(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년07월15일
(11) 등록번호 10-2000383
(24) 등록일자 2019년07월09일

(51) 국제특허분류(Int. Cl.)
C12Q 1/34 (2006.01) *C07K 14/00* (2006.01)
C12N 15/80 (2006.01) *C12N 9/00* (2006.01)
C12N 9/24 (2006.01)
(52) CPC특허분류
C12Q 1/34 (2013.01)
C07K 14/00 (2013.01)
(21) 출원번호 10-2017-7028666(분할)
(22) 출원일자(국제) 2010년09월29일
심사청구일자 2017년11월08일
(85) 번역문제출일자 2017년10월11일
(65) 공개번호 10-2017-0116254
(43) 공개일자 2017년10월18일
(62) 원출원 특허 10-2012-7011254
원출원일자(국제) 2010년09월29일
심사청구일자 2015년07월10일
(86) 국제출원번호 PCT/IB2010/002589
(87) 국제공개번호 WO 2011/039634
국제공개일자 2011년04월07일
(30) 우선권주장
61/246,847 2009년09월29일 미국(US)
(56) 선행기술조사문헌
JP2002369692 A*
Merkle 등, Biochinica et Biophysica Acta. 1997, 1336, 132-146
Chiba 등, Glycobiology. 2002, 12, 12, 821-828
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
유니버시테이트 겐트
벨기에 비-9000 겐트 신트-피에터스니유스트라트 25
브이아이비 브이지더블유
벨기에 비-9052 겐트 리비스셰스트라트 120
(뒷면에 계속)
(72) 발명자
캘러워트, 니코, 루크, 마크
벨기에 비-9850 한스베케-네베레 베지노플란 15
베르베켄, 와우터
벨기에 9860 랜드스카우터 스투이벤버그 35
(뒷면에 계속)
(74) 대리인
특허법인 무한

전체 청구항 수 : 총 57 항 심사관 : 이예리

(54) 발명의 명칭 **만노스-1-포스포-6-만노스 결합의 포스포-6-만노스로의 가수분해**

(57) 요 약

올리고당에서의 만노스-6-인산 잔기(mannose-6-phosphate residue)의 캡핑제거 (uncapping)에 유용한 방법 및 유전적으로 조작된 세포(genetically engineered cells)를 본원에서 기재하였다.

(52) CPC특허분류
C12N 15/80 (2013.01)
C12N 9/00 (2013.01)
C12N 9/24 (2013.01)
C12N 9/2488 (2013.01)
(73) 특허권자
옥시레인 유케이 리미티드
영국 엠1 2에이피 맨체스터 피카딜리 49 피카딜리 하우스 수트 313
브리제 유니버시타이트 브루셀
벨기에, 비-1050 플루셀, 블레인란 2
(72) 발명자
티엘스, 페트라, 소피
벨기에 비-9050 젠트-젠트브루그 줄스 드 상트-게노이스트라트 54
리마우트, 한, 카렐
벨기에 비-3370 부터셈 카스틸스트라트 10에이
피엔스, 캐슬린, 카밀라, 텔레스포르, 알리다, 마리아
벨기에 비-9000 젠트 시타델란 54

명 세 서

청구범위

청구항 1

올리고당(oligosaccharide)에서의 만노스-6-인산 잔기(mannose-6-phosphate residue)의 캡핑제거(uncapping) 방법으로서, 상기 방법은

a) 만노스-1-포스포-6-만노스 잔기(mannose-1-phospho-6-mannose residue)를 갖는 올리고당을 제공하는 단계; 및

b) 상기 만노스-1-포스포-6-만노스 잔기를 포스포-6-만노스로 가수분해시킬 수 있는 만노시다제(mannosidase)와 상기 올리고당을 접촉시키는 단계로서, 상기 만노시다제는 글리코실 히드롤라제 패밀리 92(glycosyl hydrolase family 92)의 멤버(member)이고, 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않으며, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 95% 동일성(identity)을 갖는 아미노산 서열을 포함하는 것이고, 단, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않는 것인, 단계를 포함하는, 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

제1항에 있어서, 상기 만노시다제는, SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 98% 동일성(identity)을 갖는 아미노산 서열을 포함하는 것이고, 단, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않는 것인, 방법.

청구항 13

제1항에 있어서, 상기 만노시다제는 하기를 갖는 아미노산 서열을 포함하는 것인 방법:

(i) GVGXXGXGG 모티프(motif)(이 식에서, X 는 Gly, Ala, Ser, Thr, 또는 Cys 이다);

(ii) VRXE 모티프(이 식에서, X 는 Pro 외의 모든 아미노산이다);

(iii) X_1YQGX_2 모티프(이 식에서, X_1 은 Leu, Ile, Val, Ala, Phe, Tyr 또는 Met 이고, X_2 은 Thr, Ser, 또는 Asn 이다); 또는

(iv) GDXGN(이 식에서, X 는 Pro 외의 임의의 아미노산일 수 있다).

청구항 14

제1항, 제12 및 제13항 중 어느 한 항에 있어서,

상기 접촉 단계(contacting step)는 정제된 만노시다제(purified mannosidase), 재조합 만노시다제(recombinant mannosidase), 상기 재조합 만노시다제를 포함하는 세포 용해액(cell lysate), 또는 상기 재조합 만노시다제를 포함하는 균류 세포(fungal cell)를 사용하여 실행하는 방법.

청구항 15

제1항, 제12항 및 제13항 중 어느 한 항에 있어서,

상기 올리고당은 단백질에 부착된(attached) 것인, 방법.

청구항 16

제15항에 있어서,

상기 단백질은 균류 유기체(fungal organism)에서 발현된 인간 단백질(human protein)인 것인, 방법.

청구항 17

제16항에 있어서,

상기 균류 유기체는 *야로위아 리폴리티카(Yarrowia lipolytica)* 또는 *아르술라 아데니니보란스(Arxula adeninivorans)*인 방법.

청구항 18

제16항에 있어서,

상기 균류 유기체는 메탄올자화효모(methylotrophic yeast)인 방법.

청구항 19

제18항에 있어서,

상기 메탄올자화효모는 *피치아 파스토리스(Pichia pastoris), 피치아 메탄올리카(Pichia methanolica), 오가테아 미뉴타(Oogataea minuta)* 또는 *한세눌라 폴리모르파(Hansenula polymorpha)*인 방법.

청구항 20

제16항에 있어서,

상기 균류 유기체는 사상 균류(filamentous fungus)인 방법.

청구항 21

제20항에 있어서,

상기 사상 균류는 *아스페르길루스 카에시엘루스(Aspergillus caesiellus), 아스페르길루스 칸디두스(Aspergillus candidus), 아스페르길루스 카르네우스(Aspergillus carneus), 아스페르길루스 클라바투수(Aspergillus clavatus), 아스페르길루스 데플렉투스(Aspergillus deflectus), 아스페르길루스 플래브스(Aspergillus flavus), 아스페르길루스 푸미가투스(Aspergillus fumigates), 아스페르길루스 글라우크스(Aspergillus glaucus), 아스페르길루스 니둘란스(Aspergillus nidulans), 아스페르길루스 니게르(Aspergillus niger), 아스페르길루스 오크라세우스(Aspergillus ochraceus), 아스페르길루스 오라자에(Aspergillus oryzae), 아스페르길루스 파라시티커스(Aspergillus parasiticus), 아스페르길루스 페니실로이데스(Aspergillus penicilloides), 아스페르길루스 레스트릭투스(Aspergillus restrictus), 아스페르길루스 소예(Aspergillus sojae), 아스페르길루스 시도위(Aspergillus sydowi), 아스페르길루스 타마리(Aspergillus tamari), 아스페르길루스 테레우스(Aspergillus terreus), 아스페르길루스 우스투스(Aspergillus ustus),* 및 *아스페르길루스 베르시컬러(Aspergillus versicolor)*로 이루어진 군으로부터 선택된 방법.

청구항 22

제15항에 있어서,

상기 단백질은 병원체 단백질(pathogen protein), 리소좀 단백질(lysosomal protein), 성장 인자(growth factor), 사이토카인(cytokine), 케모카인(chemokine), 항체 또는 이의 항원-결합 단편(antigen-binding fragment), 또는 융합 단백질(fusion protein)인 방법.

청구항 23

제22항에 있어서,

상기 리소좀 단백질은 리소좀 효소(lysosomal enzyme)인 방법.

청구항 24

제23항에 있어서,

상기 리소좀 효소는 리소좀 축적 질환(lysosomal storage disorder, LSD)과 관련된 방법.

청구항 25

제24항에 있어서,

상기 LSD 는 파브리병(Fabry's disease), 점액다당류증 I(mucopolysaccharidosis I), 파버 질병(Farber disease), 고세병(Gaucher disease), GM1-강글리오시드증(GM1-gangliosidosis), 테이-샥스병(Tay-Sachs disease), 샌드호프병(Sandhoff disease), GM2 활성제 질병(GM2 activator disease), 크라베병(Krabbe disease), 이염성백질이영양증(metachromatic leukodystrophy), 니만-피크병(Niemann-Pick disease), 샤이에 질병(Scheie disease), 헌터 질병(Hunter disease), 산필립포 질병(Sanfilippo disease), 모르키오병(Morquio disease), 마로토-라미 질병(Maroteaux-Lamy disease), 히알루로니다아제 결핍증 (hyaluronidase deficiency), 아스파르틸글루코스아민뇨증 (aspartylglucosaminuria), 푸코시드축적증(fucosidosis), 만노시도시스(mannosidosis), 쉰들러 병(Schindler disease), 사이알산축적증 유형 1(sialidosis type 1), 폼피병(Pompe disease), 피크노디소토시스(Pycnodysostosis), 세로이드 리포푸신증(ceroid lipofuscinosis), 콜레스테롤 에스테르 축적 질병(cholesterol ester storage disease), 월만병(Wolman disease), 다종 술파타아제 결손증(Multiple sulfatase deficiency), 갈락토시알리도시스(galactosialidosis), 뮤코리피드증(mucolipidosis), 시스틴축적증(cystinosis), 시알산 축적 질병(sialic acid storage disorder), 마리네스코-쉐글렌 증후군(Marinesco- Sjögren syndrome)을 갖는 킬로미크론 보유 질병(chylomicron retention disease), 헤르만스키-푸드락 증후군(Hermansky-Pudlak syndrome), 체디아크-히가시 증후군(Chediak-Higashi syndrome), 다논병(Danon disease), 또는 겔레오피직 이형성증(Geleophysic dysplasia)인 방법.

청구항 26

제24항에 있어서,

상기 LSD 는 폼피병(Pompe disease) 또는 파브리병(Fabry's disease)인 방법.

청구항 27

제1항, 제12항 및 제13항 중 어느 한 항에 있어서,

상기 만노시다제는 타겟팅 서열(targeting sequence)을 포함하는 방법.

청구항 28

캡핑이 제거된(uncapped) 포스포-6-만노스 잔기를 갖는 타겟 단백질(target protein)을 생산하는 방법으로서, 상기 방법은,

만노스-1-포스포-6-만노스 잔기를 포스포-6-만노스로 가수분해할 수 있는 만노시다제를 코드화하는(encoding) 핵산을 포함하고 발현되도록(express) 유전적으로 조작된 균류 세포를 제공하는 단계로서, 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버이고, 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않으며, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 95% 동일성(identity)을 갖는 아미노산 서열을 포함하는 것이고, 단, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않는 것인, 단계; 및

타겟 단백질을 코드화하는 핵산을 상기 세포 내로 도입하는(introducing) 단계로서, 상기 세포는 상기 타겟 단백질을 생산하고, 타겟 단백질 상의 만노스-1-포스포-6-만노스 잔기는 세포에서 발현된 상기 발현된 만노시다제에 의해 포스포-6-만노스 잔기로 전환된 것인, 단계;를 포함하는 방법.

청구항 29

제28항에 있어서, 상기 만노시다제는 제12항 또는 제13항에 기재된 것인, 방법.

청구항 30

제28항에 있어서, 상기 균류 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하는 것인, 방법.

청구항 31

제28항에 있어서, 상기 균류 세포는 OCH1[외부 사슬 연장(Outer Chain elongation)] 활성도가 결핍되도록 유전적으로 조작된 것인, 방법.

청구항 32

제28항에 있어서, 상기 만노시다제는 *C. 셀룰란스(C. cellulans)* 만노시다제인, 방법.

청구항 33

다음의 단계를 포함하는, 균류 유기체에서 캡핑이 제거된(uncapped) 포스포-6-만노스 잔기를 갖는 타겟 단백질을 생산하는 방법:

a) 만노스-1-포스포-6-만노스를 포스포-6-만노스로 가수분해할 수 있는 만노시다제를 코드화하는 핵산을 포함하고 발현되도록 유전적으로 조작된 균류 세포를 제공하는 단계로서, 여기서, 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버이고 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않으며, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 95% 동일성(identity)을 갖는 아미노산 서열을 포함하는 것이고, 단, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않으며, 상기 균류 세포는 타겟 단백질을 코드화하는 핵산을 추가적으로 포함하는데, 이때 상기 타겟 단백질 상의 만노스-1-포스포-6-만노스 잔기는 세포에서 발현된 상기 발현된 만노시다제에 의해 포스포-6-만노스 잔기로 전환되어 있는 것인, 단계; 및

b) 상기 캡핑이 제거된 포스포-6-만노스 잔기를 갖는 상기 타겟 단백질을 분리하는 단계.

청구항 34

제33항에 있어서, 상기 만노시다제는 제12항 또는 제13항에 기재된 것인, 방법.

청구항 35

제33항에 있어서, 상기 타겟 단백질과 만노시다제는 상기 세포에 의해 분비되는 것인, 방법.

청구항 36

캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 당단백질(glycoproteins)을 생산하도록 유전적으로 조작된 분리된 균류 세포(isolated fungal cell)로서, 상기 균류 세포는 만노시다제를 코드화하는 핵산을 포함하고, 상기 균류 세포에서의 상기 만노시다제의 발현(expression)은 상기 포스포-6-만노스 잔기를 포함하는 당단백질을 결과적으로 생산하고, 여기서 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버이고 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않으며, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 95% 동일성(identity)을 갖는 아미노산 서열을 포함하는 것이고, 단, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않는 것인, 분리된 균류 세포.

청구항 37

제36항에 있어서, 상기 만노시다제는 제12항 또는 제13항에 기재된 것과 같은 것인, 분리된 균류 세포.

청구항 38

제36항에 있어서, 상기 균류 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하는, 분리된 균류 세포.

청구항 39

제36항에 있어서, 상기 균류 세포는 OCH1[외부 사슬 연장(Outer CHain elongation)] 활성도가 결핍되도록 유전적으로 조작된, 분리된 균류 세포.

청구항 40

제36항에 있어서, 상기 균류 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하고, 여기서 상기 균류 세포는 OCH1 활성도가 결핍되도록 유전적으로 조작된, 분리된 균류 세포.

청구항 41

제36항에 있어서,

상기 균류 세포는 당단백질인 타겟 단백질을 코드화하는 핵산을 추가적으로 포함하는, 분리된 균류 세포.

청구항 42

제41항에 있어서,

상기 타겟 단백질은 인간 단백질인, 분리된 균류 세포.

청구항 43

제41항에 있어서,

상기 타겟 단백질은 병원체 단백질(pathogen protein), 리소좀 단백질(lysosomal protein), 성장 인자(growth factor), 사이토카인(cytokine), 케모카인(chemokine), 항체 또는 이의 항원-결합 단편(antigen-binding fragment), 또는 융합 단백질(fusion protein)인, 분리된 균류 세포.

청구항 44

제43항에 있어서,

상기 리소좀 단백질은 리소좀 효소인, 분리된 균류 세포.

청구항 45

제44항에 있어서,

상기 리소좀 효소는 산성의 알파 글루코시다아제(acid alpha glucosidase) 또는 알파 갈락토시다아제(alpha galactosidase)인, 분리된 균류 세포.

청구항 46

제41항에 있어서,

상기 타겟 단백질은 리소좀 축적 질환(lysosomal storage disorder, LSD)과 관련된 단백질인, 분리된 균류 세포.

청구항 47

제46항에 있어서,

상기 LSD 는 파브리병(Fabry's disease), 점액다당류증 I(mucopolysaccharidosis I), 파버 질병(Farber disease), 고세병(Gaucher disease), GM1-강글리오시드증(GM1-gangliosidosis), 테이-삭스병(Tay-Sachs disease), 샌드호프병(Sandhoff disease), GM2 활성제 질병(GM2 activator disease), 크라베병(Krabbe disease), 이염성백질이영양증(metachromatic leukodystrophy), 니만-피크병(Niemann-Pick disease), 샤이에 질병(Scheie disease), 헌터 질병(Hunter disease), 산필립포 질병(Sanfilippo disease), 모르키오병(Morquio disease), 마로토-라미 질병(Maroteaux-Lamy disease), 히알루로니다아제 결핍증 (hyaluronidase deficiency), 아스파르틸글루코스아민뇨증 (aspartylglucosaminuria), 푸코시드축적증(fucosidosis), 만노시도시스(mannosidosis), 쉰들러 병(Schindler disease), 사이알산축적증 유형 1(sialidosis type 1), 폼피병(Pompe disease), 피크노디소토시스(Pycnodysostosis), 세로이드 리포푸신증(ceroid lipofuscinosis), 콜레스테롤 에스테르 축적 질병(cholesterol ester storage disease), 월만병(Wolman disease), 다종 술파타아제 결손증(Multiple sulfatase deficiency), 갈락토시알리도시스(galactosialidosis), 뮤코리피드증(mucolipidosis), 시스틴축적증(cystinosis), 시알산 축적 질병(sialic acid storage disorder), 마리네스코-쉐글렌 증후군(Marinesco-Sjögren syndrome)을 갖는 킬로미크론 보유 질병(chylomicron retention disease), 헤르만스키-푸드락 증후군(Hermansky-Pudlak syndrome), 체디아크-히가시 증후군(Chediak-Higashi syndrome), 다논병(Danon disease), 또는 겔레오피직 이형성증(Geleophysic dysplasia)인, 분리된 균류 세포.

청구항 48

제36항에 있어서, 상기 균류 세포(fungal cell)는 *야로위아 리폴리티카(Yarrowia lipolytica)* 또는 *아르술라 아데닌니보란스(Arxula adeninivorans)* 세포인, 분리된 균류 세포.

청구항 49

제38항에 있어서, 만노실 인산화반응을 촉진시킬 수 있는 상기 폴리펩티드는 MNN4 폴리펩티드인, 분리된 균류 세포.

청구항 50

제49항에 있어서, 상기 MNN4 폴리펩티드는 *야로위아 리폴리티카, S.세레비시아(S. cerevisiae), 오가테아 미누타(Ogataea minuta), 피키아 파스토리스(Pichia pastoris)* 또는 *C. 알비칸스(C. albicans)* 폴리펩티드인, 분리된 균류 세포.

청구항 51

제38항에 있어서, 만노실 인산화반응을 촉진시킬 수 있는 상기 폴리펩티드는 *P. 파스토리스* PNO1 폴리펩티드인, 분리된 균류 세포.

청구항 52

제36항에 있어서, 상기 만노시다제는 *C. 셀룰란스(C. cellulans)* 만노시다제인, 분리된 균류 세포.

청구항 53

제36항에 있어서, 상기 만노시다제는 분비 신호(secretion signal)를 포함하는, 분리된 균류 세포.

청구항 54

제36항에 있어서, 상기 만노시다제는 세포내 구간(intracellular compartment)에 상기 만노시다제를 타겟하기 위한(target) 타겟팅 신호(targeting signal)를 포함하는, 분리된 균류 세포.

청구항 55

제36항에 있어서, 상기 만노시다제는 분비 신호 및 세포내 구간에 상기 만노시다제를 타겟하기 위한 타겟팅 신호를 포함하는, 분리된 균류 세포.

청구항 56

야로위아 리폴리티카(Yarrowia lipolytica), 피키아 파스토리스(Pichia pastoris), 한세눌라 폴리모르파(Hansenula polymorpha), 아르술라 아데니니보란스(Arxula adeninivorans), 피치아 메탄올리카(Pichia methanolica), 오가타아 미뉴타(Oogataea minuta), 또는 *아스페르길루스 니게르(Aspergillus niger)* 세포의 실질적인 순수 배양물(pure culture)로서, 이들 중 상당수는 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 당단백질을 생산하도록 유전적으로 조작되고, 상기 세포는 만노스-1-포스포-6-만노스를 포스포-6-만노스로 가수분해시킬 수 있는 만노시다제를 코드화하는 핵산을 포함하고, 여기서 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버이고 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않으며, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 95% 동일성(identity)을 갖는 아미노산 서열을 포함하는 것이고, 단, 상기 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않는 것인, 배양물.

청구항 57

제56항에 있어서, 상기 만노시다제는 제12항 또는 제13항에 기재된 바와 같은 것인, 배양물.

청구항 58

제56항에 있어서, 상기 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하는 배양물.

청구항 59

제56항에 있어서, 상기 세포는 OCH1 활성도가 결핍되도록 유전적으로 조작된 배양물.

청구항 60

삭제

청구항 61

SEQ ID NO: 50 또는 SEQ ID NO: 50 의 잔기 1 내지 774와 적어도 95%의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드로서, 상기 폴리펩티드는 만노스-1-포스포-6-만노스 잔기를 포스포-6-만노스로 가수분해할 수 있고, 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않는 것이며, 단 상기 폴리펩티드는 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 을 포함하지는 않는 것인, 폴리펩티드.

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

제61항에 있어서, 상기 아미노산 서열이 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 98% 동일성(identity)을 갖는 것인, 폴리펩티드.

청구항 66

제61항에 있어서, 상기 아미노산 서열이 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50을 포함하고, 추가로 SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50의 N-말단에 SEQ ID NO: 15 의 잔기 1 내지 15를 포함하는 것인, 폴리펩티드.

청구항 67

제61항, 제65항 및 제66항 중 어느 한 항에 있어서, 상기 폴리펩티드는 하기를 갖는 아미노산 서열을 포함하는 폴리펩티드:

(i) GVGXXGXGG 모티프(motif)(이 식에서, X 는 Gly, Ala, Ser, Thr, 또는 Cys 이다);

(ii) VRXE 모티프(이 식에서, X 는 Pro 외의 모든 아미노산이다);

(iii) X_1YQGX_2 모티프(이 식에서, X_1 은 Leu, Ile, Val, Ala, Phe, Tyr 또는 Met 이고, X_2 은 Thr, Ser, 또는 Asn 이다); 또는

(iv) GDXGN(이 식에서, X 는 Pro 외의 임의의 아미노산일 수 있다).

청구항 68

핵산으로서, 상기 핵산은,

(a) 제61항, 제65항 및 제66항 중 어느 한 항에 기재된 폴리펩티드를 코드화하는 뉴클레오티드 서열; 또는

(b) (a)에 기재된 뉴클레오티드 서열의 상보체(complement of the nucleotide sequence)를 포함하고, 단, 상기

뉴클레오티드 서열은 SEQ ID NO:50의 잔기 1 내지 774 또는 SEQ ID NO:50 을 코드화하지는 않는 것인, 핵산.

청구항 69

제68항에 기재된 핵산을 포함하는 발현 벡터로서, 상기 핵산은 (a)에 기재된 것을 포함하는, 발현 벡터.

청구항 70

제69항의 발현 벡터를 포함하는 숙주 세포.

청구항 71

제70항의 숙주 세포를 배양하고, 상기 배양물로부터 폴리펩티드를 분리하는 것을 포함하는, 폴리펩티드를 제조하는 방법.

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 전후 참조(CROSS-REFERENCE TO RELATED APPLICATIONS)**

[0002] 본 출원서(application)는 2009년 9월 29일에 출원된 미국 출원 번호(U.S. Application Serial) No. 61/246,847 에 대한 우선권(priority)을 주장한다. 상기 우선된 출원서(the prior application)의 내용은 이의 전체가 참고문헌으로서 포함된다.

[0004] 본 발명은 당단백질(glycoproteins)에서의 만노스-1-포스포-6-만노스 결합을 가수분해 방법, 및 특히 당단백질에서의 포스포-6-만노스 잔기의 캡을 제거하기 위해(to uncap the phospho-6-mannose residues on the glycoprotein), 만노스-1-포스포-6-만노스 결합을 가수분해하기 위해 만노시다제를 사용하는 방법에 관한 것이다.

배 경 기 술

[0006] 고성능의 발현 시스템(High performance expression systems)은 성장(development) 하에 현재 대부분의 생물약제학적 약품(biopharmaceuticals)[예를 들어, 재조합 단백질(recombinant proteins)]을 생산하기 위해 필요하다. 이러한 생물약제학적 약품 중의 많은 것들의 생물학적 활성도는 이들의 번역 후 변형(post-translational modification)[예를 들어, 인산화반응(phosphorylation) 또는 글리코실화반응(glycosylation)]에 의존한다. 효모-기반 발현 시스템(yeast-based expression system)은 미생물 유기체(microbial organism)의 유전적인 조작(genetic manipulation) 및 발효의 편이성(ease)과 단백질의 분비 및 변형을 위한 능력을 겸비하고 있다. 그러나, 효모 세포에서 생산된 재조합 당단백질은 주로 이종의 고-만노스(heterogeneous high-mannose) 및 하이퍼-만노스 글리칸 구조(hyper-mannose glycan structures)를 나타내고, 이는 단백질의 기능(protein function), 하류 공정(downstream processing), 및 특히 당단백질이 생물학적으로 중요한 역할을 하는, 후속의 치료상의 용도(subsequent therapeutic use)에 해로울 수 있다.

[0007] 미국 출원 번호 No. 12/062,469 는 이의 전체가 참고문헌으로서 포함된다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0009] 본 발명은 적어도 그 일부에서, 당단백질에서의 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 만노시다제의 발견(discovery)에 근거한다. 예를 들어, 만노시다제는 캡핑이 제거된 말단 만노스-6-인산 잔기(uncapped terminal mannose-6-phosphate residues)를 포함하는 당단백질을 수득하는데 사용될 수 있다. 이러한 당단백질을 수득하는 *생체 외(In vitro)* 및 *생체 내(in vivo)* 방법은 본원에 기재되어 있다. 유전적으로 조작된 세포(Genetically engineered cells)는 캡핑이 제거된 말단 만노스-6-인산 잔기를 갖는 타겟 분자(target molecules)를 생산하기 위한 방법에서 사용될 수 있다.

[0011] 하나의 양상에서, 본 문서는 올리고당(oligosaccharide)에서 만노스-6-인산 잔기의 캡핑을 제거하기 위한 방법을 특징으로 한다. 상기 방법은 만노스-1-포스포-6-만노스 결합을 갖는 올리고당을 제공하는 단계; 및 만노스-1-포스포-6-만노스 결합(mannose-1-phospho-6-mannose linkage)을 포스포-6-만노스로 가수분해할 수 있는 만노

시다제(mannosidase)와 상기 올리고당을 접촉시키는 단계를 포함한다. 상기 접촉 단계는 정제된 만노시다제(purified mannosidase), 재조합 만노시다제(recombinant mannosidase), 상기 재조합 만노시다제를 포함하는 세포 용해액(cell lysate), 또는 상기 재조합 만노시다제를 포함하는 균류 세포(fungal cell)를 사용하여 실행할 수 있다. 만노시다제는 타겟팅 신호(targeting sequence)를 포함할 수 있다. 올리고당은 단백질[예를 들어, 균류 유기체(fungal organism)에서 발현된 인간 단백질(human protein)]에 부착될 수 있다.

[0013] 또 다른 양상에서, 본 문서는 말단 포스포-6-만노스 잔기(terminal phospho-6-mannose residues)를 갖는 타겟 단백질(target protein)을 생산하는 방법을 특징으로 한다. 상기 방법은, 만노스-1-포스포-6-만노스 결합을 포스포-6-만노스로 가수분해할 수 있는 만노시다제를 코드화하는(encoding) 핵산을 포함하기 위해 유전적으로 조작된 균류 세포를 제공하는 단계; 및 타겟 단백질을 코드화하는 핵산을 세포 내로 도입시키는(introducing) 단계를 포함하고, 여기에서 상기 세포는 말단 포스포-6-만노스 잔기를 포함하는 타겟 단백질을 생산한다.

[0015] 본 문서는 균류 유기체(fungal organism)에서 말단 포스포-6-만노스 잔기를 갖는 타겟 단백질을 생산하기 위한 방법을 또한 특징으로 한다. 상기 방법은, 만노스-1-포스포-6-만노스 결합을 포스포-6-만노스로 가수분해할 수 있는 만노시다제를 코드화하는 핵산을 포함하기 위해 유전적으로 조작된, 타겟 단백질을 코드화하는 핵산을 추가적으로 포함하는 균류 세포를 제공하는 단계; 및 말단 포스포-6-만노스 잔기를 갖는 타겟 단백질을 분리하는 단계를 포함한다. 상기 균류 세포는 만노실 인산화반응(mannosyl phosphorylation)을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 포함할 수 있고/포함할 수 있거나, OCH1 활성도(OCH1 activity)가 결핍되게 유전적으로 조작될 수 있다.

[0017] 본 문서는, 말단 포스포-6-만노스 잔기를 포함하는 당단백질을 생산하기 위해 유전적으로 조작된 분리된 균류 세포(isolated fungal cell)를 또한 특징으로 한다. 상기 균류 세포는 만노시다제를 코드화하는 핵산을 포함하고, 여기에서 균류 세포에서 만노시다제의 발현(expression)은 말단 포스포-6-만노스 잔기를 포함하는 당단백질을 생산한다. 상기 균류 세포는 타겟 당단백질 단백질(target glycoprotein protein)을 코드화하는 핵산을 추가적으로 포함할 수 있다.

[0019] 또 다른 양상에서, 본 문서는 *야로위아 리폴리티카(Yarrowia lipolytica)*, 피키아 파스토리스(Pichia pastoris), 한세눌라 폴리모르파(Hansenula polymorpha), *아르술라 아데니니보란스(Arxula adeninivorans)*, *피치아 메탄올리카(Pichia methanolica)*, *오가태아 미뉴타(Oogataea minuta)*, 또는 *아스페르길루스 니게르(Aspergillus niger)* 세포의 실질적인 순수 배양물(substantially pure culture)로서, 이들 중 상당수는 말단 포스포-6-만노스 잔기를 포함하는 당단백질을 생산하기 위해 유전적으로 조작되었고, 상기 세포는 만노스-1-포스포-6-만노스 결합을 포스포-6-만노스로 가수분해시킬 수 있는 만노시다제를 코드화하는 핵산을 포함한다.

[0021] 본원에 기재된 실시형태 중 하나에서, 균류 유기체(fungal organism)는 야로위아 리폴리티카(Yarrowia lipolytica) 또는 *아르술라 아데니니보란스(Arxula adeninivorans)*일 수 있다. 균류 유기체는 *피치아 파스토리스(Pichia pastoris)*, *피치아 메탄올리카(Pichia methanolica)*, *오가태아 미뉴타(Oogataea minuta)* 또는 *한세눌라 폴리모르파(Hansenula polymorpha)*와 같은 메탄올자화효모(methylotrophic yeast)일 수 있다. 균류 유기체는, 사상 균류(filamentous fungus)[예를 들어, *아스페르길루스 카에시엘루스(Aspergillus caesiellus)*, *아스페르길루스 칸디두스(Aspergillus candidus)*, *아스페르길루스 카르네우스(Aspergillus carneus)*, *아스페르길루스 클라바투수(Aspergillus clavatus)*, *아스페르길루스 데플렉투스(Aspergillus deflectus)*, *아스페르길루스 플래브스(Aspergillus flavus)*, *아스페르길루스 푸미가투스(Aspergillus fumigates)*, *아스페르길루스 글라우크스(Aspergillus glaucus)*, *아스페르길루스 니둘란스(Aspergillus nidulans)*, *아스페르길루스 니게르(Aspergillus niger)*, *아스페르길루스 오크라세우스(Aspergillus ochraceus)*, *아스페르길루스 오라자에(Aspergillus oryzae)*, *아스페르길루스 파라시티커스(Aspergillus parasiticus)*, *아스페르길루스 페니실로이데스(Aspergillus penicilloides)*, *아스페르길루스 레스트릭투스(Aspergillus restrictus)*, *아스페르길루스 소예(Aspergillus sojae)*, *아스페르길루스 시도위(Aspergillus sydowi)*, *아스페르길루스 타마리(Aspergillus tamari)*, *아스페르길루스 테레우스(Aspergillus terreus)*, *아스페르길루스 우스투스(Aspergillus ustus)*, 및 *아스페르길루스 베르시컬러(Aspergillus versicolor)*로 이루어진 군으로부터 선택된 사상 균류]일 수 있다.

[0023] 본원에 기재된 실시형태 중 하나에서, 상기 단백질은, 병원체 단백질(pathogen protein), 리소좀 단백질(lysosomal protein), 성장 인자(growth factor), 사이토카인(cytokine), 케모카인(chemokine), 항체 또는 이의 항원-결합 단편(antigen-binding fragment), 또는 융합 단백질(fusion protein)일 수 있다. 상기 리소좀 단백질은 리소좀 효소[예를 들어, 산성 알파 글루코시다아제(acid alpha glucosidase) 또는 알파 갈락토시다아제(alpha galactosidase)와 같은 리소좀 축적 질환(lysosomal storage disorder, LSD)과 관련된 리소좀 효소]

일 수 있다. 상기 LSD는 파브리병(Fabry's disease), 점액다당류증 I(mucopolysaccharidosis I), 파버 질병(Farber disease), 고세병(Gaucher disease), GM1-강글리오시드증(GM1-gangliosidosis), 테이-삭스병(Tay-Sachs disease), 샌드호프병(Sandhoff disease), GM2 활성제 질병(GM2 activator disease), 크라베병(Krabbe disease), 이염성백질이영양증(metachromatic leukodystrophy), 니만-피크병(Niemann-Pick disease), 샤이에 질병(Scheie disease), 헌터 질병(Hunter disease), 산필립포 질병(Sanfilippo disease), 모르키오병(Morquio disease), 마로토-라미 질병(Maroteaux-Lamy disease), 히알루로니다아제 결핍증 (hyaluronidase deficiency), 아스파르틸글루코스아민뇨증 (aspartylglucosaminuria), 푸코시드축적증(fucosidosis), 만노시도시스(mannosidosis), 쉰들러 병(Schindler disease), 사이알산축적증 유형 1(sialidosis type 1), 폼피병(Pompe disease), 피크노디소토시스(Pycnodysostosis), 세로이드 리포푸신증(ceroid lipofuscinosis), 콜레스테롤 에스테르 축적 질병(cholesterol ester storage disease), 월만병(Wolman disease), 다종 술파타아제 결손증(Multiple sulfatase deficiency), 갈락토시알리도시스(galactosialidosis), 뮤코리피드증(mucolipidosis), 시스틴축적증(cystinosis), 시알산 축적 질병(sialic acid storage disorder), 마리네스코-쉐글렌 증후군(Marinesco-Sjögren syndrome)을 갖는 킬로미크론 보유 질병(chylomicron retention disease), 헤르만스키-푸드락 증후군(Hermansky-Pudlak syndrome), 체디아크-히가시 증후군(Chediak-Higashi syndrome), 다논병(Danon disease), 또는 겔레오피직 이형성증(Geleophysic dysplasia)일 수 있다. 예를 들어, LSD 는 폼피병(Pompe disease) 또는 파브리병(Fabry's disease)일 수 있다.

[0025] 본원에 기재된 실시형태 중 하나에서, 상기 만노시다제에 대해서, 아미노산 곁사슬(amino acid side chains)에서의 원자의 삼차원의 단백질 좌표(three dimensional protein coordinates)는 도 33 에서의 등가의 원자(equivalent atoms)의 좌표의 1.5 Å 편차(deviation) 내의 범위에 들어가는 최소의 촉매 중심(minimal catalytic center)에 위치한다.

[0027] 본원에 기재된 실시형태 중 하나에서, 만노시다제는 SEQ ID NO: 50 의 잔기 1 내지 774 에 나타낸 아미노산 서열 또는 SEQ ID NO: 50 에 나타낸 아미노산 서열과 적어도 90 % 동일성(identity)[예를 들어, 적어도 95 % 또는 98 % 동일성]을 갖는 아미노산 서열을 포함할 수 있다.

[0029] 본원에 기재된 실시형태 중 하나에서, 만노시다제는 하기를 갖는 아미노산 서열을 포함할 수 있다: (i) GVGXXGXGG 모티프(motif), 이 식에서, X 는 Gly, Ala, Ser, Thr, 또는 Cys 이다; (ii) VRXE 모티프, 이 식에서, X 는 Pro 외의 모든 아미노산이다; (iii) X_1YQGX_2 모티프, 이 식에서 X_1 은 Leu, Ile, Val, Ala, Phe, Tyr 또는 Met 이고, X_2 는 Thr, Ser, 또는 Asn 이다; 또는 (iv) GDXGN 모티프, 이 식에서 X 는 Pro 외의 모든 아미노산일 수 있다.

[0031] 본원에 기재된 실시형태 중 하나에서, 상기 만노시다제는 *C. 셀룰란스(C.* cellulans), 스트렙토마이세스 코엘리컬러(Streptomyces coelicolor), 또는 *스트렙토마이세스 리비단스(Streptomyces lividans)* 만노시다제일 수 있다.

[0033] 본원에 기재된 실시형태 중 하나에서, 균류 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드[예를 들어, *야로위아 리폴리티카(Yarrowia lipolytica), S.세레비시아(S. cerevisiae), 오가테아 미누타(Ogataea minuta), 피키아 파스토리스(Pichia pastoris)* 또는 *C. 알비칸스(C. albicans)* 폴리펩티드와 같은 MNN4 폴리펩티드]를 코드화하는 핵산을 추가적으로 포함할 수 있고 및/또는 OCH1 활성도가 결핍되게 유전적으로 조작될 수 있다. 예를 들어, 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드는 *P. 파스토리스* PNO1 폴리펩티드일 수 있다.

[0035] 본원에 기재된 실시형태 중 하나에서, 상기 만노시다제는, 세포내 구간(intracellular compartment)에 상기 만노시다제를 타겟시키기 위해 분비 신호(secretion signal) 및/또는 타겟팅 신호(targeting signal)를 포함할 수 있다. 타겟 단백질 및 만노시다제는 공동 분비될(co-secreted) 수 있다.

[0037] 본 문서는 말단 포스포-6-만노스 잔기를 포함하는 분리된 당단백질을 특징으로 하고, 여기에서 상기 단백질은 본원에 기재된 방법으로 생산되었다.

[0039] 또 다른 양상에서, 본 문서는 당단백질을 포함하는 조성물을 특징으로 하고, 여기에서 당단백질에서의 N-글리칸(N-glycans)의 적어도 47 % 는 말단 포스포-6-만노스 잔기를 갖는다. 예를 들어, 당단백질에서의 N-글리칸(N-glycans)의 적어도 50 %, 75 %, 80 %, 85 %, 또는 90 % 는 말단 포스포-6-만노스 잔기를 갖는다.

[0041] 본 문서는, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, 또는 SEQ ID NO:14 에 나타낸 뉴클레오티

드 서열(nucleotide sequence), 또는 SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, 또는 SEQ ID NO:20 와 적어도 90 % 동일한 뉴클레오티드 서열을 포함하는 분리된 핵산(isolated nucleic acid)을 또한 특징으로 한다. 본 문서는 앞에 언급한 핵산에 실시가능하게 연결된 프로모터(promoter operably linked to such a nucleic acid)를 포함하는 벡터를 또한 특징으로 하고, 여기에서 핵산은 만노시다제를 코드화한다. 핵산은 세포내 구간에 만노시다제를 타겟시키기 위해 분비 신호(secretion signal) 또는 타겟팅 서열(targeting sequence)을 포함할 수 있다.

[0043] 별도로 나타내지 않았다면, 본원에서 사용된 모든 기술 및 과학 용어는 이러한 본 발명에 속하는 본 분야에서의 통상적인 기술 중의 하나로 보통 이해된 바와 같은 동일한 의미를 갖는다. 본원이 기재된 이러한 것들과 유사하거나 또는 동등한 방법 및 물질은 본 발명의 실행(practice) 또는 테스트에 사용될 수 있긴 하지만, 대표적인 방법 및 물질은 하기에 기재되어 있다. 모든 출판물(publications), 특허 출원서(patent applications), 특허(patents), Genbank® 수납 번호(Accession Nos) , 및 본원에서 언급된 그 밖의 참고문헌은 이들의 전체가 참고문헌으로서 포함된다. 상충되는 경우에(In case of conflict), 정의(definitions)를 포함하는 본 특허 출원서가 조절할 것이다(control). 상기 물질, 방법 및 예는 설명적일 뿐, 이로 제한하려는 의도는 아니다.

[0044]

[0045] 본 발명의 그 밖의 특징 및 장점은 하기의 상세한 설명 및 청구의 범위로부터 분명할 것이다.

도면의 간단한 설명

[0047] 도 1 은 pYLTmAX 및 pYLTmAXMnn4 구조물(constructs)의 도식도(schematic)이다.

도 2 는, MTLY60△och1[Mnn4 의 1 야생형 복제물(wild type copy)], MTLY60△och1 + Hp4dMnn4(1WT + Mnn4 의 여분의 복제물) 및 MTLY60△och1 + Hp4dMnn4 + TEFMnn4의 당 분석(sugar analysis)을 나타내는 일렉트로페로그램 (electroferograms)의 시리즈(series)이다. P 는 단일인산화된 피크(monophosphorylated peak)를 나타내고, PP 는 이인산화된 피크(diphosphorylated peak)를 나타내고, Man8 은 $Man_8GlcNAc_2$ 피크를 나타낸다.

도 3 은, 포유동물 및 효모 글리칸 인산화반응 경로(mammalian and yeast glycan phosphorylation pathways)의 도식도이다. 포유동물의 글리칸 인산화반응 경로는 $Man_8GlcNAc_2$ 글리칸에 GlcNAc-인산전달효소(phosphotransferase)에 의해 촉진된(catalyzed) 포스포-GlcNAc를 첨가한 다음에, 언커버링 효소(uncovering enzyme)에 의해 인산염(phosphate)을 노출시키기 위해 GlcNAc의 디캡핑(decapping)을 포함한다. 이와 대조적으로, 효모 글리칸 인산화반응은 $Man_8GlcNAc_2$ 글리칸에 포스포-만노스의 첨가를 포함하지만, 어떠한 내생 효소(endogenous enzyme)는 인산염을 노출시키기 위해 만노스의 캡을 제거하기(uncap) 위해 제시되지 않았다.

도 4 는 *C. 셀룰란스(C. cellulans)* 배지로부터의 상청액(supernatants)과 함께 상이한 시간 프레임(different time frames)[7 hrs, 8 hrs, 또는 밤새(overnight, ON)] 동안 처리된 스트레인(strain) MTLY60△och1 + Hp4dMnn4 + TEFMnn4 으로부터 유도된 N-글리칸을 나타낸 일렉트로페로그램의 시리즈이다.

도 5 는 포스파타아제(CIP) 배양 없이 및 포스파타아제(phosphatase)와 함께 *C. 셀룰란스* 상청액(SN)으로 처리된 MNN4 과잉발현된 스트레인(overexpressing strain)으로부터 유도된 N-글리칸을 나타내는 일렉트로페로그램의 시리즈이다.

도 6 은 나타낸 MW 에서의 용리액 일부(elution fractions)의 흡광 단위(absorbance units)의 그래프이다. 각각의 용리액 일부는 ~ 500 μl 이다.

도 7 은, 실리카-기초 겔 여과(silica-based gel filtration)[각각의 분획(each fraction)의 250 μl 은 침전된(precipitated) DOC/TCA 이다]로부터의 용리액 일부(elution fractions)의 전기영동법(electrophoresis) 후의 SDS-폴리아크릴아미드 겔(polyacrylamide gel)을 나타낸 것이다. 박스표시된 밴드(boxed band)는 질량 분광분석기(tandem mass spectrometry)(MS/MS)를 사용한 펩티드 질량 지문분석(peptide mass fingerprinting) 및 드 노보 시퀀싱(de novo sequencing)에 적합하다.

도 8A 는 CcMan1[즉, *C. 셀룰란스*로부터의 만노시다제 후보물질(candidate) 1]을 코드화하는 뉴클레오티드 서열(SEQ ID NO:6)[콘틱(contig) 1003 에서]이고, 이는 MS/MS 드 노보 시퀀싱에서 확인되었다. 도 8B 는 신호 서열을 포함하는[볼드체(in bold)], CcMan 1 의 아미노산 서열(SEQ ID NO:7)이다. 신호 서열 없는 CcMan 1 폴리펩티드의 예측된 분자량은 92.6 kDa 이다.

도 9A 는 CcMan2 를 코드화하는 뉴클레오티드 서열(SEQ ID NO:8)(콘틱 774에서)이고, 도 9B 는 신호 서열(signal sequence)(볼드체)을 갖는 CcMan2의 아미노산 서열(SEQ ID NO:9)이다. 신호 서열 없는 CcMan2 폴리펩티드의 예측된 분자량은 121.6 kDa 이다.

도 10A 는 CcMan3 을 코드화하는 뉴클레오티드 서열(SEQ ID NO:10)(콘틱 774에서)이고, 도 10B 는 신호 서열(볼드체)을 갖는 CcMan3 의 아미노산 서열(SEQ ID NO:11)이다. 신호 서열 없는 CcMan3 폴리펩티드의 예측된 분자량은 116 kDa 이다.

도 11A 는 CcMan4 를 코드화하는 뉴클레오티드 서열(SEQ ID NO:12)(콘틱 1237에서)이고, 도 11B 는 신호 서열(볼드체)을 갖는 CcMan4 의 아미노산 서열(SEQ ID NO:13)이다. 신호 서열 없는 CcMan4 폴리펩티드의 예측된 분자량은 184 kDa 이다.

도 12A 는 CcMan5 를 코드화하는 뉴클레오티드 서열(SEQ ID NO:14)(콘틱 896에서)이다. 도 12B 는 신호 서열(볼드체)을 갖는 CcMan5 의 아미노산 서열(SEQ ID NO:15)이고, 도 12C 는 신호서열이 없는 CcMan5 의 아미노산 서열(SEQ ID NO:50)이다. 신호 서열 없는 CcMan5 폴리펩티드의 예측된 분자량은 173 kDa 이다.

도 13 은, N-말단에 태그된(tagged)(pYLPNtCcMan1-5) 또는 C-말단에 태그된(pYLPCtCcMan1-5), *야로위아 리폴리티카*(pYLPSecCcMan1-5)에서의 분비된 단백질로서, *야로위아 리폴리티카*의 분비경로(secretory pathway)에 타겟된(targeted) 단백질로서, *E. coli* (pET25-Man)의 주변 세포질(periplasm)에서의 CcMan1-5 의 발현을 위한 발현 플라스미드(expression plasmid)의 예를 포함한다.

도 14 는, *E. coli* 에서의 발현에 적합한(optimized) 코돈인 CcMan1의 뉴클레오티드 서열이다(SEQ ID NO:16).

도 15 는, *E. coli* 에서의 발현에 적합한 코돈인 CcMan2의 뉴클레오티드 서열이다(SEQ ID NO:17).

도 16 은, *E. coli* 에서의 발현에 적합한 코돈인 CcMan3의 뉴클레오티드 서열이다(SEQ ID NO:18).

도 17 는, *E. coli* 에서의 발현에 적합한 코돈인 CcMan4의 뉴클레오티드 서열이다(SEQ ID NO:19).

도 18 는, *E. coli* 에서의 발현에 적합한 코돈인 CcMan5의 뉴클레오티드 서열이다(SEQ ID NO:20).

도 19 는, pLSAH36 및 pLSH36 벡터의 도식도 및 벡터 내로 *C. 셀룰란스* 유전자를 도입하기 위한 클로닝 전략(cloning strategy)이다.

도 20 은, *E. coli* 세포를 발현시키는 CcMan4 및 CcMan5의 주변세포질의 분획(periplasmic fraction)의 분석을 나타낸 일렉트로페로그램의 시리즈이다. 분석은 DNA 시퀀서-지원된, 형광단-지원된 탄수화물 전기영동(DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis, DSA-FACE)을 사용하여 실행하였다. 첫 번째 및 두 번째 패널(panel)은 각각 RNaseB로부터의 덱스트란 레더(dextran ladder) 및 당을 나타낸다. 세 번째 패널은, 모노 만노인산화된(mono mannophosphorylated) $Man_8GlcNAc_2$ 피크(peak)에 해당하는 "P", 더블 모노만노인산화된(double mannophosphorylated) $Man_8GlcNAc_2$ 피크에 해당하는 "PP", 및 $Man_8GlcNAc_2$ 에 해당하는 "Man8" 을 갖는 처리되지 않은 Mnn4 당이다. 패널 4 내지 9 는 차후에 송아지 장내의 포스파타제(calf intestinal phosphatase, CIP) 소화(digest)시키거나 시키지 않은, 나타낸 주변 세포질(indicated periplasm)과 함께 배양시킨 Mnn4 글리칸으로 수득된 결과이다.

도 21 은, Zhu *et al.*, *Nat. Chem. Biol.*, 6(2):125-32. Epub 2009 Dec 27 (2010)에 기재된 Bt3990(744 AA) 및 Bt2199(739 AA)과 함께 CcMan4(1759 AA) 및 CcMan5(1650 AA)의 도식적인 평면도(schematic alignment)이다.

도 22 는, *E. coli* 세포의 발현으로부터 수득된 CcMan4 및 CcMan5 효소의 분석을 나타낸 일렉트로페로그램의 시리즈이다. 분석은 기질(substrate)로서의 글리칸(MNN4 글리칸 또는 MNN4 당을 나타냄)을 유도하는(derived) MNN4 과잉발현된 스트레인을 사용한 DSA-FACE 를 사용하여 실행하였다. 첫 번째 패널은 덱스트란 레더를 나타내고, 두 번째 패널은 처리되지 않은 Mnn4 당을 나타낸다. 세 번째 내지 여섯 번째 패널에서, CcMan4domain 주변 세포질 분획(periplasmic fraction)과 함께 배양시킨 당은 유도되지 않았고, IPTG 와 함께 18 ℃에서 밤새 유도시켰고, CcMan5domain 주변 세포질 분획은 유도되지 않았고, IPTC 와 함께 18 ℃ 에서 밤새 유도시켰다. 마지막 패널은 RNaseB 로부터의 당을 나타낸다.

도 23 은, N-말단 β-샌드위치 도메인(잔기 8-271: 회백색), α-나선형 링커(α-helical linker)(잔기 272-

290; 흑색) 및 (α α)6 바렐 도메인(barrel domain)(잔기 291-771; 암회색)으로 이루어진 $CcMan5_{1-774}$. $CcMan5_{1-744}$ 을 나타낸 리본(ribbon)이다. 촉매의 Ca2+ 는 구체(sphere)로서 나타내었다.

도 24 는, 막대기 표시(stick representation)에 나타낸 기질 결합 부위(substrate binding site)를 선으로 그은(lining) 곁사슬을 갖는 $CcMan5_{1-774}$ 단백질 백본(backbone)을 나타낸 리본(ribbon)이다. 탄소, 산소 및 질소 원자는 각각 회백색, 희색 및 암회색으로 나타내었다. 촉매 중심(catalytic center)에서의 Ca2+ 이온 및 물 W1, W2, W3 및 W4 는 구체로서 나타내었다.

도 25 는, 막대기로 나타낸 만노스-1-포스포-6-만노스의 모델이 되는 위치(modeled position) 및 기질 결합 부위를 선으로 그은 곁사슬을 갖는 $CcMan5_{1-774}$ 단백질 백본(backbone)을 나타낸 리본이다. 탄소, 산소 및 질소 원자는 각각 회백색, 희색 및 암회색으로 나타내었다. 촉매 중심(catalytic center)에서의 Ca^{2+} 이온 및 물 W1, W2, W3 및 W4 는 구체로서 나타내었다[비교를 위해, 기질 O2 및 O3 히드록실 기로 치환될 수 있는 W2 및 W3 의 위치를 여전히 나타냄]. 황색, 적색 및 흑색의 점선(dashed line)은 Ca2+ 를 갖는 배위 결합(coordination bond), 제안된 친핵성 물(proposed nucleophilic water)(W4)을 갖는 H-결합(H-bonds) 및 각각 -1 부위 만노스(site mannose) 및 포스페이트(phosphate)를 갖는 H-결합을 나타낸다. -1 부위 만노스는 이의 바닥 상태의 의자 입체형태(ground state chair conformation)를 모델로 한다. 촉매 작용 동안에, 이의 O2 히드록실은 적도 좌표면(equatorial coordination plane)에서, W2 에 대해 나타낸 보다 가까운 위치를 점령할 수 있고, 이로 인하여 만노스 -1 고리(ring)에서의 반의자 형태(half-chair conformation)에 대한 뒤틀림(distortion)을 유도하고, C1 탄소[화살표(arrow)]에서의 친핵성 물(W4)의 라인 공격(line attack)을 가능하게 한다.

도 26A 는, 볼드체에서의 lip2 pre 서열 및 밑줄처진 Myc His 태그(tag)를 갖는 α-갈락토시다아제A(GalactosidaseA)를 코드화하는 뉴클레오티드 서열이 최적화된 *Y. 리폴리티카* 코돈이다(SEQ ID NO:22). 도 26B 는, 볼드체에서의 lip2 pre 서열 밑줄처진 Myc His 태그(tag)를 갖는 I-갈락토시다아제A(GalactosidaseA)의 아미노산 서열이다(SEQ ID NO:23).

도 27A 는, 볼드체에서의 lip2 pre 서열을 갖는 인간 알파 갈락토시다아제(GAA)의 뉴클레오티드 서열이 최적화된 코돈이다(SEQ ID NO:24). 도 27B 는, 볼드체에서의 lip2 pre 서열을 갖는 인간 GAA의 아미노산 서열이고(SEQ ID NO:25), 여기에서 * 는 종결 코돈(stop codon)을 나타낸다.

도 28 은, huGAA 의 클로닝에 사용된 *Y. 리폴리티카* 발현 벡터(expression vector)의 도식도이다.

도 29 는, *E. coli* 세포의 주변 세포질 분획으로부터 유도된 CcMan5 와 함께 huGAA의 처리의 분석을 나타낸 일렉트로페로그램의 시리즈이다. 분석은 DSA-FACE를 사용하여 실행하였다.

도 30 은, CcMan5의 최소의 촉매 중심(minimal catalytic center)을 나타낸 것이다. SEQ ID NO:50 에서의 등가의 잔기(equivalent residues)의 넘버링(numbering)은 괄호에 나타내었다. 1: Q (Q536); 2: N/D-E/Q (N588-Q589); 3: D/E (D355); 4: R (R405); 5: D/E-X-D/E (D660-X-D662); 6: G-G (G71-G72); 및 7: T/S/G (T626).

도 31 은, MUSCLE(MUltiple Sequence Comparison by Log-Expectation)을 사용한 CcMan5 의 아미노산 서열[SEQ ID NO:50, 신호 펩티드가 없는 SEQ ID NO:15 에 나타낸 아미노산 서열] 및 이의 동족체(homologs) 중의 10의 배열(alignment)이다. NP_630514 *스트렙토마이세스(Streptomyces)*, SEQ ID NO:26; ZP_02866543 *클로스트리디움(Clostridium)*, SEQ ID NO:27; NP_812442 *박테로이데스(Bacteroides)*, SEQ ID NO:28; YP_003584502 *주농왕기아(Zunongwangia)*, SEQ ID NO:29; YP_003120664 *키티노파가(Chitinophaga)*, SEQ ID NO:30; AAK22560 *카울로박터(Caulobacter)*, SEQ ID NO:31; ACL94075 *카울로박터(Caulobacter)* , SEQ ID NO:32; ACT03290 *패니바실러스(Paenibacillus)*, SEQ ID NO:33; ACU59240 *키티노파가(Chitinophaga)*, SEQ ID NO:34; ACU05553 *페도박터(Pedobacter)*, SEQ ID NO:35.

도 32 는, MUSCLE을 사용한 CcMan5의 아미노산 서열(SEQ ID NO:50) 및 이의 동족체 중의 19 의 배열이다. *스트렙토마이세스(Streptomyces)* NP_630514, SEQ ID NO:26; *스트렙토마이세스(Streptomyces)* ZP_02866543, SEQ ID NO:36, ZP_06527366 *스트렙토마이세스(Streptomyces)*, SEQ ID NO:37; YP_003013376 *패니바실러스(Paenibacillus)*, SEQ ID NO:38; NP_812442 *박테로이데스(Bacteroides)*, SEQ ID NO:28; ZP_04848482 *박테로이데스(Bacteroides)*, SEQ ID NO:39; ZP_03677957 *박테로이데스(Bacteroides)*, SEQ ID NO:40; YP_003584502 *주농왕기아(Zunongwangia)*, SEQ ID NO:29; ZP_01061975 *레이우엔훅키엘라(Leeuwenhoekiella)*, SEQ ID NO:41; ZP_07083984 *스핑고박테리움(Sphingobacterium)*, SEQ ID NO:42; YP_003120664 *키티노파가(Chitinophaga)* , SEQ ID NO:30; ZP_01885202 *페도박터(Pedobacter)*, SEQ ID NO:43; ZP_02866543 *클로스트리디움(Clostridium)*, SEQ

ID NO:27; XP_367221 *마그나포르테(Magnaporthe)*, SEQ ID NO:44; ZP_07042437 *박테로이데스(Bacteroides)*, SEQ ID NO:45; ZP_05759807 *박테로이데스(Bacteroides)*, SEQ ID NO:46; ZP_05287524 *박테로이데스(Bacteroides)*, SEQ ID NO:47; ZP_06076108 *박테로이데스(Bacteroides)*, SEQ ID NO:48; YP_001302992 *파라박테로이데스(Parabacteroides)*, SEQ ID NO:49.

도 33 은, $CcMan5_{1-774}$ 의 활성 부위(active site)를 둘러싸는 잔기의 구조적인 좌표(structural coordinates)를 포함한다.

도 34 는, PDB entry 2xsg에서의 비대칭의 유닛(asymmetric unit)에서의 두 가지의 $CcMan5_{1-774}$ 분자의 촉매의 Ca2+ 원자 및 단백질 C 알파 원자를 포함하고, 단백질의 전체적인 접힌 부분(fold)을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0048] 일반적으로, 본 문서는 캡핑이 제거된(uncapped) 포스포-6-만노스 잔기(M6P) 잔기를 갖는 타겟 분자(target molecules)(예를 들어, 타겟 단백질)를 생산하도록 당단백질에서의 만노스-1-포스포-6-만노스 결합을 가수분해하기 위한 방법 및 물질을 제공한다. 본원에 기재된 방법 및 물질은 리소좀 축적 질환(lysosomal storage disorders, LSDs)을 갖는 환자를 치료하기 위한 제제를 생산하는데 특히 유용하고, 유전적인 대사 질환(hereditary metabolic disorders)의 다양한 군은, 이들의 분해를 포함하는 이화작용 효소(catabolic enzymes)의 손상된 활성도 때문에 리소좀에서의 저장 생산물(storage products)의 축적에 의해 특징지어진다. 저장 생산물의 증가(build-up)는 세포의 역기능(cell dysfunction) 및 점진적인 임상적인 징후(progressive clinical manifestations)로 이어진다. 투여된 효소가 질병에 걸린 세포의 리소좀을 타겟될(targeted) 수 있는 경우에, 이화작용 효소에서의 결핍증(Deficiencies)은 효소 보충 요법(enzyme replacement therapy)에 의해 고쳐질 수 있다. 리소좀 효소(Lysosomal enzymes)는 일반적으로, 소포체(endoplasmic reticulum)에서 합성되고, 골지(Golgi)에서의 분비 경로(secretory pathway)를 통해 수송된 다음에 리소좀에 채용되는(recruited) 당단백질이다. 리소좀 효소가 리소좀에 전달되는 하나의 방식은 양이온-의존(cation-dependent, CD) 만노스 6-인산 수용체(mannose 6-phosphate receptor, MPR)를 통한다. M6P 말단 글리칸(terminal glycans)은 분비 경로로부터의 리소좀 효소의 분류(sorting)를 중재하고, 리소좀에 효소를 수송하는 두 가지의 MPRs 에 의한 후기골지망(trans-Golgi network, TGN)에서 인지된다. 본원에서 기재된 방법 및 물질을 사용하여, 미생물의 기저 생산 공정(microbial based production process)은 캡핑이 제거된(uncapped) M6P 글리칸을 갖는 치료상의 단백질(therapeutic proteins)을 수득하는데 사용될 수 있고, 이는 동일한 M6P 의존 경로(M6P dependent pathway)를 이용하여 리소좀으로 수송할 수 있다. 따라서, 본원에 기재된 방법 및 물질은 LSDs 와 같은 대사성 질환의 치료를 위한 당단백질을 제조하는데 유용하다.

[0050] *만노시다제(Mannosidases)*

[0051] 본 문서는 올리고당에서 말단의 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 만노시다제 폴리펩티드를 코드화하는 분리된 핵산 뿐만 아니라 올리고당에서의 말단 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 분리된 만노시다제를 제공한다. 용어 "핵산" 및 "폴리뉴클레오티드"는 본원에서 서로 교체가능하게 사용되었고, 핵산 유사체(nucleic acid analogs)를 포함하는 cDNA, 게놈의(genomic) DNA, 합성(synthetic) DNA, 및 DNA (또는 RNA)를 나타낸다. 폴리뉴클레오티드는 모든 삼차원 구조(three-dimensional structure)를 가질 수 있다. 핵산은 2중 가닥(double-stranded) 또는 단일 가닥(single-stranded)[즉, 센스 가닥(sense strand) 또는 안티센스 가닥(antisense strand)]일 수 있다. 폴리뉴클레오티드의 이로 제한되지 않는 예는 유전자(genes), 유전자 단편(gene fragments), 엑손(exons), 인트론(introns), 전령(messenger) RNA(mRNA), 운반(transfer) RNA, 리보솜(ribosomal) RNA, siRNA, 마이크로(micro)-RNA, 리보자임(ribozymes), cDNA, 재조합 폴리뉴클레오티드, 가지형 폴리뉴클레오티드(branched polynucleotides), 플라스미드(plasmids), 벡터(vectors), 어떠한 서열 중 분리된(isolated) DNA, 어떠한 서열 중 분리된 RNA, 핵산 프로브(nucleic acid probes), 및 프라이머(primers), 뿐만 아니라 핵산 유사체(analogs)를 포함한다.

[0053] "폴리펩티드" 및 "단백질"은 본원에서 서로 교차가능하게 사용되고, 길이 또는 번역후 변형(post-translational modification)과 상관없이 아미노산의 어떠한 펩티드-결합 사슬(peptide-linked chain)을 의미한다. 일반적으로, 표본(preparation)에서의 전체 단백질의 적어도 60 중량%, 예를 들어 샘플에서의 전체 단백질의 60 % 를 구성하였을 때, 본원에서 기재된 폴리펩티드[예를 들어, 만노시다제 또는 캡핑이 제거된(uncapped) M6P 잔기를 갖는 타겟 단백질]는 분리된다. 몇몇 실시형태에서, 본원에 기재된 폴리펩티드는 표본(preparation)에서의 전체 단백질의 적어도 75 중량%, 적어도 90 중량% 또는 적어도 99 중량% 로 이루어져 있다.

[0055] "분리된 핵산(isolated nucleic acid)"은, 자연적으로 발생된 게놈(naturally-occurring genome)(예를 들어, 효모 게놈)에서의 핵산의 한쪽 또는 양쪽 면(one or both sides)의 측면에 배치된(flank) 핵산을 포함하는, 자연적으로 발생된 게놈에 나타낸 그 밖의 핵산 분자로부터 분리된 핵산을 나타낸다. 이러한 자연적으로 발생되지 않은 핵산은 자연에서 발견되지 않고, 자연적으로 발생된 게놈에서의 근접한 서열(contiguous sequences)을 직접적으로 갖지 않기 때문에, 핵산에 대해서 본원에서 사용된 바와 같은, 용어 "분리된(isolated)"은 어떠한 자연적으로 발생되지 않는(non-naturally-occurring) 핵산 서열을 또한 포함한다.

[0057] 분리된 핵산은, 예를 들어 자연적으로 발생한 게놈에서의 DNA 분자가 제거되거나 부재한 측면(flanking)의 바로 옆에서 발견된 일반적으로 발견된 핵산 서열 중의 하나가 제공된 DNA 분자일 수 있다(An isolated nucleic acid can be, for example, a DNA molecule, provided one of the nucleic acid sequences normally found immediately flanking that DNA molecule in a naturally-occurring genome is removed or absent). 따라서, 분리된 핵산은, 벡터, 자체적인 복제 플라스미드(autonomously replicating plasmid), 바이러스[예를 들어, 모든 파라믹소바이러스(paramyxovirus), 레트로바이러스(retrovirus), 렌티바이러스(lentivirus), 아데노바이러스(adenovirus), 또는 헤르페스 바이러스(herpes virus)] 또는 원핵 생물(prokaryote) 또는 진핵 생물(eukaryote)의 게놈 DNA 내로 포함되는 DNA 뿐만 아니라 그 밖의 서열과는 관계없는(independent of) 분리된 분자(separate molecule)[예를 들어, 화학적으로 합성된 핵산, 또는 PCR 또는 제한 효소 처리(restriction endonuclease treatment)에 의해 생성된 cDNA 또는 게놈 DNA 단편(genomic DNA fragment)]로서 존재하는 DNA 분자를 제한없이 포함한다. 게다가, 분리된 핵산은 잡종(hybrid) 또는 융합 핵산(fusion nucleic acid)의 일부인 DNA 분자와 같은 조작된 핵산(engineered nucleic acid)을 포함할 수 있다. 예를 들어, cDNA 라이브러리 또는 게놈 라이브러리(genomic libraries), 또는 게놈 DNA 제한효소 분해(genomic DNA restriction digest)를 포함하는 겔 슬라이스(gel slices) 내의 수백 내지 수백만의 다른 핵산 중에 존재하는 핵산은 분리된 핵산으로 간주되지 않는다.

[0059] 핵산 및 특정한 숙주 세포와 관련하여 본원에서 사용된 바와 같은 용어 "외인성(exogenous)"은 자연에서 발견된 바와 같은 특정한 세포에서 발생되지 않는(특정한 세포로부터 획득될 수 없는) 어떠한 핵산을 나타낸다. 따라서, 자연적으로 발생되지 않는 핵산(non-naturally-occurring nucleic acid)은 숙주 세포 내로 일단(once) 도입된 숙주 세포에 대해서 외인성일 수 있는 것으로 간주된다. 자연적으로 발생되지 않는 핵산은, 핵산이 전체로서 자연에서 존재하지 않는다면 자연에서 발견된 핵산 서열 또는 핵산 서열의 단편을 포함할 수 있음을 주의해야 한다. 예를 들어, 발현 벡터 내의 게놈 DNA 서열을 포함하는 핵산 분자는 자연적으로 발생하지 않는 핵산이고, 따라서 전체로서(게놈 DNA 뿐만 아니라 벡터 DNA) 핵산 분자가 자연에 존재하지 않기 때문에, 숙주 세포 내로 일단 도입된 숙주 세포에 대해서 외인성이다. 따라서, 어떠한 벡터, 자체적으로 복제하는 플라스미드, 또는 전체로서 자연에 존재하지 않는 바이러스(예를 들어, 레트로바이러스, 아데노 바이러스 또는 헤르페스 바이러스)가 자연적으로 발생하지 않는 핵산일 것으로 간주된다. PCR 또는 제한 엔도뉴클레아제 처리(restriction endonuclease treatment)에 의해 생산된 게놈 DNA 단편 뿐만 아니라 cDNA 는, 이들이 자연에서 분리된 분자로서 발견되지 않기 때문에 자연적으로 발생하지 않는 핵산일 것으로 간주된다는 결과가 된다. 또한, 배열(arrangement)에서의 프로모터 서열(promoter sequence) 및 폴리펩티드-코드화하는 서열(polypeptide-encoding sequence)[예를 들어, cDNA 또는 게놈의 DNA]을 포함하는, 자연에서 발견되지 않는 어떠한 핵산은 자연적으로 발생하지 않는 핵산인 결과가 된다. 자연적으로 발생된 핵산은 특정한 세포에 대해서 외인성일 수 있다. 예를 들어, 염색체가 효모의 세포 내로 일단 도입된다면, 효소 x 의 세포로부터 분리된 전체의 염색체가 효모 y의 세포에 대하여 외인성 핵산이다.

[0061] 만노시다제를 코드화하는 핵산은, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, 또는 SEQ ID NO: 14 에 나타낸 뉴클레오티드 서열과 적어도 70 % 의 서열 동일성(sequence identity)(예를 들어, 적어도 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100% 의 서열 동일성)을 가질 수 있다. 몇몇 실시형태에서, 본원에 기재된 핵산은, SEQ ID NOs: 7, 9, 11, 13, 15, 50 에 나타낸 아미노산 서열과 적어도 70 % (예를 들어, 적어도 75, 80, 85, 90, 95, 99, 또는 100 %)의 동일성을 갖는 만노시다제 폴리펩티드를 코드화할 수 있다. 예를 들어, 핵산은, SEQ ID NO:15 또는 SEQ ID NO:50, 또는 이의 부분(portion)에 나타낸 아미노산 서열과 적어도 90 % (예를 들어, 적어도 95 % 또는 98 %)동일성을 갖는 만노시다제를 코드화할 수 있다. 예를 들어, 핵산은 SEQ ID NO: 50의 아미노산 잔기 1 내지 774 와 적어도 90 % 동일성을 갖는 만노시다제를 코드화할 수 있다. 특정한 아미노산 서열 및 SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, or SEQ ID NO:50 에 나타낸 아미노산 서열 사이의 퍼센트 동일성(percent identity)은 하기와 같이 결정된다. 첫째로, 아미노산 서열은 BLASTP 버전(version) 2.0.14 를 포함하는 BLASTZ의 독립형 버전(stand-alone version)으로부터

의 BLAST 2 Sequences(Bl2seq) 프로그램을 사용하여 나란히 하였다(aligned). BLASTZ의 이러한 독립형 버전은 Fish & Richardson's 웹사이트(예를 들어, www.fr.com/blast/) 또는 the U.S. government's National Center for Biotechnology Information 웹사이트(www.ncbi.nlm.nih.gov)로부터 획득할 수 있다. Bl2seq 프로그램을 사용하는 방법을 설명하는 지시(Instructions)는 BLASTZ 를 수반하는 리드미 파일(readme file)에서 발견될 수 있다. Bl2seq 는 BLASTP 알고리즘(algorithm)을 사용한 두 개의 아미노산 서열 사이를 비교하여 수행한다. 두 개의 아미노산 서열을 비교하기 위해, Bl2seq 의 옵션(options)은 하기에 나타내었다: -i 는 비교될 첫 번째 아미노산 서열을 포함하는 파일로 설정된다(예를 들어, C:₩seq1.txt) ; -j 는 비교될 두 번째 아미노산 서열을 포함하는 파일로 설정된다(예를 들어, C:₩seq2.txt) ; -p 는 blastp 로 설정된다 ; -o 는 어떠한 원하는 파일 이름이다(예를 들어, C:₩output.txt); 및 모든 다른 옵션은 이들의 디폴트 설정(default setting)에 둔다. 예를 들어, 하기의 명령어(command)는 두 개의 아미노산 서열 사이 비교를 포함하는 출력 파일(output file)을 발생시키기 위해 사용될 수 있다: C:\Bl2seq -i c:\seq1.txt -j c:\seq2.txt -p blastp -o c:\output.txt. 만약 두 개의 비교된 서열이 상동 관계(homology)를 공유한다면, 그때 지정된 출력 파일(designated output file)은 정렬된 서열(aligned)로서의 상동(homology)의 이러한 영역(those regions)을 나타낼 것이다. 만약 두 개의 비교된 서열이 상동 관계를 공유하지 않는다면, 그때 지정된 출력 파일은 정렬된 서열을 나타내지 않을 것이다. 유사한 절차는 blastn이 사용된 경우를 제외하고는 핵산 서열에 대해서 다음일 수 있다(Similar procedures can be following for nucleic acid sequences except that blastn is used).

[0063] 정렬하자마자, 매치된 수(the number of matches)는 동일한 아미노산 잔기가 둘 다의 서열에 나타낸 위치(positions)의 수를 집계하여 결정된다. 퍼센트 동일성(percent identity)은, 매치된 수를 전장(full-length)의 만노시다제 폴리펩티드 아미노산 서열의 길이(length)로 나눈 다음에 결과적인 수치를 100과 곱하여 결정된다. 예를 들어, SEQ ID NO:7 에 나타낸 서열과 함께 정렬하였을 때 700 매치(matches)를 갖는 아미노산 서열은 SEQ ID NO:7 에 나타낸 서열과 77.8 퍼센트 동일성을 갖는다(즉, 700 ÷ 900 × 100 = 77.8).

[0065] 퍼센트 동일성 수치는 소수점 1 자리에서 반올림함을 명심할 필요가 있다. 예를 들어, 78.11, 78.12, 78.13, 및 78.14 는 반올림 내려 78.1이고, 반면에 78.15, 78.16, 78.17, 78.18, 및 78.19 는 반올림 올려 78.2이다. 길이 값(length value)은 항상 정수(integer)임을 또한 명심할 필요가 있다.

[0067] 많은 핵산은 특정한 아미노산 서열을 갖는 폴리펩티드를 코드화할 수 있음을 인식할 수 있을 것이다. 유전학적인 코드(genetic code)의 퇴보(degeneracy)는 본 분야에서 알려져 있다; 즉, 많은 아미노산에 대해서, 아미노산에 대한 코돈(codons)으로서 제공되는 하나의 뉴클레오티드 삼중자(one nucleotide triplet)가 있다. 예를 들어, 주어진 만노시다제 폴리펩티드에 대한 코딩 서열(coding sequence)에서의 코돈은, 그러한 종에 대한 적절한 코돈 바이어스 테이블(codon bias tables)을 사용하여, 특정한 종(particular species)[예를 들어, 박테리아(bacteria) 또는 균류(fungus)]에서의 최적의 발현을 수득할 수 있도록 변형될 수 있다. 예를 들어, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, 또는 SEQ ID NO:14 에 나타낸 핵산은 도 14 내지 18 에 나타낸 바와 같은 *E. coli* 발현에 대해서 최적화된 코돈일 수 있다(SEQ ID NOs:16-20 을 참고하라).

[0069] 혼성화(Hybridization)는 두 개의 핵산 서열 사이의 상동성(homology)을 평가하기 위해 또한 사용될 수 있다. 본원에 기재된 핵산 서열 또는 이의 단편 또는 변이체(variant)는 표준 혼성화 기술(standard hybridization techniques)에 따른 혼성화 프로브(hybridization probe)로서 사용될 수 있다. 테스트 근원(test source)으로부터의 DNA 또는 RNA 에 대한 관심 프로브(probe of interest)(예를 들어, CcMan5 뉴클레오티드 서열의 일부를 포함하는 프로브)의 혼성화는 테스트 근원에서 프로브에 해당하는 DNA 또는 RNA(예를 들어, CcMan5 뉴클레오티드 서열)의 존재의 표시이다. 혼성화 조건은 본 분야에서 숙련자에게 알려져 있고, 이는 Molecular Biology, John Wiley & Sons, N.Y., 6.3.1-6.3.6, 1991 에 현재의 프로토콜에서 발견할 수 있다. 중간의 혼성화 조건(Moderate hybridization conditions)은 30 ℃에서 2X 염화나트륨/시트르산나트륨(SSC)에서의 혼성화한 다음에 50 ℃ 에서 1 X SSC, 0.1% SDS 에서의 세척(wash)과 동등한 것으로 나타났다. 매우 엄격한 조건(Highly stringent conditions)은 45 ℃에서의 6X 염화나트륨/시트르산 나트륨(SSC)에서의 혼성화한 다음에 65 ℃에서의 0.2 X SSC, 0.1% SDS 에서의 세척과 동등한 것으로 나타났다.

[0071] 올리고당에서 말단 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 만노시다제 폴리펩티드는 *C. 셀룰란스* (SEQ ID NO:50 의 잔기 1 내지 774, $CcMan5_{1-774}$ 로서 또한 나타냄)로부터의 만노시다제의 일부에 대한 본원에서 기재된 삼차원 구조(three dimensional structure)에 기초하여 또한 나타낼 수 있다. 삼차원 구조는, 예를 들어 $CcMan5_{1-774}$ 결정의 X선 회절(X-ray diffraction)에 의해 결정될 수 있다. $CcMan5_{1-774}$의 구조적인 좌표(Structural coordinates)[예를 들어, $CcMan5_{1-774}$의 좌표를 Protein Data Bank(PDB ID No. 2xs하의 PDB 에서의

월드와이드웹(world wide web))에 맡겼고, CcMan5의 촉매 중심(catalytic center)에 대한 좌표를 도 33 에 나타내거나, 또는 PDB entry 2xsg에서의 비대칭의 유닛(asymmetric unit)에서의 두 개의 $CcMan5_{1-774}$ 분자의 단백질 C 알파 원자 및 촉매 Ca2+ 원자에 대한 좌표를 도 34 에 나타내었다]는, 올리고당에서의 말단 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 만노시다제의 삼차원 구조의 특성 뿐만 아니라, 말단 포스포-6-만노스를 생산하기 위해 Man-P-Man 을 가수분해하기 위한 이의 능력을 부여하고, 기질로서의 만노스-6-포스페이트-알파, 1-만노스(mannose-6-phosphate-alpha,1-mannose)[이후로 Man-P-Man로서 나타냄]의 수용(acceptance)을 포함하는 만노시다제의 영역(regions)의 시각화(visualization), 확인(identification) 및 특성을 이로 제한하지 않고 포함하는 수많은 응용에 유용하다. "구조적인 좌표(Structural coordinates)"는 분자 또는 분자로 된 복합체에서의 그 밖의 원자에 대한 원자의 공간 관계(atom's spatial relationship)에 해당하는 데카르트좌표(Cartesian coordinates)이다. 구조적인 좌표는 X선 결정학 기술 또는 NMR 기술을 사용하여 수득할 수 있거나, 또는 분자 교체 기술(replacement) 또는 상동성 모델링(homology modeling)을 사용하여 유도될 수 있다. 다양한 소프트웨어 프로그램(software programs)은 분자 또는 분자의 복합체의 삼차원의 표시(representation)를 수득하기 위한 구조적인 좌표의 세트의 그래픽의 표시(graphical representation)를 가능하게 한다. 본원에 기재된 구조의 구조적인 좌표는, 전도(inversion) 또는 정수 더하기(integer additions) 또는 빼기(subtractions)와 같은 수학적인 처리(mathematical manipulation)에 의한 도 33 또는 도 34 에 제공된 원본의 세트(original set)로부터 변형될 수 있다. 이와 같이(As such), 본 발명의 구조적인 좌표가 상대적임을 인식할 수 있고, 도 33 또는 도 34 의 실질적인 x, y, z 좌표에 의해 결코 명확하게 제한되지 않는다.

[0073] 실시예 8 에 나타낸 바와 같이, $CcMan5_{1-774}$의 구조는 두 개의 도메인, α-나선형 링커(α-helical linker)(SEQ ID NO:50의 잔기 272 내지 290)를 통해 연결된 N-말단 β-샌드위치 도메인(N-terminal β-sandwich domain)(SEQ ID NO:50의 잔기 8 내지 271) 및 C-말단 (α α)6 바렐(barrel) 도메인(SEQ ID NO:50의 잔기 291 내지 2771)으로 이루어져 있다. 두 개의 도메인 사이의 공유 영역(interface)는, 보존된(conserved) 촉매의 Ca^{2+} 이온을 포함하는 쉘로우 캐비티(shallow cavity)를 구체화하고, -1 기질 결합 부위(substrate binding site)[Davies *et al.*, *Biochem. J.* 321:557-9 (1997)에 기재된 바와 같은 명명법(nomenclature)]를 구체화한다. SEQ ID NO: 50 의 잔기 22, 25, 71, 72, 195, 196, 354, 405, 535, 536, 588, 589, 626, 658, 660, 및 662 는 기질 결합 부위를 형성한다.

[0075] $CcMan5_{1-774}$ 의 삼차원 구조는, PDB ID No. 2xs의 구조적인 좌표, 또는 도 33 에 나타낸 발췌(extract)를 이용하는, $CcMan5_{1-774}$ 의 활성 부위(active site)를 둘러싸는(surrounding) 잔기, 또는 도 34 에 나타낸 발췌를 포함하는, PDB entry 2xsg 에서 비대칭의 유닛에서의 두 개의 $CcMan5_{1-774}$ 분자의 단백질 C 알파 원자 및 촉매 Ca2+ 원자를 포함하고, 단백질의 전체적인 접힌 부분(fold)를 기재한, 이의 부분 또는 이의 모두를 특징으로 할 수 있다. 예를 들어, $CcMan5_{1-774}$ 의 삼차원의 구조는, 2Å 보다 많지 않은 상기 아미노산의 보존된 백본 원자(conserved backbone atoms)로부터의 ± 평균 평방근편차(root mean square deviation), PDB ID No. 2xs에 따른 아미노산 잔기 7 내지 771의 구조적인 좌표에 의해 특징지어 질 수 있다. 몇몇 실시형태에서, $CcMan5_{1-774}$ 의 삼차원 구조는, 2Å 보다 많지 않은(예를 들어, 1.5Å, 1.0Å, 또는 0.5Å 보다 많지 않은) 상기 아미노산의 보존된 백본 원자(conserved backbone atoms)로부터의 ± 평균 평방근편차(root mean square deviation), PDB ID No. 2xs에 따른 아미노산의 완전한 구조적인 좌표를 포함한다. 본원에 사용된 바와 같이, "평균 평방근편차(root mean square deviation)"는 평균으로부터의 편차의 제곱의 산술평균의 제곱근이고(the square root of the arithmetic mean of the squares of the deviations from the mean), 이는 본원에서 기재된 구조적인 좌표로부터의 편차(deviation) 또는 변화(variation)를 표현하는 방법이다. 본원의 사실은 평균 평방근편차 내의 동일한 구조적인 좌표를 야기하는 언급된 아미노산 잔기의 보존적인 치환(conservative substitutions)을 포함하는 모든 실시형태를 포함한다.

[0077] 본원에서 제공된 구조적인 좌표는 만노시다제 폴리펩티드의 삼차원 구조를 나타내는데에 사용될 수 있다. 이러한 구조로부터, 예를 들어, 기질 결합 부위(substrate binding sites)는, 컴퓨터를 사용하여 시각화할 수 있고, 확인할 수 있고, 분자의 표면 주고, 표면 전하, 입체의 배열(steric arrangement), 반응성 아미노산, 소수성(hydrophobicity) 또는 친수성(hydrophilicity) 등의 영역의 존재를 기초로 특징지을 수 있다.

[0079] 도 33, 도 34, 또는 PDB ID No. 2xs 에 나타낸 바와 같은 본원에서 기재된 구조에 대해 발생된 구조적인 좌표를 사용하기 위해서, 관련 있는 좌표(relevant coordinates)는 삼차원 모양(dimensional shape) 또는 그래픽 표시(graphical representation)로서 나타낼 수 있거나, 또는 전환시킬 수 있다. 구조적인 좌표의 세트로부터 분자

또는 이의 일부의 삼차원의 그래픽 표시를 생성가능한 소프트웨어 프로그램을 상업적으로 입수할 수 있다. 상업적으로 입수가능한 소프트웨어 프로그램의 예는 하기를 포함하지만 이로 제한되지 않는다: GRID (Oxford University, Oxford, UK) ; MCSS (Molecular Simulations, San Diego, CA); AUTODOCK (Scripps Research Institute, La Jolla, CA); DOCK (University of California, San Francisco, CA); Flo99 (Thistlesoft, Morris Township, NJ); Ludi (Molecular Simulations, San Diego, CA); QUANTA (Molecular Simulations, San Diego, CA); Insight (Molecular Simulations, San Diego, CA); SYBYL (TRIPOS, Inc., St. Louis. MO); 및 LEAPFROG (TRIPOS, Inc., St. Louis, MO).

[0081] 본원에서 기재된 구조적인 좌표는, 분자 또는 분자 복합체의 알려지지 않은 삼차원의 구조를 결정하기 위해 표준 상동성 모델링 기술(standard homology modeling techniques)과 함께 사용될 수 있다. 상동성 모델링은, 하나 또는 그 이상의 관련된 단백질 분자, 분자 복합체 또는 이의 일부의 구조적인 좌표를 사용한 알려지지 않은 구조의 모델을 구성하는 것을 포함한다. 상동성 모델링은 단백질의 공통 또는 상동의 부분(common or homologous portions)을 맞춤으로써 처리할 수 있고, 이의 삼차원 구조는, 명확하게 본원의 도 33 및 도 34 에 제공된 관련된[즉, 상동의(homologous)] 구조적인 좌표를 사용하여, 알려진 분자에서의 상동의 구조적인 요인의 삼차원 구조를 해결될 것이다. 상동성은, 아미노산 서열 동일성, 상동성 2차 구조 요소(homologous secondary structure elements), 및/또는 상동성 제3의 접힘(homologous tertiary folds)을 사용하여 결정될 수 있다. 상동성 모델링은, 해결된 관련된 구조의 아미노산에 의해 아미노산(또는 그 밖의 성분)의 교체(replacement)와 함께 삼차원 구조의 모두 또는 재조립된 부분(rebuilding part)을 포함할 수 있다. 이에 따라서, 알려지지 않은 분자의 삼차원 구조는 본원에 기재된 $CcMan5_{1-774}$ 의 삼차원 구조를 사용하여 생성될 수 있고, 본원에 잘 알려진 기술을 사용하여 개량할 수 있다(refined).

[0083] 본원에 기재된 삼차원 구조에 근거하여, 치환(substitutions)은, 이들의 선택성(selectivity)을 개선하거나 또는 변경시키기 위해 $CcMan5_{1-774}$ 또는 그 밖의 만노시다제의 몇몇의 원자 또는 사이드 기(side groups)를 만들 수 있다. 예를 들어, CcMan5 는 위치 536 및 588 에서의 비-산성 잔기(non-acidic residue)를 포함하고, 이는 만노시다제가 Man-P-Man 기질에서의 아노머 산소(anomeric oxygen)에 인산염 결합(phosphate linkage)을 허용하는(tolerate) 것을 가능하게 한다. 이와 마찬가지로, 그 밖의 만노시다제에서의 이에 상응하는 잔기는, Man-P-Man 기질을 수용하기 위해(accept) 만노시다제의 능력을 증가시키도록 비-산성 잔기로 변경될 수 있다.

[0085] 본원에서 사용하기에 적절한 그 밖의 만노시다제 폴리펩티드 후보물질은 뉴클레오티드 및 폴리펩티드 서열의 정렬(alignments)의 분석으로 확인될 수 있다. 예를 들어, 뉴클레오티드 또는 폴리펩티드 서열의 데이터베이스(database)에서의 의문(query)을 실행하는 것은 만노시다제 폴리펩티드의 상동성 및/또는 오솔로그(orthologs)을 확인할 수 있다. 서열 분석은 알려진 만노시다제 아미노산 서열을 사용한 불필요하지 않는 데이터베이스(nonredundant databases)의 BLAST, 상호간의(Reciprocal) BLAST, 또는 PSI-BLAST 분석을 포함할 수 있다. 40 % 이상의 서열 동일성을 갖는 데이터베이스에서의 이러한 폴리펩티드는 만노시다제 폴리펩티드로서의 적합성(suitability)을 위한 추가적인 평가를 위한 후보물질로서 확인할 수 있다. 아미노산 서열의 유사성(similarity)은, 하나의 소수성 잔기를 다른 잔기로의 치환, 또는 하나의 극성 잔기를 다른 잔기로의 치환과 같은, 보존적 아미노산 치환(conservative amino acid substitutions)을 가능하게 한다. 만약 원한다면, 이러한 후보물질의 수동 조사(manual inspection)는 추가적으로 평가될 수 있는 후부물질의 수를 줄이기 위해 실행될 수 있다. 수동 조사는, 말단 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 만노시다제에 존재하는 것으로 혐의를 받는 도메인, 예를 들어 하나 또는 그 이상(예를 들어, 1, 2, 3, 4 또는 그 이상)의 보존된 도메인 또는 기능적인 영역[예를 들어, 기질 결합 캐비티(substrate binding cavity)]을 갖는 것으로 나타낸 이러한 후보물질을 선별하여 수행될 수 있다. 이러한 도메인은 글리신-리치 모티프(glycine-rich motif) GVGXXGXGG를 포함할 수 있고, 여기에서 X 는 Gly, Ser, Thr, Val, Ala, Cys 또는 Gln [또는 작은 곁사슬(small side chain)을 갖는 그 밖의 아미노산]이다. 이러한 모티프는 SEQ ID NO:50 의 잔기 69 내지 77 에서 발견되었다. 이러한 영역은 효소의 활성 부위에서의 -1 만노스 및 포스페이트-결합 서브 사이트(phosphate-binding subsite)에 대한 필수적인 수소 결합을 제공하는 루프(loop)를 만든다.

[0087] 보전된 모티프의 또 다른 예는 VRXE 모티프를 포함하고, 여기에서 Arg (R)은 -1 고리(ring) 및 가능한 +1 고리에 수소 결합을 생성하고, Glu (E)은, 아마도 이러한 모티프의 모양을 형성하면서 이러한 R 잔기에 대한 염다리(salt bridge)에 있고; 및 X 는 Trp 또는 Pro 를 제외한 20 개의 아미노산 중 하나이다. 이러한 모티프는 SEQ ID NO:50 의 잔기 404 내지 407 에서 발견된다.

[0089] 적절한 모티프는 또한 X_1YQGX_2 모티프일 수 있고, 여기에서 X_1 는 Leu, Ile, Val Ala, Phe, Tyr 또는 Met 일 수

있고, X_2 는 Thr, Ser 또는 Asn 일 수 있다. 이러한 모티프는 SEQ ID NO:50 의 잔기 534 내지 538 에서 발견된다. E 가 올리고당에서의 말단 만노스-1-포스포-6-만노스 결합을 가수분해하는 능력을 갖지 않는 만노시다제에 존재하기 때문에, 이러한 모티프에서의 Gln(Q)는 중요하다. 이러한 모티프에서의 Tyr (Y)는 또한 +1 사이트 형성에 중요한 것으로 예상된다.

[0091] 게다가, SEQ ID NO:50 의 잔기 22, 25, 71, 72, 195, 196, 354, 405, 535, 536, 588, 589, 626, 658, 660, 및 662 에 나타낸 영역은 CcMan5 의 기질 결합 캐비티(substrate binding cavity)를 형성한다. 최소한의 필요조건으로서, G71, G72, D355, R405, Q536, N588, Q589, T626, D660, D662 는 촉매 중심을 형성하고, 여기에서 N588, Q589 및 D660 은 촉매 Ca2+ 이온을 조정하는 것(coordinating)을 포함하고, D662 및 D660 은 친핵성 물(nucleophilic water)을 활성화시키는 것을 포함하고, Q536 은 전이 상태(transition state)동안 아노머 산소(anomeric oxygen)를 안정화시키고, G71, G71, D355, R405 및 T626 은 -1 사이트에서의 기질 결합을 포함한다. 이러한 최소의 촉매 중심을 나타내는 도 30 을 참고하라. 이와 같이, 아미노산 곁사슬에서의 원자의 삼차원 단백질 좌표가 도 33 에서의 등가의 원자(equivalent atoms)의 좌표의 1.5 Å 편차(deviation) 내의 범위에 들어가는 최소의 촉매중심(minimal catalytic center)(예를 들어, 도 30에 나타냄)에 위치하는 경우에, 만노시다제는 말단 만노스-1-포스포-6-만노스 결합을 가수분해할 수 있는 후보물질 만노시다제로서 선별될 수 있다.

[0093] 보존된 모티프는 또한 단백질의 N-말단 도메인에 GDXGN 모티프일 수 있고, 여기에서 X 는 P 를 제외한 모든 아미노산 일 수 있다. 이러한 모티프는 SEQ ID NO:50 의 잔기 21 내지 25 에서 발견되고, 도 24 에 나타낸 바와 같은 효소의 기질 결합 포켓(substrate binding pocket)의 일부를 형성한다. 특히, D 및 N의 곁사슬은 기질 결합 캐비티와 라인을 형성하고(line), 이는 +1 만노스에 결합하는 대체가능한 서브포켓(alternative subpocket)을 형성할 수 있다.

[0095] 실시예 14 에 나타낸 바와 같이, 폴리펩티드 서열의 데이터베이스에서의 의문(query)의 실행(performing)은, 하기의 유기체에서의 CcMan5 의 상동성을 확인하였다: *스트렙토마이세스 코엘리컬러(Streptomyces coelicolor)*[GenBank 수납 번호Accession No. NP_630514], *스트렙토마이세스 리비단스(Streptomyces lividans)*(GenBank 수납 번호 ZP_05522540); *스트렙토마이세스 리비단스(Streptomyces lividans)*(GenBank 수납 번호 ZP_06527366); *클로스트리디움 스피로포르메(Clostridium spiroforme)*(GenBank 수납 번호 ZP_02866543), *박테로이데스 테타이오타오미크론(Bacteroides thetaiotaomicron)*(GenBank 수납 번호 NP_812442), *주농왕기아 프로푼다(Zunongwangia profunda)*(GenBank 수납 번호 YP_003584502); *키티노파가 피넨시스(Chitinophaga pinensis)*(GenBank 수납 번호 YP_003120664); *페니바실러스 sp(Paenibacillus sp)*(GenBank 수납 번호 YP_003013376); *박테로이데스 sp(Bacteroides sp.)*(GenBank 수납 번호 ZP_04848482); *박테로이데스 셀룰로실리티쿠스(Bacteroides cellulosilyticus)* (GenBank 수납 번호 ZP_03677957); *레이우엔훅키엘라 블란데니시스(Leeuwenhoekiella blandensis)*(GenBank 수납 번호 ZP_01061975); *스핑고박테리움 스피리티보럼(Sphingobacterium spiritivorum)*(GenBank 수납 번호 ZP_07083984); 및 *페도박터 sp.(Pedobacter sp.)*(GenBank 수납 번호 ZP_01885202). *스트렙토마이세스 코엘리컬러(Streptomyces coelicolor)* 및 *스트렙토마이세스 리비단스(Streptomyces lividans)*로부터의 만노시다제는, 상기의 모티프에서 뿐만 아니라 삼차원 구조의 많은 루프에서 유사하다(BLASTP 에 의해 765 배열된 잔기 상에서의 501 동일성을 갖는, CcMan5 GH92 도메인과 66 % 의 서열 동일성).

[0097] 만노시다제 폴리펩티드를 코드화하는 분리된 핵산 분자는 표준 기술에 의해 생산될 수 있다. 예를 들어, 중합효소 연쇄 반응(polymerase chain reaction, PCR)은 본원에 기재된 뉴클레오티드 서열을 포함하는 분리된 핵산을 수득하는데 사용될 수 있다. PCR 은 전체 게놈의(total genomic) DNA 또는 전체 세포의(total cellular) RNA 로부터의 서열을 포함하는, RAN 뿐만 아니라 DNA 로부터의 특정한 서열을 증폭시키기 위해 사용될 수 있다. 일반적으로, 관심 영역의 끝(beyond) 또는 그 이상(beyond)으로부터의 서열 정보는, 증폭시킬 주형(template)의 반대 가닥(opposite strands)에 대한 서열에서 동일하거나 유사한 올리고뉴클레오티드 프라이머(oligonucleotide primers)를 설계하는데 사용되었다. 부위-특이적 뉴클레오티드 서열 변형(site-specific nucleotide sequence modifications)에 의해 또한 가능한 다양한 PCR 전략(strategies)은 주형 핵산 내로 도입될(introduced) 수 있다. 분리된 핵산은 또한, 단일 핵산 분자[포스포라미디트 기술(phosphoramidite technology)을 사용한 3' 에서 5' 방향으로의 자동화된 DNA 합성을 사용함] 또는 올리고뉴클레오티드의 시리즈(series)로서 화학적으로 합성될 수 있다. 예를 들어, 긴 올리고뉴클레오티드의 하나 또는 그 이상의 한 쌍(pairs)(예를 들어, >100 뉴클레오티드)은, 올리고뉴클레오티드 쌍(pair)이 풀렸을 때(anneal) 듀플렉스(duplex)가 형성되는 것과 같이 상보성(complementarity)의 짧은 세그먼트(short segment)(예를 들어, 약 15 뉴클레오티드)를 포함하는 각각의 쌍을 갖는, 원하는 서열을 포함하면서 합성할 수 있다. DNA 중합효소는, 벡

터 내로 연결시킬(ligated) 수 있는, 뉴클레오티드 쌍 당 단일, 이중-가닥 핵산 분자(single, double-stranded nucleic acid molecule per oligonucleotide pair)를 야기하는, 올리고뉴클레오티드를 연장시키는데 사용된다. 본 발명의 분리된 핵산은 예를 들어 자연적으로 발생하는 DNA 의 돌연변이유발에 의해 또한 수득될 수 있다.

[0099] 본 문서는, 본원에 기재된 만노시다제의 (i) 생물학적으로 유효한 변이체(biologically active variants) 및 (ii) 이의 생물학적으로 유효한 단편 또는 생물학적으로 유효한 변이체를 또한 제공한다. 만노시다제의 생물학적으로 유효한 변이체는 SEQ ID NOs: 7, 9, 11, 13, 15, 또는 50 에 나타낸 서열에 관하여 첨가(additions), 결실(deletions) 또는 치환을 포함할 수 있다. 치환을 갖는 단백질은 50 보다 많지 않은(예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20, 25, 30, 35, 40, 또는 50 보다 많지 않은) 보존적 아미노산 치환(conservative amino acid substitutions)을 일반적으로 가질 수 있을 것이다. 보존적 치환은 하나의 아미노산이 유사한 특성을 갖는 다른 아미노산으로의 치환이다. 보존적 치환은 하기의 군 내의 치환을 포함한다: 발린(valine), 알라닌(alanine) 및 글리신(glycine); 류신(leucine), 발린(valine), 및 이소류신(isoleucine); 아스파르산(aspartic acid) 및 글루탐산(glutamic acid); 아스파라진(asparagine) 및 글루타민(glutamine); 세린(serine), 시스테인(cysteine), 및 트레오닌(threonine); 리신(lysine) 및 아르기닌(arginine); 및 페닐알라닌(phenylalanine) 및 티로신(tyrosine). 무극성(non-polar) 소수성 아미노산은 알라닌, 류신, 이소류신, 발린, 프롤린, 페닐알라닌, 트립토판 및 메티오닌을 포함한다. 극성 중성(polar neutral) 아미노산은 글리신, 세린, 트레오닌, 시스테인, 티로신, 아스파라진(asparagine) 및 글루타민을 포함한다. 양전하를 띄는(positively charged) [염기(basic)] 아미노산은 아르기닌, 리신 및 히스티딘을 포함한다. 음전하를 띄는 (산성의) 아미노산은 아르파르트산(aspartic acid) 및 글루탐산(glutamic acid)을 포함한다. 상기에 언급된 극성, 염기 또는 산성 그룹의 하나의 멤버(member)가 동일한 그룹의 또 다른 멤버로의 어떠한 치환은 보존적 치환(conservative substitution)으로 여길 수 있다. 대조적으로, 비-보존적 치환(non-conservative substitution)은 하나의 아미노산을 다른 특성을 갖는 또 다른 아미노산으로의 치환이다. 도 31 및 32 에 나타낸 서열 배열은 생산될 수 있는 수많은 아미노산 치환의 예를 제공한다.

[0101] 결실 변이체(Deletion variants)는 (둘 또는 그 이상의 아미노산의) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20 아미노산 세그먼트(segments) 또는 근접하지 않은(non-contiguous) 단일 아미노산이 결실될 수 있다.

[0103] 첨가[첨가 변이체(addition variants)]는 하기를 포함하는 융합 단백질(fusion proteins)을 포함한다: (a) SEQ ID NOs: 7, 9, 11, 13, 또는 15, 또는 이의 단편에 나타낸 만노시다제; 및 (b) 내부의(internal) 또는 말단 (C 또는 N)의 무관한(irrelevant) 또는 이종의(heterologous) 아미노산 서열. 이러한 융합 단백질의 문맥(context)에서, 용어 "이종의 아미노산 서열(heterologous amino acid sequences)"은 (a) 이외의 아미노산 서열을 나타낸다. 이종의 서열은, 예를 들어 재조합 단백질[예를 들어, FLAG, 폴리히스티딘(polyhistidine)(예를 들어 헥사히스티딘(hexahistidine), 헤마글루타닌(hemagluttanin, HA), 글루타티온 S-전달효소(glutathione-S-transferase, GST), 또는 말토스-결합 단백질(maltose-binding protein, MBP)]의 정제에 사용된 서열일 수 있다. 이종의 서열은 또한, 진단의 또는 발견될 수 있는 마커(diagnostic or detectable markers), 예를 들어 루시페라이제(luciferase), 녹색 형광 단백질(green fluorescent protein, GFP), 또는 클로람페니콜아세틸전달효소(chloramphenicol acetyl transferase, CAT)로서 유용한 단백질일 수 있다. 몇몇 실시형태에서, 융합 단백질은 또 다른 단백질로부터의 신호 서열을 포함한다. 특정 숙주 세포(예를 들어, 효모 숙주 세포)에서의, 발현 및/또는 타겟 단백질의 발현 및/또는 분비는 이종의 신호 서열의 사용을 통해 증가시킬 수 있다. 몇몇 실시형태에서, 융합 단백질은 운반체(carrier)(예를 들어, KLH) 예를 들어 항체 생성을 위한 면역 반응을 이끌어내는데에 유용한) 또는 소포체(endoplasmic reticulum) 또는 골지체 보유 신호(Golgi apparatus retention signals)를 포함할 수 있다. 이종의 서열은 다양한 길이일 수 있고, 몇몇의 경우에 이종의 서열은, 이종의 서열에 부착된 전장의 타겟 단백질보다 더 긴 서열일 수 있다.

[0105] 만노시다제의 생물학적으로 유효한 단편 또는 생물학적으로 유효한 변이체는, 야생형, 전장, 숙성된 단백질(wild-type, full-length, mature protein)의 만노시다제 활성도[예를 들어, M6P 잔기의 캡핑이 제거됨(uncapping)]의 적어도 40 % (예를 들어, 적어도 50 %; 60 %; 70 %; 75 %; 80 %; 85 %; 90 %; 95 %; 97 %; 98 %; 99 %; 99.5 %, 또는 100 % 또는 보다 많은)를 갖는다. 예를 들어, 만노시다제의 생물학적으로 유효한 단편은 SEQ ID NO:50 의 잔기 1 내지 774 를 포함할 수 있다.

[0107] 본원에 기재된 만노시다제는 캡핑이 제거된(Uncapping) 말단 포스포-6-만노스(M6P) 잔기를 갖는 분자(예를 들어, 타겟 단백질)을 생산하는데 사용될 수 있다. 방법은 *생체외* 또는 *생체 내*에서 실행될 수 있다.

[0109] *캡핑이 제거된 M6P 잔기의 생체외 방법*

[0110] 본원에 기재된 만노시다제는 올리고당에서의 말단 M6P 잔기의 캡을 제거하도록(to uncap terminal M6P residues on oligosaccharides) *생체외*에서 재조합적으로 생산될 수 있고, 사용될 수 있다. 재조합적으로 만노시다제를 생산하기 위해, 벡터는, 만노시다제 폴리펩티드를 코드화하는 핵산에 실시가능하게 연결된 프로모트를 포함하는 것을 사용한다. 본원에서 사용된 바와 같은, "프로모터"는 유전자가 전사될 수 있은 DNA 서열을 나타낸다. 프로모터는 RNA 중합효소에 의해 인지된 다음에 전사(transcription)를 개시한다. 따라서, RNA 중합효소의 DNA 서열에 직접적으로 결합하거나 또는 유입(recruitment)에 포함된 DNA 서열을 포함한다. 프로모터 서열은 "인핸서 영역(enhancer regions)"을 또한 포함하고, 이는 유전자 군(gene-cluster)에서 유전자의 전사 수준을 증진시키기 위해[이름이 여기서 유래함], 단백질[즉, 전사 인자(transcription factors)의 세트와 같은, 작용 전달 인자(trans-acting factors)]과 함께 결합될 수 있는 DNA의 하나 또는 그 이상의 영역이다. 일반적으로 코딩 영역(coding region)의 5' 말단에 있는 인핸서(enhancer)는 프로모터 서열로부터 또한 분리될 수 있고, 즉, 유전자의 인트론 영역(intronic region) 또는 유전자의 코딩 영역(coding region)의 3' 일 수 있다.

[0112] 본원에 사용된 바와 같이, "실시가능하게 연결된(operably linked)"은, 발현 조절 서열(expression control sequences)이 관심 코딩 서열(coding sequence)의 발현을 효율적으로 조절하기 위해, 유전적 구조(genetic construct)(예를 들어 벡터) 내로 포함됨을 의미한다.

[0114] 발현 벡터는, 코드화된(encoded) 폴리펩티드의 발현을 위해 숙주 세포 내로 도입될 수 있고[예를 들어, 형질전환(transformation) 또는 트랜스펙션(transfection)], 그 다음에 정제될 수 있다. 만노시다제 폴리펩티드의 소규모 또는 대규모 생산(small or large scale production)을 위해 사용될 수 있는 발현 시스템(Expression systems)은, 재조합 박테리오파지(recombinant bacteriophage) DNA, 플라스미드 DNA 또는 핵산 분자를 포함하는 코스미드(cosmid) DNA 발현 벡터로 형질전환된(transformed) 박테리아와 같은 박테리아(예를 들어, *E. coli*)와 같은 미생물, 및 핵산 분자를 포함하는 재조합 균류 발현 벡터(recombinant fungal expression vectors)로 형질전환된 균류(fungal)[예를 들어, *S. 세레비시아(S. cerevisiae)*, *야로위아 리폴리티카(Yarrowia lipolytica)*, *아르술라 아데니니보란스(Arxula adeninivorans)*, *피키아 파스토리스(Pichia pastoris)*, *한세눌라 폴리모르파(Hansenula polymorpha)*, 또는 *아스페르길루스(Aspergillus)*]를 포함하지만 이로 한정되지 않는다. 유용한 발현 시스템은, 핵산 분자를 포함하는 재조합 바이러스 발현 벡터로 전파된(infected)[예를 들어, 배큘로 바이러스(baculovirus)] 곤충 세포 시스템, 및 핵산 분자를 포함하는 재조합 바이러스 발현 벡터[예를 들어, 담배 모자이크 바이러스(tobacco mosaic virus)]로 전파되거나(infected) 또는 재조합 플라스미드 발현 벡터(예를 들어 Ti 플라스미드)로 형질전환된 식물 세포 시스템(plant cell systems)을 또한 포함한다. 만노시다제 폴리펩티드는 포유동물 발현 시스템을 사용하여 또한 생산될 수 있고, 이는, 본원에서 기재된 핵산과 함께, 포유동물 세포의 게놈(genome)[예를 들어, 메탈로치오네인 프로모터(metallothionein promoter)] 또는 포유동물 바이러스(mammalian viruses)[예를 들어, 아데노바이러스 후발 프로모터(adenovirus late promoter) 및 사이토메가로바이러스 프로모터(cytomegalovirus promoter)]로부터 유도된 프로모터를 포함하는 재조합 발현 구조물(recombinant expression constructs)을 포함하는(harboring) 세포[예를 들어, COS 세포, 중국 햄프서 난소 세포주(Chinese hamster ovary cells), HeLa 세포, 인간 배아 콩팥 293 세포(human embryonic kidney 293 cells) 및 3T3 L1 세포와 같은 불멸화된 세포주(immortalized cell lines)]를 포함한다.

[0116] 일반적으로, 재조합 만노시다제 폴리펩티드는, 단백질의 정제를 지원하도록 FLAG, 폴리히스티딘(예를 들어, 헥사히스티딘), 헤마글루타닌(HA), 글루타티온 S-전달효소(GST), 또는 말토스-결합 단백질(MBP)과 같은 이종의 아미노산 서열로 태그를 붙였다(tagged). 정제된 단백질에 대한 그 밖의 방법은, 이온 교환(ion exchange), 소수성(hydrophobic) 및 역상(reverse phase), 사이즈 배제(size exclusion), 친화도(affinity), 소수성 전하-도입 크로미토그래피(hydrophobic charge-induction chromatography)와 같은 크로마토그래피 기술 및 이와 같은 것(예를 들어, Scopes, Protein Purification: Principles and Practice, third edition, Springer-Verlag, New York (1993); Burton and Harding, *J. Chromatogr. A* 814:71-81 (1998)을 참고하라)을 포함한다.

[0118] 생체외 캡핑이 제거된(uncapped) 말단 M6P 잔기를 갖는 분자를 생산하기 위해, 만노스-1-포스포-6 만노스 결합을 포함하는 타겟 분자는, 재조합적으로 생산된 만노시다제를 포함하는 세포 용해액(cell lysate) 또는 정제된 만노시다제와 함께 적절한 조건 하에서 접촉된다. 세포 용해액은 균류 세포, 식물 세포 또는 동물 세포를 포함하는 어떠한 유전적으로 조작된 세포일 수 있다. 동물 세포의 한정되지 않는 예는 선충(nematode), 곤충(insect), 식물(plant), 새(bird), 파충류(reptile), 및 마우스(mouse), 랫(rat), 토끼, 햄스터, 모래쥐(gerbil), 개, 고양이, 염소, 돼지, 소, 말, 고래(whale), 원숭이, 또는 인간과 같은 포유동물을 포함한다. 정제된 만노시다제 또는 세포 용해액과 타겟 분자(예를 들어, 올리고당 또는 당단백질)를 접촉할 때, 만노시다제

는 만노스-1-포스포-6 만노스 결합을 가수분해하고, 하나 또는 그 이상의 캡핑이 제거된 말단 M6P 잔기를 갖는 타겟 분자를 생산한다. 실시예 2 에 기재된 방법은, 말단 M6P 잔기가 캡이 제거되는지를(uncapped) 측정하기 위해 사용될 수 있다. 만노시다제에 의한 공정 다음에, 캡핑이 제거된 말단 M6P 잔기를 갖는 타겟 분자를 분리할 수 있다.

[0120] 용해액(lysate)에서의 만노시다제 활성도의 온전함(integrity) 또는 활성도의 보전된 세포 용해액을 수득하기 위한 적절한 방법은, 세포 용해액에서의 N-글리코실화 활성도(N-glycosylation activities)에서의 변화를 최소화하거나 보존하는 적절한 완충용액(buffers) 및/또는 뉴클레아제(nuclease), 프로테아제(protease) 및 포스파타제 저해제(phosphatase inhibitors)를 포함하는 저해제(inhibitors)의 사용을 포함한다. 이러한 저해제는, 예를 들어 에틸렌디아민테트라아세트산(ethylenediamine tetraacetic acid, EDTA)와 같은 킬레이터(chelator), 에틸렌 글리콜 비스(P-아미노에틸 에테르) N,N,N1,N1-테트라아세트산(ethylene glycol bis(P-aminoethyl ether) N,N,N1,N1-tetraacetic acid, EGTA), 페닐메틸술포닐 플로라이드(phenylmethylsulfonyl fluoride, PMSF)와 같은 프로테아제 저해제(protease inhibitors), 아프로티닌(aprotinin), 류펩틴(leupeptin), 안티팬(antipain) 등, 및 포스페이트와 같은 포스파타제 저해제(phosphatase inhibitors), 플루오린화나트륨(sodium fluoride), 바나듐산염(vanadate) 등을 포함한다. 효소적 활성도를 포함하는 용해액(lysates)을 수득하기 위한 적절한 완충용액 및 조건은, 예를 들어 Ausubel et al. Current Protocols in Molecular Biology (Supplement 47), John Wiley & Sons, New York (1999); Harlow and Lane, Antibodies: A Laboratory Manual Cold Spring Harbor Laboratory Press (1988); Harlow and Lane, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Press (1999); Tietz Textbook of Clinical Chemistry, 3rd ed. Burtis and Ashwood, eds. W.B. Saunders, Philadelphia, (1999)에 기재되어 있다.

[0122] 세포 용해액은, 적절히, 물질(substances)의 간섭(interfering)의 존재를 제거하거나 또는 최소화하기 위해 추가적으로 처리할 수 있다. 만약 원한다면, 세포 용해액은, 초원심세포분획법(subcellular fractionation) 및 이온 교환(ion exchange), 소수성(hydrophobic) 및 역상(reverse phase), 사이즈 배제(size exclusion), 친화도(affinity), 소수성 전하-도입 크로미토그래피(hydrophobic charge-induction chromatography)와 같은 크로마토그래피 기술 등을 포함하는 본 분야의 숙련자에서 알려진 다양한 방법으로 분류될(fractionated) 수 있다.

[0124] 몇몇 실시형태에서, 세포 용해액은, 전체 세포의 세포소기관(whole cellular organelles)이 온전하고 및/또는 기능적으로 잔존하는데에서 제조될 수 있다. 예를 들어, 용해액은 하나 또는 그 이상의 온전한 조면 소포체(rough endoplasmic reticulum), 온전한 활면소포체(smooth endoplasmic reticulum) 또는 온전한 골지체(Golgi apparatus)를 포함할 수 있다. 온전한 세포의 세포 소기관을 포함하는 용해액을 제조하기 위한 적절한 방법 및 세포소기관의 기능성을 테스트하기 위한 적절한 방법은, 예를 들어 Moreau *et al.* (1991) *J. Biol. Chem.* 266(7):4329-4333; Moreau *et al.* (1991) *J. Biol. Chem.* 266(7):4322-4328; Rexach *et al.* (1991) *J. Cell Biol.* 114(2):219-229; and Paulik *et al.* (1999) *Arch. Biochem. Biophys.* 367(2):265-273에 기재되어 있다.

[0126] 본원에 기재된 바와 같은 타겟 분자는, 균류 기원(fungal origin)의 세포에서 발현되었을 때, 만노스-1-포스포-6 만노스 결합을 포함하는 어떠한 분자, 또는 말단 만노스-1-포스포-6 만노스 결합을 포함하는 어떠한 분자를 나타낸다. 적절한 타겟 단백질은, 파상풍톡소이드(tetanus toxoid) 또는 디프테리아 톡소이드(diptheria toxoid)와 같은 병원체 단백질(pathogen proteins); 사이트메가로바이러스(CMV) 당단백질 B, H 및 gCIII, 인간 면역 결핍 바이러스 1(human immunodeficiency virus 1, HIV-1) 외막 당단백질(envelope glycoproteins), 라우스 육종 바이러스(Rous sarcoma virus, RSV) 외막 당단백질, 단순헤르페스 바이러스(herpes simplex virus, HSV) 외막 당단백질, 엡스타인 바 바이러스(Epstein Barr virus, EBV) 외막 당단백질, 수두대상포진 바이러스(varicella-zoster virus, VZV) 외막 당단백질, 인체유두종 바이러스(human papilloma virus, HPV) 외막 당단백질, 인플루엔자바이러스 당단백질(Influenza virus glycoproteins), 및 간염 패밀리 표면 항원(Hepatitis family surface antigen)과 같은 바이러스성 표면 단백질(viral surface proteins); 리소좀 단백질(lysosomal proteins)[예를 들어, 산성 알파 글리코시다제(acid alpha glucosidase), 알파 갈락토시다제(alpha galatosidase), 글루코세레브로시다제(glucocerebrosidase), 셀레브로시다제(cerebrosidase), 또는 갈락토세레브로시다제(galactocerebrosidase)]; 인슐린; 글루카곤(glucagons) ; 성장 인자(growth factors) ; 사이토카인(cytokines) ; 케모카인(chemokines); 및 항체 또는 이의 단편을 포함한다. 성장 인자는, 예를 들어 혈관내피 성장인자(vascular endothelial growth factor, VEGF), 인슐린-유사 성장 인자(Insulin-like growth factor, IGF), 골 형성 단백질(bone morphogenic protein, BMP), 과립구-집락 자극 인자(Granulocyte-colony

stimulating factor, G-CSF), 과립대식세포-집락 자극 인자(Granulocyte-macrophage colony stimulating factor, GM-CSF), 신경 성장 인자(Nerve growth factor, NGF); 뉴로트로핀(Neurotrophin), 혈소판-유래 성장 인자(Platelet-derived growth factor, PDGF), 에리스로포에틴(Erythropoietin, EPO), 트롬보포이에틴(Thrombopoietin, TPO), 미오스타틴(Myostatin) (GDF-8), 성장 분화 인자-9(Growth Differentiation factor-9, GDF9), 염기성 섬유모세포 성장 인자(basic fibroblast growth factor, bFGF 또는 FGF2), 표피 성장 인자(Epidermal growth factor, EGF), 간세포 성장 인자(Hepatocyte growth factor, HGF)를 포함한다. 사이토카인은, 예를 들어 IL-1 내지 IL-33 (예를 들어, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IL-13, 또는 IL-15)과 같은 인터류킨(interleukins)을 포함한다. 케모카인은, 예를 들어 I-309, TCA-3, MCP-1, MIP-1α, MIP-1β, RANTES, C10, MRP-2, MARC, MCP-3, MCP-2, MRP-2, CCF18, MIP-1ɣ, 이오탁신(Eotaxin), MCP-5, MCP-4, NCC-1, Ckβ10, HCC-1, 루코탁신-1(Leukotactin-1), LEC, NCC-4, TARC, PARC, 또는 이오탁신-2(Eotaxin-2)를 포함한다. 또한 종양 당단백질[예를 들어, 종양-관련 항원(tumor-associated antigens)]은, 예를 들어 발암배아성 항원(carcinoembryonic antigen, CEA), 인간 무친(human mucins), HER-2/neu, 및 전립선-특이 항원(prostate-specific antigen, PSA)을 포함한다[Henderson and Finn, *Advances in Immunology*, 62, pp. 217-56 (1996)].

[0128] 몇몇 실시형태에서, 타겟 단백질은 리소좀 축적 질환(lysosomal storage disorder)과 관련된 하나일 수 있고, 타겟 단백질은, 예를 들어 산성 알파 글루코시다아제(acid alpha glucosidase), 알파 갈락토시다아제(alpha galactosidase), 알파-L-이두로니다아제(alpha-L-iduronidase), 베타-D-갈락토시다아제(beta-D-galactosidase), 베타-갈락토시다아제(beta-glucosidase), 베타-헥소사미니다아제(beta-hexosaminidase), 베타-D-만노시다제(beta-D-mannosidase), 알파-L-푸코시다아제(alpha-L-fucosidase), 아릴설파타아제 B(arylsulfatase B), 아릴설파타아제A(arylsulfatase A), 알파-N-아세틸갈락토사미니다아제(alpha-N-acetylgalactosaminidase), 아스파르틸글루코사미니다아제(aspartylglucosaminidase), 인두로네이트-2-술파타아제(iduronate-2-sulfatase), 알파-글루코사미니드-N-아세틸전달효소(alpha-glucosaminide-N-acetyltransferase), 베타-D-글루코로니다아제(beta-D-glucoronidase), 히알루로니다아제(hyaluronidase), 알파-L-만노시다제(alpha-L-mannosidase), 알파-뉴라미니다아제(alpha-neuraminidase), 인산전달효소(phosphotransferase), 산성 리파아제(acid lipase), 산성 세라미다아제(acid ceramidase), 스핑고마이엘리나제(sphingomyelinase), 티오에스테라제(thioesterase), 카텝신 K(cathepsin K), 및 지단백질지방분해효소(lipoprotein lipase)를 포함한다.

[0130] 몇몇 실시형태에서, 타겟 단백질은 융합 단백질이고, 타겟 단백질은 그 밖의 폴리펩티드 서열에 융합되거나, 또는 중합체, 운반체(carrier), 아쥬반트(adjuvant), 면역독소(immunotoxin) 또는 측정가능한[예를 들어, 형광성(fluorescent), 발광성(luminescent) 또는 방사성(radioactive)] 일부(detectable moiety)에 융합된다. 예를 들어, 표적 단백질은, 소 단백질의 분자량을 증가시키기 위해 및/또는 순환 체류 시간(circulation residence time)을 증가시키기 위해, 폴리에틸렌글리콜과 같은 중합체에 합쳐질 수 있다.

[0132] *캡핑이 제거된 M6P 잔기의 생체내 방법*

[0133] 본원에서 기재된 유전적으로 조작된 세포는 캡핑이 제거된(uncapped) M6P 잔기를 포함하는 타겟 분자를 생산하는데 사용될 수 있다. 예를 들어, 세포 기저 방법(cell based method)은, 만노시다아제를 코드화하는 핵산, 타겟 분자를 코드화하는 핵산을 포함하도록 유전적으로 조작된 균류 세포 내로 도입하는 단계를 포함할 수 있고, 여기에서 세포는 캡핑이 제거된 말단 M6P 잔기를 포함하는 타겟 분자를 생산한다. 몇몇 실시형태에서, 만노시다제 및 타겟 분자를 코드화하는 핵산은, 만노시다제 및 타겟 분자가 공동-분비되는(co-secreted) 것과 같은 분비 신호(secretion sequence)를 포함한다.

[0135] 본원에 기재된 유전적으로 조작된 세포는 만노시다제를 코드화하는 핵산을 포함하고, 캡핑이 제거된(uncapped) 말단 M6P 잔기를 갖는 하나 또는 그 이상의 타겟 분자를 생산하는데 유용하다. 캡핑이 제거된 M6P 잔기의 *생체내* 생산에 적절한 세포는, *야로위아 리폴리티카(Yarrowia lipolytica), 아르술라 아데니니보란스(Arxula adeninivorans),* 메탄올자화효모(methylotrophic yeast)[*칸디다, 한세눌라, 오가테아, 피치아* 또는 *토룰롭시스(Torulopsis)* 속(genus)의 메탄올자화효모]를 포함하는, 균류 기원(fungal origin), 또는 *아스페르길루스, 트리코데르마(Trichoderma), 네우로스포라(Neurospora), 푸사리움(Fusarium),* 또는 *크라이소스포리움(Chrysosporium)* 속(genus)의 곰팡이(filamentous fungi)를 포함할 수 있다. 대표적인 균종(fungal species)은 하기를 포함하지만, 이로 한정되지 않는다: *피치아 아노말라(Pichia anomala), 피치아 보비스(Pichia bovis), 피치아 카나데니시스(Pichia canadensis), 피치아 카르소니(Pichia carsonii), 피치아 파리노스(Pichia farinose), 피치아 페르멘탄스(Pichia fermentans), 피치아 플룩숨(Pichia fluxuum), 피치아 멤브라네*

파시엔스(Pichia membranaefaciens), 피치아 멤브라네파시엔스(Pichia membranaefaciens), 칸디다 발리다(Candida valida), 칸디다 알비칸스(Candida albicans), 칸디다 아스칼라라피다룸(Candida ascalaphidarum), 칸디다 암피시아(Candida amphixiae), 칸디다 안타르티카(Candida Antarctica), 칸디다 아틀란티카(Candida atlantica), 칸디다 아트모스피리카(Candida atmosphaerica), 칸디다 블라테(Candida blattae), 칸디다 카르포릴라(Candida carpophila), 칸디다 세람비시다룸(Candida cerambycidarum), 칸디다 차울리오데스(Candida chauliodes), 칸디다 코릴달리스(Candida corydalis), 칸디다 도세이(Candida dosseyi), 칸디다 두블리니엔시스(Candida dubliniensis), 칸디다 에르가텐시스(Candida ergatensis), 칸디다 프록투스(Candida fructus), 칸디다 글라브라타(Candida glabrata), 칸디다 페르멘타티(Candida fermentati), 칸디다 퀠리어몬디(Candida guilliermondii), 칸디다 헤물로니(Candida haemulonii), 칸디다 인섹타멘스(Candida insectamens), 칸디다 인섹토룸(Candida insectorum), 칸디다 인터메디아(Candida intermedia), 칸디다 제프레시(Candida jeffresii), 칸디다 케피르(Candida kefyr), 칸디다 크루세이(Candida krusei), 칸디다 루시타니에(Candida lusitaniae), 칸디다 릭소소필라(Candida lyxosophila), 칸디다 말토사(Candida maltosa), 칸디다 멤브라니파시엔스(Candida membranifaciens), 칸디다 밀레리(Candida milleri), 칸디다 오레오필라(Candida oleophila), 칸디다 오레고넨시스(Candida oregonensis), 칸디다 파라프실로시스(Candida parapsilosis), 칸디다 퀘르시트루사(Candida quercitrusa), 칸디다 쉬하테(Candida shehatea), 칸디다 테모칠레(Candida temnochilae), 칸디다 테누이스(Candida tenuis), 칸디다 트로피칼리스(Candida tropicalis), 칸디다 트수치예(Candida tsuchiyae), 칸디다 시놀라보란티움(Candida sinolaborantium), 칸디다 소예(Candida sojae), 칸디다 비스와나티(Candida viswanathii), 칸디다 우틸리스(Candida utilis), 오가테아 미뉴타(Oogataea minuta), 피치아 멤브라네파시엔스(Pichia membranaefaciens), 피치아 실베스트리스(Pichia silvestris), 피치아 멤브라네파시엔스(Pichia membranaefaciens), 피치아 콘다티(Pichia chodati), 피치아 멤브라네파시엔스(Pichia membranaefaciens), 피치아 멤브라네파시엔스(Pichia menbranaefaciens), 피치아 미누스큘(Pichia minuscule), 피치아 파스토리스(Pichia pastoris), 피치아 프세우도폴리모르파(Pichia pseudopolymorpha), 피치아 퀘르쿰(Pichia quercuum), 피치아 로베르트시(Pichia robertsii), 피치아 사이토이(Pichia saitoi), 피치아 실베스트리시(Pichia silvestrisi), 피치아 스트라스부르겐시스(Pichia strasburgensis), 피치아 테리콜라(Pichia terricola), 피치아 반리지(Pichia vanriji), 프슈도지마 안타르시티카(Pseudozyma Antarctica), 로도스포리디움 토룰로이데스(Rhodosporidium toruloides), 로도토룰라 글루티니스(Rhodotorula glutinis), 사카로미세스 바야누스(Saccharomyces bayanus), 사카로미세스 바야누스(Saccharomyces bayanus), 사카로미세스 몸드쉬루쿠스(Saccharomyces momdshuricus), 사카로마이세스 우바룸(Saccharomyces uvarum), 사카로마이세스 바야누스(Saccharomyces bayanus), 사카로마이세스 세레비시아(Saccharomyces cerevisiae), 사카로마이세스 비스포루스(Saccharomyces bisporus), 사카로마이세스 체발리에리(Saccharomyces chevalieri), 사카로마이세스 델브뤠키(Saccharomyces delbrueckii), 사카로마이세스 에구구우스(Saccharomyces exiguous), 사카로마이세스 페르멘타티(Saccharomyces fermentati), 사카로마이세스 프라길리스(Saccharomyces fragilis), 사카로마이세스 마르시아누스(Saccharomyces marxianus), 사카로마이세스 멜리스(Saccharomyces mellis), 사카로마이세스 로세(Saccharomyces rosei), 사카로마이세스 로우실(Saccharomyces rouxii), 사카로마이세스 우바룸(Saccharomyces uvarum), 사카로마이세스 웰리아누스(Saccharomyces willianus), 사카로마이코데스 루드윌기(Saccharomycodes ludwigii), 사카로마이코프시스 카프술라리스(Saccharomycopsis capsularis), 사카로마이코프시스 피블리게라(Saccharomycopsis fibuligera), 사카로마이코프시스 피블리게라(Saccharomycopsis fibuligera), 엔도마이세스 호르데이(Endomyces hordei), 엔도마이코프시스 포불리게라(Endomycopsis fobuligera). 사투리니스포라 사이토이(Saturnispora saitoi), 스키조사키로마이세스 옥토스포러스(Schizosaccharomyces octosporus), 스키조사키로마이세스 폼베(Schizosaccharomyces pombe), 슈완니오마이세스 옥시덴탈리스(Schwanniomyces occidentalis), 토룰라스포라 델브루에키(Torulaspora delbrueckii), 토룰라스포라 델부루에키(Torulaspora delbrueckii), 사카로마이세스 다이렌시스(Saccharomyces dairensis), 토룰라스포라 델부루에키(Torulaspora delbrueckii), 토룰라스포라 페르멘타티(Torulaspora fermentati), 사카로마이세스 페르멘타티(Saccharomyces fermentati), 토룰라스포라 델부루에키(Torulaspora delbrueckii), 토룰라스포랄 로세이(Torulaspora rosei), 사카로마이세스 로세이(Saccharomyces rosei), 토룰라스포라 델부루에키(Torulaspora delbrueckii), 사카로마이세스 로세이(Saccharomyces rosei), 토루라스포라 델부루에키(Torulaspora delbrueckii), 토루라스포라 델부루에키(Saccharomyces delbrueckii), 토룰라스포랄 델부루에키(Torulaspora delbrueckii), 사카로마에스 델부루에키(Saccharomyces delbrueckii), 자이고사카로마이세스 몬골리쿠스(Zygosaccharomyces mongolicus), 도룰라스포라 글로보사(Dorulaspora globosa), 데바르요마이세스 글로보수스(Debaryomyces globosus), 토룰로프시스 글로보사(Torulopsis globosa), 트리콘스포론 쿠타네움(Trichosporon cutaneum), 트리고노프시스 바르아빌리스(Trigonopsis variabilis), 윌리오프시스 칼리포르니카(Williopsis californica), 윌리오프시스 사투르누스

(Williopsis saturnus), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 데바르마이세스 디스포루아(Debaryomyces disporua). 사카로마이세스 비스포라스(Saccharomyces bisporas), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 사카로마이세스 비스포루스(Saccharomyces bisporus), 자이고사카로마이세스 멜리스(Zygosaccharomyces mellis), 자이코사카로마이세스 프리오리아누스(Zygosaccharomyces priorianus), 자이코사카로마이세스 로우심(Zygosaccharomyces rouxiim), 자이코사카로마이세스 로우시(Zygosaccharomyces rouxii), 자이코사카로마이세스 바르케리(Zygosaccharomyces barkeri), 사카로마이세스 로우심(Saccharomyces rouxii), 자이코사카로마이세스 로우심(Zygosaccharomyces rouxii), 자이코사카로마이세스 마조르(Zygosaccharomyces major), 사카로마이세스 로우시(Saccharomyces rousii), 피치아 아노말라(Pichia anomala), 피치아 보비스(Pichia bovis), 피치아 카나데니시스(Pichia Canadensis), 피치아 카르소니(Pichia carsonii), 피리아 파리노세(Pichia farinose), 피치아 페르멘탄스(Pichia fermentans), 피치아 플루숨(Pichia fluxuum), 피치아 멤브라나에파시엔스(Pichia membranaefaciens), 피치아 프세우도폴리모르파(Pichia pseudopolymorpha), 피치아 퀴에르쿰(Pichia quercuum), 피치아 로베르티시(Pichia robertsii), 프쉐도즈마 안타르크티카(Pseudozyma Antarctica), 로도스포리디움 토룰로이데스(Rhodosporidium toruloides), 로도스포리디움 토룰로이데스(Rhodosporidium toruloides), 로도토룰라 굴루티니스(Rhodotorula glutinis), 사카로마이세스 바야누스(Saccharomyces bayanus), 사카로마이세스 바야누스(Saccharomyces bayanus), 사카로마이세스 비스포루스(Saccharomyces bisporus), 사카로마이세스 세레비시아(Saccharomyces cerevisiae), 사카로마이세스 쉐발리에리(Saccharomyces chevalieri), 사카로마이세스 델브릭키(accharomyces delbrueckii), 사카로마이세스 페르멘타티(Saccharomyces fermentati), 사카로마이세스 프라길리스(Saccharomyces fragilis), 사카로마이세스 루드웨기(Saccharomycodes ludwigii), 시조사카로마이세스 폼베(Schizosaccharomyces pombe), 슈완니오마이세스 오시덴탈리스(Schwanniomyces occidentalis), 토룰라스포라 델브릭키이(Torulaspora delbrueckii), 토룰라스포라 글로보사(Torulaspora globosa), 트리고노프시스 바리아빌리스(Trigonopsis variabilis), 윌리오프시스 칼리포르니카(Williopsis californica), 윌리오프시스 사투루누스(Williopsis saturnus), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 자이고사카로마이세스 멜리스(Zygosaccharomyces mellis), 또는 자이고사카로마이세스 로우시(Zygosaccharomyces rouxii). 대표적인 곰팡이는 *아스페르길루스*의 다양한 종, *아스페르길루스 카시엘루스(Aspergillus caesiellus), 아스페르길루스 칸디두스(Aspergillus candidus), 아스페르길루스 카르네우스(Aspergillus carneus), 아스페르길루스 클라바투스(Aspergillus clavatus), 아스페르길루스 데플렉투스(Aspergillus deflectus), 아스페르길루스 플라부스(Aspergillus flavus), 아스페르길루스 푸미가투스(Aspergillus fumigatus), 아스페르길루스 글라우쿠스(Aspergillus glaucus), 아스페르길루스 니둘란스(Aspergillus nidulans), 아스페르길루스 니게르(Aspergillus niger), 아스페르길루스 오크라세우스(Aspergillus ochraceus), 아스페르길루스 오라자에(Aspergillus oryzae), 아스페르길루스 파라시티쿠스(Aspergillus parasiticus), 아스페르길루스 페니실로이데스(Aspergillus penicilloides), 아스페르길루스 레스트리쿠투스(Aspergillus restrictus), 아스페르길루스 소자에(Aspergillus sojae), 아스페르길루스 시도위(Aspergillus sydowi), 아스페르길루스 타마리(Aspergillus tamari), 아스페르길루스 테레우스(Aspergillus terreus), 아스페르길루스 우스투스(Aspergillus ustus),* 또는 *아스페르길루스 베르시컬러(Aspergillus versicolor)*를 포함하지만 이로 한정되지 않는다. 본원에서 특정된 바와 같은 유전공학(genetic engineering)에 앞서, 이러한 세포는 다양한 상업적인 출처(commercial sources) 및 예를 들어, American Type Culture Collection (Rockville, MD)와 같은 연구 자원 시설(commercial sources)로부터 수득할 수 있다. 타겟 분자는 본원에 기재된 타겟 단백질 중의 어떠한 것과 같은 단백질을 포함한다(상기를 참고하라).

[0137] 세포의 유전공학은, 하기와 같은 하나 또는 그 이상의 유전적인 변형(genetic modifications), 만노시다제를 코드화하는 외인성(exogenous) 핵산을 포함할 수 있다: (i) 외부 사슬 연장(Outer CHain elongation, OCH1) 단백질을 코드화하는 내생 유전자(endogenous gene)의 삭제; (ii) 만노스 잔기의 인산화반응을 증가시키기 위한 만노실 인산화반응[예를 들어, *야로위와 리폴리티카, S. 세레비시아, 오가태아 미뉴타, 피치아 파스토리스* 또는 *C. 알비칸스*로부터의 MNN4 폴리펩티드, *P. 파스토리스*로부터의 PNO1 폴리펩티드]을 촉진시킬 수 있는 폴리펩티드를 코드화하는 재조합 핵산의 도입(introduction); (iii) OCH1 단백질의 기능적인 발현을 간섭하는 RNA 분자의 도입 또는 발현; (iv) N-글리코실화 활성도(N-glycosylation activity)를 갖는 야생형 [예를 들어, 내생(endogenous) 또는 외인성(exogenous)] 단백질[예를 들어, N-글리코실화 활성도를 갖는 단백질을 발현시키는]을 코드화하는 재조합 핵산의 도입; (v) 상기에 기재된 타겟 분자를 코드화하는 재조합 핵산의 도입; (v) 이들의 코드화된 단백질의 발현을 변경시키도록, N-글리코실화 활성도를 갖는 단백질을 코드화하는 하나 또는 그 이상의 내인성 유전자의 프로모터 또는 인핸서 요소(promoter or enhancer elements)를 변경시킴. RNA 분자는, 예

를 들어 작은-간섭 RNA(small-interfering RNA, siRNA), 짧은 헤어핀 RNA(short hairpin RNA, shRNA), 안티-센스 RNA, 또는 미크로 RNA(miRNA)를 포함한다. 유전공학은, 첨가[예를 들어, 이종의 서열(heterologous sequence)], 삭제 또는 치환[예를 들어, 점 돌연변이(point mutations)와 같은 돌연변이; 보존적 또는 비-보존적 돌연변이(conservative or non-conservative mutations)]를 갖는 단백질을 생산하기 위한 N-글리코실화 활성도를 갖는 단백질을 코드화하는 내인성 유전자의 변경을 또한 포함한다. 돌연변이는 특별하게 도입될 수 있거나[예를 들어, 부위 특이적 변이(site-directed mutagenesis) 또는 상동재조합], 또는 무작위로 도입될 수 있다[예를 들어, 세포를, 예를 들어 Newman and Ferro-Novick (1987) *J. Cell Biol.* 105(4):1587에 기재된 바와 같이 화학적으로 돌연변이를 일으킬 수 있다].

[0139] 본원에 기재된 유전적인 변형(Genetic modifications)은 하나 또는 그 이상의, (i) 유전적으로 변형된 세포에서의 하나 또는 그 이상의 활성도(activities)에서의 증가, (ii) 유전적으로 변형된 세포에서의 하나 또는 그 이상의 활성도에서의 감소, 또는 (iii) 유전적으로 변형된 세포에서의 하나 또는 그 이상의 활성도의 위치(localization) 또는 세포 내의 분포(intracellular distribution)에서의 변화를 야기할 수 있다. 특정한 활성도[예를 들어, 만노실 인산화반응(mannosyl phosphorylation)의 촉진]의 양에서의 증가는, 만노실 인산화반응을 촉진시킬 수 있는 하나 또는 그 이상의 단백질의 과잉발현(overexpressing), 내인성 유전자(endogenous gene)의 복제 수(copy number)에서의 증가[예를 들어, 유전자 복제(gene duplication)], 유전자에 의해 코드화된 단백질의 발현에서의 증가를 자극하는 내인성 유전자의 프로모터(promoter) 또는 인핸서(enhancer)에서의 변경(alteration) 때문일 수 있음을 인지할 수 있다. 하나 또는 그 이상의 특정한 활성도에서의 감소는, 돌연변이의 형태의 과잉발현[예를 들어, 우성 음성 형태(dominant negative form)], 특정한 활성도를 갖는 하나 또는 그 이상의 단백질의 발현을 감소시키는 하나 또는 그 이상의 간섭 RNA 분자(interfering RNA molecules)의 도입(introduction) 또는 발현, 또는 특정한 활성도를 갖는 단백질을 코드하는 하나 또는 그 이상의 내인성 유전자의 삭제 때문일 수 있다.

[0141] 상동 재조합(homologous recombination)에 의한 유전자를 방해하도록(disrupt), "유전자 대체(gene replacement)" 벡터는 선별가능한 마커 유전자(selectable marker gene)를 포함하는 이러한 방식으로 구성될 수 있다. 선별가능한 마커 유전자는, 상동 재조합을 조정하도록(mediate) 충분한 길이의 유전자의 일부에, 5' 및 3' 둘 다에 실시가능하게 연결될 수 있다. 선별가능한 마커는, URA3, LEU2 및 HIS3 유전자를 포함하는, 숙주 세포 영양요구성(host cell auxotrophy)을 보충하거나 항생물질 내성(antibiotic resistance)을 제공하는 많은 유전자 중의 하나일 수 있다. 그 밖의 적절한 선택가능한 마커는, 효모 세포에 클로람페니콜 내성(chloramphenicol resistance)을 부여하는 CAT 유전자, 또는 β-갈락토시다아제의 발현 때문에 블루 콜로니(blue colonies)를 야기하는 lacZ 유전자를 포함한다. 유전자 대체 벡터(gene replacement vector)의 선으로 된(Linearized) DNA 단편은, 그리고 난 다음에 본 분야에서 잘 알려진 방법을 사용한 세포 내로 도입된다(하기를 참고하라). 게놈 내로의 선으로 된 단편(linear fragments)의 통합(Integration) 및 유전자의 분열(disruption)은 선별 마커(selection marker)를 기초로 결정될 수 있고, 예를 들어 서전블롯 분석(Southern blot analysis)에 의해 확인도리 수 있다. 선별가능한 마커는, 예를 들어 Cre-loxP 시스템에 의해 숙주 세포의 게놈으로부터 제거될 수 있다(하기를 참고하라).

[0143] 그 대신에, 유전자 대체 벡터는, 일부가 어떠한 내인성 유전자 프로모터 서열이 전혀 없고, 유전자의 코딩 서열이 없거나, 유전자의 코딩 서열의 불활성 단편(inactive fragment)을 코드화하는, 방해되도록 유전자의 일부를 포함하는 이러한 방식으로 구성될 수 있다. "불활성 단편(inactive fragment)"은, 유전자의 전장(full-length) 코딩 서열(coding sequence)로부터 생산된 단백질의 활성도의, 예를 들어 약 10 % 미만[예를 들어, 9 % 미만, 8 % 미만, 7 % 미만, 6 % 미만, 5 % 미만, 4 % 미만, 3 % 미만, 2 % 미만, 1 % 미만, 0 %]을 갖는 단백질을 코드화하는 유전자의 단편이다. 유전자의 이러한 일부는, 어떤 알려진 프로모터 서열도 유전자 서열에 실시가능하게 연결되지 않지만, 종결 코돈(stop codon) 및 전사 종결 서열(transcription termination sequence)은 유전자 서열의 일부에 실시가능하게 연결되는 이러한 방식으로 벡터에 삽입된다. 이러한 벡터는, 유전자 서열의 일부에서 나중에 선으로 만들 수 있고, 세포 내로 형질전환된다(transformed). 단일 상동 재조합(single homologous recombination)을 거쳐, 이러한 선형화된 벡터(this linearized vector)는 그리고 난 다음에 유전자의 내인성 대응물(endogenous counterpart)에 통합된다.

[0145] 발현 벡터는 독립적이거나(autonomous) 또는 통합될 수 있다. 재조합 핵산(예를 들어, 하나의 코드화된 만노시다제)은, 플라스미드, 파지(phage), 전이인자(transposon), 코스미드(cosmid) 또는 바이러스 입자(virus particle)와 같은 발현 벡터의 형태에 세포 내로 도입될 수 있다. 재조합 핵산은 염색체 외적으로(extrachromosomally) 유지될 수 있거나 또는, 효모 세포의 염색체의 DNA 내로 통합될 수 있다. 발현 벡터는,

원하는 핵산으로 형질전환된 이러한 세포의 삭제 및/또는 선택을 가능하도록, 선택된 조건 하에서 세포의 생존 능력(cell viability)을 위해 요구되는 단백질[예를 들어, 우라실 생합성(uracil biosynthesis)에 필요한 효소를 코드화하는 URA3, 또는 트립토판 생합성에 필요한 효소를 코드화하는 TRP1]을 코드화하는 선별 마커 유전자를 포함할 수 있다(예를 들어, U.S. Pat. No. 4,704,362를 참고하라). 발현 벡터는 독립적인 복제 서열(autonomous replication sequence, ARS)을 또한 포함할 수 있다. 예를 들어, U.S. Pat. No. 4,837,148 은 *피치아 파스토리스*에서 플라스미드를 유지하기 위한 적절한 수단(suitable means)을 제공하는 독립적인 복제 서열을 기재하였다.

[0147] 통합 벡터(Integrative vectors)는, 예를 들어 U.S. Pat. No. 4,882,279 에 기재되어 있다. 통합 벡터는, 적어도 첫 번째로 삽입가능한 DNA 단편, 선별가능한 마커 유전자 및 두 번째로 삽입가능한 DNA 단편의 연속적으로 배열된 서열을 일반적으로 포함한다. 첫 번째 및 두 번째 삽입가능한 DNA 단편은, 길이에 있어서 각각 약 200 (예를 들어, 약 250, 약 300, 약 350, 약 400, 약 450, 약 500, 또는 약 1000 또는 그 이상) 뉴클레오티드이고, 형질전환될 수 있는 종의 게놈의(genomic) DNA 의 일부에 대해 상동성인 뉴클레오티드 서열을 갖는다. 발현을 위한 관심 유전자를 포함하는 뉴클레오티드 서열(예를 들어, N-글리코실화 활성도를 갖는 단백질을 코드화하는 유전자)은, 마커 유전자의 전 또는 후이든 첫 번째 및 두 번째 삽입가능한 DNA 단편 사이의 이러한 벡터 내에 삽입된다. 통합 벡터를, 숙주 세포 게놈 내로 관심 뉴클레오티드 서열의 통합을 가능하도록 효모 형질전환 전에 선형화시킬 수 있다.

[0149] 발현 벡터는, 효모[예를 들어,*야로위아 리폴리티카(Yarrowia lipolytica),* 아르술라 아데니니보란스(Arxula adeninivorans), P. 파스토리스, 또는 그 밖의 적절한 균종(fungal species)]의 조절 하에서 재조합 핵산을 특징으로 삼을 수 있고, 이는 이들이 균류 세포에서 발현될 수 있다. 적절한 효모 프로모터는, 예를 들어 ADC1, TPI1, ADH2, hp4d, POX, 및 Gal10 프로모터[예를 들어, Guarente *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79(23):7410 을 참고하라]를 포함한다. 추가적으로 적절한 프로모터는, 예를 들어 Zhu and Zhang (1999) *Bioinformatics* 15(7-8):608-611 및 U.S. Patent No. 6,265,185 에 기재되어 있다.

[0151] 프로모터는 구성적(constitutive) 또는 유도성(inducible)[조건부의(conditional)]일 수 있다. 구성적 프로모터(constitutive promoter)는, 이의 발현이 표준 배양 조건 하에서 일정할 수 있는 것으로 이해될 수 있다. 유도성 프로모터(inducible promoter)는 하나 또는 그 이상의 유도 단서(induction cues)에 반응하는 프로모터이다. 예를 들어, 유도성 프로모터는, 화학적으로 조절된 프로모터[예를 들어, 이의 전사 활성도(transcriptional activity)가 알코올, 테트라시이클린(tetracycline), 스테로이드, 금속 또는 그 밖의 소분자와 같은 화학적 유도성 제제(chemical inducing agent)의 존재 또는 부재에 의해 조절되는 프로모터], 또는 물리적으로 조절된 프로모터[예를 들어, 이의 전사 활성도가 빛 또는 높거나 낮은 온도와 같은 물리적인 유도인자(physical inducer)의 존재 또는 부재에 의해 조절되는 프로모터]일 수 있다. 유도가능한 프로모터는 또한, 이들 자체가 화학적 또는 물리적 단서에 의해 직접적으로 조절되는 하나 또는 그 이상의 전사 인자에 의해 간접적으로 조절될 수 있다.

[0153] 그 밖의 유전적으로 조작된 변형물(other genetically engineered modifications)은 또한 조건적일 수 있는 것으로 이해된다. 예를 들어, 유전자는, 예를 들어 Cre-loxP 시스템과 같은 부위-특이적 DNA 재조합 효소(recombinase)를 사용하여 조건부로 삭제될 수 있다[예를 들어, Gossen *et al.* (2002) *Ann. Rev. Genetics* 36:153-173 and U.S. Application Publication No. 20060014264를 참고하라].

[0155] 재조합 핵산은, 스페로플라스트 기술(spheroplast technique) 또는 전체-세포 염화리튬 효모 형질전환 방법(whole-cell lithium chloride yeast transformation method)과 같은 다양한 방법을 사용하여 본원에 기재된 세포 내로 도입시킬 수 있다. 세포 내로의 플라스미드 또는 선형의 핵산 벡터의 형질전환에 유용한 그 밖의 방법은, 각각 이의 전체가 참고문헌으로 본원에 포함되는, 예를 들어 U.S. Patent No. 4,929,555; Hinnen *et al.* (1978) *Proc. Nat. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163; U.S. Patent No. 4,879,231; 및 Sreekrishna *et al.* (1987) *Gene* 59:115 에 기재되었다. 일렉트로포레이션(Electroporation) 및 PEG1000 전체 세포 형질전환 방법은, Cregg and Russel, Methods in Molecular Biology: Pichia Protocols, Chapter 3, Humana Press, Totowa, N.J., pp. 27-39 (1998)에 기재된 바와 같이 또한 사용될 수 있다.

[0157] 형질전환된 균류 세포는, 요구된 생화학적 생산물의 부재[세포의 영양요구성(cell's auxotrophy) 때문에]에서 형질전환 후에 영양 요구성 세포를 배양, 새로운 표현형의 선별(selection) 및 검출(detection), 또는 형질전환체(transformant)에 포함된 내성 유전자(resistance gene)의 부재 하에서 효모에 대해 유독성의 항생 물질(antibiotic)의 존재 하에서의 배양을 포함하지만 이로 제한되지 않는 적절한 기술을 사용하여 선택될 수 있다.

형질전환체는, 게놈 내로 발현 카세트(expression cassette)의 통합에 의해 선별될 수 있고/있거나 확인될 수 있고, 이는, 예를 들어 서전 블롯(Southern blot) 또는 PCR 분석에 의해 평가될 수 있다.

[0159] 관심 타겟 세포 내로 백터의 도입 전에, 벡터를, 상기에 나타낸 바와 같이, *Escherichia coli* (*E. coli*)와 같은 세균 세포(bacterial cells)에서 성장시킬 수 예를 들어, 증폭됨)있다. 벡터 DNA 는, 세균의 환경(bacterial milieu)으로부터의 벡터 DNA 의 정제를 야기하는, 본 분야에서 알려진 방법의 하나에 의해 세균 세포로부터 분리될 수 있다. 정제된 벡터 DNA 는, *E. coli* 단백질이 표유동물 세포에 유독할 수 있기 때문에, 어떠한 *E. coli* 단백질이 플라스미드 DNA 제조에 존재하지 않음을 보장하도록, 페놀, 클로로포름 및 에테르과 함께 광범위하게 추출될 수 있다.

[0161] 몇몇 실시형태에서, 유전적으로 조작된 균류 세포는, OCH1 유전자 또는 이의 유전자 생산물(예를 들어 mRNA 또는 단백질)이 결핍되고, OCH1 활성도에 있어서 결핍된다. 몇몇 실시형태에서, 유전적으로 조작된 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 발현시킨다[예를 들어, *야로위아 리폴리티카(Yarrowia lipolytica), S. 세레비시아, 오가테아 미뉴타(Oogataea minuta), 피치아 파스토리스(Pichia pastoris), 또는 C. 알비칸스(C. albicans)*로부터의 MNN4 폴리펩티드, 또는 *P.파스토리스(P. pastoris)*로부터의 PNO1 폴리펩티드]. 예를 들어, 균류 세포는 *Y. 리폴리티카*(Genbank® 수납 번호: XM_503217, Genolevures Ref: YALI0D24101g)로부터의 MNN4 폴리펩티드를 발현시킬 수 있다. 몇몇 실시형태에서, 유전적으로 조작된 세포는 OCH1 활성도가 결핍되고, 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 발현시킨다.

[0163] M6P 잔기의 캡핑이 제거된 후에, 타겟 분자는 분리될 수 있다. 몇몇 실시형태에서, 타겟 분자는 효모 세포 내에 유지되고, 세포 용해액 상에 방출된다. 몇몇 실시형태에서, 타겟 분자는, 세포로부터의 분자의 분비를 유도하는, 코딩 서열(coding sequence)[고유의 외인성 핵산 또는 발현 벡터 내로 조작됨(either native to the exogenous nucleic acid or engineered into the expression vector)]에 의해 제공된 메커니즘(mechanism)을 통해 배양 배지 내로 분비된다. 세포 용해액 또는 배양 배지에서의 캡핑이 제거된 타겟 분자의 존재는, 분자의 존재를 감지하기 위해 다양한 표준 프로토콜에 의해 확인될 수 있다. 예를 들어, 변경된 타겟 분자가 단백질일 경우에, 이러한 프로토콜은, 변경된 타겟 단백질(또는 타겟 단백질 이 자체)에 대한 특정한 항체와 함께 면역블러팅(immunoblotting) 또는 방사성 면역침강법(radioimmunoprecipitation), 변경된 타겟 단백질(또는 타겟 단백질 이 자체)에 특이적인 리간드(ligand)의 결합, 또는 변경된 타겟 단백질(또는 타겟 단백질 이 자체)의 특이적인 효소 활성도에 대한 테스트를 포함할 수 있지만, 이로 제한되지 않는다.

[0165] 본원에서 기재된 방법을 사용하여 생산된 타겟 분자에서, 당단백질에서의 적어도 47 %(예를 들어, 적어도 50, 55, 60, 65, 70, 75, 80, 85, 또는 90 %)의 N-글리칸은 말단 포스포-6-만노스 잔기를 갖는다. 말단 포스포-6-만노스 잔기를 갖는 N-글리칸의 퍼센트는 DSA-FACE 일렉트로페로그램(electropherograms)에서의 피크 영역(peak areas)으로부터 추정될 수 있다. 실시예 13 을 참고하라.

[0167] 몇몇 실시형태에서, 분리한 후에, 캡핑이 제거된 타겟 분자는, 이종의 일부(heterologous moiety), 예를 들어 효소적인 또는 화학적인 수단을 사용하여 부착될 수 있다. "이종의 일부(heterologous moiety)"는 변경된 타겟 분자에 연결된[예를 들어, 공유결합으로(covalently) 또는 비-공유결합으로(non-covalently)]는 어떠한 구성성분(constituent)을 나타내고, 구성성분(constituent)은, 변경된 타겟 분자에 원래 존재하는 구성성분과는 다르다. 이종의 일부는, 예를 들어 중합체, 운반체(carriers), 아쥬반트(adjuvants), 항체독소(immunotoxins), 또는 검출가능한[예를 들어, 형광성(fluorescent), 발광성(luminescent) 또는 방사성(radioactive)] 일부를 포함한다. 몇몇 실시형태에서, 추가적인 N-글리칸은 변경된 타겟 분자에 첨가될 수 있다.

[0169] 타겟 분자의 글리코실화 반응을 검출하기 위한 방법은, DNA 염기서열 분석기-지원된[DNA sequencer-assisted, DSA)], 형광단-지원된 탄수화물 전기영동법[fluorophore-assisted carbohydrate electrophoresis, FACE)] 또는 표면-증감 레이저 이탈/이온화 비행시간 질량 분석법[Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, SELDI-TOF MS]을 포함한다. 예를 들어, 당단백질이 변형된 다음에, 예를 들어 막에서 고정화(immobilization)시키는, DSA-FACE 를 사용할 수 있다. 당단백질이 디티오트레이톨(dithiothreitol, DTT) 또는 J-메르캅토에탄올(J-mercaptoethanol)과 같은 적절한 환원 제제(reducing agent)와 함께 환원될 수 있다. 단백질의 술프히드릴기(sulfhydryl groups)가 요오드아세트산(iodoacetic acid)과 같은 산을 사용하여 카르복실화될 수 있다. 그 다음에, N-글리칸은 N-글리코시다아제 F(N-glycosidase F)와 같은 효소를 사용하여 단백질로부터 방출될 수 있다. N-글리칸은 임의적으로 재구성될 수 있고, 환원-아미노화 반응(reductive amination)에 의해 유도체를 합성시킬 수 있다(derivatized). 유도체 합성된 N-글리칸은 그리고 난 다음에 농축될 수 있다. N-글리칸 분석에 적절한 기기 장치(Instrumentation)는, 예를 들어 ABI PRISM®

377 DNA 염기서열 분석기(Applied Biosystems)를 포함한다. 데이터 분석은, 예를 들어 GENESCAN® 3.1 소프트웨어(Applied Biosystems)를 사용하여 실행될 수 있다. 임의적으로, 분리된 만노단백질(mannoproteins)은 이들의 N-글리칸 상황을 확인하기 위해 하나 또는 그 이상의 효소로 추가적으로 처리될 수 있다. N-글리칸 분석의 추가적인 방법은, 예를 들어, 질량분석법(예를 들어, MALDI-TOF-MS), 일반적인 상(normal phase), 역상에서의 고성능액체크로마토그래피(high-pressure liquid chromatography, HPLC), 이온교환 크로마토그래피[예를 들어, 글리칸이 표지되지 않은 경우, 펄스 전류 검출(pulsed amperometric detection)과 함께, 및 글리칸이 적절하게 표지된 경우, UV 흡광도(absorbance) 또는 형광(fluorescence)과 함께]를 포함한다. 또한 Callewaert *et al.* (2001) *Glycobiology* 11(4):275-281 and Freire *et al.* (2006) *Bioconjug. Chem.* 17(2):559-564 를 참고하라.

[0171] *조작된 세포(Engineered Cells)의 배양*

[0172] 본 문서는 또한 본원에 기재된 유전적으로 조작된 세포 중의 하나의 실질적인 순수 배양물(substantially pure culture)을 제공한다. 본원에서 사용된 바와 같이, 유전적으로 조작된 세포의 "실질적인 순수 배양물"은, 배양물에서 생존가능한 세포(viable cells)의 전체 수의 약 40 % 미만(즉, 약 35%; 30%; 25%; 20%; 15%; 10%; 5%; 2%; 1%; 0.5%; 0.25%; 0.1%; 0.01%; 0.001%; 0.0001% 미만; 또는 더 적게)이 유전적으로 조작된 세포, 예를 들어, 박테리아의(bacterial), 균류(효모를 포함)의, 마이코플라스마(mycoplasmal) 또는 원생동물 세포 외의 생존가능한 세포인, 이러한 세포의 배양물이다. 이러한 문맥에서의 용어 "약(about)"은, 관련된 퍼센트는 특정한 퍼센트 이상 또는 이하의 특정한 백분률의 15 % 퍼센트일 수 있음을 의미한다. 따라서, 예를 들어, 약 20 % 는 17 % 내지 23 % 일 수 있다. 유전적으로 조작된 세포의 이러한 배양물은 세포 및 성장, 저장 또는 수송배지(transport medium)를 포함한다. 배지는 액체, 반고체(semi-solid)[예를 들어, 젤리와 같은 배지(gelatinous media)], 또는 동결된 것일 수 있다. 배양물은, 액체 또는 반고체 배지에서 성장한 세포, 또는 동결 저장 또는 수송 배지(frozen storage or transport medium)를 포함하는, 저장 또는 수송 배지에 저장되거나 또는 수송된 배지에서 성장한 세포를 포함한다. 배양물은 배양 용기(culture vessel) 또는 저장 용기(storage vessel) 또는 기질(substrate)[예를 들어, 배양 접시(culture dish), 플라스크, 또는 튜브(tube) 또는 저장 유리병(storage vial) 또는 튜브]에 있다.

[0174] 본원에 기재된 유전적으로 조작된 세포는, 예를 들어 동결된 세포 현탁액(frozen cell suspensions)으로서, 예를 들어 글리세롤 또는 당과 같은 동해방지제(cryoprotectant)를 포함하는 완충용액에, 동결건조된 세포(lyophilized cells)로서, 저장될 수 있다. 선택적으로, 이들은 예를 들어, 유동층 건조(fluidized bed drying) 또는 분무 건조(spray drying), 또는 그 밖의 적절한 건조 방법 중 하나로 수득된, 예를 들어 건조된 세포 표본(dried cell preparations)으로서 저장될 수 있다.

[0176] *대사 질환(Metabolic Disorders)*

[0177] 캡핑이 제거된 말단 M6P 잔기를 갖는 분자는 다양한 대사 질환을 치료하는데 사용될 수 있다. 대사 질환은 개별적인 인간(또는 동물) 세포 내에 에너지의 생산에 영향을 미치는 것 중의 하나이다. 몇몇의 대사 질환은 식단, 독소, 감염(infections)의 결과로서 획득될 수 있지라도, 대부분의 대사 질환은 유전적이다. 유전적인 대사 질환은 선천성 대사이상(inborn errors of metabolism)으로서 또한 알려졌다. 일반적으로, 유전적 대사 질환은 세포의 대사 과정에서의 몇몇의 단계에 필수적인 효소가 결핍되거나 또는 적절하지 않게 구성되어 야기되는 유전적 결함(genetic defects)에 의해 초래된다. 대사 질환의 가장 큰 클래스(the largest classes)는, 탄수화물 대사(carbohydrate metabolism)의 질환, 아미노산 대사의 질환, 유기산 대사의 질환[유기 산뇨증(organic acidurias)], 지방산 산화 및 미토콘드리아 대사(fatty acid oxidation and mitochondrial metabolism)의 질환, 포르피린 대사의 질환(disorders of porphyrin metabolism), 푸린 또는 피리미딘 대사의 질환, 스테로이드 대사의 질환, 미토콘드리아 작용의 질환(disorders of mitochondrial function), 페르옥시좀 작용의 질환(disorders of peroxisomal function), 및 리소좀 축적 질환(lysosomal storage disorders, LSDs)이다.

[0179] 캡핑이 제거된 말단 M6P 잔기를 갖는 하나 또는 그 이상의 분자(또는 동일한 약제학적 조성물)의 투여를 통해 치료될 수 있는 대사 질환의 예는, 유전성 혈액색소침착증(hereditary hemochromatosis), 안피부백피증(oculocutaneous albinism), 단백질 C 결핍증(protein C deficiency), 제1형 유전적 혈관부종(type I hereditary angioedema), 선천적인 백당분해효소 결핍증(congenital sucrase-isomaltase deficiency), 크리글러-나자르 제2형(Crigler-Najjar type II), 라론 증후군(Laron syndrome), 유전성 미엘로페옥시다제(hereditary Myeloperoxidase), 일차 감상선 기능 저하증(primary hypothyroidism), 선천적 긴 QT 증후군

(congenital long QT syndrome), 트롭신 결합 글로블린 결핍증(tyroxine binding globulin deficiency), 가족성 과콜레스테롤혈증(familial hypercholesterolemia), 가족성 카일로마이크론혈증(familial chylomicronemia), 알베타-리포프로테이네마(abeta-lipoproteinema), 낮은 플라즈마 리포단백질 A 수준(low plasma lipoprotein A levels), 간 장애(liver injury)를 갖는 유전적 폐기종(hereditary emphysema), 선천성 갑상샘기능저하증(congenital hypothyroidism), 골형성부전증(osteogenesis imperfecta), 유전성 저섬유소원혈증(hereditary hypofibrinogenemia), 알파-1 안티키모트립신 결핍증(alpha-1antichymotrypsin deficiency), 신성요붕증(nephrogenic diabetes insipidus), 신경하수체 당뇨 요붕증(neurohypophyseal diabetes insipidus), 아데노신탈아미노화효소 결핍증(adenosine deaminase deficiency), 펠리쩨우스-메르쯔바하 병(Pelizaeus Merzbacher disease), 폰빌레브란트병 제IIA형(von Willebrand disease type IIA), 복합 인자 V 및 VIII 결핍증(combined factors V and VIII deficiency), 선천 척추뼈끝 형성이상(spondylo-epiphyseal dysplasia tarda), 맥락막 결여(choroideremia), I 세포 질병(I cell disease), 거식증(Batten disease), 모세혈관확장성 운동실조증(ataxia telangiectasias), ADPKD-상염색체 우성다낭성 신질환(ADPKD-autosomal dominant polycystic kidney disease), 미세융모 포접 질병(microvillus inclusion disease), 결절성 경화증(tuberous sclerosis), 로웬의 안뇌신증후군(oculocerebro-renal syndrome of Lowe), 근위축성 측삭경화증(amyotrophic lateral sclerosis), 골수이형성증후군(myelodysplastic syndrome), 베어림프구 증후군(Bare lymphocyte syndrome), 탄지에르질환(Tangier disease), 가족성 간내 담즙정체증(familial intrahepatic cholestasis), X-연관 부신백질이영양증(X-linked adreno-leukodystrophy), 스콧 증후군(Scott syndrome), 헤르만스키-푸드락증후군 제1형 및 제2형(Hermansky-Pudlak syndrome types 1 and 2), 젤웨거증후군(Zellweger syndrome), 점상연골 이형성증(rhizomelic chondrodysplasia puncta), 상염색체 열성 일차성 고수산뇨증(autosomal recessive primary hyperoxaluria), 무어 트라네베르그 증후군(Mohr Tranebjaerg syndrome), 척추성 근위측증(spinal and bullar muscular atrophy), 일차성 섬모의 디스키네시아(primary ciliary diskenesia)[카르타게너 증후군(Kartagener's syndrome)], 말단비대증 및 거인증(giantism and acromegaly), 유즙 분비증(galactorrhea), 에디슨 병(Addison's disease), 부신성 웅성화(adrenal virilism), 쿠싱 증후군(Cushing's syndrome), 케토애시도시스(ketoacidosis), 일차 또는 이차성 알도스테론증(primary or secondary aldosteronism), 밀러 디커 증후군(Miller Dieker syndrome), 활뇌증(lissencephaly), 운동 신경 질환(motor neuron disease), 우셔 증후군(Usher's syndrome), 비스코트-알드리치 증후군(Wiskott-Aldrich syndrome), 옵티즈 증후군(Optiz syndrome), 헌팅턴병(Huntington's disease), 유전성 췌장염(hereditary pancreatitis), 항-인산화지질 증후군(anti-phospholipid syndrome), 중복성 결합 조직 질환(overlap connective tissue disease), 쇼그렌 증후군(Sjögren syndrome), 강직인간증후군(stiff-man syndrome), 브루가다증후군(Brugada syndrome), 핀란드 유형의 선천적인 신증후군(congenital nephritic syndrome of the Finnish type), 듀빈-존슨 증후군(Dubin-Johnson syndrome), X-연관 하이포스포스파테미아(X-linked hypophosphosphatemia), 펜드레드 증후군(Pendred syndrome), 유아기의 저혈당증(persistent hyperinsulinemic hypoglycemia of infancy), 유전성 구상적혈구증(hereditary spherocytosis), 선천성철분대사이상증(aceruloplasminemia), 영아기의 신경 세로이드 리포푸신증(infantile neuronal ceroid lipofuscinosis), 가연골무형성증 및 골단이형성증(pseudoachondroplasia and multiple epiphyseal), 스타가르트-유사 근육 영양실조(Stargardt-like macular dystrophy), X-연관 샤리코-마리-듀스 병(X-linked Charcot-Marie-Tooth disease), 상염색체 우성 망막색소변성증(autosomal dominant retinitis pigmentosa), 울콧-랠리슨 증후군(Wolcott-Rallison syndrome), 쿠싱병(Cushing's disease), 지대근이영양증(limb-girdle muscular dystrophy), 뮤코다당체침착증 제IV형(mucoploy-saccharidosis type IV), 피니시의 유전성 가족성 아밀로이드증(hereditary familial amyloidosis of Finish), 아데손 질환(Anderson disease), 육종(sarcoma), 만성 골수-단핵구성 백혈병(chronic myelomonocytic leukemia), 심근증(cardiomyopathy), 안면 성기형성 장애(faciogenital dysplasia), 토르시온 질병(Torsion disease), 헌팅턴 및 유전성 실조증(Huntington and spinocerebellar ataxias), 유전성 하이퍼호모시테이네미아(hereditary hyperhomosyteinemia), 말초신경병증(polyneuropathy), 말초운동신경 질환(lower motor neuron disease), 색소성 망막염(pigmented retinitis), 다발성 관절염(seronegative polyarthritis), 간질성 폐섬유증(interstitial pulmonary fibrosis), 레이노 현상(Raynaud's phenomenon), 베게너 육종종증(Wegner's granulomatosis), 프레오테이누리아(preoteinuria), CDG-Ia, CDG-Ib, CDG-Ic, CDG-Id, CDG-Ie, CDG-If, CDG-IIa, CDG-IIb, CDG-IIc, CDG-IId, 엘러스-단로스 증후군(Ehlers-Danlos syndrome), 다발성 뼈돌출증(multiple exostoses), 그리셀리 증후군(Griscelli syndrome)(제1형 또는 제2형), 또는 X-연관 비-특이적인 정신지체(non-specific mental retardation)을 포함할 수 있다. 게다가, 대사 질환은 이로 제한되지 않지만, 하기와 같은 리소좀 축적 질환을 또한 포함할 수 있다: 파브리병(Fabry's disease), 점액다당류증 I(mucopolysaccharidosis I), 파버 질병

(Farber disease), 고세병(Gaucher disease), GM1-강글리오시드증(GM1-gangliosidosis), 테이-삭스병(Tay-Sachs disease), 샌드호프병(Sandhoff disease), GM_2 활성제 질병(GM_2 activator disease), 크라베병(Krabbe disease), 이염성백질이영양증(metachromatic leukodystrophy), 니만-피크병(Niemann-Pick disease)(제A형, 제B형, 제C형), 샤이에 질병(Scheie disease), 헌터 질병(Hunter disease), 산필립포 질병(Sanfilippo disease), 모르키오병(Morquio disease), 마로토-라미 질병(Maroteaux-Lamy disease), 히알루로니다아제 결핍증(hyaluronidase deficiency), 아스파르틸글루코스아민뇨증 (aspartylglucosaminuria), 푸코시드축적증(fucosidosis), 만노시도시스(mannosidosis), 쉰들러 병(Schindler disease), 사이알산축적증 유형 1(sialidosis type 1), 폼피병(Pompe disease), 피크노디소토시스(Pycnodysostosis), 세로이드 리포푸신증(ceroid lipofuscinosis), 콜레스테롤 에스테르 축적 질병(cholesterol ester storage disease), 월만병(Wolman disease), 다종 술파타아제 결손증(Multiple sulfatase deficiency), 갈락토시알리도시스(galactosialidosis), 뮤코리피드증(mucolipidosis)(제Ⅰ형, 제Ⅱ형, 제Ⅲ형), 시스틴축적증(cystinosis), 시알산 축적 질병(sialic acid storage disorder), 마리네스코-쉐글렌 증후군(Marinesco- Sjögren syndrome)을 갖는 킬로미크론 보유 질병(chylomicron retention disease), 헤르만스키-푸드락 증후군(Hermansky-Pudlak syndrome), 체디아크-히가시 증후군(Chediak-Higashi syndrome), 다논병(Danon disease), 또는 겔레오피직 이형성증(Geleophysic dysplasia).

[0181] 대사 질환의 증후(Symptoms)는 수많고 다양하고, 하나 또는 그 이상의, 예를 들어, 빈혈(anemia), 피로(fatigue), 쉽게 힘듦(bruising easily), 저혈소판(low blood platelets), 간 비대(liver enlargement), 비장 거대(spleen enlargement), 골격 약화(skeletal weakening), 폐손상(lung impairment), 감염(infections) [예를 들어, 흉부 감염(chest infections) 또는 폐렴], 신장 손상(kidney impairment), 점진적인 뇌 손상(progressive brain damage), 발작(seizures), 추가적인 두꺼운 태반(extra thick meconium), 기침(coughing), 천명(wheezing), 과량의 침 또는 점액 생성(excess saliva or mucous production), 짧은 호흡(shortness of breath), 복통(abdominal pain), 폐쇄성 창자 또는 내장(occluded bowel or gut), 생식력 문제(fertility problems), 코에서의 폴립(polyps in the nose), 손톱/발톱 및 피부의 곤봉지(clubbing of the finger/toe nails and skin), 손 또는 발의 통증(pain in the hands or feet), 혈관각화종(angiokeratoma), 감소된 땀(decreased perspiration), 각막 및 렌티큘라의 불투명함(corneal and lenticular opacities), 백내장(cataracts), 승모판 탈줄증 및/또는 역류(mitral valve prolapse and/or regurgitation), 심장비대(cardiomegaly), 온도 비내성(temperature intolerance), 걷기의 어려움(difficulty walking), 연하의 어려움(difficulty swallowing), 점진적인 시력 손상(progressive vision loss), 점진적인 청력 손상(progressive hearing loss), 저혈압(hypotonia), 대설증(macroglossia), 반사소실(areflexia), 하부 요통(lower back pain), 수면 무호흡(sleep apnea), 기좌호흡(orthopnea), 졸림(somnolence), 척추전만증(lordosis), 또는 척추측만증(scoliosis)을 포함한다. 결합이 있거나 또는 부재한 단백질의 다양한 유형 및 결과적으로 생성된 질병의 표현형[예를 들어, 대사 질환의 징후적인 표시(symptomatic presentation)] 때문에, 주어진 질환은 특정한 질환에 대해 오직 증후의 특성을 일반적으로 나타낸 것으로 이해할 수 있다. 예를 들어, 파브리병을 가진 환자는, 이로 제한되지 않지만, 온도 과민증(temperature intolerance), 각막의 휠링(corneal whirling), 통증(pain), 피부 발진(skin rashes), 구역(nausea) 또는 설사(dirarrhea)와 같은 상기에 언급한 증상의 특정한 부분집합(subset)을 나타낼 수 있다. 고셔 증후군을 갖는 환자는 비종(splenomegaly), 간경변(cirrhosis), 경련(convulsions), 고긴장(hypertonia), 무호흡(apnea), 골다공증(osteoporosis), 또는 피부 변색(skin discoloration)으로 나타낼 수 있다.

[0183] 본원에서 기재된 하나 또는 그 이상의 캡핑이 제거된 분자의 투여 뿐만 아니라, 대사 질환은 적절한 영양분 및 비타민[예를 들어, 보조인자 치료(cofactor therapy)], 물리적인 치료 및 통증 치료(medication)로 또한 치료될 수 있다.

[0185] 주어진 대사 질환의 특정한 성질에 따라서, 환자는 나이를 불문하고 이러한 증후가 나타날 수 있다. 많은 경우에, 증후는 어린 시절(childhood) 또는 성년기 초반(early adulthood)에 나타난다. 예를 들어, 피브리병의 증후는 젊었을 때, 예를 들어 10 살 또는 11 살에 나타날 수 있다.

[0187] 본원에 사용된 바와 같이, "점진적인 대사 질환의 위험에 처한(at risk of developing a metabolic disorder)" 피검자(subject)는, 질환이 진행되기 위한 소인, 즉, 하기와 같은 효소에서의 돌연변이의 결과로서 대사 질환이 진행되기 위한 유전(predisposition)의 소인(genetic predisposition)을 갖는 피검자이다: 산성 알파 글루코시다제(acid alpha glucosidase), 알파 갈락토시다아제(alpha galactosidase), 알파-L-이두로니다제(alpha-L-iduronidase), 베타-D-갈락토시다아제(beta-D-galactosidase), 베타-글루코시다아제(beta-glucosidase), 베타-

헥소사미니다아제(beta-hexosaminidase), 베타-D-만노시다제(beta-D-mannosidase), 알파-L-푸코시다아제(alpha-L-fucosidase), 아릴설파타아제 B(arylsulfatase B), 아릴설파타아제 A(arylsulfatase A), 알파-N-아세틸갈락토사미다아제(alpha-N-acteylgalactosaminidase), 아스파르틸글루코사미니다아제(aspartylglucosaminidase), 인두로네이트-2-술파타아제(iduronate-2-sulfatase), 알파-글루코사미니데-N-아세틸트랜스퍼라아제(alpha-glucosaminide-N-acetyltransferase), 베타-D-글루코로니다아제(beta-D-glucoronidase), 히알루론다아제(hyaluronidase), 알파-L-만노시다제(alpha-L-mannosidase), 알파-뉴로미니다아제(alpha-neurominidase), 포스포트랜스퍼라아제(phosphotransferase), 산성 리파아제(acid lipase), 산성 세라미다이제(acid ceramidase), 스피노그미엘리나아제(sphinogmyelinase), 티오에스테라제(thioesterase), 카텝신 K(cathepsin K), 또는 지단백질지방분해효소(lipoprotein lipase). 명확하게, "점진적인 대사 질환의 위험에 처한(at risk of developing a metabolic disorder)" 피검자(subject)는, 관심 종(species of interest) 내의 모든 피검자는 아니다.

[0189] "질환을 갖는 것으로 혐의를 갖는(suspected of having a disorder)" 피검자는 본원에 기재된 바와 같은 대사 질환의 하나 또는 그 이상의 징후를 갖는 사람이다.

[0191] *치료의 약제학적 조성물 및 방법*

[0192] 캡핑이 제거된 M6P 잔기를 갖는 타겟 분자는, 치료학적으로 유요한 양의 분자 및 하나 또는 그 이상의 아쥬반트, 부형제(excipients), 담체(carriers) 및/또는 희석제(diluents)를 포함하는 약제학적 조성물 내로 포함될 수 있다. 허용가능한 희석제, 담체 및 부형제는 일반적으로, 수용자의 항상성(recipient's homeostasis)[예를 들어, 전해질 균형(electrolyte balance)]에 분리하게 영향을 미치지 않는다. 허용가능한 담체는, 생체에 적합한, 불활성(inert) 또는 생체흡수가능한 염(bioabsorbable salts), 완충화제(buffering agents), 올리고(oligo)- 또는 다당류, 중합체, 점성-증진제(viscosity-improving agents), 보존제 등을 포함한다. 하나의 대표적인 담체는 생리 식염수(physiologic saline)(0.15 M NaCl, pH 7.0 내지 7.4)이다. 또 다른 대표적인 담체는 50 mM 의 인산나트롬, 100 mM 의 염화나트륨이다. 약제학적 조성물의 제형 및 투여를 위한 기술에서의 추가적인 사항은 예를 들어 Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.)에서 발견할 수 있다. 추가적인 활성 성분은 조성물 내에 포함될수 있다.

[0194] 캡핑이 제거된 M6P 잔기를 갖는 분자를 포함하는 약제학적 조성물의 투여는 전신 또는 국소(systemic or local)일 수 있다. 약제학적 조성물은, 비경구 및/또는 비경구외(non-parenteral) 투여와 같이 제형화될(formulated) 수 있다. 특정한 투여 양상(Specific administration modalities)은 피하의(subcutaneous), 정맥주사의(intravenous), 근육내(intramuscular), 복강내(intraperitoneal), 피부를 통한(transdermal), 척추강내(intrathecal), 경구(oral), 직장의(rectal), 구강의(buccal), 국부의(topical), 코의(nasal), 눈의(ophthalmic), 관절 내의(intra-articular), 동맥 내의(intra-arterial), 지주막 아래의(sub-arachnoid), 기관지의(bronchial), 림프액의(lymphatic), 질의(vaginal), 및 자궁 내의(intra-uterine) 투여를 포함한다.

[0196] 투여는, 약제학적 조성물의 한 회분(bolus)의 주기적인 주사제일 수 있거나, 또는 외부[예를 들어, IV 백(bag)] 또는 내부[예를 들어, 생체침식성 임플란트(bioerodable implant), 생체 인공 장기(bioartificial organ), 또는 주입된 변경된 N-글리코실화 분자 생산 세포의 콜로니(colony of implanted altered N-glycosylation molecule production cells)]인 저장소(reservoir)로부터의 정맥내(intravenous) 또는 복강내(intraperitoneal) 투여에 의해 연속 또는 반복될 수 있다. 예를 들어, U.S. Pat. Nos. 4,407,957, 5,798,113, 및 5,800,828 를 참고하라. 약제학적 조성물의 투여는 하기와 같은 적절한 수송 수단을 사용하여 수행될 수 있다: 펌프(pump)(예를 들어, Annals of Pharmacotherapy, 27:912 (1993); Cancer, 41:1270 (1993); Cancer Research, 44:1698 (1984)를 참고하라 ; 마이크로캡슐화(microencapsulation)(예를 들어, U.S. Pat. Nos. 4,352,883; 4,353,888; 및 5,084,350를 참고하라); 지속적인 방출 중합체 임플란트(continuous release polymer implants)(예를 들어, Sabel, U.S. Pat. No. 4,883,666를 참고하라); 매크로캡슐화(macroencapsulation)(예를 들어, U.S. Pat. Nos. 5,284,761, 5,158,881, 4,976,859 및 4,968,733 및 published PCT patent applications WO92/19195, WO 95/05452를 참고하라); 주사(injection), either 피하의(subcutaneously), 정맥 내(intravenously), 동맥 내(intra-arterially), 근육 내(intramuscularly), 또는 그 밖의 적절한 위치에 ; 또는 경구 투여, 캡슐, 액체, 정제(tablet), 알약(pill), 또는 지속적인 방출 제형(prolonged release formulation).

[0198] 비경구적인 수송 시스템(parenteral delivery systems)의 예는, 에틸렌-비닐 아세테이트 공중합체 입자(ethylene-vinyl acetate copolymer particles), 삼투 펌프(osmotic pumps), 주입가능한 주입 시스템

(implantable infusion systems), 펌프 수송(pump delivery), 캡슐화된 세포 수송(encapsulated cell delivery), 리포솜의 수송(liposomal delivery), 바늘-전달된 주사(needle-delivered injection), 바늘을 제외한 주사(needle-less injection), 분무기(nebulizer), 에어로졸(aerosolizer), 일렉트로포레이션(electroporation), 및 경피투여 패치(transdermal patch)를 포함한다.

[0200] 비경구 투여에 적절한 제형은 일반적으로, 바람직하게 수용자의 혈액과 등장인, 변경된 N-글루코실화 분자(altered N-glycosylation molecule)의 멸균된 수용액(sterile aqueous preparation)[예를 들어, 생리 식염수(physiological saline solution)]을 포함한다. 제형은 단일-투여(unit-dose) 또는 다중-투여 형태로 나타낼 수 있다.

[0202] 경구 투여에 적절한 제형은, 각각 변경된 N-글루코실화 분자의 예정된 양(predetermined amount)을 포함하는, 캡슐, 카시에(cachets), 정제 또는 로젠지(lozenges) ; 또는 시럽(syrup), 엘릭시르(elixir), 또는 물약(draught)과 같은 수성의 액체 또는 비-수성의 액체에서의 현탁액과 같은 별개의 단위(discrete units)로서 나타낼 수 있다.

[0204] 국소 투여(topical administration)에 적절한 캡핑이 제거된 M6P 잔기를 갖는 분자는, 예를 들어 크림(cream), 스프레이(spray), 폼(foam), 겔(gel), 연고(ointment), 고약(salve), 또는 건조 루브(dry rub)로서 포유동물(예를 들어, 인간 환자)에게 투여될 수 있다. 건조 루브는 또한 투여 부위(site of administration)에서 다시 수화될(rehydrated) 수 있다. 이러한 분자는 붕대(bandage), 거즈(gauze), 또는 패치(patch) 내[예를 들어, 스며들게(soaked into)하고 건조시킴]에 직접적으로 스며들게 할 수 있고, 그리고 난 다음에 국소적으로 적용할 수 있다. 이러한 분자는 또한, 국소 투여를 위한 붕대, 거즈 또는 패치에 반액체(semi-liquid), 겔화(gelled) 또는 완전한 액체 상태로 유지시킬 수 있다(예를 들어, U.S. Patent No. 4,307,717을 참고하라).

[0206] 약제학적 조성물의 치료학적 유효량(Therapeutically effective amounts)을, 본 분야의 기술 중의 하나로 확인할 수 있는 투여량 처방(dosage regimen)에서 이를 필요로 하는 피검자에게 투여할 수 있다. 예를 들어, 온몸에(systemically) 투여 당(per dose), 피검자의 몸무게당 0.01 ㎍/kg 내지 10,000 ㎍/kg의 투여량으로, 피검자에게 조성물을 투여할 수 있다. 또 다른 예에서, 투여량은, 투여 당 피검자의 몸무게 당 1 ㎍/kg 내지 100 ㎍/kg 이다. 또 다른 예에서, 투여량은, 투여당 피검자의 몸무게 당 1 ㎍/kg 내지 30 ㎍/kg, 예를 들어 투여당 피검자의 몸무게 당 1 ㎍/kg 내지 10 ㎍/kg 이다.

[0208] 치료학적 효율(therapeutic efficacy)이 최적이 되도록, 캡핑이 제거된 M6P 잔기를 갖는 분자는, 상이한 투여 치료 계획(different dosing regimens)에서 첫 번째로 투여될 수 있다. 단위 투여 및 치료 계획은, 예를 들어 포유동물의 종(pecies of mammal), 이의 면역 상태(immune status), 포유동물의 몸무게를 포함하는 요인에 의존한다. 일반적으로, 조직에서의 이러한 분자의 수준을, 예를 들어 주어진 치료 계획의 효율성을 결정하기 위해, 임상적인 테스트 절차(clinical testing procedure)의 일부로서 적절한 스크리닝 검정(screening assays)을 사용하여 모니터할 수 있다.

[0210] 캡핑이 제거된 M6P 잔기를 갖는 분자에 대한 투여 횟수는, 의료의 전문가(예를 들어, 의사나 간호사)의 임상적인 판단 및 기술 내에 있다. 일반적으로, 투여 치료 계획은 최적의 투여 파라미터(optimal administration parameters)를 설정할 수 있는 임상적인 시도(clinical trials)에 의해 설정된다. 그러나, 전문가는, 피검자의 나이, 건강, 몸무게, 성별 및 의학적인 상태에 따라 이러한 투여 치료계획은 다양할 수 있다. 투여량의 횟수는 치료가 예방적(prophylactic) 또는 치료학적(therapeutic) 인지에 따라서 다양할 수 있다.

[0212] 이러한 분자 또는 약제학적 조성물의 독성(Toxicity) 및 치료학적 효과(therapeutic efficacy)는, 예를 들어 세포 배양액 또는 실험 동물(experimental animals)에서의 공지된 약제학적 절차에 의해 결정될 수 있다. 이러한 절차는, 예를 들어, LD50[50 % 의 개체군에게 치명적인 투여량 (dose lethal to 50% of the population)] 및 ED50[50 %의 개체군에서의 치료학적 유효한 투여량(dose therapeutically effective in 50% of the population)]를 측정하기 위해 사용될 수 있다. 독성 및 치료학적 효과 사이의 투여량 비율은 치료 지수(therapeutic index)이고, 이는 비율 LD50/ED50로서 표현될 수 있다. 높은 치료학적 지수(high therapeutic indices)를 나타내는 약제학적 조성물이 바람직하다. 독성 부작용을 나타내는 약제학적 조성물이 사용될 수 있는 반면에, 주의(care)는, 정상 세포[예를 들어, 비-타겟 세포(non-target cells)]에 대한 잠재적인 손상을 최소화하기 위해 영향을 받은 조직(affected tissue)의 부위(site)에 이러한 화합물을 타겟하는(targets) 전달 시스템(delivery system)을 설계하기 위한 조치를 취해야 하고, 이로 인해 부작용을 줄일 수 있다.

[0214] 세포의 배양 검정(cell culture assays) 및 동물 연구(animal studies)로부터 획득된 데이터는 적절한 피검자

(예를 들어, 인간 환자)에 사용하기 위한 투여량의 범위를 공식화하는데에 사용될 수 있다. 이러한 조성물의 투여량은, 독성이 적거나 없는 ED50을 포함하는 순환 농도(circulating concentrations)의 범위 내에 일반적으로 놓인다. 투여량은 사용된 제형(dosage form) 및 이용된 투여의 경로에 따라 이러한 범위 내에 다양할 수 있다. 본원에 사용된 바와 같은 약제학적 조성물에 대해서[예를 들어, 피검자에서의 대사 질환을 치료하기 위해], 치료학적 유효량은 세포 배양 검정으로부터 처음에 평가될 수 있다. 투여량은, 세포 배양에서의 결정된 바와 같은 IC50 [즉, 징후의 최대한의 저해의 반(half-maximal inhibition)을 실행하는 약제학적 조성물의 농도]을 포함하는 순환 플라즈마 농도 범위(circulating plasma concentration range)를 달성하기 위해 동물 모델에서 제형화될 수 있다. 이러한 정보는 인간에서의 유효한 투여량을 보다 정확하게 측정하기 위해 사용될 수 있다. 플라즈마(plasma)에서의 수준은, 예를 들어 고성능 액체 크로마토그래피(high performance liquid chromatography)에서 측정될 수 있다.

[0216] 본원에서 나타낸 바와 같이, 캡핑이 제거된 M6P 잔기를 갖는 분자의 "치료학적 유효량(therapeutically effective amount)"은 처리된 피검자(treated subject)에게 의학상으로 원하는 결과[예를 들어, 대사 질환의 하나 또는 그 이상의 징후의 개선(amelioration)]를 생산할 수 있는 분자의 양이다. 치료학적으로 유효량(즉, 유효한 투여량)은 피검자 또는 샘플 무게의 킬로그램 당 화합물의 밀리그램(milligram) 또는 마이크로그램(microgram)의 양[예를 들어, 킬로그램당 약 1 마이크로그램 내지 킬로그램당 약 500 밀리그램, 킬로그램당 약 100 마이크로그램 내지 킬로그램당 약 5 밀리그램, 킬로그램당 약 1 마이크로그램 내지 킬로그램당 약 50 마이크로 그램]을 포함할 수 있다.

[0217]

[0218] 피검자는, 어떠한 동물, 예를 들어 인간(예를 들어, 인간 환자) 또는 비-인간 영장류(non-human primate)[예를 들어, 침팬치, 개코원숭이(baboon), 또는 원숭이], 마우스, 랫, 토끼, 기니피그(guinea pig), 게르빌루스쥐(gerbil), 햄스터(hamster), 말, 가축의 종류(type)(예를 들어, 소, 돼지, 양, 또는 염소), 개, 고양이 또는 고래(whale)일 수 있다.

[0220] 본원에 기재된 분자 또는 이의 약제학적 조성물은, 다른 치료, 예를 들어 대사 질환(예를 들어, 리소좀 축적 질환)의 치료를 갖는 병용요법(combination therapy)으로서 피검자에게 투여될 수 있다. 예를 들어, 병용요법은, 대사 질환(예를 들어, 리소좀 축적 질환)이 진행될, (또는 대사 질환을 갖는 혐의를 가질) 위험이 있거나 또는 이를 갖는 피검자에게 치료학적 이익을 제공하는 하나 또는 그 이상의 추가적인 제제를 피검자(예를 들어, 인간 환자)에게 투여하는 것을 포함할 수 있다. 따라서, 화합물 또는 약제학적 조성물 및 하나 또는 그 이상의 추가적인 제제를 동사에 투여될 수 있다. 선택적으로, 분자는 첫 번째로 투여될 수 있고, 하나 또는 그 이상의 부가적인 제제는 두 번째로 투여될 수 있고, 또는 반대로 투여될 수 있다.

[0221]

[0222] 이전의 치료(previous)가 특히 독성이 있다면[예를 들어, 현저한 부작용 프로파일(side-effect profiles)을 갖는 대사 질환를 위한 치료], 본원에 기재된 분자의 투여는, 독성은 없지만 동일하거나 또는 개선된 치료 이익이 주어진 충분한 수준으로, 이전의 치료(previously therapy)의 양을 상쇄(offset) 및/또는 줄이는데 사용될 수 있음을 인지할 수 있는 것이다.

[0224] 본원에 기재된 약제학적 조성물의 어떠한 것은, 투여를 위한 지시(instructions)를 함께 갖는 용기(container), 팩(pack) 또는 디스펜서(dispenser)를 포함할 수 있다.

[0226] 하기의 내용은 본 발명의 실행의 예이다. 이들은 어떠한 방식으로든 본 발명의 범위를 한정하는 것으로 이해되지 않는다.

[0228] **[실시예]**

[0229] **[실시예 1]**

[0230] **인산화된 N-글리칸(phosphorylated N-glycans)의 보다 높은 정도(higher** degree)를 갖는 *야로위아 리폴리티카*의 조성(Creation)

[0231] *Y. 리폴리티카*에서의 글리칸의 인산화반응을 상향조절하기(upregulate) 위해, 스트레인 MTLY60 을 MNN4 유전자의 2 개의 여분의 복제물과 함께, 분리된 발현 벡터(separate expression vector)에서 각각을 형질전환시켰다. MNN4 유전자는 효모에서의 글리칸 인산화반응을 증가시키는데 포함되었다. 도 1 은 TEF 프로모터의 조절 하에서 MNN4 오픈 리딩 프레임(MNN4 open reading frame)을 포함하는, pYLTmAXMnn4 을 생산하기 위해 MNN4 유전자

가 클로닝된 pYLTmAX 플라즈마의 도식도를 포함한다. 스트레인은, MNN4 유전자의 두 개의 여분의 복제물, hp4d 프로모터의 조절 하의 하나 및 TEF1 프로모터의 조절 하의 하나를 포함하는 것으로 만들어졌다. N-글리칸은, 스트레인 MTLY60Doch1[MNN4의 1 야생형 카피(1 wild type copy)], 스트레인 MTLY60Doch1+Hp4dMNN4[MNN4의 1WT +1 여분의 카피(1WT +1 extra copy of MNN4)] 및 스트레인 MTLY60Doch1+Hp4dMNN4+TEFMNN4 [Mnn4의 1WT + 2 여분의 카피(1WT +2 extra copies of Mnn4)]로부터 제조되었고, DNA 염기서열 분석기-지원된[DNA sequencer-assisted, DSA)], 형광단-지원된 탄수화물 전기영동법[fluorophore-assisted carbohydrate electrophoresis, FACE)]로 검정하였다. Callewaert *et al.*, *Glycobiology* 11(4):275-281 (2001)를 참고하라. 도 2 에서의 결과를 기초로 하여, 모노 인산화된 피크(mono phosphorylated peak)는 1 여분의 카피를 갖는 스트레인에서 상향 조절되었고(upregulated), 더블 인산화의 피크(peak of double phosphorylation)가 나타남을 추정할 수 있다. 2 여분의 카피를 갖는 스트레인에서, 더블 인산화된 피크(double phosphorylated peak)는 보다 높았고, 중립의 $Man_8GlcNAc_2$ 당의 피크는 보다 낮았다.

[0233] **[실시예 2]**

[0234] **균류 기원(fungal origin)의 인산화된 글리칸에 나타난 캡핑된(capping) 만**노스의 캡핑을 제거할 수 있는 만노시다제의 확인

[0235] 효모 및 곰팡이(filamentous fungi)에 의한 당의 인산화반응은, 만노스-포스포-만노스 디-에스테르 결합(mannose-phospho-mannose di-ester linkage)을 야기한다(도 3). 인산염이 모노-에스테르 결합에 있는 구조를 수득하기 위해, 만노시다제는, 높은 만노스 글리칸 구조의 만노스의 6 위치에 부착된 인산염을 남기는(leaving), 만노스-포스페이트 결합(mannose-phosphate linkage)을 가수분해할 수 있음을 요구한다. Chiba *et al.*, *Glycobiology*, 12(12):821-8 (2002)는, *Cellulomonas* 종으로부터의 만노시다제가 만노스의 캡핑을 제거할 수 있음을 나타내었다. 그러나, Chiba *et al.* 오직 부분적으로 만노시다제 단백질을 정제하였고, 단백질을 코드화하는 유전자를 확인할 수 없었다.

[0237] *셀로로시미크로비움 셀룰란스(Cellulosimicrobium cellulans)*[또한 *오엘스코*비아 잔티네올리티카(Oerskovia xanthineolytica) 및 *아르스로박터 루테우스(Arthrobacter luteus)*로서 알려짐] 분리물(isolate)을 LMG 박테리아 수집물(collection)로부터 수득하였고, 만노시다제 활성도의 생산에 대해서 테스트하였다. 배지에서 만노시다제를 분비하게 하기 위해, 박테리아를 30 ℃ 로 배지를 포함하는 만난(mannan) 성장시켰다. 박테리아성 상청액(Bacterial supernatants)(SN)을 배양물로부터 수득하였고, 실시예1 에 기재된 MNN4 과잉발현된 스트레인(overexpressing strain)으로부터 유도된 분리된 N-글리칸과 함께 SN 을 배양하여 원하는 만노시다제 활성도를 테스트하였다. 배양 후에, 글리칸을 DSA-FACE 로 검정하였다(도 4).

[0239] SN 으로 처리한 후에, 글리칸은 추가적인 전하(charge)를 얻었고, 전기장(electric field)에서 보다 빨리 이동하였고, 일렉트로페로그램(electroferogram)의 좌측편(left hand side)으로 이동하였다. 만약 이러한 빠른-러닝 구조물(fast-running structures)이 실제로 포스포모노에스테르-치환된 높은 만노스 글리칸(phosphomonoester-substituted high mannose glycans)이라면, 이들은 동일한 위치에서 러닝한(running) 중성의 산물(neutral products)보다 크기에서 보다 클 것이다. 포스파타아제(phosphatase)로 이러한 글리칸의 처리는 보다 천천히 움직이는(run) 중성의 올리고당을 야기할 것이다. 도 5 에 나타낸 바와 같이, 송아지 장내 포스파타제(calf intestine phosphatase, CIP)로 처리는, 포스파타제가 말단이고, 만노스의 캡핑이 제거됨을 증명하는, 보다 낮은 전기영동이동도(lower electrophoretic mobility)를 나타내는 피크(peaks)를 결과로서 나타내었다.

[0241] **[실시예 3]**

[0242] **만노시다제의 부분적인 정제 및 추가적인 확인**

[0243] 만노시다제를 정제하기 위해, *C. 셀룰란스*을 1 L의 배지 B (Bagiyan *et al.*, *Eur. J. Biochem.* 249(1):286-92 (1997))또는 배지 A (Chiba et al., 2002, *supra*)에서 성장시켰다. 표 1 을 참고하라. 그 후에, 배지를 40 % 및 80 %의 황산암모늄(ammonium sulphate)으로 침전시키고(precipitated), 샘플을 SDS-PAGE 로 분석하였다. 황산암모늄 분획을 1 mM $CaCl_2$ 을 갖는 20mM Na-인산 완충용액(phosphate buffer) pH 6.5 에 반하여(against) 투석한 다음에, MNN4 과잉발현된 스트레인(실시예 1)으로부터 유도된 올리고당에서 활성도에 대해서 테스트하였다.

[0245] **[표 1]**

배지 성분

배지 A (1 liter)	**배지 B (1 liter)**
2 g 만난(mannan)	2 g 만난
0.5 g $(NH_4)_2SO_4$	2 g $(NH_4)_2SO_4$
0.4 g $MgSO_4.7H_2O$	0.02 g $MgSO_4.7H_2O$
20 mg $FeSO_4.7H_2O$	1 mg $FeSO_4.7H_2O$
60 mg $CaCl_2.2H_2O$	1 g 효모 추출물
1 g 효모 추출물(yeast extract)	4.2 g KOH
7.54 g K_2HPO_4	14 g KH_2PO_4
2.32 g KH_2PO_4	

[0246]

[0248] 둘 다의 배양 조건(cultivation conditions)은 캡핑을 제거하는 활성도(uncapping activity)의 생성을 결과로 나타낸다. 배지 B 로부터 유도된 40 %의 황산암모늄 분획(fractions)만이 활성도를 나타낸 반면에, 배지 A 상청액의 모든 분획은 활성도를 나타내었다.

[0250] 배지 A 배양으로부터 유도된 40 % 의 황산암모늄 샘플은 실리카-기저 겔 여과 컬럼(silica-based gel filtration column) 상에서 추가적으로 정제되었다(도 6). 이는 약 670 kDa 의 숄더(shoulder)를 갖는 피크를 결과로 나타내었다.

[0252] 인산염의 캡핑이 제거된 활성도를 테스트하기 위해, 모든 용리 분획(elution fractions)을, [하기의 CIP-소화(digest)와 함께 또는 없이] MNN4 과잉발현된 *야로위아 리폴리티카* 스트레인(실시예 1에 기재됨)으로부터 유도된 올리고당과 함께 배양하였다. 캡핑 제거(decapping) 및 만노시다제의 활성도는 샘플의 모두에서 관찰되었다. 샘플을 SDS-PAGE 에서 또한 분석하였고(도 7), 이는 단지 하나의 단백질 밴드(protein band)가 아닌 여러 가지의 단백질 밴드를 나타내었다. 몇몇의 밴드는 겔(gel)로부터 절단되었고(cut out), 이들 서열의 일부를 질량 분석법(Mass Spectrometry)을 사용한 드 노보 펩티드 시퀀싱(de novo peptide sequencing)에 의해 분석하였다.

[0254] 드 노보 시퀀싱은 몇몇의 펩티드 서열을 결과로 나타내었고, 이는 BLAST를 사용하여 불필요하지 않는 데이터베이스(non redundant database)에서 상기 서열에 대해서 비교하였다. 하기의 단백질에 대해 상동을 갖는 펩티드를 확인하였다: 포스포디에스테라아제(phosphodiesterase), 가상 단백질(hypothetical protein), 추정상의 알파-1,2 만노시다제(putative alpha-1,2 mannosidase)(표 2에 나타낸 확인된 펩티드)[*마그네토스피릴룸(Magnetospirillum)*으로부터의 만노시다제와 상동], 및 아미노펩티다아제 Y(aminopeptidase Y). 포스포디에스테라아제는 가능한 후보물질이지만, 6 개의 펩티드 중의 2 개만이 히트(hit)를 주었다. 만노시다제는 또한 2 개의 상이한 만노시다제에 대한 3/5 및 5/5 히트를 갖는 후보물질(candidate)이다.

[0256] **[표 2]**

펩티드 서열

펩티드 서열	SEQ ID NO
SAYQSFTTR	1
VWGFSHR	2
VEGGWLPR	3
TQGNNFALLLPER	4
DVHAELTAMAR	5

[0257]

[0259] **[실시예 4]**

[0260] **<u>전체 게놈 시퀀싱(whole genome sequencing)을 기초로 한 원하는 서열을 갖</u>**는 만노시다제의 확인

[0262] *원하는 활성도에 대해 코딩하는 만노시다제 유전자를 확인하기 위해, *C. 셀*룰란스의 게놈을 Titanium 454 sequencing (Eurofins MWG Operon)을 사용하여 서열화하였다(sequenced). 높은 GC 함량 때문에, 시퀀싱(sequencing)이 오직 부분적(1.96 Mbases)이고, 낮은 품질[오직 낮은 평균 콘틱 크기(low average contig size)]을 나타내었다. 에멀젼 PCR(emulsion PCR, emPCR) 동안에 루프 형성을 야기하는 게놈의 높은 GC 함량은

삭제 및 매우 짧은 서열을 결과로 나타내었다.

[0264] 이러한 문제는, Roche에 의한 베타 테스트(beta testing)를 가능하게 만들 수 있는 emPCR에 대한 새로운 시퀀싱 화학(new sequencing chemistry)을 사용하여 극복하였다. 이는 보다 개선된 서열(4.7 Mbases)을 제공하였고, 글리코실 히드롤라제 패밀리 92(glycosyl hydrolase family 92)에 속하는 5 만노시다제 유전자의 확인을 가능하게 하였고, 실시예3 에 기재된 펩티드로부터의 서열에 해당하는 이들 중 하나(CcMan1, SEQ ID NO:6)를 수득하였다. 패밀리 38 및 47 로부터 어떠한 만노시다제가 발견되지 않았다. 각각의 CcMan1-CcMan4 의 개시 코돈(start codon)은 MetaGeneAnnotator(metagene.cb.k.u-tokyo.ac.jp/metagene/ 에서의 월드와이드 웹을 참고하라)에 의해 예측되었고, 알려진 유전자와 함께 Blast 결과를 비교하였다. CcMan5의 개시 코돈은, 이는 서열로부터의 누락(missing) 때문에 예측될 수 없었다. 각각의 유전자의 신호 서열은 두 가지 방법[뉴럴 네트워크(neural networks) 및 히덴 마르코프 모델(hidden markov models)]으로 신호 P (cbs.dtu.dk/services/SignalP/ 에서의 월드와이드웹을 참고하라)로 예측되었다.

[0266] 도 8 내지 12는, *C. 셀룰란스*로부터의 5 만노시다제 유전자의 뉴클레오티드 및 코드화된 아미노산 서열을 포함한다.

[0268] **[실시예 5]**

[0269] **캡핑이 제거된(decapping)** *생체외* **또는** *생체내* **만노스의 이형 발**현(Heterologous expression)

[0270] 만노시다제에 의한 효모 유형 인산화반응(yeast type phosphorylation)의 캡핑 제거(decapping)를 가능하게 하기 위해, 이는 치료학적 용도의 단백질이 발현되는 상이한 숙주(different host) 또는 동일한 균류 숙주(same fungal host)에서 이종으로(heterologously) 발현되어야 한다. 후자의 경우에, 이는 공동 분비될(co-secreted) 수 있거나 또는 세포내 구간[예를 들어, 골지체 또는 소포체(endoplasmic reticulum)]에 타겟될(targeted) 수 있다. 이는 발현 벡터에서 프로모터 다음에 실시가능하게 연결된 유전자[타겟 숙주에 대해 최적화된 코돈 또는 아님(be it codon optimised for the target host or not)]를 클로닝(cloning)하여 실행할 수 있다. 만노시다제는, 쉬운 검출 및 정제를 가능하게 하는 에피토프 태그(epitope tag)와 함께 태그될 수 있거나, 이와 같이 발현될 수 있다. 세균 세포(bacterial cell)의 주변세포질(periplasm)에 분비될 수 있거나 또는 세포 내에 발현될 수 있다. 균류 숙주(fungal host)에서의 발현의 경우에, 상기 서열은 세포내 구간 또는 둘 다에 상기 단백질을 타겟하기 타겟팅 신호(targeting signal) 또는 분비신호(secretion signal)를 포함할 수 있다. 표 3 은 균류 유기체(fungal organisms)에서 발현을 위한 분비 및 타겟팅 신호의 목록을 포함한다. 이러한 발현 벡터의 예는 도 13 에 나타내었다.

[0272] **[표 3]**

균류 유기체에서 발현을 위한 분비 및 타겟팅 신호

분비 신호	**골지 타겟팅 신호 (Golgi targeting signal)**	
	N-말단	**C-말단**
LIP2 prepro	*MNN2*	*KEX2*
LIP2 pre	*MNN4*	
S.c. α mating factor	*MNN6*	
XPR2 prepro	*MNN1*	
XPR2 pre	*MNN9*	
	OCH1	
	SEC12	
	KRE2	

[0273]

[0275] CcMan1-Man5 유전자는 *E. coli* 에서의 발현에 최적화된 코돈이다. 코돈 최적화된 서열(codon optimized sequences)에 대한 도 14 내지 18 을 참고하라. 표 4 는 각각의 코돈 최적화된 뉴클레오티드 서열의 길이 및 신호 서열이 없는 각각의 폴리펩티드의 예측된 분자량을 포함한다.

[0277] **[표 4]**

코돈 최적화된 유전자

	길이 (bp)	SEQ ID NO	코드화된 생산물의 크기(kDa)
CcMan1	2613	16	92.6
CcMan2	3483	17	121.6
CcMan3	3363	18	116
CcMan4	5283	19	184
CcMan5	4956	20	173

[0278]

[0280] **[실시예 6]**

[0281] *C. 셀룰란스* **글리코실 히드롤라제(Glycosyl Hydrolase, GH) 패밀리 92 효소**의 클로닝 및 활성도

[0282] CcMan1-CcMan5 코돈 최적화된 핵산(codon optimized nucleic acids)을, 주변 세포질 발현을 위해, Spy 신호 서열을 포함하는 *E. coli* 벡터 pLSH36 , 및/또는 DsbA 신호 서열을 포함하는 pLSAH36 내로 클로닝하였다. pLSH36 및 pLSAH36 둘 다는 폴리히스티딘 태그(polyhistidine tag) 및 쥣과 카스파아제-3 사이트(murine caspase-3 site)를 갖는 코드화된 폴리펩티드를 결과적으로 나타내고, 이는 정제하는 동안에 His6-tag의 제거에 사용될 수 있다. 도 19 는, *C. 셀룰란스* GH92 유전자를 벡터 내로 도입하기 위한 클로닝 전략 및pLSH36 및 pLSAH36 벡터의 도식도를 포함한다. 클로닝 후에, 상이한 만노시다제를 *E. coli* BL21 + pICa2 발현 스트레인 내로 형질전환시켰다. 형질전환된 스트레인은, 0.5 내지 1 의 광밀도(optical density)로 성장시켰고, 1 mM의 이소프로필 β-D-1-티오칼록토피라노시드(isopropyl β-D-1-thiogalactopyranoside, IPTG)와 유도되었다. 상이한 세포 분획물(cell fractions)[배지, 주변세포질, 가용성 및 불용성 분획물(soluble and insoluble fraction)]을 분리하였고, 항-His 6 항체와 함께 SDS PAGE 및 웨스턴 블로팅(Western blotting)으로 분석하였다. CcMan1, CcMan2, 및 CcMan3에 대해서, 발현은 모든 분획물에서 발견되었다. CcMan4 및 CcMan5에 대해서, 발현은 가용성 분획물에서 가장 높았지만, 몇몇 발현은 그 밖의 분획물에서 발견되었다.

[0284] CcMan1-CcMan5 단백질의 활성도를 측정하기 위해, 활성도 테스트를, Chiba *et al.*, 2002, *supra* 에 나타낸 바와 같이 메틸움벨리페릴 알파 만노시드(methylumbelliferyl alpha mannoside, MUM)을 사용하여 수행하였다. CcMan1 및 CcMan2 에 대해서, 배지 및 주변 세포질 샘플은 MUM을 약하게 가수분해하였지만, CcMan3 및 CcMan5 는 MUM 을 가수분해할 수 없었다. CcMan4 의 가용성 분획물은, CcMan4 가 α1,2-만노시다제 활성도를 갖는 유일한 만노시다제임을 나타내는, 가장 높은 형광성 신호(highest fluorescent signal)를 제공한다.

[0286] 5의 상이한 *C. 셀룰란스* 만노시다제의 모든 배지 및 주변 세포질의 샘플은, 이들이 당을 분해할 수 있고, 만노스-6-인산염(mannose-6-phosphate)의 만노스의 캡핑을 제거(uncap) 할 수 있는지를 보기 위해서, 실시예 1의 MNN4 과잉발현되는 스트레인으로부터 유도된 당(본원에서 MNN4 당으로서 본원에서 언급됨)에서 또한 테스트되었다. 당을 밤새 배양시켰고, DNA 시퀀서-지원된, 형광단-지원된 탄수화물 전기영동(DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis, DSA-FACE)에 의해 분석하였다. 배지 샘플의 당 프로파일(sugar profiles)은, 일렉트로페로그램에서의 무관한 피크(irrelevant peaks)를 결과로서 나타낸 배지 프리젠테이션(medium presentation)에서의 형광성 분자(fluorophoric molecules) 때문에 분석할 수 없다. CcMan1, CcMan2 및 CcMan3 의 주변세포질의 당 프로파일은 분해(degradation) 및 캡핑 제거를 나타내지 않았고, CcMan4 은 분해를 나타내었고, CcMan5 는 캡핑 제거 활성도를 나타내었다(도 20). 캡핑이 제거된 당에서의 CIP-소화(digest)는 중성의 Man8 로 이동한 탈인산화된 피크(dephosphorylated peaks)로서의 CcMan5의 캡핑 제거된 활성도를 확인해주었다.

[0288] 유효한 만노시다제 CcMan4 및 CcMan5 는 Bt3990 (744 AA) 및 Bt2199 (739 AA), 알려진 구조를 갖는 패밀리 92 만노시다제와 함께 배열되었다[Zhu *et al.*, *Nat. Chem. Biol.*, 6(2):125-32. Epub 2009 Dec 27 (2010)를 참고하라]. 도 21 을 참고하라. CcMan4 및 CcMan5 의 첫 번째 부분만이 Bt3990 및 Bt2199 과 배열되었기 때문에, 이들이 큰 단백질이기 때문에, 이는 각각의 단백질을 별도로 첫 번째 도메인을 클로닝하고, 활성도를 테스트하기로 결정하였다. CcMan4domain (1-3357 bp, 즉, SEQ ID NO:20의 뉴클레오티드 1-3357) and CcMan5domain (1-2322 bp, 즉, SEQ ID NO:2의 뉴클레오티드 1-2322)을 pLSAH36 *E. coli* 발현 벡터 내로 클로닝하였다. pLSAH36 클로닝 벡터의 도식도에 대한 도 19를 참고하라. 발현 벡터를 *E. coli* BL21 + pICa2 발현 스트레인 내로 형질

전환시키고, 이를 0.5 내지 1 의 OD 로 성장시키고, 1 mM IPTG 와 함께 유도시켰다. 상이한 세포 분획물(cell fractions)[배지, 주변세포질, 가용성 및 불용성 분획물(soluble and insoluble fraction)]을 분리하였고, 항-His 6 항체와 함께 SDS PAGE 및 웨스턴 블로팅(Western blotting)으로 분석하였다. 발현은 모든 4 세포 분획물에서 발견되었다.

[0290] 도메인의 활성도는 Mnn4 당에서 테스트하였다. 여기에, 각각의 CcMan4domain 및 CcMan5domain 의 주변세포질 분획물을 Mnn4 당의 존재 하에서 배양하였고(도 22), DSA-FACE로 분석하였다. 이러한 실험은, 분해를 발견할 수 없기 때문에, CcMan4domain 은 이의 만노시다제 활성도를 상실함을 나타내었다[도 22, 패널(panel) 4]. 이와 대조적으로, CcMan5domain 은 이의 캡핑이 제거된 활성도를 유지하였다(도 6, 패널 5).

[0292] **[실시예 7]**

[0293] **CcMan5 및 이의 패밀리 92 상동 도메인(homologous domain)의 생산 및 정제**

[0294] 재조합 CCman5 도메인(SEQ ID NO:20의 뉴클레오티드 1-4995) 및 CCMan5 도메인(SEQ ID NO:20의 뉴클레오티드 1-2322)은, 발현 벡터 pLSAHCcMan5 및 pLSAHCcMan5domain 로 형질전환시킨 *E. coli* 스트레인 BL21codon + pICA2 에서 발현되었다. 발현을 λpL-프로모터(λpL-promotor)의 조절 하에서 IPTG 에 의해 유도되었다(WO 98/48025 및 WO 04/074488를 참고하라). pLSAH의 설명(description)에 대한 실시예 6 및 도 19 를 참고하라. 형질전환된 박테리아를, 암피실린(100 μg/ml) 및 1 % 글리세롤로 보충된 LB 배지가 제공된 20 리터의 발효조(fermenter)에서 1/100 접종(inoculation) 전에, 암피실린(100 μg/ml) 및 카나마이신(50 μg/ml)로 보충된 Luria Bertani(LB) 배지에서 28 ℃로 밤새 성장시켰다. 처음의 교반 및 공기 흐름(airflow)은 각각 200 rpm 및 1.5 l/min이고, 30 % 로 pO_2 를 유지되도록 자동적으로 조종되었다. 온도를 28 ℃로 유지시켰다. 세포를 $A_{600\ nm}$ = 1.0의 광밀도(optical density)로 성장시키고, 20 ℃ 에서 이동시키고, 발현은 1 mM IPTG 의 첨가로 밤새 유도되었다. 세포를 재취한 다음에 -20 ℃ 로 동결시켰다. 해동(thawing) 후에, 세포를, 50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 1 mM PMSF 및 10 μg/ml DNaseI 에서 3 ml/g 의 농도로 서서히 재현탁시켰다(resuspended). 주변세포질 분획물(periplasmic fraction)을 4 ℃로 1 h 동안 세포 현탁액(cell suspension)을 교반시켜 제조하였고, 30 min 동안 18,000 x g 로 원심분리하여 분리하였다. 모든 단계를 4 ℃에서 처리하였다. 투명한 상청액을 20 ml Ni-Sepharose 6 FF column (GE Healthcare)에 적용시키고, 20 mM NaH_2PO_4 pH 7.4, 300 mM NaCl, 20 mM 이미다졸, 0.1 % CHAPS 와 함께 평형을 유지시켰다. 컬럼을, 동일한 완충용액에서 50 mM의 이미다졸과 함께 추가의 세척 단계 후에, 20 mM NaH_2PO_4 pH 7.4, 20 mM NaCl, 400 mM 이미다졸, 0.1 % CHAPS 와 용리시켰다(eluted). 용리 분획물(elution fraction)을 20 mM Tris pH 8.0, 0.1 % CHAPS 와 1/10으로 희석시키고, 오염물질(contaminant)을 제거하기 위해 14 ml Source 15Q column (GE Healthcare)에 로딩시켰다(loaded). 평형(equilibration) 후에, 관심 단백질을, 20 mM 트리스, 0.1 % CHAPS 에 0 내지 1 M 의 10 컬럼 부피(column volumes)의 NaCl 상에서 선형 구배(linear gradient)로 용리시켰다. 분획물을 포함하는 CcMan5 및 CcMan5 도메인을, 러닝 용액(running solution)으로서의 PBS 와 함께 HiLoad 26/60 Superdex 200 prep grade 에 추가적으로 주입시켰다. 수득한 분획물을 항-His6 항체와 함께 SDS-PAGE 및 웨스턴 블로팅(western blotting)으로 분석하였다. 마지막으로, 농도를 BCA 검정(Pierce)을 사용하여 측정하였다. 전장 CcMan5 단백질에 대한 정제된 수득물은 5.7 mg이고, CcMan5 패밀리 92 도메인에 대해서, 이들의 20 L 발효물(fermentations)로부터 이는 110 mg 이었고, 이는 패밀리 92 도메인만이 생산될 수 있고, 보다 높은 수득률로 정제될 수 있음을 나타낸다. 정제된 CcMan5 도메인의 활성도를, 실시예 6 에 나타낸 바와 같은 Mnn4 분리된 당에서 테스트하였다. 캡핑이 제거된 당 프로파일을 수득하였다.

[0296] **[실시예 8]**

[0297] CcMan 5 도메인의 구조

[0298] $CcMan5_{1-774}$ [SEQ ID NO:20의 뉴클레오티드 1 내지 2322에 의해 코드화된, SEQ ID NO:50의 잔기 1 내지 774; 이의 자연 발생적인 선도 서열이 제거된 후(after removal its natural leader sequence)의 성숙 단백질(mature protein)에 상응함]는, DsbA 선도 서열(leader sequence) 후에 N-말단 6xHis tag에서 시작해서 그 다음에 9 아미노산 링커(VGPGSDEVD, SEQ ID NO:21)를 갖는 융합단백질로서 *E. coli* BL21 (DE3) 주변 세포질에 발현되었다. 세포를, 28 ℃ 에서 100 μg/ml 의 카나마이신 및 100 μg/ml 의 암피실린을 포함하는 M9 배지에서 배양하였다. 0.4 의 OD_{600} 에서, $CcMan5_{1-774}$ 발현은 1mM IPTG 의 첨가에 의해 유도되었고, 배양물을 18 ℃ 로 밤새 추가적으로 성장시켰다. 밤새 배양(overnight culture)으로부터의 세포를 원심분리에 의해 채취하였고, 세척하고, 스페

로플라스크(spheroplast)를 만들기 위해, 20 mM Tris/HCl pH 8.0, 20% 수크로스, 5 mM EDTA, 및 0.1 mg/ml 라이소자임(lysozyme)을 포함하는 완충용액과 함께 4 ℃에서 20 min 동안 배양하였다. 주변 세포질 단백질을 20 min 동안 20,000 x*g* 로 원심분리에 의해 스페로플라스크로부터 분리하였다. $CcMan5_{1-774}$ 를, 금속 이온 친화성 크로마토그래피(metal ion affinity chromatography)[HisTrap HP, GE Healthcare, 50 mM Tris-HCl pH 8.0, 150 mM NaCl를 포함하는 완충용액 하에서 로딩(loading), 및 400 mM 까지의 이미다졸 구배(imidazole gradient up)를 사용하여 용리시킴], 이온 교환 크로마토그래피[HiTrap Q FF, GE Healthcare, 완충용액: 20 mM Tris-HCl pH 8.0, 40 mM NaCl 및 1 M 까지의 NaCl 구배], 및 소수성 상호작용 크로마토그래피(hydrophobic interaction chromatography)[HiTrap Phenyl HP, GE Healthcare, 로딩 완충용액(loading buffer): 20 mM Tris-HCl pH 8.0, 10 mM NaCl, 1 M $(NH_4)_2SO_4$ 및 0 mM 까지의 $(NH_4)_2SO_4$ 구배를 사용하여 용리시킴]로 주변 세포질 추출물로부터 정제하였다.

[0300] 정제된 $CcMan5_{1-774}$ 를 10 mM Tris-HCl pH 8.0, 10 mM NaCl에서 130 mg/ml로 농축시키고, 접시-모양의 결정(plate-like crystals)(0.2x0.07x0.01 mm^3)을, 0.2 M Na fluoride, 0.1 M Bis-Tris propane pH 7.5 and 20% PEG 3350 을 포함하는 결정화 용액(crystallization solution)을 사용하여 수증기 발산(vapor diffusion)에 의해 성장시켰다. 결정을, 10%(v/v) 글리세롤로 보충된 결정화 용액(crystallization solution)을 포함하는 냉동보존 용액(crystallization solution) 내로 간단하게 이동시켰다. 단일 결정 회절 데이타를 Swiss Light Source에서의 PXIII 빔라인(SLS, Villigen, Swiss) 및 European Synchrotron Radiation Facility에서의 빔라인(beamline) BM30A(ESRF, Grenoble, France)에서 수집하였다. $CcMan5_{1-774}$ 의 구조를, Au L-III 흡수단(absorption edge)에 대응하는, 11.958 keV에서의 SAD 실험으로부터의 실험 단계(experimental phases)의 추정(calculation)에 대한 KAuCl4-흠뻑 젖은(soaked) 결정을 사용하여 해결하였다. 도 33 은, 촉매 중심(catalytic center)의 구조적인 좌표(structural coordinates)를 포함한다. 실험단계로부터의 만들어진 CcMan5 모델(model)은, 각각 19.3 및 23.9 % 의 최종의 R- 및 유리R-인자(final R- and freeR-factor)에 대한 천연의 결정(native crystal)으로 수집된 2Å 해상도 데이터(resolution data)에 대비한 최우추정법(maximum likelihood methods)에 의해 개선되었다. 최종의 모델은, 분자당 11.513 단백질 원소(protein atoms), 860 용매 원소(solvent atoms), 2 Ca^{2+} 이온 및 1 비스-트리스-프로판(bis-tris-propane) 및 글리세롤 분자를 포함하는, 비대칭 유닛(asymmetric unit)당 2 $CcMan5_{1-774}$ 분자(잔기 8 내지 771)를 포함한다.

[0302] 서열 유사성(sequence similarity)을 기초로, CcMan5 는 글리코실 히드롤라아제의 패밀리 92 (family 92 of glycosyl hydrolases, GH92)의 부류에 속하고, 이는 엑소-액팅 알파-만노시다제(exo-acting alpha-mannosidases)로서 나타낸다. α1을 갖는 2 GH92 패밀리 멤버 , 2-만노시다제 활성도에 대한 X-선 구조가 가능하였다: Bt3990 및 Bt2199[각각, PDB 접근 코드(access codes) 2WVX 및 2WVY]. 전체적인 폴딩(overall fold)을 여기에서 해결된 $CcMan5_{1-774}$ 구조로부터의 볼 수 있었고, PDB entry 2xsg 로서 두었고, 각각 Cα원자와 매치된 624 및 621 상에서의 1.99 Å 및 2.12 Å의 r.m.s.d (근평균 표준 편차, root mean standard deviation) 수치를 갖는 Bt3990 및 Bt2199 둘 다에서 나타낸 것에 대응한다. $CcMan5_{1-774}$는 α-나선형 링커(α-helical linker)(잔기 272 내지 290)를 통해 연결된, N-말단 β-샌드위치 도메인(잔기 8 내지 271) 및 C-말단 (αα)6 배럴 도메인(잔기 291 내지 771)인, 두 개의 도메인으로 구성되어 있다. 둘 다의 도메인 사이의 인터페이스(interface)는, 보존된 촉매 Ca^{2+} 이온을 포함하는 스웰로 캐비티(shallow cavity)를 구체화하고, -1 기질 결합 부위(-1 substrate binding site)[명명법: Davies *et al.*, *Biochem. J.* 321:557-9 (1997)] 및 촉매 중심(catalytic center)을 구체화한다(도 23 및 24).

[0304] GH92 패밀리 글리코실 히드롤라제는, 방출된 만노스에서의 아노머 입체배치(anomeric configuration)의 역위(inversion)를 유도하는, 단일 변위 매커니즘(single displacement mechanism)을 통해 글리코시드 결합 가수분해(glycosidic bond hydrolysis)를 촉매하는 Ca^{2+}-의존적인 알파-만노시다제이다(Zhu *et. al.*, 2010, *supra*). $CcMan5_{1-774}$ 에서, 촉매 Ca^{2+}는, 등가의 좌표 평면(equatorial coordination plane)에 놓여진, Asn 588의 카르보닐 산소(carbonyl oxygen), 각각의 Glu589 및 Asp662 의 카르복실 산소, 및 세 가지의 물 원자(W1, W2, W3 -도 23을 참고하라)을 통해 팔면체적으로 편성되었다(octahedrally coordinated). 추가적인 물 분자(W4)는 촉매 중심에 가까운 것으로 나타났고, 보존된 쌍(conserved pair) Asp 660 및 Asp 662 의 카르복실기에 결합한다. 촉매 Ca^{2+} 를 둘러싸는 기질 결합 캐비티(substrate binding cavity)는 잔기 Asn 588, Gln 589, Thr 626, Thr

658, Asp 22, Asn 25, Gly 71, Gly 72, Phe 195, Tyr 196, Arg 405, Trp 354, Tyr 535, 및 Gln536에 의해 줄이 형성된다(lined)(도 23).

[0306] CcMan5 는, 기질로서 만노스-알파-1-포스포-6-만노스 (Man-P-Man)를 받아들이고 알파-1,2-, 알파-1,3-, 알파-1,4- 또는 알파-1,6-만노시다제 활성도가 결여되는 이의 독특한 능력 때문에 GH92 패밀리에서 그 밖의 알파-만노시다제와는 구분된다. 이러한 독특한 기질 특이성이 생기게 하는 CcMan5 활성 부위(active site)에서의 식별할 수 있는 잔기에서의 통찰력을 얻기 위해서, Man-P-Man을 비대칭 유닛(the asymmetric unit)의 분자 B 의 $CcMan5_{1-774}$ 활성 부위 내로 설계하였다(도 25). -1 만노스의 위치(Positioning)는 Bt3990 에서 관찰된 그로스 결합 입체배치(gross binding conformation)를 기초로하고, 아포 활성 부위(apo active site)에 나타난 두 개의 물 분자(W2 및 W3) 및 글리세롤 분자의 위치에 의해 안내되었다. 이러한 방식에서, -1 만노스의 O2, O3, O4 및 O6 히드록실기는, 각각 물 분자 W2, W3 및 글리세롤 분자의 O1 및 O3 히드록실기에 대해 관찰된 이러한 것들에 대한 동등한 위치를 차지한다. 따라서, 만노스 -1 O2 및 O3 히드록실기는 actohedral Ca^{2+} 배위권(coordination sphere)의 적도면에 위치한다. O3 은 Asp 355 카르복실기에 대한 추가적인 수소결합을 형성한다. 후자는, Arg 405 구아니디듐기(guanidinium group)의 H-결합 거리(H-bonding distance) 내로 또한 오는, O4 히드록실에 대한 H-결합을 추가적으로 제공한다. O6 히드록실 및 O5 산소는 Gly 71 아미드를 갖는 H-결합을 포함할 수 있다. 모델링(modeling)에 대해서, -1 만노스는 이의 바닥상태 의자 입체배치에서 유지시켰다. Bt3990 에 대해 관찰된 바와 같이, Ca^{2+} 와 이상적인 배위(idealized coordination) 내로 오기 위한 O2 히드록실기의 위치는, 반 의자 입체배치에 대한 당 고리의 뒤틀림(distortion)을 유도할 수 있다(도 25 를 참고하라). 촉매 동안에 당 고리의 뒤틀림(distortion)이, O2 히드록실과 함께 친핵체(nucleophile)의 유입(incoming)의 1,2-이축방향 상호작용(1,2-diaxial interaction)을 끊기 위해, α-만노시다제에서의 아세틸 중심(acetal)에서의 친핵성 치환(nucleophilic substitution)을 필요로 함은 일반적인 용인되는 내용이다(Vocadlo *et al.*, *Curr. Opin. Chem. Biol.* 12:539-55 (2008)). -1 부위에서의 기질 결합에 대한 수득된 모델은, 물 분자 W4가 아세탈 탄소(acetal carbon)에서의 라인 어택(line attack)에서 친핵체(nucleophile)로서 작용하기 위한 좋은 위치에 놓임을 추가적으로 나타낸다. W4 는, GH92 효소를 도처에 보존된 Asp 660 및 Asp 662 의 카르복실기와 H-결합 상호작용(H-bond interaction)에 있고, 친핵체(nucleophile)의 활성화를 위한 염기 촉매(base catalyst(s))를 형성하도록 제한되었다. 따라서, -1 부위에 결합하는 모델된 기질 결합 및 촉매 잔기 및 친핵체의 위치는, 방출된 만노스에서의 아노머 중심(anomeric center)의 역위와 함께 친핵성 치환(nucleophilic substitution)에 기계론적인 요구(mechanistic requirements)와 일관된다. 상기에 기재된 바와 같이, CcMan5 는 Man-P-Man 을 결합시키고, 가수분해하기 위한 능력에 의해 이 자체로 구별된다. CcMan5 활성 부위에 결합하는 Man-P-Man 에 대해 수득된 모델은 이러한 관점에 대한 이유를 제공한다. 알려진 GH92 패밀리 멤버에서, 글리코시드 결합(glycosidic bond)을 형성하는 아노머 산소는 보존된 글루탐산 잔기(Bt3990에서의 Glu 533)의 카르복실기와 정진기 상호작용(electrostatic interaction)에 있다. 글루탐산 잔기는, 아노머 산소를 결합시키고, 탈리기(leaving group)에 양성자를 부가함으로써 전이 중간물질을 안정화시키는, 촉매 산으로서 역할을 하는 것으로 나타났다(Zhu *et. al.*, 2010, *supra*). CcMan5에서, Bt3990 Glu 533에 대한 등가의 잔기는 글루타민으로 돌연변이되고, 이는 양성자 공여체(proton donor)로서 작용할 수 없고 따라서 만노바이오시드(mannobioside)의 가수분해를 위한 CcMan5 에서의 기능의 손실(loss-of-function)을 설명한다. 그러나, Man-P-Man 기질에서 아노머 산소(anomeric oxygen)에 결합한 인산염은 보다 강한 탈리기를 구성하고, 이는, CcMan5 같은 효소들이 Man-P-Man 기질에 대한 촉매 활성도(catalytic activity)를 왜 유지하는지를 설명하면서, 아노머 산소에 양성자를 가하는 것을 산 촉매에게 요구하지 않는다. 촉매산(catalytic acid)의 치환과 수반되는, Bt3990에서의 Glu 585의 등가물은 CcMan5 에서 Thr (Thr 626)로 교체된다. Bt3990 에서, Glu 585 는 Glu 533 와 상호작용하고, 후자의 pK_a 를 조절하고/조절하거나 2-연결된 만노시다제에서의 탈리기(leaving group)를 결합하는 역할을 하는 것으로 제안되었다(Zhu *et. al.*, 2010, *supra*).

[0308] Gln 536 and Thr 626 쌍에서의 비-산성 잔기(non-acidic residues)로의 돌연변이가 결합 부위에서의 음전기의 정전기 전위(electrostatic potential)의 일부를 완화시키고, 이로 인하여 Man-P-Man 기질에서의 아노머 산소에 대한 인산염 결합을 용인(tolerating)시키는 것 같다. CcMan5 에서, 모델된 인산염 결합 부위(도 25 에서의 P)는 Gly 72의 아미드(amide) 및 Thr 626 에 의해 형성되고, 이들 둘 다는 인산염에서의 비-글리코시드 산소(non-glycosidic oxygens)에 대한 H-결합을 제공할 수 있는 것 같다.

[0310] 마지막으로, $CcMan5_{1-774}$ 활성 부위에서의 Man-p-Man의 모델된 결합을 기초로 하여, 환원된 말단 만노스는, 두 개의 티로신 잔기, Tyr 535 및 Tyr 196의 인근에 오고, 후자의 잔기가 +1 만노스 결합 부위의 일부를 형성함

을 암시한다. 둘 다의 잔기는, 글리칸 트리(glycan tree)의 환원된 말단(reducing end)에서의 글리칸과 추가적인 상호작용을 포함할 수 있는, 쉘로우 클레프트(shallow cleft)의 가장자리에 놓인다.

[0312] **[실시예 9]**

[0313] *Y. 리폴리티카*에서의 α갈락토시다제의 발현

[0314] pre 및 pro 서열 없이, 인간 α-갈락토시다제 A 를 코드화하는 핵산을, *Y.* 리폴리티카에 대한 코돈 최적화(codon optimization) 및 Myc-His 태그(Myc-His tag)의 첨가로 합성하였다. 획득된 서열은 lip2 유전자의 pre 서열(pre sequence) 후의 프레임 내에 클론되었다(cloned). 코돈 최적화된 뉴클레오티드 서열(codon optimized nucleotide sequence)의 뉴클레오티드 서열은 도 26A 에 나타내었고; 아미노산 서열 (SEQ ID NO:23)은 도 26 B 에 나타내었다.

[0316] 2 여분의 복제물(2 extra copies)의 MNN4 및 하나의 복제물의 α-갈락토시다제A 를 갖는 *Y.리폴리티카* MTLY60 는 Ni-NTA 컬럼 상에서 정제시키기 위해 보다 큰 배양물에서 유도되었다. 따라서, 이들을 YTG 에서 성장시키고, 48 시간 동안 2 x 225 ml [2L 쉐이크 플라스크(shake flask)]에서의 올레산 배지(oleic acid medium)에서 유도되었다. 배양물을 원심분리 시킨 다음에, 0.22 ㎛ 필터(filter) 상에서 배지를 여과시켰다. 여과된 배지를, Ni-세파로오스(sepharose) 6 FF 상에서 정제하기 전에 비-단백질 교란 오염물질(non-protein disturbing contaminants)을 제거하기 위해, 20 mM NaH_2PO_4 pH 7.4, 0.5 M NaCl, 20 mM 이미다졸에 대한 세파덱스(sephadex) G25 XK50/100 컬럼(GE Healthcare)에서 탈염시켰다. 탈염된 단백질 분획물을 4.3 ml Ni-세파로스 6 FF 컬럼(GE Healthcare)에 로딩하고(loaded), 20 mM NaH_2PO_4 pH 7.4, 0.5 M NaCl, 20 mM 이미다졸로 평형시키고, 동일한 완충용액에서 50 mM 이미다졸로 세척하고, 20 mM NaH_2PO_4 pH 7.4, 20 mM NaCl, 400 mM 이미다졸로 용리시켰다. Ni-세파로스 컬럼 뒤에 샘플 3-10 및 36-49 를 항-His6 항체를 사용하여, SDS-PAGE 및 웨스턴 블로팅에서 분석하였다. 샘플 40 및 41 에서 쿠마씨(coomassie)에 나타난 약 50 kDa 및 65 kDA 의 단백질 밴드(protein band)는, SDS-PAGE 겔의 쿠마씨 블루 염색(Coomassie blue staining)으로 나타났다. 웨스턴 블로트에서, 50 kDa의 밴드만이 검출되었고, 아마도 α-갈락토시다제 A 이다. 정제된 α-갈락토시다제 A의 추정된 수득률(estimated yield)은 100-125 ㎍/L 배양 배지이다.

[0318] 정제된 샘플은 재조합 α-갈락토시다제A 에서의 당의 유형을 결정하는데 사용되었다. 당을 용액에서 제거하고, 그 뒤에 APTS 로 라벨하였다(labelled). 겔 여과에 의해 샘플을 청소한(cleaning) 후에, 당을 DSA-FACE 로 분석하였다. 예상된 당, 모노 만노인산화된(mono mannophosphorylated) $Man_8GlcNAc_2$ 피크(peak)(P) 및 더블 만노인산화된(double mannophosphorylated) $Man_8GlcNAc_2$ 피크(PP)가 주요한 피크로서 나타났다.

[0320] **[실시예 10]**

[0321] *Y. 리폴리카*에서의 인간 알파 글루코시다아제(glucosidase)의 발현

[0322] *Y. 리폴리카* 스트레인 OXYY1589 는, 세 개의 복제물의 인간 알파 글루코시다아제[산성 알파 글루코시다아제(acid alpha glucosidase)(GAA) 또는 산성 말타아제(acid maltase) EC3.2.1.3] 및 두 개의 복제물의 *Y. 리폴리카* MNN4 유전자를 포함하는 것으로 구성되었다. 스트레인 OXY1589 의 유전자형은 하기와 같다 :

[0323] *MatA, leu2-958, ura3-302, xpr2-322,*

[0324] *gut2-744, ade2-844*

[0326] *POX2-Lip2pre-huGAA:URA3Ex::zeta*

[0327] *POX2-Lip2pre-huGAA:LEU2Ex::zeta*

[0328] *POX2-Lip2pre-hGM-CSF:GUTEx::zeta*

[0329] YlMNN4-POX2-hp4d-YLMNN4 :ADE2*::PT targeted*

[0331] 모든 형질전환은 상이한 선별적인 마커(different selective markers)에 대한 변형(modifications)을 갖는 잘 확립된 프로토콜에 따라 실행되었다. 모든 경우에(별도로 기술되어 있지 않다면), huGAA 통합 단편(integration fragment)은, 발현 플라스미드로부터의 카나마이신 저항 유전자(kanamycin resistance gene)를 제거하기 위해 NotI 제한효소 절단(restriction digestion)에 의해 수득되었다. 결과적으로 생성된 단편은, 아가로스 겔 전기영동(agarose gel electrophoresis)으로 모두 분리한 다음에, 정확한 huGAA 단편의 Qiagen 컬

럼 정제하였다. Y. 리포리티카 발현 벡터 내로 human GAA (huGAA)를 첫 번째 클로닝하고, Y. 리포리티카 MNN4 탠덤 발현 벡터(tandem expression vector)를 구성하여 스트레인 OXYY1589 를 구성하였다. 그리고 난 다음에 세 번의 안정된 통합하는 형질전환(Three stable integrative transformations)을 최종의 huGAA 생산 스트레인 OXYY1589을 수득하기 위해 실시하였다.

[0333] *Y. 리폴리카* 코돈 최적화된 huGAA 발현 벡터(*Y. lipolytica* codon optimized huGAA expression vector): 110 kDA 인간 GAA(huGAA) 전구체를 코드화하는 뉴클레오티드 서열을 화학적으로 합성하였고, 코돈은 *Y. 리폴리카* 발현에 최적화되었다. 합성의 구성물(synthetic construct)에서, pre- 및 pro-huGAA 신호 펩티드는, 단백질이 아미노산 57 에서 시작되는 것과 같이 제거되었다. 발현 벡터에 클로닝하기 위해, huGAA의 합성 ORF(도 27 A)는, 5' 말단 내지 3' 말단의 *Y.리폴리카* LIP2 신호 서열(pre) 다음에, 두 개의 Xxx-Ala 절단 부위(cleavage sites)의 코딩 서열에서의 프레임(frame)에 융합시키고, BamHI 및 AvrII 제한 부위의 측면에 배치되었다(flanked). 구조물을 유도될 수 있는(inducible) POX2 프로모터의 조절 하에 두었다. 융합 구조물(fusion construct)의 완전한 아미노산 서열(complete amino acid sequence)을 도 27B 에 나타내었다.

[0335] 발현 벡터의 일반적인 도식을 도 28 에 나타내었다. 세균성의 일부는 플라스미드 pHSS6 로부터 유도되었고, 세균성 복제의 기원(origin of replication)(ori) 및 카나마이신에 대한 내성(KanR)을 부여하는 카나마이신-내성 유전자(kanamycin-resistant gene)를 포함한다. 통합 카세트(integration cassette)는 a) *야로위아 리폴리티카*에 대한 형질전환을 위한 선별 마커(selection marker)(URA3; LEU2; GUT2), b) 프로모터로 구성된 발현 카세트(expression cassette), c) 신호 서열을 갖는 프레임에서의 huGAA 를 삽입하는 다중 클로닝 사이트(multiple cloning site, MCS), 및 d) LIP2 유전자의 종결자(terminator)를 포함한다. 통합 카세트는 *Y. 리폴리티카* 게놈 내로 안정된 비-상동성 통합(stable non-homologous integration)을 위한 제타 서열(zeta sequences) 옆에 배치되었다. 두 개의 NotI 제한효소 절단 부위(restriction sites)는 형질전환 전에 발현 카세트(expression cassette)의 분리를 가능하게 한다. 각각 URA3, LEU2 및 GUT2 형질전환 마커(transformation markers)를 포함하는 플라스미드 pRAN034, pRAN036 및 OXYP183 는, 각각 huGAA 발현 벡터 pRAN058, pRAN059 및 pRAN060를 생성하는데 사용되었다.

[0337] 탬덤 Y1MNN4 발현 벡터 : Y1MNN4 유전자는 유도성 pPOX2 프로모터 및 (세미)구성요소[(semi)constitutive]인 hp4d 프로모터의 조절 하에서 클론되었다. Y1MNN4의 이러한 두 가지의 발현 카세트는, 게놈의 ADE2 유전자 자리(locus) 및 선별 마커로서의 ADE2 유전자 내로 타겟된 통합(targeted integration)에 대한 *ADE2* 유전자의 플랭킹 영역(flanking regions)(PT)을 포함하는 탬덤 구조물로서의 하나의 벡터 내에 서브클론되었다(subcloned).

[0339] 중간물질(Intermediate) 스트레인 OXYY1569: 첫 번째 형질전환은, 중간의 재조합 스트레인 OXYY1569 를 생산하기 위해 URA3 및 LEU2 마커를 사용한 pRAN058 및 pRAN059 벡터로부터 정제된 발현 카세트의 공동-형질전환(co-transformation)이다. OXYY1569 는 스트레인 GO14 의 게놈에 임의적으로 통합된 pPOX2 프로모터의 조절 하의 huGAA 의 두 가지의 발현 구성물(two expression constructs)을 운반한다(carries).

[0341] OXYY1569 는 하기와 같이 선택되었다. 게놈 DNA 의 PCR 스크리닝(screening)을, *Y. 리폴리티카*의 게놈 내로 이질적인(foreign) huGAA DNA 의 통합을 확인하기 위해 실행하였다. 프라이머(Primers)는 huGAA 뉴클레오티드 서열로부터의 2552 bp 의 단편을 증폭시키기 위해 설계하였다. 게놈 DNA 의 서전 블롯 분석(Southern blot analysis)을, 적어도 2 복제물의 huGAA DNA 의 통합을 확인하기 위해 실행하였다. 특히, OXYY1569 클론으로부터의 게놈 DNA 를 Hind III 와 함께 절단하였고(digest), huGAA DIG 표지된 특정한 프로브(huGAA DIG labeled specific probe)와 함께 조사하였다(probed).

[0343] 높은 수준의 huGAA 를 분비하는 클론을 선별하기 위해서, PCR 스크리닝 및 서전 블롯(Southern blot)에서 양성으로 식별된 몇몇의 임의적으로 선택된 클론을, 표준 절차에 따라 POX2 유도된 조건(inducing conditions) 하에서 진탕 플라스크(shake flasks)에서 성장시켰다. 모든 경우에, 배양 성청액을 72 h 후의-유도물(72h post-induction)을 수집하였고, 표준 웨스턴 블롯 및 효소 활성도 검정 분석으로 스크린하였다(screened). OXYY1569 에서의 우세한 구조물(predominant structure)로서 나타낸 OXYY1569의 N-글리칸 분석은 $Man_8GlcNAc_2$ 이다.

[0345] 중간물질(intermediate) 스트레인 OXYY1584: 재조합 스트레인 OXYY1569 를, OXYY1584를 생산하기 위한 이의 게놈 내로 *Y.리폴리티카* MNN4 유전자의 두 개의 복제물을 통합시키기 위해서 형질전환시켰다. 형질전환을, 플라스미드 OXYP1479B로부터 삭제된(excised) SacII/XmaI 유도된 발현 카세트(SacII/XmaI derived expression cassette)와 함께 실행하였다. 발현 카세트를, *Y.리폴리티카* 게놈의 ADE2 유전자 자리(locus) 내로의 통합을 타겟하도록 설계하였다. 재조합 스트레인은, 증가된 인산화반응(increased phosphorylation)에 대하여 스트레

인 작용(behavior)을 평가하기 위해, 서던 블로팅(Southern blotting) 및 글리칸 분석 후에 선별되었다. 몇몇의 임의적으로 선택된 형질전환체의 게놈 DNA 는 SpeI 절단된 것이고(digested), MNN4 특정한 DIG 표지된 프로브(labeled probe)로 조사하였다. Y. 리폴리티카 게놈의 ADE2 유전자 자리 내로 MNN4 발현 카세트의 정확하게 타겟된 통합(Correct targeted integration)은 4207bp 및 5683bp 밴드를 제공해야한다. 서든 블롯 양성 클론(Southern blot positive clones)을, 표준 진탕 플라스크 절차(standard shake flask procedure)에서 성장시켰다. 분비된 단백질의 N-글리칸 분석을, 중간물질 클론(intermediate clone) OXYY1584 를 선별하기 위해 실행하였다. 모 스트레인(parent stain) OXXY1569과 비교하여, MNN4 과잉-발현 후의 우세한 구조(predominant structures)는 $Man_8GlcNAc_2(PMan)_1$ 및 $Man_8GlcNAc_2$ $(PMan)_2$ 이다.

[0347] 생산 스트레인 OXYY1589 : 최종의 독립 영양 생산 스트레인(final prototrophic production strain) OXYY1589를 생산하기 위해, huGAA의 세 번째 복제물을 재조합 OXYY1584 스트레인의 게놈 내로 통합되었다. 형질전환을 pRAN069 부터의 Not I 절단된 발현 카세트(excised expression cassette)로 실행하였다. 형질전환을, huGAA의 추가적인 복제물의 존재를 위한 gDNA 에서의 PCR로 첫 번째 스크린하였다(screened). huGAA 생산을 평가하기 위해, 임의적으로 선별된 PCR 양성 클론(arbitrary selected PCR positive clones)을 표준 진탕 플라스크 배양(standard shake flask cultivation) 후에 발현에 대해서 추가적으로 분석하였다. 가장 높은 수준의 huGAA (OXYY1589)이 발현된 클론을, 웨스턴 블롯 분석(Western blot analysis) 및 효소적 활성도 검정(enzymatic activity assay) 후에 선택하였다. MP-M8 N-글리칸 및 MP2-M8에 대한 M8의 전환 수준(conversion levels)을 추가적인 huGAA 발현 카세트의 존재에 의해 영향을 미치치 않는다.

[0349] **[실시예 11]**

[0350] **스트레인 OXYY1589의 패드 배치(Fed Batch Cultivation)**

[0351] 스트레인 OXYY1589로부터의 huGAA를 생산하기 위해(실시예 10), 패드 배치 공정(fed batch process)을, 6 내지 8 리터의 작업량(working volume)을 갖는 10 L 교반된 탱크를 사용하여 확립하였다(established). 상기 공정은 두 가지의 단계로 나뉘었다:

[0352] 1) 바이오매스(biomass) 형성을 위한 글루코스에서의 배치 성장(Batch growth);

[0353] 2) 제한된 올레산 공급(limited oleic acid feed)의 도움으로 유도(induction)에 의한 생산물의 형성.

[0355] 일반적으로, 배치 단계(batch phase)는 약 20 시간(h)이고, 생산 단계는 대략 72 시간이다. 공정의 끝에서, 배양액(culture broth)을 원심분리하고, 상청액을 수집하였다. 상청액을 GAA 의 정제를 위한 출발 물질(starting material)로서 사용하였다(실시예 12를 참고하라).

[0357] 하기의 파라미터를 발효 동안에 조절하였다. 에어레이션(Aeration)을 1.5 vvm air[분 당 부피당 부피(volume per volume per minute)]의 상수값(constant value)에서 유지시켰다. 용해된 산소(Dissolved oxygen, DO)를 30 % 로 초기에 유지시켰다. 교반(stirring)을 DO 수준에 따라 600 내지 1200 rpm 으로 증가시켰다. 최대 1200 rpm 에 도달하였을 때, 이러한 속도를 일정하게 유지시켰고, DO-세트포인트(setpoint)를 10 % 로 설정하였다. 10% DO 를 유지시키기 위해, 산소는 50 % 의 최대 백분률로 반응기(reactor) 내에 섞었다(spiked). 거품 발생(Foam evolution)을 폼 프로브(foam probe)로 관리하였다. 거품 검출(foam detection)의 경우에, 거품 억제제를 바이오리액터(bioreactor)에 첨가하였다. pH 는, pH 6.8 의 상수값을 유지시키기 위해 14% (v/v) 암모니아 (염기) 또는 10 % 인산(phosphoric acid)을 첨가하여 조절하였다. 온도를 전체 공정을 거쳐 28 ℃로 일정하게 유지시켰다.

[0359] 바이오매스를 600 nm(OD600)의 광밀도(optical density) 측정에 의해 모니터하였다. 샘플을, 분광 광도계(spectrophotometer)의 직선 범위(linear range)에서의 수치를 수득하기 위해, 2-1000 배로 희석시켰다. 생산물 형성을 웨스턴 블롯 분석 및 특정한 효소 활성도 테스트로 측정하였다.

[0361] **[실시예 12]**

[0362] **재조합 huGAA(rhGAA)의 정제**

[0363] 배양 후의 상청액(실시예 11 을 참고하라)을 심층 여과(depth filtration)를 통해 명확하게 하였다. 결과적으로 생성된 물질을 TFF 를 통해 20 번 농축시킨 다음에, 10kDa MWCO 막(Millipore)에서의 20 mM 인산나트륨 pH 6 및 100 mM NaCl에 대항하여 초여과되었다(diafiltere).

[0365] rhGAA 의 정제를, 1 M 의 농도까지 황산암모늄(ammonium sulphate)을 첨가하면서 시작하였다. 원심분리 후에,

상청액을 XK16/40 컬럼을 가진 Toyopearl-Phenyl 650M (Tosoh Biosciences)에 로딩시켰다(loaded). 1 에서 0 M 까지의 황산 암모늄의 선형 구배(linear gradient)를 용리액(elution)에 적용시켰다. rhGAA 를 포함하는 이러한 분획물을 모은 다음에, 10 mM BIS-TRIS pH 6 내로 완충용액 교환(buffer exchange)을 하였다. 0 에서 1 M 까지의 NaCl 의 선형 염 구배(linear salt gradient)를 사용한 소스 30Q packed Tricorn 10/50 또는 XK25/20 컬럼(GE Healthcare) 상에서 음이온 교환 크로마토그래피를 통해 추가적인 정제를 수행하였다. 그리고 난 다음에 결과적으로 생성된 GAA-함유 분획물을, 200 mM NaCl 및 50 mM 염화 나트륨 pH 6 으로 예비-평형시킨(pre-equilibrated), 최종의 Hiload 16/60 superdex 200 겔 여과 컬럼(GE Healthcare) 상에 로딩시키기 전에, 농축시켰다. 분획물을, 특정한 활성도를 기초로 선택하였고, 쿠마씨-염색된 SDS-PAGE 겔에서 정제시킨 다음에, 5 내지 10 mg/ml 의 최종 농도로 농축시켰다. 단백질 농도를, 10 kDa 의 MWCO 와 함께 15 ml 의 Amicon Ultra centrifugal devices (Millipore)에서 측정하였다.

[0367] rhGAA에 대해 질적인 스크리닝(qualitative screening)을 위한 반응을, 10 또는 5 ㎕의 용리 분획물(elution fraction)에 대한 10:1 또는 20:1의 부피 비율로, 0.35 mM 4-MUG, 0.1% BSA 및 100 mM 아세트산 나트륨 pH 4 으로 이루어진 반응 완충용액을 첨가하여 시작하였다. 모든 반응을 96-웰 플랫-보텀 마이크로타이터 플레이트(flat-bottom microtiter plates)에서 실시하였다. 37 ℃에서 30 분 내지 1 시간의 배양기간(incubation period) 후에, 동등한 부피의 100 mM 글리신 pH 11을 반응을 정지시키기 위해 첨가하였고, 형광성 반응 생산물 4-메틸움벨리페론(fluorogenic reaction product 4-methylumbelliferone)의 방출을 UV-선 하에서 관찰하였다. 특정한 활성도(Specific activities)(units/mg 단백질)를, 황색의 p-니트로페놀레이트 반응 생산물의 효소적 방출을 측정하는, 합성 기질(Specific activities) p-니트로페닐-α-D-글루코피라노시드(p-nitrophenyl-α-D-glucopyranoside, PNPG)와 함께 비색 검정(colorimetric assay)을 사용하여 측정하였다. 반응물을, 마이크로타이터플레이트(microtiterplate)의 반응 웰(reaction wells)에서 10 ㎕ 의 효소 용액 및 90 ㎕ 의 기질 반응 완충용액(substrate reaction buffer)[150mM 시트레이트-포스페이트 완충용액 pH 4 에서의 2 mM PNPG, 1% BSA]을 혼합하여 시작하였고, 그 후에 37 ℃에서 배양시켰다. 1 내지 2 시간 후에, 동일한 부피의 정지 완충용액(stop buffer), 10 % 탄산나트륨 pH 12 를, 반응물을 퀀칭시키고, 이의 이온화된 상태에서의 방출된 p-니트로페놀(PNP)를 가져오기 위해(bring) 첨가하였다. 백그라운드-수정된 흡광도(Background-corrected absorbances) 및 p-니트로페놀레이트 표준(p-nitrophenolate standards)을 405 nm 의 파장에서 측정하였고, 특정한 활성도를 계산하였다, 단백질 농도를 BCA 방법(bicinchoninic acid method)으로 측정하였다. 하나의 유닛(unit)은, 시트레이트-포스페이트 완충용액 pH 4.0 에서의 2 mM 의 최종의 기질 농도에서의 37 ℃에서의 min 당 1 nmol의 PNPG 를 1 nmol PNP 및 D-글루코스로의 전환(conversion)을 촉매하는 효소의 양을 나타낸다.

[0369] **[실시예 13]**

[0370] **보다 높은 정도의 인산화된 N-글리칸을 갖는 *Y. 리폴리티카*에서 발현된 당단**백질에서의 이종 발현된 CcMan5 의 포스페이트 캡핑제거 활성도(Phosphate uncapping activity)

[0371] huGAA는, 보다 높은 정도의 인산화된 N-글리칸 구조를 갖는 당단백질을 수득하기 위해, *Y.리폴리티카* 스트레인 OXYY1589 에서 발현되었다(실시예 10 을 참고하라). huGAA 를 실시예 12에 기재된 바와 같이 정제하였다.

[0373] CcMan5 (70 ㎍/ml의 농도로 각각 1 및 5 ㎕)를, 2 mM $CaCl_2$ 와 함께 100 mM HEPES 완충용액에서의 4 ㎍ huGAA 의 용액에 첨가하였다. 20 ㎕ 의 반응 혼합물을 실온으로 밤새 배양하였다. Laroy W. *et al.*, *Nature Protocols*, 1: 397-405 (2006)에 기재된 바와 같이, N-글리칸을 PNGaseF 와 함께 방출시키고, APTS 로 표지시킨 다음에, DSA-FACE 에서 분석하였다. CcMan5 처리의 전 및 후(before and after)에 N-글리칸 프로파일을 도 29 에 나타내었다. 정제된 huGAA 로부터 방출된 N-글리칸 혼합물은 주로 ManP-Man8GlcNAc$_2$ 및 (ManP)2-Man8GlcNAc$_2$ 로 이루어져 있다(도 29, 패널 B). ManP-Man8GlcNAc$_2$ 보다 더 빠르게 움직이는 피크를 ManP-Man7GlcNAc$_2$에 배치되었다. Man$_8$GlcNAc$_2$ 및 Man$_7$GlcNAc$_2$ 의 매우 최소한의 양만이 나타났다. CcMan5와 함께 huGAA 의 배양 후에, 각각의 ManP-Man8GlcNAc$_2$ 및 (ManP)2-Man8GlcNAc의 P-Man8GlcNAc$_2$ 및 P2-Man8GlcNAc$_2$ 의 전환이 관찰되었다(도 29, 패널 C 및 D). P-Man8GlcNAc$_2$ 및 P2-Man8GlcNAc$_2$ 사이의 러닝(running) 일렉트로페로그램에서의 피크는, 포스포디에스테르- 및 포스포모노에스테르- 결합(phosphodiester- and a phosphomonoester-linkage)을 갖는 부분적으로 캡핑되지 않은 비-인산화된(bi-phosphorylated) (ManP)2-Man8GlcNAc$_2$ 와 일치한다((MP)-M8-P, 도 29 에서의 패널 C 및 D). 이러한 생산물은, 보다 높은 농도의 CcMan5 또는 보다 긴 배양 시간을 사용하였을 때, 완전하게 캡핑이 제거된 P2-Man8GlcNAc$_2$ 로 추가적으로 가수분해되었

다.

[0375] 인산화된 N-글리칸 대 중성의(neutral) N-글리칸의 백분율을, DSA-FACE 일렉트로페로그램에서의 피크 영역을 측정으로부터 평가되었다(도 29). 곡선 아래의 영역과 관련된 도면은 CcMan5 처리 전(패널 B) 및 후(패널 D)에서의 huGAA 에 나타난 상이한 N-글리칸에 대해서 나타내었다. huGAA(패널 B)에서, (ManP)2-Man8GlcNAc$_2$ (11597), ManP-Man6GlcNAc$_2$ (1261), ManP-Man7GlcNAc$_2$ (5901), ManP-Man8GlcNAc$_2$ (15576), Man6GlcNAc$_2$ (680), Man7GlcNAc$_2$ (1716), Man8GlcNAc$_2$ (1572) 을 나타내었다. 재조합 huGAA에서 대략 90 % 의 N-글리칸을 구조(structures)를 포함하는 만노스-인산염(mannose-phosphate)으로 구성되었다.

[0377] 재조합 huGAA 와 CcMan5 (패널 D)의 밤새 처리한 후에, P2-Man8GlcNAc$_2$ (16182), (ManP)P-Man8GlcNAc$_2$ (1997), P-Man7GlcNAc$_2$ (8254), P-Man8GlcNAc$_2$ (17893), ManP-Man6GlcNAc$_2$ (500), ManP-Man7GlcNAc$_2$ (2495), ManP-Man8GlcNAc$_2$ (1326), Man6GlcNAc$_2$ (1097), Man7GlcNAc$_2$ (2143), Man8GlcNAc$_2$ (1599)를 나타내었다. huGAA 로부터 방출된 N-글리칸은 83 % 의 캡핑이 제거되지 않은 인산화된 구조로 구성되어 있고, 8 % 는 여전히 만노스-인산염 캡핑되어(mannose-phosphate capped) 있었고, 9 % 중성의 N-글리칸을 나타내었다. 캡핑이 제거된 인산화된 구조물의 백분율은, 높은 농도의 CcMan5 또는 보다 긴 배양 시간일 때 증가될 수 있다.

[0379] **[실시예 14]**

[0380] **캡핑 제거 활성도(Uncapping Activity)를 갖을 가능성이 있는 상동체(Homologs)의 확인**

[0381] 동일하게 예측된 촉매 부위 토폴로지(topology) 및 기능성(functionality)을 갖는 다른 GH92 패밀리 멤버(GH92 family members)를 확인하기 위해, 월드와이드 웹 cazy.org/GH92_all.html 으로부터 획득된, 큐레이트된(curated) GH92 패밀리 멤버를 분석하였고, NCBI 에서의 Non Redundant Protein Sequences database 에서의 CcMan5 도메인 서열로 Blastp 검색을 통해 수득된 상위 500개의 히트(hits)를 수득하였다. 그 다음에, 이러한 392개의 서열을, 다중 서열 정렬 패키지 MUSCLE(multiple sequence alignment package MUSCLE)(MUltiple Sequence Comparison by Log-Expectation)에 대한 입력(input)으로서 사용하였고, 이는 '계통발생론의(phylogenetic)' 거리(distance)[관련된 가장 가까운 것(closest related)부터 관련된 가장 먼 것(furthest related)까지]의 순으로 서열을 또한 평가하였다.

[0383] Cazy 데이터베이스로부터의 큐레이트된 GH92 패밀리 멤버를 기초로, 모든 GH92 단백질 서열(392) 및 CcMan5 도메인 서열의 MUSCLE 정렬을 CcMan5의 가장 가까운 상동으로서 하기를 확인하였다 :

[0384] *스트렙토마이세스 코엘리컬러(Streptomyces coelicolor)*CAA18915 (GenBank Accession No. NP_630514)

[0385] *클로스트리디움 스피로포르메(Clostridium spiroforme)*(GenBank Accession No. ZP_02866543)

[0386] *박테로이데스 테타이오타오미크론(Bacteroides thetaiotaomicron)* AAO78636 (GenBank Accession No. NP_812442)

[0387] *주농왕기아 프로푼다(Zunongwangia profunda)* ADF52306 (GenBank Accession No. YP_003584502)

[0388] *키티노파가 피넨시스(Chitinophaga pinensis)* ACU58463 (GenBank Accession No. YP_003120664).

[0390] 이러한 서열 및 그 다음의 5 개의 가장 가까운 상동체의 서열을 도 31 에 정렬하였다.

[0391]

[0392] 500 개의 기록한 blastp 단백질 히트 대 CcMan5 도메인의 MUSCLE 정렬을 기초로, CcMan5 의 가장 가까운 동족체(homologs)로 간주되는 것을 하기에 나타내었다:

[0393] *스트렙토마이세스 코엘리컬러(Streptomyces coelicolor)*(GenBank Accession No. NP_630514)

[0394] *스트렙토마이세스 리비단스(Streptomyces lividans)*(GenBank Accession No. ZP_05522540)

[0395] *스트렙토마이세스 리비단스(Streptomyces lividans)*(GenBank Accession No. ZP_06527366)

[0396] *페니바실러스 sp(Paenibacillus sp)*(GenBank Accession No. YP_003013376)

[0397] *박테로이데스 테타이오타오미크론(Bacteroides thetaiotaomicron)*(GenBank Accession No. NP_812442)

[0398] *박테로이데스 sp.(Bacteroides sp.)*(GenBank Accession No. ZP_04848482)

[0399] *박테로이데스 셀룰로실티구스(Bacteroides cellulosilyticus)*(GenBank Accession No. ZP_03677957)

[0400] *주농왕기아 프로푼다(Zunongwangia profunda)*(GenBank Accession No. YP_003584502)

[0401] *레이우엔훅키엘라 블란덴시스(Leeuwenhoekiella blandensis)*(GenBank Accession No. ZP_01061975)

[0402] *스핑고박테리움 스피리티보룸(Sphingobacterium spiritivorum)*(GenBank Accession No. ZP_07083984)

[0403] *키티노파가 피넨시스(Chitinophaga pinensis)*(GenBank Accession No. YP_003120664)

[0404] *페도박터 sp.(Pedobacter sp.)*(GenBank Accession No. ZP_01885202)

[0405] *클로스트리디움 스피로포르메(Clostridium spiroforme)*(GenBank Accession No. ZP_02866543).

[0407] 이러한 것들 및 5 개의 다음의-베스트 동족체의 정렬을 도 32 에서 발견할 수 있다. 주석을 단(annotated)GH92 데이터베이스로부터의 모든 5 개의 베스트 히트(best hits)는 전제적인 서열 데이터베이스에서의 Blast 검색으로부터의 이러한 13 개의 베스트 히트를 또한 발견하였다.

[0409] 도 31 에 최고의 5 개의 히트 및 도 32 에 최고의 13 개의 히트는 하기의 세 개의 모티프를 유일하게 공유하고, Zhu *et. al.*, 2010, *supra* 에 보고된 알파-1,2-만노시다제 GH92 패밀리 멤버의 구조와 상이한 실시예 8의 결정 구조(crystal structure)에서 나타내었다.

[0411] 1) 글리신-리치 모티프(glycine-rich motif) GVGxxGxGG, 각각의 X 는 G, S, T, V, A, C 또는 Q (작은 곁사슬)이고, CcMan5 도메인의 결정 구조 잔기의 넘버링: 69-77(with each X being G, S, T, V, A, C or Q (small side chains), numbering of crystal structure residues of CcMan5 domain: 69-77). 이러한 영역은, 효소의 활성 부위에서 -1 및 포스페이트--결합 서브사이트(phosphate--binding subsite)에 필수적인 수소 결합을 제공하는 루프(loop)를 형성한다.

[0413] 2) VRxE 모티프. R 은 -1 고리 및 가능한 +1 고리에 수소 결합을 형성한다. E 는, 아미도 이러한 모티프를 형성하는, 이러한 R 잔기에 대한 염 다리(salt bridge)에 있다. x 는 가장 가까운 관련된 서브패밀리(the closest-related subfamily)[CcMan 5 에 대한 상위 3 의 상동체(top 3 homologs to CcMan5)]에서의 W 이거나, P 를 제외한 모든 20 개의 아미노산 중의 하나 일 수 있다. 이러한 모티프는 SEQ ID NO:50 의 잔기 404 내지 407 에서 발견되었다.

[0415] 3) LYQGT 모티프, 만노시다제에서 E 인 Q 를 포함하고 및 이는 +1 사이트 형성에 중요한 Y535 를 포함한다[containing the Q which is an E in the mannosidases (proton donor), and which contains Y535, which is important for the +1 site formation]. 몇몇의 서열에서, L은 A 또는 Y 이고, 이는 또는 I, V, A, F 또는 M 일 것으로 합리적으로 예측될 수 있고, 이들 중의 몇몇에서, T 는 N 이고, S를 용인할(tolerate) 것으로 예측될 수 있다. 두 개의 *카울로박터(Caulobacter)* 서열은 Q 대신에 E 를 갖고, 따라서 이는 인산화된 글리칸에 작용하지 않을 것으로 예측될 것이다(would thus be predicted not to work on phosphorylated glycans).

[0417] 4) GDXGN 모티프. D 및 N 은 기질 결합 캐비티(substrate binding cavity)의 일부를 형성하고, + 만노스레 결합하기 위한 대체가능한 서브포켓(alternative subpocket)을 형성할 수 있다. X 는 P 외의 어떠한 아미노산일 수 있다. 이러한 모티프는 SEQ ID NO:50 의 잔기 21 내지 25 에서 발견되었다.

[0419] 구조(structure)를 기초로 한 상기의 생물정보학 워크플로우(bioinformatics workflow) 및 모티프 검색(motif search)에 기초해, CcMan 5 에 대한 동일한 기질 특이성, 즉 Man-6-Pi-Man 구조의 캡핑을 제거할 수 있는 특이성을 갖는 좋은 후보물질인 이러한 희귀한 패밀리 멤버(those rare family members)를 위한, 비-중복 단백질 서열 데이터베이스(non-redundant proteins sequence database)(현재 1220 서열 이상을 포함)에 나타낸 GH92 서열을 걸러내는(filter) 것이 가능하다. 특히, *스트렙토마이세스 코엘리컬러(Streptomyces coelicolor),* 및 *스트렙토마이세스 리비단스(Streptomyces lividans)*으로부터의 3 개의 서열은, 상기의 모티프 뿐만 아니라 구조의 많은 루프(loops)에서 CcMan5 와 유사하다.

[0421] CcMan5 및 이의 가장 가까운 상동체(homologs)의 독특한 서열 요소를 기초로 한 Hidden Markov Models 을 갖는 검색은, 이러한 모든 요소를 포함하는 더 이상의 추가적인 GH92 서열을 나타내지 않았고, 이는 이러한 GH92 멤버가 수렴 진화(convergent evolution)를 통해 이러한 요소를 획득할 수 없음을 강하게 나타낸다[이들은, 다중 서열 정렬에서 상위-나열되지 않은 것들이다(these are the ones that would not be top-ranked in multiple-

sequence alignments)].

[0423] **[실시예 15]**

[0424] *박테로이데스 테타이오타오미크론(Bacteroides thetaiotaomicron)***으로부터의** GH92 글리코시다아제에서의 인산염 캡핑 제거 활성도(phosphate uncapping activity)의 존재

[0425] *박테로이데스 테타이오타오미크론*으로부터의의 23 의 패밀리 GH92 α-만노시다제의 효소적 분석(enzymatic analysis)은 Zhu, Y. *et al*, 2010, *supra* 에 보고되어 있다. 몇몇의 변이체(variants)가 매우 낮은 활성도를 나타낼 지라도, α1,2-, α1,4-, α1,3- 또는 α1,6-만노시다제 활성도를 갖는 효소는 효소의 이러한 그룹에서 나타난다. 두 개의 α-1,2 만노시다제(Bt3990 및 Bt2199)의 삼차원 구조(three-dimensional structure)는 중요한 아미노산 잔기의 확인을 가능하게 하고, 이는 α-1,2 만노시다제 활성도에 대한 중요한 모티프, 즉 Bt3990 에서의 His584-Glu585 및 Trp99 인 것처럼 보인다. B. 테타이오타오미크론으로부터의 세 개의 GH92 효소, Bt3530 (Genbank nr AAO78636.1), Bt3965 (Genbank nr AAO79070.1) 및 Bt3994 (Genbank nr AAO79099.1)의 인산화된 N-글리칸(실시예 1 에 나타낸 MNN4 당)에서의 활성도를 테스트하였다. 이러한 효소는 낮은 α1,4-만노시다제 활성도를 나타내었고,His-Glu 및 Pro-Trp 모티프가 결여되어 있다.

[0427] Bt3530, Bt3965 및 Bt3994를 *E.coli* 에서 발현되었고, Zhu *et al*, 2010, *supra* 에 기재된 바와 같이 정제하였다. 샘플(0.1 mg/ml의 농도에서의 1 ㎕ 효소)을, 실온에서 오버나이트 검정(overnight assay)에서 2 mM $CaCl_2$ 와 함께 10 mM HEPES 완충용액 pH 7.0 에 용해시킨 7 ml APTS-표지된 MNN4 당과 함께 배양하였다. CcMan5 를 갖는 대조 검정(control assay)이 포함되었다. 말단 인산염(terminal phosphate)의 존재를 확인하기 위해, 반응 혼합물을 CIP 와 함께 배양하였다. $Man8GlcNAc_2$ (M8) 및 모노인산화된(monophosphorylated) ManP-$Man8GlcNAc_2$ (MP-M8)를 포함하는 N-글리칸 제조물(preparation)을 기질로서 사용하였다. Bt3530, Bt3965 및 Bt3994에 대한 어떠한 캡핑이 제거되는 활성도(uncapping activity)도 상기의 검정 조건 하에서 검출되지 않았다. 피크에서의 전기영동의 이동성(electrophoretic mobility)에서의 어떠한 시프트(shift)도 CcMan5 대조 반응(control reaction)[빠른 러닝(fast running) P-M8 피크의 출현(appearance)], 다음에 CIP 처리[P-M8 의 사라짐(disappearance)]와 비교하여 관찰되지 않았다.

[0429] 추가적인 실험에서,1 ㎕ 의 효소, 즉 각각의 Bt3530 (0.1 mg/ml), Bt3965 (4.75 mg/ml) 및 Bt3994 (1.37 mg/ml)를 실온에서 60 시간 동안 pH 7.0 (2 mM $CaCl_2$ 을 갖는 10 mM HEPES 완충용액 pH 7.0)및 pH 5.0 에서의 MNN4 N-글리칸(2 mM $CaCl_2$ 을 갖는 10 mM 아세트산 암모늄 pH 5.0)과 함께 배양하였다. 매우 작은 α1,2-만노시다제 활성도가, 작은 $Man5GlcNAc_2$ (M5) 피크가 일렉트로페로그램에서 나타남으로써, pH 7.0 에서 Bt3530 에서 관찰되었다. 반면에 pH 5.0 에서, 어떠한 α1,2-만노시다제 활성도도 나타나지 않았지만, 일렉트로페로그램의 좌측편(left hand side)에서 빠른 러닝 피크(fast running peak)가 나타났다. 이러한 피크는 P-$Man8GlcNAc_2$ (P-M8)로서 동일한 전기영동 이동성(electrophoretic mobility)을 갖고, 말단 인산염은 CIP 와 함께 배양한 후에 가수분해되었다. CcMan5 (Bt3530으로서 동일한 농도로 사용됨)는 pH 7.0 으로 실온에서 20 시간 배양하면서 완전하게 캡핑이 제거된 ManP-$Man8GlcNAc_2$이다; 따라서, Bt3530 의 관찰된 활성도는 상당히 낮다. 정제한 후에, Bt3530 샘플은, 300 mM NaCl와 함께 20 mM TRIS 완충용액, pH 8.0 에서 저장했을 때, 서서히 침전되었다. 따라서, Bt3530 단백질의 불안정성(instability)은 사용된 검정 조건 하에서 활성도에 영향을 줄 가능성이 있다. 40 배 더 높은 농도에서 사용된 Bt3965 는 pH 7.0 (패널 G 및 H) 및 pH 5.0 (패널 I 및 J)에서 Bt3530 와 동일한 결과를 나타내었다. 어떠한 활성도도 동일한 반응 조건 하에서 Bt3994 와 함께 관찰되지 않았다(패널 K 에서 N 까지).

[0431] 이러한 실험으로부터, 인산염 캡핑이 제거된 활성도는 MNN4 당에서 테스트된 세 개의 *B. 테타이오타오미크론* GH 92 효소 중의 2 개만이 최소한의 사이드 활성도(minor side activity)임을 결론으로 내릴 수 있다.

[0433] **그 밖의 실시형태**

[0434] 본 발명은 이의 상세한 설명과 함께 기재하긴 하였지면, 상기에 나타낸 상세한 설명은 본 발명을 설명하기 위한 의도이며, 이로 본 발명의 범위를 한정하지 않으며, 이는 청구의 범위의 범위에 의해 분명히 나타내었다. 그 밖의 측면, 장점 및 변형물은 하기의 청구의 범위의 범위 내에 있다.

[0436] 청구하는 사항

[0437] [청구사항 1]

[0438] 다음의 단계를 포함하는, 올리고당(oligosaccharide)에서의 만노스-6-인산 잔기(mannose-6-phosphate residue)의 캡핑제거(uncapping) 방법:

[0439] a) 만노스-1-포스포-6-만노스 잔기(mannose-1-phospho-6-mannose residue)를 갖는 상기 올리고당을 제공하는 단계; 및

[0440] b) 상기 만노스-1-포스포-6-만노스 잔기를 포스포-6-만노스로 가수분해시킬 수 있는 만노시다제(mannosidase)와 상기 올리고당을 접촉시키는 단계(여기에서, 상기 만노시다제는 글리코실 히드롤라제 패밀리 92(glycosyl hydrolase family 92)의 멤버(member)이다).

[0442] [청구사항 2]

[0443] 제1항에 있어서,

[0444] 상기 만노시다제는 SEQ ID NO:50의 잔기 1 내지 774 또는 SEQ ID NO:50와 적어도 70 % 동일성을 가지는 아미노산 서열을 포함하는 것인, 방법.

[0446] [청구사항 3]

[0447] 제1항에 있어서,

[0448] 상기 만노시다제는 아노머 산소에 양성자를 가할 수 있는 촉매산 잔기를 함유하지 않는 것인(the mannosidase does not contain a catalytic acid residue capable of protonating the anomeric oxygen), 방법.

[0450] [청구사항 4]

[0451] 제3항에 있어서,

[0452] *박테로이데스 테타이오타오미크론* 알파-1,2-만노시다제 Bt3990의 글루탐산 잔기 Glu 533에 대응하는 만노시다제의 잔기는 비-산성 잔기인 것인, 방법.

[0454] [청구사항 5]

[0455] 제4항에 있어서,

[0456] *박테로이데스 테타이오타오미크론* 알파-1,2-만노시다제 Bt3990의 글루탐산 잔기 Glu 533에 대응하는 만노시다제의 잔기는, 글루타민 또는 아스파라긴 잔기인 것인, 방법.

[0458] [청구사항 6]

[0459] 제5항에 있어서,

[0460] *박테로이데스 테타이오타오미크론* 알파-1,2-만노시다제 Bt3990의 글루탐산 잔기 Glu 533에 대응하는 만노시다제의 잔기는 글루타민 잔기이고, 상기 글루타민 잔기는 X_1YQGX_2 모티프에 함유되고, 이 식에서 X_1 은 Leu, Ile, Val, Ala, Phe, Tyr 또는 Met이고, X_2 는 Thr, Ser, 또는 Asn 인 것인, 방법.

[0462] [청구사항 7]

[0463] 제3항에 있어서,

[0464] *박테로이데스 테타이오타오미크론* 알파-1,2-만노시다제 Bt3990의 글루탐산 잔기 Glu 585에 대응하는 만노시다제의 잔기는 비-산성 잔기인 것인, 방법.

[0466] [청구사항 8]

[0467] 제1항에 있어서,

[0468] 상기 만노시다제는, SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 90 % 동일성(identity)을 갖는 아미노산 서열을 포함하는 방법.

[0470] [청구사항 9]

[0471] 제1항에 있어서,

[0472] 상기 만노시다제는, SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 95 % 동일성을 갖는 아미노산 서열을 포함하는 방법.

[0474] [청구사항 10]

[0475] 제1항에 있어서,

[0476] 상기 만노시다제는, SEQ ID NO: 50 의 잔기 1 내지 774 또는 SEQ ID NO: 50 과 적어도 98 % 동일성을 갖는 아미노산 서열을 포함하는 방법.

[0478] [청구사항 11]

[0479] 제1항에 있어서,

[0480] 상기 만노시다제는 SEQ ID NO:50 의 잔기 1 내지 774 또는 SEQ ID NO:50 을 포함하는 방법.

[0482] [청구사항 12]

[0483] 제11항에 있어서,

[0484] 상기 만노시다제는 SEQ ID NO:50 의 잔기 1 내지 774의 N-말단 또는 SEQ ID NO:50의 N-말단의 SEQ ID NO:15 의 잔기 1 내지 15 를 추가적으로 포함하는 방법.

[0486] [청구사항 13]

[0487] 제1항에 있어서,

[0488] 상기 만노시다제에 대해서, 아미노산 곁사슬(amino acid side chains)에서의 원자의 삼차원의 단백질 좌표(three dimensional protein coordinates)는, 도 33 에서의 등가의 원자(equivalent atoms)의 좌표의 1.5 Å 편차(deviation) 내의 범위에 들어가는 최소의 촉매 중심(minimal catalytic center)에 위치하는 방법.

[0490] [청구사항 14]

[0491] 제1항에 있어서,

[0492] 상기 만노시다제는 하기를 갖는 아미노산 서열을 포함하는 방법:

[0493] (i) GVGXXGXGG 모티프(motif)(이 식에서, X 는 Gly, Ala, Ser, Thr, 또는 Cys 이다);

[0494] (ii) VRXE 모티프(이 식에서, X 는 Pro 외의 모든 아미노산이다);

[0495] (iii) X_1YQGX_2 모티프(이 식에서, X_1 은 Leu, Ile, Val, Ala, Phe, Tyr 또는 Met 이고, X_2 은 Thr, Ser, 또는 Asn 이다); 또는

[0496] (iv) GDXGN(이 식에서, X 는 Pro 외의 모든 아미노산일 수 있다).

[0498] [청구사항 15]

[0499] 제1항 내지 제14항 중 어느 한 항에 있어서,

[0500] 상기 접촉 단계(contacting step)는 정제된 만노시다제(purified mannosidase), 재조합 만노시다제(recombinant mannosidase), 상기 재조합 만노시다제를 포함하는 세포 용해액(cell lysate), 또는 상기 재조합 만노시다제를 포함하는 균류 세포(fungal cell)를 사용하여 실행하는 방법.

[0502] [청구사항 16]

[0503] 제1항 내지 제14항 중 어느 한 항에 있어서,

[0504] 상기 올리고당은 단백질에 부착된(attached) 방법.

[0506] [청구사항 17]

[0507] 제16항에 있어서,

[0508] 상기 단백질은 균류 유기체(fungal organism)에서 발현된 인간 단백질(human protein)인 방법.

[0510] [청구사항 18]

[0511] 제17항에 있어서,

[0512] 상기 균류 유기체는 *야로위아 리폴리티카(Yarrowia lipolytica)* 또는 *아르술라 아데니니보란스(Arxula adeninivorans)*인 방법.

[0514] [청구사항 19]

[0515] 제17항에 있어서,

[0516] 상기 균류 유기체는 메탄올자화효모(methylotrophic yeast)인 방법.

[0518] [청구사항 20]

[0519] 제19항에 있어서,

[0520] 상기 메탄올자화효모는 *피치아 파스토리스(Pichia pastoris)*, *피치아 메탄올리카(Pichia methanolica)*, *오가태아 미뉴타(Oogataea minuta)* 또는 *한세눌라 폴리모르파(Hansenula polymorpha)*인 방법.

[0522] [청구사항 21]

[0523] 제17항에 있어서,

[0524] 상기 균류 유기체는 사상 균류(filamentous fungus)인 방법.

[0526] [청구사항 22]

[0527] 제21항에 있어서,

[0528] 상기 사상 균류는 *아스페르길루스 카에시엘루스(Aspergillus caesiellus)*, *아스페르길루스 칸디두스(Aspergillus candidus)*, *아스페르길루스 카르네우스(Aspergillus carneus)*, *아스페르길루스 클라바투수(Aspergillus clavatus)*, *아스페르길루스 데플렉투스(Aspergillus deflectus)*, *아스페르길루스 플래브스(Aspergillus flavus)*, *아스페르길루스 푸미가투스(Aspergillus fumigates)*, *아스페르길루스 글라우크스(Aspergillus glaucus)*, *아스페르길루스 니둘란스(Aspergillus nidulans)*, *아스페르길루스 니게르(Aspergillus niger)*, *아스페르길루스 오크라세우스(Aspergillus ochraceus)*, *아스페르길루스 오라자에(Aspergillus oryzae)*, *아스페르길루스 파라시티커스(Aspergillus parasiticus)*, *아스페르길루스 페니실로이데스(Aspergillus penicilloides)*, *아스페르길루스 레스트릭투스(Aspergillus restrictus)*, *아스페르길루스 소예(Aspergillus sojae)*, *아스페르길루스 시도위(Aspergillus sydowi)*, *아스페르길루스 타마리(Aspergillus tamari)*, *아스페르길루스 테레우스(Aspergillus terreus)*, *아스페르길루스 우스투스(Aspergillus ustus)*, 및 *아스페르길루스 베르시컬러(Aspergillus versicolor)*로 이루어진 군으로부터 선택된 방법.

[0530] [청구사항 23]

[0531] 제16항에 있어서,

[0532] 상기 단백질은 병원체 단백질(pathogen protein), 리소좀 단백질(lysosomal protein), 성장 인자(growth factor), 사이토카인(cytokine), 케모카인(chemokine), 항체 또는 이의 항원-결합 단편(antigen-binding fragment), 또는 융합 단백질(fusion protein)인 방법.

[0534] [청구사항 24]

[0535] 제23항에 있어서,

[0536] 상기 리소좀 단백질은 리소좀 효소(lysosomal enzyme)인 방법.

[0538] [청구사항 25]

[0539] 제24항에 있어서,

[0540] 상기 리소좀 효소는 리소좀 축적 질환(lysosomal storage disorder, LSD)과 관련된 방법.

[0542] [청구사항 26]

[0543] 제25항에 있어서,

[0544] 상기 LSD 는 파브리병(Fabry's disease), 점액다당류증 I(mucopolysaccharidosis I), 파버 질병(Farber

disease), 고세병(Gaucher disease), GM1-강글리오시드증(GM1-gangliosidosis), 테이-삭스병(Tay-Sachs disease), 샌드호프병(Sandhoff disease), GM2 활성제 질병(GM2 activator disease), 크라베병(Krabbe disease), 이염성백질이영양증(metachromatic leukodystrophy), 니만-피크병(Niemann-Pick disease), 샤이에 질병(Scheie disease), 헌터 질병(Hunter disease), 산필립포 질병(Sanfilippo disease), 모르키오병(Morquio disease), 마로토-라미 질병(Maroteaux-Lamy disease), 히알루로니다아제 결핍증 (hyaluronidase deficiency), 아스파르틸글루코스아민뇨증 (aspartylglucosaminuria), 푸코시드축적증(fucosidosis), 만노시도시스(mannosidosis), 쉰들러 병(Schindler disease), 사이알산축적증 유형 1(sialidosis type 1), 폼피병(Pompe disease), 피크노디소토시스(Pycnodysostosis), 세로이드 리포푸신증(ceroid lipofuscinosis), 콜레스테롤 에스테르 축적 질병(cholesterol ester storage disease), 월만병(Wolman disease), 다종 술파타아제 결손증(Multiple sulfatase deficiency), 갈락토시알리도시스(galactosialidosis), 뮤코리피드증(mucolipidosis), 시스틴축적증(cystinosis), 시알산 축적 질병(sialic acid storage disorder), 마리네스코-쉐글렌 증후군(Marinesco- Sjögren syndrome)을 갖는 킬로미크론 보유 질병(chylomicron retention disease), 헤르만스키-푸드락 증후군(Hermansky-Pudlak syndrome), 체디아크-히가시 증후군(Chediak-Higashi syndrome), 다논병(Danon disease), 또는 겔레오피직 이형성증(Geleophysic dysplasia)인 방법.

[0546] [청구사항 27]

[0547] 제25항에 있어서,

[0548] 상기 LSD 는 폼피병(Pompe disease) 또는 파브리병(Fabry's disease)인 방법.

[0550] [청구사항 28]

[0551] 제1항 내지 제14항 중 어느 한 항에 있어서,

[0552] 상기 만노시다제는 타겟팅 서열(targeting sequence)을 포함하는 방법.

[0554] [청구사항 29]

[0555] 다음의 단계를 포함하는, 캡핑이 제거된(uncapped) 포스포-6-만노스 잔기를 갖는 타겟 단백질(target protein)을 생산하는 방법:

[0556] 만노스-1-포스포-6-만노스 잔기를 포스포-6-만노스로 가수분해할 수 있는 만노시다제를 코드화하는(encoding) 핵산을 포함하고 발현되도록(express) 유전적으로 조작된 균류 세포를 제공하는 단계(여기에서, 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버이다); 및

[0557] 타겟 단백질을 코드화하는 핵산을 상기 세포 내로 도입하는(introducing) 단계(여기에서, 상기 세포는 상기 타겟 단백질을 생산하고, 타겟 단백질 상의 만노스-1-포스포-6-만노스 잔기는 세포에서 발현된 상기 발현된 만노시다제에 의해 포스포-6-만노스 잔기로 전환된다).

[0559] [청구사항 30]

[0560] 제29항에 있어서,

[0561] 상기 만노시다제는 제2항 내지 제14항 중 어느 한 항에 정의된 바와 같은 것인, 방법.

[0563] [청구사항 31]

[0564] 제29항에 있어서,

[0565] 상기 균류 세포는 만노실 인산화반응(mannosyl phosphorylation)을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하는 방법.

[0567] [청구사항 32]

[0568] 제29항에 있어서,

[0569] 상기 균류 세포는 OCH1 활성도(OCH1 activity)가 결핍되도록(deficient) 유전적으로 조작된 방법.

[0571] [청구사항 33]

[0572] 제29항에 있어서,

[0573] 상기 만노시다제는 *C. 셀룰란스(C. cellulans), 스트렙토마이세스 코엘리컬러(Streptomyces coelicolor),* 또는 *스트렙토마이세스 리비단스(Streptomyces lividans)* 만노시다제인 방법.

[0575] [청구사항 34]

[0576] 다음의 단계를 포함하는, 균류 유기체에서 캡핑이 제거된(uncapped) 포스포-6-만노스 잔기를 갖는 타겟 단백질을 생산하는 방법:

[0577] a) 만노스-1-포스포-6-만노스를 포스포-6-만노스로 가수분해할 수 있는 만노시다제를 코드화하는 핵산을 포함하고 발현되도록 유전적으로 조작된 균류 세포를 제공하는 단계(여기에서, 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버이고, 상기 균류 세포는 타겟 단백질을 코드화하는 핵산을 추가적으로 포함하고, 타겟(target) 상의 만노스-1-포스포-6-만노스 잔기는 세포에서 발현된 상기 발현된 만노시다제에 의해 포스포-6-만노스 잔기로 전환된다); 및

[0578] b) 상기 캡핑이 제거된 포스포-6-만노스 잔기를 갖는 상기 타겟 단백질을 분리하는 단계.

[0580] [청구사항 35]

[0581] 제34항에 있어서,

[0582] 상기 만노시다제는 제2항 내지 제14항 중 어느 한 항에 정의된 바와 같은 것인, 방법.

[0584] [청구사항 36]

[0585] 제34항에 있어서,

[0586] 상기 타겟 단백질 및 상기 만노시다제는 세포에 의해 분비되는 방법.

[0588] [청구사항 37]

[0589] 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 당단백질(glycoproteins)을 생산하도록 유전적으로 조작된 분리된 균류 세포(isolated fungal cell)로서, 상기 균류 세포는 만노시다제를 코드화하는 핵산을 포함하고, 상기 균류 세포에서의 상기 만노시다제의 발현(expression)은 상기 포스포-6-만노스 잔기를 포함하는 당단백질을 결과적으로 생산하고, 여기에서 상기 만노시다제는 글리코실 히드롤라제 패밀리 92 의 멤버인, 균류 세포.

[0591] [청구사항 38]

[0592] 제37항에 있어서,

[0593] 상기 만노시다제는 제2항 내지 제14항 중 어느 한 항에 정의된 것과 같은 것인, 균류 세포.

[0595] [청구사항 39]

[0596] 제37항에 있어서,

[0597] 상기 균류 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하는 균류 세포.

[0599] [청구사항 40]

[0600] 제37항에 있어서,

[0601] 상기 균류 세포는 OCH1[외부 사슬 연장(Outer CHain elongation)] 활성도가 결핍되도록 유전적으로 조작된 균류 세포.

[0603] [청구사항 41]

[0604] 제37항에 있어서,

[0605] 상기 균류 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하고, 여기에서 상기 균류 세포는 OCH1 활성도가 결핍되도록 유전적으로 조작된 균류 세포.

[0607] [청구사항 42]

[0608] 제37항에 있어서,

[0609] 상기 균류 세포는 당단백질인 타겟 단백질을 코드화하는 핵산을 추가적으로 포함하는 균류 세포.

[0611] [청구사항 43]

[0612] 제42항에 있어서,

[0613] 상기 타겟 단백질은 인간 단백질인 균류 세포.

[0615] [청구사항 44]

[0616] 제42항에 있어서,

[0617] 상기 타겟 단백질은 병원체 단백질(pathogen protein), 리소좀 단백질(lysosomal protein), 성장 인자(growth factor), 사이토카인(cytokine), 케모카인(chemokine), 항체 또는 이의 항원-결합 단편(antigen-binding fragment), 또는 융합 단백질(fusion protein)인 균류 세포.

[0619] [청구사항 45]

[0620] 제44항에 있어서,

[0621] 상기 리소좀 단백질은 리소좀 효소인 균류 세포.

[0623] [청구사항 46]

[0624] 제45항에 있어서,

[0625] 상기 리소좀 효소는 산성의 알파 글루코시다아제(acid alpha glucosidase) 또는 알파 갈락토시다아제(alpha galactosidase)인 균류 세포.

[0627] [청구사항 47]

[0628] 제42항에 있어서,

[0629] 상기 타겟 단백질은 리소좀 축적 질환(lysosomal storage disorder, LSD)과 관련된 단백질인 균류 세포.

[0631] [청구사항 48]

[0632] 제47항에 있어서,

[0633] 상기 LSD 는 파브리병(Fabry's disease), 점액다당류증 I(mucopolysaccharidosis I), 파버 질병(Farber disease), 고세병(Gaucher disease), GM1-강글리오시드증(GM1-gangliosidosis), 테이-삭스병(Tay-Sachs disease), 샌드호프병(Sandhoff disease), GM2 활성제 질병(GM2 activator disease), 크라베병(Krabbe disease), 이염성백질이영양증(metachromatic leukodystrophy), 니만-피크병(Niemann-Pick disease), 샤이에 질병(Scheie disease), 헌터 질병(Hunter disease), 산필립포 질병(Sanfilippo disease), 모르키오병(Morquio disease), 마로토-라미 질병(Maroteaux-Lamy disease), 히알루로니다아제 결핍증 (hyaluronidase deficiency), 아스파르틸글루코스아민뇨증 (aspartylglucosaminuria), 푸코시드축적증(fucosidosis), 만노시도시스(mannosidosis), 쉰들러 병(Schindler disease), 사이알산축적증 유형 1(sialidosis type 1), 폼피병(Pompe disease), 피크노디소토시스(Pycnodysostosis), 세로이드 리포푸신증(ceroid lipofuscinosis), 콜레스테롤 에스테르 축적 질병(cholesterol ester storage disease), 월만병(Wolman disease), 다종 술파타아제 결손증(Multiple sulfatase deficiency), 갈락토시알리도시스(galactosialidosis), 뮤코리피드증(mucolipidosis), 시스틴축적증(cystinosis), 시알산 축적 질병(sialic acid storage disorder), 마리네스코-쉐글렌 증후군(Marinesco- Sjögren syndrome)을 갖는 킬로미크론 보유 질병(chylomicron retention disease), 헤르만스키-푸드락 증후군(Hermansky-Pudlak syndrome), 체디아크-히가시 증후군(Chediak-Higashi syndrome), 다논병(Danon disease), 또는 겔레오피직 이형성증(Geleophysic dysplasia)인 균류 세포.

[0635] [청구사항 49]

[0636] 제37항에 있어서,

[0637] 상기 균류 세포(fungal cell)는 *야로위아 리폴리티카(Yarrowia lipolytica)* 또는 *아르술라 아데니니보란스(Arxula adeninivorans)* 세포인 균류 세포.

[0639] [청구사항 50]

[0640] 제39항에 있어서,

[0641] 만노실 인산화반응을 촉진시킬 수 있는 상기 폴리펩티드는 MNN4 폴리펩티드인 균류 세포.

[0643] [청구사항 51]

[0644] 제50항에 있어서,

[0645] 상기 MNN4 폴리펩티드는 *야로위아 리폴리티카, S.세레비시아(S. cerevisiae), 오가테아 미누타(Ogataea minuta), 피키아 파스토리스(Pichia pastoris)* 또는 *C. 알비칸스(C. albicans)* 폴리펩티드인 균류 세포.

[0647] [청구사항 52]

[0648] 제39항에 있어서,

[0649] 만노실 인산화반응을 촉진시킬 수 있는 상기 폴리펩티드는 *P. 파스토리스* PNO1 폴리펩티드인 균류 세포.

[0651] [청구사항 53]

[0652] 제37항에 있어서,

[0653] 상기 만노시다제는 *C. 셀룰란스(C. cellulans), 스트렙토마이세스 코엘리컬러(Streptomyces coelicolor),* 또는 *스트렙토마이세스 리비단스(Streptomyces lividans)* 만노시다제인 균류 세포.

[0655] [청구사항 54]

[0656] 제37항에 있어서,

[0657] 상기 만노시다제는 분비 신호(secretion signal)를 포함하는 균류 세포.

[0659] [청구사항 55]

[0660] 제37항에 있어서,

[0661] 상기 만노시다제는 세포내 구간(intracellular compartment)에 상기 만노시다제를 타겟하기 위한(target) 타겟팅 신호(targeting signal)를 포함하는 균류 세포.

[0663] [청구사항 56]

[0664] 제37항에 있어서,

[0665] 상기 만노시다제는 분비 신호 및 세포내 구간에 상기 만노시다제를 타겟하기 위한 타겟팅 신호를 포함하는 균류 세포.

[0667] [청구사항 57]

[0668] *야로위아 리폴리티카(Yarrowia lipolytica), 피키아 파스토리스(Pichia pastoris), 한세눌라 폴리모르파(Hansenula polymorpha), 아르술라 아데니니보란스(Arxula adeninivorans), 피치아 메탄올리카(Pichia methanolica), 오가테아 미뉴타(Oogataea minuta),* 또는 *아스페르길루스 니게르(Aspergillus niger)* 세포의 실질적인 순수 배양물(pure culture)로서, 이들 중 상당수는 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 당단백질을 생산하도록 유전적으로 조작되고, 상기 세포는 만노스-1-포스포-6-만노스를 포스포-6-만노스로 가수분해시킬 수 있는 만노시다제를 코드화하는 핵산을 포함하고, 여기에서 상기 만노시다제는 글리코실 히드롤라제 패밀리 92 의 멤버인, 배양물.

[0670] [청구사항 58]

[0671] 제57항에 있어서,

[0672] 상기 만노시다제는 제2항 내지 제14항 중 어느 한 항에 정의된 바와 같은 것인, 배양물.

[0674] [청구사항 59]

[0675] 제57항에 있어서,

[0676] 상기 세포는 만노실 인산화반응을 촉진시킬 수 있는 폴리펩티드를 코드화하는 핵산을 추가적으로 포함하는 배양물.

[0678] [청구사항 60]

[0679] 제57항에 있어서,

[0680] 상기 세포는 OCH1 활성도가 결핍되도록 유전적으로 조작된 배양물.

[0682] [청구사항 61]

[0683] 분리된 당단백질로서,

[0684] 상기 당단백질에서 적어도 47 %의 N-글리칸은 캡핑이 제거된 포스포-6-만노스 잔기를 포함하고, 상기 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 상기 적어도 47 % N-글리칸은, P_2-$Man_8GlcNAc_2$, (ManP)P-$Man_8GlcNAc_2$, P-$Man_7GlcNAc_2$, 및 P$Man_8GlcNAc_2$ 로부터 선택된 것인, 분리된 당단백질.

[0686] [청구사항 62]

[0687] 제61항에 있어서,

[0688] 상기 당단백질에서 적어도 50 %의 N-글리칸은, 캡핑이 제거된 포스포-6-만노스 잔기를 포함하고, 상기 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 상기 적어도 50 % N-글리칸은 P_2-$Man_8GlcNAc_2$, (ManP)P-$Man_8GlcNAc_2$, P-$Man_7GlcNAc_2$, 및 P$Man_8GlcNAc_2$로부터 선택된 것인, 분리된 당단백질.

[0690] [청구사항 63]

[0691] 제61항에 있어서,

[0692] 상기 당단백질에서 적어도 75 %의 N-글리칸은, 캡핑이 제거된 포스포-6-만노스 잔기를 포함하고, 상기 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 상기 적어도 75 % N-글리칸은 P_2-$Man_8GlcNAc_2$, (ManP)P-$Man_8GlcNAc_2$, P-$Man_7GlcNAc_2$, 및 P$Man_8GlcNAc_2$로부터 선택된 것인, 분리된 당단백질.

[0694] [청구사항 64]

[0695] 제61항에 있어서,

[0696] 상기 당단백질에서 적어도 90 %의 N-글리칸은, 캡핑이 제거된 포스포-6-만노스 잔기를 포함하고, 상기 캡핑이 제거된 포스포-6-만노스 잔기를 포함하는 상기 적어도 90 % N-글리칸은 P_2-$Man_8GlcNAc_2$, (ManP)P-$Man_8GlcNAc_2$, P-$Man_7GlcNAc_2$, 및 P$Man_8GlcNAc_2$로부터 선택된 것인, 분리된 당단백질.

[0698] [청구사항 65]

[0699] 만노스-1-포스포-6-만노스 잔기를 포스포-6-만노스로 가수분해할 수 있는 만노시다제 폴리펩티드로서,

[0700] 상기 만노시다제는 글리코실 히드롤라제 패밀리 92의 멤버인 것인, 만노시다제 폴리펩티드.

[0702] [청구사항 66]

[0703] 제65항에 있어서,

[0704] 상기 만노시다제는 제2항 내지 제14항 중 어느 한 항에 정의된 것과 같은 것인, 만노시다제 폴리펩티드.

[0706] [청구사항 67]

[0707] 하기를 포함하는 핵산:

[0708] (a) 제65항의 폴리펩티드를 코드화하는 뉴클레오티드 서열; 또는

[0709] (b) (a)에 기재된 바와 같은 뉴클레오티드 서열의 상보체(complement of the nucleotide sequence).

[0711] [청구사항 68]

[0712] 제67항에 있어서,

[0713] 상기 뉴클레오티드 서열은 SEQ ID NO:50의 잔기 1 내지 774 또는 SEQ ID NO:50 을 코드화하는 핵산.

[0715] [청구사항 69]

[0716] 제67항에 있어서,

[0717] 상기 뉴클레오티드 서열은 SEQ ID NO:14의 뉴클레오티드 46 내지 2322 또는 SEQ ID NO:14의 뉴클레오티드 46 내지 4995 를 포함하는 핵산.

[0719] [청구사항 70]

[0720] 제69항에 있어서,

[0721] SEQ ID NO:14의 뉴클레오티드 46 내지 2322의 5' 또는 SEQ ID NO:14의 뉴클레오티드 46 내지 4995의 5'의 SEQ ID NO:14의 뉴클레오티드 1 내지 45를 더 포함하는 핵산.

[0723] [청구사항 71]

[0724] 제67항 내지 제70항 중 어느 한 항의 핵산을 포함하는 발현 벡터.

[0726] [청구사항 72]

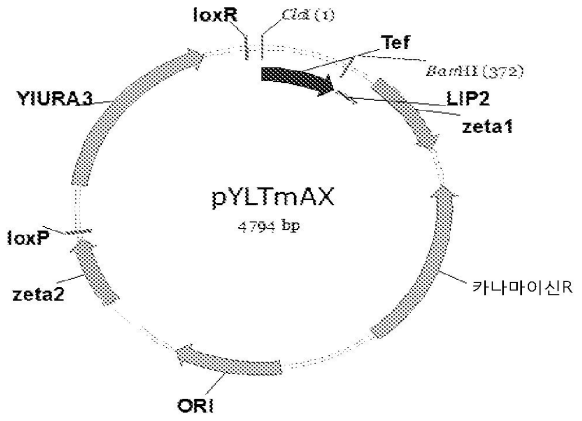
[0727] 제71항의 발현 벡터를 포함하는 숙주 세포.

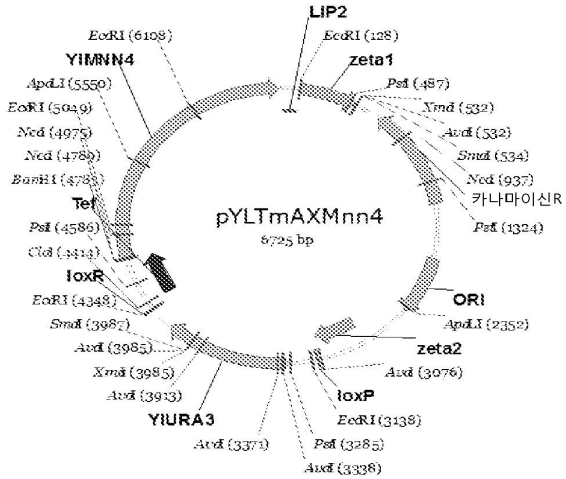
[0729] [청구사항 73]

[0730] 제72항의 숙주 세포를 배양하고, 상기 배양물로부터 폴리펩티드를 분리하는 것을 포함하는, 폴리펩티드를 제조하는 방법.

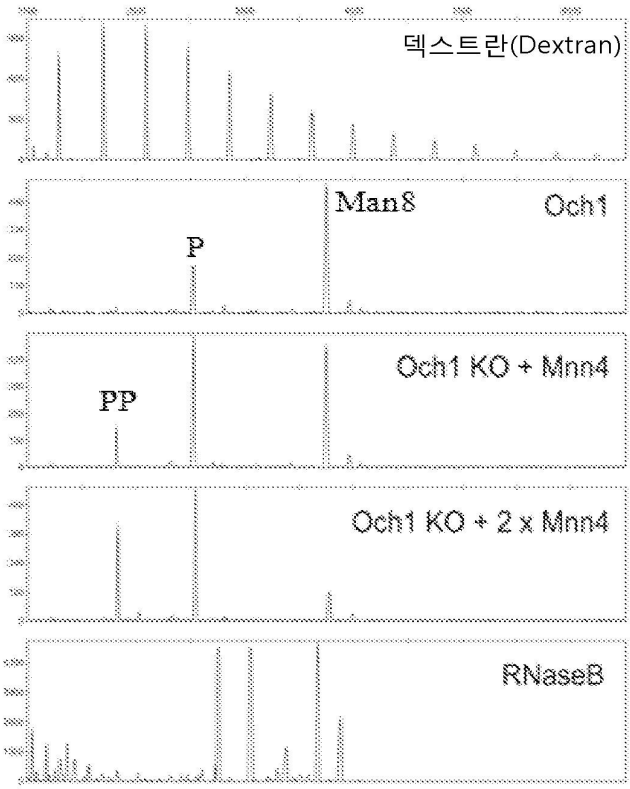
도면

도면1

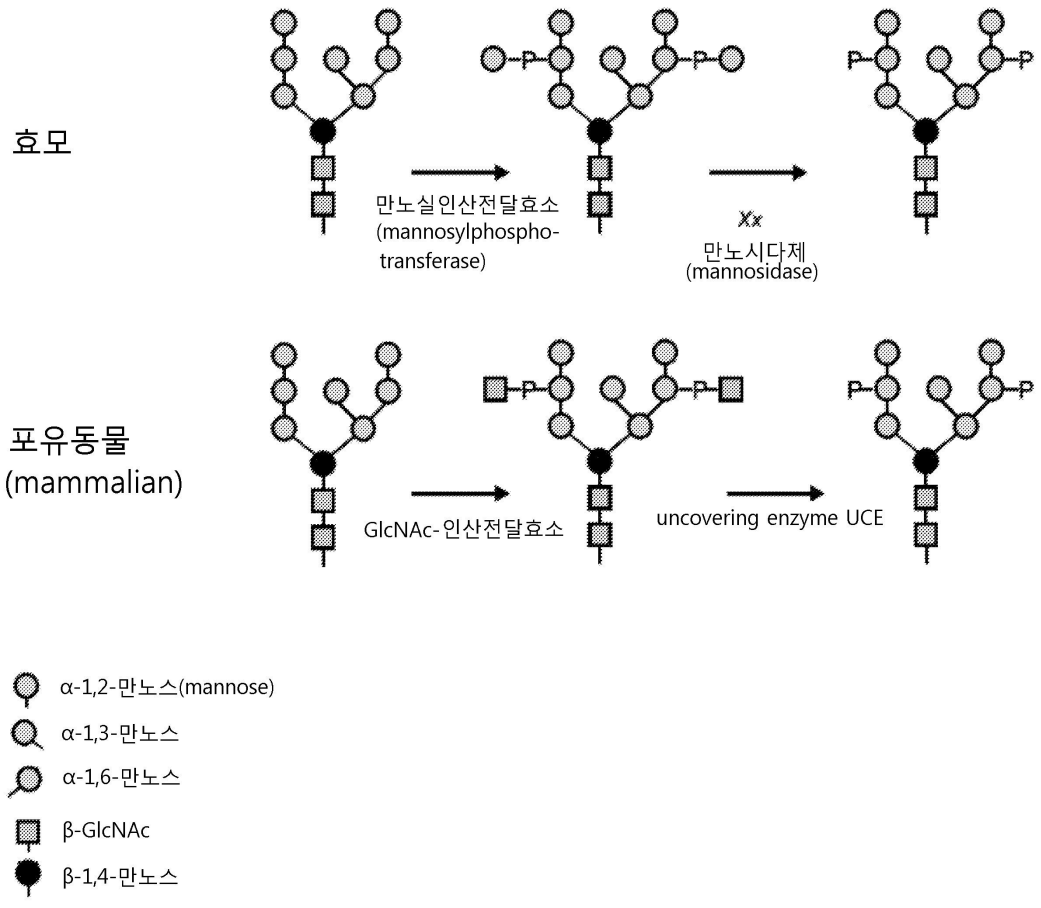




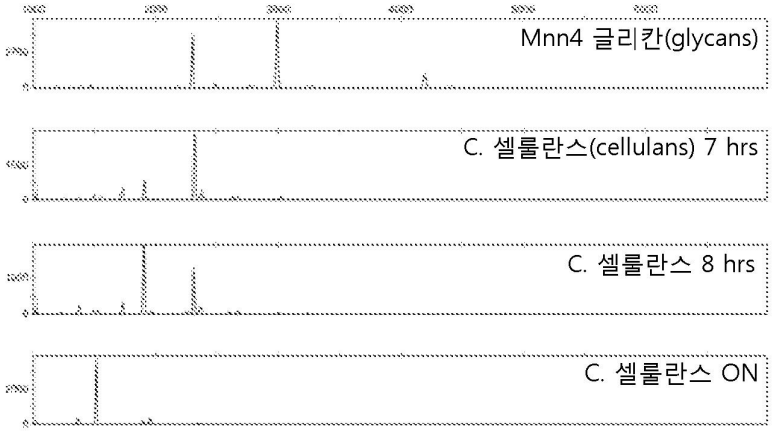
도면2


덱스트란(Dextran)
Man8
Och1
P
Och1 KO + Mnn4
PP
Och1 KO + 2 x Mnn4
RNaseB

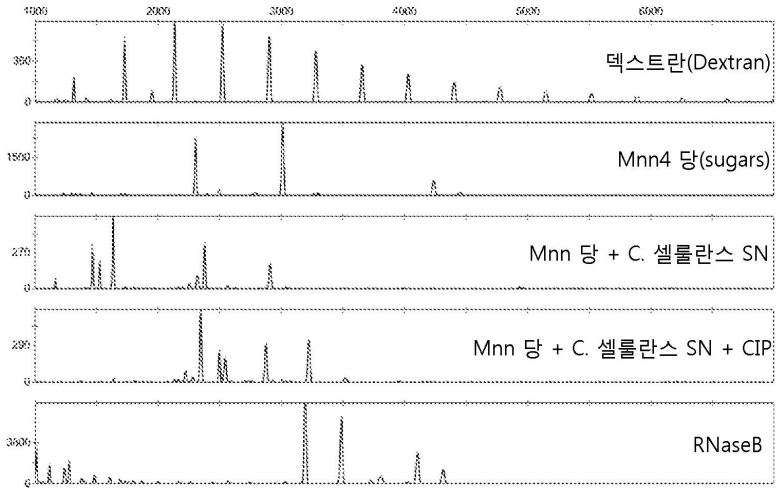
도면3


효모
만노실인산전달효소
(mannosylphospho-
transferase)
Xx
만노시다제
(mannosidase)
포유동물
(mammalian)
GlcNAc-인산전달효소
uncovering enzyme UCE
α-1,2-만노스(mannose)
α-1,3-만노스
α-1,6-만노스
β-GlcNAc
β-1,4-만노스

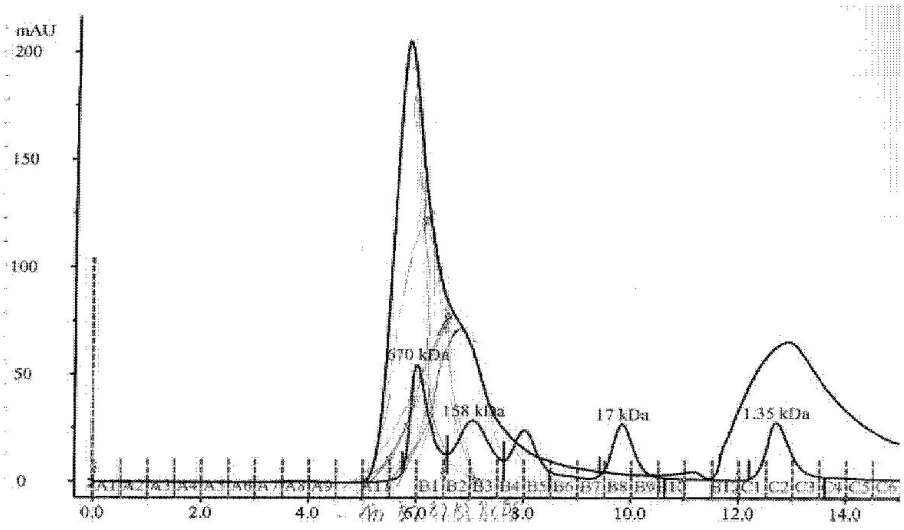
도면4


Mnn4 글리칸(glycans)
C. 셀룰란스(cellulans) 7 hrs
C. 셀룰란스 8 hrs
C. 셀룰란스 ON

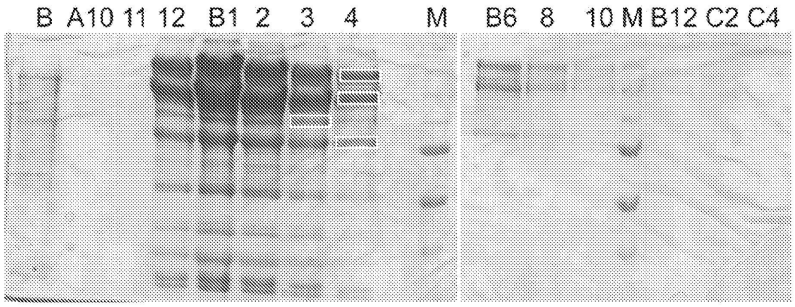
도면5


덱스트란(Dextran)
Mnn4 당(sugars)
Mnn 당 + C. 셀룰란스 SN
Mnn 당 + C. 셀룰란스 SN + CIP
RNaseB

도면6


mAU
200
150
100
50
0
670 kDa
158 kDa
17 kDa
1.35 kDa
0.0
2.0
4.0
6.0
8.0
10.0
12.0
14.0

도면7



도면8a

8A

ATGAGACGACCACGACTCGCCCTGCTCGCCGCGGGGCTCGCGCTCGCCGTCGCACCGGG
CACGCTGCTGCCCGTCGCCGCGGGCGCCGCCCCCGCCGACGAGGGCACCGTCACCGCCG
CCGCGGGCGACGACCTCACGCTCGAGGTCAACCCGTTCGTCGGCACCGAGAGCGAGGGC
AACGCCTACCCGGGCGCGACCGTGCCGTTCGGCATGGTCCAGCTCAGCCCGGACAACAC
GAACTCCTACGCCTCGACGTCGTACAGCACGAACGCGGGGCGCGTGTGGGGCTTCAGCC
ACCGGCACGTGAACAGCGCGGGCTGCCCCGCGGCGGGCGAGCTGCTCGTCACGCCGGAC
ACGAGCGCGACCCCGCGCACGTCGCGCTCCTTCATCGCCATCAAGGACCAGAAGAGCAC
CGAGCGCGCGTCGGCCGGGTTCTACGAGGTGACCCTCGCGAACGACGTGCACGCCGAGC
TCACCGCGACCACGCGCGTCGGCGCGCACCGCTACACGTTCCCCGCCTCGACGACGTCG
CACCTGTCGTTCAACGTGGGCCAGACCCTGCGCGACGCCGGCGCGAGCTCGGTGACGTG
GGTCGACGACCGCACGCTCGAGGGCTGGGTCGACAACGGCGGCTTCTGCGGCGGCACGC
CGGACAAGCAGCGGTACTTCTTCAGCGCGACGTTCGACCGCCCGGTCGCGTCGAGCGGC
ACGTGGGGGACCGATGCGCGCTACGTCGCGGGCTCCACGACGAGCGAGGTCGCGGGCGG
CAACAACGGCGCCGTCGCGGTGTTCGACACCACGACCGACCGCGACGTCGAGGTGAGCG
TGGGCGTGTCCTTCGTGAGCGTCGACGGCGCGCGCGCCAACCGCGAGGCCGAGGCCACC
GACGAGGGCGGGCAGGTCGCGTTCGACACCGTGCGTGAGGAGGCCCGCGACGCGTGGAA
CGCGGAGCTGGGCCGTGCCGCGATCGACGCGTCGCCCGACCAGCGCCGGATCTTCTACA
CCCAGCTCTACAAGACGCTGCTGTCCCCGACGATCGGCAGCGACGTCGACGGCCGGTAC
CGCGGCATGGACCTCGAGGTCCACCAGGCCGACGGCTGGGACTACTACCAGAACTTCTC
GCTCTGGGACACGTACCGCACGCAGGCGACGCTGCACGCCCTCCTGCTGCCCGAGCGCG
CGCAGGACATCGTGCGCTCGATGTACCAGCACCGCGTCGAGGGCGGCTGGCTGCCGCGC
TGGTCCCTCGGTGCACTGGAGACCAACATCATGGCGGGCGACCCCGTCACGCCGTGGCT
CGCGGAGAACTTCGCGCTCGGCACCGTCCCCGACGACATCGCGGACGAGCTGTGGGACT
ACCTCGTCGAGAACGCCACGACGACCCCGCCGGACGACGTCGCGTCCGTCGGGCGGCGC
AGCACCGAGTTCTACGCCGAGCACGGCCACGTGCCGTTCTACCCCGAGAACGAGGGCGG
CCTCGGCGGCCAGTTCGAGGAGTACCGCCACGGCGGCTCGGCGACGCTCGAGCTCGCGC
TCGCCGACGCGAGCCTCGGCGCTGCGGCCGAGCGCACGGGTCGCGAGGGCGGCCAGGCG
TTCCTCGACAAGGGTCGCAACTGGCGCAACCTCTGGAACCCGGACGTCGAGCTCTCGGG
TGGCTTCCAGGGCATGGTCAACGCGAAGCGCCCGACGGGCGAGTTCGTCACGCTGCCCG
AGCTGACGGACGTCACGCGCTCCGGCTTCCACGAGGGCGTGCCGTGGCAGTACCAGTGG
ATGGTGCCGCAGGACGTCACGGGCCTCCAGGAGGTCATGGGCGGCGAGGACGGCTTCGT
CGAGCGTCTCGACTACTACTTCGACCAGCCGGCGCTCGCCGCGAACCCCGGCGTCTCGC
CGAGCACGTGGGCCAAGGGCGGCAGCTCGTACTACACGACCATCCGCTACAACCCGGGC
AACGAGCCGACGATCATGAACGCGTGGCTCTACGGCTACGTGGGCCAGCCGTGGAAGAC
GAACGACGTCCTCGCCGCGAACCTCAACCGCTTCCCGGACACCCCGGGCGGCGGCGTCG
GGAACGACGACCTCGGCACGCTTGCCGCCTGGTACGTCATGGCGTCGCTCGGGTTCGAG
CCCGTCATGCCGGGCTCGGGGATCCTCGCGCTCAACGCGCCGAAGGTGCAGGCCGCGAC
GCTCACGACCGATGCCGGGGCGACGCTGCGCATCGACGCGGCGGGCGCGAACGAGAAGC
TCCCGAGCTACGTCGCCGGCCTGGAGGTCGACGGCGTCGCGCACACCGCCGCGTGGCTC
GACGTCGCGGCGCTGCAGGACGGCGGCACGCTCGACTTCGACCTGTCCGGCACGAGCGC
GGGCCTCACGTGGGGCACCGGCGCGGCCGACCGCATCCCGTCGGTCTCCGCCGTCGCCC
CGCCCGCGCCGGTCGAGGTCGAGGCGAGCGCGCGCTGCCTCGGCGGCCGGGCGTTCGTC
GCGGTCCGCGCGACCAGCACGGCCGACGCGCCGGTGGACGTGACTCTCACGACGCCGTT
CGGCGAGCGGACGGTCCGGCACGTGCAGCCGGGCAGGAGCGCCTACCAGTCGTTCACGA
CGCGCACGACGTCCGTCGAGGCCGGGACGGCGACCGTCACGGTCGTCGCCGCGGACGGC
ACGACGTCGACGGTCGACGCGGCGTACGAGGCGCTGGCCTGCGGC (SEQ ID NO:6)

도면8b

8B

MRRPRLALLAAGLALAVAPGTLLPVAAGAAPADEGTVTAAAGDDLTLEVNPFVGTESEG
NAYPGATVPFGMVQLSPDNTNSYASTSYSTNAGRVWGFSHRHVNSAGCPAAGELLVTPD
TSATPRTSRSFIAIKDQKSTERASAGFYEVTLANDVHAELTATTRVGAHRYTFPASTTS
HLSFNVGQTLRDAGASSVTWVDDRTLEGWVDNGGFCGGTPDKQRYFFSATFDRPVASSG
TWGTDARYVAGSTTSEVAGGNNGAVAVFDTTTDRDVEVSVGVSFVSVDGARANREAEAT
DEGGQVAFDTVREEARDAWNAELGRAAIDASPDQRRIFYTQLYKTLLSPTIGSDVDGRY
RGMDLEVHQADGWDYYQNFSLWDTYRTQATLHALLLPERAQDIVRSMYQHRVEGGWLPR
WSLGALETNIMAGDPVTPWLAENFALGTVPDDIADELWDYLVENATTTPPDDVASVGRR
STEFYAEHGHVPFYPENEGGLGGQFEEYRHGGSATLELALADASLGAAAERTGREGGQA
FLDKGRNWRNLWNPDVELSGGFQGMVNAKRPTGEFVTLPELTDVTRSGFHEGVPWQYQW
MVPQDVTGLQEVMGGEDGFVERLDYYFDQPALAANPGVSPSTWAKGGSSYYTTIRYNPG
NEPTIMNAWLYGYVGQPWKTNDVLAANLNRFPDTPGGGVGNDDLGTLAAWYVMASLGFE
PVMPGSGILALNAPKVQAATLTTDAGATLRIDAAGANEKLPSYVAGLEVDGVAHTAAWL
DVAALQDGGTLDFDLSGTSAGLTWGTGAADRIPSVSAVAPPAPVEVEASARCLGGRAFV
AVRATSTADAPVDVTLTTPFGERTVRHVQPGRSAYQSFTTRTTSVEAGTATVTVVAADG
TTSTVDAAYEALACG (SEQ ID NO:7)

도면9a

9A

GTGAGCCTCGCGCTCCCGCTGGCGGCGTACGCGGCGCCCGGGATCGGGGCGTCGCCCGC
GACCGCCGCCGGGACGGAGGCAGCGACGGGGTCCGATGCCGCCGCCGTCGACGGCCCGC
TGGTCGACTACGTCAACCCGTTCATCGGGACCAAGGACGACGGCAACACCTACCCGGGC
GCTGCCGTGCCGTTCGGCATGGTGCAACTCTCGCCGGACAACGGCCACAACGTCGGGTA
CGACTACGACCGCACGTCGGTGCGCGGGTTCTCGCTCGTGCACCTGTCCGGCGTCGGCT
GCGGCCTCGGCGGTCCGCTCCCGACCCTGCCGACGACGGGCGCGATCACCTCGACCGAC
TACGGCCAGTACGCGCTCGGTTTCTCGCACGACGACGAGGAGGCCTCGCCGGGGTACTA
CCGCGTGGGTCTCCAGGCGCCGGCGGGCACGATCGAGGCCGAGCTCACCGCGACCGAGC
GCACGGGCGTCCAGCGGTACACGTTCCCCGCGACGGCGCAGGCCAACGTCCTGCTCAAC
GCCGGCCAGGCGCTCAACCGGGTGACGGAGTCCGACGTGCGCGTCGTGGACGACCGCAC
GGTCGAGACGCGCATCACCGTCCGCGGCTTCTGCCAGGACACCGAGCCGCAGACGATCT
GGACCCGCACGACCTTCGACCGGCCGTTCGTCGCGCACGGCACGTGGGACGGCCAGGTC
GTCACCGCGGGCGCGGACGCCGCGTCCGGCGGCGAGGGCCGTCGCGGCGCGTACGTCAC
GTTCGACACGACCGGCGGCGACCTCGACGTCGAGGCCGTCACCGCGATGAGCTACGTGG
GCGCCGACGGCGCCGCGGCGAACCTCGCCGCGGAGGCCGGCACGTTCGACGCCGTGCAC
GACGCCGCGCGCTCGGCCTGGGAGGAGCGGCTCGGCCTCGTGCGGGTCGCGCAGGGCGA
CCCGGACGACCTGCGCACCTTCTACTCCTCGCTCTACCGCAGCTTCCTCGCGCCGAACG
TCGGCTCCGACGTCGACGGGCGCTACCGCGGCTGGGACCAGGAGGTCCACGCCGCGGAA
CCGGACTTCACCTACTACCAGAACTACTCGCTCTGGGACACGTACCGCACCCAGCAGCA
GCTCCTGTACCTGCTCGCGCCCGACGAGTCGGCCGACATGGCGCTCTCGCTCGTGCGCC
AGGGCCAGCAGGGCGGGTGGCTCCCGCGCTGGGGCTACGGCACGGTCGAGACGAACATC
ATGACCGGCGACCCGGCGACGCCGTTCCTCGTCAGCGCCTGGCGCCAGGGCCTGCTCGC
GGGCCACGAGGAGGAGGCGTACGCGGTCCTGAGGGAGAACGCCGACGGCGTCCCGCCCG
CCGACTCGCCCTTCAACGGGCGCGCGGCGAACGTCGAGTACCTGCGCGACGGGTTCGTC
CCGCACGAGCCGGCGCGCTCGGGCAAGCCCGGCGACTACGACCTCCAGCACGGCGCCTC
GGCGACCATGGAGTACGCCCTCGCCGACGCGATGCTCTCGACCATGGCGCGCGGCCTCG
GCCACGACGAGGACGCCGACCGGTACGCCGCCCGCGGCCAGAGCTACCGCAACGTGTTC
GACCCGCGCACGGGCAACTTCCGGGCGCGTAACGCGGACGGCTTCTTCGTGGGCGACGC
GGACCCCGCGCACTCCGACGGGTTCCACGAGGGCACGGCGGTGCAGTACCAGTGGCTCG
TGCCCCAGGACGTGCCGGGCCTGTTCGACCTCATGGGCGGCACCGACGCCGCGGTCGAC
CGCCTCGATGCGTTCTTCGCGTACGACGAGCTCGTCGCCGACCCCCCGCACGTCGCGAG
CGAGGTGTGGGTCAACGGCACGTACGACTACTACGGCTGGGAGACCTACAACCCGAACA
ACGAGCCCAACCTCCATGCGCCGTACGTCTACCTGTGGACCGGGCAGCCCTGGAAGACG
ACGGACGTCGTGCGCGCCGCGTCGACCCTCTTCACCGACGGCCCCGACGGCGTCACGGG
CAACGACGACCTCGGCACGATGTCCGCGTGGCACGTGCTGTCGTCGATCGGCGTGTACC
CGATCGTGCCGGGCGCCGATCTGTGGGGCCTGACGACGCCGCTCTTCGACGACGTGACG
ATCACGCTCGACCCGGAGGTCTTCGGTCGGGACTCCCTGCGCCTCACGGCGGACGGCGT
CGCGCCCGACACGCACTACACGCAGTCCGTGTCGCTCGGCGGCGAGCCGCTCGATCGCG
CCTGGGTCACGGGCGACGAGCTCACCGCGGCCGGCACGCTCGACGTGACCGTCGGCACC
GAGCCGTCCGCGTGGGCGACCGACCCCGCGGCCTCGCCGGGCGCCGTCGTGCCTGCGGA
CGGCACGGTCGAGCGCCTGTTCGTCGGCGCGACGCCGCGGCAGCCGGTCCTCGCCCCGG
GCGGGCGGACCGAGGTCGCAGTCCAGGTCGTCGCCCAGGGCGCGGGGACGTCCAGCGGG
ACGCTCGAGGTGACGTCCGACGGCGCGGTCACCGCGACGACCGACCTCGCCGAGTGGAC
CGCCGAGTCCGACGGCCTGCCGGCCACGGTCGAGGGCACGGTGACGATCGAGGCTCCCG
CCGACGCCGAGCCGGGTCTGCACACGGTGCGGCTCGTCGTGCGCGACGCCGCGGGGACC

도면9b

GAGGCGGTCCGCGAGGTCTCGGTCGTCGTGTCCGGGGAGTCGTGGATCGCCGACGCGTT
CGACAACGTCGGCATCGGCGACGCCGGGGCGGCCAACGCGAACCTCGACGGCTCGGGCG
CCTACCTCCTGCGCGACCTGCTCGCGGACCTCGGCGCCGTCCAGGGCCTGGAGCTCACC
GTGCCGGGCACGGACCTCACCTACACGCTCGGGGCCCCGCGGGCGGGCGCGCCCGACAA
CGTCGCCGCGAGCGGCGAGGTCCTCGAGGTGCCCGAGCACCTGCGCTCGGCCCGCCACC
TCTCGGTGGTCGGGACGAGCACGCACGGCACGCACGGGGGCGGCCTCGTCCTCGGGTTC
GCCGACGGCTCGTCGCAGACCGTCGACGTGCGCCTCAGCGACTGGTGCACGGGCTCGCC
CGAGCCCGGCAACATCACGGTCGCGAAGGCCGGGGCGCGCGGCGACCGCGAGAACGTGC
AGAAGATCGGCTGCGGCCTCTACGCCACCGCGCCCGTCGCGATCCCCGAGGGCAAGGTC
CTGACGTCGGTCACGCTGCCGAGCGACGAGCGGTTCCACGTGTTCGCGATCGCGACCGA
CGCGACGGGGGACGTCCCCGCGCCGCAGGTCGAGGTCACGGCGCAGGCCCGCTGCCTCG
GCGGCAAGGCGTTCGTCGCGGTGCGCGCGCTCAACACGGGCGAGCAGCCCGCCGCGATC
GAGCTCGCGACCCCGTACGGCTCCAAGCTCTTCGGTGACGTCGCTCCCGGGGCGAACGC
GTACCAGTCGTTCGCCACCCGCGCCGCCGCCGTCGAGGCGGGCGAGGTCACGGTGACCG
TGACGACGCCCGACGGCGAGCCCCAGCAGGTCACGGCCGCGTACGACGCCGCCGCCTGC
TCC (SEQ ID NO:8)

도면9c

9B
VSLALPLAAYAAPGIGASPATAAGTEAATGSDAAAVDGPLVDYVNPFIGTKDDGNTYPG
AAVPFGMVQLSPDNGHNVGYDYDRTSVRGFSLVHLSGVGCGLGGPLPTLPTTGAITSTD
YGQYALGFSHDDEEASPGYYRVGLQAPAGTIEAELTATERTGVQRYTFPATAQANVLLN
AGQALNRVTESDVRVVDDRTVETRITVRGFCQDTEPQTIWTRTTFDRPFVAHGTWDGQV
VTAGADAASGGEGRRGAYVTFDTTGGDLDVEAVTAMSYVGADGAAANLAAEAGTFDAVH
DAARSAWEERLGLVRVAQGDPDDLRTFYSSLYRSFLAPNVGSDVDGRYRGWDQEVHAAE
PDFTYYQNYSLWDTYRTQQQLLYLLAPDESADMALSLVRQGQQGGWLPRWGYGTVETNI
MTGDPATPFLVSAWRQGLLAGHEEEAYAVLRENADGVPPADSPFNGRAANVEYLRDGFV
PHEPARSGKPGDYDLQHGASATMEYALADAMLSTMARGLGHDEDADRYAARGQSYRNVF
DPRTGNFRARNADGFFVGDADPAHSDGFHEGTAVQYQWLVPQDVPGLFDLMGGTDAAVD
RLDAFFAYDELVADPPHVASEVWVNGTYDYYGWETYNPNNEPNLHAPYVYLWTGQPWKT
TDVVRAASTLFTDGPDGVTGNDDLGTMSAWHVLSSIGVYPIVPGADLWGLTTPLFDDVT
ITLDPEVFGRDSLRLTADGVAPDTHYTQSVSLGGEPLDRAWVTGDELTAAGTLDVTVGT
EPSAWATDPAASPGAVVPADGTVERLFVGATPRQPVLAPGGRTEVAVQVVAQGAGTSSG
TLEVTSDGAVTATTDLAEWTAESDGLPATVEGTVTIEAPADAEPGLHTVRLVVRDAAGT
EAVREVSVVVSGESWIADAFDNVGIGDAGAANANLDGSGAYLLRDLLADLGAVQGLELT
VPGTDLTYTLGAPRAGAPDNVAASGEVLEVPEHLRSARHLSVVGTSTHGTHGGGLVLGF
ADGSSQTVDVRLSDWCTGSPEPGNITVAKAGARGDRENVQKIGCGLYATAPVAIPEGKV
LTSVTLPSDERFHVFAIATDATGDVPAPQVEVTAQARCLGGKAFVAVRALNTGEQPAAI
ELATPYGSKLFGDVAPGANAYQSFATRAAAVEAGEVTVTVTTPDGEPQQVTAAYDAAAC
S

도면10a

10A

GTGCGGCGCTCCGTCGCGGCGCTCTCTGCCACGGCGGTCCTGGCCGCCGGACTCTCGAT
CGCGCCCGCCGTCGGGCTCGCGGTCCCGGCGGTCGCGGCCGCACCCGACCTCGTTGAGG
ACCCCGTCTCCTTCGTCGACCCGTTCGTCGGGACCGGCCAGGCGACGGGCGTCGTCGGG
GAGATCAACAACTTCCCCGGGCCGTCGATGCCGTTCGGCATGATGCAGCTCTCGCCCGA
CACCCAGGTCTCCGTGGGCAACGGCGACAAGGCGTACGCGGGCTACCGCTACTCGCACC
AGGCGATCCGCGGCTTCTCCATGACGCACGCGGCCGCCGGGTGCTGGATCTTCGGCGAC
GTCCCGATCCTCCCCGTGACGGGCGACGTCGGGCAGTACCCGTGGGACCGCAAGGAGGC
GTTCAGCCACGACGCGGAGAGCGCCGAGGTCGGCCGGTACGCGGTCACGCTCCAGTCGT
CGGGGATCGATGCGGAGGTGTCGGCCGCGACCCGCTCGGGCGGACTGACGTTCGACTAC
CCCGAGGGCGGTGCCGCGTCGCAGGTGATCGTCAACGCCGCGGGCTCGCTCGCGAGCGT
GCGCAACGCGACGGTCGAGGTCGAGGACGCGCGCACGGTCACCGGCTCGGTGACGAGCG
GCGGGTTCTGCGGCAAGAACAACACGCACACGACGTACTTCGCGATCGAGCTCGACCAG
GACGCGCAGGCGTTCGGCACGTGGCAGGGCTCGACCGTCTCGCCCGGCGACCCGTCGGC
CGACGGCAACGGCGCGGGCGCGTGGCTCACCTTCGCGCCCGGCGCGACGGTGCACGCGA
AGGTCGGCATGTCCTACGTGAGCGTCGAGGGCGCGCGCGCCAACCTCGCGGCCGAGATC
CCGGGCTTCGACTTCGACGCCGTCCGGGACGCCAACCGCGCCGCCTGGTCCGACCTGCT
CGGCAAGGTCCGCGTCGCGGGGCAGGACGCCGACGACCTCACCATGTTCTACACGTCGC
TCTACCACTCGCTGCTGCACCCGAACACGTTCACCGACGTGGACGGCCGGTACGTCGGG
TTCGACGGGGAGATCCACCAGGCCCCCGAGGGGCACGAGCGGTACGCGAACTTCTCCGA
CTGGGACACGTACCGGTCGCTCGGCGCGCTCCAGGCGCTGCTGGCGCCCGACCAGGCGT
CGGACATGGCGCAGTCGCTCGTCGAGGTCGCCGACCAGTCCGGCTGGCTGCCGCGCTGG
CCCGTCGCGAACCAGCACACGGGCCAGATGACCGGTGACTCCTCGGTGCCGCTCATCGC
GAGCATGTACGCGTTCGGGGCGCGCGACTTCGACGCGGAGTCGGCGCTCGCGCACATGG
TCAAGGGTGCGACGAGCGCCGCCCCGACCGCGAACGGCTACGTGCAGCGGCGCGGGATC
GAGACGTACCTCGAGCGCGGCTACGCGCCCCAGACCGAGGAGTTCCGGGGCGACCACCG
CGTCGTCGGCGCGTCGATCACGCTCGAGTGGTCGATCGCCGACTTCGCGATCGGGCAGC
TCGCGGCCGCGCTCGGCCAGGACGACGTCGCCACCGAGTACGCCGCCCGCGGCCAGTGG
TGGCAGAACGTCCACGACCCCGTGACCCGCACGGCGGGCGCCCGGAACGACGACGGCAC
GTTCGTGCGGTCGCAGGGCGGCGGCGGGTTCGGGCAGGAGGGCTTCGACGAGGGCAACG
CCGAGCAGTACACGTGGCTCGTGCCGCAGAACGTCGCGGGGCTCACCGACGCGCTCGGC
GGGCGTGAGGCCGTCGCGGAGCGGCTCGATGCCTTCACGGTGCAGCACAACGCCGGCCC
GAACGAGCCGTACCTGTGGATCGGCAACGAGCCGAACTTCGGCGTCCCGTGGCTGTACG
ACTACGTGGGCCAGCCGTGGCGGACGAGCGAGCTCGTGGACGAGCTTACGTCCACGCTG
TTCCGGCCCGAGCCGAACGGCAAGCCCGGCAACGACGACCTCGGCGCCCAGGCCGGCTG
GTACGTGTGGGCCGCGATGGGCCTGTACCCCACCACGCCGGGCACGGACGTGCTCGCGC
TCAACGCGCCGCGCTTCGACCGCGTCGTGGTCGACCTCGGCGAGGGCGACACCCTCGAC
CTGCGCGCCCCCGGCGCCTCGACCGGCGCCCGCTACATCAGCGGCGTCACCATCGACGG
CGCAGCCTGGGACGGGACCTCCCTGCCGCGCCACGTCGCGCACGACGGCGGCGTCGTCG
AGCTCGCGATGTCGACCGCACGCGACACGACGTGGGGGACCGCAGCCGAGGACGCCCCG
CCGTCGTGGCGCGACGGCGAGTCCGCCGTGGTCGCCGCCGCGGACCCGGGCCTCGTGAC
GGTCGCCCCCGGCGGGACCGCCGACGCGTCGGTGGCCGTGCAGCTCTTCGGCGCCGACG
CCGCCGACGTGCGCGTCGCGGTCGACGCGCCCGGGGGCATCGGGGTCGGTGAGCCCGCG
CTCGTCGACGACGGCTCGGGCCACCTCACCGGGACGGTCCCCGTCCAGGTGGGTGCCGG
CGTCGCGTCCGGCTACCACGACGCGCGCCTCGTGCTCTCGGCCGGGGACGACGACGTCG
AGGTGCCCCTCACCGTCCTCGTCGCCGCGCCCGGGTCGCTCGTTGCGGCCTACGACACG

도면10b

GTCGGCACTGCGCCCGAGGCGAACCGCGGCGTCGGGAACTTCGACGCGGCCGGCAACTC
GTTCTCGCGCGAGGCGCTCGCCGACGCGGGTCTCACGCCCGGGTCGGCGCACGACGTCG
ACGGCCTGGCGTTCACGTGGCCGTCCTCACCCGTGGGGCGCCCGGACTCGGTCACGCTC
ACCGGCGAGACCGTGCGGCTCGACGCGCCGACGAGCCGGCTCGCGTTCGTGGGCGCCGC
GACCGACGGGACCCATCGCGGGACCGCGGTCGTGACGTTCGACGACGGCAGCACCGCGA
CCACGACGATCGGCTTCGGCGACTGGGTGCTGCCGAGCGCGGACGGCTCGCCGGTCGAG
GGCAACTCGGTCGTCGCGCAGATGAACCGGCGCAACGGCGACAAGGACAGCGCGTTCGT
GTTCGCCACCGCCCCGTACACCGCGCCCGAGGACCGCCGCGTGGTCGCGGTGAGGTTCC
CCGACGTCGACGACCTGCACGTCTTTGCGATCGCGACCGAGCCGGCCGCGGACGTGCAC
CTCGTGGACGTGACGGTCTCCCTGCGCTGCCTCGCCGGGACCCCGTACGTGGCGGTGCG
CGCGGCGAACGTCTCCGCCGGGGCCGTCGACGTCGACCTCACGACGGGCGTGGGCTCGC
GGTCCTTCACGGCCGTCGCCCCCGGCGCCAACGCCTACCAGTCGTTCGCCGCCCGCGGC
GCGACCGGGAACGTCGACGTCACCGTCACGGCCACGGGGGAGGAGGGGACGCAGACGGT
CGCGCGGACCGTCGTCGTCCCGCGCTGCTCC (SEQ ID NO:10)

도면10c

10B
VRRSVAALSATAVLAAGLSIAPAVGLAVPAVAAAPDLVEDPVSFVDPFVGTGQATGVVG
EINNFPGPSMPFGMMQLSPDTQVSVGNGDKAYAGYRYSHQAIRGFSMTHAAAGCWIFGD
VPILPVTGDVGQYPWDRKEAFSHDAESAEVGRYAVTLQSSGIDAEVSAATRSGGLTFDY
PEGGAASQVIVNAAGSLASVRNATVEVEDARTVTGSVTSGGFCGKNNTHTTYFAIELDQ
DAQAFGTWQGSTVSPGDPSADGNGAGAWLTFAPGATVHAKVGMSYVSVEGARANLAAEI
PGFDFDAVRDANRAAWSDLLGKVRVAGQDADDLTMFYTSLYHSLLHPNTFTDVDGRYVG
FDGEIHQAPEGHERYANFSDWDTYRSLGALQALLAPDQASDMAQSLVEVADQSGWLPRW
PVANQHTGQMTGDSSVPLIASMYAFGARDFDAESALAHMVKGATSAAPTANGYVQRRGI
ETYLERGYAPQTEEFRGDHRVVGASITLEWSIADFAIGQLAAALGQDDVATEYAARGQW
WQNVHDPVTRTAGARNDDGTFVRSQGGGGFGQEGFDEGNAEQYTWLVPQNVAGLTDALG
GREAVAERLDAFTVQHNAGPNEPYLWIGNEPNFGVPWLYDYVGQPWRTSELVDELTSTL
FRPEPNGKPGNDDLGAQAGWYVWAAMGLYPTTPGTDVLALNAPRFDRVVVDLGEGDTLD
LRAPGASTGARYISGVTIDGAAWDGTSLPRHVAHDGGVVELAMSTARDTTWGTAAEDAP
PSWRDGESAVVAAADPGLVTVAPGGTADASVAVQLFGADAADVRVAVDAPGGIGVGEPA
LVDDGSGHLTGTVPVQVGAGVASGYHDARLVLSAGDDDVEVPLTVLVAAPGSLVAAYDT
VGTAPEANRGVGNFDAAGNSFSREALADAGLTPGSAHDVDGLAFTWPSSPVGRPDSVTL
TGETVRLDAPTSRLAFVGAATDGTHRGTAVVTFDDGSTATTTIGFGDWVLPSADGSPVE
GNSVVAQMNRRNGDKDSAFVFATAPYTAPEDRRVVAVRFPDVDDLHVFAIATEPAADVH
LVDVTVSLRCLAGTPYVAVRAANVSAGAVDVDLTTGVGSRSFTAVAPGANAYQSFAARG
ATGNVDVTVTATGEEGTQTVARTVVVPRCS (SEQ ID NO:11)

도면11a

11A

ATGACCAGACCACTCCCGCCCGGACGCGCGGTCGCGCGGTCCGGCAGCGGCCGCGCCCG
GCCCCTCGGCCTCGTGCTCGCCGCCGCACTCGCCGTCCCGCTCGGGGTGCCTCTCGCGG
CCCCCGCGGGAGCCCTCGCTGCCGCGCCCGCCGCGGCCGCCGAGCCCGGCGACTTCTCG
TCCTCGTTCGAGTCCGGCGACCCGGCCGCGCTGCCCACCACCGTGGCGGAGCGCGACGG
CGCGCCCTGGCAGGCGAACGTCGGCTCGTTCACGGCCGGCCTGCCCGGGAGCGTCCTCG
GGCAGCTGAAGGGCGTCACGGCGAGCGCGCAGAACCTGCCCAACGAGGGCGCGGCGAAC
CTCGCCGACGGCAGCTCGGGCACCAAGTGGCTCGCGTTCGCGTCGACCGGCTGGGTCCG
GTACGAGTTCGCCGAGCCCGTCTCGTTCGTCGCGTACACGATGACCTCCGGCGACGACG
CCGCCGGTCGCGACCCGAAGACCTGGACGGTCGAGGGGTCGAACGACGGGTCCACGTGG
GCCGCGCTCGACCGCCGGACGGACGAGGACTTCCCGAACCGCCAGCAGACGCGCACGTT
CGAGCTCGAGGCGCCCACCGCGGCGTACACGTACCTGCGCCTCAACGTCACGGCGAACT
CGGGCGACTCCATCGTCCAGCTCGCCGGGTGGGACCTCTCGGCCGACCTGAGCGCCGGC
CCGTCCGCGGCCCCCATGACGACGAAGGTCGGCACCGGGCCGCGCGTCAGCTTCACCAA
CAAGGCGGGCGTCGGGTTCTCCGGCCTGCACTCGCTCCGGTACGACGGCTCGCACCTCG
CCGACGGCGAGACGTACGCGACGAACGTGCTCTACGACGACGTGGACGTCGTCGTCGGC
GAGGACACGCGCCTGAGCTACACGATCTTCCCCGAGCTGCTCGACGATCTGCAGTACCC
GTCGACGTACGCGGCGGTGGACGTCCTGTTCACCGACGGGACCTACCTGTCCGACCTCG
GCGCGCGCGACGCGCACGAGACGGTCGCGACCGCGCAGGCGCAGGGCGAGGGCAAGATC
CTCTACGCCGACCAGTGGAACTCGGTGCGGGTCGACCTCGGCGACGTCGCCGAGGGCAA
GACCGTGGACCAGGTGCTGCTCGGGTACGACAACCCGGGCGGTCACGCCGGGACGAAGT
TCGCGGGCTGGCTCGACGACGTCGAGATCACGGCGGAGCCGGCCACGATCGACGGGTCG
AGCCTCGCCAACTACGTGGACACGCGCCGCGGCACGCTCGCGTCGGGCAGCTTCTCGCG
CGGGAACAACATCCCCGCGACGGCGACGCCGAACGGGTTCAACTTCTGGACGCCGTACA
CGAACGCCTCCTCGCAGAGCTGGCTGTACGAGTACCACAAGGCCAACAACGCCAACAAC
AAGCCCGTCCTCCAGGGCTTCGGGATCTCGCACGAGCCGAGCCCGTGGATGGGCGACCG
CAACCAGCTGACGTTCCTCCCGTCGACGGCGTCGGGGACGCCCGACGCCACGCTCTCGA
CGCGCGGCCTCGAGTTCGACCACGCGGACGAGACGGCGCGGCCGGACTACTACGGGGTC
ACGTTCACCAACGGGTCCGCGATCGAGGCGACGCCCACCGACCACGGCGCGGTGCTCCG
CTTCAGCTACCCCGGAGCCAAGGGCCACGTGCTCGTGGACAAGGTGGACGGCTCCTCCA
AGCTCACGTACGACCAGGCCACGGGCACGATCTCCGGCTGGGTCGAGAACGGCTCGGGC
CTGTCCGTGGGCCGCACGCGCATGTTCGTCGCCGGCACCTTCGACCGTAGTCCGACGGC
GGTCGGGACGGCGGCGGGCAACCGTGCGGACGCGCGCTTCGCGACGTTCGAGACGTCGT
CCGACAAGACGGTCGAGCTGCGCGTCGCGACGTCGTTCATCAGCCTCGACCAGGCGCGC
AAGAACCTCGACCTGGAGGTGACGGGCAAGACCTTCACGGAGGTCAAGGCCGCCGCCGC
GCAGGCGTGGAACGACCGCCTGGGGGTCATCGAGGTCGAGGGCGCGAGCGAGGACCAGC
TCGTCACGCTGTACTCGAACCTCTACCGCCTCAACCTGTACCCGAACTCGCAGTTCGAG
AACACGGGCACGGCGCAGGAGCCGGTGTACAGGTACGCGAGCCCGGTCTCCGCGACCAC
GGGCTCCGCGACGGACACGCAGACCAACGCGAAGATCGTCGACGGCAAGATCTACGTGA
ACAACGGGTTCTGGGACACGTACCGCACGGCCTGGCCGGCGTACTCGCTCCTCTACCCG
GAGCTCGCGGCCGAGCTGGTCGACGGGTTCGTCCAGCAGTACCGCGACGGCGGCTGGAT
CGCGCGCTGGTCCTCGCCGGGCTACGCCGACCTCATGACGGGCACGAGCTCCGACGTGG
CGTTCGCCGACGCGTACCTCAAGGGCTCGCTCCCCACGGGCACGGCGCTCGAGGCGTAC
GACGCCGCGCTGCGCAACGCGACCGTCGCGCCGCCGAGCAACGCCGTGGGCCGCAAGGG
CCTGCAGACCTCGCCGTTCCTCGGGTTCACGCCGGAGTCCACGCACGAGTCCGTGTCGT

도면11b

GGGGCCTGGAGGGCCTGGTCAACGACTTCGGCATCGGCAACATGGCCGCCGCCCTCGCG
GAGGACCCGGCGACGCCGGAGGAGCGCCGCGAGACGCTGCGCGAGGAGTCCGCGTACTT
CCTCGAGCGGGCCACGCACTACGTCGAGCTGTTCGACCCCGAGGTCGACTTCTTCGTGC
CGCGGCACGAGGACGGCACGTGGGCCGTCGACCCCGAGACGTACGACCCGGAGGCCTGG
GGCGGCGGGTACACCGAGACGAACGGCTGGAACTTCGCGTTCCACGCCCCGCAGGACGG
CCAGGGCCTCGCCAACCTCTACGGCGGCAAGCAGGGCCTCGAGGACAAGCTCGACGAGT
TCTTCTCCACGCCGGAGAAGGGCGCCGGCAACGGCGGCATCCACGAGCAGCGCGAGGCG
CGCGACGTCCGCATGGGCCAGTGGGGCATGAGCAACCAGGTGTCGCACCACATCCCGTG
GCTCTACGACGCCGCGGGCGCGCCGTCGAAGGCGCAGGAGAAGGTCCGCGAGGTCACCC
GCCGCCTGTTCGTGGGCAGCGAGATCGGCCAGGGCTACCCGGGCGACGAGGACAACGGC
GAGATGTCGTCGTGGTGGATCTTCGCCTCGCTCGGCTTCTACCCGCTCCAGGTCGGCTC
GGACCAGTACGCGGTCGGTTCGCCGCTGTTCGACAAGGCGACCGTGCACCTGCCGGACG
GCGACCTCGTCGTCAACGCCGAGAACAACTCGGTCGACAACGTCTACGTGCAGTCCCTC
GCGGTGGACGGCGAGGCCCGCACCTCGACGTCACTCTCCCAGGCGGACCTCTCGGGCGG
CACGACTCTGGACTTCGTCATGGGTCCGGAGCCGTCGGACTGGGGCACGGGCGAGGACG
ACGCGCCGCCGTCGCTCACCGAGGGCGACGAGCCCCCGACGCCGGTGCAGGACGCGACG
ACCGCGGGCCTCGGCACCACCACCGTCGCCGACGGCGACGCCACCACGAGCGCCGCGGC
GCTCACGGACAACACGTCCGGGACGCGCACGACGTTCGCCACCACGACGCCGTCGATCA
CGTGGGCGGGCAACGGCATCCGCCCGACCGTCGGGTCGTACACGCTGACCTCCGGGGCG
AGCGGGACGGCGTCACCGTCCGCATGGACTCTCGAGGGTTCCGACGACGGCGAGACGTG
GACGACGCTCGACGAGCGGTCCGGCGAGCAGTTCCGCTGGGCCCTGCAGACGCGGCCGT
TCACGGTCGCGGAGCCGACGGCGTTCGCGCGGTACCGGGTCACGGTCACCGCGACGTCG
GGCTCCGGCGCGCTGTCGCTCGCCGAGGTCGAGCTCCTCGCCGACCCGAAGGAGTCGGG
GGCCGAGGAGCTCACCCTCTCGGCCGCGCCGGACCGTGACGGCGTCACGGGCCGCGAGG
TCTCGGGCTCGTTCGCGACCCTCACCGGGGTCGAGGGCGACGTCGCGGCGCTCGACGTG
CAGGTCGCGTTCGGCGACGGCTCCGAGCCGGTCGCCGGGACGCTGCGGGCGGGCGCGTT
CGGCGGGTACGCGGTGGACGCCGCGCACACGTGGACCGCACCCGGCGTCTACCCCGTGA
CCGTCACGGTCTCGGGCGAGGGGATCGAGACCGTCTCGGCCTCCTCGTACGTCAGCGTC
TCGCTCCTGCGCGAGGGCTCGCTGCTCGCCGCGTACGACAACGTCTGCATCGGCGACGC
CGGGACGACGGTCGGCTCGTGCGACGGCCAGGGCGTGTTCTTCGACCGGGCGCAGCTCG
CGGCGAAGGGCTTCGTCCAGGGCGAGCGCGCGACGGTGCCGGGCACGGACCTCGCGTTC
GACGTCCCGGCGGTCCCCGCCGGGCAGCCGGACAACGCCACGGGCGACGGGCAGACCAT
CGAGCTCGACGTCCCCGCGGACGCGGAGCAGCTCTCGGTGATCGGCACGGGCACGGAGA
AGAACCAGCAGGCCACCGGCACGCTGACCTTCGACGACGGCTCGACCCAGCCGATCGAC
CTGAGCTTCGGCGACTGGTCGGGCGCGGCCCGCAACCCCGTGTTCGGCAACATCCCCGT
CGCGGTGACGGACAGCCGCCTCCGCGGCGGCAGCCCGCAGACCGGCACCCCCGCCGCGT
TCTTCGCGACGGCGCCGATCACCCTCCCCGAGGGCAAGCGGCCCGTGAGCCTCACGCTC
CCGGACCAGCCGGGCGAGCTCTCGCGCGACGGCCGCATCCACGTGGTCGCGGTCGCGCA
CGACGGCACGTTCGCCGAGCACCCCGCGCTCGAGGTCACGGCCGCGGAGGGCGTGACGC
TCGCCGTCGGGCAGACCTCGGACGTGGCGCTCGCCCAGGTGGCGGGCGGCCGCGAGGGC
GCCGACCTCCGGGCGGCGGTCACGTGGGGCGACGGCTCCGACGTCGCGGCCGGCGCGGT
GACCGACGGGTCGGTCTCCGGCTCGCACGCCTACACGGCGGCCGGGACGTACACGGCGT
ACGTCGTGGTCGACGACGGCTGGACCAGCCAGGTGGTCGAGGTCCCCGTGACCGTGACC
GAGGCGGAGCCGGCCCTCGCCGTCGACGTCACGGTGAGCACACGCTGCCTCGCCGGCAA
GGCGTACGTCGCGGTCCGCGCCGAGAACGGCGAGGACGTGCCGCTCGCGATCCGGCTCG
TCACGCCGTTCGGCACCAAGGAGGTCGCGGCCGTCGCGCCGGGCGCCAACGCCTACCAG
TCGTTCGCGACGCGGGTCACGGCGGTCGAGGCCGGCACCGTCACCGTCGAGGCGACGCG
CGGCACCGGCGACGAGGAGGTGACGGCGTCGATCCAGGCCGACTACGCCGCCGTGACCT
GCGGC (SEQ ID NO:12)

도면11c

11B

MTRPLPPGRAVARSGSGRARPLGLVLAAALAVPLGVPLAAPAGALAAAPAAAAEPGDFS
SSFESGDPAALPTTVAERDGAPWQANVGSFTAGLPGSVLGQLKGVTASAQNLPNEGAAN
LADGSSGTKWLAFASTGWVRYEFAEPVSFVAYTMTSGDDAAGRDPKTWTVEGSNDGSTW
AALDRRTDEDFPNRQQTRTFELEAPTAAYTYLRLNVTANSGDSIVQLAGWDLSADLSAG
PSAAPMTTKVGTGPRVSFTNKAGVGFSGLHSLRYDGSHLADGETYATNVLYDDVDVVVG
EDTRLSYTIFPELLDDLQYPSTYAAVDVLFTDGTYLSDLGARDAHETVATAQAQGEGKI
LYADQWNSVRVDLGDVAEGKTVDQVLLGYDNPGGHAGTKFAGWLDDVEITAEPATIDGS
SLANYVDTRRGTLASGSFSRGNNIPATATPNGFNFWTPYTNASSQSWLYEYHKANNANN
KPVLQGFGISHEPSPWMGDRNQLTFLPSTASGTPDATLSTRGLEFDHADETARPDYYGV
TFTNGSAIEATPTDHGAVLRFSYPGAKGHVLVDKVDGSSKLTYDQATGTISGWVENGSG
LSVGRTRMFVAGTFDRSPTAVGTAAGNRADARFATFETSSDKTVELRVATSFISLDQAR
KNLDLEVTGKTFTEVKAAAAQAWNDRLGVIEVEGASEDQLVTLYSNLYRLNLYPNSQFE
NTGTAQEPVYRYASPVSATTGSATDTQTNAKIVDGKIYVNNGFWDTYRTAWPAYSLLYP
ELAAELVDGFVQQYRDGGWIARWSSPGYADLMTGTSSDVAFADAYLKGSLPTGTALEAY
DAALRNATVAPPSNAVGRKGLQTSPFLGFTPESTHESVSWGLEGLVNDFGIGNMAAALA
EDPATPEERRETLREESAYFLERATHYVELFDPEVDFFVPRHEDGTWAVDPETYDPEAW
GGGYTETNGWNFAFHAPQDGQGLANLYGGKQGLEDKLDEFFSTPEKGAGNGGIHEQREA
RDVRMGQWGMSNQVSHHIPWLYDAAGAPSKAQEKVREVTRRLFVGSEIGQGYPGDEDNG
EMSSWWIFASLGFYPLQVGSDQYAVGSPLFDKATVHLPDGDLVVNAENNSVDNVYVQSL
AVDGEARTSTSLSQADLSGGTTLDFVMGPEPSDWGTGEDDAPPSLTEGDEPPTPVQDAT
TAGLGTTTVADGDATTSAAALTDNTSGTRTTFATTTPSITWAGNGIRPTVGSYTLTSGA
SGTASPSAWTLEGSDDGETWTTLDERSGEQFRWALQTRPFTVAEPTAFARYRVTVTATS
GSGALSLAEVELLADPKESGAEELTLSAAPDRDGVTGREVSGSFATLTGVEGDVAALDV
QVAFGDGSEPVAGTLRAGAFGGYAVDAAHTWTAPGVYPVTVTVSGEGIETVSASSYVSV
SLLREGSLLAAYDNVCIGDAGTTVGSCDGQGVFFDRAQLAAKGFVQGERATVPGTDLAF
DVPAVPAGQPDNATGDGQTIELDVPADAEQLSVIGTGTEKNQQATGTLTFDDGSTQPID
LSFGDWSGAARNPVFGNIPVAVTDSRLRGGSPQTGTPAAFFATAPITLPEGKRPVSLTL
PDQPGELSRDGRIHVVAVAHDGTFAEHPALEVTAAEGVTLAVGQTSDVALAQVAGGREG
ADLRAAVTWGDGSDVAAGAVTDGSVSGSHAYTAAGTYTAYVVVDDGWTSQVVEVPVTVT
EAEPALAVDVTVSTRCLAGKAYVAVRAENGEDVPLAIRLVTPFGTKEVAAVAPGANAYQ
SFATRVTAVEAGTVTVEATRGTGDEEVTASIQADYAAVTCG (SEQ ID NO:13)

도면12a

12A

GCGCTCGCCGTCGTCGGCCTCGCGCCCGCGACCGCCGCGAGCGCCGCCCCCGAGCCGCC
GTCGGCCGACTACGCGTCCCTGGTCGACGTCTTCGTCGGCACCGAGGGCGACTTCGGCA
ACGACATGCCCGCCGCGCAGGCGCCGAACGGCCTCGCGAAGGTCAACCCGCGCACGACC
CCGGGCCGCAACAACACCGGGTACGACTACGCGCAGTCGAAGATCTCGGGCTTCACGCA
CACCAACCTCGACGGGGTCGGGGGCTCCGGCGGCGGTGGTGACCTCCTCGTGGTGCCGA
CGTCCGGGTCGTACACGGCGCGCCCCGGCACGGGCACGTACGCGCACCCGTTCTCGCAC
GACGACGAGGACGCCGGACCGGGCTTCTACTCCGTCGGGCTCGGCAACGTCGCGGGCAC
GGACGGCGCGATCACCGGCGCGCCGGGCACGATCGAGGCCGAGGTCGCGGCGGCCACGC
GCTCGGGCGTGCACCGCTACGCGTTCCCCGCGGGCTCGACGCCGAGCCTCGTCGTGGAC
CTCGAGACGAACAACACGAGCCGCCGGTCGTCCTCGGTGCAGGTCGAGACGCGCGCGGA
CGGCACCGTGGAGCTGTCCGGACAGGTCACGGGCTACTTCTACAACGCGGCCTACACGC
TGTACTACACCGCGCGCACGCTCCAGCCCGCGACGGTGCAGACGTGGGGCGACGACGAC
CGGCTCGTCGACGCCACGGCCCAGGACGGCGTCGACACCGGCGCGATCCTCACGTTCGA
CCCGGCGGACGCCGGGGAGATCGGGCTCCAGGTCACCCTGTCGCCGGTGAGCGTCGAGC
AGGCGCGGATCGACCAGCAGGTCGAGCTCGGCGACCTGTCGTTCGACGCGATCCGTGAC
CGCACCCGCGCGGAGTGGAACGCGACGCTCGGGCGGGTCGCGATCGACGCCTCGACGGC
GACGGACCCGACGGGCGAGCTCCAGCGGCTCTTCTACACGCACCTCTACCGCATGTTCG
CGATGCCGATGAACGCGACGAGCACCTCGGGCACGTACCGCGGCGTCGACGGGGCGGTG
CACGCCGCGCAGGGCTTCACGTACTACGACTCGTGGGCCACGTGGGACGACTTCCGCAA
GTTCTCCGTCATCGCGTACATCGACCCGGCGCTGTACCGGGACATGGTGCAGTCGCTGG
TCTACCTGTTCGCGGACGCCGAGGCGACGGGCACCGGCGGCGGCCTCGGCGGGTTCGTG
CACTCGGTCCCGACGGTGCGCTGGGAGCGGTCGTCGGTCGTGGTCGCGGACGCGATCGC
CAAGGGCTTCGACGGGTTCGACCGCCTCGACGAGGCGTACCCGGCGCTCCAGCGGCTCG
TCGGGCAGTACAGCGCGGACGAGCTCCGGCGCGGCTACGTGGCGGGCAACCCCGGCGCG
TCCGTGCAGCGCGGCTACGACCAGTACGGCCTGTCCGTGATCGCGGACGAGCTCGGCCT
GACCGAGGAGGCCGAGACGCTGCGCGAGCAGGCGTCGTGGCCGATCGAGAAGCTCACCA
AGCCGGGCGCGTGGACCGCCGCCGACGGCACGCAGGTCGGCCTCCTCACCCCGCGCGCC
GCGGACGGGTCGTGGCAGAGCGCCGACCACGCGAAGTTCGAGGCCGCCGGCCTCTACCA
GGGCACGCTCTGGCAGTACCACTGGTACGACGCGTACGACATGGACGCGCTCGTCGAGG
CGATGGGCGGCCACGAGGCGGCGCGCCTCGGCATGCGCCACATGTTCGGTGAGCACGCG
CCGGACGACGGCAAGGCCATGCTCCACTCGAACGCCAACGAGATCGACCTCCAGGCGCC
GTACCTCTTCAACTACACGGGCGAGCCGAGCCTCACGCAGAAGTGGGCGCGCGCGATCT
ACACGAAGGAGACCTGGAACCGGTACATCGCGACCGGCTCCTCCAGCGCCGTGCCGAGC
GGCGGCGGCGAGTTCACGCCGCCCTTGAAGACGAAGGTGTACCGGCTCGACCCCCGCGG
GATGCTCCCCACGATGGACAACGACGCGGGCACGATGTCGACGATGTTCGTCGCCGCGG
CCGTCGGGCTGTTCCCGGTGACCGCGGGCTCGTCCCAGTTCCAGGTCGGGTCGCCGTTC
TTCGACTCGACGACCATCACCTACGACGACGGCAGCGCCTTCACGGTCACGGCCGACGG
CGTCTCCGAGGACGCGTTCTACGTCCAGTCCGCGACGCTCGACGGCGCGACGTTCGGCA
ACACGTGGGTCGACTACGCCACCGTGGTCGGGGGAGCCGACCTCGCGTTCCGCATGGGC
GAGCAGCCGAGCGACTGGGGCACGGACACCGCGCCCGCGTTCTCGATGAGTACCGCGAC
CGACGAGCCGGCCGAGGGACCGCGCGTCAGCGCCGAACCGACCACCGTGCAGACCGGCG
ACGGCGGCGCGCTCGACGCGACCGTGACGCTCACGCTCGACGGCGCCCGCCTCGCCGCG
CCCGCCGGCACGGACCTCGTCACGAGCGGGGCGGCGAGCGTCGTCGGGCTGCCCGACGG
CGTCACGGCGGCCGTGACGGTCGCGTCGCCGACCGCGCTGACCGTCTCCCTGACGGGGA
CGGCGTCCGCCGACGCGCGCTTCTTCGTGCACCTGCGCGACGCCGCGCTCGCCGACGGC

도면12b

GTCGCCGCGGCGTCGCTCCAGGGACAGGGCGTCTCGGTGCGCTCGCCCCTGCGGCTGTC
CGTGGCGTCCGCCGAGCGCGACGCGCTCGCCGCGCTCGTCGACGACGCCGTGCTCGTGC
GGCACGGGAACTACTCCTCGGTGACGTTCGACCGGTTCTCCACCGCGCTGACGAAGGCG
CAGGAGGCCCTCGGTGACGAGGCCGCGACGAGCATCGCGCTGCGGTTCGCGGCCGACCG
GCTCGGTGCGGCGGCCGACGCGCTCGACCTCACGGGCGGCGGGTACCGCACGCTCGAGG
CCGAGCAGTCCGAGGCGTGGTCGGGCGGGGAGCTGAAGAACGAGGCGAACAGCTCGTCC
GGCAACCTCGGCGGCGTGCGCTCCGGGTCGTGGGTGCAGTACCGCGACATGACCTTCGA
GACCGCCGCCGGGGACACCCCGCCGCGCTTCCTCACGGTCCGGTACGACACGAGCTTCG
CCCCGACGGACACGCCGAGCACCGTGCGCGTGCACGCGGGCGACGTGAGCGGCCCTGTG
GTCGCGACCGTCGACCTGAAGGGCACGAGCGGCTGGGGCAAGTACACCGAGGTCACGGC
GGAGCTCGGCGACGTGCAGGCGCTCGTCGACGCGCAGGTCGTCACGTTCGAGCTGCTCG
CGCCGTCCGGGCGGAGCTGGGTCGGCAACTTCGACTGGTTCCGGTTCAGCGCCGAGGAC
CCGGCTGCCCCAGGTCAGCCGGGCGAGTCCCCGACGGTGACGATCGAGGCCGAGGACTG
GACCGCGAGCTCCGGTCGCGGGCTCAAGAAGGAGTCCTCGACGTGGACGAGCGGTCCGG
TGACGAACGTCGGCGGCACCGCGGACGGCGACTGGATCGCCTACGGCGAGGTCGACCTG
GGTGAGCTCCCGCTCGGCGAGCTGTCGGTCCACTACGTGCACAACTCCAACCGGTCCGG
GAACAACTCCGCGCTGTCGGTGTACCTCGACGCGTTCGACCCGGCGAACCCGGGCGAGC
CGTTCGTCACCGTGCCGCTGCCGACGACCGGGTCGAGCTGGACCGCGGACGGGACCGCG
ACCGTCGTCCTGCCCGAGACGGTGCAGGGGACGCACGAGGTGTTCGTGCGCCTGTCGAC
CGAGCCGTACGCCGACCACCCGTACGTCGCGAACCTCGACAGCCTGACGTTCGCGCCGG
GCGGCCCGACGTCGGTCGTCGTCGAGTCCGAGGCCTGGACGTCGAACTCCGGCCGCGGG
CTGAAGAACGAGAGCTCGACGTGGACGAGCGGTCCGGTGACGAACGTCGGCGGCACCGC
GGACGGCGACTGGCTCGCCTACGGCGAGATCGACCTCGGCTCCGCCGCGCTCGACCAGC
TCTCGGTCCACTACGTGCACAACTCCAACCGGTCCGGGCGGAACTCCGCGCTGTCGGTG
TACCTCGACGCGTTCGACCCGGCGAACCCGGGCGAGCCGTTCGTCACCGTCCCGCTGGC
CAACACCGGGTCGAGCTGGACGACGGACGGGACCGCCGTCGTCGACCTGCCGAGCACGG
TGCGCGGCAAGCACCAGGTGTGGGTGCGCCTGTCCACCGAGGCGTACGCCGACCACCCG
TACGTCGCCAACCTCGACAGCATGCGCTTCTTCACCGACGCGTACGACGTCGAGGTCCC
GCCGACCGACACCGCGGCGCTCGCGGCGGTGGTCGACGCGGCCGGGACGCCCGAGGCGG
AGATCGCGCGGTACGGCCGGATCGACGCGCGCGTCTTCACACGCGAGCTCGCGGCGGCA
CGGTCCGTGCTCGCCGACGCCGGCGCCACCCAGGCGCAGGCCGACGAGCGGGCGCGGCG
CCTCGGCCTGGCGACCGACCAGCTCGTGCCCGCCGAGCGCCGTCGGCTCGAGAACCTCG
TGGCGAGCGCCGAGGCCCTGACCGACGAGGGGTACAGCCCCGAGTCCTGGCAGGCCTTC
CGCACGGCTCTCGCCGCGGCGACCGGGACGCTCGACGACGCGGCGGCGTCCGACGAGGC
GCTGCACGACGCGCGGCTCGCGCTCCAGGGCGCCGTCGACGCCCTGGAGGAGCCGGCCG
ACGTCGTGCTCGTCGAGGTCGAGGTCAGCCCGCGCTGCCTCGCCGGCAAGCCCTACGTC
GCGGTCCGCGCGGTGAACGTCTCCGACGCGGCCGTCGACGTCGAGCTGGCGTCGTCACT
GGGCACGAGGTCGTTCGTCGGCGTCGCGCCGGGGGCGAGCGCGTACCAGTCGTTCGCCG
CGCGGTCCGCGACGGGCGACCTGGACGTCACCGTCACGGCGACGGGGGCGGACGGCACC
CAGACGGTCGAGCAGGTCGTCACCGTCCCGTCCTGCTCC (SEQ ID NO:14)

도면12c

12B

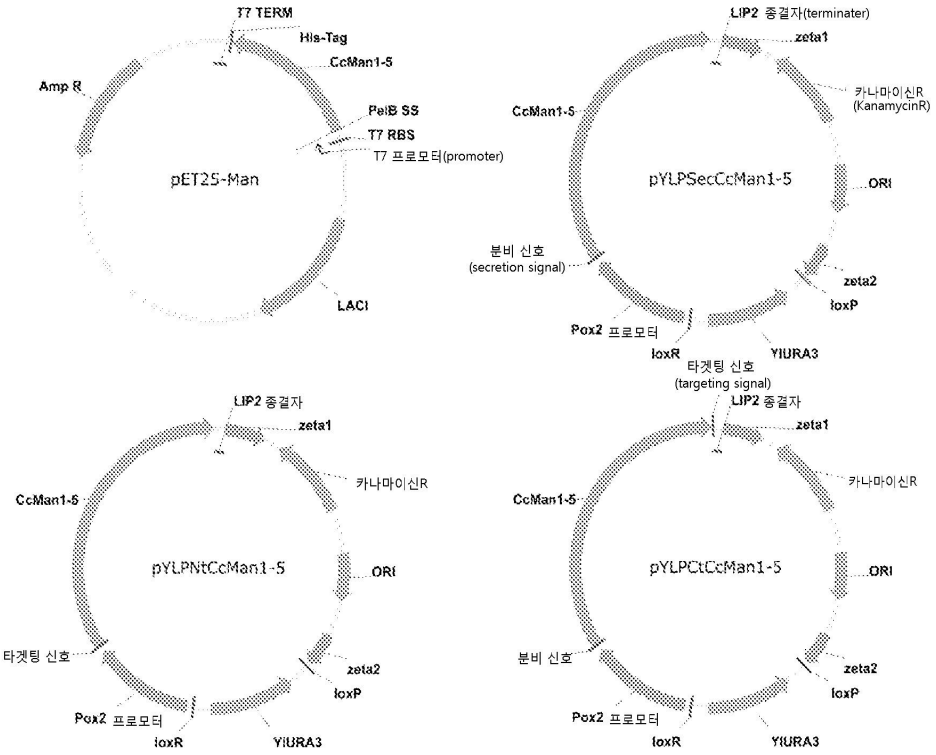
ALAVVGLAPATAASAAPEPPSADYASLVDVFVGTEGDFGNDMPAAQAPNGLAKVNPRTT
PGRNNTGYDYAQSKISGFTHTNLDGVGGSGGGGDLLVVPTSGSYTARPGTGTYAHPFSH
DDEDAGPGFYSVGLGNVAGTDGAITGAPGTIEAEVAAATRSGVHRYAFPAGSTPSLVVD
LETNNTSRRSSSVQVETRADGTVELSGQVTGYFYNAAYTLYYTARTLQPATVQTWGDDD
RLVDATAQDGVDTGAILTFDPADAGEIGLQVTLSPVSVEQARIDQQVELGDLSFDAIRD
RTRAEWNATLGRVAIDASTATDPTGELQRLFYTHLYRMFAMPMNATSTSGTYRGVDGAV
HAAQGFTYYDSWATWDDFRKFSVIAYIDPALYRDMVQSLVYLFADAEATGTGGGLGGFV
HSVPTVRWERSSVVVADAIAKGFDGFDRLDEAYPALQRLVGQYSADELRRGYVAGNPGA
SVQRGYDQYGLSVIADELGLTEEAETLREQASWPIEKLTKPGAWTAADGTQVGLLTPRA
ADGSWQSADHAKFEAAGLYQGTLWQYHWYDAYDMDALVEAMGGHEAARLGMRHMFGEHA
PDDGKAMLHSNANEIDLQAPYLFNYTGEPSLTQKWARAIYTKETWNRYIATGSSSAVPS
GGGEFTPPLKTKVYRLDPRGMLPTMDNDAGTMSTMFVAAAVGLFPVTAGSSQFQVGSPF
FDSTTITYDDGSAFTVTADGVSEDAFYVQSATLDGATFGNTWVDYATVVGGADLAFRMG
EQPSDWGTDTAPAFSMSTATDEPAEGPRVSAEPTTVQTGDGGALDATVTLTLDGARLAA
PAGTDLVTSGAASVVGLPDGVTAAVTVASPTALTVSLTGTASADARFFVHLRDAALADG
VAAASLQGQGVSVRSPLRLSVASAERDALAALVDDAVLVRHGNYSSVTFDRFSTALTKA
QEALGDEAATSIALRFAADRLGAAADALDLTGGGYRTLEAEQSEAWSGGELKNEANSSS
GNLGGVRSGSWVQYRDMTFETAAGDTPPRFLTVRYDTSFAPTDTPSTVRVHAGDVSGPV
VATVDLKGTSGWGKYTEVTAELGDVQALVDAQVVTFELLAPSGRSWVGNFDWFRFSAED
PAAPGQPGESPTVTIEAEDWTASSGRGLKKESSTWTSGPVTNVGGTADGDWIAYGEVDL
GELPLGELSVHYVHNSNRSGNNSALSVYLDAFDPANPGEPFVTVPLPTTGSSWTADGTA
TVVLPETVQGTHEVFVRLSTEPYADHPYVANLDSLTFAPGGPTSVVVESEAWTSNSGRG
LKNESSTWTSGPVTNVGGTADGDWLAYGEIDLGSAALDQLSVHYVHNSNRSGRNSALSV
YLDAFDPANPGEPFVTVPLANTGSSWTTDGTAVVDLPSTVRGKHQVWVRLSTEAYADHP
YVANLDSMRFFTDAYDVEVPPTDTAALAAVVDAAGTPEAEIARYGRIDARVFTRELAAA
RSVLADAGATQAQADERARRLGLATDQLVPAERRRLENLVASAEALTDEGYSPESWQAF
RTALAAATGTLDDAAASDEALHDARLALQGAVDALEEPADVVLVEVEVSPRCLAGKPYV
AVRAVNVSDAAVDVELASSLGTRSFVGVAPGASAYQSFAARSATGDLDVTVTATGADGT
QTVEQVVTVPSCS (SEQ ID NO:15)

도면12d

12C

APEPPSADYASLVDVFVGTEGDFGNDMPAAQAPNGLAKVNPRTTPGRNNTGYDYAQSKISGF
THTNLDGVGGSGGGGDLLVVPTSGSYTARPGTGTYAHPFSHDDEDAGPGFYSVGLGNVAGTD
GAITGAPGTIEAEVAAATRSGVHRYAFPAGSTPSLVVDLETNNTSRRSSSVQVETRADGTVE
LSGQVTGYFYNAAYTLYYTARTLQPATVQTWGDDDRLVDATAQDGVDTGAILTFDPADAGEI
GLQVTLSPVSVEQARIDQQVELGDLSFDAIRDRTRAEWNATLGRVAIDASTATDPTGELQRL
FYTHLYRMFAMPMNATSTSGTYRGVDGAVHAAQGFTYYDSWATWDDFRKFSVIAYIDPALYR
DMVQSLVYLFADAEATGTGGGLGGFVHSVPTVRWERSSVVVADAIAKGFDGFDRLDEAYPAL
QRLVGQYSADELRRGYVAGNPGASVQRGYDQYGLSVIADELGLTEEAETLREQASWPIEKLT
KPGAWTAADGTQVGLLTPRAADGSWQSADHAKFEAAGLYQGTLWQYHWYDAYDMDALVEAMG
GHEAARLGMRHMFGEHAPDDGKAMLHSNANEIDLQAPYLFNYTGEPSLTQKWARAIYTKETW
NRYIATGSSSAVPSGGGEFTPPLKTKVYRLDPRGMLPTMDNDAGTMSTMFVAAAVGLFPVTA
GSSQFQVGSPFFDSTTITYDDGSAFTVTADGVSEDAFYVQSATLDGATFGNTWVDYATVVGG
ADLAFRMGEQPSDWGTDTAPAFSMSTATDEPAEGPRVSAEPTTVQTGDGGALDATVTLTLDG
ARLAAPAGTDLVTSGAASVVGLPDGVTAAVTVASPTALTVSLTGTASADARFFVHLRDAALA
DGVAAASLQGQGVSVRSPLRLSVASAERDALAALVDDAVLVRHGNYSSVTFDRFSTALTKAQ
EALGDEAATSIALRFAADRLGAAADALDLTGGGYRTLEAEQSEAWSGGELKNEANSSSGNLG
GVRSGSWVQYRDMTFETAAGDTPPRFLTVRYDTSFAPTDTPSTVRVHAGDVSGPVVATVDLK
GTSGWGKYTEVTAELGDVQALVDAQVVTFELLAPSGRSWVGNFDWFRFSAEDPAAPGQPGES
PTVTIEAEDWTASSGRGLKKESSTWTSGPVTNVGGTADGDWIAYGEVDLGELPLGELSVHYV
HNSNRSGNNSALSVYLDAFDPANPGEPFVTVPLPTTGSSWTADGTATVVLPETVQGTHEVFV
RLSTEPYADHPYVANLDSLTFAPGGPTSVVVESEAWTSNSGRGLKNESSTWTSGPVTNVGGT
ADGDWLAYGEIDLGSAALDQLSVHYVHNSNRSGRNSALSVYLDAFDPANPGEPFVTVPLANT
GSSWTTDGTAVVDLPSTVRGKHQVWVRLSTEAYADHPYVANLDSMRFFTDAYDVEVPPTDTA
ALAAVVDAAGTPEAEIARYGRIDARVFTRELAAARSVLADAGATQAQADERARRLGLATDQL
VPAERRRLENLVASAEALTDEGYSPESWQAFRTALAAATGTLDDAAASDEALHDARLALQGA
VDALEEPADVVLVEVEVSPRCLAGKPYVAVRAVNVSDAAVDVELASSLGTRSFVGVAPGASA
YQSFAARSATGDLDVTVTATGADGTQTVEQVVTVPSCS (SEQ ID NO:50)

도면13



도면14

GCACCGGCAGATGAAGGCACCGTTACCGCAGCAGCCGGTGATGATCTGACCCTGGAAGTTAATCCGTTTGT
TGGCACCGAAAGCGAAGGTAATGCATATCCGGGTGCAACCGTTCCGTTTGGTATGGTTCAGCTGTCTCCGG
ATAATACCAATAGCTATGCCAGCACCAGCTATAGCACCAATGCAGGTCGTGTTTGGGGTTTTAGCCATCGT
CATGTTAATAGCGCAGGTTGTCCGGCAGCCGGTGAACTGCTGGTTACACCGGATACCAGCGCAACACCGCG
TACCAGCCGTAGCTTTATTGCCATCAAAGATCAGAAAAGCACCGAACGTGCAAGCGCAGGTTTTTATGAAG
TTACCCTGGCAAATGATGTTCATGCAGAACTGACCGCAACCACCCGTGTTGGTGCACATCGTTATACCTTT
CCGGCAAGCACCACCTCTCATCTGAGCTTTAATGTTGGTCAGACCCTGCGTGATGCCGGTGCAAGCAGCGT
TACCTGGGTTGATGATCGTACACTGGAAGGTTGGGTTGATAATGGTGGTTTTTGTGGTGGTACACCGGATA
AACAGCGCTATTTTTTTAGCGCAACCTTTGATCGTCCGGTTGCCAGCAGCGGTACATGGGGCACCGATGCA
CGTTATGTTGCAGGTAGCACCACCAGTGAAGTTGCCGGTGGTAATAATGGTGCAGTTGCCGTTTTTGATAC
CACCACCGATCGTGATGTTGAAGTTAGCGTTGGTGTTAGCTTTGTTAGCGTTGATGGTGCACGTGCAAATC
GTGAAGCCGAAGCAACCGATGAAGGTGGTCAGGTTGCATTTGATACCGTTCGTGAAGAAGCACGCGACGCC
TGGAATGCAGAACTGGGTCGTGCAGCAATTGATGCATCTCCGGATCAGCGTCGTATCTTTTATACCCAGCT
GTATAAAACCCTGCTGAGCCCGACCATTGGTTCTGATGTTGATGGTCGTTATCGTGGTATGGATCTGGAAG
TTCATCAGGCAGATGGCTGGGATTATTATCAGAACTTTAGCCTGTGGGATACCTATCGTACCCAGGCAACC
CTGCATGCACTGCTGCTGCCGGAACGTGCACAGGATATTGTTCGTAGCATGTATCAGCATCGTGTTGAAGG
TGGTTGGCTGCCTCGTTGGTCTCTGGGTGCACTGGAAACCAATATCATGGCAGGCGATCCGGTTACCCCGT
GGCTGGCAGAAAATTTTGCACTGGGCACCGTTCCGGATGATATTGCAGATGAACTGTGGGATTATCTGGTT
GAAAATGCAACCACCACCCCTCCGGATGATGTTGCCAGCGTTGGTCGTCGTAGCACCGAATTTTATGCCGA
ACATGGTCATGTTCCGTTTTATCCGGAAAACGAAGGTGGCCTGGGTGGTCAGTTTGAAGAATATCGTCATG
GTGGTAGCGCAACCCTGGAACTGGCACTGGCAGATGCAAGCCTGGGTGCCGCAGCAGAACGTACCGGTCGT
GAAGGTGGCCAGGCATTTCTGGATAAAGGTCGCAATTGGCGTAATCTGTGGAATCCGGATGTTGAACTGAG
CGGTGGTTTTCAGGGTATGGTTAATGCAAAACGTCCGACCGGTGAATTTGTTACCCTGCCGGAACTGACCG
ATGTTACCCGTAGCGGTTTTCATGAAGGTGTTCCGTGGCAGTATCAGTGGATGGTTCCGCAGGATGTTACC
GGTCTGCAAGAAGTTATGGGAGGCGAAGATGGTTTTGTGGAACGCCTGGATTATTATTTTGATCAGCCTGC
ACTGGCAGCAAATCCGGGGTGTTAGCCCGAGCACCTGGGCAAAAGGTGGTAGCAGCTATTATACCACCATTC
GCTATAATCCGGGTAATGAACCGACCATTATGAATGCATGGCTGTATGGTTATGTTGGTCAGCCGTGGAAA
ACCAATGATGTTCTGGCAGCCAATCTGAATCGTTTTCCGGATACACCGGGTGGTGGTGTTGGTAATGATGA
TCTGGGCACCCTGGCAGCATGGTATGTTATGGCCAGCCTGGGTTTTGAACCGGTTATGCCTGGTAGCGGTA
TTCTGGCACTGAATGCACCGAAAGTTCAGGCAGCAACCCTGACCACCGATGCCGGTGCCACCCTGCGTATT
GATGCAGCCGGTGCAAATGAAAAACTGCCGAGCTATGTTGCCGGTCTGGAAGTTGATGGTGTTGCACATAC
CGCAGCATGGCTGGATGTTGCAGCACTGCAGGATGGTGGCACCCTGGATTTTGATCTGAGCGGTACAAGCG
CAGGTCTGACATGGGGTACAGGTGCAGCAGATCGTATTCCGAGCGTTAGCGCAGTTGCACCGCCTGCACCG
GTTGAAGTGGAAGCAAGCGCACGTTGTCTGGGTGGTCGTGCATTTGTTGCAGTTCGTGCAACCAGCACCGC
AGATGCACCGGTGGATGTTACACTGACCACACCGTTTGGTGAACGTACCGTTCGTCATGTTCAGCCTGGTC
GTAGCGCATATCAGAGCTTTACCACCCGTACCACCTCTGTTGAAGCAGGCACCGCAACCGTTACCGTTGTT
GCAGCAGATGGCACCACCTCAACCGTTGATGCAGCATATGAAGCACTGGCATGTGGTTAATAA (SEQ ID
NO:16)

도면15

GCAGGCACCGAAGCAGCAACCGGTTCTGATGCAGCAGCAGTTGATGGTCCGCTGGTTGATTATGTGAATCC
GTTTATTGGCACCAAAGATGATGGTAATACCTATCCGGGTGCAGCAGTTCCGTTTGGTATGGTTCAGCTGT
CTCCGGATAATGGTCATAATGTGGGCTATGATTATGATCGTACCAGCGTTCGTGGTTTTAGCCTGGTTCAT
CTGAGCGGTGTTGGTTGTGGTCTGGGTGGTCCGCTGCCGACACTGCCGACCACCGGTGCAATTACCAGCAC
CGATTATGGTCAGTATGCACTGGGTTTTAGCCATGATGATGAAGAAGCATCTCCGGGTTATTATCGTGTTG
GTCTGCAGGCACCTGCAGGAACCATTGAAGCAGAACTGACCGCAACCGAACGTACCGGTGTTCAGCGTTAT
ACCTTTCCGGCAACCGCACAGGCAAATGTTCTGCTGAATGCAGGTCAGGCACTGAATCGTGTTACCGAATC
TGATGTTCGTGTTGTTGATGATCGTACCGTTGAAACCCGTATTACCGTGCGTGGTTTTTGTCAGGATACCG
AACCGCAGACCATTTGGACCCGTACCACCTTTGATCGTCCGTTTGTTGCACATGGCACCTGGGATGGTCAG
GTTGTTACCGCAGGCGCAGATGCAGCAAGCGGTGGTGAAGGTCGTCGTGGTGCATATGTTACCTTTGATAC
AACCGGTGGTGATCTGGATGTTGAAGCAGTTACCGCAATGAGCTATGTTGGTGCAGATGGTGCAGCAGCAA
ATCTGGCAGCAGAAGCAGGCACCTTTGACGCAGTTCATGATGCAGCACGTAGCGCATGGGAAGAACGTCTG
GGTCTGGTTCGTGTTGCACAGGGTGATCCGGATGATCTGCGTACCTTTTATAGCAGCCTGTATCGTAGCTT
TCTGGCACCGAATGTTGGTTCTGATGTGGATGGTCGTTATCGTGGTTGGGATCAGGAAGTTCACGCAGCAG
AACCGGATTTTACCTATTATCAGAATTATAGCCTGTGGGATACCTATCGTACCCAACAGCAACTGCTGTAT
CTGCTGGCACCGGATGAAAGCGCAGATATGGCACTGAGCCTGGTTCGTCAGGGTCAGCAGGGTGGTTGGCT
GCCTCGTTGGGGTTATGGTACAGTGGAAACCAATATTATGACCGGTGATCCGGCAACCCCGTTTCTGGTTA
GCGCATGGCGTCAGGGTCTGCTGGCAGGTCATGAAGAAGAAGCATACGCAGTCCTGCGTGAAAATGCAGAT
GGTGTTCCTCCGGCAGATAGCCCGTTTAATGGTCGTGCAGCCAATGTTGAATATCTGCGTGATGGTTTTGT
TCCGCATGAACCGGCACGTAGCGGTAAACCGGGTGATTATGATCTGCAGCATGGTGCAAGCGCAACCATGG
AATATGCACTGGCAGATGCAATGCTGAGCACCATGGCACGTGGTCTGGGTCATGATGAAGATGCAGATCGT
TATGCAGCCCGTGGTCAGAGCTATCGTAATGTTTTTGATCCGCGTACCGGTAATTTTCGTGCACGTAATGC
CGATGGTTTTTTTGTTGGTGATGCAGATCCGGCACATTCTGATGGTTTTCATGAAGGCACCGCAGTTCAGT
ATCAGTGGCTGGTTCCGCAGGATGTTCCGGGTCTGTTTGATCTGATGGGTGGCACCGATGCAGCCGTTGAT
CGTCTGGATGCATTTTTTGCCTATGATGAACTGGTTGCAGATCCTCCGCATGTTGCAAGCGAAGTTTGGGT
TAATGGCACCTATGATTATTATGGCTGGGAAACCTATAATCCGAATAATGAACCGAATCTGCATGCACCGT
ATGTTTATCTGTGGACCGGTCAGCCGTGGAAAACCACCGATGTTGTTCGTGCAGCAAGCACCCTGTTTACC
GATGGTCCGGATGGTGTTACCGGTAATGATGATCTGGGCACCATGAGCGCATGGCATGTTCTGAGCAGCAT
TGGTGTTTATCCGATTGTTCCGGGTGCCGATCTGTGGGGTCTGACCACACCGCTGTTTGATGATGTTACCA
TTACCCTGGACCCGGAAGTTTTTGGTCGTGATAGCCTGCGTCTGACCGCAGATGGTGTGGCACCGGATACC
CATTATACCCAGAGCGTTAGCCTGGGTGGTGAACCGCTGGATCGTGCATGGGTTACAGGTGATGAACTGAC
CGCTGCAGGCACCCTGGATGTTACCGTTGGCACCGAACCGAGCGCATGGGCAACCGATCCGGCAGCATCAC
CGGGTGCAGTTGTTCCGGCTGATGGCACCGTTGAACGTCTGTTTGTTGGTGCAACACCGCGTCAGCCGGTT
CTGGCACCGGGTGGTCGTACCGAAGTTGCAGTTCAGGTTGTTGCCCAGGGTGCAGGCACCTCTAGCGGCAC
CCTGGAAGTGACCTCTGATGGTGCAGTTACCGCCACCACCGATCTGGCAGAATGGACCGCAGAATCTGATG
GTCTGCCTGCCACCGTTGAAGGAACCGTTACCATTGAAGCTCCGGCAGATGCCGAACCGGGTCTGCATACC
GTTCGTCTGGTTGTTCGTGATGCAGCCGGTACAGAAGCAGTTCGCGAAGTTAGCGTTGTTGTTAGCGGTGA
AAGCTGGATTGCAGATGCCTTTGATAATGTGGGTATTGGTGATGCCGGTGCAGCAAATGCAAATCTGGATG
GTAGCGGTGCCTATCTGCTGCGTGATCTGCTGGCCGATCTGGGTGCAGTTCAGGGTCTGGAACTGACCGTT
CCGGGTACTGATCTGACCTATACCCTGGGTGCACCGCGTGCTGGTGCACCGGATAATGTTGCAGCCAGCGG
TGAAGTTCTGGAAGTTCCGGAACATCTGCGTAGCGCACGTCATCTGAGCGTTGTGGGCACCAGCACCCATG
GTACACATGGTGGTGGTCTGGTTCTGGGTTTTGCCGATGGTAGCAGCCAGACCGTTGATGTTCGTCTGAGC
GATTGGTGTACCGGTTCTCCGGAACCGGGTAATATTACCGTTGCAAAAGCCGGTGCACGTGGTGATCGTGA
AAATGTGCAGAAAATTGGCTGTGGTCTGTATGCAACCGCACCGGTGGCAATTCCGGAAGGTAAAGTTCTGA
CCAGCGTTACCCTGCCGTCTGATGAACGTTTTCATGTGTTTTGCAATTGCAACCGATGCAACCGGTGATGTT
CCGGCACCGCAGGTTGAAGTTACCGCACAGGCTCGTTGTCTGGGTGGTAAAGCATTTGTTGCAGTTCGTGC
ACTGAATACCGGTGAACAGCCTGCAGCAATTGAACTGGCAACCCCGTATGGTAGCAAACTGTTTGGTGATG
TTGCTCCGGGTGCAAATGCATATCAGAGCTTTGCAACCCGTGCAGCAGCCGTTGAAGCCGGTGAAGTTACC
GTTACCGTGACCACACCGGATGGTGAACCGCAGCAGGTTACCGCAGCATATGATGCAGCGGCATGTAGCTA
ATAA (SEQ ID NO:17)

도면16

GCAGGCACCGAAGCAGCAACCGGTTCTGATGCAGCAGCAGTTGATGGTCCGCTGGTTGATTATGTGAATCC
GTTTATTGGCACCAAAGATGATGGTAATACCTATCCGGGTGCAGCAGTTCCGTTTGGTATGGTTCAGCTGT
CTCCGGATAATGGTCATAATGTGGGCTATGATTATGATCGTACCAGCGTTCGTGGTTTTAGCCTGGTTCAT
CTGAGCGGTGTTGGTTGTGGTCTGGGTGGTCCGCTGCCGACACTGCCGACCACCGGTGCAATTACCAGCAC
CGATTATGGTCAGTATGCACTGGGTTTTAGCCATGATGATGAAGAAGCATCTCCGGGTTATTATCGTGTTG
GTCTGCAGGCACCTGCAGGAACCATTGAAGCAGAACTGACCGCAACCGAACGTACCGGTGTTCAGCGTTAT
ACCTTTCCGGCAACCGCACAGGCAAATGTTCTGCTGAATGCAGGTCAGGCACTGAATCGTGTTACCGAATC
TGATGTTCGTGTTGTTGATGATCGTACCGTTGAAACCCGTATTACCGTGCGTGGTTTTTGTCAGGATACCG
AACCGCAGACCATTTGGACCCGTACCACCTTTGATCGTCCGTTTGTTGCACATGGCACCTGGGATGGTCAG
GTTGTTACCGCAGGCGCAGATGCAGCAAGCGGTGGTGAAGGTCGTCGTGGTGCATATGTTACCTTTGATAC
AACCGGTGGTGATCTGGATGTTGAAGCAGTTACCGCAATGAGCTATGTTGGTGCAGATGGTGCAGCAGCAA
ATCTGGCAGCAGAAGCAGGCACCTTTGACGCAGTTCATGATGCAGCACGTAGCGCATGGGAAGAACGTCTG
GGTCTGGTTCGTGTTGCACAGGGTGATCCGGATGATCTGCGTACCTTTTATAGCAGCCTGTATCGTAGCTT
TCTGGCACCGAATGTTGGTTCTGATGTGGATGGTCGTTATCGTGGTTGGGATCAGGAAGTTCACGCAGCAG
AACCGGATTTTACCTATTATCAGAATTATAGCCTGTGGGATACCTATCGTACCCAACAGCAACTGCTGTAT
CTGCTGGCACCGGATGAAAGCGCAGATATGGCACTGAGCCTGGTTCGTCAGGGTCAGCAGGGTGGTTGGCT
GCCTCGTTGGGGTTATGGTACAGTGGAAACCAATATTATGACCGGTGATCCGGCAACCCCGTTTCTGGTTA
GCGCATGGCGTCAGGGTCTGCTGGCAGGTCATGAAGAAGAAGCATACGCAGTCCTGCGTGAAAATGCAGAT
GGTGTTCCTCCGGCAGATAGCCCGTTTAATGGTCGTGCAGCCAATGTTGAATATCTGCGTGATGGTTTTGT
TCCGCATGAACCGGCACGTAGCGGTAAACCGGGTGATTATGATCTGCAGCATGGTGCAAGCGCAACCATGG
AATATGCACTGGCAGATGCAATGCTGAGCACCATGGCACGTGGTCTGGGTCATGATGAAGATGCAGATCGT
TATGCAGCCCGTGGTCAGAGCTATCGTAATGTTTTTGATCCGCGTACCGGTAATTTTCGTGCACGTAATGC
CGATGGTTTTTTTGTTGGTGATGCAGATCCGGCACATTCTGATGGTTTTCATGAAGGCACCGCAGTTCAGT
ATCAGTGGCTGGTTCCGCAGGATGTTCCGGGTCTGTTTGATCTGATGGGTGGCACCGATGCAGCCGTTGAT
CGTCTGGATGCATTTTTTGCCTATGATGAACTGGTTGCAGATCCTCCGCATGTTGCAAGCGAAGTTTGGGT
TAATGGCACCTATGATTATTATGGCTGGGAAACCTATAATCCGAATAATGAACCGAATCTGCATGCACCGT
ATGTTTATCTGTGGACCGGTCAGCCGTGGAAAACCACCGATGTTGTTCGTGCAGCAAGCACCCTGTTTACC
GATGGTCCGGATGGTGTTACCGGTAATGATGATCTGGGCACCATGAGCGCATGGCATGTTCTGAGCAGCAT
TGGTGTTTATCCGATTGTTCCGGGTGCCGATCTGTGGGGTCTGACCACACCGCTGTTTGATGATGTTACCA
TTACCCTGGACCCGGAAGTTTTTGGTCGTGATAGCCTGCGTCTGACCGCAGATGGTGTGGCACCGGATACC
CATTATACCCAGAGCGTTAGCCTGGGTGGTGAACCGCTGGATCGTGCATGGGTTACAGGTGATGAACTGAC
CGCTGCAGGCACCCTGGATGTTACCGTTGGCACCGAACCGAGCGCATGGGCAACCGATCCGGCAGCATCAC
CGGGTGCAGTTGTTCCGGCTGATGGCACCGTTGAACGTCTGTTTGTTGGTGCAACACCGCGTCAGCCGGTT
CTGGCACCGGGTGGTCGTACCGAAGTTGCAGTTCAGGTTGTTGCCCAGGGTGCAGGCACCTCTAGCGGCAC
CCTGGAAGTGACCTCTGATGGTGCAGTTACCGCCACCACCGATCTGGCAGAATGGACCGCAGAATCTGATG
GTCTGCCTGCCACCGTTGAAGGAACCGTTACCATTGAAGCTCCGGCAGATGCCGAACCGGGTCTGCATACC
GTTCGTCTGGTTGTTCGTGATGCAGCCGGTACAGAAGCAGTTCGCGAAGTTAGCGTTGTTGTTAGCGGTGA
AAGCTGGATTGCAGATGCCTTTGATAATGTGGGTATTGGTGATGCCGGTGCAGCAAATGCAAATCTGGATG
GTAGCGGTGCCTATCTGCTGCGTGATCTGCTGGCCGATCTGGGTGCAGTTCAGGGTCTGGAACTGACCGTT
CCGGGTACTGATCTGACCTATACCCTGGGTGCACCGCGTGCTGGTGCACCGGATAATGTTGCAGCCAGCGG
TGAAGTTCTGGAAGTTCCGGAACATCTGCGTAGCGCACGTCATCTGAGCGTTGTGGGCACCAGCACCCATG
GTACACATGGTGGTGGTCTGGTTCTGGGTTTTGCCGATGGTAGCAGCCAGACCGTTGATGTTCGTCTGAGC
GATTGGTGTACCGGTTCTCCGGAACCGGGTAATATTACCGTTGCAAAAGCCGGTGCACGTGGTGATCGTGA
AAATGTGCAGAAAATTGGCTGTGGTCTGTATGCAACCGCACCGGTGGCAATTCCGGAAGGTAAAGTTCTGA
CCAGCGTTACCCTGCCGTCTGATGAACGTTTTCATGTGTTTGCAATTGCAACCGATGCAACCGGTGATGTT
CCGGCACCGCAGGTTGAAGTTACCGCACAGGCTCGTTGTCTGGGTGGTAAAGCATTTGTTGCAGTTCGTGC
ACTGAATACCGGTGAACAGCCTGCAGCAATTGAACTGGCAACCCCGTATGGTAGCAAACTGTTTGGTGATG
TTGCTCCGGGTGCAAATGCATATCAGAGCTTTGCAACCCGTGCAGCAGCCGTTGAAGCCGGTGAAGTTACC
GTTACCGTGACCACACCGGATGGTGAACCGCAGCAGGTTACCGCAGCATATGATGCAGCGGCATGTAGCTA
ATAA (SEQ ID NO:18)

도면17a

GCAGAACCGGGTGATTTTAGCAGCAGCTTTGAATCTGGCGATCCGGCAGCACTGCCGACCACCGTTGCAGA
ACGTGATGGTGCACCGTGGCAGGCAAATGTTGGTAGCTTTACCGCAGGTCTGCCTGGTAGCGTTCTGGGTC
AGCTGAAAGGTGTTACCGCAAGCGCACAGAATCTGCCGAATGAAGGTGCAGCAAATCTGGCAGATGGTAGC
AGCGGCACCAAATGGCTGGCATTTGCAAGCACCGGTTGGGTTCGTTATGAATTTGCAGAACCGGTTAGCTT
TGTTGCATATACCATGACCAGCGGTGATGATGCCGCAGGTCGTGATCCGAAAACCTGGACCGTTGAAGGTA
GCAATGATGGTTCTACCTGGGCAGCACTGGATCGTCGTACCGATGAAGATTTTCCGAATCGTCAGCAGACC
CGTACCTTTGAACTGGAAGCACCGACCGCAGCATATACCTATCTGCGTCTGAATGTTACCGCAAATAGCGG
TGATAGCATTGTTCAGCTGGCAGGTTGGGATCTGAGCGCAGATCTGTCTGCAGGTCCGAGCGCAGCACCGA
TGACCACCAAAGTTGGCACCGGTCCGCGTGTTAGCTTTACCAATAAAGCCGGTGTTGGTTTTAGCGGTCTG
CATAGCCTGCGTTATGATGGTAGCCATCTGGCCGATGGTGAAACCTATGCAACCAATGTGCTGTATGATGA
TGTTGATGTTGTGGTTGGTGAAGATACCCGTCTGAGCTATACCATTTTTCCGGAACTGCTGGATGATCTGC
AGTATCCGAGCACCTATGCAGCAGTTGATGTTCTGTTTACCGATGGCACCTATCTGAGCGATCTGGGTGCA
CGTGATGCACATGAAACCGTTGCAACCGCACAGGCACAGGGTGAAGGTAAAATTCTGTATGCCGATCAGTG
GAATAGCGTTCGTGTTGATCTGGGTGATGTTGCAGAAGGTAAAACCGTTGATCAGGTTCTGCTGGGTTATG
ATAATCCGGGTGGTCATGCAGGCACCAAATTTGCAGGTTGGCTGGATGATGTTGAAATTACCGCAGAACCG
GCAACCATTGATGGTAGCTCACTGGCAAATTATGTTGATACCCGTCGTGGCACCCTGGCAAGCGGTAGCTT
TAGCCGTGGTAATAATATTCCGGCAACCGCAACCCCGAATGGTTTTAATTTTTGGACCCCGTATACCAATG
CAAGCAGCCAGAGCTGGCTGTATGAATATCATAAAGCCAATAATGCGAATAATAAACCGGTTCTGCAGGGT
TTTGGTATTAGCCATGAACCGAGCCCGTGGATGGGTGATCGTAATCAGCTGACCTTTCTGCCGAGCACCGC
AAGCGGTACACCGGATGCAACCCTGAGCACCCGTGGTCTGGAATTTGATCATGCAGATGAAACCGCACGTC
CGGATTATTATGGTGTGACCTTTACCAATGGTAGCGCAATTGAAGCAACCCCGACCGATCATGGTGCAGTT
CTGCGTTTTAGCTATCCGGGTGCAAAAGGTCATGTTCTGGTGGATAAAGTTGATGGTAGCAGTAAACTGAC
CTATGATCAGGCAACCGGCACCATTAGCGGTTGGGTTGAAAATGGTAGCGGTCTGAGCGTTGGTCGTACCC
GTATGTTTGTTGCAGGCACCTTTGATCGTAGCCCGACCGCAGTTGGCACAGCAGCAGGTAATCGTGCAGAT
GCACGTTTTGCAACCTTTGAAACCAGCAGCGATAAAACCGTGGAACTGCGTGTTGCAACCAGCTTTATTAG
CCTGGATCAGGCACGTAAAAATCTGGATCTGGAAGTTACCGGTAAAACCTTTACCGAAGTTAAAGCAGCAG
CAGCACAGGCATGGAATGATCGTCTGGGTGTTATTGAAGTTGAAGGTGCAAGCGAAGATCAGCTGGTTACC
CTGTATAGCAATCTGTATCGCCTGAATCTGTATCCGAATAGCCAGTTTGAAAATACCGGCACCGCACAGGA
ACCGGTTTATCGTTACGCATCTCCGGTTAGCGCAACCACCGGTAGCGCAACCGATACCCAGACCAATGCCA
AAATTGTGGATGGCAAAATTTATGTGAATAATGGCTTTTGGGATACCTATCGTACCGCATGGCCTGCATAT
AGCCTGCTGTATCCGGAACTGGCAGCAGAACTGGTTGATGGTTTTGTTCAGCAGTATCGTGATGGTGGTTG
GATTGCACGTTGGAGCAGTCCGGGTTATGCAGATCTGATGACCGGTACAAGCTCTGATGTTGCATTTGCAG
ATGCCTATCTGAAAGGTAGCCTGCCGACCGGTACAGCACTGGAAGCATATGATGCAGCACTGCGTAATGCA
ACCGTTGCACCTCCGAGCAATGCAGTTGGTCGTAAAGGTCTGCAGACAAGCCCGTTTCTGGGTTTTACACC
GGAAAGCACCCATGAAAGCGTTAGCTGGGGTCTGGAAGGTCTGGTTAATGATTTTGGCATTGGCAATATGG
CTGCAGCACTGGCAGAAGATCCGGCAACACCGGAAGAACGTCGTGAAACCCTGCGTGAAGAAAGCGCATAT
TTTCTGGAACGTGCCACCCATTATGTTGAACTGTTTGATCCGGAAGTGGATTTTTTTGTTCCGCGTCATGA
AGATGGTACATGGGCAGTTGATCCGGAAACCTATGATCCGGAAGCATGGGGTGGTGGTTATACCGAAACCA
ATGGCTGGAATTTTGCATTTCATGCACCGCAGGATGGTCAGGGTCTGGCAAATCTGTATGGTGGTAAACAG
GGTCTGGAAGATAAACTGGATGAATTTTTTAGCACACCGGAAAAAGGTGCAGGTAATGGTGGTATTCATGA
ACAGCGTGAAGCACGTGATGTTCGTATGGGTCAGTGGGGTATGAGCAATCAGGTTAGCCATCATATTCCGT
GGCTGTATGATGCAGCCGGTGCTCCGAGCAAAGCACAGGAAAAAGTTCGCGAAGTTACCCGTCGTCTGTTT
GTTGGTAGCGAAATTGGTCAGGGTTATCCGGGTGATGAAGATAATGGTGAAATGTCCTCCTGGTGGATTTT
TGCAAGCCTGGGTTTTTATCCGCTGCAGGTTGGTAGCGATCAGTATGCAGTTGGTTCTCCGCTGTTTGATA
AAGCAACCGTTCATCTGCCGGATGGTGATCTGGTTGTTAATGCCGAAAATAATAGCGTGGATAATGTGTAT
GTTCAGAGCCTGGCAGTTGATGGTGAAGCACGTACCAGCACCAGCCTGAGCCAGGCAGATCTGAGCGGTGG
CACCACCCTGGATTTTGTTATGGGTCCGGAACCGAGCGATTGGGGCACCGGTGAAGATGATGCACCTCCGT
CACTGACCGAAGGTGATGAACCTCCGACACCGGTTCAGGATGCAACCACCGCAGGCCTGGGCACCACCACC
GTTGCCGATGGTGATGCCACCACCTCTGCAGCAGCCCTGACCGATAATACCAGCGGCACCCGTACCACCTT
TGCAACCACCACCCCGAGCATTACATGGGCAGGTAATGGCATTCGTCCGACCGTTGGTAGCTATACCCTGA
CCTCTGGTGCAAGCGGCACCGCAAGCCCGTCTGCATGGACCCTGGAAGGTTCTGATGATGGCGAAACCTGG
ACCACACTGGATGAACGTAGCGGTGAACAGTTTCGTTGGGCACTGCAGACCCGTCCGTTTACCGTTGCCGA
ACCGACCGCATTTGCACGTTATCGTGTTACCGTTACCGCAACCAGCGGTTCTGGTGCACTGAGCCTGGCAG
AAGTTGAACTGCTGGCAGATCCGAAAGAAAGCGGTGCAGAAGAACTGACCCTGTCTGCAGCACCGGATCGT

도면17b

GATGGCGTTACCGGTCGTGAAGTTAGCGGTTCTTTTGCAACCCTGACCGGTGTTGAAGGTGATGTTGCCGC
ACTGGATGTTCAGGTTGCATTTGGTGATGGTAGCGAACCGGTTGCAGGTACACTGCGTGCCGGTGCATTTG
GTGGTTATGCAGTTGATGCAGCACATACCTGGACCGCACCGGGTGTTTATCCGGTTACCGTGACCGTTAGC
GGTGAAGGTATTGAAACCGTTAGCGCAAGCAGCTATGTTAGCGTTAGCCTGCTGCGTGAAGGTTCTCTGCT
GGCAGCATATGATAATGTGTGCATTGGTGATGCAGGTACAACCGTTGGTTCTTGTGATGGTCAGGGCGTTT
TTTTTGATCGTGCACAGCTGGCAGCAAAAGGTTTTGTGCAGGGTGAACGTGCAACCGTTCCGGGTACAGAT
CTGGCATTTGATGTTCCGGCAGTTCCGGCTGGTCAGCCTGATAATGCAACCGGTGATGGTCAGACCATTGA
ACTGGATGTTCCGGCTGATGCAGAACAGCTGAGCGTTATTGGCACCGGCACCGAAAAAAATCAGCAGGCAA
CCGGTACACTGACCTTTGATGATGGTTCTACCCAGCCGATTGATCTGAGCTTTGGTGATTGGAGCGGTGCA
GCACGTAATCCGGTGTTTGGTAATATTCCGGTTGCAGTTACCGATAGCCGTCTGCGTGGTGGTTCTCCGCA
GACCGGTACACCGGCAGCATTTTTTGCCACCGCACCGATTACCCTGCCGGAAGGTAAACGTCCGGTTAGCC
TGACCCTGCCGGATCAGCCTGGTGAACTGAGCCGTGATGGTCGTATTCATGTTGTTGCAGTTGCACATGAT
GGCACCTTTGCAGAACATCCTGCACTGGAAGTGACCGCAGCAGAAGGTGTTACCCTGGCAGTTGGTCAGAC
CTCAGATGTTGCACTGGCACAGGTTGCCGGTGGTCGTGAAGGTGCAGATCTGCGTGCCGCAGTTACCTGGG
GTGATGGTTCTGATGTGGCAGCCGGTGCCGTTACCGATGGTAGCGTTAGCGGTAGCCATGCATATACCGCA
GCAGGCACCTATACCGCATATGTTGTTGTGGATGATGGTTGGACCAGCCAGGTTGTTGAAGTTCCGGTGAC
CGTTACAGAAGCCGAACCGGCACTGGCCGTTGATGTCACCGTTAGCACCCGTTGCCTGGCAGGTAAAGCAT
ATGTTGCAGTGCGTGCAGAAAATGGTGAAGATGTTCCGCTGGCAATTCGTCTGGTTACCCCGTTTGGCACC
AAAGAAGTTGCAGCAGTTGCTCCGGGAGCCAATGCATATCAGAGCTTTGCAACCCGTGTTACCGCAGTTGA
AGCAGGCACCGTTACCGTTGAAGCCACCCGTGGCACCGGTGATGAAGAAGTTACCGCCAGCATTCAGGCAG
ATTATGCAGCCGTTACCTGCGGTTAATAA (SEQ ID NO:19)

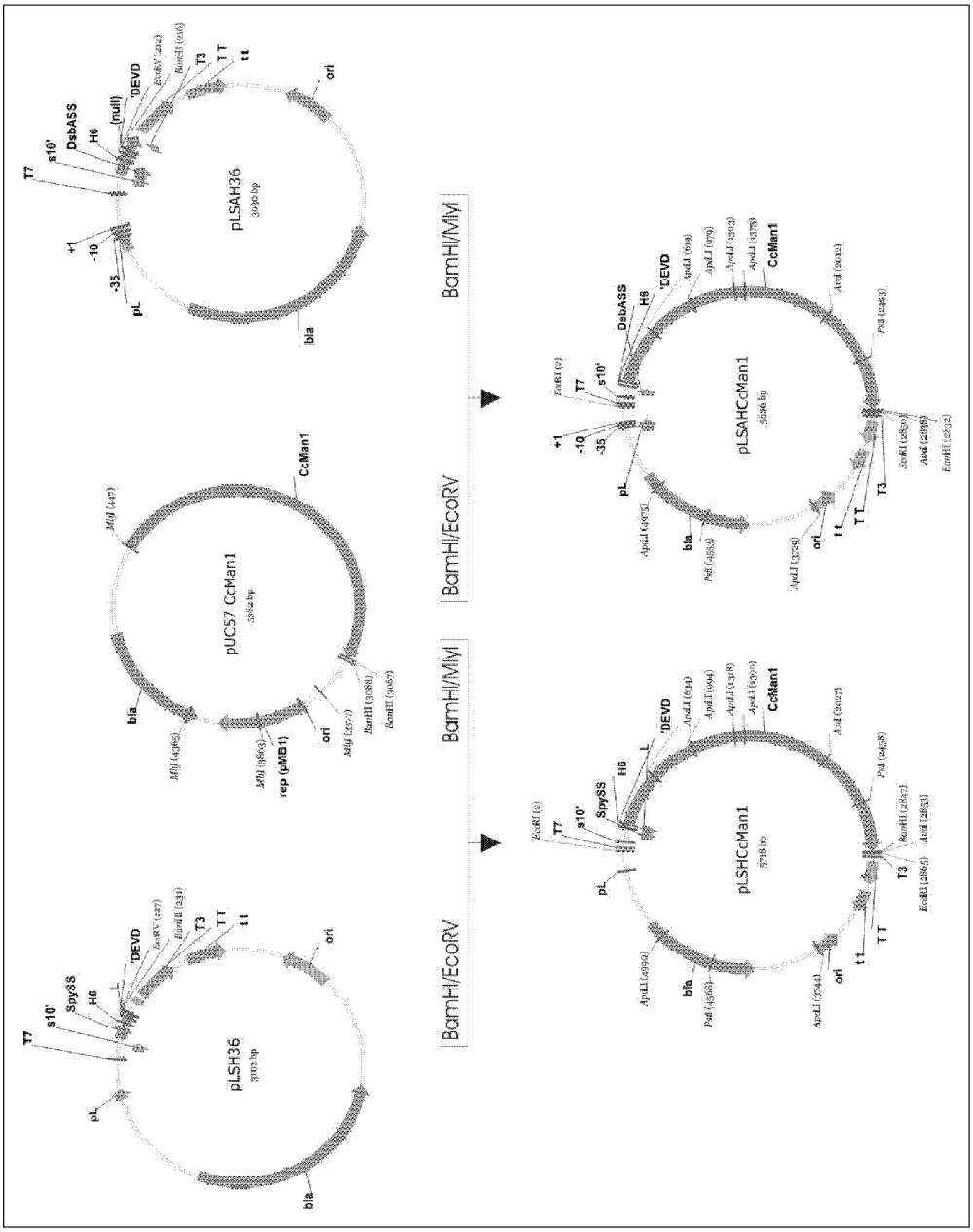
도면18a

GATATGCCTGCAGCACAGGCACCGAATGGTCTGGCAAAAGTTAATCCGCGTACCACACCGGGTCGTAATAA
TACCGGTTATGATTATGCCCAGAGCAAAATTAGCGGTTTTACCCATACCAATCTGGATGGTGTTGGTGGTA
GCGGTGGTGGTGGTGATCTGCTGGTTGTTCCGACCAGCGGTAGCTATACCGCACGTCCGGGTACAGGCACC
TATGCACATCCGTTTAGCCATGATGATGAAGATGCAGGTCCGGGTTTTTATAGCGTTGGTCTGGGTAATGT
TGCAGGCACCGATGGTGCAATTACCGGTGCTCCGGGTACAATTGAAGCAGAAGTTGCAGCAGCAACCCGTA
GCGGTGTTCATCGTTATGCATTTCCGGCAGGTAGCACCCCGAGCCTGGTTGTTGATCTGGAAACCAATAAT
ACCAGCCGTCGTAGCAGCAGCGTTCAGGTTGAAACCCGTGCAGATGGCACCGTTGAACTGAGCGGTCAGGT
TACCGGCTATTTTTATAATGCAGCCTATACCCTGTATTATACCGCACGCACCCTGCAGCCTGCAACCGTTC
AGACCTGGGGTGATGATGATCGTCTGGTTGATGCAACCGCACAGGATGGTGTTGATACCGGTGCAATTCTG
ACCTTTGATCCGGCAGATGCCGGTGAAATTGGTCTGCAGGTTACCCTGTCTCCGGTTAGCGTTGAACAGGC
ACGTATTGATCAGCAGGTTGAACTGGGTGATCTGAGCTTTGATGCAATTCGTGATCGTACCCGTGCAGAAT
GGAATGCAACCCTGGGTCGTGTTGCAATTGATGCAAGCACCGCAACCGATCCGACCGGTGAACTGCAGCGT
CTGTTTTATACCCATCTGTATCGCATGTTTGCAATGCCGATGAATGCAACCAGCACCAGCGGCACCTATCG
TGGTGTTGATGGTGCAGTTCATGCAGCACAGGGCTTTACCTATTATGATAGCTGGGCAACCTGGGATGATT
TTCGCAAATTTAGCGTGATTGCCTATATTGATCCGGCACTGTATCGTGATATGGTTCAGAGCCTGGTTTAC
CTGTTTGCAGATGCAGAAGCAACCGGTACAGGCGGTGGTCTGGGTGGTTTTGTTCATAGCGTTCCGACCGT
TCGTTGGGAACGTAGCAGCGTTGTTGTTGCAGATGCAATTGCCAAAGGCTTTGATGGTTTTGATCGTCTGG
ATGAAGCATATCCGGCACTGCAGCGCCTGGTTGGTCAGTATAGCGCAGATGAACTGCGTCGTGGTTATGTT
GCAGGTAATCCGGGTGCAAGCGTTCAGCGTGGTTATGATCAGTATGGTCTGAGCGTTATTGCCGATGAACT
GGGTCTGACCGAAGAAGCAGAAACCCTGCGCGAACAGGCAAGCTGGCCGATTGAAAAACTGACCAAACCGG
GTGCATGGACCGCAGCAGATGGTACACAGGTTGGTCTGCTGACACCGCGTGCAGCCGATGGTAGCTGGCAG
AGCGCAGATCATGCCAAATTTGAAGCAGCAGGTCTGTATCAGGGCACCCTGTGGCAGTATCATTGGTATGA
TGCCTATGATATGGATGCACTGGTTGAAGCAATGGGTGGTCATGAAGCAGCCCGTCTGGGTATGCGTCATA
TGTTTGGTGAACATGCACCGGATGATGGTAAAGCAATGCTGCATAGCAATGCCAATGAAATTGATCTGCAG
GCACCGTACCTGTTTAATTATACCGGTGAACCGAGCCTGACCCAGAAATGGGCACGTGCAATTTATACCAA
AGAAACCTGGAATCGCTATATTGCAACCGGTAGCAGCTCTGCAGTTCCGTCAGGTGGTGGTGAATTTACAC
CTCCGCTGAAAACCAAAGTTTATCGTCTGGACCCTCGTGGTATGCTGCCGACCATGGATAATGATGCAGGT
ACAATGAGCACCATGTTTGTTGCAGCAGCCGTTGGTCTGTTTCCGGTTACCGCAGGTAGCAGCCAGTTTCA
GGTTGGTAGCCCGTTTTTTGATAGCACCACCATTACCTATGATGATGGTAGCGCATTTACCGTTACCGCAG
ATGGTGTTAGCGAAGATGCCTTTTATGTTCAGAGCGCAACCCTGGATGGTGCAACCTTTGGTAATACCTGG
GTTGATTATGCAACCGTTGTTGGTGGTGCAGATCTGGCATTTCGTATGGGTGAACAGCCGAGCGATTGGGG
CACCGATACCGCACCGGCATTTAGCATGAGCACCGCCACCGATGAACCGGCAGAAGGTCCTCGCGTTAGCG
CAGAACCGACCACCGTGCAGACCGGTGATGGTGGTGCACTGGATGCAACCGTTACCCTGACACTGGATGGC
GCACGTCTGGCAGCACCGGCAGGTACAGATCTGGTTACCAGCGGTGCAGCAAGCGTTGTTGGTCTGCCGGA
TGGTGTTACCGCAGCAGTTACCGTTGCAAGCCCGACCGCACTGACCGTTAGCCTGACCGGCACCGCATCAG
CAGATGCACGTTTTTTTGTGCATCTGCGTGATGCAGCACTGGCCGATGGTGTTGCAGCCGCAAGCCTGCAG
GGTCAGGGTGTTAGCGTTCGTTCTCCGCTGCGTCTGAGCGTTGCAAGCGCAGAACGTGATGCACTGGCAGC
ACTGGTTGATGATGCCGTTCTGGTTCGTCATGGTAATTATAGCAGCGTTACCTTTGATCGTTTTAGCACCG
CTCTGACAAAAGCACAGGAAGCACTGGGCGACGAAGCAGCAACCAGCATTGCACTGCGTTTTGCAGCAGAT
CGTCTGGGTGCAGCAGCAGATGCACTGGATCTGACCGGTGGTGGTTATCGTACCCTGGAAGCAGAACAGAG
CGAAGCATGGTCTGGTGGTGAACTGAAAAATGAAGCCAATAGCAGCAGCGGTAATCTGGGTGGTGTTCGTA
GCGGTAGCTGGGTTCAGTATCGCGATATGACCTTTGAAACCGCAGCCGGTGATACACCTCCGCGTTTTCTG
ACCGTTCGTTATGATACCAGCTTTGCACCGACCGATACCCCGAGCACCGTTCGTGTTCATGCCGGTGATGT
TTCTGGTCCGGTTGTTGCAACCGTTGATCTGAAAGGCACCAGCGGTTGGGGTAAATATACCGAAGTTACCG
CAGAACTGGGTGATGTTCAGGCCCTGGTTGATGCCCAGGTTGTTACCTTTGAACTGCTGGCACCGAGCGGT
CGTAGCTGGGTTGGTAATTTTGATTGGTTTCGCTTTAGCGCAGAAGATCCGGCAGCACCGGGTCAGCCTGG
TGAAAGCCCGACCGTTACCATTGAAGCCGAAGATTGGACCGCAAGCAGCGGTCGTGGTCTGAAAAAAGAAA
GCAGCACCTGGACCAGCGGTCCGGTGACCAATGTTGGTGGTACAGCAGATGGTGATTGGATTGCCTATGGT
GAAGTTGATCTGGGTGAACTGCCGCTGGGCGAACTGAGCGTTCATTATGTGCATAATAGCAATCGCAGCGG
TAATAATAGCGCACTGAGCGTTTATCTGGATGCATTTGATCCGGCTAATCCGGGTGAACCGTTTGTTACCG
TTCCGCTGCCGACCACCGGTAGCAGTTGGACCGCAGATGGCACAGCCACCGTTGTTCTGCCGGAAACCGTG
CAGGGCACCCATGAAGTTTTTGTTCGTCTGAGCACCGAACCGTATGCAGATCATCCGTATGTTGCAAATCT
GGATAGCCTGACCTTTGCACCGGGTGGTCCGACCAGCGTTGTGGTTGAAAGCGAAGCCTGGACCAGCAATT
CTGGTCGTGGCCTGAAAAATGAATCTTCTACCTGGACCTCTGGTCCGGTTACAAATGTGGGTGGCACCGCT

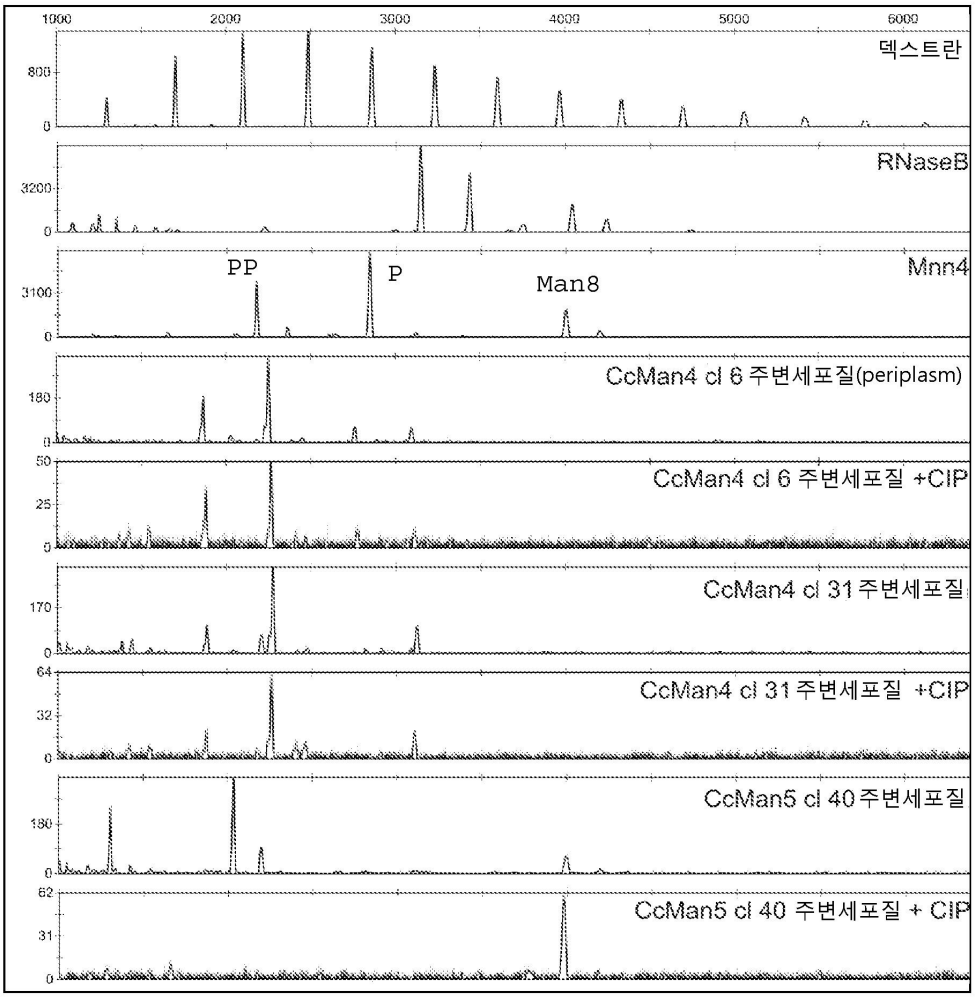
도면18b

GATGGCGATTGGCTGGCATATGGCGAAATTGATCTGGGCAGCGCAGCACTGGATCAGCTGTCTGTGCATTA
TGTTCATAATTCTAATCGCTCTGGTCGTAATTCTGCACTGTCTGTGTATCTGGATGCCTTTGATCCGGCAA
ATCCGGGTGAACCGTTTGTGACAGTGCCGCTGGCAAATACCGGTAGCTCTTGGACCACCGATGGTACTGCA
GTTGTGGATCTGCCGTCTACCGTTCGTGGTAAACATCAGGTTTGGGTTCGTCTGTCTACCGAAGCATATGC
CGATCATCCGTATGTGGCCAATCTGGATTCTATGCGCTTTTTTACCGATGCATATGATGTTGAAGTTCCTC
CGACCGATACAGCAGCACTGGCAGCCGTTGTTGATGCAGCAGGTACACCGGAAGCAGAAATTGCACGTTAT
GGTCGTATTGATGCCCGTGTTTTTACCCGTGAACTGGCAGCAGCACGTAGCGTTCTGGCCGATGCCGGTGC
AACACAGGCACAGGCAGATGAACGTGCTCGTCGTCTGGGTCTGGCAACCGATCAGCTGGTTCCGGCAGAAC
GTCGTCGTCTGGAAAATCTGGTTGCCAGCGCAGAAGCACTGACCGACGAAGGTTATTCTCCGGAAAGCTGG
CAGGCATTTCGTACCGCACTGGCTGCTGCAACCGGCACCCTGGATGATGCAGCAGCATCTGATGAAGCACT
GCATGATGCACGTCTGGCGCTGCAGGGTGCAGTTGATGCACTGGAAGAACCGGCAGATGTTGTTCTGGTTG
AAGTTGAAGTTTCTCCGCGTTGTCTGGCAGGTAAACCGTATGTTGCCGTTCGTGCAGTTAATGTTTCTGAT
GCAGCCGTTGATGTTGAACTGGCAAGCTCTCTGGGCACCCGTAGCTTTGTTGGTGTGGCACCGGGTGCGAG
CGCATATCAGAGCTTTGCAGCCCGTAGCGCAACCGGTGATCTGGATGTTACCGTGACCGCAACCGGTGCAG
ATGGTACTCAGACCGTTGAACAGGTTGTGACCGTTCCGAGCTGTAGCTAATAA (SEQ ID NO:20)

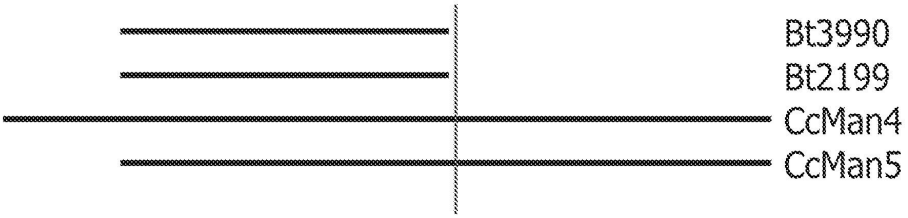
도면19


pLSAH36
pUC57 CcMan1
pLSH36
pLSAHCcMan1
pLSHCcMan1
BamHI/EcoRV
BamHI/MluI
BamHI/EcoRV
BamHI/MluI
T7
s10'
DsbASS
SpySS
H6
DEVD
CcMan1
T3
T7
pL
-35
-10
+1
bla
ori
rep (pMB1)

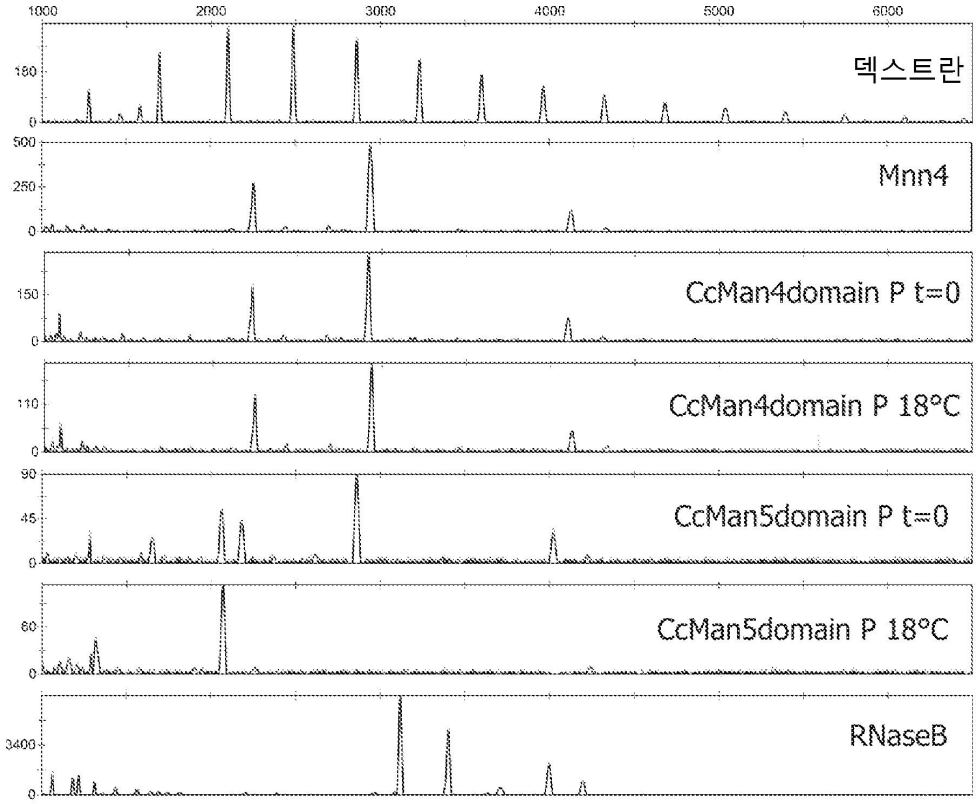
도면20


1000
2000
3000
4000
5000
6000
덱스트란
800
0
RNaseB
3200
0
PP
P
Man8
Mnn4
3100
0
CcMan4 cl 6 주변세포질(periplasm)
180
0
CcMan4 cl 6 주변세포질 +CIP
50
25
0
CcMan4 cl 31 주변세포질
170
0
CcMan4 cl 31 주변세포질 +CIP
64
32
0
CcMan5 cl 40 주변세포질
180
0
CcMan5 cl 40 주변세포질 + CIP
62
31
0

도면21


Bt3990
Bt2199
CcMan4
CcMan5

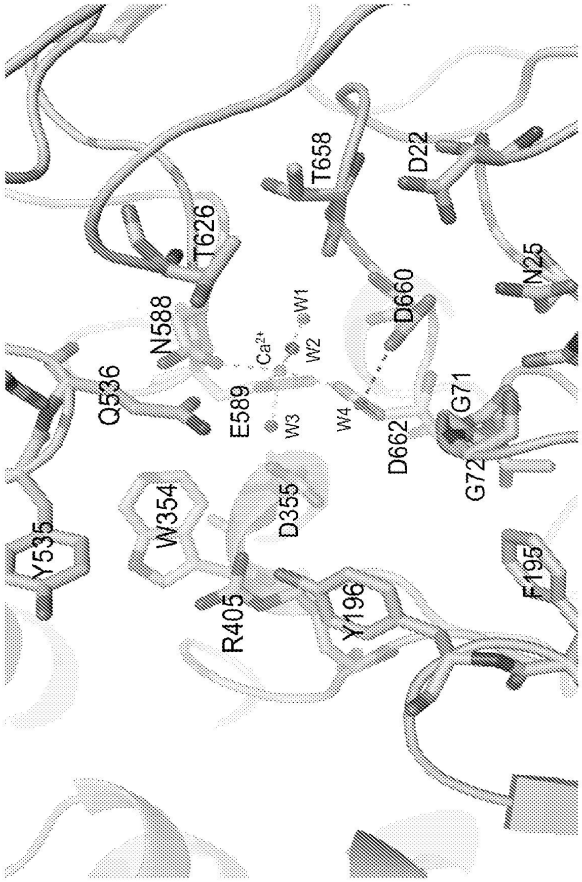
도면22


1000
2000
3000
4000
5000
6000
180
0
덱스트란
500
250
0
Mnn4
150
0
CcMan4domain P t=0
110
0
CcMan4domain P 18°C
90
45
0
CcMan5domain P t=0
60
0
CcMan5domain P 18°C
3400
0
RNaseB

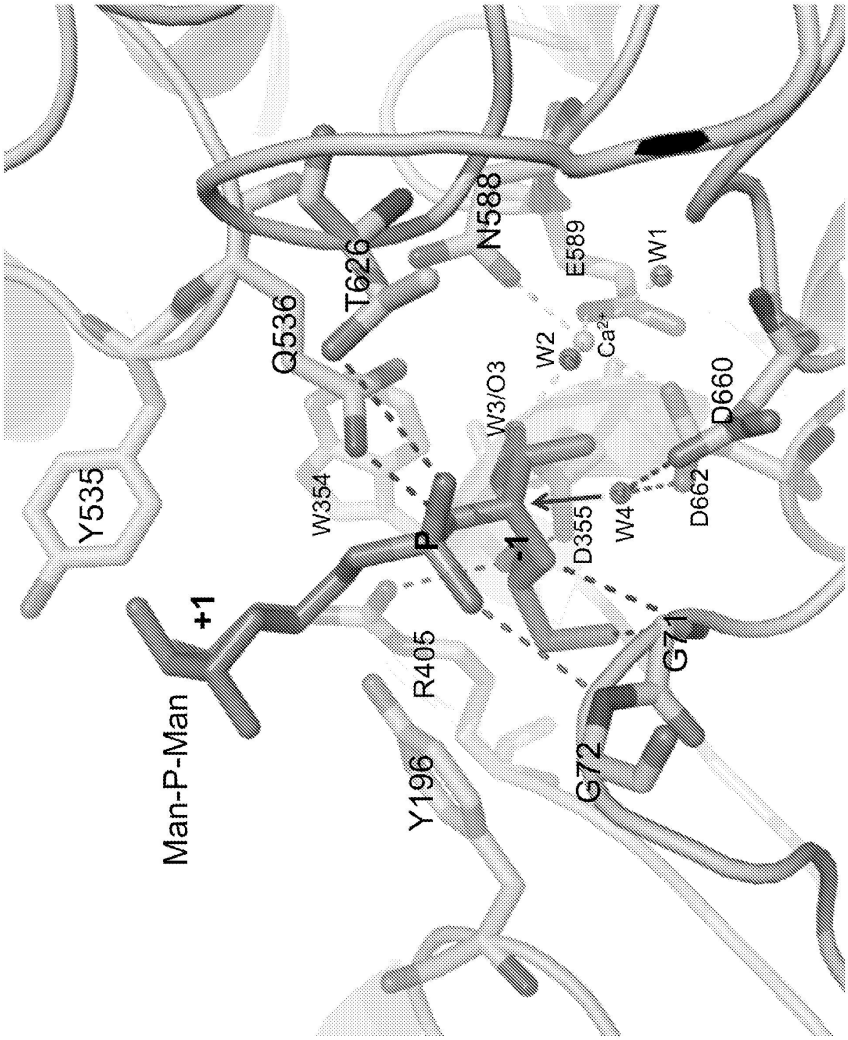
도면23



도면24


T658
D22
T626
N25
N588
W1
D660
Ca2+
W2
Q536
E589
W3
W4
G71
D662
G72
W354
D355
Y535
F195
R405
Y196

도면25



도면26a

26A

ATGAAGCTTTCCACCATCCTCTTCACAGCCTGCGCTACCCTGGCCCTGGACAACGGCCTGGCCCGAACCCC
CACCATGGGCTGGCTGCACTGGGAGCGATTCATGTGTAACCTGGACTGTCAGGAAGAGCCCGACTCTTGTA
TCTCTGAGAAGCTGTTCATGGAAATGGCCGAGCTGATGGTGTCTGAGGGCTGGAAGGACGCCGGCTACGAG
TACCTGTGTATCGACGACTGTTGGATGGCCCCCCAGCGAGACTCTGAGGGCCGACTCCAGGCCGACCCCCA
GCGATTCCCCCACGGCATCCGACAGCTCGCCAACTACGTGCACTCTAAGGGCCTGAAGCTGGGCATCTACG
CCGACGTGGGCAACAAGACCTGTGCCGGCTTCCCCGGCTCTTTCGGCTACTACGACATCGACGCCCAGACC
TTCGCCGACTGGGGCGTGGACCTGCTGAAGTTCGACGGCTGTTACTGTGACTCTCTCGAGAACCTGGCCGA
CGGCTACAAGCACATGTCTCTGGCCCTGAACCGAACCGGCCGATCTATCGTGTACTCTTGTGAGTGGCCCC
TGTACATGTGGCCCTTCCAGAAGCCCAACTACACCGAGATCCGACAGTACTGTAACCACTGGCGAAACTTC
GCCGACATCGACGACTCGTGGAAGTCTATCAAGTCTATTCTGGACTGGACCTCTTTCAACCAGGAGCGAAT
CGTCGACGTCGCCGGACCCGGCGGATGGAACGACCCCGACATGCTGGTGATCGGCAACTTCGGCCTGTCTT
GGAACCAGCAGGTGACCCAGATGGCCCTGTGGGCTATCATGGCTGCCCCCCTGTTCATGTCTAACGACCTG
CGACACATCTCTCCCCAGGCCAAGGCCCTGCTCCAGGACAAGGACGTGATCGCCATCAACCAGGACCCCCT
GGGCAAGCAGGGCTACCAGCTCCGACAGGGCGACAACTTCGAGGTGTGGGAGCGACCCCTGTCTGGCCTGG
CCTGGGCCGTGGCCATGATCAACCGACAGGAGATCGGCGGACCCCGATCTTACACCATCGCCGTGGCCTCC
CTGGGAAAGGGCGTGGCCTGTAACCCCGCCTGTTTCATCACCCAGCTCCTGCCCGTGAAGCGAAAGCTGGG
ATTCTACGAGTGGACCTCTCGACTGCGATCTCACATCAACCCCACCGGCACCGTGCTGCTCCAGCTCGAGA
ACACCATGCAGATGTCTCTGAAGGACCTGCTGACGCGTGAACAAAAACTCATCTCAGAAGAGGATCTGAAT
AGCGCCGTCGACCATCATCATCATCATCAT (SEQ ID NO:22)

도면26b

26B

MKLSTILFTACATLALDNGLARTPTMGWLHWERFMCNLDCQEEPDSCISEKLFMEMAELMVSEGWKDAGYE
YLCIDDCWMAPQRDSEGRLQADPQRFPHGIRQLANYVHSKGLKLGIYADVGNKTCAGFPGSFGYYDIDAQT
FADWGVDLLKFDGCYCDSLENLADGYKHMSLALNRTGRSIVYSCEWPLYMWPFQKPNYTEIRQYCNHWRNF
ADIDDSWKSIKSILDWTSFNQERIVDVAGPGGWNDPDMLVIGNFGLSWNQQVTQMALWAIMAAPLFMSNDL
RHISPQAKALLQDKDVIAINQDPLGKQGYQLRQGDNFEVWERPLSGLAWAVAMINRQEIGGPRSYTIAVAS
LGKGVACNPACFITQLLPVKRKLGFYEWTSRLRSHINPTGTVLLQLENTMQMSLKDLLTREQKLISEEDLN
SAVDHHHHHH (SEQ ID NO:23)

도면27a

27A

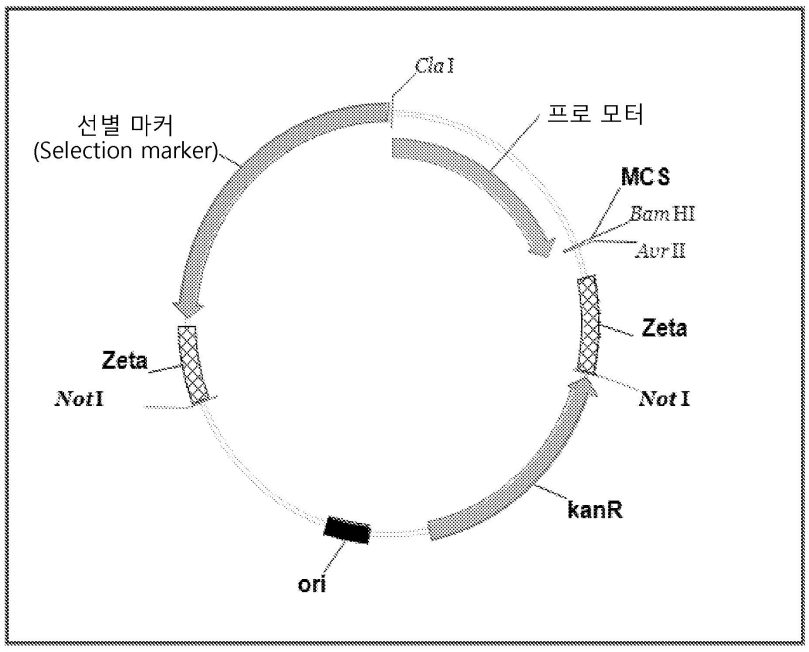
ATGAAGCTTTCCACCATCCTCTTCACAGCCTGCGCTACCCTGGCTGCCGCCCAGCAGGGAGCCT
CTCGACCCGGACCCCGAGATGCCCAGGCTCACCCCGGACGACCTCGAGCTGTGCCCACCCAGTGTG
ACGTGCCCCCCAACTCTCGATTCGACTGTGCCCCCGACAAGGCCATCACCCAGGAGCAGTGCGAGG
CCCGAGGCTGTTGTTACATCCCCGCTAAGCAGGGCCTGCAGGGCGCTCAGATGGGCCAGCCCTGGT
GTTTCTTCCCCCCCTCTTACCCCTCCTACAAGCTGGAGAACCTGTCCTCTTCGGAGATGGGCTACAC
CGCCACCCTGACCCGAACCACCCCCACCTTTTTCCCCAAGGACATCCTGACCCTGCGACTGGACGTG
ATGATGGAGACCGAGAACCGACTGCACTTCACCATCAAGGACCCCGCCAACCGACGATACGAGGT
GCCCCTGGAGACCCCCCACGTGCACTCTCGAGCCCCTTCCCCCCTGTACTCTGTGGAGTTCTCTGAG
GAGCCCTTCGGCGTGATCGTGCGACGACAGCTGGACGGCCGAGTGCTGCTGAACACCACCGTGGCC
CCCCTGTTCTTCGCCGACCAGTTCCTGCAGCTGTCTACCTCTCTGCCCTCTCAGTACATCACCGGCCT
GGCCGAGCACCTGTCCCCCCTGATGCTGTCCACCTCTTGGACTCGAATCACCCTGTGGAACCGAGA
CCTGGCCCCCACCCCCGGTGCCAACCTGTACGGCTCTCACCCCTTCTACCTGGCCCTGGAGGACGGC
GGCTCTGCCCACGGCGTGTTTCTGCTGAACTCTAACGCCATGGACGTGGTGCTGCAGCCCTCTCCCG
CCCTGTCTTGGCGATCTACCGGCGGCATCCTGGACGTGTACATCTTCCTGGGCCCTGAGCCCAAGTC
TGTGGTCCAGCAGTACCTGGACGTGGTCGGATACCCCTTCATGCCCCCCTACTGGGGCCTGGGCTTC
CACCTGTGTCGATGGGGCTACTCTTCTACCGCCATCACCCGACAGGTGGTGGAGAACATGACCCGA
GCCCACTTCCCCCTGGACGTGCAATGGAACGACCTGGACTACATGGACTCTCGACGAGACTTCACC
TTCAACAAGGACGGCTTCCGAGACTTCCCCGCCATGGTCCAGGAGCTGCACCAGGGAGGACGACG
ATACATGATGATCGTGGACCCCGCCATCTCTTCTTCCGGACCCGCCGGATCTTACCGACCCTACGAC
GAGGGCCTGCGACGAGGCGTGTTCATCACCAACGAGACCGGCCAGCCCCTGATCGGCAAGGTGTG
GCCCGGCTCTACCGCCTTCCCCGACTTCACCAACCCCACCGCCCTGGCTTGGTGGGAGGACATGGT
GGCCGAGTTCCACGACCAGGTGCCCTTCGACGGCATGTGGATCGACATGAACGAGCCCTCTAACTT
CATCCGAGGCTCTGAGGACGGCTGTCCCAACAACGAGCTGGAGAACCCCCCCCTACGTGCCCGGCGT
GGTGGGCGGAACCCTGCAGGCCGCCACCATCTGTGCCTCTTCGCACCAGTTTCTGTCTACCCACTAC
AACCTGCACAACCTGTACGGACTGACCGAGGCCATTGCCTCTCACCGAGCCCTGGTGAAGGCCCGA
GGCACCCGACCCTTCGTGATCTCTCGATCTACCTTCGCCGGCCACGGCCGATACGCCGGACACTGG
ACCGGCGATGTGTGGTCCTCTTGGGAGCAGCTGGCCTCTTCTGTGCCCGAGATCCTGCAGTTCAACC
TGCTGGGCGTGCCCCTGGTGGGCGCCGACGTGTGTGGCTTCCTGGGCAACACCTCTGAGGAGCTGT
GTGTTCGATGGACCCAGCTCGGCGCCTTCTACCCTTTCATGCGAAACCACAACTCCCTGCTGTCTCT
GCCCCAGGAGCCCTACTCGTTCTCTGAGCCCGCTCAGCAGGCCATGCGAAAGGCTCTGACCCTGCG
ATACGCCCTGCTGCCCCACCTGTACACCCTGTTCCACCAGGCCCACGTGGCTGGAGAGACCGTGGC
CCGACCCCTGTTCCTGGAGTTCCCTAAGGACTCTTCTACCTGGACCGTGGACCATCAGCTGCTGTGG
GGCGAGGCCCTCCTGATCACCCCCGTGCTGCAGGCCGGCAAGGCTGAGGTGACCGGCTACTTCCCT
CTGGGCACCTGGTACGACCTGCAGACCGTGCCTGTGGAGGCCCTGGGATCTCTGCCCCCTCCTCCCG
CCGCTCCCCGAGAGCCCGCCATCCACTCTGAGGGCCAGTGGGTGACCCTGCCCGCTCCCCTGGACA
CCATCAACGTGCACCTGCGAGCCGGCTACATCATCCCTCTGCAGGGACCCGGCCTGACCACCACCG
AGTCTCGACAGCAGCCCATGGCCCTGGCCGTGGCTCTGACCAAGGGCGGAGAGGCCCGAGGCGAG
CTGTTCTGGGACGATGGCGAGTCTCTGGAGGTGCTGGAGCGAGGCGCCTACACCCAGGTGATCTTT
CTGGCCCGAAACAACACCATCGTGAACGAGCTGGTGCGAGTGACCTCTGAGGGCGCTGGTCTGCAG
CTCCAGAAGGTGACCGTCCTGGGCGTGGCCACCGCTCCCCAGCAGGTCCTGTCTAACGGCGTGCCC
GTGTCTAACTTCACCTACTCTCCCGACACCAAGGTGCTGGACATCTGTGTGTCTCTGCTGATGGGCG
AGCAGTTCCTGGTGTCTTGGTGTTAAC (SEQ ID NO:24)

도면27b

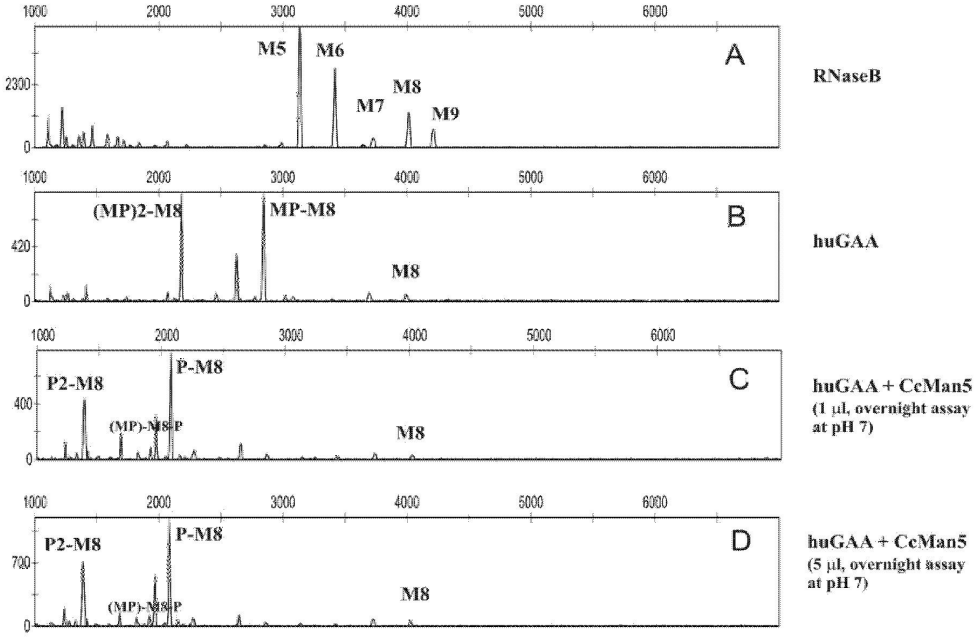
27B

MKLSTILFTACATLAAAQQGASRPGPRDAQAHPGRPRAVPTQCDVPPNSRFDCAPDK
AITQEQCEARGCCYIPAKQGLQGAQMGQPWCFFPPSYPSYKLENLSSSEMGYTATLTRT
TPTFFPKDILTLRLDVMMETENRLHFTIKDPANRRYEVPLETPHVHSRAPSPLYSVEFSE
EPFGVIVRRQLDGRVLLNTTVAPLFFADQFLQLSTSLPSQYITGLAEHLSPLMLSTSWTR
ITLWNRDLAPTPGANLYGSHPFYLALEDGGSAHGVFLLNSNAMDVVLQPSPALSWRST
GGILDVYIFLGPEPKSVVQQYLDVVGYPFMPPYWGLGFHLCRWGYSSTAITRQVVENM
TRAHFPLDVQWNDLDYMDSRRDFTFNKDGFRDFPAMVQELHQGGRRYMMIVDPAISS
SGPAGSYRPYDEGLRRGVFITNETGQPLIGKVWPGSTAFPDFTNPTALAWWEDMVAEF
HDQVPFDGMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVVGGTLQAATICASSHQF
LSTHYNLHNLYGLTEAIASHRALVKARGTRPFVISRSTFAGHGRYAGHWTGDVWSSWE
QLASSVPEILQFNLLGVPLVGADVCGFLGNTSEELCVRWTQLGAFYPFMRNHNSLLSLP
QEPYSFSEPAQQAMRKALTLRYALLPHLYTLFHQAHVAGETVARPLFLEFPKDSSTWT
VDHQLLWGEALLITPVLQAGKAEVTGYFPLGTWYDLQTVPVEALGSLPPPPAAPREPAI
HSEGQWVTLPAPLDTINVHLRAGYIIPLQGPGLTTTESRQQPMALAVALTKGGEARGEL
FWDDGESLEVLERGAYTQVIFLARNNTIVNELVRVTSEGAGLQLQKVTVLGVATAPQQ
VLSNGVPVSNFTYSPDTKVLDICVSLLMGEQFLVSWC* (SEQ ID NO:25)

도면28

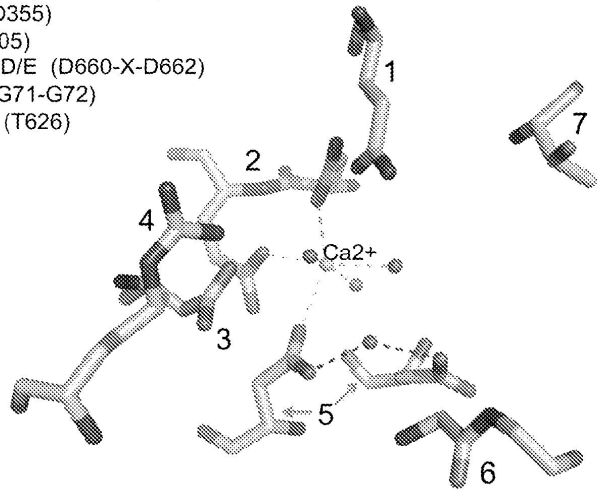


도면29



도면30

1: Q (Q536)
2: N/D-E/Q (N588-Q589)
3: D/E (D355)
4: R (R405)
5: D/E-X-D/E (D660-X-D662)
6: G-G (G71-G72)
7: T/S/G (T626)

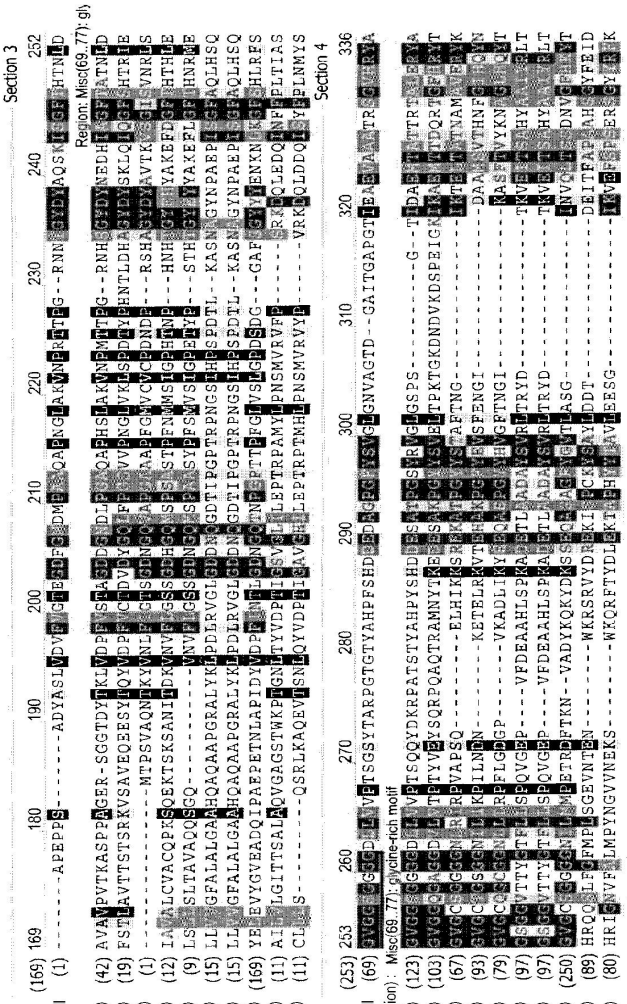


SEQ ID NO:50에서의 등가의 잔기(equivalent residiues)에 대한 넘버링(numbering)을 괄호(parenthesis)에 나타내었다.

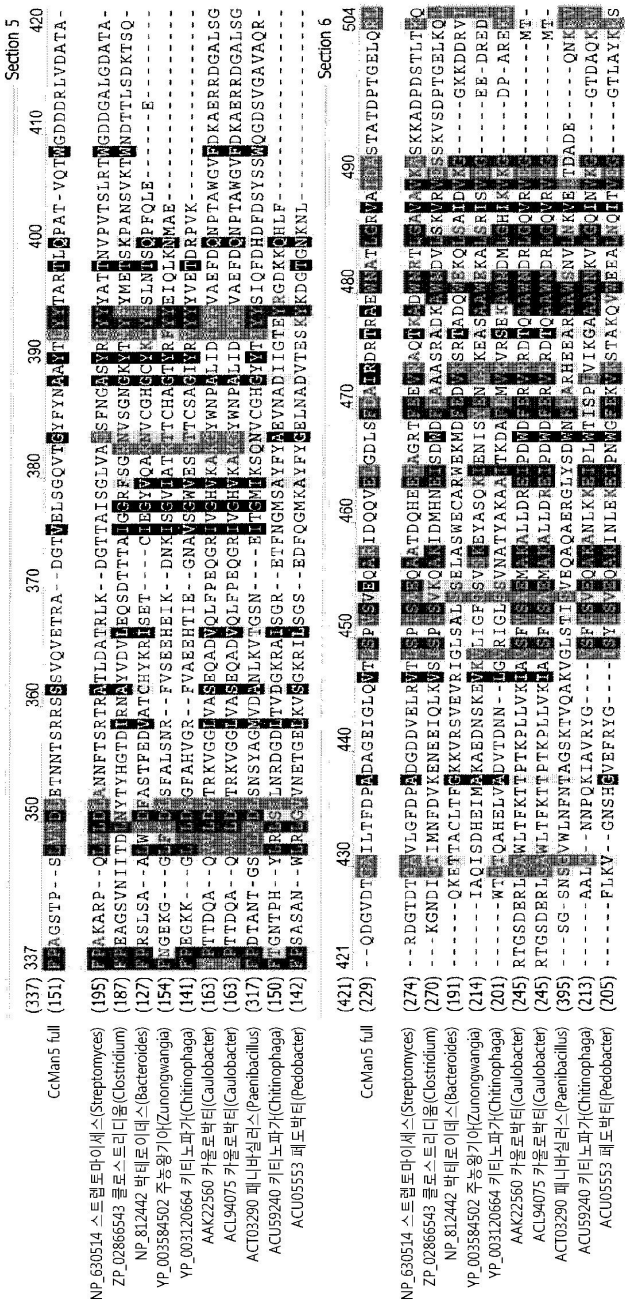
도면31a

Section 1
CcMan5 full (1) 1 10 20 30 40 50 60 70 84
(1)
NP_630514 스트렙토마이세스(Streptomyces) (1)
ZP_02866543 클로스트리디움(Clostridium) (1)
NP_812442 박테로이데스(Bacteroides) (1)
YP_003584502 주눙왕기아(Zunongwangia) (1)
YP_003120664 키티노파가(Chitinophaga) (1)
AAK22560 카울로박터(Caulobacter) (1)
ACL94075 카울로박터(Caulobacter) (1)
ACT03290 페니바실러스(Paenibacillus) (1) MFKKLFAVAMTVMCLTGILVPVGSNAAFAAAAAEGITTRNVAINAAAATASGQCNANESASNAVDGKTDTKWCDNTSAQKKWLKLD
ACU59240 키티노파가(Chitinophaga) (1)
ACU05553 페도박터(Pedobacter) (1)
Section 2
(85) 85 90 100 110 120 130 140 150 168
CcMan5 full (1)
NP_630514 스트렙토마이세스(Streptomyces) (1) ----------MPDRSKRPPIRSSSPRAALRATVAAVLAGALGLAALTGGGT
ZP_02866543 클로스트리디움(Clostridium) (1) ----------MNKKINRLLKGALAAFTVA
NP_812442 박테로이데스(Bacteroides) (1) ----------MTLIMKRMIAA
YP_003584502 주눙왕기아(Zunongwangia) (1) ----------MKKSLIYL
YP_003120664 키티노파가(Chitinophaga) (1) ----------MNPGKGPVLRFRVQ
AAK22560 카울로박터(Caulobacter) (1) ----------MNPGKGPVLRFRVQ
ACL94075 카울로박터(Caulobacter) (1) ----------MNPGKGPVLRFRVQ
ACT03290 페니바실러스(Paenibacillus) (85) LGKEYLVNEWVLQNAAINESGNSPFWNTKNFRLQKSDDGETWTDVDIVTNNAQTIVDRFVTPFTTRYLRLYIDKAAYDSNIARI
ACU59240 키티노파가(Chitinophaga) (1) ----------MKRNRYLIST
ACU05553 페도박터(Pedobacter) (1) ----------MKNSIKLMLL

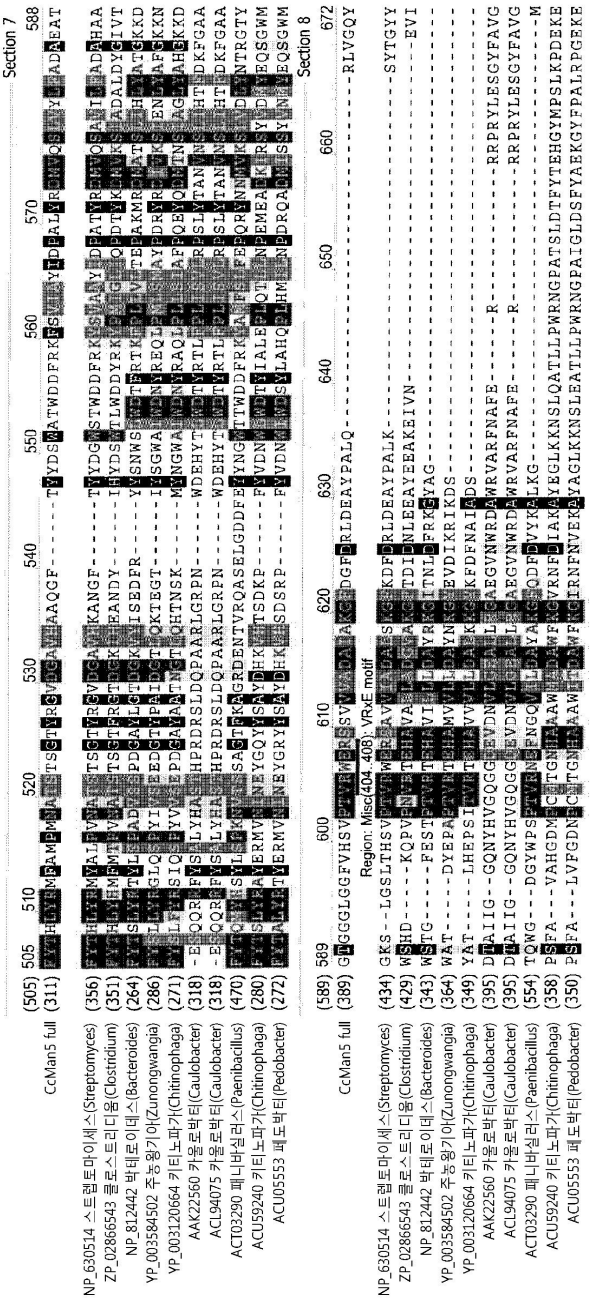
도면31b


Section 3
Section 4
Region: Misc(69..77): gly
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주능왕기아(Zunongwangia)
YP_003120664 카티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 카티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)
영역(Region) : Misc(69..77): glycine-rich motif

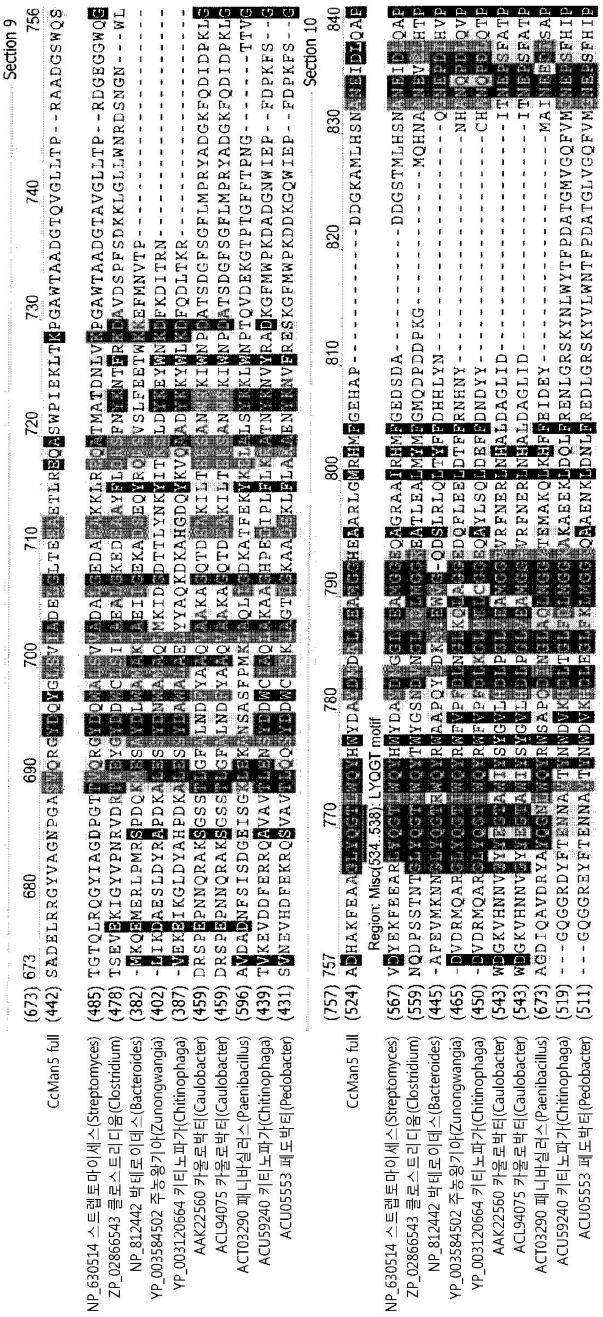
도면31c



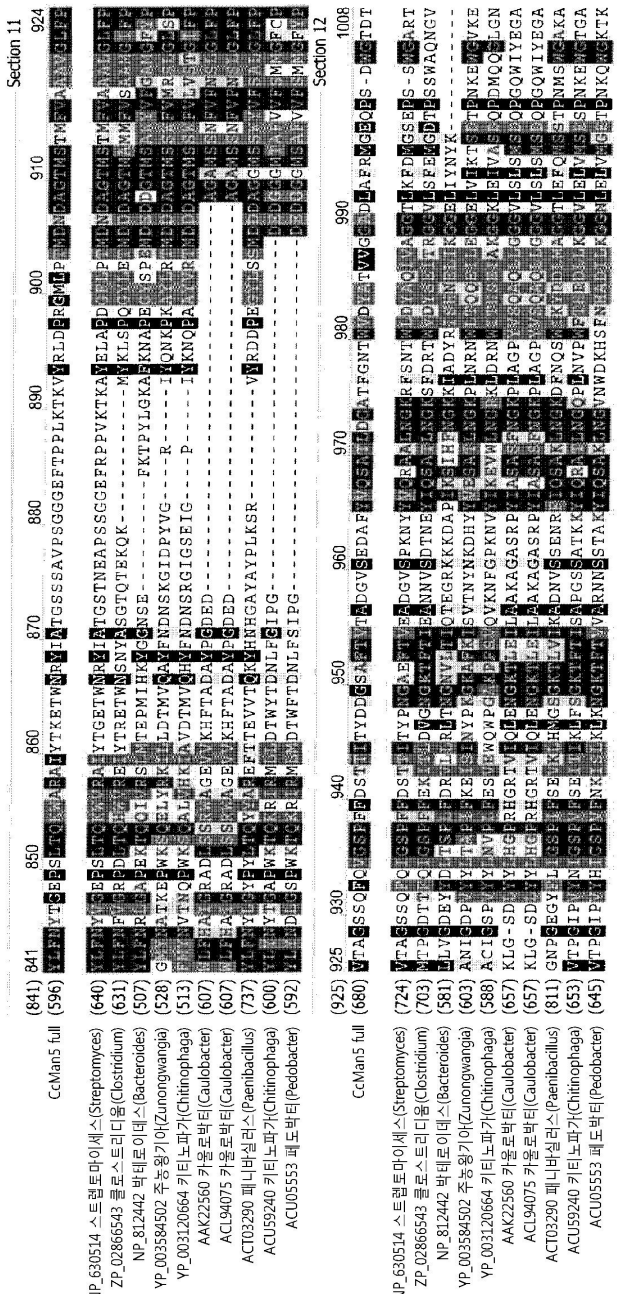
도면31d


Section 7
Section 8
Region: Misc(404..408) VRxE motif
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주능왕기아(Zunongwangia)
YP_003120664 키티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 키티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)
CcMan5 full

도면31e

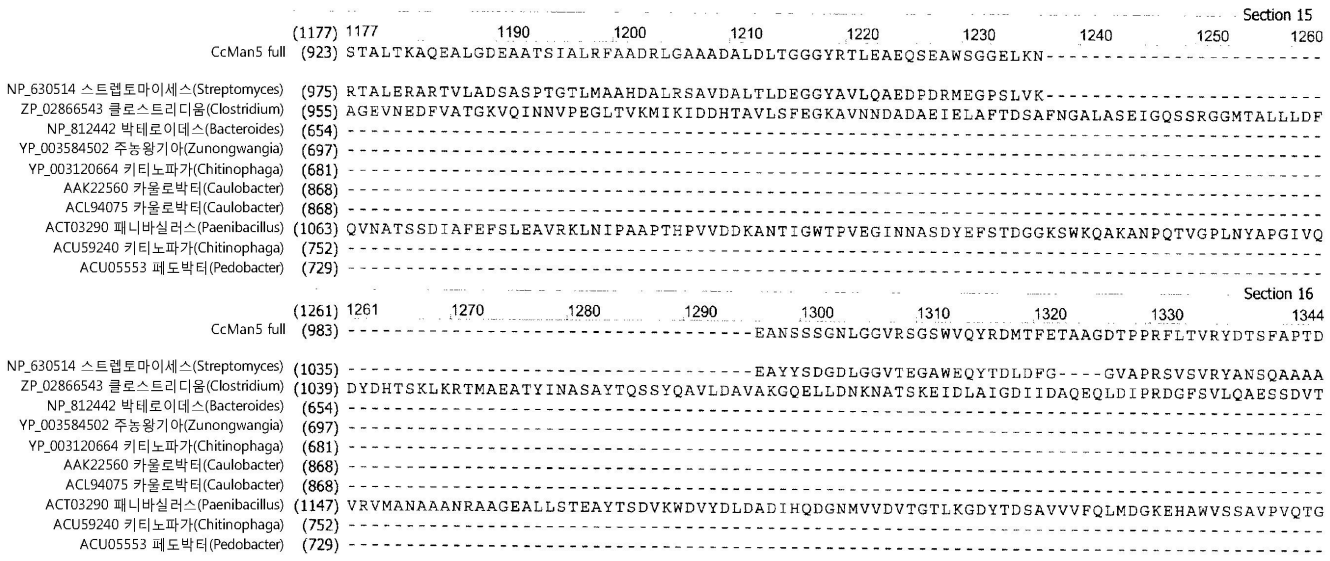

Section 9
Section 10
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주농잉기아(Zunongwangia)
YP_003120664 키티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 키티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)
Region: Misc(534..538): LYQGT motif

도면31f


Section 11
Section 12
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주능왕기아(Zunongwangia)
YP_003120664 키티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 키티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)

도면31g

Section 13
(1009) 1009 1020 1030 1040 1050 1060 1070 1080 1092
CcMan5 full (763) APAFSMSTATDEPAEGPR------VSAEPTTVQTGDGGALDATVTLTLDG-ARLAAPAGTDLVTSGAASVVGLPDGVTAAVTV
NP_630514 스트렙토마이세스(Streptomyces) (807) EPAYSLNTDSGDGDDEHAPGRGTTVVSARPETVRTAADGTVDASVELRLSGRASFAARKGTSLTPTGAASVTGLPDGVTADLRV
ZP_02866543 클로스트리디움(Clostridium) (787) TAKSSSDNADTSTYDDDEIAYSSAMPEESKANDGSFDQKITITLKTKEPAGEIGEDLVATGKINITNIPEGLEASAIKTEANKV
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주동왕기아(Zunongwangia) (687) AWVSSIRQYL----------
YP_003120664 키티노파가(Chitinophaga) (672) KWIADITRQ----------
AAK22560 카울로박터(Caulobacter) (740) VLTVRADQPSLVDGKTSTGWRAASGQSVTFSLKAPACIAAYSVSVGPDQADPSHWTLQAYDGRAWVSVDQQSNVVFDHRHATRT
ACL94075 카울로박터(Caulobacter) (740) VLTVRADQPSLVDGKTSTGWRAASGQSVTFSLKAPACIAAYSVSVGPDQADPSHWTLQAYDGRAWVSVDQQSNVVFDHRHATRT
ACT03290 페니바실러스(Paenibacillus) (895) SAAEPTVDYNADMDNDFNHEQLIPEKSTWKYDDKGKEAGEGWTQVDFDDSSWSSGKAMLGYDSYGKPATTVSYGPNANNKYVTT
ACU59240 키티노파가(Chitinophaga) (737) QAAEPPSSLNYSPAGK----------
ACU05553 페도박터(Pedobacter) (729) ----------
Section 14
(1093) 1093 1100 1110 1120 1130 1140 1150 1160 1176
CcMan5 full (839) ASPTALTVSLTGTASADAREFVHLRDAALADGVAAASLQGQGVSVRSPLRLSVASAERDALAALVDDAVLVRHGNYSSVTFDRF
NP_630514 스트렙토마이세스(Streptomyces) (891) TGKRTASLRLTGTTRTDARFGITFRDRAFPHGIPASTVTGTGVSVTDPLIVSAAAVHRGSLAALVDEASLVREGNYSDGSYGIF
ZP_02866543 클로스트리디움(Clostridium) (871) EVSLNGKAKNHTLNDSISNLTIEITDGATNEPIKDSIRKTKDNVKVMFIDNQLTYSQSEFKESESDDGAILETSTITLTGDTTF
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주동왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (824) FPLAPGRYARLRWVLDGGSEASVSEVELIAGASCAAPTSGAPLL----------
ACL94075 카울로박터(Caulobacter) (824) FPLAPGRYARLRWVLDGGSEASVSEVELIAGASCAAPTSGAPLL----------
ACT03290 페니바실러스(Paenibacillus) (979) YFRKTFDAKDLDGILELDGSLIRDDGAIVYLNGHEIFRTNMPTGAVNYSTFANATVGDERDKNGFIIDPSYLVEGKNVLTAEVH
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------


Section 15
(1177) 1177 1190 1200 1210 1220 1230 1240 1250 1260
CcMan5 full (923) STALTKAQEALGDEAATSIALRFAADRLGAAADALDLTGGGYRTLEAEQSEAWSGGELKN----------
NP_630514 스트렙토마이세스(Streptomyces) (975) RTALERARTVLADSASPTGTLMAAHDALRSAVDALTLDEGGYAVLQAEDPDRMEGPSLVK----------
ZP_02866543 클로스트리디움(Clostridium) (955) AGEVNEDFVATGKVQINNVPEGLTVKMIKIDDHTAVLSFEGKAVNNDADAEIELAFTDSAFNGALASEIGQSSRGGMTALLLDF
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 패니바실러스(Paenibacillus) (1063) QVNATSSDIAFEFSLEAVRKLNIPAAPTHPVVDDKANTIGWTPVEGINNASDYEFSTDGGKSWKQAKANPQTVGPLNYAPGIVQ
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------
Section 16
(1261) 1261 1270 1280 1290 1300 1310 1320 1330 1344
CcMan5 full (983) ----------EANSSSGNLGGVRSGSWVQYRDMTFETAAGDTPPRFLTVRYDTSFAPTD
NP_630514 스트렙토마이세스(Streptomyces) (1035) ----------EAYYSDGDLGGVTEGAWEQYTDLDFG----GVAPRSVSVRYANSQAAAA
ZP_02866543 클로스트리디움(Clostridium) (1039) DYDHTSKLKRTMAEATYINASAYTQSSYQAVLDAVAKGQELLDNKNATSKEIDLAIGDIIDAQEQLDIPRDGFSVLQAESSDVT
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 패니바실러스(Paenibacillus) (1147) VRVMANAAANRAAGEALLSTEAYTSDVKWDVYDLDADIHQDGNMVVDVTGTLKGDYTDSAVVVFQLMDGKEHAWVSSAVPVQTG
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------

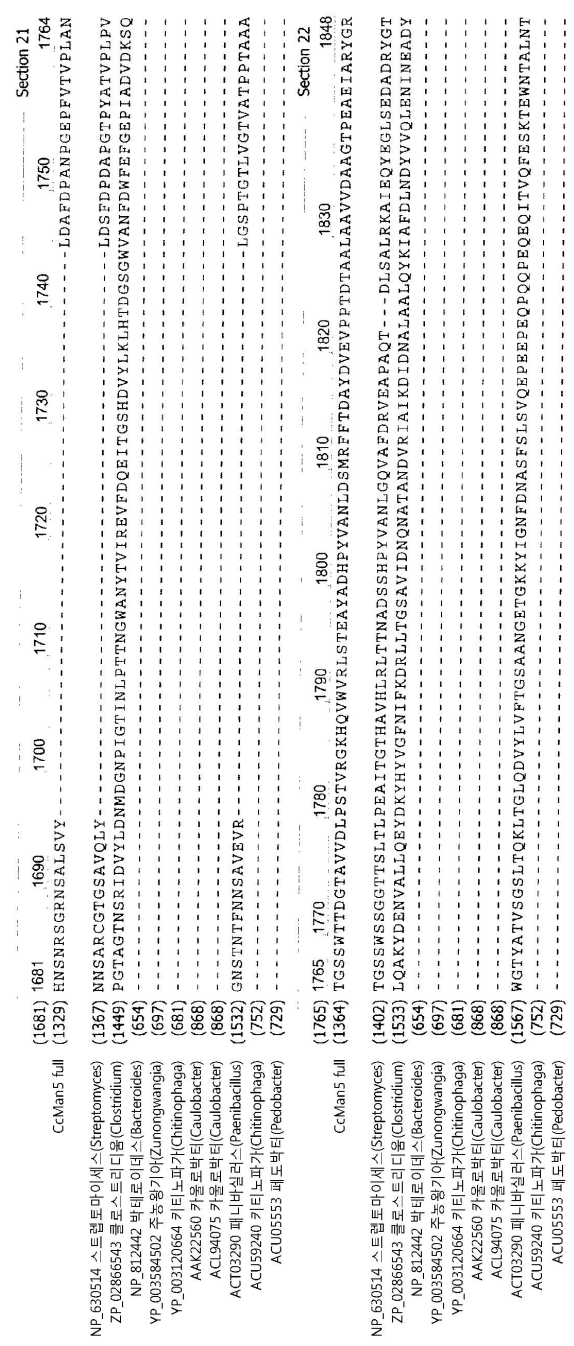
도면31i

Section 17
Section 18
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
YP_003120664 키티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 키티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)

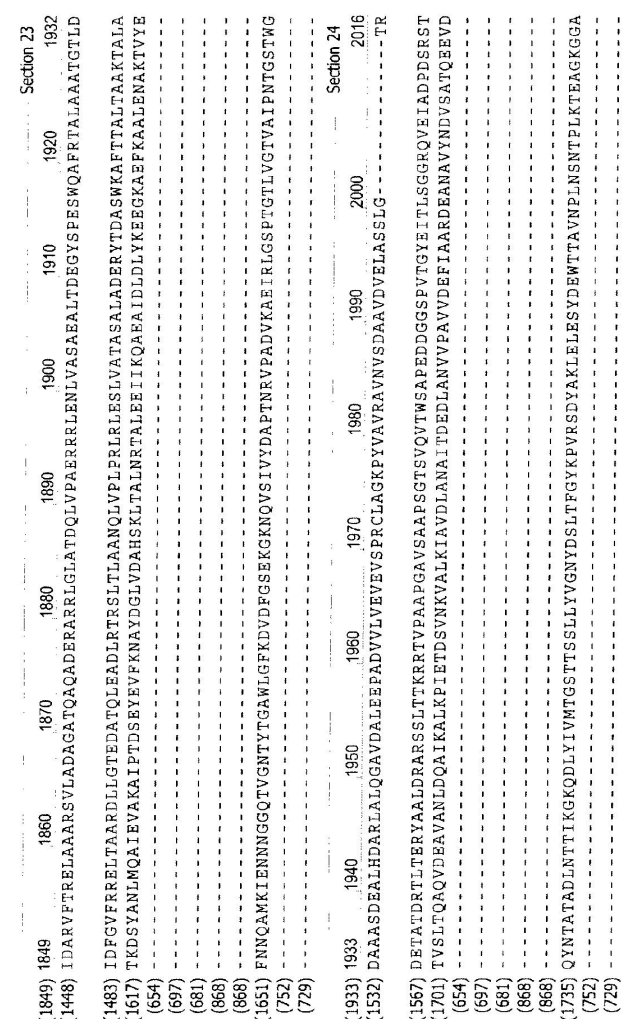
도면31j

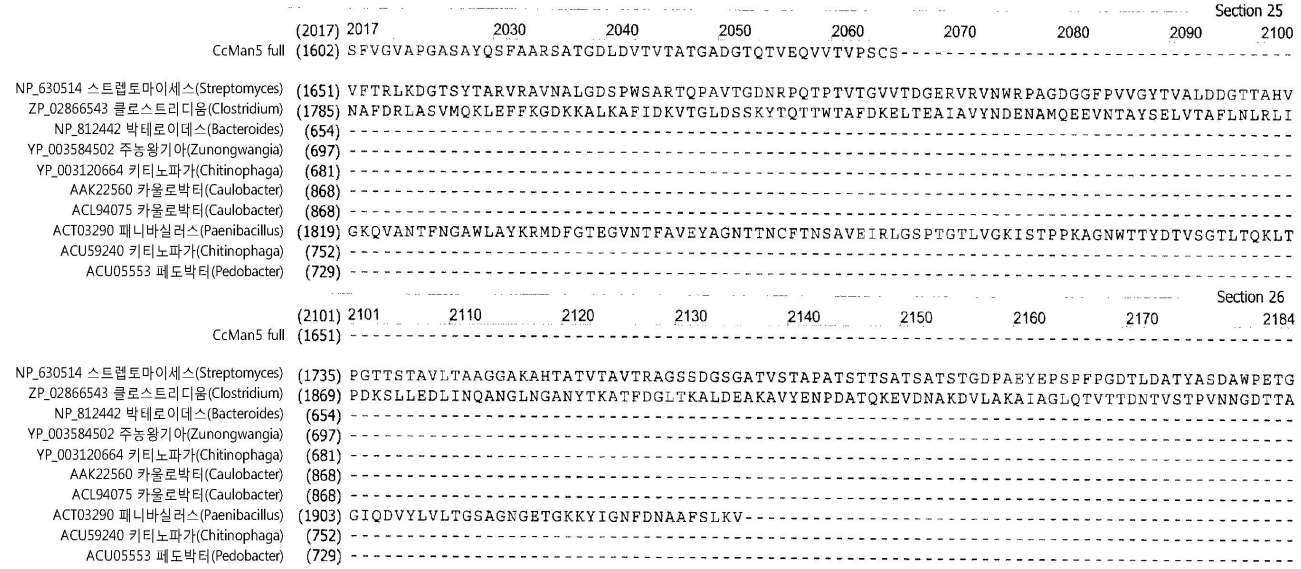
(1513) 1513 1520 1530 1540 1550 1560 1570 1580 Section 19 1596
CcMan5 full (1165) LGELPLGELSVHYVHNSNRSGNNSALSVYLDAFDPANPGEPFVTVPLPTTGSSWTADGTATVVLPETVQG--THEVFVRLSTEP
NP_630514 스트렙토마이세스(Streptomyces) (1205) LGDGADTVTVSYDKPRSR-----AASDSHIELRPGAKDGPTAVTVPLDYTGSGWGTVASTSVRLDPDVFEGTQDVYAVFVSSTQ
ZP_02866543 클로스트리디움(Clostridium) (1291) FNNLEADTINVHYSTRVDACALDARIEIRKDNKDGELLGTIMLPLTGGWSDYQTVSTKLDTSVTGVQDICFVLRGTNDGG---
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 페니바실러스(Paenibacillus) (1364) FGTTGYNNVTVQYDAPTDKVPAGSKLEFRLGSVSGELVGTVNMEDKNAGWGSYITTKANLTRTLTGQQKLYVVMVAGTPNNLPY
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------
(1597) 1597 1610 1620 1630 1640 1650 1660 1670 Section 20 1680
CcMan5 full (1247) YADHPYVANLDSLTPAPGGPTSVVVESEAWTSNSGRGLKNESSTWT--SGPVTNVGGTADGDWLAYGEIDLGSAALDQLSVHYV
NP_630514 스트렙토마이세스(Streptomyces) (1284) TDAQPYVANVHSLTLTRQADAPVVFDATAFEGSSGGGLKSEPATWSG-AGSATSLGGTYDGAWLDYGDVDFGDSPKNTVTLTYV
ZP_02866543 클로스트리디움(Clostridium) (1371) ---RP---YVANIDYMEFVNSGVNHIEAENKDDWSGAELKVENSTDNTGKSLTNIGGARNDAWLRYNGVEFNGKTEMTVRYSHN
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 페니바실러스(Paenibacillus) (1448) IGNFDWFKFDYEKIRSDYAKLELESYDEWTTDVNTGNNNTPLKTEAGKGGVGQQVANTFNGAWLAYKRMDFGSEGVDKFSIEYA
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------

도면31k

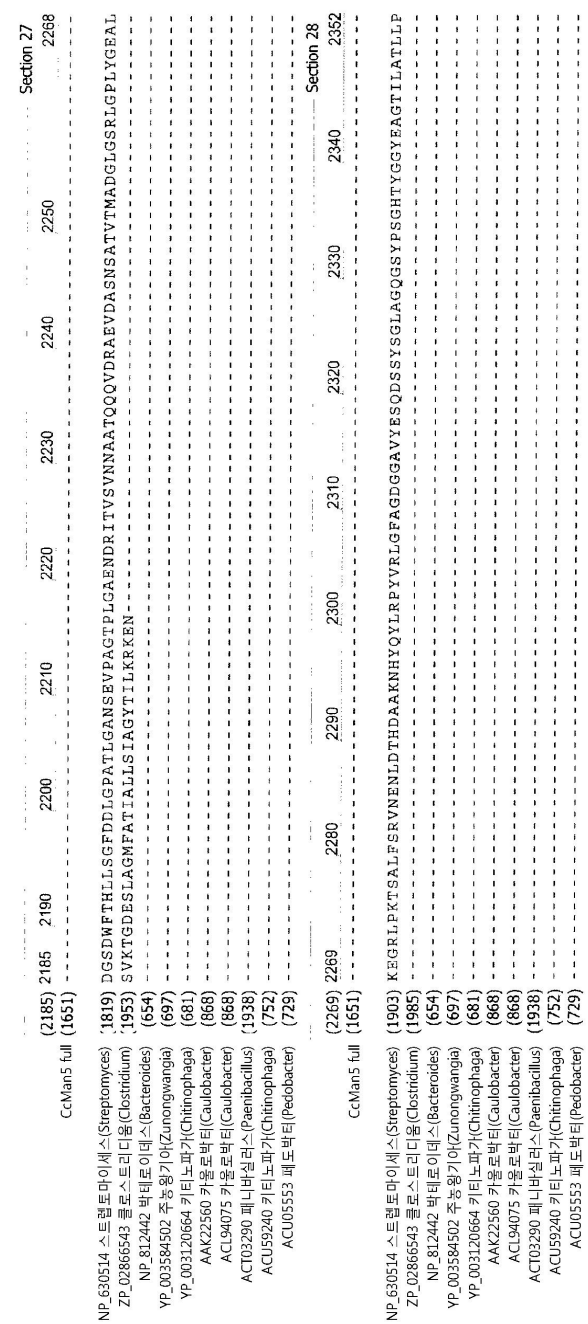

Section 21
Section 22
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주농엉기아(Zunongwangia)
YP_003120664 키티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 키티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)

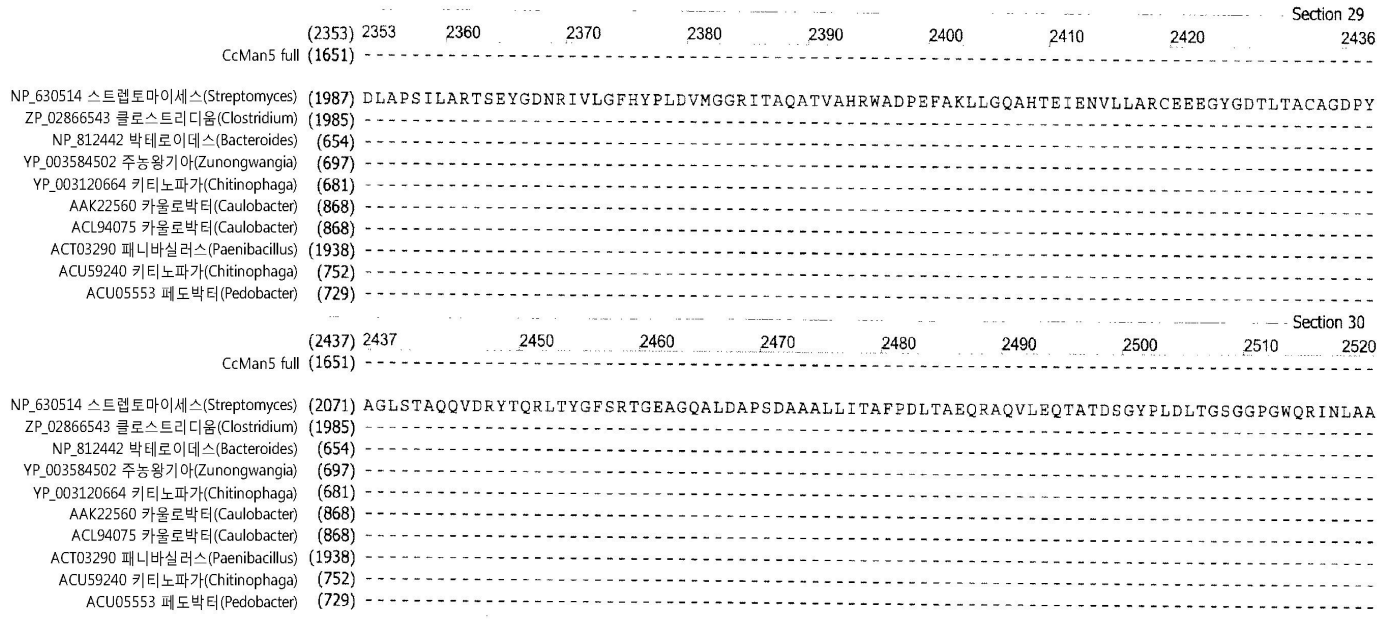
도면311

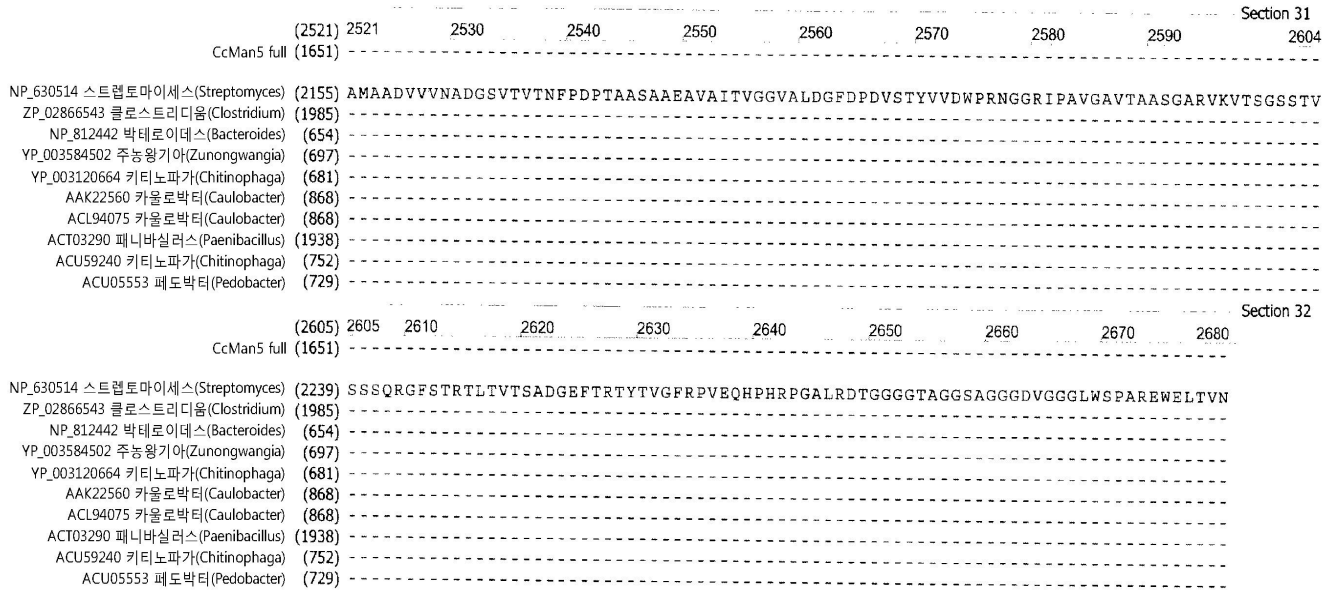

Section 23
Section 24
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_02866543 클로스트리디움(Clostridium)
NP_812442 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
YP_003120664 키티노파가(Chitinophaga)
AAK22560 카울로박터(Caulobacter)
ACL94075 카울로박터(Caulobacter)
ACT03290 페니바실러스(Paenibacillus)
ACU59240 키티노파가(Chitinophaga)
ACU05553 페도박터(Pedobacter)


Section 25
(2017) 2017 2030 2040 2050 2060 2070 2080 2090 2100
CcMan5 full (1602) SFVGVAPGASAYQSFAARSATGDLDVTVTATGADGTQTVEQVVTVPSCS
NP_630514 스트렙토마이세스(Streptomyces) (1651) VFTRLKDGTSYTARVRAVNALGDSPWSARTQPAVTGDNRPQTPTVTGVVTDGERVRVNWRPAGDGGFPVVGYTVALDDGTTAHV
ZP_02866543 클로스트리디움(Clostridium) (1785) NAFDRLASVMQKLEFFKGDKKALKAFIDKVTGLDSSKYTQTTWTAFDKELTEAIAVYNDENAMQEEVNTAYSELVTAFLNLRLI
NP_812442 박테로이데스(Bacteroides) (654)
YP_003584502 주농왕기아(Zunongwangia) (697)
YP_003120664 키티노파가(Chitinophaga) (681)
AAK22560 카울로박터(Caulobacter) (868)
ACL94075 카울로박터(Caulobacter) (868)
ACT03290 패니바실러스(Paenibacillus) (1819) GKQVANTFNGAWLAYKRMDFGTEGVNTFAVEYAGNTTNCFTNSAVEIRLGSPTGTLVGKISTPPKAGNWTTYDTVSGTLTQKLT
ACU59240 키티노파가(Chitinophaga) (752)
ACU05553 페도박터(Pedobacter) (729)
Section 26
(2101) 2101 2110 2120 2130 2140 2150 2160 2170 2184
CcMan5 full (1651)
NP_630514 스트렙토마이세스(Streptomyces) (1735) PGTTSTAVLTAAGGAKAHTATVTAVTRAGSSDGSGATVSTAPATSTTSATSATSTGDPAEYEPSPFPGDTLDATYASDAWPETG
ZP_02866543 클로스트리디움(Clostridium) (1869) PDKSLLEDLINQANGLNGANYTKATFDGLTKALDEAKAVYENPDATQKEVDNAKDVLAKAIAGLQTVTTDNTVSTPVNNGDTTA
NP_812442 박테로이데스(Bacteroides) (654)
YP_003584502 주농왕기아(Zunongwangia) (697)
YP_003120664 키티노파가(Chitinophaga) (681)
AAK22560 카울로박터(Caulobacter) (868)
ACL94075 카울로박터(Caulobacter) (868)
ACT03290 패니바실러스(Paenibacillus) (1903) GIQDVYLVLTGSAGNGETGKKYIGNFDNAAFSLKV
ACU59240 키티노파가(Chitinophaga) (752)
ACU05553 페도박터(Pedobacter) (729)

도면31n


Section 27
2185
2190
2200
2210
2220
2230
2240
2250
2268
CcMan5 full (2185) (1651)
NP_630514 스트렙토마이세스(Streptomyces) (1819) DGSDWFTHLLSGFDDLGPATLGANSEVPAGTPLGAENDRITVSVNNAATQQQVDRAEVDASNSATVTMADGLGSRLGPLYGEAL
ZP_02866543 클로스트리디움(Clostridium) (1953) SVKTGDESLAGMFATIALLSIAGYTILKRKEN
NP_812442 박테로이데스(Bacteroides) (654)
YP_003584502 주농왕기아(Zunongwangia) (697)
YP_003120664 키티노파가(Chitinophaga) (681)
AAK22560 카울로박터(Caulobacter) (868)
ACL94075 카울로박터(Caulobacter) (868)
ACT03290 페니바실러스(Paenibacillus) (1938)
ACU59240 키티노파가(Chitinophaga) (752)
ACU05553 페도박터(Pedobacter) (729)
Section 28
2269
2280
2290
2300
2310
2320
2330
2340
2352
CcMan5 full (2269) (1651)
NP_630514 스트렙토마이세스(Streptomyces) (1903) KEGRLLPKTSALPSRVNENLDTHDAAKNHYQYLRPYVRLGPAGDGGAVYESQDSSYSGLAGQGSYPSGHTYGGYEAGTILATLLP
ZP_02866543 클로스트리디움(Clostridium) (1985)
NP_812442 박테로이데스(Bacteroides) (654)
YP_003584502 주농왕기아(Zunongwangia) (697)
YP_003120664 키티노파가(Chitinophaga) (681)
AAK22560 카울로박터(Caulobacter) (868)
ACL94075 카울로박터(Caulobacter) (868)
ACT03290 페니바실러스(Paenibacillus) (1938)
ACU59240 키티노파가(Chitinophaga) (752)
ACU05553 페도박터(Pedobacter) (729)

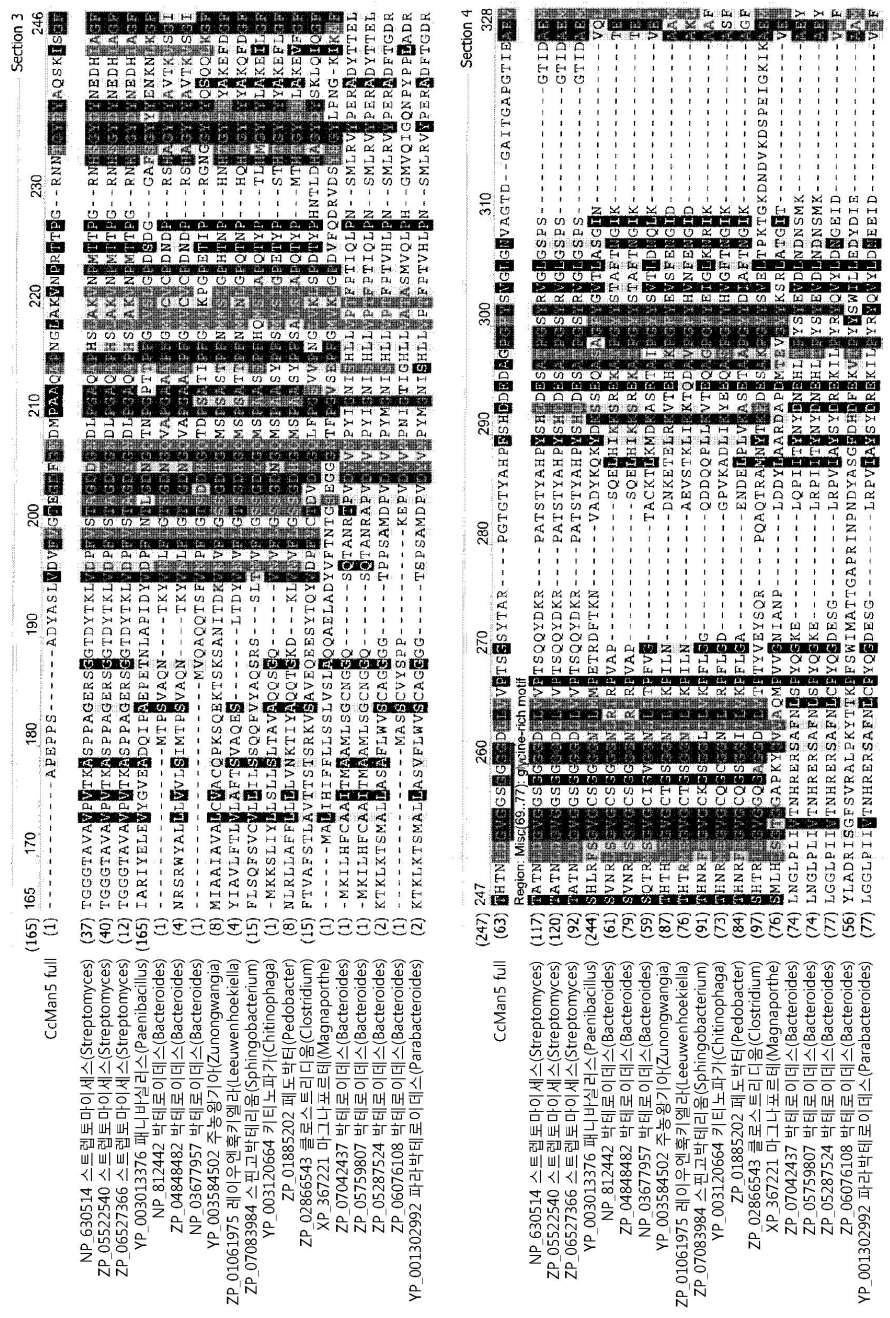

Section 29
(2353) 2353 2360 2370 2380 2390 2400 2410 2420 2436
CcMan5 full (1651) ----------
NP_630514 스트렙토마이세스(Streptomyces) (1987) DLAPSILARTSEYGDNRIVLGFHYPLDVMGGRITAQATVAHRWADPEFAKLLGQAHTEIENVLLARCEEEGYGDTLTACAGDPY
ZP_02866543 클로스트리디움(Clostridium) (1985) ----------
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농윙기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 패니바실러스(Paenibacillus) (1938) ----------
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------
Section 30
(2437) 2437 2450 2460 2470 2480 2490 2500 2510 2520
CcMan5 full (1651) ----------
NP_630514 스트렙토마이세스(Streptomyces) (2071) AGLSTAQQVDRYTQRLTYGFSRTGEAGQALDAPSDAAALLITAFPDLTAEQRAQVLEQTATDSGYPLDLTGSGGPGWQRINLAA
ZP_02866543 클로스트리디움(Clostridium) (1985) ----------
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농윙기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 패니바실러스(Paenibacillus) (1938) ----------
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------


Section 31
(2521) 2521 2530 2540 2550 2560 2570 2580 2590 2604
CcMan5 full (1651) ----------
NP_630514 스트렙토마이세스(Streptomyces) (2155) AMAADVVVNADGSVTVTNFPDPTAASAAEAVAITVGGVALDGFDPDVSTYVVDWPRNGGRIPAVGAVTAASGARVKVTSGSSTV
ZP_02866543 클로스트리디움(Clostridium) (1985) ----------
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 패니바실러스(Paenibacillus) (1938) ----------
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------
Section 32
(2605) 2605 2610 2620 2630 2640 2650 2660 2670 2680
CcMan5 full (1651) ----------
NP_630514 스트렙토마이세스(Streptomyces) (2239) SSSQRGFSTRTLTVTSADGEFTRTYTVGFRPVEQHPHRPGALRDTGGGGTAGGSAGGGDVGGGLWSPAREWELTVN
ZP_02866543 클로스트리디움(Clostridium) (1985) ----------
NP_812442 박테로이데스(Bacteroides) (654) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
AAK22560 카울로박터(Caulobacter) (868) ----------
ACL94075 카울로박터(Caulobacter) (868) ----------
ACT03290 패니바실러스(Paenibacillus) (1938) ----------
ACU59240 키티노파가(Chitinophaga) (752) ----------
ACU05553 페도박터(Pedobacter) (729) ----------

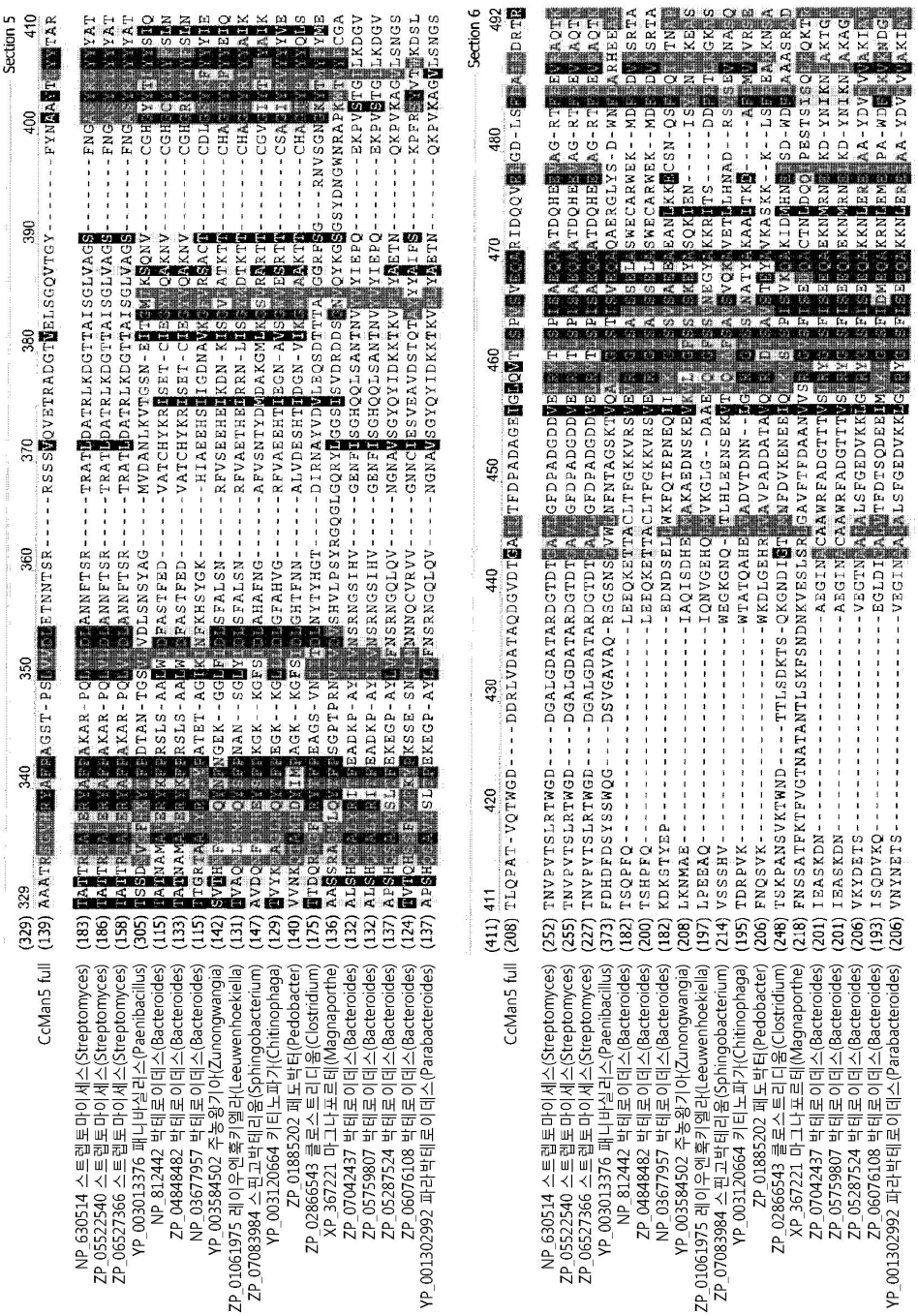
도면32a

Section 1
Section 2
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실러스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핑고박테리움(Sphingobacterium)
YP_003120664 키토파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)

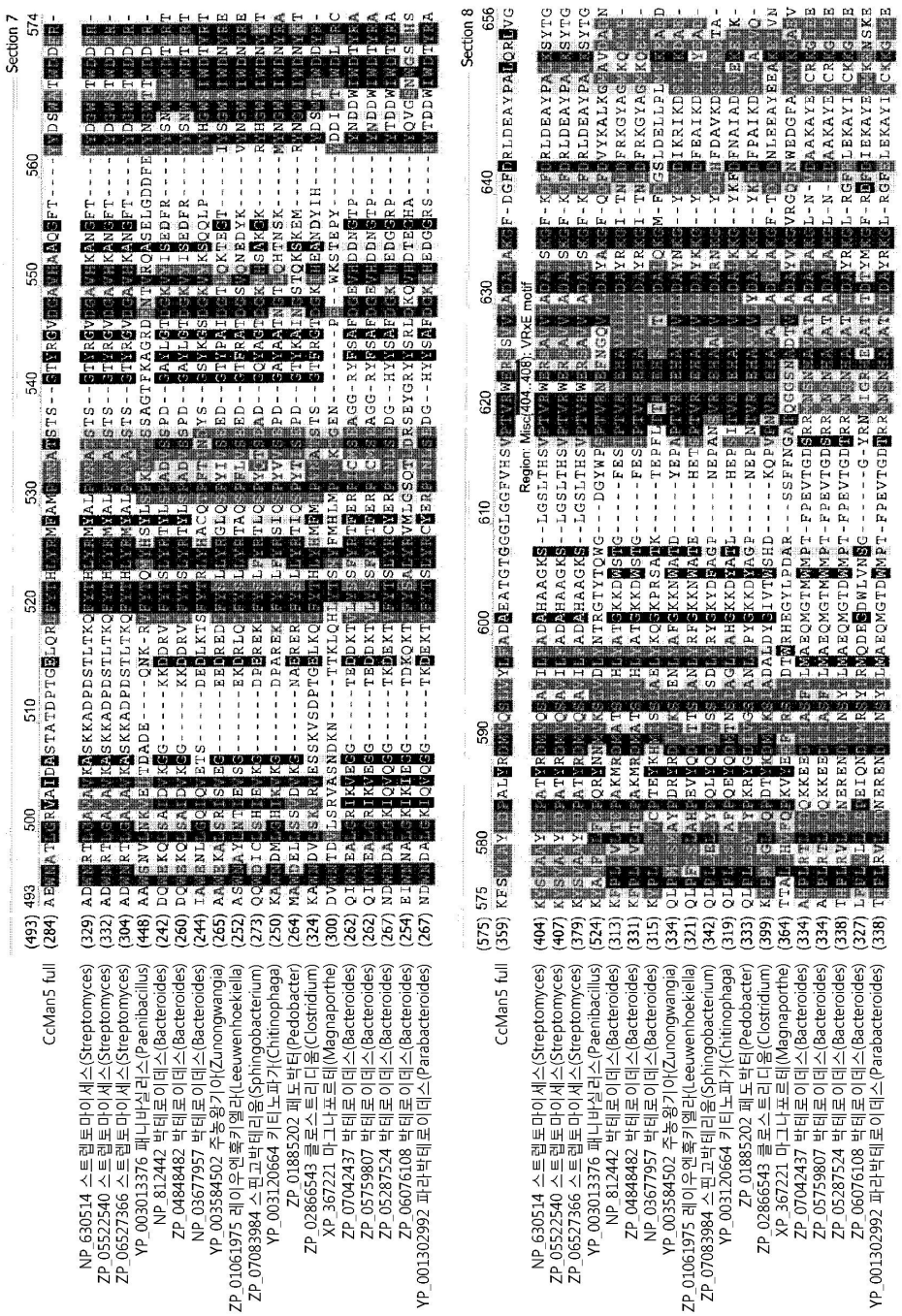
도면32b



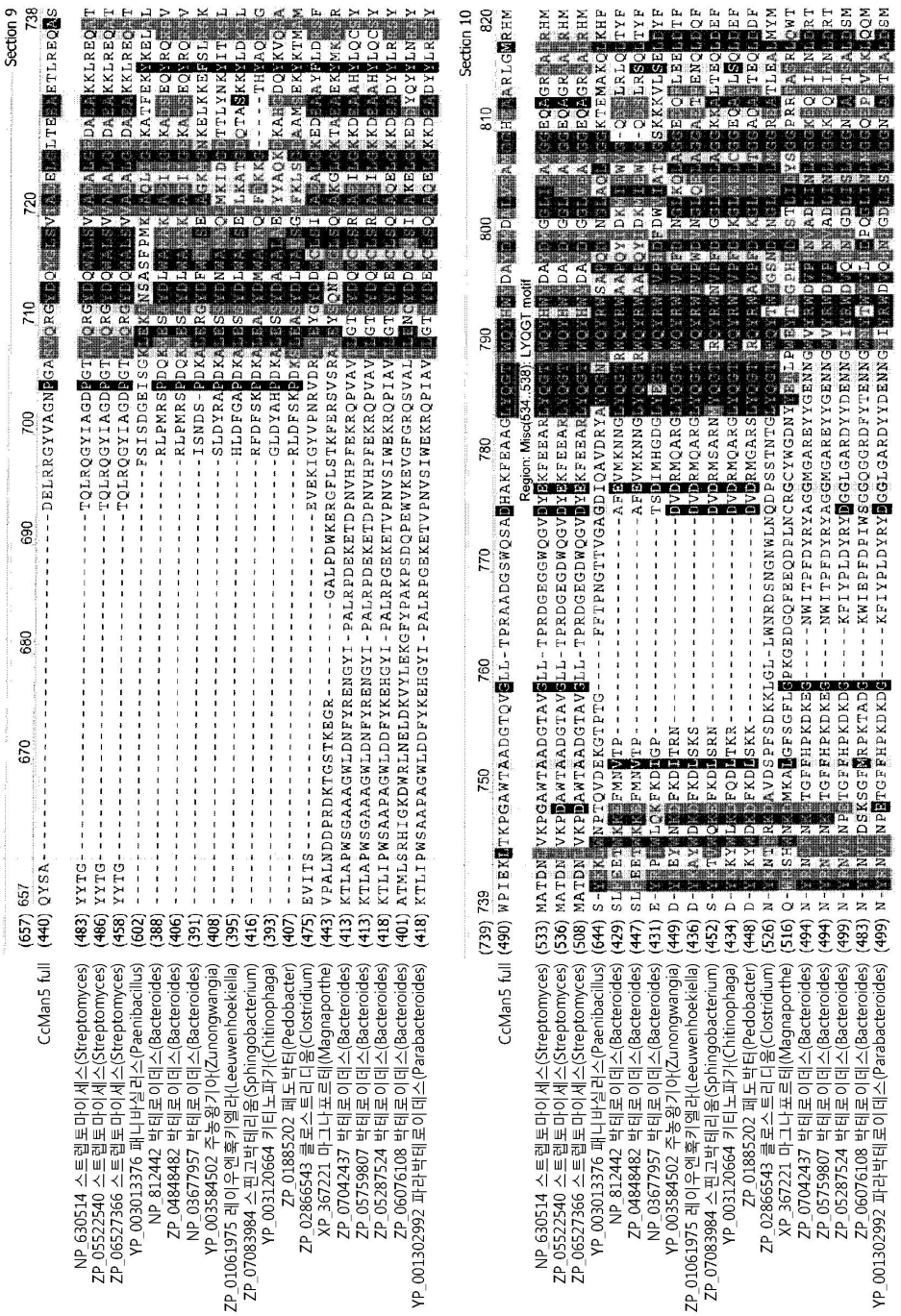
도면32c



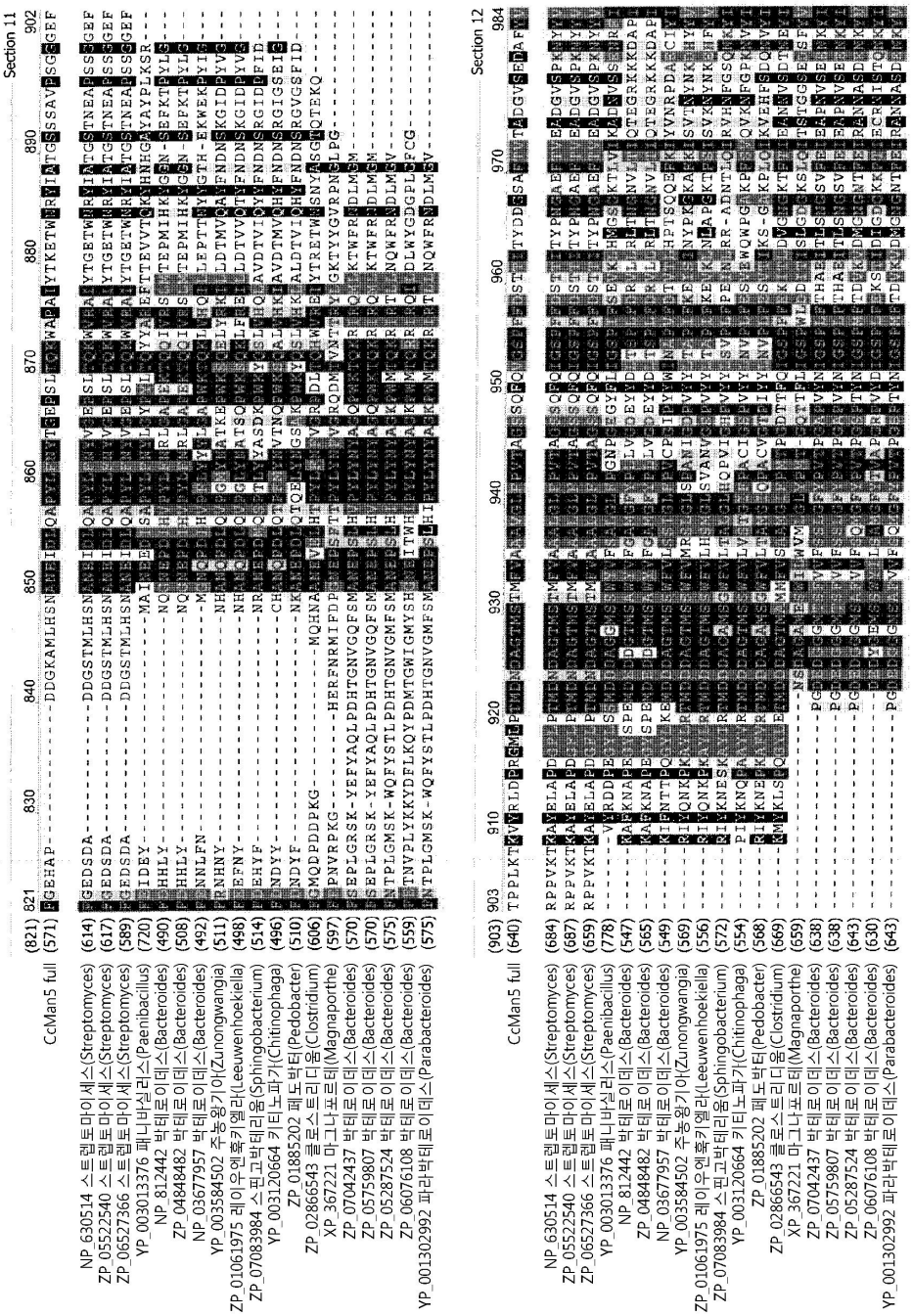
도면32d


Section 7
Section 8
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실리스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핑고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)
Region: Misc(404..408): VRXE motif

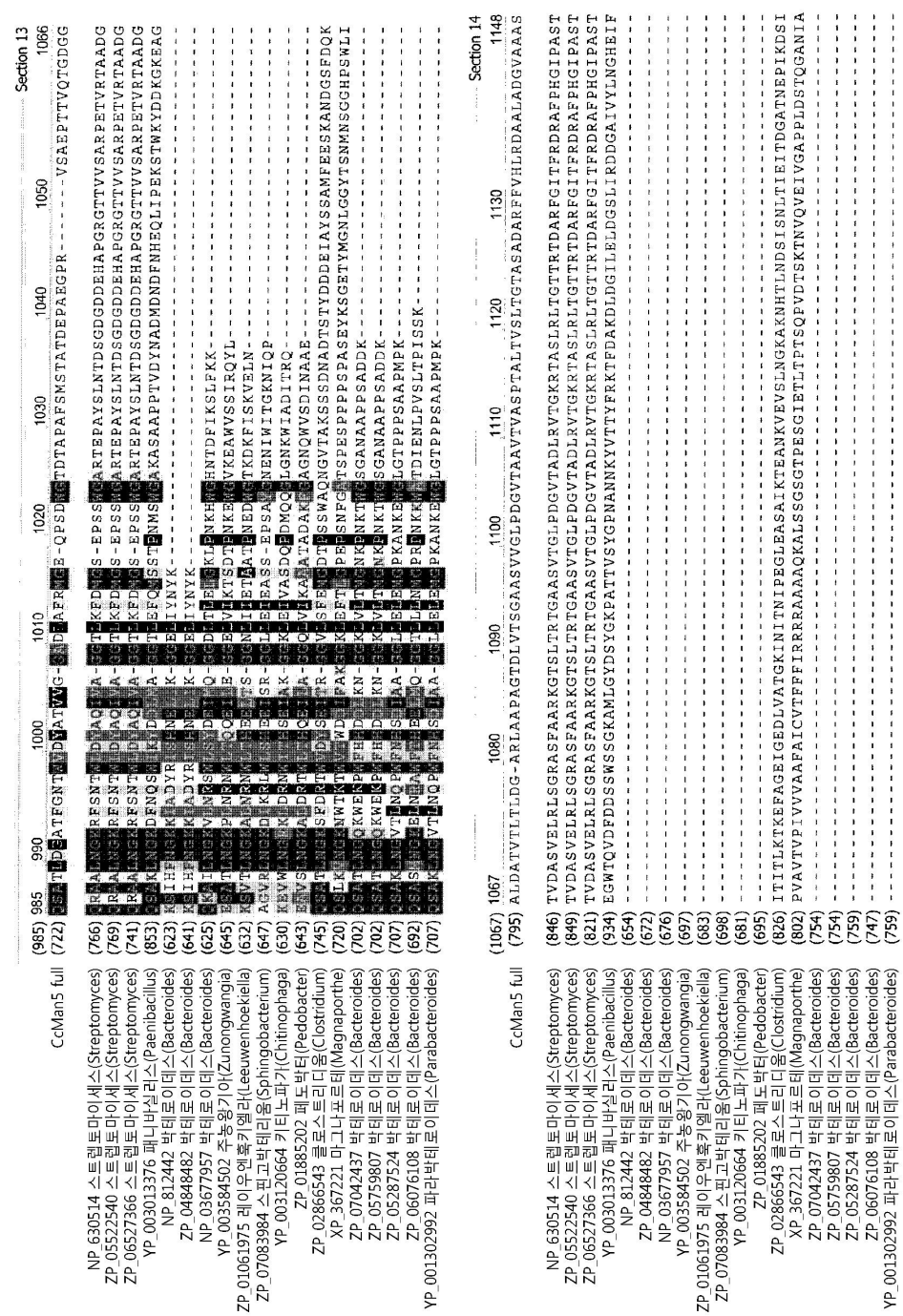
도면32e



도면32f



도면32g



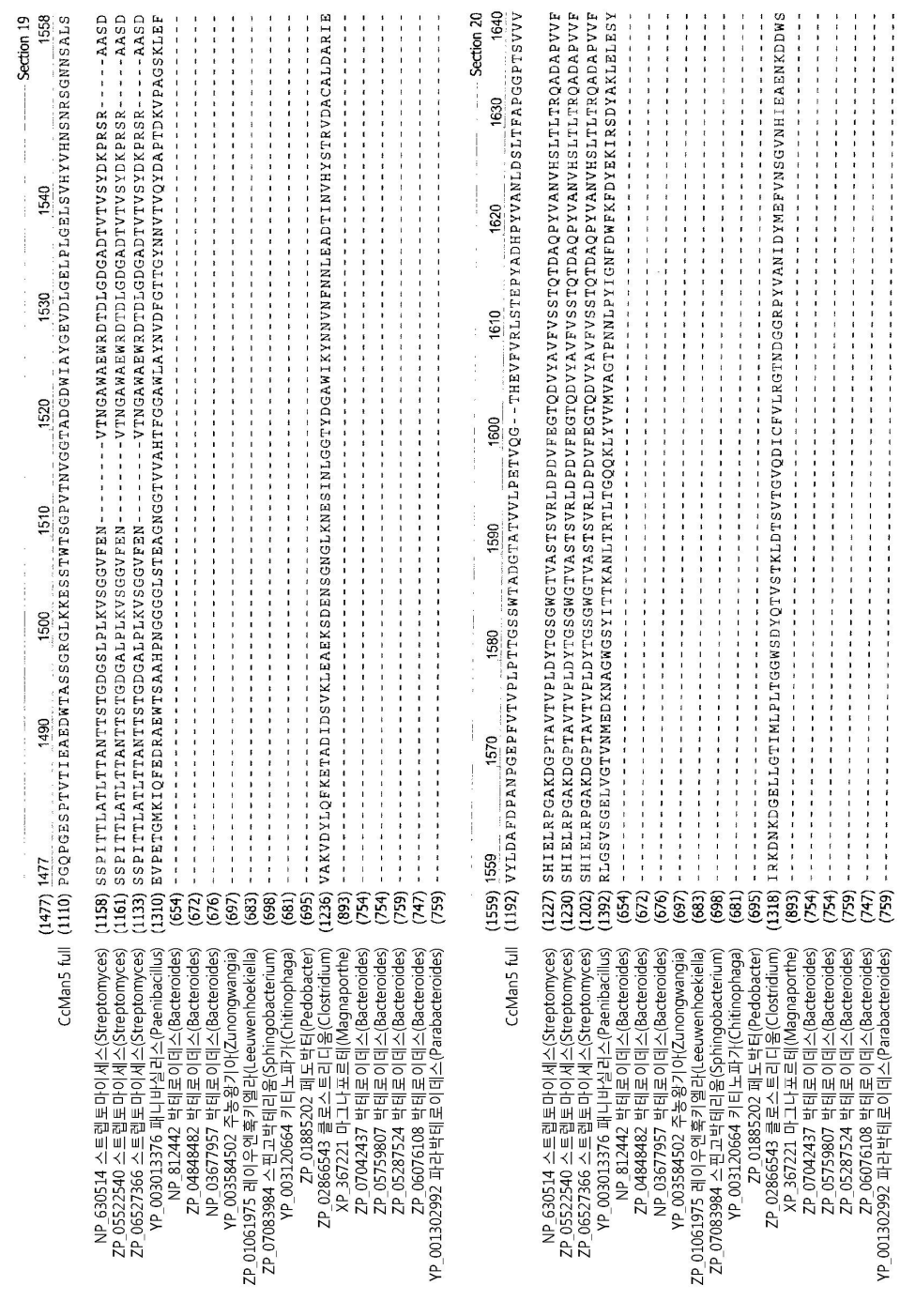
도면32h

Section 15
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실리스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔홀키엘라(Leeuwenhoekiella)
ZP_07083984 스핀고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)
Section 16

도면32i

Section 17
Section 18
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실리스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핑고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)

도면32j

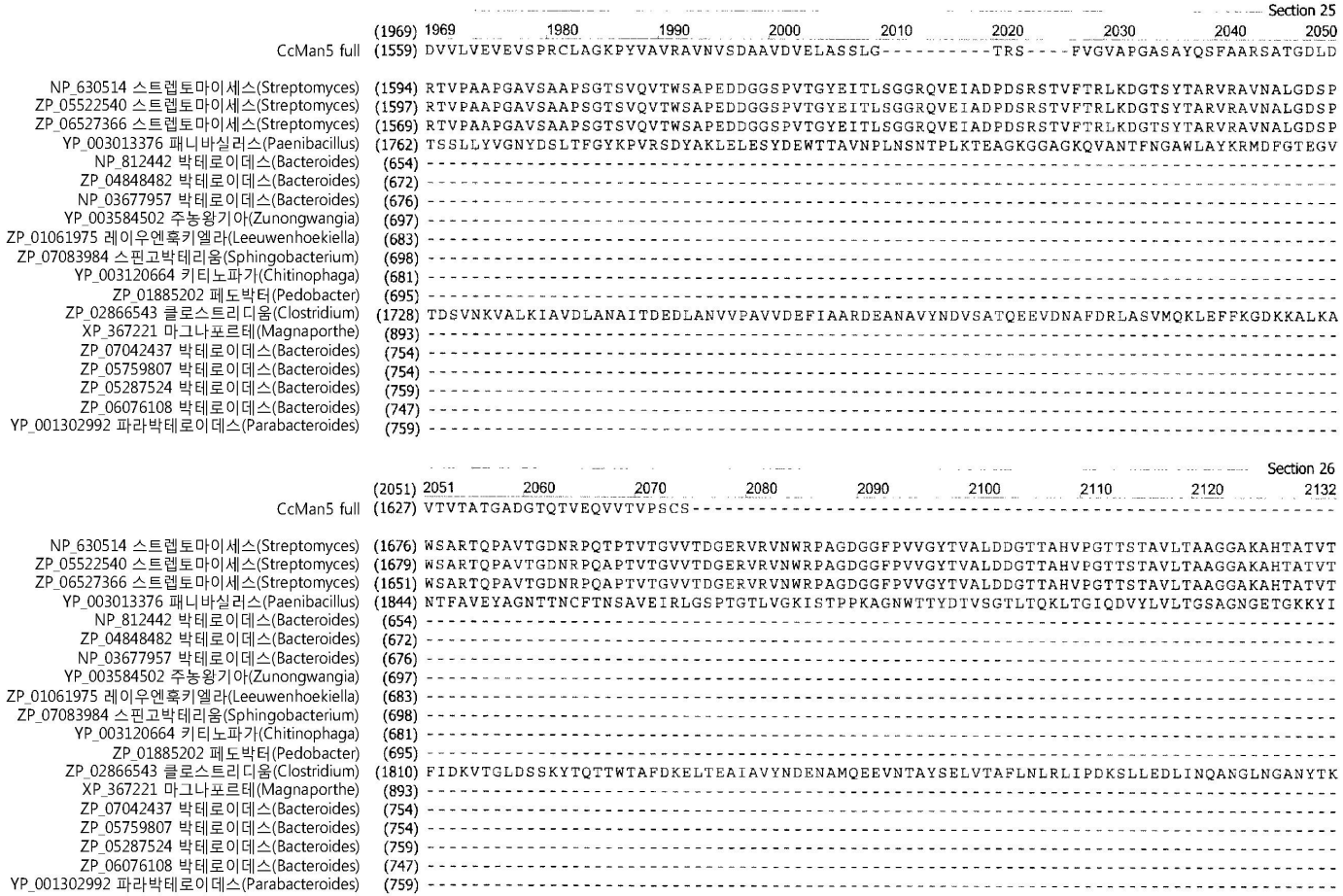

Section 19
Section 20
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실러스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핑고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)

도면32k

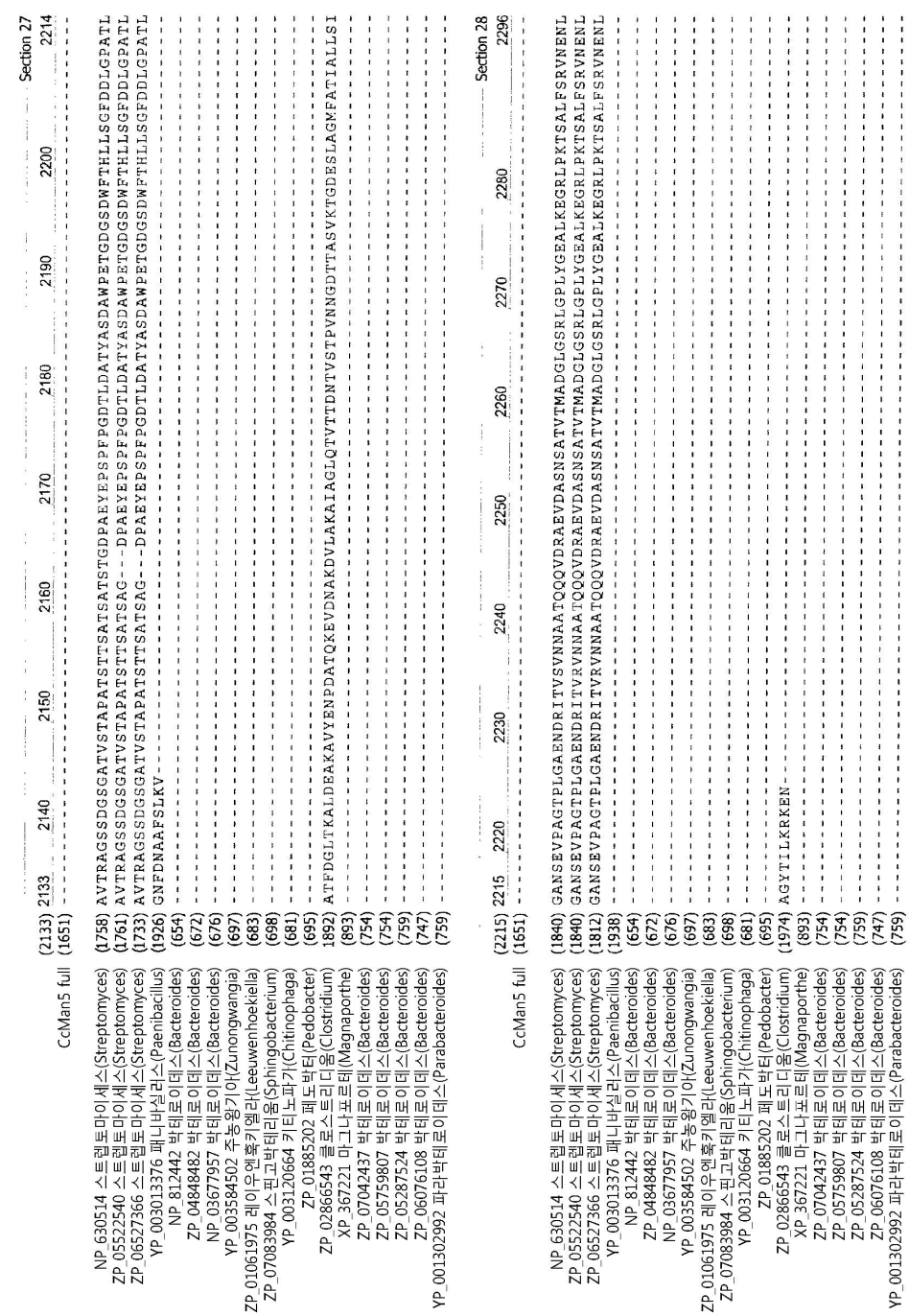
Section 21
(1641) 1641 1650 1660 1670 1680 1690 1700 1710 1722
CcMan5 full (1272) ESRAWTSNSGRGLKNESSTWT-SGPVTNVGG----------TADGDWLAYGE
NP_630514 스트렙토마이세스(Streptomyces) (1309) DATAFEGSSGGGLKSEPATWSGAGSATSLGG----------TYDGAWLDYGD
ZP_05522540 스트렙토마이세스(Streptomyces) (1312) DATAFEGSSGGGLKSEPATWSGAGSATSLGG----------TYDGAWLDYGD
ZP_06527366 스트렙토마이세스(Streptomyces) (1284) DATAFEGSSGGGLKSEPATWSGAGSATSLGG----------TYDGAWLDYGD
YP_003013376 패니바실러스(Paenibacillus) (1474) DEWTTDVNTGNNNTPLKTEAGKGGVGQQVAN----------TFNGAWLAYKR
NP_812442 박테로이데스(Bacteroides) (654) ----------
ZP_04848482 박테로이데스(Bacteroides) (672) ----------
NP_03677957 박테로이데스(Bacteroides) (676) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella) (683) ----------
ZP_07083984 스핑고박테리움(Sphingobacterium) (698) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
ZP_01885202 페도박터(Pedobacter) (695) ----------
ZP_02866543 클로스트리디움(Clostridium) (1400) GAELKVENSTDNTGKSLTNIGGARNDAWLRYNGVEENGKTEMTVRYSHNPGTAGTNSRIDVYLDNMDGNPIGTINLPTTNGW
XP_367221 마그나포르테(Magnaporthe) (893) ----------
ZP_07042437 박테로이데스(Bacteroides) (754) ----------
ZP_05759807 박테로이데스(Bacteroides) (754) ----------
ZP_05287524 박테로이데스(Bacteroides) (759) ----------
ZP_06076108 박테로이데스(Bacteroides) (747) ----------
YP_001302992 파라박테로이데스(Parabacteroides) (759) ----------
Section 22
(1723) 1723 1730 1740 1750 1760 1770 1780 1790 1804
CcMan5 full (1313) IDLGSAALDQLSVHYVHNSNRSGRNSALSVYLDAFDPANPGEPFVTVPLANTGSSWTTDGTAVVDLPSTVRGKHQVWVRLST
NP_630514 스트렙토마이세스(Streptomyces) (1351) VDFGDSFKNTVTLTYVNNSARCGTGSAVQLYLDSFDPDAPGTPYATVPLPVTGSSWSSGGTTSLTLPEAITGTHAVHLRLTT
ZP_05522540 스트렙토마이세스(Streptomyces) (1354) VDFGDSFKNTVTLTYVNNSARCGTGSAVQLYLDSFDPDAPGTPYATVPLPVTGSSWSSGGTTSLTLPEAITGTHAVHLRLTT
ZP_06527366 스트렙토마이세스(Streptomyces) (1326) VDFGDSFKNTVTLTYVNNSARCGTGSAVQLYLDSFDPDAPGTPYATVPLPVTGSSWSSGGTTSLTLPEAITGTHAVHLRLTT
YP_003013376 패니바실러스(Paenibacillus) (1516) MDFGSEGVDKFSIEYAGNSTNTFNNSAVEVRLGSPTGTLVGTVATPPTAAAWGTYATVSGSLTQKLTGLQDVYLVFTGSAAN
NP_812442 박테로이데스(Bacteroides) (654) ----------
ZP_04848482 박테로이데스(Bacteroides) (672) ----------
NP_03677957 박테로이데스(Bacteroides) (676) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella) (683) ----------
ZP_07083984 스핑고박테리움(Sphingobacterium) (698) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
ZP_01885202 페도박터(Pedobacter) (695) ----------
ZP_02866543 클로스트리디움(Clostridium) (1482) ANYTVIREVFDQEITGSHDVYLKLHTDGSGWVANFDWFEFGEPIADVDKSQLQAKYDENVALLQEYDKYHYVGFNIFKDRLL
XP_367221 마그나포르테(Magnaporthe) (893) ----------
ZP_07042437 박테로이데스(Bacteroides) (754) ----------
ZP_05759807 박테로이데스(Bacteroides) (754) ----------
ZP_05287524 박테로이데스(Bacteroides) (759) ----------
ZP_06076108 박테로이데스(Bacteroides) (747) ----------
YP_001302992 파라박테로이데스(Parabacteroides) (759) ----------

도면321

Section 23
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실리스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핑고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)
Section 24


Section 25
(1969) 1969 1980 1990 2000 2010 2020 2030 2040 2050
CcMan5 full (1559) DVVLVEVEVSPRCLAGKPYVAVRAVNVSDAAVDVELASSLG----------TRS----FVGVAPGASAYQSFAARSATGDLD
NP_630514 스트렙토마이세스(Streptomyces) (1594) RTVPAAPGAVSAAPSGTSVQVTWSAPEDDGGSPVTGYEITLSGGRQVEIADPDSRSTVFTRLKDGTSYTARVRAVNALGDSP
ZP_05522540 스트렙토마이세스(Streptomyces) (1597) RTVPAAPGAVSAAPSGTSVQVTWSAPEDDGGSPVTGYEITLSGGRQVEIADPDSRSTVFTRLKDGTSYTARVRAVNALGDSP
ZP_06527366 스트렙토마이세스(Streptomyces) (1569) RTVPAAPGAVSAAPSGTSVQVTWSAPEDDGGSPVTGYEITLSGGRQVEIADPDSRSTVFTRLKDGTSYTARVRAVNALGDSP
YP_003013376 패니바실러스(Paenibacillus) (1762) TSSLLYVGNYDSLTFGYKPVRSDYAKLELESYDEWTTAVNPLNSNTPLKTEAGKGGAGKQVANTFNGAWLAYKRMDFGTEGV
NP_812442 박테로이데스(Bacteroides) (654) ----------
ZP_04848482 박테로이데스(Bacteroides) (672) ----------
NP_03677957 박테로이데스(Bacteroides) (676) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella) (683) ----------
ZP_07083984 스핑고박테리움(Sphingobacterium) (698) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
ZP_01885202 페도박터(Pedobacter) (695) ----------
ZP_02866543 클로스트리디움(Clostridium) (1728) TDSVNKVALKIAVDLANAITDEDLANVVPAVVDEFIAARDEANAVYNDVSATQEEVDNAFDRLASVMQKLEFFKGDKKALKA
XP_367221 마그나포르테(Magnaporthe) (893) ----------
ZP_07042437 박테로이데스(Bacteroides) (754) ----------
ZP_05759807 박테로이데스(Bacteroides) (754) ----------
ZP_05287524 박테로이데스(Bacteroides) (759) ----------
ZP_06076108 박테로이데스(Bacteroides) (747) ----------
YP_001302992 파라박테로이데스(Parabacteroides) (759) ----------
Section 26
(2051) 2051 2060 2070 2080 2090 2100 2110 2120 2132
CcMan5 full (1627) VTVTATGADGTQTVEQVVTVPSCS----------
NP_630514 스트렙토마이세스(Streptomyces) (1676) WSARTQPAVTGDNRPQTPTVTGVVTDGERVRVNWRPAGDGGFPVVGYTVALDDGTTAHVPGTTSTAVLTAAGGAKAHTATVT
ZP_05522540 스트렙토마이세스(Streptomyces) (1679) WSARTQPAVTGDNRPQAPTVTGVVTDGERVRVNWRPAGDGGFPVVGYTVALDDGTTAHVPGTTSTAVLTAAGGAKAHTATVT
ZP_06527366 스트렙토마이세스(Streptomyces) (1651) WSARTQPAVTGDNRPQAPTVTGVVTDGERVRVNWRPAGDGGFPVVGYTVALDDGTTAHVPGTTSTAVLTAAGGAKAHTATVT
YP_003013376 패니바실러스(Paenibacillus) (1844) NTFAVEYAGNTTNCFTNSAVEIRLGSPTGTLVGKISTPPKAGNWTTYDIVSGTLTQKLTGIQDVYLVLTGSAGNGETGKKYI
NP_812442 박테로이데스(Bacteroides) (654) ----------
ZP_04848482 박테로이데스(Bacteroides) (672) ----------
NP_03677957 박테로이데스(Bacteroides) (676) ----------
YP_003584502 주농왕기아(Zunongwangia) (697) ----------
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella) (683) ----------
ZP_07083984 스핑고박테리움(Sphingobacterium) (698) ----------
YP_003120664 키티노파가(Chitinophaga) (681) ----------
ZP_01885202 페도박터(Pedobacter) (695) ----------
ZP_02866543 클로스트리디움(Clostridium) (1810) FIDKVTGLDSSKYTQTTWTAFDKELTEAIAVYNDENAMQEEVNTAYSELVTAFLNLRLIPDKSLLEDLINQANGLNGANYTK
XP_367221 마그나포르테(Magnaporthe) (893) ----------
ZP_07042437 박테로이데스(Bacteroides) (754) ----------
ZP_05759807 박테로이데스(Bacteroides) (754) ----------
ZP_05287524 박테로이데스(Bacteroides) (759) ----------
ZP_06076108 박테로이데스(Bacteroides) (747) ----------
YP_001302992 파라박테로이데스(Parabacteroides) (759) ----------

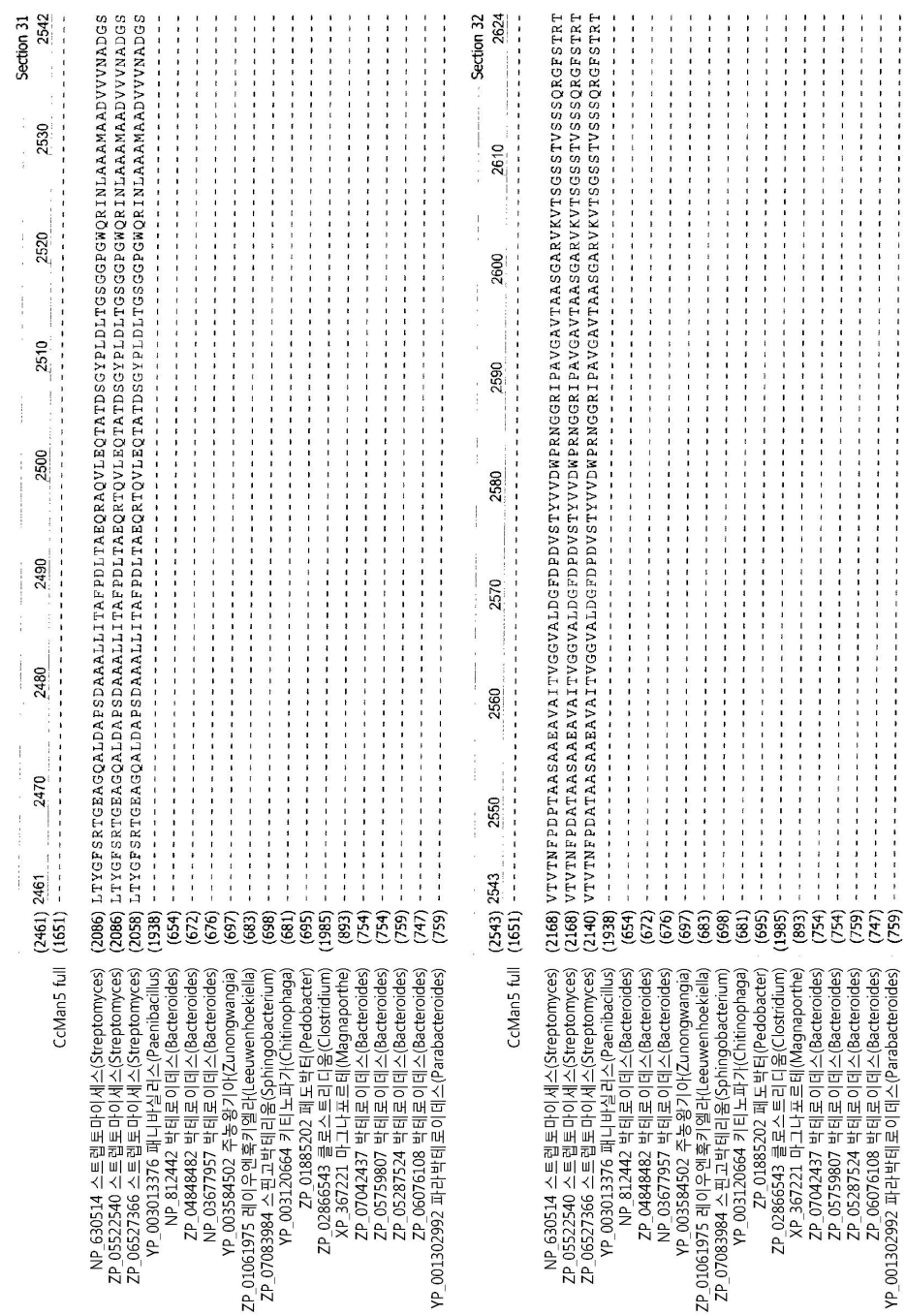
도면32n


Section 27
Section 28
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실러스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핀고박테리움(Sphingobacterium)
YP_003120664 키틴노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)

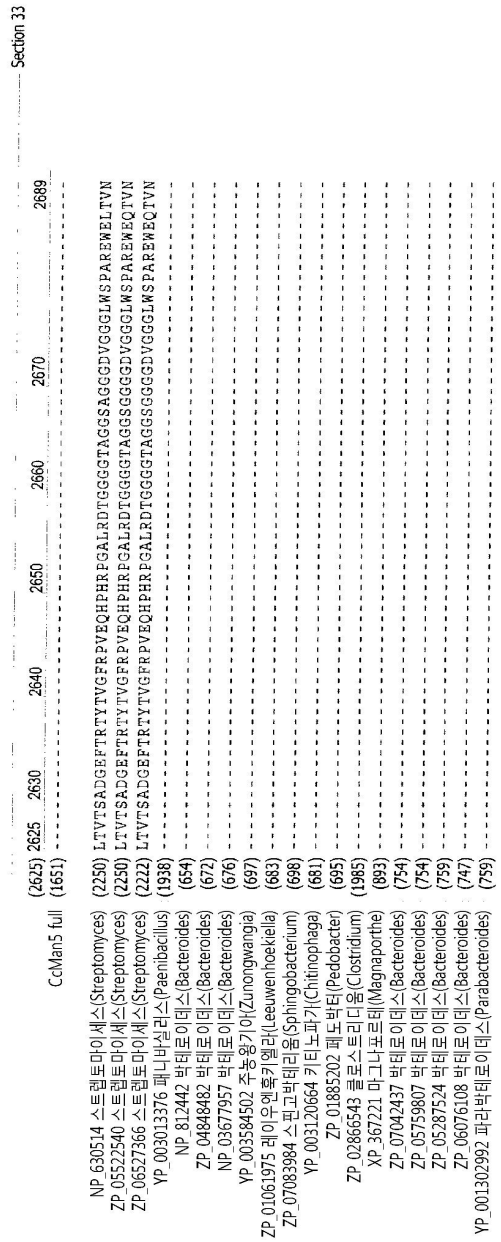
도면32o

Section 29
Section 30
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실리스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핑고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)

도면32p


Section 31
Section 32
CcMan5 full
NP_630514 스트렙토마이세스(Streptomyces)
ZP_05522540 스트렙토마이세스(Streptomyces)
ZP_06527366 스트렙토마이세스(Streptomyces)
YP_003013376 패니바실러스(Paenibacillus)
NP_812442 박테로이데스(Bacteroides)
ZP_04848482 박테로이데스(Bacteroides)
NP_03677957 박테로이데스(Bacteroides)
YP_003584502 주농왕기아(Zunongwangia)
ZP_01061975 레이우엔훅키엘라(Leeuwenhoekiella)
ZP_07083984 스핀고박테리움(Sphingobacterium)
YP_003120664 키티노파가(Chitinophaga)
ZP_01885202 페도박터(Pedobacter)
ZP_02866543 클로스트리디움(Clostridium)
XP_367221 마그나포르테(Magnaporthe)
ZP_07042437 박테로이데스(Bacteroides)
ZP_05759807 박테로이데스(Bacteroides)
ZP_05287524 박테로이데스(Bacteroides)
ZP_06076108 박테로이데스(Bacteroides)
YP_001302992 파라박테로이데스(Parabacteroides)

도면32q


Section 33
(2625) 2625 2630 2640 2650 2660 2670 2689
CcMan5 full (1651)
NP_630514 스트렙토마이세스(Streptomyces) (2250)
ZP_05522540 스트렙토마이세스(Streptomyces) (2250)
ZP_06527366 스트렙토마이세스(Streptomyces) (2222)
YP_003013376 패니바실러스(Paenibacillus) (1938)
NP_812442 박테로이데스(Bacteroides) (654)
ZP_04848482 박테로이데스(Bacteroides) (672)
NP_03677957 박테로이데스(Bacteroides) (676)
YP_003584502 주농왕기아(Zunongwangia) (697)
ZP_01061975 데이우엔훅키엘라(Leeuwenhoekiella) (683)
ZP_07083984 스핑고박테리움(Sphingobacterium) (698)
YP_003120664 키티노파가(Chitinophaga) (681)
ZP_01885202 페도박터(Pedobacter) (695)
ZP_02866543 클로스트리디움(Clostridium) (1985)
XP_367221 마그나포르테(Magnaporthe) (893)
ZP_07042437 박테로이데스(Bacteroides) (754)
ZP_05759807 박테로이데스(Bacteroides) (754)
ZP_05287524 박테로이데스(Bacteroides) (759)
ZP_06076108 박테로이데스(Bacteroides) (747)
YP_001302992 파라박테로이데스(Parabacteroides) (759)

도면33a

```
ATOM   5859  N   GLY B  21       8.798  41.888 151.532  1.00  6.14       B    N
ATOM   5860  CA  GLY B  21       7.349  41.901 151.577  1.00  6.87       B    C
ATOM   5861  C   GLY B  21       6.718  40.662 150.976  1.00  7.36       B    C
ATOM   5862  O   GLY B  21       7.111  39.526 151.272  1.00  7.83       B    O
ATOM   5863  N   ASP B  22       5.749  40.896 150.115  1.00  6.97       B    N
ATOM   5864  CA  ASP B  22       5.019  39.846 149.439  1.00  6.61       B    C
ATOM   5865  CB  ASP B  22       3.516  40.156 149.545  1.00  6.75       B    C
ATOM   5866  CG  ASP B  22       2.645  38.964 149.165  1.00  8.00       B    C
ATOM   5867  OD1 ASP B  22       3.121  37.817 149.252  1.00  8.66       B    O
ATOM   5868  OD2 ASP B  22       1.479  39.168 148.781  1.00 10.69       B    O
ATOM   5869  C   ASP B  22       5.471  39.814 147.976  1.00  6.58       B    C
ATOM   5870  O   ASP B  22       4.647  39.889 147.070  1.00  6.88       B    O
ATOM   5871  N   PHE B  23       6.785  39.719 147.753  1.00  5.72       B    N
ATOM   5872  CA  PHE B  23       7.350  39.730 146.400  1.00  5.51       B    C
ATOM   5873  CB  PHE B  23       8.139  41.020 146.136  1.00  5.56       B    C
ATOM   5874  CG  PHE B  23       7.315  42.266 146.256  1.00  6.28       B    C
ATOM   5875  CD1 PHE B  23       7.127  42.875 147.497  1.00  7.33       B    C
ATOM   5876  CE1 PHE B  23       6.344  44.032 147.621  1.00  6.32       B    C
ATOM   5877  CZ  PHE B  23       5.755  44.584 146.490  1.00  7.38       B    C
ATOM   5878  CE2 PHE B  23       5.941  43.979 145.235  1.00  7.39       B    C
ATOM   5879  CD2 PHE B  23       6.713  42.824 145.132  1.00  7.45       B    C
ATOM   5880  C   PHE B  23       8.243  38.517 146.175  1.00  5.47       B    C
ATOM   5881  O   PHE B  23       9.345  38.640 145.612  1.00  5.14       B    O
ATOM   5882  N   GLY B  24       7.750  37.349 146.600  1.00  4.56       B    N
ATOM   5883  CA  GLY B  24       8.458  36.090 146.362  1.00  4.07       B    C
ATOM   5884  C   GLY B  24       8.397  35.090 147.499  1.00  3.76       B    C
ATOM   5885  O   GLY B  24       8.310  33.871 147.263  1.00  4.04       B    O
ATOM   5886  N   ASN B  25       8.467  35.594 148.734  1.00  2.92       B    N
ATOM   5887  CA  ASN B  25       8.571  34.746 149.915  1.00  2.81       B    C
ATOM   5888  CB  ASN B  25       7.253  34.007 150.179  1.00  3.39       B    C
ATOM   5889  CG  ASN B  25       6.227  34.911 150.842  1.00  3.67       B    C
ATOM   5890  OD1 ASN B  25       6.478  35.461 151.917  1.00  6.16       B    O
ATOM   5891  ND2 ASN B  25       5.090  35.098 150.199  1.00  5.91       B    N
ATOM   5892  C   ASN B  25       9.804  33.822 149.943  1.00  2.73       B    C
ATOM   5893  O   ASN B  25       9.747  32.691 150.423  1.00  2.49       B    O
ATOM   5894  N   ASP B  26      10.924  34.343 149.450  1.00  3.00       B    N
ATOM   5895  CA  ASP B  26      12.206  33.663 149.535  1.00  3.13       B    C
ATOM   5896  CB  ASP B  26      12.975  33.829 148.235  1.00  3.67       B    C
ATOM   5897  CG  ASP B  26      12.481  32.904 147.140  1.00  4.56       B    C
ATOM   5898  OD1 ASP B  26      12.308  31.693 147.408  1.00  6.75       B    O
ATOM   5899  OD2 ASP B  26      12.281  33.385 146.007  1.00  5.43       B    O
ATOM   5900  C   ASP B  26      13.041  34.166 150.702  1.00  3.37       B    C
ATOM   5901  O   ASP B  26      12.747  35.212 151.278  1.00  3.44       B    O
ATOM   6036  N   GLY B  46       7.869  42.799 161.557  1.00  5.19       B    N
ATOM   6037  CA  GLY B  46       7.346  42.212 160.323  1.00  4.86       B    C
ATOM   6038  C   GLY B  46       8.311  41.184 159.770  1.00  5.18       B    C
ATOM   6039  O   GLY B  46       9.481  41.143 160.175  1.00  5.13       B    O
ATOM   6040  N   ARG B  47       7.818  40.352 158.848  1.00  5.25       B    N
ATOM   6041  CA  ARG B  47       8.581  39.247 158.248  1.00  4.58       B    C
ATOM   6042  CB  ARG B  47       8.650  38.046 159.210  1.00  4.70       B    C
ATOM   6043  CG  ARG B  47       7.282  37.403 159.599  1.00  5.20       B    C
ATOM   6044  CD  ARG B  47       7.495  36.267 160.596  1.00  6.03       B    C
ATOM   6045  NE  ARG B  47       6.339  35.388 160.817  1.00  5.27       B    N
ATOM   6046  CZ  ARG B  47       5.957  34.374 160.030  1.00  7.87       B    C
ATOM   6047  NH1 ARG B  47       6.586  34.104 158.889  1.00  6.65       B    N
ATOM   6048  NH2 ARG B  47       4.910  33.627 160.378  1.00  6.18       B    N
ATOM   6049  C   ARG B  47       7.861  38.854 156.975  1.00  4.46       B    C
ATOM   6050  O   ARG B  47       6.656  39.124 156.837  1.00  3.68       B    O
ATOM   6051  N   ASN B  48       8.584  38.257 156.029  1.00  4.25       B    N
ATOM   6052  CA  ASN B  48       7.898  37.485 154.983  1.00  4.09       B    C
ATOM   6053  CB  ASN B  48       8.627  37.525 153.605  1.00  3.81       B    C
ATOM   6054  CG  ASN B  48       9.824  36.572 153.502  1.00  3.63       B    C
ATOM   6055  OD1 ASN B  48      10.191  35.864 154.455  1.00  4.51       B    O
ATOM   6056  ND2 ASN B  48      10.441  36.556 152.325  1.00  2.00       B    N
ATOM   6057  C   ASN B  48       7.626  36.082 155.535  1.00  4.15       B    C
ATOM   6058  O   ASN B  48       7.928  35.813 156.710  1.00  4.08       B    O
```

도면33b

```
ATOM   6059  N   ASN B  49       7.049  35.200 154.730  1.00  3.96       B    N
ATOM   6060  CA  ASN B  49       6.685  33.858 155.217  1.00  4.78       B    C
ATOM   6061  CB  ASN B  49       5.892  33.072 154.155  1.00  4.58       B    C
ATOM   6062  CG  ASN B  49       4.528  33.683 153.864  1.00  6.22       B    C
ATOM   6063  OD1 ASN B  49       3.827  33.270 152.922  1.00  8.25       B    O
ATOM   6064  ND2 ASN B  49       4.151  34.685 154.655  1.00  4.90       B    N
ATOM   6065  C   ASN B  49       7.840  33.013 155.763  1.00  4.50       B    C
ATOM   6066  O   ASN B  49       7.632  32.168 156.622  1.00  5.46       B    O
ATOM   6067  N   THR B  50       9.054  33.243 155.273  1.00  5.02       B    N
ATOM   6068  CA  THR B  50      10.246  32.518 155.744  1.00  4.39       B    C
ATOM   6069  CB  THR B  50      11.413  32.651 154.748  1.00  4.34       B    C
ATOM   6070  OG1 THR B  50      11.882  34.011 154.742  1.00  3.44       B    O
ATOM   6071  CG2 THR B  50      10.950  32.254 153.325  1.00  4.57       B    C
ATOM   6072  C   THR B  50      10.745  33.031 157.096  1.00  4.84       B    C
ATOM   6073  O   THR B  50      11.450  32.306 157.804  1.00  4.54       B    O
ATOM   6074  N   GLY B  51      10.432  34.293 157.416  1.00  4.48       B    N
ATOM   6075  CA  GLY B  51      10.802  34.910 158.707  1.00  4.48       B    C
ATOM   6076  C   GLY B  51      11.560  36.232 158.553  1.00  5.18       B    C
ATOM   6077  O   GLY B  51      11.862  36.918 159.542  1.00  5.14       B    O
ATOM   6200  N   LEU B  67      10.876  26.584 154.563  1.00  4.27       B    N
ATOM   6201  CA  LEU B  67      10.335  26.228 153.272  1.00  4.24       B    C
ATOM   6202  CB  LEU B  67       8.858  25.822 153.413  1.00  3.74       B    C
ATOM   6203  CG  LEU B  67       8.558  24.462 154.063  1.00  2.90       B    C
ATOM   6204  CD1 LEU B  67       7.076  24.167 154.019  1.00  2.92       B    C
ATOM   6205  CD2 LEU B  67       9.361  23.328 153.400  1.00  2.00       B    C
ATOM   6206  C   LEU B  67      10.436  27.457 152.371  1.00  4.38       B    C
ATOM   6207  O   LEU B  67       9.926  28.517 152.720  1.00  4.30       B    O
ATOM   6208  N   ASP B  68      11.070  27.282 151.216  1.00  4.40       B    N
ATOM   6209  CA  ASP B  68      11.312  28.351 150.252  1.00  4.81       B    C
ATOM   6210  CB  ASP B  68      12.398  27.937 149.254  1.00  4.70       B    C
ATOM   6211  CG  ASP B  68      13.677  27.502 149.932  1.00  5.37       B    C
ATOM   6212  OD1 ASP B  68      13.702  26.364 150.444  1.00  4.55       B    O
ATOM   6213  OD2 ASP B  68      14.636  28.300 149.964  1.00  5.37       B    O
ATOM   6214  C   ASP B  68      10.079  28.727 149.459  1.00  4.93       B    C
ATOM   6215  O   ASP B  68       9.469  27.878 148.818  1.00  5.94       B    O
ATOM   6216  N   GLY B  69       9.723  30.005 149.507  1.00  5.17       B    N
ATOM   6217  CA  GLY B  69       8.769  30.598 148.558  1.00  4.56       B    C
ATOM   6218  C   GLY B  69       7.327  30.161 148.689  1.00  4.60       B    C
ATOM   6219  O   GLY B  69       6.573  30.236 147.713  1.00  4.30       B    O
ATOM   6220  N   VAL B  70       6.932  29.725 149.887  1.00  3.48       B    N
ATOM   6221  CA  VAL B  70       5.589  29.149 150.102  1.00  3.51       B    C
ATOM   6222  CB  VAL B  70       5.627  27.927 151.063  1.00  3.54       B    C
ATOM   6223  CG1 VAL B  70       6.419  26.772 150.442  1.00  3.34       B    C
ATOM   6224  CG2 VAL B  70       6.209  28.314 152.422  1.00  2.91       B    C
ATOM   6225  C   VAL B  70       4.609  30.183 150.644  1.00  3.44       B    C
ATOM   6226  O   VAL B  70       5.033  31.257 151.119  1.00  3.41       B    O
ATOM   6227  N   GLY B  71       3.317  29.856 150.590  1.00  3.21       B    N
ATOM   6228  CA  GLY B  71       2.243  30.766 151.033  1.00  4.39       B    C
ATOM   6229  C   GLY B  71       1.896  30.610 152.517  1.00  5.23       B    C
ATOM   6230  O   GLY B  71       2.732  30.168 153.313  1.00  5.01       B    O
ATOM   6231  N   GLY B  72       0.653  30.950 152.869  1.00  5.67       B    N
ATOM   6232  CA  GLY B  72       0.165  30.889 154.256  1.00  5.93       B    C
ATOM   6233  C   GLY B  72       0.949  31.862 155.107  1.00  6.07       B    C
ATOM   6234  O   GLY B  72       1.165  33.016 154.692  1.00  6.64       B    O
ATOM   6235  N   SER B  73       1.390  31.404 156.282  1.00  5.79       B    N
ATOM   6236  CA  SER B  73       2.360  32.159 157.086  1.00  5.71       B    C
ATOM   6237  CB  SER B  73       1.884  32.320 158.534  1.00  5.72       B    C
ATOM   6238  OG  SER B  73       1.572  31.062 159.111  1.00  5.04       B    O
ATOM   6239  C   SER B  73       3.751  31.512 157.055  1.00  5.79       B    C
ATOM   6240  O   SER B  73       4.621  31.864 157.855  1.00  5.53       B    O
ATOM   6241  N   GLY B  74       3.943  30.560 156.136  1.00  5.50       B    N
ATOM   6242  CA  GLY B  74       5.232  29.886 155.962  1.00  4.98       B    C
ATOM   6243  C   GLY B  74       5.389  28.660 156.834  1.00  4.36       B    C
ATOM   6244  O   GLY B  74       4.702  28.512 157.834  1.00  4.36       B    O
ATOM   6245  N   GLY B  75       6.284  27.765 156.443  1.00  4.59       B    N
ATOM   6246  CA  GLY B  75       6.527  26.565 157.216  1.00  4.40       B    C
```

도면33c

```
ATOM   6247  C   GLY B  75       7.889  26.594 157.880  1.00  4.18       B     C
ATOM   6248  O   GLY B  75       8.717  27.435 157.575  1.00  4.41       B     O
ATOM   6249  N   GLY B  76       8.116  25.653 158.788  1.00  4.82       B     N
ATOM   6250  CA  GLY B  76       9.371  25.580 159.508  1.00  4.57       B     C
ATOM   6251  C   GLY B  76       9.475  26.759 160.434  1.00  4.29       B     C
ATOM   6252  O   GLY B  76       8.504  27.108 161.114  1.00  4.51       B     O
ATOM   6253  N   GLY B  77      10.647  27.390 160.433  1.00  3.80       B     N
ATOM   6254  CA  GLY B  77      10.953  28.443 161.386  1.00  3.39       B     C
ATOM   6255  C   GLY B  77      11.241  27.958 162.792  1.00  3.21       B     C
ATOM   6256  O   GLY B  77      11.099  28.729 163.743  1.00  3.66       B     O
ATOM   7082  N   TYR B 194       3.643  22.233 162.238  1.00  3.12       B     N
ATOM   7083  CA  TYR B 194       2.613  21.951 161.250  1.00  3.30       B     C
ATOM   7084  CB  TYR B 194       1.229  21.842 161.911  1.00  3.34       B     C
ATOM   7085  CG  TYR B 194       0.110  21.512 160.940  1.00  3.71       B     C
ATOM   7086  CD1 TYR B 194      -0.461  22.504 160.135  1.00  4.84       B     C
ATOM   7087  CE1 TYR B 194      -1.490  22.204 159.240  1.00  6.19       B     C
ATOM   7088  CZ  TYR B 194      -1.954  20.908 159.146  1.00  4.77       B     C
ATOM   7089  OH  TYR B 194      -2.964  20.602 158.255  1.00  6.66       B     O
ATOM   7090  CE2 TYR B 194      -1.400  19.908 159.936  1.00  5.56       B     C
ATOM   7091  CD2 TYR B 194      -0.370  20.213 160.817  1.00  2.00       B     C
ATOM   7092  C   TYR B 194       2.618  23.039 160.183  1.00  3.39       B     C
ATOM   7093  O   TYR B 194       2.833  24.217 160.484  1.00  2.89       B     O
ATOM   7094  N   PHE B 195       2.402  22.634 158.934  1.00  3.57       B     N
ATOM   7095  CA  PHE B 195       2.176  23.583 157.852  1.00  4.28       B     C
ATOM   7096  CB  PHE B 195       3.495  24.128 157.273  1.00  3.41       B     C
ATOM   7097  CG  PHE B 195       3.307  25.061 156.101  1.00  2.84       B     C
ATOM   7098  CD1 PHE B 195       2.857  26.368 156.296  1.00  2.00       B     C
ATOM   7099  CE1 PHE B 195       2.688  27.240 155.214  1.00  2.78       B     C
ATOM   7100  CZ  PHE B 195       2.970  26.810 153.920  1.00  2.01       B     C
ATOM   7101  CE2 PHE B 195       3.431  25.487 153.713  1.00  3.28       B     C
ATOM   7102  CD2 PHE B 195       3.588  24.636 154.801  1.00  2.14       B     C
ATOM   7103  C   PHE B 195       1.340  22.977 156.739  1.00  4.75       B     C
ATOM   7104  O   PHE B 195       1.700  21.949 156.158  1.00  4.94       B     O
ATOM   7105  N   TYR B 196       0.234  23.655 156.460  1.00  6.22       B     N
ATOM   7106  CA  TYR B 196      -0.598  23.421 155.296  1.00  7.19       B     C
ATOM   7107  CB  TYR B 196       0.228  23.437 153.995  1.00  7.83       B     C
ATOM   7108  CG  TYR B 196      -0.616  23.795 152.800  1.00 10.54       B     C
ATOM   7109  CD1 TYR B 196      -1.022  25.116 152.596  1.00 12.24       B     C
ATOM   7110  CE1 TYR B 196      -1.824  25.468 151.513  1.00 15.08       B     C
ATOM   7111  CZ  TYR B 196      -2.227  24.487 150.619  1.00 16.10       B     C
ATOM   7112  OH  TYR B 196      -3.015  24.861 149.550  1.00 19.44       B     O
ATOM   7113  CE2 TYR B 196      -1.840  23.147 150.800  1.00 15.05       B     C
ATOM   7114  CD2 TYR B 196      -1.038  22.812 151.890  1.00 11.35       B     C
ATOM   7115  C   TYR B 196      -1.432  22.161 155.434  1.00  7.32       B     C
ATOM   7116  O   TYR B 196      -2.592  22.244 155.848  1.00  8.08       B     O
ATOM   7117  N   ASN B 197      -0.848  21.010 155.111  1.00  6.91       B     N
ATOM   7118  CA  ASN B 197      -1.557  19.737 155.149  1.00  7.00       B     C
ATOM   7119  CB  ASN B 197      -1.448  19.010 153.804  1.00  6.93       B     C
ATOM   7120  CG  ASN B 197      -2.381  19.573 152.756  1.00  8.25       B     C
ATOM   7121  OD1 ASN B 197      -3.492  19.991 153.063  1.00  6.81       B     O
ATOM   7122  ND2 ASN B 197      -1.932  19.577 151.499  1.00  9.68       B     N
ATOM   7123  C   ASN B 197      -1.097  18.779 156.242  1.00  6.55       B     C
ATOM   7124  O   ASN B 197      -1.841  17.875 156.589  1.00  6.69       B     O
ATOM   7125  N   ALA B 198       0.104  18.977 156.793  1.00  5.68       B     N
ATOM   7126  CA  ALA B 198       0.738  17.907 157.606  1.00  5.17       B     C
ATOM   7127  CB  ALA B 198       1.519  16.952 156.715  1.00  3.99       B     C
ATOM   7128  C   ALA B 198       1.644  18.391 158.727  1.00  4.90       B     C
ATOM   7129  O   ALA B 198       2.316  19.416 158.594  1.00  5.24       B     O
ATOM   8311  N   ALA B 352       5.634  20.319 147.023  1.00  5.44       B     N
ATOM   8312  CA  ALA B 352       4.594  21.317 146.769  1.00  6.09       B     C
ATOM   8313  CB  ALA B 352       5.135  22.730 146.962  1.00  5.99       B     C
ATOM   8314  C   ALA B 352       4.223  21.089 145.329  1.00  6.35       B     C
ATOM   8315  O   ALA B 352       4.288  21.990 144.503  1.00  6.13       B     O
ATOM   8316  N   THR B 353       3.821  19.855 145.044  1.00  7.15       B     N
ATOM   8317  CA  THR B 353       3.818  19.352 143.690  1.00  7.57       B     C
ATOM   8318  CB  THR B 353       3.887  17.837 143.696  1.00  7.65       B     C
```

도면33d

```
ATOM   8319  OG1 THR B 353       4.943  17.447 144.580  1.00  8.17      B    O
ATOM   8320  CG2 THR B 353       4.184  17.295 142.296  1.00  6.93      B    C
ATOM   8321  C   THR B 353       2.656  19.914 142.859  1.00  8.28      B    C
ATOM   8322  O   THR B 353       2.750  19.992 141.631  1.00  8.21      B    O
ATOM   8323  N   TRP B 354       1.581  20.325 143.531  1.00  8.91      B    N
ATOM   8324  CA  TRP B 354       0.511  21.060 142.861  1.00 10.28      B    C
ATOM   8325  CB  TRP B 354      -0.559  21.535 143.859  1.00 10.88      B    C
ATOM   8326  CG  TRP B 354      -1.795  22.053 143.189  1.00 13.77      B    C
ATOM   8327  CD1 TRP B 354      -2.904  21.331 142.861  1.00 15.75      B    C
ATOM   8328  NE1 TRP B 354      -3.829  22.131 142.242  1.00 18.84      B    N
ATOM   8329  CE2 TRP B 354      -3.338  23.407 142.172  1.00 19.33      B    C
ATOM   8330  CD2 TRP B 354      -2.050  23.396 142.759  1.00 17.76      B    C
ATOM   8331  CE3 TRP B 354      -1.317  24.588 142.806  1.00 19.35      B    C
ATOM   8332  CZ3 TRP B 354      -1.887  25.751 142.273  1.00 22.00      B    C
ATOM   8333  CH2 TRP B 354      -3.175  25.730 141.696  1.00 22.73      B    C
ATOM   8334  CZ2 TRP B 354      -3.914  24.569 141.634  1.00 22.09      B    C
ATOM   8335  C   TRP B 354       1.082  22.248 142.090  1.00  9.94      B    C
ATOM   8336  O   TRP B 354       0.679  22.498 140.957  1.00 10.17      B    O
ATOM   8337  N   ASP B 355       2.027  22.961 142.708  1.00  9.68      B    N
ATOM   8338  CA  ASP B 355       2.607  24.171 142.126  1.00  9.25      B    C
ATOM   8339  CB  ASP B 355       3.020  25.161 143.224  1.00  9.39      B    C
ATOM   8340  CG  ASP B 355       1.920  25.410 144.250  1.00 11.18      B    C
ATOM   8341  OD1 ASP B 355       2.026  24.884 145.380  1.00 12.17      B    O
ATOM   8342  OD2 ASP B 355       0.957  26.138 143.927  1.00 14.96      B    O
ATOM   8343  C   ASP B 355       3.832  23.883 141.261  1.00  8.63      B    C
ATOM   8344  O   ASP B 355       4.010  24.505 140.198  1.00  8.49      B    O
ATOM   8345  N   ASP B 356       4.679  22.961 141.721  1.00  7.73      B    N
ATOM   8346  CA  ASP B 356       6.027  22.821 141.162  1.00  7.97      B    C
ATOM   8347  CB  ASP B 356       7.084  22.835 142.282  1.00  8.30      B    C
ATOM   8348  CG  ASP B 356       7.193  21.517 143.028  1.00 10.67      B    C
ATOM   8349  OD1 ASP B 356       6.378  20.590 142.800  1.00 12.81      B    O
ATOM   8350  OD2 ASP B 356       8.130  21.405 143.856  1.00 13.11      B    O
ATOM   8351  C   ASP B 356       6.290  21.648 140.225  1.00  6.85      B    C
ATOM   8352  O   ASP B 356       7.432  21.434 139.843  1.00  6.23      B    O
ATOM   8702  N   THR B 403       6.656  20.211 154.931  1.00  2.24      B    N
ATOM   8703  CA  THR B 403       6.223  19.841 153.605  1.00  2.00      B    C
ATOM   8704  CB  THR B 403       6.095  18.282 153.520  1.00  2.05      B    C
ATOM   8705  OG1 THR B 403       5.760  17.892 152.198  1.00  3.15      B    O
ATOM   8706  CG2 THR B 403       5.030  17.770 154.453  1.00  2.00      B    C
ATOM   8707  C   THR B 403       4.913  20.578 153.289  1.00  2.00      B    C
ATOM   8708  O   THR B 403       4.316  21.193 154.178  1.00  2.00      B    O
ATOM   8709  N   VAL B 404       4.491  20.536 152.026  1.00  2.00      B    N
ATOM   8710  CA  VAL B 404       3.295  21.247 151.582  1.00  2.02      B    C
ATOM   8711  CB  VAL B 404       3.633  22.358 150.571  1.00  2.00      B    C
ATOM   8712  CG1 VAL B 404       2.406  23.239 150.320  1.00  2.00      B    C
ATOM   8713  CG2 VAL B 404       4.800  23.187 151.064  1.00  2.00      B    C
ATOM   8714  C   VAL B 404       2.175  20.345 151.023  1.00  2.65      B    C
ATOM   8715  O   VAL B 404       1.183  20.118 151.702  1.00  2.92      B    O
ATOM   8716  N   ARG B 405       2.332  19.834 149.803  1.00  3.29      B    N
ATOM   8717  CA  ARG B 405       1.236  19.105 149.134  1.00  3.66      B    C
ATOM   8718  CB  ARG B 405       0.170  20.101 148.633  1.00  3.90      B    C
ATOM   8719  CG  ARG B 405       0.641  20.982 147.480  1.00  3.00      B    C
ATOM   8720  CD  ARG B 405      -0.271  22.201 147.301  1.00  2.02      B    C
ATOM   8721  NE  ARG B 405      -1.643  21.828 146.974  1.00  3.88      B    N
ATOM   8722  CZ  ARG B 405      -2.594  22.700 146.644  1.00  5.53      B    C
ATOM   8723  NH1 ARG B 405      -2.324  24.001 146.580  1.00  5.29      B    N
ATOM   8724  NH2 ARG B 405      -3.812  22.271 146.379  1.00  6.56      B    N
ATOM   8725  C   ARG B 405       1.717  18.221 147.981  1.00  3.69      B    C
ATOM   8726  O   ARG B 405       2.837  18.353 147.494  1.00  3.54      B    O
ATOM   8727  N   TRP B 406       0.861  17.315 147.540  1.00  4.26      B    N
ATOM   8728  CA  TRP B 406       1.343  16.193 146.726  1.00  4.40      B    C
ATOM   8729  CB  TRP B 406       1.380  14.920 147.575  1.00  4.23      B    C
ATOM   8730  CG  TRP B 406       2.088  15.094 148.872  1.00  4.31      B    C
ATOM   8731  CD1 TRP B 406       3.417  14.922 149.110  1.00  4.94      B    C
ATOM   8732  NE1 TRP B 406       3.696  15.167 150.429  1.00  5.82      B    N
ATOM   8733  CE2 TRP B 406       2.537  15.516 151.067  1.00  4.54      B    C
```

도면33e

```
ATOM   8734  CD2 TRP B 406       1.500  15.472 150.112  1.00  4.03       B     C
ATOM   8735  CE3 TRP B 406       0.192  15.805 150.507  1.00  2.57       B     C
ATOM   8736  CZ3 TRP B 406      -0.024  16.148 151.840  1.00  4.33       B     C
ATOM   8737  CH2 TRP B 406       1.032  16.167 152.764  1.00  4.13       B     C
ATOM   8738  CZ2 TRP B 406       2.316  15.859 152.397  1.00  3.80       B     C
ATOM   8739  C   TRP B 406       0.565  15.933 145.433  1.00  4.65       B     C
ATOM   9696  N   GLY B 533     -15.847  27.725 140.342  1.00  9.53       B     N
ATOM   9697  CA  GLY B 533     -15.211  26.855 141.335  1.00  9.25       B     C
ATOM   9698  C   GLY B 533     -13.705  26.691 141.187  1.00  8.71       B     C
ATOM   9699  O   GLY B 533     -13.098  25.875 141.883  1.00  8.16       B     O
ATOM   9700  N   LEU B 534     -13.099  27.477 140.293  1.00  8.58       B     N
ATOM   9701  CA  LEU B 534     -11.653  27.388 140.006  1.00  8.07       B     C
ATOM   9702  CB  LEU B 534     -11.286  28.145 138.734  1.00  7.63       B     C
ATOM   9703  CG  LEU B 534     -11.866  27.619 137.420  1.00  6.93       B     C
ATOM   9704  CD1 LEU B 534     -11.597  28.602 136.307  1.00  3.89       B     C
ATOM   9705  CD2 LEU B 534     -11.303  26.245 137.079  1.00  7.25       B     C
ATOM   9706  C   LEU B 534     -10.747  27.869 141.134  1.00  8.88       B     C
ATOM   9707  O   LEU B 534     -11.033  28.870 141.822  1.00  7.62       B     O
ATOM   9708  N   TYR B 535      -9.637  27.152 141.288  1.00  9.49       B     N
ATOM   9709  CA  TYR B 535      -8.638  27.486 142.280  1.00 10.81       B     C
ATOM   9710  CB  TYR B 535      -8.010  26.197 142.861  1.00 11.89       B     C
ATOM   9711  CG  TYR B 535      -8.753  25.749 144.106  1.00 16.97       B     C
ATOM   9712  CD1 TYR B 535      -8.130  25.772 145.361  1.00 20.37       B     C
ATOM   9713  CE1 TYR B 535      -8.832  25.389 146.520  1.00 23.75       B     C
ATOM   9714  CZ  TYR B 535     -10.184  25.009 146.424  1.00 23.62       B     C
ATOM   9715  OH  TYR B 535     -10.865  24.649 147.562  1.00 24.62       B     O
ATOM   9716  CE2 TYR B 535     -10.827  24.987 145.193  1.00 22.59       B     C
ATOM   9717  CD2 TYR B 535     -10.111  25.365 144.038  1.00 20.68       B     C
ATOM   9718  C   TYR B 535      -7.604  28.447 141.703  1.00 10.16       B     C
ATOM   9719  O   TYR B 535      -6.865  28.094 140.801  1.00 10.24       B     O
ATOM   9720  N   GLN B 536      -7.600  29.677 142.211  1.00  9.98       B     N
ATOM   9721  CA  GLN B 536      -6.568  30.681 141.905  1.00 10.33       B     C
ATOM   9722  CB  GLN B 536      -5.234  30.334 142.598  1.00 10.30       B     C
ATOM   9723  CG  GLN B 536      -5.313  30.141 144.128  1.00 13.24       B     C
ATOM   9724  CD  GLN B 536      -3.939  29.900 144.772  1.00 16.30       B     C
ATOM   9725  OE1 GLN B 536      -2.909  29.886 144.095  1.00 17.97       B     O
ATOM   9726  NE2 GLN B 536      -3.929  29.725 146.091  1.00 18.13       B     N
ATOM   9727  C   GLN B 536      -6.337  30.936 140.402  1.00 10.03       B     C
ATOM   9728  O   GLN B 536      -5.217  31.217 139.967  1.00 10.11       B     O
ATOM   9729  N   GLY B 537      -7.397  30.850 139.608  1.00 10.06       B     N
ATOM   9730  CA  GLY B 537      -7.264  31.159 138.187  1.00  8.79       B     C
ATOM   9731  C   GLY B 537      -8.584  31.268 137.454  1.00  8.30       B     C
ATOM   9732  O   GLY B 537      -9.659  31.075 138.037  1.00  7.71       B     O
ATOM  10119  N   ASN B 586       1.158  35.774 134.842  1.00  4.61       B     N
ATOM  10120  CA  ASN B 586       2.561  35.487 135.130  1.00  4.87       B     C
ATOM  10121  CB  ASN B 586       3.446  36.671 134.695  1.00  4.94       B     C
ATOM  10122  CG  ASN B 586       3.354  36.964 133.194  1.00  4.99       B     C
ATOM  10123  OD1 ASN B 586       2.930  38.041 132.796  1.00  8.76       B     O
ATOM  10124  ND2 ASN B 586       3.754  36.017 132.371  1.00  4.46       B     N
ATOM  10125  C   ASN B 586       2.791  35.182 136.618  1.00  5.09       B     C
ATOM  10126  O   ASN B 586       3.852  34.691 136.997  1.00  3.71       B     O
ATOM  10127  N   ALA B 587       1.777  35.465 137.446  1.00  5.34       B     N
ATOM  10128  CA  ALA B 587       1.917  35.442 138.923  1.00  6.16       B     C
ATOM  10129  CB  ALA B 587       1.006  36.491 139.546  1.00  5.10       B     C
ATOM  10130  C   ALA B 587       1.688  34.108 139.635  1.00  6.83       B     C
ATOM  10131  O   ALA B 587       2.084  33.957 140.784  1.00  6.45       B     O
ATOM  10132  N   ASN B 588       0.997  33.179 138.977  1.00  8.41       B     N
ATOM  10133  CA  ASN B 588       0.684  31.857 139.540  1.00  9.70       B     C
ATOM  10134  CB  ASN B 588      -0.790  31.765 139.954  1.00 10.13       B     C
ATOM  10135  CG  ASN B 588      -1.026  30.752 141.077  1.00 12.26       B     C
ATOM  10136  OD1 ASN B 588      -0.090  30.198 141.622  1.00 15.20       B     O
ATOM  10137  ND2 ASN B 588      -2.286  30.513 141.418  1.00 15.01       B     N
ATOM  10138  C   ASN B 588       1.027  30.786 138.509  1.00 10.25       B     C
ATOM  10139  O   ASN B 588       1.296  31.109 137.356  1.00 10.70       B     O
ATOM  10140  N   GLU B 589       1.014  29.523 138.913  1.00 10.86       B     N
ATOM  10141  CA  GLU B 589       1.574  28.445 138.079  1.00 11.52       B     C
```

도면33f

```
ATOM  10142  CB  GLU B 589       2.198  27.370 138.976  1.00 11.45       B     C
ATOM  10143  CG  GLU B 589       3.501  27.805 139.630  1.00  9.96       B     C
ATOM  10144  CD  GLU B 589       3.311  28.386 141.028  1.00 10.68       B     C
ATOM  10145  OE1 GLU B 589       2.146  28.629 141.434  1.00 10.43       B     O
ATOM  10146  OE2 GLU B 589       4.335  28.606 141.722  1.00  8.00       B     O
ATOM  10147  C   GLU B 589       0.608  27.804 137.071  1.00 11.76       B     C
ATOM  10148  O   GLU B 589       1.027  27.008 136.207  1.00 11.79       B     O
ATOM  10149  N   ILE B 590      -0.670  28.170 137.171  1.00 11.86       B     N
ATOM  10150  CA  ILE B 590      -1.724  27.523 136.394  1.00 11.76       B     C
ATOM  10151  CB  ILE B 590      -3.097  28.106 136.760  1.00 11.88       B     C
ATOM  10152  CG1 ILE B 590      -3.531  27.471 138.085  1.00 12.53       B     C
ATOM  10153  CD1 ILE B 590      -4.228  28.396 138.985  1.00 13.66       B     C
ATOM  10154  CG2 ILE B 590      -4.143  27.847 135.665  1.00 12.30       B     C
ATOM  10155  C   ILE B 590      -1.461  27.444 134.877  1.00 11.77       B     C
ATOM  10156  O   ILE B 590      -1.563  26.362 134.286  1.00 11.66       B     O
ATOM  10444  CA  ILE B 624      -2.166  38.601 140.777  1.00  6.95       B     C
ATOM  10445  CB  ILE B 624      -2.394  39.529 142.009  1.00  7.43       B     C
ATOM  10446  CG1 ILE B 624      -3.881  39.873 142.189  1.00  6.21       B     C
ATOM  10447  CD1 ILE B 624      -4.256  40.384 143.612  1.00  6.93       B     C
ATOM  10448  CG2 ILE B 624      -1.589  40.820 141.848  1.00  5.38       B     C
ATOM  10449  C   ILE B 624      -2.529  37.158 141.097  1.00  8.00       B     C
ATOM  10450  O   ILE B 624      -3.395  36.597 140.423  1.00  7.71       B     O
ATOM  10451  N   ALA B 625      -1.876  36.566 142.107  1.00  8.28       B     N
ATOM  10452  CA  ALA B 625      -2.036  35.146 142.429  1.00  9.17       B     C
ATOM  10453  CB  ALA B 625      -0.672  34.532 142.890  1.00  8.73       B     C
ATOM  10454  C   ALA B 625      -3.121  34.848 143.476  1.00  9.67       B     C
ATOM  10455  O   ALA B 625      -3.305  33.688 143.871  1.00  9.55       B     O
ATOM  10456  N   THR B 626      -3.813  35.889 143.941  1.00 10.19       B     N
ATOM  10457  CA  THR B 626      -4.883  35.743 144.940  1.00  9.61       B     C
ATOM  10458  CB  THR B 626      -4.311  35.718 146.381  1.00 10.06       B     C
ATOM  10459  OG1 THR B 626      -5.328  35.322 147.300  1.00  9.01       B     O
ATOM  10460  CG2 THR B 626      -3.722  37.088 146.791  1.00 10.40       B     C
ATOM  10461  C   THR B 626      -5.890  36.882 144.760  1.00  9.89       B     C
ATOM  10462  O   THR B 626      -5.682  37.753 143.914  1.00  9.41       B     O
ATOM  10463  N   GLY B 627      -6.954  36.880 145.563  1.00 10.02       B     N
ATOM  10464  CA  GLY B 627      -8.057  37.828 145.428  1.00 11.00       B     C
ATOM  10465  C   GLY B 627      -7.684  39.294 145.557  1.00 12.06       B     C
ATOM  10466  O   GLY B 627      -8.185  40.135 144.815  1.00 12.64       B     O
ATOM  10467  N   SER B 628      -6.793  39.604 146.489  1.00 13.03       B     N
ATOM  10468  CA  SER B 628      -6.410  40.988 146.753  1.00 14.00       B     C
ATOM  10469  CB  SER B 628      -7.482  41.650 147.614  1.00 14.26       B     C
ATOM  10470  OG  SER B 628      -7.266  43.039 147.649  1.00 15.26       B     O
ATOM  10471  C   SER B 628      -5.047  41.111 147.452  1.00 14.28       B     C
ATOM  10472  O   SER B 628      -4.640  40.217 148.191  1.00 14.61       B     O
ATOM  10473  N   SER B 629      -4.362  42.229 147.219  1.00 14.43       B     N
ATOM  10474  CA  SER B 629      -3.055  42.511 147.827  1.00 14.96       B     C
ATOM  10475  CB  SER B 629      -1.950  42.366 146.779  1.00 14.81       B     C
ATOM  10476  OG  SER B 629      -0.708  42.826 147.280  1.00 15.07       B     O
ATOM  10477  C   SER B 629      -2.997  43.923 148.390  1.00 15.25       B     C
ATOM  10478  O   SER B 629      -3.433  44.866 147.747  1.00 15.02       B     O
ATOM  10680  N   PRO B 657       4.050  41.388 143.073  1.00  6.53       B     N
ATOM  10681  CA  PRO B 657       3.402  41.236 144.385  1.00  6.71       B     C
ATOM  10682  CB  PRO B 657       2.633  42.551 144.560  1.00  6.75       B     C
ATOM  10683  CG  PRO B 657       2.497  43.125 143.179  1.00  6.81       B     C
ATOM  10684  CD  PRO B 657       3.695  42.650 142.401  1.00  6.65       B     C
ATOM  10685  C   PRO B 657       2.446  40.037 144.343  1.00  7.09       B     C
ATOM  10686  O   PRO B 657       1.785  39.823 143.317  1.00  7.31       B     O
ATOM  10687  N   THR B 658       2.382  39.270 145.438  1.00  6.82       B     N
ATOM  10688  CA  THR B 658       1.659  37.972 145.528  1.00  7.41       B     C
ATOM  10689  CB  THR B 658       0.160  38.025 145.092  1.00  7.22       B     C
ATOM  10690  OG1 THR B 658       0.091  38.108 143.663  1.00  8.46       B     O
ATOM  10691  CG2 THR B 658      -0.562  39.183 145.720  1.00  7.56       B     C
ATOM  10692  C   THR B 658       2.281  36.761 144.802  1.00  6.95       B     C
ATOM  10693  O   THR B 658       1.850  35.614 145.012  1.00  7.00       B     O
ATOM  10694  N   MET B 659       3.266  37.002 143.944  1.00  6.43       B     N
ATOM  10695  CA  MET B 659       3.892  35.925 143.184  1.00  5.95       B     C
```

도면33g

```
ATOM  10696  CB  MET B 659       4.588  36.511 141.944  1.00  5.41       B    C
ATOM  10697  CG  MET B 659       5.187  35.490 140.977  1.00  5.34       B    C
ATOM  10698  SD  MET B 659       5.754  36.339 139.483  1.00  5.63       B    S
ATOM  10699  CE  MET B 659       6.973  35.196 138.813  1.00  4.33       B    C
ATOM  10700  C   MET B 659       4.865  35.073 144.034  1.00  5.78       B    C
ATOM  10701  O   MET B 659       6.092  35.284 144.006  1.00  6.02       B    O
ATOM  10702  N   ASP B 660       4.327  34.118 144.791  1.00  4.92       B    N
ATOM  10703  CA  ASP B 660       5.190  33.199 145.543  1.00  4.96       B    C
ATOM  10704  CB  ASP B 660       4.357  32.285 146.424  1.00  5.01       B    C
ATOM  10705  CG  ASP B 660       3.584  33.032 147.473  1.00  6.04       B    C
ATOM  10706  OD1 ASP B 660       4.035  34.111 147.913  1.00  7.34       B    O
ATOM  10707  OD2 ASP B 660       2.522  32.524 147.864  1.00  7.88       B    O
ATOM  10708  C   ASP B 660       6.037  32.337 144.596  1.00  4.93       B    C
ATOM  10709  O   ASP B 660       5.508  31.791 143.621  1.00  4.69       B    O
ATOM  10710  N   ASN B 661       7.340  32.240 144.873  1.00  4.84       B    N
ATOM  10711  CA  ASN B 661       8.255  31.368 144.118  1.00  5.08       B    C
ATOM  10712  CB  ASN B 661       9.694  31.455 144.681  1.00  5.10       B    C
ATOM  10713  CG  ASN B 661      10.780  31.140 143.629  1.00  6.08       B    C
ATOM  10714  OD1 ASN B 661      10.526  30.495 142.608  1.00  5.96       B    O
ATOM  10715  ND2 ASN B 661      11.992  31.636 143.872  1.00  6.67       B    N
ATOM  10716  C   ASN B 661       7.761  29.903 144.132  1.00  5.04       B    C
ATOM  10717  O   ASN B 661       7.831  29.211 143.113  1.00  4.57       B    O
ATOM  10718  N   ASP B 662       7.240  29.473 145.287  1.00  4.48       B    N
ATOM  10719  CA  ASP B 662       6.702  28.114 145.513  1.00  4.85       B    C
ATOM  10720  CB  ASP B 662       5.483  27.819 144.630  1.00  5.02       B    C
ATOM  10721  CG  ASP B 662       4.285  28.656 144.989  1.00  5.82       B    C
ATOM  10722  OD1 ASP B 662       3.900  28.697 146.173  1.00  7.25       B    O
ATOM  10723  OD2 ASP B 662       3.725  29.259 144.068  1.00  5.44       B    O
ATOM  10724  C   ASP B 662       7.733  27.009 145.308  1.00  4.49       B    C
ATOM  10725  O   ASP B 662       7.726  26.335 144.268  1.00  4.33       B    O
ATOM  10726  N   ALA B 663       8.587  26.804 146.304  1.00  4.58       B    N
ATOM  10727  CA  ALA B 663       9.647  25.783 146.215  1.00  5.19       B    C
ATOM  10728  CB  ALA B 663       9.032  24.354 146.294  1.00  4.88       B    C
ATOM  10729  C   ALA B 663      10.523  25.941 144.958  1.00  5.61       B    C
ATOM  10730  O   ALA B 663      10.951  24.953 144.355  1.00  6.58       B    O
ATOM  10731  N   GLY B 664      10.798  27.191 144.580  1.00  6.10       B    N
ATOM  10732  CA  GLY B 664      11.674  27.503 143.453  1.00  5.63       B    C
ATOM  10733  C   GLY B 664      11.067  27.352 142.068  1.00  5.88       B    C
ATOM  10734  O   GLY B 664      11.783  27.416 141.077  1.00  6.30       B    O
ATOM  12383 CA    CA C   6       1.503  30.134 143.317  1.00 32.32       C   CA
ATOM  11905  O   HOH C   7       2.949  31.974 142.705  1.00  3.45       C    O
ATOM  12345  O   HOH C   8       1.937  29.983 147.278  1.00 20.75       C    O
ATOM  12381  O   HOH C   9       0.607  31.785 144.943  1.00 23.64       C    O
ATOM  12380  O   HOH C  10       0.075  28.500 144.774  1.00 23.09       C    O
ATOM  12403  O3  GOL E   1       2.366  27.082 149.806  1.00 23.71       E    O
ATOM  12404  C3  GOL E   1       0.966  27.273 149.833  1.00 24.21       E    C
ATOM  12405  C2  GOL E   1       0.495  27.587 148.416  1.00 22.96       E    C
ATOM  12406  O2  GOL E   1      -0.173  28.816 148.396  1.00 22.95       E    O
ATOM  12407  C1  GOL E   1      -0.470  26.521 147.940  1.00 23.83       E    C
ATOM  12408  O1  GOL E   1       0.108  25.826 146.867  1.00 22.02       E    O
END
```

도면34a

```
CRYST1   82.202   91.158  224.523  90.00  90.00  90.00 P 2 2 21
SCALE1      0.012165  0.000000  0.000000        0.00000
SCALE2     -0.000000  0.010970  0.000000        0.00000
SCALE3      0.000000 -0.000000  0.004454        0.00000
ATOM      2  CA  ALA A   7       74.405  12.350  86.541  1.00 19.71      A    C
ATOM      7  CA  ASP A   8       72.571  11.287  83.330  1.00 16.72      A    C
ATOM     15  CA  TYR A   9       68.912  11.738  84.261  1.00 12.13      A    C
ATOM     27  CA  ALA A  10       68.003  12.081  80.575  1.00  8.97      A    C
ATOM     32  CA  SER A  11       68.548   8.297  80.172  1.00  8.34      A    C
ATOM     38  CA  LEU A  12       65.749   7.661  82.703  1.00  6.43      A    C
ATOM     46  CA  VAL A  13       63.113   9.628  80.736  1.00  5.75      A    C
ATOM     53  CA  ASP A  14       60.886   7.538  78.450  1.00  6.27      A    C
ATOM     61  CA  VAL A  15       59.534  10.004  75.843  1.00  4.66      A    C
ATOM     68  CA  PHE A  16       57.177   7.311  74.498  1.00  3.87      A    C
ATOM     79  CA  VAL A  17       55.121   7.160  77.667  1.00  3.47      A    C
ATOM     86  CA  GLY A  18       51.656   8.454  76.810  1.00  3.67      A    C
ATOM     90  CA  THR A  19       52.191   8.387  72.991  1.00  3.13      A    C
ATOM     97  CA  GLU A  20       49.720   5.556  72.200  1.00  4.27      A    C
ATOM    106  CA  GLY A  21       46.188   6.265  70.980  1.00  5.40      A    C
ATOM    110  CA  ASP A  22       43.990   8.104  73.483  1.00  5.27      A    C
ATOM    118  CA  PHE A  23       46.225   7.338  76.520  1.00  4.39      A    C
ATOM    129  CA  GLY A  24       47.673  10.774  77.312  1.00  3.08      A    C
ATOM    133  CA  ASN A  25       48.095  12.630  73.991  1.00  2.47      A    C
ATOM    141  CA  ASP A  26       51.797  13.270  74.588  1.00  3.20      A    C
ATOM    149  CA  MET A  27       54.547  13.334  71.981  1.00  3.36      A    C
ATOM    157  CA  PRO A  28       57.941  11.662  71.502  1.00  3.20      A    C
ATOM    164  CA  ALA A  29       59.066  14.935  69.784  1.00  3.06      A    C
ATOM    169  CA  ALA A  30       62.625  16.198  70.105  1.00  3.69      A    C
ATOM    174  CA  GLN A  31       62.423  19.283  72.321  1.00  5.14      A    C
ATOM    183  CA  ALA A  32       64.432  21.430  74.735  1.00  4.55      A    C
ATOM    188  CA  PRO A  33       63.094  22.317  78.204  1.00  4.64      A    C
ATOM    195  CA  ASN A  34       60.021  24.568  77.625  1.00  4.19      A    C
ATOM    203  CA  GLY A  35       60.994  24.689  73.959  1.00  3.44      A    C
ATOM    207  CA  LEU A  36       59.041  26.309  71.169  1.00  3.34      A    C
ATOM    215  CA  ALA A  37       60.477  23.938  68.541  1.00  3.52      A    C
ATOM    220  CA  LYS A  38       58.932  20.498  68.910  1.00  3.01      A    C
ATOM    229  CA  VAL A  39       60.209  18.174  66.213  1.00  2.00      A    C
ATOM    236  CA  ASN A  40       57.543  15.463  66.213  1.00  2.24      A    C
ATOM    244  CA  PRO A  41       57.420  12.468  63.910  1.00  3.24      A    C
ATOM    251  CA  ARG A  42       53.989  12.144  62.244  1.00  4.43      A    C
ATOM    262  CA  THR A  43       52.366   8.720  62.032  1.00  5.82      A    C
ATOM    269  CA  THR A  44       49.941   7.692  59.271  1.00  7.54      A    C
ATOM    276  CA  PRO A  45       47.070   6.985  58.732  1.00  7.64      A    C
ATOM    283  CA  GLY A  46       46.363   7.283  62.482  1.00  5.22      A    C
ATOM    287  CA  ARG A  47       47.876   9.677  65.005  1.00  4.33      A    C
ATOM    298  CA  ASN A  48       47.297  10.971  68.487  1.00  2.87      A    C
ATOM    306  CA  ASN A  49       46.469  14.686  68.971  1.00  3.59      A    C
ATOM    314  CA  THR A  50       50.174  15.553  68.695  1.00  2.96      A    C
ATOM    321  CA  GLY A  51       50.582  13.739  65.384  1.00  2.16      A    C
ATOM    325  CA  TYR A  52       52.147  10.463  66.622  1.00  2.75      A    C
ATOM    337  CA  ASP A  53       50.289   7.308  67.663  1.00  3.03      A    C
ATOM    345  CA  TYR A  54       52.612   4.563  69.034  1.00  3.42      A    C
ATOM    357  CA  ALA A  55       50.295   1.835  67.698  1.00  3.80      A    C
ATOM    362  CA  GLN A  56       50.932   2.905  64.064  1.00  4.22      A    C
ATOM    371  CA  SER A  57       53.533   1.527  61.658  1.00  6.71      A    C
ATOM    377  CA  LYS A  58       54.051   4.328  59.145  1.00  6.64      A    C
ATOM    386  CA  ILE A  59       55.576   7.777  59.431  1.00  5.57      A    C
ATOM    394  CA  SER A  60       55.344  10.584  56.904  1.00  4.00      A    C
ATOM    400  CA  GLY A  61       57.749  13.214  58.346  1.00  3.12      A    C
ATOM    404  CA  PHE A  62       58.483  15.778  61.088  1.00  4.01      A    C
ATOM    415  CA  THR A  63       56.355  18.741  62.206  1.00  3.06      A    C
ATOM    422  CA  HIS A  64       57.952  21.581  64.185  1.00  2.46      A    C
ATOM    432  CA  THR A  65       55.187  23.059  66.366  1.00  2.07      A    C
ATOM    439  CA  ASN A  66       52.932  21.418  68.894  1.00  2.86      A    C
ATOM    447  CA  LEU A  67       51.019  21.341  72.178  1.00  3.42      A    C
ATOM    455  CA  ASP A  68       51.501  18.692  74.851  1.00  2.88      A    C
```

도면34b

```
ATOM    463  CA  GLY A  69      48.598  16.584  76.017  1.00  3.16      A    C
ATOM    467  CA  VAL A  70      45.743  18.719  74.687  1.00  2.95      A    C
ATOM    474  CA  GLY A  71      42.319  17.506  73.447  1.00  4.24      A    C
ATOM    478  CA  GLY A  72      40.310  18.076  70.238  1.00  5.07      A    C
ATOM    482  CA  SER A  73      42.497  16.981  67.319  1.00  5.30      A    C
ATOM    488  CA  GLY A  74      45.570  18.647  68.860  1.00  4.44      A    C
ATOM    492  CA  GLY A  75      47.211  22.012  68.196  1.00  4.03      A    C
ATOM    496  CA  GLY A  76      50.230  23.149  66.192  1.00  3.47      A    C
ATOM    500  CA  GLY A  77      51.645  20.417  63.910  1.00  2.77      A    C
ATOM    504  CA  ASP A  78      53.003  23.083  61.502  1.00  3.19      A    C
ATOM    512  CA  LEU A  79      55.816  22.729  58.975  1.00  3.46      A    C
ATOM    520  CA  LEU A  80      55.960  19.099  57.867  1.00  3.19      A    C
ATOM    528  CA  VAL A  81      59.330  17.935  56.541  1.00  2.38      A    C
ATOM    535  CA  VAL A  82      59.151  14.633  54.633  1.00  2.66      A    C
ATOM    542  CA  PRO A  83      62.241  12.791  53.307  1.00  3.10      A    C
ATOM    549  CA  THR A  84      61.709  10.950  50.007  1.00  3.64      A    C
ATOM    556  CA  SER A  85      63.546   9.484  47.006  1.00  3.93      A    C
ATOM    562  CA  GLY A  86      60.474  10.171  44.894  1.00  4.58      A    C
ATOM    566  CA  SER A  87      59.775  13.000  42.432  1.00  6.80      A    C
ATOM    572  CA  TYR A  88      56.783  15.353  42.023  1.00  6.34      A    C
ATOM    584  CA  THR A  89      55.158  17.278  39.133  1.00  6.35      A    C
ATOM    591  CA  ALA A  90      52.313  18.731  41.257  1.00  5.37      A    C
ATOM    596  CA  ARG A  91      51.505  19.377  44.932  1.00  5.13      A    C
ATOM    607  CA  PRO A  92      52.147  16.058  46.755  1.00  5.88      A    C
ATOM    614  CA  GLY A  93      49.327  13.739  47.757  1.00  6.88      A    C
ATOM    618  CA  THR A  94      49.802  13.201  51.509  1.00  7.19      A    C
ATOM    625  CA  GLY A  95      49.477   9.410  51.104  1.00  6.49      A    C
ATOM    629  CA  THR A  96      52.770   9.459  49.173  1.00  4.50      A    C
ATOM    636  CA  TYR A  97      54.650  10.619  52.300  1.00  3.46      A    C
ATOM    648  CA  ALA A  98      54.035   7.311  54.100  1.00  3.39      A    C
ATOM    653  CA  HIS A  99      57.136   5.240  55.089  1.00  4.47      A    C
ATOM    663  CA  PRO A 100      57.203   1.916  56.975  1.00  5.44      A    C
ATOM    670  CA  PHE A 101      58.803   2.128  60.401  1.00  5.59      A    C
ATOM    681  CA  SER A 102      59.416  -0.191  63.342  1.00  7.19      A    C
ATOM    687  CA  HIS A 103      59.895   0.563  67.048  1.00  6.56      A    C
ATOM    697  CA  ASP A 104      62.859  -1.852  66.789  1.00 10.20      A    C
ATOM    705  CA  ASP A 105      64.517   0.789  64.587  1.00 10.01      A    C
ATOM    713  CA  GLU A 106      63.589   3.886  66.611  1.00  9.26      A    C
ATOM    722  CA  ASP A 107      65.354   5.697  69.473  1.00  9.20      A    C
ATOM    730  CA  ALA A 108      64.518   8.899  71.371  1.00  5.45      A    C
ATOM    735  CA  GLY A 109      65.127  10.798  74.597  1.00  4.60      A    C
ATOM    739  CA  PRO A 110      65.502  14.365  75.895  1.00  3.95      A    C
ATOM    746  CA  GLY A 111      66.404  16.609  72.933  1.00  4.26      A    C
ATOM    750  CA  PHE A 112      66.407  13.997  70.163  1.00  4.28      A    C
ATOM    761  CA  TYR A 113      64.446  11.529  68.079  1.00  5.36      A    C
ATOM    773  CA  SER A 114      65.729   8.957  65.540  1.00  5.75      A    C
ATOM    779  CA  VAL A 115      64.016   6.387  63.260  1.00  4.53      A    C
ATOM    786  CA  GLY A 116      64.781   4.199  60.259  1.00  5.51      A    C
ATOM    790  CA  LEU A 117      62.194   4.893  57.576  1.00  5.77      A    C
ATOM    798  CA  GLY A 118      61.496   2.539  54.687  1.00  6.00      A    C
ATOM    802  CA  ASN A 119      62.614   4.169  51.445  1.00  5.77      A    C
ATOM    810  CA  VAL A 120      59.965   5.458  49.014  1.00  5.26      A    C
ATOM    817  CA  ALA A 121      60.384   6.350  45.306  1.00  6.56      A    C
ATOM    822  CA  GLY A 122      58.455   6.925  42.071  1.00  6.12      A    C
ATOM    826  CA  THR A 123      56.572   9.913  40.635  1.00  7.50      A    C
ATOM    833  CA  ASP A 124      53.544  11.573  42.245  1.00  7.71      A    C
ATOM    841  CA  GLY A 125      50.654   9.176  43.087  1.00  7.97      A    C
ATOM    845  CA  ALA A 126      52.803   6.284  41.903  1.00  8.07      A    C
ATOM    850  CA  ILE A 127      55.283   6.769  44.795  1.00  8.60      A    C
ATOM    858  CA  THR A 128      55.591   3.528  46.766  1.00  9.40      A    C
ATOM    865  CA  GLY A 129      58.010   1.444  48.876  1.00  9.09      A    C
ATOM    869  CA  ALA A 130      61.528   1.221  47.442  1.00  8.74      A    C
ATOM    874  CA  PRO A 131      64.650  -0.773  48.532  1.00  9.45      A    C
ATOM    881  CA  GLY A 132      66.719   0.579  51.405  1.00  9.13      A    C
ATOM    885  CA  THR A 133      66.380   2.820  54.435  1.00  8.85      A    C
ATOM    892  CA  ILE A 134      66.150   6.548  55.020  1.00  7.14      A    C
```

도면34c

```
ATOM    900  CA  GLU A 135      67.868   6.967  58.400  1.00  7.03      A    C
ATOM    909  CA  ALA A 136      66.326  10.008  60.046  1.00  5.25      A    C
ATOM    914  CA  GLU A 137      67.736  11.833  63.043  1.00  3.46      A    C
ATOM    923  CA  VAL A 138      66.270  15.029  64.552  1.00  2.00      A    C
ATOM    930  CA  ALA A 139      67.283  17.116  67.586  1.00  2.00      A    C
ATOM    935  CA  ALA A 140      66.021  20.401  69.089  1.00  3.29      A    C
ATOM    940  CA  ALA A 141      67.192  23.587  70.725  1.00  5.41      A    C
ATOM    945  CA  THR A 142      64.774  26.112  72.324  1.00  3.98      A    C
ATOM    952  CA  ARG A 143      63.747  27.685  69.007  1.00  3.05      A    C
ATOM    963  CA  SER A 144      65.230  25.253  66.475  1.00  3.15      A    C
ATOM    969  CA  GLY A 145      64.814  21.822  64.969  1.00  3.61      A    C
ATOM    973  CA  VAL A 146      67.745  20.138  63.259  1.00  3.66      A    C
ATOM    980  CA  HIS A 147      67.601  17.212  60.859  1.00  4.27      A    C
ATOM    990  CA  ARG A 148      70.193  14.697  59.628  1.00  5.28      A    C
ATOM   1001  CA  TYR A 149      69.133  12.182  56.930  1.00  5.88      A    C
ATOM   1013  CA  ALA A 150      70.978   9.286  55.295  1.00  6.54      A    C
ATOM   1018  CA  PHE A 151      69.444   8.138  51.980  1.00  6.34      A    C
ATOM   1029  CA  PRO A 152      70.329   4.881  50.196  1.00  8.11      A    C
ATOM   1036  CA  ALA A 153      73.284   5.138  47.799  1.00  9.27      A    C
ATOM   1041  CA  GLY A 154      72.114   6.052  44.300  1.00  9.43      A    C
ATOM   1045  CA  SER A 155      69.016   7.860  45.553  1.00  8.94      A    C
ATOM   1051  CA  THR A 156      68.111  11.379  44.480  1.00  5.81      A    C
ATOM   1058  CA  PRO A 157      67.363  12.771  47.997  1.00  5.54      A    C
ATOM   1065  CA  SER A 158      64.501  15.181  48.463  1.00  5.58      A    C
ATOM   1071  CA  LEU A 159      62.757  16.902  51.348  1.00  4.75      A    C
ATOM   1079  CA  VAL A 160      59.189  18.092  50.985  1.00  3.15      A    C
ATOM   1086  CA  VAL A 161      58.408  21.099  53.170  1.00  2.91      A    C
ATOM   1093  CA  ASP A 162      54.597  21.001  53.529  1.00  2.92      A    C
ATOM   1101  CA  LEU A 163      53.165  24.242  54.913  1.00  4.33      A    C
ATOM   1109  CA  GLU A 164      49.518  23.198  54.922  1.00  4.22      A    C
ATOM   1118  CA  THR A 165      49.840  20.473  57.636  1.00  3.08      A    C
ATOM   1125  CA  ASN A 166      48.201  21.399  60.948  1.00  3.13      A    C
ATOM   1133  CA  ASN A 167      46.701  19.356  63.804  1.00  3.44      A    C
ATOM   1141  CA  THR A 168      43.479  21.366  63.648  1.00  4.11      A    C
ATOM   1148  CA  SER A 169      43.334  24.170  61.080  1.00  2.77      A    C
ATOM   1154  CA  ARG A 170      45.644  25.892  58.648  1.00  3.20      A    C
ATOM   1165  CA  ARG A 171      44.592  29.537  58.189  1.00  2.85      A    C
ATOM   1176  CA  SER A 172      47.387  30.959  56.027  1.00  2.36      A    C
ATOM   1182  CA  SER A 173      50.992  30.215  55.065  1.00  3.42      A    C
ATOM   1188  CA  SER A 174      53.845  31.585  52.935  1.00  4.21      A    C
ATOM   1194  CA  VAL A 175      57.213  30.528  51.618  1.00  4.87      A    C
ATOM   1201  CA  GLN A 176      60.036  32.533  50.050  1.00  7.10      A    C
ATOM   1210  CA  VAL A 177      63.000  30.700  48.499  1.00  8.23      A    C
ATOM   1217  CA  GLU A 178      66.645  31.960  48.440  1.00 10.90      A    C
ATOM   1226  CA  THR A 179      69.583  30.187  46.804  1.00  9.85      A    C
ATOM   1233  CA  ARG A 180      72.927  31.368  48.203  1.00 10.39      A    C
ATOM   1244  CA  ALA A 181      76.419  31.485  46.632  1.00 11.40      A    C
ATOM   1249  CA  ASP A 182      77.367  28.014  47.948  1.00 11.35      A    C
ATOM   1257  CA  GLY A 183      74.201  26.531  46.431  1.00  9.02      A    C
ATOM   1261  CA  THR A 184      72.333  25.942  49.723  1.00  6.71      A    C
ATOM   1268  CA  VAL A 185      68.670  26.993  50.031  1.00  5.35      A    C
ATOM   1275  CA  GLU A 186      66.959  29.077  52.696  1.00  5.35      A    C
ATOM   1284  CA  LEU A 187      63.146  29.005  52.987  1.00  3.46      A    C
ATOM   1292  CA  SER A 188      61.169  31.481  55.086  1.00  2.87      A    C
ATOM   1298  CA  GLY A 189      57.690  32.795  55.787  1.00  2.92      A    C
ATOM   1302  CA  GLN A 190      54.805  32.417  58.222  1.00  2.71      A    C
ATOM   1311  CA  VAL A 191      52.276  29.832  59.268  1.00  2.84      A    C
ATOM   1318  CA  THR A 192      48.954  30.792  60.850  1.00  3.69      A    C
ATOM   1325  CA  GLY A 193      46.998  28.051  62.583  1.00  3.21      A    C
ATOM   1329  CA  TYR A 194      44.116  27.757  64.995  1.00  3.16      A    C
ATOM   1341  CA  PHE A 195      43.413  25.739  68.120  1.00  3.26      A    C
ATOM   1352  CA  TYR A 196      40.408  25.930  70.460  1.00  3.84      A    C
ATOM   1364  CA  ASN A 197      39.874  29.693  71.045  1.00  4.26      A    C
ATOM   1372  CA  ALA A 198      42.510  31.491  69.033  1.00  3.45      A    C
ATOM   1377  CA  ALA A 199      44.567  31.753  65.914  1.00  3.47      A    C
ATOM   1382  CA  TYR A 200      48.369  32.096  66.243  1.00  2.36      A    C
```

도면34d

```
ATOM   1394  CA  THR A 201      51.191  33.045  63.858  1.00  2.69      A    C
ATOM   1401  CA  LEU A 202      54.816  31.843  63.819  1.00  2.11      A    C
ATOM   1409  CA  TYR A 203      57.621  32.892  61.467  1.00  2.45      A    C
ATOM   1421  CA  TYR A 204      60.210  30.396  60.264  1.00  2.82      A    C
ATOM   1433  CA  THR A 205      63.573  30.086  58.500  1.00  2.58      A    C
ATOM   1440  CA  ALA A 206      64.904  26.723  57.213  1.00  4.47      A    C
ATOM   1445  CA  ARG A 207      68.253  26.305  55.508  1.00  5.09      A    C
ATOM   1456  CA  THR A 208      69.773  23.209  53.912  1.00  4.88      A    C
ATOM   1463  CA  LEU A 209      73.404  22.419  54.654  1.00  5.26      A    C
ATOM   1471  CA  GLN A 210      73.983  20.900  51.175  1.00  6.01      A    C
ATOM   1480  CA  PRO A 211      73.274  22.391  47.732  1.00  6.91      A    C
ATOM   1487  CA  ALA A 212      69.689  21.962  46.560  1.00  7.18      A    C
ATOM   1492  CA  THR A 213      67.307  23.100  43.847  1.00  5.83      A    C
ATOM   1499  CA  VAL A 214      63.724  23.957  44.710  1.00  5.43      A    C
ATOM   1506  CA  GLN A 215      60.145  24.166  43.470  1.00  5.10      A    C
ATOM   1515  CA  THR A 216      57.130  25.576  45.286  1.00  4.16      A    C
ATOM   1522  CA  TRP A 217      53.372  25.146  45.243  1.00  3.72      A    C
ATOM   1536  CA  GLY A 218      50.356  27.288  46.105  1.00  3.16      A    C
ATOM   1540  CA  ASP A 219      46.606  26.736  46.322  1.00  3.69      A    C
ATOM   1548  CA  ASP A 220      46.180  25.735  42.659  1.00  3.99      A    C
ATOM   1556  CA  ASP A 221      48.330  22.647  43.374  1.00  5.99      A    C
ATOM   1564  CA  ARG A 222      50.800  23.609  40.598  1.00  6.95      A    C
ATOM   1575  CA  LEU A 223      54.426  22.727  41.314  1.00  6.51      A    C
ATOM   1583  CA  VAL A 224      56.418  25.552  39.779  1.00  7.29      A    C
ATOM   1590  CA  ASP A 225      59.828  27.203  39.638  1.00  8.47      A    C
ATOM   1598  CA  ALA A 226      58.537  30.421  41.281  1.00  8.13      A    C
ATOM   1603  CA  THR A 227      60.445  31.382  44.417  1.00  7.82      A    C
ATOM   1610  CA  ALA A 228      57.480  32.859  46.386  1.00  7.30      A    C
ATOM   1615  CA  GLN A 229      54.054  31.642  47.513  1.00  5.21      A    C
ATOM   1624  CA  ASP A 230      51.457  33.169  49.802  1.00  5.36      A    C
ATOM   1632  CA  GLY A 231      48.063  31.590  50.521  1.00  3.52      A    C
ATOM   1636  CA  VAL A 232      46.718  28.704  52.595  1.00  2.62      A    C
ATOM   1643  CA  ASP A 233      48.105  25.470  51.111  1.00  3.49      A    C
ATOM   1651  CA  THR A 234      51.701  26.281  50.149  1.00  3.46      A    C
ATOM   1658  CA  GLY A 235      54.937  24.312  50.148  1.00  3.13      A    C
ATOM   1662  CA  ALA A 236      58.416  23.735  48.763  1.00  3.15      A    C
ATOM   1667  CA  ILE A 237      60.286  20.673  47.495  1.00  3.80      A    C
ATOM   1675  CA  LEU A 238      64.059  20.549  47.934  1.00  4.57      A    C
ATOM   1683  CA  THR A 239      66.059  18.244  45.649  1.00  6.16      A    C
ATOM   1690  CA  PHE A 240      69.681  17.295  46.371  1.00  8.00      A    C
ATOM   1701  CA  ASP A 241      72.537  15.737  44.401  1.00 10.79      A    C
ATOM   1709  CA  PRO A 242      72.834  11.907  44.637  1.00 10.26      A    C
ATOM   1716  CA  ALA A 243      76.443  12.489  45.768  1.00  9.54      A    C
ATOM   1721  CA  ASP A 244      75.082  13.939  49.047  1.00  9.68      A    C
ATOM   1729  CA  ALA A 245      72.792  10.984  49.778  1.00  9.01      A    C
ATOM   1734  CA  GLY A 246      75.008   9.953  52.749  1.00  8.32      A    C
ATOM   1738  CA  GLU A 247      74.244  13.241  54.603  1.00  8.02      A    C
ATOM   1747  CA  ILE A 248      71.325  15.626  54.100  1.00  6.42      A    C
ATOM   1755  CA  GLY A 249      71.023  18.320  56.806  1.00  5.33      A    C
ATOM   1759  CA  LEU A 250      68.338  20.892  57.577  1.00  3.68      A    C
ATOM   1767  CA  GLN A 251      67.982  23.542  60.272  1.00  3.91      A    C
ATOM   1776  CA  VAL A 252      64.613  25.109  61.086  1.00  4.45      A    C
ATOM   1783  CA  THR A 253      64.148  28.007  63.522  1.00  4.13      A    C
ATOM   1790  CA  LEU A 254      60.850  29.555  64.744  1.00  3.15      A    C
ATOM   1798  CA  SER A 255      60.066  33.075  65.957  1.00  3.13      A    C
ATOM   1804  CA  PRO A 256      56.835  34.755  67.030  1.00  2.32      A    C
ATOM   1811  CA  VAL A 257      58.462  38.034  65.945  1.00  2.47      A    C
ATOM   1818  CA  SER A 258      59.667  37.831  62.318  1.00  3.37      A    C
ATOM   1824  CA  VAL A 259      61.516  35.892  59.699  1.00  3.32      A    C
ATOM   1831  CA  GLU A 260      64.595  38.024  60.217  1.00  4.54      A    C
ATOM   1840  CA  GLN A 261      64.385  37.533  63.976  1.00  2.88      A    C
ATOM   1849  CA  ALA A 262      64.235  33.748  63.289  1.00  3.15      A    C
ATOM   1854  CA  ARG A 263      67.445  34.085  61.196  1.00  4.21      A    C
ATOM   1865  CA  ILE A 264      69.210  35.928  64.012  1.00  4.33      A    C
ATOM   1873  CA  ASP A 265      67.932  33.358  66.544  1.00  4.14      A    C
ATOM   1881  CA  GLN A 266      69.289  30.623  64.311  1.00  6.08      A    C
```

도면34e

```
ATOM   1890  CA  GLN A 267      72.772  32.164  64.108  1.00  9.62      A    C
ATOM   1899  CA  VAL A 268      72.748  32.711  67.907  1.00  8.96      A    C
ATOM   1906  CA  GLU A 269      71.365  29.282  68.815  1.00  8.04      A    C
ATOM   1915  CA  LEU A 270      73.089  27.042  66.271  1.00  8.55      A    C
ATOM   1923  CA  GLY A 271      76.001  29.052  64.802  1.00 10.23      A    C
ATOM   1927  CA  ASP A 272      78.321  26.600  63.009  1.00 10.69      A    C
ATOM   1935  CA  LEU A 273      77.764  23.840  65.566  1.00  8.02      A    C
ATOM   1943  CA  SER A 274      77.383  20.316  64.224  1.00  5.99      A    C
ATOM   1949  CA  PHE A 275      74.219  18.210  64.582  1.00  4.80      A    C
ATOM   1960  CA  ASP A 276      75.912  16.069  67.289  1.00  5.78      A    C
ATOM   1968  CA  ALA A 277      76.925  19.139  69.337  1.00  5.08      A    C
ATOM   1973  CA  ILE A 278      73.442  20.669  69.150  1.00  5.84      A    C
ATOM   1981  CA  ARG A 279      71.924  17.315  70.206  1.00  6.83      A    C
ATOM   1992  CA  ASP A 280      74.549  16.902  72.944  1.00  7.77      A    C
ATOM   2000  CA  ARG A 281      74.023  20.429  74.293  1.00  7.74      A    C
ATOM   2011  CA  THR A 282      70.270  19.953  74.599  1.00  6.38      A    C
ATOM   2018  CA  ARG A 283      70.839  16.638  76.431  1.00  5.81      A    C
ATOM   2029  CA  ALA A 284      73.047  18.504  78.940  1.00  5.99      A    C
ATOM   2034  CA  GLU A 285      70.461  21.247  79.310  1.00  6.57      A    C
ATOM   2043  CA  TRP A 286      67.904  18.490  80.139  1.00  5.67      A    C
ATOM   2057  CA  ASN A 287      70.291  16.909  82.648  1.00  8.00      A    C
ATOM   2065  CA  ALA A 288      70.657  20.275  84.422  1.00  8.86      A    C
ATOM   2070  CA  THR A 289      66.862  20.740  84.464  1.00  7.09      A    C
ATOM   2077  CA  LEU A 290      66.138  17.180  85.630  1.00  6.19      A    C
ATOM   2085  CA  GLY A 291      69.085  17.378  88.069  1.00  6.65      A    C
ATOM   2089  CA  ARG A 292      67.121  19.931  90.104  1.00  6.77      A    C
ATOM   2100  CA  VAL A 293      65.530  16.844  91.707  1.00  8.09      A    C
ATOM   2107  CA  ALA A 294      67.845  14.070  92.912  1.00 10.40      A    C
ATOM   2112  CA  ILE A 295      66.318  10.811  94.046  1.00 13.23      A    C
ATOM   2120  CA  ASP A 296      67.534   7.879  96.121  1.00 14.69      A    C
ATOM   2128  CA  ALA A 297      65.001   5.043  95.685  1.00 12.88      A    C
ATOM   2133  CA  SER A 298      65.612   1.701  97.466  1.00 11.58      A    C
ATOM   2139  CA  THR A 299      64.657  -1.726  96.152  1.00  9.13      A    C
ATOM   2146  CA  ALA A 300      61.955  -1.728  98.846  1.00  7.00      A    C
ATOM   2151  CA  THR A 301      59.955   1.150  97.314  1.00  6.08      A    C
ATOM   2158  CA  ASP A 302      61.022   0.632  93.655  1.00  6.93      A    C
ATOM   2166  CA  PRO A 303      62.048  -3.016  93.116  1.00  7.98      A    C
ATOM   2173  CA  THR A 304      62.107  -2.745  89.277  1.00  8.90      A    C
ATOM   2180  CA  GLY A 305      63.438   0.816  88.893  1.00  8.26      A    C
ATOM   2184  CA  GLU A 306      60.285   1.808  86.935  1.00  8.98      A    C
ATOM   2193  CA  LEU A 307      59.007   4.206  89.571  1.00  6.57      A    C
ATOM   2201  CA  GLN A 308      62.236   6.260  89.308  1.00  7.75      A    C
ATOM   2210  CA  ARG A 309      61.928   6.228  85.486  1.00  7.83      A    C
ATOM   2221  CA  LEU A 310      58.276   7.252  85.828  1.00  6.95      A    C
ATOM   2229  CA  PHE A 311      59.288  10.140  88.071  1.00  5.21      A    C
ATOM   2240  CA  TYR A 312      61.933  11.475  85.641  1.00  4.28      A    C
ATOM   2252  CA  THR A 313      59.676  10.962  82.611  1.00  3.87      A    C
ATOM   2259  CA  HIS A 314      56.982  13.080  84.288  1.00  4.29      A    C
ATOM   2269  CA  LEU A 315      59.509  15.636  85.519  1.00  3.00      A    C
ATOM   2277  CA  TYR A 316      60.512  15.912  81.827  1.00  2.15      A    C
ATOM   2289  CA  ARG A 317      56.827  16.467  80.909  1.00  3.71      A    C
ATOM   2300  CA  MET A 318      56.497  19.172  83.611  1.00  6.35      A    C
ATOM   2308  CA  PHE A 319      58.967  21.419  81.712  1.00  7.83      A    C
ATOM   2319  CA  ALA A 320      57.045  21.807  78.404  1.00  7.53      A    C
ATOM   2324  CA  MET A 321      54.323  24.487  78.650  1.00  6.18      A    C
ATOM   2332  CA  PRO A 322      54.063  27.434  78.638  1.00  5.02      A    C
ATOM   2339  CA  MET A 323      56.762  27.568  75.935  1.00  4.44      A    C
ATOM   2347  CA  ASN A 324      59.593  30.069  75.619  1.00  3.49      A    C
ATOM   2355  CA  ALA A 325      57.989  32.716  73.341  1.00  2.49      A    C
ATOM   2360  CA  THR A 326      61.012  35.098  73.515  1.00  2.79      A    C
ATOM   2367  CA  SER A 327      63.597  35.861  70.819  1.00  3.68      A    C
ATOM   2373  CA  THR A 328      67.381  35.744  71.383  1.00  4.17      A    C
ATOM   2380  CA  SER A 329      67.191  39.550  71.284  1.00  4.95      A    C
ATOM   2386  CA  GLY A 330      64.860  39.602  74.327  1.00  4.52      A    C
ATOM   2390  CA  THR A 331      61.681  40.533  72.420  1.00  2.64      A    C
ATOM   2397  CA  TYR A 332      58.212  38.956  72.097  1.00  2.99      A    C
```

도면34f

```
ATOM   2409  CA  ARG A 333      55.039  39.623  70.103  1.00  4.36      A    C
ATOM   2420  CA  GLY A 334      52.050  40.974  71.967  1.00  4.85      A    C
ATOM   2424  CA  VAL A 335      48.352  40.362  71.328  1.00  5.79      A    C
ATOM   2431  CA  ASP A 336      48.362  43.923  69.904  1.00  8.91      A    C
ATOM   2439  CA  GLY A 337      50.391  42.488  67.010  1.00  8.05      A    C
ATOM   2443  CA  ALA A 338      53.464  44.524  67.878  1.00  7.17      A    C
ATOM   2448  CA  VAL A 339      56.967  43.729  69.093  1.00  6.49      A    C
ATOM   2455  CA  HIS A 340      57.792  44.405  72.780  1.00  8.62      A    C
ATOM   2465  CA  ALA A 341      60.735  44.008  75.137  1.00  9.13      A    C
ATOM   2470  CA  ALA A 342      60.679  41.217  77.711  1.00 11.07      A    C
ATOM   2475  CA  GLN A 343      62.952  43.339  79.907  1.00 14.61      A    C
ATOM   2484  CA  GLY A 344      64.997  41.352  82.439  1.00 14.36      A    C
ATOM   2488  CA  PHE A 345      62.907  38.172  82.163  1.00 11.18      A    C
ATOM   2499  CA  THR A 346      61.913  35.493  79.671  1.00  8.35      A    C
ATOM   2506  CA  TYR A 347      58.429  35.786  78.204  1.00  5.52      A    C
ATOM   2518  CA  TYR A 348      56.503  32.511  78.127  1.00  4.55      A    C
ATOM   2530  CA  ASP A 349      53.244  31.880  76.254  1.00  3.32      A    C
ATOM   2538  CA  SER A 350      50.780  29.020  75.499  1.00  3.95      A    C
ATOM   2544  CA  TRP A 351      48.175  28.588  78.184  1.00  3.88      A    C
ATOM   2558  CA  ALA A 352      45.515  26.000  79.064  1.00  3.39      A    C
ATOM   2563  CA  THR A 353      44.865  27.385  82.495  1.00  3.82      A    C
ATOM   2570  CA  TRP A 354      41.278  26.050  82.947  1.00  4.26      A    C
ATOM   2584  CA  ASP A 355      42.871  22.556  82.994  1.00  4.99      A    C
ATOM   2592  CA  ASP A 356      46.325  23.205  84.330  1.00  5.97      A    C
ATOM   2600  CA  PHE A 357      46.209  25.899  87.012  1.00  5.02      A    C
ATOM   2611  CA  ARG A 358      47.423  23.675  89.886  1.00  4.11      A    C
ATOM   2622  CA  LYS A 359      50.623  22.752  87.999  1.00  2.82      A    C
ATOM   2631  CA  PHE A 360      52.171  26.201  88.467  1.00  2.32      A    C
ATOM   2642  CA  SER A 361      52.008  25.888  92.261  1.00  2.68      A    C
ATOM   2648  CA  VAL A 362      54.293  22.814  91.937  1.00  3.61      A    C
ATOM   2655  CA  ILE A 363      56.639  24.444  89.399  1.00  3.43      A    C
ATOM   2663  CA  ALA A 364      57.066  27.142  92.096  1.00  3.80      A    C
ATOM   2668  CA  TYR A 365      58.855  24.692  94.414  1.00  4.70      A    C
ATOM   2680  CA  ILE A 366      60.869  22.863  91.729  1.00  6.01      A    C
ATOM   2688  CA  ASP A 367      62.106  25.826  89.627  1.00  5.57      A    C
ATOM   2696  CA  PRO A 368      61.194  29.124  91.362  1.00  3.96      A    C
ATOM   2703  CA  ALA A 369      62.922  31.367  88.734  1.00  3.74      A    C
ATOM   2708  CA  LEU A 370      61.088  29.740  85.833  1.00  3.88      A    C
ATOM   2716  CA  TYR A 371      57.804  30.009  87.729  1.00  4.80      A    C
ATOM   2728  CA  ARG A 372      58.430  33.689  88.384  1.00  5.45      A    C
ATOM   2739  CA  ASP A 373      59.035  34.310  84.651  1.00  6.19      A    C
ATOM   2747  CA  MET A 374      55.843  32.441  83.776  1.00  3.88      A    C
ATOM   2755  CA  VAL A 375      53.730  34.523  86.171  1.00  2.79      A    C
ATOM   2762  CA  GLN A 376      55.358  37.744  84.962  1.00  3.18      A    C
ATOM   2771  CA  SER A 377      54.574  36.606  81.413  1.00  3.14      A    C
ATOM   2777  CA  LEU A 378      50.953  35.908  82.336  1.00  3.14      A    C
ATOM   2785  CA  VAL A 379      50.799  39.454  83.766  1.00  3.57      A    C
ATOM   2792  CA  TYR A 380      52.193  40.950  80.508  1.00  4.06      A    C
ATOM   2804  CA  LEU A 381      49.694  38.961  78.456  1.00  4.18      A    C
ATOM   2812  CA  PHE A 382      46.697  40.475  80.286  1.00  4.57      A    C
ATOM   2823  CA  ALA A 383      48.397  43.881  80.475  1.00  6.03      A    C
ATOM   2828  CA  ASP A 384      48.684  43.708  76.664  1.00  7.44      A    C
ATOM   2836  CA  ALA A 385      44.974  42.882  76.223  1.00  9.01      A    C
ATOM   2841  CA  GLU A 386      44.199  45.935  78.362  1.00 10.91      A    C
ATOM   2850  CA  ALA A 387      46.670  48.106  76.365  1.00 12.41      A    C
ATOM   2855  CA  THR A 388      44.776  47.471  73.087  1.00 13.68      A    C
ATOM   2862  CA  GLY A 389      41.944  49.514  74.624  1.00 14.39      A    C
ATOM   2866  CA  THR A 390      39.452  47.191  72.889  1.00 15.63      A    C
ATOM   2873  CA  GLY A 391      38.019  45.522  76.019  1.00 14.99      A    C
ATOM   2877  CA  GLY A 392      37.896  42.282  73.993  1.00 13.79      A    C
ATOM   2881  CA  GLY A 393      38.759  38.874  75.470  1.00 11.07      A    C
ATOM   2885  CA  LEU A 394      42.017  37.084  74.740  1.00  7.81      A    C
ATOM   2893  CA  GLY A 395      40.142  34.863  72.238  1.00  7.24      A    C
ATOM   2897  CA  GLY A 396      39.767  37.812  69.847  1.00  7.80      A    C
ATOM   2901  CA  PHE A 397      43.479  38.330  69.213  1.00  7.11      A    C
ATOM   2912  CA  VAL A 398      46.038  36.458  67.145  1.00  5.79      A    C
```

도면34g

```
ATOM   2919  CA  HIS A 399      48.445  34.799  69.624  1.00  3.71      A    C
ATOM   2929  CA  SER A 400      52.295  34.534  69.502  1.00  3.42      A    C
ATOM   2935  CA  VAL A 401      52.383  30.703  69.819  1.00  2.51      A    C
ATOM   2942  CA  PRO A 402      49.744  27.945  69.645  1.00  3.03      A    C
ATOM   2949  CA  THR A 403      47.621  28.327  72.753  1.00  2.63      A    C
ATOM   2956  CA  VAL A 404      44.423  27.083  74.389  1.00  2.59      A    C
ATOM   2963  CA  ARG A 405      42.502  28.941  77.121  1.00  2.87      A    C
ATOM   2974  CA  TRP A 406      43.027  31.341  80.033  1.00  3.31      A    C
ATOM   2988  CA  GLU A 407      40.430  30.754  82.763  1.00  4.40      A    C
ATOM   2997  CA  ARG A 408      41.810  30.579  86.344  1.00  3.53      A    C
ATOM   3008  CA  SER A 409      44.708  32.938  85.445  1.00  3.56      A    C
ATOM   3014  CA  SER A 410      43.670  35.101  88.462  1.00  4.17      A    C
ATOM   3020  CA  VAL A 411      44.366  32.099  90.707  1.00  3.23      A    C
ATOM   3027  CA  VAL A 412      47.816  31.532  89.197  1.00  2.11      A    C
ATOM   3034  CA  VAL A 413      48.819  35.209  89.648  1.00  2.29      A    C
ATOM   3041  CA  ALA A 414      47.478  35.051  93.214  1.00  2.19      A    C
ATOM   3046  CA  ASP A 415      49.593  31.906  93.728  1.00  3.52      A    C
ATOM   3054  CA  ALA A 416      52.816  33.913  93.216  1.00  3.54      A    C
ATOM   3059  CA  ILE A 417      51.560  36.792  95.351  1.00  4.96      A    C
ATOM   3067  CA  ALA A 418      50.526  34.493  98.238  1.00  7.10      A    C
ATOM   3072  CA  LYS A 419      53.976  32.864  97.904  1.00  7.37      A    C
ATOM   3081  CA  GLY A 420      55.694  36.259  98.483  1.00  8.91      A    C
ATOM   3085  CA  PHE A 421      56.789  37.036  94.914  1.00  9.86      A    C
ATOM   3096  CA  ASP A 422      56.741  40.741  94.170  1.00 13.61      A    C
ATOM   3104  CA  GLY A 423      57.809  43.423  91.676  1.00 13.34      A    C
ATOM   3108  CA  PHE A 424      55.491  42.086  88.966  1.00 11.76      A    C
ATOM   3119  CA  ASP A 425      55.749  44.833  86.387  1.00 12.47      A    C
ATOM   3127  CA  ARG A 426      52.408  45.949  84.848  1.00 10.18      A    C
ATOM   3138  CA  LEU A 427      50.173  44.147  87.370  1.00  9.19      A    C
ATOM   3146  CA  ASP A 428      48.058  47.323  87.662  1.00 10.66      A    C
ATOM   3154  CA  GLU A 429      47.378  47.038  83.880  1.00 10.07      A    C
ATOM   3163  CA  ALA A 430      46.738  43.286  83.995  1.00  8.16      A    C
ATOM   3168  CA  TYR A 431      44.091  43.690  86.747  1.00  8.25      A    C
ATOM   3180  CA  PRO A 432      41.138  45.159  84.775  1.00  8.05      A    C
ATOM   3187  CA  ALA A 433      41.670  42.554  82.030  1.00  7.78      A    C
ATOM   3192  CA  LEU A 434      41.696  39.872  84.724  1.00  7.30      A    C
ATOM   3200  CA  GLN A 435      38.424  41.359  86.060  1.00  9.52      A    C
ATOM   3209  CA  ARG A 436      36.760  41.034  82.610  1.00  8.83      A    C
ATOM   3220  CA  LEU A 437      38.067  37.458  82.349  1.00  8.34      A    C
ATOM   3228  CA  VAL A 438      36.448  36.553  85.664  1.00  9.83      A    C
ATOM   3235  CA  GLY A 439      33.227  38.399  84.746  1.00  9.87      A    C
ATOM   3239  CA  GLN A 440      30.557  39.769  87.076  1.00 11.48      A    C
ATOM   3248  CA  TYR A 441      27.387  38.162  88.403  1.00  9.32      A    C
ATOM   3260  CA  SER A 442      24.326  39.771  86.747  1.00  9.01      A    C
ATOM   3266  CA  ALA A 443      21.930  41.944  88.804  1.00 10.00      A    C
ATOM   3271  CA  ASP A 444      19.645  38.962  89.504  1.00 11.36      A    C
ATOM   3279  CA  GLU A 445      22.683  36.814  90.345  1.00 10.65      A    C
ATOM   3288  CA  LEU A 446      24.057  39.420  92.752  1.00 12.13      A    C
ATOM   3296  CA  ARG A 447      20.688  39.484  94.519  1.00 14.84      A    C
ATOM   3307  CA  ARG A 448      20.411  35.689  94.956  1.00 13.37      A    C
ATOM   3318  CA  GLY A 449      24.148  34.933  95.289  1.00 11.12      A    C
ATOM   3322  CA  TYR A 450      24.570  32.356  92.478  1.00 11.01      A    C
ATOM   3334  CA  VAL A 451      24.024  31.514  88.794  1.00 10.44      A    C
ATOM   3341  CA  ALA A 452      20.597  29.813  88.469  1.00 11.06      A    C
ATOM   3346  CA  GLY A 453      20.832  26.118  87.493  1.00 10.92      A    C
ATOM   3350  CA  ASN A 454      24.597  26.415  86.920  1.00  9.68      A    C
ATOM   3358  CA  PRO A 455      26.517  25.184  90.002  1.00  9.01      A    C
ATOM   3365  CA  GLY A 456      29.710  24.813  87.923  1.00  7.30      A    C
ATOM   3369  CA  ALA A 457      29.902  28.429  86.804  1.00  5.98      A    C
ATOM   3374  CA  SER A 458      29.216  29.530  90.428  1.00  5.00      A    C
ATOM   3380  CA  VAL A 459      32.021  27.538  92.097  1.00  3.77      A    C
ATOM   3387  CA  GLN A 460      34.377  28.579  89.287  1.00  3.69      A    C
ATOM   3396  CA  ARG A 461      33.738  32.298  89.725  1.00  3.05      A    C
ATOM   3407  CA  GLY A 462      34.143  31.742  93.464  1.00  2.45      A    C
ATOM   3411  CA  TYR A 463      37.707  30.519  92.994  1.00  3.51      A    C
ATOM   3423  CA  ASP A 464      38.419  33.243  90.372  1.00  3.80      A    C
```

도면34h

```
ATOM   3431  CA  GLN A 465      37.202  35.909  92.790  1.00  4.82      A    C
ATOM   3440  CA  TYR A 466      39.300  34.526  95.621  1.00  5.23      A    C
ATOM   3452  CA  GLY A 467      42.325  34.871  93.265  1.00  5.67      A    C
ATOM   3456  CA  LEU A 468      41.385  38.365  92.150  1.00  6.17      A    C
ATOM   3464  CA  SER A 469      40.991  39.304  95.828  1.00  5.97      A    C
ATOM   3470  CA  VAL A 470      44.682  38.455  96.482  1.00  6.48      A    C
ATOM   3477  CA  ILE A 471      45.705  40.579  93.461  1.00  6.17      A    C
ATOM   3485  CA  ALA A 472      43.411  43.420  94.600  1.00  8.07      A    C
ATOM   3490  CA  ASP A 473      45.060  43.606  98.080  1.00 10.08      A    C
ATOM   3498  CA  GLU A 474      48.496  43.569  96.443  1.00 10.69      A    C
ATOM   3507  CA  LEU A 475      47.508  46.603  94.342  1.00 11.15      A    C
ATOM   3515  CA  GLY A 476      46.177  48.346  97.453  1.00 13.34      A    C
ATOM   3519  CA  LEU A 477      42.538  48.006  96.344  1.00 15.39      A    C
ATOM   3527  CA  THR A 478      41.434  47.034  99.867  1.00 17.62      A    C
ATOM   3534  CA  GLU A 479      37.723  47.742  99.394  1.00 18.74      A    C
ATOM   3543  CA  GLU A 480      37.595  45.763  96.110  1.00 16.33      A    C
ATOM   3552  CA  ALA A 481      39.372  42.864  97.813  1.00 13.75      A    C
ATOM   3557  CA  GLU A 482      36.715  42.720 100.595  1.00 14.63      A    C
ATOM   3566  CA  THR A 483      33.842  42.673  98.101  1.00 12.95      A    C
ATOM   3573  CA  LEU A 484      35.614  39.932  96.121  1.00 11.73      A    C
ATOM   3581  CA  ARG A 485      36.127  37.758  99.233  1.00 11.05      A    C
ATOM   3592  CA  GLU A 486      32.458  38.265 100.146  1.00 10.61      A    C
ATOM   3601  CA  GLN A 487      31.331  37.147  96.664  1.00  8.60      A    C
ATOM   3610  CA  ALA A 488      33.786  34.225  96.759  1.00  6.52      A    C
ATOM   3615  CA  SER A 489      32.029  32.893  99.875  1.00  7.12      A    C
ATOM   3621  CA  TRP A 490      28.661  32.726  98.010  1.00  7.33      A    C
ATOM   3635  CA  PRO A 491      28.980  29.421  96.072  1.00  6.73      A    C
ATOM   3642  CA  ILE A 492      29.748  27.439  99.257  1.00  8.58      A    C
ATOM   3650  CA  GLU A 493      27.034  29.091 101.368  1.00 11.10      A    C
ATOM   3659  CA  LYS A 494      24.283  29.217  98.744  1.00 10.38      A    C
ATOM   3668  CA  LEU A 495      24.772  25.992  96.727  1.00  8.08      A    C
ATOM   3676  CA  THR A 496      25.642  23.390  99.397  1.00  9.94      A    C
ATOM   3683  CA  LYS A 497      22.269  21.696  99.874  1.00 10.32      A    C
ATOM   3692  CA  PRO A 498      21.844  20.282 103.387  1.00  9.82      A    C
ATOM   3699  CA  GLY A 499      21.074  16.559 103.567  1.00 10.14      A    C
ATOM   3703  CA  ALA A 500      21.838  16.042  99.846  1.00 10.82      A    C
ATOM   3708  CA  TRP A 501      23.386  12.728 100.863  1.00 11.71      A    C
ATOM   3722  CA  THR A 502      22.581  10.696 104.006  1.00 12.62      A    C
ATOM   3729  CA  ALA A 503      25.365   8.647 105.597  1.00 13.78      A    C
ATOM   3734  CA  ALA A 504      24.951   4.994 106.737  1.00 15.02      A    C
ATOM   3739  CA  ASP A 505      24.357   6.130 110.325  1.00 14.83      A    C
ATOM   3747  CA  GLY A 506      21.856   8.861 109.305  1.00 12.52      A    C
ATOM   3751  CA  THR A 507      24.298  11.810 109.343  1.00 10.12      A    C
ATOM   3758  CA  GLN A 508      23.119  14.588 106.989  1.00  8.56      A    C
ATOM   3767  CA  VAL A 509      25.765  15.604 104.451  1.00  6.31      A    C
ATOM   3774  CA  GLY A 510      25.468  18.758 102.362  1.00  5.03      A    C
ATOM   3778  CA  LEU A 511      26.594  18.750  98.722  1.00  4.48      A    C
ATOM   3786  CA  LEU A 512      26.884  21.120  95.793  1.00  5.45      A    C
ATOM   3794  CA  THR A 513      23.470  21.012  94.071  1.00  7.09      A    C
ATOM   3801  CA  PRO A 514      21.941  23.109  91.250  1.00  7.88      A    C
ATOM   3808  CA  ARG A 515      19.517  25.737  92.573  1.00  9.63      A    C
ATOM   3819  CA  ALA A 516      16.670  27.438  90.743  1.00 11.40      A    C
ATOM   3824  CA  ALA A 517      16.201  31.213  90.509  1.00 13.83      A    C
ATOM   3829  CA  ASP A 518      13.518  31.125  93.261  1.00 16.58      A    C
ATOM   3837  CA  GLY A 519      15.865  29.277  95.648  1.00 16.47      A    C
ATOM   3841  CA  SER A 520      14.202  25.852  95.069  1.00 15.99      A    C
ATOM   3847  CA  TRP A 521      16.785  23.089  94.870  1.00 13.14      A    C
ATOM   3861  CA  GLN A 522      16.878  21.329  91.494  1.00 12.01      A    C
ATOM   3870  CA  SER A 523      16.780  17.550  91.127  1.00 12.74      A    C
ATOM   3876  CA  ALA A 524      20.170  15.972  90.267  1.00 11.18      A    C
ATOM   3881  CA  ASP A 525      22.087  12.671  90.488  1.00  8.84      A    C
ATOM   3889  CA  HIS A 526      25.271  13.887  92.225  1.00  6.57      A    C
ATOM   3899  CA  ALA A 527      27.293  11.003  90.717  1.00  6.14      A    C
ATOM   3904  CA  LYS A 528      26.151  11.481  87.106  1.00  7.50      A    C
ATOM   3913  CA  PHE A 529      28.818  12.767  84.690  1.00  7.13      A    C
ATOM   3924  CA  GLU A 530      27.867  16.202  83.304  1.00  7.91      A    C
```

도면34i

```
ATOM   3933  CA  ALA A 531      24.533  16.513  85.106  1.00  8.44      A    C
ATOM   3938  CA  ALA A 532      23.397  20.032  86.170  1.00  7.77      A    C
ATOM   3943  CA  GLY A 533      25.194  21.889  83.332  1.00  6.76      A    C
ATOM   3947  CA  LEU A 534      28.619  20.734  84.578  1.00  5.97      A    C
ATOM   3955  CA  TYR A 535      31.682  20.411  82.372  1.00  8.46      A    C
ATOM   3967  CA  GLN A 536      33.627  17.123  82.410  1.00  6.34      A    C
ATOM   3976  CA  GLY A 537      32.521  16.050  85.901  1.00  4.04      A    C
ATOM   3980  CA  THR A 538      29.936  15.281  88.564  1.00  3.50      A    C
ATOM   3987  CA  LEU A 539      28.500  17.232  91.498  1.00  3.71      A    C
ATOM   3995  CA  TRP A 540      30.500  15.072  93.986  1.00  3.80      A    C
ATOM   4009  CA  GLN A 541      33.680  16.012  92.073  1.00  3.42      A    C
ATOM   4018  CA  TYR A 542      32.808  19.714  91.708  1.00  4.03      A    C
ATOM   4030  CA  HIS A 543      31.606  19.795  95.352  1.00  4.44      A    C
ATOM   4040  CA  TRP A 544      35.221  20.315  96.537  1.00  4.40      A    C
ATOM   4054  CA  TYR A 545      36.045  23.019  93.950  1.00  4.18      A    C
ATOM   4066  CA  ASP A 546      36.276  26.024  96.237  1.00  5.25      A    C
ATOM   4074  CA  ALA A 547      39.546  24.793  97.671  1.00  5.51      A    C
ATOM   4079  CA  TYR A 548      40.358  28.217  99.183  1.00  6.10      A    C
ATOM   4091  CA  ASP A 549      37.753  27.586 101.917  1.00  6.22      A    C
ATOM   4099  CA  MET A 550      37.554  24.017 103.177  1.00  6.61      A    C
ATOM   4107  CA  ASP A 551      36.549  25.382 106.586  1.00  8.26      A    C
ATOM   4115  CA  ALA A 552      33.422  27.067 105.225  1.00  7.80      A    C
ATOM   4120  CA  LEU A 553      32.800  24.021 103.028  1.00  7.77      A    C
ATOM   4128  CA  VAL A 554      32.983  21.675 106.066  1.00  8.48      A    C
ATOM   4135  CA  GLU A 555      30.421  23.906 107.886  1.00 10.05      A    C
ATOM   4144  CA  ALA A 556      28.101  24.172 104.869  1.00  8.14      A    C
ATOM   4149  CA  MET A 557      28.262  20.355 104.586  1.00  7.75      A    C
ATOM   4157  CA  GLY A 558      26.833  20.080 108.112  1.00  7.77      A    C
ATOM   4161  CA  GLY A 559      29.995  19.989 110.230  1.00  8.36      A    C
ATOM   4165  CA  HIS A 560      33.086  17.891 110.775  1.00 10.24      A    C
ATOM   4175  CA  GLU A 561      31.496  14.448 111.036  1.00 10.61      A    C
ATOM   4184  CA  ALA A 562      29.420  15.118 107.900  1.00  9.17      A    C
ATOM   4189  CA  ALA A 563      32.594  16.099 105.943  1.00  7.47      A    C
ATOM   4194  CA  ARG A 564      34.534  13.129 107.304  1.00  6.15      A    C
ATOM   4205  CA  LEU A 565      31.813  10.683 106.236  1.00  5.27      A    C
ATOM   4213  CA  GLY A 566      31.408  12.469 102.857  1.00  4.45      A    C
ATOM   4217  CA  MET A 567      35.136  12.086 102.137  1.00  5.05      A    C
ATOM   4225  CA  ARG A 568      35.000   8.412 103.020  1.00  6.55      A    C
ATOM   4236  CA  HIS A 569      31.985   7.943 100.689  1.00  7.31      A    C
ATOM   4246  CA  MET A 570      33.786   9.924  97.906  1.00  7.92      A    C
ATOM   4254  CA  PHE A 571      36.388   7.150  97.840  1.00  7.04      A    C
ATOM   4265  CA  GLY A 572      34.058   4.168  98.426  1.00  8.17      A    C
ATOM   4269  CA  GLU A 573      35.790   3.316 101.736  1.00 10.14      A    C
ATOM   4278  CA  HIS A 574      32.693   1.377 102.888  1.00 11.27      A    C
ATOM   4288  CA  ALA A 575      32.914  -0.845  99.770  1.00 11.91      A    C
ATOM   4293  CA  PRO A 576      36.615  -1.198  98.736  1.00 13.14      A    C
ATOM   4300  CA  ASP A 577      35.923  -3.909  96.131  1.00 13.69      A    C
ATOM   4308  CA  ASP A 578      32.994  -2.108  94.404  1.00 12.41      A    C
ATOM   4316  CA  GLY A 579      34.247  -0.097  91.423  1.00 11.12      A    C
ATOM   4320  CA  LYS A 580      30.905   1.748  91.259  1.00 11.21      A    C
ATOM   4329  CA  ALA A 581      31.370   3.276  94.757  1.00  8.19      A    C
ATOM   4334  CA  MET A 582      34.358   5.303  93.506  1.00  5.86      A    C
ATOM   4342  CA  LEU A 583      33.022   8.889  93.198  1.00  4.23      A    C
ATOM   4350  CA  HIS A 584      36.434  10.466  92.671  1.00  3.37      A    C
ATOM   4360  CA  SER A 585      37.816  10.665  89.125  1.00  3.12      A    C
ATOM   4366  CA  ASN A 586      41.508  10.091  88.359  1.00  3.90      A    C
ATOM   4374  CA  ALA A 587      41.103  10.824  84.638  1.00  5.72      A    C
ATOM   4379  CA  ASN A 588      40.425  14.577  84.853  1.00  7.03      A    C
ATOM   4387  CA  GLU A 589      41.649  17.614  86.824  1.00  7.17      A    C
ATOM   4396  CA  ILE A 590      38.295  18.647  88.338  1.00  7.24      A    C
ATOM   4404  CA  ASP A 591      38.863  16.848  91.650  1.00  7.24      A    C
ATOM   4412  CA  LEU A 592      42.668  16.421  91.471  1.00  6.70      A    C
ATOM   4420  CA  GLN A 593      42.981  17.626  95.053  1.00  4.59      A    C
ATOM   4429  CA  ALA A 594      40.356  15.206  96.428  1.00  4.71      A    C
ATOM   4434  CA  PRO A 595      42.784  12.503  97.633  1.00  5.03      A    C
ATOM   4441  CA  TYR A 596      44.506  15.069  99.845  1.00  6.02      A    C
```

도면34j

```
ATOM   4453  CA  LEU A 597      41.304  16.226 101.591  1.00  4.94      A    C
ATOM   4461  CA  PHE A 598      41.417  13.474 104.213  1.00  5.47      A    C
ATOM   4472  CA  ASN A 599      44.037  15.684 105.929  1.00  6.53      A    C
ATOM   4480  CA  TYR A 600      41.176  18.148 106.491  1.00  7.92      A    C
ATOM   4492  CA  THR A 601      38.801  15.556 107.938  1.00  9.25      A    C
ATOM   4499  CA  GLY A 602      41.334  14.393 110.555  1.00  9.27      A    C
ATOM   4503  CA  GLU A 603      42.259  11.247 108.634  1.00  8.63      A    C
ATOM   4512  CA  PRO A 604      45.684  11.982 107.132  1.00  7.11      A    C
ATOM   4519  CA  SER A 605      46.475   8.224 106.915  1.00  5.46      A    C
ATOM   4525  CA  LEU A 606      43.781   7.928 104.201  1.00  4.96      A    C
ATOM   4533  CA  THR A 607      45.390  10.729 102.175  1.00  4.57      A    C
ATOM   4540  CA  GLN A 608      48.682   8.824 102.399  1.00  5.02      A    C
ATOM   4549  CA  LYS A 609      47.008   5.602 101.265  1.00  5.44      A    C
ATOM   4558  CA  TRP A 610      45.214   7.208  98.331  1.00  5.42      A    C
ATOM   4572  CA  ALA A 611      48.211   9.266  97.205  1.00  6.31      A    C
ATOM   4577  CA  ARG A 612      50.436   6.184  97.099  1.00  5.71      A    C
ATOM   4588  CA  ALA A 613      47.643   4.074  95.492  1.00  4.20      A    C
ATOM   4593  CA  ILE A 614      46.445   6.379  92.691  1.00  4.67      A    C
ATOM   4601  CA  TYR A 615      49.973   7.212  91.557  1.00  4.85      A    C
ATOM   4613  CA  THR A 616      51.897   3.953  92.071  1.00  4.62      A    C
ATOM   4620  CA  LYS A 617      49.365   1.088  92.368  1.00  6.36      A    C
ATOM   4629  CA  GLU A 618      46.132  -0.236  90.829  1.00  7.91      A    C
ATOM   4638  CA  THR A 619      42.939   1.583  91.795  1.00  6.88      A    C
ATOM   4645  CA  TRP A 620      39.227   1.357  91.008  1.00  7.46      A    C
ATOM   4659  CA  ASN A 621      38.296   3.692  88.170  1.00  6.73      A    C
ATOM   4667  CA  ARG A 622      34.663   4.574  87.576  1.00  4.87      A    C
ATOM   4678  CA  TYR A 623      35.087   7.709  85.473  1.00  5.03      A    C
ATOM   4690  CA  ILE A 624      36.716   8.661  82.149  1.00  5.64      A    C
ATOM   4698  CA  ALA A 625      37.316  12.291  81.127  1.00  6.35      A    C
ATOM   4703  CA  THR A 626      34.590  12.516  78.472  1.00  6.19      A    C
ATOM   4710  CA  GLY A 627      31.208  10.937  77.589  1.00  5.93      A    C
ATOM   4714  CA  SER A 628      32.592   7.840  75.913  1.00  7.00      A    C
ATOM   4720  CA  SER A 629      35.720   6.302  74.355  1.00  9.77      A    C
ATOM   4726  CA  SER A 630      36.292   3.632  71.688  1.00 11.33      A    C
ATOM   4732  CA  ALA A 631      39.492   2.767  73.559  1.00 10.90      A    C
ATOM   4737  CA  VAL A 632      37.659   1.102  76.492  1.00  8.53      A    C
ATOM   4744  CA  PRO A 633      34.094  -0.107  77.097  1.00  7.28      A    C
ATOM   4751  CA  SER A 634      32.472   3.099  78.345  1.00  5.04      A    C
ATOM   4757  CA  GLY A 635      29.414   5.345  78.326  1.00  4.96      A    C
ATOM   4761  CA  GLY A 636      27.832   8.244  80.176  1.00  6.17      A    C
ATOM   4765  CA  GLY A 637      31.304   9.416  81.298  1.00  5.99      A    C
ATOM   4769  CA  GLU A 638      32.099   6.071  82.923  1.00  5.01      A    C
ATOM   4778  CA  PHE A 639      34.109   2.868  82.500  1.00  4.24      A    C
ATOM   4789  CA  THR A 640      31.342   0.261  81.956  1.00  5.80      A    C
ATOM   4796  CA  PRO A 641      32.109  -1.741  83.969  1.00  6.76      A    C
ATOM   4803  CA  PRO A 642      34.393   0.103  86.417  1.00  8.12      A    C
ATOM   4810  CA  LEU A 643      38.016  -0.971  85.969  1.00  9.82      A    C
ATOM   4818  CA  LYS A 644      40.825  -1.805  88.435  1.00 11.06      A    C
ATOM   4827  CA  THR A 645      43.998  -0.565  86.793  1.00  8.37      A    C
ATOM   4834  CA  LYS A 646      47.067   1.637  87.273  1.00  7.54      A    C
ATOM   4843  CA  VAL A 647      46.307   5.272  86.488  1.00  5.75      A    C
ATOM   4850  CA  TYR A 648      49.922   5.925  85.527  1.00  5.56      A    C
ATOM   4862  CA  ARG A 649      51.964   3.328  83.533  1.00  7.76      A    C
ATOM   4873  CA  LEU A 650      55.512   3.322  82.177  1.00  6.76      A    C
ATOM   4881  CA  ASP A 651      54.026   2.407  78.809  1.00  6.42      A    C
ATOM   4889  CA  PRO A 652      53.043   4.126  75.570  1.00  5.92      A    C
ATOM   4896  CA  ARG A 653      49.549   3.680  77.020  1.00  4.60      A    C
ATOM   4907  CA  GLY A 654      50.724   6.058  79.774  1.00  4.89      A    C
ATOM   4911  CA  MET A 655      47.517   7.356  81.356  1.00  4.93      A    C
ATOM   4919  CA  LEU A 656      43.963   6.075  81.567  1.00  3.04      A    C
ATOM   4927  CA  PRO A 657      42.105   6.134  78.251  1.00  3.10      A    C
ATOM   4934  CA  THR A 658      40.754   9.734  77.676  1.00  3.53      A    C
ATOM   4941  CA  MET A 659      43.088  11.046  80.398  1.00  3.65      A    C
ATOM   4949  CA  ASP A 660      44.895  13.859  78.555  1.00  3.95      A    C
ATOM   4957  CA  ASN A 661      48.031  15.161  80.276  1.00  3.15      A    C
ATOM   4965  CA  ASP A 662      46.843  18.761  79.522  1.00  2.87      A    C
```

도면34k

```
ATOM   4973  CA  ALA A 663      50.075  20.739  79.098  1.00  3.12       A    C
ATOM   4978  CA  GLY A 664      51.894  18.459  81.548  1.00  2.70       A    C
ATOM   4982  CA  THR A 665      49.328  18.721  84.376  1.00  2.81       A    C
ATOM   4989  CA  MET A 666      48.910  14.993  85.026  1.00  2.55       A    C
ATOM   4997  CA  SER A 667      52.753  14.620  84.897  1.00  2.87       A    C
ATOM   5003  CA  THR A 668      53.193  17.564  87.294  1.00  3.50       A    C
ATOM   5010  CA  MET A 669      50.719  16.032  89.771  1.00  2.90       A    C
ATOM   5018  CA  PHE A 670      52.751  12.810  89.639  1.00  2.90       A    C
ATOM   5029  CA  VAL A 671      55.921  14.716  90.519  1.00  3.29       A    C
ATOM   5036  CA  ALA A 672      54.049  16.446  93.385  1.00  3.04       A    C
ATOM   5041  CA  ALA A 673      52.740  13.126  94.733  1.00  3.21       A    C
ATOM   5046  CA  ALA A 674      56.289  11.665  94.548  1.00  4.16       A    C
ATOM   5051  CA  VAL A 675      57.770  14.572  96.510  1.00  3.68       A    C
ATOM   5058  CA  GLY A 676      54.729  14.329  98.773  1.00  3.81       A    C
ATOM   5062  CA  LEU A 677      53.628  17.976  98.689  1.00  3.63       A    C
ATOM   5070  CA  PHE A 678      50.401  18.865  96.910  1.00  4.98       A    C
ATOM   5081  CA  PRO A 679      48.736  22.236  96.308  1.00  4.89       A    C
ATOM   5088  CA  VAL A 680      45.131  21.802  97.449  1.00  3.54       A    C
ATOM   5095  CA  THR A 681      44.815  25.594  97.517  1.00  3.01       A    C
ATOM   5102  CA  ALA A 682      46.410  27.062  94.421  1.00  2.78       A    C
ATOM   5107  CA  GLY A 683      46.558  30.790  95.067  1.00  4.10       A    C
ATOM   5111  CA  SER A 684      47.895  30.234  98.585  1.00  5.30       A    C
ATOM   5117  CA  SER A 685      51.390  29.528  99.913  1.00  6.30       A    C
ATOM   5123  CA  GLN A 686      50.476  26.083 101.310  1.00  5.66       A    C
ATOM   5132  CA  PHE A 687      51.101  22.447 100.320  1.00  5.37       A    C
ATOM   5143  CA  GLN A 688      49.438  19.356 101.748  1.00  4.94       A    C
ATOM   5152  CA  VAL A 689      51.517  16.390 102.936  1.00  5.79       A    C
ATOM   5159  CA  GLY A 690      50.928  13.088 101.117  1.00  6.69       A    C
ATOM   5163  CA  SER A 691      52.974   9.889 100.906  1.00  7.27       A    C
ATOM   5169  CA  PRO A 692      56.528  10.626  99.669  1.00  6.85       A    C
ATOM   5176  CA  PHE A 693      57.581   7.765  97.347  1.00  6.34       A    C
ATOM   5187  CA  PHE A 694      61.346   7.747  97.796  1.00  6.87       A    C
ATOM   5198  CA  ASP A 695      63.891   7.202 100.575  1.00  7.56       A    C
ATOM   5206  CA  SER A 696      65.365  10.632  99.830  1.00  6.91       A    C
ATOM   5212  CA  THR A 697      64.278  13.400  97.424  1.00  7.09       A    C
ATOM   5219  CA  THR A 698      66.465  16.572  97.123  1.00  7.46       A    C
ATOM   5226  CA  ILE A 699      65.459  19.752  95.308  1.00  6.95       A    C
ATOM   5234  CA  THR A 700      68.553  21.914  94.670  1.00  8.40       A    C
ATOM   5241  CA  TYR A 701      67.913  25.630  94.091  1.00  9.41       A    C
ATOM   5253  CA  ASP A 702      69.971  28.028  91.946  1.00 11.48       A    C
ATOM   5261  CA  ASP A 703      72.068  29.199  94.915  1.00 13.97       A    C
ATOM   5269  CA  GLY A 704      73.018  25.617  95.847  1.00 13.70       A    C
ATOM   5273  CA  SER A 705      70.765  25.419  98.907  1.00 12.46       A    C
ATOM   5279  CA  ALA A 706      68.419  22.383  98.854  1.00 10.18       A    C
ATOM   5284  CA  PHE A 707      65.052  21.131 100.137  1.00  7.55       A    C
ATOM   5295  CA  THR A 708      65.351  17.581 101.387  1.00  7.30       A    C
ATOM   5302  CA  VAL A 709      62.397  15.352 102.091  1.00  7.65       A    C
ATOM   5309  CA  THR A 710      63.206  11.874 103.424  1.00  8.06       A    C
ATOM   5316  CA  ALA A 711      60.926   8.889 103.973  1.00  8.76       A    C
ATOM   5321  CA  ASP A 712      62.785   6.541 106.270  1.00 10.03       A    C
ATOM   5329  CA  GLY A 713      61.790   2.931 105.687  1.00  7.18       A    C
ATOM   5333  CA  VAL A 714      59.420   3.754 102.770  1.00  5.21       A    C
ATOM   5340  CA  SER A 715      58.461   0.754 100.621  1.00  5.75       A    C
ATOM   5346  CA  GLU A 716      55.621  -0.434  98.399  1.00  7.40       A    C
ATOM   5355  CA  ASP A 717      53.911  -1.492 101.619  1.00  7.77       A    C
ATOM   5363  CA  ALA A 718      55.257   1.143 103.999  1.00  5.28       A    C
ATOM   5368  CA  PHE A 719      53.472   4.224 102.665  1.00  3.86       A    C
ATOM   5379  CA  TYR A 720      51.935   5.573 105.926  1.00  5.07       A    C
ATOM   5391  CA  VAL A 721      53.647   8.305 107.943  1.00  6.16       A    C
ATOM   5398  CA  GLN A 722      54.378   7.196 111.544  1.00  8.68       A    C
ATOM   5407  CA  SER A 723      56.107  10.362 112.750  1.00  9.44       A    C
ATOM   5413  CA  ALA A 724      57.936  13.385 111.311  1.00  8.65       A    C
ATOM   5418  CA  THR A 725      60.561  16.054 112.006  1.00  8.30       A    C
ATOM   5425  CA  LEU A 726      61.170  19.486 110.467  1.00  8.74       A    C
ATOM   5433  CA  ASP A 727      64.804  20.662 110.723  1.00  9.93       A    C
ATOM   5441  CA  GLY A 728      65.330  18.182 113.597  1.00 11.60       A    C
```

도면341

```
ATOM   5445  CA  ALA A 729      62.241  19.122 115.635  1.00 11.25      A    C
ATOM   5450  CA  THR A 730      58.993  17.157 116.102  1.00 10.67      A    C
ATOM   5457  CA  PHE A 731      56.579  18.068 113.320  1.00  9.05      A    C
ATOM   5468  CA  GLY A 732      52.846  17.319 113.474  1.00  8.76      A    C
ATOM   5472  CA  ASN A 733      51.196  19.402 110.742  1.00  8.99      A    C
ATOM   5480  CA  THR A 734      49.869  17.843 107.528  1.00  9.35      A    C
ATOM   5487  CA  TRP A 735      50.725  20.969 105.477  1.00  8.91      A    C
ATOM   5501  CA  VAL A 736      53.901  22.986 104.955  1.00  8.70      A    C
ATOM   5508  CA  ASP A 737      54.483  26.555 103.781  1.00  8.20      A    C
ATOM   5516  CA  TYR A 738      56.186  27.338 100.476  1.00  6.61      A    C
ATOM   5528  CA  ALA A 739      58.753  29.427 102.391  1.00  6.41      A    C
ATOM   5533  CA  THR A 740      59.787  26.356 104.396  1.00  7.06      A    C
ATOM   5540  CA  VAL A 741      60.397  24.354 101.242  1.00  6.62      A    C
ATOM   5547  CA  VAL A 742      62.422  26.857  99.172  1.00  7.73      A    C
ATOM   5554  CA  GLY A 743      64.172  27.894 102.414  1.00  9.90      A    C
ATOM   5558  CA  GLY A 744      65.997  24.543 102.207  1.00  9.66      A    C
ATOM   5562  CA  ALA A 745      64.348  22.686 105.078  1.00  9.46      A    C
ATOM   5567  CA  ASP A 746      64.950  19.052 106.107  1.00  9.46      A    C
ATOM   5575  CA  LEU A 747      61.574  17.315 106.252  1.00  7.98      A    C
ATOM   5583  CA  ALA A 748      62.092  13.775 107.535  1.00  8.59      A    C
ATOM   5588  CA  PHE A 749      59.312  11.203 107.659  1.00  9.99      A    C
ATOM   5599  CA  ARG A 750      59.298   7.807 109.398  1.00 11.09      A    C
ATOM   5610  CA  MET A 751      57.231   5.392 107.265  1.00  8.36      A    C
ATOM   5618  CA  GLY A 752      55.426   2.175 108.214  1.00  8.95      A    C
ATOM   5622  CA  GLU A 753      53.025  -0.439 106.937  1.00 10.36      A    C
ATOM   5631  CA  GLN A 754      50.077   0.533 109.170  1.00 10.20      A    C
ATOM   5640  CA  PRO A 755      48.009   3.743 109.500  1.00  9.59      A    C
ATOM   5647  CA  SER A 756      49.021   5.965 112.430  1.00 10.96      A    C
ATOM   5653  CA  ASP A 757      47.455   8.970 114.180  1.00 12.99      A    C
ATOM   5661  CA  TRP A 758      50.369  11.238 113.073  1.00 11.54      A    C
ATOM   5675  CA  GLY A 759      49.073  14.673 112.067  1.00 12.17      A    C
ATOM   5679  CA  THR A 760      45.608  14.275 113.614  1.00 12.83      A    C
ATOM   5686  CA  ASP A 761      46.374  17.270 115.851  1.00 13.13      A    C
ATOM   5694  CA  THR A 762      47.486  19.473 112.939  1.00 10.45      A    C
ATOM   5701  CA  ALA A 763      47.395  23.279 113.167  1.00  9.30      A    C
ATOM   5706  CA  PRO A 764      44.545  24.274 110.760  1.00  8.86      A    C
ATOM   5713  CA  ALA A 765      45.404  24.788 107.092  1.00  8.11      A    C
ATOM   5718  CA  PHE A 766      44.495  27.881 105.084  1.00  9.09      A    C
ATOM   5729  CA  SER A 767      40.845  28.777 104.561  1.00  9.56      A    C
ATOM   5735  CA  MET A 768      39.792  32.266 103.433  1.00 11.12      A    C
ATOM   5743  CA  SER A 769      37.000  32.501 106.061  1.00 14.28      A    C
ATOM   5749  CA  THR A 770      39.217  31.636 109.046  1.00 17.50      A    C
ATOM   5756  CA  ALA A 771      42.527  33.224 107.955  1.00 19.68      A    C
ATOM   5761  CA  ASP B   8      33.160  32.500 141.172  1.00 18.72      B    C
ATOM   5769  CA  TYR B   9      29.570  31.506 140.304  1.00 14.20      B    C
ATOM   5781  CA  ALA B  10      28.422  31.895 143.925  1.00 10.68      B    C
ATOM   5786  CA  SER B  11      29.320  35.614 143.693  1.00 10.07      B    C
ATOM   5792  CA  LEU B  12      26.696  36.062 140.938  1.00  8.36      B    C
ATOM   5800  CA  VAL B  13      23.923  34.801 143.242  1.00  6.94      B    C
ATOM   5807  CA  ASP B  14      21.880  37.475 145.012  1.00  6.81      B    C
ATOM   5815  CA  VAL B  15      20.179  35.726 147.955  1.00  4.17      B    C
ATOM   5822  CA  PHE B  16      18.230  38.889 148.802  1.00  3.40      B    C
ATOM   5833  CA  VAL B  17      16.289  38.686 145.547  1.00  3.53      B    C
ATOM   5840  CA  GLY B  18      12.666  38.026 146.520  1.00  4.03      B    C
ATOM   5844  CA  THR B  19      13.196  38.689 150.262  1.00  5.26      B    C
ATOM   5851  CA  GLU B  20      10.979  41.797 150.516  1.00  5.53      B    C
ATOM   5860  CA  GLY B  21       7.349  41.901 151.577  1.00  6.87      B    C
ATOM   5864  CA  ASP B  22       5.019  39.846 149.439  1.00  6.61      B    C
ATOM   5872  CA  PHE B  23       7.350  39.730 146.400  1.00  5.51      B    C
ATOM   5883  CA  GLY B  24       8.458  36.090 146.362  1.00  4.07      B    C
ATOM   5887  CA  ASN B  25       8.571  34.746 149.915  1.00  2.81      B    C
ATOM   5895  CA  ASP B  26      12.206  33.663 149.535  1.00  3.13      B    C
ATOM   5903  CA  MET B  27      14.936  33.714 152.197  1.00  4.16      B    C
ATOM   5911  CA  PRO B  28      18.521  35.051 152.473  1.00  3.52      B    C
ATOM   5918  CA  ALA B  29      19.101  32.077 154.801  1.00  2.82      B    C
ATOM   5923  CA  ALA B  30      22.436  30.302 154.790  1.00  2.98      B    C
```

도면34m

```
ATOM   5928  CA  GLN B  31      21.896  26.852 153.161  1.00  4.61      B     C
ATOM   5937  CA  ALA B  32      23.671  24.106 151.121  1.00  5.14      B     C
ATOM   5942  CA  PRO B  33      22.280  22.821 147.805  1.00  6.54      B     C
ATOM   5949  CA  ASN B  34      18.900  21.202 148.694  1.00  5.61      B     C
ATOM   5957  CA  GLY B  35      19.876  21.616 152.349  1.00  4.62      B     C
ATOM   5961  CA  LEU B  36      17.780  20.633 155.336  1.00  5.14      B     C
ATOM   5969  CA  ALA B  37      19.361  23.299 157.533  1.00  3.87      B     C
ATOM   5974  CA  LYS B  38      18.219  26.746 156.507  1.00  3.49      B     C
ATOM   5983  CA  VAL B  39      19.767  29.259 158.889  1.00  3.32      B     C
ATOM   5990  CA  ASN B  40      17.444  32.228 158.346  1.00  3.22      B     C
ATOM   5998  CA  PRO B  41      17.676  35.552 160.107  1.00  3.73      B     C
ATOM   6005  CA  ARG B  42      14.314  36.534 161.600  1.00  3.05      B     C
ATOM   6016  CA  THR B  43      13.161  40.161 161.230  1.00  2.90      B     C
ATOM   6023  CA  THR B  44      10.738  41.938 163.588  1.00  5.21      B     C
ATOM   6030  CA  PRO B  45       8.030  43.161 163.977  1.00  6.45      B     C
ATOM   6037  CA  GLY B  46       7.346  42.212 160.323  1.00  4.86      B     C
ATOM   6041  CA  ARG B  47       8.581  39.247 158.248  1.00  4.58      B     C
ATOM   6052  CA  ASN B  48       7.898  37.485 154.983  1.00  4.09      B     C
ATOM   6060  CA  ASN B  49       6.685  33.858 155.217  1.00  4.78      B     C
ATOM   6068  CA  THR B  50      10.246  32.518 155.744  1.00  4.39      B     C
ATOM   6075  CA  GLY B  51      10.802  34.910 158.707  1.00  4.48      B     C
ATOM   6079  CA  TYR B  52      12.692  37.754 157.012  1.00  5.15      B     C
ATOM   6091  CA  ASP B  53      11.399  40.842 155.301  1.00  5.15      B     C
ATOM   6099  CA  TYR B  54      14.064  43.028 153.656  1.00  5.43      B     C
ATOM   6111  CA  ALA B  55      11.984  46.169 154.374  1.00  5.97      B     C
ATOM   6116  CA  GLN B  56      12.446  45.769 158.155  1.00  6.73      B     C
ATOM   6125  CA  SER B  57      15.268  47.086 160.349  1.00  7.18      B     C
ATOM   6131  CA  LYS B  58      15.306  44.770 163.351  1.00  4.81      B     C
ATOM   6140  CA  ILE B  59      16.408  41.138 163.680  1.00  4.41      B     C
ATOM   6148  CA  SER B  60      15.847  38.827 166.673  1.00  4.63      B     C
ATOM   6154  CA  GLY B  61      17.915  35.758 165.791  1.00  3.36      B     C
ATOM   6158  CA  PHE B  62      18.275  32.810 163.426  1.00  3.48      B     C
ATOM   6169  CA  THR B  63      15.870  29.935 162.748  1.00  3.46      B     C
ATOM   6176  CA  HIS B  64      17.063  26.576 161.377  1.00  3.52      B     C
ATOM   6186  CA  THR B  65      14.187  25.010 159.402  1.00  4.13      B     C
ATOM   6193  CA  ASN B  66      12.141  26.467 156.576  1.00  4.29      B     C
ATOM   6201  CA  LEU B  67      10.335  26.228 153.272  1.00  4.24      B     C
ATOM   6209  CA  ASP B  68      11.312  28.351 150.252  1.00  4.81      B     C
ATOM   6217  CA  GLY B  69       8.769  30.598 148.558  1.00  4.56      B     C
ATOM   6221  CA  VAL B  70       5.589  29.149 150.102  1.00  3.51      B     C
ATOM   6228  CA  GLY B  71       2.243  30.766 151.033  1.00  4.39      B     C
ATOM   6232  CA  GLY B  72       0.165  30.889 154.256  1.00  5.93      B     C
ATOM   6236  CA  SER B  73       2.360  32.159 157.086  1.00  5.71      B     C
ATOM   6242  CA  GLY B  74       5.232  29.886 155.962  1.00  4.98      B     C
ATOM   6246  CA  GLY B  75       6.527  26.565 157.216  1.00  4.40      B     C
ATOM   6250  CA  GLY B  76       9.371  25.580 159.508  1.00  4.57      B     C
ATOM   6254  CA  GLY B  77      10.953  28.443 161.386  1.00  3.39      B     C
ATOM   6258  CA  ASP B  78      11.986  26.048 164.188  1.00  3.33      B     C
ATOM   6266  CA  LEU B  79      14.888  26.547 166.612  1.00  5.32      B     C
ATOM   6274  CA  LEU B  80      15.348  30.284 167.110  1.00  5.08      B     C
ATOM   6282  CA  VAL B  81      18.851  31.213 168.309  1.00  4.86      B     C
ATOM   6289  CA  VAL B  82      19.084  34.745 169.745  1.00  5.18      B     C
ATOM   6296  CA  PRO B  83      22.375  36.414 170.798  1.00  5.18      B     C
ATOM   6303  CA  THR B  84      22.055  38.823 173.750  1.00  4.92      B     C
ATOM   6310  CA  SER B  85      24.059  40.445 176.538  1.00  3.51      B     C
ATOM   6316  CA  GLY B  86      20.865  40.604 178.572  1.00  3.93      B     C
ATOM   6320  CA  SER B  87      19.752  38.579 181.582  1.00  5.35      B     C
ATOM   6326  CA  TYR B  88      16.498  36.652 182.221  1.00  5.05      B     C
ATOM   6338  CA  THR B  89      14.650  35.335 185.336  1.00  5.80      B     C
ATOM   6345  CA  ALA B  90      11.608  33.987 183.442  1.00  4.94      B     C
ATOM   6350  CA  ARG B  91      10.687  32.820 179.933  1.00  5.95      B     C
ATOM   6361  CA  PRO B  92      11.913  35.587 177.603  1.00  5.93      B     C
ATOM   6368  CA  GLY B  93       9.414  38.047 176.166  1.00  5.97      B     C
ATOM   6372  CA  THR B  94       9.767  37.813 172.371  1.00  6.48      B     C
ATOM   6379  CA  GLY B  95      10.124  41.588 172.055  1.00  4.67      B     C
ATOM   6383  CA  THR B  96      13.359  41.564 174.059  1.00  4.62      B     C
```

도면34n

```
ATOM   6390  CA  TYR B  97      15.098  39.627 171.228  1.00  4.32      B    C
ATOM   6402  CA  ALA B  98      15.015  42.620 168.846  1.00  4.26      B    C
ATOM   6407  CA  HIS B  99      18.311  44.130 167.648  1.00  4.45      B    C
ATOM   6417  CA  PRO B 100      18.684  47.027 165.199  1.00  4.56      B    C
ATOM   6424  CA  PHE B 101      20.304  46.145 161.878  1.00  5.62      B    C
ATOM   6435  CA  SER B 102      21.236  47.834 158.588  1.00  6.73      B    C
ATOM   6441  CA  HIS B 103      21.603  46.324 155.111  1.00  7.29      B    C
ATOM   6451  CA  ASP B 104      24.879  48.336 155.037  1.00  9.91      B    C
ATOM   6459  CA  ASP B 105      26.290  45.946 157.703  1.00 10.20     B    C
ATOM   6467  CA  GLU B 106      24.917  42.703 156.229  1.00  9.23      B    C
ATOM   6476  CA  ASP B 107      26.499  40.286 153.765  1.00  9.46      B    C
ATOM   6484  CA  ALA B 108      25.242  36.970 152.370  1.00  6.84      B    C
ATOM   6489  CA  GLY B 109      25.559  34.488 149.524  1.00  4.80      B    C
ATOM   6493  CA  PRO B 110      25.491  30.714 148.861  1.00  4.09      B    C
ATOM   6500  CA  GLY B 111      26.276  29.034 152.198  1.00  4.03      B    C
ATOM   6504  CA  PHE B 112      26.451  32.010 154.507  1.00  4.63      B    C
ATOM   6515  CA  TYR B 113      24.841  35.051 156.062  1.00  4.43      B    C
ATOM   6527  CA  SER B 114      26.381  37.783 158.200  1.00  5.09      B    C
ATOM   6533  CA  VAL B 115      24.962  40.865 159.954  1.00  4.89      B    C
ATOM   6540  CA  GLY B 116      25.874  43.467 162.573  1.00  6.35      B    C
ATOM   6544  CA  LEU B 117      23.215  43.427 165.294  1.00  6.43      B    C
ATOM   6552  CA  GLY B 118      22.939  46.299 167.802  1.00  6.20      B    C
ATOM   6556  CA  ASN B 119      23.960  44.928 171.238  1.00  5.86      B    C
ATOM   6564  CA  VAL B 120      21.162  44.472 173.803  1.00  4.83      B    C
ATOM   6571  CA  ALA B 121      21.260  44.132 177.613  1.00  4.23      B    C
ATOM   6576  CA  GLY B 122      19.365  44.559 180.892  1.00  4.26      B    C
ATOM   6580  CA  THR B 123      16.919  42.174 182.582  1.00  5.43      B    C
ATOM   6587  CA  ASP B 124      13.711  40.546 181.298  1.00  5.91      B    C
ATOM   6595  CA  GLY B 125      11.132  43.006 179.823  1.00  6.02      B    C
ATOM   6599  CA  ALA B 126      13.463  45.973 180.415  1.00  5.90      B    C
ATOM   6604  CA  ILE B 127      16.092  44.620 177.976  1.00  5.52      B    C
ATOM   6612  CA  THR B 128      17.063  47.192 175.369  1.00  4.75      B    C
ATOM   6619  CA  GLY B 129      19.991  48.703 173.407  1.00  4.79      B    C
ATOM   6623  CA  ALA B 130      23.437  48.378 175.037  1.00  5.28      B    C
ATOM   6628  CA  PRO B 131      26.932  49.536 173.894  1.00  5.84      B    C
ATOM   6635  CA  GLY B 132      28.389  48.225 170.633  1.00  5.70      B    C
ATOM   6639  CA  THR B 133      27.548  45.592 168.062  1.00  6.06      B    C
ATOM   6646  CA  ILE B 134      27.013  41.848 168.174  1.00  7.11      B    C
ATOM   6654  CA  GLU B 135      28.833  40.654 165.045  1.00  8.84      B    C
ATOM   6663  CA  ALA B 136      26.820  37.651 163.802  1.00  6.10      B    C
ATOM   6668  CA  GLU B 137      27.966  35.088 161.226  1.00  4.63      B    C
ATOM   6677  CA  VAL B 138      26.200  31.897 160.196  1.00  3.25      B    C
ATOM   6684  CA  ALA B 139      26.905  29.212 157.622  1.00  3.25      B    C
ATOM   6689  CA  ALA B 140      25.244  25.933 156.657  1.00  4.37      B    C
ATOM   6694  CA  ALA B 141      26.177  22.444 155.558  1.00  5.98      B    C
ATOM   6699  CA  THR B 142      23.483  19.964 154.381  1.00  4.34      B    C
ATOM   6706  CA  ARG B 143      22.140  19.123 157.865  1.00  4.38      B    C
ATOM   6717  CA  SER B 144      23.954  21.728 159.957  1.00  4.49      B    C
ATOM   6723  CA  GLY B 145      23.839  25.419 160.895  1.00  5.05      B    C
ATOM   6727  CA  VAL B 146      26.994  26.868 162.414  1.00  4.62      B    C
ATOM   6734  CA  HIS B 147      27.242  30.186 164.272  1.00  4.38      B    C
ATOM   6744  CA  ARG B 148      30.103  32.540 165.208  1.00  4.25      B    C
ATOM   6755  CA  TYR B 149      29.335  35.704 167.264  1.00  4.46      B    C
ATOM   6767  CA  ALA B 150      31.522  38.561 168.553  1.00  5.60      B    C
ATOM   6772  CA  PHE B 151      30.094  40.472 171.553  1.00  6.42      B    C
ATOM   6783  CA  PRO B 152      31.380  43.820 172.738  1.00  8.35      B    C
ATOM   6790  CA  ALA B 153      34.133  43.586 175.369  1.00  8.99      B    C
ATOM   6795  CA  GLY B 154      32.667  43.385 178.865  1.00  8.42      B    C
ATOM   6799  CA  SER B 155      29.390  41.759 177.743  1.00  8.54      B    C
ATOM   6805  CA  THR B 156      28.171  38.643 179.484  1.00  7.06      B    C
ATOM   6812  CA  PRO B 157      27.319  36.804 176.247  1.00  7.08      B    C
ATOM   6819  CA  SER B 158      24.232  34.650 176.044  1.00  6.71      B    C
ATOM   6825  CA  LEU B 159      22.370  32.582 173.502  1.00  4.87      B    C
ATOM   6833  CA  VAL B 160      18.646  31.941 173.847  1.00  3.09      B    C
ATOM   6840  CA  VAL B 161      17.419  28.777 172.147  1.00  3.34      B    C
ATOM   6847  CA  ASP B 162      13.682  29.337 171.720  1.00  4.07      B    C
```

도면340

```
ATOM   6855  CA  LEU B 163      11.876  26.065 170.870  1.00  3.79       B     C
ATOM   6863  CA  GLU B 164       8.402  27.636 170.688  1.00  4.18       B     C
ATOM   6872  CA  THR B 165       9.046  29.648 167.474  1.00  2.64       B     C
ATOM   6879  CA  ASN B 166       7.417  28.351 164.312  1.00  2.42       B     C
ATOM   6887  CA  ASN B 167       6.159  30.090 161.193  1.00  2.44       B     C
ATOM   6895  CA  THR B 168       2.700  28.469 161.520  1.00  2.77       B     C
ATOM   6902  CA  SER B 169       2.191  26.088 164.470  1.00  3.44       B     C
ATOM   6908  CA  ARG B 170       4.367  24.663 167.209  1.00  3.61       B     C
ATOM   6919  CA  ARG B 171       2.777  21.329 168.092  1.00  4.09       B     C
ATOM   6930  CA  SER B 172       5.367  20.136 170.618  1.00  3.92       B     C
ATOM   6936  CA  SER B 173       9.026  20.436 171.605  1.00  4.52       B     C
ATOM   6942  CA  SER B 174      11.644  19.095 174.010  1.00  5.26       B     C
ATOM   6948  CA  VAL B 175      15.144  19.964 175.194  1.00  7.61       B     C
ATOM   6955  CA  GLN B 176      17.707  17.942 177.116  1.00  9.87       B     C
ATOM   6964  CA  VAL B 177      20.827  19.575 178.479  1.00 10.75       B     C
ATOM   6971  CA  GLU B 178      24.266  17.972 178.853  1.00 11.75       B     C
ATOM   6980  CA  THR B 179      27.408  19.621 180.271  1.00 10.76       B     C
ATOM   6987  CA  ARG B 180      30.637  17.869 179.161  1.00 11.07       B     C
ATOM   6998  CA  ALA B 181      34.052  17.533 180.867  1.00 12.22       B     C
ATOM   7003  CA  ASP B 182      35.476  20.584 179.033  1.00 11.74       B     C
ATOM   7011  CA  GLY B 183      32.508  22.707 180.228  1.00  9.15       B     C
ATOM   7015  CA  THR B 184      30.755  22.961 176.837  1.00  5.80       B     C
ATOM   7022  CA  VAL B 185      27.013  22.306 176.541  1.00  5.32       B     C
ATOM   7029  CA  GLU B 186      24.991  20.053 174.245  1.00  6.37       B     C
ATOM   7038  CA  LEU B 187      21.225  20.551 173.801  1.00  5.44       B     C
ATOM   7046  CA  SER B 188      19.138  17.900 172.108  1.00  4.09       B     C
ATOM   7052  CA  GLY B 189      15.468  17.073 171.486  1.00  3.12       B     C
ATOM   7056  CA  GLN B 190      12.682  17.282 168.938  1.00  2.37       B     C
ATOM   7065  CA  VAL B 191      10.502  19.973 167.357  1.00  2.26       B     C
ATOM   7072  CA  THR B 192       7.128  19.112 165.863  1.00  2.53       B     C
ATOM   7079  CA  GLY B 193       5.491  21.715 163.683  1.00  2.57       B     C
ATOM   7083  CA  TYR B 194       2.613  21.951 161.250  1.00  3.30       B     C
ATOM   7095  CA  PHE B 195       2.176  23.583 157.852  1.00  4.28       B     C
ATOM   7106  CA  TYR B 196      -0.598  23.421 155.296  1.00  7.19       B     C
ATOM   7118  CA  ASN B 197      -1.557  19.737 155.149  1.00  7.00       B     C
ATOM   7126  CA  ALA B 198       0.738  17.907 157.606  1.00  5.17       B     C
ATOM   7131  CA  ALA B 199       2.648  17.834 160.876  1.00  4.90       B     C
ATOM   7136  CA  TYR B 200       6.409  17.069 160.789  1.00  4.22       B     C
ATOM   7148  CA  THR B 201       9.101  16.185 163.362  1.00  5.31       B     C
ATOM   7155  CA  LEU B 202      12.799  16.977 163.328  1.00  3.48       B     C
ATOM   7163  CA  TYR B 203      15.415  15.973 165.896  1.00  3.48       B     C
ATOM   7175  CA  TYR B 204      18.277  18.294 166.724  1.00  4.54       B     C
ATOM   7187  CA  THR B 205      21.584  18.490 168.504  1.00  5.64       B     C
ATOM   7194  CA  ALA B 206      23.308  21.765 169.341  1.00  5.72       B     C
ATOM   7199  CA  ARG B 207      26.654  22.253 170.998  1.00  5.49       B     C
ATOM   7210  CA  THR B 208      28.601  25.232 172.226  1.00  4.80       B     C
ATOM   7217  CA  LEU B 209      32.264  25.598 171.305  1.00  4.71       B     C
ATOM   7225  CA  GLN B 210      33.041  27.516 174.554  1.00  6.35       B     C
ATOM   7234  CA  PRO B 211      32.110  26.663 178.148  1.00  7.67       B     C
ATOM   7241  CA  ALA B 212      28.543  27.611 179.143  1.00  7.25       B     C
ATOM   7246  CA  THR B 213      26.099  27.434 182.004  1.00  7.17       B     C
ATOM   7253  CA  VAL B 214      22.470  26.720 181.121  1.00  6.69       B     C
ATOM   7260  CA  GLN B 215      18.911  27.268 182.283  1.00  6.38       B     C
ATOM   7269  CA  THR B 216      15.794  25.836 180.631  1.00  4.98       B     C
ATOM   7276  CA  TRP B 217      12.111  26.702 180.466  1.00  3.64       B     C
ATOM   7290  CA  GLY B 218       8.884  24.866 180.016  1.00  5.63       B     C
ATOM   7294  CA  ASP B 219       5.244  25.649 179.622  1.00  8.12       B     C
ATOM   7302  CA  ASP B 220       4.977  27.146 183.127  1.00 12.50       B     C
ATOM   7310  CA  ASP B 221       7.168  30.093 181.953  1.00 12.91       B     C
ATOM   7318  CA  ARG B 222       9.634  29.500 184.787  1.00 11.75       B     C
ATOM   7329  CA  LEU B 223      13.374  29.633 184.047  1.00  9.26       B     C
ATOM   7337  CA  VAL B 224      15.093  26.886 186.002  1.00  9.70       B     C
ATOM   7344  CA  ASP B 225      18.254  24.870 186.542  1.00  9.71       B     C
ATOM   7352  CA  ALA B 226      16.446  21.645 185.479  1.00  8.94       B     C
ATOM   7357  CA  THR B 227      18.172  19.922 182.585  1.00  8.94       B     C
ATOM   7364  CA  ALA B 228      15.027  18.498 180.872  1.00  8.41       B     C
```

도면34p

```
ATOM   7369  CA  GLN B 229      11.887  19.991 179.343  1.00  7.11      B     C
ATOM   7378  CA  ASP B 230       9.154  18.361 177.262  1.00  7.62      B     C
ATOM   7386  CA  GLY B 231       5.982  20.180 176.272  1.00  5.49      B     C
ATOM   7390  CA  VAL B 232       5.019  22.814 173.687  1.00  3.91      B     C
ATOM   7397  CA  ASP B 233       6.717  26.124 174.666  1.00  3.59      B     C
ATOM   7405  CA  THR B 234      10.207  25.219 175.870  1.00  3.22      B     C
ATOM   7412  CA  GLY B 235      13.760  26.456 175.485  1.00  2.52      B     C
ATOM   7416  CA  ALA B 236      17.165  27.054 176.932  1.00  2.01      B     C
ATOM   7421  CA  ILE B 237      19.409  30.014 177.730  1.00  2.62      B     C
ATOM   7429  CA  LEU B 238      23.188  29.544 177.393  1.00  3.80      B     C
ATOM   7437  CA  THR B 239      25.439  31.941 179.352  1.00  5.81      B     C
ATOM   7444  CA  PHE B 240      29.137  32.391 178.663  1.00  7.71      B     C
ATOM   7455  CA  ASP B 241      32.053  34.026 180.461  1.00 10.49      B     C
ATOM   7463  CA  PRO B 242      32.976  37.585 179.419  1.00 10.21      B     C
ATOM   7470  CA  ALA B 243      36.482  36.270 178.652  1.00 10.56      B     C
ATOM   7475  CA  ASP B 244      34.898  34.545 175.608  1.00 10.22      B     C
ATOM   7483  CA  ALA B 245      32.993  37.584 174.265  1.00  9.19      B     C
ATOM   7488  CA  GLY B 246      35.260  37.861 171.199  1.00  9.07      B     C
ATOM   7492  CA  GLU B 247      34.117  34.471 169.883  1.00  8.40      B     C
ATOM   7501  CA  ILE B 248      30.976  32.504 170.743  1.00  6.26      B     C
ATOM   7509  CA  GLY B 249      30.419  29.317 168.700  1.00  5.26      B     C
ATOM   7513  CA  LEU B 250      27.365  27.148 168.128  1.00  4.75      B     C
ATOM   7521  CA  GLN B 251      26.729  24.080 165.940  1.00  4.54      B     C
ATOM   7530  CA  VAL B 252      23.221  22.815 165.229  1.00  4.09      B     C
ATOM   7537  CA  THR B 253      22.517  19.573 163.382  1.00  3.14      B     C
ATOM   7544  CA  LEU B 254      19.085  18.325 162.254  1.00  3.66      B     C
ATOM   7552  CA  SER B 255      17.710  14.818 161.568  1.00  3.72      B     C
ATOM   7558  CA  PRO B 256      14.350  13.322 160.581  1.00  3.90      B     C
ATOM   7565  CA  VAL B 257      15.675  10.122 162.247  1.00  4.49      B     C
ATOM   7572  CA  SER B 258      16.553  10.685 165.919  1.00  5.09      B     C
ATOM   7578  CA  VAL B 259      18.811  12.681 168.270  1.00  6.85      B     C
ATOM   7585  CA  GLU B 260      21.309   9.841 168.225  1.00  8.56      B     C
ATOM   7594  CA  GLN B 261      21.425  10.162 164.416  1.00  7.14      B     C
ATOM   7603  CA  ALA B 262      21.764  13.965 164.573  1.00  6.00      B     C
ATOM   7608  CA  ARG B 263      24.840  13.567 166.780  1.00  6.59      B     C
ATOM   7619  CA  ILE B 264      26.315  11.008 164.363  1.00  7.36      B     C
ATOM   7627  CA  ASP B 265      25.571  13.292 161.370  1.00  6.50      B     C
ATOM   7635  CA  GLN B 266      27.201  16.188 163.205  1.00  7.37      B     C
ATOM   7644  CA  GLN B 267      30.403  14.178 163.786  1.00 11.20      B     C
ATOM   7653  CA  VAL B 268      30.481  13.075 160.124  1.00 10.88      B     C
ATOM   7660  CA  GLU B 269      29.430  16.369 158.505  1.00 11.16      B     C
ATOM   7669  CA  LEU B 270      31.314  18.882 160.704  1.00 10.96      B     C
ATOM   7677  CA  GLY B 271      33.869  16.943 162.740  1.00 11.80      B     C
ATOM   7681  CA  ASP B 272      36.639  19.186 164.002  1.00 11.87      B     C
ATOM   7689  CA  LEU B 273      36.374  21.606 161.059  1.00  9.30      B     C
ATOM   7697  CA  SER B 274      36.585  25.316 161.684  1.00  8.10      B     C
ATOM   7703  CA  PHE B 275      33.692  27.729 161.001  1.00  7.57      B     C
ATOM   7714  CA  ASP B 276      35.633  29.179 158.049  1.00  9.11      B     C
ATOM   7722  CA  ALA B 277      36.288  25.679 156.588  1.00  8.17      B     C
ATOM   7727  CA  ILE B 278      32.607  24.690 156.805  1.00  6.31      B     C
ATOM   7735  CA  ARG B 279      31.510  27.935 155.180  1.00  7.50      B     C
ATOM   7746  CA  ASP B 280      34.193  27.585 152.447  1.00  8.14      B     C
ATOM   7754  CA  ARG B 281      33.248  23.952 151.718  1.00  7.53      B     C
ATOM   7765  CA  THR B 282      29.590  24.914 151.207  1.00  7.03      B     C
ATOM   7772  CA  ARG B 283      30.749  27.741 148.908  1.00  8.06      B     C
ATOM   7783  CA  ALA B 284      32.685  25.192 146.809  1.00  7.18      B     C
ATOM   7788  CA  GLU B 285      29.647  22.838 146.826  1.00  8.04      B     C
ATOM   7797  CA  TRP B 286      27.513  25.684 145.409  1.00  6.44      B     C
ATOM   7811  CA  ASN B 287      30.194  26.541 142.822  1.00  7.77      B     C
ATOM   7819  CA  ALA B 288      30.163  22.844 141.765  1.00  8.01      B     C
ATOM   7824  CA  THR B 289      26.340  22.924 141.481  1.00  7.57      B     C
ATOM   7831  CA  LEU B 290      25.928  26.258 139.687  1.00  6.68      B     C
ATOM   7839  CA  GLY B 291      28.902  25.412 137.486  1.00  8.30      B     C
ATOM   7843  CA  ARG B 292      26.832  22.721 135.807  1.00  8.97      B     C
ATOM   7854  CA  VAL B 293      25.594  25.636 133.725  1.00 10.54      B     C
ATOM   7861  CA  ALA B 294      28.121  27.918 132.046  1.00 12.46      B     C
```

도면34q

```
ATOM   7866  CA  ILE B 295      26.911  31.025 130.202  1.00 15.01      B    C
ATOM   7874  CA  ASP B 296      28.503  33.444 127.742  1.00 16.68      B    C
ATOM   7882  CA  ALA B 297      26.492  36.686 127.506  1.00 16.02      B    C
ATOM   7887  CA  SER B 298      27.400  39.653 125.271  1.00 15.25      B    C
ATOM   7893  CA  THR B 299      26.635  43.276 126.121  1.00 14.42      B    C
ATOM   7900  CA  ALA B 300      24.046  43.110 123.326  1.00 11.48      B    C
ATOM   7905  CA  THR B 301      21.783  40.582 125.103  1.00  9.84      B    C
ATOM   7912  CA  ASP B 302      22.704  41.528 128.706  1.00  9.77      B    C
ATOM   7920  CA  PRO B 303      23.868  45.197 128.715  1.00 10.54      B    C
ATOM   7927  CA  THR B 304      23.489  45.589 132.512  1.00 11.07      B    C
ATOM   7934  CA  GLY B 305      24.761  42.154 133.524  1.00 11.11      B    C
ATOM   7938  CA  GLU B 306      21.483  41.714 135.432  1.00 10.64      B    C
ATOM   7947  CA  LEU B 307      20.268  38.736 133.293  1.00 10.30      B    C
ATOM   7955  CA  GLN B 308      23.365  36.685 134.127  1.00 10.19      B    C
ATOM   7964  CA  ARG B 309      22.914  37.468 137.835  1.00  8.46      B    C
ATOM   7975  CA  LEU B 310      19.207  36.600 137.644  1.00  6.90      B    C
ATOM   7983  CA  PHE B 311      20.109  33.341 135.894  1.00  5.42      B    C
ATOM   7994  CA  TYR B 312      22.509  32.208 138.679  1.00  4.65      B    C
ATOM   8006  CA  THR B 313      20.288  33.567 141.446  1.00  3.72      B    C
ATOM   8013  CA  HIS B 314      17.411  31.465 140.078  1.00  3.77      B    C
ATOM   8023  CA  LEU B 315      19.597  28.424 139.419  1.00  4.98      B    C
ATOM   8031  CA  TYR B 316      20.467  28.753 143.144  1.00  4.55      B    C
ATOM   8043  CA  ARG B 317      16.731  28.839 144.007  1.00  5.20      B    C
ATOM   8054  CA  MET B 318      16.085  25.817 141.789  1.00  7.54      B    C
ATOM   8062  CA  PHE B 319      18.175  23.558 144.063  1.00  8.02      B    C
ATOM   8073  CA  ALA B 320      16.316  24.249 147.345  1.00  7.25      B    C
ATOM   8078  CA  MET B 321      13.403  21.781 147.467  1.00  5.88      B    C
ATOM   8086  CA  PRO B 322      12.863  18.952 147.903  1.00  4.25      B    C
ATOM   8093  CA  MET B 323      15.380  18.850 150.760  1.00  3.69      B    C
ATOM   8101  CA  ASN B 324      17.944  16.177 151.447  1.00  3.38      B    C
ATOM   8109  CA  ALA B 325      16.107  14.123 154.077  1.00  3.72      B    C
ATOM   8114  CA  THR B 326      18.714  11.357 154.443  1.00  3.39      B    C
ATOM   8121  CA  SER B 327      21.071  10.805 157.383  1.00  4.75      B    C
ATOM   8127  CA  THR B 328      24.861  10.478 156.953  1.00  5.56      B    C
ATOM   8134  CA  SER B 329      24.175   6.783 157.673  1.00  5.03      B    C
ATOM   8140  CA  GLY B 330      21.990   6.586 154.509  1.00  5.56      B    C
ATOM   8144  CA  THR B 331      18.681   6.308 156.385  1.00  4.67      B    C
ATOM   8151  CA  TYR B 332      15.380   8.194 156.221  1.00  4.08      B    C
ATOM   8163  CA  ARG B 333      12.110   8.234 158.156  1.00  5.90      B    C
ATOM   8174  CA  GLY B 334       8.940   7.045 156.393  1.00  7.33      B    C
ATOM   8178  CA  VAL B 335       5.362   8.276 156.776  1.00  9.86      B    C
ATOM   8185  CA  ASP B 336       4.715   4.745 158.108  1.00 11.83      B    C
ATOM   8193  CA  GLY B 337       6.472   5.889 161.292  1.00 12.67      B    C
ATOM   8197  CA  ALA B 338       9.546   3.712 160.784  1.00 10.29      B    C
ATOM   8202  CA  VAL B 339      13.181   4.400 159.827  1.00  8.36      B    C
ATOM   8209  CA  HIS B 340      14.277   2.783 156.523  1.00  8.22      B    C
ATOM   8219  CA  ALA B 341      17.345   2.468 154.295  1.00  9.30      B    C
ATOM   8224  CA  ALA B 342      17.738   4.699 151.232  1.00 10.96      B    C
ATOM   8229  CA  GLN B 343      19.736   1.973 149.549  1.00 14.92      B    C
ATOM   8238  CA  GLY B 344      22.061   3.243 146.798  1.00 14.18      B    C
ATOM   8242  CA  PHE B 345      20.330   6.658 146.470  1.00  9.94      B    C
ATOM   8253  CA  THR B 346      19.663   9.851 148.454  1.00  5.70      B    C
ATOM   8260  CA  TYR B 347      16.132  10.341 149.769  1.00  3.08      B    C
ATOM   8272  CA  TYR B 348      14.568  13.786 149.336  1.00  2.49      B    C
ATOM   8284  CA  ASP B 349      11.394  15.091 150.938  1.00  2.40      B    C
ATOM   8292  CA  SER B 350       9.337  18.340 151.066  1.00  3.13      B    C
ATOM   8298  CA  TRP B 351       6.904  18.658 148.223  1.00  3.44      B    C
ATOM   8312  CA  ALA B 352       4.594  21.317 146.769  1.00  6.09      B    C
ATOM   8317  CA  THR B 353       3.818  19.352 143.690  1.00  7.57      B    C
ATOM   8324  CA  TRP B 354       0.511  21.060 142.861  1.00 10.28      B    C
ATOM   8338  CA  ASP B 355       2.607  24.171 142.126  1.00  9.25      B    C
ATOM   8346  CA  ASP B 356       6.027  22.821 141.162  1.00  7.97      B    C
ATOM   8354  CA  PHE B 357       5.476  19.675 139.069  1.00  5.45      B    C
ATOM   8365  CA  ARG B 358       7.086  21.219 135.922  1.00  4.54      B    C
ATOM   8376  CA  LYS B 359      10.204  22.174 137.907  1.00  3.86      B    C
ATOM   8385  CA  PHE B 360      11.251  18.513 137.948  1.00  3.85      B    C
```

도면34r

```
ATOM   8396  CA  SER B 361      11.230  18.254 134.132  1.00  4.19      B    C
ATOM   8402  CA  VAL B 362      13.848  21.051 134.057  1.00  4.66      B    C
ATOM   8409  CA  ILE B 363      15.981  19.656 136.859  1.00  3.60      B    C
ATOM   8417  CA  ALA B 364      15.986  16.487 134.724  1.00  4.12      B    C
ATOM   8422  CA  TYR B 365      18.152  18.190 132.040  1.00  5.46      B    C
ATOM   8434  CA  ILE B 366      20.412  20.182 134.397  1.00  6.51      B    C
ATOM   8442  CA  ASP B 367      21.232  17.592 137.055  1.00  6.76      B    C
ATOM   8450  CA  PRO B 368      19.936  14.178 135.858  1.00  6.15      B    C
ATOM   8457  CA  ALA B 369      21.272  12.238 138.895  1.00  5.16      B    C
ATOM   8462  CA  LEU B 370      19.595  14.581 141.391  1.00  5.66      B    C
ATOM   8470  CA  TYR B 371      16.368  14.351 139.411  1.00  5.00      B    C
ATOM   8482  CA  ARG B 372      16.537  10.538 139.507  1.00  6.13      B    C
ATOM   8493  CA  ASP B 373      16.991  10.482 143.321  1.00  4.65      B    C
ATOM   8501  CA  MET B 374      14.069  12.884 143.690  1.00  3.74      B    C
ATOM   8509  CA  VAL B 375      11.706  10.681 141.648  1.00  3.94      B    C
ATOM   8516  CA  GLN B 376      12.853   7.548 143.481  1.00  3.66      B    C
ATOM   8525  CA  SER B 377      12.139   9.421 146.732  1.00  3.93      B    C
ATOM   8531  CA  LEU B 378       8.564  10.288 145.559  1.00  4.51      B    C
ATOM   8539  CA  VAL B 379       8.064   6.634 144.704  1.00  3.56      B    C
ATOM   8546  CA  TYR B 380       9.204   5.663 148.259  1.00  4.32      B    C
ATOM   8558  CA  LEU B 381       6.959   8.247 149.897  1.00  3.49      B    C
ATOM   8566  CA  PHE B 382       3.838   6.774 148.265  1.00  3.82      B    C
ATOM   8577  CA  ALA B 383       5.141   3.221 148.777  1.00  4.94      B    C
ATOM   8582  CA  ASP B 384       5.471   3.995 152.503  1.00  6.79      B    C
ATOM   8590  CA  ALA B 385       1.923   5.413 152.611  1.00  9.08      B    C
ATOM   8595  CA  GLU B 386       0.765   2.141 150.976  1.00 10.89      B    C
ATOM   8604  CA  ALA B 387       2.810   0.119 153.498  1.00 13.19      B    C
ATOM   8609  CA  THR B 388       0.740   1.552 156.445  1.00 15.38      B    C
ATOM   8616  CA  GLY B 389      -2.271  -0.495 155.224  1.00 15.84      B    C
ATOM   8620  CA  THR B 390      -4.587   2.321 156.337  1.00 14.91      B    C
ATOM   8627  CA  GLY B 391      -5.415   3.542 152.829  1.00 12.96      B    C
ATOM   8631  CA  GLY B 392      -5.069   7.090 154.245  1.00 10.55      B    C
ATOM   8635  CA  GLY B 393      -3.986  10.168 152.254  1.00  8.29      B    C
ATOM   8639  CA  LEU B 394      -0.516  11.679 152.818  1.00  6.39      B    C
ATOM   8647  CA  GLY B 395      -1.926  14.514 154.938  1.00  6.61      B    C
ATOM   8651  CA  GLY B 396      -2.978  12.076 157.705  1.00  7.77      B    C
ATOM   8655  CA  PHE B 397       0.603  11.030 158.585  1.00  7.05      B    C
ATOM   8666  CA  VAL B 398       3.457  12.825 160.317  1.00  6.28      B    C
ATOM   8673  CA  HIS B 399       6.082  13.832 157.779  1.00  4.37      B    C
ATOM   8683  CA  SER B 400       9.868  13.714 158.080  1.00  4.62      B    C
ATOM   8689  CA  VAL B 401      10.519  17.328 157.074  1.00  3.81      B    C
ATOM   8696  CA  PRO B 402       8.279  20.398 156.732  1.00  2.81      B    C
ATOM   8703  CA  THR B 403       6.223  19.841 153.605  1.00  2.00      B    C
ATOM   8710  CA  VAL B 404       3.295  21.247 151.582  1.00  2.02      B    C
ATOM   8717  CA  ARG B 405       1.236  19.105 149.134  1.00  3.66      B    C
ATOM   8728  CA  TRP B 406       1.343  16.193 146.726  1.00  4.40      B    C
ATOM   8742  CA  GLU B 407      -1.154  16.632 143.859  1.00  4.95      B    C
ATOM   8751  CA  ARG B 408       0.354  16.056 140.358  1.00  4.63      B    C
ATOM   8762  CA  SER B 409       2.805  13.445 141.798  1.00  4.41      B    C
ATOM   8768  CA  SER B 410       1.577  10.955 139.158  1.00  5.65      B    C
ATOM   8774  CA  VAL B 411       2.871  13.412 136.534  1.00  4.55      B    C
ATOM   8781  CA  VAL B 412       6.352  13.733 138.090  1.00  3.62      B    C
ATOM   8788  CA  VAL B 413       6.857   9.945 138.202  1.00  3.87      B    C
ATOM   8795  CA  ALA B 414       5.438   9.658 134.653  1.00  5.13      B    C
ATOM   8800  CA  ASP B 415       7.928  12.368 133.664  1.00  4.82      B    C
ATOM   8808  CA  ALA B 416      10.908  10.154 134.619  1.00  4.44      B    C
ATOM   8813  CA  ILE B 417       9.427   7.020 132.987  1.00  5.16      B    C
ATOM   8821  CA  ALA B 418       8.712   8.909 129.732  1.00  6.23      B    C
ATOM   8826  CA  LYS B 419      12.303  10.194 129.851  1.00  8.29      B    C
ATOM   8835  CA  GLY B 420      13.794   6.686 129.922  1.00 10.82      B    C
ATOM   8839  CA  PHE B 421      14.569   6.285 133.624  1.00 11.95      B    C
ATOM   8850  CA  ASP B 422      13.949   2.885 135.136  1.00 14.48      B    C
ATOM   8858  CA  GLY B 423      14.815   0.654 138.084  1.00 14.15      B    C
ATOM   8862  CA  PHE B 424      12.489   2.574 140.366  1.00 12.52      B    C
ATOM   8873  CA  ASP B 425      12.381   0.289 143.387  1.00 12.13      B    C
ATOM   8881  CA  ARG B 426       8.874  -0.058 144.898  1.00  9.62      B    C
```

도면34s

```
ATOM   8892  CA  LEU B 427       6.950   1.477 141.964  1.00  9.74       B     C
ATOM   8900  CA  ASP B 428       4.503  -1.436 142.300  1.00 11.74       B     C
ATOM   8908  CA  GLU B 429       3.837  -0.485 145.942  1.00 10.85       B     C
ATOM   8917  CA  ALA B 430       3.557   3.228 145.139  1.00  9.56       B     C
ATOM   8922  CA  TYR B 431       0.970   2.649 142.393  1.00  9.69       B     C
ATOM   8934  CA  PRO B 432      -2.178   2.019 144.553  1.00  8.88       B     C
ATOM   8941  CA  ALA B 433      -1.368   5.086 146.736  1.00  7.60       B     C
ATOM   8946  CA  LEU B 434      -0.914   7.129 143.566  1.00  7.98       B     C
ATOM   8954  CA  GLN B 435      -4.332   5.893 142.378  1.00  9.79       B     C
ATOM   8963  CA  ARG B 436      -5.876   6.985 145.714  1.00  8.63       B     C
ATOM   8974  CA  LEU B 437      -4.175  10.398 145.345  1.00  8.82       B     C
ATOM   8982  CA  VAL B 438      -5.651  10.832 141.863  1.00 10.72       B     C
ATOM   8989  CA  GLY B 439      -9.074   9.520 142.973  1.00 11.53       B     C
ATOM   8993  CA  GLN B 440     -11.952   7.945 141.034  1.00 13.71       B     C
ATOM   9002  CA  TYR B 441     -14.865   9.454 139.124  1.00 12.01       B     C
ATOM   9014  CA  SER B 442     -18.189   8.549 140.760  1.00 12.74       B     C
ATOM   9020  CA  ALA B 443     -20.666   6.339 138.839  1.00 12.68       B     C
ATOM   9025  CA  ASP B 444     -22.553   9.434 137.687  1.00 12.69       B     C
ATOM   9033  CA  GLU B 445     -19.307  11.199 136.788  1.00 10.37       B     C
ATOM   9042  CA  LEU B 446     -18.282   8.164 134.766  1.00 10.75       B     C
ATOM   9050  CA  ARG B 447     -21.517   8.211 132.730  1.00 12.18       B     C
ATOM   9061  CA  ARG B 448     -21.331  11.954 131.904  1.00 11.28       B     C
ATOM   9072  CA  GLY B 449     -17.521  12.199 131.769  1.00  9.56       B     C
ATOM   9076  CA  TYR B 450     -16.840  15.171 134.104  1.00  8.69       B     C
ATOM   9088  CA  VAL B 451     -17.362  16.664 137.586  1.00  8.42       B     C
ATOM   9095  CA  ALA B 452     -20.609  18.655 137.398  1.00  9.63       B     C
ATOM   9100  CA  GLY B 453     -19.943  22.350 137.910  1.00  9.76       B     C
ATOM   9104  CA  ASN B 454     -16.242  21.775 138.624  1.00  8.50       B     C
ATOM   9112  CA  PRO B 455     -14.156  22.194 135.439  1.00  8.32       B     C
ATOM   9119  CA  GLY B 456     -11.107  22.756 137.692  1.00  8.02       B     C
ATOM   9123  CA  ALA B 457     -11.159  19.292 139.280  1.00  6.90       B     C
ATOM   9128  CA  SER B 458     -11.963  17.688 135.889  1.00  6.05       B     C
ATOM   9134  CA  VAL B 459      -8.939  18.989 133.896  1.00  5.70       B     C
ATOM   9141  CA  GLN B 460      -6.681  18.365 136.882  1.00  5.79       B     C
ATOM   9150  CA  ARG B 461      -7.804  14.732 137.121  1.00  6.42       B     C
ATOM   9161  CA  GLY B 462      -7.221  14.406 133.370  1.00  5.35       B     C
ATOM   9165  CA  TYR B 463      -3.539  15.278 133.771  1.00  5.04       B     C
ATOM   9177  CA  ASP B 464      -3.270  13.035 136.851  1.00  5.57       B     C
ATOM   9185  CA  GLN B 465      -4.804  10.198 134.841  1.00  5.78       B     C
ATOM   9194  CA  TYR B 466      -2.483  10.769 131.919  1.00  6.49       B     C
ATOM   9206  CA  GLY B 467       0.439  10.511 134.378  1.00  5.94       B     C
ATOM   9210  CA  LEU B 468      -1.027   7.423 136.006  1.00  7.08       B     C
ATOM   9218  CA  SER B 469      -1.447   5.806 132.563  1.00  7.35       B     C
ATOM   9224  CA  VAL B 470       2.324   6.119 131.916  1.00  7.31       B     C
ATOM   9231  CA  ILE B 471       3.054   4.452 135.299  1.00  7.55       B     C
ATOM   9239  CA  ALA B 472       0.432   1.758 134.635  1.00  8.69       B     C
ATOM   9244  CA  ASP B 473       2.077   0.765 131.318  1.00 10.11       B     C
ATOM   9252  CA  GLU B 474       5.524   0.684 132.942  1.00 11.08       B     C
ATOM   9261  CA  LEU B 475       4.081  -1.770 135.471  1.00 10.77       B     C
ATOM   9269  CA  GLY B 476       2.557  -3.863 132.643  1.00 12.28       B     C
ATOM   9273  CA  LEU B 477      -1.005  -2.946 133.703  1.00 14.48       B     C
ATOM   9281  CA  THR B 478      -1.957  -2.607 130.034  1.00 17.13       B     C
ATOM   9288  CA  GLU B 479      -5.736  -2.506 130.474  1.00 18.35       B     C
ATOM   9297  CA  GLU B 480      -5.722   0.045 133.355  1.00 15.71       B     C
ATOM   9306  CA  ALA B 481      -3.447   2.289 131.256  1.00 13.30       B     C
ATOM   9311  CA  GLU B 482      -5.993   2.198 128.387  1.00 13.38       B     C
ATOM   9320  CA  THR B 483      -8.838   3.007 130.776  1.00 11.91       B     C
ATOM   9327  CA  LEU B 484      -6.802   5.857 132.279  1.00  9.11       B     C
ATOM   9335  CA  ARG B 485      -6.030   7.235 128.765  1.00  8.25       B     C
ATOM   9346  CA  GLU B 486      -9.735   7.030 127.831  1.00  8.59       B     C
ATOM   9355  CA  GLN B 487     -10.721   8.982 130.961  1.00  7.84       B     C
ATOM   9364  CA  ALA B 488      -7.917  11.568 130.491  1.00  6.89       B     C
ATOM   9369  CA  SER B 489      -9.413  12.554 127.143  1.00  6.88       B     C
ATOM   9375  CA  TRP B 490     -12.765  13.388 128.798  1.00  7.07       B     C
ATOM   9389  CA  PRO B 491     -12.065  16.904 130.180  1.00  7.75       B     C
ATOM   9396  CA  ILE B 492     -11.008  18.232 126.748  1.00 10.03       B     C
```

도면34t

```
ATOM   9404  CA  GLU B 493     -13.896  16.510 124.905  1.00 12.06      B    C
ATOM   9413  CA  LYS B 494     -16.671  17.061 127.427  1.00 10.91      B    C
ATOM   9422  CA  LEU B 495     -15.887  20.513 128.833  1.00  9.26      B    C
ATOM   9430  CA  THR B 496     -14.827  22.544 125.800  1.00  9.84      B    C
ATOM   9437  CA  LYS B 497     -17.972  24.524 124.878  1.00 10.36      B    C
ATOM   9446  CA  PRO B 498     -18.062  25.503 121.171  1.00 10.16      B    C
ATOM   9453  CA  GLY B 499     -18.441  29.251 120.587  1.00 11.73      B    C
ATOM   9457  CA  ALA B 500     -17.533  30.238 124.169  1.00 12.89      B    C
ATOM   9462  CA  TRP B 501     -15.579  33.120 122.622  1.00 13.32      B    C
ATOM   9476  CA  THR B 502     -16.087  34.739 119.196  1.00 14.82      B    C
ATOM   9483  CA  ALA B 503     -13.072  36.086 117.234  1.00 17.18      B    C
ATOM   9488  CA  ALA B 504     -12.857  39.344 115.232  1.00 19.86      B    C
ATOM   9493  CA  ASP B 505     -13.256  36.963 112.234  1.00 21.28      B    C
ATOM   9501  CA  GLY B 506     -16.585  35.755 113.609  1.00 20.03      B    C
ATOM   9505  CA  THR B 507     -14.731  32.439 114.087  1.00 18.03      B    C
ATOM   9512  CA  GLN B 508     -16.251  30.532 117.010  1.00 15.34      B    C
ATOM   9521  CA  VAL B 509     -13.708  29.336 119.570  1.00 11.34      B    C
ATOM   9528  CA  GLY B 510     -14.366  26.537 122.049  1.00  9.07      B    C
ATOM   9532  CA  LEU B 511     -13.197  27.026 125.631  1.00  7.86      B    C
ATOM   9540  CA  LEU B 512     -13.219  25.124 128.921  1.00  7.01      B    C
ATOM   9548  CA  THR B 513     -16.601  25.890 130.550  1.00  8.85      B    C
ATOM   9555  CA  PRO B 514     -18.393  24.554 133.677  1.00 10.28      B    C
ATOM   9562  CA  ARG B 515     -21.081  22.039 132.717  1.00 12.26      B    C
ATOM   9573  CA  ALA B 516     -24.111  20.973 134.766  1.00 13.16      B    C
ATOM   9578  CA  ALA B 517     -24.951  17.342 135.665  1.00 14.35      B    C
ATOM   9583  CA  ASP B 518     -27.642  17.167 132.923  1.00 16.53      B    C
ATOM   9591  CA  GLY B 519     -25.191  18.281 130.210  1.00 16.33      B    C
ATOM   9595  CA  SER B 520     -26.240  21.945 129.971  1.00 15.96      B    C
ATOM   9601  CA  TRP B 521     -23.559  24.648 130.008  1.00 15.21      B    C
ATOM   9615  CA  GLN B 522     -23.432  26.929 133.060  1.00 15.39      B    C
ATOM   9624  CA  SER B 523     -22.733  30.599 132.427  1.00 15.49      B    C
ATOM   9630  CA  ALA B 524     -19.406  32.190 133.361  1.00 13.08      B    C
ATOM   9635  CA  ASP B 525     -16.958  34.949 132.667  1.00 10.42      B    C
ATOM   9643  CA  HIS B 526     -13.996  33.026 131.216  1.00  9.11      B    C
ATOM   9653  CA  ALA B 527     -11.679  35.887 132.307  1.00  9.10      B    C
ATOM   9658  CA  LYS B 528     -12.918  36.154 135.899  1.00 10.32      B    C
ATOM   9667  CA  PHE B 529     -10.563  35.052 138.686  1.00  8.73      B    C
ATOM   9678  CA  GLU B 530     -11.981  32.026 140.571  1.00  8.70      B    C
ATOM   9687  CA  ALA B 531     -15.175  31.800 138.513  1.00  9.23      B    C
ATOM   9692  CA  ALA B 532     -16.810  28.339 138.183  1.00 10.00      B    C
ATOM   9697  CA  GLY B 533     -15.211  26.855 141.335  1.00  9.25      B    C
ATOM   9701  CA  LEU B 534     -11.653  27.388 140.006  1.00  8.07      B    C
ATOM   9709  CA  TYR B 535      -8.638  27.486 142.280  1.00 10.81      B    C
ATOM   9721  CA  GLN B 536      -6.568  30.681 141.905  1.00 10.33      B    C
ATOM   9730  CA  GLY B 537      -7.264  31.159 138.187  1.00  8.79      B    C
ATOM   9734  CA  THR B 538      -9.626  31.771 135.278  1.00  7.44      B    C
ATOM   9741  CA  LEU B 539     -11.154  29.505 132.643  1.00  8.31      B    C
ATOM   9749  CA  TRP B 540      -8.976  31.006 129.900  1.00  6.79      B    C
ATOM   9763  CA  GLN B 541      -6.001  30.062 132.087  1.00  5.72      B    C
ATOM   9772  CA  TYR B 542      -7.247  26.567 133.014  1.00  6.07      B    C
ATOM   9784  CA  HIS B 543      -8.336  25.987 129.378  1.00  6.44      B    C
ATOM   9794  CA  TRP B 544      -4.804  24.912 128.489  1.00  5.93      B    C
ATOM   9808  CA  TYR B 545      -4.279  22.715 131.546  1.00  4.79      B    C
ATOM   9820  CA  ASP B 546      -4.407  19.324 129.817  1.00  4.56      B    C
ATOM   9828  CA  ALA B 547      -0.984  19.835 128.277  1.00  5.44      B    C
ATOM   9833  CA  TYR B 548      -0.609  16.142 127.393  1.00  6.08      B    C
ATOM   9845  CA  ASP B 549      -3.092  16.487 124.510  1.00  5.83      B    C
ATOM   9853  CA  MET B 550      -2.830  19.820 122.688  1.00  7.27      B    C
ATOM   9861  CA  ASP B 551      -3.867  18.042 119.495  1.00  9.92      B    C
ATOM   9869  CA  ALA B 552      -7.181  17.034 121.101  1.00  9.29      B    C
ATOM   9874  CA  LEU B 553      -7.610  20.544 122.602  1.00  8.74      B    C
ATOM   9882  CA  VAL B 554      -6.969  22.270 119.281  1.00  9.28      B    C
ATOM   9889  CA  GLU B 555      -9.660  20.056 117.696  1.00 10.79      B    C
ATOM   9898  CA  ALA B 556     -12.151  20.721 120.539  1.00  8.98      B    C
ATOM   9903  CA  MET B 557     -11.629  24.480 120.201  1.00  9.27      B    C
ATOM   9911  CA  GLY B 558     -12.831  24.096 116.592  1.00 11.18      B    C
```

도면34u

```
ATOM   9915  CA  GLY B 559      -9.648  23.449 114.651  1.00 12.57      B    C
ATOM   9919  CA  HIS B 560      -6.240  24.925 113.907  1.00 13.76      B    C
ATOM   9929  CA  GLU B 561      -7.665  28.397 113.047  1.00 13.40      B    C
ATOM   9938  CA  ALA B 562      -9.852  28.561 116.156  1.00 11.93      B    C
ATOM   9943  CA  ALA B 563      -6.829  27.744 118.342  1.00 11.17      B    C
ATOM   9948  CA  ARG B 564      -4.548  30.202 116.523  1.00 10.14      B    C
ATOM   9959  CA  LEU B 565      -6.933  33.107 117.193  1.00  9.78      B    C
ATOM   9967  CA  GLY B 566      -7.517  31.781 120.725  1.00  8.93      B    C
ATOM   9971  CA  MET B 567      -3.792  31.948 121.520  1.00  9.45      B    C
ATOM   9979  CA  ARG B 568      -3.398  35.497 120.126  1.00  9.92      B    C
ATOM   9990  CA  HIS B 569      -6.452  36.629 122.170  1.00  9.91      B    C
ATOM  10000  CA  MET B 570      -4.961  34.969 125.270  1.00 10.22      B    C
ATOM  10008  CA  PHE B 571      -1.992  37.376 124.985  1.00  9.43      B    C
ATOM  10019  CA  GLY B 572      -3.956  40.469 123.877  1.00 11.08      B    C
ATOM  10023  CA  GLU B 573      -2.195  40.540 120.475  1.00 13.30      B    C
ATOM  10032  CA  HIS B 574      -5.029  42.598 118.931  1.00 14.86      B    C
ATOM  10042  CA  ALA B 575      -4.622  45.304 121.645  1.00 15.90      B    C
ATOM  10047  CA  PRO B 576      -0.972  45.390 122.696  1.00 17.21      B    C
ATOM  10054  CA  ASP B 577      -1.328  48.595 124.695  1.00 18.80      B    C
ATOM  10062  CA  ASP B 578      -4.343  47.440 126.680  1.00 16.37      B    C
ATOM  10070  CA  GLY B 579      -3.447  45.773 129.999  1.00 14.31      B    C
ATOM  10074  CA  LYS B 580      -6.893  44.273 130.536  1.00 14.35      B    C
ATOM  10083  CA  ALA B 581      -6.401  42.289 127.289  1.00 11.55      B    C
ATOM  10088  CA  MET B 582      -3.732  40.150 128.992  1.00  9.08      B    C
ATOM  10096  CA  LEU B 583      -5.611  36.913 129.818  1.00  6.75      B    C
ATOM  10104  CA  HIS B 584      -2.406  35.066 130.755  1.00  5.26      B    C
ATOM  10114  CA  SER B 585      -1.183  35.314 134.331  1.00  5.08      B    C
ATOM  10120  CA  ASN B 586       2.561  35.487 135.130  1.00  4.87      B    C
ATOM  10128  CA  ALA B 587       1.917  35.442 138.923  1.00  6.16      B    C
ATOM  10133  CA  ASN B 588       0.684  31.857 139.540  1.00  9.70      B    C
ATOM  10141  CA  GLU B 589       1.574  28.445 138.079  1.00 11.52      B    C
ATOM  10150  CA  ILE B 590      -1.724  27.523 136.394  1.00 11.76      B    C
ATOM  10158  CA  ASP B 591      -0.787  28.589 132.835  1.00 11.65      B    C
ATOM  10166  CA  LEU B 592       3.036  28.784 133.060  1.00 10.00      B    C
ATOM  10174  CA  GLN B 593       3.220  26.855 129.797  1.00  7.99      B    C
ATOM  10183  CA  ALA B 594       0.955  29.258 127.891  1.00  7.42      B    C
ATOM  10188  CA  PRO B 595       3.697  31.434 126.264  1.00  6.47      B    C
ATOM  10195  CA  TYR B 596       5.040  28.300 124.524  1.00  5.99      B    C
ATOM  10207  CA  LEU B 597       1.684  27.269 123.043  1.00  5.41      B    C
ATOM  10215  CA  PHE B 598       2.175  29.492 119.962  1.00  6.82      B    C
ATOM  10226  CA  ASN B 599       4.570  26.728 118.788  1.00  7.99      B    C
ATOM  10234  CA  TYR B 600       1.469  24.494 118.559  1.00  8.95      B    C
ATOM  10246  CA  THR B 601      -0.592  27.029 116.565  1.00  9.34      B    C
ATOM  10253  CA  GLY B 602       2.026  27.493 113.815  1.00  9.48      B    C
ATOM  10257  CA  GLU B 603       3.299  30.812 115.139  1.00  9.51      B    C
ATOM  10266  CA  PRO B 604       6.593  29.991 116.946  1.00  8.50      B    C
ATOM  10273  CA  SER B 605       7.845  33.578 116.478  1.00  7.57      B    C
ATOM  10279  CA  LEU B 606       5.226  34.638 119.060  1.00  7.18      B    C
ATOM  10287  CA  THR B 607       6.464  32.030 121.578  1.00  5.66      B    C
ATOM  10294  CA  GLN B 608       9.959  33.487 121.141  1.00  5.96      B    C
ATOM  10303  CA  LYS B 609       8.701  37.059 121.593  1.00  6.51      B    C
ATOM  10312  CA  TRP B 610       6.714  36.241 124.769  1.00  7.00      B    C
ATOM  10326  CA  ALA B 611       9.360  33.957 126.314  1.00  7.18      B    C
ATOM  10331  CA  ARG B 612      11.852  36.768 125.860  1.00  7.07      B    C
ATOM  10342  CA  ALA B 613       9.502  39.547 127.021  1.00  7.16      B    C
ATOM  10347  CA  ILE B 614       8.036  37.923 130.160  1.00  7.15      B    C
ATOM  10355  CA  TYR B 615      11.442  37.037 131.623  1.00  8.13      B    C
ATOM  10367  CA  THR B 616      13.573  39.966 130.409  1.00  8.81      B    C
ATOM  10374  CA  LYS B 617      11.308  42.977 129.716  1.00  9.89      B    C
ATOM  10383  CA  GLU B 618       8.250  44.801 130.976  1.00 10.62      B    C
ATOM  10392  CA  THR B 619       4.925  43.183 130.129  1.00 10.85      B    C
ATOM  10399  CA  TRP B 620       1.277  43.848 130.780  1.00 10.87      B    C
ATOM  10413  CA  ASN B 621       0.104  42.242 133.955  1.00  8.54      B    C
ATOM  10421  CA  ARG B 622      -3.614  41.858 134.675  1.00  6.42      B    C
ATOM  10432  CA  TYR B 623      -3.638  39.092 137.312  1.00  6.84      B    C
ATOM  10444  CA  ILE B 624      -2.166  38.601 140.777  1.00  6.95      B    C
```

도면34v

```
ATOM  10452  CA  ALA B 625      -2.036  35.146 142.429  1.00  9.17      B    C
ATOM  10457  CA  THR B 626      -4.883  35.743 144.940  1.00  9.61      B    C
ATOM  10464  CA  GLY B 627      -8.057  37.828 145.428  1.00 11.00      B    C
ATOM  10468  CA  SER B 628      -6.410  40.988 146.753  1.00 14.00      B    C
ATOM  10474  CA  SER B 629      -3.055  42.511 147.827  1.00 14.96      B    C
ATOM  10480  CA  SER B 630      -2.186  45.392 150.152  1.00 15.87      B    C
ATOM  10486  CA  ALA B 631       1.099  45.793 148.230  1.00 15.15      B    C
ATOM  10491  CA  VAL B 632      -0.451  46.993 144.954  1.00 12.68      B    C
ATOM  10498  CA  PRO B 633      -3.837  48.543 144.078  1.00 11.90      B    C
ATOM  10505  CA  SER B 634      -5.790  45.403 143.232  1.00 10.24      B    C
ATOM  10511  CA  GLY B 635      -9.047  43.424 143.655  1.00 10.87      B    C
ATOM  10515  CA  GLY B 636     -10.965  40.434 142.243  1.00 11.37      B    C
ATOM  10519  CA  GLY B 637      -7.618  38.717 141.514  1.00 10.74      B    C
ATOM  10523  CA  GLU B 638      -6.354  41.552 139.331  1.00 10.06      B    C
ATOM  10532  CA  PHE B 639      -4.034  44.542 139.288  1.00 10.96      B    C
ATOM  10543  CA  THR B 640      -6.497  47.479 139.174  1.00 13.37      B    C
ATOM  10550  CA  PRO B 641      -5.496  48.960 136.843  1.00 14.12      B    C
ATOM  10557  CA  PRO B 642      -3.275  46.456 134.964  1.00 13.72      B    C
ATOM  10564  CA  LEU B 643       0.396  47.404 135.238  1.00 14.65      B    C
ATOM  10572  CA  LYS B 644       3.137  47.317 132.691  1.00 13.96      B    C
ATOM  10581  CA  THR B 645       6.095  45.936 134.647  1.00 11.78      B    C
ATOM  10588  CA  LYS B 646       8.969  43.450 134.688  1.00 11.14      B    C
ATOM  10597  CA  VAL B 647       7.800  40.166 136.218  1.00  8.36      B    C
ATOM  10604  CA  TYR B 648      11.303  39.198 137.406  1.00  7.74      B    C
ATOM  10616  CA  ARG B 649      13.591  41.824 139.020  1.00  9.04      B    C
ATOM  10627  CA  LEU B 650      17.093  41.707 140.433  1.00  7.53      B    C
ATOM  10635  CA  ASP B 651      15.807  43.347 143.621  1.00  5.17      B    C
ATOM  10643  CA  PRO B 652      14.545  42.350 147.055  1.00  5.40      B    C
ATOM  10650  CA  ARG B 653      11.139  42.853 145.381  1.00  6.53      B    C
ATOM  10661  CA  GLY B 654      12.108  40.059 142.996  1.00  6.41      B    C
ATOM  10665  CA  MET B 655       8.705  38.795 141.823  1.00  6.72      B    C
ATOM  10673  CA  LEU B 656       5.334  40.460 141.196  1.00  6.70      B    C
ATOM  10681  CA  PRO B 657       3.402  41.236 144.385  1.00  6.71      B    C
ATOM  10688  CA  THR B 658       1.659  37.972 145.528  1.00  7.41      B    C
ATOM  10695  CA  MET B 659       3.892  35.925 143.184  1.00  5.95      B    C
ATOM  10703  CA  ASP B 660       5.190  33.199 145.543  1.00  4.96      B    C
ATOM  10711  CA  ASN B 661       8.255  31.368 144.118  1.00  5.08      B    C
ATOM  10719  CA  ASP B 662       6.702  28.114 145.513  1.00  4.85      B    C
ATOM  10727  CA  ALA B 663       9.647  25.783 146.215  1.00  5.19      B    C
ATOM  10732  CA  GLY B 664      11.674  27.503 143.453  1.00  5.63      B    C
ATOM  10736  CA  THR B 665       9.108  26.987 140.680  1.00  6.11      B    C
ATOM  10743  CA  MET B 666       9.161  30.568 139.404  1.00  7.32      B    C
ATOM  10751  CA  SER B 667      12.973  30.544 139.595  1.00  7.14      B    C
ATOM  10757  CA  THR B 668      13.107  27.187 137.738  1.00  8.41      B    C
ATOM  10764  CA  MET B 669      10.894  28.674 134.975  1.00  8.61      B    C
ATOM  10772  CA  PHE B 670      13.293  31.613 134.474  1.00  5.87      B    C
ATOM  10783  CA  VAL B 671      16.232  29.234 134.165  1.00  5.19      B    C
ATOM  10790  CA  ALA B 672      14.275  27.232 131.532  1.00  5.77      B    C
ATOM  10795  CA  ALA B 673      13.327  30.366 129.576  1.00  6.53      B    C
ATOM  10800  CA  ALA B 674      17.004  31.392 129.591  1.00  6.52      B    C
ATOM  10805  CA  VAL B 675      18.093  27.945 128.278  1.00  6.08      B    C
ATOM  10812  CA  GLY B 676      15.236  28.143 125.757  1.00  5.59      B    C
ATOM  10816  CA  LEU B 677      13.801  24.680 126.537  1.00  4.36      B    C
ATOM  10824  CA  PHE B 678      10.456  24.513 128.310  1.00  4.27      B    C
ATOM  10835  CA  PRO B 679       8.476  21.441 129.468  1.00  4.23      B    C
ATOM  10842  CA  VAL B 680       4.935  22.169 128.193  1.00  4.71      B    C
ATOM  10849  CA  THR B 681       4.133  18.524 128.725  1.00  4.02      B    C
ATOM  10856  CA  ALA B 682       5.534  17.401 132.028  1.00  2.99      B    C
ATOM  10861  CA  GLY B 683       5.162  13.611 132.013  1.00  4.05      B    C
ATOM  10865  CA  SER B 684       6.552  13.443 128.463  1.00  5.13      B    C
ATOM  10871  CA  SER B 685      10.143  13.369 127.150  1.00  6.66      B    C
ATOM  10877  CA  GLN B 686       9.863  16.641 125.200  1.00  6.38      B    C
ATOM  10886  CA  PHE B 687      10.738  20.344 125.639  1.00  5.59      B    C
ATOM  10897  CA  GLN B 688       9.427  23.232 123.520  1.00  6.03      B    C
ATOM  10906  CA  VAL B 689      11.793  25.796 122.043  1.00  6.55      B    C
ATOM  10913  CA  GLY B 690      11.579  29.419 123.274  1.00  7.04      B    C
```

도면34w

```
ATOM  10917  CA  SER B 691      14.106  32.286 122.899  1.00  7.79      B    C
ATOM  10923  CA  PRO B 692      17.484  31.426 124.521  1.00  7.84      B    C
ATOM  10930  CA  PHE B 693      18.864  34.550 126.227  1.00  7.82      B    C
ATOM  10941  CA  PHE B 694      22.594  34.064 125.890  1.00  7.33      B    C
ATOM  10952  CA  ASP B 695      25.205  33.655 123.125  1.00  9.43      B    C
ATOM  10960  CA  SER B 696      26.210  30.332 124.574  1.00  9.32      B    C
ATOM  10966  CA  THR B 697      24.781  28.206 127.383  1.00  8.98      B    C
ATOM  10973  CA  THR B 698      26.586  24.967 128.241  1.00  8.91      B    C
ATOM  10980  CA  ILE B 699      25.297  22.278 130.616  1.00  9.45      B    C
ATOM  10988  CA  THR B 700      28.086  19.908 131.647  1.00 10.12      B    C
ATOM  10995  CA  TYR B 701      27.061  16.493 132.887  1.00  9.88      B    C
ATOM  11007  CA  ASP B 702      28.712  14.246 135.442  1.00 11.45      B    C
ATOM  11015  CA  ASP B 703      30.657  12.267 132.819  1.00 13.40      B    C
ATOM  11023  CA  GLY B 704      32.161  15.461 131.345  1.00 11.98      B    C
ATOM  11027  CA  SER B 705      29.921  15.485 128.268  1.00 10.92      B    C
ATOM  11033  CA  ALA B 706      27.904  18.613 127.584  1.00 10.52      B    C
ATOM  11038  CA  PHE B 707      24.798  20.036 125.938  1.00  9.26      B    C
ATOM  11049  CA  THR B 708      25.464  23.375 124.241  1.00  8.88      B    C
ATOM  11056  CA  VAL B 709      22.737  25.748 123.064  1.00  9.09      B    C
ATOM  11063  CA  THR B 710      24.043  28.747 121.128  1.00 11.17      B    C
ATOM  11070  CA  ALA B 711      22.134  31.834 120.050  1.00 12.24      B    C
ATOM  11075  CA  ASP B 712      24.141  33.564 117.322  1.00 14.25      B    C
ATOM  11083  CA  GLY B 713      23.722  37.334 117.271  1.00 12.57      B    C
ATOM  11087  CA  VAL B 714      21.303  37.259 120.194  1.00 12.10      B    C
ATOM  11094  CA  SER B 715      20.477  40.687 121.622  1.00 11.79      B    C
ATOM  11100  CA  GLU B 716      17.633  42.447 123.438  1.00 13.43      B    C
ATOM  11109  CA  ASP B 717      16.112  43.108 119.987  1.00 12.74      B    C
ATOM  11117  CA  ALA B 718      17.337  39.953 118.227  1.00 11.10      B    C
ATOM  11122  CA  PHE B 719      15.169  37.456 120.135  1.00 10.51      B    C
ATOM  11133  CA  TYR B 720      13.602  35.772 117.045  1.00 10.60      B    C
ATOM  11145  CA  VAL B 721      14.971  32.509 115.687  1.00  9.51      B    C
ATOM  11152  CA  GLN B 722      15.866  32.757 111.949  1.00  9.69      B    C
ATOM  11161  CA  SER B 723      17.260  29.236 111.377  1.00 10.67      B    C
ATOM  11167  CA  ALA B 724      18.749  26.372 113.367  1.00 10.44      B    C
ATOM  11172  CA  THR B 725      21.057  23.385 113.295  1.00 10.19      B    C
ATOM  11179  CA  LEU B 726      21.285  20.299 115.497  1.00  9.79      B    C
ATOM  11187  CA  ASP B 727      24.740  18.682 115.470  1.00 10.78      B    C
ATOM  11195  CA  GLY B 728      25.486  20.479 112.168  1.00  9.59      B    C
ATOM  11199  CA  ALA B 729      22.291  19.355 110.407  1.00  9.37      B    C
ATOM  11204  CA  THR B 730      19.314  21.561 109.427  1.00  9.54      B    C
ATOM  11211  CA  PHE B 731      16.911  21.620 112.380  1.00  9.46      B    C
ATOM  11222  CA  GLY B 732      13.291  22.799 111.894  1.00 10.01      B    C
ATOM  11226  CA  ASN B 733      11.327  21.345 114.825  1.00 10.28      B    C
ATOM  11234  CA  THR B 734       9.967  23.567 117.612  1.00 10.95      B    C
ATOM  11241  CA  TRP B 735      10.620  20.889 120.262  1.00  9.83      B    C
ATOM  11255  CA  VAL B 736      13.510  18.634 121.295  1.00  9.57      B    C
ATOM  11262  CA  ASP B 737      13.581  15.277 123.044  1.00  9.41      B    C
ATOM  11270  CA  TYR B 738      15.214  14.889 126.458  1.00  7.09      B    C
ATOM  11282  CA  ALA B 739      17.644  12.232 125.051  1.00  6.73      B    C
ATOM  11287  CA  THR B 740      19.020  14.795 122.576  1.00  6.66      B    C
ATOM  11294  CA  VAL B 741      19.805  17.297 125.352  1.00  6.54      B    C
ATOM  11301  CA  VAL B 742      21.387  14.908 127.847  1.00  6.53      B    C
ATOM  11308  CA  GLY B 743      23.160  13.136 124.956  1.00  8.60      B    C
ATOM  11312  CA  GLY B 744      25.256  16.293 124.535  1.00  8.96      B    C
ATOM  11316  CA  ALA B 745      23.948  17.869 121.301  1.00  8.77      B    C
ATOM  11321  CA  ASP B 746      24.948  21.245 119.824  1.00  9.99      B    C
ATOM  11329  CA  LEU B 747      21.727  23.156 119.256  1.00  9.20      B    C
ATOM  11337  CA  ALA B 748      22.723  26.298 117.347  1.00  9.57      B    C
ATOM  11342  CA  PHE B 749      20.245  29.087 116.636  1.00  9.42      B    C
ATOM  11353  CA  ARG B 750      20.864  32.051 114.353  1.00 10.76      B    C
ATOM  11364  CA  MET B 751      18.965  35.015 115.906  1.00 10.21      B    C
ATOM  11372  CA  GLY B 752      17.572  38.240 114.387  1.00 12.06      B    C
ATOM  11376  CA  GLU B 753      15.318  41.247 114.999  1.00 14.15      B    C
ATOM  11385  CA  GLN B 754      12.309  40.094 112.922  1.00 15.06      B    C
ATOM  11394  CA  PRO B 755       9.930  37.094 113.128  1.00 14.14      B    C
ATOM  11401  CA  SER B 756      10.685  34.248 110.709  1.00 14.80      B    C
```

도면34x

```
ATOM  11407  CA  ASP B 757       8.804  31.176 109.465  1.00 14.96      B    C
ATOM  11415  CA  TRP B 758      11.407  28.912 111.150  1.00 13.21      B    C
ATOM  11429  CA  GLY B 759       9.689  25.918 112.773  1.00 13.37      B    C
ATOM  11433  CA  THR B 760       6.357  26.434 111.014  1.00 13.94      B    C
ATOM  11440  CA  ASP B 761       6.914  23.033 109.382  1.00 13.96      B    C
ATOM  11448  CA  THR B 762       7.829  21.295 112.640  1.00 10.47      B    C
ATOM  11455  CA  ALA B 763       7.173  17.563 113.052  1.00  9.77      B    C
ATOM  11460  CA  PRO B 764       4.213  17.436 115.464  1.00  9.45      B    C
ATOM  11467  CA  ALA B 765       4.950  17.568 119.216  1.00  7.60      B    C
ATOM  11472  CA  PHE B 766       3.648  14.939 121.659  1.00  7.84      B    C
ATOM  11483  CA  SER B 767      -0.100  14.561 122.179  1.00  8.05      B    C
ATOM  11489  CA  MET B 768      -1.670  11.546 123.865  1.00 10.37      B    C
ATOM  11497  CA  SER B 769      -4.411  11.138 121.212  1.00 13.56      B    C
ATOM  11503  CA  THR B 770      -1.982  11.233 118.242  1.00 17.85      B    C
ATOM  11510  CA  ALA B 771       0.792   9.094 119.762  1.00 20.44      B    C
ATOM  12382 CA    CA C   1      41.425  16.915  81.268  1.00 30.05      C   CA
ATOM  12383 CA    CA C   6       1.503  30.134 143.317  1.00 32.32      C   CA
END
```

서 열 목 록

<110> MUHANN Patent & Law Firm

<120> Hydrolysis of mannose-1-phospho-6-mannose linkage to phospho-6-mannose

<130> APC-2012-0161-DIV

<150> US 61/246,847

<151> 2009-09-29

<160> 58

<170> KoPatentIn 3.0

<210> 1

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<400> 1

Ser Ala Tyr Gln Ser Phe Thr Thr Arg

1 5

<210> 2

<211> 7

<212> PRT

<213> Artificial Sequence

<

220><223> Synthetic peptide

<400> 2

Val Trp Gly Phe Ser His Arg

1 5

<210> 3

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<400> 3

Val Glu Gly Gly Trp Leu Pro Arg

1 5

<210> 4

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<400> 4

Thr Gln Gly Asn Asn Phe Ala Leu Leu Leu Pro Glu Arg

1 5 10

<210> 5

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<400> 5

Asp Val His Ala Glu Leu Thr Ala Met Ala Arg

1 5 10

<210> 6

<211> 2700

<212> DNA

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 6

atgagacgac cacgactcgc cctgctcgcc gcggggctcg cgctcgccgt cgcaccgggc 60

acgctgctgc ccgtcgccgc gggcgccgcc cccgccgacg agggcaccgt caccgccgcc 120

gcgggcgacg acctcacgct cgaggtcaac ccgttcgtcg gcaccgagag cgagggcaac 180

gcctacccgg gcgcgaccgt gccgttcggc atggtccagc tcagcccgga caacacgaac 240
tcctacgcct cgacgtcgta cagcacgaac gcggggcgcg tgtggggctt cagccaccgg 300
cacgtgaaca gcgcgggctg ccccgcggcg ggcgagctgc tcgtcacgcc ggacacgagc 360
gcgaccccgc gcacgtcgcg ctccttcatc gccatcaagg accagaagag caccgagcgc 420
gcgtcggccg ggttctacga ggtgaccctc gcgaacgacg tgcacgccga gctcaccgcg 480
accacgcgcg tcggcgcgca ccgctacacg ttccccgcct cgacgacgtc gcacctgtcg 540
ttcaacgtgg gccagaccct gcgcgacgcc ggcgcgagct cggtgacgtg ggtcgacgac 600

cgcacgctcg agggctgggt cgacaacggc ggcttctgcg gcggcacgcc ggacaagcag 660
cggtacttct tcagcgcgac gttcgaccgc ccggtcgcgt cgagcggcac gtgggggacc 720
gatgcgcgct acgtcgcggg ctccacgacg agcgaggtcg cgggcggcaa caacggcgcc 780
gtcgcggtgt tcgacaccac gaccgaccgc gacgtcgagg tgagcgtggg cgtgtccttc 840
gtgagcgtcg acggcgcgcg cgccaaccgc gaggccgagg ccaccgacga gggcgggcag 900
gtcgcgttcg acaccgtgcg tgaggaggcc cgcgacgcgt ggaacgcgga gctgggccgt 960
gccgcgatcg acgcgtcgcc cgaccagcgc cggatcttct acacccagct ctacaagacg 1020

ctgctgtccc cgacgatcgg cagcgacgtc gacggccggt accgcggcat ggacctcgag 1080
gtccaccagg ccgacggctg ggactactac cagaacttct cgctctggga cacgtaccgc 1140
acgcaggcga cgctgcacgc cctcctgctg cccgagcgcg cgcaggacat cgtgcgctcg 1200
atgtaccagc accgcgtcga gggcggctgg ctgccgcgct ggtccctcgg tgcactggag 1260
accaacatca tggcgggcga ccccgtcacg ccgtggctcg cggagaactt cgcgctcggc 1320
accgtccccg acgacatcgc ggacgagctg tgggactacc tcgtcgagaa cgccacgacg 1380
accccgccgg acgacgtcgc gtccgtcggg cggcgcagca ccgagttcta cgccgagcac 1440

ggccacgtgc cgttctaccc cgagaacgag ggcggcctcg gcggccagtt cgaggagtac 1500
cgccacggcg gctcggcgac gctcgagctc gcgctcgccg acgcgagcct cggcgctgcg 1560
gccgagcgca cgggtcgcga gggcggccag gcgttcctcg acaagggtcg caactggcgc 1620
aacctctgga acccggacgt cgagctctcg ggtggcttcc agggcatggt caacgcgaag 1680
cgcccgacgg gcgagttcgt cacgctgccc gagctgacgg acgtcacgcg ctccggcttc 1740
cacgagggcg tgccgtggca gtaccagtgg atggtgccgc aggacgtcac gggcctccag 1800
gaggtcatgg gcggcgagga cggcttcgtc gagcgtctcg actactactt cgaccagccg 1860

gcgctcgccg cgaaccccgg cgtctcgccg agcacgtggg ccaagggcgg cagctcgtac 1920

tacacgacca tccgctacaa cccgggcaac gagccgacga tcatgaacgc gtggctctac 1980

ggctacgtgg gccagccgtg gaagacgaac gacgtcctcg ccgcgaacct caaccgcttc 2040

ccggacaccc cgggcggcgg cgtcgggaac gacgacctcg gcacgcttgc cgcctggtac 2100

gtcatggcgt cgctcgggtt cgagcccgtc atgccgggct cggggatcct cgcgctcaac 2160

gcgccgaagg tgcaggccgc gacgctcacg accgatgccg gggcgacgct gcgcatcgac 2220

gcggcgggcg cgaacgagaa gctcccgagc tacgtcgccg gcctggaggt cgacggcgtc 2280

gcgcacaccg ccgcgtggct cgacgtcgcg gcgctgcagg acggcggcac gctcgacttc 2340

gacctgtccg gcacgagcgc gggcctcacg tggggcaccg gcgcggccga ccgcatcccg 2400

tcggtctccg ccgtcgcccc gcccgcgccg gtcgaggtcg aggcgagcgc gcgctgcctc 2460

ggcggccggg cgttcgtcgc ggtccgcgcg accagcacgg ccgacgcgcc ggtggacgtg 2520

actctcacga cgccgttcgg cgagcggacg gtccggcacg tgcagccggg caggagcgcc 2580

taccagtcgt tcacgacgcg cacgacgtcc gtcgaggccg ggacggcgac cgtcacggtc 2640

gtcgccgcgg acggcacgac gtcgacggtc gacgcggcgt acgaggcgct ggcctgcggc 2700

2700

<210> 7

<211> 900

<212> PRT

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 7

Met Arg Arg Pro Arg Leu Ala Leu Leu Ala Ala Gly Leu Ala Leu Ala

1 5 10 15

Val Ala Pro Gly Thr Leu Leu Pro Val Ala Ala Gly Ala Ala Pro Ala

20 25 30

Asp Glu Gly Thr Val Thr Ala Ala Ala Gly Asp Asp Leu Thr Leu Glu

35 40 45

Val Asn Pro Phe Val Gly Thr Glu Ser Glu Gly Asn Ala Tyr Pro Gly

50 55 60

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

65 70 75 80

Ser Tyr Ala Ser Thr Ser Tyr Ser Thr Asn Ala Gly Arg Val Trp Gly

85 90 95

Phe Ser His Arg His Val Asn Ser Ala Gly Cys Pro Ala Ala Gly Glu

100 105 110

Leu Leu Val Thr Pro Asp Thr Ser Ala Thr Pro Arg Thr Ser Arg Ser

115 120 125

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

130 135 140

Phe Tyr Glu Val Thr Leu Ala Asn Asp Val His Ala Glu Leu Thr Ala

145 150 155 160

Thr Thr Arg Val Gly Ala His Arg Tyr Thr Phe Pro Ala Ser Thr Thr

165 170 175

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

180 185 190

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

195 200 205

Asn Gly Gly Phe Cys Gly Gly Thr Pro Asp Lys Gln Arg Tyr Phe Phe

210 215 220

Ser Ala Thr Phe Asp Arg Pro Val Ala Ser Ser Gly Thr Trp Gly Thr

225 230 235 240

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

245 250 255

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

260 265 270

Glu Val Ser Val Gly Val Ser Phe Val Ser Val Asp Gly Ala Arg Ala

275 280 285

Asn Arg Glu Ala Glu Ala Thr Asp Glu Gly Gly Gln Val Ala Phe Asp

290 295 300

Thr Val Arg Glu Glu Ala Arg Asp Ala Trp Asn Ala Glu Leu Gly Arg

305 310 315 320

Ala Ala Ile Asp Ala Ser Pro Asp Gln Arg Arg Ile Phe Tyr Thr Gln

325 330 335

Leu Tyr Lys Thr Leu Leu Ser Pro Thr Ile Gly Ser Asp Val Asp Gly
340 345 350

Arg Tyr Arg Gly Met Asp Leu Glu Val His Gln Ala Asp Gly Trp Asp
355 360 365

Tyr Tyr Gln Asn Phe Ser Leu Trp Asp Thr Tyr Arg Thr Gln Ala Thr
370 375 380

Leu His Ala Leu Leu Leu Pro Glu Arg Ala Gln Asp Ile Val Arg Ser
385 390 395 400

Met Tyr Gln His Arg Val Glu Gly Gly Trp Leu Pro Arg Trp Ser Leu
405 410 415

Gly Ala Leu Glu Thr Asn Ile Met Ala Gly Asp Pro Val Thr Pro Trp
420 425 430

Leu Ala Glu Asn Phe Ala Leu Gly Thr Val Pro Asp Asp Ile Ala Asp
435 440 445

Glu Leu Trp Asp Tyr Leu Val Glu Asn Ala Thr Thr Thr Pro Pro Asp
450 455 460

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

465 470 475 480

Gly His Val Pro Phe Tyr Pro Glu Asn Glu Gly Gly Leu Gly Gly Gln
485 490 495

Phe Glu Glu Tyr Arg His Gly Gly Ser Ala Thr Leu Glu Leu Ala Leu
500 505 510

Ala Asp Ala Ser Leu Gly Ala Ala Ala Glu Arg Thr Gly Arg Glu Gly
515 520 525

Gly Gln Ala Phe Leu Asp Lys Gly Arg Asn Trp Arg Asn Leu Trp Asn
530 535 540

Pro Asp Val Glu Leu Ser Gly Gly Phe Gln Gly Met Val Asn Ala Lys
545 550 555 560

Arg Pro Thr Gly Glu Phe Val Thr Leu Pro Glu Leu Thr Asp Val Thr
565 570 575

Arg Ser Gly Phe His Glu Gly Val Pro Trp Gln Tyr Gln Trp Met Val
580 585 590

Pro Gln Asp Val Thr Gly Leu Gln Glu Val Met Gly Gly Glu Asp Gly
595 600 605
Phe Val Glu Arg Leu Asp Tyr Tyr Phe Asp Gln Pro Ala Leu Ala Ala

610 615 620
Asn Pro Gly Val Ser Pro Ser Thr Trp Ala Lys Gly Gly Ser Ser Tyr
625 630 635 640
Tyr Thr Thr Ile Arg Tyr Asn Pro Gly Asn Glu Pro Thr Ile Met Asn
645 650 655
Ala Trp Leu Tyr Gly Tyr Val Gly Gln Pro Trp Lys Thr Asn Asp Val
660 665 670
Leu Ala Ala Asn Leu Asn Arg Phe Pro Asp Thr Pro Gly Gly Gly Val
675 680 685

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

755 760 765
Val Ala Ala Leu Gln Asp Gly Gly Thr Leu Asp Phe Asp Leu Ser Gly
770 775 780
Thr Ser Ala Gly Leu Thr Trp Gly Thr Gly Ala Ala Asp Arg Ile Pro
785 790 795 800
Ser Val Ser Ala Val Ala Pro Pro Ala Pro Val Glu Val Glu Ala Ser
805 810 815
Ala Arg Cys Leu Gly Gly Arg Ala Phe Val Ala Val Arg Ala Thr Ser
820 825 830

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

835 840 845

Arg Thr Val Arg His Val Gln Pro Gly Arg Ser Ala Tyr Gln Ser Phe

850 855 860

Thr Thr Arg Thr Thr Ser Val Glu Ala Gly Thr Ala Thr Val Thr Val

865 870 875 880

Val Ala Ala Asp Gly Thr Thr Ser Thr Val Asp Ala Ala Tyr Glu Ala

885 890 895

Leu Ala Cys Gly

900

<210> 8

<211> 3543

<212> DNA

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 8

gtgagcctcg cgctcccgct ggcggcgtac gcggcgcccg ggatcggggc gtcgcccgcg 60

accgccgccg ggacggaggc agcgacgggg tccgatgccg ccgccgtcga cggcccgctg 120

gtcgactacg tcaacccgtt catcgggacc aaggacgacg gcaacaccta cccgggcgct 180

gccgtgccgt tcggcatggt gcaactctcg ccggacaacg gccacaacgt cgggtacgac 240

tacgaccgca cgtcggtgcg cgggttctcg ctcgtgcacc tgtccggcgt cggctgcggc 300

ctcggcggtc cgctcccgac cctgccgacg acgggcgcga tcacctcgac cgactacggc 360

cagtacgcgc tcggtttctc gcacgacgac gaggaggcct cgccggggta ctaccgcgtg 420

ggtctccagg cgccggcggg cacgatcgag gccgagctca ccgcgaccga gcgcacgggc 480

gtccagcggt acacgttccc cgcgacggcg caggccaacg tcctgctcaa cgccggccag 540

gcgctcaacc gggtgacgga gtccgacgtg cgcgtcgtgg acgaccgcac ggtcgagacg 600

cgcatcaccg tccgcggctt ctgccaggac accgagccgc agacgatctg gacccgcacg 660

accttcgacc ggccgttcgt cgcgcacggc acgtgggacg gccaggtcgt caccgcgggc 720

gcggacgccg cgtccggcgg cgagggccgt cgcggcgcgt acgtcacgtt cgacacgacc 780

ggcggcgacc tcgacgtcga ggccgtcacc gcgatgagct acgtgggcgc cgacggcgcc 840

gcggcgaacc tcgccgcgga ggccggcacg ttcgacgccg tgcacgacgc cgcgcgctcg 900

gcctgggagg agcggctcgg cctcgtgcgg gtcgcgcagg gcgacccgga cgacctgcgc 960

accttctact cctcgctcta ccgcagcttc ctcgcgccga acgtcggctc cgacgtcgac 1020

gggcgctacc gcggctggga ccaggaggtc cacgccgcgg aaccggactt cacctactac 1080

cagaactact cgctctggga cacgtaccgc acccagcagc agctcctgta cctgctcgcg 1140

cccgacgagt cggccgacat ggcgctctcg ctcgtgcgcc agggccagca gggcgggtgg 1200

ctcccgcgct ggggctacgg cacggtcgag acgaacatca tgaccggcga cccggcgacg 1260

ccgttcctcg tcagcgcctg gcgccagggc ctgctcgcgg gccacgagga ggaggcgtac 1320

gcggtcctga gggagaacgc cgacggcgtc ccgcccgccg actcgccctt caacgggcgc 1380

gcggcgaacg tcgagtacct gcgcgacggg ttcgtcccgc acgagccggc gcgctcgggc 1440

aagcccggcg actacgacct ccagcacggc gcctcggcga ccatggagta cgccctcgcc 1500

gacgcgatgc tctcgaccat ggcgcgcggc ctcggccacg acgaggacgc cgaccggtac 1560

gccgcccgcg gccagagcta ccgcaacgtg ttcgacccgc gcacgggcaa cttccgggcg 1620

cgtaacgcgg acggcttctt cgtgggcgac gcggaccccg cgcactccga cgggttccac 1680

gagggcacgg cggtgcagta ccagtggctc gtgccccagg acgtgccggg cctgttcgac 1740

ctcatgggcg gcaccgacgc cgcggtcgac cgcctcgatg cgttcttcgc gtacgacgag 1800

ctcgtcgccg acccccccgca cgtcgcgagc gaggtgtggg tcaacggcac gtacgactac 1860

tacggctggg agacctacaa cccgaacaac gagcccaacc tccatgcgcc gtacgtctac 1920

ctgtggaccg ggcagccctg gaagacgacg gacgtcgtgc gcgccgcgtc gaccctcttc 1980

accgacggcc ccgacggcgt cacgggcaac gacgacctcg gcacgatgtc cgcgtggcac 2040

gtgctgtcgt cgatcggcgt gtacccgatc gtgccgggcg ccgatctgtg gggcctgacg 2100

acgccgctct tcgacgacgt gacgatcacg ctcgacccgg aggtcttcgg tcgggactcc 2160

ctgcgcctca cggcggacgg cgtcgcgccc gacacgcact acacgcagtc cgtgtcgctc 2220

ggcggcgagc cgctcgatcg cgcctgggtc acgggcgacg agctcaccgc ggccggcacg 2280

ctcgacgtga ccgtcggcac cgagccgtcc gcgtgggcga ccgaccccgc ggcctcgccg 2340

ggcgccgtcg tgcctgcgga cggcacggtc gagcgcctgt tcgtcggcgc gacgccgcgg 2400

cagccggtcc tcgccccggg cgggcggacc gaggtcgcag tccaggtcgt cgcccagggc 2460

gcggggacgt ccagcgggac gctcgaggtg acgtccgacg gcgcggtcac cgcgacgacc 2520

gacctcgccg agtggaccgc cgagtccgac ggcctgccgg ccacggtcga gggcacggtg 2580

acgatcgagg ctcccgccga cgccgagccg ggtctgcaca cggtgcggct cgtcgtgcgc 2640

gacgccgcgg ggaccgaggc ggtccgcgag gtctcggtcg tcgtgtccgg ggagtcgtgg 2700

atcgccgacg cgttcgacaa cgtcggcatc ggcgacgccg gggcggccaa cgcgaacctc 2760

gacggctcgg gcgcctacct cctgcgcgac ctgctcgcgg acctcggcgc cgtccagggc 2820

ctggagctca ccgtgccggg cacggacctc acctacacgc tcggggcccc gcgggcgggc 2880
gcgcccgaca acgtcgccgc gagcggcgag gtcctcgagg tgcccgagca cctgcgctcg 2940
gcccgccacc tctcggtggt cgggacgagc acgcacggca cgcacggggg cggcctcgtc 3000
ctcgggttcg ccgacggctc gtcgcagacc gtcgacgtgc gcctcagcga ctggtgcacg 3060
ggctcgcccg agcccggcaa catcacggtc gcgaaggccg gggcgcgcgg cgaccgcgag 3120
aacgtgcaga agatcggctg cggcctctac gccaccgcgc ccgtcgcgat ccccgagggc 3180
aaggtcctga cgtcggtcac gctgccgagc gacgagcggt tccacgtgtt cgcgatcgcg 3240

accgacgcga cgggggacgt ccccgcgccg caggtcgagg tcacggcgca ggcccgctgc 3300
ctcggcggca aggcgttcgt cgcggtgcgc gcgctcaaca cgggcgagca gcccgccgcg 3360
atcgagctcg cgaccccgta cggctccaag ctcttcggtg acgtcgctcc cggggcgaac 3420
gcgtaccagt cgttcgccac ccgcgccgcc gccgtcgagg cgggcgaggt cacggtgacc 3480
gtgacgacgc ccgacggcga gccccagcag gtcacggccg cgtacgacgc cgccgcctgc 3540
tcc 3543

<210> 9

<211> 1181

<212> PRT

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 9

Val Ser Leu Ala Leu Pro Leu Ala Ala Tyr Ala Ala Pro Gly Ile Gly
1 5 10 15

Ala Ser Pro Ala Thr Ala Ala Gly Thr Glu Ala Ala Thr Gly Ser Asp
20 25 30

Ala Ala Ala Val Asp Gly Pro Leu Val Asp Tyr Val Asn Pro Phe Ile
35 40 45

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

50 55 60

Gly Met Val Gln Leu Ser Pro Asp Asn Gly His Asn Val Gly Tyr Asp
65 70 75 80

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

85 90 95

Val Gly Cys Gly Leu Gly Gly Pro Leu Pro Thr Leu Pro Thr Thr Gly

100 105 110

Ala Ile Thr Ser Thr Asp Tyr Gly Gln Tyr Ala Leu Gly Phe Ser His

115 120 125

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

130 135 140

Pro Ala Gly Thr Ile Glu Ala Glu Leu Thr Ala Thr Glu Arg Thr Gly

145 150 155 160

Val Gln Arg Tyr Thr Phe Pro Ala Thr Ala Gln Ala Asn Val Leu Leu

165 170 175

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

180 185 190

Val Asp Asp Arg Thr Val Glu Thr Arg Ile Thr Val Arg Gly Phe Cys

195 200 205

Gln Asp Thr Glu Pro Gln Thr Ile Trp Thr Arg Thr Thr Phe Asp Arg

210 215 220

Pro Phe Val Ala His Gly Thr Trp Asp Gly Gln Val Val Thr Ala Gly

225 230 235 240

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

245 250 255

Phe Asp Thr Thr Gly Gly Asp Leu Asp Val Glu Ala Val Thr Ala Met

260 265 270

Ser Tyr Val Gly Ala Asp Gly Ala Ala Ala Asn Leu Ala Ala Glu Ala

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

Ser Asp Val Asp Gly Arg Tyr Arg Gly Trp Asp Gln Glu Val His Ala

340 345 350

Ala Glu Pro Asp Phe Thr Tyr Tyr Gln Asn Tyr Ser Leu Trp Asp Thr

355 360 365

Tyr Arg Thr Gln Gln Gln Leu Leu Tyr Leu Leu Ala Pro Asp Glu Ser

370 375 380

Ala Asp Met Ala Leu Ser Leu Val Arg Gln Gly Gln Gln Gly Gly Trp

385 390 395 400

Leu Pro Arg Trp Gly Tyr Gly Thr Val Glu Thr Asn Ile Met Thr Gly

405 410 415

Asp Pro Ala Thr Pro Phe Leu Val Ser Ala Trp Arg Gln Gly Leu Leu

420 425 430

Ala Gly His Glu Glu Glu Ala Tyr Ala Val Leu Arg Glu Asn Ala Asp

435 440 445

Gly Val Pro Pro Ala Asp Ser Pro Phe Asn Gly Arg Ala Ala Asn Val

450 455 460

Glu Tyr Leu Arg Asp Gly Phe Val Pro His Glu Pro Ala Arg Ser Gly

465 470 475 480

Lys Pro Gly Asp Tyr Asp Leu Gln His Gly Ala Ser Ala Thr Met Glu

485 490 495

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

500 505 510

His Asp Glu Asp Ala Asp Arg Tyr Ala Ala Arg Gly Gln Ser Tyr Arg

515 520 525

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

530 535 540

Gly Phe Phe Val Gly Asp Ala Asp Pro Ala His Ser Asp Gly Phe His

545 550 555 560

Glu Gly Thr Ala Val Gln Tyr Gln Trp Leu Val Pro Gln Asp Val Pro

565 570 575

Gly Leu Phe Asp Leu Met Gly Gly Thr Asp Ala Ala Val Asp Arg Leu

580 585 590

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

595 600 605

Ala Ser Glu Val Trp Val Asn Gly Thr Tyr Asp Tyr Tyr Gly Trp Glu

610 615 620

Thr Tyr Asn Pro Asn Asn Glu Pro Asn Leu His Ala Pro Tyr Val Tyr

625 630 635 640

Leu Trp Thr Gly Gln Pro Trp Lys Thr Thr Asp Val Val Arg Ala Ala

645 650 655

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

660 665 670

Leu Gly Thr Met Ser Ala Trp His Val Leu Ser Ser Ile Gly Val Tyr

675 680 685

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

690 695 700

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

705 710 715 720

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

725 730 735

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

740 745 750

Asp Glu Leu Thr Ala Ala Gly Thr Leu Asp Val Thr Val Gly Thr Glu

755 760 765

Pro Ser Ala Trp Ala Thr Asp Pro Ala Ala Ser Pro Gly Ala Val Val

770 775 780

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

785 790 795 800

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

805 810 815

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

820 825 830

Asp Gly Ala Val Thr Ala Thr Thr Asp Leu Ala Glu Trp Thr Ala Glu

835 840 845

Ser Asp Gly Leu Pro Ala Thr Val Glu Gly Thr Val Thr Ile Glu Ala
850 855 860
Pro Ala Asp Ala Glu Pro Gly Leu His Thr Val Arg Leu Val Val Arg
865 870 875 880
Asp Ala Ala Gly Thr Glu Ala Val Arg Glu Val Ser Val Val Val Ser
885 890 895
Gly Glu Ser Trp Ile Ala Asp Ala Phe Asp Asn Val Gly Ile Gly Asp
900 905 910
Ala Gly Ala Ala Asn Ala Asn Leu Asp Gly Ser Gly Ala Tyr Leu Leu

915 920 925
Arg Asp Leu Leu Ala Asp Leu Gly Ala Val Gln Gly Leu Glu Leu Thr
930 935 940
Val Pro Gly Thr Asp Leu Thr Tyr Thr Leu Gly Ala Pro Arg Ala Gly
945 950 955 960
Ala Pro Asp Asn Val Ala Ala Ser Gly Glu Val Leu Glu Val Pro Glu
965 970 975
His Leu Arg Ser Ala Arg His Leu Ser Val Val Gly Thr Ser Thr His
980 985 990

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

1060 1065 1070
Arg Phe His Val Phe Ala Ile Ala Thr Asp Ala Thr Gly Asp Val Pro
1075 1080 1085

Ala Pro Gln Val Glu Val Thr Ala Gln Ala Arg Cys Leu Gly Gly Lys
1090 1095 1100

Ala Phe Val Ala Val Arg Ala Leu Asn Thr Gly Glu Gln Pro Ala Ala
1105 1110 1115 1120

Ile Glu Leu Ala Thr Pro Tyr Gly Ser Lys Leu Phe Gly Asp Val Ala
1125 1130 1135

Pro Gly Ala Asn Ala Tyr Gln Ser Phe Ala Thr Arg Ala Ala Ala Val
1140 1145 1150

Glu Ala Gly Glu Val Thr Val Thr Val Thr Thr Pro Asp Gly Glu Pro
1155 1160 1165

Gln Gln Val Thr Ala Ala Tyr Asp Ala Ala Ala Cys Ser
1170 1175 1180

<210> 10

<211> 3453

<212> DNA

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 10

gtgcggcgct ccgtcgcggc gctctctgcc acggcggtcc tggccgccgg actctcgatc 60

gcgcccgccg tcgggctcgc ggtcccggcg gtcgcggccg cacccgacct cgttgaggac 120

cccgtctcct tcgtcgaccc gttcgtcggg accggccagg cgacgggcgt cgtcggggag 180

atcaacaact tccccgggcc gtcgatgccg ttcggcatga tgcagctctc gcccgacacc 240

caggtctccg tgggcaacgg cgacaaggcg tacgcgggct accgctactc gcaccaggcg 300

atccgcggct tctccatgac gcacgcggcc gccgggtgct ggatcttcgg cgacgtcccg 360

atcctccccg tgacgggcga cgtcgggcag tacccgtggg accgcaagga ggcgttcagc 420

cacgacgcgg agagcgccga ggtcggccgg tacgcggtca cgctccagtc gtcggggatc 480

gatgcggagg tgtcggccgc gacccgctcg ggcggactga cgttcgacta ccccgagggc 540

ggtgccgcgt cgcaggtgat cgtcaacgcc gcgggctcgc tcgcgagcgt gcgcaacgcg 600

acggtcgagg tcgaggacgc gcgcacggtc accggctcgg tgacgagcgg cgggttctgc 660

ggcaagaaca acacgcacac gacgtacttc gcgatcgagc tcgaccagga cgcgcaggcg 720

ttcggcacgt ggcagggctc gaccgtctcg cccggcgacc cgtcggccga cggcaacggc 780

gcgggcgcgt ggctcacctt cgcgcccggc gcgacggtgc acgcgaaggt cggcatgtcc 840

tacgtgagcg tcgagggcgc gcgcgccaac ctcgcggccg agatcccggg cttcgacttc 900

gacgccgtcc gggacgccaa ccgcgccgcc tggtccgacc tgctcggcaa ggtccgcgtc 960

gcggggcagg acgccgacga cctcaccatg ttctacacgt cgctctacca ctcgctgctg 1020

cacccgaaca cgttcaccga cgtggacggc cggtacgtcg ggttcgacgg ggagatccac 1080

caggcccccg aggggcacga gcggtacgcg aacttctccg actgggacac gtaccggtcg 1140

ctcggcgcgc tccaggcgct gctggcgccc gaccaggcgt cggacatggc gcagtcgctc 1200

gtcgaggtcg ccgaccagtc cggctggctg ccgcgctggc ccgtcgcgaa ccagcacacg 1260

ggccagatga ccggtgactc ctcggtgccg ctcatcgcga gcatgtacgc gttcggggcg 1320

cgcgacttcg acgcggagtc ggcgctcgcg cacatggtca agggtgcgac gagcgccgcc 1380

ccgaccgcga acggctacgt gcagcggcgc gggatcgaga cgtacctcga gcgcggctac 1440

gcgccccaga ccgaggagtt ccggggcgac caccgcgtcg tcggcgcgtc gatcacgctc 1500

gagtggtcga tcgccgactt cgcgatcggg cagctcgcgg ccgcgctcgg ccaggacgac 1560

gtcgccaccg agtacgccgc ccgcggccag tggtggcaga acgtccacga ccccgtgacc 1620

cgcacggcgg gcgcccggaa cgacgacggc acgttcgtgc ggtcgcaggg cggcggcggg 1680

ttcgggcagg agggcttcga cgagggcaac gccgagcagt acacgtggct cgtgccgcag 1740

aacgtcgcgg ggctcaccga cgcgctcggc gggcgtgagg ccgtcgcgga gcggctcgat 1800

gccttcacgg tgcagcacaa cgccggcccg aacgagccgt acctgtggat cggcaacgag 1860

ccgaacttcg gcgtcccgtg gctgtacgac tacgtgggcc agccgtggcg gacgagcgag 1920

ctcgtggacg agcttacgtc cacgctgttc cggcccgagc cgaacggcaa gcccggcaac 1980

gacgacctcg gcgcccaggc cggctggtac gtgtgggccg cgatgggcct gtaccccacc 2040

acgccgggca cggacgtgct cgcgctcaac gcgccgcgct tcgaccgcgt cgtggtcgac 2100

ctcggcgagg gcgacaccct cgacctgcgc gcccccggcg cctcgaccgg cgcccgctac 2160

atcagcggcg tcaccatcga cggcgcagcc tgggacggga cctccctgcc gcgccacgtc 2220

gcgcacgacg gcggcgtcgt cgagctcgcg atgtcgaccg cacgcgacac gacgtggggg 2280

accgcagccg aggacgcccc gccgtcgtgg cgcgacggcg agtccgccgt ggtcgccgcc 2340

gcggacccgg gcctcgtgac ggtcgccccc ggcgggaccg ccgacgcgtc ggtggccgtg 2400

cagctcttcg gcgccgacgc cgccgacgtg cgcgtcgcgg tcgacgcgcc cgggggcatc 2460

ggggtcggtg agcccgcgct cgtcgacgac ggctcgggcc acctcaccgg gacggtcccc 2520

gtccaggtgg gtgccggcgt cgcgtccggc taccacgacg cgcgcctcgt gctctcggcc 2580

ggggacgacg acgtcgaggt gcccctcacc gtcctcgtcg ccgcgcccgg gtcgctcgtt 2640

gcggcctacg acacggtcgg cactgcgccc gaggcgaacc gcggcgtcgg gaacttcgac 2700

gcggccggca actcgttctc gcgcgaggcg ctcgccgacg cgggtctcac gcccgggtcg 2760

gcgcacgacg tcgacggcct ggcgttcacg tggccgtcct cacccgtggg gcgcccggac 2820

tcggtcacgc tcaccggcga gaccgtgcgg ctcgacgcgc cgacgagccg gctcgcgttc 2880

gtgggcgccg cgaccgacgg gacccatcgc gggaccgcgg tcgtgacgtt cgacgacggc 2940

agcaccgcga ccacgacgat cggcttcggc gactgggtgc tgccgagcgc ggacggctcg 3000

ccggtcgagg gcaactcggt cgtcgcgcag atgaaccggc gcaacggcga caaggacagc 3060

gcgttcgtgt tcgccaccgc cccgtacacc gcgcccgagg accgccgcgt ggtcgcggtg 3120

aggttccccg acgtcgacga cctgcacgtc tttgcgatcg cgaccgagcc ggccgcggac 3180

gtgcacctcg tggacgtgac ggtctccctg cgctgcctcg ccgggacccc gtacgtggcg 3240

gtgcgcgcgg cgaacgtctc cgccggggcc gtcgacgtcg acctcacgac gggcgtgggc 3300

tcgcggtcct tcacggccgt cgcccccggc gccaacgcct accagtcgtt cgccgcccgc 3360

ggcgcgaccg ggaacgtcga cgtcaccgtc acggccacgg gggaggaggg gacgcagacg 3420

gtcgcgcgga ccgtcgtcgt cccgcgctgc tcc 3453

<210> 11

<211> 1151

<212> PRT

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 11

Val Arg Arg Ser Val Ala Ala Leu Ser Ala Thr Ala Val Leu Ala Ala
1 5 10 15

Gly Leu Ser Ile Ala Pro Ala Val Gly Leu Ala Val Pro Ala Val Ala
20 25 30

Ala Ala Pro Asp Leu Val Glu Asp Pro Val Ser Phe Val Asp Pro Phe
35 40 45

Val Gly Thr Gly Gln Ala Thr Gly Val Val Gly Glu Ile Asn Asn Phe
50 55 60

Pro Gly Pro Ser Met Pro Phe Gly Met Met Gln Leu Ser Pro Asp Thr
65 70 75 80

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

85 90 95

Ser His Gln Ala Ile Arg Gly Phe Ser Met Thr His Ala Ala Ala Gly

100 105 110

Cys Trp Ile Phe Gly Asp Val Pro Ile Leu Pro Val Thr Gly Asp Val

115 120 125

Gly Gln Tyr Pro Trp Asp Arg Lys Glu Ala Phe Ser His Asp Ala Glu

130 135 140

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

145 150 155 160

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

165 170 175

Tyr Pro Glu Gly Gly Ala Ala Ser Gln Val Ile Val Asn Ala Ala Gly

180 185 190

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

195 200 205

Thr Val Thr Gly Ser Val Thr Ser Gly Gly Phe Cys Gly Lys Asn Asn

210 215 220

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

225 230 235 240

Phe Gly Thr Trp Gln Gly Ser Thr Val Ser Pro Gly Asp Pro Ser Ala

245 250 255

Asp Gly Asn Gly Ala Gly Ala Trp Leu Thr Phe Ala Pro Gly Ala Thr

260 265 270

Val His Ala Lys Val Gly Met Ser Tyr Val Ser Val Glu Gly Ala Arg

275 280 285

Ala Asn Leu Ala Ala Glu Ile Pro Gly Phe Asp Phe Asp Ala Val Arg

290 295 300

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

305 310 315 320

Ala Gly Gln Asp Ala Asp Asp Leu Thr Met Phe Tyr Thr Ser Leu Tyr

325 330 335

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

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

465                 470                 475                 480
Ala Pro Gln Thr Glu Glu Phe Arg Gly Asp His Arg Val Val Gly Ala
                485                 490                 495
Ser Ile Thr Leu Glu Trp Ser Ile Ala Asp Phe Ala Ile Gly Gln Leu
            500                 505                 510
Ala Ala Ala Leu Gly Gln Asp Asp Val Ala Thr Glu Tyr Ala Ala Arg
        515                 520                 525
Gly Gln Trp Trp Gln Asn Val His Asp Pro Val Thr Arg Thr Ala Gly
    530                 535                 540

Ala Arg Asn Asp Asp Gly Thr Phe Val Arg Ser Gln Gly Gly Gly Gly
545                 550                 555                 560
Phe Gly Gln Glu Gly Phe Asp Glu Gly Asn Ala Glu Gln Tyr Thr Trp
                565                 570                 575
Leu Val Pro Gln Asn Val Ala Gly Leu Thr Asp Ala Leu Gly Gly Arg
            580                 585                 590
```

Glu Ala Val Ala Glu Arg Leu Asp Ala Phe Thr Val Gln His Asn Ala
595 600 605
Gly Pro Asn Glu Pro Tyr Leu Trp Ile Gly Asn Glu Pro Asn Phe Gly

610 615 620
Val Pro Trp Leu Tyr Asp Tyr Val Gly Gln Pro Trp Arg Thr Ser Glu
625 630 635 640
Leu Val Asp Glu Leu Thr Ser Thr Leu Phe Arg Pro Glu Pro Asn Gly
645 650 655
Lys Pro Gly Asn Asp Asp Leu Gly Ala Gln Ala Gly Trp Tyr Val Trp
660 665 670
Ala Ala Met Gly Leu Tyr Pro Thr Thr Pro Gly Thr Asp Val Leu Ala
675 680 685

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

755 760 765
Ser Trp Arg Asp Gly Glu Ser Ala Val Val Ala Ala Ala Asp Pro Gly
770 775 780
Leu Val Thr Val Ala Pro Gly Gly Thr Ala Asp Ala Ser Val Ala Val
785 790 795 800
Gln Leu Phe Gly Ala Asp Ala Ala Asp Val Arg Val Ala Val Asp Ala
805 810 815
Pro Gly Gly Ile Gly Val Gly Glu Pro Ala Leu Val Asp Asp Gly Ser
820 825 830

Gly His Leu Thr Gly Thr Val Pro Val Gln Val Gly Ala Gly Val Ala

835 840 845

Ser Gly Tyr His Asp Ala Arg Leu Val Leu Ser Ala Gly Asp Asp Asp

850 855 860

Val Glu Val Pro Leu Thr Val Leu Val Ala Ala Pro Gly Ser Leu Val

865 870 875 880

Ala Ala Tyr Asp Thr Val Gly Thr Ala Pro Glu Ala Asn Arg Gly Val

885 890 895

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

900 905 910

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

915 920 925

Phe Thr Trp Pro Ser Ser Pro Val Gly Arg Pro Asp Ser Val Thr Leu

930 935 940

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

945 950 955 960

Val Gly Ala Ala Thr Asp Gly Thr His Arg Gly Thr Ala Val Val Thr

965 970 975

Phe Asp Asp Gly Ser Thr Ala Thr Thr Thr Ile Gly Phe Gly Asp Trp

980 985 990

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

995 1000 1005

Ala Gln Met Asn Arg Arg Asn Gly Asp Lys Asp Ser Ala Phe Val Phe

1010 1015 1020

Ala Thr Ala Pro Tyr Thr Ala Pro Glu Asp Arg Arg Val Val Ala Val

1025 1030 1035 1040

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

1045 1050 1055

Pro Ala Ala Asp Val His Leu Val Asp Val Thr Val Ser Leu Arg Cys

1060 1065 1070

Leu Ala Gly Thr Pro Tyr Val Ala Val Arg Ala Ala Asn Val Ser Ala

1075 1080 1085

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

1090 1095 1100

Thr Ala Val Ala Pro Gly Ala Asn Ala Tyr Gln Ser Phe Ala Ala Arg

1105 1110 1115 1120

Gly Ala Thr Gly Asn Val Asp Val Thr Val Thr Ala Thr Gly Glu Glu

1125 1130 1135

Gly Thr Gln Thr Val Ala Arg Thr Val Val Val Pro Arg Cys Ser

1140 1145 1150

<210> 12

<211> 5433

<212> DNA

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 12

atgaccagac cactcccgcc cggacgcgcg gtcgcgcggt ccggcagcgg ccgcgcccgg 60

cccctcggcc tcgtgctcgc cgccgcactc gccgtcccgc tcggggtgcc tctcgcggcc 120

cccgcgggag ccctcgctgc cgcgcccgcc gcggccgccg agcccggcga cttctcgtcc 180

tcgttcgagt ccggcgaccc ggccgcgctg cccaccaccg tggcggagcg cgacggcgcg 240

ccctggcagg cgaacgtcgg ctcgttcacg gccggcctgc ccgggagcgt cctcgggcag 300

ctgaagggcg tcacggcgag cgcgcagaac ctgcccaacg agggcgcggc gaacctcgcc 360

gacggcagct cgggcaccaa gtggctcgcg ttcgcgtcga ccggctgggt ccggtacgag 420

ttcgccgagc ccgtctcgtt cgtcgcgtac acgatgacct ccggcgacga cgccgccggt 480

cgcgacccga agacctggac ggtcgagggg tcgaacgacg ggtccacgtg ggccgcgctc 540

gaccgccgga cggacgagga cttcccgaac cgccagcaga cgcgcacgtt cgagctcgag 600

gcgcccaccg cggcgtacac gtacctgcgc ctcaacgtca cggcgaactc gggcgactcc 660

atcgtccagc tcgccgggtg ggacctctcg gccgacctga gcgccggccc gtccgcggcc 720

cccatgacga cgaaggtcgg caccgggccg cgcgtcagct tcaccaacaa ggcgggcgtc 780

gggttctccg gcctgcactc gctccggtac gacggctcgc acctcgccga cggcgagacg 840

tacgcgacga acgtgctcta cgacgacgtg gacgtcgtcg tcggcgagga cacgcgcctg 900

agctacacga tcttccccga gctgctcgac gatctgcagt acccgtcgac gtacgcggcg 960

gtggacgtcc tgttcaccga cgggacctac ctgtccgacc tcggcgcgcg cgacgcgcac 1020

gagacggtcg cgaccgcgca ggcgcagggc gagggcaaga tcctctacgc cgaccagtgg 1080

aactcggtgc gggtcgacct cggcgacgtc gccgagggca agaccgtgga ccaggtgctg 1140

ctcgggtacg acaacccggg cggtcacgcc gggacgaagt tcgcgggctg gctcgacgac 1200

gtcgagatca cggcggagcc ggccacgatc gacgggtcga gcctcgccaa ctacgtggac 1260

acgcgccgcg gcacgctcgc gtcgggcagc ttctcgcgcg ggaacaacat ccccgcgacg 1320

gcgacgccga acgggttcaa cttctggacg ccgtacacga acgcctcctc gcagagctgg 1380

ctgtacgagt accacaaggc caacaacgcc aacaacaagc ccgtcctcca gggcttcggg 1440

atctcgcacg agccgagccc gtggatgggc gaccgcaacc agctgacgtt cctcccgtcg 1500

acggcgtcgg ggacgcccga cgccacgctc tcgacgcgcg gcctcgagtt cgaccacgcg 1560

gacgagacgg cgcggccgga ctactacggg gtcacgttca ccaacgggtc cgcgatcgag 1620

gcgacgccca ccgaccacgg cgcggtgctc cgcttcagct accccggagc caagggccac 1680

gtgctcgtgg acaaggtgga cggctcctcc aagctcacgt acgaccaggc cacgggcacg 1740

atctccggct gggtcgagaa cggctcgggc ctgtccgtgg gccgcacgcg catgttcgtc 1800

gccggcacct tcgaccgtag tccgacggcg gtcgggacgg cggcgggcaa ccgtgcggac 1860

gcgcgcttcg cgacgttcga gacgtcgtcc gacaagacgg tcgagctgcg cgtcgcgacg 1920

tcgttcatca gcctcgacca ggcgcgcaag aacctcgacc tggaggtgac gggcaagacc 1980

ttcacggagg tcaaggccgc cgccgcgcag gcgtggaacg accgcctggg ggtcatcgag 2040

gtcgagggcg cgagcgagga ccagctcgtc acgctgtact cgaacctcta ccgcctcaac 2100

ctgtacccga actcgcagtt cgagaacacg ggcacggcgc aggagccggt gtacaggtac 2160

gcgagcccgg tctccgcgac cacgggctcc gcgacggaca cgcagaccaa cgcgaagatc 2220

gtcgacggca agatctacgt gaacaacggg ttctgggaca cgtaccgcac ggcctggccg 2280

gcgtactcgc tcctctaccc ggagctcgcg gccgagctgg tcgacgggtt cgtccagcag 2340

taccgcgacg gcggctggat cgcgcgctgg tcctcgccgg gctacgccga cctcatgacg 2400

ggcacgagct ccgacgtggc gttcgccgac gcgtacctca agggctcgct ccccacgggc 2460

acggcgctcg aggcgtacga cgccgcgctg cgcaacgcga ccgtcgcgcc gccgagcaac 2520

gccgtgggcc gcaagggcct gcagacctcg ccgttcctcg ggttcacgcc ggagtccacg 2580

cacgagtccg tgtcgtgggg cctggagggc ctggtcaacg acttcggcat cggcaacatg 2640

gccgccgccc tcgcggagga cccggcgacg ccggaggagc gccgcgagac gctgcgcgag 2700

gagtccgcgt acttcctcga gcgggccacg cactacgtcg agctgttcga ccccgaggtc 2760

gacttcttcg tgccgcggca cgaggacggc acgtgggccg tcgaccccga gacgtacgac 2820

ccggaggcct ggggcggcgg gtacaccgag acgaacggct ggaacttcgc gttccacgcc 2880

ccgcaggacg gccagggcct cgccaacctc tacggcggca agcagggcct cgaggacaag 2940

ctcgacgagt tcttctccac gccggagaag ggcgccggca acggcggcat ccacgagcag 3000

cgcgaggcgc gcgacgtccg catgggccag tggggcatga gcaaccaggt gtcgcaccac 3060

atcccgtggc tctacgacgc cgcgggcgcg ccgtcgaagg cgcaggagaa ggtccgcgag 3120

gtcacccgcc gcctgttcgt gggcagcgag atcggccagg gctacccggg cgacgaggac 3180

aacggcgaga tgtcgtcgtg gtggatcttc gcctcgctcg gcttctaccc gctccaggtc 3240

ggctcggacc agtacgcggt cggttcgccg ctgttcgaca aggcgaccgt gcacctgccg 3300

gacggcgacc tcgtcgtcaa cgccgagaac aactcggtcg acaacgtcta cgtgcagtcc 3360

ctcgcggtgg acggcgaggc ccgcacctcg acgtcactct cccaggcgga cctctcgggc 3420

ggcacgactc tggacttcgt catgggtccg gagccgtcgg actggggcac gggcgaggac 3480

gacgcgccgc cgtcgctcac cgagggcgac gagcccccga cgccggtgca ggacgcgacg 3540

accgcgggcc tcggcaccac caccgtcgcc gacggcgacg ccaccacgag cgccgcggcg 3600

ctcacggaca acacgtccgg gacgcgcacg acgttcgcca ccacgacgcc gtcgatcacg 3660

tgggcgggca acggcatccg cccgaccgtc gggtcgtaca cgctgacctc cggggcgagc 3720

gggacggcgt caccgtccgc atggactctc gagggttccg acgacggcga gacgtggacg 3780

acgctcgacg agcggtccgg cgagcagttc cgctgggccc tgcagacgcg gccgttcacg 3840

gtcgcggagc cgacggcgtt cgcgcggtac cgggtcacgg tcaccgcgac gtcgggctcc 3900

ggcgcgctgt cgctcgccga ggtcgagctc ctcgccgacc cgaaggagtc gggggccgag 3960

gagctcaccc tctcggccgc gccggaccgt gacggcgtca cgggccgcga ggtctcgggc 4020

tcgttcgcga ccctcaccgg ggtcgagggc gacgtcgcgg cgctcgacgt gcaggtcgcg 4080

ttcggcgacg gctccgagcc ggtcgccggg acgctgcggg cgggcgcgtt cggcgggtac 4140

gcggtggacg ccgcgcacac gtggaccgca cccggcgtct accccgtgac cgtcacggtc 4200

tcgggcgagg ggatcgagac cgtctcggcc tcctcgtacg tcagcgtctc gctcctgcgc 4260

gagggctcgc tgctcgccgc gtacgacaac gtctgcatcg gcgacgccgg gacgacggtc 4320

ggctcgtgcg acggccaggg cgtgttcttc gaccgggcgc agctcgcggc gaagggcttc 4380

gtccagggcg agcgcgcgac ggtgccgggc acggacctcg cgttcgacgt cccggcggtc 4440

cccgccgggc agccggacaa cgccacgggc gacgggcaga ccatcgagct cgacgtcccc 4500

gcggacgcgg agcagctctc ggtgatcggc acgggcacgg agaagaacca gcaggccacc 4560

ggcacgctga ccttcgacga cggctcgacc cagccgatcg acctgagctt cggcgactgg 4620

tcgggcgcgg cccgcaaccc cgtgttcggc aacatccccg tcgcggtgac ggacagccgc 4680

ctccgcggcg gcagcccgca gaccggcacc cccgccgcgt tcttcgcgac ggcgccgatc 4740

accctccccg agggcaagcg gcccgtgagc ctcacgctcc cggaccagcc gggcgagctc 4800
tcgcgcgacg gccgcatcca cgtggtcgcg gtcgcgcacg acggcacgtt cgccgagcac 4860
cccgcgctcg aggtcacggc cgcggagggc gtgacgctcg ccgtcgggca gacctcggac 4920
gtggcgctcg cccaggtggc gggcggccgc gagggcgccg acctccgggc ggcggtcacg 4980
tggggcgacg gctccgacgt cgcggccggc gcggtgaccg acgggtcggt ctccggctcg 5040
cacgcctaca cggcggccgg gacgtacacg gcgtacgtcg tggtcgacga cggctggacc 5100
agccaggtgg tcgaggtccc cgtgaccgtg accgaggcgg agccggccct cgccgtcgac 5160

gtcacggtga gcacacgctg cctcgccggc aaggcgtacg tcgcggtccg cgccgagaac 5220
ggcgaggacg tgccgctcgc gatccggctc gtcacgccgt tcggcaccaa ggaggtcgcg 5280
gccgtcgcgc cgggcgccaa cgcctaccag tcgttcgcga cgcgggtcac ggcggtcgag 5340
gccggcaccg tcaccgtcga ggcgacgcgc ggcaccggcg acgaggaggt gacggcgtcg 5400
atccaggccg actacgccgc cgtgacctgc ggc 5433

<210> 13
<211> 1811
<212> PRT
<213> Unknown
<220><223> Cellulosimicrobium cellulans
<400> 13

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

65 70 75 80
Pro Trp Gln Ala Asn Val Gly Ser Phe Thr Ala Gly Leu Pro Gly Ser
85 90 95

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

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

210 215 220
Ala Gly Trp Asp Leu Ser Ala Asp Leu Ser Ala Gly Pro Ser Ala Ala
225 230 235 240
Pro Met Thr Thr Lys Val Gly Thr Gly Pro Arg Val Ser Phe Thr Asn
245 250 255
Lys Ala Gly Val Gly Phe Ser Gly Leu His Ser Leu Arg Tyr Asp Gly
260 265 270
Ser His Leu Ala Asp Gly Glu Thr Tyr Ala Thr Asn Val Leu Tyr Asp
275 280 285

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

Lys Ile Leu Tyr Ala Asp Gln Trp Asn Ser Val Arg Val Asp Leu Gly

355 360 365

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

370 375 380

Asn Pro Gly Gly His Ala Gly Thr Lys Phe Ala Gly Trp Leu Asp Asp

385 390 395 400

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

405 410 415

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

420 425 430

Arg Gly Asn Asn Ile Pro Ala Thr Ala Thr Pro Asn Gly Phe Asn Phe

435 440 445

Trp Thr Pro Tyr Thr Asn Ala Ser Ser Gln Ser Trp Leu Tyr Glu Tyr

450 455 460

His Lys Ala Asn Asn Ala Asn Asn Lys Pro Val Leu Gln Gly Phe Gly

465 470 475 480

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

485 490 495

Phe Leu Pro Ser Thr Ala Ser Gly Thr Pro Asp Ala Thr Leu Ser Thr

500 505 510

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

515 520 525

Tyr Gly Val Thr Phe Thr Asn Gly Ser Ala Ile Glu Ala Thr Pro Thr

530 535 540

Asp His Gly Ala Val Leu Arg Phe Ser Tyr Pro Gly Ala Lys Gly His

545 550 555 560

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

565 570 575

Ala Thr Gly Thr Ile Ser Gly Trp Val Glu Asn Gly Ser Gly Leu Ser

580 585 590

Val Gly Arg Thr Arg Met Phe Val Ala Gly Thr Phe Asp Arg Ser Pro

595 600 605

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

610 615 620

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

625 630 635 640

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

645 650 655

Thr Gly Lys Thr Phe Thr Glu Val Lys Ala Ala Ala Ala Gln Ala Trp

660 665 670

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

675 680 685

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

690 695 700

Ser Gln Phe Glu Asn Thr Gly Thr Ala Gln Glu Pro Val Tyr Arg Tyr

705 710 715 720

Ala Ser Pro Val Ser Ala Thr Thr Gly Ser Ala Thr Asp Thr Gln Thr

725 730 735

Asn Ala Lys Ile Val Asp Gly Lys Ile Tyr Val Asn Asn Gly Phe Trp

740 745 750

Asp Thr Tyr Arg Thr Ala Trp Pro Ala Tyr Ser Leu Leu Tyr Pro Glu

755 760 765

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

770 775 780

Gly Trp Ile Ala Arg Trp Ser Ser Pro Gly Tyr Ala Asp Leu Met Thr

785 790 795 800

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

805 810 815

Leu Pro Thr Gly Thr Ala Leu Glu Ala Tyr Asp Ala Ala Leu Arg Asn

820 825 830

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

835 840 845

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

850 855 860

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

930 935 940
Gly Gly Gly Tyr Thr Glu Thr Asn Gly Trp Asn Phe Ala Phe His Ala
945 950 955 960
Pro Gln Asp Gly Gln Gly Leu Ala Asn Leu Tyr Gly Gly Lys Gln Gly
965 970 975
Leu Glu Asp Lys Leu Asp Glu Phe Phe Ser Thr Pro Glu Lys Gly Ala
980 985 990
Gly Asn Gly Gly Ile His Glu Gln Arg Glu Ala Arg Asp Val Arg Met
995 1000 1005

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

1075 1080 1085
Ser Pro Leu Phe Asp Lys Ala Thr Val His Leu Pro Asp Gly Asp Leu
1090 1095 1100

Val Val Asn Ala Glu Asn Asn Ser Val Asp Asn Val Tyr Val Gln Ser
1105 1110 1115 1120
Leu Ala Val Asp Gly Glu Ala Arg Thr Ser Thr Ser Leu Ser Gln Ala
1125 1130 1135
Asp Leu Ser Gly Gly Thr Thr Leu Asp Phe Val Met Gly Pro Glu Pro
1140 1145 1150

Ser Asp Trp Gly Thr Gly Glu Asp Asp Ala Pro Pro Ser Leu Thr Glu
1155 1160 1165
Gly Asp Glu Pro Pro Thr Pro Val Gln Asp Ala Thr Thr Ala Gly Leu
1170 1175 1180
Gly Thr Thr Thr Val Ala Asp Gly Asp Ala Thr Thr Ser Ala Ala Ala
1185 1190 1195 1200
Leu Thr Asp Asn Thr Ser Gly Thr Arg Thr Thr Phe Ala Thr Thr Thr
1205 1210 1215
Pro Ser Ile Thr Trp Ala Gly Asn Gly Ile Arg Pro Thr Val Gly Ser

1220 1225 1230
Tyr Thr Leu Thr Ser Gly Ala Ser Gly Thr Ala Ser Pro Ser Ala Trp
1235 1240 1245
Thr Leu Glu Gly Ser Asp Asp Gly Glu Thr Trp Thr Thr Leu Asp Glu
1250 1255 1260
Arg Ser Gly Glu Gln Phe Arg Trp Ala Leu Gln Thr Arg Pro Phe Thr
1265 1270 1275 1280
Val Ala Glu Pro Thr Ala Phe Ala Arg Tyr Arg Val Thr Val Thr Ala
1285 1290 1295

Thr Ser Gly Ser Gly Ala Leu Ser Leu Ala Glu Val Glu Leu Leu Ala
1300 1305 1310
Asp Pro Lys Glu Ser Gly Ala Glu Glu Leu Thr Leu Ser Ala Ala Pro
1315 1320 1325
Asp Arg Asp Gly Val Thr Gly Arg Glu Val Ser Gly Ser Phe Ala Thr
1330 1335 1340
Leu Thr Gly Val Glu Gly Asp Val Ala Ala Leu Asp Val Gln Val Ala
1345 1350 1355 1360

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

1365 1370 1375

Phe Gly Gly Tyr Ala Val Asp Ala Ala His Thr Trp Thr Ala Pro Gly

1380 1385 1390

Val Tyr Pro Val Thr Val Thr Val Ser Gly Glu Gly Ile Glu Thr Val

1395 1400 1405

Ser Ala Ser Ser Tyr Val Ser Val Ser Leu Leu Arg Glu Gly Ser Leu

1410 1415 1420

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

1425 1430 1435 1440

Gly Ser Cys Asp Gly Gln Gly Val Phe Phe Asp Arg Ala Gln Leu Ala

1445 1450 1455

Ala Lys Gly Phe Val Gln Gly Glu Arg Ala Thr Val Pro Gly Thr Asp

1460 1465 1470

Leu Ala Phe Asp Val Pro Ala Val Pro Ala Gly Gln Pro Asp Asn Ala

1475 1480 1485

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

1490 1495 1500

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

1505 1510 1515 1520

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

1525 1530 1535

Phe Gly Asp Trp Ser Gly Ala Ala Arg Asn Pro Val Phe Gly Asn Ile

1540 1545 1550

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

1555 1560 1565

Gly Thr Pro Ala Ala Phe Phe Ala Thr Ala Pro Ile Thr Leu Pro Glu

1570 1575 1580

Gly Lys Arg Pro Val Ser Leu Thr Leu Pro Asp Gln Pro Gly Glu Leu

1585 1590 1595 1600

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

1605 1610 1615

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

1620 1625 1630

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

1635 1640 1645

Gly Arg Glu Gly Ala Asp Leu Arg Ala Ala Val Thr Trp Gly Asp Gly

1650 1655 1660

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

1665 1670 1675 1680

His Ala Tyr Thr Ala Ala Gly Thr Tyr Thr Ala Tyr Val Val Val Asp

1685 1690 1695

Asp Gly Trp Thr Ser Gln Val Val Glu Val Pro Val Thr Val Thr Glu

1700 1705 1710

Ala Glu Pro Ala Leu Ala Val Asp Val Thr Val Ser Thr Arg Cys Leu

1715 1720 1725

Ala Gly Lys Ala Tyr Val Ala Val Arg Ala Glu Asn Gly Glu Asp Val

1730 1735 1740

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

1745 1750 1755 1760

Ala Val Ala Pro Gly Ala Asn Ala Tyr Gln Ser Phe Ala Thr Arg Val

1765 1770 1775

Thr Ala Val Glu Ala Gly Thr Val Thr Val Glu Ala Thr Arg Gly Thr

1780 1785 1790

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

1795 1800 1805

Thr Cys Gly

1810

<210> 14

<211> 4995

<212> DNA

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 14

gcgctcgccg tcgtcggcct cgcgcccgcg accgccgcga gcgccgcccc cgagccgccg 60

tcggccgact acgcgtccct ggtcgacgtc ttcgtcggca ccgagggcga cttcggcaac 120

gacatgcccg ccgcgcaggc gccgaacggc ctcgcgaagg tcaacccgcg cacgaccccg 180

ggccgcaaca acaccgggta cgactacgcg cagtcgaaga tctcgggctt cacgcacacc 240

aacctcgacg gggtcggggg ctccggcggc ggtggtgacc tcctcgtggt gccgacgtcc 300

gggtcgtaca cggcgcgccc cggcacgggc acgtacgcgc acccgttctc gcacgacgac 360

gaggacgccg gaccgggctt ctactccgtc gggctcggca acgtcgcggg cacggacggc 420

gcgatcaccg gcgcgccggg cacgatcgag gccgaggtcg cggcggccac gcgctcgggc 480

gtgcaccgct acgcgttccc cgcgggctcg acgccgagcc tcgtcgtgga cctcgagacg 540

aacaacacga gccgccggtc gtcctcggtg caggtcgaga cgcgcgcgga cggcaccgtg 600

gagctgtccg gacaggtcac gggctacttc tacaacgcgg cctacacgct gtactacacc 660

gcgcgcacgc tccagcccgc gacggtgcag acgtggggcg acgacgaccg gctcgtcgac 720

gccacggccc aggacggcgt cgacaccggc gcgatcctca cgttcgaccc ggcggacgcc 780

ggggagatcg ggctccaggt caccctgtcg ccggtgagcg tcgagcaggc gcggatcgac 840

cagcaggtcg agctcggcga cctgtcgttc gacgcgatcc gtgaccgcac ccgcgcggag 900

tggaacgcga cgctcgggcg ggtcgcgatc gacgcctcga cggcgacgga cccgacgggc 960

gagctccagc ggctcttcta cacgcacctc taccgcatgt tcgcgatgcc gatgaacgcg 1020

acgagcacct cgggcacgta ccgcggcgtc gacggggcgg tgcacgccgc gcagggcttc 1080

acgtactacg actcgtgggc cacgtgggac gacttccgca agttctccgt catcgcgtac 1140

atcgacccgg cgctgtaccg ggacatggtg cagtcgctgg tctacctgtt cgcggacgcc 1200

gaggcgacgg gcaccggcgg cggcctcggc gggttcgtgc actcggtccc gacggtgcgc 1260

tgggagcggt cgtcggtcgt ggtcgcggac gcgatcgcca agggcttcga cgggttcgac 1320

cgcctcgacg aggcgtaccc ggcgctccag cggctcgtcg ggcagtacag cgcggacgag 1380

ctccggcgcg gctacgtggc gggcaacccc ggcgcgtccg tgcagcgcgg ctacgaccag 1440

tacggcctgt ccgtgatcgc ggacgagctc ggcctgaccg aggaggccga gacgctgcgc 1500

gagcaggcgt cgtggccgat cgagaagctc accaagccgg gcgcgtggac cgccgccgac 1560

ggcacgcagg tcggcctcct caccccgcgc gccgcggacg ggtcgtggca gagcgccgac 1620

cacgcgaagt tcgaggccgc cggcctctac cagggcacgc tctggcagta ccactggtac 1680

gacgcgtacg acatggacgc gctcgtcgag gcgatgggcg gccacgaggc ggcgcgcctc 1740

ggcatgcgcc acatgttcgg tgagcacgcg ccggacgacg gcaaggccat gctccactcg 1800

aacgccaacg agatcgacct ccaggcgccg tacctcttca actacacggg cgagccgagc 1860

ctcacgcaga agtgggcgcg cgcgatctac acgaaggaga cctggaaccg gtacatcgcg 1920

accggctcct ccagcgccgt gccgagcggc ggcggcgagt tcacgccgcc cttgaagacg 1980

aaggtgtacc ggctcgaccc ccgcgggatg ctccccacga tggacaacga cgcgggcacg 2040

atgtcgacga tgttcgtcgc cgcggccgtc gggctgttcc cggtgaccgc gggctcgtcc 2100

cagttccagg tcgggtcgcc gttcttcgac tcgacgacca tcacctacga cgacggcagc 2160

gccttcacgg tcacggccga cggcgtctcc gaggacgcgt tctacgtcca gtccgcgacg 2220

ctcgacggcg cgacgttcgg caacacgtgg gtcgactacg ccaccgtggt cgggggagcc 2280

gacctcgcgt tccgcatggg cgagcagccg agcgactggg gcacggacac cgcgcccgcg 2340

ttctcgatga gtaccgcgac cgacgagccg gccgagggac cgcgcgtcag cgccgaaccg 2400

accaccgtgc agaccggcga cggcggcgcg ctcgacgcga ccgtgacgct cacgctcgac 2460

ggcgcccgcc tcgccgcgcc cgccggcacg gacctcgtca cgagcggggc ggcgagcgtc 2520

gtcgggctgc ccgacggcgt cacggcggcc gtgacggtcg cgtcgccgac cgcgctgacc 2580

gtctccctga cggggacggc gtccgccgac gcgcgcttct tcgtgcacct gcgcgacgcc 2640

gcgctcgccg acggcgtcgc cgcggcgtcg ctccagggac agggcgtctc ggtgcgctcg 2700

cccctgcggc tgtccgtggc gtccgccgag cgcgacgcgc tcgccgcgct cgtcgacgac 2760

gccgtgctcg tgcggcacgg gaactactcc tcggtgacgt tcgaccggtt ctccaccgcg 2820

ctgacgaagg cgcaggaggc cctcggtgac gaggccgcga cgagcatcgc gctgcggttc 2880

gcggccgacc ggctcggtgc ggcggccgac gcgctcgacc tcacgggcgg cgggtaccgc 2940

acgctcgagg ccgagcagtc cgaggcgtgg tcgggcgggg agctgaagaa cgaggcgaac 3000

agctcgtccg gcaacctcgg cggcgtgcgc tccgggtcgt gggtgcagta ccgcgacatg 3060

accttcgaga ccgccgccgg ggacaccccg ccgcgcttcc tcacggtccg gtacgacacg 3120

agcttcgccc cgacggacac gccgagcacc gtgcgcgtgc acgcgggcga cgtgagcggc 3180

cctgtggtcg cgaccgtcga cctgaagggc acgagcggct ggggcaagta caccgaggtc 3240

acggcggagc tcggcgacgt gcaggcgctc gtcgacgcgc aggtcgtcac gttcgagctg 3300

ctcgcgccgt ccgggcggag ctgggtcggc aacttcgact ggttccggtt cagcgccgag 3360

gacccggctg ccccaggtca gccgggcgag tccccgacgg tgacgatcga ggccgaggac 3420

tggaccgcga gctccggtcg cgggctcaag aaggagtcct cgacgtggac gagcggtccg 3480

gtgacgaacg tcggcggcac cgcggacggc gactggatcg cctacggcga ggtcgacctg 3540

ggtgagctcc cgctcggcga gctgtcggtc cactacgtgc acaactccaa ccggtccggg 3600

aacaactccg cgctgtcggt gtacctcgac gcgttcgacc cggcgaaccc gggcgagccg 3660
ttcgtcaccg tgccgctgcc gacgaccggg tcgagctgga ccgcggacgg gaccgcgacc 3720
gtcgtcctgc ccgagacggt gcaggggacg cacgaggtgt tcgtgcgcct gtcgaccgag 3780
ccgtacgccg accacccgta cgtcgcgaac ctcgacagcc tgacgttcgc gccgggcggc 3840
ccgacgtcgg tcgtcgtcga gtccgaggcc tggacgtcga actccggccg cgggctgaag 3900
aacgagagct cgacgtggac gagcggtccg gtgacgaacg tcggcggcac cgcggacggc 3960
gactggctcg cctacggcga gatcgacctc ggctccgccg cgctcgacca gctctcggtc 4020

cactacgtgc acaactccaa ccggtccggg cggaactccg cgctgtcggt gtacctcgac 4080
gcgttcgacc cggcgaaccc gggcgagccg ttcgtcaccg tcccgctggc caacaccggg 4140
tcgagctgga cgacggacgg gaccgccgtc gtcgacctgc cgagcacggt gcgcggcaag 4200
caccaggtgt gggtgcgcct gtccaccgag gcgtacgccg accacccgta cgtcgccaac 4260
ctcgacagca tgcgcttctt caccgacgcg tacgacgtcg aggtcccgcc gaccgacacc 4320
gcggcgctcg cggcggtggt cgacgcggcc gggacgcccg aggcggagat cgcgcggtac 4380
ggccggatcg acgcgcgcgt cttcacacgc gagctcgcgg cggcacggtc cgtgctcgcc 4440

gacgccggcg ccacccaggc gcaggccgac gagcgggcgc ggcgcctcgg cctggcgacc 4500
gaccagctcg tgcccgccga gcgccgtcgg ctcgagaacc tcgtggcgag cgccgaggcc 4560
ctgaccgacg aggggtacag ccccgagtcc tggcaggcct tccgcacggc tctcgccgcg 4620
gcgaccggga cgctcgacga cgcggcggcg tccgacgagg cgctgcacga cgcgcggctc 4680
gcgctccagg gcgccgtcga cgccctggag gagccggccg acgtcgtgct cgtcgaggtc 4740
gaggtcagcc cgcgctgcct cgccggcaag ccctacgtcg cggtccgcgc ggtgaacgtc 4800
tccgacgcgg ccgtcgacgt cgagctggcg tcgtcactgg gcacgaggtc gttcgtcggc 4860

gtcgcgccgg gggcgagcgc gtaccagtcg ttcgccgcgc ggtccgcgac gggcgacctg 4920
gacgtcaccg tcacggcgac gggggcggac ggcacccaga cggtcgagca ggtcgtcacc 4980
gtcccgtcct gctcc 4995

<210> 15

<211> 1665

<212> PRT

<213> Unknown

<220><223> Cellulosimicrobium cellulans

<400> 15

Ala Leu Ala Val Val Gly Leu Ala Pro Ala Thr Ala Ala Ser Ala Ala
1 5 10 15
Pro Glu Pro Pro Ser Ala Asp Tyr Ala Ser Leu Val Asp Val Phe Val

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

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

165 170 175
Asp Leu Glu Thr Asn Asn Thr Ser Arg Arg Ser Ser Ser Val Gln Val
180 185 190
Glu Thr Arg Ala Asp Gly Thr Val Glu Leu Ser Gly Gln Val Thr Gly
195 200 205
Tyr Phe Tyr Asn Ala Ala Tyr Thr Leu Tyr Tyr Thr Ala Arg Thr Leu
210 215 220
Gln Pro Ala Thr Val Gln Thr Trp Gly Asp Asp Asp Arg Leu Val Asp
225 230 235 240

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

245 250 255

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

260 265 270

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

275 280 285

Ser Phe Asp Ala Ile Arg Asp Arg Thr Arg Ala Glu Trp Asn Ala Thr

290 295 300

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

305 310 315 320

Glu Leu Gln Arg Leu Phe Tyr Thr His Leu Tyr Arg Met Phe Ala Met

325 330 335

Pro Met Asn Ala Thr Ser Thr Ser Gly Thr Tyr Arg Gly Val Asp Gly

340 345 350

Ala Val His Ala Ala Gln Gly Phe Thr Tyr Tyr Asp Ser Trp Ala Thr

355 360 365

Trp Asp Asp Phe Arg Lys Phe Ser Val Ile Ala Tyr Ile Asp Pro Ala

370 375 380

Leu Tyr Arg Asp Met Val Gln Ser Leu Val Tyr Leu Phe Ala Asp Ala

385 390 395 400

Glu Ala Thr Gly Thr Gly Gly Gly Leu Gly Gly Phe Val His Ser Val

405 410 415

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

420 425 430

Ala Lys Gly Phe Asp Gly Phe Asp Arg Leu Asp Glu Ala Tyr Pro Ala

435 440 445

Leu Gln Arg Leu Val Gly Gln Tyr Ser Ala Asp Glu Leu Arg Arg Gly

450 455 460

Tyr Val Ala Gly Asn Pro Gly Ala Ser Val Gln Arg Gly Tyr Asp Gln

465 470 475 480

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

485 490 495

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

500 505 510

Pro Gly Ala Trp Thr Ala Ala Asp Gly Thr Gln Val Gly Leu Leu Thr

515 520 525

Pro Arg Ala Ala Asp Gly Ser Trp Gln Ser Ala Asp His Ala Lys Phe

530 535 540

Glu Ala Ala Gly Leu Tyr Gln Gly Thr Leu Trp Gln Tyr His Trp Tyr

545 550 555 560

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

565 570 575

Ala Ala Arg Leu Gly Met Arg His Met Phe Gly Glu His Ala Pro Asp

580 585 590

Asp Gly Lys Ala Met Leu His Ser Asn Ala Asn Glu Ile Asp Leu Gln

595 600 605

Ala Pro Tyr Leu Phe Asn Tyr Thr Gly Glu Pro Ser Leu Thr Gln Lys

610 615 620

Trp Ala Arg Ala Ile Tyr Thr Lys Glu Thr Trp Asn Arg Tyr Ile Ala

625 630 635 640

Thr Gly Ser Ser Ser Ala Val Pro Ser Gly Gly Gly Glu Phe Thr Pro

645 650 655

Pro Leu Lys Thr Lys Val Tyr Arg Leu Asp Pro Arg Gly Met Leu Pro

660 665 670

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

675 680 685

Ala Val Gly Leu Phe Pro Val Thr Ala Gly Ser Ser Gln Phe Gln Val

690 695 700

Gly Ser Pro Phe Phe Asp Ser Thr Thr Ile Thr Tyr Asp Asp Gly Ser

705 710 715 720

Ala Phe Thr Val Thr Ala Asp Gly Val Ser Glu Asp Ala Phe Tyr Val

725 730 735

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

740 745 750

Tyr Ala Thr Val Val Gly Gly Ala Asp Leu Ala Phe Arg Met Gly Glu
755 760 765

Gln Pro Ser Asp Trp Gly Thr Asp Thr Ala Pro Ala Phe Ser Met Ser
770 775 780

Thr Ala Thr Asp Glu Pro Ala Glu Gly Pro Arg Val Ser Ala Glu Pro
785 790 795 800

Thr Thr Val Gln Thr Gly Asp Gly Gly Ala Leu Asp Ala Thr Val Thr
805 810 815

Leu Thr Leu Asp Gly Ala Arg Leu Ala Ala Pro Ala Gly Thr Asp Leu
820 825 830

Val Thr Ser Gly Ala Ala Ser Val Val Gly Leu Pro Asp Gly Val Thr
835 840 845

Ala Ala Val Thr Val Ala Ser Pro Thr Ala Leu Thr Val Ser Leu Thr
850 855 860

Gly Thr Ala Ser Ala Asp Ala Arg Phe Phe Val His Leu Arg Asp Ala
865 870 875 880

Ala Leu Ala Asp Gly Val Ala Ala Ala Ser Leu Gln Gly Gln Gly Val
885 890 895

Ser Val Arg Ser Pro Leu Arg Leu Ser Val Ala Ser Ala Glu Arg Asp
900 905 910

Ala Leu Ala Ala Leu Val Asp Asp Ala Val Leu Val Arg His Gly Asn
915 920 925

Tyr Ser Ser Val Thr Phe Asp Arg Phe Ser Thr Ala Leu Thr Lys Ala
930 935 940

Gln Glu Ala Leu Gly Asp Glu Ala Ala Thr Ser Ile Ala Leu Arg Phe
945 950 955 960

Ala Ala Asp Arg Leu Gly Ala Ala Ala Asp Ala Leu Asp Leu Thr Gly
965 970 975

Gly Gly Tyr Arg Thr Leu Glu Ala Glu Gln Ser Glu Ala Trp Ser Gly
980 985 990

Gly Glu Leu Lys Asn Glu Ala Asn Ser Ser Ser Gly Asn Leu Gly Gly
995 1000 1005

Val Arg Ser Gly Ser Trp Val Gln Tyr Arg Asp Met Thr Phe Glu Thr

1010 1015 1020

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

1025 1030 1035 1040

Ser Phe Ala Pro Thr Asp Thr Pro Ser Thr Val Arg Val His Ala Gly

1045 1050 1055

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

1060 1065 1070

Gly Trp Gly Lys Tyr Thr Glu Val Thr Ala Glu Leu Gly Asp Val Gln

1075 1080 1085

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

1090 1095 1100

Gly Arg Ser Trp Val Gly Asn Phe Asp Trp Phe Arg Phe Ser Ala Glu

1105 1110 1115 1120

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

1125 1130 1135

Glu Ala Glu Asp Trp Thr Ala Ser Ser Gly Arg Gly Leu Lys Lys Glu

1140 1145 1150

Ser Ser Thr Trp Thr Ser Gly Pro Val Thr Asn Val Gly Gly Thr Ala

1155 1160 1165

Asp Gly Asp Trp Ile Ala Tyr Gly Glu Val Asp Leu Gly Glu Leu Pro

1170 1175 1180

Leu Gly Glu Leu Ser Val His Tyr Val His Asn Ser Asn Arg Ser Gly

1185 1190 1195 1200

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

1205 1210 1215

Pro Gly Glu Pro Phe Val Thr Val Pro Leu Pro Thr Thr Gly Ser Ser

1220 1225 1230

Trp Thr Ala Asp Gly Thr Ala Thr Val Val Leu Pro Glu Thr Val Gln

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275 1280

Pro Thr Ser Val Val Val Glu Ser Glu Ala Trp Thr Ser Asn Ser Gly

1285 1290 1295

Arg Gly Leu Lys Asn Glu Ser Ser Thr Trp Thr Ser Gly Pro Val Thr

1300 1305 1310

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

1315 1320 1325

Asp Leu Gly Ser Ala Ala Leu Asp Gln Leu Ser Val His Tyr Val His

1330 1335 1340

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

1345 1350 1355 1360

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

1365 1370 1375

Ala Asn Thr Gly Ser Ser Trp Thr Thr Asp Gly Thr Ala Val Val Asp

1380 1385 1390

Leu Pro Ser Thr Val Arg Gly Lys His Gln Val Trp Val Arg Leu Ser

1395 1400 1405

Thr Glu Ala Tyr Ala Asp His Pro Tyr Val Ala Asn Leu Asp Ser Met

1410 1415 1420

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

1425 1430 1435 1440

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

1445 1450 1455

Ile Ala Arg Tyr Gly Arg Ile Asp Ala Arg Val Phe Thr Arg Glu Leu

1460 1465 1470

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

1475 1480 1485

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

1490 1495 1500

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

1505 1510 1515 1520

Leu Thr Asp Glu Gly Tyr Ser Pro Glu Ser Trp Gln Ala Phe Arg Thr

1525 1530 1535

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

1540 1545 1550

Glu Ala Leu His Asp Ala Arg Leu Ala Leu Gln Gly Ala Val Asp Ala

1555 1560 1565

Leu Glu Glu Pro Ala Asp Val Val Leu Val Glu Val Glu Val Ser Pro

1570 1575 1580

Arg Cys Leu Ala Gly Lys Pro Tyr Val Ala Val Arg Ala Val Asn Val

1585 1590 1595 1600

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

1605 1610 1615

Ser Phe Val Gly Val Ala Pro Gly Ala Ser Ala Tyr Gln Ser Phe Ala

1620 1625 1630

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

1635 1640 1645

Ala Asp Gly Thr Gln Thr Val Glu Gln Val Val Thr Val Pro Ser Cys

1650 1655 1660

Ser

166

<210> 16

<211> 2619

<212> DNA

<213> Artificial Sequence

<220

><223> Synthetic polynucleotide

<400> 16

gcaccggcag atgaaggcac cgttaccgca gcagccggtg atgatctgac cctggaagtt 60

aatccgtttg ttggcaccga aagcgaaggt aatgcatatc cgggtgcaac cgttccgttt 120

ggtatggttc agctgtctcc ggataatacc aatagctatg ccagcaccag ctatagcacc 180

aatgcaggtc gtgtttgggg ttttagccat cgtcatgtta atagcgcagg ttgtccggca 240

gccggtgaac tgctggttac accggatacc agcgcaacac cgcgtaccag ccgtagcttt 300

attgccatca aagatcagaa aagcaccgaa cgtgcaagcg caggttttta tgaagttacc 360

ctggcaaatg atgttcatgc agaactgacc gcaaccaccc gtgttggtgc acatcgttat 420

acctttccgg caagcaccac ctctcatctg agctttaatg ttggtcagac cctgcgtgat 480

gccggtgcaa gcagcgttac ctgggttgat gatcgtacac tggaaggttg ggttgataat 540

ggtggttttt gtggtggtac accggataaa cagcgctatt tttttagcgc aacctttgat 600

cgtccggttg ccagcagcgg tacatggggc accgatgcac gttatgttgc aggtagcacc 660

accagtgaag ttgccggtgg taataatggt gcagttgccg tttttgatac caccaccgat 720

cgtgatgttg aagttagcgt tggtgttagc tttgttagcg ttgatggtgc acgtgcaaat 780

cgtgaagccg aagcaaccga tgaaggtggt caggttgcat ttgataccgt tcgtgaagaa 840

gcacgcgacg cctggaatgc agaactgggt cgtgcagcaa ttgatgcatc tccggatcag 900

cgtcgtatct tttataccca gctgtataaa accctgctga gcccgaccat tggttctgat 960

gttgatggtc gttatcgtgg tatggatctg gaagttcatc aggcagatgg ctgggattat 1020

tatcagaact ttagcctgtg ggatacctat cgtacccagg caaccctgca tgcactgctg 1080

ctgccggaac gtgcacagga tattgttcgt agcatgtatc agcatcgtgt tgaaggtggt 1140

tggctgcctc gttggtctct gggtgcactg gaaaccaata tcatggcagg cgatccggtt 1200

accccgtggc tggcagaaaa ttttgcactg ggcaccgttc cggatgatat tgcagatgaa 1260

ctgtgggatt atctggttga aaatgcaacc accacccctc cggatgatgt tgccagcgtt 1320

ggtcgtcgta gcaccgaatt ttatgccgaa catggtcatg ttccgtttta tccggaaaac 1380

gaaggtggcc tgggtggtca gtttgaagaa tatcgtcatg gtggtagcgc aaccctggaa 1440

ctggcactgg cagatgcaag cctgggtgcc gcagcagaac gtaccggtcg tgaaggtggc 1500

caggcatttc tggataaagg tcgcaattgg cgtaatctgt ggaatccgga tgttgaactg 1560

agcggtggtt ttcagggtat ggttaatgca aaacgtccga ccggtgaatt tgttaccctg 1620

ccggaactga ccgatgttac ccgtagcggt tttcatgaag gtgttccgtg gcagtatcag 1680

tggatggttc cgcaggatgt taccggtctg caagaagtta tgggaggcga agatggtttt 1740

gtggaacgcc tggattatta ttttgatcag cctgcactgg cagcaaatcc gggtgttagc 1800

ccgagcacct gggcaaaagg tggtagcagc tattatacca ccattcgcta taatccgggt 1860

aatgaaccga ccattatgaa tgcatggctg tatggttatg ttggtcagcc gtggaaaacc 1920

aatgatgttc tggcagccaa tctgaatcgt tttccggata caccgggtgg tggtgttggt 1980

aatgatgatc tgggcaccct ggcagcatgg tatgttatgg ccagcctggg ttttgaaccg 2040

gttatgcctg gtagcggtat tctggcactg aatgcaccga aagttcaggc agcaaccctg 2100

accaccgatg ccggtgccac cctgcgtatt gatgcagccg gtgcaaatga aaaactgccg 2160

agctatgttg ccggtctgga agttgatggt gttgcacata ccgcagcatg gctggatgtt 2220

gcagcactgc aggatggtgg caccctggat tttgatctga gcggtacaag cgcaggtctg 2280

acatggggta caggtgcagc agatcgtatt ccgagcgtta gcgcagttgc accgcctgca 2340

ccggttgaag tggaagcaag cgcacgttgt ctgggtggtc gtgcatttgt tgcagttcgt 2400

gcaaccagca ccgcagatgc accggtggat gttacactga ccacaccgtt tggtgaacgt 2460

accgttcgtc atgttcagcc tggtcgtagc gcatatcaga gctttaccac ccgtaccacc 2520

tctgttgaag caggcaccgc aaccgttacc gttgttgcag cagatggcac cacctcaacc 2580

gttgatgcag catatgaagc actggcatgt ggttaataa 2619

<210> 17

<211> 3483

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 17

gcaggcaccg aagcagcaac cggttctgat gcagcagcag ttgatggtcc gctggttgat 60

tatgtgaatc cgtttattgg caccaaagat gatggtaata cctatccggg tgcagcagtt 120

ccgtttggta tggttcagct gtctccggat aatggtcata atgtgggcta tgattatgat 180

cgtaccagcg ttcgtggttt tagcctggtt catctgagcg gtgttggttg tggtctgggt 240

ggtccgctgc cgacactgcc gaccaccggt gcaattacca gcaccgatta tggtcagtat 300

gcactgggtt ttagccatga tgatgaagaa gcatctccgg gttattatcg tgttggtctg 360

caggcacctg caggaaccat tgaagcagaa ctgaccgcaa ccgaacgtac cggtgttcag 420

cgttatacct ttccggcaac cgcacaggca aatgttctgc tgaatgcagg tcaggcactg 480

aatcgtgtta ccgaatctga tgttcgtgtt gttgatgatc gtaccgttga aacccgtatt 540

accgtgcgtg gtttttgtca ggataccgaa ccgcagacca tttggacccg taccaccttt 600

gatcgtccgt ttgttgcaca tggcacctgg gatggtcagg ttgttaccgc aggcgcagat 660

gcagcaagcg gtggtgaagg tcgtcgtggt gcatatgtta cctttgatac aaccggtggt 720

gatctggatg ttgaagcagt taccgcaatg agctatgttg gtgcagatgg tgcagcagca 780

aatctggcag cagaagcagg cacctttgac gcagttcatg atgcagcacg tagcgcatgg 840

gaagaacgtc tgggtctggt tcgtgttgca cagggtgatc cggatgatct gcgtaccttt 900

tatagcagcc tgtatcgtag ctttctggca ccgaatgttg gttctgatgt ggatggtcgt 960

tatcgtggtt gggatcagga agttcacgca gcagaaccgg attttaccta ttatcagaat 1020

tatagcctgt gggataccta tcgtacccaa cagcaactgc tgtatctgct ggcaccggat 1080

gaaagcgcag atatggcact gagcctggtt cgtcagggtc agcagggtgg ttggctgcct 1140

cgttggggtt atggtacagt ggaaaccaat attatgaccg gtgatccggc aaccccgttt 1200

ctggttagcg catggcgtca gggtctgctg gcaggtcatg aagaagaagc atacgcagtc 1260

ctgcgtgaaa atgcagatgg tgttcctccg gcagatagcc cgtttaatgg tcgtgcagcc 1320

aatgttgaat atctgcgtga tggttttgtt ccgcatgaac cggcacgtag cggtaaaccg 1380

ggtgattatg atctgcagca tggtgcaagc gcaaccatgg aatatgcact ggcagatgca 1440

atgctgagca ccatggcacg tggtctgggt catgatgaag atgcagatcg ttatgcagcc 1500

cgtggtcaga gctatcgtaa tgtttttgat ccgcgtaccg gtaattttcg tgcacgtaat 1560

gccgatggtt tttttgttgg tgatgcagat ccggcacatt ctgatggttt tcatgaaggc 1620

accgcagttc agtatcagtg gctggttccg caggatgttc cgggtctgtt tgatctgatg 1680

ggtggcaccg atgcagccgt tgatcgtctg gatgcatttt ttgcctatga tgaactggtt 1740

gcagatcctc cgcatgttgc aagcgaagtt tgggttaatg gcacctatga ttattatggc 1800

tgggaaacct ataatccgaa taatgaaccg aatctgcatg caccgtatgt ttatctgtgg 1860

accggtcagc cgtggaaaac caccgatgtt gttcgtgcag caagcaccct gtttaccgat 1920

ggtccggatg gtgttaccgg taatgatgat ctgggcacca tgagcgcatg gcatgttctg 1980

agcagcattg gtgtttatcc gattgttccg ggtgccgatc tgtggggtct gaccacaccg 2040

ctgtttgatg atgttaccat taccctggac ccggaagttt ttggtcgtga tagcctgcgt 2100

ctgaccgcag atggtgtggc accggatacc cattataccc agagcgttag cctgggtggt 2160

gaaccgctgg atcgtgcatg ggttacaggt gatgaactga ccgctgcagg caccctggat 2220

gttaccgttg gcaccgaacc gagcgcatgg gcaaccgatc cggcagcatc accgggtgca 2280

gttgttccgg ctgatggcac cgttgaacgt ctgtttgttg gtgcaacacc gcgtcagccg 2340

gttctggcac cgggtggtcg taccgaagtt gcagttcagg ttgttgccca gggtgcaggc 2400

acctctagcg gcaccctgga agtgacctct gatggtgcag ttaccgccac caccgatctg 2460

gcagaatgga ccgcagaatc tgatggtctg cctgccaccg ttgaaggaac cgttaccatt 2520

gaagctccgg cagatgccga accgggtctg cataccgttc gtctggttgt tcgtgatgca 2580

gccggtacag aagcagttcg cgaagttagc gttgttgtta gcggtgaaag ctggattgca 2640

gatgcctttg ataatgtggg tattggtgat gccggtgcag caaatgcaaa tctggatggt 2700

agcggtgcct atctgctgcg tgatctgctg gccgatctgg gtgcagttca gggtctggaa 2760

ctgaccgttc cgggtactga tctgacctat accctgggtg caccgcgtgc tggtgcaccg 2820

gataatgttg cagccagcgg tgaagttctg gaagttccgg aacatctgcg tagcgcacgt 2880

catctgagcg ttgtgggcac cagcacccat ggtacacatg gtggtggtct ggttctgggt 2940

tttgccgatg gtagcagcca gaccgttgat gttcgtctga gcgattggtg taccggttct 3000

ccggaaccgg gtaatattac cgttgcaaaa gccggtgcac gtggtgatcg tgaaaatgtg 3060

cagaaaattg gctgtggtct gtatgcaacc gcaccggtgg caattccgga aggtaaagtt 3120

ctgaccagcg ttaccctgcc gtctgatgaa cgttttcatg tgtttgcaat tgcaaccgat 3180

gcaaccggtg atgttccggc accgcaggtt gaagttaccg cacaggctcg ttgtctgggt 3240

ggtaaagcat ttgttgcagt tcgtgcactg aataccggtg aacagcctgc agcaattgaa 3300

ctggcaaccc cgtatggtag caaactgttt ggtgatgttg ctccgggtgc aaatgcatat 3360

cagagctttg caacccgtgc agcagccgtt gaagccggtg aagttaccgt taccgtgacc 3420

acaccggatg gtgaaccgca gcaggttacc gcagcatatg atgcagcggc atgtagctaa 3480

taa 3483

<210> 18

<211> 3483

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 18

gcaggcaccg aagcagcaac cggttctgat gcagcagcag ttgatggtcc gctggttgat 60

tatgtgaatc cgtttattgg caccaaagat gatggtaata cctatccggg tgcagcagtt 120

ccgtttggta tggttcagct gtctccggat aatggtcata atgtgggcta tgattatgat 180

cgtaccagcg ttcgtggttt tagcctggtt catctgagcg gtgttggttg tggtctgggt 240

ggtccgctgc cgacactgcc gaccaccggt gcaattacca gcaccgatta tggtcagtat 300

gcactgggtt ttagccatga tgatgaagaa gcatctccgg gttattatcg tgttggtctg 360

caggcacctg caggaaccat tgaagcagaa ctgaccgcaa ccgaacgtac cggtgttcag 420

cgttatacct ttccggcaac cgcacaggca aatgttctgc tgaatgcagg tcaggcactg 480

aatcgtgtta ccgaatctga tgttcgtgtt gttgatgatc gtaccgttga aacccgtatt 540

accgtgcgtg gtttttgtca ggataccgaa ccgcagacca tttggacccg taccaccttt 600

gatcgtccgt ttgttgcaca tggcacctgg gatggtcagg ttgttaccgc aggcgcagat 660

gcagcaagcg gtggtgaagg tcgtcgtggt gcatatgtta cctttgatac aaccggtggt 720

gatctggatg ttgaagcagt taccgcaatg agctatgttg gtgcagatgg tgcagcagca 780

aatctggcag cagaagcagg cacctttgac gcagttcatg atgcagcacg tagcgcatgg 840

gaagaacgtc tgggtctggt tcgtgttgca cagggtgatc cggatgatct gcgtaccttt 900

tatagcagcc tgtatcgtag ctttctggca ccgaatgttg gttctgatgt ggatggtcgt 960

tatcgtggtt gggatcagga agttcacgca gcagaaccgg attttaccta ttatcagaat 1020

tatagcctgt gggataccta tcgtacccaa cagcaactgc tgtatctgct ggcaccggat 1080

gaaagcgcag atatggcact gagcctggtt cgtcagggtc agcagggtgg ttggctgcct 1140

cgttggggtt atggtacagt ggaaaccaat attatgaccg gtgatccggc aaccccgttt 1200

ctggttagcg catggcgtca gggtctgctg gcaggtcatg aagaagaagc atacgcagtc 1260

ctgcgtgaaa atgcagatgg tgttcctccg gcagatagcc cgtttaatgg tcgtgcagcc 1320

aatgttgaat atctgcgtga tggttttgtt ccgcatgaac cggcacgtag cggtaaaccg 1380

ggtgattatg atctgcagca tggtgcaagc gcaaccatgg aatatgcact ggcagatgca 1440

atgctgagca ccatggcacg tggtctgggt catgatgaag atgcagatcg ttatgcagcc 1500

cgtggtcaga gctatcgtaa tgtttttgat ccgcgtaccg gtaattttcg tgcacgtaat 1560

gccgatggtt tttttgttgg tgatgcagat ccggcacatt ctgatggttt tcatgaaggc 1620

accgcagttc agtatcagtg gctggttccg caggatgttc cgggtctgtt tgatctgatg 1680

ggtggcaccg atgcagccgt tgatcgtctg gatgcatttt ttgcctatga tgaactggtt 1740

gcagatcctc cgcatgttgc aagcgaagtt tgggttaatg gcacctatga ttattatggc 1800

tgggaaacct ataatccgaa taatgaaccg aatctgcatg caccgtatgt ttatctgtgg 1860

accggtcagc cgtggaaaac caccgatgtt gttcgtgcag caagcaccct gtttaccgat 1920

ggtccggatg gtgttaccgg taatgatgat ctgggcacca tgagcgcatg gcatgttctg 1980

agcagcattg gtgtttatcc gattgttccg ggtgccgatc tgtggggtct gaccacaccg 2040

ctgtttgatg atgttaccat taccctggac ccggaagttt ttggtcgtga tagcctgcgt 2100

ctgaccgcag atggtgtggc accggatacc cattataccc agagcgttag cctgggtggt 2160

gaaccgctgg atcgtgcatg ggttacaggt gatgaactga ccgctgcagg caccctggat 2220

gttaccgttg gcaccgaacc gagcgcatgg gcaaccgatc cggcagcatc accgggtgca 2280

gttgttccgg ctgatggcac cgttgaacgt ctgtttgttg gtgcaacacc gcgtcagccg 2340

gttctggcac cgggtggtcg taccgaagtt gcagttcagg ttgttgccca gggtgcaggc 2400

acctctagcg gcaccctgga agtgacctct gatggtgcag ttaccgccac caccgatctg 2460

gcagaatgga ccgcagaatc tgatggtctg cctgccaccg ttgaaggaac cgttaccatt 2520

gaagctccgg cagatgccga accgggtctg cataccgttc gtctggttgt tcgtgatgca 2580

gccggtacag aagcagttcg cgaagttagc gttgttgtta gcggtgaaag ctggattgca 2640

gatgcctttg ataatgtggg tattggtgat gccggtgcag caaatgcaaa tctggatggt 2700

agcggtgcct atctgctgcg tgatctgctg gccgatctgg gtgcagttca gggtctggaa 2760

ctgaccgttc cgggtactga tctgacctat accctgggtg caccgcgtgc tggtgcaccg 2820

gataatgttg cagccagcgg tgaagttctg gaagttccgg aacatctgcg tagcgcacgt 2880

catctgagcg ttgtgggcac cagcacccat ggtacacatg gtggtggtct ggttctgggt 2940

tttgccgatg gtagcagcca gaccgttgat gttcgtctga gcgattggtg taccggttct 3000

ccggaaccgg gtaatattac cgttgcaaaa gccggtgcac gtggtgatcg tgaaaatgtg 3060

cagaaaattg gctgtggtct gtatgcaacc gcaccggtgg caattccgga aggtaaagtt 3120

ctgaccagcg ttaccctgcc gtctgatgaa cgttttcatg tgtttgcaat tgcaaccgat 3180

gcaaccggtg atgttccggc accgcaggtt gaagttaccg cacaggctcg ttgtctgggt 3240

ggtaaagcat ttgttgcagt tcgtgcactg aataccggtg aacagcctgc agcaattgaa 3300

ctggcaaccc cgtatggtag caaactgttt ggtgatgttg ctccgggtgc aaatgcatat 3360

cagagctttg caacccgtgc agcagccgtt gaagccggtg aagttaccgt taccgtgacc 3420

acaccggatg gtgaaccgca gcaggttacc gcagcatatg atgcagcggc atgtagctaa 3480

taa 3483

<210> 19

<211> 5283

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 19

gcagaaccgg gtgattttag cagcagcttt gaatctggcg atccggcagc actgccgacc 60

accgttgcag aacgtgatgg tgcaccgtgg caggcaaatg ttggtagctt taccgcaggt 120

ctgcctggta gcgttctggg tcagctgaaa ggtgttaccg caagcgcaca gaatctgccg 180

aatgaaggtg cagcaaatct ggcagatggt agcagcggca ccaaatggct ggcatttgca 240

agcaccggtt gggttcgtta tgaatttgca gaaccggtta gctttgttgc atataccatg 300

accagcggtg atgatgccgc aggtcgtgat ccgaaaacct ggaccgttga aggtagcaat 360

gatggttcta cctgggcagc actggatcgt cgtaccgatg aagattttcc gaatcgtcag 420

cagacccgta cctttgaact ggaagcaccg accgcagcat atacctatct gcgtctgaat 480

gttaccgcaa atagcggtga tagcattgtt cagctggcag gttgggatct gagcgcagat 540

ctgtctgcag gtccgagcgc agcaccgatg accaccaaag ttggcaccgg tccgcgtgtt 600

agctttacca ataaagccgg tgttggtttt agcggtctgc atagcctgcg ttatgatggt 660

agccatctgg ccgatggtga aacctatgca accaatgtgc tgtatgatga tgttgatgtt 720

gtggttggtg aagatacccg tctgagctat accatttttc cggaactgct ggatgatctg 780

cagtatccga gcacctatgc agcagttgat gttctgttta ccgatggcac ctatctgagc 840

gatctgggtg cacgtgatgc acatgaaacc gttgcaaccg cacaggcaca gggtgaaggt 900

aaaattctgt atgccgatca gtggaatagc gttcgtgttg atctgggtga tgttgcagaa 960

ggtaaaaccg ttgatcaggt tctgctgggt tatgataatc cgggtggtca tgcaggcacc 1020

aaatttgcag gttggctgga tgatgttgaa attaccgcag aaccggcaac cattgatggt 1080

agctcactgg caaattatgt tgatacccgt cgtggcaccc tggcaagcgg tagctttagc 1140

cgtggtaata atattccggc aaccgcaacc ccgaatggtt ttaatttttg gaccccgtat 1200

accaatgcaa gcagccagag ctggctgtat gaatatcata aagccaataa tgcgaataat 1260

aaaccggttc tgcagggttt tggtattagc catgaaccga gcccgtggat gggtgatcgt 1320

aatcagctga cctttctgcc gagcaccgca agcggtacac cggatgcaac cctgagcacc 1380

cgtggtctgg aatttgatca tgcagatgaa accgcacgtc cggattatta tggtgtgacc 1440

tttaccaatg gtagcgcaat tgaagcaacc ccgaccgatc atggtgcagt tctgcgtttt 1500

agctatccgg gtgcaaaagg tcatgttctg gtggataaag ttgatggtag cagtaaactg 1560

acctatgatc aggcaaccgg caccattagc ggttgggttg aaaatggtag cggtctgagc 1620

gttggtcgta cccgtatgtt tgttgcaggc acctttgatc gtagcccgac cgcagttggc 1680

acagcagcag gtaatcgtgc agatgcacgt tttgcaacct ttgaaaccag cagcgataaa 1740

accgtggaac tgcgtgttgc aaccagcttt attagcctgg atcaggcacg taaaaatctg 1800

gatctggaag ttaccggtaa aacctttacc gaagttaaag cagcagcagc acaggcatgg 1860

aatgatcgtc tgggtgttat tgaagttgaa ggtgcaagcg aagatcagct ggttaccctg 1920

tatagcaatc tgtatcgcct gaatctgtat ccgaatagcc agtttgaaaa taccggcacc 1980

gcacaggaac cggtttatcg ttacgcatct ccggttagcg caaccaccgg tagcgcaacc 2040

gatacccaga ccaatgccaa aattgtggat ggcaaaattt atgtgaataa tggcttttgg 2100

gatacctatc gtaccgcatg gcctgcatat agcctgctgt atccggaact ggcagcagaa 2160

ctggttgatg gttttgttca gcagtatcgt gatggtggtt ggattgcacg ttggagcagt 2220

ccgggttatg cagatctgat gaccggtaca agctctgatg ttgcatttgc agatgcctat 2280

ctgaaaggta gcctgccgac cggtacagca ctggaagcat atgatgcagc actgcgtaat 2340

gcaaccgttg cacctccgag caatgcagtt ggtcgtaaag gtctgcagac aagcccgttt 2400

ctgggtttta caccggaaag cacccatgaa agcgttagct ggggtctgga aggtctggtt 2460

aatgattttg gcattggcaa tatggctgca gcactggcag aagatccggc aacaccggaa 2520

gaacgtcgtg aaaccctgcg tgaagaaagc gcatattttc tggaacgtgc cacccattat 2580

gttgaactgt ttgatccgga agtggatttt tttgttccgc gtcatgaaga tggtacatgg 2640

gcagttgatc cggaaaccta tgatccggaa gcatggggtg gtggttatac cgaaaccaat 2700

ggctggaatt ttgcatttca tgcaccgcag gatggtcagg gtctggcaaa tctgtatggt 2760

ggtaaacagg gtctggaaga taaactggat gaatttttta gcacaccgga aaaaggtgca 2820

ggtaatggtg gtattcatga acagcgtgaa gcacgtgatg ttcgtatggg tcagtggggt 2880

atgagcaatc aggttagcca tcatattccg tggctgtatg atgcagccgg tgctccgagc 2940

aaagcacagg aaaaagttcg cgaagttacc cgtcgtctgt ttgttggtag cgaaattggt 3000

cagggttatc cgggtgatga agataatggt gaaatgtcct cctggtggat ttttgcaagc 3060

ctgggttttt atccgctgca ggttggtagc gatcagtatg cagttggttc tccgctgttt 3120

gataaagcaa ccgttcatct gccggatggt gatctggttg ttaatgccga aaataatagc 3180

gtggataatg tgtatgttca gagcctggca gttgatggtg aagcacgtac cagcaccagc 3240

ctgagccagg cagatctgag cggtggcacc accctggatt ttgttatggg tccggaaccg 3300

agcgattggg gcaccggtga agatgatgca cctccgtcac tgaccgaagg tgatgaacct 3360

ccgacaccgg ttcaggatgc aaccaccgca ggcctgggca ccaccaccgt tgccgatggt 3420

gatgccacca cctctgcagc agccctgacc gataatacca gcggcacccg taccaccttt 3480

gcaaccacca ccccgagcat tacatgggca ggtaatggca ttcgtccgac cgttggtagc 3540

tataccctga cctctggtgc aagcggcacc gcaagcccgt ctgcatggac cctggaaggt 3600

tctgatgatg gcgaaacctg gaccacactg gatgaacgta gcggtgaaca gtttcgttgg 3660

gcactgcaga cccgtccgtt taccgttgcc gaaccgaccg catttgcacg ttatcgtgtt 3720

accgttaccg caaccagcgg ttctggtgca ctgagcctgg cagaagttga actgctggca 3780

gatccgaaag aaagcggtgc agaagaactg accctgtctg cagcaccgga tcgtgatggc 3840

gttaccggtc gtgaagttag cggttctttt gcaaccctga ccggtgttga aggtgatgtt 3900

gccgcactgg atgttcaggt tgcatttggt gatggtagcg aaccggttgc aggtacactg 3960

cgtgccggtg catttggtgg ttatgcagtt gatgcagcac atacctggac cgcaccgggt 4020

gtttatccgg ttaccgtgac cgttagcggt gaaggtattg aaaccgttag cgcaagcagc 4080

tatgttagcg ttagcctgct gcgtgaaggt tctctgctgg cagcatatga taatgtgtgc 4140

attggtgatg caggtacaac cgttggttct tgtgatggtc agggcgtttt ttttgatcgt 4200

gcacagctgg cagcaaaagg ttttgtgcag ggtgaacgtg caaccgttcc gggtacagat 4260

ctggcatttg atgttccggc agttccggct ggtcagcctg ataatgcaac cggtgatggt 4320

cagaccattg aactggatgt tccggctgat gcagaacagc tgagcgttat tggcaccggc 4380

accgaaaaaa atcagcaggc aaccggtaca ctgacctttg atgatggttc tacccagccg 4440

attgatctga gctttggtga ttggagcggt gcagcacgta atccggtgtt tggtaatatt 4500

ccggttgcag ttaccgatag ccgtctgcgt ggtggttctc cgcagaccgg tacaccggca 4560

gcatttttttg ccaccgcacc gattaccctg ccggaaggta aacgtccggt tagcctgacc 4620

ctgccggatc agcctggtga actgagccgt gatggtcgta ttcatgttgt tgcagttgca 4680

catgatggca cctttgcaga acatcctgca ctggaagtga ccgcagcaga aggtgttacc 4740

ctggcagttg gtcagacctc agatgttgca ctggcacagg ttgccggtgg tcgtgaaggt 4800

gcagatctgc gtgccgcagt tacctggggt gatggttctg atgtggcagc cggtgccgtt 4860

accgatggta gcgttagcgg tagccatgca tataccgcag caggcaccta taccgcatat 4920

gttgttgtgg atgatggttg gaccagccag gttgttgaag ttccggtgac cgttacagaa 4980

gccgaaccgg cactggccgt tgatgtcacc gttagcaccc gttgcctggc aggtaaagca 5040

tatgttgcag tgcgtgcaga aaatggtgaa gatgttccgc tggcaattcg tctggttacc 5100

ccgtttggca ccaaagaagt tgcagcagtt gctccgggag ccaatgcata tcagagcttt 5160

gcaacccgtg ttaccgcagt tgaagcaggc accgttaccg ttgaagccac ccgtggcacc 5220

ggtgatgaag aagttaccgc cagcattcag gcagattatg cagccgttac ctgcggttaa 5280

taa 5283

<210> 20

<211> 4881

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 20

gatatgcctg cagcacaggc accgaatggt ctggcaaaag ttaatccgcg taccacaccg 60

ggtcgtaata ataccggtta tgattatgcc cagagcaaaa ttagcggttt tacccatacc 120

aatctggatg gtgttggtgg tagcggtggt ggtggtgatc tgctggttgt tccgaccagc 180

ggtagctata ccgcacgtcc gggtacaggc acctatgcac atccgtttag ccatgatgat 240

gaagatgcag gtccgggttt ttatagcgtt ggtctgggta atgttgcagg caccgatggt 300

gcaattaccg gtgctccggg tacaattgaa gcagaagttg cagcagcaac ccgtagcggt 360

gttcatcgtt atgcatttcc ggcaggtagc accccgagcc tggttgttga tctggaaacc 420

aataatacca gccgtcgtag cagcagcgtt caggttgaaa cccgtgcaga tggcaccgtt 480

gaactgagcg gtcaggttac cggctatttt tataatgcag cctataccct gtattatacc 540

gcacgcaccc tgcagcctgc aaccgttcag acctggggtg atgatgatcg tctggttgat 600

gcaaccgcac aggatggtgt tgataccggt gcaattctga cctttgatcc ggcagatgcc 660

ggtgaaattg gtctgcaggt taccctgtct ccggttagcg ttgaacaggc acgtattgat 720

cagcaggttg aactgggtga tctgagcttt gatgcaattc gtgatcgtac ccgtgcagaa 780

tggaatgcaa ccctgggtcg tgttgcaatt gatgcaagca ccgcaaccga tccgaccggt 840

gaactgcagc gtctgtttta tacccatctg tatcgcatgt ttgcaatgcc gatgaatgca 900

accagcacca gcggcaccta tcgtggtgtt gatggtgcag ttcatgcagc acagggcttt 960

acctattatg atagctgggc aacctgggat gattttcgca aatttagcgt gattgcctat 1020

attgatccgg cactgtatcg tgatatggtt cagagcctgg tttacctgtt tgcagatgca 1080

gaagcaaccg gtacaggcgg tggtctgggt ggttttgttc atagcgttcc gaccgttcgt 1140

tgggaacgta gcagcgttgt tgttgcagat gcaattgcca aaggctttga tggttttgat 1200

cgtctggatg aagcatatcc ggcactgcag cgcctggttg gtcagtatag cgcagatgaa 1260

ctgcgtcgtg gttatgttgc aggtaatccg ggtgcaagcg ttcagcgtgg ttatgatcag 1320

tatggtctga gcgttattgc cgatgaactg ggtctgaccg aagaagcaga aaccctgcgc 1380

gaacaggcaa gctggccgat tgaaaaactg accaaaccgg gtgcatggac cgcagcagat 1440

ggtacacagg ttggtctgct gacaccgcgt gcagccgatg gtagctggca gagcgcagat 1500

catgccaaat ttgaagcagc aggtctgtat cagggcaccc tgtggcagta tcattggtat 1560

gatgcctatg atatggatgc actggttgaa gcaatgggtg gtcatgaagc agcccgtctg 1620

ggtatgcgtc atatgtttgg tgaacatgca ccggatgatg gtaaagcaat gctgcatagc 1680

aatgccaatg aaattgatct gcaggcaccg tacctgttta attataccgg tgaaccgagc 1740

ctgacccaga aatgggcacg tgcaatttat accaaagaaa cctggaatcg ctatattgca 1800

accggtagca gctctgcagt tccgtcaggt ggtggtgaat ttacacctcc gctgaaaacc 1860

aaagtttatc gtctggaccc tcgtggtatg ctgccgacca tggataatga tgcaggtaca 1920

atgagcacca tgtttgttgc agcagccgtt ggtctgtttc cggttaccgc aggtagcagc 1980

cagtttcagg ttggtagccc gttttttgat agcaccacca ttacctatga tgatggtagc 2040

gcatttaccg ttaccgcaga tggtgttagc gaagatgcct tttatgttca gagcgcaacc 2100

ctggatggtg caacctttgg taatacctgg gttgattatg caaccgttgt tggtggtgca 2160

gatctggcat ttcgtatggg tgaacagccg agcgattggg gcaccgatac cgcaccggca 2220

tttagcatga gcaccgccac cgatgaaccg gcagaaggtc ctcgcgttag cgcagaaccg 2280

accaccgtgc agaccggtga tggtggtgca ctggatgcaa ccgttaccct gacactggat 2340

ggcgcacgtc tggcagcacc ggcaggtaca gatctggtta ccagcggtgc agcaagcgtt 2400

gttggtctgc cggatggtgt taccgcagca gttaccgttg caagcccgac cgcactgacc 2460

gttagcctga ccggcaccgc atcagcagat gcacgttttt ttgtgcatct gcgtgatgca 2520

gcactggccg atggtgttgc agccgcaagc ctgcagggtc agggtgttag cgttcgttct 2580

ccgctgcgtc tgagcgttgc aagcgcagaa cgtgatgcac tggcagcact ggttgatgat 2640

gccgttctgg ttcgtcatgg taattatagc agcgttacct ttgatcgttt tagcaccgct 2700

ctgacaaaag cacaggaagc actgggcgac gaagcagcaa ccagcattgc actgcgtttt 2760

gcagcagatc gtctgggtgc agcagcagat gcactggatc tgaccggtgg tggttatcgt 2820

accctggaag cagaacagag cgaagcatgg tctggtggtg aactgaaaaa tgaagccaat 2880

agcagcagcg gtaatctggg tggtgttcgt agcggtagct gggttcagta tcgcgatatg 2940

acctttgaaa ccgcagccgg tgatacacct ccgcgttttc tgaccgttcg ttatgatacc 3000

agctttgcac cgaccgatac cccgagcacc gttcgtgttc atgccggtga tgtttctggt 3060

ccggttgttg caaccgttga tctgaaaggc accagcggtt ggggtaaata taccgaagtt 3120

accgcagaac tgggtgatgt tcaggccctg gttgatgccc aggttgttac ctttgaactg 3180

ctggcaccga gcggtcgtag ctgggttggt aattttgatt ggtttcgctt tagcgcagaa 3240

gatccggcag caccgggtca gcctggtgaa agcccgaccg ttaccattga agccgaagat 3300

tggaccgcaa gcagcggtcg tggtctgaaa aaagaaagca gcacctggac cagcggtccg 3360

gtgaccaatg ttggtggtac agcagatggt gattggattg cctatggtga agttgatctg 3420

ggtgaactgc cgctgggcga actgagcgtt cattatgtgc ataatagcaa tcgcagcggt 3480

aataatagcg cactgagcgt ttatctggat gcatttgatc cggctaatcc gggtgaaccg 3540

tttgttaccg ttccgctgcc gaccaccggt agcagttgga ccgcagatgg cacagccacc 3600

gttgttctgc cggaaaccgt gcagggcacc catgaagttt ttgttcgtct gagcaccgaa 3660

ccgtatgcag atcatccgta tgttgcaaat ctggatagcc tgacctttgc accgggtggt 3720

ccgaccagcg ttgtggttga aagcgaagcc tggaccagca attctggtcg tggcctgaaa 3780
aatgaatctt ctacctggac ctctggtccg gttacaaatg tgggtggcac cgctgatggc 3840
gattggctgg catatggcga aattgatctg ggcagcgcag cactggatca gctgtctgtg 3900
cattatgttc ataattctaa tcgctctggt cgtaattctg cactgtctgt gtatctggat 3960
gcctttgatc cggcaaatcc gggtgaaccg tttgtgacag tgccgctggc aaataccggt 4020
agctcttgga ccaccgatgg tactgcagtt gtggatctgc cgtctaccgt tcgtggtaaa 4080
catcaggttt gggttcgtct gtctaccgaa gcatatgccg atcatccgta tgtggccaat 4140

ctggattcta tgcgcttttt taccgatgca tatgatgttg aagttcctcc gaccgataca 4200
gcagcactgg cagccgttgt tgatgcagca ggtacaccgg aagcagaaat tgcacgttat 4260
ggtcgtattg atgcccgtgt ttttacccgt gaactggcag cagcacgtag cgttctggcc 4320
gatgccggtg caacacaggc acaggcagat gaacgtgctc gtcgtctggg tctggcaacc 4380
gatcagctgg ttccggcaga acgtcgtcgt ctggaaaatc tggttgccag cgcagaagca 4440
ctgaccgacg aaggttattc tccggaaagc tggcaggcat ttcgtaccgc actggctgct 4500
gcaaccggca ccctggatga tgcagcagca tctgatgaag cactgcatga tgcacgtctg 4560

gcgctgcagg gtgcagttga tgcactggaa gaaccggcag atgttgttct ggttgaagtt 4620
gaagtttctc cgcgttgtct ggcaggtaaa ccgtatgttg ccgttcgtgc agttaatgtt 4680
tctgatgcag ccgttgatgt tgaactggca agctctctgg gcacccgtag ctttgttggt 4740
gtggcaccgg gtgcgagcgc atatcagagc tttgcagccc gtagcgcaac cggtgatctg 4800
gatgttaccg tgaccgcaac cggtgcagat ggtactcaga ccgttgaaca ggttgtgacc 4860
gttccgagct gtagctaata a 4881

<210> 21
<211> 9
<212> PRT

<213> Artificial Sequence
<220><223> Synthetic polypeptide
<400> 21
Val Gly Pro Gly Ser Asp Glu Val Asp
1 5

<210> 22
<211> 1308

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 22

atgaagcttt ccaccatcct cttcacagcc tgcgctaccc tggccctgga caacggcctg 60

gcccgaaccc ccaccatggg ctggctgcac tgggagcgat tcatgtgtaa cctggactgt 120

caggaagagc ccgactcttg tatctctgag aagctgttca tggaaatggc cgagctgatg 180

gtgtctgagg gctggaagga cgccggctac gagtacctgt gtatcgacga ctgttggatg 240

gccccccagc gagactctga gggccgactc caggccgacc cccagcgatt cccccacggc 300

atccgacagc tcgccaacta cgtgcactct aagggcctga agctgggcat ctacgccgac 360

gtgggcaaca agacctgtgc cggcttcccc ggctctttcg gctactacga catcgacgcc 420

cagaccttcg ccgactgggg cgtggacctg ctgaagttcg acggctgtta ctgtgactct 480

ctcgagaacc tggccgacgg ctacaagcac atgtctctgg ccctgaaccg aaccggccga 540

tctatcgtgt actcttgtga gtggcccctg tacatgtggc ccttccagaa gcccaactac 600

accgagatcc gacagtactg taaccactgg cgaaacttcg ccgacatcga cgactcgtgg 660

aagtctatca agtctattct ggactggacc tctttcaacc aggagcgaat cgtcgacgtc 720

gccggacccg gcggatggaa cgaccccgac atgctggtga tcggcaactt cggcctgtct 780

tggaaccagc aggtgaccca gatggccctg tgggctatca tggctgcccc cctgttcatg 840

tctaacgacc tgcgacacat ctctccccag gccaaggccc tgctccagga caaggacgtg 900

atcgccatca accaggaccc cctgggcaag cagggctacc agctccgaca gggcgacaac 960

ttcgaggtgt gggagcgacc cctgtctggc ctggcctggg ccgtggccat gatcaaccga 1020

caggagatcg gcggaccccg atcttacacc atcgccgtgg cctccctggg aaagggcgtg 1080

gcctgtaacc ccgcctgttt catcacccag ctcctgcccg tgaagcgaaa gctgggattc 1140

tacgagtgga cctctcgact gcgatctcac atcaacccca ccggcaccgt gctgctccag 1200

ctcgagaaca ccatgcagat gtctctgaag gacctgctga cgcgtgaaca aaaactcatc 1260

tcagaagagg atctgaatag cgccgtcgac catcatcatc atcatcat 1308

<210> 23

<211> 436

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic polypeptide

<400>

23

Met Lys Leu Ser Thr Ile Leu Phe Thr Ala Cys Ala Thr Leu Ala Leu
1 5 10 15

Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp Glu
20 25 30

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

Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu Gly
50 55 60

Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp Met

65 70 75 80

Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln Arg
85 90 95

Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys Gly
100 105 110

Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys Ala Gly
115 120 125

Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe Ala
130 135 140

Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp Ser
145 150 155 160

Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu Asn
165 170 175

Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr Met
180 185 190

Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys Asn
195 200 205

His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile Lys

210 215 220

Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp Val
225 230 235 240

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

245 250 255

Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp Ala

260 265 270

Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile Ser

275 280 285

Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala Ile Asn

290 295 300

Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp Asn

305 310 315 320

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

325 330 335

Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile Ala

340 345 350

Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe Ile

355 360 365

Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp Thr

370 375 380

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

385 390 395 400

Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu Thr Arg Glu

405 410 415

Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser Ala Val Asp His His

420 425 430

His His His His

435

<210> 24

<211> 2743

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic polynucleotide

<400> 24

atgaagcttt ccaccatcct cttcacagcc tgcgctaccc tggctgccgc ccagcaggga 60

gcctctcgac ccggaccccg agatgcccag gctcaccccg gacgacctcg agctgtgccc 120

acccagtgtg acgtgccccc caactctcga ttcgactgtg cccccgacaa ggccatcacc 180

caggagcagt gcgaggcccg aggctgttgt tacatccccg ctaagcaggg cctgcagggc 240

gctcagatgg gccagccctg gtgtttcttc ccccccctct acccctccta caagctggag 300

aacctgtcct cttcggagat gggctacacc gccaccctga cccgaaccac ccccaccttt 360

ttccccaagg acatcctgac cctgcgactg gacgtgatga tggagaccga gaaccgactg 420

cacttcacca tcaaggaccc cgccaaccga cgatacgagg tgcccctgga gacccccccac 480

gtgcactctc gagccccttc cccctgtac tctgtggagt tctctgagga gcccttcggc 540

gtgatcgtgc gacgacagct ggacggccga gtgctgctga acaccaccgt ggcccccctg 600

ttcttcgccg accagttcct gcagctgtct acctctctgc cctctcagta catcaccggc 660

ctggccgagc acctgtcccc cctgatgctg tccacctctt ggactcgaat caccctgtgg 720

aaccgagacc tggcccccac ccccggtgcc aacctgtacg gctctcaccc cttctacctg 780

gccctggagg acggcggctc tgcccacggc gtgtttctgc tgaactctaa cgccatggac 840

gtggtgctgc agccctctcc cgccctgtct tggcgatcta ccggcggcat cctggacgtg 900

tacatcttcc tgggccctga gcccaagtct gtggtccagc agtacctgga cgtggtcgga 960

taccccttca tgcccccccta ctggggcctg ggcttccacc tgtgtcgatg gggctactct 1020

tctaccgcca tcacccgaca ggtggtggag aacatgaccc gagcccactt ccccctggac 1080

gtgcaatgga acgacctgga ctacatggac tctcgacgag acttcacctt caacaaggac 1140

ggcttccgag acttccccgc catggtccag gagctgcacc agggaggacg acgatacatg 1200

atgatcgtgg accccgccat ctcttcttcc ggacccgccg gatcttaccg accctacgac 1260

gagggcctgc gacgaggcgt gttcatcacc aacgagaccg gccagcccct gatcggcaag 1320

gtgtggcccg gctctaccgc cttccccgac ttcaccaacc ccaccgccct ggcttggtgg 1380

gaggacatgg tggccgagtt ccacgaccag gtgcccttcg acggcatgtg gatcgacatg 1440

aacgagccct ctaacttcat ccgaggctct gaggacggct gtcccaacaa cgagctggag 1500

aacccccccct acgtgcccgg cgtggtgggc ggaaccctgc aggccgccac catctgtgcc 1560

tcttcgcacc agtttctgtc tacccactac aacctgcaca acctgtacgg actgaccgag 1620

gccattgcct ctcaccgagc cctggtgaag gcccgaggca cccgaccctt cgtgatctct 1680

cgatctacct tcgccggcca cggccgatac gccggacact ggaccggcga tgtgtggtcc 1740

tcttgggagc agctggcctc ttctgtgccc gagatcctgc agttcaacct gctgggcgtg 1800

cccctggtgg gcgccgacgt gtgtggcttc ctgggcaaca cctctgagga gctgtgtgtt 1860

cgatggaccc agctcggcgc cttctaccct ttcatgcgaa accacaactc cctgctgtct 1920

ctgccccagg agccctactc gttctctgag cccgctcagc aggccatgcg aaaggctctg 1980

accctgcgat acgccctgct gccccacctg tacaccctgt tccaccaggc ccacgtggct 2040
ggagagaccg tggcccgacc cctgttcctg gagttcccta aggactcttc tacctggacc 2100
gtggaccatc agctgctgtg gggcgaggcc ctcctgatca cccccgtgct gcaggccggc 2160
aaggctgagg tgaccggcta cttccctctg ggcacctggt acgacctgca gaccgtgcct 2220
gtggaggccc tgggatctct gccccctcct cccgccgctc cccgagagcc cgccatccac 2280
tctgagggcc agtgggtgac cctgcccgct cccctggaca ccatcaacgt gcacctgcga 2340
gccggctaca tcatccctct gcagggaccc ggcctgacca ccaccgagtc tcgacagcag 2400

cccatggccc tggccgtggc tctgaccaag ggcggagagg cccgaggcga gctgttctgg 2460
gacgatggcg agtctctgga ggtgctggag cgaggcgcct acacccaggt gatctttctg 2520
gcccgaaaca acaccatcgt gaacgagctg gtgcgagtga cctctgaggg cgctggtctg 2580
cagctccaga aggtgaccgt cctgggcgtg gccaccgctc cccagcaggt cctgtctaac 2640
ggcgtgcccg tgtctaactt cacctactct cccgacacca aggtgctgga catctgtgtg 2700
tctctgctga tgggcgagca gttcctggtg tcttggtgtt aac 2743

<210> 25
<211> 913
<212> PRT

<213> Artificial Sequence
<220><223> Synthetic polypeptide
<400> 25

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

85 90 95

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

100 105 110

Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu

115 120 125

Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile

130 135 140

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

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

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

210 215 220

Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp

225 230 235 240

Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His

245 250 255

Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe

260 265 270

Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala

275 280 285

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

290 295 300

Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly

305 310 315 320

Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg

325 330 335

```
Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met

            340                 345                 350
Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr
        355                 360                 365
Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp
    370                 375                 380
Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met
385                 390                 395                 400
Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr
                405                 410                 415

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

                485                 490                 495
Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr
            500                 505                 510
Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr
        515                 520                 525
His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser
    530                 535                 540
His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser
545                 550                 555                 560

Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly
                565                 570                 575
Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile
```

580 585 590

Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys

595 600 605

Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln

610 615 620

Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser

625 630 635 640

Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met

645 650 655

Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr

660 665 670

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

675 680 685

Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln

690 695 700

Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly

705 710 715 720

Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu

725 730 735

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

740 745 750

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

755 760 765

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

770 775 780

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

785 790 795 800

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

805 810 815

Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly

820 825 830

Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn

835 840 845

Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys
850 855 860
Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn
865 870 875 880
Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu
885 890 895
Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp
900 905 910
Cys

<

210> 26

<211> 2314

<212> PRT

<213> Streptomyces coelicolor

<400> 26

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

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

115 120 125

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

195 200 205
Asn Asn Phe Thr Ser Arg Thr Arg Ala Thr Leu Asp Ala Thr Arg Leu
210 215 220
Lys Asp Gly Thr Thr Ala Ile Ser Gly Leu Val Ala Gly Ser Phe Asn
225 230 235 240
Gly Ala Ser Tyr Arg Leu Tyr Tyr Tyr Ala Thr Thr Asn Val Pro Val
245 250 255
Thr Ser Leu Arg Thr Trp Gly Asp Asp Gly Ala Leu Gly Asp Ala Thr
260 265 270

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

340 345 350
Thr Lys Gln Phe Tyr Thr His Leu Tyr Arg Met Tyr Ala Leu Pro Val
355 360 365

Asn Ala Thr Ser Thr Ser Gly Thr Tyr Arg Gly Val Asp Gly Ala Val
370 375 380
His Lys Ala Asn Gly Phe Thr Tyr Tyr Asp Gly Trp Ser Thr Trp Asp
385 390 395 400
Asp Phe Arg Lys Tyr Ser Val Ala Ala Tyr Ile Asp Pro Ala Thr Tyr
405 410 415

Arg Asp Met Val Gln Ser Ala Val Ile Leu Phe Ala Asp Ala His Ala
420 425 430
Ala Gly Lys Ser Leu Gly Ser Leu Thr His Ser Val Pro Thr Val Arg
435 440 445
Trp Glu Arg Ser Ala Val Val Ile Ala Asp Ala Leu Ser Lys Gly Phe
450 455 460
Lys Asp Phe Asp Arg Leu Asp Glu Ala Tyr Pro Ala Leu Lys Ser Tyr
465 470 475 480
Thr Gly Tyr Tyr Thr Gly Thr Gln Leu Arg Gln Gly Tyr Ile Ala Gly
485 490 495
Asp Pro Gly Thr Thr Val Gln Arg Gly Tyr Asp Gln Trp Ala Leu Ser
500 505 510
Val Val Ala Asp Ala Leu Gly Glu Asp Ala Glu Ala Lys Lys Leu Arg
515 520 525
Glu Gln Ala Thr Met Ala Thr Asp Asn Leu Val Lys Pro Gly Ala Trp
530 535 540
Thr Ala Ala Asp Gly Thr Ala Val Gly Leu Leu Thr Pro Arg Asp Gly
545 550 555 560

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

Thr Met Leu His Ser Asn Ala Asn Glu Ile Asp Leu Gln Ala Pro Tyr

625 630 635 640

Leu Phe Asn Tyr Val Gly Glu Pro Ser Leu Thr Gln Lys Trp Val Arg

645 650 655

Ala Ile Tyr Thr Gly Glu Thr Trp Asn Arg Tyr Ile Ala Thr Gly Ser

660 665 670

Thr Asn Glu Ala Pro Ser Ser Gly Gly Glu Phe Arg Pro Pro Val Lys

675 680 685

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

690 695 700

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

705 710 715 720

Leu Phe Pro Val Thr Ala Gly Ser Ser Gln Phe Gln Ile Gly Ser Pro

725 730 735

Phe Phe Asp Ser Thr Thr Ile Thr Tyr Pro Asn Gly Ala Glu Phe Thr

740 745 750

Val Glu Ala Asp Gly Val Ser Pro Lys Asn Tyr Tyr Val Gln Arg Ala

755 760 765

Ala Leu Asn Gly Lys Arg Phe Ser Asn Thr Trp Leu Asp Tyr Ala Gln

770 775 780

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

785 790 795 800

Ser Trp Gly Ala Arg Thr Glu Pro Ala Tyr Ser Leu Asn Thr Asp Ser

805 810 815

Gly Asp Gly Asp Asp Glu His Ala Pro Gly Arg Gly Thr Thr Val Val

820 825 830

Ser Ala Arg Pro Glu Thr Val Arg Thr Ala Ala Asp Gly Thr Val Asp

835 840 845

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

850 855 860

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

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

Val Arg Glu Gly Asn Tyr Ser Asp Gly Ser Tyr Gly Ile Phe Arg Thr

965 970 975

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

980 985 990

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

995 1000 1005

Leu Thr Leu Asp Glu Gly Gly Tyr Ala Val Leu Gln Ala Glu Asp Pro

1010 1015 1020

Asp Arg Met Glu Gly Pro Ser Leu Val Lys Glu Ala Tyr Tyr Ser Asp

1025 1030 1035 1040

Gly Asp Leu Gly Gly Val Thr Glu Gly Ala Trp Glu Gln Tyr Thr Asp

1045 1050 1055

Leu Asp Phe Gly Gly Val Ala Pro Arg Ser Val Ser Val Arg Tyr Ala

1060 1065 1070

Asn Ser Gln Ala Ala Ala Ala Glu Pro Ser Ser Val Asp Ile His Ala

1075 1080 1085

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

1090 1095 1100

Gly Gly Trp Gln Tyr Tyr Thr Thr Val Arg Ala Ala Val Ser Asp Pro

1105 1110 1115 1120

Gln Ala Leu Leu Lys Ala Ser Ser Ala Thr Phe Val Phe His Ala Pro

1125 1130 1135

Ser Gly Arg Gln Trp Val Ser Asn Phe Asp Trp Tyr Gln Phe Ser Pro

1140 1145 1150

Glu Ala Ala Pro Ser Ser Ser Pro Ile Thr Thr Leu Ala Thr Leu Thr

1155 1160 1165

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

1170 1175 1180

Ser Gly Gly Val Phe Glu Asn Val Thr Asn Gly Ala Trp Ala Glu Trp

1185 1190 1195 1200

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

1205 1210 1215

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

1220 1225 1230

Pro Gly Ala Lys Asp Gly Pro Thr Ala Val Thr Val Pro Leu Asp Tyr

1235 1240 1245

Thr Gly Ser Gly Trp Gly Thr Val Ala Ser Thr Ser Val Arg Leu Asp

1250 1255 1260

Pro Asp Val Phe Glu Gly Thr Gln Asp Val Tyr Ala Val Phe Val Ser

1265 1270 1275 1280

Ser Thr Gln Thr Asp Ala Gln Pro Tyr Val Ala Asn Val His Ser Leu

1285 1290 1295

Thr Leu Thr Arg Gln Ala Asp Ala Pro Val Val Phe Asp Ala Thr Ala

1300 1305 1310

Phe Glu Gly Ser Ser Gly Gly Gly Leu Lys Ser Glu Pro Ala Thr Trp

1315 1320 1325

Ser Gly Ala Gly Ser Ala Thr Ser Leu Gly Gly Thr Tyr Asp Gly Ala

1330 1335 1340

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

1345 1350 1355 1360

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

1365 1370 1375

Ala Val Gln Leu Tyr Leu Asp Ser Phe Asp Pro Asp Ala Pro Gly Thr
1380 1385 1390
Pro Tyr Ala Thr Val Pro Leu Pro Val Thr Gly Ser Ser Trp Ser Ser
1395 1400 1405
Gly Gly Thr Thr Ser Leu Thr Leu Pro Glu Ala Ile Thr Gly Thr His
1410 1415 1420

Ala Val His Leu Arg Leu Thr Thr Asn Ala Asp Ser Ser His Pro Tyr
1425 1430 1435 1440
Val Ala Asn Leu Gly Gln Val Ala Phe Asp Arg Val Glu Ala Pro Ala
1445 1450 1455
Gln Thr Asp Leu Ser Ala Leu Arg Lys Ala Ile Glu Gln Tyr Glu Gly
1460 1465 1470
Leu Ser Glu Asp Ala Asp Arg Tyr Gly Thr Ile Asp Phe Gly Val Phe
1475 1480 1485
Arg Arg Glu Leu Thr Ala Ala Arg Asp Leu Leu Gly Thr Glu Asp Ala

1490 1495 1500
Thr Gln Leu Glu Ala Asp Leu Arg Thr Arg Ser Leu Thr Leu Ala Ala
1505 1510 1515 1520
Asn Gln Leu Val Pro Leu Pro Arg Leu Arg Leu Glu Ser Leu Val Ala
1525 1530 1535
Thr Ala Ser Ala Leu Ala Asp Glu Arg Tyr Thr Asp Ala Ser Trp Lys
1540 1545 1550
Ala Phe Thr Thr Ala Leu Thr Ala Ala Lys Thr Ala Leu Ala Asp Glu
1555 1560 1565

Thr Ala Thr Asp Arg Thr Leu Thr Glu Arg Tyr Ala Ala Leu Asp Arg
1570 1575 1580
Ala Arg Ser Ser Leu Thr Thr Lys Arg Arg Thr Val Pro Ala Ala Pro
1585 1590 1595 1600
Gly Ala Val Ser Ala Ala Pro Ser Gly Thr Ser Val Gln Val Thr Trp
1605 1610 1615
Ser Ala Pro Glu Asp Asp Gly Gly Ser Pro Val Thr Gly Tyr Glu Ile
1620 1625 1630

Thr Leu Ser Gly Gly Arg Gln Val Glu Ile Ala Asp Pro Asp Ser Arg

1635 1640 1645

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

1650 1655 1660

Val Arg Ala Val Asn Ala Leu Gly Asp Ser Pro Trp Ser Ala Arg Thr

1665 1670 1675 1680

Gln Pro Ala Val Thr Gly Asp Asn Arg Pro Gln Thr Pro Thr Val Thr

1685 1690 1695

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

1700 1705 1710

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

1715 1720 1725

Gly Thr Thr Ala His Val Pro Gly Thr Thr Ser Thr Ala Val Leu Thr

1730 1735 1740

Ala Ala Gly Gly Ala Lys Ala His Thr Ala Thr Val Thr Ala Val Thr

1745 1750 1755 1760

Arg Ala Gly Ser Ser Asp Gly Ser Gly Ala Thr Val Ser Thr Ala Pro

1765 1770 1775

Ala Thr Ser Thr Thr Ser Ala Thr Ser Ala Thr Ser Thr Gly Asp Pro

1780 1785 1790

Ala Glu Tyr Glu Pro Ser Pro Phe Pro Gly Asp Thr Leu Asp Ala Thr

1795 1800 1805

Tyr Ala Ser Asp Ala Trp Pro Glu Thr Gly Asp Gly Ser Asp Trp Phe

1810 1815 1820

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

1825 1830 1835 1840

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

1845 1850 1855

Arg Ile Thr Val Ser Val Asn Asn Ala Ala Thr Gln Gln Gln Val Asp

1860 1865 1870

Arg Ala Glu Val Asp Ala Ser Asn Ser Ala Thr Val Thr Met Ala Asp

1875 1880 1885

Gly Leu Gly Ser Arg Leu Gly Pro Leu Tyr Gly Glu Ala Leu Lys Glu

1890 1895 1900

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

1905 1910 1915 1920

Leu Asp Thr His Asp Ala Ala Lys Asn His Tyr Gln Tyr Leu Arg Pro

1925 1930 1935

Tyr Val Arg Leu Gly Phe Ala Gly Asp Gly Gly Ala Val Tyr Glu Ser

1940 1945 1950

Gln Asp Ser Ser Tyr Ser Gly Leu Ala Gly Gln Gly Ser Tyr Pro Ser

1955 1960 1965

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

1970 1975 1980

Leu Pro Asp Leu Ala Pro Ser Ile Leu Ala Arg Thr Ser Glu Tyr Gly

1985 1990 1995 2000

Asp Asn Arg Ile Val Leu Gly Phe His Tyr Pro Leu Asp Val Met Gly

2005 2010 2015

Gly Arg Ile Thr Ala Gln Ala Thr Val Ala His Arg Trp Ala Asp Pro

2020 2025 2030

Glu Phe Ala Lys Leu Leu Gly Gln Ala His Thr Glu Ile Glu Asn Val

2035 2040 2045

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

2050 2055 2060

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

2065 2070 2075 2080

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

2085 2090 2095

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

2100 2105 2110

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

2115 2120 2125

Thr Ala Thr Asp Ser Gly Tyr Pro Leu Asp Leu Thr Gly Ser Gly Gly

2130 2135 2140

Pro Gly Trp Gln Arg Ile Asn Leu Ala Ala Ala Met Ala Ala Asp Val
2145 2150 2155 2160
Val Val Asn Ala Asp Gly Ser Val Thr Val Thr Asn Phe Pro Asp Pro
2165 2170 2175
Thr Ala Ala Ser Ala Ala Glu Ala Val Ala Ile Thr Val Gly Gly Val
2180 2185 2190
Ala Leu Asp Gly Phe Asp Pro Asp Val Ser Thr Tyr Val Val Asp Trp
2195 2200 2205
Pro Arg Asn Gly Gly Arg Ile Pro Ala Val Gly Ala Val Thr Ala Ala

2210 2215 2220
Ser Gly Ala Arg Val Lys Val Thr Ser Gly Ser Ser Thr Val Ser Ser
2225 2230 2235 2240
Ser Gln Arg Gly Phe Ser Thr Arg Thr Leu Thr Val Thr Ser Ala Asp
2245 2250 2255
Gly Glu Phe Thr Arg Thr Tyr Thr Val Gly Phe Arg Pro Val Glu Gln
2260 2265 2270
His Pro His Arg Pro Gly Ala Leu Arg Asp Thr Gly Gly Gly Gly Thr
2275 2280 2285

Ala Gly Gly Ser Ala Gly Gly Gly Asp Val Gly Gly Gly Leu Trp Ser
2290 2295 2300
Pro Ala Arg Glu Trp Glu Leu Thr Val Asn
2305 2310
<210> 27
<211> 1984
<212> PRT
<213> Clostridium spiroforme
<400> 27
Met Asn Lys Lys Ile Asn Arg Leu Leu Lys Gly Ala Leu Ala Phe Thr
1 5 10 15
Val Ala Phe Ser Thr Leu Ala Val Thr Thr Ser Thr Ser Arg Lys Val

20 25 30

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

100 105 110
Val Leu Val Thr Pro Thr Tyr Val Glu Tyr Ser Gln Arg Pro Gln Ala
115 120 125
Gln Thr Arg Ala Met Asn Tyr Thr Lys Glu Asp Glu Ser Ala Lys Pro
130 135 140
Gly Tyr Tyr Ser Val Glu Leu Thr Pro Lys Thr Gly Lys Asp Asn Asp
145 150 155 160
Val Lys Asp Ser Pro Glu Ile Gly Lys Ile Lys Ala Glu Met Thr Thr
165 170 175

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

245 250 255
Thr Trp Asn Asp Thr Thr Leu Ser Asp Lys Thr Ser Gln Lys Gly Asn
260 265 270

Asp Ile Gly Thr Ile Met Asn Phe Asp Val Lys Glu Asn Glu Glu Ile
275 280 285

Gln Leu Lys Val Ser Ile Ser Pro Ile Ser Val Lys Gln Ala Lys Ile
290 295 300

Asp Met His Asn Glu Ile Ser Asp Trp Asp Phe Asp Ala Ala Ala Ser
305 310 315 320

Arg Ala Asp Lys Ala Trp Asn Asp Val Leu Ser Lys Val Arg Val Glu
325 330 335

Ser Ser Lys Val Ser Asp Pro Thr Gly Glu Leu Lys Gln Leu Phe Tyr
340 345 350

Thr His Leu Tyr His Met Phe Met Thr Pro Val Asn Ala Thr Ser Thr
355 360 365

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

Tyr Ile His Tyr Asp Ser Trp Thr Leu Trp Asp Asp Tyr Arg Lys Tyr

385 390 395 400

Pro Met Ile Gly Leu Ile Gln Pro Asp Thr Tyr Lys Asp Met Val Lys
405 410 415

Ser Ile Ala Asp Ala Leu Asp Tyr Gly Ile Val Thr Trp Ser His Asp
420 425 430

Lys Gln Pro Val Pro Asn Val Arg Thr Glu His Ala Val Ala Leu Leu
435 440 445

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

Ala Tyr Glu Glu Ala Lys Glu Ile Val Asn Glu Val Ile Thr Ser Glu
465 470 475 480

Val Glu Lys Ile Gly Tyr Val Pro Asn Arg Val Asp Arg Thr Ile Glu
485 490 495

Tyr Gly Tyr Asp Asp Trp Cys Leu Ser Ile Ile Ala Glu Ala Leu Gly
500 505 510

Lys Glu Asp Glu Ala Ala Tyr Phe Leu Asp Arg Ser Phe Asn Tyr Lys
515 520 525

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

530 535 540

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

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

675 680 685
Met Ala Met Met Phe Val Ser Ala Ala Met Gly Ile Phe Pro Met Thr
690 695 700
Pro Gly Asp Thr Thr Phe Gln Ile Gly Ser Pro Phe Phe Glu Lys Ile
705 710 715 720
Thr Leu Asp Val Gly Asn Gly Lys Thr Phe Thr Ile Glu Ala Asn Asn
725 730 735
Val Ser Asp Thr Asn Glu Tyr Ile Gln Ser Ala Thr Leu Asn Gly Lys
740 745 750

Ser Phe Asp Arg Thr Trp Val Asp Tyr Ser Glu Ile Thr Arg Gly Gly
755 760 765
Val Leu Ser Phe Glu Met Gly Asp Thr Pro Ser Ser Trp Ala Gln Asn

770 775 780

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

785 790 795 800

Asp Asp Asp Glu Ile Ala Tyr Ser Ser Ala Met Phe Glu Glu Ser Lys

805 810 815

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

820 825 830

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

835 840 845

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

850 855 860

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

865 870 875 880

His Thr Leu Asn Asp Ser Ile Ser Asn Leu Thr Ile Glu Ile Thr Asp

885 890 895

Gly Ala Thr Asn Glu Pro Ile Lys Asp Ser Ile Arg Lys Thr Lys Asp

900 905 910

Asn Val Lys Val Met Phe Ile Asp Asn Gln Leu Thr Tyr Ser Gln Ser

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

Leu Thr Val Lys Met Ile Lys Ile Asp Asp His Thr Ala Val Leu Ser

980 985 990

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

995 1000 1005

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

1010 1015 1020

Gln Ser Ser Arg Gly Gly Met Thr Ala Leu Leu Leu Asp Phe Asp Tyr

1025 1030 1035 1040

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

1105 1110 1115 1120
Val Thr Ser Gly Gly Ser Leu Arg Val Glu Gly Ser Val Leu His Gly
1125 1130 1135
Thr Tyr Asp Gly Ala Trp Ile Arg Tyr Asp Ala Leu Asp Phe Asn Gly
1140 1145 1150
Leu Ser Pro Lys Tyr Leu Glu Leu Arg Tyr Asp Asn Ala Ser Asn Arg
1155 1160 1165
Cys Ala Ser Asp Ser His Leu Glu Val Arg Leu Asp Gly Val Asp Gly
1170 1175 1180

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

1250 1255 1260
Gly Asn Gly Leu Lys Asn Glu Ser Ile Asn Leu Gly Gly Thr Tyr Asp
1265 1270 1275 1280

Gly Ala Trp Ile Lys Tyr Asn Asn Val Asn Phe Asn Asn Leu Glu Ala
1285 1290 1295
Asp Thr Ile Asn Val His Tyr Ser Thr Arg Val Asp Ala Cys Ala Leu
1300 1305 1310
Asp Ala Arg Ile Glu Ile Arg Lys Asp Asn Lys Asp Gly Glu Leu Leu
1315 1320 1325

Gly Thr Ile Met Leu Pro Leu Thr Gly Gly Trp Ser Asp Tyr Gln Thr
1330 1335 1340
Val Ser Thr Lys Leu Asp Thr Ser Val Thr Gly Val Gln Asp Ile Cys
1345 1350 1355 1360
Phe Val Leu Arg Gly Thr Asn Asp Gly Gly Arg Pro Tyr Val Ala Asn
1365 1370 1375
Ile Asp Tyr Met Glu Phe Val Asn Ser Gly Val Asn His Ile Glu Ala
1380 1385 1390
Glu Asn Lys Asp Asp Trp Ser Gly Ala Glu Leu Lys Val Glu Asn Ser

1395 1400 1405
Thr Asp Asn Thr Gly Lys Ser Leu Thr Asn Ile Gly Gly Ala Arg Asn
1410 1415 1420
Asp Ala Trp Leu Arg Tyr Asn Gly Val Glu Phe Asn Gly Lys Thr Glu
1425 1430 1435 1440
Met Thr Val Arg Tyr Ser His Asn Pro Gly Thr Ala Gly Thr Asn Ser
1445 1450 1455
Arg Ile Asp Val Tyr Leu Asp Asn Met Asp Gly Asn Pro Ile Gly Thr
1460 1465 1470

Ile Asn Leu Pro Thr Thr Asn Gly Trp Ala Asn Tyr Thr Val Ile Arg
1475 1480 1485
Glu Val Phe Asp Gln Glu Ile Thr Gly Ser His Asp Val Tyr Leu Lys
1490 1495 1500
Leu His Thr Asp Gly Ser Gly Trp Val Ala Asn Phe Asp Trp Phe Glu
1505 1510 1515 1520
Phe Gly Glu Pro Ile Ala Asp Val Asp Lys Ser Gln Leu Gln Ala Lys
1525 1530 1535

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

1540 1545 1550

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

1555 1560 1565

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

1570 1575 1580

Asp Ile Asp Asn Ala Leu Ala Ala Leu Gln Tyr Lys Ile Ala Phe Asp

1585 1590 1595 1600

Leu Asn Asp Tyr Val Val Gln Leu Glu Asn Ile Asn Glu Ala Asp Tyr

1605 1610 1615

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

1620 1625 1630

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

1635 1640 1645

Gly Leu Val Asp Ala His Ser Lys Leu Thr Ala Leu Asn Arg Thr Ala

1650 1655 1660

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

1665 1670 1675 1680

Lys Glu Glu Gly Lys Ala Glu Phe Lys Ala Ala Leu Glu Asn Ala Lys

1685 1690 1695

Thr Val Tyr Glu Thr Val Ser Leu Thr Gln Ala Gln Val Asp Glu Ala

1700 1705 1710

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

1715 1720 1725

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

1730 1735 1740

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

1745 1750 1755 1760

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

1765 1770 1775

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

1780 1785 1790

Val Met Gln Lys Leu Glu Phe Phe Lys Gly Asp Lys Lys Ala Leu Lys

1795 1800 1805

Ala Phe Ile Asp Lys Val Thr Gly Leu Asp Ser Ser Lys Tyr Thr Gln

1810 1815 1820

Thr Thr Trp Thr Ala Phe Asp Lys Glu Leu Thr Glu Ala Ile Ala Val

1825 1830 1835 1840

Tyr Asn Asp Glu Asn Ala Met Gln Glu Glu Val Asn Thr Ala Tyr Ser

1845 1850 1855

Glu Leu Val Thr Ala Phe Leu Asn Leu Arg Leu Ile Pro Asp Lys Ser

1860 1865 1870

Leu Leu Glu Asp Leu Ile Asn Gln Ala Asn Gly Leu Asn Gly Ala Asn

1875 1880 1885

Tyr Thr Lys Ala Thr Phe Asp Gly Leu Thr Lys Ala Leu Asp Glu Ala

1890 1895 1900

Lys Ala Val Tyr Glu Asn Pro Asp Ala Thr Gln Lys Glu Val Pro Arg

1905 1910 1915 1920

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

1925 1930 1935

Thr Asp Asn Thr Val Ser Thr Pro Val Asn Asn Gly Asp Thr Thr Ala

1940 1945 1950

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

1955 1960 1965

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

1970 1975 1980

<210> 28

<211> 653

<212> PRT

<213> Bacteroides thetaiotaomicron

<400> 28

Met Thr Pro Ser Val Ala Gln Asn Thr Lys Tyr Val Asn Leu Phe Ile

1 5 10 15

Gly Thr Ser Gly Asp Asn Gly Gln Val Ala Pro Gly Ala Ala Ala Pro
20 25 30
Phe Gly Met Val Cys Val Cys Pro Asp Asn Asp Pro Arg Ser His Ala
35 40 45

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

115 120 125
Arg Ser Leu Ser Ala Ala Leu Trp Ile Asp Phe Ala Ser Thr Phe Glu
130 135 140
Asp Val Ala Thr Cys His Tyr Lys Arg Ile Ser Glu Thr Cys Ile Glu
145 150 155 160
Gly Tyr Val Gln Ala Lys Asn Val Cys Gly His Gly Cys Tyr Lys Leu
165 170 175
Tyr Phe Ser Leu Asn Thr Ser Gln Pro Phe Gln Leu Glu Glu Gln Lys
180 185 190

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

260 265 270

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

275 280 285

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

290 295 300

Ser Leu Trp Asp Thr Phe Arg Thr Lys Phe Pro Leu Leu Val Leu Thr

305 310 315 320

Glu Pro Ala Lys Met Arg Asp Met Ala Thr Ser Leu Ile His Leu Tyr

325 330 335

Ala Thr Gly Lys Lys Asp Trp Ser Thr Gly Phe Glu Ser Thr Pro Thr

340 345 350

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

355 360 365

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

370 375 380

Glu Met Glu Arg Leu Pro Met Arg Ser Pro Asp Gln Lys Met Glu Ser

385 390 395 400

Ala Tyr Asp Leu Trp Ala Met Ala Lys Ile Ala Glu Ile Ile Gly Glu

405 410 415

Lys Ala Asp Ser Glu Gln Tyr Arg Gln Arg Ser Val Ser Leu Phe Glu

420 425 430

Glu Thr Trp Lys Lys Glu Phe Met Asn Val Thr Pro Ala Phe Glu Val

435 440 445

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

450 455 460

Ala Ala Pro Gln Tyr Ile Asp Lys Met Ile Glu Trp Val Gly Gln Asp

465 470 475 480

Ser Leu Arg Leu Gln Leu Thr Tyr Phe Phe Asp His His Leu Tyr Asn

485 490 495

Gln Gly Asn Glu Pro Asp Ile His Val Pro Tyr Leu Phe Asn Arg Leu

500 505 510

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

515 520 525

Pro Met Ile His Lys Tyr Gly Gly Asn Ser Glu Phe Lys Thr Pro Tyr

530 535 540

Leu Gly Lys Ala Phe Lys Asn Ala Pro Glu Gly Tyr Ser Pro Glu Met

545 550 555 560

Asp Glu Asp Asp Gly Thr Met Ser Ala Trp Tyr Val Phe Gly Ala Met

565 570 575

Gly Phe Tyr Pro Leu Leu Val Gly Asp Glu Tyr Tyr Asp Leu Thr Ser

580 585 590

Pro Leu Phe Asp Arg Val Leu Leu Arg Leu Thr Asn Gly Asn Val Leu

595 600 605

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

610 615 620

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

625 630 635 640

Glu Leu Ile Lys Gly Gly Glu Leu Ile Tyr Asn Tyr Lys

645 650

<210> 29

<211> 696

<212> PRT

<213> Unknown

<220><223> Zunongwangia profunda

<400> 29

Met Thr Leu Ile Met Lys Arg Met Ile Ala Ala Ile Ala Val Ala Leu

1 5 10 15

Cys Val Ala Cys Gln Pro Lys Ser Gln Glu Lys Thr Ser Lys Ser Ala

20 25 30

Asn Ile Thr Asp Lys Val Asn Val Phe Leu Gly Ser Ser Gly Asp His

35 40 45

Gly Gln Met Ser Pro Ser Ala Ser Thr Pro Phe Asn Met Met Ser Ile

50 55 60

Gly Pro His Thr Asn Pro His Asn His Thr Gly Tyr Glu His Tyr Ala

65 70 75 80

Lys Glu Phe Asp Gly Phe Thr His Thr His Leu Glu Gly Val Gly Cys

85 90 95

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

100 105 110

Lys Glu Thr Glu Leu Arg Lys Val Thr Glu His Ala Lys Pro Gly Phe

115 120 125

Tyr Glu Val Ser Phe Glu Asn Gly Ile Asp Ala Ala Met Ser Val Thr

130 135 140

His Asn Phe Gly Ile His Gln Tyr Asn Phe Asn Gly Glu Lys Gly Gly

145 150 155 160

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

165 170 175

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

180 185 190

Thr Thr Cys His Ala Gly Thr Tyr Arg Phe Tyr Tyr Glu Ile Gln Leu

195 200 205

Lys Asn Met Ala Glu Ile Ala Gln Ile Ser Asp His Glu Ile Met Ala

210 215 220

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

225 230 235 240

Ser Val Ser Lys Glu Tyr Ala Ser Gln Lys Ile Glu Asn Ile Ser Tyr

245 250 255

Glu Asn Leu Lys Lys Glu Ala Ser Ala Ala Trp Glu Lys Ala Leu Ser

260 265 270

Arg Ile Ser Val Glu Gly Glu Glu Asp Arg Glu Asp Leu Phe Tyr Ser

275 280 285

Leu Leu Tyr Arg Gly Leu Gln Ser Pro Tyr Ile Val Ser Glu Glu Asp

290 295 300

Gly Thr Tyr Pro Ala Ile Asp Gly Thr Leu Gln Lys Thr Glu Gly Thr

305 310 315 320

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

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

450 455 460
Asp Val Asp Arg Met Gln Ala Arg Gly Leu Tyr Gln Gly Thr Ile Trp
465 470 475 480
Gln Tyr Arg Trp Phe Val Pro Phe Asp Leu Asn Gly Leu Lys Gln Leu
485 490 495
Ala Gly Gly Glu Asp Gln Phe Leu Glu Glu Leu Asp Thr Phe Phe Arg
500 505 510
Asn His Asn Tyr Asn His Ala Asn Gln Pro Asp Leu Gln Val Pro Gly
515 520 525

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

Ala Tyr Leu Arg Thr Met Asp Asp Asp Ala Gly Thr Met Ser Ser Trp
580 585 590
Phe Val Met Arg Ser Leu Gly Leu Ser Pro Ala Asn Ile Gly Asp Pro

595 600 605
Val Tyr Tyr Leu Thr Ala Pro Ile Phe Lys Glu Ile Ser Ile Asn Tyr
610 615 620
Pro Lys Gly Lys Ala Phe Lys Ile Ser Val Thr Asn Tyr Asn Lys Asp
625 630 635 640
His Tyr Tyr Val Glu Ser Ala Thr Leu Asn Gly Lys Pro Leu Asn Arg
645 650 655
Asn Trp Leu Thr Gln Gln Glu Ile Leu Glu Gly Gly Glu Leu Val Ile
660 665 670

Lys Thr Ser Asp Thr Pro Asn Lys Glu Trp Gly Val Lys Glu Ala Trp
675 680 685
Val Ser Ser Ile Arg Gln Tyr Leu
690 695

<210> 30
<211> 680
<212> PRT
<213> Unknown
<220><223> Chitinophaga pinensis
<400> 30

Met Lys Lys Ser Leu Ile Tyr Leu Leu Ser Leu Leu Ser Leu Thr Ala
1 5 10 15
Val Ala Gln Gln Ser Gly Gln Val Asn Val Phe Leu Gly Ser Ser Gly
20 25 30

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

Gly Cys Gln Gly Cys Gly Gly Asn Leu Leu Leu Arg Pro Phe Leu Gly
85 90 95
Asp Gly Pro Val Lys Ala Asp Leu Ile Lys Tyr Glu Glu Gln Ala Ser

100 105 110
Pro Gly Tyr Tyr His Val Gly Phe Thr Asn Gly Ile Lys Ala Ser Phe
115 120 125
Thr Val Tyr Lys Asn Ala Gly Leu His Gln Tyr Thr Phe Pro Glu Gly
130 135 140
Lys Lys Gly Leu Leu Leu Asp Leu Gly Phe Ala His Val Gly Arg Phe
145 150 155 160
Val Ala Glu Glu His Thr Ile Glu Gly Asn Ala Val Ser Gly Trp Val
165 170 175

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

245 250 255
Gly His Ile Lys Val Lys Gly Asp Pro Ala Arg Glu Lys Leu Phe Tyr
260 265 270
Ser Leu Phe Tyr Arg Ser Ile Gln Ser Pro Tyr Val Val Ser Glu Pro
275 280 285
Asp Gly Ala Tyr Ala Ala Thr Asn Gly Thr Leu Gln His Thr Asn Ser
290 295 300
Lys Met Tyr Asn Gly Trp Ala Ile Trp Asp Asn Tyr Arg Ala Gln Leu
305 310 315 320

Pro Leu Leu Ser Ile Ala Phe Pro Gln Glu Tyr Gln Asp Met Thr Asn

325 330 335

Ser Ile Ala Gly Leu Tyr Ala His Gly Lys Lys Asp Tyr Ala Thr Leu

340 345 350

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

355 360 365

Leu Asp Ala Leu Lys Lys Gly Tyr Lys Phe Asp Phe Asn Ala Ile Ala

370 375 380

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

385 390 395 400

Lys Ala Leu Glu Ser Ser Tyr Asp Ala Trp Ala Leu Ala Glu Leu Tyr

405 410 415

Tyr Ala Gln Lys Asp Lys Ala His Gly Asp Gln Tyr Lys Val Gln Ala

420 425 430

Ala Asp Tyr Lys Lys Tyr Trp Leu Lys Asp Phe Gln Asp Leu Thr Lys

435 440 445

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

450 455 460

Trp Gln Tyr Arg Trp Phe Val Pro Phe Asp Leu Lys Gly Leu Met Glu

465 470 475 480

Leu Cys Gly Gly Glu Gln Ala Tyr Leu Ser Gln Leu Asp Glu Phe Phe

485 490 495

Asp Asn Asp Tyr Tyr Cys His Ala Asn Gln Pro Asp Leu Gln Thr Pro

500 505 510

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

515 520 525

His Lys Ile Ala Val Asp Thr Met Val Gln His Tyr Phe Asn Asp Asn

530 535 540

Ser Arg Gly Ile Gly Ser Glu Ile Gly Pro Ile Tyr Lys Asn Gln Pro

545 550 555 560

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

565 570 575

Trp Phe Val Leu Val Ser Thr Gly Ile Phe Pro Ala Cys Ile Gly Ser

580 585 590

Pro Val Tyr Tyr Leu Asn Val Pro Leu Phe Glu Ser Val Glu Trp Gln

595 600 605

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

610 615 620

Lys Asn Val Tyr Ile Lys Glu Val Trp Leu Asn Gly Arg Lys Leu Asp

625 630 635 640

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

645 650 655

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

660 665 670

Trp Ile Ala Asp Ile Thr Arg Gln

675 680

<210> 31

<211> 867

<212> PRT

<213> Caulobacter crescentus

<400> 31

Met Asn Pro Gly Lys Gly Pro Val Leu Arg Phe Arg Val Gln Leu Leu

1 5 10 15

Val Val Gly Phe Ala Leu Ala Leu Gly Ala Ala His Gln Ala Gln Ala

20 25 30

Ala Pro Gly Arg Ala Leu Tyr Lys Leu Pro Asp Leu Arg Val Gly Leu

35 40 45

Gly Asp Asp Asn Gln Gly Asp Thr Ile Pro Gly Pro Thr Arg Pro Asn

50 55 60

Gly Ser Ile His Pro Ser Pro Asp Thr Leu Lys Ala Ser Asn Ala Gly

65 70 75 80

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

85 90 95

Gly Ser Gly Gly Val Thr Thr Tyr Gly Thr Phe Leu Leu Ser Pro Gln

100 105 110

Val Gly Glu Pro Val Phe Asp Glu Ala Ala His Leu Ser Pro Lys Ala
115 120 125

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

195 200 205
Trp Asn Pro Ala Leu Ile Asp Ile Trp Phe Val Ala Glu Phe Asp Gln
210 215 220
Asn Pro Thr Ala Trp Gly Val Phe Asp Lys Ala Glu Arg Arg Asp Gly
225 230 235 240
Ala Leu Ser Gly Arg Thr Gly Ser Asp Glu Arg Leu Gly Ala Trp Leu
245 250 255
Thr Phe Lys Thr Thr Pro Thr Lys Pro Leu Leu Val Lys Ile Ala Val
260 265 270

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

340 345 350
Asp Glu His Tyr Thr Leu Trp Asp Thr Tyr Arg Thr Leu Phe Pro Leu

355 360 365

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

370 375 380

Ile His Thr Phe Asp Lys Phe Gly Ala Ala Asp Thr Ala Ile Ile Gly

385 390 395 400

Gly Gln Asn Tyr His Val Gly Gln Gly Gly Asp Glu Val Asp Asn Val

405 410 415

Leu Gly Glu Ala Leu Leu Arg Gly Ala Glu Gly Val Asn Trp Arg Asp

420 425 430

Ala Trp Arg Val Ala Arg Phe Asn Ala Phe Glu Arg Arg Arg Pro Arg

435 440 445

Tyr Leu Glu Ser Gly Tyr Phe Ala Val Gly Asp Arg Ser Pro Glu Pro

450 455 460

Asn Asn Gln Arg Ala Lys Ser Gly Ser Ser Thr Leu Gly Phe Ala Leu

465 470 475 480

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

485 490 495

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

500 505 510

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

515 520 525

Tyr Ala Asp Gly Lys Phe Gln Asp Ile Asp Pro Lys Leu Gly Trp Asp

530 535 540

Gly Lys Val His Asn Asn Val Gly Tyr Tyr Glu Gly Thr Ala Trp Ile

545 550 555 560

Tyr Ser Tyr Gly Val Leu His Asp Leu Pro Gly Leu Val Glu Ala Met

565 570 575

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

580 585 590

Gly Leu Ile Asp Ile Thr Asn Glu Pro Ser Phe Ala Thr Pro Trp Leu

595 600 605

Phe His Ala Ile Gly Arg Ala Asp Leu Ser Ser Arg Trp Ala Gly Glu

610 615 620

Val Val Lys His Phe Thr Ala Asp Ala Tyr Pro Gly Asp Glu Asp Ala

625 630 635 640

Gly Ala Met Ser Ser Asn Phe Val Phe Asn Ser Leu Gly Leu Phe Pro

645 650 655

Lys Leu Gly Ser Asp Leu Tyr Tyr Leu His Gly Pro Arg His Gly Arg

660 665 670

Thr Val Ile Gln Leu Glu Asn Gly Lys Thr Leu Glu Ile Leu Ala Ala

675 680 685

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

690 695 700

Lys Pro Leu Ala Gly Pro Tyr Val Ser Gln Ala Gln Leu Leu Gly Gly

705 710 715 720

Gly Val Leu Ser Leu Ser Met Ser Asp Gln Pro Gly Gln Trp Ile Tyr

725 730 735

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

740 745 750

Gly Lys Thr Ser Thr Gly Trp Arg Ala Ala Ser Gly Gln Ser Val Thr

755 760 765

Phe Ser Leu Lys Ala Pro Ala Cys Ile Ala Ala Tyr Ser Val Ser Val

770 775 780

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

785 790 795 800

Gly Arg Ala Trp Val Ser Val Asp Gln Gln Ser Asn Val Val Phe Asp

805 810 815

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

820 825 830

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

835 840 845

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

850 855 860

Pro Leu Leu

865

<210> 32

<211> 867

<212> PRT

<213> Caulobacter crescentus

<400> 32

Met Asn Pro Gly Lys Gly Pro Val Leu Arg Phe Arg Val Gln Leu Leu

1 5 10 15

Val Val Gly Phe Ala Leu Ala Leu Gly Ala Ala His Gln Ala Gln Ala

20 25 30

Ala Pro Gly Arg Ala Leu Tyr Lys Leu Pro Asp Leu Arg Val Gly Leu

35 40 45

Gly Asp Asp Asn Gln Gly Asp Thr Ile Pro Gly Pro Thr Arg Pro Asn

50 55 60

Gly Ser Ile His Pro Ser Pro Asp Thr Leu Lys Ala Ser Asn Ala Gly

65 70 75 80

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

85 90 95

Gly Ser Gly Gly Val Thr Thr Tyr Gly Thr Phe Leu Leu Ser Pro Gln

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

Leu Thr Tyr Pro Thr Thr Asp Gln Ala Gln Val Val Leu Asp Val Thr

165 170 175

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

180 185 190

Phe Pro Glu Gln Gly Arg Ile Val Gly His Val Lys Ala Lys Gly Tyr

195 200 205

Trp Asn Pro Ala Leu Ile Asp Ile Trp Phe Val Ala Glu Phe Asp Gln

210 215 220

Asn Pro Thr Ala Trp Gly Val Phe Asp Lys Ala Glu Arg Arg Asp Gly

225 230 235 240

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

245 250 255

Thr Phe Lys Thr Thr Pro Thr Lys Pro Leu Leu Val Lys Ile Ala Val

260 265 270

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

275 280 285

Pro Asp Trp Asp Phe Glu Arg Val Arg Arg Asp Thr Gln Ala Ala Trp

290 295 300

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

305 310 315 320

Gln Arg Arg Phe Tyr Ser Ala Leu Tyr His Ala Ser Thr His Pro Arg

325 330 335

Asp Arg Ser Leu Asp Gln Pro Ala Ala Arg Leu Gly Arg Pro Asn Trp

340 345 350

Asp Glu His Tyr Thr Leu Trp Asp Thr Tyr Arg Thr Leu Phe Pro Leu

355 360 365

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

370 375 380

Ile His Thr Phe Asp Lys Phe Gly Ala Ala Asp Thr Ala Ile Ile Gly

385 390 395 400

Gly Gln Asn Tyr His Val Gly Gln Gly Gly Asp Glu Val Asp Asn Val

405 410 415

Leu Gly Glu Ala Leu Leu Arg Gly Ala Glu Gly Val Asn Trp Arg Asp

420 425 430

Ala Trp Arg Val Ala Arg Phe Asn Ala Phe Glu Arg Arg Arg Pro Arg

435 440 445

Tyr Leu Glu Ser Gly Tyr Phe Ala Val Gly Asp Arg Ser Pro Glu Pro

450 455 460

Asn Asn Gln Arg Ala Lys Ser Gly Ser Ser Thr Leu Gly Phe Ala Leu

465 470 475 480

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

485 490 495

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

500 505 510

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

515 520 525

Tyr Ala Asp Gly Lys Phe Gln Asp Ile Asp Pro Lys Leu Gly Trp Asp

530 535 540

Gly Lys Val His Asn Asn Val Gly Tyr Tyr Glu Gly Thr Ala Trp Ile

545 550 555 560

Tyr Ser Tyr Gly Val Leu His Asp Leu Pro Gly Leu Val Glu Ala Met

565 570 575

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

580 585 590

Gly Leu Ile Asp Ile Thr Asn Glu Pro Ser Phe Ala Thr Pro Trp Leu

595 600 605

Phe His Ala Ile Gly Arg Ala Asp Leu Ser Ser Arg Trp Ala Gly Glu

610 615 620

Val Val Lys His Phe Thr Ala Asp Ala Tyr Pro Gly Asp Glu Asp Ala

625 630 635 640

Gly Ala Met Ser Ser Asn Phe Val Phe Asn Ser Leu Gly Leu Phe Pro

645 650 655

Lys Leu Gly Ser Asp Leu Tyr Tyr Leu His Gly Pro Arg His Gly Arg

660 665 670

Thr Val Ile Gln Leu Glu Asn Gly Lys Thr Leu Glu Ile Leu Ala Ala

675 680 685

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

690 695 700

Lys Pro Leu Ala Gly Pro Tyr Val Ser Gln Ala Gln Leu Leu Gly Gly
705 710 715 720
Gly Val Leu Ser Leu Ser Met Ser Asp Gln Pro Gly Gln Trp Ile Tyr
725 730 735
Glu Gly Ala Val Leu Thr Val Arg Ala Asp Gln Pro Ser Leu Val Asp
740 745 750
Gly Lys Thr Ser Thr Gly Trp Arg Ala Ala Ser Gly Gln Ser Val Thr

755 760 765
Phe Ser Leu Lys Ala Pro Ala Cys Ile Ala Ala Tyr Ser Val Ser Val
770 775 780
Gly Pro Asp Gln Ala Asp Pro Ser His Trp Thr Leu Gln Ala Tyr Asp
785 790 795 800
Gly Arg Ala Trp Val Ser Val Asp Gln Gln Ser Asn Val Val Phe Asp
805 810 815
His Arg His Ala Thr Arg Thr Phe Pro Leu Ala Pro Gly Arg Tyr Ala
820 825 830

Arg Leu Arg Trp Val Leu Asp Gly Gly Ser Glu Ala Ser Val Ser Glu
835 840 845
Val Glu Leu Ile Ala Gly Ala Ser Cys Ala Ala Pro Thr Ser Gly Ala
850 855 860
Pro Leu Leu
865

<210> 33
<211> 1937
<212> PRT
<213> Paenibacillus sp.
<400> 33

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

20 25 30
Glu Gly Ile Thr Thr Arg Asn Val Ala Ile Asn Ala Ala Ala Thr Ala

35 40 45

Ser Gly Gln Cys Asn Ala Asn Glu Ser Ala Ser Asn Ala Val Asp Gly

50 55 60

Lys Thr Asp Thr Lys Trp Cys Asp Asn Thr Ser Ala Gln Lys Lys Trp

65 70 75 80

Leu Lys Leu Asp Leu Gly Lys Glu Tyr Leu Val Asn Glu Trp Val Leu

85 90 95

Gln Asn Ala Ala Ile Asn Glu Ser Gly Asn Ser Pro Phe Trp Asn Thr

100 105 110

Lys Asn Phe Arg Leu Gln Lys Ser Asp Asp Gly Glu Thr Trp Thr Asp

115 120 125

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

130 135 140

Thr Pro Phe Thr Thr Arg Tyr Leu Arg Leu Tyr Ile Asp Lys Ala Ala

145 150 155 160

Tyr Asp Ser Asn Ile Ala Arg Ile Tyr Glu Leu Glu Val Tyr Gly Val

165 170 175

Glu Ala Asp Gln Ile Pro Ala Glu Pro Glu Thr Asn Leu Ala Pro Ile

180 185 190

Asp Tyr Val Asp Pro Phe Ile Asn Thr Leu Gly Asp Asn Gly Gln Thr

195 200 205

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

210 215 220

Ser Asp Gly Gly Ala Phe Ser Gly Tyr Tyr Tyr Glu Asn Lys Asn Leu

225 230 235 240

Lys Gly Phe Ser His Leu Arg Phe Ser Gly Val Gly Cys Ser Gly Gly

245 250 255

Gly Gly Asn Ile Leu Met Met Pro Glu Thr Arg Asp Phe Thr Lys Asn

260 265 270

Val Ala Asp Tyr Lys Gln Lys Tyr Asp Lys Ser Ser Glu Gln Ala Ser

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

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

340 345 350

Met Ile Lys Ser Gln Asn Val Cys Gly His Gly Tyr Tyr Thr Ile Tyr

355 360 365

Tyr Ser Ile Gln Phe Asp His Asp Phe Asp Ser Tyr Ser Ser Trp Gln

370 375 380

Gly Asp Ser Val Gly Ala Val Ala Gln Arg Ser Gly Ser Asn Ser Gly

385 390 395 400

Val Trp Leu Asn Phe Asn Thr Ala Gly Ser Lys Thr Val Gln Ala Lys

405 410 415

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

420 425 430

Leu Tyr Ser Asp Trp Asn Phe Asp Ala Arg His Glu Glu Ala Arg Ala

435 440 445

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

450 455 460

Gln Asn Lys Arg Val Phe Tyr Thr Gln Met Tyr His Ser Tyr Leu Ser

465 470 475 480

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

485 490 495

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

500 505 510

Tyr Asn Gly Trp Thr Thr Trp Asp Asp Phe Arg Lys Tyr Ala Met Phe

515 520 525

Ser Leu Phe Glu Pro Gln Arg Tyr Asn Asn Met Val Lys Ser Leu Val

530 535 540

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

595 600 605
Ile Ser Gly Lys Leu Glu Lys Ala Asn Ser Ala Ser Phe Pro Met Lys
610 615 620
Leu Ala Gln Leu Ile Gly Asp Lys Ala Thr Phe Glu Lys Tyr Lys Glu
625 630 635 640
Leu Ala Leu Ser Tyr Lys Lys Leu Trp Asn Pro Thr Gln Val Asp Glu
645 650 655
Lys Gly Thr Pro Thr Gly Phe Phe Thr Pro Asn Gly Thr Thr Val Gly
660 665 670

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

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

Tyr Val Phe Ser Ala Leu Gly Leu Phe Pro Gly Asn Pro Gly Glu Gly
805 810 815

Tyr Phe Leu Ile Gly Ser Pro Ile Phe Ser Glu Val Lys Leu His Met
820 825 830

Gly Ser Gly Lys Thr Leu Val Ile Lys Ala Asp Asn Val Ser Ser Glu
835 840 845

Asn Arg Phe Ile Gln Ser Ala Lys Leu Asn Gly Lys Asp Phe Asn Gln
850 855 860

Ser Trp Ile Lys Tyr Asp Asp Leu Met Ala Gly Gly Thr Leu Glu Phe
865 870 875 880

Gln Met Ser Ser Thr Pro Asn Met Ser Trp Gly Ala Lys Ala Ser Ala
885 890 895

Ala Pro Pro Thr Val Asp Tyr Asn Ala Asp Met Asp Asn Asp Phe Asn
900 905 910

His Glu Gln Leu Ile Pro Glu Lys Ser Thr Trp Lys Tyr Asp Asp Lys
915 920 925

Gly Lys Glu Ala Gly Glu Gly Trp Thr Gln Val Asp Phe Asp Asp Ser
930 935 940

Ser Trp Ser Ser Gly Lys Ala Met Leu Gly Tyr Asp Ser Tyr Gly Lys
945 950 955 960

Pro Ala Thr Thr Val Ser Tyr Gly Pro Asn Ala Asn Asn Lys Tyr Val
965 970 975

Thr Thr Tyr Phe Arg Lys Thr Phe Asp Ala Lys Asp Leu Asp Gly Ile
980 985 990

Leu Glu Leu Asp Gly Ser Leu Ile Arg Asp Asp Gly Ala Ile Val Tyr
995 1000 1005

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

Asn Tyr Ser Thr Phe Ala Asn Ala Thr Val Gly Asp Glu Arg Asp Lys
1025 1030 1035 1040

Asn Gly Phe Ile Ile Asp Pro Ser Tyr Leu Val Glu Gly Lys Asn Val

1045 1050 1055

Leu Thr Ala Glu Val His Gln Val Asn Ala Thr Ser Ser Asp Ile Ala

1060 1065 1070

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

1075 1080 1085

Pro Thr His Pro Val Val Asp Asp Lys Ala Asn Thr Ile Gly Trp Thr

1090 1095 1100

Pro Val Glu Gly Ile Asn Asn Ala Ser Asp Tyr Glu Phe Ser Thr Asp

1105 1110 1115 1120

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

1125 1130 1135

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

1140 1145 1150

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

1155 1160 1165

Tyr Thr Ser Asp Val Lys Trp Asp Val Tyr Asp Leu Asp Ala Asp Ile

1170 1175 1180

His Gln Asp Gly Asn Met Val Val Asp Val Thr Gly Thr Leu Lys Gly

1185 1190 1195 1200

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

1205 1210 1215

Glu His Ala Trp Val Ser Ser Ala Val Pro Val Gln Thr Gly Ser Phe

1220 1225 1230

Asp Ile Ser Gln Ile Tyr Asn Val Asp Ala Ser Lys Tyr Lys Val Asn

1235 1240 1245

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

1250 1255 1260

Trp Leu Ala Asp Pro Ile Val Gln Gln Ser Glu Pro Gly Ser Leu Pro

1265 1270 1275 1280

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

1285 1290 1295

Pro Leu Pro Asp Pro Lys Pro Asp Glu Pro Glu Glu Pro Glu Val Pro

1300 1305 1310

Glu Thr Gly Met Lys Ile Gln Phe Glu Asp Arg Ala Glu Trp Thr Ser

1315 1320 1325

Ala Ala His Pro Asn Gly Gly Gly Gly Leu Ser Thr Glu Ala Gly Asn

1330 1335 1340

Gly Gly Thr Val Val Ala His Thr Phe Gly Gly Ala Trp Leu Ala Tyr

1345 1350 1355 1360

Asn Val Asp Phe Gly Thr Thr Gly Tyr Asn Asn Val Thr Val Gln Tyr

1365 1370 1375

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

1380 1385 1390

Leu Gly Ser Val Ser Gly Glu Leu Val Gly Thr Val Asn Met Glu Asp

1395 1400 1405

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

1410 1415 1420

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

1425 1430 1435 1440

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

1445 1450 1455

Asp Tyr Glu Lys Ile Arg Ser Asp Tyr Ala Lys Leu Glu Leu Glu Ser

1460 1465 1470

Tyr Asp Glu Trp Thr Thr Asp Val Asn Thr Gly Asn Asn Asn Thr Pro

1475 1480 1485

Leu Lys Thr Glu Ala Gly Lys Gly Gly Val Gly Gln Gln Val Ala Asn

1490 1495 1500

Thr Phe Asn Gly Ala Trp Leu Ala Tyr Lys Arg Met Asp Phe Gly Ser

1505 1510 1515 1520

Glu Gly Val Asp Lys Phe Ser Ile Glu Tyr Ala Gly Asn Ser Thr Asn

1525 1530 1535

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

1540 1545 1550

Thr Leu Val Gly Thr Val Ala Thr Pro Pro Thr Ala Ala Ala Trp Gly
1555 1560 1565
Thr Tyr Ala Thr Val Ser Gly Ser Leu Thr Gln Lys Leu Thr Gly Leu
1570 1575 1580
Gln Asp Val Tyr Leu Val Phe Thr Gly Ser Ala Ala Asn Gly Glu Thr
1585 1590 1595 1600
Gly Lys Lys Tyr Ile Gly Asn Phe Asp Asn Ala Ser Phe Ser Leu Ser

1605 1610 1615
Val Gln Glu Pro Glu Glu Pro Glu Gln Pro Gln Gln Pro Glu Gln Glu
1620 1625 1630
Gln Ile Thr Val Gln Phe Glu Ser Lys Thr Glu Trp Asn Thr Ala Leu
1635 1640 1645
Asn Thr Phe Asn Asn Gln Ala Met Lys Ile Glu Asn Asn Asn Gly Gly
1650 1655 1660
Gln Thr Val Gly Asn Thr Tyr Thr Gly Ala Trp Leu Gly Phe Lys Asp
1665 1670 1675 1680

Val Asp Phe Gly Ser Glu Lys Gly Lys Asn Gln Val Ser Ile Val Tyr
1685 1690 1695
Asp Ala Pro Thr Asn Arg Val Pro Ala Asp Val Lys Ala Glu Ile Arg
1700 1705 1710
Leu Gly Ser Pro Thr Gly Thr Leu Val Gly Thr Val Ala Ile Pro Asn
1715 1720 1725
Thr Gly Ser Thr Trp Gly Gln Tyr Asn Thr Ala Thr Ala Asp Leu Asn
1730 1735 1740
Thr Thr Ile Lys Gly Lys Gln Asp Leu Tyr Ile Val Met Thr Gly Ser

1745 1750 1755 1760
Thr Thr Ser Ser Leu Leu Tyr Val Gly Asn Tyr Asp Ser Leu Thr Phe
1765 1770 1775
Gly Tyr Lys Pro Val Arg Ser Asp Tyr Ala Lys Leu Glu Leu Glu Ser
1780 1785 1790
Tyr Asp Glu Trp Thr Thr Ala Val Asn Pro Leu Asn Ser Asn Thr Pro
1795 1800 1805

Leu Lys Thr Glu Ala Gly Lys Gly Gly Ala Gly Lys Gln Val Ala Asn
1810 1815 1820

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

1890 1895 1900
Gln Asp Val Tyr Leu Val Leu Thr Gly Ser Ala Gly Asn Gly Glu Thr
1905 1910 1915 1920
Gly Lys Lys Tyr Ile Gly Asn Phe Asp Asn Ala Ala Phe Ser Leu Lys
1925 1930 1935
Val

<210> 34
<211> 751
<212> PRT
<213> Unknown
<220><223> Chitinophaga pinensis
<400> 34
Met Lys Arg Asn Arg Tyr Leu Ile Ser Thr Ala Ile Leu Val Leu Gly
1 5 10 15

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

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

85 90 95
Phe Met Pro Leu Ser Gly Glu Val Asn Thr Glu Asn Trp Lys Arg Ser
100 105 110
Arg Val Tyr Asp Arg Glu Lys Ile Ser Pro Cys Lys Tyr Ser Ala Tyr
115 120 125
Leu Asp Asp Thr Asp Glu Ile Thr Phe Ala Pro Ala Ala His Ser Gly
130 135 140
Tyr Phe Glu Ile Asp Phe Thr Gly Asn Thr Pro His Tyr Leu Arg Leu
145 150 155 160

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

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

Asp His Lys Val His Thr Ser Asp Lys Pro Phe Tyr Val Asp Asn Trp

305 310 315 320

Leu Trp Asp Thr Tyr Ile Ala Leu Glu Pro Leu Gln Thr Leu Leu Asn

325 330 335

Pro Glu Met Glu Ala Asp Lys Ile Arg Ser Tyr Val Asp Met Tyr Glu

340 345 350

Gln Ser Gly Trp Met Pro Ser Phe Ala Val Ala His Gly Asp Met Pro

355 360 365

Cys Met Thr Gly Asn His Ala Ala Ala Trp Met Ala Asp Ala Trp Phe

370 375 380

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

385 390 395 400

Lys Asn Ser Leu Gln Ala Thr Leu Leu Pro Trp Arg Asn Gly Pro Ala

405 410 415

Thr Ser Leu Asp Thr Phe Tyr Thr Glu His Gly Tyr Met Pro Ser Leu

420 425 430

Lys Pro Asp Glu Lys Glu Thr Val Lys Glu Val Asp Asp Phe Glu Arg

435 440 445

Arg Gln Ala Val Ala Val Thr Leu Glu Asn Ser Tyr Asp Asp Trp Cys

450 455 460

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

465 470 475 480

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

485 490 495

Gly Phe Met Trp Pro Lys Asp Ala Asp Gly Asn Trp Ile Glu Pro Phe

500 505 510

Asp Pro Lys Phe Ser Gly Gly Gln Gly Gly Arg Asp Tyr Phe Thr Glu

515 520 525

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

530 535 540

Leu Phe Asp Leu Met Gly Gly Lys Ala Lys Ala Glu Glu Lys Leu Asp

545 550 555 560

Gln Leu Phe Arg Glu Asn Leu Gly Arg Ser Lys Tyr Asn Leu Trp Tyr

565 570 575

Thr Phe Pro Asp Ala Thr Gly Met Val Gly Gln Phe Val Met Gly Asn

580 585 590

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

595 600 605

Trp Lys Thr Gln Lys Arg Ile Arg Met Leu Leu Asp Thr Trp Tyr Thr

610 615 620

Asp Asn Leu Phe Gly Ile Pro Gly Asp Glu Asp Gly Gly Gly Met Thr

625 630 635 640

Ala Phe Val Val Phe Ser Met Met Gly Phe Cys Pro Val Thr Pro Gly

645 650 655

Ile Pro Val Tyr Asn Ile Gly Ser Pro Val Phe Ser Glu Ile Thr Ile

660 665 670

Lys Leu Phe Ser Gly Lys Thr Phe Thr Ile Ser Ala Pro Gly Ser Ser

675 680 685

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

690 695 700

Asn Val Pro Trp Phe Thr His Glu Asp Leu Leu Lys Gly Gly Val Leu

705 710 715 720

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

725 730 735

Gln Ala Ala Pro Pro Ser Ser Leu Asn Tyr Ser Pro Ala Gly Lys

740 745 750

<210> 35

<211> 728

<212> PRT

<213> Unknown

<220><223> Pedobacter heparinus

<400> 35

Met Lys Asn Ser Ile Lys Leu Met Leu Leu Cys Leu Leu Leu Ser Gln

1 5 10 15

Ser Arg Leu Lys Ala Gln Glu Val Thr Ser Asn Leu Gln Tyr Val Asp

20 25 30

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

35 40 45

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

50 55 60

Gln Leu Asp Asp Gln Ile Ser Tyr Phe Pro Leu Asn Met Tyr Ser His

65 70 75 80

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

85 90 95

Glu Lys Ser Trp Lys Gln Arg Phe Thr Tyr Asp Leu Glu Lys Thr Ala

100 105 110

Pro His Tyr Tyr Thr Ala Val Leu Glu Glu Ser Gly Ile Lys Val Glu

115 120 125

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

130 135 140

Ala Ser Ala Asn Trp Leu Arg Leu Gly Val Val Asn Glu Thr Gly Glu

145 150 155 160

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

165 170 175

Gly Met Lys Ala Tyr Phe Tyr Gly Glu Leu Asn Ala Asp Val Thr Glu

180 185 190

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

195 200 205

Gly Asn Ser His Gly Val Glu Phe Arg Tyr Gly Ile Ser Tyr Ile Ser

210 215 220

Val Glu Gln Ala Lys Ile Asn Leu Glu Lys Glu Ile Pro Asn Trp Gly

225 230 235 240

Phe Glu Lys Val Lys Ser Thr Ala Lys Gln Val Trp Glu Glu Ala Leu

245 250 255

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

260 265 270

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

325 330 335
Ser Ser Tyr Val Asn Met Tyr Glu Gln Ser Gly Trp Met Pro Ser Phe
340 345 350
Ala Leu Val Phe Gly Asp Asn Pro Cys Met Thr Gly Asn His Ala Ala
355 360 365
Ala Trp Ile Thr Asp Ala Trp Phe Lys Gly Ile Arg Asn Phe Asn Val
370 375 380
Glu Lys Ala Tyr Ala Gly Leu Lys Lys Asn Ser Leu Glu Ala Thr Leu
385 390 395 400

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

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

Asp Val Lys His Asp Leu Glu Gly Leu Phe Lys Leu Met Gly Gly Lys
530 535 540

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

610 615 620
Asp Glu Asp Gly Gly Gly Met Ser Ala Phe Val Val Phe Ser Met Met
625 630 635 640
Gly Phe Tyr Pro Val Thr Pro Gly Ile Pro Val Tyr His Ile Gly Ser
645 650 655
Pro Val Phe Asn Lys Ile Ser Leu Lys Leu Lys Asn Gly Lys Thr Phe
660 665 670
Thr Val Val Ala Arg Asn Asn Ser Ser Thr Ala Lys Tyr Ile Gln Ser
675 680 685

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

<210> 36
<211> 2234
<212> PRT
<213> Unknown
<220><223> Streptomyces
<400> 36

Met Leu Val Met Pro Asp Arg Ser Lys Arg Pro Pro Ile Arg Ser Ser
1 5 10 15

Ser Pro Arg Ala Ala Leu Arg Ala Thr Val Ala Ala Val Leu Ala Gly
20 25 30

Ala Leu Gly Leu Ala Ala Leu Thr Gly Gly Gly Thr Ala Val Ala Val
35 40 45

Pro Val Thr Lys Ala Ser Pro Pro Ala Gly Glu Arg Ser Gly Gly Thr
50 55 60

Asp Tyr Thr Lys Leu Val Asp Pro Phe Val Ser Thr Ala Gly Asp Tyr
65 70 75 80

Gly Asn Asp Leu Pro Gly Ala Gln Ala Pro His Ser Leu Ala Lys Val

85 90 95

Asn Pro Met Thr Thr Pro Gly Arg Asn His Ser Gly Tyr Asp Tyr Asn
100 105 110

Glu Asp His Ile Ala Gly Phe Thr Ala Thr Asn Leu Asp Gly Val Gly
115 120 125

Gly Ser Gly Gly Gly Gly Asp Leu Leu Val Val Pro Thr Ser Gln Gln
130 135 140

Tyr Asp Lys Arg Pro Ala Thr Ser Thr Tyr Ala His Pro Tyr Ser His
145 150 155 160

Asp Asp Glu Ser Ala Thr Pro Gly Ser Tyr Arg Val Gly Leu Gly Ser
165 170 175

Pro Ser Gly Thr Ile Asp Ala Glu Met Thr Ala Thr Thr Arg Thr Ala
180 185 190

Leu Glu Arg Tyr Ala Phe Pro Ala Lys Ala Arg Pro Gln Leu Val Leu
195 200 205

Asp Leu Ala Asn Asn Phe Thr Ser Arg Thr Arg Ala Thr Leu Asp Ala
210 215 220

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

225 230 235 240

Ser Phe Asn Gly Ala Ser Tyr Arg Leu Tyr Tyr Tyr Ala Thr Thr Asn

245 250 255

Val Pro Val Thr Ser Leu Arg Thr Trp Gly Asp Asp Gly Ala Leu Gly

260 265 270

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

275 280 285

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

290 295 300

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

305 310 315 320

Arg Thr Phe Glu Glu Val Arg Ala Gln Thr Lys Ala Asp Trp Asn Arg

325 330 335

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

340 345 350

Ser Thr Leu Thr Lys Gln Phe Tyr Thr His Leu Tyr Arg Met Tyr Ala

355 360 365

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

370 375 380

Gly Ala Val His Lys Ala Asn Gly Phe Thr Tyr Tyr Asp Gly Trp Ser

385 390 395 400

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

405 410 415

Ala Thr Tyr Arg Asp Met Val Gln Ser Ala Val Ile Leu Phe Ala Asp

420 425 430

Ala His Ala Ala Gly Lys Ser Leu Gly Ser Leu Thr His Ser Val Pro

435 440 445

Thr Val Arg Trp Glu Arg Ser Ala Val Val Ile Ala Asp Ala Leu Ser

450 455 460

Lys Gly Phe Lys Asp Phe Asp Arg Leu Asp Glu Ala Tyr Pro Ala Leu

465 470 475 480

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

485 490 495

Ile Ala Gly Asp Pro Gly Thr Thr Val Gln Arg Gly Tyr Asp Gln Trp

500 505 510

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

515 520 525

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

530 535 540

Asp Ala Trp Thr Ala Ala Asp Gly Thr Ala Val Gly Leu Leu Thr Pro

545 550 555 560

Arg Asp Gly Glu Gly Asp Trp Gln Gly Val Asp Tyr Glu Lys Phe Glu

565 570 575

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

580 585 590

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

595 600 605

Gly Arg Ala Ala Ile Arg His Met Phe Gly Glu Asp Ser Asp Ala Asp

610 615 620

Asp Gly Ser Thr Met Leu His Ser Asn Ala Asn Glu Ile Asp Leu Gln

625 630 635 640

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

645 650 655

Trp Val Arg Ala Ile Tyr Thr Gly Glu Thr Trp Asn Arg Tyr Ile Ala

660 665 670

Thr Gly Ser Thr Asn Glu Ala Pro Ser Ser Gly Gly Glu Phe Arg Pro

675 680 685

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

690 695 700

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

705 710 715 720

Ala Leu Gly Leu Phe Pro Val Thr Ala Gly Ser Ser Gln Phe Gln Ile

725 730 735

Gly Ser Pro Phe Phe Asp Ser Thr Thr Ile Thr Tyr Pro Asn Gly Ala

740 745 750

Glu Phe Thr Val Glu Ala Asp Gly Val Ser Pro Lys Asn Tyr Tyr Val
755 760 765

Thr Val Asp Ala Ser Val Glu Leu Arg Leu Ser Gly Arg Ala Ser Phe
770 775 780

Ala Ala Arg Lys Gly Thr Ser Leu Thr Arg Thr Gly Ala Ala Ser Val
785 790 795 800

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

805 810 815

Arg Thr Ala Ser Leu Arg Leu Thr Gly Thr Thr Arg Thr Asp Ala Arg
820 825 830

Phe Gly Ile Thr Phe Arg Asp Arg Ala Phe Pro His Gly Ile Pro Ala
835 840 845

Ser Thr Val Thr Gly Thr Gly Val Ser Val Thr Asp Pro Leu Ile Val
850 855 860

Ser Ala Ala Ala Val His Arg Gly Ser Leu Ala Ala Leu Val Asp Glu
865 870 875 880

Ala Ser Leu Val Arg Glu Gly Asn Tyr Ser Asp Gly Ser Tyr Gly Ile

885 890 895

Phe Arg Thr Ala Leu Glu Arg Ala Arg Thr Val Leu Ala Asp Ser Ala
900 905 910

Ser Pro Thr Gly Thr Leu Met Ala Ala His Asp Ala Leu Arg Ser Ala
915 920 925

Val Asp Ala Leu Thr Leu Asp Glu Gly Gly Tyr Ala Val Leu Gln Ala
930 935 940

Glu Asp Pro Asp Arg Met Glu Gly Pro Ser Leu Val Lys Glu Ala Tyr

945 950 955 960

Tyr Ser Asp Gly Asp Leu Gly Gly Val Thr Glu Gly Ala Trp Glu Gln

965 970 975

Tyr Thr Asp Leu Asp Phe Gly Gly Val Pro Pro Arg Ser Val Ser Val
980 985 990

Arg Tyr Ala Asn Ser Gln Ala Ala Ala Ala Glu Pro Ser Ser Val Asp
995 1000 1005

Ile His Ala Gly Asp Ala Asp Gly Pro Val Val Ala Thr Val Ser Leu
1010 1015 1020

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

1090 1095 1100
Leu Lys Val Ser Gly Gly Val Phe Glu Asn Val Thr Asn Gly Ala Trp
1105 1110 1115 1120
Ala Glu Trp Arg Asp Thr Asp Leu Gly Asp Gly Ala Asp Thr Val Thr
1125 1130 1135
Val Ser Tyr Asp Lys Pro Arg Ser Arg Ala Ala Ser Asp Ser His Ile
1140 1145 1150
Glu Leu Arg Pro Gly Ala Lys Asp Gly Pro Thr Ala Val Thr Val Pro
1155 1160 1165

Leu Asp Tyr Thr Gly Ser Gly Trp Gly Thr Val Ala Ser Thr Ser Val
1170 1175 1180
Arg Leu Asp Pro Asp Val Phe Glu Gly Thr Gln Asp Val Tyr Ala Val
1185 1190 1195 1200
Phe Val Ser Ser Thr Gln Thr Asp Ala Gln Pro Tyr Val Ala Asn Val
1205 1210 1215
His Ser Leu Thr Leu Thr Arg Gln Ala Asp Ala Pro Val Val Phe Asp
1220 1225 1230
Ala Thr Ala Phe Glu Gly Ser Ser Gly Gly Gly Leu Lys Ser Glu Pro

1235 1240 1245
Ala Thr Trp Ser Gly Ala Gly Ser Ala Thr Ser Leu Gly Gly Thr Tyr

1250 1255 1260

Asp Gly Ala Trp Leu Asp Tyr Gly Asp Val Asp Phe Gly Asp Ser Pro

1265 1270 1275 1280

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

1285 1290 1295

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

1300 1305 1310

Pro Gly Thr Pro Tyr Ala Thr Val Pro Leu Pro Val Thr Gly Ser Ser

1315 1320 1325

Trp Ser Ser Gly Gly Thr Thr Ser Leu Thr Leu Pro Glu Ala Ile Thr

1330 1335 1340

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

1345 1350 1355 1360

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

1365 1370 1375

Ala Pro Ala Gln Thr Asp Leu Ser Ala Leu Arg Lys Ala Ile Glu Gln

1380 1385 1390

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

1395 1400 1405

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

1410 1415 1420

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

1425 1430 1435 1440

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

1445 1450 1455

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

1460 1465 1470

Ser Trp Lys Ala Phe Thr Thr Ala Leu Thr Ala Ala Lys Thr Ala Val

1475 1480 1485

Ala Asp Glu Thr Ala Thr Asp Arg Thr Leu Thr Glu Arg Tyr Ala Ala

1490 1495 1500

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

1505 1510 1515 1520

Ala Ala Pro Gly Ala Val Ser Ala Ala Pro Ser Gly Thr Ser Val Gln

1525 1530 1535

Val Thr Trp Ser Ala Pro Glu Asp Asp Gly Gly Ser Pro Val Thr Gly

1540 1545 1550

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

1555 1560 1565

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

1570 1575 1580

Thr Ala Arg Val Arg Ala Val Asn Ala Leu Gly Asp Ser Pro Trp Ser

1585 1590 1595 1600

Ala Arg Thr Gln Pro Ala Val Thr Gly Asp Asn Arg Pro Gln Ala Pro

1605 1610 1615

Thr Val Thr Gly Val Val Thr Asp Gly Glu Arg Val Arg Val Asn Trp

1620 1625 1630

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

1635 1640 1645

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

1650 1655 1660

Val Leu Thr Ala Ala Gly Gly Ala Lys Ala His Thr Ala Thr Val Thr

1665 1670 1675 1680

Ala Val Thr Arg Ala Gly Ser Ser Asp Gly Ser Gly Ala Thr Val Ser

1685 1690 1695

Thr Ala Pro Ala Thr Ser Thr Thr Ser Ala Thr Ser Ala Gly Asp Pro

1700 1705 1710

Ala Glu Tyr Glu Pro Ser Pro Phe Pro Gly Asp Thr Leu Asp Ala Thr

1715 1720 1725

Tyr Ala Ser Asp Ala Trp Pro Glu Thr Gly Asp Gly Ser Asp Trp Phe

1730 1735 1740

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

1745 1750 1755 1760

Ala Asn Ser Glu Val Pro Ala Gly Thr Pro Leu Gly Ala Glu Asn Asp
1765 1770 1775
Arg Ile Thr Val Arg Val Asn Asn Ala Ala Thr Gln Gln Gln Val Asp
1780 1785 1790
Arg Ala Glu Val Asp Ala Ser Asn Ser Ala Thr Val Thr Met Ala Asp
1795 1800 1805
Gly Leu Gly Ser Arg Leu Gly Pro Leu Tyr Gly Glu Ala Leu Lys Glu

1810 1815 1820
Gly Arg Leu Pro Lys Thr Ser Ala Leu Phe Ser Arg Val Asn Glu Asn
1825 1830 1835 1840
Leu Asp Thr His Asp Ala Ala Lys Asn His Tyr Gln Tyr Leu Arg Pro
1845 1850 1855
Tyr Val Arg Leu Gly Phe Ala Gly Asp Gly Gly Ala Val Tyr Glu Ser
1860 1865 1870
Gln Asp Ser Ser Tyr Ser Gly Leu Ala Gly Gln Gly Ser Tyr Pro Ser
1875 1880 1885

Gly His Thr Tyr Gly Gly Tyr Glu Ala Gly Thr Ile Leu Ala Thr Leu
1890 1895 1900
Leu Pro Asp Leu Ala Pro Ser Ile Leu Ala Arg Thr Ser Glu Tyr Gly
1905 1910 1915 1920
Asp Asn Arg Ile Val Leu Gly Phe His Tyr Pro Leu Asp Val Met Gly
1925 1930 1935
Gly Arg Ile Thr Ala Gln Ala Thr Val Ala His Arg Trp Ala Asp Pro
1940 1945 1950
Glu Phe Ala Lys Leu Leu Gly Gln Ala His Thr Glu Ile Glu Asn Val

1955 1960 1965
Leu Leu Ala Arg Cys Glu Glu Glu Gly Tyr Gly Asp Thr Leu Thr Ala
1970 1975 1980
Cys Ala Gly Asp Pro Tyr Ala Gly Leu Ser Thr Ala Gln Gln Val Asp
1985 1990 1995 2000
Arg Tyr Thr Gln Arg Leu Thr Tyr Gly Phe Ser Arg Thr Gly Glu Ala
2005 2010 2015

Gly Gln Ala Leu Asp Ala Pro Ser Asp Ala Ala Ala Leu Leu Ile Thr
2020 2025 2030

Ala Phe Pro Asp Leu Thr Ala Glu Gln Arg Thr Gln Val Leu Glu Gln
2035 2040 2045
Thr Ala Thr Asp Ser Gly Tyr Pro Leu Asp Leu Thr Gly Ser Gly Gly
2050 2055 2060
Pro Gly Trp Gln Arg Ile Asn Leu Ala Ala Ala Met Ala Ala Asp Val
2065 2070 2075 2080
Val Val Asn Ala Asp Gly Ser Val Thr Val Thr Asn Phe Pro Asp Ala
2085 2090 2095
Thr Ala Ala Ser Ala Ala Glu Ala Val Ala Ile Thr Val Gly Gly Val

2100 2105 2110
Ala Leu Asp Gly Phe Asp Pro Asp Val Ser Thr Tyr Val Val Asp Trp
2115 2120 2125
Pro Arg Asn Gly Gly Arg Ile Pro Ala Val Gly Ala Val Thr Ala Ala
2130 2135 2140
Ser Gly Ala Arg Val Lys Val Thr Ser Gly Ser Ser Thr Val Ser Ser
2145 2150 2155 2160
Ser Gln Arg Gly Phe Ser Thr Arg Thr Leu Thr Val Thr Ser Ala Asp
2165 2170 2175

Gly Glu Phe Thr Arg Thr Tyr Thr Val Gly Phe Arg Pro Val Glu Gln
2180 2185 2190
His Pro His Arg Pro Gly Ala Leu Arg Asp Thr Gly Gly Gly Gly Thr
2195 2200 2205
Ala Gly Gly Ser Gly Gly Gly Gly Asp Val Gly Gly Gly Leu Trp Ser
2210 2215 2220
Pro Ala Arg Glu Trp Glu Gln Thr Val Asn
2225 2230
<210> 37
<211> 2206
<212> PRT

<213> Streptomyces lividans

<400> 37

Met Leu Ala Gly Ala Leu Gly Leu Ala Ala Leu Thr Gly Gly Gly Thr
1 5 10 15

Ala Val Ala Val Pro Val Thr Lys Ala Ser Pro Pro Ala Gly Glu Arg
20 25 30

Ser Gly Gly Thr Asp Tyr Thr Lys Leu Val Asp Pro Phe Val Ser Thr
35 40 45

Ala Gly Asp Tyr Gly Asn Asp Leu Pro Gly Ala Gln Ala Pro His Ser
50 55 60

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

65 70 75 80

Tyr Asp Tyr Asn Glu Asp His Ile Ala Gly Phe Thr Ala Thr Asn Leu
85 90 95

Asp Gly Val Gly Gly Ser Gly Gly Gly Gly Asp Leu Leu Val Val Pro
100 105 110

Thr Ser Gln Gln Tyr Asp Lys Arg Pro Ala Thr Ser Thr Tyr Ala His
115 120 125

Pro Tyr Ser His Asp Asp Glu Ser Ala Thr Pro Gly Ser Tyr Arg Val
130 135 140

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

Thr Arg Thr Ala Leu Glu Arg Tyr Ala Phe Pro Ala Lys Ala Arg Pro
165 170 175

Gln Leu Val Leu Asp Leu Ala Asn Asn Phe Thr Ser Arg Thr Arg Ala
180 185 190

Thr Leu Asp Ala Thr Arg Leu Lys Asp Gly Thr Thr Ala Ile Ser Gly
195 200 205

Leu Val Ala Gly Ser Phe Asn Gly Ala Ser Tyr Arg Leu Tyr Tyr Tyr

210 215 220

Ala Thr Thr Asn Val Pro Val Thr Ser Leu Arg Thr Trp Gly Asp Asp

225 230 235 240

Gly Ala Leu Gly Asp Ala Thr Ala Arg Asp Gly Thr Asp Thr Gly Ala

245 250 255

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

260 265 270

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

275 280 285

Glu Val Ala Gly Arg Thr Phe Glu Glu Val Arg Ala Gln Thr Lys Ala

290 295 300

Asp Trp Asn Arg Thr Leu Gly Ala Val Ala Val Lys Ala Ser Lys Lys

305 310 315 320

Ala Asp Pro Asp Ser Thr Leu Thr Lys Gln Phe Tyr Thr His Leu Tyr

325 330 335

Arg Met Tyr Ala Leu Pro Val Asn Ala Thr Ser Thr Ser Gly Thr Tyr

340 345 350

Arg Gly Val Asp Gly Ala Val His Lys Ala Asn Gly Phe Thr Tyr Tyr

355 360 365

Asp Gly Trp Ser Thr Trp Asp Asp Phe Arg Lys Tyr Ser Val Ala Ala

370 375 380

Tyr Ile Asp Pro Ala Thr Tyr Arg Asp Met Val Gln Ser Ala Val Ile

385 390 395 400

Leu Phe Ala Asp Ala His Ala Ala Gly Lys Ser Leu Gly Ser Leu Thr

405 410 415

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

420 425 430

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

435 440 445

Tyr Pro Ala Leu Lys Ser Tyr Thr Gly Tyr Tyr Thr Gly Thr Gln Leu

450 455 460

Arg Gln Gly Tyr Ile Ala Gly Asp Pro Gly Thr Thr Val Gln Arg Gly

465 470 475 480

Tyr Asp Gln Trp Ala Leu Ser Val Val Ala Asp Ala Leu Gly Glu Asp

485 490 495

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

500 505 510

Leu Val Lys Pro Asp Ala Trp Thr Ala Ala Asp Gly Thr Ala Val Gly

515 520 525

Leu Leu Thr Pro Arg Asp Gly Glu Gly Asp Trp Gln Gly Val Asp Tyr

530 535 540

Glu Lys Phe Glu Glu Ala Arg Leu Tyr Gln Gly Thr Leu Trp Gln Tyr

545 550 555 560

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

565 570 575

Gly Glu Gln Ala Gly Arg Ala Ala Ile Arg His Met Phe Gly Glu Asp

580 585 590

Ser Asp Ala Asp Asp Gly Ser Thr Met Leu His Ser Asn Ala Asn Glu

595 600 605

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

610 615 620

Leu Thr Gln Lys Trp Val Arg Ala Ile Tyr Thr Gly Glu Thr Trp Asn

625 630 635 640

Arg Tyr Ile Ala Thr Gly Ser Thr Asn Glu Ala Pro Ser Ser Gly Gly

645 650 655

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

660 665 670

Gly Phe Leu Pro Thr Met Asp Asn Asp Ala Gly Thr Met Ser Thr Met

675 680 685

Phe Val Ala Ala Ala Leu Gly Leu Phe Pro Val Thr Ala Gly Ser Ser

690 695 700

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

705 710 715 720

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

725 730 735

Asn Tyr Tyr Val Thr Val Asp Ala Ser Val Glu Leu Arg Leu Ser Gly
740 745 750

Arg Ala Ser Phe Ala Ala Arg Lys Gly Thr Ser Leu Thr Arg Thr Gly
755 760 765

Ala Ala Ser Val Thr Gly Leu Pro Asp Gly Val Thr Ala Asp Leu Arg
770 775 780

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

785 790 795 800

Thr Asp Ala Arg Phe Gly Ile Thr Phe Arg Asp Arg Ala Phe Pro His
805 810 815

Gly Ile Pro Ala Ser Thr Val Thr Gly Thr Gly Val Ser Val Thr Asp
820 825 830

Pro Leu Ile Val Ser Ala Ala Ala Val His Arg Gly Ser Leu Ala Ala
835 840 845

Leu Val Asp Glu Ala Ser Leu Val Arg Glu Gly Asn Tyr Ser Asp Gly
850 855 860

Ser Tyr Gly Ile Phe Arg Thr Ala Leu Glu Arg Ala Arg Thr Val Leu
865 870 875 880

Ala Asp Ser Ala Ser Pro Thr Gly Thr Leu Met Ala Ala His Asp Ala
885 890 895

Leu Arg Ser Ala Val Asp Ala Leu Thr Leu Asp Glu Gly Gly Tyr Ala
900 905 910

Val Leu Gln Ala Glu Asp Pro Asp Arg Met Glu Gly Pro Ser Leu Val
915 920 925

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

930 935 940

Ala Trp Glu Gln Tyr Thr Asp Leu Asp Phe Gly Gly Val Pro Pro Arg
945 950 955 960

Ser Val Ser Val Arg Tyr Ala Asn Ser Gln Ala Ala Ala Ala Glu Pro
965 970 975

Ser Ser Val Asp Ile His Ala Gly Asp Ala Asp Gly Pro Val Val Ala
980 985 990

Thr Val Ser Leu Pro Gly Thr Gly Gly Trp Gln Tyr Tyr Thr Thr Val
995 1000 1005

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

1075 1080 1085
Asn Gly Ala Trp Ala Glu Trp Arg Asp Thr Asp Leu Gly Asp Gly Ala
1090 1095 1100
Asp Thr Val Thr Val Ser Tyr Asp Lys Pro Arg Ser Arg Ala Ala Ser
1105 1110 1115 1120
Asp Ser His Ile Glu Leu Arg Pro Gly Ala Lys Asp Gly Pro Thr Ala
1125 1130 1135
Val Thr Val Pro Leu Asp Tyr Thr Gly Ser Gly Trp Gly Thr Val Ala
1140 1145 1150

Ser Thr Ser Val Arg Leu Asp Pro Asp Val Phe Glu Gly Thr Gln Asp
1155 1160 1165
Val Tyr Ala Val Phe Val Ser Ser Thr Gln Thr Asp Ala Gln Pro Tyr
1170 1175 1180
Val Ala Asn Val His Ser Leu Thr Leu Thr Arg Gln Ala Asp Ala Pro
1185 1190 1195 1200
Val Val Phe Asp Ala Thr Ala Phe Glu Gly Ser Ser Gly Gly Gly Leu
1205 1210 1215
Lys Ser Glu Pro Ala Thr Trp Ser Gly Ala Gly Ser Ala Thr Ser Leu

1220 1225 1230
Gly Gly Thr Tyr Asp Gly Ala Trp Leu Asp Tyr Gly Asp Val Asp Phe

1235 1240 1245

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

1250 1255 1260

Ala Arg Cys Gly Thr Gly Ser Ala Val Gln Leu Tyr Leu Asp Ser Phe

1265 1270 1275 1280

Asp Pro Asp Ala Pro Gly Thr Pro Tyr Ala Thr Val Pro Leu Pro Val

1285 1290 1295

Thr Gly Ser Ser Trp Ser Ser Gly Gly Thr Thr Ser Leu Thr Leu Pro

1300 1305 1310

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

1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360

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

1365 1370 1375

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

1380 1385 1390

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

1395 1400 1405

Arg Ser Leu Thr Leu Ala Ala Asn Gln Leu Val Pro Leu Pro Arg Leu

1410 1415 1420

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

1425 1430 1435 1440

Tyr Thr Asp Ala Ser Trp Lys Ala Phe Thr Thr Ala Leu Thr Ala Ala

1445 1450 1455

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

1460 1465 1470

Arg Tyr Ala Ala Leu Asp Arg Ala Arg Ser Ser Leu Thr Thr Lys Arg

1475 1480 1485

Arg Thr Val Pro Ala Ala Pro Gly Ala Val Ser Ala Ala Pro Ser Gly

1490 1495 1500

Thr Ser Val Gln Val Thr Trp Ser Ala Pro Glu Asp Asp Gly Gly Ser

1505 1510 1515 1520

Pro Val Thr Gly Tyr Glu Ile Thr Leu Ser Gly Gly Arg Gln Val Glu

1525 1530 1535

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

1540 1545 1550

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

1555 1560 1565

Ser Pro Trp Ser Ala Arg Thr Gln Pro Ala Val Thr Gly Asp Asn Arg

1570 1575 1580

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

1585 1590 1595 1600

Arg Val Asn Trp Arg Pro Ala Gly Asp Gly Gly Phe Pro Val Val Gly

1605 1610 1615

Tyr Thr Val Ala Leu Asp Asp Gly Thr Thr Ala His Val Pro Gly Thr

1620 1625 1630

Thr Ser Thr Ala Val Leu Thr Ala Ala Gly Gly Ala Lys Ala His Thr

1635 1640 1645

Ala Thr Val Thr Ala Val Thr Arg Ala Gly Ser Ser Asp Gly Ser Gly

1650 1655 1660

Ala Thr Val Ser Thr Ala Pro Ala Thr Ser Thr Thr Ser Ala Thr Ser

1665 1670 1675 1680

Ala Gly Asp Pro Ala Glu Tyr Glu Pro Ser Pro Phe Pro Gly Asp Thr

1685 1690 1695

Leu Asp Ala Thr Tyr Ala Ser Asp Ala Trp Pro Glu Thr Gly Asp Gly

1700 1705 1710

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

1715 1720 1725

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

1730 1735 1740

Ala Glu Asn Asp Arg Ile Thr Val Arg Val Asn Asn Ala Ala Thr Gln
1745 1750 1755 1760
Gln Gln Val Asp Arg Ala Glu Val Asp Ala Ser Asn Ser Ala Thr Val
1765 1770 1775
Thr Met Ala Asp Gly Leu Gly Ser Arg Leu Gly Pro Leu Tyr Gly Glu
1780 1785 1790
Ala Leu Lys Glu Gly Arg Leu Pro Lys Thr Ser Ala Leu Phe Ser Arg

1795 1800 1805
Val Asn Glu Asn Leu Asp Thr His Asp Ala Ala Lys Asn His Tyr Gln
1810 1815 1820
Tyr Leu Arg Pro Tyr Val Arg Leu Gly Phe Ala Gly Asp Gly Gly Ala
1825 1830 1835 1840
Val Tyr Glu Ser Gln Asp Ser Ser Tyr Ser Gly Leu Ala Gly Gln Gly
1845 1850 1855
Ser Tyr Pro Ser Gly His Thr Tyr Gly Gly Tyr Glu Ala Gly Thr Ile
1860 1865 1870

Leu Ala Thr Leu Leu Pro Asp Leu Ala Pro Ser Ile Leu Ala Arg Thr
1875 1880 1885
Ser Glu Tyr Gly Asp Asn Arg Ile Val Leu Gly Phe His Tyr Pro Leu
1890 1895 1900
Asp Val Met Gly Gly Arg Ile Thr Ala Gln Ala Thr Val Ala His Arg
1905 1910 1915 1920
Trp Ala Asp Pro Glu Phe Ala Lys Leu Leu Gly Gln Ala His Thr Glu
1925 1930 1935
Ile Glu Asn Val Leu Leu Ala Arg Cys Glu Glu Glu Gly Tyr Gly Asp

1940 1945 1950
Thr Leu Thr Ala Cys Ala Gly Asp Pro Tyr Ala Gly Leu Ser Thr Ala
1955 1960 1965
Gln Gln Val Asp Arg Tyr Thr Gln Arg Leu Thr Tyr Gly Phe Ser Arg
1970 1975 1980
Thr Gly Glu Ala Gly Gln Ala Leu Asp Ala Pro Ser Asp Ala Ala Ala
1985 1990 1995 2000

Leu Leu Ile Thr Ala Phe Pro Asp Leu Thr Ala Glu Gln Arg Thr Gln
2005 2010 2015

Val Leu Glu Gln Thr Ala Thr Asp Ser Gly Tyr Pro Leu Asp Leu Thr
2020 2025 2030
Gly Ser Gly Gly Pro Gly Trp Gln Arg Ile Asn Leu Ala Ala Ala Met
2035 2040 2045
Ala Ala Asp Val Val Val Asn Ala Asp Gly Ser Val Thr Val Thr Asn
2050 2055 2060
Phe Pro Asp Ala Thr Ala Ala Ser Ala Ala Glu Ala Val Ala Ile Thr
2065 2070 2075 2080
Val Gly Gly Val Ala Leu Asp Gly Phe Asp Pro Asp Val Ser Thr Tyr

2085 2090 2095
Val Val Asp Trp Pro Arg Asn Gly Gly Arg Ile Pro Ala Val Gly Ala
2100 2105 2110
Val Thr Ala Ala Ser Gly Ala Arg Val Lys Val Thr Ser Gly Ser Ser
2115 2120 2125
Thr Val Ser Ser Ser Gln Arg Gly Phe Ser Thr Arg Thr Leu Thr Val
2130 2135 2140
Thr Ser Ala Asp Gly Glu Phe Thr Arg Thr Tyr Thr Val Gly Phe Arg
2145 2150 2155 2160

Pro Val Glu Gln His Pro His Arg Pro Gly Ala Leu Arg Asp Thr Gly
2165 2170 2175
Gly Gly Gly Thr Ala Gly Gly Ser Gly Gly Gly Gly Asp Val Gly Gly
2180 2185 2190
Gly Leu Trp Ser Pro Ala Arg Glu Trp Glu Gln Thr Val Asn
2195 2200 2205

<210> 38
<211> 1856
<212> PRT
<213> Paenibacillus sp.
<400> 38

Met Phe Lys Lys Leu Phe Ala Val Ala Met Thr Val Met Cys Leu Thr

1 5 10 15

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

20 25 30

Glu Gly Ile Thr Thr Arg Asn Val Ala Ile Asn Ala Ala Ala Thr Ala

35 40 45

Ser Gly Gln Cys Asn Ala Asn Glu Ser Ala Ser Asn Ala Val Asp Gly

50 55 60

Lys Thr Asp Thr Lys Trp Cys Asp Asn Thr Ser Ala Gln Lys Lys Trp

65 70 75 80

Leu Lys Leu Asp Leu Gly Lys Glu Tyr Leu Val Asn Glu Trp Val Leu

85 90 95

Gln Asn Ala Ala Ile Asn Glu Ser Gly Asn Ser Pro Phe Trp Asn Thr

100 105 110

Lys Asn Phe Arg Leu Gln Lys Ser Asp Asp Gly Glu Thr Trp Thr Asp

115 120 125

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

130 135 140

Thr Pro Phe Thr Thr Arg Tyr Leu Arg Leu Tyr Ile Asp Lys Ala Ala

145 150 155 160

Tyr Asp Ser Asn Ile Ala Arg Ile Tyr Glu Leu Glu Val Tyr Gly Val

165 170 175

Glu Ala Asp Gln Ile Pro Ala Glu Pro Glu Thr Asn Leu Ala Pro Ile

180 185 190

Asp Tyr Val Asp Pro Phe Ile Asn Thr Leu Gly Asp Asn Gly Gln Thr

195 200 205

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

210 215 220

Ser Asp Gly Gly Ala Phe Ser Gly Tyr Tyr Tyr Glu Asn Lys Asn Leu

225 230 235 240

Lys Gly Phe Ser His Leu Arg Phe Ser Gly Val Gly Cys Ser Gly Gly

245 250 255

Gly Gly Asn Ile Leu Met Met Pro Glu Thr Arg Asp Phe Thr Lys Asn

260 265 270

Val Ala Asp Tyr Lys Gln Lys Tyr Asp Lys Ser Ser Glu Gln Ala Ser

275 280 285

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

290 295 300

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

305 310 315 320

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

325 330 335

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

340 345 350

Met Ile Lys Ser Gln Asn Val Cys Gly His Gly Tyr Tyr Thr Ile Tyr

355 360 365

Tyr Ser Ile Gln Phe Asp His Asp Phe Asp Ser Tyr Ser Ser Trp Gln

370 375 380

Gly Asp Ser Val Gly Ala Val Ala Gln Arg Ser Gly Ser Asn Ser Gly

385 390 395 400

Val Trp Leu Asn Phe Asn Thr Ala Gly Ser Lys Thr Val Gln Ala Lys

405 410 415

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

420 425 430

Leu Tyr Ser Asp Trp Asn Phe Asp Ala Arg His Glu Glu Ala Arg Ala

435 440 445

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

450 455 460

Gln Asn Lys Arg Val Phe Tyr Thr Gln Met Tyr His Ser Tyr Leu Ser

465 470 475 480

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

485 490 495

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

500 505 510

Tyr Asn Gly Trp Thr Thr Trp Asp Asp Phe Arg Lys Tyr Ala Met Phe
515 520 525
Ser Leu Phe Glu Pro Gln Arg Tyr Asn Asn Met Val Lys Ser Leu Val
530 535 540
Asp Leu Tyr Asn Thr Arg Gly Thr Tyr Thr Gln Trp Gly Asp Gly Tyr
545 550 555 560
Trp Pro Ser Pro Thr Val Arg Asn Glu Phe Asn Gly Gln Val Ile Leu
565 570 575
Asp Ala Tyr Ala Lys Gly Phe Gln Asp Phe Asp Val Tyr Lys Ala Leu

580 585 590
Lys Gly Met Ala Val Asp Ala Asp Asn Phe Ser Ile Ser Asp Gly Glu
595 600 605
Ile Ser Gly Lys Leu Glu Lys Ala Asn Ser Ala Ser Phe Pro Met Lys
610 615 620
Leu Ala Gln Leu Ile Gly Asp Lys Ala Thr Phe Glu Lys Tyr Lys Glu
625 630 635 640
Leu Ala Leu Ser Tyr Lys Lys Leu Trp Asn Pro Thr Gln Val Asp Glu
645 650 655

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

725 730 735
Tyr Leu Phe Asn Tyr Leu Gly Tyr Pro Tyr Leu Thr Gln Tyr Tyr Ala
740 745 750

Arg Glu Phe Thr Thr Glu Val Val Thr Gln Lys Tyr His Asn His Gly
755 760 765

Ala Tyr Ala Tyr Pro Leu Lys Ser Arg Val Tyr Arg Asp Asp Pro Glu
770 775 780

Gly Tyr Leu Ser Ser Met Asp Asp Asp Ala Gly Gly Met Ser Ser Trp
785 790 795 800

Tyr Val Phe Ser Ala Leu Gly Leu Phe Pro Gly Asn Pro Gly Glu Gly
805 810 815

Tyr Phe Leu Ile Gly Ser Pro Ile Phe Ser Glu Val Lys Leu His Met
820 825 830

Gly Ser Gly Lys Thr Leu Val Ile Lys Ala Asp Asn Val Ser Ser Glu
835 840 845

Asn Arg Phe Ile Glu Gly Trp Thr Gln Val Asp Phe Asp Asp Ser Ser
850 855 860

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

865 870 875 880

Ala Thr Thr Val Ser Tyr Gly Pro Asn Ala Asn Asn Lys Tyr Val Thr
885 890 895

Thr Tyr Phe Arg Lys Thr Phe Asp Ala Lys Asp Leu Asp Gly Ile Leu
900 905 910

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

Asn Gly His Glu Ile Phe Arg Thr Asn Met Pro Thr Gly Ala Val Asn
930 935 940

Tyr Ser Thr Phe Ala Asn Ala Thr Val Gly Asp Glu Arg Asp Lys Asn
945 950 955 960

Gly Phe Ile Ile Asp Pro Ser Tyr Leu Val Glu Gly Lys Asn Val Leu
965 970 975

Thr Ala Glu Val His Gln Val Asn Ala Thr Ser Ser Asp Ile Ala Phe
980 985 990

Glu Phe Ser Leu Glu Ala Val Arg Lys Leu Asn Ile Pro Ala Ala Pro
995 1000 1005

Thr His Pro Val Val Asp Asp Lys Ala Asn Thr Ile Gly Trp Thr Pro

1010 1015 1020

Val Glu Gly Ile Asn Asn Ala Ser Asp Tyr Glu Phe Ser Thr Asp Gly

1025 1030 1035 1040

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

1045 1050 1055

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

1060 1065 1070

Ala Ala Asn Arg Ala Ala Gly Glu Ala Leu Leu Ser Thr Glu Ala Tyr

1075 1080 1085

Thr Ser Asp Val Lys Trp Asp Val Tyr Asp Leu Asp Ala Asp Ile His

1090 1095 1100

Gln Asp Gly Asn Met Val Val Asp Val Thr Gly Thr Leu Lys Gly Asp

1105 1110 1115 1120

Tyr Thr Asp Ser Ala Val Val Val Phe Gln Leu Met Asp Gly Lys Glu

1125 1130 1135

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

1140 1145 1150

Ile Ser Gln Ile Tyr Asn Val Asp Ala Ser Lys Tyr Lys Val Asn Val

1155 1160 1165

Tyr Leu Val Asn Glu Phe Asn Gly Asp Ile Tyr Glu Ser Pro Leu Trp

1170 1175 1180

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

1185 1190 1195 1200

Pro Glu Gly Pro Pro Val Thr Glu Glu Pro Leu Pro Glu Pro Ile Pro

1205 1210 1215

Leu Pro Asp Pro Lys Pro Asp Glu Pro Glu Glu Pro Glu Val Pro Glu

1220 1225 1230

Thr Gly Met Lys Ile Gln Phe Glu Asp Arg Ala Glu Trp Thr Ser Ala

1235 1240 1245

Ala His Pro Asn Gly Gly Gly Gly Leu Ser Thr Glu Ala Gly Asn Gly

1250 1255 1260

Gly Thr Val Val Ala His Thr Phe Gly Gly Ala Trp Leu Ala Tyr Asn

1265 1270 1275 1280

Val Asp Phe Gly Thr Thr Gly Tyr Asn Asn Val Thr Val Gln Tyr Asp

1285 1290 1295

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

1300 1305 1310

Gly Ser Val Ser Gly Glu Leu Val Gly Thr Val Asn Met Glu Asp Lys

1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360

Pro Asn Asn Leu Pro Tyr Ile Gly Asn Phe Asp Trp Phe Lys Phe Asp

1365 1370 1375

Tyr Glu Lys Ile Arg Ser Asp Tyr Ala Lys Leu Glu Leu Glu Ser Tyr

1380 1385 1390

Asp Glu Trp Thr Thr Asp Val Asn Thr Gly Asn Asn Asn Thr Pro Leu

1395 1400 1405

Lys Thr Glu Ala Gly Lys Gly Gly Val Gly Gln Gln Val Ala Asn Thr

1410 1415 1420

Phe Asn Gly Ala Trp Leu Ala Tyr Lys Arg Met Asp Phe Gly Ser Glu

1425 1430 1435 1440

Gly Val Asp Lys Phe Ser Ile Glu Tyr Ala Gly Asn Ser Thr Asn Thr

1445 1450 1455

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

1460 1465 1470

Leu Val Gly Thr Val Ala Thr Pro Pro Thr Ala Ala Ala Trp Gly Thr

1475 1480 1485

Tyr Ala Thr Val Ser Gly Ser Leu Thr Gln Lys Leu Thr Gly Leu Gln

1490 1495 1500

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

1505 1510 1515 1520

Lys Lys Tyr Ile Gly Asn Phe Asp Asn Ala Ser Phe Ser Leu Ser Val

1525 1530 1535

Gln Glu Pro Glu Glu Pro Glu Gln Pro Gln Gln Pro Glu Gln Glu Gln

1540 1545 1550

Ile Thr Val Gln Phe Glu Ser Lys Thr Glu Trp Asn Thr Ala Leu Asn

1555 1560 1565

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

1570 1575 1580

Thr Val Gly Asn Thr Tyr Thr Gly Ala Trp Leu Gly Phe Lys Asp Val

1585 1590 1595 1600

Asp Phe Gly Ser Glu Lys Gly Lys Asn Gln Val Ser Ile Val Tyr Asp

1605 1610 1615

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

1620 1625 1630

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

1635 1640 1645

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

1650 1655 1660

Thr Ile Lys Gly Lys Gln Asp Leu Tyr Ile Val Met Thr Gly Ser Thr

1665 1670 1675 1680

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

1685 1690 1695

Tyr Lys Pro Val Arg Ser Asp Tyr Ala Lys Leu Glu Leu Glu Ser Tyr

1700 1705 1710

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

1715 1720 1725

Lys Thr Glu Ala Gly Lys Gly Gly Ala Gly Lys Gln Val Ala Asn Thr

1730 1735 1740

Phe Asn Gly Ala Trp Leu Ala Tyr Lys Arg Met Asp Phe Gly Thr Glu

1745 1750 1755 1760

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

Tyr Asp Thr Val Ser Gly Thr Leu Thr Gln Lys Leu Thr Gly Ile Gln
1810 1815 1820
Asp Val Tyr Leu Val Leu Thr Gly Ser Ala Gly Asn Gly Glu Thr Gly
1825 1830 1835 1840
Lys Lys Tyr Ile Gly Asn Phe Asp Asn Ala Ala Phe Ser Leu Lys Val
1845 1850 1855

<210> 39
<211> 640
<212> PRT
<213> Bacteroides sp.
<400> 39

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

Asn Arg Leu Ser Gly Val Gly Cys Ser Gly Gly Gly Gly Asn Leu Arg
85 90 95
Ile Arg Pro Val Ala Pro Ser Gln Glu Leu His Ile Lys Lys Ser Arg
100 105 110
Glu Lys Ala Thr Pro Gly Tyr Tyr Ser Thr Ala Phe Thr Asn Gly Ile

115 120 125

Lys Thr Glu Leu Thr Ala Thr Asn Ala Met Ala Val Glu Arg Tyr Lys

130 135 140

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

145 150 155 160

Phe Glu Asp Val Ala Thr Cys His Tyr Lys Arg Ile Ser Glu Thr Cys

165 170 175

Ile Glu Gly Tyr Val Gln Ala Lys Asn Val Cys Gly His Gly Arg Tyr

180 185 190

Lys Leu Tyr Phe Ser Leu Asn Thr Ser His Pro Phe Gln Leu Glu Glu

195 200 205

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

210 215 220

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

225 230 235 240

Trp Glu Cys Ala Arg Trp Glu Lys Met Asp Phe Glu Asp Val Lys Ser

245 250 255

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

260 265 270

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

275 280 285

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

290 295 300

Gly Thr Asp Gly Lys Val Tyr Ile Ser Glu Asp Phe Arg Tyr Tyr Ser

305 310 315 320

Asn Trp Ser Leu Trp Asp Thr Phe Arg Thr Lys Phe Pro Leu Leu Val

325 330 335

Leu Thr Glu Pro Ala Lys Met Arg Asp Met Ala Thr Ser Leu Ile His

340 345 350

Leu Tyr Ala Thr Gly Lys Lys Asp Trp Ser Thr Gly Phe Glu Ser Thr

355 360 365

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

435 440 445
Phe Glu Glu Thr Trp Lys Lys Glu Phe Met Asn Val Thr Pro Ala Phe
450 455 460
Glu Val Met Lys Asn Asn Gly Leu Tyr Gln Gly Thr Arg Trp Gln Tyr
465 470 475 480
Arg Trp Ala Ala Pro Gln Tyr Ile Asp Lys Met Ile Glu Trp Val Gly
485 490 495
Gln Asp Ser Leu Arg Ser Gln Leu Thr Tyr Phe Phe Asp His His Leu
500 505 510

Tyr Asn Gln Gly Asn Glu Pro Asp Ile His Val Pro Tyr Leu Phe Asn
515 520 525
Arg Leu Gly Ala Pro Glu Lys Thr Gln Gln Ile Val Arg Ser Leu Met
530 535 540
Thr Glu Pro Met Ile His Lys Tyr Gly Gly Asn Ser Glu Phe Lys Thr
545 550 555 560
Pro Tyr Leu Gly Lys Ala Phe Lys Asn Ala Pro Glu Gly Tyr Ser Pro
565 570 575
Glu Met Asp Glu Asp Asp Gly Thr Met Ser Ala Trp Tyr Val Phe Gly

580 585 590
Ala Met Gly Phe Tyr Pro Leu Leu Val Gly Asp Glu Tyr Tyr Asp Leu
595 600 605
Thr Ser Pro Leu Phe Asp Arg Val Leu Leu Arg Leu Thr Asn Gly Asn
610 615 620

Val Leu Thr Ile Gln Thr Glu Gly Arg Lys Lys Lys Asp Ala Pro Ile
625 630 635 640

<210> 40

<211> 624

<212> PRT

<213> Unknown

<220><223> Bacteroides cellulosilyticus

<400> 40

Met Val Gln Ala Gln Gln Thr Ser Phe Val Asn Pro Phe Ile Gly Thr
1 5 10 15

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

Met Ile Lys Pro Gly Pro Glu Thr Ile Pro Arg Gly Asn Gly Gly Tyr
35 40 45

Asp Tyr Gln Ser Gln Gln Leu Lys Gly Phe Ser Gln Thr Arg Met Ser
50 55 60

Gly Val Gly Cys Ile Gly Val Gly Gly Asn Leu Leu Ile Thr Pro Phe

65 70 75 80

Val Gly Thr Ala Cys Lys Thr Leu Lys Met Asp Lys Ala Ser Glu Thr
85 90 95

Ala Ile Pro Gly Tyr Tyr Ser Val Thr Leu Asp Asn Gln Leu Lys Val
100 105 110

Glu Ile Thr Thr Gly Arg Thr Ala Ala Ile Tyr Arg Phe Thr Tyr Pro
115 120 125

Ala Thr Glu Thr Ala Gly Ile Lys Ile Asn Phe Lys His Ser Tyr Gly
130 135 140

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

Gly Phe Val Arg Ser Ala Cys Thr Cys Asp Leu Gly Ser Tyr Lys Phe
165 170 175

Tyr Tyr Tyr Ile Glu Lys Asp Lys Ser Thr Tyr Glu Pro Glu Asp Asn
180 185 190

Asp Ser Glu Leu Leu Trp Lys Phe Gln Thr Glu Pro Asn Glu Gln Ile
195 200 205
Ile Leu Lys Ile Gly Leu Ser Ser Val Ser Ala Glu Glu Ala Glu Ala

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

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

355 360 365
Gln Lys Gly Met Phe Asp Gly Ser Leu Asp Glu Leu Leu Pro Leu Met
370 375 380
Leu Lys Glu Ala Glu Asp Ile Ser Asn Asp Ser Pro Asp Lys Ala Leu
385 390 395 400
Glu Arg Gly Tyr Asp Phe Trp Gly Val Ser Glu Leu Ala Gly Lys Met
405 410 415
Gly Asn Lys Glu Leu Lys Lys Glu Phe Ser Leu Arg Ser Lys Glu Tyr
420 425 430

Arg Pro Ile Trp Leu Gln Lys Phe Lys Asp Ile Gly Pro Thr Ser Asp

435 440 445

Ile Met His Gly Asp Gly Leu Tyr Glu Gly Thr Ile Trp Gln Tyr Arg

450 455 460

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

465 470 475 480

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

485 490 495

Phe Asn Met Gly Asn Gln Pro Asp Ile His Val Pro Phe Leu Tyr Tyr

500 505 510

Tyr Leu Gly Ala Pro Trp Lys Thr Gln Lys Leu Val His Gln Ile Leu

515 520 525

Leu Glu Pro Thr Thr Asn Tyr Tyr Gly Thr His Glu Lys Trp Glu Lys

530 535 540

Pro Tyr Ile Gly Lys Ile Phe Asn Thr Thr Pro Gln Gly Tyr Leu Lys

545 550 555 560

Glu Met Asp Asp Asp Ala Gly Thr Met Ser Ser Trp Phe Val Leu Ser

565 570 575

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

580 585 590

Asn Ala Pro Val Phe Asp Thr Val Thr Leu His Pro Thr Ser Gln Gln

595 600 605

Glu Phe Lys Ile Tyr Val Asn Arg Pro Asp Ala Glu Cys Ile Tyr Ile

610 615 620

<210> 41

<211> 631

<212> PRT

<213> Unknown

<220><223> Leeuwenhoekiella blandensis

<400> 41

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

1 5 10 15

Ser Val Ala Gln Glu Ser Leu Thr Asp Tyr Val Asn Val Phe Leu Gly

20 25 30

Thr Ser Gly Asp His Gly Gln Met Ser Pro Ser Ala Thr Thr Pro Phe

35 40 45

Asn Met Met Asn Leu Gly Pro Gln Thr Asn Pro His Gln His Thr Gly

50 55 60

Tyr Glu Tyr Tyr Ala Lys Gln Phe Asp Gly Phe Thr His Thr Arg Met

65 70 75 80

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

85 90 95

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

100 105 110

Ala Val Pro Gly Phe Tyr His Val Asn Phe Glu Asn Gly Ile Asp Ala

115 120 125

Lys Leu Thr Val Ala Gln Asn Leu Gly Met His Gln Tyr Ser Phe Pro

130 135 140

Asn Ala Asn Ser Gly Leu Tyr Ile Asp Leu Ser Phe Ala Leu Ser Asn

145 150 155 160

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

165 170 175

Tyr Met Asp Thr Lys Thr Thr Cys His Ala Gly Thr Tyr Arg Ile Tyr

180 185 190

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

195 200 205

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

210 215 220

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

225 230 235 240

Asp Asp Phe Glu Thr Leu Lys Gly Lys Ala Ser Ala Ser Trp Asn Ala

245 250 255

Tyr Leu Asn Thr Ile Glu Val Ser Gly Glu Lys Asp Arg Leu Gln Leu

260 265 270

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

275 280 285

Glu Glu Asp Gly Thr Phe Arg Ala Thr Asp Gly Ser Val Gln Asn Glu

290 295 300

Asp Tyr Lys Val Tyr Asn Gly Trp Ala Ile Trp Asp Asn Tyr Arg Glu

305 310 315 320

Gln Leu Pro Phe Leu Ser Leu Ala His Pro Glu Val Tyr Gln Asp Ile

325 330 335

Thr Thr Ser Ile Ala Asn Leu Tyr Arg Phe Gly Lys Lys Asn Trp Ala

340 345 350

Thr Glu His Glu Thr Ser Pro Thr Val Arg Thr Glu His Ala Met Val

355 360 365

Val Leu Leu Asp Ala Tyr Lys Lys Gly Tyr Asp Val Asp Phe Glu Ala

370 375 380

Ile Lys Asp Ser Leu Leu Tyr Glu Ala Asp His Leu Asp Phe Gly Ala

385 390 395 400

Pro Asp Lys Ala Leu Glu Ser Ser Tyr Asp Leu Trp Ala Met Ser Glu

405 410 415

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

420 425 430

Lys Ala Leu Asp Tyr Lys Ala Tyr Trp Asp Lys Asp Phe Lys Asp Leu

435 440 445

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

450 455 460

Thr Ile Trp Gln Tyr Arg Trp Phe Val Pro Trp Asp Val Asn Gly Leu

465 470 475 480

Gln Glu Leu Ala Gly Gly Ala Glu Thr Phe Glu Asn Gln Leu Asp Gln

485 490 495

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

500 505 510

Val Pro Gly Leu Tyr Asn Ala Thr Ser Gln Pro Trp Lys Ser Gln Lys

515 520 525

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

580 585 590
Gly Ser Pro Val Tyr Tyr Leu Thr Ala Pro Ile Phe Lys Glu Val Lys
595 600 605
Leu Asn Leu Ala Pro Gly Lys Thr Phe Ser Ile Ser Val Lys Asn Tyr
610 615 620
Asn Lys Glu His Phe Tyr Val
625 630

<210> 42
<211> 646
<212> PRT
<213> Sphingobacterium spiritivorum
<400> 42

Met Gln Val Leu Thr Asp Leu Leu Leu Val Ser Met Lys Asn Phe Leu

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

Glu Tyr Leu Ala Lys Glu Ile Leu Gly Phe Thr His Asn Arg Phe Glu
85 90 95
Gly Val Gly Cys Lys Gly Ser Gly Gly Leu Ile Leu Val Lys Pro Phe

100 105 110

Leu Gly Gly Gln Asp Asp Gln Gln Pro Leu Leu Lys Val Thr Glu Gln

115 120 125

Ala Gly Pro Gly Phe Tyr Glu Ile Gly Leu Lys Asn Arg Ile Lys Ala

130 135 140

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

145 150 155 160

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

165 170 175

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

180 185 190

Gly Ser Val Arg Ala Arg Thr Thr Cys Gly Val Gly Ile Tyr Thr Ile

195 200 205

His Tyr Ala Ile Lys Val Ser Ser Ser His Val Trp Glu Gly Lys Gly

210 215 220

Asn Gln Leu Thr Leu His Leu Glu Glu Asn Ser Glu Lys Val Thr Ile

225 230 235 240

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

245 250 255

Leu His Asn Ala Asp Arg Ser Tyr Ser Glu Val Arg Asn Ala Ser Gln

260 265 270

Gln Gln Trp Asp Ile Cys Leu Ser His Ile Glu Val Lys Gly Asp Pro

275 280 285

Glu Arg Glu Lys Leu Phe Tyr Ser Leu Phe Tyr Arg Thr Leu Gln Ser

290 295 300

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

305 310 315 320

Lys Met His Ser Ala Lys Gly Lys Arg Tyr His Gly Trp Ala Ile Trp

325 330 335

Asp Asn Tyr Lys Thr Gln Leu Pro Leu Leu Glu Leu Leu Tyr Pro Gln

340 345 350

Leu Tyr Gln Asp Val Val Ser Ser Ile Ser Asp Leu Tyr Arg Tyr Gly

355 360 365

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

435 440 445
Arg Ala Gly Ser Tyr Lys Thr Val Trp Gln Lys Glu Phe Lys Asp Leu
450 455 460
Ser Arg Asn Asp Val Asp Arg Met Ser Ala Arg Asn Met Tyr Gln Gly
465 470 475 480
Thr Ile Arg Gln Tyr Arg Trp Asn Val Pro Phe Asp Ile Gly Gly Leu
485 490 495
Val Asp Leu Ala Gly Gly Lys Lys Ala Leu Thr Glu Gln Leu Asp Glu
500 505 510

Phe Phe Asp Glu His Tyr Phe Asn Arg Ala Asn Glu Pro Asp Met Gln
515 520 525
Ser Pro Thr Leu Tyr Tyr Ala Ser Asp Lys Pro Trp Lys Tyr Gln Ser
530 535 540
Leu Val His Gln Leu Ala Val Asp Thr Val Ile Gln Tyr Tyr Phe Asn
545 550 555 560
Asp Asn Ser Arg Gly Ile Asp Pro Phe Ile Asp Arg Ile Tyr Lys Asn
565 570 575
Glu Ser Lys Ala Tyr Ile Arg Thr Met Asp Asp Asp Ala Gly Ala Met

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

Gly His Pro Val Tyr Tyr Leu Ser Val Pro Leu Phe Pro Glu Ile Asn
610 615 620
Leu Arg Arg Ala Asp Asn Thr Leu Gln Ile Arg Val His Asn Phe Gly
625 630 635 640
Ser Gln Asn Lys Tyr Ile
645

<210> 43
<211> 642
<212> PRT

<213> Unknown
<220><223> Pedobacter sp.
<400> 43
Met Phe Phe Met Leu Lys Met Asn Leu Arg Leu Leu Ala Phe Phe Leu
1 5 10 15
Leu Leu Leu Val Asn Lys Thr Ile Tyr Ala Gln Gln Thr Gly Lys Asp
20 25 30
Lys Leu Val Asn Val Phe Leu Gly Ser Ser Gly Asp His Gly Gln Met
35 40 45
Ser Pro Ala Ala Ser Tyr Pro Phe Ser Ala Leu Ser Ile Ala Pro Gln
50 55 60

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

130 135 140
Thr Ala Met His Asp Tyr Ile Met Pro Ala Gly Lys Lys Gly Phe Ser
145 150 155 160

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

Ser Val Lys Trp Lys Asp Leu Gly Glu His Arg Leu Val Ala Val Pro
210 215 220
Ala Asp Asp Ala Thr Ala Val Gln Leu Arg Val Asp Ile Ser Ala Val
225 230 235 240
Ser Thr Glu Tyr Ala Val Lys Ala Ser Lys Lys Lys Leu Ser Phe Glu
245 250 255
Glu Ala Lys Lys Asn Ser Ala Met Ala Trp Asp Glu Leu Leu Ser Ser
260 265 270
Val Asp Val Lys Gly Asn Ala Glu Arg Glu Arg Leu Phe Tyr Ser Leu
275 280 285
Leu Tyr Arg Thr Ile Gln Ser Pro Tyr Thr Ile Ser Glu Pro Asp Gly
290 295 300
Thr Tyr Lys Ala Ile Asn Gly Ser Thr Gln Lys Ser Lys Glu Met Arg
305 310 315 320
Tyr Asn Gly Trp Ala Ile Trp Asp Asn Tyr Lys Thr Gln Leu Pro Leu
325 330 335
Leu Ser Val Leu Tyr Pro Lys Arg Tyr Gly Asp Ile Val Gly Ser Ile
340 345 350

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

Leu Glu Ala Ser Tyr Asp Leu Trp Ala Val Ser Gly Met Phe Lys Leu

420 425 430

Ser Gly Asp Ala Ala Met Ser Glu Lys Tyr Lys Thr Met Ala Met Asp

435 440 445

Tyr Lys Lys Tyr Trp Asp Lys Asp Phe Lys Asp Leu Ser Lys Lys Asp

450 455 460

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

465 470 475 480

Tyr Arg Trp Ala Val Pro Phe Asp Val Lys Gly Leu Val Glu Leu Thr

485 490 495

Gly Gly Ala Gln Ala Phe Thr Glu Gln Leu Asp Asp Phe Phe Asp Asn

500 505 510

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

515 520 525

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

530 535 540

Leu Ala Leu Asp Thr Val Ile Gln His Tyr Phe Asn Asp Asn Ser Arg

545 550 555 560

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

565 570 575

Phe Val Arg Thr Met Asp Asp Asp Ala Gly Ala Met Ser Gly Trp Phe

580 585 590

Val Leu Thr Ala Met Gly Ile Gln Pro Ala Cys Val Gly Thr Pro Ile

595 600 605

Tyr Tyr Leu Asn Val Pro Leu Phe Glu Ser Val Thr Ile Lys Ser Gly

610 615 620

Ala Lys Pro Leu Gln Ile Lys Val Glu His Phe Ser Asp Gln Asn Val

625 630 635 640

Tyr Ile

<210> 44

<211> 810

<212> PRT

<213> Magnaporthe grisea

<400> 44

Met Ala Leu Ile His Ile Phe Phe Leu Leu Ser Ser Leu Val Ser Leu
1 5 10 15

Ala Gln Gln Ala Glu Leu Ala Asp Tyr Val Phe Thr Asn Thr Gly Ser
20 25 30

Glu Gly Gly Gly Asn Thr Phe Pro Gly Val Ser Glu Pro Phe Gly Met
35 40 45

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

50 55 60

Gly Tyr Leu Pro Asn Gly Lys Ile Lys Ala Phe Ser Met Leu His Leu
65 70 75 80

Ser Gly Thr Gly Gly Ala Pro Lys Tyr Gly Val Val Ala Gln Met Pro
85 90 95

Val Val Gly Asn Ile Ala Asn Pro Leu Asp Asp Tyr Leu Ala Ala Arg
100 105 110

Asp Ala Pro Asp Met Thr Glu Val Gly Tyr Tyr Lys Ser Arg Leu Ala
115 120 125

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

Gln Tyr Thr Phe Pro Ser Gly Pro Thr Pro Arg Asn Val Ile Val Asp
145 150 155 160

Val Ser His Val Leu Pro Ser Tyr Arg Gly Gln Gly Leu Gly Gln Arg
165 170 175

Tyr Leu Gly Gly Ser Ile Ser Val Asp Arg Asp Asp Ser Gly Asn Val
180 185 190

Gln Tyr Lys Gly Ser Gly Ser Tyr Asp Asn Gly Trp Asn Arg Ala Pro

195 200 205

Lys Trp Thr Val Tyr Phe Cys Gly Ala Phe Asn Ser Ser Ala Thr Phe
210 215 220

Lys Thr Phe Val Gly Thr Asn Ala Thr Ala Asn Thr Leu Ser Lys Phe
225 230 235 240
Ser Asn Asp Asn Lys Val Glu Ser Leu Ser Arg Leu Gly Ala Val Phe
245 250 255
Thr Phe Asp Ala Ala Asn Val Val Ser Arg Val Gly Val Ser Phe Ile
260 265 270

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

340 345 350
Asp Ile Phe Thr Leu Trp Asp Leu Phe Arg Cys Thr Thr Ala Leu Leu
355 360 365
His Val Phe Gln Pro Lys Val Tyr Glu Glu Phe Ile Arg Ser Leu Ile
370 375 380
Asp Thr Trp Arg His Glu Gly Tyr Leu Pro Asp Ala Arg Ser Ser Phe
385 390 395 400
Phe Asn Gly Ala Thr Gln Gly Gly Ser Asn Ala Asp Thr Val Leu Ala
405 410 415

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

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

485 490 495

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

500 505 510

Arg Ser Arg Gln Trp Arg Ser His Trp Asn Lys Asp Met Lys Ala Leu

515 520 525

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

530 535 540

Glu Gln Asp Pro Leu Asn Cys Arg Gly Cys Tyr Trp Gly Asp Asn Tyr

545 550 555 560

Tyr Glu Ala Leu Pro Trp Glu Tyr Thr Phe Gly Pro His His Asp Ile

565 570 575

Ser Thr Leu Ile Asp Tyr Ser Gly Gly Pro Arg Arg Phe Ala Ser Arg

580 585 590

Leu Gln Trp Thr Phe Glu Pro Asn Val Arg Pro Lys Gly His Glu Arg

595 600 605

Phe Asn Arg Met Ile Phe Asp Pro Gly Asn Glu Pro Ser Phe Thr Thr

610 615 620

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

625 630 635 640

Thr Arg Tyr Leu Gly Lys Thr Tyr Tyr Gly Val Arg Pro Asn Gly Leu

645 650 655

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

660 665 670

Met Leu Gly Leu Tyr Pro Met Thr Gly Gln Thr Thr Phe Leu Ile Gly

675 680 685

Ser Pro Trp Leu Asp Asp Ile Thr Ile Ser Leu Gly Asp Gly Lys Ser

690 695 700

Leu Gln Ile Thr Ser Thr Gly Gly Ser Glu Asp Ser Phe Tyr Val Pro

705 710 715 720

Val Ala Val Thr Val Pro Ile Val Val Val Ala Ala Phe Ala Ile Cys

725 730 735

Val Thr Phe Phe Phe Ile Arg Arg Arg Arg Ala Ala Ala Ala Gln Lys

740 745 750

Ala Leu Ser Ser Gly Ser Gly Thr Pro Glu Ser Gly Ile Glu Thr Leu

755 760 765

Thr Pro Thr Ser Gln Pro Val Asp Thr Ser Lys Thr Asn Val Gln Val

770 775 780

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

785 790 795 800

Ala Gln Leu Pro Pro Pro Val Ala Asn Arg

805 810

<210> 45

<211> 701

<212> PRT

<213> Unknown

<220><223> Bacteroides sp.

<400> 45

Met Lys Ile Leu His Phe Cys Ala Ala Ile Thr Met Ala Ala Met Leu

1 5 10 15

Ser Gly Cys Asn Gly Gly Gln Ser Gln Thr Ala Asn Arg Thr Pro Val

20 25 30

Asp Tyr Val Asn Pro Tyr Ile Gly Asn Ile Ser His Leu Leu Val Pro

35 40 45

Thr Phe Pro Thr Ile Gln Leu Pro Asn Ser Met Leu Arg Val Tyr Pro

50 55 60

Glu Arg Ala Asp Tyr Thr Thr Glu Leu Leu Asn Gly Leu Pro Leu Ile

65 70 75 80

Val Thr Asn His Arg Glu Arg Ser Ala Phe Asn Leu Ser Pro Tyr Gln

85 90 95

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

100 105 110

Leu Thr Pro Tyr Ser Tyr Glu Val Asp Leu Asn Asp Asn Ser Met Lys

115 120 125

Ala Glu Tyr Ala Leu Ser His Gln Ser Ala Leu Tyr Arg Ile Thr Phe

130 135 140

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

145 150 155 160

Ile His Val Gly Glu Asn Phe Ile Ser Gly His Gln Gln Leu Ser Ala

165 170 175

Asn Thr Asn Val Tyr Val Tyr Ile Glu Pro Gln Glu Lys Pro Val Ser

180 185 190

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

195 200 205

Glu Gly Ile Asn Ala Cys Ala Ala Trp Arg Phe Ala Asp Gly Thr Thr

210 215 220

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

225 230 235 240

Glu Lys Asn Met Arg Asn Glu Leu Lys Asp Tyr Asn Ile Lys Asn Leu

245 250 255

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

260 265 270

Val Glu Gly Gly Thr Glu Asp Asp Lys Thr Val Leu Tyr Ser Ser Phe

275 280 285

Tyr Arg Thr Phe Glu Arg Pro Ile Cys Met Ser Glu Ala Gly Gly Arg

290 295 300

Tyr Phe Ser Ala Phe Asp Gly Glu Val His Asp Asp Asn Gly Thr Pro

305 310 315 320

Phe Tyr Asn Asp Asp Trp Ile Trp Asp Thr Tyr Arg Ala Ala His Pro

325 330 335

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

340 345 350

Phe Leu Leu Met Ala Glu Gln Met Gly Thr Met Trp Met Pro Thr Phe

355 360 365

Pro Glu Val Thr Gly Asp Ser Arg Arg Met Asn Ser Asn His Ala Val

370 375 380

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

450 455 460
Thr Ser Tyr Asp Gln Trp Cys Leu Ser Arg Ile Ala Glu Ile Leu Gly
465 470 475 480
Lys Lys Asp Glu Ala Ala His Tyr Leu Gln Cys Ser Tyr Asn Tyr Arg
485 490 495
Asn Leu Phe Asn Lys Glu Thr Gly Phe Phe His Pro Lys Asp Lys Glu
500 505 510
Gly Asn Trp Ile Thr Pro Phe Asp Tyr Arg Tyr Ala Gly Gly Met Gly
515 520 525

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

595 600 605
Tyr Asn Tyr Ala Gly Gln Pro Trp Lys Thr Gln Lys Arg Ile Arg Gln
610 615 620

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

Leu Phe Thr His Ala Glu Ile Thr Leu Ser Asn Gly Ser Val Phe Glu
675 680 685
Ile Glu Ala Pro Asn Val Ser Glu Glu Asn Lys Tyr Ile
690 695 700

<210> 46
<211> 701
<212> PRT
<213> Unknown
<220><223> Bacteroides sp.
<400> 46

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

Gly Lys Glu Leu Arg Pro Ile Ile Thr Tyr Asn Tyr Asp Asn Glu His
100 105 110
Leu Thr Pro Tyr Ser Tyr Glu Val Asp Leu Asn Asp Asn Ser Met Lys
115 120 125

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

165 170 175
Asn Thr Asn Val Tyr Val Tyr Ile Glu Pro Gln Glu Lys Pro Val Ser
180 185 190
Thr Gly Ile Leu Lys Asp Gly Val Ile Glu Ala Ser Lys Asp Asn Ala
195 200 205
Glu Gly Ile Asn Ala Cys Ala Ala Trp Arg Phe Ala Asp Gly Thr Thr
210 215 220
Thr Val Ser Leu Arg Tyr Gly Ile Ser Phe Ile Ser Glu Glu Gln Ala
225 230 235 240

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

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

370 375 380

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

450 455 460
Thr Ser Tyr Asp Gln Trp Cys Leu Ser Arg Ile Ala Glu Ile Leu Gly
465 470 475 480
Lys Lys Asp Glu Ala Ala His Tyr Leu Gln Cys Ser Tyr Asn Tyr Arg
485 490 495
Asn Leu Phe Asn Lys Glu Thr Gly Phe Phe His Pro Lys Asp Lys Glu
500 505 510
Gly Asn Trp Ile Thr Pro Phe Asp Tyr Arg Tyr Ala Gly Gly Met Gly
515 520 525

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

595 600 605
Tyr Asn Tyr Ala Gly Gln Pro Trp Lys Thr Gln Lys Arg Ile Arg Gln
610 615 620

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

Leu Phe Thr His Ala Glu Ile Thr Leu Ser Asn Gly Ser Val Phe Glu
675 680 685
Ile Glu Ala Pro Asn Val Ser Glu Glu Asn Lys Tyr Ile
690 695 700

<210> 47
<211> 706
<212> PRT
<213> Unknown
<220><223> Bacteroides sp.
<400> 47

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

Pro Tyr Gln Gly Asp Glu Ser Gly Leu Arg Pro Val Ile Ala Tyr Ser
100 105 110
Tyr Asp Arg Glu Lys Ile Leu Pro Tyr Arg Tyr Gln Val Tyr Leu Asp
115 120 125

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

165 170 175
Gln Tyr Ile Asp Lys Lys Thr Lys Val Phe Leu Tyr Ala Glu Thr Asn
180 185 190
Gln Lys Pro Val Lys Ala Gly Val Leu Ser Asn Gly Ser Val Lys Tyr
195 200 205
Asp Glu Thr Ser Val Glu Gly Thr Asn Ala Ala Ile Ala Leu Ser Phe
210 215 220
Gly Glu Asp Val Lys Lys Leu Gly Val Arg Tyr Gly Ile Ser Phe Ile
225 230 235 240

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

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

370 375 380

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

450 455 460
Ala Val Thr Leu Gly Thr Ser Tyr Asp Glu Trp Cys Leu Ser Gln Ile
465 470 475 480
Ala Gln Glu Leu Gly Lys Lys Asp Glu Ala Asp Tyr Tyr Leu Arg Arg
485 490 495
Ser Tyr Asn Tyr Arg Asn Val Phe Asn Pro Glu Thr Gly Phe Phe His
500 505 510
Pro Lys Asp Lys Asp Gly Lys Phe Ile Tyr Pro Leu Asp Tyr Arg Tyr
515 520 525

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

595 600 605
His Ile Pro Tyr Leu Tyr Asn Tyr Ala Gly Lys Pro Trp Met Thr Gln
610 615 620

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

Asn Ile Gly Ser Pro Met Phe Thr Asp Val Lys Val Asp Met Gly Asn
675 680 685
Gly Asn Thr Phe Glu Ile Arg Ala Asn Asn Ala Ser Asp Glu Asn Lys
690 695 700
Tyr Val
705

<210> 48
<211> 691
<212> PRT
<213> Unknown
<220><223> Bacteroides sp.
<400> 48
Met Ala Ser Ser Cys Val Tyr Ser Pro Pro Lys Glu Pro Val Asp Tyr
1 5 10 15
Val Asn Pro Asn Ile Gly Thr Ile Gly His Leu Leu Val Ala Thr Ala

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

Gly Phe Asp His Asp Phe Glu Lys Val Thr Pro Tyr Tyr Ser Trp Ile
100 105 110

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

165 170 175
Ile Phe Ser Lys Pro Phe Arg Ser Tyr Val Thr Trp Lys Asp Ser Leu
180 185 190
Ile Ser Gln Asp Val Lys Gln Glu Gly Leu Asp Ile Gly Ala Leu Val
195 200 205
Thr Phe Asp Thr Ser Gln Asp Glu Glu Ile Met Val Lys Ile Gly Val
210 215 220
Ser Phe Ile Asp Met Glu Gln Ala Lys Arg Asn Leu Glu Met Glu Ile
225 230 235 240

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

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

His Glu Val Ala Thr Ile Thr Asp Thr Tyr Met Lys Gly Phe Arg Asp
370 375 380

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

450 455 460
Ala Lys Glu Leu Gly Lys Glu Asp Asp Tyr Gln Tyr Tyr Leu Asn Arg
465 470 475 480
Ala Tyr Asn Tyr Arg Asn Val Phe Asp Ser Lys Ser Gly Phe Met Arg
485 490 495
Pro Lys Thr Ala Asp Gly Lys Trp Ile Glu Pro Phe Asp Pro Ile Trp
500 505 510
Ser Gly Gly Gln Gly Gly Arg Asp Phe Tyr Thr Glu Asn Asn Gly Trp
515 520 525

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

595 600 605
Gln Arg Arg Ile Arg Gln Ile Leu Asp Leu Trp Tyr Gly Asp Gly Pro

610 615 620

Leu Gly Phe Cys Gly Asp Glu Asp Tyr Gly Glu Met Ser Ser Trp Tyr
625 630 635 640

Ile Leu Ser Ala Met Gly Phe Tyr Thr Val Ala Pro Gly Arg Pro Val
645 650 655

Tyr Asp Ile Gly Ser Pro Leu Phe Glu Lys Ser Thr Ile Asp Ile Gly
660 665 670

Asp Gly Lys Lys Phe Thr Ile Glu Cys Arg Asn Ile Ser Thr Gln Asn
675 680 685

Lys Tyr Ile
690

<210> 49

<211> 706

<212> PRT

<213> Unknown

<220><223> Parabacteroides distasonis

<400> 49

Met Lys Thr Lys Leu Lys Thr Ser Met Ala Leu Leu Ala Ser Val Phe
1 5 10 15

Leu Trp Val Ser Cys Ala Gly Gly Gly Gly Thr Ser Pro Ser Ala Met
20 25 30

Asp Pro Val Asp Tyr Val Asn Pro Tyr Met Gly Asn Ile Ser His Leu
35 40 45

Leu Val Pro Thr Phe Pro Thr Val His Leu Pro Asn Ser Met Leu Arg
50 55 60

Val Tyr Pro Glu Arg Ala Asp Phe Thr Gly Asp Arg Leu Gly Gly Leu
65 70 75 80

Pro Ile Ile Val Thr Asn His Arg Glu Arg Ser Ala Phe Asn Leu Cys
85 90 95

Pro Tyr Gln Gly Asp Glu Ser Gly Leu Arg Pro Val Ile Ala Tyr Ser

100 105 110

Tyr Asp Arg Glu Lys Ile Leu Pro Tyr Arg Tyr Gln Val Tyr Leu Asp

115 120 125

Asn Glu Glu Ile Asp Val Asp Phe Ala Pro Ser His Gln Ser Ala Val

130 135 140

Tyr Ser Leu Thr Phe Glu Lys Glu Gly Pro Ala Tyr Leu Val Phe Asn

145 150 155 160

Ser Arg Asn Gly Gln Leu Gln Val Asn Gly Asn Ala Val Ser Gly Tyr

165 170 175

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

180 185 190

Gln Lys Pro Val Lys Ala Gly Val Leu Ser Asn Gly Ser Val Asn Tyr

195 200 205

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

210 215 220

Gly Glu Asp Val Lys Lys Leu Gly Val Arg Tyr Gly Ile Ser Phe Ile

225 230 235 240

Ser Glu Glu Gln Ala Lys Lys Asn Leu Glu Arg Glu Ile Ala Ala Tyr

245 250 255

Asp Val Asp Val Val Ala Lys Ile Ala Arg Asn Asp Trp Asn Asp Ala

260 265 270

Leu Gly Lys Ile Gln Val Gln Gly Gly Thr Lys Asp Glu Lys Thr Val

275 280 285

Phe Tyr Thr Ser Leu Tyr Arg Cys Tyr Glu Arg Pro Ile Asn Leu Ser

290 295 300

Glu Asp Gly His Tyr Tyr Ser Ala Phe Asp Gly Lys Ile His Glu Asp

305 310 315 320

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

325 330 335

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

340 345 350

Ile Ile Asn Ser Tyr Leu Leu Met Ala Glu Gln Met Gly Thr Asp Trp

355 360 365

Met Pro Thr Phe Pro Glu Val Thr Gly Asp Thr Arg Arg Met Asn Ser

370 375 380

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

385 390 395 400

Gly Phe Glu Leu Glu Lys Ala Tyr Ile Ala Cys Lys Lys Gly Ile Glu

405 410 415

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

420 425 430

Asp Phe Tyr Lys Glu His Gly Tyr Ile Pro Ala Leu Arg Pro Gly Glu

435 440 445

Lys Glu Thr Val Pro Asn Val Ser Ile Trp Glu Lys Arg Gln Pro Ile

450 455 460

Ala Val Thr Leu Gly Thr Ser Tyr Asp Glu Trp Cys Leu Ser Gln Ile

465 470 475 480

Ala Gln Glu Leu Gly Lys Lys Asp Glu Ala Asp Tyr Tyr Leu Arg Arg

485 490 495

Ser Tyr Asn Phe Asn Thr Pro Leu Gly Met Ser Lys Trp Gln Phe Tyr

500 505 510

Ser Thr Leu Pro Asp His Thr Gly Asn Val Gly Met Phe Ser Met Ala

515 520 525

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

530 535 540

Pro Trp Met Thr Gln Lys Arg Ile Arg Thr Leu Leu Asn Gln Trp Phe

545 550 555 560

Arg Asn Asp Leu Met Gly Val Tyr Arg Asn Val Phe Asn Pro Glu Thr

565 570 575

Gly Phe Phe His Pro Lys Asp Lys Asp Gly Lys Phe Ile Tyr Pro Leu

580 585 590

Asp Tyr Arg Tyr Asp Gly Gly Leu Gly Ala Arg Asp Tyr Tyr Asp Glu

595 600 605

Asn Asn Gly Tyr Ile Tyr Arg Trp Asp Val Gln His Asn Ile Gly Asp

610 615 620

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

675 680 685
Gly Asn Thr Phe Glu Ile Arg Ala Asn Asn Ala Ser Asp Glu Asn Lys
690 695 700
Tyr Val
705

<210> 50
<211> 1650
<212> PRT
<213> Unknown
<220><223> Cellulosimicrobium cellulans
<400> 50

Ala Pro Glu Pro Pro Ser Ala Asp Tyr Ala Ser Leu Val Asp Val Phe
1 5 10 15
Val Gly Thr Glu Gly Asp Phe Gly Asn Asp Met Pro Ala Ala Gln Ala
20 25 30

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

100 105 110

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

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

245 250 255
Val Ser Val Glu Gln Ala Arg Ile Asp Gln Gln Val Glu Leu Gly Asp
260 265 270
Leu Ser Phe Asp Ala Ile Arg Asp Arg Thr Arg Ala Glu Trp Asn Ala
275 280 285
Thr Leu Gly Arg Val Ala Ile Asp Ala Ser Thr Ala Thr Asp Pro Thr
290 295 300
Gly Glu Leu Gln Arg Leu Phe Tyr Thr His Leu Tyr Arg Met Phe Ala
305 310 315 320

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

Ala Leu Tyr Arg Asp Met Val Gln Ser Leu Val Tyr Leu Phe Ala Asp

370 375 380

Ala Glu Ala Thr Gly Thr Gly Gly Gly Leu Gly Gly Phe Val His Ser

385 390 395 400

Val Pro Thr Val Arg Trp Glu Arg Ser Ser Val Val Val Ala Asp Ala

405 410 415

Ile Ala Lys Gly Phe Asp Gly Phe Asp Arg Leu Asp Glu Ala Tyr Pro

420 425 430

Ala Leu Gln Arg Leu Val Gly Gln Tyr Ser Ala Asp Glu Leu Arg Arg

435 440 445

Gly Tyr Val Ala Gly Asn Pro Gly Ala Ser Val Gln Arg Gly Tyr Asp

450 455 460

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

465 470 475 480

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

485 490 495

Lys Pro Gly Ala Trp Thr Ala Ala Asp Gly Thr Gln Val Gly Leu Leu

500 505 510

Thr Pro Arg Ala Ala Asp Gly Ser Trp Gln Ser Ala Asp His Ala Lys

515 520 525

Phe Glu Ala Ala Gly Leu Tyr Gln Gly Thr Leu Trp Gln Tyr His Trp

530 535 540

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

545 550 555 560

Glu Ala Ala Arg Leu Gly Met Arg His Met Phe Gly Glu His Ala Pro

565 570 575

Asp Asp Gly Lys Ala Met Leu His Ser Asn Ala Asn Glu Ile Asp Leu

580 585 590

Gln Ala Pro Tyr Leu Phe Asn Tyr Thr Gly Glu Pro Ser Leu Thr Gln

595 600 605

Lys Trp Ala Arg Ala Ile Tyr Thr Lys Glu Thr Trp Asn Arg Tyr Ile

610 615 620

Ala Thr Gly Ser Ser Ser Ala Val Pro Ser Gly Gly Gly Glu Phe Thr

625 630 635 640

Pro Pro Leu Lys Thr Lys Val Tyr Arg Leu Asp Pro Arg Gly Met Leu

645 650 655

Pro Thr Met Asp Asn Asp Ala Gly Thr Met Ser Thr Met Phe Val Ala

660 665 670

Ala Ala Val Gly Leu Phe Pro Val Thr Ala Gly Ser Ser Gln Phe Gln

675 680 685

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

690 695 700

Ser Ala Phe Thr Val Thr Ala Asp Gly Val Ser Glu Asp Ala Phe Tyr

705 710 715 720

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

725 730 735

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

740 745 750

Glu Gln Pro Ser Asp Trp Gly Thr Asp Thr Ala Pro Ala Phe Ser Met

755 760 765

Ser Thr Ala Thr Asp Glu Pro Ala Glu Gly Pro Arg Val Ser Ala Glu

770 775 780

Pro Thr Thr Val Gln Thr Gly Asp Gly Gly Ala Leu Asp Ala Thr Val

785 790 795 800

Thr Leu Thr Leu Asp Gly Ala Arg Leu Ala Ala Pro Ala Gly Thr Asp

805 810 815

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

820 825 830

Thr Ala Ala Val Thr Val Ala Ser Pro Thr Ala Leu Thr Val Ser Leu

835 840 845

Thr Gly Thr Ala Ser Ala Asp Ala Arg Phe Phe Val His Leu Arg Asp

850 855 860

Ala Ala Leu Ala Asp Gly Val Ala Ala Ala Ser Leu Gln Gly Gln Gly

865 870 875 880

Val Ser Val Arg Ser Pro Leu Arg Leu Ser Val Ala Ser Ala Glu Arg

885 890 895

Asp Ala Leu Ala Ala Leu Val Asp Asp Ala Val Leu Val Arg His Gly

900 905 910

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

915 920 925

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

930 935 940

Phe Ala Ala Asp Arg Leu Gly Ala Ala Ala Asp Ala Leu Asp Leu Thr

945 950 955 960

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

965 970 975

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

980 985 990

Gly Val Arg Ser Gly Ser Trp Val Gln Tyr Arg Asp Met Thr Phe Glu

995 1000 1005

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

1010 1015 1020

Thr Ser Phe Ala Pro Thr Asp Thr Pro Ser Thr Val Arg Val His Ala

1025 1030 1035 1040

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

1045 1050 1055

Ser Gly Trp Gly Lys Tyr Thr Glu Val Thr Ala Glu Leu Gly Asp Val

1060 1065 1070

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

1075 1080 1085

Ser Gly Arg Ser Trp Val Gly Asn Phe Asp Trp Phe Arg Phe Ser Ala

1090 1095 1100

Glu Asp Pro Ala Ala Pro Gly Gln Pro Gly Glu Ser Pro Thr Val Thr

1105 1110 1115 1120

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

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

1250 1255 1260
Gly Pro Thr Ser Val Val Val Glu Ser Glu Ala Trp Thr Ser Asn Ser
1265 1270 1275 1280
Gly Arg Gly Leu Lys Asn Glu Ser Ser Thr Trp Thr Ser Gly Pro Val
1285 1290 1295
Thr Asn Val Gly Gly Thr Ala Asp Gly Asp Trp Leu Ala Tyr Gly Glu
1300 1305 1310
Ile Asp Leu Gly Ser Ala Ala Leu Asp Gln Leu Ser Val His Tyr Val
1315 1320 1325

His Asn Ser Asn Arg Ser Gly Arg Asn Ser Ala Leu Ser Val Tyr Leu
1330 1335 1340
Asp Ala Phe Asp Pro Ala Asn Pro Gly Glu Pro Phe Val Thr Val Pro
1345 1350 1355 1360
Leu Ala Asn Thr Gly Ser Ser Trp Thr Thr Asp Gly Thr Ala Val Val
1365 1370 1375

Asp Leu Pro Ser Thr Val Arg Gly Lys His Gln Val Trp Val Arg Leu

1380 1385 1390

Ser Thr Glu Ala Tyr Ala Asp His Pro Tyr Val Ala Asn Leu Asp Ser

1395 1400 1405

Met Arg Phe Phe Thr Asp Ala Tyr Asp Val Glu Val Pro Pro Thr Asp

1410 1415 1420

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

1425 1430 1435 1440

Glu Ile Ala Arg Tyr Gly Arg Ile Asp Ala Arg Val Phe Thr Arg Glu

1445 1450 1455

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

1460 1465 1470

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

1475 1480 1485

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

1490 1495 1500

Ala Leu Thr Asp Glu Gly Tyr Ser Pro Glu Ser Trp Gln Ala Phe Arg

1505 1510 1515 1520

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

1525 1530 1535

Asp Glu Ala Leu His Asp Ala Arg Leu Ala Leu Gln Gly Ala Val Asp

1540 1545 1550

Ala Leu Glu Glu Pro Ala Asp Val Val Leu Val Glu Val Glu Val Ser

1555 1560 1565

Pro Arg Cys Leu Ala Gly Lys Pro Tyr Val Ala Val Arg Ala Val Asn

1570 1575 1580

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

1585 1590 1595 1600

Arg Ser Phe Val Gly Val Ala Pro Gly Ala Ser Ala Tyr Gln Ser Phe

1605 1610 1615

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

1620 1625 1630

Gly Ala Asp Gly Thr Gln Thr Val Glu Gln Val Val Thr Val Pro Ser

1635 1640 1645

Cys Ser

1650

<210> 51

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<220><221> VARIANT

<222> (4)

<223> Xaa = Gly, Ala, Ser, Thr, or Cys

<220><221> VARIANT

<222> (5)

<223> Xaa = Gly, Ala, Ser, Thr, or Cys

<220><221> VARIANT

<222> (7)

<223> Xaa = Gly, Ala, Ser, Thr, or Cys

<400> 51

Val Arg Xaa Glu

1

<210> 52

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<220><221> VARIANT

<222> (3)

<223> Any amino acid except Pro

<400> 52

Val Arg Xaa Glu

1

<210> 53

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<220><221> VARIANT

<222> (1)

<

223> Xaa = Leu, Ile, Val, Ala, Phe, Tyr, or Met

<400> 53

Val Arg Xaa Glu

1

<210> 54

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<220><221> VARIANT

<222> (3)

<223> Xaa = Any amino acid except Pro

<400> 54

Gly Asp Xaa Gly Asn

1 5

<210> 55

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<220><221> VARIANT

<222> (4)

<223> Xaa = Gly, Ser, Thr, Val, Ala, Cys, or Gln

<220><221> VARIANT

<222> (5)

<223> Xaa = Gly, Ser, Thr, Val, Ala, Cys, or Gln

<220><221> VARIANT

<222> (7)

<223> Xaa = Gly, Ser, Thr, Val, Ala, Cys, or Gln

<400> 55

Gly Val Gly Xaa Xaa Gly Xaa Gly Gly

1 5

<210> 56

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic 6xHis tag

<400> 56

His His His His His His

1 5

<210> 57

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<400> 57

Leu Tyr Gln Gly Thr

1 5

<210> 58

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic peptide

<400> 58

Asp Glu Val Asp

1