JOURNAL OF BACTERIOLOGY, May 1986, p. 635–643
0021-9193/86/050635-09$02.00/0

Vol. 166, No. 2

Structural Genes Encoding the Thermophilic α-Amylases of *Bacillus stearothermophilus* and *Bacillus licheniformis*

GREGORY L. GRAY,[1]* STANLEY E. MAINZER,[1] MICHAEL W. REY,[1] MICHAEL H. LAMSA,[1] KAREN L. KINDLE,[2] CYNTHIA CARMONA,[1] AND CAROL REQUADT[1]

Genencor, Inc., South San Francisco, California 94080,[1] and Department of Biochemistry, Cornell University, Ithaca, New York 14831[2]

Received 18 October 1985/Accepted 29 January 1986

The genes encoding the thermostable α-amylases of *Bacillus stearothermophilus* and *B. licheniformis* were cloned in *Escherichia coli*, and their DNA sequences were determined. The coding and deduced polypeptide sequences are 59 and 62% homologous to each other, respectively. The *B. stearothermophilus* protein differs most significantly from that of *B. licheniformis* in that it possesses a 32-residue COOH-terminal tail. Transformation of *E. coli* with vectors containing either gene resulted in the synthesis and secretion of active enzymes similar to those produced by the parental organisms. A plasmid was constructed in which the promoter and the NH_2-terminal two-thirds of the *B. stearothermophilus* coding sequence was fused out of frame to the entire mature coding sequence of the *B. licheniformis* gene. Approximately 1 in 5,000 colonies transformed with this plasmid was found to secrete an active amylase. Hybridization analysis of plasmids isolated from these amylase-positive colonies indicated that the parental coding sequences had recombined by homologous recombination. DNA sequence analysis of selected hybrid genes revealed symmetrical, nonrandom distribution of loci at which the crossovers had resolved. Several purified hybrid α-amylases were characterized and found to differ with respect to thermostability and specific activity.

The α-amylases secreted by a variety of *Bacillus* species have been intensively studied. Interest has been focused on their mode of secretion, regulation of synthesis, protein structure, and industrial applications. In recent years, the amylase genes of *B. coagulans* (4), *B. subtilis* (26), *B. amyloliquefaciens* (22), *B. licheniformis* (21), and *B. stearothermophilus* (15, 16) have been cloned and expressed in either *B. subtilis* or *Escherichia coli*.

The mesophile *B. licheniformis* and the thermophile *B. stearothermophilus* produce amylases which are active at temperatures in excess of 75°C (7). It is therefore of interest to determine their primary structures and compare them with each other and with those known for other amylases to ascertain which sequences are associated with the unusual thermophilicity of these enzymes. In this study we showed that the *B. stearothermophilus* and *B. licheniformis* enzymes differ markedly in their specific activities and thermostabilities. Primary structure analysis might also offer clues to these differences.

It has been known for some time that the mesophilic amylase of *B. amyloliquefaciens* has considerable amino acid homology with the *B. licheniformis* amylase but no homology to the *B. subtilis* enzyme (21). Recently, comparison of the *B. stearothermophilus* and *B. amyloliquefaciens* primary structures has revealed that these proteins are also evolutionarily related (16). In addition, strong similarities between the restriction endonuclease cleavage maps of the amylase genes of *B. coagulans* and *B. licheniformis* indicate that these genes may also show homology (21).

Since DNA sequence divergence has led to differing chemical properties of the encoded proteins, it is expected that further diversity in this enzyme family might be found in additional natural *Bacillus* isolates. Alternatively, methods to generate amylase DNA sequence divergence in the laboratory might also lead to enzymes with new properties.

* Corresponding author.

We cloned and determined the DNA sequences of the amylase genes of *B. stearothermophilus* and *B. licheniformis*. To obtain further diversity in this gene family, a rapid method for generating single crossovers between these genes in vivo has been devised. Selected hybrid genes created by this method were analyzed in detail, and their encoded proteins were characterized. A preliminary account of some of the data presented in this paper was previously prepared (M. W. Rey, C. Requadt, S. E. Mainzer, M. H. Lamsa, E. Ferrari, P. J. Lad, and G. L. Gray, *in* A. T. Ganesan and J. A. Hoch, ed., *Genetics and Biotechnology of Bacilli—1985*, in press).

MATERIALS AND METHODS

Bacterial strains and plasmids. *B. stearothermophilus* NZ-3 was isolated by us from the Rotorua thermal area of New Zealand. This strain has tentatively been assigned this species name because it is an obligately aerobic, gram-positive, spore-forming rod which grows optimally at 70 to 75°C. The sequence of its amylase gene is essentially identical to that from another *B. stearothermophilus* strain (16). *B. licheniformis* NCIB 8061 (14) was obtained from the National Collection of Industrial Bacteria, Aberdeen, Scotland. *B. subtilis* IA289 (*aroI906 metB5 sacA321 amyE*) was obtained from the *Bacillus* Genetic Stock Center, Ohio State University, Columbus. *E. coli* 294 (*endA1 thi-1 hsdR17*) (1) was obtained from Genentech, Inc. (South San Francisco, Calif.). Transformation of competent bacteria with plasmids was carried out by standard methods for *E. coli* (12) and *B. subtilis* (26). Transformants were selected on LB agar plates supplemented with 20 μg of ampicillin or 10 μg of chloramphenicol per ml. In experiments in which it was desired to detect amylase production, the agar was supplemented with 0.5% potato starch. Brief exposure to iodine vapor was used to reveal zones of starch hydrolysis (4). Cells were cultured in antibiotic-supplemented LB medium

for the preparation of plasmid DNA. Plasmid DNA was prepared by a cleared-lysate method (13).

Isolation of amylase clones. High-molecular-weight genomic DNA was prepared from *B. licheniformis* or *B. stearothermophilus* cells cultured in Saito medium (18). The DNA was purified from lysozyme lysates as previously described (15), except that the lysate was digested with proteinase K (10 μg/ml) prior to centrifugation in cesium chloride gradients. The DNA was partially digested with *Sau*3A, and DNA fragments larger than 6 kilobases (kb) were separated on 10 to 40% sucrose gradients. The bacteriophage vector λ1059 (8) was digested with *Bam*HI and treated with calf intestinal alkaline phosphatase to minimize self-ligation. We ligated and packaged vector and bacterial DNA fragments in vitro by using a commercial (Promega Biotec, Madison, Wis.) packaging extract and then used them to infect *E. coli* Q358 and Q359 (8). The number of recombinant plaques was approximately 2.5×10^3/μg of DNA in typical reactions. Approximately 5.0×10^3 plaques were screened on LB-starch agar plates for amylase activity. Five positive plaques were found in the λ1059-*B. licheniformis* library, and three positive plaques were found in the λ1059-*B. stearothermophilus* library.

Plasmid constructions. DNA was prepared from one of the λ1059-*B. licheniformis* amylase bacteriophage by standard methods (13). The DNA was digested with *Bam*HI and *Eco*RI and subcloned into similarly digested pBR322. Restriction endonuclease analysis of plasmids isolated from amylase-producing colonies showed that the entire amylase gene (*amyL*) was contained on a 9.4-kb *Bam*HI fragment or a 3.2-kb *Eco*RI fragment. One *Eco*RI subclone, designated pBR322BL, was saved for further analysis.

DNA prepared from one of the λ1059-*B. stearothermophilus* amylase bacteriophage was digested with a variety of restriction enzymes, separated on a 1.0% agarose gel, transferred to a nitrocellulose filter, and subjected to DNA-DNA hybridization. The probe used was a *Pst*I-*Sal*I fragment containing most of the *B. licheniformis* amylase coding sequence (see Fig. 2C). This fragment was radioactively labeled with [α-^{32}P]ATP by the nick translation method (13). Hybridization and washing were performed under standard high-stringency conditions (6). A 1.8-kb *Bam*HI fragment and a 1.8-kb *Sal*I fragment were among the positively hybridizing bands. These fragments were separately subcloned into *Bam*HI- and *Sal*I-cleaved pBR322, yielding pBR322BS-B and pBR322BS-S, respectively. Shown in Fig. 1 are the steps in which these plasmids were used to derive pUC13BS, which contains the entire *B. stearothermophilus* amylase gene (*amyS*).

We joined the *E. coli-B. subtilis* shuttle vector (Cm^r) pBS42 (25) cleaved with *Eco*RI to the *amyL*-containing *Eco*RI fragment of pBR322BL to produce the *amyL* expression vector pBS42BL. We joined pBS42 cleaved with *Bam*HI and *Xba*I to the *amyS* containing *Bam*HI-*Xba*I fragment of pUC13BS to produce the *amyS* expression vector pBS42BS. (The *Xba*I site at the 3′ end of the *amyS* gene is derived from the pUC13 polylinker sequence [23].)

DNA sequencing. Restriction endonuclease maps of the *amy* gene containing inserts of plasmids pBR322BS-S, pBR322BS-B, and pBR322BL were generated, and various subfragments were subcloned into the M13 cloning vectors mp18 and mp19 (17). Preliminary sequencing by the dideoxy chain termination method was performed (19). It was found that, for both the *amyS* and *amyL* genes, sequences near the *Kpn*I site were highly homologous to those of the *amy* gene of *B. amyloliquefaciens* (*amyA*) (22). The assumption that all three of these *amy* genes were closely similar allowed prediction of which regions of the *B. licheniformis* and *B. stearothermophilus* subclones were likely to comprise the *amy* genes. These regions were sequenced completely. We sequenced the crossover regions in the recombined *amy* genes of plasmids pα1.hyb1 through pα1.hyb18 by collapsing the supercoils with sodium hydroxide followed by enzymatic sequencing (3) using as primers the synthetic oligonucleotides described below.

Generation of hybrid amylase plasmids. A single pα1 transformant colony of *E. coli* 294 was grown in ampicillin-supplemented LB broth. This passage of plasmid pα1 in *E. coli* 294 resulted in a low level of recombination between the regions of homology between the *amyS* and *amyL* genes. To reduce the proportion of unrecombined plasmid in the plasmid preparation, we digested the DNA with *Pst*I prior to retransformation. Transformants containing recombinant *amy* genes were detected by their ability to generate zones of starch hydrolysis on LB-starch agar plates.

Mapping of crossovers in plasmids pα1.hyb1 through pα1.hyb18. Nine sequences of 16 to 22 bases of the *amyS* gene separated by intervals of about 150 to 200 base pairs (bp) were selected. Corresponding oligonucleotides were synthesized by the triester method (5) and then end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase for use as hybridization probes to detect *amyS* sequences in hybrid *amy* genes. Colonies postulated to contain plasmids with

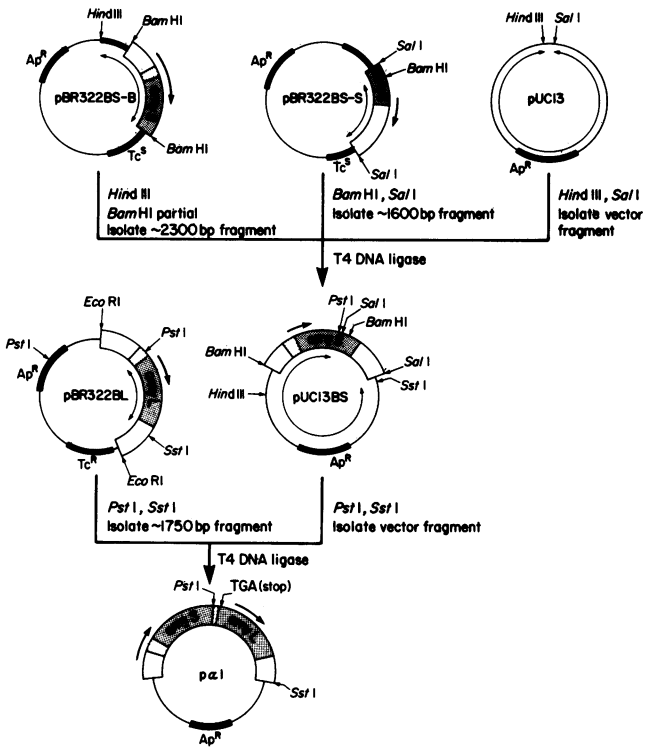


FIG. 1. Construction of pUC13BS and pα1. The coding regions for mature amylase polypeptides are indicated by the shaded boxes; cross-hatched boxes indicate signal peptide codons; 5′ and 3′ noncoding sequences are shown as open boxes. Single-headed arrows show the direction of transcription. The fragments isolated are shown by double-headed arrows. The TGA stop codon occurring early in the out-of-frame *amyL* coding region of pα1 is shown. This figure has been published elsewhere (Rey et al., in press; published here by permission from Academic Press, Inc.).

hybrid *amy* genes were inoculated onto each of nine nitrocellulose filter strips. We placed these strips on an LB agar plate to allow colony growth. The filters were then prepared for colony hybridization by standard methods (6). In separate vessels, the nine strips were incubated with each of the probes under conditions for low-stringency hybridization (6). The filters were washed in 0.3 M NaCl–0.03 M sodium citrate (pH 7.0)–0.1% sodium dodecyl sulfate at 37°C and then autoradiographed.

Cellular localization of amylase in *E. coli*. Transformants were grown in LB medium supplemented with 20 μg of ampicillin per ml. We added isopropyl β-D-thiogalactoside (1 mM) to induce the cytoplasmic marker enzyme β-galactosidase. Cells were pelleted by centrifugation, and the culture supernatants were saved. The cells were then fractionated by the osmotic shock method (9). In this procedure, cells are first washed in a buffer which sensitizes them to osmotic shock. We have found that this wash often results in release of a significant portion of the periplasmic enzyme β-lactamase but very little of the cytoplasmic enzyme β-galactosidase. Thus, we consider the sum of activities of the wash buffer and the actual osmotic shock fluid to represent the periplasmic component. The shocked cells contain cytoplasmic and membrane-bound enzymes and are assayed after the cells are disrupted by sonication.

Characterization of parental and hybrid amylases. For production of amylase in the native (*Bacillus*) hosts, cells were grown in starch-supplemented Saito medium. The enzymes from cloned DNA were obtained from the periplasmic fractions of *E. coli* transformants as described above. To purify the enzymes, we adjusted the periplasmic fractions to 50 mM morpholinepropanesulfonic acid (MOPS; pH 7.0)–5 mM $CaCl_2$ (MC buffer) and poured them at 4°C over a column containing insoluble starch, which resulted in amylase binding to the starch. The enzymes were eluted from the column in MC buffer by raising the temperature to 50°C. Starch was removed from the enzymes by gel permeation chromatography on a P-2 column (Bio-Rad Laboratories, Richmond, Calif.) followed by DEAE chromatography. The proteins were judged pure by their homogeneity on Coomassie blue-stained sodium dodecyl sulfate-polyacrylamide gels (11).

Purified amylase concentrations were determined by the dye binding method of Bradford (2). To measure amylase specific activities, we assayed equal concentrations of purified enzymes over 60-min periods by the Phaedebas amylase assay (Pharmacia, Inc., Piscataway, N.J.). Specific activities are expressed as units per microgram of enzyme. One unit is the amount of enzyme that catalyzes the hydrolysis of 1 μmol of glucosidic linkages per min. To measure amylase thermostabilities, we incubated equal concentrations of purified enzyme at 90°C for 0 to 120 min in MC buffer and then assayed them using starch as a substrate. The reducing equivalents generated in 10 min were measured.

RESULTS

Nucleotide sequence of *amy* genes. The DNA sequence of the cloned *amyS* gene is shown in Fig. 2B. The first 16 amino acids of the deduced mature coding sequence correspond completely to the NH_2-terminal sequence as determined by amino acid sequence analysis of the purified secreted protein (K. Hayenga, personal communication). The GTG, which is inferred to be the initiation codon and corresponds to methionine, is located 102 bp upstream of the mature coding sequence and begins the code for a larger precursor containing a highly hydrophobic sequence of 34 amino acids (positions −34 through −1), presumed to be the amylase signal peptide. The mature form of the amylase contains 515 residues.

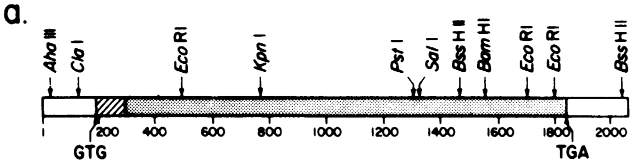
The nucleotide sequence of the cloned *amyL* gene is shown in Fig. 2D. The deduced NH_2 terminus of the mature enzyme (beginning with position +1) is in agreement with that determined by direct amino acid sequencing (10). Located 87 bp upstream of the mature coding sequence is an ATG, presumed to be the initiation codon, which begins the code for a larger precursor containing a highly hydrophobic sequence of 29 residues presumed to be the signal peptide. The mature form of the amylase contains 483 residues.

The *amyS* and *amyL* nucleotide sequences were compared (Fig. 3). Overall homology was 59%, clearly indicating evolution from a common ancestral gene. The inferred primary amino acid sequences are very similar (62% homology), with two exceptions (Fig. 4). First, there is no obvious homology in the signal peptides per se, although their coding sequences are 42% homologous. This emphasizes the extreme amino acid sequence divergence which can occur without impairing the function of a signal peptide. Second, the COOH terminus of the *amyS* gene product extends 32 amino acid residues beyond that of the *amyL* protein. The product of the *B. amyloliquefaciens amy* gene (*amyA*) ends in a sequence nearly identical to that of *amyL*. It thus appears likely that the *amyS*-encoded COOH-terminal tail may be a distinctive feature of the *amyS* protein. The DNA sequence homologies of the various functional regions of the *amy* genes vary greatly. The coding sequences exclusive of the *amyS* region encoding the COOH-terminal tail are 63% homologous, the adjacent 3′ sequences are 43% homologous, and the 5′ flanking sequences are 32% homologous. The relatively low degree of homology in the untranslated sequences is perhaps reflective of the great latitude of allowable sequences for promoters and transcription terminators.

As noted above, the *amyL* and *amyA* proteins are similar in sequence at their COOH termini. High homology (79%; data not shown) between the entire *amyL* and *amyA* mature polypeptides indicates that the *amyL* protein is structurally much closer to that of *amyA* than to that of *amyS* (62% homology).

Expression of *amy* genes in *E. coli* and *B. subtilis*. *E. coli* transformed with plasmid pUC13BS, which contains the *B. stearothermophilus* amylase gene (*amyS*), or plasmid pBR322BL, which contains the *B. licheniformis* amylase gene (*amyL*), produced amylase as indicated by the production of halos of starch hydrolysis on starch-containing agar plates (data not shown). The presence of halos indicated that at least a fraction of the amylases were produced as extracellular products, either by active secretion or as a result of partial cell lysis. To further localize the amylase activities, we separated cells from late-logarithmic-phase liquid cultures from the culture supernatants and then further divided them into periplasmic and cytoplasm-membrane cell fractions (9). In both 294(pUC13BS) and 294(pBR322BL) cells, most of the amylase activities were found in the periplasmic fractions (Table 1). Smaller amounts were present in the cytoplasm-membrane and culture supernatant fractions. The presence of most of the periplasmic enzyme β-lactamase in the periplasmic fraction indicated that periplasmic protein contents were released by the procedure. Significant cell lysis during culture or as a result of the cell fractionation procedure was ruled out by the near absence of the cytoplasmic enzyme β-galactosidase in the culture supernatant fractions. Amylase-negative *B. subtilis* IA289 transformed

a.



b.

TTCTATTTTCGATTCGTGTCAAACTCAAAATTGTTTAAATTCGATATTGAAAACGATTACAAATAAAAATTATAATGGG

ACGTAAACGTTCGAGGGTTTTGGCTTCCTTTTTAGCTCTTTTTATGCAATCGATTTCCTTAATTTTTTGGAAGCCAAAC

-34 -30
fmet leu thr phe his arg ile ile arg lys
CGTCGAATGCAACATTTGATTAAGGGGGAAGGGCATT GTG CTA ACG TTT CAC CGC ATC ATT CGA AAA

-20 -10
gly trp val phe leu leu ala phe trp leu thr ala ser leu phe cys pro thr gly gln
GGG TGG GTG TTC CTG CTC GCG TTT TGG CTC ACT GCC TCG CTG TTC TGC CCG ACA GGA CAG

1 10
pro ala lys ala ala ala pro phe asn gly thr met met gln tyr phe glu trp tyr leu
CCC GCC AAG GCT GCC GCA CCG TTT AAC GGC ACC ATG ATG CAG TAT TTT GAA TGG TAC TTG

20 30
pro asp asp gly thr leu trp thr lys val ala asn glu ala asn asn leu ser ser leu
CCG GAT GAT GGC ACG TTA TGG ACC AAA GTG GCC AAT GAA GCC AAC AAC TTA TCC AGC CTT

40 50
gly ile thr ala leu trp leu pro pro ala tyr lys gly thr ser arg ser asp val gly
GGC ATC ACC GCT CTT TGG CTG CCG CCC GCT TAT AAA GGA ACA AGC CGC AGC GAC GTA GGG

60 70
tyr gly val tyr asp leu tyr asp leu gly glu phe asn gln lys gly thr val arg thr
TAC GGA GTA TAC GAC TTG TAT GAC CTC GGC GAA TTC AAT CAA AAA GGG ACC GTC CGC ACA

80 90
lys tyr gly thr lys ala gln tyr leu gln ala ile gln ala ala his ala ala gly met
AAA TAC GGA ACA AAA GCT CAA TAT CTT CAA GCC ATT CAA GCC GCC CAC GCC GCT GGA ATG

100 110
gln val tyr ala asp val val phe asp his lys gly gly ala asp gly thr glu trp val
CAA GTG TAC GCC GAT GTC GTG TTC GAC CAT AAA GGC GGC GCC GAC GGC ACG GAA TGG GTG

120 130
asp ala val glu val asn pro ser asp arg asn gln glu ile ser gly thr tyr gln ile
GAC GCC GTC GAA GTC AAT CCG TCC GAC CGC AAC CAA GAA ATC TCG GGC ACC TAT CAA ATC

140 150
gln ala trp thr lys phe asp phe asn gly arg gly asn thr tyr ser ser phe lys trp
CAA GCA TGG ACG AAA TTT GAT TTT AAC GGG CGG GGC AAC ACC TAC TCC AGC TTT AAG TGG

160 170
arg trp tyr his phe asp gly val asp trp asp glu ser arg lys leu ser arg ile tyr
CGC TGG TAC CAT TTT GAC GGC GTT GAC TGG GAC GAA AGC CGA AAA TTA AGC CGC ATT TAC

180 190
lys phe arg gly ile gly lys ala trp asp trp glu val asp thr glu asn gly asn tyr
AAA TTC CGC GGC ATC GGC AAA GCG TGG GAT TGG GAA GTA GAC ACG GAA AAC GGA AAC TAT

200 210
asp tyr leu met tyr ala asp leu asp met asp his pro glu val val thr glu leu lys
GAC TAC TTA ATG TAT GCC GAC CTT GAT ATG GAT CAT CCC GAA GTC GTG ACC GAG CTG AAA

220 230
asn trp gly lys trp tyr val asn thr thr asn ile asp gly phe arg leu asp ala val
AAC TGG GGG AAA TGG TAT GTC AAC ACA ACG AAC ATT GAT GGG TTC CGG CTT GAT GCT GTC

240 250
lys his ile lys phe ser phe phe pro asp trp leu ser tyr val arg ser gln thr gly
AAG CAT ATT AAG TTC AGT TTT TTT CCT GAT TGG TTG TCG TAT GTG CGT TCT CAG ACT GGC

260 270
lys pro leu phe thr val gly glu tyr trp ser tyr asp ile asn lys leu his asn tyr
AAG CCG CTA TTT ACC GTC GGG GAA TAT TGG AGC TAT GAC ATC AAC AAG TTG CAC AAT TAC

280 290
ile thr lys thr asn gly thr met ser leu phe asp ala pro leu his asn lys phe tyr
ATT ACG AAA ACA AAC GGA ACG ATG TCT TTG TTT GAT GCC CCG TTA CAC AAC AAA TTT TAT

300 310
thr ala ser lys ser gly gly ala phe asp met ser thr leu met asn asn thr leu met
ACT GCT TCC AAA TCG GGG GGC GCA TTT GAT ATG AGC ACG TTA ATG AAC AAT ACT CTC ATG

320 330
lys asp gln pro thr leu ala val thr phe val asp asn his asp thr glu pro gly gln
AAA GAT CAA CCG ACA TTG GCC GTC ACC TTC GTT GAT AAT CAT GAC ACC GAA CCC GGC CAA

340 350
ala leu gln ser trp val asp pro trp phe lys pro leu ala tyr ala phe ile leu thr
GCG CTG CAG TCA TGG GTC GAC CCA TGG TTC AAA CCG TTG GCT TAC GCC TTT ATT CTA ACT

360 370
arg gln glu gly tyr pro cys val phe tyr gly asp tyr tyr gly ile pro gln tyr asn
CGG CAG GAA GGA TAC CCG TGC GTC TTT TAT GGT GAC TAT TAT GGC ATT CCA CAA TAT AAC

380 390
ile pro ser leu lys ser lys ile asp pro leu leu ile ala arg arg asp tyr ala tyr
ATT CCT TCA CTG AAA AGC AAA ATC GAT CCG CTC CTC ATC GCG CGC AGG GAT TAT GCT TAT

400 410
gly thr gln his asp tyr leu asp his ser asp ile ile gly trp thr arg glu gly val
GGA ACG CAA CAT GAT TAT CTT GAT CAC TCC GAC ATC ATC GGG TGG ACA AGG GAA GGC GTT

420 430
thr glu lys pro gly ser gly leu ala ala leu ile thr asp gly pro gly gly ser lys
ACC GAA AAA CCA GGA TCC GGA CTG GCC GCA CTG ATC ACC GAT GGG CCG GGA GGA AGC AAA

440 450
trp met tyr val gly lys gln his ala gly lys val phe tyr asp leu thr gly asn arg
TGG ATG TAC GTT GGC AAA CAA CAC GCC GGA AAA GTG TTC TAT GAC CTT ACC GGC AAC CGG

460 470
ser asp thr val thr ile asn ser asp gly trp gly glu phe lys val asn gly gly ser
AGT GAC ACC GTC ACC ATC AAC AGT GAT GGA TGG GGG GAA TTC AAA GTC AAT GGC GGT TCG

480 490
val ser val trp val pro arg lys thr thr val ser thr ile ala trp pro ile thr thr
GTT TCG GTT TGG GTT CCT AGA AAA ACG ACC GTC TCT ACT ATC GCT TGG CCG ATC ACA ACC

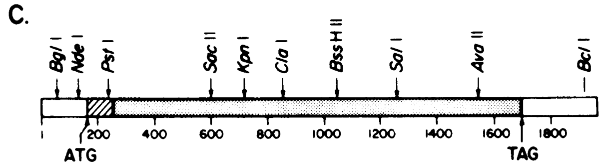
500 510 515
arg pro trp thr gly glu phe val arg trp thr glu pro arg leu val ala trp pro OP
CGA CCG TGG ACT GGT GAA TTC GTC CGT TGG ACC GAA CCA CGG TTG GTG GCA TGG CCT TGA

TGCCTGCGATCGCGTTGTAAAGACATTCCGCTCTATCATTGAGGCAAAAACATGGCCTTGTCCGCCATGAATGGCGG

CACAAGGCCGTGTTTGATGTTACCATCCATTTGCTTGCTTCAACTTTTCCTTCGACGGCGTTTCGTAGCGGATGTGCG

TGTCGATGTCGGTCACGTAATACCCGCCGCCGACCGCGTATTGCCCGCGAAGCGCGCGTCATACC

c.



d.

AGCTTGAAGAAGTGAAGAAGCAGAGAGGCTATTGAATAAATGAGTAGAAAGCGCCATATCGGCGCTTTTCTTTTGGAAG

AAAATATAGGGAAAATGGTACTTGTTAAAAATTCGGAATATTTATACAACATCATATGTTTCACATTGAAAGGGGAGGA

[illegible] [illegible]
fmet lys glu gln lys arg leu tyr ala arg leu leu thr [illegible] leu phe ala leu
GAATC ATG AAA CAA CAA AAA CGG CTT TAC GCT CGA TTG CTG ACG CTG TTA TTT GCG CTC

-10 1
ile phe leu leu pro his ser ala ala ala ala ala asn leu asn gly thr leu met gln
ATC TTC TTG CTG CCT CAT TCT GCA GCA GCG GCG GCA AAT CTT AAT GGG ACG CTG ATG CAG

10 20
tyr phe glu trp tyr met pro asn asp gly gln his trp lys arg leu gln asn asp ser
TAT TTT GAA TGG TAC ATG CCC AAT GAC GGC CAA CAT TGG AAG CGT TTG CAA AAC GAC TCG

30 40
ala tyr leu ala glu his gly ile thr ala val trp ile pro pro ala tyr lys gly thr
GCA TAT TTG GCT GAA CAC GGT ATT ACT GCC GTC TGG ATT CCC CCG GCA TAT AAG GGA ACG

50 60
ser gln ala asp val gly tyr gly ala tyr asp leu tyr asp leu gly glu phe his gln
AGC CAA GCG GAT GTG GGC TAC GGT GCT TAC GAC CTT TAT GAT TTA GGG GAG TTT CAT CAA

70 80
lys gly thr val arg thr lys tyr gly thr lys gly glu leu gln ser ala ile lys ser
AAA GGG ACG GTT CGG ACA AAG TAC GGC ACA AAA GGA GAG CTG CAA TCT GCG ATC AAA AGT

90 100
leu his ser arg asp ile asn val tyr gly asp val val ile asn his lys gly gly ala
CTT CAT TCC CGC GAC ATT AAC GTT TAC GGG GAT GTG GTC ATC AAC CAC AAA GGC GGC GCT

110 120
asp ala thr glu asp val thr ala val glu val asp pro ala asp arg asn arg val ile
GAT GCG ACC GAA GAT GTA ACC GCG GTT GAA GTC GAT CCC GCT GAC CGC AAC CGC GTA ATT

130 140
ser gly glu his leu ile lys ala trp thr his phe his phe pro gly arg gly ser thr
TCA GGA GAA CAC CTA ATT AAA GCC TGG ACA CAT TTT CAT TTT CCG GGG CGC GGC AGC ACA

150 160
tyr ser asp phe lys trp his trp tyr his phe asp gly thr asp trp asp glu ser arg
TAC AGC GAT TTT AAA TGG CAT TGG TAC CAT TTT GAC GGA ACC GAT TGG GAC GAG TCC CGA

170 180
lys leu asn arg ile tyr lys phe gln gly lys ala trp asp trp glu val ser asn glu
AAG CTG AAC CGC ATC TAT AAG TTT CAA GGA AAG GCT TGG GAT TGG GAA GTT TCC AAT GAA

190 200
asn gly asn tyr asp tyr leu met tyr ala asp ile asp tyr asp his pro asp val ala
AAC GGC AAC TAT GAT TAT TTG ATG TAT GCC GAC ATC GAT TAT GAC CAT CCT GAT GTC GCA

210 220
ala glu ile lys arg trp gly thr trp tyr ala asn glu leu gln leu asp gly phe arg
GCA GAA ATT AAG AGA TGG GGC ACT TGG TAT GCC AAT GAA CTG CAA TTG GAC GGT TTC CGT

230 240
leu asp ala val lys his ile lys phe ser phe leu arg asp trp val asn his val arg
CTT GAT GCT GTC AAA CAC ATT AAA TTT TCT TTT TTG CGG GAT TGG GTT AAT CAT GTC AGG

250 260
glu lys thr gly lys glu met phe thr val ala glu tyr trp gln asn asp leu gly ala
GAA AAA ACG GGG AAG GAA ATG TTT ACG GTA GCT GAA TAT TGG CAG AAT GAC TTG GGC GCG

270 280
leu glu asn tyr leu asn lys thr asn phe asn his ser val phe asp val pro leu his
CTG GAA AAC TAT TTG AAC AAA ACA AAT TTT AAT CAT TCA GTG TTT GAC GTG CCG CTT CAT

290 300
tyr gln phe his ala ala ser thr gln gly gly gly tyr asp met arg lys leu leu asn
TAT CAG TTC CAT GCT GCA TCG ACA CAG GGA GGC GGC TAT GAT ATG AGG AAA TTG CTG AAC

310 320
gly thr val val ser lys his pro leu lys ser val thr phe val asp asn his asp thr
GGT ACG GTC GTT TCC AAG CAT CCG TTG AAA TCG GTT ACA TTT GTC GAT AAC CAT GAT ACA

330 340
gln pro gly gln ser leu glu ser thr val gln thr trp phe lys pro leu ala tyr ala
CAG CCG GGG CAA TCG CTT GAG TCG ACT GTC CAA ACA TGG TTT AAG CCG CTT GCT TAC GCT

350 360
phe ile leu thr arg glu ser gly tyr pro gln val phe tyr gly asp met tyr gly thr
TTT ATT CTC ACA AGG GAA TCT GGA TAC CCT CAG GTT TTC TAC GGG GAT ATG TAC GGG ACG

370 380
lys gly asp ser gln arg glu ile pro ala leu lys his lys ile glu pro ile leu lys
AAA GGA GAC TCC CAG CGC GAA ATT CCT GCC TTG AAA CAC AAA ATT GAA CCG ATC TTA AAA

390 400
ala arg lys gln tyr ala tyr gly ala gln his asp tyr phe asp his his asp ile val
GCG AGA AAA CAG TAT GCG TAC GGA GCA CAG CAT GAT TAT TTC GAC CAC CAT GAC ATT GTC

410 420
gly trp thr arg glu gly asp ser ser val ala asn ser gly leu ala ala leu ile thr
GGC TGG ACA AGG GAA GGC GAC AGC TCG GTT GCA AAT TCA GGT TTG GCG GCA TTA ATA ACA

430 440
asp gly pro gly gly ala lys arg met tyr val gly arg gln asn ala gly glu thr trp
GAC GGA CCC GGT GGG GCA AAG CGA ATG TAT GTC GGC CGG CAA AAC GCC GGT GAG ACA TGG

450 460
his asp ile thr gly asn arg ser glu pro val val ile asn ser glu gly trp gly glu
CAT GAC ATT ACC GGA AAC CGT TCG GAG CCG GTT GTC ATC AAT TCG GAA GGC TGG GGA GAG

470 480 483
phe his val asn gly gly ser val ser ile tyr val gln arg AM
TTT CAC GTA AAC GGC GGG TCG GTT TCA ATT TAT GTT CAA AGA TAG AAGAGCAGAGAGGACGGA

TTTCCTGAAGGAAATCCGTTTTTTTATTTTGCCCGTCTTATAAATTTCTTTGATTACATTTTATAATTAATTTTAACA

AAGTGTCATCAGCCCTCAGGAAGGACTTGCTGACAGTTTGAATCGCATAGGTAAGGCGGGGATGAAATGGCAACGTTA

TCTGATGTAGCAAAGAAAGCAAATGTGTCGAAAATGACGGTATCGCGGGTGATCAATCATCCTGAGACTGTGACGGAT

GAATTGAAAAAGCT

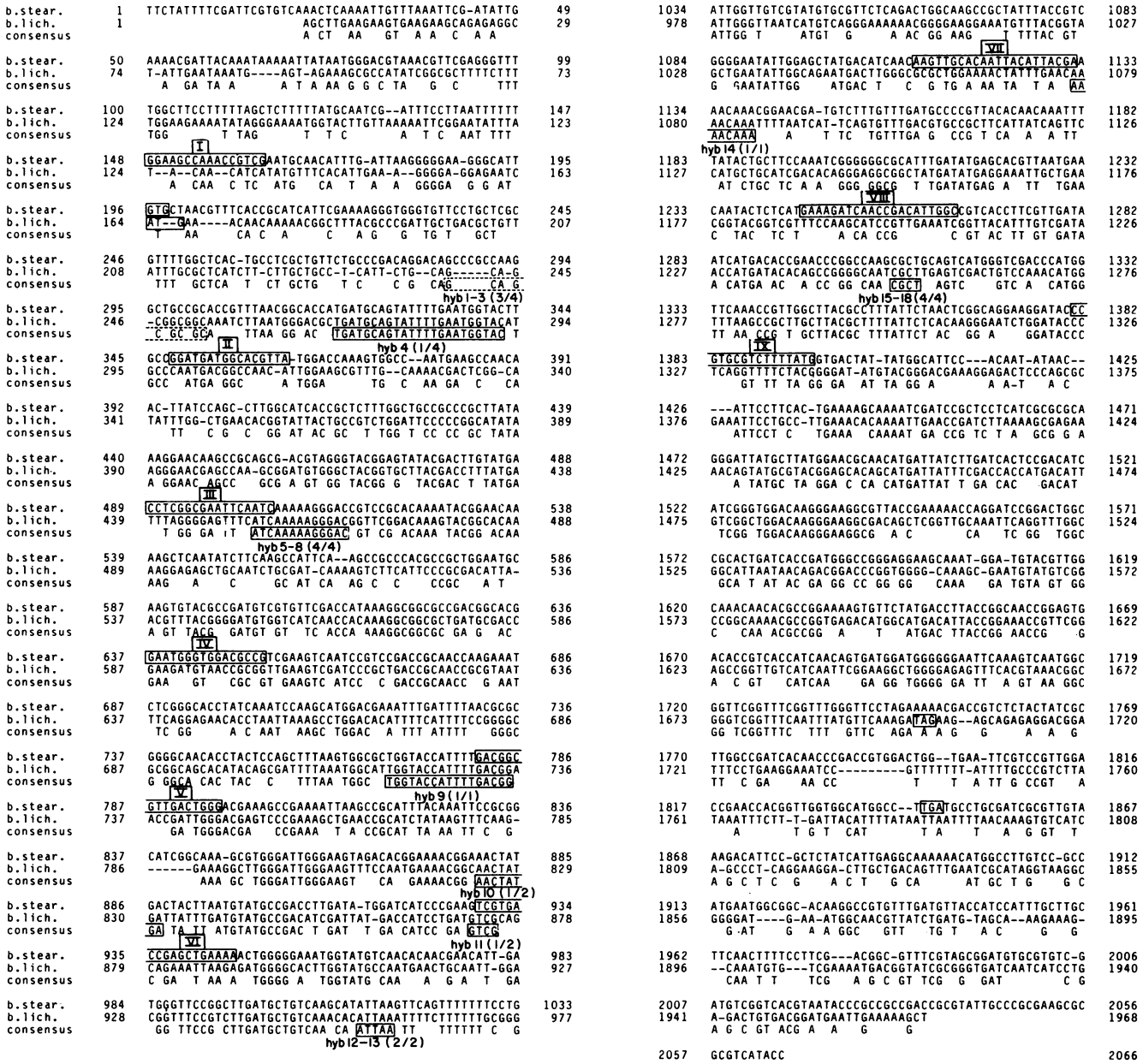


FIG. 3. Comparison of *amyS* and *amyL* DNA sequences. The upper lines represent the *amyS* gene, the *amyL* gene is shown in the middle lines, and the sequences common to both genes appear in the lower lines. The *amyS* initiation codon (positions 196 to 198), *amyL* initiation codon (positions 164 to 166), *amyS* termination codon (positions 1700 to 1702), and *amyL* termination codon (positions 1843 to 1845) are boxed. The sequences of oligonucleotide probes I to IX are also shown in boxes on the upper lines. The crossover positions for plasmids pα1.hyb1 through pα1.hyb18 and their frequencies within the probe region in which they occur are shown boxed on the lower (homology) lines. The box corresponding to pα1.hyb1 through pα1.hyb3 is stippled to indicate that each of these three crossover sequences is slightly different.

FIG. 2 (facing page). (a) Restriction endonuclease map of the *B. stearothermophilus* amylase gene (*amyS*). (b) Nucleotide sequence of the *amyS* gene. (c) Restriction endonuclease map of the *B. licheniformis* gene (*amyL*). (d) Nucleotide sequence of the *amyL* gene. The shaded regions of a and c represent the coding regions for the mature polypeptides. The leader peptides are indicated by cross-hatched regions, and the open regions show the 5′ and 3′ noncoding sequences. In b and d, the signal peptides are represented by residues with negative numbers, and the mature polypeptides are represented by residues with positive numbers. The boxed sequences are the postulated Shine-Delgarno regions, the underlined sequences represent possible −10 and −35 sequences, and potential transcription terminator sequences are overlined.

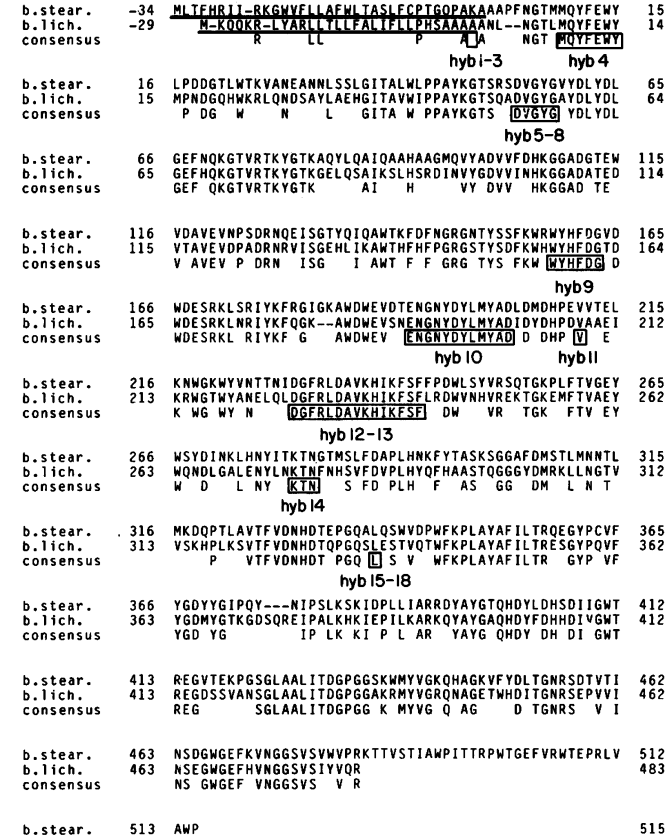


FIG. 4. Comparison of the *amyS* and *amyL* amino acid sequences. The upper lines represent the *amyS* protein, the *amyL* protein is shown in the middle lines, and the lower lines show the residues which the two amylases have in common. The signal peptides are underlined. The residues corresponding to the crossover positions are shown in boxes on the lower lines. For hybrid proteins 1 to 3 a nonhomologous residue is boxed, indicating crossovers just 5′ of the homologous regions (see Results).

with the *amyS*-containing plasmid pBS42BS or the *amyL*-containing plasmid pBS42BL was also found to produce α-amylase (data not shown).

We purified the *amyS* and *amyL* gene products to homogeneity (see Materials and Methods) from *E. coli* periplasmic extracts to compare them to the enzymes derived from the parent species. We have shown that the cloned enzymes had the same mobilities on sodium dodecyl sulfate-polyacrylamide gels as the natural enzymes, indicating that correct processing probably occurred in *E. coli* (Rey et al., in press). However, proof of this must await determination of the NH_2-terminal amino acid sequences of the *E. coli*-derived amylases. To further compare the cloned and natural products, we also measured the relative enzyme activities as a function of temperature over the range 20 to 90°C. The profiles of the amylases derived from *E. coli* were indistinguishable from those of the authentic proteins (data not shown). These and other results indicated that the amylases synthesized in *E. coli* transformants are similar to those secreted by their natural hosts.

Generation of hybrid *amy* genes in vivo. The construction of the hybrid precursor plasmid pα1 is outlined in Fig. 1. It contains the promoter and codons −34 to +338 of the *amyS* gene joined at a naturally occurring *Pst*I site to codons −4 to +484 of the *amyL* gene. This *Pst*I site is unique in pα1. Since the two *amy* fragments are joined in the same orientation but out of frame, the translation product of pα1 is expected to contain the *amyS* residues followed by a nonsense *amyL* region terminating prematurely at the first stop codon (TGA contained in codons +7 and +8) and therefore to be enzymatically inactive.

Homologous recombination can occur over the homologous region of the *amyS* and *amyL* genes present in pα1, that is, between codons −4 and +338. Such recombination events result in deletion of the unique *Pst*I site between the *amy* gene fragments. Since only recombined plasmids should be resistant to *Pst*I cleavage, predigestion with this enzyme should enrich for recombinant plasmids in a plasmid preparation containing only a small proportion of recombined plasmids. In one experiment, after transformation of *E. coli* 294 with 0.1 ng of pα1 DNA, approximately 800 Ap^r colonies arose. None of these produced amylase. Indeed, only about 1 in 5,000 such colonies was amylase positive. In contrast, transformation with 10 ng of *Pst*I-digested pα1 DNA resulted in only 49 Ap^r colonies, but 14 of these were amylase positive. Thus, *Pst*I digestion reduced the background of unrecombined pα1 over 1,000-fold in this experiment, thereby allowing the appearance of numerous well-isolated amylase-positive colonies on a single plate. It is possible that a portion of the amylase-positive colonies which arose from *Pst*I-digested pα1 may have resulted from transformation

TABLE 1. Localization of amylases in *E. coli* 294 cells

Plasmid and fraction	% of total		
	α-Amylase	β-Lactamase	β-Galactosidase
pUC13BS			
Culture supernatant	0.7	18.8	0.0
Periplasm	87.7	81.2	1.2
Cytoplasm-membranes	11.6	0.0	98.8
pBR322BL			
Culture supernatant	0.4	8.1	0.0
Periplasm	83.0	91.9	4.4
Cytoplasm-membranes	16.6	0.0	95.6
pα1.hyb4			
Culture supernatant	0.5	3.6	0.0
Periplasm	96.5	96.4	5.4
Cytoplasm-membranes	3.0	0.0	94.6
pα1.hyb6			
Culture supernatant	0.3	0.0	0.0
Periplasm	89.2	96.3	3.7
Cytoplasm-membranes	10.5	3.7	96.3
pα1.hyb.9			
Culture supernatant	0.3	0.0	0.0
Periplasm	91.3	99.5	3.5
Cytoplasm-membranes	8.4	0.5	96.5

[a] Cells were removed from culture supernatants by centrifugation and separated into periplasmic and cytoplasm-membrane fractions as described in Materials and Methods. Activities are given as percentages of whole culture extracts. β-Lactamase and β-galactosidase were measured as previously described (6). Total β-lactamase activities for the various transformants were 0.8 to 1.2 μmol of 7-(thienyl-2-acetamido)-3[2(4-*N*,*N*-dimethylaminophenylazo)pyridinium methyl]-3-cephem-4-carboxylic acid hydrolyzed per min per mg of protein at 37°C (pH7.0). Total β-galactosidase activities were 55 to 95 μmol of *o*-nitrophenol generated per min per mg of protein at 28°C (pH 7.0). The Phaedebas colorimetric assay (Pharmacia) was used to measure activity. Total amylase activities for the various transformant cultures were more variable at 3.4×10^3 to 7.2×10^4 U/mg of protein. These differences were generally reflective of the different specific activities of the various amylases.

with linear DNA which recircularized in vivo by homologous recombination (24) between the *amyS* and *amyL* sequences.

Mapping of *amy* crossovers. We subjected plasmid DNA extracted from several amylase-positive transformants to restriction enzyme mapping to verify that single crossovers had occurred. It was found that the *amy* genes contained the restriction sites expected for the 5′ *amyS* and 3′ *amyL* sequences (data not shown) and therefore occurred as a result of single crossovers. To study the crossovers in more detail, we developed a fine-structure mapping and sequencing strategy as follows. Nine regions of 16 to 22 nucleotides from the *amyS* sense strand at intervals of around 150 bp, which showed low homology to the *amyL* gene, were selected (indicated by boxes on top lines in Fig. 3). Corresponding oligonucleotides were synthesized and radioactively labeled (probes I to IX). Colony hybridization with a particular probe was expected to indicate the presence of the corresponding *amyS* sequence in that region, whereas nonhybridization would indicate the presence of *amyL* sequences. Using this strategy, we mapped the *amy* crossovers in the plasmids of 96 amylase-positive colonies. Crossovers were found in all intervals except between probe regions II and III (II-III crossover). Therefore we examined another 192 colonies using these probes to search further for II-III crossovers. One was found. The results (further detailed in Rey et al., in press) show that the crossovers were well distributed over the *amy* homologous region in pα1 but that II-III and VII-VIII crossovers were relatively rare.

DNA sequences of *amy* crossover loci. To precisely locate the positions of individual *amy* crossovers and to study the distribution of crossovers within a particular probe interval, we directly sequenced 18 recombined pα1 plasmids, pα1.hyb1 through pα1.hyb18, by the supercoil method. The most 3′ probe which hybridized was used to prime synthesis. The crossover loci thus determined are indicated by boxed regions on the lower lines of Fig. 3.

The results for the most 3′ (VIII-IX) and most 5′ (I-II) crossovers were particularly striking. The crossover points in the four plasmids with VIII-IX crossovers, pα1.hyb15 to pα1.hyb18, all mapped to the same CGCT tetranucleotide. Interestingly, this tetranucleotide represents the most 3′ region of *amy* homology on the pα1 plasmid. It is possible that these four plasmids are siblings derived from the same crossover event. However, sequencing of a single VIII-IX crossover generated in an independent experiment yielded an identical result. An apparent bias for extreme 5′ crossovers was also observed. Three of four I-II crossovers mapped to the nucleotide region represented by the stippled box of Fig. 3, which is just 5′ of the extreme 5′ region of *amy* homology of the pα1 plasmid. Their sequences were slightly different, indicating that they were generated by independent crossover events. Most of the other crossovers mapped to relatively long stretches of perfect homology (8 to 23 bp), although crossovers occurred with some frequency in shorter regions of homology, e.g., crossing over in pα1.hyb11 occurred at a tetranucleotide. To determine whether any crossovers resulted in the formation of hybrid genes encoding inactive amylases, we mapped 96 amylase-negative colonies from the *Pst*I enrichment experiment. All contained unrecombined pα1 plasmids.

Characteristics of *amy* hybrid gene products. The *amy* gene products encoded by pα1.hyb4, pα1.hyb6, and pα1.hyb9 *amy* genes are predicted to contain the *amyS* signal peptide and NH_2-terminal *amyS* mature polypeptide regions of 15, 58, and 163 residues, respectively (see the boxed regions on the lower lines of Fig. 4), the remaining residues being

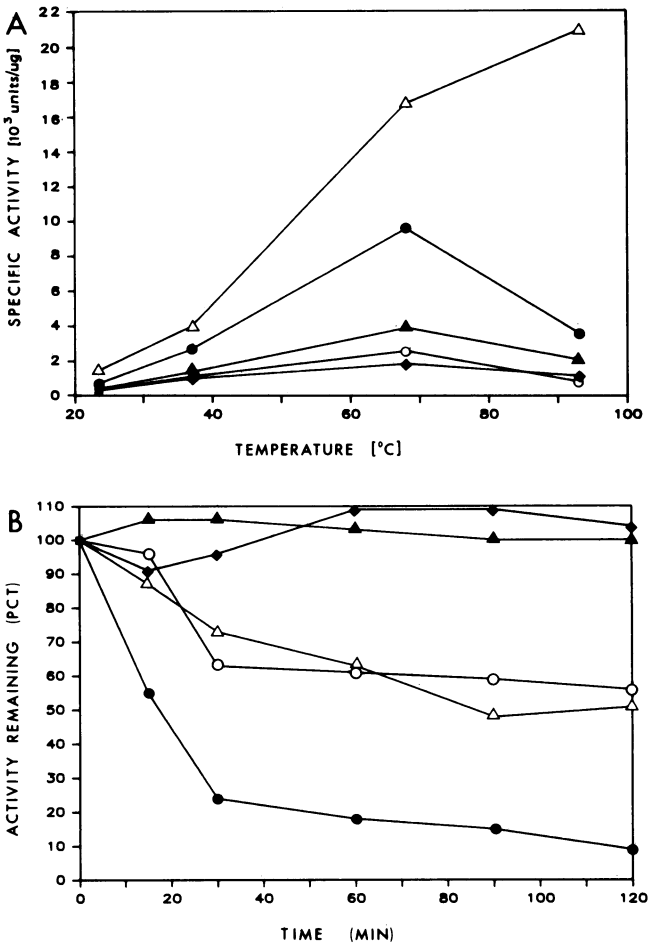


FIG. 5. Comparison of purified hybrid amylases. (A) The specific activities at elevated temperatures. (B) The residual activities after incubation at 90°C. Symbols: ▲, *E. coli* 294(pBR322BL); △, *E. coli* 294(pUC13BS); ◆, *E. coli* 294(pα1.hyb4); ○, *E. coli* 294(pα1.hyb6); ●, *E. coli* 294(pα1.hyb9). This figure has been published elsewhere (Rey et al., in press; published here by permission from Academic Press, Inc.).

derived from the *amyL* gene. It was therefore of interest to determine whether these were secreted and processed normally. Cell fractionation experiments (Table 1) showed that each of the hybrid amylases, like the parental enzymes, were secreted into the periplasmic space of *E. coli*. In addition, the purified pα1.hyb4, pα1.hyb6, and pα1.hyb9 proteins have mobilities similar to those of the parental enzymes on sodium dodecyl sulfate-polyacrylamide gels (Rey et al., in press), suggesting that they may also be properly processed from their signal peptide-containing precursors. Of course, NH_2-terminal amino acid sequencing is needed to confirm this.

Altered intramolecular contacts within the various hybrid amylases might result in novel temperature sensitivities. Thus, the specific activities of the purified pα1.hyb4-, pα1.hyb6-, and pα1.hyb9-derived *amy* gene products as a function of temperature were determined and compared with those of the *E. coli*-derived parental enzymes. The *amyS* gene product had a substantially higher specific activity than the *amyL* gene product at all temperatures tested (Fig. 5A). The pα1.hyb4 gene product, which differs from the *amyL* gene product only at its 15 NH_2-terminal residues, had an

activity profile similar to that of the *amyL* gene product. The pα1.hyb6 and pα1.hyb9 gene products, which contain increasingly longer *amyS*-derived NH_2-terminal regions (58 and 163 residues, respectively), were more *amyS*-like in that they had increasingly higher specific activities than the *amyL* gene product. On the basis of this small sampling of *amy* hybrid gene products, it appears that the very high specific activity of the *amyS* protein is contributed by continuous regions of its primary structure.

The thermostabilities of the three purified hybrid gene products were also compared with those of the parental enzymes (Fig. 5B). The *amyL* product retained all of its activity after heating at 90°C for 120 min, whereas the *amyS* gene lost about half of its activity. The pα1.hyb4 product, which differs only slightly in primary structure from that of the *amyL* parent, had a similar high thermostability. However, the pα1.hyb6 and pα1.hyb9 products, which contain longer *amyS*-derived NH_2-terminal sequences (58 and 163 residues, respectively), had lower residual activities (65 and 10%, respectively). Thus, based on these limited data, it appears that increases in the length of the NH_2-terminal regions contributed by the *amyS* gene product decrease hybrid enzyme stability with respect to the *amyL* enzyme, whereas these same increases enhance specific activity relative to *amyL*.

DISCUSSION

The amylases of *B. licheniformis* (*amyL*) and *B. stearothermophilus* (*amyS*) are related as indicated by homology at the DNA and protein levels. They belong to an enzyme family with members which also include the amylases of *B. coagulans* and *B. amyloliquefaciens* (*amyA*) but not that of *B. subtilis*. It is intriguing that the homology relationships among the alkaline proteases of *B. subtilis*, *B. amyloliquefaciens*, and *B. licheniformis* contrast strongly with those of the amylases. The *amyL* and *amyA* amylases are highly homologous to each other but unrelated to the *B. subtilis* amylase, whereas the alkaline proteases of *B. subtilis* (20) and *B. amyloliquefaciens* (25) are extremely similar to each other and also show low but obvious homology to that of *B. licheniformis* (J. Wells, personal communication).

The method described here for the generation and sequence analysis of hybrid amylase genes should be applicable to any DNA sequences which have homology. Indeed, we have applied a modification of it to the *Bacillus* alkaline protease gene family (manuscript in preparation). In this case, the precursor plasmid contains a polylinker with three unique restriction sites separating the homologous genes. Predigestion with three enzymes allowed reduction of the background of unrecombined plasmids to an insignificant level (2%). A low background is expected to be of particular importance when very low recombination frequencies due to low homology are anticipated. The method may also be used to generate multiple crossovers by construction of a precursor plasmid in which one of the homology regions is derived from a previous crossover.

Using a somewhat different approach, Weber and Weissmann (24) have generated hybrid leukocyte interferons by recombination in *E. coli*. The methods they describe for enrichment for particular classes of recombinants should be equally applicable to our approach for amylases.

The distribution of crossover loci that we observed in one recombined plasmid population may be strongly biased by enrichment for hybrid genes generated early in culture growth. This may also account for multiple plasmids containing identical crossover loci. However, the plasmid preparation isolated from a second independent pα1 transformation experiment produced a crossover distribution also characterized by a rarity of II-III and VII-VIII crossovers. In addition, sequencing of one each of I-II and VIII-IX crossovers revealed terminal crossover loci identical to those found for pα1.hyb2 and pα1.hyb15-18, respectively. It is possible that the crossovers observed are from a narrow population which gave rise to amylase-positive colonies, whereas other crossovers occurred but were undetected because they yielded amylase-negative colonies. However, this appears not to be the case, as all of the amylase-negative colonies which arose after the *Pst*I enrichment step were found to contain unrecombined pα1 rather than plasmids with recombined *amy* genes. Finally, it is possible that those crossovers which gave rise to improperly folded proteins were lethal to the cells and thus went undetected. We are investigating this idea by examining the effect of deletion of the amylase promoter on crossover distribution.

The tertiary structures of proteins are in part stabilized by noncovalent interactions between amino acid residues having positions widely separated in the primary structure. It would be expected that in the hybrid amylases some of these bonds might be weakened or strengthened by the presence of nonhomologous residues in some of these positions. This might lead to alterations in various amylase physical properties such as thermostability and temperature optimum. Thus it is not surprising that our data on the specific activities and thermostabilities of three hybrid amylases show dramatic differences with the parental enzymes. We anticipate that further characterization of other hybrid amylases will reveal enzymes which show a wide spectrum of values for these and other properties.

Site-directed mutagenesis in recent years has been the preferred method of changing the catalytic properties of enzymes. This approach has yielded significant results but depends on detailed X-ray crystallographic data to obtain a picture of the three-dimensional structure of a protein. Such data are unavailable for the *Bacillus* amylases and many other enzyme families. Random generation of hybrid genes followed by characterization of the encoded proteins is a promising method for alteration of such enzyme families until information on their tertiary structures becomes available.

ACKNOWLEDGMENTS

We thank Kirk Hayenga and Bill Hensel for determination of the *amyS* NH_2-terminal amino acid sequence. We thank the Genentech Organic Synthesis Group for preparation of synthetic oligonucleotides. For helpful discussions of this work, we thank Herb Heyneker and Eugenio Ferrari. We thank Alane Gray and Leslie Flint for assistance in manuscript preparation.

LITERATURE CITED

1. **Backman, K., M. Ptashne, and W. Gilbert.** 1976. Construction of plasmids carrying the cI gene of bacteriophage λ. Proc. Natl. Acad. Sci. USA **73**:4174–4178.
2. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
3. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
4. **Cornelis, P., C. Digneffe, and K. Willemot.** 1982. Cloning and expression of the *Bacillus coagulans* amylase gene in *Escherichia coli*. Mol. Gen. Genet. **186**:507–511.
5. **Crea, R., A. Kraszewski, T. Hirose, and K. Itakura.** 1978. Chemical synthesis of genes for human insulin. Proc. Natl.

Acad. Sci. USA **75**:5765–5769.
6. **Gray, G. L., D. H. Smith, J. S. Baldridge, R. N. Harkins, M. L. Vasil, E. Y. Chen, and H. L. Heyneker.** 1984. Cloning, nucleotide sequence, and expression in *Escherichia coli* of the exotoxin A structural gene of *Pseudomonas aeruginosa*. Proc. Natl. Acad. Sci. USA **81**:2645–2649.
7. **Ingle, M. B., and R. J. Erickson.** 1978. Bacterial α-amylases. Adv. Appl. Microbiol. **24**:257–278.
8. **Karn, J., S. Brenner, L. Barnett, and G. Cesareni.** 1980. Novel bacteriophage cloning vector. Proc. Natl. Acad. Sci. USA **77**:5172–5176.
9. **Koshland, D., and D. Botstein.** 1980. Secretion of beta-lactamase requires the carboxy end of the protein. Cell **20**:749–760.
10. **Kuhn, H., P. P. Fietzek, and J. O. Lampen.** 1982. N-terminal amino acid sequence of *Bacillus licheniformis* α-amylase: comparison with *Bacillus amyloliquefaciens* and *Bacillus subtilis* enzymes. J. Bacteriol. **149**:372–373.
11. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
12. **Lederberg, E. M., and S. N. Cohen.** 1974. Transformation of *Salmonella typhimurium* by plasmid deoxyribonucleic acid. J. Bacteriol. **119**:1072–1074.
13. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Meers, J. L., and L. K. Peterson.** 1972. Nitrogen assimilation by *Bacillus licheniformis* organisms growing in chemostat cultures. J. Gen. Microbiol. **70**:277–286.
15. **Mielenz, J. R.** 1983. *Bacillus stearothermophilus* contains a plasmid-borne gene for α-amylase. Proc. Natl. Acad. Sci. USA **80**:5975–5979.
16. **Nakajima, R., T. Imanaka, and S. Aiba.** 1985. Nucleotide sequence of the *Bacillus stearothermophilus* α-amylase gene. J. Bacteriol. **163**:401–406.
17. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
18. **Saito, N.** 1973. A thermophilic extracellular α-amylase from *Bacillus licheniformis*. Anal. Biochem. Biophys. **155**:290–298.
19. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Stahl, M. L., and E. Ferrari.** 1984. Replacement of the *Bacillus subtilis* subtilisin structural gene with an in vitro-derived deletion mutation. J. Bacteriol. **158**:411–418.
21. **Stephens, M. A., S. A. Ortlepp, J. F. Ollington, and D. J. McConnell.** 1984. Nucleotide sequence of the 5′ region of the *Bacillus licheniformis* α-amylase gene: comparison with the *B. amyloliquefaciens* gene. J. Bacteriol. **158**:369–372.
22. **Takkinen, K., R. F. Pettersson, N. Kalkkinen, I. Palva, H. Soderlund, and L. Kaariainen.** 1983. Amino acid sequence of α-amylase from *B. amyloliquefaciens* deduced from the nucleotide sequence of the cloned gene. J. Biol. Chem. **258**:1007–1013.
23. **Veira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259–268.
24. **Weber, H., and C. Weissmann.** 1983. Formation of genes coding for hybrid proteins by recombination between related cloned genes in *E. coli*. Nucleic Acids Res. **11**:5661–5669.
25. **Wells, J. A., E. Ferrari, D. J. Henner, D. A. Estell, and E. Y. Chen.** 1983. Cloning, sequencing, and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*. Nucleic Acids Res. **11**:7911–7925.
26. **Yang, M., A. Gelizzi, and D. Henner.** 1983. Nucleotide sequence of the amylase gene from *Bacillus subtilis*. Nucleic Acids Res. **11**:237–249.