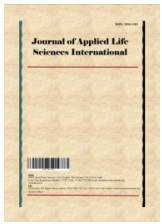


Journal of Applied Life Sciences International

24(7): 26-35, 2021; Article no.JALSI.74719
ISSN: 2394-1103

In silico Structural and Functional Annotation of Tomato Chocolate spot Virus

Ftepti B. Jelani [a*]

[a] *Department of Biotechnology, School of Life Sciences, Modibbo Adama University of Technology, P. M. B 1079, Yola, Nigeria.*

Author's contribution

The sole author designed, analyzed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/JALSI/2021/v24i730249
Editor(s):
(1) Dr. J. Rodolfo Rendón Villalobos, National Polytechnic Institute, México.
Reviewers:
(1) Ali Abdel-hadi Mahoud Alsudani, University of Al-Qadisiyah, Iraq.
(2) S. Priyadharshini, The American College, India.
Complete Peer review History: https://www.sdiarticle4.com/review-history/74719

Original Research Article

Received 01 August 2021
Accepted 04 October 2021
Published 17 November 2021

ABSTRACT

Aims: The study aims to predict *in-silico* the structural and functional annotation of Tomato Chocolate Spot Virus (TCSV) retrieved from Uniprotkb with the accession number C7EXM3.
Study design: To use the *In-silico* approach for the structural and functional annotation of the Tomato Chocolate Spot Virus.
Place and Duration of Study: The research was conducted at the Bioinformatics Unit, Chevron Biotechnology Centre, Modibbo Adama University Yola, Nigeria. Between August 2021 to September 2021.
Methodology: The sequence of the Tomato Chocolate Spot Virus was retrieved from Uniprotkb with accession number C7EXM3, Physicochemical characteristics were computed using the ProtParam tool. The sever SOPMA was used for secondary structure analysis (Helix, Sheets and Coils). The tool CELLO v2.5 was used to predict the subcellular localization of the protein. Four different Homology Modelling tools (trRosetta, Lomet, RaptorX and IntFOLD5) were used to predict the 3D structure of the protein, the quality of the predicted proteins was assessed used PROCHECK. Three tools (InterProScan, NCBI conserved domains and Phobius) were used to get the possible function(s) of the protein.
Results: ProtParam tool computed various Physical and Chemical properties such as Molecular weight (MW) 20396.96 Daltons, isoelectric point (pI) of 6.92. Instability Index 41.94, and Grand Average Hydropathy (GRAVY) -0.503. SOPMA was used for calculating the secondary structure

**Corresponding author: E-mail: fteptibenson@mautech.edu.ng;*

parameters of the protein as Helices (Hh) 43.48%, Extended strands (Ee) 18.48%, Random coils (Cc) 38.04%. CELLO v2.5 was used for subcellular localization of the protein, it predicted that the protein can be both Nuclear and Cytoplasmic with the reliability of 1.653 and 1.504 respectively. Different Homology modelling tools were used to obtain the best 3D structure of the protein. Furthermore, PROCHECK was used to assess the quality of the models obtained. Model from trRosetta was found to be the best because of the quality of the Ramachandran Plot obtained from PROCHECK which has more than 90% of amino acid in the most favourable regions. NCBI-CDD and interproScan predicted that protein is a DNA double-strand break repair Rad50 ATPase, which is involved in the early steps of DNA double-strand break (DSB) repair. Furthermore, the Phobius server predicted the protein to be non-cytoplasmic in its domain, which means they help target proteins to their final destinations.

Conclusion: The study has helped in obtaining the 3D structure of the protein Tomato Chocolate Spot Virus from different Modelling tools, as well as the possible function of the protein.

Keywords: Functional annotation; Protein structure prediction; Homology modeling; Tomato chocolate spot virus.

1. INTRODUCTION

Tomato chocolate spot virus is a new member of torradovirus species in the family Secoviridae [1]. A chocolate-spot virus is a novel group of picorna-like viruses inciting necrosis-related diseases of tomatoes in Mexico [tomato apex necrosis virus (ToANV) and tomato marchitez virus (ToMarV)] and Spain [tomato torrado virus (ToTV)] [2]. Nucleic Acids analysis of virions of the chocolate-spot associated virus showed the genome consist of two single-stranded RNAs of ~7.5 and ~5.1 kb which corresponds to the torradovirus RNA1 and RNA2 [3].

The disease Tomato chocolate spot virus is defined by unique necrotic spots on stems, leaves and petioles that finally enlarge and result in a dieback of apical tissues. This virus is transmissible through sap and graft and causes disease symptoms in a variety of solonaceous plants [2]. The purified virions are icosahedral (~28–30 nm) and composed of two single stranded RNAs genome similar to the described torradoviruses [4]. Thus, the name tomato chocolate spot virus (ToCSV) is proposed.
Determination of ToCSV coat protein three-dimensional structure is vital for understanding its function. However, Experimental determination for structural elucidation of the virus coat protein structure through Nuclear Magnetic resonance (NMR) spectroscopy or X-ray crystallography is expensive and time consuming [5]. Therefore, viable alternatives through computational techniques for building structural models is very paramount. Protein Data Bank (PDB) is a repository for three-dimensional structural data of large biomolecules deposited by Scientists from across the world. These structures provide an excellent foundation for the functional analysis of experimentally derived crystal structures. Thus, help in understanding the protein structure function [6].

There are three types of computational modelling for predicting structural protein models: homology modelling/comparative modelling, by threading and by ab initio/De novo. Among computational modelling techniques, homology modelling is a renowned silico tool for obtaining protein structure models [7]. Homology modelling is on the basis that the three-dimensional structure of a protein is more conserved than its primary structure. Hence, changes in the sequence do not always change the structural domains of a protein, therefore maintaining its original function. It is based on the premise that proteins from the same functional family maintain their structural domains, which allows for comparative modelling by homology [6]. Being homologous means that the proteins belong to the same genetic and functional family, and hypothetically, have the same structural motifs. If a specific protein does not have an elucidated three-dimensional structure but is homologous to a protein with a solved structure, a 3-D structure for the sequence can be built using the known structure as a template [8].

The present study is focused on Insilico Structural and Functional Annotation of the Tomato Chocolate spot Virus.

2. MATERIALS AND METHODS

2.1 Sequence Retrieval

The Amino acids sequence of Tomato Chocolate Spot Virus (TCSV) was retrieved from uniprotkb

(https://www.uniprot.org/uniprot/C7EXM3) with the accession number C7EXM3

2.2 Physicochemical Analysis

The Physicochemical characteristics of the Tomato Chocolate Spot Virus (TCSV) such as the Molecular weight, atomic composition, amino acid composition, theoretical pI, instability index, extinction coefficient and grand average of hydropathicity was determined using Protparam tool (https://web.expasy.org/cgi-bin/protparam/protparam [9].

2.3 Secondary Structure Analysis

The server Sopma was used for secondary structure prediction (helix, sheets and coils) of the protein (https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl) [10].

2.4 Subcellular Localization Prediction

Prediction of Subcellular Localization was carried out using CELLO v2.5 (http://cello.life.nctu.edu.tw/cgi/main.cgi).[11].

2.5 Homology Modelling and Quality Assessment

The 3D structure of the hypothetical protein Tomato Chocolate Spot Virus was obtained using servers such as trRosetta [12], Lomet [13], RaptorX [14], IntFold5 [15]. The Ramachandran plots were checked using PROCHECK server [16] and visualized using pymol tool.

2.6 Functional Annotation

TCSV hypothetical protein C7EXM3 was analysed for the function. Three bioinformatics tools and databases including InterProScan [17], NCBI Conserved Domains Database (NCBI-CDD) [18] and Phobius server [19] were utilized for this reason.

3. RESULTS

3.1 Sequence Retrieval

The sequence of the hypothetical protein Tomato Chocolate Spot Virus (TCSV) was retrieved from Uniprotkb with accession number C7EXM3.

3.2 Physicochemical Characteristics

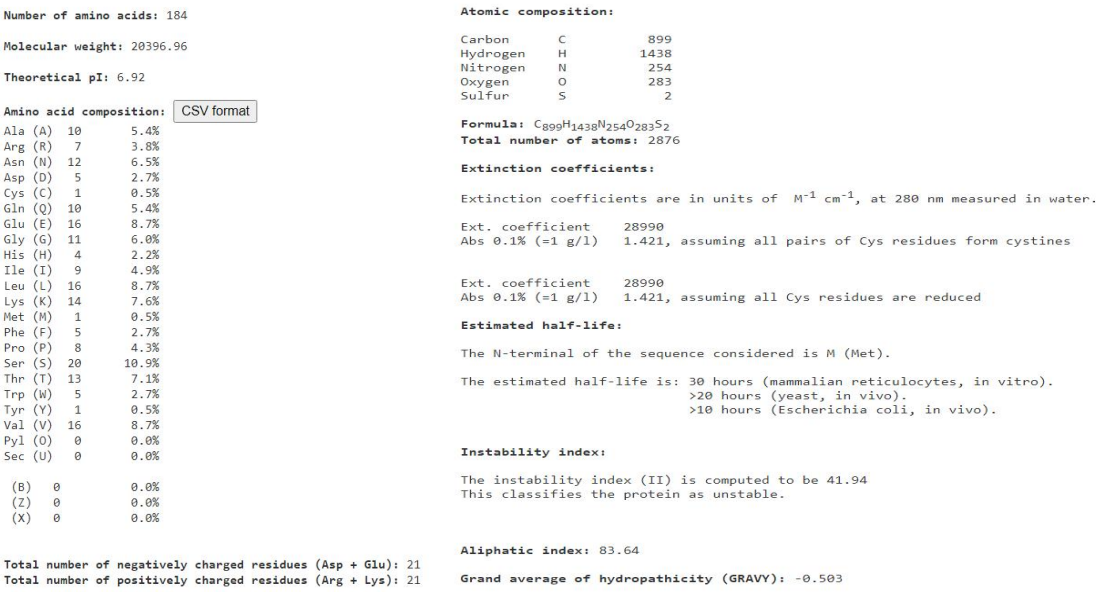
The results presented in Fig. 2 below shows the physicochemical properties of the Tomato Chocolate Spot Virus amino acid sequence. The Molecular weight (MW), the total number of positively (+R), negative charged residues (-R), theoretical isoelectric point (pI), extinction coefficient (EC), aliphatic index (AI) and grand average hydropathy (GRAVY) were computed.

3.3 Secondary Structure Analysis

The results presented in Fig. 3 shows the SOPMA server which was used for the prediction of the secondary structure such as Alpha helix, 310 helix, Pi helix, Beta bridge, Extended strand, Beta turn, Bend region, Random coil, Ambiguous states and other states.

```
        10         20         30         40         50
MSFIGRLNTA EEEKAFHQQV ASSNWICSVD VGSGVINSNP TLDFKVIPPT
        60         70         80         90        100
GGAVSVLTVS WENSTPQLVP GHYLLRSGTW PVKNVKLSGL LVHRSVRLET
       110        120        130        140        150
TRKVLEQNKV SIAQQTENSV SDKGKTTETA WKFKEEIAHL NAELERARKE
       160        170        180
IAEKQSEISK LQLQLSNQPS NNDIFTGWSE DGPK
```

Fig. 1. Amino acids sequence of TCSV generated from uniprotkb



Number of amino acids: 184

Molecular weight: 20396.96

Theoretical pI: 6.92

Amino acid composition: CSV format

Ala (A)	10	5.4%
Arg (R)	7	3.8%
Asn (N)	12	6.5%
Asp (D)	5	2.7%
Cys (C)	1	0.5%
Gln (Q)	10	5.4%
Glu (E)	16	8.7%
Gly (G)	11	6.0%
His (H)	4	2.2%
Ile (I)	9	4.9%
Leu (L)	16	8.7%
Lys (K)	14	7.6%
Met (M)	1	0.5%
Phe (F)	5	2.7%
Pro (P)	8	4.3%
Ser (S)	20	10.9%
Thr (T)	13	7.1%
Trp (W)	5	2.7%
Tyr (Y)	1	0.5%
Val (V)	16	8.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 21
Total number of positively charged residues (Arg + Lys): 21

Atomic composition:

Carbon	C	899
Hydrogen	H	1438
Nitrogen	N	254
Oxygen	O	283
Sulfur	S	2

Formula: $C_{899}H_{1438}N_{254}O_{283}S_2$
Total number of atoms: 2876

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 28990
Abs 0.1% (=1 g/l) 1.421, assuming all pairs of Cys residues form cystines

Ext. coefficient 28990
Abs 0.1% (=1 g/l) 1.421, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 41.94
This classifies the protein as unstable.

Aliphatic index: 83.64

Grand average of hydropathicity (GRAVY): -0.503

Fig. 2. Physicochemical Characteristics of TCSV generated from ProtParam tool

```
          10         20         30         40         50         60         70
           |          |          |          |          |          |          |
MSFIGRLNTAEEEKAFHQQVASSNWICSVDVGSGVINSNPTLDFKVIPPTGGAVSVLTVSWENSTPQLVP
hhhhhhcccchhhhhhhhhhhhhcccceeeehhhcccccccceeeeecccccccceeeeeeecccccccccc
GHYLLRSGTWPVKNVKLSGLLVHRSVRLETTRKVLEQNKVSIAQQTENSVSDKGKTTETAWKFKEEIAHL
ceeeecccccccceeeeccheehhhhhhhhhhhhhhhhcccceecccccchhhhhhcccchhhhhhhhhhhhhh
NAELERARKEIAEKQSEISKLQLQLSNQPSNNDIFTGWSEDGPK
hhhhhhhhhhhhhhhhhhhhhhhhecccccccccceeeeeccccccc

Sequence length :    184

SOPMA :
    Alpha helix      (Hh) :     80 is   43.48%
    3_10  helix      (Gg) :      0 is    0.00%
    Pi helix         (Ii) :      0 is    0.00%
    Beta bridge      (Bb) :      0 is    0.00%
    Extended strand  (Ee) :     34 is   18.48%
    Beta turn        (Tt) :      0 is    0.00%
    Bend region      (Ss) :      0 is    0.00%
    Random coil      (Cc) :     70 is   38.04%
    Ambiguous states (?)  :      0 is    0.00%
    Other states          :      0 is    0.00%
```

Fig. 3. Secondary structure prediction obtained from Sopma

3.4 Subcellular Localization

Fig. 4. showed the Subcellular localization of TCSV was predicted using CELLO shows the various percentage of protein distributions that the protein could fall into.

3.5 Homology Modelling

The results presented in Fig. 5 below, shows different Homology Modelling results obtained from different Modelling tools.

Among the Models obtained from the different tools, model 1 of trRosetta is considered the best model due to the quality that was accessed using the Ramachandran plot.

3.6 Functional Annotation

A systematic workflow consisting of several Bioinformatics tools and databases was defined and used to perform functional annotation of TCSV i.e., NCBI-CDD, and InterProScan. It suggested that TCSV is a DNA double-strand break repair Rad50 ATPase. On a further study,

it was analysed that the protein is non-cytoplasmic in its domain.

4. DISCUSSION

The present study is focused on structural and functional annotation of Hypothetical protein Tomato Chocolate Spot Virus by searching using Uniprotkb server with accession number C7EXM3. Fig. 1 shows the amino acids sequence retrieved from uniprotkb.

ProtParam tools were used to analyze different physicochemical characteristics using the amino acid sequence. The hypothetical protein TCSV was analysed to have 184 amino acids with a molecular weight of 20396.96 Daltons and an isoelectric point of 6.92. An isoelectric point below 7 indicates a negatively charged protein. The instability index was predicted to be 41.94 which suggests an unstable protein. The negative GRAVY index of -0.503 indicates hydrophilic and soluble protein [20]. The protein sequence was found to be rich in amino acid serine which suggests a preference for alpha-helices in 3D structure [21].

Secondary structure characteristics were predicted using the Sopma server. The protein was analysed to contain a high proportion of Helices (Hh) at 43.48%, Extended strands (Ee) at 18.48% and Random coils (Cc) at 38.04%. The percentage of Helices (Hh0 in the structure makes the protein more flexible for folding, which might increase protein interactions [22,23].

Subcellular localization is a very important property of proteins. Cellular functions are often localized in specific compartments. Therefore, predicting the subcellular localization of unknown proteins could be used to obtained useful information about their functions as well as understanding disease mechanisms and developing drugs (Wang, 2004). The cello tool was used to determine the subcellular localization of the hypothetical protein. It suggests that the protein could both be a Nuclear and Cytoplasmic protein [24].

The main goal of protein modelling is to predict a 3D structure from the protein's sequence based on templates with an accuracy that is comparable to the best results achieved experimentally. For this purpose, four (4) different homology tools were used to predict the 3D structure which includes trRosetta, Lomet, RaptorX and IntFOLD5. The quality of each model was checked by the Ramachandran plot obtained from PROCHECK [20]. For a model to be considered as a good model, it must have at least 90% of its residues in the most favourable regions. Ramachandran plot for trRosetta showed 92.6% of residues in the most favourable region and 7.2% residues in additional allowed regions. Suggesting the model of trRosetta is the best among the models obtained. The final model was deposited in PDMB and is available under I.D: PM0084181.

CELLO RESULTS

SeqID: tr|C7EXM3|

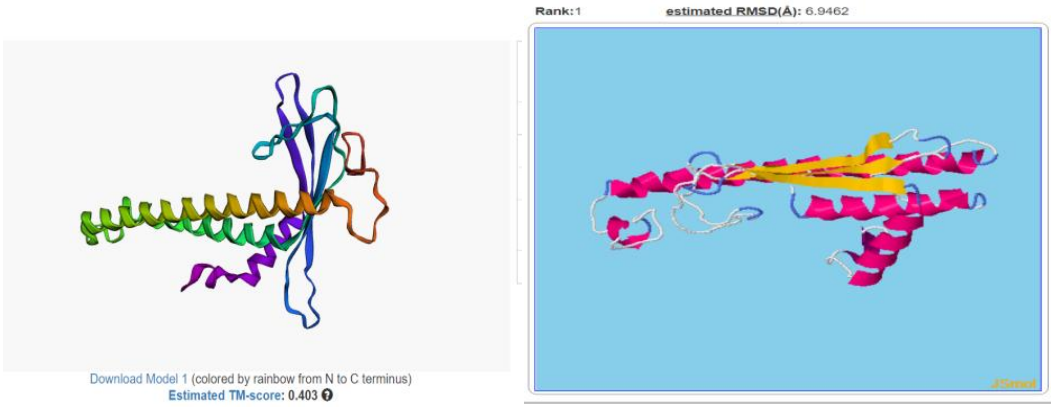
Analysis Report:

SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	Nuclear	0.480
N-peptide Comp.	Cytoplasmic	0.652
Partitioned seq. Comp.	Nuclear	0.533
Physico-chemical Comp.	Chloroplast	0.315
Neighboring seq. Comp.	Cytoplasmic	0.457

CELLO Prediction:

Nuclear	1.635 *
Cytoplasmic	1.504 *
Mitochondrial	0.635
Chloroplast	0.615
Extracellular	0.204
PlasmaMembrane	0.108
Peroxisomal	0.093
Golgi	0.077
Vacuole	0.054
Cytoskeletal	0.029
ER	0.026
Lysosomal	0.019

Fig. 4. CELLO prediction of Subcellular Localization



(a) TrRosetta model (C) RaptorX model

(b) Lomet model d) IntFOLD5 model

Fig. 5. Models obtained from different Modelling tools

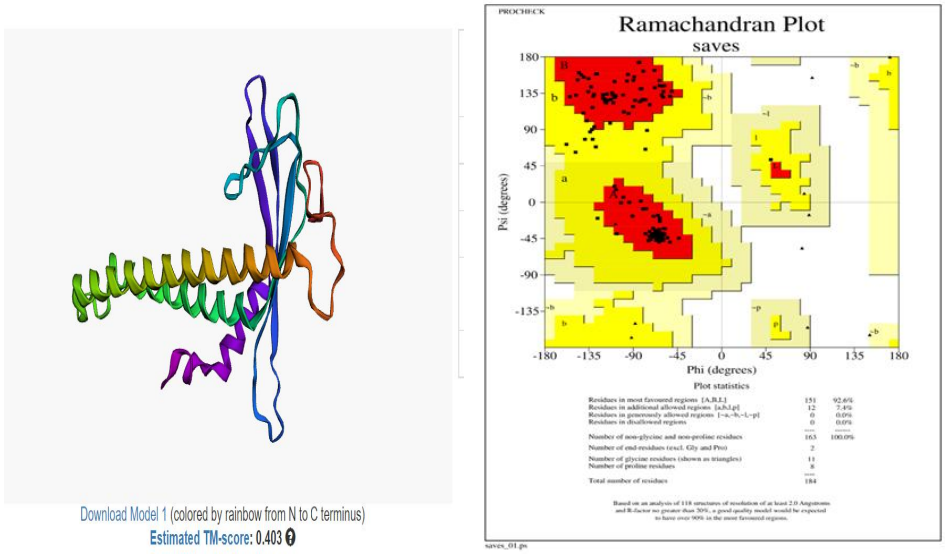
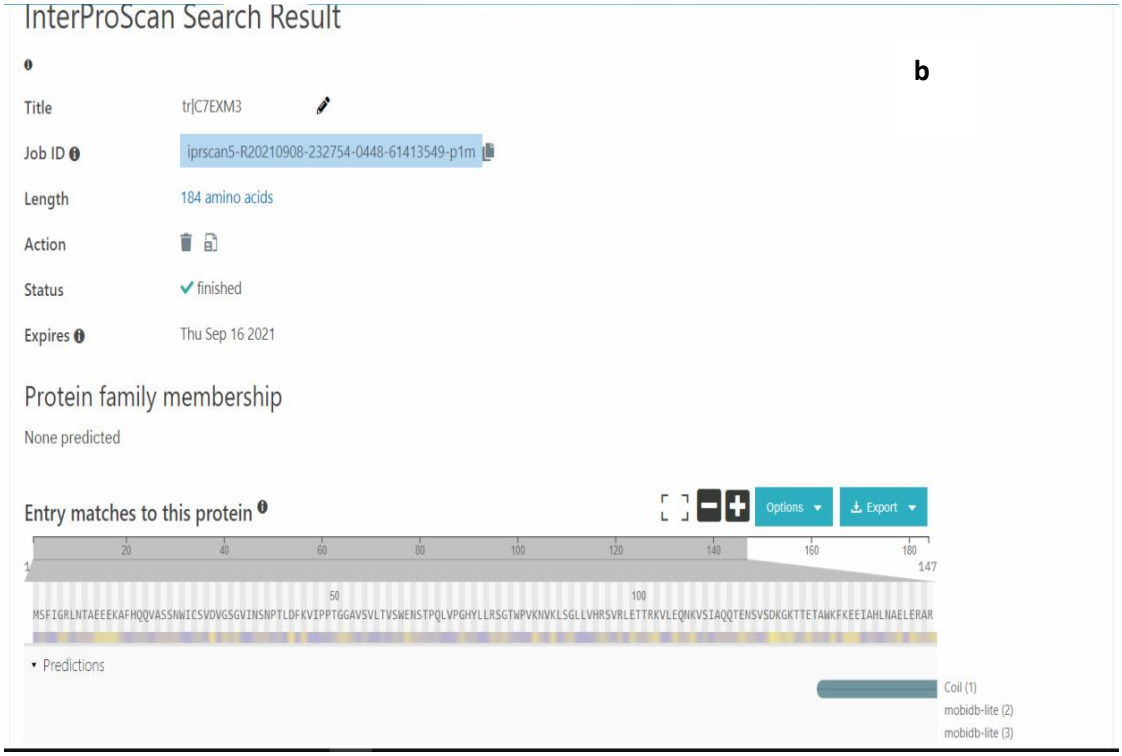


Fig. 6. Model of trRosetta and Ramachandran Plot obtained using PROCHECK


NCBI
Conserved Domains
SH3
SH2
a
HOME | SEARCH | GUIDE
NewSearch
Structure Home
3D Macromolecular Structures
Conserved Domains
Pubchem
BioSystems
Conserved domains on [tr|C7EXM3]
Local query sequence
View Concise Results
Graphical summary
Zoom to residue level
hide extra options «
Show site features
Horizontal zoom: ×
Update graph
Query seq.
Superfamilies
PRK02224
Search for similar domain architectures
Refine search
List of domain hits
Name
Accession
Description
Interval
E-value
PRK02224 super family
cl32023
DNA double-strand break repair Rad50 ATPase;
97-165
4.16e-04
DNA double-strand break repair Rad50 ATPase;
The actual alignment was detected with superfamily member PRK02224:
Pssm-ID: 179385 [Multi-domain] Cd Length: 880 Bit Score: 40.02 E-value: 4.16e-04
tr|C7EXM3 97 RLETTRKVLEQNKVSIAQQTEN--SVSDKGKTTET-AWKFKEEIAHLNAELERARKEIAEKQSEISKLQLQL 165
Cdd:PRK02224 322 RDEELRDRLEECRVAAQAHNEEaeSLREDADDLEErAEELREEAAELESELEEAREAVEDRREEIEELEEEI 393
References:
Marchler-Bauer A et al. (2017), "CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.", Nucleic Acids Res.45(D)200-3.
Marchler-Bauer A et al. (2015), "CDD: NCBI's conserved domain database.", Nucleic Acids Res.43(D)222-6.
Marchler-Bauer A et al. (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins.", Nucleic Acids Res.39(D)225-9.
Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", Nucleic Acids Res.32(W)327-331.
Help | Disclaimer | Write to the Help Desk
NCBI | NLM | NIH


InterProScan Search Result
b
Title
tr|C7EXM3
Job ID
iprscan5-R20210908-232754-0448-61413549-p1m
Length
184 amino acids
Action
Status
finished
Expires
Thu Sep 16 2021
Protein family membership
None predicted
Entry matches to this protein
Options
Export
MSFIGRLNTAEEEKAFHQQVASSNWICSVDVGSGVINSNPTLDFKVIPPTGGAVSVLTVSWENSTPQLVPGHYLLRSGTWPVKNVKLSGLLVHRSVRLETTRKVLEQNKVSIAQQTENSVSDKGKTTETAWKFKEEIAHLNAELERAR
Predictions
Coil (1)
mobidb-lite (2)
mobidb-lite (3)

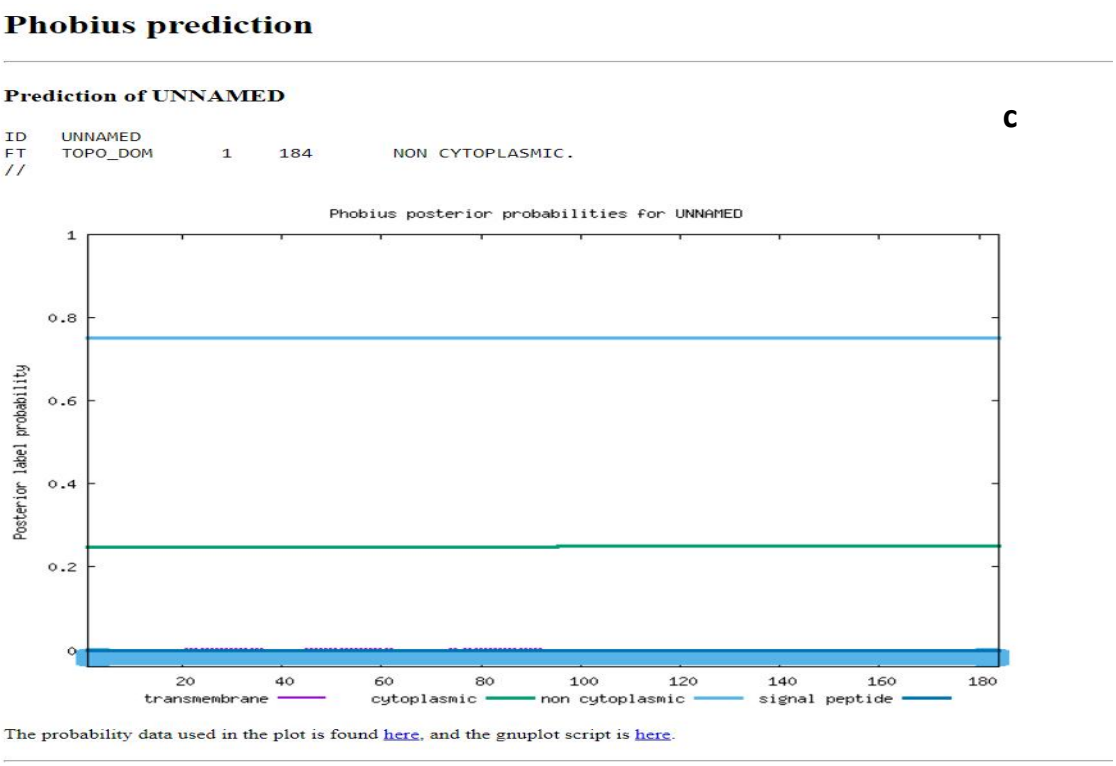


Fig. 7. The potential function of TCSV

Currently, there is no known function of TCSV. However, three web tools were utilized ~~to~~ to search the conserved domains and potential function of TCSV. Based on consensus prediction made by NCBI-CDD and interproScan (Fig 7a and 7b) suggested that TCSV is a DNA double-strand break repair Rad50 ATPase, which is involved in the early steps of DNA double-strand break (DSB) repair [25]. The complex may facilitate the opening of the processed DNA ends to aid in the recruitment of HerA and NurA [26]. Furthermore, the Phobius server (Fig 7c) predicted the protein to be a non-cytoplasmic in its domain, which means they help target proteins to their final destinations, as they contain a range of targeting signals that function to direct them along a targeting pathway [27].

5. CONCLUSION

The study has helped in obtaining the 3D structure of the protein from different Modelling tools. Model from trRosetta was selected as the best model based on the quality that assessed using PROCHECK of having more than 90% of amino acids in the most favourable region. It has also helped in understanding the function of the protein as well as the main target of the protein. In addition, the workflow described in this study can be used for other Hypothetical proteins.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Verbeek M, Dullemans AM. First Report of Tomato torrado virus infecting tomato in Colombia. Plant disease. 2012;96(4):592.
2. Batuman O, Kuo YW, Palmieri M, Rojas MR, Gilbertson RL. Tomato chocolate spot virus, a member of a new torradovirus species that causes a necrosis-associated

disease of tomato in Guatemala. Archives of virology. 2010;155(6):857-69.
3. Budziszewska M, Wieczorek P, Obrępalska-Stęplowska A. One-step reverse transcription loop-mediated isothermal amplification (RT-LAMP) for detection of tomato torrado virus. Archives of virology. 2016;161(5):1359-64..
4. Ferriol I, Junior DS, Nigg JC, Zamora-Macorra EJ, Falk BW. Identification of the cleavage sites of the RNA2-encoded polyproteins for two members of the genus Torradovirus by N-terminal sequencing of the virion capsid proteins. Virology. 2016; 498:109-15.
5. Iuzzolino L, McCabe P, Price SL, Brandenburg JG. Crystal structure prediction of flexible pharmaceutical-like molecules: density functional tight-binding as an intermediate optimisation method and for free energy estimation. Faraday discussions. 2018;211:275-96.
6. Maia RT, de Araújo Campos M, de Moraes Filho RM. Introductory Chapter: Homology Modeling. Homology Molecular Modeling: Perspectives and Applications. 2021:3.
7. Haddad Y, Adam V, Heger Z. Ten quick tips for homology modeling of high-resolution protein 3D structures. PLoS computational biology. 2020;16(4): e1007449.
8. Verli H. Bioinformática: da biologia à flexibilidade molecular; 2014.
9. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. Protein identification and analysis tools in the ExPASy server. Methods Molecular Biology. 1999;112:531-52
10. Geourjon C, Deleage G. SOPMA: Significant improvements in protein secondary structure prediction by prediction from multiple alignments. Comput Applic Bioci. 1995;11:681-684.
11. Yu NY, et al. Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 2010;26:13.
12. Ivan A, Hahnbeom P, Naozumi H, David E et al. Protein Tertiary Structure Prediction and Refinement using Deep and Rosetta in CASP14. Protein structures. 2021
13. Sitao W and Yang Z. LOMETS: A Local Meta-threading-server for Protein Structure Prediction. Nucleic Acid Research. 2007; 35(10):3375-82
14. Morten K, Haipeng W, Sheng W, Jian P, Hui L and Jinbo X. Template-based Protein Structure Modelling Using the RaptorX web. Nature Protocols. 2012;7,1511-22
15. Mc Guffin LJ, Adiyaman R, Maghrabi AHA, Shuis AN, Brackenridge DA, Nealon JO, Philomina LS. IntFOLD: an integrated web resource for high performance protein structure and function prediction. Nucleic Acids Research. 2019;47(1): 408–13
16. Laskowski RA, MacArthur MW, Thornton JM. PROCHECK: validation of protein-structure coordinates. Wiley Online Library. 2006;25(2):722-25
17. Finn R. D. et al. The Pfam protein families database. Nucleic Acids Residues. 2010; 38:211
18. Marchler-Bauer A. et al. NCBI's conserved domain database. Nucleic Acids Residues. 2011;39:D225.
19. Wang G, Dunbrack V. Scoring profile-to-profile sequence alignments. Protein Science. 2004;13:1612.
20. Butt AM, Batool M, Tong Y. Homology modeling, comparative genomics and functional annotation of Mycoplasma genitalium hypothetical protein MG_237. Bioinformation. 2011;7(6):299-303
21. Guruprasad K. Reddy B. V., Pandit M. W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Engineering. 2010;4:155-61.
22. Ikai A. J. Thermo stability and aliphatic index of globular proteins. Journal of Biochemistry. 2012;88:1895-98.
23. Gali IA. In-silico Structural Annotation of Methylthioadenosine Nucleosidase Protein Zm00014a_031618 in Maize (Zea mays L). Journal of Applied Science International. 2019;21(3):1-8
24. Xiong E, Zheng C, Wu X, Wang W. Protein Subcellular Location: The Gap between Prediction and Experimentation. Plant Mol Biol.

DOI 10.1007/s11105-015-0898-2. 2015
25. Hopfner K P, Karcher A, Shin D, Fairley C, Tainer J A, Carney J P. Mre11 and Rad50 from Pyrococcus furiosus: cloning and biochemical characterization reveal an evolutionarily conserved multiprotein machine. Journal of Bacteriology. 2000; 182(21):6036-41.

26. Ben, B. Hopkins, Tanya T Paull. The P. furiosus mre11/rad50 complex promotes 5' strand resection at a DNA double-strand break. Cell. 2008;135(2):250-60.

27. Kall L, Krogh A, Erik L, Sonnhammer L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. Journal of Molecular Biology. 2004;338(5):1027-36

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle4.com/review-history/74719