

De-Novo Protein Design Strategies for Targeting Rotavirus Proteins: Evolutionary Conservation and Therapeutic Implications



Noura B. Abdulrahman¹, Noor M. Ahmed¹ and Ahmed AbdulJabbar Suleiman^{2,*} 

¹Biology Department, College of Science, Tikrit University, Tikrit, Saladin, Iraq

²Biotechnology Department, College of Science, University of Anbar, Ramadi, Anbar, Iraq

© 2025 The Author(s). Published by Bentham Open.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



CrossMark

Received: June 26, 2025

Revised: August 18, 2025

Accepted: September 01, 2025

Published: November 11, 2025



Send Orders for Reprints to
reprints@benthamscience.net

*Address correspondence to this author at the Biotechnology Department, College of Science, University of Anbar, Ramadi, Anbar, Iraq; E-mail: ahmed.suleiman@uoanbar.edu.iq

Cite as: Abdulrahman N, Ahmed N, Suleiman A. De-Novo Protein Design Strategies for Targeting Rotavirus Proteins: Evolutionary Conservation and Therapeutic Implications. Open Biotechnol J, 2025; 19: e18740707424062. <http://dx.doi.org/10.2174/0118740707424062251028072626>

Multiple Sequence Alignment

The MSA analysis results of 11 proteins.



Fig. (S1). MSA result of outer capsid VP4.

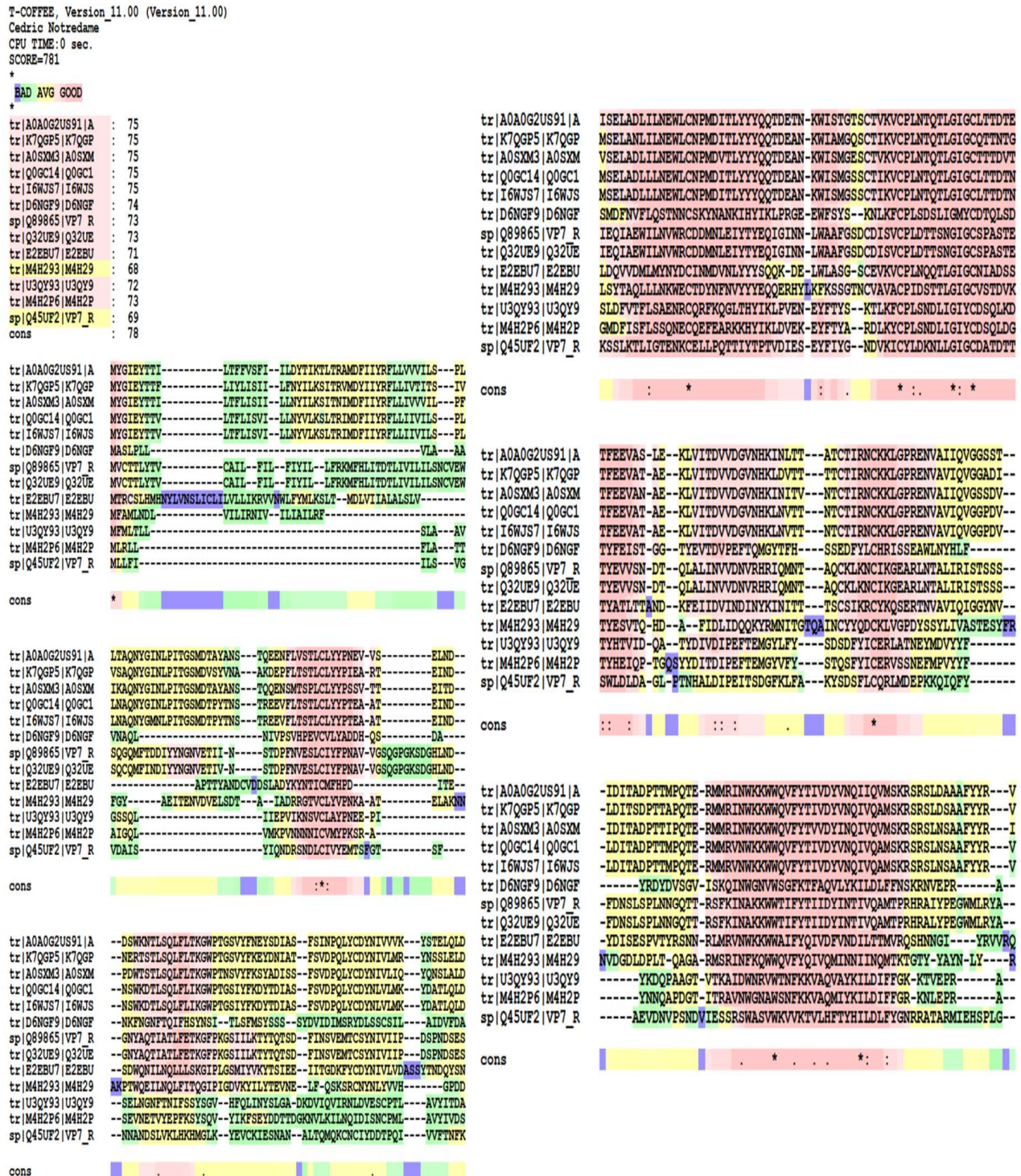


Fig. (S2). MSA result of outer capsid VP7.

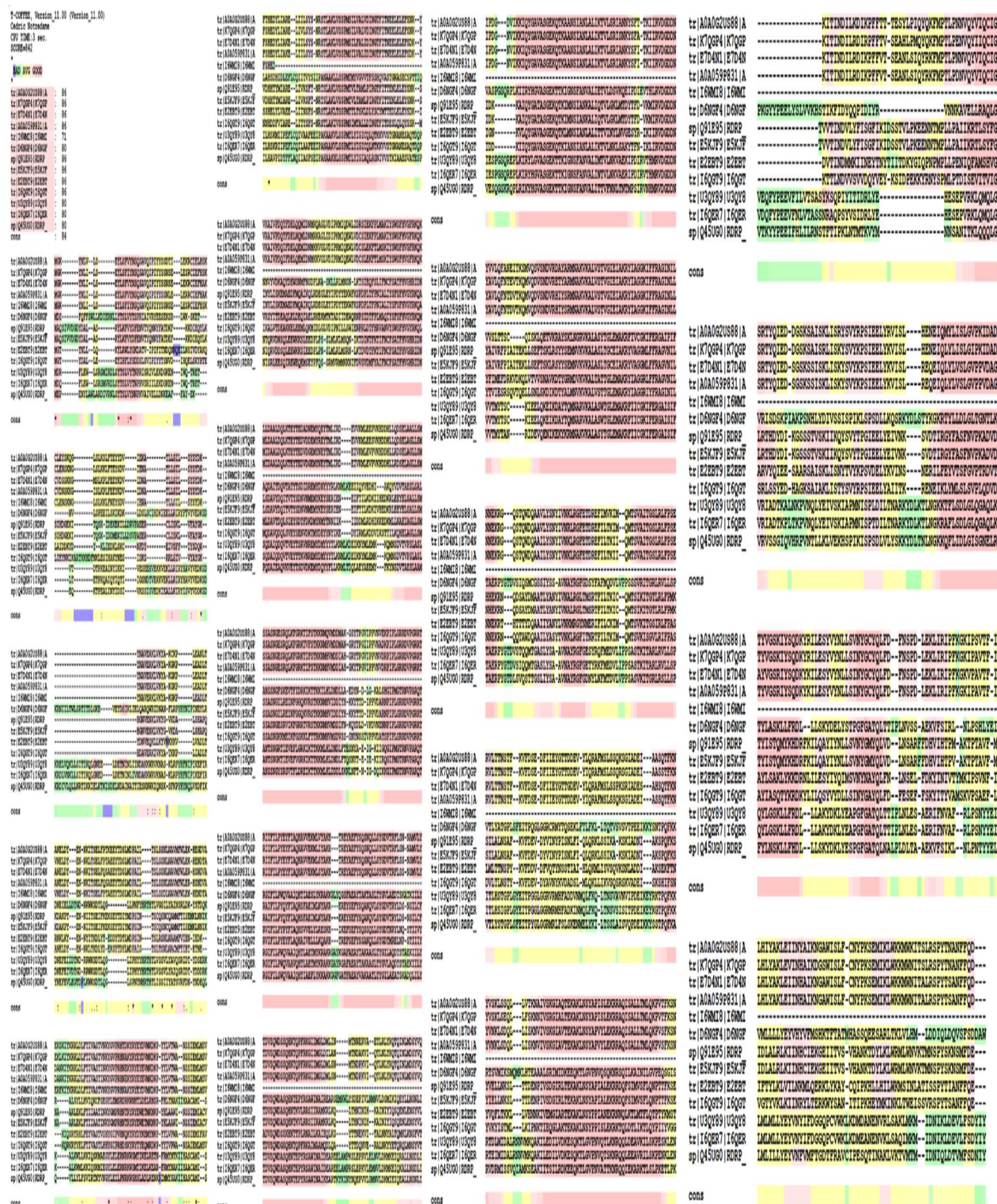


Fig. (S3). MSA result of RNA-directed RNA polymerase.

T-COFFEE, Version 11.00 (Version 11.00)
Cedric Notredame
CPU TIME: 0 sec.
SCORE=519

Bad Avg Good

tr|A0A0G2M37|A 66
tr|K7Q676|K7Q67 66
tr|E7D4N7|E7D4N 62
tr|A0A059P770|A 65
tr|I6WA25|I6WA2 62
tr|D6NGG0|D6NGG 50
tr|D6NGG1|D6NGG 40
sp|Q9PY94|NSP1 56
tr|Q32Y79|Q32Y7 55
tr|E2EBU3|E2EBU 48
tr|M4H2L7|M4H2L 57
tr|U3QY86|U3QY8 44
tr|U3R082|U3R08 40
tr|M4H298|M4H29 47
tr|M4H2L9|M4H2L 32
sp|Q45UF6|NSP1N 46
cons 51

tr|A0A0G2M37|A MAT FKDAC F-H YKRI TKIN
tr|K7Q676|K7Q67 MAT FKDAC Y-H YKRI NKIN
tr|E7D4N7|E7D4N MKSLVEMATFKDAC Y-Y YKRI NKIN
tr|A0A059P770|A MAT FKDAC Y-Y YKRI NKIN
tr|I6WA25|I6WA2 MESLVEAMATFKDAC Y-Q YKRI NKIN
tr|D6NGG0|D6NGG MCV RQSSA Q-N KSHLTHTNSQ NSN
tr|D6NGG1|D6NGG MEY RHFLS E-L SKDTLLPY ML NOSCPA
sp|Q9PY94|NSP1 MEY SFRRM LVM YKRI IDR
tr|Q32Y79|Q32Y7 MAN SFRRM LVM YKRI IDR
tr|E2EBU3|E2EBU FBU LKLVK YY LRLGLVTVS HIKSKSDYSGOMKRL
tr|M4H2L7|M4H2L FBU LKLVK YY LRLGLVTVS HIKSKSDYSGOMKRL
tr|U3QY86|U3QY8 MGS HOSSY Q-V N
tr|U3R082|U3R08 MEK TIMEISSLEADADIL K-MWNTNCPVPHED LNKQI
tr|M4H298|M4H29 MCV SNTHI Q-V SCQNTHTQS DS
tr|M4H2L9|M4H2L MVI SNEKE KFPISGTQWINSY LIND
sp|Q45UF6|NSP1N MMS LQRAA Q-I N

cons

tr|A0A0G2M37|A RELLKIGANSIWPVPSDK IKGWC ECOL
tr|K7Q676|K7Q67 SSVLKIGANDENPAPITK FKWCGL DCOGY
tr|E7D4N7|E7D4N BAVLKIGANDENPAPITK FKWCGL DCOGY
tr|A0A059P770|A BAVLKIGANDENPAPITK FKWCGL DCOGY
tr|I6WA25|I6WA2 BAVLKIGANDENPAPITK FKWCGL DCOGY
tr|D6NGG0|D6NGG FIDSKATVP
tr|D6NGG1|D6NGG QSSKPKKESIPIMONSLECPMP DSIILSDICCP
sp|Q9PY94|NSP1 KLCVNVNMRREIEXA NGICL NCINE
tr|Q32Y79|Q32Y7 D KKPMTLWMTLMPKMSKPCGNCL TQCF
tr|E2EBU3|E2EBU KTLGLVPTDDEWSETCTQCS LVTGCL SGCV
tr|M4H2L7|M4H2L NWTLIS NSNLD
tr|U3QY86|U3QY8 KEIKMNLGNKKKPEKLE NUDVM
tr|M4H298|M4H29 KLEHLNOKTALQSOTQL LSTI GALII
tr|M4H2L9|M4H2L MISPPVYVYV XQ DIADLGMIT PFNCS
sp|Q45UF6|NSP1N RQFMSLEKLNAPFLH KPGDGL DMFQNGDTPFMGCIQKASKANK

cons

tr|A0A0G2M37|A N-MCHWMSHY NEMFTICQYKNNVW HCSIEIRV
tr|K7Q676|K7Q67 INEPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|E7D4N7|E7D4N VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|A0A059P770|A VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|I6WA25|I6WA2 VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|D6NGG0|D6NGG VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|D6NGG1|D6NGG VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
sp|Q9PY94|NSP1 VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|Q32Y79|Q32Y7 VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|E2EBU3|E2EBU VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|M4H2L7|M4H2L VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|U3QY86|U3QY8 VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|U3R082|U3R08 VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|M4H298|M4H29 VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
tr|M4H2L9|M4H2L VNOPIERIV NNVLFIRALSNGAG HCSIEIRV
sp|Q45UF6|NSP1N VNOPIERIV NNVLFIRALSNGAG HCSIEIRV

tr|A0A0G2M37|A YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|K7Q676|K7Q67 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|E7D4N7|E7D4N YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|A0A059P770|A YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|I6WA25|I6WA2 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|D6NGG0|D6NGG YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|D6NGG1|D6NGG YPMVNVNMTSTYID ETLEKIPNMTVNI
sp|Q9PY94|NSP1 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|Q32Y79|Q32Y7 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|E2EBU3|E2EBU YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|M4H2L7|M4H2L YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|U3QY86|U3QY8 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|U3R082|U3R08 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|M4H298|M4H29 YPMVNVNMTSTYID ETLEKIPNMTVNI
tr|M4H2L9|M4H2L YPMVNVNMTSTYID ETLEKIPNMTVNI
sp|Q45UF6|NSP1N YPMVNVNMTSTYID ETLEKIPNMTVNI

tr|A0A0G2M37|A DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|K7Q676|K7Q67 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|E7D4N7|E7D4N DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|A0A059P770|A DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|I6WA25|I6WA2 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|D6NGG0|D6NGG DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|D6NGG1|D6NGG DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
sp|Q9PY94|NSP1 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|Q32Y79|Q32Y7 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|E2EBU3|E2EBU DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|M4H2L7|M4H2L DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|U3QY86|U3QY8 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|U3R082|U3R08 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|M4H298|M4H29 DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
tr|M4H2L9|M4H2L DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW
sp|Q45UF6|NSP1N DQRY ISFW QIDLS LYLN NM LNT KALPMLKHEFGVIGLIN QW

tr|A0A0G2M37|A CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|K7Q676|K7Q67 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|E7D4N7|E7D4N CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|A0A059P770|A CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|I6WA25|I6WA2 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|D6NGG0|D6NGG CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|D6NGG1|D6NGG CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
sp|Q9PY94|NSP1 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|Q32Y79|Q32Y7 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|E2EBU3|E2EBU CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|M4H2L7|M4H2L CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|U3QY86|U3QY8 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|U3R082|U3R08 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|M4H298|M4H29 CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
tr|M4H2L9|M4H2L CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308
sp|Q45UF6|NSP1N CDF SVYS E-LICVST DT LNMVQKSL E-RELLI308

cons

tr|A0A0G2M37|A FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|K7Q676|K7Q67 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|E7D4N7|E7D4N FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|A0A059P770|A FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|I6WA25|I6WA2 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|D6NGG0|D6NGG FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|D6NGG1|D6NGG FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
sp|Q9PY94|NSP1 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|Q32Y79|Q32Y7 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|E2EBU3|E2EBU FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H2L7|M4H2L FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|U3QY86|U3QY8 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|U3R082|U3R08 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H298|M4H29 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H2L9|M4H2L FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
sp|Q45UF6|NSP1N FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR

cons

tr|A0A0G2M37|A FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|K7Q676|K7Q67 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|E7D4N7|E7D4N FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|A0A059P770|A FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|I6WA25|I6WA2 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|D6NGG0|D6NGG FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|D6NGG1|D6NGG FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
sp|Q9PY94|NSP1 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|Q32Y79|Q32Y7 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|E2EBU3|E2EBU FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H2L7|M4H2L FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|U3QY86|U3QY8 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|U3R082|U3R08 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H298|M4H29 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H2L9|M4H2L FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
sp|Q45UF6|NSP1N FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR

cons

tr|A0A0G2M37|A FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|K7Q676|K7Q67 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|E7D4N7|E7D4N FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|A0A059P770|A FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|I6WA25|I6WA2 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|D6NGG0|D6NGG FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|D6NGG1|D6NGG FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
sp|Q9PY94|NSP1 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|Q32Y79|Q32Y7 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|E2EBU3|E2EBU FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H2L7|M4H2L FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|U3QY86|U3QY8 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|U3R082|U3R08 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H298|M4H29 FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
tr|M4H2L9|M4H2L FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR
sp|Q45UF6|NSP1N FLDG QPRLS KIRTPRAT TQKQV VIMYQMLF P1-NKINR

cons

tr|A0A0G2M37|A GYPRM VWC IDVD RLL LSVNFM SALPS
tr|K7Q676|K7Q67 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|E7D4N7|E7D4N GYPRM VWC IDVD RLL LSVNFM SALPS
tr|A0A059P770|A GYPRM VWC IDVD RLL LSVNFM SALPS
tr|I6WA25|I6WA2 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|D6NGG0|D6NGG GYPRM VWC IDVD RLL LSVNFM SALPS
tr|D6NGG1|D6NGG GYPRM VWC IDVD RLL LSVNFM SALPS
sp|Q9PY94|NSP1 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|Q32Y79|Q32Y7 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|E2EBU3|E2EBU GYPRM VWC IDVD RLL LSVNFM SALPS
tr|M4H2L7|M4H2L GYPRM VWC IDVD RLL LSVNFM SALPS
tr|U3QY86|U3QY8 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|U3R082|U3R08 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|M4H298|M4H29 GYPRM VWC IDVD RLL LSVNFM SALPS
tr|M4H2L9|M4H2L GYPRM VWC IDVD RLL LSVNFM SALPS
sp|Q45UF6|NSP1N GYPRM VWC IDVD RLL LSVNFM SALPS

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

tr|A0A0G2M37|A D G D
tr|K7Q676|K7Q67 D G D
tr|E7D4N7|E7D4N D G D
tr|A0A059P770|A D G D
tr|I6WA25|I6WA2 D G D
tr|D6NGG0|D6NGG D G D
tr|D6NGG1|D6NGG D G D
sp|Q9PY94|NSP1 D G D
tr|Q32Y79|Q32Y7 D G D
tr|E2EBU3|E2EBU D G D
tr|M4H2L7|M4H2L D G D
tr|U3QY86|U3QY8 D G D
tr|U3R082|U3R08 D G D
tr|M4H298|M4H29 D G D
tr|M4H2L9|M4H2L D G D
sp|Q45UF6|NSP1N D G D

cons

Fig. (S4). MSA result of non-structural protein 1.

T-COFFEE, Version_11.00 (Version_11.00)
 Cedric Notredame
 CPU TIME:0 sec.
 SCORE=700

*
 BAD AVG GOOD

tr|A0A0G2UNK7|A : 72
 tr|A0A0G2UNK7|A : 72
 tr|D3K8X8|D3K8X : 72
 tr|X4Y6V3|X4Y6V : 72
 tr|I6VRE5|I6VRE : 72
 tr|D6NGG2|D6NGG : 59
 sp|Q9PY93|NSP2 : 74
 tr|Q32Y77|Q32Y7 : 74
 tr|E2EBU6|E2EBU : 71
 tr|M4H1X8|M4H1X : 66
 tr|U3QY07|U3QY0 : 58
 tr|M4H1Y4|M4H1Y : 57
 sp|Q45UF4|NSP2 : 62
 cons : 70

tr|A0A0G2UNK7|A MAELACFCYPHL---ENDSYKFIFFNSL-AIKCMLTAKVDKQDQKFFN-----SI--
 tr|A0A0G2UNK7|A MAELACFCYPHL---ENDSYKFIFFNSL-AIKCMLTAKVDKQDQKFFN-----SI--
 tr|D3K8X8|D3K8X MAELACFCYPHL---ENDSYRFIFFNSL-AIKCMLTAKVDKQDQKFFN-----SI--
 tr|X4Y6V3|X4Y6V MAELACFCYPHL---ENDSYKFIFFNSL-AIKCMLTAKVDKQDQKFFN-----SI--
 tr|I6VRE5|I6VRE MAELACFCYPHL---ENDSYKFIFFNSL-AIKCMLTAKVDKQDQKFFN-----SI--
 tr|D6NGG2|D6NGG MTQSVLSLSD-FI---VKTEGDMPSDRECVLDRLSKQKELRETFKDGKNDRSALRIKMFEL
 sp|Q9PY93|NSP2 MAELACFVSLSL---TEKGVWFFPINKK-AVQMLCAKVEKQDSNYD-----TI--
 tr|Q32Y77|Q32Y7 MAELACFVSLSL---TEKGVWFFPINKK-AVQMLCAKVEKQDSNYD-----TI--
 tr|E2EBU6|E2EBU MAELGCFVNVVE---SDSTFTFIPKSK-AINILLIT---KNDNDNPID-----TI--
 tr|M4H1X8|M4H1X MAELGCFVNVVE---SDSTFTFIPKSK-AINILLIT---KNDNDNPID-----TI--
 tr|U3QY07|U3QY0 MTQSVSIAD-FI---VKTEGDMPSDRECVLDRLSKQKELRETFKDGKNDRSALRIKMFEL
 tr|M4H1Y4|M4H1Y MTQSVSIAD-FI---VKTEGDMPSDRECVLDRLSKQKELRETFKDGKNDRSALRIKMFEL
 sp|Q45UF4|NSP2 MVSISVLADFI---VKTEEGWIPSDNC-PALDRFKTKTEKELLSIKKEGADRASIRKQLEL

cons *..

tr|A0A0G2UNK7|A -----IYGIAPPQPKKRYNTNDSRGMNYETDMFKVATLICEALNSV-K---VTQ
 tr|A0A0G2UNK7|A -----IYGIAPPQPKKRYNTNDSRGMNYETDMFKVATLICEALNSV-K---VTQ
 tr|D3K8X8|D3K8X -----IYGIAPPQPKKRYNTNDSRGMNYETDMFKVATLICEALNSI-K---VTQ
 tr|X4Y6V3|X4Y6V -----VYGIAPPQPKKRYNTNDSRGMNYETDMFKVATLICEALNSI-K---VTQ
 tr|I6VRE5|I6VRE -----IYGIAPPQPKKRYNTNDSRGMNYETDMFKVATLICEALNSI-K---VTQ
 tr|D6NGG2|D6NGG SEPSRSRRFTQHGVPVPMREIKTN-----TDIPSTLWLTVDWLLNLQDE-ENQEMFD
 sp|Q9PY93|NSP2 -----LYGVAPPEFERNRFTKNE-RYGLDYSEQYTELANLADVLAMV-S---MPT
 tr|Q32Y77|Q32Y7 -----LYGVAPPEFERNRFTKNE-RYGLDYSEQYTELANLADVLAMV-S---MPT
 tr|E2EBU6|E2EBU -----IYGLTPVEKYRRFASDMSPSGINDNELFDKVAIKLAEILNFGFKK-MSV
 tr|M4H1X8|M4H1X -----IYGTPTPRKHLRRFRTKRTNKGSGFHDNDVIGCCQMLATVMTA-H---LKG
 tr|U3QY07|U3QY0 INAPSRRRFTQEGVPMKELSK-----TDIPSTLWLTVDWLLNLQDE-ENQEMFD
 tr|M4H1Y4|M4H1Y LAAPSRRRFTQEGVPMKELSK-----TDIPSTLWLTVDWLLNLQDE-ENQEMFD
 sp|Q45UF4|NSP2 TSISNKRLLQLGVPVPRDITRS-----TTIPSSRNLITDMLNIPNDE-ESGEVES

cons ..*

tr|A0A0G2UNK7|A SD-IANVLSRVSVRHLNVLIRKENYQDVLPHS-KELLIKSVLIAIGHIKEIETTATAEGGE
 tr|A0A0G2UNK7|A SD-IANVLSRVSVRHLNVLIRKENYQDVLPHS-KELLIKSVLIAIGHIKEIETTATAEGGE
 tr|D3K8X8|D3K8X SD-VASVLSKVVSVRHLNVLIRKENYQDVLPHS-KELLIKSVLIAIGHIKEIETTATAEGGE
 tr|X4Y6V3|X4Y6V SE-VANVLSRVSVRHLNVLIRKENYQDVLPHS-KELLIKSVLIAIGHIKEIETTATAEGGE
 tr|I6VRE5|I6VRE AN-VSNVLSRVSVRHLNVLIRKENYQDVLPHS-KOLLKSTLIAIGHIKEIETTATAEGGE
 tr|D6NGG2|D6NGG FI-SSKFPDVLASADKLARFAQLREDRKQVLIHQN-FSKAMNAFACF---NAIKPFTATEGKC
 sp|Q9PY93|NSP2 EKFQDFIVKTVQVRHLNLLCRIDKQVNDILNAN-VKLRVKAVMIACNLVNETETPTLESND
 tr|Q32Y77|Q32Y7 EKFQDFIVKTVQVRHLNLLCRIDKQVNDILNAN-VKLRVKAVMIACNLVNETETPTLESND
 tr|E2EBU6|E2EBU QQ-AMMLTRVSVRHLNLLCRIDKQVNDILNAN-VKLRVKAVMIACNLVNETETPTLESND
 tr|M4H1X8|M4H1X EQ-AKKLLNSVSVRHLNLLCRIDKQVNDILNAN-VKLRVKAVMIACNLVNETETPTLESND
 tr|U3QY07|U3QY0 VE-ENKFPDIFASSDKISRFAMLNKKDILHQN-MSKAMNAFACF---NAIKPFTATEGKC
 tr|M4H1Y4|M4H1Y VE-ENKFPDIFASSDKISRFAMLNKKDILHQN-MSKAMNAFACF---NAIKPFTATEGKC
 sp|Q45UF4|NSP2 AI-ASKYDIPFCISADKISRFAMLNKKDILHQN-MSKAMNAFACF---NAIKPFTATEGKC

cons ..*

tr|A0A0G2UNK7|A IV-----FQNVAFITMMKLTYLEHQMLPILDQNFVEYKVTLNEDK-PISESHIKEL
 tr|A0A0G2UNK7|A IV-----FQNVAFITMMKLTYLEHQMLPILDQNFVEYKVTLNEDK-PISESHIKEL
 tr|D3K8X8|D3K8X VV-----FQNAAFITMMKLTYLEHRLPILDQNFIEYKITVNEDK-PISESHVKEI
 tr|X4Y6V3|X4Y6V IV-----FQNAAFITMMKLTYLEHRLPILDQNFIEYKITVNEDK-PISDVCVKEI
 tr|I6VRE5|I6VRE IV-----FQNAAFITMMKLTYLEHRLPILDQNFIEYKITVNEDK-PISDGHVKEL
 tr|D6NGG2|D6NGG NVVRATDDSMILEFPQVPEYFRGSKATFYKLYPLSD-----EQ-P-----VNG
 sp|Q9PY93|NSP2 IV-----YQDSYFTITKLDYSNHKLLPLMA---DEYKITINTKT-DIPDRNQTA
 tr|Q32Y77|Q32Y7 IV-----YQDSYFTITKLDYSNHKLLPLMA---DEYKITINTKT-DIPDRNQTA
 tr|E2EBU6|E2EBU LL-----YVDKTYALNEDLYRKYKLMFVKA---VEYKITLNSAT-EDADS-ERKM
 tr|M4H1X8|M4H1X IE-----YMKYFTITKLDYSNHKLLPLMA---IEYKITFNSDSNDIPDDAFKKL
 tr|U3QY07|U3QY0 TVARACEDSIILEFPVPIEHLRIGQARGVYKLYPLSD-----DL-P-----TQG
 tr|M4H1Y4|M4H1Y TVARACEDSIILEFPVPIEHLRIGQARGVYKLYPLSD-----DL-P-----TQG
 sp|Q45UF4|NSP2 EIVRATEDAIKAFEPVSEHLICIGNRPGVYKAFPIKK-----EQ-P-----MVY

cons ..*

tr|A0A0G2UNK7|A VAELEWQYNKFAAITHGKGHYRIVKYSSVANHADRVYATFKSNSKNSMLSEFNLLDQRIWQN
 tr|A0A0G2UNK7|A VAELEWQYNKFAAITHGKGHYRIVKYSSVANHADRVYATFKSNSKNSMLSEFNLLDQRIWQN
 tr|D3K8X8|D3K8X IAELEWQYNKFAAITHGKGHYRIVKYSSVANHADRVYATFKSNSKNSMLSEFNLLDQRIWQN
 tr|X4Y6V3|X4Y6V VAELEWQYNKFAAITHGKGHYRIVKYSSVANHADRVYATFKSNSKNSMLSEFNLLDQRIWQN
 tr|I6VRE5|I6VRE VAELEWQYNKFAAITHGKGHYRIVKYSSVANHADRVYATFKSNSKNSMLSEFNLLDQRIWQN
 tr|D6NGG2|D6NGG MLAKAVAGNQFFMYHGHGHRTVPYHELAD-AIKSFARKDKETLESISK-----S
 sp|Q9PY93|NSP2 AAYIRYFNKFAAISHGKRHWRLVLSQMSHAERLDRKIKSKDKHGRQFSYDDGDMFVHPG
 tr|Q32Y77|Q32Y7 AAYIRYFNKFAAISHGKRHWRLVLSQMSHAERLDRKIKSKDKHGRQFSYDDGDMFVHPG
 tr|E2EBU6|E2EBU AAYIRYFNKFAAISHGKRHWRLVLSQMSHAERLDRKIKSKDKHGRQFSYDDGDMFVHPG
 tr|M4H1X8|M4H1X GGYIKFYNKFAAISHGKRHWRLVLSQMSHAERLDRKIKSKDKHGRQFSYDDGDMFVHPG
 tr|U3QY07|U3QY0 FLALKHVSNNQFQMYHGHGHRTVPYSEIPD-AVRSFAKKQKDELEKITK-----D
 tr|M4H1Y4|M4H1Y FLALKHVSNNQFQMYHGHGHRTVPYSEIPD-AVRSFAKKQKDELEKITK-----D
 sp|Q45UF4|NSP2 VKALLGISNRDFIMNHGHGHRTVPYSEINN-AVRSFAKKNEAEIKRIRS-----D

cons : .. : ** * * * . . * :

tr|A0A0G2UNK7|A WYAFSTSMKQGNS-LDICKNT---FFKMKQDKSSFKGLSTDRKMDVSVQV-G
 tr|A0A0G2UNK7|A WYAFSTSMKQGNS-LDICKNT---FFKMKQDKSSFKGLSTDRKMDVSVQV-G
 tr|D3K8X8|D3K8X WYAFSTSMKQGNT-LETCKKL-LFQKMKRESNPFKGLSTDRKMDVSVQIG-I
 tr|X4Y6V3|X4Y6V WYAFSTSMKQGNT-IDVCKKL-LFQKMKQEKNFKGLSTDRKMDVSVHVG-I
 tr|I6VRE5|I6VRE WYAFSTSMKQGNT-LDVCKKL-LFQKMKQEKNFKGLSTDRKMDVSVQV-G
 tr|D6NGG2|D6NGG PLA---AQCGSKFLDMLDGI---RSQKQIEDVLKAKIFEKRS-----
 sp|Q9PY93|NSP2 WKTICIGQLCGTT-FEVAKTS---LYSIKPSK-TVTRATNKIESDLISMIG-N
 tr|Q32Y77|Q32Y7 WKTICIGQLCGTT-FEVAKTS---LYRIKPSK-TVTRATNKIESDLISMIG-N
 tr|E2EBU6|E2EBU WQDFMNAVEYSGKS-MSECRNE-LFKSSKTSN-PVKHVNILQDEMSTV--N
 tr|M4H1X8|M4H1X WMEFMAIESGMD-VQKAKDQCLFKRVQFTK-EVKAHARELAISSMSVINGN
 tr|U3QY07|U3QY0 QLA---VQCGQKFIKMIDDL---RAGRKIEEIVSDVMKFDKK-Q-----
 tr|M4H1Y4|M4H1Y QLA---VQCGQKFIKMIDDL---RAGRKIEEIVSDVMKFDKK-Q-----
 sp|Q45UF4|NSP2 SLS---PNAGEKFINMCDML---LQKEKIEVTIAKIMKSDKN-----

cons * .. :

Fig. (S5). MSA result of non-structural protein 2.

tr>CPFF2, Version 11.00 (Version 11.00)
 Cedric Notredame
 CPU TIME: 0 sec.
 SCORE: -697

BAD AVG GOOD

tr>ADA0G02UL3(A) : 75
 tr>KIQG60 KIQG6 : 75
 tr>E7D4P0 E7D4P : 75
 tr>X4YJ76 X4YJ7 : 75
 tr>I6M0J4 I6M0J : 75
 tr>D6NG63 D6NG6 : 67
 sp>Q9P195 NSP3 : 74
 tr>Q29088 Q2908 : 74
 tr>E2E805 E2E80 : 74
 tr>M4H254 M4H25 : 70
 tr>U3R2P1 U3R2P : 66
 tr>M4H258 M4H25 : 66
 sp>Q5Y19A9 NSP3 : 66
 cons : 69

tr>ADA0G02UL3(A) MES-ITQL
 tr>KIQG60 KIQG6 M4H254ITQL
 tr>E7D4P0 E7D4P MES-ITQL
 tr>X4YJ76 X4YJ7 MES-ITQL
 tr>I6M0J4 I6M0J M4H254ITQL
 tr>D6NG63 D6NG6 M4L--D-LI
 sp>Q9P195 NSP3 MAT--Q-AS
 tr>Q29088 Q2908 MAT--Q-AS
 tr>E2E805 E2E80 MQ--FRK--NTL--CVLLIPDTSAPFLPQRKNNKLLIKTVLAKRQKSSSYLPDASH
 tr>M4H254 M4H25 MAT--E-AS
 tr>U3R2P1 U3R2P M4B--L-LK
 tr>M4H258 M4H25 M4B--L-LK
 sp>Q5Y19A9 NSP3 M4B--L-LK
 cons

tr>ADA0G02UL3(A) ANSI-MN
 tr>KIQG60 KIQG6 ANSI-IN
 tr>E7D4P0 E7D4P VSSI-IN
 tr>X4YJ76 X4YJ7 VSSI-IN
 tr>I6M0J4 I6M0J ANSL-ID
 tr>D6NG63 D6NG6 AS--
 sp>Q9P195 NSP3 VENT-FN
 tr>Q29088 Q2908 VENT-FN
 tr>E2E805 E2E80 YERQDYLKQDQVHEHTEIQDQFES--RUSLEIKETSLOQDPPKQDQNSGSSGS
 tr>M4H254 M4H25 M4B-LV-LK
 tr>U3R2P1 U3R2P VC--
 tr>M4H258 M4H25 VC--
 sp>Q5Y19A9 NSP3 DA--
 cons

tr>ADA0G02UL3(A) SPEA--AAATSTLEMGIDQYD-EITVRK--SKDPVMDSGVNNLIGKAITDQ
 tr>KIQG60 KIQG6 SPEA--AAATSTLEMGIDQYD-EVITVRK--SKDPVMDSGVNNLIGKAITDQ
 tr>E7D4P0 E7D4P SPEA--AAATSTLEMGIDQYD-EVITVRK--SKDPVMDSGVNNLIGKAITDQ
 tr>X4YJ76 X4YJ7 SPEA--AAATSTLEMGIDQYD-EVITVRK--SKDPVMDSGVNNLIGKAITDQ
 tr>I6M0J4 I6M0J SPEA--AAATSALEMGIDQYD-EITVRK--SKDPVMDSGVNNLIGKAITDQ
 tr>D6NG63 D6NG6 TLTPVLKQGVN-P-ISRIT--TKPTEALDDQGMVDWRA
 sp>Q9P195 NSP3 AGSA--ASSSLKADKADGGSNFSKVLTKRY--DNVQCIDDSGVNACIGKARTDK
 tr>Q29088 Q2908 AGSA--ASSSLKADKADGGSNFSKVLTKRY--DNVQCIDDSGVNACIGKARTDK
 tr>E2E805 E2E80 NLEALMPKGLCKRLKQPPVQK-SIHDDA--MKPSMLKSSGSVPSKQVQVDE
 tr>M4H254 M4H25 TQCE--V--VQLLTGKLVG--BAR--RRLVLDVDTGKALICKARTDKH
 tr>U3R2P1 U3R2P E--AB-M-KSNGSPSPERKODAGINDAWDA
 tr>M4H258 M4H25 DE-M-KSNGSPSPERKODAGINDAWDA
 sp>Q5Y19A9 NSP3 LATUTNITGN-DESARFC--RMPRTDQSLGIVTNR--A
 cons

tr>ADA0G02UL3(A) IRQO--EYINLIQY
 tr>KIQG60 KIQG6 SQOO--SQADLQY
 tr>E7D4P0 E7D4P SQOO--NQADLQY
 tr>X4YJ76 X4YJ7 SQOO--NQADLQY
 tr>I6M0J4 I6M0J SQOO--NQADLQY
 tr>D6NG63 D6NG6 EREK--EYVNFMLCKSNKRLKQKQCEBZDRLALV
 sp>Q9P195 NSP3 NE--SQCLPMQY
 tr>Q29088 Q2908 SE--SQCPWQYV
 tr>E2E805 E2E80 DQKO--ELISKLRG
 tr>M4H254 M4H25 DTKRDSRLKMRQLSTVYKJLZKDRMLREERS
 tr>U3R2P1 U3R2P NMCK--EYVIMQNEFEM--SS
 tr>M4H258 M4H25 KQCE--RMSRLQENEM--SS
 sp>Q5Y19A9 NSP3 KTOR--DEATCLGL
 cons

tr>ADA0G02UL3(A) MELP-HEIKDITQDQNS-IDQNFIND--DL--ELITA
 tr>KIQG60 KIQG6 MELP-HEIKDITQDQNS-IDQNFIND--DL--ESLIR
 tr>E7D4P0 E7D4P MELP-HEIKDITQDQNS-IDQNFIND--DL--ESLIR
 tr>X4YJ76 X4YJ7 MELP-HEIKDITQDQNS-IDQNFIND--DL--ESLIR
 tr>I6M0J4 I6M0J MELP-HEIKDITQDQNS-IDQNFIND--DL--ESLIR
 tr>D6NG63 D6NG6 SQOS-DEIKDITQDQNS-IDQNFIND--DL--ELIR
 sp>Q9P195 NSP3 NSSP-OTQDQNSH--K--TSTNCHACICCTKEDMDCTGRKQKQKQVPLVA
 tr>Q29088 Q2908 DESP-OTQDQNSH--K--TSTNCHACICCTKEDMDCTGRKQKQKQVPLVA
 tr>E2E805 E2E80 RQIS-XEL-SAGQVIT-LSTIK-OTPS--EL--EERVT
 tr>U3R2P1 U3R2P VIIE-SQNDQNTDIN-DIN--DQDQS--
 tr>M4H258 M4H25 RQIS--YSSDLSGL--DQDQ--QDQ--ES--
 sp>Q5Y19A9 NSP3 KKKRGRKRLSTVSGALP-GLPQGLSL--DY--
 cons

tr>ADA0G02UL3(A) HILSYDRTY
 tr>KIQG60 KIQG6 HILSYDRTY
 tr>E7D4P0 E7D4P HILSYDRTY
 tr>X4YJ76 X4YJ7 HILSYDRTY
 tr>I6M0J4 I6M0J HILSYDRTY
 tr>D6NG63 D6NG6 HAKDERPHDEKTLNATDIL
 sp>Q9P195 NSP3 NSKQPTIKLISLSDILAMPDILTY--HDTYR--VQLK
 tr>Q29088 Q2908 NSKQPTIKLISLSDILAMPDILTY--HDTYR--VQLK
 tr>E2E805 E2E80 HLEK--II--AGLA--
 tr>M4H254 M4H25 --GX--FKTIT
 tr>U3R2P1 U3R2P RNEI--TDLD--DLKQYQDQ--HEHRLAKLQDQDQNSQDQ
 tr>M4H258 M4H25 DEET--FDLD--DLKQYQDQ--HEHRLAKLQDQDQNSQDQ
 sp>Q5Y19A9 NSP3
 cons

tr>ADA0G02UL3(A) CHYETTY
 tr>KIQG60 KIQG6 CHYETTY
 tr>E7D4P0 E7D4P CHYETAC
 tr>X4YJ76 X4YJ7 CHYETTY
 tr>I6M0J4 I6M0J CHYETAC
 tr>D6NG63 D6NG6 CHYETAC
 sp>Q9P195 NSP3 CTDTFLDQDQSTVYCYHKLSS--KALQSGSKDQKQKQYDYE--LQGLQ
 tr>Q29088 Q2908 CTDTFLDQDQSTVYCYHKLSS--KALQSGSKDQKQKQYDYE--LQGLQ
 tr>E2E805 E2E80 COLVTEY
 tr>M4H254 M4H25 COLVTEY
 tr>U3R2P1 U3R2P YAEVLYG
 tr>M4H258 M4H25 NEYAYNI
 sp>Q5Y19A9 NSP3
 cons

tr>ADA0G02UL3(A) CHYETTY
 tr>KIQG60 KIQG6 CHYETTY
 tr>E7D4P0 E7D4P CHYETAC
 tr>X4YJ76 X4YJ7 CHYETTY
 tr>I6M0J4 I6M0J CHYETAC
 tr>D6NG63 D6NG6 CHYETAC
 sp>Q9P195 NSP3 CTDTFLDQDQSTVYCYHKLSS--KALQSGSKDQKQKQYDYE--LQGLQ
 tr>Q29088 Q2908 CTDTFLDQDQSTVYCYHKLSS--KALQSGSKDQKQKQYDYE--LQGLQ
 tr>E2E805 E2E80 COLVTEY
 tr>M4H254 M4H25 COLVTEY
 tr>U3R2P1 U3R2P YAEVLYG
 tr>M4H258 M4H25 NEYAYNI
 sp>Q5Y19A9 NSP3
 cons

T-COFFEE, Version_11.00 (Version_11.00)

Cedric Notredame

CPU TIME:0 sec.

SCORE=528

*

BAD AVG GOOD

*

```

tr|A0A0G2U594|A : 57
tr|K7QGP6|K7QGP : 58
tr|E7D4P3|E7D4P : 58
tr|Q8V9B7|Q8V9B : 58
tr|I6WJ78|I6WJ7 : 58
tr|Q6XD85|Q6XD8 : 29
sp|Q00682|NSP5 : 49
tr|Q32Y78|Q32Y7 : 49
tr|E2EBV0|E2EBV : 38
tr|M4H296|M4H29 : 51
tr|U3R085|U3R08 : 35
tr|M4H2A1|M4H2A : 38
sp|Q45UF0|NSP5 : 28
cons : 52

```

```

tr|A0A0G2U594|A MSL--TEFSLPS-----ISSSVYKNESN
tr|K7QGP6|K7QGP MSLSDVTSPLS-----ISSSIYKNESN
tr|E7D4P3|E7D4P MSLSDVTSPLS-----ISSSIYKNESN
tr|Q8V9B7|Q8V9B MSLSDVTSPLS-----ISSSIYKNESN
tr|I6WJ78|I6WJ7 MSLSDVTSPLS-----ISSSIYKNESN
tr|Q6XD85|Q6XD8 MAEASEP-NFTTKRAQIM-SD--RRTRDTRQKKIEEKSD--VDLVDSASVYSQESA
sp|Q00682|NSP5 MSDT-GI-NLDA-----ICDNVKKGQTE
tr|Q32Y78|Q32Y7 MSDT-GI-NLDA-----ICDNVKKGQTE
tr|E2EBV0|E2EBV MMDLDF-NFESNLPEISL-IS-----
tr|M4H296|M4H29 MSMOLDI-DLANCVIDSSSIIG-----G-----SN-TGSRLPPTQAAS
tr|U3R085|U3R08 MAEVSEF-DFNIKKTKKKI-DTTTRQKNFKNFNDTATEITN--DNLIETKSNYSESEF
tr|M4H2A1|M4H2A MAEVSEF-DFNIKKTKKKI-EK---TKSKRMVKDNETVWTH--EERSEKSGVSESEF
sp|Q45UF0|NSP5 MSEVPRF-ELRSKKIKGK-QK---VD---IPGDKDDESMQIDCETDSLISESVSSHTSY

```

```

cons * . . . . .

```

```

tr|A0A0G2U594|A S-ATSTLSGKSGRSEQYISPDADFKNKMLSKSPEDIGPSDASNDPLTS-PSIRSNVAKT-
tr|K7QGP6|K7QGP S-TTSTLSGKSGRSEQYISPDADFKNKMLSKSPEDIGPSDASNDPLTS-PSIRSNVAKT-
tr|E7D4P3|E7D4P S-TTSTLSGKSGRSEQYISPDADFKNKMLSKSPEDIGPSDASNDPLTS-PSIRSNVAKT-
tr|Q8V9B7|Q8V9B S-TTSTLSGKSGRSEQYISPDADFKNKMLSKSPEDIGPSDASNDPLTS-PSIRSNVAKT-
tr|I6WJ78|I6WJ7 S-TTSTLSGKSGRSEQYISPDADFKNKMLSKSPEDIGPSDASNDPLTS-PSIRSNVAKT-
tr|Q6XD85|Q6XD8 R---SNY-----SDAIDKLKR-EP---IV--EESNDARYNFEFSEDEEV-
sp|Q00682|NSP5 SRTGSQLSNRS-SRRMDFV--DDEELSTYFNASKA--SVTQSDSCSNDLEIK-HSIITEAVVC
tr|Q32Y78|Q32Y7 SRTGSQLSNRS-SRRMDFV--DDEELSTYFNASKA--SVTQSDSCSNDLEIK-HSIITEAVVC
tr|E2EBV0|E2EBV ---SRAGTTY-----TKIDYDEDML-L-DDITPSDASSQDTNQ-RTFREKSPKS
tr|M4H296|M4H29 Y-TTTLAAGLEEAENERR--KAIEYSKMLEK--QDLGPNDASNDGME-WSVSSRSFST-
tr|U3R085|U3R08 H-TWTSY-----SDAINDLKK-EL---ES--DESNDIKCK-KTIENWAE--
tr|M4H2A1|M4H2A S-HSSSNY-----ABAYERLQR-EL---NA--SESDNCKCK-RTIRNWADE--
sp|Q45UF0|NSP5 ---EDY-----SKAIEKLT-L-ETPADVNSASTIVDSVCE-ESWYDKTIKDE

```

```

cons . . . . .

```

```

tr|A0A0G2U594|A -NADAGVSM-----S-----STQSRPS--SNVGCQDQDFSGKAININANDSSI
tr|K7QGP6|K7QGP -NADAGVSM-----S-----STQSRPS--SNVGCQDQDFSGKIRVANANDSSV
tr|E7D4P3|E7D4P -NADAGVSM-----S-----STQSRPS--SNVGCQDQDFSLTKGINVANANDSCI
tr|Q8V9B7|Q8V9B -NADAGVSM-----S-----STQSRPS--SNVGCQDQDFSLTKGINVANANDSCI
tr|I6WJ78|I6WJ7 -NADAGVSM-----S-----STQSRPS--SNVGCQDQDFSLTKGKIVSANLDSSV
tr|Q6XD85|Q6XD8 YRPSSKASDK-----
sp|Q00682|NSP5 -DESAHVSAD-----A-----IQEKDETVPQIDHRIMKWLSDHSGVSLNGGINFTK
tr|Q32Y78|Q32Y7 -DESAHVSAD-----A-----IQEKDETVPQIDHRIMKWLSDHSGVSLNGGINFTK
tr|E2EBV0|E2EBV -SSMVSQCDDEDDIASQEMNKLETIVNSACADEBQ---NIDWNEYLEENSGIKIMEGRVSTN
tr|M4H296|M4H29 -NESNMEVSN-----N-----FEINLPS---DHSCVSVKSSNSMNSQNFKSAV
tr|U3R085|U3R08 MDREDIEDLC-----DM-----
tr|M4H2A1|M4H2A VEKQSESEN-----
sp|Q45UF0|NSP5 QTKEDKKTDK-----

```

```

cons . . . . .

```

```

tr|A0A0G2U594|A SIAT-----KPKKRSKDSRSRK-----HY--PKI--E-----
tr|K7QGP6|K7QGP SIST-----SIKKRSKDSRSRK-----HY--PKI--E-----
tr|E7D4P3|E7D4P SIST-----DHKKK--KSKKRSRK-----HY--PRI--E-----
tr|Q8V9B7|Q8V9B SIST-----DHKKK--KSKKRSRK-----HY--PRI--E-----
tr|I6WJ78|I6WJ7 SIST-----NVKKRSKDSRSRK-----HY--PRI--E-----
tr|Q6XD85|Q6XD8 -----SYR-----EM--KKK--Y-----
sp|Q00682|NSP5 AKSK-----LKETENEITEMSKTNLLVNASVGINSNVGAPE--NFINOT-----
tr|Q32Y78|Q32Y7 AKSK-----VKETENEITEMSKTNLLVNASVGINSNVGAPE--NFINOT-----
tr|E2EBV0|E2EBV EVDLNGVT-----ES-KILNRNSII-----NKNIDSAV--K-----
tr|M4H296|M4H29 QSIHQHSRIRENPKPQ--RQIQCKKKK-----HK--EKA--V-----
tr|U3R085|U3R08 -----QDT-----EE--ISO--KISVKITDNM
tr|M4H2A1|M4H2A -----EYD-----VP--DTE--FIPKKT--NI
sp|Q45UF0|NSP5 -----KIK-----RI--EKV--K-----

```

```

cons . . . . .

```

```

tr|A0A0G2U594|A --AES-----DSDEY-----VLDDSDSDGCKCK--NCKYKKKYFIIIRGRMQ
tr|K7QGP6|K7QGP --AES-----DSDEY-----VLDDSDSDGCKCK--NCKYKKKYFIIIRGRMQ
tr|E7D4P3|E7D4P --ADS-----DSDEY-----VLDDSDSDGCKCK--NCKYKKKYFIIIRGRMQ
tr|Q8V9B7|Q8V9B --ADS-----DSDEY-----VLDDSDSDGCKCK--NCKYKKKYFIIIRGRMQ
tr|I6WJ78|I6WJ7 --ADS-----DSDEY-----VLDDSDSDGCKCK--NCKYKKKYFIIIRGRMQ
tr|Q6XD85|Q6XD8 --DGTSTSDSILEKLESLMLEIKIKQ--MNQPVITDAAPIM--ILRNVDNLTIRQK
sp|Q00682|NSP5 --IKT-----EAVSD--MFEDIEDGICCK--NCPYREKRYKRLSRMRN
tr|Q32Y78|Q32Y7 --VKT-----EAVSD--MFEDIEDGICCK--NCPYREKRYKRLSRMRN
tr|E2EBV0|E2EBV --KKA-----NINRM--NMHDTSSDEECNR--NCKCKKRLKRLKRMSTI
tr|M4H296|M4H29 --IDA-----ISDDEWGNRVESPDSSDSDTCNNSCCKCKRYKRLKRSVKH
tr|U3R085|U3R08 RANGDYITTEMLNEISRLRVLDVVKK-----VSKINSIDAFNT--ILRNVDNLTIRQK
tr|M4H2A1|M4H2A IDWGSRAKQIMNEISKIRMEMDVKE-----AMKPGVDAAPNL--ILRNVDNLTIRQK
sp|Q45UF0|NSP5 --E--NNQNDSSMLQIALSLRIQIES-----ETKLTSLDSIWT--IITQADNLTTPCKK

```

```

cons . . . . .

```

```

tr|A0A0G2U594|A VAMQL-IEDLLK
tr|K7QGP6|K7QGP VAMQL-IEDL--
tr|E7D4P3|E7D4P VAMQL-IEDL--
tr|Q8V9B7|Q8V9B VAMQL-IEDL--
tr|I6WJ78|I6WJ7 VAMQL-IEDL--
tr|Q6XD85|Q6XD8 ALTNALINSM-N
sp|Q00682|NSP5 VLIDM-INEM--
tr|Q32Y78|Q32Y7 VLIDM-INEM--
tr|E2EBV0|E2EBV LIAES---Y---
tr|M4H296|M4H29 TVARL-ISDL--
tr|U3R085|U3R08 ALVNAIVTMM-K
tr|M4H2A1|M4H2A ALVNAIVMSM-K
sp|Q45UF0|NSP5 SLISAILATM-R

```

```

cons . . . . .

```

Fig. (S7). MSA result of non-structural protein 5.

T-COFFEE, Version_11.00 (Version_11.00)

Cedric Notredame

CPU TIME:0 sec.

SCORE=409

*

BAD AVG GOOD

*

```

tr|A0A0G2UUF8|A : 53
tr|K7QG9|K7QGK : 53
tr|A0SX5|A0SX5M : 53
tr|Q5KR57|Q5KR5 : 53
tr|I6VGG2|I6VGG : 53
tr|D6NGG4|D6NGG : 28
sp|Q82055|NSP4 : 38
tr|Q32Y80|Q32Y8 : 39
tr|E2EBU8|E2EBU : 38
tr|M4H2P5|M4H2P : 42
tr|U3R5W0|U3R5W : 37
tr|M4H2P7|M4H2P : 37
sp|Q45UF1|NSP4 : 31
cons : 40

```

```

tr|A0A0G2UUF8|A MDKLADEL-----NYT-----SNVIDLMNSTLH--SIIDDPGMAYFPPIASVLTILF
tr|K7QG9|K7QGK MEKLTDL-----NYT-----LSVITLMNSTLH--TILEDPMAYFPPIASVLTILF
tr|A0SX5|A0SX5M MDKLADEL-----NYT-----LSVITLMNSTLH--SIIDDPGMAYFPPIASVLTILF
tr|Q5KR57|Q5KR5 MDKLADEL-----NYT-----LSVITLMNSTLH--SIIDDPGMAYFPPIASVLTILF
tr|I6VGG2|I6VGG MEKLTDL-----NYT-----LSVITLMNSTLH--TILEDPMAYFPPIASVLTILF
tr|D6NGG4|D6NGG MAESSNM-----QDLFVQAYEEILKLASSVNH--EQIRETISSSSP-Q-KILRTGM
sp|Q82055|NSP4 MDFINQ-----TLE-----SKYTESNVDITIPYLLGLI
tr|Q32Y80|Q32Y8 MDFINQ-----TLE-----SKYTESNVDITIPYLLGLI
tr|E2EBU8|E2EBU MSLLESIG-----NINVT-----EVIAGLNINEIG-T-----QIWIYITVIALLSII
tr|M4H2P5|M4H2P MDASSIMANINITGNDTAAN-----GSVELTLMINN--YITNPGTFLYT-IITFLSTMF
tr|U3R5W0|U3R5W MEKEATH-----SMT-----KVLITSLTENIDYEKVVQI-----VITALIAFF
tr|M4H2P7|M4H2P MEETIQ-----NMT-----KMLMATLTDHIDENKILQT-----FVTAIVAF
sp|Q45UF1|NSP4 MEGTSES-----FVLDE-----FEVNNNDYDNDFFIS--RFSQNPINAFSLFTDGNLQEFF

```

```

cons *

```

```

tr|A0A0G2UUF8|A T-----LHKASVPTMKI--ALKTSKCSY--KVIKYCIV-TIFNALLKLA--GYKEQI
tr|K7QG9|K7QGK T-----LHKASVPTMKI--ALKTSKCSY--KVIKYCIV-TIFNALLKLA--GYKEQI
tr|A0SX5|A0SX5M T-----LHKASVPTMKI--ALKTSKCSY--KVIKYCIV-TIFNALLKLA--GYKEQV
tr|Q5KR57|Q5KR5 T-----LHKASVPTMKI--ALKTSKCSY--KVIKYCIV-TIFNALLKLA--GYKEQV
tr|I6VGG2|I6VGG T-----LHKASVPTMKI--ALKTSKCSY--KVIKYCIV-TIFNALLKLA--GYKEQI
tr|D6NGG4|D6NGG L-----TLGALLTTLIV--KKGSTKFLT--TKMSNIVY-----LAEMLV--WKAQ
sp|Q82055|NSP4 -----LALTNGSRILRFINSFI--IICKHIVT-TSKSAIDKMR--KINNSE
tr|Q32Y80|Q32Y8 -----LALTNGSRILRFINSFI--IICKHIVT-TSKSAIDKMR--KINNSE
tr|E2EBU8|E2EBU S-----LIKAKAY--KT--LPLLCFFV--TTVKQITI-LINDKILRIF--GIDVSI
tr|M4H2P5|M4H2P A-----ISKIGV--IKV--ISKELVSI--KKIKELVM-VCDRLNKNV--GVDVAI
tr|U3R5W0|U3R5W F-----NKKVATPTIKK--IL--TRVKG--KINEKIDSIIDQSMRLRLGTQNSKEMV
tr|M4H2P7|M4H2P F-----NKKVATPTIKK--VL--SKVKE--KINEKIDSVIIDQSMRLRLGTQNSKEMV
sp|Q45UF1|NSP4 MNNSEKIVIRVVLIVISLCSIK--AQTSKLIIVVRLLEFKIY--NVINNLVKNV--INREKI

```

```

cons

```

```

tr|A0A0G2UUF8|A STKDEVE--K-----QMDRIKEMRQLEMIEKLT-----
tr|K7QG9|K7QGK TTKDEIE--K-----QMDRVVKEMRQLEMIDKLT-----
tr|A0SX5|A0SX5M TTKDEIE--R-----QMDRVVKEMRQLEMIDKLT-----
tr|Q5KR57|Q5KR5 TTKDEIE--Q-----QMDRVVKEMRQLEMIDKLT-----
tr|I6VGG2|I6VGG TTKDEIE--K-----QMDRVVKEMRQLEMIDKLT-----
tr|D6NGG4|D6NGG TVKTIQDEVLHQRKVIKQLQCLDKMCLTEKLYDIEHFQMDVTKELIQCMERKMIDSKV
sp|Q82055|NSP4 HNTKNAH--E-----EYEEVMKQIREMR-----
tr|Q32Y80|Q32Y8 HNTKNAH--E-----EYEEVMKQIREMR-----
tr|E2EBU8|E2EBU TDTQVIE--S-----HFAYIRDDLKIRLEIQS-----
tr|M4H2P5|M4H2P TDQMRLN--M-----DLQYIKNELNELKMLVVKVT-----
tr|U3R5W0|U3R5W TSLAQE--M-Q-----KMDAIVNIETESMNFVMA-----RFDEKSIDDKRI
tr|M4H2P7|M4H2P TNCIAQE--M-R-----KFDTIMQLEDTTLDVFMV-----RFDEKTVIDDKKI
sp|Q45UF1|NSP4 INHQVVD--N-----RFREFEERFLL-----

```

```

cons

```

```

tr|A0A0G2UUF8|A -----TREVEQVELLKRIYD-K-LIIQRTDDIDMS-----
tr|K7QG9|K7QGK -----TREIEQVELLKRIHD-R-LMVRPIDEIDMT-----
tr|A0SX5|A0SX5M -----TREIEQVELLKRIHD-N-LITRPADIVDMT-----
tr|Q5KR57|Q5KR5 -----TREIEQVELLKRIHD-N-LITRPDIVDMT-----
tr|I6VGG2|I6VGG -----TREIEQVELLKRIYD-R-LMVRSTDEIDMT-----
tr|D6NGG4|D6NGG KEVEKSCDRRIRDYDWKIATLTAHTQ-QFPTHIDIINQHIE-----
sp|Q82055|NSP4 -----IHMALFNSLHDDN-VKWRMSIESIRRE-----
tr|Q32Y80|Q32Y8 -----IHMALFNSLHDDN-VKWRMSIESIRRE-----
tr|E2EBU8|E2EBU -----LKTIKKLINE-T-I-----
tr|M4H2P5|M4H2P -----LCRTV-FVDRA-----
tr|U3R5W0|U3R5W KEVEESMERMRDPEIKMNALIKSVTQ-C-VKKGKIPVQMK-----
tr|M4H2P7|M4H2P RGIEESMERMRDPEIKMNALIKSVTQ-C-TKKKETPIVHIA-----
sp|Q45UF1|NSP4 -----LLQHDK--NIAKQDDIVQYNKLDNFAESIKSEFNLKVAEME

```

```

cons

```

```

tr|A0A0G2UUF8|A ---KEFNQKN--I-KSIEE-W--EEGRNPYPESKVTIA---SM-----
tr|K7QG9|K7QGK ---KEINQKN--I-KTIEE-W--GNGKNPYEPKEVIA---AMCEVELPSTVSGGGVLYSK
tr|A0SX5|A0SX5M ---KEFNQKN--I-KTIDE-W--ESGKNPYEPIEVIA---SM-----
tr|Q5KR57|Q5KR5 ---KEFNQKN--V-KTIDE-W--ENGKNPYEPIEVIA---SM-----
tr|I6VGG2|I6VGG ---KEINQKN--V-KTLEE-W--ENGKNPYEPIEVIA---AM-----
tr|D6NGG4|D6NGG ---DD-----AETQIIQQHMNKQARVKLNSRNRL-----
sp|Q82055|NSP4 ---KKHEMOMS---DNRNE---FKHSHNDTNICEKSGLETEVCL
tr|Q32Y80|Q32Y8 ---KKHEIKMS---DNRNE---FKHSHDDTNICEKSGLETEVCL
tr|E2EBU8|E2EBU -----EQ---NVSG-----
tr|M4H2P5|M4H2P ---QQTDKQDTS-KVSK-L--DSTANATHVDVTMT---E-TN-----
tr|U3R5W0|U3R5W ---QQ-----LDMQP---YTETKKEYDPHDKL-----
tr|M4H2P7|M4H2P ---QQ-----DKVOL---YTGPKEYPDPRDKL-----
sp|Q45UF1|NSP4 RRFQELKWRCD---MIANK---AMNTIVLANVTSDNNKDEKIVFD-----EGSVVQYN

```

```

cons

```

```

tr|A0A0G2UUF8|A ----
tr|K7QG9|K7QGK PHRT
tr|A0SX5|A0SX5M ----
tr|Q5KR57|Q5KR5 ----
tr|I6VGG2|I6VGG ----
tr|D6NGG4|D6NGG ----
sp|Q82055|NSP4 ----
tr|Q32Y80|Q32Y8 ----
tr|E2EBU8|E2EBU ----
tr|M4H2P5|M4H2P ----
tr|U3R5W0|U3R5W ----
tr|M4H2P7|M4H2P ----
sp|Q45UF1|NSP4 RE--

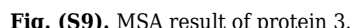
```

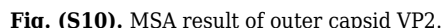
```

cons

```

Fig. (S8). MSA result of non-structural glycoprotein 4.





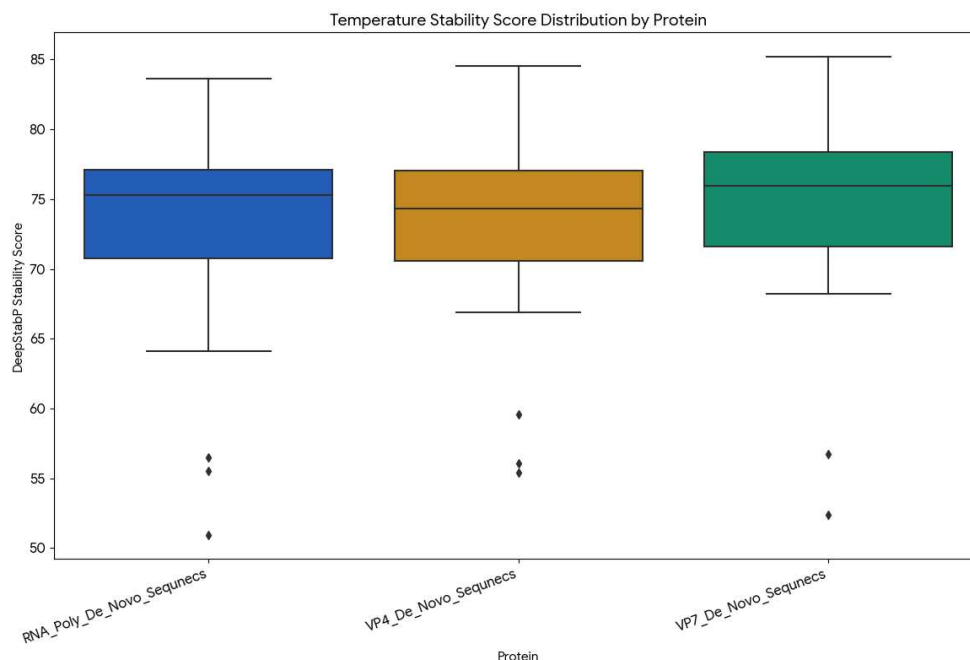


Fig. (S12). Thermal stability analysis of de novo protein sequences.

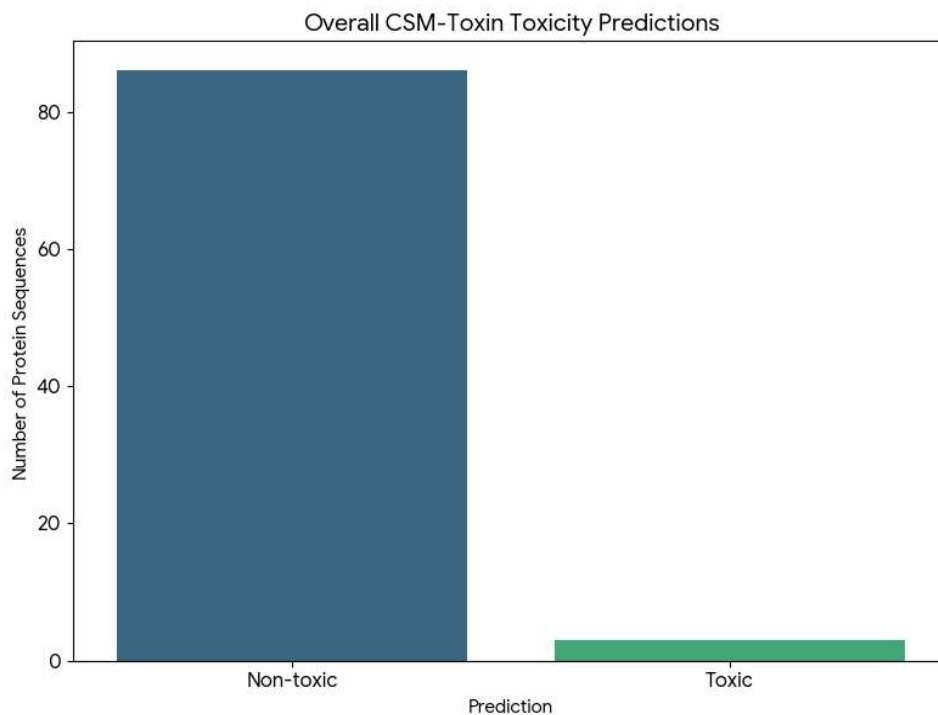


Fig. (S13). Thermal stability analysis of de novo protein sequences.

DISCLAIMER: The above article has been published, as is, ahead-of-print, to provide early visibility but is not the final version. Major publication processes like copyediting, proofing, typesetting and further review are still to be done and may lead to changes in the final published version, if it is eventually published. All legal disclaimers that apply to the final published article also apply to this ahead-of-print version.