



RVPSD: A comprehensive and user-friendly web database for RNA viral protein structures

Qiangzhen Yang^{1,#}, Zhongshuai Tian^{2,#}, Tao Hu^{3,4,#}, Jiangrong Lou¹, Hengcong Liu², Edward C. Holmes⁵, Yongyong Shi^{1,*}, Juan Li^{3,4,6,*}, Weifeng Shi^{2,7,8,*}

¹Bio-X Institutes, Shanghai Jiao Tong University,

² Shanghai Institute of Virology, Shanghai Jiao Tong University School of Medicine,

³ Department of Central Lab, The Second Affiliated Hospital of Shandong First Medical University

⁴Key Laboratory of Emerging Infectious Diseases in Universities of Shandong,
Shandong First Medical University and Shandong Academy of Medical Sciences

⁵School of Medical Sciences, The University of Sydney

⁶School of Clinical and Basic Medical Sciences,
Shandong First Medical University & Shandong Academy of Medical Sciences

⁷Ruijin Hospital, Shanghai Jiao Tong University School of Medicine

⁸School of Life Sciences and Biotechnology, Shanghai Jiao Tong University



Qiangzhen Yang, Zhongshuai Tian, Tao Hu, Jiangrong Lou, Hengcong Liu, Edward C. Holmes, Yongyong Shi, et al. 2026.

RVPSD: A comprehensive and user-friendly web database for RNA viral protein structures.

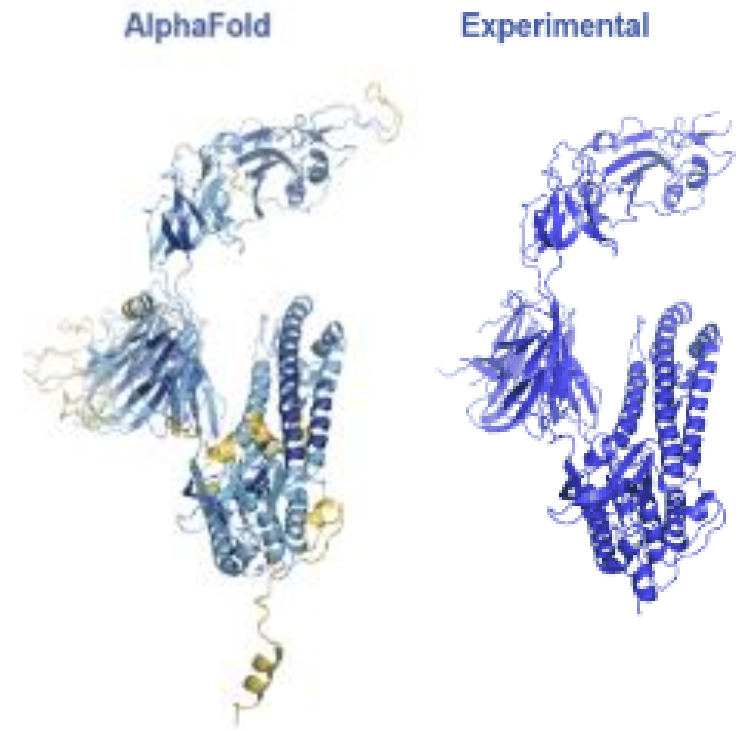
iMetaOmics 3: e70120. <https://doi.org/10.1002/imo2.70120>



Introduction

Key Gap: No comprehensive database for RNA viral protein structures across diverse lineages

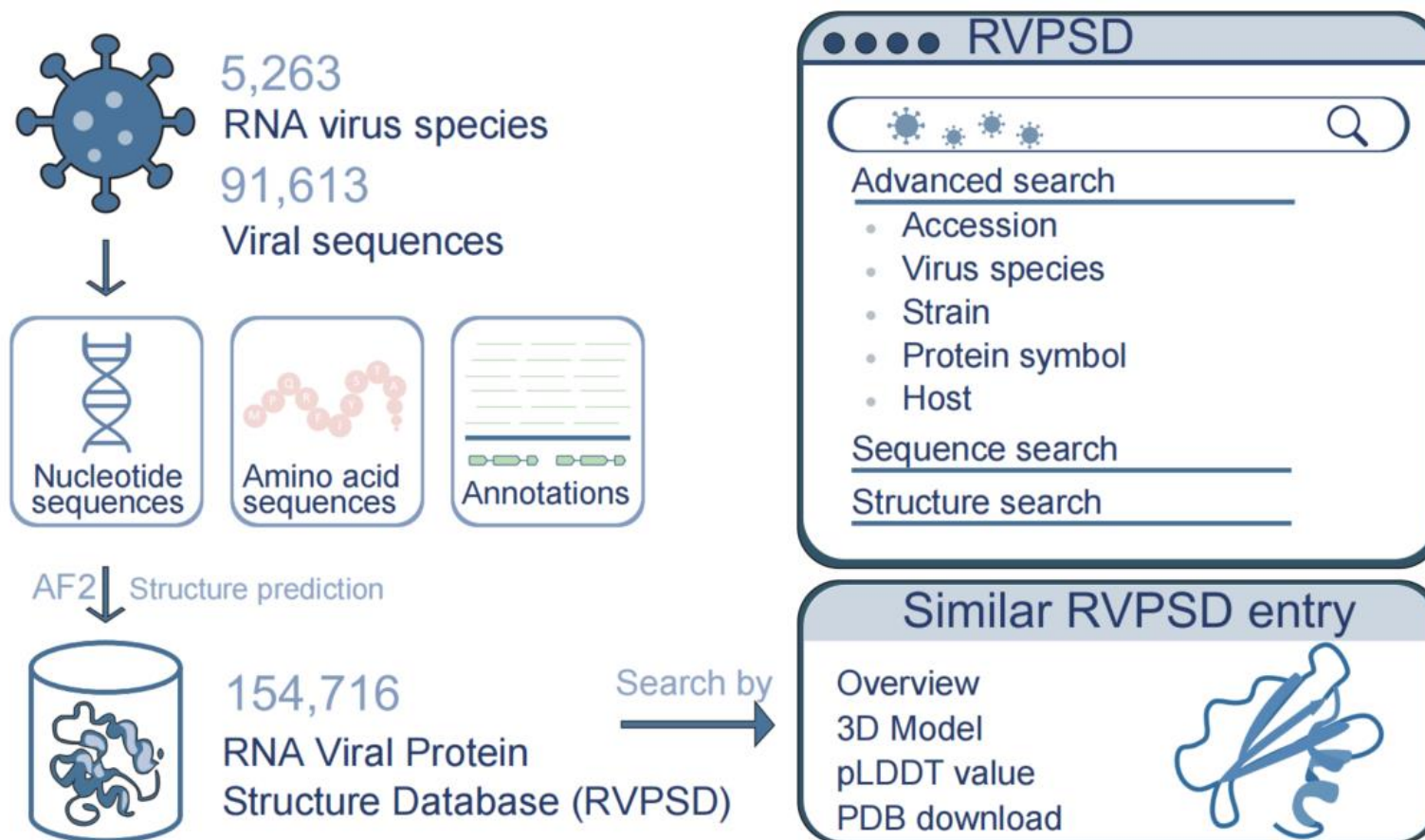
- **RNA viruses** are highly diverse, rapidly evolving pathogens posing major threats to global health
- Protein structures are **essential** for understanding infection mechanisms and guiding therapeutic design
- Experimentally determined structures remain **limited** due to technical and biosafety constraints
- **AI-driven methods** (AlphaFold2) enable large-scale structure prediction
- Existing resources are **fragmented** or family-specific



RVPSD fills this gap as an integrated resource



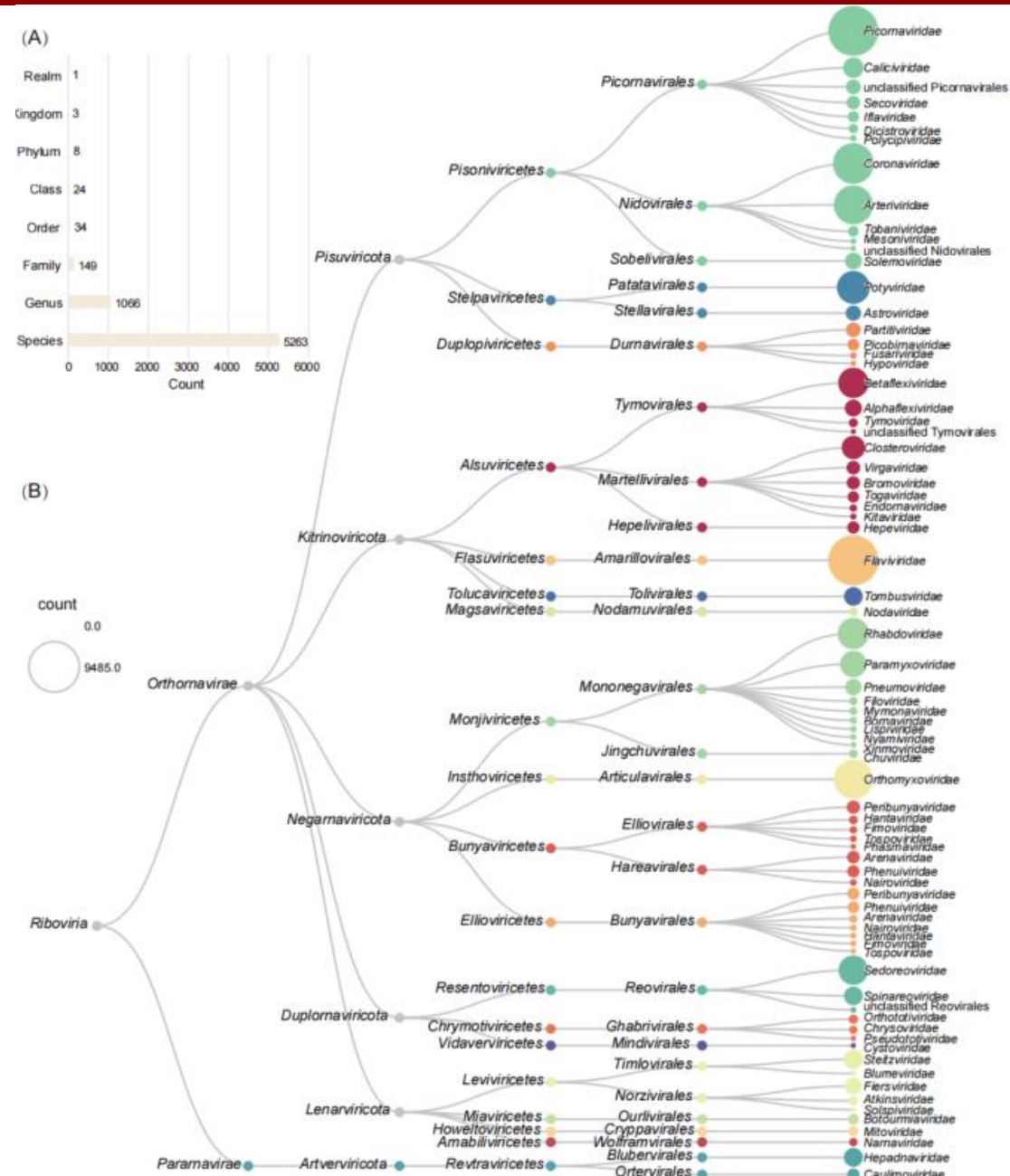
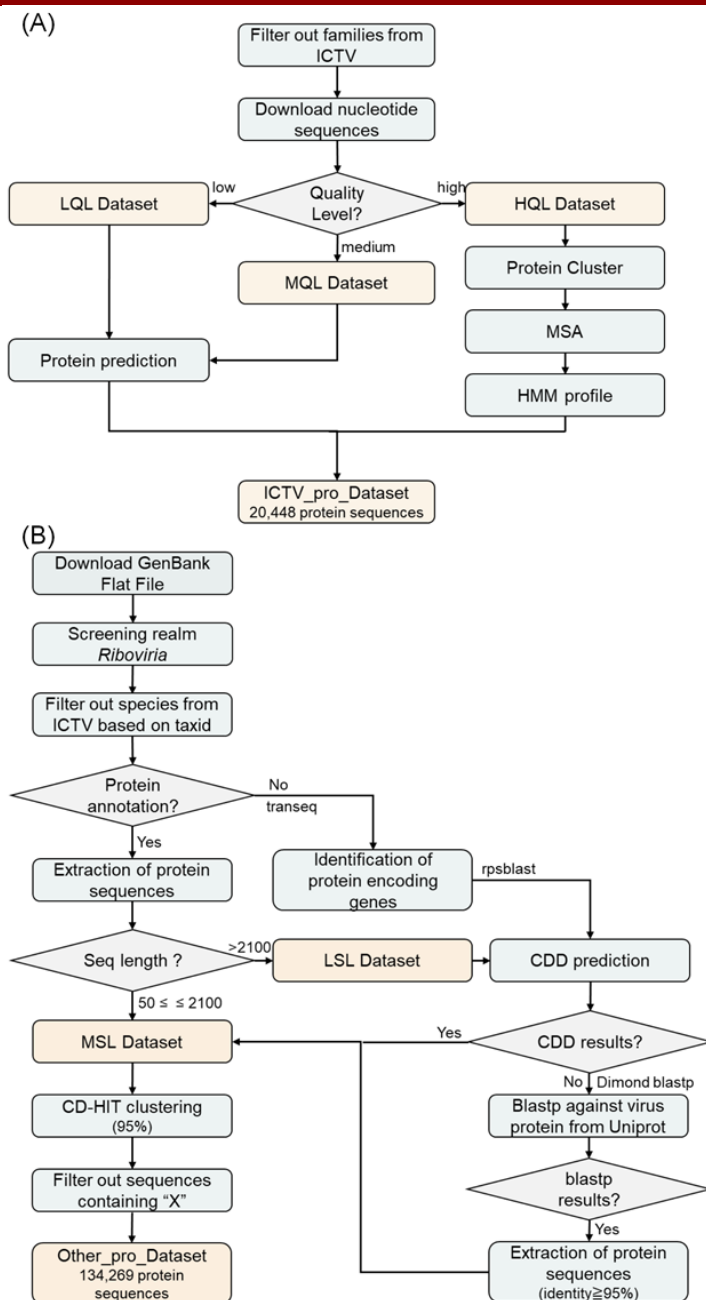
Highlights



- RVPSD provides 154,716 AlphaFold2 predicted protein structures for > 5,200 RNA virus species.
- RVPSD integrates viral protein structures and rich meta data into a comprehensive resource.
- RVPSD enables sequence- and structure-based analyses for RNA viral dark matter discovery.

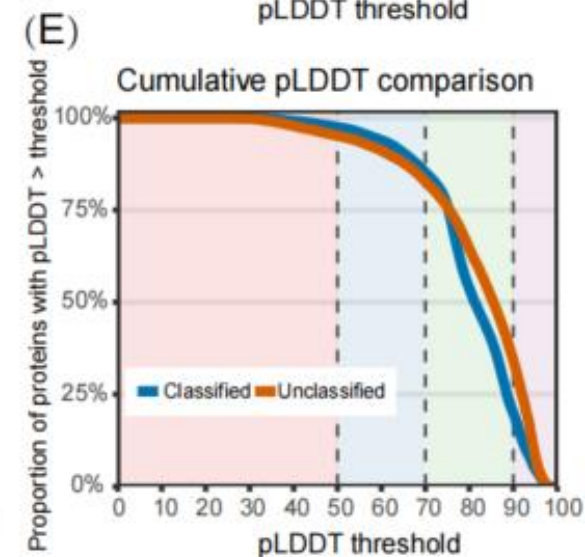
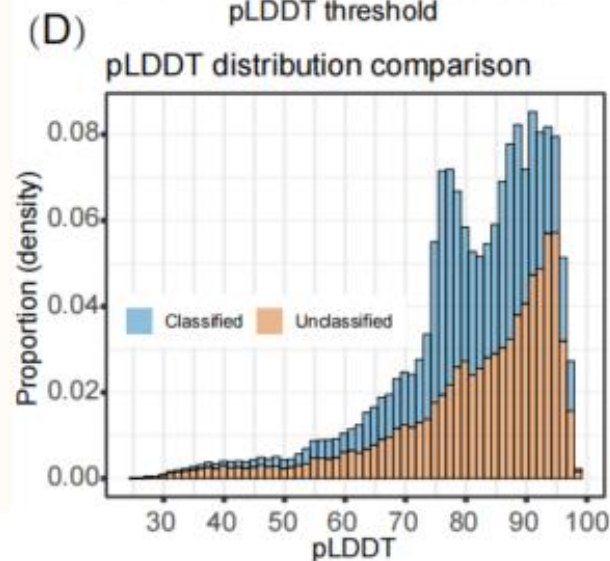
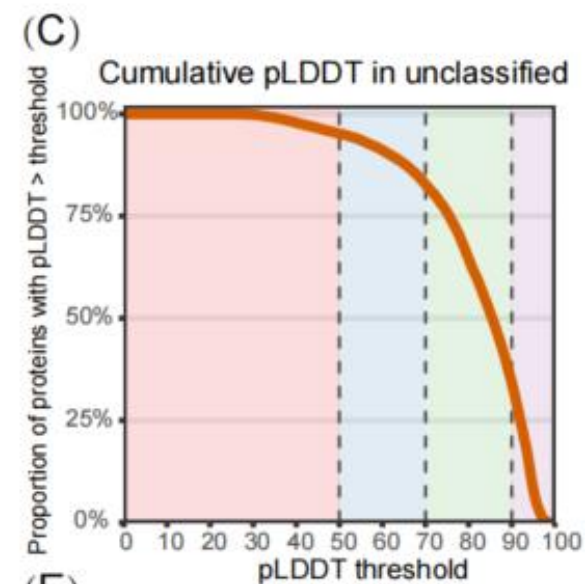
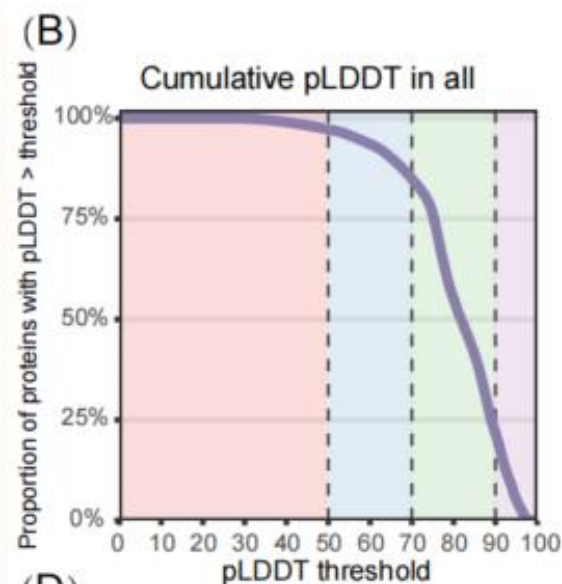
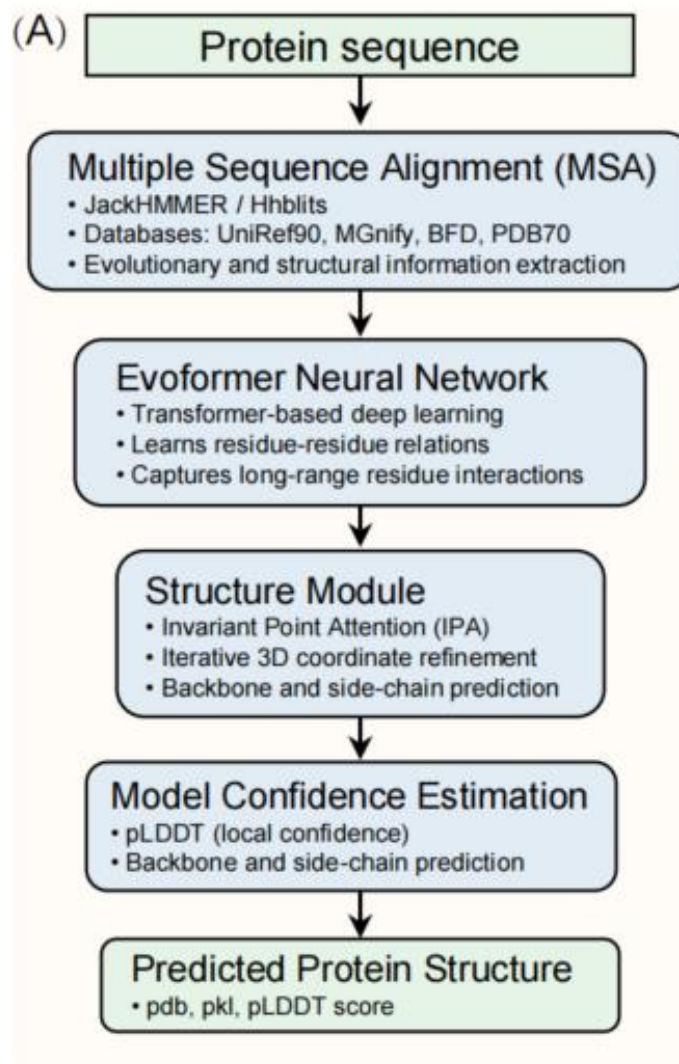


Database Overview & Sequence Collection





Predictive Modeling & Quality Assessment





Database Architecture and Functional Implementation

RNA Viral Protein Structure Database

Advanced Search

You can refine your queries using GenBank accession numbers for viral proteins or nucleotides, virus species names, strain names, protein symbols, taxonomy, viral hosts, collection site or date.

- GenBank Accession Numbers (Protein): Remove
- Virus Name: Remove
- Abbreviated Virus Name: Remove
- Protein Name: Remove
- GenBank Accession Numbers (Nucleotide): Remove
- Taxonomy: Remove
- Strain Name or Isolate Name: Remove
- Host: Remove
- Location: Remove
- Collection date: Remove

Cancel Search

Structure Alignment

* Job Title

Structure 1

- Upload a file [Download the test file](#)
- Or, input content

ATOM	1	N	MET	A	1	1.186	-5.674	-43.192	1.00	39.83	N
ATOM	2	CA	MET	A	1	2.145	-5.769	-42.073	1.00	39.83	C
ATOM	3	C	MET	A	1	1.364	-5.358	-40.945	1.00	39.83	C
ATOM	4	O	MET	A	1	1.296	-4.175	-40.542	1.00	39.83	O
ATOM	5	CB	MET	A	1	3.392	-4.884	-42.283	1.00	39.83	C

Structure 2

- Upload a file [Download the test file](#)
- Or, input content

ATOM	1	N	MET	A	1	1.186	-5.674	-43.192	1.00	39.83	N
ATOM	2	CA	MET	A	1	2.145	-5.769	-42.073	1.00	39.83	C
ATOM	3	C	MET	A	1	1.364	-5.358	-40.945	1.00	39.83	C
ATOM	4	O	MET	A	1	1.296	-4.175	-40.542	1.00	39.83	O
ATOM	5	CB	MET	A	1	3.392	-4.884	-42.283	1.00	39.83	C

* Email [Note: This result will be available for two weeks and please check them immediately.](#)

Cancel Align

Sequence Search

Enter your job title and your protein sequence by uploading a .fasta, .fna, .fa, .fas, or .faa file, or by inserting a sequence in FASTA format.

* Job Title

Enter Query Sequence

- Please upload a file [Download the test file](#)
- Or, input content

```
>ABC123456 1 virionlike RNA-binding protein NSP2 [Betavirus A]  
MALACPCYHLESDYDFIPFNLAKMLTAKTRKSNKPFYDIIYGAAPPFQKRYSTNSDNGM  
SPFTIMPTKXAMLCALNQLKTYQVYNNLSRYVSIRKLSLYIRKNSGDLPRKDKLLKSLTIAI  
GQSKETITITAGGGTYFGAAPTWMILTLEKLMPLDGSFIEYKTLSEKPISSYKELVLELR  
FYQNFATITGGKGVETKTSSTANARVYATPFENKTYVSNFNLDRITQNTAFTSEMGGN  
TLVYCKLLFGQMFENKPGSLSTRNDKESGVGV
```

* Max Target Structures:

* Email [Note: This result will be available for two weeks and please check them immediately.](#)

Cancel Search

Structure Search

Enter your job title and upload your protein structure file in .pdb format, or paste the content in PDB format.

* Job Title

Enter Query Sequence

- Please upload a file [Download the test file](#)
- Or, input content

ATOM	1	N	MET	A	1	1.186	-5.674	-43.192	1.00	39.83	N
ATOM	2	CA	MET <td>A</td> <td>1</td> <td>2.145</td> <th>-5.769</th> <th>-42.073</th> <th>1.00</th> <th>39.83</th> <th>C</th>	A	1	2.145	-5.769	-42.073	1.00	39.83	C
ATOM	3	C	MET	A	1	1.364	-5.358	-40.945	1.00	39.83	C
ATOM	4	O	MET	A	1	1.296	-4.175	-40.542	1.00	39.83	O
ATOM	5	CB	MET	A	1	3.392	-4.884	-42.283	1.00	39.83	C
ATOM	6	CG	MET	A	1	4.399	-5.461	-43.297	1.00	39.83	C
ATOM	7	SD	MET	A	1	5.321	-6.896	-42.675	1.00	39.83	S

* Max Target Structures:

* Email [Note: This result will be available for two weeks and please check them immediately.](#)

Query History

Cancel Search

Search Summary | This query matches 1 structures
1 to 10 of 1 Structures

Title	Type	Email	Upload time	Completion time	Status	Controls
blast	Sequence Search	809078811@qq.com	2026-05-11 12:53:38	2026-05-11 12:53:42	Processing is successful	view

Sequence of AXP33558.1

Structure Tools

- Structure: AXP33558.1, Type: Model, Nothing Focused
- Quick Styles: Apply Representation (Default, Cartoon, Spacefill, Surface), Apply Style (Default, Illustrative)
- Components: Preset, + Add
- Measurements: + Add
- Export Animation, Export Geometry

Information

GenBank Accession Numbers (Protein): AXP33558.1

Virus Name: Abras_virus

Abbreviated Virus Name: ABRV

Protein Name: RNA-dependent_RNA_polymerase

GenBank Accession Numbers (Nucleotide): MH017275.1

Taxonomy: [Riboviria](#) [Orthornavirae](#) [Negarnaviricota](#) [Bunyaviricetes](#) [Eliovirales](#) [Peribunyaviridae](#) [Orthobunyavirus](#) [Orthobunyavirus abrasense](#)

Strain Name or Isolate Name: 75V1183

Host: Culex_paracrybda

Location: Ecuador

Predicted Local-Distance Difference Test: 90.39530959

Collection date: 1974-08-21

PDB download FASTA download

Search

- Advanced Search**
Search using virus metadata.
- Sequence Search**
Find structures using protein sequences via BLASTP.
- Structure Search**
Identify similar protein structures using Foldseek.
- Structure Alignment**
Perform protein structure alignment.
- History**
View your Sequence Search, Structure Search and structure alignment history.

[Complete tasks](#)

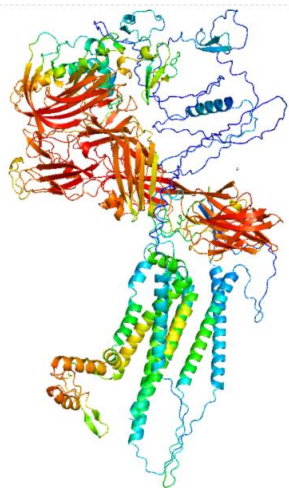


Examples of the utility of RVPSD

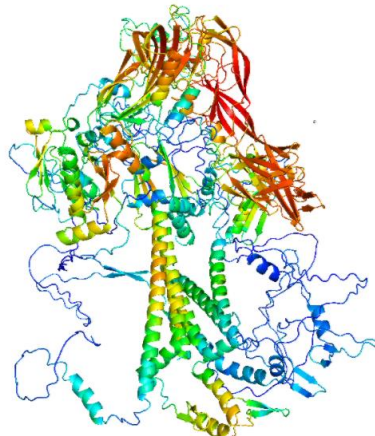
Case I: Functional inference of a “dark” viral protein



Glycoprotein precursor

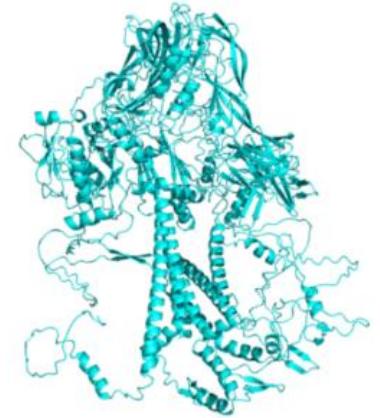


BSVV1 PP480790.1

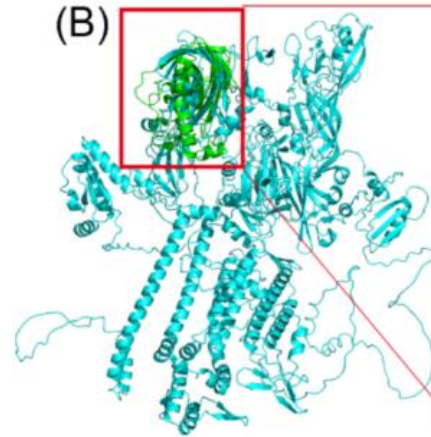


WhIV2 AJG39294.1

(A)



(B)



BSVV1 r576-806

WhIV2 r855-1073



Examples of the utility of RVPSD

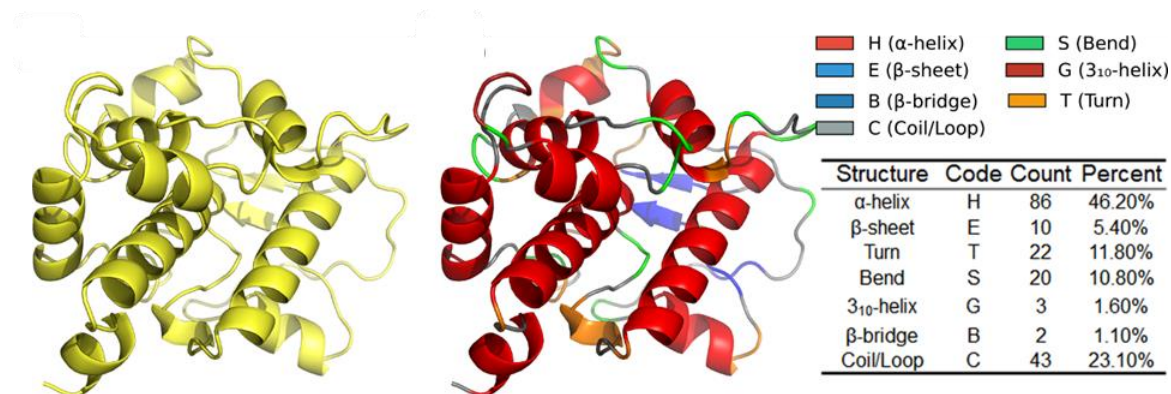
Case II: Structure-based identification of conserved regions in bunyavirus NSs



Structure-based comparison of NSs from multiple bunyaviruses

SFTSV NSs AVV65325.1	MSLSKCSNVDLKSVAMNANTVRLE-----PSLGEYPT-----LRR-----	35
Guertu virus NSs QBQ64951.1	MANSKHRDILNLSVALCTNTVKLE-----PNLGEFPT-----VRR-----	35
Heartland virus NSs UDE22460.1	MSLSKAS-----QPSVKSAQVRL-----PIVVLPENLAELESTSYV-----	35
Hunter Island virus NSs AHI10996.1	MACCL-----REQPTIVLM-----PMRLELP-----LSRK-----	25
Mukawa virus NSs BBE15817.1	MSSERV-----LYFSTLNWRCAVRKKKNCKMGRSHISAVHDASMSLCDADHFFMYNIYP	54
SFTSV NSs AVV65325.1	-----DLVECSCSVLTLMSVVRMGKMTNTVWL-----FGNPKNPLHQLEPGLEQLL	81
Guertu virus NSs QBQ64951.1	-----GLVECSCSVLTVAMVKKMGKMTNTVWL-----FGNPNPLHQLEPAIEHLL	81
Heartland virus NSs UDE22460.1	-----GLVSCKCSVLTCMMRKMKAFTNTVWL-----FGNPNPLHALPEAVEQLL	81
Hunter Island virus NSs AHI10996.1	-----YSNQTTYNFKSWAEYKRRNLI PSEFWI-----LGDPKNELHQTNGHLMSLF	71
Mukawa virus NSs BBE15817.1	NFGFSLKVEQRTVVGCENTVYKFLDHGVMPARVCAWESSGSCVI PRRSWLDCGISWQFCN	114
SFTSV NSs AVV65325.1	DMYYKDMRCYSQRELSALRWP SGKPSVWFLQAAHMFSSI KNSWAMETGRENWRGLFHRI T	141
Guertu virus NSs QBQ64951.1	EKYYVDLNCYSQREKSALRWP SGKPSVWFLQAAHMFSSVKNTWAMETGKENWRGLFHRI T	141
Heartland virus NSs UDE22460.1	DEYSGDLGSYSQQEKSALRWP SGKPSVHFLQAAHFFSLKNTWAVETGQENWRGFFHRI T	141
Hunter Island virus NSs AHI10996.1	RRYHSSFNYSYCLEDAQI MSWPSGFGQNYFLQAAATYFGEELTWEI NDGRVSPGLYLKRI A	131
Mukawa virus NSs BBE15817.1	WTIENLMNNGEVEVLRALSWPTGRPTLQFIR- EFCRSKSFSTTSGFRSVRERVAKLIFKAA	173
SFTSV NSs AVV65325.1	KGRRYLFEGDMI LDSLEAI EKRRRLRLGLPDLIITGLSPI LDVALLQI ESLARLRGMSLNH	201
Guertu virus NSs QBQ64951.1	RGCQYPFEGDMI LDALDII EKRRIRLGLPDSLIITGLSPI LDVALLQI EALARLRGMSLNH	201
Heartland virus NSs UDE22460.1	PGKKYKFEGDMVIDSCYKIDERRRRMGLPDFTITGLNPI MDVALLQI ESLLRVRGLTLNY	201
Hunter Island virus NSs AHI10996.1	KHPKVSAGDLII LACERI EAKRLEMRMPETFITGLSPI RDVALIQI EMMCRSRSQCLNV	191
Mukawa virus NSs BBE15817.1	KSEEGESLGEAI TKCWRKVRSEAMRLGVPEFDVPGVDLIRDI SLIQIERA-----YMRI	227
SFTSV NSs AVV65325.1	HLFTSSSLRKPLL-DCWDFFIPIRKKKTGSGYS-----VLDEDEDEPGVL-----	244
Guertu virus NSs QBQ64951.1	HMYTASALRKPLL-DCFNFIPIKRRKTSDDSYR-----DDEHGED-FL-----	243
Heartland virus NSs UDE22460.1	HLFTSSSFLDKPLL-DSL YFAIWRDKKKDDGSGYS-----QDKGARQDDPLNPLDELLY-	252
Hunter Island virus NSs AHI10996.1	PLLSCLCLATEEYLSHEPILTLTDSSEHDYESP-----HPSLTTSRSS-----	234
Mukawa virus NSs BBE15817.1	NKCGSKWDNL PATGWDEPGSGSMYEGVI PFFNLSQKLSMKCASKVRNEPESRVWMRRKLKG	287
SFTSV NSs AVV65325.1	-----QGYPYLMAHYLNRC--PFHNLI RFDEELRTAALNTI WGRDWPALGDPPKEV	296
Guertu virus NSs QBQ64951.1	-----KGYPLLMAHYLNKC--PFHNLI KYDEELREAALNTI WGDWPKL-----	288
Heartland virus NSs UDE22460.1	-----LSDLPKPLAHYLNKC--PLHNIIMHDEEVREAYLNPI WGDWPAALSSSP--	302
Hunter Island virus NSs AHI10996.1	-----LPEPDY-----ESDLELREAMLNTAWGDWPSLGEPSK--	270
Mukawa virus NSs BBE15817.1	KKVESKSQLIDLQEGITNIWMLDISDHFNMFPLDIRVL DKSFGKDWPA LGAP----	343

Structure-derived sequence alignment of bunyavirus NSs proteins



The conserved core domain of NSs proteins



Summary

- ❑ In this study, we introduced RVPSPD: A comprehensive and user-friendly web database for RNA viral protein structures.
- ❑ Currently, RVPSPD provides 154,716 predicted structures with 5-to-7-fold coverage over existing resources.
- ❑ The platform integrates multiple search and analysis tools, and case studies demonstrate its utility for functional inference and conserved feature identification.
- ❑ Website: <https://virus.9itsg.net>

Qiangzhen Yang, Zhongshuai Tian, Tao Hu, Jiangrong Lou, Hengcong Liu, Edward C. Holmes, Yongyong Shi, et al. 2026.
RVPSPD: A comprehensive and user-friendly web database for RNA viral protein structures.
iMetaOmics 3: e70120. <https://doi.org/10.1002/imo2.70120>



“***iMeta***” launched in 2022 by iMeta Science Society, **impact factor (IF) 44.4**, **ranking top 47/22643 in the world**. It aims to publish innovative and high-quality papers with broad and diverse audiences. Its scope is similar to *Cell*, *Nature* and *Science*. Its unique features include video abstract, bilingual publication, and social media with 800,000 followers. Indexed by [SCIE/ESI](#), [PubMed](#), [Google Scholar](#) etc.

“***iMetaOmics***” launched in 2024, indexed by [ESCI](#), [PubMed](#), [Google Scholar](#), with a **target IF>15**, and its scope is similar to *Nature Communications*, *Science Advances*, *Advanced Science*, *Nucleic Acids Research*, etc.

“***iMetaMed***” launched in 2025, with a target IF>15, similar to *Med*, *Cell Reports Medicine*, *eBioMedicine*, *eClinicalMedicine* etc.



Society: <http://www.imeta.science>

Publisher: <https://wileyonlinelibrary.com/journal/imeta>

iMeta: <https://wiley.atyponrex.com/journal/IMT2>

Submission: iMetaOmics: <https://wiley.atyponrex.com/journal/IMO2>

iMetaMed: <https://wiley.atyponrex.com/journal/IMM3>



[iMetaScience](#)



[iMetaScience](#)



office@imeta.science

imetaomics@imeta.science



[Promotion Video](#)

Update
2026/6/17