

# SEQUENCE LISTING

<110> Daiichi Sankyo Company, Limited

<120> NOVEL ENDOS MUTANT ENZYME

<130> SP102171SGM

<150> JP2015-141901

<151> 2015-07-16

<160> 3

<170> PatentIn version 3.5

<210> 1

<211> 995

<212> PRT

<213> Artificial Sequencec

<220>

<223> D233Q mutant of Endo-S

<400> 1

Met Asp Lys His Leu Leu Val Lys Arg Thr Leu Gly Cys Val Cys Ala  
1 5 10 15

Ala Thr Leu Met Gly Ala Ala Leu Ala Thr His His Asp Ser Leu Asn  
20 25 30

Thr Val Lys Ala Glu Glu Lys Thr Val Gln Val Gln Lys Gly Leu Pro  
35 40 45

Ser Ile Asp Ser Leu His Tyr Leu Ser Glu Asn Ser Lys Lys Glu Phe  
50 55 60

Lys Glu Glu Leu Ser Lys Ala Gly Gln Glu Ser Gln Lys Val Lys Glu  
65 70 75 80

Ile Leu Ala Lys Ala Gln Gln Ala Asp Lys Gln Ala Gln Glu Leu Ala  
85 90 95

Lys Met Lys Ile Pro Glu Lys Ile Pro Met Lys Pro Leu His Gly Pro  
100 105 110

Leu Tyr Gly Gly Tyr Phe Arg Thr Trp His Asp Lys Thr Ser Asp Pro  
115 120 125

Thr Glu Lys Asp Lys Val Asn Ser Met Gly Glu Leu Pro Lys Glu Val  
130 135 140

Asp Leu Ala Phe Ile Phe His Asp Trp Thr Lys Asp Tyr Ser Leu Phe  
145 150 155 160

Trp Lys Glu Leu Ala Thr Lys His Val Pro Lys Leu Asn Lys Gln Gly  
165 170 175

Thr Arg Val Ile Arg Thr Ile Pro Trp Arg Phe Leu Ala Gly Gly Asp  
 180 185 190

Asn Ser Gly Ile Ala Glu Asp Thr Ser Lys Tyr Pro Asn Thr Pro Glu  
 195 200 205

Gly Asn Lys Ala Leu Ala Lys Ala Ile Val Asp Glu Tyr Val Tyr Lys  
 210 215 220

Tyr Asn Leu Asp Gly Leu Asp Val Gln Val Glu His Asp Ser Ile Pro  
 225 230 235 240

Lys Val Asp Lys Lys Glu Asp Thr Ala Gly Val Glu Arg Ser Ile Gln  
 245 250 255

Val Phe Glu Glu Ile Gly Lys Leu Ile Gly Pro Lys Gly Val Asp Lys  
 260 265 270

Ser Arg Leu Phe Ile Met Asp Ser Thr Tyr Met Ala Asp Lys Asn Pro  
 275 280 285

Leu Ile Glu Arg Gly Ala Pro Tyr Ile Asn Leu Leu Leu Val Gln Val  
 290 295 300

Tyr Gly Ser Gln Gly Glu Lys Gly Gly Trp Glu Pro Val Ser Asn Arg  
 305 310 315 320

Pro Glu Lys Thr Met Glu Glu Arg Trp Gln Gly Tyr Ser Lys Tyr Ile  
 325 330 335

Arg Pro Glu Gln Tyr Met Ile Gly Phe Ser Phe Tyr Glu Glu Asn Ala  
 340 345 350

Gln Glu Gly Asn Leu Trp Tyr Asp Ile Asn Ser Arg Lys Asp Glu Asp  
 355 360 365

Lys Ala Asn Gly Ile Asn Thr Asp Ile Thr Gly Thr Arg Ala Glu Arg  
 370 375 380

Tyr Ala Arg Trp Gln Pro Lys Thr Gly Gly Val Lys Gly Gly Ile Phe  
 385 390 395 400

Ser Tyr Ala Ile Asp Arg Asp Gly Val Ala His Gln Pro Lys Lys Tyr  
 405 410 415

Ala Lys Gln Lys Glu Phe Lys Asp Ala Thr Asp Asn Ile Phe His Ser  
 420 425 430

Asp Tyr Ser Val Ser Lys Ala Leu Lys Thr Val Met Leu Lys Asp Lys  
 435 440 445

Ser Tyr Asp Leu Ile Asp Glu Lys Asp Phe Pro Asp Lys Ala Leu Arg  
 450 455 460

Glu Ala Val Met Ala Gln Val Gly Thr Arg Lys Gly Asp Leu Glu Arg  
 465 470 475 480

Phe Asn Gly Thr Leu Arg Leu Asp Asn Pro Ala Ile Gln Ser Leu Glu  
 485 490 495

Gly Leu Asn Lys Phe Lys Lys Leu Ala Gln Leu Asp Leu Ile Gly Leu  
 500 505 510

Ser Arg Ile Thr Lys Leu Asp Arg Ser Val Leu Pro Ala Asn Met Lys  
 515 520 525

Pro Gly Lys Asp Thr Leu Glu Thr Val Leu Glu Thr Tyr Lys Lys Asp  
 530 535 540

Asn Lys Glu Glu Pro Ala Thr Ile Pro Pro Val Ser Leu Lys Val Ser  
 545 550 555 560

Gly Leu Thr Gly Leu Lys Glu Leu Asp Leu Ser Gly Phe Asp Arg Glu  
 565 570 575

Thr Leu Ala Gly Leu Asp Ala Ala Thr Leu Thr Ser Leu Glu Lys Val  
 580 585 590

Asp Ile Ser Gly Asn Lys Leu Asp Leu Ala Pro Gly Thr Glu Asn Arg  
 595 600 605

Gln Ile Phe Asp Thr Met Leu Ser Thr Ile Ser Asn His Val Gly Ser  
 610 615 620

Asn Glu Gln Thr Val Lys Phe Asp Lys Gln Lys Pro Thr Gly His Tyr  
 625 630 635 640

Pro Asp Thr Tyr Gly Lys Thr Ser Leu Arg Leu Pro Val Ala Asn Glu  
 645 650 655

Lys Val Asp Leu Gln Ser Gln Leu Leu Phe Gly Thr Val Thr Asn Gln  
 660 665 670

Gly Thr Leu Ile Asn Ser Glu Ala Asp Tyr Lys Ala Tyr Gln Asn His  
 675 680 685

Lys Ile Ala Gly Arg Ser Phe Val Asp Ser Asn Tyr His Tyr Asn Asn  
 690 695 700

Phe Lys Val Ser Tyr Glu Asn Tyr Thr Val Lys Val Thr Asp Ser Thr  
 705 710 715 720

Leu Gly Thr Thr Thr Asp Lys Thr Leu Ala Thr Asp Lys Glu Glu Thr  
 725 730 735  
 Tyr Lys Val Asp Phe Phe Ser Pro Ala Asp Lys Thr Lys Ala Val His  
 740 745 750  
 Thr Ala Lys Val Ile Val Gly Asp Glu Lys Thr Met Met Val Asn Leu  
 755 760 765  
 Ala Glu Gly Ala Thr Val Ile Gly Gly Ser Ala Asp Pro Val Asn Ala  
 770 775 780  
 Arg Lys Val Phe Asp Gly Gln Leu Gly Ser Glu Thr Asp Asn Ile Ser  
 785 790 795 800  
 Leu Gly Trp Asp Ser Lys Gln Ser Ile Ile Phe Lys Leu Lys Glu Asp  
 805 810 815  
 Gly Leu Ile Lys His Trp Arg Phe Phe Asn Asp Ser Ala Arg Asn Pro  
 820 825 830  
 Glu Thr Thr Asn Lys Pro Ile Gln Glu Ala Ser Leu Gln Ile Phe Asn  
 835 840 845  
 Ile Lys Asp Tyr Asn Leu Asp Asn Leu Leu Glu Asn Pro Asn Lys Phe  
 850 855 860  
 Asp Asp Glu Lys Tyr Trp Ile Thr Val Asp Thr Tyr Ser Ala Gln Gly  
 865 870 875 880  
 Glu Arg Ala Thr Ala Phe Ser Asn Thr Leu Asn Asn Ile Thr Ser Lys  
 885 890 895  
 Tyr Trp Arg Val Val Phe Asp Thr Lys Gly Asp Arg Tyr Ser Ser Pro  
 900 905 910  
 Val Val Pro Glu Leu Gln Ile Leu Gly Tyr Pro Leu Pro Asn Ala Asp  
 915 920 925  
 Thr Ile Met Lys Thr Val Thr Thr Ala Lys Glu Leu Ser Gln Gln Lys  
 930 935 940  
 Asp Lys Phe Ser Gln Lys Met Leu Asp Glu Leu Lys Ile Lys Glu Met  
 945 950 955 960  
 Ala Leu Glu Thr Ser Leu Asn Ser Lys Ile Phe Asp Val Thr Ala Ile  
 965 970 975  
 Asn Ala Asn Ala Gly Val Leu Lys Asp Cys Ile Glu Lys Arg Gln Leu  
 980 985 990

Leu Lys Lys  
995

<210> 2  
<211> 995  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Novel mutants of Endo-S D233Q having at least one mutation other than D233Q.

<220>  
<221> mutation  
<222> (122)..(122)  
<223> The amino acid is His in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (184)..(184)  
<223> The amino acid is Pro in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (187)..(187)  
<223> The amino acid is Phe in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (279)..(279)  
<223> The amino acid is Asp in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (281)..(281)  
<223> The amino acid is Thr in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (282)..(282)  
<223> The amino acid is Tyr in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (303)..(303)  
<223> The amino acid is Gln in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (348)..(348)  
<223> The amino acid is Tyr in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (350)..(350)  
<223> The amino acid is Glu in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (402)..(402)  
<223> The amino acid is Tyr in wild type EndoS and can be mutated.

<220>  
<221> mutation  
<222> (405)..(405)

<223> The amino acid is Asp in wild type Endos and can be mutated.

<220>

<221> mutation

<222> (406)..(406)

<223> The amino acid is Arg in wild type Endos and can be mutated.

<400> 2

Met Asp Lys His Leu Leu Val Lys Arg Thr Leu Gly Cys Val Cys Ala  
1 5 10 15

Ala Thr Leu Met Gly Ala Ala Leu Ala Thr His His Asp Ser Leu Asn  
20 25 30

Thr Val Lys Ala Glu Glu Lys Thr Val Gln Val Gln Lys Gly Leu Pro  
35 40 45

Ser Ile Asp Ser Leu His Tyr Leu Ser Glu Asn Ser Lys Lys Glu Phe  
50 55 60

Lys Glu Glu Leu Ser Lys Ala Gly Gln Glu Ser Gln Lys Val Lys Glu  
65 70 75 80

Ile Leu Ala Lys Ala Gln Gln Ala Asp Lys Gln Ala Gln Glu Leu Ala  
85 90 95

Lys Met Lys Ile Pro Glu Lys Ile Pro Met Lys Pro Leu His Gly Pro  
100 105 110

Leu Tyr Gly Gly Tyr Phe Arg Thr Trp Xaa Asp Lys Thr Ser Asp Pro  
115 120 125

Thr Glu Lys Asp Lys Val Asn Ser Met Gly Glu Leu Pro Lys Glu Val  
130 135 140

Asp Leu Ala Phe Ile Phe His Asp Trp Thr Lys Asp Tyr Ser Leu Phe  
145 150 155 160

Trp Lys Glu Leu Ala Thr Lys His Val Pro Lys Leu Asn Lys Gln Gly  
165 170 175

Thr Arg Val Ile Arg Thr Ile Xaa Trp Arg Xaa Leu Ala Gly Gly Asp  
180 185 190

Asn Ser Gly Ile Ala Glu Asp Thr Ser Lys Tyr Pro Asn Thr Pro Glu  
195 200 205

Gly Asn Lys Ala Leu Ala Lys Ala Ile Val Asp Glu Tyr Val Tyr Lys  
210 215 220

Tyr Asn Leu Asp Gly Leu Asp Val Gln Val Glu His Asp Ser Ile Pro  
225 230 235 240

Lys Val Asp Lys Lys Glu Asp Thr Ala Gly Val Glu Arg Ser Ile Gln  
 245 250 255  
 Val Phe Glu Glu Ile Gly Lys Leu Ile Gly Pro Lys Gly Val Asp Lys  
 260 265 270  
 Ser Arg Leu Phe Ile Met Xaa Ser Xaa Xaa Met Ala Asp Lys Asn Pro  
 275 280 285  
 Leu Ile Glu Arg Gly Ala Pro Tyr Ile Asn Leu Leu Leu Val Xaa Val  
 290 295 300  
 Tyr Gly Ser Gln Gly Glu Lys Gly Gly Trp Glu Pro Val Ser Asn Arg  
 305 310 315 320  
 Pro Glu Lys Thr Met Glu Glu Arg Trp Gln Gly Tyr Ser Lys Tyr Ile  
 325 330 335  
 Arg Pro Glu Gln Tyr Met Ile Gly Phe Ser Phe Xaa Glu Xaa Asn Ala  
 340 345 350  
 Gln Glu Gly Asn Leu Trp Tyr Asp Ile Asn Ser Arg Lys Asp Glu Asp  
 355 360 365  
 Lys Ala Asn Gly Ile Asn Thr Asp Ile Thr Gly Thr Arg Ala Glu Arg  
 370 375 380  
 Tyr Ala Arg Trp Gln Pro Lys Thr Gly Gly Val Lys Gly Gly Ile Phe  
 385 390 395 400  
 Ser Xaa Ala Ile Xaa Xaa Asp Gly Val Ala His Gln Pro Lys Lys Tyr  
 405 410 415  
 Ala Lys Gln Lys Glu Phe Lys Asp Ala Thr Asp Asn Ile Phe His Ser  
 420 425 430  
 Asp Tyr Ser Val Ser Lys Ala Leu Lys Thr Val Met Leu Lys Asp Lys  
 435 440 445  
 Ser Tyr Asp Leu Ile Asp Glu Lys Asp Phe Pro Asp Lys Ala Leu Arg  
 450 455 460  
 Glu Ala Val Met Ala Gln Val Gly Thr Arg Lys Gly Asp Leu Glu Arg  
 465 470 475 480  
 Phe Asn Gly Thr Leu Arg Leu Asp Asn Pro Ala Ile Gln Ser Leu Glu  
 485 490 495  
 Gly Leu Asn Lys Phe Lys Lys Leu Ala Gln Leu Asp Leu Ile Gly Leu  
 500 505 510

Ser Arg Ile Thr Lys Leu Asp Arg Ser Val Leu Pro Ala Asn Met Lys  
 515 520 525

Pro Gly Lys Asp Thr Leu Glu Thr Val Leu Glu Thr Tyr Lys Lys Asp  
 530 535 540

Asn Lys Glu Glu Pro Ala Thr Ile Pro Pro Val Ser Leu Lys Val Ser  
 545 550 555 560

Gly Leu Thr Gly Leu Lys Glu Leu Asp Leu Ser Gly Phe Asp Arg Glu  
 565 570 575

Thr Leu Ala Gly Leu Asp Ala Ala Thr Leu Thr Ser Leu Glu Lys Val  
 580 585 590

Asp Ile Ser Gly Asn Lys Leu Asp Leu Ala Pro Gly Thr Glu Asn Arg  
 595 600 605

Gln Ile Phe Asp Thr Met Leu Ser Thr Ile Ser Asn His Val Gly Ser  
 610 615 620

Asn Glu Gln Thr Val Lys Phe Asp Lys Gln Lys Pro Thr Gly His Tyr  
 625 630 635 640

Pro Asp Thr Tyr Gly Lys Thr Ser Leu Arg Leu Pro Val Ala Asn Glu  
 645 650 655

Lys Val Asp Leu Gln Ser Gln Leu Leu Phe Gly Thr Val Thr Asn Gln  
 660 665 670

Gly Thr Leu Ile Asn Ser Glu Ala Asp Tyr Lys Ala Tyr Gln Asn His  
 675 680 685

Lys Ile Ala Gly Arg Ser Phe Val Asp Ser Asn Tyr His Tyr Asn Asn  
 690 695 700

Phe Lys Val Ser Tyr Glu Asn Tyr Thr Val Lys Val Thr Asp Ser Thr  
 705 710 715 720

Leu Gly Thr Thr Thr Asp Lys Thr Leu Ala Thr Asp Lys Glu Glu Thr  
 725 730 735

Tyr Lys Val Asp Phe Phe Ser Pro Ala Asp Lys Thr Lys Ala Val His  
 740 745 750

Thr Ala Lys Val Ile Val Gly Asp Glu Lys Thr Met Met Val Asn Leu  
 755 760 765

Ala Glu Gly Ala Thr Val Ile Gly Gly Ser Ala Asp Pro Val Asn Ala  
 770 775 780

Arg Lys Val Phe Asp Gly Gln Leu Gly Ser Glu Thr Asp Asn Ile Ser  
 785 790 795 800  
 Leu Gly Trp Asp Ser Lys Gln Ser Ile Ile Phe Lys Leu Lys Glu Asp  
 805 810 815  
 Gly Leu Ile Lys His Trp Arg Phe Phe Asn Asp Ser Ala Arg Asn Pro  
 820 825 830  
 Glu Thr Thr Asn Lys Pro Ile Gln Glu Ala Ser Leu Gln Ile Phe Asn  
 835 840 845  
 Ile Lys Asp Tyr Asn Leu Asp Asn Leu Leu Glu Asn Pro Asn Lys Phe  
 850 855 860  
 Asp Asp Glu Lys Tyr Trp Ile Thr Val Asp Thr Tyr Ser Ala Gln Gly  
 865 870 875 880  
 Glu Arg Ala Thr Ala Phe Ser Asn Thr Leu Asn Asn Ile Thr Ser Lys  
 885 890 895  
 Tyr Trp Arg Val Val Phe Asp Thr Lys Gly Asp Arg Tyr Ser Ser Pro  
 900 905 910  
 Val Val Pro Glu Leu Gln Ile Leu Gly Tyr Pro Leu Pro Asn Ala Asp  
 915 920 925  
 Thr Ile Met Lys Thr Val Thr Thr Ala Lys Glu Leu Ser Gln Gln Lys  
 930 935 940  
 Asp Lys Phe Ser Gln Lys Met Leu Asp Glu Leu Lys Ile Lys Glu Met  
 945 950 955 960  
 Ala Leu Glu Thr Ser Leu Asn Ser Lys Ile Phe Asp Val Thr Ala Ile  
 965 970 975  
 Asn Ala Asn Ala Gly Val Leu Lys Asp Cys Ile Glu Lys Arg Gln Leu  
 980 985 990  
 Leu Lys Lys  
 995

<210> 3  
 <211> 2985  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> cDNA of EndoS D233Q

<400> 3  
 atggacaaac atctgctggt taaacgtacc ctgggttgtg tttgtgcagc aaccctgatg

60

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ggtgcagcac tggcaacca tcatgatagc ctgaataccg ttaaagcaga agaaaaaacc    | 120  |
| gttcaggttc agaaaggtct gccgagcatt gatagtctgc attatctgag cgaaaacagc   | 180  |
| aaaaaagaat ttaaagaaga actgagcaaa gccggtcaag aaagccagaa agttaaagaa   | 240  |
| attctggcaa aagcacagca ggcagataaa caggcacaag aactggccaa aatgaaaatc   | 300  |
| ccggaaaaaa ttccgatgaa accgctgcat ggtccgctgt atggtgggta ttttcgtacc   | 360  |
| tggcatgata aaaccagcga tccgaccgaa aaagataaag ttaatagcat ggggtgaactg  | 420  |
| ccgaaagaag ttgatctggc atttatcttt cagcactgga ccaaagatta tagcctgttt   | 480  |
| tggaaagaac tggcgaccaa acatgttccg aaactgaata aacagggcac ccgtgttatt   | 540  |
| cgtaccattc cgtggcgttt tctggcaggc ggtgataata gcggtattgc agaagatacc   | 600  |
| agcaaataatc cgaatacacc ggaaggtaat aaagcactgg cgaaagcaat tgtggatgag  | 660  |
| tatgtgtata aatacaatct ggatggcctg gatgttcaag ttgaacatga tagcattccg   | 720  |
| aaagtggaca aaaaagagga taccgcaggc gttgaacgta gcattcagggt ttttgaagaa  | 780  |
| atcggcaaac tgattgggtcc gaaagggtgt gataaaagcc gtctgtttat tatggatagc  | 840  |
| acctatatgg ccgataaaaaa cccgctgatt gaacgtgggtg caccgtatat taacctgctg | 900  |
| ctggttcagg tttatggtag ccagggtgaa aaagggtggt gggaaaccggt tagcaatcgt  | 960  |
| cctgaaaaaa ccatggaaga acgttggcag ggttatagca aatatatccg tccggaacag   | 1020 |
| tatatgatcg gctttagctt ttatgaagaa aatgcccaag agggcaatct gtggtatgat   | 1080 |
| attaacagcc gtaaagatga ggataaagcc aatggcatta acaccgatat caccggtaca   | 1140 |
| cgtgcagaac gttatgcccg ttggcagccg aaaaccgggtg gtgttaaagg tggatTTTT   | 1200 |
| agctatgcc a ttgatcgtga tgggtgttgca catcaaccga aaaaatacgc caaacagaaa | 1260 |
| gagttcaaag acgccaccga taacattttc catagcgatt atagcgttag caaagccctg   | 1320 |
| aaaaccgtga tgctgaaaga caaaagctat gatctgatcg acgagaaaga ttttccggat   | 1380 |
| aaagcgctgc gtgaagcagt tatggcacag gttggcaccg gtaaagggtga tctggaacgt  | 1440 |
| tttaattggca ccctgcgtct ggataatccg gcaattcaga gcctggaagg tctgaacaaa  | 1500 |
| ttcaaaaaac tggcacagct ggatctgatt ggcctgagcc gtattacaa actggatcgt    | 1560 |
| agcgttctgc ctgcaaatat gaaaccgggt aaagataccc tggaaccgt tctggaacc     | 1620 |
| tacaaaaaag acaacaaaga agaaccggca accattccgc ctgttagcct gaaagttagc   | 1680 |
| ggtctgaccg gtctgaaaga gctggacctg tcaggttttg atcgcgaaac cctggcagggt  | 1740 |
| ctggatgcag ccacactgac cagcctggaa aaagttgata ttagcggcaa taaactggac   | 1800 |
| ctggcaccgg gtacagaaaa tcgtcagatc ttcgatacca tgctgagcac cattagcaat   | 1860 |
| catgttggtg gcaatgaaca gaccgtgaaa tttgataaac agaaaccgac cggtcattat   | 1920 |
| cctgatacct atggtaaaac cagtctgcgt ctgccgggtg caaatgaaaa agtggatctg   | 1980 |
| cagagccagc tgctgtttgg caccgttacc aatcagggca cactgattaa tagcgaagca   | 2040 |
| gattataaag cctatcagaa ccataaaatt gccggtcgta gctttgtgga tagcaactac   | 2100 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cattacaaca acttcaaagt gagctacgag aactataccg tgaaagttac cgatagcacc  | 2160 |
| ctgggcacca ccacagataa aacactggcc acggataaag aggaaaccta taaagtcgat  | 2220 |
| tttttcagcc ctgccgacaa aaccaaagca gttcataccg caaaagttat cgtgggtgat  | 2280 |
| gaaaaaacga tgatgggttaa tctggcagaa ggtgcaaccg ttattggtgg tagcgcagat | 2340 |
| ccggttaatg cacgtaaagt ttttgatggt cagctgggta gcgaaaccga taatatcagc  | 2400 |
| ctgggttggg atagcaaaca gagcattatc tttaaactga aagaggacgg cctgattaaa  | 2460 |
| cattggcgct tttttaacga tagcgcacgt aatccgaaa ccacaaataa accgattcaa   | 2520 |
| gaagcaagcc tgagatttt taacatcaaa gattacaacc tggacaacct gctggaaaac   | 2580 |
| ccgaacaaat ttgatgatga gaaatactgg attaccgtgg atacctatag cgacaggggt  | 2640 |
| gaacgtgcaa ccgcatttag caataccctg aataatatca cgagcaaata ttggcgtgtg  | 2700 |
| gtgtttgata ccaaaggcga tcgttatagc agtccggttg tgccggaact gcagattctg  | 2760 |
| ggttatccgc tgccgaatgc agataccatt atgaaaacag ttaccaccgc caaagaactg  | 2820 |
| tcacagcaga aagacaaatt tagccagaaa atgctggacg aactgaaaat caaagaaatg  | 2880 |
| gcactggaaa ccagcctgaa ctccaaaatt ttcgatgtta ccgccattaa tgcaaatgcc  | 2940 |
| ggtgttctga aagattgcat tgaaaaacgt cagctgctga aaaaa                  | 2985 |