

Virus Report

20-01-2015 15:28:06

Secondary structure results

Virus name:
HI-V

Sequence:

```
TGGTATCAGAGCAAGGTTAAACATGGTCATGTCCGGCTAAAAACCTAGTGTTAGGTTTCGG
GGAAAGAGGTCTACTGTGTGTTTCAAATATCCATTTGTTATTATCCAAACAGTGTCTGTTT
CTTGTGTTTAGGATGGAAACCTGAGGTTGCGAACTACGTGTTCAAAGGAGTGATCATTTT
AAAAGGAATGTCATTCTAACTATCGTATGCTCTAGACTTGCCCTGTTATAGCAACAGAGGA
TCCTGTCTATCAGAAAATAGTTTGAACAATTCCGATGATCTCTCATACTTTAGGTATAACAA
CGCCACGTATAAAGGAGCTAATTCTTTTGTAAGGACGTAACCCAGAAGAGAGTACCCAGA
GAGGGGAAGACTTGGGGAGAAGGTGAACAGAAGAGAGAGCTTACTGTACCTGGTACTGT
TATAGGAAGATTTTGTATGTCCAGTCGCTGGGAGAATAGTATACAAGAGTGGTATGAGA
AGTCACACACAGCAAACCTTGAGTACCTTGACCTAGCATCTACTAGTAAAGTGACCAACA
ACCAGTTAGCACATAACCTTGACGTAACCTTTGATAGAGTAAATTTAGGTAACCGAGTTTT
TATTAATAAACCTGAAGCAGATTCAAGAGTCTATTCTTGAATTAACACCCCGTGTGATACTA
TAGAAGTAGCTCTAAGGAGGTTAACCAAGACCTTCAGAGAAAGCAAACCACTTACCGAAA
GTGAAGTCCGTAAGTTGGTTGAAGAAATTGCTCAACAACCTAAGATCGTCGAGAAACAAG
CACTCGAGATCTCTCAACAACCTTGAGCAGAAACTAGAGAAGGTGGAAAAGCTCTTGACACA
AACTTGACCAGTGGGTTGGTCAATGACTGATAGCCCAGCCTATCAAGAAGCTTTAAAGGA
AGCAGAAAAGGTTGACCCACCAGCAGTTGGGATAACAACCTCCACCGGAGTAACAGCGG
TTCAAGGGTTTTAAACCGTTATTAAGCAGAACAACGTTGAGATCTGCTTACTCGCTGTCAT
AGCAGATAAACTAGAAGAGCTGATACAGGATCAGAAGAAGGCCAGGAAGGATAAGGCAA
AGGAAGTTGCCATTCCAGAAGACTTAATCACAAAACCTCCAAGGATTATCTATTCAGGAGAA
AGGAGAGGCCAAAAGTCACAAGAAAACCCGAGCCAAAAGGAACGCTGTTTGGATTCAAAG
ATCCTTACAAGATCCTAGCAGCAGAAAAGGCTAAGATCACACTTAAGCCTGTGAAAGAGT
AGAAAGATGAGTCGAGCACGACCGCAACACCCAGTTCCTAGTGTGACTACAACCACTAGT
GAGCAAAAACAGGGAAGGACCTCTATACGAGGATCAGATCAGAGACTACCGAAGAGGACA
AAGAAGGATCTTTAACCTTCGAAGAAGAGCCAGAAGGTTAAGAAGATCAATGATGGGGTC
TAGATATCAAGAGACCCTAGAACAAGAAATAGATCCACAGACAACACTGAGGTTGTCCAT
GCAAGAACGAGCAAGACTAGTACCAGCTGAGGTACTGTACAGATCCAGGCGAGACACTG
TTCATCACAGAGTCTACACACACCGATCTGAAGAGTCAGTCCTATGTGTTGGCGGAAGCC
AAGTGGATAGGGCGTTTTATACAGCCTGAAAGTCTGGAACAACCTCCAAAGGACAGGAATGT
CCTTTATTCATATCGGAATTCTGCAAGTGAGGATTCAGATTCTGCACCGACAAGAAGAAG
GCACCATGGCCTTGGTAGTATTCAGGGATAACAGATGGTTCGGGAGATCAGTCTATTTTCG
CTCAAATGGAAATAGACCTAACAAAAGGAAGCCAATTGGTGTTCGTTATACCAGACACCAT
GATGACCATCGGAGACTTTGCCCGGAACGTACAACCTATCCATCCTCACAAGAGGGGTATGA
AAATTGGCAGAACGGGAGAAGCCAATCTCCTGATAACACGTGGCATGACAGGGGCGACTGT
CCAACACTCCCAATGTCGCCTTTGCTTACCAAATTGCCAGCGCAACAGATTACTTGGCAA
GTCACGGTGTAAGGCTATCGCAGGAAAGAAAATGAATCTGCAACACCTGCGAAACCAAC
AGTGGATACTACGTCCACCACAAACGGACATCACTCCAATGCAACCAAGATCGGTTGAAA
CAAGAAATCTCGTAGATGGAAGTATCTCCATCAGATTCCATGATTATGAGGCAGCTACTTC
AGCTTCAAGACCTCACTACAATGAGGAAGATGAAGAAGTGGAATCAGAAACGGAGTCAGA
AATAAGGGGAGCATACTATTGCAGTCTGGATAGGAGAAGAAGAAATTCAGACCAGACAGG
GAGAAAGAAGGTATGGGAAGAATCTAGTAATGGAAATGGAAGATTCTTCCGGTATTACAC
TCCTCCACCAACATTTGAAGGGGCAAATCATTGCTACAGGATGGGGAAAGTGATGATGATAA
TGAAAAAACTCCTCCAAAATGGGATGAAAGCCCAGATGAAGAAGGACCCACAGAACCCAT
ATGGGATCAAGAAGAAGAAGAAGATGAATATGATCCCAATGTCTATAGGGGCATACTTACA
AAAGGAGGAAGATGAGTGGCAAGAAATCACAGCTAGTCTCAGGGAAGAAATGGAGTACC
CAAAAAGACGACCACAAACAGAGATGGCGTTCTCTGAAACAGTCGACTATACTCCACCTG
GTGACACTATGATGACACCTGTCGGATACCCCCCGGCCTCGTCATCAAGATCAACAGTCA
CAACGCCAAGTAGACCGCCACTTTTTGAAGGAAGGACCACACACGTGCCACGATTCTTAA
AAAGGGATGAGTACACAGAATGGTGGCAATTACCATCATCGCAAGGCACAACAGGGGGCG
TTATTTGTAATGCCTAAACAAATGGGCCTATTCCATGAAGTCTTCTCGAGATGGGAATCCA
TCACCAAAAACCTACGTTGCGGCCCAAGGTTTTACGGACCCAACAGAAAAGATGGAGTTCA
TGGAATCTGTTAGGAGAAACAGAAAAGCTAACCTGGATCCAATGGAGAATGAATTACG
AGGCTGAATACCAGCAGCTGCTAACCCAAGCAGATGGGCGACAAGGAACCCAAAATATC
TTGTCCCAGATCAAGAGAATCTTTTCTTTAGAAGACCCAGCCTCTGGATCCACGAGAATAC
AAGATGCTGCATACAGAGACCTGGAGAGATTAACTGCCACAACATAAAGACATAGTTC
```

AATTTCTAAATGACTATGGAAGGCTTGCAGCAAAGAGTGGGCGACTGTTCTTAGGAACTG
AGCTCAGTGAAAAGTTATGGATGAAGATGCCACCAGAAGTGGGCATCGAATGAAAGAAG
CATTTCAAAAGGAGTATTCAGGCAATGAAGTAGGAGTCTTCCCGCGCATCTTGTTTGCCT
ACAGATATCTGGAGCAAGAGTGTAAAGATGCAGCTTTTAAAGAGAAGCCTGAAGTCATTAA
GCTTCTGCAAAGACATGCCATTAACAGGATACTATGATAAAACCTCCAAGTATGGCATGA
GAAAGTCAAGAACCCTATAAAGGAAAACCACATGCATCACATGCAAGGGTGGAGAAGAGG
AAGCACTTGATCAGAAATAAAAAGTGCAAATGTTATCTCTGCGGGGATGAAGGACACTTT
GCCCCGGAATGCCCCAATCAGAGGAGAGATGTCAAGAGAGTAGCCATTTTTGAAGGAAT
TGATCTTCCTGAAGGCTTTGATATCGTCTCAGTAGAAGAAGGAGAAGAAGAGTCAGATGC
TATTTATAGCATATCTGAAAATGAAGACGGAGAAGTGGACACTGAAGTAGTCCATGAGAAA
GTCTTCATGATGAGAGAAGAAGACCAGTCCTATTGGTTAGGAAAAGACAAACCATTGGACG
GCTATGGTGCGAGTCAGCAGTCAGCAATATCACTGCATGCACCAATGGGAGCACAAATAAA
GAAATATTAGTGGTGGCCACATCAACTGCCACTTCTGTAAGCAACCCACTCAGTTAAGG
AGTCGAATACACTGTCCACAGTGTCAACTCACCAGTTGCTTCATGTGTGCCCCAATATACT
GCAATATGATAGTCCAGCAGCAGCCTAAACCACCAGTACCGTTTAATACTCACACACTGC
TCCAACAACAAGCGGCTTATATCCAGTGGTTGGAAAAAGAAAATCAGCGGTTAACTGAAG
CCGTTGAATTCTATAAAAAGGAGGCTGAAGAATTAAGGCTCGAAAGAGACTTAGAGCAAG
ATAGAAGGAGTCTGGAACCTACGTTGTTAGACAAAGGAAAGAAGGTTTCAGATTCTTGATC
CAGATGAAGATCAGCACACAGCGTATCTTGAAGAAGATACCATCAGCCGTGTTATCGGCC
ATACTGTGGAACAACAAGAGGTTAGAAAGCCCGTTAAAAAGGGAAACATGCTCTATAACC
TCGATGTGGTGTACATATCCAGAGGTTAGGAAGACCTATCAAAGTCAAAGCAATTCTAG
ACACTGGAGCAACCACATGCTGTATAAACATCAACTCTGTACCACAAACAGCAATTGAACA
GAACACTTTTTCTGGTACAATTCCGAGGCATAAATTCCACGCAATCTGTGGATAAGAAACTC
AAATATGGGAGGATGACTATCAGCAATCACCAGTTCAGAATCCCGTACTGTTATGCCTTTC
CTCTATCCCTTGGTGATGGAATAGAGATGATCCTAGGGTGTAATTTTCATCCGTGGGATGT
ATGGCGGTTTGCCTATTGAAGGTACACACAATCACCTTCTACAAAAATGTCACTACAATCCA
AACCCGCTTGTGCTGCCGTAATGGTTGGTGGTACAACCGCTTCTGAGTTAGGGGGAGGGG
AGGAGTCCAAATCCGATTCTGAATCCATGTTTGACCTCTCAGAAACAGAAGAATTTGACTC
AGAAACCCACCAGCAGATTGTGAGTCATGTTGCAGCCCAAGCCCAACAACAAAAATTGGA
TCCAAAACCTCCAACAACATAATGGTCCAACCTTCAGGATCAGGGCTTTATTGGGGAAAATCCT
ATGCAACATTGGGCTAAAAACAAGATCCTATGTGCTAGATATCAAGAATCCTGATCTTA
TAATAGAAGACAAGCCCATCAAGCATCTAACACCGGCCATGGAGAAACAGTTCCAGAAGC
ACATCAAAGCACTCCTGGACATTGGTGTATCAGGCCTAGTAAGTCAAAACACAGGACTA
CGGCCTTCATTGTGAATCAGGCACTGTTATTGATCCAGTCACGAAAAAGACTATACACG
GTAAAGAACGTTTGGTCTTCAACTACAAACGCCTGAACGACAACACTGAAAAGGACCAGT
ACTCGCTACCCGGTATACAGACCATCCTAAAGCGGGTGGGCAACAAGAAGGTCTTCAGC
AAGTTTCGATCTAAAATCGGGCTTCCATCAAGTTGCCATGGCGGAAGAGTCCATCCCTTGG
ACTGCTTTCTGGGTACCGCAGGGCCTTTATGAATGGTTAGTGATGCCCTTTGGGCTCAAA
AACGCTCCTGCAGTATTTCAAAGAAAAATGGACCAATGCTTCAAAGGTACAGAAGAATTCA
TTGCAGTGTACATTGATGACATTTTGGTCTTCAGCGAAAATATGGCAGAGCATACCAAGCA
CATTGGAATCATGCTCAAGATCTGCCAAGAAAATGGGCTGGTACTAAGCCCAAGTAAGAT
ATGTCTTGCCCAACGCGAGATTGAATTTTTGGGCACAGTCATCTCACAAGGACAAATGAA
GCTTCAGGCCCCATGTAATCAAAAAGATAGTCAACAAAGCCAACATAGAGCTAGAAACAAC
AAAAGGCCTGAGATCCTTTTTGGGCCTCCTCAACTATGCCCGTATCTACATACCCAACCT
GGGTAGAAAGCTAAGTCCACTATATGCCAAGACTAGTCCCACCGGAGAGAAAAGGTTTAA
TCGACAAGATTGGCATCTGATAAAGGAAATCAAAGATATGGTCCAAAAGCTCCCAAACCT
CGCTATCCCACCAGCAAGATGCTACATTATCATTGAAAGCGATGGCTGCATGGAAGGATG
GGGGGCCGTATGTAAGTGGAATTAGCAAAAGAAGATTCCCGCACTACTGAAAAGATTG
TGCGTACGCTAGTGGGAAATTGCGGTGTTGTCAAGTCTACCATCGACGCCGAAATTTACGC
ACTCATAAAAGCATTGGAATCTTTTAAAATCTTCTATCTGGACAAAAAACATTTGGTGGTGC
GAACAGACTGTCAAGCGATAGTGACATTTTACAACAAGACAAGTACTCATAAACCCAGTA
GAATACGTTGGATCACCTTTTCCGACTATATAACGGGGTTAGGAGTTCCGGTTACTATCG
AACACATAGACGGAAAAGAGAACCAGTTGGCTGATACTAAGCAGACTCGTGTATACAA
CATGGAACCAGTCCCAAACCTACCAACCGGAGGAAGAAGAGCTGGAGAAGTCCCAACAT
CTCAGCTTCGCGGGGCTAGCTATCCCTATAGCTTGGCCTATGATGGGCTCCTACAACAAA
AGAAGGACGCCATTACTCACGGGACAATCACTCTGGCAACGGAACAAGCCATCTCAGCA
CAGCTCTACCGCATCGAAGAGCAGGCAGCCAAGAAAGCACTATTGGCCCTACGTGACCT
ACAGGGCGTACTCCACTTCAAGAGAGACTATTTGGCCGCTACTGCCACTAGAGACAACCTG
GGCTAGCGACAGACTGCCAGCTGCCCAACAAGACTCAGCCGCCCTAGACCAACATGCTG
GCGTGATTAACGCCATTATTGAAAGGGCTGTCCAACCCTAGTTTGGACGGTAGTAGTAGG
TGTAATAATAATAGTTAGGTGTGCTTTACTTTTCCAAGCTGTCACTCATTATAGAGTAGACA
TGATGATCGACGATGGGGCCAATGAGCACCCGGATCACCATTCTCCCATCTATAAATGA
GAGTTTGTAAGGCTTAGCCATCAGAGAGTGAAAACCTCAACTGATCCTAAGTGTTAGA
GTTTGTATTTTCTAAGAGTTTGTAAAGATTTTATGAAATAAAGAGTCTACTTTGTGTTTATC

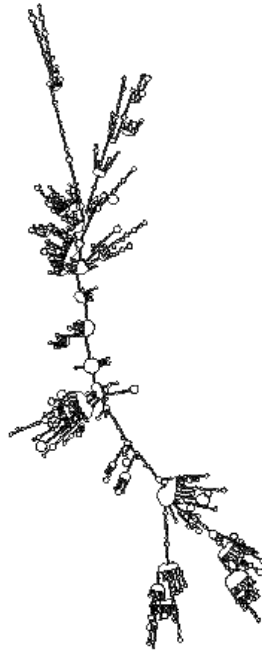
TCTTTGTTTCACCTGGGATTTAAACAGTTTTTGTTCCTCGCACCATCGGTTTGCGCCCGA
TCGATGT

Structure in dot-bracket format:

[illegible]

Energy:
-2056.9

RNA-Structure:



Virus Domain(s)

Number of proteins:
3

Protein 1

Number of domains:
1

Aminoacidsequence:
VTNNQLAHNLA VTFDRVNLGNRVFIKNLKQIQESILELNTRVDTIEVALRRLTKTFRESKPLTES
EVRKLVEEIAQQPKIVEKQALEISQQLEQKLEKVEKLLHKLDQWVGQ

Starting nucleotide position:
533

Ending nucleotide position:
872



Domain 1

Starting aminoacid position:
1

Ending aminoacid position:
99

Domain description:
Protein of unknown function (DUF1319)

Identifier:
DUF1319

Protein 2

Number of domains:
0

Aminoacidsequence:

MTDSPAYQEALKEAEKVDPPAVGITTSTGVTAVQGFKTVIKQNNVQICLLAVIADKLEELIQDQ
KKARKDKAKEVAIPEDLITKLQGLSIQEKGEAKVTRKPEPKGTLFGFKDPYKILAAEKAKITLKP
VKE

Starting nucleotide position:

868

Ending nucleotide position:

1267



Protein 3

Number of domains:

4

Aminoacidsequence:

MSRARPPQHPVPSVTTTTSEQNREGPLYEDQIRDYRRGQRRIFNLRRRARRLRRSMMGSRYQ
ETLEQEIDPQTTLRLSMQERARLVPAEVLYRSRRDTVHHRVYTHRSEESVLCVGGGSQVDRAF
IQPESLEQLQRTGMSFIHIGILQVRIQILHRQEEGTALVVFRDNRWSGDQSIFAQMEIDLTGK
SQLVFVIPDTMMTIGDFARNVQLSILTRGYENWQNGEANLLITRGMTGRLSNTPNVAFAYQIA
SATDYLASHGVKAIAGKKMNLQHLRNQQWILRPPQTDITPMQPRSVETRNLVDGSISIRFHDY
EAATSASRPHYNEEDEEVESETESEIREHTIAVWIGEEEEIPDQTGRKKVWEESSNGNGRFFRY
YTPPPTFEGQIIATGWGSDDDNEKTPPKWDESPDEEGPTEPIWDQEEEEDEYDPNVYRAYL
QKEEDEWQEITASLREEMEYPKRRPQTEMAFSETVDYTPPGDTMMTPVGYPASSSRSTVT
TPSRPPLFEGRTTHVPRFLKRDEYTEWWQLPSSQGTGALFVMPKQMGLFHEVFSRWESIT
KNYVAAQGFTDPTEKMEFMENLLGETEKLTIQWRMNYEAEYQQLLTQADGRQGTQNILSQI
KRIFSLEDPASGSTRIQDAAYRDLERLTCHNIKDIVQFLNDYGRLAAKSGRLFLGTELSEKLWM
KMPPPELGHRMKEAFQKEYSGNEVGVPFRILFAYRYLEQECKDAAFKRSLSLSFCKDMPLTG
YYDKTSKYGMRKSRTYKKGPHASHARVEKRKHLIRNKKCKCYLCGDEGHFARECPNQRRDV
KRVAIFEGIDLPEGFDIVSVEEGEEESDAIYSISENEDGELDTEVVHEKVFMREEDQSYWLG
KTNHWTAMVRVSSQQYHCMHQWEHNKEILVVAHINCHFCQPTQLRSRIHCPTCQLTSCFM
CAPIYCNMIVQQQPKPPVPFNTHLLQQQAAYIQWLEKENQRLTEAVEFYKKEAEELRLERDL
EQDRRSLEPTLLDKGKKVQILDPEDEDQHTAYLEEDTISRVIIGHTVEQQEVRKPKVKKGNMLYNL
DVVLHIPEVGRPIKVKAILDTGATTCCININSVPQTAIEQNTFLVQFRGINSTQSVDKKLKYGRM
TISNHQFRIPYCYAFPLSLGDGIEMILGCNFIRGMYGGLRIEGHTITFYKNVTTIQTRLAAMVG
GTTASELGGGEESKSDSESMFDLSETEEFDSETHQQIVSHVAAQAQQQKLDPKLQQLMVQL
QDQGFIGENPMQHWAKNKILCRLDIKNPDLIEDKPIKHLTPAMEKQFQKHIALLDIGVIRPSK
SKHRTTAFIVESGTVIDPVTKKTIHGKERLVFNKYRLNDNTEKDQYSLPGIQTILKRVGNKKVFS
KFDLKSGFHQVAMAEESIPWTAFWVPQGLYEWLVMPFGLKNAPAVFQRKMDQCFKGTEEFI
AVYIDDILVFSENMAEHTKHIGIMLKICQENGLVLSPSKICLAQREIEFLGTVISQGMKLQAHVI
KKIVNKANIELETTKGLRSFLGLLNYARIYIPNLGRKLSPLYAKTSPTGEKRFNRQDWHLIKEIKD
MVQKLPNLAIPPARCYIIIESDGCMEGWGAVCKWKLAKEDSRTTEKICAYASGKFGVVKSTIDA
EIYALIKALESFKIFYLDKKHLVVRTDCQAIVTFYNKTSTHKPSRIRWITFSYITGLGVPVTIEHI
DGKENQLADTL SRLVYTTWNQSQTHQPEEEEELEKSQHLSFAGLAIPAWPMMGSYNKRRTPL
LTGQSLWQRNKPSQHSSTASKSRQPRKHYPYVVTYRAYSTSRETIWPLLPLETTGLATDCQL
PNKTQPP

Starting nucleotide position:

1271

Ending nucleotide position:

6776



Domain 1

Starting aminoacid position:
86

Ending aminoacid position:
241

Domain description:
Viral movement protein (MP)

Identifier:
MP

Domain 2

Starting aminoacid position:
786

Ending aminoacid position:
803

Domain description:
Zinc knuckle

Identifier:
zf-CCHC

Domain 3

Starting aminoacid position:
1046

Ending aminoacid position:
1163

Domain description:
Retroviral aspartyl protease

Identifier:
RVP

Domain 4

Starting aminoacid position:
1333

Ending aminoacid position:
1494

Domain description:
Reverse transcriptase (RNA-dependent DNA polymerase)

Identifier:
RVT_1