SEQUENCE LISTING

<110> GlycoVaxyn AG

<120> NOVEL POLYSACCHARIDE AND USES THEREOF

<130> 103749PC

<140> TBA
<141> On even date herewith

<150> 61/943,710
<151> 2014-02-24

<160> 15

<170> PatentIn version 3.5

<210> 1
<211> 1086
<212> DNA
<213> E. Coli

<220>
<223> rmlB (upec138)

<400> 1
gtgaagatac ttgttactgg tggcgcagga tttattggtt ctgctgttgt tcgtcacata 60

ataaataata cgcaagatag tgttgttaat gtcgataaat taacatacgc cggaaacctg 120

gaatcacttg cagatgtttc tgattctgaa cgctatttct ttgaacatgc ggatatttgt 180

gatgcagctg caatggcacg gatttttgct cagcatcagc cggatgcagt gatgcacctg 240

gcagctgaaa gccatgttga ccgttcaatt acaggccctg cggcatttat tgaaaccaat 300

attgtgggta cttatgtcct tttagaagcg gctcggaatt attggtctgg tctggatgat 360

gaaaagaaaa aaaacttccg ttttcatcat atttctactg atgaggtgta tggtgactta 420

ccccatccgg atgaagtaaa tagcaatgaa acgttgccgc tatttacgga aacgacagca 480

tacgcgccaa gtagtccata ttctgcttct aaagcttcca gcgatcattt ggttcgcgca 540

tggaaacgta cttatggttt accgaccatt gtgactaatt gctcgaacaa ctatggtcct 600

tatcatttcc cggaaaagct tattccactg gttattctta attcactgga aggtaaggca 660

ttacctattt atggcaaagg agatcagatc cgcgactggt tgtatgtaga ggatcatgct 720

cgagcgttat ataccgtcgt aaccgaaggt aaagcgggcg aaacttataa cattggtgga 780

cacaacgaaa agaaaaacat cgacgtagtg ttcactattt gtgatttgtt ggatgagata 840

gtcccgaaag agaaatctta ccgcgagcaa attacttatg ttaccgatcg tccgggacac 900

gatcgccgtt atgcgattga tgctgagaag attggtcgcg aattgggatg gaaaccacag 960

gaaacgtttg agagtgggat tcgtaaaacg gtggaatggt acctgtccaa tacaaaatgg 1020

gttgataatg tgaaaagtgg tgcctatcaa tcgtggattg aacagaacta tgagggccgc 1080

cagtaa 1086

<210> 2
<211> 900

<212> DNA
<213> E. Coli

<220>
<223> rmlD (upec138)

<400> 2
atgaatatcc tccttttgg caaaacaggg caggtaggtt gggaactaca gcgtgctctg 60

gcacctctgg gtaatttgat tgctcttgat gttcactcca ctgattactg tggtgatttt 120

agtaatcctg aaggtgtagc tgaaaccgta agaagcattc ggcctgatat tattgtcaac 180

gcagccgctc acaccgcagt agacaaagca gaatcagaac cgaagtttgc acaattactg 240

aacgcgacga gtgtcgaagc gatcgcgaaa gcagccaatg aagtcggcgc ctgggttatt 300

cactactcta ctgactacgt atttccgggg accggtgaaa taccatggca ggaggaggat 360

gcaaccgcac cgctaaatgt ttacggtgaa accaagttag cgggagaaaa agcattacaa 420

gagcattgtg cgaagcacct tattttccgg accagctggg tctatgcagg taaaggaaat 480

aacttcgcca aaacaatgtt gcgtctggca aaagagcgtg aagaattagc cgttattaat 540

gatcagtttg gtgcgccaac tggcgcagag ttactggctg attgtacggc acatgctatt 600

cgtgtggcac tgaataaacc ggaagtcgca ggcttgtacc atctggtagc tagtggtacc 660

acaacgtggc acgattatgc tgcgctggtt tttgaagagg cgcgcaaagc aggcattccc 720

cttgcactca acaagctcaa cgcagtacca acaacagcct atcctacacc agctcgtcgt 780

ccacataact ctcgccttaa tacagaaaaa tttcagcaga actttgcgct tgtcttgcct 840

gactggcagg ttggcgtgaa acgaatgctt aacgaattat ttacgactac agcaatttaa 900

<210> 3
<211> 879
<212> DNA
<213> E. Coli

<220>
<223> rmlA (upec138)

<400> 3
atgaaaacgc gtaaaggtat tattttggcg ggtggttctg gtactcgtct ttatcctgtg 60

acgatggccg tcagtaaaca gctgttaccg atttatgata aaccgatgat ctattacccg 120

ctctctacac tgatgttagc gggtattcgc gatattctga ttatcagtac accacaggat 180

actcctcgtt ttcaacaact gctgggtgac gggagccagt ggggcctgaa tcttcagtac 240

aaagtgcaac cgagtccgga tggtcttgcg caggcgttta ttatcggtga agagtttatt 300

ggtggtgatg attgtgcttt ggtacttggt gataatatct tctacggcca cgacctgccg 360

aagttaatgg acgtagctgt taacaaagaa agtggtgcaa cggtatttgc ctatcacgtt 420

aatgatcctg aacgttatgg tgtcgtggag tttgataata acggtactgc aattagcctg 480

gaagaaaaac cgctggaacc aaaaagtaac tatgcggtta ctgggcttta tttctatgac 540

aatgacgttg tggaaatggc gaaaaacctt aagccttctg cccgaggtga actggaaatt 600

accgatatta accgtattta tatggaacaa ggacgtttgt ctgtcgctat gatggggcgt 660

ggctatgcat ggctggatac agggacgcat caaagtctta ttgaagcaag caacttcatt 720

gccaccattg aagagcgcca gggactaaag gtttcctgtc cggaagaaat tgcttatcgt 780

aaagggttta ttgatgctga gcaggtaaaa gtattagccg aaccgttgaa gaaaaatgct 840

tatggtcagt atctgctcaa aatgattaaa ggttattaa 879

<210> 4
<211> 543
<212> DNA
<213> E. Coli

<220>
<223> rmlC (upec138)

<400> 4
atgaacgtaa ttaaaactga aattcctgat gtgctgattt ttgaaccaaa agtttttggg 60

gatgaacgtg gcttcttttt tgagagtttt aatcagagga tttttgaaga agcagtaggt 120

cgtaaggttg agtttgttca ggataaccat tctaagtcca gtaaaggtgt tttacgtggt 180

cttcattatc agttagaacc ttatgctcaa ggaaaactgg tgcgctgtgt tgttggcgag 240

gtttttgatg ttgcggttga tattcgtaaa tcgtcaccta catttgggaa atgggttggg 300

gtgaatttgt ctgctgagaa taagcgtcag ttgtggattc ctgagggatt tgcacatggt 360

ttttggtgc tgagtgattt agcagaagtt ttatataaaa cgaatcaata ttatgctcca 420

tcacatgaaa aaaatattat atggaatgac ctcttgctta atattaaatg gccgagcaca 480

gcactgatca ctctgtctga taaggatgca aatggggaaa gatttgaact aagtgagttt 540

tga 543

<210> 5
<211> 1260
<212> DNA
<213> E. Coli

<220>
<223> wzx (upec138)

<400> 5
atgtctctct taaaacatag tatatggaat gttgcgggct actttatacc aacattaatt 60

gcaattcccg cctttggatt aattgcgagg aaaattggtg tagaactatt tggtttgtat 120

acgttagcaa tgatttttat agggtatgca agtatatttg atgctgggtt aacaagagct 180

gttgtgcgtg aaatagcatt actaaaaaac agagtggacg attgtaatac gataatagta 240

acttctatta tcgctgtgat atttttaggg tttatcggag gcgggggagt gtttctgctt 300

aaaggcgata ttattgaact gttaaatatc tcaccaatat attacgccga ttcgataaag 360

tctctagtat tattatcatc tctgatacct gtattcttag tcacgcaaat actattagca 420

gagcttgagg gtcgggaata ttttgggatt ctaaatatac aaaaaagtgt agggaattct 480

ttaattgcag ggttacctgc attatttgtt ttaattaatc aaacgctttt ttctgcaatt 540

attggtgtag cgattgcaag agttatatgc ttgtggttaa gctacattat gagcagggaa 600

agaataacta tcgatatctc atttttttca ataactgttt taaagcggtt atttagatat 660

ggcgggtggg taactataag taacataata tctcctatat tagcgagtat ggatagattt 720

attctatccc atatccaggg agcatcaaaa atatcattct atacagtccc taatgagctg 780

gtaactaggc ttggaatagt tccaggctct cttgggaaag ctgtttttcc aaaattaagt 840

catgcaagga attttacagc gtcatatgca gagcaaaaaa aagcttatat attaatgact 900

gtcattgtaa tgcctttggt tttatttgta tattattacg caaagtttat tttaacattg 960

tggatggggg ctgagtatgc agggatttcg gtcgaaatat tacggattat gcttataggg 1020

tatattttta actgttattc acaaatctct ttgccaaca tacaggcctt tggaaaagca 1080

aaatacactg catacatcca tatgatggaa tttattcctt atttgataat gttatatata 1140

atttcaaagg aatatggggt tattggtgtt gcgtggttat ggacaattcg agtaataatt 1200

gatttttga tgcttttata tatgagttat cgttgtaata atcttatgaa aaaagggtag 1260

<210> 6
<211> 966
<212> DNA
<213> E. Coli

<220>
<223> wekA (upec138)

<400> 6
atgatatata ttgtggtatt aaattggaat ggggctatag ataccattaa ttgtgttaaa 60

agtttaatgg atttaaatgt tagcgattat aaaattatca ttgttgataa ctgttctatg 120

gataactcat atgatactat aaaagaaaat cttaattcat tatatattgc tgataaaagt 180

atcattgagg tgaagtatga ggatagaaat aaatataaaa ccttagaaaa cgataaaatc 240

atattaatac aatctccgca aaataatggg tacgcaagtg gtaataatat tggcatagag 300

ttcgctctta atcaggagaa tatgaaatac gtctgggttc tgaataatga tactgaagtg 360

gataaagagg ctttaactca tttaattagt aaatgtgatt cagataaaag tatagggatt 420

tgcggttctc gtttagtcta ttttgccgac agagagatgc agcaaggact aggtggggtg 480

cataacaaat ggttatgcac tacaaaaaat tatgaaatgg gaagattagt ttccaaaaaa 540

tatgatgatg aagtcattag taatgatata gattatataa ttggcgcatc gatgtttttc 600

tctagagaat gtttggaaac agttggattg atgaatgaag aatatttttt atactatgaa 660

gagttagata tttgcctcag agcaaaagca aagaacttta aattaggtat ttgctcagaa 720

agtttggttt atcataaaat aggtgcaagt actgatgggg gaaagagcat gatggctgat 780

ctttgctcaa taaaaaatag gctggtcatt acagaaaggt tttatcccca atattattgg 840

acggtatggt tgtcactttt tgttgtagca tttaaccgtg ctagaagagg tgagtttaat 900

aagatgaaaa gatgtttgaa tgttatgttt aacttcaaac gaaacaaagg tagcaaatgc 960

cattag 966

<210> 7
<211> 1149
<212> DNA
<213> E. Coli

<220>
<223> wekB (upec138)

<400> 7
atgaaagtgg ctttttatc tgcttatgat ccactatcta catccagttg gtctggcaca 60

ccttattata tgctaaaggc attatcgaag agaaatattt ccattgaaat attaggaccg 120

gtaaatagct atatgatata catgttaaaa gtatataaat taatattaag gtgtttcgga 180

aaagaatatg attatagtca ttcgaagttg ctttccaggt attacggtag aatattcggt 240

aggaaattaa aaaaaattga tggtttggat tttattatcg cacctgcagg ttcctcacaa 300

attgcttttt taaaaacaac cataccaata atatatctat cggatacaac atatgatcaa 360

ttaaaaagct attatccgaa tttaaataaa aaaacaatta taaatgatga ggatgcaagt 420

ttaatcgaac gcaaggctat tgaaaaagca acagtagtat ctttcccatc taaatgggca 480

atggattttt gcaggaatta ttacagatta gattttgata aattagttga aataccatgg 540

ggggctaatt tatttgatga tattcacttt gctaataaaa atataattca aaagaatagt 600

tatacttgtc ttttcttggg agttgattgg gaaagaaaag gtgggaaaac agccttgaaa 660

gcaattgaat atgtaaggca gttatatggg atcgatgtta gactaaaaat ttgtggatgt 720

actccgaatc aaaagatttt acctacttgg gttgaattaa ttgataaagt agataaaaat 780

aacgttgacg aatatcagaa attcatcgat gtgttatcta acgctgatat acttctttta 840

ccaaccattg ctgaatgtta tggaatggta ttttgtgaag ctgctgcttt tggattgcct 900

gttgtcgcta cagatacagg tggagtcagt tctatagtta tcaacgaaag gacggggata 960

ttaattaaag acccgttaga ctataagcac tttggaaatg caattcataa aataattagt 1020

tccgtagaga cttatcaaaa ctactcccaa aacgcaagaa ttagatataa taatatattg 1080

cattgggaca attgggctaa aaagataatt gagattatgt atgagcataa gaatagaaga 1140

atcaaatag 1149

<210> 8
<211> 1218
<212> DNA
<213> E. Coli

<220>
<223> wzy (upec138)

<400> 8
atgagcataa gaatagaaga atcaaatagc acaaaaagaa ttatatgttt atttatactt 60

tttcttgttt tccctgattt tttgttttat acattagggg ttgataattt tagcatttca 120

acgataatct caattacatt gctttttgtt tttttaagag ctaaaaatat ttgcaaagat 180

aattttctaa taatagtagc gttattcata ttgttgtgtt ttaactgttt gttaagtatg 240

ctatttaata ttgaacaggc tttaacattt aaagttgtac tttcaatata tagcatctta 300

ataatggcat acgtctcctc ttgttatgca cagacgttgt ggttatgttc tgaagaaata 360

cttaagagat ccgtctttta tttgttcgca tttctttgcc ttattggcat tataagtatt 420

cttttacaga agactgagat tatacatgat aaaagtatga ttctttttcc tgaaccatca 480

gcatttgcat tggtttttat acctatcttt tcattttgtt tatactatac aagaggggg 540

gggctactat tgctctatat attatctttg ggtattgcgt taggtatcca gaatttaaca 600

atgttggtag gcattgtgat tagtgttttt gtgatgaaaa aaataactat aaggcaaact 660

attgttatac ttttgggggc atggattttt tccatgatat taagtgattt agacatttct 720

tactatacat cgcggcttga ttttaaaaat actacgaacc tatcagtgct tgtatatctt 780

tcaggaattg aaagagcttt cttgaatttt attacaagtt atggtcttgg tattggtttt 840

caacaaatgg gagtgaatgg ggagatagga atatatcaac aaattttagc tgaacttgat 900

gcccctatgt taaatatata cgatggctca tttatttctt ctaagttaat atctgagttt 960

ggggttattg gtgcattaat gtgtattttc tatttttttt atttttcccg attttatctg 1020

cgtttcaaaa aaagtaagag atattcaccg cagtatattt tagcatatag cttctacatg 1080

tgtttcttca tccctctttt tatacgtggt gctggttata taaaccccta tgtgtttatg 1140

ttattttcat caatattttt gtgcaaatat cacgctaaaa atatcttgat gaaatctaat 1200

gtccagatag ctatataa 1218

<210> 9
<211> 606
<212> DNA
<213> E. Coli

<220>
<223> wbbJ (upec138)

<400> 9
atgtgcatta aaaaaaaact taagttaatt aaacgatatg gcctttatgg tggtcttagg 60

cttcttaaag atatattctt aacaaaattt ttattttgtt caaatgttag gattattaga 120

tttccatgtt atattagaaa agatggaagt gttagttttg gaaaaggttt tacatcaggt 180

gtaggattac gagttgatgc atttatggat gccgtagttt ccattggaga aaatgttcaa 240

attaatgact atgttcacat cgcggctatt aataatgtca ttattggtag agatacatta 300

atagcaagta aagtatttat tagtgatcat aatcatggta ttttttctaa atccgatatc 360

catagttcac caactattat tccttcgtct aggccccttg aatctgcacc tgtgtatatt 420

ggagagcgtg tgtggattgg cgaaaatgtg acaatattac caggtgcgtg tataggtaat 480

ggtgtagtta ttggcgcaaa cagtgttgtt cgtggtgaga ttcctaataa tgtgatcatt 540

gctggtgttc cagctaaaat tgttaaaaaa tataactatg agcgtatgca atgggaaaga 600

atatag 606

<210> 10
<211> 1101

<212> DNA
<213> E. Coli

<220>
<223> wbbK (upec138)

<400> 10
atgggaaaga atatagttgt aatatcggct gttaatttta caaccggagg ccccttacc 60

gtactaaaaa atgtgcttac agcaactaaa gatagagccg aatgtaaatt tattgcactg 120

gttcatagct ctgctgaact aatggaatta tttccgtggg ttgaatttat agagtatcca 180

gaagtcaagt cttcgtgggt taaaagatta tatttcgaat atataacttg caatagatta 240

tctaaggtga ttaaggcaac tcattgggta tgcttacatg atattacagc aaatgttagt 300

gtaccctata gatttgttta ttgccacaat cctgcaccgt tctataaata tttaagctat 360

cgagatatta taggagaacc taaattttat ctttttttatc ttttttatgg gcttttatac 420

aatatcaata taaaaaagaa cacagcagtt tttgttcagc agcagtggct aaaaaaagaa 480

ttcgaaaaaa aatataagtt aaagaatgtt gttgttagtc gccctgaaga tatttgccct 540

tttgaaagtg atggtttggt aagaaataat aataaaaagg atgtgaggat attttaccca 600

gcagtgcccc gtatatttaa aaactttgaa gttatcatac gtgctgcaca aatattacaa 660

gataaaaata ttcattttta tcttactttt gatggtactg aaaataagta tgcaaaaaga 720

atatataaat tagcttccga actgaaaaat gtacatttcc tcggttacct taatgcaacc 780

gagatggtta acttttatca agattcagat attatttgtt tcccatcgaa actagaaacg 840

tggggattac cattatcaga agctaaaaca tacaaaaaat ggatatttgc ggcagactta 900

ccttatgctc atgaagtttt atataactat tcaaaaacta gatattttcc atttgacgat 960

gagaaaatac ttgttcgcta catattagag tacacaagta aaaatatgca tgaagatata 1020

aaaaatagta gggtgaattt taataatgat gcattgactg gttttgaaca gtttattgaa 1080

tatatcctca aggggaactg a 1101

<210> 11
<211> 564
<212> DNA
<213> E. Coli

<220>
<223> wbbL (upec138)

<400> 11
atgattatga ataatgatta ttttctcttt cttaaccccg atgtattcat aaccagtgaa 60

agtttgatta attatgttga ttatataatt agtaatgatt ataagtttag cacattatgt 120

ctttatcgag attttactaa aagcaaacat gattattcaa tacggagttt tccaacttta 180

tatgattttc tttgttcttt tttattgggg gtgaataaaa gtaaaattaa gaaggaaaat 240

atactttctg atactgtagt tgattggtgt gctggctcat ttatgcttat tcatgcttta 300

agtttcttaa atgtgaatgg ttttgatcaa aaatatttta tgtattgtga agatattgac 360

ctttgtatgc gtttaaaatt aagtggagta gatctttact atactcccca ttttgatgct 420

attcattatg cgcagcatga aaatagaaga atatttacta aagcatttcg atggcatata 480

aggagtatta cgcgctacat attacggaaa ccaattcttt cttataaaaa ctatagaaaa 540

attacatccg aactggtaaa gtga 564

<210> 12
<211> 10653
<212> DNA
<213> E. Coli

<220>
<223> E. coli rfb(upec138) gene cluster

<400> 12
gtgaagatac ttgttactgg tggcgcagga tttattggtt ctgctgttgt tcgtcacata 60

ataaataata cgcaagatag tgttgttaat gtcgataaat taacatacgc cggaaacctg 120

gaatcacttg cagatgtttc tgattctgaa cgctatttct ttgaacatgc ggatatttgt 180

gatgcagctg caatggcacg gatttttgct cagcatcagc cggatgcagt gatgcacctg 240

gcagctgaaa gccatgttga ccgttcaatt acaggccctg cggcatttat tgaaaccaat 300

attgtgggta cttatgtcct tttagaagcg gctcggaatt attggtctgg tctggatgat 360

gaaaagaaaa aaaacttccg ttttcatcat atttctactg atgaggtgta tggtgactta 420

ccccatccgg atgaagtaaa tagcaatgaa acgttgccgc tatttacgga aacgacagca 480

tacgcgccaa gtagtccata ttctgcttct aaagcttcca gcgatcattt ggttcgcgca 540

tggaaacgta cttatggttt accgaccatt gtgactaatt gctcgaacaa ctatggtcct 600

tatcatttcc cggaaaagct tattccactg gttattctta attcactgga aggtaaggca 660

ttacctattt atggcaaagg agatcagatc cgcgactggt tgtatgtaga ggatcatgct 720

cgagcgttat ataccgtcgt aaccgaaggt aaagcgggcg aaacttataa cattggtgga 780

cacaacgaaa agaaaaacat cgacgtagtg ttcactattt gtgatttgtt ggatgagata 840

gtcccgaaag agaaatctta ccgcgagcaa attacttatg ttaccgatcg tccgggacac 900

gatcgccgtt atgcgattga tgctgagaag attggtcgcg aattgggatg gaaaccacag 960

gaaacgtttg agagtgggat tcgtaaaacg gtggaatggt acctgtccaa tacaaaatgg 1020

gttgataatg tgaaaagtgg tgcctatcaa tcgtggattg aacagaacta tgagggccgc 1080

cagtaatgaa tatcctcctt tttggcaaaa cagggcaggt aggttgggaa ctacagcgtg 1140

ctctggcacc tctgggtaat ttgattgctc ttgatgttca ctccactgat tactgtggtg 1200

attttagtaa tcctgaaggt gtagctgaaa ccgtaagaag cattcggcct gatattattg 1260

tcaacgcagc cgctcacacc gcagtagaca aagcagaatc agaaccgaag tttgcacaat 1320

tactgaacgc gacgagtgtc gaagcgatcg cgaaagcagc caatgaagtc ggcgcctggg 1380

ttattcacta ctctactgac tacgtatttc cggggaccgg tgaaatacca tggcaggagg 1440

aggatgcaac cgcaccgcta aatgtttacg gtgaaaccaa gttagcggga gaaaaagcat 1500

tacaagagca ttgtgcgaag caccttattt tccggaccag ctgggtctat gcaggtaaag 1560

gaaataactt cgccaaaaca atgttgcgtc tggcaaaaga gcgtgaagaa ttagccgtta 1620

ttaatgatca gtttggtgcg ccaactggcg cagagttact ggctgattgt acggcacatg 1680

ctattcgtgt ggcactgaat aaaccggaag tcgcaggctt gtaccatctg gtagctagtg 1740

gtaccacaac gtggcacgat tatgctgcgc tggtttttga agaggcgcgc aaagcaggca 1800

ttccccttgc actcaacaag ctcaacgcag taccaacaac agcctatcct acaccagctc 1860

gtcgtccaca taactctcgc cttaatacag aaaaatttca gcagaacttt gcgcttgtct 1920

tgcctgactg gcaggttggc gtgaaacgaa tgcttaacga attatttacg actacagcaa 1980

tttaatagtt tttgcatctt gttcgtaatg gtggagcaag atgtattaaa aggaatgatg 2040

aaatgaaaac gcgtaaaggt attattttgg cgggtggttc tggtactcgt ctttatcctg 2100

tgacgatggc cgtcagtaaa cagctgttac cgatttatga taaaccgatg atctattacc 2160

cgctctctac actgatgtta gcgggtattc gcgatattct gattatcagt acaccacagg 2220

atactcctcg ttttcaacaa ctgctgggtg acgggagcca gtggggcctg aatcttcagt 2280

acaaagtgca accgagtccg gatggtcttg cgcaggcgtt tattatcggt gaagagttta 2340

ttggtggtga tgattgtgct ttggtacttg gtgataatat cttctacggc cacgacctgc 2400

cgaagttaat ggacgtagct gttaacaaag aaagtggtgc aacggtattt gcctatcacg 2460

ttaatgatcc tgaacgttat ggtgtcgtgg agtttgataa taacggtact gcaattagcc 2520

tggaagaaaa accgctggaa ccaaaaagta actatgcggt tactgggctt tatttctatg 2580

acaatgacgt tgtggaaatg gcgaaaaacc ttaagccttc tgcccgaggt gaactggaaa 2640

ttaccgatat taaccgtatt tatatggaac aaggacgttt gtctgtcgct atgatggggc 2700

gtggctatgc atggctggat acagggacgc atcaaagtct tattgaagca agcaacttca 2760

ttgccaccat tgaagagcgc cagggactaa aggtttcctg tccggaagaa attgcttatc 2820

gtaaagggtt tattgatgct gagcaggtaa aagtattagc cgaaccgttg aagaaaaatg 2880

cttatggtca gtatctgctc aaaatgatta aaggttatta ataagatgaa cgtaattaaa 2940

actgaaattc ctgatgtgct gatttttgaa ccaaaagttt ttggggatga acgtggcttc 3000

ttttttgaga gttttaatca gaggattttt gaagaagcag taggtcgtaa ggttgagttt 3060

gttcaggata accattctaa gtccagtaaa ggtgttttac gtggtcttca ttatcagtta 3120

gaaccttatg ctcaaggaaa actggtgcgc tgtgttgttg gcgaggtttt tgatgttgcg 3180

gttgatattc gtaaatcgtc acctacattt gggaaatggg ttggggtgaa tttgtctgct 3240

gagaataagc gtcagttgtg gattcctgag ggatttgcac atggtttttt ggtgctgagt 3300

gatttagcag aagttttata taaaacgaat caatattatg ctccatcaca tgaaaaaaat 3360

attatatgga atgacctctt gcttaatatt aaatggccga gcacagcact gatcactctg 3420

tctgataagg atgcaaatgg ggaaagattt gaactaagtg agttttgaaa tgtctctctt 3480

aaaacatagt atatggaatg ttgcgggcta ctttatacca acattaattg caattcccgc 3540

ctttggatta attgcgagga aaattggtgt agaactattt ggtttgtata cgttagcaat 3600

gatttttata gggtatgcaa gtatatttga tgctgggtta acaagagctg ttgtgcgtga 3660

aatagcatta ctaaaaaaca gagtggacga ttgtaatacg ataatagtaa cttctattat 3720

cgctgtgata tttttagggt ttatcggagg cgggggagtg tttctgctta aaggcgatat 3780

tattgaactg ttaaatatct caccaatata ttacgccgat tcgataaagt ctctagtatt 3840

attatcatct ctgatacctg tattcttagt cacgcaaata ctattagcag agcttgaggg 3900

tcgggaatat tttgggattc taaatataca aaaaagtgta gggaattctt taattgcagg 3960

gttacctgca ttatttgttt taattaatca aacgcttttt tctgcaatta ttggtgtagc 4020

gattgcaaga gttatatgct tgtggttaag ctacattatg agcagggaaa gaataactat 4080

cgatatctca ttttttttcaa taactgtttt aaagcggtta tttagatatg gcgggtgggt 4140

aactataagt aacataatat ctcctatatt agcgagtatg gatagattta ttctatccca 4200

tatccaggga gcatcaaaaa tatcattcta tacagtccct aatgagctgg taactaggct 4260

tggaatagtt ccaggctctc ttgggaaagc tgtttttcca aaattaagtc atgcaaggaa 4320

ttttacagcg tcatatgcag agcaaaaaaa agcttatata ttaatgactg tcattgtaat 4380

gcctttggtt ttatttgtat attattacgc aaagtttatt ttaacattgt ggatgggggc 4440

tgagtatgca gggatttcgg tcgaaatatt acggattatg cttatagggt atatttttaa 4500

ctgttattca caaatctctt ttgccaacat acaggccttt ggaaaagcaa aatacactgc 4560

atacatccat atgatggaat ttattcctta tttgataatg ttatatataa tttcaaagga 4620

atatggggtt attggtgttg cgtggttatg gacaattcga gtaataattg atttttgat 4680

gcttttatat atgagttatc gttgtaataa tcttatgaaa aaagggtagc ctgatgatat 4740

atattgtggt attaaattgg aatggggcta tagataccat taattgtgtt aaaagtttaa 4800

tggatttaaa tgttagcgat tataaaatta tcattgttga taactgttct atggataact 4860

catatgatac tataaaagaa aatcttaatt cattatatat tgctgataaa agtatcattg 4920

aggtgaagta tgaggataga aataaatata aaaccttaga aaacgataaa atcatattaa 4980

tacaatctcc gcaaaataat gggtacgcaa gtggtaataa tattggcata gagttcgctc 5040

ttaatcagga gaatatgaaa tacgtctggg ttctgaataa tgatactgaa gtggataaag 5100

aggctttaac tcatttaatt agtaaatgtg attcagataa aagtataggg atttgcggtt 5160

ctcgtttagt ctattttgcc gacagagaga tgcagcaagg actaggtggg gtgcataaca 5220

aatggttatg cactacaaaa aattatgaaa tgggaagatt agtttccaaa aaatatgatg 5280

atgaagtcat tagtaatgat atagattata taattggcgc atcgatgttt ttctctagag 5340

aatgtttgga aacagttgga ttgatgaatg aagaatattt tttatactat gaagagttag 5400

atatttgcct cagagcaaaa gcaaagaact ttaaattagg tatttgctca gaaagtttgg 5460

tttatcataa aataggtgca agtactgatg ggggaaagag catgatggct gatctttgct 5520

caataaaaaa taggctggtc attacagaaa ggttttatcc ccaatattat tggacggtat 5580

ggttgtcact ttttgttgta gcatttaacc gtgctagaag aggtgagttt aataagatga 5640

aaagatgttt gaatgttatg tttaacttca aacgaaacaa aggtagcaaa tgccattaga 5700

atatgcactt aatcatggtg ttaataaatc tatagtttga tatgttatta aagggtattt 5760

aatgaaagtg gctttttat ctgcttatga tccactatct acatccagtt ggtctggcac 5820

accttattat atgctaaagg cattatcgaa gagaaatatt tccattgaaa tattaggacc 5880

ggtaaatagc tatatgatat acatgttaaa agtatataaa ttaatattaa ggtgtttcgg 5940

aaaagaatat gattatagtc attcgaagtt gctttccagg tattacggta gaatattcgg 6000

taggaaatta aaaaaaattg atggtttgga ttttattatc gcacctgcag gttcctcaca 6060

aattgctttt ttaaaaacaa ccataccaat aatatatcta tcggatacaa catatgatca 6120

attaaaaagc tattatccga atttaaataa aaaaacaatt ataaatgatg aggatgcaag 6180

tttaatcgaa cgcaaggcta ttgaaaaagc aacagtagta tctttcccat ctaaatgggc 6240

aatggatttt tgcaggaatt attacagatt agattttgat aaattagttg aaataccatg 6300

gggggctaat ttatttgatg atattcactt tgctaataaa aatataattc aaaagaatag 6360

ttatacttgt cttttcttgg gagttgattg ggaaagaaaa ggtgggaaaa cagccttgaa 6420

agcaattgaa tatgtaaggc agttatatgg gatcgatgtt agactaaaaa tttgtggatg 6480

tactccgaat caaaagattt tacctacttg ggttgaatta attgataaag tagataaaaa 6540

taacgttgac gaatatcaga aattcatcga tgtgttatct aacgctgata tacttctttt 6600

accaaccatt gctgaatgtt atggaatggt attttgtgaa gctgctgctt ttggattgcc 6660

tgttgtcgct acagatacag gtggagtcag ttctatagtt atcaacgaaa ggacggggat 6720

attaattaaa gacccgttag actataagca ctttggaaat gcaattcata aaataattag 6780

ttccgtagag acttatcaaa actactccca aaacgcaaga attagatata ataatatatt 6840

gcattgggac aattgggcta aaaagataat tgagattatg tatgagcata agaatagaag 6900

aatcaaatag cacaaaaaga attatatgtt tatttatact ttttcttgtt ttccctgatt 6960

ttttgtttta tacattaggg gttgataatt ttagcatttc aacgataatc tcaattacat 7020

tgctttttgt ttttttaaga gctaaaaata tttgcaaaga taattttcta ataatagtag 7080

cgttattcat attgttgtgt tttaactgtt tgttaagtat gctatttaat attgaacagg 7140

ctttaacatt taaagttgta ctttcaatat atagcatctt aataatggca tacgtctcct 7200

cttgttatgc acagacgttg tggttatgtt ctgaagaaat acttaagaga tccgtctttt 7260

atttgttcgc atttctttgc cttattggca ttataagtat tcttttacag aagactgaga 7320

ttatacatga taaaagtatg attctttttc ctgaaccatc agcatttgca ttggttttta 7380

tacctatctt ttcattttgt ttatactata caagaggggg ggggctacta ttgctctata 7440

tattatcttt gggtattgcg ttaggtatcc agaatttaac aatgttggta ggcattgtga 7500

ttagtgtttt tgtgatgaaa aaaataacta taaggcaaac tattgttata cttttggggg 7560

catggatttt ttccatgata ttaagtgatt tagacatttc ttactataca tcgcggcttg 7620

attttaaaaa tactacgaac ctatcagtgc ttgtatatct ttcaggaatt gaaagagctt 7680

tcttgaattt tattacaagt tatggtcttg gtattggttt tcaacaaatg ggagtgaatg 7740

gggagatagg aatatatcaa caaattttag ctgaacttga tgcccctatg ttaaatatat 7800

acgatggctc atttatttct tctaagttaa tatctgagtt tggggttatt ggtgcattaa 7860

tgtgtatttt ctatttttt tatttttccc gattttatct gcgtttcaaa aaaagtaaga 7920

gatattcacc gcagtatatt ttagcatata gcttctacat gtgtttcttc atccctcttt 7980

ttatacgtgg tgctggttat ataaacccct atgtgtttat gttattttca tcaatatttt 8040

tgtgcaaata tcacgctaaa aatatcttga tgaaatctaa tgtccagata gctatataat 8100

agtagattat attatcatta tcacgtaaat tacatattaa tagcatatat gataactagg 8160

acataaataa tgtgcattaa aaaaaaactt aagttaatta aacgatatgg cctttatggt 8220

ggtcttaggc ttcttaaaga tatattctta acaaaatttt tattttgttc aaatgttagg 8280

attattagat ttccatgtta tattagaaaa gatggaagtg ttagttttgg aaaaggtttt 8340

acatcaggtg taggattacg agttgatgca tttatggatg ccgtagtttc cattggagaa 8400

aatgttcaaa ttaatgacta tgttcacatc gcggctatta ataatgtcat tattggtaga 8460

gatacattaa tagcaagtaa agtatttatt agtgatcata atcatggtat tttttctaaa 8520

tccgatatcc atagttcacc aactattatt ccttcgtcta ggccccttga atctgcacct 8580

gtgtatattg gagagcgtgt gtggattggc gaaaatgtga caatattacc aggtgcgtgt 8640

ataggtaatg gtgtagttat tggcgcaaac agtgttgttc gtggtgagat tcctaataat 8700

gtgatcattg ctggtgttcc agctaaaatt gttaaaaaat ataactatga gcgtatgcaa 8760

tgggaaagaa tatagttgta atatcggctg ttaattttac aaccggaggc ccctttaccg 8820

tactaaaaaa tgtgcttaca gcaactaaag atagagccga atgtaaattt attgcactgg 8880

ttcatagctc tgctgaacta atggaattat ttccgtgggt tgaatttata gagtatccag 8940

aagtcaagtc ttcgtgggtt aaaagattat atttcgaata tataacttgc aatagattat 9000

ctaaggtgat taaggcaact cattgggtat gcttacatga tattacagca aatgttagtg 9060

taccctatag atttgtttat tgccacaatc ctgcaccgtt ctataaatat ttaagctatc 9120

gagatattat aggagaacct aaattttatc ttttttatct tttttatggg cttttataca 9180

atatcaatat aaaaaagaac acagcagttt ttgttcagca gcagtggcta aaaaaagaat 9240

tcgaaaaaaa atataagtta aagaatgttg ttgttagtcg ccctgaagat atttgccctt 9300

ttgaaagtga tggtttggta agaaataata ataaaaagga tgtgaggata ttttacccag 9360

cagtgccccg tatatttaaa aactttgaag ttatcatacg tgctgcacaa atattacaag 9420

ataaaaatat tcatttttat cttacttttg atggtactga aaataagtat gcaaaaagaa 9480

tatataaatt agcttccgaa ctgaaaaatg tacatttcct cggttacctt aatgcaaccg 9540

agatggttaa cttttatcaa gattcagata ttatttgttt cccatcgaaa ctagaaacgt 9600

ggggattacc attatcagaa gctaaaacat acaaaaaatg gatatttgcg gcagacttac 9660

cttatgctca tgaagtttta tataactatt caaaaactag atattttcca tttgacgatg 9720

agaaaatact tgttcgctac atattagagt acacaagtaa aaatatgcat gaagatataa 9780

aaaatagtag ggtgaatttt aataatgatg cattgactgg ttttgaacag tttattgaat 9840

atatcctcaa ggggaactga cgtggtttat attataatcg tttcacatgg ccatgatgac 9900

tatatagaaa atcttttatt aaatttaaag ttgccctctg gaagatttaa aataatagtt 9960

cgtgataaca aaagttcaat ggttttaaaa aaaacatgcg aaaaaaattg cgtaacctat 10020

ttgcatggag ggcaatatgg atttggacat aataataaca tagcagtgtc atatataatt 10080

aataacttca tgattatgaa taatgattat tttctctttc ttaaccccga tgtattcata 10140

accagtgaaa gtttgattaa ttatgttgat tatataatta gtaatgatta taagtttagc 10200

acattatgtc tttatcgaga ttttactaaa agcaaacatg attattcaat acggagtttt 10260

ccaactttat atgattttct ttgttctttt ttattggggg tgaataaaag taaaattaag 10320

aaggaaaata tactttctga tactgtagtt gattggtgtg ctggctcatt tatgcttatt 10380

catgctttaa gtttcttaaa tgtgaatggt tttgatcaaa aatattttat gtattgtgaa 10440

gatattgacc tttgtatgcg tttaaaatta agtggagtag atctttacta tactccccat 10500

tttgatgcta ttcattatgc gcagcatgaa aatagaagaa tatttactaa agcatttcga 10560

tggcatataa ggagtattac gcgctacata ttacggaaac caattctttc ttataaaaac 10620

tatagaaaaa ttacatccga actggtaaag tga 10653

<210> 13
<211> 652
<212> PRT
<213> Artificial Sequence

<220>
<223> Detoxified EPA protein comprising 4 optimized N-glycosylation sequences

<400> 13

Gly Ser Gly Gly Gly Asp Gln Asn Ala Thr Gly Ser Gly Gly Gly Lys
1 5 10 15

Leu Ala Glu Glu Ala Phe Asp Leu Trp Asn Glu Cys Ala Lys Ala Cys
20 25 30

Val Leu Asp Leu Lys Asp Gly Val Arg Ser Ser Arg Met Ser Val Asp
35 40 45

Pro Ala Ile Ala Asp Thr Asn Gly Gln Gly Val Leu His Tyr Ser Met
50 55 60

Val Leu Glu Gly Gly Asn Asp Ala Leu Lys Leu Ala Ile Asp Asn Ala
65 70 75 80

Leu Ser Ile Thr Ser Asp Gly Leu Thr Ile Arg Leu Glu Gly Gly Val
85 90 95

```
Glu Pro Asn Lys Pro Val Arg Tyr Ser Tyr Thr Arg Gln Ala Arg Gly
            100                 105                 110

Ser Trp Ser Leu Asn Trp Leu Val Pro Ile Gly His Glu Lys Pro Ser
        115                 120                 125

Asn Ile Lys Val Phe Ile His Glu Leu Asn Ala Gly Asn Gln Leu Ser
    130                 135                 140

His Met Ser Pro Ile Tyr Thr Ile Glu Met Gly Asp Glu Leu Leu Ala
145                 150                 155                 160

Lys Leu Ala Arg Asp Ala Thr Phe Phe Val Arg Ala His Glu Ser Asn
                165                 170                 175

Glu Met Gln Pro Thr Leu Ala Ile Ser His Ala Gly Val Ser Val Val
            180                 185                 190

Met Ala Gln Ala Gln Pro Arg Arg Glu Lys Arg Trp Ser Glu Trp Ala
        195                 200                 205

Ser Gly Lys Val Leu Cys Leu Leu Asp Pro Leu Asp Gly Val Tyr Asn
    210                 215                 220

Tyr Leu Ala Gln Gln Arg Cys Asn Leu Asp Asp Thr Trp Glu Gly Lys
225                 230                 235                 240

Ile Tyr Arg Val Leu Ala Gly Asn Pro Ala Lys His Asp Leu Asp Ile
                245                 250                 255

Lys Asp Asn Asn Asn Ser Thr Pro Thr Val Ile Ser His Arg Leu His
            260                 265                 270

Phe Pro Glu Gly Gly Ser Leu Ala Ala Leu Thr Ala His Gln Ala Cys
        275                 280                 285

His Leu Pro Leu Glu Ala Phe Thr Arg His Arg Gln Pro Arg Gly Trp
    290                 295                 300

Glu Gln Leu Glu Gln Cys Gly Tyr Pro Val Gln Arg Leu Val Ala Leu
305                 310                 315                 320

Tyr Leu Ala Ala Arg Leu Ser Trp Asn Gln Val Asp Gln Val Ile Arg
                325                 330                 335

Asn Ala Leu Ala Ser Pro Gly Ser Gly Gly Asp Leu Gly Glu Ala Ile
            340                 345                 350

Arg Glu Gln Pro Glu Gln Ala Arg Leu Ala Leu Thr Leu Ala Ala Ala
        355                 360                 365
```

```
Glu Ser Glu Arg Phe Val Arg Gln Gly Thr Gly Asn Asp Glu Ala Gly
    370                 375                 380

Ala Ala Ser Ala Asp Val Val Ser Leu Thr Cys Pro Val Ala Lys Asp
385                 390                 395                 400

Gln Asn Arg Thr Lys Gly Glu Cys Ala Gly Pro Ala Asp Ser Gly Asp
                405                 410                 415

Ala Leu Leu Glu Arg Asn Tyr Pro Thr Gly Ala Glu Phe Leu Gly Asp
            420                 425                 430

Gly Gly Asp Val Ser Phe Ser Thr Arg Gly Thr Gln Asn Trp Thr Val
        435                 440                 445

Glu Arg Leu Leu Gln Ala His Arg Gln Leu Glu Glu Arg Gly Tyr Val
    450                 455                 460

Phe Val Gly Tyr His Gly Thr Phe Leu Glu Ala Ala Gln Ser Ile Val
465                 470                 475                 480

Phe Gly Gly Val Arg Ala Arg Ser Gln Asp Leu Asp Ala Ile Trp Arg
                485                 490                 495

Gly Phe Tyr Ile Ala Gly Asp Pro Ala Leu Ala Tyr Gly Tyr Ala Gln
            500                 505                 510

Asp Gln Glu Pro Asp Ala Arg Gly Arg Ile Arg Asn Gly Ala Leu Leu
        515                 520                 525

Arg Val Tyr Val Pro Arg Trp Ser Leu Pro Gly Phe Tyr Arg Thr Gly
    530                 535                 540

Leu Thr Leu Ala Ala Pro Glu Ala Ala Gly Glu Val Glu Arg Leu Ile
545                 550                 555                 560

Gly His Pro Leu Pro Leu Arg Leu Asp Ala Ile Thr Gly Pro Glu Glu
                565                 570                 575

Glu Gly Gly Arg Val Thr Ile Leu Gly Trp Pro Leu Ala Glu Arg Thr
            580                 585                 590

Val Val Ile Pro Ser Ala Ile Pro Thr Asp Pro Arg Asn Val Gly Gly
        595                 600                 605

Asp Leu Asp Pro Ser Ser Ile Pro Asp Lys Glu Gln Ala Ile Ser Ala
    610                 615                 620

Leu Pro Asp Tyr Ala Ser Gln Pro Gly Lys Pro Pro Arg Glu Asp Leu
625                 630                 635                 640
```

Lys Leu Gly Ser Gly Gly Gly Asp Gln Asn Ala Thr
645 650

<210> 14
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> N-glycosylation consensus sequence

<220>
<221> VARIANT
<222> (2)..(2)
<223> Xaa = any amino acid

<220>
<221> VARIANT
<222> (3)..(3)
<223> Xaa = Ser or Thr

<400> 14

Asn Xaa Xaa
1

<210> 15
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> N-glycosylation consensus sequence

<220>
<221> VARIANT
<222> (1)..(1)
<223> Xaa = Asp or Glu

<220>
<221> VARIANT
<222> (2)..(2)
<223> Xaa = independently selected from any natural amino acid except Pro

<220>
<221> VARIANT
<222> (4)..(4)
<223> Xaa = independently selected from any natural amino acid except Pro

<220>
<221> VARIANT
<222> (5)..(5)
<223> Xaa = Ser or Thr

<400> 15

Xaa Xaa Asn Xaa Xaa
1 5