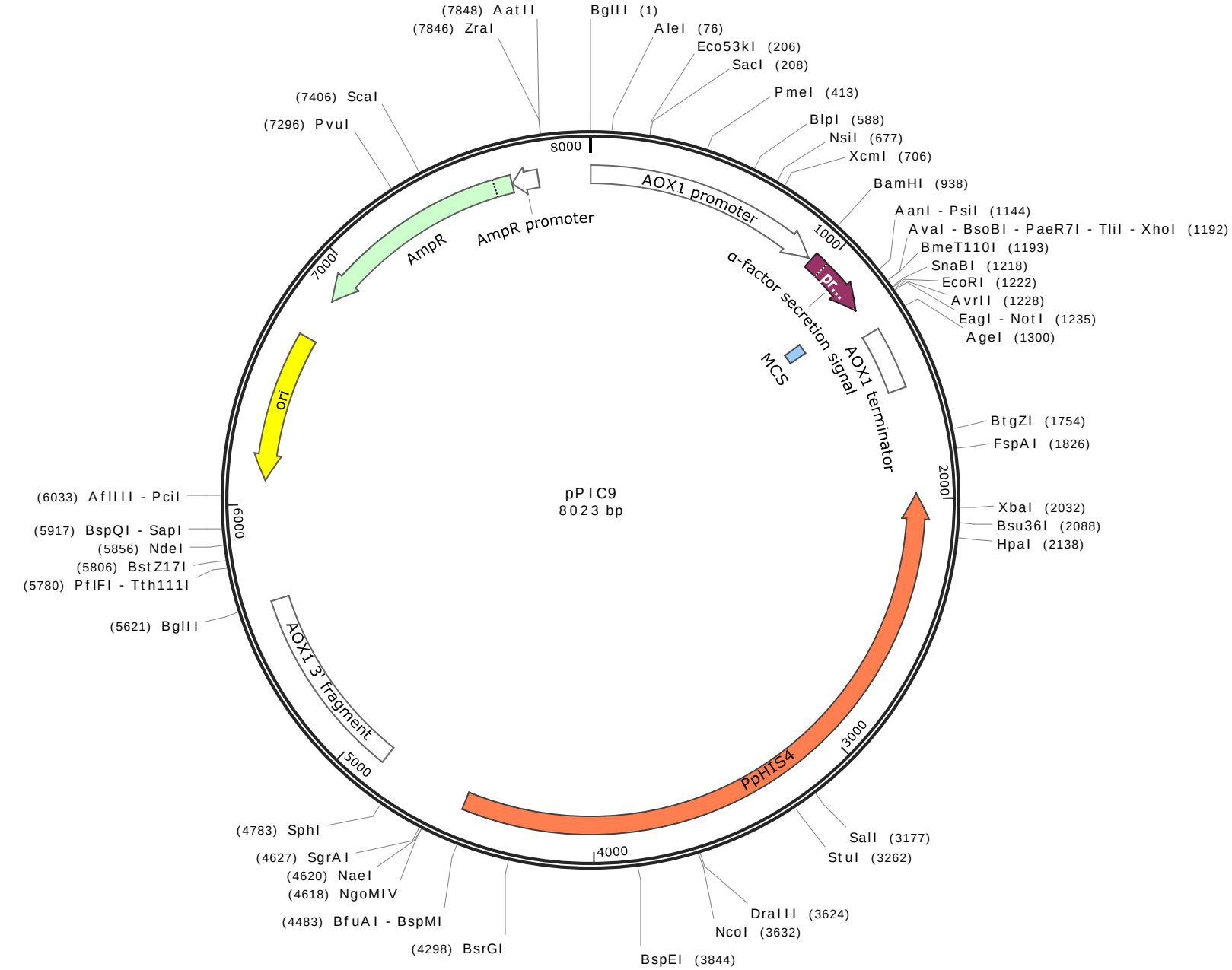
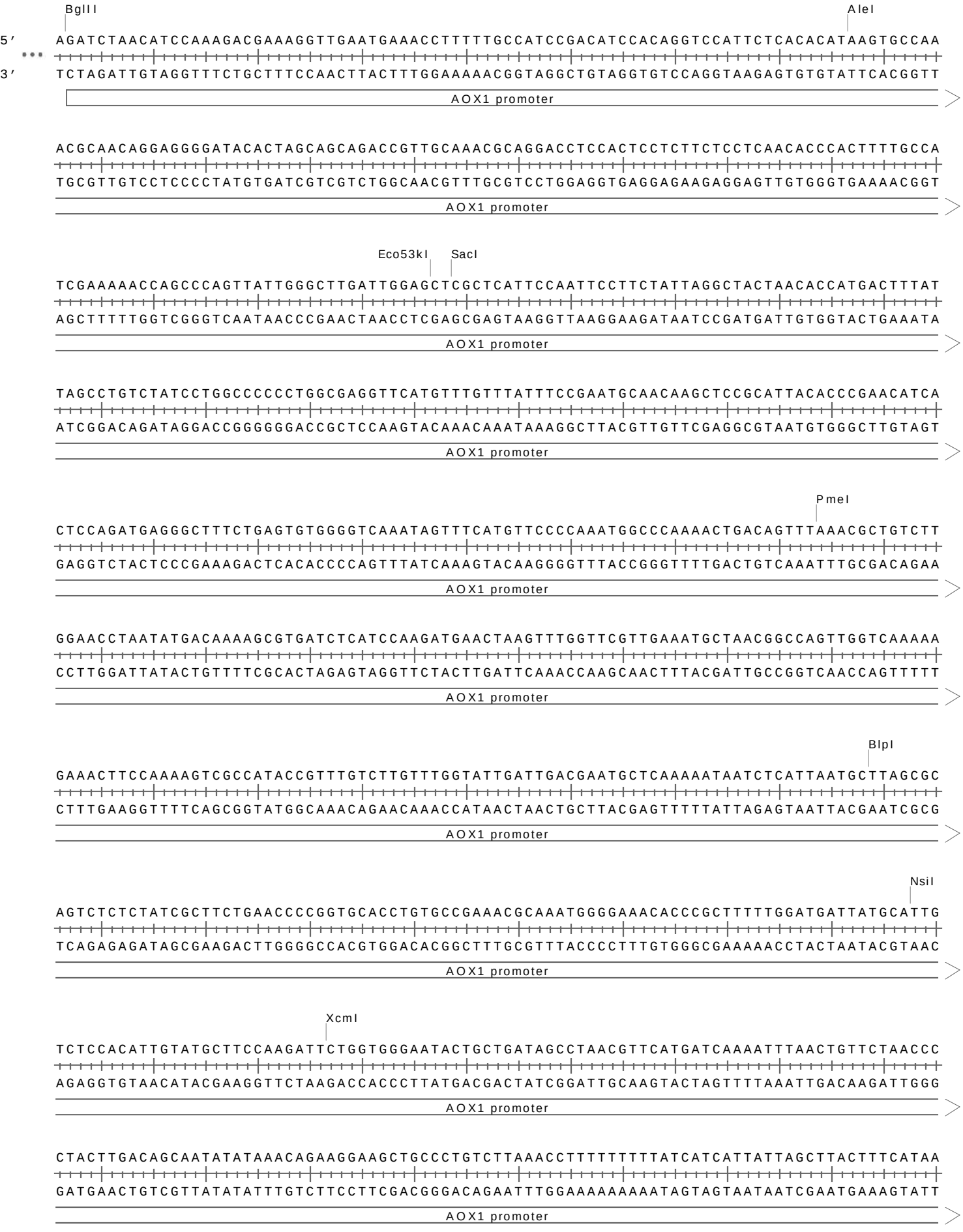
Pichia pastoris HIS4 vector for methanol-inducible expression of a secreted protein.

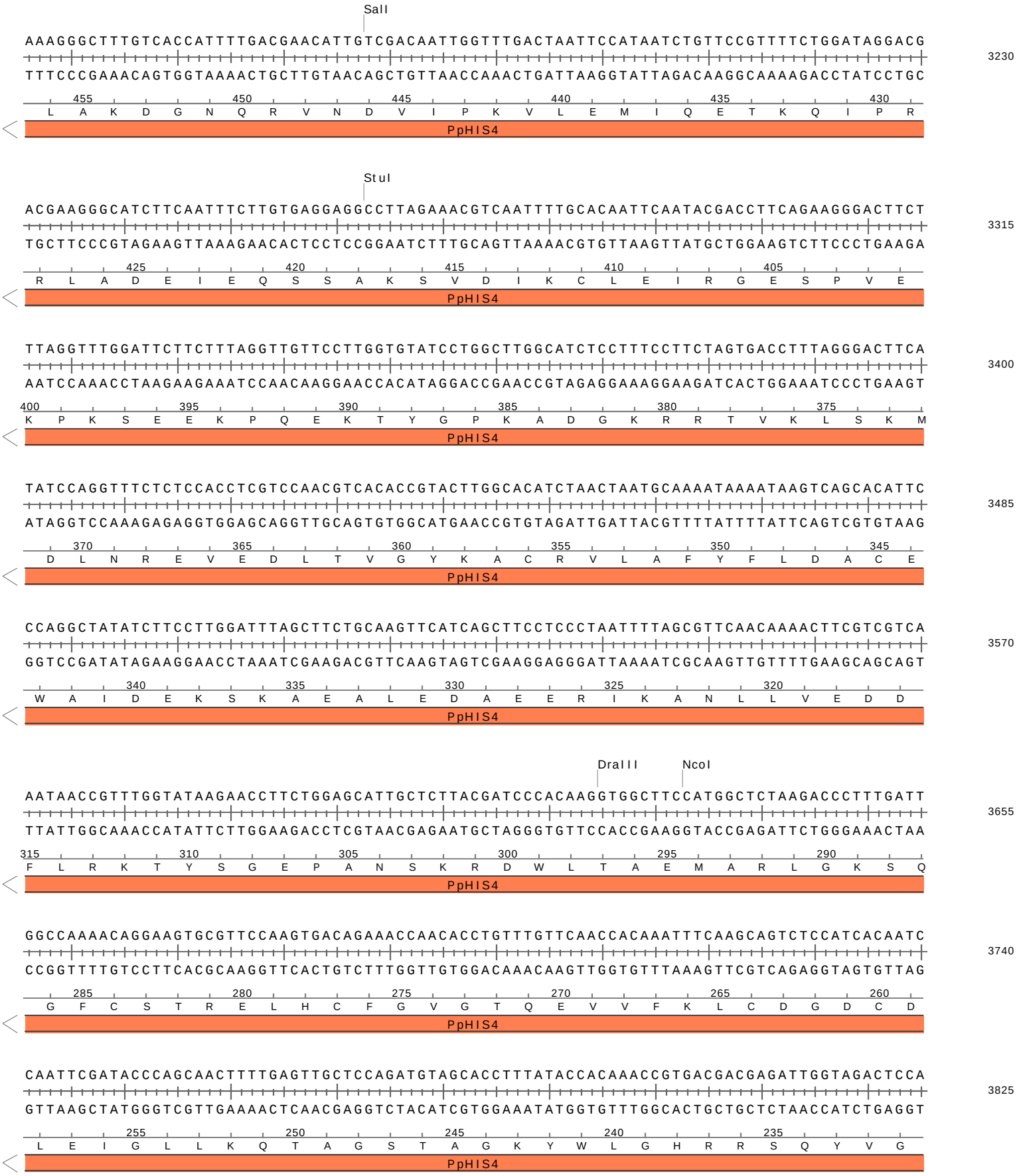


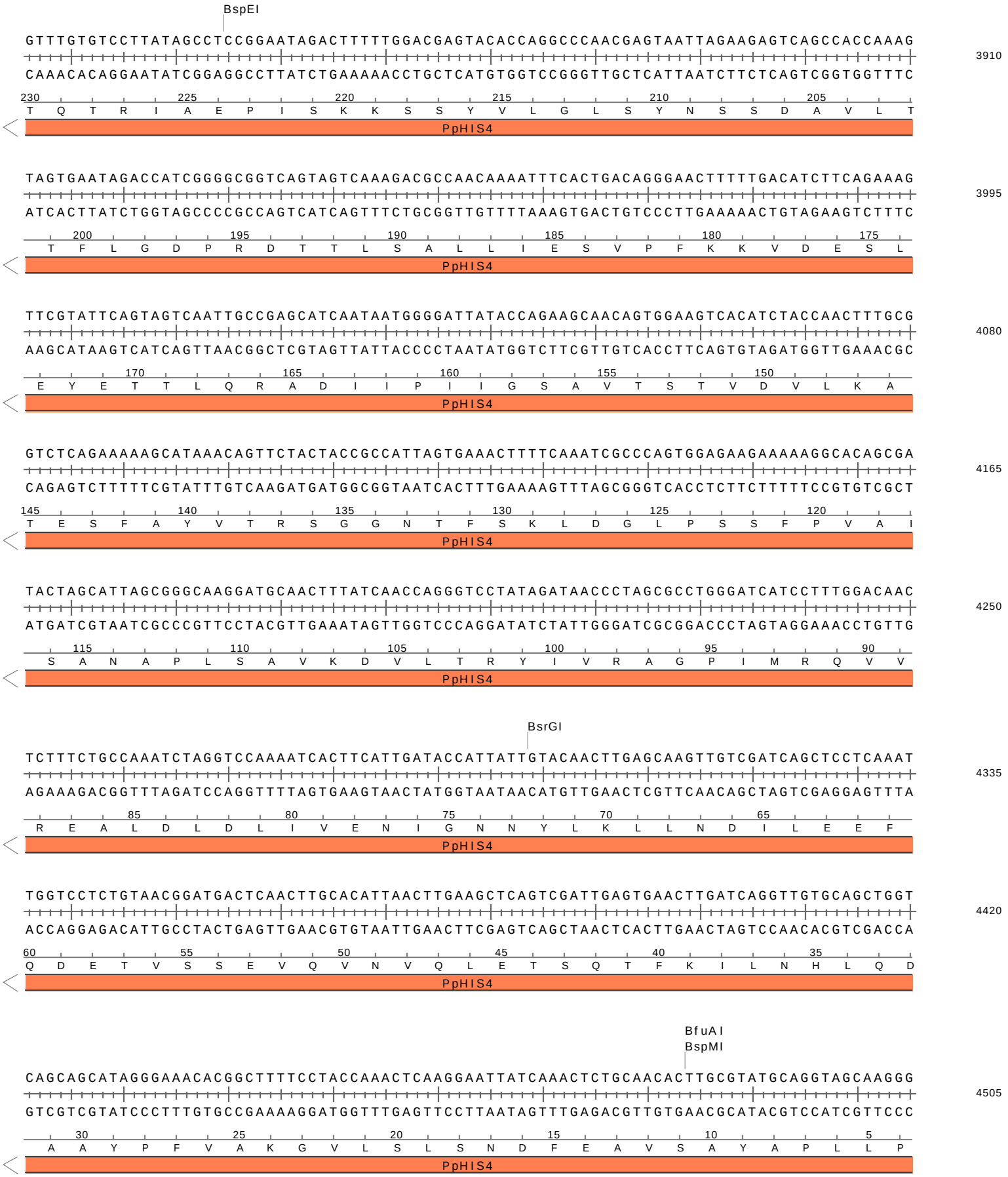

BglII
AleI
5′
3′
AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACACATAAGTGCCAA
TCTAGATTGTAGGTTTCTGCTTTCCAACTTACTTTGGAAAAACGGTAGGCTGTAGGTGTCCAGGTAAGAGTGTGTATTCACGGTT
85
AOX1 promoter
ACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACCCACTTTTGCCA
TGCGTTGTCCTCCCCTATGTGATCGTCGTCTGGCAACGTTTGCGTCCTGGAGGTGAGGAGAAGAGGAGTTGTGGGTGAAAACGGT
170
AOX1 promoter
Eco53kI
SacI
TCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCTACTAACACCATGACTTTAT
AGCTTTTTGGTCGGGTCAATAACCCGAACTAACCTCGAGCGAGTAAGGTTAAGGAAGATAATCCGATGATTGTGGTACTGAAATA
255
AOX1 promoter
TAGCCTGTCTATCCTGGCCCCCCTGGCGAGGTTCATGTTTGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCGAACATCA
ATCGGACAGATAGGACCGGGGGGACCGCTCCAAGTACAAACAAATAAAGGCTTACGTTGTTCGAGGCGTAATGTGGGCTTGTAGT
340
AOX1 promoter
PmeI
CTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCCAAATGGCCCAAAACTGACAGTTTAAACGCTGTCTT
GAGGTCTACTCCCGAAAGACTCACACCCCAGTTTATCAAAGTACAAGGGGTTTACCGGGTTTTGACTGTCAAATTTGCGACAGAA
425
AOX1 promoter
GGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAA
CCTTGGATTATACTGTTTTCGCACTAGAGTAGGTTCTACTTGATTCAAACCAAGCAACTTTACGATTGCCGGTCAACCAGTTTTT
510
AOX1 promoter
BlpI
GAAACTTCCAAAAGTCGCCATACCGTTTGTCTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGC
CTTTGAAGGTTTTCAGCGGTATGGCAAACAGAACAAACCATAACTAACTGCTTACGAGTTTTTATTAGAGTAATTACGAATCGCG
595
AOX1 promoter
NsiI
AGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTGGATGATTATGCATTG
TCAGAGAGATAGCGAAGACTTGGGGCCACGTGGACACGGCTTTGCGTTTACCCCTTTGTGGGCGAAAAACCTACTAATACGTAAC
680
AOX1 promoter
XcmI
TCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTCATGATCAAAATTTAACTGTTCTAACCC
AGAGGTGTAACATACGAAGGTTCTAAGACCACCCTTATGACGACTATCGGATTGCAAGTACTAGTTTTAAATTGACAAGATTGGG
765
AOX1 promoter
CTACTTGACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTATCATCATTATTAGCTTACTTTCATAA
GATGAACTGTCGTTATATATTTGTCTTCCTTCGACGGGACAGAATTTGGAAAAAAAAAATAGTAGTAATAATCGAATGAAAGTATT
850
AOX1 promoter

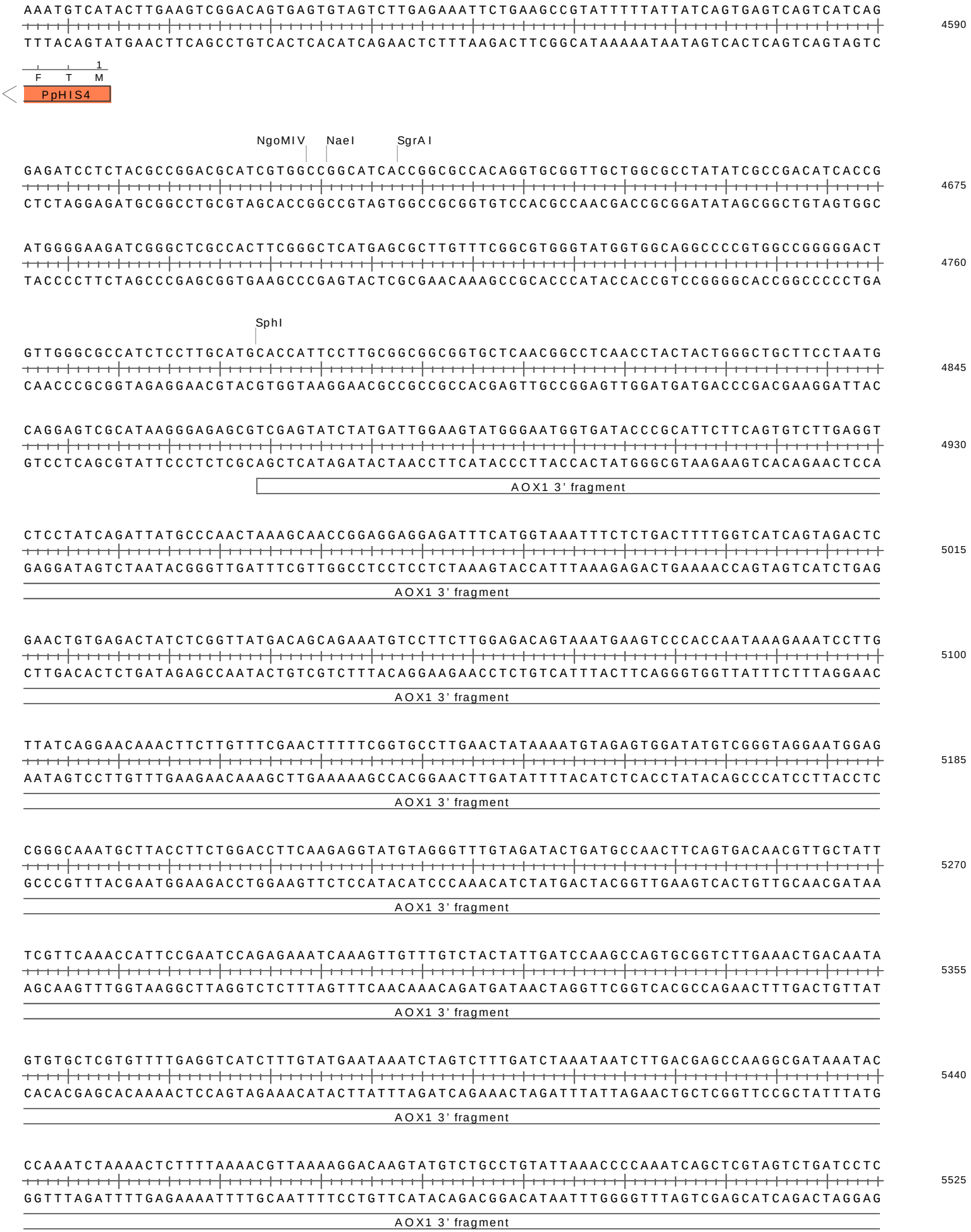

TTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGACAACTTGAGAAGATCAAAAAACAACTAATTATTCG
AACGCTGACCAAGGTTAACTGTTCGAAAACTAAAATTGCTGAAAATTGCTGTTGAACTCTTCTAGTTTTTTGTTGATTAATAAGC
935
AOX1 promoter
BamHI
AAGGATCCAAACGATGAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACT
TTCCTAGGTTTGCTACTCTAAAGGAAGTTAAAAATGACGTCAAAATAAGCGTCGTAGGAGGCGTAATCGACGAGGTCAGTTGTGA
1020
1 5 10 15 20
M R F P S I F T A V L F A A S S A L A A P V N T
presequence
pro region
AOX1 promoter
α-factor secretion signal
ACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTT
TGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCAATGAGTCTAAATCTTCCCCTAAAGCTACAACGACAAA
1105
25 30 35 40 45 50
T T E D E T A Q I P A E A V I G Y S D L E G D F D V A V
pro region
α-factor secretion signal
AanI
PsiI
TGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATC
ACGGTAAAAGGTTGTCGTGTTTATTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTTCTTCTTCCCCATAG
1190
55 60 65 70 75 80
L P F S N S T N N G L L F I N T T I A S I A A K E E G V S
pro region
α-factor secretion signal
TliI
XhoI
AvaI
BsoBI
PaeR7I
BmeT110I
SnaBI
EcoRI
AvrII
NotI
EagI
TCTCGAGAAAAGAGAGGCTGAAGCTTACGTAGAATTCCCTAGGGCGGCCGCGAATTAATTCGCCTTAGACATGACTGTTCCTCAG
AGAGCTCTTTTCTCTCCGACTTCGAATGCATCTTAAGGGATCCCGCCGGCGCTTAATTAAGCGGAATCTGTACTGACAAGGAGTC
1275
85
L E K R E A E A Y V E F P R A A A N *
pro region
(in frame with α-factor secretion signal)
α-factor secretion signal
MCS
AgeI
TTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAG
AAGTTCAACCCGTGAATGCTCTTCTGGCCAGAACGATCTAAGATTAGTTCTCCTACAGTCTTACGGTAAACGGACTCTCTACGTC
1360
AOX1 terminator
GCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCATTTTGTTTCTTCTCGTACGAGCTTGCT
CGAAGTAAAAACTATGAAAAAATAAACATTGGATATATCATATCCTAAAAAAAACAGTAAAACAAAGAAGAGCATGCTCGAACGA
1445
AOX1 terminator
CCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAGGGGTTTGGGAAAATCATTCGAGTTTGATGTTTTTCTTGGTATT
GGACTAGTCGGATAGAGCGTCGACTACTTATAGAACACCATCCCCAAACCCTTTTAGTAAGCTCAAACTACAAAAAGAACCATAA
1530
AOX1 terminator


TCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGAAGTTCGTTTGTGCAAGCTTATCGATAAGCTTTAATGCGGTAGTTTATC
AGGGTGAGGAGAAGTCTCATGTCTTCTAATTCACTCTTCAAGCAAACACGTTCGAATAGCTATTCGAAATTACGCCATCAAATAG
1615
AOX1 terminator
ACAGTTAAATTGCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGAT
TGTCAATTTAACGATTGCGTCAGTCCGTGGCACATACTTTAGATTGTTACGCGAGTAGCAGTAGGAGCCGTGGCAGTGGGACCTA
1700
BtgZI
GCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCCTCTTGCGGGATATCGTCCATTCCGACAGCATCGCCAGTCACTATG
CGACATCCGTATCCGAACCAATACGGCCATGACGGCCCGGAGAACGCCCTATAGCAGGTAAGGCTGTCGTAGCGGTCAGTGATAC
1785
FspAI
GCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCCC
CGCACGACGATCGCGATATACGCAACTACGTTAAAGATACGCGTGGGCAAGAGCCTCGTGACAGGCTGGCGAAACCGGCGGCGGG
1870
AGTCCTGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACCACACCCGTCCTGTGGATCTATCGAATCTAA
TCAGGACGAGCGAAGCGATGAACCTCGGTGATAGCTGATGCGCTAGTACCGCTGGTGTGGGCAGGACACCTAGATAGCTTAGATT
1955
XbaI
ATGTAAGTTAAAATCTCTAAATAATTAAATAAGTCCCAGTTTCTCCATACGAACCTTAACAGCATTGCGGTGAGCATCTAGACCT
TACATTCAATTTTAGAGATTTATTAATTTATTCAGGGTCAAAGAGGTATGCTTGGAATTGTCGTAACGCCACTCGTAGATCTGGA
2040
845 840 835 830
I L G L K E M R V K V A N R H A D L G
PpHIS4
Bsu36I
TCAACAGCAGCCAGATCCATCACTGCTTGGCCAATATGTTTCAGTCCCTCAGGAGTTACGTCTTGTGAAGTGATGAACTTCTGGA
AGTTGTCGTCGGTCTAGGTAGTGACGAACCGGTTATACAAAGTCAGGGAGTCCTCAATGCAGAACACTTCACTACTTGAAGACCT
2125
825 820 815 810 805 800
E V A A L D M V A Q G I H K L G E P T V D Q S T I F K Q F
PpHIS4
HpaI
AGGTTGCAGTGTTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAAAGTGTGGTTGGTACCGGAGGAGTAATCTCCACA
TCCAACGTCACAATTGAGGCGACATAACTGCCCGTATAGGCATGCAACCGTTTCACACCAACCATGGCCTCCTCATTAGAGGTGT
2210
795 790 785 780 775 770
T A T N V G S Y Q R A Y G Y T P L T H N T G S S Y D G C
PpHIS4
ACTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTGTTGTACTTGATCAACATAAGAAGAAGCATTCTCGATTTGCAGG
TGAGAGACCTCTCATCCGTGGTTGTTTGTGTCTAGGTCGCACAACATGAACTAGTTGTATTCTTCTTCGTAAGAGCTAAACGTCC
2295
765 760 755 750 745
S E P S Y A G V F V S G A H Q V Q D V Y S S A N E I Q L
PpHIS4
ATCAAGTGTTCAGGAGCGTACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTGCAATAC
TAGTTCACAAGTCCTCGCATGACTAACCTGTAAAGGTTTCGGACGAGCATCCAACGTTGGCTATCCCAACATCTCACACGTTATG
2380
740 735 730 725 720 715
I L H E P A Y Q N S M E L A Q E Y T A V S L T T S H A I C
PpHIS4


ACTTGCGTACAATTTCAACCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTCAATTCTGGCAAGCTCCTTGTCTGTCAT
TGAACGCATGTTAAAGTTGGGAACCGTTGACGTGTCGAACCAACACTTGTCGTAGAAGTTAAGACCGTTCGAGGAACAGACAGTA
2465
710 705 700 695 690 685
K R V I E V R P L Q V A Q N H V A D E I R A L E K D T M
PpHIS4
ATCGACAGCCAACAGAATCACCTGGGAATCAATACCATGTTCAGCTTGAGACAGAAGGTCTGAGGCAACGAAATCTGGATCAGCG
TAGCTGTCGGTTGTCTTAGTGGACCCTTAGTTATGGTACAAGTCGAACTCTGTCTTCCAGACTCCGTTGCTTTAGACCTAGTCGC
2550
680 675 670 665 660
D V A L L I V Q S D I G H E A Q S L L D S A V F D P D A
PpHIS4
TATTTATCAGCAATAACTAGAACTTCAGAAGGCCCAGCAGGCATGTCAATACTACACAGGGCTGATGTGTCATTTTGAACCATCA
ATAAATAGTCGTTATTGATCTTGAAGTCTTCCGGGTCGTCCGTACAGTTATGATGTGTCCCGACTACACAGTAAAACTTGGTAGT
2635
655 650 645 640 635 630
Y K D A I V L V E S P G A P M D I S C L A S T D N Q V M M
PpHIS4
TCTTGGCAGCAGTAACGAACTGGTTTCCTGGACCAAATATTTTGTCACACTTAGGAACAGTTTCTGTTCCGTAAGCCATAGCAGC
AGAACCGTCGTCATTGCTTGACCAAAGGACCTGGTTTATAAAACAGTGTGAATCCTTGTCAAAGACAAGGCATTCGGTATCGTCG
2720
625 620 615 610 605 600
K A A T V F Q N G P G F I K D C K P V T E T G Y A M A A
PpHIS4
TACTGCCTGGGCGCCTCCTGCTAGCACGATACACTTAGCACCAACCTTGTGGGCAACGTAGATGACTTCTGGGGTAAGGGTACCA
ATGACGGACCCGCGGAGGACGATCGTGCTATGTGAATCGTGGTTGGAACACCCGTTGCATCTACTGAAGACCCCATTCCCATGGT
2805
595 590 585 580 575
V A Q A G G A L V I C K A G V K H A V Y I V E P T L T G
PpHIS4
TCCTTCTTAGGTGGAGATGCAAAAACAATTTCTTTGCAACCAGCAACTTTGGCAGGAACACCCAGCATCAGGGAAGTGGAAGGCA
AGGAAGAATCCACCTCTACGTTTTTGTTAAAGAAACGTTGGTCGTTGAAACCGTCCTTGTGGGTCGTAGTCCCTTCACCTTCCGT
2890
570 565 560 555 550 545
D K K P P S A F V I E K C G A V K A P V G L M L S T S P L
PpHIS4
GAATTGCGGTTCCACCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGCAGACTACACCAGGGCAAGTCTC
CTTAACGCCAAGGTGGTCCTTATATCTCCGGTTGAAAGAGTTATCCAGAACGTTTTGCTCTCGTCTGATGTGGTCCCGTTCAGAG
2975
540 535 530 525 520 515
I A T G G P I Y L G V K E I P R A F R S C V V G P C T E
PpHIS4
AACTTGCAACGTCTCCGTTAGTTGAGCTTCATGGAATTTCCTGACGTTATCTATAGAGAGATCAATGGCTCTCTTAACGTTATCT
TTGAACGTTGCAGAGGCAATCAACTCGAAGTACCTTAAAGGACTGCAATAGATATCTCTCTAGTTACCGAGAGAATTGCAATAGA
3060
510 505 500 495 490
V Q L T E T L Q A E H F K R V N D I S L D I A R K V N D
PpHIS4
GGCAATTGCATAAGTTCCTCTGGGAAAGGAGCTTCTAACACAGGTGTCTTCAAAGCGACTCCATCAAACTTGGCAGTTAGTTCTA
CCGTTAACGTATTCAAGGAGACCCTTTCCTCGAAGATTGTGTCCACAGAAGTTTCGCTGAGGTAGTTTGAACCGTCAATCAAGAT
3145
485 480 475 470 465 460
P L Q M L E E P F P A E L V P T K L A V G D F K A T L E L
PpHIS4

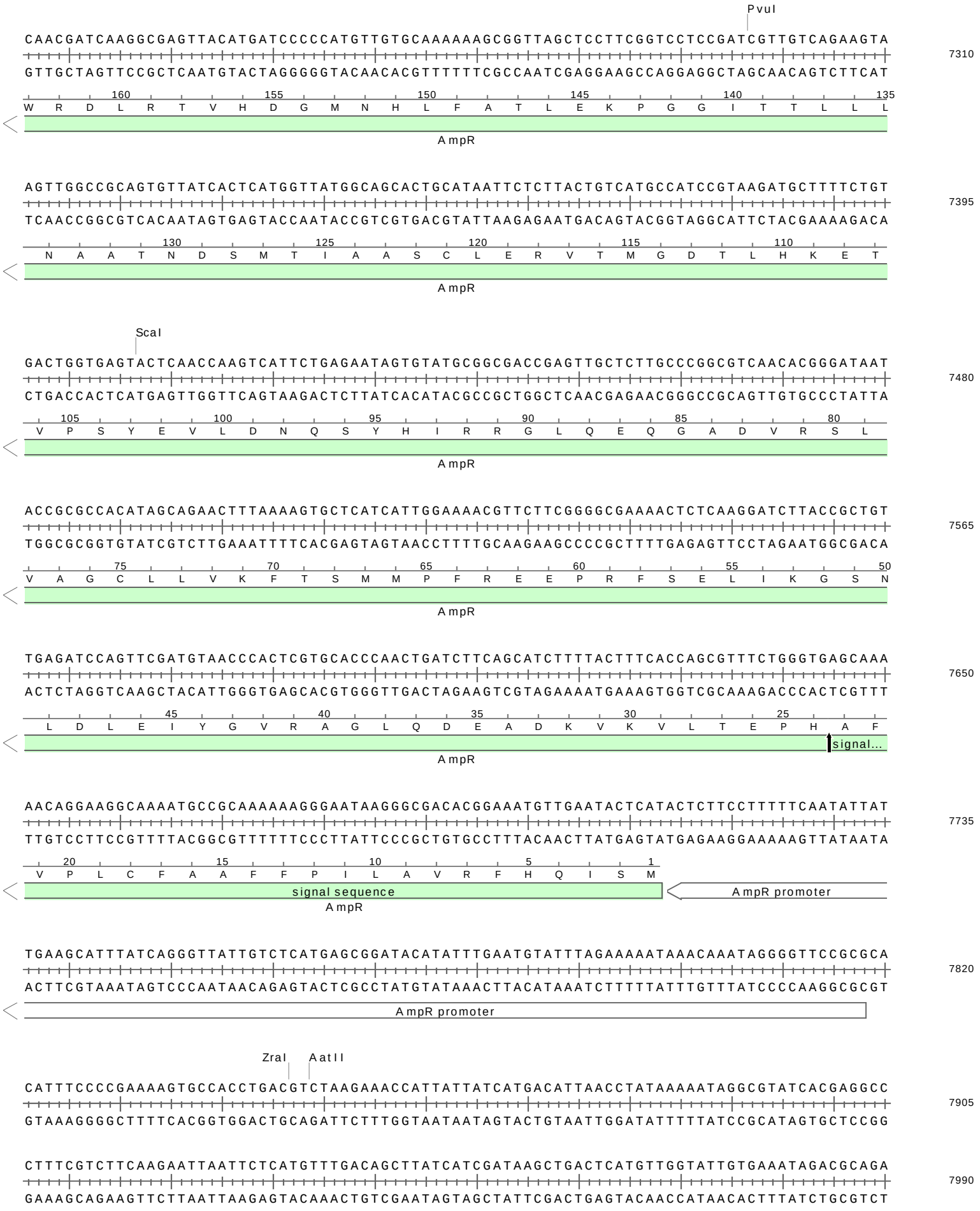

SalI
AAAGGGCTTTGTCACCATTTTGACGAACATTGTCGACAATTGGTTTGACTAATTCCATAATCTGTTCCGTTTTCTGGATAGGACG
TTTCCCGAAACAGTGGTAAAACTGCTTGTAACAGCTGTTAACCAAACTGATTAAGGTATTAGACAAGGCAAAAGACCTATCCTGC
3230
455 450 445 440 435 430
L A K D G N Q R V N D V I P K V L E M I Q E T K Q I P R
PpHIS4
StuI
ACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCCTTAGAAACGTCAATTTTGCACAATTCAATACGACCTTCAGAAGGGACTTCT
TGCTTCCCGTAGAAGTTAAAGAACACTCCTCCGGAATCTTTGCAGTTAAAACGTGTTAAGTTATGCTGGAAGTCTTCCCTGAAGA
3315
425 420 415 410 405
R L A D E I E Q S S A K S V D I K C L E I R G E S P V E
PpHIS4
TTAGGTTTGGATTCTTCTTTAGGTTGTTCCTTGGTGTATCCTGGCTTGGCATCTCCTTTCCTTCTAGTGACCTTTAGGGACTTCA
AATCCAAACCTAAGAAGAAATCCAACAAGGAACCACATAGGACCGAACCGTAGAGGAAAGGAAGATCACTGGAAATCCCTGAAGT
3400
400 395 390 385 380 375
K P K S E E K P Q E K T Y G P K A D G K R R T V K L S K M
PpHIS4
TATCCAGGTTTCTCTCCACCTCGTCCAACGTCACACCGTACTTGGCACATCTAACTAATGCAAAATAAAATAAGTCAGCACATTC
ATAGGTCCAAAGAGAGGTGGAGCAGGTTGCAGTGTGGCATGAACCGTGTAGATTGATTACGTTTTATTTTATTCAGTCGTGTAAG
3485
370 365 360 355 350 345
D L N R E V E D L T V G Y K A C R V L A F Y F L D A C E
PpHIS4
CCAGGCTATATCTTCCTTGGATTTAGCTTCTGCAAGTTCATCAGCTTCCTCCCTAATTTTAGCGTTCAACAAAACTTCGTCGTCA
GGTCCGATATAGAAGGAACCTAAATCGAAGACGTTCAAGTAGTCGAAGGAGGGATTAAAATCGCAAGTTGTTTTGAAGCAGCAGT
3570
340 335 330 325 320
W A I D E K S K A E A L E D A E E R I K A N L L V E D D
PpHIS4
DraIII
NcoI
AATAACCGTTTGGTATAAGAACCTTCTGGAGCATTGCTCTTACGATCCCACAAGGTGGCTTCCATGGCTCTAAGACCCTTTGATT
TTATTGGCAAACCATATTCTTGGAAGACCTCGTAACGAGAATGCTAGGGTGTTCCACCGAAGGTACCGAGATTCTGGGAAACTAA
3655
315 310 305 300 295 290
F L R K T Y S G E P A N S K R D W L T A E M A R L G K S Q
PpHIS4
GGCCAAAACAGGAAGTGCGTTCCAAGTGACAGAAACCAACACCTGTTTGTTCAACCACAAATTTCAAGCAGTCTCCATCACAATC
CCGGTTTTGTCCTTCACGCAAGGTTCACTGTCTTTGGTTGTGGACAAACAAGTTGGTGTTTAAAGTTCGTCAGAGGTAGTGTTAG
3740
285 280 275 270 265 260
G F C S T R E L H C F G V G T Q E V V F K L C D G D C D
PpHIS4
CAATTCGATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACCTTTATACCACAAACCGTGACGACGAGATTGGTAGACTCCA
GTTAAGCTATGGGTCGTTGAAAACTCAACGAGGTCTACATCGTGGAAATATGGTGTTTGGCACTGCTGCTCTAACCATCTGAGGT
3825
255 250 245 240 235
L E I G L L K Q T A G S T A G K Y W L G H R R S Q Y V G
PpHIS4

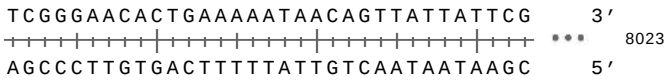

BspEI
GTTTGTGTCCTTATAGCCTCCGGAATAGACTTTTTGGACGAGTACACCAGGCCCAACGAGTAATTAGAAGAGTCAGCCACCAAAG
CAAACACAGGAATATCGGAGGCCTTATCTGAAAAACCTGCTCATGTGGTCCGGGTTGCTCATTAATCTTCTCAGTCGGTGGTTTC
3910
230 225 220 215 210 205
T Q T R I A E P I S K K S S Y V L G L S Y N S S D A V L T
PpHIS4
TAGTGAATAGACCATCGGGGCGGTCAGTAGTCAAAGACGCCAACAAAATTTCACTGACAGGGAACTTTTTGACATCTTCAGAAAG
ATCACTTATCTGGTAGCCCCGCCAGTCATCAGTTTCTGCGGTTGTTTTAAAGTGACTGTCCCTTGAAAAACTGTAGAAGTCTTTC
3995
200 195 190 185 180 175
T F L G D P R D T T L S A L L I E S V P F K K V D E S L
PpHIS4
TTCGTATTCAGTAGTCAATTGCCGAGCATCAATAATGGGGATTATACCAGAAGCAACAGTGGAAGTCACATCTACCAACTTTGCG
AAGCATAAGTCATCAGTTAACGGCTCGTAGTTATTACCCCTAATATGGTCTTCGTTGTCACCTTCAGTGTAGATGGTTGAAACGC
4080
170 165 160 155 150
E Y E T T L Q R A D I I P I I G S A V T S T V D V L K A
PpHIS4
GTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTGAAACTTTTCAAATCGCCCAGTGGAGAAGAAAAAGGCACAGCGA
CAGAGTCTTTTTCGTATTTGTCAAGATGATGGCGGTAATCACTTTGAAAAGTTTAGCGGGTCACCTCTTCTTTTTCCGTGTCGCT
4165
145 140 135 130 125 120
T E S F A Y V T R S G G N T F S K L D G L P S S F P V A I
PpHIS4
TACTAGCATTAGCGGGCAAGGATGCAACTTTATCAACCAGGGTCCTATAGATAACCCTAGCGCCTGGGATCATCCTTTGGACAAC
ATGATCGTAATCGCCCGTTCCTACGTTGAAATAGTTGGTCCCAGGATATCTATTGGGATCGCGGACCCTAGTAGGAAACCTGTTG
4250
115 110 105 100 95 90
S A N A P L S A V K D V L T R Y I V R A G P I M R Q V V
PpHIS4
BsrGI
TCTTTCTGCCAAATCTAGGTCCAAAATCACTTCATTGATACCATTATTGTACAACTTGAGCAAGTTGTCGATCAGCTCCTCAAAT
AGAAAGACGGTTTAGATCCAGGTTTTAGTGAAGTAACTATGGTAATAACATGTTGAACTCGTTCAACAGCTAGTCGAGGAGTTTA
4335
85 80 75 70 65
R E A L D L D L I V E N I G N N Y L K L L N D I L E E F
PpHIS4
TGGTCCTCTGTAACGGATGACTCAACTTGCACATTAACTTGAAGCTCAGTCGATTGAGTGAACTTGATCAGGTTGTGCAGCTGGT
ACCAGGAGACATTGCCTACTGAGTTGAACGTGTAATTGAACTTCGAGTCAGCTAACTCACTTGAACTAGTCCAACACGTCGACCA
4420
60 55 50 45 40 35
Q D E T V S S E V Q V N V Q L E T S Q T F K I L N H L Q D
PpHIS4
BfuAI
BspMI
CAGCAGCATAGGGAAACACGGCTTTTCCTACCAAACTCAAGGAATTATCAAACTCTGCAACACTTGCGTATGCAGGTAGCAAGGG
GTCGTCGTATCCCTTTGTGCCGAAAAGGATGGTTTGAGTTCCTTAATAGTTTGAGACGTTGTGAACGCATACGTCCATCGTTCCC
4505
30 25 20 15 10 5
A A Y P F V A K G V L S L S N D F E A V S A Y A P L L P
PpHIS4


AAATGTCATACTTGAAGTCGGACAGTGAGTGTAGTCTTGAGAAATTCTGAAGCCGTATTTTTATTATCAGTGAGTCAGTCATCAG
4590
TTTACAGTATGAACTTCAGCCTGTCACTCACATCAGAACTCTTTAAGACTTCGGCATAAAAATAATAGTCACTCAGTCAGTAGTC
1
F T M
PpHIS4
NgoMIV NaeI SgrA I
GAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCG
4675
CTCTAGGAGATGCGGCCTGCGTAGCACCGGCCGTAGTGGCCGCGGTGTCCACGCCAACGACCGCGGATATAGCGGCTGTAGTGGC
ATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACT
4760
TACCCCTTCTAGCCCGAGCGGTGAAGCCCGAGTACTCGCGAACAAAGCCGCACCCATACCACCGTCCGGGGCACCGGCCCCCTGA
SphI
GTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATG
4845
CAACCCGCGGTAGAGGAACGTACGTGGTAAGGAACGCCGCCGCCACGAGTTGCCGGAGTTGGATGATGACCCGACGAAGGATTAC
CAGGAGTCGCATAAGGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGGTGATACCCGCATTCTTCAGTGTCTTGAGGT
4930
GTCCTCAGCGTATTCCCTCTCGCAGCTCATAGATACTAACCTTCATACCCTTACCACTATGGGCGTAAGAAGTCACAGAACTCCA
AOX1 3' fragment
CTCCTATCAGATTATGCCCAACTAAAGCAACCGGAGGAGGAGATTTCATGGTAAATTTCTCTGACTTTTGGTCATCAGTAGACTC
5015
GAGGATAGTCTAATACGGGTTGATTTCGTTGGCCTCCTCCTCTAAAGTACCATTTAAAGAGACTGAAAACCAGTAGTCATCTGAG
AOX1 3' fragment
GAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAAATGAAGTCCCACCAATAAAGAAATCCTTG
5100
CTTGACACTCTGATAGAGCCAATACTGTCGTCTTTACAGGAAGAACCTCTGTCATTTACTTCAGGGTGGTTATTTCTTTAGGAAC
AOX1 3' fragment
TTATCAGGAACAAACTTCTTGTTTCGAACTTTTTCGGTGCCTTGAACTATAAAATGTAGAGTGGATATGTCGGGTAGGAATGGAG
5185
AATAGTCCTTGTTTGAAGAACAAAGCTTGAAAAAGCCACGGAACTTGATATTTTACATCTCACCTATACAGCCCATCCTTACCTC
AOX1 3' fragment
CGGGCAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGATGCCAACTTCAGTGACAACGTTGCTATT
5270
GCCCGTTTACGAATGGAAGACCTGGAAGTTCTCCATACATCCCAAACATCTATGACTACGGTTGAAGTCACTGTTGCAACGATAA
AOX1 3' fragment
TCGTTCAAACCATTCCGAATCCAGAGAAATCAAAGTTGTTTGTCTACTATTGATCCAAGCCAGTGCGGTCTTGAAACTGACAATA
5355
AGCAAGTTTGGTAAGGCTTAGGTCTCTTTAGTTTCAACAAACAGATGATAACTAGGTTCGGTCACGCCAGAACTTTGACTGTTAT
AOX1 3' fragment
GTGTGCTCGTGTTTTGAGGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTGACGAGCCAAGGCGATAAATAC
5440
CACACGAGCACAAAACTCCAGTAGAAACATACTTATTTAGATCAGAAACTAGATTTATTAGAACTGCTCGGTTCCGCTATTTATG
AOX1 3' fragment
CCAAATCTAAAACTCTTTTAAAACGTTAAAAGGACAAGTATGTCTGCCTGTATTAAACCCCAAATCAGCTCGTAGTCTGATCCTC
5525
GGTTTAGATTTTGAGAAAATTTTGCAATTTTCCTGTTCATACAGACGGACATAATTTGGGGTTTAGTCGAGCATCAGACTAGGAG
AOX1 3' fragment


ATCAACTTGAGGGGCACTATCTTGTTTTAGAGAAATTTGCGGAGATGCGATATCGAGAAAAAGGTACGCTGATTTTAAACGTGAA
5610
TAGTTGAACTCCCCGTGATAGAACAAAATCTCTTTAAACGCCTCTACGCTATAGCTCTTTTTCCATGCGACTAAAATTTGCACTT
AOX1 3' fragment
BglII
ATTTATCTCAAGATCTCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGC
5695
TAAATAGAGTTCTAGAGACGGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGTGTACGTCGAGGGCCTCTGCCAGTGTCG
AOX1 3' fragment
TTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACC
5780
AACAGACATTCGCCTACGGCCCTCGTCTGTTCGGGCAGTCCCGCGCAGTCGCCCACAACCGCCCACAGCCCCGCGTCGGTACTGG
PflFI
Tth111I
BstZ17I
NdeI
CAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTG
5865
GTCAGTGCATCGCTATCGCCTCACATATGACCGAATTGATACGCCGTAGTCTCGTCTAACATGACTCTCACGTGGTATACGCCAC
SapI
BspQI
TGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTC
5950
ACTTTATGGCGTGTCTACGCATTCCTCTTTTATGGCGTAGTCCGCGAGAAGGCGAAGGAGCGAGTGACTGAGCGACGCGAGCCAG
PciI
AflIII
GTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACA
6035
CAAGCCGACGCCGCTCGCCATAGTCGAGTGAGTTTCCGCCATTATGCCAATAGGTGTCTTAGTCCCCTATTGCGTCCTTTCTTGT
TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA
6120
ACACTCGTTTTCCGGTCGTTTTCCGGTCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGGACTGCT
ori
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC
6205
CGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGG
ori
CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAAT
6290
GAGCACGCGAGAGGACAAGGCTGGGACGGCGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTTA
ori
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCG
6375
CGAGTGCGACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGC
ori
CTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
6460
GACGCGGAATAGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCC
ori


ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTG
TAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTCCTGTCATAAAC
6545
ori
GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG
CATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTGTTTGGTGGCGACCATCGCC
6630
ori
TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAC
ACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACTAGAAAAGATGCCCCAGACTG
6715
ori
GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
CGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTCCTAGAAGTGGATCTAGGAAAATTTAATTT
6800
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC
TTACTTCAAAATTTAGTTAGATTTCATATATACTCATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAG
6885
285 280
W H K I L S A G I E
AmpR
AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGC
TCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCG
6970
275 270 265 260 255 250
A I Q R N R E D M T A Q S G T T Y I V V I R S P K G D P
AmpR
CCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC
GGGTCACGACGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCG
7055
245 240 235 230 225 220
G L A A I I G R S G R E G A G S K D A I F W G A P L A S R
AmpR
GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAA
CGTCTTCACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGGTCAATT
7140
215 210 205 200 195
L L P G A V K D A E M W D I L Q Q R S A L T L L E G T L
AmpR
TAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCC
ATCAAACGCGTTGCAACAACGGTAACGACGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGG
7225
190 185 180 175 170 165
L K R L T T A M A A P M T T D R E D N P I A E N L E P E
AmpR


PvuI
CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTA
7310
GTTGCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTCAT
160 155 150 145 140 135
W R D L R T V H D G M N H L F A T L E K P G G I T T L L L
AmpR
AGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT
7395
TCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAGAGAATGACAGTACGGTAGGCATTCTACGAAAAGACA
130 125 120 115 110
N A A T N D S M T I A A S C L E R V T M G D T L H K E T
AmpR
ScaI
GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAAT
7480
CTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCACATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTGTGCCCTATTA
105 100 95 90 85 80
V P S Y E V L D N Q S Y H I R R G L Q E Q G A D V R S L
AmpR
ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGT
7565
TGGCGCGGTGTATCGTCTTGAAATTTTCACGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACA
75 70 65 60 55 50
V A G C L L V K F T S M M P F R E E P R F S E L I K G S N
AmpR
TGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA
7650
ACTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGACTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTT
45 40 35 30 25
L D L E I Y G V R A G L Q D E A D K V K V L T E P H A F
signal...
AmpR
AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTAT
7735
TTGTCCTTCCGTTTTACGGCGTTTTTTCCCTTATTCCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAATA
20 15 10 5 1
V P L C F A A F F P I L A V R F H Q I S M
signal sequence
AmpR promoter
AmpR
TGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA
7820
ACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGT
AmpR promoter
ZraI AatII
CATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCC
7905
GTAAAGGGGCTTTTCACGGTGGACTGCAGATTCTTTGGTAATAATAGTACTGTAATTGGATATTTTTATCCGCATAGTGCTCCGG
CTTTCGTCTTCAAGAATTAATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTATTGTGAAATAGACGCAGA
7990
GAAAGCAGAAGTTCTTAATTAAGAGTACAAACTGTCGAATAGTAGCTATTCGACTGAGTACAACCATAACACTTTATCTGCGTCT


TCGGGAACACTGAAAAATAACAGTTATTATTCG 3′
8023
AGCCCTTGTGACTTTTTATTGTCAATAATAAGC 5′

Enzymes	Sites	
AanI	1	1144
AatII	1	7848
AflIII	1	6033
AgeI	1	1300
AleI	1	76
AvaI	1	1192
AvrII	1	1228
BamHI	1	938
BfuAI	1	4483
BglII	2	1 5621
BlpI	1	588
BmeT110I	1	1193
BsoBI	1	1192
BspEI	1	3844
BspMI	1	4483
BspQI	1	5917
BsrGI	1	4298
BstZ17I	1	5806
Bsu36I	1	2088
BtgZI	1	1754
DraIII	1	3624
EagI	1	1235
Eco53kI	1	206
EcoRI	1	1222
FspAI	1	1826
HpaI	1	2138
NaeI	1	4620
NcoI	1	3632
NdeI	1	5856
NgoMIV	1	4618
NotI	1	1235
NsiI	1	677
PaeR7I	1	1192
PciI	1	6033
PflFI	1	5780
PmeI	1	413
PsiI	1	1144
PvuI	1	7296
SacI	1	208
SalI	1	3177
SapI	1	5917
ScaI	1	7406
SgrAI	1	4627
SnaBI	1	1218
SphI	1	4783
StuI	1	3262
TliI	1	1192
Tth111I	1	5780
XbaI	1	2032
XcmI	1	706
XhoI	1	1192
ZraI	1	7846

	Feature	Location	Size (bp)	Type
✓	AOX1 promoter	2 .. 937	936	promoter

/gene = Pichia pastoris AOX1

/note = inducible promoter, regulated by methanol

✓	α-factor secretion signal	949 .. 1215	267	CDS

3 segments

/gene = MFα1

/product = N-terminal secretion signal from S. cerevisiae alpha-factor

/note = Cleavage by the Kex2 protease occurs after the dibasic KR sequence. The EA dipeptides are then removed by dipeptidyl aminopeptidase A.

/translation = MRFPSIFTAVLFAASSALA,APVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNGLLFINTTIASIAAKEEGVSLEKR,EAEA
89 amino acids = 9.3 kDa

✓	MCS	1192 .. 1241	50	misc_feature

/note = multiple cloning site

✓	AOX1 terminator	1321 .. 1567	247	terminator

/gene = Pichia pastoris AOX1

/note = transcription terminator for AOX1

✓	PpHIS4	1980 .. 4514	2535	CDS

/gene = Pichia pastoris HIS4

/product = multifunctional enzyme, required for histidine biosynthesis

/note = auxotrophic marker for Pichia pastoris

/translation = MTFPLLPAYASVAEFDNSLSLVGKAVFPYAADQLHNLIKFTQSTELQVNVQVESSVTEDQFEELIDNLLKLYNNGINEVILDLDLAE
RVVQRMIPGARVIYRTLVDKVASLPANASIAVPFSSPLGDLKSFTNGGSRTVYAFSETAKLVDVTSTVASGIIPIIDARQLTTEYELS
EDVKKFPVSEILLASLTTDRPDGLFTTLVADSSNYSLGLVYSSKKSIPEAIRTQTGVYQSRRHGLWYKGATSGATQKLLGIELDCDGD
CLKFVVEQTGVGFCHLERTSCFGQSKGLRAMEATLWDRKSNAPEGSYTKRLFDDEVLLNAKIREEADELAEAKSKEDIAWECADLF
YFALVRCAKYGVTLDEVERNLDMKSLKVTRRKGDAKPGYTKEQPKEESKPKEVPSEGRIELCKIDVSKASSQEIEDALRRPIQKTEQ
IMELVKPIVDNVRQNGDKALLELTAKFDGVALKTPVLEAPFPEELMQLPDNVKRAIDLSIDNVRKFHEAQLTETLQVETCPGVVCSR
FARPIEKVGLYIPGGTAILPSTSLMLGVPAKVAGCKEIVFASPPKKDGTLTPEVIYVAHKVGAKCIVLAGGAQAVAAMAYGTETVP
KCDKIFGPGNQFVTAAKMMVQNDTSALCSIDMPAGPSEVLVIADKYADPDFVASDLLSQAEHGIDSQVILLAVDMTDKELARIED
AVHNQAVQLPRVEIVRKCIAHSTTLSVATYEQALEMSNQYAPEHLILQIENASSYVDQVQHAGSVFVGAYSPESCGDYSSGTNH
TLPTYGYAROYSGVNTATFQKFITSQDVTPEGLKHIGQAVMDLAAVEGLDAHRNAVKVRMEKLGLL*
844 amino acids = 92.2 kDa

✓	AOX1 3' fragment	4869 .. 5625	757	misc_feature

/note = region downstream of Pichia pastoris AOX1 gene

✓	ori	6094 .. 6682	589	rep_origin

/direction = LEFT

/note = high-copy-number ColE1/pMB1/pBR322/pUC origin of replication

✓	AmpR	6853 .. 7713	861	CDS

2 segments

/gene = bla

/product = β-lactamase

/note = confers resistance to ampicillin, carbenicillin, and related antibiotics

/translation = MSIQHFRVALIPFFAAFCLPVFA,HPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRVDAGQE
QLGRRIHYSQNDLVEYSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPN
DERDTTMPAAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIY
TTGSQATMDERNRQIAEIGASLIKHW*
286 amino acids = 31.5 kDa

	Feature	Location	Size (bp)			Type
✓	AmpR promoter	7714 .. 7818	105	☐	←	promoter
	/gene = bla					

Description:	Pichia pastoris HIS4 vector for methanol-inducible expression of a secreted protein.
Created:	Tuesday, Feb 5, 2013
Last Modified:	Tuesday, Feb 5, 2013
Accession Number:	
Code Number:	
Sequence Author:	Invitrogen (Life Technologies)
DNA Type:	Synthetic DNA
Laboratory Host Organism:	Pichia pastoris
Bacterial Transformation Strain:	Unspecified Dam$^+$ Dcm$^+$ EcoKI$^+$
Comments:	
References:	