Biosynthesis of D-Alanyl-Lipoteichoic Acid: Cloning, Nucleotide Sequence, and Expression of the *Lactobacillus casei* Gene for the D-Alanine-Activating Enzyme

MICHAEL P. HEATON AND FRANCIS C. NEUHAUS*

Department of Biochemistry, Molecular Biology and Cell Biology,
Northwestern University, Evanston, Illinois 60208

Received 6 December 1991/Accepted 5 May 1992

The D-alanine-activating enzyme (Dae; EC 6.3.2.4) encoded by the *dae* gene from *Lactobacillus casei* ATCC 7469 is a cytosolic protein essential for the formation of the D-alanyl esters of membrane-bound lipoteichoic acid. The gene has been cloned, sequenced, and expressed in *Escherichia coli*, an organism which does not possess Dae activity. The open reading frame is 1,518 nucleotides and codes for a protein of 55.867 kDa, a value in agreement with the 56 kDa obtained by electrophoresis. A putative promoter and ribosome-binding site immediately precede the *dae* gene. A second open reading frame contiguous with the *dae* gene has also been partially sequenced. The organization of these genetic elements suggests that more than one enzyme necessary for the biosynthesis of D-alanyl-lipoteichoic acid may be present in this operon. Analysis of the amino acid sequence deduced from the *dae* gene identified three regions with significant homology to proteins in the following groups of ATP-utilizing enzymes: (i) the acid-thiol ligases, (ii) the activating enzymes for the biosynthesis of enterobactin, and (iii) the synthetases for tyrocidine, gramicidin S, and penicillin. From these comparisons, a common motif (GXXGXPK) has been identified that is conserved in the 19 protein domains analyzed. This motif may represent the phosphate-binding loop of an ATP-binding site for this class of enzymes. A DNA fragment (1,568 nucleotides) containing the *dae* gene and its putative ribosome-binding site has been subcloned and expressed in *E. coli*. Approximately 0.5% of the total cell protein is active Dae, whereas 21% is in the form of inclusion bodies. The isolation of this minimal fragment without a native promoter sequence provides the basis for designing a genetic system for modulating the D-alanine ester content of lipoteichoic acid.

The D-alanyl-ester content of lipoteichoic acid (LTA) appears to play a role in determining cell shape. Mutants that are deficient in these esters are characterized by aberrant morphology and defective cell separation (52). To establish further the role of these esters, genetic methods were sought for modulating their content in LTA.

In *Lactobacillus casei*, the incorporation of D-alanine into ester-linked D-alanyl-LTA is accomplished in the following two-step reaction sequence (6, 9, 11, 42, 51, 54):

In reaction 1, enzyme-bound aminoacyl adenylate is synthesized by the D-alanine-activating enzyme (Dae; EC 6.3.2.4) in the presence of ATP. In reaction 2, the activated D-alanine is covalently linked to membrane-associated LTA (mLTA) in the presence of D-alanine:membrane acceptor ligase (EC 6.3.2.16). The mechanism of LTA acylation catalyzed by the ligase has not been established.

The D-alanyl esters of LTA have been implicated in at least three functions. These include the regulation of autolytic activity (22), the binding of Mg^{2+} for enzyme function (5, 25, 30, 37), and the synthesis of wall teichoic acid (20, 33). In order to clarify the role of these esters in the growth of the organism, additional *in vivo* evidence for these functions is required (18, 19). Experiments have shown that the D-alanyl

ester content of LTA fluctuates in response to different growth conditions. For example, growth of *Staphylococcus aureus* in the presence of increasing concentrations of NaCl decreases the degree of esterification by D-alanine (21). Moreover, growth of this bacterium at pH 8.1 instead of 7.1 also results in a decrease of alanine ester content (45). Because of the multiplicity of salt and pH effects, the correlation of changes in the D-alanine ester content of LTA with aberrant morphology and cell separation is intrinsically difficult.

Thus, a genetic system for altering the D-alanyl ester content of LTA is essential for delineating the role of this macroamphiphile in the morphogenesis of *L. casei*. To accomplish this goal, we have cloned, sequenced, and expressed in *Escherichia coli* the gene for the Dae from *L. casei*. Isolation of this gene and its flanking regions represents a first step in the development of a genetic system for modulating the D-alanine ester content of LTA.

MATERIALS AND METHODS

Bacterial strains, phagemids, and bacteriophages. All bacterial strains, phagemids, and bacteriophages used in this study are listed in Table 1.

Chemicals and reagents. [γ - ^{32}P]dATP and α - ^{35}S -dATP were the products of ICN Biomedicals, Inc., and Amersham Corp., respectively. Medium supplies were obtained from Difco Laboratories. Tetracycline and ampicillin were purchased from Sigma Chemical Co. Nitrocellulose and polyvinylidene difluoride membranes were obtained from Millipore Corp. Restriction enzymes were obtained from a number of

* Corresponding author.

TABLE 1. Bacterial strains, phagemids, and bacteriophages used

Strain, phage, or phagemid	Relevant characteristics ^a	Source
<i>L. casei</i> 7469	Wild type	American Type Culture Collection
<i>E. coli</i> XL1-Blue	<i>supE hsdR17 recA1 endA1 gyrA46 thi relA1 lac F'</i> [<i>proAB</i> ⁺ <i>lacI</i> ^q <i>lacZ</i> ΔM15 Tn10(Tc ^r)]	Stratagene
Phages		
Lambda		
λZapII		Stratagene
λZapII(-)DE24	Contains 2,383-bp <i>EcoRI</i> insert with <i>dae</i> gene, minus-strand rescue	This work
λZapII(+)DE24	Plus-strand rescue	This work
R408 (M13)	Helper phage for single-strand rescue	Stratagene
Phagemids		
pBluescript SK+	Ap ^r	Stratagene
pDPP26	pBluescript SK+ with 2.6-kb <i>PstI</i> <i>L. casei</i> chromosomal fragment containing the first 413 bp of the <i>dae</i> gene	This work
pDAE1	pBluescript SK+ with promoterless 1,568-bp <i>XhoI</i> fragment cloned in sense mRNA orientation	This work

^a Abbreviations: Ap^r and Tc^r, ampicillin and tetracycline resistance, respectively.

suppliers and used according to instructions. Inorganic pyrophosphatase isopropylthiol-β-D-galactoside (IPTG), mung bean nuclease, polynucleotide kinase, and T4 DNA ligase were purchased from either GIBCO BRL, Promega, Sigma Chemical Corp., or Stratagene. Gigapack II extract and StrataClean were obtained from Stratagene. The Sequenase and GeneAmp kits were purchased from United States Biochemical Corp. and Perkin-Elmer Cetus, respectively. Oligonucleotides were synthesized by the Northwestern University Biotechnology Facility. The biconchonic acid protein reagent was purchased from Pierce Chemical Co.

Growth of bacteria. *L. casei* was grown in LC medium adapted from that of Keleman and Baddiley (31), containing 2% glucose, 2% tryptone, 1% yeast extract, 45 mM potassium phosphate (pH 6.3), and trace minerals (166 μM magnesium sulfate, 6 μM manganese sulfate, and 0.7 μM ferrous sulfate). *E. coli* was cultured in Luria-Bertani (LB) broth, plated on LB agar, and preserved by freezing in LB broth plus 10% glycerol at -80°C (47). Tetracycline and ampicillin were used to select and maintain plasmids (10 and 100 μg/ml, respectively). Competent cultures of *E. coli* were prepared by the method of Hanahan (24) and stored at -80°C. Bacteriophage were propagated by standard methods (58).

Purification of Dae from *L. casei* extracts. *L. casei* was grown in 13 liters of LC broth at 37°C in a New Brunswick Microferm fermentor (1 lb/in², 25 rpm). Late-log-phase cultures were harvested by centrifugation (4°C) and washed in cold buffer containing 100 mM Tris-HCl (pH 8.0) and 1 mM EDTA (WB). The cell paste (75 g) was suspended in 1.5 volumes of WB (4°C) with 1.5 volumes of washed glass beads (200 to 300 μm diameter). This mixture was disintegrated for 5 min at 60-s intervals with liquid carbon dioxide cooling in the Bronwill cell homogenizer at 4,000 cycles per min (Braun, model MSK). Beads were removed by low-speed centrifugation, and the supernatant fraction was clarified by centrifugation (25,000 × g, 30 min). Extracts were fractionated by the dropwise addition of 2 volumes of fresh 2% protamine sulfate to 7 volumes of extract. The mixture was stirred on ice for 15 min, and the precipitate was removed by centrifugation at 25,000 × g for 15 min. The supernatant fraction was brought to 70% saturation by the slow addition of pulverized ammonium sulfate and stirred on ice for 60 min. Precipitates were collected by centrifugation

at 25,000 × g for 15 min, dissolved in a minimum amount of WB, and dialyzed overnight against three changes of 50 mM Tris-HCl (pH 7.8) containing 1 mM EDTA.

Dialyzed extracts were clarified by centrifugation and applied to a column of freshly prepared calcium tartrate gel (CTG) (1) equilibrated with 30 mM sodium phosphate (pH 7.8). The CTG column was developed with a linear gradient of the equilibration buffer and 150 mM sodium phosphate buffer (pH 7.8). Active fractions were pooled and concentrated via ultrafiltration and applied to a DEAE-cellulose column equilibrated with 10 mM Tris-HCl (pH 7.8) containing 100 mM NaCl. This column was developed with a linear gradient of the same buffer and 10 mM Tris-HCl (pH 7.8) containing 200 mM NaCl. Active fractions were pooled and concentrated with Centriprep-30 concentrators (Amicon Corp.). The protein solution was applied to a 30-cm Waters Protein-Pak 300 SW size exclusion high-performance liquid chromatography (HPLC) column (Millipore Corp.) and eluted (1 ml/min) with 100 mM potassium phosphate buffer (pH 7.5). Fractions from multiple runs were pooled and stored at 4°C.

Assay of Dae. The activity of Dae was determined by the assay of Baddiley and Neuhaus (6). The reaction mixture (0.5 ml) contained 100 mM magnesium chloride, 100 mM Tris-HCl (pH 7.8), 100 mM D-alanine, 10 mM ATP, 1 M salt-free hydroxylamine, and 1 U of inorganic pyrophosphatase. This mixture was incubated at 37°C for 15 min before the reaction was terminated by the addition of 2 ml of 1.46 M trichloroacetic acid-0.7 M HCl. The insoluble protein was removed by low-speed centrifugation. The amount of alanine hydroxamate in a 2-ml sample was quantified by the addition of 1 ml of 1.11 M ferric chloride, and the *A*₅₄₀ of the solution was read immediately. One unit of enzyme is defined as that amount of enzyme which catalyzes the formation of 1 μmol of alanine hydroxamate per h. Specific activity is expressed as the number of units per milligram of protein.

The Dae activity expressed by *E. coli* upon infection by bacteriophage λ strains containing *L. casei* cloned inserts was measured in plate lysates. Log-phase cultures were collected by centrifugation and then washed and suspended in 10 mM magnesium sulfate to a turbidity of 1.4 (measured at 600 nm). Approximately 10⁷ PFU were used to infect 200 μl of culture. Soft agar overlays of infected cultures were

incubated overnight at 37°C. Proteins were eluted from plate lysates by the addition of WB (3 ml) and then shaken gently for 90 min (23°C). The eluate was extensively dialyzed against 1,000 volumes of WB at 4°C, and the particulate material was removed by centrifugation. The supernatant fraction was used directly to assay for Dae activity.

SDS-PAGE and limited N-terminal microsequencing. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) of proteins was performed by the method of Laemmli (36). Preparative SDS-PAGE was performed with pooled fractions (50 µg of protein) from size exclusion HPLC. The gel was electroblotted to a polyvinylidene difluoride membrane, which was then stained with Coomassie blue R-250. The 56-kDa band was excised from the membrane and used directly for N-terminal sequencing. This analysis was performed with an Applied Biosystems model 477A pulsed liquid-phase peptide sequencer (26).

DNA techniques. *L. casei* cells were lysed by the method of Chassy and Giuffrida (10). Restriction digests, agarose gel electrophoresis, small- and large-scale plasmid preparations, transformations, isolation of chromosomal DNA, Southern blots, and plaque and colony hybridizations were performed by standard techniques (8, 27, 47, 58, 62). DNA was sequenced by the dideoxy chain termination method (59). Double-stranded template DNA was prepared by equilibrium centrifugation in CsCl-ethidium bromide gradients. Single-stranded template DNA was obtained by rescue from λZapII(-)DE24 and λZapII(+)-DE24 with helper phage R408.

Computer analysis. DNA and deduced amino acid sequences were analyzed with the University of Wisconsin Genetics Computer Group (UWGCG) sequence analysis software package, version 7.1 (14). Sequence similarity searches of the GenBank, EMBL, SWISS-PROT, and PIR data bases were carried out with the TFASTA algorithm. Comparison of homologous regions I, II, and III was accomplished with PILEUP, and the percent identity and similarity were calculated with BESTFIT. The molecular mass and isoelectric point of Dae were calculated by PEPTIDE SORT.

Construction of pDPP26. *L. casei* chromosomal DNA (75 µg) digested with *Pst*I was resolved by preparative agarose gel electrophoresis, the 2.6-kb region was excised, and the DNA was recovered by electroelution. Competent *E. coli* cells were transformed with this DNA ligated into pBluescript SK+. Colonies were screened with a radiolabeled pool of oligonucleotides (nucleotides [nt] 1 through 20). Positive colonies were isolated in pure culture, and mini-preparations of phagemids were obtained for further analysis.

Construction of an *L. casei* HindIII-BamHI library in λZapII. *L. casei* chromosomal DNA fragments were purified by agarose gel electrophoresis after *Bam*HI digestion and again after *Hind*III digestion. This fraction (4 to 6 kb) was treated with mung bean nuclease for 30 min at 50°C to generate blunt ends. *Eco*RI restriction sites were introduced by blunt-end ligation of phosphorylated *Eco*RI linkers, followed by enzymatic digestion with *Eco*RI. After removal of small DNA fragments by membrane filtration, the resulting DNA was ligated to λZapII *Eco*RI arms. This preparation was packaged and used to infect *E. coli*. Phage from resulting plaques were transferred to nitrocellulose membranes and screened with a radiolabeled 15-mer probe (nt 94 through 108). Plaques from hybridizing clones were purified, and their phagemids were rescued by helper phage R408 and established in *E. coli* as described by the supplier. Miniprep-

arations of rescued phagemids were used for restriction analysis of cloned fragments.

Construction of λZapII(+)-DE24. DNA from λZapII(-)DE24 (Table 1) was digested with *Eco*RI, ligated, packaged, and used to infect *E. coli*. Clear plaques were chosen from color selection plates for purification. Phagemids were rescued and established in *E. coli*. Restriction digests of phagemid DNA were used to determine the orientation of the fragment with respect to M13 initiation and termination sites.

Construction of pDAE1. Amplification of a promoterless *dae* gene was accomplished by the polymerase chain reaction (56) in a DNA Thermal Cycler (Perkin-Elmer Cetus). The reaction was carried out with λZapII(-)DE24 DNA and the mutagenic oligonucleotides containing *Xho*I sites. The DNA was denatured at 95°C for 1 min, annealed at 45°C for 1.5 min, and extended at 72°C for 2.5 min (26 cycles). The amplified DNA was digested with *Xho*I, ligated into pBluescript SK+, and used in the transformation of competent *E. coli*. Phagemid DNA from colonies identified by Southern hybridization was digested a second time with *Xho*I, purified on an agarose gel, and ligated again into pBluescript SK+. A phagemid containing one copy of the *dae* gene in the sense mRNA orientation (pDAE1) was selected for expression of *dae* in *E. coli*.

Expression of Dae in *E. coli*. *E. coli* cultures harboring pBluescript or pDAE1 were grown in LB broth (100 ml) with ampicillin and tetracycline at 37°C (300 rpm). When the A_{600} reached 0.7, cultures were induced with 4 mM IPTG. Log-phase cells were harvested 2 h after induction, suspended in WB (4°C) with lysozyme (5 mg/ml), and incubated for 20 min at 23°C. One volume of washed glass beads was added, and the cells were disrupted by vortexing three times (15 s each). Supernatant fractions were collected by centrifugation at 4°C and dialyzed overnight against 1,000 volumes of WB. After dialysis, extracts were clarified by centrifugation prior to enzyme and protein assays. Expression of the 56-kDa protein was quantified by harvesting samples 2 h after induction, suspending whole cells in 0.1 volume of SDS-PAGE loading buffer, and performing electrophoresis by the method of Laemmli (36). Densitometry was accomplished with Coomassie blue-stained gels with an LKB Ultrosan XL laser densitometer and Gelscan XL software.

Nucleotide sequence accession number. The sequence reported has been entered into GenBank under accession number M83993.

RESULTS

Isolation and limited N-terminal sequencing of Dae. The N-terminal sequence of the Dae was established in order to synthesize an oligonucleotide probe for identifying clones containing the *dae* gene. The initial purification of this enzyme by Baddiley and Neuhaus (6) yielded a preparation with a specific activity of 29.8 U/mg. To improve the purification, fractionation on hydroxyapatite was replaced with chromatography on CTG. Further purification of the CTG eluate on DEAE-cellulose followed by HPLC gel filtration on Waters Protein Pak SW 300 yielded a preparation with a specific activity of 920 U/mg (Table 2). The apparent molecular mass was estimated from gel filtration to be 41 ± 2 kDa. Nondenaturing PAGE of the preparation (specific activity, 920 U/mg) yielded an active fraction with an apparent molecular mass of 45 ± 3 kDa. The specific activity of the enzyme from this gel was estimated to be $5,000 \pm 2,000$ U/mg. Excision of this fraction and subse-

TABLE 2. Purification of Dae from *L. casei*

Fraction	Vol (ml)	Protein (mg)	Total activity (U)	Sp act (U/mg)	Purification (fold)	Recovery (%)
Crude extract ^a	72	1,470	4,380	3.0	1.0	100
0–70% ammonium sulfate	12	289	4,300	16.2	5.4	98
Calcium tartrate gel	24	16	2,530	149	49	58
DEAE-cellulose	7.8	1.6	754	477	159	17
Size exclusion HPLC	1.0	0.13	115	920	307	2.5

^a Resulting from 75 g of washed cells.

quent resolution on SDS-PAGE yielded a 56-kDa band. The analysis of protein by SDS-PAGE at each stage of purification revealed that a 56-kDa protein was the primary candidate for the Dae. From these results, it was concluded that the 56-kDa protein (SDS-PAGE) was Dae; thus, the predicted coding region for the gene was approximately 1.5 kb long. To obtain the N-terminal amino acid sequence, the 56-kDa band was electroblotted from the SDS-PAGE gel to a polyvinylidene difluoride membrane, excised, and microsequenced (Materials and Methods). The N-terminal sequence is shown in Fig. 1.

Cloning of a chromosomal DNA fragment containing the N-terminal region of the *dae* gene. To identify clones containing the N-terminal coding region for the *dae* gene, a 20-mer oligonucleotide pool (128-fold degenerate) was synthesized with the following constraint in codon usage (Fig. 1). Since usage of the ATA codon for isoleucine is rare in lactobacilli (40), adenine was not included in the wobble position of that codon. This pool of oligonucleotides was used in Southern analyses for identifying restriction fragments containing the N-terminal region of the *dae* gene.

A single 2.6-kb fragment was identified from the *Pst*I digest of chromosomal DNA and cloned into the pBluescript phagemid vector. The restriction map of this fragment is shown in Fig. 2A. This phagemid (pDPP26) was partially sequenced to determine the direction and location of the *dae* gene. Translation of the first 63 nt in this open reading frame (ORF) yielded a deduced peptide sequence which was identical to that obtained from limited N-terminal sequencing of the 56-kDa protein. From this *Pst*I fragment, only 413 bp of the coding region were obtained. To facilitate further cloning, a map of the *dae* chromosomal region was generated by a combination of Southern analyses of *L. casei* restriction fragments and restriction analyses of pDPP26 (Fig. 2B). This map shows three restriction fragments (6.1-kb *Kpn*I, 5.9-kb *Hind*III, and 5.5-kb *Bam*HI) which contain the gene.

Cloning of a chromosomal fragment containing the *dae* gene. Attempts to clone the entire gene with the use of the pBluescript phagemid were unsuccessful. All of the isolated phagemid constructs that contained the *dae* promoter and coding region were unstable in *E. coli* XL1-Blue. Thus, cloning in λ ZapII was initiated.

The 5.5-kb *Bam*HI fragment (Fig. 2B) containing the *dae* gene was enriched by agarose gel electrophoresis. This fraction was digested with *Hind*III to yield a 4.8-kb *Hind*III-*Bam*HI fragment, which was also purified by electrophoresis. This preparation was digested with mung bean exonuclease, and the blunt ends were ligated to *Eco*RI linkers. After cleavage with *Eco*RI, DNA fragments were ligated to λ ZapII *Eco*RI arms. The ligated λ DNA was packaged in vitro, and the resulting library was screened by Southern analyses with a 15-mer oligonucleotide probe complementary to nt 94 to 108. Approximately 1% of the clones in this

library hybridized with this probe. Of the six plaques purified, at least two contained insert DNA (2.4 and 4.3 kb) and were further screened for enzyme activity in plate lysates. At least one strain [λ ZapII(–)DE24], containing the 2.4-kb fragment shown in Fig. 2C, expressed D-alanine-activating activity upon infection of *E. coli*. Because the expected fragment length was 4.8 kb, exonuclease treatment before the ligation of *Eco*RI linker DNA was apparently excessive. This 2.4-kb *Eco*RI fragment was subcloned in the opposite orientation [λ ZapII(+)DE24]. Both clones were sequenced from single-stranded templates with 20 specific synthetic oligonucleotides as primers (Fig. 3).

The first 21 codons of ORF1 correspond to the amino acid sequence determined from N-terminal sequencing of Dae. The mass calculated from the translation of ORF1 is 55.867 kDa (506 residues), which is in agreement with the 56 kDa determined by SDS-PAGE. The isoelectric point, calculated by using the UWCGC sequence analysis algorithm, is 5.18. An interesting feature of the amino acid sequence is the presence of only one cysteine (residue 269).

The DNA sequence analysis 5' to the *dae* gene reveals a putative promoter region and a ribosome-binding site (Fig. 3). The spacing between the –35 and –10 promoter sequences is 20 nt, and the distance between the ribosome-binding site and the start codon is 8 nt. The –35 sequence (TCTTGGTTTAT) is identical to that reported for the lactose permease from *Lactobacillus bulgaricus* (39). The putative ribosome-binding site sequence GGGGG is identical to that reported by Lerch et al. (41) for the *L-HicDH* gene from *Lactobacillus confusus* and similar to ribosome-binding site sequences tabulated by Alpert and Chassy (3) for lactobacilli. The DNA sequence 3' to the *dae* gene contains another ORF (ORF2) which has no apparent promoter or ribosome-binding site. The start codon of ORF2 shares the last adenine of the termination codon from the *dae* gene (TAATG). A physical map of these elements is presented in Fig. 4.

Codon usage for the *dae* gene and ORF2 is similar to that reported for other lactobacilli (40). Of the 47 codons for isoleucine and the 27 codons for arginine, neither AGA or ATA was found in either ORF. The GC content of the *dae*

FIG. 1. N-terminal sequence of Dae with the corresponding pool of oligonucleotide probes used for the identification of the *dae* gene. The sequence obtained was MIDNVITAIDRVAAEHPT RVA.

FIG. 2. Cloning strategy for the *L. casei dae* gene. (A) Restriction map of the *Pst*I insert from pDPP26. (B) Chromosomal map of the *dae* gene region. (C) Chromosomal insert from λ ZapII(-)DE24. (D) The *dae* gene from pDAE1. Abbreviations: ATG, start codon; TAA, termination codon; †, restriction sites not present on *L. casei* chromosomal DNA; B, *Bam*HI; C, *Cl*aI; D, *Dra*I; E, *Eco*RI; Hc, *Hinc*II; Hd, *Hind*III; K, *Kpn*I; N, *Nru*I; P, *Pst*I; Pv1, *Pvu*I; Pv2, *Pvu*II; Sp, *Sph*I; Ss, *Ssp*I; St, *Sty*I; X, *Xho*I.

gene and ORF2 is 46.7 and 43.6%, respectively, whereas that of the upstream region is 40.6%.

Sequence homology of Dae with other proteins. A TFASTA search of the computer data bases for proteins with significant sequence homology to Dae has identified 19 functional domains among 15 proteins. The highest overall sequence identities (27.6 to 31.5%) were observed with a group of five enzymes (Table 3, group I) that catalyze the non-ribosomal peptide synthesis of the antibiotics penicillin, gramicidin S, and tyrocidine. A second group of three proteins required for biosynthesis of the siderophores enterobactin (in *E. coli*) and anguibactin (in *Vibrio anguillarum*) (Table 3, group II) was observed to have sequence identity with Dae (23.4 to 28.1%). α -Amino adipate reductase from the lysine biosynthesis pathway of *Saccharomyces cerevisiae* (27.5% identical) represents a third category of proteins homologous to Dae. The acid-thiol ligases (Table 3, group IV) represent a fourth group of enzymes having overall sequence identity with Dae of 19.4 to 23.8%. The enzymes in each of these categories, with the exception of AngR, catalyze the activation of an amino or fatty acid via an acyl adenylate intermediate. Each of the proteins or domains in Table 3 contains a "core" peptide sequence (GXXGXPK) that is conserved (Fig. 5, region I). Although the overall sequence identity of Dae with the acid-thiol ligases is low, the high identity observed in region I further supports the participation of this motif in the formation of the acyl adenylate intermediate.

Two additional regions, II and III, from proteins listed in the first three groups in Table 3 were found to have high identity with Dae. With the exception of EntE in region II, the identity with Dae in these regions is 42.3 to 65.4% (Fig. 5). The consensus sequences for regions II and III are

GRXDXQXKXXGXRXE and PX₉PX₅KXDX₃L, respectively. The average number of residues between regions I and II and regions II and III is 238 ± 6 and 52 ± 6 , respectively. The number of residues between regions I and II and regions II and III in Dae is 232 and 57, respectively (Fig. 4). Unlike the first three groups of proteins in Table 3, the acid-thiol ligases do not have significant homology with Dae in these regions (<35%).

The following proteins were not identified by TFASTA searches of the data bases and show only low homology in regions I, II, and III when compared with Dae by BESTFIT (<30%): aminoacyl-tRNA synthetases (EC 6.1.1.X), D-alanine:D-alanine ligase (ADP) (EC 6.3.2.4), and alanine racemase (EC 5.1.1.1).

Analysis of the deduced sequence from the partial ORF2. The termination codon of ORF2 was not found on the 2.4-kb fragment; thus, a complete sequence analysis has not yet been possible. A preliminary analysis of the 188 amino acids deduced from the N terminus of ORF2 suggests that it is a basic membrane protein. A Kyte-Doolittle plot (35) shows six regularly spaced hydrophobic domains of similar lengths with intervening regions of charged residues (data not shown). The isoelectric point calculated for this partial ORF is 10.22.

Subcloning and expression of the *dae* gene. A 1,568-bp DNA fragment containing the *dae* gene and ribosome-binding site was amplified by the polymerase chain reaction and subcloned into pBluescript SK+ (Fig. 2D). This was required because of the apparent toxicity of the 2.4-kb DNA fragment in *E. coli* phagemid vectors. Two mutagenic primers were designed to each contain an *Xho*I site. One primer was complementary to the minus strand of the *dae* gene and

-291		EcoRI GG·AATTCC·TCTAACTGGCCTG	-271
-270	TTTGATTCAATTACGAAACTGTTAGGTGATTGTACGCTTTTGGCTTGGCGAGGCTTTTCTGTGTAATAATATGATGTGTCAATAAG	-181	
-180	AACGTGGGGTGACAAGATGTTTCGCAAAATTAAGACGGTGTGGGAAATCCACATTTCGGTTCATCTTTGGACCGTCACGTATTTCGT	-91	
-90	GATACTGCTGGTCTTGGTTTATCTATACGGTTACAGCGGTATTAAACAACAGTAAATTCATTATAACGAGTTTAAAGGGGAAACCGCG	-1	
1	ORF1- ATGATCGATAATGTTATTACAGCGATTGATCGGGTCGCCGAGAGCATCCAACGCGCGGTGGCTTATGATTATGAAGGAACCAATATACC	90	
1	M I D N V I T A I D R V A A E H P T R V A Y D Y E G T Q Y T	30	
91	TACGCCAGTTGAAAGAAGCAGTGACCGCTTGTCTGATTCTTTGCGGAATCATTGCCGGCTGGTGAACCCATCATTGTCTATGGTGGT	180	
31	Y A Q L K E G S D R L A G F F A E S L P A G E P I I V Y G G	60	
181	CAAACCTTTGACATGGTTGAGGTTTCTTGGACTTTCTAAATCAGGCGATGCCTACATTCCGATCGACAGCATAGTCCAAATGAACGG	270	
61	Q T F D M V E V F L G L S K S G H A Y I P I D T H S P N E R	90	
271	ATTACCCAGGTGCAAGATGTTTCCCATGCGCCTGCTGTGATTGAAGTGGCACCCTTCGGATTACGGTCCAGATGTCAAGATTATTCCG	360	
91	I T Q V Q D V A H A P A V I E V A P L P I T V P D V K I I R	120	
361	GCACCGCGGTTACATCAAGCTGAGCAGTCACACGCGCAATTCATTCTCAGTCAGCATGCGGTTGCTGGCGATGATAATTACTACATCATC	450	
121	A P A L H Q A E Q S H A P I H S L Q H A V A G D D N Y Y I I	150	
451	TTTCAAGCGGCACAACGGGAAACCAAGCGGTTAGTACGATAACCTTTTGGCTATGTGAACCTGAATATTTCCGACTTT	540	
151	F T S G T T G K P K G V Q I S H D N L L S Y V N W N I S D F	180	
541	GGGCTTGAAGAAGCGGTGTGCCCATGTCTCAGCGCGGTTATTCCTTTGACCTAAGTGTGATGGATCTTTATCCGACCTTGGTTTATGGC	630	
181	G L E E G V V A M S Q P P Y S F D L S V M D L Y P T L V L G	210	
631	GGCACCTTGAAGGCTTTACCAAAAGAGTCACTGATAACTTTAAAGAATTATTTCGACGTTGCCAAAATTGGGCTTGAACGAATGGGTA	720	
211	G T L K A L P K E V T D N F K E L F A T L P K L G L N E W V	240	
721	TCAACCCCTTCTTTTGTGAAATTCGCTTACTTGATCCGAATTTCAAGCAGGAAATATCCGAACCTGACCCATTTTCTGTTTTCGGGT	810	
241	S T P S F V E I A L L D P N F K Q E N Y P N L T H F L F C G	270	
811	GAGGAGCTAGTTAATAAAACCGCCCAAGCGCTGATCACAGTTCGCCGAAAGCAACCGTCTATAACAGTATGTCGACCGCAAGCCACC	900	
271	E E L V N K T A Q A L I T R F P K A T V Y N T Y G P T E A T	300	
901	GTTGCGGTTACGGGAATGGCGATTACCCAGCGGATTGTTGATCAATATCCGCGGTTACCGATTGGCTATGCGAAGCCTGATACCAACGTT	990	
301	V A V T G M A I T Q A I V D Q Y P R L P I G Y A K P D T N V	330	
991	TATGTGGTTGATGAGCAAGGTGAACAGGTTTCAGCTGGCACAGAAGCGAATTGATGATCGTGGGTCCAAGCGTGTGAAAGGGTATCTC	1080	
331	Y V V D E Q G E Q V S A G T E G E L M I V G P S V S K G Y L	360	
1081	AATAATCCGACAAACAGCCCGGCATTTTCAAGGCTGGCAATCAACGCGGTTATCGGTCTGGTGACTTAGTGACCATGACCCAGAT	1170	
361	N N P D K T A A A F F K A G N Q R G Y R S G D L V T M T A D	390	
1171	GGGATGGTCTTTTATCGCGGTCGAGCTGATTTCAGGTTAAGTTCATGCTGTTATCGCAITGAGCTAGAAGATGTTGATCATAATTTGAAT	1260	
391	G M V F Y R G R T D F Q V K L H G Y R I E L E D V D H N L N	420	
1261	CAGGTTTCTTATTAAGCAGGCATCAACGGTTCCGCGGTATGACAAGGATCATAAAGTTGCCAGTTGATTGCCTTTGCGGTTGCCAAA	1350	
421	Q V S Y I K Q A S T V P R Y D K D H K V A Q L I A F A V A K	450	
1351	CCAAATGACTTTGATAGCGAAATGAACTAACGCAAGCAATTAAGCGCACTAGGGAATGTTGATGGAATATATGATTCCCCAACGG	1440	
451	P N D F D S E M K L T Q A I K A E L G K M V M E Y M I P Q R	480	
1441	ATTATTTACCGGGATCAGTTACCGTTAAGCGCTAATGGCAAGGTTGATCGCAAGCACTGATTGCAGAGGTAATCATTAACTGCTTAAT	1530	
481	I I Y R D Q L P L T A N G K V D R K A L I A E V N H ***	506	
	ORF2- M L N L	4	
1531	TACAACCGTATGAGAATCCACAGTATTTTGTCTATTTAATCATTGCCTTGTCTCCCGTCATCATCGGCATGTTCAAAGGCTTTTCGGATGC	1620	
5	Q P Y E N P Q Y F V Y L I I A L L P V I I G M F K G F R M H	34	
1621	ATTGGTATGAAAGCATCTTTTCTGTTGGTTTCTGCTGACTGATTTTGTATGCGGATAAATGGCCTCAAGGAAAAGCGTTATTGGGGTATG	1710	
35	W Y E S I F S L V F L V L I F D A D K W P Q G K A L L G Y V	64	
1711	TGGTTTCAACTTATTACTGGTTTACGCGTATTTCAAATATCGAACGCGGAAGGCTCAAAAACTCCACGGCAGTCTTTTATCTTAGTG	1800	
65	V F N L L L V Y A Y F K Y R T R E G S K N S T A V F Y L S V	94	
1801	TCGCGTTAGGAATTGCGCACGTGGCAGTTGTCAAGTTTACGCGCGTGTTCACATCATGGGTCTATTCTTGGCTTCTTGGGCATTAGTT	1890	
95	A L G I A H V A V V K F T P L F Q H H G S I L G F L G I S Y	124	
1891	ATTTAACTTTCCGAGTGGTCCGACGATTATGAAATTCGTGATGGGTCGATTAAAGACCTGAACATGTGGAAGTTTATCCAATTTCTGC	1980	
125	L T F R V V G T I M E I R D G S I K D L N M W K F I Q F L L	154	
1981	TATCTTCCCAACCATCTCAAGTGGCCGATTGATCGGTATCGGCGATTGTCAAAGATTACGATCGCGTTCTGTATCCCGAGCATTACG	2070	
155	F F P T I S S G P I D R Y R R F V K D Y D R V P D P E H Y A	184	
2071	CCCAGTTGGTGACC·GG·AATTCC Q L V T	2092 188	

FIG. 3. DNA sequence of λ ZapII(-)DE24 insert from *L. casei*, with the deduced amino acid sequences of ORF1 (506 residues) and ORF2 (188 residues). Putative -35, -10, and ribosome-binding site (rbs) sequences for ORF1 are underlined in the DNA sequence, and nucleotides of the termination codon are labeled with asterisks. Restriction sites relevant to DNA sequencing are presented. Bases are numbered relative to the initiation codon of ORF1, and amino acids are numbered relative to the putative start of the ORFs. Regions of high identity with the proteins listed in Table 3 are underlined and identified in the amino acid sequence.

FIG. 4. Physical map of the 2.4-kb insert from λ ZapII(-)DE24 containing ORF1 (*dae*) and a portion of ORF2. The beginning of each ORF is represented by arrows, and regions corresponding to putative -35, -10, and ribosome-binding site (rbs) sequences are labeled with vertical lines. The horizontal lines show relative lengths and positions of amino acid (A. A.) regions I, II, and III, which are homologous to the proteins in Table 3.

positioned the *Xho*I site between the putative -10 promoter sequence and the ribosome-binding site. The second primer was complementary to the plus strand and positioned the *Xho*I site 18 nt after the termination codon for the *dae* gene. Cleavage with *Xho*I produced a 1,568-nt fragment (nt -25 to 1539) that was ligated to pBluescript SK+. An *E. coli* clone was identified that contained a stable phagemid (pDAE1) with one copy of the *dae* gene in the same orientation as the *lac* promoter.

E. coli does not contain detectable levels of Dae activity and thus provides a useful host for the expression of *dae*. The *lac* promoter was used to drive transcription of the *dae* gene with the *L. casei* ribosome-binding site. Two hours after induction with IPTG, a 56-kDa protein accumulated to 21% of the total cell protein. When viewed by phase-contrast

microscopy, large inclusion bodies were observed in the induced cultures containing pDAE1. The specific activity of soluble Dae in dialyzed *E. coli* crude extracts was 10 U/mg. This specific activity is approximately fivefold higher than that of *L. casei* extracts. It was estimated that 0.5% of the total cell protein is active Dae.

DISCUSSION

The 1,518-nt ORF1 from *L. casei* codes for the Dae. This gene has been sequenced, subcloned, and expressed in *E. coli*. Insights into the mechanism of D-alanine activation are provided from comparisons of the deduced amino acid sequence with those of other proteins. Expression of the *dae* gene in *L. casei* with foreign promoters will provide a means for modulating the D-alanyl ester content of LTA. Thus, an *L. casei* expression vector that complements an isogenic *dae* mutant will facilitate our ability to determine the function of these esters in the morphogenetic program of this organism.

The identification of ORF1 as the *dae* gene is based on the pDAE1-dependent Dae expression in *E. coli*. In addition, the calculated molecular mass (55.867 kDa) and isoelectric point (5.18) of the protein are consistent with those observed for the Dae. The exact match of the deduced N-terminal sequence to that determined by limited N-terminal sequencing of the 56-kDa protein also supports this conclusion. Examination of the upstream flanking region from ORF1 reveals at least one region with putative -35 and -10 promoter and ribosome-binding site sequences. The ribosome-binding site is 8 nt from the *dae* start codon and is composed of five

TABLE 3. Proteins with sequence homology to Dae

Group	Gene	Protein	Organism	Reference(s)	% Identity ^a	% Similarity ^a
I ^b	<i>grs2-pro</i>	Gramicidin S synthetase 2 Proline-activating domain	<i>Bacillus brevis</i>	29	31.5	53.3
	<i>acvA</i>	AcvA synthetase	<i>Penicillium chrysogenum</i>	15, 61		
		Domain A			29.7	54.4
		Domain B			29.4	52.2
		Domain C			30.0	53.7
	<i>acvA</i>	AcvA synthetase	<i>Aspergillus nidulans</i>	46		
		Domain A			29.0	53.8
		Domain B			30.1	50.3
		Domain C			28.6	50.5
	<i>gs1 (grsA)</i> <i>tycA</i>	Gramicidin S synthetase 1 (EC 5.1.1.11) Tyrocidine synthetase 1 (EC 5.1.1.11)	<i>Bacillus brevis</i> <i>Bacillus brevis</i>	28, 34 66	28.2 27.6	49.4 50.9
II ^c	<i>entF</i>	Serine-activating enzyme	<i>Escherichia coli</i>	55	28.1	52.0
	<i>angR</i>	Positive transcriptional activator	<i>Vibrio anguillarum</i>	17	28.1	49.2
	<i>entE</i>	2,3-Dihydroxybenzoate-activating enzyme	<i>Escherichia coli</i>	43	23.4	44.9
III ^d	<i>lys2</i>	α -Aminoadipate reductase	<i>Saccharomyces cerevisiae</i>	23	27.5	47.9
IV ^e	<i>4cl</i>	4-Coumarate-coenzyme A ligase (EC 6.2.1.12)	<i>Petroselinum crispum</i>	44	23.8	46.8
	<i>acs</i>	Long-chain acyl-coenzyme A ligase (EC 6.2.1.3)	<i>Rattus norvegicus</i>	64	23.3	47.7
	<i>acs</i>	Acetyl-coenzyme A ligase (EC 6.2.1.1)	<i>Methanoxanthus soehngenii</i>	16	23.2	50.3
	<i>luc</i>	Firefly luciferase	<i>Luciola cruciata</i>	48	22.4	45.4
	<i>acu-5</i>	Acetyl-coenzyme A ligase (EC 6.2.1.1)	<i>Neurospora crassa</i>	12	22.2	44.1
	<i>facA</i>	Acetyl-coenzyme A ligase (EC 6.2.1.1)	<i>Aspergillus nidulans</i>	12	19.4	43.9

^a Calculated by using BESTFIT with a gap weight of 3.0 and length weight of 0.1.

^b Nonribosomal peptide synthesis via thiolester intermediates.

^c Siderophore biosynthesis.

^d Lysine biosynthesis.

^e Acid-thiol ligases and related enzymes.

Protein	Residue	Region	Identity	Similarity
I				
Dae	145	D N Y Y I I F T S G T T G K P K G V Q I S H	100.0	100.0
Grs 2-pro	606	D L F Y I I Y T S G T T G K P K G V M L E H	72.7	86.4
AcvA (A. n.)	1559	D L A Y I I Y T S G T T G K P K G V M V E H	72.7	81.8
AcvA (P. c.)	1533	D L A Y I I Y T S G T T G R P K G V T V E H	68.2	81.8
AngR	594	Q P A Y I I Y T S G S T G T P K G V V I S H	68.2	77.3
TycA	171	D L A Y V I Y T S G T T G K P K G T M L E H	63.6	77.3
AcvA (P. c.)	2615	N L A Y I I F T S G T S G K P K G V L V E Q	63.6	72.7
AcvA (A. n.)	2639	N L A Y I I F T S G T S G K P K G V L V E Q	63.6	72.7
Gs 1	183	D L A Y V I Y T S G T T G N P K G T M L E H	59.1	72.7
EntF	596	H T A Y I I F T S G S T G R P K G V M V G Q	59.1	77.3
Acs (R. n.)	270	D L A I I C F T S G T T G N P K G A M V T H	59.1	63.6
Acu-5	236	D P L F L L Y T S G S T G K P K G V A H S T	54.5	72.7
Acs (M. s.)	282	H R L Y I L Y T S G T T G K P K G I E H A T	54.5	72.7
FacA	265	D P L F L L Y T S G S T G K P K G V M H T T	50.0	68.2
AcvA (A. n.)	473	Q L A Y V T Y T S G T T G F P K G I L K Q H	50.0	63.6
AcvA (P. c.)	446	Q L A Y V T Y T S G T T G F P K G I F K Q H	50.0	63.6
4Cl	183	D V V A L P Y S S G T T G L P K G V M L T H	50.0	63.6
Luc	193	Q V A L I M N S S G S T G L P K G V Q L T H	50.0	63.6
Lys2	414	S N P T L S F T S G S E G I P K G V L G R H	50.0	54.5
EntE	191	E V A Y F Q L S G G T T G T P K L I P R T H	36.4	59.1
II				
Dae	399	G R T D F Q V K L H G Y R I E L E D V D H N L	100.0	100.0
AcvA (P. c.)	1792	G R N D F Q V K I R G L R I E L G E I E A I L	56.5	82.6
AcvA (A. n.)	2898	G R A D L Q V K I R G Y R I E P S E V Q N V L	56.5	82.6
AngR	864	G R R D K Q V K V G G Y R I E L G E I E V A L	56.5	78.3
Gs 1	427	G R I D N Q V K I R G H R V E L E E V E S I L	56.5	78.3
AcvA (P. c.)	2874	G R G D L Q I K M R G Y R I E I S E V Q N L V	52.2	87.0
Grs 2-pro	869	G R A D H Q V K I R G H R I E L G E I E A Q L	52.2	78.3
AcvA (A. n.)	1818	G R N D F Q V K I R G L R I E L G E I E A V M	52.2	82.6
AcvA (A. n.)	729	G R A D F Q I K L R G I R I E P G E I E S T L	52.2	78.3
AcvA (P. c.)	702	G R A D F Q I K L R G I R I E P G E I E T M L	52.2	78.3
TycA	415	G R I D H Q V K I R G H R I E L G E I E S V L	52.2	78.3
Lys2	696	G R A D D Q V K I R G F R I E L G E I D T H I	52.2	89.0
EntF	854	G R S D D Q L K I R G Q R I E L G E I D R V M	47.8	82.6
III				
Dae	478	Y M I F Q R I I Y R D Q L P L T A N G K V D R K A L	100.0	100.0
Grs 2-pro	940	Y M I F S F F V P L D Q L P L T P N G K I D R K S L	65.4	80.8
AngR	923	Y M V P K R I I F L E T F P L T A N G K V D H K A L	65.4	88.5
Gs 1	498	Y M I F S Y F I Q L D K M P L T S N G K I D R K Q L	61.5	73.1
EntF	933	H M V P V V L L Q L P Q L P L S A N G K L D R K A L	57.7	73.1
TycA	486	Y M L P S Y F V K L D K M P L T P N D K I D R K A L	53.8	80.8
EntE	509	F K L P D R V E C V D S L P L T A V G K V D K K Q L	53.8	76.9
AcvA (A. n.)	806	Y M I P T R L V R V S Q I P V T V N G K A D L R A L	53.8	73.1
AcvA (A. n.)	1890	Y M I P S S F I P I S S L P V T P S G K L D T K A L	53.8	69.2
AcvA (P. c.)	1864	Y M V P S R L I L V S K F P V T P S G K L D T K A L	50.0	73.1
AcvA (P. c.)	779	Y M I P T R L V Q L S Q I P V N V N G K A D L R A L	50.0	69.2
Lys2	796	Y A M P S L I V V M D K L P L N P N G K V D K P K L	50.0	65.4
AcvA (P. c.)	2951	Y M V P S H L C L E G A L P V T I N G K L D V R R L	42.3	73.1
AcvA (A. n.)	2975	Y M V P K L Y C L E G G L P V T I N G K L D V R K L	42.3	65.4

FIG. 5. Three regions of the Dae sequence which have significant identity to other proteins. Regions I, II, and III of Dae were compared with the protein domains listed in Table 3 by BESTFIT. Regions from protein domains having an identity of >35% are presented in descending order. Residues identical to those in Dae are shaded. A. n., *A. nidulans*; P. c., *P. chrysogenum*; R. n., *R. norvegicus*; M. s., *M. soehngenii*.

guanines and flanked by multiple adenines. These promoter and ribosome-binding site sequences are similar to those reported for gram-positive bacteria and genes isolated from *L. casei* (2-4, 40, 49, 53, 65). The use of *L. casei* promoters and ribosome-binding site sequences by *E. coli* is consistent with the observed Dae activity in noninduced *E. coli* lysates of λ ZapII(-)DE24-infected cultures. Initiation of transcription in *E. coli* by *L. casei* promoters has been reported (13, 38, 40, 50, 65). The organization of these elements relative to the *dae* gene indicates the presence of a transcriptional unit (Fig. 4).

The locations of the ORFs on this transcriptional unit are such that the termination codon of *dae* overlaps the initiation codon of the adjacent ORF2 by 1 nt. The overlap of termination and initiation codons has been reported for the polycistronic tryptophan genes and folypoly- γ -glutamate synthetase genes from *L. casei* (50, 65). The lack of a putative rho-independent RNA polymerase termination sequence on the 2.4-kb fragment also indicates that a polycistronic message may be transcribed from this operon. It is conceivable that the contiguous ORF2 may be a membrane protein involved in the incorporation of D-alanine into LTA.

The mechanism of activation and transfer of activated D-alanine to LTA is unknown. Computer searches have provided insights into this mechanism by identifying 19 protein domains with similar amino acid sequences and enzyme function. These enzymes have one mechanistic feature in common, amino or fatty acid activation via an acyl adenylate intermediate. There is one motif, GXXGXPXK, common to each of these enzymes which may be essential for the formation of the acyl adenylate. As suggested by Hori et al. (29), the consensus sequence (SGTTGXPKG) in this region may represent a new class of phosphate-binding loops (60). The essential nature of this region has also been recognized by other investigators (16, 29, 46, 61).

Many of these homologous enzymes are also known to catalyze nonribosomal peptide condensation via a thiolester intermediate (32). This group of enzyme domains has two additional motifs in common with Dae, GRXDXQXXKX XGXRXE and PX₂PX₂KXDX₃L. The similarity between these enzymes and Dae suggests that Dae may catalyze the transfer of D-alanine via a thiolester intermediate. This mechanism is supported by the observation that Dae was inhibited 50% by 1.7 μ M *p*-hydroxymercuribenzoate (42). Since Dae has only one cysteine residue (Cys-269), it is proposed that this residue may form a thiolester intermediate with D-alanine. Because regions II and III are present in thiolester template enzymes but not in the acid-thiol ligases, the participation of these regions in the formation of the thiolester intermediate is suggested. All thiotemplate activation sites on multienzymes are able to accept 2'-dATP as a substrate (32). The ability of Dae to utilize this substrate (data not shown) further supports the classification of Dae with the thiotemplate enzymes.

One of the proposed functions of LTA and wall teichoic acid in gram-positive organisms is metal ion chelation, specifically the binding of Mg²⁺ (5, 30). Dae shows homology to several enzymes that participate in the biosynthesis of high-affinity metal ion chelation systems. For example, EntF and EntE from *E. coli* participate in the biosynthesis of enterobactin, a cyclic triester of 2,3-dihydroxybenzoyl serine (7). EntF, which catalyzes the formation of an acyl adenylate intermediate, has an essential sulfhydryl and 4-phosphopantetheine (55). Recently, Rusnak et al. (55) have suggested that EntF can be placed within the family of enzymes involved in the biosynthesis of cyclic peptides

through a thiotemplate mechanism. AngR from *V. anguillarum* represents another protein homologous to Dae that also participates in the biosynthesis of a high-affinity metal ion chelator. This protein is essential for the biosynthesis of anguibactin, a siderophore analogous to enterobactin (57). The sequence homology of Dae with proteins essential for Fe³⁺ chelation suggests that a functional relationship between these metal ion-scavenging systems and D-alanyl-LTA may exist. Thus, one of the proposed functions of D-alanyl-LTA, binding of Mg²⁺ for enzyme function, is sustained by these correlations.

Although LTAs are a well-characterized class of amphiphilic polymers (63), their precise biological function has yet to be elucidated. Mutants that are deficient in D-alanyl ester content of LTA show aberrant morphology and defective cell separation (52). Fischer et al. (22) proposed that the D-alanine ester content of LTA may function in the regulation of autolysins. Since the regulation of autolytic activity is one of the determinants of cell shape, a genetically defined system for modulating the D-alanyl ester content of LTA is critical for establishing the role of this macroamphiphile in the morphogenetic program of this gram-positive organism.

ACKNOWLEDGMENTS

This investigation was supported in part by Public Health Service grant AI-04615 from the National Institute of Allergy and Infectious Diseases.

We thank Bruce Chassy and Neil Welker for helpful discussions, Jacqueline White for isolation and identification of bacteriophage λ ZapII(+)DE24, and Ka-Leung Ngai and Wei Luo of the Northwestern University Biotechnology Facility for N-terminal amino acid sequencing and oligonucleotide synthesis. We are also grateful for valuable technical advice provided by Aseem Ansari, Karen Ector, and Sergei Gryaznov.

REFERENCES

1. Akhrem, A. A., and A. P. Drozhdenyuk. 1989. Calcium tartrate gel. *Anal. Biochem.* 179:86-89.
2. Alpert, C.-A., and B. M. Chassy. 1988. Molecular cloning and nucleotide sequence of the factor III^{lac} gene of *Lactobacillus casei*. *Gene* 62:277-288.
3. Alpert, C.-A., and B. M. Chassy. 1990. Molecular cloning and DNA sequence of *lacE*, the gene encoding the lactose-specific enzyme II of the phosphotransferase system of *Lactobacillus casei*. Evidence that a cysteine residue is essential for sugar phosphorylation. *J. Biol. Chem.* 265:22561-22568.
4. Andrews, J., G. M. Clore, R. W. Davies, A. M. Gronenborn, B. Gronenborn, D. Calderon, P. C. Papadopoulos, S. Schäfer, P. F. G. Sims, and R. Stancombe. 1985. Nucleotide sequence of the dihydrofolate reductase gene of methotrexate-resistant *Lactobacillus casei*. *Gene* 35:217-222.
5. Archibald, A. R., J. Baddiley, and S. Heptinstall. 1973. The alanine ester content and magnesium binding capacity of walls of *Staphylococcus aureus* H grown at different pH values. *Biochim. Biophys. Acta* 291:629-634.
6. Baddiley, J., and F. C. Neuhaus. 1960. The enzymic activation of D-alanine. *Biochem. J.* 75:579-587.
7. Bagg, A., and J. B. Neillands. 1987. Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiol. Rev.* 51:509-518.
8. Berger, S. L., and A. R. Kimmel. 1987. Guide to molecular cloning techniques. *Methods Enzymol.* 152:3-755.
9. Brautigan, V. M., W. C. Childs III, and F. C. Neuhaus. 1981. Biosynthesis of D-alanyl-lipoteichoic acid in *Lactobacillus casei*: D-alanyl-lipophilic compounds as intermediates. *J. Bacteriol.* 146:239-250.
10. Chassy, B. M., and A. Giuffrida. 1980. Method for the lysis of gram-positive, asporogenous bacteria with lysozyme. *Appl.*

- Environ. Microbiol. **39**:153–158.
11. Childs, W. C., III, and F. C. Neuhaus. 1980. Biosynthesis of D-alanyl-lipoteichoic acid: characterization of ester-linked D-alanine in the in vitro-synthesized product. *J. Bacteriol.* **143**:293–301.
 12. Connerton, I. F., J. R. S. Fincham, R. A. Sandeman, and M. J. Hynes. 1990. Comparison and cross-species expression of the acetyl-CoA synthetase genes of the Ascomycete fungi, *Aspergillus nidulans* and *Neurospora crassa*. *Mol. Microbiol.* **4**:451–460.
 13. Davies, R. W., and A. M. Gronenborn. 1982. Molecular cloning of the gene for dihydrofolate reductase from *Lactobacillus casei*. *Gene* **17**:229–233.
 14. Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
 15. Díez, B., S. Gutiérrez, J. L. Barredo, P. van Solingen, L. H. M. van der Voort, and J. F. Martín. 1990. The clusters of penicillin biosynthetic genes. Identification and characterization of the *pcbAB* gene encoding the α -aminoadipyl-cysteinyl-valine synthetase and linkage to the *pcbC* and *penDE* genes. *J. Biol. Chem.* **265**:16358–16365.
 16. Eggen, R. I. L., A. C. M. Geerling, A. B. P. Boshoven, and W. M. de Vos. 1991. Cloning, sequence analysis, and functional expression of the acetyl coenzyme A synthetase gene from *Methanoxanthus soehngenii* in *Escherichia coli*. *J. Bacteriol.* **173**:6383–6389.
 17. Farrell, D. H., P. Mikesell, L. A. Actis, and J. H. Crosa. 1990. A regulatory gene, *angR*, of the iron uptake system of *Vibrio anguillarum*: similarity with phage P22 *cro* and regulation by iron. *Gene* **86**:45–51.
 18. Fischer, W. 1988. Physiology of lipoteichoic acids in bacteria. *Adv. Microb. Physiol.* **29**:233–302.
 19. Fischer, W. 1990. Bacterial phosphoglycolipids and lipoteichoic acids, p. 123–234. In M. Kates (ed.), *Handbook of lipid research: glycolipids, phospholipids, and sulfoglycolipids*. Plenum Publishing Corp., New York.
 20. Fischer, W., H. U. Koch, P. Rösel, and F. Fiedler. 1980. Alanine ester-containing native lipoteichoic acids do not act as lipoteichoic acid carrier: isolation, structural and functional characterization. *J. Biol. Chem.* **255**:4557–4562.
 21. Fischer, W., and P. Rösel. 1980. The alanine ester substitution of lipoteichoic acid (LTA) in *Staphylococcus aureus*. *FEBS Lett.* **119**:224–226.
 22. Fischer, W., P. Rösel, and H. U. Koch. 1981. Effect of alanine ester substitution and other structural features of lipoteichoic acids on their inhibitory activity against autolysins of *Staphylococcus aureus*. *J. Bacteriol.* **146**:467–475.
 23. Fleig, U. N., R. D. Pridmore, and P. Philippsen. 1986. Construction of *LYS2* cartridges for use in genetic manipulations of *Saccharomyces cerevisiae*. *Gene* **46**:237–245.
 24. Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557–580.
 25. Heptinstall, S., A. R. Archibald, and J. Baddiley. 1970. Teichoic acids and membrane function in bacteria. *Nature (London)* **225**:519–521.
 26. Hewick, R. M., M. W. Hunkapiller, L. E. Hood, and W. J. Dreyer. 1981. A gas-liquid solid phase peptide and protein sequencer. *J. Biol. Chem.* **256**:7990–7997.
 27. Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**:193–197.
 28. Hori, K., Y. Yamamoto, T. Minetoki, T. Kurotsu, M. Kanda, S. Miura, K. Okamura, J. Furuyama, and Y. Saito. 1989. Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene. *J. Biochem.* **106**:639–645.
 29. Hori, K., Y. Yamamoto, K. Tokita, F. Saito, T. Kurotsu, M. Kanda, K. Okamura, J. Furuyama, and Y. Saito. 1991. The nucleotide sequence for a proline-activating domain of gramicidin S synthetase 2 gene from *Bacillus brevis*. *J. Biochem.* **110**:111–119.
 30. Hughes, A. H., I. C. Hancock, and J. Baddiley. 1973. The function of teichoic acids in cation control in bacterial membranes. *Biochem. J.* **132**:83–93.
 31. Keleman, M. V., and J. Baddiley. 1961. Structure of the intracellular glycerol teichoic acid from *Lactobacillus casei* A.T.C.C. 7469. *Biochem. J.* **80**:246–254.
 32. Kleinkauf, H., and H. von Döhren. 1987. Biosynthesis of peptide antibiotics. *Annu. Rev. Microbiol.* **41**:259–289.
 33. Koch, H. U., W. Fischer, and F. Fiedler. 1982. Influence of alanine ester and glycosyl substitution on the lipoteichoic acid carrier activity of lipoteichoic acids. *J. Biol. Chem.* **257**:9473–9479.
 34. Krätzschmar, J., M. Krause, and M. A. Marahiel. 1989. Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases. *J. Bacteriol.* **171**:5422–5429.
 35. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105–132.
 36. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
 37. Lambert, P. A., I. C. Hancock, and J. Baddiley. 1975. Influence of alanyl ester residues on the binding of magnesium ions to teichoic acid. *Biochem. J.* **151**:671–676.
 38. Lee, L.-J., J. B. Hansen, E. K. Jaguszyn-Krynicka, and B. M. Chassy. 1982. Cloning and expression of the β -D-phosphogalactoside galactohydrolase gene from *Lactobacillus casei* in *Escherichia coli* K-12. *J. Bacteriol.* **152**:1138–1146.
 39. Leong-Morgenthaler, P., M. C. Zwahlen, and H. Hottinger. 1991. Lactose metabolism in *Lactobacillus bulgaricus*: analysis of the primary structure and expression of the genes involved. *J. Bacteriol.* **173**:1951–1957.
 40. Lerch, H.-P., H. Blöcker, H. Kallwass, J. Hoppe, H. Tsai, and J. Collins. 1989. Cloning, sequencing and expression in *Escherichia coli* of the D-2-hydroxyisocaproate dehydrogenase gene of *Lactobacillus casei*. *Gene* **78**:47–57.
 41. Lerch, H.-P., R. Frank, and J. Collins. 1989. Cloning, sequencing and expression of the L-2-hydroxyisocaproate dehydrogenase-encoding gene of *Lactobacillus confusus* in *Escherichia coli*. *Gene* **83**:263–270.
 42. Linzer, R., and F. C. Neuhaus. 1973. Biosynthesis of membrane teichoic acid: a role for the D-alanine-activating enzyme. *J. Biol. Chem.* **248**:3196–3201.
 43. Liu, J., K. Duncan, and C. T. Walsh. 1989. Nucleotide sequence of a cluster of *Escherichia coli* enterobactin biosynthesis genes: identification of *entA* and purification of its product 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase. *J. Bacteriol.* **171**:791–798.
 44. Lozoya, E., H. Hoffmann, C. Douglas, W. Schulz, D. Scheel, and K. Hahlbrock. 1988. Primary structures and catalytic properties of isozymes encoded by the two 4-coumarate:CoA ligase genes in parsley. *Eur. J. Biochem.* **176**:661–667.
 45. MacArthur, A. E., and A. R. Archibald. 1984. Effect of culture pH on the D-alanine ester content of lipoteichoic acid in *Staphylococcus aureus*. *J. Bacteriol.* **160**:792–793.
 46. MacCabe, A. P., H. van Liempt, H. Palissa, S. E. Unkles, M. B. R. Riach, E. Pfeifer, H. van Döhren, and J. R. Kinghorn. 1991. δ -(L- α -Aminoadipyl)-L-cysteinyl-D-valine synthetase from *Aspergillus nidulans*. Molecular characterization of the *acvA* gene encoding the first enzyme of the penicillin biosynthetic pathway. *J. Biol. Chem.* **266**:12646–12654.
 47. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 48. Masuda, T., H. Tatsumi, and E. Nakano. 1989. Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly, *Luciola cruciata*. *Gene* **77**:265–270.
 49. McLaughlin, J. R., C. L. Murray, and J. C. Rabinowitz. 1981. Unique features in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* β -lactamase gene. *J. Biol. Chem.* **256**:11283–11291.
 50. Natori, Y., Y. Kano, and F. Imamoto. 1990. Nucleotide sequences and genomic constitution of five tryptophan genes of

- Lactobacillus casei*. J. Biochem. 107:248-255.
51. Neuhaus, F. C., R. Linzer, and V. M. Reusch, Jr. 1974. Biosynthesis of membrane teichoic acid: role of the D-alanine activating enzyme and D-alanine:membrane acceptor ligase. Ann. N.Y. Acad. Sci. 235:502-518.
 52. Ntamere, A. S., D. J. Taron, and F. C. Neuhaus. 1987. Assembly of D-alanyl-lipoteichoic acid in *Lactobacillus casei*: mutants deficient in the D-alanyl ester content of this amphiphile. J. Bacteriol. 169:1702-1711.
 53. Porter, E. V., and B. M. Chassy. 1988. Nucleotide sequence of the β -D-phosphogalactoside galactohydrolase gene of *Lactobacillus casei*: comparison to analogous *pbg* genes of other Gram-positive organisms. Gene 62:263-276.
 54. Reusch, V. M., Jr., and F. C. Neuhaus. 1971. D-Alanine: membrane acceptor ligase from *Lactobacillus casei*. J. Biol. Chem. 246:6136-6143.
 55. Rusnak, F., M. Sakaitani, D. Drueckhammer, J. Reichert, and C. T. Walsh. 1991. Biosynthesis of the *Escherichia coli* siderophore enterobactin: sequence of the *entF* gene, expression and purification of EntF, and analysis of covalent phosphopantetheine. Biochemistry 30:2916-2927.
 56. Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350-1354.
 57. Salinas, P. C., M. E. Tolmasky, and J. H. Crosa. 1989. Regulation of the iron uptake system in *Vibrio anguillarum*: evidence for a cooperative effect between two transcriptional activators. Proc. Natl. Acad. Sci. USA 86:3529-3533.
 58. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 59. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
 60. Saraste, M., P. R. Sibbald, and A. Wittinghofer. 1990. The P-loop—a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. 15:430-434.
 61. Smith, D. J., A. J. Earl, and G. Turner. 1990. The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillium chrysogenum* is a 421 073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. EMBO J. 9:2743-2750.
 62. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.
 63. Sutcliffe, I. C., and N. Shaw. 1991. Atypical lipoteichoic acids of gram-positive bacteria. J. Bacteriol. 173:7065-7069.
 64. Suzuki, H., Y. Kawarabayasi, J. Kondo, T. Abe, K. Nishikawa, S. Kimura, T. Hashimoto, and T. Yamamoto. 1990. Structure and regulation of rat long-chain acyl-CoA synthetase. J. Biol. Chem. 265:8681-8685.
 65. Toy, J., and A. L. Bognar. 1990. Cloning and expression of the gene encoding *Lactobacillus casei* folylpoly- γ -glutamate synthetase in *Escherichia coli* and determination of its primary structure. J. Biol. Chem. 265:2492-2499.
 66. Weckermann, R., R. Fürbass, and M. A. Marahiel. 1988. Complete nucleotide sequence of the *tycA* gene coding the tyrocidine synthetase 1 from *Bacillus brevis*. Nucleic Acids Res. 16:11841.