

# Secondary Structure-IOP

*by* Jamsari Jamsari

---

**Submission date:** 07-Aug-2020 04:00PM (UTC+0800)

**Submission ID:** 1366883916

**File name:** Hidayati\_2020\_IOP\_Conf-Naskah.pdf (654.67K)

**Word count:** 2214

**Character count:** 12218

PAPER • OPEN ACCESS

## Prediction of secondary, tertiary protein structure and interaction eIF4E from *Capsicum annum* with Rep *Geminivirus*

To cite this article: R Hidayati *et al* 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **497** 012026

View the [article online](#) for updates and enhancements.

## Prediction of secondary, tertiary protein structure and interaction eIF4E from *Capsicum annum* with Rep Geminivirus

R Hidayati<sup>1</sup>, B Nova<sup>2</sup>, and J Jamsari<sup>1,3\*</sup>

<sup>1</sup> Biotechnology Postgraduate Program, Universitas Andalas, Padang, West Sumatra, 25163, Indonesia

<sup>2</sup> Doctoral Program of Agriculture Science, Departement of Agriculture, Andalas University, Padang, West Sumatra, 25163, Indonesia

<sup>3</sup> Department of Agrotechnology, Universitas Andalas, Padang, West Sumatra, 25163, Indonesia

E-mail: ajamsari@yahoo.com

**Abstract.** The *eIF4E* gene is a gene that plays a role in the initiation of protein translation. This gene is relatively conserved and is present in all organisms. In humans, this gene functions as a proto-oncogene, its expression and activation are associated with transformation and tumorigenesis. The eIF4E protein takes a role in the translation initiation by recruiting ribosomes to the 5'-cap structure. In some conditions, it functions as an initiator in the process of translation and co-factor of the plant's defense system from virus attacks. The eIF4E protein also can bind with the genome linked viral protein to turnip mosaic virus [TuMV] in *Capsicum annum*. Their interaction is effective for virus infectivity and upregulated genome amplification. Understanding the tertiary structure of the eIF4E protein and its ligands will help in elucidating its interactions with viruses. So that it can be used to avoid spreading the virus in chili plants.

Keywords: eIF4E, *Capsicum annum*, interaction, tertiary structure, ligand

### 1. Introduction

Eukaryote initiating factor 4E [*eIF4E*] gene generally has a function as an initiation in the process of protein translation [1]. Besides, this gene plays a role in resistance to viruses. In humans, it acts as a proto-oncogene. This gene expresses and activates when bound to tumorigenesis [2]. The gene-encoded protein roles in translation initiation by recruiting ribosomes to the 5'-cap structure [3].

The *eIF4E* gene is found not only in humans but almost in all organisms. This is because the *eIF4E* gene is conserved and plays roles in many gene expressions of almost all organisms. In plants, apart from being the initiation of protein translation this gene functions as a co-factor in the plant's defense system [4]. Known to tomatoes, this gene helps in resistance to Potyvirus and Tomato Yellow Leaf Curl Virus [TYLCV] [5][6]. In chili, it is known to be associated with resistance to Potato virus Y and PVMV and others [7][8].



The virus *Replicase [Rep]* gene has an essential function in the multiplication of *geminivirus* particles in the host. Expression of this gene enables the virus to widespread plant tissues which cause leaves to change to churlish and yellowish. However, reports describing the interaction between eIF4E and the *Rep* gene have ever been carried out so far.

This paper was aimed to give an overview of the prediction of interaction between the eIF4E protein and the Rep protein *geminivirus*.

## 2. Methods

### 2.1. Collecting genetic data from NCBI.

In this study, the data was retrieved from the National Center for Biotechnology Information [NCBI] [9]. The genetic data was taken from the amino acid sequence data of chili with accession numbers CBL94660 isolated from *Capsicum annum*. The Rep protein was chosen from sequence data of *Geminivirus* collected from Tanah Datar, West Sumatra [genebank ID AMO26189.1]. All sequence data were mined in fasta format.

### 2.2. Secondary structure prediction

Prediction of secondary structure was performed to determine the estimate conformation structure of proteins. This purpose can be achieved by looking at the helix and sheet patterns. The prediction generated using the ProFunc software from EMBL-EBI [<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/profunc/>] [10].

### 2.3. Tertiary structure prediction

Prediction of tertiary structure is carried out for 3D protein modeling. The produced model is used to determine the ligand types that can interact with this protein. The tertiary structure was generated using Swiss-Model software [<https://swissmodel.expasy.org/interactive>] [11].

### 2.4. Protein interactions prediction

The obtained 3D model was saved .pdb format which is available for the next step analysis. The PatchDock [<https://bioinfo3d.cs.tau.ac.il/PatchDock/>] [12] was used in this study. This was run by uploading a protein modeling file in .pdb format for both eIF4E and Replicase.

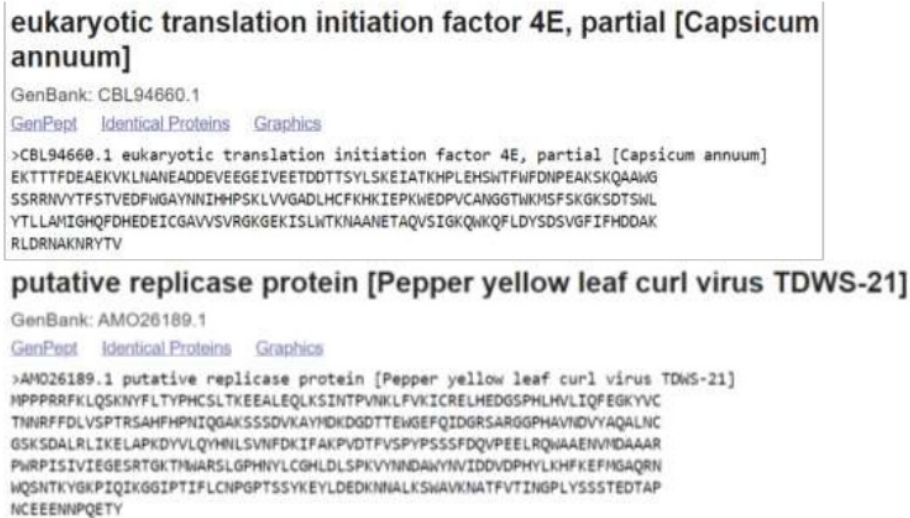
### 2.5. Interaction visualization

The interaction model obtained from the PatchDock software was visualized with the Chimera UCSF application [13]. This application enables us to see, rotate, interact with proteins and so on. Using the application also makes it possible to adjust the appearance, color, and position of proteins.

## 3. Results and discussion

### 3.1. Collecting genetic data from NCBI.

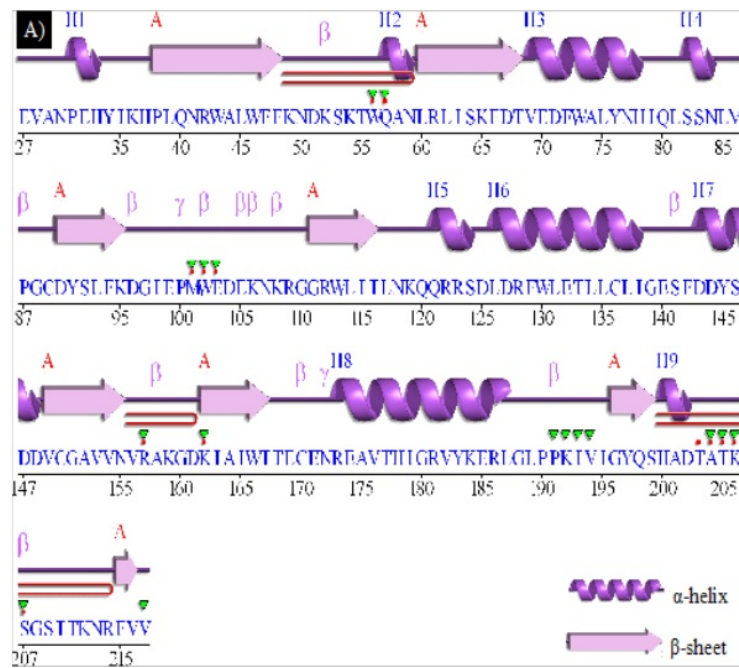
The data of *Capsicum annum* and Replicase [Rep] were successfully mined from the NCBI. They are shown in Figure 1. Protein Rep was chosen from the accession number because it encodes for Rep *Geminivirus* protein in West Sumatra [14]. Then the two data proceed to the next step; prediction of protein secondary structure.

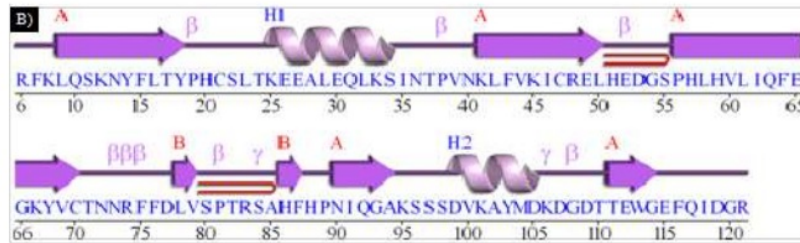


**Figure 1.** The amino acid sequence of *C. annum eIF4E* and replicase protein used in this study.

### 3.2. Secondary structure prediction

Before uploading the PDB code in the protein prediction, the inserted sequences must be checked as amino acids, instead of DNA or mRNA.



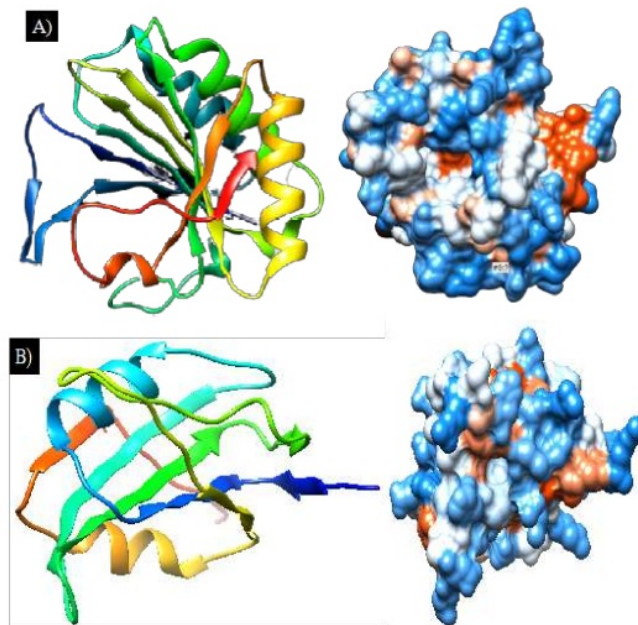


**Figure 2.** The appearance of secondary structure prediction results from ProFunc; A) eIF4E *Capsicum annum* and B) Rep Protein.

Figure 2 shows a comparison of the number of helices and sheets between the eIF4E *Capsicum annum* and the Rep Protein structure. The eIF4E protein possesses more  $\beta$ -sheet than the  $\alpha$ -helix that in turn will affect protein conformation [15] [16]. Structures contain more  $\beta$ -sheets will usually have more curves than fewer sheets. In contrast, the Rep protein has more  $\alpha$ -helix than the  $\beta$ -sheet. Further analysis will have proceeded to the modeling of tertiary [3D] protein structure of eIF4E *C. Annum* and Rep *Geminivirus* protein using Swiss-Model.

### 3.3. Tertiary structure [3D] prediction

The 3D structure prediction using Swiss-Model software is presented in Figure 2 A-B.



**Figure 3.** Results of prediction of tertiary structure [3D] protein; A) eIF4E *C. annum* and B) *Geminivirus* Rep using Swiss-Model software. Left is ribbon color; right is hydrophobicity surface display using UCSF Chimera.

Visualization of the ribbon structure shows the only structure of the  $\alpha$ -helix and  $\beta$ -sheet. The hydrophobicity visualization shows blue color representing hydrophilic sites. That site could bind with a compound, molecule or ligand having similar characteristics [hydrophilic]. Red color represents the surface of hydrophobic which is only bound with hydrophobic molecules. Site filling with white color is likely could bind with a neutral molecule [17]. Based on this data, an analysis was further continued to docking prediction using PatchDock software.

### 3.4. Prediction of protein interaction

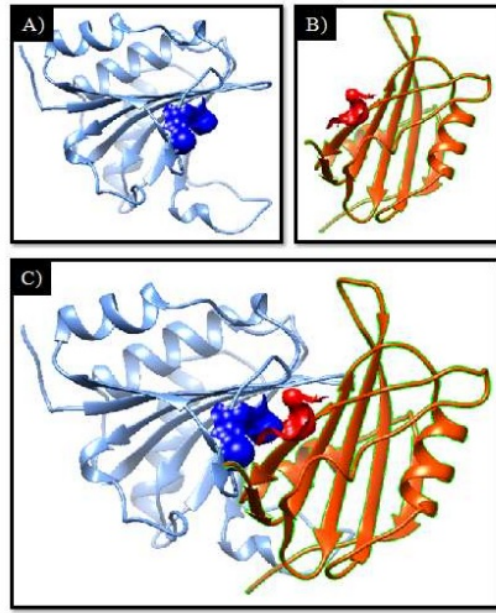
In order to predict protein interaction, the position of the binding should be performed first. The binding site prediction was performed using CASTp software by uploading the .pdb file for each protein data [18]. The interaction probability was simulated using with PatchDock software, based on binding energy in curves and caves [Figure 4].



**Figure 4.** PatchDock software interaction estimation between eIF4E and Rep protein.

Figure 4 listed score value of interaction probability between protein eIF4E *C. annuum* and Rep protein of Geminivirus. The highest score indicated the most likely interaction possibility. The data also presented a score of energy and prediction of the binding area. Based on the estimation generated by the PatchDock score with 15336 was chosen as a model of interaction.

The site with red and blue colors [Figure 4.C indicates the binding site]. The two binding site surfaces are close together, meaning interaction between both molecules is most likely could happen. The mutation on the site of both molecules could change their interaction capability. Based on this idea, plant resistance in this case *Capsicum annuum* could be improved through this hypothesis. However, this data has to be proved by empirical data using experimental research. The targeted mutation or any change could be done by applying some genome editing approach, for instance, the CRISPR/Cas9 platform or any site-directed mutagenesis technique.



**Figure 5.** Binding prediction between eIF4E and Rep. protein Geminivirus.

#### 4. Conclusion

The prediction of secondary and tertiary structures, as well as the prediction of the interaction of the eIF4E *C. annuum* protein with the Rep *Geminivirus* protein, was successfully carried out. This can be an initial prediction of the success of experimental research. Modification of the eIF4E gene could be possible to inhibit or block the replication of *Geminivirus* particles by editing the eIF4E protein binding region.

#### References

- [1] Goodfellow I G, L O Roberts. 2008. Eukaryotic initiation factor 4E. *Int J Biochem Cell Biol.* **40** 2675-80.
- [2] Feoktistova K, E Tuvshintogs, A Do, C S Fraser. 2013. Human eIF4E promotes mRNA restructuring by stimulating eIF4A helicase activity. *Proc Natl Acad Sci U S A.* **110** 13339-44.
- [3] Rhoads R E. 2009. EIF4E: New family members, new binding partners, new roles. *J Biol Chem.* **284** 16711-5.
- [4] Ruffel S, C Caranta, A Palloix, V Lefebvre, M Caboche, A Bendahmane. 2004. Structural analysis of the eukaryotic initiation factor 4E gene controlling potyvirus resistance in pepper: Exploitation of a BAC library. *Gene.* **338** 209-16.
- [5] Ruffel S, J L Gallois, M L Lesage, C Caranta. 2005. The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene. *Mol Genet Genomics.* **274** 346-53.
- [6] Ruffel S, J L Gallois, B Moury, C Robaglia, A Palloix, C Caranta. 2006. Simultaneous mutations in translation initiation factors eIF4E and eIF (iso) 4E are required to prevent pepper vein mottle virus infection of pepper. *J Gen Virol.* **87** 2089-98.



- [7] Ruffel S, M H Dussault, A Palloix, et al. 2002. A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). *Plant J.* **32** 1067-75.
- [8] Arcibal E, K M Gold, S Flaherty, J Jiang, M Jahn, A M Rakotondrafara. 2016. A Mutant eIF4E Confers Resistance to Potato Virus Y Strains and is Inherited in a Dominant Manner in the Potato Varieties Atlantic and Russet Norkotah. *Am J Potato Res.* **93** 64–71.
- [9] Arnemann J. 2018. NCBI. In: *Lexikon Der Medizinischen Laboratoriumsdiagnostik*.
- [10] Laskowski R A, J D Watson, J M Thornton. 2005. ProFunc: A server for predicting protein function from 3D structure. *Nucleic Acids Res.* **33** W89–W93
- [11] Guex N, M C Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis.* **18** 2714-23.
- [12] Schneidman-Duhovny D, Y Inbar, R Nussinov, H J Wolfson. 2005. PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**. W363-7.
- [13] Pettersen E F, T D Goddard, CC Huang, G S Couch, D M Greenblatt, E C Meng, T E Ferrin. 2004. UCSF Chimera-A Visualization System for Exploratory Research and Analysis. **25** 1605-12.
- [14] Jamsari J, J Pedri. 2013. Complete nucleotide sequence of DNA A-like genome and DNA- $\beta$  of monopartite Pepper Yellow Leaf Curl Virus, a dominant begomovirus infecting *Capsicum annuum* in West Sumatera Indonesia. *Asian J Plant Pathol.* **7** 1-14.
- [15] Laskowski R A. 2017. The proFunc function prediction server. In: *Methods in Molecular Biology.* **1611** 75-95
- [16] Fujiwara K, H Toda, M Ikeguchi. 2012. Dependence of  $\alpha$ -helical and  $\beta$ -sheet amino acid propensities on the overall protein fold type. *BMC Struct Biol.* **12** doi:10.1186/1472-6807-12-18
- [17] Giovambattista N, C F Lopez, P J Rossky, P G Debenedetti. 2008. Hydrophobicity of protein surfaces: Separating geometry from chemistry. *Proc Natl Acad Sci USA.* **105** 2274-9
- [18] Andrusier N, R Nussinov, H J Wolfson. 2007. FireDock: Fast interaction refinement in molecular docking. *Proteins Struct Funct Genet.* **69** 139-59.

#### Acknowledgments

This research was fully funded by the Andalas University Post Graduated for Thesis-Grants, through contract number: B 4/UN.16.16 DIR/PT.01.03/2019 year dated 2019.

# Secondary Structure-IOP

---

## ORIGINALITY REPORT

---

**2%**

SIMILARITY INDEX

**1%**

INTERNET SOURCES

**3%**

PUBLICATIONS

**0%**

STUDENT PAPERS

---

## MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

---

1%

★ "Advances in Computational Biology", Springer  
Nature, 2014

Publication

---

Exclude quotes      On

Exclude bibliography      On

Exclude matches      < 1%