

SEQUENCE LISTING

<110> Temaseck Life Sciences Laboratory Limited

<120> Production of Fatty Alcolhols in Rhodosporidium

<130> 2577-258PCT

<150> US 62/455,669

<151> 2017-02-07

<160> 62

<170> PatentIn version 3.5

<210> 1

<211> 1856

<212> DNA

<213> Rhodosporidium toruloides

<400> 1

atgggccagc aggcgacgct cgaggagctg tacacacgct cagaaatctc caagatcaag	60
caagtgcgagc cagctcttct cctcaccacc ccacaacata ccccgagcc cagcagcagct	120
ctcccacagc acccacagcc tgctgaccag ctcgagtgc tccacagatt tgcgcccttt	180
ggcgtccgc ggtcgcgccg gctgcagacg ttctccgtct ttgcctggac gacggcactg	240
cccatcctac tcggcgtctt cttcctgctc tggcgcgtca ggcttggcgt ggattgggag	300
tagcgggcga ctgagctgac ttgcgcatcc gctccagctc gttcccaccg ctctggccgg	360
ccgtcatcgc ctacctcacc tgggtctttt tcattgacca ggcgccgact cacggtggac	420
gggcgcagtc ttggctgcgg aagagtcgga tatgggtctg gtttgcagga tattatcccg	480
tcagggtcgt cgtcccgtct gttgcgcgtc ttgcgacctc gtcacggcc aactcggccg	540
accggctacc tccgaacttc ccgccaacag cttgatcaag gttcgtccac ctttccttca	600
gcttgagtga tctgtagagg agctgcagga tcaagccaa cccggggagg acctcggagg	660
acgacgccgc tgacttgctc tcctcctaca gagcgccgac ttgccgcctg accggaagta	720
cgtcttcggc tatcatccgc acggcgtcat aggcattgggc gccatcgcca acttcgcgac	780
cgacgcaacc ggcttctcga cactcttccc cggcttgaac cctcacctcc tcacctcca	840
aagcaacttc aagctccgc tctatcgca gttgctgctc gccctcggca tctgctccgt	900
ctcgatgaag agctgccaga acatcctgcg gcaaggtgcg ccagtcattc cgaacgggcg	960
gtcgagcgtg aactctgggg atgggaagag ctgaccttct gcctcactcc atccatgcag	1020

gtcctggctc ggctctcacc atcgttgtcg gtggcgagc cgagagcttg agtgcgcac	1080
ccggaaccgc cgacctcacg ctcaagcgac gaaaaggatt catcaagctc gcgatccggc	1140
aaggcgccga cctcgtgccc gtcttttcgt tcggcgagaa cgacgtgcgt cctctgctcg	1200
acttccgcta gcgaagccct tcgctgacgc tcccggtttc ttccccaga tcttcggcca	1260
gctgcgaaat gagcgagga cgcggtgta caagttgcag aagcgtttcc agggcgtatt	1320
cggcttcact ctccgtacgt tgcgccgtgt cgcttcaatc tgtcgagcgt ccagtcgctc	1380
acgcagctac aactcccaca gctctcttct acggtcgggg actcttcaac tgtacgccc	1440
agtctacgtg actagtctac cgtgggaggc actgaagagc acggctgacg tcccacctct	1500
ccgcgcagat aacgttggct tgatgccgta ccgccatccg atcgtctcgg tcggtgcgtc	1560
ccccctcgtc cctcctgacc tgcgggcttc agctaacaat tctcgacgac atctagtcgg	1620
tcgaccaatc tcggtgcagc agaaggacca cccaacgaca gcggatctcg aagaagtcca	1680
ggcgcggtat atcgcagaac tcaagcgtg cgttccagac gtctaccttt gcccgttgtc	1740
tcagactcgg taagacagat cactgacgct tcggtcactg gccgcgcagc atctgggaag	1800
actacaagga cgcctacgcc aaaagtcgca cgcgggagct caatattatc gcctga	1856

<210> 2
 <211> 3024
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 2	
atgagcacag tacggaggcg caaccagccg aagcgcaacg cttctgcgcc cgccaatcag	60
ctcgataccg acgattcttc gcccgtccc tctcgacccg actcgcccga tatcgacaaa	120
gcgccggtgc gagaggctcc actgttcaag cggaggctga aaaagggaga tggcaagggtg	180
gaagaccga aggtcattct ggacgagaag acgggcaggg tctccctcga ctttcctagg	240
acgccatatg accagctcga cctgagtac tcgttcctcg ccgccttgca cgaggcacc	300
ttcgagccaa agaagacgtg gacgaagcgg cggagaggct gggtcttttt gggaggattg	360
ctcggctctt gcgcagggtg cgttcgcttt cgttgttgcc gcttaaaatg cgtccagact	420
gacagctcgc tcattcctcc gtgtagctgg atgttcaccg aaggcgaccc actcgctctt	480
ctcgccaacc tcgacctcga cgccttctca tcctgggacc ttcagtcaat cctcgccgac	540

atgccttccc tctccgctct caacgtcacc gaactccttg caccaggctcg tgaatggctg	600
aacagccgag tgaacaactt tgacgtcggg cgagatgcgg ctgcgcgagg gctgaagaag	660
aagcatgcgg tcattctggt gccgggcatc atctcgtcgg tgcgtgcgct gcactttctca	720
catacgcctt acaggctgat gatccccgag tagggctctcg agtcgtggtc gaccaagcca	780
gatgcagctc ctttctttcg ctcgaaggctc tgggcgggaa cctcgatgat tcgtgccgtc	840
atcaagaaca aggaggcttg ggtcaaggcg atcagtctcg acccgtttac tggattggat	900
caggacgggt acaagatccg agctgctcag ggtgtgagtc gactcgttct ccagttcgcg	960
ctgctcgtg attgcatcct tctccagctc gacgccgctt ccgctttcat gcccggttac	1020
tgtaaatttg ccactttctca gctccttcgc tcgcgatgct gacccttcgt tactcgcagg	1080
gatctggcag aaggctcatcg agaacctcgc cgtcctggac tacgaccaca acgacctgtc	1140
tctcgcatac tacgactggc gacttgcgta cgttgcgctt gatcgctcgc ttgtcttgca	1200
acctgcttac ctttcgcgtc cttcagcttc tacaacctcg aagtccgaga tcgttacttc	1260
tctcggctca aggcttcgat agaattcaac ctcgccatca gcggccagaa gacggttctc	1320
gtctcgcaca gcatgggctc ttcggctctt ctgggtgcgtc tcggtttccc aatcttgccg	1380
cacactgtct gacccttgcg tccgttgtag tggttcttca agtaagcttg cctcgccatc	1440
cagcctcgcg aaggattgac tgacggttgc ggacttgcg aggtgggtcg agtcgccgag	1500
atacggcaac ggcgggtccgg actgggtcga gcggcacgtc tcggactggg tcaatgtcgc	1560
gggcacgatg ctcgggtgcgc tcaaactctc cttgcgatag atgctctact gaaccctccg	1620
gtcctatcag gcgttccgaa agccatggcc gccctcctct cgggcgagat gcgcgatacg	1680
gtcacgcttt ctcccgccgt catctacctc ctcgaacgct tcttctcccg ctccgaacgc	1740
gccaagctgt tccgctcgtg ggcagggtcg gcgagcatga tgcttaaggg aggaaacgac	1800
gtctggggcg acgagcagca ggcgccggac gacatggagg aatcggcaat cacgggcgga	1860
aagctgtagt gcgtccgccg tttcatacat ctcacgcatc ggctgacgcc cattcacgca	1920
gctacttccg gcctgagagt cactcgaaca cgtcgggaagt gacggaagac accgtccacc	1980
cgaacctgac gctcaacgac gcgaccagct tcctcctaga gaagggtccc ccgtcctacc	2040
agcagatgct ggcgagtaac ttctcgggtgc gtgcactgat acggactgag cgagtgagga	2100
agctcatcga ctcgcgacag ttcggcttcg agcgcgacga gcagcagctc atcaagaaca	2160

acgacgacca ttcgaagtgg agcaatccgc tcgaggtgca gctgccgaag gcgccgtcga	2220
tgaccatcta ctgcttgtac ggcgctcggca aggagactga gcgcgcgtac ttctaccagc	2280
agggtgcggc ccgccctcat cgccgtctgc gagcgttact gacgtcgaca gcggcttact	2340
acaggcggtt acgagcacga cgaaacgccg accgtcaatc tcacagactc gcgggcgtcc	2400
agcctcgaac ccgtttgtct cgagccgaac tgcaccgact cgacgccgcg cccgccgctc	2460
gacctgccgc tccagcgacg cgtatggatc gacggtagcg tgacgatgga cgagaagagc	2520
gtgccgaagg ttaggagcgg agtcgtcttc aacgacggag acgggacggt cagcttgctc	2580
agtctgggca gcatgtgtgt cgagggctgg aaggtgcgtc tccttcaacg tattctgtga	2640
gctgaaactg acgagcgccc ctcttagcga cctctgtaca atccggcagg catcaaggtc	2700
gtcacgcacg agatcttgca tgctccgctc gctttcgatc cgcgaggtgg accaacgact	2760
gcagaccatg tgcgctctgc ttctacaccg tttcccggct atcagctgac cctcatggtc	2820
gcaggtcgac atcctcggat cgtgtgtcct cctcgtcgcg ctacctccgc cagacaaggt	2880
tactgaccag cgcctcgtga ctgctacaga tccgaactga acgacgcgat cctcgacatt	2940
gctgccggcc aaggcgagcg cgtcaaggac caataccatt cgcggatcca ggacattgcg	3000
cgcaagatca ggtgggaggg ctga	3024

<210> 3

<211> 2788

<212> DNA

<213> Rhodosporidium toruloides

<400> 3

atggcctcgc tagaccgcc actcccaggc ccagccaacc tcgtcgacga cgccctccga	60
caccccgact cggcacctcc catcccgcc gactccgctc ctccttcgca tccctcgact	120
gcgactcagc cctccgcaac ttcgcgcgga cagctttcga ctgcttcgag ctatgcgagc	180
gatgtgtcga cgagggacgg gacaccggat ctggcgaatg ggtgagttga agcgttgag	240
gagaagtacc gaggtcttga tcagtgggga ggtgtgtgcg gcgaggagc gagatgcagg	300
ggctgacggg ctgcgtacgt gcagacaagg cgtaacgacg actattacga ccgtgacagg	360
caaggcgga aaggcgtca ccagacgtg cgttcccga ctcctcctcg tcgattgttc	420
ctccttcccc ccagtctttg gagactgacg acgttcctc caccgtcctt atcaatttcc	480
gaccgttctg ctccacgttc cgtgtcccc tcgtccccga cccgcctggc acctcccgt	540

cttccacctc ctaagcctca cccacgtcgg cgcctcctcc gtcgacgccc gcttctcctc	600
ctccacctcc tccatcacc cccgccccat cccagcacgc ggcggcgacc cgaagaagat	660
caaagtcctc cgttcccgtc gaaccactt cgccccacgc acctccact tcgaccgcca	720
caacctgacc tccgcctccg accccttccg cggcctctac actttattct ggatcgtaat	780
cttcgtcggg gactcaaga ctgtgtatca tcggtttgcg gagcagggtg ggtggggagg	840
tgaatggagg tttgcggcgt tgattagtcg tgatgggtgg gtgttggcgg ttagtgatgc	900
ggtgttggtt agcgcgtcgt tgttggtcgt gccatatgcc aaggtgcgta cgcgcttgcg	960
cttcggactc acagctcgaa gctcctgaac tgacggaacg atgctcccgg tctcgctacg	1020
cagctgctcg tacacggctg gatccggtac cacggcgccg gcgtaatcat ccaacacatc	1080
tgccaaacgc tctacctcgc catcgcgatc cgctggacct tccaccggtt cgtcctctcc	1140
cttccccctt cccttcctcc tctcctgtgc tgacgggaca atgagcgcag caactggccc	1200
tgggtccaaa gcggcttcat gaccctccac gccctctcga tgctcatgaa gatccatagc	1260
tactgttcgc taaacgggga gctttcggag cggcggagac agttgaggaa ggatgagggg	1320
aggctggagg aggtgctgga ggagatgggc gggaggagga gggcggagag ggaggcgagg	1380
gaggagtggg agaggcagt cggtgaggcg gcgagggcga aggagggcga ggcgggttcg	1440
cgcgagggaa agaaggagga ggtggcggcc cagtcgtcga cagacgcttc gacttcggct	1500
ctgtcgtcgg aggatgaggc ggcagcggcg ttgctgcggc atcgacagtc gactgcccga	1560
cggcgttcca tctcgccatc cgcctctcgc accactctt cttccgcttc ctctctcac	1620
cccgctccct cccgcgccga agagcccaa gaaggcgtcg agacgctcac atggcaccca	1680
tccgaccgag tcagcaaact tgctatcgcc atctgcgagg cgaaggacct tctcaccagc	1740
aacggcaaga aaccgctcac gttccccgag aacgtcacgt ttgcgaactt catcgactac	1800
ttgcttgctc cgacgctggt gtatgagctg gaataccctc ggacggactc gtgagtactc	1860
gttcgttcgc ttgggttaca ctttcgagtg gaactgacac gcttcgacgc cgtacagcat	1920
ccgccctctc tacatcctcg aaaagaccct cgcaaccttc ggcaccttct ccatacctcg	1980
cctgatcgtc gactcgttca tcctccccgt cacctcgcgc accgacaccc ccctcttcgg	2040
tttcgtcctc gacctcgcgc tcccgttcac actcgcgtac ctctcatct tttatgtcat	2100
tttcgagggc gtgtgtaatg ggtttgccga gttgacgagg tttgcggatc ggaacttggtg	2160

cgtgctggcg ctttcttccc tacttgcgaa ttgatccacc taactctctg tcttcccgt	2220
gcttcgacga ctggtggaac tcgtgcacgt tcgatgagtt ctcccgcaaa gtgcgcttcc	2280
gctcctcacc ctctcatcc taatcccggc taactctttt ctctgtctccg ctgagtgga	2340
ccgcccagtg cacgccttcc tcctccgcca cgtctacgcc gaaacgatgg cctcgtacaa	2400
gctctcgaag ctctcggccg cgttcgtcac gttcttgttc agcgcttgcg tgcacgaact	2460
cgtcatggcg gtcgtcacga agaagcttcg gctgtatttg ttctcgatgc aggtgcgtgc	2520
gatggctccg tatagtcggc tgcaaaaggg cgaaagccat catatactga tgaacttgga	2580
cgaacgtcgc acagatggcc caactcccgc tcatcatggt cggccgcgcc aagatcttcc	2640
gcaagtaccc tgccctcggc aacgtacgtt ctttctctct ctttcccccg tgcgccggac	2700
tgacgaccct cgccgctcac agctcttctt ctggctcgct cttctctcgg ggttcccgt	2760
tctcgggacg ctgtatttgc ggtactga	2788

<210> 4
 <211> 2250
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 4	
atgtccgtct cgacgaacgc tatcttcaag ctcccgaccg tcaagaacga gcccacttc	60
aactatgcgc caggcagccc cgagcgcgcg gctcttaaga aggccctcgc cgagctcgag	120
gccgccgcgc cgttcgaagt gcccgccttc gtcggcggca aggaggtgcg tcgcacgcct	180
tatcgcggac gcatgtgcat ccgagacggg agctgaccac tatacgtgtg caggtcaagg	240
tctcgggca gcgcaccgag cagcctatgc cccacaacca caagaaggcg ctctgcacct	300
tcggaacctc gactcctgag ctgccagcc aggctatcga ggctgcgctc gctgcccagc	360
ctgagtggga ggctatgccc ttcaacgacc gcgccgcat cttcttgagg gtgcgcgaga	420
ccacgctact ttagctgcga acgaatcgct cactcggctc tttgcaggct gccgacctga	480
tcagcggcaa gtaccgtgcc aagatctgcg ctgctacat gctcggccag ggcaagaacg	540
tctggcaggc cgagattgac gccgctgccg aggtgcgcat ctcgatccg tcgtccgctt	600
cgacagagcc tgaccttcgc tccgaacctg cagctcgccg acttcttccg cttcggcgcc	660
gctcaggtcg agcagctcta cgcgatgcag cccaccgaga actcgcctgg cgtctggaac	720

cgcaccgagt accgcccgt tgagggtttc gtctttgctg tcaccccctt caagtgcgtg	780
cagtcccgtt cttccttgac gtacacagct gacgagtact ttgcagcttc accgctatcg	840
gcggcaacct cgtcggcgct cccgcgctca ccggaaacgt cctcgtctgg aagccctcgc	900
ccatggcgac ctactcctcg tggctcgtct tccagatcct ccttgaggcc ggtctcccca	960
agaacgtcat ccagttcctc ccctgcccac acggcgacgc gaccatctcg ctcgttgacc	1020
gcgtcctctc gcaccgcatg ttcgccggtc tccacttcac tggctcgacc cacgtcttcc	1080
gccacctctg gcagaagatc ggcaacaaca tcaacaacta cctcagctac ccccgcatcg	1140
tcggcgagac cggcggaag aacttcagc tcgtccaccc ctctgccaac gtccgcgctg	1200
ctgtcaccgg tgccatccgc ggtgccttcg agtaccaggg tgcgtcgac gtctacctta	1260
ctcttcgagt cctaggctga cgctcctccg aatcgaggc cagaagtgct cggctctctc	1320
gcgtctttac gtcccgaagg gtctgtggga gggcgagggc aagttcaagg agatccttct	1380
ctccgaggtt gccaagatca ctctcgccc cgtcaccgag tttgagcact tcatgggtcc	1440
cgtcatgtgc gtcgcggtcg tcccgaagt ctgcttcct ccagctgacc tcgtctcgtc	1500
tcgccccgca gtcgcaggc ttcgttcgac aagtgcctca gctacgttga gaaggccaag	1560
caggcaggtg gcgaggtcct cgccggcggc aaggcgacg cgtcgagcgg ttactacgtc	1620
gagccgacca tcatcctgac caaggaccct cgctcgcta ccatggtcga cgagatcttc	1680
ggcccgggtc tactgttta catctacgag gacgaccagt tcgaggagac gtgcaagttg	1740
atcgaccaga cgacgacgta cgccctcact ggctgcatct tctcgacga ccgtgccgcg	1800
actgtcaagg ccggtgctct cctccgccac gctgcgggta actactacat caacgtgcgt	1860
tgcgcacact tgcttgaccg ttcggcgtag agttggactg gactgtttcg ctgaccttcc	1920
tcgtcccctt ctgctccct tcacccgctt tcgcactctt cttgctactg ccaatcatcc	1980
ttccgccctt tgtacgtct tccggcaaca ggacaagtcg accggtgctg ttgtcggtgc	2040
ccgtaagtct gaacatacat ctcttctcga catcgaccgg cactgacgct ccccttcac	2100
gcagagcctt tcggtggcgc acgcggatcg ggcacgaacg acaaggcggg ctcgatgacg	2160
ttcttcaccc gctggtgcca gccgcggagt gtgaaggaat ccttctgccc gccgaatct	2220
ttcccttacc cgtcgaacca gcgcgattaa	2250

<400>	5															54		
ttc	gct	cttc	aca	acc	agct	cag	ac	ggg	gct	gcg	atg	ggc	cag	cag	gcg	acg	ctc	
											Met	Gly	Gln	Gln	Ala	Thr	Leu	
											1				5			
gag	gag	ctg	tac	aca	cgc	tca	gaa	atc	tcc	aag	atc	aaa	ttt	gcg	ccc			102
Glu	Glu	Leu	Tyr	Thr	Arg	Ser	Glu	Ile	Ser	Lys	Ile	Lys	Phe	Ala	Pro			
		10					15					20						
ttt	ggc	gtc	ccg	cgg	tcg	cgc	cgg	ctg	cag	acg	ttc	tcc	gtc	ttt	gcc			150
Phe	Gly	Val	Pro	Arg	Ser	Arg	Arg	Leu	Gln	Thr	Phe	Ser	Val	Phe	Ala			
	25					30					35							
tgg	acg	acg	gca	ctg	ccc	atc	cta	ctc	ggc	gtc	ttc	ttc	ctg	ctc	tgc			198
Trp	Thr	Thr	Ala	Leu	Pro	Ile	Leu	Leu	Gly	Val	Phe	Phe	Leu	Leu	Cys			
40					45					50					55			
tcg	ttc	cca	ccg	ctc	tgg	ccg	gcc	gtc	atc	gcc	tac	ctc	acc	tgg	gtc			246
Ser	Phe	Pro	Pro	Leu	Trp	Pro	Ala	Val	Ile	Ala	Tyr	Leu	Thr	Trp	Val			
				60					65					70				
ttt	ttc	att	gac	cag	gcg	ccg	act	cac	ggc	gga	cgg	gcg	cag	tct	tgg			294
Phe	Phe	Ile	Asp	Gln	Ala	Pro	Thr	His	Gly	Gly	Arg	Ala	Gln	Ser	Trp			
			75					80					85					
ctg	cgg	aag	agt	cgg	ata	tgg	gtc	tgg	ttt	gca	gga	tat	tat	ccc	gtc			342
Leu	Arg	Lys	Ser	Arg	Ile	Trp	Val	Trp	Phe	Ala	Gly	Tyr	Tyr	Pro	Val			
		90					95					100						
agc	ttg	atc	aag	agc	gcc	gac	ttg	ccg	cct	gac	cgg	aag	tac	gtc	ttc			390
Ser	Leu	Ile	Lys	Ser	Ala	Asp	Leu	Pro	Pro	Asp	Arg	Lys	Tyr	Val	Phe			
	105					110					115							
ggc	tat	cat	ccg	cac	ggc	gtc	ata	ggc	atg	ggc	gcc	atc	gcc	aac	ttc			438
Gly	Tyr	His	Pro	His	Gly	Val	Ile	Gly	Met	Gly	Ala	Ile	Ala	Asn	Phe			
120					125				130						135			
gcg	acc	gac	gca	acc	ggc	ttc	tcg	aca	ctc	ttc	ccc	ggc	ttg	aac	cct			486
Ala	Thr	Asp	Ala	Thr	Gly	Phe	Ser	Thr	Leu	Phe	Pro	Gly	Leu	Asn	Pro			
				140					145					150				
cac	ctc	ctc	acc	ctc	caa	agc	aac	ttc	aag	ctc	ccg	ctc	tat	cg	gag			534
His	Leu	Leu	Thr	Leu	Gln	Ser	Asn	Phe	Lys	Leu	Pro	Leu	Tyr	Arg	Glu			
			155					160						165				

ttg ctg ctc gcc ctc ggc atc tgc tcc gtc tcg atg aag agc tgc cag Leu Leu Leu Ala Leu Gly Ile Cys Ser Val Ser Met Lys Ser Cys Gln 170 175 180	582
aac atc ctg cgg caa ggt cct ggc tcg gct ctc acc atc gtt gtc ggt Asn Ile Leu Arg Gln Gly Pro Gly Ser Ala Leu Thr Ile Val Val Gly 185 190 195	630
ggc gca gcc gag agc ttg agt gcg cat ccc gga acc gcc gac ctc acg Gly Ala Ala Glu Ser Leu Ser Ala His Pro Gly Thr Ala Asp Leu Thr 200 205 210 215	678
ctc aag cga cga aaa gga ttc atc aag ctc gcg atc cgg caa ggc gcc Leu Lys Arg Arg Lys Gly Phe Ile Lys Leu Ala Ile Arg Gln Gly Ala 220 225 230	726
gac ctc gtg ccc gtc ttt tcg ttc ggc gag aac gac atc ttc ggc cag Asp Leu Val Pro Val Phe Ser Phe Gly Glu Asn Asp Ile Phe Gly Gln 235 240 245	774
ctg cga aat gag cga ggg acg cgg ctg tac aag ttg cag aag cgt ttc Leu Arg Asn Glu Arg Gly Thr Arg Leu Tyr Lys Leu Gln Lys Arg Phe 250 255 260	822
cag ggc gta ttc ggc ttc act ctc cct ctc ttc tac ggt cgg gga ctc Gln Gly Val Phe Gly Phe Thr Leu Pro Leu Phe Tyr Gly Arg Gly Leu 265 270 275	870
ttc aac tat aac gtt ggc ttg atg ccg tac cgc cat ccg atc gtc tcg Phe Asn Tyr Asn Val Gly Leu Met Pro Tyr Arg His Pro Ile Val Ser 280 285 290 295	918
gtc gtc ggt cga cca atc tcg gtg cag cag aag gac cac cca acg aca Val Val Gly Arg Pro Ile Ser Val Gln Gln Lys Asp His Pro Thr Thr 300 305 310	966
gcg gat ctc gaa gaa gtc cag gcg cgg tat atc gca gaa ctc aag cgc Ala Asp Leu Glu Glu Val Gln Ala Arg Tyr Ile Ala Glu Leu Lys Arg 315 320 325	1014
atc tgg gaa gac tac aag gac gcc tac gcc aaa agt cgc acg cgg gag Ile Trp Glu Asp Tyr Lys Asp Ala Tyr Ala Lys Ser Arg Thr Arg Glu 330 335 340	1062
ctc aat att atc gcc tga cctcctcgca acggacgctt cccgccggcc Leu Asn Ile Ile Ala 345	1110
aaaagaccct ccattgcctg tattctctcc tgttgtgtgcc atccctagcc gttgtcgatt	1170
ccccttgccg ttcccctctc ctcgatatct cgatatccgt tctctgtagc ttgccttccg	1230
ctttgcacca tgccccggtt caacacg	1257

<210> 6
<211> 348
<212> PRT
<213> Rhodosporidium toruloides

<400> 6

Met Gly Gln Gln Ala Thr Leu Glu Glu Leu Tyr Thr Arg Ser Glu Ile
1 5 10 15

Ser Lys Ile Lys Phe Ala Pro Phe Gly Val Pro Arg Ser Arg Arg Leu
20 25 30

Gln Thr Phe Ser Val Phe Ala Trp Thr Thr Ala Leu Pro Ile Leu Leu
35 40 45

Gly Val Phe Phe Leu Leu Cys Ser Phe Pro Pro Leu Trp Pro Ala Val
50 55 60

Ile Ala Tyr Leu Thr Trp Val Phe Phe Ile Asp Gln Ala Pro Thr His
65 70 75 80

Gly Gly Arg Ala Gln Ser Trp Leu Arg Lys Ser Arg Ile Trp Val Trp
85 90 95

Phe Ala Gly Tyr Tyr Pro Val Ser Leu Ile Lys Ser Ala Asp Leu Pro
100 105 110

Pro Asp Arg Lys Tyr Val Phe Gly Tyr His Pro His Gly Val Ile Gly
115 120 125

Met Gly Ala Ile Ala Asn Phe Ala Thr Asp Ala Thr Gly Phe Ser Thr
130 135 140

Leu Phe Pro Gly Leu Asn Pro His Leu Leu Thr Leu Gln Ser Asn Phe
145 150 155 160

Lys Leu Pro Leu Tyr Arg Glu Leu Leu Leu Ala Leu Gly Ile Cys Ser
165 170 175

Val Ser Met Lys Ser Cys Gln Asn Ile Leu Arg Gln Gly Pro Gly Ser
180 185 190

Ala Leu Thr Ile Val Val Gly Gly Ala Ala Glu Ser Leu Ser Ala His
 195 200 205

Pro Gly Thr Ala Asp Leu Thr Leu Lys Arg Arg Lys Gly Phe Ile Lys
 210 215 220

Leu Ala Ile Arg Gln Gly Ala Asp Leu Val Pro Val Phe Ser Phe Gly
 225 230 235 240

Glu Asn Asp Ile Phe Gly Gln Leu Arg Asn Glu Arg Gly Thr Arg Leu
 245 250 255

Tyr Lys Leu Gln Lys Arg Phe Gln Gly Val Phe Gly Phe Thr Leu Pro
 260 265 270

Leu Phe Tyr Gly Arg Gly Leu Phe Asn Tyr Asn Val Gly Leu Met Pro
 275 280 285

Tyr Arg His Pro Ile Val Ser Val Val Gly Arg Pro Ile Ser Val Gln
 290 295 300

Gln Lys Asp His Pro Thr Thr Ala Asp Leu Glu Glu Val Gln Ala Arg
 305 310 315 320

Tyr Ile Ala Glu Leu Lys Arg Ile Trp Glu Asp Tyr Lys Asp Ala Tyr
 325 330 335

Ala Lys Ser Arg Thr Arg Glu Leu Asn Ile Ile Ala
 340 345

<210> 7
 <211> 2336
 <212> DNA
 <213> Rhodosporidium toruloides

<220>
 <221> CDS
 <222> (54)..(2264)

<400> 7
 ctcactagac ctcgctcacc tcgcttgact ggactgctgc tggctgagaa agg atg

Met
1

agc aca gta cgg agg cgc aac cag ccg aag cgc aac gct tct gcg ccc Ser Thr Val Arg Arg Arg Asn Gln Pro Lys Arg Asn Ala Ser Ala Pro	104
5 10 15	
gcc aat cag ctc gat acc gac gat tct tcg ccc gct ccc tct cga ccc Ala Asn Gln Leu Asp Thr Asp Asp Ser Ser Pro Ala Pro Ser Arg Pro	152
20 25 30	
gac tcg ccc gat atc gac aaa gcg ccg gtg cga gag gtc cca ctg ttc Asp Ser Pro Asp Ile Asp Lys Ala Pro Val Arg Glu Val Pro Leu Phe	200
35 40 45	
aag cgg agg ctg aaa aag gga gat ggc aag gtg gaa gac ccg aag gtc Lys Arg Arg Leu Lys Lys Gly Asp Gly Lys Val Glu Asp Pro Lys Val	248
50 55 60 65	
att ctg gac gag aag acg ggc agg gtc tcc ctc gac ttt cct agg acg Ile Leu Asp Glu Lys Thr Gly Arg Val Ser Leu Asp Phe Pro Arg Thr	296
70 75 80	
cca tat gac cag ctc gac ctg agt gac tcg ttc ctc gcc gcc ttg cac Pro Tyr Asp Gln Leu Asp Leu Ser Asp Ser Phe Leu Ala Ala Leu His	344
85 90 95	
gag gca ccc ttc gag cca aag aag acg tgg acg aag cgg cgg aga ggc Glu Ala Pro Phe Glu Pro Lys Lys Thr Trp Thr Lys Arg Arg Arg Gly	392
100 105 110	
tgg ttc ttt ttg gga gga ttg ctc ggt ctc tgc gca ggc tgg atg ttc Trp Phe Phe Leu Gly Gly Leu Leu Gly Leu Cys Ala Gly Trp Met Phe	440
115 120 125	
acc gaa ggc gac cca ctc gcc tct ctc gcc aac ctc gac ctc gac gcc Thr Glu Gly Asp Pro Leu Ala Ser Leu Ala Asn Leu Asp Leu Asp Ala	488
130 135 140 145	
ttc tca tcc tgg gac ctt cag tca atc ctc gcc gac atg cct tcc ctc Phe Ser Ser Trp Asp Leu Gln Ser Ile Leu Ala Asp Met Pro Ser Leu	536
150 155 160	
tcc gct ctc aac gtc acc gaa ctc ctt gca cca ggt cgt gaa tgg ctg Ser Ala Leu Asn Val Thr Glu Leu Leu Ala Pro Gly Arg Glu Trp Leu	584
165 170 175	
aac agc cga gtg aac aac ttt gac gtc ggg cga gat gcg gct gcg cga Asn Ser Arg Val Asn Asn Phe Asp Val Gly Arg Asp Ala Ala Ala Arg	632
180 185 190	
ggg ctg aag aag aag cat gcg gtc att ctg gtg ccg ggc atc atc tcg Gly Leu Lys Lys Lys His Ala Val Ile Leu Val Pro Gly Ile Ile Ser	680
195 200 205	

tcg ggt ctc gag tcg tgg tcg acc aag cca gat gca gct ccc ttc ttt	728
Ser Gly Leu Glu Ser Trp Ser Thr Lys Pro Asp Ala Ala Pro Phe Phe	
210 215 220 225	
cgc tcg aag gtc tgg gcg gga acc tcg atg att cgt gcc gtc atc aag	776
Arg Ser Lys Val Trp Ala Gly Thr Ser Met Ile Arg Ala Val Ile Lys	
230 235 240	
aac aag gag gct tgg gtc aag gcg atc agt ctc gac ccg ttt act gga	824
Asn Lys Glu Ala Trp Val Lys Ala Ile Ser Leu Asp Pro Phe Thr Gly	
245 250 255	
ttg gat cag gac ggg tac aag atc cga gct gct cag ggt ctc gac gcc	872
Leu Asp Gln Asp Gly Tyr Lys Ile Arg Ala Ala Gln Gly Leu Asp Ala	
260 265 270	
gct tcc gct ttc atg ccc ggg tac tgg atc tgg cag aag gtc atc gag	920
Ala Ser Ala Phe Met Pro Gly Tyr Trp Ile Trp Gln Lys Val Ile Glu	
275 280 285	
aac ctc gcc gtc ctg gac tac gac cac aac gac ctg tct ctc gca tca	968
Asn Leu Ala Val Leu Asp Tyr Asp His Asn Asp Leu Ser Leu Ala Ser	
290 295 300 305	
tac gac tgg cga ctt gcc ttc tac aac ctc gaa gtc cga gat cgt tac	1016
Tyr Asp Trp Arg Leu Ala Phe Tyr Asn Leu Glu Val Arg Asp Arg Tyr	
310 315 320	
ttc tct cgg ctc aag gct tcg ata gaa ttc aac ctc gcc atc agc ggc	1064
Phe Ser Arg Leu Lys Ala Ser Ile Glu Phe Asn Leu Ala Ile Ser Gly	
325 330 335	
cag aag acg gtt ctc gtc tcg cac agc atg ggc tct tcg gct ctt ctg	1112
Gln Lys Thr Val Leu Val Ser His Ser Met Gly Ser Ser Ala Leu Leu	
340 345 350	
tgg ttc ttc aag tgg gtc gag tcg ccg aga tac ggc aac ggc ggt ccg	1160
Trp Phe Phe Lys Trp Val Glu Ser Pro Arg Tyr Gly Asn Gly Gly Pro	
355 360 365	
gac tgg gtc gag cgg cac gtc tcg gac tgg gtc aat gtc gcg ggc acg	1208
Asp Trp Val Glu Arg His Val Ser Asp Trp Val Asn Val Ala Gly Thr	
370 375 380 385	
atg ctc ggc gtt ccg aaa gcc atg gcc gcc ctc ctc tcg ggc gag atg	1256
Met Leu Gly Val Pro Lys Ala Met Ala Ala Leu Leu Ser Gly Glu Met	
390 395 400	
cgc gat acg gtc acg ctt tct ccc gcc gtc atc tac ctc ctc gaa cgc	1304
Arg Asp Thr Val Thr Leu Ser Pro Ala Val Ile Tyr Leu Leu Glu Arg	
405 410 415	
ttc ttc tcc cgc tcc gaa cgc gcc aag ctg ttc cgc tcg tgg gca ggt	1352

Phe	Phe	Ser	Arg	Ser	Glu	Arg	Ala	Lys	Leu	Phe	Arg	Ser	Trp	Ala	Gly		
		420					425					430					
gcg	gcg	agc	atg	atg	ctt	aag	gga	gga	aac	gac	gtc	tgg	ggc	gac	gag	1400	
Ala	Ala	Ser	Met	Met	Leu	Lys	Gly	Gly	Asn	Asp	Val	Trp	Gly	Asp	Glu		
		435				440					445						
cag	cag	gcg	ccg	gac	gac	atg	gag	gaa	tcg	gca	atc	acg	ggc	gga	aag	1448	
Gln	Gln	Ala	Pro	Asp	Asp	Met	Glu	Glu	Ser	Ala	Ile	Thr	Gly	Gly	Lys		
		450			455					460					465		
ctg	tac	tac	ttc	cgg	cct	gag	agt	cac	tcg	aac	acg	tcg	gaa	gtg	acg	1496	
Leu	Tyr	Tyr	Phe	Arg	Pro	Glu	Ser	His	Ser	Asn	Thr	Ser	Glu	Val	Thr		
				470					475					480			
gaa	gac	acc	gtc	cac	ccg	aac	ctg	acg	ctc	aac	gac	gcg	acc	agc	ttc	1544	
Glu	Asp	Thr	Val	His	Pro	Asn	Leu	Thr	Leu	Asn	Asp	Ala	Thr	Ser	Phe		
			485					490					495				
ctc	cta	gag	aag	gtc	ccg	ccg	tcc	tac	cag	cag	atg	ctg	gcg	agt	aac	1592	
Leu	Leu	Glu	Lys	Val	Pro	Pro	Ser	Tyr	Gln	Gln	Met	Leu	Ala	Ser	Asn		
		500					505					510					
ttc	tcg	ttc	ggc	ttc	gag	cgc	gac	gag	cag	cag	ctc	atc	aag	aac	aac	1640	
Phe	Ser	Phe	Gly	Phe	Glu	Arg	Asp	Glu	Gln	Gln	Leu	Ile	Lys	Asn	Asn		
		515				520					525						
gac	gac	cat	tcg	aag	tgg	agc	aat	ccg	ctc	gag	gtg	cag	ctg	ccg	aag	1688	
Asp	Asp	His	Ser	Lys	Trp	Ser	Asn	Pro	Leu	Glu	Val	Gln	Leu	Pro	Lys		
					535					540					545		
gcg	ccg	tcg	atg	acc	atc	tac	tgc	ttg	tac	ggc	gtc	ggc	aag	gag	act	1736	
Ala	Pro	Ser	Met	Thr	Ile	Tyr	Cys	Leu	Tyr	Gly	Val	Gly	Lys	Glu	Thr		
				550					555					560			
gag	cgc	gcg	tac	ttc	tac	cag	cag	ggc	ggt	tac	gag	cac	gac	gaa	acg	1784	
Glu	Arg	Ala	Tyr	Phe	Tyr	Gln	Gln	Gly	Gly	Tyr	Glu	His	Asp	Glu	Thr		
			565					570					575				
ccg	acc	gtc	aat	ctc	aca	gac	tcg	cgg	gcg	tcc	agc	ctc	gaa	ccc	gtt	1832	
Pro	Thr	Val	Asn	Leu	Thr	Asp	Ser	Arg	Ala	Ser	Ser	Leu	Glu	Pro	Val		
		580					585					590					
tgt	ctc	gag	ccg	aac	tgc	acc	gac	tcg	acg	ccg	cgc	ccg	ccg	ctc	gac	1880	
Cys	Leu	Glu	Pro	Asn	Cys	Thr	Asp	Ser	Thr	Pro	Arg	Pro	Pro	Leu	Asp		
		595				600					605						
ctg	ccg	ctc	cag	cga	cgc	gta	tgg	atc	gac	ggt	agc	gtg	acg	atg	gac	1928	
Leu	Pro	Leu	Gln	Arg	Arg	Val	Trp	Ile	Asp	Gly	Ser	Val	Thr	Met	Asp		
					615					620					625		
gag	aag	agc	gtg	ccg	aag	gtt	agg	agc	gga	gtc	gtc	ttc	aac	gac	gga	1976	
Glu	Lys	Ser	Val	Pro	Lys	Val	Arg	Ser	Gly	Val	Val	Phe	Asn	Asp	Gly		
				630					635					640			

gac ggg acg gtc agc ttg ctc agt ctg ggc agc atg tgt gtc gag ggc	2024
Asp Gly Thr Val Ser Leu Leu Ser Leu Gly Ser Met Cys Val Glu Gly	
645 650 655	

tgg aag cga cct ctg tac aat ccg gca ggc atc aag gtc gtc acg cac	2072
Trp Lys Arg Pro Leu Tyr Asn Pro Ala Gly Ile Lys Val Val Thr His	
660 665 670	

gag atc ttg cat gct ccg ctc gct ttc gat ccg cga ggt gga cca acg	2120
Glu Ile Leu His Ala Pro Leu Ala Phe Asp Pro Arg Gly Gly Pro Thr	
675 680 685	

act gca gac cat gtc gac atc ctc gga tca tcc gaa ctg aac gac gcg	2168
Thr Ala Asp His Val Asp Ile Leu Gly Ser Ser Glu Leu Asn Asp Ala	
690 695 700 705	

atc ctc gac att gct gcc ggc caa ggc gag cgc gtc aag gac caa tac	2216
Ile Leu Asp Ile Ala Ala Gly Gln Gly Glu Arg Val Lys Asp Gln Tyr	
710 715 720	

cat tcg cgg atc cag gac att gcg cgc aag atc agg tgg gag ggc tga	2264
His Ser Arg Ile Gln Asp Ile Ala Arg Lys Ile Arg Trp Glu Gly	
725 730 735	

cccaccctgt acgccttgcc cagcgtttct cccctttctc gtgtagagat cgcaacagtg	2324
---	------

ttgtttcccg ct	2336
---------------	------

<210> 8
 <211> 736
 <212> PRT
 <213> Rhodosporidium toruloides

<400> 8

Met Ser Thr Val Arg Arg Arg Asn Gln Pro Lys Arg Asn Ala Ser Ala
1 5 10 15

Pro Ala Asn Gln Leu Asp Thr Asp Asp Ser Ser Pro Ala Pro Ser Arg
20 25 30

Pro Asp Ser Pro Asp Ile Asp Lys Ala Pro Val Arg Glu Val Pro Leu
35 40 45

Phe Lys Arg Arg Leu Lys Lys Gly Asp Gly Lys Val Glu Asp Pro Lys
50 55 60

Val Ile Leu Asp Glu Lys Thr Gly Arg Val Ser Leu Asp Phe Pro Arg

65	70	75	80
Thr Pro Tyr Asp Gln Leu Asp Leu Ser Asp Ser Phe Leu Ala Ala Leu	85	90	95
His Glu Ala Pro Phe Glu Pro Lys Lys Thr Trp Thr Lys Arg Arg Arg	100	105	110
Gly Trp Phe Phe Leu Gly Gly Leu Leu Gly Leu Cys Ala Gly Trp Met	115	120	125
Phe Thr Glu Gly Asp Pro Leu Ala Ser Leu Ala Asn Leu Asp Leu Asp	130	135	140
Ala Phe Ser Ser Trp Asp Leu Gln Ser Ile Leu Ala Asp Met Pro Ser	145	150	155
Leu Ser Ala Leu Asn Val Thr Glu Leu Leu Ala Pro Gly Arg Glu Trp	165	170	175
Leu Asn Ser Arg Val Asn Asn Phe Asp Val Gly Arg Asp Ala Ala Ala	180	185	190
Arg Gly Leu Lys Lys Lys His Ala Val Ile Leu Val Pro Gly Ile Ile	195	200	205
Ser Ser Gly Leu Glu Ser Trp Ser Thr Lys Pro Asp Ala Ala Pro Phe	210	215	220
Phe Arg Ser Lys Val Trp Ala Gly Thr Ser Met Ile Arg Ala Val Ile	225	230	235
Lys Asn Lys Glu Ala Trp Val Lys Ala Ile Ser Leu Asp Pro Phe Thr	245	250	255
Gly Leu Asp Gln Asp Gly Tyr Lys Ile Arg Ala Ala Gln Gly Leu Asp	260	265	270
Ala Ala Ser Ala Phe Met Pro Gly Tyr Trp Ile Trp Gln Lys Val Ile	275	280	285

Glu Asn Leu Ala Val Leu Asp Tyr Asp His Asn Asp Leu Ser Leu Ala
290 295 300

Ser Tyr Asp Trp Arg Leu Ala Phe Tyr Asn Leu Glu Val Arg Asp Arg
305 310 315 320

Tyr Phe Ser Arg Leu Lys Ala Ser Ile Glu Phe Asn Leu Ala Ile Ser
325 330 335

Gly Gln Lys Thr Val Leu Val Ser His Ser Met Gly Ser Ser Ala Leu
340 345 350

Leu Trp Phe Phe Lys Trp Val Glu Ser Pro Arg Tyr Gly Asn Gly Gly
355 360 365

Pro Asp Trp Val Glu Arg His Val Ser Asp Trp Val Asn Val Ala Gly
370 375 380

Thr Met Leu Gly Val Pro Lys Ala Met Ala Ala Leu Leu Ser Gly Glu
385 390 395 400

Met Arg Asp Thr Val Thr Leu Ser Pro Ala Val Ile Tyr Leu Leu Glu
405 410 415

Arg Phe Phe Ser Arg Ser Glu Arg Ala Lys Leu Phe Arg Ser Trp Ala
420 425 430

Gly Ala Ala Ser Met Met Leu Lys Gly Gly Asn Asp Val Trp Gly Asp
435 440 445

Glu Gln Gln Ala Pro Asp Asp Met Glu Glu Ser Ala Ile Thr Gly Gly
450 455 460

Lys Leu Tyr Tyr Phe Arg Pro Glu Ser His Ser Asn Thr Ser Glu Val
465 470 475 480

Thr Glu Asp Thr Val His Pro Asn Leu Thr Leu Asn Asp Ala Thr Ser
485 490 495

Phe Leu Leu Glu Lys Val Pro Pro Ser Tyr Gln Gln Met Leu Ala Ser

500

505

510

Asn Phe Ser Phe Gly Phe Glu Arg Asp Glu Gln Gln Leu Ile Lys Asn
 515 520 525

Asn Asp Asp His Ser Lys Trp Ser Asn Pro Leu Glu Val Gln Leu Pro
 530 535 540

Lys Ala Pro Ser Met Thr Ile Tyr Cys Leu Tyr Gly Val Gly Lys Glu
 545 550 555 560

Thr Glu Arg Ala Tyr Phe Tyr Gln Gln Gly Gly Tyr Glu His Asp Glu
 565 570 575

Thr Pro Thr Val Asn Leu Thr Asp Ser Arg Ala Ser Ser Leu Glu Pro
 580 585 590

Val Cys Leu Glu Pro Asn Cys Thr Asp Ser Thr Pro Arg Pro Pro Leu
 595 600 605

Asp Leu Pro Leu Gln Arg Arg Val Trp Ile Asp Gly Ser Val Thr Met
 610 615 620

Asp Glu Lys Ser Val Pro Lys Val Arg Ser Gly Val Val Phe Asn Asp
 625 630 635 640

Gly Asp Gly Thr Val Ser Leu Leu Ser Leu Gly Ser Met Cys Val Glu
 645 650 655

Gly Trp Lys Arg Pro Leu Tyr Asn Pro Ala Gly Ile Lys Val Val Thr
 660 665 670

His Glu Ile Leu His Ala Pro Leu Ala Phe Asp Pro Arg Gly Gly Pro
 675 680 685

Thr Thr Ala Asp His Val Asp Ile Leu Gly Ser Ser Glu Leu Asn Asp
 690 695 700

Ala Ile Leu Asp Ile Ala Ala Gly Gln Gly Glu Arg Val Lys Asp Gln
 705 710 715 720

Tyr His Ser Arg Ile Gln Asp Ile Ala Arg Lys Ile Arg Trp Glu Gly
725 730 735

<210> 9
<211> 2153
<212> DNA
<213> Rhodosporidium toruloides

<220>
<221> CDS
<222> (7)..(2043)

<400> 9
caaaga atg gcc tcg cta gac ccg cca ctc cca ggc cca gcc aac ctc 48
Met Ala Ser Leu Asp Pro Pro Leu Pro Gly Pro Ala Asn Leu
1 5 10

gtc gac gac gcc ctc cga cac ccc gac tcg gca cct ccc atc ccg ccc 96
Val Asp Asp Ala Leu Arg His Pro Asp Ser Ala Pro Pro Ile Pro Pro
15 20 25 30

gac tcc gct cct cct tcg cat ccc tcg act gcg act cag ccc tcc gca 144
Asp Ser Ala Pro Pro Ser His Pro Ser Thr Ala Thr Gln Pro Ser Ala
35 40 45

act tcg cgc gga cag ctt tcg act gct tcg agc tat gcg agc gat gtg 192
Thr Ser Arg Gly Gln Leu Ser Thr Ala Ser Ser Tyr Ala Ser Asp Val
50 55 60

tcg acg agg gac ggg aca ccg gat ctg gcg aat gga caa ggc gta acg 240
Ser Thr Arg Asp Gly Thr Pro Asp Leu Ala Asn Gly Gln Gly Val Thr
65 70 75

acg act att acg acc gtg aca ggc aag ggc gga aag gcg gtc acc cag 288
Thr Thr Ile Thr Thr Val Thr Gly Lys Gly Gly Lys Ala Val Thr Gln
80 85 90

acc ctc acc cac gtc ggc gcc tcc tcc gtc gac gcc cgc ttc tcc tcc 336
Thr Leu Thr His Val Gly Ala Ser Ser Val Asp Ala Arg Phe Ser Ser
95 100 105 110

tcc acc tcc tcc atc acc ctc cgc ccc atc cca gca cgc ggc ggc gac 384
Ser Thr Ser Ser Ile Thr Leu Arg Pro Ile Pro Ala Arg Gly Gly Asp
115 120 125

ccg aag aag atc aaa gtc ctc cgt tcc cgt cga acc cac ttc gcc cca 432
Pro Lys Lys Ile Lys Val Leu Arg Ser Arg Arg Thr His Phe Ala Pro
130 135 140

cgc acc tcc cac ttc gac cgc cac aac ctg acc tcc gcc tcc gac ccc 480
Arg Thr Ser His Phe Asp Arg His Asn Leu Thr Ser Ala Ser Asp Pro

145	150	155	
ttc cgc ggc ctc tac act tta ttc tgg atc gta atc ttc gtc gga gca Phe Arg Gly Leu Tyr Thr Leu Phe Trp Ile Val Ile Phe Val Gly Ala 160 165 170			528
ctc aag act gtg tat cat cgg ttt gcg gag cag ggt ggg tgg gga ggt Leu Lys Thr Val Tyr His Arg Phe Ala Glu Gln Gly Gly Trp Gly Gly 175 180 185 190			576
gaa tgg agg ttt gcg gcg ttg att agt cgt gat ggg tgg gtg ttg gcg Glu Trp Arg Phe Ala Ala Leu Ile Ser Arg Asp Gly Trp Val Leu Ala 195 200 205			624
gtt agt gat gcg gtg ttg gtt agc gcg tcg ttg ttg tgc gtg cca tat Val Ser Asp Ala Val Leu Val Ser Ala Ser Leu Leu Cys Val Pro Tyr 210 215 220			672
gcc aag ctg ctc gta cac ggc tgg atc cgg tac cac ggc gcc ggc gta Ala Lys Leu Leu Val His Gly Trp Ile Arg Tyr His Gly Ala Gly Val 225 230 235			720
atc atc caa cac atc tgc caa acg ctc tac ctc gcc atc gcg atc cgc Ile Ile Gln His Ile Cys Gln Thr Leu Tyr Leu Ala Ile Ala Ile Arg 240 245 250			768
tgg acc ttc cac cgc aac tgg ccc tgg gtc caa agc ggc ttc atg acc Trp Thr Phe His Arg Asn Trp Pro Trp Val Gln Ser Gly Phe Met Thr 255 260 265 270			816
ctc cac gcc ctc tcg atg ctc atg aag atc cat agc tac tgt tcg cta Leu His Ala Leu Ser Met Leu Met Lys Ile His Ser Tyr Cys Ser Leu 275 280 285			864
aac ggg gag ctt tcg gag cgg cgg aga cag ttg agg aag gat gag ggg Asn Gly Glu Leu Ser Glu Arg Arg Arg Gln Leu Arg Lys Asp Glu Gly 290 295 300			912
agg ctg gag gag gtg ctg gag gag atg ggc ggg agg agg agg gcg gag Arg Leu Glu Glu Val Leu Glu Glu Met Gly Gly Arg Arg Arg Ala Glu 305 310 315			960
agg gag gcg agg gag gag tgg gag agg cag tgc ggt gag gcg gcg agg Arg Glu Ala Arg Glu Glu Trp Glu Arg Gln Cys Gly Glu Ala Ala Arg 320 325 330			1008
gcg aag gag ggc gag gcg ggt tcg cgc gag gga aag aag gag gag gtg Ala Lys Glu Gly Glu Ala Gly Ser Arg Glu Gly Lys Lys Glu Glu Val 335 340 345 350			1056
gcg gcc cag tcg tcg aca gac gct tcg act tcg gct ctg tcg tcg gag Ala Ala Gln Ser Ser Thr Asp Ala Ser Thr Ser Ala Leu Ser Ser Glu 355 360 365			1104

gat gag gcg gca gcg gcg ttg ctg cgg cat cga cag tcg act gcc cga Asp Glu Ala Ala Ala Ala Leu Leu Arg His Arg Gln Ser Thr Ala Arg 370 375 380	1152
cgg cgt tcc atc tcg cca tcc gcc tct cgc acc cac tct tct tcc gct Arg Arg Ser Ile Ser Pro Ser Ala Ser Arg Thr His Ser Ser Ser Ala 385 390 395	1200
tcc tcc tct cac ccc gct ccc tcc cgc gcc gaa gag ccc caa gaa ggc Ser Ser Ser His Pro Ala Pro Ser Arg Ala Glu Glu Pro Gln Glu Gly 400 405 410	1248
gtc gag acg ctc aca tgg cac cca tcc gac cga gtc agc aaa ctt gct Val Glu Thr Leu Thr Trp His Pro Ser Asp Arg Val Ser Lys Leu Ala 415 420 425 430	1296
atc gcc atc tgc gag gcg aag gac ctt ctc acc agc aac ggc aag aaa Ile Ala Ile Cys Glu Ala Lys Asp Leu Leu Thr Ser Asn Gly Lys Lys 435 440 445	1344
ccc gtc acg ttc ccc gag aac gtc acg ttt gcg aac ttc atc gac tac Pro Val Thr Phe Pro Glu Asn Val Thr Phe Ala Asn Phe Ile Asp Tyr 450 455 460	1392
ttg ctt gtg ccg acg ctg gtg tat gag ctg gaa tac cct cgg acg gac Leu Leu Val Pro Thr Leu Val Tyr Glu Leu Glu Tyr Pro Arg Thr Asp 465 470 475	1440
tcc atc cgc cct ctc tac atc ctc gaa aag acc ctc gca acc ttc ggc Ser Ile Arg Pro Leu Tyr Ile Leu Glu Lys Thr Leu Ala Thr Phe Gly 480 485 490	1488
acc ttc tcc atc ctc gtc ctg atc gtc gac tcg ttc atc ctc ccc gtc Thr Phe Ser Ile Leu Val Leu Ile Val Asp Ser Phe Ile Leu Pro Val 495 500 505 510	1536
acc tcg cgc acc gac acc ccc ctc ttc ggt ttc gtc ctc gac ctc gcg Thr Ser Arg Thr Asp Thr Pro Leu Phe Gly Phe Val Leu Asp Leu Ala 515 520 525	1584
ctc ccg ttc aca ctc gcg tac ctc ctc atc ttt tat gtc att ttc gag Leu Pro Phe Thr Leu Ala Tyr Leu Leu Ile Phe Tyr Val Ile Phe Glu 530 535 540	1632
ggc gtg tgt aat ggg ttt gcc gag ttg acg agg ttt gcg gat cgg aac Gly Val Cys Asn Gly Phe Ala Glu Leu Thr Arg Phe Ala Asp Arg Asn 545 550 555	1680
ttc ttc gac gac tgg tgg aac tcg tgc acg ttc gat gag ttc tcc cgc Phe Phe Asp Asp Trp Trp Asn Ser Cys Thr Phe Asp Glu Phe Ser Arg 560 565 570	1728
aaa tgg aac cgc cca gtg cac gcc ttc ctc ctc cgc cac gtc tac gcc Lys Trp Asn Arg Pro Val His Ala Phe Leu Leu Arg His Val Tyr Ala	1776

575	580	585	590	
gaa acg atg gcc tcg tac aag ctc tcg aag ctc tcg gcc gcg ttc gtc				1824
Glu Thr Met Ala Ser Tyr Lys Leu Ser Lys Leu Ser Ala Ala Phe Val				
	595	600	605	
acg ttc ttg ttc agc gct tgc gtg cac gaa ctc gtc atg gcg gtc gtc				1872
Thr Phe Leu Phe Ser Ala Cys Val His Glu Leu Val Met Ala Val Val				
	610	615	620	
acg aag aag ctt cgg ctg tat ttg ttc tcg atg cag atg gcc caa ctc				1920
Thr Lys Lys Leu Arg Leu Tyr Leu Phe Ser Met Gln Met Ala Gln Leu				
	625	630	635	
ccg ctc atc atg gtc ggc cgc gcc aag atc ttc cgc aag tac cct gcc				1968
Pro Leu Ile Met Val Gly Arg Ala Lys Ile Phe Arg Lys Tyr Pro Ala				
	640	645	650	
ctc ggc aac ctc ttc ttc tgg ctc gct ctt ctc tcg ggg ttc ccg ctt				2016
Leu Gly Asn Leu Phe Phe Trp Leu Ala Leu Leu Ser Gly Phe Pro Leu				
	655	660	665	670
ctc ggg acg ctg tat ttg cgg tac tga gctcagggcg cagtgcacga				2063
Leu Gly Thr Leu Tyr Leu Arg Tyr				
	675			
ctgtattcac taggcagtgc cttgcaatct gcatttcgcg cttcgctcgt tgcaagcggg				2123
tcgactttca gcgcgaagtc tgcgacggcc				2153

<210> 10
 <211> 678
 <212> PRT
 <213> Rhodosporidium toruloides

<400> 10

Met	Ala	Ser	Leu	Asp	Pro	Pro	Leu	Pro	Gly	Pro	Ala	Asn	Leu	Val	Asp
1				5					10					15	

Asp	Ala	Leu	Arg	His	Pro	Asp	Ser	Ala	Pro	Pro	Ile	Pro	Pro	Asp	Ser
			20					25						30	

Ala	Pro	Pro	Ser	His	Pro	Ser	Thr	Ala	Thr	Gln	Pro	Ser	Ala	Thr	Ser
			35				40					45			

Arg	Gly	Gln	Leu	Ser	Thr	Ala	Ser	Ser	Tyr	Ala	Ser	Asp	Val	Ser	Thr
	50					55				60					

Arg Asp Gly Thr Pro Asp Leu Ala Asn Gly Gln Gly Val Thr Thr Thr
65 70 75 80

Ile Thr Thr Val Thr Gly Lys Gly Gly Lys Ala Val Thr Gln Thr Leu
85 90 95

Thr His Val Gly Ala Ser Ser Val Asp Ala Arg Phe Ser Ser Ser Thr
100 105 110

Ser Ser Ile Thr Leu Arg Pro Ile Pro Ala Arg Gly Gly Asp Pro Lys
115 120 125

Lys Ile Lys Val Leu Arg Ser Arg Arg Thr His Phe Ala Pro Arg Thr
130 135 140

Ser His Phe Asp Arg His Asn Leu Thr Ser Ala Ser Asp Pro Phe Arg
145 150 155 160

Gly Leu Tyr Thr Leu Phe Trp Ile Val Ile Phe Val Gly Ala Leu Lys
165 170 175

Thr Val Tyr His Arg Phe Ala Glu Gln Gly Gly Trp Gly Gly Glu Trp
180 185 190

Arg Phe Ala Ala Leu Ile Ser Arg Asp Gly Trp Val Leu Ala Val Ser
195 200 205

Asp Ala Val Leu Val Ser Ala Ser Leu Leu Cys Val Pro Tyr Ala Lys
210 215 220

Leu Leu Val His Gly Trp Ile Arg Tyr His Gly Ala Gly Val Ile Ile
225 230 235 240

Gln His Ile Cys Gln Thr Leu Tyr Leu Ala Ile Ala Ile Arg Trp Thr
245 250 255

Phe His Arg Asn Trp Pro Trp Val Gln Ser Gly Phe Met Thr Leu His
260 265 270

Ala Leu Ser Met Leu Met Lys Ile His Ser Tyr Cys Ser Leu Asn Gly
275 280 285

Glu Leu Ser Glu Arg Arg Arg Gln Leu Arg Lys Asp Glu Gly Arg Leu
290 295 300

Glu Glu Val Leu Glu Glu Met Gly Gly Arg Arg Arg Ala Glu Arg Glu
305 310 315 320

Ala Arg Glu Glu Trp Glu Arg Gln Cys Gly Glu Ala Ala Arg Ala Lys
325 330 335

Glu Gly Glu Ala Gly Ser Arg Glu Gly Lys Lys Glu Glu Val Ala Ala
340 345 350

Gln Ser Ser Thr Asp Ala Ser Thr Ser Ala Leu Ser Ser Glu Asp Glu
355 360 365

Ala Ala Ala Ala Leu Leu Arg His Arg Gln Ser Thr Ala Arg Arg Arg
370 375 380

Ser Ile Ser Pro Ser Ala Ser Arg Thr His Ser Ser Ser Ala Ser Ser
385 390 395 400

Ser His Pro Ala Pro Ser Arg Ala Glu Glu Pro Gln Glu Gly Val Glu
405 410 415

Thr Leu Thr Trp His Pro Ser Asp Arg Val Ser Lys Leu Ala Ile Ala
420 425 430

Ile Cys Glu Ala Lys Asp Leu Leu Thr Ser Asn Gly Lys Lys Pro Val
435 440 445

Thr Phe Pro Glu Asn Val Thr Phe Ala Asn Phe Ile Asp Tyr Leu Leu
450 455 460

Val Pro Thr Leu Val Tyr Glu Leu Glu Tyr Pro Arg Thr Asp Ser Ile
465 470 475 480

Arg Pro Leu Tyr Ile Leu Glu Lys Thr Leu Ala Thr Phe Gly Thr Phe
485 490 495

Ser Ile Leu Val Leu Ile Val Asp Ser Phe Ile Leu Pro Val Thr Ser
500 505 510

Arg Thr Asp Thr Pro Leu Phe Gly Phe Val Leu Asp Leu Ala Leu Pro
515 520 525

Phe Thr Leu Ala Tyr Leu Leu Ile Phe Tyr Val Ile Phe Glu Gly Val
530 535 540

Cys Asn Gly Phe Ala Glu Leu Thr Arg Phe Ala Asp Arg Asn Phe Phe
545 550 555 560

Asp Asp Trp Trp Asn Ser Cys Thr Phe Asp Glu Phe Ser Arg Lys Trp
565 570 575

Asn Arg Pro Val His Ala Phe Leu Leu Arg His Val Tyr Ala Glu Thr
580 585 590

Met Ala Ser Tyr Lys Leu Ser Lys Leu Ser Ala Ala Phe Val Thr Phe
595 600 605

Leu Phe Ser Ala Cys Val His Glu Leu Val Met Ala Val Val Thr Lys
610 615 620

Lys Leu Arg Leu Tyr Leu Phe Ser Met Gln Met Ala Gln Leu Pro Leu
625 630 635 640

Ile Met Val Gly Arg Ala Lys Ile Phe Arg Lys Tyr Pro Ala Leu Gly
645 650 655

Asn Leu Phe Phe Trp Leu Ala Leu Leu Ser Gly Phe Pro Leu Leu Gly
660 665 670

Thr Leu Tyr Leu Arg Tyr
675

<210> 11
<211> 1929
<212> DNA
<213> Rhodosporidium toruloides

<220>

<221> CDS

<222> (111)..(1778)

<400> 11

acccttacac ccgtatgtcc cgacgagcag cagcacgagc tctctcagga ccactgctga 60

ccctcgcacg ctcgcgcgca ctccaaaccc gcctctttcca gcccacaacc atg tcc 116
Met Ser
1

gtc tcg acg aac gct atc ttc aag ctc ccg acc gtc aag aac gag ccc 164
Val Ser Thr Asn Ala Ile Phe Lys Leu Pro Thr Val Lys Asn Glu Pro
5 10 15

aac ttc aac tat gcg cca ggc agc ccc gag cgc gcg gct ctt aag aag 212
Asn Phe Asn Tyr Ala Pro Gly Ser Pro Glu Arg Ala Ala Leu Lys Lys
20 25 30

gcc ctc gcc gag ctc gag gcc gcc gcg ccg ttc gaa gtg ccc gcc ttc 260
Ala Leu Ala Glu Leu Glu Ala Ala Ala Pro Phe Glu Val Pro Ala Phe
35 40 45 50

gtc ggc ggc aag gag gtc aag gtc tcg ggc gag cgc acc gag cag cct 308
Val Gly Gly Lys Glu Val Lys Val Ser Gly Glu Arg Thr Glu Gln Pro
55 60 65

atg ccc cac aac cac aag aag gcg ctc tgc acc ttc gga acc tcg act 356
Met Pro His Asn His Lys Lys Ala Leu Cys Thr Phe Gly Thr Ser Thr
70 75 80

cct gag ctc gcc agc cag gct atc gag gct gcg ctc gct gcc cag cct 404
Pro Glu Leu Ala Ser Gln Ala Ile Glu Ala Ala Leu Ala Ala Gln Pro
85 90 95

gag tgg gag gct atg ccc ttc aac gac cgc gcc gcc atc ttc ttg agg 452
Glu Trp Glu Ala Met Pro Phe Asn Asp Arg Ala Ala Ile Phe Leu Arg
100 105 110

gct gcc gac ctg atc agc ggc aag tac cgt gcc aag atc tgc gct gct 500
Ala Ala Asp Leu Ile Ser Gly Lys Tyr Arg Ala Lys Ile Cys Ala Ala
115 120 125 130

acc atg ctc ggc cag ggc aag aac gtc tgg cag gcc gag att gac gcc 548
Thr Met Leu Gly Gln Gly Lys Asn Val Trp Gln Ala Glu Ile Asp Ala
135 140 145

gct gcc gag ctc gcc gac ttc ttc cgc ttc ggc gcc gct cag gtc gag 596
Ala Ala Glu Leu Ala Asp Phe Phe Arg Phe Gly Ala Ala Gln Val Glu
150 155 160

cag ctc tac gcg atg cag ccc acc gag aac tcg cct ggc gtc tgg aac 644
Gln Leu Tyr Ala Met Gln Pro Thr Glu Asn Ser Pro Gly Val Trp Asn
165 170 175

cgc acc gag tac cgc ccg ctt gag ggt ttc gtc ttt gct gtc acc ccc Arg Thr Glu Tyr Arg Pro Leu Glu Gly Phe Val Phe Ala Val Thr Pro 180 185 190	692
ttc aac ttc acc gct atc ggc ggc aac ctc gtc ggc gct ccc gcg ctc Phe Asn Phe Thr Ala Ile Gly Gly Asn Leu Val Gly Ala Pro Ala Leu 195 200 205 210	740
acc gga aac gtc ctc gtc tgg aag ccc tcg ccc atg gcg acc tac tcc Thr Gly Asn Val Leu Val Trp Lys Pro Ser Pro Met Ala Thr Tyr Ser 215 220 225	788
tcg tgg ctc gtc ttc cag atc ctc ctt gag gcc ggt ctc ccc aag aac Ser Trp Leu Val Phe Gln Ile Leu Leu Glu Ala Gly Leu Pro Lys Asn 230 235 240	836
gtc atc cag ttc ctc ccc tgc cca aac ggc gac gcg acc atc tcg ctc Val Ile Gln Phe Leu Pro Cys Pro Asn Gly Asp Ala Thr Ile Ser Leu 245 250 255	884
gtt gac cgc gtc ctc tcg cac cgc atg ttc gcc ggt ctc cac ttc act Val Asp Arg Val Leu Ser His Arg Met Phe Ala Gly Leu His Phe Thr 260 265 270	932
ggc tcg acc cac gtc ttc cgc cac ctc tgg cag aag atc ggc aac aac Gly Ser Thr His Val Phe Arg His Leu Trp Gln Lys Ile Gly Asn Asn 275 280 285 290	980
atc aac aac tac ctc agc tac ccc cgc atc gtc ggc gag acc ggc gga Ile Asn Asn Tyr Leu Ser Tyr Pro Arg Ile Val Gly Glu Thr Gly Gly 295 300 305	1028
aag aac ttc cag ctc gtc cac ccc tct gcc aac gtc cgc gct gct gtc Lys Asn Phe Gln Leu Val His Pro Ser Ala Asn Val Arg Ala Ala Val 310 315 320	1076
acc ggt gcc atc cgc ggt gcc ttc gag tac cag ggc cag aag tgc tcg Thr Gly Ala Ile Arg Gly Ala Phe Glu Tyr Gln Gly Gln Lys Cys Ser 325 330 335	1124
gct ctc tcg cgt ctt tac gtc ccg aag ggt ctg tgg gag ggc gag ggc Ala Leu Ser Arg Leu Tyr Val Pro Lys Gly Leu Trp Glu Gly Glu Gly 340 345 350	1172
aag ttc aag gag atc ctt ctc tcc gag gtt gcc aag atc act ctc ggc Lys Phe Lys Glu Ile Leu Leu Ser Glu Val Ala Lys Ile Thr Leu Gly 355 360 365 370	1220
ccc gtc acc gag ttt gag cac ttc atg ggt ccc gtc atc tcg cag gct Pro Val Thr Glu Phe Glu His Phe Met Gly Pro Val Ile Ser Gln Ala 375 380 385	1268
tcg ttc gac aag tgc ctc agc tac gtt gag aag gcc aag cag gca ggt	1316

Ser	Phe	Asp	Lys	Cys	Leu	Ser	Tyr	Val	Glu	Lys	Ala	Lys	Gln	Ala	Gly		
			390					395					400				
ggc	gag	gtc	ctc	gcc	ggc	ggc	aag	ggc	gac	gcg	tcg	agc	ggg	tac	tac		1364
Gly	Glu	Val	Leu	Ala	Gly	Gly	Lys	Gly	Asp	Ala	Ser	Ser	Gly	Tyr	Tyr		
		405					410					415					
gtc	gag	ccg	acc	atc	atc	ctg	acc	aag	gac	cct	cgc	tcg	cct	acc	atg		1412
Val	Glu	Pro	Thr	Ile	Ile	Leu	Thr	Lys	Asp	Pro	Arg	Ser	Pro	Thr	Met		
	420					425					430						
gtc	gac	gag	atc	ttc	ggc	ccg	gtc	ctc	act	gtt	tac	atc	tac	gag	gac		1460
Val	Asp	Glu	Ile	Phe	Gly	Pro	Val	Leu	Thr	Val	Tyr	Ile	Tyr	Glu	Asp		
435					440					445					450		
gac	cag	ttc	gag	gag	acg	tgc	aag	ttg	atc	gac	cag	acg	acg	acg	tac		1508
Asp	Gln	Phe	Glu	Glu	Thr	Cys	Lys	Leu	Ile	Asp	Gln	Thr	Thr	Thr	Tyr		
			455					460						465			
gcc	ctc	act	ggc	tgc	atc	ttc	tcg	gac	gac	cg	gcc	gag	act	gtc	aag		1556
Ala	Leu	Thr	Gly	Cys	Ile	Phe	Ser	Asp	Asp	Arg	Ala	Ala	Thr	Val	Lys		
			470					475					480				
gcc	ggg	gct	ctc	ctc	cg	cac	gct	gag	ggg	aac	tac	tac	atc	aac	gac		1604
Ala	Gly	Ala	Leu	Leu	Arg	His	Ala	Ala	Gly	Asn	Tyr	Tyr	Ile	Asn	Asp		
		485					490					495					
aag	tcg	acc	ggg	gct	gtt	gtc	ggg	gcc	cag	cct	ttc	ggg	ggc	gca	cg		1652
Lys	Ser	Thr	Gly	Ala	Val	Val	Gly	Ala	Gln	Pro	Phe	Gly	Gly	Ala	Arg		
	500					505					510						
gga	tcg	ggc	acg	aac	gac	aag	gag	ggc	tcg	atg	acg	ttc	ttc	acc	cg		1700
Gly	Ser	Gly	Thr	Asn	Asp	Lys	Ala	Gly	Ser	Met	Thr	Phe	Phe	Thr	Arg		
515					520					525					530		
tgg	tgc	cag	ccg	cg	agt	gtg	aag	gaa	tcc	ttc	tgc	ccg	ccc	gaa	tct		1748
Trp	Cys	Gln	Pro	Arg	Ser	Val	Lys	Glu	Ser	Phe	Cys	Pro	Pro	Glu	Ser		
			535					540						545			
ttc	cct	tac	ccg	tcg	aac	cag	cg	gat	taa	atggaggagt	tgaggaggag						1798
Phe	Pro	Tyr	Pro	Ser	Asn	Gln	Arg	Asp									
			550				555										
gaggacgtcg	agggagctgg	ggaggcggag	gacgtcgagg	aggagttggg	gaggagtttg												1858
tcgaggagga	ggagaagggt	ttctcctcgc	ctgtagttgt	acaaaatcag	cacgcctttg												1918
cttccaccgc	c																1929

<210> 12
 <211> 555
 <212> PRT
 <213> Rhodosporidium toruloides

<400> 12

Met Ser Val Ser Thr Asn Ala Ile Phe Lys Leu Pro Thr Val Lys Asn
1 5 10 15

Glu Pro Asn Phe Asn Tyr Ala Pro Gly Ser Pro Glu Arg Ala Ala Leu
20 25 30

Lys Lys Ala Leu Ala Glu Leu Glu Ala Ala Ala Pro Phe Glu Val Pro
35 40 45

Ala Phe Val Gly Gly Lys Glu Val Lys Val Ser Gly Glu Arg Thr Glu
50 55 60

Gln Pro Met Pro His Asn His Lys Lys Ala Leu Cys Thr Phe Gly Thr
65 70 75 80

Ser Thr Pro Glu Leu Ala Ser Gln Ala Ile Glu Ala Ala Leu Ala Ala
85 90 95

Gln Pro Glu Trp Glu Ala Met Pro Phe Asn Asp Arg Ala Ala Ile Phe
100 105 110

Leu Arg Ala Ala Asp Leu Ile Ser Gly Lys Tyr Arg Ala Lys Ile Cys
115 120 125

Ala Ala Thr Met Leu Gly Gln Gly Lys Asn Val Trp Gln Ala Glu Ile
130 135 140

Asp Ala Ala Ala Glu Leu Ala Asp Phe Phe Arg Phe Gly Ala Ala Gln
145 150 155 160

Val Glu Gln Leu Tyr Ala Met Gln Pro Thr Glu Asn Ser Pro Gly Val
165 170 175

Trp Asn Arg Thr Glu Tyr Arg Pro Leu Glu Gly Phe Val Phe Ala Val
180 185 190

Thr Pro Phe Asn Phe Thr Ala Ile Gly Gly Asn Leu Val Gly Ala Pro
195 200 205

Ala Leu Thr Gly Asn Val Leu Val Trp Lys Pro Ser Pro Met Ala Thr
210 215 220

Tyr Ser Ser Trp Leu Val Phe Gln Ile Leu Leu Glu Ala Gly Leu Pro
225 230 235 240

Lys Asn Val Ile Gln Phe Leu Pro Cys Pro Asn Gly Asp Ala Thr Ile
245 250 255

Ser Leu Val Asp Arg Val Leu Ser His Arg Met Phe Ala Gly Leu His
260 265 270

Phe Thr Gly Ser Thr His Val Phe Arg His Leu Trp Gln Lys Ile Gly
275 280 285

Asn Asn Ile Asn Asn Tyr Leu Ser Tyr Pro Arg Ile Val Gly Glu Thr
290 295 300

Gly Gly Lys Asn Phe Gln Leu Val His Pro Ser Ala Asn Val Arg Ala
305 310 315 320

Ala Val Thr Gly Ala Ile Arg Gly Ala Phe Glu Tyr Gln Gly Gln Lys
325 330 335

Cys Ser Ala Leu Ser Arg Leu Tyr Val Pro Lys Gly Leu Trp Glu Gly
340 345 350

Glu Gly Lys Phe Lys Glu Ile Leu Leu Ser Glu Val Ala Lys Ile Thr
355 360 365

Leu Gly Pro Val Thr Glu Phe Glu His Phe Met Gly Pro Val Ile Ser
370 375 380

Gln Ala Ser Phe Asp Lys Cys Leu Ser Tyr Val Glu Lys Ala Lys Gln
385 390 395 400

Ala Gly Gly Glu Val Leu Ala Gly Gly Lys Gly Asp Ala Ser Ser Gly
405 410 415

Tyr Tyr Val Glu Pro Thr Ile Ile Leu Thr Lys Asp Pro Arg Ser Pro

420

425

430

Thr Met Val Asp Glu Ile Phe Gly Pro Val Leu Thr Val Tyr Ile Tyr
 435 440 445

Glu Asp Asp Gln Phe Glu Glu Thr Cys Lys Leu Ile Asp Gln Thr Thr
 450 455 460

Thr Tyr Ala Leu Thr Gly Cys Ile Phe Ser Asp Asp Arg Ala Ala Thr
 465 470 475 480

Val Lys Ala Gly Ala Leu Leu Arg His Ala Ala Gly Asn Tyr Tyr Ile
 485 490 495

Asn Asp Lys Ser Thr Gly Ala Val Val Gly Ala Gln Pro Phe Gly Gly
 500 505 510

Ala Arg Gly Ser Gly Thr Asn Asp Lys Ala Gly Ser Met Thr Phe Phe
 515 520 525

Thr Arg Trp Cys Gln Pro Arg Ser Val Lys Glu Ser Phe Cys Pro Pro
 530 535 540

Glu Ser Phe Pro Tyr Pro Ser Asn Gln Arg Asp
 545 550 555

<210> 13
 <211> 1542
 <212> DNA
 <213> Marinobacter aquaeolei

<220>
 <221> CDS
 <222> (1)..(1542)

<400> 13
 atg gcc atc cag cag gtc cac cac gcc gac acc tcg tcg tcg aag gtc 48
 Met Ala Ile Gln Gln Val His His Ala Asp Thr Ser Ser Ser Lys Val
 1 5 10 15

ctc ggc cag ctc cgc ggc aag cgc gtc ctc atc acc ggc acc acc ggc 96
 Leu Gly Gln Leu Arg Gly Lys Arg Val Leu Ile Thr Gly Thr Thr Gly
 20 25 30

ttc ctc ggc aag gtc gtc ctc gag cgc ctc atc cgc gcc gtc ccg gac Phe Leu Gly Lys Val Val Leu Glu Arg Leu Ile Arg Ala Val Pro Asp 35 40 45	144
atc ggc gcc atc tac ctc ctc atc cgc ggc aac aag cgc cac ccg gac Ile Gly Ala Ile Tyr Leu Leu Ile Arg Gly Asn Lys Arg His Pro Asp 50 55 60	192
gcc cgc tcg cgc ttc ctc gag gag atc gcc acc tcg tcg gtc ttc gac Ala Arg Ser Arg Phe Leu Glu Glu Ile Ala Thr Ser Ser Val Phe Asp 65 70 75 80	240
cgc ctc cgc gag gcc gac tcg gag ggc ttc gac gcc ttc ctc gag gag Arg Leu Arg Glu Ala Asp Ser Glu Gly Phe Asp Ala Phe Leu Glu Glu 85 90 95	288
cgc atc cac tgc gtc acc ggc gag gtc acc gag gcc ggc ttc ggc atc Arg Ile His Cys Val Thr Gly Glu Val Thr Glu Ala Gly Phe Gly Ile 100 105 110	336
ggc cag gag gac tac cgc aag ctc gcc acc gag ctc gac gcc gtc atc Gly Gln Glu Asp Tyr Arg Lys Leu Ala Thr Glu Leu Asp Ala Val Ile 115 120 125	384
aac tcg gcc gcc tcg gtc aac ttc cgc gag gag ctc gac aag gcc ctc Asn Ser Ala Ala Ser Val Asn Phe Arg Glu Glu Leu Asp Lys Ala Leu 130 135 140	432
gcc atc aac acc ctc tgc ctc cgc aac atc gcc ggc atg gtc gac ctc Ala Ile Asn Thr Leu Cys Leu Arg Asn Ile Ala Gly Met Val Asp Leu 145 150 155 160	480
aac ccg aag ctc gcc gtc ctc cag gtc tcg acc tgc tac gtc aac ggc Asn Pro Lys Leu Ala Val Leu Gln Val Ser Thr Cys Tyr Val Asn Gly 165 170 175	528
atg aac tcg ggc cag gtc acc gag tcg gtc atc aag ccg gcc ggt gag Met Asn Ser Gly Gln Val Thr Glu Ser Val Ile Lys Pro Ala Gly Glu 180 185 190	576
gcg gtc ccg cgc tcg ccg gac ggc ttc tac gag atc gag gag ctc gtc Ala Val Pro Arg Ser Pro Asp Gly Phe Tyr Glu Ile Glu Glu Leu Val 195 200 205	624
cgc ctc ctc cag gac aag atc gag gac gtc cag gcc cgc tac tcg ggc Arg Leu Leu Gln Asp Lys Ile Glu Asp Val Gln Ala Arg Tyr Ser Gly 210 215 220	672
aag gtc ctc gag cgc aag ctc gtc gac ctc ggc atc cgc gag gcc aac Lys Val Leu Glu Arg Lys Leu Val Asp Leu Gly Ile Arg Glu Ala Asn 225 230 235 240	720
cgc tac ggc tgg tcg gac acc tac acc ttc acc aag tgg ctc ggc gag Arg Tyr Gly Trp Ser Asp Thr Tyr Thr Phe Thr Lys Trp Leu Gly Glu	768

245										250					255					
cag	ctc	ctt	atg	aag	gcc	ctc	aac	ggc	cgc	acc	ctc	acc	atc	ctc	cgc	816				
Gln	Leu	Leu	Met	Lys	Ala	Leu	Asn	Gly	Arg	Thr	Leu	Thr	Ile	Leu	Arg					
			260					265					270							
ccg	tcg	atc	atc	gag	tcg	gcc	ctc	gag	gag	ccg	gcc	ccc	ggc	tgg	atc	864				
Pro	Ser	Ile	Ile	Glu	Ser	Ala	Leu	Glu	Glu	Pro	Ala	Pro	Gly	Trp	Ile					
		275					280					285								
gag	ggc	gtc	aag	gtc	gcc	gac	gcc	atc	atc	ctc	gcc	tac	gcc	cgc	gag	912				
Glu	Gly	Val	Lys	Val	Ala	Asp	Ala	Ile	Ile	Leu	Ala	Tyr	Ala	Arg	Glu					
	290					295					300									
aag	gtc	acc	ctc	ttc	ccc	ggc	aag	cgc	tcg	ggc	atc	atc	gac	gtc	atc	960				
Lys	Val	Thr	Leu	Phe	Pro	Gly	Lys	Arg	Ser	Gly	Ile	Ile	Asp	Val	Ile					
305					310					315					320					
ccg	gtc	gac	ctc	gtc	gcc	aac	tcg	atc	atc	ctc	tcg	ctc	gcc	gag	gcc	1008				
Pro	Val	Asp	Leu	Val	Ala	Asn	Ser	Ile	Ile	Leu	Ser	Leu	Ala	Glu	Ala					
				325					330					335						
ctc	ggc	gag	ccg	ggc	cgc	cgc	cgc	atc	tac	cag	tgc	tgc	tcg	ggc	ggc	1056				
Leu	Gly	Glu	Pro	Gly	Arg	Arg	Arg	Ile	Tyr	Gln	Cys	Cys	Ser	Gly	Gly					
			340					345					350							
ggc	aac	ccg	atc	tcg	ctc	ggc	gag	ttc	atc	gac	cac	ctc	atg	gcc	gag	1104				
Gly	Asn	Pro	Ile	Ser	Leu	Gly	Glu	Phe	Ile	Asp	His	Leu	Met	Ala	Glu					
		355					360					365								
tcg	aag	gcc	aac	tac	gcc	gcc	tac	gac	cac	ctc	ttc	tac	cgc	cag	ccg	1152				
Ser	Lys	Ala	Asn	Tyr	Ala	Ala	Tyr	Asp	His	Leu	Phe	Tyr	Arg	Gln	Pro					
	370					375					380									
tcg	aag	ccg	ttc	ctc	gcc	gtc	aac	cgc	gcc	ctc	ttc	gac	ctc	gtc	atc	1200				
Ser	Lys	Pro	Phe	Leu	Ala	Val	Asn	Arg	Ala	Leu	Phe	Asp	Leu	Val	Ile					
385					390					395					400					
tcg	ggc	gtc	cgc	ctc	ccg	ctc	tcg	ctc	acc	gac	cgc	gtc	ctc	aag	ctc	1248				
Ser	Gly	Val	Arg	Leu	Pro	Leu	Ser	Leu	Thr	Asp	Arg	Val	Leu	Lys	Leu					
			405					410					415							
ctc	ggc	aac	tcg	cgc	gac	ctc	aag	atg	ctc	cgc	aac	ctc	gac	acc	acc	1296				
Leu	Gly	Asn	Ser	Arg	Asp	Leu	Lys	Met	Leu	Arg	Asn	Leu	Asp	Thr	Thr					
			420				425					430								
cag	tcg	ctc	gcc	acc	atc	ttc	ggc	ttc	tac	acc	gcc	ccc	gac	tac	atc	1344				
Gln	Ser	Leu	Ala	Thr	Ile	Phe	Gly	Phe	Tyr	Thr	Ala	Pro	Asp	Tyr	Ile					
		435					440					445								
ttc	cgc	aac	gac	gag	ctc	atg	gcc	ctc	gcg	aac	cgc	atg	ggc	gag	gtc	1392				
Phe	Arg	Asn	Asp	Glu	Leu	Met	Ala	Leu	Ala	Asn	Arg	Met	Gly	Glu	Val					
	450					455					460									

gac aag ggc ctc ttc ccg gtc gac gcc cgc ctc atc gac tgg gag ctc 1440
Asp Lys Gly Leu Phe Pro Val Asp Ala Arg Leu Ile Asp Trp Glu Leu
465 470 475 480

gag cgc aag gtc tac tcg ctc aag act gcc cgc cag cgc aag aag gcc 1536
Glu Arg Lys Val Tyr Ser Leu Lys Thr Ala Arg Gln Arg Lys Lys Ala
500 505 510

```
<210> 14
<211> 513
<212> PRT
<213> Marinobacter aquaeolei
```

Met Ala Ile Gln Gln Val His His Ala Asp Thr Ser Ser Ser Lys Val
1 5 10 15

Phe Leu Gly Lys Val Val Leu Glu Arg Leu Ile Arg Ala Val Pro Asp
35 40 45

Ile Gly Ala Ile Tyr Leu Leu Ile Arg Gly Asn Lys Arg His Pro Asp
50 55 60

Ala Arg Ser Arg Phe Leu Glu Glu Ile Ala Thr Ser Ser Val Phe Asp
65 70 75 80

Arg Leu Arg Glu Ala Asp Ser Glu Gly Phe Asp Ala Phe Leu Glu Glu
85 90 95

Arg Ile His Cys Val Thr Gly Glu Val Thr Glu Ala Gly Phe Gly Ile
100 105 110

Gly Gln Glu Asp Tyr Arg Lys Leu Ala Thr Glu Leu Asp Ala Val Ile
115 120 125

Asn Ser Ala Ala Ser Val Asn Phe Arg Glu Glu Leu Asp Lys Ala Leu
130 135 140

Ala Ile Asn Thr Leu Cys Leu Arg Asn Ile Ala Gly Met Val Asp Leu
145 150 155 160

Asn Pro Lys Leu Ala Val Leu Gln Val Ser Thr Cys Tyr Val Asn Gly
165 170 175

Met Asn Ser Gly Gln Val Thr Glu Ser Val Ile Lys Pro Ala Gly Glu
180 185 190

Ala Val Pro Arg Ser Pro Asp Gly Phe Tyr Glu Ile Glu Glu Leu Val
195 200 205

Arg Leu Leu Gln Asp Lys Ile Glu Asp Val Gln Ala Arg Tyr Ser Gly
210 215 220

Lys Val Leu Glu Arg Lys Leu Val Asp Leu Gly Ile Arg Glu Ala Asn
225 230 235 240

Arg Tyr Gly Trp Ser Asp Thr Tyr Thr Phe Thr Lys Trp Leu Gly Glu
245 250 255

Gln Leu Leu Met Lys Ala Leu Asn Gly Arg Thr Leu Thr Ile Leu Arg
260 265 270

Pro Ser Ile Ile Glu Ser Ala Leu Glu Glu Pro Ala Pro Gly Trp Ile
275 280 285

Glu Gly Val Lys Val Ala Asp Ala Ile Ile Leu Ala Tyr Ala Arg Glu
290 295 300

Lys Val Thr Leu Phe Pro Gly Lys Arg Ser Gly Ile Ile Asp Val Ile
305 310 315 320

Pro Val Asp Leu Val Ala Asn Ser Ile Ile Leu Ser Leu Ala Glu Ala
325 330 335

Leu Gly Glu Pro Gly Arg Arg Arg Ile Tyr Gln Cys Cys Ser Gly Gly
340 345 350

Gly Asn Pro Ile Ser Leu Gly Glu Phe Ile Asp His Leu Met Ala Glu
355 360 365

Ser Lys Ala Asn Tyr Ala Ala Tyr Asp His Leu Phe Tyr Arg Gln Pro
370 375 380

Ser Lys Pro Phe Leu Ala Val Asn Arg Ala Leu Phe Asp Leu Val Ile
385 390 395 400

Ser Gly Val Arg Leu Pro Leu Ser Leu Thr Asp Arg Val Leu Lys Leu
405 410 415

Leu Gly Asn Ser Arg Asp Leu Lys Met Leu Arg Asn Leu Asp Thr Thr
420 425 430

Gln Ser Leu Ala Thr Ile Phe Gly Phe Tyr Thr Ala Pro Asp Tyr Ile
435 440 445

Phe Arg Asn Asp Glu Leu Met Ala Leu Ala Asn Arg Met Gly Glu Val
450 455 460

Asp Lys Gly Leu Phe Pro Val Asp Ala Arg Leu Ile Asp Trp Glu Leu
465 470 475 480

Tyr Leu Arg Lys Ile His Leu Ala Gly Leu Asn Arg Tyr Ala Leu Lys
485 490 495

Glu Arg Lys Val Tyr Ser Leu Lys Thr Ala Arg Gln Arg Lys Lys Ala
500 505 510

Ala

<210> 15

<211> 517

<212> DNA

<213> Rhodosporidium toruloides

<400> 15

cacgcctctg tgactcggta cggagagaga gagcgttggg ttgggttggg gatttgtggcg	60
agcgaagggc gaccagcag ccaggagag ggagttggtc tggatgcaaa ccatgcgcat	120
ctcctctcgc agttgaatcg tttttcccg cctgccctcg ctctctcttc cttctgctct	180
ttactcgctc acgaacaaca acgagccaca cagcgtgagc acacaccgct gcactcactc	240
gctgtcacgg accgcagctc acccttatcg tctctccctc tcccaccgca cagcctcact	300
ccctctctcg ctctccctca caagcacaac acacggcaca ctcgcacgca cactcgcacg	360
caatggccac cgtcaacgag aagcagcccg ccaccgacgc gccctcgcg cagcagaccg	420
ccatccaccg cgtgcgtccc catccctccc actgtcttcc tcgtgaaacc cgctcaccg	480
ttcgcaagca cacacgcagg tgtcggacta ccccg	517

<210> 16
 <211> 1429
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 16	
ctgcagaact acgccctctc acaccaact tccgactcga ccggcggtac gagcacgacc	60
tacttctact gcctgccatc gacatccggg cgggtcgctg cctaccctgt gcgttctgcg	120
ccctccctcg tctcgggagg cagtgtctga cagaagcttt gcgcgcagta ccccgtaag	180
atgcaactct acgcaacgtt cggcacagaa gtcgccaagc tccgcgcacg gccgcctcaa	240
gctctcgcgc tgcccgcagg tgtcgtctat tacgaggcgg agaagctcga gttgccggct	300
ttgccagcgg cgggtcaaggt tgaggtggag acggagaagg cgggagtagc gggggaggac	360
aatgaggcga agggtagat ggtgctggtg gagactctta cgggtggagca ggaggagatt	420
gaattgggct cgggagtcgt gcagattgag gagtcgttgc tcgtcaagct ggaggtcagc	480
ggctgatacct tccgttcgtt gcaaggatcg tctgcatgtt tcgcttctct caatgacaca	540
acctggagag cgctcccgtc agcgagaatc gaggacattc cgagctcgt gagcaagcgg	600
aggtgcgagg ctccctcgaa agctgcgcct cttcagacgg cttgttctct cctgctctgg	660
tgggctggcc tgacatgtaa tgtgctccgc cgcaagtccg tcgtcggctc caattcgacg	720
ttgaaagggc atagcgcaag gaagaaccct ctgcggacat gcagaattac tggctcgcct	780
gctccttcgt ctactggaat aagtcctgtc tcgttaaagc cccaacgtcg tttttcgacg	840
tttgtaaggc gcaagaggtg ctatgggcta cgcaggaagc tgagaggaca tagaagtcgg	900

gggaggaacg gcgcagagcg gcagttgcgg aagcatgagg aaagcgagac ggtccagcat	960
ctgcagcgcc aatccgcaat ctcttggttg agcctgcacc ggaagcgctc gaacagtatg	1020
cgagagtcg aacgcaagta agaaagacgc accctcacac tcgcttactt cgagccatac	1080
aacggatcaa agctgcgcgt atctcggctt gtaagggccg gaaagcaacc tcggagatgg	1140
acacgtcaca tcaccaactt atcgatctcg gccgtcgacg tcgcagagag ggcgagagaa	1200
gcggtgaagg agggaaacaa cccctcgaga gcatgatccg accgaatctg cagcgcagga	1260
agccgttaca agccgcctc gagcgcaggt cgggtccagc cgggggacga aacgcgcgag	1320
gctgattcgt gagcgaagga agccgcatcg acaagttcgc tcccctttgc cctctttccc	1380
atcacccgtt ctgccttac ccgctcagaa caacaccaga tcactcaca	1429

<210> 17
 <211> 1130
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 17	
gcacgcgaag cggtagaagc aatgaagcga ggcgagagcg agagaggcag ggcttcagcc	60
atgtccagct gatcggctgt aacgtcgcgc cgggccagtc tgttgaattt gttgcgtcgc	120
ctgagcgtaa tagaagtgca gtagtctact ccgcatgccg agaacgtcga agagcgcgaa	180
gtagggagtc gagggaagcg aggggtggcaa acacagcaac gacaagcggg tccgcttcgc	240
tcaaaagctc gttgacgttg ttttgacgtt ttgaagacag tacaacagca gcaagaggcg	300
tgcaagcgt tgggtggcgag agcagcgaca aggagggagg aatgaggag tgggtggcgag	360
ggctcgcaaa cgggcgtacg cctcgaatgg agacgtgcga gtcgttcttc gacgtccgag	420
ggatgccgag cgccgagacg gagcacgcaa cgagcgagag gagagcagcc gcgcaagggtg	480
attcgagtgg cgcaagcgga ggacgacgag gagacggacg agggaggagg agggatggcg	540
agcgagcatc ggacggcggg gcgcgagaga cggcgtgagg agccgggtgt ggagagtttg	600
aggaggcgcg ggatgcgaag tggctgggtg tgcggagtga gcggtggcaa agagcgcact	660
tagagtctag agcgaggcag tagtagtaga gctgtatgaa tgaatacaaa gtgtgaatac	720
aacagtttgt aatgcgattc tgagcttgga cgtgtgcgcg cgagaggcg acttgcaagc	780
cagcgcccg cgcgtcttct tccttctgca cctcgcgtca accctcgcac ctcacaccta	840

cactcgcatt caaagtgcgt acactctccc acgacacacg gggacggcgc acaccaccgc	900
gcgtcgcttg aacggcgtcg ccacttcgag ccgtcactga cttcgtcctc gtcctccctc	960
ctctactctc ttgtactgta ctgtgtactg ggggggatag atgggcaagg aaaagggaca	1020
cgtcaacgtc gtcgttatcg gtacgttcag cgtcgtcgag gcgagtctgg cgaggaggag	1080
gacgtcgagc tgacctcgcc ccgtcctccc gcgcaggcca cgtcgactcc	1130

<210> 18
 <211> 21
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 18	
ccgccaataa cctcacctca g	21

<210> 19
 <211> 21
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 19	
ggc gatggga gcgtagaata c	21

<210> 20
 <211> 29
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 20	
aaagagtca ctactggcc tcctcgttc	29

<210> 21
 <211> 30
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 21	
aaaggatcca cggtgagagc ggagaggga	30

<210> 22
 <211> 30
 <212> DNA
 <213> Rhodosporidium toruloides

<400> 22	
tttaagcttg gaccaacgac tgcagaccat	30

<210>	23	
<211>	30	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	23	
	ttaggcctg cccaacccga gaatgagctt	30
<210>	24	
<211>	28	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	24	
	aaagagctca ttgaccctgc gtgtatgc	28
<210>	25	
<211>	30	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	25	
	aaaggatccg tcttgagtgc tccgacgaag	30
<210>	26	
<211>	21	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	26	
	tcgtcacgtt cttgttcagc g	21
<210>	27	
<211>	29	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	27	
	ttaggcctc gcctctacct cactcacgt	29
<210>	28	
<211>	30	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	28	
	aaagagctcg cggaacagga gaacaaggag	30

<210>	29	
<211>	30	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	29	
	aaaggatccg ctcacgtcaa cactcccaaa	30
<210>	30	
<211>	30	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	30	
	tttaagctta gaaccactcg accgtcttca	30
<210>	31	
<211>	29	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	31	
	tttaggcctc tatcgacctc tcccaagcc	29
<210>	32	
<211>	28	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	32	
	tttaggcctc cagatcaggg tgagtcgt	28
<210>	33	
<211>	27	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	33	
	tttaagcttc ttgtggttgt ggggcat	27
<210>	34	
<211>	29	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	34	
	aaaggatcct cactggctgc atcttctcg	29

<210>	35	
<211>	20	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	35	
	gtgtgcatg actgtgtggt	20
<210>	36	
<211>	23	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	36	
	ctgtagagga gctgcaggat caa	23
<210>	37	
<211>	23	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	37	
	cgagcttgat gaatcctttt cgt	23
<210>	38	
<211>	21	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	38	
	tgcagctccc ttctttcgct c	21
<210>	39	
<211>	21	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	39	
	atgctgtgag agacgagaac c	21
<210>	40	
<211>	21	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	40	
	ctcgaccca ctcttcttcc g	21

<210> 41
<211> 21
<212> DNA
<213> Rhodosporidium toruloides

<400> 41
ttccgatccg caaacctcgt c 21

<210> 42
<211> 21
<212> DNA
<213> Rhodosporidium toruloides

<400> 42
aaccaggaca cgatgggctt g 21

<210> 43
<211> 21
<212> DNA
<213> Rhodosporidium toruloides

<400> 43
tccttctcag tcgcgccaat g 21

<210> 44
<211> 21
<212> DNA
<213> Rhodosporidium toruloides

<400> 44
catgttcgcc ggtctccact t 21

<210> 45
<211> 21
<212> DNA
<213> Rhodosporidium toruloides

<400> 45
cctgcttggc cttctcaacg t 21

<210> 46
<211> 23
<212> DNA
<213> Rhodosporidium toruloides

<400> 46
tttccatggg ccagcaggcg acg 23

<210>	47	
<211>	29	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	47	
	tttgatatcg tccgttgcga ggaggtcag	29
<210>	48	
<211>	27	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	48	
	ttttcatgag cacagtacgg aggcgca	27
<210>	49	
<211>	32	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	49	
	tttgtttaaa caggcgtaga ggggtgggtca gc	32
<210>	50	
<211>	26	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	50	
	tttccatggc ctcgctagac ccgcca	26
<210>	51	
<211>	28	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	51	
	tttgatatct gcgccctgag ctcagtac	28
<210>	52	
<211>	29	
<212>	DNA	
<213>	Rhodospiridium toruloides	
<400>	52	
	ttttcatgac cgtctcgacg aacgctatc	29

<210> 53
<211> 30
<212> DNA
<213> Rhodosporidium toruloides

<400> 53
tttgatatct ccatttaatc gcgctggttc 30

<210> 54
<211> 45
<212> DNA
<213> Rhodosporidium toruloides

<400> 54
agattgtcgt ttccgcctt cagtttttta ctagtggacg gcttg 45

<210> 55
<211> 65
<212> DNA
<213> Rhodosporidium toruloides

<400> 55
gtcgcacccc aaacgatgct gaaggctcgc aaacatgcta attcggggga tctggatttt 60
agtac 65

<210> 56
<211> 57
<212> DNA
<213> Rhodosporidium toruloides

<400> 56
gcgagccttc agcatcgttt ggggtgcgac ctagtcacgc ctctgtgact cggtacg 57

<210> 57
<211> 65
<212> DNA
<213> Rhodosporidium toruloides

<400> 57
caatcaagat gtcgttgtgc tagtgtacgc aaacatgcta attcggggga tctggatttt 60
agtac 65

<210> 58
<211> 54
<212> DNA

<213> Rhodosporidium toruloides

<400> 58

gcgtacacta gcacaacgac atcttgattg ctagtgcacg cgaagcggta gaag 54

<210> 59

<211> 45

<212> DNA

<213> Rhodosporidium toruloides

<400> 59

cgatactctc aaggtcagct cgaattgttt aaacatgcta attcg 45

<210> 60

<211> 2596

<212> DNA

<213> Rhodosporidium toruloides

<220>

<221> CDS

<222> (308)..(2509)

<400> 60

tctgcgcgag attcgcaact ggacgctcgt cgtcgccggc gaggtcgggg acgagtggcc 60

ggcagctgct cgcttctctt cctcgctcca acacaccgca gaacgtcgac ttcgcacttg 120

gatcgacgtt ggacagtcac agacaccaga ccctcgtcat cttcatctgc ctcgacacgc 180

ccgcacaccc ggctaccctc cttgcactct cacacttcac acagcgtcgc ctagccagtc 240

gccgcgcaac cacactcgca caccagcccc ctcgaaaccg catacccca ctttcctct 300

tcacacg atg ccg act ccc ctc gcc gac acg cct acg ccc ccg acg ccc 349

Met Pro Thr Pro Leu Ala Asp Thr Pro Thr Pro Pro Thr Pro

1 5 10

atg gag acg atc gca aag gag cgc ttg aac ccg ccg ttc gac ctg cgc 397

Met Glu Thr Ile Ala Lys Glu Arg Leu Asn Pro Pro Phe Asp Leu Arg

15 20 25 30

cgc atg acc tac gcc atg ggc gac gga gag aag gag gtc aag ctg cgc 445

Arg Met Thr Tyr Ala Met Gly Asp Gly Glu Lys Glu Val Lys Leu Arg

35 40 45

gag aag cac atg atg gag atc cag cgg aat ccg ctc ttc cgc ctc gac 493

Glu Lys His Met Met Glu Ile Gln Arg Asn Pro Leu Phe Arg Leu Asp

50 55 60

gac atc cac gac ttg aca aag gac cag ctc cgc gag cgt acc atg gct 541

Asp Ile His Asp Leu Thr Lys Asp Gln Leu Arg Glu Arg Thr Met Ala

65					70					75						
aag	ttc	tca	aca	atg	gtg	tcg	tac	gta	tcg	aac	gag	cga	atc	gac	gag	589
Lys	Phe	Ser	Thr	Met	Val	Ser	Tyr	Val	Ser	Asn	Glu	Arg	Ile	Asp	Glu	
	80					85					90					
ttc	cag	aag	cga	atg	aac	gtc	gtc	tcc	ctc	gcc	gac	cct	ggg	ttc	tgg	637
Phe	Gln	Lys	Arg	Met	Asn	Val	Val	Ser	Leu	Ala	Asp	Pro	Gly	Phe	Trp	
95					100					105					110	
act	cga	ttt	ggc	gtc	cat	ttc	ggg	ttg	ttc	ctc	ggc	gcc	atc	cgg	tca	685
Thr	Arg	Phe	Gly	Val	His	Phe	Gly	Leu	Phe	Leu	Gly	Ala	Ile	Arg	Ser	
				115					120					125		
ggc	gca	acc	gcc	aac	cag	atg	agc	tac	tgg	atg	tcg	aag	ggc	ctc	ctc	733
Gly	Ala	Thr	Ala	Asn	Gln	Met	Ser	Tyr	Trp	Met	Ser	Lys	Gly	Leu	Leu	
			130					135					140			
ggc	ttg	aac	ggc	atg	gtc	ggc	tgc	ttc	ggc	atg	acg	gaa	ctc	gcg	cac	781
Gly	Leu	Asn	Gly	Met	Val	Gly	Cys	Phe	Gly	Met	Thr	Glu	Leu	Ala	His	
		145					150					155				
ggc	tca	aac	gtc	gcc	gga	ctc	gag	acg	acc	gcg	act	ttt	gac	gag	cag	829
Gly	Ser	Asn	Val	Ala	Gly	Leu	Glu	Thr	Thr	Ala	Thr	Phe	Asp	Glu	Gln	
	160					165				170						
acg	gac	gag	ttt	gtg	att	cac	acg	ccg	agc	gtc	tcg	gct	acc	aag	tgg	877
Thr	Asp	Glu	Phe	Val	Ile	His	Thr	Pro	Ser	Val	Ser	Ala	Thr	Lys	Trp	
175					180					185					190	
tgg	att	ggc	ggc	gcc	gcc	cac	tcc	gcc	acg	cac	tgc	tcc	gtc	ttc	gcc	925
Trp	Ile	Gly	Gly	Ala	Ala	His	Ser	Ala	Thr	His	Cys	Ser	Val	Phe	Ala	
				195					200					205		
caa	ctg	atc	gtc	aaa	ggc	caa	gtc	tac	ggc	acg	aag	acc	ttc	atc	gta	973
Gln	Leu	Ile	Val	Lys	Gly	Gln	Val	Tyr	Gly	Thr	Lys	Thr	Phe	Ile	Val	
			210					215					220			
ccc	ctc	cgc	gaa	ccg	aag	acg	tac	cag	ctc	ctg	ccc	gga	gtc	gcg	att	1021
Pro	Leu	Arg	Glu	Pro	Lys	Thr	Tyr	Gln	Leu	Leu	Pro	Gly	Val	Ala	Ile	
		225					230					235				
ggc	gat	att	ggg	aag	aag	atg	ggc	agg	gat	ggg	atc	gat	aat	ggc	tgg	1069
Gly	Asp	Ile	Gly	Lys	Lys	Met	Gly	Arg	Asp	Gly	Ile	Asp	Asn	Gly	Trp	
	240					245					250					
atc	caa	ttc	acc	aac	gtc	cgc	atc	ccc	cgc	gcc	tac	atg	ctc	atg	aag	1117
Ile	Gln	Phe	Thr	Asn	Val	Arg	Ile	Pro	Arg	Ala	Tyr	Met	Leu	Met	Lys	
255					260					265					270	
cac	act	caa	gtc	acg	cgc	gac	ggc	gag	gtc	cgc	gag	ccg	ccc	ctc	gca	1165
His	Thr	Gln	Val	Thr	Arg	Asp	Gly	Glu	Val	Arg	Glu	Pro	Pro	Leu	Ala	
				275					280					285		

cag ctc acg tac ggc gcg ctc ctc caa ggc cgg acg gcg atg gtt gcc Gln Leu Thr Tyr Gly Ala Leu Leu Gln Gly Arg Thr Ala Met Val Ala 290 295 300	1213
gac gct gcc aac gtc gcc aag aaa gct ttg acc att gcg att cgg tat Asp Ala Ala Asn Val Ala Lys Lys Ala Leu Thr Ile Ala Ile Arg Tyr 305 310 315	1261
gcg gcg gtc agg agg caa ttc aag gtc ggc gag aac aag ctc gag tcg Ala Ala Val Arg Arg Gln Phe Lys Val Gly Glu Asn Lys Leu Glu Ser 320 325 330	1309
cag ctc ctc gac tac ccg atc cac cag cgc cgc ctc ttg ccg ctc ctc Gln Leu Leu Asp Tyr Pro Ile His Gln Arg Arg Leu Leu Pro Leu Leu 335 340 345 350	1357
tcg cag gcc gtc gcg atg ggc ttc acc tcg tac cgc atg acc gct ctc Ser Gln Ala Val Ala Met Gly Phe Thr Ser Tyr Arg Met Thr Ala Leu 355 360 365	1405
ttc gag gaa atg agc ggc caa ctc gag tcg ctc ggg tcc gac tcg gac Phe Glu Glu Met Ser Gly Gln Leu Glu Ser Leu Gly Ser Asp Ser Asp 370 375 380	1453
gaa gcc gag act aag gag gtc ctc gag aaa ctc aag gag acg cat gcg Glu Ala Glu Thr Lys Glu Val Leu Glu Lys Leu Lys Glu Thr His Ala 385 390 395	1501
acg agt gcg ggc ctc aag gcg ttt tgc acg tgg aac gcg ctc gag acg Thr Ser Ala Gly Leu Lys Ala Phe Cys Thr Trp Asn Ala Leu Glu Thr 400 405 410	1549
atc gag aag tgt cgg gcg tcg ctc ggt gga cac ggc tac tcg gcg tac Ile Glu Lys Cys Arg Ala Ser Leu Gly Gly His Gly Tyr Ser Ala Tyr 415 420 425 430	1597
tcg ggc ttg ccg ggc atg tac gcg gac cag gcg gtt cag tgc acc tgg Ser Gly Leu Pro Gly Met Tyr Ala Asp Gln Ala Val Gln Cys Thr Trp 435 440 445	1645
gag ggc gac aac acc atc ctc acg ctc cag tcc ggc cgc tcg ctg gtc Glu Gly Asp Asn Thr Ile Leu Thr Leu Gln Ser Gly Arg Ser Leu Val 450 455 460	1693
tcg tcg tac gcc gac gcc gtc aag ggc cag aag ctt ccc ggc ggc acc Ser Ser Tyr Ala Asp Ala Val Lys Gly Gln Lys Leu Pro Gly Gly Thr 465 470 475	1741
gcg tac ctg aac ggc ctc ccc ggc gtc ctg acc gcc tca tgc ccg tcg Ala Tyr Leu Asn Gly Leu Pro Gly Val Leu Thr Ala Ser Cys Pro Ser 480 485 490	1789
gac gag gcg act ctc gac ctc aag acc ctc gag gcg ggc tgg gac acc Asp Glu Ala Thr Leu Asp Leu Lys Thr Leu Glu Ala Gly Trp Asp Thr	1837

495		500		505		510	
gtc tct gcc aac gtc gtc cga cag gcg tac gaa aag ttc gag ggc gcg	1885						
Val Ser Ala Asn Val Val Arg Gln Ala Tyr Glu Lys Phe Glu Gly Ala		515		520		525	
atg aag agt ggg aag gga agg gag gag gcg ctc gag ttg tgc tcg cag	1933						
Met Lys Ser Gly Lys Gly Arg Glu Glu Ala Leu Glu Leu Cys Ser Gln		530		535		540	
gag cgg ttc gtt gcc gcg cgg gtg cac acc gct ggc tac atg ttc cgc	1981						
Glu Arg Phe Val Ala Ala Arg Val His Thr Ala Gly Tyr Met Phe Arg		545		550		555	
atg ttc cac gag gcg ctc gtc gag ctc gcc aag gac gag cca aag gac	2029						
Met Phe His Glu Ala Leu Val Glu Leu Ala Lys Asp Glu Pro Lys Asp		560		565		570	
aac ggc gtc atc aag acg ctc gac gac atc tgc cgc ctg tac ggc tgc	2077						
Asn Gly Val Ile Lys Thr Leu Asp Asp Ile Cys Arg Leu Tyr Gly Cys		580		585		590	
tgg gct att gag gag aac gcc gcg cag ttc ctc aag tac aag ttc ttc	2125						
Trp Ala Ile Glu Glu Asn Ala Ala Gln Phe Leu Lys Tyr Lys Phe Phe		595		600		605	
acc ccc aag cag atg gac atc atc tcc aac gag gtc acg agc ttg tgc	2173						
Thr Pro Lys Gln Met Asp Ile Ile Ser Asn Glu Val Thr Ser Leu Cys		610		615		620	
gcc gag ctc cgc aag tgc gca gtc ttg ctc acc gac agc ttt ggc ttc	2221						
Ala Glu Leu Arg Lys Cys Ala Val Leu Leu Thr Asp Ser Phe Gly Phe		625		630		635	
acc gac cac atc atc aac tcg cca ttc gga cgg tac gat ggc aac gtc	2269						
Thr Asp His Ile Ile Asn Ser Pro Phe Gly Arg Tyr Asp Gly Asn Val		640		645		650	
tac gag tcg tac tac aac cag gtc aag gcg gcg aac ccg cac aac ccc	2317						
Tyr Glu Ser Tyr Tyr Asn Gln Val Lys Ala Ala Asn Pro His Asn Pro		660		665		670	
gtc gcg ccg tac ttc gag cgc gtg atc cgg cct ctg atc gag cgc gag	2365						
Val Ala Pro Tyr Phe Glu Arg Val Ile Arg Pro Leu Ile Glu Arg Glu		675		680		685	
ccg ctc gag ctc ggc gac gac gcc gac tcg atg gag ctc gac gac gag	2413						
Pro Leu Glu Leu Gly Asp Asp Ala Asp Ser Met Glu Leu Asp Asp Glu		690		695		700	
att gcc gag atc cag gct gag cgg gag gag gag aag ggc gag gct gag	2461						
Ile Ala Glu Ile Gln Ala Glu Arg Glu Glu Glu Lys Gly Glu Ala Glu		705		710		715	

acg gcg gac gag gca gag aag gag ttg aag aag cct gcc gag gag tga 2509
 Thr Ala Asp Glu Ala Glu Lys Glu Leu Lys Lys Pro Ala Glu Glu
 720 725 730

tccgcctctc gccacgcatg tgtccctata ccgtgtacta gtgtctattc ctcgagctgc 2569

aactgtcgac aacaaagtat tccatcc 2596

<210> 61

<211> 733

<212> PRT

<213> Rhodosporidium toruloides

<400> 61

Met Pro Thr Pro Leu Ala Asp Thr Pro Thr Pro Pro Thr Pro Met Glu
 1 5 10 15

Thr Ile Ala Lys Glu Arg Leu Asn Pro Pro Phe Asp Leu Arg Arg Met
 20 25 30

Thr Tyr Ala Met Gly Asp Gly Glu Lys Glu Val Lys Leu Arg Glu Lys
 35 40 45

His Met Met Glu Ile Gln Arg Asn Pro Leu Phe Arg Leu Asp Asp Ile
 50 55 60

His Asp Leu Thr Lys Asp Gln Leu Arg Glu Arg Thr Met Ala Lys Phe
 65 70 75 80

Ser Thr Met Val Ser Tyr Val Ser Asn Glu Arg Ile Asp Glu Phe Gln
 85 90 95

Lys Arg Met Asn Val Val Ser Leu Ala Asp Pro Gly Phe Trp Thr Arg
 100 105 110

Phe Gly Val His Phe Gly Leu Phe Leu Gly Ala Ile Arg Ser Gly Ala
 115 120 125

Thr Ala Asn Gln Met Ser Tyr Trp Met Ser Lys Gly Leu Leu Gly Leu
 130 135 140

Asn Gly Met Val Gly Cys Phe Gly Met Thr Glu Leu Ala His Gly Ser
 145 150 155 160

Asn Val Ala Gly Leu Glu Thr Thr Ala Thr Phe Asp Glu Gln Thr Asp
165 170 175

Glu Phe Val Ile His Thr Pro Ser Val Ser Ala Thr Lys Trp Trp Ile
180 185 190

Gly Gly Ala Ala His Ser Ala Thr His Cys Ser Val Phe Ala Gln Leu
195 200 205

Ile Val Lys Gly Gln Val Tyr Gly Thr Lys Thr Phe Ile Val Pro Leu
210 215 220

Arg Glu Pro Lys Thr Tyr Gln Leu Leu Pro Gly Val Ala Ile Gly Asp
225 230 235 240

Ile Gly Lys Lys Met Gly Arg Asp Gly Ile Asp Asn Gly Trp Ile Gln
245 250 255

Phe Thr Asn Val Arg Ile Pro Arg Ala Tyr Met Leu Met Lys His Thr
260 265 270

Gln Val Thr Arg Asp Gly Glu Val Arg Glu Pro Pro Leu Ala Gln Leu
275 280 285

Thr Tyr Gly Ala Leu Leu Gln Gly Arg Thr Ala Met Val Ala Asp Ala
290 295 300

Ala Asn Val Ala Lys Lys Ala Leu Thr Ile Ala Ile Arg Tyr Ala Ala
305 310 315 320

Val Arg Arg Gln Phe Lys Val Gly Glu Asn Lys Leu Glu Ser Gln Leu
325 330 335

Leu Asp Tyr Pro Ile His Gln Arg Arg Leu Leu Pro Leu Leu Ser Gln
340 345 350

Ala Val Ala Met Gly Phe Thr Ser Tyr Arg Met Thr Ala Leu Phe Glu
355 360 365

Glu Met Ser Gly Gln Leu Glu Ser Leu Gly Ser Asp Ser Asp Glu Ala
370 375 380

Glu Thr Lys Glu Val Leu Glu Lys Leu Lys Glu Thr His Ala Thr Ser
385 390 395 400

Ala Gly Leu Lys Ala Phe Cys Thr Trp Asn Ala Leu Glu Thr Ile Glu
405 410 415

Lys Cys Arg Ala Ser Leu Gly Gly His Gly Tyr Ser Ala Tyr Ser Gly
420 425 430

Leu Pro Gly Met Tyr Ala Asp Gln Ala Val Gln Cys Thr Trp Glu Gly
435 440 445

Asp Asn Thr Ile Leu Thr Leu Gln Ser Gly Arg Ser Leu Val Ser Ser
450 455 460

Tyr Ala Asp Ala Val Lys Gly Gln Lys Leu Pro Gly Gly Thr Ala Tyr
465 470 475 480

Leu Asn Gly Leu Pro Gly Val Leu Thr Ala Ser Cys Pro Ser Asp Glu
485 490 495

Ala Thr Leu Asp Leu Lys Thr Leu Glu Ala Gly Trp Asp Thr Val Ser
500 505 510

Ala Asn Val Val Arg Gln Ala Tyr Glu Lys Phe Glu Gly Ala Met Lys
515 520 525

Ser Gly Lys Gly Arg Glu Glu Ala Leu Glu Leu Cys Ser Gln Glu Arg
530 535 540

Phe Val Ala Ala Arg Val His Thr Ala Gly Tyr Met Phe Arg Met Phe
545 550 555 560

His Glu Ala Leu Val Glu Leu Ala Lys Asp Glu Pro Lys Asp Asn Gly
565 570 575

Val Ile Lys Thr Leu Asp Asp Ile Cys Arg Leu Tyr Gly Cys Trp Ala
580 585 590

Ile Glu Glu Asn Ala Ala Gln Phe Leu Lys Tyr Lys Phe Phe Thr Pro
595 600 605

Lys Gln Met Asp Ile Ile Ser Asn Glu Val Thr Ser Leu Cys Ala Glu
610 615 620

Leu Arg Lys Cys Ala Val Leu Leu Thr Asp Ser Phe Gly Phe Thr Asp
625 630 635 640

His Ile Ile Asn Ser Pro Phe Gly Arg Tyr Asp Gly Asn Val Tyr Glu
645 650 655

Ser Tyr Tyr Asn Gln Val Lys Ala Ala Asn Pro His Asn Pro Val Ala
660 665 670

Pro Tyr Phe Glu Arg Val Ile Arg Pro Leu Ile Glu Arg Glu Pro Leu
675 680 685

Glu Leu Gly Asp Asp Ala Asp Ser Met Glu Leu Asp Asp Glu Ile Ala
690 695 700

Glu Ile Gln Ala Glu Arg Glu Glu Glu Lys Gly Glu Ala Glu Thr Ala
705 710 715 720

Asp Glu Ala Glu Lys Glu Leu Lys Lys Pro Ala Glu Glu
725 730

<210> 62
<211> 3327
<212> DNA
<213> Rhodosporidium toruloides

<400> 62
atgccgactc ccctcgccga cacgcctacg cccccgacgc ccatggagac gatcgcaaag 60
gagcgcttga acccgccgtt cgacctgcg cgcatgacct acgcatggg cgacggagag 120
aagggtgctg agcggcttac ttcttctcgc cggctttctc tttccctccg aagcttccgc 180
tttgaggaca gcacgtgac attgtgtgca cacgcagagg tcaagctgcg cgagaagcac 240
atgatggaga tccagcgga tccgctcttc cgcctcgacg acatccacga cttgacaaag 300

gaccagctcc gcgagcgtag catggctaag ttgtgcgttc agctcattag actatgcgaa	360
caggcgaaaag cgaggggaccg ccggggcgaga cggtcctgcg tcttttcgct cactcttcag	420
gacttcccgg acttcgtgcg ctgtcgctga ctttcctccg cacatataca gctcaacaat	480
ggtgtcgtag gtatcgaacg agcgaatcga cgagttccag aaggtgcgcc cctcgtcctc	540
ctcgatttcc tcccgtcgcg cccggctgat cctctcctcc tcgcagcgaa tgaacgtcgt	600
ctccctcgcc gaccctgggt tctggactcg atttgtgagt gcgcactcat tccacctcgc	660
aggaagtcga agctgaccct cttacgtttc gcagggcggt acgttgacct caggctgctt	720
cctggagacc gggcgatgat gctgactgct cttgtaatat ttccaagtcc atttcggttt	780
gtgggagtct cttttttatc gctctctcaa gctcccggga cttaccagtc gttttcggat	840
tcacagtcct cggcgccatc cggtcaggcg caaccgcaa ccagatggtc cgtcggctct	900
cattctctac tgcacggtac taaactgacg ctgtccgtcc tgcagagcta ctggatgtcg	960
aagggcctcc tcggcttgaa cggcatggtc ggctgcttcg gcatggtgcg cccgctcttc	1020
ctccccctcc tctcttcctt catccctatc agagctcgcg ctaaccctc atgcgcagac	1080
ggaactcgcg cacggctcaa acgtcgccgg actcgagacg accgcgactt ttgacgagca	1140
gacggacgag tttgtgattc acacgccgag cgtctcggct accaagtggg ggattggcgg	1200
tgcgttcgct tccgcctacc tttaggtgtc gaagagacgg tcgaactgac cttcgcaccc	1260
tcgcttccgc gcgcaggcgc cggccactcc gccacgcact gctccgtctt cgcccaactg	1320
atcgtcaaag gccaaagtcta cggcacgaag accttcacg taccctccg cgaaccgaag	1380
acgtaccagc tcctgcccgg agtcgcgatt ggcgatattg ggaagaagat gggcagggat	1440
gggatcgata atggctggat gtgagtcgcc aactcgctcc attctagagc tgtgaaggat	1500
gcgacgctgc ggactgcacg gcgctgggtg atttcgatgt ggcgacctct tcgctgacct	1560
tcggtctcca cgcacagcca attcaccaac gtccgcatcc cccgcgccta catgctcatg	1620
aagcacactc aagtcacgcg cgacggcgag gtccgcgagc cggccctcgc acagctcacg	1680
tacggcgcgc tcctccaagg ccggacggcg atggttgccg acgctgcaa cgtcgccaag	1740
aaagctttga ccattgcgat tcggtatgcg gcggtcagga ggcaattcaa ggtcggcgag	1800
aacaaggtga gtccgcgcga acactgtggc ggaggtcttg cgctgacatt cacacgatcc	1860
cacagctcga gtcgcagctc ctcgactacc cgatccacca gcgccgcctc ttgccgtcc	1920

tctcgaggc cgtcgcgatg ggcttcacct cgtaccgcat gaccgctctc ttcgaggaaa	1980
tgagcggcca actcgagtcg ctccgggtccg actcggacga agccgagact aaggaggtcc	2040
tcgagaaact caaggagacg catgcgacga gtgcgggcct caaggcgttt tgcacgtgga	2100
acgcgctcga gacgatcgag aagtgtcggg cgtcgctcgg tggacacggc tactcggcgt	2160
actcgggctt gccgggcatg tacgcggacc aggcggttca gtgcacctgg gagggcgaca	2220
acgtgcgtgc ccgcggcttt ccctcgagcg actccgtcga aacttggaca ctcacgtccc	2280
ccgcccgcag accatcctca cgctccagtc cggccgctcg ctggtctcgt cgtacgccga	2340
cgccgtcaag ggccagaagc ttcccggcgg caccgcgtac ctgaacggcc tcccggcgt	2400
cctgaccgcc tcatgcccgt cggacgaggc gactctcgac ctcaagacc tcgaggcggg	2460
ctgggacacc gtctctgcca acgtcgtccg acaggcgtac gaaaagtctg agggcgcgat	2520
gaagagtggg aagggaaggg aggaggcgct cgagttgtgc tcgcaggagc ggttcgttgc	2580
cgcgcggtg cacaccgctg gctacatgtt ccggtgcgtc gcgctccttc acaccctctc	2640
tagctctcga ctgacgcacc ccgacgcaca gcatgttcca cgaggcgctc gtcgagctcg	2700
ccaaggacga gccaaaggac aacggcgtca tcaagacgct cgacgacatc tgccgcctgt	2760
acggctgctg ggctattgag gagaacgccg cgagttcct caagtacaag ttcttcaccc	2820
ccaagcagat ggacatcatc tccaacgagg tcggtcgcgg agacttttcg tcaatcacga	2880
ggcagtcgct cacgcggttt cgtgcgcgtg caggtcacga gcttgtgcgc cgagctccgc	2940
aagtgcgcag tcttgctcac cgacagcttt ggcttcaccg accacatcgt gcgtcgcga	3000
atgtctctcg tctacgcgt tcccaagctg acgtctcgcc tctcacggca gatcaactcg	3060
ccattcggac ggtacgatgg caacgtctac gagtctgact acaaccaggt caaggcggcg	3120
aaccgcaca accccgtcgc gccgtacttc gagcgcgtga tccggcctct gatcgagcgc	3180
gagccgctcg agctcggcga cgacccgac tcgatggagc tcgacgacga gattgccgag	3240
atccaggctg agcgggagga ggagaaggc gaggtgaga cggcggacga ggcagagaag	3300
gagttgaaga agcctgccga ggagtga	3327