

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2012-0031750 (43) 공개일자 2012년04월04일
(51) 국제특허분류(Int. Cl.) C12N 9/24 (2006.01) C12N 15/56 (2006.01) C12Q 1/34 (2006.01)	(71) 출원인 고려대학교 산학협력단 서울 성북구 안암동5가 1	
(21) 출원번호 10-2010-0093316	(72) 발명자	
(22) 출원일자 2010년09월27일	최인걸	
심사청구일자 2010년09월27일	서울특별시 동작구 남부순환로263길 34 (사당동) 김경현 경기 성남시 분당구 삼평동 봇들마을 907동 1704호 이세영 경기도 의왕시 호성로 48-7, 엘지아파트 103동 1502호 (오전동)	
	(74) 대리인 특허법인다나	

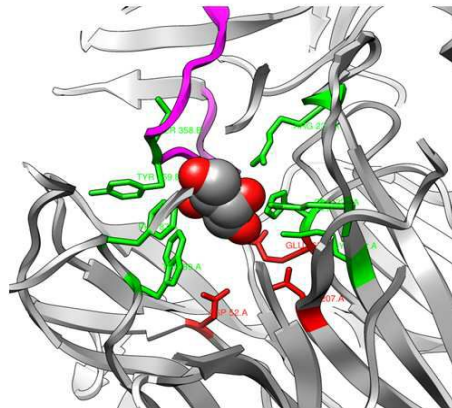
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 신규한 알파-네오 아가로바이오스 가수분해 효소 및 그 3차 구조

(57) 요약

본 발명은 신규한 알파-네오 아가로바이오스 가수분해 효소 (α -neoagarobiose hydrolase) 및 그 효소의 3차 구조에 관한 것이다.

대표도 - 도22



특허청구의 범위

청구항 1

서열번호 4의 52번 잔기의 아스파테이트,

207번 잔기의 아스파테이트, 및

265번잔기의 글루타메이트를 반드시 포함하는 알파-네오 아가로바이오스 가수분해 효소(α -neoagarobiose hydrolase).

청구항 2

제 1항에 있어서 상기 알파-네오 아가로바이오스 가수분해 효소는 서열번호 4의 222번 잔기인 라이신, 224번 잔기의 글루타메이트, 235번 잔기인(ARG), 아르기닌, 358번 잔기인 세린, 및 359번 잔기인 타이로신을 더욱 포함하는 것을 특징으로 하는 알파-네오 아가로바이오스 가수분해 효소(α -neoagarobiose hydrolase).

청구항 3

제 1항 또는 제2항에 있어서, 상기 알파-네오 아가로바이오스 가수분해 효소는 서열번호 1 내지 서열번호 11의 아미노산 서열 중 하나를 가지는 효소인 것을 특징으로 하는 알파-네오 아가로바이오스 가수분해 효소(α -neoagarobiose hydrolase).

청구항 4

제 3항에 있어서, 상기 알파-네오 아가로바이오스 가수분해 효소는 서열번호 4의 아미노산 서열로 이루어진 알파-네오 아가로바이오스 가수분해 효소(α -neoagarobiose hydrolase).

청구항 5

서열번호 4의 52번 잔기의 아스파테이트, 207번 잔기의 아스파테이트, 및

265번 잔기의 글루타메이트를 활성부위로 가지고, 공간 그룹은 C2이고, 유니트 셀 파라미터는 $a = 129.83\text{\AA}$, $b = 76.81\text{\AA}$, $c = 90.1\text{\AA}$ 이고, $\beta = 101.86^\circ$ 인 것인 알파-네오 아가로바이오스 가수분해 효소 결정.

청구항 6

제 5항에 있어서, 상기 결정은 갈락토스 결합 부위인 서열번호 4의 222번 잔기인 라이신, 224번 잔기인 글루타메이트, 235번 잔기인 아르기닌, 358번 잔기인 세린, 및 359번 잔기인 타이로신을 더욱 포함하는 것인 알파-네오 아가로바이오스 가수분해 효소 결정.

청구항 7

제 5항 또는 제 6항에 있어서, 상기 결정은 C-말단 부위인 서열번호 4의 354-366 잔기가 D-갈락토스와 상호작용하는 알파-네오 아가로바이오스 가수분해 효소 결정.

청구항 8

제 5항 또는 제6항에 있어서, 상기 결정은 체인 A와 체인 B로 구성된 상동 이량체 구조이고, 각 단량체는 5개의 날개를 갖는 베타-프로펠러 접힘 구조를 가지는 것을 특징으로 하는 알파-네오 아가로바이오스 가수분해 효소 결정.

청구항 9

a)알파-네오 아가로바이오스 가수분해 효소를 발현 및 정제하는 단계; 및

b)상기 효소 단백질과 침전용액을 혼합하여 시딩 드롭(sitting drop) 방법으로 셋업 한 후, 실온에서 결정화하는 단계를 포함하는 제 5항 또는 제6항의 알파-네오 아가로바이오스 가수분해 효소 결정화 방법.

청구항 10

제 9항에 있어서, 상기 결정은 갈락토스 결합 부위인 서열번호 4의 222번 잔기인 라이신, 224번 잔기인 글루타메이트, 235번 잔기인 아르기닌, 358번 잔기인 세린, 및 359번 잔기인 타이로신을 더욱 포함하는 것인 제 5항 또는 제 6항의 알파-네오 아가로바이오스 가수분해 효소 결정화 방법.

청구항 11

제 9항 또는 제 10항에 있어서, 상기 결정은 C-말단 부위인 서열번호 4의 354-366 잔기가 D-갈락토스와 상호작용하는 제 5항 또는 제 6항의 알파-네오 아가로바이오스 가수분해 효소 결정화 방법.

청구항 12

제 9항 또는 제 10항에 있어서, 상기 결정은 체인 A와 체인 B로 구성된 상동 이량체 구조이고, 각 단량체는 5개의 날개를 갖는 베타-프로펠러 접힘 구조를 가지는 것을 특징으로 하는 제 5항 또는 제 6항의 알파-네오 아가로바이오스 가수분해 효소 결정화 방법.

청구항 13

- a) 제 5항 또는 제 6항의 알파-네오 아가로바이오스 가수분해 효소 결정을 얻는 단계;
- b) 상기 결정의 원자 코디네이트들을 얻는 단계; 및
- c) 상기 효소와 상호작용하는 물질을 동정하기 위하여 상기 원자 코디네이트들과 하나 이상의 분자 모델링 기술을 사용하는 단계를 포함하는 알파-네오 아가로바이오스 가수분해 효소 활성 조절물질 동정 방법.

청구항 14

제 13항에 있어서, 상기 분자 모델링 기술은 그래픽 분자 모델링 또는 컴퓨터 화학(computational chemistry) 기술인 알파-네오 아가로바이오스 가수분해 효소 활성 조절물질 동정 방법.

명세서

기술분야

[0001] 본 발명은 신규한 알파-네오 아가로바이오스 가수분해 효소 (α -neoagarobiose hydrolase) 및 그 3차 구조에 관한 것이다.

배경기술

[0002] 현재 대기 중으로 배출되는 이산화탄소의 양의 증가로 인한 지구 온난화 뿐만 아니라, 가까운 미래에 일어날 수도 있는 원유의 고갈 및 가격상승에 따라 전세계적으로 새로운 이산화탄소 저감형 대체 에너지의 개발이 급격히 요구되고 있다. 다양한 대체에너지들 중 재생가능하고 풍부한 지구자원중의 하나인 바이오매스를 이용하여 생산되는 바이오에너지가 앞서의 요구들을 충족하는 주요 재생에너지로 현재 각광받고 있다. 특히 바이오에탄올은 현재 많은 수요를 보이고 있는 수송용 연료의 대안으로 고려되고 있으며, 미국을 비롯한 선진국에서는 이미 바이오에탄올의 사용을 법률로서 의무화 하고 있다. 현재 바이오 에탄올을 생산하기 위해 사용하고 있는 원료물질들은 옥수수나 기타 식량자원에서 유래된 당질, 전분 등에 제한되어 있어(제1세대 바이오 연료), 인류의 식량자원과 경쟁하며 그로 인해 국제 곡물가격 상승의 유발한다는 문제점을 안고 있다. 이러한 단점을 극복할 수 있는 대안으로 식량자원과 경쟁하지 않는 새로운 옥상 목질계 바이오매스(제2세대 바이오연료)나 해양 해조류 바이오매스(제3세대 바이오연료)가 차세대 바이오에너지 원료로서 또한 주목 받고 있으며, 이를 활용하여 바이오 에너지를 생산하는 기술들이 현재 활발하게 연구되고 있다.

[0003] 해조류 바이오매스는 목질계 바이오매스에 비해 미생물이 이용가능한 다당류의 함량이 많고, 리그닌을 함유하고 있지 않아 전처리가 비교적 간단하며, 연중 여러 번 수확 가능한 것 등 여러 가지 장점을 가지고 있다. 특히 대한민국은 삼면이 바다로 이루어져 있어 해조류를 생물자원이 용이하며, 해조류의 연간 생산량은 2006년 기준 13,754톤으로서 중국, 일본, 북한과 더불어 세계적인 해조류 생산국가에 속하고 있으나, 그 활용률면에서는 식용자원 이외로는 저조한 실정이다 (어업생산통계, 2006, 통계청 사회통계국 농어업생산통계과). 최근 들어서 해조류로부터 바이오에너지를 생산하는 연구가 일본과 대한민국 등에서 활발하게 논의되고 있다. 일본의 수산업진

홍재단에서는 'Ocean Sunrise Project'라는 이름으로 일본의 배타적 경제수역과 해양벨트의 미사용 해역 4.47 백만 km²에서 해조류를 대량 양식하여 50억 리터의 연료용 에탄올을 생산하려는 계획을 세우고 있다(Aizawa M et al., Seaweed bioethanol production in Japan - The Ocean Sunrise Project, OCEANS Conference, Sep. 29-Oct. 04, 2007, Vancouver, Canada). 또한 한국에서는 2009년 1월 최종 발표된 정부의 17개 신성장동력사업 분야 중 1개 분야인 신재생에너지 분야에 해양바이오연료가 포함되어 해조류에 대한 관심이 더욱 집중되고 있다. 2009년 농림수산식품부 해양바이오매스 확보 및 활용기술개발 연구기획 발표에 의하면 사방 71km의 정사 각형의 해역에서 해조류를 양식하여 바이오에탄올을 생산할 경우 연간 37.74억 리터의 생산이 가능하며 이는 2030년 한국의 휘발유 예상 소비량의 31.4%를 대체할 수 있는 양이다.

[0004] 현재 알려져 있는 해조류들 중에서 특히 홍조류 바이오매스(예, *Gelidium amansii*)를 원료물질로 사용하는 연구가 현재 활발하게 진행되고 있다. 홍조류는 전체 건중량의 70% 이상이 다당류로서 미생물이 이용가능한 발효성당으로의 전환이 가능하다. 특히, 홍조류 바이오매스에서 유래된 다당류의 주요 성분은 아가(agar)로서 전체 건중량의 60%정도를 함유하고 있어서 바이오 에너지생산을 위한 주요 원료로 고려되고 있다. 아가 다당체는 갈락토오스(D-galactose)와 3,6-안하이드로갈락토오스(3,6-anhydro-L-galactose; 줄여서 AHG)를 단위체로 하여, 그들이 각각 알파 1,3 결합(α -1,3-linkage) 또는 베타 1,4 결합(β -1,4-linkage)로 연결된 중합체이다(Duckworth, M. and W. Yaphe (1971) *Carbohydrate Research* 16, 435-445) (도면 1 참고).

[0005] 아가 다당체를 탄소원으로 이용하는 미생물들의 경우, 베타 아가레이즈(β -agarase) 또는 알파 아가레이즈(α -agarase)를 이용하여 아가 다당체를 작은 크기의 올리고당으로 자른 후에 최종적으로 베타 아가레이즈의 경우 알파-네오아가로바이오스(α -neoagarobiose, α -1,3-D-galactosyl-3,6-anhydro-L-galactose)로 분해하고, 알파 아가레이즈의 경우는 베타-아가로바이오스(β -agarobiose, β -1,4-anhydro-L-galactosyl-D-galactose)로 분해하는 것으로 알려져 있다. 베타 아가레이즈의 분해산물인 네오아가로바이오스의 경우, 미생물이 대사하기 위해서는 갈락토오스로 전환이 필요하며, 그를 위해 알파-1,3 결합을 끊는 효소(알파-네오아가로바이오스 가수분해 효소)가 필수적이라고 알려져 있다(Ekborg, N. A. et al (2005) *Int. J. Syst. Evol. Microbiol.* 55, 1545-1549; Ekborg, N. A. et al., (2006) *Appl. Environ. Microbiol.* 72, 3396-3405). 그러나 아직까지 *S. degradas*에서의 네오아가로바이오스의 알파-1,3 결합을 자르는 효소는 발견되지 않았다(Ekborg, N. A. et al. (2006) *Appl. Environ. Microbiol.* 72, 3396-3405).

[0006] 미생물에서 아가로스로부터 올리고아가로사카라이드를 생산하는 베타-아가레이즈는 예컨대, 슈도모나스(*Pseudomonas*) 속 균주(Ha, J. C. et al. (1997) *Biotechnol. Appl. Biochem.* 26:1-6), 알테로모나스(*Alteromonas*) 속 균주(Potin, P., et al. (1993) *Eur. J. Biochem.* 214:599-607), 아가리보란스(*Agarivorans*) 속 균주(Ohta, Y. et al. (2005) *Biotechnol. Appl. Biochem.* 41:183-191), 슈도알테로모나스(*Pseudoalteromonas*) 속 균주(Belas, R. (1989) *J. Bacteriol.* 171:602-605), 마이크로실라(*Microsilla*) 속 균주(Zhong, Z. et al. (2001) *Appl. Environ. Microbiol.* 67:5771-5779) 및 비브리오(*Vibrio*) 속 균주(Aoki, T. et al. (1990) *Eur. J. Biochem.* 187:461-465) 등 많은 미생물에 의해 생산되는 것으로 보고 되었다.

[0007] 홍조류에서 유래된 아가 다당체를 바이오 에너지생산을 위한 원료로 사용할 경우, 여러단계의 전처리과정을 거쳐 실제 미생물이 이용 가능한 발효성 당으로의 전환이 필수적이다. 아가 다당체의 발효성 단당류로의 전환은 크게 화학적 전처리와 생물학적전처리의 두 가지의 공정을 통하여 가능하다. 첫번째는 산가수분해등을 이용하는 화학적 방법으로 비교적 간단한 공정이나 다당류로 구성된 바이오매스를 고온에서 화학적으로 전처리하여 퍼피릴(furfural), HMF(하이드록시메틸퍼피릴) 등의 독성부산물을 다량 생성하고, 무작위로 절단된 단당류 및 올리고의 혼합물로 얻는다는 단점이 있다. (Pickering et al., 1993, *Journal of Applied Phycology* 5: 85-91; Armise'n, 1995). 그에 반해 아가레이즈와 같은 효소를 사용하는 생물학적 전처리 및 당화 방법은 상온에서 환경친화적인 방법으로 발효성당인 갈락토오스를 얻을 수 있다는 장점이 있으나, 현재 상업적으로 구할 수 있는 효소는 베타-아가레이즈에 국한하고 있으며, 아가레이즈의 산물도 통상의 미생물들이 사용하기 어려운 이당류(네오아가로바이오스나 아가로바이오스)를 최종산물로 한다는 단점이 있다.

[0008] 베타-아가레이즈 반응의 결과 생성된 네오아가로바이오스의 경우, 바이오에너지생산을 위해 사용하기 위해서는 발효성 단당류인 갈락토오스로의 전환이 필수적이며, 이때 알파-네오아가로바이오스 가수분해효소가 필요하다. 따라서 홍조류 바이오매스를 바이오에탄올과 같은 바이오에너지 생산의 원료로 사용하기 위한 효율적인 바이오매스 생물학적(효소적) 전처리 및 당화 공정의 최종단계에는 알파-네오아가로바이오스 가수분해효소가 필수불가결하다. 또한 네오아가로바이오스의 분해산물로 갈락토오스와 함께 얻어지는 AHG는 상업적으로도 판매되고 있지 않으며 D 형의 AHG만 구입이 가능한 실정이며 이 또한 고가에 판매되고 있다(2009년 현재 200 파운드(영

국)/100mg, dextra Laboratories). 따라서 본 효소를 이용하여 아가로스로부터 고가의 희귀성 단당류인 AHG의 대량생산이 가능하다.

발명의 내용

해결하려는 과제

- [0009] 본 발명은 상기의 문제점을 해결하고, 상기의 필요성에 의하여 안출된 것으로서 본 발명의 목적은 바이오에너지 생산 등을 위한 발효성 단당류인 갈락토오스로의 전환에 필수적인 효소를 제공하는 것이다.
- [0010] 본 발명의 다른 목적은 네오아가로바이오스 효소의 결정 및 그 결정화 방법을 제공한다.
- [0011] 본 발명의 또 다른 목적은 네오아가로바이오스 효소의 결정 및 그 결정화 방법을 이용한 그 효소와 상호작용하는 물질을 동정하는 방법을 제공한다.

과제의 해결 수단

- [0012] 상기의 목적을 달성하기 위하여 본 발명은 서열번호 4의 52번 잔기의 아스파테이트, 207번 잔기의 아스파테이트, 및 265번잔기의 글루타메이트를 반드시 포함하는 알파-네오 아가로바이오스 가수분해 효소(α -neogagarbiose hydrolase)를 제공한다.
- [0013] 본 발명의 일 구현예에 있어서 상기 알파-네오 아가로바이오스 가수분해 효소는 서열번호 4의 222번 잔기인 라이신, 224번 잔기의 글루타메이트, 235번 잔기인(ARG), 아르기닌, 358번 잔기인 세린, 및 359번 잔기인 타이로신을 더욱 포함하는 것이 바람직하나 이에 한정되지 아니한다.
- [0014] 본 발명의 또 다른 구현예에 있어서, 상기 알파-네오 아가로바이오스 가수분해 효소는 서열번호 1 내지 서열번호 11의 아미노산 서열 중 하나를 가지는 효소인 것이 바람직하나 이에 한정되지 아니한다.
- [0015] 본 발명의 바람직한 실시예에서, 상기 알파-네오 아가로바이오스 가수분해 활성을 가진 효소는 서열번호: 1 내지 서열번호 11에 개시된 아미노산 서열 뿐 아니라 상기 효소의 하나 이상의 치환, 결손, 전위, 첨가 등의 변이 단백질로서 상기 알파-네오 아가로바이오스 가수분해 활성을 가지는 단백질도 본 발명의 효소의 권리범위에 포함되며, 바람직하게는 서열번호: 서열번호 1 내지 서열번호 11에 개시된 아미노산 서열과 서열 동일성이 80% 이상, 85% 이상, 90% 이상, 93% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상 및 99% 이상인 아미노산 서열을 포함한다.
- [0016] 본 발명에서 폴리펩티드가 또 다른 서열에 대하여 특정 비율 (예컨대, 80%, 85%, 90%, 95%, 또는 99%)의 서열 동일성을 가진다는 것은, 상기 두 서열을 정렬시킬 때, 상기 서열들의 비교시에 상기 비율의 아미노산 잔기가 동일함을 의미한다. 상기 정렬 및 백분율 상동성 또는 동일성은, 당업계에 공지된 임의의 적당한 소프트웨어 프로그램, 예를 들어 문헌[CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel 등 (eds) 1987 Supplement 30 section 7.7.18)]에 기재된 것들을 사용하여 결정할 수 있다. 바람직한 프로그램으로는, GCG Pileup 프로그램, FASTA(Pearson 등 1988 Proc. Natl Acad. Sci USA85:2444-2448), 및 BLAST (BLAST Manual, Altschul 등, Natl. Cent. Biotechnol. Inf., Natl Lib. Med. (NCIB NLM NIH), Bethesda, MD, 및 Altschul 등 1997 NAR25:3389-3402)이 있다. 또 다른 바람직한 정렬 프로그램은 ALIGN Plus(Scientific and Educational Software, PA)로서, 바람직하게는 기본 매개변수를 사용하는 것이다. 사용가능한 또 다른 서열 소프트웨어 프로그램은 Sequence Software Package Version 6.0 (Genetics Computer Group, University of Wisconsin, Madison, WI)에서 이용가능한 TFASTA Data Searching Program 이다.
- [0017] 본 발명의 일 실시예에 있어서, 상기 효소는 사카로파거스 데그라단스 등으로부터 수득하는 것이 바람직하나 이에 한정되지 아니한다.
- [0018] 본 발명의 사카로파거스 데그라단스는 통상적으로 구입(*Saccharophagus degradans* ATCC 43961)할 수 있으나 이에 한정되지 아니하는 다양한 방법으로 얻을 수 있다.
- [0019] 또한 본 발명은 서열번호 4의 52번 잔기의 아스파테이트, 207번 잔기의 아스파테이트, 및 265번 잔기의 글루타메이트를 활성부위로 가지고, 공간 그룹은 C2이고, 유니트 셀 파라미터는 $a = 129.83\text{\AA}$, $b = 76.81\text{\AA}$, $c = 90.1\text{\AA}$ 이고, $\beta = 101.86^\circ$ 인 것인 알파-네오 아가로바이오스 가수분해 효소 결정을 제공한다.
- [0020] 본 발명의 일 구현예에 있어서, 상기 결정은 갈락토스 결합 부위인 서열번호 4의 222번 잔기인 라이신, 224번

잔기인 글루타메이트, 235번 잔기인 아르기닌, 358번 잔기인 세린, 및 359번 잔기인 타이로신을 더욱 포함하는 것이 바람직하나 이에 한정되지 아니한다.

- [0021] 본 발명의 또 다른 구현예에 있어서, 상기 결정은 C-말단 부위인 서열번호 4의 354-366 잔기가 D-갈락토스와 상호작용하는 것이 바람직하나 이에 한정되지 아니한다.
- [0022] 본 발명의 다른 구현예에 있어서, 상기 결정은 체인 A와 체인 B로 구성된 상동이량체 구조이고, 각 단량체는 5개의 날개를 갖는 베타-프로펠러 접힘구조를 가지는 것이 바람직하나 이에 한정되지 아니한다.
- [0023] 또한 본 발명은 a)알파-네오 아가로바이오스 가수분해 효소를 발현 및 정제하는 단계; 및 b)상기 효소 단백질과 침전용액을 혼합하여 시딩 드롭 방법으로 셋업 한 후, 실온에서 결정화하는 단계를 포함하는 알파-네오 아가로바이오스 가수분해 효소 결정화 방법을 제공한다.
- [0024] 또한 본 발명은 a)상기 본 발명의 알파-네오 아가로바이오스 가수분해 효소 결정을 얻는 단계;b) 상기 결정의 원자 코디네이트들을 얻는 단계; 및 c) 상기 효소와 상호작용하는 물질을 동정하기 위하여 상기 원자 코디네이트들과 하나 이상의 분자 모델링 기술을 사용하는 단계를 포함하는 알파-네오 아가로바이오스 가수분해 효소 활성 조절물질 동정 방법을 제공한다.
- [0025] 본 발명의 동정방법에 있어서, 상기 분자 모델링 기술은 그래픽 분자 모델링 또는 컴퓨터 화학(computational chemistry) 기술인 것이 바람직하나 이에 한정되지 아니한다.
- [0026] 또한 본 발명은 도 8에 기재된 모티프 1 내지 50(총 50개의 모티프)으로 구성된 군으로부터 선택된 모티프 중 1, 2, 3, 4, 5, 6, 7, 14, 16, 17, 20, 34, 40을 포함하는 13개의 모티프들을 포함하는 알파-네오 아가로바이오스 가수분해 효소(α -neogarobiose hydrolase)를 제공하며, 더욱 바람직하게는 도 8-9에 기재된 50개의 모티프들 중 모티프 7과 34를 반드시 포함하는 알파-네오 아가로바이오스 가수분해 효소(α -neogarobiose hydrolase)을 제공한다.
- [0027] 단백질 모티프는 규칙적으로 나타나는 패턴으로 표시할수 있으면, 통상적으로 패턴의 표현양식은 정규표현식으로 나타낸다; <http://www.expasy.ch/prosite/prosuser.html>; 예를 들어:PA [AC]-x-V-x(4)-{ED}.
- [0028] 본 패턴은 다음과 같이 해석된다: [Ala or Cys]-any-Val-any-any-any-any-{any but Glu or Asp}
- [0029] PA <A-x-[ST](2)-x(0,1)-V.
- [0030] 이 패턴은 서열의 N-말단에 있어야 하고('<'), 다음과 같이 해석된다: Ala-any-[Ser or Thr]-[Ser or Thr]-(any or none)-Val; Sigrist C.J.A., Cerutti L., Hulo N., Gattiker A., Falquet L., Pagni M., Bairoch A., Bucher P.PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform. 3:265-274(2002);Sigrist C.J.A., De Castro E., Langendijk-Genevaux P.S., Le Saux V., Bairoch A., Hulo N. ProRule: a new database containing functional and structural information on PROSITE profiles. Bioinformatics. 2005 Nov 1;21(21):4060-6. Epub 2005 Aug 9; Timothy L. Bailey, Nadya Williams, Chris Misch, and Wilfred W. Li, MEME: discovering and analyzing DNA and protein sequence motifs, Nucleic Acids Research, Vol. 34, pp. W369-W373, 2006).
- [0031] 본 발명에서 세포, 핵산, 단백질 또는 벡터와 관련하여 사용될 때 용어 "재조합"은, 상기 세포, 핵산, 단백질 또는 벡터가 이중 핵산 또는 단백질의 도입 또는 본래적 핵산 또는 단백질의 변경에 의해 변형되었거나, 또는 상기 세포가 이렇게 변형된 세포로부터 유래한 것을 가리킨다. 즉, 예를 들어, 재조합 세포는 상기 세포의 본래적 (비(非)재조합) 형태 내에서는 발견되지 않는 유전자를 발현하거나 또는, 다르게는 발현시 비정상적으로 발현되거나 또는 전혀 발현되지 않는 본래적 유전자를 발현한다.
- [0032] 용어 "단백질" 및 "폴리펩티드"는 본원에서 상호교환가능하게 사용된다. 본원에서는 아미노산 잔기에 대하여 통상의 1문자 또는 3문자 코드가 사용된다.
- [0033] 본 발명에서 용어 "유래된"은 용어 "~로부터 기원하는", "수득되는" 또는 "~로부터 수득가능한", 및 "~로부터 단리된"을 포괄하며, 본원에서 사용될 때, 이는 상기 뉴클레오티드 서열에 의해 인코딩되는 폴리펩티드가 상기 뉴클레오티드가 본래적으로 존재하거나 또는 상기 뉴클레오티드 서열이 삽입되어 있는 세포로부터 생산됨을 의미한다.
- [0034] 용어 "배양"은 액체 또는 고체 배지에서 적당한 조건 하에서 미생물 세포 집단을 성장시키는 것을 지칭한다. 바람직한 실시예에서, 배양은 (전형적으로는 용기 또는 반응기 내에서의) 아가로스를 포함한 기질을 최종 생성물

로 생물전환을 지칭한다. 발효는 미생물에 의해 유기 물질을 효소적 및 혐기적으로 분해하여 더욱 단순한 유기 화합물을 생산하는 것을 지칭한다. 발효가 혐기적 조건하에서 일어나는 것이긴 하지만, 발효는 또한 산소의 존재하에서도 일어나므로, 상기 용어를 엄격한 혐기적 조건으로만 한정하고자 하는 의도는 없다.

[0035] 본 발명에서 사용된 용어 "회수", "단리" 및 "분리"는 자연적으로 결합되어 있는 하나 이상의 성분으로부터 제거된 화합물, 단백질, 세포, 핵산 또는 아미노산을 지칭한다.

[0036] 본원에서 사용된 세포와 관련하여 사용되는 용어 "형질전환", "안정한 형질전환" 및 "유전자 이식(transgenic)"은 세포가 여러 세대를 통해 유지되는 에피솜 플라스미드로서의 또는 그의 게놈 내로 통합된 비(非)본래적(예컨대, 이종) 핵산 서열을 가지는 것을 의미한다.

[0037] 본원에서 사용된 용어 "발현"은 유전자의 핵산 서열에 기초하여 폴리펩티드가 생산되는 방법을 지칭한다. 상기 방법에는 전사 및 번역이 포함된다.

[0038] 핵산 서열을 세포 내로 삽입하는 용어 "도입"은 "트랜스펙션(transfection)", 또는 "형질전환" 또는 "형질도입(transduction)"을 의미하며, 핵산 서열의 진핵 또는 원핵 세포 내로의 통합에 대한 언급이 포함되고, 이때 상기 핵산 서열은 세포의 게놈(예컨대, 염색체, 플라스미드, 색소체, 또는 미토콘드리아 DNA) 내로 통합되어, 자율 레플리콘으로 전환되거나, 또는 일시적으로 발현된다.

[0039] 이하 본 발명을 설명한다.

[0040] 1. 홍조류를 포함한 해양바이오매스의 전처리 공정 과정에서 알파-네오아가로바이오스로부터 발효성 당인 갈락토오스를 생성하는 효소처리공정

[0041] 아가 다당체로부터 발효성당을 생성하기 위해 단당류로 분해하는 전처리 과정으로는 크게 화학적인 방법과 효소적 방법 두 가지가 사용 가능하다. 먼저, 화학적인 방법은 복합 다당류를 무작위적으로 분해하기 때문에 원하는 발효성 단당류를 선택적으로 생산하기 매우 어려울 뿐만 아니라 생성된 당의 발효를 저해하는 부산물을 생성하기도 한다. 또한, 알칼리 처리나 산처리의 경우 상당한 양의 오염물질을 배출하여 이것을 정화시키기 위해 많은 비용이 소모되어 미생물이 이용가능한 발효성 당의 생산방법에는 적합하지 않다. 효소적인 방법으로는 미생물이 생성하는 아가레이즈를 이용하여 아가 다당체를 분해시키는 것이 가능하다. 이러한 방법에 사용가능한 효소로는 알파 아가레이즈 또는 베타 아가레이즈가 있다. 아가 다당체를 분해하는 효소로서 알파 아가레이즈는 아가 다당체에 존재하는 알파-1,3과 베타-1,4 결합 중 알파 결합을 가수분해하여 최종적으로 베타-아가로바이오스 이당체를 만들고 베타 아가레이즈는 베타 결합을 끊어 알파-네오아가로바이오스 이당체를 생성한다. 이러한 효소적 전처리의 과정 중 베타 아가레이즈를 사용할 경우, 최종 산물은 알파-네오아가로바이오스로 이는 통상의 미생물이 사용할 수 없는 비발효성 당이다. 이를 발효성 당인 갈락토오스로 최종적으로 전환하는 전처리 공정을 위해서는 알파-네오아가로바이오스 가수분해효소가 필요 불가결하다.

[0042] 2. 아가다당체의 효소적 분해산물인 알파-네오아가로바이오스로부터 갈락토오스 또는 AHG의 생산

[0043] 현재 네오아가로바이오스의 가수분해로 얻어지는 3,6-Anhydro-L-galactose는 상업적으로도 판매되고 있지 않으며 D 형의 AHG만 구입이 가능한 실정이며 이 또한 고가에 판매되고 있다(2009년 현재 200 파운드(영국)/100mg, dextra Laboratories). 따라서 본 효소를 이용하여 아가로스로부터 고가의 AHG의 대량생산이 가능하다.

[0044] 본 발명의 발명자들은 해양세균 사카로파거스 테그라단스에서 알파-네오아가로바이오스 가수분해효소의 활성을 확인하고, 해당 유전자를 최초로 확인하였다. 또한 대장균에 형질전환하여 단백질을 대량생산하여 알파-네오아가로바이오스 가수분해효소의 활성을 확인하였다. 신규 알파-네오아가로바이오스 가수분해효소 유전자의 효소 활성을 이용하여 아가 다당체를 포함하는 해양바이오매스의 전처리 공정에서 적용하고, 전처리 공정에서 얻어진 알파-네오아가로바이오스로부터 갈락토오스를 포함한 발효성 당의 획득에 이용한다.

[0045] 최근 아가 다당체를 탄소원으로 사용하여 성장할 수 있는 해양세균인 사카로파거스 테그라단스(*Saccharophagus degradans*)가 발견되었고 유전체 서열이 최근에 결정되었다.

[0046] 알파-네오아가로바이오스 가수분해효소의 유전자를 발굴하기 위하여, 먼저 사카로파거스 테그라단스에서 알파-네오아가로바이오스 가수분해효소의 활성을 확인하였다. 그 다음 유전체서열정보 분석을 통하여 사카로파거스 테그라단스의 유전자중 당가수분해효소 패밀리 32에 속하는 유전자 (단백질데이터베이스 Uniprot database ID: Q21HB2) 알파-네오아가로바이오스 가수분해효소로 예측하였다. 실제 단백질이 알파-네오아가로바이오스 가수분해효소의 활성을 가지고 있는지 확인하기 위하여, 발현용 벡터에 클로닝하고 대장균에서 과발현하여 단백질을

분리 정제한 다음 최종적으로 효소의 활성을 확인하였다.

발명의 효과

[0047]

본 발명에서는 신규한 알파-네오아가로바이오스 가수분해효소의 염기서열과 단백질서열을 보고하고 있다. 특히, 앞서 서술한 바와 같이 최근 해조류 바이오매스를 이용한 바이오 에너지 생산에서 바이오매스의 발효성당으로의 효소적 전환과정 중 아가 다당체로부터 안하이드로 갈락토오스 및 갈락토오스의 생산에는 알파-네오아가로바이오스 가수분해효소의 사용이 필수적이므로 바이오매스전처리 비용절감 및 수율향상 효과를 기대할 수 있다. 또한 바이오연료를 생산하는 효모나 세균등에 알파-NABH 유전자를 도입하여 아가 또는 네오아가로바이오스로 부터 직접 바이오 연료를 생산하는 것도 기대할 수 있다. 또한 해양바이오매스로 부터 고부가가치의 단당류인 갈락토오스와 AHG를 생산하는 공정에도 알파-네오아가로바이오스 가수분해효소를 사용하여, 유용물질의 생산도 가능하다.

도면의 간단한 설명

[0048]

도 1은 아가 다당체의 구조를 나타내는 그림이다.

도 2는 *S. degradans* 2-40으로부터 얻은 세포내 조효소에 의한 반응 생성물의 박막 크로마토그래피(Thin layer chromatography)를 나타낸 사진. 레인 A, 갈락토오스의 스탠다드; 레인 B, 0.25% (w/v) 아가로스를 세포내 효소와 배양한 것; 레인 C, 0.3%(w/v) 젤라틴(홍조류) 분말을 세포내 효소와 배양한 것. 효소반응은 30, 20 mM Tris-HCl (pH6.8)에서 12시간 동안 수행하였다.

도 3은 본 발명에서 발굴한 유전자의 염기서열 및 단백질 서열을 나타낸다.

도 4는 대장균에서 발현되어 정제된 아가레이즈; 12% SDS-폴리아크릴아마이드젤 전기영동; A : 분자크기 마커, B : 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase)를 나타냄.

도 5는 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 및 베타-아가레이즈에 의한 아가로스 가수분해의 산물의 박막 크로마토그래피 사진. 효소 반응을 30, 20 mM Tris-HCl(pH6.8)에서 수행하였고, 그 기질 농도는 0.25% (w/v)이었다. (a) 갈락토오스 (b-c) 반응 혼합물. A: 갈락토오스 표준물질; B: 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 반응산물(2시간 반응); C: 네오아가로바이오스.

도 6은 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 및 베타-아가레이즈에 의한 아가로스 가수분해의 산물의 액체크로마토그래피 질량분석 (Mass Spectrometry)을 통한 반응산물의 확인을 나타내고 있다. 질량분석기를 이용한 반응산물들의 질량분석결과 (formate(HCOO-, 분자량: 45)로 음이온화),

도 6에서 나타나는 것과 같이 갈락토오스의 질량은 225.1 (180+45), 안하이드로갈락토오스는 207.1 (162+45)에서 확인되고 있다. 따라서 가수분해효소의 산물이 갈락토오스와 안하이드로갈락토오스를 확인하고 있다. 즉 기초과학지원센터 서울분소에 질량분석을 의뢰하여 얻은 결과물이 도면 6에 나타난 질량분석스펙트럼이고, 그 결과물을 해석한 결과로 효소의 가수분해산물을 확인하였음.

도 7은 알파-네오아가로바이오스 가수분해 효소를 포함한 호모로그(homolog) 단백질의 모티프(motif)분석결과를 나타낸다.

도 8은 알파-네오아가로바이오스 가수분해 효소 특이적 모티프들을 나타낸다.

도 9-10은 도 8에서 확인된 알파-네오아가로바이오스 가수분해 효소 특이적 모티프들을 모두 포함하고 있는 단백질들의 서열을 나타내고 있다.

도 11은 도 9-10에서 얻어진 10개의 서열과 알파-네오아가로바이오스 가수분해효소의 아미노산 서열을 사용한 다중서열정렬을 나타낸다.

도 12는 대장균에서 발현되어 정제된 슈도알테로모나스 아틀란티카 유래 아가레이즈; 12% SDS-폴리아크릴아마이드젤 전기영동; Lane A, 분자크기 마커; Lane B, 알파-네오아가로바이오스 가수분해효소.

도 13은 대장균에서 발현되어 정제된 스트렙토마이세스 코엘리컬러 아가레이즈; 12% SDS-폴리아크릴아마이드젤 전기영동; Lane A, 알파-네오아가로바이오스 가수분해 효소; Lane B, 분자크기 마커.

도 14는 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 및 베타-아가레이즈에 의한 아가로스 가수분해의 산물의 박막 크로마토그래피 사진이다. 효소 반응을 30, 20mM Tris-HCl(pH6.8)에서 2시간 동안

수행하였고, 그 기질 농도는 0.25%(w/v)이었다. (A) 갈락토오스 (B-D) 반응 혼합물. Lane A, 갈락토오스 표준물질; Lane B, 사카로파거스 데그라단스(*Saccharophagus degradans* 2-40) 유래 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 반응산물; Lane C, 슈도알테로모나스 아틀란티카 (*Pseudoalteromonas atlantica* T6c) 유래 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 반응산물; Lane D, 스트렙토마이세스 코엘리컬러(*Streptomyces coelicolor* A3) 유래 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 반응산물.

도 15는 알파-네오아가로바이오스 가수분해 효소(neoagarobiose hydrolase) 결정화를 위한 데이터 수집 표이다. 표에서, $R_{\text{merge}} = \sum_{kl} \sum_i | \langle i(hkl) \rangle - \langle i(hkl) \rangle | / \sum_{kl} I_i(hkl)$, 이고 여기서, $\sum_{kl}$ 는 모든 리플렉션에 대한 합을 나타내고, $\sum_i$ 는 대칭 관련 리플렉션 및 모든 이큐벨런트의 합을 나타낸다.

도 16은 2.4 M sodium malonate pH 7.0을 사용하여 결정화된 알파-네오아가로바이오스 가수분해 효소의 결정 그림이다. 그림에서 결정의 차원은 120 x 120 μm 이다.

도 17은 Advanced Light Source, Lawrence Berkeley National Laboratory에서 싱크로트론 레디에이션을 사용하여 얻은 알파-네오아가로바이오스 가수분해 효소 결정의 회절 이미지를 나타낸다.

도 18은 알파-네오아가로바이오스 가수분해 효소 결정으로부터 얻어진 효소의 3차원 구조를 나타낸 것으로 5개의 베타 병풍구조 날개를 갖는 프로펠러 모양의 접힘 갖는 것을 보여주고 있다. 도18에 표시된 그리스문자들은 5개의 프로펠러 날개를 나타내고 있다.

도 19는 알파-네오아가로바이오스 가수분해 효소 결정의 호모 다이머 구조를 나타낸 것이다. 도 19에서 알 수 있는 바와 같이 본 발명의 효소는 도 18과 같이 단량체의 독립적인 폴리펩타이드로서 2개의 단량체가 서로 상호작용하여 만든 호모 다이머의 4차구조(quaternary structure)로서 활성형태를 갖는다.

도 20은 D-Galactose와 다이머의 알파-네오아가로바이오스 가수분해 효소 구조를 나타낸 것이다. 도 20은 단백질의 결정제조시에 본 발명의 효소의 반응산물인 D-Galactose를 첨가하여, 실제로 효소가 D-Galactose와 결합하고 있는 것을 보여주고 있다. 이는 D-Galactose가 결합한 부위가 실제 활성부위로 예측가능하게 한다.

도 21은 알파-네오아가로바이오스 가수분해 효소의 C-말단 부위(354-366 잔기)의 구조를 나타낸 것이다. 도 21은 서로 다른 서브유닛(subunit, 각각의 단량체)가 활성부위(D-Galactose)와 상호작용하고 있는 것을 보여주고 있다. 따라서 본 발명의 효소활성은 반드시 호모다이머에서 확인되며, 두 개의 단량체가 함께 참여하고 있음을 보여주고 있다. 그림에서 마젠타 색은 D-Gal과 상호작용하는 부위를 나타낸다.

도 22는 알파-네오아가로바이오스 가수분해 효소의 활성부위 예측 그림이다. 마젠타 색은 도 21과 같이 다른 단량체(모노머)에서 온 C-말단 루프, 녹색은 D-Galactose와 인접하는 잔기로, 222(Lys), 224(Glu), 235(ARG), 264(HIS), 358(SER), 359(TYR), 적색은 다른 유사 구조 호모로그 (GH32, GH48)에서 공통적으로 나타나는 촉매 활성을 담당하는 잔기들로 52(ASP), 207(ASP), 265(GLU), 원자 구형 모델로 표시된 부분은 D-갈락토스를 나타낸다.

도 23은 도 7-8에 나타낸 것과 같이 알파-네오아가로바이오스 가수분해 효소의 모티프 분석을 통해 얻어진 모티프 17과 34의 위치를 3차 구조에 직접 표시하였다. 모티프 17은 파란색, 모티프 34는 오렌지색으로 표시하였다. 구조를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0049] 이하, 비한정적인 실시예를 통하여 본 발명을 더욱 상세하게 설명한다. 단 하기 실시예는 본 발명을 설명하기 위한 목적으로 기재된 것으로서 본 발명의 범위는 하기 실시예에 의하여 제한되지 아니한다.

[0050] 실시예 1: 사카로파거스 데그라단스 조추출액으로 부터의 알파-네오아가로바이오 가수분해효소의 효소활성 확인

[0051] 사카로파거스 데그라단스가 알파-네오아가로바이오스 가수분해를 통해 단당류인 갈락토오스와 AHG를 생성하는 효소 활성을 가지고 있는지 확인하기 위하여 다음과 같은 방법으로 조추출액을 얻어 활성을 확인하였다. 사카로파거스 데그라단스를 해수염을 포함하는 배지에서 대수증식기 중간까지 키운 후 40ml의 배양액을 원심분리한 후에 초음파를 이용해 세포를 파쇄하고 조추출액을 얻은 다음 아가다당체를 베타 아가레이즈로 분해하여 얻어진 아가로를 기질로 하여 반응분해산물을 TLC를 통하여 관찰하였다. A레인에서는 갈락토오스 표준물, B레인에서는 0.25%(w/v)의 아가로오스를 사카로파거스 데그라단스의 세포내 조효소액과 반응시킨 산물, C레인에서는

0.3%(w/v) 홍조류인 건조된 우뭇가사리 가루를 사카로퍼거스 테그라단스의 세포내 조효소액과 반응시킨 후 산물을 TLC로 분석한 결과이다. 사카로퍼거스 테그라단스의 세포내 조효소액과 반응을 통하여 아가로오스와 우뭇가사리로부터 모두 단당류인 AHG와 갈락토오스 그리고 이당류를 생성하는 것으로 나타났다. 따라서 사카로퍼거스 테그라단스는 세포내에 알파-네오아가로바이오스 가수분해를 통해 단당류인 갈락토오스와 AHG를 생성하는 효소를 가지고 있는 것으로 알 수 있다.

[0052] 본 발명에서 발굴한 유전자의 염기서열 및 단백질의 아미노산 서열은 도면 3과 같다.

[0053] 실시예 2: 알파-네오아가로바이오스 가수분해효소의 생화학적 활성 및 특징 조사

[0054] 정제된 알파-네오아가로바이오스 가수분해효소의 활성은 다음과 같이 확인하였다. 먼저 아가 다당체를 베타-아가레이스로 처리하여 효소적 처리의 최종산물인 네오아가로바이오스를 생성한 다음, 그것을 기질로 하여 알파-네오아가로바이오스 가수분해효소의 반응 산물을 TLC로 확인하였다 (도면 5 TLC 결과). TLC용매조건 (n-Buthanol:EtOH:Water ; 3:2:2)에서 알파-네오아가로바이오 가수분해효소로 처리된 기질은 갈락토오스와 안하이드로갈락토오스로 추정되는 물질로 분해되는 것을 확인하였다. 갈락토오스 Rf 값은 약 0.46, 네오아가로바이오스의 Rf 값은 약 0.58로 TLC상에서 확인하였다.

[0055] 네오아가로바이오스 가수분해효소에 의해 생성된 산물의 분자량을 측정하기 위하여 액체크로마토그래피-질량분석기(Liquid chromatography-mass spectrometry)를 통해 분자량을 확인하였다. 안하이드로갈락토오스(AHG)는 분자량이 162이고, 갈락토오스 분자량은 180임을 LC-MS분석결과에서 나타난 207.1m/z는 안하이드로갈락토오스에 포름산(formic acid)가 결합된 상태인 분자량으로 안하이드로갈락토오스가 검출되었음을 알 수 있고, 179.1m/z와 225.1m/z는 갈락토오스와 갈락토오스와 결합한 포름산(formic acid)의 분자량으로 갈락토오스가 검출되었음을 확인 할 수 있다. 따라서 LC-MS에서 나타난 바와 같이 네오아가로바이오스 가수분해효소의 반응산물인 두 가지 단당류인 갈락토오스와 안하이드로갈락토오스가 반응산물임을 확인할 수 있다(도 6 참조).

[0056] 실시예 3: 알파-네오아가로바이오스 분해효소 특이적 펩타이드 모티프 서열

[0057] *일반적으로 단백질 모티프 (motif)는 동일한 분자적 기능을 갖는 단백질들에서 일종의 패턴처럼 나타나는 짧은 펩타이드 서열을 의미한다. 이러한 단백질 모티프들은 전체단백질 서열중에서 대개 진화적으로 잘 보존되어 있으며 활성부위를 포함하는 분자적 기능을 대표하는 영역으로 표시되는 패턴화된 아미노산 서열로 표시된다 (Sigrist C.J.A., Cerutti L., Hulo N., Gattiker A., Falquet L., Pagni M., Bairoch A., Bucher P. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform. 3:265-274(2002); Sigrist C.J.A., De Castro E., Langendijk-Genevaux P.S., Le Saux V., Bairoch A., Hulo N. ProRule: a new database containing functional and structural information on PROSITE profiles. Bioinformatics. 2005 Nov 1;21(21):4060-6. Epub 2005 Aug 9; Timothy L. Bailey, Nadya Williams, Chris Mischel, and Wilfred W. Li, MEME: discovering and analyzing DNA and protein sequence motifs, Nucleic Acids Research, Vol. 34, pp. W369-W373, 2006).

[0058] 알파-네오아가로오스 가수분해효소를 규정지을 수 있는 단백질 모티프의 서열을 확인하기 위해, 먼저 알파-네오아가로바이오스 가수분해효소의 아미노산서열을 템플레이트로 사용하여 공공데이터베이스를 NCBI 블라스트 (blast)를 사용하여 검색하여 통계적 유의성(E-value<0.001)을 갖는 60개(알파-네오아가로바이오스 가수분해효소 포함)의 단백질의 아미노산 서열을 수집하였다. 이들 서열을 단백질 모티프 검색 프로그램인 MEME (http://meme.sdsc.edu/meme4_1/intro.html; 사용 파라미터: mode=zero or one occurrence & nsites=50, mwin=8 그 이외의 조건은 default parameter를 사용하였음; Timothy L. Bailey, Nadya Williams, Chris Mischel, and Wilfred W. Li, MEME: discovering and analyzing DNA and protein sequence motifs, Nucleic Acids Research, Vol. 34, pp. W369-W373, 2006) 을 사용하여 알파-네오아가로바이오스 분해효소에 특이적인 모티프를 검색하였다. 그결과 도면 7과 같이 사카로퍼거스 테그라단스의 알파-네오아가로바이오스 가수분해효소는 상동성을 갖는 단백질들의 분석으로부터 얻어진 총 50개의 모티프 중, 13개의 특이적 모티프 (1,2,3,4,5,6,7,14,16,17,20,34,40) 를 갖는 것으로 확인되었다. 따라서 이 13개의 특이적 모티프가 알파-네오아가로바이오스 분해 효소의 활성을 대표하는 모티프라는 것을 확인할 수 있었다. 그 중에서도 모티프 7과 34는 알파-네오아가로바이오스 가수분해 효소 활성을 대표하는 단백질들에 반드시 나타나는 필수적인 2개의 모티프로서 각각 도 8에 나타내었다 (통상적으로 단백질 모티프는 정규표현식(regular expression)을 사용하여 나타냄으로, 도면에서도 같은 방식으로 표현하였다; Sigrist C.J.A., Cerutti L., Hulo N., Gattiker A., Falquet L., Pagni M., Bairoch A., Bucher P. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform. 3:265-274(2002)).

- [0059] 따라서, 13개의 특이적 모티프중 일부를 포함하고 그 중에서도 모티프 7과 34를 반드시 포함하는 단백질들은 알파-네오아가로바이오스 가수분해활성을 가진다고 할 수 있다 (도면 8 참조). 알파-네오아가로바이오스 분해효소와 서열상동성을 보이는 단백질들 중에서 이들 13개의 특이적 모티프들의 일부를 포함하는 한편 모티프 7과 34를 반드시 포함하는 단백질은 10개로 확인되었다 (도면 9-10 참조).
- [0060] 앞서 알파-네오아가로바이오스 분해효소에 발견되는 13개의 특이적 모티프들의 일부를 포함하는 한편 모티프 7과 34를 반드시 포함하는 단백질은 현재 공공 데이터베이스상에서 10개가 발견되고 있다. 그들의 오리진은 각각, *Pseudoalteromonas atlantica* T6c, *Microscilla* sp. PRE1, *Bacteroides plebeius* DSM 17135, *Gramella forsetii* (strain KT0803), *Flavobacteriales bacterium* HTCC2170, *Paenibacillus* sp. oral taxon 786 str. D14, *Ruminococcus* sp. 5_1_39BFAA, *Streptomyces coelicolor* A3이다. 각각의 아미노산 서열은 도면 10에 나타내었다.
- [0061] 도면 11에 나와 있는 서열들은 위에서 서술한 것 처럼 알파네오아가로바이오스 가수분해효소와 서열상동성을 갖는 단백질중에서 13개의 특이적 모티프들의 일부를 포함하는 한편 모티프 7과 34를 반드시 포함하는 단백질 10개를 골라서 다중정렬한것이고, 그들 10개 단백질의 서열은 도면 9-10에 나타내었다.
- [0062] 실시예 4:알파-네오아가로바이오스 가수분해효소 유전자의 단백질 서열 상동성 확인
- [0063] 사카로파거스 데그라단스유래의 가수분해효소의 단백질 서열 상동성 검색을 통하여 유사한 기능을 갖을 것으로 예상되는 후보 유전자들을 확인하였다. 후보유전자들의 구체적인 기능은 알려져 있지 않으나 탄수화물 관련 효소/단백질 데이터베이스인 CAZy(<http://www.cazy.org>)에 의하면 모두 글리코시드 가수분해효소(glycoside hydrolase family 32, GH32)에 속하는 것으로 확인 되었다. GH32 family에 속하는 단백질들이 갖는 알려진 기능은 invertase(EC 3.2.1.26); endo-inulinase(EC 3.2.1.7); β -2,6-fructan 6-levanbiohydrolase(EC 3.2.1.64); endo-levanase(EC 3.2.1.65); exo-inulinase(EC 3.2.1.80); fructan β -(2,1)-fructosidase/1-exohydrolase(EC 3.2.1.153); fructan β -(2,6)-fructosidase/6-exohydrolase(EC 3.2.1.154); sucrose:sucrose 1-fructosyltransferase(EC 2.4.1.99); fructan:fructan 1-fructosyltransferase(EC2.4.1.100); sucrose:fructan 6-fructosyltransferase(EC 2.4.1.10); fructan:fructan 6G-fructosyltransferase(EC 2.4.1.243); levan fructosyltransferase(EC 2.4.1.-)이다. 따라서 GH32 family에 속하는 일부 단백질들이 네오아가로바이오스 가수분해능과 같은 분자기능을 갖는 다는 것은 본 발명에서 처음으로 보고하고 있다.
- [0064] 사카로파거스 데그라단스(*Saccharophagusdegradans*2-40)유래 알파-네오아가로바이오스 가수분해효소(이하 α -NABH)의 아미노산 서열의 상동성 검색(블라스트 검색-첨부화일 참고)으로 얻어진 서열 중 50%이상의 상동성을 갖는 단백질 서열을 블라스트의 E-value에 의해 나열하면 다음과 같다(다음은 Uniprot 데이터베이스 번호를 사용, 괄호안은 유래 미생물 및 % identity).
- [0065] 1. Q15UF2 (*Pseudoalteromonas atlantica* (strain T6c / BAA-1087), 70%),
- [0066] 2. Q93PB3 (*Microscilla* sp.PRE1.,59%),
- [0067] 3. B4CY74 (*Bacteroides plebeius* DSM17135.,60%),
- [0068] 4. AOM245 (*Gramella forsetii* (strainKT0803).,56%),
- [0069] 5. A4AR39 (*Flavobacteriales bacterium* HTCC2170.,57%),
- [0070] 6. C6J3P3 (*Paenibacillus* sp.oral taxon 786 str.D14.,58%),
- [0071] 7. C6JDD4 (*Ruminococcus* sp.5_1_39BFAA.,58%),
- [0072] 8. C6J313 (*Paenibacillus* sp.oral taxon 786 str.D14.,57%),
- [0073] 9. Q15XP8 (*Pseudoalteromonas atlantica* (strainT6c/BAA-1087),55%),
- [0074] 10. Q9RKF6 (*Streptomyces coelicolor*,56%)
- [0075] 실시예 5:대장균에서의 발현 및 크기 확인
- [0076] 앞서 획득한 50% identity이상의 서열 상동성을 갖는 단백질중에서 가장 높은 서열상동성을 갖는 슈도알테로모나스 아틀란티카(*Pseudoalteromonas atlantica* T6c)유래 단백질(Q15UF2,)과 가장 낮은 스트렙토마이세스 코엘

리컬러(*Streptomyces coelicolor*A3) 유래 단백질(Q9RKF6)을 각각 클로닝하여 α -NABH 활성을 갖고 있는지 확인하였다. 먼저, 이들을 암호화하는 유전자의 염기서열을 각각 대장균 발현벡터인 pET21a (노바젠, 미국)에 삽입하였다(이하 슈도알테로모나스 아틀란티카 유래 α -NABH 유전자를 포함하는 발현벡터를 pPsAGAJ, 스트렙토마이세스 코엘리컬러 유래 α -NABH 유전자를 포함하는 발현벡터를 pScAGAJ라 명명함). 재조합 α -NABH가 대장균에서 성공적으로 발현하는 것을 확인하기 위하여, pPsAGAJ와 pScAGAJ를 발현용 대장균인 *E. coli* BL21(DE3)에 형질전환시킨 후에, 50 mg/L 농도의 앰피실린 항생제를 함유한 고체배지에 도말하였다. 위의 형질전환에 의해 획득한 콜로니를 50mg/L 농도의 앰피실린 항생제가 첨가된 루리아 버타니(LB) 배지에 접종한 후, 37에서 하루 동안 진탕 배양하여 균체를 확보하였다. 그 후 발현확인을 위해, 형질전환체를 50mg/L 농도의 앰피실린 항생제가 첨가된 루리아 버타니(LB) 배지에 접종한 후, 37에서 OD₆₀₀=0.5~1.0까지 진탕배양하고, 0.5mM/L의 농도로 IPTG를 첨가하여 180rpm에서 4시간 동안 발현을 유도하였다. 배양액은 원심분리(12000rpm, 4, 10분)하여 균체를 회수하였고, 회수한 균체는 20mM 트리스 완충액(Tris-HCl pH7.4)에 현탁하여 초음파 파쇄기로 파쇄한 후, 12% SDS-PAGE에서 크기를 확인하였다. 크기가 확인된 파쇄한 현탁액을 15분간 원심분리하여 상층액을 조효소액으로 사용하였다.

[0077] 실시예 6:알파-네오아가로바이오스 가수분해효소의 생화학적 활성확인

[0078] 정제된 α -NABH의 활성은 다음과 같이 확인하였다. 먼저 아가 다당체를 베타-아가레이즈로 처리하여 효소적 처리의 최종산물인 네오아가로바이오스를 생성한 다음, 그것을 기질로 하여 α -NABH의 반응 산물을 TLC로 확인하였다. TLC를 통한 확인은 Silica gel 60 TLC 평판에 반응액을 1ul 점적하고, TLC용매조건(n-Buthanol:EtOH:Water ; 3:2:2)에서 전개하였다. 전개한 TLC 평판은 1차 처리 용액인 황산(10%(v/v) H₂SO₄ in Ethanol)으로 처리한 후 건조하고, 1차 처리된 평판은 다시 2차 처리 용액인 나프토레조시놀(0.2%(w/v) naphthoresorcinol in Ethanol)로 처리하였다. 이렇게 처리된 TLC 평판은 건조 후 가열하였다.

[0079] 상기 실시예 4 내지 6의 결과는 다음과 같다.

[0080] 발현용 균주인 *E. coli* BL21(DE3)를 사용하여 형질전환한 슈도알테로모나스 아틀란티카와 스트렙토마이세스 코엘리컬러 유래 α -NABH의 발현과 크기의 확인은 12% SDS-PAGE를 통하여 확인하였다. 슈도알테로모나스 아틀란티카와 스트렙토마이세스 코엘리컬러 유래 α -NABH의 예상 분자량은 각각 40.7kDa과 41.1kDa정도이며, 예상 분자량과 동일한 것으로 확인하였다(도면 12, 도면 13).

[0081] 또한 분해산물을 확인하기 위해서 표준물질로 D-갈락토오스를 사용하여 예상 분해산물을 확인하였다. α -NABH로 처리된 이당류인 네오아가로바이오스는 D-갈락토오스와 동일한 Rf 값에서 확인된 갈락토오스와 안하이드로갈락토오스로 추정되는 물질로 분해되는 것을 확인하였다(도 14).

[0082] 실시예 7:결정화를 위한 알파-네오아가로바이오스 가수분해효소의 발현 및 정제

[0083] 재조합 α -NABH의 성공적인 발현과 구조결정을 위해 먼저 α -NABH가 재조합된 벡터를 *E. coli* B834 (DE3) 대장균에 형질전환하였다. 형질전환시킨 대장균을 50mg/L 농도의 앰피실린 항생제가 포함된 고체배지에 도말하였고, 획득한 단일콜로니를 50mg/L 농도의 앰피실린 항생제가 포함된 루리아 버타니(LB)배지 10ml에 접종한 후, 37℃에서 180rpm으로 하루 동안 진탕배양하였다. 배양된 균체 10ml을 50mg/L 농도의 앰피실린 항생제와 L-selenomethionine (Se-Met)이 포함된 M9 최소배지에 접종 후, 37℃에서 OD₆₀₀=0.5~1.0까지 진탕배양하고, 0.1mM의 isopropyl β -D-1-thiogalactopyranoside(IPTG)를 첨가하고 180rpm으로 발현하였다.

[0084] 발현된 균체는 초음파 파쇄기(Sonifier 450, Branson, USA)를 사용하여 파쇄한 후, 4℃에서 15000rpm으로 1시간 동안 원심분리하여 0.45 μ l 여과지(Sartorius stedim biotech, Germany)를 사용하여 여과하였다. 이렇게 추출한 조추출액은 3단계 과정으로 정제하였다. 첫 번째 과정으로 HisTrap HP 5ml 컬럼(GE Healthcare)을 사용한 Immobilized nickel ion-affinity chromatography (IMAC) 방법으로 정제하고, 두번째 과정으로 HiTrap Q FF 5ml 컬럼을 사용하여 anion-exchange chromatography (AIEC) 방법으로 정제하였다. 마지막으로 HiPrep 26/10 desalting 컬럼을 사용하여 최종 완충용액을 20mM Tris-HCl, pH 7.4 완충용액으로 바꿔주었다. 최종 단백질 농도는 9mg/ml로 농축하였다.

[0085] 실시예 8:알파-네오아가로바이오스 가수분해효소의 결정화 및 구조결정

[0086] 상기의 발현된 α -NABH는 피닉스 로봇(Phoenix robot)으로 sparse-matrix 방법(Jancaric, J. and Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409-411)을 이용하여 결정화 하였다. 결정화시, 단백질과 침전용액은 각각 0.5ul씩 섞어 50ul의 저장용액을 넣은 96-well plate에 sitting drop방법으로 셋업 한 후, 실온에서 결정화하였다. α -NABH는 2.4M sodium malonate, pH7.0 조건에서 약 1주일 후에 120 X 120 μ m크기의 결정으로 확인하였다(도

16).

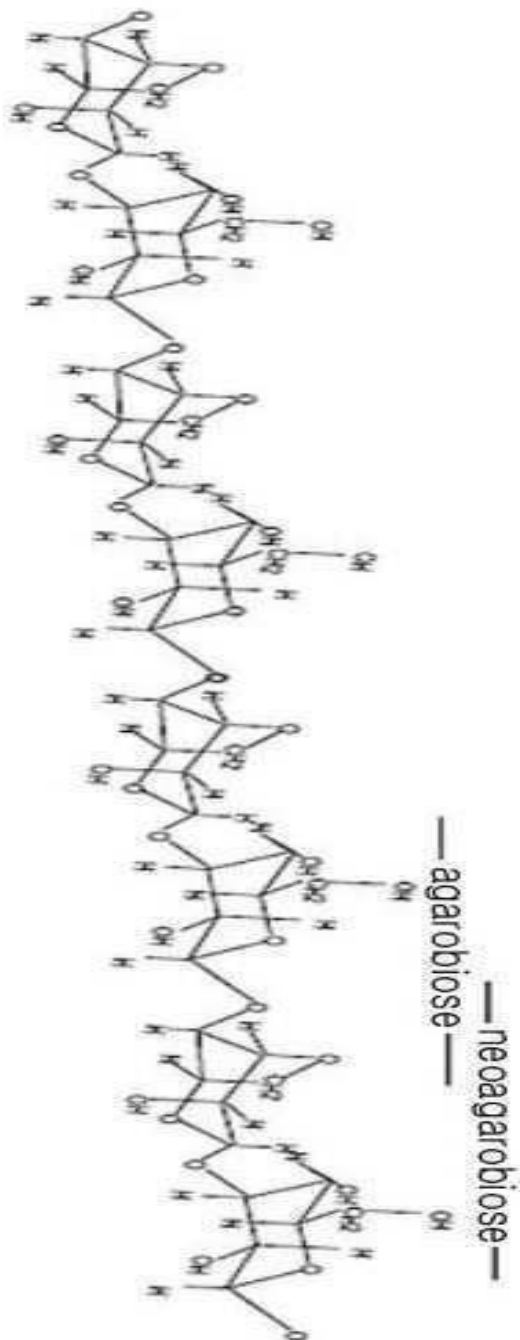
[0087] α-NABH의 SAD(singlewavelength anomalous diffraction data) 데이터는 내동제(cryoprotectant solution) 조건에서 로렌스 버클리 국립 연구소(Lawrence Berkeley National Laboratory)의 advanced Light Source (ALS)를 사용하여 수집하였고, 모든 데이터는 ADSC Q315 검사기(detector)를 사용하여 이미지를 확인하였다(도 17). 데이터 스케일은 이중 가장 좋은 데이터 세트를 사용하여 X-ray 파장이 0.9784Å에서 $\Delta\phi=1^\circ$ 씩 90° 범위의 데이터를 수집하였다. 수집한 모든 데이터는 HKL-2000프로그램을 사용하여 프로세스를 하고, CCP4(Collaborative Computational Project, Number 4, 1994)프로그램(Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760-763)을 이용하여 메튜 계수(Matthews coefficient)를 계산하였다(도 15).

[0088] 스케일된 SAD(singlewavelength anomalous diffraction data) 데이터는 CCP4(Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760-763)와 SOLVE/RESOLVE(Terwilliger TC, Berendzen J. (1999). *Acta Cryst.* **D55**, 849-861)프로그램을 이용하여 구조결정한 후, 리파인(refine)을 통해 초기에 결정된 구조를 2.0Å의 해상도로 개선하였다. 모델구축은 Chimera(Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. (2004). *J Comput Chem.* **25**(13), 1605-1612.)와 REFMAC5(Terwilliger TC, Berendzen J. (1999). *Acta Cryst.* **D55**, 849-861)을 사용하여 최종적으로 R_{free} 값이 0.22인 모델을 구축하였다.

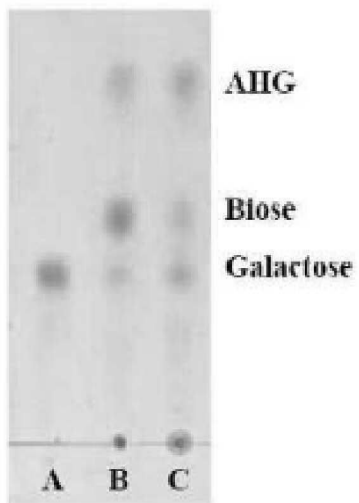
[0089] 상기와 같은 실시예 7 및 실시예 8에 의한 결정화 결과는 도 15 내지 도23에 도시하였다.

도면

도면1



도면2



도면3

```

ATG ACC GAT TGA AAA GTA AAT AAA AAA TTG ACT AAA GGT ACC CIG CCA CCC ALA GAG CGC < 60
M S D S K V N K K L S K A S I R A C E R

GGC TAC GAT GAA AAG GGC CCG GAA TGG CTG TCT GAG CTT GAC ATT ACC CCA CCA AAA GGC < 120
G Y D E K G P E W L F E F C I T P C K G

GAC TTA GGC TAC GAA GAA GGC GTA ACT CGT CGA GAC CCC AGC GCA GTA TCA AAG GTG GAC < 180
C I A Y R E G V T R R D F S A V T K V C

GAC GAA TAT CAC GTT CCG CAC ACC AAG GGC GAA GGT GAA ACA GTA GGC TTC GGC AGC GAC < 240
C E Y E V W Y T K G E G E F V G F G S C

AAG CCC GAA CAG AAA CTC CTC CCA TGG CAT AAA ACA GAA GTT TCG CAG CCC ACG TCG AAA < 300
N F S D K Y P P W D K T E V W H A T S X

GAC AAG ACT ACT TGG AAA GAA ATT GGC CCT GGC ACA GAA GGC GGC GCA GGT GGC GCA TAC < 360
C K I C W K C C S P A C Q R C A A G A Y

GAC GAC CTT CCA CTC CTC ACC CCC GAA CCG CTC GGC GAT AAG CCG ACC TAC TAC CTC GTA < 420
C C R A V F T P E V C R H N G T Y Y L V

GAC CAA AGC CCA AAA GGC CCG TAC TCA AAG CCA GGC GTA GAG CAT ATA CCC ACC GCA TAC < 480
Y Q I V K A P Y L N R S C C D C A C A Y

AGC GAT TCG CCC TTT GGC CCA TGG ACC AAA TCG GAT GGC CCA ATT TTA AGC CCA GAA AAT < 540
S C S P E G P W C K S D A F I C S P C N

GAC GGC GTT TGG GAT ACC GAC GAA GAC AAT CCA CTT TTA GTA AAA GAG AAA GGC AGT TTC < 600
C G V W D I C E D K R F C V K E K G S F

GAC ACC CAC AAA GTA CAC GAC CCC TCG TCA ACC CTT CTC AAG AAT CCT TTC TAC CTC TAC < 660
C S A K V R C P C C M F F N H R F C L Y

TGC AAA GGC GAG ACT ATG GGC GAA AGC ATG AAG ATG GGC GGC AGA GAA ATG AAA GAC GGT < 720
Y K C E T M G E S M N M G G R E I K H G

GTA GCC ATT GGC GAC CCG CCA CTT GGC CCG TAC ACC AAA AGC GAA TAC AAC CTT ATT ACC < 780
V A T A D S P I C P Y T K S P Y N P T T

AAC AGT GGC CAT GAA GTT GGC GTA TGG CCC TAC AAA GGT GGA ATG GGC ACC AGC GTA ACC < 840
N S G E E V A V W P Y K G C Y A C M L I

ACC GAC GGC CCA GAA AAA AAC ACC TCG CAG TGG GCA GAA GAC GGC ATT AAC TTT GAC ATT < 900
T C G P E K N T C Q W A E C C C N P C I

ATG TCG CMT ACA AAA GGC GCA CCA GAA GCA GCA GGT CTT TTT AGA CCA GAA AGC GAT AGC < 960
P S F T K G A P F A V G F F R P E S C S

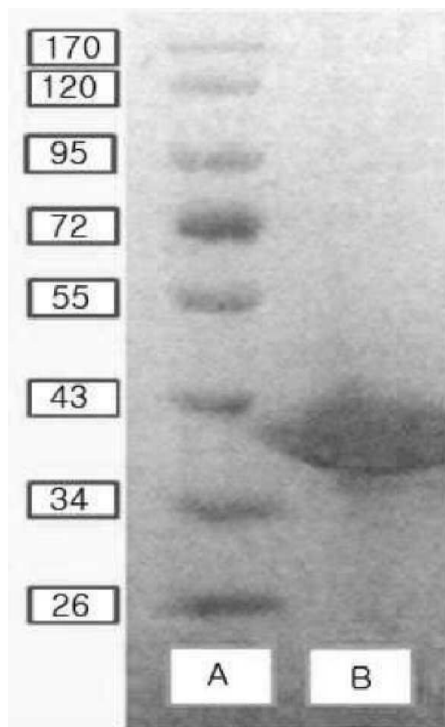
GAC GAC CTT ACA AGC GGC ATT GAA TGG GGG CCA ACC CAC AAG TAC GAC GGC AGC GGG AAC < 1020
C C P I S G I E W G C S S X Y D A S W N

TGC AAC TAT CCA TGC CTC CTC AAA AGC CTT CGA CAA CTT TTA GAT GCA GGT AGC CAC CAG < 1080
W N Y L C F F K C R R Q V L C A G S Y Q

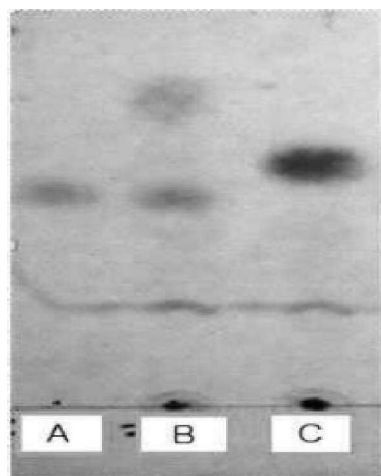
CAA ACA GGC GAT TCC GGA CCA GTA TAA < 1107
Q T S D S G A V *

```

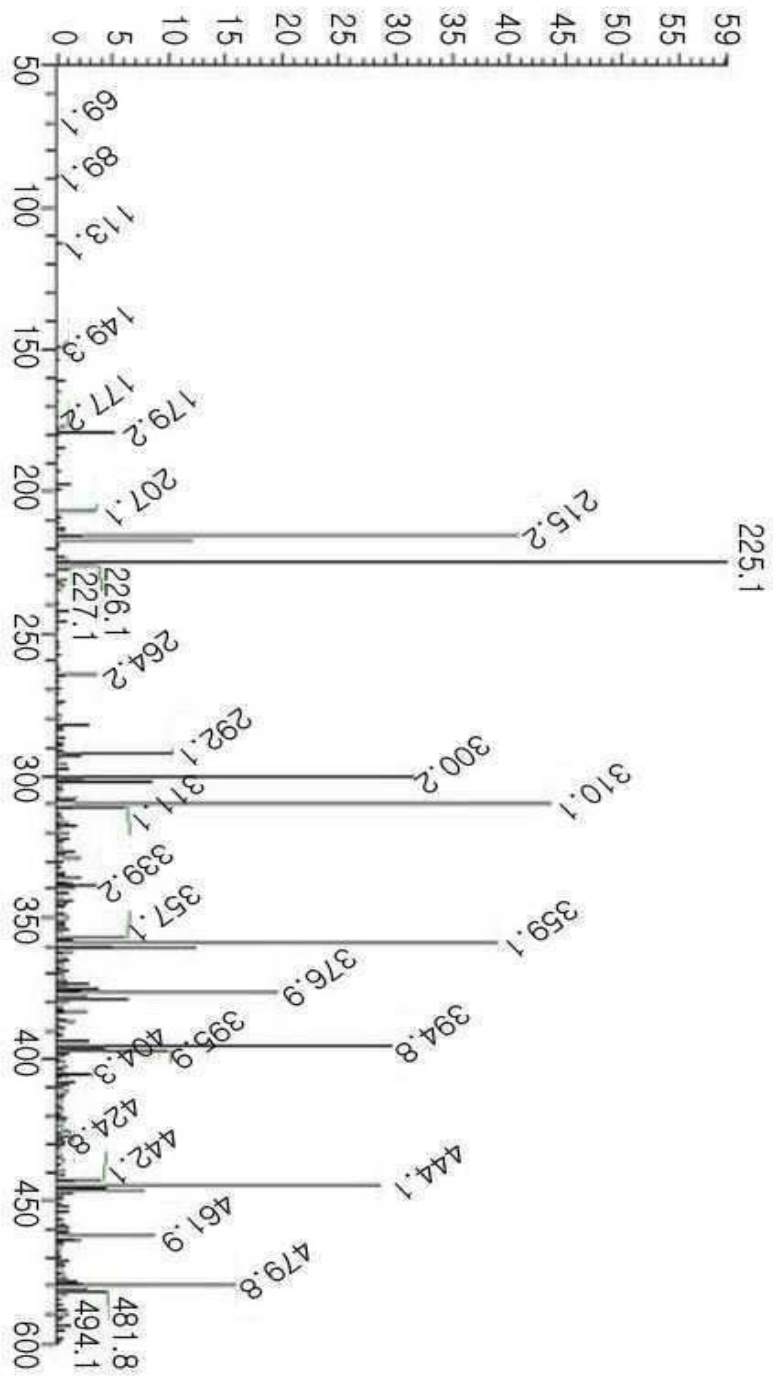
도면4



도면5



도면6



모티프 정규표현식

Motif 1	[V]GVAV[AS]DSPxGP[WFI]KET[KRI]
Motif 2	[ATN]PE[IV][LY]EH[DN]GK[YFY]LM[YFY][YF]
Motif 3	VGVATTI[AS]D[SD]PI[LE]GP[YFT]KSEXNP[IV]
Motif 4	NSGHE[TV][C][VL][WF]PQ[YFY]K[GD]G[IM]A[AS]L[LI]TTDGPENKNTQ[YWF]A[PE]DGI[V]NF[ED][MK][AS][VH]K[GN][AP]P[EH]A
Motif 5	[AY]TSKD[GL][VI][TN]WKE
Motif 6	VxRR[DE]P[SP]S[VI][LL]K[EV]D[GD]xYY[VM]WY[ST]
Motif 7	DNRI[F][LI]TV[KV][SH][KQ]G[DS]FDS[H]K[VH]DP[CT][L][F][I][PP][FY][NR][GN][KQ]RF[YW]LYYKGE[QRT]MGEEM[NT][FAM]GGRIET[IT][KR]
Motif 14	GFGLTG[IG][DI][DI][PE][DN]IDA[KVFPWDR][CS][DE][IV]W[YH]
Motif 16	[DEK][QAEK]PI[LL][RS]P[GS][DN][IDTN]G[E]S[WD][GD]
Motif 17	[VL]GRGE[KP]G[SA][YF]DDR[SA]VF
Motif 20	G[LP][YFN]R[INPST][ELP]DA[DES][DKR][ENS]P[LT][EA][G][LI][RE]WGL[CS]H[DKN][YV][INDH][SAN][SK][WD]
Motif 34	Q[TTVA][VI][OK][SA]PY[TLN][WV][RSTV]
Motif 40	PLKGDLAYEEG

면8도

도면9

```

>tr|Q9RKF6|Q9RKF6.STB00 Putative uncharacterized protein SC03481 OS=Streptomyces coelicolor GN=SC03481 PE=4 SV=1
MTMPAESAAIVRGRVPADEPFCDHFITFVQGLPEVGHRRDPSSVLTVDPRHWYT
RSTGETDGFSTGDMKAKYFPWQDSEIWHATSD0GATWVEQGHALGRGEPSSVDORSYTP
EVLHQGFVLYVDYIRSPVALRFEESIAMAKAPDQGWBPQPLLRQDADEWAGEE
DNRLSVYSQGSFDSKHVDPLTFPQGRFTLVYGEOMGEFSAGGRITFWGLAIAD0IE
GPRVGRAMPVYNSGHETCYWYRGDGIAMLLTIDGERNTIQAPDGINFEIMAHIANPP
VATGPALPDHAYAPLDVRYWGLCHDYTASWYIQGFADDERDKMFTIRGLSPETAESAE
ADALPAPHEK
>tr|Q15XP8|Q15XP8.PSEAB Putative uncharacterized protein OS=Pseudomonas atlantica (strain Tbc / BAA-1087) GN=Pat1_0812 PE=3 SV=1
MKWVLIGALSLTJACNNTMESETPAASTSVAPKAYSQADYAVLAISDPTMSAASK
BALARNYHKNEDQFGEERHETLEGFADFDPVSPRPSKYLVNDVFTWYTRITGEAYG
FGIGDPMKVPWQDAEIWYATSKDGHQDEKGRALGTDVGFDOGSYFPEVLAHQIM
FYLVDQAIKAPVLRNTKNTYAMAYSSSPDGFWRKLDAPLITASDTGEWAGTIDNRFLYKS
QGFDSKHVDPTLLFVBDKFYLWKGGEWGEKTAAGREIRWGYAIADNPQSPYIKSAF
NPITHSGHELQWYQDGIAYSSHDPGEKTIQVAPDGINFEVMSYLPSPSAIGLVES
LDKAYPTAGLEWGLYHEWYIPPGKTIWMSGIINHKEFSNAPRYSMDGDSIGNTQKNS
GEFINKQSDH
>tr|D6J3J|D6J3J.9BAOL Putative uncharacterized protein OS=Paenibacillus sp. oral taxon 786 str. D14 GN=POTG_02762 PE=4 SV=1
MKMKESAAITIRAKVYTKMAWFOEFKYEPEVEGLVVEEGLHRRDPSSYIKVQDLVYWYVT
KSYGEAYGENTIDPDSKYFPWQDSQVWYATSPDITWTERGMAYGRGKSYVDORSYFTR
ELIHHQKFYLVDYIEQVLLRKYENIAMAYAESPEGPWRKLDAPILPRANWGEWFGD0
DNRLTYKHGDFDSLKHVDPLTFVYNNQFWLVYGEQKGEENFGGRITKWGYAISDHPF
GPYKSEVNPVYNSGHETLLWHYRGMAALLTIDGERKTIQVAPDGINFEIEAVIKNPP
EAGPRTIEDDRPLEGIRWGLCHNYHTKNNHLKFRITVEDYKHWYANKINPEVLWLKQ
KAKQKQD
>tr|D6J0D4|D6J0D4.9F1PM Putative uncharacterized protein OS=Paenibacillus sp. 5.1_398FAA GN=PSAG_01951 PE=4 SV=1
MKLSASAKRAIERNYVQSCDFDQFKEGSEIGIYGERGIRHRRDPSSYIKVQDVFYWYSR
SVGRKGFHTGDEEAKYFPWQDILWYATSKDGFWRKEKAPVYRGEKAYDORSYFTR
LLEVERKYLVDYQWQRPVYNSFESIAIAYADSPHGFYSKEPILITPKDGIWEGEED
NRPAYKKGSFDSKHVDPLTFAPRAKFYLWKGEPWGEELWYGRRETKWGYAIDNILLG
PYHSEVNPVYNSGHETLWQYNGGIAARLTDYERKTIQFAEDGINFEIKSYIKQGRE
ACDPVHLESQNPULKMEWGLCHQVSDQVGFKFPDIDENWQKYVINHEWME
>tr|D6J3J|D6J3J.9BAOL Putative uncharacterized protein OS=Paenibacillus sp. oral taxon 786 str. D14 GN=POTG_02958 PE=4 SV=1
MKKESAAITKRAKQVTKNSDFQCFERISPYGLGVEEGYHRRDPSSYIKVQDLVYWYTKT
EGEAAIGFTIGDPMKVPWQDSEWYATSPDGNWYTERGAILGRPKSYVDORSYFTRPEI
LHEDKFYLVDYQSPYTLRLTLENIAMAYAESPEGPWRKLDPEILPRSNDEWFGD0N
RLTYKHGDFDSLKHVDPTLTFVYNNQFWLVYGEQKGEENFGGRITKWGYAIDHPEGP
YKSEVNPVYNSGHETLWQYRGMAALLTIDGERKTIQVAPDGINFEIEAVIKNPEEA
PGVPRTPDAPRPLEGLRWGLCHNYHSKWVYIDNYTDESFKHWYMKINPEELWLKSP
KRNQD
>tr|A4AR93|A4AR93.9FLAO Putative uncharacterized protein OS=Flavobacteriales bacterium HT022170 GN=FB2170L15408 PE=4 SV=1
MKKEKEVEKKSFEISNEQDLNLGITNDPSLSSAKRALKPNIGNEWIFERSALKPLKG
DLAVEEAYRRDPSSLIENKAYWYWSKLSGESQFGGDENDKVPWQDQIHWATSY
DGIWKEEGPAILVRGNGEFDORSYFTEIMKYENWYLCYQTVDSPTVRYKNGVGLAW
SOSPDPGWTKSKDPLISPADNNGIWKGLDNPFSEYKGFDFDSKHVDPCIIIPYNGKFLY
YKGEOMGEELTFGRHQLRHGLAIENPDQPYIKSPVPISNSGHEICWYWKDGIASLIT
TDGERKTIQWANDGINFEIMAYIKGAPHAIGLNRWYDNGKPTFVLRWGLSHWYWNDSY
QSIWRFSSQLKTVHTAKGEKAK

```

도면10

```

>tr|A0N245|A0N245_GBAFK_Secreted protein OS=Gramella forsetii (strain K10803) GN=GF0.1720 PE=4 SV=1
MKTKKLVIGLMAITLVSCHNTSEKISAEKTEKPEISQVODLGLITNPQSLSAASVR
ALDWFDVGNEMFI EFSELQPLKGLAYEEGVPRDPSALLKHDKYVWYTSVPTQDF
GGDIENOKYFPMDRDIWATSEDAWTEEDGIAVABEKGEVDORSFTYELMEDKQKY
VLSVQTIQSPVWRTKNQGLAWADSPNGPWTKSQPI LSPADNGIWKGEEDNFAVEKK
GDFDSHKYHDPCLIPYNBKYVLYKGEOMGEETFGGRQIRHGVAILDNKQPVYKSAVN
PI SNSGHEICVWKNNGGIALSLITTDPERKNTIQMAPDGNFGIKSYIKGAPHAIGLNREL
NIDSEEPGAILGNALSHIYNISDVOSIMFSSKRTSHAKGEKAE
>tr|B5CY74|B5CY74_GBACE_Putative uncharacterized protein OS=Bacteroides plebeius DSM 17135 GN=BACPLE.01671 PE=4 SV=1
MRKLI FFAIGMISLLAAGNTGNITQTIAYDITQNVDERKADSLGPKGKLSAMKPAKQW
ENHDKWFEETXMEPLKGLAYEEGVPRDPSALLKHDKYVWYTSVPTQDFAGDIE
KOKYFPMDRDIWATSKDGLTWKEEGIAVABEKGVVDORSFTTEVMEKQYVLCYQ
AYKSPVTVRKNTIQMAADSPGALWTKDQPVLEPSDTEEDNDRKRYKSGDFDS
HKYHDPCLIPYNBKYVWYKGEOMGEETWGRPEIKHGVAILDNKQPVYKSEVPI SNS
GHEVCWPKYGGIALSLITTDPERKNTLQWSPDGNFELMSYKGAPHAIGLNSDAEKE
PTEILRNGLTHIYNISDVOSIMFSTWLTQTHKGESEK
>tr|D99P93|D99P93_GSPHI_Putative uncharacterized protein OS=Microscilla sp. PBEI PE=4 SV=1
MKIMWYKSTIGAMTALLFASQDTPDEEDNEITDOKIDVGLITNPDL SAKSRA
LQMDPIGNWYGEFQYDQKGLAYEEGVPRDPSALLKHDKYFVWYTSKSTKTDGFAQ
DI EKDYFPMDRDIWATSEDAWTEEDGIAVABEKGVVDORSFTTEVMEHGTLYL
VQITKSPVWRTKNQGLAWADSPNGPWTKSQPI LSPADNGIWKGEEDNFAVEKKD
FDSKHYHDPCLIPYNBKYVLYKGEOMGEETFGGRQIRHGVAILDNKQPVYKSPVPI
SNSGHEICVWKEHGGIALSLITTDPERNTIQMAPDGNFIDIAAIPGAPHAIGLNRLNN
EDPFGIFGMLTHEVTVQWQYIPFEERKRTTHVABETGSKSSDE
>tr|D15U21|D15U21_PSEAB_Putative uncharacterized protein OS=Pseudomonas atlantica (strain Tbc / BAA-1087) GN=Pat.1368 PE=4 SV=1
MTQAKQSLATKRAIEBQVDSKADWIMETFEISPKGQFAEEGVPRDPSVLSVDQY
VQWWTGATGETVGFSTNPDKYFPMDL TEVWHAISDQETWHEGPAITRATPGAFDS
RAYTPEVLAHNKQYVLYQTVQVPTNRQIEQALIAHDSFGPWLKSTAPILSPSMDS
EMRGEONRPHYKSGDSKHYHDPCLIPYNBKYVLYKGETMGEENFGGEIKHGVAIL
DADILAPYKSEVPI SNSGHEVWVWQGVASLLITDPERKNTIQFADGINFDIKAS
HIKGAPEWALYPPSNI EDSPPGLHWLCHKYDSSNNWVLCRYALKQYLDAGTQNS
N
>tr|D21H82|D21H82_SACD2_Glycosyl hydrolase family 32, N terminal OS=Saccharophagus degradans (strain 2-40 / ATCC 43961 / DSM 17024) GN=Sde_2657 PE=4 SV=1
MSOSKYNKKLSKASLRALERATDERGPEWLFETDITPLKGLAYEEGVPRDPSVLYQD
DEVHWYTKGEETVGFSDNPEKYPMDKTEVWHAISKDITWKEIGPALIRGAGAY
DQAVFTPEVLAHNQYVLYQTVKAPYLNRSLHIALIAYDSFGPWTKSQPI LSPEN
DQWDTQENDRFLYKSGDSKHYHDPCLIPYNBKYVLYKGETMGEENMGPEIKHGS
VAIADSPLPYKSEVPI SNSGHEVWPKGMAITLITDPERKNTQMAEDGINFDI
MSHIKGAPEWGFPEDSDDPI SAILWGLSKYDASNNWVLCFTKTRQYLDAGSYQ
QIDGDSAY

```

도면11

CG13JC6313.9BAC1/1-367 1 --- MCNKKEKSAATIRAKQVTKNADHE 21
CG13JC6313.9BAC1/1-364 1 --- MRKESATHRAKQVTKNSDFE 21
Q9NKP6/Q9NKP6_STKCO/1-370 1 --- MTNPAESATVGRGRYPADPEF 21
CG10D4/CG10D4_SHRM/1-353 1 --- UKLSAETKRIERNVYQSCDF 21
Q13UJ7/Q13UJ7_P5EA6/1-361 1 --- MTQAQKQKSLATERAIERGYDSKNADNM 21
Q13UJ7/Q13UJ7_P5EA6/1-361 1 --- MSDQSVNKKLSKASIRAIERGVDEKCPDF 21
AAAR33/AAAR33_STLAC/1-382 1 --- MKKEKEVVEKKSFESIHNQIDNIGCTNRDSSSKRALKWFNTCNEF 21
AOM245/AOM245_GRAFX/1-407 1 --- MKNKTKALVIGLMTATLYSCNHTSEKISAEKTEKPEISQVQIDRLGINTPDSLSAASVRLQWPDVGCNEF 21
Q13P89/Q13P89_SPH/1-408 1 --- MKIMNVKYSTICAMTALFVFCCTQPPREEENFEITQDKIDYLCITPNOKLSAAMKRLQWPDICNEF 21
B5CY74/B5CY74_P5EA6/1-402 1 --- MRKIIIFAAQNMSSLAACTCNTQTIAVDQTNQYDEKADPSLCKPNKGLKSAKRAHKWENHONK 21
Q13XP8/Q13XP8_P5EA6/1-436 1 --- IMXYWLIGALESLTACNNTTMESETRAASTV SAPNKAYSQADYALAI SQPDMSAASKRALARNYHKNEQ 21

CG13JC6313.9BAC1/1-367 24 CHTKYP --- VEGLCYSETHHHPHSPVHKDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG13JC6313.9BAC1/1-364 24 CTRTSP --- VQGLCYSETHHHPHSPVHKDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q9NKP6/Q9NKP6_STKCO/1-370 24 CDRKIFP --- VQGLGCPVYHHPHSPVLTQDGRHMTNTRATSTCTDCTGCDPMASVYHNSCISHTSOS 21
CG10D4/CG10D4_SHRM/1-353 23 CDRKLSG --- LEGICVYSETHHHPHSPVHKDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 29 IETETSP --- LKQDTAFLECYHCDTSTVSKYQDVCWCTKGTGLTVEGSTPNDDKVTWLTVEVHST 21
Q13UJ7/Q13UJ7_P5EA6/1-361 31 FIDITP --- LKQDILAYEYHCDTSTVSKYQDVCWCTKGTGLTVEGSTPNDDKVTWLTVEVHST 21
AAAR33/AAAR33_STLAC/1-382 10 FLSALK --- LKQDILAYEYHCDTSTVSKYQDVCWCTKGTGLTVEGSTPNDDKVTWLTVEVHST 21
AOM245/AOM245_GRAFX/1-407 11 FLSLQ --- LKQDILAYEYHCDTSTVSKYQDVCWCTKGTGLTVEGSTPNDDKVTWLTVEVHST 21
Q13P89/Q13P89_SPH/1-408 72 GLEWVE --- LKQDILAYEYHCDTSTVSKYQDVCWCTKGTGLTVEGSTPNDDKVTWLTVEVHST 21
B5CY74/B5CY74_P5EA6/1-402 59 FLYKNEF --- LKQDILAYEYHCDTSTVSKYQDVCWCTKGTGLTVEGSTPNDDKVTWLTVEVHST 21
Q13XP8/Q13XP8_P5EA6/1-436 74 FGFHET --- LFGDFAFDPNSSKLSKLEKQNVYFWTTRTTEIAVSGCTCPDNMAVYHNSCISHTSOS 21

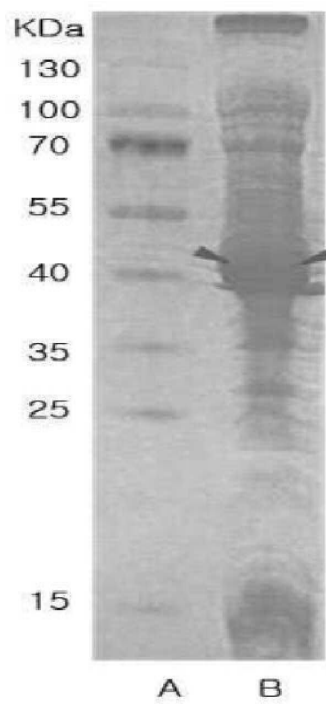
CG13JC6313.9BAC1/1-367 95 TWTTR --- TRMVEGRPKASVQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG13JC6313.9BAC1/1-364 93 VNTTR --- TRMVEGRPKASVQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q9NKP6/Q9NKP6_STKCO/1-370 95 ATNVY --- QDRLCTEPSSVQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG10D4/CG10D4_SHRM/1-353 94 YWKEK --- PPHVYAEERAAQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 101 ETWHEK --- QDRLCTEPSSVQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 101 ETWHEK --- QDRLCTEPSSVQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
AAAR33/AAAR33_STLAC/1-382 123 WTKKE --- PPAVYVYAEERAAQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
AOM245/AOM245_GRAFX/1-407 146 WTKKE --- PPAVYVYAEERAAQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13P89/Q13P89_SPH/1-408 144 WTKKE --- PPAVYVYAEERAAQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
B5CY74/B5CY74_P5EA6/1-402 141 LTKK --- QDRLCTEPSSVQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13XP8/Q13XP8_P5EA6/1-436 147 YHWK --- KRLGTDVYSGQDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21

CG13JC6313.9BAC1/1-367 165 ITF --- PANNVEFQDQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG13JC6313.9BAC1/1-364 166 IFSNN --- FEFQDQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q9NKP6/Q9NKP6_STKCO/1-370 167 LTRAD --- DQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG10D4/CG10D4_SHRM/1-353 107 LTRAD --- DQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 174 LTRAD --- DQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 174 LTRAD --- DQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
AAAR33/AAAR33_STLAC/1-382 376 IFSNN --- FEFQDQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
AOM245/AOM245_GRAFX/1-407 376 IFSNN --- FEFQDQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13P89/Q13P89_SPH/1-408 217 IFSNN --- FEFQDQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
B5CY74/B5CY74_P5EA6/1-402 214 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13XP8/Q13XP8_P5EA6/1-436 220 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21

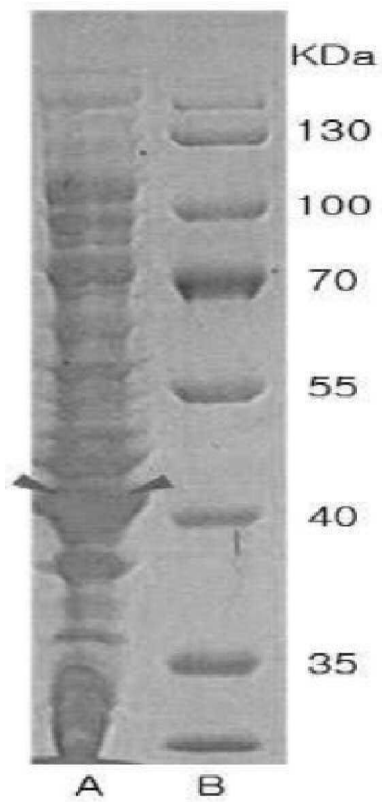
CG13JC6313.9BAC1/1-367 241 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG13JC6313.9BAC1/1-364 239 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q9NKP6/Q9NKP6_STKCO/1-370 241 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
CG10D4/CG10D4_SHRM/1-353 240 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 247 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13UJ7/Q13UJ7_P5EA6/1-361 247 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
AAAR33/AAAR33_STLAC/1-382 269 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
AOM245/AOM245_GRAFX/1-407 290 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13P89/Q13P89_SPH/1-408 292 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
B5CY74/B5CY74_P5EA6/1-402 287 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21
Q13XP8/Q13XP8_P5EA6/1-436 293 LBSOT --- EHEQDNLCTKKHAGDQVYHWTKTVAEVAINTCDPDSVYHWQDSBYVYATSP 21

CG13JC6313.9BAC1/1-367 312 DLSLE --- DLRGQDNYHKNWYIQMTYDSEKHYMYMKIWPCLWLCKKAKQK 21
CG13JC6313.9BAC1/1-364 310 DLSLE --- DLRGQDNYHKNWYIQMTYDSEKHYMYMKIWPCLWLCKKAKQK 21
Q9NKP6/Q9NKP6_STKCO/1-370 311 DLSLE --- DLRGQDNYHKNWYIQMTYDSEKHYMYMKIWPCLWLCKKAKQK 21
CG10D4/CG10D4_SHRM/1-353 311 DLSLE --- DLRGQDNYHKNWYIQMTYDSEKHYMYMKIWPCLWLCKKAKQK 21
Q13UJ7/Q13UJ7_P5EA6/1-361 318 DLSLE --- DLRGQDNYHKNWYIQMTYDSEKHYMYMKIWPCLWLCKKAKQK 21
Q13UJ7/Q13UJ7_P5EA6/1-361 320 DLSLE --- DLRGQDNYHKNWYIQMTYDSEKHYMYMKIWPCLWLCKKAKQK 21
AAAR33/AAAR33_STLAC/1-382 340 CKNTE --- EYHKLSTHYNNIDYQSMRFSSQLKVHTAKCEKAK 21
AOM245/AOM245_GRAFX/1-407 365 SEFEAT --- EYHKLSTHYNNIDYQSMRFSSQLKVHTAKCEKAK 21
Q13P89/Q13P89_SPH/1-408 360 NIDFG --- FGLMTEVYDQVYHREFEGRKRTHYARCEKSKSSDE 21
B5CY74/B5CY74_P5EA6/1-402 358 EKEF --- EYHKLSTHYNNIDYQSMRFSSQLKVHTAKCEKAK 21
Q13XP8/Q13XP8_P5EA6/1-436

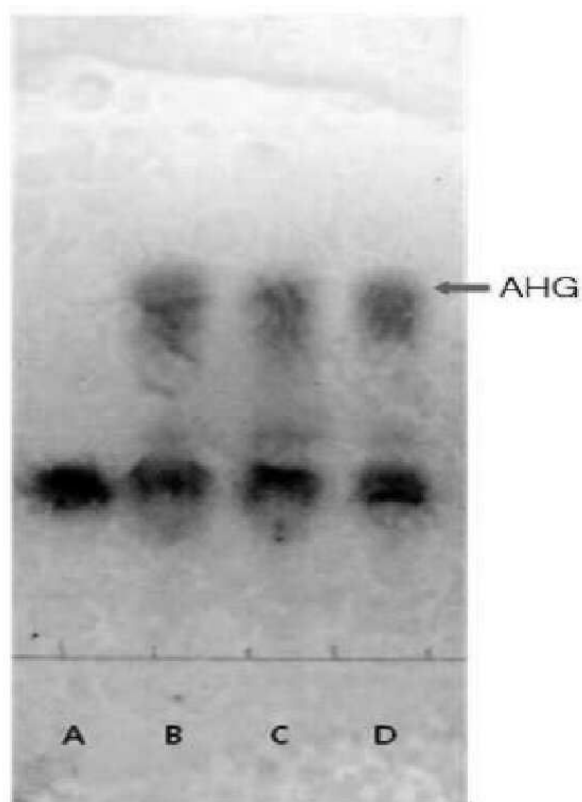
도면12



도면13



도면14



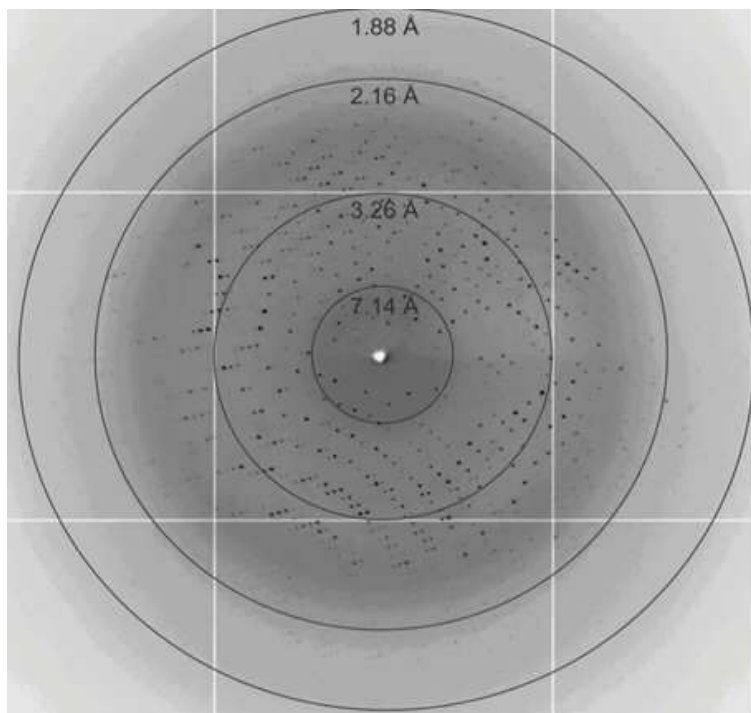
도면15

Space group	<i>C2</i>
Unit-cell parameters (Å)	$a = 129.83$, $b = 76.81$, $c = 90.1$, $\beta = 101.86$
Resolution (Å)	50 – 1.98 (2.05–1.98)
Total No. of observations	226401 (22764)
No. of unique observations	58856 (5837)
Multiplicity	3.8 (3.9)
Data completeness (%)	98.4 (98.1)
$\langle I/\sigma(I) \rangle$	25.3 (4.1)
R_{merge}^* (%)	0.086 (0.46)

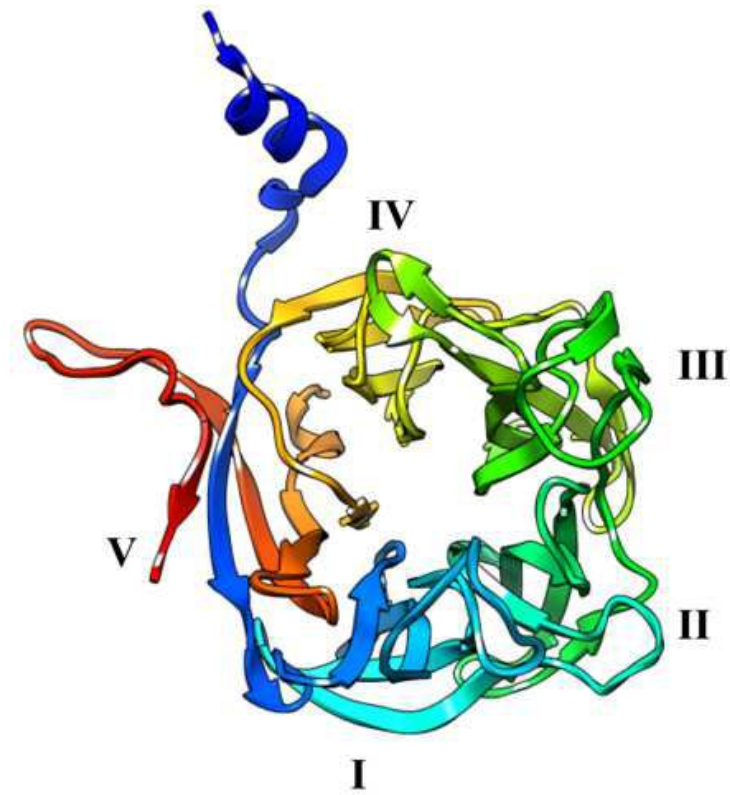
도면16



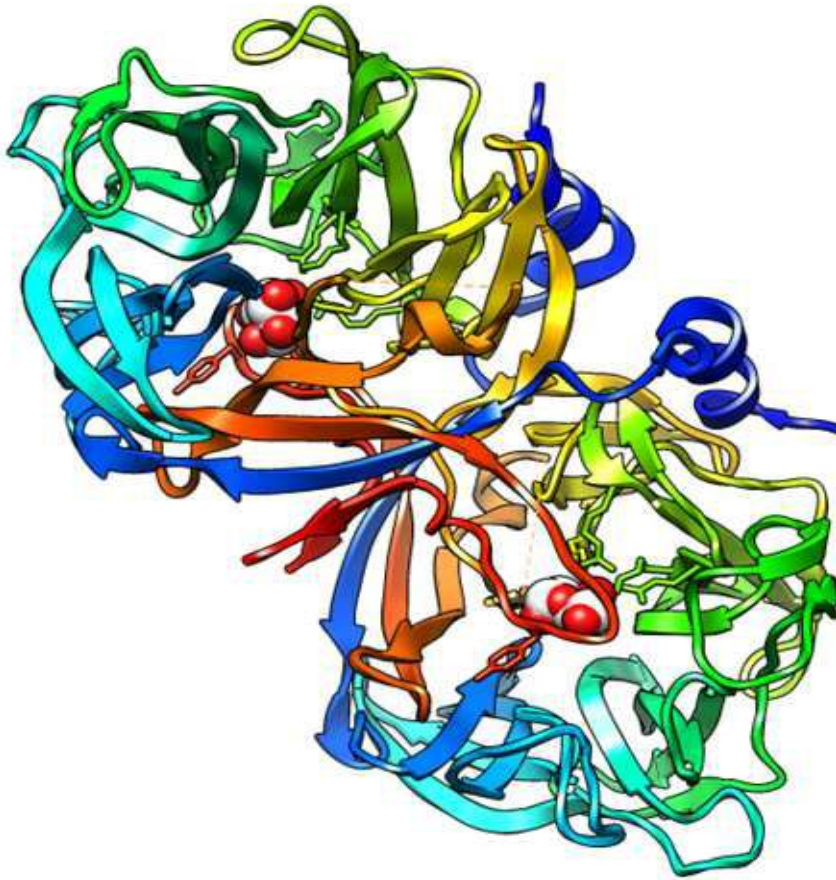
도면17



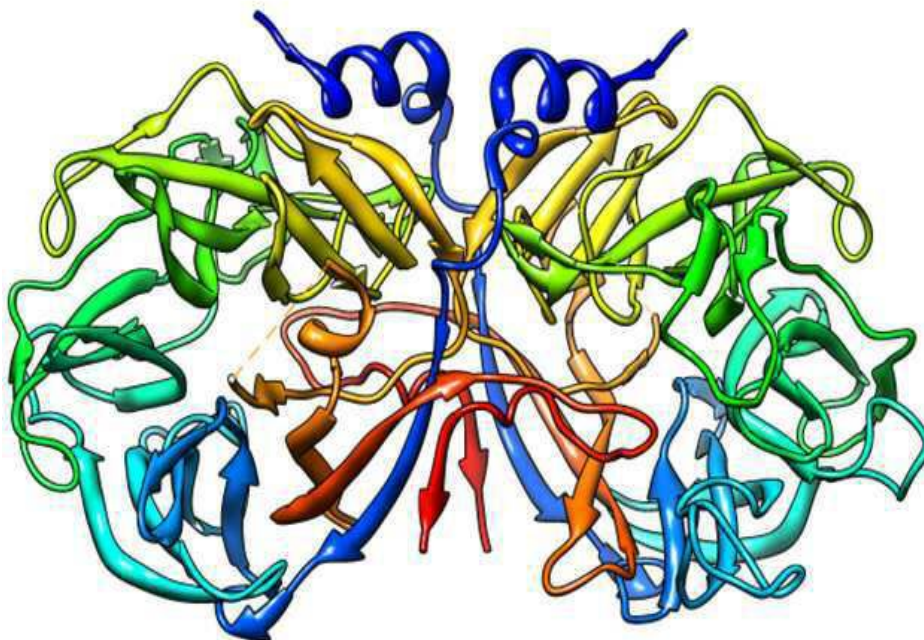
도면18



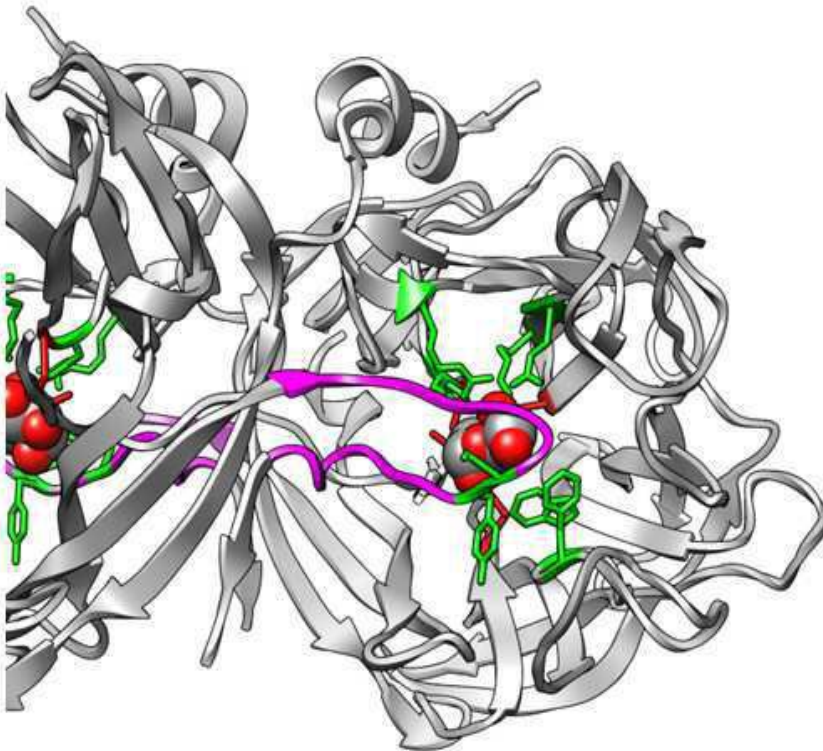
도면19



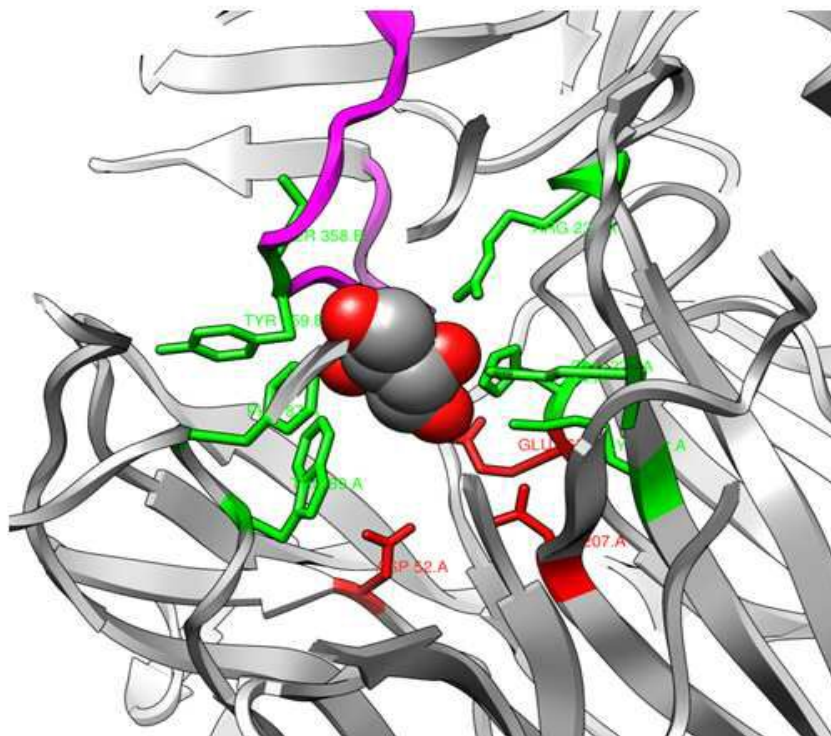
도면20



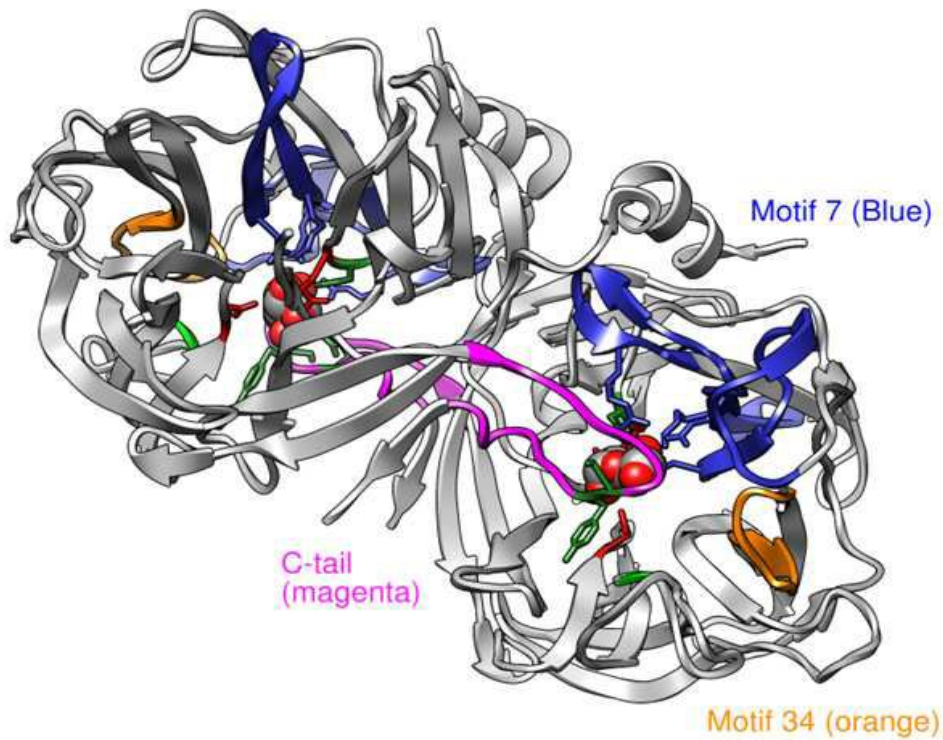
도면21



도면22



도면23



서열 목록

<110> KOREA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION

<120> A novel alpha-neoagarobiose hydrolase and and three dimensional structure thereof

<160> 12

<170> Kopatent In 1.71

<210> 1

<211> 361

<212> PRT

<213> Pseudoalteromonas atlantica T6c

<400> 1

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

1 5 10 15

Glu Arg Gly Tyr Asp Ser Lys Asn Ala Asp Trp Met Ile Glu Phe Glu

20 25 30

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

35 40 45

Arg Asp Pro Thr Ser Val Ile Ser Val Asp Asp Val Tyr His Cys Trp

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

 100 105 110
 Gly Thr Pro Gly Ala Phe Asp Asp Arg Ala Val Phe Thr Pro Glu Val
 115 120 125
 Leu Ala His Asn Gly Lys Tyr Tyr Leu Val Tyr Gln Thr Val Gln Tyr
 130 135 140
 Pro Tyr Thr Asn Arg Gln Ile Glu Gln Ile Ala Ile Ala His Ala Asp
 145 150 155 160
 Ser Pro Phe Gly Pro Trp Ile Lys Ser Thr Ala Pro Ile Leu Ser Pro
 165 170 175

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

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

305						310												320
Ser	Pro	Pro	Gly	Leu	His	Trp	Gly	Leu	Cys	His	Lys	Tyr	Asp	Ser	Ser			
				325					330						335			
Trp	Asn	Trp	Asn	Tyr	Ile	Cys	Arg	Tyr	Arg	Leu	Lys	Lys	Gln	Val	Leu			
			340						345					350				
Asp	Ala	Gly	Thr	Phe	Gln	Asn	Ser	Asn										
		355						360										
<210>	2																	
<211>	370																	
<212>	PRT																	
<213>	Streptomyces coelicolor A3																	
<400>	2																	
Met	Thr	Met	Pro	Ala	Glu	Ser	Ala	Ala	Thr	Val	Arg	Gly	Arg	Arg	Tyr			
1				5					10						15			
Pro	Ala	Asp	Pro	Glu	Trp	Phe	Cys	Asp	Phe	Arg	Thr	Phe	Pro	Val	Gln			
			20						25					30				
Gly	Leu	Gly	Pro	Glu	Val	Gly	Val	His	Arg	Arg	Asp	Pro	Ser	Ser	Val			
		35						40					45					
Leu	Thr	Val	Asp	Gly	Arg	Tyr	His	Val	Trp	Tyr	Thr	Arg	Ser	Thr	Gly			
	50						55					60						
Glu	Thr	Asp	Gly	Phe	Gly	Thr	Gly	Asp	Pro	Met	Ala	Lys	Val	Phe	Pro			
65					70					75					80			
Trp	Asp	Trp	Ser	Glu	Ile	Trp	His	Ala	Thr	Ser	Asp	Asp	Gly	Ala	Thr			
					85					90					95			
Trp	Val	Glu	Gln	Gly	Arg	Ala	Leu	Gly	Arg	Gly	Glu	Pro	Gly	Ser	Tyr			
			100						105					110				
Asp	Asp	Arg	Ser	Val	Phe	Thr	Pro	Glu	Val	Leu	Glu	His	Gln	Gly	Trp			
		115							120				125					
Phe	Tyr	Leu	Val	Tyr	Gln	Val	Ile	Arg	Ser	Pro	Tyr	Ala	Leu	Arg	Ser			
	130						135					140						
Phe	Glu	Ser	Ile	Ala	Met	Ala	Lys	Ala	Pro	Ser	Pro	Asp	Gly	Pro	Trp			

145 150 155 160
 Val Arg Ser Pro Gln Pro Ile Leu Arg Pro Gln Ala Asp Gly Glu Trp
 165 170 175
 Ala Gly Glu Glu Asp Asn Arg Leu Ser Val Val Ser Gln Gly Ser Phe
 180 185 190
 Asp Ser His Lys Val His Asp Pro Ile Leu Val Pro Phe Gln Gly Arg
 195 200 205
 Phe Phe Leu Tyr Tyr Lys Gly Glu Gln Met Gly Glu Gly Phe Ser Ala
 210 215 220
 Gly Gly Arg Thr Thr Arg Trp Gly Leu Ala Ile Ala Asp Asp Ile Glu

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

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

 370

<210> 3

<211> 407

<212> PRT

<213> Gramella forsetii (strain KT0803)

<400> 3

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

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Pro Asp Val Gly Asn Glu Trp Phe Ile Glu Phe Ser Glu Leu Gln Pro

65 70 75 80

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

85 90 95

Ser Ala Leu Leu Lys His Asp Gly Lys Tyr Tyr Val Trp Tyr Thr Lys

100 105 110

Ser Val Gly Pro Thr Gln Gly Phe Gly Gly Asp Ile Glu Asn Asp Lys

115 120 125

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

130 135 140

Gly Trp Thr Trp Glu Glu Gln Gly Ile Ala Val Ala Arg Gly Glu Lys

145 150 155 160

Gly Glu Tyr Asp Asp Arg Ser Val Phe Thr Val Glu Ile Met Glu Asp

165 170 175

Lys Gly Lys Tyr Tyr Leu Ser Tyr Gln Thr Ile Gln Ser Pro Tyr Asn

180 185 190

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

195 200 205

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

210 215 220

Gly Ile Trp Lys Gly Glu Glu Gln Asn Arg Phe Ala Val Glu Lys Lys

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

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

 340 345 350
 Ala Ile Gly Leu Asn Arg Glu Leu Asn Thr Asp Glu Ser Glu Glu Pro
 355 360 365
 Gly Ala Ile Leu Gly Trp Gly Leu Ser His Ile Tyr Asn Asn Ser Asp
 370 375 380
 Tyr Gln Ser Ile Met Arg Phe Ser Ser Lys Arg Arg Thr Ser His Val
 385 390 395 400
 Ala Lys Gly Glu Lys Ala Glu
 405

 <210> 4
 <211> 368

 <212> PRT
 <213> Saccharophagus degradans
 <400> 4
 Met Ser Asp Ser Lys Val Asn Lys Lys Leu Ser Lys Ala Ser Leu Arg
 1 5 10 15
 Ala Ile Glu Arg Gly Tyr Asp Glu Lys Gly Pro Glu Trp Leu Phe Glu
 20 25 30

Phe Asp Ile Thr Pro Leu Lys Gly Asp Leu Ala Tyr Glu Glu Gly Val
 35 40 45
 Ile Arg Arg Asp Pro Ser Ala Val Leu Lys Val Asp Asp Glu Tyr His
 50 55 60

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

 130 135 140
 Lys Ala Pro Tyr Leu Asn Arg Ser Leu Glu His Ile Ala Ile Ala Tyr
 145 150 155 160
 Ser Asp Ser Pro Phe Gly Pro Trp Thr Lys Ser Asp Ala Pro Ile Leu
 165 170 175
 Ser Pro Glu Asn Asp Gly Val Trp Asp Thr Asp Glu Asp Asn Arg Phe
 180 185 190
 Leu Val Lys Glu Lys Gly Ser Phe Asp Ser His Lys Val His Asp Pro
 195 200 205

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

275 280 285
 Cys Gln Trp Ala Glu Asp Gly Ile Asn Phe Asp Ile Met Ser His Ile
 290 295 300
 Lys Gly Ala Pro Glu Ala Val Gly Phe Phe Arg Pro Glu Ser Asp Ser
 305 310 315 320
 Asp Asp Pro Ile Ser Gly Ile Glu Trp Gly Leu Ser His Lys Tyr Asp
 325 330 335
 Ala Ser Trp Asn Trp Asn Tyr Leu Cys Phe Phe Lys Thr Arg Arg Gln
 340 345 350

 Val Leu Asp Ala Gly Ser Tyr Gln Gln Thr Gly Asp Ser Gly Ala Val
 355 360 365
 <210> 5
 <211> 430
 <212> PRT
 <213> Pseudoalteromonas atlantica
 <400> 5
 Met Lys Tyr Trp Leu Ile Gly Ala Leu Ser Leu Ser Leu Thr Ala Cys
 1 5 10 15
 Asn Asn Thr Thr Met Glu Ser Glu Thr Arg Ala Ala Ser Thr Ser Val
 20 25 30
 Ser Ala Pro Asn Lys Ala Tyr Ser Gln Ala Asp Tyr Ala Tyr Leu Ala

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

 Arg Thr Thr Gly Glu Ala Val Gly Phe Gly Thr Gly Asp Pro Asn Ala
 115 120 125

Lys Val Phe Pro Trp Asp Gln Ala Glu Ile Trp Tyr Ala Thr Ser Lys
 130 135 140
 Asp Gly Tyr His Trp Gln Glu Lys Gly Arg Ala Leu Gly Thr Gly Asp
 145 150 155 160
 Val Gly Gln Phe Asp Asp Arg Ser Val Phe Thr Pro Glu Val Leu Ala
 165 170 175
 His Gln Gly Met Phe Tyr Leu Val Tyr Gln Ala Ile Lys Ala Pro Tyr
 180 185 190
 Leu Asn Arg Thr Lys Asn Thr Val Ala Met Ala Tyr Ser Ser Ser Pro
 195 200 205
 Asp Gly Pro Trp Arg Lys Leu Asp Ala Pro Ile Leu Thr Ala Ser Asp
 210 215 220
 Thr Gly Glu Trp Ala Gly Thr Gln Asp Asn Arg Phe Leu Val Lys Ser
 225 230 235 240
 Gln Gly Asp Phe Asp Ser His Lys Val His Asp Pro Thr Leu Leu Phe
 245 250 255
 Tyr Arg Asp Lys Phe Tyr Leu Tyr Tyr Lys Gly Glu Arg Met Gly Glu
 260 265 270
 Arg Lys Thr Ala Gly Gly Arg Glu Ile Arg Trp Gly Val Ala Ile Ala
 275 280 285
 Asp Asn Pro Gln Gly Pro Tyr Ile Lys Ser Ala Phe Asn Pro Ile Thr
 290 295 300
 His Ser Gly His Glu Leu Cys Val Trp Gln Tyr Gln Asp Gly Ile Ala
 305 310 315 320
 Ile Val Ser Ser His Asp Gly Pro Glu Lys Gln Thr Ile Gln Tyr Ala
 325 330 335
 Pro Asp Gly Ile Asn Phe Glu Val Met Ser Tyr Leu Pro Ser Val Pro
 340 345 350
 Ser Ala Ile Gly Leu Val Glu Ser Leu Asp Lys Asn Ala Tyr Pro Thr
 355 360 365
 Ala Gly Leu Glu Trp Gly Leu Tyr His Glu Tyr Val Ile Pro Pro Gly
 370 375 380

Lys Thr Trp Met Ser Gly Ile Asn His Ile Lys Arg Phe Ser Phe Asn
385 390 395 400

Ala Pro Asn Arg Val Ser Met Gln Gly Asp Ser Ile Gly Asn Thr Pro
405 410 415

Gln Asn Asn Ser Gly Glu Phe Ile Asn Lys Gln Ser Asp His
420 425 430

<210> 6

<211> 367

<212> PRT

<213> Paenibacillus sp.

<400> 6

Met Gly Met Lys Lys Glu Ser Ala Ala Thr Ile Arg Ala Lys Gln Tyr
1 5 10 15

Thr Lys Asn Ala Asp Trp Phe Cys Glu Phe Lys Tyr Glu Pro Val Glu

20 25 30

Gly Leu Gly Tyr Glu Glu Gly Ile His Arg Arg Asp Pro Ser Ser Val
35 40 45

Ile Lys Val Gly Asp Leu Tyr Tyr Val Trp Tyr Thr Lys Ser Val Gly
50 55 60

Glu Ala Val Gly Phe Asn Thr Gly Asp Pro Asp Ser Lys Val Phe Pro
65 70 75 80

Trp Asp Gln Ser Asp Val Trp Tyr Ala Thr Ser Pro Asp Gly Thr Thr
85 90 95

Trp Thr Glu Arg Gly Met Ala Val Gly Arg Gly Pro Lys Gly Ser Tyr
100 105 110

Asp Asp Arg Ser Val Phe Thr Pro Glu Ile Leu His His Asp Gly Lys
115 120 125

Phe Tyr Leu Val Tyr Gln Val Ile Glu Gly Gln Tyr Leu Leu Arg Lys
130 135 140

Tyr Glu Asn Ile Ala Met Ala Val Ala Glu Ser Pro Glu Gly Pro Trp
145 150 155 160

Arg Lys Leu Asp Ala Pro Ile Leu Arg Pro Ala Met Asn Gly Glu Trp

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

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

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

<210> 7

<211> 353

<212> PRT

<213> Ruminococcus sp.

<400> 7

Met Lys Leu Ser Ser Ala Ser Lys Arg Ala Ile Glu Arg Asn Tyr Tyr
1 5 10 15
Gln Ser Cys Glu Trp Phe Cys Asp Phe Lys Glu Ser Gly Ile Glu Gly
20 25 30
Ile Gly Tyr Glu Lys Gly Ile His Arg Arg Asp Pro Ser Ser Val Ile
35 40 45
Lys Val Gly Asp Asp Tyr Phe Val Trp Tyr Ser Arg Ser Val Gly Pro
50 55 60
His Lys Gly Phe His Thr Gly Asp Glu Glu Ala Lys Val Phe Pro Trp
65 70 75 80
Asp Tyr Cys Asp Ile Trp Tyr Ala Val Ser Lys Asp Gly Tyr Arg Trp
85 90 95
Glu Glu Lys Gly Pro Ala Val Val Arg Gly Glu Arg Gly Ala Tyr Asp
100 105 110
Asp Arg Ser Val Phe Thr Pro Glu Ile Leu Glu Tyr Glu Gly Lys Tyr
115 120 125
Tyr Leu Val Tyr Gln Val Val Gln His Pro Tyr Val Asn Arg Ser Phe
130 135 140
145 150 155 160
Lys Ser Lys Glu Pro Ile Leu Thr Pro Thr Lys Asp Gly Ile Trp Glu
165 170 175
Gly Glu Glu Asp Asn Arg Phe Ala Val Lys Lys Lys Gly Ser Phe Asp
180 185 190
Ser His Lys Val His Asp Pro Ile Leu Phe Ala Phe Arg Gly Lys Phe
195 200 205
Tyr Leu Tyr Tyr Lys Gly Glu Pro Met Gly Glu Glu Leu Tyr Met Gly
210 215 220
Gly Arg Glu Thr Lys Trp Gly Val Ala Ile Ala Asp Asn Ile Leu Gly
225 230 235 240
Pro Tyr His Arg Ser Glu Tyr Asn Pro Val Thr Asn Ser Gly His Glu
245 250 255

Thr Cys Leu Trp Gln Tyr Asn Gly Gly Ile Ala Ala Phe Leu Arg Thr
260 265 270

Asp Gly Val Glu Lys Asn Thr Ile Gln Phe Ala Glu Asp Gly Ile Asn
275 280 285

Phe Glu Ile Lys Ser Val Ile Lys Gln Gly Pro Glu Ala Cys Gly Pro
290 295 300

Tyr Arg His Leu Glu Ser Asp Ser Asn Pro Leu Lys Gly Met Glu Trp
305 310 315 320

Gly Leu Cys His Asp Val Ser Gln Asp Tyr Gly Phe Ile Lys Arg Phe
325 330 335

Asp Ile Asp Glu Trp Gln Lys Lys Val Tyr Thr Asn Arg Glu Met Tyr
340 345 350

Glu

<

210> 8

<211> 364

<212> PRT

<213> Paenibacillus sp.

<400> 8

Met Arg Lys Glu Ser Ala Ala Thr Lys Arg Ala Lys Gln Tyr Thr Lys
1 5 10 15

Asn Ser Asp Trp Phe Cys Glu Phe Arg Ile Ser Pro Val Gln Gly Leu
20 25 30

Gly Tyr Glu Glu Gly Val His Arg Arg Asp Pro Ser Ser Val Ile Lys
35 40 45

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

50 55 60

Ala Gly Phe Gln Thr Gly Asp Pro Asp Ala Lys Val Phe Pro Trp Asp
65 70 75 80

Leu Ser Glu Val Trp Tyr Ala Thr Ser Pro Asp Gly Val Asn Trp Thr
85 90 95

Glu Arg Gly Arg Ala Ile Gly Arg Gly Pro Lys Gly Ser Tyr Asp Asp

100	105	110	
Arg Ser Val Phe Thr Pro Glu Ile Leu His His Glu Asp Lys Phe Tyr			
115	120	125	
Leu Val Tyr Gln Val Val Gln Ser Pro Tyr Thr Leu Arg Thr Leu Glu			
130	135	140	
Asn Ile Ala Met Ala Val Ala Glu Ser Pro Glu Gly Pro Trp Arg Lys			
145	150	155	160
Leu Asp Glu Pro Ile Leu Arg Pro Ser Asn Asn Gly Glu Trp Phe Gly			
165	170	175	
Asp Asp Asp Asn Arg Leu Thr Val Lys His Lys Gly Asp Phe Asp Ser			
180	185	190	
Leu Lys Val His Asp Pro Thr Leu Phe Phe Tyr Asn Asn Gln Phe Trp			
195	200	205	
Leu Tyr Tyr Lys Gly Glu Gln Met Gly Glu Glu Met Asn Phe Gly Gly			
210	215	220	
Arg Thr Thr Lys Trp Gly Val Ala Ile Ala Asp His Pro Glu Gly Pro			
225	230	235	240
Tyr Val Lys Ser Glu Tyr Asn Pro Ile Thr Asn Ser Gly His Glu Thr			
245	250	255	
Cys Leu Trp Gln Tyr Arg Gly Gly Met Ala Ala Leu Leu Thr Thr Asp			
260	265	270	
Gly Pro Glu Lys Asn Thr Ile Gln Tyr Ala Pro Asp Gly Ile Asn Phe			
275	280	285	
Glu Ile Glu Ala Val Ile Lys Asn Pro Pro Glu Ala Pro Gly Pro Tyr			
290	295	300	
Arg Thr Pro Asp Ala Asp Arg Ser Pro Leu Glu Gly Leu Arg Trp Gly			
305	310	315	320
Leu Cys His Asn Val His Ser Lys Trp Asn Tyr Ile Gln Met Tyr Thr			
325	330	335	
Val Asp Glu Ser Phe Lys His Tyr Tyr Met Asn Lys Ile Asn Pro Glu			
340	345	350	

Glu Leu Trp Leu Gly Lys Ser Pro Lys Arg Asn Gln

355

360

<210> 9

<211> 382

<212> PRT

<213> Flavobacteriales bacterium

<400> 9

Met Lys Lys Glu Lys Glu Val Val Glu Lys Lys Ser Phe Glu Ile Ser

1

5

10

15

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

20

25

30

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

35

40

45

Phe Ile Glu Phe Ser Ala Leu Lys Pro Leu Lys Gly Asp Leu Ala Tyr

50

55

60

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

65

70

75

80

Gly Lys Tyr Tyr Val Trp Tyr Ser Lys Ser Leu Gly Glu Ser Gln Gly

85

90

95

Phe Gly Gly Asp Ile Glu Asn Asp Lys Val Phe Pro Trp Asp Arg Cys

100

105

110

Asp Ile Trp Tyr Ala Thr Ser Val Asp Gly Trp Thr Trp Lys Glu Glu

115

120

125

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

130

135

140

Val Phe Thr Val Glu Ile Met Lys Tyr Glu Asn Met Tyr Tyr Leu Cys

145

150

155

160

Tyr Gln Thr Val Gln Ser Pro Tyr Thr Val Arg Val Lys Asn Gln Val

165

170

175

Gly Leu Ala Trp Ser Asp Ser Pro Asp Gly Pro Trp Thr Lys Ser Lys

180

185

190

Gln Pro Ile Leu Ser Pro Ala Asp Asn Gly Ile Trp Lys Gly Leu Asp

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

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

 325 330 335
 Val Asp Asn Gly Lys Asn Pro Thr Glu Val Leu Arg Trp Gly Leu Ser
 340 345 350
 His Val Tyr Asn Asn Ser Asp Tyr Gln Ser Ile Met Arg Phe Ser Ser
 355 360 365
 Gln Leu Lys Thr Val His Thr Ala Lys Gly Glu Lys Ala Lys
 370 375 380
 <210> 10
 <211> 402
 <212> PRT
 <213> Bacteroides plebeius
 <400> 10
 Met Arg Lys Ile Ile Phe Ala Ala Gly Met Met Ser Leu Leu Ala Ala

1 5 10 15
 Cys Gly Asn Thr Gly Asn Thr Gln Thr Ile Ala Val Asp Asp Thr Gln
 20 25 30

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

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

 145 150 155 160
 Arg Ser Val Phe Thr Pro Glu Val Met Glu Trp Lys Gly Lys Tyr Tyr
 165 170 175
 Leu Cys Tyr Gln Ala Val Lys Ser Pro Tyr Thr Val Arg Val Lys Asn
 180 185 190
 Thr Ile Gly Met Ala Cys Ala Asp Ser Pro Glu Gly Leu Trp Thr Lys
 195 200 205
 Thr Asp Lys Pro Val Leu Glu Pro Ser Asp Thr Gly Glu Trp Glu Gly
 210 215 220

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

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

290 295 300

Cys Val Trp Pro Tyr Lys Gly Gly Ile Ala Ser Leu Ile Thr Thr Asp

305 310 315 320

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

325 330 335

Glu Ile Met Ser Val Val Lys Gly Ala Pro His Ala Ile Gly Leu Asn

340 345 350

Arg Ser Ala Asp Ala Glu Lys Glu Pro Thr Glu Ile Leu Arg Trp Gly

355 360 365

Leu Thr His Ile Tyr Asn Ser Ser Asp Tyr Gln Ser Ile Met Arg Phe

370 375 380

Ser Thr Trp Thr Leu Gln Thr His Thr Ala Lys Gly Glu Ser Lys Glu

385 390 395 400

Arg Lys

<210> 11

<211> 408

<212> PRT

<213> Microscilla sp.

<400> 11

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

1 5 10 15

Leu Leu Phe Phe Ala Ser Cys Gln Thr Pro Cys Pro Glu Glu Gln Glu

20 25 30

Glu Asn Phe Glu Ile Thr Gln Asp Lys Ile Asp Tyr Leu Gly Ile Thr

35 40 45

Asn Pro Asp Lys Leu Ser Ala Ala Ser Lys Arg Ala Leu Gln Trp Pro

50 55 60

Asp Ile Gly Asn Glu Trp Tyr Gly Glu Phe Gln Val Tyr Asp Leu Lys

65 70 75 80

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

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

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

Arg Phe Glu Gly Lys Arg Lys Thr Thr His Val Ala Arg Gly Glu Thr
 385 390 395 400
 Gly Ser Lys Lys Ser Ser Asp Glu
 405

<210> 12

<211> 1107

<212> DNA

<213> Saccharophagus degradans

<400> 12

atgagcgatt caaaagtaaa taaaaaattg agtaaagcta gcctgcgagc catagagcgc	60
ggctacgatg aaaagggggc tgaatggctg tttgagtttg atattacccc actaaaaggc	120
gacttagcct acgaagaagg cgtaattcgt cgagacccca gcgcagtatt aaagggtggac	180
gatgaatatc acgtttggta caccaagggc gaaggtgaaa cagtaggctt cggcagcgac	240
aaccccgaaag acaaagtctt cccatgggat aaaacagaag tttggcacgc cacctctaaa	300
gataagatta ctggaaaga aattggccct gccatacaac gcggcgcagc tggggcatat	360
gatgaccgtg cagtgttcac cccgaagtc ctgcgccata acggcaccta ctacctgta	420
tatcaaacgg taaaagcgcc ctacttaaac cgatcgctag agcatatagc catcgcatac	480
agcgattccc ctttggccc atggaccaa tccgatgcgc caattttaag cccagaaaat	540
gacggcgttt gggatacggc cgaagacaat cgatttttag taaaagagaa aggcagtttc	600
gatagccaca aaglacacga cccctgctta atgtttttta acaatcgttt ctacctgtat	660
tacaaaggcg agactatggg cgaaagcatg aacatgggcg gcagagaaat aaaacacggt	720
gtagccattg ccgactcgcc acttgggccc tacacaaaa gcgaatacaa ccctattacc	780
aatagtggcc atgaagttgc cgtatggccc tacaaagtg gaatggccac catgctaacc	840
accgacgggc cagaaaaaaa cacctgccag tgggcagaag acggcattaa ctttgacatt	900
atgtcgcata taaaaggcgc accagaagca gtaggttttt ttagaccaga aagcgatagc	960

gacgacccta taagcggcat tgaatggggg ctaagccaca agtacgacgc cagctggaac 1020

tggaactatc tatgcttttt taaaacgcgt cgacaagttt tagatgcagg tagctatcag 1080

caaacaggcg attccggagc agtataa 1107