Conformational mobility of human translation elongation factor A1

Dmitry S. Kanibolotsky[1], Oleksandra V. Novosyl'na, Boris S. Negrutskii, Anna V. El'skaya

The Institute of Molecular Biology and Genetics
150 Zabolotny Str., Kyiv 03143, Ukraine

[1]Kyiv National Taras Shevchenko University
64, Volodymyrska street, 01033 Kyiv, Ukraine.

negrutskii@imbg.org.ua

A model of the eEF1A1 isoform of human translation elongation factor 1A has been proposed using a homology modelling method. The conformational mobility of eEF1A1 has been studied by means of multiple molecular dynamics simulation. The most essential correlated motions in the protein have been identified using the covariance analysis of atom trajectories. It has been determined that reciprocal flexibility of domains I and II can lead to disappearance of the gap between the domains and to formation of a "closed" conformation of the protein. The amino acid residues, which are characterised by maximal flexibility of C -atoms, have been described.

Keywords: protein synthesis, translation elongation, protein molecular dynamics

Introduction. The main translational role of eEF1A is to deliver aminocyl-tRNA to 80S ribosome and to participate in selection of correct aminoacyl-tRNA in ribosomal A site [1]. Supposedly, eEF1A is involved in channelling of deacetylated tRNA between aminocyl-tRNA synthetases and ribosomes [2]. Besides, eEF1A may take part in a number of processes, which seem not to be connected with protein biosynthesis in the cell [3, 4].

There are at least two mammalian eEF1A isoforms, which are 97% homologous and 93% identical [5]. It seems like the functions of these isoforms do not differ much, however, the data testify to the differences between them, which may be revealed in their different participation in cell signalling ways, particularly, in the course of carcinogenesis [6–8]. Therefore, it is important to investigate the spatial structures of highly homologous human eEF1A isoforms in order to understand the basic structural cause of these differences. Current work presents the study on human eEF1A1 isoform.

eEF1A1 molecule of mammals consists of 462 amino acid residues and has got seven modified residues as follows: N-trimethyllysines (M31) in positions 36, 79, and 318; N-dimethyllysines (Mly) in positions 55 and 165; L-glutamyl-5-glycerolphosphorylethanolamines (GPE) in positions 301 and 374 [9]. The spatial struc-

```
eEF1A1          1         MGKEKT HINIVVIGHV DSGKSTTTGH LIYKCGGIDK RTIEKFEKEA
Yeast           1         MGKEKS HINVVVIGHV DSGKSTTTGH LIYKCGGIDK RTIEKFEKEA
S.sulfotaricus  1         MS-QKP HLNLIVIGHV DHGKSTLVGR LLMDRGFIDE KTV-KEAEEA

eEF1A1          47   AE-MGKGSFK YAWVLDKLKA ERERGITIDI SLWKFETSKY YVTIIDAPGH
Yeast           47   AE-LGKGSFK YAWVLDKLKA ERERGITIDI ALWKFETPKY QVTVIDAPGH
S.sulfotaricus  45   AKKLGKESEK FAFLLDRLKE ERERGVTINL TFMRFETKKY FFTIIDAPGH

eEF1A1          96   RDFIKNMITG TSQADCAVLI VAAGVGEFEA GISKNGQTRE HALLAYTLGV
Yeast           96   RDFIKNMITG TSQADCAILI IAGGVGEFEA GISKDGQTRE HALLAFTLGV
S.sulfotaricus  95   RDFVKNMITG ASQADAAILV VSAKKGEYEA GMSVEGQTRE HIILAKTMGL

eEF1A1          146  KQLIVGVNKM DSTEPPYSQK RYEEIVKEVS TYIKKIGYNP DTVAFVPISG
Yeast           146  RQLIVAVNKM DSVK—WDES RFQEIVKETS NFIKKVGYNP KTVPFVPISG
S.sulfotaricus  145  DQLIVAVNKM DLTEPPYDEK RYKEIVDQVS KFMRSYGFNT NKVRFVPVVA

eEF1A1          196  WNGDNMLEPS ANMPWFKGWK VTRKDGNASG TTLLEALDCI LP-PTRPTDK
Yeast           194  WNGDNMIEAT TNAPWYKGWE KETKAGVVKG KTLLEAIDAI EQ-PSRPTDK
S.sulfotaricus  195  PSGDNITHKS ENMKWYNGPT LEE------- -----YLDQL EL-PPKPVDK

eEF1A1          245  PLRLPLQDVY KIGGIGTVPV GRVETGVLKP GMVVTF---A PVNVTTEVKS
Yeast           243  PLRLPLQDVY KIGGIGTVPV GRVETGVIKP GMVVTF---A PAGVTTEVKS
S.sulfotaricus  232  PLRIPIQDVY SISGVGTVPV GRVESGVLKV GDKIVF---M PAGKVGEVRS

eEF1A1          292  VEMHHEALSE ALPGDNVGFN VKNVSVKDVR RGNVAGDSKN DPPMEAAGFT
Yeast           290  VEMHHEQLEQ GVPGDNVGFN VKNVSVKEIR RGNVCGDAKN DPPKGCNASF
S.sulfotaricus  279  IETHHTKMDK AEPGDNIGFN VRGVEKKDIK RGDVVG-HPN NPPTVADEFT

eEF1A1          342  AQVIIL---- --NHPGQISA GYAPVLDCHT AHIACKFAEL KEKIDRRSGK
Yeast           340  ATVIVL---- --NHPGQISA GYSPVLDCHT AHIACRFDEL LEKNDRRSGK
S.sulfotaricus  328  ARIIVV---- --WHPTALAN GYTPVLHVHT ASVACRVSEL VSKLDPRTGQ

eEF1A1          386  KLEDGPKFLK SGDAAIVDMV PGKPMCVESF SDYPPLGRFA VRDMRQTVAV
Yeast           384  KLEDHPKFLK SGDAALVKFV PSKPMCVEAF SEYPPLGRFA VRDMRQTVAV
S.sulfotaricus  372  EAEKNPQFLK QGDVAIVKFK PIKPLCVEKY NEFPPLGRFA MRDMGKTVGV

eEF1A1          436  GVIKAVDKKA AGAGKVTKSA QKAQKAK
Yeast           434  GVIKSVDK-T EKAAKVTKAA QKAAK-K
S.sulfotaricus  422  GIIVDVKP-- ---AKVEIK
```

Fig.1 Alignment of human eEF1A [13], eEF1A of yeasts [14], and aEF1A of archaebacteria *Sulfolobus solfataricus* [15]

ture of eEF1A of mammalians remains unknown as it was impossible to obtain the crystals of this protein. At the same time the knowledge of 3D structure of eEF1A as well as its dynamics in the solution is necessary for both thorough understanding of functioning of ribosomal phase of protein biosynthesis in mammals and solving the question on the mechanisms of eEF1A interactions with numerous partner-proteins, which are not involved in translation directly.

Therefore, the aim of current investigation was to design the model of spatial structure of human eEF1A1 and to analyse the molecular dynamics of this protein in the solution.

eEF1A of the following yeasts *Saccharomyces cerevisiae* [10, 11] and archaebacteria *Sulfolobus solfataricus* [12] were analysed using the method of X-ray diffraction up to now. As amino acid sequence of human eEF1A1 is 80.7% and 53.1% homologous to

that of yeasts and archaebacterial factors 1A respectively (Fig.1), therefore, the data were used as templates to simulate 3D structure of the eEF1A1.

The conformational mobility of eEF1A1 was studied using the method of molecular dynamics simulation in the 10 ns time scale. Six trajectories with different initial velocities of atoms were obtained and the most movable regions of protein structure were determined. The most essential correlated motions in the protein were identified using the covariance analysis of atom trajectories. It was determined that reciprocal flexibility of domain II and the "top" part of domain I can lead to disappearance of the gap between the domains I and II and to the formation of new ("closed") conformation of eEF1A. The transition between "open" and "closed" protein conformations was shown to be reversible. This sort of changes in eEF1A conformation may be involved in the compactisation of eEF1A molecule at the interaction with tRNA earlier described [16].

Materials and Methods. 3D model of human eEF1A1 was designed using Swiss-Model server (http://swissmodel.expasy.org//SWISS-MODEL.html) [17] using crystallographic structures of eEF1A of *S. cerevisiae* (PDB codes – 1ije, 1ijf, 1g7c [11], and 1f60 [10]) and of *S. solfataricus* (1jny [12]) as templates (Fig.1). The model obtained consists of 442 amino acid residues – from Gly2 to Lys443. N-terminal amino acid residues of yeast and archaebacterial proteins were not determined using the X-ray diffraction, thus, homologous residues of human eEF1A1 were not included into the final model.

Lysine residues 36, 79, and 318 were replaced with trimethyllysines, whereas Lys55 and Lys165 were replaced with dimethyllysines, substituting corresponding hydrogenous atoms for methyl groups. NH_2-groups of Gln301 and Gln374 were replaced with glycerylphosphorylethanolamine ones [9].

The major part of simulation of molecular dynamics was performed at the computing clusters of National Taras Shevchenko University of Kyiv (www.cluster.kiev.ua/eng/) and of Glushkov Institute of Cybernetics of NAS of Ukraine (cluster.icyb.kiev.ua). Molecular dynamics simulation was performed using Gromacs 3.1.4 software [18]. Force field Gromos96 [19] was modified, adding parameters for N-methyllysine (Mls), N-dimethyllysine, N-trimethyllysine, and

L-glutamyl-5-glycerylphosphorylethanolamine residues. The topological data on non-standard amino acid residues were obtained from PRODRG2 server (http://davapc1.bioch.dun-dee.ac.uk/programs/prodrg/prodrg.html) [20]. The hydrogen atoms were added to non-carbon heavy atoms using pdb2gmx program of Gromacs 3.1.4 package. Arginine, lysine, dimethyllysine residues and N-termianal amino group were protonated and charged positively. Carboxyl groups of aspartic and glutamine acids residues and C-terminal residue were remained deprotonated with –1 charge. His7 residue was protonated by N atom, because in case of N protonation, the distance between hydrogen atom and CH -group of Tyr86 would have been less than the sum of Van der Waals radii. Histidine residues 15 and 95 were protonated by N, as N atoms of these residues form the hydrogen bonds with peptide hydrogen atoms of Asp17 and Asp97 respectively. His26, His296, and His367 were protonated by N atoms, which allows the protons to interact electrostatically with carbonyl oxygen atoms of the same residues. His136 was protonated by N, as N atom forms the hydrogen bond with amide hydrogen of Gln132 side chain. His295 was protonated by N position to allow proton to interact with N of His296. His 349 was protonated in position N rather than in N because the distance from CH_2 of Met429 would have been less than the sum of Van der Waals radii. His364 is defined as N -protonated, as in this case proton may interact with sulphur atoms of Cys363 residue.

The protein models were inserted into the virtual box of a truncated octahedron shape, 1256.34 nm³. The minimal distance between the protein and the box wall was 1.5 nm to prevent artificial periodicity and to allow the proteins to change conformations freely [21, 22]. 39557 SPC (Single Point Charge) models of water molecules were added to the box [23]. To neutralise the charge of the system and to imitate the ionic force of 0.1 mole/l, 69 and 74 molecules of water were replaced with sodium and chlorine ions. The positions of the ions were chosen by the Poisson-Boltzmann distribution. The initial atoms velocities were generated according to the Maxwell's distribution. Atoms coordinates data were updated every $2 \cdot 10^{-15}$ s. The protein bonds are constrained by the linear constraint solver (LINCS) algorithm [24]. The cut-off for electrostatic

interaction is 0.9 nm, whereas a double cut-off was set for Van der Waals interactions: the interactions between atoms within 0.9 nm were updated at each step, and the interactions within the distance between 0.9 and 1.4 nm were updated at each 10th step. The Particle-mesh Ewald (PME) algorithm [25] was applied to describe the long-range electrostatic interactions.

The levels of temperature and pressure were kept at 298 K and 1 atmosphere using the Berendsen's method [26] with relaxation times of 0.1 ps and 0.5 ps, respectively. After equilibration of solvent molecules, additional energy minimization of the protein model was carried out. Then, the main MD simulation was performed with the same parameters as the restrained simulation, except for affixment of protein coordinates to the initial values and the pressure control. The protein coordinates were recorded into the output trajectory file every 1 ps. The time of the system equilibration was defined as the time when the root-mean-square deviation (rmsd) of the C -atoms trace reached a plateau. The time of equilibration was not included during the covariance analysis and during the calculations of root-mean-square fluctuations of C -atoms.

Six MD simulations were performed. Six trajectories, different in the initial velocities of atoms are as follows: 1 (10000 ps), 2 (10466 ps), 3 (7110 ps), 4 (10514 ps), 5 (9920 ps), and 6 (10297 ps).

The trajectory analysis was performed using following parameters:

1.) C -atoms trace rmsd after fitting to the initial conformations before the main dynamics (using g_rms program of Gromacs 3.1.4 package);

2.) The distances between the centres of the domains (g_dist program) (coordinates of centres were determined as the mean value of coordinates of all C -atoms of the corresponding domain);

3.) Minimal distance and the number of contacts between residues Arg69-Leu77 and His295-Gly305, located on the opposite sides to the gap between domains I and II (g_mindist program);

4.) The root-mean-square fluctuations (rmsf) of C -atoms with respect to their average positions after fitting to the initial conformation (g_rmsf program) (as the domains moved freely relative to each other, root-mean-square fluctuations of C -atoms were calculated for the each domain separately after fitting of the

domain conformations to the initial ones; the trajectory ranges of 6000 ps after equilibration were selected for the calculations);

5.) The covariance analysis of the correlated motions of C -atoms (g_covar, g_anaeig and g_analyse programs).

The trajectory 2 was selected as the most stable one using the graphs of root-mean square fluctuations of protein conformation (Fig.3, a), then the matrix M of root-mean-square fluctuations for the selected trajectory was built. Each matrix element $M_{i,j}$ is the rmsd between the protein conformations at time moments i and j. Thus, the matrix size is n n, where n is the number of the trajectory frames (equal to the number of picoseconds in the trajectory). The matrix was visualized in such a way that every element would be presented as a dot, colour range from white to black, depending on the corresponding value. The matrix region of visualised, presented as light square and specific for low root-mean-square fluctuations (2500–10466 ps) was determined.

The covariance analysis was performed for this trajectory range and the covariance matrix C was constructed:

$$C_{k,l} \left\langle \overrightarrow{x_k(t) \langle x_k(t)\rangle} \right\rangle \left\langle \overrightarrow{x_l(t) \langle x_l(t)\rangle} \right\rangle, (1)$$

where the arrows above the expressions denote vector values, the angle brackets – average values, the sign is scalar product here, x is coordinate, t - time, k and l are one of the space dimensions (x, y or z) for one of the atoms. So the covariance matrix has $3m$ $3m$ dimensions, where m is the total number of C -atoms in the protein model, $i.e.$ 442. If the two atoms move along certain Cartesian axes absolutely asynchronously, the matrix element is equal to zero. If they move absolutely synchronously, the matrix element corresponds to the atoms rmsf. As the atoms move synchronously with themselves, the covariance matrix diagonal contains the corresponding atoms rmsf values along certain Cartesian axes.

The matrix C has been diagonalized using orthonormal matrix R:

$$C R diag(_1, _2,..._{3m}) R^T (2)$$

where the columns of matrix R are eigenvectors, which correspond to eigenvalues , $_1$ $_2$ … $_{3m}$. The first few eigenvectors (characterized by the largest eigenvalues) often reflect collective global motions of the atoms. The quality of the covariance analysis must be controlled to exclude interpretation of the random diffusion of protein atoms as correlated motion. In the case of the random diffusion, the projection of the trajectory on eigenvector k is the cosine with a period of t $k/2$, where t is the length of the analyzed part of the trajectory (ps). That is why the cosine content of the trajectory projections is calculated using the g_analyse program. The cosine content is equal to 1 if the trajectory projection is completely cosinusoid with respective period and the atom motions are completely random. If the atom motions along respective eigenvector are completely correlated, the cosine content is zero. So the cosine content can be considered as a fraction of the atoms motion randomness.

The programs VMD [28] and Swiss-PDB Viewer [29] were used for the trajectory visualization and graphical analysis of resulted conformations.

Results and Discussion. The model consists of three domains. The first domain contains 8 -strands (Thr6-Ile13, Trp78-Thr82, Tyr85-Ala92, Cys111-Ala118, Gln147-Asn153, Ala189-Ile193, Trp214-Arg218, and Gly221-Gly225), which form the -sheet surrounded by 8 -helixes (Lys20-Lys30, M3l/Lys36-Glu48, Ala57-Glu68, Asp97-Thr104, Val120-Ala125, Thr133-Leu143, Gln164-Ile181, and Leu228-Asp233).

The second domain is the -barrel formed by the -strands Leu248-Leu250, Asp252-Ile256, Gly260-Val267, Met276-Ala281, Val285-Val289, Ser291-Met294, Glu297-Leu299, Asp306-Val312 and Asn324-Ser329.

The third domain also is the -barrel consisting of -strands Gly339-Leu347, Tyr357-Cys363, Ala366-Asp380, Lys385-Gly390, Asp398-Gly407, Gly422-Asp428, and Gln431-Lys443. Domains I and III, as well as II and III are located close to each other, whereas there is a significant distance between domains I and II, limited by Arg69-Leu77 loop from the side of domain I and by His295-Gly305 from the side of domain II (Fig.2, *a*).

In the present study the residues from Gly2 to Pro238 are attributed to domain I, because Pro238 is the last residue from the continuous series of residues in the unstructured chain Cys234-Arg247, which are situated within the distance of the Van der Waals radii sum to residues from the other chain of the domain I (Asp110 for Pro238). Similarly, Pro241 is the first residue from the chain Cys234-Arg247, which is within the distance of the Van der Waals radii sum to residues of other chain of the domain II (Gly270), so Pro241 is chosen as the first residue of the domain II. Ser329 is the end of the domain II because it is the last residue in the Asn324-Ser329 -strand. Domain III starts with Met335, which is situated within the distance of the Van der Waals radii sum to residue of other chain of the domain III (Cys411). Such selection of domain borders prevents artificial mobility at the terminal ends of domains during fitting of any other domain conformation to the initial conformation.

Complete conformational sampling for a three-domain protein may require an MD trajectory of a relatively long time scale. However, the protein can adopt denatured forms during simulation, and if the only trajectory is available, the denatured state can be interpreted as native conformation in solution. Moreover, statistical errors accumulate at long MD calculation. In view of the aforesaid, an alternative, multiple MD simulation method [30, 31] has been used. The method consists in the simulation of several relatively short MD trajectories starting from the same initial protein conformation with different initial atom velocities. Multiple MD simulation permits a reduction of computational time, a widening of the statistical basis, and allows us to evaluate quality of single trajectories and minimize force-field induced artefacts [32, 33].

The time dependence of rmsd of current eEF1A conformations after fitting on the initial one is presented in Fig.3, *a*. Essential variability in data shows the presence of a significant conformational space (the set of conformations) of eEF1A in aqueous solution. Trajectory 2 is characterized by the protein conformations with minimal rmsd. The range of 2500-10466 ps of trajectory 2 was determined to be the most stable one; therefore, the covariance analysis was performed for this range.
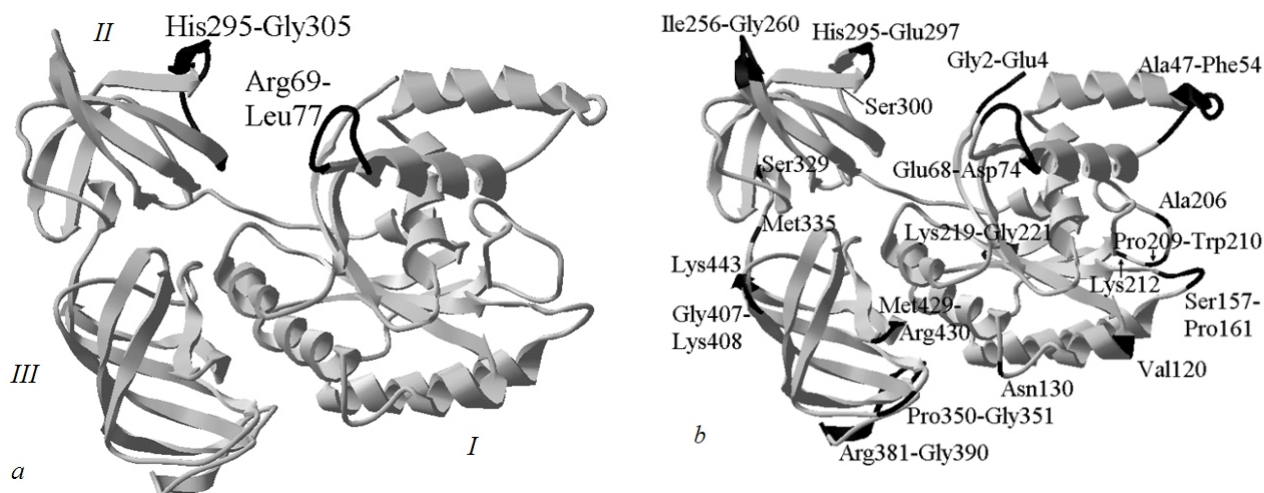
Fig.2 Three-dimensional model of human eEF1A: *a* – domains *I*, *II*, *III* (Arg69-Leu77 and His295-Gly305, located on the opposite sides of the gap between domains *I and II* are marked black); b – amino acid residues rmsf, exceeding 0.14 nm in domain *I; 0.11 nm in domain II*, and 0.09 nm in domain *III* (marked black)

The eigenvalues and cosine contents for projections of the trajectory onto first six eigenvectors are presented in Table 1. As it is seen cosine contents are acquired in the range of 0.00111–0.0337, which indicates the high quality of the covariance analysis performed. The first principal component of trajectory 2 (Fig. 4, *a*) is characterized by an increased proximity of the "top" part of the domain I (tail Gly2-Lys5, -helixes M3L36-Glu48, Ala57-Glu68 and loops Cys31-Asp32, Met49-Tyr56, Arg69-Leu77) and domain II. The projection of trajectory 2 onto the second eigenvector is described by rotation of the "top" part of domain I around horizontal axis and of domain II and the "remaining" part of domain I around vertical axes (Fig.4, *b*). In addition, the "up" and "down" motions of residues Gln352, Leu375, Lys376, Leu387, Glu388, Gly390, Met429, Arg430 occur on domain III. Leu375-Lys376 and Leu387-Gly390 compose -hairpin Leu375-Gly390, residues Met429-Arg430 are located at the tip of в-hairpin Gly422-Val437, whereas Gln352 is located in the disordered loop Asn348-Gly356.

The third principal component of the trajectory 2 (Fig.4, *c*) is consistent with the rotation of domain I, on one side, and of domains II and III, on the other side, around the same horizontal axis, but in the opposite directions.

The distance between centres of the domains was calculated (Fig.3, *b*) to analyse directly the inter-domain mobility of the isoforms. A direct link between the full protein rmsd and the distance between the domains I and II is observed for eEF1A1 (Fig. 3, *a* and *b*). The trajectories characterized by the maximal rmsd for the full protein (6 and 1) demonstrate the smallest distance between domains I and II. Similarly, the trajectories with the minimal rmsd (2 and 4) show the maximal distance between the domains I and II. Consequently, the departure of eEF1A1 from the initial conformation is accompanied by the approaching of the domains I and II, and trajectory average conformations differ in smaller distances between domains than the initial conformation. It has to be noted also that the distances between domains I and II and between domains II and III do not change essentially.

Mutual approach of domains I and II results in closing of the gap between the domains, i.e. between Arg69-Leu77 and His295-Gly305, located at the borders of the gap from the side of domains I and II respectively (Fig.2, *a*). Minimal distance and the number of interatomic contacts between the mentioned residues are presented in Fig.3, *c* and *d*. For trajectories 1 and 6, the value of minimal distance is decreased to the sum of Van der Waals radii at 6100 and 1300 ps, respectively (Fig.3, *c*), whereas the number of interatomic contacts
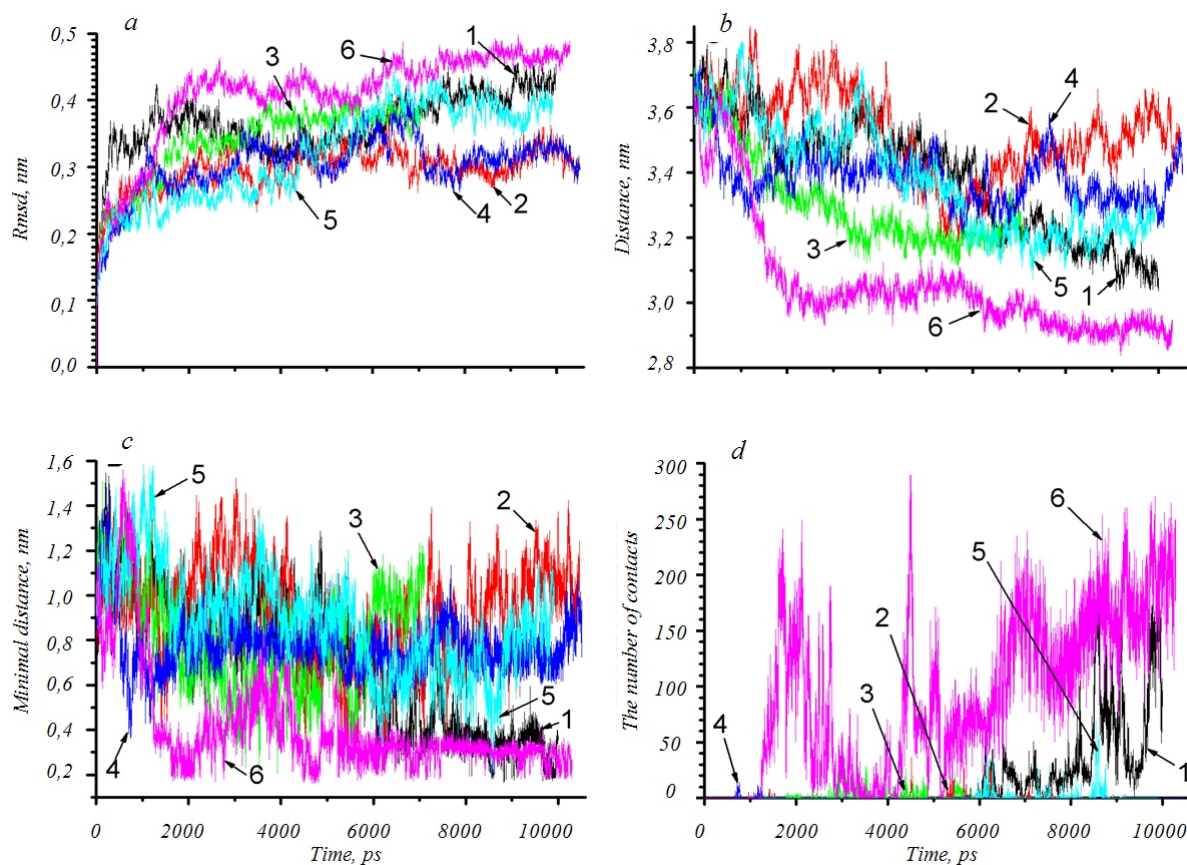
Fig.3 Results of molecular dynamics analysis of human eEF1A1: *a* – rmsd of protein conformations from the initial C -atoms conformation; *b* – distance between domains *I* and *II*; *c* – minimal distance between amino acid sequences Arg69-Leu77 and His295-Gly305; *d* – the num-

First eight eigenvalues and cosine contents for the 2500-10466 ps part of trajectory 2

Number of eigenvector	Eigenvalue	Cosine content
1	2.92577	0.0337375
2	1.68434	0.0250312
3	0.688702	0.00184525
4	0.468017	0.001109
5	0.326925	0.0284407
6	0.282979	0.0273437

increases significantly – from 0 to 20–280 (Fig. 3, *d*). Therefore, completely closed protein conformation, with no gap between domains I and II, is formed. The proteins can oscillate from an extended to a compact conformation and vice versa. For example, the trajec-

tory 6 moves to an "open" conformation at 2900 ps, then returns to a "closed" one at 4300 ps. The protein on trajectory 5 adopts the "closed" conformation for a short period of time only (8500–8600 psec). Earlier, using the method of small-angle neutron scattering we have demonstrated that eEF1A1 molecule becomes more compact in the presence of tRNA [16]. Supposedly, "closed" conformation of eEF1A1 is stabilized by the interactions with molecule of tRNA.

To characterize the random, non-correlated motions of C -atoms, the rmsfs of these atoms were calculated for the separate domains in the 6000 ps trajectory ranges: 4000-10000 ps for the trajectories 1, 2, 4, and 6; 1110-7110 ps for the trajectory 3, and 3920-9920 ps for the trajectory 5. The data were averaged for each trajectory (Fig. 5). Amino acid residues with rmsfs higher than 0.14 in domain I, 0.11 in domain II, and 0.09 in domain III are presented in Fig.2, *b* in black colour. Different values of rmsf thresholds are connected with dif-
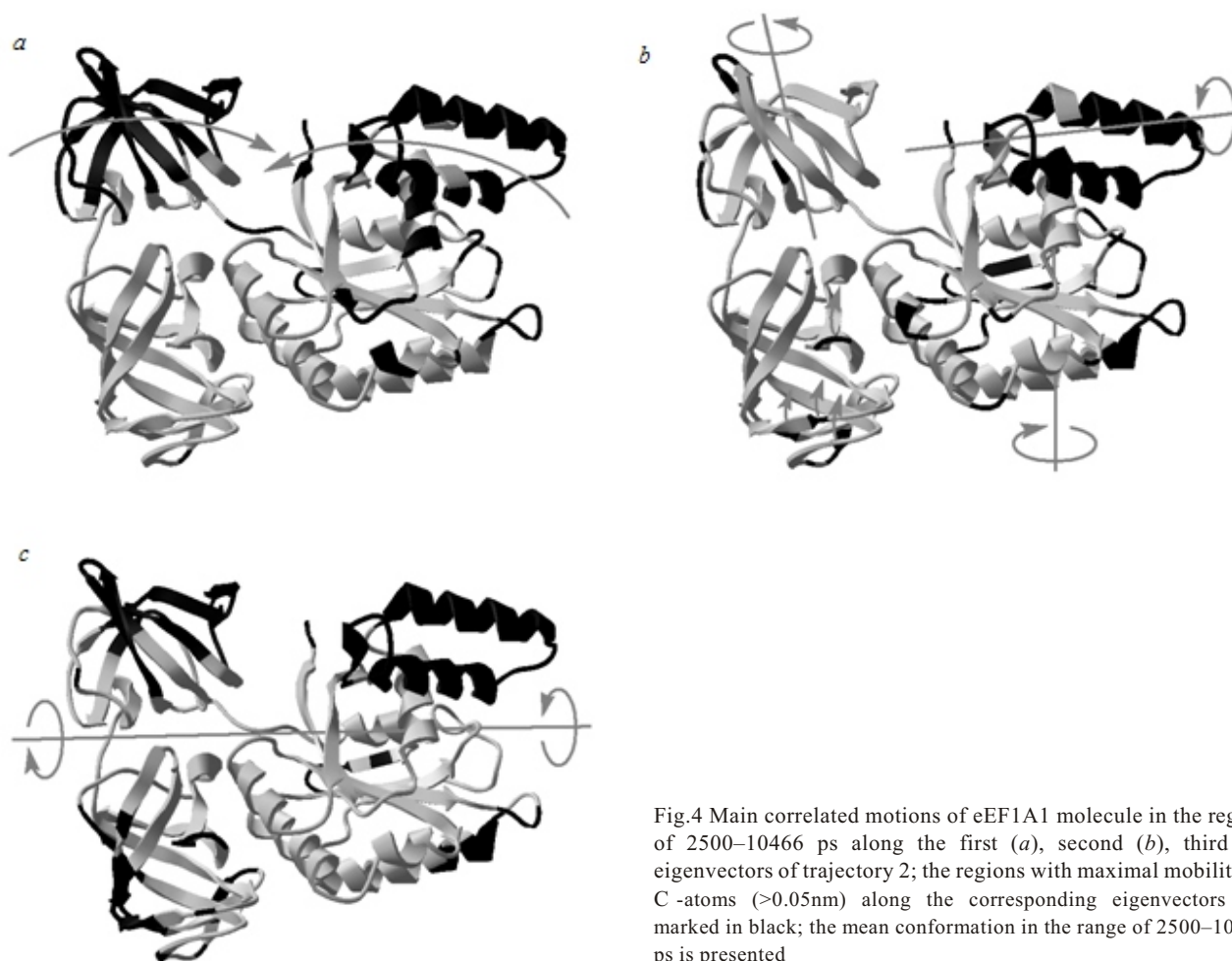
313

Fig.4 Main correlated motions of eEF1A1 molecule in the region of 2500–10466 ps along the first (*a*), second (*b*), third (*c*) eigenvectors of trajectory 2; the regions with maximal mobility of C -atoms (>0.05nm) along the corresponding eigenvectors are marked in black; the mean conformation in the range of 2500–10466 ps is presented

ferent sizes of domains – the larger domain, the higher are root-mean-square fluctuations of C -atoms. As it is seen from Figs. 5, *a* and 2, *b*, domain I is specific for the highest mobility of amino acid residues in the disordered regions (Gly2-Glu4, Met49-Phe54, Arg69-Asp74, Asn130, Ser157-Pro161, Ala206, Pro209, Trp210, Lys212), at the terminal ends of -helixes, close to disordered regions (Ala47-Glu48, Glu68, and Val120), and at the tip of -hairpin Trp214-Gly225 (Lys219-Gly221). In domain II the most movable are Ile256-Gly260 and His295-Glu297, which form the tips of -hairpins Asp252-Val264 and Val292-Leu299, respectively, and also Ser329 at the C-terminus of the domain (Fig.5, *b* and Fig.2, *b*). In domain III the most movable are Arg381-Gly390 and Met429-Arg430, located in -hairpins Leu375-Gly390, Gly422-Val437, as well as Met335, Pro350, Gly351, Lys408 in the dis-

ordered regions and Gly407, Lys443 at the termini of -strands (Fig.5, *c* and Fig.2, *b*).

Thus, we described in detail the conformational mobility of the important component of human translation system – eEF1A using the method of multiple molecular dynamics simulation. The main mobility of the molecule was shown to reveal at the level of opening-closing of the gap between structural domains I and II. The possibility for compactisation of eEF1A was confirmed by experimental data, obtained earlier. The most movable regions of eEF1A molecule, located in all three structural domains of protein were described.
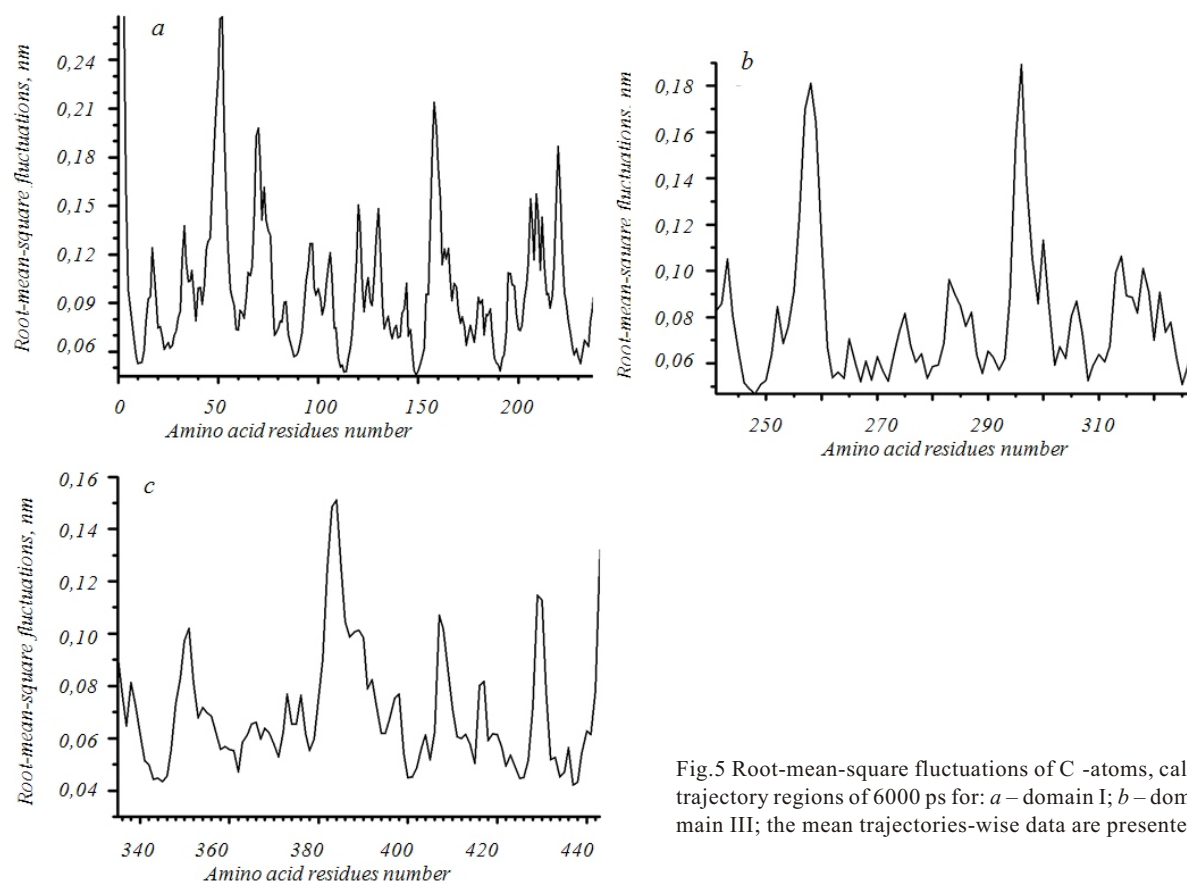
Acknowledgements

Fig.5 Root-mean-square fluctuations of C -atoms, calculated in the trajectory regions of 6000 ps for: *a* – domain I; *b* – domain II; *c* – domain III; the mean trajectories-wise data are presented

Д. С. Каниболоцкий, А. В. Новосильная, Б. С. Негруцкий, А. В. Ельская

Конформационная подвижность фактора элонгации трансляции eEF1A1 человека

Резюме

Методом моделирования по гомологии построена модель изоформы eEF1A1 фактора элонгации трансляции человека. Конформационная подвижность eEF1A1 изучена методом многократного моделирования молекулярной динамики. С использованием ковариационного анализа выделены наиболее существенные согласованные движения атомов. Определено, что взаимные движения первого и второго доменов могут приводить к закрытию щели между доменами и образованию новой «закрытой» конформации. Описаны аминокислотные остатки, обладающие наибольшей подвижностью С -атомов.

Ключевые слова: биосинтез белка, элонгация трансляции, молекулярная динамика белков.

REFERENCES

1. *Negrutskii B. S., El'skaya A. V.* Eukaryotic translation elongation factor 1 alpha: structure, expression, functions, and possible role in aminoacyl-tRNA channeling // Progr. Nucl. Acids Res. Mol. Biol.–1998.–**60**.–P. 47–78.

2. *Petrushenko Z. M., Budkevich T. V., Shalak V. F., Negrutskii B. S., El'skaya A. V.* Novel complexes of mammalian translation elongation factor eEF1A·GDP with uncharged tRNA and aminoacyl-tRNA synthetase. Implications for tRNA channeling // Eur. J. Biochem.–2002.–**269**.–P. 4811–4818.

3. *Thornton S., Anand N., Purcell D., Lee J.* Not just for housekeeping: protein initiation and elongation factors in cell growth and tumorigenesis // J. Mol. Med.–2003.–**18**.–P. 536–548.

4. *Ejiri S.* Moonlighting functions of polypeptide elongation factor 1: from actin bundling to zinc finger protein R1-associated nuclear localization // Biosci. Biotechnol. Biochem.–2002.–**66**.–P. 1–21.

5. *Kahns S., Lund A., Kristensen P., Knudsen C. R., Clark B. F., Cavallius J., Merrick W. C.* The elongation factor 1 A-2 isoform from rabbit: cloning of the cDNA and characterization of the protein // Nucl. Acids Res.–1998.–**26**.–P. 1884–1890.

6. *Anand N., Murthy S., Amann G., Wernick M., Porter L. A., Cukier I. H., Collins C., Gray J. W., Diebold J., Demetrick D. J., Lee J. M.* Protein elongation factor EEF1A2 is a putative oncogene in ovarian cancer // Nat. Genet.–2002.–**31**.–P. 301–305.

7. *Tomlinson V. A., Newbery H. J., Wray N. R., Jackson J., Larionov A., Miller W. R., Dixon J. M., Abbott C. M.* Translation elongation factor eEF1A2 is a potential oncoprotein that is overexpressed in two-thirds of breast tumours // BMC Cancer.–2005.–**5**.–P. 113.

8. *Kulkarni G., Turbin D. A., Amiri A., Jeganathan S., Andrade-Navarro M. A., Wu T. D, Huntsman D. G., Lee J. M.* Expression of protein elongation factor eEF1A2 predicts favorable outcome in breast cancer // Breast Cancer Res. Treat.–2007.–**102**.–P. 31–41.

9. *Dever T. E., Costello C. E., Owens C. L., Rosenberry T. L., Merrick W. C.* Location of seven post-translational modifications in rabbit elongation factor 1 alpha including dimethyllysine, trimethyllysine, and glycerylphosphorylethanolamine // J. Biol. Chem.–1989.–**264**.–P. 20518–20525.

10. *Andersen G. R., Pedersen L. Valente L., Chatterjee I., Kinzy T. G., Kjeldgaard M., Nyborg J.* Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex Eef1A:Eef1Ba // Mol. Cell.–2000.–**6**.–P. 1261.

11. *Andersen G. R., Valente L., Pedersen L., Kinzy T. G., Nyborg J.* Crystal structures of nucleotide exchange intermediates in the Eef1A-Eef1Balpha complex // Nat. Struct. Biol.–2001.–**8**.–P. 531–534.

12. *Vitagliano L., Masullo M., Sica F., Zagari A., Bocchini V.* The crystal structure of *Sulfolobus solfataricus* elongation factor 1 alpha in complex with Gdp reveals novel features in nucleotide binding and exchange // EMBO J.–2001.–**20**.–P. 5305.

13. *Uetsuki T., Naito A., Nagata S., Kaziro Y.* Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1 alpha // J. Biol. Chem.–1989.–**264**.–P. 5791–5798.

14. *Norbeck J., Blomberg A.* Two-dimensional electrophoretic separation of yeast proteins using a non-linear wide range (pH 3–10) immobilized pH gradient in the first dimension; reproducibility and evidence for isoelectric focusing of alkaline (pI > 7) proteins // Yeast.–1997.–**13**.–P. 1519–1534.

15. *She Q., Singh R. K., Confalonieri F., Zivanovic Y., Allard G., Awayez M. J., Chan-Weiher C. C., Clausen I. G., Curtis B. A., De Moors A., Erauso G., Fletcher C., Gordon P. M., Heikamp-de Jong I., Jeffries A. C., Kozera C. J., Medina N., Peng X., Thi-Ngoc H. P., Redder P., Schenk M. E., Theriault C., Tolstrup N., Charlebois R. L., Doolittle W. F., Duguet M., Gaasterland T., Garrett R. A., Ragan M. A., Sensen C. W., Van der Oost J.* The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2 // Proc. Nat. Acad. Sci. USA.–2001.–**98**.–P. 7835–7840.

16. *Budkevich T. V., Timchenko A. A., Tiktopulo E. I., Negrutskii B. S., Shalak V. F., Petrushenko Z. M., Aksenov V. L., Willumeit R., Kohlbrecher J., Serdyuk I. N., El'skaya A. V.* Extended conformation of mammalian translation elongation factor 1A in solution // Biochemistry.–2002.–**41**.–P. 15342–15349.

17. *Schwede T., Kopp J., Guex N., Peitsch M. C.* SWISS-MODEL: an automated protein homology-modeling server // Nucl. Acids Res.–2003.–**31**.–P. 3381–3385.

18. *Lindahl E., Hess B., van der Spoel D.* GROMACS. 3.0: a package for molecular simulation and trajectory. Analysis // J. Mol. Mod.–2001.–**7**.–P. 306–317.

19. *Schuler L. D., Daura X., van Gunsteren W. F.* An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase // J. Comput. Chem.–2001.–**22**.–P. 1205–1218.

20. *Schuettelkopf A. W., van Aalten D. M. F.* PRODRG – a tool for high-throughput crystallography of protein-ligand complexes // Acta Crystallogr. D.–2004.–**60**.–P. 1355–1363.

21. *Hunenberger P. H., McCammon J. A.* Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatics study // Biophys. Chem.–1999.–**78**.–P. 69–88.

22. *Weber W., Hunenberger P. H., McCammon J. A.* Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: influence of artificial periodicity on peptide conformation // J. Phys. Chem. B.–2000.–**104**.–P. 3668–3675.

23. *Berendsen H. J. C., Postma J. P. M., van Gunsteren W. F., Hermans J.* Interaction models for water in relation to protein hydration // Intermolecular Forces / Ed. B. Pullman.–Dordrecht: D. Reidel Publ. Company, 1981.– P. 331–342.

24. *Hess B., Bekker H., Berendsen H. J. C., Fraaije J. G. E. M.* LINCS: A linear constraint solver for molecular simulations // J. Comp. Chem.–1997.–**18**.–P. 1463–1472.

25. *Essman U., Perela L., Berkowitz M. L., Darden T., Lee H., Pedersen L. G.* A smooth particle mesh Ewald method // J. Chem Phys.–1995.–**103**.–P. 8577–8592.

26. *Berendsen H. J. C., Postma J. P. M., DiNola A., Haak J. R.* Molecular dynamics with coupling to an external bath // J. Chem. Phys.–1984.–**81**.–P. 3684–3690.

27. *Hess B.* Convergence of sampling in protein simulations // Phys. Rev. E.–2002.–**65**.–P. 1–10.

28. *Humphrey W., Dalke A., Schulten K.* VMD – visual molecular dynamics // J. Mol. Graphics.–1996.–**14**.–P. 33–38.

29. *Guex N., Peitsch M. C.* SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling // Electrophoresis.–1997.–**18**.–P. 2714–2723.

30. *Louise-May S., Auffinger P., Westhof E.* Calculation of nucleic acid conformation // Curr. Opin. Struct. Biol.–1996.–**6**.–P. 289–298.

31. *Auffinger P., Louise-May S., Westhof E.* Multiple molecular dynamics simulations of the anticodon loop of tRNAAsp in aqueous solution with counterions // J. Amer. Chem. Soc.–1995.–**117**.–P. 6720–6726.

32. *Auffinger P., Louise-May S., Westhof E.* Molecular dynamics simulations of the anticodon hairpin of tRNAAsp: structuring effects of C–H O hydrogen bonds and of long-range hydration forces // J. Amer. Chem. Soc.–1996.–**118**.–P. 1181–1189.

33. *Vaiana A. C., Westhof E., Auffinger P.* A molecular dynamics simulation study of an aminoglycoside/A-site RNA complex: conformational and hydration patterns // Biochimie.–2006.–**88**.–P. 1061–1073.