

Cloning and Sequence Analysis of cDNA encoding *Rhizopus niveus* Lipase

Wataru KUGIMIYA, Yasuo OTANI, Mitsutaka KOHNO, and Yukio HASHIMOTO

Central Research Institute, Fuji Oil Co., Ltd., 4–3 Kinunodai, Yawara, Tsukuba-gun, Ibaraki 300–24, Japan

Received September 6, 1991

Complementary DNA encoding *Rhizopus niveus* lipase (RNL) was isolated from the *R. niveus* IF04759 cDNA library using a synthetic oligonucleotide corresponding to the amino acid sequence of the enzyme. A clone, which had an insert of 1.0 kilobase pairs, was found to contain the coding region of the enzyme. The lipase gene was expressed in *Escherichia coli* as a *lacZ* fusion protein. The mature RNL consisted of 297 amino acid residues with a molecular mass of 32 kDa. The RNL sequence showed significant overall homology to *Rhizomucor miehei* lipase and the putative active site residues were strictly conserved.

Lipase (EC 3.1.1.3) catalyzes the hydrolysis of fats to give free fatty acid, partial glycerides, and glycerol, and in reverse the formation of glycerides from glycerol and free fatty acid under certain conditions.¹⁾ Because of these characteristics, this enzyme is widely used in dairy or other food processing. Especially, lipase has recently gained much attention as an enzyme that is useful for transesterification, synthesis of esters, and resolution of racemic mixtures to produce optically active compounds.²⁾

A filamentous fungus, *Rhizopus niveus*, produces an extracellular lipase having 1·3 positional specificity.³⁾ This enzyme is especially useful in producing cocoa butter substitutes.³⁾ Lipases from various *Rhizopus* strains have been isolated and biochemically analyzed.^{4–8)} However, little is known of the primary structure of lipases. We have started molecular cloning of the gene for *R. niveus* lipase to get information on the structure–function relationship of lipases and to improve the properties of the enzyme by protein engineering techniques.

In this paper, cloning of the *R. niveus* lipase cDNA, and analysis of the nucleotide and amino acid sequences are described. The mature RNL sequence shows significant overall homology to the lipase from *Rhizomucor miehei*.

Materials and Methods

Strains, plasmids and media. *R. niveus* IF04759 was used as a strain producing lipase and a messenger RNA donor. *Escherichia coli* strains JM109,⁹⁾ Y1088,¹⁰⁾ Y1090,¹⁰⁾ and MV1184,¹¹⁾ plasmid vectors pUC8⁹⁾ and pUC118,¹¹⁾ and the phage vector λ gt11¹⁰⁾ were used for DNA manipulation.

R. niveus was cultivated at 27°C in peptone medium (5% peptone, 2% glucose, 0.02% yeast extract, 0.1% KH₂PO₄, and 0.05% MgSO₄, pH 6.5) and *E. coli* in LB medium¹²⁾ at 37°C.

Purification of RNL. RNL protein was isolated by the purification procedure for *Rhizopus delemar* C-lipase described by Iwai and Tsujisaka⁵⁾ with minor modifications. *R. niveus* was grown in peptone medium at 27°C for 4 days with shaking. The culture filtrate was adjusted to pH 4.0 with 1 N HCl and kept at 30°C for 30 min. The precipitate was separated from the supernatant by centrifugation. The supernatant was treated with (NH₄)₂SO₄ at 0.7 saturation. The precipitate was collected by centrifugation and dissolved with deionized water. The solution was put on a column of Sephadex G-25 (1.8 × 25 cm) to remove ammonium sulfate. To the desalted solution, 1 M acetate buffer (pH 4.0) and 1 M NaCl were added to give 0.01 M and 0.05 M, respectively. The solution was then put on a column of DEAE-Sephadex A-50 (4.8 × 70 cm) previously equilibrated with the same buffer. Eluate from the column with lipase activity was salted out with (NH₄)₂SO₄ at 0.7 saturation. The precipitate was dissolved in 2 ml of deionized water, and the solution

was desalted by Sephadex G-25 column chromatography. One M acetate buffer (pH 4.7) was added to give 0.05 M to the desalted solution. The solution was then put on a column of CM-Sephadex C-50 (2.2 × 30 cm) pre-equilibrated with 0.05 M acetate buffer (pH 4.7). After washing with the same buffer, substances on the column were eluted with a linear gradient of increasing NaCl concentrations (0–0.5 M). The main fraction with lipase activity was collected and precipitated with (NH₄)₂SO₄ at 0.7 saturation. The precipitate was dissolved with deionized water, and the solution was desalted with a Sephadex G-25 column. To the desalted solution, 1 M NaCl and 1 M acetate buffer (pH 6.0) were added to give 0.02 M, and this was chromatographed on a Cellulofine GC-200-m column (1.8 × 70 cm) equilibrated with the same buffer. From 5 l of culture filtrate, about 50 mg of purified enzyme was isolated. The lipase fraction was finally purified by reverse-phase HPLC (column: RP300 C-8, 300A, Applied Biosystems) using a linear gradient of acetonitrile (0–70%) containing 0.1% trifluoroacetic acid. The main peak was used as a purified RNL.

Amino acid analysis of RNL. NH₂-terminal amino acid sequencing of the purified RNL was done with a gas-phase protein sequencer (Model 473A, Applied Biosystems). For analysis of the amino acid composition, the purified RNL was hydrolyzed in 6 N HCl at 110°C for 24, 48, and 72 hr, and the hydrolyzate was then analyzed with an amino acid analyzer (Model L-8500, Hitachi).

Construction of a cDNA library. *R. niveus* was grown to late log phase in the peptone medium at 27°C for 4 days with shaking. The culture broth was centrifuged to obtain mycelia. The total cellular RNA was prepared by the method of Chirgwin *et al.*¹³⁾ From 15 g of wet mycelia, 30 mg of total RNA was isolated. Ten μ g of poly(A)⁺ RNA was then purified using 5 mg of total RNA by oligo(dT)-cellulose column chromatography as described by Maniatis *et al.*¹⁴⁾ From 5 μ g of poly(A)⁺ RNA, double stranded cDNA was synthesized using a cDNA synthesis kit (Amersham). Then, *Eco*RI linkers (dpCCGAATTCCGG) were attached to both ends of the cDNA and ligated with *Eco*RI-digested λ gt11 as described by Young and Davis.¹⁰⁾ After ligation with T4 DNA ligase, DNA was packed using an *in vitro* packaging kit (Promega Biotec), giving a cDNA library consisting of approximately 1.0×10^5 independent clones.

Screening of the RNL cDNA clone. Four 14-mer-oligonucleotides (Fig. 2) corresponding to the amino acid sequence of the lipase were synthesized by the phosphoramidite method using a DNA synthesizer (Model 381A, Applied Biosystems), radiolabeled with [γ -³²P]ATP (Amersham), and their mixture was used as the probe for finding RNL cDNA from the library. Plaques were transferred onto a Hybond-N nylon membrane (Amersham) and then fixed by UV-irradiation for 5 min. Hybridization was done at 45°C for 4 hr in a solution containing 6 × SSC, 0.5% SDS, and 100 μ g/ml salmon sperm DNA. After hybridization, the membranes were washed for 1 hr in a solution of 6 × SSC with shaking at room temperature. A positive clone was detected by autoradiography at –70°C with Fuji RX film and intensified screen.

DNA Sequencing. The cDNA insert of λ RL106B was subcloned into plasmid pUC118, and terminally deleted by a deletion kit for kilo-

sequencing (Takara Shuzo). A series of plasmids containing deleted cDNA inserts was used as templates for DNA sequencing. DNA sequencing was done with the dideoxynucleotide chain termination method.¹⁵⁾ The sequences obtained were analyzed with the GENETYX Program from Software Development Co., Ltd., Japan.

Results and Discussion

NH₂-Terminal amino acid sequence and molecular weight of RNL

When the purified RNL was analyzed by SDS-PAGE, a single band was obtained as shown in Fig. 1. Its molecular mass was estimated to be about 34 kDa from the mobility relative to markers. The *NH₂*-terminal amino acid sequence of RNL was found by Edman degradation to be Asp-Asp-Asn-Leu-Val-Gly-Gly-Met-Thr-Leu-Asp-Leu-Pro-Ser-Asp-Ala-Pro-Pro-Ile-Ser-, and used to design oligonucleotide probes for finding RNL cDNA from the library.

Cloning of the lipase cDNA

A cDNA library in λ gt11 was constructed and a clone having RNL cDNA was screened with four oligonucleotide probes corresponding to the amino acids from 5 to 9 of

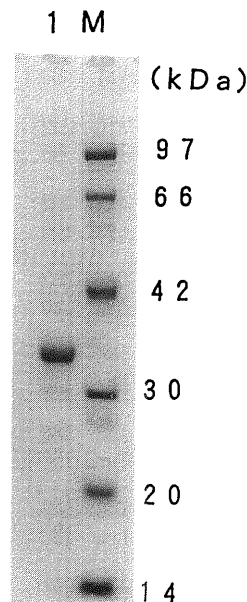


Fig. 1. SDS-PAGE of RNL.

Molecular weight markers (lane M) and the purified RNL (lane 1) were electrophoresed on a gradient SDS-polyacrylamide gel (10–20%) and protein bands were stained with Coomassie brilliant blue R250. Numbers by the side of lane M are molecular masses of marker proteins (kilodaltons).

	5	6	7	8	9
Amino acid sequence	-Val	-Gly	-Gly	-Met	-Thr-
mRNA	5'-GUU	-GGU	-GGU	-AUG	-AC-3'
	C	C	C		
	A	A	A		
	G	G	G		
Probe 1	3'-CAX-CCX-CCZ-TAC-TG-5'				
2	3'-CAX-CCY-CCZ-TAC-TG-5'				
3	3'-CAY-CCX-CCZ-TAC-TG-5'				
4	3'-CAY-CCY-CCZ-TAC-TG-5'				

(X:A, G Y:T, C Z:A, G, T, C)

Fig. 2. Synthetic Oligonucleotide Probes Used for Isolating RNL cDNA.

RNL sequence (Fig. 2) as described in Materials and Methods. A positive plaque (λ RL106B) was obtained from the 3,000 plaques screened. λ RL106B had an insert of about 1.0 kilobase pairs that hybridized with the probes.

Expression of the lipase cDNA in Escherichia coli

The 1.0-kb *Eco*RI fragment of λ RL106B was ligated into the *Eco*RI site of pUC8, and used to transform *E. coli* JM109. The resulting transformant, *E. coli* JM109 containing pRL1068B (Fig. 3), showed the lipase activity on a tributyrin agar plate as shown in Fig. 4. The nucleotide sequence around the junction of the pUC8 and cDNA in pRL1068B was analyzed. As shown in Fig. 3, a nucleotide sequence corresponding to the *NH₂*-terminal sequence of RNL was found in the same reading frame as *lacZ*. These results showed that the lipase gene existed in the insert DNA of λ RL106B, and the lipase cDNA was expressed in *E. coli* JM109 harboring pRL1068B as a *lacZ*-lipase fusion protein.

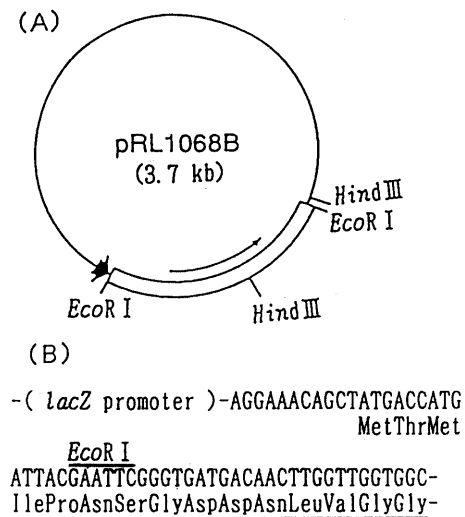


Fig. 3. Restriction Map of pRL1068B.

(A): The thin line represents the pUC8 moiety and the double line the cloned cDNA moiety. The transcription direction of the lipase gene is indicated by the arrow. The solid arrow represents the *lacZ* promoter region.

(B): Nucleotide and predicted amino acid sequences around the junction of the *lacZ* promoter and cDNA in the pRL1068B. The nucleotide sequence was determined using oligodeoxyribonucleotide, d(CAGGAAACAGCTATGAC), (Takara Shuzo) as a primer. The nucleotide sequence corresponding to the primer is cited from the published sequence.⁹⁾

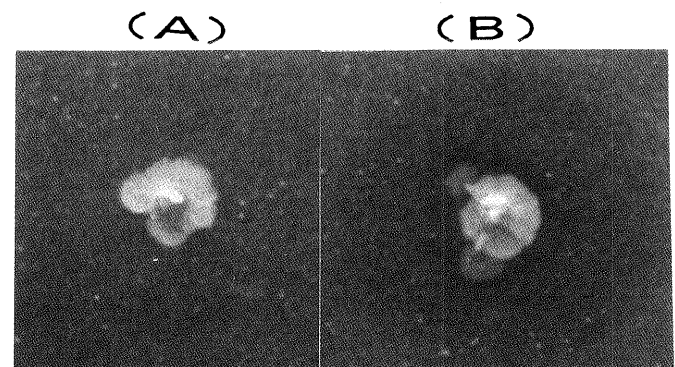


Fig. 4. Expression of the Lipase cDNA in *E. coli* on a Tributyrin Agar Plate.¹⁸⁾

(A): *E. coli* JM109 containing pUC8.

(B): *E. coli* JM109 harboring pRL1068B.

Sequencing of the lipase cDNA

The cDNA insert was sequenced as described in Materials and Methods. The nucleotide sequence of lipase cDNA and the amino acid sequence deduced from its sequence are shown in Fig. 5. It consisted of 958 nucleotides included a 10-nucleotide-long poly(A)⁺ tail and *EcoRI* linker sequences. A putative polyadenylation signal, AATAAA, was located at position 925–930, 19 bp upstream from the poly(A)⁺ tail.

The NH₂-terminal sequence deduced from the Edman degradation of the protein was found to start at residue 1 in Fig. 5 (starting at nucleotide 3). The COOH-terminal position was found at Leu297 (nucleotides 891–893), which immediately proceeded a termination codon, TAA. The predicted mature RNL consisted of 297 amino acids with a calculated molecular mass of 32,349 daltons. This

data coincided well with a molecular mass of 34 kDa on SDS-PAGE. The measurement of the sugar content of the purified RNL by the phenol–sulfuric acid method¹⁹⁾ showed that there was no detectable amount of the sugar. The amino acid composition of RNL from the amino acid analysis agreed well with that calculated from the sequence data (Table 1). From these results, it is reasonable to conclude that though this cDNA clone does not contain the full length of the lipase gene, it contains the coding region of the mature enzyme.

Amino acid sequences homology between the lipases of *R. niveus* and *Rhizomucor miehei*

A comparison of the amino acid sequences of RNL and lipase from *Rhizomucor miehei* (RML) is shown in Fig. 6. The sequence of mature RNL was 28 amino acids

	linker ccgaattcggGT	2
GAT GAC AAC TTG GTT GGT GGC ATG ACT TTG GAC TTA CCC AGC GAT GCT CCT CCT ATC AGC		62
Asp Asp Asn Leu Val Gly Gly Met Thr Leu Asp Leu Pro Ser Asp Ala Pro Pro Ile Ser		20
CTC TCT AGC TCT ACC AAC AGC GCC TCT GAT GGT GGT AAG GTT GTT GCT GCT ACT ACT GCT		122
Leu Ser Ser Ser Thr Asn Ser Ala Ser Asp Gly Gly Lys Val Val Ala Ala Thr Thr Ala		40
CAG ATC CAA GAG TTC ACC AAG TAT GCT GGT ATC GCT GCC ACT GCC TAC TGT CGT TCT GTT		182
Gln Ile Gln Glu Phe Thr Lys Tyr Ala Gly Ile Ala Ala Thr Ala Tyr Cys Arg Ser Val		60
GTC CCT GGT AAC AAG TGG GAT TGT GTC CAA TGT CAA AAG TGG GTT CCT GAT GGC AAG ATC		242
Val Pro Gly Asn Lys Trp Asp Cys Val Gln Cys Gln Lys Trp Val Pro Asp Gly Lys Ile		80
ATC ACT ACC TTT ACC TCC TTG CTT TCC GAT ACA AAT GGT TAC GTC TTG AGA AGT GAT AAA		302
Ile Thr Thr Phe Thr Ser Leu Leu Ser Asp Thr Asn Gly Tyr Val Leu Arg Ser Asp Lys		100
CAA AAG ACC ATT TAT CTT GTT TTC CGT GGT ACC AAC TCC TTC AGA AGT GCC ATC ACT GAT		362
Gln Lys Thr Ile Tyr Leu Val Phe Arg Gly Thr Asn Ser Phe Arg Ser Ala Ile Thr Asp		120
ATC GTC TTC AAC TTT TCT GAC TAC AAG CCT GTC AAG GGC GCC AAA GTT CAT GCT GGT TTC		422
Ile Val Phe Asn Phe Ser Asp Tyr Lys Pro Val Lys Gly Ala Lys Val His Ala Gly Phe		140
CTT TCC TCT TAT GAG CAA GTT GTC AAT GAC TAT TTC CCT GTC GTC CAA GAA CAA TTG ACC		482
Leu Ser Ser Tyr Glu Gln Val Val Asn Asp Tyr Phe Pro Val Val Gln Glu Gln Leu Thr		160
GCC CAC CCT ACT TAT AAG GTC ATC GTT ACC GGT CAC TCA CTC GGT GGT GCA CAA GCT TTG		542
Ala His Pro Thr Tyr Lys Val Ile Val Thr Gly His Ser Leu Gly Gly Ala Gln Ala Leu		180
CTT GCC GGT ATG GAT CTC TAC CAA CGT GAA CCA AGA TTG TCT CCC AAG AAT TTG AGC ATC		602
Leu Ala Gly Met Asp Leu Tyr Gln Arg Glu Pro Arg Leu Ser Pro Lys Asn Leu Ser Ile		200
TTC ACT GTC GGT GGT CCT CGT GTT GGT AAC CCC ACC TTT GCT TAC TAT GTT GAA TCC ACC		662
Phe Thr Val Gly Gly Pro Arg Val Gly Asn Pro Thr Phe Ala Tyr Tyr Val Glu Ser Thr		220
GGT ATC CCT TTC CAA CGT ACC GTT CAC AAG AGA GAT ATC GTT CCT CAC GTT CCT CCT CAA		722
Gly Ile Pro Phe Gln Arg Thr Val His Lys Arg Asp Ile Val Pro His Val Pro Pro Gln		246
TCC TTC GGA TTC CTT CAT CCC GGT GTT GAA TCT TGG ATG AAG TCT GGT ACT TCC AAC GTT		782
Ser Phe Gly Phe Leu His Pro Gly Val Glu Ser Trp Ile Lys Ser Gly Thr Ser Asn Val		260
CAA ATC TGT ACT TCT GAA ATT GAA ACC AAG GAT TGC AGT AAC TCT ATC GTT CCT TTC ACC		842
Gln Ile Cys Thr Ser Glu Ile Glu Thr Lys Asp Cys Ser Asn Ser Ile Val Pro Phe Thr		280
TCT ATC CTT GAC CAC TTG AGT TAC TTT GAT ATC AAC GAA GGA AGC TGT TTG TAA AACACTT		903
Ser Ile Leu Asp His Leu Ser Tyr Phe Asp Ile Asn Glu Gly Ser Cys Leu *		297
GACGTGTACTCTAATTTTATATAAAATTAAAGTTTTATACAATAAAAAAAAAAccgaattcgg		958
	linker	

Fig. 5. Nucleotide Sequence of the Lipase cDNA and Deduced Amino Acid Sequence of RNL.

The amino acid confirmed by protein sequencing are underlined. The putative poly(A)⁺ additional signal is boxed. The nucleotide sequence hybridized with the synthetic oligonucleotide probe is overlined. * is a stop codon.

```

1' DDNLVGGMTLDLPSDAPPISLSSSTNSASDGGKVVAAATTAQIQEFTKYAGIAATAYCRSV
      * * * * *
1" SIDGGIRAATSQEINELTYTTLSANSYCRTV
      * * * * *
61' VPGNKWDVQCQKWVPDGGKIITFTSLLSDTNGYVLRSDKQKTIYLVFRGTNSFRSAITD
      * * * * *
33" NPGATWDCIHCDA-TEDLKIITWSTLIYDTNAMVARGDSEKTIYIVFRGSSSIRNWIAD
      * * * * *
121' IVFNFSDYKPKVKGAHVHAGFLSSYEQVVDYFPVVQEQLTAHPTYKVIIVTGHSLGGAQAL
      * * * * *
92" LTFVPVSYPVSGTKVHKGFLDSYGEVQNELVATVLDQFKQPSYKVAVTGHSLGGATAL
      * * * * *
181' LAGMDLYQREPRLSPKNLSIFTVGGPRVGNPTFAYYVESTGIPFQRTVHKRDIVPHVPPQ
      * * * * *
152" LCALGLYQREEGLSSSNLFLYTQGGPRVGDPAFANYVVSTGIPYRRTVNERNRDIIVPHLPPA
      * * * * *
241' SFGFLHPGVESWIKSGTSN-VQICTSEIETKDCNSIVPFTSILDHLSYFDINEGSC
      * * * * *
212" AFGFLHAGEEYWITDNSPETVQVCTSDLETDCNSIVPFTSVLDHLSYFGINTGLCT
      * * * * *

```

The sequence of the mature RNL (upper line) is aligned with that of the mature RML using the GENETYX Program HOMOGAPP. The closed circles represent the putative active site residues. Identical cystein residues are boxed.

Amino acid	DNA sequence	Protein analysis*
Asp (Asx)	17	(30.1)
Asn	12	
Thr	24	23.8
Ser	30	32.0
Glu (Glx)	9	(25.5)
Gln	13	
Pro	18	21.1
Gly	23	24.9
Ala	17	17.4
Cys	6	ND
Val	27	30.1
Met	2	1.4
Ile	18	16.8
Leu	20	20
Tyr	12	12.1
Phe	15	15.8
Lys	15	15.4
His	7	6.4
Arg	9	9.0
Trp	3	ND

The X-ray structure of RML was recently reported and showed that the active center of RML contains an Asp-His-Ser triad that is characteristics of serine proteinase.¹⁷⁾ The putative active site residues, Ser144, Asp203, and His257 of RML were strictly conserved in RNL. RML has three disulfide bridges in the positions of Cys29-Cys268, Cys40-Cys43, and Cys235-Cys244.¹⁷⁾ These six cystein residues were also completely conserved in RNL. Therefore, Ser173, Asp232, and His285 of RNL may be the amino acid residues of the catalytic triad, and three disulfide bridges at Cys57-Cys296, Cys68-Cys71, and

References

- 1) Y. Tsujisaka, S. Okumura, and M. Iwai, *Biochim. Biophys. Acta*, **489**, 415—422 (1977).
- 2) F. Xavier, H. R. Reyes, H. S. Garcia, C. G. Hill, and C. H. Amundson, *J. Am. Oil. Chem. Soc.*, **67**, 890—910 (1990).
- 3) T. Matsuo, N. Sawamura, Y. Hashimoto, and W. Hashida, U. K. Patent 35359 (1980).
- 4) M. Iwai and Y. Tsujisaka, in "Lipases," ed. by B. Borgstrom and H. L. Brockman, Elsevier, New York, 1984, pp. 443—469.
- 5) M. Iwai and Y. Tsujisaka, *Agric. Biol. Chem.*, **38**, 1241—1247 (1974).
- 6) M. Semeriva, G. Benzonana, and P. Desnuelle, *Biochim. Biophys. Acta*, **191**, 598—610 (1969).
- 7) K. Aisaka and D. Terada, *J. Biochem.*, **89**, 817—822 (1981).
- 8) K. D. Davranov and M. M. Rakhimov, *Biokhimiya*, **46**, 250—255 (1981).
- 9) C. Yianisch-Perron, J. Vieira, and J. Messing, *Gene*, **33**, 103—119 (1985).
- 10) R. A. Young and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 1194—1198 (1983).
- 11) J. Vieira and J. Messing, *Methods Enzymol.*, **153**, 3—11 (1987).
- 12) R. W. Davis, D. Botstein, and J. R. Roth, "Advanced Bacterial Genetics," Cold Spring Harbor Laboratory, New York, 1980, pp. 201.
- 13) J. M. Chirgwin, A. E. Przybyla, R. J. MacDonald, and W. J. Rutter, *Biochemistry*, **18**, 5294—5299 (1979).
- 14) T. Maniatis, E. F. Fritsch, and J. Sambrook, "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982.
- 15) F. Sanger, *Science*, **214**, 1205—1210 (1981).
- 16) E. Boel, B. Høge-Jensen, M. Christensen, L. Thim, and N. P. Fill, *Lipids*, **23**, 701—706 (1988).
- 17) L. Brady, A. M. Brzozowski, Z. S. Derewenda, E. Dodson, G. Dodson, S. Tolley, J. P. Turkenburg, L. Chritiansen, B. Høge-Jensen, L. Nørskov, L. Thim, and U. Menge, *Nature*, **343**, 767—770 (1990).
- 18) W. Kugimiya, Y. Otami, Y. Hashimoto, and Y. Takagi, *Biochem. Biophys. Res. Commun.*, **141**, 185—190 (1986).
- 19) M. Dubois, K. A. Gilles, J. K. Hamilton, P. A. Rebers, and F. Smith, *Anal. Chem.*, **28**, 350 (1956).