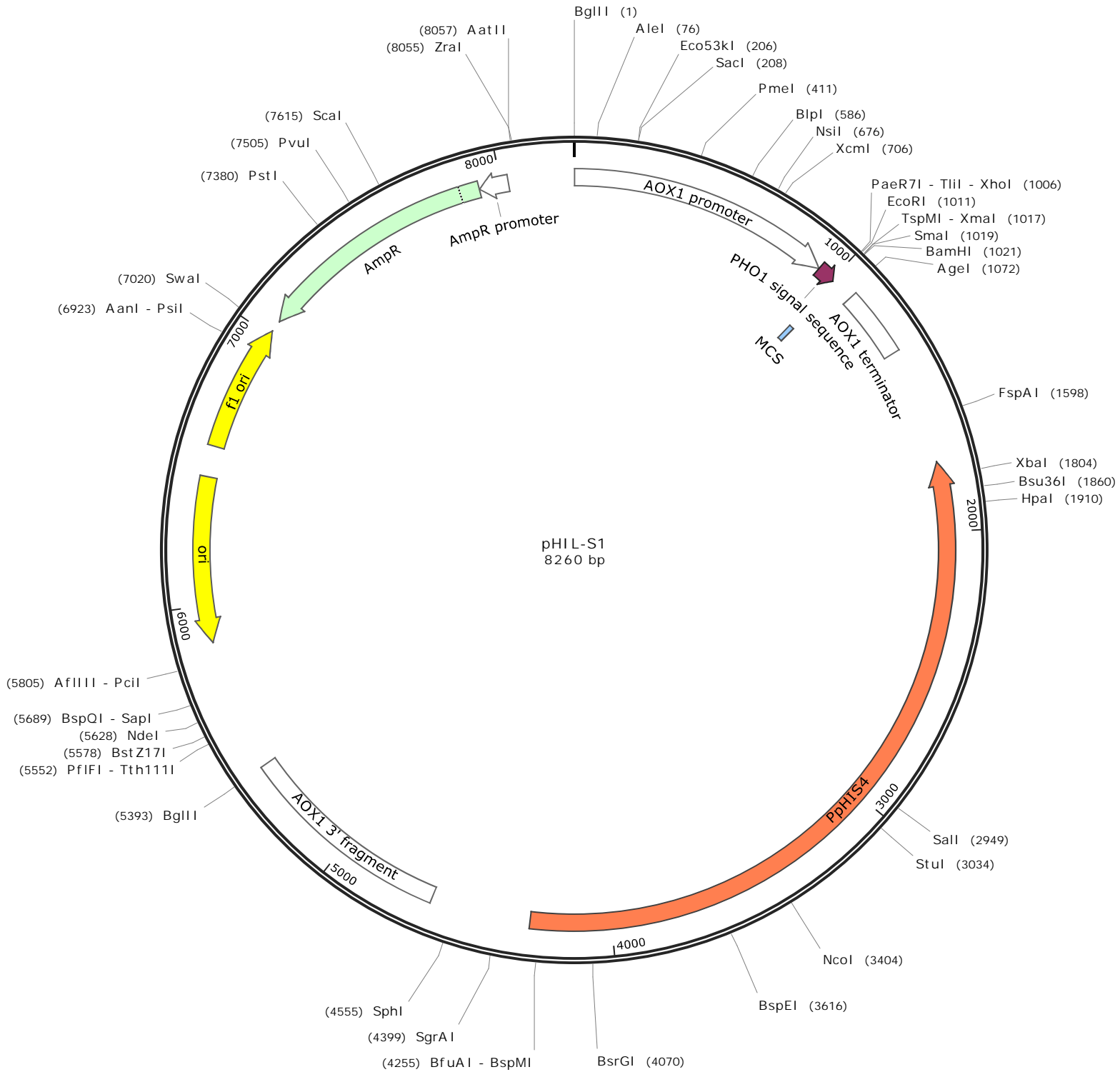


Pichia pastoris HIS4 vector for methanol-inducible expression of a secreted protein.





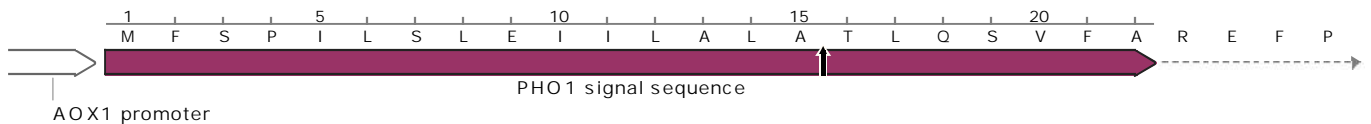
ATTGCGACTGGTTCCAATTGACAAGCTTTT GATTTT AACGACTTTT AACGACAAC TTGAGAAGATCAAAAAACA ACTAATTATTC  
TAACGCTGACCAAGGTTAACTGTTTCGAAAAC TAAAATTGCTGAAAATTGCTGTTGA ACTCTTCTAGTTTTTTGTTGATTAATAAG

935

AOX1 promoter

GAAACGATGTTCTCTCCAATTTTGTCTTGGAAATTATTTTAGCTTTGGCTACTTTGCAATCTGTCTTCGCTCGAGAATTCCCCG  
CTTTGCTACAAGAGAGGTTAAAACAGGAACCTTTAATAAAATCGAAAACCGATGAAACGTTAGACAGAAGCGAGCTCTTAAGGGGC

1020



MCS

GGATCCTTAGACATGACTGTTCCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAGAGGATGTC  
CCTAGGAATCTGTA CTGACAAGGAGTCAAGTTCAACCCGTGAATGCTCTTCTGGCCAGAACGATCTAAGATTAGTTCTCCTACAG  
G I L R H D C S S V Q V G H L R E D R S C (in frame with PHO 1 signal sequence)

1105

AOX1 terminator

MCS

AGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTTGTCA  
TCTTACGGTAAACGGACTCTCTACGTCCGAAGTAAAACCTATGAAAAAATAAACATTGGATATATCATATCCTAAAAAAAACAGT

1190

AOX1 terminator

TTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAGGGGTTTGGGAAAATCA  
AAAACAAAGAAGAGCATGCTCGAACGAGGACTAGTCGGATAGAGCGTCGACTACTTATAGAACACCATCCCCAAACCCTTTTAGT

1275

AOX1 terminator

TTCGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGAAGTTCGTTTGTGCAAGCTTAT  
AAGCTCAAAC TACAAAAGAACCATAAAGGGTGAGGAGAAGTCTCATGTCTTCTAATTCACTCTTCAAGCAAACACGTTCTGAATA

1360

AOX1 terminator

CGATAAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATC  
GCTATTCGAAATTACGCCATCAAATAGTGCAATTTAACGATTGCGTCAGTCCGTGGCACATACTTTAGATTGTTACGCGAGTAG

1445

GTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACTGCCGGGCTCTTGGGGATATCGTCC  
CAGTAGGAGCCGTGGCAGTGGGACCTACGACATCCGTATCCGAACCAATACGGCCATGACGGCCCCGAGAACGCCCTATAGCAGG

1530

FspAI

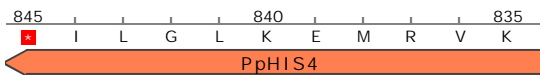
ATTCCGACAGCATCGCCAGTCACTATGGCGTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCGGAGC  
TAAGGCTGTCGTAGCGGTCAAGTACCGCACGACGATCGCGATATACGCAACTACGTTAAAGATACGCGTGGGCAAGAGCCTCG

1615

ACTGTCCGACCGCTTTGGCCGCCGCCAGTCCTGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGGATCATGGCGACCACA  
TGACAGGCTGGCGAAACCGGCGGGTCAAGGACGAGCGAAGCGATGAACCTCGGTGATAGCTGATGCGCTAGTACCGCTGGTGT

1700

CCCGTCTGTGGATCTATCGAATCTAAATGTAAGTTAAAATCTCTAAATAATTAATAAGTCCCAGTTTCTCCATACGAACCTTA  
 GGGCAGGACACCTAGATAGCTTAGATTTACATTCAATTTTAGAGATTTATTAATTTATTTCAGGGTCAAAGAGGTATGCTTGGAAAT



ACAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATCCATCACTGCTTGGCCAATATGTTTCAGTCCCTCAGGAGTTA  
 TGTCGTAACGCCACTCGTAGATCTGGAAGTTGTCGTCGGTCTAGGTAGTGACGAACCGGTTATACAAAGTCAGGGAGTCTCAAT

V A N R H A D L G E V A A L D M V A Q G I H K L G E P T V

CGTCTTGTGAAGTGATGAACCTTCTGGAAGGTTGCAGTGTTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAAAGTGTG  
 GCAGAACACTTCACTACTTGAAGACCTTCCAACGTCACAATTGAGGCGACATAACTGCCCGTATAGGCATGCAACCGTTTCACAC

D Q S T I F K Q F T A T N V G S Y Q R A Y G Y T P L T H

GTTGGTACCGGAGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTGTGTACTTGATCAACA  
 CAACCATGGCCTCCTCATTAGAGGTGTTGAGAGACCTCTCATCCGTGGTTGTTTGTGTCTAGGTGCGACAACATGAACTAGTTGT

N T G S S Y D G C S E P S Y A G V F V S G A H Q V Q D V

TAAGAAGAAGCATTCTCGATTTGCAGGATCAAGTGTTGAGGAGCGTACTGATTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAA  
 ATTCTTCTTCGTAAGAGCTAAACGTCCTAGTTTCAAGGTCCTCGCATGACTAACCTGTAAAGGTTTTCGGACGAGCATCCAACGTT

Y S S A N E I Q L I L H E P A Y Q N S M E L A Q E Y T A V

CCGATAGGGTTGTAGAGTGTGCAATACACTTGCCTACAATTTCAACCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTC  
 GGCTATCCCAACATCTCACACGTTATGTGAACGCATGTTAAAGTTGGGAACCGTTGACGTGTGCAACCAACACTTGTCTGTAAGAAG

S L T T S H A I C K R V I E V R P L Q V A Q N H V A D E

AATTCTGGCAAGCTCCTTGTCTGTCAATATCGACAGCCAACAGAATCACCTGGGAATCAATACCATGTTTCAGCTTGAGACAGAAGG  
 TTAAGACCGTTTCGAGGAACAGACAGTATAGCTGTGCGTTGTCTTAGTGGACCTTAGTTATGGTACAAGTCGAACTCTGTCTTCC

I R A L E K D T M D V A L L I V Q S D I G H E A Q S L L

TCTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAACTAGAACTTCAAGAGGCCAGCAGGCATGTCAATACTACACA  
 AGACTCCGTTGCTTTAGACCTAGTTCGCATAAATAGTCTGTTATTGATCTTGAAGTCTTCCGGGTCGTCGTTACAGTTATGATGTGT

D S A V F D P D A Y K D A I V L V E S P G A P M D I S C L

GGGCTGATGTGTCATTTGAACCATCATCTTGGCAGCAGTAACGAACCTGGTTTCTGGACCAAAATATTTTGTACACTTAGGAAC  
 CCCGACTACACAGTAAAACCTTGGTAGTAGAACCGTCGTCATTGCTTGACCAAAGGACCTGGTTTATAAAACAGTGTGAATCCTTG

A S T D N Q V M M K A A T V F Q N G P G F I K D C K P V

AGTTTCTGTTCCGTAAGCCATAGCAGCTACTGCCTGGGCGCCTCCTGCTAGCACGATACACTTAGCACCAACCTTGTGGGCAACG  
TCAAAGACAAGGCATTTCGGTATCGTCGATGACGGACCCGCGGAGGACGATCGTGCTATGTGAATCGTGGTTGGAACACCCGTTGC

2550

T E T G Y A M A A V A Q A G G A L V I C K A G V K H A V  
PpHIS4

TAGATGACTTCTGGGGTAAGGGTACCATCCTTCTTAGGTGGAGATGCAAAAACAATTTCTTTGCAACCAGCAACTTTGGCAGGAA  
ATCTACTGAAGACCCCATTCATGTTAGGAAGAATCCACCTCTACGTTTTTGTAAAGAAACGTTGGTCGTTGAAACCGTCTCTT

2635

Y I V E P T L T G D K K P P S A F V I E K C G A V K A P V  
PpHIS4

CACCCAGCATCAGGGAAGTGGAAAGGCAGAATTGCGGTTCCACCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTGCAAAAACG  
GTGGGTCGTAGTCCCTTCCACCTTCCGTCCTAACGCCAAGGTGGTCCTTATATCTCCGGTTGAAAGAGTTATCCAGAACGTTTTGC

2720

G L M L S T S P L I A T G G P I Y L G V K E I P R A F R  
PpHIS4

AGAGCAGACTACACCAGGGCAAGTCTCAACTTGCAACGTCTCCGTTAGTTGAGCTTCATGGAATTTCTGACGTTATCTATAGAG  
TCTCGTCTGATGTGGTCCCGTTCCAGAGTTGAACGTTGCAGAGGCCAATCAACTCGAAGTACCTTAAAGGACTGCAATAGATATCTC

2805

S C V V G P C T E V Q L T E T L Q A E H F K R V N D I S  
PpHIS4

AGATCAATGGCTCTCTTAACGTTATCTGGCAATTGCATAAGTTCTCTGGGAAAGGAGCTTCTAACACAGGTGTCTTCAAAGCGA  
TCTAGTTACCGAGAGAATTGCAATAGACCGTTAACGTATTCAAGGAGACCCCTTCTCGAAGATTGTGTCCACAGAAGTTTCGCT

2890

L D I A R K V N D P L Q M L E E P F P A E L V P T K L A V  
PpHIS4

CTCCATCAAACCTGGCAGTTAGTTCTAAAAGGGCTTTGTACCATTGACGAACATTGTCGACAATTTGGTTTACTAATTCCAT  
GAGGTAGTTTGAACCGTCAATCAAGATTTTCCCGAAACAGTGGTAAAACCTGCTTGTAAACAGCTGTTAACCAAACCTGATTAAGGTA

2975

G D F K A T L E L L A K D G N Q R V N D V I P K V L E M  
PpHIS4

AATCTGTTCCGTTTTCTGGATAGGACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCCTTAGAAACGTCAATTTTGCACAAT  
TTAGACAAGGCAAAAAGACCTATCCTGCTGCTTCCCGTAGAAGTTAAAGAACAACCTCCTCCGGAATCTTTGCAGTTAAAACGTGTTA

3060

I Q E T K Q I P R R L A D E I E Q S S A K S V D I K C L  
PpHIS4

TCAATACGACCTTCAGAAGGGACTTCTTTAGGTTTGGATTCTTCTTTAGGTTGTTCCCTTGGTGTATCCTGGCTTGGCATCTCCTT  
AGTTATGCTGGAAGTCTTCCCTGAAGAAATCCAAACCTAAGAAGAAATCCAACAAGGAACCCACATAGGACCGAACCGTAGAGGAA

3145

E I R G E S P V E K P K S E E K P Q E K T Y G P K A D G K  
PpHIS4

TCCTTCTAGTGACCTTAGGGACTTCATATCCAGGTTTCTCTCCACCTCGTCCAACGTACACCCGTACTTGGCACATCTAACTAA  
AGGAAGATCACTGGAAATCCCTGAAGTATAGGTCCAAAGAGAGGTGGAGCAGGTTGCAGTGTGGCATGAACCGTGTAGATTGATT

3230

R R T V K L S K M D L N R E V E D L T V G Y K A C R V L  
PpHIS4

TGCAAAATAAAATAAGTCAGCACATTCCCAGGCTATATCTTCCTTGGATTAGCTTCTGCAAGTTCATCAGCTTCCTCCCTAATT  
ACGTTTTATTATTATTCAGTCGTGTAAGGGTCCGATATAGAAGGAACCTAAATCGAAGACGTTCAAGTAGTCGAAGGAGGGATTAA

3315

A F Y F L D A C E W A I D E K S K A E A L E D A E E R I  
PpHIS4

TTAGCGTTCAACAAAACCTTCGTCGTCAAATAACCGTTTGGTATAAGAACCTTCTGGAGCATTGCTCTTACGATCCCACAAGGTGG  
AATCGCAAGTTGTTTTGAAGCAGCAGTTTATTGGCAAACCATATTCTTGGGAAGACCTCGTAACGAGAATGCTAGGGTGTTCCACC

3400

K A N L L V E D D F L R K T Y S G E P A N S K R D W L T A  
PpHIS4

NcoI

CTTCCATGGCTCTAAGACCCCTTTGATTGGCCAAAACAGGAAGTGC GTTCCAAGTGACAGAAACCAACACCTGTTTGTTCACCAC  
GAAGGTACCGAGATTCTGGGAAACTAACCGGTTTTGTCTTTCACGCAAGGTTCACTGTCTTTGGTTGTGGACAAAACAAGTTGGTG

3485

E M A R L G K S Q G F C S T R E L H C F G V G T Q E V V  
PpHIS4

AAATTTCAAGCAGTCTCCATCACAATCCAATTCGATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACCTTTATAACCACAAA  
TTTAAAGTTCGTCAGAGGTAGTGTTAGGTTAAGCTATGGGTCGTTGAAAACCTCAACGAGGTCTACATCGTGGAAATATGGTGTTT

3570

F K L C D G D C D L E I G L L K Q T A G S T A G K Y W L  
PpHIS4

BspEI

CCGTGACGACGAGATTGGTAGACTCCAGTTTTGTGTCTTATAGCCTCCGGAATAGACTTTTTGGACGAGTACACCAGGCCCAACG  
GGCACTGCTGCTCTAACCATCTGAGGTCAAACACAGGAATATCGGAGGCCTTATCTGAAAAACCTGCTCATGTGGTCCGGGTTGC

3655

G H R R S Q Y V G T Q T R I A E P I S K K S S Y V L G L S  
PpHIS4

AGTAATTAGAAGAGTCAGCCACCAAAGTAGTGAATAGACCATCGGGGCGGTCAGTAGTCAAAGACGCCAACAAAATTTCACTGAC  
TCATTAATCTTCTCAGTCGGTGGTTTTATCACTTATCTGGTAGCCCCGCCAGTCATCAGTTTTCTGCGGTTGTTTTAAAGTGACTG

3740

Y N S S D A V L T T F L G D P R D T T L S A L L I E S V  
PpHIS4

AGGGAACTTTTTGACATCTTCAGAAAAGTTCGTATTCAGTAGTCAATTGCCGAGCATCAATAATGGGGATTATAACCAGAAGCAACA  
TCCCTTGAAAAACTGTAGAAGTCTTTCAAGCATAAGTCATCAGTTAACGGCTCGTAGTTATTACCCTAATATGGTCTTCGTTGT

3825

P F K K V D E S L E Y E T T L Q R A D I I P I I G S A V  
PpHIS4

GTGGAAGTCACATCTACCAACTTTGCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTGAAACTTTTTCAAATCGC  
CACCTTCAGTGATAGTGGTTGAAACGCCAGAGTCTTTTTCGTATTTGTCAAGATGATGGCGGTAATCACTTTGAAAAGTTTAGCG

3910

T S T V D V L K A T E S F A Y V T R S G G N T F S K L D G  
PpHIS4

CCAGTGGAGAAGAAAAAGGCACAGCGATACTAGCATTAGCGGGCAAGGATGCAACTTTATCAACCAGGGTCTATAGATAACCCT  
GGTCACCTCTTCTTTTTCCGTGTCGCTATGATCGTAATCGCCCGTTTCTACGTTGAAATAGTTGGTCCCAGGATATCTATTGGGA

3995

L P S S F P V A I S A N A P L S A V K D V L T R Y I V R  
PpHIS4

BsrGI

AGCGCCTGGGATCATCCTTTGGACAACCTCTTTCTGCCAAATCTAGGTCCAAAATCACTTCATTGATACCATTATTGTACAACCTTG  
 TCGCGGACCCTAGTAGGAAACCTGTTGAGAAAGACGGTTTAGATCCAGGTTTTAGTGAAGTAACTATGGTAATAACATGTTGAAC

95 90 85 80 75 70

A G P I M R Q V V R E A L D L D L I V E N I G N N Y L K

PpHIS4

4080

AGCAAGTTGTCGATCAGCTCCTCAAATTGGTCCTCTGTAACGGATGACTCAACTTGCACATTAACCTGAAGCTCAGTCGATTGAG  
 TCGTTCAACAGCTAGTCGAGGAGTTAACCAGGAGACATTGCCTACTGAGTTGAACGTGTAATTGAACTTCGAGTCAGCTAACTC

65 60 55 50 45

L L N D I L E E F Q D E T V S S E V Q V N V Q L E T S Q T

PpHIS4

4165

TGAACTTGATCAGGTTGTGCAGCTGGTCAGCAGCATAGGGAAACACGGCTTTTTCTACCAAACCTCAAGGAATTATCAAACCTCTGC  
 ACTTGAACCTAGTCCAACACGTCGACCAGTCGTCGTATCCCTTTGTGCCGAAAAGGATGGTTTGAGTTCCTTAATAGTTTGAGACG

40 35 30 25 20 15

F K I L N H L Q D A A Y P F V A K G V L S L S N D F E A

PpHIS4

4250

BfuAI  
BspMI

AACACTTGCGTATGCAGGTAGCAAGGGAAATGTCATACTTGAAGTCGGACAGTGAGTGTAGTCTTGAGAAATTCTGAAGCCGTAT  
 TTGTGAACGCATACGTCCATCGTTCCCTTTACAGTATGAACTTCAGCCTGTCACCTCACATCAGAACTCTTTAAGACTTCGGCATA

10 5 1

V S A Y A P L L P F T M

PpHIS4

4335

SgrAI

TTTTATTATCAGTGAGTCAGTCATCAGGAGATCCTCTACGCCGGACGCATCGTGGCCGGGCATCACCGGCCACAGGTGCGGTTG  
 AAAATAATAGTCACTCAGTCAGTAGTCCTCTAGGAGATGCGGCCTGCGTAGCACCGGCCGTAGTGGCCGCGGTGTCCACGCCAAC

CTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTTCGGCGTGGGTAT  
 GACCGCGGATATAGCGGCTGTAGTGGCTACCCCTTCTAGCCCGAGCGGTGAAGCCCGAGTACTCGCGAACAAAGCCGCACCCATA

4420

4505

SphI

GGTGGCAGGCCCGTGGCCGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCTTGCGGCGGCGGTGCTCAACGGCCTC  
 CCACCGTCCGGGGCACCGGCCCTGACAACCCGCGGTAGAGGAACGTACGTGGTAAGGAACGCCGCCACGAGTTGCCGGAG

AACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGGTGA  
 TTGGATGATGACCCGACGAAGGATTACGTCCTCAGCGTATTCCCTCTCGCAGCTCATAGATACTAACCTTCATAACCTTACCACT

4590

4675

AOX1 3' fragment

TACCCGCATTCTTCAGTGTCTTGAGGTCTCCTATCAGATTATGCCAACTAAAGCAACCGGAGGAGGAGATTTTCATGGTAAATTT  
 ATGGGCGTAAGAAGTCACAGAACTCCAGAGGATAGTCTAATACGGTTGATTTTCGTTGGCCTCCTCCTCTAAAGTACCATTTAAA

AOX1 3' fragment

4760

CTCTGACTTTTGGTCATCAGTAGACTCGAACTGTGAGACTATCTCGGTTATGACAGCAGAAATGTCCTTCTTGGAGACAGTAAAT  
 GAGACTGAAAACAGTAGTCATCTGAGCTTGACACTCTGATAGAGCCAATACTGTCGTCTTTACAGGAAGAACCTCTGTCAATTA

AOX1 3' fragment

4845

GAAGTCCCACCAATAAAGAAATCCTTGTTATCAGGAACAAACTTCTTGTTTCGAACTTTTTTCGGTGCCTTGAACATAAAAATGTA  
CTTCAGGGTGGTTATTTCTTTAGGAACAATAGTCCTTGTTTGAAGAACAAGCTTGAAAAAGCCACGGAACCTTGATATTTTACAT

4930

AOX1 3' fragment

GAGTGGATATGTCGGGTAGGAATGGAGCGGGCAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGAT  
CTCACCTATACAGCCCATCCTTACCTCGCCCGTTTACGAATGGAAGACCTGGAAGTTCTCCATACATCCCAAACATCTATGACTA

5015

AOX1 3' fragment

GCCAACTTCAGTGACAACGTTGCTATTTTCGTTCAAACCATTCGGAATCCAGAGAAATCAAAGTTGTTTGTCTACTATTGATCCAA  
CGGTTGAAGTCACTGTTGCAACGATAAAGCAAGTTTGGTAAGGCTTAGGTCTCTTTAGTTTCAACAAACAGATGATAACTAGGTT

5100

AOX1 3' fragment

GCCAGTGCGGTCTTGAAACTGACAATAGTGTGCTCGTGTGTTTGGAGTCATCTTTGTATGAATAAATCTAGTCTTTGATCTAAATA  
CGGTACAGCCAGAAGCTTACTGTTATCACACGAGCACAAAACCTCCAGTAGAAACATACTTATTTAGATCAGAAACTAGATTTAT

5185

AOX1 3' fragment

ATCTTGACGAGCCAAGGCGATAAATACCCAAATCTAAAACCTTTTTAAAACGTTAAAAGGACAAGTATGTCTGCCTGTATTAAC  
TAGAACTGCTCGGTTCCGCTATTTATGGGTTTAGATTTTGGAAAATTTGCAATTTTCTGTTTCATACAGACGGACATAATTTG

5270

AOX1 3' fragment

CCCAAATCAGCTCGTAGTCTGATCCTCATCAACTTGAGGGGCACTATCTTGTTTTAGAGAAATTTGCGGAGATGCGATATCGAGA  
GGGTTTAGTTCGAGCATCAGACTAGGAGTAGTTGAACTCCCCGTGATAGAACAAAATCTCTTTAAACGCCTCTACGCTATAGCTCT

5355

AOX1 3' fragment

AAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACA  
TTTTCCATGCGACTAAAATTTGCACTTTAAATAGAGTTCTAGAGACGGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGT

5440

AOX1 3' fragment

Bgl I

CATGCAGCTCCCGGAGACGGTACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT  
GTACGTCGAGGGCCTCTGCCAGTGTGCAACAGACATTCGCCTACGGCCCTCGTCTGTTTCGGGCAGTCCCAGCAGTCGCCACAA

5525

PfI  
Tth111

BstZ171

GGCGGGTGTGCGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGAT  
CCGCCCCACAGCCCCGCGTGGTACTGGGTCAGTGCATCGCTATCGCCTCACATATGACCGAATTGATACGCCGTAGTCTCGTCTA

5610

Nde I

Sap I  
BspQ I

TGTA CTGAGAGTGACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGGCCTCTTCCGCTTCC  
ACATGACTCTCACGTGGTATACGCCACACTTTATGGCGTGTCTACGCATTCCTCTTTTATGGCGTAGTCCGCGAGAAAGGCGAAGG

5695

TCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGGCAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG  
AGCGAGTGACTGAGCGACGCGAGCCAGCAAGCCGACGCCGCTCGCCATAGTCGAGTGAGTTTCCGCCATTATGCCAATAGGTGTC

5780

Pci I  
Afi III

AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTT  
TTAGTCCCCTATTGCGTCTTTCTTGTACACTCGTTTTCCGGTCGTTTTCCGGTCCTTGGCATTTCGCGCGCAACGACCGCAA

5865



TTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGA  
AAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTTCACTCCACCCTTTGGGCTGTCCTGATATTTCT

5950

ori

TACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCAGGATACCTGTCCGCCTTTCTCC  
ATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGGCGAATGGCCTATGGACAGGCGGAAAGAGG

6035

ori

CTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGT  
GAAGCCCTTCGCACCAGGAAAGAGTTACGAGTGCACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCCGACCCGACACA

6120

ori

GCACGAACCCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCG  
CGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCATTGATAGCAGAAGTCAAGTGGGCCATTCTGTGCTGAATAGC

6205

ori

CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACT  
GGTGACCGTCGTCGGTGACCATTGTCCTAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAAGTTCACCACCGGATTGA

6290

ori

ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC  
TGCCGATGTGATCTTCTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAAGTAC

6375

ori

CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT  
GCCGTTTGTGGTGGCGACCATCGCCACCAAAAAACAACGTTTCGTCGTCTAATGCGCGTCTTTTTTTTCTAGAGTCTTTCTA

6460

ori

CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA  
GGAAACTAGAAAAGATGCCCCAGACTGCGAGTACCTTGTCTTTGAGTGCAATTCCTAAAACCAAGTACTCTAATAGTTTTTCT

6545

TCTTACCTAGATCCTTTTACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGGTTGGTGGTTACGCGCAGCGTGACCGCTACAC  
AGAAGTGGATCTAGGAAAATGCGCGGGACATCGCCGCGTAATTCGCGCCGCCACACCACCAATGCGCGTTCGCACTGGCGATGTG

6630

f1 ori

TTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAA  
AACGGTTCGCGGGATCGCGGGCGAGGAAAGCGAAAGAAGGGAAGGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCAAGTTCGAGATTT

6715

f1 ori

TCGGGGGCTCCCTTTAGGGTTCGATTTAGTGCTTTACGGCACCTCGACCCAAAAAAGTGGATTAGGGTGGTTCACGTAGT  
AGCCCCCGAGGGGAAATCCCAAGGCTAAATCACGAAATGCCGTGGAGCTGGGGTTTTTTGAACTAATCCCACTACCAAGTGCATCA

6800

f1 ori

GGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCAAAAGTGGAA  
CCCGGTAGCGGGACTATCTGCCAAAAAGCGGGAAACTGCAACCTCAGGTGCAAGAAATTATCACCTGAGAACAAGTTTTGACCTT

6885

f1 ori

AanI  
PstI

CAACTCAACCTATCTCGGCTATTCTTTTATTATAAGGGATTTTGGCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGAT  
GTTGTGAGTTGGGATAGAGCCAGATAAGAAAATAAATATTCCTAAAACGGCTAAAGCCGGATAACCAATTTTTTACTCGACTA

6970

f1 ori

Swal

TTAACAAAAATTTAACGCGAATTTTAAACAAAATATTAACGTTTACAATTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTC  
AATTGTTTTTAAATTGCGCTTAAAATTGTTTTATAATTGCAAATGTTAAATTTAGTTAGATTTTCATATATACTCATTTGAACCAG

7055

f1 ori

TGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTC  
ACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTTCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTGAGGGGACG

7140

285 280 275 270 265  
W H K I L S A G I E A I Q R N R E D M T A Q S G T  
AmpR

GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG  
CACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGGTGCGAGTGGCCGAGGTC

7225

260 255 250 245 240 235  
T Y I V V I R S P K G D P G L A A I I G R S G R E G A G S  
AmpR

ATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAA  
TAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTACCAGGACGTTGAAATAGGCGGAGGTAGGTTCAGATAATT

7310

230 225 220 215 210 205  
K D A I F W G A P L A S R L L P G A V K D A E M W D I L  
AmpR

PstI

TTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCA  
AACAAACGGCCCTTCGATCTCATTTCATCAAGCGGTCAATTATCAAACGCGTTGCAACAACGGTAACGACGTCCTGATGACCCACAGT

7395

200 195 190 185 180  
Q Q R S A L T L L E G T L L K R L T T A M A A P M T T D  
AmpR

CGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAG  
GCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCGCGCTCAATGTAAGGGGGTACAACACGTTTTTTTC

7480

175 170 165 160 155 150  
R E D N P I A E N L E P E W R D L R T V H D G M N H L F A  
AmpR

PvuI

CGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAA  
GCCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATT

7565

145 140 135 130 125 120  
T L E K P G G I T T L L L N A A T N D S M T I A A S C L  
AmpR

ScaI

TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGG  
 AAGAGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACTCATGAGTTGGTTTCAGTAAGACTCTTATCACATACGCC

115 110 105 100 95

E R V T M G D T L H K E T V P S Y E V L D N Q S Y H I R

7650

AmpR

CGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAAC  
 GCTGGCTCAACGAGAACGGGCCGAGTTGTGCCCTATTATGGCGCGGTGTATCGTCTTGAATTTTACAGAGTAGTAACCTTTTG

90 85 80 75 70 65

R G L Q E Q G A D V R S L V A G C L L V K F T S M M P F R

7735

AmpR

GTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACCTGATCTTC  
 CAAGAAGCCCCGCTTTTGTAGAGTTCTAGAAATGGCGACAACCTTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGACTAGAAG

60 55 50 45 40 35

E E P R F S E L I K G S N L D L E I Y G V R A G L Q D E

7820

AmpR

AGCATCTTTTACTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGG  
 TCGTAGAAAATGAAAGTGGTCGCAAAAGACCCACTCGTTTTTGTCTTCCGTTTTACGGCGTTTTTCCCTTATTCCCGCTGTGCC

30 25 20 15 10

A D K V K V L T E P H A F V P L C F A A F F P I L A V R

7905

signal sequence

AmpR

AAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTG  
 TTTACAACCTTATGAGTATGAGAAGGAAAAAGTTATAATACTTCGTAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAAC

5 1

F H Q I S M

7990

signal sequence

AmpR promoter

AmpR

ZraI AatII

AATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT  
 TTACATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGTGTAAAGGGGCTTTTTACGGTGGACTGCAGATTCTTTGGTAATAATA

8075

AmpR promoter

CATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGATTTCTCATGTTTGACAGCTTATCATCGAATT  
 GTACTGTAATTGGATATTTTTATCCGCATAGTGCTCCGGGAAAGCAGAAGTTCTAAAGAGTACAACTGTGCAATAGTAGCTTAA

8160





















AATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTATTGTGAAATAGACGCAGATCGGGAACACTGAAAAAT  
 TTAAGAGTACAACTGTGCAATAGTAGCTATTGACTGAGTACAACCATAACACTTTTATCTGCGTCTAGCCCTTGTGACTTTTTTA




8245

AACAGTTATTATTCG 3'  
 TTTGCAATAATAAGC 5'

8260

Enzymes	Sites	
AanI	1	6923
AatII	1	8057
AflIII	1	5805
AgeI	1	1072
AleI	1	76
BamHI	1	1021
BfuAI	1	4255
BglII	2	1 5393
BipI	1	586
BspEI	1	3616
BspMI	1	4255
BspQI	1	5689
BsrGI	1	4070
BstZ17I	1	5578
Bsu36I	1	1860
Eco53kI	1	206
EcoRI	1	1011
FspAI	1	1598
HpaI	1	1910
NcoI	1	3404
NdeI	1	5628
NsiI	1	676
PaeR7I	1	1006
PciI	1	5805
PfIFI	1	5552
PmeI	1	411
PsiI	1	6923
PstI	1	7380
PvuI	1	7505
SacI	1	208
SalI	1	2949
SapI	1	5689
ScaI	1	7615
SgrAI	1	4399
SmaI	1	1019
SphI	1	4555
StuI	1	3034
Swal	1	7020
TiiI	1	1006
TspMI	1	1017
Tth111I	1	5552
XbaI	1	1804
XcmI	1	706
XhoI	1	1006
XmaI	1	1017
ZraI	1	8055

Feature	Location	Size (bp)			Type
✓ AOX1 promoter	2 .. 941	940			promoter
/gene	= Pichia pastoris AOX1				
/note	= inducible promoter, regulated by methanol				
✓ PHO1 signal sequence	942 .. 1007	66			CDS
/product	= signal sequence from a secreted acid phosphatase of Pichia pastoris				
/note	= synthetic gene fragment				
/translation	= MFSPILSLEIILALATLQSVFA 22 amino acids = 2.4 kDa				
✓ MCS	1006 .. 1026	21			misc_feature
/note	= multiple cloning site				
✓ AOX1 terminator	1093 .. 1339	247			terminator
/gene	= Pichia pastoris AOX1				
/note	= transcription terminator for AOX1				
✓ PpHIS4	1752 .. 4286	2535			CDS
/gene	= Pichia pastoris HIS4				
/product	= multifunctional enzyme, required for histidine biosynthesis				
/note	= auxotrophic marker for Pichia pastoris				
/translation	= MTFPLLPAAYASVAEFDNSLSLVGKAVFPYAADQLHNLIKFTQSTELQVNVQVESSVTEDEQFEELIDNLLKLYNNGINEVILDLDLAE RVVQRMIPGARVIYRTLVDKVASLPANASIAVPFSSPLGDLKSFTNGGSRVYAFSETAKLVDVTSTVASGIIPIIDARQLTTEYELS EDVKKFPVSEILLASLTTDRPDGLFTTLVADSSNYSLGLVYSSKKSIEAIRTOTGVYQSRRHGLWYKSGATQKLLGIELDCDGD CLKFVVEQTGVGFCHLERTSCFGQSKGLRAMEATLWDRKSNAPESYTKRLEFDDEVLLNAKIREEAEDELAEAKSKEDIAWECADLF YFALVRCAYKGVTLDEVERNLDMSLKVTRRKGDAKPGYTKEQPKESKPKVEPSEGRIELCKIDVSKASSQEI EDALRRPIQKTEQ IMELVKPIVDNVRQNGDKALLELTAKFDGVALKTPVLEAPFPEELMQLPDNVKRAIDLSIDNVRKFHEAQLTETLQVETCPGVVCSR FARPIEKVGLYIPGGTALPSTSLMLGVPKAVAGCKEIVFASPPKKGDTLTPVEIYVAHKVGAACKIVLAGGAQAVAAMAYGTETVP KCDKIFGPGNQFVTAAKMMVQNDTSALCSIDMPAGPSEVLVIADKYADPDFVASDLLSQAHEGIDSQVILLAVDMTDKELARIED AVHNQAVQLPRVEIVRKCIAHSTTSLVATYEQALEMSNQYAPEHLILOIENASSYVDQVQHAGSVFVGAYSPESC GDYSSGTNH TIPTYGYARQYSGVNTATFOKFITSDVTPFGIKHIGOAVMDIAAVFVGI DAHRNAVKVRMFKI GII * 844 amino acids = 92.2 kDa				
✓ AOX1 3' fragment	4641 .. 5397	757			misc_feature
/note	= region downstream of Pichia pastoris AOX1 gene				
✓ ori	5866 .. 6454	589			rep_origin
/direction	= LEFT				
/note	= high-copy-number CoIE 1/pMB1/pBR322/pUC origin of replication				
✓ f1 ori	6565 .. 7020	456			rep_origin
/direction	= RIGHT				
/note	= f1 bacteriophage origin of replication; arrow indicates direction of (+) strand synthesis				
✓ AmpR	7062 .. 7922	861			CDS
▶ 2 segments					
/gene	= bla				
/product	= -lactamase				
/note	= confers resistance to ampicillin, carbenicillin, and related antibiotics				
/translation	= MSIQHFRVALIPFFAAFCPLPVFA,HPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERFPMMSSTFKVLLCGAVLSRV DAGQE QLGRRIHYSQNDLVEYSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTITIGGPKELTALFLNHMGDHSVTRLD RWEPELNEAIPN DERDTTTPAAMATTLRKLTTGELLTLASRQQLIDWMEADKVAAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIY TTGSQATMDFERNROIAFIGASIKHW* 286 amino acids = 31.5 kDa				

Feature	Location	Size (bp)			Type
✓ AmpR promoter	7923 .. 8027	105	<input type="checkbox"/>		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<sup>+</sup> Dcm<sup>+</sup> EcoKI<sup>+</sup>

Comments:

References: