg Leu Leu Leu Ser Asp Ser Tyr Asp Pro Trp Phe

1 5 10 15
 aac ctt gcc gta gag gag tcc att ttc cgc cag atg cca gcg aca cag 96
 Asn Leu Ala Val Glu Glu Ser Ile Phe Arg Gln Met Pro Ala Thr Gln

20 25 30
 cgt gtc ttg ttt ttg tgg cgt aac gcc gat acc gta gtt atc gga cgc 144
 Arg Val Leu Phe Leu Trp Arg Asn Ala Asp Thr Val Val Ile Gly Arg

35 40 45

 gca cag aat cca tgg aag gag tgt aac aca cgt cgc atg gag gag gac 192
 Ala Gln Asn Pro Trp Lys Glu Cys Asn Thr Arg Arg Met Glu Glu Asp

50 55 60
 aat gtg cgt ctg gct cgt cgc tct tcc ggg gga ggt gct gtg ttt cat 240

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

225	230	235	240	
tcg tgg gaa tgg aac ttc ggc caa gct cca gcc ttt agt cac ctg ctt	768			
Ser Trp Glu Trp Asn Phe Gly Gln Ala Pro Ala Phe Ser His Leu Leu				
245	250	255		
gac gaa cgc ttc aca tgg gga ggc gtt gag ctt cat ttc gac gtg gag	816			
Asp Glu Arg Phe Thr Trp Gly Gly Val Glu Leu His Phe Asp Val Glu				
260	265	270		
aag gga cac atc aca cgc acc cag att ttt acc gac agc ctt aac cca	864			
Lys Gly His Ile Thr Arg Thr Gln Ile Phe Thr Asp Ser Leu Asn Pro				
275	280	285		
gca ccg ctg gag gct ttg gcc gcc cgc tta caa ggt tgc ctt tat cgc	912			
Ala Pro Leu Glu Ala Leu Ala Ala Arg Leu Gln Gly Cys Leu Tyr Arg				
290	295	300		
gcc gac atg ctg caa caa gag tgc gat gct ctg tta gta gac ttt cca	960			
Ala Asp Met Leu Gln Gln Glu Cys Asp Ala Leu Leu Val Asp Phe Pro				
305	310	315	320	
gag cag gag aaa gca tta cgt gag ctg tca gcc tgg att gct ggt gca	1008			
Glu Gln Glu Lys Ala Leu Arg Glu Leu Ser Ala Trp Ile Ala Gly Ala				
325	330	335		
gta cgt taa	1017			
Val Arg				

<210> 125

<211> 338

<212> PRT

<213> Klebsiella oxytoca

<400> 125

Met Ser Thr Leu Arg Leu Leu Leu Ser Asp Ser Tyr Asp Pro Trp Phe

1 5 10 15

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

20	25	30
Arg Val Leu Phe Leu Trp Arg Asn Ala Asp Thr Val Val Ile Gly Arg		
35	40	45
Ala Gln Asn Pro Trp Lys Glu Cys Asn Thr Arg Arg Met Glu Glu Asp		
50	55	60
Asn Val Arg Leu Ala Arg Arg Ser Ser Gly Gly Gly Ala Val Phe His		
65	70	75
Asp Leu Gly Asn Thr Cys Phe Thr Phe Met Ala Gly Lys Pro Glu Tyr		
85	90	95
Asp Lys Thr Val Ser Thr Asn Ile Val Leu Thr Ala Leu Asn Ala Leu		
100	105	110
Gly Val Ala Ala Glu Ala Ser Gly Arg Asn Asp Leu Val Val Lys Thr		
115	120	125
Ala Glu Gly Asp Arg Lys Val Ser Gly Ser Ala Tyr Arg Glu Thr Met		
130	135	140
Asp Arg Gly Phe His His Gly Thr Leu Leu Leu Asn Ala Asp Leu Ser		
145	150	155
Arg Leu Ala Asn Tyr Leu Asn Pro Asp Lys Lys Lys Leu Gln Ala Lys		
165	170	175
Gly Ile Thr Ser Val Arg Gly Arg Val Ala Asn Leu Val Glu Leu Leu		
180	185	190
Pro Gly Ile Thr His Gln Gln Val Cys Glu Ala Ile Gln Glu Ala Phe		
195	200	205
Phe Ala His Tyr Gly Glu Arg Val Glu Ala Glu Val Ile Ser Pro Glu		
210	215	220
Lys Met Pro Asp Leu Pro Asn Phe Ala Ala Thr Phe Ala Arg Gln Ser		
225	230	235
Ser Trp Glu Trp Asn Phe Gly Gln Ala Pro Ala Phe Ser His Leu Leu		
245	250	255
Asp Glu Arg Phe Thr Trp Gly Gly Val Glu Leu His Phe Asp Val Glu		
260	265	270

Lys Gly His Ile Thr Arg Thr Gln Ile Phe Thr Asp Ser Leu Asn Pro
275 280 285

Ala Pro Leu Glu Ala Leu Ala Ala Arg Leu Gln Gly Cys Leu Tyr Arg
290 295 300

Ala Asp Met Leu Gln Gln Glu Cys Asp Ala Leu Leu Val Asp Phe Pro
305 310 315 320

Glu Gln Glu Lys Ala Leu Arg Glu Leu Ser Ala Trp Ile Ala Gly Ala
325 330 335

Val Arg

<210> 126

<211> 1854

<212> DNA

<213> Arabidopsis thaliana

<220><221> CDS

<222> (1)..(1854)

<223> Arabidopsis thaliana gene encoding TMP phosphatase [AT5G32470.1]

<400> 126

atg cgc ttc ctc ttc ccc acg cgc ctc atc aac aac tca tct ctc ggt 48

Met Arg Phe Leu Phe Pro Thr Arg Leu Ile Asn Asn Ser Ser Leu Gly
1 5 10 15

ctc ctc cga tct cca cac acc acc gcg ccg atc cgt tct ctc tgg ttt 96

Leu Leu Arg Ser Pro His Thr Thr Ala Pro Ile Arg Ser Leu Trp Phe
20 25 30

cgc acc aag tct ccg gtc ttc cga tcg gcg act act cca ata atg acg 144

Arg Thr Lys Ser Pro Val Phe Arg Ser Ala Thr Thr Pro Ile Met Thr

35 40 45

gcg gtc gct ttc tct tca tcg ttg tcg att ccc cct acc tcg gaa gaa 192

Ala Val Ala Phe Ser Ser Ser Leu Ser Ile Pro Pro Thr Ser Glu Glu
50 55 60

gca ctt cca ggg aag cta tgg atc aag ttt aac aga gag tgt ctc ttc 240

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

65	70	75	80	
tct atc tat agc ccc ttc gcc gtc tgt tta gcc gcc gga aat ctc aag				288
Ser Ile Tyr Ser Pro Phe Ala Val Cys Leu Ala Ala Gly Asn Leu Lys				
85	90	95		
atc gac aca ttt cgt cag tat att gca cag gat gtt cat ttc ctt aag				336
Ile Asp Thr Phe Arg Gln Tyr Ile Ala Gln Asp Val His Phe Leu Lys				
100	105	110		
gcc ttt gct cac gcg tat gaa ctg gcc gca gat tgt gct gat gac gat				384
Ala Phe Ala His Ala Tyr Glu Leu Ala Ala Asp Cys Ala Asp Asp Asp				
115	120	125		
gat gat aaa ttg gca att tct gat ttg agg aaa agc gtg atg gaa gaa				432
Asp Asp Lys Leu Ala Ile Ser Asp Leu Arg Lys Ser Val Met Glu Glu				
130	135	140		
ttg aaa atg cac gac tca ttt gta cag gat tgg gat tta gac atc aac				480
Leu Lys Met His Asp Ser Phe Val Gln Asp Trp Asp Leu Asp Ile Asn				
145	150	155	160	
aaa gaa gta agt gtt aac tca gca act ttg aga tac act gag ttc ttg				528
Lys Glu Val Ser Val Asn Ser Ala Thr Leu Arg Tyr Thr Glu Phe Leu				
165	170	175		
tta gct aca gca tcc gga aaa gta gaa gga tgc aaa gct ccc ggc atg				576
Leu Ala Thr Ala Ser Gly Lys Val Glu Gly Cys Lys Ala Pro Gly Met				
180	185	190		
ctt gat act cca ttt gaa aaa aca aaa gtt gct gcc tac acg ctt ggt				624
Leu Asp Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr Thr Leu Gly				
195	200	205		
gct gtg aca cct tgc atg cgg ttg tat gcc ttt ctc ggt aag gag ttt				672
Ala Val Thr Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Phe				
210	215	220		
gga tca ctt ctt gat ctg agt gat gtg aac cat ccc tac aag aaa tgg				720

Gly Ser Leu Leu Asp Leu Ser Asp Val Asn His Pro Tyr Lys Lys Trp
225 230 235 240
atc gat aat tat tct agt gat gct ttc cag gca tca gcc aag caa act 768
Ile Asp Asn Tyr Ser Ser Asp Ala Phe Gln Ala Ser Ala Lys Gln Thr
245 250 255
gaa gac ttg ctt gag aag ctt agt gtc tct atg act ggt gaa gaa ttg 816
Glu Asp Leu Leu Glu Lys Leu Ser Val Ser Met Thr Gly Glu Glu Leu
260 265 270
gac ata att gaa aaa ttg tat caa cag gct atg aaa ctt gaa gta gag 864
Asp Ile Ile Glu Lys Leu Tyr Gln Gln Ala Met Lys Leu Glu Val Glu
275 280 285
ttc ttc cat gcc cag cca ctt gcc cag cct acc ata gtt cca ctg ctc 912
Phe Phe His Ala Gln Pro Leu Ala Gln Pro Thr Ile Val Pro Leu Leu
290 295 300
aag aac cac tca aaa gat gat ctg gtg atc ttt tct gat ttt gat ctg 960
Lys Asn His Ser Lys Asp Asp Leu Val Ile Phe Ser Asp Phe Asp Leu
305 310 315 320
act tgc acc gtt gtg gat tct tct gct att tta gcg gaa ata gca att 1008
Thr Cys Thr Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala Ile
325 330 335
gta act gcc cca aaa gat gaa caa agt cga tct gga caa caa att cat 1056
Val Thr Ala Pro Lys Asp Glu Gln Ser Arg Ser Gly Gln Gln Ile His
340 345 350
cgg atg ctc tca tct gac ctt aag aac acc tgg aat cta ctt tct aaa 1104
Arg Met Leu Ser Ser Asp Leu Lys Asn Thr Trp Asn Leu Leu Ser Lys
355 360 365
caa tac aca gag cat tat gaa gaa tgc ata gag agt att ctg aat aaa 1152
Gln Tyr Thr Glu His Tyr Glu Glu Cys Ile Glu Ser Ile Leu Asn Lys
370 375 380
aag aaa gcg gac aag ttt gac tat gaa ggt tta tgt aaa gca cta gag 1200
Lys Lys Ala Asp Lys Phe Asp Tyr Glu Gly Leu Cys Lys Ala Leu Glu

385	390	395	400	
cag ctt tca gat ttt gag aaa gag gca aat aat cga gtg att gag tct				1248
Gln Leu Ser Asp Phe Glu Lys Glu Ala Asn Asn Arg Val Ile Glu Ser				
405	410	415		
ggg gta ctc aaa ggc ctg aat ctt gaa gac att aag cgc gct ggg gaa				1296
Gly Val Leu Lys Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Gly Glu				
420	425	430		
agg tta atc ctt caa gat gga tgc atc aat gtc ttc cag aaa att tta				1344
Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Val Phe Gln Lys Ile Leu				
435	440	445		
aag act gag aat ctg aat gca gaa ctt cat gtg ctt tcc tat tgt tgg				1392
Lys Thr Glu Asn Leu Asn Ala Glu Leu His Val Leu Ser Tyr Cys Trp				
450	455	460		
tgt ggt gac ctc atc agg gca gcc ttt tct gca ggc gga gta gat gca				1440
Cys Gly Asp Leu Ile Arg Ala Ala Phe Ser Ala Gly Gly Val Asp Ala				
465	470	475	480	
gtg gaa gta cat gca aat gaa ttc aca ttt gag gaa tcc atc tcg act				1488
Val Glu Val His Ala Asn Glu Phe Thr Phe Glu Glu Ser Ile Ser Thr				
485	490	495		
ggg gag atc gaa aga aag gtg gaa tcc cca att aac aaa gct caa cag				1536
Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asn Lys Ala Gln Gln				
500	505	510		
ttc aaa agt atc cta caa aac aga aag aat gag aac aat aag aaa agt				1584
Phe Lys Ser Ile Leu Gln Asn Arg Lys Asn Glu Asn Asn Lys Lys Ser				
515	520	525		
ttc ttg agt gtg tat att gga gat tcg gta ggt gac ttg ctg tgt ctc				1632
Phe Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu				
530	535	540		
ctc gaa gca gat ata gga ata gtg gtt agc tct agc tcg agt ctc agg				1680
Leu Glu Ala Asp Ile Gly Ile Val Val Ser Ser Ser Ser Ser Leu Arg				
545	550	555	560	

aga gtt gga agc cat ttt ggg gtc tca ttt gtg cct ttg ttt tct gga 1728

Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu Phe Ser Gly

565 570 575

atc gtc cag aaa cag aaa caa cac act gaa gaa tca tca tca tca gca 1776

Ile Val Gln Lys Gln Lys Gln His Thr Glu Glu Ser Ser Ser Ser Ala

580 585 590

tgg aaa gga ctc tct ggc aca ctt tac aca gtt tca agc tgg gcc gaa 1824

Trp Lys Gly Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu

595 600 605

att cat tca ttc gct ctt gga tgg gag taa 1854

Ile His Ser Phe Ala Leu Gly Trp Glu

610 615

<210> 127

<211> 617

<212> PRT

<213> Arabidopsis thaliana

<400> 127

Met Arg Phe Leu Phe Pro Thr Arg Leu Ile Asn Asn Ser Ser Leu Gly

1 5 10 15

Leu Leu Arg Ser Pro His Thr Thr Ala Pro Ile Arg Ser Leu Trp Phe

20 25 30

Arg Thr Lys Ser Pro Val Phe Arg Ser Ala Thr Thr Pro Ile Met Thr

35 40 45

Ala Val Ala Phe Ser Ser Ser Leu Ser Ile Pro Pro Thr Ser Glu Glu

50 55 60

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

65 70 75 80

Ser Ile Tyr Ser Pro Phe Ala Val Cys Leu Ala Ala Gly Asn Leu Lys

85 90 95

Ile Asp Thr Phe Arg Gln Tyr Ile Ala Gln Asp Val His Phe Leu Lys

100 105 110

Ala Phe Ala His Ala Tyr Glu Leu Ala Ala Asp Cys Ala Asp Asp Asp
115 120 125
Asp Asp Lys Leu Ala Ile Ser Asp Leu Arg Lys Ser Val Met Glu Glu
130 135 140

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

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

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

Val Thr Ala Pro Lys Asp Glu Gln Ser Arg Ser Gly Gln Gln Ile His
340 345 350
Arg Met Leu Ser Ser Asp Leu Lys Asn Thr Trp Asn Leu Leu Ser Lys

355 360 365
 Gln Tyr Thr Glu His Tyr Glu Glu Cys Ile Glu Ser Ile Leu Asn Lys
 370 375 380
 Lys Lys Ala Asp Lys Phe Asp Tyr Glu Gly Leu Cys Lys Ala Leu Glu
 385 390 395 400

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

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

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

 Trp Lys Gly Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu
 595 600 605

Ile His Ser Phe Ala Leu Gly Trp Glu

610

615

<210> 128

<211> 1806

<212> DNA

<213> Pyrus x bretschneideri

<220><221> CDS

<222> (1)..(1806)

<223> Pyrus x bretschneideri gene encoding TMP phosphatase

[XP_009379735.1]

<400> 128

atg cgc ata ctc ttc ccc cca aac cca atc aaa acc cca act ctc ttc 48

Met Arg Ile Leu Phe Pro Pro Asn Pro Ile Lys Thr Pro Thr Leu Phe

1

5

10

15

aac tcc ctc cgt ctg cga ttc aac tcg ctc cga tcc cac tgt gcc aac 96

Asn Ser Leu Arg Leu Arg Phe Asn Ser Leu Arg Ser His Cys Ala Asn

20

25

30

tca atg gcc gta cct ccg ccg aag tca gcc atg gct tcc gcc gtc gtc 144

Ser Met Ala Val Pro Pro Pro Lys Ser Ala Met Ala Ser Ala Val Val

35

40

45

ggc aac gag gtg ggt ctc gcc cgc cgc ttc tgg atc aag ttc aag cga 192

Gly Asn Glu Val Gly Leu Ala Arg Arg Phe Trp Ile Lys Phe Lys Arg

50

55

60

gaa tcg att ttc gct atg tac act ccc ttc acg ctc tgt ttg gct gct 240

Glu Ser Ile Phe Ala Met Tyr Thr Pro Phe Thr Leu Cys Leu Ala Ala

65

70

75

80

ggg aat ctc aag att gaa act ttc cgc gat tat att gcc caa gat gtt 288

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

85

90

95

cac ttt ctc aag gcc ttc gct cac gcg tat gaa ttg gca gaa gat tgt 336

His Phe Leu Lys Ala Phe Ala His Ala Tyr Glu Leu Ala Glu Asp Cys

100

105

110

gca gac gat gat gat gca aag ccc gtg att tct gag ttg agg agg gca 384
Ala Asp Asp Asp Asp Ala Lys Pro Val Ile Ser Glu Leu Arg Arg Ala

115 120 125
gtt ctg cag gag ctg aaa atg cat gat tca ttt gtg aag gaa tgg ggg 432
Val Leu Gln Glu Leu Lys Met His Asp Ser Phe Val Lys Glu Trp Gly

130 135 140
tta cag ggt gct aaa gag acc cct atc aac tcc gct gcg gtg aag tac 480
Leu Gln Gly Ala Lys Glu Thr Pro Ile Asn Ser Ala Ala Val Lys Tyr

145 150 155 160
aca gat ttc tta ttg gca aca gcc tct gga aaa gtt gaa gga gtc aag 528
Thr Asp Phe Leu Leu Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys

165 170 175
gga cct ggt aaa ctt gca act cca ttt gaa aga acc aaa gtg gct gct 576
Gly Pro Gly Lys Leu Ala Thr Pro Phe Glu Arg Thr Lys Val Ala Ala

180 185 190
tac acc ctt ggc gct atg act cct tgc atg aga ctg tat gcc ttt ctt 624

Tyr Thr Leu Gly Ala Met Thr Pro Cys Met Arg Leu Tyr Ala Phe Leu

195 200 205
ggt aag gag ttc aag gca ctt cta gat ccc agc gaa ggc agt cac ccg 672
Gly Lys Glu Phe Lys Ala Leu Leu Asp Pro Ser Glu Gly Ser His Pro

210 215 220
tac ttg aag tgg att gac agt tat tct tct aaa agt ttt cag gca tca 720
Tyr Leu Lys Trp Ile Asp Ser Tyr Ser Ser Lys Ser Phe Gln Ala Ser

225 230 235 240
gct gtg caa atc gaa gag ttg ctg gat aaa cta agt gtc tct ttg aca 768
Ala Val Gln Ile Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr

245 250 255
ggc gag gag ctt gac atc atc gaa aag ctt tac cac caa gca atg aaa 816
Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys

260	265	270	
ctt gag atc gag ttc ttc tct gct cag tct ctt gtt cag cca act gta			864
Leu Glu Ile Glu Phe Phe Ser Ala Gln Ser Leu Val Gln Pro Thr Val			
275	280	285	
gtt cct ctg atc aga gaa cat aac cct gca gaa gat cgg ctc atg ata			912
Val Pro Leu Ile Arg Glu His Asn Pro Ala Glu Asp Arg Leu Met Ile			
290	295	300	
ttt tct gat ttt gat ttg act tgt aca gtc gtt gat tca tct gcc att			960
Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser Ala Ile			
305	310	315	320
ttg gct gaa att gca ata gta aca gca cca aaa tct gat caa cat caa			1008
Leu Ala Glu Ile Ala Ile Val Thr Ala Pro Lys Ser Asp Gln His Gln			
325	330	335	
ccc gaa aat cag att gct cgg atg tct tgg gct gat ctc agg aat aca			1056
Pro Glu Asn Gln Ile Ala Arg Met Ser Ser Ala Asp Leu Arg Asn Thr			
340	345	350	
tgg ggt ctt ctt tcc agg cag tac aca gaa gag tat gag caa tgc ata			1104
Trp Gly Leu Leu Ser Arg Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile			
355	360	365	
gaa agc att gtt ccc act gaa aaa gca gtg ttt gac tat gaa aat ttg			1152
Glu Ser Ile Val Pro Thr Glu Lys Ala Val Phe Asp Tyr Glu Asn Leu			
370	375	380	
ctt aaa gca cta gag aaa ctt tca gat ttt gag agg aag gca aac aat			1200
Leu Lys Ala Leu Glu Lys Leu Ser Asp Phe Glu Arg Lys Ala Asn Asn			
385	390	395	400
aga gtc acg aag tct gaa gta ctc aag ggt ctt aat ctc gaa gat ata			1248
Arg Val Thr Lys Ser Glu Val Leu Lys Gly Leu Asn Leu Glu Asp Ile			
405	410	415	
aaa aga gct ggt gaa cgt ctc att ctt caa gat ggc tgt att aat ttc			1296

Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Phe
420 425 430
ttt cag aaa att gcc aag agt gaa aac ttg aat gca aat gtt cat gtt 1344
Phe Gln Lys Ile Ala Lys Ser Glu Asn Leu Asn Ala Asn Val His Val
435 440 445
ctt tca tac tgt tgg tgt ggt gat ctc ata aga tcg gcc ttt tca tca 1392
Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Ala Phe Ser Ser
450 455 460
ggg ggt tta aac gag ctg gat gta cat gca aat gag ttt acc ttc gag 1440
Gly Gly Leu Asn Glu Leu Asp Val His Ala Asn Glu Phe Thr Phe Glu
465 470 475 480
gaa tcc atc tcc aca ggt gat att gtt aag aag gtg gag tcc cct att 1488
Glu Ser Ile Ser Thr Gly Asp Ile Val Lys Lys Val Glu Ser Pro Ile
485 490 495
gac aag gtt aaa tct ttt aaa gat att ttg aaa aat tgc agc aat gac 1536
Asp Lys Val Lys Ser Phe Lys Asp Ile Leu Lys Asn Cys Ser Asn Asp
500 505 510
aga aag aac ttg act gtt tac att gga gac tcg gtg ggt gac tta ctt 1584
Arg Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu
515 520 525
tgt ctg ctg gag gcg gat att gga atc gta att ggg tca agt tca agc 1632
Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser
530 535 540
ctt agg aga gtg gcg act cag ttt ggg gta tct ttt gtt ccg ttg ttc 1680
Leu Arg Arg Val Ala Thr Gln Phe Gly Val Ser Phe Val Pro Leu Phe
545 550 555 560
ccg ggt tta gtt aag aaa cag aaa gaa tgc aca gat gga agg tct cct 1728
Pro Gly Leu Val Lys Lys Gln Lys Glu Cys Thr Asp Gly Arg Ser Pro
565 570 575
agt tgg aaa ggg tta act ggt att ctt tac aca gtg aat agt tgg gcg 1776
Ser Trp Lys Gly Leu Thr Gly Ile Leu Tyr Thr Val Asn Ser Trp Ala

580 585 590
gaa ata cat gcc ttc att ttg ggg tgt taa 1806
Glu Ile His Ala Phe Ile Leu Gly Cys

595 600

<210> 129

<211> 601

<212> PRT

<213> Pyrus x bretschneideri

<400> 129

Met Arg Ile Leu Phe Pro Pro Asn Pro Ile Lys Thr Pro Thr Leu Phe

1 5 10 15

Asn Ser Leu Arg Leu Arg Phe Asn Ser Leu Arg Ser His Cys Ala Asn

20 25 30

Ser Met Ala Val Pro Pro Pro Lys Ser Ala Met Ala Ser Ala Val Val

35 40 45

Gly Asn Glu Val Gly Leu Ala Arg Arg Phe Trp Ile Lys Phe Lys Arg

50 55 60

Glu Ser Ile Phe Ala Met Tyr Thr Pro Phe Thr Leu Cys Leu Ala Ala

65 70 75 80

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

85 90 95

His Phe Leu Lys Ala Phe Ala His Ala Tyr Glu Leu Ala Glu Asp Cys

100 105 110

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

115 120 125

Val Leu Gln Glu Leu Lys Met His Asp Ser Phe Val Lys Glu Trp Gly

130 135 140

Leu Gln Gly Ala Lys Glu Thr Pro Ile Asn Ser Ala Ala Val Lys Tyr

145 150 155 160

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

165 170 175

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

180 185 190

Tyr Thr Leu Gly Ala Met Thr Pro Cys Met Arg Leu Tyr Ala Phe Leu

195 200 205

Gly Lys Glu Phe Lys Ala Leu Leu Asp Pro Ser Glu Gly Ser His Pro

210 215 220

Tyr Leu Lys Trp Ile Asp Ser Tyr Ser Ser Lys Ser Phe Gln Ala Ser

225 230 235 240

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

245 250 255

Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys

260 265 270

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

275 280 285

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

290 295 300

Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser Ala Ile

305 310 315 320

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

325 330 335

Pro Glu Asn Gln Ile Ala Arg Met Ser Ser Ala Asp Leu Arg Asn Thr

340 345 350

Trp Gly Leu Leu Ser Arg Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile

355 360 365

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

370 375 380

Leu Lys Ala Leu Glu Lys Leu Ser Asp Phe Glu Arg Lys Ala Asn Asn

385 390 395 400

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

405 410 415

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

420 425 430
Phe Gln Lys Ile Ala Lys Ser Glu Asn Leu Asn Ala Asn Val His Val

435 440 445
Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Ala Phe Ser Ser

450 455 460
Gly Gly Leu Asn Glu Leu Asp Val His Ala Asn Glu Phe Thr Phe Glu
465 470 475 480

Glu Ser Ile Ser Thr Gly Asp Ile Val Lys Lys Val Glu Ser Pro Ile
485 490 495
Asp Lys Val Lys Ser Phe Lys Asp Ile Leu Lys Asn Cys Ser Asn Asp

500 505 510
Arg Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu

515 520 525
Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser
530 535 540

Leu Arg Arg Val Ala Thr Gln Phe Gly Val Ser Phe Val Pro Leu Phe
545 550 555 560
Pro Gly Leu Val Lys Lys Gln Lys Glu Cys Thr Asp Gly Arg Ser Pro

565 570 575
Ser Trp Lys Gly Leu Thr Gly Ile Leu Tyr Thr Val Asn Ser Trp Ala

580 585 590
Glu Ile His Ala Phe Ile Leu Gly Cys

595 600

<210> 130

<211> 1797

<212> DNA

<213> Brassica napus

<220><221> CDS

<222> (1)..(1797)

<223> Brassica napus gene encoding TMP phosphatase [CDY62623.1]

<400> 130

atg cgc atc ctc aac aac tcg ctc gcc ctt ctc cga tcg ccc cgc gcc	48
Met Arg Ile Leu Asn Asn Ser Leu Ala Leu Leu Arg Ser Pro Arg Ala	
1 5 10 15	
gcc gcc ccg atc cgt tct cta ctg ttc ggc agc aag aag tct tcc gtc	96
Ala Ala Pro Ile Arg Ser Leu Leu Phe Gly Ser Lys Lys Ser Ser Val	
20 25 30	
tcc cga tcg gcg gcc gcc ttc tct tcg gcg atg tcg att cct cct cct	144
Ser Arg Ser Ala Ala Ala Phe Ser Ser Ala Met Ser Ile Pro Pro Pro	
35 40 45	
agc ata tcc acc tcg gaa gaa gct ctg gcg ggg agg ctg tgg atc aag	192
Ser Ile Ser Thr Ser Glu Glu Ala Leu Ala Gly Arg Leu Trp Ile Lys	
50 55 60	
ttc aac aga gag tgc ctc ttc tct atg tac agc ccc ttc gcc gtt tct	240
Phe Asn Arg Glu Cys Leu Phe Ser Met Tyr Ser Pro Phe Ala Val Ser	
65 70 75 80	
ttg gcc gcc ggc aat ctc aag atc gag acc ttc cgg cag tat att gct	288
Leu Ala Ala Gly Asn Leu Lys Ile Glu Thr Phe Arg Gln Tyr Ile Ala	
85 90 95	
cag gat gtt cat ttc ctc aag gcc ttt gct cac gcg tat gag ttg gcc	336
Gln Asp Val His Phe Leu Lys Ala Phe Ala His Ala Tyr Glu Leu Ala	
100 105 110	
gca gag tgt gct gat gat gat gat gat aag ttg gca att tct gac ttg	384
Ala Glu Cys Ala Asp Asp Asp Asp Asp Lys Leu Ala Ile Ser Asp Leu	
115 120 125	
agg aaa agc gtc atg gat gag ttg aaa atg cac aac tca ttt gta cag	432
Arg Lys Ser Val Met Asp Glu Leu Lys Met His Asn Ser Phe Val Gln	
130 135 140	
gat tgg gat tta gac atc agc aaa gaa gta agt gtt aac tca gca aca	480
Asp Trp Asp Leu Asp Ile Ser Lys Glu Val Ser Val Asn Ser Ala Thr	

145	150	155	160	
ttg aga tac acc gag ttc tta tta gct aca tca tcc gga aaa gta gaa				528
Leu Arg Tyr Thr Glu Phe Leu Leu Ala Thr Ser Ser Gly Lys Val Glu				
	165	170	175	
gga ctc aaa gct ccc ggc atg ctt gat act cca ttt gag aaa acc aaa				576
Gly Leu Lys Ala Pro Gly Met Leu Asp Thr Pro Phe Glu Lys Thr Lys				
	180	185	190	
gtg gcc gcc tac acg ctt ggt gct gtg aca cct tgc atg aag ctg tat				624
Val Ala Ala Tyr Thr Leu Gly Ala Val Thr Pro Cys Met Lys Leu Tyr				
	195	200	205	
gcc ttt ctt ggt aag gag ttt gga gcg ctt cta gat tcg agt gaa gcg				672
Ala Phe Leu Gly Lys Glu Phe Gly Ala Leu Leu Asp Ser Ser Glu Ala				
	210	215	220	
aac cat ccc tac aag aaa tgg atc gaa aat tat tct agt gat gca ttc				720
Asn His Pro Tyr Lys Lys Trp Ile Glu Asn Tyr Ser Ser Asp Ala Phe				
225	230	235	240	
cag gca tca gct aag caa act gaa gac ttg ctt gag aag ctt agt gtg				768
Gln Ala Ser Ala Lys Gln Thr Glu Asp Leu Leu Glu Lys Leu Ser Val				
	245	250	255	
tgt atg act ggc gaa gag ctg gac atc att gaa aaa ctg tat caa cag				816
Cys Met Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr Gln Gln				
	260	265	270	
gca atg aaa ctt gaa gta gag ttc ttc cac gca caa ccg ttt gct cag				864
Ala Met Lys Leu Glu Val Glu Phe Phe His Ala Gln Pro Phe Ala Gln				
	275	280	285	
cct acc ata gtt ccg ctg ctg aag aac cat tca aaa gat gag ctg atg				912
Pro Thr Ile Val Pro Leu Leu Lys Asn His Ser Lys Asp Glu Leu Met				
	290	295	300	
ata ttt tct gat ttt gat ctg act tgc acc gtt gtt gat tct tct gct				960
Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser Ala				
305	310	315	320	

att tta gcc gaa att gca atc gta act gcc ccg aaa gat gat cag ggt	1008
Ile Leu Ala Glu Ile Ala Ile Val Thr Ala Pro Lys Asp Asp Gln Gly	
325 330 335	
caa caa att aat cgg atg ctt tcg gct gac ctt aag aac acc tgg agt	1056
Gln Gln Ile Asn Arg Met Leu Ser Ala Asp Leu Lys Asn Thr Trp Ser	
340 345 350	
cta ctt tcc aaa cag tat aca gag cac tat gaa gag tgc ata gag agt	1104
Leu Leu Ser Lys Gln Tyr Thr Glu His Tyr Glu Glu Cys Ile Glu Ser	
355 360 365	
att ctg aat aag gaa aaa gcg gac aag ttt gac tac gag ggt ttg tgt	1152
Ile Leu Asn Lys Glu Lys Ala Asp Lys Phe Asp Tyr Glu Gly Leu Cys	
370 375 380	
gaa gca cta gag cag ctg tca gag ttt gag aag aaa gca aac gac cga	1200
Glu Ala Leu Glu Gln Leu Ser Glu Phe Glu Lys Lys Ala Asn Asp Arg	
385 390 395 400	
gtg ata gag tct ggt gta ctc aag ggc ctg aat ctc gat gac atc aag	1248
Val Ile Glu Ser Gly Val Leu Lys Gly Leu Asn Leu Asp Asp Ile Lys	
405 410 415	
cga gct ggg gaa agg ttg att ctt caa gat ggc tgc atc aat gtc ttc	1296
Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Val Phe	
420 425 430	
cag aaa att ttg aag act cag gat gtg aat gca aaa ctc cac gtg ctt	1344
Gln Lys Ile Leu Lys Thr Gln Asp Val Asn Ala Lys Leu His Val Leu	
435 440 445	
tcg tat tgt tgg tgt ggt gac ctc atc aga gca gcc ttt tct gca cgg	1392
Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ala Ala Phe Ser Ala Arg	
450 455 460	
gga gta gat gca gtg gaa gta cat gca aat gaa ttc aca ttc gag gaa	1440
Gly Val Asp Ala Val Glu Val His Ala Asn Glu Phe Thr Phe Glu Glu	
465 470 475 480	
tcc atc tct act gga gaa ata gaa aga aaa gtg gaa tcc cca atc gac	1488

Ser Ile Ser Thr Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asp

485 490 495
aag gct caa cag ttc aag agc atc cta caa aac aga aag aag gat gag 1536

Lys Ala Gln Gln Phe Lys Ser Ile Leu Gln Asn Arg Lys Lys Asp Glu

500 505 510
gag aaa agc atc ctc act gtt tac att gga gat tca gta ggt gac ttg 1584

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

515 520 525
ctc tgt ctc ctg gag gca gac att gga ata gtg gtc gcc tct agc tcg 1632

Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Ala Ser Ser Ser

530 535 540
agc ctc agg aga gtg gga agc cat ttc ggg gtc tca ttt gtg cct ttg 1680

Ser Leu Arg Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu

545 550 555 560
ttc tct gga att gtg caa aaa cag aaa caa gaa gaa acc tgg aag ggg 1728

Phe Ser Gly Ile Val Gln Lys Gln Lys Gln Glu Glu Thr Trp Lys Gly

565 570 575
ctc tct ggc aca ctt tac acg gta tca agc tgg gct gaa ata cat tcc 1776

Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu Ile His Ser

580 585 590
ttc gct ctt gga tgg gag taa 1797

Phe Ala Leu Gly Trp Glu

595

<210> 131

<211> 598

<212> PRT

<213> Brassica napus

<400> 131

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

1 5 10 15

Ala Ala Pro Ile Arg Ser Leu Leu Phe Gly Ser Lys Lys Ser Ser Val

20 25 30
 Ser Arg Ser Ala Ala Ala Phe Ser Ser Ala Met Ser Ile Pro Pro Pro
 35 40 45

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

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

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

 Gln Ala Ser Ala Lys Gln Thr Glu Asp Leu Leu Glu Lys Leu Ser Val
 245 250 255
 Cys Met Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr Gln Gln
 260 265 270

Ala Met Lys Leu Glu Val Glu Phe Phe His Ala Gln Pro Phe Ala Gln
275 280 285

Pro Thr Ile Val Pro Leu Leu Lys Asn His Ser Lys Asp Glu Leu Met
290 295 300

Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser Ala
305 310 315 320

Ile Leu Ala Glu Ile Ala Ile Val Thr Ala Pro Lys Asp Asp Gln Gly
325 330 335

Gln Gln Ile Asn Arg Met Leu Ser Ala Asp Leu Lys Asn Thr Trp Ser
340 345 350

Leu Leu Ser Lys Gln Tyr Thr Glu His Tyr Glu Glu Cys Ile Glu Ser
355 360 365

Ile Leu Asn Lys Glu Lys Ala Asp Lys Phe Asp Tyr Glu Gly Leu Cys
370 375 380

Glu Ala Leu Glu Gln Leu Ser Glu Phe Glu Lys Lys Ala Asn Asp Arg
385 390 395 400

Val Ile Glu Ser Gly Val Leu Lys Gly Leu Asn Leu Asp Asp Ile Lys
405 410 415

Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Val Phe
420 425 430

Gln Lys Ile Leu Lys Thr Gln Asp Val Asn Ala Lys Leu His Val Leu
435 440 445

Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ala Ala Phe Ser Ala Arg
450 455 460

Gly Val Asp Ala Val Glu Val His Ala Asn Glu Phe Thr Phe Glu Glu
465 470 475 480

Ser Ile Ser Thr Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asp
485 490 495

Lys Ala Gln Gln Phe Lys Ser Ile Leu Gln Asn Arg Lys Lys Asp Glu
500 505 510

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

515 520 525
 Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Ala Ser Ser Ser
 530 535 540
 Ser Leu Arg Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu
 545 550 555 560

 Phe Ser Gly Ile Val Gln Lys Gln Lys Gln Glu Glu Thr Trp Lys Gly
 565 570 575
 Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu Ile His Ser
 580 585 590
 Phe Ala Leu Gly Trp Glu
 595
 <210> 132
 <211> 1815
 <212> DNA
 <213> Glycine max
 <220><221> CDS
 <222> (1)..(1815)
 <223> Glycine max gene encoding TMP phosphatase [XP_003536133.1]
 <400> 132
 atg cgc atg cgg tgg ttc ctc cga agc cca atc atc aaa acc tcg ctg 48

 Met Arg Met Arg Trp Phe Leu Arg Ser Pro Ile Ile Lys Thr Ser Leu
 1 5 10 15
 ctg aat ctg agc cct cca att tcg ttt aga cct cac tgg gcg agg agg 96
 Leu Asn Leu Ser Pro Pro Ile Ser Phe Arg Pro His Trp Ala Arg Arg
 20 25 30
 acc ttc act tct tcg aga ttg tca atg gcg gcc atc cac aac cac agc 144
 Thr Phe Thr Ser Ser Arg Leu Ser Met Ala Ala Ile His Asn His Ser

 35 40 45
 aac agc aac agc gaa acc gga ctc gcg aga cgg ttt tgg atc aag ttc 192
 Asn Ser Asn Ser Glu Thr Gly Leu Ala Arg Arg Phe Trp Ile Lys Phe
 50 55 60
 act cgt gaa tcc atc ttc gcc atg tac act ccc ttc gcc atc gcc ttg 240

Thr	Arg	Glu	Ser	Ile	Phe	Ala	Met	Tyr	Thr	Pro	Phe	Ala	Ile	Ala	Leu	
65				70				75				80				
gcc	tcc	ggt	aat	ttg	cac	att	gat	tcc	ttc	cac	cat	tac	atc	gcc	caa	288
Ala	Ser	Gly	Asn	Leu	His	Ile	Asp	Ser	Phe	His	His	Tyr	Ile	Ala	Gln	
85				90				95								
gac	gtt	cat	ttc	cta	cgc	gcc	ttt	gct	caa	gcg	tat	gag	ttg	gct	gaa	336
Asp	Val	His	Phe	Leu	Arg	Ala	Phe	Ala	Gln	Ala	Tyr	Glu	Leu	Ala	Glu	
100				105				110								
gag	tgt	gct	gat	gac	gac	gat	gcg	aaa	ctt	gga	atc	tgt	gag	ttg	agg	384
Glu	Cys	Ala	Asp	Asp	Asp	Asp	Ala	Lys	Leu	Gly	Ile	Cys	Glu	Leu	Arg	
115				120				125								
aag	gca	gtt	cta	gag	gag	ctg	aag	atg	cac	aac	ttg	ctg	gta	cag	gaa	432
Lys	Ala	Val	Leu	Glu	Glu	Leu	Lys	Met	His	Asn	Leu	Leu	Val	Gln	Glu	
130				135				140								
cgg	gag	ttg	gac	ctt	gcc	aaa	gag	cat	ggt	att	aat	tct	gca	act	gtt	480
Arg	Glu	Leu	Asp	Leu	Ala	Lys	Glu	His	Gly	Ile	Asn	Ser	Ala	Thr	Val	
145				150				155				160				
aag	tac	aca	gag	ttc	ctg	ctg	gct	aca	gcc	tct	ggg	aag	att	gaa	gga	528
Lys	Tyr	Thr	Glu	Phe	Leu	Leu	Ala	Thr	Ala	Ser	Gly	Lys	Ile	Glu	Gly	
165				170				175								
cta	aaa	ggt	cct	ggt	aaa	ctt	gct	aca	cca	ttt	gag	aaa	aca	aaa	att	576
Leu	Lys	Gly	Pro	Gly	Lys	Leu	Ala	Thr	Pro	Phe	Glu	Lys	Thr	Lys	Ile	
180				185				190								
gct	gct	tat	act	tta	ggt	gcc	atg	act	cct	tgc	atg	agg	ctt	tat	gcc	624
Ala	Ala	Tyr	Thr	Leu	Gly	Ala	Met	Thr	Pro	Cys	Met	Arg	Leu	Tyr	Ala	
195				200				205								
gtt	atg	gga	aag	aag	ttc	cag	gaa	ctt	ttg	gat	tcc	aat	gaa	agt	act	672
Val	Met	Gly	Lys	Lys	Phe	Gln	Glu	Leu	Leu	Asp	Ser	Asn	Glu	Ser	Thr	
210				215				220								

cac	cca	tat	aac	aag	tgg	atc	aac	aac	tat	tcc	tct	gat	ggg	ttc	cag		720
His	Pro	Tyr	Asn	Lys	Trp	Ile	Asn	Asn	Tyr	Ser	Ser	Asp	Gly	Phe	Gln		
225					230					235					240		
gct	act	act	ctg	caa	act	gaa	gat	ttg	ctc	gac	aaa	cta	agt	gtc	tct		768
Ala	Thr	Thr	Leu	Gln	Thr	Glu	Asp	Leu	Leu	Asp	Lys	Leu	Ser	Val	Ser		
				245					250					255			
ttg	act	ggg	gaa	gaa	ctt	gat	gtc	att	gaa	aag	ctt	tat	tac	caa	gca		816
Leu	Thr	Gly	Glu	Glu	Leu	Asp	Val	Ile	Glu	Lys	Leu	Tyr	Tyr	Gln	Ala		
			260					265					270				
atg	aag	ctt	gaa	ata	gag	ttc	ttc	tct	gct	cag	cca	ctc	ttc	cag	cca		864
Met	Lys	Leu	Glu	Ile	Glu	Phe	Phe	Ser	Ala	Gln	Pro	Leu	Phe	Gln	Pro		
	275						280					285					
act	ata	gta	ccc	ttg	act	aaa	gga	cat	aag	cct	gtg	gaa	gat	cat	ctc		912
Thr	Ile	Val	Pro	Leu	Thr	Lys	Gly	His	Lys	Pro	Val	Glu	Asp	His	Leu		
290						295					300						
att	att	ttt	tct	gat	ttt	gat	tta	aca	tgc	acc	gta	gtt	gat	tcg	tcc		960
Ile	Ile	Phe	Ser	Asp	Phe	Asp	Leu	Thr	Cys	Thr	Val	Val	Asp	Ser	Ser		
305					310					315					320		
gcc	atc	ttg	gct	gaa	att	gcc	ata	gtg	acg	gca	cca	aaa	tct	gat	cag		1008
Ala	Ile	Leu	Ala	Glu	Ile	Ala	Ile	Val	Thr	Ala	Pro	Lys	Ser	Asp	Gln		
				325					330					335			
aat	cag	cct	gaa	gat	caa	att	gtt	cgg	atg	tta	tct	tct	gac	ctc	agg		1056
Asn	Gln	Pro	Glu	Asp	Gln	Ile	Val	Arg	Met	Leu	Ser	Ser	Asp	Leu	Arg		
		340						345					350				
aat	aca	tgg	ggg	ttt	cta	tct	aaa	cag	tat	acg	gag	gag	tat	gag	caa		1104
Asn	Thr	Trp	Gly	Phe	Leu	Ser	Lys	Gln	Tyr	Thr	Glu	Glu	Tyr	Glu	Gln		
		355						360					365				
tgt	ata	gaa	agc	att	atg	cct	ccc	gat	aga	ttg	aac	aat	ttc	gat	tac		1152
Cys	Ile	Glu	Ser	Ile	Met	Pro	Pro	Asp	Arg	Leu	Asn	Asn	Phe	Asp	Tyr		

370	375	380	
aaa gaa ttg tcg atg gcc ctt gag caa ctt tca aaa ttt gag aac act			1200
Lys Glu Leu Ser Met Ala Leu Glu Gln Leu Ser Lys Phe Glu Asn Thr			
385	390	395	400
gca aat aat agg gtt atc gag tca ggg gta ctc aag ggt ata agt cta			1248
Ala Asn Asn Arg Val Ile Glu Ser Gly Val Leu Lys Gly Ile Ser Leu			
405	410	415	
gaa gat ata aag cgt gct gga gag cgt ctg ata cta caa gat ggt tgc			1296
Glu Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys			
420	425	430	
cct aac ttc ttt cag agc att gtt aag aat gaa aat ttg aat gcc aac			1344
Pro Asn Phe Phe Gln Ser Ile Val Lys Asn Glu Asn Leu Asn Ala Asn			
435	440	445	
gtg cat gtt ctt tca tac tgc tgg tgt ggt gac ctc att agg tct act			1392
Val His Val Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Thr			
450	455	460	
ttc tct tcc gct gat tta aat gag ttg aat gtt cat gct aat gag ttc			1440
Phe Ser Ser Ala Asp Leu Asn Glu Leu Asn Val His Ala Asn Glu Phe			
465	470	475	480
act tat gag gga tct gtt tcc acg ggt gaa att gtt aag aaa gtg gag			1488
Thr Tyr Glu Gly Ser Val Ser Thr Gly Glu Ile Val Lys Lys Val Glu			
485	490	495	
tct ccc att gac aag gtt gaa gct ttt cgt aac ata ttg aaa aat tgc			1536
Ser Pro Ile Asp Lys Val Glu Ala Phe Arg Asn Ile Leu Lys Asn Cys			
500	505	510	
aat gat gac aaa aag aaa tta act gtt tac att ggc gat tca gtg ggt			1584
Asn Asp Asp Lys Lys Lys Leu Thr Val Tyr Ile Gly Asp Ser Val Gly			
515	520	525	
gat tta ctt tgc cta ctt gaa gct gat gta gga att gtg att ggt tca			1632
Asp Leu Leu Cys Leu Leu Glu Ala Asp Val Gly Ile Val Ile Gly Ser			
530	535	540	

agt tca agc ctt aga agt gta ggg acg cag ttt ggt att tca ttt gtc 1680
 Ser Ser Ser Leu Arg Ser Val Gly Thr Gln Phe Gly Ile Ser Phe Val
 545 550 555 560
 cca ttg tat tct ggc ttg gtt aag aaa cag aaa gaa tat gtt gaa gga 1728

 Pro Leu Tyr Ser Gly Leu Val Lys Lys Gln Lys Glu Tyr Val Glu Gly
 565 570 575
 agc act tct gat tgg aag ggt tta tct ggc att ctt tac aca gtc tct 1776
 Ser Thr Ser Asp Trp Lys Gly Leu Ser Gly Ile Leu Tyr Thr Val Ser
 580 585 590
 agt tgg gct gaa gtg cat gct ttt att ttg ggt tgc tag 1815
 Ser Trp Ala Glu Val His Ala Phe Ile Leu Gly Cys

 595 600
 <210> 133
 <211> 604
 <212> PRT
 <213> Glycine max
 <400> 133
 Met Arg Met Arg Trp Phe Leu Arg Ser Pro Ile Ile Lys Thr Ser Leu
 1 5 10 15
 Leu Asn Leu Ser Pro Pro Ile Ser Phe Arg Pro His Trp Ala Arg Arg
 20 25 30
 Thr Phe Thr Ser Ser Arg Leu Ser Met Ala Ala Ile His Asn His Ser
 35 40 45

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

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

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

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

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

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

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

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

 Pro Leu Tyr Ser Gly Leu Val Lys Lys Gln Lys Glu Tyr Val Glu Gly
 565 570 575
 Ser Thr Ser Asp Trp Lys Gly Leu Ser Gly Ile Leu Tyr Thr Val Ser
 580 585 590
 Ser Trp Ala Glu Val His Ala Phe Ile Leu Gly Cys
 595 600

<210> 134
 <211> 1845
 <212> DNA
 <213> Nicotiana tomentosiformis
 <220><221> CDS
 <222> (1)..(1845)
 <223> Nicotiana tomentosiformi gene encoding TMP phosphatase

[XP_009615535.1]

<400> 134
 atg cgc ttc tca tta tta tcg ccc ctt gtt ctt aac cca gtc atc aga 48
 Met Arg Phe Ser Leu Leu Ser Pro Leu Val Leu Asn Pro Val Ile Arg
 1 5 10 15
 ttc tcc aat tcc aac gcg ctt ttt ggg tta cga ttc caa tta tac cct 96
 Phe Ser Asn Ser Asn Ala Leu Phe Gly Leu Arg Phe Gln Leu Tyr Pro
 20 25 30
 cgt tac tct cgg tat tta cga tcg ccc gtt aca atg gcg tcg gcg aaa 144

 Arg Tyr Ser Arg Tyr Leu Arg Ser Pro Val Thr Met Ala Ser Ala Lys
 35 40 45
 cca aag ccg gcg gcg gcg gtg aac aag ttt ccg gta gag gag gaa tgt 192
 Pro Lys Pro Ala Ala Ala Val Asn Lys Phe Pro Val Glu Glu Glu Cys
 50 55 60
 gtg ggt ata gcg agg aag tgt tgg atc aag ttc aag aga gag tct act 240
 Val Gly Ile Ala Arg Lys Cys Trp Ile Lys Phe Lys Arg Glu Ser Thr

 65 70 75 80
 ttc gct ctg tac act ccg ttt gtg gtt agt ttg gca tca gga acc cta 288
 Phe Ala Leu Tyr Thr Pro Phe Val Val Ser Leu Ala Ser Gly Thr Leu
 85 90 95
 aat ctg gac act ttc cgc cat tac att gct cag gat gtt cac ttc etc 336
 Asn Leu Asp Thr Phe Arg His Tyr Ile Ala Gln Asp Val His Phe Leu
 100 105 110

 aaa tcc ttc gct caa gcg tat gaa gct gca gaa gag tgt act gac gat 384

Lys Ser Phe Ala Gln Ala Tyr Glu Ala Ala Glu Glu Cys Thr Asp Asp
115 120 125
gac gat gcg aag gtt ggc att agt gag ttg cgg aag aat gtt att gaa 432
Asp Asp Ala Lys Val Gly Ile Ser Glu Leu Arg Lys Asn Val Ile Glu
130 135 140
gaa ctt aaa atg cat gat gca gtt tta aaa gag tgg ggc att gat ctg 480

Glu Leu Lys Met His Asp Ala Val Leu Lys Glu Trp Gly Ile Asp Leu
145 150 155 160
gtc aaa gag tcc agt ctt aac cct gca acg gcc aag tac aca gat ttt 528
Val Lys Glu Ser Ser Leu Asn Pro Ala Thr Ala Lys Tyr Thr Asp Phe
165 170 175
tta tca gct aca gct tca gga aag gtg gaa gga gta aaa gct gct aaa 576
Leu Ser Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys Ala Ala Lys

180 185 190
ctt gcc aca cca ttt gag aga acg aag ttg gca gct tat act cta ggt 624
Leu Ala Thr Pro Phe Glu Arg Thr Lys Leu Ala Ala Tyr Thr Leu Gly
195 200 205
gct atg act cct tgc atg agg ctt tac gcc tac att ggt aag gag ctg 672
Ala Met Thr Pro Cys Met Arg Leu Tyr Ala Tyr Ile Gly Lys Glu Leu
210 215 220

caa gtg ttc ctc gag gga gag aaa att cat cca tac aag aag tgg att 720
Gln Val Phe Leu Glu Gly Glu Lys Ile His Pro Tyr Lys Lys Trp Ile
225 230 235 240
gac agt tat gcc tct gaa agt ttc cag gca tca gct ctt caa acc gag 768
Asp Ser Tyr Ala Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Thr Glu
245 250 255
gac ttg ttg gat aaa ctg agt gtc cct ttg aca ggc gag gag ctt gac 816

Asp Leu Leu Asp Lys Leu Ser Val Pro Leu Thr Gly Glu Glu Leu Asp
260 265 270
atc att gaa aag ctt tat cat caa gca atg aaa ctt gaa att gat ttc 864
Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys Leu Glu Ile Asp Phe

275	280	285	
ttc tta acc cag cca ctt gtt cag aaa gct gtc atc cct ttg tca aaa			912
Phe Leu Thr Gln Pro Leu Val Gln Lys Ala Val Ile Pro Leu Ser Lys			
290	295	300	
gat cac aac cct gct gaa cac cgg ctt aca ata ttt tct gat ttc gat			960
Asp His Asn Pro Ala Glu His Arg Leu Thr Ile Phe Ser Asp Phe Asp			
305	310	315	320
ttg acg tgc act gtt gtt gat tct tct gcc atc ttg gct gaa att gca			1008
Leu Thr Cys Thr Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala			
325	330	335	
att ata aca gca ccg aga tct gat caa aat cga cca gag aat caa att			1056
Ile Ile Thr Ala Pro Arg Ser Asp Gln Asn Arg Pro Glu Asn Gln Ile			
340	345	350	
gcg cgg atg ttg tcg gct gat ttg agg aat aca tgg gga gat ctc tct			1104
Ala Arg Met Leu Ser Ala Asp Leu Arg Asn Thr Trp Gly Asp Leu Ser			
355	360	365	
aag cag tac act gaa gag tat gag caa tgt ata gag aag atg tta ctt			1152
Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Lys Met Leu Leu			
370	375	380	
act gaa aaa gcg gaa aaa ttt gat tat gaa aga ctg cat aaa aca ctt			1200
Thr Glu Lys Ala Glu Lys Phe Asp Tyr Glu Arg Leu His Lys Thr Leu			
385	390	395	400
gag gaa ctt tct gat ttt gag aaa aga gca aat act agg gtg act gaa			1248
Glu Glu Leu Ser Asp Phe Glu Lys Arg Ala Asn Thr Arg Val Thr Glu			
405	410	415	
tct ggg gta ctg aaa ggt tta aac ctt gaa gac ata aaa cga gct ggg			1296
Ser Gly Val Leu Lys Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Gly			
420	425	430	
cag cga ttg att ctc cag gat ggt tgc acc aac ttc ttc cag agc ata			1344
Gln Arg Leu Ile Leu Gln Asp Gly Cys Thr Asn Phe Phe Gln Ser Ile			

435	440	445	
ata aga aat gaa aat ctg aac gca gac att cat gtc ctc tcc tat tgc			1392
Ile Arg Asn Glu Asn Leu Asn Ala Asp Ile His Val Leu Ser Tyr Cys			
450	455	460	
tgg tgt ggc gac ctt att agg tct tcc ttt tca tca ggg ggt ata gac			1440
Trp Cys Gly Asp Leu Ile Arg Ser Ser Phe Ser Ser Gly Gly Ile Asp			
465	470	475	480
gct ctg aat gtg cat gcc aat gag ttt atg ttt caa gaa tct cta tcc			1488
Ala Leu Asn Val His Ala Asn Glu Phe Met Phe Gln Glu Ser Leu Ser			
485	490	495	
act ggt gaa att gtt aag aaa gtt gaa tcc ccc att gac aag gtt caa			1536
Thr Gly Glu Ile Val Lys Lys Val Glu Ser Pro Ile Asp Lys Val Gln			
500	505	510	
gca ttc agt aaa att cga atg aac tgt ggc aat gac caa aaa aat ctg			1584
Ala Phe Ser Lys Ile Arg Met Asn Cys Gly Asn Asp Gln Lys Asn Leu			
515	520	525	
act ctt tat att ggg gat tca gtc ggc gac tta ctt tgc ttg ctt gaa			1632
Thr Leu Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu			
530	535	540	
gca gat gtt ggc ata gtg ctt ggt acg agc tca agt cta agg acg gtg			1680
Ala Asp Val Gly Ile Val Leu Gly Thr Ser Ser Ser Leu Arg Thr Val			
545	550	555	560
ggg aat cat ttt ggt gtt tct ttt gtt cct ctg ttt cca ggt gtt gtc			1728
Gly Asn His Phe Gly Val Ser Phe Val Pro Leu Phe Pro Gly Val Val			
565	570	575	
cag aaa cag aag atg tgc act ggg gta gac tcg tca agt tgt tgg aag			1776
Gln Lys Gln Lys Met Cys Thr Gly Val Asp Ser Ser Ser Cys Trp Lys			
580	585	590	
gga cta tct ggt gtt ctc tat act gcc tct agc tgg gct gag ata cat			1824

Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ala Glu Ile His
595 600 605

gct ttt gta ttg ggg tca tga 1845

Ala Phe Val Leu Gly Ser
610

<210> 135

<211> 614

<212> PRT

<213> Nicotiana tomentosiformis

<400> 135

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

1 5 10 15
Phe Ser Asn Ser Asn Ala Leu Phe Gly Leu Arg Phe Gln Leu Tyr Pro
20 25 30

Arg Tyr Ser Arg Tyr Leu Arg Ser Pro Val Thr Met Ala Ser Ala Lys
35 40 45

Pro Lys Pro Ala Ala Ala Val Asn Lys Phe Pro Val Glu Glu Glu Cys
50 55 60

Val Gly Ile Ala Arg Lys Cys Trp Ile Lys Phe Lys Arg Glu Ser Thr

65 70 75 80
Phe Ala Leu Tyr Thr Pro Phe Val Val Ser Leu Ala Ser Gly Thr Leu
85 90 95

Asn Leu Asp Thr Phe Arg His Tyr Ile Ala Gln Asp Val His Phe Leu
100 105 110

Lys Ser Phe Ala Gln Ala Tyr Glu Ala Ala Glu Glu Cys Thr Asp Asp
115 120 125

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

130 135 140
Glu Leu Lys Met His Asp Ala Val Leu Lys Glu Trp Gly Ile Asp Leu
145 150 155 160

Val Lys Glu Ser Ser Leu Asn Pro Ala Thr Ala Lys Tyr Thr Asp Phe
165 170 175

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

420 425 430
 Gln Arg Leu Ile Leu Gln Asp Gly Cys Thr Asn Phe Phe Gln Ser Ile
 435 440 445
 Ile Arg Asn Glu Asn Leu Asn Ala Asp Ile His Val Leu Ser Tyr Cys

 450 455 460
 Trp Cys Gly Asp Leu Ile Arg Ser Ser Phe Ser Ser Gly Gly Ile Asp
 465 470 475 480
 Ala Leu Asn Val His Ala Asn Glu Phe Met Phe Gln Glu Ser Leu Ser
 485 490 495
 Thr Gly Glu Ile Val Lys Lys Val Glu Ser Pro Ile Asp Lys Val Gln
 500 505 510
 Ala Phe Ser Lys Ile Arg Met Asn Cys Gly Asn Asp Gln Lys Asn Leu

 515 520 525
 Thr Leu Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu
 530 535 540
 Ala Asp Val Gly Ile Val Leu Gly Thr Ser Ser Ser Leu Arg Thr Val
 545 550 555 560
 Gly Asn His Phe Gly Val Ser Phe Val Pro Leu Phe Pro Gly Val Val
 565 570 575
 Gln Lys Gln Lys Met Cys Thr Gly Val Asp Ser Ser Ser Cys Trp Lys

 580 585 590
 Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ala Glu Ile His
 595 600 605
 Ala Phe Val Leu Gly Ser
 610
 <210> 136
 <211> 1779
 <212> DNA
 <213> Populus trichocarpa
 <220><221> CDS
 <222> (1)..(1779)
 <223> Populus trichocarpa gene encoding TMP phosphatase

[XP_002325785.2]

<400> 136

atg cgc cta ctc ttg ttt act tct cca aac cca atc aaa acc tct tca 48

Met Arg Leu Leu Leu Phe Thr Ser Pro Asn Pro Ile Lys Thr Ser Ser

1 5 10 15

tca cta tat ttc ctc aac tcg ctc cga tcc aac tta acc aaa cgc acc 96

Ser Leu Tyr Phe Leu Asn Ser Leu Arg Ser Asn Leu Thr Lys Arg Thr

20 25 30

ttg cca act cgg aga tct ttc atc cct gca aga atg gca atc cct cca 144

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

35 40 45

cga tca ata gca tca gcg cca tct tgc act aca aca tca ggc aga agt 192

Arg Ser Ile Ala Ser Ala Pro Ser Cys Thr Thr Thr Ser Gly Arg Ser

50 55 60

aac atc aac att gaa gag ggt ctt gct agt aaa ttc tgg atc aag ttt 240

Asn Ile Asn Ile Glu Glu Gly Leu Ala Ser Lys Phe Trp Ile Lys Phe

65 70 75 80

aga aga gaa tcc gtt ttt gct atg tac act cct ttt gtc atc tct ttg 288

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

85 90 95

gct tct ggc act ctc aag att gat tct ttc agg cat tat atc tct caa 336

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

100 105 110

gat tct cac ttt ctc aaa tct ttt gct cat gcg ttt gaa tta gcg gaa 384

Asp Ser His Phe Leu Lys Ser Phe Ala His Ala Phe Glu Leu Ala Glu

115 120 125

gag tgt gct gat gat gat gaa gca aag cta gca atc tcc gag ttg agg 432

Glu Cys Ala Asp Asp Asp Glu Ala Lys Leu Ala Ile Ser Glu Leu Arg

130 135 140

aag ggt gtc tta gag gag ctg aag atg cac aat tca ttt gta cag gaa 480

Lys Gly Val Leu Glu Glu Leu Lys Met His Asn Ser Phe Val Gln Glu

145 150 155 160

tgg ggt ata gac cca ggt aaa gag ggg act atc aat tct gct act gta 528

Trp Gly Ile Asp Pro Gly Lys Glu Gly Thr Ile Asn Ser Ala Thr Val

 165 170 175

aaa tac aca gat ttc ttg ttg gct aca gct tct ggg aag gtt gaa gga 576

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

 180 185 190

gtg aaa ggt ctt ggt aaa ctt gca act cct ttt gaa aga aca aaa gtt 624

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

 195 200 205

gca gcc tat act ctg ggt gcc atg aca cct tgc atg cgg ctg tat tcc 672

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

 210 215 220

ttt cta ggc aag gaa ctc cag gca gtt tta gat ccg gag gaa gat ggg 720

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

225 230 235 240

cac cct tac aag aag tgg att gac agt tat tcg tct gag agt ttt cag 768

His Pro Tyr Lys Lys Trp Ile Asp Ser Tyr Ser Ser Glu Ser Phe Gln

 245 250 255

gca tca gct ctg caa act gaa gac ttg ctg gat aaa ctt agt gtc tcc 816

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

 260 265 270

ttg aca ggc gag gag ctt gac atc att gaa aag ctt tat cac cag gcc 864

Leu Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr His Gln Ala

 275 280 285

atg aaa ctt gaa ata gaa ttc ttc ctt gct cag cca att gct cag aca 912

Met Lys Leu Glu Ile Glu Phe Phe Leu Ala Gln Pro Ile Ala Gln Thr

 290 295 300

act tta gct ccc ctg aca aaa ggg cat aac cct gaa gaa gac cgg ctt	960
Thr Leu Ala Pro Leu Thr Lys Gly His Asn Pro Glu Glu Asp Arg Leu	
305 310 315 320	
gtc ata ttt tct gat ttt gat ttg aca tgc act gtt gtt gac tct tct	1008
Val Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser	
325 330 335	
gcc att ttg gca gaa att gca ata cta aca gca cca aaa tct gat gtg	1056
Ala Ile Leu Ala Glu Ile Ala Ile Leu Thr Ala Pro Lys Ser Asp Val	
340 345 350	
gtt caa cct gag act caa att gct cga atg tca tca gct gat ctg agg	1104
Val Gln Pro Glu Thr Gln Ile Ala Arg Met Ser Ser Ala Asp Leu Arg	
355 360 365	
aac aca tgg ggt ctt ctt tct gga cag tac acg gaa gag tat gaa caa	1152
Asn Thr Trp Gly Leu Leu Ser Gly Gln Tyr Thr Glu Glu Tyr Glu Gln	
370 375 380	
tgt att gaa agc att atg cca tct gca aaa gtg gaa ttc aac tat gaa	1200
Cys Ile Glu Ser Ile Met Pro Ser Ala Lys Val Glu Phe Asn Tyr Glu	
385 390 395 400	
gct ctt tgt aaa gca ctt gaa caa ctt tca gac ttt gag cga agg gca	1248
Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp Phe Glu Arg Arg Ala	
405 410 415	
aat tct aga gtg att gat tct gga gtt ctc aaa ggt ttg aat ctt gaa	1296
Asn Ser Arg Val Ile Asp Ser Gly Val Leu Lys Gly Leu Asn Leu Glu	
420 425 430	
gat gta aaa cga gcg ggt gaa cgt ttg att ctt cag gat ggt tgc att	1344
Asp Val Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile	
435 440 445	
ggg ttc ttt cag aaa att gtg aag aat gaa aat ttg aac act aat gtc	1392
Gly Phe Phe Gln Lys Ile Val Lys Asn Glu Asn Leu Asn Thr Asn Val	
450 455 460	
cat gtg ctc tca tac tgc tgg tgt ggt gat ctc atc aga tca gct ttc	1440

His Val Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Ala Phe
 465 470 475 480
 tcc tca ggg ggt ttg gat gct cta aat att cat gca aat gag tta att 1488
 Ser Ser Gly Gly Leu Asp Ala Leu Asn Ile His Ala Asn Glu Leu Ile
 485 490 495
 ttt gaa gaa tca atc tcc acg gga gag att aac ttg act gtt tac att 1536
 Phe Glu Glu Ser Ile Ser Thr Gly Glu Ile Asn Leu Thr Val Tyr Ile
 500 505 510
 gga gat tca gtt ggt gac ttg ctt tgt cta ctt cag gca gat att ggt 1584
 Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Gln Ala Asp Ile Gly
 515 520 525
 att gta gtt gga tct agt gca agc tta agg agc gtg gga agt caa tat 1632
 Ile Val Val Gly Ser Ser Ala Ser Leu Arg Ser Val Gly Ser Gln Tyr
 530 535 540
 ggt gtt tct ttt gta cca ctg ttc cct ggc ttg gta aga aaa cag aaa 1680
 Gly Val Ser Phe Val Pro Leu Phe Pro Gly Leu Val Arg Lys Gln Lys
 545 550 555 560
 gaa tct gat gga gaa tct cct aat tgg aaa ggg cta tct ggc ata cta 1728
 Glu Ser Asp Gly Glu Ser Pro Asn Trp Lys Gly Leu Ser Gly Ile Leu
 565 570 575
 tat aca gtc tcc agt tgg tca gaa ata cat gcc ttc att ttg ggg tgg 1776
 Tyr Thr Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Trp
 580 585 590
 tag 1779
 <210> 137
 <211> 592
 <212> PRT
 <213> Populus trichocarpa
 <400> 137
 Met Arg Leu Leu Leu Phe Thr Ser Pro Asn Pro Ile Lys Thr Ser Ser
 1 5 10 15

Ser Leu Tyr Phe Leu Asn Ser Leu Arg Ser Asn Leu Thr Lys Arg Thr
 20 25 30
 Leu Pro Thr Arg Arg Ser Phe Ile Pro Ala Arg Met Ala Ile Pro Pro
 35 40 45
 Arg Ser Ile Ala Ser Ala Pro Ser Cys Thr Thr Thr Ser Gly Arg Ser
 50 55 60

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

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

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

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

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

Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Gln Ala Asp Ile Gly
515 520 525
Ile Val Val Gly Ser Ser Ala Ser Leu Arg Ser Val Gly Ser Gln Tyr
530 535 540
Gly Val Ser Phe Val Pro Leu Phe Pro Gly Leu Val Arg Lys Gln Lys
545 550 555 560
Glu Ser Asp Gly Glu Ser Pro Asn Trp Lys Gly Leu Ser Gly Ile Leu
565 570 575

Tyr Thr Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Trp
580 585 590

<210> 138

<211> 1860

<212> DNA

<213> *Jatropha curcas*

<220><221> CDS

<222> (1)..(1860)

<223> *Jatropha curcas* gene encoding TMP phosphatase [KDP23738.1]

<400> 138

atg gcg atc cct cca aag cta gct tcc tca tcg tct tcc atg gcc gcc 48
Met Ala Ile Pro Pro Lys Leu Ala Ser Ser Ser Ser Met Ala Ala
1 5 10 15

tcc cct act tct gct ggt gga acc aac gag gaa ggc ctc gct agt aaa 96
Ser Pro Thr Ser Ala Gly Gly Thr Asn Glu Glu Gly Leu Ala Ser Lys
20 25 30

ttc tgg atc aag ttt cgc cga gaa tcg gtt ctc gct atg tac act cct 144
Phe Trp Ile Lys Phe Arg Arg Glu Ser Val Leu Ala Met Tyr Thr Pro
35 40 45

ttc gtc gtc tct ttt gcc gcc ggc aac ctc aag att gag agt ttt agg 192

Phe Val Val Ser Phe Ala Ala Gly Asn Leu Lys Ile Glu Ser Phe Arg
50 55 60

cat tac atc gct cag gat ttt cac ttc ctc aaa gcc ttc gct cac gcg 240
His Tyr Ile Ala Gln Asp Phe His Phe Leu Lys Ala Phe Ala His Ala

65	70	75	80	
tat gaa ttg gca gaa gag tgt gct gat gat gat gcc aag cta gct				288
Tyr Glu Leu Ala Glu Glu Cys Ala Asp Asp Asp Asp Ala Lys Leu Ala				
	85	90	95	
att gcc gcg ttg agg aag ggg gtc tta gag gag ctg aag ttg cat aaa				336
Ile Ala Ala Leu Arg Lys Gly Val Leu Glu Glu Leu Lys Leu His Lys				
	100	105	110	
tca ttt gta cag gaa tgg ggt atg gac cct tcc aaa gag gtg act atc				384
Ser Phe Val Gln Glu Trp Gly Met Asp Pro Ser Lys Glu Val Thr Ile				
	115	120	125	
aat tct gca act gca aaa tac aca gat ttc ttg ttg gct aca gct tct				432
Asn Ser Ala Thr Ala Lys Tyr Thr Asp Phe Leu Leu Ala Thr Ala Ser				
	130	135	140	
gga aag gtt gaa gga gtg aaa ggt cct ggt aaa ctt gca act cct ttt				480
Gly Lys Val Glu Gly Val Lys Gly Pro Gly Lys Leu Ala Thr Pro Phe				
	145	150	155	160
gaa aga aca aaa gtt gca gct tac act ctt ggt acc atg aca ccc tgt				528
Glu Arg Thr Lys Val Ala Ala Tyr Thr Leu Gly Thr Met Thr Pro Cys				
	165	170	175	
atg agg ttg tat gcc ttt cta gct aag gag ctg caa gca cta ata gat				576
Met Arg Leu Tyr Ala Phe Leu Ala Lys Glu Leu Gln Ala Leu Ile Asp				
	180	185	190	
gca gaa gct ggt att cat cct tac cag aag tgg att gac aat tac tca				624
Ala Glu Ala Gly Ile His Pro Tyr Gln Lys Trp Ile Asp Asn Tyr Ser				
	195	200	205	
tct gag agt ttt cag gca tca gct ctg caa act gaa gac ttg ctg gat				672
Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Thr Glu Asp Leu Leu Asp				
	210	215	220	
aaa ctt agt gtc cct ttg aca ggc gaa gag ctt gac atc att gaa aag				720
Lys Leu Ser Val Pro Leu Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys				

225	230	235	240	
ctt tat cac caa gcc atg aaa ctt gaa ata gag ttc ttc aat gcg cag				768
Leu Tyr His Gln Ala Met Lys Leu Glu Ile Glu Phe Phe Asn Ala Gln				
245	250	255		
cca ctt gat cag ccc act gtg gtt cct ctg aca aaa gag cat aac cct				816
Pro Leu Asp Gln Pro Thr Val Val Pro Leu Thr Lys Glu His Asn Pro				
260	265	270		
cta gaa gat cgc ctc gtg ata ttt tct gat ttt gat ttg aca tgc aca				864
Leu Glu Asp Arg Leu Val Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr				
275	280	285		
gtt gtt gat tcc tct gcc att ttg gca gag att gca att tta aca gca				912
Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala Ile Leu Thr Ala				
290	295	300		
tca aaa tct gat cag tca caa tct gat aat caa aat gct agg atg tca				960
Ser Lys Ser Asp Gln Ser Gln Ser Asp Asn Gln Asn Ala Arg Met Ser				
305	310	315	320	
tca act gag cta agg aac aca tgg gtt ctt ctc tct gga cag tat act				1008
Ser Thr Glu Leu Arg Asn Thr Trp Val Leu Leu Ser Gly Gln Tyr Thr				
325	330	335		
gaa gaa tat gag caa tgc att gaa agc att ctg ccc tct gaa aaa atg				1056
Glu Glu Tyr Glu Gln Cys Ile Glu Ser Ile Leu Pro Ser Glu Lys Met				
340	345	350		
gag ttc aac ttt gaa gct ttg tgt aaa gca ctc gaa caa ctc tca gac				1104
Glu Phe Asn Phe Glu Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp				
355	360	365		
ttt gag cga agg gca aat gct aga gtt atc aaa tct gga gtt ctt aag				1152
Phe Glu Arg Arg Ala Asn Ala Arg Val Ile Lys Ser Gly Val Leu Lys				
370	375	380		
ggt ttg aat ctt gaa gac ata aaa cga gct gtg gag ttc aac ttt gaa				1200

Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Val Glu Phe Asn Phe Glu
385 390 395 400
gct ttg tgt aaa gca ctc gaa caa ctc tca gac ttt gag cga agg gca 1248
Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp Phe Glu Arg Arg Ala
405 410 415
aat gct aga gtt atc aaa tct gga gtt ctt aag ggt ttg aat ctt gaa 1296
Asn Ala Arg Val Ile Lys Ser Gly Val Leu Lys Gly Leu Asn Leu Glu
420 425 430
gac ata aaa cga gct ggt gaa aga ctg att ctt caa gat ggc tgc acc 1344
Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Thr
435 440 445
agt ttt ttt cag aaa atc tcg aag aat gaa aat ctg aat gct aat ata 1392
Ser Phe Phe Gln Lys Ile Ser Lys Asn Glu Asn Leu Asn Ala Asn Ile
450 455 460
cat ttc ctc tca tat tgt tgg tgt gct gat ctg atc aga tct gct ttc 1440
His Phe Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe
465 470 475 480
tca tca ggg ggt ttg gat gtt ctg aat ata cat gcg aat gag ttt gat 1488
Ser Ser Gly Gly Leu Asp Val Leu Asn Ile His Ala Asn Glu Phe Asp
485 490 495
ttc gta gaa tca att tca acg ggt gag att att atg aag gtg gaa acc 1536
Phe Val Glu Ser Ile Ser Thr Gly Glu Ile Ile Met Lys Val Glu Thr
500 505 510
cct aca gac aaa gcc caa gct ttt aat aat att tta atg aac tac agc 1584
Pro Thr Asp Lys Ala Gln Ala Phe Asn Asn Ile Leu Met Asn Tyr Ser
515 520 525
cct gac aaa aag aat ttg act gtt tat att gga gac tca gtt ggg gac 1632
Pro Asp Lys Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp
530 535 540
ttg ctt tgt ctg ctt gcg gca gat ata ggc atc gtg atc gga tca agc 1680
Leu Leu Cys Leu Leu Ala Ala Asp Ile Gly Ile Val Ile Gly Ser Ser

545 550 555 560
tcc agc cta agg aga gtc gga agt cag ttt ggt gta aca ttt tta cca 1728
Ser Ser Leu Arg Arg Val Gly Ser Gln Phe Gly Val Thr Phe Leu Pro
 565 570 575

ttg tat cct ggc ttg gtt aaa aaa cag aga gag tat act gaa gga agc 1776
Leu Tyr Pro Gly Leu Val Lys Lys Gln Arg Glu Tyr Thr Glu Gly Ser
 580 585 590
tct tgg aat tgg aag ggt caa tct ggc gtt ctg tac aca gtt tct agt 1824
Ser Trp Asn Trp Lys Gly Gln Ser Gly Val Leu Tyr Thr Val Ser Ser
 595 600 605
tgg gct gaa ata cat tcc ttc gtt ttg gga tgg tag 1860

Trp Ala Glu Ile His Ser Phe Val Leu Gly Trp

610

615

<210> 139

<211> 619

<212> PRT

<213> Jatropha curcas

<400> 139

Met Ala Ile Pro Pro Lys Leu Ala Ser Ser Ser Ser Ser Met Ala Ala

1 5 10 15

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

20 25 30

Phe Trp Ile Lys Phe Arg Arg Glu Ser Val Leu Ala Met Tyr Thr Pro

35

40

45

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

50

55

60

His Tyr Ile Ala Gln Asp Phe His Phe Leu Lys Ala Phe Ala His Ala

65 70 75 80

Tyr Glu Leu Ala Glu Glu Cys Ala Asp Asp Asp Ala Lys Leu Ala

85

90

95

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

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

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

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

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

Glu Phe Asn Phe Glu Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp

355 360 365

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

370 375 380

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

385 390 395 400

Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp Phe Glu Arg Arg Ala

405 410 415

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

420 425 430

Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Thr

435 440 445

Ser Phe Phe Gln Lys Ile Ser Lys Asn Glu Asn Leu Asn Ala Asn Ile

450 455 460

His Phe Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe

465 470 475 480

Ser Ser Gly Gly Leu Asp Val Leu Asn Ile His Ala Asn Glu Phe Asp

485 490 495

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

500 505 510

Pro Thr Asp Lys Ala Gln Ala Phe Asn Asn Ile Leu Met Asn Tyr Ser

515 520 525

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

530 535 540

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

545 550 555 560

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

565 570 575

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

580 585 590

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

595 600 605

Trp Ala Glu Ile His Ser Phe Val Leu Gly Trp

610 615

<210> 140

<211> 1842

<212> DNA

<213> Citrus sinensis

<220><221> CDS

<222> (1)..(1842)

<223> Citrus sinensi gene encoding TMP phosphatase [XP_006484613.1]

<400> 140

atg cgc ttc ctt ttc aca aac cca atc aaa acc cca tta ctc tct tct 48

Met Arg Phe Leu Phe Thr Asn Pro Ile Lys Thr Pro Leu Leu Ser Ser

1 5 10 15

att ctt ttc cat tgt ccc aac tcg ccc cga ctc ggc ctt ctt gac tca 96

Ile Leu Phe His Cys Pro Asn Ser Pro Arg Leu Gly Leu Leu Asp Ser

20 25 30

gtc cga gtc aac tca cct tct tct ttg aca act caa aga tcg tca ctt 144

Val Arg Val Asn Ser Pro Ser Ser Leu Thr Thr Gln Arg Ser Ser Leu

35 40 45

tcg atg gcg gcg att ccc cca aaa tcg ccg agc cct gag gag gag gga 192

Ser Met Ala Ala Ile Pro Pro Lys Ser Pro Ser Pro Glu Glu Glu Gly

50 55 60

ctc gcg agg agg ttg tgg atc aag ttt aag aga gaa tct gtg ttt gcc 240

Leu Ala Arg Arg Leu Trp Ile Lys Phe Lys Arg Glu Ser Val Phe Ala

65 70 75 80

atg tac tcc ccg ttt acg gtt tgt ttg gct tct ggg aac cta aag ctt 288

Met Tyr Ser Pro Phe Thr Val Cys Leu Ala Ser Gly Asn Leu Lys Leu

85 90 95

gaa acc ttc agg cat tac atc gcc caa gat ttt cat ttt ctc aaa gct 336

Glu Thr Phe Arg His Tyr Ile Ala Gln Asp Phe His Phe Leu Lys Ala

100	105	110	
ttc gcc caa gcg tat gaa ctg gcg gaa gaa tgt gct gat gat gat gat			384
Phe Ala Gln Ala Tyr Glu Leu Ala Glu Glu Cys Ala Asp Asp Asp Asp			
115	120	125	
gca aag tta tct atc tct gaa ttg agg aag ggt gta ctt gag gag tta			432
Ala Lys Leu Ser Ile Ser Glu Leu Arg Lys Gly Val Leu Glu Glu Leu			
130	135	140	
aaa atg cat gat tcc ttt gtg aag gag tgg ggt aca gat ctt gct aaa			480
Lys Met His Asp Ser Phe Val Lys Glu Trp Gly Thr Asp Leu Ala Lys			
145	150	155	160
atg gct act gtt aac tct gca act gta aag tat aca gag ttc ttg ttg			528
Met Ala Thr Val Asn Ser Ala Thr Val Lys Tyr Thr Glu Phe Leu Leu			
165	170	175	
gca aca gct tcc ggg aag gtc gaa ggt gtt aaa ggt cct gga aaa ctt			576
Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys Gly Pro Gly Lys Leu			
180	185	190	
gca acc cca ttt gag aaa act aaa gtt gcc gct tac aca ttg ggt gcc			624
Ala Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr Thr Leu Gly Ala			
195	200	205	
atg tca cct tgt atg agg ctc tat gct ttc ctt gga aag gaa ttc cat			672
Met Ser Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Phe His			
210	215	220	
ggc ctc cta aat gct aat gaa ggc aat cat cct tac aag aag tgg att			720
Gly Leu Leu Asn Ala Asn Glu Gly Asn His Pro Tyr Lys Lys Trp Ile			
225	230	235	240
gac aat tat tct tct gaa agt ttt cag gcc tca gct ctg caa aat gag			768
Asp Asn Tyr Ser Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Asn Glu			
245	250	255	
gac ttg ctg gat aaa ctt agt gtc tct ttg aca ggc gaa gaa cta gac			816
Asp Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu Leu Asp			
260	265	270	

ata ata gaa aag ctc tat cac caa gcc atg aaa ctt gaa gta gag ttc	864
Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys Leu Glu Val Glu Phe	
275 280 285	
ttc tgt gct cag cca ctt gct cag ccc act gta gtt cct ctg att aaa	912
Phe Cys Ala Gln Pro Leu Ala Gln Pro Thr Val Val Pro Leu Ile Lys	
290 295 300	
ggg cat aat cct gca gga gac cgt cta att ata ttt tct gat ttc gat	960
Gly His Asn Pro Ala Gly Asp Arg Leu Ile Ile Phe Ser Asp Phe Asp	
305 310 315 320	
ttg act tgc acc att gtt gat tcc tct gcc att ttg gca gag atc gca	1008
Leu Thr Cys Thr Ile Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala	
325 330 335	
ata gtg aca gca cca aaa tct gac cag aat caa cct gaa aat caa ctt	1056
Ile Val Thr Ala Pro Lys Ser Asp Gln Asn Gln Pro Glu Asn Gln Leu	
340 345 350	
ggg cgg atg tca tca ggt gag ctg agg aac aca tgg ggt ctt ctt tcc	1104
Gly Arg Met Ser Ser Gly Glu Leu Arg Asn Thr Trp Gly Leu Leu Ser	
355 360 365	
aaa cag tac aca gag gag tac gaa caa tgc att gaa agc ttc atg ccc	1152
Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Ser Phe Met Pro	
370 375 380	
tct gag aaa gtg gag aat ttc aac tat gaa act ttg cat aaa gca ctt	1200
Ser Glu Lys Val Glu Asn Phe Asn Tyr Glu Thr Leu His Lys Ala Leu	
385 390 395 400	
gag caa ctc tca cac ttt gag aag agg gca aat tct aga gtg atc gaa	1248
Glu Gln Leu Ser His Phe Glu Lys Arg Ala Asn Ser Arg Val Ile Glu	
405 410 415	
tct gga gtt ctc aag ggt ata aat ctt gaa gat att aaa aaa gct ggt	1296
Ser Gly Val Leu Lys Gly Ile Asn Leu Glu Asp Ile Lys Lys Ala Gly	
420 425 430	

gaa cgc ctg agt ctt caa gat ggt tgc act acc ttc ttt cag aaa gtt	1344
Glu Arg Leu Ser Leu Gln Asp Gly Cys Thr Thr Phe Phe Gln Lys Val	
435 440 445	
gta aag aat gaa aat ttg aat gct aat gtc cat gtg ctt tca tac tgt	1392
Val Lys Asn Glu Asn Leu Asn Ala Asn Val His Val Leu Ser Tyr Cys	
450 455 460	
tgg tgt ggt gat ctc atc aga gca tct ttt tct tca gca ggt tta aat	1440
Trp Cys Gly Asp Leu Ile Arg Ala Ser Phe Ser Ser Ala Gly Leu Asn	
465 470 475 480	
gca ctg aat gta cat gcg aat gag ttc tca ttc aaa gaa tct att tca	1488
Ala Leu Asn Val His Ala Asn Glu Phe Ser Phe Lys Glu Ser Ile Ser	
485 490 495	
acg ggt gaa att att gag aaa gtg gag tcc ccc att gac aaa gtt caa	1536
Thr Gly Glu Ile Ile Glu Lys Val Glu Ser Pro Ile Asp Lys Val Gln	
500 505 510	
gct ttc aac aat act tta gag aaa tac gga act gac aga aag aac ttg	
Ala Phe Asn Asn Thr Leu Glu Lys Tyr Gly Thr Asp Arg Lys Asn Leu	1584
515 520 525	
agt gtt tac att gga gac tct gtg ggt gac ttg ctt tgt ctg ctt gag	1632
Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu	
530 535 540	
gct gat ata ggc att gta atc ggg tct agc tca agc tta agg aga gtg	1680
Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser Leu Arg Arg Val	
545 550 555 560	
gga tct caa ttt ggt gtt aca ttt atc ccg ttg tac cct ggc ttg gtt	1728
Gly Ser Gln Phe Gly Val Thr Phe Ile Pro Leu Tyr Pro Gly Leu Val	
565 570 575	
aag aaa cag aag gag tac act gaa gga agc tct tct aac tgg aag gag	1776
Lys Lys Gln Lys Glu Tyr Thr Glu Gly Ser Ser Ser Asn Trp Lys Glu	

580 585 590
 aaa tct ggc ata ctt tac aca gtc tca agt tgg gct gaa gta cat gcc 1824
 Lys Ser Gly Ile Leu Tyr Thr Val Ser Ser Trp Ala Glu Val His Ala

595 600 605
 ttt atc ttg ggg tgg tag 1842
 Phe Ile Leu Gly Trp
 610

<210> 141

<211> 613

<212> PRT

<213> Citrus sinensis

<400> 141

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

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

115 120 125
 Ala Lys Leu Ser Ile Ser Glu Leu Arg Lys Gly Val Leu Glu Glu Leu
 130 135 140
 Lys Met His Asp Ser Phe Val Lys Glu Trp Gly Thr Asp Leu Ala Lys
 145 150 155 160

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

405 410 415
Ser Gly Val Leu Lys Gly Ile Asn Leu Glu Asp Ile Lys Lys Ala Gly

420 425 430
Glu Arg Leu Ser Leu Gln Asp Gly Cys Thr Thr Phe Phe Gln Lys Val

435 440 445
Val Lys Asn Glu Asn Leu Asn Ala Asn Val His Val Leu Ser Tyr Cys

450 455 460
Trp Cys Gly Asp Leu Ile Arg Ala Ser Phe Ser Ser Ala Gly Leu Asn

465 470 475 480
Ala Leu Asn Val His Ala Asn Glu Phe Ser Phe Lys Glu Ser Ile Ser

485 490 495
Thr Gly Glu Ile Ile Glu Lys Val Glu Ser Pro Ile Asp Lys Val Gln

500 505 510
Ala Phe Asn Asn Thr Leu Glu Lys Tyr Gly Thr Asp Arg Lys Asn Leu

515 520 525
Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu

530 535 540
Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser Leu Arg Arg Val

545 550 555 560
Gly Ser Gln Phe Gly Val Thr Phe Ile Pro Leu Tyr Pro Gly Leu Val

565 570 575
Lys Lys Gln Lys Glu Tyr Thr Glu Gly Ser Ser Ser Asn Trp Lys Glu

580 585 590
Lys Ser Gly Ile Leu Tyr Thr Val Ser Ser Trp Ala Glu Val His Ala

595 600 605
Phe Ile Leu Gly Trp

610

<210> 142

<211> 1728

<212> DNA

<213> Prunus persica

<220><221> CDS

<222> (1)..(1728)

<223> Prunus persica gene encoding TMP phosphatase [XP_007199656.1]

<400> 142

atg gcg gca ttg gct cgt cat agc att gtt aga ctc aat cac gaa gga 48

Met Ala Ala Leu Ala Arg His Ser Ile Val Arg Leu Asn His Glu Gly

1 5 10 15

ggc cta gcc aga cgg ctg tgg ttc aag ttc aga gac gac tct gtt ttc 96

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

20 25 30

tct ctc tac act ccc ttc ttc gtt ggc tta gcc tct gct act ctg cac 144

Ser Leu Tyr Thr Pro Phe Phe Val Gly Leu Ala Ser Ala Thr Leu His

35 40 45

tct gaa act acc ttt cgc cat ttc atc tct cag gac ctc cat ttt ctc 192

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

50 55 60

aaa gcc ttc gtt ctc gca tat gaa ttg gcg gaa gat tgt gct gat gat 240

Lys Ala Phe Val Leu Ala Tyr Glu Leu Ala Glu Asp Cys Ala Asp Asp

65 70 75 80

gag gac gac aag aat ggt tta cgc gat ttg aga aaa cgt gcc gtc ggc 288

Glu Asp Asp Lys Asn Gly Leu Arg Asp Leu Arg Lys Arg Ala Val Gly

85 90 95

agg ctt caa atg cac gac aca ttt gtc cga gaa tgg ggt ttt gaa ttc 336

Arg Leu Gln Met His Asp Thr Phe Val Arg Glu Trp Gly Phe Glu Phe

100 105 110

cca aat gag gac att tct aaa gac att gca aca acc aaa tac aca gat 384

Pro Asn Glu Asp Ile Ser Lys Asp Ile Ala Thr Thr Lys Tyr Thr Asp

115 120 125

ttc ttg ctt gca aca gca tca ggg aaa att gaa gga gaa aga tcg gtt 432

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

130 135 140

ctg gac aaa atc gca acc cct ttc gaa aag acc aag gtt gct gca tat 480

Leu Asp Lys Ile Ala Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr
145 150 155 160

aca ctt gct gct ctg gct cct tgt atg aga ctc tat gcc ttc atc agt 528

Thr Leu Ala Ala Leu Ala Pro Cys Met Arg Leu Tyr Ala Phe Ile Ser
165 170 175

act gag atc caa ggc att ata aat cct gat caa gat agc act cac att 576

Thr Glu Ile Gln Gly Ile Ile Asn Pro Asp Gln Asp Ser Thr His Ile

180 185 190

tac aaa agc tgg ata gaa aat tat tcg tct caa gtt ttc gag gaa ata 624

Tyr Lys Ser Trp Ile Glu Asn Tyr Ser Ser Gln Val Phe Glu Glu Ile
195 200 205

gcc ctg caa aat gaa gac atg cta gat aaa ctt agt gtt tct ttg act 672

Ala Leu Gln Asn Glu Asp Met Leu Asp Lys Leu Ser Val Ser Leu Thr
210 215 220

ggt gag gag ctt gag att ata gag aag ctc tat cat caa gct atg aag 720

Gly Glu Glu Leu Glu Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys
225 230 235 240

ctt caa gta gat ttt att gct gct caa cca att tct gat cag caa tct 768

Leu Gln Val Asp Phe Ile Ala Ala Gln Pro Ile Ser Asp Gln Gln Ser
245 250 255

gta gtt cct ttg tct cgg gtg cat gac ttt agc aaa cgc cat ctt acg 816

Val Val Pro Leu Ser Arg Val His Asp Phe Ser Lys Arg His Leu Thr
260 265 270

ata ctt tgt gac ttt gat ttg gca tgc act gct ttt gat tct gct gcc 864

Ile Leu Cys Asp Phe Asp Leu Ala Cys Thr Ala Phe Asp Ser Ala Ala
275 280 285

ata ttg gct gag att gcg atc ata aca gca cca aag gct gat atg gat 912

Ile Leu Ala Glu Ile Ala Ile Ile Thr Ala Pro Lys Ala Asp Met Asp

290	295	300	
gga tct gat caa acc caa ctt gct cgg atg cca tca gca gac tta agg			960
Gly Ser Asp Gln Thr Gln Leu Ala Arg Met Pro Ser Ala Asp Leu Arg			
305	310	315	320
agc aca tgg gat gtt ctt tca acc caa tac act gaa caa ttt gaa caa			1008
Ser Thr Trp Asp Val Leu Ser Thr Gln Tyr Thr Glu Gln Phe Glu Gln			
	325	330	335
tgt gta gaa agc att gtg gcc agt gag aga gtg gaa gaa ttc gat tat			1056
Cys Val Glu Ser Ile Val Ala Ser Glu Arg Val Glu Glu Phe Asp Tyr			
	340	345	350
gaa cgt ctg tgt agc gcg ctt gaa caa ctt gcg gag ttt gag aga aag			1104
Glu Arg Leu Cys Ser Ala Leu Glu Gln Leu Ala Glu Phe Glu Arg Lys			
	355	360	365
gca aat gaa agg gtg gtt cag tca gga gtg ttg aag ggt tta aat gcg			1152
Ala Asn Glu Arg Val Val Gln Ser Gly Val Leu Lys Gly Leu Asn Ala			
	370	375	380
gag gat ata aaa agg gct gga cag agc ctc att ctg caa gat ggt tgc			1200
Glu Asp Ile Lys Arg Ala Gly Gln Ser Leu Ile Leu Gln Asp Gly Cys			
385	390	395	400
aga agc ttc ttt cag aag att gtg aaa aat aaa aat ctg aaa act gat			1248
Arg Ser Phe Phe Gln Lys Ile Val Lys Asn Lys Asn Leu Lys Thr Asp			
	405	410	415
gtt cat gtg ctt tca tac tgc tgg tgt aat gac ctc att gta tca gct			1296
Val His Val Leu Ser Tyr Cys Trp Cys Asn Asp Leu Ile Val Ser Ala			
	420	425	430
ttc tct tca gga gat ttg aat gtc ttg aat gta cat tca aat gag ttg			1344
Phe Ser Ser Gly Asp Leu Asn Val Leu Asn Val His Ser Asn Glu Leu			
	435	440	445
gtt tat caa gaa tct gtc aca act ggt gaa att gta aag aag atg gag			1392
Val Tyr Gln Glu Ser Val Thr Thr Gly Glu Ile Val Lys Lys Met Glu			
450	455	460	

tct ccc atg gaa aag ctt caa gtc ttc aac gac gtc cta atc gac cgc 1440
 Ser Pro Met Glu Lys Leu Gln Val Phe Asn Asp Val Leu Ile Asp Arg
 465 470 475 480
 agg ggc gaa ggc aat aaa cac ttg aca gtt tac att gga ggc tca gtg 1488
 Arg Gly Glu Gly Asn Lys His Leu Thr Val Tyr Ile Gly Gly Ser Val
 485 490 495
 ggt gac ttg ctt tgc ctg ctt gaa gca gat ata ggc att gta gtt ggt 1536
 Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Gly
 500 505 510
 tca agt tca agc cta agg aga cta ggt gat cat ttt ggt gtt tcc ttt 1584
 Ser Ser Ser Ser Leu Arg Arg Leu Gly Asp His Phe Gly Val Ser Phe
 515 520 525
 gtc cca ttg ttc tct ggc ttg gtg aag agg cag aaa gaa ctt gct gat 1632
 Val Pro Leu Phe Ser Gly Leu Val Lys Arg Gln Lys Glu Leu Ala Asp
 530 535 540
 caa gat tgt gcc tct aat tgg tgg aaa cca ttg tct ggt gtt ctt tat 1680
 Gln Asp Cys Ala Ser Asn Trp Trp Lys Pro Leu Ser Gly Val Leu Tyr
 545 550 555 560
 acg gtg tct agt tgg gct gaa ata cag gca ttc att ttg ggt aca tag 1728
 Thr Val Ser Ser Trp Ala Glu Ile Gln Ala Phe Ile Leu Gly Thr
 565 570 575
 <210> 143
 <211> 575
 <212> PRT
 <213> Prunus persica
 <400> 143
 Met Ala Ala Leu Ala Arg His Ser Ile Val Arg Leu Asn His Glu Gly
 1 5 10 15
 Gly Leu Ala Arg Arg Leu Trp Phe Lys Phe Arg Asp Asp Ser Val Phe
 20 25 30
 Ser Leu Tyr Thr Pro Phe Phe Val Gly Leu Ala Ser Ala Thr Leu His

35	40	45	
Ser Glu Thr Thr Phe Arg His Phe Ile Ser Gln Asp Leu His Phe Leu			
50	55	60	
Lys Ala Phe Val Leu Ala Tyr Glu Leu Ala Glu Asp Cys Ala Asp Asp			
65	70	75	80
Glu Asp Asp Lys Asn Gly Leu Arg Asp Leu Arg Lys Arg Ala Val Gly			
85	90	95	
Arg Leu Gln Met His Asp Thr Phe Val Arg Glu Trp Gly Phe Glu Phe			
100	105	110	
Pro Asn Glu Asp Ile Ser Lys Asp Ile Ala Thr Thr Lys Tyr Thr Asp			
115	120	125	
Phe Leu Leu Ala Thr Ala Ser Gly Lys Ile Glu Gly Glu Arg Ser Val			
130	135	140	
Leu Asp Lys Ile Ala Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr			
145	150	155	160
Thr Leu Ala Ala Leu Ala Pro Cys Met Arg Leu Tyr Ala Phe Ile Ser			
165	170	175	
Thr Glu Ile Gln Gly Ile Ile Asn Pro Asp Gln Asp Ser Thr His Ile			
180	185	190	
Tyr Lys Ser Trp Ile Glu Asn Tyr Ser Ser Gln Val Phe Glu Glu Ile			
195	200	205	
Ala Leu Gln Asn Glu Asp Met Leu Asp Lys Leu Ser Val Ser Leu Thr			
210	215	220	
Gly Glu Glu Leu Glu Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys			
225	230	235	240
Leu Gln Val Asp Phe Ile Ala Ala Gln Pro Ile Ser Asp Gln Gln Ser			
245	250	255	
Val Val Pro Leu Ser Arg Val His Asp Phe Ser Lys Arg His Leu Thr			
260	265	270	
Ile Leu Cys Asp Phe Asp Leu Ala Cys Thr Ala Phe Asp Ser Ala Ala			
275	280	285	

Ile Leu Ala Glu Ile Ala Ile Ile Thr Ala Pro Lys Ala Asp Met Asp
290 295 300

Gly Ser Asp Gln Thr Gln Leu Ala Arg Met Pro Ser Ala Asp Leu Arg
305 310 315 320

Ser Thr Trp Asp Val Leu Ser Thr Gln Tyr Thr Glu Gln Phe Glu Gln
325 330 335

Cys Val Glu Ser Ile Val Ala Ser Glu Arg Val Glu Glu Phe Asp Tyr
340 345 350

Glu Arg Leu Cys Ser Ala Leu Glu Gln Leu Ala Glu Phe Glu Arg Lys
355 360 365

Ala Asn Glu Arg Val Val Gln Ser Gly Val Leu Lys Gly Leu Asn Ala
370 375 380

Glu Asp Ile Lys Arg Ala Gly Gln Ser Leu Ile Leu Gln Asp Gly Cys
385 390 395 400

Arg Ser Phe Phe Gln Lys Ile Val Lys Asn Lys Asn Leu Lys Thr Asp
405 410 415

Val His Val Leu Ser Tyr Cys Trp Cys Asn Asp Leu Ile Val Ser Ala
420 425 430

Phe Ser Ser Gly Asp Leu Asn Val Leu Asn Val His Ser Asn Glu Leu
435 440 445

Val Tyr Gln Glu Ser Val Thr Thr Gly Glu Ile Val Lys Lys Met Glu
450 455 460

Ser Pro Met Glu Lys Leu Gln Val Phe Asn Asp Val Leu Ile Asp Arg
465 470 475 480

Arg Gly Glu Gly Asn Lys His Leu Thr Val Tyr Ile Gly Gly Ser Val
485 490 495

Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Gly
500 505 510

Ser Ser Ser Ser Leu Arg Arg Leu Gly Asp His Phe Gly Val Ser Phe
515 520 525

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

530 535 540

Gln Asp Cys Ala Ser Asn Trp Trp Lys Pro Leu Ser Gly Val Leu Tyr
545 550 555 560

Thr Val Ser Ser Trp Ala Glu Ile Gln Ala Phe Ile Leu Gly Thr
565 570 575

<210> 144
<211> 1944
<212> DNA
<213> Phoenix dactylifera
<220><221> CDS
<222> (1)..(1944)
<223> Phoenix_dactylifera gene encoding TMP phosphatase [XP_008796407]
<400> 144

atg cga ttc ctc tcc cct ctt ctc ccc ctc cgc cga aac cca aac cct 48

Met Arg Phe Leu Ser Pro Leu Leu Pro Leu Arg Arg Asn Pro Asn Pro
1 5 10 15

agc cct agg ttc ttc tcg ctc tcc cct ccc ata tcc ctc gcc tcc gcc 96

Ser Pro Arg Phe Phe Ser Leu Ser Pro Pro Ile Ser Leu Ala Ser Ala
20 25 30

tgc ccc cga ttc ggt ttc ttg aat cga gat cgc ccc cgg cgc cgc ctt 144

Cys Pro Arg Phe Gly Phe Leu Asn Arg Asp Arg Pro Arg Arg Arg Leu

35 40 45

cca aag ggg ttc cga tcg atc gcc gcg gcg aat cag cgg gcg tcg cct 192

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

50 55 60

cca aga ttg gtg ccg gag agg gcg gcc gcc acg agt tct tgg cct tct 240

Pro Arg Leu Val Pro Glu Arg Ala Ala Ala Thr Ser Ser Trp Pro Ser

65 70 75 80

tca gcc gga cga gcc atg gca gtg gtg gcg acg gcg gtt gaa gaa ggc 288

Ser Ala Gly Arg Ala Met Ala Val Val Ala Thr Ala Val Glu Glu Gly

85 90 95

tcc gcg gcg aag cgg ttc tgg atc agg tcc cgg aag gag gcg gtg ttc	336
Ser Ala Ala Lys Arg Phe Trp Ile Arg Ser Arg Lys Glu Ala Val Phe	
100 105 110	
gcg gag tac acc ccg ttc gtg gtg tgc ctg gcg gcg ggg aga ctg gag	384
Ala Glu Tyr Thr Pro Phe Val Val Cys Leu Ala Ala Gly Arg Leu Glu	
115 120 125	
atg gag gcc ttc cgc gac tac att gct cag gac gtg cac ttc ctc aat	432
Met Glu Ala Phe Arg Asp Tyr Ile Ala Gln Asp Val His Phe Leu Asn	
130 135 140	
act ttt gcc caa gcg tat gag atg gcg gaa gag tgt gct gat gat gat	480
Thr Phe Ala Gln Ala Tyr Glu Met Ala Glu Glu Cys Ala Asp Asp Asp	
145 150 155 160	
gat gcg aag gct gca ata act gat ctg agg aaa gct gtt ttg gag gaa	528
Asp Ala Lys Ala Ala Ile Thr Asp Leu Arg Lys Ala Val Leu Glu Glu	
165 170 175	
ctg aaa atg cat agt tca ttt gtc caa gaa tgg gga ata gac ccc act	576
Leu Lys Met His Ser Ser Phe Val Gln Glu Trp Gly Ile Asp Pro Thr	
180 185 190	
aaa gaa atc att cct ttc cct gca aca gta aag tac acc gac ttc ctg	624
Lys Glu Ile Ile Pro Phe Pro Ala Thr Val Lys Tyr Thr Asp Phe Leu	
195 200 205	
ctt gct aca gct gca gga aaa gtt gaa gga ggg aaa gat cct ggg aaa	672
Leu Ala Thr Ala Ala Gly Lys Val Glu Gly Gly Lys Asp Pro Gly Lys	
210 215 220	
att gtc act cct ttt gag aag aca aaa att gct gct tat act gta ggt	720
Ile Val Thr Pro Phe Glu Lys Thr Lys Ile Ala Ala Tyr Thr Val Gly	
225 230 235 240	
gcc atg gct cct tgc atg agg ctt tat gca ttc ttg gga aaa gag ctc	768
Ala Met Ala Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Leu	
245 250 255	
cag acg tgt ctg caa ctt gac gaa aat tgt cat ccc tac aaa aag tgg	816

Gln Thr Cys Leu Gln Leu Asp Glu Asn Cys His Pro Tyr Lys Lys Trp

260 265 270

att gat aat tat tcc tct gaa agt ttt gag aca gct gct gtg caa ata 864

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

275 280 285

gaa gaa ttg ctt gac aaa ttg agt gtt tca ttg act ggg gag gag ctt 912

Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu Leu

290 295 300

gaa gac ata gaa aag ctt tac cgc caa gct atg aaa ctt gaa att gaa 960

Glu Asp Ile Glu Lys Leu Tyr Arg Gln Ala Met Lys Leu Glu Ile Glu

305 310 315 320

ttt ttt ctt gct cag cca att gtc cga cca gct gta gtt cct ttg aca 1008

Phe Phe Leu Ala Gln Pro Ile Val Arg Pro Ala Val Val Pro Leu Thr

325 330 335

aga ctg cat gat ccg gca aat tgc ctt gtc att ttt tct gat ttt gac 1056

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

340 345 350

ttg aca tgc agt gta gtt gat tcc tct gcc att tta gca gag att gca 1104

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

355 360 365

ata tta agt gca cca aag act gat aag act ggg act gat aat tta gat 1152

Ile Leu Ser Ala Pro Lys Thr Asp Lys Thr Gly Thr Asp Asn Leu Asp

370 375 380

gct cga agg tct tct tca gaa atg aga aac tca tgg gat gct ctt tct 1200

Ala Arg Arg Ser Ser Ser Glu Met Arg Asn Ser Trp Asp Ala Leu Ser

385 390 395 400

aaa cag tat aca gaa gag tat gag cag tgc ata gaa agc tta ctt cca 1248

Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Ser Leu Leu Pro

405 410 415

tta gaa gaa gct aaa aca ttt gat tat gaa ggc ctt tgc aaa agt ttg	1296
Leu Glu Glu Ala Lys Thr Phe Asp Tyr Glu Gly Leu Cys Lys Ser Leu	
420 425 430	
ggc cag ctc tct gag ttt gag aaa cga gca aat tcc agg gtt att gag	1344
Gly Gln Leu Ser Glu Phe Glu Lys Arg Ala Asn Ser Arg Val Ile Glu	
435 440 445	
tct ggg gtg cta aag gga atg aat cta gat gac ata aaa aga gct ggg	1392
Ser Gly Val Leu Lys Gly Met Asn Leu Asp Asp Ile Lys Arg Ala Gly	
450 455 460	
gaa cgt ttg atc ctc caa gat ggt tgt ata gat ttt ttt cag aag gtt	1440
Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asp Phe Phe Gln Lys Val	
465 470 475 480	
gta aag gaa aag gaa aat cta aat tta gat ctc cat gta ctt tct tat	1488
Val Lys Glu Lys Glu Asn Leu Asn Leu Asp Leu His Val Leu Ser Tyr	
485 490 495	
tgt tgg tgt gcg gat cta ata agg tca gct ttt tca tca gta ggt tgc	1536
Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe Ser Ser Val Gly Cys	
500 505 510	
cta aat gat ttg aac ata cac tca aat gag ttc aac tat caa gaa tct	1584
Leu Asn Asp Leu Asn Ile His Ser Asn Glu Phe Asn Tyr Gln Glu Ser	
515 520 525	
att tca acg ggt gaa att gtt agg aag atg gaa tca ccc atg gac aag	1632
Ile Ser Thr Gly Glu Ile Val Arg Lys Met Glu Ser Pro Met Asp Lys	
530 535 540	
gtt gaa gca ttc aaa agt atc tta agc aac ctt gga agc aat gag aag	1680
Val Glu Ala Phe Lys Ser Ile Leu Ser Asn Leu Gly Ser Asn Glu Lys	
545 550 555 560	
cgc tta tct gtg tac att gga gat tcg gtt ggt gac ttg ctt tgc ctg	1728
Arg Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu	
565 570 575	
ttg gaa gca gat gtt ggt att gtg att gga tca agc act agc tta agg	1776

Leu Glu Ala Asp Val Gly Ile Val Ile Gly Ser Ser Thr Ser Leu Arg
580 585 590
aga atc ggg aag cag ttt ggt gtt tct ttc att cca ctc ttc cgt ggt 1824
Arg Ile Gly Lys Gln Phe Gly Val Ser Phe Ile Pro Leu Phe Arg Gly
595 600 605
ttg gta aac aag caa aga caa ctt aat gaa aaa gac tca tct atc tgg 1872
Leu Val Asn Lys Gln Arg Gln Leu Asn Glu Lys Asp Ser Ser Ile Trp
610 615 620
aag ggg ttg tct ggt gtt ctt tat aca gca tca agc tgg tca gaa ata 1920
Lys Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ser Glu Ile
625 630 635 640
caa gct ttt att ttg ggg gca taa 1944
Gln Ala Phe Ile Leu Gly Ala
645
<210> 145
<211> 647
<212> PRT
<213> Phoenix dactylifera
<400> 145
Met Arg Phe Leu Ser Pro Leu Leu Pro Leu Arg Arg Asn Pro Asn Pro
1 5 10 15
Ser Pro Arg Phe Phe Ser Leu Ser Pro Pro Ile Ser Leu Ala Ser Ala
20 25 30
Cys Pro Arg Phe Gly Phe Leu Asn Arg Asp Arg Pro Arg Arg Arg Leu
35 40 45
Pro Lys Gly Phe Arg Ser Ile Ala Ala Ala Asn Gln Arg Ala Ser Pro
50 55 60
Pro Arg Leu Val Pro Glu Arg Ala Ala Ala Thr Ser Ser Trp Pro Ser
65 70 75 80
Ser Ala Gly Arg Ala Met Ala Val Val Ala Thr Ala Val Glu Glu Gly
85 90 95

Ser Ala Ala Lys Arg Phe Trp Ile Arg Ser Arg Lys Glu Ala Val Phe
100 105 110

Ala Glu Tyr Thr Pro Phe Val Val Cys Leu Ala Ala Gly Arg Leu Glu
115 120 125

Met Glu Ala Phe Arg Asp Tyr Ile Ala Gln Asp Val His Phe Leu Asn
130 135 140

Thr Phe Ala Gln Ala Tyr Glu Met Ala Glu Glu Cys Ala Asp Asp Asp
145 150 155 160

Asp Ala Lys Ala Ala Ile Thr Asp Leu Arg Lys Ala Val Leu Glu Glu
165 170 175

Leu Lys Met His Ser Ser Phe Val Gln Glu Trp Gly Ile Asp Pro Thr
180 185 190

Lys Glu Ile Ile Pro Phe Pro Ala Thr Val Lys Tyr Thr Asp Phe Leu
195 200 205

Leu Ala Thr Ala Ala Gly Lys Val Glu Gly Gly Lys Asp Pro Gly Lys
210 215 220

Ile Val Thr Pro Phe Glu Lys Thr Lys Ile Ala Ala Tyr Thr Val Gly
225 230 235 240

Ala Met Ala Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Leu
245 250 255

Gln Thr Cys Leu Gln Leu Asp Glu Asn Cys His Pro Tyr Lys Lys Trp
260 265 270

Ile Asp Asn Tyr Ser Ser Glu Ser Phe Glu Thr Ala Ala Val Gln Ile
275 280 285

Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu Leu
290 295 300

Glu Asp Ile Glu Lys Leu Tyr Arg Gln Ala Met Lys Leu Glu Ile Glu
305 310 315 320

Phe Phe Leu Ala Gln Pro Ile Val Arg Pro Ala Val Val Pro Leu Thr
325 330 335

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

340					345					350						
Leu	Thr	Cys	Ser	Val	Val	Asp	Ser	Ser	Ala	Ile	Leu	Ala	Glu	Ile	Ala	
355					360					365						
Ile	Leu	Ser	Ala	Pro	Lys	Thr	Asp	Lys	Thr	Gly	Thr	Asp	Asn	Leu	Asp	
370					375					380						
Ala	Arg	Arg	Ser	Ser	Ser	Glu	Met	Arg	Asn	Ser	Trp	Asp	Ala	Leu	Ser	
385					390					395					400	
Lys	Gln	Tyr	Thr	Glu	Glu	Tyr	Glu	Gln	Cys	Ile	Glu	Ser	Leu	Leu	Pro	
405					410					415						
Leu	Glu	Glu	Ala	Lys	Thr	Phe	Asp	Tyr	Glu	Gly	Leu	Cys	Lys	Ser	Leu	
420					425					430						
Gly	Gln	Leu	Ser	Glu	Phe	Glu	Lys	Arg	Ala	Asn	Ser	Arg	Val	Ile	Glu	
435					440					445						
Ser	Gly	Val	Leu	Lys	Gly	Met	Asn	Leu	Asp	Asp	Ile	Lys	Arg	Ala	Gly	
450					455					460						
Glu	Arg	Leu	Ile	Leu	Gln	Asp	Gly	Cys	Ile	Asp	Phe	Phe	Gln	Lys	Val	
465					470					475					480	
Val	Lys	Glu	Lys	Glu	Asn	Leu	Asn	Leu	Asp	Leu	His	Val	Leu	Ser	Tyr	
485					490					495						
Cys	Trp	Cys	Ala	Asp	Leu	Ile	Arg	Ser	Ala	Phe	Ser	Ser	Val	Gly	Cys	
500					505					510						
Leu	Asn	Asp	Leu	Asn	Ile	His	Ser	Asn	Glu	Phe	Asn	Tyr	Gln	Glu	Ser	
515					520					525						
Ile	Ser	Thr	Gly	Glu	Ile	Val	Arg	Lys	Met	Glu	Ser	Pro	Met	Asp	Lys	
530					535					540						
Val	Glu	Ala	Phe	Lys	Ser	Ile	Leu	Ser	Asn	Leu	Gly	Ser	Asn	Glu	Lys	
545					550					555					560	
Arg	Leu	Ser	Val	Tyr	Ile	Gly	Asp	Ser	Val	Gly	Asp	Leu	Leu	Cys	Leu	
565					570					575						
Leu	Glu	Ala	Asp	Val	Gly	Ile	Val	Ile	Gly	Ser	Ser	Thr	Ser	Leu	Arg	
580					585					590						

Arg Ile Gly Lys Gln Phe Gly Val Ser Phe Ile Pro Leu Phe Arg Gly
595 600 605

Leu Val Asn Lys Gln Arg Gln Leu Asn Glu Lys Asp Ser Ser Ile Trp
610 615 620

Lys Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ser Glu Ile
625 630 635 640

Gln Ala Phe Ile Leu Gly Ala
645

<210> 146

<211> 1860

<212> DNA

<213> Zea mays

<220><221> CDS

<222> (1)..(1860)

<223> Zea mays gene encoding TMP phosphatase [XP_008678418.1]

<400> 146

atg ctt gtt ctc cgc cgt ctc cgc ctc cgc ctc cca ctg cca cgc cct 48

Met Leu Val Leu Arg Arg Leu Arg Leu Arg Leu Pro Leu Pro Arg Pro
1 5 10 15

ctt ctc gtc tcc tcc ttc tcc tcc acc tcc ccc tcc tcc tca ccc tcg 96

Leu Leu Val Ser Ser Phe Ser Ser Thr Ser Pro Ser Ser Ser Pro Ser
20 25 30

acc tct agc tcc tcc tcc tgt tgg tcg tcg aca ggc gaa agt aga agg 144

Thr Ser Ser Ser Ser Ser Cys Trp Ser Ser Thr Gly Glu Ser Arg Arg

35 40 45

gcc atg gcg tca tct cct tct ccc gat tcg gcc gcg gtc gtt gcc gag 192

Ala Met Ala Ser Ser Pro Ser Pro Asp Ser Ala Ala Val Val Ala Glu
50 55 60

ggc tcc gcg gct cgc cgc ttc tgg atc gct gcc tcc acg cgc gag gcc 240

Gly Ser Ala Ala Arg Arg Phe Trp Ile Ala Ala Ser Thr Arg Glu Ala
65 70 75 80

gcc ttc gcc gca tac acg ccc ttc ctc ctc tcc ctc gcc gcc ggc aat 288
Ala Phe Ala Ala Tyr Thr Pro Phe Leu Leu Ser Leu Ala Ala Gly Asn
85 90 95

ctg cgg ctc aac gtg ttt cgc cac tac atc gcg cag gac gcg cac ttc 336
Leu Arg Leu Asn Val Phe Arg His Tyr Ile Ala Gln Asp Ala His Phe
100 105 110

ctt cac gcc ttc gct cgc gcg tac gaa atg gcc gag gac tgc gct gat 384

Leu His Ala Phe Ala Arg Ala Tyr Glu Met Ala Glu Asp Cys Ala Asp
115 120 125

gat gac gac gac atg gcc acc ata gcc gcc ctc agg aag gcc atc ctc 432
Asp Asp Asp Asp Met Ala Thr Ile Ala Ala Leu Arg Lys Ala Ile Leu
130 135 140

caa gag ctc aac ctc cac tcc tcc gtt ctg aag gag tgg gga gtt gat 480
Gln Glu Leu Asn Leu His Ser Ser Val Leu Lys Glu Trp Gly Val Asp

145 150 155 160

cct acc aaa gag ata cct cca agt gca gct aca acc aaa tat act gat 528
Pro Thr Lys Glu Ile Pro Pro Ser Ala Ala Thr Thr Lys Tyr Thr Asp
165 170 175

ttc cta ctt gca act gcg gct gga aaa gtt gat ggc aca aaa ggt tct 576
Phe Leu Leu Ala Thr Ala Ala Gly Lys Val Asp Gly Thr Lys Gly Ser
180 185 190

gac aaa atg gtt act cca ttt gag aag act aaa att gct gca tac act 624
Asp Lys Met Val Thr Pro Phe Glu Lys Thr Lys Ile Ala Ala Tyr Thr
195 200 205

gtt ggg gcc atg act cca tgc atg agg ctt tat gca tat cta ggc aaa 672
Val Gly Ala Met Thr Pro Cys Met Arg Leu Tyr Ala Tyr Leu Gly Lys
210 215 220

gaa ctc atg gtt ttc ctt aaa caa gat gaa aat cat cca tac aag aaa 720

Glu Leu Met Val Phe Leu Lys Gln Asp Glu Asn His Pro Tyr Lys Lys
225 230 235 240

tgg att aac aca tat gca tcc agt gat ttt gag gac acc aca ctc caa 768

Trp Ile Asn Thr Tyr Ala Ser Ser Asp Phe Glu Asp Thr Thr Leu Gln
245 250 255
ata gaa gaa ttg cta gac aaa cta agt gtc tca tta act ggt gag gaa 816
Ile Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu
260 265 270
ctt gag att att ggc aag ctc tac cag caa gct atg aaa ctg gaa gtg 864
Leu Glu Ile Ile Gly Lys Leu Tyr Gln Gln Ala Met Lys Leu Glu Val
275 280 285
gag ttc ttt tct tct cag ctt ata gac caa cct gtt gta gct cca ctt 912
Glu Phe Phe Ser Ser Gln Leu Ile Asp Gln Pro Val Val Ala Pro Leu
290 295 300
tca aga tac tgt gat cca aaa tat aaa ctc ttg atc ttt tct gat ttt 960
Ser Arg Tyr Cys Asp Pro Lys Tyr Lys Leu Leu Ile Phe Ser Asp Phe
305 310 315 320
gat ttg acg tgc act att gtt gat tca tct gcc att ttg gcg gag att 1008
Asp Leu Thr Cys Thr Ile Val Asp Ser Ser Ala Ile Leu Ala Glu Ile
325 330 335
gca att ttg tca ttc caa aag gca aat caa agt ggg att gat aat aac 1056
Ala Ile Leu Ser Phe Gln Lys Ala Asn Gln Ser Gly Ile Asp Asn Asn
340 345 350
ctc gac cgt gca aaa tcg gga gac ctg aga agt tcg tgg aac atg ctc 1104
Leu Asp Arg Ala Lys Ser Gly Asp Leu Arg Ser Ser Trp Asn Met Leu
355 360 365
tct aag caa tac atg gaa gag tat gag aaa tgc atg gaa aga cta ctt 1152
Ser Lys Gln Tyr Met Glu Glu Tyr Glu Lys Cys Met Glu Arg Leu Leu
370 375 380
cct cca gaa gaa tcg aag tca cta gat tat gat aaa ctg tat aaa ggc 1200
Pro Pro Glu Glu Ser Lys Ser Leu Asp Tyr Asp Lys Leu Tyr Lys Gly
385 390 395 400
ctg gag gtg cta gct gag ttt gag aag ctt gca aat tct agg gtt gtc 1248
Leu Glu Val Leu Ala Glu Phe Glu Lys Leu Ala Asn Ser Arg Val Val

405	410	415	
gac tct ggt gtg ctg agg gga atg aat ttg gaa gac atc agg aaa gct			1296
Asp Ser Gly Val Leu Arg Gly Met Asn Leu Glu Asp Ile Arg Lys Ala			
420	425	430	
ggt gag cgt ctt att ctt caa ggt ggc tgt aaa aat ttc ttt cag aag			1344
Gly Glu Arg Leu Ile Leu Gln Gly Gly Cys Lys Asn Phe Phe Gln Lys			
435	440	445	
att gta aaa aca agg gag aac ctc aat ttg gat gtc cat att ctt tcc			1392
Ile Val Lys Thr Arg Glu Asn Leu Asn Leu Asp Val His Ile Leu Ser			
450	455	460	
tat tgc tgg tgt gca gaa ctt ata aga tca gcc ttc tca tca gcc ggt			1440
Tyr Cys Trp Cys Ala Glu Leu Ile Arg Ser Ala Phe Ser Ser Ala Gly			
465	470	475	480
tgt cta gat ggt ttg aac ata cat tca aat gag ttt gcc ttt gag gat			1488
Cys Leu Asp Gly Leu Asn Ile His Ser Asn Glu Phe Ala Phe Glu Asp			
485	490	495	
tct gtt tca act ggt gag atc gac aga aag atg cag tct ccg cta gac			1536
Ser Val Ser Thr Gly Glu Ile Asp Arg Lys Met Gln Ser Pro Leu Asp			
500	505	510	
aaa gtt gaa aag ttc aag agc atc aga agt gac gtg gac agt aca gtg			1584
Lys Val Glu Lys Phe Lys Ser Ile Arg Ser Asp Val Asp Ser Thr Val			
515	520	525	
cca ttc cta tct gtt tat att gga gac tcg gtt gga gat ttg ctc tgc			1632
Pro Phe Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys			
530	535	540	
tta ttg gag gct gat att ggt ata gtc att ggg tca acc aca agt ttg			1680
Leu Leu Glu Ala Asp Ile Gly Ile Val Ile Gly Ser Thr Thr Ser Leu			
545	550	555	560
cgt agg gtg ggc aaa cag ttt ggt gtt tct ttt gtc cca ttg ttc cct			1728

Arg Arg Val Gly Lys Gln Phe Gly Val Ser Phe Val Pro Leu Phe Pro
565 570 575
ggg cta gta gag aag cag agg caa ctg gcg gag gaa gat gca tcc gta 1776
Gly Leu Val Glu Lys Gln Arg Gln Leu Ala Glu Glu Asp Ala Ser Val
580 585 590
ttc aag gca cgg tct gga gtc ctc tat acg gtt tct agc tgg tca gaa 1824
Phe Lys Ala Arg Ser Gly Val Leu Tyr Thr Val Ser Ser Trp Ser Glu
595 600 605
ata cac gcc ttc gta ctg gga agt gat ttc agc tga 1860
Ile His Ala Phe Val Leu Gly Ser Asp Phe Ser
610 615
<210> 147
<211> 619
<212> PRT
<213> Zea mays
<400> 147
Met Leu Val Leu Arg Arg Leu Arg Leu Arg Leu Pro Leu Pro Arg Pro
1 5 10 15
Leu Leu Val Ser Ser Phe Ser Ser Thr Ser Pro Ser Ser Ser Pro Ser
20 25 30
Thr Ser Ser Ser Ser Ser Cys Trp Ser Ser Thr Gly Glu Ser Arg Arg
35 40 45
Ala Met Ala Ser Ser Pro Ser Pro Asp Ser Ala Ala Val Val Ala Glu
50 55 60
Gly Ser Ala Ala Arg Arg Phe Trp Ile Ala Ala Ser Thr Arg Glu Ala
65 70 75 80
Ala Phe Ala Ala Tyr Thr Pro Phe Leu Leu Ser Leu Ala Ala Gly Asn
85 90 95
Leu Arg Leu Asn Val Phe Arg His Tyr Ile Ala Gln Asp Ala His Phe
100 105 110
Leu His Ala Phe Ala Arg Ala Tyr Glu Met Ala Glu Asp Cys Ala Asp
115 120 125

Asp Asp Asp Asp Met Ala Thr Ile Ala Ala Leu Arg Lys Ala Ile Leu
 130 135 140
 Gln Glu Leu Asn Leu His Ser Ser Val Leu Lys Glu Trp Gly Val Asp

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

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

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

 340 345 350
 Leu Asp Arg Ala Lys Ser Gly Asp Leu Arg Ser Ser Trp Asn Met Leu
 355 360 365
 Ser Lys Gln Tyr Met Glu Glu Tyr Glu Lys Cys Met Glu Arg Leu Leu

370 375 380
 Pro Pro Glu Glu Ser Lys Ser Leu Asp Tyr Asp Lys Leu Tyr Lys Gly
 385 390 395 400
 Leu Glu Val Leu Ala Glu Phe Glu Lys Leu Ala Asn Ser Arg Val Val

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

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

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

 595 600 605
 Ile His Ala Phe Val Leu Gly Ser Asp Phe Ser
 610 615

<210> 148
 <211> 1830
 <212> DNA
 <213> Oryza sativa
 <220><221> CDS
 <222> (1)..(1830)
 <223> Oryza sativa gene encoding TMP phosphatase [NP_001062539.1]
 <400> 148

atg cgc ggc ctc ctc cgc cgt gtc tac ctc cgc ctc ccc cct ttc cct 48
 Met Arg Gly Leu Leu Arg Arg Val Tyr Leu Arg Leu Pro Pro Phe Pro
 1 5 10 15

cct gcc acc tct ctt tat tat tgg tca aga aca aga cct gca gct gca 96
 Pro Ala Thr Ser Leu Tyr Tyr Trp Ser Arg Thr Arg Pro Ala Ala Ala
 20 25 30

ggg ccc aac cac ccc atc cct agg cgc atg tcg acg tcc tct act gcc 144
 Gly Pro Asn His Pro Ile Pro Arg Arg Met Ser Thr Ser Ser Thr Ala
 35 40 45

gcg gcg gtc gtt gcc gag ggc tcc gcc gct cgc cgc ttc tgg atc gcc 192

Ala Ala Val Val Ala Glu Gly Ser Ala Ala Arg Arg Phe Trp Ile Ala
 50 55 60

gcc gcc tcg agg gag gcc gcc ttc gcc gcc tac acg ccc ttc ctc gtc 240
 Ala Ala Ser Arg Glu Ala Ala Phe Ala Ala Tyr Thr Pro Phe Leu Val
 65 70 75 80

tcc ctc gcc gcc ggg gcc ctc cgc ctg gat tcc ttc cgc caa tac atc 288
 Ser Leu Ala Ala Gly Ala Leu Arg Leu Asp Ser Phe Arg Gln Tyr Ile
 85 90 95

gcc cag gat gcc tac ttc ctc cac gcc ttc gcc cgc gcc tat gag atg 336
 Ala Gln Asp Ala Tyr Phe Leu His Ala Phe Ala Arg Ala Tyr Glu Met
 100 105 110

gcc gag gag tgc gcc gat gac gac gac gac aag gcc acc atc gtc gtc 384
 Ala Glu Glu Cys Ala Asp Asp Asp Asp Asp Lys Ala Thr Ile Val Val

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

Ala Met Arg Leu Glu Val Glu Phe Phe Ser Ala Gln Pro Val Asp Gln
275 280 285
cct gtt gta gct cca ctc tca aga tat tgt ggt ccg aaa gat aag ctc 912
Pro Val Val Ala Pro Leu Ser Arg Tyr Cys Gly Pro Lys Asp Lys Leu
290 295 300
ttg ata ttt tgt gat ttt gat ttg aca tgc act gtt gtt gat tca tct 960
Leu Ile Phe Cys Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser
305 310 315 320
gcc att ttg gcg gag att gca atc ttg tca cac caa aag gct agt caa 1008
Ala Ile Leu Ala Glu Ile Ala Ile Leu Ser His Gln Lys Ala Ser Gln
325 330 335
ggg ggt gct gat agt tcc ctt gat cgt aca aaa tca gcg gac ttg aga 1056
Gly Gly Ala Asp Ser Ser Leu Asp Arg Thr Lys Ser Ala Asp Leu Arg
340 345 350
aat tca tgg aac atg ctc tca aat caa tac atg gaa gag tat gag caa 1104
Asn Ser Trp Asn Met Leu Ser Asn Gln Tyr Met Glu Glu Tyr Glu Gln
355 360 365
tgc ata gca agc ttg ctt cct cca gaa gaa gca agg tca cta gac tat 1152
Cys Ile Ala Ser Leu Leu Pro Pro Glu Glu Ala Arg Ser Leu Asp Tyr
370 375 380
gat caa ctg tat aaa ggt ttg gag gtg cta tcg cag ttt gag aaa ctt 1200
Asp Gln Leu Tyr Lys Gly Leu Glu Val Leu Ser Gln Phe Glu Lys Leu
385 390 395 400
gca aac tct agg gtg gtt gat tct ggt gtc ctg agg gga atg aat tta 1248
Ala Asn Ser Arg Val Val Asp Ser Gly Val Leu Arg Gly Met Asn Leu
405 410 415
gat gac atc cga aaa gct gga gag agg ctt att ctg caa gat gga tgc 1296
Asp Asp Ile Arg Lys Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys
420 425 430
aaa att ttt ttt caa aag att ggc aaa aca agg gag aac ctc aat tta 1344
Lys Ile Phe Phe Gln Lys Ile Gly Lys Thr Arg Glu Asn Leu Asn Leu

435	440	445	
gat gtc cat att ctt tcc tat tgc tgg tgc gca gat ctt ata agg tca			1392
Asp Val His Ile Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser			
450	455	460	
gct ttt tca tca gtt ggt tgt cta gac ggg ctg aac ata cat tca aat			1440
Ala Phe Ser Ser Val Gly Cys Leu Asp Gly Leu Asn Ile His Ser Asn			
465	470	475	480
gag ttt gct ttt gag gga tct gtt tca act ggt cat att aac aga caa			1488
Glu Phe Ala Phe Glu Gly Ser Val Ser Thr Gly His Ile Asn Arg Gln			
485	490	495	
atg gag tct cct ctg gac aaa gct gaa aag ttc aag agc atc aaa agc			1536
Met Glu Ser Pro Leu Asp Lys Ala Glu Lys Phe Lys Ser Ile Lys Ser			
500	505	510	
gac gtg ggt agt aca ggg aca tta ttg tca gtc tat att ggg gac tcg			1584
Asp Val Gly Ser Thr Gly Thr Leu Leu Ser Val Tyr Ile Gly Asp Ser			
515	520	525	
gtt gga gat ttg ctt tgc ttg ttg gag gca gat att ggt att gtt gtt			1632
Val Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val			
530	535	540	
gga tca agc aca acc ttg cgg aga gtg ggc aaa cag ttt ggt gtt tca			1680
Gly Ser Ser Thr Thr Leu Arg Arg Val Gly Lys Gln Phe Gly Val Ser			
545	550	555	560
ttt gtt cct ctg ttc act ggg ttg gta gag aag cag agg cga ata gaa			1728
Phe Val Pro Leu Phe Thr Gly Leu Val Glu Lys Gln Arg Arg Ile Glu			
565	570	575	
aag gaa gaa tca tcc atc ttc aag gca cgg tct gga att ctt tat acg			1776
Lys Glu Glu Ser Ser Ile Phe Lys Ala Arg Ser Gly Ile Leu Tyr Thr			
580	585	590	
gtt tct agc tgg tcg gag gta cag gct ttc atc ctg gga aat gat ttc			1824
Val Ser Ser Trp Ser Glu Val Gln Ala Phe Ile Leu Gly Asn Asp Phe			
595	600	605	

agc tga

1830

Ser

<210> 149

<211> 609

<212> PRT

<213> Oryza sativa

<400> 149

Met Arg Gly Leu Leu Arg Arg Val Tyr Leu Arg Leu Pro Pro Phe Pro

1 5 10 15

Pro Ala Thr Ser Leu Tyr Tyr Trp Ser Arg Thr Arg Pro Ala Ala Ala

20 25 30

Gly Pro Asn His Pro Ile Pro Arg Arg Met Ser Thr Ser Ser Thr Ala

35 40 45

Ala Ala Val Val Ala Glu Gly Ser Ala Ala Arg Arg Phe Trp Ile Ala

50 55 60

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

65 70 75 80

Ser Leu Ala Ala Gly Ala Leu Arg Leu Asp Ser Phe Arg Gln Tyr Ile

85 90 95

Ala Gln Asp Ala Tyr Phe Leu His Ala Phe Ala Arg Ala Tyr Glu Met

100 105 110

Ala Glu Glu Cys Ala Asp Asp Asp Asp Lys Ala Thr Ile Val Val

115 120 125

Leu Arg Lys Ala Ile Leu Arg Glu Leu Asn Leu His Ala Ser Val Leu

130 135 140

Gln Glu Trp Gly Val Asp Pro Asn Lys Glu Ile Pro Pro Ile Pro Ala

145 150 155 160

Thr Thr Lys Tyr Thr Asp Phe Leu Leu Ala Thr Ser Thr Gly Lys Val

165 170 175

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

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

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

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

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

 420 425 430

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

<210> 150

<211> 1683

<212> DNA

<213> Picea sitchensis

<220><221> CDS

<222> (1)..(1683)

<223> Picea_sitchensis gene encoding TMP phosphatase [ABR16455]

<400> 150

atg ggg gtc gcc gat gaa gca gga gtc gcc aga agg cta tgg aca aag	48
Met Gly Val Ala Asp Glu Ala Gly Val Ala Arg Arg Leu Trp Thr Lys	
1 5 10 15	
ttc aag aaa gac acc gcg ctt gca cag tat aat tcc ttt gtt gtt gct	96
Phe Lys Lys Asp Thr Ala Leu Ala Gln Tyr Asn Ser Phe Val Val Ala	
20 25 30	
ttg gcg gcc ggg acg ctc aac atg acg tct ttt cag cag tac atg gcg	144
Leu Ala Ala Gly Thr Leu Asn Met Thr Ser Phe Gln Gln Tyr Met Ala	
35 40 45	
cag gat gct tat ttt ctc aaa gca ttt gct cag gcg tac aca atg gca	192
Gln Asp Ala Tyr Phe Leu Lys Ala Phe Ala Gln Ala Tyr Thr Met Ala	
50 55 60	
gag gat tgc gca gat gat gac gac gac aaa gca tcg atc cgt gaa cta	240
Glu Asp Cys Ala Asp Asp Asp Asp Asp Lys Ala Ser Ile Arg Glu Leu	
65 70 75 80	
cga aaa gcc gct gag gaa gag ctc aat ctg cac aat tcc ttg gct gag	288
Arg Lys Ala Ala Glu Glu Glu Leu Asn Leu His Asn Ser Leu Ala Glu	
85 90 95	
gac tgg gac gtt gaa ttt gca aaa gag tgc tct ccc aat atg gca aca	336
Asp Trp Asp Val Glu Phe Ala Lys Glu Cys Ser Pro Asn Met Ala Thr	
100 105 110	
gtc aag tac aca gaa ttt tta ttg gca aca gct gct ggc aag gtg gaa	384
Val Lys Tyr Thr Glu Phe Leu Leu Ala Thr Ala Ala Gly Lys Val Glu	
115 120 125	
gga ggg aag gga cca agc aga agt gtg act cct ttt gag aaa aca aaa	432
Gly Gly Lys Gly Pro Ser Arg Ser Val Thr Pro Phe Glu Lys Thr Lys	
130 135 140	
ata gca gca tac aca gtg ggt gcc atg acc ccg tgc atg agg ctt tat	480
Ile Ala Ala Tyr Thr Val Gly Ala Met Thr Pro Cys Met Arg Leu Tyr	
145 150 155 160	
gct ttc ttg ggc caa gaa att gtc aaa gcc ctg gaa cct gat tgc agt	528

Ala Phe Leu Gly Gln Glu Ile Val Lys Ala Leu Glu Pro Asp Cys Ser	
165	170
175	
aat cat cca tat aag cag tgg att gaa aca tac tct tct gca aag ttt	576
Asn His Pro Tyr Lys Gln Trp Ile Glu Thr Tyr Ser Ser Ala Lys Phe	
180	185
190	
gag gca tcg gca tta caa act gaa gag ttg ctt gac aaa ctg gct att	624
Glu Ala Ser Ala Leu Gln Thr Glu Glu Leu Leu Asp Lys Leu Ala Ile	
195	200
205	
tcg cta act ggg gaa gag ctt gaa gtg ctg cgg agg ttg tat tat cat	672
Ser Leu Thr Gly Glu Glu Leu Glu Val Leu Arg Arg Leu Tyr Tyr His	
210	215
220	
gcc tta aaa cta gaa ata gaa ttc ttt tcc gct cag cct ttc tct cag	720
Ala Leu Lys Leu Glu Ile Glu Phe Phe Ser Ala Gln Pro Phe Ser Gln	
225	230
235	240
aga aca tta gtt ccg atg ttg aaa ctg ggt gat tca gcc agc cgc cga	768
Arg Thr Leu Val Pro Met Leu Lys Leu Gly Asp Ser Ala Ser Arg Arg	
245	250
255	
tat acc att gtc tca gat ttc gat ttg tct tgc act gtc ttg gat tct	816
Tyr Thr Ile Val Ser Asp Phe Asp Leu Ser Cys Thr Val Leu Asp Ser	
260	265
270	
tca gca gta tta gca gaa att gca ata ttg act act ctc aaa act gag	864
Ser Ala Val Leu Ala Glu Ile Ala Ile Leu Thr Thr Leu Lys Thr Glu	
275	280
285	
caa aat ggt gct gaa aac tta agt gat cac aag tca tca tcg gag ttg	912
Gln Asn Gly Ala Glu Asn Leu Ser Asp His Lys Ser Ser Ser Glu Leu	
290	295
300	
aga aaa act tgg gat gca ctt tct agt caa tat tct gaa gaa tgt gaa	960
Arg Lys Thr Trp Asp Ala Leu Ser Ser Gln Tyr Ser Glu Glu Cys Glu	
305	310
315	320

gaa tgc tta agg aag act ctg cca cct gaa gaa gtg ggc tct ttt gat	1008
Glu Cys Leu Arg Lys Thr Leu Pro Pro Glu Glu Val Gly Ser Phe Asp	
325 330 335	
tat gaa ggc cta cac caa tct ctt gag cat ctg tct cag ttt gaa atg	1056
Tyr Glu Gly Leu His Gln Ser Leu Glu His Leu Ser Gln Phe Glu Met	
340 345 350	
gag gca aac tct aaa gtt gtc gag tca ggt gtc ctt gag ggc att aat	1104
Glu Ala Asn Ser Lys Val Val Glu Ser Gly Val Leu Glu Gly Ile Asn	
355 360 365	
ata gat gac att aaa aag gca gga gag cgt ctt gca ttt cag gat gga	1152
Ile Asp Asp Ile Lys Lys Ala Gly Glu Arg Leu Ala Phe Gln Asp Gly	
370 375 380	
tgc gca aac ttt ttt gaa caa atc cta acg aaa atg gac agc tta aat	1200
Cys Ala Asn Phe Phe Glu Gln Ile Leu Thr Lys Met Asp Ser Leu Asn	
385 390 395 400	
gtg gat gtg cac ata att tct gtt tgt tgg agt gga gat atc atc agg	1248
Val Asp Val His Ile Ile Ser Val Cys Trp Ser Gly Asp Ile Ile Arg	
405 410 415	
gct gct ttt tca tca agc ggt ttg gat ggt tta cag gtt cat tca aat	1296
Ala Ala Phe Ser Ser Ser Gly Leu Asp Gly Leu Gln Val His Ser Asn	
420 425 430	
gaa ctc acc ttt gtg gaa tca gtc tct act ggt ggt att gat agg cgt	1344
Glu Leu Thr Phe Val Glu Ser Val Ser Thr Gly Gly Ile Asp Arg Arg	
435 440 445	
gtt gag tcc cca gtt gac aag ttg aaa atc ttc aat aat att tgg agt	1392
Val Glu Ser Pro Val Asp Lys Leu Lys Ile Phe Asn Asn Ile Trp Ser	
450 455 460	
tct tca aag gac cag gac acg gaa cat atc tct ata tac att ggg gac	1440
Ser Ser Lys Asp Gln Asp Thr Glu His Ile Ser Ile Tyr Ile Gly Asp	

465													470													475													480																																		
ggg tta ggt gac ttg ctt tgt ctt ctt cag gca gat att gga ata gtg																																																1488																									
Gly	Leu	Gly	Asp	Leu	Leu	Cys	Leu	Leu	Gln	Ala	Asp	Ile	Gly	Ile	Val																																																										
485																490																495																																									
att ggt aca agc tca acg cta aga agg gtt gga aaa cgt ttt gga gta																																																1536																									
Ile	Gly	Thr	Ser	Ser	Thr	Leu	Arg	Arg	Val	Gly	Lys	Arg	Phe	Gly	Val																																																										
500																505																510																																									
tcc ttt gtt cct ttg ttt tct ggt ctt ctc aaa cag gag aga gca tat																																																1584																									
Ser	Phe	Val	Pro	Leu	Phe	Ser	Gly	Leu	Leu	Lys	Gln	Glu	Arg	Ala	Tyr																																																										
515																520																525																																									
gta gaa ggt tct agt tgt tgg aca aaa caa agt ggt att ctt tat acc																																																1632																									
Val	Glu	Gly	Ser	Ser	Cys	Trp	Thr	Lys	Gln	Ser	Gly	Ile	Leu	Tyr	Thr																																																										
530																535																540																																									
gtc tct agt tgg agt gaa ata cat gct ttt att ttg ggc tct tcc aat																																																1680																									
Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Ser Ser Asn																																																																									
545																550																555																560																									
tga																																																1683																									
<210> 151																																																																									
<211> 560																																																																									
<212> PRT																																																																									
<213> Picea sitchensis																																																																									
<400> 151																																																																									
Met Gly Val Ala Asp Glu Ala Gly Val Ala Arg Arg Leu Trp Thr Lys																																																																									
1																5																10																15																									
Phe Lys Lys Asp Thr Ala Leu Ala Gln Tyr Asn Ser Phe Val Val Ala																																																																									
20																25																30																																									
Leu Ala Ala Gly Thr Leu Asn Met Thr Ser Phe Gln Gln Tyr Met Ala																																																																									
35																40																45																																									
Gln Asp Ala Tyr Phe Leu Lys Ala Phe Ala Gln Ala Tyr Thr Met Ala																																																																									
50																55																60																																									
Glu Asp Cys Ala Asp Asp Asp Asp Asp Lys Ala Ser Ile Arg Glu Leu																																																																									
65																70																75																80																									

Arg Lys Ala Ala Glu Glu Glu Leu Asn Leu His Asn Ser Leu Ala Glu
85 90 95

Asp Trp Asp Val Glu Phe Ala Lys Glu Cys Ser Pro Asn Met Ala Thr
100 105 110

Val Lys Tyr Thr Glu Phe Leu Leu Ala Thr Ala Ala Gly Lys Val Glu
115 120 125

Gly Gly Lys Gly Pro Ser Arg Ser Val Thr Pro Phe Glu Lys Thr Lys
130 135 140

Ile Ala Ala Tyr Thr Val Gly Ala Met Thr Pro Cys Met Arg Leu Tyr
145 150 155 160

Ala Phe Leu Gly Gln Glu Ile Val Lys Ala Leu Glu Pro Asp Cys Ser
165 170 175

Asn His Pro Tyr Lys Gln Trp Ile Glu Thr Tyr Ser Ser Ala Lys Phe
180 185 190

Glu Ala Ser Ala Leu Gln Thr Glu Glu Leu Leu Asp Lys Leu Ala Ile
195 200 205

Ser Leu Thr Gly Glu Glu Leu Glu Val Leu Arg Arg Leu Tyr Tyr His
210 215 220

Ala Leu Lys Leu Glu Ile Glu Phe Phe Ser Ala Gln Pro Phe Ser Gln
225 230 235 240

Arg Thr Leu Val Pro Met Leu Lys Leu Gly Asp Ser Ala Ser Arg Arg
245 250 255

Tyr Thr Ile Val Ser Asp Phe Asp Leu Ser Cys Thr Val Leu Asp Ser
260 265 270

Ser Ala Val Leu Ala Glu Ile Ala Ile Leu Thr Thr Leu Lys Thr Glu
275 280 285

Gln Asn Gly Ala Glu Asn Leu Ser Asp His Lys Ser Ser Ser Glu Leu
290 295 300

Arg Lys Thr Trp Asp Ala Leu Ser Ser Gln Tyr Ser Glu Glu Cys Glu
305 310 315 320

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

325 330 335
Tyr Glu Gly Leu His Gln Ser Leu Glu His Leu Ser Gln Phe Glu Met
340 345 350

Glu Ala Asn Ser Lys Val Val Glu Ser Gly Val Leu Glu Gly Ile Asn
355 360 365

Ile Asp Asp Ile Lys Lys Ala Gly Glu Arg Leu Ala Phe Gln Asp Gly
370 375 380

Cys Ala Asn Phe Phe Glu Gln Ile Leu Thr Lys Met Asp Ser Leu Asn
385 390 395 400

Val Asp Val His Ile Ile Ser Val Cys Trp Ser Gly Asp Ile Ile Arg
405 410 415

Ala Ala Phe Ser Ser Ser Gly Leu Asp Gly Leu Gln Val His Ser Asn
420 425 430

Glu Leu Thr Phe Val Glu Ser Val Ser Thr Gly Gly Ile Asp Arg Arg
435 440 445

Val Glu Ser Pro Val Asp Lys Leu Lys Ile Phe Asn Asn Ile Trp Ser
450 455 460

Ser Ser Lys Asp Gln Asp Thr Glu His Ile Ser Ile Tyr Ile Gly Asp
465 470 475 480

Gly Leu Gly Asp Leu Leu Cys Leu Leu Gln Ala Asp Ile Gly Ile Val
485 490 495

Ile Gly Thr Ser Ser Thr Leu Arg Arg Val Gly Lys Arg Phe Gly Val
500 505 510

Ser Phe Val Pro Leu Phe Ser Gly Leu Leu Lys Gln Glu Arg Ala Tyr
515 520 525

Val Glu Gly Ser Ser Cys Trp Thr Lys Gln Ser Gly Ile Leu Tyr Thr
530 535 540

Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Ser Ser Asn
545 550 555 560

<210> 152

<211> 1641

<212> DNA

<213> Physcomitrella patens

<220><221> CDS

<222> (1)..(1641)

<223> Physcomitrella patens gene encoding TMP phosphatase

[XP_001769831]

<400> 152

atg aat ttg agc acg caa gct aac aca ggg ttg gcg aag agc ttc tgg 48

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

1 5 10 15

gct agt tgt aag aga gag gct tat gca tca ctc tac cat ccg ttt gtg 96

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

20 25 30

gtt gcg tta gcg gct ggc acc ttg cca aaa caa act ttt caa cgt tac 144

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

35 40 45

atg gca cag gat gcc tat ttc ttg gag gcg ttc aag aat gcg tat caa 192

Met Ala Gln Asp Ala Tyr Phe Leu Glu Ala Phe Lys Asn Ala Tyr Gln

50 55 60

ctg gct atg gaa acc act aca gac gaa gag gca aag gcc atc att gag 240

Leu Ala Met Glu Thr Thr Thr Asp Glu Glu Ala Lys Ala Ile Ile Glu

65 70 75 80

tcc ctt cag aga gat gtg cag gaa gag ctc aat ttg cac tcg tcg atc 288

Ser Leu Gln Arg Asp Val Gln Glu Glu Leu Asn Leu His Ser Ser Ile

85 90 95

atg cag tct ttg gat gct acc gat cag aat tgc ttt gaa cca aac atg 336

Met Gln Ser Leu Asp Ala Thr Asp Gln Asn Cys Phe Glu Pro Asn Met

100 105 110

gca aca aca gcg tat tgt gat ttt ctg cta gcc aca gct aca gga agt 384

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

115 120 125

aac gaa gca caa aaa ttt gga agc aca agt gct caa atc ata acc gct 432
 Asn Glu Ala Gln Lys Phe Gly Ser Thr Ser Ala Gln Ile Ile Thr Ala
 130 135 140
 atg act cct tgc atg cgg cta tat gca ttt ttg ggg cag gag ctc aaa 480
 Met Thr Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Gln Glu Leu Lys
 145 150 155 160
 aaa cac gtt gat cat gtt gct gac cat cct tac cag gag tgg att gat 528

 Lys His Val Asp His Val Ala Asp His Pro Tyr Gln Glu Trp Ile Asp
 165 170 175
 act tac tct gct gca gag ttc gag gct gca gct tcg aag att gag cag 576
 Thr Tyr Ser Ala Ala Glu Phe Glu Ala Ala Ala Ser Lys Ile Glu Gln
 180 185 190
 ctg cta gac aag tta act gct act ttg act gga aag cat gaa ata gca 624
 Leu Leu Asp Lys Leu Thr Ala Thr Leu Thr Gly Lys His Glu Ile Ala

 195 200 205
 ttc tta gaa agt ctc tat ctt caa gcc atg aac ttg gag gtg gat ttc 672
 Phe Leu Glu Ser Leu Tyr Leu Gln Ala Met Asn Leu Glu Val Asp Phe
 210 215 220
 ttc ggt gct cag ctg tta ggg cct gtg ctc gta ccc ttc ctc aaa tgc 720
 Phe Gly Ala Gln Leu Leu Gly Pro Val Leu Val Pro Phe Leu Lys Cys
 225 230 235 240

 caa ccg gct cca gag agc tat ata tta ctt gcg tct gac ttt gat tcc 768
 Gln Pro Ala Pro Glu Ser Tyr Ile Leu Leu Ala Ser Asp Phe Asp Ser
 245 250 255
 acg tgc acg ata tct gat tca tgc ccc ata ttg gca gac ctg acc gtg 816
 Thr Cys Thr Ile Ser Asp Ser Cys Pro Ile Leu Ala Asp Leu Thr Val
 260 265 270
 caa act gcg cga aaa tct cac ggt ggt cgt tca gtt ggt gaa tca ggg 864

 Gln Thr Ala Arg Lys Ser His Gly Gly Arg Ser Val Gly Glu Ser Gly
 275 280 285
 gcc agc ttg ttg aaa aaa aga tgg gat gat ctc gtc atg cag tat atg 912

Ala Ser Leu Leu Lys Lys Arg Trp Asp Asp Leu Val Met Gln Tyr Met
290 295 300
gac gag tat gag gac gtt ctg aag cga agc ctg gtg aaa aaa gat aat 960
Asp Glu Tyr Glu Asp Val Leu Lys Arg Ser Leu Val Lys Lys Asp Asn
305 310 315 320
ggc agt gtt aat gcg ctc agt gca gag aat ctc caa gag ttt ctg aag 1008
Gly Ser Val Asn Ala Leu Ser Ala Glu Asn Leu Gln Glu Phe Leu Lys
325 330 335
gaa atg tcc aac ttc gaa cag aag gcc aat gcg agg gtc gaa gag gct 1056
Glu Met Ser Asn Phe Glu Gln Lys Ala Asn Ala Arg Val Glu Glu Ala
340 345 350
gca gtt cta aaa ggc tta tct ctg gct tcg att caa gaa gct gga aaa 1104
Ala Val Leu Lys Gly Leu Ser Leu Ala Ser Ile Gln Glu Ala Gly Lys
355 360 365
tcc atg cct ctt cgt gag ggc tgt tct gac ttt ttt aag cgt ctg gaa 1152
Ser Met Pro Leu Arg Glu Gly Cys Ser Asp Phe Phe Lys Arg Leu Glu
370 375 380
tca gga gag gtt ctt gtt gac aca tgt ata ttg tct gtg tgc tgg agc 1200
Ser Gly Glu Val Leu Val Asp Thr Cys Ile Leu Ser Val Cys Trp Ser
385 390 395 400
aaa acc ttc atc gaa gct gtc ttg gaa aag gtt cgt att cca aac atc 1248
Lys Thr Phe Ile Glu Ala Val Leu Glu Lys Val Arg Ile Pro Asn Ile
405 410 415
aat gcc aac gag ctc gtt ttc gaa gga cgc att tcc acc ggt gct att 1296
Asn Ala Asn Glu Leu Val Phe Glu Gly Arg Ile Ser Thr Gly Ala Ile
420 425 430
atc aaa aac gtc gaa acg gct ctt gac aag caa aga cac ttc gtt cag 1344
Ile Lys Asn Val Glu Thr Ala Leu Asp Lys Gln Arg His Phe Val Gln
435 440 445
ttg ctg gat aat cta aaa cca act caa gac gtg ctg tcc att tat gtt 1392
Leu Leu Asp Asn Leu Lys Pro Thr Gln Asp Val Leu Ser Ile Tyr Val

450	455	460										
ggt gat agt ctg act gat ctt ctc tgc cta atc aga gca gac ctg ggt												1440
Gly Asp Ser Leu Thr Asp Leu Leu Cys Leu Ile Arg Ala Asp Leu Gly												
465	470	475	480									
ata gtt ctc ggt gac agc agc gct ctg aag cag gtg tat ggg cca aaa												1488
Ile Val Leu Gly Asp Ser Ser Ala Leu Lys Gln Val Tyr Gly Pro Lys												
485	490	495										
atg gcc ccc ctc ttc atg aaa gcc ata ctc ttg gag cag gca aac atg												1536
Met Ala Pro Leu Phe Met Lys Ala Ile Leu Leu Glu Gln Ala Asn Met												
500	505	510										
cga ggc agg cag caa ccc aca ggt tac gtc ttc act gtc tcc agt tgg												1584
Arg Gly Arg Gln Gln Pro Thr Gly Tyr Val Phe Thr Val Ser Ser Trp												
515	520	525										
tat gag gtg gaa gcc ttt ctg ttg ggt cct gct aga aac aga cct ttg												1632
Tyr Glu Val Glu Ala Phe Leu Leu Gly Pro Ala Arg Asn Arg Pro Leu												
530	535	540										
tac atc tag												1641
Tyr Ile												
545												
<210> 153												
<211> 546												
<212> PRT												
<213> Physcomitrella patens												
<400> 153												
Met Asn Leu Ser Thr Gln Ala Asn Thr Gly Leu Ala Lys Ser Phe Trp												
1 5 10 15												
Ala Ser Cys Lys Arg Glu Ala Tyr Ala Ser Leu Tyr His Pro Phe Val												
20 25 30												
Val Ala Leu Ala Ala Gly Thr Leu Pro Lys Gln Thr Phe Gln Arg Tyr												
35 40 45												
Met Ala Gln Asp Ala Tyr Phe Leu Glu Ala Phe Lys Asn Ala Tyr Gln												

50				55				60							
Leu	Ala	Met	Glu	Thr	Thr	Thr	Asp	Glu	Glu	Ala	Lys	Ala	Ile	Ile	Glu
65				70				75				80			
Ser	Leu	Gln	Arg	Asp	Val	Gln	Glu	Glu	Leu	Asn	Leu	His	Ser	Ser	Ile
85				90				95							
Met	Gln	Ser	Leu	Asp	Ala	Thr	Asp	Gln	Asn	Cys	Phe	Glu	Pro	Asn	Met
100				105				110							
Ala	Thr	Thr	Ala	Tyr	Cys	Asp	Phe	Leu	Leu	Ala	Thr	Ala	Thr	Gly	Ser
115				120				125							
Asn	Glu	Ala	Gln	Lys	Phe	Gly	Ser	Thr	Ser	Ala	Gln	Ile	Ile	Thr	Ala
130				135				140							
Met	Thr	Pro	Cys	Met	Arg	Leu	Tyr	Ala	Phe	Leu	Gly	Gln	Glu	Leu	Lys
145				150				155				160			
Lys	His	Val	Asp	His	Val	Ala	Asp	His	Pro	Tyr	Gln	Glu	Trp	Ile	Asp
165				170				175							
Thr	Tyr	Ser	Ala	Ala	Glu	Phe	Glu	Ala	Ala	Ala	Ser	Lys	Ile	Glu	Gln
180				185				190							
Leu	Leu	Asp	Lys	Leu	Thr	Ala	Thr	Leu	Thr	Gly	Lys	His	Glu	Ile	Ala
195				200				205							
Phe	Leu	Glu	Ser	Leu	Tyr	Leu	Gln	Ala	Met	Asn	Leu	Glu	Val	Asp	Phe
210				215				220							
Phe	Gly	Ala	Gln	Leu	Leu	Gly	Pro	Val	Leu	Val	Pro	Phe	Leu	Lys	Cys
225				230				235				240			
Gln	Pro	Ala	Pro	Glu	Ser	Tyr	Ile	Leu	Leu	Ala	Ser	Asp	Phe	Asp	Ser
245				250				255							
Thr	Cys	Thr	Ile	Ser	Asp	Ser	Cys	Pro	Ile	Leu	Ala	Asp	Leu	Thr	Val
260				265				270							
Gln	Thr	Ala	Arg	Lys	Ser	His	Gly	Gly	Arg	Ser	Val	Gly	Glu	Ser	Gly
275				280				285							
Ala	Ser	Leu	Leu	Lys	Lys	Arg	Trp	Asp	Asp	Leu	Val	Met	Gln	Tyr	Met
290				295				300							

Asp Glu Tyr Glu Asp Val Leu Lys Arg Ser Leu Val Lys Lys Asp Asn
 305 310 315 320
 Gly Ser Val Asn Ala Leu Ser Ala Glu Asn Leu Gln Glu Phe Leu Lys
 325 330 335

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

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

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

 Tyr Glu Val Glu Ala Phe Leu Leu Gly Pro Ala Arg Asn Arg Pro Leu
 530 535 540
 Tyr Ile

545

<210> 154

<211> 1593

<212> DNA

<213> Selaginella moellendorffii

<220><221> CDS

<222> (1)..(1593)

<223> Selaginella_moellendorffii gene encoding TMP phosphatase
[XP_002990363]

<400> 154

atg tcg tgt ttg ctt aga aat gta gtg gcc aga gga ttg agg agc ttg 48
Met Ser Cys Leu Leu Arg Asn Val Val Ala Arg Gly Leu Arg Ser Leu

1 5 10 15
gca agc gcc cag gcg atg gag cca tcc att tca aag cgc ttg tgg cag 96
Ala Ser Ala Gln Ala Met Glu Pro Ser Ile Ser Lys Arg Leu Trp Gln

20 25 30
caa tcc aag cgc gag gca atg gta tgt ctg tat cat cca ttt gtg gtg 144
Gln Ser Lys Arg Glu Ala Met Val Cys Leu Tyr His Pro Phe Val Val

35 40 45
tcc atc gct gct ggg acg ctg gat ctt cac agc ttc cag cga ttc ata 192
Ser Ile Ala Ala Gly Thr Leu Asp Leu His Ser Phe Gln Arg Phe Ile

50 55 60
gcg cag gat tcc ttc ttc ctg acg gca ttc gcg aaa gcc tat ggt ttg 240
Ala Gln Asp Ser Phe Phe Leu Thr Ala Phe Ala Lys Ala Tyr Gly Leu

65 70 75 80
gcc ata gag cgc agc gat gat cga gaa gtt aaa tct gag att tgc aag 288

Ala Ile Glu Arg Ser Asp Asp Arg Glu Val Lys Ser Glu Ile Cys Lys
85 90 95

ctc caa cag gct gtg tac gag gaa ctt gag ctc cat tct tcc ctc atg 336
Leu Gln Gln Ala Val Tyr Glu Glu Leu Glu Leu His Ser Ser Leu Met

100 105 110
aag gct tgg aac ttc gat cat aca cca cca tcg cca gca act tgt gct 384

Lys Ala Trp Asn Phe Asp His Thr Pro Pro Ser Pro Ala Thr Cys Ala

115 120 125

tac aca gat ttt ctc ctc gca gtg gct gct ggg aag aaa att gaa tgc 432

Tyr Thr Asp Phe Leu Leu Ala Val Ala Ala Gly Lys Lys Ile Glu Cys

130 135 140

gag aaa act aag gtg ccg atg ctc gct ctg gca gca atg gct ccg tgc 480

Glu Lys Thr Lys Val Pro Met Leu Ala Leu Ala Ala Met Ala Pro Cys

145 150 155 160

atg cgt ctc tac gct ttc cta ggc caa gag acg aga gtt ttc tct cga 528

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

165 170 175

gaa aat catcca tat cgc gac tgg att tcg act tac tcg tcg cct ggt 576

Glu Asn His Pro Tyr Arg Asp Trp Ile Ser Thr Tyr Ser Ser Pro Gly

180 185 190

ttc gag act gct gct act cga ctc gag cag ctt ctc gat agc ctc tcg 624

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

195 200 205

gaa gct caa gag act acg gca gcg gaa ttt cag agt atg caa agt ttg 672

Glu Ala Gln Glu Thr Thr Ala Ala Glu Phe Gln Ser Met Gln Ser Leu

210 215 220

tat cac cgt gcc ata gcg tac gag gtg agc ttc ttc gat gcc cag gaa 720

Tyr His Arg Ala Ile Ala Tyr Glu Val Ser Phe Phe Asp Ala Gln Glu

225 230 235 240

gtg cgt ggc agc aac gct ttt gtc ccg ctg cta gag agt gta gca ctc 768

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

245 250 255

aag gat cgc aac ttc gtc ctc atc tct gat ttt gat tct act tgc acc 816

Lys Asp Arg Asn Phe Val Leu Ile Ser Asp Phe Asp Ser Thr Cys Thr

260 265 270

gtc tct gat tca tcc cca gtt cta gcg gag ctg gct atg gcg gtc gat	864
Val Ser Asp Ser Ser Pro Val Leu Ala Glu Leu Ala Met Ala Val Asp	
275 280 285	
cca aat gta agg agg aaa tgg agc agc ctc tcg gac gag tat ttc agg	912
Pro Asn Val Arg Arg Lys Trp Ser Ser Leu Ser Asp Glu Tyr Phe Arg	
290 295 300	
gac tac tcc aaa ctc ctg gaa gaa gtt gtt ctt cgt gag tac gac tac	960
Asp Tyr Ser Lys Leu Leu Glu Glu Val Val Leu Arg Glu Tyr Asp Tyr	
305 310 315 320	
gat gcg atc aaa gag gct ctc caa gtt ctt tcc gag ttt gag aag caa	1008
Asp Ala Ile Lys Glu Ala Leu Gln Val Leu Ser Glu Phe Glu Lys Gln	
325 330 335	
ggg aac gcg aaa atc gac gcc tcc cgc gtt ttg caa ggc att aag atc	1056
Gly Asn Ala Lys Ile Asp Ala Ser Arg Val Leu Gln Gly Ile Lys Ile	
340 345 350	
gat gat atc aag caa gcc gga caa aac atg gca ctt caa gct ggc tgt	1104
Asp Asp Ile Lys Gln Ala Gly Gln Asn Met Ala Leu Gln Ala Gly Cys	
355 360 365	
gcc agt gtg ctt tgc agg cta agt tcc aaa atc tct tgt caa atc ctc	1152
Ala Ser Val Leu Cys Arg Leu Ser Ser Lys Ile Ser Cys Gln Ile Leu	
370 375 380	
tcg gtt tgc tgg agc cgg acc ttc atc gaa gca gct ttc tcc aaa gag	1200
Ser Val Cys Trp Ser Arg Thr Phe Ile Glu Ala Ala Phe Ser Lys Glu	
385 390 395 400	
aat atc acc aat gtt cct gtc cat tcc aac gaa ctc gaa aac gat ggg	1248
Asn Ile Thr Asn Val Pro Val His Ser Asn Glu Leu Glu Asn Asp Gly	
405 410 415	
aac ttt aca acc ggg agc ttg atc aga cgc gtc gag aca ccg att gac	1296
Asn Phe Thr Thr Gly Ser Leu Ile Arg Arg Val Glu Thr Pro Ile Asp	
420 425 430	
aag gaa gag acg atg ttt cgt gag att cta cac gct ccg gac gac aag	1344

Lys Glu Glu Thr Met Phe Arg Glu Ile Leu His Ala Pro Asp Asp Lys
435 440 445
ttt gtg att ttc att gga gac agc ctc acg gat ctg cta gcc ttg ctc 1392
Phe Val Ile Phe Ile Gly Asp Ser Leu Thr Asp Leu Leu Ala Leu Leu
450 455 460
cga gct gac att gga att gtt cta gga acg agc tcc agc ctc gat cga 1440
Arg Ala Asp Ile Gly Ile Val Leu Gly Thr Ser Ser Ser Leu Asp Arg
465 470 475 480
gcc tcc aaa gcc ttt gga gtg aag atc gtg cca ctc ttt tcc ggc ctc 1488
Ala Ser Lys Ala Phe Gly Val Lys Ile Val Pro Leu Phe Ser Gly Leu
485 490 495
gtc cag cgg cag caa agc tct cga tca gcg tgg aga aaa gag gaa gga 1536
Val Gln Arg Gln Gln Ser Ser Arg Ser Ala Trp Arg Lys Glu Glu Gly
500 505 510
gtt ttg tat cga gct tct gga tgg ctg gag ata gaa gcg ttt cta gct 1584
Val Leu Tyr Arg Ala Ser Gly Trp Leu Glu Ile Glu Ala Phe Leu Ala
515 520 525
ggt aat tag 1593

Gly Asn

530

<210> 155

<211> 530

<212> PRT

<213> Selaginella moellendorffii

<400> 155

Met Ser Cys Leu Leu Arg Asn Val Val Ala Arg Gly Leu Arg Ser Leu

1 5 10 15

Ala Ser Ala Gln Ala Met Glu Pro Ser Ile Ser Lys Arg Leu Trp Gln

20 25 30

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

35

40

45

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

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

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

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

 485 490 495
 Val Gln Arg Gln Gln Ser Ser Arg Ser Ala Trp Arg Lys Glu Glu Gly
 500 505 510
 Val Leu Tyr Arg Ala Ser Gly Trp Leu Glu Ile Glu Ala Phe Leu Ala
 515 520 525
 Gly Asn
 530
 <210> 156

<211> 648
 <212> DNA
 <213> Anaerotruncus colihominis
 <220><221> CDS
 <222> (1)..(648)
 <223> Anaerotruncus colihominis gene encoding TMP phosphatase

[WP_006874980]

<400> 156
 atg atc aaa ggc gcg att ttt gat atg gac ggt acg ctg att gat tcc 48
 Met Ile Lys Gly Ala Ile Phe Asp Met Asp Gly Thr Leu Ile Asp Ser
 1 5 10 15
 atg cct cta tgg gag gac tgc gga cgg gcc ttt tta tcc gcg cgc ggc 96
 Met Pro Leu Trp Glu Asp Cys Gly Arg Ala Phe Leu Ser Ala Arg Gly
 20 25 30
 att act gcg cgt gac gat ctg ggc gaa acg ctc aaa tcc ctg tcg atg 144

 Ile Thr Ala Arg Asp Asp Leu Gly Glu Thr Leu Lys Ser Leu Ser Met
 35 40 45
 gag caa acg gct aat tat ttg cgg gac gca tac ggt att tcc gag aca 192
 Glu Gln Thr Ala Asn Tyr Leu Arg Asp Ala Tyr Gly Ile Ser Glu Thr
 50 55 60
 acc tct gaa atc att gag atg atc aat gga atg gtt act gac gca tat 240
 Thr Ser Glu Ile Ile Glu Met Ile Asn Gly Met Val Thr Asp Ala Tyr

 65 70 75 80
 cag cgc acc atc ccg ctt aaa cgt gac att gcc gcg ttt ctc gag cgc 288
 Gln Arg Thr Ile Pro Leu Lys Arg Asp Ile Ala Ala Phe Leu Glu Arg
 85 90 95
 ctc agg cag gcg gat gtg cgc atg tgt gtc gca acg gca acg gac cgt 336
 Leu Arg Gln Ala Asp Val Arg Met Cys Val Ala Thr Ala Thr Asp Arg
 100 105 110

 cca ctg gtg gag gcg gcg ctt gga cgc ctt gac ctc ctg ccc ttt ttt 384
 Pro Leu Val Glu Ala Ala Leu Gly Arg Leu Asp Leu Leu Pro Phe Phe

115	120	125	
gaa cgg att ttc acc tgt tgc gag gtg ggg gcc ggc aag gac cgc ccc			432
Glu Arg Ile Phe Thr Cys Ser Glu Val Gly Ala Gly Lys Asp Arg Pro			
130	135	140	
gat atc ttt gag cag gcg tgc gcc gcg ctt ggc acg ccg cgc ggc gaa			480
Asp Ile Phe Glu Gln Ala Cys Ala Ala Leu Gly Thr Pro Arg Gly Glu			
145	150	155	160
acc gtc atc ttt gag gat gct ctt tat gcg att gaa aca gct cgg cgc			528
Thr Val Ile Phe Glu Asp Ala Leu Tyr Ala Ile Glu Thr Ala Arg Arg			
165	170	175	
gcc ggg ttc cgc gtt gtc gca atc gcg gac gac gcc tcc gcc ggc gac			576
Ala Gly Phe Arg Val Val Ala Ile Ala Asp Asp Ala Ser Ala Gly Asp			
180	185	190	
gag gcg cgc ata gcc gca ctg tct gag caa tat ata cat aac tat gag			624
Glu Ala Arg Ile Ala Ala Leu Ser Glu Gln Tyr Ile His Asn Tyr Glu			
195	200	205	
gaa tgc gag gta aac agt tta tga			648
Glu Cys Glu Val Asn Ser Leu			
210	215		
<210> 157			
<211> 215			
<212> PRT			
<213> Anaerotruncus colihominis			
<400> 157			
Met Ile Lys Gly Ala Ile Phe Asp Met Asp Gly Thr Leu Ile Asp Ser			
1	5	10	15
Met Pro Leu Trp Glu Asp Cys Gly Arg Ala Phe Leu Ser Ala Arg Gly			
20	25	30	
Ile Thr Ala Arg Asp Asp Leu Gly Glu Thr Leu Lys Ser Leu Ser Met			
35	40	45	
Glu Gln Thr Ala Asn Tyr Leu Arg Asp Ala Tyr Gly Ile Ser Glu Thr			

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

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

180 185 190
 Glu Ala Arg Ile Ala Ala Leu Ser Glu Gln Tyr Ile His Asn Tyr Glu
 195 200 205
 Glu Cys Glu Val Asn Ser Leu
 210 215

<210> 158
 <211> 666
 <212> DNA
 <213> Eubacterium ventriosum
 <220><221> CDS
 <222> (1)..(666)
 <223> Eubacterium ventriosum gene encoding TMP phosphatase
 [WP_005362972]

<400> 158
 atg tca aca gga ttt ata ttt gat gta gat gga aca ata cta gac tca

48

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

1	5	10	15	
atg gga ata tgg atg aac gta gga gaa cta tat cta aaa gat atg gga				96
Met Gly Ile Trp Met Asn Val Gly Glu Leu Tyr Leu Lys Asp Met Gly				
	20	25	30	
ata aag gcg gaa cca aat ctt gga gaa att cta ttc gaa atg aca atg				144
Ile Lys Ala Glu Pro Asn Leu Gly Glu Ile Leu Phe Glu Met Thr Met				
	35	40	45	
aat gaa ggt gca gaa tac ata caa aaa aag tat aat cta aac ctt aca				192
Asn Glu Gly Ala Glu Tyr Ile Gln Lys Lys Tyr Asn Leu Asn Leu Thr				
	50	55	60	
aca gaa gaa ata tgc acc gga ata aac aac cgt gta tac aaa ttc tac				240
Thr Glu Glu Ile Cys Thr Gly Ile Asn Asn Arg Val Tyr Lys Phe Tyr				
65	70	75	80	
gaa aaa gaa gca atg cca aaa cca aaa gtt atc gac ttt ata gaa caa				288
Glu Lys Glu Ala Met Pro Lys Pro Lys Val Ile Asp Phe Ile Glu Gln				
	85	90	95	
gcc tac gag aac aaa atc cca atg aca ata gca acg tca aca gac aga				336
Ala Tyr Glu Asn Lys Ile Pro Met Thr Ile Ala Thr Ser Thr Asp Arg				
	100	105	110	
cca atg ata gaa gca gct ttc aaa aga ctg cac ata gac aaa tat ttt				384
Pro Met Ile Glu Ala Ala Phe Lys Arg Leu His Ile Asp Lys Tyr Phe				
	115	120	125	
aaa aaa ata ttt acc acg aca gag gtt ggg tat gga aaa gac aaa ccg				432
Lys Lys Ile Phe Thr Thr Thr Glu Val Gly Tyr Gly Lys Asp Lys Pro				
	130	135	140	
gac atc ttc ata aaa gca atg gaa gaa atg gga aca aca cca aag caa				480
Asp Ile Phe Ile Lys Ala Met Glu Glu Met Gly Thr Thr Pro Lys Gln				
145	150	155	160	
aca tgg cta ttt gaa gat gga gca tac tca ata gaa aca gcc aaa caa				528
Thr Trp Leu Phe Glu Asp Gly Ala Tyr Ser Ile Glu Thr Ala Lys Gln				
	165	170	175	

cta ggc ata aaa aca ata gga atc tac gat cct gca agc gaa aaa gac 576
 Leu Gly Ile Lys Thr Ile Gly Ile Tyr Asp Pro Ala Ser Glu Lys Asp
 180 185 190

cag gaa aaa ata aga aac cta aca aac atc tac ata aaa aat tgg aca 624
 Gln Glu Lys Ile Arg Asn Leu Thr Asn Ile Tyr Ile Lys Asn Trp Thr
 195 200 205

gaa cac aaa acc cta ctt aac caa ata caa aac aac aag tag 666
 Glu His Lys Thr Leu Leu Asn Gln Ile Gln Asn Asn Lys
 210 215 220

<210> 159

<211> 221

<212> PRT

<213> Eubacterium ventriosum

<400> 159

Met Ser Thr Gly Phe Ile Phe Asp Val Asp Gly Thr Ile Leu Asp Ser
 1 5 10 15

Met Gly Ile Trp Met Asn Val Gly Glu Leu Tyr Leu Lys Asp Met Gly
 20 25 30

Ile Lys Ala Glu Pro Asn Leu Gly Glu Ile Leu Phe Glu Met Thr Met
 35 40 45

Asn Glu Gly Ala Glu Tyr Ile Gln Lys Lys Tyr Asn Leu Asn Leu Thr
 50 55 60

Thr Glu Glu Ile Cys Thr Gly Ile Asn Asn Arg Val Tyr Lys Phe Tyr
 65 70 75 80

Glu Lys Glu Ala Met Pro Lys Pro Lys Val Ile Asp Phe Ile Glu Gln
 85 90 95

Ala Tyr Glu Asn Lys Ile Pro Met Thr Ile Ala Thr Ser Thr Asp Arg
 100 105 110

Pro Met Ile Glu Ala Ala Phe Lys Arg Leu His Ile Asp Lys Tyr Phe
 115 120 125

Lys Lys Ile Phe Thr Thr Thr Glu Val Gly Tyr Gly Lys Asp Lys Pro

130 135 140
 Asp Ile Phe Ile Lys Ala Met Glu Glu Met Gly Thr Thr Pro Lys Gln
 145 150 155 160
 Thr Trp Leu Phe Glu Asp Gly Ala Tyr Ser Ile Glu Thr Ala Lys Gln
 165 170 175
 Leu Gly Ile Lys Thr Ile Gly Ile Tyr Asp Pro Ala Ser Glu Lys Asp
 180 185 190
 Gln Glu Lys Ile Arg Asn Leu Thr Asn Ile Tyr Ile Lys Asn Trp Thr
 195 200 205
 Glu His Lys Thr Leu Leu Asn Gln Ile Gln Asn Asn Lys
 210 215 220
 <210> 160
 <211> 1482
 <212> DNA
 <213> Coprococcus eutactus
 <220><221> CDS
 <222> (1)..(1482)
 <223> Coprococcus eutactus ATCC 27759 gene encoding TMP phosphatase
 [EDP27707]
 <400> 160
 atg aaa aag ata gtt atc agc gat ata aaa ggt gcg ata ttt gac atg 48
 Met Lys Lys Ile Val Ile Ser Asp Ile Lys Gly Ala Ile Phe Asp Met
 1 5 10 15
 gat gga gtt ctg ctg gac tct atg ccg atg tgg gac cat gcg ggc gag 96
 Asp Gly Val Leu Leu Asp Ser Met Pro Met Trp Asp His Ala Gly Glu
 20 25 30
 atg tac ctt gca gga cag ggg ata gag gct gag cct gat ctt gaa aaa 144
 Met Tyr Leu Ala Gly Gln Gly Ile Glu Ala Glu Pro Asp Leu Glu Lys
 35 40 45
 gtc ttg ttt aca atg act atg caa aag ggc gct gaa tat ata cgt gat 192
 Val Leu Phe Thr Met Thr Met Gln Lys Gly Ala Glu Tyr Ile Arg Asp
 50 55 60

cat	tat	ggg	tta	aaa	ctc	acg	gcg	gat	gag	atc	ata	gat	ggc	ata	aat	240
His	Tyr	Gly	Leu	Lys	Leu	Thr	Ala	Asp	Glu	Ile	Ile	Asp	Gly	Ile	Asn	
65			70			75			80							
gag	act	gtg	aga	gat	ttc	tat	gca	aat	aag	gtt	gtg	cct	aag	aat	gga	288
Glu	Thr	Val	Arg	Asp	Phe	Tyr	Ala	Asn	Lys	Val	Val	Pro	Lys	Asn	Gly	
85			90			95										
gtc	ctt	aag	ttc	ctc	agg	ctg	ttg	aag	agt	cac	aat	ata	cct	gta	acc	336
Val	Leu	Lys	Phe	Leu	Arg	Leu	Leu	Lys	Ser	His	Asn	Ile	Pro	Val	Thr	
100			105			110										
gtt	gca	act	tcg	acc	gac	aga	tgc	cat	gtg	gag	gct	gct	ctt	tca	aga	384
Val	Ala	Thr	Ser	Thr	Asp	Arg	Cys	His	Val	Glu	Ala	Ala	Leu	Ser	Arg	
115			120			125										
aat	gga	ctt	atg	gaa	tat	gta	gac	aag	ata	ttt	acg	tgt	tcg	gaa	gtt	432
Asn	Gly	Leu	Met	Glu	Tyr	Val	Asp	Lys	Ile	Phe	Thr	Cys	Ser	Glu	Val	
130			135			140										
ggc	gtt	gga	aag	gct	gcc	tct	cca	aag	ata	tat	gag	ctt	gcg	gcc	gaa	480
Gly	Val	Gly	Lys	Ala	Ala	Ser	Pro	Lys	Ile	Tyr	Glu	Leu	Ala	Ala	Glu	
145			150			155			160							
ttt	atg	ggg	acg	aaa	gtc	ggc	gag	tca	ttt	gtg	ttc	gag	gat	gcc	tat	528
Phe	Met	Gly	Thr	Lys	Val	Gly	Glu	Ser	Phe	Val	Phe	Glu	Asp	Ala	Tyr	
165			170			175										
cat	gcg	gcc	gag	aca	gct	cag	aat	gcg	gga	ttt	aca	gtt	gta	gga	ctc	576
His	Ala	Ala	Glu	Thr	Ala	Gln	Asn	Ala	Gly	Phe	Thr	Val	Val	Gly	Leu	
180			185			190										
tat	gac	gag	tca	agc	cgt	gac	atg	caa	gca	gaa	ctt	aag	gtt	cac	tgc	624
Tyr	Asp	Glu	Ser	Ser	Arg	Asp	Met	Gln	Ala	Glu	Leu	Lys	Val	His	Cys	
195			200			205										
aat	tat	tac	tat	ttg	gga	ttt	gcc	gag	ctt	ata	gat	gag	ctg	ctg	cct	672
Asn	Tyr	Tyr	Tyr	Leu	Gly	Phe	Ala	Glu	Leu	Ile	Asp	Glu	Leu	Leu	Pro	
210			215			220										

gac	aga	agc	cag	ctt	gca	ccg	gtt	ctt	acc	atc	gcg	ggc	agt	gat	tca	720
Asp	Arg	Ser	Gln	Leu	Ala	Pro	Val	Leu	Thr	Ile	Ala	Gly	Ser	Asp	Ser	
225					230					235					240	
tcg	gga	ggt	gcg	gga	ata	cag	gca	gat	ctt	aag	acc	atg	cag	gca	aat	768
Ser	Gly	Gly	Ala	Gly	Ile	Gln	Ala	Asp	Leu	Lys	Thr	Met	Gln	Ala	Asn	
				245					250					255		
gga	gtg	ttt	ggc	atg	agc	gca	gta	act	gcc	ttg	acg	gcg	cag	aat	acc	816
Gly	Val	Phe	Gly	Met	Ser	Ala	Val	Thr	Ala	Leu	Thr	Ala	Gln	Asn	Thr	
				260					265					270		
aca	ggt	gtg	aca	tcc	atc	atg	aat	gtg	aca	cct	gac	ata	ctt	gca	gat	864
Thr	Gly	Val	Thr	Ser	Ile	Met	Asn	Val	Thr	Pro	Asp	Ile	Leu	Ala	Asp	
		275						280					285			
cag	ata	gat	gca	gta	ttt	aca	gat	ata	aga	cca	cag	gcg	gtc	aag	ata	912
Gln	Ile	Asp	Ala	Val	Phe	Thr	Asp	Ile	Arg	Pro	Gln	Ala	Val	Lys	Ile	
	290						295						300			
ggt	atg	gtg	tct	gtg	cca	gaa	ctt	ata	aat	gtg	atc	gca	gac	aag	ctt	960
Gly	Met	Val	Ser	Val	Pro	Glu	Leu	Ile	Asn	Val	Ile	Ala	Asp	Lys	Leu	
305					310					315					320	
gaa	ttt	tac	agg	gcg	gag	aat	gtg	gtg	ctt	gat	cct	gtg	atg	gtt	gcg	1008
Glu	Phe	Tyr	Arg	Ala	Glu	Asn	Val	Val	Leu	Asp	Pro	Val	Met	Val	Ala	
				325						330				335		
aca	agc	ggt	gct	aaa	ctc	ata	agc	gat	gat	gct	gtg	gac	gtt	ttg	aca	1056
Thr	Ser	Gly	Ala	Lys	Leu	Ile	Ser	Asp	Asp	Ala	Val	Asp	Val	Leu	Thr	
				340					345					350		
gga	agg	ctg	ttc	cca	ctt	gca	aag	ctg	atc	acc	cca	aat	att	cca	gag	1104
Gly	Arg	Leu	Phe	Pro	Leu	Ala	Lys	Leu	Ile	Thr	Pro	Asn	Ile	Pro	Glu	
		355						360					365			
aca	gag	gcc	ctc	aca	ggt	atg	agt	atc	cgg	tct	aag	gaa	gat	atg	gaa	1152
Thr	Glu	Ala	Leu	Thr	Gly	Met	Ser	Ile	Arg	Ser	Lys	Glu	Asp	Met	Glu	

370 375 380
 agt gca gca agg aaa ata tat gaa aaa tat ggc tgc tca gtt ctt gtg 1200
 Ser Ala Ala Arg Lys Ile Tyr Glu Lys Tyr Gly Cys Ser Val Leu Val
 385 390 395 400
 aag ggc gga cat agc ata aac gat gcg aat gat atg ctg ttt gat gga 1248
 Lys Gly Gly His Ser Ile Asn Asp Ala Asn Asp Met Leu Phe Asp Gly
 405 410 415
 gag aat gta tca tgg ttt tca ggt gag aga ata gaa aat ccg aat acc 1296
 Glu Asn Val Ser Trp Phe Ser Gly Glu Arg Ile Glu Asn Pro Asn Thr
 420 425 430
 cat gga acg ggg tgt aca ctc tca agt gca ata gcc tcc aac ctt gca 1344
 His Gly Thr Gly Cys Thr Leu Ser Ser Ala Ile Ala Ser Asn Leu Ala
 435 440 445
 aag gga tat gat ata gaa act tct gtg cag aga gca aaa gcg tac atc 1392
 Lys Gly Tyr Asp Ile Glu Thr Ser Val Gln Arg Ala Lys Ala Tyr Ile
 450 455 460
 tca gga gcc ctg gct gcg atg ctt gat cta gga aga gga agc ggc ccg 1440
 Ser Gly Ala Leu Ala Ala Met Leu Asp Leu Gly Arg Gly Ser Gly Pro
 465 470 475 480
 tta aac cat ggc ttt gat ata gac agc aga ttc atg ata taa 1482
 Leu Asn His Gly Phe Asp Ile Asp Ser Arg Phe Met Ile

485 490
 <210> 161
 <211> 493
 <212> PRT
 <213> Coprococcus eutactus
 <400> 161
 Met Lys Lys Ile Val Ile Ser Asp Ile Lys Gly Ala Ile Phe Asp Met
 1 5 10 15
 Asp Gly Val Leu Leu Asp Ser Met Pro Met Trp Asp His Ala Gly Glu
 20 25 30
 Met Tyr Leu Ala Gly Gln Gly Ile Glu Ala Glu Pro Asp Leu Glu Lys

35 40 45
 Val Leu Phe Thr Met Thr Met Gln Lys Gly Ala Glu Tyr Ile Arg Asp
 50 55 60
 His Tyr Gly Leu Lys Leu Thr Ala Asp Glu Ile Ile Asp Gly Ile Asn
 65 70 75 80
 Glu Thr Val Arg Asp Phe Tyr Ala Asn Lys Val Val Pro Lys Asn Gly
 85 90 95
 Val Leu Lys Phe Leu Arg Leu Leu Lys Ser His Asn Ile Pro Val Thr
 100 105 110

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

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

 Ser Gly Gly Ala Gly Ile Gln Ala Asp Leu Lys Thr Met Gln Ala Asn
 245 250 255
 Gly Val Phe Gly Met Ser Ala Val Thr Ala Leu Thr Ala Gln Asn Thr
 260 265 270
 Thr Gly Val Thr Ser Ile Met Asn Val Thr Pro Asp Ile Leu Ala Asp
 275 280 285

Gln Ile Asp Ala Val Phe Thr Asp Ile Arg Pro Gln Ala Val Lys Ile
290 295 300

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

Glu Phe Tyr Arg Ala Glu Asn Val Val Leu Asp Pro Val Met Val Ala
325 330 335

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

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

Thr Glu Ala Leu Thr Gly Met Ser Ile Arg Ser Lys Glu Asp Met Glu
370 375 380

Ser Ala Ala Arg Lys Ile Tyr Glu Lys Tyr Gly Cys Ser Val Leu Val
385 390 395 400

Lys Gly Gly His Ser Ile Asn Asp Ala Asn Asp Met Leu Phe Asp Gly
405 410 415

Glu Asn Val Ser Trp Phe Ser Gly Glu Arg Ile Glu Asn Pro Asn Thr
420 425 430

His Gly Thr Gly Cys Thr Leu Ser Ser Ala Ile Ala Ser Asn Leu Ala
435 440 445

Lys Gly Tyr Asp Ile Glu Thr Ser Val Gln Arg Ala Lys Ala Tyr Ile
450 455 460

Ser Gly Ala Leu Ala Ala Met Leu Asp Leu Gly Arg Gly Ser Gly Pro
465 470 475 480

Leu Asn His Gly Phe Asp Ile Asp Ser Arg Phe Met Ile
485 490

<210> 162

<211> 663

<212> DNA

<213> Ruminococcus bromii

<220><221> CDS

<222> (1)..(663)

<223> Ruminococcus bromii L2-63 gene encoding TMP phosphatase
[CBL14666]

<400> 162

atg att aaa tct gca ata ttt gat gtt gac ggc aca ctt ctc gat tca 48

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

1 5 10 15

atg aag ata tgg gat gat gca gga gag cgt tac ctc tcg tct gtc ggc 96

Met Lys Ile Trp Asp Asp Ala Gly Glu Arg Tyr Leu Ser Ser Val Gly

20 25 30

aaa aca gcc gaa aac gga ctt tcc gaa aag ctc tgt gat atg agt ctg 144

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

35 40 45

acg gag ggt gcg gag tat atg aaa aag cag tat gct ctt tcc ttt tca 192

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

50 55 60

act gat gaa ata gtt tcg ggt gtg ctg aaa atc att gaa gat ttt tac 240

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

65 70 75 80

ttt tat gag gtc ggt tta aaa aac gat gca aaa gaa att ttg cag ttt 288

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

85 90 95

ttg gaa tcg aac aat atc aaa atg att att gca aca tca agc gac aaa 336

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

100 105 110

acg cat att aaa aag gca ttt gaa agg ctc ggt att cta aaa tat ttt 384

Thr His Ile Lys Lys Ala Phe Glu Arg Leu Gly Ile Leu Lys Tyr Phe

115 120 125

acg gat att gtg acc tgt tca caa gtc gga aaa ggc aaa aca agc ccc 432

Thr Asp Ile Val Thr Cys Ser Gln Val Gly Lys Gly Lys Thr Ser Pro

130 135 140

gac att tac ctt gtc tgt gca gat aaa ctc gga aca gct ccg agt gaa 480
 Asp Ile Tyr Leu Val Cys Ala Asp Lys Leu Gly Thr Ala Pro Ser Glu
 145 150 155 160
 acg ctt gta ttc gag gac gct gtt ttt gcc gca gaa act gct cac aag 528
 Thr Leu Val Phe Glu Asp Ala Val Phe Ala Ala Glu Thr Ala His Lys
 165 170 175
 gca ggt ttc aaa acg gtg gga gtg tat gac gaa ttg agc agg aat aat 576
 Ala Gly Phe Lys Thr Val Gly Val Tyr Asp Glu Leu Ser Arg Asn Asn
 180 185 190
 aaa aac aga ata aaa gcc gtt tgc gat tac tac gca gac agc ttt gaa 624
 Lys Asn Arg Ile Lys Ala Val Cys Asp Tyr Tyr Ala Asp Ser Phe Glu
 195 200 205
 aaa gcg gca gat tgg ggg cac cac ctt ttg tcg ctg taa 663

Lys Ala Ala Asp Trp Gly His His Leu Leu Ser Leu
 210 215 220

<210> 163

<211> 220

<212> PRT

<213> Ruminococcus bromii

<400> 163

Met Ile Lys Ser Ala Ile Phe Asp Val Asp Gly Thr Leu Leu Asp Ser
 1 5 10 15

Met Lys Ile Trp Asp Asp Ala Gly Glu Arg Tyr Leu Ser Ser Val Gly
 20 25 30

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

35 40 45

Thr Glu Gly Ala Glu Tyr Met Lys Lys Gln Tyr Ala Leu Ser Phe Ser
 50 55 60

Thr Asp Glu Ile Val Ser Gly Val Leu Lys Ile Ile Glu Asp Phe Tyr
 65 70 75 80

Phe Tyr Glu Val Gly Leu Lys Asn Asp Ala Lys Glu Ile Leu Gln Phe
 85 90 95

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

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

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

<210> 164

<211> 1434

<212> DNA

<213> Dorea longicatena

<220><221> CDS

<222> (1)..(1434)

<223> Dorea longicatena DSM13814 gene encoding TMP phosphatase
 [EDM62146]

<400> 164

atg ata aaa gga gca ata ttt gat gta gac gga acc ctt ctg gat tcc 48
 Met Ile Lys Gly Ala Ile Phe Asp Val Asp Gly Thr Leu Leu Asp Ser
 1 5 10 15
 atg gag atc tgg gaa gac gta gga gtc cgt tat ctg aac agt atc ggt 96
 Met Glu Ile Trp Glu Asp Val Gly Val Arg Tyr Leu Asn Ser Ile Gly
 20 25 30

ata gag gca gag ccg gat ctt ggg acg gtg tta ttt aca atg agc atc	144
Ile Glu Ala Glu Pro Asp Leu Gly Thr Val Leu Phe Thr Met Ser Ile	
35 40 45	
cag gaa ggt gca gca tat gta aaa gaa cat tat cat ctg tcc cag gag	192
Gln Glu Gly Ala Ala Tyr Val Lys Glu His Tyr His Leu Ser Gln Glu	
50 55 60	
ccg gaa gaa att gtg cag gga gtt ctg gac atc atc agc aat tat tat	240
Pro Glu Glu Ile Val Gln Gly Val Leu Asp Ile Ile Ser Asn Tyr Tyr	
65 70 75 80	
aag aaa acc gca cta tta aag agt gga gtg aag gaa ctt ctg gaa aag	288
Lys Lys Thr Ala Leu Leu Lys Ser Gly Val Lys Glu Leu Leu Glu Lys	
85 90 95	
ctt gat aag cat aat atc cca atg acg gtt gca tca tcc aat aat aaa	336
Leu Asp Lys His Asn Ile Pro Met Thr Val Ala Ser Ser Asn Asn Lys	
100 105 110	
aaa gag ata gag atg gca ttt gag cgt ctg gga att gca aaa tat ttt	384
Lys Glu Ile Glu Met Ala Phe Glu Arg Leu Gly Ile Ala Lys Tyr Phe	
115 120 125	
gac cgg atc ttt acc tgt gaa gag gtc ggt gcg gga aag acg aag ccg	432
Asp Arg Ile Phe Thr Cys Glu Glu Val Gly Ala Gly Lys Thr Lys Pro	
130 135 140	
gat att tat ctg cgg gca gca gaa tat ctc gga acc cgt ccg gag gag	480
Asp Ile Tyr Leu Arg Ala Ala Glu Tyr Leu Gly Thr Arg Pro Glu Glu	
145 150 155 160	
acg gtt gta ttc gaa gat gtc att cat gca atc cgt act gca aag cag	528
Thr Val Val Phe Glu Asp Val Ile His Ala Ile Arg Thr Ala Lys Gln	
165 170 175	
gca ggg ttc cag gtt gta gga atc tat gat gaa gca agt aag gat gac	576
Ala Gly Phe Gln Val Val Gly Ile Tyr Asp Glu Ala Ser Lys Asp Asp	
180 185 190	
cag gaa gag gtt cag aga gaa gta gac tgg tat tgt aga gag tgg gca	624

Gln Glu Glu Val Gln Arg Glu Val Asp Trp Tyr Cys Arg Glu Trp Ala
195 200 205
gaa ctt atg aaa aaa aag aca gca att aca atc gcc gga agt gat tca 672
Glu Leu Met Lys Lys Lys Thr Ala Ile Thr Ile Ala Gly Ser Asp Ser
210 215 220
agt gga ggt gca gga att cag gca gac atc aag acg atg cag gca aac 720
Ser Gly Gly Ala Gly Ile Gln Ala Asp Ile Lys Thr Met Gln Ala Asn
225 230 235 240
gga gtc tac gca atg agt gca atc acc gca ctg aca gcc cag aat aca 768
Gly Val Tyr Ala Met Ser Ala Ile Thr Ala Leu Thr Ala Gln Asn Thr
245 250 255
acc gga gta acc gga atc atg gaa gta tct ccg gaa ttt cta gaa caa 816
Thr Gly Val Thr Gly Ile Met Glu Val Ser Pro Glu Phe Leu Glu Gln
260 265 270
cag ttg gac gca gtt atc aca gac atc cgt ccg gat gca gtg aaa atc 864
Gln Leu Asp Ala Val Ile Thr Asp Ile Arg Pro Asp Ala Val Lys Ile
275 280 285
ggt atg gtg tca tca gaa gag tta ata aaa atg ata tca aag aaa cta 912
Gly Met Val Ser Ser Glu Glu Leu Ile Lys Met Ile Ser Lys Lys Leu
290 295 300
aaa gag tac cat ctg gag aat atc gta gtt gat cca gtg atg gta gca 960
Lys Glu Tyr His Leu Glu Asn Ile Val Val Asp Pro Val Met Val Ala
305 310 315 320
aca agc gga tcc aga ctg atc agt gaa acc gcg att gat aca tta aaa 1008
Thr Ser Gly Ser Arg Leu Ile Ser Glu Thr Ala Ile Asp Thr Leu Lys
325 330 335
aca cag ctg ctg cca atg gca act gtg atc aca ccg aat atc cca gag 1056
Thr Gln Leu Leu Pro Met Ala Thr Val Ile Thr Pro Asn Ile Pro Glu
340 345 350
gca gaa gtt ctt gca gaa atg gag att aga tca gaa gat gat atg gtg 1104
Ala Glu Val Leu Ala Glu Met Glu Ile Arg Ser Glu Asp Asp Met Val

355	360	365														
gaa gca gca aag aag att cat gaa atg tat cac tgt gca gtc tta tgc																1152
Glu Ala Ala Lys Lys Ile His Glu Met Tyr His Cys Ala Val Leu Cys																
370				375				380								
aaa ggc gga cac agc ctg aat gat gcg aat gat ctc cta tac cag gat																1200
Lys Gly Gly His Ser Leu Asn Asp Ala Asn Asp Leu Leu Tyr Gln Asp																
385				390				395				400				
gga gaa aca aca tgg ttc cac gga aaa aga atc aac aac ccg aac act																1248
Gly Glu Thr Thr Trp Phe His Gly Lys Arg Ile Asn Asn Pro Asn Thr																
				405				410				415				
cac gga acc ggc tgt acc tta tcc agc gca atc gca tcc aat ctg gca																1296
His Gly Thr Gly Cys Thr Leu Ser Ser Ala Ile Ala Ser Asn Leu Ala																
				420				425				430				
aaa gga tat tct ctg gaa gaa tct att cac cgc gcg aaa gag tat atc																1344
Lys Gly Tyr Ser Leu Glu Glu Ser Ile His Arg Ala Lys Glu Tyr Ile																
				435				440				445				
agc ggg gcg ttg gaa gcc atg tta gat ctg gga aaa gga agc gga ccg																1392
Ser Gly Ala Leu Glu Ala Met Leu Asp Leu Gly Lys Gly Ser Gly Pro																
450				455				460								
atg gat cat ggg ttt gag atg cgg ggg aga ttt tct att taa																1434
Met Asp His Gly Phe Glu Met Arg Gly Arg Phe Ser Ile																
465				470				475								
<210> 165																
<211> 477																
<212> PRT																
<213> Dorea longicatena																
<400> 165																
Met Ile Lys Gly Ala Ile Phe Asp Val Asp Gly Thr Leu Leu Asp Ser																
1				5				10				15				
Met Glu Ile Trp Glu Asp Val Gly Val Arg Tyr Leu Asn Ser Ile Gly																
20				25				30								

Ile Glu Ala Glu Pro Asp Leu Gly Thr Val Leu Phe Thr Met Ser Ile
35 40 45
Gln Glu Gly Ala Ala Tyr Val Lys Glu His Tyr His Leu Ser Gln Glu
50 55 60
Pro Glu Glu Ile Val Gln Gly Val Leu Asp Ile Ile Ser Asn Tyr Tyr
65 70 75 80
Lys Lys Thr Ala Leu Leu Lys Ser Gly Val Lys Glu Leu Leu Glu Lys
85 90 95
Leu Asp Lys His Asn Ile Pro Met Thr Val Ala Ser Ser Asn Asn Lys
100 105 110
Lys Glu Ile Glu Met Ala Phe Glu Arg Leu Gly Ile Ala Lys Tyr Phe
115 120 125
Asp Arg Ile Phe Thr Cys Glu Glu Val Gly Ala Gly Lys Thr Lys Pro
130 135 140
Asp Ile Tyr Leu Arg Ala Ala Glu Tyr Leu Gly Thr Arg Pro Glu Glu
145 150 155 160
Thr Val Val Phe Glu Asp Val Ile His Ala Ile Arg Thr Ala Lys Gln
165 170 175
Ala Gly Phe Gln Val Val Gly Ile Tyr Asp Glu Ala Ser Lys Asp Asp
180 185 190
Gln Glu Glu Val Gln Arg Glu Val Asp Trp Tyr Cys Arg Glu Trp Ala
195 200 205
Glu Leu Met Lys Lys Lys Thr Ala Ile Thr Ile Ala Gly Ser Asp Ser
210 215 220
Ser Gly Gly Ala Gly Ile Gln Ala Asp Ile Lys Thr Met Gln Ala Asn
225 230 235 240
Gly Val Tyr Ala Met Ser Ala Ile Thr Ala Leu Thr Ala Gln Asn Thr
245 250 255
Thr Gly Val Thr Gly Ile Met Glu Val Ser Pro Glu Phe Leu Glu Gln
260 265 270
Gln Leu Asp Ala Val Ile Thr Asp Ile Arg Pro Asp Ala Val Lys Ile

275 280 285
Gly Met Val Ser Ser Glu Glu Leu Ile Lys Met Ile Ser Lys Lys Leu

290 295 300
Lys Glu Tyr His Leu Glu Asn Ile Val Val Asp Pro Val Met Val Ala

305 310 315 320
Thr Ser Gly Ser Arg Leu Ile Ser Glu Thr Ala Ile Asp Thr Leu Lys

325 330 335
Thr Gln Leu Leu Pro Met Ala Thr Val Ile Thr Pro Asn Ile Pro Glu

340 345 350
Ala Glu Val Leu Ala Glu Met Glu Ile Arg Ser Glu Asp Asp Met Val

355 360 365
Glu Ala Ala Lys Lys Ile His Glu Met Tyr His Cys Ala Val Leu Cys

370 375 380
Lys Gly Gly His Ser Leu Asn Asp Ala Asn Asp Leu Leu Tyr Gln Asp

385 390 395 400
Gly Glu Thr Thr Trp Phe His Gly Lys Arg Ile Asn Asn Pro Asn Thr

405 410 415
His Gly Thr Gly Cys Thr Leu Ser Ser Ala Ile Ala Ser Asn Leu Ala

420 425 430
Lys Gly Tyr Ser Leu Glu Glu Ser Ile His Arg Ala Lys Glu Tyr Ile

435 440 445
Ser Gly Ala Leu Glu Ala Met Leu Asp Leu Gly Lys Gly Ser Gly Pro

450 455 460
Met Asp His Gly Phe Glu Met Arg Gly Arg Phe Ser Ile

465 470 475

<210> 166

<211> 1305

<212> DNA

<213> Lachnospiraceae bacterium

<220><221> CDS

<222> (1)..(1305)

<223> Lachnospiraceae_bacterium_3_1_57FAA_CT1 gene encoding TMP

phosphatase [EPC05128]

<400> 166

atg aaa tgt gac aga aag aca atg ctt ctt tat gcg gtg acc gat cgg 48

Met Lys Cys Asp Arg Lys Thr Met Leu Leu Tyr Ala Val Thr Asp Arg

1 5 10 15

gcc tgg aca gga gaa aag aca ctg ctt atg cag gtc gag gaa gcg ctg 96

Ala Trp Thr Gly Glu Lys Thr Leu Leu Met Gln Val Glu Glu Ala Leu

20 25 30

gca gga ggt gtg acc tgt gtc cag ctt cgt gaa aag gat atg cca aag 144

Ala Gly Gly Val Thr Cys Val Gln Leu Arg Glu Lys Asp Met Pro Lys

35 40 45

gag cag ttc ctg gaa gaa gcg gag agt ata aaa aga ctt tgt cat aaa 192

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

50 55 60

tat ggg atc cct ttt ata att gac gat gat gtg gag ctg gcc gta cgc 240

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

65 70 75 80

tgc ggc gcg gac ggg gtg cat gtg gga cag cat gat atg gag gca ggc 288

Cys Gly Ala Asp Gly Val His Val Gly Gln His Asp Met Glu Ala Gly

85 90 95

gcg gtc cgc cgg aaa atc gga gac ggc atg ctg ctg ggc gta tca gtc 336

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

100 105 110

cag act gtg gaa cag gca gtg gaa gcc gaa aaa aag gga gcg gat tac 384

Gln Thr Val Glu Gln Ala Val Glu Ala Glu Lys Lys Gly Ala Asp Tyr

115 120 125

ctt ggt gtg ggc gct gtg ttt tcc act tcc acg aaa acg gac gca cag 432

Leu Gly Val Gly Ala Val Phe Ser Thr Ser Thr Lys Thr Asp Ala Gln

130 135 140

gag gtt tcc ctg gat acc ctc cgg gaa atc tgc cgg gcg gtg tcc gta 480

Glu Val Ser Leu Asp Thr Leu Arg Glu Ile Cys Arg Ala Val Ser Val
145 150 155 160

ccc gtc tgt gca atc gga ggg ata cac aaa gga aat atg cat ttg ctg 528

Pro Val Cys Ala Ile Gly Gly Ile His Lys Gly Asn Met His Leu Leu
165 170 175

cag gat acg gga atc gat ggg gtg gct ttg gtg tgc gcc atc ttt tcc 576

Gln Asp Thr Gly Ile Asp Gly Val Ala Leu Val Ser Ala Ile Phe Ser
180 185 190

agt ccc tgc ata cag aag gaa tgc agg gag ctg cgg gtc ctg gca gag 624

Ser Pro Cys Ile Gln Lys Glu Cys Arg Glu Leu Arg Val Leu Ala Glu
195 200 205

aga ctg aaa agg aaa ggg gct att ttt gat gcg gac gga acc ctg ctg 672

Arg Leu Lys Arg Lys Gly Ala Ile Phe Asp Ala Asp Gly Thr Leu Leu
210 215 220

gat tcc atg tcc gtt tgg gat act ctg ggt gaa aaa tat ctg cgg aaa 720

Asp Ser Met Ser Val Trp Asp Thr Leu Gly Glu Lys Tyr Leu Arg Lys
225 230 235 240

aag ggt att gtt ccg gaa aag aac atc agg gaa aca ata aaa aat atg 768

Lys Gly Ile Val Pro Glu Lys Asn Ile Arg Glu Thr Ile Lys Asn Met
245 250 255

agt ctt cct cag gct gcg gtc tat ttt cag act gct tat ggg att gcg 816

Ser Leu Pro Gln Ala Ala Val Tyr Phe Gln Thr Ala Tyr Gly Ile Ala
260 265 270

gat gca gaa gac aag att ata gag gat att aat gga ata gcg gcg tcc 864

Asp Ala Glu Asp Lys Ile Ile Glu Asp Ile Asn Gly Ile Ala Ala Ser
275 280 285

ttt tac atc aat gag gtg aag ctg aag gaa ggc gtg aaa acg gtt ctg 912

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

290	295	300	
gac aag ctg aag cag aaa aac gta aag atg tgt gtg gcg acg gct acg			960
Asp Lys Leu Lys Gln Lys Asn Val Lys Met Cys Val Ala Thr Ala Thr			
305	310	315	320
gac aaa ggg ctg att gaa aag gca ctt gag aga aac gga atc aga gat			1008
Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp			
	325	330	335
tat ttt gag gct gtc ctc acc tgc acg gat gtg ggc gcg gga aag gat			1056
Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp			
	340	345	350
gag ccg gtt atc ttc cgt aag gcc ggg cag ctt ctc gga aca gca aaa			1104
Glu Pro Val Ile Phe Arg Lys Ala Gly Gln Leu Leu Gly Thr Ala Lys			
	355	360	365
gag gat acc att gta att gaa gat gcc ttg tat gct gtt aag aca gcg			1152
Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala			
	370	375	380
aaa gag gac ggt ttc ctg gtg gcg gct gtt tat gat ccg tca gca gaa			1200
Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu			
385	390	395	400
aag gag gaa ccg gag atc cgg gag atc tct gac ttc tat ttc cgg tca			1248
Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser			
	405	410	415
ttt aat gaa atg gag agt tat ctg aat gaa aaa agt tct tac gat agc			1296
Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser			
	420	425	430
ggg ctc tga			1305
Gly Leu			

<210> 167

<211> 434

<212> PRT

<213> Lachnospiraceae bacterium

<400> 167

```

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

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

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

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

```


Lys Gly Ile Val Pro Glu Lys Asn Ile Arg Glu Thr Ile Lys Asn Met

245 250 255

Ser Leu Pro Gln Ala Ala Val Tyr Phe Gln Thr Ala Tyr Gly Ile Ala

260 265 270

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

275 280 285

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

290 295 300

Asp Lys Leu Lys Gln Lys Asn Val Lys Met Cys Val Ala Thr Ala Thr

305 310 315 320

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

325 330 335

Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp

340 345 350

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

355 360 365

Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala

370 375 380

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

385 390 395 400

Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser

405 410 415

Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser

420 425 430

Gly Leu

<210> 168

<211> 1305

<212> DNA

<213> Fusicatenibacter

<220><221> CDS

<222> (1)..(1305)

<223> Fusicatenibacter gene encoding TMP phosphatase [CUQ30753]

<400> 168

atg aaa tgt aac aga aag aca atg ctt ctt tat gcg gtg acc gac cgg 48

Met Lys Cys Asn Arg Lys Thr Met Leu Leu Tyr Ala Val Thr Asp Arg

1 5 10 15

gcc tgg aca gga gaa aag aca ctg ctt acg cag gtc gag gaa gcg ctg 96

Ala Trp Thr Gly Glu Lys Thr Leu Leu Thr Gln Val Glu Glu Ala Leu

20 25 30

gca gga ggt gta acc tgt gtc cag ctt cgt gaa aag gat atg cca aag 144

Ala Gly Gly Val Thr Cys Val Gln Leu Arg Glu Lys Asp Met Pro Lys

35 40 45

gag cag ttc ctg gaa gaa gcg gag agt ata aaa aga ctt tgc cat aaa 192

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

50 55 60

tat ggt gtc cct ttt ata att gac gat gat gtg gag ctg gcc gta cgc 240

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

65 70 75 80

tgc ggt gcg gac ggg gta cat gtg gga cag cat gat atg gag gca ggc 288

Cys Gly Ala Asp Gly Val His Val Gly Gln His Asp Met Glu Ala Gly

85 90 95

gcg gtc cgc cgg aaa atc gga gac ggc atg ctg ctg ggc gta tca gtc 336

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

100 105 110

cag act gtg gaa cag gca gtg gaa gcc gag aaa aag gga gcg gat tac 384

Gln Thr Val Glu Gln Ala Val Glu Ala Glu Lys Lys Gly Ala Asp Tyr

115 120 125

ctt ggt gtg ggc gct gtg ttt tcc act tcc acg aaa acg gac gca cag 432

Leu Gly Val Gly Ala Val Phe Ser Thr Ser Thr Lys Thr Asp Ala Gln

130 135 140

gag gtt tcc ctg gat acc ctc cgg gaa atc tgc cgg gcg gtg tcc gta 480

305 310 315 320
gac aag ggg ctg att gaa aag gca ctt gag aga aac gga atc aga gat 1008

Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp
325 330 335
tat ttt gag gct gtc ctc acc tgc acg gat gtg ggc gcg gga aag gat 1056

Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp
340 345 350

gag ccg gtt atc ttc cgt aag gcc ggg cag ctt ctc gga acc gca aaa 1104
Glu Pro Val Ile Phe Arg Lys Ala Gly Gln Leu Leu Gly Thr Ala Lys

355 360 365
gag gat acc att gta att gaa gat gcc ttg tat gct gtt aag aca gcg 1152
Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala

370 375 380
aaa gag gac ggt ttc ctg gtg gcg gct gtt tat gat ccg tca gca gaa 1200
Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu

385 390 395 400

aag gag gaa ccg gag atc cgg gag atc tct gac ttc tat ttc cgg tca 1248
Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser

405 410 415
ttt aat gaa atg gag agt tat ctg aat gaa aaa agt tct tac gat agc 1296
Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser

420 425 430
ggg ctc tga 1305

Gly Leu

<210> 169

<211> 434

<212> PRT

<213> Fusicatenibacter

<400> 169

Met Lys Cys Asn Arg Lys Thr Met Leu Leu Tyr Ala Val Thr Asp Arg

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

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

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

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

 225 230 235 240
 Lys Gly Ile Val Pro Glu Lys Asn Ile Arg Glu Thr Ile Lys Asn Met
 245 250 255

Ser Leu Pro Gln Ala Ala Val Tyr Phe Gln Thr Ala Tyr Gly Ile Thr
260 265 270
Asp Ala Glu Asp Lys Ile Ile Glu Asp Ile Asn Gly Ile Ala Ala Ser
275 280 285
Phe Tyr Ile Asn Glu Val Lys Leu Lys Glu Gly Val Lys Thr Val Leu
290 295 300
Asp Lys Leu Lys Gln Lys Asn Val Lys Met Cys Val Ala Thr Ala Thr
305 310 315 320
Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp
325 330 335
Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp
340 345 350
Glu Pro Val Ile Phe Arg Lys Ala Gly Gln Leu Leu Gly Thr Ala Lys
355 360 365
Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala
370 375 380
Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu
385 390 395 400
Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser
405 410 415
Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser
420 425 430
Gly Leu

<210> 170

<211> 1296

<212> DNA

<213> Clostridium species

<220><221> CDS

<222> (1)..(1296)

<223> Clostridium sp KLE1755 gene encoding TMP phosphatase [ERI68966]:

<400> 170

atg aaa tgt gac aga agc atg ctg ctc ctc tat gcc gta acc gac cgt	48
Met Lys Cys Asp Arg Ser Met Leu Leu Leu Tyr Ala Val Thr Asp Arg	
1 5 10 15	
gcc tgg acg ggt aaa aaa aca ctg ctg cag cag gtg gag gaa gcc ctg	96
Ala Trp Thr Gly Lys Lys Thr Leu Leu Gln Gln Val Glu Glu Ala Leu	
20 25 30	
gca ggc ggc gcc acc tgc atc cag ctt cgg gaa aag gag ctg ccg gag	144
Ala Gly Gly Ala Thr Cys Ile Gln Leu Arg Glu Lys Glu Leu Pro Glu	
35 40 45	
gaa gaa ttc cgg cag gaa gcc ctg gct gtg aaa gaa ctt tgc cgc aga	192
Glu Glu Phe Arg Gln Glu Ala Leu Ala Val Lys Glu Leu Cys Arg Arg	
50 55 60	
tac cat gtc cct ttc ctc att aac gac aac gta gag ctg gct gtc agc	240
Tyr His Val Pro Phe Leu Ile Asn Asp Asn Val Glu Leu Ala Val Ser	
65 70 75 80	
tgc ggc gcg gac ggc gtc cat gtg ggc cag cac gac atg tct gcg gcg	288
Cys Gly Ala Asp Gly Val His Val Gly Gln His Asp Met Ser Ala Ala	
85 90 95	
gat gtg cgc cgc aga atc ggc ccc ggc aaa ata ctg gga gta tcc gcg	336
Asp Val Arg Arg Arg Ile Gly Pro Gly Lys Ile Leu Gly Val Ser Ala	
100 105 110	
cag acg gtg gag cag gcc cgc cag gcg gaa gaa gac ggc gca gat tat	384
Gln Thr Val Glu Gln Ala Arg Gln Ala Glu Glu Asp Gly Ala Asp Tyr	
115 120 125	
ctg ggc gtg ggc gct gtt ttt tcc acc tcc acc aaa tcc gac gca gac	432
Leu Gly Val Gly Ala Val Phe Ser Thr Ser Thr Lys Ser Asp Ala Asp	
130 135 140	
gcg gta tcc cat gag acc ctg caa aag atc tgc gcc gca gta tcc atc	480
Ala Val Ser His Glu Thr Leu Gln Lys Ile Cys Ala Ala Val Ser Ile	
145 150 155 160	

ccc gtc tgc gcc ata ggc ggc atc cat aaa gaa aat ctg cat ttg ctc 528

Pro Val Cys Ala Ile Gly Gly Ile His Lys Glu Asn Leu His Leu Leu
165 170 175

aaa ggc aca ggc atc gcc ggc gtg gcc ctt gtt tcc gcc atc ttc gca 576
Lys Gly Thr Gly Ile Ala Gly Val Ala Leu Val Ser Ala Ile Phe Ala
180 185 190

agc ccg gat atc cgt aag tcc tgc gaa gac ctg aaa aaa ctg gcc ctg 624
Ser Pro Asp Ile Arg Lys Ser Cys Glu Asp Leu Lys Lys Leu Ala Leu
195 200 205

cag ata aac gcg cag gac aca ctg gaa gca ctg ctg cat aca aac atc 672
Gln Ile Asn Ala Gln Asp Thr Leu Glu Ala Leu Leu His Thr Asn Ile
210 215 220

cgc gga gcc atc ttt gac gcg gac ggc acc ctt tta gac tcc atg ggc 720
Arg Gly Ala Ile Phe Asp Ala Asp Gly Thr Leu Leu Asp Ser Met Gly
225 230 235 240

atc tgg gat act ctg ggg gaa gat tac ctg cgt aca aaa ggg aaa atc 768
Ile Trp Asp Thr Leu Gly Glu Asp Tyr Leu Arg Thr Lys Gly Lys Ile
245 250 255

ccc cgg gaa aac ctg cgt gaa acc ttc cgc gac atg agc ctt ctc cag 816
Pro Arg Glu Asn Leu Arg Glu Thr Phe Arg Asp Met Ser Leu Leu Gln
260 265 270

gcc gcc tgc tat tac cgg gaa aat tac gcc ctt acg gaa agc cct gaa 864

Ala Ala Cys Tyr Tyr Arg Glu Asn Tyr Ala Leu Thr Glu Ser Pro Glu
275 280 285

aaa ata gtg gaa gag ctt aac gcc atg atc gcc tcc ttc tat gaa aaa 912
Lys Ile Val Glu Glu Leu Asn Ala Met Ile Ala Ser Phe Tyr Glu Lys
290 295 300

gaa gcc ccc ctg aag gaa gga gcc gcc gcc ttc ctg gaa gcg ctt tgc 960
Glu Ala Pro Leu Lys Glu Gly Ala Ala Ala Phe Leu Glu Ala Leu Cys

305 310 315 320
 caa aga aac ata aaa atg tgc att gca aca gcc acc gat cac agc ctt 1008
 Gln Arg Asn Ile Lys Met Cys Ile Ala Thr Ala Thr Asp His Ser Leu
 325 330 335
 atc cgg gcc gcc ctg aag cga tgc gga gtg ctg cat tac ttt act ttt 1056
 Ile Arg Ala Ala Leu Lys Arg Cys Gly Val Leu His Tyr Phe Thr Phe
 340 345 350

 ata ctt acc tgc gga caa gca gga gcg gga aaa gac acc ccc gcc att 1104
 Ile Leu Thr Cys Gly Gln Ala Gly Ala Gly Lys Asp Thr Pro Ala Ile
 355 360 365
 tat gaa gaa gcc ctg gcc ctg ctt gga acc gga aaa aaa gaa acc ttc 1152
 Tyr Glu Glu Ala Leu Ala Leu Leu Gly Thr Gly Lys Lys Glu Thr Phe
 370 375 380
 gtt ttt gaa gat gcc ctg tac gcc ctg aaa acg gcg aaa aca gcc ggc 1200

 Val Phe Glu Asp Ala Leu Tyr Ala Leu Lys Thr Ala Lys Thr Ala Gly
 385 390 395 400
 ttt cct aca gtc ggt gta aaa gac ccc tcc tcc gcc gga cag gaa ggg 1248
 Phe Pro Thr Val Gly Val Lys Asp Pro Ser Ser Ala Gly Gln Glu Gly
 405 410 415
 gag att ata aaa caa gcc gat tac tat ctt tat acc ttc acg aaa tga 1296
 Glu Ile Ile Lys Gln Ala Asp Tyr Tyr Leu Tyr Thr Phe Thr Lys

 420 425 430
 <210> 171
 <211> 431
 <212> PRT
 <213> Clostridium species
 <400> 171
 Met Lys Cys Asp Arg Ser Met Leu Leu Leu Tyr Ala Val Thr Asp Arg
 1 5 10 15
 Ala Trp Thr Gly Lys Lys Thr Leu Leu Gln Gln Val Glu Glu Ala Leu
 20 25 30
 Ala Gly Gly Ala Thr Cys Ile Gln Leu Arg Glu Lys Glu Leu Pro Glu

35 40 45
 Glu Glu Phe Arg Gln Glu Ala Leu Ala Val Lys Glu Leu Cys Arg Arg
 50 55 60
 Tyr His Val Pro Phe Leu Ile Asn Asp Asn Val Glu Leu Ala Val Ser
 65 70 75 80
 Cys Gly Ala Asp Gly Val His Val Gly Gln His Asp Met Ser Ala Ala
 85 90 95
 Asp Val Arg Arg Arg Ile Gly Pro Gly Lys Ile Leu Gly Val Ser Ala
 100 105 110

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

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

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

Lys Ile Val Glu Glu Leu Asn Ala Met Ile Ala Ser Phe Tyr Glu Lys
290 295 300

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

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

<210> 172
<211> 1452
<212> DNA
<213> Eubacterium hallii
<220><221> CDS
<222> (1)..(1452)
<223> Eubacterium hallii gene encoding TMP phosphatase [EEG35494]
<400> 172

atg ata aaa gga gca atc ttt gat att gat gga act tta ctt gat tcc 48
Met Ile Lys Gly Ala Ile Phe Asp Ile Asp Gly Thr Leu Leu Asp Ser
1 5 10 15
atg ccc atc tgg gaa aat gca gga gcg aga tat ctt gct act ctt ggc 96
Met Pro Ile Trp Glu Asn Ala Gly Ala Arg Tyr Leu Ala Thr Leu Gly

20	25	30	
att aag gca aag cca gat tta aaa gaa cgg ctg gat gct tta tct ttg			144
Ile Lys Ala Lys Pro Asp Leu Lys Glu Arg Leu Asp Ala Leu Ser Leu			
35	40	45	
cca gaa gga gcc atc tat atg caa aaa gag tat ggc ctt tcg gta tca			192
Pro Glu Gly Ala Ile Tyr Met Gln Lys Glu Tyr Gly Leu Ser Val Ser			
50	55	60	
gca gaa gac att tta gaa gga gtc aat cag gtt gta aaa gat ttt tac			240
Ala Glu Asp Ile Leu Glu Gly Val Asn Gln Val Val Lys Asp Phe Tyr			
65	70	75	80
tat aaa gaa gcg gtc atg aag ccg gga gcc tat gcc tta gta aaa cgt			288
Tyr Lys Glu Ala Val Met Lys Pro Gly Ala Tyr Ala Leu Val Lys Arg			
85	90	95	
ctg aaa gaa aat ggt gtg aag tta att ata gcc aca gcg aca gat aag			336
Leu Lys Glu Asn Gly Val Lys Leu Ile Ile Ala Thr Ala Thr Asp Lys			
100	105	110	
gag atg gca aag gcg gcg ctt att cgt aac ggc ata tgg cag gac ttt			384
Glu Met Ala Lys Ala Ala Leu Ile Arg Asn Gly Ile Trp Gln Asp Phe			
115	120	125	
acg gga atg att acc tgc gag gaa gcc gga gcc gga aag aca agc ccg			432
Thr Gly Met Ile Thr Cys Glu Glu Ala Gly Ala Gly Lys Thr Ser Pro			
130	135	140	
aag gta ttt gag ctt gca agg caa aag cta ggc act aaa aaa gag gaa			480
Lys Val Phe Glu Leu Ala Arg Gln Lys Leu Gly Thr Lys Lys Glu Glu			
145	150	155	160
aca tgg gta ttt gaa gat tct tta tat gcg gtg aaa act gct act gaa			528
Thr Trp Val Phe Glu Asp Ser Leu Tyr Ala Val Lys Thr Ala Thr Glu			
165	170	175	
gct gga ttt cca gta tgc agt atc tac gat acc tac agt gtg gga aat			576
Ala Gly Phe Pro Val Cys Ser Ile Tyr Asp Thr Tyr Ser Val Gly Asn			
180	185	190	

gcg aaa gaa atc cag aaa ctt tct aat att tat gtg aga gat ttt tcg	624
Ala Lys Glu Ile Gln Lys Leu Ser Asn Ile Tyr Val Arg Asp Phe Ser	
195 200 205	
gag ata ggt gat tat tct ttt tca aat atg aaa aca gtt ctt aca att	672
Glu Ile Gly Asp Tyr Ser Phe Ser Asn Met Lys Thr Val Leu Thr Ile	
210 215 220	
gca ggc agt gat tcg agc gga gga gca ggt att caa gcg gat atc aag	720
Ala Gly Ser Asp Ser Ser Gly Gly Ala Gly Ile Gln Ala Asp Ile Lys	
225 230 235 240	
act tta act gtt cat aaa gta tat gcc atg act tgt atc acc gca ctt	768
Thr Leu Thr Val His Lys Val Tyr Ala Met Thr Cys Ile Thr Ala Leu	
245 250 255	
acc gca caa aat aca gtc gga att acc ggg att atg cca gta cca gca	816
Thr Ala Gln Asn Thr Val Gly Ile Thr Gly Ile Met Pro Val Pro Ala	
260 265 270	
gaa ttt ttt aaa aaa cag atg gaa agc att ttc aca gat ata aag cca	864
Glu Phe Phe Lys Lys Gln Met Glu Ser Ile Phe Thr Asp Ile Lys Pro	
275 280 285	
gat gcg gtg aaa att gga atg att gct tca aag gaa cag gca gag att	912
Asp Ala Val Lys Ile Gly Met Ile Ala Ser Lys Glu Gln Ala Glu Ile	
290 295 300	
atc gca gaa tac ctg gaa aaa tat tct atc aaa aat gta gtg gca gac	960
Ile Ala Glu Tyr Leu Glu Lys Tyr Ser Ile Lys Asn Val Val Ala Asp	
305 310 315 320	
ccg gtg atg att tcg aca agc ggt acg gtt tta gta gaa gaa aca acg	1008
Pro Val Met Ile Ser Thr Ser Gly Thr Val Leu Val Glu Glu Thr Thr	
325 330 335	
aga aag ata tta tat gag aaa tta tat cca aaa gtt tcc ctg cta acc	1056
Arg Lys Ile Leu Tyr Glu Lys Leu Tyr Pro Lys Val Ser Leu Leu Thr	
340 345 350	
ccg aac att cca gaa acc gaa ttt tta tcc ggg ata aaa att acc gat	1104

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

355 360 365
aaa aaa aca agg gaa gaa gca gca aaa gtc att gca gac agg tgg aat 1152

Lys Lys Thr Arg Glu Glu Ala Ala Lys Val Ile Ala Asp Arg Trp Asn

370 375 380
tgt gcg gtc tta agt aag ggc ggt cac agc gaa gaa aat gcg gac gat 1200

Cys Ala Val Leu Ser Lys Gly Gly His Ser Glu Glu Asn Ala Asp Asp

385 390 395 400

ttg ctt tat gag agt ttt ttg cag gaa gaa aaa aaa gaa aaa gcc gtt 1248

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

405 410 415
tgg ttt cca gaa gaa aga att gat aat cca aac aca cac gga acc ggc 1296

Trp Phe Pro Glu Glu Arg Ile Asp Asn Pro Asn Thr His Gly Thr Gly

420 425 430
tgt aca ctt tca agt gcg gta gcg gca aat ctg gca aag gga ttt cct 1344

Cys Thr Leu Ser Ser Ala Val Ala Ala Asn Leu Ala Lys Gly Phe Pro

435 440 445
gta gaa gaa tcc gta aaa aag gca aaa gta tac atc agc gga gca att 1392

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

450 455 460
aga gca atg ctg aat ctt gga cag gga aat ggc ccg cta aat cat atg 1440

Arg Ala Met Leu Asn Leu Gly Gln Gly Asn Gly Pro Leu Asn His Met

465 470 475 480
tgg gat ttg taa 1452

Trp Asp Leu

<210> 173

<211> 483

<212> PRT

<213> Eubacterium hallii

<400> 173

Met Ile Lys Gly Ala Ile Phe Asp Ile Asp Gly Thr Leu Leu Asp Ser
1 5 10 15

Met Pro Ile Trp Glu Asn Ala Gly Ala Arg Tyr Leu Ala Thr Leu Gly
20 25 30

Ile Lys Ala Lys Pro Asp Leu Lys Glu Arg Leu Asp Ala Leu Ser Leu
35 40 45

Pro Glu Gly Ala Ile Tyr Met Gln Lys Glu Tyr Gly Leu Ser Val Ser
50 55 60

Ala Glu Asp Ile Leu Glu Gly Val Asn Gln Val Val Lys Asp Phe Tyr
65 70 75 80

Tyr Lys Glu Ala Val Met Lys Pro Gly Ala Tyr Ala Leu Val Lys Arg
85 90 95

Leu Lys Glu Asn Gly Val Lys Leu Ile Ile Ala Thr Ala Thr Asp Lys
100 105 110

Glu Met Ala Lys Ala Ala Leu Ile Arg Asn Gly Ile Trp Gln Asp Phe
115 120 125

Thr Gly Met Ile Thr Cys Glu Glu Ala Gly Ala Gly Lys Thr Ser Pro
130 135 140

Lys Val Phe Glu Leu Ala Arg Gln Lys Leu Gly Thr Lys Lys Glu Glu
145 150 155 160

Thr Trp Val Phe Glu Asp Ser Leu Tyr Ala Val Lys Thr Ala Thr Glu
165 170 175

Ala Gly Phe Pro Val Cys Ser Ile Tyr Asp Thr Tyr Ser Val Gly Asn
180 185 190

Ala Lys Glu Ile Gln Lys Leu Ser Asn Ile Tyr Val Arg Asp Phe Ser
195 200 205

Glu Ile Gly Asp Tyr Ser Phe Ser Asn Met Lys Thr Val Leu Thr Ile
210 215 220

Ala Gly Ser Asp Ser Ser Gly Gly Ala Gly Ile Gln Ala Asp Ile Lys
225 230 235 240

Thr Leu Thr Val His Lys Val Tyr Ala Met Thr Cys Ile Thr Ala Leu

245 250 255
 Thr Ala Gln Asn Thr Val Gly Ile Thr Gly Ile Met Pro Val Pro Ala
 260 265 270

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

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

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

Arg Ala Met Leu Asn Leu Gly Gln Gly Asn Gly Pro Leu Asn His Met
 465 470 475 480
 Trp Asp Leu

<210> 174
 <211> 1362
 <212> DNA
 <213> Eubacterium species
 <220><221> CDS
 <222> (1)..(1362)
 <223> Eubacterium sp. CAG:252 gene encoding TMP phosphatase [CDB67556]
 <400> 174

atg aaa aat aaa ttt ttc aca cgc gag att tgt gtc tgc gtg cac ttg 48
 Met Lys Asn Lys Phe Phe Thr Arg Glu Ile Cys Val Cys Val His Leu
 1 5 10 15

aca caa act cgt tat gcg caa aaa acg tgc gca gaa atg agg aat agt 96
 Thr Gln Thr Arg Tyr Ala Gln Lys Thr Cys Ala Glu Met Arg Asn Ser
 20 25 30

gtg aag gtt aaa gct gag gat atg cag cta tac gct gtt aca gat aca 144
 Val Lys Val Lys Ala Glu Asp Met Gln Leu Tyr Ala Val Thr Asp Thr
 35 40 45

cag tgg ctt aat gga cgt gac ttt ctt gaa gta ata gaa agc gtt ctt 192

Gln Trp Leu Asn Gly Arg Asp Phe Leu Glu Val Ile Glu Ser Val Leu
 50 55 60

gca aat gga gct aca ttt tta cag tta agg gaa aaa aat gcc aca cat 240
 Ala Asn Gly Ala Thr Phe Leu Gln Leu Arg Glu Lys Asn Ala Thr His
 65 70 75 80

gag gaa ata gtg gca aag gcg aag gca ata aag cca ata gct aag aag 288
 Glu Glu Ile Val Ala Lys Ala Lys Ala Ile Lys Pro Ile Ala Lys Lys
 85 90 95

tac gga gtg cct ttt gtc ata gat gat gac ata tat gca gct aaa gag 336
 Tyr Gly Val Pro Phe Val Ile Asp Asp Asp Ile Tyr Ala Ala Lys Glu
 100 105 110

gca gac gtg gat ggt gtc cac ata ggg cag aat gat gca agc tat gag 384
 Ala Asp Val Asp Gly Val His Ile Gly Gln Asn Asp Ala Ser Tyr Glu

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

Ala Ile Leu Arg Gly Ala Asp Ile Pro Asp Asn Phe Asp Glu Ile Thr
275 280 285
ggc gtt atg gac ctt aat gat tat gct gag tat gtt aaa aat gtt ctt 912
Gly Val Met Asp Leu Asn Asp Tyr Ala Glu Tyr Val Lys Asn Val Leu
290 295 300
ggc ata gat act aat ctt gag cag ata aca gaa gcg gct gtc gag att 960
Gly Ile Asp Thr Asn Leu Glu Gln Ile Thr Glu Ala Ala Val Glu Ile
305 310 315 320
atg aat aaa cat tac gaa aaa gat ata cct gca aag gac ggt atg aca 1008
Met Asn Lys His Tyr Glu Lys Asp Ile Pro Ala Lys Asp Gly Met Thr
325 330 335
gag ctt gtc acg aga gaa tat aag gcc gga agc aga ctt gtt gtg ttt 1056
Glu Leu Val Thr Arg Glu Tyr Lys Ala Gly Ser Arg Leu Val Val Phe
340 345 350
acg gct tca gat aga aga agt gtt gaa att ctt ctt tca cac ctt gga 1104
Thr Ala Ser Asp Arg Arg Ser Val Glu Ile Leu Leu Ser His Leu Gly
355 360 365
ata aga gaa tgt ttt tat gat ata tat aca gtc tat gat gta gga ctt 1152
Ile Arg Glu Cys Phe Tyr Asp Ile Tyr Thr Val Tyr Asp Val Gly Leu
370 375 380
aag aag agt gat aag aac agc tat ctt aag gtg gca gag ctt gca ggc 1200
Lys Lys Ser Asp Lys Asn Ser Tyr Leu Lys Val Ala Glu Leu Ala Gly
385 390 395 400
atg aaa gat aca tca cag gta tgg gta tat gag gat ata tta aga ggt 1248
Met Lys Asp Thr Ser Gln Val Trp Val Tyr Glu Asp Ile Leu Arg Gly
405 410 415
gta aag gca gcg aaa gag gcc gga ctt aat gtg tgt gca gtg tat gat 1296
Val Lys Ala Ala Lys Glu Ala Gly Leu Asn Val Cys Ala Val Tyr Asp
420 425 430
gag gac tcc gca ggc gac tgg gag gac ata aaa gag ctt gcg gat aag 1344
Glu Asp Ser Ala Gly Asp Trp Glu Asp Ile Lys Glu Leu Ala Asp Lys

435 440 445

acc ctt gaa ctt gtg taa 1362

Thr Leu Glu Leu Val

450

<210> 175

<211> 453

<212> PRT

<213> Eubacterium species

<400> 175

Met Lys Asn Lys Phe Phe Thr Arg Glu Ile Cys Val Cys Val His Leu

1 5 10 15

Thr Gln Thr Arg Tyr Ala Gln Lys Thr Cys Ala Glu Met Arg Asn Ser

20 25 30

Val Lys Val Lys Ala Glu Asp Met Gln Leu Tyr Ala Val Thr Asp Thr

35 40 45

Gln Trp Leu Asn Gly Arg Asp Phe Leu Glu Val Ile Glu Ser Val Leu

50 55 60

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

65 70 75 80

Glu Glu Ile Val Ala Lys Ala Lys Ala Ile Lys Pro Ile Ala Lys Lys

85 90 95

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

100 105 110

Ala Asp Val Asp Gly Val His Ile Gly Gln Asn Asp Ala Ser Tyr Glu

115 120 125

Lys Ala Arg Glu Val Leu Gly Glu Gly Lys Ile Ile Gly Met Thr Val

130 135 140

Lys Thr Arg Gln Gln Ala Glu Asn Ala Ile Arg Leu Gly Ala Asp Tyr

145 150 155 160

Val Gly Met Gly Ala Val Phe His Thr Ser Thr Lys Lys Asp Ala Lys

165 170 175

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

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

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

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

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

420 425 430
Glu Asp Ser Ala Gly Asp Trp Glu Asp Ile Lys Glu Leu Ala Asp Lys

435 440 445
Thr Leu Glu Leu Val

450
<210> 176
<211> 1260
<212> DNA
<213> Lachnospiraceae pectinoschiza
<220><221> CDS
<222> (1)..(1260)
<223> Lachnospiraceae pectinoschiza gene encoding TMP phosphatase
[CUQ76318]
<400> 176

atg aaa gtt acc cgt gaa gat atg cag ctt tac gcc gtt aca gat acg 48
Met Lys Val Thr Arg Glu Asp Met Gln Leu Tyr Ala Val Thr Asp Thr
1 5 10 15

caa tgg ctt aat ggc agg gat ttc tat gaa gag att gag aaa gtc ctt 96
Gln Trp Leu Asn Gly Arg Asp Phe Tyr Glu Glu Ile Glu Lys Val Leu
20 25 30

gcg gca gga gct aca ttt ttg cag tta aga gaa aag gat tcg aca cac 144
Ala Ala Gly Ala Thr Phe Leu Gln Leu Arg Glu Lys Asp Ser Thr His
35 40 45

gag gag att gta aaa aaa gca ttg gca att aaa ccg ata gca aga aga 192

Glu Glu Ile Val Lys Lys Ala Leu Ala Ile Lys Pro Ile Ala Arg Arg
50 55 60

tat ggt gtg cca ttt gtt ata gat gat gat ata tac gcg gcg tta gag 240
Tyr Gly Val Pro Phe Val Ile Asp Asp Asp Ile Tyr Ala Ala Leu Glu
65 70 75 80

gca gat gtt gac gga gtt cat ata gga caa agt gat gca agc tac gaa 288
Ala Asp Val Asp Gly Val His Ile Gly Gln Ser Asp Ala Ser Tyr Glu

85	90	95	
aca gca aga gag ctt cta gga cct gac aag ata ata gga atg aca gta			336
Thr Ala Arg Glu Leu Leu Gly Pro Asp Lys Ile Ile Gly Met Thr Val			
100	105	110	
aag aca cca gag cag gcg gca aat gcg gca aga ctt ggt gct gat tat			384
Lys Thr Pro Glu Gln Ala Ala Asn Ala Ala Arg Leu Gly Ala Asp Tyr			
115	120	125	
gtt gga atg gga gct gta ttt cat aca agc acg aag aaa gat gcc aaa			432
Val Gly Met Gly Ala Val Phe His Thr Ser Thr Lys Lys Asp Ala Lys			
130	135	140	
gat tta agc agg gat aat ctt ctt aag ctt aca gct atg ctt gat atg			480
Asp Leu Ser Arg Asp Asn Leu Leu Lys Leu Thr Ala Met Leu Asp Met			
145	150	155	160
ccg ata gtt gca att ggc ggc att aat tat gac aac tgt gat tat tta			528
Pro Ile Val Ala Ile Gly Gly Ile Asn Tyr Asp Asn Cys Asp Tyr Leu			
165	170	175	
aaa gat aca ggc gtg gac gga att gct gtt gta tcg gcg ata ttt gca			576
Lys Asp Thr Gly Val Asp Gly Ile Ala Val Val Ser Ala Ile Phe Ala			
180	185	190	
agt gat gac tgc gcg gag gcg aca cga aag ctt tat aag aag aca aga			624
Ser Asp Asp Cys Ala Glu Ala Thr Arg Lys Leu Tyr Lys Lys Thr Arg			
195	200	205	
aag ctg ttt aat tat aat aag aac ata ata ttt gat atg gac gga aca			672
Lys Leu Phe Asn Tyr Asn Lys Asn Ile Ile Phe Asp Met Asp Gly Thr			
210	215	220	
ctt gtt gac tct atg ccg ttc tgg aag aat agt gca agg gaa tat gcc			720
Leu Val Asp Ser Met Pro Phe Trp Lys Asn Ser Ala Arg Glu Tyr Ala			
225	230	235	240
att tta aga ggt gct aag ctt cca aag aat ttt gat gag ata aca gga			768
Ile Leu Arg Gly Ala Lys Leu Pro Lys Asn Phe Asp Glu Ile Thr Gly			
245	250	255	

gtt atg gac ctt tcg gaa tat gcg gct tat ctg caa aat gtt ctt ggg	816
Val Met Asp Leu Ser Glu Tyr Ala Ala Tyr Leu Gln Asn Val Leu Gly	
260 265 270	
att gat aca tcg cta gaa cag ata aca gag gca gca gtt gat att atg	864
Ile Asp Thr Ser Leu Glu Gln Ile Thr Glu Ala Ala Val Asp Ile Met	
275 280 285	
aat aag cat tat gca agt gat att cct gca aag aag gga atg ata aag	912
Asn Lys His Tyr Ala Ser Asp Ile Pro Ala Lys Lys Gly Met Ile Lys	
290 295 300	
ctt ata aga aga gaa tat gag gct gga agc aag ctt gta ata ttc agt	960
Leu Ile Arg Arg Glu Tyr Glu Ala Gly Ser Lys Leu Val Ile Phe Ser	
305 310 315 320	
gct tcc gat act tcc agt gtg gaa att ctt ctt aaa agg tta gaa ata	1008
Ala Ser Asp Thr Ser Ser Val Glu Ile Leu Leu Lys Arg Leu Glu Ile	
325 330 335	
tat gaa tgt ttt gag gga ata tac aca gta tat gat gtc ggc ata gga	1056
Tyr Glu Cys Phe Glu Gly Ile Tyr Thr Val Tyr Asp Val Gly Ile Gly	
340 345 350	
aag agt gat aag gaa agc tat aaa aag gtt gcc agg tca gca gga atg	1104
Lys Ser Asp Lys Glu Ser Tyr Lys Lys Val Ala Arg Ser Ala Gly Met	
355 360 365	
gat ata tct gat acg tgg gtg tat gag gat att cta aga ggc gtt cgg	1152
Asp Ile Ser Asp Thr Trp Val Tyr Glu Asp Ile Leu Arg Gly Val Arg	
370 375 380	
gcg gca cat aat gct gga ttg aaa gtg tgt gcg gta tat gat aaa gac	1200
Ala Ala His Asn Ala Gly Leu Lys Val Cys Ala Val Tyr Asp Lys Asp	
385 390 395 400	
tcg gca gat gac tgg gat gag ata tgc agt att gca gat aaa tgt ata	1248
Ser Ala Asp Asp Trp Asp Glu Ile Cys Ser Ile Ala Asp Lys Cys Ile	
405 410 415	
ata acc gga taa	1260

Ile Thr Gly

<210> 177

<211> 419

<212> PRT

<213> Lachnospiraceae pectinoschiza

<400> 177

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

1 5 10 15

Gln Trp Leu Asn Gly Arg Asp Phe Tyr Glu Glu Ile Glu Lys Val Leu

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

Ala Asp Val Asp Gly Val His Ile Gly Gln Ser Asp Ala Ser Tyr Glu

85 90 95

Thr Ala Arg Glu Leu Leu Gly Pro Asp Lys Ile Ile Gly Met Thr Val

100 105 110

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

115 120 125

Val Gly Met Gly Ala Val Phe His Thr Ser Thr Lys Lys Asp Ala Lys

130 135 140

Asp Leu Ser Arg Asp Asn Leu Leu Lys Leu Thr Ala Met Leu Asp Met

145 150 155 160

Pro Ile Val Ala Ile Gly Gly Ile Asn Tyr Asp Asn Cys Asp Tyr Leu

165 170 175

Lys Asp Thr Gly Val Asp Gly Ile Ala Val Val Ser Ala Ile Phe Ala

180 185 190

Ser Asp Asp Cys Ala Glu Ala Thr Arg Lys Leu Tyr Lys Lys Thr Arg
195 200 205

Lys Leu Phe Asn Tyr Asn Lys Asn Ile Ile Phe Asp Met Asp Gly Thr
210 215 220

Leu Val Asp Ser Met Pro Phe Trp Lys Asn Ser Ala Arg Glu Tyr Ala
225 230 235 240

Ile Leu Arg Gly Ala Lys Leu Pro Lys Asn Phe Asp Glu Ile Thr Gly
245 250 255

Val Met Asp Leu Ser Glu Tyr Ala Ala Tyr Leu Gln Asn Val Leu Gly
260 265 270

Ile Asp Thr Ser Leu Glu Gln Ile Thr Glu Ala Ala Val Asp Ile Met
275 280 285

Asn Lys His Tyr Ala Ser Asp Ile Pro Ala Lys Lys Gly Met Ile Lys
290 295 300

Leu Ile Arg Arg Glu Tyr Glu Ala Gly Ser Lys Leu Val Ile Phe Ser
305 310 315 320

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

Tyr Glu Cys Phe Glu Gly Ile Tyr Thr Val Tyr Asp Val Gly Ile Gly
340 345 350

Lys Ser Asp Lys Glu Ser Tyr Lys Lys Val Ala Arg Ser Ala Gly Met
355 360 365

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

Ala Ala His Asn Ala Gly Leu Lys Val Cys Ala Val Tyr Asp Lys Asp
385 390 395 400

Ser Ala Asp Asp Trp Asp Glu Ile Cys Ser Ile Ala Asp Lys Cys Ile
405 410 415

Ile Thr Gly

<210> 178

<211> 1296

<212> DNA

<213> Peptostreptococcaceae bacterium

<220>

<221> CDS

<222> (1)..(1296)

<223> Peptostreptococcaceae bacterium OBRC8 gene encoding TMP
phosphatase[WP_009530263]

<400> 178

atg aaa aat att gac tat aca atg tat tac gtc acc gat gaa gac ctt 48

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

1 5 10 15

ttg agc agt aat cat acc ttg gaa aca tct gta caa gat gcc att tta 96

Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu

20 25 30

ggt ggc tgt aca atg ata cag ctt cga gaa aaa cat tca tcc act ctc 144

Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu

35 40 45

gat ttt tat aac aaa gcc ata aaa att aaa gcc att tgc gac aag tac 192

Asp Phe Tyr Asn Lys Ala Ile Lys Ile Lys Ala Ile Cys Asp Lys Tyr

50 55 60

aac ata cct ctt ata ata aat gac aga ata gat gta gct ctt gca ata 240

Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile

65 70 75 80

aac gca gac gga gta cat ctc gga caa gac gat atg cct ctt gat att 288

Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile

85 90 95

gca aga aaa att atg gga gat ggc aaa att ata gga ata tca act gca 336

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

100 105 110

act tta gat gaa gct cta atc gct caa caa ggc ggt gca gat tat gta 384

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

115 120 125

gga gta ggt gct atg tac agc aca aac aca aaa acc gat gcc aat ttg 432
Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu

130 135 140
aca act ata aac gag ctt aca aaa ata aaa aac aat cta aaa ata cct 480
Thr Thr Ile Asn Glu Leu Thr Lys Ile Lys Asn Asn Leu Lys Ile Pro

145 150 155 160
gta gtt gca atc ggc ggt ata aac ctt gac aca ata cct gct cta aaa 528
Val Val Ala Ile Gly Gly Ile Asn Leu Asp Thr Ile Pro Ala Leu Lys

165 170 175
cct gca caa ata gac gga gtt gca ata gta tcc gct ata tct atg cag 576
Pro Ala Gln Ile Asp Gly Val Ala Ile Val Ser Ala Ile Ser Met Gln

180 185 190
gaa gat acc gta tct gca aca aga aaa tta aaa aat act ttt ttg aaa 624
Glu Asp Thr Val Ser Ala Thr Arg Lys Leu Lys Asn Thr Phe Leu Lys

195 200 205
caa tat caa act aaa ggc gta ata ttc gat att gac ggt act ctg ctt 672

Gln Tyr Gln Thr Lys Gly Val Ile Phe Asp Ile Asp Gly Thr Leu Leu
210 215 220
gaa act atg aac ata tgg gac aat gta ctt cta aac ctt atg aat aca 720
Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr

225 230 235 240
ctt aat atc agc tat acc gaa gat gaa ata caa aaa ata tgg aat atg 768
Leu Asn Ile Ser Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met

245 250 255
ggt ttt gca gag ctt gca cag ttc agc ata aaa aaa ttc aag ctt gat 816
Gly Phe Ala Glu Leu Ala Gln Phe Ser Ile Lys Lys Phe Lys Leu Asp

260 265 270
atg agt gta aaa gaa ttt tgg caa ctt ata aaa aaa tta tca gtc gaa 864
Met Ser Val Lys Glu Phe Trp Gln Leu Ile Lys Lys Leu Ser Val Glu

275	280	285	
gag tat aaa aat agc aaa ata cac tta aaa aaa ggt gca aaa aaa ctg			912
Glu Tyr Lys Asn Ser Lys Ile His Leu Lys Lys Gly Ala Lys Lys Leu			
290	295	300	
ctt gag tat ctc aaa gaa aaa ggt gta aaa tta gcc ata gca act gcc			960
Leu Glu Tyr Leu Lys Glu Lys Gly Val Lys Leu Ala Ile Ala Thr Ala			
305	310	315	320
ctt tgc aaa gaa cag tat gaa ata gtg ctt aca aag aca ggt atc ata			1008
Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile			
325	330	335	
gac tat ttt gac ata ata gca tca agc gta gat tta aaa atg gaa aaa			1056
Asp Tyr Phe Asp Ile Ile Ala Ser Ser Val Asp Leu Lys Met Glu Lys			
340	345	350	
tca gac aga caa ata ttt gac tat ata gca aaa aat cta caa gtt cca			1104
Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro			
355	360	365	
aac aaa aat ctt att ttc ttt gaa gac gac ata aac tcg tca aca ggt			1152
Asn Lys Asn Leu Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly			
370	375	380	
gcc aag ttg gca gga cta aaa ctg tgc att gta tca aac aag aaa tat			1200
Ala Lys Leu Ala Gly Leu Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr			
385	390	395	400
aac ggt aac agc aaa ttt gac gct ctc ata gat tat aaa ata gat gat			1248
Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp			
405	410	415	
ttt gaa aat aaa ttg ata tat gat gaa ata ata gtg gag aaa aat tag			1296
Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn			
420	425	430	
<210> 179			
<211> 431			
<212> PRT			

<213> Peptostreptococcaceae bacterium

<400> 179

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

1 5 10 15

Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu

20 25 30

Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu

35 40 45

Asp Phe Tyr Asn Lys Ala Ile Lys Ile Lys Ala Ile Cys Asp Lys Tyr

50 55 60

Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile

65 70 75 80

Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile

85 90 95

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

100 105 110

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

115 120 125

Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu

130 135 140

Thr Thr Ile Asn Glu Leu Thr Lys Ile Lys Asn Asn Leu Lys Ile Pro

145 150 155 160

Val Val Ala Ile Gly Gly Ile Asn Leu Asp Thr Ile Pro Ala Leu Lys

165 170 175

Pro Ala Gln Ile Asp Gly Val Ala Ile Val Ser Ala Ile Ser Met Gln

180 185 190

Glu Asp Thr Val Ser Ala Thr Arg Lys Leu Lys Asn Thr Phe Leu Lys

195 200 205

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

210 215 220

Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr

225 230 235 240
 Leu Asn Ile Ser Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met
 245 250 255

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

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

 Ala Lys Leu Ala Gly Leu Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr
 385 390 395 400
 Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp
 405 410 415
 Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn
 420 425 430

 <210> 180
 <211> 1296
 <212> DNA
 <213> Peptostreptococcaceae bacterium
 <220><221> CDS
 <222> (1)..(1296)
 <223>

> Peptostreptococcaceae bacterium CM2 gene encoding TMP

phosphatase[WP_009527854]

<400> 180

atg aaa aat att gac tat aca atg tat tac gtc acc gat gaa gac ctt 48

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

1 5 10 15

ttg agc agt aat cac acc ttg gaa aca tct gtg caa gat gcc att tta 96

Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu

20 25 30

ggt ggc tgt aca atg ata cag ctt cga gaa aaa cat tca tcc act ctc 144

Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu

35 40 45

gat ttt tat aac aaa gcc gta aaa att aaa gct att tgc gac aag tac 192

Asp Phe Tyr Asn Lys Ala Val Lys Ile Lys Ala Ile Cys Asp Lys Tyr

50 55 60

aac ata cct ctt ata ata aat gac aga ata gac gta gct ctt gca ata 240

Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile

65 70 75 80

aat gca gac gga gta cat ctc gga caa gac gat atg cct ctt gat att 288

Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile

85 90 95

gca aga aaa att atg gga gat ggc aaa att ata gga ata tca acc tca 336

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

100 105 110

act tta gat gaa gct cta atc gct caa caa ggc ggt gca gat tat gta 384

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

115 120 125

ggt gta ggt gct atg tac agc aca aac aca aaa act gat gcc aat ttg 432

Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu

130 135 140

aca act ata gac gag ctt aca aaa ata aaa aac aat tta aaa ata cct 480

Thr	Thr	Ile	Asp	Glu	Leu	Thr	Lys	Ile	Lys	Asn	Asn	Leu	Lys	Ile	Pro	
145				150				155				160				
gtt	gtt	gca	atc	ggc	ggt	ata	aac	ctt	gac	act	ata	ccc	gct	cta	aaa	528
Val	Val	Ala	Ile	Gly	Gly	Ile	Asn	Leu	Asp	Thr	Ile	Pro	Ala	Leu	Lys	
165				170				175								
cct	gcg	caa	ata	gac	gga	gtt	gca	ata	gta	tcc	gct	ata	tct	atg	cag	576
Pro	Ala	Gln	Ile	Asp	Gly	Val	Ala	Ile	Val	Ser	Ala	Ile	Ser	Met	Gln	
180				185				190								
gaa	gat	acc	gta	tct	gca	aca	aga	aaa	tta	aaa	aat	act	ttt	ttg	aaa	624
Glu	Asp	Thr	Val	Ser	Ala	Thr	Arg	Lys	Leu	Lys	Asn	Thr	Phe	Leu	Lys	
195				200				205								
caa	tat	caa	act	aaa	ggc	gta	ata	ttc	gat	att	gac	ggg	act	ctg	ctt	672
Gln	Tyr	Gln	Thr	Lys	Gly	Val	Ile	Phe	Asp	Ile	Asp	Gly	Thr	Leu	Leu	
210				215				220								
gaa	act	atg	aac	ata	tgg	gac	aat	gta	ctt	cta	aat	ctt	atg	aat	acg	720
Glu	Thr	Met	Asn	Ile	Trp	Asp	Asn	Val	Leu	Leu	Asn	Leu	Met	Asn	Thr	
225				230				235				240				
ctt	aat	atc	cgc	tat	acc	gaa	gat	gaa	ata	caa	aag	ata	tgg	aat	atg	768
Leu	Asn	Ile	Arg	Tyr	Thr	Glu	Asp	Glu	Ile	Gln	Lys	Ile	Trp	Asn	Met	
245				250				255								
ggg	ttt	gca	gag	ctt	gca	cag	ttc	agc	ata	aaa	aaa	ttc	aag	ctt	gat	816
Gly	Phe	Ala	Glu	Leu	Ala	Gln	Phe	Ser	Ile	Lys	Lys	Phe	Lys	Leu	Asp	
260				265				270								
atg	agt	gta	aaa	gaa	ttt	tgg	caa	ctt	ata	aaa	aaa	tta	tca	gtc	gaa	864
Met	Ser	Val	Lys	Glu	Phe	Trp	Gln	Leu	Ile	Lys	Lys	Leu	Ser	Val	Glu	
275				280				285								
gag	tat	aaa	aat	agc	aaa	ata	cac	tta	aaa	aaa	ggg	gca	aaa	aaa	ctg	912
Glu	Tyr	Lys	Asn	Ser	Lys	Ile	His	Leu	Lys	Lys	Gly	Ala	Lys	Lys	Leu	
290				295				300								
ctt	gag	tat	ctc	aaa	gaa	aaa	ggg	gta	aaa	tta	gcc	ata	gca	act	gcc	960
Leu	Glu	Tyr	Leu	Lys	Glu	Lys	Gly	Val	Lys	Leu	Ala	Ile	Ala	Thr	Ala	

305 310 315 320
 ctt tgc aaa gaa cag tat gaa ata gtg ctt aca aag aca ggt atc ata 1008
 Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile

 325 330 335
 gac tat ttt gac ata ata gca tca agc gta gat tta aaa atg gaa aaa 1056
 Asp Tyr Phe Asp Ile Ile Ala Ser Ser Val Asp Leu Lys Met Glu Lys

 340 345 350
 tca gat aga caa ata ttt gac tat ata gca aaa aat cta caa gtt cca 1104
 Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro

 355 360 365

 aac aaa aat ttt att ttc ttt gaa gac gac ata aac tcg tca aca ggt 1152
 Asn Lys Asn Phe Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly

 370 375 380
 gca aaa cgt gca gga gta aaa ctg tgc att gta tca aac aag aaa tat 1200
 Ala Lys Arg Ala Gly Val Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr

385 390 395 400
 aat ggt aac agc aaa ttt gac gct ctc ata gat tat aaa ata gat gat 1248

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

 405 410 415
 ttt gaa aat aaa ttg ata tat gat gaa ata ata gtg gag aaa aat tag 1296
 Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn

 420 425 430

<210> 181

<211> 431

<212> PRT

<213> Peptostreptococcaceae bacterium

<400> 181

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

1 5 10 15
 Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu

 20 25 30

Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu
 35 40 45
 Asp Phe Tyr Asn Lys Ala Val Lys Ile Lys Ala Ile Cys Asp Lys Tyr
 50 55 60
 Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile
 65 70 75 80
 Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile
 85 90 95
 Ala Arg Lys Ile Met Gly Asp Gly Lys Ile Ile Gly Ile Ser Thr Ser
 100 105 110
 Thr Leu Asp Glu Ala Leu Ile Ala Gln Gln Gly Gly Ala Asp Tyr Val
 115 120 125
 Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu
 130 135 140
 Thr Thr Ile Asp Glu Leu Thr Lys Ile Lys Asn Asn Leu Lys Ile Pro
 145 150 155 160
 Val Val Ala Ile Gly Gly Ile Asn Leu Asp Thr Ile Pro Ala Leu Lys
 165 170 175
 Pro Ala Gln Ile Asp Gly Val Ala Ile Val Ser Ala Ile Ser Met Gln
 180 185 190
 Glu Asp Thr Val Ser Ala Thr Arg Lys Leu Lys Asn Thr Phe Leu Lys
 195 200 205
 Gln Tyr Gln Thr Lys Gly Val Ile Phe Asp Ile Asp Gly Thr Leu Leu
 210 215 220
 Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr
 225 230 235 240
 Leu Asn Ile Arg Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met
 245 250 255
 Gly Phe Ala Glu Leu Ala Gln Phe Ser Ile Lys Lys Phe Lys Leu Asp
 260 265 270
 Met Ser Val Lys Glu Phe Trp Gln Leu Ile Lys Lys Leu Ser Val Glu

275 280 285
 Glu Tyr Lys Asn Ser Lys Ile His Leu Lys Lys Gly Ala Lys Lys Leu
 290 295 300
 Leu Glu Tyr Leu Lys Glu Lys Gly Val Lys Leu Ala Ile Ala Thr Ala
 305 310 315 320
 Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile
 325 330 335
 Asp Tyr Phe Asp Ile Ile Ala Ser Ser Val Asp Leu Lys Met Glu Lys
 340 345 350
 Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro
 355 360 365
 Asn Lys Asn Phe Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly
 370 375 380
 Ala Lys Arg Ala Gly Val Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr

385 390 395 400
 Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp
 405 410 415
 Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn
 420 425 430

<210> 182

<211> 1365

<212> DNA

<213> Atopobium species

<220><221> CDS

<222> (1)..(1365)

<223> Atopobium sp. ICM42b gene encoding TMP phosphatase [WP_035427744]

<400> 182

atg cag gtg acc ggt gca att ttt gat tgc gat gga act ctt gtt gat 48
 Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp
 1 5 10 15
 tca atg cgc gtt tgg cat aac gtt ttt ggc gct gtt ctt cct aaa tat 96
 Ser Met Arg Val Trp His Asn Val Phe Gly Ala Val Leu Pro Lys Tyr

20	25	30	
ggc aag act att gat tcg gat att ttt gac cgc gta gag gct gtt tcc			144
Gly Lys Thr Ile Asp Ser Asp Ile Phe Asp Arg Val Glu Ala Val Ser			
35	40	45	
ctc att ggt gga tgt cag att tgc gtt gat gaa ctg gat ttg cct att			192
Leu Ile Gly Gly Cys Gln Ile Cys Val Asp Glu Leu Asp Leu Pro Ile			
50	55	60	
aca gcg gaa gct ttg tat gaa gag ttc tgc gcg tac gta att gat cag			240
Thr Ala Glu Ala Leu Tyr Glu Glu Phe Cys Ala Tyr Val Ile Asp Gln			
65	70	75	80
tac caa cat cat gtt tca atc att ccc ggt gca aag gag ttc tta cag			288
Tyr Gln His His Val Ser Ile Ile Pro Gly Ala Lys Glu Phe Leu Gln			
85	90	95	
gag ctc tac gat gca ggt att cct atg gcc gtt gct tcg tca act ccc			336
Glu Leu Tyr Asp Ala Gly Ile Pro Met Ala Val Ala Ser Ser Thr Pro			
100	105	110	
gtg cga gaa gtt cgt gca gct ctg gca gct caa ggt att gag cac ctc			384
Val Arg Glu Val Arg Ala Ala Leu Ala Ala Gln Gly Ile Glu His Leu			
115	120	125	
ttc aaa aca gtg gtc tca aca gaa gat gtg ggg gga gtg gac aag gtt			432
Phe Lys Thr Val Val Ser Thr Glu Asp Val Gly Gly Val Asp Lys Val			
130	135	140	
gag cct gat gtt tat ctt gag gct ctt cgc cgt ctt ggc acc gat aag			480
Glu Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys			
145	150	155	160
gca act acc tgg gtc ttc gag gat gcc ccg ttt ggc gca cag aca gca			528
Ala Thr Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Ala Gln Thr Ala			
165	170	175	
caa aat gcg ggc ttt cct gtg gta gcg ctc tac aat gat cat gac ggc			576

Gln Asn Ala Gly Phe Pro Val Val Ala Leu Tyr Asn Asp His Asp Gly

180 185 190

cgc gac ccc gtc ttt atg cgc gag cac tct aac atc ttt gcc cac acc 624

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

195 200 205

tac ggc gag ctg tcg ctt ctg cgc ctt cag gac tac gag cgc cct ctg 672

Tyr Gly Glu Leu Ser Leu Leu Arg Leu Gln Asp Tyr Glu Arg Pro Leu

210 215 220

acc gca gcg cct tct ggc gag aaa ccc ctt gag gtc ctt gtt gtg ggc 720

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

225 230 235 240

gga tcc cca gag gcg gtt tca cac acg acg ctg tct acc tgc gcc caa 768

Gly Ser Pro Glu Ala Val Ser His Thr Thr Leu Ser Thr Cys Ala Gln

245 250 255

agc gct gac tac ctg ata gcg gtt gac cat ggt gca gat gca tgt cac 816

Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Ala Cys His

260 265 270

gct gcc ggc gtg att cca cag ctt gcg ctt gga gac ttt gac tcg gct 864

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

275 280 285

aca cca gaa act ctg gct tgg ctc aaa gag cag cag gta cct tgc atg 912

Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met

290 295 300

aag ttt aat gcg gac aag tac gat acc gac ctg gct ctt gct tta aag 960

Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys

305 310 315 320

tcc gcc gag tac gag gct att cgt aga gat agc aag ctc tct ctt acg 1008

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

325 330 335

gtt gtc tcc aca tct ggc gga cac ctt gat cac cag ctt gta gtg ctt 1056
Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu
340 345 350

ggc ctt ctc gcc acg tgg gca aag acg ggc aag gca agt gtt cga gtt 1104
Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val
355 360 365

att gaa aat gac ttt gag atg cgc ttt tta act gca ggc cag gtt gat 1152

Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Thr Ala Gly Gln Val Asp
370 375 380

tct tgg cag ctg agc gca act gat gta ggt aaa aag atg tcc ctt gtg 1200
Ser Trp Gln Leu Ser Ala Thr Asp Val Gly Lys Lys Met Ser Leu Val
385 390 395 400

gct ttg tca gag gag tgc gag gtt tct gag gcc ggc atg aag tgg aat 1248
Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn
405 410 415

ctt gat cac cag aag ttc acc ttg ctg gga gac gac ggt att tca aat 1296
Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn
420 425 430

atc gtc gaa tca gac aat tcc tgg gta agg tgc gag aag ggc tgt ctt 1344
Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu
435 440 445

ttg gtg cag ctt tgg aac taa 1365
Leu Val Gln Leu Trp Asn
450

<210> 183
<211> 454
<212> PRT
<213> Atopobium species
<400> 183

Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp
1 5 10 15
Ser Met Arg Val Trp His Asn Val Phe Gly Ala Val Leu Pro Lys Tyr

				20				25				30			
Gly	Lys	Thr	Ile	Asp	Ser	Asp	Ile	Phe	Asp	Arg	Val	Glu	Ala	Val	Ser
35				40				45							
Leu	Ile	Gly	Gly	Cys	Gln	Ile	Cys	Val	Asp	Glu	Leu	Asp	Leu	Pro	Ile
50				55				60							
Thr	Ala	Glu	Ala	Leu	Tyr	Glu	Glu	Phe	Cys	Ala	Tyr	Val	Ile	Asp	Gln
65				70				75				80			
Tyr	Gln	His	His	Val	Ser	Ile	Ile	Pro	Gly	Ala	Lys	Glu	Phe	Leu	Gln
85				90				95							
Glu	Leu	Tyr	Asp	Ala	Gly	Ile	Pro	Met	Ala	Val	Ala	Ser	Ser	Thr	Pro
100				105				110							
Val	Arg	Glu	Val	Arg	Ala	Ala	Leu	Ala	Ala	Gln	Gly	Ile	Glu	His	Leu
115				120				125							
Phe	Lys	Thr	Val	Val	Ser	Thr	Glu	Asp	Val	Gly	Gly	Val	Asp	Lys	Val
130				135				140							
Glu	Pro	Asp	Val	Tyr	Leu	Glu	Ala	Leu	Arg	Arg	Leu	Gly	Thr	Asp	Lys
145				150				155				160			
Ala	Thr	Thr	Trp	Val	Phe	Glu	Asp	Ala	Pro	Phe	Gly	Ala	Gln	Thr	Ala
165				170				175							
Gln	Asn	Ala	Gly	Phe	Pro	Val	Val	Ala	Leu	Tyr	Asn	Asp	His	Asp	Gly
180				185				190							
Arg	Asp	Pro	Val	Phe	Met	Arg	Glu	His	Ser	Asn	Ile	Phe	Ala	His	Thr
195				200				205							
Tyr	Gly	Glu	Leu	Ser	Leu	Leu	Arg	Leu	Gln	Asp	Tyr	Glu	Arg	Pro	Leu
210				215				220							
Thr	Ala	Ala	Pro	Ser	Gly	Glu	Lys	Pro	Leu	Glu	Val	Leu	Val	Val	Gly
225				230				235				240			
Gly	Ser	Pro	Glu	Ala	Val	Ser	His	Thr	Thr	Leu	Ser	Thr	Cys	Ala	Gln
245				250				255							
Ser	Ala	Asp	Tyr	Leu	Ile	Ala	Val	Asp	His	Gly	Ala	Asp	Ala	Cys	His
260				265				270							

Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala
275 280 285

Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met
290 295 300

Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys
305 310 315 320

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

Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu
340 345 350

Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val
355 360 365

Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Thr Ala Gly Gln Val Asp
370 375 380

Ser Trp Gln Leu Ser Ala Thr Asp Val Gly Lys Lys Met Ser Leu Val
385 390 395 400

Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn
405 410 415

Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn
420 425 430

Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu
435 440 445

Leu Val Gln Leu Trp Asn
450

<210> 184

<211> 1365

<212> DNA

<213> Atopobium parvulum

<220><221> CDS

<222> (1)..(1365)

<223> Atopobium parvulum gene encoding TMP phosphatase [WP_035433109]

<400> 184

atg cag gtg acc ggt gca att ttt gat tgc gat gga act ctt gtt gat 48

Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp
1 5 10 15

tca atg cac gtt tgg cac aac gtt ttt ggc gct gtt ctt cct aaa tac 96

Ser Met His Val Trp His Asn Val Phe Gly Ala Val Leu Pro Lys Tyr
20 25 30

ggc aag act att gat tgc gat att ttt gac cgc gta gag gct gtt tcc 144

Gly Lys Thr Ile Asp Ser Asp Ile Phe Asp Arg Val Glu Ala Val Ser
35 40 45

ctc att ggt gga tgt cag att tgc gtt gat gag ctg gat ttg cct att 192

Leu Ile Gly Gly Cys Gln Ile Cys Val Asp Glu Leu Asp Leu Pro Ile
50 55 60

aca gcg gaa gct tta tat gaa gag ttc tgc gcg tac gta act gat cag 240

Thr Ala Glu Ala Leu Tyr Glu Glu Phe Cys Ala Tyr Val Thr Asp Gln
65 70 75 80

tac cga cat cat gtt tca atc att ccc ggt gca aag gag ttc tta cag 288

Tyr Arg His His Val Ser Ile Ile Pro Gly Ala Lys Glu Phe Leu Gln
85 90 95

gaa ctc cac gac gca ggc att cct atg gcc gtt gct tgc tca act ccc 336

Glu Leu His Asp Ala Gly Ile Pro Met Ala Val Ala Ser Ser Thr Pro
100 105 110

gtg cga gaa gtt cgt gca gct ctg gca gct caa ggt att gag cac ctc 384

Val Arg Glu Val Arg Ala Ala Leu Ala Ala Gln Gly Ile Glu His Leu
115 120 125

ttt aaa aca gtg gtc tca acg gaa gat gtg ggg gga gtg gac aag gtt 432

Phe Lys Thr Val Val Ser Thr Glu Asp Val Gly Gly Val Asp Lys Val
130 135 140

gag cca gat gtt tac ctt gag gct ctt cgc cgt ctt ggc act gat aag 480

Glu Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys

145	150	155	160	
gca act acc tgg gtc ttc gag gat gct ccg ttt ggc gca cag aca gca				528
Ala Thr Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Ala Gln Thr Ala				
	165	170	175	
caa aat gca ggc ttt cct gtg gct gta ctc tac aac gac cac gat ggc				576
Gln Asn Ala Gly Phe Pro Val Ala Val Leu Tyr Asn Asp His Asp Gly				
	180	185	190	
cgc gac ccc gtc ttt atg cgc gag cac tct aac atc ttt gcc cac acc				624
Arg Asp Pro Val Phe Met Arg Glu His Ser Asn Ile Phe Ala His Thr				
	195	200	205	
tac ggc gag ctg tcg ctt ctg cgc ctt cag gac tac gag cgc cct ctg				672
Tyr Gly Glu Leu Ser Leu Leu Arg Leu Gln Asp Tyr Glu Arg Pro Leu				
	210	215	220	
acc gca gcg cct tct ggc gag aaa ccc ctt gag gtc ctt gtt gtg ggc				720
Thr Ala Ala Pro Ser Gly Glu Lys Pro Leu Glu Val Leu Val Val Gly				
	225	230	235	240
gga tcc cca gag gcg gtt tcg cac acg acg ctg tct acc tgc gcc caa				768
Gly Ser Pro Glu Ala Val Ser His Thr Thr Leu Ser Thr Cys Ala Gln				
	245	250	255	
agc gct gac tac ctg ata gcg gtt gac cat ggc gca gat gtc tgt cac				816
Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Val Cys His				
	260	265	270	
gct gcc ggc gtg att cca caa ctt gcg ctt gga gac ttt gac tcc gct				864
Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala				
	275	280	285	
aca cca gaa act ctg gct tgg ctc aaa gag cag cag gta cct tgc atg				912
Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met				
	290	295	300	
aag ttt aat gcg gac aag tac gat acc gac ctg gcg cta gca ttg aaa				960
Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys				
	305	310	315	320

tca gct gaa tat gag gca att cgt aga gat agc aag ctc tct ctg acg 1008
 Ser Ala Glu Tyr Glu Ala Ile Arg Arg Asp Ser Lys Leu Ser Leu Thr
 325 330 335
 gtt gtc tcc aca tct ggc ggc cac ctt gat cac cag ctt gta gtg ctt 1056
 Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu
 340 345 350
 ggt ctt ctc gcc acg tgg gca aag acg ggc aag gca agc gtt cga gtt 1104
 Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val
 355 360 365
 att gag aat gac ttt gag atg cgc ttt tta gtt gct ggc cag gtg gat 1152
 Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Val Ala Gly Gln Val Asp
 370 375 380
 tct tgg cag ctg aac act atc aat gta ggt aaa aag att tct ctt gta 1200
 Ser Trp Gln Leu Asn Thr Ile Asn Val Gly Lys Lys Ile Ser Leu Val
 385 390 395 400
 gct ttg tca gag gag tgc gag gtt tct gag gcc ggc atg aag tgg aat 1248
 Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn
 405 410 415
 ctt gat cac cag aag ttc acc ttg ctg gga gac gac ggt att tca aac 1296
 Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn
 420 425 430
 ata gtt gaa tca gac aat tcc tgg gta agg tgc gag aag ggc tgt ctt 1344
 Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu
 435 440 445
 ttg gtg cag ctt tgg aac taa 1365
 Leu Val Gln Leu Trp Asn
 450
 <210> 185
 <211> 454
 <212> PRT
 <213> Atopobium parvulum

<400> 185

Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp

1 5 10 15

Ser Met His Val Trp His Asn Val Phe Gly Ala Val Leu Pro Lys Tyr

20 25 30

Gly Lys Thr Ile Asp Ser Asp Ile Phe Asp Arg Val Glu Ala Val Ser

35 40 45

Leu Ile Gly Gly Cys Gln Ile Cys Val Asp Glu Leu Asp Leu Pro Ile

50 55 60

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

65 70 75 80

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

85 90 95

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

100 105 110

Val Arg Glu Val Arg Ala Ala Leu Ala Ala Gln Gly Ile Glu His Leu

115 120 125

Phe Lys Thr Val Val Ser Thr Glu Asp Val Gly Gly Val Asp Lys Val

130 135 140

Glu Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys

145 150 155 160

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

165 170 175

Gln Asn Ala Gly Phe Pro Val Ala Val Leu Tyr Asn Asp His Asp Gly

180 185 190

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

195 200 205

Tyr Gly Glu Leu Ser Leu Leu Arg Leu Gln Asp Tyr Glu Arg Pro Leu

210 215 220

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

225 230 235 240

Gly Ser Pro Glu Ala Val Ser His Thr Thr Leu Ser Thr Cys Ala Gln
 245 250 255
 Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Val Cys His
 260 265 270
 Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala
 275 280 285
 Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met
 290 295 300
 Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys
 305 310 315 320
 Ser Ala Glu Tyr Glu Ala Ile Arg Arg Asp Ser Lys Leu Ser Leu Thr
 325 330 335
 Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu
 340 345 350
 Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val
 355 360 365
 Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Val Ala Gly Gln Val Asp
 370 375 380
 Ser Trp Gln Leu Asn Thr Ile Asn Val Gly Lys Lys Ile Ser Leu Val
 385 390 395 400
 Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn
 405 410 415
 Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn
 420 425 430
 Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu
 435 440 445
 Leu Val Gln Leu Trp Asn
 450
 <210> 186
 <211> 1383
 <212> DNA
 <213> Atopobium rimae

<220><221> CDS

<222> (1)..(1383)

<223> Atopobium rimae gene encoding TMP phosphatase [WP_003148415]

<400> 186

atg cag ata acg ggt gca atc ttt gat ctt gat ggg aca ctg gtt gac 48

Met Gln Ile Thr Gly Ala Ile Phe Asp Leu Asp Gly Thr Leu Val Asp

1 5 10 15

tcc atg tgg atg tgg aga aga tcg ttc gga gat gtt tta gaa gac ctg 96

Ser Met Trp Met Trp Arg Arg Ser Phe Gly Asp Val Leu Glu Asp Leu

20 25 30

cat atc aat atg act ccg gat ttt ttt aaa agg gtc gag gcc att tcg 144

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

35 40 45

ctt tac gat ggt tgc gta gcg tgt att gag gaa ttt aat ctt cct tta 192

Leu Tyr Asp Gly Cys Val Ala Cys Ile Glu Glu Phe Asn Leu Pro Leu

50 55 60

tcc gca gaa gag ctg tat gaa aag ttc ctt ttg tat gta caa acg gta 240

Ser Ala Glu Glu Leu Tyr Glu Lys Phe Leu Leu Tyr Val Gln Thr Val

65 70 75 80

tat tcg cac gat att aaa agc att gcg ggg gct acc gac ttt ctc cag 288

Tyr Ser His Asp Ile Lys Ser Ile Ala Gly Ala Thr Asp Phe Leu Gln

85 90 95

gaa ctt ttt gac gca gga ata cct ctt gcc att gct tct tct acg cca 336

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

100 105 110

tct cgt gcc ata cat gtt gct ctt gaa gcc caa ggt atg gag aag ttt 384

Ser Arg Ala Ile His Val Ala Leu Glu Ala Gln Gly Met Glu Lys Phe

115 120 125

ttt aaa gcg gtt gtg tgt acc gaa gac gtc ggg ggt gtc gat aaa gca 432

Phe Lys Ala Val Val Cys Thr Glu Asp Val Gly Gly Val Asp Lys Ala

130 135 140

aaa ccc gat gtc tat ctt gag gct ctc aga cgc ctg ggc acc gat aaa 480
Lys Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys

145 150 155 160
gca cac acg tgg gtc ttt gag gac gct gag ttt ggt gta cat acg gca 528
Ala His Thr Trp Val Phe Glu Asp Ala Glu Phe Gly Val His Thr Ala

165 170 175
caa acc gag ggc ttt ccc gtt gtt gcg ctg ttc aat ggc aaa gac ggc 576
Gln Thr Glu Gly Phe Pro Val Val Ala Leu Phe Asn Gly Lys Asp Gly

180 185 190

cgt gat ctt gag tat atg aag gcg cac tct gat ctt ctc gca cat gat 624
Arg Asp Leu Glu Tyr Met Lys Ala His Ser Asp Leu Leu Ala His Asp

195 200 205
tat cga gaa ctc tct ctt gcc cgc att tac gat tat gaa cgg gtg acg 672
Tyr Arg Glu Leu Ser Leu Ala Arg Ile Tyr Asp Tyr Glu Arg Val Thr

210 215 220
aat cag cca cat ctg ggc gcc tca tcg gct cag aag gtc ttt tcg gtt 720

Asn Gln Pro His Leu Gly Ala Ser Ser Ala Gln Lys Val Phe Ser Val

225 230 235 240
ctc gtt gtt gat gga tct ccc acg cca agt tca gcc gcg ctg gtt tca 768
Leu Val Val Asp Gly Ser Pro Thr Pro Ser Ser Ala Ala Leu Val Ser

245 250 255
gaa ctt tca tca tgc tcg gat tat gtc gtt gct gca gat cgc ggg gca 816
Glu Leu Ser Ser Cys Ser Asp Tyr Val Val Ala Ala Asp Arg Gly Ala

260 265 270
tat atc tgc aag gag gcc ggt gtc gtt cct gat att gcg tgc gga gac 864
Tyr Ile Cys Lys Glu Ala Gly Val Val Pro Asp Ile Ala Cys Gly Asp

275 280 285
ttt gat tcc gtg gga gaa gag aca ctc tct tgg atc cat gca caa aag 912
Phe Asp Ser Val Gly Glu Glu Thr Leu Ser Trp Ile His Ala Gln Lys

290	295	300	
gtg cac acg att gct tat cct caa gat aag tac gag acc gat ttg tct			960
Val His Thr Ile Ala Tyr Pro Gln Asp Lys Tyr Glu Thr Asp Leu Ser			
305	310	315	320
ctt gca ctc aat gcc gct tgc cat gaa gca acc cgt caa gca ctt ccg			1008
Leu Ala Leu Asn Ala Ala Cys His Glu Ala Thr Arg Gln Ala Leu Pro			
	325	330	335
ctg tca ctg aca ctt acc tgc gct tcc ggc ggc agg ctt gat cat gag			1056
Leu Ser Leu Thr Leu Thr Cys Ala Ser Gly Gly Arg Leu Asp His Glu			
	340	345	350
ctt ggt gta gtg ggg ctt ctg gct cga tta agc act gcc tca tgg agg			1104
Leu Gly Val Val Gly Leu Leu Ala Arg Leu Ser Thr Ala Ser Trp Arg			
	355	360	365
gtg cgg att gtt gag gat gcc ttt gaa gca agg att ctt tcg gca gat			1152
Val Arg Ile Val Glu Asp Ala Phe Glu Ala Arg Ile Leu Ser Ala Asp			
	370	375	380
acg tat gcg gcg tgg agg ctc tca gaa aaa gat cga gga aag aca ctg			1200
Thr Tyr Ala Ala Trp Arg Leu Ser Glu Lys Asp Arg Gly Lys Thr Leu			
385	390	395	400
tcg gtg ctt ccg ctt cag gaa gaa acg gtg att acc gag atc ggt atg			1248
Ser Val Leu Pro Leu Gln Glu Glu Thr Val Ile Thr Glu Ile Gly Met			
	405	410	415
caa tgg gac ctt gcc tca cga act ttg ctg ctc ctg tct gat gaa gga			1296
Gln Trp Asp Leu Ala Ser Arg Thr Leu Leu Leu Leu Ser Asp Glu Gly			
	420	425	430
att tcc aat gtg gta caa acg gat gtg gca caa ata cat tgc gag aag			1344
Ile Ser Asn Val Val Gln Thr Asp Val Ala Gln Ile His Cys Glu Lys			
	435	440	445
ggc aag gcg ctc gtg gtg ctt ctc gca aat gaa tcg tga			1383

Gly Lys Ala Leu Val Val Leu Leu Ala Asn Glu Ser

450 455 460

<210> 187

<211> 460

<212> PRT

<213> Atopobium rimae

<400> 187

Met Gln Ile Thr Gly Ala Ile Phe Asp Leu Asp Gly Thr Leu Val Asp

1 5 10 15

Ser Met Trp Met Trp Arg Arg Ser Phe Gly Asp Val Leu Glu Asp Leu

20 25 30

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

35 40 45

Leu Tyr Asp Gly Cys Val Ala Cys Ile Glu Glu Phe Asn Leu Pro Leu

50 55 60

Ser Ala Glu Glu Leu Tyr Glu Lys Phe Leu Leu Tyr Val Gln Thr Val

65 70 75 80

Tyr Ser His Asp Ile Lys Ser Ile Ala Gly Ala Thr Asp Phe Leu Gln

85 90 95

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

100 105 110

Ser Arg Ala Ile His Val Ala Leu Glu Ala Gln Gly Met Glu Lys Phe

115 120 125

Phe Lys Ala Val Val Cys Thr Glu Asp Val Gly Gly Val Asp Lys Ala

130 135 140

Lys Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys

145 150 155 160

Ala His Thr Trp Val Phe Glu Asp Ala Glu Phe Gly Val His Thr Ala

165 170 175

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

180 185 190

Arg Asp Leu Glu Tyr Met Lys Ala His Ser Asp Leu Leu Ala His Asp

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

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

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

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

 420 425 430
 Ile Ser Asn Val Val Gln Thr Asp Val Ala Gln Ile His Cys Glu Lys
 435 440 445

Gly Lys Ala Leu Val Val Leu Leu Ala Asn Glu Ser

450 455 460

<210> 188

<211> 1380

<212> DNA

<213> *Olsenella uli*

<220><221> CDS

<222> (1)..(1380)

<223> *Olsenella uli* gene encoding TMP phosphatase [WP_013251930]

<400> 188

atg ccc atc aag gcc gcc atc ttc gac tgt gac gga acg ctg gtc gac 48

Met Pro Ile Lys Ala Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp

1 5 10 15

tcc atg ccc ctg tgg cat gac gtg acg gtc gaa ctg ctg cgc cgc cac 96

Ser Met Pro Leu Trp His Asp Val Thr Val Glu Leu Leu Arg Arg His

20 25 30

cat gtc gcc gac gcc gag gag gcg ttc gtc cgc acc gag tgc ctt ccc 144

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

35 40 45

atg gtc gag atg tgc cat gcc ttc cac gac gag tgg ggc gtt gag gcc 192

Met Val Glu Met Cys His Ala Phe His Asp Glu Trp Gly Val Glu Ala

50 55 60

gag ggc gag gag ctg gtg cgc gag ctg gtc gat atg gtc cgc gag ggg 240

Glu Gly Glu Glu Leu Val Arg Glu Leu Val Asp Met Val Arg Glu Gly

65 70 75 80

tat cgc agc cgg gtt agc ctg ctg ccg ggc tgc cgg gcg ttt ctg gac 288

Tyr Arg Ser Arg Val Ser Leu Leu Pro Gly Cys Arg Ala Phe Leu Asp

85 90 95

gag ctg gcg tct gcg ggc gtc cgc atg gtc gtc gcg tgc tgc acg gct 336

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

100 105 110

ccg gag gag ctc tcc gtc gcg cta tcg gcg cag ggg gtc gac ggc tac 384

Pro Glu Glu Leu Ser Val Ala Leu Ser Ala Gln Gly Val Asp Gly Tyr
115 120 125

ttc gag cgg gtc ttc tcc acg gga ggc ccc ata cgc agc aag gac tac 432

Phe Glu Arg Val Phe Ser Thr Gly Gly Pro Ile Arg Ser Lys Asp Tyr
130 135 140

ccg gac atc tgg gag ctg gtc ctg gac tac ctg ggc acc gac ccg gct 480

Pro Asp Ile Trp Glu Leu Val Leu Asp Tyr Leu Gly Thr Asp Pro Ala

145 150 155 160

gac acc tgg gtc ttc gag gac gcc ccg ttt ggg atg cgg acg gcc cga 528

Asp Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Met Arg Thr Ala Arg
165 170 175

tcg gtc ggc gcc aac acc gtc tgc ctg ttc agc cca cac ggg gac cgc 576

Ser Val Gly Ala Asn Thr Val Cys Leu Phe Ser Pro His Gly Asp Arg
180 185 190

gac ctt gcg gcc tgc gag cgc tac gct gac ata ctg gtc cac agc tac 624

Asp Leu Ala Ala Cys Glu Arg Tyr Ala Asp Ile Leu Val His Ser Tyr
195 200 205

cac gag cta tcg ctc gcc ctg ctg gac gac tac gcc cgt ccg ccg caa 672

His Glu Leu Ser Leu Ala Leu Leu Asp Asp Tyr Ala Arg Pro Pro Gln
210 215 220

gcg tcc ccc tcg gcc cac cct cgc ctc gcg ccg ctt cgc gtc ctc gtc 720

Ala Ser Pro Ser Ala His Pro Arg Leu Ala Pro Leu Arg Val Leu Val
225 230 235 240

gtg ggc gcc tcg ccc gag cgc ccg tct tcg gcg ctg ctc cgc tcc ctg 768

Val Gly Ala Ser Pro Glu Arg Pro Ser Ser Ala Leu Leu Arg Ser Leu
245 250 255

gcc gcc agt acc gac tac gtc atc gcc gcc gac gcc ggg gcc gac gcg 816

Ala Ala Ser Thr Asp Tyr Val Ile Ala Ala Asp Ala Gly Ala Asp Ala

260	265	270	
ctg cgc tcc tgt ggc atc gcc ccc gac gtc ttc tgc ggc gac gcc gac			864
Leu Arg Ser Cys Gly Ile Ala Pro Asp Val Phe Cys Gly Asp Ala Asp			
275	280	285	
tcg gca acg ggc gaa tcg gct gcg tgg gcc cgc tcg gtc gcc cgt gcg			912
Ser Ala Thr Gly Glu Ser Ala Ala Trp Ala Arg Ser Val Ala Arg Ala			
290	295	300	
gac ata gag ttt ccc tcc gag aag tac gcg acc gac ctc gcc ctc gcc			960
Asp Ile Glu Phe Pro Ser Glu Lys Tyr Ala Thr Asp Leu Ala Leu Ala			
305	310	315	320
atc tcc tgc gcc cgc cat gag gcc gct cga cgc aac gcg cgg ctg gag			1008
Ile Ser Cys Ala Arg His Glu Ala Ala Arg Arg Asn Ala Arg Leu Glu			
325	330	335	
ctc acg ctg acc ggc gtc acg ggc ggc agg ccc gac cac gcc ctt gcc			1056
Leu Thr Leu Thr Gly Val Thr Gly Gly Arg Pro Asp His Ala Leu Ala			
340	345	350	
gtc gtg ggt cag ctc gcg cgg aac gct gac gcc tcg ccg cgc atc gtg			1104
Val Val Gly Gln Leu Ala Arg Asn Ala Asp Ala Ser Pro Arg Ile Val			
355	360	365	
gag gac ggc ttc gag tgc cga ctg ctc agc ccc tct ggc act gcg tgc			1152
Glu Asp Gly Phe Glu Cys Arg Leu Leu Ser Pro Ser Gly Thr Ala Cys			
370	375	380	
tgg gag ctg ggt ggg gcc cac gtg cca gcc gcc ggg gtc gag ggg acg			1200
Trp Glu Leu Gly Gly Ala His Val Pro Ala Ala Gly Val Glu Gly Thr			
385	390	395	400
ctc ttc tcg gcc att ccc gtg gca gag ggg acc atg ctc tcc gag cgg			1248
Leu Phe Ser Ala Ile Pro Val Ala Glu Gly Thr Met Leu Ser Glu Arg			
405	410	415	
ggc ttc aag tgg gag ctg gat cat cgt gag ctg ccc ctt ctg ggg gat			1296
Gly Phe Lys Trp Glu Leu Asp His Arg Glu Leu Pro Leu Leu Gly Asp			
420	425	430	

gag gga atc tcg aac gtg gtc acg tcc gcg acg gcc agc gtc gag tgc 1344
 Glu Gly Ile Ser Asn Val Val Thr Ser Ala Thr Ala Ser Val Glu Cys
 435 440 445
 cat gcc ggc gca gtt gcg gcg ttc ctg ttg gca tag 1380

His Ala Gly Ala Val Ala Ala Phe Leu Leu Ala
 450 455

<210> 189

<211> 459

<212> PRT

<213> Olsenella uli

<400> 189

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

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

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

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

165 170 175

Ser Val Gly Ala Asn Thr Val Cys Leu Phe Ser Pro His Gly Asp Arg

180 185 190

Asp Leu Ala Ala Cys Glu Arg Tyr Ala Asp Ile Leu Val His Ser Tyr

195 200 205

His Glu Leu Ser Leu Ala Leu Leu Asp Asp Tyr Ala Arg Pro Pro Gln

210 215 220

Ala Ser Pro Ser Ala His Pro Arg Leu Ala Pro Leu Arg Val Leu Val

225 230 235 240

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

245 250 255

Ala Ala Ser Thr Asp Tyr Val Ile Ala Ala Asp Ala Gly Ala Asp Ala

260 265 270

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

275 280 285

Ser Ala Thr Gly Glu Ser Ala Ala Trp Ala Arg Ser Val Ala Arg Ala

290 295 300

Asp Ile Glu Phe Pro Ser Glu Lys Tyr Ala Thr Asp Leu Ala Leu Ala

305 310 315 320

Ile Ser Cys Ala Arg His Glu Ala Ala Arg Arg Asn Ala Arg Leu Glu

325 330 335

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

340 345 350

Val Val Gly Gln Leu Ala Arg Asn Ala Asp Ala Ser Pro Arg Ile Val

355 360 365

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

370 375 380

Trp Glu Leu Gly Gly Ala His Val Pro Ala Ala Gly Val Glu Gly Thr

385 390 395 400

Leu Phe Ser Ala Ile Pro Val Ala Glu Gly Thr Met Leu Ser Glu Arg

405 410 415
 Gly Phe Lys Trp Glu Leu Asp His Arg Glu Leu Pro Leu Leu Gly Asp

420 425 430
 Glu Gly Ile Ser Asn Val Val Thr Ser Ala Thr Ala Ser Val Glu Cys

435 440 445
 His Ala Gly Ala Val Ala Ala Phe Leu Leu Ala

450 455
 <210> 190
 <211> 762
 <212> DNA
 <213> Atopobium minutum
 <220><221> CDS
 <222> (1)..(762)
 <223> Atopobium minutum gene encoding TMP phosphatase [KRN55115]
 <400> 190
 atg tgg gct aaa acc tct cga cat tgt acg caa aaa ggc ttt acc atg 48

Met Trp Ala Lys Thr Ser Arg His Cys Thr Gln Lys Gly Phe Thr Met
 1 5 10 15
 aac cct gca cgc att tta ttt gat gga gga act tgt atg gca ata agc 96
 Asn Pro Ala Arg Ile Leu Phe Asp Gly Gly Thr Cys Met Ala Ile Ser
 20 25 30
 ggc gca atc ttt gac tgt gac ggc acg ctg gtt gat tct atg tat atg 144
 Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp Ser Met Tyr Met

35 40 45
 tgg tgg gac gcc ttt ccc cgc ctg ctt gcc agc cat ggc ttt gct atg 192
 Trp Trp Asp Ala Phe Pro Arg Leu Leu Ala Ser His Gly Phe Ala Met
 50 55 60
 acg cct cag atc gag aaa atc ttg cat gag tgt gag gcg gtc agc ttg 240
 Thr Pro Gln Ile Glu Lys Ile Leu His Glu Cys Glu Ala Val Ser Leu
 65 70 75 80
 gat gaa gag atc cat acg ctg cgc aac gct ctt gct att ccc gct tct 288

Asp Glu Glu Ile His Thr Leu Arg Asn Ala Leu Ala Ile Pro Ala Ser	
85 90 95	
gcc gag cag cta gca caa gaa tta tcc cag aat att agc aat gcg tat	336
Ala Glu Gln Leu Ala Gln Glu Leu Ser Gln Asn Ile Ser Asn Ala Tyr	
100 105 110	
gcc tca gag atc aaa gca tgg cct gcc gtt aag ccg ttc ttg gat cag	384
Ala Ser Glu Ile Lys Ala Trp Pro Ala Val Lys Pro Phe Leu Asp Gln	
115 120 125	
ctc aaa gac gca ggt atc ccc atg atc att tgt act tct acc gga gcc	432
Leu Lys Asp Ala Gly Ile Pro Met Ile Ile Cys Thr Ser Thr Gly Ala	
130 135 140	
aaa gaa gtt ggt ctg tgc atg gat cat ctt ggt ttg tcc aag ttt ttt	480
Lys Glu Val Gly Leu Cys Met Asp His Leu Gly Leu Ser Lys Phe Phe	
145 150 155 160	
gta gat att gtc agc gcg gaa gaa aac aat ttc acc aaa act gag cca	528
Val Asp Ile Val Ser Ala Glu Glu Asn Asn Phe Thr Lys Thr Glu Pro	
165 170 175	
gat atc tat tac tat gcg cta aaa aag ctt ggt acc act aaa gag aca	576
Asp Ile Tyr Tyr Tyr Ala Leu Lys Lys Leu Gly Thr Thr Lys Glu Thr	
180 185 190	
acc tgg gta ttt gag gat gct ccg ttt ggc ctt act acc tct gag cgt	624
Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Leu Thr Thr Ser Glu Arg	
195 200 205	
gca gga ttt cct aat gtg tgc gtc ttt aat gcg cac gat aag cgc gat	672
Ala Gly Phe Pro Asn Val Cys Val Phe Asn Ala His Asp Lys Arg Asp	
210 215 220	
gag gac ttt ttg cgt ctt cat gct acg ttg ttt acg cac ata tat gag	720
Glu Asp Phe Leu Arg Leu His Ala Thr Leu Phe Thr His Ile Tyr Glu	
225 230 235 240	
gat att tcc ctt gcg gat ttg cag tgc tac ccc acc aag taa	762
Asp Ile Ser Leu Ala Asp Leu Gln Ser Tyr Pro Thr Lys	

245 250

<210> 191

<211> 253

<212> PRT

<213> Atopobium minutum

<400> 191

Met Trp Ala Lys Thr Ser Arg His Cys Thr Gln Lys Gly Phe Thr Met

1 5 10 15

Asn Pro Ala Arg Ile Leu Phe Asp Gly Gly Thr Cys Met Ala Ile Ser

20 25 30

Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp Ser Met Tyr Met

35 40 45

Trp Trp Asp Ala Phe Pro Arg Leu Leu Ala Ser His Gly Phe Ala Met

50 55 60

Thr Pro Gln Ile Glu Lys Ile Leu His Glu Cys Glu Ala Val Ser Leu

65 70 75 80

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

85 90 95

Ala Glu Gln Leu Ala Gln Glu Leu Ser Gln Asn Ile Ser Asn Ala Tyr

100 105 110

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

115 120 125

Leu Lys Asp Ala Gly Ile Pro Met Ile Ile Cys Thr Ser Thr Gly Ala

130 135 140

Lys Glu Val Gly Leu Cys Met Asp His Leu Gly Leu Ser Lys Phe Phe

145 150 155 160

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

165 170 175

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

180 185 190

Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Leu Thr Thr Ser Glu Arg

195	200	205	
Ala Gly Phe Pro Asn Val Cys Val Phe Asn Ala His Asp Lys Arg Asp			
210	215	220	
Glu Asp Phe Leu Arg Leu His Ala Thr Leu Phe Thr His Ile Tyr Glu			
225	230	235	240
Asp Ile Ser Leu Ala Asp Leu Gln Ser Tyr Pro Thr Lys			
245	250		
<210>	192		
<211>	648		
<212>	DNA		
<213>	Syntrophomonas wolfei		
<220>			
><221>	CDS		
<222>	(1)..(648)		
<223>	Syntrophomonas wolfei gene encoding TMP phosphatase		
	[WP_011640074]		
<400>	192		
atg gga gag aaa tta ata att ttt atg gat ttc gat ggc act att tct			48
Met Gly Glu Lys Leu Ile Ile Phe Met Asp Phe Asp Gly Thr Ile Ser			
1 5 10 15			
cgg gag gat gtc tgc aat aag atg gca gcc agg tat gcc ggc agg gac			96
Arg Glu Asp Val Cys Asn Lys Met Ala Ala Arg Tyr Ala Gly Arg Asp			
20 25 30			
tgg gag gaa ata aac cgc ctc tgg gaa gag gga ggt att act act gga			144
Trp Glu Glu Ile Asn Arg Leu Trp Glu Glu Gly Gly Ile Thr Thr Gly			
35 40 45			
gag tgc gcc agt cgt att ctt tca tca atg gag gta ggg gcg gct gaa			192
Glu Cys Ala Ser Arg Ile Leu Ser Ser Met Glu Val Gly Ala Ala Glu			
50 55 60			
ttg gag gcc ttt ttt cag gct cag gaa gta gac ccc ggc ttt tcc cct			240
Leu Glu Ala Phe Phe Gln Ala Gln Glu Val Asp Pro Gly Phe Ser Pro			
65 70 75 80			
ttc ctg gac tgg gta caa aaa aat cag cac ctc ccc att ata ttg agc			288

Phe Leu Asp Trp Val Gln Lys Asn Gln His Leu Pro Ile Ile Leu Ser
85 90 95
gat ggt tat gac cgc tat ata aaa agc ata tta cgg ggc cag ggc tgg 336

Asp Gly Tyr Asp Arg Tyr Ile Lys Ser Ile Leu Arg Gly Gln Gly Trp
100 105 110
gaa atc gag ttt tat gcc aat aaa tta tac tgg gat gac gcc tgg cgg 384
Glu Ile Glu Phe Tyr Ala Asn Lys Leu Tyr Trp Asp Asp Ala Trp Arg
115 120 125
atg gaa tcg ccc tac ctg gat gaa gaa tgc ttt aaa tgt ggg gta tgc 432
Met Glu Ser Pro Tyr Leu Asp Glu Glu Cys Phe Lys Cys Gly Val Cys

130 135 140
aag agc aag ata atc cag gaa aga agt tta ccc ggc tat ctc aca gta 480
Lys Ser Lys Ile Ile Gln Glu Arg Ser Leu Pro Gly Tyr Leu Thr Val
145 150 155 160
tat atc gga gat ggc tac tcc gat ttc tgc ccg gcg gcc tct tgt gat 528
Tyr Ile Gly Asp Gly Tyr Ser Asp Phe Cys Pro Ala Ala Ser Cys Asp
165 170 175

att gtt ttt gcc aaa aat gaa ctg gcc ggc tac tgc cag aaa gag ggt 576
Ile Val Phe Ala Lys Asn Glu Leu Ala Gly Tyr Cys Gln Lys Glu Gly
180 185 190
tta act tac tac ccc tac cgg gat ttt cac gat att ctc cag caa ctg 624
Leu Thr Tyr Tyr Pro Tyr Arg Asp Phe His Asp Ile Leu Gln Gln Leu
195 200 205
ccg agg att gtt agc agg atg tag 648

Pro Arg Ile Val Ser Arg Met
210 215
<210> 193
<211> 215
<212> PRT
<213> Syntrophomonas wolfei
<400> 193

Met Gly Glu Lys Leu Ile Ile Phe Met Asp Phe Asp Gly Thr Ile Ser
1 5 10 15
Arg Glu Asp Val Cys Asn Lys Met Ala Ala Arg Tyr Ala Gly Arg Asp
20 25 30
Trp Glu Glu Ile Asn Arg Leu Trp Glu Glu Gly Gly Ile Thr Thr Gly
35 40 45
Glu Cys Ala Ser Arg Ile Leu Ser Ser Met Glu Val Gly Ala Ala Glu
50 55 60
Leu Glu Ala Phe Phe Gln Ala Gln Glu Val Asp Pro Gly Phe Ser Pro
65 70 75 80
Phe Leu Asp Trp Val Gln Lys Asn Gln His Leu Pro Ile Ile Leu Ser
85 90 95
Asp Gly Tyr Asp Arg Tyr Ile Lys Ser Ile Leu Arg Gly Gln Gly Trp
100 105 110
Glu Ile Glu Phe Tyr Ala Asn Lys Leu Tyr Trp Asp Asp Ala Trp Arg
115 120 125
Met Glu Ser Pro Tyr Leu Asp Glu Glu Cys Phe Lys Cys Gly Val Cys
130 135 140
Lys Ser Lys Ile Ile Gln Glu Arg Ser Leu Pro Gly Tyr Leu Thr Val
145 150 155 160
Tyr Ile Gly Asp Gly Tyr Ser Asp Phe Cys Pro Ala Ala Ser Cys Asp
165 170 175
Ile Val Phe Ala Lys Asn Glu Leu Ala Gly Tyr Cys Gln Lys Glu Gly
180 185 190
Leu Thr Tyr Tyr Pro Tyr Arg Asp Phe His Asp Ile Leu Gln Gln Leu
195 200 205
Pro Arg Ile Val Ser Arg Met
210 215
<210> 194
<211> 666
<212> DNA
<213> Desulfitobacterium hafniense

<220><221> CDS

<222> (1)..(666)

<223>

Desulfitobacterium hafniense gene encoding TMP phosphatase

[WP_018212876]

<400> 194

atg gag gaa ttg aac agt att ttc ttc gtg gat ttt gac ggc acc atc 48

Met Glu Glu Leu Asn Ser Ile Phe Phe Val Asp Phe Asp Gly Thr Ile

1 5 10 15

gtc act cag gat atg tgt gca gtc ctc gtt gaa acc ttg gcc ggg gaa 96

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

20 25 30

gga tgg cgg gag att aat gaa ctt tgg gaa aga aaa gag ctt tcc acc 144

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

35 40 45

ctg gag tgc gcc cgc cgg acc ttt aaa ctc ttt aac agc aat gac ccg 192

Leu Glu Cys Ala Arg Arg Thr Phe Lys Leu Phe Asn Ser Asn Asp Pro

50 55 60

gaa gtt ttt cgc cag ctt atc ggg cag gcg gtg ttc gat ccc gga ttt 240

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

65 70 75 80

tta gat ttt gcc gct ttt tgt gaa cag aga gga ttt ccc ctc atc att 288

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

85 90 95

ctc agc gac gga tat gat ttc tat att gag tac ctc ttg caa aga gag 336

Leu Ser Asp Gly Tyr Asp Phe Tyr Ile Glu Tyr Leu Leu Gln Arg Glu

100 105 110

gga ttg aac ctg cca tac tat gcc aac aaa ttg ctg ttt gct ccc caa 384

Gly Leu Asn Leu Pro Tyr Tyr Ala Asn Lys Leu Leu Phe Ala Pro Gln

115 120 125

ctt gac gta gaa acc ccc tac agc tcc ggc gaa tgt gat cta tgc ggg 432

Leu Asp Val Glu Thr Pro Tyr Ser Ser Gly Glu Cys Asp Leu Cys Gly

130	135	140	
gtc tgc aaa ctg cag ctg atg gaa aaa ttg ctt aaa ccc ggt tgc cga			480
Val Cys Lys Leu Gln Leu Met Glu Lys Leu Leu Lys Pro Gly Cys Arg			
145	150	155	160
tcc gtc tat atc gga gat ggg act tcc gat ttt tgc ccg gcg gaa agg			528
Ser Val Tyr Ile Gly Asp Gly Thr Ser Asp Phe Cys Pro Ala Glu Arg			
	165	170	175
gcg gat aag gtc ttt gcc agg agc agg ctt tat cag cat tgc cag gag			576
Ala Asp Lys Val Phe Ala Arg Ser Arg Leu Tyr Gln His Cys Gln Glu			
	180	185	190
gtg ggc aaa gaa gcc cag cta ttc caa tgc ttt cag gat att ctt cag			624
Val Gly Lys Glu Ala Gln Leu Phe Gln Ser Phe Gln Asp Ile Leu Gln			
	195	200	205
aca gtt gaa cat tgg gga agg gaa gag gag gaa ggg act tga			666
Thr Val Glu His Trp Gly Arg Glu Glu Glu Glu Gly Thr			
210	215	220	
<210> 195			
<211> 221			
<212> PRT			
<213> Desulfitobacterium hafniense			
<400> 195			
Met Glu Glu Leu Asn Ser Ile Phe Phe Val Asp Phe Asp Gly Thr Ile			
1	5	10	15
Val Thr Gln Asp Met Cys Ala Val Leu Val Glu Thr Leu Ala Gly Glu			
	20	25	30
Gly Trp Arg Glu Ile Asn Glu Leu Trp Glu Arg Lys Glu Leu Ser Thr			
	35	40	45
Leu Glu Cys Ala Arg Arg Thr Phe Lys Leu Phe Asn Ser Asn Asp Pro			
50	55	60	
Glu Val Phe Arg Gln Leu Ile Gly Gln Ala Val Phe Asp Pro Gly Phe			
65	70	75	80

Leu Asp Phe Ala Ala Phe Cys Glu Gln Arg Gly Phe Pro Leu Ile Ile
85 90 95

Leu Ser Asp Gly Tyr Asp Phe Tyr Ile Glu Tyr Leu Leu Gln Arg Glu
100 105 110

Gly Leu Asn Leu Pro Tyr Tyr Ala Asn Lys Leu Leu Phe Ala Pro Gln
115 120 125

Leu Asp Val Glu Thr Pro Tyr Ser Ser Gly Glu Cys Asp Leu Cys Gly
130 135 140

Val Cys Lys Leu Gln Leu Met Glu Lys Leu Leu Lys Pro Gly Cys Arg
145 150 155 160

Ser Val Tyr Ile Gly Asp Gly Thr Ser Asp Phe Cys Pro Ala Glu Arg
165 170 175

Ala Asp Lys Val Phe Ala Arg Ser Arg Leu Tyr Gln His Cys Gln Glu
180 185 190

Val Gly Lys Glu Ala Gln Leu Phe Gln Ser Phe Gln Asp Ile Leu Gln
195 200 205

Thr Val Glu His Trp Gly Arg Glu Glu Glu Glu Gly Thr
210 215 220

<210> 196

<211> 642

<212> DNA

<213> Pelotomaculum thermopropionicum

<220><221> CDS

<222> (1)..(642)

<223> Pelotomaculum thermopropionicum gene encoding TMP phosphatase

[WP_012032097]

<400> 196

atg gaa aaa gtt ttt ttt gtt gat ttt gac ggg acg gta acc aaa aag 48

Met Glu Lys Val Phe Phe Val Asp Phe Asp Gly Thr Val Thr Lys Lys

1 5 10 15

gat acc tgc gtg gcc atg atc gag gcc ttt gcc ggc ggc aac tgg aga 96

Asp Thr Cys Val Ala Met Ile Glu Ala Phe Ala Gly Gly Asn Trp Arg

20	25	30	
gag att aac gag gcg tgg gaa aga aaa gaa att tcc acg gaa gaa tgt			144
Glu Ile Asn Glu Ala Trp Glu Arg Lys Glu Ile Ser Thr Glu Glu Cys			
35	40	45	
gca aac atg atc ttc agg ctt ttc cgc gcc ggc att gaa gac atc agg			192
Ala Asn Met Ile Phe Arg Leu Phe Arg Ala Gly Ile Glu Asp Ile Arg			
50	55	60	
aag ctt ttg gac ggt atc gag ata gac ggc cat ttt aaa gat ttt ctt			240
Lys Leu Leu Asp Gly Ile Glu Ile Asp Gly His Phe Lys Asp Phe Leu			
65	70	75	80
tct ttt tgc cgg gaa aga ggc tat aaa ata tac atc ctc agc gac ggt			288
Ser Phe Cys Arg Glu Arg Gly Tyr Lys Ile Tyr Ile Leu Ser Asp Gly			
85	90	95	
tac gac ttt tgc att gag acg gtg ttt aaa aaa cac gga ata gag ctg			336
Tyr Asp Phe Cys Ile Glu Thr Val Phe Lys Lys His Gly Ile Glu Leu			
100	105	110	
ccg tac tat gcc aac aaa atg gtt tac ggc aat ggt ttt aaa ata gaa			384
Pro Tyr Tyr Ala Asn Lys Met Val Tyr Gly Asn Gly Phe Lys Ile Glu			
115	120	125	
tgc ttc agg ccc aac ccg gcc tgc ggt att tgc ggg acc tgc aag acc			432
Cys Phe Arg Pro Asn Pro Ala Cys Gly Ile Cys Gly Thr Cys Lys Thr			
130	135	140	
aag ctg att gag gag ctt aaa ggg gac ggc agc cag gtt att tac att			480
Lys Leu Ile Glu Glu Leu Lys Gly Asp Gly Ser Gln Val Ile Tyr Ile			
145	150	155	160
ggc gac gga tat tcg gac aca tgc ccg gcc atg aaa gcc gat gtg gtt			528
Gly Asp Gly Tyr Ser Asp Thr Cys Pro Ala Met Lys Ala Asp Val Val			
165	170	175	
ttt gcc aag gga gta ttg tac agg cat tgc cgg gaa aac ggc aaa aag			576

Phe Ala Lys Gly Val Leu Tyr Arg His Cys Arg Glu Asn Gly Lys Lys

180	185	190	
gct att tat tat aat aac ttt ggt gat att att aat tat ttt ttc caa	624		
Ala Ile Tyr Tyr Asn Asn Phe Gly Asp Ile Ile Asn Tyr Phe Phe Gln			
195	200	205	
ata aaa aaa agt ttg taa	642		
Ile Lys Lys Ser Leu			
210			

<210> 197

<211> 213

<212> PRT

<213> Pelotomaculum thermopropionicum

<400> 197

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

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

115	120	125
Cys Phe Arg Pro Asn Pro Ala Cys Gly Ile Cys Gly Thr Cys Lys Thr		
130	135	140

Lys Leu Ile Glu Glu Leu Lys Gly Asp Gly Ser Gln Val Ile Tyr Ile
 145 150 155 160
 Gly Asp Gly Tyr Ser Asp Thr Cys Pro Ala Met Lys Ala Asp Val Val
 165 170 175
 Phe Ala Lys Gly Val Leu Tyr Arg His Cys Arg Glu Asn Gly Lys Lys

180 185 190
 Ala Ile Tyr Tyr Asn Asn Phe Gly Asp Ile Ile Asn Tyr Phe Phe Gln
 195 200 205
 Ile Lys Lys Ser Leu

210

<210> 198

<211> 651

<212> DNA

<213> Desulfotomaculum ruminis

<220><221> CDS

<222> (1)..(651)

<223> Desulfotomaculum ruminis gene encoding TMP phosphatase
 [WP_013840216]

<400> 198

atg gaa acc att ctt ttt ctg gat ttt gac ggc acc att acc gag cag 48

Met Glu Thr Ile Leu Phe Leu Asp Phe Asp Gly Thr Ile Thr Glu Gln
 1 5 10 15

gat acc tgc gat atg ctg atg gag cgc tac ggc aat gcg gaa tgt ctg 96

Asp Thr Cys Asp Met Leu Met Glu Arg Tyr Gly Asn Ala Glu Cys Leu
 20 25 30

gaa ttg aac cgg cgc tgg gaa cgc aag gaa att tcc acc atg gaa tgt 144

Glu Leu Asn Arg Arg Trp Glu Arg Lys Glu Ile Ser Thr Met Glu Cys

35 40 45
 gcc cgg cag tcc ttc cgg caa atg cag gta act ccc gag gtt cta aag 192

Ala Arg Gln Ser Phe Arg Gln Met Gln Val Thr Pro Glu Val Leu Lys
 50 55 60

cgg ttg gtg cag gag gtg aag gta gac cct cat ttg aaa gaa ttg ctc 240

Arg	Leu	Val	Gln	Glu	Val	Lys	Val	Asp	Pro	His	Leu	Lys	Glu	Leu	Leu	
65			70			75			80							
cgt	ttc	tgt	gag	cag	gag	aat	tac	ccc	gcc	tat	att	ttg	agc	gat	ggg	288
Arg	Phe	Cys	Glu	Gln	Glu	Asn	Tyr	Pro	Ala	Tyr	Ile	Leu	Ser	Asp	Gly	
85			90			95										
tat	gaa	ccc	atc	att	cag	ggg	gta	ctg	cag	cgg	gaa	gga	ata	aaa	ata	336
Tyr	Glu	Pro	Ile	Ile	Gln	Gly	Val	Leu	Gln	Arg	Glu	Gly	Ile	Lys	Ile	
100			105			110										
tct	tgt	ttt	tgc	aac	ggg	ttg	tcc	ttt	gac	ggc	cag	tac	cgg	gtc	atg	384
Ser	Cys	Phe	Cys	Asn	Gly	Leu	Ser	Phe	Asp	Gly	Gln	Tyr	Arg	Val	Met	
115			120			125										
gcg	cct	cac	tat	aat	ccc	cgg	tgc	ggc	cgg	tgc	gga	acc	tgt	aaa	caa	432
Ala	Pro	His	Tyr	Asn	Pro	Arg	Cys	Gly	Arg	Cys	Gly	Thr	Cys	Lys	Gln	
130			135			140										
aag	ctg	gtg	gaa	cgc	ctg	ggt	cag	ccg	ggc	gcc	cgg	aag	att	ttt	gtg	480
Lys	Leu	Val	Glu	Arg	Leu	Gly	Gln	Pro	Gly	Ala	Arg	Lys	Ile	Phe	Val	
145			150			155			160							
gga	gac	ggg	tat	tcg	gat	ttc	tgt	gcc	gca	gag	tcc	tgc	agt	aag	gtc	528
Gly	Asp	Gly	Tyr	Ser	Asp	Phe	Cys	Ala	Ala	Glu	Ser	Cys	Ser	Lys	Val	
165			170			175										
ttt	gct	aaa	aaa	aat	tta	ttg	aag	tat	tgc	ctg	gaa	aac	cag	att	ccg	576
Phe	Ala	Lys	Lys	Asn	Leu	Leu	Lys	Tyr	Cys	Leu	Glu	Asn	Gln	Ile	Pro	
180			185			190										
gcc	cac	ccc	tat	gaa	acc	ctg	gga	gag	gtt	tta	cag	tgg	ctg	aga	gga	624
Ala	His	Pro	Tyr	Glu	Thr	Leu	Gly	Glu	Val	Leu	Gln	Trp	Leu	Arg	Gly	
195			200			205										
gag	gct	gaa	cat	gga	cat	ccg	gtt	taa								651
Glu	Ala	Glu	His	Gly	His	Pro	Val									
210			215													
<210>	199															
<211>	216															

<212> PRT

<213> Desulfotomaculum ruminis

<400> 199

Met Glu Thr Ile Leu Phe Leu Asp Phe Asp Gly Thr Ile Thr Glu Gln

1 5 10 15

Asp Thr Cys Asp Met Leu Met Glu Arg Tyr Gly Asn Ala Glu Cys Leu

20 25 30

Glu Leu Asn Arg Arg Trp Glu Arg Lys Glu Ile Ser Thr Met Glu Cys

35 40 45

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

50 55 60

Arg Leu Val Gln Glu Val Lys Val Asp Pro His Leu Lys Glu Leu Leu

65 70 75 80

Arg Phe Cys Glu Gln Glu Asn Tyr Pro Ala Tyr Ile Leu Ser Asp Gly

85 90 95

Tyr Glu Pro Ile Ile Gln Gly Val Leu Gln Arg Glu Gly Ile Lys Ile

100 105 110

Ser Cys Phe Cys Asn Gly Leu Ser Phe Asp Gly Gln Tyr Arg Val Met

115 120 125

Ala Pro His Tyr Asn Pro Arg Cys Gly Arg Cys Gly Thr Cys Lys Gln

130 135 140

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

145 150 155 160

Gly Asp Gly Tyr Ser Asp Phe Cys Ala Ala Glu Ser Cys Ser Lys Val

165 170 175

Phe Ala Lys Lys Asn Leu Leu Lys Tyr Cys Leu Glu Asn Gln Ile Pro

180 185 190

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

195 200 205

Glu Ala Glu His Gly His Pro Val

210 215

<210> 200
 <211> 1896
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(1896)
 <223> ThiC gene from E. coli encoding HMP-P synthase
 <400> 200

atg tct gca aca aaa ctg acc cgc cgc gaa caa cgc gcc cgg gcc caa 48
 Met Ser Ala Thr Lys Leu Thr Arg Arg Glu Gln Arg Ala Arg Ala Gln

1 5 10 15
 cat ttt atc gac acc ctg gaa ggc acc gcc ttt ccc aac tca aaa cgc 96
 His Phe Ile Asp Thr Leu Glu Gly Thr Ala Phe Pro Asn Ser Lys Arg

20 25 30
 att tat atc act ggc aca cac ccc ggc gtg cgc gtg ccg atg cgt gag 144
 Ile Tyr Ile Thr Gly Thr His Pro Gly Val Arg Val Pro Met Arg Glu

35 40 45
 atc cag ctt agc ccg acg cta att ggc ggt agc aaa gaa cag ccg cag 192
 Ile Gln Leu Ser Pro Thr Leu Ile Gly Gly Ser Lys Glu Gln Pro Gln

50 55 60
 tac gaa gaa aac gaa gcg att ccg gtc tac gac acc tcc ggc ccg tat 240
 Tyr Glu Glu Asn Glu Ala Ile Pro Val Tyr Asp Thr Ser Gly Pro Tyr

65 70 75 80
 ggt gat ccg cag att gcc att aac gtg cag caa ggg ctg gca aaa cta 288

Gly Asp Pro Gln Ile Ala Ile Asn Val Gln Gln Gly Leu Ala Lys Leu

85 90 95
 cgc cag ccg tgg atc gat gcg cgc ggc gat acc gaa gaa ctt acc gtg 336
 Arg Gln Pro Trp Ile Asp Ala Arg Gly Asp Thr Glu Glu Leu Thr Val

100 105 110
 cgc agt tcc gat tac act aaa gcg cgg ctg gca gat gat ggc ctc gac 384
 Arg Ser Ser Asp Tyr Thr Lys Ala Arg Leu Ala Asp Asp Gly Leu Asp

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

gtg ccg atc ggt aca gtg ccg atc tac cag gcg ctg gag aag gtt aac	912
Val Pro Ile Gly Thr Val Pro Ile Tyr Gln Ala Leu Glu Lys Val Asn	
290 295 300	
ggg atc gcc gaa gat ctt acc tgg gaa gcg ttc cgc gac acg ctg ctg	960
Gly Ile Ala Glu Asp Leu Thr Trp Glu Ala Phe Arg Asp Thr Leu Leu	
305 310 315 320	
gaa cag gcc gag caa ggt gtg gat tac ttc act atc cat gcg ggc gta	1008
Glu Gln Ala Glu Gln Gly Val Asp Tyr Phe Thr Ile His Ala Gly Val	
325 330 335	
ctg ctg cgc tat gtg ccg atg acc gcg aaa cgc ctg acc ggt atc gtc	1056
Leu Leu Arg Tyr Val Pro Met Thr Ala Lys Arg Leu Thr Gly Ile Val	
340 345 350	
tct cgc ggc ggt tcg att atg gcg aaa tgg tgc ctc tcc cat cat cag	1104
Ser Arg Gly Gly Ser Ile Met Ala Lys Trp Cys Leu Ser His His Gln	
355 360 365	
gaa aat ttc ctc tat caa cac ttc cgc gaa att tgt gaa atc tgt gcc	1152
Glu Asn Phe Leu Tyr Gln His Phe Arg Glu Ile Cys Glu Ile Cys Ala	
370 375 380	
gct tat gac gtt tcg ctg tcg ctg ggc gac ggt ctg cgc ccc ggt tct	1200
Ala Tyr Asp Val Ser Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser	
385 390 395 400	
att cag gac gcc aac gat gaa gcg cag ttt gcc gag ctg cat acg ctg	1248
Ile Gln Asp Ala Asn Asp Glu Ala Gln Phe Ala Glu Leu His Thr Leu	
405 410 415	
ggc gaa ctg acc aaa att gcc tgg gaa tat gac gtg cag gtg atg att	1296
Gly Glu Leu Thr Lys Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile	
420 425 430	
gaa ggc cca ggc cac gtg ccg atg cag atg atc cgc cgc aat atg acc	1344
Glu Gly Pro Gly His Val Pro Met Gln Met Ile Arg Arg Asn Met Thr	
435 440 445	
gag gag tta gag cac tgc cac gaa gcg ccg ttt tac act ctg ggg ccg	1392

Glu Glu Leu Glu His Cys His Glu Ala Pro Phe Tyr Thr Leu Gly Pro

450	455	460	
cta act acc gat att gcg ccg ggc tat gac cac ttc acg tcg ggg att			1440
Leu Thr Thr Asp Ile Ala Pro Gly Tyr Asp His Phe Thr Ser Gly Ile			
465	470	475	480
ggg gcg gcg atg att ggc tgg ttt ggc tgc gcg atg ctc tgt tac gta			1488
Gly Ala Ala Met Ile Gly Trp Phe Gly Cys Ala Met Leu Cys Tyr Val			
485	490	495	
acg cca aaa gag cat ctg ggt ctg ccc aat aaa gaa gat gtt aag cag			1536
Thr Pro Lys Glu His Leu Gly Leu Pro Asn Lys Glu Asp Val Lys Gln			
500	505	510	
ggg ctt atc acc tat aag att gct gcc cac gcc gct gac ctg gcg aaa			1584
Gly Leu Ile Thr Tyr Lys Ile Ala Ala His Ala Ala Asp Leu Ala Lys			
515	520	525	
ggg cat ccg ggc gcg caa att cgc gat aac gcc atg tcg aaa gcc cgc			1632
Gly His Pro Gly Ala Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg			
530	535	540	
ttc gaa ttt cgc tgg gaa gac cag ttt aat ctg gcc ctc gac ccg ttt			1680
Phe Glu Phe Arg Trp Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe			
545	550	555	560
acc gcc cgc gct tat cac gat gaa acc ctg ccg caa gag tca ggt aaa			1728
Thr Ala Arg Ala Tyr His Asp Glu Thr Leu Pro Gln Glu Ser Gly Lys			
565	570	575	
gtc gcc cat ttt tgc tcc atg tgt ggg ccg aaa ttc tgc tcg atg aaa			1776
Val Ala His Phe Cys Ser Met Cys Gly Pro Lys Phe Cys Ser Met Lys			
580	585	590	
atc agc cag gaa gtg cgt gat tac gcc gcc acg caa act att gaa atg			1824
Ile Ser Gln Glu Val Arg Asp Tyr Ala Ala Thr Gln Thr Ile Glu Met			
595	600	605	

gga atg gcg gat atg tcg gag aac ttc cgt gcc aga ggc gga gaa atc 1872
 Gly Met Ala Asp Met Ser Glu Asn Phe Arg Ala Arg Gly Gly Glu Ile
 610 615 620

tac ctg cgt aag gag gaa gcg tga 1896
 Tyr Leu Arg Lys Glu Glu Ala
 625 630

<210> 201
 <211> 631
 <212> PRT
 <213> Escherichia coli

<400> 201

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

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

Glu Leu Arg Phe Ser Gly Val Leu Thr Pro Lys Arg Ala Lys Ala Gly
 130 135 140
 Arg Arg Val Thr Gln Leu His Tyr Ala Arg Gln Gly Ile Ile Thr Pro
 145 150 155 160
 Glu Met Glu Phe Ile Ala Ile Arg Glu Asn Met Gly Arg Glu Arg Ile

165 170 175
 Arg Ser Glu Val Leu Arg His Gln His Pro Gly Met Ser Phe Gly Ala
 180 185 190

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

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

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

 Ala Tyr Asp Val Ser Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser
 385 390 395 400
 Ile Gln Asp Ala Asn Asp Glu Ala Gln Phe Ala Glu Leu His Thr Leu
 405 410 415

Gly Glu Leu Thr Lys Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile
420 425 430

Glu Gly Pro Gly His Val Pro Met Gln Met Ile Arg Arg Asn Met Thr
435 440 445

Glu Glu Leu Glu His Cys His Glu Ala Pro Phe Tyr Thr Leu Gly Pro
450 455 460

Leu Thr Thr Asp Ile Ala Pro Gly Tyr Asp His Phe Thr Ser Gly Ile
465 470 475 480

Gly Ala Ala Met Ile Gly Trp Phe Gly Cys Ala Met Leu Cys Tyr Val
485 490 495

Thr Pro Lys Glu His Leu Gly Leu Pro Asn Lys Glu Asp Val Lys Gln
500 505 510

Gly Leu Ile Thr Tyr Lys Ile Ala Ala His Ala Ala Asp Leu Ala Lys
515 520 525

Gly His Pro Gly Ala Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg
530 535 540

Phe Glu Phe Arg Trp Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe
545 550 555 560

Thr Ala Arg Ala Tyr His Asp Glu Thr Leu Pro Gln Glu Ser Gly Lys
565 570 575

Val Ala His Phe Cys Ser Met Cys Gly Pro Lys Phe Cys Ser Met Lys
580 585 590

Ile Ser Gln Glu Val Arg Asp Tyr Ala Ala Thr Gln Thr Ile Glu Met
595 600 605

Gly Met Ala Asp Met Ser Glu Asn Phe Arg Ala Arg Gly Gly Glu Ile
610 615 620

Tyr Leu Arg Lys Glu Glu Ala
625 630

<210> 202

<211> 1371

<212> DNA

<213> Synechococcus elongatus

<220><221> CDS

<222> (1)..(1371)

<223> ThiC gene from Synechococcus_elongatus_PCC_7942:[NC_007604]
encoding a HMP-P synthase

<400> 202

atg cgc agc gac tgg atc gca ccc cgc cga ggc caa gcc aac gtc act 48

Met Arg Ser Asp Trp Ile Ala Pro Arg Arg Gly Gln Ala Asn Val Thr

1 5 10 15

caa atg cac tac gcc cgc caa ggc gtg atc acc gaa gaa atg gac ttc 96

Gln Met His Tyr Ala Arg Gln Gly Val Ile Thr Glu Glu Met Asp Phe

20 25 30

gtg gcg cgg cgc gaa aat ctg cca gcc gat cta att cgg gat gaa gtg 144

Val Ala Arg Arg Glu Asn Leu Pro Ala Asp Leu Ile Arg Asp Glu Val

35 40 45

gca cgg ggt cgg atg att atc ccc gcc aac atc aac cac acc aat ttg 192

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

50 55 60

gag ccg atg gcg atc ggc att gcc tcc aag tgc aag gtc aac gcc aac 240

Glu Pro Met Ala Ile Gly Ile Ala Ser Lys Cys Lys Val Asn Ala Asn

65 70 75 80

atc ggt get tcg cct aac gcc tcc aac atc gat gaa gaa gtc gag aag 288

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

85 90 95

ctg aag ctc gcg gtc aaa tac ggt gcc gat acc gtc atg gac ctc tcg 336

Leu Lys Leu Ala Val Lys Tyr Gly Ala Asp Thr Val Met Asp Leu Ser

100 105 110

acc ggc ggc ggc aac ctc gat gag att cgc acc gcg atc atc aat gct 384

Thr Gly Gly Gly Asn Leu Asp Glu Ile Arg Thr Ala Ile Ile Asn Ala

115 120 125

tcg ccg gta ccg atc ggc acc gtg ccg gtc tac caa gcc ctg gaa tcc 432

Ser Pro Val Pro Ile Gly Thr Val Pro Val Tyr Gln Ala Leu Glu Ser

130	135	140	
gtt cac ggg cgc atc gaa aaa ctc agc gcc gac gac ttc ttg cat gtg			480
Val His Gly Arg Ile Glu Lys Leu Ser Ala Asp Asp Phe Leu His Val			
145	150	155	160
atc gaa aag cac tgc gaa cag ggc gtc gac tac caa acc atc cac gcc			528
Ile Glu Lys His Cys Glu Gln Gly Val Asp Tyr Gln Thr Ile His Ala			
165	170	175	
ggg ctg ctg att gaa cac ctg ccc aag gtc aag agc cgg atc acc ggg			576
Gly Leu Leu Ile Glu His Leu Pro Lys Val Lys Ser Arg Ile Thr Gly			
180	185	190	
att gtt tcg cgg ggc ggc ggc atc att gcc cag tgg atg ctc tac cac			624
Ile Val Ser Arg Gly Gly Gly Ile Ile Ala Gln Trp Met Leu Tyr His			
195	200	205	
cac aag caa aac ccg ctc tat acc cac ttt cgc gac atc atc gaa atc			672
His Lys Gln Asn Pro Leu Tyr Thr His Phe Arg Asp Ile Ile Glu Ile			
210	215	220	
ttc aag cgc tac gac tgt agc ttc agc ttg ggt gac tcg ctg cgg ccg			720
Phe Lys Arg Tyr Asp Cys Ser Phe Ser Leu Gly Asp Ser Leu Arg Pro			
225	230	235	240
ggg tgc ctg cac gat gct agc gac gat gcc cag ctc agc gag ctg aag			768
Gly Cys Leu His Asp Ala Ser Asp Asp Ala Gln Leu Ser Glu Leu Lys			
245	250	255	
act ctc ggt caa ctg acg cgg gtt gct tgg gaa cac gac gtg caa gtc			816
Thr Leu Gly Gln Leu Thr Arg Val Ala Trp Glu His Asp Val Gln Val			
260	265	270	
atg gtc gaa ggg cca ggc cac gtt ccc atg gac cag atc gag ttc aac			864
Met Val Glu Gly Pro Gly His Val Pro Met Asp Gln Ile Glu Phe Asn			
275	280	285	

gtc cgc aag caa atg gaa gag tgc tca gaa gct ccc ttc tac gtc ttg	912
Val Arg Lys Gln Met Glu Glu Cys Ser Glu Ala Pro Phe Tyr Val Leu	
290 295 300	
ggt ccc ctc gtg acc gac att gca ccg ggc tat gac cac atc acc agc	960
Gly Pro Leu Val Thr Asp Ile Ala Pro Gly Tyr Asp His Ile Thr Ser	
305 310 315 320	
gcg atc ggg gca gca atg gcg ggc tgg tat ggc acg gca atg ctc tgc	1008
Ala Ile Gly Ala Ala Met Ala Gly Trp Tyr Gly Thr Ala Met Leu Cys	
325 330 335	
tac gtc acg ccc aaa gag cac ttg ggt ctg ccc aat gcg gaa gat gtg	1056
Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Ala Glu Asp Val	
340 345 350	
cgc aat ggt ttg atc gcc tac aaa att gcg gct cat gca gca gat atc	1104
Arg Asn Gly Leu Ile Ala Tyr Lys Ile Ala Ala His Ala Ala Asp Ile	
355 360 365	
gct cgc cac cgt ccg ggt gct cgc gat cgc gat gat gaa ctg agt cgg	1152
Ala Arg His Arg Pro Gly Ala Arg Asp Arg Asp Asp Glu Leu Ser Arg	
370 375 380	
gca cgc tac gcc ttc gac tgg aac aag caa ttt gac ttg agc ctc gat	1200
Ala Arg Tyr Ala Phe Asp Trp Asn Lys Gln Phe Asp Leu Ser Leu Asp	
385 390 395 400	
cca gag cgg gcg cgg gaa tac cac gac gaa act ctg cca gca gat atc	1248
Pro Glu Arg Ala Arg Glu Tyr His Asp Glu Thr Leu Pro Ala Asp Ile	
405 410 415	
tac aaa acg gca gaa ttc tgt tgc atg tgt gga ccg aag cac tgt ccg	1296
Tyr Lys Thr Ala Glu Phe Cys Ser Met Cys Gly Pro Lys His Cys Pro	
420 425 430	
atg caa acc aag atc acc gag gaa gat cta acc gag ttg gaa aaa ttc	1344
Met Gln Thr Lys Ile Thr Glu Glu Asp Leu Thr Glu Leu Glu Lys Phe	
435 440 445	
ctc gag aaa gat agc gct ctg gcg tag	1371

Leu Glu Lys Asp Ser Ala Leu Ala

450

455

<210> 203

<211> 456

<212> PRT

<213> *Synechococcus elongatus*

<400> 203

Met Arg Ser Asp Trp Ile Ala Pro Arg Arg Gly Gln Ala Asn Val Thr

1

5

10

15

Gln Met His Tyr Ala Arg Gln Gly Val Ile Thr Glu Glu Met Asp Phe

20

25

30

Val Ala Arg Arg Glu Asn Leu Pro Ala Asp Leu Ile Arg Asp Glu Val

35

40

45

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

50

55

60

Glu Pro Met Ala Ile Gly Ile Ala Ser Lys Cys Lys Val Asn Ala Asn

65

70

75

80

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

85

90

95

Leu Lys Leu Ala Val Lys Tyr Gly Ala Asp Thr Val Met Asp Leu Ser

100

105

110

Thr Gly Gly Gly Asn Leu Asp Glu Ile Arg Thr Ala Ile Ile Asn Ala

115

120

125

Ser Pro Val Pro Ile Gly Thr Val Pro Val Tyr Gln Ala Leu Glu Ser

130

135

140

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

145

150

155

160

Ile Glu Lys His Cys Glu Gln Gly Val Asp Tyr Gln Thr Ile His Ala

165

170

175

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

180

185

190

Ile Val Ser Arg Gly Gly Gly Ile Ile Ala Gln Trp Met Leu Tyr His

195 200 205

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

210 215 220

Phe Lys Arg Tyr Asp Cys Ser Phe Ser Leu Gly Asp Ser Leu Arg Pro

225 230 235 240

Gly Cys Leu His Asp Ala Ser Asp Asp Ala Gln Leu Ser Glu Leu Lys

245 250 255

Thr Leu Gly Gln Leu Thr Arg Val Ala Trp Glu His Asp Val Gln Val

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

Ala Ile Gly Ala Ala Met Ala Gly Trp Tyr Gly Thr Ala Met Leu Cys

325 330 335

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

340 345 350

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

355 360 365

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

370 375 380

Ala Arg Tyr Ala Phe Asp Trp Asn Lys Gln Phe Asp Leu Ser Leu Asp

385 390 395 400

Pro Glu Arg Ala Arg Glu Tyr His Asp Glu Thr Leu Pro Ala Asp Ile

405 410 415

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

420 425 430

Met Gln Thr Lys Ile Thr Glu Glu Asp Leu Thr Glu Leu Glu Lys Phe

435 440 445
 Leu Glu Lys Asp Ser Ala Leu Ala
 450 455

<210> 204
 <211> 1764
 <212> DNA
 <213> Corynebacterium glutamicum
 <220><221> CDS
 <222> (1)..(1764)
 <223> ThiC gene from Corynebacterium glutamicum encoding an HMP-P
 synthase

<400> 204

atg acg cct acc caa aat gag atc cac ccg aaa cat agc tac tcc ccc 48
 Met Thr Pro Thr Gln Asn Glu Ile His Pro Lys His Ser Tyr Ser Pro
 1 5 10 15
 atc cgc aag gac ggt ctc gag gtc ccg gag acc gaa atc cgc ctc gat 96

Ile Arg Lys Asp Gly Leu Glu Val Pro Glu Thr Glu Ile Arg Leu Asp
 20 25 30
 gac tcg cca agc ggc ccc aac gaa ccc ttc cgc atc tac cgc acc cgt 144
 Asp Ser Pro Ser Gly Pro Asn Glu Pro Phe Arg Ile Tyr Arg Thr Arg
 35 40 45
 ggc cca gaa acc aac ccc aag cag gga ctt ccg cgg ctg cgc gag tca 192
 Gly Pro Glu Thr Asn Pro Lys Gln Gly Leu Pro Arg Leu Arg Glu Ser

50 55 60
 tgg atc acc gcc cgc ggc gac gtt gcc acc tat cag ggg cgc gag cgt 240
 Trp Ile Thr Ala Arg Gly Asp Val Ala Thr Tyr Gln Gly Arg Glu Arg
 65 70 75 80
 ttg ctt atc gac gac ggc cgc tcg gca atg cgt cga ggt caa gct tcg 288
 Leu Leu Ile Asp Asp Gly Arg Ser Ala Met Arg Arg Gly Gln Ala Ser
 85 90 95
 gct gag tgg aaa ggc caa aaa cca gct cct ttg aag gcg cta cct ggc 336

Ala Glu Trp Lys Gly Gln Lys Pro Ala Pro Leu Lys Ala Leu Pro Gly	
100 105 110	
aaa aga gtc acc caa atg gcc tat gca cgt gct ggc gtg att act cgt	384
Lys Arg Val Thr Gln Met Ala Tyr Ala Arg Ala Gly Val Ile Thr Arg	
115 120 125	
gaa atg gag ttt gta gcg ctg cgc gaa cac gtt gat gcg gag ttt gtg	432
Glu Met Glu Phe Val Ala Leu Arg Glu His Val Asp Ala Glu Phe Val	
130 135 140	
cgc tct gag gtg gcg cgc ggt cgg gcc att att ccc aac aac gtc aac	480
Arg Ser Glu Val Ala Arg Gly Arg Ala Ile Ile Pro Asn Asn Val Asn	
145 150 155 160	
cac ccc gaa tct gaa ccg atg att att ggt cgc aaa ttt ttg acc aaa	528
His Pro Glu Ser Glu Pro Met Ile Ile Gly Arg Lys Phe Leu Thr Lys	
165 170 175	
atc aac gcc aat att ggc aat tct gcg gtc acc tct tca atc gag gaa	576
Ile Asn Ala Asn Ile Gly Asn Ser Ala Val Thr Ser Ser Ile Glu Glu	
180 185 190	
gag gtg tcc aag ctg cag tgg gcc acg cgc tgg ggt gcc gat acc gtg	624
Glu Val Ser Lys Leu Gln Trp Ala Thr Arg Trp Gly Ala Asp Thr Val	
195 200 205	
atg gat cta tcc acc ggc gat gat att cac acc acc cgc gaa tgg att	672
Met Asp Leu Ser Thr Gly Asp Asp Ile His Thr Thr Arg Glu Trp Ile	
210 215 220	
atc cgc aac tcc ccc gtt cct atc ggc acc gtc ccg atc tac caa gcg	720
Ile Arg Asn Ser Pro Val Pro Ile Gly Thr Val Pro Ile Tyr Gln Ala	
225 230 235 240	
ctg gaa aaa gta aat ggc gtg gcc gca gac ctt aac tgg gaa gta ttc	768
Leu Glu Lys Val Asn Gly Val Ala Ala Asp Leu Asn Trp Glu Val Phe	
245 250 255	
cgc gat acc atc att gag cag tgt gaa caa ggc gtg gac tat atg acc	816
Arg Asp Thr Ile Ile Glu Gln Cys Glu Gln Gly Val Asp Tyr Met Thr	

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

420	425	430	
atg ctg tgt tat gtc acc ccg aaa gaa cac ctt ggc ctg ccc aac cgt			1344
Met Leu Cys Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Arg			
435	440	445	
gac gac gtc aaa acc ggc gta atc acc tac aag ctc gct gcc cac gca			1392
Asp Asp Val Lys Thr Gly Val Ile Thr Tyr Lys Leu Ala Ala His Ala			
450	455	460	
gca gat gtg gcc aag ggt cat ccc ggc gcg cgt gcc tgg gac gac gcc			1440
Ala Asp Val Ala Lys Gly His Pro Gly Ala Arg Ala Trp Asp Asp Ala			
465	470	475	480
atg agt aaa gcg cgt ttt gaa ttc cgt tgg aat gat cag ttt gcg ctc			1488
Met Ser Lys Ala Arg Phe Glu Phe Arg Trp Asn Asp Gln Phe Ala Leu			
485	490	495	
tcc ctc gac ccc gac act gca atc gct tat cac gac gaa acc ctg ccg			1536
Ser Leu Asp Pro Asp Thr Ala Ile Ala Tyr His Asp Glu Thr Leu Pro			
500	505	510	
gca gag cct gcg aaa acc gca cac ttc tgt tca atg tgt ggc ccg aag			1584
Ala Glu Pro Ala Lys Thr Ala His Phe Cys Ser Met Cys Gly Pro Lys			
515	520	525	
ttc tgc tcc atg cga att agc cag gac att cgc gat atg ttt ggc gat			1632
Phe Cys Ser Met Arg Ile Ser Gln Asp Ile Arg Asp Met Phe Gly Asp			
530	535	540	
caa atc gcg gaa ttg ggg atg cct ggg gtt ggg gat tct tct agt gct			1680
Gln Ile Ala Glu Leu Gly Met Pro Gly Val Gly Asp Ser Ser Ser Ala			
545	550	555	560
gtt gct tct agt ggg gca cgg gag ggg atg gct gag aaa tcc cgg gaa			1728
Val Ala Ser Ser Gly Ala Arg Glu Gly Met Ala Glu Lys Ser Arg Glu			
565	570	575	
ttt att gct ggt ggt gcg gag gtt tat cgg cgt tag			1764

Phe Ile Ala Gly Gly Ala Glu Val Tyr Arg Arg

580

585

<210> 205

<211> 587

<212> PRT

<213> Corynebacterium glutamicum

<400> 205

Met Thr Pro Thr Gln Asn Glu Ile His Pro Lys His Ser Tyr Ser Pro

1 5 10 15

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

20 25 30

Asp Ser Pro Ser Gly Pro Asn Glu Pro Phe Arg Ile Tyr Arg Thr Arg

35

40

45

Gly Pro Glu Thr Asn Pro Lys Gln Gly Leu Pro Arg Leu Arg Glu Ser

50 55 60

Trp Ile Thr Ala Arg Gly Asp Val Ala Thr Tyr Gln Gly Arg Glu Arg

65 70 75 80

Leu Leu Ile Asp Asp Gly Arg Ser Ala Met Arg Arg Gly Gln Ala Ser

85 90 95

Ala Glu Trp Lys Gly Gln Lys Pro Ala Pro Leu Lys Ala Leu Pro Gly

100

105

110

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

115 120 125

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

130 135 140

Arg Ser Glu Val Ala Arg Gly Arg Ala Ile Ile Pro Asn Asn Val Asn

145 150 155 160

His Pro Glu Ser Glu Pro Met Ile Ile Gly Arg Lys Phe Leu Thr Lys

165

170

175

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

180 185 190

Glu Val Ser Lys Leu Gln Trp Ala Thr Arg Trp Gly Ala Asp Thr Val

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

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

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

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

 420 425 430
 Met Leu Cys Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Arg
 435 440 445

Asp Asp Val Lys Thr Gly Val Ile Thr Tyr Lys Leu Ala Ala His Ala
450 455 460

Ala Asp Val Ala Lys Gly His Pro Gly Ala Arg Ala Trp Asp Asp Ala
465 470 475 480

Met Ser Lys Ala Arg Phe Glu Phe Arg Trp Asn Asp Gln Phe Ala Leu

485 490 495

Ser Leu Asp Pro Asp Thr Ala Ile Ala Tyr His Asp Glu Thr Leu Pro
500 505 510

Ala Glu Pro Ala Lys Thr Ala His Phe Cys Ser Met Cys Gly Pro Lys
515 520 525

Phe Cys Ser Met Arg Ile Ser Gln Asp Ile Arg Asp Met Phe Gly Asp
530 535 540

Gln Ile Ala Glu Leu Gly Met Pro Gly Val Gly Asp Ser Ser Ser Ala

545 550 555 560

Val Ala Ser Ser Gly Ala Arg Glu Gly Met Ala Glu Lys Ser Arg Glu
565 570 575

Phe Ile Ala Gly Gly Ala Glu Val Tyr Arg Arg
580 585

<210> 206

<211> 1869

<212> DNA

<213> Candidatus Baumannia cicadellinicola

<220><221> CDS

<222> (1)..(1869)

<223> ThiC gene from Candidatus Baumannia cicadellinicola
[WP_011520252] encoding HMP-P synthase

<400> 206

atg tca aga tca tca ata cct gct tca cgc cga gtg agc cgt gca aaa 48

Met Ser Arg Ser Ser Ile Pro Ala Ser Arg Arg Val Ser Arg Ala Lys

1 5 10 15

gca cag gct ttt atg gat agc tta aca ggt agt agc tat ttt cct aac 96

Ala Gln Ala Phe Met Asp Ser Leu Thr Gly Ser Ser Tyr Phe Pro Asn

20	25	30	
tca aga agg ata tat tta caa ggt aaa aca cct tca gta cat gta cca			144
Ser Arg Arg Ile Tyr Leu Gln Gly Lys Thr Pro Ser Val His Val Pro			
35	40	45	
atg cgt gaa att aag cta cat cct aca ttg atc ggt aaa aac ggt gaa			192
Met Arg Glu Ile Lys Leu His Pro Thr Leu Ile Gly Lys Asn Gly Glu			
50	55	60	
cat tat gag gat aat caa cct ata cca gtt tat gat act tca ggt cct			240
His Tyr Glu Asp Asn Gln Pro Ile Pro Val Tyr Asp Thr Ser Gly Pro			
65	70	75	80
tac ggt gat cct act ata gca att aac gta cgt aca ggt ctt aac cgg			288
Tyr Gly Asp Pro Thr Ile Ala Ile Asn Val Arg Thr Gly Leu Asn Arg			
85	90	95	
tta cgc gag ata tgg att ctt gca cga caa gat agt gag cca ata agt			336
Leu Arg Glu Ile Trp Ile Leu Ala Arg Gln Asp Ser Glu Pro Ile Ser			
100	105	110	
aat aat aat aac gat cgt cag agt tca gat aaa cag tta agt ttt act			384
Asn Asn Asn Asn Asp Arg Gln Ser Ser Asp Lys Gln Leu Ser Phe Thr			
115	120	125	
act aac tat aat cca cgc cga gct agc tat gga cgc tgt att aca caa			432
Thr Asn Tyr Asn Pro Arg Arg Ala Ser Tyr Gly Arg Cys Ile Thr Gln			
130	135	140	
tta cat tac gca cgt gcc ggt atc ata acg cca gaa atg gag ttt ata			480
Leu His Tyr Ala Arg Ala Gly Ile Ile Thr Pro Glu Met Glu Phe Ile			
145	150	155	160
gct tta cgt gaa aat atg ggc cga gaa cgt att agt agc aac gtg cta			528
Ala Leu Arg Glu Asn Met Gly Arg Glu Arg Ile Ser Ser Asn Val Leu			
165	170	175	
cat cag cag cat tta ggt tct aac ttt ggt gct aaa aaa gct gat cat			576

His Gln Gln His Leu Gly Ser Asn Phe Gly Ala Lys Lys Ala Asp His

180 185 190
att aca gca gaa ttt gtc cgg cag gaa gta gca gca gga cgt gct att 624
Ile Thr Ala Glu Phe Val Arg Gln Glu Val Ala Ala Gly Arg Ala Ile

195 200 205
ata cct agt aat att aat cat cca gaa tct gag cca atg atc att ggc 672
Ile Pro Ser Asn Ile Asn His Pro Glu Ser Glu Pro Met Ile Ile Gly

210 215 220
cgt aat ttt ctc gta aaa gta aat gca aat att ggt aac tca gca gta 720
Arg Asn Phe Leu Val Lys Val Asn Ala Asn Ile Gly Asn Ser Ala Val
225 230 235 240

aca tct tct att gag gaa gaa gtc gaa aag tta gta tgg gct act cgt 768
Thr Ser Ser Ile Glu Glu Glu Val Glu Lys Leu Val Trp Ala Thr Arg
245 250 255
tgg gga gct gat aca gtc atg gac tta tct act ggt agt tat att cac 816

Trp Gly Ala Asp Thr Val Met Asp Leu Ser Thr Gly Ser Tyr Ile His
260 265 270
gaa act aga gaa tgg ata tta cgt aat agc cca gta cct ata ggt act 864
Glu Thr Arg Glu Trp Ile Leu Arg Asn Ser Pro Val Pro Ile Gly Thr
275 280 285
gta cct atc tat caa gcg tta gaa aaa gta aat gga gtc ata gaa aat 912
Val Pro Ile Tyr Gln Ala Leu Glu Lys Val Asn Gly Val Ile Glu Asn

290 295 300
ctt aat tgg gat att ttc tac gag aca tta tta gaa caa gct aac caa 960
Leu Asn Trp Asp Ile Phe Tyr Glu Thr Leu Leu Glu Gln Ala Asn Gln
305 310 315 320
gga gta gat tat ttt acg att cat gct ggc gta tta aaa cgt tat gtt 1008
Gly Val Asp Tyr Phe Thr Ile His Ala Gly Val Leu Lys Arg Tyr Val
325 330 335

cta cta aca gct agt agg tta act ggt atc gta tcg cgt ggt ggc tct	1056
Leu Leu Thr Ala Ser Arg Leu Thr Gly Ile Val Ser Arg Gly Gly Ser	
340 345 350	
att atg gct caa tgg agt tta gta cat aat cag gaa aac ttc ctt tat	1104
Ile Met Ala Gln Trp Ser Leu Val His Asn Gln Glu Asn Phe Leu Tyr	
355 360 365	
gag cat ttt agt gaa att tgc aag ctt tgt gct gct tat gat att gct	1152
Glu His Phe Ser Glu Ile Cys Lys Leu Cys Ala Ala Tyr Asp Ile Ala	
370 375 380	
cta tct ctt gga gat ggt cta aga ccc ggt tcc gta caa gat gct aat	1200
Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser Val Gln Asp Ala Asn	
385 390 395 400	
gat gaa gca caa ttt tct gag tta cat aca cta ggc gaa tta act aaa	1248
Asp Glu Ala Gln Phe Ser Glu Leu His Thr Leu Gly Glu Leu Thr Lys	
405 410 415	
att gcc tgg gaa tat gat gtg caa gta atg atc gaa gga cct ggt cat	1296
Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile Glu Gly Pro Gly His	
420 425 430	
att cca cta cat atg att gag cgt aat atg act gat caa ctt aaa tat	1344
Ile Pro Leu His Met Ile Glu Arg Asn Met Thr Asp Gln Leu Lys Tyr	
435 440 445	
tgc cac gaa gca cca ttc tac act ctc gga cca ctc aca aca gat att	1392
Cys His Glu Ala Pro Phe Tyr Thr Leu Gly Pro Leu Thr Thr Asp Ile	
450 455 460	
gct cct ggt tat gac cac ttt act tca ggt att ggt gcc gca cta ata	1440
Ala Pro Gly Tyr Asp His Phe Thr Ser Gly Ile Gly Ala Ala Leu Ile	
465 470 475 480	
ggc tgg ttt gga tgt gct atg ctg tgc tat gta act cct aaa gag cat	1488
Gly Trp Phe Gly Cys Ala Met Leu Cys Tyr Val Thr Pro Lys Glu His	
485 490 495	
cta ggt tta cct aat aag gaa gac gta aaa cag ggt tta att gcc tat	1536

Leu Gly Leu Pro Asn Lys Glu Asp Val Lys Gln Gly Leu Ile Ala Tyr
500 505 510
aaa att gcc gca cat gct gca gat cta gct aaa gga cat cct ggt gct 1584
Lys Ile Ala Ala His Ala Ala Asp Leu Ala Lys Gly His Pro Gly Ala
515 520 525
caa ata cgt gat aat gct atg tca aaa gct cgt ttc gaa ttt cgc tgg 1632
Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg Phe Glu Phe Arg Trp
530 535 540
gaa gat caa ttt aac tta gct tta gat cct ttt acg gcg cgt atg tat 1680
Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe Thr Ala Arg Met Tyr
545 550 555 560
cac gat gaa act ata ccg caa aca gca gga aaa tta gca aat ttt tgc 1728
His Asp Glu Thr Ile Pro Gln Thr Ala Gly Lys Leu Ala Asn Phe Cys
565 570 575
tcg atg tgt ggt cct aag ttt tgt tct atg aag cta tca aaa aaa ata 1776
Ser Met Cys Gly Pro Lys Phe Cys Ser Met Lys Leu Ser Lys Lys Ile
580 585 590
cgt aat tac act aat atg aaa aat ata aaa act att agt aat agt ttc 1824
Arg Asn Tyr Thr Asn Met Lys Asn Ile Lys Thr Ile Ser Asn Ser Phe
595 600 605
atg aat aaa tta gat aat agc ggt att aaa aat gct gac cga taa 1869
Met Asn Lys Leu Asp Asn Ser Gly Ile Lys Asn Ala Asp Arg
610 615 620
<210> 207
<211> 622
<212> PRT
<213> Candidatus Baumannia cicadellinicola
<400> 207
Met Ser Arg Ser Ser Ile Pro Ala Ser Arg Arg Val Ser Arg Ala Lys
1 5 10 15
Ala Gln Ala Phe Met Asp Ser Leu Thr Gly Ser Ser Tyr Phe Pro Asn

20 25 30
 Ser Arg Arg Ile Tyr Leu Gln Gly Lys Thr Pro Ser Val His Val Pro
 35 40 45
 Met Arg Glu Ile Lys Leu His Pro Thr Leu Ile Gly Lys Asn Gly Glu
 50 55 60
 His Tyr Glu Asp Asn Gln Pro Ile Pro Val Tyr Asp Thr Ser Gly Pro

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

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

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

 260 265 270

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

515 520 525
 Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg Phe Glu Phe Arg Trp
 530 535 540
 Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe Thr Ala Arg Met Tyr
 545 550 555 560
 His Asp Glu Thr Ile Pro Gln Thr Ala Gly Lys Leu Ala Asn Phe Cys
 565 570 575
 Ser Met Cys Gly Pro Lys Phe Cys Ser Met Lys Leu Ser Lys Lys Ile

580 585 590
 Arg Asn Tyr Thr Asn Met Lys Asn Ile Lys Thr Ile Ser Asn Ser Phe
 595 600 605
 Met Asn Lys Leu Asp Asn Ser Gly Ile Lys Asn Ala Asp Arg
 610 615 620

<210> 208
 <211> 636
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(636)
 <223> ThiE gene from E. coli encoding thiamine phosphate synthase
 <400> 208

atg tat cag cct gat ttt cct cct gta cct ttt cgt tca gga ctg tac 48

Met Tyr Gln Pro Asp Phe Pro Pro Val Pro Phe Arg Ser Gly Leu Tyr
 1 5 10 15
 ccg gtg gtg gac agc gta cag tgg atc gaa cgt ctg ttg gat gca ggc 96
 Pro Val Val Asp Ser Val Gln Trp Ile Glu Arg Leu Leu Asp Ala Gly
 20 25 30
 gta cgt act ctc cag cta cgc atc aaa gat cgg cgc gat gaa gag gtg 144
 Val Arg Thr Leu Gln Leu Arg Ile Lys Asp Arg Arg Asp Glu Glu Val

35 40 45
 gaa gcc gat gtc gtg gcg gca att gcg ctg ggc cgc cgc tat aac gcg 192
 Glu Ala Asp Val Val Ala Ala Ile Ala Leu Gly Arg Arg Tyr Asn Ala

50	55	60	
cga ttg ttt atc aac gat tac tgg cgg ctg gcg atc aag cat cag gcg			240
Arg Leu Phe Ile Asn Asp Tyr Trp Arg Leu Ala Ile Lys His Gln Ala			
65	70	75	80
tat ggc gtc cat ttg ggg cag gaa gat ttg caa gcc acc gat ctc aat			288
Tyr Gly Val His Leu Gly Gln Glu Asp Leu Gln Ala Thr Asp Leu Asn			
85	90	95	
gcc atc cgc gcg gca ggc ctg cgg ctg ggc gtt tcg aca cat gac gat			336
Ala Ile Arg Ala Ala Gly Leu Arg Leu Gly Val Ser Thr His Asp Asp			
100	105	110	
atg gaa atc gac gtc gcg ctg gca gca cgc ccc tct tat atc gcg ctg			384
Met Glu Ile Asp Val Ala Leu Ala Ala Arg Pro Ser Tyr Ile Ala Leu			
115	120	125	
gga cat gtg ttc ccg acg caa acc aaa cag atg cct tct gca ccg cag			432
Gly His Val Phe Pro Thr Gln Thr Lys Gln Met Pro Ser Ala Pro Gln			
130	135	140	
ggg ctg gaa cag ctg gca cgg cat gtt gag cga ctg gcg gat tat ccc			480
Gly Leu Glu Gln Leu Ala Arg His Val Glu Arg Leu Ala Asp Tyr Pro			
145	150	155	160
acc gtg gcg att ggc ggt atc agt ctg gca cgc gcg cct gcg gtg ata			528
Thr Val Ala Ile Gly Gly Ile Ser Leu Ala Arg Ala Pro Ala Val Ile			
165	170	175	
gca acg ggt gtc ggc agt atc gcc gtc gtc agc gcc att act caa gcc			576
Ala Thr Gly Val Gly Ser Ile Ala Val Val Ser Ala Ile Thr Gln Ala			
180	185	190	
gca gac tgg cgt ttg gca acg gca cag ttg ctg gaa att gca gga gtt			624
Ala Asp Trp Arg Leu Ala Thr Ala Gln Leu Leu Glu Ile Ala Gly Val			
195	200	205	
ggc gat gaa tga			636
Gly Asp Glu			
210			

<210> 209

<211> 211

<212> PRT

<213> Escherichia coli

<400> 209

Met Tyr Gln Pro Asp Phe Pro Pro Val Pro Phe Arg Ser Gly Leu Tyr

1 5 10 15

Pro Val Val Asp Ser Val Gln Trp Ile Glu Arg Leu Leu Asp Ala Gly

20 25 30

Val Arg Thr Leu Gln Leu Arg Ile Lys Asp Arg Arg Asp Glu Glu Val

35 40 45

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

50 55 60

Arg Leu Phe Ile Asn Asp Tyr Trp Arg Leu Ala Ile Lys His Gln Ala

65 70 75 80

Tyr Gly Val His Leu Gly Gln Glu Asp Leu Gln Ala Thr Asp Leu Asn

85 90 95

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

100 105 110

Met Glu Ile Asp Val Ala Leu Ala Ala Arg Pro Ser Tyr Ile Ala Leu

115 120 125

Gly His Val Phe Pro Thr Gln Thr Lys Gln Met Pro Ser Ala Pro Gln

130 135 140

Gly Leu Glu Gln Leu Ala Arg His Val Glu Arg Leu Ala Asp Tyr Pro

145 150 155 160

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

165 170 175

Ala Thr Gly Val Gly Ser Ile Ala Val Val Ser Ala Ile Thr Gln Ala

180 185 190

Ala Asp Trp Arg Leu Ala Thr Ala Gln Leu Leu Glu Ile Ala Gly Val

195 200 205

Gly Asp Glu

210

<210> 210

<211> 756

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(756)

<223> ThiF gene from E.coli encoding a ThiS adenylyltransferase

<400> 210

atg aat gac cgt gac ttt atg cgt tat agc cgc caa atc ctg ctc gac 48

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

1 5 10 15

gat atc gct ctg gac ggg cag caa aaa ctg ctc gac agc cag gtg ctg 96

Asp Ile Ala Leu Asp Gly Gln Gln Lys Leu Leu Asp Ser Gln Val Leu

20 25 30

att atc ggt ctg ggc ggg ctg ggt aca cct gct gcg ctg tac ctg gcg 144

Ile Ile Gly Leu Gly Gly Leu Gly Thr Pro Ala Ala Leu Tyr Leu Ala

35 40 45

ggc gct ggc gtc ggg acg ctg gta ctg gca gat gac gac gat gtg cat 192

Gly Ala Gly Val Gly Thr Leu Val Leu Ala Asp Asp Asp Asp Val His

50 55 60

tta agc aat ctg caa cga caa atc ctc ttt acc act gaa gat atc gat 240

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

65 70 75 80

cgc ccg aaa tcg cag gtc agc caa cag cga ctg aca cag ttg aat ccc 288

Arg Pro Lys Ser Gln Val Ser Gln Gln Arg Leu Thr Gln Leu Asn Pro

85 90 95

gac att caa ctg aca gca tta caa caa cgg tta acg ggt gag gcg tta 336

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

100 105 110

aaa gat gcg gtt gca cgg gcc gat gtg gtg ctc gac tgt acc gac aat 384

Lys Asp Ala Val Ala Arg Ala Asp Val Val Leu Asp Cys Thr Asp Asn
115 120 125

atg gcg act cgc cag gag att aat gcc gcc tgc gtg gca ctc aac acg 432
Met Ala Thr Arg Gln Glu Ile Asn Ala Ala Cys Val Ala Leu Asn Thr
130 135 140

ccg ctt atc acc gcc agc gcg gtc gga ttt ggc ggt cag ttg atg gta 480
Pro Leu Ile Thr Ala Ser Ala Val Gly Phe Gly Gly Gln Leu Met Val
145 150 155 160

ctg acg ccg ccc tgg gag cag ggg tgt tac cgc tgc ctg tgg cca gat 528

Leu Thr Pro Pro Trp Glu Gln Gly Cys Tyr Arg Cys Leu Trp Pro Asp
165 170 175

aac cag gag cca gaa cgc aac tgc cgc acg gcg ggc gtg gtt ggc ccg 576
Asn Gln Glu Pro Glu Arg Asn Cys Arg Thr Ala Gly Val Val Gly Pro
180 185 190

gtg gtc ggg gtt atg ggc act ttg cag gca ctg gaa gcc att aag tta 624
Val Val Gly Val Met Gly Thr Leu Gln Ala Leu Glu Ala Ile Lys Leu

195 200 205

tta agc ggt ata gag aca cct gcg gga gaa ctc cga ctg ttc gac ggt 672
Leu Ser Gly Ile Glu Thr Pro Ala Gly Glu Leu Arg Leu Phe Asp Gly
210 215 220

aaa tcg agc cag tgg cgc agc ctg gcg ttg cgc cgc gcc agt ggt tgc 720
Lys Ser Ser Gln Trp Arg Ser Leu Ala Leu Arg Arg Ala Ser Gly Cys
225 230 235 240

ccg gta tgc gga gga agc aat gca gat cct gtt taa 756
Pro Val Cys Gly Gly Ser Asn Ala Asp Pro Val
245 250

<210> 211
<211> 251
<212> PRT
<213> Escherichia coli
<400> 211

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

245

250

<210> 212

<211> 201

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(201)

<223> ThiS gene from E. coli encoding Sulfur-carrier protein

<400> 212

atg cag atc ctg ttt aac gat caa gcg atg cag tgc gcc gcc ggg caa 48
Met Gln Ile Leu Phe Asn Asp Gln Ala Met Gln Cys Ala Ala Gly Gln

1 5 10 15
act gtt cac gaa cta ctg gag caa ctc gac caa cga caa gcg ggc gcg 96
Thr Val His Glu Leu Leu Glu Gln Leu Asp Gln Arg Gln Ala Gly Ala

20 25 30
gct ctg gcg att aat cag caa atc gtc ccg cgt gag cag tgg gcg caa 144
Ala Leu Ala Ile Asn Gln Gln Ile Val Pro Arg Glu Gln Trp Ala Gln

35 40 45
cat atc gtg cag gat ggc gac cag atc ctg ctt ttt cag gtt att gca 192
His Ile Val Gln Asp Gly Asp Gln Ile Leu Leu Phe Gln Val Ile Ala

50 55 60
ggg ggt tga 201
Gly Gly

65

<210> 213

<211> 66

<212> PRT

<213> Escherichia coli

<400> 213

Met Gln Ile Leu Phe Asn Asp Gln Ala Met Gln Cys Ala Ala Gly Gln

1 5 10 15
Thr Val His Glu Leu Leu Glu Gln Leu Asp Gln Arg Gln Ala Gly Ala

20 25 30
Ala Leu Ala Ile Asn Gln Gln Ile Val Pro Arg Glu Gln Trp Ala Gln
35 40 45
His Ile Val Gln Asp Gly Asp Gln Ile Leu Leu Phe Gln Val Ile Ala
50 55 60

Gly Gly

65

<210> 214

<211> 771

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(771)

<223> ThiG gene from E. coli encoding Thiazole synthase

<400> 214

atg tta cgt att gcg gac aaa acg ttt gat tca cat ctg ttt acc ggc 48

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

1 5 10 15

aca ggg aaa ttc gct tct tca caa ctg atg gtg gag gcg atc cgc gct 96

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

20 25 30
tcc ggc agc cag ctg gtg aca ctg gcg atg aaa cgt gtc gac ttg cgc 144

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

35 40 45

cag cac aac gac gct atc ctc gaa ccg ctt atc gcg gcg ggt gtg acc 192

Gln His Asn Asp Ala Ile Leu Glu Pro Leu Ile Ala Ala Gly Val Thr

50 55 60

ctg ctg cca aat aca tcc ggg gcg aaa aca gcg gaa gaa gcc att ttc 240

Leu Leu Pro Asn Thr Ser Gly Ala Lys Thr Ala Glu Glu Ala Ile Phe

65 70 75 80

gcc gcc cat ctg gct cgt gaa gcg tta ggc aca aac tgg tta aaa tta 288

Ala Ala His Leu Ala Arg Glu Ala Leu Gly Thr Asn Trp Leu Lys Leu

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

Phe Ala His Ala Thr Ser Pro Leu Thr Gly Phe Leu Glu Ala Ser Ala

245 250 255

tga 771

<210> 215

<211> 256

<212> PRT

<213> Escherichia coli

<400> 215

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

1 5 10 15

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

20 25 30

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

35 40 45

Gln His Asn Asp Ala Ile Leu Glu Pro Leu Ile Ala Ala Gly Val Thr

50 55 60

Leu Leu Pro Asn Thr Ser Gly Ala Lys Thr Ala Glu Glu Ala Ile Phe

65 70 75 80

Ala Ala His Leu Ala Arg Glu Ala Leu Gly Thr Asn Trp Leu Lys Leu

85 90 95

Glu Ile His Pro Asp Ala Arg Trp Leu Leu Pro Asp Pro Ile Glu Thr

100 105 110

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

115 120 125

Tyr Cys Gly Ala Asp Pro Val Leu Cys Lys Arg Leu Glu Glu Val Gly

130 135 140

Cys Ala Ala Val Met Pro Leu Gly Ala Pro Ile Gly Ser Asn Gln Gly

145 150 155 160

Leu Glu Thr Arg Ala Met Leu Glu Ile Ile Ile Gln Gln Ala Thr Val

165 170 175

Pro Val Val Val Asp Ala Gly Ile Gly Val Pro Ser His Ala Ala Gln

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

225 230 235 240
Phe Ala His Ala Thr Ser Pro Leu Thr Gly Phe Leu Glu Ala Ser Ala
245 250 255
<210> 216
<211> 1134
<212> DNA
<213> Escherichia coli
<220><221> CDS
<222> (1)..(1134)
<223> ThiH gene from E. coli encoding 2-iminoacetate synthase
<400> 216
atg aaa acc ttc agc gat cgc tgg cga caa ctg gac tgg gac gac atc 48
Met Lys Thr Phe Ser Asp Arg Trp Arg Gln Leu Asp Trp Asp Asp Ile

1 5 10 15
cgc ctg cgt atc aac ggc aaa acg gct gct gac gta gag cgg gcg cta 96
Arg Leu Arg Ile Asn Gly Lys Thr Ala Ala Asp Val Glu Arg Ala Leu
20 25 30
aat gcc tcg caa ctc acc cgc gac gac atg atg gcg ctg tta tcg cct 144
Asn Ala Ser Gln Leu Thr Arg Asp Asp Met Met Ala Leu Leu Ser Pro
35 40 45

gcc gcc agt ggc tat ctg gaa caa ctg gcc caa cgg gcg cag cgt ctg 192
Ala Ala Ser Gly Tyr Leu Glu Gln Leu Ala Gln Arg Ala Gln Arg Leu
50 55 60
acc cgt cag cga ttt ggc aac aca gtt agt ttc tac gtc ccg ctt tat 240
Thr Arg Gln Arg Phe Gly Asn Thr Val Ser Phe Tyr Val Pro Leu Tyr
65 70 75 80

ctt tcc aat ctt tgc gct aac gac tgc acg tac tgt gga ttt tcc atg 288

Leu Ser Asn Leu Cys Ala Asn Asp Cys Thr Tyr Cys Gly Phe Ser Met

85 90 95

agt aat cgc atc aag cgc aaa acg ctg gat gaa gcg gat att gcc agg 336

Ser Asn Arg Ile Lys Arg Lys Thr Leu Asp Glu Ala Asp Ile Ala Arg

100 105 110

gaa agt gcc gct ata cgg gag atg ggc ttt gaa cat ctg ctg tta gtc 384

Glu Ser Ala Ala Ile Arg Glu Met Gly Phe Glu His Leu Leu Leu Val

115 120 125

act ggt gaa cat cag gcg aaa gtg ggg atg gat tac ttt cgt cgt cat 432

Thr Gly Glu His Gln Ala Lys Val Gly Met Asp Tyr Phe Arg Arg His

130 135 140

ctc cct gcc ctt cgt gaa cag ttc tct tca cta cag atg gaa gtg caa 480

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

145 150 155 160

ccg ctg gcg gag acg gaa tac gcc gag tta aag caa ctt ggt ctg gat 528

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

165 170 175

ggc gtg atg gtt tat cag gag aca tat cac gag gcg act tat gcc cgc 576

Gly Val Met Val Tyr Gln Glu Thr Tyr His Glu Ala Thr Tyr Ala Arg

180 185 190

cat cat ctg aaa ggc aaa aaa cag gac ttc ttc tgg cgg ctg gaa acg 624

His His Leu Lys Gly Lys Lys Gln Asp Phe Phe Trp Arg Leu Glu Thr

195 200 205

ccg gat cgg ctg ggg cgt gcg ggg att gat aag ata ggc ctc ggc gcg 672

Pro Asp Arg Leu Gly Arg Ala Gly Ile Asp Lys Ile Gly Leu Gly Ala

210 215 220

cta att ggc ctt tcc gac aac tgg cgc gtt gac agc tat atg gtt gcc 720

Leu Ile Gly Leu Ser Asp Asn Trp Arg Val Asp Ser Tyr Met Val Ala

225	230	235	240	
gaa cat ttg cta tgg ctg caa cag cat tac tgg caa agc cgt tac tct				768
Glu His Leu Leu Trp Leu Gln Gln His Tyr Trp Gln Ser Arg Tyr Ser				
	245	250	255	
gtc tcc ttt ccg cgc ctg cgc ccg tgt act ggc ggc att gag cct gcg				816
Val Ser Phe Pro Arg Leu Arg Pro Cys Thr Gly Gly Ile Glu Pro Ala				
	260	265	270	
tcg att atg gat gaa cgc cag tta gtg caa acc atc tgc gcc ttc cga				864
Ser Ile Met Asp Glu Arg Gln Leu Val Gln Thr Ile Cys Ala Phe Arg				
	275	280	285	
ctg ctt gca ccg gag att gaa ctg tca ctc tcc acg cgg gaa tca ccg				912
Leu Leu Ala Pro Glu Ile Glu Leu Ser Leu Ser Thr Arg Glu Ser Pro				
	290	295	300	
tgg ttt cgc gat cgc gtt att ccg ctg gcg atc aat aac gtc agc gcc				960
Trp Phe Arg Asp Arg Val Ile Pro Leu Ala Ile Asn Asn Val Ser Ala				
305	310	315	320	
ttc tcg aaa acg cag cca ggt ggc tat gcc gat aat cac ccc gag ttg				1008
Phe Ser Lys Thr Gln Pro Gly Gly Tyr Ala Asp Asn His Pro Glu Leu				
	325	330	335	
gaa cag ttc tca ccg cac gac gat cgc aga ccg gaa gcg gtt gct gcc				1056
Glu Gln Phe Ser Pro His Asp Asp Arg Arg Pro Glu Ala Val Ala Ala				
	340	345	350	
gcg tta acc gct cag ggt ttg cag ccg gta tgg aaa gac tgg gac agc				1104
Ala Leu Thr Ala Gln Gly Leu Gln Pro Val Trp Lys Asp Trp Asp Ser				
	355	360	365	
tat ctg gga cgc gcc tcg caa aga cta tga				1134
Tyr Leu Gly Arg Ala Ser Gln Arg Leu				
	370	375		

<210> 217

<211> 377

<212> PRT

<213> Escherichia coli

<400> 217

Met Lys Thr Phe Ser Asp Arg Trp Arg Gln Leu Asp Trp Asp Asp Ile

1 5 10 15

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

20 25 30

Asn Ala Ser Gln Leu Thr Arg Asp Asp Met Met Ala Leu Leu Ser Pro

35 40 45

Ala Ala Ser Gly Tyr Leu Glu Gln Leu Ala Gln Arg Ala Gln Arg Leu

50 55 60

Thr Arg Gln Arg Phe Gly Asn Thr Val Ser Phe Tyr Val Pro Leu Tyr

65 70 75 80

Leu Ser Asn Leu Cys Ala Asn Asp Cys Thr Tyr Cys Gly Phe Ser Met

85 90 95

Ser Asn Arg Ile Lys Arg Lys Thr Leu Asp Glu Ala Asp Ile Ala Arg

100 105 110

Glu Ser Ala Ala Ile Arg Glu Met Gly Phe Glu His Leu Leu Leu Val

115 120 125

Thr Gly Glu His Gln Ala Lys Val Gly Met Asp Tyr Phe Arg Arg His

130 135 140

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

145 150 155 160

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

165 170 175

Gly Val Met Val Tyr Gln Glu Thr Tyr His Glu Ala Thr Tyr Ala Arg

180 185 190

His His Leu Lys Gly Lys Lys Gln Asp Phe Phe Trp Arg Leu Glu Thr

195 200 205

Pro Asp Arg Leu Gly Arg Ala Gly Ile Asp Lys Ile Gly Leu Gly Ala

210 215 220

Leu Ile Gly Leu Ser Asp Asn Trp Arg Val Asp Ser Tyr Met Val Ala

225 230 235 240

Glu His Leu Leu Trp Leu Gln Gln His Tyr Trp Gln Ser Arg Tyr Ser

245 250 255

Val Ser Phe Pro Arg Leu Arg Pro Cys Thr Gly Gly Ile Glu Pro Ala

260 265 270

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

275 280 285

Leu Leu Ala Pro Glu Ile Glu Leu Ser Leu Ser Thr Arg Glu Ser Pro

290 295 300

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

305 310 315 320

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

325 330 335

Glu Gln Phe Ser Pro His Asp Asp Arg Arg Pro Glu Ala Val Ala Ala

340 345 350

Ala Leu Thr Ala Gln Gly Leu Gln Pro Val Trp Lys Asp Trp Asp Ser

355 360 365

Tyr Leu Gly Arg Ala Ser Gln Arg Leu

370 375

<210> 218

<211> 1110

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(1110)

<223> Thi0 gene from E. coli encoding Glycine oxidase

<400> 218

atg aaa agg cat tat gaa gca gtg gtg att gga ggc gga att atc ggt 48

Met Lys Arg His Tyr Glu Ala Val Val Ile Gly Gly Gly Ile Ile Gly

1 5 10 15

tcc gca att gct tat tat ttg gca aag gaa aac aaa aac acc gca ttg 96

Ser Ala Ile Ala Tyr Tyr Leu Ala Lys Glu Asn Lys Asn Thr Ala Leu

20	25	30	
ttt gaa agc gga aca atg ggc ggc aga acg aca agt gcc gct gcc gga			144
Phe Glu Ser Gly Thr Met Gly Gly Arg Thr Thr Ser Ala Ala Ala Gly			
35	40	45	
atg ctg ggc gcc cat gcc gaa tgc gag gaa cgt gac gcg ttt ttt gat			192
Met Leu Gly Ala His Ala Glu Cys Glu Glu Arg Asp Ala Phe Phe Asp			
50	55	60	
ttc gct atg cac agt cag cgt ctg tac aaa ggt ctt gga gaa gag ctt			240
Phe Ala Met His Ser Gln Arg Leu Tyr Lys Gly Leu Gly Glu Glu Leu			
65	70	75	80
tat gca tta tcc ggt gtg gat atc agg cag cat aac ggc ggt atg ttt			288
Tyr Ala Leu Ser Gly Val Asp Ile Arg Gln His Asn Gly Gly Met Phe			
85	90	95	
aag ctt gca ttt tct gaa gaa gat gtg ctg cag ctg aga cag atg gac			336
Lys Leu Ala Phe Ser Glu Glu Asp Val Leu Gln Leu Arg Gln Met Asp			
100	105	110	
gat ttg gac tct gtc agc tgg tat tca aaa gaa gag gtg tta gaa aaa			384
Asp Leu Asp Ser Val Ser Trp Tyr Ser Lys Glu Glu Val Leu Glu Lys			
115	120	125	
gag ccg tat gcg tct ggt gac atc ttt ggt gca tct ttt att cag gat			432
Glu Pro Tyr Ala Ser Gly Asp Ile Phe Gly Ala Ser Phe Ile Gln Asp			
130	135	140	
gat gtg cat gtg gag cct tat ttt gtt tgc aag gca tat gtg aaa gca			480
Asp Val His Val Glu Pro Tyr Phe Val Cys Lys Ala Tyr Val Lys Ala			
145	150	155	160
gca aaa atg ctt ggg gcg gag att ttt gag cat acg ccc gtc ctg cat			528
Ala Lys Met Leu Gly Ala Glu Ile Phe Glu His Thr Pro Val Leu His			
165	170	175	
gtc gaa cgt gac ggt gaa gcc ctg ttc atc aag acc cct agc gga gac			576
Val Glu Arg Asp Gly Glu Ala Leu Phe Ile Lys Thr Pro Ser Gly Asp			
180	185	190	

gta tgg gct aat cat gtt gtc gtt gcc agc ggg gtg tgg agc gga atg	624
Val Trp Ala Asn His Val Val Val Ala Ser Gly Val Trp Ser Gly Met	
195 200 205	
ttt ttt aaa cag ctt gga ctg aac aat gct ttt ctc cct gta aaa ggg	672
Phe Phe Lys Gln Leu Gly Leu Asn Asn Ala Phe Leu Pro Val Lys Gly	
210 215 220	
gag tgc ctg tcc gtt tgg aat gat gat atc ccg ctg aca aaa acg ctt	720
Glu Cys Leu Ser Val Trp Asn Asp Asp Ile Pro Leu Thr Lys Thr Leu	
225 230 235 240	
tac cat gat cac tgc tat atc gta ccg aga aaa agc gga aga ctg gtt	768
Tyr His Asp His Cys Tyr Ile Val Pro Arg Lys Ser Gly Arg Leu Val	
245 250 255	
gtc ggc gcg aca atg aag ccg ggg gac tgg agt gaa aca ccg gat ctt	816
Val Gly Ala Thr Met Lys Pro Gly Asp Trp Ser Glu Thr Pro Asp Leu	
260 265 270	
ggc gga ttg gaa tct gtt atg aaa aaa gca aaa acg atg ctg ccg gct	864
Gly Gly Leu Glu Ser Val Met Lys Lys Ala Lys Thr Met Leu Pro Ala	
275 280 285	
ata cag aat atg aag gtg gat cgt ttt tgg gcg gga ctc cgt ccg gga	912
Ile Gln Asn Met Lys Val Asp Arg Phe Trp Ala Gly Leu Arg Pro Gly	
290 295 300	
aca aag gat gga aaa ccg tac atc ggc aga cat cct gag gac agc cgt	960
Thr Lys Asp Gly Lys Pro Tyr Ile Gly Arg His Pro Glu Asp Ser Arg	
305 310 315 320	
att tta ttt gcg gct ggc cat ttc aga aac ggg atc ctg ctt gct ccc	1008
Ile Leu Phe Ala Ala Gly His Phe Arg Asn Gly Ile Leu Leu Ala Pro	
325 330 335	
gca acg ggc gct ttg atc agt gat ctc atc atg aat aaa gag gtc aac	1056
Ala Thr Gly Ala Leu Ile Ser Asp Leu Ile Met Asn Lys Glu Val Asn	
340 345 350	
caa gac tgg ctg cac gca ttc cga att gat cgc aag gag gcg gtt cag	1104

Gln Asp Trp Leu His Ala Phe Arg Ile Asp Arg Lys Glu Ala Val Gln

355 360 365
ata tga 1110
Ile

<210> 219

<211> 369

<212> PRT

<213> Escherichia coli

<400> 219

Met Lys Arg His Tyr Glu Ala Val Val Ile Gly Gly Gly Ile Ile Gly
1 5 10 15

Ser Ala Ile Ala Tyr Tyr Leu Ala Lys Glu Asn Lys Asn Thr Ala Leu
20 25 30

Phe Glu Ser Gly Thr Met Gly Gly Arg Thr Thr Ser Ala Ala Ala Gly
35 40 45

Met Leu Gly Ala His Ala Glu Cys Glu Glu Arg Asp Ala Phe Phe Asp
50 55 60

Phe Ala Met His Ser Gln Arg Leu Tyr Lys Gly Leu Gly Glu Glu Leu
65 70 75 80

Tyr Ala Leu Ser Gly Val Asp Ile Arg Gln His Asn Gly Gly Met Phe
85 90 95

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

Asp Leu Asp Ser Val Ser Trp Tyr Ser Lys Glu Glu Val Leu Glu Lys
115 120 125

Glu Pro Tyr Ala Ser Gly Asp Ile Phe Gly Ala Ser Phe Ile Gln Asp
130 135 140

Asp Val His Val Glu Pro Tyr Phe Val Cys Lys Ala Tyr Val Lys Ala
145 150 155 160

Ala Lys Met Leu Gly Ala Glu Ile Phe Glu His Thr Pro Val Leu His

165 170 175
 Val Glu Arg Asp Gly Glu Ala Leu Phe Ile Lys Thr Pro Ser Gly Asp
 180 185 190
 Val Trp Ala Asn His Val Val Val Ala Ser Gly Val Trp Ser Gly Met
 195 200 205

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

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

 Ala Thr Gly Ala Leu Ile Ser Asp Leu Ile Met Asn Lys Glu Val Asn
 340 345 350
 Gln Asp Trp Leu His Ala Phe Arg Ile Asp Arg Lys Glu Ala Val Gln
 355 360 365
 Ile

<210> 220

<211> 1098

<212> DNA

<213> *Pseudomonas putida*

<220><221> CDS

<222> (1)..(1098)

<223> Thi0 gene from Pseudomonas putida encoding a Glycine oxidase

<400> 220

atg agc aag caa gta gtg gtg gtc ggt ggc ggg gtc att ggc ctg ctg 48

Met Ser Lys Gln Val Val Val Val Gly Gly Gly Val Ile Gly Leu Leu

1 5 10 15

acg gca ttc aac ctg gcg gcg agc gtc gac cag gtg gtg gta tgc gac 96

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

20 25 30

cag ggc gaa gta ggg cgc gag tcc tcc tgg gct ggg ggc ggt atc gtc 144

Gln Gly Glu Val Gly Arg Glu Ser Ser Trp Ala Gly Gly Gly Ile Val

35 40 45

tcg ccc ctg tat cct tgg cgc tac agc ccg gca gtg acc gcc ctg gcg 192

Ser Pro Leu Tyr Pro Trp Arg Tyr Ser Pro Ala Val Thr Ala Leu Ala

50 55 60

cat tgg tcg cag gac ttt tac cca cag ttg ggc gag cgc ttg ttc gcc 240

His Trp Ser Gln Asp Phe Tyr Pro Gln Leu Gly Glu Arg Leu Phe Ala

65 70 75 80

agc acg ggc ctg gat cct gag gtg cat acc acc ggg ctt tac tgg ctc 288

Ser Thr Gly Leu Asp Pro Glu Val His Thr Thr Gly Leu Tyr Trp Leu

85 90 95

gac ctg gat gac caa gcc cag gcc ttg gcg tgg gca ggc cgt cag cag 336

Asp Leu Asp Asp Gln Ala Gln Ala Leu Ala Trp Ala Gly Arg Gln Gln

100 105 110

cgt ccg ctg agc gcc gtg gat att tca gcg gtg tac gac gca gtc cct 384

Arg Pro Leu Ser Ala Val Asp Ile Ser Ala Val Tyr Asp Ala Val Pro

115 120 125

gtg ctg ggg cca ggc ttt gag cga gcc ctc tac atg gaa ggc gtg gcc 432

Val Leu Gly Pro Gly Phe Glu Arg Ala Leu Tyr Met Glu Gly Val Ala

130 135 140

aat gtg cgc aac ccg cgc ctg gtc aaa tcg ctg aag gcg gcg ttg ctg 480

Asn Val Arg Asn Pro Arg Leu Val Lys Ser Leu Lys Ala Ala Leu Leu

145 150 155 160
gca ttg ccc aat gtg agc gtg cgc gag cac tgc cag atc acg ggg ttc 528

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

 165 170 175
gtg cag cag ggc gct cgt atc att ggg gtg agc acc gct gaa ggc gag 576

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

 180 185 190
ctg gcc gcc gac gaa gtc gta ctg agc gcc ggt gcc tgg agc ggc gaa 624

Leu Ala Ala Asp Glu Val Val Leu Ser Ala Gly Ala Trp Ser Gly Glu

 195 200 205
ctg ctg cgc cac ttg ggc ctt gag ctt cca gtc gag ccg gta aaa ggg 672

Leu Leu Arg His Leu Gly Leu Glu Leu Pro Val Glu Pro Val Lys Gly

 210 215 220
cag atg atc ctg ttc aaa tgc gct gaa gat ttt ctg cca agc atg gtg 720

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

225 230 235 240
ctt gcc aaa ggt cgt tat gca att ccg cgt cgg gat ggt cac att ctg 768

Leu Ala Lys Gly Arg Tyr Ala Ile Pro Arg Arg Asp Gly His Ile Leu

 245 250 255
gtg ggc agc acg ctg gag cat gcc ggc tac gac aag aca ccc acc gat 816

Val Gly Ser Thr Leu Glu His Ala Gly Tyr Asp Lys Thr Pro Thr Asp

 260 265 270
gag gcg ttg gcc agc ctc aag gca tcg gcg gtg gat ctg ctg ccc ggc 864

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

 275 280 285
ctg gaa ggc gcg cac gtg gtt gcc cac tgg gcc ggg ctg cgg cca ggt 912

Leu Glu Gly Ala His Val Val Ala His Trp Ala Gly Leu Arg Pro Gly

 290 295 300
gag ggc ttg gcc agc ctc aag gca tcg gcg gtg gat ctg ctg ccc ggc 864

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

 275 280 285
ctg gaa ggc gcg cac gtg gtt gcc cac tgg gcc ggg ctg cgg cca ggt 912

Leu Glu Gly Ala His Val Val Ala His Trp Ala Gly Leu Arg Pro Gly

 290 295 300
gag ggc ttg gcc agc ctc aag gca tcg gcg gtg gat ctg ctg ccc ggc 864

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

tcg cca gaa ggc gtt ccg ttt atc ggg ccg gta ccc ggc ttc gat ggg 960
 Ser Pro Glu Gly Val Pro Phe Ile Gly Pro Val Pro Gly Phe Asp Gly
 305 310 315 320
 tta tgg ctg aac tgc ggc cat tac cga aac ggg ctg gtg ctg gcg ccc 1008
 Leu Trp Leu Asn Cys Gly His Tyr Arg Asn Gly Leu Val Leu Ala Pro
 325 330 335
 gct tcg tgc caa ctg ctg gcc gat ttg ctc aat ggc gcc gag ccc atc 1056

Ala Ser Cys Gln Leu Leu Ala Asp Leu Leu Asn Gly Ala Glu Pro Ile
 340 345 350
 atc gac ccg tca ccc tac gcc ccg tct ggg cgc ctt ggc taa 1098
 Ile Asp Pro Ser Pro Tyr Ala Pro Ser Gly Arg Leu Gly
 355 360 365

<210> 221

<211> 365

<212> PRT

<213> Pseudomonas putida

<400> 221

Met Ser Lys Gln Val Val Val Val Gly Gly Gly Val Ile Gly Leu Leu

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

65 70 75 80
 Ser Thr Gly Leu Asp Pro Glu Val His Thr Thr Gly Leu Tyr Trp Leu
 85 90 95
 Asp Leu Asp Asp Gln Ala Gln Ala Leu Ala Trp Ala Gly Arg Gln Gln
 100 105 110
 Arg Pro Leu Ser Ala Val Asp Ile Ser Ala Val Tyr Asp Ala Val Pro

115 120 125
Val Leu Gly Pro Gly Phe Glu Arg Ala Leu Tyr Met Glu Gly Val Ala

130 135 140
Asn Val Arg Asn Pro Arg Leu Val Lys Ser Leu Lys Ala Ala Leu Leu
145 150 155 160
Ala Leu Pro Asn Val Ser Val Arg Glu His Cys Gln Ile Thr Gly Phe

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

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

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

260 265 270
Glu Ala Leu Ala Ser Leu Lys Ala Ser Ala Val Asp Leu Leu Pro Gly
275 280 285
Leu Glu Gly Ala His Val Val Ala His Trp Ala Gly Leu Arg Pro Gly

290 295 300
Ser Pro Glu Gly Val Pro Phe Ile Gly Pro Val Pro Gly Phe Asp Gly
305 310 315 320
Leu Trp Leu Asn Cys Gly His Tyr Arg Asn Gly Leu Val Leu Ala Pro

325 330 335
Ala Ser Cys Gln Leu Leu Ala Asp Leu Leu Asn Gly Ala Glu Pro Ile
340 345 350
Ile Asp Pro Ser Pro Tyr Ala Pro Ser Gly Arg Leu Gly

355 360 365

<210> 222
 <211> 1140
 <212> DNA
 <213> *Synechococcus elongatus*
 <220><221> CDS
 <222> (1)..(1140)
 <223> Thi0 gene from *Synechococcus elongatus* encoding a Glycine oxidase
 <400> 222

atg gcg ttc gag gta gcc gtc ttt ggg ggc ggc gtc att ggc ttg gcg 48

Met Ala Phe Glu Val Ala Val Phe Gly Gly Gly Val Ile Gly Leu Ala
 1 5 10 15

atc gcg cta gaa ctg cga tcg cga ggc gcg atg gtg cag gtc tac agt 96
 Ile Ala Leu Glu Leu Arg Ser Arg Gly Ala Met Val Gln Val Tyr Ser
 20 25 30

caa aac act cag gcg gcg gca ggt cgt gtg gca gca ggg atg ttg gcg 144
 Gln Asn Thr Gln Ala Ala Ala Gly Arg Val Ala Ala Gly Met Leu Ala
 35 40 45

ccc cag tcg gaa ggc atc gaa gtc ggg ccc atg ctg gat ctg ggg ctg 192
 Pro Gln Ser Glu Gly Ile Glu Val Gly Pro Met Leu Asp Leu Gly Leu
 50 55 60

cgc agc cga tcg ctc tac gcc cgc tgg acc cag caa ctc gaa caa ctc 240
 Arg Ser Arg Ser Leu Tyr Ala Arg Trp Thr Gln Gln Leu Glu Gln Leu
 65 70 75 80

agc ggt caa gac agt ggc tac tgg ccc tgc ggc att ttg gtg ccc ctg 288
 Ser Gly Gln Asp Ser Gly Tyr Trp Pro Cys Gly Ile Leu Val Pro Leu
 85 90 95

agt gag gcc aaa aat cgc gat cgc tat cct cat cca gca gaa tct ccg 336
 Ser Glu Ala Lys Asn Arg Asp Arg Tyr Pro His Pro Ala Glu Ser Pro
 100 105 110

ggg caa tgg ctc tcg gca gcg gac tta cga gac ttt cag ccc gca cta 384

Gly Gln Trp Leu Ser Ala Ala Asp Leu Arg Asp Phe Gln Pro Ala Leu

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

gtt gaa act tgg tgg gga tac cgt ccc gcg aca cca gat gaa tgg ccg 912
Val Glu Thr Trp Trp Gly Tyr Arg Pro Ala Thr Pro Asp Glu Trp Pro
290 295 300

atc ctg ggg caa ggc ccc gct gag aac tta ttc ttg gcg acc ggc cac 960
Ile Leu Gly Gln Gly Pro Ala Glu Asn Leu Phe Leu Ala Thr Gly His
305 310 315 320

tac cgc aac ggt atg ctg ctc gcc cca att acc gct cag cta ctc gct 1008
Tyr Arg Asn Gly Met Leu Leu Ala Pro Ile Thr Ala Gln Leu Leu Ala
325 330 335

gac caa att ctc gac cac tgc acg gat caa ctg ctt cat gcc ttc cgt 1056

Asp Gln Ile Leu Asp His Cys Thr Asp Gln Leu Leu His Ala Phe Arg
340 345 350

tac gac cgc ttc tcc agc cat gac tcc agc acc cat caa ccc tta ccc 1104
Tyr Asp Arg Phe Ser Ser His Asp Ser Ser Thr His Gln Pro Leu Pro
355 360 365

gct ctt gca ggc ttg tca gcg tca acg ggt cag tga 1140
Ala Leu Ala Gly Leu Ser Ala Ser Thr Gly Gln

370 375

<210> 223
<211> 379
<212> PRT
<213> Synechococcus elongatus
<400> 223

Met Ala Phe Glu Val Ala Val Phe Gly Gly Gly Val Ile Gly Leu Ala
1 5 10 15

Ile Ala Leu Glu Leu Arg Ser Arg Gly Ala Met Val Gln Val Tyr Ser
20 25 30

Gln Asn Thr Gln Ala Ala Ala Gly Arg Val Ala Ala Gly Met Leu Ala
35 40 45

Pro Gln Ser Glu Gly Ile Glu Val Gly Pro Met Leu Asp Leu Gly Leu
50 55 60

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

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

 Ala Leu Ala Gly Leu Ser Ala Ser Thr Gly Gln
 370 375
 <210> 224
 <211> 801
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(801)
 <223> ThiD gene from E. coli encoding phosphohydroxymethylpyrimidine
 kinase
 <400> 224

 atg aaa cga att aac gct ctg acg att gcc ggt act gat ccg agt ggt 48
 Met Lys Arg Ile Asn Ala Leu Thr Ile Ala Gly Thr Asp Pro Ser Gly
 1 5 10 15

 ggt gcg ggg att cag gcc gat ctt aaa acc ttc tcg gca ctt ggc gct 96
 Gly Ala Gly Ile Gln Ala Asp Leu Lys Thr Phe Ser Ala Leu Gly Ala
 20 25 30
 tat ggt tgc tca gtt att act gca ctg gtg gcg caa aat acc cgt ggc 144
 Tyr Gly Cys Ser Val Ile Thr Ala Leu Val Ala Gln Asn Thr Arg Gly
 35 40 45
 gta cag tcg gtg tat cgc att gag cct gat ttt gtc gcc gcc cag etc 192

 Val Gln Ser Val Tyr Arg Ile Glu Pro Asp Phe Val Ala Ala Gln Leu
 50 55 60
 gat tcg gtg ttc agc gat gtg cga atc gat acc act aaa atc ggt atg 240
 Asp Ser Val Phe Ser Asp Val Arg Ile Asp Thr Thr Lys Ile Gly Met

65	70	75	80	
ctg gcg gaa acc gat att gtt gaa gcg gtg gca gaa cgg ttg caa cgt				288
Leu Ala Glu Thr Asp Ile Val Glu Ala Val Ala Glu Arg Leu Gln Arg				
	85	90	95	
tat cag atc caa aac gtg gta ctc gac acc gtt atg ctg gca aaa agc				336
Tyr Gln Ile Gln Asn Val Val Leu Asp Thr Val Met Leu Ala Lys Ser				
	100	105	110	
ggc gac ccg ctg ctt tca cct tgc gcg gtt gct acg ctg cgc agt cga				384
Gly Asp Pro Leu Leu Ser Pro Ser Ala Val Ala Thr Leu Arg Ser Arg				
	115	120	125	
tta ttg cca cag gtt tca tta ata acg cca aac ttg ccc gaa gct gcc				432
Leu Leu Pro Gln Val Ser Leu Ile Thr Pro Asn Leu Pro Glu Ala Ala				
	130	135	140	
gcc ttg ctc gac gcg cca cac gcg cgc acc gaa cag gaa atg ctg gaa				480
Ala Leu Leu Asp Ala Pro His Ala Arg Thr Glu Gln Glu Met Leu Glu				
	145	150	155	160
caa ggg cga tgc ctg ttg gcg atg ggc tgt ggc gca gtg cta atg aaa				528
Gln Gly Arg Ser Leu Leu Ala Met Gly Cys Gly Ala Val Leu Met Lys				
	165	170	175	
ggg ggt cat ctg gat gat gag caa agc ccg gac tgg ctg ttt acc cgc				576
Gly Gly His Leu Asp Asp Glu Gln Ser Pro Asp Trp Leu Phe Thr Arg				
	180	185	190	
gag ggt gaa caa cgg ttt acc gca ccg cgc att atg acc aaa aac acc				624
Glu Gly Glu Gln Arg Phe Thr Ala Pro Arg Ile Met Thr Lys Asn Thr				
	195	200	205	
cac ggc act ggt tgt aca ctc tct gcg gcg ttg gct gca cta cgc ccg				672
His Gly Thr Gly Cys Thr Leu Ser Ala Ala Leu Ala Ala Leu Arg Pro				
	210	215	220	
cgc cat aca aac tgg gct gac acc gta cag gag gca aaa agc tgg ctt				720
Arg His Thr Asn Trp Ala Asp Thr Val Gln Glu Ala Lys Ser Trp Leu				

225	230	235	240	
tca tcg gcg tta gcc cag gcc gac acg ctg gaa gtt ggt cac ggt att				768
Ser Ser Ala Leu Ala Gln Ala Asp Thr Leu Glu Val Gly His Gly Ile				
245	250	255		
ggg ccg gtt cac cac ttc cac gcc tgg tgg tga				801
Gly Pro Val His His Phe His Ala Trp Trp				
260	265			
<210>	225			
<211>	266			
<212>	PRT			
<213>	Escherichia coli			
<400>	225			
Met Lys Arg Ile Asn Ala Leu Thr Ile Ala Gly Thr Asp Pro Ser Gly				
1	5	10	15	
Gly Ala Gly Ile Gln Ala Asp Leu Lys Thr Phe Ser Ala Leu Gly Ala				
20	25	30		
Tyr Gly Cys Ser Val Ile Thr Ala Leu Val Ala Gln Asn Thr Arg Gly				
35	40	45		
Val Gln Ser Val Tyr Arg Ile Glu Pro Asp Phe Val Ala Ala Gln Leu				
50	55	60		
Asp Ser Val Phe Ser Asp Val Arg Ile Asp Thr Thr Lys Ile Gly Met				
65	70	75	80	
Leu Ala Glu Thr Asp Ile Val Glu Ala Val Ala Glu Arg Leu Gln Arg				
85	90	95		
Tyr Gln Ile Gln Asn Val Val Leu Asp Thr Val Met Leu Ala Lys Ser				
100	105	110		
Gly Asp Pro Leu Leu Ser Pro Ser Ala Val Ala Thr Leu Arg Ser Arg				
115	120	125		
Leu Leu Pro Gln Val Ser Leu Ile Thr Pro Asn Leu Pro Glu Ala Ala				
130	135	140		
Ala Leu Leu Asp Ala Pro His Ala Arg Thr Glu Gln Glu Met Leu Glu				

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

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

Gly Pro Val His His Phe His Ala Trp Trp
 260 265

<210> 226
 <211> 789
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(789)
 <223> ThiM gene from E. coli encoding a Hydroxyethylthiazole kinase
 <400> 226

atg caa gtc gac ctg ctg ggt tca gcg caa tct gcg cac gcg tta cac 48
 Met Gln Val Asp Leu Leu Gly Ser Ala Gln Ser Ala His Ala Leu His
 1 5 10 15
 ctt ttt cac caa cat tcc cct ctt gtg cac tgc atg acc aat gat gtg 96

Leu Phe His Gln His Ser Pro Leu Val His Cys Met Thr Asn Asp Val
 20 25 30
 gtg caa acc ttt acc gcc aat acc ttg ctg gcg ctc ggt gca tcg cca 144
 Val Gln Thr Phe Thr Ala Asn Thr Leu Leu Ala Leu Gly Ala Ser Pro
 35 40 45

gcg atg gtt atc gaa acc gaa gag gcc agt cag ttt gcg gct atc gcc 192
Ala Met Val Ile Glu Thr Glu Glu Ala Ser Gln Phe Ala Ala Ile Ala

50 55 60
agt gcc ttg ttg att aac gtt ggc aca ctg acg cag cca cgc gct cag 240
Ser Ala Leu Leu Ile Asn Val Gly Thr Leu Thr Gln Pro Arg Ala Gln

65 70 75 80
gcg atg cgt gct gcc gtt gag caa gca aaa agc tct caa aca ccc tgg 288
Ala Met Arg Ala Ala Val Glu Gln Ala Lys Ser Ser Gln Thr Pro Trp

85 90 95
acg ctt gat cca gta gcg gtg ggt gcg ctc gat tat cgc cgc cat ttt 336
Thr Leu Asp Pro Val Ala Val Gly Ala Leu Asp Tyr Arg Arg His Phe

100 105 110
tgt cat gaa ctt tta tct ttt aaa ccg gca gcg ata cgt ggt aat gct 384
Cys His Glu Leu Leu Ser Phe Lys Pro Ala Ala Ile Arg Gly Asn Ala

115 120 125
tcg gaa atc atg gca tta gct ggc att gct aat ggc gga cgg gga gtg 432

Ser Glu Ile Met Ala Leu Ala Gly Ile Ala Asn Gly Gly Arg Gly Val

130 135 140
gat acc act gac gcc gca gct aac gcg ata ccc gct gca caa aca ctg 480
Asp Thr Thr Asp Ala Ala Ala Asn Ala Ile Pro Ala Ala Gln Thr Leu

145 150 155 160
gca cgg gaa act ggc gca atc gtc gtg gtc act ggc gag atg gat tat 528
Ala Arg Glu Thr Gly Ala Ile Val Val Val Thr Gly Glu Met Asp Tyr

165 170 175
gtt acc gat gga cat cgt atc att ggt att cac ggt ggt gat ccg tta 576
Val Thr Asp Gly His Arg Ile Ile Gly Ile His Gly Gly Asp Pro Leu

180 185 190
atg acc aaa gtg gta gga act ggc tgt gca tta tcg gcg gtt gtc gct 624
Met Thr Lys Val Val Gly Thr Gly Cys Ala Leu Ser Ala Val Val Ala

195	200	205										
gcc tgc tgt gcg tta cca ggc gat acg ctg gaa aat gtc gca tct gcc												672
Ala Cys Cys Ala Leu Pro Gly Asp Thr Leu Glu Asn Val Ala Ser Ala												
210	215	220										
tgt cac tgg atg aaa caa gcc gga gaa cgc gca gtc gcc aga agc gag												720
Cys His Trp Met Lys Gln Ala Gly Glu Arg Ala Val Ala Arg Ser Glu												
225	230	235	240									
ggg cca ggc agt ttt gtt cca cat ttc ctt gat gcg ctc tgg caa ttg												768
Gly Pro Gly Ser Phe Val Pro His Phe Leu Asp Ala Leu Trp Gln Leu												
245	250	255										
acg cag gag gtg cag gca tga												789
Thr Gln Glu Val Gln Ala												
260												
<210>	227											
<211>	262											
<212>	PRT											
<213>	Escherichia coli											
<400>	227											
Met Gln Val Asp Leu Leu Gly Ser Ala Gln Ser Ala His Ala Leu His												
1	5	10	15									
Leu Phe His Gln His Ser Pro Leu Val His Cys Met Thr Asn Asp Val												
20	25	30										
Val Gln Thr Phe Thr Ala Asn Thr Leu Leu Ala Leu Gly Ala Ser Pro												
35	40	45										
Ala Met Val Ile Glu Thr Glu Glu Ala Ser Gln Phe Ala Ala Ile Ala												
50	55	60										
Ser Ala Leu Leu Ile Asn Val Gly Thr Leu Thr Gln Pro Arg Ala Gln												
65	70	75	80									
Ala Met Arg Ala Ala Val Glu Gln Ala Lys Ser Ser Gln Thr Pro Trp												
85	90	95										
Thr Leu Asp Pro Val Ala Val Gly Ala Leu Asp Tyr Arg Arg His Phe												

100 105 110
 Cys His Glu Leu Leu Ser Phe Lys Pro Ala Ala Ile Arg Gly Asn Ala
 115 120 125
 Ser Glu Ile Met Ala Leu Ala Gly Ile Ala Asn Gly Gly Arg Gly Val

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

195 200 205
 Ala Cys Cys Ala Leu Pro Gly Asp Thr Leu Glu Asn Val Ala Ser Ala
 210 215 220
 Cys His Trp Met Lys Gln Ala Gly Glu Arg Ala Val Ala Arg Ser Glu
 225 230 235 240
 Gly Pro Gly Ser Phe Val Pro His Phe Leu Asp Ala Leu Trp Gln Leu
 245 250 255
 Thr Gln Glu Val Gln Ala
 260

<210> 228

<211> 978

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(978)

<223> ThiL gene from E. coli encoding thiamine-phosphate kinase (Note:
 mutation at nucleotides 133-135 (GGT to GAC) encodes a G133D
 substitution)

<400> 228

atg gca tgt ggc gag ttc tcc ctg att gcc cgt tat ttt gac cgt gta

48

ccg ggc gat agc gcc gcc ggg ctg gcg att ttg caa aac cgt ttg cag 528

Pro Gly Asp Ser Ala Ala Gly Leu Ala Ile Leu Gln Asn Arg Leu Gln
165 170 175

gtt gcc gat gct aaa gat gcg gac tac ttg atc aaa cgt cat ctc cgt 576

Val Ala Asp Ala Lys Asp Ala Asp Tyr Leu Ile Lys Arg His Leu Arg
180 185 190

cca tcg ccg cgt att tta cag ggg cag gca ctg cgc gat ctg gca aat 624

Pro Ser Pro Arg Ile Leu Gln Gly Gln Ala Leu Arg Asp Leu Ala Asn
195 200 205

tca gcc atc gat ctc tct gac ggt ttg att tcc gat ctc ggg cat atc 672

Ser Ala Ile Asp Leu Ser Asp Gly Leu Ile Ser Asp Leu Gly His Ile
210 215 220

gtg aaa gcc agc gac tgc ggc gca cgt att gac ctg gca ttg ctg ccg 720

Val Lys Ala Ser Asp Cys Gly Ala Arg Ile Asp Leu Ala Leu Leu Pro
225 230 235 240

ttt tct gat gcg ctt tct cgc cat gtt gaa ccg gaa cag gcg ctg cgc 768

Phe Ser Asp Ala Leu Ser Arg His Val Glu Pro Glu Gln Ala Leu Arg
245 250 255

tgg gcg ctc tct ggc ggt gaa gat tac gag ttg tgt ttc act gtg ccg 816

Trp Ala Leu Ser Gly Gly Glu Asp Tyr Glu Leu Cys Phe Thr Val Pro
260 265 270

gaa ctg aac cgt ggc gcg ctg gat gtg gct ctc gga cac ctg ggc gta 864

Glu Leu Asn Arg Gly Ala Leu Asp Val Ala Leu Gly His Leu Gly Val
275 280 285

ccg ttt acc tgt atc ggg caa atg acc gcc gat atc gaa ggg ctt tgt 912

Pro Phe Thr Cys Ile Gly Gln Met Thr Ala Asp Ile Glu Gly Leu Cys
290 295 300

ttt att cgt gac ggc gaa cct gtt aca tta gac tgg aaa gga tat gac 960

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

305	310										315					320					
cat ttt gcc acg cca taa																				978	
His Phe Ala Thr Pro																					
325																					
<210> 229																					
<211> 325																					
<212> PRT																					
<213> Escherichia coli																					
<400> 229																					
Met	Ala	Cys	Gly	Glu	Phe	Ser	Leu	Ile	Ala	Arg	Tyr	Phe	Asp	Arg	Val						
1				5					10						15						
Arg	Ser	Ser	Arg	Leu	Asp	Val	Glu	Leu	Gly	Ile	Gly	Asp	Asp	Cys	Ala						
			20					25					30								
Leu	Leu	Asn	Ile	Pro	Glu	Lys	Gln	Thr	Leu	Ala	Ile	Ser	Thr	Asp	Thr						
		35					40					45									
Leu	Val	Ala	Gly	Asn	His	Phe	Leu	Pro	Asp	Ile	Asp	Pro	Ala	Asp	Leu						
	50					55				60											
Ala	Tyr	Lys	Ala	Leu	Ala	Val	Asn	Leu	Ser	Asp	Leu	Ala	Ala	Met	Gly						
65					70					75					80						
Ala	Asp	Pro	Ala	Trp	Leu	Thr	Leu	Ala	Leu	Thr	Leu	Pro	Asp	Val	Asp						
			85						90					95							
Glu	Ala	Trp	Leu	Glu	Ser	Phe	Ser	Asp	Ser	Leu	Phe	Asp	Leu	Leu	Asn						
		100							105				110								
Tyr	Tyr	Asp	Met	Gln	Leu	Ile	Gly	Gly	Asp	Thr	Thr	Arg	Gly	Pro	Leu						
		115						120				125									
Ser	Met	Thr	Leu	Gly	Ile	His	Gly	Phe	Val	Pro	Met	Gly	Arg	Ala	Leu						
	130					135					140										
Thr	Arg	Ser	Gly	Ala	Lys	Pro	Gly	Asp	Trp	Ile	Tyr	Val	Thr	Gly	Thr						
145					150				155					160							
Pro	Gly	Asp	Ser	Ala	Ala	Gly	Leu	Ala	Ile	Leu	Gln	Asn	Arg	Leu	Gln						
			165						170				175								
Val	Ala	Asp	Ala	Lys	Asp	Ala	Asp	Tyr	Leu	Ile	Lys	Arg	His	Leu	Arg						

180 185 190
 Pro Ser Pro Arg Ile Leu Gln Gly Gln Ala Leu Arg Asp Leu Ala Asn
 195 200 205

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

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

<210> 230

<211> 47

<212> DNA

<213> Escherichia coli

<220><221

> promoter

<222> (1)..(47)

<223> apFAB46 promoter

<400> 230

aaaaagagta ttgacttcgc atctttttgt acctataata gattcat

47

<210> 231

<211> 37

<212> DNA

<213> Escherichia coli

<220><221> promoter
 <222> (1)..(37)
 <223> apFAB70 promoter
 <400> 231
 ttgacatcgc atctttttgt acctataatg tgtggat 37
 <210> 232
 <211> 37
 <212> DNA
 <213> Escherichia coli
 <220><221> promoter
 <222> (1)..(37)
 <223> apFAB71 promoter
 <400> 232

 ttgacatcgc atctttttgt acctataata gattcat 37
 <210> 233
 <211> 29
 <212> DNA
 <213> Escherichia coli
 <220><221> promoter
 <222> (1)..(29)
 <223> pBAD Ara promoter
 <400> 233
 ctgacgcttt ttatcgcaac tctctactg 29
 <210> 234
 <211> 74
 <212> DNA
 <213> Escherichia coli
 <220><221> promoter
 <222> (1)..(74)
 <223> lac Promoter with lacO operator site
 <400> 234
 agctgtttcc tigtgtgaaat tggttatccgc tcacaattcc acacaacata cgagccggaa 60

 gcataaagtg taaa 74

<210> 235
 <211> 91
 <212> DNA
 <213> Escherichia coli
 <220><221> terminator
 <222> (1)..(91)
 <223> apFAB378 terminator
 <400> 235
 gagttggttag ctcttgatcc ggcaaacaaa ccaccgttgg tagcgggtggt ttttttgttt 60
 gcaagcagca gattacgcgc agaaaaaaag g 91
 <210> 236
 <211> 91
 <212> DNA
 <213> Escherichia coli
 <220><221> terminator
 <222> (1)..(91)
 <223> apFAB377 terminator
 <400> 236
 atgaccatct acattactga gctaataaca ggcttctggtg taatcgcagg cttttttatt 60
 tgggggagag ggaagtcattg aaaaaactaa c 91
 <210> 237
 <211> 90
 <212> DNA
 <213> Escherichia coli
 <220><221> terminator
 <222> (1)..(90)
 <223> apFAB381 terminator
 <400> 237
 accctcaaga gaaaatgtaa ccaactcact ggctcacctt cacgggtggg cttttcttcg 60
 ttccgggcat taaccctcac taacaggaga 90
 <210> 238
 <211> 258
 <212> DNA
 <213>

> Synthetic

<220><221> CDS

<222> (1)..(258)

<223> Synthetic gene encoding a E2 hybrid polypeptide (subunit of pyruvate dehydrogenase)

<400> 238

atg gct atc gaa atc aaa gta ccg gac atc ggg gct gat gaa gtt gaa 48

Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu

1 5 10 15

atc acc gag atc ctg gtc aaa gtg ggc gac aaa gtt gaa gcc gaa cag 96

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

20 25 30

tcg ctg atc acc gta gaa ggc gac aaa gct tct atg gaa gtt ccg gcg 144

Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ala

35 40 45

ccg ttt gca ggc gtc gtg aag gaa ctg aaa gtc aac gtt ggc gat aaa 192

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

50 55 60

gtg aaa act ggc tcg ctg att atg atc ttc gaa gtt gaa ggc gca gcg 240

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

65 70 75 80

cct gcg gca gct cct gcg 258

Pro Ala Ala Ala Pro Ala

85

<210> 239

<211> 86

<212> PRT

<213> Synthetic

<400> 239

Met Ala Ile Glu Ile Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu

1 5 10 15

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

20 25 30
 Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ala
 35 40 45
 Pro Phe Ala Gly Val Val Lys Glu Leu Lys Val Asn Val Gly Asp Lys
 50 55 60

 Val Lys Thr Gly Ser Leu Ile Met Ile Phe Glu Val Glu Gly Ala Ala
 65 70 75 80
 Pro Ala Ala Ala Pro Ala
 85
 <210> 240
 <211> 747
 <212> DNA
 <213> Escherichia coli
 <220><221> CDS
 <222> (1)..(747)
 <223> fpr gene_E. coli encoding a flavodoxin/ferredoxin reductase
 <400> 240

 atg gct gat tgg gta aca ggc aaa gtc act aaa gtg cag aac tgg acc 48
 Met Ala Asp Trp Val Thr Gly Lys Val Thr Lys Val Gln Asn Trp Thr

 1 5 10 15
 gac gcc ctg ttt agt ctc acc gtt cac gcc ccc gtg ctt ccg ttt acc 96
 Asp Ala Leu Phe Ser Leu Thr Val His Ala Pro Val Leu Pro Phe Thr
 20 25 30
 gcc ggg caa ttt acc aag ctt ggc ctt gaa atc gac ggc gaa cgc gtc 144
 Ala Gly Gln Phe Thr Lys Leu Gly Leu Glu Ile Asp Gly Glu Arg Val
 35 40 45

 cag cgc gcc tac tcc tat gta aac tcg ccc gat aat ccc gat ctg gag 192
 Gln Arg Ala Tyr Ser Tyr Val Asn Ser Pro Asp Asn Pro Asp Leu Glu
 50 55 60
 ttt tac ctg gtc acc gtc ccc gat ggc aaa tta agc cca cga ctg gcg 240
 Phe Tyr Leu Val Thr Val Pro Asp Gly Lys Leu Ser Pro Arg Leu Ala
 65 70 75 80

gca ctg aaa cca ggc gat gaa gtg cag gtg gtt agc gaa gcg gca gga 288

Ala Leu Lys Pro Gly Asp Glu Val Gln Val Val Ser Glu Ala Ala Gly
85 90 95

ttc ttt gtg ctc gat gaa gtg ccg cac tgc gaa acg cta tgg atg ctg 336

Phe Phe Val Leu Asp Glu Val Pro His Cys Glu Thr Leu Trp Met Leu
100 105 110

gca acc ggt aca gcg att ggc cct tat tta tcg att ctg caa cta ggt 384

Ala Thr Gly Thr Ala Ile Gly Pro Tyr Leu Ser Ile Leu Gln Leu Gly
115 120 125

aaa gat tta gat cgc ttc aaa aat ctg gtc ctg gtg cac gcc gca cgt 432

Lys Asp Leu Asp Arg Phe Lys Asn Leu Val Leu Val His Ala Ala Arg
130 135 140

tat gcc gcc gac tta agc tat ttg cca ctg atg cag gaa ctg gaa aaa 480

Tyr Ala Ala Asp Leu Ser Tyr Leu Pro Leu Met Gln Glu Leu Glu Lys
145 150 155 160

cgc tac gaa gga aaa ctg cgc att cag acg gtg gtc agt cgg gaa acg 528

Arg Tyr Glu Gly Lys Leu Arg Ile Gln Thr Val Val Ser Arg Glu Thr
165 170 175

gca gcg ggg tcg ctc acc gga cgg ata ccg gca tta att gaa agt ggg 576

Ala Ala Gly Ser Leu Thr Gly Arg Ile Pro Ala Leu Ile Glu Ser Gly
180 185 190

gaa ctg gaa agc acg att ggc ctg ccg atg aat aaa gaa acc agc cat 624

Glu Leu Glu Ser Thr Ile Gly Leu Pro Met Asn Lys Glu Thr Ser His
195 200 205

gtg atg ctg tgc ggc aat cca cag atg gtg cgc gat aca caa cag ttg 672

Val Met Leu Cys Gly Asn Pro Gln Met Val Arg Asp Thr Gln Gln Leu
210 215 220

ctg aaa gag acc cgg cag atg acg aaa cat tta cgt cgc cga ccg ggc 720

Leu Lys Glu Thr Arg Gln Met Thr Lys His Leu Arg Arg Arg Pro Gly

225	230	235	240	
cat atg aca gcg gag cat tac tgg taa				747
His Met Thr Ala Glu His Tyr Trp				
	245			
<210>	241			
<211>	248			
<212>	PRT			
<213>	Escherichia coli			
<400>	241			
Met Ala Asp Trp Val Thr Gly Lys Val Thr Lys Val Gln Asn Trp Thr				
1	5	10	15	
Asp Ala Leu Phe Ser Leu Thr Val His Ala Pro Val Leu Pro Phe Thr				
	20	25	30	
Ala Gly Gln Phe Thr Lys Leu Gly Leu Glu Ile Asp Gly Glu Arg Val				
	35	40	45	
Gln Arg Ala Tyr Ser Tyr Val Asn Ser Pro Asp Asn Pro Asp Leu Glu				
50	55	60		
Phe Tyr Leu Val Thr Val Pro Asp Gly Lys Leu Ser Pro Arg Leu Ala				
65	70	75	80	
Ala Leu Lys Pro Gly Asp Glu Val Gln Val Val Ser Glu Ala Ala Gly				
	85	90	95	
Phe Phe Val Leu Asp Glu Val Pro His Cys Glu Thr Leu Trp Met Leu				
	100	105	110	
Ala Thr Gly Thr Ala Ile Gly Pro Tyr Leu Ser Ile Leu Gln Leu Gly				
	115	120	125	
Lys Asp Leu Asp Arg Phe Lys Asn Leu Val Leu Val His Ala Ala Arg				
130	135	140		
Tyr Ala Ala Asp Leu Ser Tyr Leu Pro Leu Met Gln Glu Leu Glu Lys				
145	150	155	160	
Arg Tyr Glu Gly Lys Leu Arg Ile Gln Thr Val Val Ser Arg Glu Thr				
	165	170	175	
Ala Ala Gly Ser Leu Thr Gly Arg Ile Pro Ala Leu Ile Glu Ser Gly				

180 185 190
Glu Leu Glu Ser Thr Ile Gly Leu Pro Met Asn Lys Glu Thr Ser His
195 200 205

Val Met Leu Cys Gly Asn Pro Gln Met Val Arg Asp Thr Gln Gln Leu
210 215 220
Leu Lys Glu Thr Arg Gln Met Thr Lys His Leu Arg Arg Arg Pro Gly
225 230 235 240
His Met Thr Ala Glu His Tyr Trp
245

<210> 242
<211> 999
<212> DNA
<213> Bacillus subtilis
<220><221> CDS
<222> (1)..(999)
<223> YumC gene_B. subtilis encoding a flavodoxin/ferredoxin reductase
<

400> 242

atg cga gag gat aca aag gtt tat gat att aca att ata ggc ggg gga 48
Met Arg Glu Asp Thr Lys Val Tyr Asp Ile Thr Ile Ile Gly Gly Gly
1 5 10 15
ccg gtc ggc tta ttc acc gct ttt tac ggc ggg atg aga cag gca agc 96
Pro Val Gly Leu Phe Thr Ala Phe Tyr Gly Gly Met Arg Gln Ala Ser
20 25 30
gtc aaa att atc gaa agc ctg cct cag ctc ggc gga cag ctt agc gcc 144

Val Lys Ile Ile Glu Ser Leu Pro Gln Leu Gly Gly Gln Leu Ser Ala
35 40 45
cta tac cct gag aag tat ata tat gat gta gcg gga ttc ccg aaa atc 192
Leu Tyr Pro Glu Lys Tyr Ile Tyr Asp Val Ala Gly Phe Pro Lys Ile
50 55 60
cgc gcg caa gag ctt atc aat aac cta aaa gag caa atg gcg aaa ttc 240
Arg Ala Gln Glu Leu Ile Asn Asn Leu Lys Glu Gln Met Ala Lys Phe

65	70	75	80	
gac caa acc att tgt ctg gag caa gcg gtt gaa tct gtt gag aaa caa				288
Asp Gln Thr Ile Cys Leu Glu Gln Ala Val Glu Ser Val Glu Lys Gln				
	85	90	95	
gcg gac ggc gtg ttt aag ctt gta aca aat gaa gaa acc cac tac tct				336
Ala Asp Gly Val Phe Lys Leu Val Thr Asn Glu Glu Thr His Tyr Ser				
	100	105	110	
aaa acg gtc atc ata act gca gga aac ggc gca ttc aaa ccg aga aag				384
Lys Thr Val Ile Ile Thr Ala Gly Asn Gly Ala Phe Lys Pro Arg Lys				
	115	120	125	
ctg gaa ctt gaa aat gcc gag cag tat gaa ggc aaa aac ctc cat tac				432
Leu Glu Leu Glu Asn Ala Glu Gln Tyr Glu Gly Lys Asn Leu His Tyr				
	130	135	140	
ttc gtt gat gat ctg caa aaa ttc gcc ggc aga cgc gtt gcg atc ctt				480
Phe Val Asp Asp Leu Gln Lys Phe Ala Gly Arg Arg Val Ala Ile Leu				
145	150	155	160	
ggc ggt gga gat tcc gcg gtt gac tgg gcg ctt atg ctt gag cca atc				528
Gly Gly Gly Asp Ser Ala Val Asp Trp Ala Leu Met Leu Glu Pro Ile				
	165	170	175	
gca aaa gaa gta tcg atc att cac cgc cgc gac aag ttc cga gcg cac				576
Ala Lys Glu Val Ser Ile Ile His Arg Arg Asp Lys Phe Arg Ala His				
	180	185	190	
gag cac agt gtg gaa aac ctt cat gcg tcg aag gtt aat gtc ctg aca				624
Glu His Ser Val Glu Asn Leu His Ala Ser Lys Val Asn Val Leu Thr				
	195	200	205	
cca ttc gtc cct gcg gag ctg atc ggc gaa gac aaa atc gaa cag cta				672
Pro Phe Val Pro Ala Glu Leu Ile Gly Glu Asp Lys Ile Glu Gln Leu				
	210	215	220	
gtg ctt gaa gaa gtg aaa ggc gac cgc aaa gag att tta gaa att gat				720
Val Leu Glu Glu Val Lys Gly Asp Arg Lys Glu Ile Leu Glu Ile Asp				
225	230	235	240	

gac tta atc gtc aac tac ggt ttc gtt tca tct ctt gga ccg atc aaa 768

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

245 250 255

aac tgg ggc ctg gac atc gag aaa aat tcc att gtc gtg aaa tca aca 816

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

260 265 270

atg gaa aca aat atc gaa ggc ttc ttt gca gca ggt gac att tgt aca 864

Met Glu Thr Asn Ile Glu Gly Phe Phe Ala Ala Gly Asp Ile Cys Thr

275 280 285

tac gaa gga aaa gtc aac ctg att gcc agc ggc ttc ggc gag gca ccg 912

Tyr Glu Gly Lys Val Asn Leu Ile Ala Ser Gly Phe Gly Glu Ala Pro

290 295 300

aca gca gtg aac aac gcc aag gct tac atg gac ccg aaa gcc cgc gta 960

Thr Ala Val Asn Asn Ala Lys Ala Tyr Met Asp Pro Lys Ala Arg Val

305 310 315 320

cag cct ctt cac tca aca agt ctt ttt gaa aat aaa taa 999

Gln Pro Leu His Ser Thr Ser Leu Phe Glu Asn Lys

325 330

<210> 243

<211> 332

<212> PRT

<213> Bacillus subtilis

<400> 243

Met Arg Glu Asp Thr Lys Val Tyr Asp Ile Thr Ile Ile Gly Gly Gly

1 5 10 15

Pro Val Gly Leu Phe Thr Ala Phe Tyr Gly Gly Met Arg Gln Ala Ser

20 25 30

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

35 40 45

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

50 55 60

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

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

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

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

305 310 315 320
Gln Pro Leu His Ser Thr Ser Leu Phe Glu Asn Lys
325 330
<210> 244
<211> 780
<212> DNA
<213> Pseudomonas pudita
<220><221> CDS
<222> (1)..(780)
<223> fpr-1 gene from Pseudomonas putida KT2440 encoding a
flavodoxin/ferredoxin reductase
<400> 244
atg agc aac atg aac cac gaa cgt gtc ctc agt gtg cac cac tgg aac 48
Met Ser Asn Met Asn His Glu Arg Val Leu Ser Val His His Trp Asn
1 5 10 15
gac acc ctg ttc agc ttc aag tgc acc cgc gac ccg ggc ctg cgc ttc 96
Asp Thr Leu Phe Ser Phe Lys Cys Thr Arg Asp Pro Gly Leu Arg Phe
20 25 30
gag aac ggt cag ttc gtg atg atc ggc ctg cag cag gac aac ggc cgt 144
Glu Asn Gly Gln Phe Val Met Ile Gly Leu Gln Gln Asp Asn Gly Arg
35 40 45
ccg ctc atg cgt gcc tac tcc atc gct tgc cca aac tgg gaa gag cac 192
Pro Leu Met Arg Ala Tyr Ser Ile Ala Ser Pro Asn Trp Glu Glu His
50 55 60
ctt gaa ttc ttc agc atc aag gtg ccg gac ggc ccg ctg acc tgc cag 240
Leu Glu Phe Phe Ser Ile Lys Val Pro Asp Gly Pro Leu Thr Ser Gln
65 70 75 80
ctg cag cac ctg aag gaa ggc gat gag atc atc atc agc aag aag cct 288
Leu Gln His Leu Lys Glu Gly Asp Glu Ile Ile Ile Ser Lys Lys Pro
85 90 95
acc ggc acc ctg gtc ctc gac gac ctg aat cct ggc aag cac ctg tac 336
Thr Gly Thr Leu Val Leu Asp Asp Leu Asn Pro Gly Lys His Leu Tyr

100	105	110	
ctg ctg agc acc ggc act ggt ctg gcg ccg ttc atg agc gtc atc cag			384
Leu Leu Ser Thr Gly Thr Gly Leu Ala Pro Phe Met Ser Val Ile Gln			
115	120	125	
gac ccg gaa acc tac gag cgc ttt gaa aaa gtg atc ctg gtg cac ggc			432
Asp Pro Glu Thr Tyr Glu Arg Phe Glu Lys Val Ile Leu Val His Gly			
130	135	140	
gtg cgc tat gtg aac gaa gtg gcc tac cgc gag ttc atc acc gag cac			480
Val Arg Tyr Val Asn Glu Val Ala Tyr Arg Glu Phe Ile Thr Glu His			
145	150	155	160
ctg ccg cag aac gag ttc ttc ggt gag tcg gtt cgc gac aag ctg atc			528
Leu Pro Gln Asn Glu Phe Phe Gly Glu Ser Val Arg Asp Lys Leu Ile			
165	170	175	
tac tac ccg acc gtg acc cgc gag ccg ttc gaa aac cag ggc cgt ctg			576
Tyr Tyr Pro Thr Val Thr Arg Glu Pro Phe Glu Asn Gln Gly Arg Leu			
180	185	190	
acc gac ctg atg cgc agc ggc aag ctg ttc agc gac atc ggc ctg ccg			624
Thr Asp Leu Met Arg Ser Gly Lys Leu Phe Ser Asp Ile Gly Leu Pro			
195	200	205	
ccg atc aac ccg caa gac gac cgc gcg atg atc tgc ggc agc ccg agc			672
Pro Ile Asn Pro Gln Asp Asp Arg Ala Met Ile Cys Gly Ser Pro Ser			
210	215	220	
atg ctc gac gag acc agc gaa gtg ctg gac agc ttc ggc ctg aag atc			720
Met Leu Asp Glu Thr Ser Glu Val Leu Asp Ser Phe Gly Leu Lys Ile			
225	230	235	240
tcc ccg cgc atg cgc gag ccg ggt gac tac ctg atc gaa cgt gcc ttc			768
Ser Pro Arg Met Arg Glu Pro Gly Asp Tyr Leu Ile Glu Arg Ala Phe			
245	250	255	
gtc gag aag taa			780
Val Glu Lys			

<210> 245

<211> 259

<212> PRT

<213> *Pseudomonas pudita*

<400> 245

Met Ser Asn Met Asn His Glu Arg Val Leu Ser Val His His Trp Asn

1 5 10 15

Asp Thr Leu Phe Ser Phe Lys Cys Thr Arg Asp Pro Gly Leu Arg Phe

20 25 30

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

35 40 45

Pro Leu Met Arg Ala Tyr Ser Ile Ala Ser Pro Asn Trp Glu Glu His

50 55 60

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

65 70 75 80

Leu Gln His Leu Lys Glu Gly Asp Glu Ile Ile Ile Ser Lys Lys Pro

85 90 95

Thr Gly Thr Leu Val Leu Asp Asp Leu Asn Pro Gly Lys His Leu Tyr

100 105 110

Leu Leu Ser Thr Gly Thr Gly Leu Ala Pro Phe Met Ser Val Ile Gln

115 120 125

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

130 135 140

Val Arg Tyr Val Asn Glu Val Ala Tyr Arg Glu Phe Ile Thr Glu His

145 150 155 160

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

165 170 175

Tyr Tyr Pro Thr Val Thr Arg Glu Pro Phe Glu Asn Gln Gly Arg Leu

180 185 190

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

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

245 250 255
 Val Glu Lys

<210> 246

<211> 936

<212> DNA

<213> Streptomyces venezuelae

<220><221> CDS

<222> (1)..(936)

<223> SVEN_0113 gene from Streptomyces venezuelae encoding a
 flavodoxin/ferredoxin reductase

<400> 246

atg agc gag aac ccg ctg caa ctg atc gtc cac cgc atg aca cgg gag 48
 Met Ser Glu Asn Pro Leu Gln Leu Ile Val His Arg Met Thr Arg Glu
 1 5 10 15

gcc gag ggc gta ctg tcc gtc gaa ctc gcc cac ccc gac ggc aag ccg 96
 Ala Glu Gly Val Leu Ser Val Glu Leu Ala His Pro Asp Gly Lys Pro
 20 25 30

ctg ccc gcc tgg acg ccg ggc gcc cac atc gac gtc cac gtc ggg ggc 144
 Leu Pro Ala Trp Thr Pro Gly Ala His Ile Asp Val His Val Gly Gly
 35 40 45

cac gtc cgc cag tac agc ctg tgc ggc gac ccg cac gac cag ggc gcg 192

His Val Arg Gln Tyr Ser Leu Cys Gly Asp Pro His Asp Gln Gly Ala
 50 55 60
 tac cgg atc ggc gtc ctc gac gaa ccc gcc tca cgc ggc ggt tcg cgc 240
 Tyr Arg Ile Gly Val Leu Asp Glu Pro Ala Ser Arg Gly Gly Ser Arg

65	70	75	80	
ttc gtg cac acc gca ctg cgc ccc ggc cag acc ctc acg gtc tcc gca				288
Phe Val His Thr Ala Leu Arg Pro Gly Gln Thr Leu Thr Val Ser Ala				
	85	90	95	
ccc cgc aac cac ttc gcc ctc gag gac gcc gcc gcg tac gtc ctc gtc				336
Pro Arg Asn His Phe Ala Leu Glu Asp Ala Ala Ala Tyr Val Leu Val				
	100	105	110	
gcc ggc ggc atc ggc atc acg ccc ctg ctc gcc atg gcc cgc gag gcg				384
Ala Gly Gly Ile Gly Ile Thr Pro Leu Leu Ala Met Ala Arg Glu Ala				
	115	120	125	
gcc cgc cgg ggc gcc gag tgg cgc ctg gtc tac ggc ggc cgg agc cgg				432
Ala Arg Arg Gly Ala Glu Trp Arg Leu Val Tyr Gly Gly Arg Ser Arg				
	130	135	140	
gcg tcg atg gcc ttc acc gcc gaa ctg gcc ctg ctc ggc ggc gag gtg				480
Ala Ser Met Ala Phe Thr Ala Glu Leu Ala Leu Leu Gly Gly Glu Val				
	145	150	155	160
acc ctc gtc ccg cag gac gaa cgc ggc cac atc gac ctg gac gcc gag				528
Thr Leu Val Pro Gln Asp Glu Arg Gly His Ile Asp Leu Asp Ala Glu				
	165	170	175	
ctg tcc cgg ctg ccc gac ggc gcc ctc gtc tac gcc tgc ggc ccg gaa				576
Leu Ser Arg Leu Pro Asp Gly Ala Leu Val Tyr Ala Cys Gly Pro Glu				
	180	185	190	
ccc ctc ctc gcg gcc gtc gag gaa cgc tgt ccg caa gga cag ctg cgc				624
Pro Leu Leu Ala Ala Val Glu Glu Arg Cys Pro Gln Gly Gln Leu Arg				
	195	200	205	
acc gaa cgg ttc acc gcc ccc acc gtc gaa cgc gca gaa gac gac gga				672
Thr Glu Arg Phe Thr Ala Pro Thr Val Glu Arg Ala Glu Asp Asp Gly				
	210	215	220	
gag ttc gag gtc gag tgc cgc acc tcg ggc ctg acg ctc cgg gtc gac				720
Glu Phe Glu Val Glu Cys Arg Thr Ser Gly Leu Thr Leu Arg Val Asp				

225	230	235	240	
gca cac tcc tgc atc ctc gac gcc gcc gag aac gcc ggg atc gcc gtc				768
Ala His Ser Ser Ile Leu Asp Ala Ala Glu Asn Ala Gly Ile Ala Val				
245 250 255				
gac agc tcc tgc cgc gac ggc atc tgc ggc tcc tgc gag acc cgc gtc				816
Asp Ser Ser Cys Arg Asp Gly Ile Cys Gly Ser Cys Glu Thr Arg Val				
260 265 270				
ctc gaa ggc acc ccg gac cac cgc gac ttc ctc ctc agc gag gcg gaa				864
Leu Glu Gly Thr Pro Asp His Arg Asp Phe Leu Leu Ser Glu Ala Glu				
275 280 285				
cag gcc gcc ggc gcc acc atg atg atc tgc gtc tgc cgg tgc gcc tcc				912
Gln Ala Ala Gly Ala Thr Met Met Ile Cys Val Ser Arg Cys Ala Ser				
290 295 300				
ggc cgg ctc gtc ctc gac ctg tga				936
Gly Arg Leu Val Leu Asp Leu				
305 310				
<210> 247				
<211> 311				
<212> PRT				
<213> Streptomyces venezuelae				
<400> 247				
Met Ser Glu Asn Pro Leu Gln Leu Ile Val His Arg Met Thr Arg Glu				
1 5 10 15				
Ala Glu Gly Val Leu Ser Val Glu Leu Ala His Pro Asp Gly Lys Pro				
20 25 30				
Leu Pro Ala Trp Thr Pro Gly Ala His Ile Asp Val His Val Gly Gly				
35 40 45				
His Val Arg Gln Tyr Ser Leu Cys Gly Asp Pro His Asp Gln Gly Ala				
50 55 60				
Tyr Arg Ile Gly Val Leu Asp Glu Pro Ala Ser Arg Gly Gly Ser Arg				
65 70 75 80				

Phe Val His Thr Ala Leu Arg Pro Gly Gln Thr Leu Thr Val Ser Ala
 85 90 95
 Pro Arg Asn His Phe Ala Leu Glu Asp Ala Ala Ala Tyr Val Leu Val
 100 105 110

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

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

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

 Gly Arg Leu Val Leu Asp Leu
 305 310
 <210> 248

<211> 978

<212> DNA

<213> Corynebacterium glutamicum

<220><221> CDS

<222> (1)..(978)

<223> Cgl2384 gene from Corynebacterium glutamicum encoding a
flavodoxin/ferredoxin reductase

<400> 248

atg aac tcg caa tgg caa gat gca cat gtt gtt tcc agc gaa atc atc 48

Met Asn Ser Gln Trp Gln Asp Ala His Val Val Ser Ser Glu Ile Ile

1 5 10 15

gct gca gac att cgg cga ata gaa cta tcc ccg aaa ttt gcg att cca 96

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

20 25 30

gta aaa ccc ggc gaa cat ctc aag atc atg gtg ccc cta aaa act gga 144

Val Lys Pro Gly Glu His Leu Lys Ile Met Val Pro Leu Lys Thr Gly

35 40 45

cag gaa aag aga tcg tac tcc atc gtt gac gct cgt cac gac ggt tcg 192

Gln Glu Lys Arg Ser Tyr Ser Ile Val Asp Ala Arg His Asp Gly Ser

50 55 60

act ctc gcc ctg agc gta ctc aaa acc aga aac tcc cgt gga gga tct 240

Thr Leu Ala Leu Ser Val Leu Lys Thr Arg Asn Ser Arg Gly Gly Ser

65 70 75 80

gag ttc atg cat acg ctt cga gct gga gac aca gtt act gtc tcc agg 288

Glu Phe Met His Thr Leu Arg Ala Gly Asp Thr Val Thr Val Ser Arg

85 90 95

ccg tct cag gat ttt cct ctc cgc gtg ggt gcg cct gag tat gta ctt 336

Pro Ser Gln Asp Phe Pro Leu Arg Val Gly Ala Pro Glu Tyr Val Leu

100 105 110

gtt gcc ggc gga att gga atc aca gcg atc cgt tca atg gca tct tta 384

Val Ala Gly Gly Ile Gly Ile Thr Ala Ile Arg Ser Met Ala Ser Leu

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

Gln Val Arg Val Leu Glu Val Asp Gly Gln Val Asp His Arg Asp Val
275 280 285
ttc ttc tct gat cgt caa aaa gaa tcc gac gca aag gca tgc gcc tgc 912
Phe Phe Ser Asp Arg Gln Lys Glu Ser Asp Ala Lys Ala Cys Ala Cys
290 295 300
gtg tct cga gta gtc tcc tcc cct tcc tgc tcc cca acc tgc acc att 960
Val Ser Arg Val Val Ser Ser Pro Ser Ser Ser Pro Thr Ser Thr Ile
305 310 315 320
acg gtc gcc ctc tcc taa 978
Thr Val Ala Leu Ser
325
<210> 249
<211> 325
<212> PRT
<213> Corynebacterium glutamicum
<400> 249
Met Asn Ser Gln Trp Gln Asp Ala His Val Val Ser Ser Glu Ile Ile
1 5 10 15
Ala Ala Asp Ile Arg Arg Ile Glu Leu Ser Pro Lys Phe Ala Ile Pro
20 25 30
Val Lys Pro Gly Glu His Leu Lys Ile Met Val Pro Leu Lys Thr Gly
35 40 45
Gln Glu Lys Arg Ser Tyr Ser Ile Val Asp Ala Arg His Asp Gly Ser
50 55 60
Thr Leu Ala Leu Ser Val Leu Lys Thr Arg Asn Ser Arg Gly Gly Ser
65 70 75 80
Glu Phe Met His Thr Leu Arg Ala Gly Asp Thr Val Thr Val Ser Arg
85 90 95
Pro Ser Gln Asp Phe Pro Leu Arg Val Gly Ala Pro Glu Tyr Val Leu
100 105 110
Val Ala Gly Gly Ile Gly Ile Thr Ala Ile Arg Ser Met Ala Ser Leu
115 120 125

Leu Lys Lys Leu Gly Ala Asn Tyr Arg Ile His Phe Ala Ala Arg Ser
130 135 140

Leu Asp Ala Met Ala Tyr Lys Asp Glu Leu Val Ala Glu His Gly Asp
145 150 155 160

Lys Leu His Leu His Leu Asp Ser Glu Gly Thr Thr Ile Asp Val Pro
165 170 175

Ala Leu Ile Glu Thr Leu Asn Pro His Thr Glu Leu Tyr Met Cys Gly
180 185 190

Pro Ile Arg Leu Met Asp Ala Ile Arg Arg Ala Trp Asn Thr Arg Gly
195 200 205

Leu Asp Pro Thr Asn Leu Arg Phe Glu Thr Phe Gly Asn Ser Gly Trp
210 215 220

Phe Ser Pro Glu Val Phe His Ile Gln Val Pro Glu Leu Gly Leu His
225 230 235 240

Ala Thr Val Asn Lys Asp Glu Ser Met Leu Glu Ala Leu Gln Lys Ala
245 250 255

Gly Ala Asn Met Met Phe Asp Cys Arg Lys Gly Glu Cys Gly Leu Cys
260 265 270

Gln Val Arg Val Leu Glu Val Asp Gly Gln Val Asp His Arg Asp Val
275 280 285

Phe Phe Ser Asp Arg Gln Lys Glu Ser Asp Ala Lys Ala Cys Ala Cys
290 295 300

Val Ser Arg Val Val Ser Ser Pro Ser Ser Pro Thr Ser Thr Ile
305 310 315 320

Thr Val Ala Leu Ser
325

<210> 250

<211> 921

<212> DNA

<213> Sphingobacterium sp

<220

><221> CDS

<222> (1)..(921)

<223> SJN15614.1 from *Sphingobacterium* sp. JB170 encoding a
flavodoxin/ferredoxin reductase

<400> 250

atg ttt ggc gat cgc gag gtt cgt aga tcg tat tcc ttt agt agc tcg 48

Met Phe Gly Asp Arg Glu Val Arg Arg Ser Tyr Ser Phe Ser Ser Ser

1 5 10 15

cct gca gtt gtt gag ccg cta gcc att acc gta aaa aga gtg gat aat 96

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

20 25 30

ggg gaa att tcc cgc ctg ttg cat cat cgt aca cgc gtt ggg gat ctt 144

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

35 40 45

gtt gat gtt cta gcc ccg cag gga tta ttt aca tac gaa cct gac ccc 192

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

50 55 60

act aca gct cga aca tta ttt tta ttt ggc gcc ggc gtt ggt gtt act 240

Thr Thr Ala Arg Thr Leu Phe Leu Phe Gly Ala Gly Val Gly Val Thr

65 70 75 80

ccg tta ttt tcc atc ctg aaa act gcg ctg tcc aca gaa ccc aaa acg 288

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

85 90 95

aag gtt gtc ctc att tat agc aac agt tca ccc gat agg aca gtt ttt 336

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

100 105 110

aaa gtt gaa ctt gaa cat tgg caa caa ctg tat gcc gat cgg ctt gaa 384

Lys Val Glu Leu Glu His Trp Gln Gln Leu Tyr Ala Asp Arg Leu Glu

115 120 125

att ata tgg att tac tcc aat tca aaa aat ctg tta aat gca cac cta 432

Ile Ile Trp Ile Tyr Ser Asn Ser Lys Asn Leu Leu Asn Ala His Leu

130 135 140

aac	cgc	gag	aac	tta	tta	cgc	ttt	gtc	aat	gaa	cgc	atg	cct	gag	gat	
Asn	Arg	Glu	Asn	Leu	Leu	Arg	Phe	Val	Asn	Glu	Arg	Met	Pro	Glu	Asp	
145				150				155				160				
aac	aat	gct	ata	ttt	ttc	acc	tgt	ggc	ccg	gta	ttt	tac	atg	gac	tta	
Asn	Asn	Ala	Ile	Phe	Phe	Thr	Cys	Gly	Pro	Val	Phe	Tyr	Met	Asp	Leu	
				165				170				175				
gta	cgc	ttc	acg	tta	ctc	ggc	ctt	gga	atc	ccc	gac	gag	gat	atc	cgc	
Val	Arg	Phe	Thr	Leu	Leu	Gly	Leu	Gly	Ile	Pro	Asp	Glu	Asp	Ile	Arg	
180								185				190				
aag	gag	aca	ttt	cat	ttt	cct	gaa	gaa	gaa	gat	gat	gaa	gat	gag	aaa	
Lys	Glu	Thr	Phe	His	Phe	Pro	Glu	Glu	Glu	Asp	Asp	Glu	Asp	Glu	Lys	
195								200				205				
gaa	gac	gat	ccg	gtt	gat	acg	aca	gcc	tac	aac	ata	ttg	ctc	agg	ttt	
672																
Glu	Asp	Asp	Pro	Val	Asp	Thr	Thr	Ala	Tyr	Asn	Ile	Leu	Leu	Arg	Phe	
210				215				220								
caa	gga	caa	gaa	tac	ccg	ttg	aca	att	cca	tac	aac	aaa	aca	atc	ttg	
720																
Gln	Gly	Gln	Glu	Tyr	Pro	Leu	Thr	Ile	Pro	Tyr	Asn	Lys	Thr	Ile	Leu	
225	230				235				240							
cag	gcc	gga	ctg	gat	aat	aat	att	aaa	ctt	ccc	tat	tct	tgt	aaa	tcg	
768																
Gln	Ala	Gly	Leu	Asp	Asn	Asn	Ile	Lys	Leu	Pro	Tyr	Ser	Cys	Lys	Ser	
245				250				255								
ggg	atg	tgt	agt	act	tgt	atc	tca	caa	tgt	tcc	agc	ggg	tcc	gtc	cgg	
816																
Gly	Met	Cys	Ser	Thr	Cys	Ile	Ser	Gln	Cys	Ser	Ser	Gly	Ser	Val	Arg	
260				265				270								
atg	gat	tac	aat	gag	gtt	cta	aca	gat	cgt	gag	gtt	gaa	aat	ggc	cgt	
864																
Met	Asp	Tyr	Asn	Glu	Val	Leu	Thr	Asp	Arg	Glu	Val	Glu	Asn	Gly	Arg	
275				280				285								
tgt	ttg	att	tgt	act	tcc	cac	ccg	tta	gaa	gat	ggc	acg	acg	att	gat	
912																
Cys	Leu	Ile	Cys	Thr	Ser	His	Pro	Leu	Glu	Asp	Gly	Thr	Thr	Ile	Asp	
290				295				300								
gta	gtt	taa														
921																

Val Val

305

<210> 251

<211> 306

<212> PRT

<213> Sphingobacterium sp

<400> 251

Met Phe Gly Asp Arg Glu Val Arg Arg Ser Tyr Ser Phe Ser Ser Ser

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Thr Thr Ala Arg Thr Leu Phe Leu Phe Gly Ala Gly Val Gly Val Thr

65 70 75 80

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

85 90 95

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

100 105 110

Lys Val Glu Leu Glu His Trp Gln Gln Leu Tyr Ala Asp Arg Leu Glu

115 120 125

Ile Ile Trp Ile Tyr Ser Asn Ser Lys Asn Leu Leu Asn Ala His Leu

130 135 140

Asn Arg Glu Asn Leu Leu Arg Phe Val Asn Glu Arg Met Pro Glu Asp

145 150 155 160

Asn Asn Ala Ile Phe Phe Thr Cys Gly Pro Val Phe Tyr Met Asp Leu

165 170 175

Val Arg Phe Thr Leu Leu Gly Leu Gly Ile Pro Asp Glu Asp Ile Arg

180 185 190

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

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

Val Val

305

<210> 252

<211> 3525

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(3525)

<223> ydbK gene from E. coli encoding pyruvate-flavodoxin/ferredoxin
 oxidoreductase

<400> 252

atg att act att gac ggt aat ggc gcg gtt gct tcg gtc gca ttt cgc 48

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

1 5 10 15

acc agt gaa gtt atc gcc atc tac cct att acc ccc agt tcc acg atg 96

Thr Ser Glu Val Ile Ala Ile Tyr Pro Ile Thr Pro Ser Ser Thr Met

20 25 30

gca gaa cag gct gat gcc tgg gcc gga aac ggc tta aag aac gtt tgg 144

Ala Glu Gln Ala Asp Ala Trp Ala Gly Asn Gly Leu Lys Asn Val Trp	
35 40 45	
gga gac aca cca cgc gtg gtt gaa atg cag tcg gaa gcg ggt gct atc	192
Gly Asp Thr Pro Arg Val Val Glu Met Gln Ser Glu Ala Gly Ala Ile	
50 55 60	
gct acc gtg cat ggc gct ttg cag acg ggt gcc ctt tca aca tcg ttt	240
Ala Thr Val His Gly Ala Leu Gln Thr Gly Ala Leu Ser Thr Ser Phe	
65 70 75 80	
acg tca tcg cag ggt ttg ctg ctg atg atc ccg acg ctg tac aaa ctg	288
Thr Ser Ser Gln Gly Leu Leu Leu Met Ile Pro Thr Leu Tyr Lys Leu	
85 90 95	
gca ggc gaa cta aca ccg ttt gtc ctg cat gta gcg gca cgt acc gtt	336
Ala Gly Glu Leu Thr Pro Phe Val Leu His Val Ala Ala Arg Thr Val	
100 105 110	
gcc aca cat gca ctc tct att ttt ggc gat cat tcc gac gtt atg gcg	384
Ala Thr His Ala Leu Ser Ile Phe Gly Asp His Ser Asp Val Met Ala	
115 120 125	
gtg cgc cag acg ggt tgc gcg atg ttg tgt gca gca aac gtc cag gaa	432
Val Arg Gln Thr Gly Cys Ala Met Leu Cys Ala Ala Asn Val Gln Glu	
130 135 140	
gcg caa gac ttt gct ctc att tcg caa atc gcg acg ctg aaa agc cgc	480
Ala Gln Asp Phe Ala Leu Ile Ser Gln Ile Ala Thr Leu Lys Ser Arg	
145 150 155 160	
gtg cca ttt att cat ttc ttt gat ggt ttc cgc acg tcc cac gaa atc	528
Val Pro Phe Ile His Phe Phe Asp Gly Phe Arg Thr Ser His Glu Ile	
165 170 175	
aat aaa att gtc ccg ctg gcc gat gac acg att ctt gat ctc atg ccg	576
Asn Lys Ile Val Pro Leu Ala Asp Asp Thr Ile Leu Asp Leu Met Pro	
180 185 190	
cag gtc gaa att gat gct cat cgc gcc cgg gca ctc aac ccg gaa cat	624
Gln Val Glu Ile Asp Ala His Arg Ala Arg Ala Leu Asn Pro Glu His	

195	200	205	
ccg gtg atc cgc ggt acg tcc gcc aat cct gac act tat ttc cag tct			672
Pro Val Ile Arg Gly Thr Ser Ala Asn Pro Asp Thr Tyr Phe Gln Ser			
210	215	220	
cgc gaa gcc acc aac cca tgg tac aac gcg gtc tat gac cat gtt gaa			720
Arg Glu Ala Thr Asn Pro Trp Tyr Asn Ala Val Tyr Asp His Val Glu			
225	230	235	240
cag gcg atg aat gat ttc tct gcc gcg aca ggt cgt cag tat cag ccg			768
Gln Ala Met Asn Asp Phe Ser Ala Ala Thr Gly Arg Gln Tyr Gln Pro			
245	250	255	
ttt gaa tat tac ggg cat ccg caa gcg gaa cgg gtg att atc ctg atg			816
Phe Glu Tyr Tyr Gly His Pro Gln Ala Glu Arg Val Ile Ile Leu Met			
260	265	270	
ggc tct gcc att ggc acc tgt gaa gaa gtg gtt gat gaa ttg cta acc			864
Gly Ser Ala Ile Gly Thr Cys Glu Glu Val Val Asp Glu Leu Leu Thr			
275	280	285	
cgt ggc gaa aaa gtt ggc gtg ctg aaa gtt cgc ctg tac cgc ccc ttc			912
Arg Gly Glu Lys Val Gly Val Leu Lys Val Arg Leu Tyr Arg Pro Phe			
290	295	300	
tcc gct aaa cat tta ctg caa gct ctg ccg gga tcc gta cgc agc gtg			960
Ser Ala Lys His Leu Leu Gln Ala Leu Pro Gly Ser Val Arg Ser Val			
305	310	315	320
gcg gta ctg gac aga acc aaa gaa ccc ggt gcc cag gca gaa ccg ctc			1008
Ala Val Leu Asp Arg Thr Lys Glu Pro Gly Ala Gln Ala Glu Pro Leu			
325	330	335	
tat ctg gat gta atg acc gca ctg gca gaa gcc ttt aat aat ggc gag			1056
Tyr Leu Asp Val Met Thr Ala Leu Ala Glu Ala Phe Asn Asn Gly Glu			
340	345	350	
cgc gaa act ctg ccc cgt gtc att ggt ggg cgc tat ggt ctt tca tcc			1104
Arg Glu Thr Leu Pro Arg Val Ile Gly Gly Arg Tyr Gly Leu Ser Ser			

355	360	365	
aaa gaa ttt ggc cca gac tgt gta ctg gcg gta ttt gcc gag ctc aac			1152
Lys Glu Phe Gly Pro Asp Cys Val Leu Ala Val Phe Ala Glu Leu Asn			
370	375	380	
gcg gct aaa ccg aaa gcg cgc ttt acg gtt ggt att tac gat gat gtg			1200
Ala Ala Lys Pro Lys Ala Arg Phe Thr Val Gly Ile Tyr Asp Asp Val			
385	390	395	400
acc aat ctg tca ctg ccg ttg ccg gaa aac acc ctg cca aac tcg gcg			1248
Thr Asn Leu Ser Leu Pro Leu Pro Glu Asn Thr Leu Pro Asn Ser Ala			
405	410	415	
aaa ctg gaa gcc ttg ttt tat ggc ctt ggt agt gat ggc agc gtt tcc			1296
Lys Leu Glu Ala Leu Phe Tyr Gly Leu Gly Ser Asp Gly Ser Val Ser			
420	425	430	
gcg acc aaa aac aat atc aag att atc ggt aat tcc acg ccg tgg tac			1344
Ala Thr Lys Asn Asn Ile Lys Ile Ile Gly Asn Ser Thr Pro Trp Tyr			
435	440	445	
gca cag ggc tat ttt gtt tac gac tcc aaa aag gcg ggc ggc ctg acg			1392
Ala Gln Gly Tyr Phe Val Tyr Asp Ser Lys Lys Ala Gly Gly Leu Thr			
450	455	460	
gtt tct cac ctt cga gtg agc gaa cag ccg att cgt tcc gct tat ctc			1440
Val Ser His Leu Arg Val Ser Glu Gln Pro Ile Arg Ser Ala Tyr Leu			
465	470	475	480
att tcc cag gct gat ttt gtt ggc tgc cac cag ttg cag ttt atc gat			1488
Ile Ser Gln Ala Asp Phe Val Gly Cys His Gln Leu Gln Phe Ile Asp			
485	490	495	
aaa tat cag atg gct gag cgt tta aaa cct ggc ggc att ttc ctg ctc			1536
Lys Tyr Gln Met Ala Glu Arg Leu Lys Pro Gly Gly Ile Phe Leu Leu			
500	505	510	
aac acg ccg tac agc gca gat gaa gtg tgg tcg cgc ttg ccg caa gaa			1584

Asn Thr Pro Tyr Ser Ala Asp Glu Val Trp Ser Arg Leu Pro Gln Glu
515 520 525
gtt cag gcc gtg tta aac cag aaa aaa gcg cgc ttc tat gtg att aac 1632
Val Gln Ala Val Leu Asn Gln Lys Lys Ala Arg Phe Tyr Val Ile Asn
530 535 540
gcg gcg aaa atc gcc cgc gaa tgt ggc ctg gcg gcc cgt att aat acc 1680
Ala Ala Lys Ile Ala Arg Glu Cys Gly Leu Ala Ala Arg Ile Asn Thr
545 550 555 560
gtc atg cag atg gct ttt ttc cat ctg acg caa att ctg cct ggc gat 1728
Val Met Gln Met Ala Phe Phe His Leu Thr Gln Ile Leu Pro Gly Asp
565 570 575
agc gcc ctc gca gaa ttg cag ggt gcg att gcc aaa agt tac agt agc 1776
Ser Ala Leu Ala Glu Leu Gln Gly Ala Ile Ala Lys Ser Tyr Ser Ser
580 585 590
aaa ggc cag gat ctg gtg gaa cgc aac tgg cag gct ctg gcg ctg gcg 1824
Lys Gly Gln Asp Leu Val Glu Arg Asn Trp Gln Ala Leu Ala Leu Ala
595 600 605
cgt gaa tcc gta gaa gaa gtt ccg ttg caa ccg gta aat ccg cac agc 1872
Arg Glu Ser Val Glu Glu Val Pro Leu Gln Pro Val Asn Pro His Ser
610 615 620
gcc aat cga ccg cca gtg gtt tcc gat gcc gcc cct gat ttc gtg aaa 1920
Ala Asn Arg Pro Pro Val Val Ser Asp Ala Ala Pro Asp Phe Val Lys
625 630 635 640
acc gta acc gct gcg atg ctc gcc ggg ctt ggt gac gcc ctc ccc gtt 1968
Thr Val Thr Ala Ala Met Leu Ala Gly Leu Gly Asp Ala Leu Pro Val
645 650 655
tcg gcg ctg ccg cca gac ggc acc tgg ccg atg ggc act acg cgc tgg 2016
Ser Ala Leu Pro Pro Asp Gly Thr Trp Pro Met Gly Thr Thr Arg Trp
660 665 670
gaa aaa cgc aat atc gcc gaa gag atc ccc atc tgg aaa gag gaa ctc 2064
Glu Lys Arg Asn Ile Ala Glu Glu Ile Pro Ile Trp Lys Glu Glu Leu

675	680	685	
tgt acc caa tgt aac cac tgc gtt gcc gct tgc cca cac tca gct att			2112
Cys Thr Gln Cys Asn His Cys Val Ala Ala Cys Pro His Ser Ala Ile			
690	695	700	
cgc gca aaa gtg gtg ccg cct gaa gcg atg gaa aac gcc cct gcc agc			2160
Arg Ala Lys Val Val Pro Pro Glu Ala Met Glu Asn Ala Pro Ala Ser			
705	710	715	720
ctg cat tcg ctg gat gtg aaa tcg cgt gat atg cgc ggg cag aaa tat			2208
Leu His Ser Leu Asp Val Lys Ser Arg Asp Met Arg Gly Gln Lys Tyr			
725	730	735	
gtc ttg cag gtg gca ccg gaa gat tgc acc ggt tgt aac ctg tgc gtc			2256
Val Leu Gln Val Ala Pro Glu Asp Cys Thr Gly Cys Asn Leu Cys Val			
740	745	750	
gaa gtt tgc ccg gcg aaa gac cgt cag aat cca gag att aaa gcc atc			2304
Glu Val Cys Pro Ala Lys Asp Arg Gln Asn Pro Glu Ile Lys Ala Ile			
755	760	765	
aat atg atg tct cgc ctg gaa cat gtc gaa gaa gag aaa atc aat tac			2352
Asn Met Met Ser Arg Leu Glu His Val Glu Glu Glu Lys Ile Asn Tyr			
770	775	780	
gat ttc ttc ctc aac ctg cca gaa atc gac cgt agc aaa ctg gaa cgt			2400
Asp Phe Phe Leu Asn Leu Pro Glu Ile Asp Arg Ser Lys Leu Glu Arg			
785	790	795	800
att gat att cgt aca tcg cag ctg att aca ccg ctg ttt gaa tat tca			2448
Ile Asp Ile Arg Thr Ser Gln Leu Ile Thr Pro Leu Phe Glu Tyr Ser			
805	810	815	
ggt gct tgc tcc ggt tgt ggc gag acg ccg tat att aaa tta ctg act			2496
Gly Ala Cys Ser Gly Cys Gly Glu Thr Pro Tyr Ile Lys Leu Leu Thr			
820	825	830	
cag ctc tat ggc gac cgg atg ttg atc gct aac gcc act ggc tgt tct			2544
Gln Leu Tyr Gly Asp Arg Met Leu Ile Ala Asn Ala Thr Gly Cys Ser			
835	840	845	

tca att tat ggc ggt aac ctg ccc tct aca ccg tat acc acc gat gcc	2592
Ser Ile Tyr Gly Gly Asn Leu Pro Ser Thr Pro Tyr Thr Thr Asp Ala	
850 855 860	
aac ggt cgt ggg ccg gca tgg gcg aac tct cta ttt gaa gat aat gcc	2640
Asn Gly Arg Gly Pro Ala Trp Ala Asn Ser Leu Phe Glu Asp Asn Ala	
865 870 875 880	
gaa ttt ggc ctt ggt ttc cgc ctg acg gtc gat caa cac cgt gtc cgc	2688
Glu Phe Gly Leu Gly Phe Arg Leu Thr Val Asp Gln His Arg Val Arg	
885 890 895	
gtg ctg cgt ctg ctg gat caa ttt gcc gat aaa atc ccg gcg gaa tta	2736
Val Leu Arg Leu Leu Asp Gln Phe Ala Asp Lys Ile Pro Ala Glu Leu	
900 905 910	
ctg acg gcg ttg aaa tca gac gcc acg cca gag gtt cgt cgt gaa cag	2784
Leu Thr Ala Leu Lys Ser Asp Ala Thr Pro Glu Val Arg Arg Glu Gln	
915 920 925	
gtt gca gct tta cgc cag caa ctc aac gat gtt gcc gaa gca cat gaa	2832
Val Ala Ala Leu Arg Gln Gln Leu Asn Asp Val Ala Glu Ala His Glu	
930 935 940	
ctg cta cgt gat gca gat gca ctg gtg gaa aaa tca atc tgg ctg att	2880
Leu Leu Arg Asp Ala Asp Ala Leu Val Glu Lys Ser Ile Trp Leu Ile	
945 950 955 960	
ggt ggt gat ggc tgg gct tac gat atc ggc ttt ggc ggt ctg gat cat	2928
Gly Gly Asp Gly Trp Ala Tyr Asp Ile Gly Phe Gly Gly Leu Asp His	
965 970 975	
gta ttg agt ttg acg gaa aac gtc aac att ctg gtg ctg gat acg caa	2976
Val Leu Ser Leu Thr Glu Asn Val Asn Ile Leu Val Leu Asp Thr Gln	
980 985 990	
tgc tat tcc aac acc ggt ggt cag gcg tcg aaa gcg aca ccg ctg ggt	3024
Cys Tyr Ser Asn Thr Gly Gly Gln Ala Ser Lys Ala Thr Pro Leu Gly	

995	1000	1005	
gca gta act aaa ttt ggc gag cac ggc aaa cgt aaa gcg cgt aaa			3069
Ala Val Thr Lys Phe Gly Glu His Gly Lys Arg Lys Ala Arg Lys			
1010	1015	1020	
gat ctt ggc gtc agt atg atg atg tac ggt cat gtt tat gtg gcg			3114
Asp Leu Gly Val Ser Met Met Met Tyr Gly His Val Tyr Val Ala			
1025	1030	1035	
cag att tct ctc ggc gcg cag ctg aac cag acg gtg aaa gcg att			3159
Gln Ile Ser Leu Gly Ala Gln Leu Asn Gln Thr Val Lys Ala Ile			
1040	1045	1050	
cag gaa gcg gaa gcg tat ccg ggg cca tcg ctg atc att gct tat			3204
Gln Glu Ala Glu Ala Tyr Pro Gly Pro Ser Leu Ile Ile Ala Tyr			
1055	1060	1065	
agc ccg tgt gaa gag cat ggt tac gat ctg gca ctc agc cac gac			3249
Ser Pro Cys Glu Glu His Gly Tyr Asp Leu Ala Leu Ser His Asp			
1070	1075	1080	
cag atg cgc caa ctc aca gct acc ggc ttc tgg ccg cta tat cgc			3294
Gln Met Arg Gln Leu Thr Ala Thr Gly Phe Trp Pro Leu Tyr Arg			
1085	1090	1095	
ttt gat ccg cgt cgt gcc gat gaa ggc aaa ctg ccg ctg gcc ttg			3339
Phe Asp Pro Arg Arg Ala Asp Glu Gly Lys Leu Pro Leu Ala Leu			
1100	1105	1110	
gat tca cgc ccg ccg tca gaa gca ccg gaa gaa acg tta ctt cac			3384
Asp Ser Arg Pro Pro Ser Glu Ala Pro Glu Glu Thr Leu Leu His			
1115	1120	1125	
gag caa cgt ttc cgt cgg ctg aat tcg cag cag cca gaa gtg gca			3429
Glu Gln Arg Phe Arg Arg Leu Asn Ser Gln Gln Pro Glu Val Ala			
1130	1135	1140	
gaa cag tta tgg aaa gat gct gca gct gat ttg caa aaa cgc tat			3474
Glu Gln Leu Trp Lys Asp Ala Ala Ala Asp Leu Gln Lys Arg Tyr			
1145	1150	1155	

gac ttc ctg gca caa atg gcc gga aaa gcg gaa aaa agc aac acc 3519
 Asp Phe Leu Ala Gln Met Ala Gly Lys Ala Glu Lys Ser Asn Thr
 1160 1165 1170
 gat taa 3525

Asp

<210> 253

<211> 1174

<212> PRT

<213> Escherichia coli

<400> 253

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

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

165 170 175

Asn Lys Ile Val Pro Leu Ala Asp Asp Thr Ile Leu Asp Leu Met Pro

180 185 190

Gln Val Glu Ile Asp Ala His Arg Ala Arg Ala Leu Asn Pro Glu His

195 200 205

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

210 215 220

Arg Glu Ala Thr Asn Pro Trp Tyr Asn Ala Val Tyr Asp His Val Glu

225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

Arg Gly Glu Lys Val Gly Val Leu Lys Val Arg Leu Tyr Arg Pro Phe

290 295 300

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

305 310 315 320

Ala Val Leu Asp Arg Thr Lys Glu Pro Gly Ala Gln Ala Glu Pro Leu

325 330 335

Tyr Leu Asp Val Met Thr Ala Leu Ala Glu Ala Phe Asn Asn Gly Glu

340 345 350

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

355 360 365

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

370 375 380

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

385 390 395 400

Thr Asn Leu Ser Leu Pro Leu Pro Glu Asn Thr Leu Pro Asn Ser Ala

405 410 415
Lys Leu Glu Ala Leu Phe Tyr Gly Leu Gly Ser Asp Gly Ser Val Ser

420 425 430
Ala Thr Lys Asn Asn Ile Lys Ile Ile Gly Asn Ser Thr Pro Trp Tyr

435 440 445
Ala Gln Gly Tyr Phe Val Tyr Asp Ser Lys Lys Ala Gly Gly Leu Thr

450 455 460
Val Ser His Leu Arg Val Ser Glu Gln Pro Ile Arg Ser Ala Tyr Leu

465 470 475 480
Ile Ser Gln Ala Asp Phe Val Gly Cys His Gln Leu Gln Phe Ile Asp

485 490 495
Lys Tyr Gln Met Ala Glu Arg Leu Lys Pro Gly Gly Ile Phe Leu Leu

500 505 510
Asn Thr Pro Tyr Ser Ala Asp Glu Val Trp Ser Arg Leu Pro Gln Glu

515 520 525
Val Gln Ala Val Leu Asn Gln Lys Lys Ala Arg Phe Tyr Val Ile Asn

530 535 540
Ala Ala Lys Ile Ala Arg Glu Cys Gly Leu Ala Ala Arg Ile Asn Thr

545 550 555 560
Val Met Gln Met Ala Phe Phe His Leu Thr Gln Ile Leu Pro Gly Asp

565 570 575
Ser Ala Leu Ala Glu Leu Gln Gly Ala Ile Ala Lys Ser Tyr Ser Ser

580 585 590
Lys Gly Gln Asp Leu Val Glu Arg Asn Trp Gln Ala Leu Ala Leu Ala

595 600 605
Arg Glu Ser Val Glu Glu Val Pro Leu Gln Pro Val Asn Pro His Ser

610 615 620
Ala Asn Arg Pro Pro Val Val Ser Asp Ala Ala Pro Asp Phe Val Lys

625 630 635 640
Thr Val Thr Ala Ala Met Leu Ala Gly Leu Gly Asp Ala Leu Pro Val

645 650 655

Ser Ala Leu Pro Pro Asp Gly Thr Trp Pro Met Gly Thr Thr Arg Trp
660 665 670

Glu Lys Arg Asn Ile Ala Glu Glu Ile Pro Ile Trp Lys Glu Glu Leu
675 680 685

Cys Thr Gln Cys Asn His Cys Val Ala Ala Cys Pro His Ser Ala Ile
690 695 700

Arg Ala Lys Val Val Pro Pro Glu Ala Met Glu Asn Ala Pro Ala Ser
705 710 715 720

Leu His Ser Leu Asp Val Lys Ser Arg Asp Met Arg Gly Gln Lys Tyr
725 730 735

Val Leu Gln Val Ala Pro Glu Asp Cys Thr Gly Cys Asn Leu Cys Val
740 745 750

Glu Val Cys Pro Ala Lys Asp Arg Gln Asn Pro Glu Ile Lys Ala Ile
755 760 765

Asn Met Met Ser Arg Leu Glu His Val Glu Glu Glu Lys Ile Asn Tyr
770 775 780

Asp Phe Phe Leu Asn Leu Pro Glu Ile Asp Arg Ser Lys Leu Glu Arg
785 790 795 800

Ile Asp Ile Arg Thr Ser Gln Leu Ile Thr Pro Leu Phe Glu Tyr Ser
805 810 815

Gly Ala Cys Ser Gly Cys Gly Glu Thr Pro Tyr Ile Lys Leu Leu Thr
820 825 830

Gln Leu Tyr Gly Asp Arg Met Leu Ile Ala Asn Ala Thr Gly Cys Ser
835 840 845

Ser Ile Tyr Gly Gly Asn Leu Pro Ser Thr Pro Tyr Thr Thr Asp Ala
850 855 860

Asn Gly Arg Gly Pro Ala Trp Ala Asn Ser Leu Phe Glu Asp Asn Ala
865 870 875 880

Glu Phe Gly Leu Gly Phe Arg Leu Thr Val Asp Gln His Arg Val Arg
885 890 895

Val Leu Arg Leu Leu Asp Gln Phe Ala Asp Lys Ile Pro Ala Glu Leu

900	905	910	
Leu Thr Ala Leu Lys Ser Asp Ala Thr Pro Glu Val Arg Arg Glu Gln			
915	920	925	
Val Ala Ala Leu Arg Gln Gln Leu Asn Asp Val Ala Glu Ala His Glu			
930	935	940	
Leu Leu Arg Asp Ala Asp Ala Leu Val Glu Lys Ser Ile Trp Leu Ile			
945	950	955	960
Gly Gly Asp Gly Trp Ala Tyr Asp Ile Gly Phe Gly Gly Leu Asp His			
965	970	975	
Val Leu Ser Leu Thr Glu Asn Val Asn Ile Leu Val Leu Asp Thr Gln			
980	985	990	
Cys Tyr Ser Asn Thr Gly Gly Gln Ala Ser Lys Ala Thr Pro Leu Gly			
995	1000	1005	
Ala Val Thr Lys Phe Gly Glu His Gly Lys Arg Lys Ala Arg Lys			
1010	1015	1020	
Asp Leu Gly Val Ser Met Met Met Tyr Gly His Val Tyr Val Ala			
1025	1030	1035	
Gln Ile Ser Leu Gly Ala Gln Leu Asn Gln Thr Val Lys Ala Ile			
1040	1045	1050	
Gln Glu Ala Glu Ala Tyr Pro Gly Pro Ser Leu Ile Ile Ala Tyr			
1055	1060	1065	
Ser Pro Cys Glu Glu His Gly Tyr Asp Leu Ala Leu Ser His Asp			
1070	1075	1080	
Gln Met Arg Gln Leu Thr Ala Thr Gly Phe Trp Pro Leu Tyr Arg			
1085	1090	1095	
Phe Asp Pro Arg Arg Ala Asp Glu Gly Lys Leu Pro Leu Ala Leu			
1100	1105	1110	
Asp Ser Arg Pro Pro Ser Glu Ala Pro Glu Glu Thr Leu Leu His			
1115	1120	1125	
Glu Gln Arg Phe Arg Arg Leu Asn Ser Gln Gln Pro Glu Val Ala			
1130	1135	1140	

Glu Gln Leu Trp Lys Asp Ala Ala Ala Asp Leu Gln Lys Arg Tyr
 1145 1150 1155
 Asp Phe Leu Ala Gln Met Ala Gly Lys Ala Glu Lys Ser Asn Thr
 1160 1165 1170
 Asp

<210> 254

<211> 3588

<212> DNA

<

213> Geobacter sulfurreducens

<220><221> CDS

<222> (1)..(3588)

<223> por gene from Geobacter sulfurreducens AM-1 encoding
 pyruvate-flavodoxin/ferredoxin oxidoreductase

<400> 254

atg agt cgc aaa atg gta acc atc gac ggc aat acc gcg gct gcc cac 48

Met Ser Arg Lys Met Val Thr Ile Asp Gly Asn Thr Ala Ala Ala His
 1 5 10 15

gtg gcg cac gcc acc aac gag gtc att gcc atc tac ccc att acc cct 96

Val Ala His Ala Thr Asn Glu Val Ile Ala Ile Tyr Pro Ile Thr Pro

20 25 30

tcg tcg gtc atg ggt gag att tcc gac atc aag agc gcc atg ggc gag 144

Ser Ser Val Met Gly Glu Ile Ser Asp Ile Lys Ser Ala Met Gly Glu
 35 40 45

aaa aac atc tgg gga acc gta ccg tcg gtt gtc gag atg cag tcg gaa 192

Lys Asn Ile Trp Gly Thr Val Pro Ser Val Val Glu Met Gln Ser Glu
 50 55 60

ggc ggc gct gcc ggt gcc gtg cac ggt gcc ctc cag gca ggt gcg ctg 240

Gly Gly Ala Ala Gly Ala Val His Gly Ala Leu Gln Ala Gly Ala Leu
 65 70 75 80

acc acc act ttt acc gcc agc cag ggt ctg ctc ctg atg atc ccg aac 288

Thr Thr Thr Phe Thr Ala Ser Gln Gly Leu Leu Leu Met Ile Pro Asn

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

Glu Ile Val Ala Glu Glu Met Asp Lys Phe Ala Lys Ile Thr Gly Arg

245	250	255	
cag tac aag ctg gtg gac tac gtg ggc gcc ccc gac gcc gac cgg gtc			816
Gln Tyr Lys Leu Val Asp Tyr Val Gly Ala Pro Asp Ala Asp Arg Val			
260	265	270	
atc gtc atc atg gga tcc ggc gcc gac acg gtg cag gag acc gtg gag			864
Ile Val Ile Met Gly Ser Gly Ala Asp Thr Val Gln Glu Thr Val Glu			
275	280	285	
cac ctg aac acc aag ggt gag aag atc ggc gtg gtg aag gtc cac ctc			912
His Leu Asn Thr Lys Gly Glu Lys Ile Gly Val Val Lys Val His Leu			
290	295	300	
tac cgg ccg ttc ccc atc gat gcc ttc att gcc gcc ctg ccc aag acc			960
Tyr Arg Pro Phe Pro Ile Asp Ala Phe Ile Ala Ala Leu Pro Lys Thr			
305	310	315	320
gtg aag aag atc gcg gtc ctc gac cgg acc aag gag ccc ggc gcc ctg			1008
Val Lys Lys Ile Ala Val Leu Asp Arg Thr Lys Glu Pro Gly Ala Leu			
325	330	335	
ggc gag ccc ctg tac ctg gat gtc cgc act gcc atc ggc gag gcc atg			1056
Gly Glu Pro Leu Tyr Leu Asp Val Arg Thr Ala Ile Gly Glu Ala Met			
340	345	350	
gcc gac ggg aag tgc cag ttc gac ggc tac ccg gtc atc gtg ggc ggt			1104
Ala Asp Gly Lys Cys Gln Phe Asp Gly Tyr Pro Val Ile Val Gly Gly			
355	360	365	
cgc tac ggc ctt ggt tcc aag gag ttc acc ccg gcc cag gcc aag gcg			1152
Arg Tyr Gly Leu Gly Ser Lys Glu Phe Thr Pro Ala Gln Ala Lys Ala			
370	375	380	
gtg ttc gat aac cta gcc act gcc aag ccg cag aac aag ttc gtg gtc			1200
Val Phe Asp Asn Leu Ala Thr Ala Lys Pro Gln Asn Lys Phe Val Val			
385	390	395	400

ggc atc acc gag gac gtg acc aac agc agc ctc ccg tgt gat ccg tcc	1248
Gly Ile Thr Glu Asp Val Thr Asn Ser Ser Leu Pro Cys Asp Pro Ser	
405 410 415	
ttc ttc aac ccg atg gaa ggg gcc tac cag gcc atg ttc ttc ggc ctc	1296
Phe Phe Asn Pro Met Glu Gly Ala Tyr Gln Ala Met Phe Phe Gly Leu	
420 425 430	
ggc tcc gac ggt acc gtg ggc gcc aac aag aac tcc atc aag atc atc	1344
Gly Ser Asp Gly Thr Val Gly Ala Asn Lys Asn Ser Ile Lys Ile Ile	
435 440 445	
ggc gag atg acc gat aac aac gcc cag gcc tac ttc gtc tac gac tcc	1392
Gly Glu Met Thr Asp Asn Asn Ala Gln Ala Tyr Phe Val Tyr Asp Ser	
450 455 460	
aag aag gcc ggc tcc atg acc acc tcg cac ctg cgc ttc ggc aag aag	1440
Lys Lys Ala Gly Ser Met Thr Thr Ser His Leu Arg Phe Gly Lys Lys	
465 470 475 480	
tac atc aga gcg ccg tac ctg gtg cag gag gcc gac ttc gtg gcc tgc	1488
Tyr Ile Arg Ala Pro Tyr Leu Val Gln Glu Ala Asp Phe Val Ala Cys	
485 490 495	
cac aac ttc gcc ttc gtg gaa aag tac gac atg ctg gcc aag gcc aag	1536
His Asn Phe Ala Phe Val Glu Lys Tyr Asp Met Leu Ala Lys Ala Lys	
500 505 510	
cag ggt gcc acg ttc ctc ctg aac gcc cct tac gac cac aac gag gtg	1584
Gln Gly Ala Thr Phe Leu Leu Asn Ala Pro Tyr Asp His Asn Glu Val	
515 520 525	
tgg gac agg ctc ccc gcc gac atg cag cag cag atc atc gac aag aag	1632
Trp Asp Arg Leu Pro Ala Asp Met Gln Gln Gln Ile Ile Asp Lys Lys	
530 535 540	
ctc aag ttc ttc gtg atc gat ggg gta cgc ctc ggc aac gag atc ggg	1680
Leu Lys Phe Phe Val Ile Asp Gly Val Arg Leu Gly Asn Glu Ile Gly	
545 550 555 560	
ctc ggt ccc cgg atc aac gtg atc atg cag acc gcc ttc ttc aag ata	1728

Leu Gly Pro Arg Ile Asn Val Ile Met Gln Thr Ala Phe Phe Lys Ile	
565 570 575	
tcc aac atc atc ccg ctg gat cag gcc att gac gag atc aag gac gct	1776
Ser Asn Ile Ile Pro Leu Asp Gln Ala Ile Asp Glu Ile Lys Asp Ala	
580 585 590	
atc aag aaa acc tat ggc aag gca ggc gag aag gtc gtg gag atg aac	1824
Ile Lys Lys Thr Tyr Gly Lys Ala Gly Glu Lys Val Val Glu Met Asn	
595 600 605	
tac aag gcg gtt gag gcc ggc ctc aac aac ttc tac gag gta acg gta	1872
Tyr Lys Ala Val Glu Ala Gly Leu Asn Asn Phe Tyr Glu Val Thr Val	
610 615 620	
ccg gca acg gca acc agt acc ctc cag aag cct ccc gtg gtc agc gcc	1920
Pro Ala Thr Ala Thr Ser Thr Leu Gln Lys Pro Pro Val Val Ser Ala	
625 630 635 640	
agg gcc ccc cag ttc gtg cag gag acc acc gcc ccc atc atc gcc ggc	1968
Arg Ala Pro Gln Phe Val Gln Glu Thr Thr Ala Pro Ile Ile Ala Gly	
645 650 655	
ctc ggc gac gac ctg ccg gtg tcc aag atg ccg gcc gac ggc acc ttc	2016
Leu Gly Asp Asp Leu Pro Val Ser Lys Met Pro Ala Asp Gly Thr Phe	
660 665 670	
ccg acg gcg acc tcc cag ttc gaa aag cgg aac atc gcc gtg gag atc	2064
Pro Thr Ala Thr Ser Gln Phe Glu Lys Arg Asn Ile Ala Val Glu Ile	
675 680 685	
ccc gtg tgg gac gag cag ctc tgc atc cag tgc ggc atc tgc tcc ttc	2112
Pro Val Trp Asp Glu Gln Leu Cys Ile Gln Cys Gly Ile Cys Ser Phe	
690 695 700	
gtc tgc ccc cac gcc acc atc agg atg aag gcc tat gac gcc tcc gcc	2160
Val Cys Pro His Ala Thr Ile Arg Met Lys Ala Tyr Asp Ala Ser Ala	
705 710 715 720	
ctt gcc ggc gcc ccg gca gcg ttc aag tgc gtt gac tgc aag att ccc	2208
Leu Ala Gly Ala Pro Ala Ala Phe Lys Ser Val Asp Cys Lys Ile Pro	

725	730	735	
gag ttc aag ggg cag aag ttc acc atc cag gta gcc ccg gaa gac tgc			2256
Glu Phe Lys Gly Gln Lys Phe Thr Ile Gln Val Ala Pro Glu Asp Cys			
740	745	750	
acc ggc tgc ggc gcc tgc gtc cac aac tgc ccg gcc aag agt aag gaa			2304
Thr Gly Cys Gly Ala Cys Val His Asn Cys Pro Ala Lys Ser Lys Glu			
755	760	765	
gac ccg aac cac aag gcc atc aat atg gca tac cag ccg ccc ctg cgt			2352
Asp Pro Asn His Lys Ala Ile Asn Met Ala Tyr Gln Pro Pro Leu Arg			
770	775	780	
tct caa gag gtc gag aac tgg gac ttc ttc ctc acc atc ccg gac gtg			2400
Ser Gln Glu Val Glu Asn Trp Asp Phe Phe Leu Thr Ile Pro Asp Val			
785	790	795	800
gac ccc acc gta gcc aag ctg gac acg gtc cgc ggt tcc cag ttg gtg			2448
Asp Pro Thr Val Ala Lys Leu Asp Thr Val Arg Gly Ser Gln Leu Val			
805	810	815	
cgg ccg ctg ttc gaa ttc tcc ggc gcc tgc ctc ggc tgc ggc gag acc			2496
Arg Pro Leu Phe Glu Phe Ser Gly Ala Cys Leu Gly Cys Gly Glu Thr			
820	825	830	
ccg tac ctg aag ctc ctg acc cag ctc ttc ggc gac cgg acc gtc att			2544
Pro Tyr Leu Lys Leu Leu Thr Gln Leu Phe Gly Asp Arg Thr Val Ile			
835	840	845	
gcc aac gcc acc ggc tgc tcc tcc atc tac ggc gga aac ctg ccc acc			2592
Ala Asn Ala Thr Gly Cys Ser Ser Ile Tyr Gly Gly Asn Leu Pro Thr			
850	855	860	
acc ccc tat gcc cag cgg gcc gac ggg ctt ggg ccg gca tgg tcg aac			2640
Thr Pro Tyr Ala Gln Arg Ala Asp Gly Leu Gly Pro Ala Trp Ser Asn			
865	870	875	880
tcc ctg ttc gag gac aac gcc gag ttc ggc tac ggc atg cgt ctg gcc			2688

Ser Leu Phe Glu Asp Asn Ala Glu Phe Gly Tyr Gly Met Arg Leu Ala	
885 890 895	
gtg gat aaa ttc aac gcc atg gcc ctt gag ctg gtc gac aag ctt tcg	2736
Val Asp Lys Phe Asn Ala Met Ala Leu Glu Leu Val Asp Lys Leu Ser	
900 905 910	
tct tcc tgc tcc tgc tct tcc tgc acg agc gcg gtg ccc ctc atg aac	2784
Ser Ser Cys Ser Cys Ser Ser Cys Thr Ser Ala Val Pro Leu Met Asn	
915 920 925	
gag atc aag ggc gcc gac cag tcg agc cag gcc ggc atc gag gcc cag	2832
Glu Ile Lys Gly Ala Asp Gln Ser Ser Gln Ala Gly Ile Glu Ala Gln	
930 935 940	
cgg gcc cgg gtg gcg gag ctg aag aag acc ctt gct tcc tgt ccc gag	2880
Arg Ala Arg Val Ala Glu Leu Lys Lys Thr Leu Ala Ser Cys Pro Glu	
945 950 955 960	
ccg gat gcc aag cgc ctg ctc acc gtg gcc gac tac ctg gtc aag aag	2928
Pro Asp Ala Lys Arg Leu Leu Thr Val Ala Asp Tyr Leu Val Lys Lys	
965 970 975	
tcc gtc tgg tgc atc ggc ggc gac ggc tgg gcg tac gat atc ggc tac	2976
Ser Val Trp Cys Ile Gly Gly Asp Gly Trp Ala Tyr Asp Ile Gly Tyr	
980 985 990	
ggc ggc ctc gac cac gtc atc gcc agc ggc aag aac atc aat ctg ctg	3024
Gly Gly Leu Asp His Val Ile Ala Ser Gly Lys Asn Ile Asn Leu Leu	
995 1000 1005	
gtg ctc gac acc gag gtc tac tcc aac acc ggc ggc cag gct tcc	3069
Val Leu Asp Thr Glu Val Tyr Ser Asn Thr Gly Gly Gln Ala Ser	
1010 1015 1020	
aag tcg acc ccg ctg ggc gcc gtg gcc cag ttc gcc gcc ggc ggt	3114
Lys Ser Thr Pro Leu Gly Ala Val Ala Gln Phe Ala Ala Gly Gly	
1025 1030 1035	
aag ccg gtc tcc aag aag gat ctc ggc atg atg gcc atg gcc tac	3159
Lys Pro Val Ser Lys Lys Asp Leu Gly Met Met Ala Met Ala Tyr	

1040	1045	1050	
ggg tcg gtc tac gtg gcc act	gtc tcc ctc gcc aat	ccg gcc cag	3204
Gly Ser Val Tyr Val Ala Thr	Val Ser Leu Ala Asn	Pro Ala Gln	
1055	1060	1065	
tgc atc aag gcg ttc ctg gag	gcc gag gcc tat gac	ggt ccg tcg	3249
Cys Ile Lys Ala Phe Leu Glu	Ala Glu Ala Tyr Asp	Gly Pro Ser	
1070	1075	1080	
ctc atc atc gcc tat gcc cac	tgc atc gcc cac ggc	atc gac atg	3294
Leu Ile Ile Ala Tyr Ala His	Cys Ile Ala His Gly	Ile Asp Met	
1085	1090	1095	
acc agc ggc gtg gat gcc cag	aag cgg gcg gtt cag	tcc ggt tac	3339
Thr Ser Gly Val Asp Ala Gln	Lys Arg Ala Val Gln	Ser Gly Tyr	
1100	1105	1110	
tgg ccc ctc tac cgc tat aat	ccg cag ctg gcc gcc	gag tgc aag	3384
Trp Pro Leu Tyr Arg Tyr Asn	Pro Gln Leu Ala Ala	Glu Cys Lys	
1115	1120	1125	
aac ccg ctc cag ctc gac agc	aag gcc ccg acc atc	gcc ttt gaa	3429
Asn Pro Leu Gln Leu Asp Ser	Lys Ala Pro Thr Ile	Ala Phe Glu	
1130	1135	1140	
gag tac gtc aac agc gag aac	cgc tac cgc gtc ctc	aag aag aac	3474
Glu Tyr Val Asn Ser Glu Asn	Arg Tyr Arg Val Leu	Lys Lys Asn	
1145	1150	1155	
aac ccg aaa ggg tac gag gat	ctc atg aga aaa gcg	gcg gca tgg	3519
Asn Pro Lys Gly Tyr Glu Asp	Leu Met Arg Lys Ala	Ala Ala Trp	
1160	1165	1170	
tcc aag gcc cac ttc agc tac	tac cag aag ctg gcg	gcc ctc aac	3564
Ser Lys Ala His Phe Ser Tyr	Tyr Gln Lys Leu Ala	Ala Leu Asn	
1175	1180	1185	
ttc gag gat acc tgc gag aag	tag		3588
Phe Glu Asp Thr Cys Glu Lys			
1190	1195		

<210> 255

<211> 1195

<212> PRT

<213> *Geobacter sulfurreducens*

<400> 255

Met Ser Arg Lys Met Val Thr Ile Asp Gly Asn Thr Ala Ala Ala His

1 5 10 15

Val Ala His Ala Thr Asn Glu Val Ile Ala Ile Tyr Pro Ile Thr Pro

20 25 30

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

35 40 45

Lys Asn Ile Trp Gly Thr Val Pro Ser Val Val Glu Met Gln Ser Glu

50 55 60

Gly Gly Ala Ala Gly Ala Val His Gly Ala Leu Gln Ala Gly Ala Leu

65 70 75 80

Thr Thr Thr Phe Thr Ala Ser Gln Gly Leu Leu Leu Met Ile Pro Asn

85 90 95

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

100 105 110

Ala Arg Ala Ile Ala Ala Gln Ala Leu Ser Ile Phe Gly Asp His Ser

115 120 125

Asp Val Met Ser Cys Arg Ser Thr Gly Trp Ala Met Leu Cys Ser Asn

130 135 140

Asn Ser Gln Glu Val Met Asp Phe Ala Leu Ile Ala Gln Ser Ala Thr

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

Ser Pro Asp His Pro Val Met Arg Gly Thr Ala Gln Asn Pro Asp Val
210 215 220

Tyr Phe Gln Gly Arg Glu Thr Val Asn Pro Phe Tyr Pro Lys Cys Ile
225 230 235 240

Glu Ile Val Ala Glu Glu Met Asp Lys Phe Ala Lys Ile Thr Gly Arg
245 250 255

Gln Tyr Lys Leu Val Asp Tyr Val Gly Ala Pro Asp Ala Asp Arg Val
260 265 270

Ile Val Ile Met Gly Ser Gly Ala Asp Thr Val Gln Glu Thr Val Glu
275 280 285

His Leu Asn Thr Lys Gly Glu Lys Ile Gly Val Val Lys Val His Leu
290 295 300

Tyr Arg Pro Phe Pro Ile Asp Ala Phe Ile Ala Ala Leu Pro Lys Thr
305 310 315 320

Val Lys Lys Ile Ala Val Leu Asp Arg Thr Lys Glu Pro Gly Ala Leu
325 330 335

Gly Glu Pro Leu Tyr Leu Asp Val Arg Thr Ala Ile Gly Glu Ala Met
340 345 350

Ala Asp Gly Lys Cys Gln Phe Asp Gly Tyr Pro Val Ile Val Gly Gly
355 360 365

Arg Tyr Gly Leu Gly Ser Lys Glu Phe Thr Pro Ala Gln Ala Lys Ala
370 375 380

Val Phe Asp Asn Leu Ala Thr Ala Lys Pro Gln Asn Lys Phe Val Val
385 390 395 400

Gly Ile Thr Glu Asp Val Thr Asn Ser Ser Leu Pro Cys Asp Pro Ser
405 410 415

Phe Phe Asn Pro Met Glu Gly Ala Tyr Gln Ala Met Phe Phe Gly Leu
420 425 430

Gly Ser Asp Gly Thr Val Gly Ala Asn Lys Asn Ser Ile Lys Ile Ile
435 440 445

Gly Glu Met Thr Asp Asn Asn Ala Gln Ala Tyr Phe Val Tyr Asp Ser

450 455 460
 Lys Lys Ala Gly Ser Met Thr Thr Ser His Leu Arg Phe Gly Lys Lys
 465 470 475 480
 Tyr Ile Arg Ala Pro Tyr Leu Val Gln Glu Ala Asp Phe Val Ala Cys
 485 490 495
 His Asn Phe Ala Phe Val Glu Lys Tyr Asp Met Leu Ala Lys Ala Lys
 500 505 510

 Gln Gly Ala Thr Phe Leu Leu Asn Ala Pro Tyr Asp His Asn Glu Val
 515 520 525
 Trp Asp Arg Leu Pro Ala Asp Met Gln Gln Gln Ile Ile Asp Lys Lys
 530 535 540
 Leu Lys Phe Phe Val Ile Asp Gly Val Arg Leu Gly Asn Glu Ile Gly
 545 550 555 560
 Leu Gly Pro Arg Ile Asn Val Ile Met Gln Thr Ala Phe Phe Lys Ile
 565 570 575

 Ser Asn Ile Ile Pro Leu Asp Gln Ala Ile Asp Glu Ile Lys Asp Ala
 580 585 590
 Ile Lys Lys Thr Tyr Gly Lys Ala Gly Glu Lys Val Val Glu Met Asn
 595 600 605
 Tyr Lys Ala Val Glu Ala Gly Leu Asn Asn Phe Tyr Glu Val Thr Val
 610 615 620
 Pro Ala Thr Ala Thr Ser Thr Leu Gln Lys Pro Pro Val Val Ser Ala
 625 630 635 640

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

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

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

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

Val Asp Lys Phe Asn Ala Met Ala Leu Glu Leu Val Asp Lys Leu Ser
900 905 910
Ser Ser Cys Ser Cys Ser Ser Cys Thr Ser Ala Val Pro Leu Met Asn
915 920 925
Glu Ile Lys Gly Ala Asp Gln Ser Ser Gln Ala Gly Ile Glu Ala Gln
930 935 940
Arg Ala Arg Val Ala Glu Leu Lys Lys Thr Leu Ala Ser Cys Pro Glu

945	950	955	960
Pro Asp Ala Lys Arg Leu Leu Thr Val Ala Asp Tyr Leu Val Lys Lys			
	965	970	975
Ser Val Trp Cys Ile Gly Gly Asp Gly Trp Ala Tyr Asp Ile Gly Tyr			
	980	985	990
Gly Gly Leu Asp His Val Ile Ala Ser Gly Lys Asn Ile Asn Leu Leu			
	995	1000	1005
Val Leu Asp Thr Glu Val Tyr Ser Asn Thr Gly Gly Gln Ala Ser			
1010	1015	1020	
Lys Ser Thr Pro Leu Gly Ala Val Ala Gln Phe Ala Ala Gly Gly			
1025	1030	1035	
Lys Pro Val Ser Lys Lys Asp Leu Gly Met Met Ala Met Ala Tyr			
1040	1045	1050	
Gly Ser Val Tyr Val Ala Thr Val Ser Leu Ala Asn Pro Ala Gln			
1055	1060	1065	
Cys Ile Lys Ala Phe Leu Glu Ala Glu Ala Tyr Asp Gly Pro Ser			
1070	1075	1080	
Leu Ile Ile Ala Tyr Ala His Cys Ile Ala His Gly Ile Asp Met			
1085	1090	1095	
Thr Ser Gly Val Asp Ala Gln Lys Arg Ala Val Gln Ser Gly Tyr			
1100	1105	1110	
Trp Pro Leu Tyr Arg Tyr Asn Pro Gln Leu Ala Ala Glu Cys Lys			
1115	1120	1125	
Asn Pro Leu Gln Leu Asp Ser Lys Ala Pro Thr Ile Ala Phe Glu			
1130	1135	1140	
Glu Tyr Val Asn Ser Glu Asn Arg Tyr Arg Val Leu Lys Lys Asn			
1145	1150	1155	
Asn Pro Lys Gly Tyr Glu Asp Leu Met Arg Lys Ala Ala Ala Trp			
1160	1165	1170	
Ser Lys Ala His Phe Ser Tyr Tyr Gln Lys Leu Ala Ala Leu Asn			
1175	1180	1185	

Phe Glu Asp Thr Cys Glu Lys

1190 1195

<210> 256

<211> 1956

<212> DNA

<213> Streptomyces pratensis

<220><221> CDS

<222> (1)..(1956)

<223> Sfla_2592 gene from Streptomyces pratensis ATCC 33331 encoding
pyruvate-flavodoxin/ferredoxin oxidoreductase

<400> 256

atg acc agc cag gtc agt agc cca gcc gga aag tcc gat gag gcc agc 48

Met Thr Ser Gln Val Ser Ser Pro Ala Gly Lys Ser Asp Glu Ala Ser

1 5 10 15

gag gct gtc gtc ggg gaa cag cgc gcc ccg cac atc gcc ggt gcg ggt 96

Glu Ala Val Val Gly Glu Gln Arg Ala Pro His Ile Ala Gly Ala Gly

20 25 30

ggc acg gag aag gaa atc cgc cgt ctg gac cgg gtg atc atc cgt ttc 144

Gly Thr Glu Lys Glu Ile Arg Arg Leu Asp Arg Val Ile Ile Arg Phe

35 40 45

gcg ggt gac tcg ggt gac ggt atg cag ttg acg ggc gac cgt ttc acg 192

Ala Gly Asp Ser Gly Asp Gly Met Gln Leu Thr Gly Asp Arg Phe Thr

50 55 60

tcg gag acg gcg tcg ttc ggg aac gac ctg tcg aca ctg ccc aac ttc 240

Ser Glu Thr Ala Ser Phe Gly Asn Asp Leu Ser Thr Leu Pro Asn Phe

65 70 75 80

ccg gcc gag atc cgg gca ccc gcc ggc acc ctg ccc ggg gtg tcg tcg 288

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

85 90 95

ttc cag ctg cat ttc gcg gac cac gac atc ctg aca ccg ggc gac gcg 336

Phe Gln Leu His Phe Ala Asp His Asp Ile Leu Thr Pro Gly Asp Ala

100	105	110	
ccg aac gtc ctg gtc gcg atg aat ccc gcc gcg ctg aag gcg aat atc			384
Pro Asn Val Leu Val Ala Met Asn Pro Ala Ala Leu Lys Ala Asn Ile			
115	120	125	
gcc gat gtg ccg cgc ggg gcc gac atc atc gtg aac acg gac gag ttc			432
Ala Asp Val Pro Arg Gly Ala Asp Ile Ile Val Asn Thr Asp Glu Phe			
130	135	140	
acg aag cgc ccg atg gcg aaa gtc gga tat gcg gaa tcc cct ttg gag			480
Thr Lys Arg Pro Met Ala Lys Val Gly Tyr Ala Glu Ser Pro Leu Glu			
145	150	155	160
gac ggt tcc ctc gag gcg tac aac gtg cat ccg gtg ccg ttg acg acg			528
Asp Gly Ser Leu Glu Ala Tyr Asn Val His Pro Val Pro Leu Thr Thr			
165	170	175	
ttg acg atc gag gct ttg aag gag ttc ggg ctt tcc cgc aag gag gcc			576
Leu Thr Ile Glu Ala Leu Lys Glu Phe Gly Leu Ser Arg Lys Glu Ala			
180	185	190	
gag cgg tcg aag aac atg ttc gcg ctc ggg ctt ctg tcc tgg atg tac			624
Glu Arg Ser Lys Asn Met Phe Ala Leu Gly Leu Leu Ser Trp Met Tyr			
195	200	205	
aac cgt ccg acc gag ggt acg gag aag ttc ctg cgg tcg aag ttc gcc			672
Asn Arg Pro Thr Glu Gly Thr Glu Lys Phe Leu Arg Ser Lys Phe Ala			
210	215	220	
agg aag ccg gag atc gcc gag gcc aat gtg gcg gct ttc cgc gcg ggc			720
Arg Lys Pro Glu Ile Ala Glu Ala Asn Val Ala Ala Phe Arg Ala Gly			
225	230	235	240
tgg aat ttc ggt gag acg acg gag gat ttc gct gtc tcc tac gag gtc			768
Trp Asn Phe Gly Glu Thr Thr Glu Asp Phe Ala Val Ser Tyr Glu Val			
245	250	255	
gca ccg gcg tca cag gat ttc ccg acg ggc acc tac cgc aat atc tcc			816
Ala Pro Ala Ser Gln Asp Phe Pro Thr Gly Thr Tyr Arg Asn Ile Ser			
260	265	270	

ggg aat ctc gca ctg tcg tac ggg ctg atc gcg gcg gga cgg cag gcc	864
Gly Asn Leu Ala Leu Ser Tyr Gly Leu Ile Ala Ala Gly Arg Gln Ala	
275 280 285	
gat ctg ccg gtg tat ctc ggc tcg tat ccg atc act ccg gcg tcc gac	912
Asp Leu Pro Val Tyr Leu Gly Ser Tyr Pro Ile Thr Pro Ala Ser Asp	
290 295 300	
atc ctg cac gag ctc agc aag cac aag aac ttc ggt gtg cgg acc ttc	960
Ile Leu His Glu Leu Ser Lys His Lys Asn Phe Gly Val Arg Thr Phe	
305 310 315 320	
cag gcg gag gac gag atc gcc ggg atc ggt gcg gcc ctg ggc gcg tcg	1008
Gln Ala Glu Asp Glu Ile Ala Gly Ile Gly Ala Ala Leu Gly Ala Ser	
325 330 335	
ttc ggc ggt tca ctg ggt gtg acg acg acg tcg ggc ccg ggt gtg gcg	1056
Phe Gly Gly Ser Leu Gly Val Thr Thr Thr Ser Gly Pro Gly Val Ala	
340 345 350	
ctg aag tcg gag acg atc ggc ctg gcg gtg tca ctg gaa ctg ccg ctg	1104
Leu Lys Ser Glu Thr Ile Gly Leu Ala Val Ser Leu Glu Leu Pro Leu	
355 360 365	
ctg atc atc gac atc cag cgc ggc ggc ccc tcc acc ggc ctg ccg acc	1152
Leu Ile Ile Asp Ile Gln Arg Gly Gly Pro Ser Thr Gly Leu Pro Thr	
370 375 380	
aag acc gag cag gcc gac ctg ctc cag gcc atg tac ggg cgc aac ggc	1200
Lys Thr Glu Gln Ala Asp Leu Leu Gln Ala Met Tyr Gly Arg Asn Gly	
385 390 395 400	
gag gcc ccg gtc ccg atc gtg gca ccg agg act ccg gcg gac tgc ttc	1248
Glu Ala Pro Val Pro Ile Val Ala Pro Arg Thr Pro Ala Asp Cys Phe	
405 410 415	
gac gcc gcc ctg gac gcg gcg cgg atc gcg ctg acc tac cgc acc ccg	1296
Asp Ala Ala Leu Asp Ala Ala Arg Ile Ala Leu Thr Tyr Arg Thr Pro	
420 425 430	

gtc ttc ctg ctc tcg gac ggg tac ctc gcg aac ggc tcc gag ccg tgg 1344

Val Phe Leu Leu Ser Asp Gly Tyr Leu Ala Asn Gly Ser Glu Pro Trp
435 440 445

cgg atc ccc gag gcc gac agc ctc ccc gac ctg cgg aca cgg ttc gcg 1392

Arg Ile Pro Glu Ala Asp Ser Leu Pro Asp Leu Arg Thr Arg Phe Ala
450 455 460

acc ggc ccg aat cac gaa ctc gcg gac ggc acc gag gtg ttc tgg ccc 1440

Thr Gly Pro Asn His Glu Leu Ala Asp Gly Thr Glu Val Phe Trp Pro

465 470 475 480

tac aag agg gac ccc gag acg ctg gcc cgc ccg tgg gcg gtg ccc ggc 1488

Tyr Lys Arg Asp Pro Glu Thr Leu Ala Arg Pro Trp Ala Val Pro Gly
485 490 495

acc ccg ggt ctg gag cac cgg atc ggc ggg atc gag aag cag gac ggc 1536

Thr Pro Gly Leu Glu His Arg Ile Gly Gly Ile Glu Lys Gln Asp Gly
500 505 510

acg ggg aac atc tcc tac gat ccg gcc aac cac gac ttc atg gtc cgc 1584

Thr Gly Asn Ile Ser Tyr Asp Pro Ala Asn His Asp Phe Met Val Arg
515 520 525

acc cgc cag gcc aag atc gac ggc atc cgg gtc ccc gac ctg gag gtc 1632

Thr Arg Gln Ala Lys Ile Asp Gly Ile Arg Val Pro Asp Leu Glu Val
530 535 540

gac gac ccg gcc ggc gcg acg acc ctg gtc ctg ggc tgg ggt tcg acg 1680

Asp Asp Pro Ala Gly Ala Thr Thr Leu Val Leu Gly Trp Gly Ser Thr
545 550 555 560

tac ggg ccg atc acc gcc gcc gtg cgc cgt ctc cgc gcg gcc ggc gag 1728

Tyr Gly Pro Ile Thr Ala Ala Val Arg Arg Leu Arg Ala Ala Gly Glu
565 570 575

acg atc gca cag gca cat ctg cgc cac ctc aat ccc ttc ccc gcc aat 1776

Thr Ile Ala Gln Ala His Leu Arg His Leu Asn Pro Phe Pro Ala Asn

580	585	590	
ctc ggt gag gta ctg cgg cgc tac gac aag gtc gtc gtc ccc gag atg			1824
Leu Gly Glu Val Leu Arg Arg Tyr Asp Lys Val Val Val Pro Glu Met			
595	600	605	
aac ctc ggt cag ctc gcc ctg ctg ctg aga gcc aag tac ctc gtg gac			1872
Asn Leu Gly Gln Leu Ala Leu Leu Leu Arg Ala Lys Tyr Leu Val Asp			
610	615	620	
gcg cag agt ttc aac cag gtc aac gga atg ccc ttc aag gcg gag cag			1920
Ala Gln Ser Phe Asn Gln Val Asn Gly Met Pro Phe Lys Ala Glu Gln			
625	630	635	640
ctc gcc aca gcc ctc aag gag gcc atc gat gcc tga			1956
Leu Ala Thr Ala Leu Lys Glu Ala Ile Asp Ala			
645	650		
<210> 257			
<211> 651			
<212> PRT			
<213> Streptomyces pratensis			
<400> 257			
Met Thr Ser Gln Val Ser Ser Pro Ala Gly Lys Ser Asp Glu Ala Ser			
1	5	10	15
Glu Ala Val Val Gly Glu Gln Arg Ala Pro His Ile Ala Gly Ala Gly			
20	25	30	
Gly Thr Glu Lys Glu Ile Arg Arg Leu Asp Arg Val Ile Ile Arg Phe			
35	40	45	
Ala Gly Asp Ser Gly Asp Gly Met Gln Leu Thr Gly Asp Arg Phe Thr			
50	55	60	
Ser Glu Thr Ala Ser Phe Gly Asn Asp Leu Ser Thr Leu Pro Asn Phe			
65	70	75	80
Pro Ala Glu Ile Arg Ala Pro Ala Gly Thr Leu Pro Gly Val Ser Ser			
85	90	95	
Phe Gln Leu His Phe Ala Asp His Asp Ile Leu Thr Pro Gly Asp Ala			
100	105	110	

Pro Asn Val Leu Val Ala Met Asn Pro Ala Ala Leu Lys Ala Asn Ile
115 120 125

Ala Asp Val Pro Arg Gly Ala Asp Ile Ile Val Asn Thr Asp Glu Phe
130 135 140

Thr Lys Arg Pro Met Ala Lys Val Gly Tyr Ala Glu Ser Pro Leu Glu
145 150 155 160

Asp Gly Ser Leu Glu Ala Tyr Asn Val His Pro Val Pro Leu Thr Thr
165 170 175

Leu Thr Ile Glu Ala Leu Lys Glu Phe Gly Leu Ser Arg Lys Glu Ala
180 185 190

Glu Arg Ser Lys Asn Met Phe Ala Leu Gly Leu Leu Ser Trp Met Tyr
195 200 205

Asn Arg Pro Thr Glu Gly Thr Glu Lys Phe Leu Arg Ser Lys Phe Ala
210 215 220

Arg Lys Pro Glu Ile Ala Glu Ala Asn Val Ala Ala Phe Arg Ala Gly
225 230 235 240

Trp Asn Phe Gly Glu Thr Thr Glu Asp Phe Ala Val Ser Tyr Glu Val
245 250 255

Ala Pro Ala Ser Gln Asp Phe Pro Thr Gly Thr Tyr Arg Asn Ile Ser
260 265 270

Gly Asn Leu Ala Leu Ser Tyr Gly Leu Ile Ala Ala Gly Arg Gln Ala
275 280 285

Asp Leu Pro Val Tyr Leu Gly Ser Tyr Pro Ile Thr Pro Ala Ser Asp
290 295 300

Ile Leu His Glu Leu Ser Lys His Lys Asn Phe Gly Val Arg Thr Phe
305 310 315 320

Gln Ala Glu Asp Glu Ile Ala Gly Ile Gly Ala Ala Leu Gly Ala Ser
325 330 335

Phe Gly Gly Ser Leu Gly Val Thr Thr Thr Ser Gly Pro Gly Val Ala
340 345 350

Leu Lys Ser Glu Thr Ile Gly Leu Ala Val Ser Leu Glu Leu Pro Leu

355 360 365
 Leu Ile Ile Asp Ile Gln Arg Gly Gly Pro Ser Thr Gly Leu Pro Thr
 370 375 380

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

 Arg Ile Pro Glu Ala Asp Ser Leu Pro Asp Leu Arg Thr Arg Phe Ala
 450 455 460
 Thr Gly Pro Asn His Glu Leu Ala Asp Gly Thr Glu Val Phe Trp Pro
 465 470 475 480
 Tyr Lys Arg Asp Pro Glu Thr Leu Ala Arg Pro Trp Ala Val Pro Gly
 485 490 495
 Thr Pro Gly Leu Glu His Arg Ile Gly Gly Ile Glu Lys Gln Asp Gly
 500 505 510

 Thr Gly Asn Ile Ser Tyr Asp Pro Ala Asn His Asp Phe Met Val Arg
 515 520 525
 Thr Arg Gln Ala Lys Ile Asp Gly Ile Arg Val Pro Asp Leu Glu Val
 530 535 540
 Asp Asp Pro Ala Gly Ala Thr Thr Leu Val Leu Gly Trp Gly Ser Thr
 545 550 555 560
 Tyr Gly Pro Ile Thr Ala Ala Val Arg Arg Leu Arg Ala Ala Gly Glu
 565 570 575

 Thr Ile Ala Gln Ala His Leu Arg His Leu Asn Pro Phe Pro Ala Asn
 580 585 590
 Leu Gly Glu Val Leu Arg Arg Tyr Asp Lys Val Val Val Pro Glu Met
 595 600 605

Asn Leu Gly Gln Leu Ala Leu Leu Arg Ala Lys Tyr Leu Val Asp
610 615 620
Ala Gln Ser Phe Asn Gln Val Asn Gly Met Pro Phe Lys Ala Glu Gln
625 630 635 640
Leu Ala Thr Ala Leu Lys Glu Ala Ile Asp Ala
645 650
<210> 258
<211> 3768
<212> DNA
<213> Propionibacterium freudenreichii
<220><221> CDS
<222> (1)..(3768)
<223> RM25_0186 gene from Propionibacterium freudenreichii DSM
20271 encoding pyruvate-flavodoxin/ferredoxin oxidoreductase
<400> 258
atg act aca act acc cgt ggg ccg gtt ccc ggc tcg aat ggc atg ccc 48
Met Thr Thr Thr Thr Arg Gly Pro Val Pro Gly Ser Asn Gly Met Pro
1 5 10 15
gcc aat cca ggt ctg agc ggc gag gcc gcc acc gca acc ccg tca ccc 96
Ala Asn Pro Gly Leu Ser Gly Glu Ala Ala Thr Ala Thr Pro Ser Pro
20 25 30
gtt gac gtc gct gcc ggc gcc aag gac gct gcc gat gag ctg gcc cag 144
Val Asp Val Ala Ala Gly Ala Lys Asp Ala Ala Asp Glu Leu Ala Gln
35 40 45
tca cga cgc gag cag gac atc acc cat cag atg atc tgc gac ggc aac 192
Ser Arg Arg Glu Gln Asp Ile Thr His Gln Met Ile Cys Asp Gly Asn
50 55 60
acc gcc gcc tct gat gtg gcc ttc cgc atc aat gag ctg tgc tcg atc 240
Thr Ala Ala Ser Asp Val Ala Phe Arg Ile Asn Glu Leu Cys Ser Ile
65 70 75 80
tac ccg atc acg ccg agc tcc ccg atg gcc gaa ctg gcc gac gag tgg 288

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

245	250	255	
cgc gag cgg gcc ctg tgc ccc gac aac ccg ttc atc cgt ggc acc gcc			816
Arg Glu Arg Ala Leu Ser Pro Asp Asn Pro Phe Ile Arg Gly Thr Ala			
260	265	270	
cag aac gcc gac gtg tac ttc cag ggc cgc gag gcc ggc aac aag tac			864
Gln Asn Ala Asp Val Tyr Phe Gln Gly Arg Glu Ala Gly Asn Lys Tyr			
275	280	285	
tac gac tgc gtt ccg ggc atc gtg cag gac gcg atg gac gag ttc gcc			912
Tyr Asp Ser Val Pro Gly Ile Val Gln Asp Ala Met Asp Glu Phe Ala			
290	295	300	
gcc atg acc ggc cgc cag tac cac ctg gcc gac tac tac ggc gcg ccc			960
Ala Met Thr Gly Arg Gln Tyr His Leu Ala Asp Tyr Tyr Gly Ala Pro			
305	310	315	320
gac gcc gat cgc gtc atc gtg atc atg ggc tgc ggt gcc gag acc gtg			1008
Asp Ala Asp Arg Val Ile Val Ile Met Gly Ser Gly Ala Glu Thr Val			
325	330	335	
cag cag acc gtc agc aag ctc aat gag cag ggc gag aag gtc ggc ctg			1056
Gln Gln Thr Val Ser Lys Leu Asn Glu Gln Gly Glu Lys Val Gly Leu			
340	345	350	
gtg gtc atc cgc ctg tac cgt ccg ttc ccg acg cag gcc gtg ctg gac			1104
Val Val Ile Arg Leu Tyr Arg Pro Phe Pro Thr Gln Ala Val Leu Asp			
355	360	365	
tgc att ccc gca tgc gtc aag aag atc gcc gtg ctc gac cgc acc aag			1152
Cys Ile Pro Ala Ser Val Lys Lys Ile Ala Val Leu Asp Arg Thr Lys			
370	375	380	
gag ccg ggc tcc aac ggt gag ccc ctg ttc ctc gac gtg gtc tgc gca			1200
Glu Pro Gly Ser Asn Gly Glu Pro Leu Phe Leu Asp Val Val Ser Ala			
385	390	395	400
gtc tcc gag gcc tat tgc aac ggc gag cgc gac aac ctg ccc gcc atc			1248
Val Ser Glu Ala Tyr Ser Asn Gly Glu Arg Asp Asn Leu Pro Ala Ile			
405	410	415	

atc ggt ggc cgc tac ggc ctg tcg agc aag gag ttc acg ccg ggc atg 1296

Ile Gly Gly Arg Tyr Gly Leu Ser Ser Lys Glu Phe Thr Pro Gly Met
420 425 430

tgc gcc gcc gtg tac gac gag ctc gcc aag gac aag ccg aag cgt cgc 1344
Cys Ala Ala Val Tyr Asp Glu Leu Ala Lys Asp Lys Pro Lys Arg Arg
435 440 445

ttc acc gtc ggc atc acc gac gat gtg acg cac ctg tcg atc ccg tgg 1392
Phe Thr Val Gly Ile Thr Asp Asp Val Thr His Leu Ser Ile Pro Trp

450 455 460

gac gcc tcg ctc gac ctg gag gac ccc gag acc tcg cgc gca gtg ttc 1440
Asp Ala Ser Leu Asp Leu Glu Asp Pro Glu Thr Ser Arg Ala Val Phe
465 470 475 480

tac ggc atc ggt gct gac ggc acc gtc ggc gcc aac aag aac acc atc 1488
Tyr Gly Ile Gly Ala Asp Gly Thr Val Gly Ala Asn Lys Asn Thr Ile
485 490 495

aag atc ctc ggc tcc gag ccg ggc acc tac gcg cag ggc tac ttc gtc 1536
Lys Ile Leu Gly Ser Glu Pro Gly Thr Tyr Ala Gln Gly Tyr Phe Val
500 505 510

tac gac tcg aag aag tcc ggc ggc cgc acc acc tcg cac ctt cgc ttc 1584
Tyr Asp Ser Lys Lys Ser Gly Gly Arg Thr Thr Ser His Leu Arg Phe
515 520 525

gga ccc gat ccg atc aag gcc ccc tac ctg gtg aac cag gcc ggc ttc 1632

Gly Pro Asp Pro Ile Lys Ala Pro Tyr Leu Val Asn Gln Ala Gly Phe
530 535 540

atc ggc gtg cac cac tgg gcc gac ctt gag cgc atc gac gtg ctg gcg 1680
Ile Gly Val His His Trp Ala Asp Leu Glu Arg Ile Asp Val Leu Ala
545 550 555 560

ttc gcc cgc aag ggc acc acg gtg ctg atc aac agc ccg tac ccc gcc 1728
Phe Ala Arg Lys Gly Thr Thr Val Leu Ile Asn Ser Pro Tyr Pro Ala

565	570	575	
gag gac gtc tgg ggc cat ctg ccg gcc ccg atg cag aag aag atc atc			1776
Glu Asp Val Trp Gly His Leu Pro Ala Pro Met Gln Lys Lys Ile Ile			
580	585	590	
gac ctc gac ctg cag gtg tat gcg atc gac gcc ggt gag gtg gcc cgt			1824
Asp Leu Asp Leu Gln Val Tyr Ala Ile Asp Ala Gly Glu Val Ala Arg			
595	600	605	
tcg gtg ggc ctg ggc aac cgc acc aac acg gtg ctg cag acc tgc tac			1872
Ser Val Gly Leu Gly Asn Arg Thr Asn Thr Val Leu Gln Thr Cys Tyr			
610	615	620	
ttc aag atc agt ggc gtg ctt ccc gag gac cac gcg atc gag gcc atc			1920
Phe Lys Ile Ser Gly Val Leu Pro Glu Asp His Ala Ile Glu Ala Ile			
625	630	635	640
aag aac tcg atc acc aag acc tac gcg aag aag tcg atg gag atc gtg			1968
Lys Asn Ser Ile Thr Lys Thr Tyr Ala Lys Lys Ser Met Glu Ile Val			
645	650	655	
gag aag aac cac gcc gcc gtc gac gcc gcc ctg gag cac ctg cac aag			2016
Glu Lys Asn His Ala Ala Val Asp Ala Ala Leu Glu His Leu His Lys			
660	665	670	
atc gac gtg ccg gcc aag gtc acc tcc acc gag gac tac ctg ccg ccc			2064
Ile Asp Val Pro Ala Lys Val Thr Ser Thr Glu Asp Tyr Leu Pro Pro			
675	680	685	
gtg ccg tcg ttc gcg cct gac ttc gtc aag gac gtc acc gcg gcc atg			2112
Val Pro Ser Phe Ala Pro Asp Phe Val Lys Asp Val Thr Ala Ala Met			
690	695	700	
atg acc gag cag ggc gag tcg ctg ccg gtg agc aag ctg ccg gcc gat			2160
Met Thr Glu Gln Gly Glu Ser Leu Pro Val Ser Lys Leu Pro Ala Asp			
705	710	715	720
ggt tcg ttc ccc tcg ggc acc acg cag tac gag aag cgc aat gtg tcc			2208
Gly Ser Phe Pro Ser Gly Thr Thr Gln Tyr Glu Lys Arg Asn Val Ser			
725	730	735	

gag atc atc gcg gtc tgg gac cag gac aac tgc atc cag tgc ggc aac	2256
Glu Ile Ile Ala Val Trp Asp Gln Asp Asn Cys Ile Gln Cys Gly Asn	
740 745 750	
tgc gcc ttc gtc tgc ccg cac ggc gtg ctg agg gcc aag tac tac aag	2304
Cys Ala Phe Val Cys Pro His Gly Val Leu Arg Ala Lys Tyr Tyr Lys	
755 760 765	
ccc gat gtg ctc gac gat gcg ccg aag tgc ttc cag gcg gtt ccg ctg	2352
Pro Asp Val Leu Asp Asp Ala Pro Lys Ser Phe Gln Ala Val Pro Leu	
770 775 780	
aat gcg gcc ggc ctg ccc gac gag atg tac acc ctg cag gtg ttc gcc	2400
Asn Ala Ala Gly Leu Pro Asp Glu Met Tyr Thr Leu Gln Val Phe Ala	
785 790 795 800	
gag gac tgc acc ggt tgt ggc ctg tgc gtc gag gcc tgc ccc gtg cat	2448
Glu Asp Cys Thr Gly Cys Gly Leu Cys Val Glu Ala Cys Pro Val His	
805 810 815	
ccc atc ggt ggc gac ccc gaa tgc aag gcg atc aac ctg gat tcc gtg	2496
Pro Ile Gly Gly Asp Pro Glu Cys Lys Ala Ile Asn Leu Asp Ser Val	
820 825 830	
ctc gac cgc acc aac gag cgg gcg aac gtg gag ttc ttc cag aag atc	2544
Leu Asp Arg Thr Asn Glu Arg Ala Asn Val Glu Phe Phe Gln Lys Ile	
835 840 845	
ccc gag ccc ccg cgc acc cgc gtg aac tac ggt gcc gtg cgt ggc gcc	2592
Pro Glu Pro Pro Arg Thr Arg Val Asn Tyr Gly Ala Val Arg Gly Ala	
850 855 860	
cag ttc ctg cag ccg ctg ttc gag ttc agc ggt gcc tgc ccg ggt tgt	2640
Gln Phe Leu Gln Pro Leu Phe Glu Phe Ser Gly Ala Cys Pro Gly Cys	
865 870 875 880	
ggc gag acg ccg tac ctc aag ctg ctc acc cag ctg ttc ggc gac cgc	2688
Gly Glu Thr Pro Tyr Leu Lys Leu Leu Thr Gln Leu Phe Gly Asp Arg	
885 890 895	
gcc acc gtg gcg aat gcc acc ggc tgc tgc tcc atc tac ggc ggc aac	2736

Ala Thr Val Ala Asn Ala Thr Gly Cys Ser Ser Ile Tyr Gly Gly Asn

900 905 910
ctg ccg acc acc ccg tgg gcg aag aac aag gag gga cgc ggc ccg gcc 2784

Leu Pro Thr Thr Pro Trp Ala Lys Asn Lys Glu Gly Arg Gly Pro Ala

915 920 925
tgg agc aac tca ttg ttc gag gac aac gcc gag ttc ggc ctt ggc atg 2832

Trp Ser Asn Ser Leu Phe Glu Asp Asn Ala Glu Phe Gly Leu Gly Met

930 935 940

cgc ctg gcg gcc gac ctg cac aac gaa ctg gcc cgt cag cgc gtt gac 2880

Arg Leu Ala Ala Asp Leu His Asn Glu Leu Ala Arg Gln Arg Val Asp

945 950 955 960

gag ctg tcc gat gcg atc aac gac ccc gag ctg gtc gat cag ctg ctg 2928

Glu Leu Ser Asp Ala Ile Asn Asp Pro Glu Leu Val Asp Gln Leu Leu

965 970 975
aac gcc ccg cag gcg cag gag tcc gat ctg cac gcc cag gcc gag cgc 2976

Asn Ala Pro Gln Ala Gln Glu Ser Asp Leu His Ala Gln Ala Glu Arg

980 985 990

gtc gac gcc ctg cag gat cgc ctg acc gac ctg gtc aac gat ccg aac 3024

Val Asp Ala Leu Gln Asp Arg Leu Thr Asp Leu Val Asn Asp Pro Asn

995 1000 1005

gtg gac gcc gac acc aag gcc aag gtc gag gac ctg cgg tcg gtg 3069

Val Asp Ala Asp Thr Lys Ala Lys Val Glu Asp Leu Arg Ser Val

1010 1015 1020

gcc gac aac ctg ctg cgt cgt tcc gtg tgg atc gtc ggc ggc gac 3114

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

1025 1030 1035

ggt tgg gcc tac gac atc ggt tcg ggc ggc ctt gac cat gtg ctg 3159

Gly Trp Ala Tyr Asp Ile Gly Ser Gly Gly Leu Asp His Val Leu

1040 1045 1050

tcc acc gga cgc aat gtc aat gtg ctg gtg ctc gac acc gag gtc	3204
Ser Thr Gly Arg Asn Val Asn Val Leu Val Leu Asp Thr Glu Val	
1055 1060 1065	
tac tcc aat acc ggt ggc cag gcc tcg aag tcg tcg ccc atg ggt	3249
Tyr Ser Asn Thr Gly Gly Gln Ala Ser Lys Ser Ser Pro Met Gly	
1070 1075 1080	
gcg atc gcg aag ttc gcg acc gcc ggc aag cgc acg aac aag aag	3294
Ala Ile Ala Lys Phe Ala Thr Ala Gly Lys Arg Thr Asn Lys Lys	
1085 1090 1095	
gac atc gcc atg cag gcc gtg tcc tac ggc gac gtc tat gtc gcc	3339
Asp Ile Ala Met Gln Ala Val Ser Tyr Gly Asp Val Tyr Val Ala	
1100 1105 1110	
cgc gtg gcg ttc ggt gcc gac ccg gag cag acg ctg aag gca ttc	3384
Arg Val Ala Phe Gly Ala Asp Pro Glu Gln Thr Leu Lys Ala Phe	
1115 1120 1125	
cgt gag gcc gag gcc tac ccc ggc ccc agc ctg atc atc gcc tac	3429
Arg Glu Ala Glu Ala Tyr Pro Gly Pro Ser Leu Ile Ile Ala Tyr	
1130 1135 1140	
agc cac tgc atc agc cat ggc tac aac ctg cgc aag ggc ctg gac	3474
Ser His Cys Ile Ser His Gly Tyr Asn Leu Arg Lys Gly Leu Asp	
1145 1150 1155	
cag cag tac aag gca gtg gcc tcc ggt cac tgg ccg ctg atc cgg	3519
Gln Gln Tyr Lys Ala Val Ala Ser Gly His Trp Pro Leu Ile Arg	
1160 1165 1170	
tac aac ccg gag gtt cgc gac tcg ggt ggc aac ccg ttc ctg ctc	3564
Tyr Asn Pro Glu Val Arg Asp Ser Gly Gly Asn Pro Phe Leu Leu	
1175 1180 1185	
gac tcg gcc cgt ccg cgc atc tcg ctg atg gac tac cgc aag acc	3609
Asp Ser Ala Arg Pro Arg Ile Ser Leu Met Asp Tyr Arg Lys Thr	
1190 1195 1200	
gag ctg cgc ttc aag atg ctg atg gtc aag gat ccg gaa gag gcc	3654

Glu Leu Arg Phe Lys Met Leu Met Val Lys Asp Pro Glu Glu Ala
1205 1210 1215
aag cac ctc aat gac ctc agc cag gag cag gtg acc agg cgt ttc 3699
Lys His Leu Asn Asp Leu Ser Gln Glu Gln Val Thr Arg Arg Phe
1220 1225 1230
gcc gac tac gag gaa atg gcc tca cgt ccg gcc gag atg ttc gcc 3744
Ala Asp Tyr Glu Glu Met Ala Ser Arg Pro Ala Glu Met Phe Ala
1235 1240 1245
acc gac gca cgg agg gat gtc tga 3768
Thr Asp Ala Arg Arg Asp Val
1250 1255

<210> 259
<211> 1255
<212> PRT
<213> Propionibacterium freudenreichii
<400> 259

Met Thr Thr Thr Thr Arg Gly Pro Val Pro Gly Ser Asn Gly Met Pro
1 5 10 15
Ala Asn Pro Gly Leu Ser Gly Glu Ala Ala Thr Ala Thr Pro Ser Pro
20 25 30
Val Asp Val Ala Ala Gly Ala Lys Asp Ala Ala Asp Glu Leu Ala Gln
35 40 45
Ser Arg Arg Glu Gln Asp Ile Thr His Gln Met Ile Cys Asp Gly Asn
50 55 60
Thr Ala Ala Ser Asp Val Ala Phe Arg Ile Asn Glu Leu Cys Ser Ile
65 70 75 80
Tyr Pro Ile Thr Pro Ser Ser Pro Met Ala Glu Leu Ala Asp Glu Trp
85 90 95
Ser Ala Arg Asp Arg Met Asn Ile Trp Gly Gln Val Pro His Val Met
100 105 110
Glu Met Gln Ser Glu Ala Gly Ala Ala Gly Ala Met His Gly Ser Leu

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

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

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

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

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

370 375 380

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

385 390 395 400

Val Ser Glu Ala Tyr Ser Asn Gly Glu Arg Asp Asn Leu Pro Ala Ile

405 410 415

Ile Gly Gly Arg Tyr Gly Leu Ser Ser Lys Glu Phe Thr Pro Gly Met

420 425 430

Cys Ala Ala Val Tyr Asp Glu Leu Ala Lys Asp Lys Pro Lys Arg Arg

435 440 445

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

450 455 460

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

465 470 475 480

Tyr Gly Ile Gly Ala Asp Gly Thr Val Gly Ala Asn Lys Asn Thr Ile

485 490 495

Lys Ile Leu Gly Ser Glu Pro Gly Thr Tyr Ala Gln Gly Tyr Phe Val

500 505 510

Tyr Asp Ser Lys Lys Ser Gly Gly Arg Thr Thr Ser His Leu Arg Phe

515 520 525

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

530 535 540

Ile Gly Val His His Trp Ala Asp Leu Glu Arg Ile Asp Val Leu Ala

545 550 555 560

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

565 570 575

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

580 585 590

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

595 600 605

Ser Val Gly Leu Gly Asn Arg Thr Asn Thr Val Leu Gln Thr Cys Tyr

610 615 620
 Phe Lys Ile Ser Gly Val Leu Pro Glu Asp His Ala Ile Glu Ala Ile

 625 630 635 640
 Lys Asn Ser Ile Thr Lys Thr Tyr Ala Lys Lys Ser Met Glu Ile Val

 645 650 655
 Glu Lys Asn His Ala Ala Val Asp Ala Ala Leu Glu His Leu His Lys

 660 665 670
 Ile Asp Val Pro Ala Lys Val Thr Ser Thr Glu Asp Tyr Leu Pro Pro

 675 680 685
 Val Pro Ser Phe Ala Pro Asp Phe Val Lys Asp Val Thr Ala Ala Met

 690 695 700
 Met Thr Glu Gln Gly Glu Ser Leu Pro Val Ser Lys Leu Pro Ala Asp

 705 710 715 720
 Gly Ser Phe Pro Ser Gly Thr Thr Gln Tyr Glu Lys Arg Asn Val Ser

 725 730 735
 Glu Ile Ile Ala Val Trp Asp Gln Asp Asn Cys Ile Gln Cys Gly Asn

 740 745 750
 Cys Ala Phe Val Cys Pro His Gly Val Leu Arg Ala Lys Tyr Tyr Lys

 755 760 765
 Pro Asp Val Leu Asp Asp Ala Pro Lys Ser Phe Gln Ala Val Pro Leu

 770 775 780
 Asn Ala Ala Gly Leu Pro Asp Glu Met Tyr Thr Leu Gln Val Phe Ala

 785 790 795 800
 Glu Asp Cys Thr Gly Cys Gly Leu Cys Val Glu Ala Cys Pro Val His

 805 810 815
 Pro Ile Gly Gly Asp Pro Glu Cys Lys Ala Ile Asn Leu Asp Ser Val

 820 825 830
 Leu Asp Arg Thr Asn Glu Arg Ala Asn Val Glu Phe Phe Gln Lys Ile

 835 840 845
 Pro Glu Pro Pro Arg Thr Arg Val Asn Tyr Gly Ala Val Arg Gly Ala

 850 855 860

Gln Phe Leu Gln Pro Leu Phe Glu Phe Ser Gly Ala Cys Pro Gly Cys
865 870 875 880
Gly Glu Thr Pro Tyr Leu Lys Leu Leu Thr Gln Leu Phe Gly Asp Arg
885 890 895
Ala Thr Val Ala Asn Ala Thr Gly Cys Ser Ser Ile Tyr Gly Gly Asn
900 905 910
Leu Pro Thr Thr Pro Trp Ala Lys Asn Lys Glu Gly Arg Gly Pro Ala
915 920 925
Trp Ser Asn Ser Leu Phe Glu Asp Asn Ala Glu Phe Gly Leu Gly Met
930 935 940
Arg Leu Ala Ala Asp Leu His Asn Glu Leu Ala Arg Gln Arg Val Asp
945 950 955 960
Glu Leu Ser Asp Ala Ile Asn Asp Pro Glu Leu Val Asp Gln Leu Leu
965 970 975
Asn Ala Pro Gln Ala Gln Glu Ser Asp Leu His Ala Gln Ala Glu Arg
980 985 990
Val Asp Ala Leu Gln Asp Arg Leu Thr Asp Leu Val Asn Asp Pro Asn
995 1000 1005
Val Asp Ala Asp Thr Lys Ala Lys Val Glu Asp Leu Arg Ser Val
1010 1015 1020
Ala Asp Asn Leu Leu Arg Arg Ser Val Trp Ile Val Gly Gly Asp
1025 1030 1035
Gly Trp Ala Tyr Asp Ile Gly Ser Gly Gly Leu Asp His Val Leu
1040 1045 1050
Ser Thr Gly Arg Asn Val Asn Val Leu Val Leu Asp Thr Glu Val
1055 1060 1065
Tyr Ser Asn Thr Gly Gly Gln Ala Ser Lys Ser Ser Pro Met Gly
1070 1075 1080
Ala Ile Ala Lys Phe Ala Thr Ala Gly Lys Arg Thr Asn Lys Lys
1085 1090 1095
Asp Ile Ala Met Gln Ala Val Ser Tyr Gly Asp Val Tyr Val Ala

1100 1105 1110
 Arg Val Ala Phe Gly Ala Asp Pro Glu Gln Thr Leu Lys Ala Phe
 1115 1120 1125
 Arg Glu Ala Glu Ala Tyr Pro Gly Pro Ser Leu Ile Ile Ala Tyr

 1130 1135 1140
 Ser His Cys Ile Ser His Gly Tyr Asn Leu Arg Lys Gly Leu Asp
 1145 1150 1155
 Gln Gln Tyr Lys Ala Val Ala Ser Gly His Trp Pro Leu Ile Arg
 1160 1165 1170
 Tyr Asn Pro Glu Val Arg Asp Ser Gly Gly Asn Pro Phe Leu Leu
 1175 1180 1185
 Asp Ser Ala Arg Pro Arg Ile Ser Leu Met Asp Tyr Arg Lys Thr

1190 1195 1200
 Glu Leu Arg Phe Lys Met Leu Met Val Lys Asp Pro Glu Glu Ala
 1205 1210 1215
 Lys His Leu Asn Asp Leu Ser Gln Glu Gln Val Thr Arg Arg Phe
 1220 1225 1230
 Ala Asp Tyr Glu Glu Met Ala Ser Arg Pro Ala Glu Met Phe Ala
 1235 1240 1245
 Thr Asp Ala Arg Arg Asp Val
 1250 1255

<210> 260
 <211> 3600
 <212> DNA
 <213> Synechocystis PCC6803
 <220><221> CDS
 <222> (1)..(3600)
 <223> nifJ gene from Synechocystis sp. PCC 6803 encoding
 pyruvate-flavodoxin/ferredoxin oxidoreductase
 <400> 260

atg agt tta cct acc tat gcc acc ctc gac ggt aat gaa gcg gtg gcc
 Met Ser Leu Pro Thr Tyr Ala Thr Leu Asp Gly Asn Glu Ala Val Ala

48

1	5	10	15	
cgt gtg gcc tac ctg ctc agt gaa gtg att gcc att tat ccc atc acc				96
Arg Val Ala Tyr Leu Leu Ser Glu Val Ile Ala Ile Tyr Pro Ile Thr				
20	25	30		
cct tcc tcg ccc atg ggg gaa tgg tcc gat gct tgg gca gca gaa cac				144
Pro Ser Ser Pro Met Gly Glu Trp Ser Asp Ala Trp Ala Ala Glu His				
35	40	45		
cgg ccc aat ttg tgg ggc acc gta cca ttg gtg gtg gaa atg caa agc				192
Arg Pro Asn Leu Trp Gly Thr Val Pro Leu Val Val Glu Met Gln Ser				
50	55	60		
gag ggg gga gcc gcc ggt act gtc cat ggc gct ctg caa tcg gga gct				240
Glu Gly Gly Ala Ala Gly Thr Val His Gly Ala Leu Gln Ser Gly Ala				
65	70	75	80	
ttg acc aca aca ttt acc gct tcc cag ggc tta atg ttg atg ttg ccc				288
Leu Thr Thr Thr Phe Thr Ala Ser Gln Gly Leu Met Leu Met Leu Pro				
85	90	95		
aat atg cac aaa att gct ggg gaa tta aca gcc atg gtt ttg cat gtg				336
Asn Met His Lys Ile Ala Gly Glu Leu Thr Ala Met Val Leu His Val				
100	105	110		
gcg gcc cgt tct tta gcg gcc cag ggc cta tct att ttt ggg gat cac				384
Ala Ala Arg Ser Leu Ala Ala Gln Gly Leu Ser Ile Phe Gly Asp His				
115	120	125		
agt gat gtg atg gcg gcc aga aat acg ggc ttt gcc atg tta agt tcc				432
Ser Asp Val Met Ala Ala Arg Asn Thr Gly Phe Ala Met Leu Ser Ser				
130	135	140		
aat tct gtc cag gaa gcc cac gat ttt gcc ctc att gcc acg gcc acc				480
Asn Ser Val Gln Glu Ala His Asp Phe Ala Leu Ile Ala Thr Ala Thr				
145	150	155	160	
agc ttt gcc acc agg ata ccg gga ctg cac ttt ttt gat ggt ttt cgc				528

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

165 170 175

act tcc cac gaa gaa caa aaa att gag ctt tta ccc cag gaa gta ctc 576

Thr Ser His Glu Glu Gln Lys Ile Glu Leu Leu Pro Gln Glu Val Leu

180 185 190

cgt ggt ttg att aag gat gag gat gtg cta gcc cac cgg gga cgg gct 624

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

195 200 205

ttg acc ccc gat cgc ccg aag ttg cgg ggg acg gcc caa aat ccg gat 672

Leu Thr Pro Asp Arg Pro Lys Leu Arg Gly Thr Ala Gln Asn Pro Asp

210 215 220

gtc tat ttc caa gct agg gaa acg gtt aat ccc ttt tat gcc agt tat 720

Val Tyr Phe Gln Ala Arg Glu Thr Val Asn Pro Phe Tyr Ala Ser Tyr

225 230 235 240

ccc aac gtg ctg gag cag gtg atg gaa caa ttt ggc cag cta acc ggc 768

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

245 250 255

cgc cat tac cgt ccc tat gaa tat tgt ggc cat ccg gaa gcg gaa cgg 816

Arg His Tyr Arg Pro Tyr Glu Tyr Cys Gly His Pro Glu Ala Glu Arg

260 265 270

gtg att gtg ctg atg ggt tct ggt gcg gaa acg gcc cag gaa acg gtg 864

Val Ile Val Leu Met Gly Ser Gly Ala Glu Thr Ala Gln Glu Thr Val

275 280 285

gat ttt cta act gcc caa ggg gaa aag gtt ggt tta ctg aaa gta cgc 912

Asp Phe Leu Thr Ala Gln Gly Glu Lys Val Gly Leu Leu Lys Val Arg

290 295 300

ctc tat cgg ccc ttt gct ggc gat cgc ctg gtt aat gct cta cca aaa 960

Leu Tyr Arg Pro Phe Ala Gly Asp Arg Leu Val Asn Ala Leu Pro Lys

305 310 315 320

acg gtg caa aaa ata gcg gtg ctg gac cgg tgt aag gaa ccg ggg agc	1008
Thr Val Gln Lys Ile Ala Val Leu Asp Arg Cys Lys Glu Pro Gly Ser	
325 330 335	
att ggg gaa ccc ctc tat cag gat gtg ctg acg gcc ttt ttt gaa gcg	1056
Ile Gly Glu Pro Leu Tyr Gln Asp Val Leu Thr Ala Phe Phe Glu Ala	
340 345 350	
ggc atg atg ccg aaa att att ggt ggc cgt tac ggt ctg tca tcc aag	1104
Gly Met Met Pro Lys Ile Ile Gly Gly Arg Tyr Gly Leu Ser Ser Lys	
355 360 365	
gaa ttt acc ccc gcc atg gtt aaa ggg gtg ttg gac cat tta aat caa	1152
Glu Phe Thr Pro Ala Met Val Lys Gly Val Leu Asp His Leu Asn Gln	
370 375 380	
acc aac ccc aaa aac cat ttc acc gta ggc att aac gat gat ttg agc	1200
Thr Asn Pro Lys Asn His Phe Thr Val Gly Ile Asn Asp Asp Leu Ser	
385 390 395 400	
cac acc agc atc gac tat gac ccc agt ttt tcc acg gaa gca gat tct	1248
His Thr Ser Ile Asp Tyr Asp Pro Ser Phe Ser Thr Glu Ala Asp Ser	
405 410 415	
gtc gtc cgg gca att ttc tac ggt ctc ggt tcc gac ggt acg gtg ggg	1296
Val Val Arg Ala Ile Phe Tyr Gly Leu Gly Ser Asp Gly Thr Val Gly	
420 425 430	
gcc aat aag aac tcc atc aaa atc att ggc gaa gat acg gat aac tac	1344
Ala Asn Lys Asn Ser Ile Lys Ile Ile Gly Glu Asp Thr Asp Asn Tyr	
435 440 445	
gcc cag ggt tat ttt gtt tac gac tcg aaa aaa tcc ggt tct gta acc	1392
Ala Gln Gly Tyr Phe Val Tyr Asp Ser Lys Lys Ser Gly Ser Val Thr	
450 455 460	
gtt tcc cat ctg cgc ttt ggc cct aat ccc atc ctg tcc act tac ctg	1440
Val Ser His Leu Arg Phe Gly Pro Asn Pro Ile Leu Ser Thr Tyr Leu	
465 470 475 480	
att agc caa gcc aat ttt gtc gcc tgt cac cag tgg gaa ttt ttg gaa	1488

Ile Ser Gln Ala Asn Phe Val Ala Cys His Gln Trp Glu Phe Leu Glu	
485 490 495	
cag ttt gaa gtc ttg gaa cca gcc gtt gat ggc ggc gtt ttc ctg gtc	1536
Gln Phe Glu Val Leu Glu Pro Ala Val Asp Gly Gly Val Phe Leu Val	
500 505 510	
aat agc ccc tac ggc cca gag gaa att tgg cga gag ttt ccc cgc aaa	1584
Asn Ser Pro Tyr Gly Pro Glu Glu Ile Trp Arg Glu Phe Pro Arg Lys	
515 520 525	
gta caa cag gaa att att gac aaa aat ctc aag gtt tac acc atc aat	1632
Val Gln Gln Glu Ile Ile Asp Lys Asn Leu Lys Val Tyr Thr Ile Asn	
530 535 540	
gcc aat gac gta gcc agg gat gcg ggc atg ggc cgc cgc acc aac aca	1680
Ala Asn Asp Val Ala Arg Asp Ala Gly Met Gly Arg Arg Thr Asn Thr	
545 550 555 560	
gtc atg caa acc tgt ttc ttt gcc cta gcg gga gtg tta ccc cgg gaa	1728
Val Met Gln Thr Cys Phe Phe Ala Leu Ala Gly Val Leu Pro Arg Glu	
565 570 575	
gag gcg atc gcc aaa att aag cag tcg gtc caa aaa acc tac ggc aaa	1776
Glu Ala Ile Ala Lys Ile Lys Gln Ser Val Gln Lys Thr Tyr Gly Lys	
580 585 590	
aag ggt cag gaa att gtc gag atg aat att aaa gcg gtg gat tcc acc	1824
Lys Gly Gln Glu Ile Val Glu Met Asn Ile Lys Ala Val Asp Ser Thr	
595 600 605	
ctg gcc cat ctc tat gaa gtg tcc gta ccg gaa acg gtg agc gac gat	1872
Leu Ala His Leu Tyr Glu Val Ser Val Pro Glu Thr Val Ser Asp Asp	
610 615 620	
gcc cct gct atg cgg ccg gtg gtg cct gat aac gcc ccg gtg ttt gtg	1920
Ala Pro Ala Met Arg Pro Val Val Pro Asp Asn Ala Pro Val Phe Val	
625 630 635 640	
cgg gaa gtg tta gga aaa atc atg gcc cgg caa ggg gat gat ctc ccg	1968
Arg Glu Val Leu Gly Lys Ile Met Ala Arg Gln Gly Asp Asp Leu Pro	

645	650	655	
gtc agt gct tta ccc tgc gat ggc acc tat ccc acc gcc act acc caa			2016
Val Ser Ala Leu Pro Cys Asp Gly Thr Tyr Pro Thr Ala Thr Thr Gln			
660	665	670	
tgg gaa aaa cgc aac gtg ggc cac gaa att ccc gtt tgg gac ccc gat			2064
Trp Glu Lys Arg Asn Val Gly His Glu Ile Pro Val Trp Asp Pro Asp			
675	680	685	
gtt tgt gtg caa tgc ggc aaa tgc gtc att gtt tgt ccc cat gct gtg			2112
Val Cys Val Gln Cys Gly Lys Cys Val Ile Val Cys Pro His Ala Val			
690	695	700	
att cgg ggc aaa gtt tac gag gag gca gaa ttg gcc aat gct ccg gtc			2160
Ile Arg Gly Lys Val Tyr Glu Glu Ala Glu Leu Ala Asn Ala Pro Val			
705	710	715	720
agt ttc aaa ttt acc aat gcc aaa gac cat gat tgg caa ggt tct aag			2208
Ser Phe Lys Phe Thr Asn Ala Lys Asp His Asp Trp Gln Gly Ser Lys			
725	730	735	
ttc acc atc cag gta gcc ccg gaa gat tgc acc ggt tgc ggc atc tgt			2256
Phe Thr Ile Gln Val Ala Pro Glu Asp Cys Thr Gly Cys Gly Ile Cys			
740	745	750	
gtg gac gta tgc ccg gct aaa aat aaa tcc cag cct cgt tta agg gcg			2304
Val Asp Val Cys Pro Ala Lys Asn Lys Ser Gln Pro Arg Leu Arg Ala			
755	760	765	
att aat atg gct ccc cag tta ccc ttg cgg gaa cag gaa cgg gag aat			2352
Ile Asn Met Ala Pro Gln Leu Pro Leu Arg Glu Gln Glu Arg Glu Asn			
770	775	780	
tgg gac ttt ttc cta gat ttg ccc aac ccc gat cgc ctc agt ttg aat			2400
Trp Asp Phe Phe Leu Asp Leu Pro Asn Pro Asp Arg Leu Ser Leu Asn			
785	790	795	800
ttg aac aaa atc agc cat caa cag atg cag gag ccg tta ttt gaa ttt			2448

Leu Asn Lys Ile Ser His Gln Gln Met Gln Glu Pro Leu Phe Glu Phe	
805 810 815	
tct gga gcc tgt gcc ggt tgt ggg gaa acc cct tat ttg aaa ctg gtc	2496
Ser Gly Ala Cys Ala Gly Cys Gly Glu Thr Pro Tyr Leu Lys Leu Val	
820 825 830	
agt caa tta ttt ggc gat cgc atg tta gtg gcc aac gcc acc ggt tgc	2544
Ser Gln Leu Phe Gly Asp Arg Met Leu Val Ala Asn Ala Thr Gly Cys	
835 840 845	
tct tcc atc tat ggc ggc aac tta ccg aca act ccc tgg gcc caa aat	2592
Ser Ser Ile Tyr Gly Gly Asn Leu Pro Thr Thr Pro Trp Ala Gln Asn	
850 855 860	
gct gag ggt cgc ggt ccc gct tgg tcc aat tcc ctg ttt gaa gat aac	2640
Ala Glu Gly Arg Gly Pro Ala Trp Ser Asn Ser Leu Phe Glu Asp Asn	
865 870 875 880	
gct gaa ttt ggc ctt ggt ttc cga gtg gcg atc gac aag caa acg gaa	2688
Ala Glu Phe Gly Leu Gly Phe Arg Val Ala Ile Asp Lys Gln Thr Glu	
885 890 895	
ttt gca ggg gaa ttg cta aaa acc ttt gct ggg gag ttg gga gac agt	2736
Phe Ala Gly Glu Leu Leu Lys Thr Phe Ala Gly Glu Leu Gly Asp Ser	
900 905 910	
ttg gta agt gaa att ctc aac aat gcc caa acc act gaa gcg gat att	2784
Leu Val Ser Glu Ile Leu Asn Asn Ala Gln Thr Thr Glu Ala Asp Ile	
915 920 925	
ttt gaa caa cgg caa ttg gta gaa cag gtt aag caa cgt ttg caa aat	2832
Phe Glu Gln Arg Gln Leu Val Glu Gln Val Lys Gln Arg Leu Gln Asn	
930 935 940	
ctg gaa act ccc caa gcc caa atg ttc ctt tct gta gcg gat tac ctc	2880
Leu Glu Thr Pro Gln Ala Gln Met Phe Leu Ser Val Ala Asp Tyr Leu	
945 950 955 960	
gtg aag aaa agc gtt tgg att att ggt ggc gat ggc tgg gcc tac gac	2928
Val Lys Lys Ser Val Trp Ile Ile Gly Gly Asp Gly Trp Ala Tyr Asp	

965	970	975	
att ggg tac ggc ggt ttg gat cac gtc ctc gcc agt ggg cgt aat gtc			2976
Ile Gly Tyr Gly Gly Leu Asp His Val Leu Ala Ser Gly Arg Asn Val			
980	985	990	
aat atc ttg gtg atg gat acg gaa gtc tat tcc aac acc ggg ggc caa			3024
Asn Ile Leu Val Met Asp Thr Glu Val Tyr Ser Asn Thr Gly Gly Gln			
995	1000	1005	
gcc tcc aaa gcc act ccc cgg gcc gct gta gct aaa ttc gcc gct			3069
Ala Ser Lys Ala Thr Pro Arg Ala Ala Val Ala Lys Phe Ala Ala			
1010	1015	1020	
ggg ggt aaa ccc tct ccc aaa aaa gat ttg ggc tta atg gcc atg			3114
Gly Gly Lys Pro Ser Pro Lys Lys Asp Leu Gly Leu Met Ala Met			
1025	1030	1035	
acc tac ggc aac gtc tat gtg gcc agt atc gcc atg gga gcc aaa			3159
Thr Tyr Gly Asn Val Tyr Val Ala Ser Ile Ala Met Gly Ala Lys			
1040	1045	1050	
aat gag cag tcc att aaa gcc ttt atg gaa gcg gaa gcc tat ccc			3204
Asn Glu Gln Ser Ile Lys Ala Phe Met Glu Ala Glu Ala Tyr Pro			
1055	1060	1065	
ggt gtc tcg tta att att gcc tac tcc cac tgc att gcc cac ggc			3249
Gly Val Ser Leu Ile Ile Ala Tyr Ser His Cys Ile Ala His Gly			
1070	1075	1080	
att aat atg acc acc gcg atg aac cat caa aaa gag ttg gtg gac			3294
Ile Asn Met Thr Thr Ala Met Asn His Gln Lys Glu Leu Val Asp			
1085	1090	1095	
agc ggt cgt tgg ttg ctc tac cgc tat aac cct ttg ttg gcg gat			3339
Ser Gly Arg Trp Leu Leu Tyr Arg Tyr Asn Pro Leu Leu Ala Asp			
1100	1105	1110	
gaa ggt aaa aat ccc ctg caa ttg gat atg gga tcg cca aaa gta			3384
Glu Gly Lys Asn Pro Leu Gln Leu Asp Met Gly Ser Pro Lys Val			
1115	1120	1125	

gcc att gac aaa acg gtc tat tcg gaa aat cgc ttt gcc atg ctc 3429

Ala Ile Asp Lys Thr Val Tyr Ser Glu Asn Arg Phe Ala Met Leu
1130 1135 1140

acc cgc agt caa cca gag gag gcc aaa cgc tta atg aag tta gct 3474

Thr Arg Ser Gln Pro Glu Glu Ala Lys Arg Leu Met Lys Leu Ala
1145 1150 1155

caa ggg gat gtg aac act cgc tgg gcc atg tac gaa tat ctg gcg 3519

Gln Gly Asp Val Asn Thr Arg Trp Ala Met Tyr Glu Tyr Leu Ala

1160 1165 1170

aaa cgt tct ctg ggt ggg gaa att aac ggt aac aac cat ggt gtt 3564

Lys Arg Ser Leu Gly Gly Glu Ile Asn Gly Asn Asn His Gly Val

1175 1180 1185

tcc cca tct ccg gag gta att gct aaa tct gtt tag 3600

Ser Pro Ser Pro Glu Val Ile Ala Lys Ser Val

1190 1195

<210> 261

<211> 1199

<212> PRT

<213> Synechocystis PCC6803

<400> 261

Met Ser Leu Pro Thr Tyr Ala Thr Leu Asp Gly Asn Glu Ala Val Ala

1 5 10 15

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

20 25 30

Pro Ser Ser Pro Met Gly Glu Trp Ser Asp Ala Trp Ala Ala Glu His

35 40 45

Arg Pro Asn Leu Trp Gly Thr Val Pro Leu Val Val Glu Met Gln Ser

50 55 60

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

65 70 75 80

Leu Thr Thr Thr Phe Thr Ala Ser Gln Gly Leu Met Leu Met Leu Pro

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

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

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

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

 305 310 315 320
 Thr Val Gln Lys Ile Ala Val Leu Asp Arg Cys Lys Glu Pro Gly Ser
 325 330 335

Ile Gly Glu Pro Leu Tyr Gln Asp Val Leu Thr Ala Phe Phe Glu Ala
340 345 350
Gly Met Met Pro Lys Ile Ile Gly Gly Arg Tyr Gly Leu Ser Ser Lys
355 360 365
Glu Phe Thr Pro Ala Met Val Lys Gly Val Leu Asp His Leu Asn Gln
370 375 380
Thr Asn Pro Lys Asn His Phe Thr Val Gly Ile Asn Asp Asp Leu Ser
385 390 395 400
His Thr Ser Ile Asp Tyr Asp Pro Ser Phe Ser Thr Glu Ala Asp Ser
405 410 415
Val Val Arg Ala Ile Phe Tyr Gly Leu Gly Ser Asp Gly Thr Val Gly
420 425 430
Ala Asn Lys Asn Ser Ile Lys Ile Ile Gly Glu Asp Thr Asp Asn Tyr
435 440 445
Ala Gln Gly Tyr Phe Val Tyr Asp Ser Lys Lys Ser Gly Ser Val Thr
450 455 460
Val Ser His Leu Arg Phe Gly Pro Asn Pro Ile Leu Ser Thr Tyr Leu
465 470 475 480
Ile Ser Gln Ala Asn Phe Val Ala Cys His Gln Trp Glu Phe Leu Glu
485 490 495
Gln Phe Glu Val Leu Glu Pro Ala Val Asp Gly Gly Val Phe Leu Val
500 505 510
Asn Ser Pro Tyr Gly Pro Glu Glu Ile Trp Arg Glu Phe Pro Arg Lys
515 520 525
Val Gln Gln Glu Ile Ile Asp Lys Asn Leu Lys Val Tyr Thr Ile Asn
530 535 540
Ala Asn Asp Val Ala Arg Asp Ala Gly Met Gly Arg Arg Thr Asn Thr
545 550 555 560
Val Met Gln Thr Cys Phe Phe Ala Leu Ala Gly Val Leu Pro Arg Glu
565 570 575
Glu Ala Ile Ala Lys Ile Lys Gln Ser Val Gln Lys Thr Tyr Gly Lys

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

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

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

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

 820 825 830

Ser Gln Leu Phe Gly Asp Arg Met Leu Val Ala Asn Ala Thr Gly Cys
835 840 845

Ser Ser Ile Tyr Gly Gly Asn Leu Pro Thr Thr Pro Trp Ala Gln Asn
850 855 860

Ala Glu Gly Arg Gly Pro Ala Trp Ser Asn Ser Leu Phe Glu Asp Asn
865 870 875 880

Ala Glu Phe Gly Leu Gly Phe Arg Val Ala Ile Asp Lys Gln Thr Glu
885 890 895

Phe Ala Gly Glu Leu Leu Lys Thr Phe Ala Gly Glu Leu Gly Asp Ser
900 905 910

Leu Val Ser Glu Ile Leu Asn Asn Ala Gln Thr Thr Glu Ala Asp Ile
915 920 925

Phe Glu Gln Arg Gln Leu Val Glu Gln Val Lys Gln Arg Leu Gln Asn
930 935 940

Leu Glu Thr Pro Gln Ala Gln Met Phe Leu Ser Val Ala Asp Tyr Leu
945 950 955 960

Val Lys Lys Ser Val Trp Ile Ile Gly Gly Asp Gly Trp Ala Tyr Asp
965 970 975

Ile Gly Tyr Gly Gly Leu Asp His Val Leu Ala Ser Gly Arg Asn Val
980 985 990

Asn Ile Leu Val Met Asp Thr Glu Val Tyr Ser Asn Thr Gly Gly Gln
995 1000 1005

Ala Ser Lys Ala Thr Pro Arg Ala Ala Val Ala Lys Phe Ala Ala
1010 1015 1020

Gly Gly Lys Pro Ser Pro Lys Lys Asp Leu Gly Leu Met Ala Met
1025 1030 1035

Thr Tyr Gly Asn Val Tyr Val Ala Ser Ile Ala Met Gly Ala Lys
1040 1045 1050

Asn Glu Gln Ser Ile Lys Ala Phe Met Glu Ala Glu Ala Tyr Pro
1055 1060 1065

Gly Val Ser Leu Ile Ile Ala Tyr Ser His Cys Ile Ala His Gly

1070 1075 1080
 Ile Asn Met Thr Thr Ala Met Asn His Gln Lys Glu Leu Val Asp
 1085 1090 1095
 Ser Gly Arg Trp Leu Leu Tyr Arg Tyr Asn Pro Leu Leu Ala Asp
 1100 1105 1110
 Glu Gly Lys Asn Pro Leu Gln Leu Asp Met Gly Ser Pro Lys Val
 1115 1120 1125
 Ala Ile Asp Lys Thr Val Tyr Ser Glu Asn Arg Phe Ala Met Leu

1130 1135 1140
 Thr Arg Ser Gln Pro Glu Glu Ala Lys Arg Leu Met Lys Leu Ala
 1145 1150 1155
 Gln Gly Asp Val Asn Thr Arg Trp Ala Met Tyr Glu Tyr Leu Ala
 1160 1165 1170
 Lys Arg Ser Leu Gly Gly Glu Ile Asn Gly Asn Asn His Gly Val
 1175 1180 1185
 Ser Pro Ser Pro Glu Val Ile Ala Lys Ser Val
 1190 1195

<210> 262

<211> 531

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(531)

<223> fldA gene from E. coli encoding flavodoxin

<400> 262

atg gct atc act ggc atc ttt ttc ggc agc gac acc ggt aat acc gaa 48

Met Ala Ile Thr Gly Ile Phe Phe Gly Ser Asp Thr Gly Asn Thr Glu

1 5 10 15

aat atc gca aaa atg att caa aaa cag ctt ggt aaa gac gtt gcc gat 96

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

20 25 30

gtc cat gac att gca aaa agc agc aaa gaa gat ctg gaa gct tat gac 144

Val His Asp Ile Ala Lys Ser Ser Lys Glu Asp Leu Glu Ala Tyr Asp	
35 40 45	
att ctg ctg ctg ggc atc cca acc tgg tat tac ggc gaa gcg cag tgt	192
Ile Leu Leu Leu Gly Ile Pro Thr Trp Tyr Tyr Gly Glu Ala Gln Cys	
50 55 60	
gac tgg gat gac ttc ttc ccg act ctc gaa gag att gat ttc aac ggc	240
Asp Trp Asp Asp Phe Phe Pro Thr Leu Glu Glu Ile Asp Phe Asn Gly	
65 70 75 80	
aaa ctg gtt gcg ctg ttt ggt tgt ggt gac cag gaa gat tac gcc gaa	288
Lys Leu Val Ala Leu Phe Gly Cys Gly Asp Gln Glu Asp Tyr Ala Glu	
85 90 95	
tat ttc tgc gac gca ttg ggc acc atc cgc gac atc att gaa ccg cgc	336
Tyr Phe Cys Asp Ala Leu Gly Thr Ile Arg Asp Ile Ile Glu Pro Arg	
100 105 110	
ggt gca acc atc gtt ggt cac tgg cca act gcg ggc tat cat ttc gaa	384
Gly Ala Thr Ile Val Gly His Trp Pro Thr Ala Gly Tyr His Phe Glu	
115 120 125	
gca tca aaa ggt ctg gca gat gac gac cac ttt gtc ggt ctg gct atc	432
Ala Ser Lys Gly Leu Ala Asp Asp Asp His Phe Val Gly Leu Ala Ile	
130 135 140	
gac gaa gac cgt cag ccg gaa ctg acc gct gaa cgt gta gaa aaa tgg	480
Asp Glu Asp Arg Gln Pro Glu Leu Thr Ala Glu Arg Val Glu Lys Trp	
145 150 155 160	
gtt aaa cag att tct gaa gag ttg cat ctc gac gaa att ctc aat gcc	528
Val Lys Gln Ile Ser Glu Glu Leu His Leu Asp Glu Ile Leu Asn Ala	
165 170 175	
tga	531
<210> 263	
<211> 176	
<212> PRT	
<213> Escherichia coli	

<400> 263

Met Ala Ile Thr Gly Ile Phe Phe Gly Ser Asp Thr Gly Asn Thr Glu

1 5 10 15

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

20 25 30

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

35 40 45

Ile Leu Leu Leu Gly Ile Pro Thr Trp Tyr Tyr Gly Glu Ala Gln Cys

50 55 60

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

65 70 75 80

Lys Leu Val Ala Leu Phe Gly Cys Gly Asp Gln Glu Asp Tyr Ala Glu

85 90 95

Tyr Phe Cys Asp Ala Leu Gly Thr Ile Arg Asp Ile Ile Glu Pro Arg

100 105 110

Gly Ala Thr Ile Val Gly His Trp Pro Thr Ala Gly Tyr His Phe Glu

115 120 125

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

130 135 140

Asp Glu Asp Arg Gln Pro Glu Leu Thr Ala Glu Arg Val Glu Lys Trp

145 150 155 160

Val Lys Gln Ile Ser Glu Glu Leu His Leu Asp Glu Ile Leu Asn Ala

165 170 175

<210> 264

<211> 522

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(522)

<223> fldB gene from E. coli encoding flavodoxin

<400> 264

atg aat atg ggt ctt ttt tac ggt tcc agc acc tgt tac acc gaa atg

48

165

170

<210> 265

<211> 173

<212> PRT

<213> Escherichia coli

<400> 265

Met Asn Met Gly Leu Phe Tyr Gly Ser Ser Thr Cys Tyr Thr Glu Met

1 5 10 15

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

20 25 30

His Asn Leu Lys Asp Asp Ser Pro Lys Leu Met Glu Gln Tyr Asp Val

35 40 45

Leu Ile Leu Gly Ile Pro Thr Trp Asp Phe Gly Glu Ile Gln Glu Asp

50 55 60

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

65 70 75 80

Ile Val Ala Leu Tyr Gly Leu Gly Asp Gln Leu Gly Tyr Gly Glu Trp

85 90 95

Phe Leu Asp Ala Leu Gly Met Leu His Asp Lys Leu Ser Thr Lys Gly

100 105 110

Val Lys Phe Val Gly Tyr Trp Pro Thr Glu Gly Tyr Glu Phe Thr Ser

115 120 125

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

130 135 140

Asp Glu Thr Asn Gln Tyr Asp Leu Ser Asp Glu Arg Ile Gln Ser Trp

145 150 155 160

Cys Glu Gln Ile Leu Asn Glu Met Ala Glu His Tyr Ala

165 170

<210> 266

<211> 477

<212> DNA

<213> Bacillus subtilis

<220><

221> CDS

<222> (1)..(477)

<223> ykuN gene from Bacillus subtilis encoding flavodoxin

<400> 266

atg gct aaa gcc ttg att aca tat gcc agc atg tca gga aat aca gaa 48

Met Ala Lys Ala Leu Ile Thr Tyr Ala Ser Met Ser Gly Asn Thr Glu

1 5 10 15

gac att gcc ttc ata ata aaa gat acg ctt cag gaa tat gag ttg gat 96

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

20 25 30

atc gat tgt gtc gag ata aat gat atg gat gcg tct tgt tta acc tcc 144

Ile Asp Cys Val Glu Ile Asn Asp Met Asp Ala Ser Cys Leu Thr Ser

35 40 45

tat gat tat gta ctg att ggc acc tat aca tgg ggg gac ggc gat ttg 192

Tyr Asp Tyr Val Leu Ile Gly Thr Tyr Thr Trp Gly Asp Gly Asp Leu

50 55 60

ccc tac gaa gcg gag gat ttt ttc gaa gag gtc aaa cag att cag ctt 240

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

65 70 75 80

aat ggt tta aaa aca gcc tgc ttc ggg tct ggc gat tat tct tat cca 288

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

85 90 95

aag ttt tgc gaa gcg gtg aat ttg ttc aat gtc atg ctg caa gag gcg 336

Lys Phe Cys Glu Ala Val Asn Leu Phe Asn Val Met Leu Gln Glu Ala

100 105 110

gga gct gct gtt tac cag gaa aca cta aaa att gaa tta gcg cct gaa 384

Gly Ala Ala Val Tyr Gln Glu Thr Leu Lys Ile Glu Leu Ala Pro Glu

115 120 125

aca gat gaa gat gtg gaa agc tgc cga gcg ttt gcg aga ggt ttt ctt 432

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

130 135 140

gca tgg gca gat tat atg aac aag gaa aaa atc cat gtt tca taa 477

Ala Trp Ala Asp Tyr Met Asn Lys Glu Lys Ile His Val Ser

145 150 155

<210> 267

<211> 158

<212> PRT

<213> Bacillus subtilis

<400> 267

Met Ala Lys Ala Leu Ile Thr Tyr Ala Ser Met Ser Gly Asn Thr Glu

1 5 10 15

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

20 25 30

Ile Asp Cys Val Glu Ile Asn Asp Met Asp Ala Ser Cys Leu Thr Ser

35 40 45

Tyr Asp Tyr Val Leu Ile Gly Thr Tyr Thr Trp Gly Asp Gly Asp Leu

50 55 60

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

65 70 75 80

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

85 90 95

Lys Phe Cys Glu Ala Val Asn Leu Phe Asn Val Met Leu Gln Glu Ala

100 105 110

Gly Ala Ala Val Tyr Gln Glu Thr Leu Lys Ile Glu Leu Ala Pro Glu

115 120 125

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

130 135 140

Ala Trp Ala Asp Tyr Met Asn Lys Glu Lys Ile His Val Ser

145 150 155

<210> 268

<211> 513

<212> DNA

<213> Synechocystis PCC6803

<220><221> CDS

<222> (1)..(513)

<223> isiB gene from Synechocystis encoding flavodoxin

<400> 268

atg aca aaa att gga ctt ttt tac ggt act caa acc ggc aac act gaa 48

Met Thr Lys Ile Gly Leu Phe Tyr Gly Thr Gln Thr Gly Asn Thr Glu

1 5 10 15

acc att gct gaa ctg att caa aaa gaa atg ggc ggc gat agt gtg gtc 96

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

20 25 30

gat atg atg gat ata tcc cag gct gat gtt gat gat ttt agg caa tat 144

Asp Met Met Asp Ile Ser Gln Ala Asp Val Asp Asp Phe Arg Gln Tyr

35 40 45

agt tgc ctg att atc ggt tgt ccc acc tgg aat gtg ggg gaa ctc cag 192

Ser Cys Leu Ile Ile Gly Cys Pro Thr Trp Asn Val Gly Glu Leu Gln

50 55 60

agt gat tgg gaa ggc ttt tat gac caa tta gac gaa att gat ttt aat 240

Ser Asp Trp Glu Gly Phe Tyr Asp Gln Leu Asp Glu Ile Asp Phe Asn

65 70 75 80

ggc aaa aaa gta gcc tat ttt ggt gct ggc gat cag gtt ggt tat gca 288

Gly Lys Lys Val Ala Tyr Phe Gly Ala Gly Asp Gln Val Gly Tyr Ala

85 90 95

gat aat ttt caa gac gcc atg ggc att tta gaa gaa aaa atc agt gga 336

Asp Asn Phe Gln Asp Ala Met Gly Ile Leu Glu Glu Lys Ile Ser Gly

100 105 110

tta ggc ggt aaa aca gtg ggg ttt tgg ccc acc gct ggc tat gat ttt 384

Leu Gly Gly Lys Thr Val Gly Phe Trp Pro Thr Ala Gly Tyr Asp Phe

115 120 125

gac gaa tca aaa gcg gtg aaa aat ggg aaa ttt gtt ggt tta gct ttg 432

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

130	135	140	
gac gaa gat aat cag cca gag tta aca gaa tta aga gta aag aca tgg			480
Asp Glu Asp Asn Gln Pro Glu Leu Thr Glu Leu Arg Val Lys Thr Trp			
145	150	155	160
gta agt gaa att aaa cca att ttg caa tcc tag			513
Val Ser Glu Ile Lys Pro Ile Leu Gln Ser			
	165	170	

<210> 269

<211> 170

<212> PRT

<213> Synechocystis PCC6803

<400> 269

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

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

115	120	125
Asp Glu Ser Lys Ala Val Lys Asn Gly Lys Phe Val Gly Leu Ala Leu		
130	135	140

Asp Glu Asp Asn Gln Pro Glu Leu Thr Glu Leu Arg Val Lys Thr Trp
 145 150 155 160
 Val Ser Glu Ile Lys Pro Ile Leu Gln Ser
 165 170

<210> 270

<211> 585

<212> DNA

<213> Streptomyces venezuelae

<220><221> CDS

<222> (1)..(585)

<223> wrbA gene from Streptomyces venezuelae encoding flavodoxin

<400> 270

atg acc acc ccc gtc gtc tcc atc gcc tac cac tcc ggc tac ggc cac 48

Met Thr Thr Pro Val Val Ser Ile Ala Tyr His Ser Gly Tyr Gly His

1 5 10 15

acc gcg gtc ctg gcc gag gcc gtc cgt gac ggc gcc gcc gac gcg ggc 96

Thr Ala Val Leu Ala Glu Ala Val Arg Asp Gly Ala Ala Asp Ala Gly

20 25 30

gcc acc gtc cac ctg atc aag gtc gac ggg atc acc gag gcg gag tgg 144

Ala Thr Val His Leu Ile Lys Val Asp Gly Ile Thr Glu Ala Glu Trp

35 40 45

gag ctg ctc gac gcc tcc gac gcg atc gtc ttc ggc tcc ccg acc tac 192

Glu Leu Leu Asp Ala Ser Asp Ala Ile Val Phe Gly Ser Pro Thr Tyr

50 55 60

atg ggc acc gcc tcc ggt gcc ttc cac cag ttc gcc gag gac tcc tcg 240

Met Gly Thr Ala Ser Gly Ala Phe His Gln Phe Ala Glu Asp Ser Ser

65 70 75 80

aag gcg tgg ttc ggc gac gtc tgg ctg gac aag ctc gcc gcc ggc ttc 288

Lys Arg Trp Phe Gly Asp Val Trp Leu Asp Lys Leu Ala Ala Gly Phe

85 90 95

acc aac tcc ggc tcc aag agc ggc gac aag ctg cac acc ctg cag tac 336

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

100 105 110
ttc cag atc ctc gcc ggc cag cac ggc atg cac tgg gtc aac ctc ggc 384

Phe Gln Ile Leu Ala Gly Gln His Gly Met His Trp Val Asn Leu Gly

115 120 125
ctg aag ccc ggc tgg aac acc agc gag gcc tcc gag aac gac atc aac 432

Leu Lys Pro Gly Trp Asn Thr Ser Glu Ala Ser Glu Asn Asp Ile Asn

130 135 140

cgc ctc ggc ttc ttc tcc ggc gcc gcc ggc cag acc ccc gcg gac ctg 480

Arg Leu Gly Phe Phe Ser Gly Ala Ala Gly Gln Thr Pro Ala Asp Leu

145 150 155 160
ggc ccc gag gcc gtc cac aag gcc gac gtc gcc acc gcc gaa cac ctc 528

Gly Pro Glu Ala Val His Lys Ala Asp Val Ala Thr Ala Glu His Leu

165 170 175
ggc cgc cgc gtc gcc gag acc gcc cgc acc ttc gcg gcc ggc aag gcc 576

Gly Arg Arg Val Ala Glu Thr Ala Arg Thr Phe Ala Ala Gly Lys Ala

180 185 190
gcc gcc tga 585

Ala Ala

<210> 271

<211> 194

<212> PRT

<213> Streptomyces venezuelae

<400> 271

Met Thr Thr Pro Val Val Ser Ile Ala Tyr His Ser Gly Tyr Gly His

1 5 10 15
Thr Ala Val Leu Ala Glu Ala Val Arg Asp Gly Ala Ala Asp Ala Gly

20 25 30
Ala Thr Val His Leu Ile Lys Val Asp Gly Ile Thr Glu Ala Glu Trp

35 40 45

Glu Leu Leu Asp Ala Ser Asp Ala Ile Val Phe Gly Ser Pro Thr Tyr

50 55 60

Met Gly Thr Ala Ser Gly Ala Phe His Gln Phe Ala Glu Asp Ser Ser

65 70 75 80

Lys Arg Trp Phe Gly Asp Val Trp Leu Asp Lys Leu Ala Ala Gly Phe

85 90 95

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

100 105 110

Phe Gln Ile Leu Ala Gly Gln His Gly Met His Trp Val Asn Leu Gly

115 120 125

Leu Lys Pro Gly Trp Asn Thr Ser Glu Ala Ser Glu Asn Asp Ile Asn

130 135 140

Arg Leu Gly Phe Phe Ser Gly Ala Ala Gly Gln Thr Pro Ala Asp Leu

145 150 155 160

Gly Pro Glu Ala Val His Lys Ala Asp Val Ala Thr Ala Glu His Leu

165 170 175

Gly Arg Arg Val Ala Glu Thr Ala Arg Thr Phe Ala Ala Gly Lys Ala

180 185 190

Ala Ala

<210> 272

<211> 459

<

212> DNA

<213> Methanococcus aeolicus

<220><221> CDS

<222> (1)..(459)

<223> PRK06242 gene from Methanococcus aeolicus encoding flavodoxin

<400> 272

atg aaa ata tta att att tgt aaa tcc gta cac cat gga aac act aaa 48

Met Lys Ile Leu Ile Ile Cys Lys Ser Val His His Gly Asn Thr Lys

1 5 10 15

aaa ata gca gat gcc atg gca gag gtt tta aat gca gag gtt att gca 96

Lys Ile Ala Asp Ala Met Ala Glu Val Leu Asn Ala Glu Val Ile Ala

20 25 30
cct gaa aat gta agt tcc gaa gat atc aaa aaa tat gat ttg gtg gga 144

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

35 40 45
ttt ggc tct gga ata tat att ggg aaa cat cat aaa aag cta tta aaa 192

Phe Gly Ser Gly Ile Tyr Ile Gly Lys His His Lys Lys Leu Leu Lys

50 55 60
ctt gcg gat aat ctt cca aat gga gaa aat aaa aca gta ttt gta ttt 240

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

65 70 75 80
tcc aca agc gat aac tgg aag caa aat tac cat aag cca tta atg gat 288

Ser Thr Ser Asp Asn Trp Lys Gln Asn Tyr His Lys Pro Leu Met Asp

85 90 95
aaa cta aat tcc aga gga tat aaa aca gta gga gaa ttc aac tgt aaa 336

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

100 105 110
ggg ttt gat gac tgg ttt ata ttt aaa tta att ggt ggt aga aat aaa 384

Gly Phe Asp Asp Trp Phe Ile Phe Lys Leu Ile Gly Gly Arg Asn Lys

115 120 125
gga cat cca aat aaa aaa gat att gaa aat gca aaa aaa ttt gct gaa 432

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

130 135 140
aat ata aag aat ata gaa aat ata tag 459

Asn Ile Lys Asn Ile Glu Asn Ile

145 150
<210> 273

<211> 152
<212> PRT

<213> Methanococcus aeolicus
<400> 273

Met Lys Ile Leu Ile Ile Cys Lys Ser Val His His Gly Asn Thr Lys
1 5 10 15

Lys Ile Ala Asp Ala Met Ala Glu Val Leu Asn Ala Glu Val Ile Ala
20 25 30

Pro Glu Asn Val Ser Ser Glu Asp Ile Lys Lys Tyr Asp Leu Val Gly
35 40 45

Phe Gly Ser Gly Ile Tyr Ile Gly Lys His His Lys Lys Leu Leu Lys
50 55 60

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

Ser Thr Ser Asp Asn Trp Lys Gln Asn Tyr His Lys Pro Leu Met Asp
85 90 95

Lys Leu Asn Ser Arg Gly Tyr Lys Thr Val Gly Glu Phe Asn Cys Lys
100 105 110

Gly Phe Asp Asp Trp Phe Ile Phe Lys Leu Ile Gly Gly Arg Asn Lys
115 120 125

Gly His Pro Asn Lys Lys Asp Ile Glu Asn Ala Lys Lys Phe Ala Glu
130 135 140

Asn Ile Lys Asn Ile Glu Asn Ile
145 150

<210> 274

<211> 336

<212> DNA

<213> Escherichia coli

<220><221> CDS

<222> (1)..(336)

<223> fdx gene from E. coli encoding ferredoxin

<400> 274

atg cca aag att gtt att ttg cct cat cag gat ctc tgc cct gat ggc 48

Met Pro Lys Ile Val Ile Leu Pro His Gln Asp Leu Cys Pro Asp Gly
1 5 10 15

gct gtt ctg gaa gct aat agc ggt gaa acc att ctc gac gca gct ctg 96

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

20 25 30

cgt aac ggt atc gag att gaa cac gcc tgt gaa aaa tcc tgt gct tgc 144

Arg Asn Gly Ile Glu Ile Glu His Ala Cys Glu Lys Ser Cys Ala Cys

35 40 45

acc acc tgc cac tgc atc gtt cgt gaa ggt ttt gac tca ctg ccg gaa 192

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

50 55 60

agc tca gag cag gaa gac gac atg ctg gac aaa gcc tgg gga ctg gag 240

Ser Ser Glu Gln Glu Asp Asp Met Leu Asp Lys Ala Trp Gly Leu Glu

65 70 75 80

ccg gaa agc cgt tta agc tgc cag gcg cgc gtt acc gac gaa gat tta 288

Pro Glu Ser Arg Leu Ser Cys Gln Ala Arg Val Thr Asp Glu Asp Leu

85 90 95

gta gtc gaa atc ccg cgt tac act atc aac cat gcg cgt gag cat taa 336

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

100 105 110

<210> 275

<211> 111

<212> PRT

<213> Escherichia coli

<400> 275

Met Pro Lys Ile Val Ile Leu Pro His Gln Asp Leu Cys Pro Asp Gly

1 5 10 15

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

20 25 30

Arg Asn Gly Ile Glu Ile Glu His Ala Cys Glu Lys Ser Cys Ala Cys

35 40 45

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

50 55 60

Ser Ser Glu Gln Glu Asp Asp Met Leu Asp Lys Ala Trp Gly Leu Glu
65 70 75 80
Pro Glu Ser Arg Leu Ser Cys Gln Ala Arg Val Thr Asp Glu Asp Leu
85 90 95

Val Val Glu Ile Pro Arg Tyr Thr Ile Asn His Ala Arg Glu His
100 105 110

<210> 276

<211> 249

<212> DNA

<213> Bacillus subtilis

<220><221> CDS

<222> (1)..(249)

<223> fer gene from B. subtilis encoding ferredoxin

<400> 276

atg gca aag tac aca atc gta gac aaa gat aca tgt att gca tgc ggc 48

Met Ala Lys Tyr Thr Ile Val Asp Lys Asp Thr Cys Ile Ala Cys Gly

1 5 10 15

gct tgc gga gct gct gca cca gac att tac gat tac gat gat gaa ggc 96

Ala Cys Gly Ala Ala Ala Pro Asp Ile Tyr Asp Tyr Asp Asp Glu Gly

20 25 30

atc gcg ttc gta acg ctt gat gaa aac aaa ggt gtt gtc gaa gtt cct 144

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

35 40 45

gag gta ctg gaa gaa gat atg att gac gca ttt gaa gga tgc cct act 192

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

50 55 60

gat tcc atc aaa gtg gcg gat gag cca ttt gaa ggc gac ccg ctt aaa 240

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

65 70 75 80

ttt gaa tag 249

Phe Glu

<210> 277

<211> 82

<212> PRT

<213> *Bacillus subtilis*

<400> 277

Met Ala Lys Tyr Thr Ile Val Asp Lys Asp Thr Cys Ile Ala Cys Gly

1 5 10 15

Ala Cys Gly Ala Ala Ala Pro Asp Ile Tyr Asp Tyr Asp Asp Glu Gly

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

Phe Glu

<210> 278

<211> 321

<212> DNA

<213> *Corynebacterium glutamicum*

<220><221> CDS

<222> (1)..(321)

<223> fdxB gene from *Corynebacterium glutamicum* encoding ferredoxin

<400> 278

atg tct act att cat ttc att gat cat gct ggc aaa acc cgc acc atc 48

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

1 5 10 15

gag gcg act gtt ggt gat tca gta atg gag acc gca gtc cga aac gga 96

Glu Ala Thr Val Gly Asp Ser Val Met Glu Thr Ala Val Arg Asn Gly

20 25 30

gtg cct gga att gtt gct gaa tgc ggc ggt tcc tta tcg tgt gca acc 144

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

35 40 45
tgc cat gtg ttt gtt gac cct gca cag tat gat gcg ctt ccc cca atg 192

Cys His Val Phe Val Asp Pro Ala Gln Tyr Asp Ala Leu Pro Pro Met

50 55 60
gag gag atg gaa gat gaa atg ctg tgg ggt gct gcc gtg gac cgt gag 240

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

65 70 75 80

gat tgc tcc cgt ttg tct tgc caa atc aag gtc acc gaa ggc atg gat 288

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

85 90 95
ctt tcg ttg acc acg cca gaa acg caa gtg tga 321

Leu Ser Leu Thr Thr Pro Glu Thr Gln Val

100 105

<210> 279

<211> 106

<212> PRT

<213> Corynebacterium glutamicum

<400> 279

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

1 5 10 15

Glu Ala Thr Val Gly Asp Ser Val Met Glu Thr Ala Val Arg Asn Gly

20 25 30

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

35 40 45

Cys His Val Phe Val Asp Pro Ala Gln Tyr Asp Ala Leu Pro Pro Met

50 55 60

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

65 70 75 80

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

85 90 95

Leu Ser Leu Thr Thr Pro Glu Thr Gln Val

100

105

<210> 280

<211> 561

<212> DNA

<213> Synechocystis PCC6803

<220><221> CDS

<222> (1)..(561)

<223> fdx gene from Synechocystis encoding ferredoxin

<400> 280

atg acc atg cca cca tta tgg aat tgc tct gtc gcc aac agg gtt aat 48

Met Thr Met Pro Pro Leu Trp Asn Cys Ser Val Ala Asn Arg Val Asn

1 5 10 15

gcc att gtt gcc agt act aag gag gat tgt gtg gct aaa act att aag 96

Ala Ile Val Ala Ser Thr Lys Glu Asp Cys Val Ala Lys Thr Ile Lys

20 25 30

ctc gac ccc att gat tta aaa gtc gcc atc gag acc aac gat aac ctg 144

Leu Asp Pro Ile Asp Leu Lys Val Ala Ile Glu Thr Asn Asp Asn Leu

35 40 45

ctc tcg ggg ttg ctc ggt cag gat tta cgg atc atg aag gag tgt ggt 192

Leu Ser Gly Leu Leu Gly Gln Asp Leu Arg Ile Met Lys Glu Cys Gly

50 55 60

ggg cgg ggt atg tgt gcc act tgt cac gtt tac atc acc gct ggg atg 240

Gly Arg Gly Met Cys Ala Thr Cys His Val Tyr Ile Thr Ala Gly Met

65 70 75 80

gag agt ctt tct ccc ctc aac cgt cgg gag cag cgc acc cta gag gtg 288

Glu Ser Leu Ser Pro Leu Asn Arg Arg Glu Gln Arg Thr Leu Glu Val

85 90 95

atc acc acc cac aat cgt tat tcc cgt ttg gct tgc caa gcc cgg gtg 336

Ile Thr Thr His Asn Arg Tyr Ser Arg Leu Ala Cys Gln Ala Arg Val

100 105 110

ttg gat gaa ggc gtg gtg gtg gaa ttg ccc gct ggg atg tac gtc agt 384
 Leu Asp Glu Gly Val Val Val Glu Leu Pro Ala Gly Met Tyr Val Ser
 115 120 125

gaa att gag gac atc gag gag ctg att ggc cgt cga gcg gag gaa aat 432
 Glu Ile Glu Asp Ile Glu Glu Leu Ile Gly Arg Arg Ala Glu Glu Asn
 130 135 140

att ctc aat cct cgg gat ggg agc atc cta gtg gaa aaa ggt aag tta 480

Ile Leu Asn Pro Arg Asp Gly Ser Ile Leu Val Glu Lys Gly Lys Leu
 145 150 155 160

att acc cgt tcc atg att agt caa cta gat gac cag tta cag gcg gcc 528
 Ile Thr Arg Ser Met Ile Ser Gln Leu Asp Asp Gln Leu Gln Ala Ala
 165 170 175

aaa att cag att gtc aac gat acc gat gaa taa 561
 Lys Ile Gln Ile Val Asn Asp Thr Asp Glu

180 185

<210> 281
 <211> 186
 <212> PRT
 <213> Synechocystis PCC6803
 <400> 281

Met Thr Met Pro Pro Leu Trp Asn Cys Ser Val Ala Asn Arg Val Asn
 1 5 10 15

Ala Ile Val Ala Ser Thr Lys Glu Asp Cys Val Ala Lys Thr Ile Lys
 20 25 30

Leu Asp Pro Ile Asp Leu Lys Val Ala Ile Glu Thr Asn Asp Asn Leu
 35 40 45

Leu Ser Gly Leu Leu Gly Gln Asp Leu Arg Ile Met Lys Glu Cys Gly
 50 55 60

Gly Arg Gly Met Cys Ala Thr Cys His Val Tyr Ile Thr Ala Gly Met
 65 70 75 80

Glu Ser Leu Ser Pro Leu Asn Arg Arg Glu Gln Arg Thr Leu Glu Val
 85 90 95

Ile Thr Thr His Asn Arg Tyr Ser Arg Leu Ala Cys Gln Ala Arg Val
100 105 110

Leu Asp Glu Gly Val Val Val Glu Leu Pro Ala Gly Met Tyr Val Ser
115 120 125

Glu Ile Glu Asp Ile Glu Glu Leu Ile Gly Arg Arg Ala Glu Glu Asn
130 135 140

Ile Leu Asn Pro Arg Asp Gly Ser Ile Leu Val Glu Lys Gly Lys Leu
145 150 155 160

Ile Thr Arg Ser Met Ile Ser Gln Leu Asp Asp Gln Leu Gln Ala Ala
165 170 175

Lys Ile Gln Ile Val Asn Asp Thr Asp Glu
180 185

<210> 282

<211> 294

<212> DNA

<213> Streptomyces venezuelae

<220><221> CDS

<222> (1)..(294)

<223> SVEN_7039 gene from Streptomyces venezuelae encoding ferredoxin

<400> 282

atg gcg tac gtc gtc acc gac gag tgc atc ggc tgc aag tac acg gac 48

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

1 5 10 15

tgt gtg gac gtc tgc ccc gtg agc tgt ttc cac gag ggc ccc gag atg 96

Cys Val Asp Val Cys Pro Val Ser Cys Phe His Glu Gly Pro Glu Met

20 25 30

ctc tac atc aac ccc gag gaa tgc atc gac tgc aac gcg tgc gtc gcc 144

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

35 40 45

gag tgc ccg ccc gag gcc atc tgg gcg gac gtc gac ctg ccg gag gac 192

Glu Cys Pro Pro Glu Ala Ile Trp Ala Asp Val Asp Leu Pro Glu Asp

50 55 60
aag ctc cag tgg atc gag atc aac gga gag atg agt gcc aag tac ccg 240
Lys Leu Gln Trp Ile Glu Ile Asn Gly Glu Met Ser Ala Lys Tyr Pro
65 70 75 80
gtt ctc cac gag agc cgg ggc ccc cac gga cag ccc tcc agc cag cct 288
Val Leu His Glu Ser Arg Gly Pro His Gly Gln Pro Ser Ser Gln Pro
85 90 95
tcc tga 294
Ser

<210> 283

<211> 97

<212> PRT

<213> Streptomyces venezuelae

<400> 283

Met Ala Tyr Val Val Thr Asp Glu Cys Ile Gly Cys Lys Tyr Thr Asp
1 5 10 15
Cys Val Asp Val Cys Pro Val Ser Cys Phe His Glu Gly Pro Glu Met

20 25 30
Leu Tyr Ile Asn Pro Glu Glu Cys Ile Asp Cys Asn Ala Cys Val Ala
35 40 45
Glu Cys Pro Pro Glu Ala Ile Trp Ala Asp Val Asp Leu Pro Glu Asp
50 55 60
Lys Leu Gln Trp Ile Glu Ile Asn Gly Glu Met Ser Ala Lys Tyr Pro
65 70 75 80
Val Leu His Glu Ser Arg Gly Pro His Gly Gln Pro Ser Ser Gln Pro
85 90 95
Ser

<210> 284

<211> 1746

<212> DNA

<213> Methanococcus aeolicus

<220><221> CDS

<222> (1)..(1746)

<223> fdx gene from Methanococcus aeolicus encoding ferredoxin

<400> 284

atg ggg ggt gtt atg atg tat aat att aca tac ata aaa gag gat gga 48

Met Gly Gly Val Met Met Tyr Asn Ile Thr Tyr Ile Lys Glu Asp Gly

1 5 10 15

act aaa aaa tca att aaa gtt aaa gaa gga acc aca ata ctt gaa gga 96

Thr Lys Lys Ser Ile Lys Val Lys Glu Gly Thr Thr Ile Leu Glu Gly

20 25 30

gcg ata aaa gcg gga gtt tat att gat gct cca tgt gga acg ggg aaa 144

Ala Ile Lys Ala Gly Val Tyr Ile Asp Ala Pro Cys Gly Thr Gly Lys

35 40 45

tgt ggt aag tgt aaa gtt tta gtg gag aaa ggt tta gaa aat att gat 192

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

50 55 60

aag gat agt att gtg gaa gat gag tat gca ctg gca tgt gtg gca aaa 240

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

65 70 75 80

gtt tat ggg gac ata tca att aat gtt cca aat ttc caa ggt gtg gtt 288

Val Tyr Gly Asp Ile Ser Ile Asn Val Pro Asn Phe Gln Gly Val Val

85 90 95

tgt aag gat atc acc aac gaa gtt ggt gag cta caa act cga aga gtt 336

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

100 105 110

tgt tca att acc gaa caa tgt aaa ggt gag cta caa aat ctc gag gga 384

Cys Ser Ile Thr Glu Gln Cys Lys Gly Glu Leu Gln Asn Leu Glu Gly

115 120 125

ttt cat ccg atg tct tta aac ccc gat att gga ata aat aaa att act 432

Phe His Pro Met Ser Leu Asn Pro Asp Ile Gly Ile Asn Lys Ile Thr

130	135	140	
aca aca gta ttg gaa tca tct aac tat aac cta aca tta gat gca ata			480
Thr Thr Val Leu Glu Ser Ser Asn Tyr Asn Leu Thr Leu Asp Ala Ile			
145	150	155	160
aat aag ctt aat tct atg aag tta tcg gac gaa gta act tta ata tta			528
Asn Lys Leu Asn Ser Met Lys Leu Ser Asp Glu Val Thr Leu Ile Leu			
	165	170	175
aag gga gat aat gtc gtt aat gta gaa aaa gat ttt tct gga att tat			576
Lys Gly Asp Asn Val Val Asn Val Glu Lys Asp Phe Ser Gly Ile Tyr			
	180	185	190
ggg ctt tca att gat att ggg act aca tct gtt gtt gta tat ctt gtt			624
Gly Leu Ser Ile Asp Ile Gly Thr Thr Ser Val Val Val Tyr Leu Val			
	195	200	205
gat att tct aaa ggt att gtt tta gat aat att tct ttt tta aat cct			672
Asp Ile Ser Lys Gly Ile Val Leu Asp Asn Ile Ser Phe Leu Asn Pro			
210	215	220	
cag agg cag ttt ggg gca gat gtt gtt tca aga ata gca tac aac aac			720
Gln Arg Gln Phe Gly Ala Asp Val Val Ser Arg Ile Ala Tyr Asn Asn			
225	230	235	240
gga att tta ctg caa aaa aca ctt ata act gaa tta aac gat tct ata			768
Gly Ile Leu Leu Gln Lys Thr Leu Ile Thr Glu Leu Asn Asp Ser Ile			
	245	250	255
tca aaa tta tgt tca aac aat aac ata aaa atg gat aat att tat gaa			816
Ser Lys Leu Cys Ser Asn Asn Asn Ile Lys Met Asp Asn Ile Tyr Glu			
	260	265	270
gtt agt gtg gta gga aac act gct atg ata cac ttc ttt tat gga ata			864
Val Ser Val Val Gly Asn Thr Ala Met Ile His Phe Phe Tyr Gly Ile			
	275	280	285
gtc cca aaa aat ctt gca acc cat cct tat gtt cca aca ttt aaa aac			912
Val Pro Lys Asn Leu Ala Thr His Pro Tyr Val Pro Thr Phe Lys Asn			
290	295	300	

tca cca tat ctt cct gca aaa gag ttg ggg cta aac cta aga aac gca	960
Ser Pro Tyr Leu Pro Ala Lys Glu Leu Gly Leu Asn Leu Arg Asn Ala	
305 310 315 320	
tac att tac aca ctt ccg ata ata gga ggt tat gtt ggg gca gac aca	1008
Tyr Ile Tyr Thr Leu Pro Ile Ile Gly Gly Tyr Val Gly Ala Asp Thr	
325 330 335	
gtt gga gca att tta tca tct gaa atg cat aaa aaa gat gat ata agt	1056
Val Gly Ala Ile Leu Ser Ser Glu Met His Lys Lys Asp Asp Ile Ser	
340 345 350	
ctc ctt ata gat att ggc aca aat ggg gaa att gtt tta ggg aat aaa	1104
Leu Leu Ile Asp Ile Gly Thr Asn Gly Glu Ile Val Leu Gly Asn Lys	
355 360 365	
gaa aag tta tta acc tgt tca tgt gca gca ggt cct gca ttt gag ggt	1152
Glu Lys Leu Leu Thr Cys Ser Cys Ala Ala Gly Pro Ala Phe Glu Gly	
370 375 380	
gtc agc ata gag cat ggg aca aat gct aga gag ggg gca gta tgt aga	1200
Val Ser Ile Glu His Gly Thr Asn Ala Arg Glu Gly Ala Val Cys Arg	
385 390 395 400	
gta aaa ata gat gaa aat aac ata tac tat gag acc ata gga aat aaa	1248
Val Lys Ile Asp Glu Asn Asn Ile Tyr Tyr Glu Thr Ile Gly Asn Lys	
405 410 415	
acg ccc cct att gga ata tgc ggg tct gga ata ata gat att gta gct	1296
Thr Pro Pro Ile Gly Ile Cys Gly Ser Gly Ile Ile Asp Ile Val Ala	
420 425 430	
gaa ttt tta aaa tcc gga tta att aat aaa acc ggt aga ttt act gga	1344
Glu Phe Leu Lys Ser Gly Leu Ile Asn Lys Thr Gly Arg Phe Thr Gly	
435 440 445	
gaa cat aaa aac tta aag gaa aat aaa ttt atc att gaa gat tct att	1392
Glu His Lys Asn Leu Lys Glu Asn Lys Phe Ile Ile Glu Asp Ser Ile	
450 455 460	

tat ttc aca cag ggc gat att agg gaa gta cag ctt gca aaa ggg gca 1440

Tyr Phe Thr Gln Gly Asp Ile Arg Glu Val Gln Leu Ala Lys Gly Ala

465 470 475 480

ata tat gca gga ata aaa att ctc tgt tat gaa tat gga ata agt atg 1488

Ile Tyr Ala Gly Ile Lys Ile Leu Cys Tyr Glu Tyr Gly Ile Ser Met

485 490 495

gaa gat ata tct aat gta tat gtt act gga gca ttt gga tgt cat atc 1536

Glu Asp Ile Ser Asn Val Tyr Val Thr Gly Ala Phe Gly Cys His Ile

500 505 510

gat gtt gaa aat gca aag att atc gga ctt tta ccg gat ttg gat aat 1584

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

515 520 525

ata ttg agt att gat aat gct gct gga agg ggg act ata atg gct tta 1632

Ile Leu Ser Ile Asp Asn Ala Ala Gly Arg Gly Thr Ile Met Ala Leu

530 535 540

cta tct aaa aaa att aga aat gaa gcc gat aag ttg gca aaa aat acg 1680

Leu Ser Lys Lys Ile Arg Asn Glu Ala Asp Lys Leu Ala Lys Asn Thr

545 550 555 560

aaa tat att gaa tta agt agt cat gat aat ttt gaa agt gag ttc ata 1728

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

565 570 575

tct gcc ctt ggg ttt taa 1746

Ser Ala Leu Gly Phe

580

<210> 285

<211> 581

<212> PRT

<213> Methanococcus aeolicus

<400> 285

Met Gly Gly Val Met Met Tyr Asn Ile Thr Tyr Ile Lys Glu Asp Gly

1 5 10 15

Thr Lys Lys Ser Ile Lys Val Lys Glu Gly Thr Thr Ile Leu Glu Gly
20 25 30
Ala Ile Lys Ala Gly Val Tyr Ile Asp Ala Pro Cys Gly Thr Gly Lys
35 40 45
Cys Gly Lys Cys Lys Val Leu Val Glu Lys Gly Leu Glu Asn Ile Asp
50 55 60
Lys Asp Ser Ile Val Glu Asp Glu Tyr Ala Leu Ala Cys Val Ala Lys
65 70 75 80
Val Tyr Gly Asp Ile Ser Ile Asn Val Pro Asn Phe Gln Gly Val Val
85 90 95
Cys Lys Asp Ile Thr Asn Glu Val Gly Glu Leu Gln Thr Arg Arg Val
100 105 110
Cys Ser Ile Thr Glu Gln Cys Lys Gly Glu Leu Gln Asn Leu Glu Gly
115 120 125
Phe His Pro Met Ser Leu Asn Pro Asp Ile Gly Ile Asn Lys Ile Thr
130 135 140
Thr Thr Val Leu Glu Ser Ser Asn Tyr Asn Leu Thr Leu Asp Ala Ile
145 150 155 160
Asn Lys Leu Asn Ser Met Lys Leu Ser Asp Glu Val Thr Leu Ile Leu
165 170 175
Lys Gly Asp Asn Val Val Asn Val Glu Lys Asp Phe Ser Gly Ile Tyr
180 185 190
Gly Leu Ser Ile Asp Ile Gly Thr Thr Ser Val Val Val Tyr Leu Val
195 200 205
Asp Ile Ser Lys Gly Ile Val Leu Asp Asn Ile Ser Phe Leu Asn Pro
210 215 220
Gln Arg Gln Phe Gly Ala Asp Val Val Ser Arg Ile Ala Tyr Asn Asn
225 230 235 240
Gly Ile Leu Leu Gln Lys Thr Leu Ile Thr Glu Leu Asn Asp Ser Ile
245 250 255
Ser Lys Leu Cys Ser Asn Asn Asn Ile Lys Met Asp Asn Ile Tyr Glu

260 265 270
Val Ser Val Val Gly Asn Thr Ala Met Ile His Phe Phe Tyr Gly Ile

275 280 285
Val Pro Lys Asn Leu Ala Thr His Pro Tyr Val Pro Thr Phe Lys Asn

290 295 300
Ser Pro Tyr Leu Pro Ala Lys Glu Leu Gly Leu Asn Leu Arg Asn Ala

305 310 315 320
Tyr Ile Tyr Thr Leu Pro Ile Ile Gly Gly Tyr Val Gly Ala Asp Thr

325 330 335
Val Gly Ala Ile Leu Ser Ser Glu Met His Lys Lys Asp Asp Ile Ser

340 345 350
Leu Leu Ile Asp Ile Gly Thr Asn Gly Glu Ile Val Leu Gly Asn Lys

355 360 365
Glu Lys Leu Leu Thr Cys Ser Cys Ala Ala Gly Pro Ala Phe Glu Gly

370 375 380
Val Ser Ile Glu His Gly Thr Asn Ala Arg Glu Gly Ala Val Cys Arg

385 390 395 400
Val Lys Ile Asp Glu Asn Asn Ile Tyr Tyr Glu Thr Ile Gly Asn Lys

405 410 415
Thr Pro Pro Ile Gly Ile Cys Gly Ser Gly Ile Ile Asp Ile Val Ala

420 425 430
Glu Phe Leu Lys Ser Gly Leu Ile Asn Lys Thr Gly Arg Phe Thr Gly

435 440 445
Glu His Lys Asn Leu Lys Glu Asn Lys Phe Ile Ile Glu Asp Ser Ile

450 455 460
Tyr Phe Thr Gln Gly Asp Ile Arg Glu Val Gln Leu Ala Lys Gly Ala

465 470 475 480
Ile Tyr Ala Gly Ile Lys Ile Leu Cys Tyr Glu Tyr Gly Ile Ser Met

485 490 495
Glu Asp Ile Ser Asn Val Tyr Val Thr Gly Ala Phe Gly Cys His Ile

500 505 510

Asp Val Glu Asn Ala Lys Ile Ile Gly Leu Leu Pro Asp Leu Asp Asn
515 520 525

Ile Leu Ser Ile Asp Asn Ala Ala Gly Arg Gly Thr Ile Met Ala Leu
530 535 540

Leu Ser Lys Lys Ile Arg Asn Glu Ala Asp Lys Leu Ala Lys Asn Thr

545 550 555 560
Lys Tyr Ile Glu Leu Ser Ser His Asp Asn Phe Glu Ser Glu Phe Ile
565 570 575

Ser Ala Leu Gly Phe
580

<210> 286

<211> 720

<212> DNA

<213> Synthetic

<220><221> CDS

<222> (1)..(720)

<223> Synthetic gpf gene encoding GFP

<400> 286

atg cgt aaa ggc gaa gag ctg ttc act ggt gtc gtc cct att ctg gtg 48
Met Arg Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val

1 5 10 15
gaa ctg gat ggt gat gtc aac ggt cat aag ttt tcc gtg cgt ggc gag 96
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly Glu
20 25 30

ggt gaa ggt gac gca act aat ggt aaa ctg acg ctg aag ttc atc tgt 144
Gly Glu Gly Asp Ala Thr Asn Gly Lys Leu Thr Leu Lys Phe Ile Cys
35 40 45

act act ggt aaa ctg ccg gta cct tgg ccg act ctg gta acg acg ctg 192
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
50 55 60

act tat ggt gtt cag tgc ttt gct cgt tat ccg gac cat atg aag cag 240
Thr Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys Gln

65	70	75	80	
cat gac ttc ttc aag tcc gcc atg ccg gaa ggc tat gtg cag gaa cgc				288
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg				
85	90	95		
acg att tcc ttt aag gat gac ggc acg tac aaa acg cgt gcg gaa gtg				336
Thr Ile Ser Phe Lys Asp Asp Gly Thr Tyr Lys Thr Arg Ala Glu Val				
100	105	110		
aaa ttt gaa ggc gat acc ctg gta aac cgc att gag ctg aaa ggc att				384
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile				
115	120	125		
gac ttt aaa gaa gac ggc aat atc ctg ggc cat aag ctg gaa tac aat				432
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn				
130	135	140		
ttt aac agc cac aat gtt tac atc acc gcc gat aaa caa aaa aat ggc				480
Phe Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly				
145	150	155	160	
att aaa gcg aat ttt aaa att cgc cac aac gtg gag gat ggc agc gtg				528
Ile Lys Ala Asn Phe Lys Ile Arg His Asn Val Glu Asp Gly Ser Val				
165	170	175		
cag ctg gct gat cac tac cag caa aac act cca atc ggt gat ggt cct				576
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro				
180	185	190		
gtt ctg ctg cca gac aat cac tat ctg agc acg caa agc gtt ctg tct				624
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Val Leu Ser				
195	200	205		
aaa gat ccg aac gag aaa cgc gat cat atg gtt ctg ctg gag ttc gta				672
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val				
210	215	220		
acc gca gcg ggc atc acg cat ggt atg gat gaa ctg tac aaa tga tga				720

Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys

225 230 235

<210> 287

<211> 238

<212> PRT

<213> Synthetic

<400> 287

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

1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly Glu

20 25 30

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

35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu

50 55 60

Thr Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys Gln

65 70 75 80

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

85 90 95

Thr Ile Ser Phe Lys Asp Asp Gly Thr Tyr Lys Thr Arg Ala Glu Val

100 105 110

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

115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn

130 135 140

Phe Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly

145 150 155 160

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

165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro

180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Val Leu Ser

195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val

210 215 220

Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys

225 230 235

<210> 288

<211> 1486

<212> DNA

<213> Escherichia coli

<220><221> promoter

<222> (1)..(36)

<223> apFAB306 promoter

<220><221> CDS

<222> (73)..(603)

<223> fldA gene encoding flavodoxin

<220><221> CDS

<222> (620)..(1366)

<223> flpr gene encoding flavodoxin/ferredoxin reductase

<220><221> terminator

<222> (1400)..(1486)

<400> 288

ttgacaatta atcatccggc tcgtagtggt tgtggatggc agtggctagc ggccgcagag 60

gttatttcac tc atg gct atc act ggc atc ttt ttc ggc agc gac acc ggt 111

Met Ala Ile Thr Gly Ile Phe Phe Gly Ser Asp Thr Gly

1 5 10

aat acc gaa aat atc gca aaa atg att caa aaa cag ctt ggt aaa gac 159

Asn Thr Glu Asn Ile Ala Lys Met Ile Gln Lys Gln Leu Gly Lys Asp

15 20 25

gtt gcc gat gtc cat gac att gca aaa agc agc aaa gaa gat ctg gaa 207

Val Ala Asp Val His Asp Ile Ala Lys Ser Ser Lys Glu Asp Leu Glu

30 35 40 45

gct tat gac att ctg ctg ctg ggc atc cca acc tgg tat tac ggc gaa 255
Ala Tyr Asp Ile Leu Leu Leu Gly Ile Pro Thr Trp Tyr Tyr Gly Glu
50 55 60

gcg cag tgt gac tgg gat gac ttc ttc ccg act ctc gaa gag att gat 303
Ala Gln Cys Asp Trp Asp Asp Phe Phe Pro Thr Leu Glu Glu Ile Asp
65 70 75

ttc aac ggc aaa ctg gtt gcg ctg ttt ggt tgt ggt gac cag gaa gat 351
Phe Asn Gly Lys Leu Val Ala Leu Phe Gly Cys Gly Asp Gln Glu Asp
80 85 90

tac gcc gaa tat ttc tgc gac gca ttg ggc acc atc cgc gac atc att 399
Tyr Ala Glu Tyr Phe Cys Asp Ala Leu Gly Thr Ile Arg Asp Ile Ile
95 100 105

gaa ccg cgc ggt gca acc atc gtt ggt cac tgg cca act gcg ggc tat 447

Glu Pro Arg Gly Ala Thr Ile Val Gly His Trp Pro Thr Ala Gly Tyr
110 115 120 125

cat ttc gaa gca tca aaa ggt ctg gca gat gac gac cac ttt gtc ggt 495
His Phe Glu Ala Ser Lys Gly Leu Ala Asp Asp Asp His Phe Val Gly
130 135 140

ctg gct atc gac gaa gac cgt cag ccg gaa ctg acc gct gaa cgt gta 543
Leu Ala Ile Asp Glu Asp Arg Gln Pro Glu Leu Thr Ala Glu Arg Val
145 150 155

gaa aaa tgg gtt aaa cag att tct gaa gag ttg cat ctc gac gaa att 591
Glu Lys Trp Val Lys Gln Ile Ser Glu Glu Leu His Leu Asp Glu Ile
160 165 170

ctc aat gcc tga aaaacaggag aaaaac atg gct gat tgg gta aca ggc aaa 643
Leu Asn Ala Met Ala Asp Trp Val Thr Gly Lys
175 180

gtc act aaa gtg cag aac tgg acc gac gcc ctg ttt agt ctc acc gtt 691
Val Thr Lys Val Gln Asn Trp Thr Asp Ala Leu Phe Ser Leu Thr Val
185 190 195 200

cac gcc ccc gtg ctt ccg ttt acc gcc ggg caa ttt acc aag ctt ggc 739

His Ala Pro Val Leu Pro Phe Thr Ala Gly Gln Phe Thr Lys Leu Gly	
205 210 215	
ctt gaa atc gac ggc gaa cgc gtc cag cgc gcc tac tcc tat gta aac	787
Leu Glu Ile Asp Gly Glu Arg Val Gln Arg Ala Tyr Ser Tyr Val Asn	
220 225 230	
tcg ccc gat aat ccc gat ctg gag ttt tac ctg gtc acc gtc ccc gat	835
Ser Pro Asp Asn Pro Asp Leu Glu Phe Tyr Leu Val Thr Val Pro Asp	
235 240 245	
ggc aaa tta agc cca cga ctg gcg gca ctg aaa cca ggc gat gaa gtg	883
Gly Lys Leu Ser Pro Arg Leu Ala Ala Leu Lys Pro Gly Asp Glu Val	
250 255 260	
cag gtg gtt agc gaa gcg gca gga ttc ttt gtg ctc gat gaa gtg ccg	931
Gln Val Val Ser Glu Ala Ala Gly Phe Phe Val Leu Asp Glu Val Pro	
265 270 275 280	
cac tgc gaa acg cta tgg atg ctg gca acc ggt aca gcg att ggc cct	979
His Cys Glu Thr Leu Trp Met Leu Ala Thr Gly Thr Ala Ile Gly Pro	
285 290 295	
tat tta tcg att ctg caa cta ggt aaa gat tta gat cgc ttc aaa aat	1027
Tyr Leu Ser Ile Leu Gln Leu Gly Lys Asp Leu Asp Arg Phe Lys Asn	
300 305 310	
ctg gtc ctg gtg cac gcc gca cgt tat gcc gcc gac tta agc tat ttg	1075
Leu Val Leu Val His Ala Ala Arg Tyr Ala Ala Asp Leu Ser Tyr Leu	
315 320 325	
cca ctg atg cag gaa ctg gaa aaa cgc tac gaa gga aaa ctg cgc att	1123
Pro Leu Met Gln Glu Leu Glu Lys Arg Tyr Glu Gly Lys Leu Arg Ile	
330 335 340	
cag acg gtg gtc agt cgg gaa acg gca gcg ggg tcg ctc acc gga cgg	1171
Gln Thr Val Val Ser Arg Glu Thr Ala Ala Gly Ser Leu Thr Gly Arg	
345 350 355 360	
ata ccg gca tta att gaa agt ggg gaa ctg gaa agc acg att ggc ctg	1219

Ile Pro Ala Leu Ile Glu Ser Gly Glu Leu Glu Ser Thr Ile Gly Leu

365 370 375
ccg atg aat aaa gaa acc agc cat gtg atg ctg tgc ggc aat cca cag 1267

Pro Met Asn Lys Glu Thr Ser His Val Met Leu Cys Gly Asn Pro Gln

380 385 390
atg gtg cgc gat aca caa cag ttg ctg aaa gag acc cgg cag atg acg 1315

Met Val Arg Asp Thr Gln Gln Leu Leu Lys Glu Thr Arg Gln Met Thr

395 400 405
aaa cat tta cgt cgc cga ccg ggc cat atg aca gcg gag cat tac tgg 1363

Lys His Leu Arg Arg Arg Pro Gly His Met Thr Ala Glu His Tyr Trp

410 415 420
taa tagcttcata tgggccacag gacactcggt gctttcacca tgcgtaaagc 1416

aatcagatac ccagcccgcc taatgagcgg gctttttttt gaacaaaatt agagaataac 1476

aatgcaaaca 1486

<210> 289

<211> 176

<212> PRT

<213> Escherichia coli

<400> 289

Met Ala Ile Thr Gly Ile Phe Phe Gly Ser Asp Thr Gly Asn Thr Glu

1 5 10 15

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

20 25 30

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

35 40 45

Ile Leu Leu Leu Gly Ile Pro Thr Trp Tyr Tyr Gly Glu Ala Gln Cys

50 55 60

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

65 70 75 80

Lys Leu Val Ala Leu Phe Gly Cys Gly Asp Gln Glu Asp Tyr Ala Glu

85 90 95

Tyr Phe Cys Asp Ala Leu Gly Thr Ile Arg Asp Ile Ile Glu Pro Arg
 100 105 110
 Gly Ala Thr Ile Val Gly His Trp Pro Thr Ala Gly Tyr His Phe Glu
 115 120 125

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

 <210> 290
 <211> 248
 <212> PRT
 <213> Escherichia coli
 <400> 290

 Met Ala Asp Trp Val Thr Gly Lys Val Thr Lys Val Gln Asn Trp Thr

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

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

Lys Asp Leu Asp Arg Phe Lys Asn Leu Val Leu Val His Ala Ala Arg

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

195 200 205
Val Met Leu Cys Gly Asn Pro Gln Met Val Arg Asp Thr Gln Gln Leu
210 215 220
Leu Lys Glu Thr Arg Gln Met Thr Lys His Leu Arg Arg Arg Pro Gly
225 230 235 240
His Met Thr Ala Glu His Tyr Trp

245

<210> 291

<211> 36

<212> DNA

<213> Synthetic

<220><221> promoter

<222> (1)..(36)

<223> apFAB309 promoter

<400> 291

ttgacaatta atcatccggc tcgtagtgtc tgtgga

36

<210> 292

<211> 87

<212> DNA

<213> Synthetic

<220><221> terminator

<222> (1)..(87)

<223> apFAB378 terminator

<400> 292

ttcaccatgc gtaaagcaat cagataccca gcccgccctaa tgagcgggct tttttttgaa 60

caaaattaga gaataacaat gcaaaca 87