WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

The crystal structure of the estA protein, a virulence factor from *Streptococcus pneumoniae*

Myung Hee Kim,[1] Beom Sik Kang,[2] Sujin Kim,[1] Kyung-Jin Kim,[3] Choong Hwan Lee,[4] Byung-Chul Oh,[5] Seung-Chun Park,[6] and Tae-Kwang Oh[1]*

[1] Systems Microbiology Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Korea

[2] School of Life Science and Biotechnology, Kyungpook National University, Daegu 702-701, Korea

[3] Beamline Division, Pohang Accelerator Laboratory, Pohang, Kyungbuk 790-784, Korea

[4] Division of Bioscience and Biotechnology, IBST, Konkuk University, Seoul 143-701, Korea

[5] Lee Gil Ya Cancer and Diabetes Institute, Gachon University of Medicine and Science, Incheon 406-840, Korea

[6] College of Veterinary Medicine, Kyungpook National University, Daegu 702-701, Korea

INTRODUCTION

Streptococcus pneumoniae is currently the leading cause of a variety of severe infections, such as pneumonia, sepsis, and meningitis.[1] The infections are associated with high mortality rates of 5–35% depending on the site of infection, age, and comorbidity.[2,3] Pneumococcal infections have been treated with penicillin G and β-lactam antibiotics for many years. However, the prevalence of antibiotic resistant strains has progressively increased during the past decade. Accordingly, efforts to generate new drug candidates against invasive pathogens need to be carried out. Pneumococcal genes essential for lung infection were recently identified by using a signature-tagged mutagenesis technique, which showed an attenuation of virulence in a murine model of pneumonia.[4] A gene (*estA*), encoding a putative 259-residue esterase (estA), was found to be a virulence factor for lung infection, although its function to promote virulence is unknown. The *estA* gene is also found in other pathogenic bacteria, such as *S. sanguinis*, *S. suis*, *S. pyogene*, *S. agalactiae*, *S. mutans*, and the new pathogen *Lactococcus lactis*, with 53–75% identity, which suggests its involvement in virulence [Fig. 1(D)]. In this study, a high resolution 3D structure of the estA protein was determined, providing information on the functional aspects of this protein.

METHODS

Protein production and crystallization

The *estA* gene coding for the estA protein was amplified from the chromosomal DNA of *Streptococcus pneumoniae* (ATCC BAA255D) by polymerase chain reaction (PCR) using the following primers: 5′-CGCGGATCCGG-CAGTAATGAAAATCG-3′ and 5′-CCCAAGCTTTCAAGT-CAGTCTCTCTTCTAATTTG-3′, which carried *Bam*HI and *Hin*dIII restriction sites, respectively. The *estA* gene was subcloned into the *Bam*HI and *Hin*dIII sites of the pHis-Parallel1 expression vector, which is a hexahistidine (6×His)-fusion protein expression vector containing a TEV protease cleavage site.[5] *Escherichia coli* Rosetta-gami (DE3) cells harboring the recombinant plasmid were grown in an LB medium-containing ampicillin at 37°C until reaching an OD_{600} between 0.6 and 0.8. The temperature was then lowered to 25°C, and expression of estA was induced with 0.5 m*M* isopropyl-β-D-thiogalactopyranoside

Grant sponsor: 21C Frontier Microbial Genomics and Application Center Program, Ministry of Science and Technology, Republic of Korea.
*Correspondence to: Dr. Tae-Kwang Oh, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Korea. E-mail: otk@kribb.re.kr
Received 24 April 2007; Revised 12 June 2007; Accepted 17 June 2007
Published online 11 October 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21680

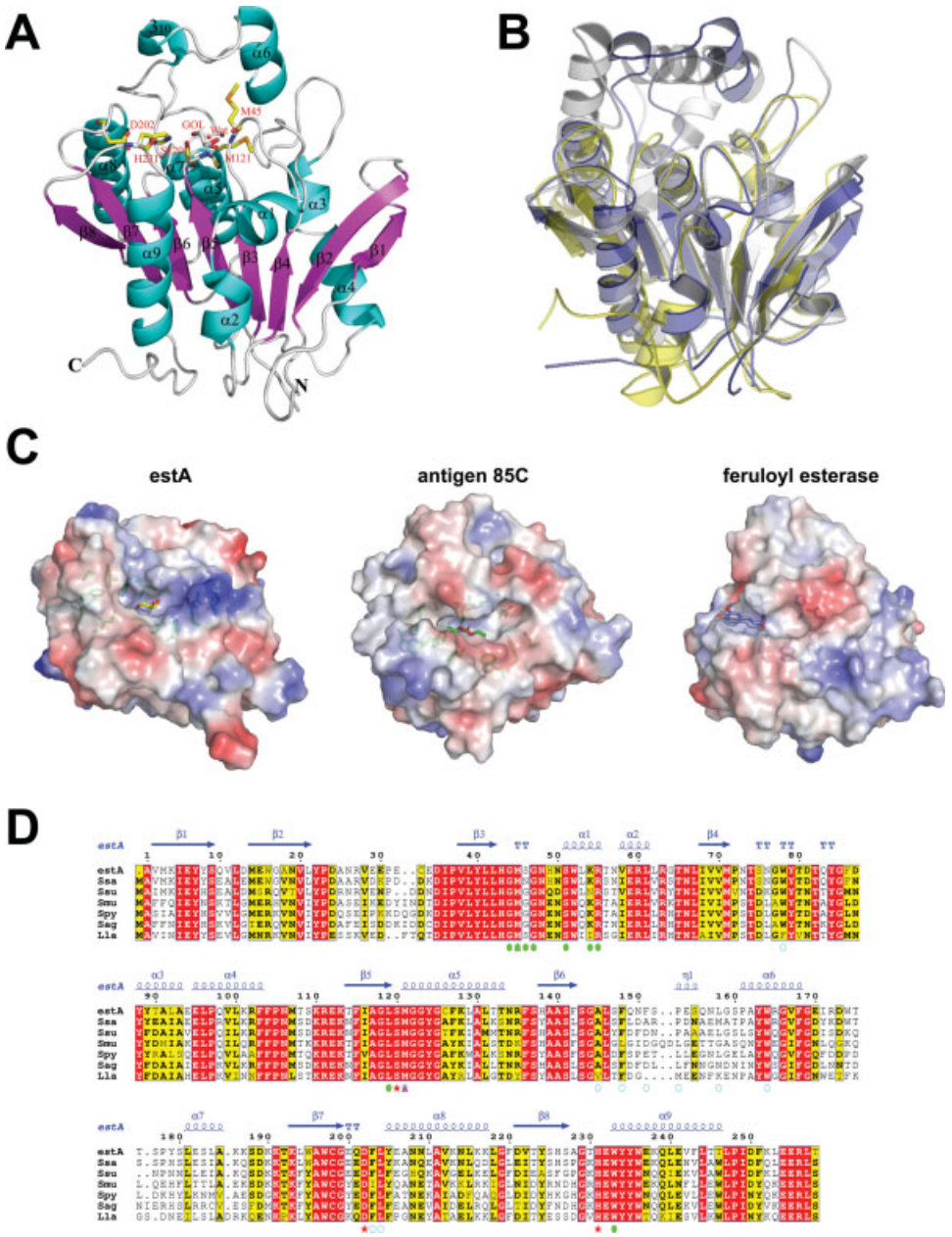

Figure 1

Structural characterization of the estA protein. ***A:*** *Ribbon representation of the estA protein.* α*-Helices and* β*-sheets are shown in cyan and purple, respectively. Residues of the catalytic triad Ser120, His 231, Asp202, and putative oxyanion hole forming backbone amides of Met45 and Met121 are shown (carbon, yellow; oxygen, red; nitrogen, blue; and sulfur, orange). Carbon and oxygen atoms for the glycerol molecule (GOL) are shown in white and red, respectively. Water molecule (Wat) is shown in red.* ***B:*** *Superposition of estA with the antigen 85C (PDB ID code 1DQY) and feruloyl esterase (PDB ID code 1JT2) proteins. The estA, antigen 85C, and feruloyl esterase proteins are shown in blue, white, and yellow, respectively.* ***C:*** *Comparison of substrate-binding sites. The surface representations are colored according to electrostatic potential in a range where red is negative (−110 kT/e) and blue is positive (+110 kT/e). Residues forming substrate-binding sites are displayed. Glycerol, diethyl phosphate, and ferulic acid are bound to estA, antigen 85C, and feruloyl esterase, respectively. Unless otherwise noted, figures were prepared using PyMOL (http://www.pymol.org).* ***D:*** *Structure-related amino-acid sequence conservation in putative esterases from pathogenic bacteria. The above-mentioned alignments are the elements of secondary structure of estA. Numbering shown is from estA. Red stars represent catalytic triad residues, closed purple triangles represent residues involved in oxyanion hole, open cyan circles represent residues involved in the hydrophobic pocket, and residues forming the long and narrow cavity are indicated by closed green circles. Strictly conserved residues are highlighted with red boxes. Biological sources and accession codes for the sequences are as follows: Ssa, S. sanguinis SK36 (gi125718929); Ssu, S. suis 89/1591 (gi81096828); Smu, S. mutans UA159 (gi24379842); Spy, S. pyogenes MGAS8232 (gi19746007); Sag, S. agalactiae 515 (gi77412854); and Lla, L. lactis subsp. lactis Il1403 (gi15673753). Sequence alignments were assembled using T-COFFEE software and visualized using ESPript software, both located on the ExPASy Proteomics Server (http://au.expasy.org/).*

(IPTG) for 14 h. The cells were harvested by centrifugation at 5,000*g* for 20 min at 4°C. The cell pellets were resuspended in buffer A (50 m*M* Tris–HCl pH 8.0 and 300 m*M* NaCl) and disrupted by ultrasonication. The crude cell extracts were centrifuged at 11,000*g* for 1 h at 4°C. The cell lysate containing the 6×His-tagged estA protein was then bound to Ni-NTA agarose (QIAGEN) equilibrated with buffer A for 2 h at 4°C. After the resin was washed with buffer A, the bound proteins were eluted with buffer A containing 200 m*M* imidazole. The 6×His-tag was then cleaved from the estA by incubating with rTEV protease (GIBCO), followed by Ni-NTA agarose chromatography and size exclusion chromatography. After purification, the protein retained a five-residue cloning artifact (Gly-Ala-Met-Asp-Pro) at its N-termini. The homogeneity of the protein was assessed by 10% SDS-PAGE and Coomassie blue staining. The purified protein was then dialyzed against 20 m*M* Tris–HCl pH 7.5, concentrated to 15 mg/mL using a YM-10 membrane (Amicon), and stored at −80°C for use in the crystallization trials. The SeMet-substituted protein was expressed in the methionine auxotroph strain *E. coli* B834 (DE3) in a minimal medium supplemented with 50 mg/mL SeMet under the same conditions as the native plasmid. The purification procedure of the SeMet substituted protein was identical to that of the native protein, except for the addition of 5 m*M* methionine to all of the buffers. Gel-filtration analysis of the estA protein was carried out on Superdex-75 10/30 column (Amersham Biosciences). The column was calibrated using the following molecular-mass standards: ribonuclease (13.7 kDa), chymotrypsinogen (25 kDa), ovalbumin (43 kDa), and bovine serum albumin (67 kDa). The protein crystals were initially screened using Crystal Screen and Crystal Screen 2 from Hampton Research using a sitting drop technique at 21°C. The crystals appeared in a week under the following conditions: 10% (v/v) PEG 8000, 8% (v/v) ethylene glycol, and 0.1 *M* HEPES pH 7.5. Optimization of these conditions to 10% (v/v) PEG 6000, 8% (v/v) ethylene glycol, 4% (v/v) 2-methyl-2,4-pentanediol (MPD), 10 m*M* cysteine, and 0.1 *M* Tris–HCl pH 7.0 improved the quality of the crystals. X-ray diffraction-quality crystals appeared in 2 weeks. The SeMet substituted protein crystals were also obtained under the same crystallization conditions.

Metal ion determination

The purified estA protein was incubated in 20 m*M* Tris–HCl pH 8.0 and 5 m*M* $CaCl_2$ at 4°C overnight. The Ca^{2+}-replenished protein was then concentrated and extensively dialyzed against a 20 m*M* Tris–HCl buffer at pH 8.0 to remove unbound Ca^{2+}. Before Ca^{2+} content measurement, the protein was applied to a PD-10 column (Amersham Biosciences) equilibrated with a metal-free Tris–HCl buffer at pH 8.0, which was made by passing over a Chelex 100 column (Bio-Rad). The metal content in estA was confirmed using inductively coupled plasma atomic emission spectrometry (ICP-AES), where 1 mL of the sample (1 mg/mL concentration) was analyzed. A PD-10 column-treated metal-free Tris–HCl buffer was used as the control.

Thermal stability

The thermal stability of the estA protein was evaluated by incubating the protein in the presence of either 5 m*M* $CaCl_2$ or 5 m*M* EDTA in 20 m*M* Tris–HCl pH 8.0 at different temperatures ranging from 10 to 60°C for 30 min. Subsequently, an activity assay was carried out by measuring the hydrolysis of the chromogenic *p*-nitrophenyl butyrate substrate. The reaction mixture contained a 50 m*M* Tris–HCl buffer at pH 8.0, 4% (v/v) ethanol, 0.1 m*M* *p*-nitrophenyl butyrate, and 2.5 μL of the purified estA (0.044 mg/mL) in a total volume of 1 mL. The residual activity was measured by addition of the estA protein at 30°C. The reaction rate was calculated by recording the change in the absorbance at 405 nm.

Data collection, structure determination, and refinement

The crystals were transferred to a cryo-protecting solution containing 12% (v/v) PEG 6000, 8% (v/v) ethylene glycol, 4% (v/v) MPD, 0.1 *M* Tris–HCl pH 7.0, and 20% (v/v) glycerol, fished, and placed immediately in a −173°C nitrogen-gas stream. The single-wavelength anomalous diffraction (SAD) data for the SeMet crystals were collected on beamline 6B at the Pohang Accelerator Laboratory (PAL), Pohang, Korea, at a 1.9 Å resolution. Since the SeMet-substituted crystals diffracted better than the native crystals, SeMet crystals were used to collect high resolution data at 1.7 Å on beamline 4A at PAL. All data were processed using the HKL2000.[6] The SeMet-estA crystal belongs to the space group $P2_1$. There were four molecules in the asymmetric unit with a packing density of 2.12 $Å^3$/Da corresponding to an estimated solvent content of 42%. The structure was determined by SAD phasing utilizing the anomalous signal from Se atoms using SOLVE,[7] which identified 26 sites. Density modification and subsequent automated model building were performed using the program RESOLVE.[8] The initial RESOLVE-built model was used as a guide to manually build the remainder of the protein into density-modified electron density maps with the programs COOT[9] and O.[10] Translation liberation screw-rotation (TLS) parameters and restrained refinement options in REFMAC5[11] were used for the final refinement cycles. The SeMet-substituted crystal structure was solved at 1.7 Å resolution by molecular replacement with AMORE[12] using the partially refined model of the SeMet crystal at 1.9 Å resolution. The refinement again included the TLS

DOI 10.1002/prot

procedure. The cycles of manual rebuilding and refinement resulted in R_{free} and R_{cryst} of 0.213 and 0.183, respectively. No density was visible for the N-terminal nine amino acid residues (Ala25 to Cys33) of chain A, C, and D and 10 amino acid residues (Ala25 to Glu34) of chain B, indicating significant flexibility in these regions. In addition, a disordered map was observed between residues Phe151 and Leu158 of chain D. They were not included in the model. The side chains of residues Met6 and Met17 of each chain were modeled using alternate conformations. The model contained 971 amino acid residues, including the five artificial amino acids at the N-termini derived from the cloning, 2 calcium ions, 6 glycerols, 4 ethylene glycols, and 563 water molecules. The model satisfied the quality criteria limits of the program PROCHECK.[13] The crystallographic data statistics are summarized in Table I. The atomic coordinates and structure factor amplitudes of SeMet-estA have been deposited in the PDB under the accession code 2UZ0.

Table I
Data Collection and Refinement Statistics

Dataset	SeMet_EstA_peak	SeMet_EstA
Beamline (PAL)	6B (MX)	4A (MXW)
Wavelength	0.97894	1.00000
Space group	$P2_1$	$P2_1$
Cell dimensions (Å)		
A	71.58	71.69
B	72.03	72.13
C	101.23	101.83
β (°)	94.65	94.31
Resolution (Å)	1.9 (1.97–1.9)	1.7 (1.76–1.7)
No. of total reflections	301,512	424,530
No. of unique reflections	79,927	113,097
Redundancy	3.8	3.8
Completeness (%)	98.9 (98.1)	99.9 (99.1)
R_{sym} (%)[a]	9.3 (44.1)	7.8 (45.5)
$I/\sigma(I)$	25.0 (3.1)	15.7 (2.4)
Refinement		
Resolution (Å)		30.0–1.70
Reflections in work/test sets		107,413/5,662
R_{work}/R_{free} (%)[b,c]		18.3/21.3
R.m.s. deviations		
Bond lengths (Å)		0.011
Bond angles (°)		1.307
Model composition		
		971 residues
		563 waters
		6 glycerols
		4 ethylene glycols
		2 Ca^{2+}
Geometry		
Most favored regions (%)		88.9
Additional allowed regions (%)		11.1
PDB accession code		2UZ0

The numbers in parentheses describe the relevant value for the highest resolution shell.
[a] $R_{sym} = \sum|I_i - \langle I\rangle|/\sum I$ where I_i is the intensity of the *i*th observation and $\langle I\rangle$ is the mean intensity of the reflections.
[b] $R_{work} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$ where F_{calc} and F_{obs} are the calculated and observed structure factor amplitude, respectively.
[c] $R_{free} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$ where all reflections belong to a test set of randomly selected data.

RESULTS AND DISCUSSION

The crystal structure of estA was refined at 1.7 Å resolution. The asymmetric unit contains four protein molecules, which are arranged as a dimer of homodimers. A dimer with two molecules by the twofold pseudosymmetry axis binds the other dimer related to the second twofold pseudosymmetric axis perpendicular to the first axis. The structure of each estA molecule in the asymmetric unit is essentially the same, and the RMSDs of the C_α atoms of the molecules are all less than 0.36 Å. The estA monomer comprises two domains, a large catalytic domain containing the catalytic triad and a small cap domain covering the substrate-binding site located at the interface between two domains [Fig. 1(A)]. The catalytic domain displays a canonical α/β hydrolase topology. It contains typical eight-stranded β-sheet consisted of a β-hairpin structure and six-stranded parallel β-sheet and helices forming the outer, solvent-exposed layer surrounding the β-sheets. The smaller domain (Ala145 to Pro178) is composed of an α-helix and a 3_{10} helix that form a small cap over the active site. The α/β-hydrolase fold of estA embraces the classical Ser-His-Asp "catalytic triad" comprised residues Ser120, His231, and Asp202 [Fig. 1(A)]. The serine residue is located at the tip of a β-strand-turn-helix structural motif termed a "nucleophilic elbow," which is formed by strand β5 and helix α5. To confirm the position of the catalytic serine residue, Ser120 was mutated to an alanine. As expected, the mutant showed no detectable activity against *p*-nitrophenyl butyrate (data not shown). His231 located on a loop between β8 and α9 is juxtaposed to shuttle protons between the serine nucleophile and the general base Asp202. All protein molecules show that the Nε2 is hydrogen bonded to the hydroxyl group (~3.19 Å) of Ser120 and that the Nδ1 atom is hydrogen bonded to the Oδ1 atom (~3.23 Å) and Oδ2 atom (~2.70 Å) of Asp202. This stereochemical structure is consistent with those of the catalytic triads found in other α/β-hydrolase families. Thus, it is believed that the estA protein maintains the conserved hydrolysis mechanism seen in the α/β-hydrolase family, although different substrate specificities may be expected. The catalytic serine is completely exposed to the solvent in a long and narrow cavity, which is lined by Gly44, Met45, Ser46, Gly47, Ser51, Lys54, Arg55, Leu119, and Trp233 [Fig. 1(C)]. The small cap domain forms a hydrophobic pocket generated by Trp78, Met121, Ala145, Phe148, Phe151, Pro153, Leu158, Trp164, Phe203, and Leu204. This pocket is connected to the opening cavity that extends through the center of the protein [Fig. 1(C)]. The calculated volume of the cavity and hydrophobic pocket are only 251 $Å^3$.[14] In the estA structure, three glycerol molecules (802, 803, and 804)

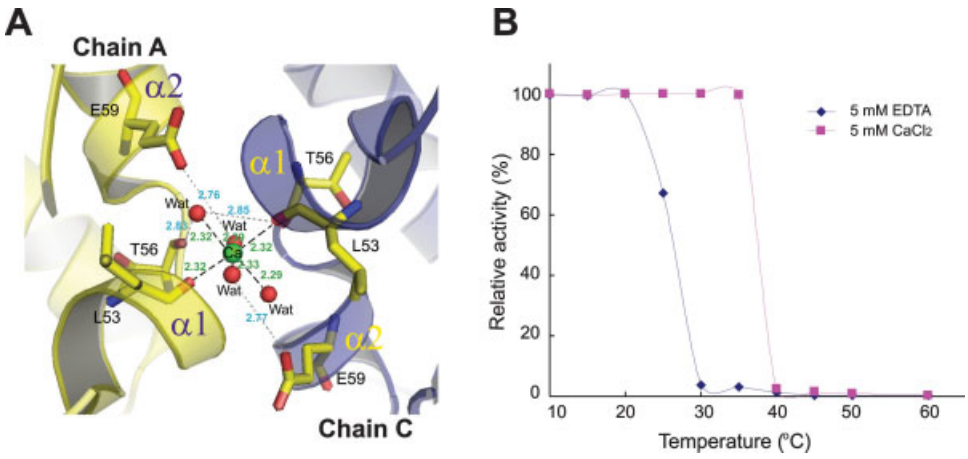

Figure 2

Calcium-mediated dimerization of the estA protein. ***A:*** *Calcium-mediated dimer interface between estA molecules. Interactions for calcium coordination are outlined by black long-dashed lines, and hydrogen bonds are depicted as gray short-dashed lines. Interactions were measured in angstroms. Carbons, oxygens, and nitrogens are shown in yellow, red, and blue, respectively. Water molecules are indicated as red spheres, and calcium ion is shown as green sphere.* ***B:*** *Stabilizing effect of calcium ion on the estA protein.*

and four ethylene glycol molecules (901, 902, 903, and 904) are hydrogen bonded between crystallographically related protein molecules. Three other glycerols 801, 805, and 806 are directly bound over the hydroxyl group of the catalytic residue Ser120 in molecules A, C, and D, respectively. An ambiguous electron density is observed in this region in molecule B, which may suggest the location of a glycerol molecule. The glycerols may mimic part of a natural substrate of estA [Fig. 1(C)]. They are positioned in the cavity and also adjacent to a hydrophobic pocket suitable for accommodating substrates. A water molecule, which may also be a part of an estA protein substrate, is positioned near to the glycerols that are bound to the serine hydrogen bonds with the Met45 and Met121 backbone nitrogen atoms at distance of ~2.91 and ~3.09 Å [Fig. 1(A)]. These two residues form the oxyanion hole that may be critical for the substrate binding and transition state stabilization for subsequent catalysis by the estA protein. A structural fold similarity search conducted using DALI[15] reveals that estA has 52 structural homologues with a *Z*-score greater than 12. The rms deviation for these structures ranges from 2.1 to 3.8 Å for 164 to 232 aligned C_α atoms, respectively. These similarities could not be easily inferred from the sequence alone, because the pairwise sequence identities with estA are 9–28%. Interestingly, the estA protein shows high structural similarity to the *Mycobacterium tuberculosis* antigen 85C (PDB ID code 1DQY; $Z = 25.8$)[16] and anaerobic bacterium *Clostridium thermocellum* feruloyl esterase (PDB ID code 1JJF and 1JT2; $Z = 23.7$) proteins [Fig. 1(B)].[17] The catalytic triad residues and the oxyanion hole of the proteins are structurally well conserved. The most striking structural difference between the proteins is the presence of the small cap domain in the estA protein, which generates a hydrophobic pocket. The antigen 85C also has a similar cap domain that together with a long α-helix (α5), which is much longer than the corresponding α-helix (α8) of estA, forms the wall of a large hydrophobic pocket [Fig. 1(B,C)]. There is no equivalent domain in the feruloyl esterase structure and as a result, this enzyme does not contain a hydrophobic pocket. In addition to the hydrophobic pocket, estA has a long, narrow cavity as mentioned earlier. The cavity is a different shape from that of the antigen 85C and feruloyl esterase proteins. The cavity of the antigen 85C forms a longer channel than estA, whereas the feruloyl esterase has a wider cavity than estA. The different shapes of the substrate-binding sites of these proteins may lead to different specificities.

The size exclusion chromatography revealed that the estA forms a dimer in solution (data not shown). Analysis of the lattice packing interactions in the crystal shows two distinct dimer interfaces. One type (between monomers A and C or B and D) of the interface is mediated by a calcium ion. The calcium ion is coordinated by the carbonyl oxygen atoms of Leu53 from chains A and C, and four water molecules with the Ca—O bond distances ranging from 2.29 to 2.39 Å. Thus, the Ca^{2+} ion shows an octahedral coordination [Fig. 2(A)]. In addition, the calcium-coordinated water molecules are stabilized by hydrogen bonding with residues Thr56 and Glu59 from each chain [Fig. 2(A)]. The same coordination of the calcium ion is seen between molecules B and D. The second dimer interface (between monomers A and D or B and

 DOI 10.1002/prot

C), which is an antiparallel association of a strand involving residues Val2 to Glu6, seems to be derived from crystal packing. The calcium ion of the estA protein was independently verified by ICP-AES. The analysis showed a mean calcium content of 0.62 ± 0.06 mol of Ca^{2+}/mol of protein, indicating the presence of one Ca^{2+} ion per two estA protein molecules. These results suggest that the calcium-mediated dimer may represent the actual functional unit of the estA protein. However, the Ca^{2+} ion is 20.26 Å away from the serine residue in the active site, ruling out any direct involvement of the Ca^{2+} ion in the enzyme catalysis. This implies that the Ca^{2+} in the estA protein may play a role in stabilization. Thermostability of the estA protein was evaluated by incubating the protein at different temperatures ranging from 10 to 60°C in the presence of either Ca^{2+} or EDTA. The protein gained full activity when incubated at 35°C in the presence of calcium whereas the activity was lower at 25°C and completely absent at 30°C in the presence of EDTA [Fig. 2(B)]. These results indicate that the calcium ion plays a critical role in maintaining estA protein stability under physiological conditions, although further studies are required to validate its precise function.

ACKNOWLEDGMENTS

We thank the beamline scientists in PAL (MXW 4A and MX 6B) for their help in the X-ray diffraction experiments.

REFERENCES

1. Kalin M, Örtqvist Å, Almela M, Aufwerber E, Dwyer B, Henriques B, Jorup C, Julander I, Marrie TJ, Mufson MA, Riquelme R, Thalme A, Torres A, Woodhead MA. Prospective study of prognostic factors in community-acquired bacteremic pneumococcal disease in 5 countries. J Infect Dis 2000;182:840–847.
2. Moroney JF, Fiore AE, Harrison LH, Patterson J, Farley MM, Jorgenson JH, Phelan M, Facklam RR, Cetron MS, Breiman RF, Kolczak M, Schuchat Al. Clinical outcomes of bacteremic pneumococcal pneumonia in the era of antibiotic resistance. Clin Infect Dis 2001;33:797–805.
3. Robinson KA, Baughman W, Rothrock G, Barrett NL, Pass M, Lexau C, Damaske B, Stefonek K, Barnes B, Patterson J, Zell ER, Schuchat A, Whitney CG. Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995-1998: opportunities for prevention in the conjugate vaccine era. JAMA 2001;285:1729–1735.
4. Hava DL, Camilli A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol Microbiol 2002;45:1389–1406.
5. Sheffield P, Garrard S, Derewenda Z. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr Purif 1999;15:34–39.
6. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
7. Terwilliger TC, Berendzen J. Bayesian correlated MAD phasing. Acta Crystallogr D Biol Crystallogr 1997;53:571–579.
8. Terwilliger TC. Maximum-likehood density modification with pattern recognition of structural motifs. Acta Crystallogr D Biol Crystallogr 2001;57:1755–1762.
9. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
10. Jones AT, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 1991;47:110–119.
11. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53:240–255.
12. Navaza J. AMoRe: an automated package for molecular replacement. Acta Crystallogr A 1994;50:157–163.
13. Laskowski RA, MacArthur MW, Moss DS, Thornton JM, PROCHECK—a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
14. Binkowski TA, Naghibzadeg S, Liang J. CASTp: computed atlas of surface topography of proteins. Nucleic Acid Res 2003;31:3352–3355.
15. Holm L, Sander C. Mapping the protein universe. Science 1996;273:595–602.
16. Ronning DR, Klabunde T, Besra GS, Vissa VD, Belisle JT, Sacchettini JC. Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. Nat Struct Biol 2000;7:141–146.
17. Schubot FD, Kataeva IA, Blum DL, Shah AK, Ljungdahl LG, Rose JP, Wang BC. Structural basis for the substrate specificity of the feruloyl esterase domain of the cellulosomal xylanase Z from *Clostridium thermocellum*. Biochemistry 2001;40:12524–12532.