An esterase from *Escherichia coli* with a sequence similarity to hormone-sensitive lipase

Shigenori KANAYA*, Tomoyoshi KOYANAGI* and Eiko KANAYA†

*Department of Material and Life Science, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565, Japan, and †Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan

An esterase from *Escherichia coli* that is a member of the hormone-sensitive lipase (HSL) family was overproduced, purified and characterized. It is encoded by the *ybaC* gene and composed of 319 amino acid residues with an M_r of 36038. The enzymic activity was determined by using various *p*-nitrophenyl esters of fatty acids as a substrate at 25 °C and pH 7.1. The enzyme showed hydrolytic activity towards substrates with an acyl chain length of less than 8, whereas it showed little hydrolytic activity towards those with an acyl chain length of more than 10. In addition, it showed little hydrolytic activity towards trioleoylglycerol and cholesterol oleate. Determination of the kinetic parameters for the hydrolyses of the substrates from C_2 to C_8 indicates that C_4 and C_5 substrates are the most preferred. Close agreement between the M_r determined by SDS/PAGE (37000) and column chromatography (38000) suggests that the enzyme exists in a monomeric form. It is an acidic protein with a pI value of 4.1. The far-UV CD spectrum suggests that its helical content is 26.1 %. Comparison of the amino acid sequence of this enzyme with those involved in the HSL family allows us to propose that Ser^{165}, Asp^{262} and His^{292} constitute the catalytic triad of *E. coli* esterase.

INTRODUCTION

Proteins with similar amino acid sequences usually have the same folding conformation, even if their similarities are minimal [1,2]. They are generally related with one another in function as well, and therefore grouped into a family. Occasionally, distant organisms produce proteins that are the members of the same family. In this case, comparison of the properties of these proteins provides valuable information on their evolutionary relationships. In addition, detailed analyses of the structure and function of a given protein may help us to understand a fundamental mechanism for protein function in the family. For this purpose, prokaryotic proteins are preferred to eukaryotic ones for analysis because of their simple structure (function) and ease of preparation.

It has previously been shown that *Escherichia coli* chromosome has a gene which encodes a hormone-sensitive lipase (HSL) homologue [3,4]. This gene was originally shown to encode a protein with 203 amino acid residues with unknown function and therefore had been designated *ORF203* [5]. Later, the complete genome sequence of *E. coli* K12 showed that this gene (*ybaC*) encodes a protein with 319 amino acid residues (M_r 36038) [6]. The location of this gene on the map of the *E. coli* chromosome is shown in Figure 1. This gene is located close to the *hemH* (*visA*) gene, which encodes ferrochelatase [7]. The termination codons for *ybaC* and *hemH* overlap in opposite directions. In addition to this *ybaC* gene product, several bacterial proteins, such as lipase 2 from *Moraxella* TA144, acetyl hydrolases from *Streptomyces hygroscopicus* and *Streptomyces viridochromogenes*, and the ORF3 protein from *Bacillus acidocaldarius*, have been shown to be closely related to HSL in amino acid sequence [3,4]. Therefore, these proteins have been proposed to belong to the same superfamily (the HSL family). None of the proteins in this family shows sequence similarity to other well-known lipases and esterases.

HSL is a unique enzyme that hydrolyses triacylglycerols stored in adipose tissue under acute hormonal and neural controls (for a review see [8]). It is a multi-functional enzyme and is much larger than the bacterial proteins in a HSL family in size. Among various HSLs, rat HSL with 768 amino acid residues [9,10], which is highly similar to human HSL with 786 amino acid residues [11], has been most extensively studied for structure and function. Analyses of sequence similarities to bacterial proteins, as well as the biochemical characterizations, suggest that HSL is composed of domain structures [12,13]. They are the N-terminal domains with unknown function, the catalytic domain, and the regulatory domain that contains two phosphorylation sites (Ser^{563} and Ser^{565}). Site-directed-mutagenesis studies suggest that Ser^{423}, Asp^{703} and His^{733} constitute the catalytic triad of the enzyme

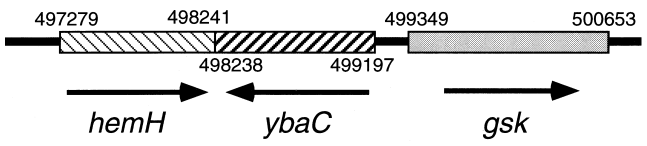

Figure 1 Location of the *ybaC* gene on the map of the *E. coli* chromosome

Numbers along the gene represent the positions of the base-pairs in the *E. coli* K-12 genome [6]. The *hemH* gene encodes ferrochelatase with 310 amino acid residues. The *gsk* gene encodes a protein with 434 amino acid residues that is involved in nucleotide biosynthesis and metabolism. The direction of the transcription for each gene is shown by an arrow.

Abbreviations used: HSL, hormone-sensitive lipase; LB, Luria–Bertani; PNPB, *p*-nitrophenyl butyrate; IPTG, isopropyl β-D-thiogalactoside; AChE, acetylcholinesterase.

[1] To whom correspondence should be addressed (e-mail kanaya@chem.eng.osaka-u.ac.jp).

[10,14]. Furthermore, on the basis of the similarity of the secondary structures of acetylcholinesterase, a three-dimensional model for the structure of the catalytic domain of the enzyme has been constructed [15]. A distant relationship between the HSL and acetylcholinesterase sequences was originally proposed by Hemila et al. [4].

The bacterial proteins in the HSL family are only homologous to the catalytic domain of HSL [4,12]. Their sizes are roughly comparable with the size of this catalytic domain as well. These results suggest that the three-dimensional structures of these bacterial proteins are similar to that of the catalytic domain of HSL and may represent a minimum structural element responsible for catalytic function. Interestingly, in the primary structure, the catalytic domain of HSL is composed of two regions that are separated by the insertion of the regulatory domain (residues 521–669) [12]. They are region 1 (residues ~ 330–~ 500), which contains a G-X-S-X-G motif that is conserved in most lipases/esterases and includes the active-site Ser residue [16], and region 2 (residues ~ 670–768), which contains other two active-site residues (Asp and His). How does the HSL molecule acquire functions other than the catalytic one during evolution? Comparative studies for the structures and functions of the proteins in the HSL family will be necessary to answer this question.

Since the *ybaC* gene product from *E. coli* is one of the smallest proteins among those in the HSL family, it may be suitable for structural and functional studies at the atomic level. However, no biochemical or enzymic characterization has been attempted for this protein. Here we report that the purified *ybaC* gene product exhibits an esterase activity, because it hydrolyses only the *p*-nitrophenyl esters of fatty acids with a relatively short acyl chain length. Therefore we propose hereafter to designate this enzyme '*E. coli* esterase'.

EXPERIMENTAL

Cells and plasmids

Plasmid pFC1, which contains the *ybaC* gene [5], was kindly donated by Dr. K. Miyamoto (Osaka University, Japan). Plasmid pDR600 for the overproduction of *E. coli* RNase HI [17] was previously constructed by one of us (S. K.). Competent cells of *E. coli* JM109 [recA1, endA1, gyrA96, thi, hsdR17, SupE44, relA1, λ^-, Δ(lac-ProAB)/F′, traD36, $ProA^+B^+$, $lacI^qZ\Delta M15$.] were obtained from Toyobo Co. (Osaka, Japan). Cells were grown in Luria–Bertani (LB) medium [18] containing 100 mg/l ampicillin.

Overproduction

The 1 kbp DNA fragment containing the *ybaC* gene was amplified by PCR with plasmid pFC1 as a template. Two DNA oligomers with sequences of 5′-CCGGATCCACAATGACGT-AATTTGA*AAGGAG*TTTTTGT-3′ and 5′-CCGTCGACTTA-AAGCTGAGCCGTAAAGAAC-3′, which were synthesized by Sawady Technology Co. (Tokyo, Japan) were used as 5′- and 3′-primers respectively. In these sequences the positions of the *Bam*HI and *Sal*I sites are underlined, and the Shine–Dalgarno sequence is shown in *italics*. The downsteam sequence of the *Bam*HI site of the 5′-primer is complementary to the genome sequence of *E. coli* from nt 499198 to nt 499227 [6]. Likewise, the downstream sequence of the *Sal*I site of the 3′-primer is identical with the genome sequence of *E. coli* from 498239 to 498259. PCR was performed in 24 cycles with a Perkin–Elmer DNA thermal cycler (model PJ2000) using a Gene Amp kit of Takara Shuzo Co. (Kyoto, Japan) according to the procedure recommended by the supplier. The resultant DNA fragment was digested with *Bam*HI and *Sal*I, and ligated to the large *Bam*HI–*Sal*I fragment of plasmid pDR600 to generate plasmid pDR319. In this plasmid, the *ybaC* gene is put under the control of the *tac* promoter. The nucleotide sequence of this gene was confirmed by the dideoxy-chain-termination method [19].

An overproducing strain was constructed by transforming *E. coli* JM109 with plasmid pDR319. Cultivation was carried out at 37 °C. When the A_{600} of the culture reached 0.8, 2 mM isopropyl β-D-thiogalactoside (IPTG) was added to the culture medium and cultivation was continued for additional 12 h. Cells were then harvested by centrifugation and subjected to the purification procedures described below. The production of *E. coli* esterase in cells was examined by analysing the whole-cell extract by SDS/PAGE [20], which was prepared by dissolving cells in 50 mM Tris/HCl, pH 6.8, containing 1 % SDS and 0.5 M 2-mercaptoethanol at 100 °C.

Purification

All purification procedures were carried out at 4 °C. Cells from a 1-litre culture were suspended in 25 ml of 40 mM Tris/HCl, pH 7.8, containing 25 % sucrose and lysed with lysozyme by the method of Nakamura and Yura [21]. The supernatant (crude extract) was adjusted to 20 % saturation with $(NH_4)_2SO_4$. After stirring for 15 min, the solution was centrifuged at 12000 ***g*** for 15 min. The resulting supernatant was 40 %-satd. with $(NH_4)_2SO_4$. After stirring for 15 min, the solution was again centrifuged at 12000 ***g*** for 15 min. The precipitate was dissolved in 10 ml of 10 mM Tris/HCl, pH 8.0, containing 1 mM EDTA (TE buffer) and applied to a column (2 ml) of DE-52 (Whatman) equilibrated with the same buffer. After washing the column, *E. coli* esterase was eluted from the column with a linear gradient of NaCl from 0 to 0.5 M in the TE buffer. The enzyme fractions were pooled, concentrated to 2 ml with Centricon 10 centrifugal concentrators (Amicon Corp.) and applied to a column (1.6 cm × 100 cm) of Sephacryl S-300 (Pharmacia) equilibrated with the TE buffer containing 0.1 M NaCl. Fractions (2.5 ml each) were collected. The flow rate was 10 ml/h. BSA, ovalbumin, chymotrypsinogen A and RNase A with M_r values of 67000, 43000, 25000 and 13700 respectively, were also applied individually to this column as standard proteins in order to determine the apparent M_r of *E. coli* esterase. The purity of the protein was analysed by SDS/PAGE.

Assay for enzymic activity

The esterase/lipase activity was determined by using various *p*-nitrophenyl esters of fatty acids with acyl chain lengths from 2 to 18 as a substrate. The hydrolysis of the substrate was carried out at 25 °C for 15 min in 100 μl of 20 mM phosphate buffer, pH 7.1, containing 5 % acetonitrile. The reaction was terminated by the addition of SDS to a final concentration of 0.2 %. The amount of *p*-nitrophenol produced by the reaction was detemined from the molar-absorption-coefficient value of 14200 $M^{-1} \cdot cm^{-1}$ at 412 nm. One unit of the enzymic activity was defined as the amount of the enzyme that produced 1 μmol of *p*-nitrophenol/min at 25 °C. The specific activity is defined as the enzymic activity per mg of protein. For the kinetic analyses, the substrate concentration was varied so that at least the highest and lowest concentrations of the substrate were larger and smaller than the K_m. The hydrolysis of the substrate by the enzyme followed Michaelis–Menten kinetics, and the kinetic parameters, K_m and V_{max}, were determined from Lineweaver–Burk plots.

The enzymic activities for the hydrolyses of tributyrylglycerol, trioleoylglycerol and cholesterol oleate were measured by

titrating fatty acids from these substrates with 0.05 M KOH as described previously [22]. The reaction was carried out in 50 mM sodium acetate, pH 5.6, containing 0.1 % Triton X-100 at 30 °C for 60 min with stirring at 500 rev./min. One unit of the enzymic activity was defined as the amount of the enzyme that produced 1 μmol of fatty acid/min at 30 °C. To compare the enzymic activities for the hydrolyses of these substrates with that for the hydrolysis of *p*-nitrophenyl butyrate (PNPB), the enzymic activity for the hydrolysis of PNPB was determined in 50 mM sodium acetate, pH 5.6, containing, in addition, 0.1 % Triton X-100 at 30 °C. The enzymic activity for the hydrolysis of tributyrylglycerol was also measured by examining whether an *E. coli* JM109 transformant with pDR319 forms a halo of lipolysis on a tributyrylglycerol/L-agar plate at 37 °C. For preparation of the plate, 0.5 % tributyrylglycerol was emulsified by vigorous sonication prior to autoclaving.

Protein concentration

The protein concentration of *E. coli* esterase was determined from the UV absorption. Because 17.3 nmol of protein per A_{280} unit of *E. coli* esterase was recovered by amino acid analysis (corresponding to 0.63 mg of protein with an M_r of 36038), we used an $A_{280}^{0.1\%}$ value of 1.60. This value is comparable with an $A_{280}^{0.1\%}$ value of 1.51 calculated by using 1,576 $M^{-1}\cdot cm^{-1}$ for tyrosine ($\times 18$) and 5225 $M^{-1}\cdot cm^{-1}$ for tryptophan ($\times 5$) at 280 nm [23].

Analyses for amino acid sequence and composition

Amino-acid-sequence analysis was carried out with a gas-phase automated sequencer (Applied Biosystems model 477A) equipped with an on-line HPLC apparatus (Applied Biosystems model 120A). Amino acid analysis was carried out with a Beckman system 6300E automatic amino acid analyser. Samples were hydrolysed at 150 °C for 1.5 h by using vapour-phase hydrolysis technique with constant-boiling HCl containing 0.5 % (v/v) phenol.

CD

The far-UV CD spectrum (200–260 nm) was measured on a J-720 automatic spectropolarimeter (JASCO). The spectrum was obtained by using a solution containing the protein at 0.2 mg/ml in the TE buffer in a cell with an optical path of 2 mm at 25 °C. The mean residue ellipticity, [θ], which has the units of degrees$\cdot cm^2\cdot dmol^{-1}$, was calculated by using an average amino acid M_r of 110.

Isoelectric focusing

Isoelectric focusing was carried out at 4 °C for 16 h at 700 V on a 110 ml column (Pharmacia) by the method of Vesterberg [24]. The sample was added in the middle of a 0–50 %-(w/v)-sucrose gradient that contained 1 % carrier Ampholine (pH 3–9).

RESULTS AND DISCUSSION

Overproduction and purification

To facilitate the purification of an esterase from *E. coli*, which is an HSL homologue, its overproducing strain (JM109/pDR319) was constructed. Upon induction of the *ybaC* gene by the addition of IPTG, the enzyme accumulated in cells as one of the most abundant proteins (results not shown). The production level was estimated to be $\sim$ 30 mg/l of culture from the intensity of the band revealed with Coomassie Brilliant Blue staining after SDS/PAGE. Because the amount of the enzyme in the crude extract, which includes only soluble components of cells, was nearly identical with that in the whole-cell extract, *E. coli* esterase accumulated in cells should exist in a soluble form. From the crude extract, the enzyme was easily purified to give a single band on SDS/PAGE by three purification steps (Figure 2). Consequently, 6.0 mg of the pure protein was obtained from a 1-litre culture. Determination of enzymic activity by using PNPB as a substrate indicated that the specific activity of the enzyme was 50 units/mg. By using this value, the amount of the enzyme in the crude extract prepared from a 1-litre culture was calculated as 29.5 mg, which is nearly identical with that estimated from SDS/PAGE. Likewise, the yield of the enzyme in each purification step was calculated as 77 % for $(NH_4)_2SO_4$ precipitation, 38 % for DE-52 column chromatography and 69 % for S-300 column chromatography. Thus the overall purification yield of the enzyme was nearly 20 %. This yield was relatively low, because only the fractions that contained the enzyme with high purity were pooled in each purification step.

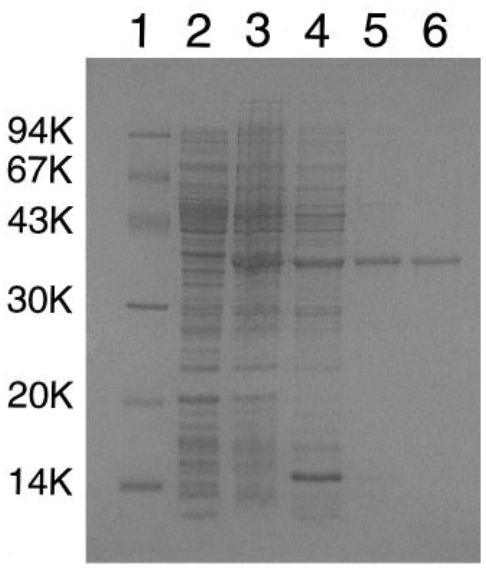

Figure 2 Comparison of the purity of *E. coli* esterase by SDS/PAGE

Samples were subjected to electrophoresis on a 15% polyacrylamide gel in the presence of SDS. After electrophoresis, the gel was stained with Coomassie Brilliant Blue. Lane 1, a low-M_r marker kit (Pharmacia LKB Biotechnology Inc.) containing phosphorylase *b*, BSA, ovalbumin, carbonic anhydrase, trypsin inhibitor and α-lactalbumin (94K = 94000 M_r etc.); lane 2, crude extract from *E. coli* JM109 transformant with pUC19; lane 3, crude extract from *E. coli* JM109 transformant with pDR319; and lanes 4–6, fractions pooled after $(NH_4)_2SO_4$ precipitation, DE-52 column chromatography and Sephacryl S-300 column chromatography respectively.

Enzymic activity

The enzymic activity of the purified *E. coli* esterase was determined by using various *p*-nitrophenyl esters of fatty acids as a substrate at pH 7.1 and 25 °C. These conditions are not optimum for the enzyme. For example, an apparent optimum pH for the hydrolysis of PNPB was 9.0, and the enzymic activity determined at pH 7.1 was roughly half as that determined at pH 9.0. However, this condition was chosen because the stability of the substrates dramatically decreases as the pH increases beyond 7.5 or the temperature increases beyond 30 °C. The kinetic parameters of the enzyme for the hydrolysis of the *p*-nitrophenyl esters of fatty acids with chain length from 2 to 18 are summarized

Table 1 Kinetic parameters of *E. coli* esterase for the hydrolysis of various *p*-nitrophenyl esters of fatty acids

The hydrolysis with the enzyme of various *p*-nitrophenyl esters of fatty acids was carried out at 25 °C for 15 min in 100 μl of 20 mM phosphate buffer, pH 7.1, containing 5% acetonitrile. Kinetic parameters were determined by a least-squares fit of the data obtained from Lineweaver–Burk plots. Errors, which represent the 67% confidence limits, are within 10% of the values reported.

Substrate*		K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m
Acetate	C_2	1.5	30.8	20.5
Propionate	C_3	0.7	30.9	44.1
Butyrate	C_4	0.5	30.9	61.8
Valerate	C_5	0.26	20.5	78.8
Hexanoate	C_6	0.22	6.2	28.2
Octanoate	C_8	0.16	0.21	1.2
Decanoate	C_{10}	–	$\leqslant 0.01$	–
Oleate	C_{18}	–	$\leqslant 0.05$	–

* *p*-Nitrophenyl ester of the fatty acid cited.

in Table 1. Clearly, the K_m value decreased as the aliphatic chain length of the substrate increased, suggesting that the longer the aliphatic chain of the substrate is, the more the enzyme strongly binds to the substrate. The $k_{cat.}$ value decreased as the aliphatic chain length of the substrate increased as well, but only when it increased beyond C_5. The enzymic activity was only poorly detected when *p*-nitrophenyl octanoate (C_8) was used as a substrate, and very little enzymic activity was detected when *p*-nitrophenyl decanoate (C_{10}) or oleate (C_{18}) was used as a substrate. As a C_{18} substrate, *p*-nitrophenyl ester of oleic acid, instead of that of stearic acid, was used, because the former was soluble in the reaction mixture for assay, whereas the latter was not. Comparison of the k_{cat}/K_m values for various substrates indicated that these values were dependent on the aliphatic chain length of the substrate and increased in the following order: $C_8 < C_2 < C_6 < C_3 < C_4 < C_5$. Thus, *p*-nitrophenyl valerate (C_5) was the most preferable substrate for the enzyme. Because lipases prefer substrates with relatively long aliphatic chains, these results suggest that *E. coli* esterase exhibits little lipase activity.

HSL is implicated in a variety of metabolic processes and can cleave various substrates, including triacylglycerols, cholesterol esters and steroid esters [8]. It hydrolyses diacylglycerols, such as 1(3)-oleoyl-2-oleylglycerol, and fatty acid esters, such as PNPB, as well [10,14]. To examine whether *E. coli* esterase can hydrolyse these HSL substrates, we have determined the enzymic activity by using tributyrylglycerol, trioleoylglycerol and cholesterol oleate as a substrate at 30 °C and pH 5.6 in the presence of 0.1% Triton X-100. The enzymic activity towards PNPB determined under these conditions was roughly 60% of that determined at 25 °C and pH 7.1 in the presence of 5% acetonitrile. The results are summarized in Table 2. The enzyme could hydrolyse tributyrylglycerol with the catalytic efficiency of 10.7% of that for the hydrolysis of PNPB. In contrast, it showed little activity for the hydrolyses of trioleoylglycerol and cholesterol oleate. These results indicate that the enzyme cannot hydrolyse the major HSL substrates effectively. Lipases have been shown to be activated at the lipid/water interface (interfacial activation), because a surface loop ('lid'), which covers the active site of the enzyme, changes its conformation, so that the active site is exposed when it interacts with the substrate [25–27]. In contrast, esterases do not have this 'lid'. HSL may have this 'lid' or a related one. Because *E. coli* esterase is small in size compared with HSL, a 'lid' or a related one may be missing in this enzyme.

Table 2 Enzymic activities of *E. coli* esterase for the hydrolysis of tributyrylglycerol, trioleoylglycerol and cholesterol oleate

The hydrolysis of PNPB by the enzyme was carried out at 30 °C for 15 min in 100 μl of 50 mM sodium acetate, pH 5.6, containing 0.1% Triton X-100. The hydrolyses of tributyrylglycerol, trioleoylglycerol and cholesterol oleate were carried out at 30 °C for 60 min with stirring at 500 rev./min in 50 mM sodium acetate, pH 5.6, containing 0.1% Triton X-100. Relative activity was calculated by dividing the specific activity of the enzyme for the hydrolysis of tributyrylglycerol, trioleoylglycerol or cholesterol oleate by that for the hydrolysis of PNPB.

Substrate	Specific activity (units/mg)	Relative activity (%)
PNPB	34.2	100
Tributyrylglycerol	3.67	10.7
Trioleoylglycerol	0.22	0.6
Cholesterol oleate	$\leqslant 0.1$	$\leqslant 0.3$

Because *E. coli* esterase has no signal sequence, it must be localized in the cytoplasm. In fact, when the overproducing strain of *E. coli* esterase (JM109/pDR319) was grown on a tributyrylglycerol L-agar plate, which is often used to screen for micro-organisms that produce lipases, no halo of lipolysis was detected around the colonies. However, its function remains to be determined. No difference was detected in the phenotype of *E. coli* cells when the *ybaC* gene was inactivated by the insertion of the kanamycin-resistance gene, indicating that *E. coli* esterase is not essential for cell growth (K. Miyamoto and H. Inokuchi, unpublished work).

Biochemical characterizations

The M_r of *E. coli* esterase was determined to be 37000 by SDS/PAGE, which was nearly identical with that (36038) calculated from the amino acid sequence predicted from the DNA sequence. The M_r of the enzyme was also estimated to be 38000 on Sephacryl S-300 column chromatography. The good agreement between these values indicates that *E. coli* esterase exists in a monomeric form. To confirm the amino acid sequence of *E. coli* esterase predicted from the DNA sequence, the N-terminal amino acid sequence and amino acid composition of the protein were determined. The sequence of the N-terminal ten amino acid residues was MKPENKLPVL, which was identical with that predicted from the DNA sequence. In addition, there is a good agreement between the amino acid composition predicted from the DNA sequence and that experimentally determined (results not shown). *E. coli* esterase contains 38 acidic residues, 26 basic amino acid residues and six His residues. PI determination indicated that this enzyme is an acidic protein with a pI value of 4.1. This value is comparable with the calculated one (4.83). The far-UV CD spectrum of *E. coli* esterase exhibited a broad trough with two minimum [θ] values of -7200 at 211 nm and -7700 degrees·cm^2·dmol^{-1} at 220 nm (Figure 3). The helical content of this enzyme was calculated as 26.1% by the method of Wu et al. [28].

Active site

Multiple alignment of the amino acid sequences of the proteins involved in the same superfamily is effective in extracting the candidates for the active-site residues, because highly conserved residues are likely to be functionally or structurally important for enzymic activity. When the *E. coli* esterase sequence was compared with that of human HSL, *Moraxella* TA144 lipase 2, and *B. acidocaldarius* ORF3 sequences, conserved residues are

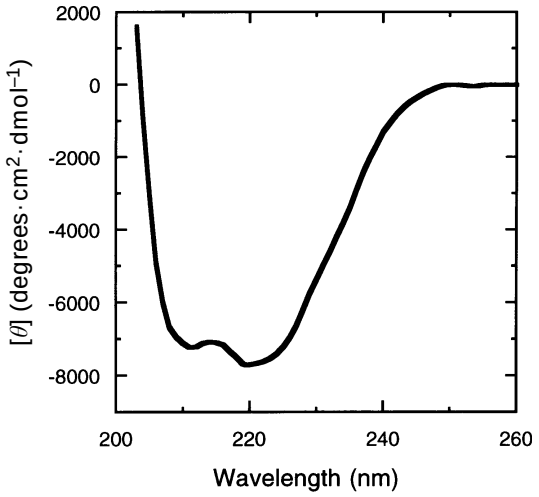

Figure 3 Far-UV CD spectrum of *E. coli* esterase

The spectrum was measured in 10 mM Tris/HCl, pH 8.0 containing 1 mM EDTA at 25 °C as described in the Experimental section.

localized in two regions (Figure 4). Region 1 consists of 95 amino acid residues (residues 83–177 for *E. coli* esterase) and contains HGGG and GDSAG motifs. These motifs have been shown to be characteristic of the HSL family [4]. Because a structural motif GXSXG (where X represents any amino acid), which contains the active-site serine residue, is conserved in most esterases and lipases [16], the GDSAG motif must contain the active-site serine residue. In fact, the corresponding serine residue in rat HSL (Ser^{423}) (Ser^{424} for human HSL) has been shown to be essential for the enzymic activity [10]. Therefore it seems likely that Ser^{165} of *E. coli* esterase, which is located at the centre of the GDSAG motif, is involved in the catalytic triad of the enzyme. Region 2 consists of 59 amino acid residues (residues 250–308 for *E. coli* esterase). This region contains no sequence motif that is characteristic of the HSL family, but seems to contain catalytically essential aspartic acid and histidine residues. Asp^{703} and His^{733} of rat HSL (Asp^{693} and His^{723} for human HSL), both of which are involved in the region 2, have been shown to be essential for the enzymic activity [14]. Therefore it seems likely that the corresponding residues, Asp^{262} and His^{292}, together with Ser^{165}, form the catalytic triad of *E. coli* esterase. We are currently undertaking site-directed-mutagenesis experiments to confirm that these residues are essential for enzymic activity.

A model for the three-dimensional structure of the catalytic domain of HSL has recently been constructed on the basis of the similarity in the secondary-structure elements between HSL and acetylcholinesterase (AChE) [15]. AChE from *Torpedo californica*, in which Ser^{200}, Glu^{327} and His^{440} form the catalytic triad, has been shown to have an α/β hydrolase fold by X-ray-crystallographic analysis [29]. A strong similarity in the geometry of the catalytic triad (Ser-Asp/Glu-His) between these two structures suggests that this geometry is conserved in the structures of *E. coli* esterase and also in the other proteins involved in the HSL family. In addition, it is likely that the main-chain nitrogen atoms of the second and third glycine residues in the HGGG motif are involved in the formation of the oxyanion hole [30]. The amino acid sequence around this motif is LHGGGFI for *E. coli* esterase, FHGGGFV for human HSL, and IYGGGFY for AChE from *T. californica*, where the conserved residues in these

region 1 (95 residues)

```
EcE     (82 aa)--SPATLFYL HGGGFILGNL DTHDRIMRLL ASYSQCTVIG IDYTLSPEAR FPQAIEEIVA
hHSL   (341 aa)--SRSLIVHF HGGGFVAQTS RSHEPYLKSW AQELGAPIIS IDYSLAPEAP FPRALEECFF
MoL2   (156 aa)--DEAAMLFF HGGGFCIGDI DTHHEFCHTV CAQTGWAVVS VDYRMAPEYP APTALKDCLA
Ba3     (72 aa)--PYPALVYY HGGGWVVGDL ETHDPVCRVL AKDGRAVVFS VDYRLAPEHK FPAAVEDAYD

                                    *                                   *
EcE    ACCYFHQQAE DYQINMSRIG FAGDSAGAML ALASALW---(72 aa)--EVPPCFIAGA EFDPLLDDSR
hHSL   AYCWAIKHCA LLGSTGERIC LAGDSAGGNL CFTVALR--(244 aa)--SLPPVHIVAC ALDPMLDDSV
MoL2   AYAWLAEHSQ SLGASPSRIV LSGDSAGGCL AALVAQQ---(97 aa)--QLCPSYIVVA ELDILRDEGL
Ba3    ALQWIAERAA DFHLDPARIA VGGDSAGGNL AAVTSIL---(72 aa)--GLPPAYIATA QYDPLRDVGK

       region 2 (59 residues)      *
EcE    LLYQTLAAHQ QPCEFKLYPG TLHAFLHYSR MMKTADEAL--(11 aa)
hHSL   MLARRLRNLG QPVTLRVVED LPHGFITLAA LCRETRQAA--(36 aa)
MoL2   AYAELLQKEG VQVQTYTVLG APHGFINLMS VHQGLGNQT--(26 aa)
Ba3    LYAEALNKAG VKVEIENFED LIHGFAQFYS LSPGATKAL--(12 aa)
```

Figure 4 The alignment of the *E. coli* esterase, human HSL, *Moraxella* TA144 lipase 2 and *B. acidocaldarius* ORF3 sequences

Only relatively well conserved sequences, which are located in two different regions (regions 1 and 2) of each sequence, are shown. Fully conserved residues in these regions are indicated in **bold**. The putative catalytic residues that are expected to form the catalytic triad (Ser-Asp-His) are denoted by asterisks (*). The numbers of amino acid residues present at the N-terminus of the region 1, those between two regions, and those at the C-terminus of the region 2 are shown in parentheses. The sequences have been deposited in the Swiss-Prot database with code names of YBAC_ECOLI for *E. coli* esterase (EcE), LIPS_HUMAN for human HSL (hHSL), and LIP2_MORSP for *Moraxella* TA144 lipase 2 (MoL2). The nucleotide sequence encoding *B. acidocaldarius* ORF3 (Ba3) has been deposited in the EMBL, GenBank™ and DDBJ Nucleotide Sequence Databases under the accession number X62835.

sequences are underlined. Three glycine residues in a row have been proposed to make the chain flexible enough to allow amide nitrogen atoms from the second and third glycine residues to be part of the oxyanion hole for AChE [29]. However, none of the three-dimensional structures of the HSL family has been so far determined. *E. coli* esterase may be an ideal target for X-ray-crystallographic studies because it is a relatively small globular protein with a monomeric form, and because its overproduction system has been established. Such structural information will help us to understand the evolutionary relationship not only between proteins within the HSL family, but also between proteins in the HSL family and those in the AChE family.

We thank Dr. K. Miyamoto for the plasmid pFC1 and for helpful discussions. We thank Dr. Y. Shimada for measurement of the lipase activity by titration and for helpful discussions. We also thank Dr. H. Toh, Dr. Y. Matuo, Dr. M. Haruki and Dr. H. Inokuchi for helpful discussions.

REFERENCES

1 Bajaj, M. and Blundell, T. (1984) Annu. Rev. Biophys. Bioeng. **13**, 453–492
2 Chothia, C. and Lesk, A. M. (1986) EMBO J. **5**, 823–826
3 Langin, D. and Holm, C. (1993) Trends Biochem. Sci. **18**, 466–467
4 Hemila, H., Koivula, T. T. and Palva, I. (1994) Biochim. Biophys. Acta **1210**, 249–253
5 Miyamoto, K., Nakahigashi, K., Nishimura, K. and Inokuchi, H. (1991) J. Mol. Biol. **219**, 393–398
6 Blattner, F. R., Plunkett, III, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F. et al. (1997) Science **277**, 1453–1461
7 Miyamoto, K., Kanaya, S., Morikawa, K. and Inokuchi, H. (1994) J. Biochem. (Tokyo) **115**, 545–551
8 Yeaman, S. J. (1990) Biochim. Biophys. Acta **1052**, 128–132
9 Holm, C., Kirchgessner, T. G., Svenson, K. L., Fredrikson, G., Nilsson, S., Miller, C. G., Shively, J. E., Heinzmann, C., Sparkes, R. S., Mohandas, T. et al. (1988) Science **241**, 1503–1506
10 Holm, C., Davis, R. C., Osterlund, T., Schotz, M. C. and Fredrikson, G. (1994) FEBS Lett. **344**, 234–238
11 Langin, D., Laurell, H., Stenson Holst, L., Belfrage, P. and Holm, C. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 4897–4901
12 Osterlund, T., Danielsson, B., Degerman, E., Contreras, J. A., Edgren, G., Davis, R. C., Schotz, M. C. and Holm, C. (1996) Biochem. J. **319**, 411–420
13 Smith, G. M., Garton, A. J., Aitken, A. and Yeaman, S. J. (1996) FEBS Lett. **396**, 90–94
14 Osterlund, T., Contreras, J. A. and Holm, C. (1997) FEBS Lett. **403**, 259–262
15 Contreras, J. A., Karlsson, M., Osterlund, T., Laurell, H., Svensson, A. and Holm, C. (1996) J. Biol. Chem. **271**, 31426–31430
16 Brenner, S. (1988) Nature (London) **334**, 528–530
17 Kanaya, S., Kohara, A., Miura, Y., Sekiguchi, A., Iwai, S., Inoue, H., Ohtuka, E. and Ikehara, M. (1990) J. Biol. Chem. **265**, 4615–4621
18 Miller, J. H. (ed.) (1972) Experiments in Molecular Genetics, p. 433, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
19 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
20 Laemmli, U. K. (1970) Nature (London) **227**, 680–685
21 Nakamura, Y. and Yura, T. (1975) J. Mol. Biol. **97**, 621–642
22 Sugihara, A., Tani, T. and Tominaga, Y. (1991) J. Biochem. (Tokyo) **109**, 211–216
23 Goodwin, T. W. and Morton, R. A. (1946) Biochem. J. **40**, 628–632
24 Vesterberg, O. (1971) Methods Enzymol. **22**, 389–412
25 Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G. G., Tolley, S., Turkenburg, J. P., Christiansen, L., Huge-Jensen, B., Norskov, L. et al. (1990) Nature (London) **343**, 767–770
26 Winkler, F. K., D'Arcy, A. and Hunziker, W. (1990) Nature (London) **343**, 771–774
27 Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A. and Thim, L. (1991) Nature (London) **351**, 491–494
28 Wu, C.-S. C., Ikeda, K. and Yang, J. T. (1981) Biochemistry **20**, 566–570
29 Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. and Silman, I. (1991) Science **253**, 872–879
30 Steitz, T. A. and Shulman, R. G. (1982) Annu. Rev. Biophys. Bioeng. **11**, 419–444

Received 2 December 1997/9 February 1998; accepted 13 February 1998