Identification of Residues Essential for Differential Fatty Acyl Specificity of *Geotrichum candidum* Lipases I and II[†]

Mats Holmquist,[‡] Daniel C. Tessier, and Miroslaw Cygler*

Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Quebec, Canada H4P 2R2

Received June 10, 1997; Revised Manuscript Received August 25, 1997[⊗]

ABSTRACT: The fungus *Geotrichum candidum* produces two lipase isoenzymes, GCL I and GCL II, with distinct differences in substrate specificity despite their 86% identical primary structure. GCL I prefers ester substrates with long-chain *cis* (Δ-9) unsaturated fatty acid moieties, whereas GCL II also accepts medium-length (C8−C14) acyl moieties in the substrate. To reveal structural elements responsible for differences in substrate differentiating ability of these isoenzymes, we designed, expressed, and characterized 12 recombinant lipase variants. Three chimeric lipases containing unique portions of the N-terminal and the C-terminal part of GCL I and GCL II, respectively, were constructed and enzymatically characterized. Activities were measured against mixed triglyceride−poly(dimethyl siloxane) particles. Our results indicate that residues within sequence positions 349−406 are essential for GCL I's high triolein/trioctanoin activity ratio of 20. The substitution of that segment in the specific GCL I to the corresponding residues in the nonspecific GCL II resulted in an enzyme with a triolein/trioctanoin activity ratio of 1.4, identical to that of GCL II. The reverse mutation in GCL II increased its specificity for triolein by a factor of 2, thus only in part restoring the high specificity seen with GCL I. In further experiments, the point mutations at the active site entrance of the GCL I, Leu358Phe and Ile357Ala/Leu358Phe, lowered the triolein/trioctanoin activity ratio from 20 to 4 and 2.5, respectively. The substitutions Cys379Phe/Ser380Tyr at the bottom of the active site cavity of GCL I decreased its specificity to a value of 3.6. Measurements of lipase activity with substrate particles composed of pure triglycerides or ethyl esters of oleic and octanoic acids resulted in qualitatively similar results as reported above. Our data reveal for the first time the identity of residues essential for the unusual substrate preference of GCL I and show that the anatomy, both at the entrance and the bottom of the active site cavity, plays a key role in substrate discrimination.

Triacylglycerol lipases (EC 3.1.1.3) constitute a large family of enzymes that catalyze the cleavage of ester bonds in their lipid substrates in aqueous emulsions. Lipases show low activity toward monomeric substrate molecules whereas the presence of substrate aggregates usually is associated with a dramatically increased activity of the lipase. This phenomenon, referred to as interfacial activation, was proposed to be due to conformational rearrangements in the lipase molecule (Sarda & Desnuelle, 1958). This suggestion has been substantiated for some lipases by the analysis of their 3D structures free (Brady et al., 1990; Winkler et al., 1990; Grochulski et al., 1994a) and in complex with transition-state analogs (Brzozowski et al., 1991; van Tilbeurg et al., 1993; Grochulski et al., 1994b). These structures revealed that a surface loop covering the active site of the lipase molecule must be displaced to allow the substrate to access the active site. This yields an open conformation that is believed to represent the catalytically active species of the enzyme. Lipases from various sources have been isolated and characterized. They all possess the α/β-hydrolase fold (Ollis et al., 1992) and a catalytic machinery defined by a Ser-His-Asp(Glu) triad. The 3D structures of numerous lipases have been elucidated in recent years [for a review, see Cygler and Schrag (1997)]. As a result, many structural aspects of lipase interfacial activation have been clarified. However, the basis of lipase substrate specificities remains poorly understood. An X-ray crystallographic investigation (Cygler et al., 1994) and a few molecular modeling studies (Norin et al., 1994; Uppenberg et al., 1995; Holmquist et al., 1996; Hult & Holmquist, 1997) of lipases have shed light on the basis of their enantioselectivities. Lipases also show different triglyceride and fatty acyl preferences. Mutational analysis of human pancreatic lipase showed that the flap has a role in substrate discrimination (Dugi et al., 1992). Point mutantions in lipases from *Rhizopus delemar* (Joerger & Haas, 1994) and *Humicola lanuginosa* (Martinelle et al., 1996) have been shown to affect triglyceride and fatty acyl specificities.

Among the triacylglycerol lipases with known 3D structures, a group of large (M_w ∼ 60 kDa) lipases can be distinguished that show high structural similarity to acetylcholinesterases (Cygler et al., 1993). These include lipases from the fungus *Geotrichum candidum* and yeast *Candida rugosa* in which two (Shimada et al. 1989, 1990; Bertolini et al., 1994) and five (Longhi et al., 1992; Lotti et al., 1993) isoenzymes, respectively, have been identified. The two *G.*

[†]M.H. gratefully acknowledges a fellowship from the Swedish Research Council for Engineering Sciences and financial support from Bengt Rydin's Foundation.

* Corresponding author: tel, 514-496-6387; fax, 514-496-5143; e-mail, mirek.cygler@bri.nrc.ca.

[‡] Permanent address: Department of Biochemistry and Biotechnology, Royal Institute of Technology SE-100 44 Stockholm, Sweden.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1997.

S0006-2960(97)01390-1 CCC: $14.00

candidum lipases (GCLs)[1] have 86% identical primary structure. Despite the high level of sequence similarity, they show distinctly different substrate preferences (Bertolini et al., 1995). Numerous reports dealing with the isolation, purification, characterization, and substrate specificities of *G. candidum* lipases have appeared during the past three decades [for a review, see Charton (1991)]. However, the reason for the high preference of a *G. candidum* lipase for lipids with *cis*-octadecenoic fatty acyl moieties early reported (Alford & Pierce, 1961; Jensen et al., 1965) still remained unknown. Among the GCLs that were studied, differences in amino acid sequence, extent of glycosylation, and substrate specificities were observed depending on factors such as the GC strain and culturing conditions. Microheterogeneity and insufficient separation of various isoforms in the studied lipase preparations may explain why ambiguity arose between the reported substrate specificities of GCLs.

A few years ago the GCL II gene from strain ATCC 34614 was cloned and expressed in *Saccharomyces cerevisiae* (Vernet et al., 1993). Only recently, the high oleyl ester preference was unambiguously traced to a separate gene product GCL I (Bertolini et al., 1995). The high-resolution (1.8-Å) 3D structure of GCL II has been determined, and the substrate binding site has been located (Schrag et al., 1991; Schrag & Cygler, 1993). The structures of the closely related *C. rugosa* lipase complexed with inhibitors mimicking different acyl chain lenghts have been elucidated (Grochulski et al., 1994b). Despite this knowledge, the structural basis of the different substrate preferences of GCLs remained obscure, in part due to the fact that the structure represents the closed, inactive conformation of the lipase. The characterization of a recombinant chimeric lipase molecule between GCL I and GCL II revealed that the 194 N-terminal residues encompassing the flap covering the active site of the lipase were not responsible for the unique fatty acyl preference of GCL I (Bertolini et al., 1995). In our pursuit to identify substrate specificity determinants in GCL I, we recently reported the high-level expression of GCLs in yeast *Pichia pastoris* (Holmquist et al., 1997).

In this paper, we describe the cloning, expression, and characterization of several recombinant GCL variants. The lipases were enzymatically characterized with triglycerides and ethyl esters of oleic and octanoic acids. Lipase activities were determined under different physicochemical conditions; substrate emulsions consisting of mixed triglycerides−poly(dimethyl siloxane) (PDMS) particles or pure lipid particles. The analysis of substrate specificities and amino acid sequences of wild-type GCLs and lipase chimeras between GCL I and GCL II revealed a region in the GCL I lipase molecule with structural elements essential for its high oleyl ester preference. In subsequent experiments, we targeted the identified region by site-directed mutagenesis to locate individual amino acids essential for the high oleate ester preference of GCL I.

[1] Abbreviations: BMGY, buffered glycerol-complex medium; BMMY, buffered methanol-complex medium; BSA, bovine serum albumin; GCL, *Geotrichum candidum* lipase; PDMS, polydimethyl siloxane; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; SD-His, dextrose-based medium without histidine supplementation.

MATERIALS AND METHODS

Yeast and Escherichia coli Strains. *P. pastoris* GS115 (his4) (Cregg et al., 1985; Invitrogen Corporation, San Diego, CA) was used for expression of *G. candidum* lipase variants. *E. coli* strains MC1061 (Casadaban & Cohen, 1980) and CJ236 (Bio-Rad, Hercules, CA) were used for the construction of all plasmids and the production of single-stranded DNA, respectively.

Yeast Culture Media. One liter of BMGY medium contained the following: yeast extract (10 g), peptone (20 g), glycerol (10 mL), biotin (400 μg), yeast nitrogen base with ammonium sulfate (13.4 g), and 1 M potassium phosphate buffer pH 6.0 (100 mL). BMMY medium had the same composition as BMGY medium with the exception that 5 mL of methanol/L was added instead of glycerol.

Construction of Lipase Hybrid Genes. (a) Lipase I-348-II. A *P. pastoris* expression vector carrying a gene coding for this hybrid lipase molecule (the 348 N-terminal amino acid residues of GCL I fused to residues 349−544 of GCL II) was constructed using the unique *Bst*EII site in both lipase genes. The GCL II cDNA portion was excised as a *Bst*EII−*Not*I fragment (756 bp) from plasmid YpDC523 (Holmquist et al., 1997) and cloned into plasmid YpDC521 (GCL I fused to the α-factor secretion signal and under the control of the alcohol oxidase promoter; Holmquist et al., 1997) yielding plasmid YpDC535. The constructed lipase gene included a segment coding for an N-terminal $(His)_8$LeuValProArg extension. The $(His)_8$ tag allows for the purification of the recombinant lipase on a nickel nitrilotriacetic acid (Ni^{2+}-NTA) agarose column (Qiagen) (Bertolini et al., 1995; Holmquist et al., 1997). The LeuValPro Arg is a thrombin cleavage site. The presence of the $(His)_8$ tag extension does not affect the catalytic properties of the lipase (Schrag et al. 1995).

(b) Lipase I-406-II. A lipase gene coding for the 406 N-terminal amino acid residues of GCL I and residues 407−544 of GCL II was created using the single *Bsm*I site in both lipase genes. The GCL I cDNA portion was isolated as a *Pst*I−*Bsm*I fragment (1496 bp) from plasmid YpDC521, and the GCL II portion was excised as a *Bsm*I−*Not*I fragment (581 bp) from plasmid YpDC523. These DNA fragments were simultaneously cloned into the *Pst*I−*Not*I digested plasmid YpDC519 [*P. pastoris* expression vector pPIC9 (Invitrogen Corporation, San Diego, CA) with its *Pst*I site in the ampicillin resistance gene deleted (Holmquist et al., 1997)] yielding plasmid YpDC544.

(c) Lipase I-462-II. To construct a lipase gene coding for the 462 N-terminal amino acids of lipase I and residues 463−544 of lipase II, the unique *Nco*I site in both lipase genes was used. The GCL I cDNA portion was isolated as a *Pst*I−*Nco*I fragment (1662 bp) from YpDC521, and the GCL II cDNA portion was excised as a *Nco*I−*Not*I fragment (415 bp) from YpDC523. The isolated fragments were simultaneously cloned into the *Pst*I−*Not*I digested plasmid YpDC519 yielding plasmid YpDC546.

(d) Lipase I-348-II-406-I. A lipase gene coding for GCL I with amino acid residues 349−406 exchanged for the corresponding residues of GCL II was created from plasmids YpDC535, YpDC521, and YpDC519. The *Pst*I−*Bsm*I fragment (1496 bp) from YpDC535 (N-terminal portion of GCL I-348-II) and the *Bsm*I−*Not*I fragment (421 bp) from

Table 1: Oligonucleotides Used in This Study for Site-Directed Mutagensis and Sequencing of *Geotrichum candidum* Lipases

oligo	oligonucleotide sequence	mutation (diagnostic)
DTo310	5′GGATGAGGGTACTATT**T**T**C**GC**T**CCCGTGGCTATTAATGC3′	L358F (destroys *Bgl*I)
DTo311	5′GGATGAAGGTACTGCT**C**TTGC**C**CCTGT**G**GCTCTCAACGCTACC3′	F358L (creates *Bgl*I)
DTo315	5′CAGGAGGATGAGGGTACT**GC**T**T**T**C**GC**T**CCCGTGGCTATTAATGC3′	I357A/L358F (destroys *Bgl*I)
DTo334	5′GTGGTTGAAGTACATTT**TTA**CGA**A**GCTTCTGACGCTTCG3′	C379F/S380Y (creates *Hin*dIII)
DTo338	5′CAACATTATTCCCGATGC3′	for sequencing of mutations

[a] Mutated nucleotides are shown in bold, and the diagnostic restriction site destroyed or deleted is underlined.

YpDC521 (C-terminal portion of GCL I) were simultaneously cloned into the *Pst*I−*Not*I digested plasmid YpDC519 yielding plasmid YpDC569.

(e) Lipase II-348-I-406-II. A lipase gene coding for GCL II with amino acids residues 349−406 substituted for those of GCL I was created from plasmids YpDC523 and YpDC544. The *Pst*I−*Bst*EII fragment (1321 bp) from YpDC523 (N-terminal portion of GCL II) was cloned into the *Pst*I−*Bst*EII digested plasmid YpDC544 (GCL I-406-II) yielding plasmid YpDC566.

Construction of Parent Plasmid Used for Site-Directed Mutagenesis. The plasmid YpDC519 (Holmquist et al., 1997) was used to introduce the kanamycin-resistance gene of commercial pPIC9K (Invitrogen Corporation, San Diego, CA) in the form of a *Pst*I cassette (1241 bp) treated with T4 DNA polymerase within the unique *Nae*I site of YpDC519. The resulting plasmid YpDC525 was then digested with *Bst*1107I to introduce an *Spe*I−*Nru*I blunt-ended fragment (623 bp) of pVT102-U (Vernet et al., 1987) containing the f1 intergenic region. The orientation of the f1 intergenic region was chosen such that single-stranded DNA prepared from this plasmid would be complementary to the mRNA. Finally, this plasmid contained an extra *Xho*I site within the kanamycin-resistance gene. To remove it, the plasmid was digested with *Pae*R7I (an isoschizomer of *Xho*I), which does not cut if its recognition sequence is preceeded by CT. The *Xho*I site within the α-factor pro region is not digested by *Pae*R7I whereas the one within the kanamycin-resistance gene is. Following digestion, the plasmid was blunt-ended with T4 DNA polymerase and treated with Taq DNA polymerase in the presence of dATP and dTTP only, to favor its template-independent DNA polymerase activity (Clark, 1988). The molar ratio of dTTP exceeded that of dATP by 100-fold. The addition of a single base to the blunt-ended *Xho*I site would thus restore the reading frame of the kanamycin-resistance gene. The plasmid was then ligated, transformed in *E. coli*, and plated on 2YT + ampicillin + kanamycin. The absence of digestion of the selected recombinant plasmid with *Eae*I indicated that an A was added to the coding strand. The only change in the amino acid sequence of the kanamycin-resistance gene was the deletion of a serine residue at position 11. The final plasmid was called YpDC541 and is 9895bp in length.

Site-Directed Mutagenesis of Lipase Genes. The *G. candidum* lipase genes I and II were excised as the *Sca*I−*Not*I fragments from plasmids YpDC521 and YpDC523 (Holmquist et al., 1997), respectively, and cloned into plasmid YpDC541. This yielded plasmids YpDC550 (GCL I) and YpDC552 (GCL II). These plasmids were transformed into *E. coli* strain CJ236, and uracil-containing single-stranded template DNAs were prepared for in vitro site-directed mutagenesis (Kunkel, 1985; Kunkel et al., 1987). Lipase mutant genes were constructed using the mutagenic oligonucleotides described in Table 1. The mutations were confirmed by analysis of the presence or the absence of a diagnostic restriction site created or deleted, respectively, in the mutagenic oligonucleotides (Table 1). Introduced mutations were also confirmed by DNA sequencing according to the method of Sanger (1977).

Transformation of Plasmid DNAs into Yeast P. pastoris. The constructed *P. pastoris* expression plasmids carrying the GCL genes were linearized by *Stu*I digestion, and 5 mg of plasmid DNA was transformed into competent *P. pastoris* cells by electroporation using a Bio-Rad Gene Pulser instrument. Aliquots were spread on SD-His plates and incubated at 30 °C. After 3 days of incubation, colonies of recombinant *P. pastoris* appeared on the plate. These were screened with the previously described Yeastern method (Holmquist et al., 1997) to identify lipase secreting transformants. In short, this method involves the replication of methanol-induced yeast clones to a nitrocellulose membrane. High-level secreting clones can subsequently be identified by means of GCL-specific antibodies and an IgG-alkaline phosphatase conjugate.

Cultivation of Pichia pastoris. Transformed *P. pastoris* cells were grown at 30 °C in 25 mL BMGY until a cell density of $OD_{260} = 2-6$ was reached. The culture was centrifuged, the supernatant was discarded, and the cells were resuspended in 100 mL of BMMY to an $OD_{260} \sim 1$ to start induction. The culture was grown for 4 days with the addition of 500 μL of methanol once a day to maintain induction.

SDS−PAGE and Western Blot Analyses. The culture medium (10 μL) containing the produced lipase was analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) on a 8% gel. Western blot analysis was performed by transferring the SDS−PAGE-separated sample to a nitrocellulose membrane using a Transblot apparatus (Bio-Rad) according to the company's specifications.

Lipase Purification. The culture medium containing recombinant lipase was concentrated five times by means of a Centriprep-30 (Amicon) into 20 mM MOPS buffer pH 7.0. Lipase concentration was determined spectrophotometrically (Bradford, 1976) using BSA as a standard. The concentrated lipases were enzymatically characterized without further purification. All recombinant lipases migrated as a single band on a Coomassie stained SDS−PAGE gel and after Western blot analysis.

Lipase Activity Determination. Lipase activity was measured titrimetrically. Triglyceride−poly(dimethyl siloxane) (PDMS) emulsions were prepared by mixing triglyceride dissolved in chloroform (20 μL, 0.5 M) with PDMS (0.5 mL, viscosity 50 cs) and vigorously vortexed to allow the solvent to evaporate. A solution of gum arabic (9 mL, 2% w/v) and calcium chloride (0.5 mL, 2 M) was added to the substrate. Alternatively ester (50 mg) was mixed with gum

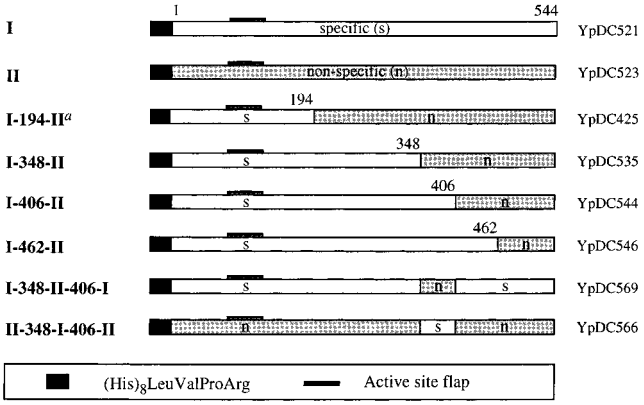


FIGURE 1: Schematic representation of wild-type and chimeras of *Geotrichum candidum* lipases used for the analysis of substrate specificity determinants. The names of the constructs are listed to the left, and numbers indicate the amino acid residue at the N-terminal side of the junction between the lipase isoenzymes. The high (specific) and low (nonspecific) substrate specificity of lipases I and II, respectively, for oleate esters is pointed out. The plasmid numbers of the constructs are listed to the right. Footnote *a* indicates that the data are Bertolini et al. (1995).

arabic (10 mL, 2% w/v) and calcium chloride (0.1 M). The substrates were emulsified in a sonicator (Heat System, Ultrasonics Inc.) for 2 minutes using a macroprobe (1.27 cm diameter) with the power output set at 50%. Lipase was added, and the enzymatic activity was determined with a pH-stat (RTS822 Recording Titration System, Radiometer) at pH 7.0 by the titration of the released fatty acid with sodium hydroxide (50 mM). Lipase activity was expressed in millimoles of released fatty acid minute^{-1} (milligram of protein)$^{-1}$. The results shown represent mean values of at least three separate measurements.

RESULTS AND DISCUSSION

Construction and Expression of Lipase Variants. The constructed chimeric *G. candidum* lipase genes are shown in Figure 1. We recently reported the successful high-level expression of wild-type GCLs in yeast *P. pastoris* (Holmquist et al., 1997). This prompted us to use the same system for the constructs described here. All lipase genes were expressed and secreted into the culture medium in high yield and purity. Expression levels for all constructs were similar to what was previously obtained with wild-type GCLs (~60 mg/L; Holmquist et al., 1997). All recombinant lipases migrated as a single band on a Coomassie-stained SDS−PAGE and after Western blot analysis. As a first step toward identifying amino acid residues responsible for the unusual substrate specificity of GCL I, hybrid lipase molecules containing the N-terminal part of GCL I and the C-terminal part of GCL II were designed, cloned, and expressed (Figure 1). By using the single common restriction sites *Bst*EII, *Bsm*I, or *Nco*I in the GCL I and GCL II genes, three GCL I/II chimeras were synthesized. The expression of these genes resulted in a set of lipases where the position of the junction between the GCL I and GCL II part of the hybrid progressed toward the C-terminus in steps of 56 residues or more.

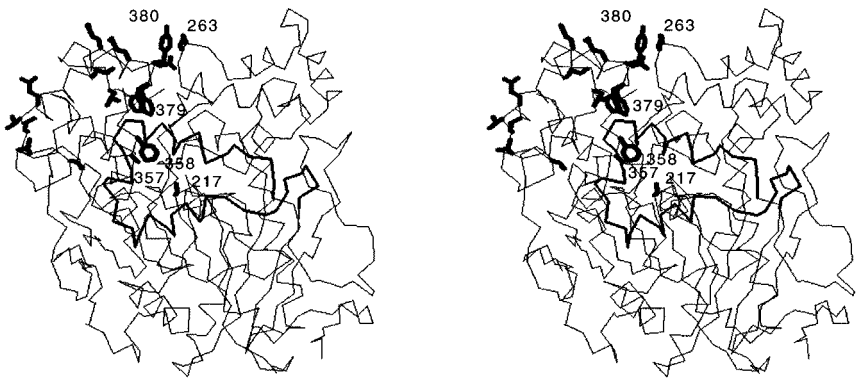
Activities and Substrate Specificities against Mixed Triglyceride−PDMS Particles. The specific activities and substrate specificities of GCL I, GCL II, and the lipase chimeras were first determined toward mixed triglyceride−PDMS emulsions. With this type of emulsion, the triglyceride substrate accumulates at the surface of the PDMS droplets of defined radii. As a consequence, the surface pressure of different substrates will be normalized at the lipid/water interface (Ziomek et al., 1996). As model triglyceride substrates, we chose triolein and trioctanoin. These lipids have previously been reported to be diagnostic for the different substrate preferences of the naturally occurring GCL isoenzymes (Sugihara et al., 1993; Bertolini et al., 1995). As can be seen in Table 2, the specific activities of GCL I and GCL II differ clearly for both substrates. The most pronounced difference occurs for trioctanoin, for which the GCL I isoform has 10 times lower specific activity than GCL II. At the same time, the GCL I has slightly higher (1.4 times) activity toward triolein. The triolein/trioctanoin activity ratio, indicative of the degree of lipase specificity, has a value of 20 for GCL I and 1.4 for GCL II. The GCL I-194-II hybrid described previously (Bertolini et al., 1995) showed properties similar to GCL II (Table 2), demonstrating that the N-terminal fragment including the flap, has little effect on substrate specificity. Due to a different expression system, purification protocol, and some differences in assay procedures, we interpret these results only qualitatively in the present context. Chimeric lipases described here have been characterized more thoroughly, and a quantitative comparison was possible. The GCL I-348-II hybrid, containing the 348 N-terminal residues of GCL I and the C-terminal residues from GCL II, also had properties similar to GCL II, but with a slightly higher preference for triolein over trioctanoin (Table 2) resulting from an increased specific activity toward triolein as compared to GCL II and identical to that of GCL I. The activity toward trioctanoin was the same as GCL II. By contrast, the next hybrid, GCL I-406-II, showed enzymatic properties identical to GCL I (Table 2). The GCL I-462-II hybrid with the junction moved even further toward the C-terminus had, as expected, the same substrate specificity as the wild-type GCL I. The 10-fold difference in specific activity of GCL I-348-II and GCL I-406-II toward trioctanoin indicates that residues essential for the activity toward this substrate are located within the 349−406 fragment of the lipase isoforms (Table 2).

Table 2: Specific Activities and Substrate Specificities of *Geotrichum candidum* Lipase Variants against Triglyceride−PDMS Emulsions

lipase	triolein (μmol min^{-1} mg^{-1})[a]	trioctanoin (μmol min^{-1} mg^{-1})[a]	triolein/trioctanion activity ratio[b]
I	380 ± 34	19 ± 6	20 ± 7
II	270 ± 31	190 ± 20	1.4 ± 0.2
I-194-II[c]	680 ± 100	700 ± 40	1.0 ± 0.2
I-348-II	380 ± 26	180 ± 31	2.1 ± 0.4
I-406-II	390 ± 14	19 ± 2	21 ± 2
I-462-II	410 ± 9	15 ± 3	27 ± 6
I-348-II-406-I	290 ± 9	190 ± 20	1.5 ± 0.2
II-348-I-406-II	290 ± 28	98 ± 20	3.0 ± 0.7

[a] Standard deviations are calculated from three separate experiments. [b] Standard deviations are estimated from the determined standard deviations in specific activities. [c] Bertolini et al., 1995.

Subsequent work focused on the 349−406 region in an effort to define more specifically the residues that determine the lipase substrate specificity. We constructed a swap mutant, GCL I-348-II-406-I, in which residues 349−406 in GCL I were substituted by the corresponding sequence of




```
                 349                                                  406
GCL I:   GN QEDEGTILAP VAINATTTPH VKKWLKYICSE ASDASLDRVL SLYPGSWSE GAPFRT
GCL II:  -- ------AF-- --L------- -----Q--FYD --E--I---- ----QTL-V -S----
```

FIGURE 2: (Top) Stereoview of *Geotrichum candidum* lipase II (Schrag et al., 1993) showing residues differing between lipase isoenzymes I and II within positions 349−406. The flap (residues 61−105) covering the active site is shown in thicker line α-trace, and the active site Ser217 (center) is shown. Phe379 is illustrated in the two conformations observed in the crystal structure, and the position of Pro263 (conserved in GCLs) is indicated. (Bottom) Comparison of amino acid residues within positions 349−406 in *Geotrichum candidum* lipases I and II. Residues targeted by site-directed mutagenesis (357, 358, 379, and 380) are shown in bold.

the less specific GCL II (Figure 2). As seen in Table 2, this resulted in an enzyme with the identical specific activities for both substrates and the same specificity as GCL II. This shows that the most important elements controlling the substrate differentiating ability of GCL I are located within this targeted region of the lipase molecule. To test if this region is solely responsible for specificity, a reverse swap introducing the GCL I segment into the nonspecific GCL II was made (GCL II-348-I-406-II construct). This chimeric lipase showed an increased preference for triolein over trioctanoin, but only by a factor of 2. This increased specificity was due to an exclusively decreased specific activity toward trioctanoin (Table 2). The exchange of residues 349−406 in either GCL isoform thus affected the activity toward trioctanoin most, while these two chimeras interacted in a similar manner with the triolein substrate. The GCL II-348-I-406-II lipase shows lower substrate specificity than GCL I. This indicates the involvement of residues outside of the 349−406 region in the discrimination against trioctanoin (short fatty acyl chains). The specific activity of various constructs toward triolein, shown in Table 2, give no clear indication of a particular region of the lipase that might be responsible for the observed variations. Although the variations are only of the order of 30%, the activities are consistently observed at two levels: ∼290 or ∼400 μmol min^{-1} mg^{-1}. Although the results for the chimeras could be interpreted as indicating that the higher level of activity is associated with the ∼350 N-terminal residues, this is not supported by the activities of mutants with swapped 349−406 segment.

To further define key elements within residues 349−406 in GCL I that are critical for its substrate differentiating ability, we targeted this region by site-directed mutagenesis. To identify target amino acids, we analyzed the sequence differences between GCL I and GCL II within residues 349−406. Within this sequence portion, 14 amino acid residues differ between GCL I and GCL II, most of them located at the surface of the lipase molecule (Figure 2). Three of these residues (357, 358, and 379) likely contribute directly to the formation of the substrate binding cavity of the enzyme (Bertolini et al., 1995). Inspection of the 3D structure of GCL II suggests that only two of them (357 and 358) appeared to be located in positions capable of interactions with a bound octanoic acid moiety of a substrate. The corresponding residues in the closely related *C. rugosa* lipase are located at the entrance of the acyl-binding tunnel (Grochulski et al., 1994b). Since our aim at this point was to identify rapidly a small number of key residues determining substrate specificity, we decided to investigate several mutants with multiple substitutions rather than to explore in depth all individual single mutants. The most promising was position 358 at the entrance of the active site, and, indeed, the single Leu358Phe mutation in GCL I resulted in a 5-fold increase in specific activity toward trioctanoin without affecting triolein activity (Table 3). Thus, the specificity for triolein decreased to a value of 4, 5-fold lower as compared to wild type GCL I (Table 3). The Ile357Ala/Leu358Phe double substitution in GCL I lowered its specificity for triolein from 20 to 2.5, almost reaching the low value of 1.5 seen with the nonspecific GCL II (Table 3). This additional effect relative to the single Leu358Phe mutant was due to a decrease of activity toward triolein and an increase of activity toward trioctanoin. Similar to the results seen with the GCL I-348-II-406-I construct, these two mutations had primarily affected the activities toward trioctanion. These results show that both targeted residues, Ile357 and Leu358, play a significant role in creating the unusual substrate preference of GCL I.

The Cys379Phe/Ser380Tyr double substitution at the bottom of the active site cavity of GCL I had an unexpected effect on activity and specificity. The specific activities for both substrates increased significantly, making this mutant by far the most active toward triolein, while at the same time its triolein/trioctanoin activity ratio was lowered from 20 to 3.5 (Table 3). The results for trioctanoin are intriguing. According to the structural model based on the homologous *C. rugosa* lipase, residues 379−380 are at the bottom of the substrate binding site (tunnel) and should not be able to contact a short fatty acyl chain of eight carbon atoms. This result indicates subtle effects of the residues at the bottom of the "tunnel" that may be related to the dynamics of the molecule and influence the catalysis. In the crystal structure of GCL II, the side chain of Phe379 is observed in two different conformations while Tyr380 forms ring stacking interactions with Pro263 (conserved in GCLs) (Schrag & Cygler, 1993). The GCL molecule consists of two subdo-

Table 3: Specific Activities and Substrate Specificties of *Geotrichum candidum* Lipase Mutants against Triglycerides and Ethyl Esters of Oleic and Octanoic Acids

	mixed triglyceride−PDMS particles			pure triglyceride particles			pure ethyl ester particles		
lipase	oleate (μmol min^{-1} mg^{-1})	octanoate (μmol min^{-1} mg^{-1})	oleate/ octanoate activity ratio[b]	oleate (μmol min^{-1} mg^{-1})	octanoate (μmol min^{-1} mg^{-1})	oleate/ octanoate activity ratio[b]	oleate (μmol min^{-1} mg^{-1})	octanoate (μmol min^{-1} mg^{-1})	oleate/ octanoate activity ratio[b]
I	380 ± 34	19 ± 6	20 ± 7	1200 ± 180	380 ± 10	3.2 ± 0.5	370 ± 18	2.4 ± 0.8	154 ± 52
I-348-II-406-I	290 ± 9	190 ± 20	1.5 ± 0.2	630 ± 20	600 ± 20	1.1 ± 0.1	330 ± 27	19 ± 2	17 ± 2
I I357A/L358F	320 ± 25	130 ± 36	2.5 ± 0.7	1100 ± 170	470 ± 70	2.3 ± 0.5	540 ± 45	16 ± 1	33 ± 3
I L358F	400 ± 23	100 ± 35	4 ± 1	1200 ± 180	520 ± 80	2.2 ± 0.5	630 ± 11	13 ± 1	52 ± 5
I C379F/S380Y	920 ± 100	260 ± 41	3.5 ± 0.7	1600 ± 100	740 ± 170	2.2 ± 0.5	200 ± 13	7 ± 1	29 ± 4
I I357A/L358F/ C379F/S380Y	350 ± 17	160 ± 17	2.2 ± 0.3	510 ± 80	270 ± 15	1.9 ± 0.3	140 ± 13	15 ± 3	9 ± 2
II	270 ± 31	190 ± 20	1.4 ± 0.2	770 ± 170	770 ± 30	1.0 ± 0.2	400 ± 25	30 ± 2	13 ± 1
II-348-I-406-II	290 ± 28	98 ± 20	3.0 ± 0.6	910 ± 81	440 ± 35	2.1 ± 0.3	130 ± 3	6.5 ± 0.7	21 ± 2
II F358L	370 ± 21	230 ± 15	1.6 ± 0.1	1000 ± 150	510 ± 40	1.9 ± 0.3	220 ± 29	4.8 ± 1.6	45 ± 16

[a] Standard deviations are calculated from three separate experiments. [b] Standard deviations are estimated from the determined standard deviations in specific activities.

mains of approximately equal size (Cygler et al., 1993). The Tyr380/Pro263 interaction may control the flexibility between theses part of the lipase molecule (Figure 2). In the *C. rugosa* lipase, such movements have been suggested to be implicated in the process of binding the scissile fatty acyl chain of the substrate into the active site tunnel (Grochulski et al., 1994b). However, to address this question in detail, further studies are needed. Combining all four mutations, Ile357Ala/Leu358Phe/Cys379Phe/Ser380Tyr, in GCL I led to specific activities and specificity very similar, although still not identical, to that of GCL II (Table 3). Thus, these four residues appear to constitute the major elements that control substrate specificity of GCL I. The reverse mutation at position 358 in GCL II, Phe358Leu, did not significantly affect substrate specificity. Instead, an overall increased enzyme activity was seen (Table 3). Further studies are needed to address the question of how to introduce substrate specificity into GCL II with a minimum number of mutations.

Activities and Substrate Specificities against Pure Triglyceride Particles. It is known that lipase activity and specificity can depend on the physicochemical properties of the substrate particle (Verger, 1980; Rogalska et al., 1993). For that reason, we also characterized the lipases with emulsions of pure triglyceride substrate. With this assay, GCL I and GCL II showed triolein/trioctanoin activity ratios of 3 and 1, respectively (Table 3). The substrate specificities of the GCL isoforms thus depend on the composition of the substrate-containing emulsion. Nevertheless, the GCL I and GCL II still show clearly distinguishable substrate differentiating abilities. The GCL I-348-II-406-I mutant showed the substrate specificity and specific activity close to the nonspecific GCL II, similar to the results seen with the mixed triglyceride−PDMS emulsion (Table 3). Furthermore, the GCL II-348-I-406-II construct had a 2-fold increased preference for triolein as compared to wild-type GCL II. The mutations in the GCL II thus resulted in a lipase that preferentially hydrolyzed triolein. In fact, this represents an example where a modest level of specificity was truly engineered into GCL II, which does not have the ability to discriminate between the triglycerides under these conditions (Table 3). This proves again the importance of structural elements within residues 349−406 in conferring substrate specificity to the lipase molecule.

The Leu358Phe, Ile357Ala/Leu358Phe, and Cys379Phe/Ser380Tyr mutants of GCL I all showed triolein/trioctanoin activity ratios of ∼2, intermediate to what was seen with wild-type GCLs. Under these assay conditions, the role of residue 358 as a substrate specificity determinant appeared to depend on the identity of residue 357. The single mutation Phe358Leu in GCL II was sufficient to achieve very similar catalytic properties as when the segment 349−406 was exchanged (Table 3). The nonspecific GCL II had thus been engineered to augment preferential hydrolysis of triolein by the single mutation Phe358Leu at the active site entrance of the lipase.

Activities and Substrate Specificities against Pure Ethyl Ester Particles. The fatty acyl specificities of the recombinant lipases were determined with ethyl esters of oleic and octanoic acids. GCL I and GCL II showed oleate/octanoate activity ratios of 154 and 13, respectively. Both lipases had activities toward ethyl oleate in the similar range (100−150%) as seen toward the triolein−PDMS emulsion. By contrast, their activities against ethyl octanoate were 6−8 times lower than seen with trioctanoate−PDMS substrate particles (Table 3). The GCL I-348-II-406-I construct had the similar enzymatic selectivities as GCL II. This proves that the mutated residues dictate fatty acyl specificty and not only triglyceride preference as revealed by the experiments described above. The mutations in GCL I affected more significantly the specific activity toward the octanoate ester than the activity against ethyl oleate. The GCL II-348-I-406-II mutant again showed a small (1.5-fold) but significant, within experimental error, increase in oleate ester preference as compared to wild-type GCL II (Table 3).

CONCLUSIONS

We have shown that the major specificity determinants of GCL I lipase are located within a 58 residue long region, spanning residues 349−406, and that modifications within this segment affect primarily the activity toward short-chain fatty acyl substrates. Furthermore, we have narrowed down the crucial elements of this region to four residues at positions 357, 358, 379, and 380. Their replacement in GCL I by equivalent amino acids from GCL II led to a profile very similar to that of GCL II. While this study showed that the specificity can be lowered by a few mutations, introducing specificity into the nonspecific GCL II turned out to be more challenging. The reverse mutations in GCL II recovered only a fraction of specificity observed in GCL I. It became obvious that there are residues in other parts of the enzyme

also contributing to substrate specificity. Nevertheless, some selectivity between triolein and trioctanoin has been engineered. Importantly, among the mutants investigated here, one has significantly higher specific activities for both triglyceride substrates. Thus, an improved lipase has been engineered by a rational, 3D structure-based approach.

ACKNOWLEDGMENT

We would like to thank Avak Kahwajian for making two mutants, Daniel Dignard for performing the DNA sequencing, and Drs. J. D. Schrag and E. Ziomek for helpful discussions.

REFERENCES

Alford, J. A., & Pierce, D. A. (1961) *J. Food Sci. 26*, 518−524.
Bertolini, M. C., Laramée, L., Thomas, D. Y., Cygler, M., Schrag, J. D., & Vernet, T. (1994) *Eur. J. Biochem. 219*, 119−125.
Bertolini, M. C., Schrag, J. D., Cygler, M., Ziomek, E., Thomas, D. Y., & Vernet, T. (1995) *Eur. J. Biochem. 228*, 863−869.
Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
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., Thim, L., & Menge, U. (1990) *Nature 343*, 767−770.
Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Björkling, F., Huge-Jensen, B., Patkar, S. A., & Thim, L. (1991) *Nature 351*, 491−494.
Casadaban, M., & Cohen, S. N. (1980). *J. Mol. Biol. 138*, 179−207.
Charton, E. (1991) Ph.D. Thesis, Institut National Agronomique Paris-Grignon, France.
Clark, J. M. (1988) *Nucleic Acids Res. 16*, 9677−9686.
Cregg, J. M., Barringer, K. J., Heseler, A. Y., & Madden, K. R. (1985) *Mol. Cell. Biol. 5*, 3376−3385.
Cygler, M., & Schrag, J. D. (1997) *Methods Enzymol. 284*, 3−27.
Cygler, M., Schrag, J. D., Sussman, J. L., Harel, M., Silman, I., Gentry, M. K., & Doctor, B. P. (1993) *Protein Sci. 2*, 366−382.
Cygler, M., Grochulski, P., Kazlauskas, R. J., Schrag, J. D., Bouthillier, F., Rubin., B., Serreqi, A. N., & Gupta, A. K. (1994) *J. Am. Chem. Soc. 116*, 3180−3186.
Dugi, K. A., Dichek, H. L., Talley, G. D., Brewer, H. B., Jr., & Santamarina-Fojo, S. (1992). *J. Biol. Chem. 267*, 25086−25091.
Grochulski, P., Li, Y., Schrag, J. D., & Cygler, M. (1994a) *Protein Sci. 3*, 82−91.
Grochulski, P., Bouthillier, F., Kazlauskas, R. J., Serreqi, A. N., Schrag, J. D., Ziomek, E., & Cygler, M. (1994b) *Biochemistry 33*, 3494−3500.
Holmquist, M., Hæffner, F., Norin, T., & Hult, K. (1996) *Protein Sci. 5*, 83−88.
Holmquist, M., Tessier, D. C., & Cygler, M. (1997) *Protein Expression Purif.* (in press).
Hult, K., & Holmquist, M. (1997) *Methods Enzymol.* (in press).
Jensen, R. G., Sampugna, J., Guinn, J. G., Carpenter, D. L., & Marks, T. A. (1965) *J. Am. Chem. Soc. 42*, 1029−1032.
Joerger, R. D., & Haas, M. J. (1994) *Lipids 29*, 377−384.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
Longhi, S., Fusetti, F., Grandori, R., Lotti, M., Vanoni, M., & Alberghina, L. (1992*) Biochim. Biophys. Acta 1131*, 227−232.
Lotti, M., Grandori, R., Fusetti, F. Longhi, S., Brocca, S., Tramontano, A., & Alberghina, L. (1993) *Gene 124*, 45−55.
Martinelle, M., Holmquist. M., Clausen, I. G., Patkar, S., Svendsen, A., & Hult, K. (1996) *Protein Eng. 9*, 519−524.
Norin, M., Hæffner, F., Achour, A., Norin, T., & Hult, K. (1994) *Protein Sci. 3*, 1493−1503.
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., Verscheuren, K. H. G., & Goldman, A. (1992) *Protein Eng. 5*, 197−211.
Rogalska, E., Ransac, S., & Verger, R. (1993) *J. Biol. Chem. 268*, 792−794.
Sanger, F., Niklen, S., & Coulson, R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Sarda, L., & Desnuelle, P. (1958) *Biochim. Biophys. Acta 30*, 513−521.
Schrag, J. D., & Cygler, M. (1993) *J. Mol. Biol. 230*, 575−591.
Schrag, J. D., Li, Y., Wu, S., & Cygler, M. (1991) *Nature 351*, 761−764.
Schrag, J. D., Vernet, T., Laramée, L., Thomas, D. Y., Recktenwald, A., Okoniewska, M., Ziomek, E., & Cygler, M. (1995) *Protein Eng. 8*, 835−842.
Shimada, Y., Sugihara, A., Tominaga, Y., Iizumi, T., & Tsunasawa, S. (1989) *J. Biochem. 106*, 383−388.
Shimada, Y., Sugihara, A., Iizumi, T., & Tominaga, Y. (1990) *J. Biochem. 107*, 703−707.
Sugihara, A., Hata, S., Shimada, Y., Goto, K., Tsunasawa, S., & Tominaga, Y. (1993) *Appl. Microbiol. Biotechnol. 40*, 279−283.
Uppenberg, J., Öhrner, N., Norin, M., Hult, K., Kleywegt, G. J., Patkar, S., Waagen, V., Anthonsen, T., & Jones, T. A. (1995) *Biochemistry 34*, 16838−16851.
van Tilbeurgh, H., Egloff, M.-P., Martinez, C., Rugani, N., Verger, R., & Cambillau, C. (1993) *Nature 362*, 814−820.
Verger, R. (1980) *Methods Enzymol. 64*, 340−392.
Vernet, T., Dignard, D., & Thomas, D. Y (1987) *Gene 52*, 225−233.
Vernet, T., Ziomek, E., Recktenwald, A., Schrag, J. D., de Montigny, C., Tessier, D. C., Thomas, D. Y., & Cygler, M. (1993) *J. Biol. Chem. 268*, 26212−26219.
Winkler, F. K., D'Arcy, A., & Hunziker, W. (1990) *Nature 343*, 771−774.
Ziomek, E., Douchet, I., Ivanova, M., & Verger, R. (1996) *Chem. Phys. Lipids 81*, 1−9.

BI971390D