Supporting Information

Stereoselective Synthesis of Chiral δ-Lactones via an Engineered Carbonyl Reductase

Tao Wang,[†,a] Xiao-Yan Zhang,[†,a] Yu-Cong Zheng,[a] and Yun-Peng Bai*[,a]

[a] State Key Laboratory of Bioreactor Engineering, Shanghai Collaborative Innovation Center for Biomanufacturing, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China.

[†] These authors contributed equally.

Contents

1. Experimental Section

General information

All chemicals were purchased from TCI (Japan) and Aladdin and Shaoyuan (Shanghai, China), and used without further treatment unless otherwise indicated. NADPH and $NADP^+$ were purchased from Bontac Bioengineering (Shenzhen, China). Luria–Bertani (LB) medium was used to culture *E. coli* cells. GC was performed with a Shimadzu GC-2014 system with a CP-Chirasil-Dex CB column (25 m × 0.25 mm × 0.39 mm, Varian) or Rxi®-5Sil MS column (30 m× 0.25 mm × 0.25 μm) and a flame ionization detector. ^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) spectra were recorded in $CDCl_3$. Column chromatography was performed with silica gel of particle size 200-300 mesh. *n*-Dodecane was used as an internal standard.

Directed evolution

A recombination plasmid containing the *Sm*CR_{V4} gene was used as the template for random mutagenesis. $MnCl_2$ (0.25 mM) was used to obtain the desired mutagenesis rate (one to two amino acid substitutions). The amplified PCR products were extracted, digested with *Eco*R I and *Hin*d III, ligated into the *Eco*R I and *Hin*d III sites of pET-28a, and then transformed into chemically competent *E. coli* BL21 (DE3) cells. The recombination plasmid containing the *Sm*CR_{V4} gene used was amplified by PCR with NNK codon degeneracy. The resulting PCR products were digested with Dpn I (20 U) at 37 °C for 3 h and then transformed into chemically competent *E. coli* BL21 (DE3) cells and plated on LB agar plates containing 50 μg/mL kanamycin. The colonies were picked with sterile toothpicks to inoculate LB medium (300 μL) containing 50 μg/mL kanamycin into 96-well plates. The cultures were grown overnight at 37 °C prior to inoculating LB medium (600 μL) into new 96-well plates. The plates were incubated at 37 °C for 2 h and protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (0.2 mM) and incubating at 16 °C for 24 h. Cells were lyzed by adding buffer (300 μL) containing lysozyme (1 mg/mL) and DNase I (0.01 mg/mL), and incubating at 37 °C for 2 h. The plates were centrifuged at 3500 rpm for 15 min at 4 °C. A sample (100 μL) from each well was transferred to a microtiter plate and then the reduction reaction was initiated by adding a mixture of 100 mM phosphate buffer (100 μL, pH 6.0), 0.3 mM NADPH, and 8 mM **1h**. The variant activities were determined by measuring the change in the NADPH absorbance at 340 nm for 10 min at 35 °C by using a microplate spectrophotometer (BioTek, USA). Variants with higher activities were chosen for re-screening in 96-deep-well plates, and the best performers were grown on a 100 mL scale. The best variants were selected for sequencing and

purified by N-terminal His-tagged nickel affinity chromatography for further characterization.

Protein Crystallization

The sitting-drop vapor-diffusion technique was used for *Sm*CR_{V4}. The purified *Sm*CR_{V4} protein was concentrated to 15 mg mL^{-1} and incubated with 0.6 mM $NADP^+$. Crystallization of *Sm*CR_{V4} was performed by mixing the incubated protein (1 μL) with an equal volume of reservoir solution containing 0.1 M sodium formate, 0.1 M ammonium acetate, 0.1 M sodium citrate tribasic dihydrate, 0.1 M potassium sodium tartrate tetrahydrate, 0.1 M sodium oxamate, 0.1 M imidazole, 0.1 M MES monohydrate (pH 6.50), 12.5% v/v 2-methyl-1,3-propanediol, 12.5% polyethylene glycol 1000 and 12.5% w/v polyethylene glycol 3350 at 18 °C. Crystals were observed after 3-4 days. The crystals were then cryoprotected by transient soaking in the corresponding reservoir solution with additional 30% glycerol and transferred into liquid nitrogen. Data collection experiments were performed at the BL19U1 beamline at the Shanghai Synchrotron Radiation Facility (SSRF, China), and the data set was indexed and processed by using the HKL-2000 program suite. The *Sm*CR_{V4} crystal form diffracted in the 2.49-2.45 Å resolution range and belongs to the hexagonal space group $C222_1$. The structure was determined by CCP4 by using the coordinates of a β-ketoacyl-acyl carrier protein reductase (PDB ID: 1Q7B) as the model. The final *R*-work/*R*-free values for the model were 18.7/24.9 after modification by CCP4. The atomic coordinates of *Sm*CR_{V4} in a complex with $NADP^+$ have been uploaded to the Protein Data Bank (PDB) under accession number 7EMG.

Enzyme assay and kinetic analysis

The reductase activity was assayed at 40 °C by monitoring the decrease in the absorbance of NADPH at 340 nm with a UV spectrophotometer (Shimadzu UV-1900). The standard assay mixture (1 mL) consisted of sodium phosphate buffer (100 mM, pH 6.0, 960 μL), 5-oxodecanoic acid (200 mM, 20 μL), NADPH (10 mM, 10 μL), and pure enzyme (10 μL) at an appropriate concentration. One unit of enzyme activity (1 U) is defined as the amount of enzyme that can catalyze the oxidation of 1 μmol of NADPH per minute under the above conditions. The substrates were dissolved in dimethyl sulfoxide to give various final concentrations and then the specific activities were determined with a UV spectrophotometer. The data were processed using Origin 2018 software, based on the Michaelis-Menten equation.

Analytical methods

The stereoconfigurations of the products were determined in 1 mL of reaction

solution. The solution was prepared by mixing 4 mM substrate, 10 mM glucose, 0.2 mM $NADP^+$, an appropriate amount of enzyme solution, and *Bm*GDH lyophilized crude enzymes (5 mg), and stirring at 30°C. The reaction was terminated by adding 2 M sulfuric acid solution and the reaction mixture was extracted with ethyl acetate after heating at 80°C for 1 h. Gas chromatographic analyses (GC) was conducted on a Shimadzu GC-2010 chromatograph equipped with a flame ionization detector (FID) and a CP-Chirasil-Dex CB column (25 m×0.25 mm×0.39 μm, Varian). Nitrogen was used as the carrier gas. Temperatures of the inlet and detector were set to 280°C. The enantioselectivities of lactone products were determined by chiral GC analysis.

Preparation of structurally diverse δ-lactones

The reaction mixture (50 mL) was prepared by dissolving the substrate (50-400 mmol/L), glucose (1.5 equiv with respect to the substrate), crude *Sm*CR_{M5} enzyme, lyophilized crude *Bm*GDH enzymes, 5% dimethyl sulfoxide (v/v), and $NADP^+$ (0.2 mM) in a sodium phosphate buffer (100 mM, pH 6.0). The pH of the mixture was maintained at 6.0 by using 1 M Na_2CO_3. The reaction was performed at 30°C and 300 rpm for an appropriate time, which depended on the reaction conditions. The reaction was terminated by acidification to pH 2.0 and the mixture was extracted with ethyl acetate. The organic layer was dried over anhydrous Na_2SO_4. Solvent removal afforded the hydroxy acid, which was heated at 80°C for 1 h to complete lactonization. The residue was purified by flash column chromatography with a mixture of ethyl acetate and petroleum ether (1:10) as the eluent, to provide the pure δ-lactone. ^1H NMR (400 MHz) and ^{13}C NMR (100 MHz) spectra of the pure δ-lactones in $CDCl_3$ were obtained.

2. Additional tables and figures

Table S1. Beneficial variants obtained in the first round of random mutagenesis

Entry	Enzyme	Variants	Specific activity ($U\ mg^{-1}$)[a]	*ee* (%)	Fold
1	*Sm*CR_{V4}	-	0.33±0.011	95 (*R*)	1
2	*Sm*CR_{M1}	E23K	0.635±0.007	93 (*R*)	1.9
3	*Sm*CR_{M1-1}	K127E	0.538±0.016	93 (*R*)	1.6

[a] Specific activity was determined spectrophotometrically at 40°C in sodium phosphate buffer (100 mM, pH 6.0) containing 4 mM substrate, 0.1 mM NADPH and a proper amount of purified enzymes.

Table S2. X-ray data collection and refinement statistic of *Sm*CR$_{V4}$

		*Sm*CR$_{V4}$
PDB code		7EMG
Space group		C222$_1$
Cell dimens		
	a, b, c (Å)	63.8, 120.3, 120.7
	α, β, γ (deg)	90, 90, 90
Resolution range (Å)		41.2–2.5
Highest resolution shell (Å)		2.49–2.45
R_{merge} (%)		17.0 (9.4/90)
I/σ(I)		15.0 (4.1/1.0)
No. of unique reflections		17156 (849)
Completeness (%)		99.2
Redundancy		12.8
Refinement		
Resolution (Å)		2.45
R_{work}/R_{free} (%)		18.7/24.9
No. of water molecules		61
B factors		
	protein (Å^2)	36.0
	NDP (Å^2)	56.7
	water (Å^2)	31.8
RMS deviations		
	bond lengths (Å)	0.016
	bond angles (deg)	1.01
Ramachandran plot		
	favored (%)	95.3
	allowed (%)	4.7
	outliers (%)	0

Table S3. Beneficial variants obtained by site-saturation mutagenesis

Entry	Enzyme	Variants	Specific activity (U mg^{-1})[a]	*ee* (%)	Fold
1	*Sm*CR_{V4}	-	0.33±0.011	95 (*R*)	1
2	*Sm*CR_{M2}	N145R	1.16±0.013	98 (*R*)	3.5
3	*Sm*CR_{M2-1}	K61V	1.16±0.029	99 (*R*)	3.5

[a] Specific activity was determined spectrophotometrically at 40°C in sodium phosphate buffer (100 mM, pH 6.0) containing 4 mM substrate, 0.1 mM NADPH and a proper amount of purified enzymes.

Table S4. Comparison of iterative combination variants

Entry	Enzyme	Variants	Specific activity (U mg^{-1})[a]	*ee* (%)	Fold
1	*Sm*CR_{V4}	-	0.33±0.011	95 (*R*)	1
2	*Sm*CR_{M3-1}	N145R/E23K	1.66±0.076	99 (*R*)	5.1
3	*Sm*CR_{M3-2}	N145/K127E	1.44±0.012	99 (*R*)	4.4
4	*Sm*CR_{M3-3}	N145R/E23K/K127E	1.83±0.077	99 (*R*)	5.6
5	*Sm*CR_{M3}	E23K/K61V/K127E/ N145R	1.94±0.12	99 (*R*)	5.9

[a] Specific activity was determined spectrophotometrically at 40°C in sodium phosphate buffer (100 mM, pH 6.0) containing 4 mM substrate, 0.1 mM NADPH and a proper amount of purified enzymes.

Table S5. Beneficial variants obtained in the second round of random mutagenisis

Entry	Enzyme	Variants	Specific activity (U mg^{-1})[a]	*ee* (%)	Fold
1	*Sm*CR_{V4}	-	0.33±0.011	95 (*R*)	1
2	*Sm*CR_{M4}	*Sm*CR_{M3}-A149T	3.85±0.28	99 (*R*)	11.7
3	*Sm*CR_{M4-1}	*Sm*CR_{M3}-L107I	2.08±0.086	99 (*R*)	6.3
4	*Sm*CR_{M4-2}	*Sm*CR_{M3}-E104K	2.53±0.07	99 (*R*)	7.7

[a] Specific activity was determined spectrophotometrically at 40°C in sodium phosphate buffer (100 mM, pH 6.0) containing 4 mM substrate, 0.1 mM NADPH and a proper amount of purified enzymes.

Table S6. Comparison of iterative combination variants

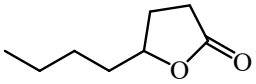
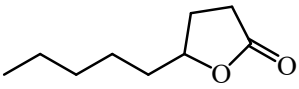
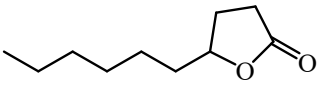
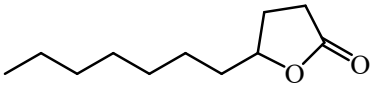
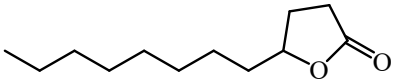
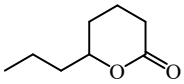
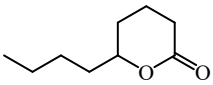
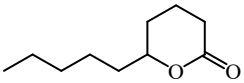
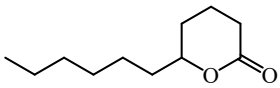
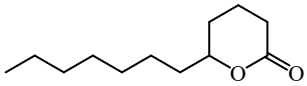
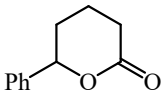
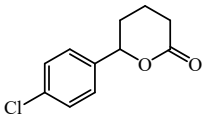
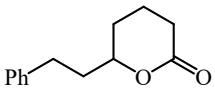
Entry	Enzyme	Variants	Specific activity (U mg^{-1})[a]	*ee* (%)	Fold
1	*Sm*CR_{V4}	-	0.33±0.011	95 (*R*)	1
2	*Sm*CR_{M5}	*Sm*CR_{M4}-E104K	4.55±0.23	99 (*R*)	13.9
3	*Sm*CR_{M5-1}	*Sm*CR_{M4}-L107I	3.82±0.12	99 (*R*)	11.6
4	*Sm*CR_{M5-2}	*Sm*CR_{M4}-E104K/L107I	4.18±0.11	99 (*R*)	12.7

[a] Specific activity was determined spectrophotometrically at 40°C in sodium phosphate buffer (100 mM, pH 6.0) containing 4 mM substrate, 0.1 mM NADPH and a proper amount of purified enzymes.

Table S7. The half-life of different variants at different temperatures

Enzyme	30°C		40°C		50°C	
	k_d/h^{-1}	$t_{1/2}$/h	k_d/h^{-1}	$t_{1/2}$/h	k_d/min^{-1}	$t_{1/2}$/min
*Sm*CR$_{M5}$	0.0041	170	0.0085	81	0.0704	9.8
*Sm*CR$_{M4}$	0.0034	202	0.0143	48	0.1326	5.2
*Sm*CR$_{M3}$	0.0029	239	0.0132	53	0.1026	6.5
*Sm*CR$_{M0}$[6]	0.0056	124	0.0096	72	0.0048	144

Table S8. GC conditions used to determine the enantioselectivity of corresponding products

Entry	Product	Method	Retention time (min)	
			R	*S*
1		100/8/2/120/5/2/130/1	23.48	24.13
2		110/0/1.5/120/5/0.8/130/10	23.09	23.73
3		110/0/2/140/10	22.77	23.19
4		90/1/4/130/3/1/150/5	35.39	35.79
5		140/50	38.83	39.74
6		110/40	29.04	28.12
7		110/60	50.25	49.01
8		80/10/1/100/30/0.5/130/10/10/180/10	116.74	116.0
9		130/60	47.33	46.53
10		135/70	59.41	58.52
11		110/0/3/140/10/2/170/5	37.04	36.56
12		110/1/10/160/5/10/180/25	32.09	32.95
13		110/1/10/160/5/10/180/20	26.69	26.39

[a] GC program: initial temp. (°C)/ time (min)/ slope (°C/min)/ temp. (°C)/ time (min)/ slope (°C/min)/ final temp. (°C)/ time (min).

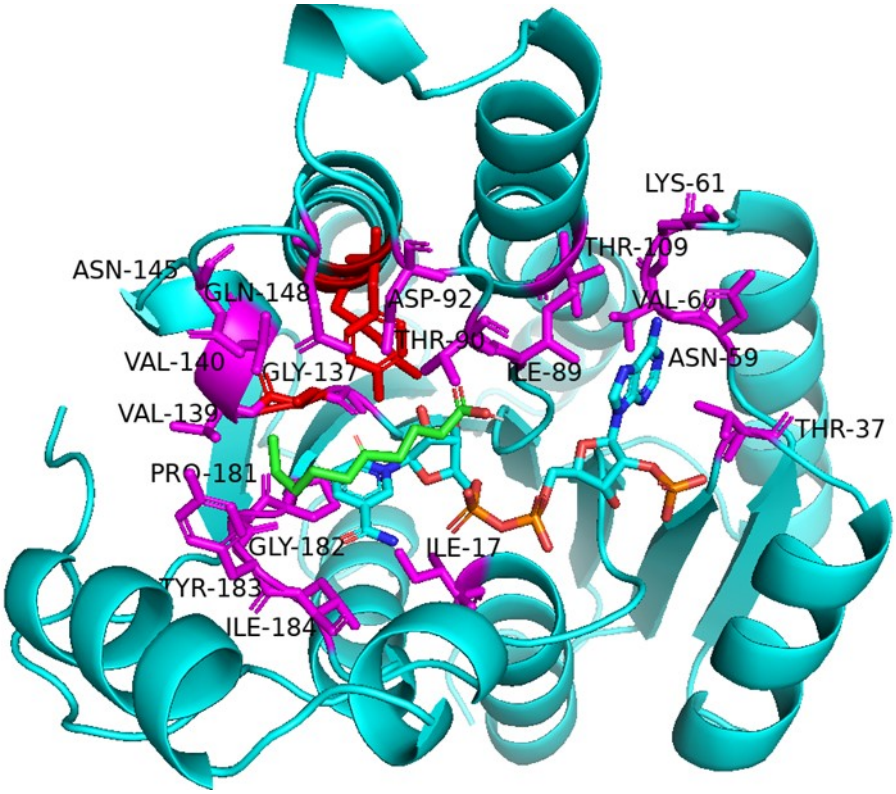


Figure S1. Structure model of *Sm*CR$_{V4}$. 18 Amino acid residues (in pink) were selected for site-saturation mutagenesis.

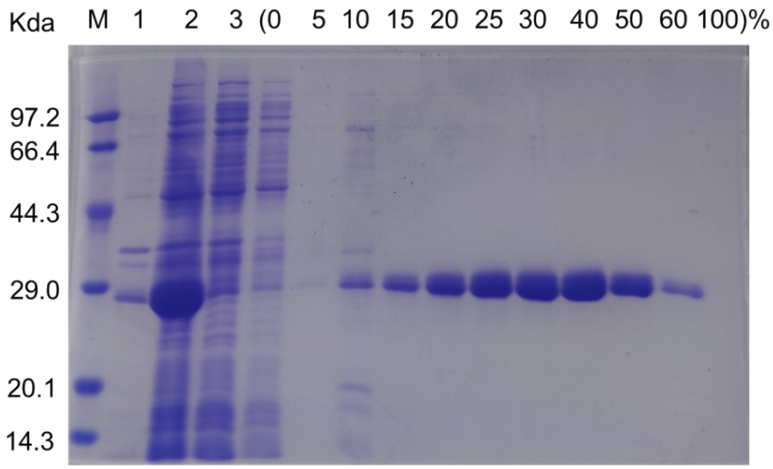


Figure S2. SDS-PAGE of pure *Sm*CR_{M5} by nickel ion affinity chromatography. M: Protein standard marker; line 1: precipitant of cell lysate; line 2: supernatant of cell lysate; line 3: flow-through of protein; other line: elution conditions with different volume ratio (buffer B: buffer A+B).

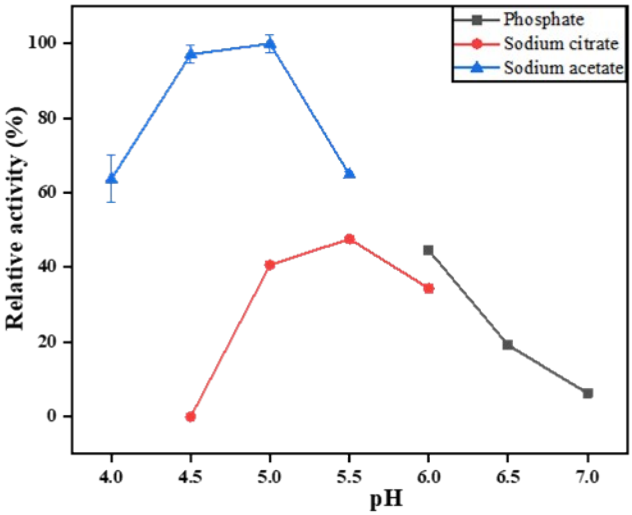


Figure S3. Effect of pH on the activity of purified *Sm*CR_{M5}. Error bars (standard deviation) were obtained from triplicate experiments at each point.

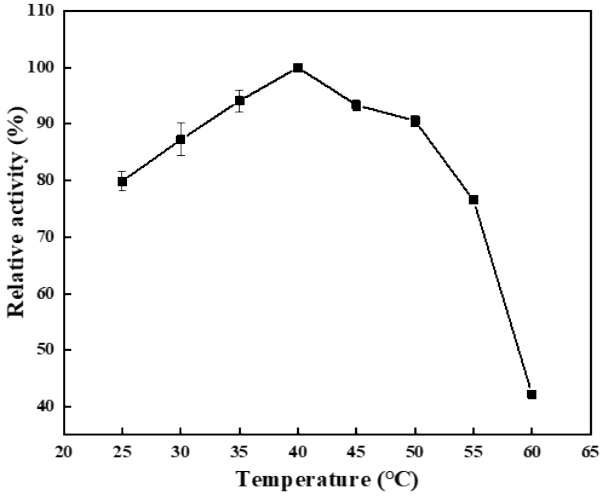


Figure S4. Effect of temperature on the activity of purified *Sm*CR_{M5}. Error bars (standard deviation) were obtained from triplicate experiments at each point.

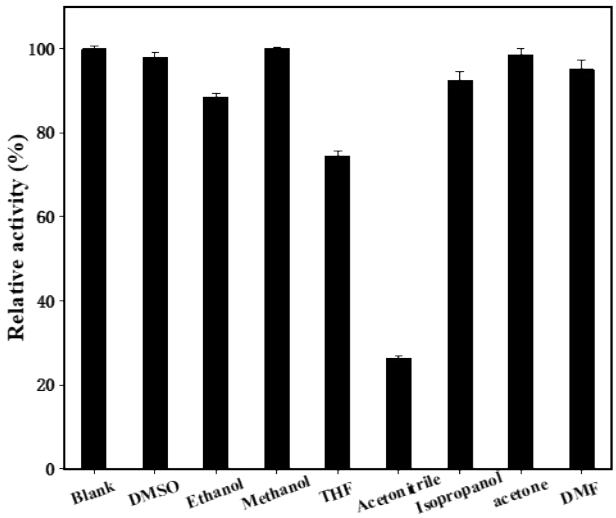


Figure S5. Effect of organic solvents on the activity of purified *Sm*CR_{M5}. Error bars (standard deviation) were obtained from triplicate experiments at each point.

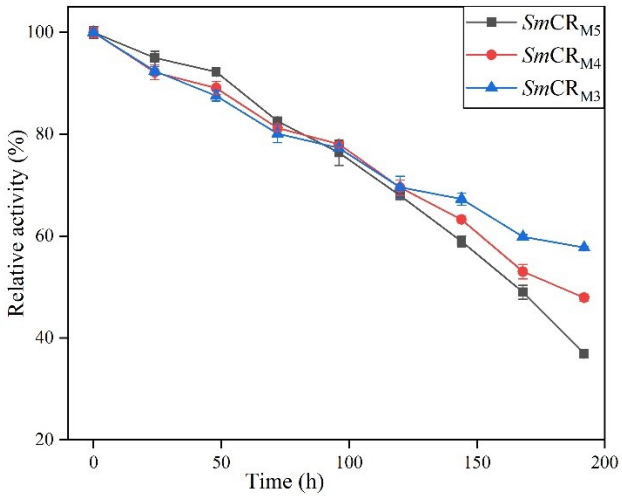


Figure S6. Thermostability of purified *Sm*CR$_{M5}$, *Sm*CR$_{M4}$ and *Sm*CR$_{M3}$ at 30℃. Error bars (standard deviation) were obtained from triplicate experiments at each point.

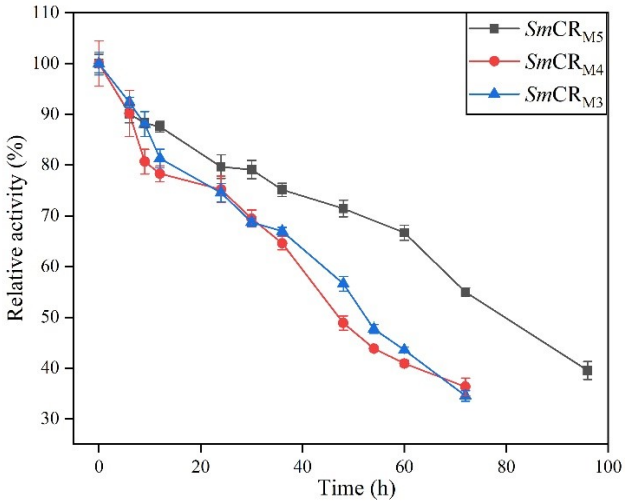


Figure S7. Thermostability of purified *Sm*CR$_{M5}$, *Sm*CR$_{M4}$ and *Sm*CR$_{M3}$ at 40°C. Error bars (standard deviation) were obtained from triplicate experiments at each point.

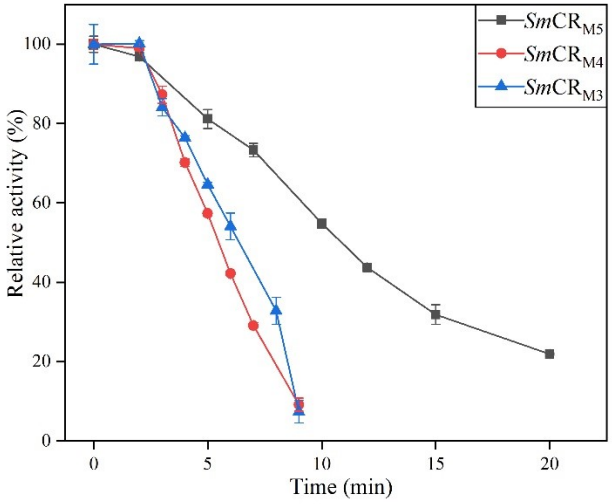


Figure S8. Thermostability of purified *Sm*CR$_{M5}$, *Sm*CR$_{M4}$ and *Sm*CR$_{M3}$ at 50°C. Error bars (standard deviation) were obtained from triplicate experiments at each point.

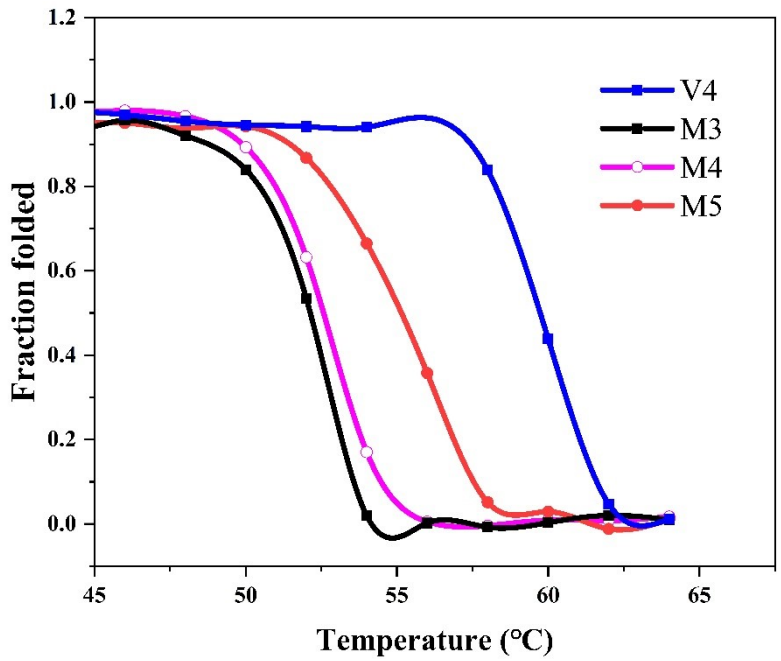


Figure S9. Thermal unfolding curves of V4 and variants analyzed by circular dichroism spectroscopy.

Enzyme	T_m(°C)
*Sm*CR$_{v4}$	59.8
*Sm*CR$_{M3}$	52.1
*Sm*CR$_{M4}$	52.6
*Sm*CR$_{M5}$	55.2

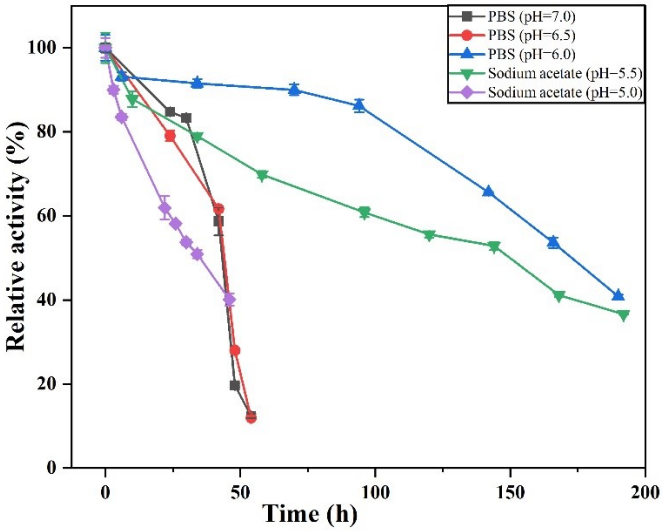
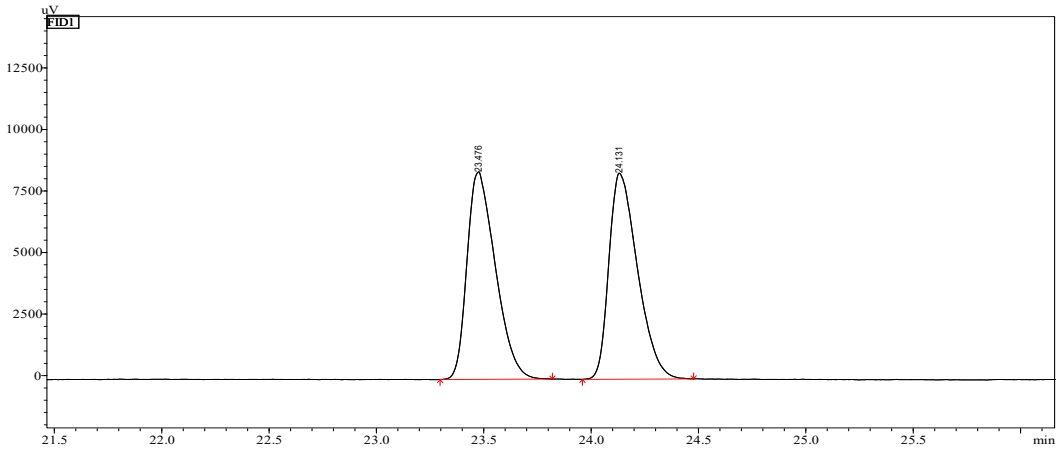


Figure S10. Thermostability of purified *Sm*CR_{M5} at different buffers. Error bars (standard deviation) were obtained from triplicate experiments at each point.

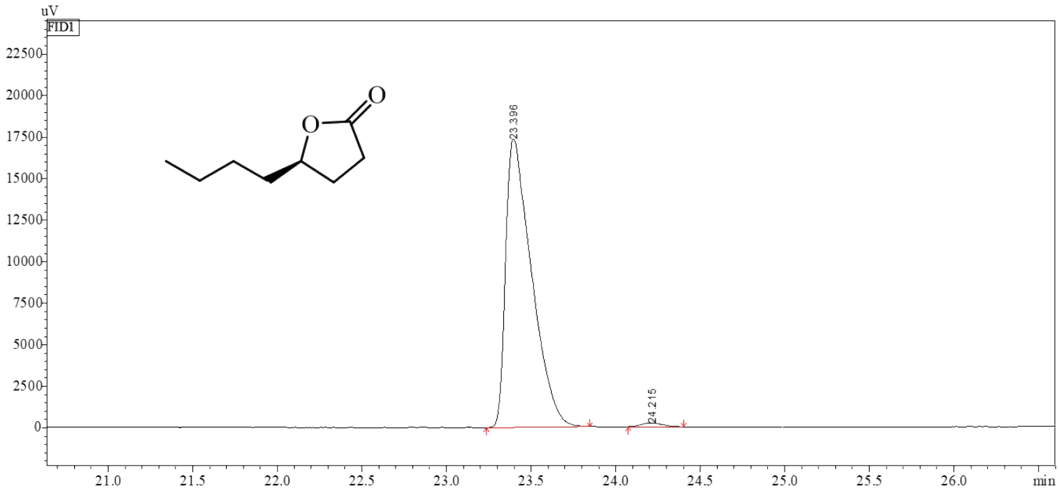
3. Chiral GC data of enantiopure lactones

Reduction of **1a** with $NaBH_4$



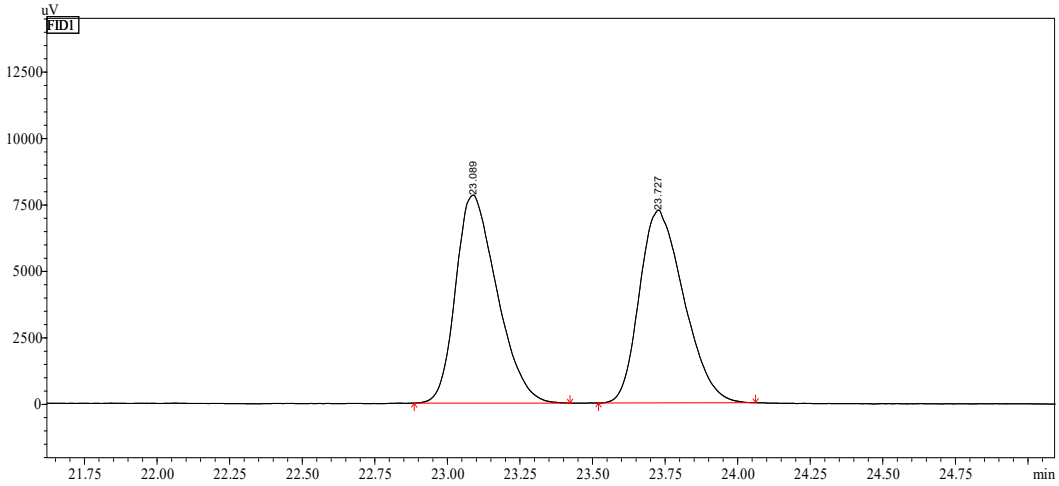
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	23.476	77493	50.0
2	24.131	77555	50.0

Bioreduction of **1a** with *Sm*CR_{M5}



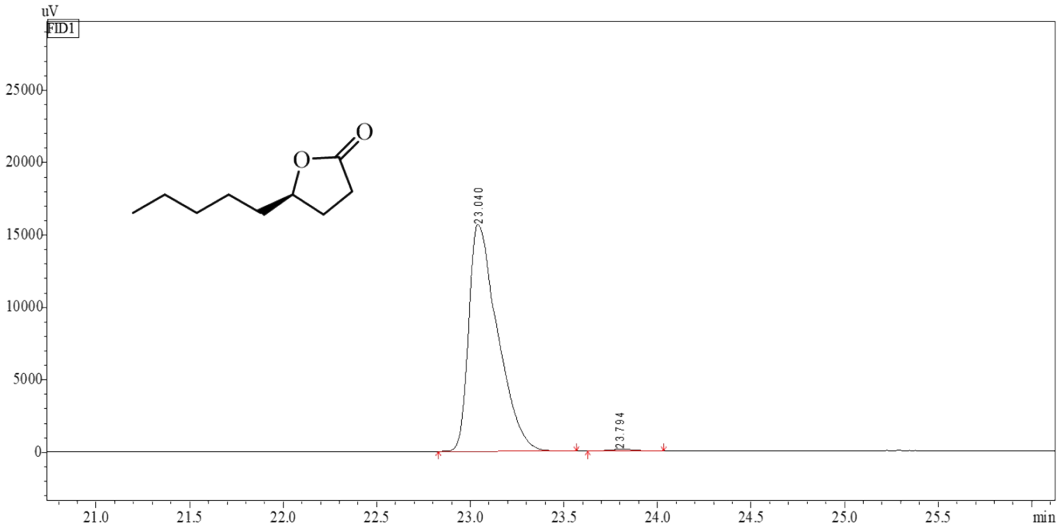
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	23.396	184995	98.9
2	24.215	2006	1.1

Reduction of **1b** with $NaBH_4$



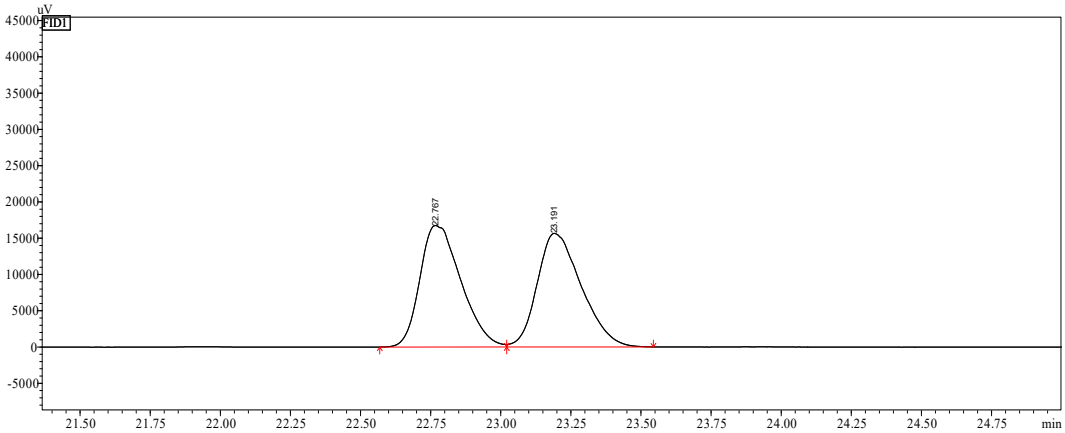
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	23.089	77523	50.1
2	23.727	77107	49.9

Bioreduction of **1b** with *Sm*CR_{M5}



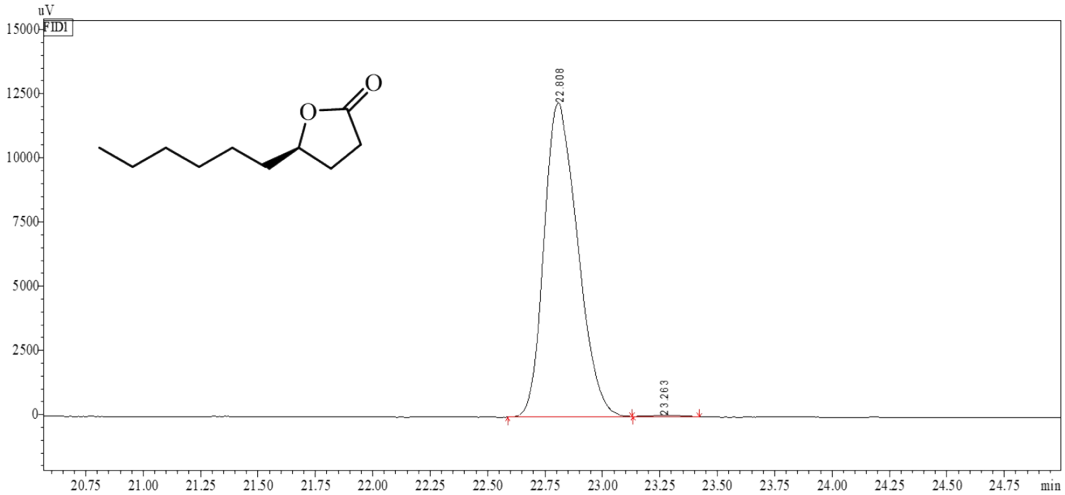
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	23.04	167850	99.3
2	23.794	1190	0.7

Racemic γ-decalactone



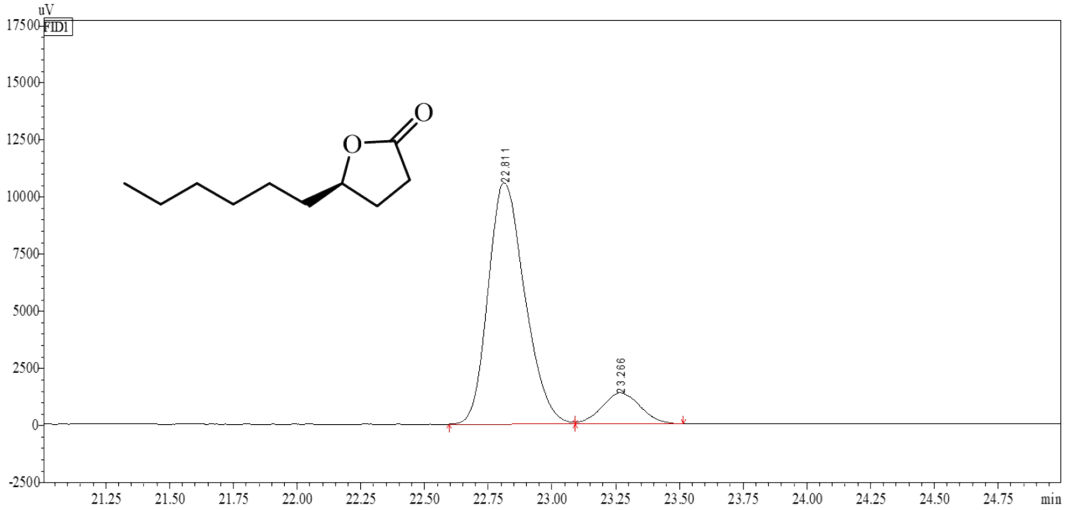
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	22.767	170564	49.9
2	23.191	170934	50.1

Bioreduction of **1c** with *Sm*CR$_{M5}$



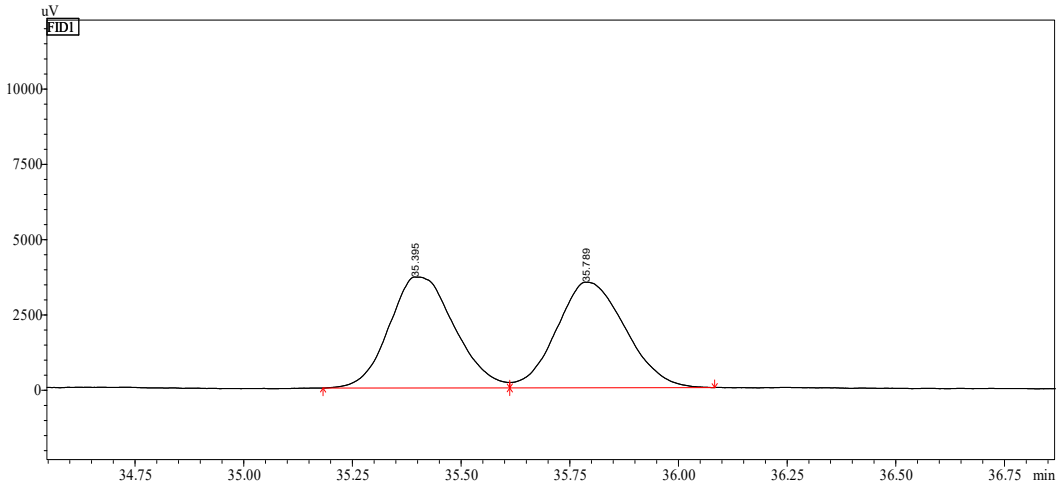
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	22.808	124332	99.5
2	23.263	620	0.5

Bioreduction of **1k** with *Sm*CR$_{M5}$



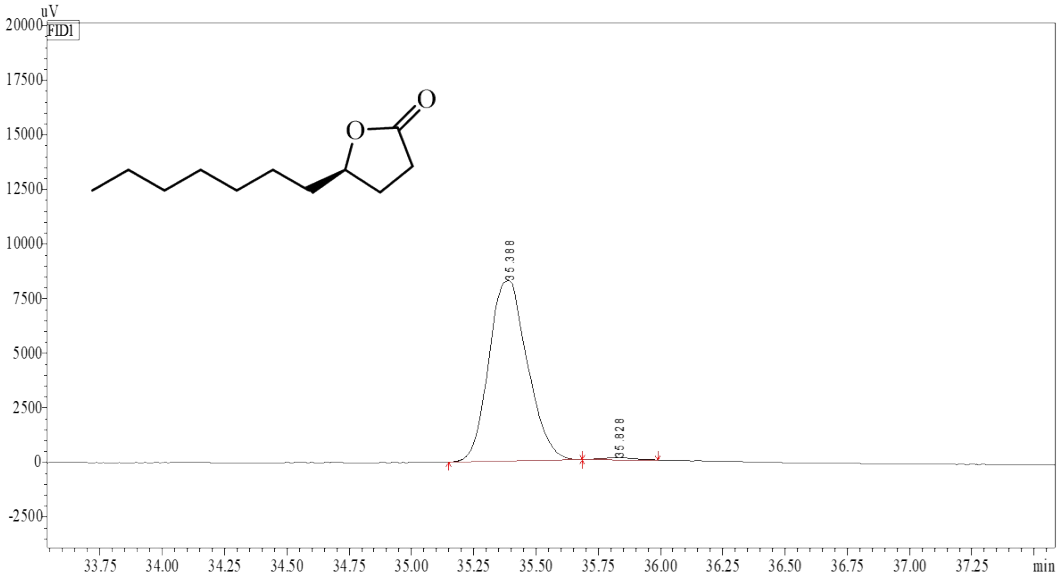
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	22.811	105674	88.2
2	23.266	14136	11.8

Reduction of **1d** with $NaBH_4$



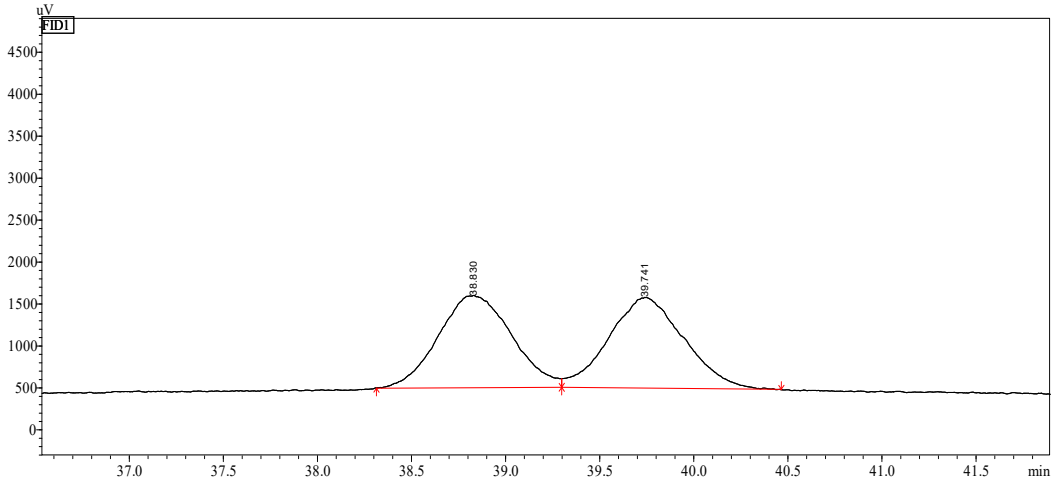
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	35.395	38606	50.0
2	35.789	38639	50.0

Bioreduction of **1d** with *Sm*CR_{M5}



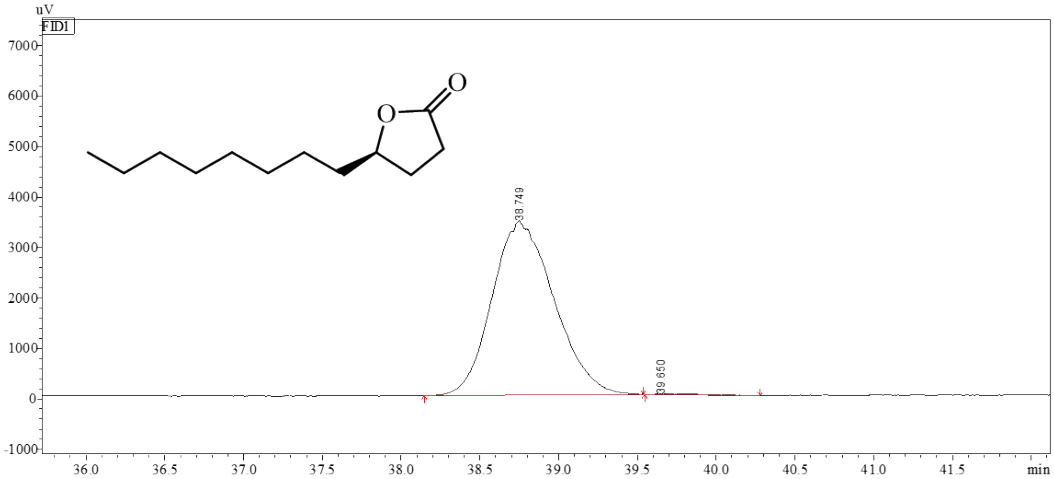
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	35.388	88746	98.9
2	35.828	1008	1.1

Reduction of **1e** with $NaBH_4$



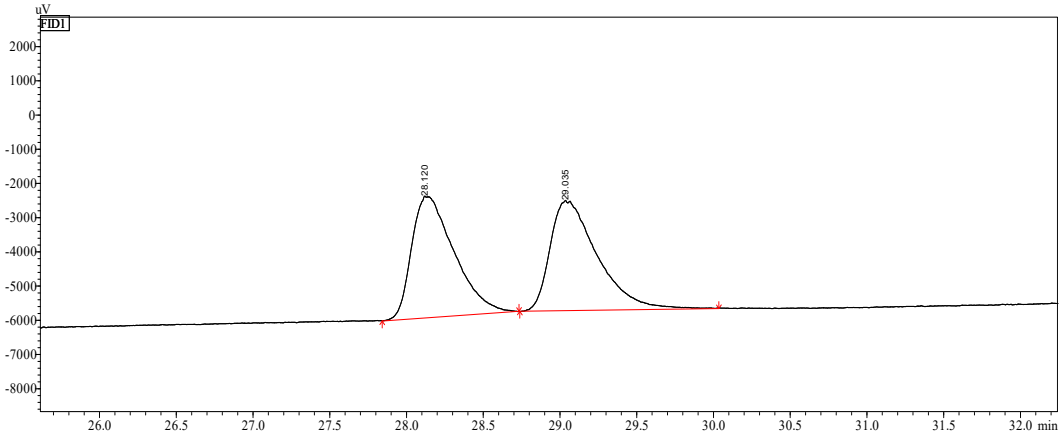
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	38.83	29925	50.0
2	39.741	29893	50.0

Bioreduction of **1e** with *Sm*CR_{M5}



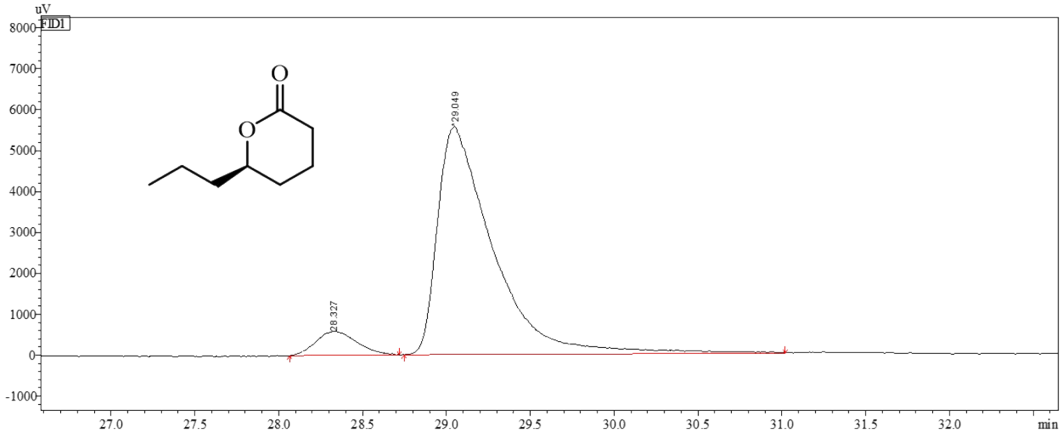
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	38.749	94080	99.5
2	39.65	441	0.5

Reduction of **1f** with $NaBH_4$



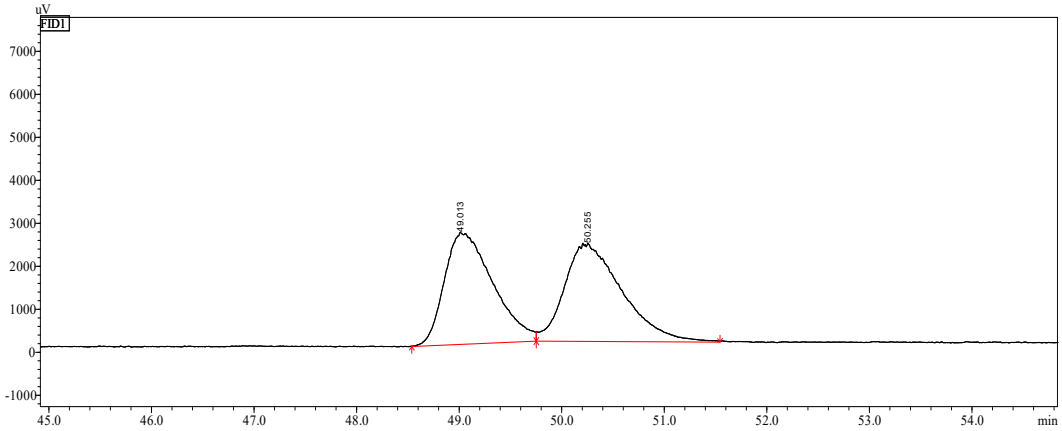
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	28.12	68215	49.9
2	29.035	68621	50.1

Bioreduction of **1f** with *Sm*CR_{M5}



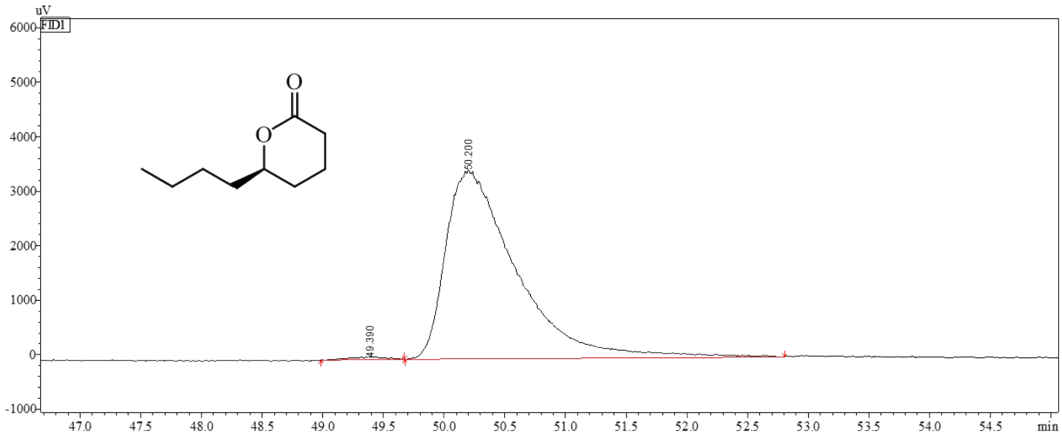
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	28.327	9952	7.2
2	29.049	128551	92.8

Reduction of **1g** with $NaBH_4$



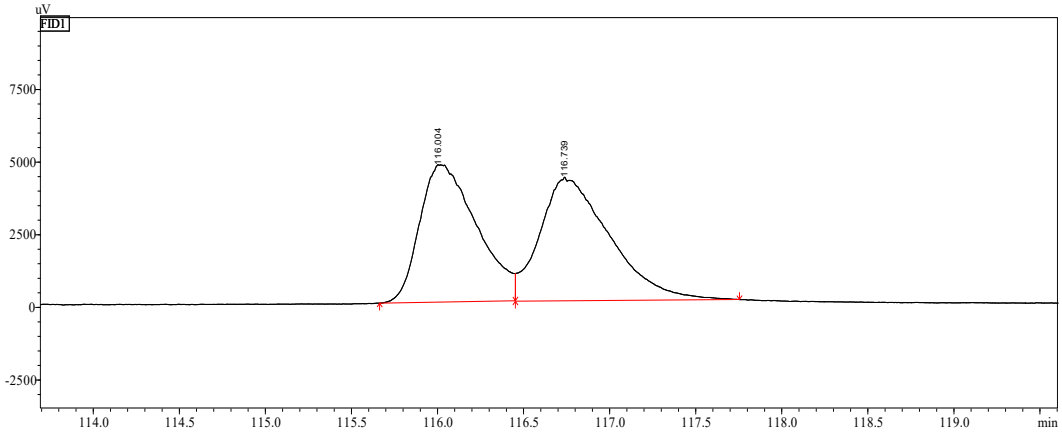
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	49.013	85306	49.0
2	50.255	88742	51.0

Bioreduction of **1g** with *Sm*CR_{M5}



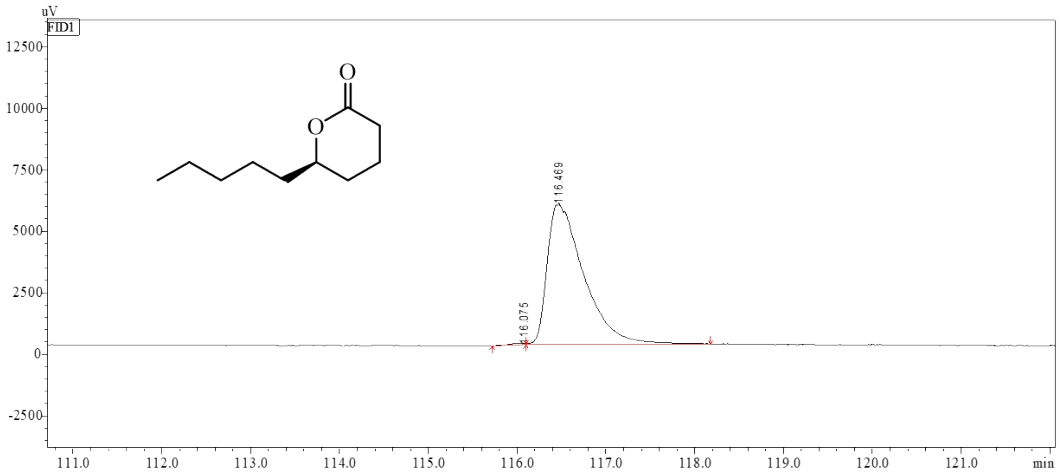
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	49.39	1095	0.8
2	50.2	139894	99.2

Racemic δ-decalactone



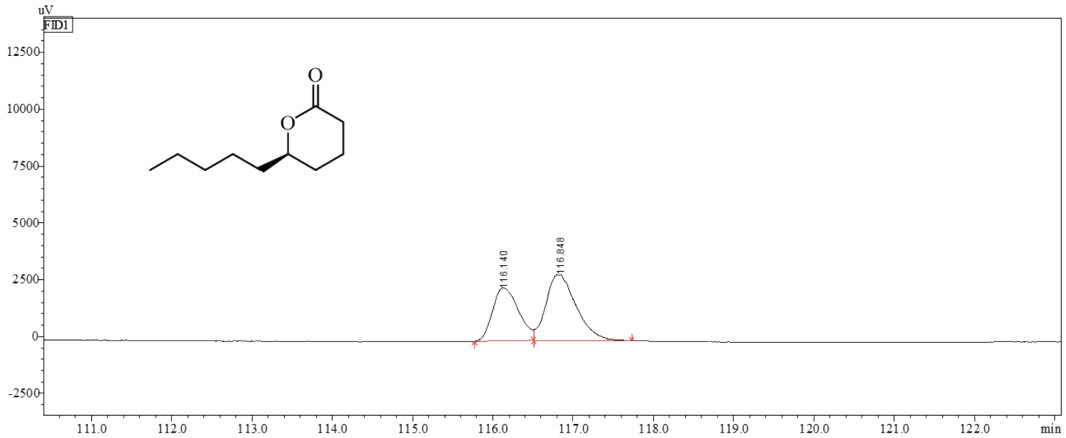
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	116.004	109281	47.8
2	116.739	119264	52.2

Bioreduction of **1h** with *Sm*CR$_{M5}$



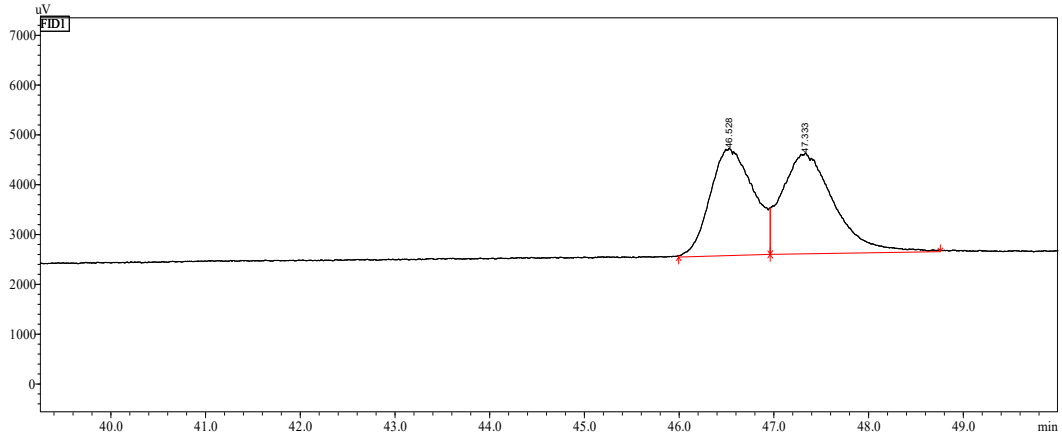
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	116.075	780	0.5
2	116.469	163222	99.5

Bioreduction of **1l** with *Sm*CR$_{M5}$



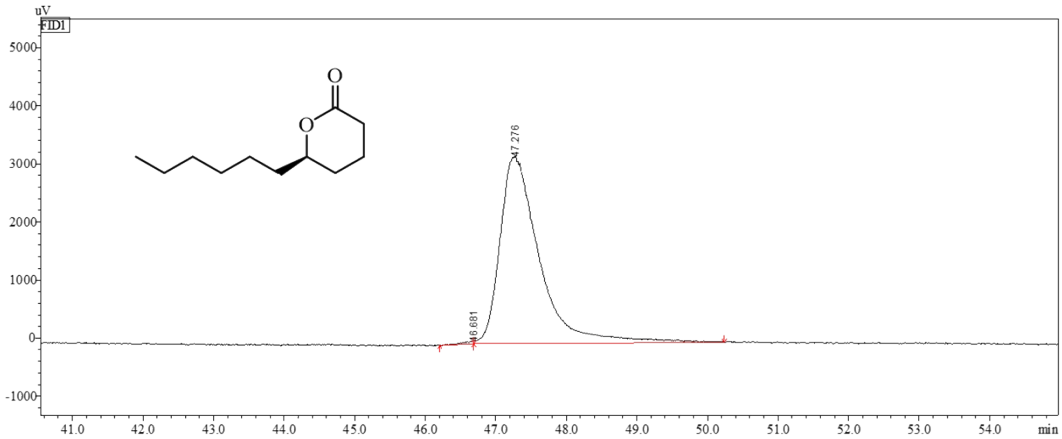
Peak #	Ret. Time (min)	Area (µV*s)	Area%
1	116.14	52492	40.7
2	116.848	76561	59.3

Reduction of **1i** with $NaBH_4$



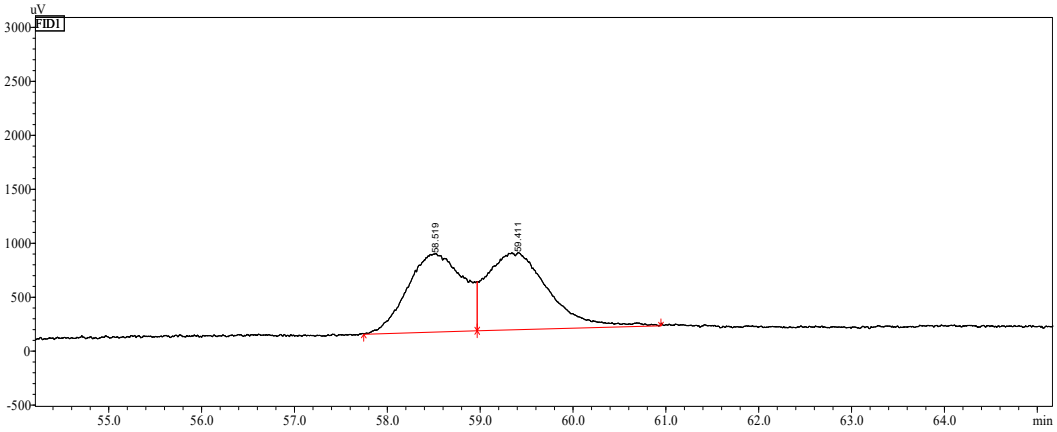
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	46.528	66871	46.7
2	47.333	76214	53.3

Bioreduction of **1i** with *Sm*CR_{M5}



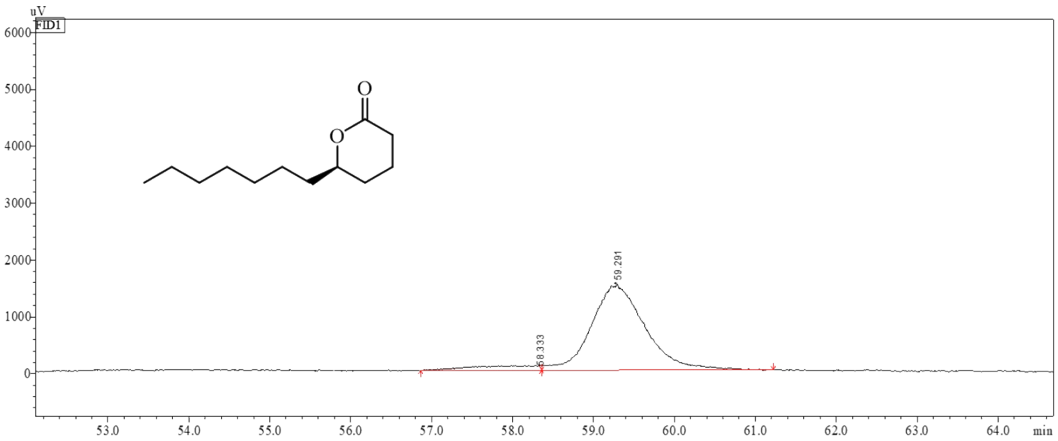
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	46.681	587	0.5
2	47.276	125220	99.5

Reduction of **1j** with $NaBH_4$



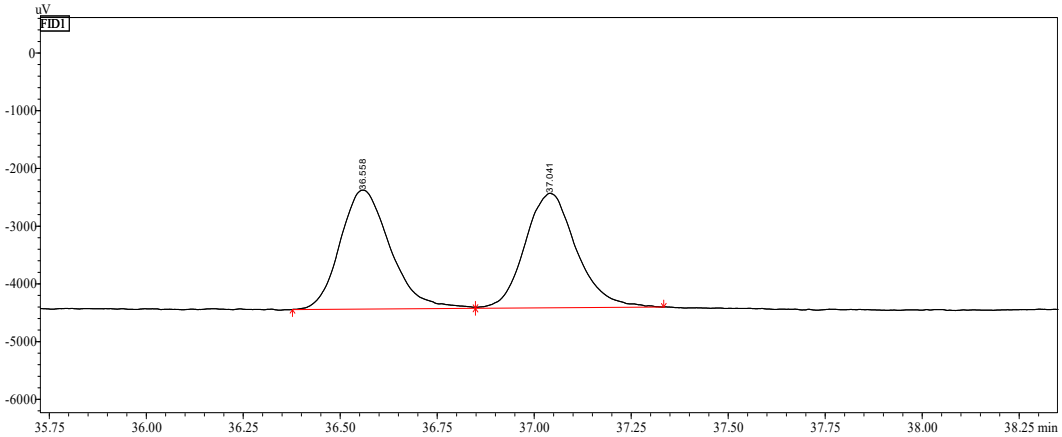
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	58.519	30538	48.1
2	59.411	32890	51.9

Bioreduction of **1j** with *Sm*CR_{M5}



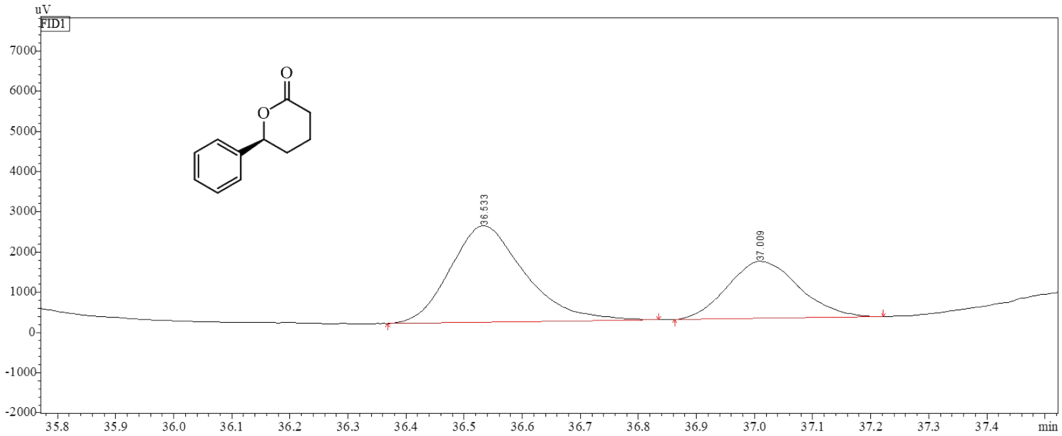
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	58.149	3784	5.2
2	59.291	69191	94.8

Reduction of **1m** with $NaBH_4$



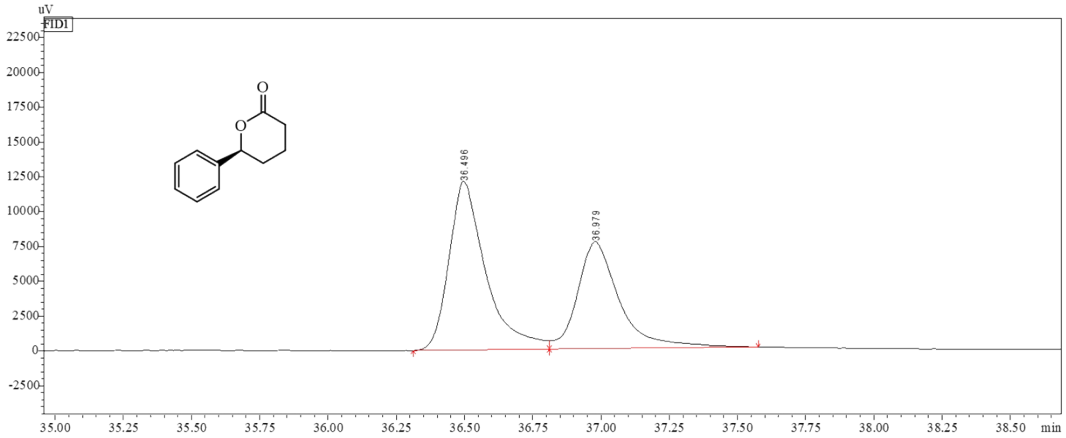
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	36.558	18498	50.4
2	37.041	18237	49.6

Bioreduction of **1m** with *Sm*CR_{M5}



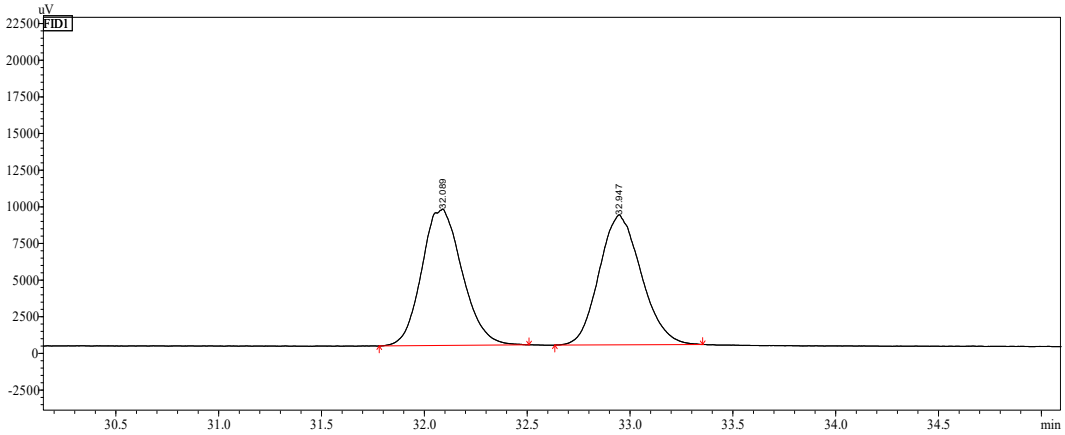
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	36.533	21416	63.9
2	37.009	12084	36.1

Bioreduction of **1n** with *Sm*CR$_{M5}$



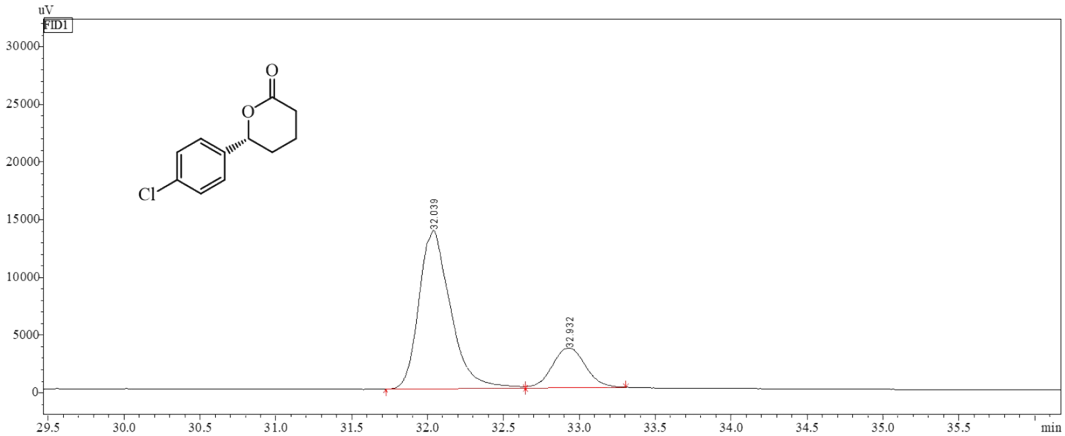
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	36.496	113572	58.1
2	36.979	82006	41.9

Reduction of **1o** with $NaBH_4$



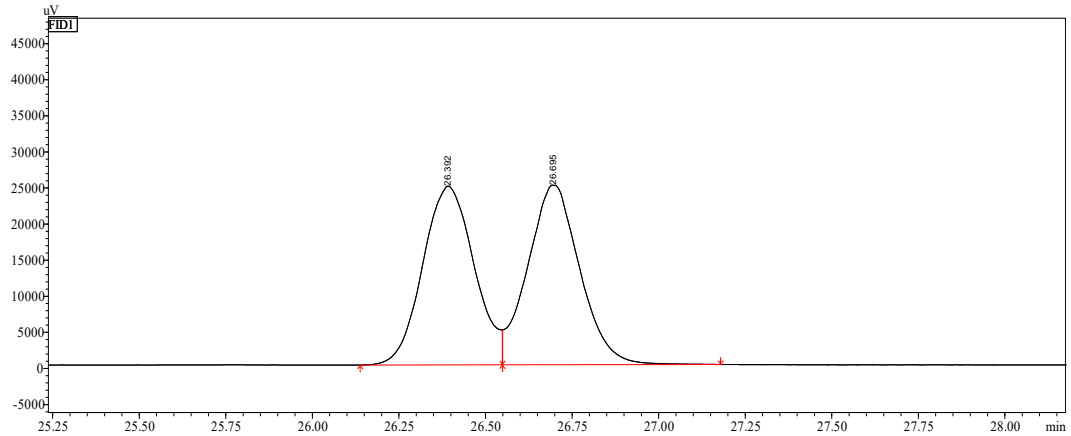
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	32.089	127355	50.2
2	32.947	126253	49.8

Bioreduction of **1o** with *Sm*CR_{M5}



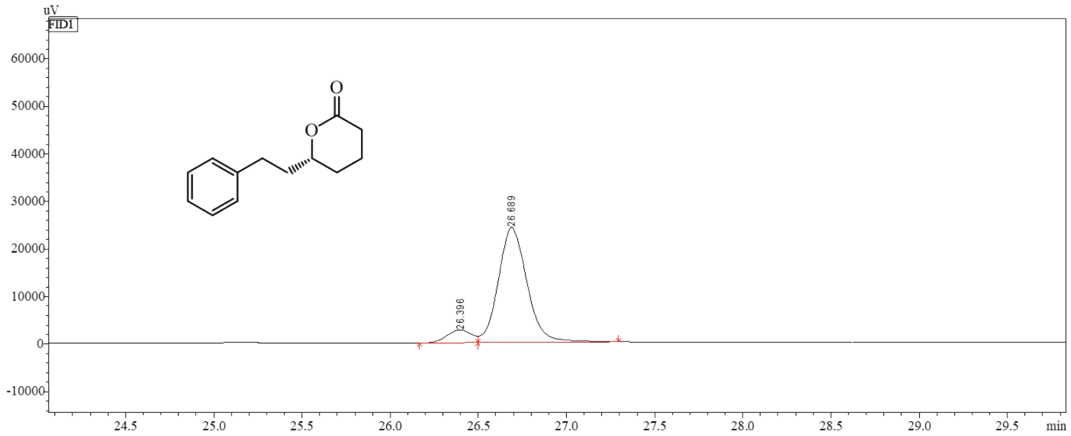
Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	32.039	198877	79.1
2	32.932	52569	20.9

Reduction of **1p** with $NaBH_4$



Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	26.392	255666	49.6
2	26.695	260165	50.4

Bioreduction of **1p** with *Sm*CR_{M5}



Peak #	Ret. Time (min)	Area (μV*s)	Area%
1	26.396	26040	8.7
2	26.689	272691	91.3

4. NMR spectras of enantiopure lactones

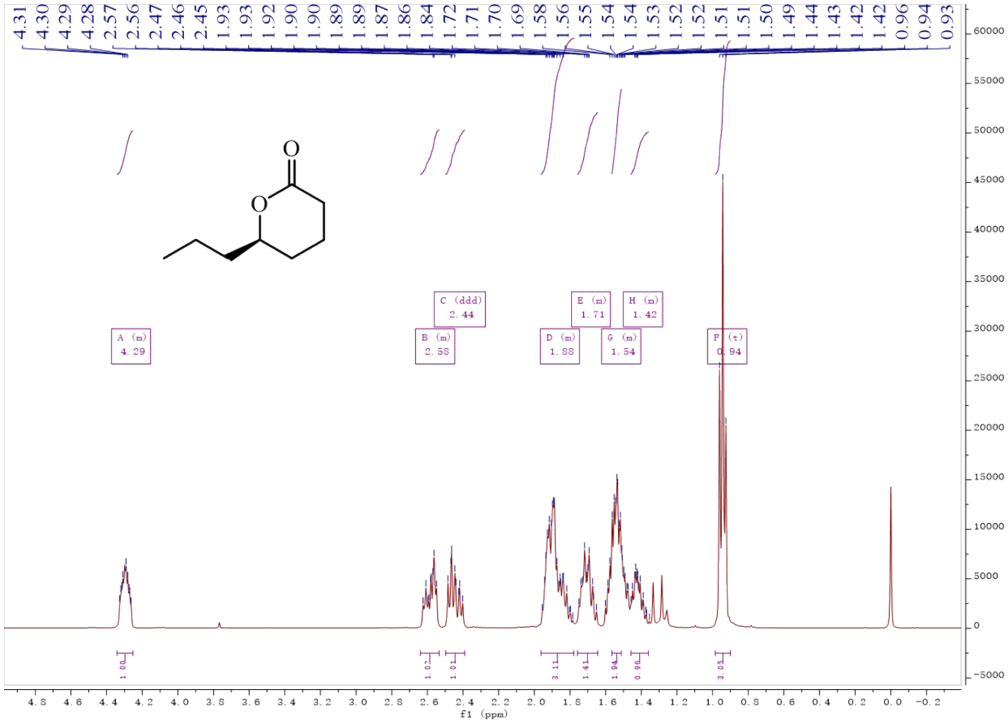


Figure S9. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectra of (*R*)-3f in $CDCl_3$.

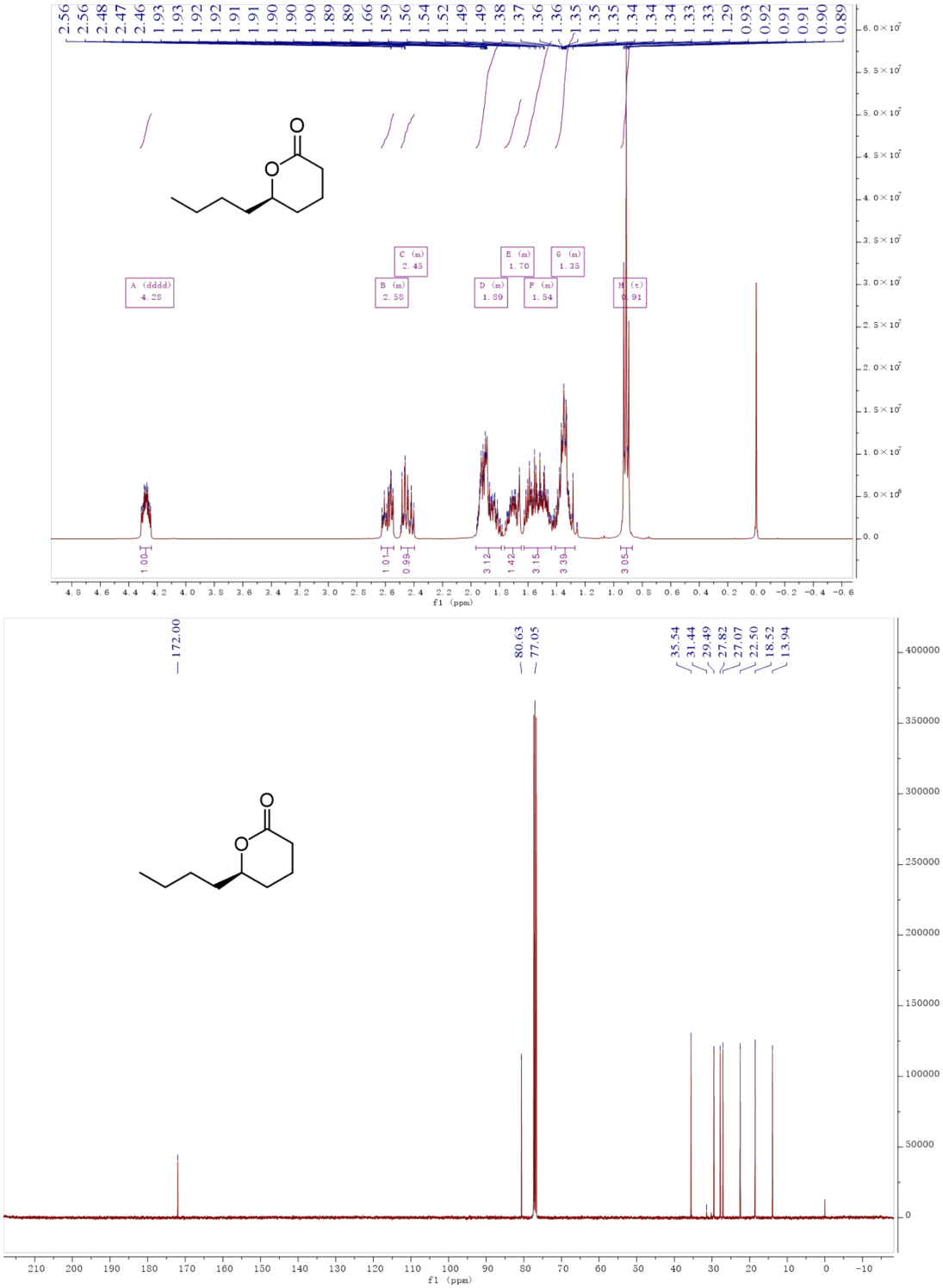
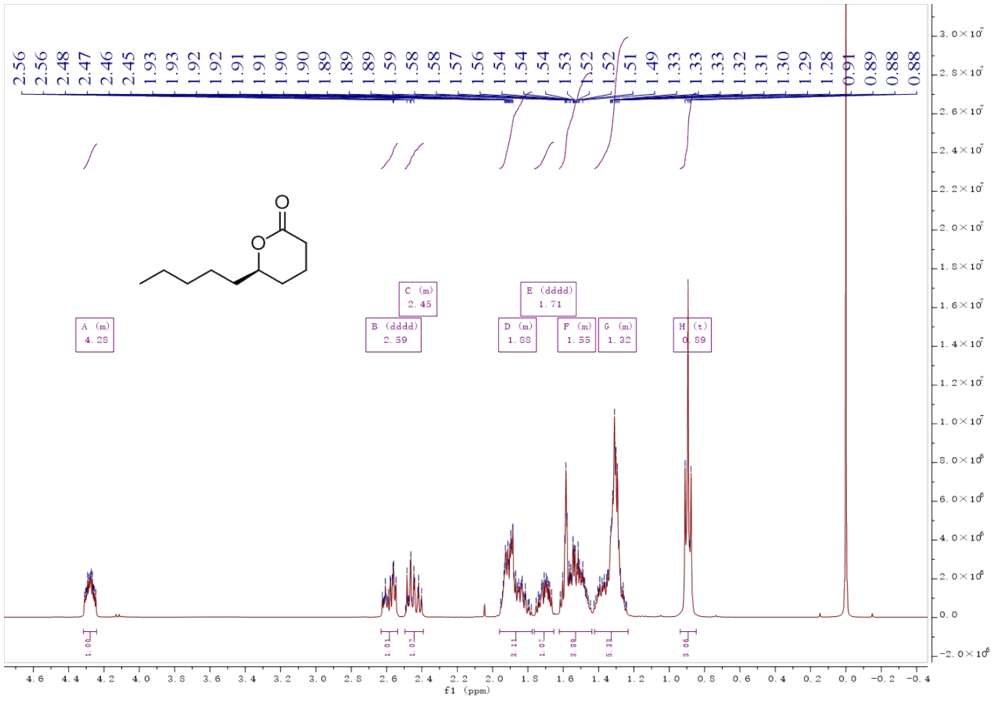


Figure S10. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectra of (*R*)-3g in $CDCl_3$.



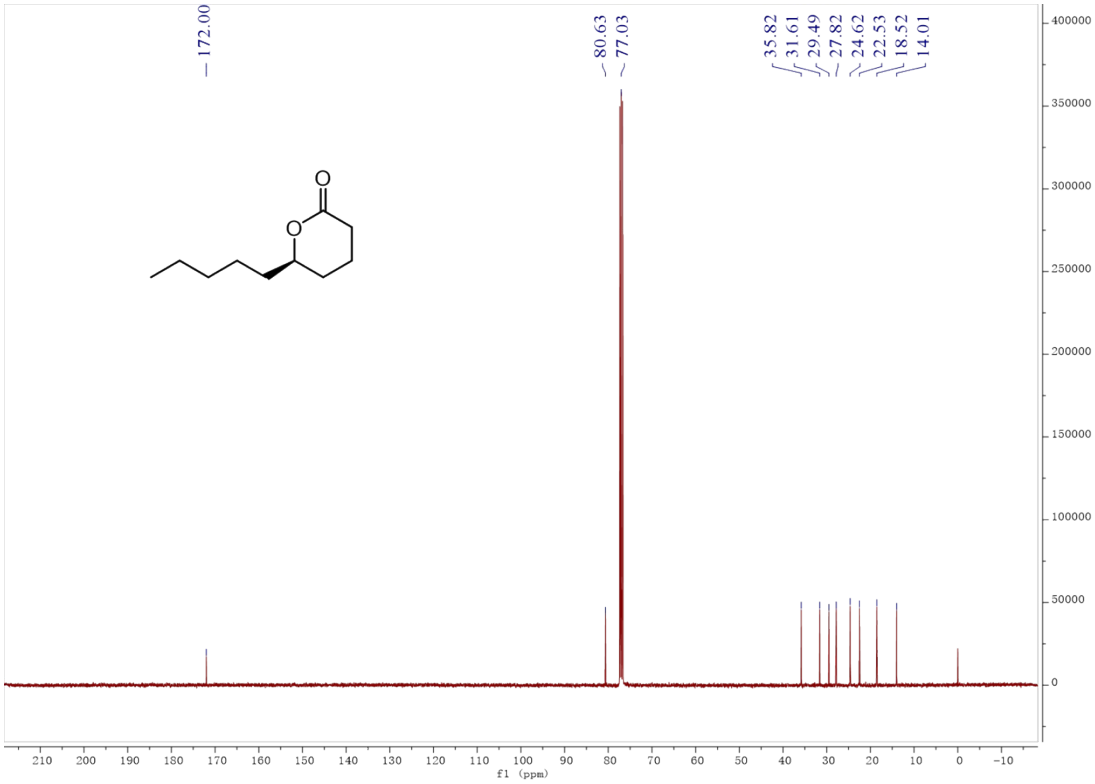


Figure S11. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectra of (*R*)-3h in $CDCl_3$.

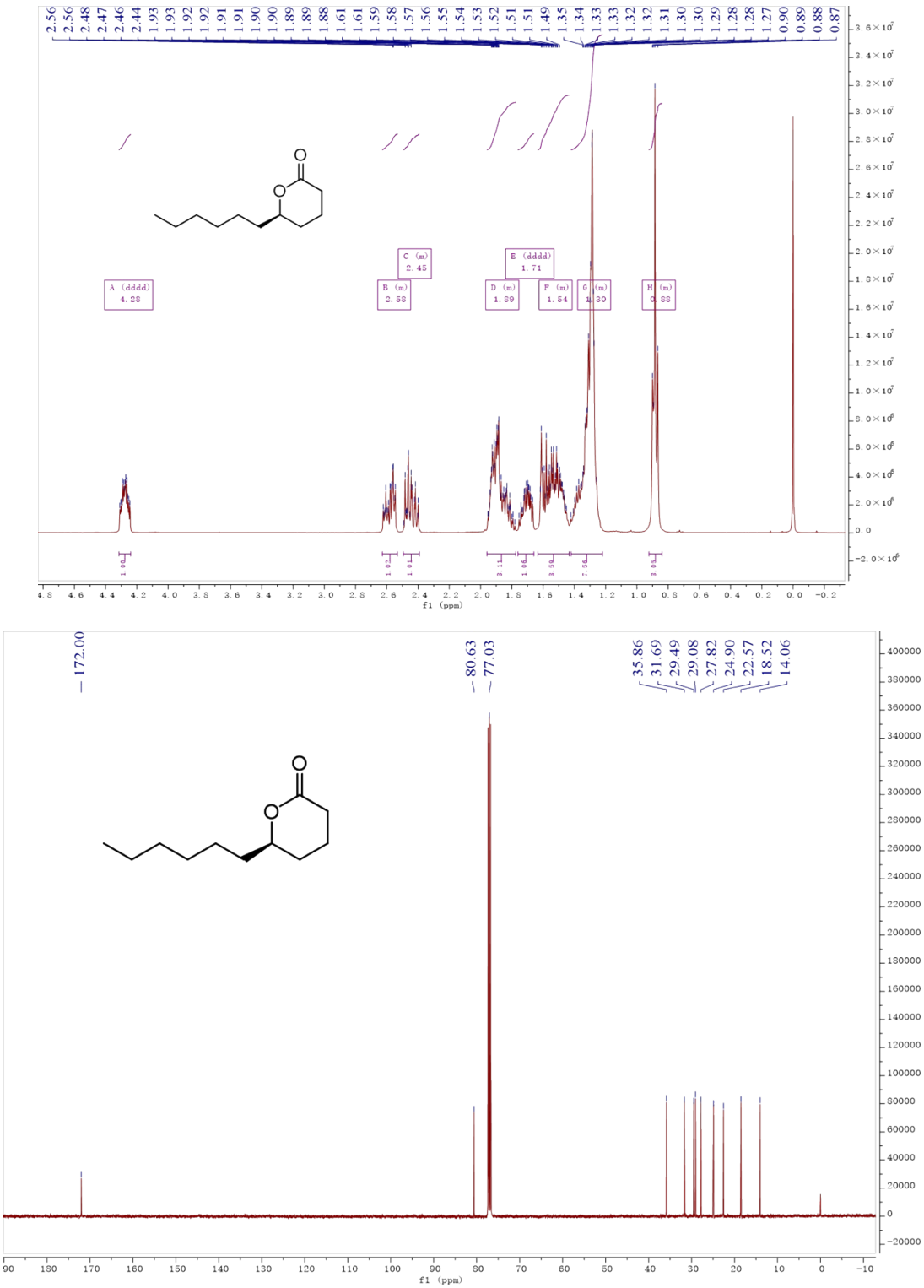


Figure S12. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectra of (*R*)-3i in $CDCl_3$.

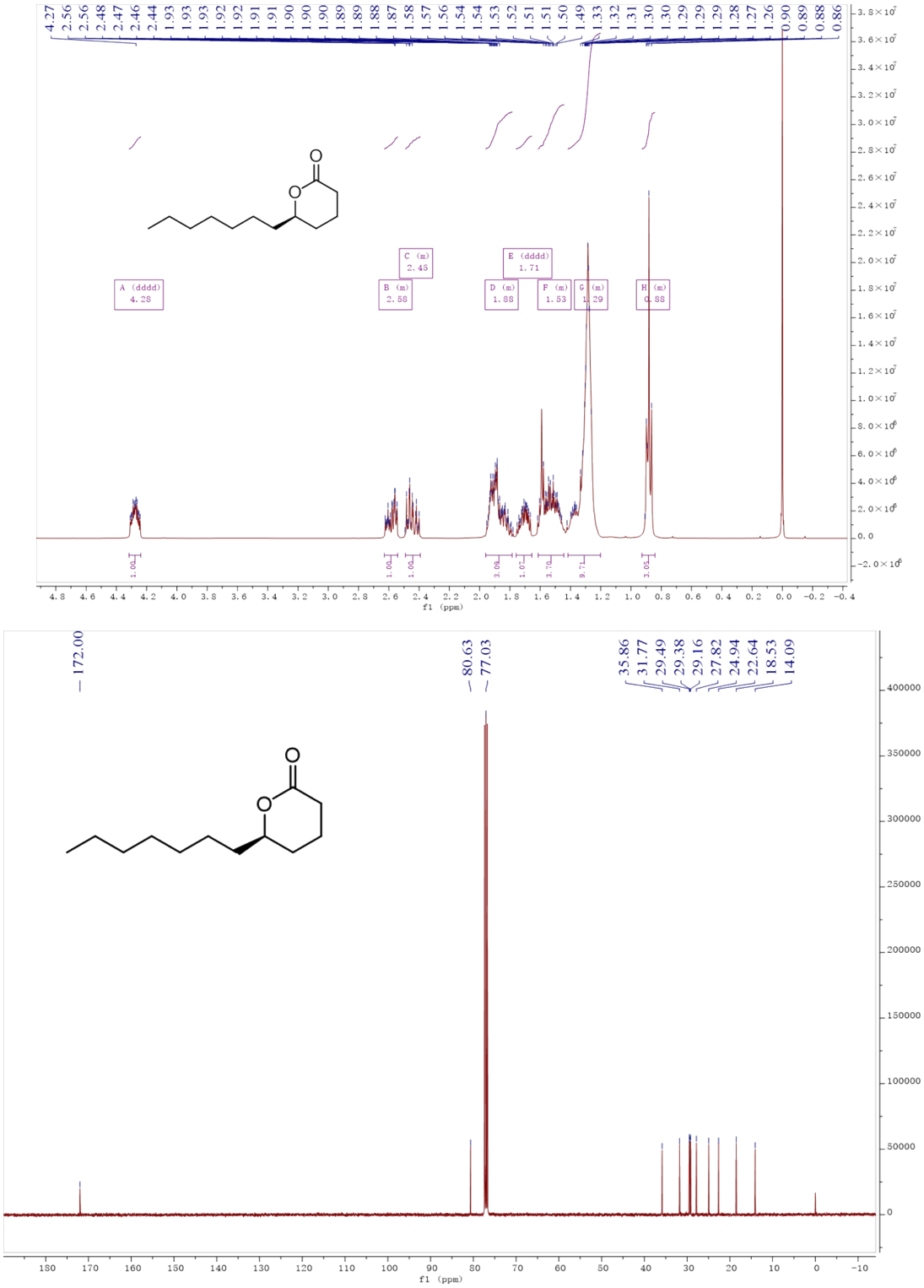


Figure S13. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectra of (*R*)-3j in $CDCl_3$.

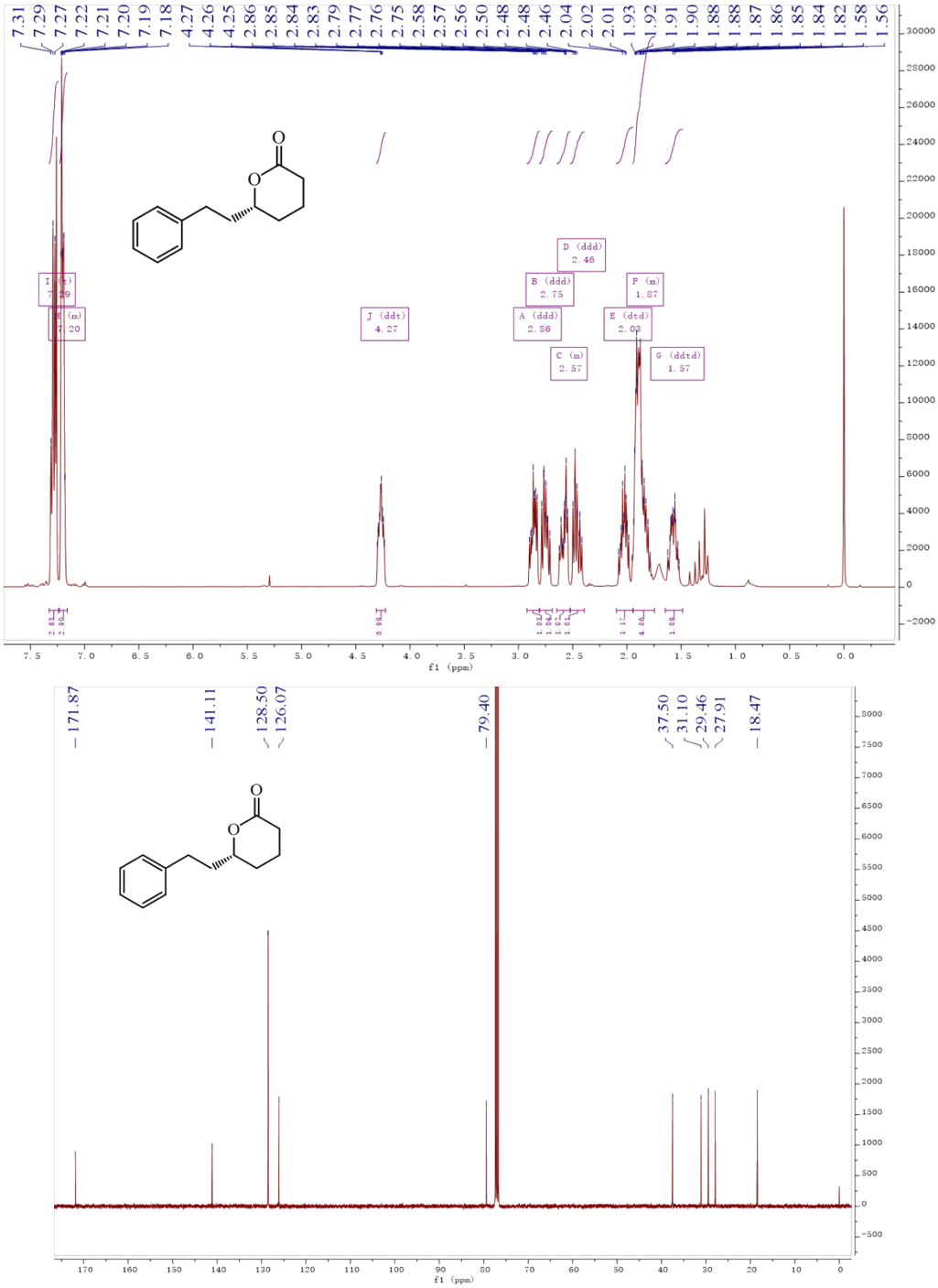


Figure S14. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectra of (*R*)-3m in $CDCl_3$.