

Bacillus stearothermophilus Acetylxylan Esterase 유전자(estI)의 염기 서열 결정

이정숙 · 최용진*
고려대학교 생명공학원

Nucleotide Sequence of the estI Gene Coding for Bacillus stearothermophilus Acetylxylan Esterase. Lee, Jeong-Sook and Yong-Jin Choi*. Graduate School of Biotechnology, Korea University, Seoul, 136-701, Korea - The nucleotide sequence of the estI gene encoding acetylxylan esterase I of *Bacillus stearothermophilus* was determined and analyzed. The estI gene was found to consist of a 810 base pair open reading frame coding for a polypeptide of 270 amino acids with a deduced molecular weight of 30 kDa. This was in well agreement with the molecular weight (29 kDa) estimated by SDS-PAGE of the purified esterase. The coding sequence was preceded by a putative ribosome binding site 10 bp upstream of the ATG codon. Further 53 bp upstream, the transcription initiation signals were identified. The putative -10 sequence (TCCAAT) and -35 sequence (TTGAAT) corresponded closely to the respective consensus sequences for the *Bacillus subtilis* major RNA polymerase. The G+C content of the coding region of the estI was 51% whereas that of the third position of codons was 60.2%. The N-terminal amino acid sequence of the EstI deduced from the nucleotide sequence perfectly matched the corresponding region of the purified esterase described previously. Comparison with the amino acid sequence of other esterases and lipases reported so far allowed us to identify a sequence, GLSMG at positions 123 to 127 of the EstI which was reported to be the highly conserved active site sequence for those enzymes. The nucleotide sequence of the estI revealed 55.7% homology to that of the xylC coding for the acetylxylan esterase of *Caldocellum saccharolyticum*.

Xylan은 자연계에서의 전체 분포량이 cellulose 다음을 차지하고 있는 중요 biomass 자원인 hemicellulose의 주성분으로 β -1,4-polyxylose 주 탄소쇄에 acetic acid, arabinose 및 glucuronic acid 등의 측쇄를 가지고 있는 복잡한 구조의 복합 다당류이다(1, 2).

그러므로 xylan의 효율적 완전 분해 이용을 위해서는 xylanase와 β -xylosidase의 주 xylan 분해계 효소와 acetylxylan esterase, α -arabinofuranosidase 및 α -glucuronidase 등의 탈측쇄 효소와의 협동 작용이 절대적인 것으로 보고되고 있다(1).

실제로 softwood xylan을 제외한 대다수 xylan 분자는 그 출처에 따라 다소의 차이는 있으나 구성 xylose 잔기가 아세틸화되어 있는 것이 일반적이며 특히 hardwood xylan과 forage xylan은 구성 xylose 잔기의 50~70%가 아세틸화되어 있으면서 주 xylan 분해계 효소의 공격으로부터 xylan 분자를 보호하고 있는 것으로 보고되고 있다(3, 4).

본 연구자 등도 강력한 xylan 분해 분리 균주인 *Bacillus stearothermophilus* 생산 xylanase, β -xylosidase 및

acetylxylan esterase를 활용하여 acetylxylan 기질에 대한 이들 효소의 협동 작용에 의한 가수분해 상승효과를 직접 관찰함으로써 xylan 분해에 있어서의 acetylxylan esterase의 역할을 확인한 바 있다(5).

본 연구에서는 상기 *B. stearothermophilus* genomic DNA로부터 분리하여 클로닝한 두 acetylxylan esterase 생산 유전자(estI과 estII)(5, 6) 중의 하나인 estI 유전자의 염기 서열을 결정, estI 유전자의 분자 수준의 특성을 조사하여 그 결과를 보고한다.

재료 및 방법

사용 균주 및 플라스미드

형질 전환의 숙주 세포로 사용한 균주는 *E. coli* JM 109(recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, (lac-proAB), F, traD36, proAB, lacIqZM15)(7)와 DH 5 α (F- Φ , 80dlacZ, (lac-ZYA-argF)U169, deoR, recA1, endA, hsd17(r_K⁻, m_K⁺), phoA, supE44, thi-1, gyrA96, relA1)(7)이었다. Subcloning 실험에서 사용한 vector는 pUC119(7)와 pT7T3 19u이었고 single-stranded DNA 분리를 위한 helper phage로는 M13K07(7)을 사용하였다.

또한, 염기 서열 결정을 위한 *B. stearothermophilus* estI 유전자 DNA 시료는 pKMG6 재조합 플라스미드(6)

*Corresponding author

Tel. 82-2-920-1356, Fax. 82-2-923-9923

Key words: *Bacillus stearothermophilus*, Acetylxylan esterase, Nucleotide sequence

로부터 얻었다.

사용 배지 및 배양 조건

플라스미드 분리를 위한 *E. coli* 세포 배양은 분리 방법에 따라 50 µg/ml의 ampicillin을 첨가한 LB 배지(0.5% yeast extract, 1% NaCl, 1% tryptone), 또는 TB 배지(1.2% tryptone, 2.4% yeast extract, glycerol 4 ml/l, 0.17 M KH₂PO₄, 0.72 M K₂HPO₄)를 사용하여 37°C에서 진탕 배양하였다. 그리고, single-stranded DNA 분리를 위한 *E. coli* 숙주 세포의 배양은 ampicillin이 첨가된 2×YT 배지(1.6% tryptone, 1% yeast extract, 0.5% NaCl)를 사용 37°C에서 진탕 배양하였다.

사용 시약 및 효소

Agarose, lysozyme, TEMED, acrylamide, N,N'-methylenebisacrylamide 등은 Sigma Chemical Co., 기타 일반 시약류는 시판 일급 이상의 분석용을 사용하였다.

각 종 제한 효소, CIP 및 T4 DNA ligase는 Promega Corp. 또는 New England Biolabs(NEB), 그리고 [³⁵S]-α-dATP는 Amersham(U.K.)에서 구입하였다.

DNA sequencing kit는 United States Biochemical (USB)회사 제품인 sequenase version 2.0을, Silver Sequence DNA Sequencing System은 Promega Corp. 제품을 구입 사용하였다.

플라스미드 DNA 분리 및 형질 전환

플라스미드 DNA는 alkaline lysis 법에 따라 플라스미드 DNA를 분리한 다음 PEG를 사용하여 DNA를 침전 분리, 정제하였다(7). 형질 전환은 Hanahan 방법(8)에 따라 준비한 competent cell을 이용, Simanis 법에 준하여 실시하였다.

estI 유전자 클론의 선별

형질 전환체를 Ampicillin(50 µg/ml)과 tributyrin(1.0%)을 첨가한 LB 한천 배지에서 37°C, 24시간 배양했을 때 estI 유전자 발현 산물인 acetylxylan esterase에 의한 tributyrin의 분해로 형성되는 colony 주위의 투명환을 확인함으로써 estI 유전자 클론을 선별하였다.

Deletion subclone의 분리

pKMG6 DNA 상의 acetylxylan esterase 활성과 관련된 3.5 kb DNA 단편을 pUC119에 클로닝하여 pKMG 116을 분리한 다음 Maniatis 등의 방법에 따라 pKMG116으로부터 single-stranded phagemid DNA를 얻고 RD 29 primer와 T4 DNA polymerase를 활용하여 각 종 크기의 deleted mutant plasmid를 얻었다.

상기 mutant plasmid 중에서 esterase 활성을 유지하

고 있으면서 가장 작은 삽입 DNA를 가지고 있는 플라스미드를 분리하고 이 플라스미드의 삽입 DNA(*EcoRI-HindIII* 단편)를 다시 pT7T3 19u vector에 subcloning 함으로써 염기 서열 결정을 위한 estI 유전자 DNA를 최종 분리하였다.

염기 서열 결정

Sanger법에 근거한 USB manual에 따라 Sequenase version 2.0 kit를 이용하여 염기 서열을 결정하였다. 이때 사용한 DNA template는 double-stranded DNA를 분리하여 사용하였다. 또한, Silver Sequence DNA sequencing system을 이용한 경우는 공급회사의 manual에 따라 d/ddNTP 혼합물에 double-stranded template DNA와 primer 등을 첨가한 뒤 PCR로 DNA 증폭 반응을 시켰다. 반응이 끝난 반응 혼합물은 8.3 M urea가 포함된 6% polyacrylamide gel을 이용, 전기영동을 실시한 다음 silver staining을 하는 과정을 거쳐 염기 서열을 결정하였다.

염기 서열 및 아미노산 서열의 computer analysis

정한 estI 유전자 염기 서열로부터 G+C 함량, codon usage 및 EstI 효소의 아미노산 서열 분석 등은 HIBIO DNASIS™ 07-00(Eighth version), EstI의 아미노산 서열로부터 단백질 2차 구조 예측 내지는 hydrophathy plot의 작성은 PROSIS, 다른 관련 유전자들과의 염기 서열의 비교 분석은 Gene Bank R85.0 Sequence Database, 그리고 EstI의 아미노산 서열 비교 분석은 CD-DATA™ Gene/Protein Sequence Database를 활용하였다.

결과 및 고찰

Acetylxylan esterase 유전자(estI)의 염기 서열 결정

Subcloning 과정에서 얻은 *B. stearothermophilus* estI 유전자를 포함하고 있는 약 2.2 kb의 *SmaI-HindIII* DNA 단편을 *EcoRI* 처리하여 1.6 kb DNA 단편을 일차 분리한 다음 *BamHI*을 이용, 300 bp의 *EcoRI-BamHI*과 1300 bp 정도의 *BamHI-HindIII* 두 단편을 얻었다.

pT7T3 19u vector에 상기 두 DNA 단편을 각각 별도로 클로닝하여 크기가 작은 *EcoRI-BamHI* 단편의 경우는 직접 염기 서열을 결정하였으며 *BamHI-HindIII* 단편은 Fig. 1에 표시되어 있는 바와 같이 T3 promoter 및 T7 promoter primers 그리고 각각의 promoter 방향으로 작용하는 별도의 primer A와 B(A: 5'-TCCG-TCCATTCATCAGG-3', B: 5'-GTCAACTGAATCG-TCCTCT-3') 등을 활용하여 염기 서열을 결정하였다.

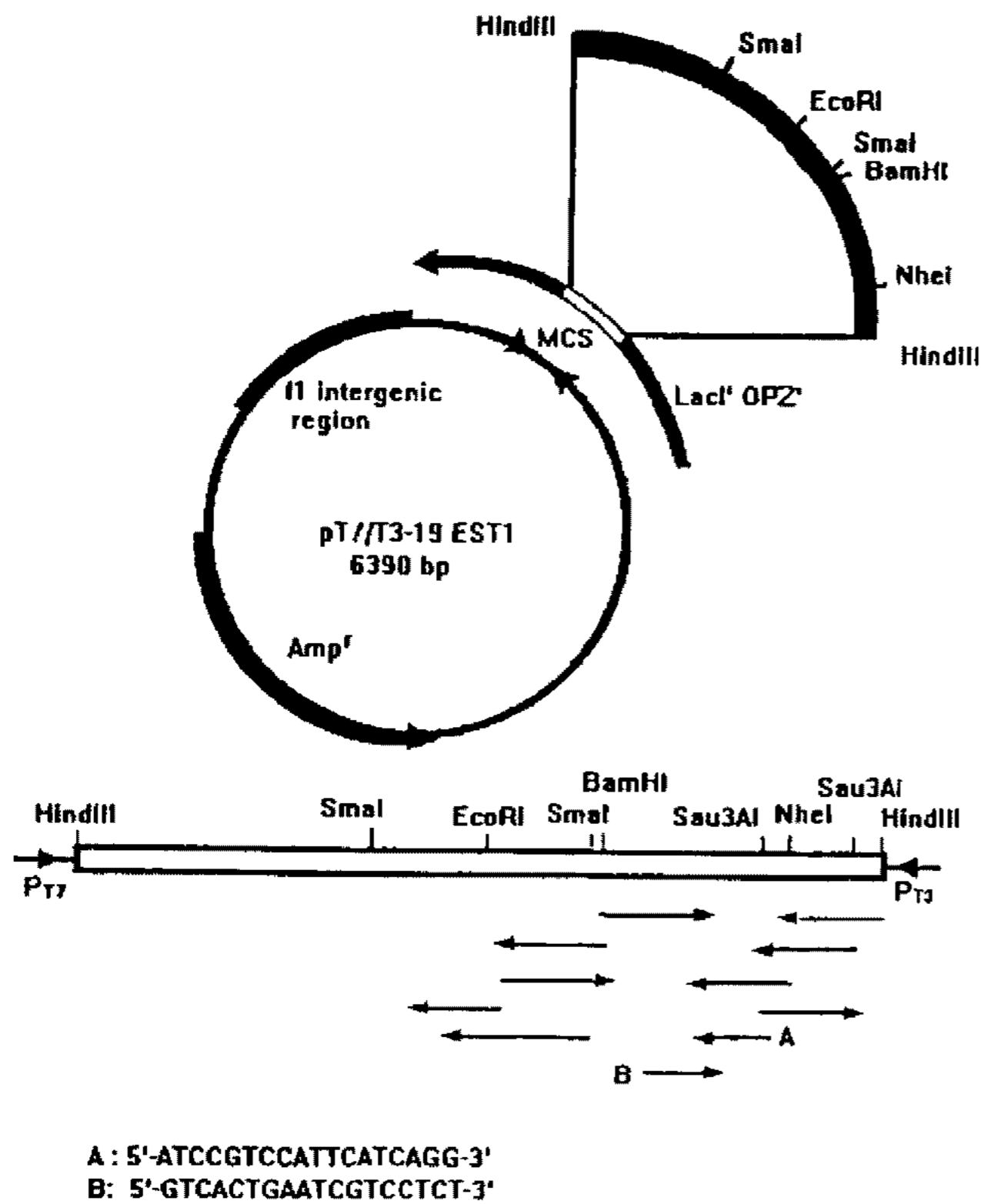


Fig. 1. Fine restriction map of the 2.2 kb insert DNA and pT7T3-19EST1 vector used in this sequencing experiment.
 A: a synthetic oligonucleotide primer used in T3 promoter direction, B: a synthetic oligonucleotide primer used in T7 promoter direction, Arrows indicate the length and direction of sequencing for each fragment used in this experiment.

이와 별도로 상기 1.6 kb *EcoRI-HindIII* 단편을 T3 promoter primer를 이용, *EcoRI* site로부터 약 350 base를 읽음으로써 *EcoRI-BamHI*과 *BamHI-HindIII* 두 단편의 염기 서열을 중복 결정함으로써 상기 두 단편을 연결하였다.

또한, 전체 삽입 DNA를 *SmaI*으로 처리하여 분리한 약 900 bp의 단편과 이 단편을 다시 *EcoRI*으로 처리하여 얻은 600 bp의 *EcoRI-SmaI* 단편을 각각 pT7T3 19u vector에 클로닝하여 마찬가지로 T3, T7 primer를 이용, *EcoRI* site로부터 추가로 340 base를 더 읽음으로써 *estI* 유전자를 포함하고 있는 전체 1965 bp의 DNA 단편의 염기 서열을 일차 결정하였다.

다음은 *BamHI-HindIII* DNA 단편 상에 존재하는 *Sau3AI*과 *NheI* site를 이용하여 일차 결정한 염기 서열을 확인함으로써 Fig. 2에 표시되어 있는 최종 염기 서열을 결정하였다.

염기 서열 분석

앞 항에서 결정한 염기 서열을 토대로하여 *estI* 유전자를 포함하고 있는 DNA 단편상의 유전정보를 분석해 본 결과 상기 DNA상에는 ATG initiation codon으로 시작되는 810 bp의 ORF 한 개가 있음을 알 수 있었다. 본 ORF는 270개의 아미노산으로 구성되어 있으면서 N-말단 아미노산 서열이 Met-Ala-Met-Leu-Gln-Ile-Gln-

1	AA ATT CGA CTG CGC TGG AGC AAG GAG GAC AAA AGC TGT TTG CTC ACC	47	864	AAG GAA AGC GCA GGG ACC CAT GAC TGG AAT TTC TGG AAT GAG CAT ATC	911
48	GTC GAT TTG AAT AGC AAC CAT TCG GTG ATC CAA TAC ATC GAC GAG CAC	95	239	K E S A G T H D W N F W N E H I	254
96	GGG CTG TTG GCT GAT TAC CGG ATT TGA TTT CAG AAA AAA AGG AGG AAC	143	912	GGC CCG GCC ATA GCA TGG GCT GTC ATG GAT CTC GAT TCG GTC AAC TGA	959
144	CGC AGC ATG GCT ATG TTA CAG ATC CAA TTT TTC TCG CAA TCG ATG AGA	191	255	G P A I A W A V M D L D S V N *	270
1	M A M L Q I Q F F S Q S M R	14	960	ATC GTC CTC TGC CCT CAT CTG CCC CTG AAG AGC AGA GAG CAG CGG CCC	1007
192	AGG GAA GTT GCG CTC AGC GCG CTG CTC CCC TTA GAC GCC CCG CCG ATT	239	1008	ATT CGC CGG TTG GCT CCC GCT TCG GGG GCC TCC TCA TTC GAA CCA TCG	1055
15	R E V A L S A L L P L D A P P I	30	1056	TCC AGA TAA AGG AAG AGG ATA CTT GAT GAA AGT CGG AGG ATT ATT TAC	1103
240	CCG GAT CAA CCT GAG AAG GAG GTC AAG CCT CTC AAG ACG TTA TAT CTG	287	1104	TAG AAA AAG CAT CGT TTT CAG GTT TTT CGT ACC CTT CAC AAT CAT CTT	1151
31	P D Q P E K E V K P L K T L Y L	46	1152	GGT CAT CTT GTT TAC GGT ATT ATT GCT CAG CAA CAT TTA TTC CTT GGA	1199
288	TTG CAC GGG TAT TCC GGC AAT CAT ACG GAT TGG CTT CAC TTC TCC CGA	335	1200	TGT GGT TCG GAC CCA ATC GCT CAG CAA TTC GCG TAA TAC GCT CAG AAT	1247
47	L H G Y S G N H T D W L H F S R	62	1248	CTA CGT GGA GAA CAT ACA CAG GCA AGC TGA ATC TAT ATG CAA AGG ATC	1295
336	ATT CGG GAA TTG TCC GAT AAA TAC AAT ATC GCT GTA TTT ATG CCT TCG	383	1296	TTA TCG AGG TGT TCG AGA ACC AGG TGG ATG CGG CGA CAG AAT ATG AGC	1343
63	I R E L S D K Y N I A V F M P S	78	1344	GAC TAA ACG GGA GCG ATC GCT ATC TCA AAG AAA TTG GCT TTG CTG GAT	1391
384	GGA GAA AAT CAT TTC TAC CTT GAC GAT GAG GAT AAG GGG CAG CTA TAC	431	1392	GCC CTG AAG GCC AAA GTG TCG AAT AAT GAC TCC AGC GAC GGG GTC TTT	1439
79	G E N H F Y L D D E D K G Q L Y	94	1440	ATC AAA TTG CCC TGG AGA AGG AGC GAC GCT GGT TCA GTT CGG AAA TCG	1487
432	GGG GAG TTC CTG GGA AGG GAA TTG ATT CAT TTC ACA CGC TCG ATG TTC	479	1488	AAC CCG ATC CGA GGA TAA GCT GGG GCT GAT AGA CTT TCT GAA GAA GCA	1535
95	G E F L G R E L I H F T R S M F	110	1536	CGA TTT CCC ACC CTA TTC TAA CCG ACC CCC TGA TGA ATG GAC GGA TTA	1583
480	CCG CTA TCC GCA AAT CGC AAG GAT ACA TTC ATC GGT GCA CTT TCG ATG	527	1584	TGA AGT TCA GGG TGC GCA TTA TTT GTT TCG ATA CAT CAC CTA TGA GAA	1631
111	P L S A N R K D T F I G G L S M	126	1632	GGC TAG CTT TGG CAC CTT CGT CAA AGC CGA TTC CTT GCT TGC CGT TGC	1679
528	GGC GGG TAT GGA GCG ATT CGG AAC GGT CTG AAG TAT GCC GAC CAG TTT	575	1680	CAG CGG GAA AGG ACA AGA GCT CGG ACC AAC CCT CCT AAG CTC TCG GGA	1727
127	G G Y G A I R N G L K Y A D Q F	142	1728	AGG CGG TAT TCT TGC TTC GAC CGA TGG GAG CCT GTT AGC AGC CGC ATC	1775
576	GGC CGA ATC ATC GCT CTC TCT TCA GCG CTC ATT CCT TAC ACC ATT GCC	623	1776	GTC TCT GTC GGA TTT GAA AGA TCG GTC CAA CAA GAG CTA TCT GTT CAT	1823
143	G R I I A L S S A L I P Y T I A	158	1824	TTC TGA ACC GAT CGC TGA ATT TGG CGA AAT CAC CCA CAT GAT GGC CAA	1871
624	GGG TTG CCC CCG GGC TAT CAG GAC GCG ATT GCC GAC TAC AAT TAC TAC	671	1872	GCA TAG TAT TTT CTC AGG ACT CAA GCT AAT TCA ATG GCT TAT TAT CTC	1919
159	G L P P G Y Q D A I A D Y N Y Y	174	1920	CTT GGC TGT GAT TTC GAT GGT GGC GAT TCC CCT GAT GCT GAA GCT T	1965
672	AAG CGT GTT TTC GGG GAC TTG AAT CAG CTT CTT GGC AGC GAC AAG GAT	719			
175	K R V F G D L N Q L L G S D K D	190			
720	CCG GAA GCA CTG GTC CGA CAG TTG AAG GAG CAG GCG GCG CCG ATT CCA	767			
191	P E A L V R Q L K E G G G A P I P	206			
768	GGG ATT TAC ATG GCA TGC GGT GAC GAG GAT TTT TTG CTT GAT GTT AAC	815			
207	G I Y N A C G D E D F L L D V N	222			
816	CGG CGG TTC CAT GAT TAT TTA ACC CAA GAG GAT ATC CCT CAT CAG TAT	863			
223	R R F H D Y L T Q E D I P H Q Y	238			

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the 1965 bp insert in pT7T3 19u-EST1. Putative promoter sequence (-35 region, -10 region), Shine-Dalgarno (SD) sequence, N-terminal sequence (1), and catalytic residues (2) are underlined. The translational termination codon is indicated by asterisk (*).

Phe-Phe-Ser-Gln이고 추산 분자량이 30 kDa인 한 polypeptide 분자의 유전 정보에 해당되었다.

이와 같은 염기 서열 분석 결과는 정제 EstI 효소를 이용 분석한 N-말단 아미노산 서열과 아마도 post-translational processing에 의한 첫번째 Met 잔기가 제거된 것 외에 완전 일치하였으며 29 kDa으로 측정된 분자량도 매우 근사한 값을 보였다(5).

또한 Table 1에 표시하고 있는 바와 같이 ATG initiation codon의 9 base 윗쪽에는 *Bacillus subtilis* 16S rRNA 3'-region의 consensus sequence(9)와 높은 상보성을 나타내는 Shine-Dalgarno sequence로 추정되는 sequence가 있었다. 따라서, translation initiation 효율과 직접적인 관계를 가지고 있는 것으로 알려진 SD sequence의 염기 서열, initiation codon의 종류 내지는 SD sequence와 initiation codon 사이의 거리 등의 관점에서 볼 때 본 mRNA는 매우 효율적인 translation template인 것으로 판단된다(10).

한편, *estI* 유전자 관련 promoter region을 조사해 본 결과 translation initiation codon으로부터 각각 68 및 90 base 윗쪽에 -10 element와 -35 element로 추정되는 sequence의 존재를 확인할 수 있었다.

TCCAAT의 염기 서열을 가진 Pribnow box의 경우는 *B. subtilis*의 consensus sequence(9)와 4개의 염기가, 또 TTGAAT의 recognition sequence 역시 consensus sequence와 6개 중 4개의 염기가 일치되고 있었다 (Table 2 참조).

또한 상기 두 element 사이의 거리도 17 bp로 매우 효율적인 promoter 구조임을 보여주고 있으며 이와 같은 regulatory element는 *E. coli* host에서도 효과적으로 작용할 수 있을 것으로 판단되었다(10).

그러나, 본 *estI* 유전자의 3'-말단에서 transcription termination signal로 작용할 수 있는 stem-loop 구조는 발견되지 않았다.

G+C 함량

estI 유전자의 염기 서열 분석에 의한 G+C 함량은 51%, 그리고 codon의 세번째 즉 wobble position의 G+C 분포는 60.2%로 분석되었다.

일반적으로 세균 계놈의 G+C 함량 특히 wobble position의 G+C 분포량은 그 세균의 최적 생육 온도와 밀접한 상관 관계가 있는 것으로 분석되고 있다(14). 이와 관련하여 고온균(moderate thermophile)으로 분류되고 있는 *B. stearothermophilus* 균주의 *estI* 유전자의 G+C 함량 역시 Table 3과 같이 상기 생육 최적 온도와의 상관 관계에 잘 부합되는 값을 보여주고 있다.

Codon usage

다음은 *estI* 유전자 발현에 있어서 각 codon의 이용 빈도를 *E. coli* 경우(21)와 비교해 보았다. Table 4와 같이 61.2%로 산출된 wobble position의 비교적 높은 G+C 분포량은 codon의 세번째 염기가 G나 혹은 C인 codon을 선호하고 있는 특징을 보여주고 있다.

Table 1. Nucleotide sequence of putative ribosome-binding sites

Gene	Nucleotide sequence	References
<i>B. subtilis</i> α -amylase	AAAUGAAGGGAGAGG	9
<i>B. subtilis</i> β -lactamase	AACGGAGGGAGACG	9
<i>B. stearothermophilus</i> <i>xylA</i>	GAACAGGAGGTAATC	10
<i>B. stearothermophilus</i> <i>xynA</i>	TCTGAGGAGGAACAT	11
<i>B. stearothermophilus estI</i>	AAAGGAGGAACCG	in this study
<i>E. coli</i> consensus	CTACTGGAGGAAT	9
<i>B. subtilis</i> consensus	AGAAAGGAGGGTGATC	9
<i>E. coli</i> 16S rRNA	3'-GAUGACCUCCUUA-5'	9
<i>B. subtilis</i> 16S rRNA	3'-UCUUUCCUCCACUAG-5'	9

Table 2. Promoter sequences in *Bacillus* sp

Gene	-35 Region	-10 Region	Spacer length	References
<i>B. subtilis</i> σ^{55} , consensus	TTGACA	TATAAT	17-19	9
<i>B. subtilis pen</i>	TTGCAT	AATACT	17	9
<i>B. subtilis veg</i>	TTGACA	TACAAT	17	9
<i>B. subtilis tms</i>	TTGAAA	TATATT	17	9
<i>B. stearothermophilus nprT</i>	TITTCC	TATTTT	18	13
<i>B. stearothermophilus xylA</i>	TTGTTA	CATAAT	13	10
<i>B. stearothermophilus xynA</i>	TTGACC	AAGAAT	19	11
<i>B. stearothermophilus estI</i>	TTGAAT	TCCAAT	17	in this study

Table 3. Comparison of G+C content of the *B. stearothermophilus estI* gene with those of other genes of thermophile and mesophile

Strains	Gene	% G+C	% G+C at the third base of codon	Reference
Extreme thermophile <i>T. thermophilus</i>	Isopropylmalate dehydrogenase	79	89	14
Moderate thermophile <i>B. stearothermophilus</i>	Xylanase (<i>xynA</i>)	53	64	11
	β -xylosidase (<i>xylA</i>)	56	63	10
	Acetylxy lan esterase (<i>estI</i>)	51	60	in this study
	Neutral protease	58	72	13
	α -amylase	50	57	15
	Tyrosyl-tRNA synthetase	54	70	16
Mesophile <i>B. subtilis</i>	Neutral protease	44	42	17
	Subtilisin	46	41	18
<i>B. amyloliquefaciens</i>	Neutral protease	46	49	19
	Alkaline protease	49	46	19
	Average of 64 <i>E. coli</i> genes	53	55	20

그러나, CTG leu codon은 *E. coli*에서의 평균 이용 빈도인 6.8%에 비해 *estI*에서는 1.9%로 5% 전후의 큰 차이를 보이고 있으며 GTG val codon의 경우도 *estI*에서는 오히려 이용되고 있지 않는데 반해 *E. coli*에서는 2.2%의 이용 빈도를 보이는 극히 예외적인 경우도 있었다.

EstI acetylxy lan esterase의 catalytic site와 2차 구조의 분석

현재까지 보고되고 있는 대다수의 esterase와 lipase 그리고 serine protease의 경우, active site에 GX SXG의 conserved sequence를 가지고 있는 것으로 알려지고 있다(22). 그러므로, *estI* 유전자 산물인 acetylxy lan esterase도 상기 conserved sequence를 가지고 있을 것으로 예상, EstI의 아미노산 서열을 분석해 본 결과 Fig. 3과 같이 123-127 codon 사이에 GLSMG의 존재를 확인하였다. 따라서, 이 부분이 EstI 효소의 촉매 활성 부위일 가능성이 크다고 생각되며 앞으로 정확한 확인 실험을 실시할 계획이다.

또한, EstI 단백질의 hydropathy plot과 Chou와 Fasman 방법에 따라 예측해 본 2차 구조를 각각 Fig. 4와 Fig. 5에 나타내었다. EstI 효소 단백질은 친수성 부분과 소수성 부분이 교대로 반복되고 있는 일차구조를 보이고 있으며, 예상 2차 구조 역시 α -helix와 β -sheet 구조가 반복되고있어 대표적인 globular protein의 일반적인 구조를 나타내고 있음을 알 수 있었다.

또한, 본 EstI 효소는 앞 연구에서의 gel-filtration 방

Table 4. Comparison of codon usage frequency

CODON	#	%AGE	CODON	#	%AGE
TTT-Phe	4	1.5% (1.3%)	TAT-Tyr	7	2.6% (1.0%)
TTC-Phe	10	3.7% (2.2%)	TAC-Tyr	8	3.0% (1.5%)
TTA-Leu	4	1.5% (0.7%)	TAA-***	0	0.0%
TTG-Leu	7	2.6% (0.9%)	TAG-***	0	0.0%
CTT-Leu	6	2.2% (0.8%)	CAT-His	7	2.6% (0.7%)
CTC-Leu	6	2.2% (0.8%)	CAC-His	2	0.7% (1.2%)
CTA-Leu	2	0.7% (0.2%)	CAA-Gln	4	1.5% (1.0%)
CTG-Leu	5	1.9% (6.8%)	CAG-Gln	8	3.0% (3.2%)
ATT-Ile	9	3.3% (2.2%)	AAT-Asn	8	3.0% (1.0%)
ATC-Ile	7	2.6% (3.7%)	AAC-Asn	3	1.1% (2.8%)
ATA-Ile	1	0.4% (0.2%)	AAA-Lys	1	0.4% (4.1%)
ATG-Met	8	3.0% (2.8%)	AAG-Lys	10	3.7% (1.3%)
GTT-Val	3	1.1% (2.9%)	GAT-Asp	13	4.8% (2.5%)
GTC-Val	4	1.5% (1.2%)	GAC-Asp	9	3.3% (3.0%)
GTA-Val	1	0.4% (1.8%)	GAA-Glu	6	2.2% (4.9%)
GTG-Val	0	0.0% (2.2%)	GAG-Glu	8	3.0% (1.8%)
TCT-Ser	1	0.4% (1.3%)	TGT-Cys	0	0.0% (0.4%)
TCC-Ser	4	1.5% (2.0%)	TGC-Cys	1	0.4% (0.5%)
TCA-Ser	1	0.4% (0.4%)	TGA-***	1	0.4%
TCG-Ser	6	2.2% (0.6%)	TGG-Trp	4	1.5% (0.7%)
CCT-Pro	5	1.9% (0.5%)	CGT-Arg	1	0.4% (3.1%)
CCC-Pro	2	0.7% (0.3%)	CGC-Arg	2	0.7% (2.0%)
CCA-Pro	1	0.4% (0.7%)	CGA-Arg	3	1.1% (0.2%)
CCG-Pro	8	3.0% (2.5%)	CGG-Arg	4	1.5% (0.2%)
ACT-Thr	0	0.0% (1.1%)	AGT-Ser	0	0.0% (0.3%)
ACC-Thr	3	1.1% (2.4%)	AGC-Ser	3	1.1% (1.4%)
ACA-Thr	2	0.7% (0.3%)	AGA-Arg	1	0.4% (0.1%)
ACG-Thr	2	0.7% (0.8%)	AGG-Arg	2	0.7% (0.1%)
GCT-Ala	4	1.5% (2.6%)	GGT-Gly	3	1.1% (3.8%)
GCC-Ala	5	1.9% (2.2%)	GGC-Gly	7	2.6% (3.1%)
GCA-Ala	5	1.9% (2.3%)	GGA-Gly	4	1.5% (0.4%)
GCG-Ala	6	2.2% (3.2%)	GGG-Gly	8	3.0% (0.6%)

The number in parenthesis represents the codon usage frequency of *E. coli* (21). Three asterisks represent translation termination codon.

법에 의한 분자량이 약 60 kDa으로 측정되었으므로 functional acetylxy lan esterase는 *estI* 유전자 산물인 약 30 kDa의 globular protein subunit 두 개로 구성된 homodimer의 3차 구조를 가진 것으로 추측된다.

다른 유전자들과의 유사성

지금까지의 보고에 의하면 acetylxy lan esterase 생산 미생물 군종은 곰팡이와 세균을 포함해서 불과 몇 종에 지나지 않으며 더욱이 관계 유전자를 분리하여 염기 서열을 결정 보고하고 있는 예는 드물다. 따라서 현재 각 종 Data Bank에 등록되어 있는 염기 서열 자료에는 *Caldocellum saccharolyticum* acetylxy lan esterase 유전자(23) 관련 자료 외엔 다른 자료를 찾을 수 없었으며 본 *B. stearothermophilus estI* 유전자와는 염기 서열의 경우 55.7%, 아미노산 서열은 57%의 동일성을 나타내고 있어 진화적 측면에서 볼 때 두 acetylxy lan esterase 생산 유전자는 비교적

EST1.AMI	51	60	70	80	90	100
CDCXYNAB.AMI	51	A**YH**CCG	A**V**MR*	*V**L**S*	*S**E*	EYF**V**N
PSEEST5.AMI	51	GECIFPIAGT	ASDKSMW*GL	RQELSGKYRT	VA*DNR*S	STICDQPYTH
BACPBH7.AMI	51	LLL*GPTGNS	ADV**GRFL	ESKGYTCHAP	IYKGHGVFFE	ELVHTGDDW
PFXYND.AMI	51	RPPACTYIID	*EWSTGFTAN	ITLKNOTGAA	INNWNVNWQY	SSNRMTSGWN
EST1.AMI	101	110	120	130	140	150
CDCXYNAB.AMI	101	*I**E**V*	*IPQK*EK*	*****	*L*****NK	N*VG*****
PSEEST5.AMI	101	LDLAKDALS	MDAEGQLKAK	*V*H*L**MI	*QELAILAP*	RVSTLSLVNT
BACPBH7.AMI	101	WQDVMNGYE	LKNKGYEKIA	VA**L**VF	SLKL*YTVPI	EGIVTHCAPH
PFXYND.AMI	101	ANFSG*NPYN	ATNMSWNGSI	AP*Q*ISFGL	QGEKXGST*E	RPTVTG*ACN
EST1.AMI	151	160	170	180	190	200
CDCXYNAB.AMI	151	***IHK**I	*KD*RN*Y*S	*****	**S*I*****	IN**TK**Q
PSEEST5.AMI	151	*SRIDNMYRS	VIELARDWSK	TITDQ*LLNR	SLYF*ALGSK	ALGSDIFNQV
BACPBH7.AMI	151	YIKSEETMYE	GVLEYAREYK	KREG*SEEQI	EQEMEKFKQT	*MHTLKA*Q*
PFXYND.AMI	151	SATTSVV*SS	SSTPTTSSSS	ASSVASALLL	QEAQA*FCRV	DGTIDNNHTG
EST1.AMI	201	210	220	230	240	250
CDCXYNAB.AMI	201	EKGS**K**	***RD**VQ	E**DLFNF*K	N*G*DVV*E*	DE*G**D**
PSEEST5.AMI	201	VDFASGSQSQ	PREALIRQWE	IDLTVDITDR	LSL*NAKTHV	IWASE*KIVT
BACPBH7.AMI	201	LI*DVRDHL	LIYAPT*VVQ	ARHDEMHPD	SAN*IYNEI*	*PVKQKWE
PFXYND.AMI	201	FTGSGFANTN	NAQGAAVVWA	IDATSSGRR*	LTIRVANGGT	ANRNGSLVIN

Fig. 3. Amino acid sequence alignment of the *B. stearothermophilus* acetylxylylan esterase and other bacterial esterases.

*denotes identical amino acids, ESTLAMI; acetylxylylan esterase I from *B. stearothermophilus*, CDCXYNAB.AMI; acetyl esterase from *C. saccharolyticum*, PSEEST5.AMI; esterase V from *Pseudomonas* sp. KWI-56, BACPBH7.AMI; esterase from *B. stearothermophilus* IFO12550, FXYND.AMI; esterase D from *P. fluorescens* subsp. *cellulosa*

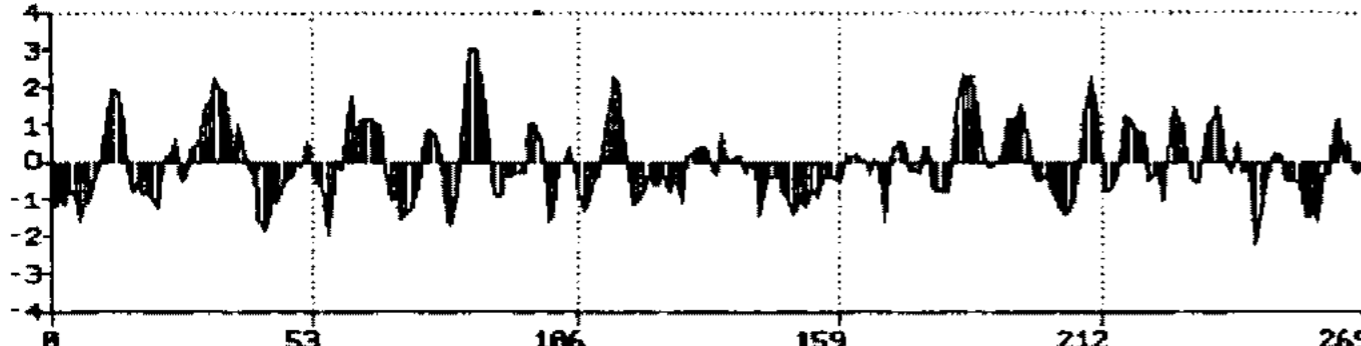


Fig. 4. Hydropathy plot of the EstI. The abscissa of the panel shows amino acid number. The ordinate shows the average hydrophobicity (positive ordinate) or hydrophilicity (negative ordinate) of amino acid residues.

근연의 관계에 있는 것으로 생각할 수 있겠다. 한편, 본 연구실에서 분리, 클로닝한 *B. stearothermophilus*의 또 다른 acetylxylylan esterase 생산 유전자인 *estII*의 경우 *estI* 효소와 매우 유사한 특성을 가지고 있는 효소를 생산하고 있으나(6), 현재 염기서열 결정이 끝나지 않아 본 *estI* 유전자와 염기서열 비교 분석은 하지 못 하였으나 매우 흥미 있는 과제로 가까운 시일 안에 실험을 착수할 예정이다.

요 약

강력한 xylan 분해 분리 균주인 *Bacillus stearothermophilus*의 acetylxylylan esterase 유전자(*estI*)의 염기서열을 결정 분석하였다. *estI* 유전자는 810 bp의 270개의 아미노산을 지시하는 ORF로 구성되어 있으며 ATG initiation codon의 9 bp 윗쪽에 SD-sequence로 추정되는 AGGAGG, 그리고 53 bp 더 윗쪽에는 promoter로 추정되는 염기서열의 존재를 확인할 수 있었다. 추정 -10 element(TCCAAT)와 -35 element(TTGAAT)는 *Bacillus subtilis* promoter consensus sequence와 높은 유

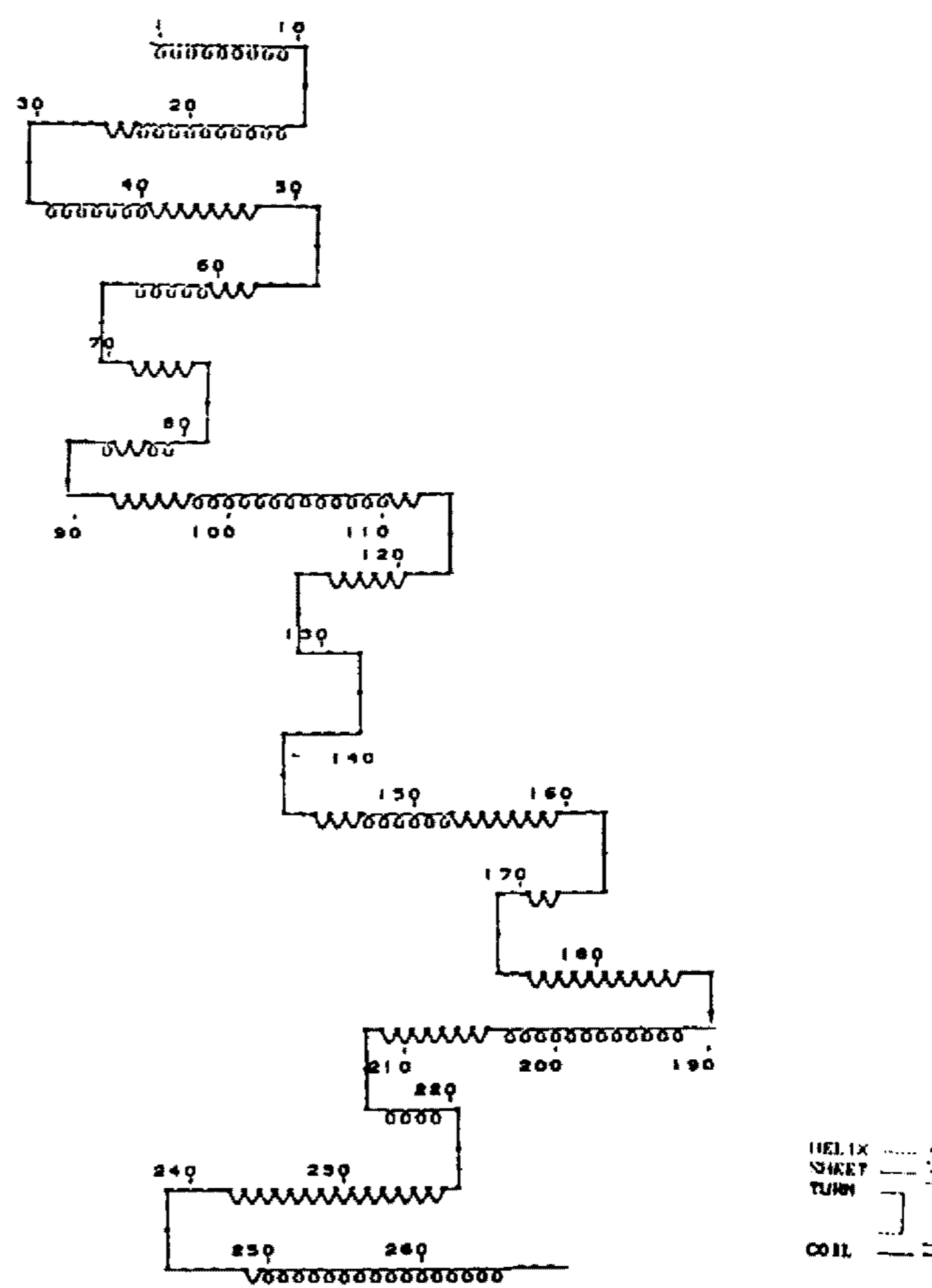


Fig. 5. Putative secondary structure of the EstI predicted by using PROSIS program This analysis was originally developed by Chou and Fasman.

사성을 나타내고 있다. *estI* 유전자의 G+C 함량은 51%, codon의 세번째 자리의 G+C 분포량은 60.2%로 고온성 세균의 G+C 함량의 일반적인 경향을 보이고 있다. 한편, 염기서열로부터 산출된 EstI의 분자량은 30 kDa으로 정제 효소의 SDS-PAGE로부터 측정된 분자량인 29 kDa과 큰 차이를 보이지 않았다. 또한 상기 ORF의 N-말단 아미노산 서열은 MAMIQIQFFSQ로서 정제 EstI 단백질의 N-말단 서열과 완전 일치되고 있다. 또 EstI 효소의 아미노산 서열을 분석해 본 결과 지금까지 보고되고 있는 다른 esterase와 lipase 등의 촉매 부위의 conserved sequence(GXSXG)와 일치되는 GLSMG(123-127 codons)를 가지고 있어 이 부분이 EstI 효소의 촉매 부위일 것으로 추정된다. 또한, *estI* 유전자의 염기서열은 지금까지 보고되고 있는 유일한 acetylxylylan esterase 유전자인 *Caldocellum saccharolyticum* xylC와 55.7%, 그리고 아미노산 서열은 57%의 동일성을 나타내고 있다.

감사의 말

이 연구는 통산산업부 에너지 자원기술개발 지원센터

의 연구비 지원을 받아 수행되었으며 이에 감사드립니다.

참고문헌

1. Lee, Yong-Eok, Lowe, Sue E., and Jeikus, J. Gregory. 1993. Regulation and Characterization of Xylanolytic Enzymes of *Thermoanaerobacterium saccharolyticum* B6A-RI. *Appl. Environ. Microbiol.* **59**: 763-771.
2. McDermid, Kevin P., Forsberg, Cecil W., and MacKenzie, C. Roger. 1990. Purification and Properties of an Acetylxylan Esterase from *Fibrobacter succinogenes* S85. *Appl. Environ. Microbiol.* **56**: 3805-3810.
3. Poutaner, K. and Sunberg, M. 1988. An Acetyl Esterase of *Trichoderma reesei* and its role in the Hydrolysis of Acetyl Xylans. *Appl. Microbiol. Biotechnol.* **28**: 419-424.
4. Gormann, K, Mitchell, D. J., Himmel, M. E., Dale, B. E. and Schroeder. H. A. 1989. The Role of Ester Groups in Resistance of Plant Cell Wall Polysaccharides to Enzymatic Hydrolysis. *Appl. Biotech.* **20/21**: 45-61.
5. Kim, H. S., Suh, J. H. and Choi, Y. J. 1995. Purification and Characterization of Acetylxylan Esterase II from *E. coli* Cells Harboring the Recombinant Plasmid pKMG6. *Kor. J. Appl. Microbiol. Biotech.* **23**(4): 454-460.
6. Kim, H. S., Cho S. G. and Choi, Y. J. 1993. Molecular Cloning and Expression of the Acetylxylan Esterase Gene of *Bacillus stearothermophilus* in *Escherichia coli*. *Kor. J. Appl. Microbiol. Biotech.* **21**: 542-548.
7. Maniatis, T., Fritsch, E. F. and Sambrook, J. 1989. Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory. New York.
8. Hanahan, D. 1985. Techniques for Transformation of *E. coli*. pp 109-135. In D. M. Glover (ed.), DNA Cloning, Vol. I. IRL Press, Oxford. 9. Moran Jr. C. P., Lang, N., LeGrice, S. F. J., Lee, G., Stephens, M. A., Sonenshein, L., Pero, J., and Losick, R. 1982. Nucleotide Sequences that signal the Initiation of Transcription and Translation in *Bacillus subtilis*. *Mol. Gen. Genet.* **186**: 339-346.
10. Oh, H. J. and Choi, Y. J. 1995. Sequence Analysis of β -xylosidase Gene from *Bacillus stearothermophilus*. *Kor. J. Appl. Microbiol. Biotech.* **22**: 134-142.
11. Cho, S. G., Park, S. S., Park, Y. I., and Choi, Y. J. 1992. Molecular Cloning and Expression of an Endo-xylanase from *Bacillus stearothermophilus* into *Escherichia coli*. *Kor. J. Appl. Microbiol. Biotech.* **20**(3): 271-279.
12. Sarry Gold. 1988. Posttranscriptional Regulatory mechanisms in *Escherichia coli*. *Ann. Rev. Biochem.* **57**: 199-233.
13. Takagi, M., T. Imanaka, and Aiba, S. 1985. Nucleotide Sequence and Promoter Region for the Neutral Protease Gene from *Bacillus stearothermophilus*. *J. Bacteriol.* **163**: 824-831.
14. Kagaqa, Y., Nojima, H., Isgizuka, K., Makajima, R., Yasuhara, R., Ranaka, R. and Oishima, R. 1984. High Guanine Plus Cytosine Content in the Third Letter of Codons of an Extreme Thermophile. *J. Biol. Chem.* **259**: 2956-2960.
15. Nakajima, T. R. Imanaka, and Aiba, S. 1985. Nucleotide Sequence of *Bacillus stearothermophilus* α -amylase Gene. *J. Bacteriol.* **163**: 404-409.
16. Winter, G., Koch, G. L. E., Hartley, B., and Barker, D. 1983. The Amino Acid Sequence of the Tyrosyl-tRNA Synthetase from *Bacillus stearothermophilus*. *Eur. J. Biochem.* **132**: 383-387.
17. Yang, M. Y., Ferrari, E., and Henner, D. J. 1984. Cloning of the Neutral Protease Gene of *Bacillus subtilis* and the Use of the Cloned Gene to Create an *In Vitro*-derived Deletion Mutation. *J. Bacteriol.* **160**: 15-21.
18. Stahl, M. L. and Ferrari, E. 1984. Replacement of the *Bacillus subtilis* Subtilisin Structural Gene with an *In Vitro*-derived Deletion Mutation. *J. Bacteriol.* **158**: 411-418.
19. Vasantha, M. L., Thomson, D., Rhodes, C., Banner, C., Nagle, J., and Filpula, D. 1984. Genes for Alkaline Protease and Neutral Protease from *Bacillus amyloliquefaciens* Contain an Large Open Reading Frame between the Regions Coding for Signal Sequence and Mature Protein. *J. Bacteriol.* **159**: 811-819.
20. Grosjean, H. and Fiers, W. 1982. Preferential Codon Usage in Prokaryotic Genes: the Optimal Codon-anticodon Interaction Energy and the Selective Usage in Efficiently Expressed Genes. *Gene* **18**: 199-209.
21. Allf-Steinberger, C. 1984. Evidence for Coding Pattern on the Noncoding Strand of the *Escherichia coli* Genome. *Nucleic Acids Res.* **12**: 235-241.
22. Sydney, Brenner. 1988. The Molecular Evolution of Genes and Proteins: a Tale of Two Serines. *Nature.* **334**: 528-530.
23. Luethi, E., Love, D. R., McAnulty, J., Wallace, C., Caughey, P. A., Saul, D., and Bergquist, P. L. 1990. Cloning, Sequence Analysis, and Expression of Genes Encoding Xylan-Degrading Enzymes from the Thermophile *Caldocellum saccharolyticum*. *Appl. Environ. Microbiol.* **56**: 1017-1024.

(Received 24 September 1996)