RESEARCH ARTICLE

OPEN ACCESS

Impact of the N-terminal Fragment on the Solubility of eIF4E1 *Solanum tuberosum*

Viktoria V. Kolesnikova[1], Victoria Yu. Korchinskaya[2,3], Oleg S. Nikonov[1], Vladimir V. Andreitsev[1], T. Phat[4] and Ekaterina Yu. Nikonova[1,*]

[1]Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290, Russia
[2]All-Russia Research Institute of Agricultural Biotechnology, Russian Academy of Sciences, Moscow, 127550, Russia
[3]K.A. Timiryazev Institute of Plant Physiology, Russian Academy of Sciences, Moscow, 127276, Russia
[4]Institute of Biotechnology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

Abstract:

Introduction: Obtaining soluble eukaryotic proteins using bacterial expression systems remains a significant challenge. Despite the availability of various techniques to optimize protein expression in *E. coli*, eukaryotic proteins are frequently expressed in an insoluble form when produced in prokaryotic cells.

Methods: Genes of interest were cloned into expression vectors, pET22b, and modified pET32a, using the restriction-ligation method. BL21(DE3), BL21(DE3) Star, BL21(DE3)pRARE, Tuner(DE3), Origami, SHuffle®T7, and C41(DE3) strains were used for analytical induction. Molecular modeling and molecular dynamics simulations were employed to design the experiments and to interpret the resulting data.

Results and Discussion: To prevent the accumulation of a truncated form of potato eIF4E in inclusion bodies, we tested various *E. coli* strains and repositioned affinity tags from the C-terminus to the N-terminus of the protein. Only the full-length eIF4E was found to be soluble in the prokaryotic expression system. Based on the eIF4E model and molecular dynamics simulations, we proposed a potential explanation for the impact of the N-terminal fragment on protein solubility.

Conclusion: The interaction between the N-terminal fragment and the dorsal surface of eIF4E may prevent protein aggregation. This shielding of hydrophobic regions appears to be a key factor in reducing aggregation, thereby facilitating the expression of eIF4E in a soluble form.

Keywords: Protein purification, Cap-dependent translation initiation, Translation initiation factors eIF4E, eIF4F, Potyviruses.

*Address correspondence to this author at the Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290, Russia; E-mail: katya_nik@vega.protres.ru

1. INTRODUCTION

The accumulation of proteins in inclusion bodies remains a significant challenge for many recombinant proteins produced in *E.coli*. Numerous protocols have been developed for the solubilization of proteins from inclusion bodies [1, 2, 3, 4, 5]. The isolation of proteins from inclusion bodies is a process of denaturation and their subsequent renaturation. Cotranslational protein folding is much more efficient than *in vitro* refolding. It is not clear whether the protein retains its functional activity after refolding or not. Additional tests should be performed to confirm the activity of the obtained protein. There are several reasons why proteins accumulate in inclusion bodies during synthesis, such as protein aggregation, improper formation of inter- or intramolecular disulfide bonds, toxicity of target proteins, and overexpression due to strong induction. Different strategies have been developed to obtain proteins in a soluble form [6, 7, 8].

The amino acid sequence at the N-terminus of a recombinant protein can determine its propensity for aggregation and accumulation in inclusion bodies. For example, charged amino acids may enhance protein solubility by forming hydrogen bonds, whereas hydrophobic residues tend to promote aggregation. To prevent the accumulation of protein in the inclusion bodies, various techniques can be used, for example, adding a signal peptide to the N-end of the protein [9, 10]. The signaling peptide promotes protein transport to a specific location in the cell. Various modifications at the N-terminal part, such as acetylation or methylation, can lead to changes in the interaction of the molecule with the solvent, which also affects the solubility of the protein [11, 12, 13]. However, since posttranslational modifications occur predominantly in eukaryotic systems, their relevance is limited in bacterial cells. The co-expression of target protein and N- or C-terminal tags, such as thioredoxin, SUMO, or MBP, can also affect the accumulation of proteins in soluble form [14, 15, 16, 17]. Moreover, sometimes, the movement of the tag from the C to the N-terminus can have a positive effect on protein solubility [18, 19, 20]. The presence of 6His at the N- or C-terminus can affect the thermal stability of the protein, which is an important indicator when working with recombinant proteins [21].

During the heterologous expression of the translation initiation factor 4E (eIF4E) from Solanum tuberosum in a bacterial system, we encountered a major obstacle: the accumulation of the protein in inclusion bodies. eIF4E is a cap-binding protein that is involved in the initiation of translation in eukaryotic cells. It is a component of the eIF4F complex, which binds to the mRNA cap and attracts the 43S pre-initiation complex [22, 23, 24, 25]. eIF4E is represented by a family of proteins. This family is one of the factors that makes Solanaceae plants susceptible to potyviruses. However, there are no experimentally determined structures of eIF4E proteins of Solanaceae. The interaction of proteins of the eIF4E from the Solanaceae family with VPg of potato Y virus is a necessary condition for the development of viral infection [26, 27, 28]. There is no experimental data confirming the mechanism of interaction between eIF4E and VPg, making structural studies crucial.

The significant impact of potyviruses on crop yields also makes it an urgent priority to study the mechanism of potyviral RNA translation. Solving the structure of the eIF4E-VPg PVY complex would give an obvious interpretation of the previously obtained data, explain the mechanism of interaction between the two proteins, and predict mutations that could disrupt this interaction. We need to obtain a large amount of functionally active recombinant eIF4E for protein crystallization, as well as for functional studies. However, the tendency of eIF4E to accumulate in inclusion bodies makes obtaining active protein multistage and inefficient. To address this challenge, we have conducted analytical induction of different variants of potato protein eIF4E1 in various *E. coli* expression strains in order to obtain this protein in a soluble form.

2. MATERIALS AND METHODS

2.1. Nucleotide Sequences

The nucleotide sequence of cDNA eIF4E1 was used to create expression vectors (Genbank accession: MT828879).

2.2. Molecular Modeling and Molecular Dynamics

The eIF4E1-49aa model was obtained as previously described [29, 30]. Alphafold3 was used to obtain full-length eIF4E1 models [31].

Molecular dynamics studies of the obtained theoretical model of the full-size eIF4E1 were carried out using the Gromacs 2022.3 software package [31, 32, 33, 34] and the all-atomic force field Charmm36 (July 2022 version: CHARMM all-atom force field (July 2022)) [35]. The solution was modeled using a three-point CHARMM TIP3P (TIPS3P) water model. The system was equilibrated to achieve a constant temperature of 310 K and a constant pressure of 1 bar. The molecular dynamic trajectory was calculated in a time interval of 200 ns at a given constant pressure and temperature with a time step of 2 fs. Manual editing of the models was carried out in the Coot program [36].

2.3. Data Analysis and Visualization

The superposition and comparative analysis of experimental and theoretical eIF4E models were carried out in the software packages Coot and PyMOL (Version 2.5.0a0. OPEN-SOURCE). Clustering and analysis of the molecular dynamic trajectory were performed in the Gromacs and VMD software packages [37]. Clustering of MD trajectory frames was carried out for the Cα atoms of the model using the gmx cluster module and the GROMOS algorithm, applying a threshold value of rmsd=1.5A. As a result, 81 clusters were obtained with a maximum occupancy of 6240 frames out of 20,000 frames of the trajectory. For a full-length eIF4E1 potato model, a molecular surface with coloring according to the Wimley-White ΔG water-membrane hydrophobicity scale [38] was constructed and visualized in the MolStar program [39]. Visualization of the obtained data was carried out using the Inkscape vector graphics program (https://inkscape.org/) and the software complexes of molecular graphics PyMOL and MolStar.

2.4. Plasmid Design

Sequences of eIF4E1 (1-696 b.p.) and eIF4E1-49aa (145-696 b.p.) were amplified from cDNA of the *Solanum tuberosum* cultivar Zhukovskiy ranniy. A modified pET32a-GG-ΔTrx vector was used for cloning. The thioredoxin located between the two NdeI sites was removed; the BsaI recognition site after the enterokinase site was also added to the plasmid, while preserving the triplet encoding alanine between the enterokinase site and the BsaI site. The amplicons and vector were treated with BsaIHF-v2 (NEB) restrictase and then ligated using a highly active T4 DNA ligase by the Golden Gate method (according to the manufacturer's protocol). The introduction of restriction sites was carried out to obtain a more user-friendly vector. The nucleotide sequence of eIF4E1-49aa (145-696 b.p.) was subcloned into the expression vector pET22b at the restriction sites, NcoI and XhoI, before the sequence of the signal peptide. This signaling peptide is intended to facilitate the movement of the protein into the periplasmic space [9]. All plasmids were sequenced. Each of the resulting genetic constructions carries a 6His-tag to facilitate affine purification.

2.5. Analytical Induction and Protein Isolation

E. coli strains BL21(DE3) (Novagene), BL21(DE3) Star (Novagene), BL21(DE3)pRARE(Novagene) (Cm), Tuner (DE3) (Novagene), Origami (Novagene) (Kn+Tet), SHuffle ®T7(Novagene) (Str), and C41(DE3) (Novagene) were transformed by plasmid pET22b-eIF4E1-49aa(Amp) and grown on the selective agar-media during 16 h +37 °C. One colony from the agar plate was transferred into 10 ml selective Luria-Bertani (LB) medium and incubated in a shaker at +37 °C 180 vol.\min. When the optical density of OD_{590} reached ≈0.6–0.8, the inductor was added to the final concentration of 0.5 mM isopropyl-β-D-thiogalactoside (IPTG). Cells were incubated during 3 h at +37 °C or 20 h at +20 °C. The expression level was checked using SDS-PAGE. The cells were precipitated by centrifugation and then resuspended in buffer (10 mM Tris HCL pH7.5. 250 mM NaCl, 10% glycerol) with the addition of DNase and PMSF. The cell suspension was homogenized by sonication. To avoid overheating during sonication, a container with a suspension of cells was placed into watered ice. After homogenization, the suspension was centrifuged and the presence of protein in soluble and insoluble fractions was determined by electrophoresis. The target protein was verified by mass spectrometry.

The calculation of the expected molecular weight of the protein, the coefficient of extinction, and the isoelectric point was carried out using the Biochemistry Online tool (https://vitalonic.narod.ru/biochem/algorithm.htm).

3. RESULTS AND DISCUSSION

3.1. Overview of Protocols of eIF4E Isolation for Crystallization

To begin with, we analyzed purification protocols of human, mouse, yeast, nematode, wheat, pea, and melon eIF4E, which were successfully crystallized [40, 41, 42, 43,

44]. All mentioned eIF4E proteins were N-terminally truncated: 1-51aa (amino acids) were removed in melon eIF4E (UniProtKB Q00LS8), 1-50 aa were removed in pea eIF4E (UniProtKB Q0GRC4), 1-36 aa were removed in wheat eIF4E (UniProtKB P29557), 1-28 aa were removed in mouse eIF4E (UniProtKB P63073), and 1-30 aa were removed in nematode eIF4E (UniProtKB O61955). Purification protocols for these proteins are very similar. The *E. coli* strain Rosetta (DE3) pLysS (Novagen) was used for the expression of melon, pea, and wheat eIF4E, and the strains HB101(DE3) and BL21(DE3) Star were used for the expression of mouse and nematode eIF4E, accordingly. Expression strains with plasmids that encode rare tRNAs are often used to induce the synthesis of eukaryotic proteins in the prokaryotic system. All protocols used LB medium, IPTG at a concentration between 0.4 and 0.6 mM, and a sonication or a French press as the homogenization method. Nematode and melon eIF4E were coexpressed with the eIF4G fragment. Accumulation of the target protein in inclusion bodies and its isolation under denaturing conditions was reported only for the mouse N-terminally truncated eIF4E. Renaturation of this protein does not prevent its crystallization [42, 45].

3.2. eIF4E Modeling and Design of Expression Vectors

Mobile elements in the structure of proteins often prevent their crystallization. The N-terminal fragment, which was deleted in the mentioned eIF4E, is unstructured and mobile. It was shown in the structure of full-length eIF4E obtained in the NMR experiment (PDB code: 4B6U). The absence of this fragment in the recombinant protein should not affect its functional activity or the integrity of the globular part, but at the same time, the possible role of this fragment in the interaction of eIF4E with eIF4G or other molecules cannot be excluded.
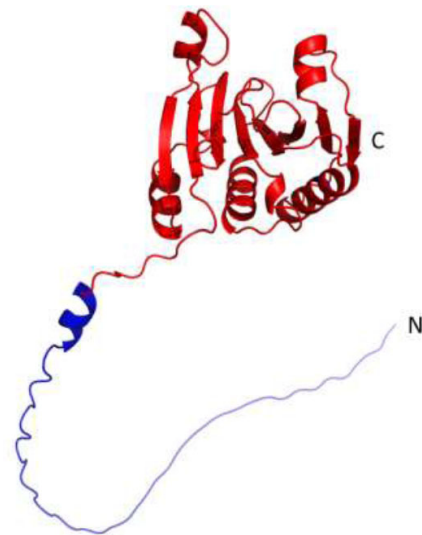

Fig. (1). The model of *Solanum tuberosum* eIF4E1. The blue color indicates the N-terminal fragment (1-48 aa), which is missing in the shortened version of the protein (eIF4E1-49aa).

In this regard, we also started working with a shortened version of eIF4E1 for the following crystallization. Using the previously obtained model, we determined the boundaries of the N-terminal mobile unstructured part and the globular part of eIF4E1: 1-48 and 49-233, accordingly (Fig. 1). The shortened version of eIF4E is referred to as eIF4E1-49aa in the following text. At the same time, for functional experiments, we needed both a full-length and a shortened version of the target protein. Therefore, we obtained several variants of vectors encoding both the full-length and shortened versions of eIF4E.

The sequences encoding eIF4E1-49aa and full-length eIF4E1 were cloned into the expression vector pET32a-GG-ΔTrx. As a result, pET32a-GG-ΔTrx-eIF4E1-49aa and pET32a-GG-ΔTrx-eIF4E1 plasmids were obtained. Using the first plasmid enabled us to get the target protein with an estimated molecular weight of 25.8 kDa. This variant includes a 6His-tag, thrombin cleavage site, S-tag, and ente-

rokinase cleavage site at the N-terminus (Fig. 2). Additionally, using the pET32a-GG-ΔTrx-eIF4E1 plasmid allowed us to get the target protein with an estimated molecular weight of 30.8 kDa. This variant of the protein contains the same N-terminal sequence as previously mentioned (Fig. 2). This design allows the removal of the tag during one of the purification steps.

For the expression of full-length eIF4E1 and eIF4E1-49aa, we used the BL21(DE3)-pRARE strain that has an additional plasmid encoding rare tRNAs [46]. One potential cause of protein accumulation in inclusion bodies is the misincorporation of amino acids, which can result in improper folding. Expression of full-length eIF4E1 in this strain during 3 h at +37 °C and 20 h at +20 °C led to the formation of a soluble form of full-length eIF4E1 (Fig. 3). In contrast, under the same conditions, eIF4E1-49aa accumulated in inclusion bodies.
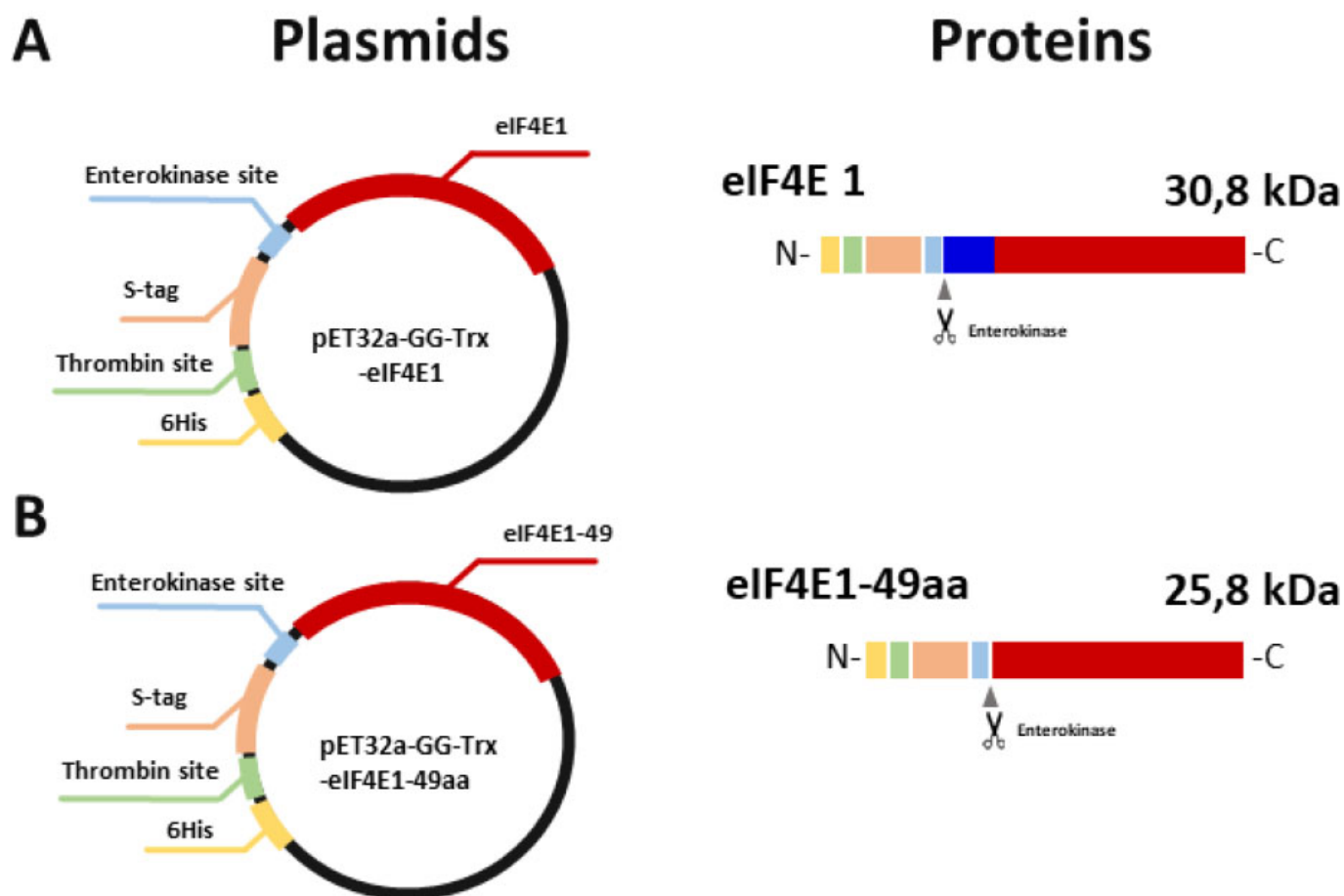

Fig. (2). Schematic representation of the expression plasmids and the primary sequence of the target proteins. Different colors on the circle indicate the main elements, located between the T7 promoter and the terminator. The same colors indicate the corresponding elements on the primary structure of the protein. Scissors indicate the place where the protease cut the protein. A – Scheme of the plasmid encoding full-length eIF4E1 and the protein primary sequence; B – Scheme of the plasmid encoding eIF4E1-49aa and the protein primary sequence.
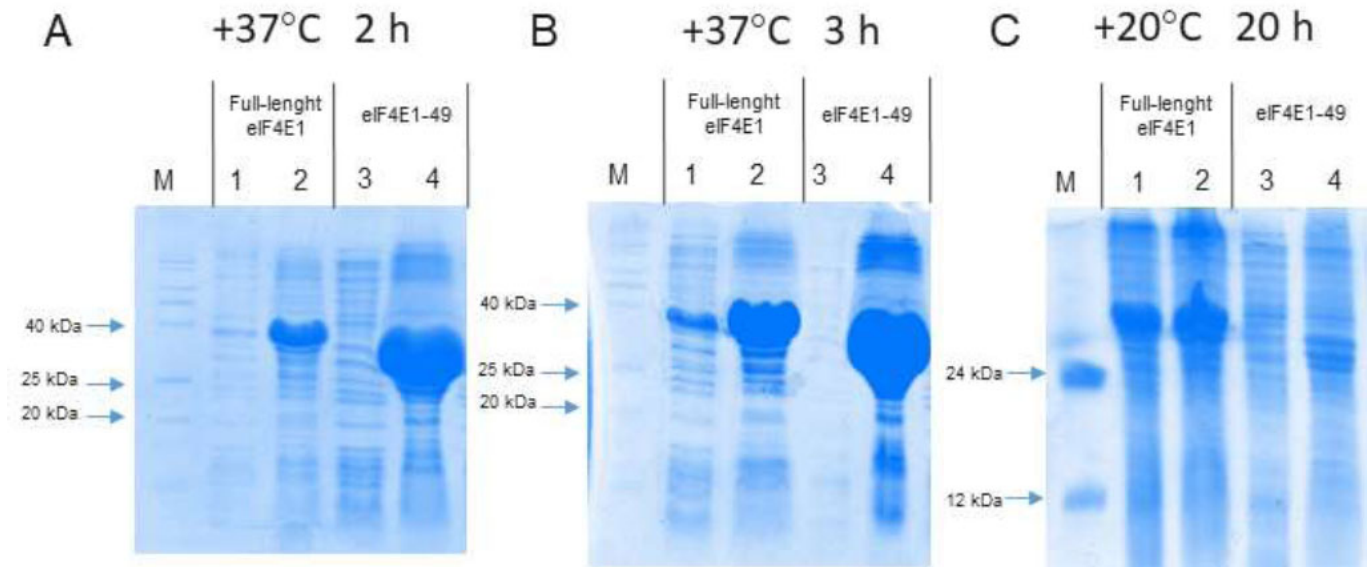
Fig. (3). Electrophoregrams of the samples obtained after analytical induction in the strain BL21(DE3) pRARE. Odd numbers correspond to samples of supernatant after cell sonication; even numbers correspond to samples of debris after cell sonication. 1 and 2 correspond to full-length eIF4E1; 3 and 4 correspond to eIF4E1-49aa; M – protein molecular weight markers. A – protein expression during 2 h at +37 °C; B – protein expression during 3 h at +37 °C; C – protein expression during 20 h at +20 °C.

Isolation of proteins from inclusion bodies is not an impossible task. However, this complicates the isolation process and does not guarantee that the target protein will refold correctly and will be functionally active. In this regard, we next modified the genetic construct to obtain eIF4E1-49aa in a soluble form. Since the presence of additional tags (despite the possibility of getting rid of them) is undesirable for crystallization experiments, we selected the expression vector pET22b, which does not code sequences of long tags and sites for protease digestion. The sequence encoding eIF4E1-49aa was cloned into this vector. As a result, the pET22b-eIF4E1-49aa plasmid was obtained. This construction enabled us to get the target protein with an estimated molecular weight of 22.3 kDa. This variant includes a 6His-tag at the C-terminus and 22 aa signal peptide at the N-terminus, which is expected to help the protein to be released into the periplasmic space (Fig. **4**).
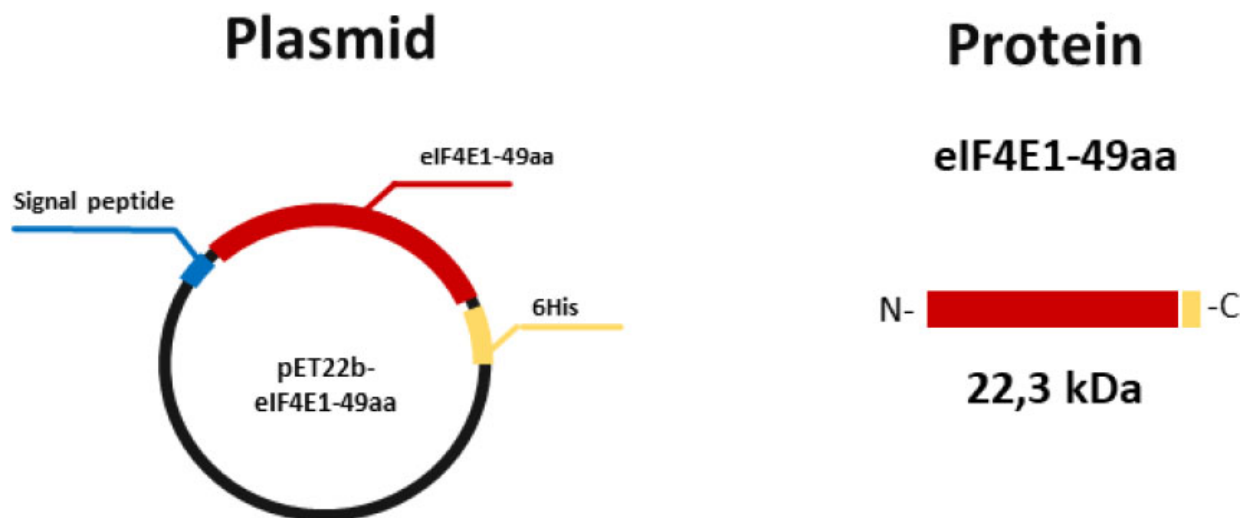

Fig. (4). Schematic representation of the expression plasmid pET22b and the primary sequence of eIF4E1-49aa. Different colors on the circle indicate the main elements, located between the T7 promoter and the terminator. The same colors indicate the corresponding elements on the primary structure of the protein.

In our attempts to get soluble SteIF4E1-49aa, we tried different *E. coli* strains and growth conditions. Analytical induction of synthesis of eIF4E1-49aa from pET22b plasmid in strains BL21(DE3) and BL21(DE3)-pRARE during 3 h at +37 °C and 20 h at +20 °C did not lead to the production of the soluble protein (Fig. 5).

Lowering the temperature slows down the process of protein synthesis, reducing the likelihood of translational errors and potentially improving protein solubility. However, in our case, it was ineffective. Another factor contributing to strong protein aggregation and, consequently, its accumulation in inclusion bodies is the presence of unpaired cysteine residues in the protein structure. The Shuffle and Origami *E. coli* strains were created to improve the accuracy of the formation of disulfide bonds and were also selected for our experiment [47, 48]. We also used the *E. coli* strain C41(DE3), which is recommended for the

expression of toxic proteins, and the strain Tuner (DE3), which is suitable for toxic and insoluble proteins. Unfortunately, all these strains did not yield a positive result for the short version of SteIF4E1 (Fig. 5).

The full-length version of eIF4E1 was produced in soluble form under standard conditions (Table 1) without the need for further optimization. The presence of any tags at the N- or C-end of the eIF4E1-49aa, lowering of the incubation temperatures, and using different *E. coli* strains had no positive effect on the solubility of eIF4E1-49aa (Table 1).

The sequence of SteIF4E1 N-terminal part (1-48 aa) (MATAEMERTTSFDAAEKLKAADAGGGEVDDELEEGEIVEE SNDTASYL) does not show any similarity with any known signaling peptides sequence. At the same time, experimental data suggest that the N-terminal fragment (1-48 amino acids) protects the protein from aggregation.
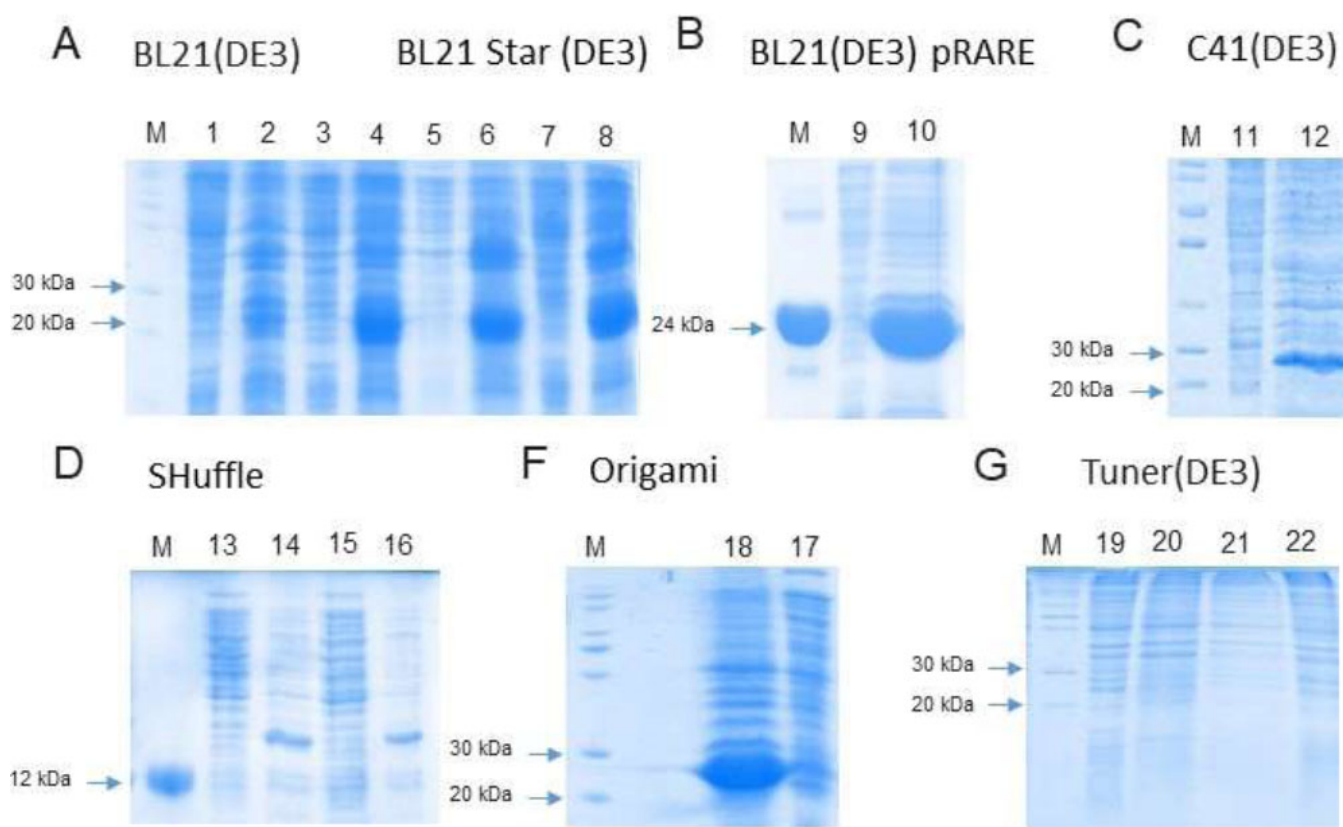

Fig. (5). Electrophoregrams of the samples obtained after analytical induction of the eIF4E1-49aa in different strains. Odd numbers correspond to samples of supernatant after cell sonication; even numbers correspond to samples of debris after cell sonication. M – protein molecular weight markers. A – 1, 2, 3, 4 – strain BL21(DE3): 1, 2 – protein expression during 3 h at +37 °C, 3, 4 – protein expression during 20 h at +20 °C. 5, 6, 7, 8, – strain BL21(DE3)Star: 5, 6 – protein expression during 3 h at +37 °C, 7, 8 – protein expression during 20 h at +20 °C, B – strain BL21(DE3) pRARE: 9, 10 – protein expression during 20 h at +20 °C; C – strain C41(DE3), 11, 12 – protein expression during 20 h at +20 °C. D – strain Shuffle: 13, 14 – protein expression during 3 h at +37 °C, 15, 16 – protein expression during 20 h at +20 °C. F – strain Origami: 18, 17 – protein expression during 20 h at +20 °C. G – strain Tuner(DE3): 19, 20 – protein expression during 3 h at +37 °C, 20, 21 – protein expression during 20 h at +20 °C.

Table 1. Comparison of the results of analytical induction.

Protein	Vector	Strain	Result	
			37 °C 3 h	20 °C 20 h
eIF4E1-49aa	pET22b	BL21(DE3)	Insoluble	Insoluble
		BL21(DE3) pRARE	Insoluble	Insoluble
		BL21(DE3)Star	Insoluble	Insoluble
		C41(DE3)	-	Insoluble
		SHuffle	There is no induction	
		Origami	-	Insoluble
		Tuner (DE3)	There is no induction	
eIF4E1-full	pET32a-GG-ΔTrx	BL21(DE3) pRARE	soluble	soluble
eIF4E1-49aa	pET32a-GG-ΔTrx	BL21(DE3) pRARE	Insoluble	Insoluble

Note: "-" – undetermined.

3.3. Structural Interpretation of the Influence of the N-Terminal Fragment on Protein Solubility

These data can be interpreted from a structural point of view using molecular modeling and molecular dynamics methods. A comparative analysis of potato eIF4E1 and homology protein models, both experimental and theoretical (including those generated using the latest version of the Alphafold3 algorithms), revealed the ability of different regions of the unfolded N-terminal fragment to form alpha helices.

According to the results of molecular modeling and molecular dynamics studies, this fragment is able to interact with the globular part of the protein, forming stable contact with a section of its dorsal surface (Fig. **6**). The dorsal surface of eIF4E forms part of the eIF4G binding site and is rather hydrophobic. The exposition of such kind of hydrophobic surfaces to a solvent (which is the case for the models with the truncated N-terminal part) could cause the aggregation of proteins in a water solution. The cluster analysis of the 200ns MD trajectory for the intact factor gives us the most stable conformation for the N-terminal part in this experiment. This example shows the ability of the N-terminal fragment to shield the hydrophobic region of the eIF4E globule surface, extending the area of the hydrophilic one. This could be the reason for the increased solubility of the intact protein compared to the truncated one. Thus, the above-mentioned interaction of a part of the N-terminal fragment with the eIF4E dorsal surface can prevent aggregation, preserving the protein in a soluble form.

If we take a look at the same region of eIF4E from other organisms, we will also find the hydrophobic spot (Fig. **1S**). The intensity of these regions can explain the tendency of the proteins to aggregate.

A. Cartoon representation of the most stable conformation of the model of full-length eIF4E1 from the 200ns molecular dynamic trajectory. The view from the cap-recognizing loop is on the left. The view from the eIF4G binding site is on the right. The gray color shows the globular part of the factor. The green color shows the N–terminal fragment of the factor, which was initially poorly structured and adopted a conformation during the MD experiment.
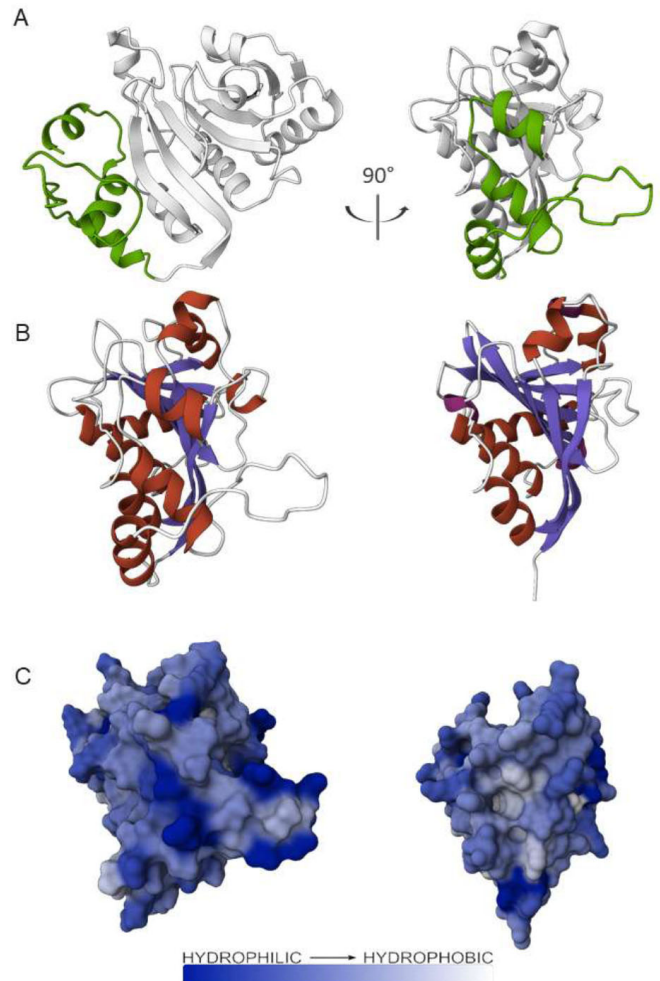

Fig. (6). Possible role of the N-terminal fragment in the extension of the hydrophilic surface of eIF4E1.

B. Cartoon representation of the models of full-length eIF4E1 (left) and eIF4E1-49aa (right), viewed from the back side. The elements of the secondary structure are colored: α-helixes are highlighted in red, β-strands are highlighted in slate, and loops are highlighted in white.

C. Surface representation of the models of full-length eIF4E1 (left) and eIF4E1-49aa (right) from the dorsal view. The surfaces are colored according to the Wimley -White ΔG water-membrane hydrophobicity scale [38].

CONCLUSION

The full-size version of eIF4E1 obtained in bacterial cells is soluble. The shortened version of eIF4E1 (eIF4E1 -49aa) accumulates in inclusion bodies. The presence of a hydrophobic region on the surface of the molecule provides a tendency for the protein to aggregate. This is due to the fact that the N–terminal fragment of the full-size eIF4E1 variant interacts with the hydrophobic surface of the globule, which makes it more hydrophilic, thereby preventing aggregation. Using the method of molecular modeling and molecular dynamics, it is possible to identify such hydrophobic interface regions and find a way to shield them. This approach will make it possible to obtain aggregation-prone proteins containing hydrophobic regions in a soluble form.

AUTHORS' CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: E.Y.N.: Study conception and design; V.V.K. and V.Y.K.: Data collection; P.T.D.: Methodology; O.S.N.: Analysis and interpretation of results; and V.V.A.: Visualization. All authors reviewed the results and approved the final version of the manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of the research.

CONSENT FOR PUBLICATION

Not applicable

AVAILABILITY OF DATA AND MATERIALS

All data generated or analyzed during this study are included in this published article.

FUNDING

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIAL

Fig. (**1S**). Surface representation of a few members of the eIF4E family from the dorsal view.

The surfaces are colored according to the Wimley–White ΔG water-membrane hydrophobicity scale [38]. A – Crystal structure of eIF4E from *Cucumis melo*; B – Crystal structure of eIF4E from *Drosophila melanogaster*; C – Crystal structure of eIF4E from *Mus musculus*; D – Crystal structure of eIF4E from *Pisum sativum*.

REFERENCES

[1]   Hashemzadeh MS, Mohammadi M, Ghaleh HEG, Sharti M, Choopani A, Panda AK. Expression, solubilization, refolding and final purification of recombinant proteins as expressed in the form of "classical inclusion bodies" in *E. coli.* Protein Pept Lett 2021; 28(2): 122-30.
http://dx.doi.org/10.2174/0929866527999200729182831   PMID: 32729411

[2]   Ryan BJ. Differential precipitation and solubilization of proteins. Methods Mol Biol 2017; 1485: 191-208.
http://dx.doi.org/10.1007/978-1-4939-6412-3_10

[3]   Ryan BJ, Kinsella GK, Henehan GT. Protein extraction and purification by differential solubilization. Methods Mol Biol 2023; 2699: 349-68.
http://dx.doi.org/10.1007/978-1-0716-3362-5_17 PMID: 37647006

[4]   Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins. J Biosci Bioeng 2005; 99(4): 303-10.
http://dx.doi.org/10.1263/jbb.99.303 PMID: 16233795

[5]   Singh A. Solubilization and refolding of inclusion body proteins. Methods Mol Biol 2015; 1258: 283-91.
http://dx.doi.org/10.1007/978-1-0716-1859-2_22 PMID: 35089569

[6]   Bochkarev A. High throughput crystallography at SGC Toronto: An overview. Methods Mol Biol 2008; 426: 515-21.
http://dx.doi.org/10.1007/978-1-60327-058-8_34 PMID: 18542887

[7]   Gopal GJ, Kumar A. Strategies for the production of recombinant protein in Escherichia coli. Protein J 2013; 32(6): 419-25.
http://dx.doi.org/10.1007/s10930-013-9502-5 PMID: 23897421

[8]   Hayat SMG, Farahani N, Golichenari B, Sahebkar A. Recombinant protein expression in escherichia coli (E.coli): What we need to know. Curr Pharm Des 2018; 24(6): 718-25.
http://dx.doi.org/10.2174/1381612824666180131121940   PMID: 29384059

[9]   Han BW, Lee JY, Yang JK, Lee BI, Suh SW. Crystallization and preliminary X-ray crystallographic analysis of deoxyuridine triphosphate nucleotidohydrolase from *Saccharomyces cerevisiae*. Acta Crystallogr D Biol Crystallogr 2001; 57(8): 1147-9.
http://dx.doi.org/10.1107/S0907444901007909 PMID: 11468401

[10]   Owji H, Nezafat N, Negahdaripour M, Hajiebrahimi A, Ghasemi Y. A comprehensive review of signal peptides: Structure, roles, and applications. Eur J Cell Biol 2018; 97(6): 422-41.
http://dx.doi.org/10.1016/j.ejcb.2018.06.003 PMID: 29958716

[11]   Kudlicki W, Odom OW, Kramer G, Hardesty B, Merrill GA, Horowitz PM. The importance of the N-terminal segment for DnaJ-mediated folding of rhodanese while bound to ribosomes as peptidyl-tRNA. J Biol Chem 1995; 270(18): 10650-7.
http://dx.doi.org/10.1074/jbc.270.18.10650 PMID: 7738002

[12]   Kramer G, Kudlicki W, McCarthy D, Tsalkova T, Simmons D, Hardesty B. N-terminal and C-terminal modifications affect folding, release from the ribosomes and stability of *in vitro* synthesized proteins. Int J Biochem Cell Biol 1999; 31(1): 231-41.
http://dx.doi.org/10.1016/S1357-2725(98)00143-5   PMID: 10216956

[13]   Ojima-Kato T, Nagai S, Nakano H. N-terminal SKIK peptide tag markedly improves expression of difficult-to-express proteins in Escherichia coli and Saccharomyces cerevisiae. J Biosci Bioeng 2017; 123(5): 540-6.
http://dx.doi.org/10.1016/j.jbiosc.2016.12.004 PMID: 28052816

[14]   Belenkaya SV, Shcherbakov DN, Chapoval AI, Esina TI, Elchaninov VV. The effect of thioredoxin and prochymosin coexpression on the refolding of recombinant alpaca chymosin. Vavilovskii Zhurnal Genet Selektsii 2023; 27(4): 421-7.

http://dx.doi.org/10.18699/VJGB-23-50 PMID: 37465195

[15]   Fox JD, Kapust RB, Waugh DS. Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins. Protein Sci 2001; 10(3): 622-30.
http://dx.doi.org/10.1110/ps.45201 PMID: 11344330

[16]   Tropea JE, Cherry S, Waugh DS. Expression and purification of soluble His(6)-tagged TEV protease. Methods Mol Biol 2009; 498: 297-307.
http://dx.doi.org/10.1007/978-1-59745-196-3_19 PMID: 18988033

[17]   Xu Y, Dong M, Wang Q, *et al.* Soluble expression of antimicrobial peptide BSN-37 from Escherichia coli by SUMO fusion technology. Protein J 2023; 42(5): 563-74.
http://dx.doi.org/10.1007/s10930-023-10144-2 PMID: 37561256

[18]   Costa S, Almeida A, Castro A, Domingues L. Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: The novel Fh8 system. Front Microbiol 2014; 5: 63.
http://dx.doi.org/10.3389/fmicb.2014.00063 PMID: 24600443

[19]   Lopez PJ, Marchand I, Joyce SA, Dreyfus M. The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing *in vivo*. Mol Microbiol 1999; 33(1): 188-99.
http://dx.doi.org/10.1046/j.1365-2958.1999.01465.x   PMID: 10411735

[20]   Waugh DS. Making the most of affinity tags. Trends Biotechnol 2005; 23(6): 316-20.
http://dx.doi.org/10.1016/j.tibtech.2005.03.012 PMID: 15922084

[21]   Booth WT, Schlachter CR, Pote S, *et al.* Impact of an N-terminal polyhistidine tag on protein thermal stability. ACS Omega 2018; 3(1): 760-8.
http://dx.doi.org/10.1021/acsomega.7b01598 PMID: 29399652

[22]   Brito Querido J, Díaz-López I, Ramakrishnan V. The molecular basis of translation initiation and its regulation in eukaryotes. Nat Rev Mol Cell Biol 2024; 25(3): 168-86.
http://dx.doi.org/10.1038/s41580-023-00624-9 PMID: 38052923

[23]   Kozak M. Initiation of translation in prokaryotes and eukaryotes. Gene 1999; 234(2): 187-208.
http://dx.doi.org/10.1016/S0378-1119(99)00210-3   PMID: 10395892

[24]   Merrick WC, Pavitt GD. Protein synthesis initiation in eukaryotic cells. Cold Spring Harb Perspect Biol 2018; 10(12): a033092.
http://dx.doi.org/10.1101/cshperspect.a033092 PMID: 29735639

[25]   Roiuk M, Neff M, Teleman AA. eIF4E-independent translation is largely eIF3d-dependent. Nat Commun 2024; 15(1): 6692.
http://dx.doi.org/10.1038/s41467-024-51027-z PMID: 39107322

[26]   Ruffel S, Dussault MH, Palloix A, *et al.* A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). Plant J 2002; 32(6): 1067-75.
http://dx.doi.org/10.1046/j.1365-313X.2002.01499.x   PMID: 12492847

[27]   Gallois JL, Charron C, Sánchez F, *et al.* Single amino acid changes in the turnip mosaic virus viral genome-linked protein (VPg) confer virulence towards Arabidopsis thaliana mutants knocked out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G. J Gen Virol 2010; 91(1): 288-93.
http://dx.doi.org/10.1099/vir.0.015321-0 PMID: 19741065

[28]   Michel V, Julio E, Candresse T, *et al.* A complex *eIF4E locus* impacts the durability of *va* resistance to *Potato virus Y* in tobacco. Mol Plant Pathol 2019; 20(8): 1051-66.
http://dx.doi.org/10.1111/mpp.12810 PMID: 31115167

[29]   Lebedeva MV, Nikonova EY, Terentiev AA, Taranov VV, Babakov AV, Nikonov OS. VPg of potato Virus Y and potato cap-binding eIF4E factors: Selective interaction and its supposed mechanism. Biochemistry 2021; 86(9): 1128-38.
http://dx.doi.org/10.1134/S000629792109008X PMID: 34565316

[30]   Lebedeva M, Nikonova E, Babakov A, *et al.* Interaction of Solanum tuberosum L. translation initiation factors eIF4E with potato virus Y VPg: Apprehend and avoid. Biochimie 2024; 219: 1-11.
http://dx.doi.org/10.1016/j.biochi.2023.08.005 PMID: 37562705

[31]   Abramson J, Adler J, Dunger J, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 2024; 630(8016): 493-500.
http://dx.doi.org/10.1038/s41586-024-07487-w PMID: 38718835

[32]   Pronk S, Páll S, Schulz R, *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 2013; 29(7): 845-54.
http://dx.doi.org/10.1093/bioinformatics/btt055 PMID: 23407358

[33]   Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. J Comput Chem 2005; 26(16): 1701-18.
http://dx.doi.org/10.1002/jcc.20291 PMID: 16211538

[34]   Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comput Phys Commun 1995; 91(1-3): 43-56.
http://dx.doi.org/10.1016/0010-4655(95)00042-E

[35]   Páll S, Abraham MJ, Kutzner C, Hess B, Lindahl E. Tackling exascale software challenges in molecular dynamics simulations with GROMACS. Solving Software Challenges for Exascale EASC. Springer, Cham, vol 8759. pp. 3-27
http://dx.doi.org/10.1007/978-3-319-15976-8_1

[36]   Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of *Coot*. Acta Crystallogr D Biol Crystallogr 2010; 66(4): 486-501.
http://dx.doi.org/10.1107/S0907444910007493 PMID: 20383002

[37]   Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph 1996; 14(1): 33-38, 27-28.
http://dx.doi.org/10.1016/0263-7855(96)00018-5 PMID: 8744570

[38]   Wimley WC, White SH. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat Struct Mol Biol 1996; 3(10): 842-8.
http://dx.doi.org/10.1038/nsb1096-842 PMID: 8836100

[39]   Sehnal D, Bittrich S, Deshpande M, *et al.* Mol* Viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res 2021; 49(W1): W431-7.
http://dx.doi.org/10.1093/nar/gkab314 PMID: 33956157

[40]   Ashby J A, Stevenson C E M, Jarvis G E, Lawson D M, Maule A J. Structure-based mutational analysis of eIF4E in relation to sbm1 resistance to pea seed-borne mosaic virus in pea. PLoS One 2011; 6(1): e15873.
http://dx.doi.org/10.1371/journal.pone.0015873 PMID: 21283665

[41]   Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cocrystal structure of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell 1997; 89(6): 951-61.
http://dx.doi.org/10.1016/S0092-8674(00)80280-9 PMID: 9200613

[42]   Miras M, Truniger V, Silva C, Verdaguer N, Aranda MA, Querol-Audí J. Structure of eIF4E in complex with an eIF4G peptide supports a universal bipartite binding mode for protein translation. Plant Physiol 2017; 174(3): 1476-91.
http://dx.doi.org/10.1104/pp.17.00193 PMID: 28522457

[43]   Monzingo AF, Dhaliwal S, Dutt-Chaudhuri A, *et al.* The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. Plant Physiol 2007; 143(4): 1504-18.
http://dx.doi.org/10.1104/pp.106.093146 PMID: 17322339

[44]   Peter D, Weber R, Köne C, *et al.* Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation. Genes Dev 2015; 29(17): 1835-49.
http://dx.doi.org/10.1101/gad.269068.115 PMID: 26294658

[45]   Niedzwiecka A, Marcotrigiano J, Stepinski J, *et al.* Biophysical studies of eIF4E cap-binding protein: Recognition of mRNA 5′ cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. J Mol Biol 2002; 319(3): 615-35.
http://dx.doi.org/10.1016/S0022-2836(02)00328-5   PMID: 12054859

[46]   Savitsky P, Bray J, Cooper CDO, *et al.* High-throughput production of human proteins for crystallization: The SGC experience. J Struct Biol 2010; 172(1): 3-13.
http://dx.doi.org/10.1016/j.jsb.2010.06.008 PMID: 20541610

[47]   Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M.

SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact 2012; 11(1): 753.
http://dx.doi.org/10.1186/1475-2859-11-56 PMID: 22569138

[48] López-Cano A, Martínez-Miguel M, Guasch J, Ratera I, Arís A, Garcia-Fruitós E. Exploring the impact of the recombinant Escherichia coli strain on defensins antimicrobial activity: BL21 *versus* Origami strain. Microb Cell Fact 2022; 21(1): 77.
http://dx.doi.org/10.1186/s12934-022-01803-7 PMID: 35527241