

Hepatitis B virus (strain ayw) genome
NCBI Reference Sequence: NC_003977.2

FASTA Graphics

Go to:

LOCUS NC_003977 3182 bp DNA circular VRL 25-MAR-2021

DEFINITION Hepatitis B virus (strain ayw) genome.

ACCESSION NC_003977

VERSION NC_003977.2

DBLINK BioProject: PRJNA485481

KEYWORDS RefSeq; antigen; core antigen; genome; signal peptide.

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus

Viruses; Riboviria; Pararnavirae; Artverviricota; Revtraviricetes;
Blubervirales; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 3182)

AUTHORS Abraham,T.M. and Loeb,D.D.

TITLE The topology of hepatitis B virus pregenomic RNA promotes its replication

JOURNAL J. Virol. 81 (21), 11577-11584 (2007)

PUBMED 17699570

REFERENCE 2 (bases 1 to 3182)

AUTHORS Chen,A., Kao,Y.F. and Brown,C.M.

TITLE Translation of the first upstream ORF in the hepatitis B virus pregenomic RNA modulates translation at the core and polymerase initiation codons

JOURNAL Nucleic Acids Res. 33 (4), 1169-1181 (2005)

PUBMED 15731337

REMARK Publication Status: Online-Only

REFERENCE 3 (bases 1773 to 2478)

AUTHORS Borisova,G.P., Pumpen,P.P., Bychko,V.V., Pushko,P.M., Kalis,Y.V., Dishler,A.V., Gren,E.Y., Tsibinogin,V.V. and Kukain,R.A.

TITLE Structure and expression of the gene of the core antigen of human hepatitis B virus (HBV) in Escherichia coli cells

JOURNAL Dokl. Biochem. 279, 386-390 (1985)

REFERENCE 4 (bases 1 to 3182)

AUTHORS Galibert,F., Mandart,E., Fitoussi,F., Tiollais,P. and Charnay,P.

TITLE Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli

JOURNAL Nature 281 (5733), 646-650 (1979)

PUBMED 399327

REFERENCE 5 (bases 1 to 3182)

CONSRTM NCBI Genome Project

TITLE Direct Submission
 JOURNAL Submitted (20-OCT-2015) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA
 COMMENT VALIDATED REFSEQ: This record has undergone validation or preliminary review. The reference sequence is identical to V01460. On Oct 23, 2015 this sequence version replaced NC_003977.1. COMPLETENESS: full length.
 FEATURES Location/Qualifiers
 source 1..3182
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /strain="ayw"
 /db_xref="taxon:10407"
 gene join(2309..3182,1..1625)
 /gene="P"
 /locus_tag="HBVgp1"
 /db_xref="GeneID:944565"
 CDS join(2309..3182,1..1625)
 /gene="P"
 /locus_tag="HBVgp1"
 /note="reverse transcriptase; DNA polymerase; P protein; P; HP; protein P; polymerase with reverse; transcriptase and RNAses H activit"
 /codon_start=1
 /product="polymerase"
 /protein_id="YP_009173866.1"
 /db_xref="GeneID:944565"

/translation="MPLSYQHFRLLLLDDEAGPLEEELPRLADEGLNRRVAEDLNLG

NLNVSI PWTHKVG NFTGLYSSTVPVFNPHWKTPSFPNIHLHQDIIKKCEQFVGPLTVN

EKRRLQLIMPARFYPKVTKYLPLDKGIKPYYPEHLVNHYFQTRHYLHTLWKAGILYKR

ETTHSASFCGSPYSWEQDLQHGAESFHQQSSGILSRPPVGSSLQSKHRKSRLGLQSQQ

GHLARRQQGRSWSIRAGFHPTARRPFGVEPSGSGHTTNFASKSASCLHQSPVRKAAYP

AVSTFEKHSSSGHAVEFHNLPPNSARSQSERPVPFCWWLQFRNSKPCSDYCLSLIVNL

LEDWGPCA EHGEH HIRPTPSRV TGGVFLVDKNPHNTAESRLV VDFSQFSRGNYRVS

WPKFAV PNLQSLTNLLSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARL

SSNSRILNNQHGTMPDLHDYCSRNLVSLLLLYQTFGRKLHLYSHPIILGFRKIPMGV
GLSPFLLAQFTSAICSVVRRAPHCFAFSYMDDVVLGAKSVQHLESFTAVTNFLLSL
GIHLNPNKTKRWGYSLNFMGYVIGCYGSLPQEHIIQKIKECFRKLPIINRPIDWKVCQR
IVGLLGFAAPFTQCGYPALMPYACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRP
LCQVFADATPTGWGLVMGHQRMRGTFSAPLPIHTAELLAACFARSRSGANIIGTDNSV

VLSRKYTSFPWLLGCAANWILRGTSFVYVPSALNPADDPSRGRGLSRPLLRLPFRPT

	TGRTSLYADSPSVSHLPDRVHFASPLHVAWRPP"
gene	join(2850..3182,1..837) /gene="S" /locus_tag="HBVgp2" /db_xref="GenelD:944569"
CDS	join(2850..3182,1..837) /gene="S" /locus_tag="HBVgp2" /note="long surface protein; PreS1; L protein; LHBs; large S protein; pre-S1/pre-S2/S regions; L glycoprotein; L-HBsAG; LHBs" /codon_start=1 /product="large envelope protein" /protein_id="YP_009173869.1" /db_xref="GenelD:944569"

/translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPD

NKVGAGAFGLGFTPPHGGLLGWSPQAQILQTLPANPPPASTNRQSGRQPTLSPPLR

NTHPQAMQWNSTTFHQLQDPRVRGLYFPAGGSSSGTVNPVLTASPLSSIFSRIGDP

ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS

PTSNHSPTSCPPTCPGYRWMCLRRFIIFLILLCLIFLLVLLDYQGMLPVCPLIPGS

STTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLWEWASARFS

WLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

CDS	join(3174..3182,1..837) /gene="S" /locus_tag="HBVgp2" /note="middle surface protein; PreS2; M protein; MHBs;
-----	---

middle S protein; pre-S2/S"
/codon_start=1
/product="middle envelope protein"
/protein_id="YP_009173870.1"
/db_xref="GenelD:944569"

/translation="MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLTTASPLSS
IFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTV
CLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLIFILLCLIFLLVLLDYQGMLP
VCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLW
EWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFC

CDS
LWVYI"
157..837
/gene="S"
/locus_tag="HBVgp2"
/note="small surface protein; S; HBs antigen; SHBs; S
protein; S glycoprotein; S-HBsAg, SHB, small S protein; S
gene product, putative"
/codon_start=1
/product="small envelope protein"
/protein_id="YP_009173871.1"
/db_xref="GenelD:944569"

/translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGG
TTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLIFILLCLIFLLVLLDYQG
MLPVCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPISSWAFGK
FLWEWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPI

FFCLWVYI"
gene 1376..1840
/gene="X"
/locus_tag="HBVgp3"
/db_xref="GenelD:944566"
CDS 1376..1840
/gene="X"
/locus_tag="HBVgp3"
/note="X protein; HBx; X; HBx; peptide X; pX"
/codon_start=1

/product="X protein"
/protein_id="YP_009173867.1"
/db_xref="GeneID:944566"

/translation="MAARLCCQLDPARDVLCLRPVGAESCGRPFSGSLGTLSSPSPA

VPTDHGAHLSLRGLPVCAFSSAGPCALRFTSARRMETTVNAHQILPKVLHKRTLGLSA

MSTTDLEAYFKDCLFKDWEELGEEIRLKVFLGGCRHKLVCAPAPCNFFTSA"

misc_difference complement(1775)

/note="G is A in ref[2];
conflict"

/citation=[3]

gene 1816..2454

/gene="C"

/locus_tag="HBVgp4"

/db_xref="GeneID:944568"

CDS 1816..2454

/gene="C"

/locus_tag="HBVgp4"

/note="precore; HBe antigen; PreC; HBeAg; precore protein;
external core antigen; HBeAg; p25; precursor"

/codon_start=1

/product="pre-capsid protein"

/protein_id="YP_009173857.1"

/db_xref="GOA:P03146"

/db_xref="InterPro:IPR002006"

/db_xref="UniProtKB/Swiss-Prot:P03146"

/db_xref="GeneID:944568"

/translation="MQLFHLCLIISCSCPTVQASKLCLGWLWGMIDIDPYKEFGATVEL

LSFLPSDFFPSVRDLLDTASALYREALSPEHCSPHHTALRQAILCWGELMTLATWVG

VNLEDPASRDLVVS YVNTNMGLKFRQLLWFHISCLTFGRET VIEYLV SFGVWIRT PPA

YRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC"

sig_peptide 1816..1902

/gene="C"

/locus_tag="HBVgp4"

/note="signal peptide (aa -29 to -1)"

mRNA join(1820..3182,1..1932)

/product="pgRNA"

/note="pre-genomic RNA; encodes core and polymerase;

```

                                template for reverse transcription"
                                /citation=[1]
CDS      1903..2454
                                /gene="C"
                                /locus_tag="HBVgp4"
                                /note="core; capsid; HBc antigen; C; core protein; core
                                antigen; HBcAg; p21.5; C protein"
                                /codon_start=1
                                /product="capsid protein"
                                /protein_id="YP_009173868.3"
                                /db_xref="GenelD:944568"

/translation="MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALES

PEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPASRDLVVSYNVTNMGLKFRQLLW

FHISCLTFGRETVIEYLVSFQVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRT
                                PSPRRRRSQSPRRRRSQSRESQC"
misc_difference complement(1910)
                                /note="G is A in ref[2];
                                conflict"
                                /citation=[3]
misc_difference complement(1913)
                                /note="G is A in ref[2];
                                conflict"
                                /citation=[3]
misc_difference complement(1998..2000)
                                /note="GGT(Thr) is GTT(Asn) in ref[2];
                                conflict"
                                /citation=[3]
misc_difference complement(2078)
                                /note="A is T in ref[2];
                                conflict"
                                /citation=[3]
misc_difference complement(2081)
                                /note="A is C in ref[2];
                                conflict"
                                /citation=[3]
misc_difference complement(2123)
                                /note="AAC(Val) is ACC(Gly) in ref[2];
                                conflict"
                                /citation=[3]
misc_difference complement(2139..2141)
                                /note="CGC(Ala) is TAT(Ile) in ref[2];

```

```

conflict"
/citation=[3]
misc_difference complement(2144)
/note="A is G in ref[2];
conflict"
/citation=[3]
misc_difference complement(2147)
/note="T is C in ref[2];
conflict"
/citation=[3]
misc_difference complement(2188)
/note="C is T in ref[2];
conflict"
/citation=[3]
misc_difference complement(2201)
/note="G is T in ref[2];
conflict"
/citation=[3]
misc_difference complement(2267)
/note="G is A in ref[2];
conflict"
/citation=[3]
misc_difference complement(2319)
/note="G is A in ref[2];
conflict"
/citation=[3]

```

ORIGIN

```

1 aattccacaa cctccacca aactctgcaa gatcccagag tgagaggcct gtatttcct
61 gctggtggct ccagttcagg aacagtaaac cctgttctga ctactgcctc tcccttatcg
121 tcaatcttct cgaggattgg ggaccctgcg ctgaacatgg agaacatcac atcaggattc
181 ctaggacccc ttctcgtgtt acaggcgggg ttttcttgt tgacaagaat cctcacaata
241 ccgcagagtc tagactcgtg gtggacttct ctcaatttct tagggggaac tacctgtgtg
301 cttggccaaa attcgcagtc cccaacctcc aatcactcac caacctcttg tcctccaact
361 tgcctcgtgt atcgctggat gtgtctgcgg cgttttatca tcttctctt catcctgctg
421 ctatgcctca tcttctgtt ggttctctg gactatcaag gtatgttgcc cgtttgcct
481 ctaattccag gatcctcaac aaccagcacg ggacctgcc ggacctgcat gactactgct
541 caaggaacct ctatgtatcc ctctgttgc tgtaccaaac ctccggacgg aaattgcacc
601 tgtattccca tcccatcatc ctgggcttcc ggaaaattcc tatgggagtg ggcctcagcc
661 cgtttctcct ggctcagttt actagtcca tttgtcagt ggttcgtagg gctttccccc
721 actgtttggc ttccagttat atggatgatg tggattggg ggccaagtct gtacagcatc
781 ttgagtcctt tttaccgct gttaccaatt ttctttgtc ttgggtata catttaaacc
841 ctaacaaaac aaagagatgg ggttactctc taaattttat gggttatgtc attggatgtt
901 atgggtcctt gccacaagaa cacatcatatc aaaaaatcaa agaattttt agaaaacttc
961 ctattaacag gcctattgat tggaaagtat gtcaacgaat tgtgggtctt ttgggttttg

```

1021 ctgccccttt tacacaatgt gggtatcctg cgttgatgcc ttgtatgca tgtattcaat
 1081 ctaagcaggc ttctactttc tcgccaactt acaaggcctt tctgtgtaaa caatacctga
 1141 acctttaccc cggtgcccgg caacggccag gtctgtgcca agtggttgct gacgcaaccc
 1201 ccactggctg gggcttggtc atggggccatc agcgcatacg tggaaccttt tcggctcctc
 1261 tgccgatcca tactgcggaa ctctagccg cttgttttgc tcgcagcagg tctggagcaa
 1321 acattatcgg gactgataac tctgttgcc tatcccgcaa atatacatcg ttccatggc
 1381 tgctaggctg tgctgccaac tggatcctgc gcgggacgtc cttgttttac gtcccgtcg
 1441 cgctgaatcc tgccgacgac cttctcggg gtcgcttggg actctctcgt ccccttctc
 1501 gtctgccgtt ccgaccgacc acggggcgca cctctcttta cgccgactcc ccgtctgtgc
 1561 cttctcatct gccggaccgt gtgcacttcg cttcacctct gcacgtcgca tggagaccac
 1621 cgtgaacgcc caccaaatac tgccaaggt cttacataag aggactcttg gactctcage
 1681 aatgtcaacg accgaccttg aggcatactt caaagactgt ttgtttaaag actgggagga
 1741 gttgggggag gagattaggt taaaggtctt tgtactagga ggctgtaggc ataaattgg
 1801 ctgcgcacca gcaccatgca actttttcac ctctgcctaa tcactctctg ttcattgtct
 1861 actgttcaag cctccaagct gtgccttggg tggcttggg gcatggacat cgacccttat
 1921 aaagaatttg gagctactgt ggagtactc tcgttttgc cttctgactt ctttcttca
 1981 gtacgagatc tttagatac cgctcagct ctgtatcggg aagccttaga gtctcctgag
 2041 cattgttcac ctaccatac tgcactcagg caagcaattc ttgtctgggg ggaactaatg
 2101 actctagcta cctgggtggg tgtaatttg gaagatccag cgtctagaga cctagtagtc
 2161 agttatgtca acactaatat gggcctaaag ttcaggcaac tctgtggtt tcacatttct
 2221 tgtctcactt ttggaagaga aacagttata gagtatttgg tgtctttcgg agtggtggatt
 2281 cgcactctc cagcttatag accaccaaact gccctatcc tatcaacact tccggagact
 2341 actgttgta gacgacgagg caggtcccct agaagaagaa ctccctcgcc tcgcagacga
 2401 aggtctcaat cgccgcgtcg cagaagatct caatctcggg aatctcaatg ttagtattcc
 2461 ttggactcat aaggtgggga acttactgg gctttattct tctactgtac ctgtctttaa
 2521 tcctcatttg aaaacaccat ctttctctaa tatacattha caccaagaca ttatcaaaaa
 2581 atgtgaacag ttgtaggcc cactcacagt taatgagaaa agaagattgc aattgattat
 2641 gcctgccagg tttatccaa aggttacaa atatttacca ttggataagg gtattaaacc
 2701 ttattatcca gaacatctag ttaatcatta ctccaaact agacactatt tacacactct
 2761 atggaaggcg ggtatattat ataagagaga aacaacacat agcgcctcat ttgtgggtc
 2821 accatattct tgggaacaag atctacagca tggggcagaa tctttccacc agcaatcctc
 2881 tgggattctt tcccgaccac cagttggatc cagccttcag agcaaacacc gcaatccag
 2941 attgggactt caatcccaac aaggacacct ggccagacgc caacaaggta ggagctggag
 3001 cattcgggct gggtttcacc ccaccgcacg gaggcctttt ggggtggagc cctcaggctc
 3061 agggcatact acaaactttg ccagcaaata cgctcctgc ctccaccaat cgccagtcag
 3121 gaaggcagcc taccgcgtg tctccacctt tgagaaacac tcactctcag gccatgcagt
 3181 gg

//