

<110> CJ CheilJedang Corporation
 <120> NOVEL URIDINE DIPHOSPHATE GLYCOSYLTRANSFERASE AND USE THEREOF
 <130> OPA21401
 <150> KR 10-2021-0139474
 <151> 2021-10-19
 <160> 6
 <170> KoPatentIn 3.0
 <210> 1
 <211> 463
 <212> PRT
 <213> Unknown
 <220>
 <223> UGT-B

<400> 1
 Met Asp Gly Asp Gly Asn Ser Ser Ser Ser Ser Ser Pro Leu His Val
 1 5 10 15
 Val Ile Cys Pro Trp Leu Ala Leu Gly His Leu Leu Pro Cys Leu Asp
 20 25 30
 Ile Ala Glu Arg Leu Ala Ser Arg Gly His Arg Val Ser Phe Val Ser
 35 40 45
 Thr Pro Arg Asn Ile Ala Arg Leu Pro Pro Leu Arg Pro Ala Val Ala
 50 55 60
 Pro Leu Val Glu Phe Val Ala Leu Pro Leu Pro His Val Asp Gly Leu
 65 70 75 80
 Pro Glu Gly Ala Glu Ser Thr Asn Asp Val Pro Tyr Asp Lys Phe Glu
 85 90 95
 Leu His Arg Lys Ala Phe Asp Gly Leu Ala Ala Pro Phe Ser Glu Phe
 100 105 110
 Leu Arg Ala Ala Cys Ala Glu Gly Ala Gly Ser Arg Pro Asp Trp Leu
 115 120 125
 Ile Val Asp Thr Phe His His Trp Ala Ala Ala Ala Ala Val Glu Asn
 130 135 140

Lys Val Pro Cys Val Met Leu Leu Leu Gly Ala Ala Thr Val Ile Ala
 145 150 155 160
 Gly Phe Ala Arg Gly Val Ser Glu His Ala Ala Ala Val Gly Lys
 165 170 175
 Glu Arg Pro Ala Ala Glu Ala Pro Ser Phe Glu Thr Glu Arg Arg Lys
 180 185 190
 Leu Met Thr Thr Gln Asn Ala Ser Gly Met Thr Val Ala Glu Arg Tyr
 195 200 205
 Phe Leu Thr Leu Met Arg Ser Asp Leu Val Ala Ile Arg Ser Cys Ala
 210 215 220
 Glu Trp Glu Pro Glu Ser Val Ala Ala Leu Thr Thr Leu Ala Gly Lys
 225 230 235 240
 Pro Val Val Pro Leu Gly Leu Leu Pro Pro Ser Pro Glu Gly Gly Arg
 245 250 255
 Gly Val Ser Lys Glu Asp Ala Thr Val Arg Trp Leu Asp Ala Gln Pro
 260 265 270
 Thr Lys Ser Val Val Tyr Val Ala Leu Gly Ser Glu Val Pro Leu Gly
 275 280 285
 Ala Lys Glu Val His Glu Leu Ala Leu Gly Leu Glu Leu Ala Gly Thr
 290 295 300
 Arg Phe Leu Trp Ser Leu Arg Lys Pro Ser Gly Val Ser Asp Ala Asp
 305 310 315 320
 Ile Leu Pro Ser Gly Phe Glu Glu Arg Thr Arg Gly Arg Gly Leu Val
 325 330 335
 Thr Met Gly Trp Val Pro Gln Ile Ser Val Leu Ala His Gly Ala Val
 340 345 350
 Gly Ala Phe Leu Thr His Cys Gly Trp Asn Ser Ile Ile Glu Gly Leu
 355 360 365
 Gln Phe Gly His Pro Leu Val Met Leu Pro Ile Phe Gly Asp Gln Gly
 370 375 380
 Pro Asn Ala Arg Met Met Glu Gly Arg Lys Val Gly Val Gln Val Pro
 385 390 395 400
 Arg Asp Glu Ser Asn Gly Ser Phe Asp Arg Glu Gly Val Ala Thr Thr
 405 410 415

Val Arg Ala Val Ala Val Glu Glu Glu Gly Asn Arg Ile Phe Thr Ala
420 425 430

Asn Ala Lys Lys Met Gln Glu Ile Val Ala Asp Lys Gly Cys His Asp
435 440 445

Lys Tyr Val Asp Lys Phe Ile Gln Lys Leu Arg Ser Tyr Met Glu
450 455 460

<210> 2
<211> 458
<212> PRT
<213> Unknown

<220>
<223> UGT-A

<400> 2
Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile
1 5 10 15

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu
20 25 30

Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr
35 40 45

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg
50 55 60

Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro
65 70 75 80

Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His
85 90 95

Gly Ala Asp Glu Leu Arg Arg Glu Leu Glu Leu Leu Met Leu Ala Ser
100 105 110

Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr
115 120 125

Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu
130 135 140

Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln
 145 150 155 160
 Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu
 165 170 175
 Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser
 180 185 190
 Ala Tyr Ser Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile
 195 200 205
 Lys Gln Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu
 210 215 220
 Leu Glu Glu Ser Glu Leu Glu Thr Val Ile Arg Glu Ile Pro Ala Pro
 225 230 235 240
 Ser Phe Leu Ile Pro Leu Pro Lys His Leu Thr Ala Ser Ser Ser Ser
 245 250 255
 Leu Leu Asp His Asp Arg Thr Val Phe Gln Trp Leu Asp Gln Gln Pro
 260 265 270
 Pro Ser Ser Val Leu Tyr Val Ser Phe Gly Ser Thr Ser Glu Val Asp
 275 280 285
 Glu Lys Asp Phe Leu Glu Ile Ala Arg Gly Leu Val Asp Ser Lys Gln
 290 295 300
 Ser Phe Leu Trp Val Val Arg Pro Gly Phe Val Lys Gly Ser Thr Trp
 305 310 315 320
 Val Glu Pro Leu Pro Asp Gly Phe Leu Gly Glu Arg Gly Arg Ile Val
 325 330 335
 Lys Trp Val Pro Gln Gln Glu Val Leu Ala His Gly Ala Ile Gly Ala
 340 345 350
 Phe Trp Thr His Ser Gly Trp Asn Ser Thr Leu Glu Ser Val Cys Glu
 355 360 365
 Gly Val Pro Met Ile Phe Ser Asp Phe Gly Leu Asp Gln Pro Leu Asn
 370 375 380
 Ala Arg Tyr Met Ser Asp Val Leu Lys Val Gly Val Tyr Leu Glu Asn
 385 390 395 400
 Gly Trp Glu Arg Gly Glu Ile Ala Asn Ala Ile Arg Arg Val Met Val
 405 410 415

Asp Glu Glu Gly Glu Tyr Ile Arg Gln Asn Ala Arg Val Leu Lys Gln
420 425 430

Lys Ala Asp Val Ser Leu Met Lys Gly Gly Ser Ser Tyr Glu Ser Leu
435 440 445

Glu Ser Leu Val Ser Tyr Ile Ser Ser Leu
450 455

<210> 3
<211> 808
<212> PRT
<213> Unknown

<220>
<223> Sucrose synthase

<400> 3
Met Ala Ala Lys Leu Ala Arg Leu His Ser Leu Arg Glu Arg Leu Gly
1 5 10 15

Ala Thr Phe Ser Ser His Pro Asn Glu Leu Ile Ala Leu Phe Ser Arg
20 25 30

Tyr Val Asn Gln Gly Lys Gly Met Leu Gln Arg His Gln Leu Leu Ala
35 40 45

Glu Phe Asp Ala Leu Ile Glu Ala Asp Lys Glu Lys Tyr Ala Pro Phe
50 55 60

Glu Asp Ile Leu Arg Ala Ala Gln Glu Ala Ile Val Leu Pro Pro Trp
65 70 75 80

Val Ala Leu Ala Ile Arg Pro Arg Pro Gly Val Trp Asp Tyr Ile Arg
85 90 95

Val Asn Val Ser Glu Leu Ala Val Glu Glu Leu Ser Val Ser Glu Tyr
100 105 110

Leu Ala Phe Lys Glu Gln Leu Val Asp Gly His Thr Asn Ser Asn Phe
115 120 125

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

Ser Met Ser Lys Ser Ile Gly Asn Gly Val Gln Phe Leu Asn Arg His
145 150 155 160

Leu Ser Ser Lys Leu Phe Gln Asp Lys Glu Ser Leu Tyr Pro Leu Leu
 165 170 175
 Asn Phe Leu Lys Ala His Asn His Lys Gly Thr Thr Met Met Leu Asn
 180 185 190
 Asp Arg Ile Gln Ser Leu Arg Gly Leu Gln Ser Ser Leu Arg Lys Ala
 195 200 205
 Glu Glu Tyr Leu Met Gly Ile Pro Gln Asp Thr Pro Tyr Ser Glu Phe
 210 215 220
 Asn His Arg Phe Gln Glu Leu Gly Leu Glu Lys Gly Trp Gly Asp Cys
 225 230 235 240
 Ala Lys Arg Val Leu Asp Thr Ile His Leu Leu Leu Asp Leu Leu Glu
 245 250 255
 Ala Pro Asp Pro Ala Asn Leu Glu Lys Phe Leu Gly Thr Ile Pro Met
 260 265 270
 Met Phe Asn Val Val Ile Leu Ser Pro His Gly Tyr Phe Ala Gln Ser
 275 280 285
 Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln Val Val Tyr Ile Leu
 290 295 300
 Asp Gln Val Arg Ala Leu Glu Asn Glu Met Leu Leu Arg Ile Lys Gln
 305 310 315 320
 Gln Gly Leu Asp Ile Thr Pro Lys Ile Leu Ile Val Thr Arg Leu Leu
 325 330 335
 Pro Asp Ala Val Gly Thr Thr Cys Gly Gln Arg Val Glu Lys Val Ile
 340 345 350
 Gly Thr Glu His Thr Asp Ile Leu Arg Val Pro Phe Arg Ser Glu Asn
 355 360 365
 Gly Ile Leu Arg Lys Trp Ile Ser Arg Phe Asp Val Trp Pro Phe Leu
 370 375 380
 Glu Thr Tyr Thr Glu Asp Val Ala Asn Glu Ile Met Arg Glu Met Gln
 385 390 395 400
 Ala Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser Asp Gly Asn Leu Val
 405 410 415
 Ala Thr Leu Leu Ala His Lys Leu Gly Val Thr Gln Cys Thr Ile Ala

Thr Cys His Gly Gly Pro Ala Glu Ile Ile Val Asp Gly Val Ser Gly
 690 695 700
 Leu His Ile Asp Pro Tyr His Ser Asp Lys Ala Ala Asp Ile Leu Val
 705 710 715 720
 Asn Phe Phe Glu Lys Cys Lys Gln Asp Ser Thr Tyr Trp Asp Asn Ile
 725 730 735
 Ser Gln Gly Gly Leu Gln Arg Ile Tyr Glu Lys Tyr Thr Trp Lys Leu
 740 745 750
 Tyr Ser Glu Arg Leu Met Thr Leu Thr Gly Val Tyr Gly Phe Trp Lys
 755 760 765
 Tyr Val Ser Asn Leu Glu Arg Arg Glu Thr Arg Arg Tyr Ile Glu Met
 770 775 780
 Phe Tyr Ala Leu Lys Tyr Arg Ser Leu Ala Ser Ala Val Pro Leu Ala
 785 790 795 800
 Val Asp Gly Glu Ser Thr Ser Lys
 805

<210> 4
 <211> 1392
 <212> DNA
 <213> Unknown

<220>
 <223> UGT-B

<400> 4
 atggatggtg acggcaactc ctccagttcg agtagtccgc tgcattgtagt tatatgcccg 60
 tggtttagcct tgggccattt actgccgtgt ctggatattg ccgaacgcct ggcgagccgc 120
 ggccaccgtg tctcatttgt ctccacgccg cgcaacatcg cgcgcttgcc gccgctgcgc 180
 cccgctgtgg cgccgctcgt cgaatttgtc gccctgcctc taccgcatgt cgatggtttg 240
 cctgaaggcg cagaaagtac aaacgatgtc ccatacgaca aattcgagtt acatcgcaaa 300
 gcctttgatg gcctcgctgc acccttttca gagtttctgc gtgcagcttg tgccgagggc 360
 gctggatctc ggcccgattg gcttatagtg gatacctttc accactgggc agccgcggct 420
 gccgtcgaat ataaagttcc gtgcgttatg cttttgctgg gagctgcgac agtgatcgca 480

ggtttcgctc gaggtgtgtc tgagcacgcc gctgcagccg tcgggaaaga gcgaccggca	540
gcggaagcgc catcttttga gactgaaagg aggaagctga tgaccacgca gaatgcaagc	600
gggatgacgg tagccgaacg ctacttctta acgttaatgc gtagcgatct cgtggccatt	660
cggagctgtg ctgagtggga gccagagagc gtcgccgctg taaccacctt agcgggcaag	720
ccagtcgtcc ctttgggttt gctaccaccg tcgcccgaag gaggtcgcgg tgtgagcaag	780
gaggatgcaa ccgtgcgttg gctcgatgcc cagccgacca aatcagtggg ttatgtcgcg	840
ctaggctcgg aagtcccact gggagccaag gaagtacatg aacttgcctt aggccttgag	900
ttagcaggga cacgcttcct ttggtctctg cgcaaaccct ctggcgtaag tgatgcagac	960
atcctcccct cgggttttga agagcggacc cgcgcccgctg gtcttgtgac gatgggctgg	1020
gttcctcaga ttagcgtact ggcacacggg gcagtaggtg cattcctgac tcattgtgga	1080
tggaactcta tcatagaggg gttacaattt gggcatcctt tagtgatgtt gcctatcttt	1140
ggtgaccaag gtcccaatgc gcgtatgatg gaggggagaa aagttggagt gcaggtgcca	1200
agagatgaat ccaatggaag tttcgacaga gaaggcgtcg ctacaaccgt tcgggcggtc	1260
gctgtcgagg aagaaggtaa tagaattttc actgctaatt ccaaaaaaat gcaagagatc	1320
gttgcggaca aaggatgcca tgacaagtat gtcgataaat ttattcagaa actgagatct	1380
tatatggagt ga	1392

<210> 5
 <211> 1377
 <212> DNA
 <213> Unknown

<220>
 <223> UGT-A

<400> 5	
atggaaaata aaacgggagac caccgttcgc cggcgccgga gaataatatt attcccggta	60
ccatttcaag gccacattaa cccaattctt cagctagcca atgtgttgta ctctaaagga	120
ttcagtatca ccatctttca caccaacttc aacaaacca aaacatctaa ttaccctcac	180

ttcactttca gattcatcct cgacaacgac ccacaagacg aacgcatttc caatctaccg	240
actcatggtc cgctcgctgg tatgcggatt ccgattatca acgaacacgg agctgacgaa	300
ttacgacgcg aactggaact gttgatgtta gcttctgaag aagatgaaga ggtatcgtgt	360
ttaatcacgg atgctctttg gtacttcgcg caatctgttg ctgacagtct taacctccga	420
cggcttgttt tgatgacaag cagcttgttt aattttcatg cacatgtttc acttcctcag	480
tttgatgagc ttggttacct cgatcctgat gacaaaaccc gtttgggaaga acaagcgagt	540
gggtttccta tgctaaaagt gaaagacatc aagtctgcgt attcgaactg gcaaatactc	600
aaagagatat tagggaagat gataaaacaa acaaaagcat cttcaggagt catctggaac	660
tcatttaagg aactcgaaga gtctgagctc gaaactgtta tccgtgagat cccggctcca	720
agtttcttga taccactccc caagcatttg acagcctctt ccagcagctt actagaccac	780
gatcgaaccg tttttcaatg gttagaccaa caaccgcaa gttcgggtact gtatgttagt	840
tttggtagta ctagtgaagt ggatgagaaa gatttcttgg aaatagctcg tgggttggtt	900
gatagcaagc agtcgttttt atgggtgggt cgacctgggt ttgtcaaggg ttcgacgtgg	960
gtcgaaccgt tgccagatgg gttcttgggt gaaagaggac gtattgtgaa atgggttcca	1020
cagcaagaag tgctagctca tggagcaata ggcgcattct ggactcatag cggatggaac	1080
tctacgttgg aaagcgtttg tgaaggtgtt cctatgattt tctcggattt tgggctcgat	1140
caaccgttga atgctagata catgagtgat gttttgaagg taggggtgta tttggaaaat	1200
gggtgggaaa gaggagagat agcaaatgca ataagaagag ttatggtgga tgaagaagga	1260
gaatacatta gacagaatgc aagagttttg aaacaaaagg cagatgtttc tttgatgaag	1320
ggtggttcgt cttacgaatc attagagtct ctagtttctt acatttcatc gttgtaa	1377

<210> 6
 <211> 2427
 <212> DNA
 <213> Unknown

<220>
 <223> Sucrose synthase

<400> 6

atggctgcca agctagctcg cctccacagt ctccgcgaac gcctcgggtgc caccttctcg	60
tctcatccca atgagttgat tgcactcttc tctaggtatg ttaaccaggg aaagggaatg	120
ctccagcgtc accagctgct tgcggagttc gatgccttga tcgaagctga caaagagaaa	180
tatgctccct ttgaagacat tctccgggct gctcaggaag ccattgtgct gccgccctgg	240
gttgactgg ccatcaggcc aaggcctggg gtctgggact acattcgggt gaatgtaagt	300
gagttggcag tggaagagct gagtgtttct gagtacttgg cattcaagga acagcttggt	360
gatggacaca ccaacagcaa ctttgttctt gagcttgatt ttgagccctt caatgcctcc	420
ttcccgcgcc cgtccatgtc caagtccatc ggaaatgggg tgcaagttcct taaccgtcac	480
ctgtcgtcca agttgttcca ggacaaggag agcctctacc ccttgctgaa cttcctgaaa	540
gcccataacc acaagggcac gacaatgatg ctgaatgaca gaattcagag ccttcgtggg	600
ctccaatcat cccttagaaa ggcaagaaga tatctgatgg gcattcctca agacacgccc	660
tactcggagt tcaaccacag gttccaagag ctcggtttgg agaagggttg gggtgactgt	720
gcaaagcgtg tgcttgacac catccacttg cttcttgacc ttcttgaggc ccctgatccg	780
gccaacttgg agaagttcct tggaactatt ccaatgatgt tcaatgttgt tatcctgtct	840
ccgcatggat actttgccc atccaatgtg ttgggatacc ctgatactgg tggtcaggtt	900
gtgtacattt tggaccaagt ccgcgctttg gagaatgaga tgcttttgag gatcaagcag	960
caaggccttg atatcacacc taagatcctc attgtaacca ggctgttgcc tgatgctggt	1020
gggtactacat gcggccagcg tgtggagaag gttattggaa ctgagcacac tgacattctt	1080
cgtgttccat tcaggagtga gaatggtatc ctccgcaagt ggatctcccg ttttgatgtc	1140
tggccattcc tggaacata cactgaggat gttgcaaacg aaattatgag ggaaatgcaa	1200
gcaaacctg atctcatcat tggcaattac agtgatggaa accttgttgc cactctgctg	1260
gctcaciaat taggagttac ccagtgtacc attgctcatg ccttgagaga aaccaatac	1320
cccaactcag acatatactt ggacaagttt gacagccagt accacttctc atgccaattc	1380
actgctgatc ttatcgccat gaatcacact gatttcatca tcaccagtac attccaagaa	1440
attgctggaa gcaaggacac tgtggggcag tatgaatcac acattgcatt cacccttctt	1500

gggctttacc gagttgtgca tggcatagat gtttttgatc ccaagttcaa cattgtctct	1560
cctggagctg acatgagtgt ctacttcccg tacaccgagg ctgacaagag gctcactgct	1620
ttccaccctg aaattgagga gcttctctac agtgaagtcg agaacgatga acacaagttt	1680
gtattgaagg acaagaacaa gccaatcatc ttctccatgg ctctgtcttga ccgagtgaag	1740
aacatgacag gtctgggtga gatgtatggt aagaatgcac atctcaggga tttggcaaac	1800
cttgtgattg tttgtgggtga ccacggcaat cagtccaagg acagggagga gcaggctgag	1860
ttcaagaaga tgtacggtct cattgaccag tacaagttga aggggcatat ccgctggatc	1920
tcagctcaga tgaaccgtgt tcgtaacggg gagttgtacc gatacatttg tgacaccaag	1980
ggagtctttg tccagcctgc attctatgaa gcgtttgggtc tgactgtcat cgaagccatg	2040
acatgtgggtt tgccaacaat cgcaacatgc catggtggcc ctgctgagat tattgttgat	2100
ggggtgtctg gtctgcacat tgatccttac cacagtgaca aggctgctga tatcttggtc	2160
aacttctttg agaagtgcaa gcaggattca acctactggg acaatatttc acagggaggt	2220
ctgcagagga tttacgagaa gtacacctgg aagctgtact ctgagaggct gatgaccttg	2280
actggtgtat acggattctg gaagtacgta agcaaccttg agaggcgcgga gactcgccgt	2340
tacattgaga tgttctatgc tctgaaatac cgcagcctgg ccagcgccgt cccattggct	2400
gtcgatggag agagcacatc caagtaa	2427