Substrate Specificity of *Staphylococcus hyicus* Lipase and *Staphylococcus aureus* Lipase As Studied by in Vivo Chimeragenesis[†]

Muriel D. van Kampen, Niek Dekker, Maarten R. Egmond, and Hubertus M. Verheij*

Department of Enzymology and Protein Engineering, Centre for Biomembranes and Lipid Enzymology, Institute of Biomembranes, Utrecht University, P.O. Box 80054, 3508 TB Utrecht, The Netherlands

Received October 14, 1997; Revised Manuscript Received December 24, 1997

ABSTRACT: *Staphylococcus hyicus* lipase (SHL) and *Staphylococcus aureus* lipase (SAL) are highly homologous enzymes, yet they show remarkable differences in their biochemical characteristics. SHL displays a high phospholipase activity, hydrolyses neutral lipids, and has no chain length preference, whereas SAL only degrades short-chain fatty acid esters. To identify the regions in the primary sequence of SHL responsible for phospholipase activity and chain length selectivity, a set of histidine-tagged SAL/SHL chimeras was generated by in vivo recombination in *Escherichia coli*. Several classes of chimeric enzymes were identified on the basis of restriction site analysis. All chimeras were well-expressed as active enzymes. They were characterized for their specific activities on both phospholipids and *p*-nitrophenyl esters of various chain lengths. Phospholipase activity appeared to be determined by three regions, all located in the C-terminal domain of SHL. Testing of the enzymatic activity of the chimeras toward *p*-nitrophenyl esters showed that chain length selectivity is defined by elements within the region of residues 180−253. Moreover, also residues along the stretch 275−358 contribute to the binding of acyl chains. Interestingly, several chimeras were even more active than the parent enzymes on long-chain *p*-nitrophenyl esters.

Lipases (glycerol ester hydrolase EC 3.1.1.3) and bacterial lipases in particular are widespread in nature. These enzymes are active at a lipid−water interface and have as their primary activity the hydrolysis of neutral lipids such as triglycerides. They generally display a broad substrate specificity and also degrade water-soluble and insoluble esters, Tweens, and exceptionally phospholipids. Many Staphylococci are known to produce extracellular lipases, which are all expressed as pre−pro enzymes and have molecular masses of approximately 70 kDa. After secretion into the medium, the proenzyme is processed by an extracellular protease generating the 40−46 kDa mature lipase form. Thus far, the lipase genes of *S. hyicus*, *S. aureus* strains *PS54* and *NCTC8530* and *S. epidermidis* have been cloned and sequenced (*1*−*4*). Their mature parts are closely related, suggesting a comparable tertiary structure.

Alignment of *S. hyicus* lipase (SHL)[1] and *S. aureus* strain NCTC8530 lipase (SAL) reveals 50% identity and 64% sequence similarity at the amino acid level. Despite this homology, comparative studies of SAL and SHL reveal remarkable differences in pH optimum, in Ca^{2+} affinity, and, most importantly, in substrate specificity and chain length selectivity (*4, 5*). SAL readily degrades short-chain triacylglycerols and *p*-nitrophenyl esters, whereas medium- and long-chain lipids, as well as phospholipids, are poor substrates. In contrast, SHL is unique because it has a high phospholipase activity, which distinguishes the enzyme not only from the other staphylococcal enzymes but also from all other bacterial lipases. Moreover, SHL hydrolyzes neutral lipids almost irrespective of their chain length. In literature, a number of structurally homologous lipases has been described displaying different substrate selectivities. Examples are the couples hepatic lipase/lipoprotein lipase (*6*), pancreatic lipase related proteins/classical human pancreatic lipase (*7*), and lipase I/lipase II from *Geotrichum candidum* (*8*). In lipases, surface loops and helixes are present that block the entrance to the active site in aqueous environment. It is generally accepted that these regions (often referred to as the lid) move away upon interaction of the lipase with an interface thus explaining interfacial activation. Moreover, it has been proposed that the lid region is involved in interfacial interaction and substrate recognition. Based on X-ray structural data, an approach of domain exchange has been used to investigate the origin of substrate selectivity for hepatic lipase/lipoprotein lipase and pancreatic lipase related proteins/classical human pancreatic lipase (*9, 10*). Although these studies primarily focused on the role of the lipase lid domain and on the C-terminus of the enzyme, the outcome of these studies is that although the lid plays a major role in substrate selectivity, contribution from other parts of the protein cannot be ruled out.

[†] This research has been supported by a grant from the EC (Grant No. BIO2 CT94-3013). The work was carried out under auspices of the Dutch Foundation for Chemical Research (SON) with financial aid from the Dutch Organization for Scientific Research (NWO).

* To whom correspondence should be addressed.

[1] Abbreviations: C_{16}thioPC, [2-(hexadecanoylthio)ethyl]-1-phosphocholine; diC_6thioPC, *rac*-1,2-dihexanoyldithiopropyl-3-phosphocholine; pNPC_4, *p*-nitrophenylbutyrate; SHL, *Staphylococcus hyicus* lipase; SAL, *Staphylococcus aureus* lipase.

S0006-2960(97)02543-9 CCC: $15.00
Published on Web 02/20/1998

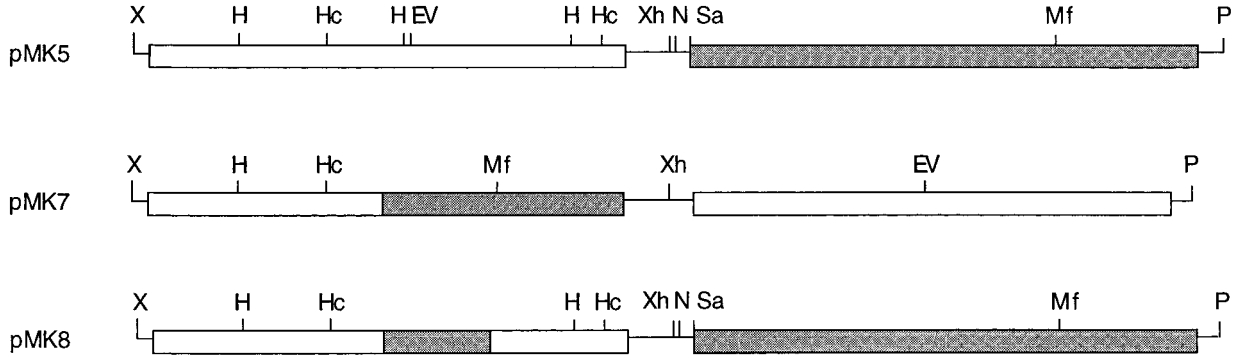

FIGURE 1: Tandem plasmids for chimeragenesis. The plasmids are composed of SAL sequence (white boxes) and SHL sequence (grey boxes): EV, *EcoR* V; H, *Hind* III; Hc, *Hinc* II; Mf, *Mfe* I; N, *Nhe* I; P, *Pst* I; Sa, *Sal* I; X, *Xba* I*; Xh, Xho* I.

For Staphylococcal lipases, no X-ray structure is available and sequence homology with other lipase families is too low for modeling of the complete structure. Therefore, a directed approach for structure−function analysis of these enzymes based on an X-ray model is impossible. On the other hand, it can be questioned to which extent this lack of structural information is a real drawback. During the opening of the lid several interactions of the lid residues with the core of the protein are lost, while other interactions with the lipase surface, with the interface, and with the substrate molecule are established. Detailed engineering of substrate specificity might thus be beyond our present level of understanding of these interactions. We have chosen for a different approach which does not depend on structural information and which takes advantage of the fact that two enzymes have evolved differently. The homology between SAL and SHL enabled us to use a random recombination method for the generation of SAL/SHL chimeras, which were analyzed to identify structural regions that define phospholipase activity and chain length selectivity. This technique makes use of in vivo recombination events between homologous regions of the two genes and has proven to be useful for structure−function analysis of proteins (*11*−*14*). All SAL/SHL chimeras were tested for activity on phospholipids and *p*-nitrophenyl esters. The results indicate that several regions within the C-terminal half of SHL contain essential elements for the high phospholipase activity of this enzyme. Moreover, a limited number of residues located in the central part of the sequence appears to be important for chain length selectivity.

MATERIALS AND METHODS

Chemicals. Oligonucleotides were purchased from Pharmacia. [2-(Hexadecanoylthio)ethyl]-1-phosphocholine (C_{16}-thioPC) and *rac*-1,2-dihexanoyldithiopropyl-3-phosphocholine (diC_6thioPC) were synthesized as described before (*15, 16*). *p*-Nitrophenyl esters were obtained from Sigma. Nickel nitrilotriacetate was obtained from Invitrogen. All other chemicals were of analytical grade.

Bacterial Strains and Media. Bacterial strains used were DH5α (F$^-$, *supE*44, Δ*lacU*169 (*ϕ*80*lacZ*Δ*M*15), *hsd*R17, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *rel*A1 (*17*)); JC8679 (F$^-$, *thr-1*, *ara-14*, *leuB*6, *thi*-1, *lac*Y1, *gal*K2, *xyl*-5, *mtl*-1, *pro*A2, *his*4, *arg*E3, *tsx*-33, *sup*E44, *λ*$^-$, *λ*^{s}, *rec*B21, *rec*C22, *sbc*A23 (*18*)), and JM103 (*F′*, *tra*D36, *lacI*^{q}, Δ(*lacZ*)*M*15, *pro*A$^+$B$^+$/*end*A1, *sup*E, *sbc*B15, *thi*-1, *rps*L(*str*R), Δ(*lac-pro*)(*P*1), (*rK*$^+$*mK*$^+$*rP*$_1^+$*mP*$_1^+$) (*19*)). Cells were grown in Luria−Bertani medium supplemented with ampicillin (0.1 mg/mL) at 37 °C.

DNA Manipulations and Construction of Plasmids. All enzymes for DNA manipulations were purchased from New England Biolabs and applied according to the manufacturer's instructions. *E. coli* strain DH5α was used for all plasmid constructions. The construction of plasmids pH6SHT7, pH6SAB7, and pSHT7 has been described before (*5, 20*). The *Xba* I/*PstI* I fragment of plasmid pH6SHT7, encoding the histidine-tagged mature SHL, was cloned into the corresponding sites of plasmid pUC18, resulting in plasmid pMK1. Plasmid pUC18 was digested with *Hind* III, and the resulting 3′ overhangs were filled in. The plasmid was religated, thereby removing the *Hind* III site and introducing a *Nhe* I site, resulting in pUC18-H. The *Xba* I/*Pst* I fragment of plasmid pH6SAB7, encoding the histidine-tagged mature SAL, was cloned into the corresponding sites of this plasmid, resulting in pMK2. A synthetic linker of oligonucleotides 5′-CTCGAGCTAGCAACTGCAGC-3′ and 5′-CTGGCTG-CAGTTGCTAGCTCGAGTGCA-3′ was inserted into *Pst* I/*Nhe* I digested pMK2, thereby introducing an *Xho* I site and *Nhe* I and *Pst* I sites in reversed order. This plasmid was designated pMK3. After digestion of pMK3 with *Nhe* I and *Pst* I, the *Xba* I/*Pst* I fragment of plasmid pSHT7, encoding the mature SHL, was inserted. The resulting plasmid was digested with *Sal* I and an additional linker of oligonucleotides 5′-TCGACTGACTAAGCTAGCTTAG-3′ and 5′-TCGACTAAGCTAGCTTAGTCAG-3′ was inserted in order to introduce stop codons in all reading frames, yielding plasmid pMK5 (Figure 1). The *Mlu* I/*Pst* I fragment of plasmid 13 (class II, obtained from in vivo chimeragenesis) was replaced by the *Mlu* I/*Pst* I fragment of plasmid pMK3, resulting in plasmid pMK3b. Subsequently, the *Xba* I/*Pst* I fragment of plasmid 13 (class II) was inserted into the corresponding sites of the plasmid pUC18-H. Additionally, the same procedure as that described for pMK5 was initially followed, i.e. insertion of the first synthetic linker and digestion with *Nhe* I/*Pst* I. The *Xba* I/*Pst* I fragment of plasmid pH6SAB7 was inserted into these sites, resulting in plasmid pMK7 (Figure 1). The *Hind* III fragment of plasmid 29 (class V) was subcloned into the corresponding sites of plasmid pMK5. The correct orientation of the fragment was checked by *Hinc* II digestion. The new plasmid was designated pMK8 (Figure 1).

Chimeragenesis. Plasmids were linearized by digestion with *Xho* I/*Nhe* I (pMK5), *Xho* I/*EcoR* V (pMK7), and *Xho* I/*Mfe* I (pMK8) for 5 h at 37 °C to ensure complete digestion.

The linearized plasmid pMK5 (1 μg) was transformed to *E. coli* strain DH5α using the calcium chloride method (*21*). Linearized plasmids pMK7 and pMK8 were purified from agarose gel prior to transformation using QIAEX according to manufacturer's instructions. The recombination sites of the chimeras were determined by DNA sequencing using the dideoxy chain termination method (*22*).

Expression. *E. coli* strain DH5α was transformed with the appropriate chimeric construct. Transformants were used to inoculate overnight cultures (10 mL). Cells were collected by centrifugation and dissolved in 1 mL of the application buffer (6 M guanidine·HCl, 10 mM Tris/HCl, pH 8.2, 5 mM imidazole). After centrifugation of the lysate (10 min, 16000*g*), 10 μL aliquots of the clear supernatant were taken for activity measurements. Nickel nitrilotriacetate resin (150 μL of 50% slurry (QIAGEN), equilibrated in this buffer) was added to the remaining solution followed by 4 h incubation at 4 °C on a rotating wheel. Subsequently, the resin was washed twice for 15 min at 4 °C with 1 mL of wash buffer (10 mM Tris/HCl, 5 mM imidazole). Finally, the protein was eluted by addition of 300 μL of sodium dodecyl sulfate polyacrylamide gel electrophoresis loading buffer, containing 0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol, and by subsequent boiling of the samples for 3 min. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS−PAGE) was performed with 5−10 μL aliquots of the protein samples using 15% acrylamide gels with Coomassie brilliant blue staining (*21*). Lipase amounts were determined by scanning densitometry, and these numbers were used to calculate specific activities. We prefer to give specific activities rather than absolute activity numbers per 10 mL of culture because the former values are independent of expression levels and cell densities.

Enzymatic Activity Measurements. The phospholipase activity of cells harboring the appropriate construct was qualitatively detected as a zone of white deposits around the colony on the plate, due to the precipitation of liberated fatty acids essentially as described before (*2*). LB agar medium plates were prepared containing 1% lysophosphatidylcholine, added from a 25% stock solution in 25% ethanol. Enzymatic activities of cell lysates were determined spectrophotometrically at pH 8.0 (50 mM Tris/HCl) in the presence of 10 mM $CaCl_2$ at room temperature. Activity toward *p*-nitrophenyl butyrate was determined in mixed micelles with Triton X-100 (0.75 mM substrate in 5 mM Triton X-100) in initial experiments. For chain length selectivity measurements the substrates *p*-nitrophenyl butyrate, octanoate, laurate, and palmitate were used at a concentration of 1 mM in 100 mM Triton X-100. The formation of the *p*-nitrophenoxide ion was monitored at 400 nm. Activity on long-chain phospholipids was quantitatively determined in mixed micelles using C_{16}thioPC as a substrate (0.25 mM in 0.2 mM Triton X-100). Activity on short-chain phospholipids was assayed on pure micelles of diC_6thioPC (3mM). In both phospholipase assays, 0.2 mM 5,5′-dithiobis(2-nitrobenzoic acid) was included and the increase in absorption was followed at 412 nm. Activities are given in Units (1 Unit is the conversion of 1 μmol of substrate/min) using molar absorption coefficients of 16 888 M^{-1} cm^{-1} (*p*-nitrophenoxide) and 13 600 M^{-1} cm^{-1} (2-nitro-5-thiobenzoic acid anion).

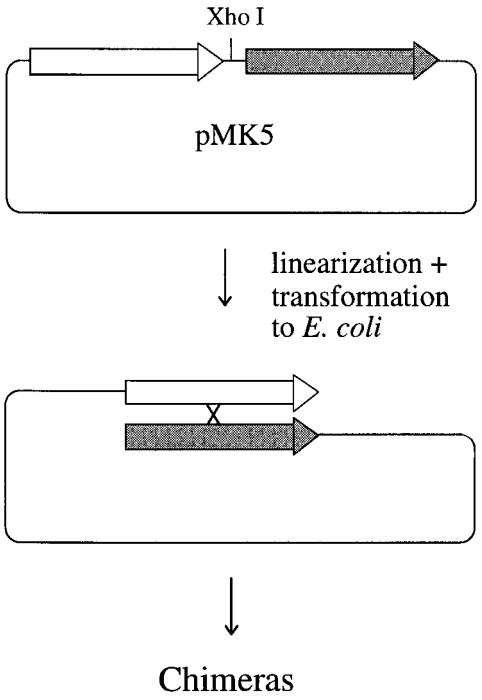

FIGURE 2: Generation of SAL/SHL chimeric genes by in vivo chimeragenesis. Plasmid pMK5 containing SAL (white arrow) and SHL (grey arrow) in a tandem array was linearized with *Xho* I and was used to transform *E. coli*. Recombination between homologous regions resulted in a set of plasmids encoding chimeras with fusion sites distributed throughout the gene sequence.

RESULTS

In Vivo Chimeragenesis. To identify regions that confer phospholipase activity and chain length selectivity to SHL, chimeras of SAL and SHL were generated by in vivo chimeragenesis. The principle of this method is depicted in Figure 2. First, a plasmid was constructed containing the two genes tandemly arranged in the same orientation under the control of the *lac* promoter. After digestion of the plasmid at restriction sites located between the two genes, the linear DNA was used for transformation of an *E. coli* host. Since the SAL and SHL sequences are highly homologous, recombination events can take place between the two genes, resulting in the formation of SAL/SHL chimeric genes. The mechanism of the recombination process is believed to be mediated by an *E. coli* correction/repair system but is not clearly understood. The process of in vivo chimeragenesis has been described as occuring in both $recA^+$ and $recA^-$ backgrounds (*23*). Furthermore, it has been shown that intramolecular recombination of linear DNA can be catalyzed by the *E. coli* recE recombination system (*24*). For the generation of SAL/SHL chimeras two contrasting requirements are important: on one hand the bacterial strain should be efficient in recombination, and on the other hand it should produce chimeras with fusion sites randomly distributed throughout the gene sequence. We examined *E. coli* strains DH5α ($recA^-$), JM103 ($recA^+$), and JC8679 ($recA^+recE^+$) as chimeragenesis hosts. Each strain was transformed with *Xho* I/*Nhe* I linearized plasmid pMK5. As expected, the transformation efficiency of the linear DNA was greatly reduced for all strains to values of 0.2−1.3% compared to that of the intact circular plasmid. Thirty transformants of each strain were arbitrarily selected, and their plasmid DNA was analyzed by digestion with *Xba* I/*Pst*

Table 1: Transformation Efficiency of *E. coli* Strains and Analysis of Recombination Products

strain	no. of transformants (1 μg of linear DNA)	true recombination (% of transformants)
DH5α	640	30
JM103	610	67
JC8679	1330	93

I. A fragment with the size of one gene (i.e. 1250 bp) indicated true chimera formation, while the remaining plasmids were associated with larger products. The results are summarized in Table 1. Evidently, in all three strains recombination took place, and the efficiency depended on the genotype of the host. To address the question of whether the recombination site was randomly distributed, restriction analysis with *Msc* I, which cuts about in the middle of the SHL gene (Figure 3) was performed. These results showed that recombination is biased in strains JC8679 and JM103. We found fusion sites within the 5′-half of the gene in respectively 75% and 65% of the products, whereas this phenomenon is less strict in DH5α (58% of the products). Recombination events in strain JC8679 were biased to occur within regions of relatively high homology of eight or more contiguous nucleotides. On the other hand, DH5α required shorter stretches (five or more nucleotides) for recombination, thus generating a larger diversity of products. Therefore we decided to use DH5α as the host strain for later experiments.

The true chimeric products from the three strains were further analyzed by digestion with restriction enzymes that cleave either in SHL (*Msc* I, *EcoR* I, *Ssp* I) or SAL (*Hind* III, *Hinc* II, *EcoR* V). On the basis of their restriction patterns, the chimeras were arbitrarily divided into four classes. The composition of each class is shown in Figure 3, where the approximate position of the fusion site is indicated by arrows. The precise recombination sites of a number of representatives of the four classes were determined by DNA sequencing. The composition of these individual products, which were used for further studies, is shown in Table 2, in which the numbering refers to the amino acid numbering of the SHL/SAL alignment shown in Figure 4.

Small-scale expression of the chimeric constructs and SDS−PAGE showed that the cells produced 150−900 μg of lipase/(10 mL of culture). To test if these chimeras were enzymatically active, the activity in the lysate of cells expressing the chimeric constructs was determined toward *p*-nitrophenyl butyrate (0.75 mM $pNPC_4$ in 5 mM Triton X-100). Lysates of pMK1- (SHL), pMK2- (SAL), and pUC18-transformed cells were also tested for activity. No activity could be detected in cells harboring pUC18, whereas SHL, SAL, and all chimeric enzymes showed a clear activity toward $pNPC_4$. Specific activities were calculated on the basis of these activities and the estimates of the amount of lipase determined by SDS−PAGE (see Materials and Methods). The results are summarized in Table 2. It should be stressed that all class I−IV chimeras displayed a reduction in specific activity as compared to both parent enzymes. Such a reduction could be the result of a less good packing of regions originating from two different parents. However, the fact that all clones show activity indicates that there is sufficient homology to generate chimeras which are well-folded.

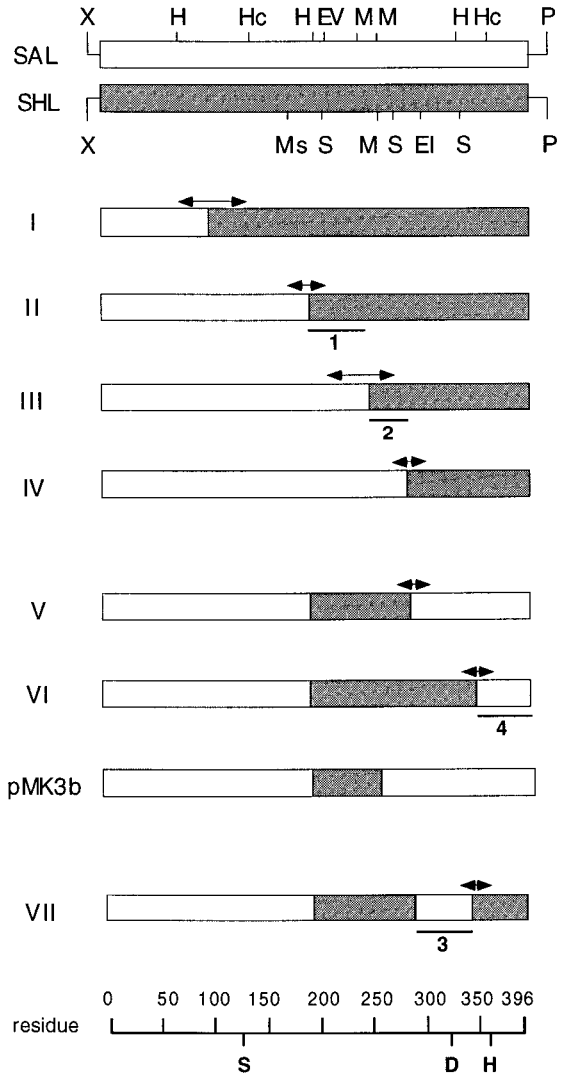

FIGURE 3: Classification of SAL/SHL chimeric genes. On the basis of their restriction pattern, the chimeric genes were divided into classes. The arrows indicate the approximate position of the fusion site. Fusion sites are located between H/Hc (I), Ms/EV (II), EV/S (III), S/EI (IV), S/EI (V), S/Hc (VI), and S/Hc (VII). EI, *EcoR* I; EV, *EcoR* V; H, *Hind* III; Hc, *Hinc* II, M, *Mlu* I; Mf, *Mfe* I; Ms, *Msc* I; P, *Pst* I; S, *Ssp* I; X, *Xba* I; *Xh*, *Xho* I. Regions 1−4, indicated by the lines, are in the corresponding protein sequence: 1, residues 180−253; 2, 254−274; 3, 275−358; 4, 359−399. Active site residues are indicated in bold at the bottom of the figure.

Although it is important that the chimeras retain lipase activity, a more interesting property is their phospholipase activity. The phospholipase activity of the chimeras was first qualitatively determined by an in vivo plate assay with lysophosphatidylcholine as the substrate. Class I and class II chimeras, but not class III and IV chimeras, displayed phospholipase activity on the long-chain lysophospholipids present in the plate (data not shown). Subsequently, a quantitative determination of phospholipase activity was performed by the spectrophotometric assay using C_{16}thioPC as a substrate on which SHL displays a high phospholipase activity, while SAL does not. In agreement with the results of the plate assay, class I and II chimeras appeared to be highly active on this chromogenic phospholipid. However, class III and IV chimeras, which were negative in the plate assay, exhibited a low phospholipase activity in this sensitive assay (0.4−2.6% compared to SHL). Specific activities and

Table 2: Composition, Enzymatic Activities, and Phospholipase/Lipase Ratios of SHL, SAL, and Chimeras[a,b]

class	plasmid	composition[c] (SHL residues)	$pNPC_4$ (U/mg)	C_{16}thioPC (U/mg)	C_{16}thioPC/ $pNPC_4$	diC_6thioPC (U/mg)	diC_6thioPC/$pNPC_4$ (×1000)
SHL	pMK1	1−399	132	230	2	6	47
SAL	pMK2		48	<0.01	0	0.3	6
I	67	63−399	10	74	7	0.5	55
	46	73−399	15	27	2	0.8	53
	43	108−399	10	55	6	0.7	67
II	13	180−399	23	95	4	1	43
	34	180−399	28	137	5	1	50
	105	180−399	35	157	5	2	45
III	2	254−399	12	4	0.3	0.5	43
IV	107	275−399	17	6	0.4	0.1	8
	123	280−399	8	1	0.1	0.1	7
V	25	180−274	87	<0.01	0	0.4	5
	28	180−274	89	<0.01	0	0.4	4
	29	180−279	87	<0.01	0	0.5	6
VI	32	180−349	35	1	0.03	0.1	3
	89	180−358	33	3	0.09	0.4	11
	94	180−351	34	2	0.06	0.1	2
pMK 3b		180−253	96	<0.01	0	0.1	5
VII	6	180−279 + 346−399	157	1	0.007	ND	ND
	11	180−279 + 346−399	157	1	0.007	ND	ND
	14	180−279 + 350−399	195	1	0.006	ND	ND

[a] ND: not determined. [b] Accuracy of specific activities was about 25%; accuracy of phospholipase/lipase ratios was about 10%. [c] Composition of individual chimeras as determined by DNA sequencing. The numbering refers to the amino acid numbering used in the alignment shown in Figure 4 and indicates which part of the chimeras is made up of SHL.

phospholipase/lipase ratios were calculated, and the results are shown in Table 2.

Inspection of the data in Table 2 shows that the phospholipase/lipase ratio (C_{16}thioPC/$pNPC_4$) has increased in class I and class II chimeras and has decreased in class III and IV chimeras as compared to the ratio for SHL. This means that there is no linear reduction in this ratio when increasingly more SAL residues are introduced into the SHL sequence. This phenomenon is particularly clear when one looks at the results obtained with clones 67, 46, and 43. Although they all have been assigned to class I, clone 46 differs from clones 67 and 43 by four residues and clone 67 differs from 43 by eight residues (cf. Table 2). It should also be noted that chimeras with phospholipase activity result from recombination either well before (all three class I clones) or well after (class II) the domain carrying the active site serine residue. The addition of another 74 or more amino acids from SAL (classes III and IV) causes a dramatic decrease in the phospholipase/lipase ratio. It is interesting to note that clones 107 and 123 of class IV differ by only a single substitution (Val/Glu), and this substitution is reflected in a 4-fold decrease in the C_{16}thioPC/$pNPC_4$ ratio.

Recently it has been shown that apart from the difference in phospholipase activities, SAL and SHL are quite different in chain length selectivity with SAL being mainly active on short-chain substrates (*5*). Thus, it can be questioned whether the observed C_{16}thioPC/$pNPC_4$ ratios really reflect the recognition of the polar head group or that they merely reflect chain length selectivity. To test this possibility we made use of the short-chain phospholipid diC_6thioPC. In contrast to the results obtained with the long-chain phospholipid C_{16}thioPC, class I−III chimeras now behave similarly (Table 2). However, a break in phospholipase activity with diC_6thioPC is observed upon going from class III to class IV chimeras. In summary, our data suggest that the residues within the region 180−253 (region 1, Figure 3) contain important information for chain length selectivity. The fact that the phospholipase activity is lost in going from class III to class IV suggests that one or more residues along the stretch 254−274 in SHL (region 2, Figure 3) are crucial for phospholipase activity.

Second Generation Chimeragenesis. At this point it was clear that the C-terminal region of SHL contains essential elements for phospholipase activity. To obtain insight in the residues at the C-terminal end which are essential, a strategy was developed to interchange specifically this part of SHL with SAL elements and *vice versa*. To this end, using direct cloning procedures, plasmid pMK3b (Figure 3) was constructed which contains the elements for recognition of long-chain substrates. In addition, starting from class II chimera 13, which displays phospholipase activity, tandem plasmid pMK7 was constructed. By in vivo chimeragenesis, part of the SAL sequence was incorporated at the C-terminus of this chimera, resulting in the generation of class V and VI chimeras (Figure 3). In a "complementary" approach, tandem plasmid pMK8 was constructed to yield class VII after in vivo chimeragenesis (Figure 3). Representatives of the obtained chimeric genes of classes V−VII were sequenced, and their composition is shown in Table 2.

Subsequently, the activity of the chimeras in cell lysates was determined, showing that all of these second generation chimeras were active on $pNPC_4$. It should be noticed that compared to class I−IV chimeras now the lipase activities are high and even exceed those of the parent enzymes. However, phospholipase activity was found to be very low or absent, resulting in phospholipase/lipase ratios below 0.1 (Table 2). Both for long-chain and for short-chain phospholipids these ratios resemble those of SAL. From the observation that class VI chimeras have lost phospholipase activity, it is clear that elements within the C-terminus of SHL (residues 359−399, region 4, Figure 3) are required for phospholipase activity. Similarly, class VII chimeras are hardly active on long-chain phospholipid substrates, implying that also residues along the stretch 275−358 (region 3, Figure

```
SHL   vkAAPEAVQN PENPKNKDPF VFVHGFTGFV GEVAAKG-EN
SAL   kdDQTNKVAK QGQYKNQDPI VLVHGFNGFT DDINPSVLAH   40
            *     o ** **  * **** **   oo

SHL   HWGGTKANLR NHLRKAGYET YEASVSALAS NHERAVELYY
SAL   YWGGNKMNIR QDLEENGYKA YEASISAFGS NYDRAVELYY   80
       *** * *o* o *   **   ****o** o* * o*******

SHL   YLKGGRVDYG AAHSEKYGHE RYGKTYEGVL KDWKPGHPVH
SAL   YIKGGRVDYG AAHAAKYGHE RYGKTYEGIY KDWKPGQKVH  120
      *o********* ***  ***** ********o  ******  **

SHL   FIGHSMGGQT IRLLEHYLRF GDKAEIAYQQ QHGGIISELF
SAL   LVGHSMGGQT IRQLEELLRN GNREEIEYQK KHGGEISPLF  160
       o********* ** **  **  * o ** **   *** ** **

SHL   KGGQDNMVTS ITTIATPHNG THASDDIGNT PTIRNILYSF
SAL   KGNHDNMISS ITTLGTPHNG THASDLAGNE ALVRQIVFDI  200
      **  ***oo* ***oo***** *****  **    o*o*o

SHL   AQMSSHLGT- IDFGMDHWGF KRKDGESLTD YNKRIAESKI
SAL   GKMFGNKNSR VDFGLAQWGL KQKPNESYID YVKRVKQSNL  240
      o *     o  o***o  **  * *  **  * * **o  * o

SHL   WDSEDTGLYD LTREGAEKIN QKTELNPNIY YKTYTGVATH
SAL   WKSKDNGFYD LTREGATDLN RKTSLNPNIV YKTYTGEATH  280
      * * * * ** ******  o*  ** *****  ****** ***

SHL   ET-QLGKHIA DLGMEFTKIL TGNYIGSVDD ILWRPNDGLV
SAL   KALNSDRQKA DLNMFFPFVI TGNLIGKATE KEWRENDGLV  320
         o  o  * ** * *  oo *** **   o   ** *****

SHL   SEISSQHPSD EKNISVDENS ELHKGTWQVM PTMKGWDHSD
SAL   SVISSQHPFN QAYTKATDK- -IQKGIWQVT PTKHDWDHVD  360
      * ******          o    o ** ***  **   *** *

SHL   FIGNDALDTK HSAIELTNFY HSISDYLMRI EKAESTKNA
SAL   FVGQDSSDTV RTREELQDFW HHLADDLVKT EKLTDTKQA   399
      *o*o*  **   o  **  *o * o * *oo  **   **o*
```

FIGURE 4: Sequence alignment of SHL and SAL. Identical amino acids are indicated by an asterisk, conservative replacements by an open circle (M = L = I = V, Y = F = W, S = T, A = G, K = R, N = Q, E = D). Active site residues (Ser125, Asp317, and His358) are in bold letters. Plasmids pMK1 and pMK2 encode the histidine-tagged mature lipase starting with Ala3 (SHL) and Asp3 (SAL). For ease of comparison the same numbering of amino acids is used for SHL and SAL.

3) play a role in the high phospholipase activity of SHL.

Chain Length Selectivity. On the basis of the different activities on long-chain and on short-chain phospholipids we were able to localize a binding region for long acyl chains with the class I−IV enzymes. To investigate chain length selectivity in more detail, the activity of the chimeras was tested with *p*-nitrophenyl esters of various chain length which have been shown to be good substrates for lipases in general and, as it turned out, also for most chimeras. Since the longer chain esters *p*-nitrophenyl laurate and palmitate only dissolve at high Triton X-100 concentrations, all activity measurements were now performed with 1 mM substrate in 100 mM Triton X-100. Activity ratios of octanoate/butyrate, laurate/butyrate, and palmitate/butyrate and the percentages of activity compared to SHL were calculated, and the results are shown in Table 3. It is clear that these ratios decrease for SHL and SAL in the order octanoate, laurate, to palmitate. SHL is only 4 times less active on the long-chain palmitate compared to the short-chain butyrate, whereas SAL shows a 37-fold decrease for the long-chain substrate. All chimeric classes containing the central region of SHL are highly active on $pNPC_{16}$, whereas classes III and IV (not containing this region) are almost inactive on this substrate. Obviously,

Table 3: Chain Length Selectivity of SHL, SAL, and Chimeras toward Acyl *p*-Nitrophenyl Esters[a]

class	% activity	octanoate/ butyrate	laurate/ butyrate	palmitate/ butyrate
SHL	100	0.625	0.357	0.250
SAL	36	0.333	0.120	0.027
I	8	0.357	0.270	0.277
II	21	0.213	0.120	0.132
III	9	0.071	0.050	0.037
IV	13	0.059	0.053	0.030
V	67	0.625	0.476	0.476
VI	26	0.278	0.167	0.147
pMK3b	73	1.000	0.833	0.769
VII	121	0.434	0.335	0.318

[a] Accuracy was about 10% for each given value. [b] Based on $pNPC_4$ activities (see Table 2).

residues within the central region are essential for the recognition of long acyl chains. Furthermore, going from class VI to V and pMK3b shows that stepwise introduction of the SAL sequence results in an increasing palmitate/butyrate ratio. Remarkably, class V and pMK3b, which consist of 77 and 81% of the SAL sequence, respectively, not only do show a better relative activity compared to SHL, but even their absolute activities on *p*-nitrophenyl palmitate are 130 and 220% of that of SHL, respectively. These data indicate that in the presence of the SHL central region, mainly SAL region 3 and to a less extent SAL region 2, contribute to a better recognition of long acyl chains.

DISCUSSION

In general, specificity is determined by several residues which assemble in the tertiary structure. Consequently, the use of techniques such as in vitro evolution/random mutagenesis is limited by the amount of clones and the number of mutations per clone which have to be screened for identification of all the important residues. Therefore, to identify the regions in the primary sequence of SHL responsible for its substrate specificity, a set of SAL/SHL chimeric enzymes was generated by in vivo chimeragenesis. Three *E. coli* strains (DH5α, JM103, and JC8679) were tested as hosts. Our results show that strain DH5α is useful as a chimeragenesis host, although the recombination is somewhat less efficient than in the other strains. On the other hand, even with DH5α, the chimeras that were obtained from constructs pMK7 and pMK8 show that recombination frequently occurred in the rather highly conserved region corresponding to residues 344−358 and not in the highly variable region coding for the last 41 amino acids (compare Table 2 and Figure 3). For the introduction of random modifications into such a hypervariable region other methods such as e.g. localized random mutagenesis/error-prone PCR methods (*25*, *26*) might be more promising.

The products of the SAL/SHL chimeric genes were first tested for lipase activity. The chimeric classes I−IV showed a clear but reduced activity on *p*-nitrophenyl butyrate. Despite the fact that this observation is indicative of the generation of folded enzymes, the precise interactions between parts of the structure might be less precise in the chimeras than in the parent enzymes. We assume that SHL and SAL share the α/β hydrolase fold (*27*). Within this highly conserved fold of the central core, the diversity among SHL and SAL mainly arises from additional structural

extensions of low homology. Recombination occurs in the highly but not absolutely conserved regions of the fold, and it is conceivable that induced mismatches give rise to folded but somewhat distorted structures. According to the analysis of class II chimeras, it is clear that the N-terminal half of SHL (residues 1−179) can be replaced by the corresponding part of SAL, without destroying phospholipase activity. This region includes the active site Ser125 and a conserved region that most likely constitutes the oxyanion hole (residues 19−26). Interestingly, the phospholipase/lipase ratio of class II chimeras for long-chain phospholipid substrates is 2−3-fold increased as compared to SHL. This higher ratio is due to a decrease in lipase activity rather than an increase in phospholipase activity. When the fusion site shifts up to the C-terminus as is the case in class III chimeras, the enzyme is significantly less active on long-chain substrates, yet is still able to degrade short-chain substrates with the same phospholipase/lipase ratio as SHL. This finding indicates the importance of the central part of SHL (region 1) for the hydrolysis of long-chain substrates (discussed below). In contrast to class III chimeras, class IV chimeras are characterized by a ratio comparable to SAL for short-chain phospholipid hydrolysis. Thus, amino acids within region 2, which are additionally present in class III, appear to be essential for a high phospholipase activity. Site-directed mutagenesis targeted at residues that differ between SAL and SHL in this region can now be used to reveal the individual contributions of each of these residues to substrate specificity.

The presence of region 2 is not solely responsible for phospholipase activity, which follows from analysis of class V and VI chimeras. These data reveal the additional contribution of amino acids within the C-terminal region 4 for this function. Moreover, analysis of class VII chimeras shows that elements within region 3, containing the conserved active site residues Asp317 and His358, are required as well. Apparently, several elements which are separated in the primary structure assemble in the tertiary structure, thereby creating the polar binding pocket for the phospholipid head group.

Chain length selectivity constitutes the second main difference between SHL and SAL. Chain length selectivity as well as selectivity for unsaturation has been described for several lipases (*8, 28*−*30*). For the understanding of chain length selectivity it is important to realize that in triglycerides three acyl chains occur which may each contribute in a different way to substrate selectivity of the enzyme, giving rise to a complex picture. The complexity of these interactions has been shown by two research groups. Mannesse and co-workers showed that both the activity and the enantioselectivity of cutinase is determined by the chain length distribution on all three positions (*31*). The Paltauf group has shown that chiral selectivity is strongly influenced by the polarity at *sn*-2 of the substrates (*32*−*34*). In theory, X-ray structures should yield the information to explain the underlying molecular mechanisms of the observations made by these authors. However, most X-ray structures contain an inhibitor molecule with only one single alkyl chain. On the basis of such structures the binding pockets of the scissile acyl chain have been rather well-defined in several lipases of mammalian, fungal, and bacterial origin (*35*, *36*), but these structures do not yield information about the binding pockets for the two other chains. Only two structures of a lipase complexed with an inhibitor with three acyl chains are known: a covalent cutinase complex (*37*) and a human pancreatic lipase/colipase/phospholipid complex (*38*). For these two enzymes again a well-defined binding pocket for the scissile acyl chain at *sn*-3 was observed, but these structures yielded only limited information about the binding of the *sn*-1 and the *sn*-2 chains.

On the basis of the observations described above, it is obvious that the choice of the substrate is highly important for the characterization of chain length selectivity. Monoacyl compounds such as *p*-nitrophenyl esters and lysophospholipid analogues mainly yield information on the size of the binding pocket of the scissile chain. On the other hand, substrates with two or three chains might yield complex information reflecting the preference of two or three subsites of the enzyme. Our results with the *p*-nitrophenyl esters suggest that the binding pockets for the scissile acyl chain differ in size in SAL and in SHL as reflected by the strong preference of SAL for short-chain substrates (Table 3). The fact that SHL is highly active on long-chain substrates does not mean that this binding pocket in SHL is optimized for long chains, since chimeras such as V and pMK3b are better lipases than SHL when acting on long-chain substrates. The two latter chimeras differ from SHL by the presence of the SAL region 3, a domain that was identified as being essential for phospholipase activity in SHL. The most likely explanation for the phospholipase activity in SHL is an evolution of one of the subsites, *id est* the binding pockets for *sn*-1 or *sn*-2, from an apolar site to a more polar site capable of interacting with the phospholipid head group, although this site is still able to bind an acyl chain. As elaborated above, there is limited structural information available on these subsites, and hence the modifications in region 3 that change the interaction with either polar head groups or acyl chains are not easily predicted.

Previously, Nikoleit and co-workers (*4*) have shown that the C-terminal 146 amino acids of SHL are required for phospholipase activity. From our results, it is now clear that three regions within this C-terminal sequence each confer phospholipase activity to SHL. Moreover, at least two regions confer chain length selectivity to the enzyme. Our data also show that there is an overlap in at least region 3 in the contribution to either selectivity. Of course it is only the assembly of these regions in the tertiairy structure that gives full expression to the selectivity of the enzyme. At present we try to improve our understanding of lipase selectivity by localizing the residues rather than the regions which are important for substrate selectivity. Insight in these residues may help model building of the staphylococcal enzymes on the basis of the known X-ray structures of other bacterial lipases, because it shows which residues are close in space.

A surprising outcome of this study is that at least two enzymes were isolated with higher activity than either of the parent enzymes. Because we did not specifically screen for increased activity, the actual number of improved enzymes might have been higher. We feel that the homologous recombination of related enzymes with different properties, in addition to the well-known approaches such as screening or site-directed mutagenesis, might be a useful approach to find enzymes with desired properties. This

approach might have a wider applicability than the lipases as exemplified by the observation, though in a different context, that chimeric proteins can exceed the sum of their parts (*39*).

ACKNOWLEDGMENT

We would like to thank Mr. Ruud Cox and Mr. Ruud Dijkman for synthesis of the substrates and technical help, Dr. Marcel J. W. Janssen for introduction to DNA sequencing, Dr. Jan-Willem F. A. Simons for construction of WT-plasmids, and Cécile Lemette for secretarial help.

REFERENCES

1. Farrel, A. M., Foster, T. J., and Holland, K. T. (1993) *J. Gen. Microbiol. 139*, 267−277.
2. Götz, F., Popp, F., Korn, E., and Schleifer, K. H. (1985) *Nucleic Acids Res. 13*, 5895−5906.
3. Lee, C. Y., and Iandolo, J. J. (1986) *J. Bacteriol. 166*, 385−391.
4. Nikoleit, K., Rosenstein, R., Verheij, H. M., and Götz, F. (1995) *Eur. J. Biochem. 228*, 732−738.
5. Simons, J.-W. F. A., Adams, H., Cox, R. C., Dekker: N., Götz, F., Slotboom, A. J., and Verheij, H. M. (1996) *Eur. J. Biochem. 242*, 760−769.
6. Dugi, K. A., Dichek, H. L., and Santamarina-Fojo, S. (1995) *J. Biol. Chem. 270*, 25396−25401.
7. Thirstrup, K., Verger, R., and Carrière, F. (1994) *Biochemistry 33*, 2748−2756.
8. Bertolini, M. C., Schrag, J. D., Cygler, M., Ziomek, E., Thomas, D. Y., and Vernet, T. (1995) *Eur. J. Biochem. 228*, 863−869.
9. Kobayashi, J., Applebaum-Bowden, D., Dugi, K. A., Brown, D. R., Kaskyap, V. S., Parrott, C., Duarte, C., Maeda, N., and Santamarina-Fojo, S. (1996) *J. Biol. Chem. 271*, 26296−26301.
10. Carrière, F., Thirstrup, K., Hjorth, S., Ferrato, F., Nielsen, P. F., Withers-Martinez, C., Cambillau, C., Boel, E., Thim, L., and Verger, R. (1997) *Biochemistry 36*, 239- 248.
11. Tommassen, J., Van der Ley, P., Van Zeijl, M., and Agterberg, M. (1985) *EMBO J. 4*, 1583−1587.
12. Kim, J.-Y., and Devreotes, P. N. (1994) *J. Biol. Chem. 269*, 28724−28731.
13. Levin, L. R., and Reed, R. R. (1995) *J. Biol. Chem. 270*, 7573−7579.
14. Pikuleva, I. A., Björkhem, I., and Waterman, M. R. (1996) *Arch. Biochem. Biophys. 334*, 183−192.
15. Aarsman, A. J., Van Deenen, L. L. M., and Van den Bosch, H. (1976) *Bioorg. Chem. 5*, 241−253.
16. Volwerk, J. J., Dedieu, A. G. R., Verheij, H. M., Dijkman, R., and De Haas, G. H. (1979) *Recl. Trav. Chim. Pays-Bas 98*, 214−220.
17. Hanahan, D. (1983) *J. Mol. Biol. 166*, 557−580.
18. Gillen, J. R., Willis, D. K., and Clark, A. J. (1981) *J. Bacteriol. 145*, 521−532.
19. Yanish-Perron, C., Viera, J., and Messing, J. (1985) *Gene 33*, 103−119.
20. Simons, J.-W. F. A., Boots, J.-W. P., Slotboom, A. J., and Verheij, H. M. (1997) *J. Mol. Catal. B: Enzymatic 3*, 13−23.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular cloning: A laboratory manual*, 2nd ed. (Ford, N., Nolan, C., and Ferguson, M., Eds.), Cold Spring Harbor Laboratory Press, New York.
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. U.S.A. 74*, 5463−5467.
23. Caramori, T., Albertini, A. M., and Galizzi, A. (1991) *Gene 98*, 37−44.
24. Symington, L. S., Morrison, P., and Kolodner, R. (1985) *J. Mol. Biol. 186*, 515−525.
25. Huang, W., Petrosino, J., Hirsch, M., Shenkin, P. S., and Palzkill, T. (1996) *J. Mol. Biol. 258*, 688−703.
26. Moore, J. C., and Arnold, F. H. (1996) *Nature Biotechnol. 14*, 458−467.
27. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G., and Goldman, A. (1992) *Prot. Eng. 5*, 197−211.
28. Martinelle, M., Holmquist, M., Clausen, I. G., Patkar, S., Svendsen, A., and Hult, K. (1996) *Prot. Eng. 9*, 519−524.
29. Klein, R. R., King, G., Moreau, R. A., and Haas, M. J. (1997) *Lipids 32*, 123−130.
30. Holmquist, M., Tessier, D. C., and Cygler, M. (1997) *Biochemistry 36*, 15019- 15025.
31. Mannesse, M. L. M., Cox, R. C., Koops, B. C., Verheij, H. M., De Haas, G. H., Egmond, M. R., Van der Hijden, H. T. W. M., and De Vlieg, J. (1995) *Biochemistry 34*, 6400−6407.
32. Stadler, O., Kovac, A., Haalck, L., Spener, F., and Paltauf, F. (1995) *Eur. J. Biochem. 227*, 335−343.
33. Zandonella, G., Haalck, L., Spener, F., Faber, K., Paltauf, F., and Hermetter, A. (1995) *Eur. J. Biochem. 231*, 50−55.
34. Duque, M., Graupner, M., Stütz, H., Wicher, I., Zechner, R., Paltauf, F., and Hermetter, A. (1996) *J. Lipp. Res. 37*, 868−876.
35. 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 351*, 491−494.
36. Egloff, M.-P., Marguet, F., Buono, G., Verger, R., Cambillau, C., and Van Tilbeurgh, H. (1995) *Biochemistry 34*, 2751−2762.
37. Longhi, S., Mannesse, M. L. M., Verheij, H. M., De Haas, G. H., Egmond, M. R., Knoops-Mouthuy, E., and Cambillau, C. (1997) *Protein Sci. 6*, 275−286.
38. Van Tilbeurgh, H., Egloff, M.-P., Martinez, C., Rugani, N., Verger, R., and Cambillau, C. (1993) *Nature 362*, 814−819.
39. Campbell, R. K., Bergert, E. R., Wang, Y., Morris, J. C., and Moyle, W. R. (1997) *Nature Biotechnol. 15*, 439−443.

BI9725430