

SEQUENCE LISTING

5 <110> Shionogi & Co.,Ltd.
 <120> 08P00078WO
 <130> A process for the preparation of adamantine hydroxide with
 cytochrome P450
 10 <160> 11
 <170> PatentIn version 3.3
 15 <210> 1
 <211> 396
 <212> PRT
 <213> Bacillus subtilis
 20 <220>
 <222> (2)..(2)
 <223> Xaa is Asp or Asn residue.
 <220>
 25 <222> (113)..(113)
 <223> Xaa is Arg or Gln residue.
 <400> 1
 30 Met Xaa Val Leu Asn Arg Arg Gln Ala Leu Gln Arg Ala Leu Leu Asn
 1 5 10 15
 Gly Lys Asn Lys Gln Asp Ala Tyr His Pro Phe Pro Trp Tyr Glu Ser
 35 20 25 30

Met Arg Lys Asp Ala Pro Val Ser Phe Asp Glu Glu Asn Gln Val Trp
 35 40 45
 5
 Ser Val Phe Leu Tyr Asp Asp Val Lys Lys Val Val Gly Asp Lys Glu
 50 55 60
 10
 Leu Phe Ser Ser Cys Met Pro Gln Gln Thr Ser Ser Ile Gly Asn Ser
 65 70 75 80
 15 Ile Ile Asn Met Asp Pro Pro Lys His Thr Lys Ile Arg Ser Val Val
 85 90 95
 Asn Lys Ala Phe Thr Pro Arg Val Met Lys Gln Trp Glu Pro Arg Ile
 20 100 105 110
 Xaa Glu Ile Thr Asp Glu Leu Ile Gln Lys Phe Gln Gly Arg Ser Glu
 115 120 125
 25
 Phe Asp Leu Val His Asp Phe Ser Tyr Pro Leu Pro Val Ile Val Ile
 130 135 140
 30
 Ser Glu Leu Leu Gly Val Pro Ser Ala His Met Glu Gln Phe Lys Ala
 145 150 155 160
 35 Trp Ser Asp Leu Leu Val Ser Thr Pro Lys Asp Lys Ser Glu Glu Ala

	165	170	175
	Glu Lys Ala Phe Leu Glu Glu Arg Asp Lys Cys Glu Glu Glu Leu Ala		
5	180	185	190
	Ala Phe Phe Ala Gly Ile Ile Glu Glu Lys Arg Asn Lys Pro Glu Gln		
	195	200	205
10			
	Asp Ile Ile Ser Ile Leu Val Glu Ala Glu Glu Thr Gly Glu Lys Leu		
	210	215	220
15			
	Ser Gly Glu Glu Leu Ile Pro Phe Cys Thr Leu Leu Leu Val Ala Gly		
	225	230	235 240
20			
	Asn Glu Thr Thr Thr Asn Leu Ile Ser Asn Ala Met Tyr Ser Ile Leu		
	245	250	255
	Glu Thr Pro Gly Val Tyr Glu Glu Leu Arg Ser His Pro Glu Leu Met		
25	260	265	270
	Pro Gln Ala Val Glu Glu Ala Leu Arg Phe Arg Ala Pro Ala Pro Val		
	275	280	285
30			
	Leu Arg Arg Ile Ala Lys Arg Asp Thr Glu Ile Gly Gly His Leu Ile		
	290	295	300
35			



ggggataaag agttgttttc cagttgcatg cgcagcaga caagctctat tggaaattcc 240

atcattaaca tggacccgcc gaagcataca aaaatecggt cagtcgtgaa caaagccttt 300

5 actccgcgcg tgatgaagca atgggaaccg agaattcgag aaatcacaga tgaactgatt 360

caaaaatttc aggggcgcag tgagtttgac cttgttcacg atttttcata cccgcttcgg 420

10 gttattgtga tatctgagtt gctgggagtg ccttcagcgc atatggaaca gtttaaagca 480

tggtctgate ttctggtcag tacaccgaag gataaaagtg aagaagctga aaaagccttt 540

ttggaagaac gagataagtg tgaggaagaa ctggccgcgt tttttgccgg catcatagaa 600

15 gaaaagecga acaaaccgga acaggatatt atttctattt tagtggaagc ggaagaaaca 660

ggcgagaagc tgtecggtga agagctgatt ccgttttgca cgctgctgct ggtggccgga 720

20 aatgaaacca ctacaaacct gatttcaa at gcatgtaca gcatattaga aacgccaggc 780

gtttacgagg aactgcgcag ccctcctgaa ctgatgcctc aggcagtgga ggaagccttg 840

cgtttcagag cgccggcccc ggttttgagg cgcattgcca agcgggatac ggagatcggg 900

25 gggcacctga ttaaagaagg tgatatggtt ttggcggttg tggcatcggc aaatcgtgat 960

gaagcaaagt ttgacagacc gcacatgttt gatatecgcc gccatcccaa tccgcatatt 1020

30 gcgtttggcc acggcatcca tttttgctt ggggccccgc ttgcccgctct tgaagcaa at 1080

atcgcgtaa cgtctttgat ttctgctttt cctcatatgg agtgcgtcag tatcactccg 1140

attgaaaaca gtgtgatata cggattaaag agcttcggtg tgaaaatgta a 1191

35

<210> 3

<211> 727

<212> PRT

5 <213> Artificial sequence

<220>

<223> A protein containing fused CYP109B1 and and P450Rhfr reductase
domain

10

<400> 3

Met Asp Val Leu Asn Arg Arg Gln Ala Leu Gln Arg Ala Leu Leu Asn

1 5 10 15

15

Gly Lys Asn Lys Gln Asp Ala Tyr His Pro Phe Pro Trp Tyr Glu Ser

20 25 30

20

Met Arg Lys Asp Ala Pro Val Ser Phe Asp Glu Glu Asn Gln Val Trp

35 40 45

25 Ser Val Phe Leu Tyr Asp Asp Val Lys Lys Val Val Gly Asp Lys Glu

50 55 60

Leu Phe Ser Ser Cys Met Pro Gln Gln Thr Ser Ser Ile Gly Asn Ser

30 65 70 75 80

Ile Ile Asn Met Asp Pro Pro Lys His Thr Lys Ile Arg Ser Val Val

85 90 95

35

	Asn Lys Ala Phe Thr Pro Arg Val Met Lys Gln Trp Glu Pro Arg Ile		
	100	105	110
5	Arg Glu Ile Thr Asp Glu Leu Ile Gln Lys Phe Gln Gly Arg Ser Glu		
	115	120	125
10	Phe Asp Leu Val His Asp Phe Ser Tyr Pro Leu Pro Val Ile Val Ile		
	130	135	140
	Ser Glu Leu Leu Gly Val Pro Ser Ala His Met Glu Gln Phe Lys Ala		
15	145	150	155 160
	Trp Ser Asp Leu Leu Val Ser Thr Pro Lys Asp Lys Ser Glu Glu Ala		
	165	170	175
20	Glu Lys Ala Phe Leu Glu Glu Arg Asp Lys Cys Glu Glu Glu Leu Ala		
	180	185	190
25	Ala Phe Phe Ala Gly Ile Ile Glu Glu Lys Arg Asn Lys Pro Glu Gln		
	195	200	205
30	Asp Ile Ile Ser Ile Leu Val Glu Ala Glu Glu Thr Gly Glu Lys Leu		
	210	215	220
	Ser Gly Glu Glu Leu Ile Pro Phe Cys Thr Leu Leu Leu Val Ala Gly		
35	225	230	235 240

Asn Glu Thr Thr Thr Asn Leu Ile Ser Asn Ala Met Tyr Ser Ile Leu

245

250

255

5

Glu Thr Pro Gly Val Tyr Glu Glu Leu Arg Ser His Pro Glu Leu Met

260

265

270

10

Pro Gln Ala Val Glu Glu Ala Leu Arg Phe Arg Ala Pro Ala Pro Val

275

280

285

15 Leu Arg Arg Ile Ala Lys Arg Asp Thr Glu Ile Gly Gly His Leu Ile

290

295

300

Lys Glu Gly Asp Met Val Leu Ala Phe Val Ala Ser Ala Asn Arg Asp

20 305

310

315

320

Glu Ala Lys Phe Asp Arg Pro His Met Phe Asp Ile Arg Arg His Pro

325

330

335

25

Asn Pro His Ile Ala Phe Gly His Gly Ile His Phe Cys Leu Gly Ala

340

345

350

30

Pro Leu Ala Arg Leu Glu Ala Asn Ile Ala Leu Thr Ser Leu Ile Ser

355

360

365

35 Ala Phe Pro His Met Glu Cys Val Ser Ile Thr Pro Ile Glu Asn Ser

370

375

380

Val Ile Tyr Gly Leu Lys Ser Phe Arg Val Lys Met Glu Phe Val Leu

5 385 390 395 400

His Arg His Gln Pro Val Thr Ile Gly Glu Pro Ala Ala Arg Ala Val

405 410 415

10

Ser Arg Thr Val Thr Val Glu Arg Leu Asp Arg Ile Ala Asp Asp Val

420 425 430

15

Leu Arg Leu Val Leu Arg Asp Ala Gly Gly Lys Thr Leu Pro Thr Trp

435 440 445

20 Thr Pro Gly Ala His Ile Asp Leu Asp Leu Gly Ala Leu Ser Arg Gln

450 455 460

Tyr Ser Leu Cys Gly Ala Pro Asp Ala Pro Ser Tyr Glu Ile Ala Val

25 465 470 475 480

His Leu Asp Pro Glu Ser Arg Gly Gly Ser Arg Tyr Ile His Glu Gln

485 490 495

30

Leu Glu Val Gly Ser Pro Leu Arg Met Arg Gly Pro Arg Asn His Phe

500 505 510

35

Ala Leu Asp Pro Gly Ala Glu His Tyr Val Phe Val Ala Gly Gly Ile
515 520 525

5 Gly Ile Thr Pro Val Leu Ala Met Ala Asp His Ala Arg Ala Arg Gly
530 535 540

10 Trp Ser Tyr Glu Leu His Tyr Cys Gly Arg Asn Arg Ser Gly Met Ala
545 550 555 560

15 Tyr Leu Glu Arg Val Ala Gly His Gly Asp Arg Ala Ala Leu His Val
565 570 575

Ser Glu Glu Gly Thr Arg Ile Asp Leu Ala Ala Leu Leu Ala Glu Pro
580 585 590

20 Ala Pro Gly Val Gln Ile Tyr Ala Cys Gly Pro Gly Arg Leu Leu Ala
595 600 605

25 Gly Leu Glu Asp Ala Ser Arg Asn Trp Pro Asp Gly Ala Leu His Val
610 615 620

30 Glu His Phe Thr Ser Ser Leu Ala Ala Leu Asp Pro Asp Val Glu His
625 630 635 640

35 Ala Phe Asp Leu Glu Leu Arg Asp Ser Gly Leu Thr Val Arg Val Glu
645 650 655

Pro Thr Gln Thr Val Leu Asp Ala Leu Arg Ala Asn Asn Ile Asp Val

660

665

670

5

Pro Ser Asp Cys Glu Glu Gly Leu Cys Gly Ser Cys Glu Val Ala Val

675

680

685

10 Leu Asp Gly Glu Val Asp His Arg Asp Thr Val Leu Thr Lys Ala Glu

690

695

700

Arg Ala Ala Asn Arg Gln Met Met Thr Cys Cys Ser Arg Ala Cys Gly

15 705

710

715

720

Asp Arg Leu Ala Leu Arg Leu

725

20

<210> 4

<211> 2184

<212> DNA

25 <213> Artificial sequence

<220>

<223> A gene coding a protein containing fused CYP109B1 and P450Rhfr
reductase domain

30

<400> 4

atggatgtgt taaaccgccg gcaagccttg cagcgagcgc tgctcaatgg gaaaaacaaa 60

caggatgcgt atcatccgtt tccatggtat gaatcgatga gaaaggatgc gcctgtttcc 120

35

tttgatgaag aaaaccaagt gtggagcggt tttctttatg atgatgtcaa aaaagttggt 180
 ggggataaag agttgttttc cagttgcatg cgcagcaga caagctctat tggaaattcc 240
 5 atcattaaca tggaccgcc gaagcataca aaaatccgtt cagtcgtgaa caaagccttt 300
 actccgcgcg tgatgaagca atgggaaccg agaattcgag aaatcacaga tgaactgatt 360
 caaaaatttc aggggcgcag tgagtttgac cttgttcag atttttcata cccgcttcg 420
 10 gttattgtga tatctgagtt gctgggagtg cttcagcgc atatggaaca gtttaaagca 480
 tggctcgatc ttctggtcag tacaccgaag gataaaagtg aagaagctga aaaagccttt 540
 15 ttggaagaac gagataagtg tgaggaagaa ctggccgcgt tttttgccg catcatagaa 600
 gaaaagcgaa acaaaccgga acaggatatt atttctatct tagtggaagc ggaagaaaca 660
 ggcgagaagc tgccgggtga agagctgatt ccgttttgca cgctgctgct ggtggccgga 720
 20 aatgaaacca ctacaaacct gatttcaaat gcgatgtaca gcatattaga aacgccaggc 780
 gtttacgagg aactgcgcag ccactctgaa ctgatgcctc aggcagtgga ggaagccttg 840
 25 cgtttcagag cgccggcccc ggttttgagg cgcattgcca agcgggatac ggagatcggg 900
 gggcacctga ttaaagaagg tgatatggtt ttggcgtttg tggcatcggc aaatcgtgat 960
 gaagcaaagt ttgacagacc gcacatgttt gatatccgcc gccatccaa tccgcatatt 1020
 30 gcgtttggcc acggcatcca tttttgcctt ggggccccgc ttgcccgctc tgaagcaaat 1080
 atcgcgtaa cgtctttgat ttctgctttt cctcatatgg agtgcgtcag tatcactccg 1140
 35 attgaaaaca gtgtgatata cggattaaag agcttccgtg tgaaaatgga attcgtgctg 1200

caccgccatc aaccggtcac catcgagaaa cccgcccgcg gggegggtgc cgcaccgtc 1260
accgtcgagc gcctggaccg gatcgccgac gacgtgctgc gcctcgctct gcgcgacgcc 1320
5 ggcggaaga cattgcccac gtggactccc gggecccata tcgacctga cctcggcgcg 1380
ctgtcgcgcc agtactccct gtgcggcgcg cccgatgcgc cgagctacga gattgccgtg 1440
10 cactggatc ccgagagccg cggcggttcg cgctacatcc acgaacagct cgaggtggga 1500
agcccgtcc ggatgcgcgg cctcggaac catttcgcgc tcgaccccg gcgcgagcac 1560
tacgtgttcg tcgcggcgcg catcgccatc accccagtc tggccatggc cgaccacgcc 1620
15 cggecccggg ggtggagcta cgaactgcac tactgcggcc gaaaccgttc cggcatggcc 1680
tatctcgagc gtgtcgccgg gcacgggtgac cgggcgcgcg tgcacgtgtc cgaggaagge 1740
20 acccgatcg acctcgccgc cctcctcgcc gagcccgccc ccggcgccca gatctacgcg 1800
tgccggcccg ggcggtgct cgcgggactc gaggacgcga gccggaactg gcccgacggg 1860
gcgtgcacg tcgagcactt cactcgctc ctcgcggcg tcgacccgga cgtcgagcac 1920
25 gcctcgacc tcgaactgcg tgactcgggg ctgacctgc gggctgaacc caccagacc 1980
gtcctcgacg cgttcgcgc caacaacatc gacgtgccca gcgactgga ggaaggctc 2040
30 tgcggctcgt gcgaggtgc cgtcctcgac ggcgaggtcg accatcgga caccgtgctg 2100
accaaggccg agcggggcg gaaccggcag atgatgacct gctgctcgcg tgcctgtggc 2160
gaccggctgg cctcgcgact ctga 2184
35

<210> 5
 <211> 1191
 <212> DNA
 5 <213> Artificial sequence

 <220>
 <223> A gene coding a protein CYP109B1 with I77F mutation

 10 <400> 5
 atgggatgtgt taaaccgcgc gcaagccttg cagcgagcgc tgctcaatgg gaaaaacaaa 60

 caggatgcgt atcatccgtt tccatgggtat gaatcgatga gaaaggatgc gcctgtttcc 120

 15 tttgatgaag aaaaccaagt gtggagcgtt tttctttatg atgatgtcaa aaaagttggt 180

 ggggataaag agttgttttc cagttgcatg cgcgagcaga caagctcttt tggaaattcc 240

 atcattaaca tggaccegcc gaagcataca aaaatccgtt cagtcgtgaa caaagccttt 300
 20 actccgcgcg tgatgaagca atgggaaccg agaattcgag aaatcacaga tgaactgatt 360

 caaaaatttc aggggcgcag tgagtttgac ctgtttcag attttcata cccgtttccg 420

 25 gttattgtga tatctgagtt gctgggagtg cttcagcgc atatggaaca gtttaagca 480

 tggctctgat ttctggtcag tacaccgaag gataaaagtg aagaagctga aaaagccttt 540

 ttggaagaac gagataagtg tgaggaagaa ctggccgcgt ttttgccgg catcatagaa 600
 30 gaaaagcgaa acaaaccgga acaggatatt atttctatgt tagtggaagc ggaagaaaca 660

 ggcgagaagc tgtccggtga agagctgatt ccgttttgca cgctgctgct ggtggccgga 720

 35 aatgaaacca ctacaaacct gatttcaa at gcatgttaca gcatattaga aacgccaggc 780

gtttacgagg aactgcgcag ccatectgaa ctgatgcctc aggcagtgga ggaagccttg 840
 cgtttcagag cgccggcccc ggttttgagg cgcattgccca agcgggatac ggagatcggg 900
 5 gggcacctga ttaaagaagg tgatatgggt ttggcgtttg tggcatcggc aaatcgtgat 960
 gaagcaaagt ttgacagacc gcacatgttt gatatccgcc gccatcccaa tccgcatatt 1020
 10 gcgtttggcc acggcatecca tttttgcctt ggggccccgc ttgccgtct tgaagcaaatt 1080
 atcgcggttaa cgtctttgat ttctgctttt cctcatatgg agtgcgtcag tateactcgc 1140
 attgaaaaca gtgtgatata cggattaaag agcttccgtg tgaaaatgta a 1191
 15
 <210> 6
 <211> 1104
 <212> DNA
 20 <213> Escherichia coli
 <400> 6
 atggaccgca ttattcaatc accgggtaaa tacatccagg gcgctgatgt gattaatcgt 60
 25 ctgggcgaat acctgaagcc gctggcagaa cgctgggttag tggtaggtga caaatttgtt 120
 ttaggttttg ctcaatccac tgtcgagaaa agcttttaaag atgctggact ggtagtagaa 180
 attgcgccgt ttggcgggtga atgttcgcaa aatgagatcg accgtctgcg tggcatcgcg 240
 30 gagactgcgc agtgtggcgc aattctcggt atcgggtggcg gaaaaaccct cgatactgcc 300
 aaagcactgg cacatttcat ggggtgtccg gtagcgatcg caccgactat cgcctctacc 360
 35 gatgcaccgt gcagcgcatt gtctgttata tacaccgatg agggtagatt tgaccgctat 420

ctgctgttgc caaataacec gaatatgggc attgtcgaca ccaaaatcgt cgctggcgca 480
 cctgcacgtc tgtagcggc gggtatcggc gatgcgctgg caacctgggt tgaagecggt 540
 5 gectgtcttc gtagcggcgc gaccaccatg gcgggcggca agtgcacca ggctgcgctg 600
 gcactggctg aactgtgcta caacacctg ctggaagaag gcgaaaaage gatgcttgct 660
 10 gccgaacagc atgtagtac tccggcgctg gagcgctga ttgaagcgaa cacctatttg 720
 agcgggtgtg gttttgaaag tgggtgctg gctgcgggc acgcagtgc taacggcctg 780
 accgctatcc cggacgcgc tcaactattat cacgggtgaaa aagtggcatt cggtacgctg 840
 15 acgcagctgg ttctggaaaa tgcgccggtg gaggaaatcg aaaccgtagc tgcccttagc 900
 catgcggtag gtttgccaat aactctcgt caactggata ttaaagaaga tgtcccgcg 960
 20 aaaatgcgaa ttgtggcaga agcggcatgt gcagaaggtg aaaccattca caacatgcct 1020
 ggcggcgcga cgccagatca gggttacgcc gctctgctgg tagccgacca gtacggtcag 1080
 cgtttctgc aagagtggga ataa 1104

25

<210> 7

<211> 29

<212> DNA

30 <213> Artificial sequence

<220>

<223> Primer 1(FW T7)

35 <400> 7

cccgcgaaat taatacgact cactatagg

29

<210> 8

5 <211> 55

<212> DNA

<213> Artificial sequence

<220>

10 <223> Primer 2(RV-109-EcoRI-Rhf-2)

<400> 8

cggttgatgg cgggtgcagca cgaattccat tttcacacgg aagctcttta atccg

55

15

<210> 9

<211> 55

<212> DNA

<213> Artificial sequence

20

<220>

<223> Primer 3(FW Rhf-EcoRI-109-2)

<400> 9

25 cggattaaag agcttcctg tgaaaatgga attcgtgctg caccgccatc aaccg

55

<210> 10

<211> 34

30 <212> DNA

<213> Artificial sequence

<220>

<223> Primer 4(RV-Rhf-Link-Red-HindIII+EcoRV)

35

<400> 10

aagatatcaa gctttcagag tcgcagggcc agcc

34

5 <210> 11

<211> 33

<212> DNA

<213> Artificial sequence

10 <220>

<223> Primer 5(PCR cyp109B1 A11)

<400> 11

cctgcagggtt acattttcac acggaagctc ttt

33

15