

AVIDD ASAP: NS2B-GSG-NS3 fusions for assays

PAGE22-01614

Author: **Fairhead, Michael**

Date Started: **2022-Sep-22**

Experiment Started:

Projects: **Cloning;ASAP**

Related Pages:

Referenced by: **PAGE22-01756**

Tags:

Title missing - double click to edit

>U87411.1 Dengue virus type 2 (strain 16681) polyprotein mRNA, complete cds

>AF404756.1 West Nile virus isolate WN NY 2000-crow3356, complete genome

>MH158237.1 Zika virus isolate PRVABC59, complete genome

>DV2-NS2BgsgNS3

ADLELERAADVWEDQAEISGSSPILSITISEDGSMISIKNEEEEQTLGGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYS
QIGAGVYKEGTFHTMWHTVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIG
AVSLDFSPGTSGPSIIDKKGKVVGLYGNGVVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

>DV2-NS2BgsgNS3-DEAD

ADLELERAADVWEDQAEISGSSPILSITISEDGSMISIKNEEEEQTLGGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYS
QIGAGVYKEGTFHTMWAVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIG
AVSLDFSPGTSGPSIIDKKGKVVGLYGNGVVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

>WNV-NS2BgsgNS3

STDMWIERTADISWESDAEITGSSSERVDVRLDDDGNGFQLMNDPGAPWKGGGGSGGGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSY
QAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGVSVEQLRCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIG
AVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMDEPI PAGFEPEMLRKK

>WNV-NS2BgsgNS3-DEAD

STDMWIERTADISWESDAEITGSSSERVDVRLDDDGNGFQLMNDPGAPWKGGGGSGGGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSY
QAGAGVMVEGVFHTLWATTTKGAALMSGEGRLDPYWGVSVEQLRCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIG
AVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMDEPI PAGFEPEMLRKK

>ZV-NS2BgsgNS3

SGKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMREGGGGSGGGGSGALWDVPAPKEVKKGETTDGVYRVMTR
RLLGSTQVGVMQEGVFHTMWHTTKGSALRSGEGRDPYWGVDVKQLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKT
KGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE

>ZV-NS2BgsgNS3-DEAD

SGKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMREGGGGSGGGGSGALWDVPAPKEVKKGETTDGVYRVMTR
RLLGSTQVGVMQEGVFHTMWAVTKGSALRSGEGRDPYWGVDVKQLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKT
KGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE

For direct golden gate cloning of TWIST gene add to ends

AAAAAAGGTCTCACATG

TARGET

TAATGACTCGAGAGACCAAAAAA

ALIGNMENTS to guesstimate boundary constructs

AF404756.1 STDMWIERTADISWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWK
2GGV MSTDMWIERTADISWESDAEITGSSERVDVRLDDDGNFQLMNDPGA
2YOL SMTDMWIERTADIWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPW
2IJO MSTDMWIERTADISWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPW
2FP7 SMTDMWIERTADIWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWK

AF404756.1
2GGV GGGGSGGG
2YOL AGGGGSGGGGG
2IJO AGGGGSGGGGG
2FP7 GGGGSGGGGG

AF404756.1 GVLWDTSPSPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKE
2GGV RRGVLWDTSPSPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKE
2YOL VLWDTSPSPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKE
2IJO GVLWDTSPSPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKE
2FP7 DTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKE

AF404756.1 DRLCYGGPWKLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGL
2GGV DRLCYGGPWKLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGL
2YOL DRLCYGGPWKLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGL
2IJO DRLCYGGPWKLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGL
2FP7 DRLCYGGPWKLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGL

AF404756.1 YGNGVIMPNGSYISAIVQGERMDEPIPAGFEPEMLRKK
2GGV YGNGVIMPNGSYISAIVQGERMDEPIPAGKGHHHHHH
2YOL YGNGVIMPNGSYISAIVQGEP
2IJO YGNGVIMPNGSYISAIVQGERMDEPIPAGFEPEMLKGHHHHHH
2FP7 YGNGVIMPNGSYISAIVQGERMEEPAGFEPEMLRKK

U87411.1 ADLELERAADVKWEDQAEISGSSPILSITISEDGSMSEIKNEEEEQTL

4M9T SMADLELERAADVWEQAEISGSSPILSITISEDGSMSEIKNEEEEQTL

U87411.1

4M9T GGGGSGGGG

U87411.1 AGVLWDVPSPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGG

4M9T AGVLWDVPSPPPGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGG

U87411.1 WKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKNAGTIGAVSLDFSPGTSGSPIIDKKGKVVGlyGNGVVTRSGAYVSA

4M9T WKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKNAGTIGAVSLDFSPGTSGSPIIDKKGKVVGlyGNGVVTRSGAYVSA

U87411.1 IAQTEKSIEDNPEIEDDIFRK

4M9T IAQTEKSIEDNPEIEDDIFRK

MH158237.1 SGKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMRE

5TFN GSSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMA

5T1V GKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMREIIL

5GPI MTGKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMRE

MH158237.1

5TFN GGGGSGGGG

5T1V GGGGSGGGG

5GPI

MH158237.1 SGALWDVPAPKEVKKGETTDGVYRVMTRRLGSTQVGVGMQEGVFHTMWHVTKGSALRSGEGRDPYWGDKQDLVS

5TFN GSGALWDVPAPKEVKKGETTDGVYRVMTRRLGSTQVGVGMQEGVFHTMWHVTKGSALRSGEGRDPYWGDKQDLVS

5T1V GSGALWDVPAPKEVKKGETTDGVYRVMTRRLGSTQVGVGMQEGVFHTMWHVTKGAALRSGEGRDPYWGDKQDLVS

5GPI GSGALWDVPAPKEVKKGETTDGVYRVMTRRLGSTQVGVGMQEGVFHTMWHVTKGAALRSGEGRDPYWGDKQDLVS

MH158237.1 YCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVI

5TFN YCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPAGTAGSPILDKCGRVIGLYGNGVVI

5T1V YCGPWKLDAAWDGLSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVI

5GPI YCGPWKLDAAWDGLSEVQLLAVPPGERAKNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVI

MH158237.1 KNGSYVSAITQGRREEETPVE

5TFN KNGSYVSAITQGRREEETPVECFEPS

5T1V KNGSYVSAITQGKREEETPVECFEPS

Title missing - double click to edit

TWIST CLONES (DEAD ones have Catalytic His to Ala mutation)

QQ01D2VNS2B-s002

>DV2-NS2BgsNS3

AAAAAAGGTCTCACATGGCAGATTTGGAGCTGGAGCGCGCCGCGGACGTCAAATGGGAAGATCAAGCTGAGATTTCAGGTAGCTCGCCAATACTTTCTATTACTATCTCTGAGGACGGTTCTATGTCCATTAAGAACGAGGAAGAGGAACAAACGCTCGGCGGTGGTGGTAGTGGCGGTGGTGGTGCGGGTGTGTTTTGTGGGATGTACCGTCTCCACCCCATGGGGAAAGCTGAGTTAGAGGATGAGCCTATCGCATCAAGCAGAAGGGAATACTGGGCTACTCACAAATCGGGGCCGGGTATATAAAGAGGGCACTTTCCACACAATGTGGCACGTGACACGCGGTGCCGTCCTTATGCACAAAGGGAAGCGCATCGAACCTTCTTGGGCGGATGTGAAGAAGGACCTCATCTCTTACGGTGGCGGTTGGAAGTTGGAAGGTGAATGGAAAGAGGGCGAGGAAGTGCAGGTCCTTGCCCTGGAGCCCGTAAGAATCCGCGCGCAGTACAGACGAAACCCGGCCTGTTCAAGACTAACGCCGGCACTATCGGCGCCGTGTCATTAGACTTCTCGCCCGGTACTTCTGGGATCTCCGATCATAGATAAGAAGGGTAAGGTTGTTCGGGTTGTACGGTAACGGTGTCGTAACACGCGAGCGGGCGTACGTGAGTGCAATAGCCCAGACGGAGAAGAGCATCGAAGACAACCCAGAGATCGAAGATGACATATTCCGTAAGTAATGACTCGAGAGACCAAAAAA

>DV2-NS2BgsNS3

MADLELERAADV KWEDQAEISGSSPILSITISEDGSM SIKNEEEEQTLGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIGAVSLDFSPGTS GSPIIDKKGKVVG LYGNGVVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

atgcaccatcatcatcatcattcttctggtgcgagctggagtcacccacaattcgagaaaggtggtggctctggtggcggttcaggcggctccgcctggtcccatccacagtttgaaaagggttctggtgtagatctgggtaccgagaaacctgtacttccaatccATGGCAGATTTGGAGCTGGAGCGCGCCGCGGACGTCAAATGGGAAGATCAAGCTGAGATTTCAGGTAGCTCGCCAATACTTTCTATTACTATCTCTGAGGACGGTTCTATGTCCATTAAGAACGAGGAAGAGGAACAAACGCTCGGCGGTGGTGGTAGTGGCGGTGGTGGTGCGGGTGTGTTTTGTGGGATGTACCGTCTCCACCCCATGGGGAAAGCTGAGTTAGAGGATGGAGCCTATCGCATCAAGCAGAAGGGAATACTGGGCTACTCACAAATCGGGGCCGGGTATATAAAGAGGGCACTTTCCACACAATGTGGCACGTGACACGCGGTGCCGTCCTTATGCACAAAGGGAAGCGCATCGAACCTTCTTGGGCGGATGTGAAGAAGGACCTCATCTCTTACGGTGGCGGTTGGAAGTTGGAAGGTGAATGGAAAGAGGGCGAGGAAGTGCAGGTCCTTGCCCTGGAGCCCGGTAAGAATCCGCGCGCAGTACAGACGAAACCCGGCCTGTTCAAGACTAACGCCGGCACTATCGGCGCCGTGTCATTAGACTTCTCGCCCGGTACTTCTGGGATCTCCGATCATAGATAAGAAGGGTAAGGTTGTTCGGGTTGTACGGTAACGGTGTCGTAACACGCAGCGGGCGTACGTGAGTGCAATAGCCCAGACGGAGAAGAGCATCGAAGACAACCCAGAGATCGAAGATGACATATTCCGTAAGTAATGA

MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMADLELERAADV KWEDQAEISGSSPILSITISEDGSM SIKNEEEEQTLGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIGAVSLDFSPGTS GSPIIDKKGKVVG LYGNGVVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

471017301_DV2_T7F_A06.seq

NNNNN

redo

	Score	Expect	Method	Identities	Positives	Gaps
	591 bits(1524)	0.0	Compositional matrix adjust.	296/297(99%)	296/297(99%)	0/297(0%)
Query 1	MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG	TENLYFQSMADLE	60			
Sbjct 1		60			
Query 61	LERAADV KWEDQAEISGSSPILSITISEDGSM SIKNEEEEQTLGGGSGGGGAGVLWDVP	120				
Sbjct 61	120				
Query 121	SPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRGAVLMHKGKRIEPS	180				
Sbjct 121A.....	180				
Query 181	WADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIGAVSL	240				
Sbjct 181	240				

Query 241 DFSPGTSGPSIIDKKGKVVGGLYGNVGVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK 297
Sbjct 241 297

DEAD NOT ALIVE

471017301_DV2_T7F_H12_082.seq

NNNNNNNNNNNNNNNNNNNNNTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCACCATCATCATCAT
CATTCTTCTGGTGCGAGCTGGAGTCACCCACAATTCGAGAAAGGTGGTGGCTCTGGTGGCGGTTTCAGGCGGCTCCGCCTG
GTCCCATCCACAGTTTGAAAAGGGTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGGCCGACCTCGAA
TTAGAAAGAGCCGAGATGTCAAGTGGGAAGATCAGGCCGAGATCTCCGTTCTTCACCCATACTCAGCATAACTATTAGCG
AGGATGGATCCATGTCAATCAAGAACGAAGAGGAAGAGCAGACACTGGGCGGAGGAGGAAGCGGCGGTGGCGGGGCAGG
TGTATTATGGGACGTCCCAAGTCCGCCCCCAATGGGAAAAGCCGAAGTGGGAGATGGCGCCTATCGCATAAAGCAAAAGGG
AATCTTAGGCTATAGTCAGATAGGGGCGGGTGTGTATAAGGAAGGTACATTTACACCATGTGGGCAGTGACCCGCGGGGC
GGTACTCATGCATAAGGGAAAGCGGATTGAGCCAAGCTGGGCCGATGTAAAGAAAGATTTGATTTCTATGGCGGTGGATGG
AAGCTCGAAGGCGAGTGGAAGAGGGGCGAAGAGGTCCAAGTTCTGGCCCTGGAGCCTGGTAAGAACCCGCGTGCGGTGC
AAACCAAGCCAGGCCTGTTCAAGACAAACGCGGGAAGTATAGGTGCCGTGTCGCTGGACTTCTCCCCTGGCACCTCAGGTT
CCCCGATTATCGACAAGAAGGGGAAAGTCGTTGGCTTATACGAAACGGTGTGTTACGCGTTCTGGTGCCTACGTATCCGC
AATTGCCCAAACGGAGAAGAGCATCGAGGATAACCCGGAGATCGAGGACGACATTTTCAGAAAATAATGACTCGAGCACCCG
TCGACAAGCTTGC GGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCCGAAAGGAAGCTG
AGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAACGGGTCTTGAGGGGTTTTTTGCTGAA
AGGGAGGAAGTATATCCGGATTGGCGAATGGGACNNCGCCCTNTAGCGGCGCATTAAANGCCNNGGNGGNTNGGGGGGGT
TACNNCNAGGGTGGACCGCNTACNNTTGGCAAGCCCCCTTAAGNGCCNNNCTCCTTTNNNTTNTCTTCCCTTCTTTTC
TGGC

MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMADLELERAADV KWEDQAEISGSSP
ILSITISEDGSM SIKNEEEEQTLGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWA
VTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIGAVSLDFSPG
TSGSPIIDKKGKVVGGLYGNVGVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

>DV2-NS2BgsgNS3-DEAD

AAAAAAGGTCTCACATGGCCGACCTCGAATTAGAAAAGAGCCGAGATGTCAAGTGGGAAGATCAGGCCGAGATCTCCGGTTC
TTCACCCATACTCAGCATAACTATTAGCGAGGATGGATCCATGTCAATCAAGAACGAAGAGGAAGAGCAGACACTGGGCGGA
GGAGGAAGCGGCGGTGGCGGGGCAGGTGTATTATGGGACGTCCCAAGTCCGCCCCCAATGGGAAAAGCCGAAGTGAAGATG
GCGCCTATCGCATAAAGCAAAAGGGAATCTTAGGCTATAGTCAGATAGGGGCGGGTGTGTATAAGGAAGGTACATTTACAC
CATGTGGGCAGTGACCCGCGGGGCGGTACTCATGCATAAGGGAAAGCGGATTGAGCCAAGCTGGGCCGATGTAAAGAAAGAT
TTGATTTCTATGGCGGTGGATGGAAGCTCGAAGGCGAGTGGAAGAGGGGCGAAGAGGTCCAAGTTCTGGCCCTGGAGCCTG
GTAAGAACCCGCGTGCGGTGCAAACCAAGCCAGGCCTGTTCAAGACAAACGCGGGAAGTATAGGTGCCGTGTCGCTGGACTT
CTCCCCTGGCACCTCAGGTTCCCCGATTATCGACAAGAAGGGGAAAGTCGTTGGCTTATACGAAACGGTGTGTTACGCGT
TCTGGTGCGTACGTATCCGCAATTGCCCAAACGGAGAAGAGCATCGAGGATAACCCGGAGATCGAGGACGACATTTTCAGAA
AATAATGACTCGAGAGACCAAAAAA

>DV2-NS2BgsgNS3-DEAD

MADLELERAADV KWEDQAEISGSSPILSITISEDGSM SIKNEEEEQTLGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRI
KQKGILGYSQIGAGVYKEGTFHTMWAVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNP
RAVQTKPGLFKTNAGTIGAVSLDFSPGTSGPSIIDKKGKVVGGLYGNVGVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

atgcaccatcatcatcatcattcttctggtgcgagctggagtcacccacaattcgagaaaggtggtggtctggtggcggttcaggcggc
tccgcctggtcccatccacagtttgaaaagggttctggtgtagatctgggtaccgagaacctgtacttccaatccATGGCCGACCTCGAA
TTAGAAAAGAGCCGAGATGTCAAGTGGGAAGATCAGGCCGAGATCTCCGTTCTTCACCCATACTCAGCATAACTATTAGCGAGGATGGA
TCCATGTCAATCAAGAACGAAGAGGAAGAGCAGACACTGGGCGGAGGAGGAAGCGGCGGTGGCGGGGCAGGTGTATTATGGGACGTCCCA
AGTCCGCCCCCAATGGGAAAAGCCGAAGTGAAGATGGCGCCTATCGCATAAAGCAAAAGGGAATCTTAGGCTATAGTCAGATAGGGGCG
GGTGTGTATAAGGAAGGTACATTTACACCATGTGGGCAGTGACCCGCGGGGCGGTACTCATGCATAAGGGAAAGCGGATTGAGCCAAGC
TGGGCCGATGTAAAGAAAGATTTGATTTCTATGGCGGTGGATGGAAGCTCGAAGGCGAGTGGAAGAGGGGCGAAGAGGTCCAAGTTCTG
GCCCTGGAGCCTGGTAAGAACCCGCGTGCGGTGCAAACCAAGCCAGGCCTGTTCAAGACAAACGCGGGAAGTATAGGTGCCGTGTCGCTG
GACTTCTCCCCTGGCACCTCAGGTTCCCCGATTATCGACAAGAAGGGGAAAGTCGTTGGCTTATACGAAACGGTGTGTTACGCGTTCT
GGTGCGTACGTATCCGCAATTGCCCAAACGGAGAAGAGCATCGAGGATAACCCGGAGATCGAGGACGACATTTTCAGAAAATAATGA

MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMADLELERAADV KWEDQAEISGSSPILSITISEDG
SMSIKNEEEEQTLGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWAVTRGAVLMHKGKRIEPS
WADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNAGTIGAVSLDFSPGTSGPSIIDKKGKVVGGLYGNVGVTR
SGAYVSAIAQTEKSIEDNPEIEDDIFRK

471017301_DV2_DEAD_T7F_B06.seq

NNNNNNNNNNNNNNNNCTCTAGAATANTTTTGTTTAACTTTAAGAAGGAGATATACATATGCACCATCATCANCNTNATTCTTCT
GGTGCAGCTGGAGTCACCACAATTCGAGAAAGGTGGTGGCTCTGGTGGCGGTTCCAGGCGGCTCCGCCTGGTCCCATCC
ACAGTTTGAAAAGGGTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGGCAGATTTGGAGCTGGAGCGC
GCCGCGGACGTCAAATGGGAAGATCAAGCTGAGATTTAGGCTAGCTCGCCAATACTTTCTATTACTATCTCTGAGGACGGTT
CTATGTCCATTAAGAACGAGGAAGAGGAACAAACGCTCGGCGGTGGTGGTAGTGGCGGTGGTGGTGGCGGTGTTTTGTGG
GATGTACCGTCTCCACCACCCATGGGGAAAGCTGAGTTAGAGGATGGAGCCTATCGCATCAAGCAGAAGGGAATACTGGGC
TACTACAAATCGGGGCCGGGGTATATAAAGAGGGGCACTTTCCACACAATGTGGCACGTGACACGCGGTGCCGTCTTATGC
ACAAAGGGAAGCGCATCGAACCTTCTGGGCGGATGTGAAGAAGGACCTCATCTTACGGTGGCGGTTGGAAGTTGGAAG
GTGAATGGAAGAGGGCGAGGAAGTGCAGGTCCTGCCCTGGAGCCCGGTAAGAATCCGCGCGCAGTACAGACGAAACCC
GGCCTGTTCAAGACTAACGCCGGCACTATCGGCGCCGTGTCATTAGACTTCTCGCCCGGTACTTCGGGATCTCCGATCATAG
ATAAGAAGGGTAAGGTTGTCGGGTTGTACGGTAACGGTGTCTGAACACGCAGCGGGGCGTACGTGAGTGCAATAGCCCAG
ACGGAGAAGAGCATCGAAGACAACCCAGAGATCGAAGATGACATATCCGTAAGTAATGACTCGAGCACCCGTCGACAAGC
TTGCGGCCGCACTCGAGCACCACCACCACCACCTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAACCGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAAC
TATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGGNTGGTGGTTACGCGCAGNNNGAC
CGCTACCTTNGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCTTTTNCNCCACGTTGCCGGGTTTTCCCCGTC
AAGNNNNAAATCGGGGGCTCCCTTTTAGGNTTCCNAATTTANGGNTTTACNGNCCCTNCCC

MHHHXXSSGASWSHPQFEKGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMADLELERAADV KWEDQAEISGSSP
ILSITISEDGSM SIKNEEEEQTLGGGGSGGGGAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMW
HVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKNAGTIGAVSLDFSP
GTSGPSIIDKKGKVVGLYGNGVVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK

Score	Expect	Method	Identities	Positives	Gaps
582 bits(1499)	0.0	Compositional matrix adjust.	293/297(99%)	293/297(98%)	0/297(0%)
Query 1	MHHHHHHSSGASWSHPQFEKGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMADLE		60		
Sbjct 1xxx.....		60		
Query 61	LERAADV KWEDQAEISGSSPILSITISEDGSM SIKNEEEEQTLGGGGSGGGGAGVLWDVP		120		
Sbjct 61		120		
Query 121	SPPPMGKAELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWAVTRGAVLMHKGKRIEPS		180		
Sbjct 121H.....		180		
Query 181	WADVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKNAGTIGAVSL		240		
Sbjct 181		240		
Query 241	DFSPGTSGPSIIDKKGKVVGLYGNGVVTRSGAYVSAIAQTEKSIEDNPEIEDDIFRK		297		
Sbjct 241		297		

NOT DEAD BUT ALIVE

QQ01WNVNS2B-s002

>WNV-NS2BgsgNS3

AAAAAAGGTCTCACATGAGCACTGATATGTGGATTGAACGCACGGCAGACATCAGCTGGGAGTCGGATGCTGAGATCACAGG
TTCAAGCGAAAGAGTGGATGTACGATTAGACGATGACGGTAATTTTCAACTGATGAATGACCCCTGGCGCGCCGTGGAAAGGC
GGCGGCGGCTCTGGTGGAGGAGGTGGCGTTTTTGTGGGACACGCCGTCCCCCAAAGAATACAAGAAGGGTGACACGACTACCG
GTGTTTACAGAATTATGACCCGGGGTTTGCTTGGGTCGTACCAGGCCGGGGCAGGGGTGATGGTAGAGGGAGTTTTCCATAC
CTTATGGCACACTACCAAAGGTGCCGCACTTATGTCTGGTGAGGGTCGCCTTGATCCTTATTGGGGTAGCGTAAAAGAGGAT
CGGCTCTGTTACGGCGGACCATGGAAGCTTCAACACAAGTGGAACGGACAGGATGAAGTTCAGATGATTGTCGTAGAGCCGG
GCAAGAATGTCAAGAACGTGCAAACAAAGCCCGGTGTATTCAAACGCCGGAAGGCGAAATTGGCGCTGTGACCCTTGACTT
TCCAACCTGGGACCAGCGGCTCACCAATAGTCGACAAGAATGGAGATGTCATTGGGTTGTACGGAATGGTGTATTATGCC
AATGGCTCGTACATCTCCGCTATCGTACAAGGAGAGCGCATGGACGAACCGATACCTGCAGGCTTCGAACCTGAGATGTTAA
GAAAGAAGTAATGACTCGAGAGACCAAAAAA

>WNV-NS2BgsgNS3

MSTDMWIERTADISWESDAEITGSSSERVDVRLDDDNFQLMNDPGAPWKGGGGSGGGGGVLWDTPSPKEYKKGDTTGTGYRI
MTRGLLSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMI VVEPGKNVK
NVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAI VQGERMDEPI PAGFEPEMLRKK

atgcaccatcatcatcatcattcttctggtgcgagctggagtcacccacaattcgagaaaggtggtggtctctggtggcggttcaggcggc
tccgcctggtcccatccacagtttgaaaagggttctggtgtagatctgggtaccgagaaacctgtacttccaatccATGAGCACTGATATG
TGGATTGAACGCACGGCAGACATCAGCTGGGAGTCGGATGCTGAGATCACAGGTTCAAGCGAAAGAGTGGATGTACGATTAGACGATGAC
GGTAATTTTCAACTGATGAATGACCCTGGCGCGCCGTGGAAAGGCGGCGGCTCTGGTGGAGGAGGTGGCGTTTTGTGGGACACGCCG
TCCCCCAAAGAATACAAGAAGGGTGACACGACTACCGGTGTTTACAGAATTATGACCCGGGGTTTGCTTGGGTCGTACCAGGCCGGGCA
GGGGTGATGGTAGAGGGAGTTTTCCATACCTTATGGCACACTACCAAAGGTGCCGCACTTATGTCTGGTGAGGGTCGCCTTGATCCTTAT
TGGGGTAGCGTAAAAGAGGATCGGCTCTGTTACGGCGGACCATGGAAGCTTCAACACAAGTGGAAACGGACAGGATGAAGTTCAGATGATT
GTCGTAGAGCCGGGCAAGAATGTCAAGAACGTGCAAACAAAGCCCGGTGTATTCAAAACGCCGGAAGGCGAAATTTGGCGCTGTGACCCTT
GACTTTCCAACCTGGGACCAGCGGCTCACCAATAGTCGACAAGAATGGAGATGTCATTGGGTTGTACGGAAATGGTGTATTATGCCCAAT
GGCTCGTACATCTCCGCTATCGTACAAGGAGAGCGCATGGACGAACCGATACCTGCAGGCTTCGAACCTGAGATGTTAAGAAAGAAGTAA
TGA

MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMSTDMWIERTADISWESDAEITGSSSERVDVRLDDD
GNFQLMNDPGAPWKGGGSGGGGGVLWDTPSPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRDPY
WGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPN
GSYISAIVQGERMDEPI PAGFEPEMLRKK

471017301_WNV_T7F_E06.seq

NNNNNNNNNNNNNTNCCNTCTAGATAATTTTNNNTNANCTTTAAGAAGGAGATATACATATGCACCATCATCATCNNNNNNNNNNTGGTG
CGAGCTGGAGTCACCCACAATTCGAGAAAGGTGGTGGCTCTGGTGGCGGTTTCAGGCGGCTCCGCCTGGTCCCATCCACAGTTTGAAGAG
GTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGAGCACTGATATGTGGATTGAACGCACGGCAGACATCAGCTGGG
AGTCGGATGCTGAGATCACAGGTTCAAGCGAAAGAGTGGATGTACGATTAGACGATGACGGTAATTTTCAACTGATGAATGACCCTGGCG
CGCCGTGGAAAGGCGGCGGCGGCTCTGGTGGAGGAGGTGGCGTTTTGTGGGACACGCCGTCCCCCAAAGAATACAAGAAGGGTGACACGA
CTACCGGTGTTTACAGAATTATGACCCGGGGTTTGCTTGGGTGCTACCGAGCCGGGGCAGGGGTGATGGTAGAGGGAGTTTTCCATACCT
TATGGCACACTACCAAAGGTGCCGCACTTATGTCTGGTGAGGGTCGCCTTGATCCTTATTGGGGTAGCGTAAAAGAGGATCGGCTCTGTT
ACGGCGGACCATGGAAGCTTCAACACAAGTGGAAACGGACAGGATGAAGTTCAGATGATTGTCTGATAGAGCCGGGCAAGAATGTCAAGAACG
TGCAAACAAAGCCCGGTGTATTCAAAACGCCGGAAGGCGAAATTTGGCGCTGTGACCCTTGACTTTCCAACCTGGGACCAGCGGCTCACCAA
TAGTCGACAAGAATGGAGATGTCATTGGGTTGTACGGAAATGGTGTATTATGCCCAATGGCTCGTACATCTCCGCTATCGTACAAGGAG
AGCGCATGGACGAACCGATACCTGCAGGCTTCAACCTGAGATGTTAAGAAAGAAGTAATGACTCGAGCACCCGTCGACAAGCTTGCGGC
CGCACTCGAGCACCACCACCACCACCTGAGATCCGGCTGCTAACAAAGCCCCGAAAGGAAGCTGAGTTGGCTGCTGCCNCCGCTGAGCA
ATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTTGTGAAAGGAGGAACCTATATCCGGATTGGCGAATGGGAC
GCNCCCCTGTAGCGNNCCATTAAGCGCGGCGGGGTGGGGNTTACGCNN

MHHHHXXXXGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMSTDMWIERTADISWESDAEITGSSSERVDVRLDDD
GNFQLMNDPGAPWKGGGSGGGGGVLWDTPSPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRDPY
WGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPN
GSYISAIVQGERMDEPI PAGFEPEMLRKK

Score	Expect	Method	Identities	Positives	Gaps
597 bits(1540)	0.0	Compositional matrix adjust.	295/299(99%)	295/299(98%)	0/299(0%)
Query 1	MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMSTDM	60			
Sbjct 1XXXX.	60			
Query 61	WIERTADISWESDAEITGSSSERVDVRLDDDG N FQLMNDPGAPWKGGGSGGGGGVLWDTP	120			
Sbjct 61	120			
Query 121	SPKEYKKGDTTGTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRDPY	180			
Sbjct 121	180			
Query 181	WGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTL	240			
Sbjct 181	240			
Query 241	DFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMDEPI PAGFEPEMLRKK	299			
Sbjct 241	299			

>WNV-NS2BgsgNS3-DEAD

AAAAAAGGTCTCACATGAGCACGGATATGTGGATAGAACGTACAGCCGACATTAGCTGGGAAAAGTGACGCTGAAATAACCGG
ATCCAGTGAGCGCGTCGATGTGCGCTTAGATGACGACGGAATTTCCAGTTGATGAATGATCCGGGTGCTCCGTGGAAAGGT
GGCGGCGGGTCGGGTGGCGGCGGCGGTGTTCTGTGGGATACACCCTACCTAAGGAATATAAGAAAGGGGACACGACAACAG
GGGTATACAGAATCATGACGCGCGGTCTGCTTGGCTCCTACCAGGCAGGGGCTGGTGTGATGTTAGAAAGGCGTCTTCCACAC
TCTCTGGGCTACGACTAAAGGGGCGGCTCTGATGTCTGGTGAAGGTTCGGTTGGACCCGTACTGGGGCTCTGTAAAGGAAGAC

CGCCTCTGTTATGGTGGGCCTTGGAATTACAACACAAATGGAACGGTCAAGATGAGGTTCAAATGATCGTAGTGAGACCGG
GTAAGAACGTAAAGAACGTTCAAACAAAGCCAGGCGTATTTAAAACCCCCGAAGGCGAGATCGGCGCCGTCACGCTGGACTT
TCCACACAGGTACAAGCGGTTCTCCTATTGTAGACAAGAATGGTGACGTAATTGGCCTCTACGGCAATGGTGTCTATAATGCCG
AACGGCTCGTATATCAGCGCAATTGTACAAGGTGAACGGATGGACGAGCCTATCCCGGCTGGTTTTCGAACCTGAGATGTTGC
GTAAGAAGTAATGACTCGAGAGACCAAAAAA

>WNV-NS2BgsgNS3-DEAD

MSTDMWIERTADISWESDAEITGSSSERVDVRLDDDGNFQLMNDPGAPWKGGGSGGGGGVLWDTPSPKEYKKGDTTTTGVYRI
MTRGLLSYQAGAGVMVEGVFHTLWATTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKNVK
NVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMDEPIPAGEPEMLRKK

atgcaccatcatcatcatcattcttctggtgcgagctggagtcacccacaattcgagaaaggtggtggtctctggtggcggttcaggcggc
tccgcctggtcccatccacagtttgaaaagggttctggtgtagatctgggtaccgagaaacctgtacttccaatccATGAGCACGGATATG
TGGATAGAACGTACAGCCGACATTAGCTGGGAAAGTGACGCTGAAATAACCGGATCCAGTGAGCGCGTCGATGTGCGCTTAGATGACGAC
GGAAATTTCCAGTTGATGAATGATCCGGGTGCTCCGTGGAAAGGTGGCGGCGGGTCGGGTGGCGGCGGCGGTGTTCTGTGGGATACACC
TCACCTAAGGAATATAAGAAAGGGGACACGACAACAGGGGTATACAGAATCATGACGCGCGGTCTGCTTGGCTCCTACCAGGCAGGGGCT
GGTGTCTATGGTAGAAGGCGTCTTCCACACTCTCTGGGCTACGACTAAAGGGGCGGCTCTGATGTCTGGTGAAGGTCGGTTGGACCCGTAC
TGGGGCTCTGTAAAGGAAGACCGCCTCTGTTATGGTGGGCCTTGGAATTACAACACAAATGGAACGGTCAAGATGAGGTTCAAATGATC
GTAGTGGAGCCGGGTAAGAACGTAAAGAACGTTCAAACAAAGCCAGGCGTATTTAAAACCCCCGAAGGCGAGATCGGCGCCGTCACGCTG
GACTTTCCACAGGTACAAGCGGTTCTCCTATTGTAGACAAGAATGGTGACGTAATTGGCCTCTACGGCAATGGTGTCTATAATGCCGAAC
GGCTCGTATATCAGCGCAATTGTACAAGGTGAACGGATGGACGAGCCTATCCCGGCTGGTTTTCGAACCTGAGATGTTGCGTAAGAAGTAA
TGA

MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMSTDMWIERTADISWESDAEITGSSSERVDVRLDD
GNFQLMNDPGAPWKGGGSGGGGGVLWDTPSPKEYKKGDTTTTGVYRIMTRGLLSYQAGAGVMVEGVFHTLWATTKGAALMSGEGRLDPY
WGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMP
NGSYISAIVQGERMDEPIPAGEPEMLRKK

471017301_WNV_DEAD_T7F_F06.seq

NNNNNNNNNNNNNNCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGCACCATCATCATCATGCTT
CTGGTGCGAGCTGGAGTCACCCACAATTCGAGAAAGGTGGTGGCTCTGGTGGCGGTTCCAGGCGGCTCCGCCTGGTCCCAT
CCACAGTTTGAAAAGGGTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGAGCACGGATATGTGGATAG
AACGTACAGCCGACATTAGCTGGGAAAGTGACGCTGAAATAACCGGATCCAGTGAGCGCGTCGATGTGCGCTTAGATGACG
ACGGAATTTCCAGTTGATGAATGATCCGGGTGCTCCGTGGAAAGGTGGCGGCGGGTCGGGTGGCGGCGGCGGTGTTCTG
TGGGATACACCCTCACCTAAGGAATATAAGAAAGGGGACACGACAACAGGGGTATACAGAATCATGACGCGCGGTCTGCTT
GGCTCCTACCAGGCAGGGGCTGGTGTCTGGTAGAAGGCGTCTTCCACACTCTCTGGGCTACGACTAAAGGGGCGGCTCT
GATGTCTGGTGAAGGTCGGTTGGACCCGTACTGGGGCTCTGTAAAGGAAGACCGCCTCTGTTATGGTGGGCCTTGGAATT
ACAACACAAATGGAACGGTCAAGATGAGGTTCAAATGATCGTAGTGGAGCCGGGTAAGAACGTAAAGAACGTTCAAACAAA
GCCAGGCGTATTTAAAACCCCCGAAGGCGAGATCGGCGCCGTCACGCTGGACTTTCCACAGGTACAAGCGGTTCTCCTATT
GTAGACAAGAATGGTGACGTAATTGGCCTCTACGGCAATGGTGTCTATAATGCCGAACGGCTCGTATATCAGCGCAATTGTAC
AAGGTGAACGGATGGACGAGCCTATCCCGGCTGGTTTTCGAACCTGAGATGTTGCGTAAGAAGTAATGACTCGAGCACCCGT
CGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGCCTCTAAACGGGTCTTGAGGGGTTTTTGTGAAA
GGAGGAATATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGGTGTGGGGGTACGCGC
AGGNTGACCGCTANNNTTGCCNNGNCCNTAGCGCCCGCTCCTTTNNNTTTCTNCCCTTCTTTCCNNCNNGGTTNCNG
GGTTTTCCCGNNAAGCNCNAANTCGGGGGCTCCCTTNAGGGTCCAATTNNNGGCTTNANNGCCCCNNNCCCCAAAAAN
NN

MHHHHHHASGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMSTDMWIERTADISWESDAEITGSS
ERVDVRLDDDG NFQLMNDPGAPWKGGGSGGGGGVLWDTPSPKEYKKGDTTTTGVYRIMTRGLLSYQAGAGVMVEGVFH
TLWATTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPGVFKTPEGEIGAVT
LDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMDEPIPAGEPEMLRKK

Score	Expect	Method	Identities	Positives	Gaps
606 bits(1563)	0.0	Compositional matrix adjust.	298/299(99%)	299/299(100%)	0/299(0%)
Query 1	MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLG TENLYFQSMSTDM		60		
Sbjct 1 A		60		
Query 61	WIERTADISWESDAEITGSSSERVDVRLDDDG NFQLMNDPGAPWKGGGSGGGGGVLWDTP		120		

QQ01ZVNS2B-s001

NNNNNNNNNNNNNNNTCTAGANTAATTTTNTTTANCTTTAAGAAGGAGATATACATATGCACCATCATCNNCNNNATTCTTC
TGGTGCGAGCTGGAGTACCCACAATTTCGAGAAAGGTGGTGGCTCTGGTGGCGGTTGAGGCGGCTCCGCTGGTCCCATC
CACAGTTTGAAAAGGGTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGAGTGGAAAGTCTGTGATAT
GTACATCGAGCGAGCCGGCGACATCACATGGGAGAAAGACGCAGAGGTTACCGGGAATAGCCCGCGGCTTGACGTTGCGTT
GGACGAGAGTGGTGACTTCTCTGGTCGAGGACGACGGACCACCTATGCGTGAGGGAGGAGGTGGGTGAGGTGGTGGT
GGTGGCTCGGGTGCTCTTTGGGATGTCCCTGCCCCAAAGAGGTGAAGAAGGGTGAGACGACCGACGGAGTGTATCGCGT
TATGACAAGACGTTTATTAGGGTCAACACAAGTAGGTGTTGGTGTAATGCAAGAGGGCGTTTTCCACACAATGTGGCACGTT
ACTAAGGGCTCAGCTCTTCGGTCAGGTGAGGGACGCCCTTGACCCGTA CTGGGGCGACGTTAAGCAAGATCTGGTGTCTTAT
TGTGGTCCCTGGAAGTTAGACGCTGCATGGGACGGTCATAGCGAGGTCCAATTGCTGGCAGTGCCCCGGGTGAGAGAGC
GCGAAATATTCAAACGCTTCCTGGGATTTTCAAGACCAAGGACGGTGACATCGGTGCAGTGGCGCTGGATTACCCGGCGGG
AACGTCCGGCAGTCCGATTTTGGATAAGTGCGGTGCGGTAATCGGACTGTATGGTAATGGGGTGGTGATTAGAACGGCTC

ATACGTGTCTGCAATAACCCAGGGCCGTCGAGAGGAAGAACTCCTGTTGAGTAATGACTCGAGCACCCGTCGACAAGCTT
GCGGCCGCACTCGAGCACCACCACCACCCTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGC
TGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTGGCTGAAAGGAGGA
ATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGGGTGGTGGTTACNGCAGCGTGACC
GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTNCNTTTTNTTCCTTNCCTTCTCGCCACGTTGCGCGGGTTTTCCCGTC
AAGCNNAATCGGGGGCTCCCTTTAGGGTTCGATTTNNNGCTTTACGGCCNNNNNNCCCNAAAAAATTTNNATTAGGG
GTGATG

MHHHXXSSGASWSHPQFEKGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMGKSVDMYIERAGDITWEKDAEVT
GNSPRLDVALDESGDFSLVEDDGPPMREGGGSGGGGGSGALWDVPAPKEVKKGETTDGVYRVMTRLLGSTQVGVGMQ
EGVFHTMWHVTKGSALRSGEGRLDPYWGDVKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGD
GAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE

Score	Expect	Method	Identities	Positives	Gaps
587 bits(1513)	0.0	Compositional matrix adjust.	292/295(99%)	292/295(98%)	0/295(0%)
Query 1	MHHHHHHSSGASWSHPQFEKGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMGKS				60
Sbjct 1 xxx				60
Query 61	VDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMREGGGSGGGGGSG				120
Sbjct 61				120
Query 121	ALWDVPAPKEVKKGETTDGVYRVMTRLLGSTQVGVGMQEGVFHTMWHVTKGSALRSGE				180
Sbjct 121				180
Query 181	GRLDPYWGDVKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGD				240
Sbjct 181				240
Query 241	IGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE				295
Sbjct 241				295

>ZV-NS2BgsgNS3-DEAD

AAAAAAGGTCTCACATGAGCGGTAAGTCTGTGCATATGTACATTGAACGTGCCGGTGACATCACCTGGGAGAAGGATGCAGA
AGTGACGGGTAACCTCGCCAAGACTTGATGTTGCACTCGACGAGTCGGGCGATTTTAGCTTAGTTGAAGATGACGGTCCGCCT
ATGCGAGAGGGCGGCGGCGGTAGCGGCGGTGGTGGTGGTAGCGGAGCATTATGGGATGTGCCGGCACCCAAGGAAGTAAAGA
AAGGGGAGACGACAGACGGTGTATACAGAGTTATGACGAGAAGACTGTTGGGTAGCACACAAGTTGGTGGTGGCGTAATGCA
GGAAGGAGTCTTCCACACCATGTGGGCGGTACCAAGGGCAGTGCGCTTCGCAGCGGTGAGGGACGTCTGGATCCGTACTGG
GGAGACGTGAAGCAAGACCTGGTCAGTTATTGTGGTCCGTGGAAGCTTGACGCAGCCTGGGACGGCCATAGCGAGGTGCAAC
TCCTCGCCGTACCGCCGGGAGAGCGCGCCCGTAATATTCAAACCTGCCCCGGTATCTTTAAACCAAGGACGGAGATATAGG
CGCAGTGGCTCTTGATTATCCCGCAGGGACTAGCGGTAGTCCTATTCTGGATAAGTGCGGCCGAGTGATTGGCTTATACGGA
AATGGAGTGGTTATCAAGAATGGAAGCTATGTCTCAGCCATAACCCAAGGTTCGTCGGAAGAGGAGACACCCGTAGAATAAT
GACTCGAGAGACCAAAAAA

>ZV-NS2BgsgNS3-DEAD

MSGKSVDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMREGGGSGGGGGSGALWDVPAPKEVKKGETT
DGVYRVMTRLLGSTQVGVGMQEGVFHTMWAVTKGSALRSGEGRLDPYWGDVKQDLVSYCGPWKLDAAWDGHSEVQLLAVP
PGERARNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE

atgcaccatcatcatcatcattcttctggtgcgagctggagtcacccacaattcgagaaaggtggtggtctctggtggcggttcaggcggc
tccgcctggtcccacacagtttgaaaagggttctggtgtagatctgggtaccgagaaacctgtacttccaatccATGAGCGGTAAGTCT
GTCGATATGTACATTGAACGTGCCGGTGACATCACCTGGGAGAAGGATGCAGAAGTGACGGGTAACCTCGCCAAGACTTGATGTTGCACTC
GACGAGTCGGGCGATTTTAGCTTAGTTGAAGATGACGGTCCGCCTATGCGAGAGGGCGGCGGCGGTAGCGGCGGTGGTGGTGGTAGCGGA
GCATTATGGGATGTGCCGGCACCCAAGGAAGTAAAGAAAGGGGAGACGACAGACGGTGTATACAGAGTTATGACGAGAAGACTGTTGGGT
AGCACACAAGTTGGTGGTGGCGTAATGCAGGAAGGAGTCTTCCACACCATGTGGGCGGTACCAAGGGCAGTGCGCTTCGCAGCGGTGAG
GGACGTCTGGATCCGTACTGGGAGACGTGAAGCAAGACCTGGTCAGTTATTGTGGTCCGTGGAAGCTTGACGCAGCCTGGGACGGCCAT
AGCGAGGTGCAACTCCTCGCCGTACCGCCGGGAGAGCGCGCCCGTAATATTCAAACCTGCCCCGGTATCTTTAAACCAAGGACGGAGAT
ATAGGCGCAGTGGCTCTTGATTATCCCGCAGGGACTAGCGGTAGTCCTATTCTGGATAAGTGCGGCCGAGTGATTGGCTTATACGGAAT
GGAGTGGTTATCAAGAATGGAAGCTATGTCTCAGCCATAACCCAAGGTTCGTCGGAAGAGGAGACACCCGTAGAATAATGA

MHHHHHHSSGASWSHPQFEKGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMGKSVDMYIERAGDITWEKDAEVTGNSPRLDVAL
DESGDFSLVEDDGPPMREGGGSGGGGGSGALWDVPAPKEVKKGETTDGVYRVMTRLLGSTQVGVGMQEGVFHTMWAVTKGSALRSGE
GRLDPYWGDVKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPAGTSGSPILDKCGRVIGLYGN

471017301_ZV_DEAD_T7F_D06.seq

NNNNNNNNNNNNNNNNNTCTNNNNNNNATTTTNNTTAACTTTAAGAAGGAGATATACATATGCACCATCATCNNCANNNNNTCT
TCTGGTGCGAGCTGGAGTCACCCACAATTTCGAGAAAGGTGGTGGCTCTGGTGGCGGTTTCAGGCGGCTCCGCCTGGTCCCA
TCCACAGTTTGAAAAGGGTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGAGCGGTAAGTCTGTCGAT
ATGTACATTGAACGTGCCGGTGACATCACCTGGGAGAAGGATGCAGAAAGTACGGGTAACCTGCCAAGACTTGATGTTGCA
CTCGACGAGTCGGGCGATTTTAGCTTAGTTGAAGATGACGGTCCGCCTATGCGAGAGGGCGGCGGCGGTAGCGGCGGTGG
TGGTGGTAGCGGAGCATTATGGGATGTGCCGGCACCCAAGGAAGTAAAGAAAGGGGAGACGACAGACGGTGTATACAGAG
TTATGACGAGAAGACTGTTGGGTAGCACACAAGTTGGTGTGGCGTAATGCAGGAAGGAGTCTTCCACACCATGTGGGCGG
TCACCAAGGGCAGTGCGCTTCGCAGCGGTGAGGGACGTCTGGATCCGTAAGTGGGAGACGTGAAGCAAGACCTGGTCAGT
TATTGTGGTCCGTGGAAGCTTGACGCAGCCTGGGACGGCCATAGCGAGGTGCAACTCCTCGCCGTACCGCCGGGAGAGCG
CGCCCGTAATATTCAAACCTCTGCCCGGTATCTTTAAACCAAGGACGGAGATATAGGCGCAGTGGCTCTTGATTATCCCGCAG
GGACTAGCGGTAGTCTTATTCTGGATAAGTGCGGCCGAGTGATTGGCTTATACGGAAATGGAGTGGTTATCAAGAATGGAAG
CTATGTCTCAGCCATAACCCAAGGTCGTCGCGAAGAGGAGACACCCGTAGAATAATGACTCGAGCACCCGTCGACAAGCTT
GCGGCCGCACTCGAGCACCACCACCACCACCTGAGATCCGGCTGCTAACAAAGCCCCGAAAGGAAGCTGAGTTGGCTGC
TGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACT
ATTTCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGNGNATTAAGCGCGGCGGGNTGGGGGTTACGCGCAGCNTGACCG
CTACCTTTGCCAGGNCCCTAGNNCCCCTTCTTTTNNTTTTTCTTTCTTTTNGCCNNGGTTNGCCGNTTTNCCCNGT
CANGNNNTAAATCGGGGGGNNNCCTTAGGGTTC

MHHHXXSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMGKSVDMYIERAGDITWEKDAEVT
GNSPRLDVALDESGDFSLVEDDGPPMREGGGGSGGGGSGALWDVPAPKEVKKGETTDGVYRVMTRRLLGSTQVGVGMQ
EGVFHTMWAVTKGSALRSGEGRDPYWGDKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGI
GAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE

Score	Expect	Method	Identities	Positives	Gaps
586 bits(1511)	0.0	Compositional matrix adjust.	292/295(99%)	292/295(98%)	0/295(0%)
Query 1	MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGVDLGTENLYFQSMGKS				60
Sbjct 1XXX.....				60
Query 61	VDMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMREGGGGSGGGGSG				120
Sbjct 61				120
Query 121	ALWDVPAPKEVKKGETTDGVYRVMTRRLLGSTQVGVGMQEGVFHTMWAVTKGSALRSGE				180
Sbjct 121				180
Query 181	GRLDPYWGDKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGD				240
Sbjct 181				240
Query 241	IGAVALDYPAGTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVE				295
Sbjct 241				295

Title missing - double click to edit

Transformed BL21[DE3]RR with plasmid
Used to inoculate 15 mL SOC grown o/n 37C 200 rpm
Used 15 mL of o/n to inoculate 1L of AIM-TB and grew 37C 200 rpm 3 h then 21 h 18C 200 rpm
Harvested 4000 g 20 minutes
Froze pellet -80C (approx 45 g wcw/L)

Dissolved pellet in 200 mL base buffer + 1 % TX-100 and 0.5 mg/mL lysozyme and 1 ug/mL benzonase
 Centrifuge 30,000g 1 h 12C
 Apply soluble fraction to 5 mL Ni-Speharose FF
 Wash 50 mL of BaseBuffer + 30 mM Imidazole 3x
 Elute 2 x 10 mL BaseBuffer + 300 mM Imidazole

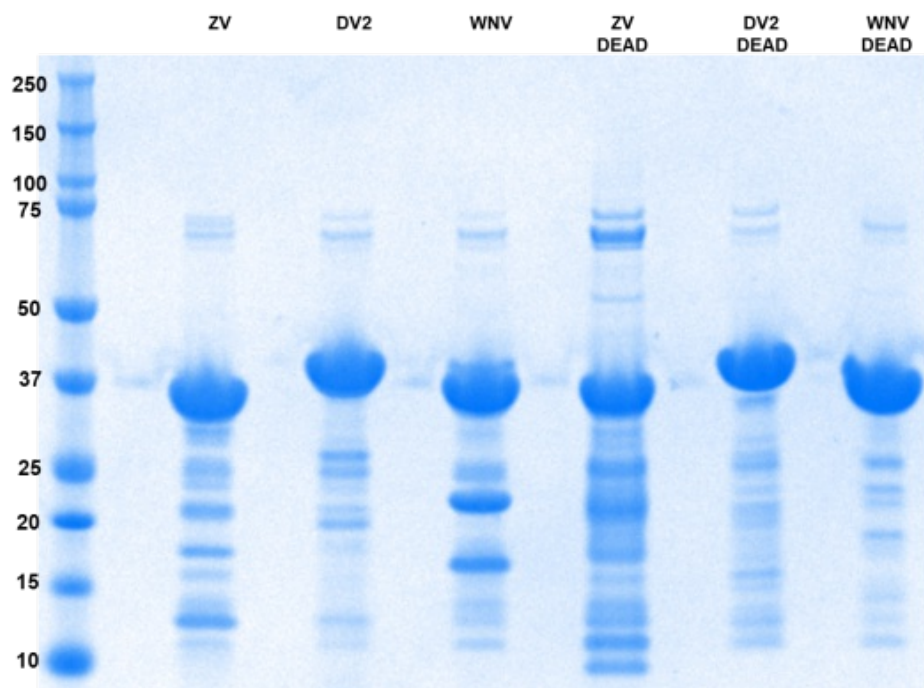
ZV A280 = 11
 DV2 A280 = 10
 WNV A280 = 12.1
 ZV DEAD A280 = 3.3
 DV2 DEAD A280 = 11.1
 WNV A280 = 15.8

pretty good yields except for ZV DEAD

WNV maybe proteolysing

Protein	Volume	A280	Yield (mg)
ZV	20	11	127
DV2	20	10	116
WNV	20	12	115
ZV DEAD	20	3.3	32
DV2 DEAD	20	11	123
WNV DEAD	20	16	178

SDS-PAGE after IMAC



ZV = 31250.5
 DV2 = 31543.0
 WNV = 32095.6
 ZV DEAD = 31184.5
 DV2 DEAD = 31477.0
 WNV DEAD = 32029.6

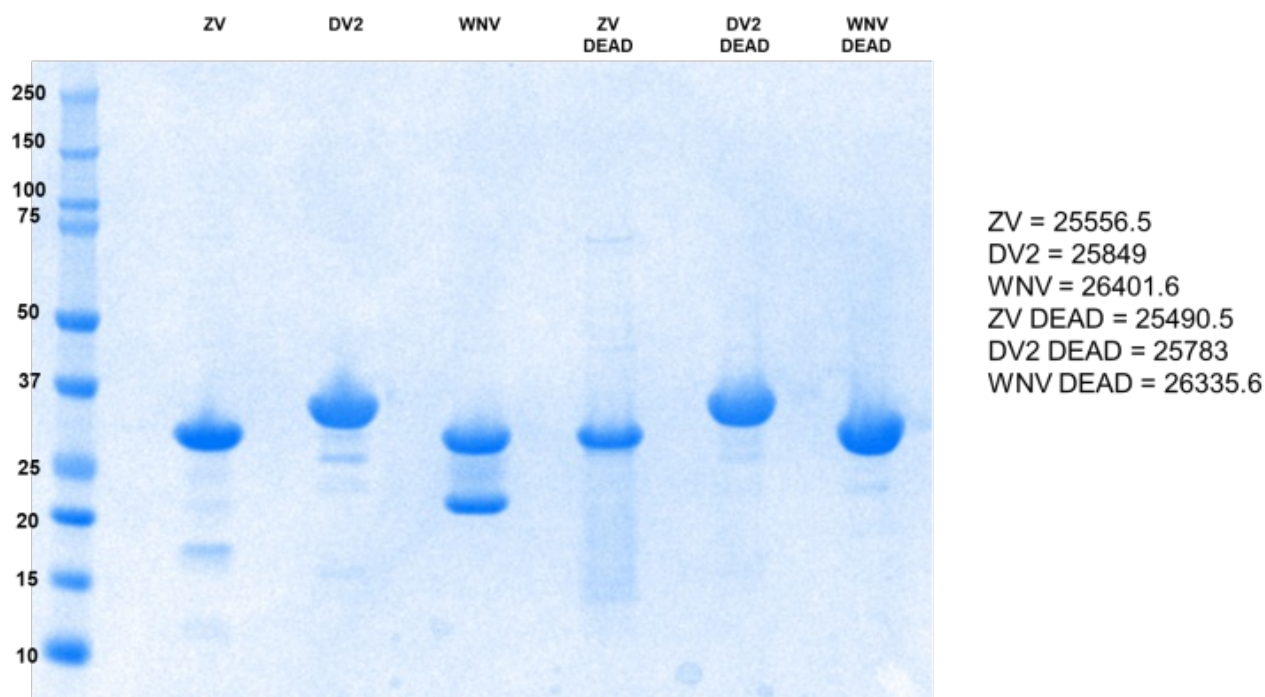
Dialyze 6 x20 mL of samples o/n against 3L of base buffer after adding 1:10 TEVsh

Reapply to 5 mL Ni-Sepharose FF (if a lot of protein (>100 mg) ran through a second 5mL column to remove TEV and uncleaved protein. Wash column with 15 mL BaseBuffer + 30 mM Imidazole

Freeze until ready to run SEC

Base Buffer = 10 mM HEPES, 500 mM NaCl, 5 % glycerol, 1 mM TCEP

SDS PAGE After TEV o/n dialysis & rev IMAC



After o/n dialysis & rev IMAC

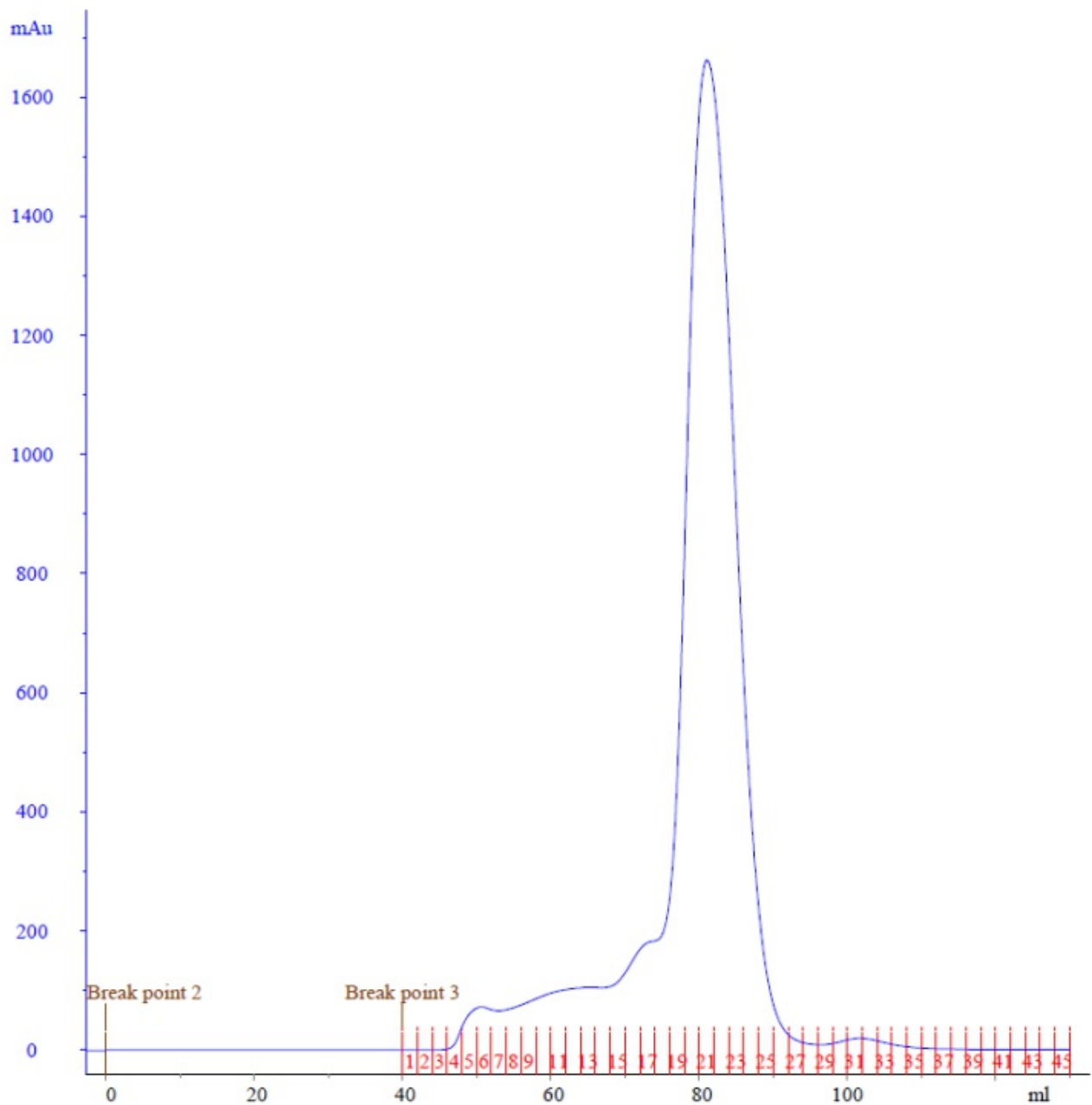
Protein	Volume	A280	Yield (mg)
ZV	30	5.2	96
DV2	30	4.7	87
WNV	30	6.1	89
ZV DEAD	30	1.2	17
DV2 DEAD	30	5.1	90
WNV DEAD	30	7.7	136

Concentrated to around 5 mL and loaded onto 125 mL Superose 12 PG column ran at 1.5 ml/min at RT using base buffer as mobile phase (10 mM HEPES pH 7.5, 500 mM NaCl, 5 % glycerol, 0.5 mM TCEP)

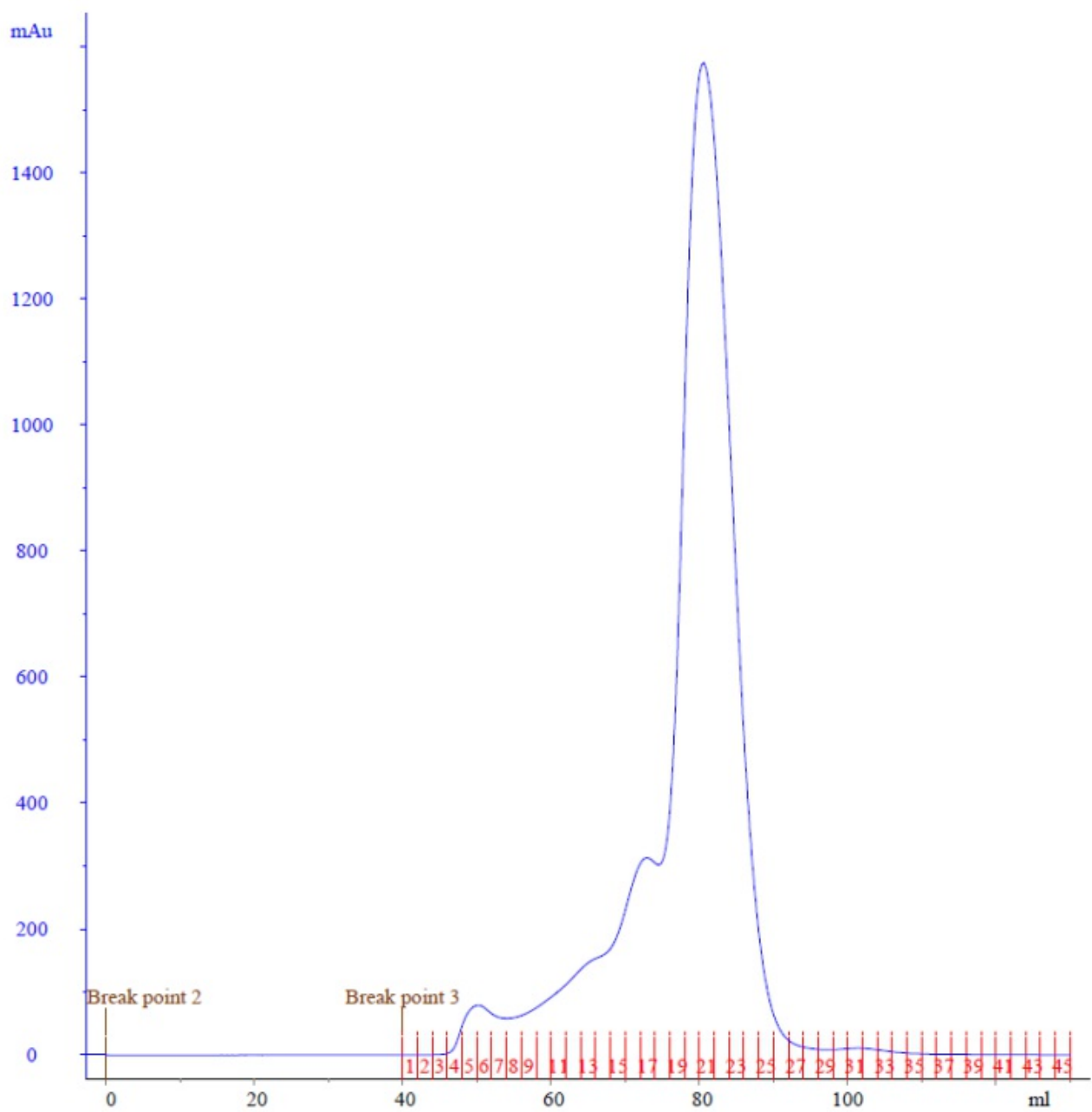
Final yields around 6 mL at 200-300 uM each except for ZV dead which is at 40 uM.

CONSTRUCT-ID	Klone ID	Expression ID	Purification Id	Notes
QQ01D2VNS2B-c001	QQ01D2VNS2B-k001	QQ01D2VNS2B-e001	QQ01D2VNS2B-p001	ALIVE
QQ01D2VNS2B-c002	QQ01D2VNS2B-k002	QQ01D2VNS2B-e002	QQ01D2VNS2B-p002	DEAD
QQ01WNVNS2B-c001	QQ01WNVNS2B-k001	QQ01WNVNS2B-e001	QQ01WNVNS2B-p001	ALIVE CLEAV
QQ01WNVNS2B-c002	QQ01WNVNS2B-k002	QQ01WNVNS2B-e002	QQ01WNVNS2B-p002	DEAD
QQ01ZVNS2B-c001	QQ01ZVNS2B-k001	QQ01ZVNS2B-e001	QQ01ZVNS2B-p001	ALIVE
QQ01ZVNS2B-c001	QQ01ZVNS2B-k002	QQ01ZVNS2B-e002	QQ01ZVNS2B-p002	DEAD

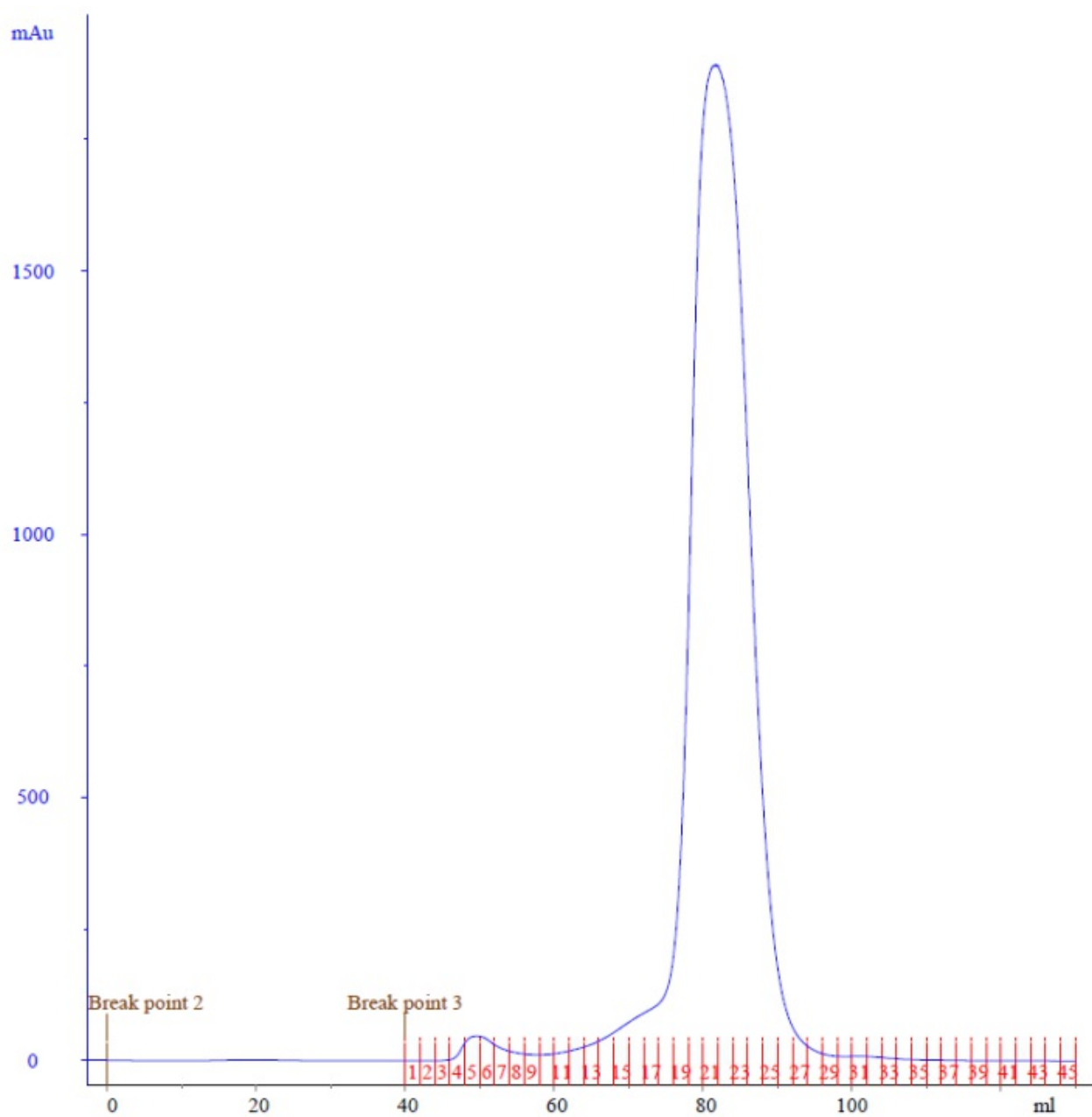
DV2_NS2B3_FP_SEC_SUPEROSE12PG



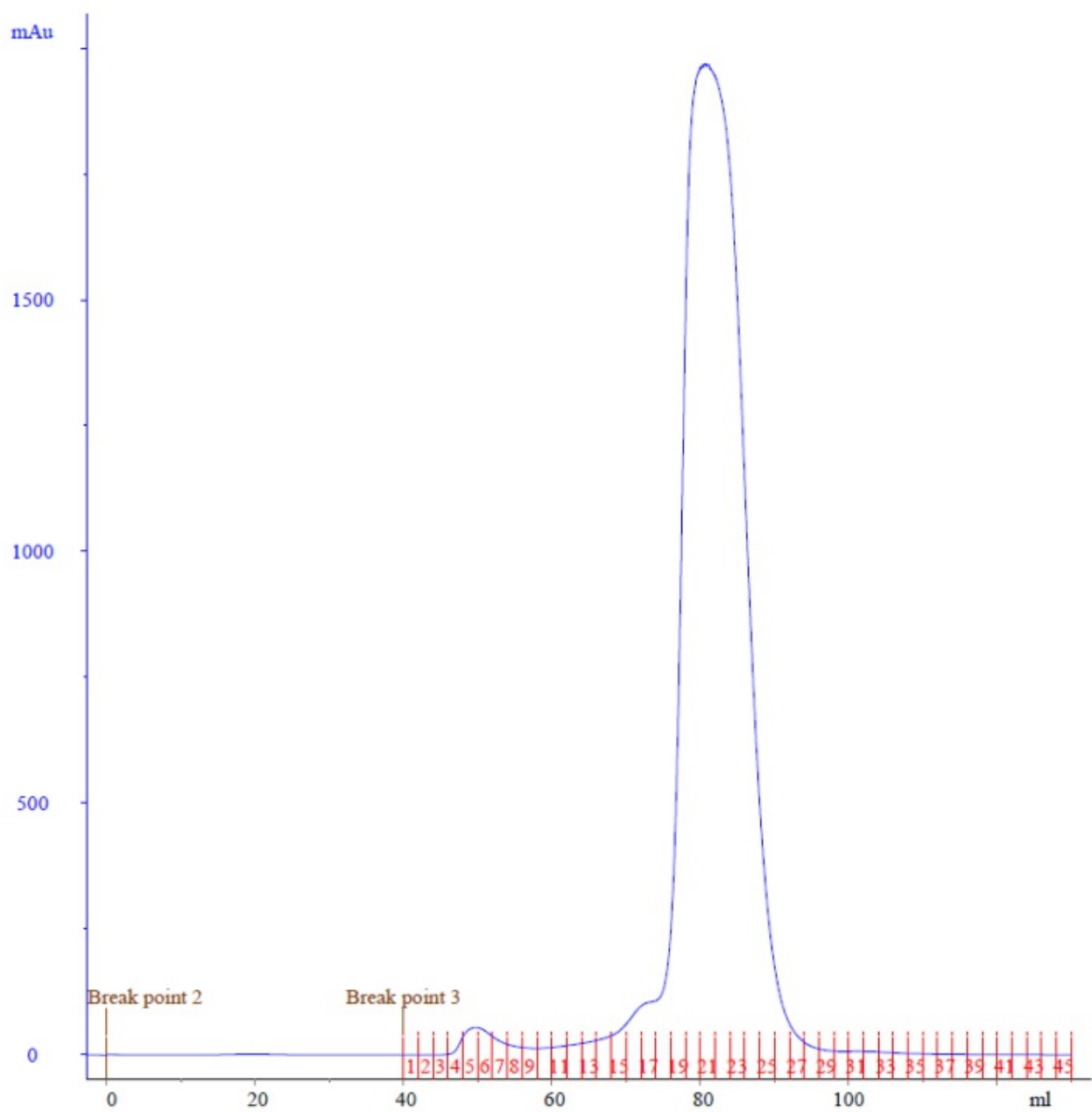
DV2_NS2B3_FP_DEAD_SEC_SUPEROSE12PG



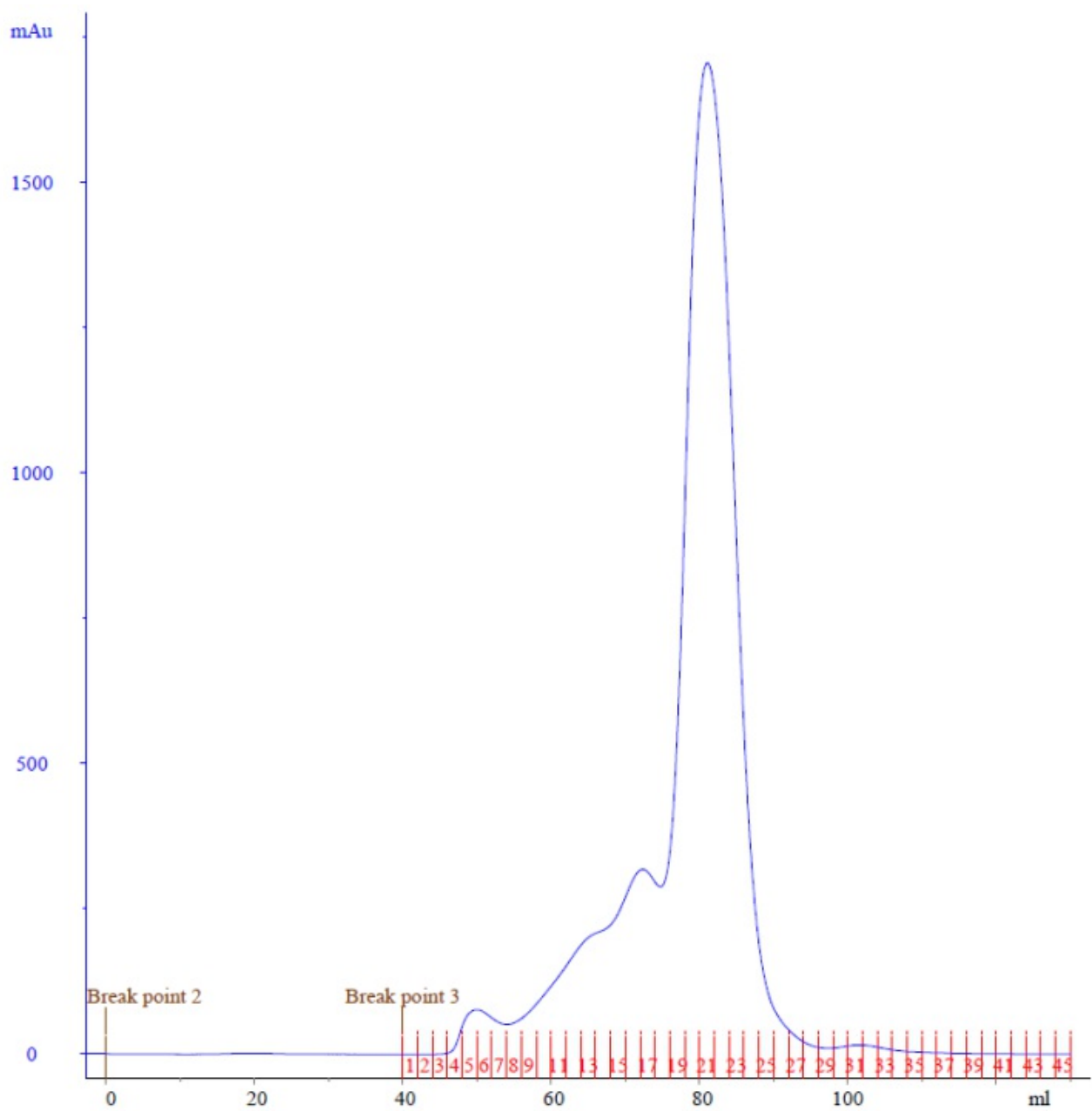
WNV_NS2B3_FP_SEC_SUPEROSE12PG



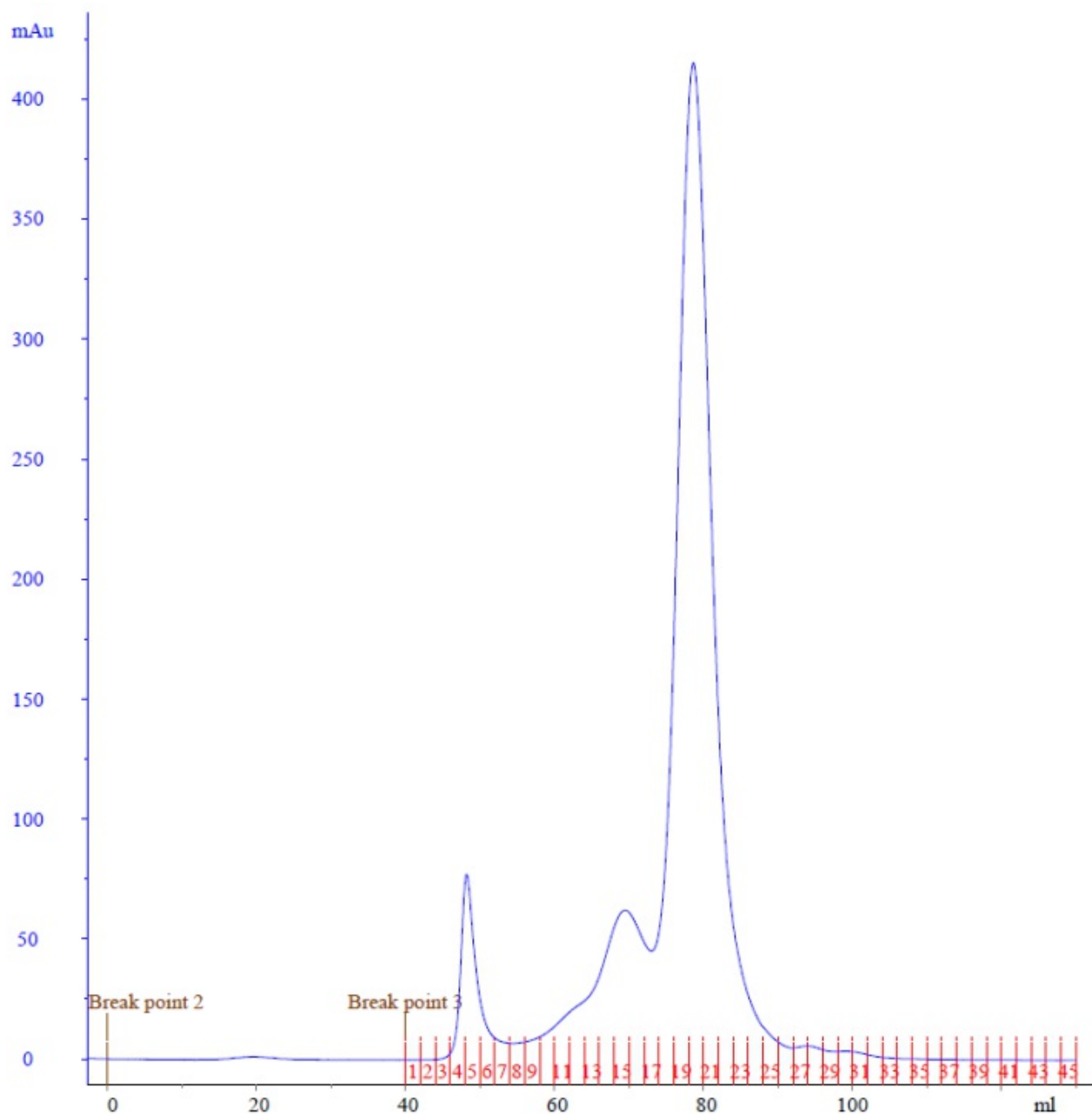
WNV_NS2B3_FP_DEAD_SEC_SUPEROSE12PG



ZV_NS2B3_FP_SEC_SUPEROSE12PG



ZV_NS2B3_FP_DEAD_SEC_SUPEROSE12PG



AKTA PRIME SEC REPORTS

WNV_NS2B3_FP_DEAD_SEC_SUPERROSE12PG.pdf
 ZV_NS2B3_FP_DEAD_SEC_SUPERROSE12PG.pdf
 WNV_NS2B3_FP_SEC_SUPERROSE12PG.pdf
 ZV_NS2B3_FP_SEC_SUPERROSE12PG.pdf
 DV2_NS2B3_FP_SEC_SUPERROSE12PG.pdf
 DV2_NS2B3_FP_DEAD_SEC_SUPERROSE12PG.pdf

Title missing - double click to edit

WNV seems to be self-cleaving linker may need redesigning

2YOL which is also a WNV fusion protein doesn't do this

comparing the sequence we can see NS2B-linker region is not the WT

>2YOL_cleaved

SMTDMWIERTADITWESDAEITGSSSERVDVRLDDDDGNFQLMNDPGAPWAGGGGSGGGGGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLG
SYQAGAGVMVEGVFHTLWHHTTKGAALMSGEGRLDPYWGSKVEDRLCYGGPWKLQHKWNGHDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGE
IGAVTLDYPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGER

A P W K G G G

but

A P W A G G G

if one looks at the cleavage sites of WNV there is a preference of K/R in P1 and G in P1'

shit

mutate to K to A to stabilize fusion

5'3' Frame 1

```
atg cac cat cat cat cat cat tct tct ggt gcg agc tgg agt cac cca caa ttc gag aaa
M H H H H H H S S G A S W S H P Q F E K
ggt ggt ggc tct ggt ggc ggt tca ggc gcc tcc gcc tgg tcc cat cca cag ttt gaa aag
G G G S G G G S G G S A W S H P Q F E K
ggt tct ggt gta gat ctg ggt acc gag aac ctg tac ttc caa tcc atg agc act gat atg
G S G V D L G T E N L Y F Q S M S T D M
tgg att gaa cgc acg gca gac atc agc tgg gag tcg gat gct gag atc aca ggt tca agc
W I E R T A D I S W E S D A E I T G S S
gaa aga gtg gat gta cga tta gac gat gac ggt aat ttt caa ctg atg aat gac cct ggc
E R V D V R L D D G N F Q L M N D P G
gcg ccg tgg aaa ggc ggc ggc ggc tct ggt gga gga ggt ggc gtt ttg tgg gac acg ccg
A P W K G G G S G G G G V L W D T P
tcc ccc aaa gaa tac aag aag ggt gac acg act acc ggt gtt tac aga att atg acc ccg
S P K E Y K K G D T T T G V Y R I M T R
ggt ttg ctt ggg tcg tac cag gcc ggg gca ggg gtg atg gta gag gga gtt ttc cat acc
G L L G S Y Q A G A G V M V E G V F H T
tta tgg cac act acc aaa ggt gcc gca ctt atg tct ggt gag ggt cgc ctt gat cct tat
L W H T T K G A A L M S G E G R L D P Y
tgg ggt agc gta aaa gag gat cgg ctc tgt tac ggc gga cca tgg aag ctt caa cac aag
W G S V K E D R L C Y G G P W K L Q H K
tgg aac gga cag gat gaa gtt cag atg att gtc gta gag ccg ggc aag aat gtc aag aac
W N G Q D E V Q M I V V E P G K N V K N
gtg caa aca aag ccc ggt gta ttc aaa acg ccg gaa ggc gaa att ggc gct gtg acc ctt
V Q T K P G V F K T P E G E I G A V T L
gac ttt cca act ggg acc agc ggc tca cca ata gtc gac aag aat gga gat gtc att ggg
D F P T G T S G S P I V D K N G D V I G
ttg tac gga aat ggt gtt att atg ccc aat ggc tcg tac atc tcc gct atc gta caa gga
L Y G N G V I M P N G S Y I S A I V Q G
gag cgc atg gac gaa ccg ata cct gca ggc ttc gaa cct gag atg tta aga aag aag taa
E R M D E P I P A G F E P E M L R K K -
```

D V R L D D D G N F Q L M N D P G A P W A G G G

GGATGTACGATTAGACGATGACGGTAATTTTCAACTGATGAATGACCCTGGCGCGCGTGGgcaGGCGGCGGCG

CCTACATGCTAATCTGCTACTGCCATTAAGTTGACTACTTACTGGGACCGCGCGGCACCCGTCCGCCGCCGCG

Top of Form

Bottom of Form

Required Primers

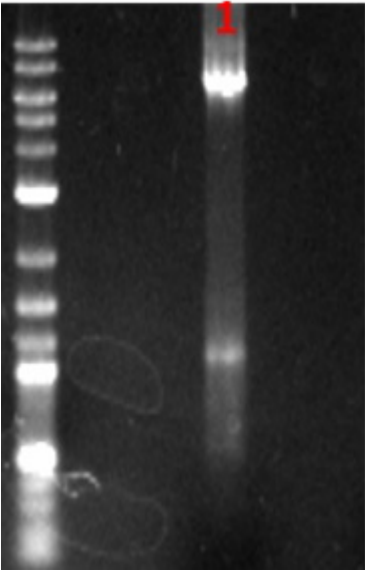
Name (F/R)	Oligo (Uppercase = target-specific primer)	Len	% GC	Tm	Ta *
WNV_K2A_FP	CGCGCCGTGGgcaGGCGGCGGCG	23	91	81°C	72°C
WNV_K2A_RP	CCAGGGTCATTTCATCAGTTGAAAATTACCGTCATCGTCTAATCGTACATCC	51	43	77°C	

* Ta (recommended annealing temperature)

```
atg cac cat cat cat cat cat tct tct ggt gcg agc tgg agt cac cca caa ttc gag aaa
M H H H H H H S S G A S W S H P Q F E K
ggt ggt ggc tct ggt ggc ggt tca ggc gcc tcc gcc tgg tcc cat cca cag ttt gaa aag
G G G S G G G S G G S A W S H P Q F E K
```

ggt tct ggt gta gat ctg ggt acc gag aac ctg tac ttc caa tcc atg agc act gat atg
G S G V D L G T E N L Y F Q S M S T D M
tgg att gaa cgc acg gca gac atc agc tgg gag tcg gat gct gag atc aca ggt tca agc
W I E R T A D I S W E S D A E I T G S S
gaa aga gtg gat gta cga tta gac gat gac ggt aat ttt caa ctg atg aat gac cct ggc
E R V D V R L D D D G N F Q L M N D P G
gcg ccg tgg gca ggc ggc ggc tct ggt gga gga ggt ggc gtt ttg tgg gac acg ccg
A P W A G G G G S G G G V L W D T P
tcc ccc aaa gaa tac aag aag ggt gac acg act acc ggt gtt tac aga att atg acc ccg
S P K E Y K K G D T T T G V Y R I M T R
ggt ttg ctt ggg tgc tac cag gcc ggg gca ggg gtg atg gta gag gga gtt ttc cat acc
G L L G S Y Q A G A G V M V E G V F H T
tta tgg cac act acc aaa ggt gcc gca ctt atg tct ggt gag ggt cgc ctt gat cct tat
L W H T T K G A A L M S G E G R L D P Y
tgg ggt agc gta aaa gag gat ccg ctc tgt tac ggc gga cca tgg aag ctt caa cac aag
W G S V K E D R L C Y G G P W K L Q H K
tgg aac gga cag gat gaa gtt cag atg att gtc gta gag ccg ggc aag aat gtc aag aac
W N G Q D E V Q M I V V E P G K N V K N
gtg caa aca aag ccc ggt gta ttc aaa acg ccg gaa ggc gaa att ggc gct gtg acc ctt
V Q T K P G V F K T P E G E I G A V T L
gac ttt cca act ggg acc agc ggc tca cca ata gtc gac aag aat gga gat gtc att ggg
D F P T G T S G S P I V D K N G D V I G
ttg tac gga aat ggt gtt att atg ccc aat ggc tgc tac atc tcc gct atc gta caa gga
L Y G N G V I M P N G S Y I S A I V Q G
gag cgc atg gac gaa ccg ata cct gca ggc ttc gaa cct gag atg tta aga aag aag taa
E R M D E P I P A G F E P E M L R K K -

INVERSE PCR Q5 25 cycles 72C AT



1 = WNV K2A

Sequencing

QQ01WNVNS2B-c003

atgcaccatcatcatcatcattcttctggtgcgagctggagtcacccacaattcgagaaaggtggtggctctggtggcggttcaggcggtccgcctggtcccatccacagtt
tgaaaagggttctggttagatctgggtaccgagaaacctgtacttccaatccatgagcactgatatgtggattgaacgcacggcagacatcagctgggagtcggatgctgaga
tcacaggttcaagcgaagagtggtgtacgattagacgatgacggtaattttcaactgatgaatgacctggcgcgccgtgggcaaggcgcgcggtctctggtggaggaggt
ggcggttttggtggacacgccgtcccccagaataacaagaaggtgacacgactaccggtgtttacagaattatgaccgggggtttgcttgggtcgtagcaggccgggacagg
ggtgatggttagagggagttttccataccttatggcacactaccaaaaggtgccgcacttatgtctggtgagggtcgccttgatccttattggggtagcgtaaaagaggtcggc
tctgttacggcgaccatggaagcttcaacacaaagtggaaacggacaggatgaagttcagatgattgtcgtagagccgggcaagaatgtcaagaacgtgcaacaaagcccggt
gtattcaaaacgccgaaggcgaaattggcgctgtgacccttgactttccaactgggaccagcggtcaccaatagtcgacaagaatggagatgtcattgggttgtagcgaaa
tggtgttattatgcccattggctcgtacatctccgctatcgtacaaggagagcgcatggacgaaccgatacctgcaggcttcgaacctgagatgttaagaaagaagtaa

MHMHHHHSSGASWSHPQFEKGGSGGSGGSAWSHPQFEKGSVDLGTENLYFQSMSTDMWIERTADISWESDAEITGSSSERVDVRLDDDGNFQLMNDPGAPWAGGGGSGGGG
GVLWDTPSPKEYKKGDTTGVYRIMTRGLLSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGFPWKLQHKWNGQDEVQMIVVEPGKNVKNVQTKPG


```

Query  361  TCCCCCAAAGAATACAAGAAGGGTGACACGACTACCGGTGTTTACAGAATTATGACCCGG  420
Sbjct  638  .....  579

Query  421  GGTTTGCTTGGGTCGTACCAGGCCGGGGCAGGGGTGATGGTAGAGGGAGTTTCCATACC  480
Sbjct  578  .....  519

Query  481  TTATGGCACACTACCAAAGGTGCCGCACTTATGTCTGGTGAGGGTCGCCTTGATCCTTAT  540
Sbjct  518  .....  459

Query  541  TGGGGTAGCGTAAAAGAGGATCGGCTCTGTTACGGCGGACCATGGAAGCTTCAACACAAG  600
Sbjct  458  .....  399

Query  601  TGGAACGGACAGGATGAAGTTCAGATGATTGTCGTAGAGCCGGGCAAGAATGTCAAGAAC  660
Sbjct  398  .....  339

Query  661  GTGCAAAACAAAGCCCGGTGTATTCAAACGCCGGAAGGCGAAATTGGCGCTGTGACCCTT  720
Sbjct  338  .....  279

Query  721  GACTTTCCAAGTGGGACCAGCGGCTCACCAATAGTCGACAAGAATGGAGATGTCATTGGG  780
Sbjct  278  .....  219

Query  781  TTGTACGGAATGGTGTTATTATGCCCAATGGCTCGTACATCTCCGCTATCGTACAAGGA  840
Sbjct  218  .....  159

Query  841  GAGCGCATGGACGAACCGATACCTGCAGGCTTCGAACCTGAGATGTTAAGAAAGAAGTAA  900
Sbjct  158  .....  99

```

MHHHHHHSSGASWSHPQFEKGGSGGGSGGSAWSHPQFEKGSVDLGTENLYFQSMSTDMWIERTADISWESDAEITGSSSERVDVRLDDDNFQLMNDPGAPWAGGGGSGGGG
GVLWDTPSPKEYKKGDTTGTGVYRIMTRGLLSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPYWGVSVKEDRLCYGGPWLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPG
VFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDIVIGLYGNVIMPNGSYISAIVQGERMDEPIIPAGFEPEMLRKK

Score	Expect	Method	Identities	Positives	Gaps
607 bits(1565)	0.0	Compositional matrix adjust.	299/299(100%)	299/299(100%)	0/299(0%)

Query	1	MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKGSVDLGTENLYFQSMSTDM	60
Sbjct	1	60
Query	61	WIERTADISWESDAEITGSSEKRVLDLDDGNFQLMNDPGAPWAGGGGSGGGGVLWDTP	120
Sbjct	61	120
Query	121	SPKEYKKGDTTGVYRIMTRGLLGSYQAGAGVMVEGVFHTLWHTTKGAALMSGEGRLDPY	180
Sbjct	121	180
Query	181	WGSVKEDRLCYGGPWKLQHKWNGQDEVQMIIVVEPGKNVKNVQTKPGVFKTPEGEIGAVTL	240
Sbjct	181	240
Query	241	DFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMDEPIPAGFEPEMLRKK	299
Sbjct	241	299

All good check to see if doesn't cut itself

<https://doi.org/10.1042/BJ20061136>

Table 2

The potential NS2B–NS3pro cleavage sites in the precursor polyprotein

The sequence of the synthesized and tested peptides which span the potential cleavage sites in the WNV and DV polyproteins is shown. The cleavage sites represent the boundaries between the individual proteins in the polyprotein precursor. The additional, dibasic, potential cleavage sites were identified via the analysis of the polyprotein peptide sequence. The efficiency of the cleavage (%) is shown in parentheses.

Proteins	WNV	DV
Capsid C protein	Q ¹⁰² KKR↓GGTA ¹⁰⁹ (62)	R ⁹⁷ RRR↓SAGM ¹⁰⁴ (64)
C/prM	C ¹²⁰ AGA↓VTLS ¹²⁷ (1)	T ¹¹¹ VMA↓FHLT ¹¹⁸ (0)
prM	H ²⁰⁹ SRR↓SRRS ²¹⁶ (6)	R ²⁰² QKR↓SVAL ²⁰⁹ (1)
prM	R ²¹² SRR↓SLTV ²¹⁹ (0)	–
prM/E	P ²⁹⁷ AYS↓FNCL ³⁰⁴ (0)	P ²⁷⁷ SMT↓MRCI ²⁸⁴ (1)
E/NS1	N ⁷⁸⁴ VHA↓DTGC ⁷⁹¹ (0)	M ⁷⁷² VQA↓DSGC ⁷⁷⁹ (2)
NS1/NS2A	R ¹¹⁴⁰ VNA↓YNAD ¹¹⁴⁷ (0)	L ¹¹²⁴ VTA↓GHGQ ¹¹³¹ (1)
NS2A	K ¹³²⁷ EKR↓SSAA ¹³³⁴ (0)	–
NS2A	A ¹³³⁴ KKK↓GACL ¹³⁴¹ (31)	–
NS2A/NS2B	N ¹³⁶⁷ RKR↓GWPA ¹³⁷⁴ (64)	S ¹³⁴² KKR↓SWPL ¹³⁴⁹ (5)
NS2B–NS3	Y ¹⁴⁹⁸ TKR↓GGVL ¹⁵⁰⁵ (73)	K ¹⁴⁷² KQR↓AGVL ¹⁴⁷⁹ (11)
NS3pro/hel	R ¹⁶⁸⁶ KKQ↓ITVL ¹⁶⁹³ (4)	R ¹⁶⁵⁹ KRR↓LTIM ¹⁶⁶⁶ (68)
NS3pro/hel	K ¹⁷⁰⁰ TRK↓ILPQ ¹⁷⁰⁷ (3)	K ¹⁶⁷⁴ TKR↓YLPA ¹⁶⁸¹ (31)
NS3pro/hel	K ¹⁷¹⁶ RLR↓TAVL ¹⁷²³ (0)	A ¹⁶⁸⁶ IKR↓GLRT ¹⁶⁹³ (6)
NS3hel	A ¹⁹⁵⁷ QRR↓GRIG ¹⁹⁶⁴ (2)	A ¹⁹³⁰ QRR↓GRIG ¹⁹³⁷ (45)
NS3/NS4A	S ²¹¹⁷ GKR↓SQIG ²¹²⁴ (0)	A ²⁰⁹⁰ GRK↓SLTL ²⁰⁹⁷ (1)
NS4A	E ²²⁴³ KQR↓SQTD ²²⁵⁰ (0)	E ²²¹⁷ KQR↓TPQD ²²²⁴ (0)
NS4A/NS4B	A ²²⁶⁶ VAA↓NEMG ²²⁷³ (0)	A ²²⁴⁰ TMA↓NEMG ²²⁴⁷ (0)
NS4B/NS5	G ²⁵²² LKR↓GGAK ²⁵²⁹ (73)	N ²⁴⁸⁸ TRR↓GTGN ²⁴⁹⁵ (1)