APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 1992, p. 4016–4025
0099-2240/92/124016-10$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 58, No. 12

Cyclization Characteristics of Cyclodextrin Glucanotransferase Are Conferred by the NH_2-Terminal Region of the Enzyme

SHINSUKE FUJIWARA,[1] HIROFUMI KAKIHARA,[1] KIM BYUNG WOO,[1†] ANDRE LEJEUNE,[1‡] MITSUHIDE KANEMOTO,[1] KENJI SAKAGUCHI,[2] AND TADAYUKI IMANAKA[1*]

Department of Biotechnology, Faculty of Engineering, Osaka University, Yamadaoka, Suita, Osaka 565,[1] and Nihon Shokuhin Kako Company, Ltd., Marunouchi, Chiyoda-ku, Tokyo 100,[2] Japan

Received 9 March 1992/Accepted 29 September 1992

Cyclodextrin glucanotransferase (CGTase; EC 2.4.1.19) is produced mainly by *Bacillus* strains. CGTase from *Bacillus macerans* IFO3490 produces α-cyclodextrin as the major hydrolysis product from starch, whereas thermostable CGTase from *Bacillus stearothermophilus* NO2 produces α- and β-cyclodextrins. To analyze the cyclization characteristics of CGTase, we cloned different types of CGTase genes and constructed chimeric genes. CGTase genes from these two strains were cloned in *Bacillus subtilis* NA-1 by using pTB523 as a vector plasmid, and their nucleotide sequences were determined. Three CGTase genes (*cgt-1*, *cgt-5*, and *cgt-232*) were isolated from *B. stearothermophilus* NO2. Nucleotide sequence analysis revealed that the three CGTase genes have different nucleotide sequences encoding the same amino acid sequence. Base substitutions were found at the third letter of five codons among the three genes. Each open reading frame was composed of 2,133 bases, encoding 711 amino acids containing 31 amino acids as a signal sequence. The molecular weight of the mature enzyme was estimated to be 75,374. The CGTase gene (*cgtM*) of *B. macerans* IFO3490 was composed of 2,142 bases, encoding 714 amino acids containing 27 residues as a signal sequence. The molecular weight of the mature enzyme was estimated to be 74,008. The sequence determined in this work was quite different from that reported previously by other workers. From data on the three-dimensional structure of a CGTase, seven kinds of chimeric CGTase genes were constructed by using *cgt-1* from *B. stearothermophilus* NO2 and *cgtM* from *B. macerans* IFO3490. We examined the characteristics of these chimeric enzymes on cyclodextrin production and thermostability. It was found that the cyclization reaction was conferred by the NH_2-terminal region of CGTase and that the thermostability of some chimeric enzymes was lower than that of the parental CGTases.

Cyclodextrins (CDs) are closed-ring structures in which six or more glucose units are joined by means of α-1,4 glucosidic bonds (39). According to the number of glucose units (G6, G7, or G8), they are named α-, β-, and γ-CDs, respectively. They are able to form inclusion complexes with many organic and inorganic molecules, thereby changing the physical and chemical properties of the included compounds. Cyclodextrin glucanotransferase (CGTase; EC 2.4.1.19), which catalyzes the degradation of starch into cyclodextrin, is produced mainly by *Bacillus* strains (2, 7, 13, 15, 16, 18, 27, 35, 37, 42, 49). CGTase is an important enzyme for the food and pharmaceutical industries. It is well known that there are at least four highly conserved amino acid sequences among many amylolytic enzymes, including CGTase and various amylases (22, 26, 34). α-Amylase and CGTase catalyze the cleavage of α-1,4 glucosidic bonds in amylose and amylopectin. It is therefore assumed that common features exist at their active sites.

Most α-amylases are composed of approximately 500 amino acids, and the majority of CGTases are composed of approximately 700 amino acids. The COOH-terminal region of CGTase contains an extra 200 amino acids in addition to the NH_2-terminal region of α-amylase. Recently, the tertiary structure of CGTase was determined (21) and found to comprise four globular domains, A, B, C, and D. Domains A and B were similar to the structure of Taka-amylase (21, 30), which is one of the α-amylases. Domains C and D were unique to the CGTase. The COOH-terminal region has homology with the raw-starch-binding site of glucoamylase, and it is assumed that this region of CGTase is related to the cyclization of linear maltooligosaccharide produced by the hydrolysis of starch with the enzyme (16, 46, 47).

Different types of CD are produced from starch by different *Bacillus* CGTases. CGTases are classified into two types. One is the *Bacillus macerans* type of enzyme, which produces mainly α-CD from starch (2, 17). The other is β-CD-forming CGTase, which is produced by *Bacillus megaterium*, *Bacillus circulans*, and alkalophilic bacilli (13, 15, 17, 27, 35). *Bacillus stearothermophilus* CGTase, which produces α-CD and β-CD as its main products, is considered intermediate between these two types (17).

In this study, we constructed and characterized chimeric enzymes to identify which regions of the proteins are important for cyclization activity.

MATERIALS AND METHODS

Bacterial strains, plasmids, and phages. *B. stearothermophilus* NO2, a producer of CGTase, was isolated from soil. *B. macerans* IFO3490 was provided from the Institute for

* Corresponding author.
† Present address: Department of Microbiology, Dongeui University, 24 Gaya Dong, Pusanjin-ku 614-715, Pusan, Korea.
‡ Present address: Laboratory of Plant Cytology, Faculty of Sciences, Catholic University of Louvain, B-1343 Louvain-la-Neuve, Belgium.

Fermentation, Osaka, Japan. *B. subtilis* NA-1 (*arg-15 hsmM hsrM* Amy$^-$ Npr$^-$) (23) was used as the host strain for plasmids. *Escherichia coli* MV1184 [*ara* Δ(*lac-proAB*) *rpsL thi* (φ80 *lacZ*Δ*M15*) (*srl-recA*)*306*::Tn*10* Tet^r F′ (*traD36 proA$^+$B$^+$ lacI^q lacZ*Δ*M15*)] (51) was used as the host strain for plasmid pBR322 (3) and phagemids pUC118 and pUC119. For phages M13mp18 and M13mp19 (31), *E. coli* TG1 [*supE hsd*Δ*5 thi* Δ(*lac-proAB*) F′ (*traD36 proA$^+$B$^+$ lacI^q lacZ*Δ*M15*)] (5) was used as the host strain. Plasmid pTB523 was constructed from pTB522 (tetracycline resistant [Tc^r]) (11) by digestion with *Eco*RI, Klenow polymerase treatment, and ligation with T4 DNA ligase. The ligated DNA was digested with *Hin*dIII, and *Eco*RI linkers were inserted. The *Eco*RI site of pTB522 was not suitable for gene cloning. Therefore, we removed the *Eco*RI site and introduced new a *Eco*RI site at the *Hin*dIII position of pTB522 by linker modification. The constructed plasmid was named pTB523 and has a unique *Eco*RI site for gene cloning.

Media. CS1 medium containing 1% soluble starch, 0.5% polypeptone, 0.5% yeast extract, 0.05% K_2HPO_4, and 0.01% $MgSO_4 \cdot 7H_2O$ was used for *B. subtilis* and *E. coli*. For agar plates (CS1 azure plates), soluble starch was replaced by starch azure (0.25%), and 1.5% (wt/vol) agar was added. Starch azure (Sigma Chemical Co., St. Louis, Mo.) was used as the substrate for the colorimetric assay for α-amylase. Potato starch is covalently linked with Remazol brilliant blue R. Amylase-producing bacteria can form clear halos around colonies on CS1 azure plates.

Cloning of the CGTase genes from *B. stearothermophilus* NO2 and *B. macerans* IFO3490. The chromosomal DNA (about 6 μg) of *B. stearothermophilus* NO2 was purified and digested with restriction enzyme *Eco*RI. DNA fragments of 2.5 to 10.0 kb were isolated and inserted into the unique *Eco*RI site of pTB523 (about 2 μg). *B. subtilis* NA-1 cells were transformed with these recombinant plasmids and plated onto CS1 azure plates containing tetracycline (15 μg/ml). From about 5.1×10^4 transformants, 40 halo-forming colonies were selected. All candidates produced CGTase. Plasmid DNA was prepared from them and digested by *Eco*RI to check insertion size. On the basis of preliminary restriction mapping, they could be divided into three groups, with insert sizes of 6.2, 7.6, and 5.8 kb, and the recombinant plasmids harboring these fragments were named pKB1, pKB5, and pKB232, respectively. It was suggested that these three fragments might encode different CGTase genes. The three CGTase genes cloned in pKB1, pKB5, and pKB232 were named *cgt-1*, *cgt-5*, and *cgt-232*, respectively. Inserts representative of each group were subcloned at the *Eco*RI site of *E. coli* plasmid pBR322, and precise maps were constructed. The restriction maps of these fragments were similar to that of an insert of pTCH201 harboring the CGTase gene of *B. stearothermophilus* TC-91 (42).

The *B. macerans* IFO3490 CGTase gene, designated *cgtM*, was also cloned into plasmid pTB523 as an 8.0-kb *Eco*RI fragment by the same procedure used to clone the CGTase gene from *B. stearothermophilus* NO2. The cloned fragment was partially digested with *Sau*3AI and subcloned into the *Bam*HI site of pUC118 in *E. coli* MV1184. The constructed plasmid in which a 3.6-kb fragment had been inserted was designated pCGT1. The *Eco*RI-*Pst*I fragment (3.65 kb) from pCGT1, containing the multiple cloning site of pUC118, was recloned into *B. subtilis* plasmid pTB523, and the recombinant plasmid was named pKBM1.

Construction of chimeric genes. Chimeric genes were constructed by polymerase chain reaction (PCR) fusion by the procedure of Yon and Fried (52). The chimeric genes are shown in Fig. 1. The amplification reaction was carried out under the following conditions: 10 ng of DNA template, 20 pmol of primer, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM mixed deoxynucleoside triphosphates, and 2.5 U of *Taq* DNA polymerase were mixed in a total volume of 100 μl. PCR was performed for 30 cycles of 1 min at 94°C, 2 min at 55°C, and 2 min at 72°C. Six different primers were used for PCR to construct pKC2 and pKC3, as follows: primer 1 was the 27-mer outer primer, 5′AAGCAA AACCCGTTTTTTGGTACCCTC3′; primer 2 was the 22-mer outer primer, 5′CCGCACCTGTATTTCACAATCC3′; primer 3 was the linking primer (45-mer) for construct pKC2, 5′CGCCGGCGACGTTTCCGGCGCACTGTATGCCCATA CACCGACTTC3′; primer 4 was the linking primer (45-mer) for construct pKC2, 5′GAAGTCGGTGTATGGGCATACA GTGCGCCGGAAACGTCGCCGGCG3′; primer 5 was the linking primer (48-mer) for construct pKC3, 5′GTCGTCGT CCCATAAGCGATGGCCGGATTGTTTCGTCGGAGAGA AGAT3′; and primer 6 was the linking primer (48-mer) for construct pKC3, 5′ATCTTCTCTCCGACGAAACAATC CGGCCATCGCTTATGGGACGACGAC3′.

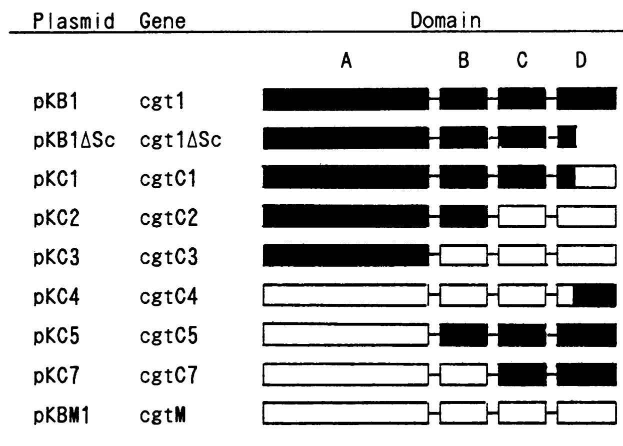

FIG. 1. Structural features of chimeric CGTases. A, B, C, and D indicate the regions of domain structures. Solid boxes and open boxes represent regions from *cgt-1* and *cgtM*, respectively.

Six different primers were used for PCR to construct pKC5 and pKC7, as follows: primer 7 was the 18-mer outer primer, 5′ACCTCTACGACCTGGCGG3′; primer 8 was the 30-mer outer primer containing a *Sal*I linker, 5′TTGCCGAAGTCGAC AGAGAAGCAGCCATAG3′; primer 9 was the linking primer (43-mer) for construct pKC5, 5′GATCCAACGCTGTTCGGT ATCACCATAAGCGATGGCCGGATTG3′; primer 10 was the linking primer (43-mer) for construct pKC5, 5′CAATC CGGCCATCGCTTATGGTGATACCGAACAGCGTTGGAT C3′; primer 11 was the linking primer (37-mer) for construct pKC7, 5′GACCAATAATTGGCGTGCTTTCCGGCGCTGT GTACTG3′; and primer 12 was the linking primer (37-mer) for construct pKC7, 5′CAGTACACAGCGCCGGAAAGCACGC CAATTATTGGTC3′.

Schematic diagrams of pKC2 construction are shown in Fig. 2. The 1.59-kb *Pvu*II fragment of *cgt-1* and the 3.65-kb *Eco*RI-*Pst*I fragment of *cgtM* were used as the template DNA for PCR. Primers 1 and 3 were annealed with the 1.59-kb *Pvu*II fragment, and an intermediate DNA (1,173 bp) was produced by PCR. This amplified DNA had the domain A and B regions from *cgt1*. Primers 2 and 4 were annealed with the 3.65-kb *Eco*RI-*Pst*I fragment, and an intermediate

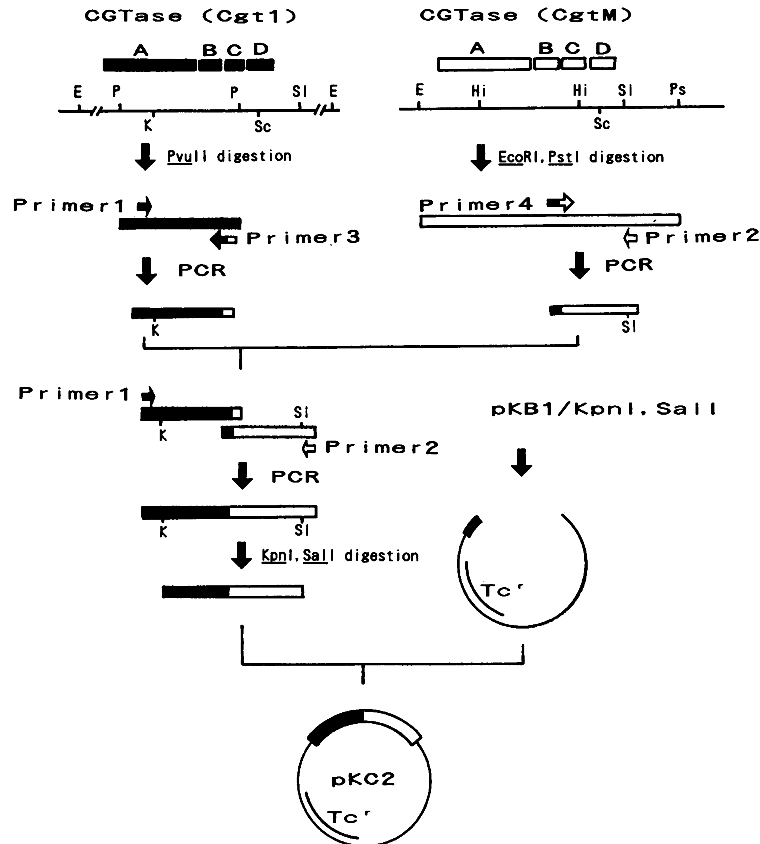

FIG. 2. Schematic diagram of the construction of a chimeric CGTase gene. The solid bar and open bar represent regions derived from pKB1 harboring *cgt-1* and from pKBM1 harboring *cgtM*, respectively. A, B, C, and D indicate domains A, B, C, and D, respectively. DNA amplified with primers 1 and 3 or primers 2 and 4 produced the domain A and B regions from *cgt-1* or the domain C and D regions from *cgtM*, respectively. Abbreviations: E, *Eco*RI; Hi, *Hin*dIII; K, *Kpn*I; P, *Pvu*II; Ps, *Pst*I; Sc, *Sac*I; Sl, *Sal*I.

DNA (716 bp) was produced by PCR. This amplified DNA had the domain C and D regions from *cgtM*. These two intermediates were joined by PCR with two outer primers (primers 1 and 2), and a fused DNA fragment (1,844 bp) was created. The synthetic DNA was digested with *Kpn*I and *Sal*I, and the 1,789-bp fragment was purified and inserted into the *Kpn*I and *Sal*I sites of pKB1 after the original fragment was removed. The resulting plasmid was named pKC2. For pKC3 construction, primers 5 and 6 were used as the linking primers instead of primers 3 and 4. Other processes were the same as for pKC2 construction. In CgtC2 and CgtC3, Ser-520 and Pro-429 of Cgt-1 were joined with Ala-522 (corresponding to Ala-521 of Cgt-1) and Ala-431 (corresponding to Ala-430 of Cgt-1) of CgtM, respectively.

For *cgtC5* construction, a 2.4-kb *Sal*I fragment of *cgtM* and a 3.65-kb *Hpa*I fragment of *cgt-1* were used as the template DNAs for PCR. Primers 7 and 9 were annealed with the 2.4-kb *Sal*I fragment, and an intermediate DNA (665 bp) was produced by PCR. Primers 8 and 10 were annealed with the 3.75-kb *Hpa*I fragment of *cgt-1*, and an intermediate DNA (943 bp) was produced by PCR. These two intermediates were joined by PCR with outer primers 7 and 8, and a fused DNA (1,568 bp) was created. The fused DNA was digested with *Hin*dIII and *Sal*I, and the 1,409-bp fragment was purified and inserted into the *Hin*dIII and *Sal*I sites of pKBM1 after the original region was removed. The resulting plasmid was named pKC5. To construct pKC7, primers 11 and 12 were used instead of primers 9 and 10. Other processes were carried out as for pKC5 construction. In CgtC5 and CgtC7, Tyr-434 and Glu-524 of CgtM were joined with Gly-434 (corresponding to Gly-435 of CgtM) and Ser-524 (corresponding to Thr-525 of CgtM) of Cgt-1, respectively.

For construction of pKC1 and pKC4, the *Sac*I site in the CGTase coding region was used, which was conserved in both *cgt-1* and *cgtM*. To construct pKC1, the *Sac*I-*Sal*I region of pKB1 was replaced with that of pKBM1. Likewise, this region of pKBM1 was replaced with that of pKB1 to construct pKC4. For construction of pKB1ΔSc, pKB1 was digested with *Sac*I, treated with Klenow polymerase to fill in the cohesive ends, and ligated with T4 DNA ligase. By these treatments, the open reading frame was changed and a new termination codon appeared 207 bp upstream of the native termination codon of *cgt-1*. The chimeric CGTase genes from pKC1, pKC2, pKC3, pKC4, pKC5, pKC7, and pKB1ΔSc were designated *cgtC1*, *cgtC2*, *cgtC3*, *cgtC4*, *cgtC5*, *cgtC7*, and *cgt1ΔSc*, respectively.

Purification of CGTase. *B. subtilis* NA-1 cells harboring pKB1, pKB5, pKB232, or pKBM1 were incubated in 500 ml of CS1 medium at 37°C for 17 h. After centrifugation (8,000 × *g*, 10 min) of the culture broth, ammonium sulfate (final concentration, 35% saturation) was added to the supernatant, and the solution was kept at 4°C for 2 h. The precipitate was removed by centrifugation (16,000 × *g*, 20 min). Am-

monium sulfate was again added to the supernatant to a final concentration of 75% saturation, and it was kept at 4°C overnight. After centrifugation (16,000 × g, 20 min), the precipitate was dissolved in 2 ml of 50 mM phosphate buffer (pH 6.0). The enzyme solution was applied to a column (1.0 by 30 cm) of Sephadex G-100 (Pharmacia Fine Chemicals, Uppsala, Sweden). This gel filtration step was repeated twice. Active fractions were concentrated and used for enzyme characterization. For determination of the amino-terminal amino acid sequence, the enzyme solution was applied to an ion-exchange high-performance liquid chromatography (HPLC) column (DEAE-5PW; Tosoh Co., Ltd., Tokyo, Japan) that had been equilibrated with the same buffer. Active enzyme was eluted with a 0 to 0.6 M NaCl gradient.

Assay of CGTase activity. α-CD-specific CGTase activity was assayed by the methyl orange method as described previously (25). The reaction mixture (3,000 μl), containing 600 μl of 5% starch in 50 mM phosphate buffer (pH 6.0), 105 μl of 1 mM methyl orange in 50 mM phosphate buffer (pH 6.0), 2,845 μl of 50 mM phosphate buffer (pH 6.0), and 50 μl of enzyme solution, was incubated for 10 min at the optimum temperature (65°C for Cgt1 and CgtC1; 55°C for CgtM, CgtC2, and CgtC4; and 45°C for CgtC7). To stop the reaction, 150 μl of 6 N HCl was added. The solution was kept at 16°C for 30 min, and the A_{505} was measured.

α-, β-, and γ-CD were identified by HPLC. The reaction mixture (pH 6.0) containing 5% soluble starch (pH 6.0) and 3 mM $CaCl_2$ was incubated with 500 U of the enzyme in a total volume of 4 ml. Samples were withdrawn and kept in a boiling-water bath for 5 min to stop the reaction. The CDs formed were analyzed by HPLC according to the procedure described by Kaneko et al. (14) with an Aminex column (HPX-42A; 300 by 7.8 mm; Bio-Rad) and a refractive index detector. The temperature of the column was kept at 70°C. The mobile phase was water at a flow rate of 0.5 ml/min. The protein concentration was measured by the method of Smith et al. (45) with a bicinchoninic acid protein assay reagent kit (Pierce Co., Rockford, Ill.).

DNA manipulation. Plasmid DNA was prepared by either the rapid alkaline extraction method or CsCl-ethidium bromide equilibrium density gradient centrifugation as described previously (10). Treatment of DNA with restriction enzyme and ligation of DNA with T4 DNA ligase were done as recommended by the manufacturer. For the analysis of DNA, gel electrophoresis with agarose or polyacrylamide was done under standard conditions (28). A GeneClean kit (Bio 101, La Jolla, Calif.) was used to recover DNA from the agarose.

DNA sequencing. Specific restriction fragments were cloned into M13mp18 and M13mp19 or pUC118 and pUC119 and sequenced by the dideoxy chain termination method of Sanger et al. (43). The oligonucleotide primers were synthesized on a Millipore Cyclone DNA synthesizer (MilliGen/Biosearch, Tokyo, Japan) by the β-cyanoethyl phosphoamidite procedure (44).

Transformation. For transformation of *B. subtilis* or *E. coli* with plasmid DNA, competent cells were prepared as described previously (12).

Determination of NH_2-terminal amino acid sequence. The amino-terminal amino acid sequence was analyzed by Edman degradation as described elsewhere (48).

Enzymes and chemicals. Restriction endonucleases and DNA modification enzymes were purchased from Takara Shuzo Co., Kyoto, Japan; Toyobo Co. Ltd., Osaka, Japan; and Bethesda Research Laboratories, Gaithersburg, Md. Tetracycline and starch azure were from Sigma Chemical Co. All other chemicals used in this study were from Wako Pure Chemical Industries, Osaka, Japan.

Other procedures. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was done at a concentration of 7.5% (wt/vol) polyacrylamide as described by Laemmli (24). Amino acid sequence homology was analyzed with an NEC PC-9801 computer (Nippon Electric Co., Tokyo, Japan) and the DNASIS system (Hitachi Co., Tokyo, Japan).

Nucleotide sequence accession numbers. The DNA sequence data have been assigned the following accession numbers: X59042 for *cgt-1*; X59043 for *cgt-5*; X59044 for *cgt-232*; and X59045 for *cgtM*.

RESULTS

Characteristics of CGTases from *B. stearothermophilus* NO2 and *B. macerans* IFO3490. *B. stearothermophilus* NO2 was originally isolated from soil as a thermostable-CGTase-producing bacterium. The CGTase from this strain produced α- and β-CD as the major hydrolysis products from starch. *B. macerans* IFO3490 (NRRL B-388) was known as a CGTase-producing bacterium, and the enzyme produces mainly α-CD (17). To further characterize the enzymes, we purified each enzyme from *B. subtilis* NA-1 cells harboring pKB1 and pKBM1, which carry the *B. stearothermophilus* NO2 CGTase gene (*cgt-1*) and the *B. macerans* IFO3490 CGTase gene (*cgtM*), respectively.

B. subtilis NA-1 harboring pKB1 was grown to the stationary phase (17 h of incubation) in CS1 medium at 37°C with tetracycline (15 μg/ml). The CGTase Cgt-1 was purified from the culture supernatant as described in Materials and Methods. Greater than 1,000-fold purification was attained, with a yield of 4%. With the purified enzyme, the optimum temperature, optimum pH, and CDs produced were analyzed. The optimum temperature of Cgt-1 was 65°C, and the optimum pH was about 6.0. Mainly α- and β-CD were produced from starch. Cgt-5 and Cgt-232 showed the same characteristics as Cgt-1.

The *cgtM* gene was also expressed in *B. subtilis* NA-1 cells, and the enzyme was secreted into the culture medium. CgtM was purified by the procedure described above. Its characteristics were similar to those found in the previous study (17). The optimum temperature and optimum pH were 55°C and 5.7, respectively, and the major hydrolysis product from starch was α-CD.

Nucleotide sequence of CGTase genes from *B. stearothermophilus* NO2. We determined the nucleotide sequences of CGTase genes *cgt-1*, *cgt-5*, and *cgt-232* by the procedure of Sanger et al. (43). The nucleotide sequence and deduced amino acid sequence of *cgt-1* are shown in Fig. 3. Analysis of the sequence showed that there was a single open reading frame of 2,133 bp which encoded a polypeptide of 711 amino acids. At 7 bases upstream from the ATG codon, there was a 9-base sequence, GAAGGAGGA (−17 to −9), which was complementary to the 3′ end of the 16S rRNAs from *B. stearothermophilus* (20) and *B. subtilis* (32). Therefore, it is the most stable ribosome-binding site (Shine-Dalgarno sequence) of *cgt-1*. A putative promoter (−35 and −10 regions) is shown in Fig. 3. It has the typical consensus sequence for a σ^A factor (TTGACA for the −35 region and TATAAT for the −10 region) in *B. subtilis* (32). The distance between the −35 region and −10 region was 17 bases, which is the same as the consensus distance in *B. subtilis* (32). Furthermore, 3 bases downstream from the termination codon at positions

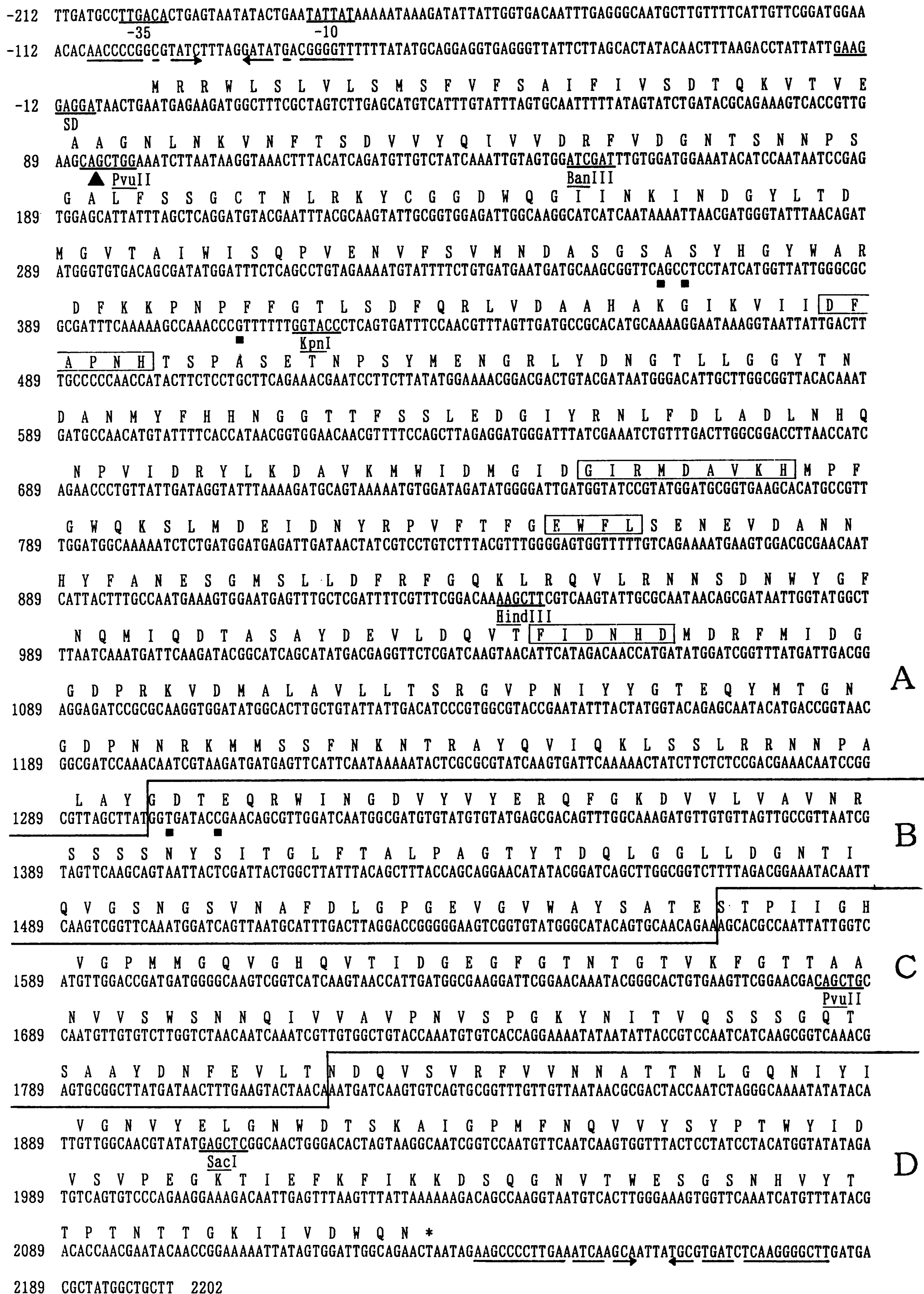

FIG. 3. Nucleotide and amino acid sequences of the *cgt-1* gene. The nucleotide sequence is numbered from the first base of the open reading frame. The amino acid sequence is shown above the nucleotide sequence. A probable Shine-Dalgarno (SD) sequence (nucleotides −17 to −9) and a putative promoter region (−35 and −10 regions) are shown by a solid line below the nucleotide sequence. Inverted repeat structures are indicated by arrows pointing toward each other. The cleavage site between the signal peptide and the extracellular mature enzyme is indicated by a solid triangle. Regions homologous with amylase are boxed. Solid squares indicate positions at which base substitutions were found. A, B, C, and D indicate domains A, B, C, and D, respectively.

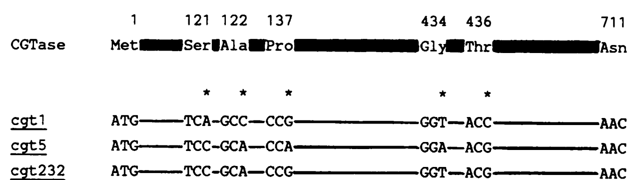

FIG. 4. Comparison of the nucleotide sequences of CGTase genes from *B. stearothermophilus* NO2. Five positions of base substitution are indicated (*). The amino acids are numbered from the N-terminal Met.

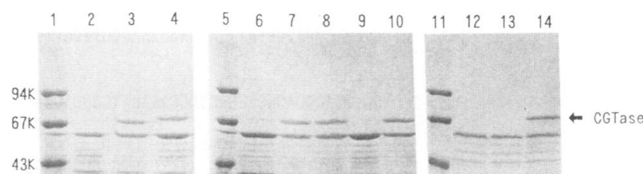

FIG. 5. SDS-PAGE pattern of extracellular fractions from a *B. subtilis* culture. *B. subtilis* NA-1 carrying chimeric CGTase genes was grown aerobically in 20 ml of CS1 medium at 37°C for 17 h. The culture supernatant was used as the extracellular fraction. A 300-μl amount of each sample was concentrated by adding acetone and applied to SDS–7.5% PAGE gels. Proteins were detected by Coomassie brilliant blue staining. Phosphorylase *b* (94,000), albumin (67,000), and ovalbumin (43,000) were used as molecular weight standards. Lanes 1, 5, and 11, molecular weight markers; lanes 2 and 12, pTB523 (control); lane 3, pKBM1; lane 4, pKB1; lane 6, pKB1ΔSc; lane 7, pKC1; lane 8, pKC4; lane 9, pKC3; lane 10, pKC2; lane 13, pKC5; lane 14, pKC7.

+2140 to +2183 was a palindromic sequence that might act as a transcriptional terminator.

The deduced amino acid sequence of Cgt-1 contains four conserved regions previously demonstrated in amylolytic enzymes. Conserved regions are indicated as boxes in Fig. 3. The amino acid sequence of Cgt-1 showed high homology (99%) with that of *B. stearothermophilus* TC-91 CGTase (42). Only two different amino acid residues were found: the Asp-108 and Arg-460 of *B. stearothermophilus* TC-91 CGTase were replaced by Val-108 and Ala-460 in Cgt-1, respectively. The CGTases from *B. stearothermophilus* TC-91 and NO2 showed the same profile of CD production and thermostability.

The nucleotide sequences of *cgt-5* and *cgt-232* were also determined. Although the three genes (*cgt-1*, *cgt-5*, and *cgt-232*) had different nucleotide sequences, they encoded the same amino acid sequence (Fig. 4). Base substitutions which did not affect the amino acid sequence were found at five different positions in these CGTase genes. All base substitutions were located at the third-letter positions of codons 121, 122, 137, 434, and 436. It is interesting that the same enzyme is encoded by three different nucleotide sequences. Sometimes, the bacterial chromosome can be rearranged by recombination (6, 9, 41). The chromosomal DNA of this strain might have been rearranged in the course of evolution. Base substitutions in the three CGTase genes were found at the same positions in the open reading frame, and they were all observed at the third letter of codons. By wobble base pairing, many of the alternative codons for an amino acid differ only in their third positions (4, 33). Our observation is in reasonable agreement with previous studies.

The five amino-terminal amino acids of the three CGTases were determined, and all the sequences were Ala-Gly-Asn-Leu-Asn. This amino acid sequence was identical to that deduced from the nucleotide sequence starting at position 94 in Fig. 3. Therefore, the 31 NH_2-terminal amino acid residues may be a signal sequence which is removed during the secretion process. The molecular weight of mature CGTase, calculated as 75,374 from the amino acid sequence (680 amino acid residues), is in reasonable agreement with the 74,000 assessed directly by SDS-PAGE (Fig. 5).

Nucleotide sequence of the CGTase gene from *B. macerans* IFO3490. The nucleotide sequence of *cgtM* was determined (Fig. 6). *cgtM* was composed of 2,142 bases, encoding 714 amino acids. However, the sequence was different from that reported previously for *B. macerans* IAM1243 (49), which corresponds to *B. macerans* IFO3490. Both *B. macerans* strains are the same as *B. macerans* NRRL B-388. Three points of insertion, three points of deletion, and four points of base substitution were found in the open reading frame, as shown in Fig. 6. Conserved region II, mentioned in the previous report (49), was not found. However, the newly determined sequence has conserved region II, which is Gly-Ile-Arg-Phe-Asp-Ala-Val-Lys-His. By determination of the amino-terminal amino acid sequence of the enzyme, it was found that 27 amino acids were used as a signal sequence, and the molecular weight of the mature enzyme was estimated to be 74,008.

Construction of chimeric genes from two different CGTase genes. Although the amino acid sequences of Cgt-1 and CgtM showed high homology (77%) under considerations of amino acid equivalence (36), the enzyme characteristics were different. CgtM produces mainly α-CD, and the optimum temperature is 55°C. On the other hand, the optimum temperature of Cgt-1 is 65°C, and it produces mainly α- and β-CD. To investigate which region of the protein is important for the specificity of the cyclization activity, we constructed seven kinds of chimeric CGTase genes from *cgt-1* and *cgtM*, as shown in Fig. 1. The tertiary structures of CGTase from *B. stearothermophilus* TC-91 and *B. circulans* no. 8 have been reported (19, 21). The amino acid sequence deduced from *cgt-1* exhibits 99% homology to that of the protein from *B. stearothermophilus* TC-91. Therefore, the molecular structure of Cgt-1 is thought to be similar to that of the *B. stearothermophilus* TC-91 enzyme, which is 80 by 70 by 50 Å (8 by 7 by 5 nm) in size and has four globular domains, A, B, C, and D, which consist of 403, 90, 82, and 105 amino acid residues, respectively. Domains A and B are similar to the structure of Taka-amylase (21, 30). Domains C and D are unique to CGTase.

From the structural analysis data, we tried to construct chimeric CGTases. Yon and Fried (52) reported a simple method for fusion gene construction by PCR between two genes at any chosen location. By this procedure with slight modifications, we constructed recombinant plasmids pKC2, pKC3, pKC5, and pKC7. Plasmids pKC1, pKC4, and pKB1ΔSc were constructed with restriction sites as described in Materials and Methods. The chimeric CGTase genes from pKC1, pKC2, pKC3, pKC4, pKC5, pKC7, and pKB1ΔSc were designated *cgtC1*, *cgtC2*, *cgtC3*, *cgtC4*, *cgtC5*, *cgtC7*, and *cgt-1ΔSc*, respectively. The structural features of the chimeric CGTases are shown in Fig. 1.

In CgtC2 and CgtC3, the Ser-520 and Pro-429 of Cgt-1 were joined with the Ala-522 (corresponding to Ala-521 of Cgt-1) and Ala-431 (corresponding to Ala-430 of Cgt-1) of CgtM, respectively. In CgtC5 and CgtC7, the Tyr-434 and Glu-524 of CgtM were joined with the Gly-434 (correspond-

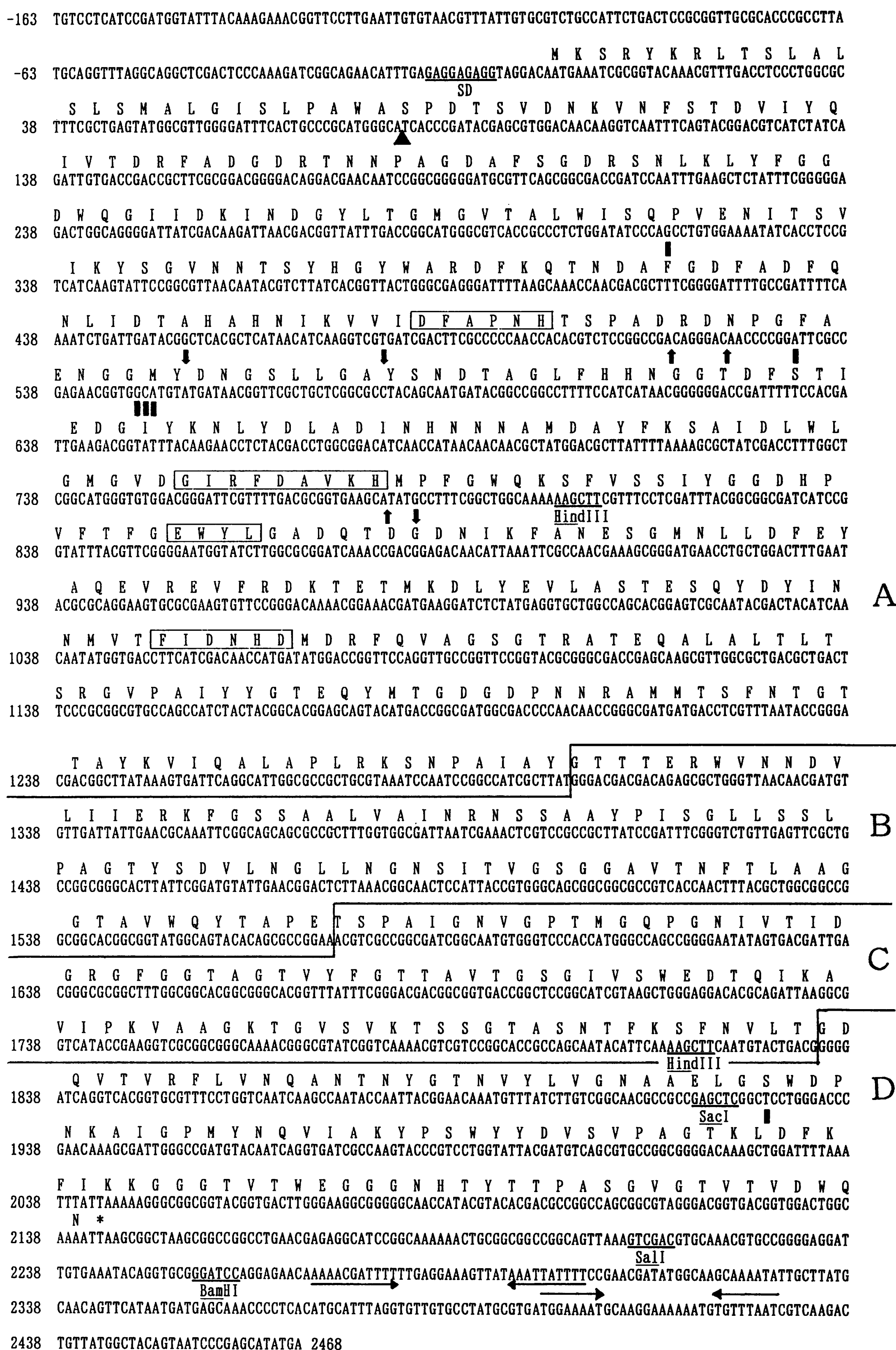

FIG. 6. Nucleotide and amino acid sequences of the *cgtM* gene. Symbols are the same as in Fig. 3. Bases that differ from those reported previously for the *B. macerans* IAM1243 CGTase gene are indicated as ↓ (insertion), ↑ (deletion), and ■ (replacement).

TABLE 1. Characteristics of chimeric CGTases

Plasmid	Gene	Enzyme secretion[a]	Optimum temp (°C)
pKB1	*cgt-1*	+	65
pKB1ΔSc	*cgt-1ΔSc*	−	ND[b]
pKC1	*cgtC1*	+	65
pKC2	*cgtC2*	+	55
pKC3	*cgtC3*	−	ND
pKC4	*cgtC4*	+	55
pKC5	*cgtC5*	−	ND
pKC7	*cgtC7*	+	45
pKBM1	*cgtM*	+	55

[a] +, secreted; −, not secreted.
[b] ND, not determined.

ing to Gly-435 of CgtM) and Ser-524 (corresponding to Thr-525 of CgtM) of Cgt-1, respectively. CgtC1 and CgtC4 were designed by exchanging the *Sac*I-*Sal*I region between *cgt-1* and *cgtM*. In *cgt-1ΔSc*, the open reading frame of *cgt-1* was changed, and a new termination codon appeared 207 bp upstream of the native termination codon of *cgt-1*.

Secretion of chimeric CGTases. It is known that protein conformation affects secretion (40). Chimeric CGTases may have different conformations, and some may not be secreted. To examine the secretion of chimeric CGTases, we attempted to isolate the enzyme fraction from the culture supernatant. The protein fraction was concentrated by acetone, and the protein was analyzed by SDS-PAGE. As shown in Fig. 5, wild-type CGTases encoded by *cgt-1* and *cgtM* were secreted into the culture medium. Chimeric CGTases CgtC1, CgtC2, CgtC4, and CgtC7 were also secreted. However, CgtC3, CgtC5, and Cgt-1ΔSc were not. This might be due to a conformational change, or these chimeric CGTases might be broken by proteolytic attack. The reason for nonsecretion remains unknown at present.

Optimum temperature and pH of chimeric CGTases. The secreted CGTases were purified and used for enzyme analysis. Their optimum temperatures were determined, and the results are summarized in Table 1. CgtC1, CgtC2, CgtC4, and CgtC7 reacted most efficiently at 65, 55, 55, and 45°C, respectively. CgtC2 and CgtC7 were more thermolabile than Cgt-1 and CgtM, respectively. This might be due to the conformational instability caused by domain exchange. A well-organized structure might be important for CGTase enzyme stability. The optimum pH for the enzyme reaction was about 6.0.

Cyclization characteristics of chimeric CGTases. The CD profiles of the chimeric CGTases are shown in Fig. 7. The CDs produced were analyzed at both the initial step (10 min) and the stationary step (60 min) of the reaction. CgtC1 and CgtC2 are α- and β-CD producers, like Cgt-1. CgtC4 and CgtC7 are α-CD producers, like CgtM. These results suggest that the CD production profile depends on domains A and B of the CGTase.

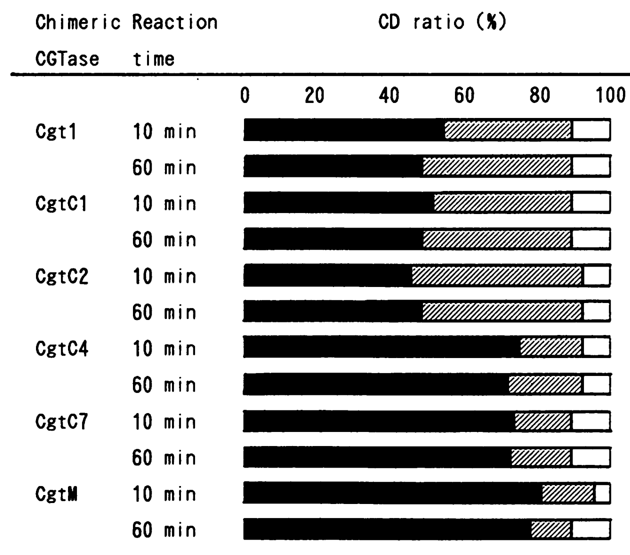

FIG. 7. Ratios of α-, β-, and γ-CD formation by chimeric CGTases. The reaction was carried out at the optimum temperature for 10 and 60 min. CD compositions were analyzed by HPLC. Experimental details are described in the text. Solid boxes, hatched boxes, and open boxes indicate α-CD, β-CD, and γ-CD, respectively.

DISCUSSION

CGTase is composed of four globular domains. Domains A and B are similar to the structure of α-amylase, and domains C and D are unique to the CGTase. Therefore, it was supposed that the COOH-terminal region of CGTase may be related to the cyclization of maltooligosaccharide produced by the hydrolysis reaction, because this region is characteristic of CGTases (16, 46, 47). However, the NH_2 domain of CGTase was considered more important for cyclization in our study.

We constructed seven chimeric CGTase genes from the newly determined nucleotide sequence. The chimeric enzymes were classified into two groups by the CDs produced. One group produces α- and β-CD, and the other group produces α-CD. α- and β-CD production by chimeric enzymes CgtC1 and CgtC2, which contained NH_2-terminal domains derived from *cgt-1*, was almost exactly equal to that by Cgt-1. CgtC4 and CgtC7, which contained NH_2-terminal domains derived from *cgtM*, produced more α-CD, like CgtM. Hence, the NH_2-terminal domain, probably domain A, is important for the cyclization function. The substrate (spiral amylose) is considered to be fixed in domains A and D of the CGTase (21). Therefore, amino acid residues surrounding the spiral amylose in domain A are supposed to affect cyclization. Substrate binding at domain D may be related to enzyme stability.

The thermostability of an enzyme is influenced by many factors (1, 8, 29, 38, 50). Disulfide linkages have been proposed as features that stabilize proteins (29, 50). The *B. stearothermophilus* CGTase has two Cys residues at positions 71 and 78 which were not found in other CGTases. It is unknown whether these Cys residues have an effect on thermostability. More-precise studies are needed to analyze the role of the NH_2- and COOH-terminal regions of CGTase.

ACKNOWLEDGMENTS

We are grateful to N. Nakamura and T. Kaneko for technical support in HPLC analysis. We also thank T. Kuriki for advice.

REFERENCES

1. **Argos, P., M. G. Rossmann, U. M. Grau, H. Zuber, G. Frank, and J. D. Tratschin.** 1979. Thermal stability and protein structure. Biochemistry **18:**5698–5703.
2. **Binder, F., O. Huber, and A. Boeck.** 1986. Cyclodextrin glycosyltransferase from *Klebsiella pneumoniae* M5a1: cloning, nu-

cleotide sequence and expression. Gene **47**:269–277.
3. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95.
4. **Crick, F. H. C.** 1966. The genetic code: III. Sci. Am. **215**:55–62.
5. **Gibson, T. J.** 1984. Studies on the Epstein-Barr virus genome. Ph.D. thesis, Cambridge University, England.
6. **Hashiguchi, K., A. Tanimoto, S. Nomura, K. Yamane, K. Yoda, S. Harada, M. Mori, T. Furusato, A. Takatsuki, M. Yamasaki, and G. Tamura.** 1985. Gene amplification of the *amyE-tmrB* region in *Bacillus subtilis*. Agric. Biol. Chem. **49**:545–550.
7. **Hill, D. E., R. Aldape, and J. D. Rozzell.** 1990. Nucleotide sequence of a cyclodextrin glucosyltransferase gene, *cgtA*, from *Bacillus licheniformis*. Nucleic Acids Res. **18**:199.
8. **Ikai, A.** 1980. Thermostability and aliphatic index of globular proteins. J. Biochem. **88**:1895–1898.
9. **Ikeda, H., and M. Shiozaki.** 1984. Nonhomologous recombination mediated by *Escherichia coli* DNA gyrase: possible involvement of DNA replication. Cold Spring Harbor Symp. Quant. Biol. **49**:401–409.
10. **Imanaka, T., M. Fujii, I. Aramori, and S. Aiba.** 1982. Transformation of *Bacillus stearothermophilus* with plasmid DNA and characterization of shuttle vector plasmids between *Bacillus stearothermophilus* and *Bacillus subtilis*. J. Bacteriol. **149**:824–830.
11. **Imanaka, T., T. Himeno, and S. Aiba.** 1985. Effect of *in vitro* DNA rearrangement in the NH_2-terminal region of the penicillinase gene from *Bacillus licheniformis* on the mode of expression in *Bacillus subtilis*. J. Gen. Microbiol. **131**:1753–1763.
12. **Imanaka, T., T. Tanaka, H. Tsunekawa, and S. Aiba.** 1981. Cloning of the genes for penicillinase, *penP* and *penI*, of *Bacillus licheniformis* in some vector plasmids and their expression in *Escherichia coli*, *Bacillus subtilis*, and *Bacillus licheniformis*. J. Bacteriol. **147**:776–786.
13. **Kaneko, T., T. Hamamoto, and K. Horikoshi.** 1988. Molecular cloning and nucleotide sequence of the cyclodextrin glucanotransferase gene from the alkalophilic *Bacillus* sp. strain no. 38-2. J. Gen. Microbiol. **134**:97–105.
14. **Kaneko, T., T. Kudo, and K. Horikoshi.** 1990. Comparison of CD composition produced by chimeric CGTase. Agric. Biol. Chem. **54**:197–201.
15. **Kimura, K., S. Kataoka, Y. Ishii, T. Takano, and K. Yamane.** 1987. Nucleotide sequence of the β-cyclodextrin glucanotransferase gene of alkalophilic *Bacillus* sp. strain 1011 and similarity of its amino acid sequence to those of α-amylase. J. Bacteriol. **169**:4399–4402.
16. **Kimura, K., S. Kataoka, A. Nakamura, T. Takano, and S. Kobayashi.** 1989. Functions of the COOH-terminal region of cyclodextrin glucanotransferase of alkalophilic *Bacillus* sp. #1011: relation to catalyzing activity and pH stability. Biochem. Biophys. Res. Commun. **161**:1273–1297.
17. **Kitahata, S., and S. Okada.** 1982. Comparison of action of cyclodextrin glucanotransferase from *Bacillus megaterium*, *B. circulans*, *B. stearothermophilus* and *B. macerans*. J. Jpn. Soc. Starch Sci. **29**:13–18.
18. **Kitahata, S., and S. Okada.** 1982. Purification and some properties of cyclodextrin glucanotransferase from *Bacillus stearothermophilus* TC-90. J. Jpn. Soc. Starch Sci. **29**:7–12.
19. **Klein, C., and G. E. Schulz.** 1991. Structure of cyclodextrin glycosyltransferase refined at 2.0 Å resolution. J. Mol. Biol. **217**:737–750.
20. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
21. **Kubota, M. Y. Matsuura, S. Sakai, and Y. Katsube.** 1991. Molecular structure of *B. stearothermophilus* cyclodextrin glucanotransferase and analysis of substrate binding site. Denpun Kagaku **38**:141–146.
22. **Kuriki, T., and T. Imanaka.** 1989. Nucleotide sequence of neopullulanase gene from *Bacillus stearothermophilus*. J. Gen. Microbiol. **135**:1521–1528.
23. **Kuriki, T., S. Okada, and T. Imanaka.** 1988. New type of pullulanase from *Bacillus stearothermophilus* and molecular cloning and expression of the gene in *Bacillus subtilis*. J. Bacteriol. **170**:1554–1559.
24. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
25. **Lejeune, A., K. Sakaguchi, and T. Imanaka.** 1989. A spectrophotometric assay for the cyclization activity of cyclomaltohexaose (α-cyclodextrin) glucanotransferase. Anal. Biochem. **181**:6–11.
26. **MacGregor, E. A., and B. Svensson.** 1989. A super-secondary structure predicted to be common to several α-1,4-D-glucan-cleaving enzymes. Biochem. J. **259**:145–152.
27. **Makela, M., P. Mattsson, E. Schinia, and T. Koppela.** 1988. Purification and properties of cyclodextrin glucanotransferase from alkalophilic *Bacillus*. Biotechnol. Appl. Biochem. **10**:414–427.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 149–185. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Matsumura, M., W. J. Becktel, M. Levitt, and B. W. Matthews.** 1989. Stabilization of phage T4 lysozyme by engineered disulfide bonds. Proc. Natl. Acad. Sci. USA **86**:6562–6566.
30. **Matsuura, Y., M. Kusunoki, W. Harada, and M. Kakudo.** 1984. Structure and possible catalytic residues of Taka-amylase A. J. Biochem. **95**:697–702.
31. **Messing, J., and J. Vieira.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
32. **Moran, C. P., Jr., N. Lang, S. F. J. LeGrice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186**:339–346.
33. **Muto, A., and S. Osawa.** 1987. The guanine and cytosine content of genomic DNA and bacterial evolution. Proc. Natl. Acad. Sci. USA **84**:1666–1669.
34. **Nakajima, R., T. Imanaka, and S. Aiba.** 1986. Comparison of amino acid sequences of eleven different α-amylases. Appl. Microbiol. Biotechnol. **23**:355–360.
35. **Nakamura, N., and K. Horikoshi.** 1976. Characterization and some cultural conditions of cyclodextrin glycosyltransferase-producing alkalophilic *Bacillus* sp. Agric. Biol. Chem. **40**:753–757.
36. **Needleman, S. B., and C. D. Wunsh.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**:443–453.
37. **Nitschke, L., K. Heeger, H. Bender, and G. E. Schulz.** 1990. Molecular cloning, nucleotide sequence and expression in *Escherichia coli* of β-cyclodextrin glycosyltransferase gene from *Bacillus circulans* strain no. 8. Appl. Microbiol. Biotechnol. **33**:542–546.
38. **Perry, L. J., and R. Wetzel.** 1984. Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation. Science **226**:555–557.
39. **Pulley, O. A., and D. French.** 1961. Studies on the Schardinger dextrins. XI. The isolation of new Schardinger dextrins. Biochem. Biophys. Res. Commun. **5**:11–15.
40. **Randall, L. L., and S. J. S. Hardy.** 1986. Correlation of competence for export with lack of tertiary structure of the mature species: a study *in vitro* of maltose-binding protein in *E. coli*. Cell **46**:921–928.
41. **Roberts, J. M., L. B. Buck, and R. Axel.** 1983. A structure for amplified DNA. Cell **33**:53–63.
42. **Sakai, S., M. Kubota, K. Yamamoto, T. Nakada, K. Torigoe, O. Ando, and T. Sugimoto.** 1987. Cloning of cyclodextrin glucanotransferase genes from *B. stearothermophilus* and *B. macerans*. J. Jpn. Soc. Starch Sci. **34**:140–147.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
44. **Sinha, N. D., J. Biernat, J. McManus, and H. Koster.** 1984. Polymer support of oligonucleotide synthesis. XVIII. Use of β-cyanoethyl-*N*,*N*-dialkylamino-*N*-morpholino phosphoamidite

of deoxynucleotides for the synthesis of DNA fragments simplifying deprotection and isolation of the final product. Nucleic Acids Res. **12:**4539–4557.

45. **Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk.** 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. **150:**76–85.
46. **Svensson, B., A. J. Clarke, and I. Svensson.** 1983. Amino acid sequence of tryptic fragments of glucoamylase G1 from *Aspergillus niger*. Carlsberg Res. Commun. **48:**517–527.
47. **Svensson, B., H. Jesperson, M. R. Sierks, and E. A. MacGregor.** 1989. Sequence homology between putative raw starch binding domains from different starch-degrading enzymes. Biochem. J. **264:**309–311.
48. **Takagi, M., T. Imanaka, and S. Aiba.** 1985. Nucleotide sequence and promoter region for the neutral protease gene from *Bacillus stearothermophilus*. J. Bacteriol. **163:**824–831.
49. **Takano, T., M. Fukada, M. Monma, S. Kobayashi, K. Kainuma, and K. Yamane.** 1986. Molecular cloning, DNA nucleotide sequencing, and expression in *Bacillus subtilis* cells of the *Bacillus macerans* cyclodextrin glucanotransferase gene. J. Bacteriol. **166:**1118–1122.
50. **Thornton, J. M.** 1981. Disulfide bridge in globular proteins. J. Mol. Biol. **151:**261–287.
51. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.
52. **Yon, J., and M. Fried.** 1989. Precise gene fusion by PCR. Nucleic Acids Res. **17:**4895.