



Modeling of hepatitis A virus polyprotein a

Manjusha Tyagi*, Soni Devi

Department of Microbiology, SGRR University, Dehradun, BTS Institute, 197 Mundia Ahmednagar, Pilibhit Road, Bareilly-243122,

*corresponding author: manjushatyagi2008@gmail.com

Abstract

polyprotein a was done using SWISS-MODEL software. Valuable structural properties could be obtained.

Key words: Modeling, hepatitis virus A, polyprotein a.

Introduction

Study of hepatitis virus protein structures will provide valuable information, hence the present work was undertaken.

Materials and Methods

Protein

Hepatitis A virus (wild-type) RNA, complete genome GenBank: M14707.1 HPA 7478 bp ss-RNA linear was downloaded and the amino acid sequence was used for modelling.

Modelling software

<https://swissmodel.expasy.org> was reached to model the protein.

Results and Discussion

Project summary

MNMSRQGIQFQTVGSGLDHILSLADIEEEQMIQSVDRTAVTGASYFTSVDQSSVHTAEVGS HQVEPLRTSVDKPGSKKTQGEKFFLIHS	12
ADWLTTHALFHEVAKL DVVKKLLYNEQFAVQGL	0
LRYHTYARFGIEIQVINPTPFQGG LICAMVPGDQSYGSIASLTVYPHGLLNCNINNVVRIKVPFIYTRGAYHFHKDPQYPVWELTIR	24
VWSELNIGTGTSAYTSLNVLARFTDLELHGLT	0
PLSTQMMRNEFRVSTTENVVNL SNYEDARAKMSFALDQEDWKS DPSQGGGIKITHFTTWT SIP TLAAQFPFNASDSVGQQIKVIPVDP	36
YFFQMTNTNPDQKCIT ALASICQMFCFWRGDL	0
VFDFQVFPTKYHSGRLLFCFVPGNELIDVSGITLKQATTAPCAVMDITGVQSTLRF RVPWISDTPYRVNRYTKSAHQKGEYTAIGKLI	48
VYCNRLTSPSNVASHVRVNVYLSAINLECF A	0
PLYHAMDVTTQVGDDSGGFSTTVSTEQNVDPDQVGITTMKDLK GKANRGKMDVSGVQAPVGAITTIEDPVLAKKVPETFPPELKPGESR	60
HTSDHMSIYKFMGRSHFLCTFTFNSNKEYTF	0
PITLSSTSNPPHGLPSTLRWFFNLFQLYRGP LDLTIIITGATDV DGMAWF TPVGLAVDTPWVEKESALSIDYKTALGAVRFNTRRTGN	72
IQIRLPWYSYLYAVSGALDGLGDKTDSTFGLV	0
SIQIANYNHSDEYLSFSCYLSVTEQSEFYFP RAPLNSNAMLSTESMMSRIAAGDLESSVDDPRSEEDKRFESHIECRKPYKELRLEV G	84
KQRLKYAQEELSNEVLP PPRKMKGLFSQAKIS	0

LFYTEEHEIMKFSWRGVTADTRALRRFGFSLAAGRSVWTTLEMDAGVLTGRILIRLNDEKWTEMKDDKIVSLIEKFTSNKYWSKVNFPHG MLDLEEIAANSKDFPNMSETDLCLLHHLNPK	96 0
KINLADRMLGLSGVQEIKEQGVGLIAECRTFLDSIAGTLKSMF GFHHSVTVEIINTVLCFVKSGILLYVIQQLNQDEHSHIIGLLRV MNYADIGCSVISCGKVF SKMLETVFNWQMSR	10 80
MMELRQTQFSNWLRLDICSGITIFKNFKDAIYWLYTKLKDFYEVNYGKKKDI LNILKDNQQKIEKAIEEADEF CILQIQDVEKFEQYQK GVDLIQKLRTVHSMQVDPNLMVHLSPLRDCI	12 00
ARVHQKLNKLSINQAMVTRCEPVV CYLYGKRGGKSLTSIALATKICKHYGVEPEKNIYTKPVASDYWDGYSGLVCIIDDIGQNTT DEDWSDFCQLVSGCPMRLNMMASLEEKGRHFSS	13 20
PFI IATSNWNSP PKTVYVKEAIDRRLHFKVEVKPASFFKNP HNDMLNVNLAKTNDAIKDMSCVDLIMDGHNVSLMDLLSSLVMTVEI RKQNMTEFMELWSQGISDDND SAVAEFFQSF	14 40
PSGEP SNSKLSGFFQSVTNHKWVAVGA AVGILGVLVGGWFVYKHSRKEEIP AEGVYHGVTKPKQVIKLDADPVESQSTLEIAGLV RKNLVQFGVGEKNGCVRWVMNALGVKDDWLLV	15 60
P SHAYKFEKDYEMMEFYFN RGGTYYSISAGNVVIQSLDVGFDVVMKVPTIPKFRDITQHF I KKGDVPRALNRLATLVTTVNGTPML ISEGPLKMEEKATYVHKKNDGTTVDLTVDQAW	16 80
RGKGEGLPGMCGGALVSSNQS IQNAILGIHVAGNSILVAKLVTQEMFQNI DKKIESQRIMKVEFTQCSMN VVSKTLFRKSPIYHHID KTMINFPAAAMPFSKAEIDPMAVMLSKYSLPIV	18 00
EEPEDYKEASIFYQNKIVGKTQLVDDFLDLDMAITGAPGIDA INMDS SPGFPYVQEKLTRDLIWLDENGLLLG VHPRIAQRILFNTV MMENCSDLVVF TTPKDELRLPEKVLKTR	19 20
AIDACPLDYSILCRMYWGP AISYFHLNPGFHTGVAIGIDPDRQWDELFTMIRFGDVGDLDFSAFDASLSPFMIREAGRIMSELSGT PSHFGTALINTIIYSKHLLYNCCYHVC GSMP	20 40
GSPCTALLNSI INNVNLYYVFSKIFGKSPVFFCQALKILCYGDDVLIVFSRDVQIDNLDLIGQKIVDEFK KLGMTATSADKNVPQLKP VSELTFLKRSFNLVEDRIRPAISEKTIWSLIA	21 60
WQRSNAEFEQNL ENAQWFAFMHGYEFYQKFYYFVQSCLEKEMIEYRLKSYDWWRMRFYDQCFICDLS	22 27

Template Results

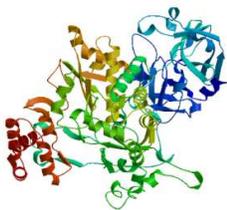
A total of 12856 templates were found to match the target sequence. This list was filtered by a heuristic down to 50. The top templates are:

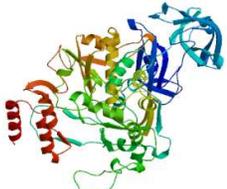
Template	Sequence Identity	Biounit Oligo State	Description
2ijd.2	26.11	homo-dimer	Picornain 3C, RNA-directed RNA polymerase Crystal Structure of the Poliovirus Precursor Protein 3CD
2ijd.1	26.11	homo-dimer	Picornain 3C, RNA-directed RNA polymerase Crystal Structure of the Poliovirus Precursor Protein 3CD
1tp7.1	27.79	monomer	Genome polyprotein Crystal Structure of the RNA-dependent RNA Polymerase from Human Rhinovirus 16
4zp7.1	27.66	monomer	Genome polyprotein Coxsackievirus B3 Polymerase - F364V mutant

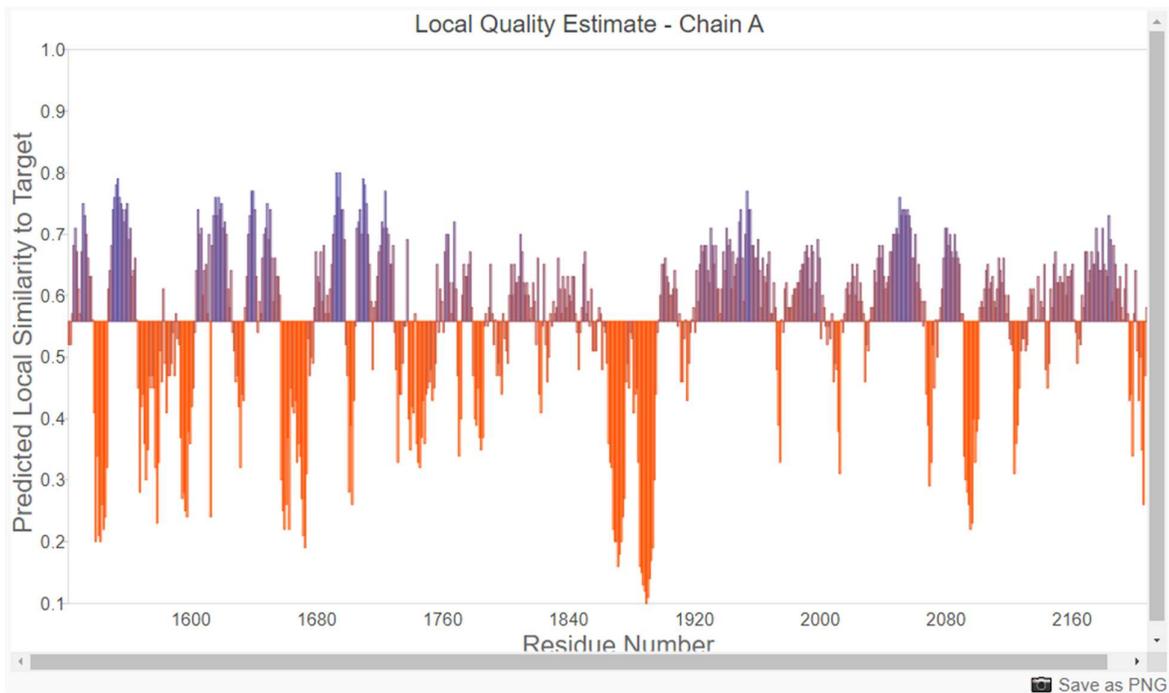
Template	Sequence Identity	Biounit Oligo State	Description
2ijd.2	29.67	homo-dimer	Picornain 3C, RNA-directed RNA polymerase Crystal Structure of the Poliovirus Precursor Protein 3CD

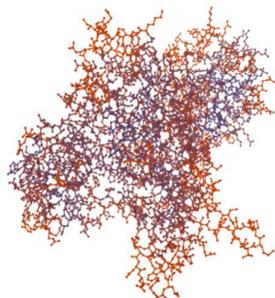
Model Results

Id **Template** **GMQE** **QMEANDisCo** **Global** **Oligo State** **Ligands**

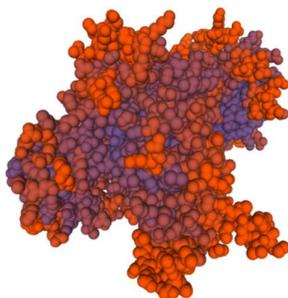
	01	2ijd.2.A	0.18	0.56 ± 0.05	monomer	-
---	----	----------	------	-------------	---------	---

	02	2ijd.2.A	0.10	± 0.05	monomer	-
--	----	----------	------	--------	---------	---





Ball + stick



Spacefill

References

- Berman H, Henrick K, Nakamura H, Markley JL. 2007. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* 35: D301-303.
- Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T. 2017. The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids Res.* 45: D313-D319.
- Bordoli L, Schwede T. 2012. Automated protein structure modeling with SWISS-MODEL Workspace and the Protein Model Portal. *Methods Mol. Biol.* 857:107-136.
- Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30: S162-S173.
- Mariani V, Biasini M, Barbato A, Schwede T. 2013. IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* 29: 2722-2728.
- Schwede T, Kopp J, Guex N, Peitsch MC. 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* 31: 3381-3385.
- Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. 2020. QMEANDisCo – distance constraints applied on model quality estimation. *Bioinformatics* 36: 1765-1771.
- Studer G, Tauriello G, Bienert S, Biasini M, Johner N, Schwede T. 2021. ProMod3 - A versatile homology modelling toolbox. *PLOS Comp. Biol.* 17: e1008667.
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46(W1): W296-W303.