



Departamento de Microbiología
Facultad de Biología
Universidad de Barcelona

Caracterización del lipopolisacárido de *Aeromonas* mesófilas

Memoria presentada por Natalia Jiménez Blasco
para optar al grado de Doctora por la Universidad de Barcelona

Programa de Doctorado
Microbiología Ambiental y Biotecnología
Bienio: 2003-2005

VºBº del director

VºBº de la codirectora

La doctoranda

Dr. Juan Tomás Magaña

Dra. Susana Merino Montero

Natalia Jiménez Blasco

Barcelona, mayo 2008

8. ANEXO

8.1 Secuencia nucleotídica de la región 1 *wa* del núcleo del LPS de *Aeromonas hydrophila* AH-3

Secuencia de nucleótidos de la región 1 *wa* del núcleo del lipopolisacárido de *A. hydrophila* AH-3, de su entorno en el genoma y de las secuencias de aminoácidos deducidas de las distintas pautas abiertas de lectura. Los codones de inicio de la traducción se indican en negrita y, los de final de lectura, en negrita y cursiva. Cuando éstos son compartidos por dos pautas de lectura adyacentes, los nucleótidos están subrayados. Las flechas señalan la dirección en la que se traducen las diferentes pautas abiertas de lectura. Los putativos promotores se indican en cajas, el posible inicio de transcripción está sombreado en amarillo, los putativos terminadores Rho-independientes se han subrayado con una línea continua y las putativas secuencias de unión a ribosomas, con una línea en acordeón.

En las secuencias aminoacídicas, se ha marcado el dominio conservado de la familia 1 de glicosiltransferasas en azul; el de la familia 9, en verde; y el de la familia 25, en rojo. El dominio polisacárido desacetilasa está indicado en marrón y los putativos dominios transmembrana, en naranja.

```

1   GCCAAGGTCG GGCAGGTCTT CGTCGTTGAC GTCAATGATC CCCCCTTCAC
    CGGTTCCAGC CCGTCCAGAA GCAGCAACTG CAGTTACTAG GGGGGAAGTG
51  GCAGGCTCCT GAGCAGGGCG ACCACGGAGG AACGCAACTG CTCCTGGGCC
    CGTCCGAGGA CTCGTCCCGC TGGTGCCTCC TTGCGTTGAC GAGGACCCGG
101 TTGAGGTATT TTTGCTGCAG CGGCTCATCG CGACTCAGGA TGTCCGGTAG
    AACTCCATAA AAACGACGTC GCCGAGTAGC GCTGAGTCCT ACAGGCCATC
151 ATTTGCGTAG AAGAAGCGGA ACTGCCACAT CAGATAGAAG ATGGCATCCA
    TAAACGCATC TTCTTCGCCT TGACGGTGTA GTCTATCTTC TACCGTAGGT
201 GGTAGCCCAT CAGGTTTTCC AGCTTGATCT CGTTGCTGCG CGCCGGGGTG
    CCATCGGGTA GTCCAAAAGG TCGAACTAGA GCAACGACGC GCGGCCCCAC
251 AAGCTCTCGC TCAGGTGGCG GCGTATTGA TCAAAAATCG AGTGGATGAT
    TTCGAGAGCG AGTCCACCGC CCGCATAACT AGTTTTTTAGC TCACCTACTA
301 GTCTTCTTTA TTCCTGAAGT GGTAATAAAG GTTGCCAGGG CTGATGCCAA
    CAGAAGAAAT AAGGACTTCA CCATTATTTT CAACGGTCCC GACTACGGTT
351 GATGCGCGGC AATGTGGTTG GTGGTGATGT TGCCTCACC CTGATCGTTG
    CTACGCGCCG TTACACCAAC CACCACTACA ACGCGAGTGG GACTAGCAAC
401 AACAGCTCGG TGGCGCTATG AATGATCCTG TCTCTGGTTT TCATCACTGT
    TTGTCGAGCC ACCGCGATAC TTAGTAGGAC AGAGACCAAA AGTAGTGACA
451 TCCCTCTGAC GGGCCCGTTG CATCAAGGTC GAATCCCCTT GCGGCCCCGTG

```

```

AGGGGAGACTG CCCGGGCAAC GTAGTTCCAG CTTAGGGGAA CGCCGGGCAC
501 ATAAACTTTT CGGCAATTTT A[AATTGGC]TC ACTATACACA GC[TTTAATAA]
TATTTGAAAA GCCGTTAAAG TTTAACCGAG TGATATGTGT CGAAATTATT

                                orf1 (hldD) →
                                MetIleValVal·
551 GCA[G]TCACAC GTGACCATCA ACGTACGCGA_GGCAAACCTCC ATGATCGTAG
CGTCAGTGTG CACTGGTAGT TGCATGCGCT CCGTTTGAGG TACTAGCATC
·VThrGlyGly AlaGlyPhe IleGlySerAsn LeuValLys GlnLeuAsn
601 TAACTGGCGG TGCTGGCTTT ATCGGCAGCA ATCTGGTCAA GCAGCTGAAC
ATTGACCGCC ACGACCGAAA TAGCCGTCGT TAGACCAGTT CGTCGACTTG
AlaGlnGlyArg ThrAspIle ValValIle AspAspLeuThr AspGlyThr·
651 GCTCAAGGAC GGACCGACAT CGTGGTCATC GATGACCTGA CCGACGGCAC
CGAGTTCCTG CCTGGCTGTA GCACCAGTAG CTACTGGACT GGCTGCCGTG
·LysPheVal AsnLeuValAsp LeuThrIle AlaAspTyr MetAspLysAsp·
701 CAAGTTCGTC AACCTGGTCG ATCTGACCAT TGCCGATTAC ATGGATAAGG
GTTCAAGCAG TTGGACCAGC TAGACTGGTA ACGGCTAATG TACCTATTCC
·AGluPheGln AlaArgIle ValSerGlyAsp GluPheGlu GluTrpAsp
751 ACGAGTTCCA GGCGCGGATC GTCTCCGGTG ACGAGTTTGA AGAGTGGGAC
TGCTCAAGGT CCGCGCCTAG CAGAGGCCAC TGCTCAAAC TCTCACCTG
AspGlyIleGlu ValIlePhe HisGluGly AlaCysSerAla ThrThrGlu·
801 GACGGCATCG AAGTGATCTT CCACGAAGGT GCCTGCTCCG CGACCACTGA
CTGCCGTAGC TCCACTAGAA GGTGCTTCCA CGGACGAGGC GCTGGTGACT
·TrpAsnGly LysPheIleMet GluValAsn TyrGluTyr SerLysAspLeu·
851 GTGGAACGGC AAGTTCATCA TGGAGGTCAA TTACGAGTAC TCCAAGGATC
CACCTTGCCG TTCAAGTAGT ACCTCCAGTT AATGCTCATG AGGTTCTAG
·LPheHisTyr CysIleGlu ArgGluIlePro PheIleTyr AlaSerSer
901 TGTTCCTACTA CTGCATCGAG CGGGAAATCC CGTTTATTTA CGCCTCCTCC
ACAAGGTGAT GACGTAGCTC GCCCTTTAGG GCAAATAAAT GCGGAGGAGG
AlaAlaThrTyr GlyGlyArg AsnAspAsn PheIleGluAsp ProLysPhe·
951 GCCGCGACCT ACGGTGGCCG CAACGACAAC TTCATCGAAG ATCCCAAGTT
CGGCGCTGGA TGCCACCGGC GTTGCTGTTG AAGTAGCTTC TAGGGTTCAA
·GluGlnPro LeuAsnValTyr GlyTyrSer LysGlnLeu PheAspGlnTyr·
1001 CGAGCAGCCG CTCAACGTCT ATGGCTACTC CAAGCAGCTG TTCGACCAGT
GCTCGTCGGC GAGTTGCAGA TACCGATGAG GTTCGTCGAC AAGCTGGTCA
·TValArgArg TrpMetPro GluIleAsnSer GlnValVal GlyLeuLys
1051 ACGTGCGCCG CTGGATGCCG GAGATCAACT CCCAGGTCGT GGGTCTGAAG
TGCACGCGGC GACCTACGGC CTCTAGTTGA GGGTCCAGCA CCCAGACTTC

```

TyrPheAsnVal TyrGlyPro ArgGluGln HisLysGlySer MetAlaSer·
 1101 TATTTCAACG TCTACGGTCC GCGCGAGCAA CACAAGGGCT CCATGGCTTC
 ATAAAGTTGC AGATGCCAGG CGCGCTCGTT GTGTTCCCGA GGTACCGAAG
 ·ValAlaPhe HisLeuAsnThr GlnValLys LysGlyGlu AsnProLysLeu·
 1151 CGTTGCCTTC CACCTCAATA CTCAGGTGAA GAAGGGCGAG AATCCCAAGC
 GCAACGGAAG GTGGAGTTAT GAGTCCACTT CTTCCCGCTC TTAGGGTTTCG
 ·LPheGluGly CysAspGly PheProAsnGly GlyGlnMet ArgAspPhe
 1201 TGTTCGAAGG CTGTGACGGT TTCCCGAACG GTGGCCAGAT GCGTGA CTTC
 ACAAGCTTCC GACACTGCCA AAGGGCTTGC CACCGGTCTA CGCACTGAAG
 IleTyrValAsp AspValCys LysValAsn LeuTrpPheTrp GlnAsnPro·
 1251 ATCTATGTGG ATGACGTCTG CAAGGTGAAT CTCTGGTTCT GGCAGAACCC
 TAGATACACC TACTGCAGAC GTTCCACTTA GAGACCAAGA CCGTCTTGGG
 ·GlnHisSer GlyIlePheAsn CysGlyThr GlyArgAla GluProPheGln·
 1301 GCAGCACTCC GGCATCTTCA ACTGCGGTAC CGGTCTGTGCC GAGCCGTTCC
 CGTCTGTGAGG CCGTAGAAGT TGACGCCATG GCCAGCACGG CTCGGCAAGG
 ·GAsnValAla GluAlaVal IleGlnHisHis GlnLysGly AlaIleGlu
 1351 AGAACGTGGC CGAAGCCGTG ATCCAGCATC ATCAGAAAGG GGCAATCGAG
 TCTTGCACCG GCTTCGGCAC TAGGTCGTAG TAGTCTTTCC CCGTTAGCTC
 TyrIleProPhe ProAspHis LeuLysGly ArgTyrGlnSer PheThrGln·
 1401 TACATTCCCT TCCCGGATCA CCTCAAGGGT CGCTATCAGA GCTTCACTCA
 ATGTAAGGGA AGGGCCTAGT GGAGTTCCCA GCGATAGTCT CGAAGTGAGT
 ·AlaAspMet ThrArgLeuArg GlyValGly TyrAspAla GluPheLysThr·
 1451 GGCAGACATG ACCAGGCTGC GCGGTGTCGG TTATGACGCC GAGTTCAAGA
 CCGTCTGTAC TGGTCCGACG CGCCACAGCC AATACTGCGG CTCAAGTTCT
 ·TValAlaGlu GlyValAla AspTyrMetAla TrpLeuAsn ArgVal***
 1501 CTGTGGCCGA AGGCGTGGCC GATTACATGG CCTGGTTGAA CCGGGTTTGA
 GACACCGGCT TCCGCACCGG CTAATGTACC GGACCAACTT GGCCCAA ACT
 1551 TCCCCCTCTG ACGGGTGGCC CCTGGCCACC CGTCGATCGG ATACGCTTCC
 AGGGGGAGAC TGCCCACCGG GGACCGGTGG GCAGCTAGCC TATGCGAAGG

orf2 (wahA) →

MetAsnIleLeu MetAlaLeu SerGlnLeu GluValThrGly AlaGluVal·
 1601 **ATGA**ACATAT TGATGGCACT CTCCAGCTC GAAGTGACGG GAGCTGAGGT
 TACTTGTATA ACTACCGTGA GAGGGTCGAG CTTCACTGCC CTCGACTCCA
 ·TyrAlaThr ThrValGlyAsn GluLeuThr ArgArgGly HisGlnValPhe·
 1651 CTATGCCACC ACGGTCGGTA ACGAATTGAC CCGACGTGGT CATCAGGTTT
 GATACGGTGG TGCCAGCCAT TGCTTAACTG GGCTGCACCA GTAGTCCAAA
 ·PTyrValSer AspThrLeu ThrLysProThr GlnGlyPro ValPheLys
 1701 TCTACGTATC GGACACGCTG ACCAAGCCTA CCCAGGGCCC CGTGTTC AAG
 AGATGCATAG CCTGTGCGAC TGGTTCGGAT GGGTCCCGGG GCACAAGTTC

LeuArgPheAsn LysArgSer IleLeuArg ArgPheTrpHis ValPheTyr·
 1751 CTGCGTTTCA ACAAGCGCAG CATACTGCGC CGCTTCTGGC ATGTGTTTTTA
 GACGCAAAGT TGTTTCGCGTC GTATGACGCG GCGAAGACCG TACACAAAAT
 ·LeuIleTyr LeuIleLysLys HisHisIle GlnLeuVal HisAlaHisSer·
 1801 TCTCATCTAC CTCATCAAGA AGCACCATAT CCAGTTAGTG CATGCACACT
 AGAGTAGATG GAGTAGTTCT TCGTGGTATA GGTC AATCAC GTACGTGTGA
 ·SArgAlaSer GlyTrpSer SerTyrValAla CysLysLeu ThrGlyThr
 1851 CTCGAGCCTC CGGCTGGAGC AGCTATGTTG CCTGCAAGCT GACCCGCACT
 GAGCTCGGAG GCCGACCTCG TCGATACAAC GGACGTTCTGA CTGGCCGTGA
 ProMetValThr ThrValHis GlyArgGln ProValHisAla SerArgLys·
 1901 CCCATGGTTA CCACGGTACA TGGCCGCCAG CCGGTGCATG CCTCCCACAA
 GGGTACCAAT GGTGCCATGT ACCGGCGGTC GGCCACGTAC GGAGGGCGTT
 ·AlaPheHis AlaLeuGlyPhe ArgAlaVal AlaValCys GluAspIleAla·
 1951 GGCTTTCCAT GCCCTCGGGT TCAGGGCCGT GGCCGTCTGC GAGGATATTG
 CCGAAAGGTA CGGGAGCCCA AGTCCCGGCA CCGGCAGACG CTCCTATAAC
 ·AHisGlnIle IleAspAsn LeuGlyValAsp ProAlaIle ValGlnVal
 2001 CCCACCAGAT CATCGACAAT CTGGGGGTGG ATCCGGCCAT AGTGCAGGTG
 GGGTGGTCTA GTAGCTGTTA GACCCCCACC TAGGCCGGTA TCACGTCCAC
 LeuArgAsnGly IleGluThr AspLysPhe GlnProValPro AlaProAsp·
 2051 CTGCGCAACG GCATCGAGAC TGACAAGTTT CAGCCAGTGC CAGCACCGGA
 GACGCGTTGC CGTAGCTCTG ACTGTTCAAA GTCGGTACAG GTCGTGGCCT
 ·AsnAlaThr ProValIleAla IleIleGly ArgLeuSer GlyProLysGly·
 2101 CAATGCAACG CCGGTGATCG CCATCATAGG GCGGCTGAGC GGCCCCAAGG
 GTTACGTTGC GGCCACTAGC GGTAGTATCC CGCCGACTCG CCGGGGTTCC
 ·GGluLeuCys TyrArgLeu LeuAspGluVal LeuAspLeu AspAlaCys
 2151 GCGAGCTCTG CTATCGTCTG CTGGACGAGG TGCTGGATCT GGATGCCTGT
 CGCTCGAGAC GATAGCAGAC GACCTGCTCC ACGACCTAGA CCTACGGACA
 GlnValArgVal ValSerGly SerGlnLeu ProGluArgPhe HisArgPhe·
 2201 CAGGTAAGGG TGGTGAGTGG CAGCCAGCTG CCTGAGCGTT TCCACCGCTT
 GTCCATTCCC ACCACTCACC GTCGGTCGAC GGACTCGCAA AGGTGGCGAA
 ·GlnSerGln ValAspPheVal GlyTyrVal AspAspVal ProThrLeuLeu·
 2251 TCAGAGTCAG GTCGATTTTCG TCGGTTATGT GGATGATGTC CCTACCTTGC
 AGTCTCAGTC CAGCTAAAGC AGCCAATACA CCTACTACAG GGATGGAACG
 ·LAlaGlyCys AspLeuVal IleGlyAlaGly ArgValAla MetGluAla
 2301 TGGCGGGATG TGATCTGGTG ATCGGTGCAG GCCGGGTGGC GATGGAGGCG
 ACCGCCCTAC ACTAGACCAC TAGCCACGTC CGGCCACCG CTACCTCCGC
 LeuLeuCysGly ArgProAla PheAlaIle GlyGluAlaLys AlaIleGly·
 2351 TTGCTGTGCG GGCGTCCCGC TTTTGCCATC GGCGAGGCCA AGGCGATTGG
 AACGACACGC CCGCAGGGCG AAAACGGTAG CCGCTCCGGT TCCGCTAACC

·LeuValThr GluGlnAsnLeu AlaGluAla LeuAlaSer AsnPheGlyAsp·
 2401 TCTGGTGACC GAACAGAATC TAGCGGAGGC GCTGGCCAGC AACTTCGGGC
 AGACCACTGG CTTGTCTTAG ATCGCCTCCG CGACCGGTCG TTGAAGCCGC
 ·AileGlyPro GlnAspLeu AlaIleAspPhe SerAlaLeu LysThrGlu
 2451 ACATAGGCCC CCAGGATCTG GCGATCGACT TCAGTGCCTT GAAGACAGAG
 TGTATCCGGG GGTCCCTAGAC CGCTAGCTGA AGTCACGGAA CTTCTGTCTC
 IleAlaProAla LeuAlaSer LysSerVal ProGluAlaIle ArgGlnArg·
 2501 ATAGCACCCG CATTGGCCAG CAAGAGTGTG CCTGAGGCCA TTCGCCAGCG
 TATCGTGGGC GTAACCGGTC GTTCTCACAC GGACTCCGGT AAGCGGTCGC
 ·IleGlnAsp GluTyrGlyLeu GluLysVal ValSerGly LeuGluSerIle·
 2551 CATTCAGGAT GAGTATGGCT TGGAGAAAGT AGTATCCGGG CTGGAATCCA
 GTAAGTCCTA CTCATACCGA ACCTCTTTCA TCATAGGCC GACCTTAGGT
 ·ITyrGlnAsp AlaValVal GluThrLeuArg ArgGluMet ProValIle
 2601 TCTATCAGGA TGCGGTGGTC GAGACCCTGC GCCGCGAGAT GCCTGTCATC
 AGATAGTCCT ACGCCACCAG CTCTGGGACG CGGCGCTCTA CGGACAGTAG
 MetTyrHisArg PheIleGlu HisAspSer GluLysGlyVal HisGlyThr·
 2651 ATGTACCACC GTTTTATCGA GCATGACAGC GAGAAGGGGG TACATGGTAC
 TACATGGTGG CAAAATAGCT CGTACTGTCTG CTCTTCCCC ATGTACCATG
 ·TrpMetPro IleAlaMetPhe GluLysHis LeuArgLeu MetLysTrpLeu·
 2701 CTGGATGCCT ATTGCCATGT TCGAGAAGCA TCTTCGTTTA ATGAAATGGC
 GACCTACGGA TAACGGTACA AGCTCTTCGT AGAAGCAAAT TACTTTACCG
 ·LGlyTyrGlu ThrLeuThr PheArgAspLeu AlaAspArg GlyPheIle
 2751 TGGGTTATGA AACTCTGACT TTCCGCGATC TGGCGGACAG GGGCTTTATC
 ACCCAATACT TTGAGACTGA AAGGCGCTAG ACCGCTGTC CCCGAAATAG
 HisArgLeuGlu TyrGlyLys LysTyrLeu MetIleThrAla AspAspGly·
 2801 CATCGTCTTG AATACGGCAA GAAGTACCTG ATGATCACCG CCGATGACGG
 GTAGCAGAAC TTATGCCGTT CTTTATGGAC TACTAGTGGC GGCTACTGCC
 ·TyrGlnAsp AsnLeuThrArg MetLeuPro LeuLeuGlu LysTyrGlyTyr·
 2851 CTATCAGGAC AACCTGACGC GCATGCTGCC GTTGCTTGAA AAATACGGTT
 GATAGTCCTG TTGGACTGCG CGTACGACGG CAACGAACTT TTTATGCCAA
 ·TLysAlaVal ValTyrVal ValThrGlyGlu AlaHisAsn ArgTrpAsp
 2901 ACAAAGCCGT GGTATATGTG GTGACCGGAG AGGCGCATAA TCGCTGGGAT
 TGTTTCGGCA CCATATACAC CACTGGCCTC TCCGCGTATT AGCGACCCTA
 ValGluHisAla ThrAsnPro AspThrGln ValSerLeuMet AsnGlyGlu·
 2951 GTGGAGCATG CCACCAACCC GGATACCCAG GTTTCCTGA TGAATGGGA
 CACCTCGTAC GGTGGTTGGG CCTATGGGTC CAAAGGGACT ACTTACCCTT
 ·GlnIleLys AlaLeuValAla SerGlyHis ValGluVal GlyGlyHisThr·
 3001 GCAGATAAAG GCCCTTGTGG CCTCTGGCCA TGTGGAAGTC GGTGGCCACA
 CGTCTATTTT CGGGAACACC GGAGACCGGT ACACCTTCAG CCACCGGTGT

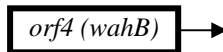
·TLeuThrHis ProArgLeu SerGluLeuSer ProGluArg GlnValHis
 3051 CCCTGACTCA TCCCAGGTTG AGCGAGCTCT CTCCTGAGCG GCAAGTCCAT
 GGGACTGAGT AGGGTCCAAC TCGCTCGAGA GAGGACTCGC CGTTCAGGTA
 GluIleArgGlu AsnLysArg GlnLeuGlu AlaLeuLeuGly HisSerLeu·
 3101 GAAATAAGAG AAAACAAACG ACAACTCGAG GCCTTGCTAG GTCACTCTCT
 CTTTATTCTC TTTTGTTCG TGTGAGCTC CGGAACGATC CAGTGAGAGA
 ·LeuSerPhe AlaTyrProTyr GlyAspMet AspGluSer AlaLysValGln·
 3151 GCTGTGCTTC GCCTACCCTT ATGGGGATAT GGACGAGAGC GCCAAGGTGC
 CGACAGCAAG CGGATGGGAA TACCCCTATA CCTGCTCTCG CGGTTCCACG
 ·GAlaGlnAla AlaGlyTyr ArgPheAlaVal AlaThrAsn SerGlyPro
 3201 AGGCACAGGC AGCAGGCTAC CGTTTTGCCG TGGCGACCAA CAGTGGTCCC
 TCCGTGTCCG TCGTCCGATG GCAAACGGC ACCGCTGGTT GTCACCAGGG
 ArgAlaMetHis GlnAspPro TyrArgIle ArgArgIleAla IlePhePro·
 3251 AGAGCCATGC ATCAGGATCC TTACCGGATC CGTCGCATCG CCATTTTCCC
 TCTCGGTACG TAGTCCTAGG AATGGCCTAG GCAGCGTAGC GGTAAAAGGG
 ·ArgThrAsp ValPheGlyLeu TrpArgLys IleArgGly AsnTyrValPhe·
 3301 TCGTACCGAT GTTTTTGGCC TGTGGCGCAA AATCCGGGGT AACTATGTCT
 AGCATGGCTA CAAAACCGG ACACCGCGTT TTAGGCCCCA TTGATACAGA

orf3 (waaL) →

 ·PArgLys*** Met ProIleArgCys LeuTyrLeu LeuAspLys
 3351 TCAGGAAATA ATGATTAATG CCCATCAGAT GCCTTTATCT GCTCGATAAA
 AGTCCTTTAT TACTAATTAC GGGTAGTCTA CGGAAATAGA CGAGCTATTT
 ThrThrThrAla LeuPheAla LeuThrIle IlePheSerPhe CysGlyLeu·
 3401 ACCACCACAG CCCTGTTTGC ATTGACCATC ATTTTTTCTT TTTGCGGCCT
 TGGTGGTGTC GGGACAAACG TAACTGGTAG TAAAAAAGGA AAACGCCGGA
 ·PheLeuVal ProAlaGlyGln SerIleGlu SerAsnLeu LeuAlaValThr·
 3451 TTTTTTGGTA CTGCCGGCC AGAGCATTGA ATCTAATCTG CTGGCCGTGA
 AAAAAACCAT GGACGGCCGG TCTCGTAACT TAGATTAGAC GACCGGCACT
 ·TGlyIlePhe GlyLeuLeu AsnTyrPheVal GlyLysGln GlnAspVal
 3501 CCGGCATATT TGGCTTGCTG AACTACTTTG TCGGCAAGCA ACAAGATGTT
 GGCCGTATAA ACCGAACGAC TTGATGAAAC AGCCGTTCTG TGTTCACAA
 GlyLeuLysAsp ArgArgIle LeuTrpGly LeuValValTyr AlaAlaMet·
 3551 GGACTGAAAG ATCGCCGCAT TCTCTGGGGA CTGGTAGTGT ATGCCGCCAT
 CCTGACTTTC TAGCGGCGTA AGAGACCCT GACCATCACA TACGGCGGTA
 ·IlePheVal AsnArgLeuIle HisGlyAsp GlnTyrGly ValMetArgGly·
 3601 GATTTTCGTC AACCGTTGA TTCATGGTGA TCAATATGGG GTCATGCGCG
 CTAAAAGCAG TTGGCCAACT AAGTACCACT AGTTATACCC CAGTACGCGC

·GLeuPheTyr ValValIle PheAlaLeuMet MetProArg LysProIle
 3651 GGCTGTTTTA TGTTGTTATC TTTGCACTGA TGATGCCGCG CAAGCCAATT
 CCGACAAAAT ACAACAATAG AAACGTGACT ACTACGGCGC GTTCGGTTAA
 LeuLeuMetLeu GlyTyrLeu AlaIleVal LeuGlyGlyMet GlyLeuGly·
 3701 CTATTGATGC TGGGTTATCT GGCTATCGTC CTGGGCGGGA TGGGACTGGG
 GATAACTACG ACCCAATAGA CCGATAGCAG GACCCGCCCT ACCCTGACCC
 ·MetLeuSer IleTrpGlnTyr GlnHisGly MetValArg ValGluGlyPhe·
 3751 AATGCTCAGC ATCTGGCAGT ATCAGCATGG TATGGTCAGG GTTGAAGGGT
 TTACGAGTCG TAGACCGTCA TAGTCGTACC ATACCAGTCC CAACTTCCCA
 ·PThrAsnAla IleLeuPhe SerGlnAlaAla LeuThrLeu AlaIleLeu
 3801 TTACCAACGC CATTCTCTTT TCACAAGCAG CATTGACGCT GGCGATCCTC
 AATGGTTGCG GTAAGAGAAA AGTGTTCGTC GTAAGTGCAG CCGCTAGGAG
 AsnTrpPheVal PheGlnGln ArgTrpLeu ProArgTrpPhe ArgArgGly·
 3851 AACTGGTTTG TCTTTCAGCA GAGATGGCTG CCTAGATGGT TTCGGAGGGG
 TTGACCAAAC AGAAAGTCGT CTCTACCGAC GGATCTACCA AAGCCTCCCC
 ·AlaLeuIle GlyLeuIleAla AlaLeuPhe AlaLeuTyr LeuSerGlnSer·
 3901 GGCTCTTATT GGCCTGATTG CGGCCCTGTT TGCCTTTAT CTCTCCAGA
 CCGAGAATAA CCGGACTAAC GCCGGGACAA ACGCGAAATA GAGAGGTCT
 ·SArgGlyVal TrpLeuAla LeuGlyValIle LeuAlaTyr ValIleCys
 3951 GCCGGGGGGT CTGGCTTGCC CTCGGGGTCA TTCTTGCCTA CGTCATCTGT
 CGGCCCCCA GACCGAACGG GAGCCCCAGT AAGAACGGAT GCAGTAGACA
 TyrLysAlaTyr PheLysPro TrpLysTyr ValAlaValAla IleLeuCys·
 4001 TACAAGGCCT ATTTCAAGCC GTGGAAATAC GTTGCGGTAG CCATATTGTG
 ATGTTCCGGA TAAAGTTCGG CACCTTTATG CAACGCCATC GTTATAACAC
 ·ThrAlaThr ThrGlyAlaIle TyrHisThr AsnGlnLeu ValGlnValArg·
 4051 TACGGCCACG ACGGGCGCCA TTTATCATA CAAATCAGCTG GTTCAGGTTC
 ATGCCGGTGC TGCCCGCGGT AAATAGTATG GTTAGTCGAC CAAGTCCAAG
 ·AVAlAlaGlu AlaValSer AspIleAsnSer AlaGluLys GlySerTyr
 4101 GGGTCGCGGA GCGGTCTCG GATATCAACT CGGCAGAAA GGGCTCGTAC
 CCCAGCGCCT CCGCCAGAGC CTATAGTTGA GCCGTCTTTT CCCGAGCATG
 AspSerSerTrp GlyLeuArg ValIleAla TrpGlnSerAla TrpLeuGly·
 4151 GATTTCATCGT GGGGTCTGCG TGTCATTGCC TGGCAGAGTG CCTGGCTGGG
 CTAAGTAGCA CCCAGACGC ACAGTAACGG ACCGTCTCAC GGACCGACCC
 ·PheLeuAsp SerProValAla GlyValGly ThrAspGly PheAspAlaLeu·
 4201 TTTTTTGGAT TCGCCTGTGG CAGGAGTGGG CACTGATGGC TTCGACGCAT
 AAAAAACCTA AGCGGACACC GTCCTCACCC GTGACTACCG AAGCTGCGTA
 ·LLysGlnThr GlnIleAsn AsnHisGlnVal SerProLeu IlePheHis
 4251 TGAAACAAAC CCAGATCAAT AATCATCAGG TCTCTCCCCT GATATTTTCAT
 ACTTTGTTTG GGTCTAGTTA TTAGTAGTCC AGAGAGGGGA CTATAAAGTA

ProAlaLeuVal HisSerHis AsnGlnTyr MetGlnAsnLeu ValValArg·
 4301 CCGGCGCTGG TCCACTCGCA TAATCAATAT ATGCAGAACC TGGTGGTTTCG
 GGCCGCGACC AGGTGAGCGT ATTAGTTATA TACGTCTTGG ACCACCAAGC
 ·GlyGlyIle GlyPheValAla LeuValIle PheLeuGly LeuProLeuTyr·
 4351 AGGCGGCATC GGATTTGTTG CATTGGTGAT ATTTCTTGGC CTGCCGCTCT
 TCCGCCGTAG CCTAAACAAC GTAACCACTA TAAAGAACCG GACGGCGAGA
 ·TPheAlaAla LysAsnIle AsnArgIleSer AlaGlyVal LeuValPro
 4401 ATTTTGCCGC GAAAAATATA AACAGGATTT CTGCAGGGGT ACTGGTGCCT
 TAAAACGGCG CTTTTTATAT TTGTCCTAAA GACGTCCCCA TGACCACGGA
 LeuAlaPheAla IleAsnSer LeuSerAsp ValPropheGlu HisGlnGly·
 4451 CTGGCGTTTG CGATTAATTC GCTAAGCGAT GTTCCGTTTG AACATCAAGG
 GACCGCAAAC GCTAATTAAG CGATTGCTA CAAGGCAAAC TTGTAGTTCC
 ·ValLeuTyr LeuTyrThrLeu SerLeuVal PheIleTrp PheAlaHisGlu·
 4501 GGTGCTCTAT TTGTATACGC TCAGCCTGGT TTTCATCTGG TTTGCCCATG
 CCACGAGATA AACATATGCG AGTCGGACCA AAAGTAGACC AAACGGGTAC
 ·GPheGluLys AspThrCys ThrSer***



MetIle ProValPhe ValIleSerLeu·
 4551 AGTTTGAAAA GGAATACCTGT ACGTCAATGAT TCCGGTTTTT GTCATCAGTT
 TCAAACTTTT CCTATGGACA TGCAGTACTA AGGCCAAAAA CAGTAGTCAA
 ·LAlaArgSer ValGluArg ArgAlaMetVal ValGlnGln MetAlaHis
 4601 TGGCTCGCAG CGTTGAGCGG CGGGCGATGG TTGTACAGCA GATGGCGCAT
 ACCGAGCGTC GCAACTCGCC GCCCGCTACC AACATGTCGT CTACCGCGTA
 LeuGlyIleAsn PheGluPhe ValAspAla ThrAspGlyLys AlaLeuSer·
 4651 CTGGGCATCA ATTTTCGAGTT CGTAGATGCG ACCGATGGGA AAGCATTGAG
 GACCCGTAGT TAAAGCTCAA GCATCTACGC TGGCTACCCT TTCGTAACCTC
 ·SerAlaGlu LeuAlaLysVal AspLeuGlu LeuAlaLys GluValCysGly·
 4701 CAGCGCGGAA CTTGCCAAGG TTGACTTGGG GCTGGCCAAA GAAGTGTGTG
 GTCGCGCCTT GAACGGTTCC AACTGAACCT CGACCGGTTT CTTACACAC
 ·GHisGluLeu SerPheGly GluIleGlyCys AlaMetSer HisIleArg
 4751 GCCATGAGCT GAGTTTTGGT GAGATTGGCT GTGCCATGAG CCATATCCGT
 CGGTACTCGA CTCAAACCA CTCTAACCGA CACGGTACTC GGTATAGGCA
 ValTyrGluArg IleValAla GluGlyIle AspArgCysVal AlaLeuGlu·
 4801 GTTTATGAGC GCATAGTCGC CGAGGGAATT GATCGCTGCG TGGCGTTGGA
 CAAATACTCG CGTATCAGCG GCTCCCTTAA CTAGCGACGC ACCGCAACCT
 ·AspAspIle TyrLeuHisMet HisSerLys AlaIleIle GluThrIleVal·
 4851 AGATGACATT TACCTGCACA TGCACAGCAA GGCGATCATC GAGACGATAG
 TCTACTGTAA ATGGACGTGT ACGTGTCTGTT CCGCTAGTAG CTCTGCTATC

·VArgSerCys HisSerGlu IleValPheLeu HisHisGly LysAlaLys
 4901 TCGCTCCTG CCATTCGGAG ATCGTCTTCC TGCATCATGG CAAAGCCAAG
 ACGCGAGGAC GGTAAGCCTC TAGCAGAAAG ACGTAGTACC GTTTCGGTTC
 ArgTrpProIle LeuSerSer LeuProGly GlyTyrArgLeu AlaArgTyr·
 4951 CGCTGGCCAA TTCTTTCTC CTTGCCGGGT GGTATCGGC TTGCCAGATA
 GCGACCGGTT AAGAAAGGAG GAACGGCCCA CCAATAGCCG AACGGTCTAT
 ·LeuAlaPro SerArgThrSer ArgArgGly IleLeuSer ThrAlaGlyTyr·
 5001 TCTGGCACCG AGTCGCACCT CCAGGAGAGG GATCCTGTCG ACCGCAGGGT
 AGACCGTGGC TCAGCGTGGG GGTCTCTTCC CTAGGACAGC TGGCGTCCCA
 ·TValLeuThr LeuAlaGly AlaLysLysLeu LeuGlnCys AlaTyrPro
 5051 ATGTTCTGAC CCTGGCCGGG GCGAAGAAAT TATTGCAGTG TGCTTATCCG
 TACAAGACTG GGACCGGCC CGCTTCTTTA ATAACGTCAC ACGAATAGGC
 LeuArgMetPro SerAspTyr LeuThrGly ArgLeuGlnLeu AsnGlyLeu·
 5101 CTGCGCATGC CTTCCGATTA TCTGACTGGT CGCCTGCAGC TTAATGGCCT
 GACGCGTACG GAAGGCTAAT AGACTGACCA GCGGACGTCG AATTACCGGA
 ·AlaAlaSer GlyValGluPro CysCysLeu AspValAsp LeuPheThrThr·
 5151 GGCTGCCTCC GGTGTTGAGC CCTGCTGCCT GGATGTGGAT CTCTTTACCA
 CCGACGGAGG CCACAACCTCG GGACGACGGA CCTACACCTA GAGAAATGGT
 ·TThrIleAsp AspArgAla TyrGlyHisTyr LeuGlu***
 5201 CCACGATTGA TGATCGGGCT TACGGCCACT ATCTGGAAATA GGCGGGTGT
 GGTGCTAACT ACTAGCCCGA ATGCCGGTGA TAGACCTTAT CCGCCCAACA
 5251 CTTGGCCGGG ATTGGGGTCA GTAAACACCC ATGCTGTCCCT GGCTGGCGAA
 GAACCGGCC TAACCCAGT CATTTGTGGG TACGACAGGA CCGACCGCTT

orf5 (wahC) →

MetAsn PheLysLys LeuLeuThr

5301 GTGGGGCCAT ATATGCTGGA TAATCTATGA ATTTCAAAAA GTTGCTTACC
 CACCCCGGTA TATACGACCT ATTAGATACT TAAAGTTTTT CAACGAATGG
 HisCysIleAla LysTrpVal IleGlySer GlnProAlaAla IleProSer·
 5351 CATTGCATCG CCAAGTGGGT CATTGGCAGC CAGCCTGCTG CCATTCCAAG
 GTAACGTAGC GGTTCACCCA GTAACCGTCG GTCGGACGAC GGTAAGGTTC
 ·ProValGly ThrGluValAla AsnValLeu LeuValGly TyrAspAlaIle·
 5401 CCCCCTCGGA ACAGAAGTGG CAAACGTGCT GCTGGTCGGT TATGACGCCA
 GGGGCAGCCT TGTCTTCACC GTTTGCACGA CGACCAGCCA ATACTGCGGT
 ·IGlyAspPhe IleLeuSer LeuProAlaIle AlaGlnLeu ArgArgLeu
 5451 TCGGCGATTT CATTCTGTCA TTGCCCGCCA TTGCGCAGCT CAGGCGCCTG
 AGCCGCTAAA GTAAGACAGT AACGGGCGGT AACCGGTCGA GTCCGCGGAC
 TyrProAlaAla ArgPheAsp LeuValCys SerGlnArgAsn AlaLeuLeu·
 5501 TATCCAGCCG CCCGATTTGA CCTCGTGTGC AGCCAGCGCA ACGCGCTGCT
 ATAGGTCGGC GGGCTAAACT GGAGCACACG TCGGTCGCGT TGCGCGACGA

·AlaAlaSer ValProGlyIle AspGlnCys HisValIle ThrLeuAsnAsp·
 5551 GGCCGCATCT GTGCCAGGTA TTGATCAATG CCATGTGATC ACCCTGAACG
 CCGCCGTAGA CACGGTCCAT AACTAGTTAC GGTACACTAG TGGGACTTGC
 ·AThrLeuLeu ProAlaGly MetTrpCysLys LeuArgGlu LeuArgGln
 5601 ACACCCTGTT GCCAGCGGGC ATGTGGTGCA AACTCCGCGA GCTCAGGCAA
 TGTGGGACAA CGGTCGCCCG TACACCACGT TTGAGGCGCT CGAGTCCGTT
 ArgHisTyrAsp ValValIle AsnLeuPhe AspGluProAsp AspIleAla·
 5651 CGGCACTATG ACGTGGTCAT CAACCTGTTC GATGAGCCCG ACGATATCGC
 GCCGTGATAC TGCACCAGTA GTTGGACAAG CTACTCGGGC TGCTATAGCG
 ·MetAlaLys LeuLeuLeuLeu AlaAsnGly ArgLeuGln SerLeuProLeu·
 5701 CATGGCCAAG CTGCTGTTGC TGGCCAATGG TCGATTGCAG TCGTTGCCCC
 GTACCGGTTT GACGACAACG ACCGGTTACC AGCTAACGTC AGCAACGGGG
 ·LArgPheLys SerGluGly GlnGlnLysLeu LeuProLeu PheAsnGlu
 5751 TGCGCTTCAA GAGCGAAGGG CAGCAGAAGT TGCTGCCGCT GTTCAACGAA
 ACGCGAAGTT CTCGCTTCCC GTCGTCTTCA ACGACGGCGA CAAGTTGCTT
 LysAlaThrIle ValProSer ArgProIle ArgAspHisPhe ValTyrArg·
 5801 AAGGCCACCA TAGTGCCATC ACGCCCCATC CGGGATCATT TCGTCTATCG
 TTCCGGTGGT ATCACGGTAG TGCGGGGTAG GCCCTAGTAA AGCAGATAGC
 ·MetLeuSer ValAlaGlyAsp GlnThrAsp ValProVal ThrIleProSer·
 5851 CATGCTCAGC GTGGCGGGGG ATCAAACGGA TGTACCGGTT ACCATACCAT
 GTACGAGTCG CACCGCCCCC TAGTTTGCCT ACATGGCCAA TGGTATGGTA
 ·SProCysAsn AlaGlnLeu AspThrSerGly TyrGlyArg TyrLeuLeu
 5901 CGCCCTGCAA TGCACAGCTG GATACCAGTG GTTATGGGCG CTATCTGCTG
 GCGGGACGTT ACGTGTGCAC CTATGGTCAC CAATACCCGC GATAGACGAC
 ValAsnLeuThr GlySerGln ValGlyAsn SerMetAlaAsp GluGlnIle·
 5951 GTCAATCTGA CGGGCAGCCA GGTGGGCAAT TCGATGGCGG ATGAGCAAAT
 CAGTTAGACT GCCCGTCGGT CCACCCGTTA AGCTACCGCC TACTCGTTTA
 ·AsnGlyIle LeuAlaGlnLeu ProValTyr ProGlyVal SerTyrLeuVal·
 6001 CAACGGCATT CTGGCCCAAT TGCCGGTGTA TCCTGGGGTT TCCTATCTGG
 GTTGCCGTAA GACCGGGTTA ACGGCCACAT AGGACCCAA AGGATAGACC
 ·VPheSerArg ArgProLeu AlaCysGlnArg GlnAspMet ThrAlaLeu
 6051 TCTTCTCCAG GCGTCCTTTG GCATGCCAGC GGCAGGATAT GACAGCCCTG
 AGAAGAGGTC CGCAGGAAAC CGTACGGTCG CCGTCCTATA CTGTGGGGAC
 PheProAspThr IleLeuAsp AlaAlaLys IleIleLysAsp AlaGlnGly·
 6101 TTTCCCACCA CCATTCTCGA TGCGGCAAAG ATTATCAAGG ATGCGCAGGG
 AAAGGGCTGT GGTAAGAGCT ACGCCGTTTC TAATAGTTCC TACGCGTCCC
 ·ValIleSer ThrAspThrSer IleIleHis IleSerSer SerPheGlyVal·
 6151 GGTACATCTG ACCGACACCT CCATCATCCA TATCTCGTCT TCGTTGGCGG
 CCAGTAGAGC TGGCTGTGGA GGTAGTAGGT ATAGAGCAGA AGCAAGCCGC

·VProThrLeu ValLeuMet AsnAsnGluSer TrpArgAsp AlaPheIle
 6201 TGCCGACTCT GGTGCTGATG AACAAACGAGT CCTGGCGCGA TGCCTTTATT
 ACGGCTGAGA CCACGACTAC TTGTTGCTCA GGACCGCGCT ACGGAAATAA
 ProLeuSerGly ArgAsnIle IleLeuArg SerThrThrAsp AsnLeuSer·
 6251 CCGCTCTCTG GCCGCAATAT CATCCTGAGA TCGACCACGG ACAACCTTTC
 GCGGAGAGAC CGGCGTTATA GTAGGACTCT AGCTGGTGCC TGTGGAAG
 ·AlaLeuSer ProValGluIle SerArgGln ValAspAla LeuMetAlaLeu·
 6301 CGCTCTTTCG CCTGTGCGAGA TCAGCCGTCA GGTCGATGCC TTGATGGCGC
 GCGAGAAAGC GGACAGCTCT AGTCGGCAGT CCAGCTACGG AACTACCGCG
 ·L***
 6351 TG**TAG**TTGCG TCAGACTGTT ATCATATCTT CCCAATTCAA CGGCCATCCT
 ACATCAACGC AGTCTGACAA TAGTATAGAA GGGTTAAGTT GCCGGTAGGA
orf6 (wahD) →
 Met SerLeuIle AsnAspLeu GlnLeuLysArg AspIleLeu·
 6401 GGTGGATTTA TGTCAATTAAT CAACGACCTG CAACTCAAGC GTGACATTCT
 CCACCTAAAT ACAGTAATTA GTTGCTGGAC GTTGAGTTCG CACTGTAAGA
 ·ArgArgLys MetGlyArgPhe LeuPheAsp ArgLysArg ProGluProThr·
 6451 TCGCCGCAAG ATGGGACGGT TCCTGTTTGA TCGAAAACGG CCTGAACCCA
 AGCGGCGTTC TACCCTGCCA AGGACAAACT AGCTTTTGCC GGACTTGGGT
 ·TArgGluSer GluGlyLys ValValLeuLeu ArgTrpAsp AlaLysLeu
 6501 CCAGGGAGAG CGAGGGGAAA GTCGTCCTGC TGCCTGGGA TGCCAAACTG
 GGTCCCTCTC GCTCCCCTTT CAGCAGGACG ACGCGACCCT ACGGTTTGAC
 GlyAspAlaIle ValSerSer TrpValAla ArgGluLeuLys ArgValAsn·
 6551 GGCGATGCGA TCGTCTCGTC CTGGGTGGCC CGCGAACTAA AGCGGGTCAA
 CCGCTACGCT AGCAGAGCAG GACCCACCGG GCGCTTGATT TCGCCCAGTT
 ·ProAlaArg GluValTrpIle ValThrThr AlaAlaMet AlaProLeuPhe·
 6601 CCCGGCACGT GAGGTGTGGA TCGTTACAAC GGCAGCCATG GCTCCCCTGT
 GGGCCGTGCA CTCCACACCT AGCAATGTTG CCGTCGGTAC CGAGGGGACA
 ·PArgAspHis PheGlyMet AspArgValIle GluValPro LysArgPro
 6651 TTCGCGATCA TTTCCGGCATG GATAGGGTCA TCGAGGTGCC CAAGCGCCCC
 AAGCGTAGT AAAGCCGTAC CTATCCCAGT AGCTCCACGG GTTCGCGGGG
 GlyTyrCysGln LeuArgArg LeuAlaLeu SerLeuGlyGlu ValGlyCys·
 6701 GGTATTGTC AACTCAGGCG GCTTGCACTG TCATTGGGGG AGGTCGGTTG
 CCAATAACAG TTGAGTCCGC CGAACGTGAC AGTAACCCCC TCCAGCCAAC
 ·LeuValHis PheGlyLysMet LeuLysMet LysAspIle PhePheIleSer·
 6751 CCTGGTTAC TTTGGCAAGA TGTTGAAGAT GAAGGATATC TTTTTCATCA
 GGACCAAGTG AAACCGTTCT ACAACTTCTA CTTCCTATAG AAAAAGTAGT

·SGlnValArg AlaAspTrp ValAlaGlyMet AspAspAla LeuGlnAla
 6801 GCCAGGTCAG GGCAGATTGG GTCGCCGGCA TGGACGACGC CTTGCAGGCC
 CGGTCCAGTC CCGTCTAACC CAGCGGCCGT ACCTGCTGCG GAACGTCCGG
 IleAsnLeuLys LeuGlyAla AlaSerArg GlyMetHisPhe SerAspLys·
 6851 ATCAATCTCA AGCTGGGAGC CGCGTCCAGA GGTATGCACT TCTCCGACAA
 TAGTTAGAGT TCGACCCTCG GCGCAGGTCT CCATACGTGA AGAGGCTGTT
 ·PheAlaLeu LeuLeuAsnArg LeuGlyVal AlaAspPro AspGlnArgTyr·
 6901 GTTTGCCCTG CTGTTGAATC GACTCGGAGT GGCAGACCCG GATCAGCGCT
 CAAACGGGAC GACAACCTAG CTGAGCCTCA CCGTCTGGGC CTAGTCGCGA
 ·TlleIlePro AlaProGly GlnGluAlaSer ValAlaSer TrpTrpPro
 6951 ACATCATCCC CGCTCCGGGT CAGGAGGCGT CCGTTGCGAG CTGGTGGCCC
 TGTAGTAGGG GCGAGGCCCA GTCCTCCGCA GCCAACGCTC GACCACCGGG
 AspAspValVal LeuCysPhe AsnProTyr GlySerGlySer SerArgArg·
 7001 GATGATGTCG TGCTCTGCTT CAACCCTTAT GGCAGCGGTA GTTCCC GCCG
 CTACTACAGC ACGAGACGAA GTTGGGAATA CCGTCGCCAT CAAGGGCGGC
 ·LeuLysPro GluLeuIleArg GlnMetLeu GluArgMet LeuAspAlaCys·
 7051 TCTGAAGCCG GAGCTGATCC GGCAGATGCT GGAGAGGATG CTGGACGCCT
 AGACTTCGGC CTCGACTAGG CCGTCTACGA CCTCTCCTAC GACCTGCGGA
 ·CProPheAla IleCysLeu LeuTyrProPro GlyMetGlu AlaGluVal
 7101 GCCCATTTCG CATCTGCCTG CTTTATCCAC CGGGTATGGA GGCTGAAGTA
 CGGGTAAACG GTAGACGGAC GAAATAGGTG GCCCATACT CCGACTTCAT
 AlaGlyIleArg AlaGlyLeu ThrArgPro GluArgMetVal LeuSerPro·
 7151 GCGGGGATCC GGGCCGGGCT CACCCGACCC GAGCGCATGG TATTGTACCC
 CGCCCCTAGG CCCGGCCCGA GTGGGCTGGG CTCGCGTACC ATAACAGTGG
 ·AspAlaPro SerLeuAlaGly LeuPheAla GlnLeuArg GlnSerGlnGly·
 7201 CGATGCGCCC TCGCTGGCGG GCCTTTTCGC CCAGTTGCGT CAGAGCCAGG
 GCTACGCGGG AGCGACCGCC CGGAAAAGCG GGTCAACGCA GTCTCGGTCC
 ·GLeuIleSer ValAspThr AlaThrIleHis IleAlaAla GlyLeuGlu
 7251 GCCTCATTTT GGTGGATACG GCGACCATAC ATATCGCGGC CGGGCTGGAG
 CGGAGTAAAG CCACCTATGC CGCTGGTATG TATAGCGCCG GCCCGACCTC
 GlnProIleLeu GlyLeuTyr AsnProAsp IleGlyGlyGly AsnGluAsn·
 7301 CAACCCATTC TGGGTCTCTA CAACCCGGAT ATAGGGGGTG GCAACGAGAA
 GTTGGGTAAG ACCCAGAGAT GTTGGGCCTA TATCCCCAC CGTTGCTCTT
 ·PheLeuGlu TrpHisProAsn SerProLeu AlaThrVal LeuPheSerGlu·
 7351 TTTCTTGAA TGGCATCCA ACTCCCCGTT GGCGACCGTG CTGTTCTCGG
 AAAGGACCTT ACCGTAGGGT TGAGGGGCAA CCGCTGGCAC GACAAGAGCC
 ·GAsnLeuArg GluGlnAsp IleAsnSerLeu AsnLeuThr GluPheAsp
 7401 AGAATCTCCG CGAACAGGAC ATCAACTCTC TCAACCTGAC CGAGTTCGAC
 TCTTAGAGGC GCTTGTCTG TAGTTGAGAG AGTTGGACTG GCTCAAGCTG

ArgLeuPheArg GlnTrpLeu LeuAlaAsn GlnAlaAlaLeu SerValGly·
 7451 CGGCTATTCA GGCAGTGGCT ACTTGCCAAT CAGGCAGCGC TGTCAGTCGG
 GCCGATAAGT CCGTCACCGA TGAACGGTTA GTCCGTCGCG ACAGTCAGCC
 *** AlaThrAlaVal GlnTrpAsp ProLeuAla ThrLeuArgAla·
 ·GlnSerPro AlaGlnAla***
 7501 GCAATCACCA GCTCAGGCGT AATGGCGCGC AACAGGGTTT CATCGATGAA
 CGTTAGTGGT CGAGTCCGCA TTACCGCGCG TTGTCCCAA GTAGCTACTT
 ..IleValLeu GluProThr IleAlaArgLeu LeuThrGlu AspIlePhe
 7551 GTCTTTCTCC GCATTTTGCG GCAGCTGGAA GTCGGCGCCC TTGGCATCGA
 CAGAAAGAGG CGTAAAACGC CGTCGACCTT CAGCCGCGGG AACCGTAGCT
 AspLysGluAla AsnGlnPro LeuGlnPhe AspAlaGlyLys AlaAspPhe·
 7601 AGATATCGAT GGCCTGATAG CGGGGTGGT TGACCGGATT GGCGATGATC
 TCTATAGCTA CCGGACTATC GCCCCACCA ACTGGCCTAA CCGCTACTAG
 .IleAspIle AlaGlnTyrArg ProHisAsn ValProAsn AlaIleIleLys·
 7651 TTGTCGCCGC CGGGAGCCAC TATGGTCAGC AGGGGCAGGT CGATGGCCTG
 AACAGCGGCG GCCCTCGGTG ATACCAGTCG TCCCCGTCCA GCTACCGGAC
 ..AspGlyGly ProAlaVal IleThrLeuLeu ProLeuAsp IleAlaGln
 7701 GGCCATGTGG CGGGGGCCGG TATCGTTGCC GACAAACAGT TGGCAATGGG
 CCGGTACACC GCCCCCGGCC ATAGCAACGG CTGTTTGTCA ACCGTTACCC
 AlaMetHisArg ProGlyThr AspAsnGly ValPheLeuGln CysHisAla·
 7751 CGAACAGCCC TACCAGCTCC CGTACCGAGC GGGTCTTGAT CTGATCAAAC
 GCTTGTCGGG ATGGTCGAGG GCATGGCTCG CCCAGAACTA GACTAGTTTG
 .PheLeuGly ValLeuGluArg ValSerArg ThrLysIle GlnAspPheVal·
 7801 ACGCTTCCCT GCAAGGCGGC AGGCAGCAGC GGCTTGAGAC CGCGGTTGTA
 TGCGAAGGGA CGTTCCGCCG TCCGTCTGTC CCGAACTCTG GCGCCAACAT
 ..SerGlyGln LeuAlaAla ProLeuLeuPro LysLeuGly ArgAsnTyr
 7851 CTCCTCTTCT CCCGGCCCGC AGTAGATCAG GATCTGGGCG TTGTGCTGTT
 GAGGAGAAGA GGGCCGGGCG TCATCTAGTC CTAGACCCGC AACACGACAA
 GluGluGluGly ProGlyCys TyrIleLeu IleGlnAlaAsn HisGlnGlu·
 7901 CGATGAGCCA CTGGGCAACC TGGGCGAAGC AGTCGATGGG CCAGCGCTTG
 GCTACTCGGT GACCCGTTGG ACCCGCTTCG TCAGCTACCC GGTCGCGAAC
 .IleLeuTrp GlnAlaValGln AlaPheCys AspIlePro TrpArgLysTyr·
 7951 TAGTGCCCCA GCGAGTTGAC GCCCATGGCG ATCAAGGGCC GTGCGAGATC
 ATCACGGGGT CGCTCAACTG CGGGTACCGC TAGTTCCCGG CACGCTCTAG
 ..HisGlyLeu SerAsnVal GlyMetAlaIle LeuProArg AlaLeuAsp
 8001 GACGCCAGCC GCCTGCAAGG TTTCCCGTCC CTGCAGCTTC TCCTCTTCGC
 CTGCGGTCCG CGGACGTTCC AAAGGGCAGG GACGTCGAAG AGGAGAAGCG
 ValGlyAlaAla GlnLeuThr GluArgGly GlnLeuLysGlu GluGluSer·

8051 TCAGCCAGAG ATGGTAATTG CGATCTTCTT TGGCCAGTTT GAGCGGTTTG
 AGTCGGTCTC TACCATTAAC GCTAGAAGAA ACCGGTCAAA CTCGCCAAAC
 .LeuTrpLeu HisTyrAsnArg AspGluLys AlaLeuLys LeuProLysLeu·

8101 AGCAGCGAGA AGCGGTCGTC GATCAGGGTA CTGTTGCCCT GGAATTCGGT
 TCGTCGCTCT TCGCCAGCAG CTAGTCCCAT GACAACGGGA CCTTAAGCCA
 ..LeuSerPhe ArgAspAsp IleLeuThrSer AsnGlyGln PheGluThr

8151 GTCTTTGCGA AAGCGCACTA TCGGTTGTG GCCGAGCCGC CAGGGGAAGG
 CAGAAACGCT TTCGCGTGAT ACGCCAACAC CGGCTCGGCG GTCCCCTTCC
 AspLysArgPhe ArgValIle ArgAsnHis GlyLeuArgTrp ProPheSer·

8201 AGTCAAATCC TATCCGCAAC CTGGCGCCGG AGAAGAGGCA GGTCAGCAGG
 TCAGTTTAGG ATAGGCGTTG GACCGCGGCC TCTTCTCCGT CCAGTCGTCC
 .AspPheGly IleArgLeuArg AlaGlySer PheLeuCys ThrLeuLeuGly·

8251 CCTATCATCT GGCCCTGGCT GTTGACTACC ACGTCATAGT GTTCGCGGCG
 GGATAGTAGA CCGGGACCGA CAACTGATGG TGCAGTATCA CAAGCGCCGC
 ..IleMetGln GlyGlnSer AsnValValVal AspTyrHis GluArgArg

8301 GATGTCCAGC AGCAGGTCTT TCATGTAGCG CCAGTCGTTG CGCCTGGCTT
 CTACAGGTCG TCGTCCAGAA AGTACATCGC GGTCAGCAAC GCGGACCGAA
 IleAspLeuLeu LeuAspLys MetTyrArg TrpAspAsnArg ArgAlaLys·

8351 TCTCGATGAT GAGGAGTCGA TCGATATTGG GATTGCCCTG CGCCATACCC
 AGAGCTACTA CTCCTCAGCT AGCTATAACC CTAACGGGAC GCGGTATGGG
 .GluIleIle LeuLeuArgAsp IleAsnPro AsnGlyGln AlaMetGlyAla·

8401 GCGCAATAAT CCATGACCAG ATAGTGA ACT TCCCCGGTTG GATCGTTCTG
 CGCGTTATTA GGTACTGGTC TATCACTTGA AGGGGCCAAC CTAGCAAGAC
 ..CysTyrAsp MetValLeu TyrHisValGlu GlyThrPro AspAsnGln

8451 CTTGAGGTTG TTGCACAGCG CGGTGGAGAT CAGTACATCC CCTATCTGTT
 GAACTCCAAC AACGTGTCGC GCCACCTCTA GTCATGTAGG GGATAGACAA
 LysLeuAsnAsn CysLeuAla ThrSerIle LeuValAspGly IleGlnLys·

8501 TGTTTTGTAC TACAAGGTAT TTCATAAGGC AGGCCTCAAT AGAGGTCGTT
 ACCAAACATG ATGTTCCATA AAGTATTCCG TCCGGAGTTA TCTCCAGCAA
 .ThrGlnVal ValLeuTyrLys Met

← orf7 (*wahE*)

8551 GTTGAGCATC AGCGTGACAG CATAACGCAC AAGCGAGTCA GATGCTGTG
 CAACTCGTAG TC[CAC]TGTC GTATGGC]GTG TTCGCTCAGT CTACGACAGC

8601 GTCAGGAATG AAAATCCGAT CTGGCTCAAG CCAATGGCGG TGTGGTGCAG
 [C]AGTCCTTAC TTTTAGGCTA GACCGAGTTC GGTTACCGCC ACACCACGTC

8651 GCATGGAGTA GAAGCCCCTC GTTGACGGGC TTCTCGCATT CATTACTCGA
 CGTACCTCAT CTTCGGGCAG CAACTGCCCG AAGAGCGTAA GTAATGAGCT

8701 TGTTCTGGAT CTGCTCACGC ATCTGCTCGA TGAGCACCTT CAACTCGACG
 ACAAGACCTA GACGAGTGCG TAGACGAGCT ACTCGTGGAA GTTGAGCTGC

8751 GCGGACTGGG TCACCTCGGC GTTGATGGAC TTGGAGCCCA GCGTGTTCGA
CGCCTGACCC AGTGGAGCCG CAACTACCTG AACCTCGGGT CGCACAAGCT
8801 CTCGCGGTTG AACTCCTGCA TCATGAAGTC CAGACGGCGA CCACAGGCGC
GAGCGCCAAC TTGAGGACGT AGTACTTCAG GTCTGCCGCT GGTGTCCGCG
8851 CACCCTTCTT CATGATCTTG ATGGTTTCGC TGATGTGCAT CTCCAGTCGA
GTGGGAAGAA GTACTAGAAC TACCAAAGCG ACTACACGTA GAGGTCAGCT

8.2 Secuencia nucleotídica de la región 2 *wa* del núcleo del LPS de *Aeromonas hydrophila* AH-3

Secuencia de nucleótidos de la región 2 *wa* del núcleo del lipopolisacárido de *A. hydrophila* AH-3, de su entorno en el genoma y de las secuencias de aminoácidos deducidas de las distintas pautas abiertas de lectura. Los codones de inicio de la traducción se indican en negrita y, los de final de lectura, en negrita y cursiva. Las flechas señalan la dirección en la que se traducen las diferentes pautas abiertas de lectura. Los putativos promotores se indican en cajas, el posible inicio de transcripción está sombreado en amarillo, los putativos terminadores Rho-independientes se han subrayado con una línea continua y las putativas secuencias de unión a ribosomas, con una línea en acordeón.

En las secuencias aminoacídicas, se ha marcado el dominio conservado de la familia 1 de glicosiltransferasas en azul; el de la familia 2, en lila; y el de la familia 9, en verde. Los putativos dominios transmembrana están indicados en naranja.

```

1   GGTACCGCAC TTCTTCGTCG GCGCGGCCGC CGGGGTATTC GGCAATGCCA
   CCATGGCGTG AAGAAGCAGC CGCGCCGGCG GCCCCATAAG CCGTTACGGT
51  CCGGTGGTCG TCGTGGCGCC ATTCTGGGCT CCTTCGCCAA CGGCCTGTTG
   GGCCACCAGC AGCACCGCGG TAAGACCCGA GGAAGCGGTT GCCGGACAAC
101 TTGCTCACCA TTCTGCCGGT GTTCTGCTG CCGGTACTGG GCGATCTTGG
   AACGAGTGGT AAGACGGCCA CAAGGACGAC GGCCATGACC CGCTAGAACC
151 TTTTGCCAAC ACGACCTTCA GCGACTCCGA TTTCGGCGTG CTGGGCATCC
   AAAACGGTTG TGCTGGAAGT CGCTGAGGCT AAAGCCGCAC GACCCGTAGG
201 TGCTGGGGCT GATTGTTCGC TAATCCCGGT CAGTGCCAAA AACAGAAACG
   ACGACCCCGA CTAACAAGCG ATTAGGGCCA GTCACGGTTT TTGTCTTTGC
251 GAGCGCATCG CGCTCCGTTT TTTATGGCTG GCCATCTTGT CTGAATGGCC
   CTCGCGTAGC GCGAGGCAAA AAATACCGAC CGGTAGAACA GACTTACCGG
301 GGATTTGAAT GGCTGGGTGT CAGTCGTTTT CCAACGCAGC GAGGCAATGA
   CCTAAACTTA CCGACCCACA GTCAGCAAAA GGTTGCGTCG CTCCGTTACT
                                     *** AspAsnGlu LeuAlaAla LeuCysHisSer
351 GACAGGGTGC GCGCCAGCGC CCCCTGATTG CGCAGCACCA CGGCACGGGC
   CTGTCCCACG CGCGGTTCGCG GGGGACTAAC GCGTCGTGGT GCCGTGCCCG
   ..LeuThrArg AlaLeuAla GlyGlnAsnArg LeuValVal AlaArgAla
401 CTGCTCCCC ATCTTGCGGC GGGCTGTCAC GTCATCAAAC AGCTGGCTCA
   GACGAGGGGG TAGAACGCCG CCCGACAGTG CAGTAGTTTG TCGACCGAGT
   GlnGluGlyMet LysArgArg AlaThrVal AspAspPheLeu GlnSerVal

```

451 CCTGCTCGCC CAGTGCGGCG GCATCGGCCA CGATAACGGC GCCCCCTGA
 GGACGAGCGG GTCACGCCGC CGTAGCCGGT GCTATTGCCG CGGGGGGACT
 .GlnGluGly LeuAlaAlaAla AspAlaVal IleValAla GlyGlyGlnAla·

501 GCCACCAGCT GCGGGTGAT ATCGCTGAAG TTGAAGTAGG CGGGGCCAGT
 CGGTGGTCTGA CCGCCCCTA TAGCGACTTC AACTTCATCC GCCCCGGTCA
 .ValLeuGln ArgThrIle AspSerPheAsn PheTyrAla ProGlyThr

551 CAGACAGGGC TTGCCGAGCG CCGCAGGTTC CAGCAGGTTG TGACCACCGA
 GTCTGTCCCG AACGGCTCGC GCGTCCAAG GTCGTCCAAC ACTGGTGGCT
 LeuCysProLys GlyLeuAla AlaProGlu LeuLeuAsnHis GlyGlyIle·

601 TCTTACCAG ACTGCCGCCG ACAAAGGCCA CATCGGCCGC CGCCAGCATC
 AGAAGTGGTC TGACGGCGGC TGTTCGGGT GTAGCCGGCG GCGGTCGTAG
 .LysValLeu SerGlyGlyVal PheAlaVal AspAlaAla AlaLeuMetLeu·

651 AGCGGCAACT CACCCATGGT ATCGCCGAGA TAGACCTTGT GCTGCGCGCC
 TCGCCGTTGA GTGGGTACCA TAGCGGCTCT ATCTGGAACA CGACGCGCGG
 .ProLeuGlu GlyMetThr AspGlyLeuTyr ValLysHis GlnAlaGly

701 TATCACGGCA TGACTGGTAC GCCGCACACA GCCATAAGGG GCACACAACCT
 ATAGTGCCGT ACTGACCATG CGGCGTGTGT CGGTATTCCC CGTGTGTTGA
 IleValAlaHis SerThrArg ArgValCys GlyTyrProAla CysLeuGlu·

751 CGGCCACCCG GTCGAACCGC TCGGGGTGGC GCGGCACCAG GATCAGCAGC
 GCCGGTGGGC CAGCTTGGCG AGCCCCACCG CGCCGTGGTC CTAGTCGTCCG
 .AlaValArg AspPheArgGlu ProHisArg ProValLeu IleLeuLeuAla·

801 GCATCGGGAT GCGCCTTGAG CAGCAGATCG AAGGCGGCCA GCACCTGCTC
 CGTAGCCCTA CCGCGAACTC GTCGTCTAGC TTCCGCCGGT CGTGGACGAG
 .AspProHis ArgLysLeu LeuLeuAspPhe AlaAlaLeu ValGlnGlu

851 GTCCTCACCC TGATGGGTGC TGGCGGCGAT CCAGACCGGC CTGCTCTGCC
 CAGGAGTGGG ACTACCCACG ACCGCCGCTA GGTCTGGCCG GACGAGACGG
 AspGluGlyGln HisThrSer AlaAlaIle TrpValProArg SerGlnGly·

901 CCAGCTGCTG ACGCAGGGCG CGACCTTGG CCTGTACCTC GTCCCCCAGC
 GGTGCGACGAC TGCGTCCCGC GCTGGGAACC GGACATGGAG CAGGGGGTCCG
 .LeuGlnGln ArgLeuAlaArg GlyLysAla GlnValGlu AspGlyLeuGln·

951 TGAATGTCGA ATTTGATGGA ACCGGTCACT GCCAGCCGCG CGCGGCTGAT
 ACTTACAGCT TAAACTACCT TGGCCAGTGA CGGTCCGGCG GCGCCGACTA
 .IleAspPhe LysIleSer GlyThrValAla LeuArgAla ArgSerIle

1001 GCCGAGGCGG GCGAAACGGT CGGCATCGTC CTGATGCTGG CAAAGCAGAT
 CGGCTCCGCC CGCTTTGCCA GCCGTAGCAG GACTACGACC GTTTCGTCTA
 GlyLeuArgAla PheArgAsp AlaAspAsp GlnHisGlnCys LeuLeuHis·

1051 GAGTCAGCGG ACGGCTCAGG GCGTCGAAGG CGCCGTGGAA ACGGGCATAA
 CTCAGTCGCC TGCCGAGTCC CGCAGCTTCC GCGGCACCTT TGCCCCGATT
 .ThrLeuPro ArgSerLeuAla AspPheAla GlyHisPhe ArgAlaTyrArg·

1101 CGTTGGCAGG ATCGCTCGGA AAGGCGGGCA TTGATGATGG TCACCGGCAG
GCAACCGTCC TAGCGAGCCT TTCCGCCCGT AACTACTACC AGTGGCCGTC
..GlnCysSer ArgGluSer LeuArgAlaAsn IleIleThr ValProLeu

1151 ATGGTGCGCC TCGCAGGCGG CCAGCCAGTT GGGCCACAGC TCTGTCTCCA
TACCACGCGG AGCGTCCGCC GGTCGGTCAA CCCGGTGTCTG AGACAGAGGT
HisHisAlaGlu CysAlaAla LeuTrpAsn ProTrpLeuGlu ThrGluMet.

1201 TCACCCAGAG GGCGCGCGGC TTGATGCGCT TGAGGAAGGC GGCCACCGCC
AGTGGGTCTC CCGCGCGCCG AACTACGCGA ACTCCTTCCG CCGGTGGCGG
.ValTrpLeu AlaArgProLys IleArgLys LeuPheAla AlaValAlaTrp.

1251 CAGGGGTAGT CGAGCGGTGC GTAGCGATGC ACCACCAGAT CTCCCAGCTT
GTCCCCATCA GCTCGCCACG CATCGCTACG TGGTGGTCTA GAGGGTCGAA
..ProTyrAsp LeuProAla TyrArgHisVal ValLeuAsp GlyLeuLys

1301 GGCGGCCTGT TCTGCTCCGG TCGGGGTGGT TGTGGTGAGC AGGATCGGCA
CCGCCGGACA AGACGAGGCC ACGCCCACCA ACACCACTCG TCCTAGCCGT
AlaAlaGlnGlu AlaGlyThr ArgThrThr ThrThrLeuLeu IleProLeu.

1351 AGTCTGGCCG CTCGGCCTTG AGGGCGCGAA TGAAGGGACT GATGGCGAGC
TCAGACCGGC GAGCCGGAAC TCCCAGCCTT ACTTCCCTGA CTACCGCTCG
.AspProArg GluAlaLysLeu AlaArgIle PheProSer IleAlaLeuThr.

1401 GTCTCACCGA CGCTCACCGC ATGGATCCAG AGCGGCGCCT CCTGCCCGCT
CAGAGTGGCT GCGAGTGGCG TACCTAGGTC TCGCCGCGGA GGACGGGCGA
..GluGlyVal SerValAla HisIleTrpLeu ProAlaGlu GlnGlySer

1451 GGCCGGCGTC CATCCCAGAT GCTCGGCCCA GCGCGCACCG AAGCCGGGCT
CCGGCCGCGAG GTAGGGTCTA CGAGCCGGGT CGCGCGTGGC TTCGGCCCCA
AlaProThrTrp GlyLeuHis GluAlaTrp ArgAlaGlyPhe GlyProLys.

1501 TGCCCTTCTT CGGCTTGTAG AGCGCCAGCA ACGCCAGCGG CAAACCCAGG
ACGGGAAGAA GCCGAACATC TCGCGGTCTG TGCGGTGCGC GTTTGGGTCC
.GlyLysLys ProLysTyrLeu AlaLeuLeu AlaLeuPro LeuGlyLeuHis.

1551 TGGATCAGCA GGTTGTAGAG CAACCGATAG CCCATGCTAT TTCGGCTCCG
ACCTAGTCGT CCAACATCTC GTTGGCTATC **GGTACGATA** AAGCCGAGGC
..IleLeuLeu AsnTyrLeu LeuArgTyrGly Met *****Lys ProGluSer.**

← orf1 (waaA)

1601 ATCGCTTGCC CGGGTCGGAG ACGAAGCGCG GCGGCTCGCC CACTTGCAGC
TAGCGAACGG GCCCAGCCTC TGCTTCGCGC CGCCGAGCGG GTGAACGTCTG
.ArgLysGly ProAspSerVal PheArgPro ProGluGly ValGlnLeuVal.

1651 ACGATCGGTT GAGCCGGCGA CTCGCCGCAC TCCAGTCGCC TGCCCGCCTC
TGCTAGCCAA CTCGGCCGCT GAGCGGCGTG AGGTCAGCGG ACGGGCGGAG
..IleProGln AlaProSer GluGlyCysGlu LeuArgArg GlyAlaGlu

1701 TATCACCTGC TCCGGCATCA GCTCGATCAG GCATTTGAGG TGACCGAACT
 ATAGTGGACG AGGCCGTAGT CGAGCTAGTC CGTAAACTCC ACTGGCTTGA
 IleValGlnGlu ProMetLeu GluIleLeu CysLysLeuHis GlyPheLys·

1751 TGCAGGTGCG CTTGAAGCAG GGGCGGCACT CTATGTCGGT ATGTACTATC
 ACGTCCACGC GAACTTCGTC CCCGCCGTGA GATACAGCCA TACATGATAG
 .CysThrArg LysPheCysPro ArgCysGlu IleAspThr HisValIleGlu·

1801 TCAACCCTGT CCGCCAGCGG CGGCGTGTAG AGCGGCGAGG TGGAGCCATA
 AGTTGGGACA GGCGGTTCGCC GCCGCACATC TCGCCGCTCC ACCTCGGTAT
 ..ValArgAsp AlaLeuPro ProThrTyrLeu ProSerThr SerGlyTyr

1851 GACGCCGATC AACGGCCGGT TGAGGGCAGC AGCTATGTGC ATCAGGCCGG
 CTGCGGCTAG TTGCCGGCCA ACTCCCGTCG TCGATACACG TAGTCCGGCC
 ValGlyIleLeu ProArgAsn LeuAlaAla AlaIleHisMet LeuGlySer·

1901 AATCATTGGC GATGACCCGC CCGGCCAGCG CCATCAGGTC GATCGCCTCA
 TTAGTAACCG CTACTIONGGCG GGCCGGTCGC GGTAGTCCAG CTAGCGGAGT
 .AspAsnAla IleValArgGly AlaLeuAla MetLeuAsp IleAlaGluTyr·

1951 TACAGGCTGG TCTTGCCCGC CAGCACATGG CAATTGGGGC GCGACAGGGG
 ATGTCCGACC AGAACGGGCG GTCGTGTACC GTTAACCCCG CGCTGTCCCC
 ..LeuSerThr LysGlyAla LeuValHisCys AsnProArg SerLeuPro

2001 ATTGATCCGG TCGCGAATGG TGTGGCCAC GGGCACATCC TTGCCGGAGC
 TAACTAGGCC AGCGCTTACC ACAACCGGTG CCCGTGTAGG AACGGCCTCG
 AsnIleArgAsp ArgIleThr AsnAlaVal ProValAspLys GlySerGly·

2051 CGAAGATCCA CACCTGCCAG CCCTCATCCA GGTGCTTCTG CGCCACCACG
 GCTTCTAGGT GTGGACGGTC GGGAGTAGGT CCACGAAGAC GCGGTGGTGC
 .PheIleTrp ValGlnTrpGly GluAspLeu HisLysGln AlaValValAla·

2101 GCGTAATGCC CTTCCGGCCA GCGCTTGGCC GGGCCGAACT CCGCCCCCGG
 CGCATTACGG GAAGGCCGGT CGCGAACCGG CCCGGCTTGA GGCGGGGGCC
 ..TyrHisGly GluProTrp ArgLysAlaPro GlyPheGlu AlaGlyPro

2151 GCAGAGCACC AGCACGGGGC GGGCGCGATC CAGCCCCAGC CGGTCCAGCG
 CGTCTCGTGG TCGTGCCCCG CCCGCGCTAG GTCGGGGTCG GCCAGTTCGC
 CysLeuValLeu ValProArg AlaArgAsp LeuGlyLeuArg AspLeuAla·

2201 CCTTGCCCTG ATTGATGAGA TCCACATTGA GCGACGGATG GGAATGACG
 GGAACGGGAC TAACTACTCT AGGTGTA ACT CGCTGCCTAC CCCTTACTGC
 .LysGlyGln AsnIleLeuAsp ValAsnLeu SerProHis ProIleValPro·

2251 GGAATGTCCG CCCGGCTCTT CATCTGCGCC TTGGGATAGG CCAGCGCCAG
 CCTTACAGCC GGGCCGAGAA GTAGACGCGG AACCTATCC GGTCGCGGTC
 ..IleAspAla ArgSerLys MetGlnAlaLys ProTyrAla LeuAlaLeu

2301 GTAGGCTTCC ACCATCAGCG GAAAGGCTGA CTTGTTGCTG CGCATGTCGT
 CATCCGAAGG TGGTAGTCGC CTTTCCGACT GAACAACGAC GCGTACAGCA
 TyrAlaGluVal MetLeuPro PheAlaSer LysAsnSerArg MetAspAsn·

2351 TGAGCAGACC GAAGCGGTGC TCCCCTTTCC AGCCAGTGCG TACCGGGATC
 ACTCGTCTGG CTTCGCCACG AGGGGAAAGG TCGGTCACGC ATGGCCCTAG
 .LeuLeuGly PheArgHisGlu GlyLysTrp GlyThrArg ValProIleArg·

2401 CTGGCAAACC AGGGGATCAG CGCCGACTTC ATGGAGTTGG GCTGGATGAT
 GACCGTTTGG TCCCCTAGTC GCGGCTGAAG TACCTCAACC CGACCTACTA
 ..AlaPheTrp ProIleLeu AlaSerLysMet SerAsnPro GlnIleIle

2451 CGCCTGATCG TACTTCTCGG CTGCCAGCTG CTTGCCGAGC CGACGGCGGG
 GCGGACTAGC ATGAAGAGCC GACGGTCGAC GAACGGCTCG GCTGCCGCCC
 AlaGlnAspTyr LysGluAla AlaLeuGln LysGlyLeuArg ArgArgAla·

2501 CCCCCAGCTT GAAGTCACCG TGGCCGAGCG GCATGGGAAT GGCTTTGTCC
 GGGGGTCGAA CTTCAGTGGC ACCGGCTCGC CGTACCCTTA CCGAAACAGG
 .GlyLeuLys PheAspGlyHis GlyLeuPro MetProIle AlaLysAspVal·

2551 ACCTCGGGCG TTCGCTCCAG CAGGGCACCA CACCAGGCGG GCGCCATCAC
 TGGAGCCCGC AAGCGAGGTC GTCCCGTGGT GTGGTCCGCC CGCGGTAGTG
 ..GluProThr ArgGluLeu LeuAlaGlyCys TrpAlaPro AlaMetVal

2601 ATGCAACTCG CTGTCCGGGT GATTGGCCTT GATCGCCTTG TAGAGGCTCT
 TACGTTGAGC GACAGCCCCA CTAACCGGAA CTAGCGGAAC ATCTCCGAGA
 HisLeuGluSer AspProHis AsnAlaLys IleAlaLysTyr LeuSerGln·

2651 GGGACATCAC CATGTCGCCG ACCCAGGAGG GGCCGATCAC CAATATCTTC
 CCCTGTAGTG GTACAGCGGC TGGGTCCTCC CCGGCTAGTG GTTATAGAAG
 .SerMetVal MetAspGlyVal TrpSerPro GlyIleVal LeuIleLysMet·

2701 ATCTCACGCG TCTCTTGGGG CCAAGCCCTG CGGTAAATCA TTACTGGCCG
 TAGAGTGCAGC AGAGAACCCC GGTTCGGGAC GCCAATTAGT AATGACGGC
 ..

← **orf2 (wahF)**

2751 AGATTATGCC CGAAGAATGG CCCGAGAAAT AACAGAAGAT CCCC GTACCA
 TCTAATACGG GCTTCTTACC GGCTCTTTTA TTGTCTTCTA GGGGCATGGT

2801 CCCGCTCTCA TGCCAAAAA GACTCCGTTT GAAACTGACG CACTGCGGGC
 GGGCGAGAGT ACGGTTTTTT CTGAGGCAAA CTTT GACTGC GTGACGCCCG

2851 GCCGGAACCA TGCTCGCCT GCCGTTTTAG GCTAGAATGG GCGCAATTC
 CGGCCTTGGT ACGAGGCGGA CGGCAAAATC CGATCTTACC CGCGGTTAAG

orf3 (waaE) →

ValLysLys·

2901 GAGCATGGCC CGCCAGCCGG GCGCCAACAA GAGCCCTTAA TCCGTGAAAA
 CTCGTACCGG GCGGTCGGCC CGCGTTTGTT CTCGGGAATT AGGCACTTTT
 ·LProThrLeu AlaAlaVal LeuIleValLys AsnGluAla AlaAsnLeu

2951 AACCGACCCT GGCGGCCGTC CTGATCGTCA AGAATGAAGC CGCAAACCTG
 TTGGCTGGGA CCGCCGGCAG GACTAGCAGT TCTTACTTCG GCGTTTGGAC

ArgAlaCysLeu AlaSerLeu AspGlyLeu ValAspGluIle ValIleMet·
 3001 AGAGCCTGCC TCGCCTCCCT CGATGGCCTG GTGGACGAGA TCGTCATCAT
 TCTCGGACGG AGCGGAGGGA GCTACCGGAC CACCTGCTCT AGCAGTAGTA
 ·AspSerGly SerGlnAspGlu ThrProAla IleAlaAla GluPheGlyAla·
 3051 GGA CTCAGGC AGCCAGGACG AAACCCAGC CATCGCCGCC GAATTTGGCG
 CCTGAGTCCG TCGGTCCCTGC TTTGGGGTCG GTAGCGGCGG CTTAAACCGC
 ·AArgPhePhe ValAsnPro ValTrpProGly PheGlyArg GlnArgArg
 3101 CCCGCTTCTT CGTCAACCCG GTATGGCCCG GCTTCGGTCG CCAGCGCCCG
 GGGCGAAGAA GCAGTTGGGC CATAACCGCC CGAAGCCAGC GGTCGCGGCG
 LeuAlaGlnSer HisValGln SerGluTrp ValLeuTrpLeu AspAlaAsp·
 3151 CTCGCCAGT CCCACGTCCA GTCAGAATGG GTGCTGTGGC TGGATGCCGA
 GAGCGGGTCA GGGTGCAGGT CAGTCTTACC CACGACACCG ACCTACGGCT
 ·GluArgLeu ThrProGluLeu LysAlaAla IleAlaThr ValMetAlaAsn·
 3201 CGAGCGGCTC ACTCCCGAGC TGAAAGCGGC CATCGCCACC GTCATGGCCA
 GCTCGCCGAG TGAGGGCTCG ACTTTCGCCG GTAGCGGTGG CAGTACCGGT
 ·AProAlaSer AspThrIle TyrSerIlePro ArgLeuSer TrpValPhe
 3251 ACCCGGCCAG CGATACCATC TACTCGATCC CGCGCCTCTC CTGGGTGTTT
 TGGGCCGGTC GCTATGGTAG ATGAGCTAGG GCGCGGAGAG GACCACAAA
 GlyArgPheIle ArgHisSer GlyTrpTyr ProAspArgVal LeuArgLeu·
 3301 GGTGCTTCA TCCGTCACAG CGGCTGGTAT CCGGATCGGG TGCTGCGCCT
 CCAGCGAAGT AGGCAGTGTC GCCGACCATA GGCCTAGCCC ACGACGCGGA
 ·TyrProLys AlaLeuThrSer TyrAsnGlu ValLeuVal HisGluLysVal·
 3351 CTACCCAAA GCGCTCACCA GCTACAACGA GGTGCTGGTG CACGAGAAAG
 GATGGGGTTT CGCGAGTGGT CGATGTTGCT CCACGACCAC GTGCTCTTTC
 ·VGluValArg AlaAsnMet LysIleValAsn IleHisGly AspLeuLeu
 3401 TGGAGGTGCG CGCCAACATG AAGATCGTCA ATATCCATGG CGACCTGCTC
 ACCTCCACGC GCGGTTGTAC TTCTAGCAGT TATAGGTACC GCTGGACGAG
 HisPheThrTyr ArgAspLeu GluHisTyr LeuValLysSer AlaGlyTyr·
 3451 CACTTACCT ATCGGGATCT GGAGCACTAT CTGGTGAAGT CCGCCGGCTA
 GTGAAGTGA TAGCCCTAGA CCTCGTGATA GACCACTTCA GGCGGCCGAT
 ·AlaArgAla TrpAlaAspGln ArgThrAla ArgGlyLys LysGlySerLeu·
 3501 TGCCAGGGCC TGGGCCGATC AGCGCACCGC CCGTGGCAAG AAGGGCTCCC
 ACGGTCCCGG ACCCGGCTAG TCGCGTGGCG GGCACCGTTC TTCCCAGGG
 ·LSerGlnGly LeuValHis AlaLeuGlyCys PheLeuLys MetTyrLeu
 3551 TCAGCCAAGG GCTGGTGCAT GCGCTGGGCT GCTTCCTCAA GATGTACCTG
 AGTCGGTTCC CGACCACGTA CGCGACCCGA CGAAGGAGTT CTACATGGAC
 LeuLysAlaGly PheLeuAsp GlyLysGln GlyLeuLeuLeu AlaIleLeu·
 3601 CTCAAGGCGG GCTTCCTCGA TGGCAAGCAG GGACTCTTGC TCGCGATCCT
 GAGTTCCGCC CGAAGGAGCT ACCGTTCTGTC CCTGAGAACG AGCGCTAGGA

·SerAlaHis SerThrPheVal LysTyrAla AspLeuTrp IleArgGlnGln·
 3651 GTCGGCCCAC TCCACCTTCG TCAAATACGC CGACCTCTGG ATCCGCCAGC
 CAGCCGGGTG AGGTGGAAGC AGTTTATGCG GCTGGAGACC TAGGCGGTTCG
 ·GProGlnAla ProAspGln ThrGlyAsn***
 3701 AGCCGCAGGC GCCGGATCAG ACTGGCAACT **AATTGCGCAG** CCGCTCAAGC
 TCGGCGTCCG CGGCCTAGTC TGACCGTTGA TTAACGCGTC GGCGAGTTTCG
 3751 CAGCATGACG AAACAGGGCC CGAAAGGGCC CTGTTTGTAT CTATCAACTC
 GTCGTACTION TTTGTCCCGG GCTTTCCCGG GACAAACATA GAT**AGT**TGAG
 ***SerGlu·
 3801 TCTGTCCGCC AGAAAGGCGG TGCCCAGCGG GCGATCGCCG CCGCCATTGT
 AGACAGGCGG TCTTTCCGCC ACGGGTCGCC CGCTAGCGGC GGCGGTAACA
 ·.ThrArgTrp PheProPro AlaTrpArgAla IleAlaAla AlaMetThr
 3851 CTTGGGATCG AGCTGGCTCA TGTCTTCCGG GTGATCGCTG TCTTCCGGCG
 GAACCCTAGC TCGACCGAGT ACAGAAGGCC CACTAGCGAC AGAAGGCCCG
 LysProAspLeu GlnSerMet AspGluPro HisAspSerAsp GluProPro·
 3901 GGCTGAACGC CAGGTGGCGC CCCTCGCTGT TGAGCGGACG CCAGCGCAGC
 CCGACTTGCG GTCCACCGCG GGGAGCGACA ACTCGCCTGC GGTCCGCTCG
 ·SerPheAla LeuHisArgGly GluSerAsn **LeuProArg TrpArgLeuPro·**
 3951 GGGGTGGCCG AGCGGTGTGC CGGGAAGAAA CCTATGGTGG GCACATCCAG
 CCCCACCGGC TCGCCACACG GCCCTTCTTT GGATAACCACC CGTGTAGGTC
·.ThrAlaSer ArgHisAla ProPhePheGly IleThrPro ValAspLeu
 4001 CGCCGCCGCT ATGTGCAGCG GCCCCGTACT GCCCGCCACG AACAGGGCGG
 GCGGCGGCGA TACACGTCGC CGGGGCATGA CGGGCGGTGC TTGTCCCGCC
AlaAlaAlaIle HisLeuPro GlyThrSer GlyAlaValPhe LeuAlaAla·
 4051 CATTGGCCAT CACCTGACAG AACTTGGGCA GCCCCTCGTC GGAGCGGTAG
 GTAACCGGTA GTGGACTGTC TTGAACCCGT CGGGGAGCAG CCTCGCCATC
·.AsnAlaMet ValGlnCysPhe LysProLeu GlyGluAsp SerArgTyrIle·
 4101 ATCCAGCCGT TGCCGCCGTG GGCCAGCAGC TCGGCCGCCA GCTGTTCGGC
 TAGGTCGGCA ACGGCGGCAC CCGGTCGTGC AGCCGGCGGT CGACAAGCCG
 ·.TrpGlyAsn GlyGlyHis AlaLeuLeuGlu AlaAlaLeu GlnGluAla
 4151 GATCTTCTCC TCCCCCGGCC CCGCCGTCAG TATGCACTGC AACTCGGGGC
 CTAGAAGAGG AGGGGGCCGG GGCGGCAGTC ATACGTGACG TTGAGCCCCG
 IleLysGluGlu GlyProGly AlaThrLeu IleCysGlnLeu GluProCys·
 4201 AAGCCTGGTT GAGCTTGATG ATGAGGCGGG CGTACTGCTC GATGGAGAGG
 TTCGGACCAA CTCGAACCTAC TACTCCGCC GCATGACGAG CTACCTCTCC
 ·AlaGlnAsn LeuLysIleIle LeuArgAla TyrGlnGlu IleSerLeuAsn·
 4251 TTGTTGGCCG AGCCGCCGCT GCCGGCATGC ACCATCAGCC AGGGGCGGCT
 AACAAACCGC TCGGCGGCGA CGGCCGTACG TGGTAGTCGG TCCCCGCCGA
 ·.AsnAlaSer GlyGlySer GlyAlaHisVal MetLeuTrp ProArgSer

4301 GGCATCCAGC TTCAAGCGGG TGGCGACCTG CTCGCGCACC TGCTGCAGAC
 CCGTAGGTCG AAGTTCGCCC ACCGCTGGAC GAGCGCGTGG ACGACGTCTG
 AlaAspLeuLys LeuArgThr AlaValGln GluArgValGln GlnLeuSer·
 4351 TCTGGGGATC GAAGCTGAGA TAGGGGCCCT GCGGTTTCGAC CACGCTCACC
 AGACCCCTAG CTTCTGACTCT ATCCCCGGA CGCCAAGCTG GTGCGAGTGG
 .GlnProAsp PheSerLeuTyr ProGlyGln ProGluVal ValSerValGln·
 4401 TGCTGGTCGG CGAGGAAGCG GCGCACCAGA TCCAGGTTGT ATTCGTACTC
 ACGACCAGCC GCTCCTTCGC CGCGTGGTCT AGGTCCAACA TAAGCATGAG
 ..GlnAspAla LeuPheArg ArgValLeuAsp LeuAsnTyr GluTyrGlu
 4451 CGGCTTCTGG GAGCGGGAAC GACGCTGGAT CAGGCGCTGG TTGTAGAGCA
 GCCGAAGACC CTCGCCCTTG CTGCGACCTA GTCCGCGACC AACATCTCGT
 ProLysGlnSer ArgSerArg ArgGlnIle LeuArgGlnAsn TyrLeuVal·
 4501 CCTGGGCCAG CTTGGTGGCA GGCGCCAACC GGTAAGGGAT GTGTGCCTTC
 GGACCCGGTC GAACCACCGT CCGCGGTTGG CCATTCCCTA CACACGGAAG
 .GlnAlaLeu LysThrAlaPro AlaLeuArg TyrProIle HisAlaLysTrp·
 4551 CACACCAGCA TGGCATTGCG AGAGTTGGAG AAGAGGCAGA TGGAAGCATC
 GTGTGGTTCGT ACCGTAACGC TCTCAACCTC TTCTCCGTCT ACCTTCGTAG
 ..ValLeuMet AlaAsnArg SerAsnSerPhe LeuCysIle SerAlaAsp
 4601 GAAGGCGGCT GCCTTGATGC GCGCCTGCAA GGCACGTTGC TGTTCCTTGT
 CTTCCGCCGA CGGAACTACG CGCGGACGTT CCGTGCAACG ACAAGGAACA
 PheAlaAlaAla LysIleArg AlaGlnLeu AlaArgGlnGln GluLysAsp·
 4651 CGGCTCCTGA ACCCGGATCC AGGATAATTT CGTCGATCCA GGGGCAGAGA
 GCCGAGGACT TGGGCCTAGG TCCTATTAAA GCAGCTAGGT CCCCCTCTCT
 .AlaGlySer GlyProAspLeu IleIleGlu AspIleTrp ProCysLeuArg·
 4701 CGGGCCAACG GCGCCGTGTA GGCAGGAACC AGGGCAGTCA CATGACAATC
 GCCCGGTTGC CGCGGCACAT CCGTCCTTGG TCCCGTCAGT GTACTGTTAG
 ..AlaLeuPro AlaThrTyr AlaProValLeu AlaThrVal HisCysAsp
 4751 CATGGAGCGC TTGAGCATGG CAAAACCTCG CCATGCCAGC ATGAAGTCAC
 GTACCTCGCG AACTCGTACC GTTTTGGAGCC GGTACGGTCG TACTTCAGTG
 MetSerArgLys LeuMetAla PheSerPro TrpAlaLeuMet PheAspGly·
 4801 CAATCTTGTC GTTGCCTACT ACCAGAATAC GTTTCATAAA AAATCCCTTT
 GTTAGAACAG CAACGCATGA TGGTCTTATG CAAAGTATTT TTTAGGGAAA
 .IleLysAsp AsnArgValVal LeuIleArg LysMet
 ← orf4 (waaF)

4851 GCTCCCTGTC ACCAATGCAA GGAGGCTTGA TGC GCGGTGC GGGTATTCTT
 CGAGGGACAG TGGTTACGTT CCTCCGAACT ACGCGCCA CG CCCA TAAGAA
 4901 GCCTGATGCA CAGCTATTCT CAAGCGAAGA CACCAGACGA TACTGATTTG
CGGACTACGT GTCGATA AGA GTTCGCTTCT GTGGTCTGCT ATGACTAAAC
 4951 GACGACCAAA CCGGGCCTCT CGCGCCAGGT CCCGCACAAA ACGGACGTCG

```

          CTGCTGGTTT  GGCCCGGAGA  GCGCGGTCCA  GGGCGTGTTT  TGCCTGCAGC
5001  TCCCGGTCGC  GACAAAACGC  CAACAGACCC  TATAATGCGA  CCAAATTCCC
          AGGGCCAGCG  CTGTTTTGCG  GTTGTCTGGG  ATATTACGCT  GGTTTAAGGG
5051  AGGTTGCAGG  AATCAACAAC  ATGACCAACA  AGGTAATTTA  TCCCGGCACC
          TCCAACGTCC  TTAGTTGTTG  TACTGGTTGT  TCCATTAAAT  AGGGCCGTGG
5101  TTCGACCCCA  TCACCAATGG  CCACACCGAC  CTCATTGGCC  GGGCCGCCAG
          AAGCTGGGGT  AGTGGTTACC  GGTGTGGCTG  GAGTAACCGG  CCCGGCGGTC
5151  ACTGTTTGAT  GAAGTGGTGG  TCGGCGTCGC  CAACAGCCCG  AGCAAGCGCC
          TGACAAACTA  CTTACCACC  AGCCGCAGCG  GTTGTGCGGC  TCGTTGCGGG
5201  CGCTGTTTCA  TCTGGCAGAA  CGTGTCTTGC  TTGCCAGCA  AGTGACCGCA
          GCGACAAGCT  AGACCGTCTT  GCACAGGACG  AACGGGTCGT  TCACTGGCGT
5251  CACCTGCCCA  ACGTCAAGGT  GGTGGGCTTC  TCCGGCCTGC  TGGTGGATTT
          GTGGACGGGT  TGCAGTTCCA  CCACCCGAAG  AGGCCGGACG  ACCACCTAAA
5301  CGCCAGGGAG  CAGCAGGCCA  ACGTGCTGAT  CCGTGGCCTG  CGCGCCGTCT
          GCGGTCCCTC  GTCGTCCGGT  TGCACGACTA  GGCACCGGAC  GCGCGGCAGA
5351  CCGACTTCGA  GTACGAATTC  CAGCTGGCCA  ACATGAACCG  CCGCCTGATG
          GGCTGAAGCT  CATGCTTAAG  GTCGACCGGT  TGTACTTGGC  GGCGGACTAC
5401  CCGGAGCTGG  AGAGCGTCTT  CCTCACTCCG  GCGGAGGAGA  ACTCCTTCAT
          GGCCTCGACC  TCTCGCAGAA  GGAGTGAGGC  CGCCTCCTCT  TGAGGAAGTA
5451  CTCCTCGACC  CTGGTGAAGG  AGGTCGCCCT  GCATGGTGGC  GATATCCGCC
          GAGGAGCTGG  GACCACTTCC  TCCAGCGGGA  CGTACCACCG  CTATAGGCGG
5501  AGTTCGTCGA  CCCAATCGTC  GCCAAGGCGA  TTGCCGCCAA  GCAAGGCAAG
          TCAAGCAGCT  GGGTTAGCAG  CGGTTCCGCT  AACGGCGGTT  CGTTCGGTTC
5551  TGATCAAAAA  CAGGCTACCC  GAAGGTAGCC  TGTTTCTTGA  GGGGCACATT
          ACTAGTTTTT  GTCCGATGGG  CTTCCATCGG  ACAAAGAACT  CCCCCTGTAA
5601  CATCCCGCCA  GCATGGACCG  TCAGCCCGCC  TGGCGAGCCA  GCACCTGCTG
          GTAGGGCGGT  CGTACCTGGC  AGTCGGGCGG  ACCGCTCGGT  CGTGGACGAC

```

8.3 Secuencia nucleotídica de la región 3 *wa* del núcleo del LPS de *Aeromonas hydrophila* AH-3

Secuencia de nucleótidos de la región 3 *wa* del núcleo del lipopolisacárido de *A. hydrophila* AH-3, de su entorno en el genoma y de las secuencias de aminoácidos deducidas de las distintas pautas abiertas de lectura. Los codones de inicio de la traducción se indican en negrita y, los de final de lectura, en negrita y cursiva. Las flechas señalan la dirección en la que se traducen las diferentes pautas abiertas de lectura. Los putativos promotores se indican en cajas, el posible inicio de transcripción está sombreado en amarillo, los putativos terminadores Rho-independientes se han subrayado con una línea continua y las putativas secuencias de unión a ribosomas, con una línea en acordeón.

En las secuencias aminoacídicas, se ha marcado el dominio conservado de la familia 9 de glicosiltransferasas en verde.

```

1   GGGAAACCGA CAGTTGCATC TTGCCTGCTG CCTTGTGTTT GGTCAGCAAG
    CCCTTTGGCT GTCAACGTAG AACGGACGAC GGAACACAAA CCAGTCGTTC
51  GCATTGAGCA GGCTGGATTT ACCCGCCCCG CTCGGGGAGG AGATGATATA
    CGTAACTCGT CCGACCTAAA TGGGCGGGGC GAGCCCCTCC TCTACTATAT
101 GAGAGTGCCT TGTTGCGCCA TGTCATTACC TTTAATCGAA ACCCGTGTTT
    CTCTCACGGA ACAACGCGGT ACAGTAATGG AAATTAGCTT TGGGCACAAA
151 TGTTTGTAGC AGAGGCCGCC GTCTTGTGCC CTAAAGGGAG ACAGCCAAAC
    ACCAACATCG TCTCCGGCGG CAGAACACGG GATTTCCCTC TGTCGGTTTTG
201 GCAAAACGCG GAAAATTCTT GGATCAGGGG GATGTAAAGG CGGCGAATTT
    CGTTTTGCGC CTTTTAAGAA CCTAGTCCCC CTACATTTCC GCCGCTTAAA
251 TAACAGAAGG CGGCGGATAA AAAAACGGCC GGAATCATAT TCCAGCCGAT
    ATTGTCTTCC GCCGCCTATT TTTTGTCCGG CCTTAGTATA AGGTCCGGCTA
301 CAAAGTATTC GTAGCTTGCT GCGGGAGCGT AACCCACATC GGCCGCCATT
    GTTTCATAAG CATCGAACGA CGCCCTCGCA TTGGGTGTAG CCGGCGGTAA
351 GTGTCTTTGG GGGCGGGTAG AGGCCAGCAG GTTATCCCTG CCTCAGGGCT
    CACAGAAACC CCCGCCCATC TCCGGTCGTC CAATAGGGAC GGAGTCCCGA
401 CGGAATTTCC GAGTTGCCTC TCAGCCTGGC AGATGGCGAT CGGCGCAGAG
    GCCTTAAAGG CTCAACGGAG AGTCGGACCG TCTACCGCTA GCCGCTCTC
                                     ***GlyProLeu HisArgAsp AlaCysLeu
451 CCGGTCAAAG GTGGCGATGA CACGTTCACT CGTGATCTTG CTCATGGCAT
    GGCCAGTTTC CACCGCTACT GTGCAAGTGA GCACTAGAAC GAGTACCGTA
    ArgAspPheThr AlaIleVal ArgGluSer ThrIleLysSer MetAlaAsp

```

501 CCGGATCCTT GAGGCGGGTG CGCCAGCTCA GCTCGGCCAG CGGCTTGCCG
GGCCTAGGAA CTCCGCCAC GCGGTGAGT CGAGCCGGTC GCCGAACGGC
.ProAspLys LeuArgThrArg TrpSerLeu GluAlaLeu ProLysGlyThr.

551 GTTTCCTGTT CGATCAGTTC CTGATAGACG CTGACCGCAT AGTCGGGCA
CAAAGGACAA GCTAGTCAAG GACTATCTGC GACTGGCGTA TCAGCGCCGT
.GluGlnGlu IleLeuGlu GlnTyrValSer ValAlaTyr AspArgCys

601 CAGATAGGGA CCGGTACGGG CCGGATTGTG ATGGGCGTAG AGGCCGATCA
GTCTATCCCT GGCCATGCCC GGCCTAACAC TACCCGCATC TCCGGCTAGT
LeuTyrProGly ThrArgAla ProAsnHis HisAlaTyrLeu GlyIleVal.

651 CCGGCGTGCC GACCAGGGTT GCCATGTGGG CAGGGCCCGT ATCCGGGGCC
GGCCGCACGG CTGGTCCCAA CGGTACACCC GTCCCGGGCA TAGGCCCCGG
.ProThrGly ValLeuThrAla MetHisAla ProGlyThr AspProAlaLeu.

701 AGCACCAGGC TCGCTTCATC AATCAGGGCC AGCAGCTGCT TGAGGTTGGT
TCGTGGTCCG AGCGAAGTAG TTAGTCCCGG TCGTCGACGA ACTCCAACCA
.ValLeuSer AlaGluAsp IleLeuAlaLeu LeuGlnLys LeuAsnThr

751 CTGCCCCACG AGATCGACCG GTTTACTCTG GCTCAGCCGC TGTATGTCCG
GACGGGGTGC TCTAGCTGGC CAAATGAGAC CGAGTCGGCG ACATACAGCC
GlnGlyValLeu AspValPro LysSerGln SerLeuArgGln IleAspAla.

801 CGGCCAGATC CCGTTCAGC TTGGCCGGGC CGCCGCACAG ATAGACCTGA
GCCGGTCTAG GGCAAGGTCG AACCGGCCCG GCGGCGTGTC TATCTGGACT
.AlaLeuAsp ArgGluLeuLys AlaProGly GlyCysLeu TyrValGlnPhe.

851 AAGCCCTTGT TGGCCGCGTG ATCTGCCAGC GCGGCATAAC CCGCGGCGGT
TTCGGGAACA ACCGGCGCAC TAGACGGTCG CGCCGTATTG GGCGCCGCA
.GlyLysAsn AlaAlaHis AspAlaLeuAla AlaTyrGly AlaAlaThr

901 CCAGTTCTTG AACGCCTTGC TGGCGGCGGC GCAGATGAGC AGGGTCCGTT
GGTCAAGAAC TTGCGGAACG ACCGCCGCCG CGTCTACTCG TCCCAGCCAA
TrpAsnLysPhe AlaLysSer AlaAlaAla CysIleLeuLeu ThrProLys.

951 TGCCGCCAAT CTTCTCCCGG GCCCAGGTTT TATGCTCGGC GCTGATGGGC
ACGGCGGTTA GAAGAGGGCC CGGGTCCAAA ATACGAGCCG CGACTACCCG
.GlyGlyIle LysGluArgAla TrpThrLys HisGluAla SerIleProLeu.

1001 AGTTGCCAGT CCGGGGTCAG ATCCTGGATG CCCAGCTCTT TGGCGAAGGC
TCAACGGTCA GGCCCCAGTC TAGGACCTAC GGGTCGAGAA ACCGCTTCCG
.GlnTrpAsp ProThrLeu AspGlnIleGly LeuGluLys AlaPheAla

1051 GAGAAAGCCA TCCAGCACGT GAGGGGAGGC GGGGGATGGC ACCCGGTGAT
CTCTTTCGGT AGGTCTGTCA CTCCCCTCCG CCCCTACCG TGGGCCACTA
LeuPheGlyAsp LeuValHis ProSerAla ProSerProVal ArgHisAsn.

1101 TGGTGAACAG CCACTGACCG TCGTTGGCGC GCGCCTTGTC GAAACCGAGT
ACCACTTGTC GGTGACTGGC AGCAACCGCG CGCGGAACAG CTTTGGCTCA
.ThrPheLeu TrpGlnGlyAsp AsnAlaArg AlaLysAsp PheGlyLeuLys.

1151 TTGACCCTGG CCTTGATACC GAGCGTCGTT ATGCTGGCGC GCATGGCCGC
 AACTGGGACC GGAACTATGG CTCGCAGCAA TACGACCGCG CGTACCGGCG
 ..ValArgAla LysIleGly LeuThrThrIle SerAlaArg MetAlaAla

1201 CTGCAGATGG AGCAGGGCAT CGAACTTGCG TCCCTTGAGG GTCTGCCAGA
 GACGTCTACC TCGTCCCGTA GCTTGAACGC AGGGAActCC CAGACGGTCT
 GlnLeuHisLeu LeuAlaAsp PheLysArg GlyLysLeuThr GlnTrpLeu.

1251 GGTACAGATA GCCGCGCCAC CCCTTGCTCT TGTCAAAGGC GATGACCTCG
 CCAGTGCTAT CGGCGCGGTG GGGAACGAGA ACAGTTTCCG CTACTGGAGC
 .AspArgTyr GlyArgTrpGly LysSerLys AspPheAla IleValGluVal.

1301 ACCCGGGCA GATCGGCGAA AAGTGTCGCC TCTATTTTGC CGGTGATCCA
 TGGGGCCCGT CTAGCCGCTT TTCACAGCGG AGATAAAACG GCCACTAGGT
 ..GlyProLeu AspAlaPhe LeuThrAlaGlu IleLysGly ThrIleTrp

1351 GGTGATGCGG GTCTGGGGCC AGGCCCGCTG AATGACGCGC ACCAGCGCCA
 CCACTACGCC CAGACCCCGG TCCGGGCGAC TTA CTGCGCG TGGTTCGGGT
 ThrIleArgThr GlnProTrp AlaArgGln IleValArgVal LeuAlaLeu.

1401 GCGCGTGACA ACAATCCCCG ATGGCGGAGA GCCGAGGAT GCAGATGGAA
 CGCGCACTGT TGTTAGGGGC TACCGCCTCT CGGCGTCCTA CGTCTACCTT
 .AlaHisCys CysAspGlyIle AlaSerLeu ArgLeuIle CysIleSerSer.

1451 GAGGGCGGCG TTTGGAACAG CGGCATGATG TGGCGCAATG GTGGCAAAT
 CTCCCGCCGC AAACCTTGTC GCCGTA CTAC ACCGCGTTAC CACCGTTTTA
 ..ProProThr GlnPheLeu ProMet

← **orf1 (waaC)**

1501 GGATGGATGG CATTATGCAA ACTCCCCTGT GGA TTGTAAA AT TGCGC CA
 CCTAC TACC GT AATACGTT TGAGGGGACA CCTA ACATTT TAACGCGCGT

1551 ACTGGCAGCC GGTACTGACT CTGATCCTCC CAAGGAACAC GGTCCGTGCC
 TGACCGTCGG CCATGACTGA GACTAGGAGG GTTCCTTG TG CCAGGCACGG

1601 GACAAGCCGC TTTATCGTGA ATTGAACCAT TGCGATTACT GACCGACCGG
 CTGTTTCGGCG AAATAGCACT TAACTTGGTA ACGCTAATGA CTGGCTGGCC

orf2 (kdkA) →

MetArgIleGln ThrGluHis AsnGlnIle CysTrpTyrAla.

1651 AATCACCAAG ATGCGAATAC AGACCGAACA CAACCAGATC TGCTGGTACG
 TTAGTGGTTC TACGCTTATG TCTGGCTTGT GTTGGTCTAG ACGACCATGC
 .AGluGlyIle PheHisAsp ProSerProGlu LeuPheAsp ProAlaTrp

1701 CCGAAGGGAT TTTTCATGAC CCTTCCCCG AGCTGTTTGA TCCGGCCTGG
 GGCTTCCCTA AAAAGTACTG GGAAGGGGGC TCGACAACT AGGCCGGACC
 TrpGlnThrHis ArgGlnVal ValGlySer SerIleGlyArg GlyValThr.

1751 TGGCAGACCC ATCGCCAGGT AGTGGGTTCC TCCATAGGTC GCGGGGTGAC
 ACCGTCTGGG TAGCGGTCCA TCACCCAAGG AGGTATCCAG CGCCCCACTG

·TrpPheVal LysAspGluSer ArgHisLeu ValLeuArg HisTyrTyrArg·
 1801 CTGGTTCGTG AAGGATGAAT CCCGCCATCT GGTGCTGCGC CATTACTATC
 GACCAAGCAC TTCCTACTTA GGGCGGTAGA CCACGACGCG GTAATGATAG
 ·AGlyGlyMet ValGlyLys ValValArgAsp ArgPheTrp PheGluGly
 1851 GCGGGGGCAT GGTGGGCAAG GTGGTGCGGG ATCGCTTCTG GTTCGAGGGG
 CGCCCCGTA CCACCCGTTT CACCACGCC TAGCGAAGAC CAAGCTCCCC
 ValGluSerSer ArgAlaMet AlaGluTyr SerLeuLeuAla LysLeuSer·
 1901 GTAGAGAGCA GCCGAGCAAT GGCCGAGTAC AGCCTGCTGG CCAAGCTGAG
 CATCTCTCGT CGGCTCGTTA CCGGCTCATG TCGGACGACC GGTTCGACTC
 ·GluGlnGly LeuProValPro ArgProPhe AlaAlaArg MetAlaLysGln·
 1951 CGAGCAGGGC CTGCCGGTGC CGCGCCCGTT TGCGGCCCGC ATGGCCAAGC
 GCTCGTCCCG GACGGCCACG GCGCGGGCAA ACGCCGGGCG TACCGGTTCCG
 ·GGlyProPhe TyrArgAla AspIleLeuIle GluArgIle ArgGlyAla
 2001 AGGGTCCCTT CTATCGGGCC GATATCCTGA TCGAGCGCAT TCGCGGCGCC
 TCCCAGGGAA GATAGCCCGG CTATAGGACT AGCTCGCGTA AGCGCCGCGG
 LysAspLeuVal AlaLeuLeu LysGlnGly ProIleAlaGly GluValTrp·
 2051 AAGGATCTGG TGGCGCTGCT CAAGCAGGGG CCGATCGCTG GCGAGGTGTG
 TTCCTAGACC ACCGCGACGA GTTCGTCCCC GGCTAGCGAC CGCTCCACAC
 ·HisLysVal GlyGlnThrVal ArgGlnLeu HisAspAla GlyValTyrHis·
 2101 GCACAAGGTC GGCCAGACGG TGCGCCAGCT GCACGATGCC GGTGTCTATC
 CGTGTTCAG CCGGTCTGCC ACGCGGTCTGA CGTGCTACGG CCACAGATAG
 ·HAlaAspLeu AsnSerHis AsnLeuLeuLeu AspLysGlu GlyLysVal
 2151 ATGCGGATCT CAACAGCCAC AACCTGCTGC TCGACAAGGA GGGCAAGGTG
 TACGCCTAGA GTTGTCTGGT TTGGACGACG AGCTGTTCTT CCCGTTCCAC
 TrpValIleAsp PheAspLys GlyAlaIle ArgSerProGly SerTrpGln·
 2201 TGGGTGATCG ATTTTCGACAA GGGAGCGATA CGTTCCCCCG GCAGCTGGCA
 ACCCACTAGC TAAAGCTGTT CCCTCGCTAT GCAAGGGGGC CGTCGACCGT
 ·GlnAlaAsn LeuGluArgLeu LeuArgSer PheThrLys GluSerGlnLeu·
 2251 GCAGGCCAAC CTGGAGCGGC TGCTGCGCTC TTTCACCAA GAGTCCCAAT
 CGTCCGGTTG GACCTCGCCG ACGACGCGAG AAAGTGGTTT CTCAGGGTTA
 ·LHisThrSer PheHisPhe ValProAspAsn TrpGlnAla LeuMetGln
 2301 TACATACCAG TTTCCACTTC GTGCCGGACA ACTGGCAGGC CTTGATGCAG
 ATGTATGGTC AAAGGTGAAG CACGGCCTGT TGACCGTCCG GAACTACGTC
 GlyTyrGlnGly LysAlaSer ***
 2351 GGTATCAAG GAAAAGCGAG **CTGA**CCCTGT TGCTCGGCGC GAGAGCCAAT
 CCAATAGTTC CTTTTGCTC GACTGGGACA ACGAGCCGCG CTCTCGGTTA
 2401 GTTACTTATT GAGAAGGTCT TCAGATTCGG TGTCGAGGCG CTCGATGGAC
 CAATGAATAA CTCTTCCAGA AGTCTAAGCC ACAGCTCCGC GAGCTACCTG
 2451 ATGCTAAAGG CTCCCCTCAG GGAGGCAATT ACCCAGTTTG ATCTGGCGCT
 TACGATTTCC GAGGGGAGTC CCTCCGTTAA TGGGTCAAAC TAGACCGCGA

2501 GATGGCGGAT GGCATCGCTT TGCTGGAGGC CGGACTCTGG AGCTG
CTACCGCCTA CCGTAGCGAA ACGACCTCCG GCCTGAGACC TCGAC