

Dynamical insight into *Caenorhabditis elegans* eIF4E recognition specificity for mono- and trimethylated structures of mRNA 5' cap

Katarzyna Rusczyńska-Bartnik · Maciej Maciejczyk · Ryszard Stolarski

Received: 8 February 2010 / Accepted: 6 May 2010 / Published online: 10 June 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Specific recognition and binding of the ribonucleic acid 5' termini (mRNA 5' cap) by the eukaryotic translation initiation factor 4E (eIF4E) is a key, rate limiting step in translation initiation. Contrary to mammalian and yeast eIF4Es that discriminate in favor of 7-methylguanosine cap, three out of five eIF4E isoforms from the nematode *Caenorhabditis elegans* as well as eIF4Es from the parasites *Schistosoma mansoni* and *Ascaris suum*, exhibit dual binding specificity for both 7-methylguanosine- and N²,N²,7-trimethylguanosine cap. To address the problem of the differences in the mechanism of the cap recognition by those highly homologous proteins, we carried out molecular dynamics simulations in water of three factors, IFE-3 and IFE-5 isoforms from *C. elegans* and murine eIF4E, in the *apo* form as well as in the complexes with 7-methyl-GDP and N²,N²,7-trimethyl-GDP. The results clearly pointed to a dynamical mechanism of discrimination between each type of the cap, *viz.* differences in mobility of the loops located at the entrance into the protein binding pockets during the cap association and dissociation. Additionally, our data showed that the hydrogen bond involving the N²-amino group of 7-

methylguanosine and the carboxylate of glutamic acid was not stable. The dynamic mechanism proposed here differs from a typical, static one in that the differences in the protein-ligand binding specificity cannot be ascribed to formation and/or disruption of well defined stabilizing contacts.

Keywords *Caenorhabditis elegans* · Cap recognition · eIF4E isoforms · Molecular dynamics · mRNA 5' cap

Introduction

The 5' terminal structure of eukaryotic RNA polymerase II transcripts (RNA 5' cap) plays a crucial role in gene expression and regulation. The cap is specifically bound to several cellular and viral proteins, including various isoforms of eukaryotic translation factor eIF4E [1], nuclear cap-binding complex CBC [2], DcpS scavenger enzyme [3], poly(A) binding protein PABP [4], poly(A)-specific ribonuclease PARN [5], pokeweed antiviral protein PAP [6], cellular mRNA cap (guanine-N7) methyltransferase [7], human paraneoplastic encephalomyelitis antigen HuD [8], vaccinia virus 2'-O-methyltransferase VP39 [9], influenza virus RNA polymerase [10], and dimethyltransferase TGS1 [11]. The eIF4E factors from vertebrates and yeast were shown to be highly selective for 7-methylguanosine cap (MMG-cap) [12], m⁷GpppN, N=G, A, U or C. In the nematodes *Caenorhabditis elegans* and *Ascaris suum* as well as in the parasitic flatworm *Schistosoma mansoni* a high population of messenger ribonucleic acids (mRNAs) contain a hypermethylated cap form, N²,N²,7-trimethylguanosine cap (TMG-cap), m₃^{2,2,7}GpppN, which is acquired along with a spliced leader during *trans*-splicing of pre-mRNA [13]. Affinity chromatography [14, 15], and

K. Rusczyńska-Bartnik · M. Maciejczyk
Nuclear Magnetic Resonance Laboratory, Institute of
Biochemistry and Biophysics, Polish Academy of Sciences,
02-106 Warsaw, Poland

R. Stolarski (✉)
Division of Biophysics, Institute of Experimental Physics,
Faculty of Physics, University of Warsaw,
02-089 Warsaw, Poland
e-mail: stolarsk@biogeo.uw.edu.pl

M. Maciejczyk
Department of Physics and Biophysics,
University of Warmia and Mazury,
Olsztyn 10-719, Poland

fluorescence titration [16, 17] experiments showed that three out of five *C. elegans* eIF4E isoforms, IFE-1, IFE-2, IFE-5, are capable of binding specifically to the MMG-cap and to the TMG-cap. Two other isoforms, IFE-3, most similar to mammalian eIF4E, and IFE-4, related to the mammalian 4E-homologous protein 4E-HP, bind only to the MMG-cap. The dual binding specificity was also observed for eIF4Es from *S. mansoni* [18] and *A. suum* [19]. Values of the equilibrium dissociation constants, K_d , for some selected complexes of the dual specificity eIF4Es with typical MMG-cap and TMG-cap analogues are shown in Table 1. The K_d values for murine eIF4E are also shown for comparison. The values reported by various groups and derived by various titration methods, mainly fluorescence and isothermal calorimetry, can differ between each other. Nevertheless, the experimental data clearly show that both *C. elegans* IFE-3 and murine eIF4E strongly discriminate in favor of MMG-cap, *ca.* 11 kJ mol⁻¹ and 16 kJ mol⁻¹, respectively, while in the case of *C. elegans* IFE-5 the binding energies are similar. The energy difference between *S. mansoni* eIF4E complexes is *ca.* 5 to 7 kJ mol⁻¹, depending on the titration experiment [18].

The TMG-cap occurs at the 5' terminus of small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and in telomerase RNA TLC1 [11]. It is specifically recognized by Snurportin1 [20], a receptor for spliceosomal small nuclear particles (snRNPs).

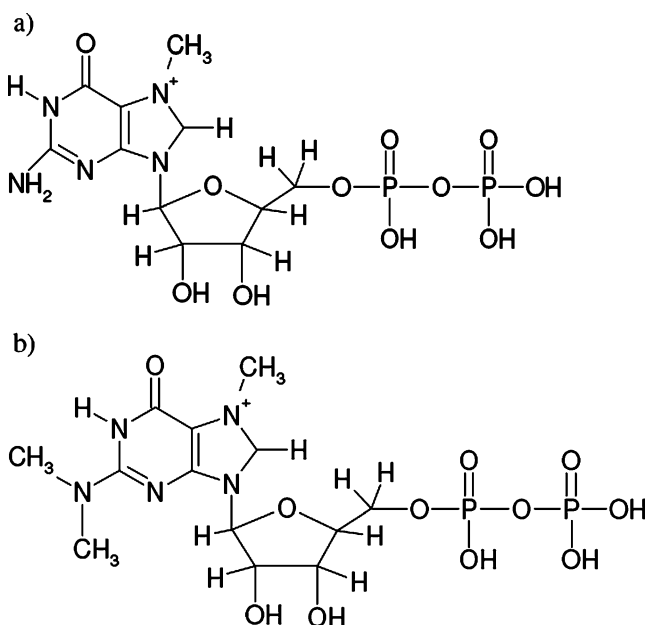
As shown by X-ray crystallography [9, 12, 21–29] and multidimensional NMR [30] most of the cap-binding proteins converged at a common mechanism of the cap recognition *via* stacking of the 7-methylguanine moiety in between two aromatic amino acid side chains. The 7-methylguanine base possesses a net positive charge, which seems indispensable for its proper recognition, *i.e.*, the 7-methylguanosine cannot be replaced by guanosine in the cap structure. In the snurportin1 complex with the dinucleotide TMG-cap analogue [20] the sandwich stacking

involves one tryptophan, and two bases of the cap, the first one, trimethylated, and the second one, unmethylated. In dimethyltransferase TGS1 the 7-methylguanine moiety is stacked only on one tryptophan and a serine polar side chain limits the binding pocket on the opposite side [11]. Except for the cation- π stacking, the cap is also stabilized by a network of hydrogen bonds, direct or water mediated salt bridges, as well as less specific van der Waals and hydrophobic contacts. Only a few exceptions have been found yet where the recognition specificity is entirely mediated through hydrogen bonds and van der Waals contacts with 7-methylguanine, *i.e.*, cap methyltransferase [7], and reovirus polymerase λ 3 [31].

In the mammalian, plant, and yeast eIF4E-cap complexes, two tryptophan aromatic residues take part in the sandwich cation- π stacking with the 7-methylguanine moiety. Two hydrogen bonds involve N1H and N²H atoms of 7-methylguanosine and the carboxyl group of a conserved glutamic acid, and one hydrogen bond is observed between O⁶ of m⁷G and the backbone amide nitrogen. Additional stabilizing interactions are between the phosphate chain of the cap and the arginine/lysine side chains of the protein. The first crystallographic structure of dual specificity eIF4E from *Schistosoma mansoni* in the complexes with two dinucleotide MMG-cap analogues [18] showed a similar binding mode as for the single specificity eIF4Es. The only difference seems to be in the conformation of E90 carboxylate of *S. mansoni* eIF4E that is rotated by $\sim 80^\circ$ in comparison to the orientation of the equivalent E103 [12, 23] in murine eIF4E. This precludes the formation of two strong hydrogen bonds with the 7-methylguanine moiety. A similar rotation was also observed in the human eIF4E ternary complex with eIF4GI peptide and a glycerol molecule located in the cap-binding center [32]. Still, the contribution of that conformational change to MMG-cap vs. TMG-cap binding specificity remains unclear. The NMR analysis of the MMG-cap and TMG-cap complexes with *S. mansoni* eIF4E [18] showed substantial chemical shift perturbation for *ca.* 15 amino acids, most of them distributed around the cap-binding pocket. Based on the crystallographic and NMR data the authors suggested that intrinsic and specific conformational flexibility of the *S. mansoni* eIF4E plays a crucial role in the TMG-cap binding, analogous to an “induced fit” mechanism. On the contrary, combined mutagenesis studies and molecular dynamics simulations of *C. elegans* dual specificity IFE-5 led to a “structural” rather than a “dynamic” model. Larger width and depth of the cap-binding pocket was postulated to be responsible for the TMG-cap binding specificity [17]. Replacement of two amino acids, N64Y/V65L decreased the size of the pocket and gave rise to discrimination against TMG-cap by steric hindrance. However, it was noted that dual specific *A. suum* eIF4E does contain Y64

Table 1 Experimentally derived dissociation constants, K_d [μ M], for the MMG- and TMG-cap analogues from *C. elegans* factors, IFE-3 and IFE-5, and *S. mansoni* eIF4E. The K_d values for murine eIF4E are shown for comparison

| | K_d [μ M] | | Ref. No |
|-------------------------|--------------------|-------------------------------------|---------|
| | m ⁷ GTP | m ₃ ^{2,2,7} GTP | |
| eIF4E murine | 0.01 | 6.99 | [12] |
| IFE-3 <i>C. elegans</i> | 0.20 | 16.7 | [16] |
| | 0.36 | 3.5 | [17] |
| IFE-5 <i>C. elegans</i> | 0.14 | 0.33 | [16] |
| | 0.54 | 0.92 | [17] |
| eIF4E <i>S. mansoni</i> | 0.58 | 11.2 | [18] |



Scheme 1 Chemical structures of the cap analogues. (a) m^7 GDP, analogue of MMG-cap, (b) $m_3^{2,2,7}$ GDP, analogue of TMG-cap

and L65 residues [18]. Unfortunately, discrimination between TMG-cap and MMG-cap by snurportin1 is based on a mechanism that differs from that expected for the eIF4E factors [20, 33].

In order to get an insight into the mechanism of dual specificity in the cap recognition by some of highly homologous eIF4Es, we performed long-lasting molecular dynamics (MD) simulations in water for three selected eIF4E homologues, murine eIF4E as well as *C. elegans* IFE-3 and IFE-5, each of them in the *apo* form and in the complexes with m^7 GDP or with $m_3^{2,2,7}$ GDP (Scheme 1). The results point to a dynamic mechanism of discrimination between the mono- and hypermethylated cap structures.

Theory and methods

Initial setup

The starting structure of the complex of truncated murine eIF4E(28–217) with m^7 GDP for molecular dynamics simulations was taken from crystallography (PDB code: 1EJ1; [23]). The missing atoms of some of the amino acid side chains were completed by SCWRL [34]. The hydrogen atoms were added in Insight II (Accelrys Software Inc., U.S.A.). The starting structures of IFE isoforms were obtained by homology modeling with murine eIF4E(28–217) bound to m^7 GDP as a template, with 51% and 42% sequence identity to IFE-3 and IFE-5, respectively. The multiple sequence alignment was performed by CLUSTAL W [35]. Ten structures for each isoform were obtained

using the program MODELLER [36]. Additional harmonic constraints were introduced for the distances between the protein and the cap atoms that were engaged in hydrogen bonds, salt bridges and van der Waals contacts. Subsequently, the resulting structures were subjected to detailed analysis regarding packing of the residues, steric hindrance, and loop conformations. Based on the analysis one representative structure was chosen for each isoform. Due to high sequence homology the modeled IFE structures were very similar to that of the eIF4E template (Fig. 1), especially regarding the main polypeptide chains. The structures of the complexes of three 4E factors with $m_3^{2,2,7}$ GDP, and of murine eIF4E with GDP, were con-



b)



Fig. 1 Structural comparison of murine eIF4E(28–217) and its two *C. elegans* isoforms, IFE-3 and IFE-5. (a) Sequence alignment. (b) Superposition of the three protein complexes with m^7 GDP. The cap and the amino acids engaged in its stabilizing are marked in bold

structed by adding two methyl groups at N² and by removing the methyl group from N7 in the protein-m⁷GDP complexes, respectively. The *apo* proteins were obtained by removing m⁷GDP from the complexes.

The ESP charges of the isolated ligands were calculated at HF/6–31G(d,p) level using Gaussian 94 (Gaussian Inc., Pittsburg PA, U.S.A.)

Molecular dynamics simulation and analysis

The MD simulations were carried out by the program Sigma [37] using CHARMM22 force field [38]. Each protein or complex was subjected to energy minimization without electrostatic interaction and immersed in an equilibrated TIP3P water box [39] keeping at least 10 Å shell thickness from the protein surface. The simulation procedure consisted of several equilibration MD runs preceded and followed by 500-step energy minimization, and a subsequent regular MD run, as follows. First, energy minimization and 48 ps dynamics of water molecules was performed keeping the protein or the complex immobilized. Second, energy minimization of the protein or of the complex with the water molecules kept fixed was followed by stepwise heating of the whole system from 50 K to 300 K for 82.56 ps. The initial velocities at each temperature were taken from Maxwell-Boltzmann distribution. The equilibration MD runs were performed in the nVT ensemble and the regular MD simulations were performed in the npT ensemble [40] at temperature T=300 K and pressure p=1 atm. The SHAKE algorithm [41] was applied to constrain the bonds. The electrostatic interactions were calculated by multiple time step [42] with a double cut-off at 6 Å and 10 Å. Short-, middle-, and long range interactions, according to particle-mesh Ewald method [43, 44], were calculated for an integration time-step of 2, 4, and 12 fs, respectively. The simulations were run for 5 ns in the case of the *apo* proteins or for 10 ns in the case of the complexes, in order to reach at least partial equilibrium according to a stability criterion for the fluctuations of root-mean-square-deviation (RMSD) of the proteins' C_α atoms.

The conformations of the solute on a simulated trajectory were written down every 0.96 ps and analyzed regarding interatomic distances and torsion angles. Essential dynamics (ED) analysis of selected, equilibrium parts of the MD trajectories was performed according to Amadei et al. [45].

Results

An experimentally observed equilibrium association constant K_{as} expressed in terms of the molar concentrations of the reactants in a protein-ligand association is related to the

standard Gibbs free energy ΔG° of the association process at temperature T, $\Delta G^\circ = RT \ln K_{as}$. Hence K_{as} is a quantitative measure of the ligand affinity for the protein. Comparison of the K_{as} values for a series of structurally modified cap analogues enabled parsing of ΔG° into separate contributions from various stabilizing contacts inside the eIF4E cap-binding pocket [12]. Bearing in mind an approximate character of the approach due to lack of additivity of the entropic terms [46, 47], combination of the crystallographic structure with such ΔG° analysis provided molecular mechanism of specific binding between the cap and eIF4E [12]. However, applying the procedure to detect the discrimination mechanism between MMG- and TMG-cap by some eIF4Es [17, 18] has failed. The structures of IFE isoforms derived by homology modeling were very similar to that of the eIF4E template due to high sequence homology (Fig. 1). Therefore, the structural differences of potential importance for the MMG- vs. TMG-cap binding selectivity have not been identified. This prompted us to evaluate a discrimination mechanism of a dynamic type. The equilibrium association constant, $K_{as} = k_{+1}/k_{-1}$, depends on the ligand ability to form and leave the complex, expressed by kinetic rate constants k₊₁ and k₋₁, respectively. Higher k₊₁ and/or lower k₋₁ values give rise to an increase of K_{as}. Since it was impossible to calculate theoretically the rate constants from all-atom MD simulations, we assumed that the MD analysis of the *apo* proteins provided some information on k₊₁ that reflects accessibility of the MMG-cap analogue and of the TMG-cap analogue for the binding sites of the three eIF4Es. Similarly, the MD analysis of the three factors, each bound to either MMG-cap or TMG-cap, might provide some hints on the stability of the complexes that influences their dissociation kinetic constants k₋₁. It should be emphasized that the simulated *apo* structures were generated by removing the ligand, and therefore they differ from those derived experimentally, at least in the case of the NMR structure of murine *apo*-eIF4E [48]. The simulated *apo* structures may be regarded as those, which approximate final conformations the ligands would dock into, especially if the cap phosphate chain anchors first as postulated by a two-state model of the cap-binding [12] (see below). It is also worth noting, that the large structural fluctuations observed in the NMR structure of *apo*-eIF4E do not occur during our MD simulations.

The MD trajectories of the *apo* eIF4Es display enhanced flexibility of the loops around the entrance to the cap-binding pocket (Fig. 2), especially S1–S2 and S7–S8 loops, while the secondary structure elements remained unchanged. This general view of the dynamic behavior is confirmed by the experimental data derived for the *apo* form of human eIF4E by multidimensional NMR [48], for the cap-free human eIF4E in the complex with glycerol

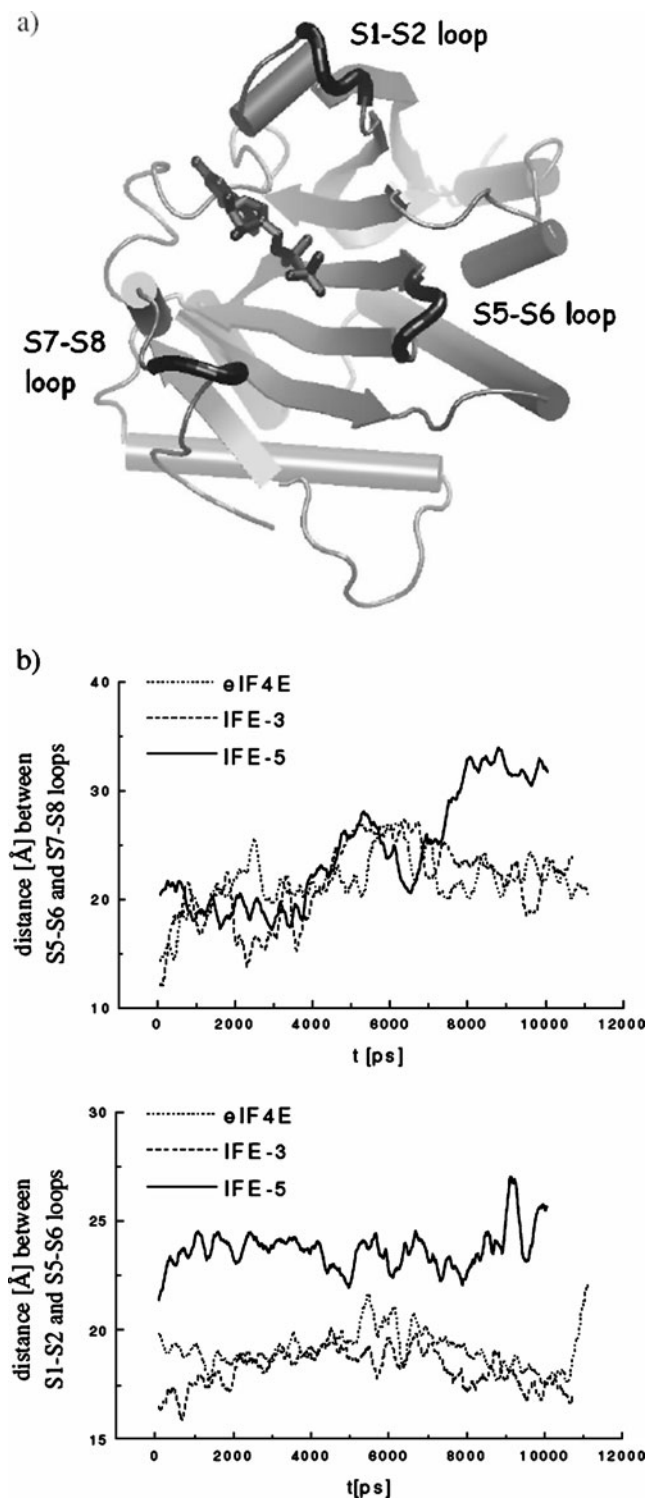


Fig. 2 Analysis of the cap accessibility into the binding pockets. **(a)** Location of the flexible loops in the eIF4E structure. **(b)** Time course of the distances between, (upper) S7–S8 and S5–S6 loops, measured for C $_{\alpha}$ atoms of S209 and K159 for apo-eIF4E and apo-IFE-3, and for C $_{\alpha}$ atoms of Q217 and K159 for apo-IFE-5, (lower) S5–S6 and S1–S2 loops, measured for C $_{\alpha}$ atoms of K159 and R52. The lines are marked as follows: apo-eIF4E dotted, apo-IFE-3 dashed, apo-IFE-5 solid

that seems to represent an intermediate state between the *apo* and the cap-bound form [32], and for the murine factor by hydrogen-deuterium exchange combined with electrospray mass spectrometry [49]. The secondary elements were preserved in *apo* eIF4E while the loops exhibited mobility on the ns-ps time scale that became abrogated upon the cap binding. The structural differences in the regions of loops S1–S2, S3–S4, S5–S6, and S7–S8 (Fig. 2a) resulted in the formation of the positively charged pocket to anchor the cap phosphate chain (loops S1–S2 and S7–S8), and the formation of the stacking triad and hydrogen bonding with the 7-methylguanine moiety *via* locking the W56 hinge (loop S1–S2) and rotating W102 (loop S3–S4) into the cap-binding site. The ability of W102 containing loop S3–S4 to rotate from the open (*apo*) state into the closed form was confirmed by both crystallography [32] and NMR studies [48]. Moreover, an alternative orientation of W102 indole ring was found by the diffraction studies of the eIF4E complexes with 7-benzyl-GMP and 7-(p-fluorobenzyl)-GMP. The W102 residue flips through 180° [50] in comparison with those complexes which contain the 7-methylguanosine caps [12, 23, 24]. All those observations agree with the two-state model of the cap-binding [12], in which anchoring of the phosphate chain was followed by a cooperative formation of the stacking triad and hydrogen bonds. This model, derived from analysis of the fluorescence titration of eIF4E with structurally modified cap analogues, was confirmed by kinetic stopped flow experiments [51, 52]. Contrary to our analysis, Slepnev et al. [53] performed and analyzed the stopped flow experiments differently, and postulated a one-step mechanism. However, our results are confirmed not only by the standard analysis of the stop-flowed kinetic traces under pseudo first-order conditions [51] but also by running the experiments under the second-order conditions combined with numerical integration of the suitable differential kinetic equations [52] by the state-of-the-art program DynaFit [54].

The flexibility of the loops seem to be crucial for the discrimination between m⁷GDP and m₃^{2,2,7}GDP by murine eIF4E and *C. elegans* IFE-3 and IFE-5. The calculated distances between S7–S8 and S5–S6, and between S5–S6 and S1–S2, on the final parts of the MD trajectories (Fig. 2b) are *ca.* 10 Å greater for IFE-5 that binds the TMG-cap than for IFE-3 and murine eIF4E that are specific for the MMG-cap only. Hence, the TMG-analogue with two additional methyl groups can easily penetrate the IFE-5 binding pocket contrary to other factors.

A similar analysis of the dynamics of the protein-cap complexes shows stable contacts between the cap phosphates and the arginine/lysine side chains, irrespective of the bound analogue. This is consistent with the anchoring character of the phosphate groups in the cap stabilization

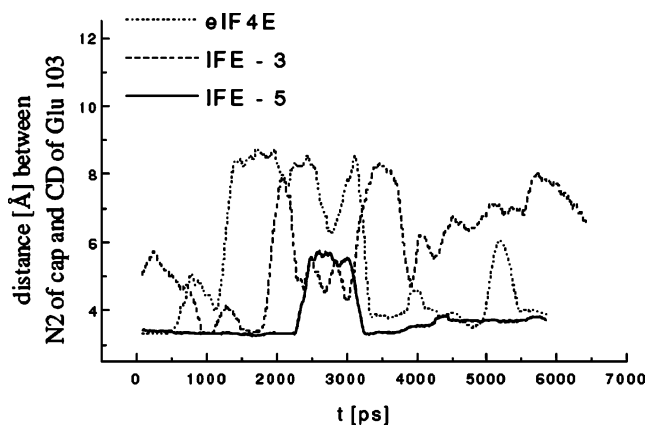


Fig. 3 Time dependence of the distance (in Å) between N² of m⁷GDP and Glu103 side chain carboxyl in murine eIF4E (dotted), IFE-3 (dashed), and IFE-5 (solid)

inside the eIF4E binding slot [12]. On the contrary, the mutual orientations of the rings in the cation- π stacking triad undergo larger fluctuations in respect to the starting structure. As expected, the largest changes are observed in the eIF4E-GDP complex that is stabilized by weaker π - π stacking, i.e., a perpendicular orientation of the W56 and G rings and a shift of W102 deeper into the binding pocket. In the case of all the other complexes the stacking triad is kept principally unchanged, with relatively larger fluctuations of W102. A temporary increase of the distance between W102 and m⁷G rings are observed in the eIF4E-m₃^{2,2,7}GDP and IFE-3-m₃^{2,2,7}GDP, but not in IFE-5-m₃^{2,2,7}GDP, complexes. On the other hand, the hydrogen bond between the N²-amino group of m⁷G and the E103 carboxylate (Fig. 3) is being broken and reformed due to shifts of the E103 side chain into the bulk solvent. Inherent “plasticity” of that part of the cap-binding center, resulting in a rotation of E103 side chain and a movement of W102 side chain, was also observed in the cap-free eIF4E ternary complex with eIF4GI peptide and glycerol in the cap-binding pocket [32]. Therefore, the low affinity of the TMG-cap for murine eIF4E cannot be explained by lack of that hydrogen bond due to the hypermethylation.

The distance between loops S7–S8 and S5–S6 in the eIF4E-cap complexes (Fig. 4) is shown to correlate with the affinity of various cap types for the eIF4E isoforms. Binding of m⁷GDP to murine eIF4E results in bend of S7–S8 toward S5–S6 that might help to keep the ligand in the binding centre by water mediated interaction between K159 and, e.g., S209. As shown previously [55] K159 is very important for binding the capped mRNA. Closing the entrance to the cap-binding pocket by decreasing the distance between S209 and K159 was also observed in MD simulations of phosphorylated eIF4E [25]. In the complexes with the low affinity ligands, m₃^{2,2,7}GDP and GDP, the distance between the loops is larger, thus making it easier for those ligands to

leave the complexes. Similarly, the distance between the loops in the IFE-3 complex with m⁷GDP is much smaller than in the IFE-3 complex with m₃^{2,2,7}GDP, while for both IFE-5 complexes it is kept fairly large, irrespective of the ligand. The analysis of the overall dynamics of the cap-bound eIF4E isoforms was also carried out by ED analysis of the covariance matrix of the atomic displacements [45]. The scalar products of the the vectors representing the normalized C _{α} displacements and the eigenvectors corresponding to the largest eigenvalue ($\lambda=1$) show non-Gaussian distributions with 2–3 maxima. This can be interpreted as correlated, long-range movements, in which the loops oscillate around several mean positions (Fig. 5). The histograms for the scalar products of the eigenvectors corresponding to lower eigenvalues ($\lambda=10$) are Gaussian, as expected for equilibrated, independent and harmonic motions.

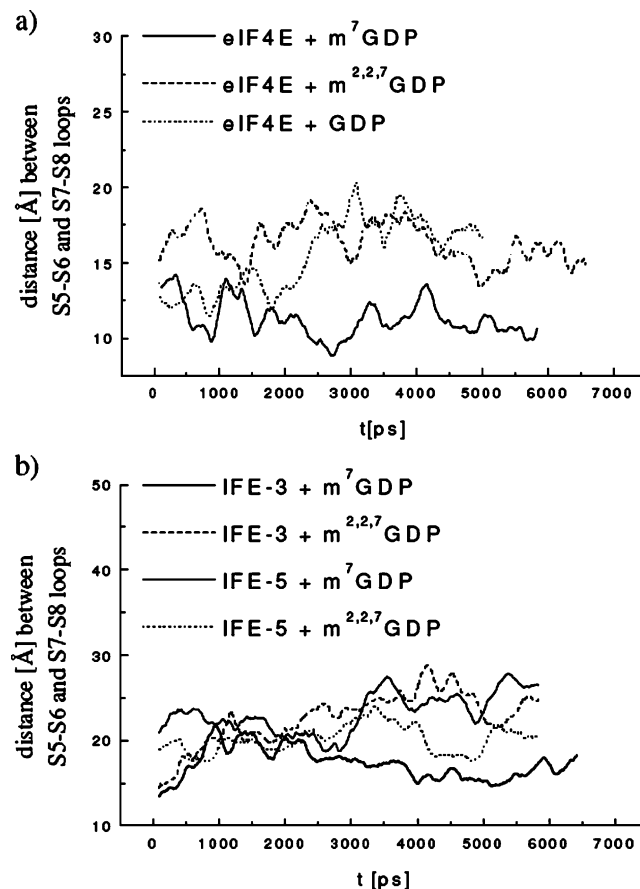


Fig. 4 Analysis of the cap propensity to leave the cap-binding pocket. Time course of the distances between S7–S8 and S5–S6 loops, measured for C _{α} atoms of S209 and K159 (IFE-3, eIF4E), and for C _{α} atoms of Q217 and K159 (IFE-5). (a) Murine eIF4E bound to m⁷GDP (solid), to m₃^{2,2,7}GDP (dashed), and to GDP (dotted). (b) IFE-3 bound to m⁷GDP (solid bold) and to m₃^{2,2,7}GDP (dashed), and IFE-5 bound to m⁷GDP (solid) and to m₃^{2,2,7}GDP (dotted)

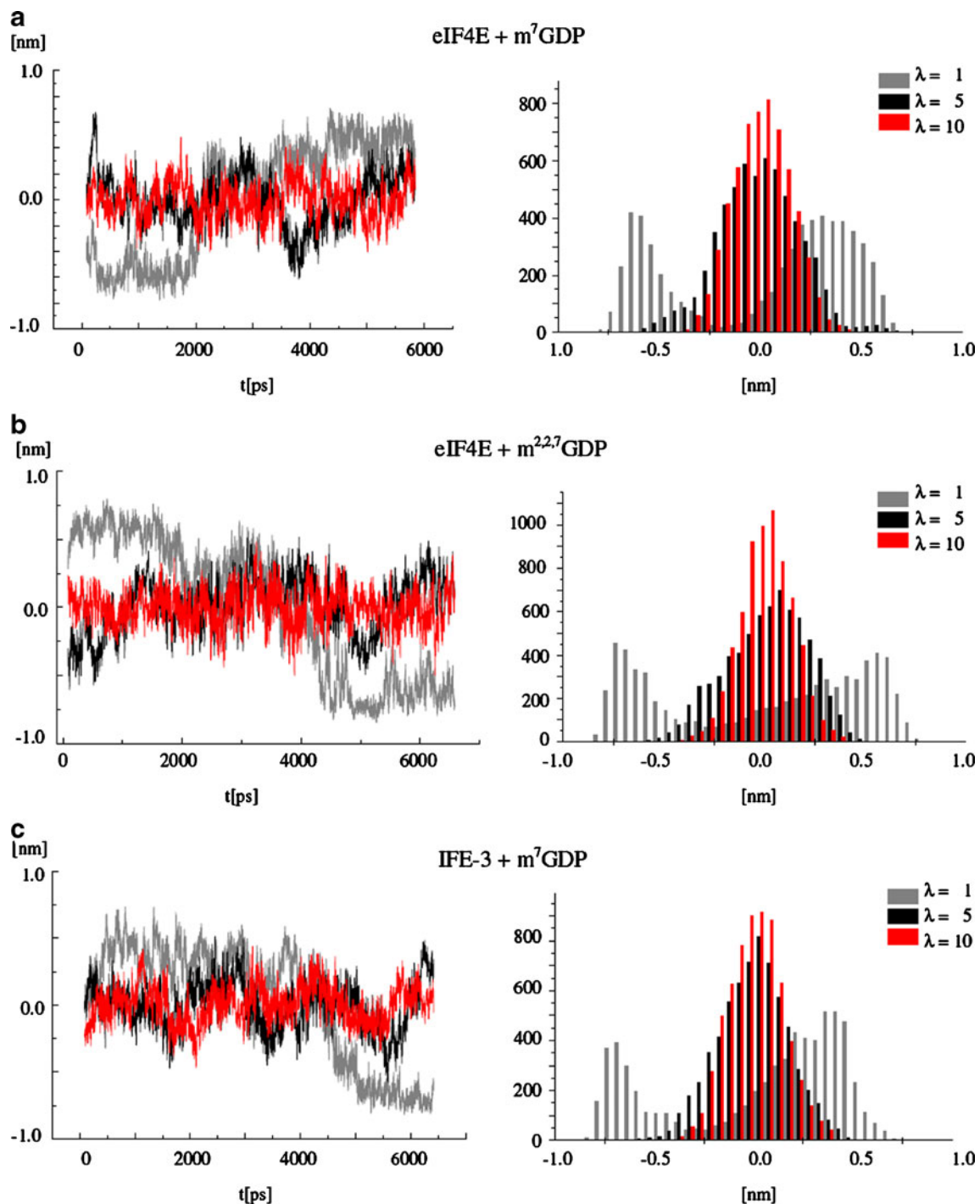


Fig. 5 Mobility of the *apo*- and cap-bound eIF4Es by essential dynamics analysis. Motions along the first ($\lambda=1$), fifth ($\lambda=5$), and tenth ($\lambda=10$) eigenvectors obtained from the C_{α} coordinates covariance matrix, and the corresponding probability distribution

for the displacements (nm), (a) eIF4E bound to m^7GDP , (b) eIF4E bound to $m_3^{2,2,7}GDP$, (c) IFE-3 bound to m^7GDP , (d) IFE-3 bound to $m_3^{2,2,7}GDP$, (e) IFE-5 bound to m^7GDP , (f) IFE-5 bound to $m_3^{2,2,7}GDP$

Discussion

The knowledge of the molecular basis of the RNA 5' cap structure recognition by the cap-binding proteins is a prerequisite for understanding possible mechanisms of

the cap functioning in various types of the gene expression processes in eukaryotes, such as translation initiation, mRNA splicing, and export of RNAs to the cytoplasm. It seems that various evolutionary unrelated cap-binding proteins converged on a similar general

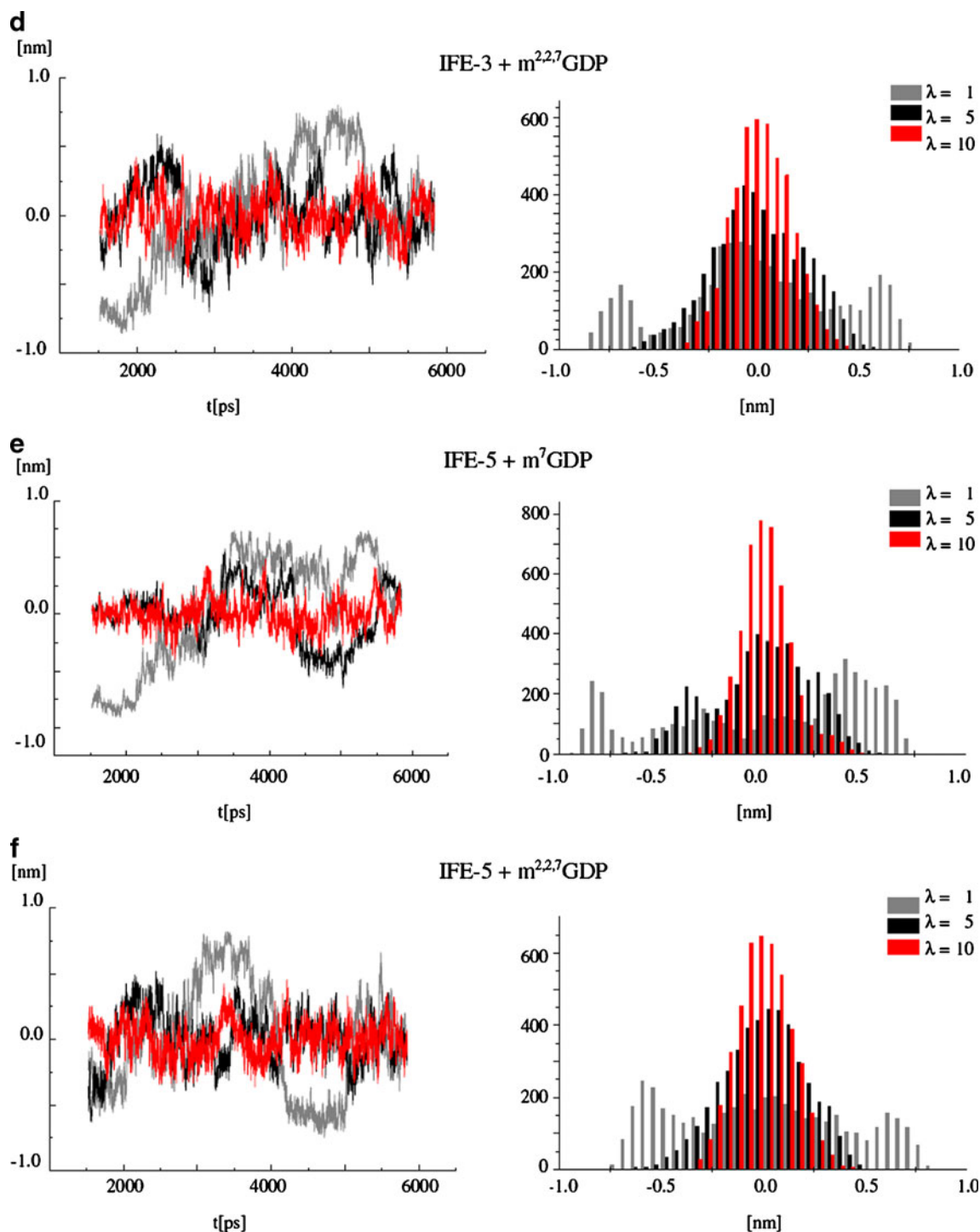


Fig. 5 (continued)

mechanism of the cap recognition. Subtle modifications of the general recognition mechanism of the cap may lead to differences in the protein functions, e.g., the diverse role of two aromatic amino acids that stack with the 7-methylguanosine moiety [56–58]. The methyl group at N7 in the cap structure imparts a net positive charge to the guanine ring, and results in more efficient stacking

compared with the unmethylated base. Quantum mechanical calculations showed that a typical energy of the cation- π stacking of the m⁷G base in the complex with the tyrosine or tryptophan aromatic ring is in a range $-11.4 \text{ kcal mol}^{-1}$ to $-16.23 \text{ kcal mol}^{-1}$, while the G/Y π - π stacking energy is *ca.* -6 kcal mol^{-1} [59]. Additional methyl groups at N² do not change the stacking ability of

the cap. The stacking energy of $m_3^{2,2,7}G/W276$ in Snurportin 1, $-12.52 \text{ kcal mol}^{-1}$, is close to typical values obtained for the m^7G/W complexes.

The presence of two methyl substituents in the amino group of 7-methylguanosine brakes at least one stabilizing hydrogen bond in the cap-binding protein pockets and may lead to a substantial decrease of the association constants observed for the complexes with TMG-cap. On the other hand, the dual specificity cap-binding proteins possess high homology and structure similarity to those that discriminate between the MMG-cap and the TMG-cap. Hence, it is a great challenge to conceive a molecular model of the discrimination vs. dual specificity for the protein-cap association. The explanations usually do not go beyond formulations like “the differences in the size of the cap-binding pocket in the *C. elegans* isoforms of eIF4E” [20]. Our approach to elucidate the specificity of the caps recognition can be expressed in terms of the ligand ability to enter or leave the *apo* protein binding pocket, since the equilibrium association constant is determined by the ratio of the two rate constants. Although we were not able to calculate the values of the rate constants by all-atom MD simulations, the comparative analysis of the MD trajectories of the *apo*- and cap-bound factors provides a more detailed explanation for the differences in the binding specificity of two *C. elegans* eIF4E isoforms, IFE-3 and IFE-5, than those published hitherto. The dynamic mechanism of the discrimination between two types of the cap may be ascribed to differences in mobility of the loops around the entrance to the protein cap-binding pockets, especially S7–S8 loop. Our results show also higher rigidity of the cation- π stacking triad and of the stabilizing interactions (salt bridges, hydrogen bonds) involving the cap phosphate chain compared with a more flexible character of the hydrogen bonds. The results of our computer modeling are generally consistent with the experimental, structural and dynamic, data [32, 48, 49].

Conclusions

Discrimination between two types of the cap, MMG-cap and TMG-cap, consist neither in the differences in the stacking energy nor in well-defined structural differences inside the cap-binding pocket. Both 7-methylguanosine and its hypermethylated form were found to stack equally well in between two amino acid aromatic rings [59], and the structure of *S. mansoni* eIF4E in complex with the MMG-cap analogues [18], showed a very similar mode of binding to that of the single specificity eIF4Es. MD simulations based on the known structures of the cap-eIF4E complexes provided means to evaluate the discrimination mechanism. Contrary to the comparative analysis of the “static” net of

stabilizing contacts inside the cap-binding pockets of highly homologous eIF4Es, we took into account the differences in the dynamics of the formation and dissociation of the eIF4E-cap complexes. An exact specification of the role of particular amino acids in the proposed dynamical mechanism, e.g., their mutual interactions and/or their interactions with various cap structures, needs further investigation.

Acknowledgments This work was supported by the Polish Ministry of Science and Higher Education, grant Nos. N N301 035 936, N N301 267 137 and BST-1434/BF. The MD simulations were performed in the Interdisciplinary Centre for Mathematical and Computational Modelling (ICM) at the University of Warsaw. We are indebted to Dr. Joanna Trylska for critical reading of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Joshi B, Cameron A, Jagus R (2004) Characterisation of mammalian eIF4E-family members. *Eur J Biochem* 271:2189–2203
- Lewis JD, Izaurralde E (1997) The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem* 247:461–469
- Liu H, Rodgers ND, Jiao X, Kiledjian M (2002) The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J* 21:4699–4708
- Kühn U, Wahle E (2004) Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* 1678:67–84
- Martinez J, Ren Y-G, Thuresson A-C, Hellman U, Åström J, Virtanen A (2000) A 54-kDa fragment of the poly(A)-specific ribonuclease is oligomeric, processive, and cap-interacting poly(A)-specific 3' exonuclease. *J Biol Chem* 275:24222–24230
- Hudak KA, Bauman JD, Tumer NE (2002) Pokeweed antiviral protein binds to the cap structure of eukaryotic mRNA and deurinates the mRNA downstream of the cap. *RNA* 8:1148–1159
- Fabrega C, Hausmann S, Shen V, Shuman S, Lima CD (2004) Structure and mechanism of mRNA cap (Guanine-N7) methyltransferase. *Mol Cell* 13:77–89
- Wang X, Hall TMT (2001) Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat Struct Biol* 8:141–145
- Hodel AE, Gershon PD, Shi X, Wang S-M, Quijcho FA (1997) Specific protein recognition of an mRNA cap through its alkylated base. *Nat Struct Biol* 4:350–354
- Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, Sehr P, Lewis J, Ruigrok RW, Ortin J, Hart DJ, Cusack S (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 15:500–506
- Monecke T, Dickmanns A, Ficner R (2009) Structural basis for m^7G -cap hypermethylation of small nuclear, small nucleolar and telomerase RNA by the dimethyltransferase TGS1. *Nucleic Acids Res* 37:3865–3877
- Niedzwiecka A, Marcotrigiano J, Stepinski J, Jankowska-Anyszka M, Wyslouch-Cieszyńska A, Dadlez M, Gingras A-C, Mak P, Darzynkiewicz E, Sonenberg N, Burley SK, Stolarski R (2002) 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* 319:615–635

13. Blumenthal T (1998) Gene clusters and polycistronic transcription in eukaryotes. *BioEssays* 20:480–487
14. Jankowska-Anyszka M, Lamphear BJ, Aamodt EJ, Harrington T, Darzynkiewicz E, Stolarski R, Rhoads RE (1998) Multiple isoforms of eukaryotic protein synthesis initiation factor 4E in *Caenorhabditis elegans* can distinguish between mono- and trimethylated mRNA cap structures. *J Biol Chem* 273:10538–10542
15. Keiper BD, Lamphear BJ, Desphande AM, Jankowska-Anyszka M, Aamodt EJ, Blumenthal T, Rhoads RE (2000) Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. *J Biol Chem* 275:10590–10596
16. Stachelska A, Wieczorek Z, Ruszczynska K, Stolarski R, Pietrzak M, Lamphear BJ, Rhoads RE, Darzynkiewicz E, Jankowska-Anyszka M (2002) Interaction of three *Caenorhabditis elegans* isoforms of translation initiation factor eIF4E with mono- and trimethylated mRNA 5' cap analogues. *Acta Biochim Pol* 49:671–682
17. Miyoshi H, Dwyer DS, Keiper BD, Jankowska-Anyszka M, Darzynkiewicz E, Rhoads RE (2002) Discrimination between mono- and trimethylated cap structures by two isoforms of *Caenorhabditis elegans* eIF4E. *EMBO J* 21:4680–4690
18. Liu W, Zhao R, McFarland C, Kieft J, Niedzwiecka A, Jankowska-Anyszka M, Stepinski J, Darzynkiewicz E, Jones DNM, Davis RE (2009) Structural insight into parasite eIF4E binding specificity for m7G and m2, 2, 7G mRNA cap. *J Biol Chem* 284:31336–31349
19. Lall S, Friedman CC, Jankowska-Anyszka M, Stepinski J, Darzynkiewicz E, Davis RE (2004) Contribution of trans-splicing, 5'-leader length, cap-poly(A) synergism, and initiation factors to nematode translation in an *Ascaris suum* embryo cell-free system. *J Biol Chem* 279:45573–45585
20. Strasser A, Dickmanns A, Luehrmann R, Ficner R (2005) Structural basis for m(3)G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1. *EMBO J* 24:2235–2243
21. Mazza C, Segref A, Mattaj JW, Cusack S (2002) Large-scale induced fit recognition of an m7G pppG cap analogue by the human nuclear cap-binding complex. *EMBO J* 21:5548–5557
22. Calero G, Wilson K, Ly T, Rios-Steiner JR, Clardy J, Cerione RA (2002) Structural basis of m(7)GpppG binding to the nuclear cap-binding protein complex. *Nat Struct Biol* 9:912–917
23. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK (1997) Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 89:951–961
24. Tomoo K, Shen X, Okabe K, Nozoe Y, Fukuhara S, Morino S, Ishida T, Taniguchi T, Hasegawa H, Terashima A, Sasaki M, Katsuya Y, Kitamura K, Miyoshi H, Ishikawa M, Miura K (2002) Crystal structures of 7-methylguanosine 5'-triphosphate (m(7)GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5', 5'-triphosphate (m(7)GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochem J* 362:539–544
25. Tomoo K, Shen X, Okabe K, Nozoe Y, Fukuhara S, Morino S, Sasaki M, Taniguchi T, Miyagawa H, Kitamura K, Miura K, Ishida T (2003) Structural features of human initiation factor 4E, studied by X-ray crystal analysis, and molecular dynamics simulations. *J Mol Biol* 328:365–383
26. Gu M, Fabrega C, Liu SW, Liu H, Kiledjian M, Lima CD (2004) Insights into the structure, mechanism, and regulation of scavenger mRNA decapping activity. *Mol Cell* 14:67–80
27. Chen N, Walsh MA, Liu Y, Parker R, Song H (2005) Crystal structures of human DcpS in ligand-free and m7GDP-bound forms suggest a dynamic mechanism for scavenger mRNA decapping. *J Mol Biol* 347:707–718
28. Fechter P, Mingay L, Sharps J, Chambers A, Fodor E, Brownlee GG (2003) Two aromatic residues in the PB2 subunit of influenza A RNA polymerase are crucial for cap binding. *J Biol Chem* 278:20381–20388
29. Monzingo AF, Dhaliwal S, Dutt-Chaudhuri A, Lyon A, Sadow JH, Hoffman DW, Robertus JD, Browning KS (2007) The structure of the translation initiation factor eIF4E from wheat reveals a novel disulfide bond. *Plant Physiol* 143:1504–1518
30. Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G (1997) Structure of translation factor eIF4E bound to m⁷GDP and interaction with 4E-binding protein. *Nat Struct Biol* 4:717–724
31. Tao Y, Faretta DL, Nibert M, Harrison SC (2002) RNA Synthesis in a cage-structural studies of reovirus polymerase λ 3. *Cell* 111:733–745
32. Brown CJ, Verma CS, Walkinshaw MD, Lane DP (2009) Crystallization of eIF4E complexed with eIF4GI peptide and glycerol reveals distinct structural differences around the cap-binding site. *Cell Cycle* 8:1905–1911
33. Goette M, Stumpe MC, Ficner R, Grubmüller H (2009) Molecular determinants of snurportin 1 ligand affinity and structural response upon binding. *Biophys J* 97:581–589
34. Brower MJ, Cohen EF, Dunbrack RL (1997) Prediction of protein side-chains rotamers from backbone dependent rotamer library: a new homology modeling tool. *J Mol Biol* 267:1268–1282
35. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
36. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815
37. Mann G, Yun RH, Nyland L, Prins J, Board J, Hermans J (2002) Sigma MD program and a generic interface applicable to multifunctional programs with complex, hierarchical command structure. In: Schlick T, Gan HH (eds) *Computational methods for macromolecules: challenges and applications*. Springer, Berlin, pp 129–145
38. MacKerell Jr AD, Bashford D, Bellot M et al (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem* 102:3586–3616
39. Jorgensen W, Chandrasekhar J, Madura J, Impey R, Klein M (1983) Comparison of simple potential function for simulating liquid water. *J Chem Phys* 79:926–936
40. Berendsen HJC, Postma JCM, van Gunsteren WE, Di Nola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
41. Ryckaert JP, Cicciotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341
42. Tuckerman NE, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* 97:1990–2001
43. Darden TA, York DM, Pedersen LG (1993) Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
44. Schlick T, Skeel R, Brünger A, Kale L, Board JA, Hermans J, Schulten K (1999) Algorithmic challenges in computational molecular biophysics. *J Comput Phys* 151:9–48
45. Amadei A, Linssen ABM, Berendsen HJC (1993) Essential dynamics of proteins. *Proteins Struct Funct Bioinf* 17:412–425
46. Mark AE, van Gunsteren WF (1994) Decomposition of the free energy of a system in terms of specific interactions. Implications for theoretical and experimental studies. *J Mol Biol* 240:167–176
47. Boresch S, Archontis G, Karplus M (1994) Free energy simulations: the meaning of the individual contributions from a component analysis. *Proteins Struct Funct Bioinf* 20:25–33
48. Volpon L, Osborne MJ, Topisirovic I, Siddiqui N, Borden KL (2006) Cap-free structure of eIF4E suggests a basis for conformational regulation by its ligands. *EMBO J* 25:5138–5149

49. Rutkowska-Wlodarczyk I, Stepinski J, Dadlez M, Darzynkiewicz E, Stolarski R, Niedzwiecka A (2008) Structural changes of eIF4E upon binding to the mRNA 5' monomethylguanosine and trimethylguanosine cap. *Biochemistry* 47:2710–2720
50. Brown CJ, McNae I, Fischer P, Walkinshaw MD (2007) Crystallographic and mass spectrometric characterisation of eIF4E with N⁷-alkylated cap derivatives. *J Mol Biol* 372:7–15
51. Blachut-Okrasinska E, Bojarska E, Niedzwiecka A, Chlebicka L, Darzynkiewicz E, Stolarski R, Stepinski J, Antosiewicz JM (2000) Stopped-flow and Brownian dynamic studies of electrostatic effects in the kinetics of binding of 7-methyl-GpppG to the protein eIF4E. *Eur Biophys J* 29:487–498
52. Niedzwiecka A, Stepinski J, Antosiewicz JM, Darzynkiewicz E, Stolarski R (2007) Biophysical approach to studies of cap-eIF4E interaction by synthetic cap analogs. *Meth Enzymol* 430:209–245
53. Slepnev SV, Darzynkiewicz E, Rhoads RE (2006) Stopped-flow kinetic analysis of eIF4E and phosphorylated eIF4E binding to cap analogs and capped oligoribonucleotides. *J Biol Chem* 281:14927–14938
54. Kuzmic P (1996) Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV protease. *Anal Biochem* 237:260–273
55. Zuberek J, Jemielity J, Jablonowska A, Stepinski J, Dadlez M, Stolarski R, darzynkiewicz E (2004) Influence of electric charge variation at residues 209 and 159 on the interaction of eIF4E with the mRNA 5' terminus. *Biochemistry* 43:5370–5379
56. Hodel AE, Gershon PD, Shi X, Wang S-M, Quioco AA (1997) Specific protein recognition of an mRNA cap through its alkylated base. *Nat Struct Biol* 4:350–354
57. Fechter P, Brownlee GG (2005) Recognition of mRNA cap structures by viral and cellular proteins. *J Gen Virol* 86:1239–1249
58. Worch R, Jankowska-Anyszka M, Niedzwiecka A, Stepinski J, Mazza K, Darzynkiewicz E, Cusack S, Stolarski R (2009) Diverse role of three tyrosines in binding of the RNA 5' cap to the human nuclear cap binding complex. *J Mol Biol* 385:618–627
59. Worch R, Stolarski R (2008) Stacking efficiency and flexibility analysis of aromatic amino acids in cap-binding proteins. *Proteins Struct Funct Bioinf* 71:2026–2037