Substrate Recognition Mechanism of the Peptidase Domain of the Quorum-Sensing-Signal−Producing ABC Transporter ComA from *Streptococcus*†

Yatsugu Kotake,‡,§ Seiji Ishii,‡ Takato Yano,‡ Yoji Katsuoka,§ and Hideyuki Hayashi*,‡

Departments of Biochemistry and Urology, Osaka Medical College, Takatsuki, Osaka 569-8686, Japan

Received November 13, 2007; Revised Manuscript Received December 21, 2007

ABSTRACT: ComA of *Streptococcus* is a member of the bacteriocin-associated ABC transporters, which is responsible for both the processing of the propeptide ComC and secretion of the mature quorum-sensing signal. The quorum-sensing system is a bacterial intercellular communication system implicated in various functions including biofilm formation. In this study, the peptidase domains (PEPs) of the ComAs from six species of *Streptococcus* and ComCs from four species were expressed, purified, and characterized to address the mechanism of the substrate recognition of PEP. PEPs specifically cleaved ComCs after the Gly-Gly site in all the PEP−ComC combinations examined. The N-terminal leader region of ComC was found to form an amphiphilic α-helix structure upon binding to the PEP. Furthermore, mutagenesis studies revealed that four conserved hydrophobic residues in this leader region of ComC extending from −15 to −4 positions are critical in the interaction with PEP. Together with the double glycine motif, these structural features of ComC would explain the strict substrate specificity of the PEP.

The genus *Streptococcus* is Gram-positive bacteria and comprises a wide variety of pathogenic and commensal species. They are present on the skin and the mucous membranes of the oral cavity or the respiratory and genitourinary tracts of humans and animals (*1*). Some of the commensal species are known to be opportunistic pathogens. Notably, oral streptococci, such as *Streptocuccus mutans*, which attach to the tooth surface as dental plaque, not only are cariogenic but also occasionally enter the human circulatory system from the gingival pocket and cause life-threatening infective endocarditis by forming a biofilm on the native or prosthetic heart valves (*2*). It is postulated that the formation of this biofilm and its inherent resistance to antibiotics are a main cause of persistent and chronic nature of this infection (*3*). The biofilm formation of *Streptococcus* is thought to be regulated by the quorum-sensing system like many other Gram-positive and Gram-negative bacteria (*4*).

Quorum sensing is a way that bacteria communicate with each other via pheromone molecules in order to properly respond to growth conditions, coordinate group behaviors, and successfully survive as a "community" (*5*). It is believed that bacterial cells may be continuously releasing inherent pheromones into their surrounding environment. As the population density of a bacterial species increases, its pheromone concentration also rises, reaching the threshold necessary to activate either the cell surface or intracellular receptors. The signal pathway then alters the gene expressions in the target bacteria (*6*, *7*). Many bacteria are known to regulate diverse physiological processes other than the activation of biofilm formation through this system. These include bioluminescence in *Vibrio fischeri* (*8*), sporulation in *Bacillus subtilus* (*9*), virulence factor expression in *Staphylococcus aureus* (*10*), antibiotics production in *Lactococcus lactis* (*10*), and competence for genetic transformation in *Streptococcus pneumoniae*, *Streptococcus gordonii*, and *S. mutans* (*4*). Although several quorum-sensing systems are known, the two most thoroughly described systems are the acyl-homoserine lactone system of Gram-negative bacteria and the peptide-based signaling system of Gram-positive bacteria (*11*).

In *S. pneumoniae*, the competence-stimulating peptide (CSP[1]) functions as a pheromone in the quorum-sensing signal pathway. CSP is cleaved from the precursor ComC peptide and concomitantly exported by ComA with help of an accessory protein, ComB. The accumulated CSP binds to the cell surface receptor ComD, which subsequently phosphorlylates ComE by its histidine kinase activity and thereby induces transcription of genes such as *comX* and *comW*, resulting in DNA uptake (competence) and recombination (*4*, *12*). Homologues of these *com* genes are found in other *Streptococcus* species and have been observed to be involved in the biofilm formation in *S. gordonii* (*13*) or *S. mutans* (*14*). Compared to the well-characterized acyl-homoserine lactone system of Gram-negative bacteria, especially that of *Pseudomonas aeruginosa* (*15*), analysis of the peptide-based signaling system of Gram-positive bacteria, including the CSP system of *Streptococcus*, has

† This work was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Category C Grant 19570141 (to S.I.)) and a grant from the Osaka Medical Research Foundation for Incurable Diseases (to T.Y.).

* To whom correspondence should be addressed: Dept. of Biochemistry, Osaka Medical College, 2-7 Daigakumachi, Takatuski, Osaka 569-8686, Japan. Tel: +81-72-683-1221. Fax: +81-72-684-6516. E-mail: hayashi@art.osaka-med.ac.jp.

‡ Department of Biochemistry.

§ Department of Urology.

[1] Abbreviations: CD, circular dichroism; CSP, competence-stimulating peptide; DTT, dithiothreitol; HPLC, high performance liquid chromatography; IPTG, isopropyl-*β*-D-thiogalactopyranoside; PEP, peptidase domain of streptococcal ComA; TFE, trifluoroethanol.

10.1021/bi702253n CCC: $40.75
Published on Web 01/31/2008

been hindered, mainly because the important components of this system are membrane proteins.

ComA, a key molecule essential for the first step of the quorum-sensing system of *Streptococcus*, is a member of a family of bacteriocin-associated ATP-binding cassette (ABC) transporters, which are composed of three domains (*16*), i.e., an N-terminal domain possessing the peptidase activity, a transmembrane domain consisting of six membrane-spanning segments, and a C-terminal ATP-binding domain located on the cytoplasmic face of the membrane. The peptidase domains of this family are thought to cleave their cognate propeptides after the consensus Gly-Gly motif. To date, peptidase domains of the family members, LagD, a transporter of lactococcin G in *Lactococcus lactis* (*16*), and CvaB, a transporter of colicin V in *Escherichia coli* (*17*), were confirmed to have proteolytic activity, for which a cysteine residue is critical. However, both peptidase domains were mainly expressed as inclusion bodies in *E. coli*, and their biochemical characterizations have remained limited. Recently, we succeeded in the heterologous overexpression of the peptidase domain (PEP) of *S. pneumoniae* ComA as a soluble protein in *E. coli*, enabling us to do a detailed biochemical analysis. The purified PEP exhibited an efficient proteolytic activity for the natural substrate ComC, which was confirmed to be cleaved after the Gly-Gly site, and the kinetic parameters of PEP for ComC were determined (*18*).

Here, we first cloned, expressed, and purified six PEPs and four ComCs from the species of *Streptococcus* and carried out catalytic and structural analyses. Then, based on these results and sequence comparison of the ComCs, a series of mutations were introduced into the *S. pneumoniae* ComC to further elucidate the substrate recognition mechanism of PEP.

EXPERIMENTAL PROCEDURES

Cloning and Plasmid Constructions of Streptococcus PEPs. In a previous report, we cloned and characterized the *S. pneumoniae* PEP (*18*). For the present study, the following eight *Streptococcus* species (two strains for *S. gordonii*) were obtained from the American Type Culture Collection, which were supposed to possess the *com* system (*19*): *S. mitis* (strain ATCC No. 49456), *S. oralis* (35037), *S. gordonii* (10558), *S. gordonii* (33399), *S. sanguis* (10556), *S. cristatus* (51100), *S. anguinosus* (33397), *S. mutans* (700610), and *S. thermophilus* (BAA-250). For *S. mutans* and *S. thermophilus*, the entire genome nucleotide sequences were available (GenBank AE014133 and CP000023, respectively). Each genome had a gene annotated as *com*A (gene ID 1027866 for the *S. mutans comA* and 3165209 for the *S. thermophilus comA*), and the PEP-encoding regions were PCR-amplified and cloned. For *S. mutans*, the second *comA*-like open reading frame, which has a 70% homology in the amino acid sequence with the *comA* gene but is located outside the *comAB* operon (*20*), was also cloned. The N-terminal hydrophobic extensions of 44 and 47 amino acids, respectively, of *S. mutans comA* and *comA*-like genes were considered to promote inclusion body formation, and were truncated from the PEP domains.

To clone the PEP regions from other *Streptococcus*, the 5′ PCR primers were designed according to a nucleotide sequence called the Box element (*21*), located 17-bp upstream from *comA*. The 3′ PCR primers were designed according to the conserved sequences in the ATP-binding domain found in the sequence comparison of the *S. pneumoniae* and *S. mutans comA* genes. The correct DNA fragments were obtained only from *S. mitis*. Then, a 5′ PCR primer with the nucleotide sequence identical to the 5′-end of the coding region of *S. pneumoniae comA* and the above 3′ primers were used to amplify the *S. oralis* PEP-coding region. The PEP regions of other *Streptococcus* species could not be obtained by these simple PCR-based cloning strategies.

For the convenience of the purification, a His_6-tag sequence was attached to the C-terminal end of each PEP. The DNA-manipulation procedures and the sequences of the PCR primers employed for the cloning and plasmid constructions are described in Supporting Information.

Cloning and Plasmid Constructions of Streptococcus ComCs. The *comCDE* operon was extensively searched for in the *Streptococcus* species (*19*), and the candidate *comC* genes were identified in the species described in this study. The nucleotide sequences for the *comC* genes were found in the GenBank, the PCR primers were designed so as to incorporate a His_6-tag sequence at the N-terminal end, and the *comC* genes of the eight species (2 strains for *S. gordonii*) were cloned. The DNA-manipulation procedures and the sequences of the PCR primers employed for the cloning and plasmid constructions are described in the Supporting Information.

Site-Directed Mutagenesis of the S. pneumoniae ComC. Each of the amino acid residues at the positions −15, −14, −12, −10, −7, −6, −4, −3, +1, and +3 of the *S. pneumoniae* ComC was replaced with alanine (for the numbering of the residues of ComC, see Figure 5A). And, the two residues Phe−15-Val−14 were mutated to Val−15-Phe−14. Mutagenesis was done using a QuikChange Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions with pSPC1 (*18*) as the template. The primer pairs used for the mutagenesis are listed in the Supporting Information.

Protein Expression and Purification. PEPs and ComCs were expressed and purified essentially as previously described (*18*). Briefly, for expression of the PEPs, an *E. coli* strain, BL21 (DE3) pLysS, carrying each expression plasmid was grown and induced by 0.2 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG) for 5 h at 30 °C for PPEP, TPEP, and MuPEP2 or for 2 h at 37 °C for MiPEP, OPEP*, and MuPEP1 (see below for the abbreviations of the PEPs). The PEPs were purified with the His•Bind Resin (Novagen, Madison, WI) and dialyzed against a buffer containing 20 mM Tris-HCl, 200 mM ammonium sulfate, and 2 mM dithiothreitol (DTT), pH 7.0.

For expression of the ComCs, an *E. coli* strain, JM109 (DE3) pLysS, carrying each expression plasmid was grown and induced by 0.2 mM IPTG for 1 h at 37 °C for PComC or 2 h at 37 °C for the other ComCs. The ComCs were purified with the His•Bind Resin followed by chromatography with a HiTrapQHP column connected to an ÄKTA fast protein liquid chromatography system (GE Healthcare, Piscataway, NJ).

The concentrations of the PEPs and ComCs were spectrophotometrically determined based on the numbers of the aromatic amino acid residues (*18*).

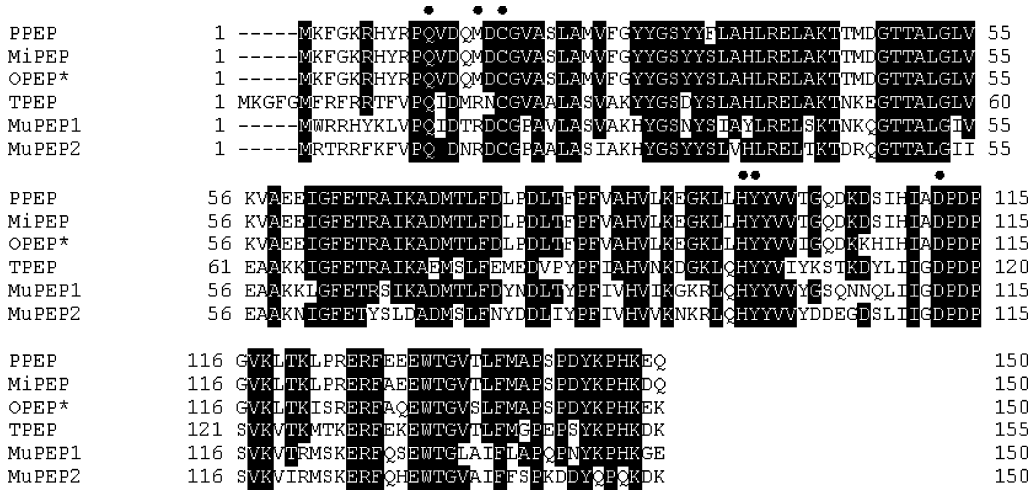


FIGURE 1: Amino acid sequence alignment of PEPs from different species of *Streptococcus*. The amino acid residues that are identical in at least four of the compared PEP sequences are shaded black. PPEP, *S. pneumoniae* PEP; MiPEP, *S. mitis* PEP; OPEP*, *S. oralis* PEP; TPEP, *S. thermophilus* PEP; MuPEP1, *S. mutans* PEP1; MuPEP2, *S. mutans* PEP2. The eight N-terminal residues of OPEP* might be different from the native *S. oralis* PEP (see text for details). MuPEP1 is a PEP domain of *S. mutans* ComA, and MuPEP2 is that of the *S. mutans* ComA-like protein, and the N-terminal highly hydrophobic sequences (44 and 47 amino acids, respectively) of the native proteins have been removed in both MuPEPs (see text). Dots indicate six amino acid residues that are supposed to be important for the catalysis.

Circular Dichroism (CD) Measurements. The CD spectra were recorded using a Jasco spectropolarimeter, model J-720WI (Jasco, Tokyo, Japan) equipped with a thermocontroller using a 0.1 cm light-path sample cell. The thermal denaturation profiles of the PEPs were measured by monitoring the CD at 222 nm while heating the samples at the rate of 1 °C/min. The buffers and protein concentrations are described in the legends of Figures 3 and 4.

Determination of Kinetic Parameters. The PEP activity was assayed in a 100 μL reaction mixture containing 50 mM Tris-HCl, pH 7.0, 150 mM ammonium sulfate, and various concentrations of a substrate. The high performance liquid chromatography (HPLC) analysis was performed as previously described (*18*). Briefly, the reaction mixtures were loaded onto a Waters μBondasphere C_8 reversed-phase column (Waters, Milford, MA) connected to a Beckman System Gold HPLC system (Beckman, Fullerton, CA) and the peptides were separated on a linear gradient from 10 to 55% (v/v) acetonitrile containing 0.1% trifluoroacetic acid over 10 min at the flow rate of 1 mL/min at ambient temperature. Typically, less than 10% of the substrates was consumed in the reactions.

RESULTS

Sequence Comparison, Expression, and Purification of Streptococcus PEPs. Figure 1 shows the primary structures of the *Streptococcus* PEPs analyzed in this study. The PEPs were designated as follows: PPEP for *S. pneumoniae* PEP; MiPEP for *S. mitis* PEP; OPEP* for *S. oralis* PEP; TPEP for *S. thermophilus* PEP; MuPEP1 and MuPEP2 for the two *S. mutans* PEPs. The asterisk of OPEP* denotes that the N-terminal eight residues might contain mutations because the 5′ PCR primer used had the sequence identical to that of PPEP. From the sequence comparisons at the amino acid level, PPEP shows a high homology to MiPEP (98%) and OPEP* (93%), and a moderate homology to TPEP (64%), MuPEP1 (57%), and MuPEP2 (51%). Comparisons of TPEP, MuPEP1, and MuPEP2 show a 60−70% homology among each other. These results are consistent with the phylogenetic relatedness of the *Streptococcus* species based on the *groESL* and *rnpB* genes (*22*, *23*). *S. mitis* is the closest to *S. pneumoniae*, and only three amino acids were different between MiPEP and PPEP. *S. oralis* is also close to *S. pneumoniae*, and OPEP* contains 10 amino acid differences located evenly throughout the 142-residue region of the PEP domain, compared to PPEP. Thus, although OPEP* might be different from the native OPEP, the difference would be at most one in the N-terminal eight residues. Anyway, results for OPEP* in this study should be regarded as those for a PEP that has the primary structure of OPEP*, not necessarily those for the *S. oralis* PEP. TPEP has a five-residue extension at the N-terminus, but the second methionine might be the initiation codon in the native *S. thermophilus* PEP. As mentioned above, the hydrophobic N-terminal regions were removed in MuPEP1 and MuPEP2.

All PEPs could be expressed as soluble proteins in *E. coli*. As in the case of PPEP (*18*), the ratio of soluble to insoluble protein amounts was higher when expressed at 30 °C than at 37 °C for TPEP or MuPEP2 (data not shown). Such an improvement was not observed for MiPEP, OPEP*, and MuPEP1, and these PEPs were expressed at 37 °C. The PEPs were purified to apparent homogeneity as judged on the SDS−polyacrylamide gels. No nonspecific peptidase activities were detected in the purified PEP samples. Routinely, the yields of the purified proteins from 100 mL of the bacterial culture were ∼2.0 mg for PPEP; ∼3.5 mg for MiPEP; ∼6.0 mg for OPEP*; ∼0.052 mg for TPEP; ∼0.80 mg for MuPEP1; and ∼0.060 mg for MuPEP2.

Expression and Purification of Streptococcus ComCs. The expression in *E. coli* was done under the conditions used for the *S. pneumoniae* ComC (*18*) and checked on SDS−polyacrylamide gels. *S. cristatus* ComC, *S. gordonii* (10558) ComC, and *S. gordonii* (33399) ComC were expressed as soluble proteins, whereas *S. anginosus* ComC was almost exclusively expressed as an inclusion body and *S. mitis* ComC, *S. oralis* ComC, and *S. mutans* ComCs were not accumulated at all in the *E. coli*. For the latter four ComCs, various expression conditions were tried, but purified ComCs

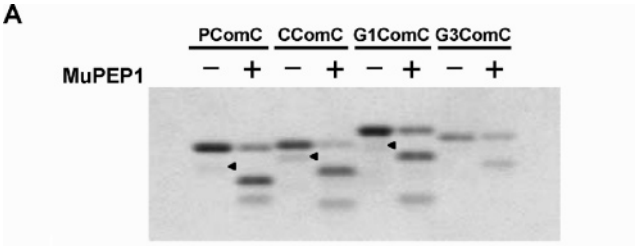
could not be obtained for analysis. The ComCs analyzed in this study were designated as follows: PComC for *S. pneumoniae* ComC; CComC for *S. cristatus* ComC; G1ComC for *S. gordonii* (10558) ComC; and G3ComC for *S. gordonii* (33399) ComC. These ComCs were purified by a two-step procedure (*18*) to apparent homogeneity as judged on SDS−polyacrylamide gels. Routinely, ∼1 mg of the purified protein was obtained from 300 mL of the bacterial culture except for G3ComC (<0.2 mg).

Stability of PEPs. The CD spectra of OPEP* in the far-UV region of 200−300 nm were measured at 30 °C (line 1) and 75 °C (line 2) in the presence of 0.15 M ammonium sulfate (shown in Supporting Information (Figure S1A)). The spectrum at 30 °C is typical of the folded, native state of a protein. At 75 °C, the spectrum showed drastic changes, and the solution became turbid, indicating that OPEP* was denatured. The denaturation was an irreversible process. The other five PEPs showed similar CD-spectral characteristics.

To monitor the heat-denaturation process, the CD value at 222 nm was monitored while the PEP solution was heated from 30 °C to 75 °C (for PPEP, MiPEP, and OPEP*) or to 90 °C (for TPEP, MuPEP1, and MuPEP2) (Figure S1B). These temperature scans of the PEPs showed denaturation curves with apparent T_m values of 47.9 °C for PPEP, 49.9 °C for MuPEP2, 55.6 °C for TPEP, 62.0 °C for MiPEP, 63.2 °C for OPEP*, and 66.7 °C for MuPEP1. The denaturation process of MiPEP, OPEP*, or MuPEP1 was highly cooperative, while PPEP, TPEP, and MuPEP2, which are not so stable as the former three PEPs, gradually denatured. The results indicate that all the PEPs are stable under the conditions employed in the following experiments.

Specific Digestion of ComCs by PEPs. The digestions of PComC, CComC, G1ComC, and G3ComC by MuPEP1 are shown in Figure 2A. Each ComC was converted into two smaller peptides, suggesting that the cleavage is specific. To verify the cleavage site of each ComC, the reaction products were separated on a C8 HPLC column, and the smaller-peak product, which is the C-terminal peptide from ComC, was collected, and its N-terminal sequence was determined: Asp-Leu-Arg-Asn-Ile for CComC, Asp-Ile-Arg-His-Arg for G1ComC, and Ser-Gln-Lys-Gly-Val for G3ComC. These sequences correspond with the N-terminal sequences of the CSPs of the respective *Streptococcus* species (Figure 5A), indicating all the ComCs were cleaved after the Gly-Gly motif. As for the PComC products, the retention times of the two peaks in the HPLC analysis of the MuPEP1−PComC reaction products were identical to those of the PPEP−PComC reaction products. It was confirmed that, by protein sequencing and mass-spectrometric analyses, PPEP correctly cleaves PComC after the Gly-Gly motif (*18*). Thus the cleavage site of PComC by MuPEP1 is also after the Gly-Gly motif.

The relative activities for all combinations of the PEPs and ComCs were roughly estimated at the fixed ComC concentration of 50 μM. The rate of substrate consumption was measured by the HPLC analysis, and each rate was compared to that of PPEP for PComC, the only cognate enzyme−substrate combination examined in this study (Figure 2B). Although the relative activities of the PEPs for the ComCs were varied, significant digestions were observed for all the enzyme−substrate combinations. PComC was the best substrate for PPEP among the ComCs tested. The



B

Relative activity (%)

	PComC	CComC	G1ComC	G3ComC
PPEP	100	26	1.8	27
MiPEP	120	22	3.1	27
OPEP*	110	23	2.3	20
TPEP	5.5	41	1.4	1.5
MuPEP1	32	130	4.8	2.9
MuPEP2	2.1	0.7	2.3	0.7

FIGURE 2: Digestions of ComCs by PEPs. (A) Specific digestions of ComCs by MuPEP1 shown by an SDS−PAGE analysis. Each ComC was incubated with or without 2 μM MuPEP1 in 100 μL of 50 mM Tris-HCl, 150 mM ammonium sulfate, 0.2 mM DTT, pH 7.0, at 25 °C for 16 h. The reaction mixtures (2.5 μL each) were separated on a 16.5% SDS−polyacrylamide gel with the Tris-tricine buffer system (*27*). Because ∼40-amino acid ComCs are stained very differently from each other, the concentrations of the ComCs in the reaction mixtures were adjusted so as to give rise to a similar visibility for each ComC band on the gel: 250 μM PComC, 250 μM CComC, 50 μM G1ComC, and 50 μM G3ComC. Arrowheads show contaminants of the purified ComCs, which are probably partially digested ComCs during the expression in *E. coli*. The smaller digestion product of G3ComC cannot be seen in this image, but was clearly seen during the destaining process. (B) Relative activities of PEPs for ComCs determined at a fixed substrate concentration of 50 μM. The experiments were carried out at 25 °C in 50 mM Tris-HCl, 150 mM ammonium sulfate, and 0.2 mM DTT, pH 7.0. The reaction mixtures were loaded onto a C_8 reversed-phase column connected to an HPLC system as described under Experimental Procedures. Relative activities were determined by comparing the initial rate of substrate consumption for each of the PEP−ComC pairs to that for the PPEP−PComC pair.

activities of PPEP for CComC and G3ComC were lower than that for PComC, and the activity for G1ComC was even lower. The profiles of the relative activities of MiPEP and OPEP* are similar to that of PPEP. TPEP and MuPEP1 prefer CComC to PComC and showed very low activities for G1ComC and G3ComC. MuPEP2 had very low activities for all the ComCs.

The retention times in the HPLC analysis of the digested products from each ComC were identical for all the reactions with the six PEPs: 7.5 and 8.8 min for PComC; 8.2 and 8.6 min for CComC; 8.2 and 9.2 min for G1ComC; and 8.5 and 9.0 min for G3ComC. These results indicated that ComCs were specifically cleaved after the Gly-Gly motif in all the 24 reactions in Figure 2B.

Kinetic Parameters of PPEP and MuPEP1 for ComCs. The kinetic parameters of PPEP and MuPEP1 for PComC, G1ComC, and CComC were determined. These two PEPs were selected because they showed different patterns of relative activities in Figure 2B. G3ComC was not included because of its low expression level and the large protein amount required for this analysis. The reaction mixtures were separated on a C_8 column, and the enzymatic activities were evaluated by measuring the initial rates of product formations at various ComC concentrations. All the reactions conformed very well to the Michaelis−Menten mechanism (typical

Table 1: Kinetic Parameters of PPEP and MuPEP1 for ComCs[a]

PEPs	ComCs	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
PPEP	PComC	1.4 ± 0.1	86 ± 6	270
	G1ComC	0.016[b]	9.4 ± 0.6	28
	CComC	0.045 ± 0.003	9.1 ± 1.8	82
MuPEP1	PComC	0.87 ± 0.04	740 ± 50	20
	G1ComC	0.050 ± 0.002	38 ± 4	22
	CComC	0.57 ± 0.03	61 ± 8	160

[a] The experiments were carried out at 25 °C in 50 mM Tris-HCl, 150 mM ammonium sulfate, and 0.2 mM DTT, pH 7.0, as described in Experimental Procedures. The concentrations of ComCs were 5−200 μM except for PComC (50−900 μM) in the MuPEP1−PComC pair. The concentrations of PEPs were 0.25−1 μM. [b] The standard deviation was <0.001.

examples are shown in Supporting Information (Figure S2)).

The kinetic parameters are summarized in Table 1. The catalytic efficiencies (k_{cat}/K_m values) for noncognate PEP−ComC pairs ranged from 10% to 60% compared to that for the cognate PPEP−PComC pair. For the pairs with low catalytic efficiencies, i.e., the PPEP−G1ComC, PPEP−CComC, and MuPEP1−G1ComC pairs, the k_{cat} values are very low, compared to that for the PPEP−PComC pair. On the other hand, the K_m values for these pairs are smaller by 2- to 9-fold than that for the cognate pair. The MuPEP1−CComC pair showed kinetic properties almost comparable to the PPEP−PComC pair. For the MuPEP1−PComC pair, the K_m value is 9-fold larger than that for the PPEP−PComC pair, although the k_{cat} value is relatively high, resulting in the reduction of the catalytic efficiency to 10% of that for the PPEP−PComC pair.

CD Spectra of ComCs. To examine if ComCs have propensities to take helical structures, CD spectra were measured in the presence of trifluoroethanol (TFE), a well-known solvent which facilitates helical formation (*24*). The amount of purified G3ComC was too low to be used for this experiment. In the aqueous buffer, PComC, CComC, or G1ComC exhibited a major negative maximum around 200 nm and a minor one around 230 nm (Figure 3A, blue lines), and these spectra did not change significantly after a 10 min incubation at 80 °C. These results indicate that the ComCs have random-coil structures in the aqueous solution. In the buffer containing 35% TFE, the CD spectra of the ComCs were dramatically changed to those with negative maxima around 210 and 220 nm (Figure 3A, red lines), typical of an α-helix structure. The changes were instantaneous, and the spectra did not show any further time-dependent changes. Thus, the ComCs are considered to have an inherent tendency to form an α-helix under appropriate circumstances.

Next, the CD spectra of ComC were measured in the aqueous buffer in the absence and presence of PEP (Figure 3B). The PPEP−CComC pair was chosen because of its smallest K_m value (Table 1). Cys17Ala PPEP, which has lost the catalytic activity and cannot cleave ComC (*18*), was used in this experiment. The spectrum of Cys17Ala PPEP (Figure 3B, green line) is the same as that of the wild-type PPEP. The yellow line in Figure 3B shows the spectrum of the mixture of 20 μM each of Cys17Ala PPEP and CComC. The spectrum below 207 nm could not be measured for this mixture because of large noise derived from the high UV absorption by the proteins. The difference spectrum between these two spectra (yellow minus green) was calculated and

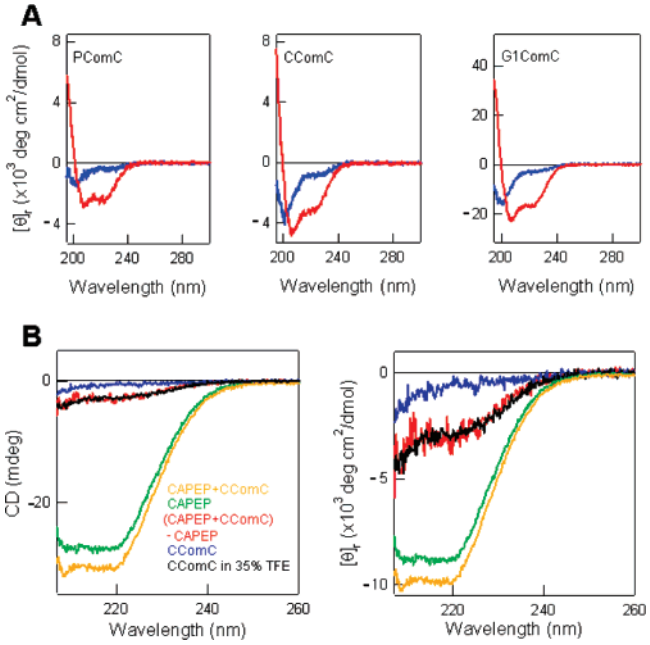


FIGURE 3: Structural transitions of ComCs in the presence of trifluoroethanol (TFE) or PEP. (A) Far-UV CD spectra of ComCs in the region from 195 to 300 nm in 0% (blue line) or 35% (red line) TFE. The CD measurements were done at a protein concentration of 40 μM (for PComC), 60 μM (CComC), or 20 μM (G1ComC) in the buffer solution containing 20 mM sodium phosphate and 150 mM sodium fluoride, pH 7.0, with or without 35% TFE. The spectra were measured at the scan speed of 100 nm/min and accumulated 10 times. (B) Far-UV CD spectra of CComC in the region 207−260 nm in the absence or presence of Cys17Ala PPEP (CAPEP). Data in the region below 207 nm could not be obtained (see text). The ordinate of the left panel is the measured CD value, while that of the right panel indicates the mean residue ellipticity. The CD measurements were done using 20 μM CComC, 20 μM CAPEP, and the mixture of 20 μM each of CComC and CAPEP. The buffer solution contains 30 mM sodium phosphate, 150 mM ammonium sulfate, 15 mM sodium fluoride, and 0.025 mM DTT, pH 7.0. The spectra were measured at the scan speed of 50 nm/min and accumulated 30 times. The spectrum of CAPEP was subtracted from that of the mixture of CAPEP and CComC, and the difference spectrum is shown ((CAPEP + CComC) − CAPEP). The spectrum of 20 μM CComC in the same buffer containing 35% TFE was measured under the same conditions and is overlaid. The mean residue ellipticity of the difference spectrum (red line on the right panel) was calculated based on the assumption that the CD changes derived from the structural changes of CComC.

shown (Figure 3B, red line). The difference spectrum is significantly different from the spectrum of 20 μM CComC (Figure 3B, blue line) but is almost identical to the spectrum of 20 μM CComC measured in 35% TFE (Figure 3B, black line).

Kinetic Parameters of PPEP for Mutant PComCs. Table 2 shows the kinetic parameters of PPEP for a series of mutant PComCs. The catalytic efficiencies are significantly decreased (30−80-fold) for the Phe−15Ala, Val−15Phe−14, Leu−12Ala, Leu−7Ala, and Ile−4Ala PComCs. For these mutant PComCs, the k_{cat} values are decreased 3−10-fold, and the K_m values are increased 10−19-fold. The parameters for the Val−14Ala, Glu−10Ala, Gln−6Ala, and Lys−3Ala PComCs are almost identical to those for the wild-type PComC. The k_{cat} value for the Glu+1Ala PComC is increased 3.3-fold, and the K_m value for the Arg+3Ala PComC is decreased 2.6-fold.

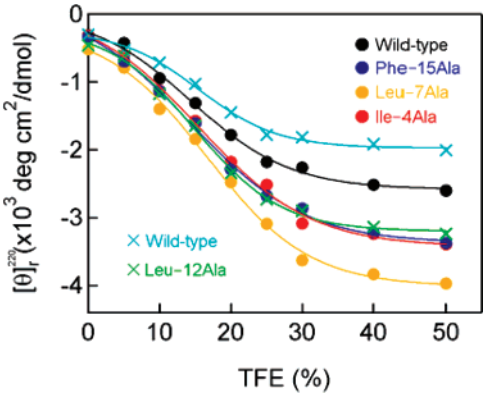


FIGURE 4: TFE-titration of the wild-type and mutant PComCs. Far-UV CD spectra of PComCs in the region from 195 to 300 nm were measured in the presence of 0, 5, 10, 15, 20, 25, 30, 40, and 50% TFE, and the mean residue ellipticities at 220 nm were plotted against the TFE concentrations. The CD measurements were done at a protein concentration of 60 μM except for the Leu−7Ala PComC (45 μM). The buffer solution contained 20 mM sodium phosphate and 150 mM sodium fluoride, pH 7.0, with TFE at the concentrations indicated (○). For the Leu−12Ala PComC, the buffer solution of 20 mM Tris-HCl, pH 7.0, was used because this mutant PComC seemed to be precipitated in the sodium phosphate/sodium fluoride buffer, and the measurements were also done for the wild-type PComC using the Tris-HCl buffer (×). The spectra were measured at the scan speed of 100 nm/min and accumulated 10 times. The solid lines are only for visual aid. Gradual linear changes in the CD values at high TFE concentrations are thought to be solvent effects by TFE, and the transitions themselves are saturated around 35% TFE.

Table 2: Kinetic Parameters of PPEP for the Wild-Type and Mutant PComCs[a]

PComCs	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
wild-type	1.4 ± 0.1	86 ± 6	270
Phe−15Ala	*b*	*b*	2.4[c]
Val−14Ala	0.92 ± 0.10	100 ± 20	150
Val−15Phe−14	0.15 ± 0.02	1400 ± 200	1.8
Lue−12Ala	0.14 ± 0.03	1600 ± 500	1.5
Glu−10Ala	1.2 ± 0.1	76 ± 14	260
Lue−7Ala	0.31 ± 0.05	1600 ± 300	3.2
Gln−6Ala	1.4 ± 0.2	89 ± 20	260
Ile−4Ala	0.50 ± 0.01	960 ± 40	8.7
Lys−3Ala	0.45 ± 0.03	120 ± 20	63
Glu+1Ala	4.6 ± 0.8	100 ± 30	760
Arg+3Ala	0.42 ± 0.03	33 ± 5	210

[a] The experiments were carried out at 25 °C in 50 mM Tris-HCl, 150 mM ammonium sulfate, and 0.2 mM DTT, pH 7.0, as described in Experimental Procedures. The concentrations of mutant PComCs were 10−500 μM, and that of PPEP was 10 μM for the PComCs with k_{cat}/K_m values of <10 M^{-1} s^{-1}. For the other mutant PComCs, the concentrations of PComCs were 10−200 μM, and that of PPEP was 0.25 μM. [b] Reactions did not show saturation kinetics at the examined substrate concentrations. [c] The standard deviation was <0.1.

TFE-Titrations of the CD Spectra of Wild-Type and Mutant PComCs. The CD spectra of the wild-type, Phe−15Ala, Leu−12Ala, Leu−7Ala, and Ile−4Ala PcomCs were measured in the presence of 5−50% TFE, and the CD values at 220 nm were plotted against TFE concentrations (Figure 4). Although the CD values in the presence of saturating amounts of TFE (CD_{max}) are different among these PComCs, the mid-transition points (~15% TFE) are almost unaltered by the mutations. If a mutation impairs the helix formation, the titration curve will be shifted rightward. The CD value of a peptide in this kind of experiment is generally interpreted as indicating the helix content of the peptide (*25*). However, it is not known whether the differences in the CD_{max} values in Figure 4 directly reflect those in the helix contents. The overall configuration of the titration curve or the mid-transition point would be more reliable to infer the helical propensity of each PComC than the CD_{max} value. Thus, it is thought that the propensity of PComC to take a helical structure is not significantly affected by these mutations.

DISCUSSION

Although little is known about the structure−function relationship of PEP, the active-site catalytic triad of Cys17, His96, and Asp112, the oxyanion hole Gln11, and the S_1 and S_2 subsites, Met15 and Tyr99, respectively, were predicted from the sequence comparison of PPEP with other cysteine proteases (*18*). All of these amino acid residues, except for Met15, are completely conserved among the PEPs from species of *Streptococcus* (Figure 1). An arginine residue is substituted for Met15 in TPEP, MuPEP1, and MuPEP2. Both Met and Arg residues have side chains bulky enough to tightly fit the glycine at the −1 position of ComC. From our experience of routinely handling PPEP, the substrate specificity of PPEP is thought to be very high. This is most obviously realized by the finding that the cytoplasmically overproduced soluble PPEP is apparently nontoxic to *E. coli*. PPEP strictly recognizes the double-glycine residues of PComC (*18*), but this alone cannot explain the nontoxicity because the Gly-Gly sequence is not uncommon in *E. coli* proteins. Information on the region far from the active site would be essential to explain the ComC-recognition mechanism of PEP.

MiPEP had only three substitutions compared to PPEP (Phe34 → Ser, Glu128 → Ala, Glu149 → Asp). The difference in only three residues would not allow MiPEP to discriminate MiComC from PComC. At this point, it was predicted that PPEP (or MiPEP) cleaves both PComC and MiComC and that in a reciprocal manner PComC (or MiComC) is cleaved by both PPEP and MiPEP. Further sequence comparisons between other combinations provided no meaningful insights into the substrate recognition mechanism of the PEPs.

The stability of the PEPs was examined before the detailed analyses to ensure that there would be no structural instability which complicates the interpretation of the obtained biochemical parameters. This is particularly important because PEP is a non-natural partial domain excised from ComA. Indeed, PPEP was structurally unstable and easily precipitated during the purification procedures, but salts such as ammonium sulfate were found to stabilize PPEP (*18*). There was a significant correlation between the expression level and thermostability of the PEPs. The expression level of TPEP or MuPEP2, more exactly the soluble/insoluble ratio, was significantly improved when induced at 30 °C, compared to that at 37 °C, as was previously observed for PPEP (*18*). These findings can be explained by the CD results measured in the absence of ammonium sulfate, which showed that these PEPs are stable at 25 °C, but become denatured at around 37 °C (data not shown). It should be noted that only the three amino acid substitutions between PPEP and MiPEP increased the apparent T_m value by 15.3 °C.

From a series of reactions using combinations of PEPs and ComCs, the specific and significant digestion was

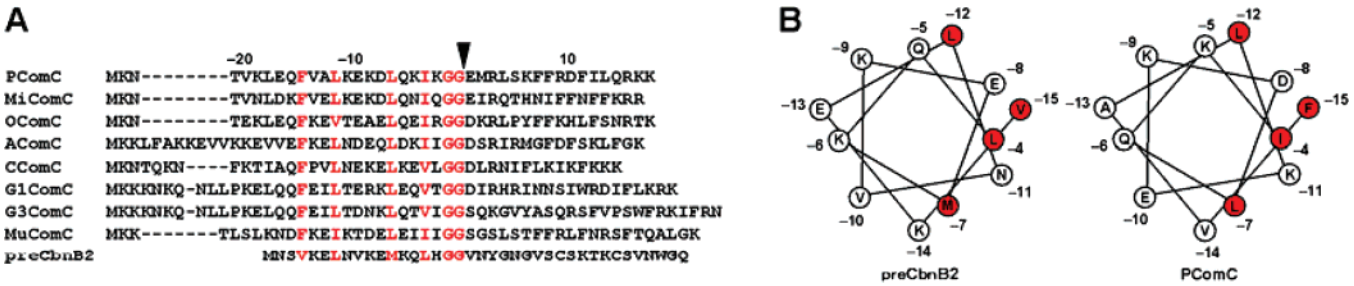


FIGURE 5: N-terminal consensus region of ComCs. (A) Sequence alignment of ComCs from different species of *Streptococcus*. Highly conserved amino acids and the consensus double glycine are colored red. The sequence of the precursor for carnobacterium B2 from *Carnobacterium piscicola*, preCbnB2 (see text), is also shown. The residues 39−66 of preCbnB2 are omitted. The arrowhead indicates a predicted processing site of ComC by PEP. PComC, *S. pneumoniae* ComC; MiComC, *S. mitis* ComC; OComC, *S. oralis* ComC; AComC, *S. anginosus* ComC; CComC, *S. cristatus* ComC; G1ComC, *S. gordonii* (strain ATCC No. 10558) ComC; G3ComC, *S. gordonii* (33399) ComC; MuComC, *S. mutans* ComC. (B) α-Helical wheel representations of the N-terminal regions of preCbnB2 and PComC. Highly conserved hydrophobic amino acids are shaded red.

observed even in the non-natural enzyme−substrate pairs (Figure 2). Based on the profiles of the relative activities for the ComCs, the PEPs are divided into three groups: PPEP, MiPEP, and OPEP*; TPEP and MuPEP1; and MuPEP2. This grouping is well correlated to that postulated from the sequence comparisons of the PEPs. Also, as predicted above, PPEP and MiPEP (and also OPEP*) have comparable activities for PComC. MuPEP2, which is the product of the *S. mutans comA*-like gene, showed low activities for all the ComCs examined. Considering that the *S. mutans comA*-like gene does not exist as the *comAB* operon like the authentic *comA* genes, this *comA*-like gene might be a pseudogene or have already been diverged into a gene with some different biological functions. Next, the kinetic parameters were determined (Table 1) for PPEP and MuPEP1, representatives of the first and second groups, respectively. The most surprising finding was that even poor substrates have high affinities for the PEPs and that the PEPs seem to discriminate one ComC from another by the rate constants rather than the binding affinities.

The promiscuous affinities of the PEPs for the ComCs without apparent proteolytic activities for unrelated proteins lead us to focus on the conserved N-terminal regions of the ComCs. As shown in Figure 5A, the sequences of the leader peptides of the *Streptococcus* ComCs are homologous, whereas those of the C-terminal CSPs are divergent. Besides the Gly-Gly motif, amino acid residues, such as Phe at the position −15, Leu/Val/Ile at −12, Leu at −7, and Ile/Val at −4, are well conserved. These amino acid residues are all hydrophobic and conserved not only among the ComCs but also among the bacteriocin precursors including the precursor of carnobacteriocin B2 (preCbn2) from *Carnobacterium piscicola* (also aligned in Figure 5A). The three-dimensional solution structure of preCbn2 was solved by NMR, and the N-terminal region of this molecule was revealed to form an α-helix structure in the presence of 70% TFE (*24*). The CD measurements showed that ComCs can also take helix structures in 35% TFE (Figure 3A).

In addition, the helix-wheel representations of the leader-peptide sequences of preCbn2 and PComC (Figure 5B) showed that the conserved Phe−15, Leu−12, Leu−7, and Ile−4 of PComC are arranged in the same way as Val−15, Leu−12, Met−7, and Leu−4 of preCbn2 and that these residues are located on one side of the α-helix. The importance of these conserved hydrophobic residues was experimentally confirmed (Table 2). The residue at the position −14, which is highly variable among the ComCs, and those at the positions −10 and −6, which are conserved to some extent, were found to play almost no role in the PPEP−PComC interaction. The residues flanking the double glycine site, i.e., those at the positions −3 and +1, are not important, and neither is the residue at the position +3, although it is conserved among the ComCs. The residues at the positions −11 and −8, which are located in the hydrophobic face of the helix, would not be important in the interaction with PEP, judging from the fact that these residues of G1ComC are different from those of PComC but still the K_m value of PPEP for G1ComC is smaller than that for the cognate PComC (Table 1).

To demonstrate that the helical transition observed in 35% TFE really happens during the PEP-catalyzed reaction process, the CD spectra of CComC were measured in the absence and presence of Cys17Ala PPEP under more physiologically relevant aqueous conditions (Figure 3B). The results clearly show that, upon binding to PEP, CComC undergoes the same transition from a random coil to α-helix as that in 35% TFE.

Based on the above findings, it is postulated that there is a hydrophobic patch on the surface of the PEP protein, which interacts with the hydrophobic face of the N-terminal helix of ComC. Two of the hydrophobic residues, Phe−15 and Leu−7, are strictly conserved among the ComCs, and another, Ile or Val at the position −4, is highly conserved. TFE-titration of the CD spectra (Figure 4) showed that the propensity of PComC to undergo a gross conformational transition was not affected by the mutations of these residues, whereas the catalytic efficiency for PPEP was very sensitive to these mutations. The catalytic efficiency of PPEP for the Val−15Phe−14 PComC, which has the same overall hydrophobicity as the wild-type PComC, was significantly decreased (Table 2). These findings indicate that the postulated hydrophobic patch of PEP might provide specific interactions with these conserved residues of ComC, not just a "greasy" surface to catch the hydrophobic face of the helix. This region of ComC extending from −15 to −4, together with the Gly-Gly motif at the positions −1 and −2, would explain both the high specificity of the PEPs for the ComCs against other protein substrates and the promiscuity of the PEPs for different ComCs. Using *in vivo* assay, it was suggested that similar double-glycine-type leader peptides are responsible for the recognition and cleavage of bacteriocin precursors by the ABC transporters and are effective

in some noncognate precursor−ABC transporter combinations (*26*). Thus, the substrate-recognition mechanism proposed here might be common in the family of bacteriocin-associated ABC transporters.

ACKNOWLEDGMENT

We thank Dr. Hideji Yoshida for his help with the peptide sequencing and Hung-Soo Park for his technical assistance.

SUPPORTING INFORMATION AVAILABLE

Cloning and plasmid constructions, Table S1 (oligonucleotides used for cloning and plasmid constructions), Table S2 (oligonucleotides used for the mutagenesis of PComC), Figure S1 (thermostability of PEPs), and Figure S2 (v/[E] versus substrate concentration plots of PPEP for G1ComC and CComC). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hardie, J. M. (1986) Genus *Streptococcus* Rosenbach 1884, 22AL., in *Bergey's Manual of Systematic Bacteriology* (Sneath, P. H. A., Mair, N. S., Sharpe, M. E., and Holt, J. G., Eds.), pp 1043−1063, Williams & Wilkins, Baltimore.
2. Meyer, D. H., and Fives-Taylor, P. M. (1998) Oral pathogens: from dental plaque to cardiac disease, *Curr. Opin. Microbiol. 1*, 88−95.
3. Costerton, J. W., Stewart, P. S., and Greenberg, E. P. (1999) Bacterial biofilms: a common cause of persistent infections, *Science 284*, 1318−1322.
4. Cvitkovitch, D. G., Li, Y. H., and Ellen, R. P. (2003) Quorum sensing and biofilm formation in Streptococcal infections, *J. Clin. Invest. 112*, 1626−1632.
5. Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators, *J. Bacteriol. 176*, 269−275.
6. Fuqua, C., Parsek, M. R., and Greenberg, E. P. (2001) Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing, *Annu. Rev. Genet. 35*, 439−468.
7. Miller, M. B., and Bassler, B. L. (2001) Quorum sensing in bacteria, *Annu. Rev. Microbiol. 55*, 165−199.
8. Greenberg, E. P. (2003) Bacterial communication and group behavior, *J. Clin. Invest. 112*, 1288−1290.
9. Lazazzera, B. A. (2000) Quorum sensing and starvation: signals for entry into stationary phase, *Curr. Opin. Microbiol. 3*, 177−182.
10. Kleerebezem, M., and Quadri, L. E. (2001) Peptide pheromone-dependent regulation of antimicrobial peptide production in Gram-positive bacteria: a case of multicellular behavior, *Peptides 22*, 1579−1596.
11. Fuqua, C., and Greenberg, E. P. (2002) Listening in on bacteria: acyl-homoserine lactone signaling, *Nat. Rev. Mol. Cell Biol. 3*, 685−695.
12. Luo, P., Li, H., and Morrison, D. A. (2004) Identification of ComW as a new component in the regulation of genetic transformation in *Streptococcus pneumoniae*, *Mol. Microbiol. 54*, 172−183.
13. Gilmore, K. S., Srinivas, P., Akins, D. R., Hatter, K. L., and Gilmore, M. S. (2003) Growth, development, and gene expression in a persistent *Streptococcus gordonii* biofilm, *Infect. Immun. 71*, 4759−4766.
14. Li, Y. H., Tang, N., Aspiras, M. B., Lau, P. C., Lee, J. H., Ellen, R. P., and Cvitkovitch, D. G. (2002) A quorum-sensing signaling system essential for genetic competence in *Streptococcus mutans* is involved in biofilm formation, *J. Bacteriol. 184*, 2699−2708.
15. Venturi, V. (2006) Regulation of quorum sensing in *Pseudomonas*, *FEMS Microbiol. Rev. 30*, 274−291.
16. Håvarstein, L. S., Diep, D. B., and Nes, I. F. (1995) A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export, *Mol. Microbiol. 16*, 229−240.
17. Wu, K. H., and Tai, P. C. (2004) Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter, *J. Biol. Chem. 279*, 901−909.
18. Ishii, S., Yano, T., and Hayashi, H. (2006) Expression and characterization of the peptidase domain of *Streptococcus pneumoniae* ComA, a bifunctional ATP-binding cassette transporter involved in quorum sensing pathway, *J. Biol. Chem. 281*, 4726−4731.
19. Håvarstein, L. S., Hakenbeck, R., and Gaustad, P. (1997) Natural competence in the genus *Streptococcus*: evidence that streptococci can change pherotype by interspecies recombinational exchanges, *J. Bacteriol. 179*, 6589−6594.
20. Hui, F. M., Zhou, L., and Morrison, D. A. (1995) Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes, *Gene 153*, 25−31.
21. Martin, B., Humbert, O., Camara, M., Guenzi, E., Walker, J., Mitchell, T., Andrew, P., Prudhomme, M., Alloing, G., Hakenbeck, R., Morrison, D. A., Boulnois, G. A., and Claverys, J. P. (1992) A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*, *Nucleic Acids Res. 20*, 3479−3483.
22. Teng, L. J., Hsueh, P. R., Tsai, J. C., Chen, P. W., Hsu, J. C., Lai, H. C., Lee, C. N., and Ho, S. W. (2002) groESL sequence determination, phylogenetic analysis, and species differentiation for viridans group streptococci, *J. Clin. Microbiol. 40*, 3172−3178.
23. Tapp, J., Thollesson, M., and Herrmann, B. (2003) Phylogenetic relationships and genotyping of the genus *Streptococcus* by sequence determination of the RNase P RNA gene, rnpB, *Int. J. Syst. Evol. Microbiol. 53*, 1861−1871.
24. Sprules, T., Kawulka, K. E., Gibbs, A. C., Wishart, D. S., and Vederas, J. C. (2004) NMR solution structure of the precursor for carnobacteriocin B2, an antimicrobial peptide from *Carnobacterium piscicola*, *Eur. J. Biochem. 271*, 1748−1756.
25. Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972) Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion, *Biochemistry 11*, 4120−4131.
26. van Belkum, M. J., Worobo, R. W., and Stiles, M. E. (1997) Double-glycine-leader peptides direct secretion of bacteriocins by ABC transporters: colicin V secrection in *Lactococcus lactis*, *Mol. Microbiol. 23*, 1293−1301.
27. Schagger, H., and von, J. G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem. 166*, 368−379.

BI702253N