UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

Corso di Laurea Magistrale

Biologia Molecolare e Applicata

Caratterizzazione delle proprietà strutturali e dinamiche dei complessi eIF4E-ligando attraverso un approccio computazionale: verso il design di un nuovo gruppo di inibitori della traduzione degli mRNA

Unveiling the structure/dynamical properties of eIF4E-ligand complexes through computational approach: towards the design of novel group of mRNAs translation inhibitors

Tesi di Laurea Magistrale
di:
Roscioni Agnese

Relatore
Chiar.mo Prof.
Di Marino Daniele

Correlatore
Dott.
Di Leva Francesco Saverio

Sessione Straordinaria, Febbraio 2022

Anno Accademico 2020-2021

Riassunto esteso in italiano

Il fattore di trascrizione eIF4E è fondamentale per molti processi fisiologici, ma la sua sovrapproduzione è stata collegata a diverse malattie come il cancro e patologie connesse allo spettro dell'autismo. Per questo motivo diversi gruppi di ricerca nel hanno sviluppando inibitori di diversa natura per bloccare questa proteina. Una delle molecole più promettenti da questo punto di vista è il 4EGI. Il nostro gruppo di ricerca guidato dal Prof. Daniele Di Marino in collaborazione con un altro gruppo di ricerca dell'Università Federico II di Napoli, guidato dal Prof. Francesco Saverio Di Leva, stanno svilupppando una serie di analoghi per migliorare l'attività farmacologica del 4EGI. Per far ciò una comprensione approfondita del legame tra 4EGI e eIF4E è essenziale.

Il complesso eIF4E-4EGI è stato cristallizzato nel 2014 (PDB id 4TPW), ma la caratterizzazione dinamica dell'interazione tra il ligando e la proteina non è mai stata eseguita. Nel 2021 un altro cristallo della struttura dell'eIF4E legato ad un'altra molecola è stato rilasciato (PDB id 7MEU). Nel nuovo cristallo la proteina eIF4E è stata cocristallizzata con una molecola che ha una struttura chimica simile al 4EGI (bifenile). A causa della similarità delle due molecole entrambi i binding mode potrebbero essere possibili e proprio per questo

motivo è necessario un esame approfondito della dinamica dei due complessi per comprendere quale sia il binding mode preferenziale.

Inoltre, negli ultimi anni diversi esperimenti volti a caratterizzare l'attività del 4EGI sono stati effettuati e sono sorti alcuni dubbi riguardo alla struttura del complesso descritta dal cristallo 4TPW.

Al fine di semplificare questo scenario poco chiaro, diverse simulazioni di dinamica molecolare classica sono state portate avanti; infatti, in collaborazione con Vincendo Maria D'Amore, studente di dottorato presso l'università di Napoli, sono stati effettuati 2μs di simulazione di diversi sistemi: l'eIF4E non legato a nessuna proteina (PDB ID 1IPC), l'eIF4E legato al 4EGI (PDB ID 4TPW), l'eIF4E legato al bifenile (PDB ID 7MEU), il 4EGI dockato nella tasca di legame del bifenile.

In questa tesi sono presentati 2μs della simulazione di dinamica molecolare effettuata sul complesso eIF4E-4EGI considerando il binding mode proposto nel cristallo pubblicato nel 2014.

Dall'analisi dell'RMSD è emerso che la proteina risulta stabile nel tempo mentre il ligando (4EGI) tende a fluttuare molto perdendo rapidamente il binding mode descritto nel cristallo.

Successivamente è stata effettuata un'analisi sul movimento dei centri di massa del ligando che ha permesso di sottolineare ancora come la molecola si muova nel tempo allontanandosi prima dalla tasca di legame e successivamente riavvicinandosi, senza tornare però alla posizione iniziale. Dall'analisi dei contatti del ligando rispetto alla proteina si nota ulteriormente come i contatti alla fine della simulazione vadano ad aumentare e non a diminuire confermando un ampio riarrangiamento del binding mode. Le ultime due analisi effettuate per confermare questo comportamento sono state la valutazione della distanza del ligando dalla tasca di legame, dove si vede come la molecola si allontani per poi riavvicinarsi, e l'RMSF dei residui, che mostra un'alta fluttuazione dei residui della tasca di legame.

In fine sono state effettuate una cluster analisi, che mostra in successione le diverse pose acquisite dal ligando nel corso della simulazione, e un'analisi delle cavità che si genera intorno al ligando. Queste due analisi ci consentono di apprezzare in modo chiaro e diretto l'ampio riarrangiamento della posizione del ligando rispetto alla proteina con la conseguente formazione di una nuova tasca di legame all'interno della proteina.

Altri studi saranno necessari sia sperimentali che teorici per confermare i dati ottenuti fino ad ora ma è possibile affermare che i risultati emersi da questo

lavoro forniscono un nuovo punto di vista nella valutazione del binding 4EGI all'interno della cavità di legame della proteina eIF4E.

INDEX

1. INTRODUCTION 1

1.1 Summary of the thesis 1

1.2 eIF4E, why is it important? 3

1.3 4EGI: an important small molecule 18

1.4 How to deeply understand the interaction between eIF4E and 4EGI: the importance of molecular dynamic 25

1.5 Previous simulations performed on eIF4E with ligands 38

1.6 Aim of the thesis 41

2.1 MATERIALS AND METHODS 43

2.1 Molecular dynamics workflow and system preparation 43

2.2 Minimization 50

2.3 Thermalization 53

2.4 Run production 59

2.5 Analysis 59

3. RESULTS 64

3.1 Molecule preparation 64

3.2 Binding modes 65

3.3 RMSD calculations 67

3.4 Evaluation of 4EGI COM and contacts with protein 69

3.5 Analysis of biding pocket residues 72

3.6 Cluster and cavity analysis 74

3.7 Pose comparison 78

4. DISCUSSION 80

4.1 Stability of the complex 80

4.2 Ligand shift and binding pocket 81

4.3 Evolution of the complex 84

5. CONCLUSIONS 87

5.1 Conclusion of the study 87

5.2 Future steps 87

Appendix II

.mdp file II

Bibliography XIII

1. INTRODUCTION

1.1 Summary of the thesis

The translational initiation factor eIF4E is fundamental for several human physiological processes, but its up-regulation has been linked to a plethora of diseases, such as cancer and neurological disease of autism spectrum disorders. Due to its importance, different kind of inhibitors are being developed (small molecules, miRNA, peptides etc,) from different research group. One of the most promising small molecules is 4EGI.

Our research group, led by Prof. Daniele Di Marino in collaboration with a research group from University Federico II of Naples, led by Prof. Francesco Saverio Di Leva, are developing a bunch of analogs molecules to improve the 4EGI pharmacological activity. To do this a deep understanding of the binding mode of the lead compound (*i.e.,* 4EGI) is crucial.

The eIF4E-4EGI complex was crystalized in 2014 (PDB id 4TPW), but its structural and dynamical features were never examined. In 2021 a different crystal structure of eIF4E bound to another small molecule was released (PDB id 7MEU). The new ligand in the 7MEU structure is biphenyl, and it has a chemical structure similar to 4EGI. Due to the similarity between the two

molecules both binding modes could be possible and, for this reason, a deep investigation of the dynamic of the complex is necessary.

In the last years several experiments have been performed on 4EGI and eIF4E, and some concerns have been raised on the stability of the binding mode observed in the crystal structure deposited in 2014.

To unveil these concenters different molecular dynamics (MD) simulations have been carried out. 2μs of classical MD simulation of: I) eIF4E protein in the unbound conformation (PDB id 1IPC); II) eIF4E-4EGI complex (PDB ID 4TPW); III) eIF4E-biphenyl complex (PDB ID 7MEU); IV) eIF4E-4EGI bind complex in which the 4EGI molecule was docked into the biphenyl binding pocket.

In this thesis is presented only the 2μs of MD simulation performed on the eIF4E-4EGI complex crystalized in 2014.

All the analyses performed point out that the 4EGI binding mode in the 4TPW crystal structure published in 2014 is highly unstable rearranging its binding in a new orientation. The new binding mode of 4EGI observed thanks to the extensive MD simulation mimics the biphenyl binding mode observed in the

7MEU structure crystallized in 2021. Other experimental and theroetical studies are necessary to confirm these preliminary data but the results reported in this thesis give a new point of view in the evaluation of the binding between 4EGI and eIF4E.

1.2 eIF4E, why is it important?

eIF4E stands for: Eukaryotic translation initiation factor 4E. It is one of the components of the eIF4F complex that is crucial for cap-dependent translation in cells. eIF4E recognizes and binds the 7-methylguanosine mRNA cap during an early stage of the initiation of protein synthesis. Moreover, it represents the quantitatively limiting factor, and consequently, it strongly influences the translation initiation rates (Duncan et al., 1987; Hiremath et al., 1985).

In mammals, eIF4E is a 25 kDa protein that has been identified by affinity chromatography (Sonenberg et al., 1979). It is 217 amino acids long and it localizes in the nucleus (Dostie et al., 2000), in the cytoplasm (Dostie et al., 2000; Shih et al., 2012), P-bodies (Ferraiuolo et al., 2005; James et al., 2010; Kamenska et al., 2014; Martínez et al., 2015) and in stress granules (Shih et al., 2012).

The role of eIF4E in protein synthesis

Gene expression involves the transcription of the genetic information encoded in the DNA into messenger RNAs (mRNAs), followed by translation of the mRNA into a polypeptide sequence (Figure 1). The translation process can be divided into four phases: initiation, elongation, termination, and ribosome recycling. Translation can be cap-dependent or cap-independent. Cap-dependent translation requires the assembly of the eukaryotic Initiation Factor 4F (eIF4F) complex and its binding to the mRNA 5'-cap, followed by the formation of the 43S preinitiation complex (43S PIC). The 43S PIC consists of the small ribosomal subunit 40S, eukaryotic translation initiation factors eIF1, eIF1A, eIF3, eIF5 and the eIF2-GTP-Met-tRNAi ternary complex (Merrick & Pavitt, 2018). The eIF4F complex then recruits the 43S PIC on the mRNA, leading to the formation of the 48S translation initiation complex (48S IC). The 48S complex, the structure of which has been recently resolved (Querido et al., 2020), scans the mRNA in the 5'-3' direction until a start codon is found.

eIF4F is a heterotrimeric complex composed by eIF4E, the DEAD-box helicase eIF4A and eIF4G (Santini & Klann, 2014). All eIF4G proteins bind eIF4E through a motif of sequence Y(X)4LF, where X is variable and F is hydrophobic (Altmann et al., 1997; Mader et al., 1995). Although the Y(X)4LF

motif may not be sufficient for full affinity binding of mammalian eIF4G to eIF4E, the fact that mutations in this sequence can abrogate binding indicates that it is necessary for this protein-protein interaction (Mader et al., 1995).

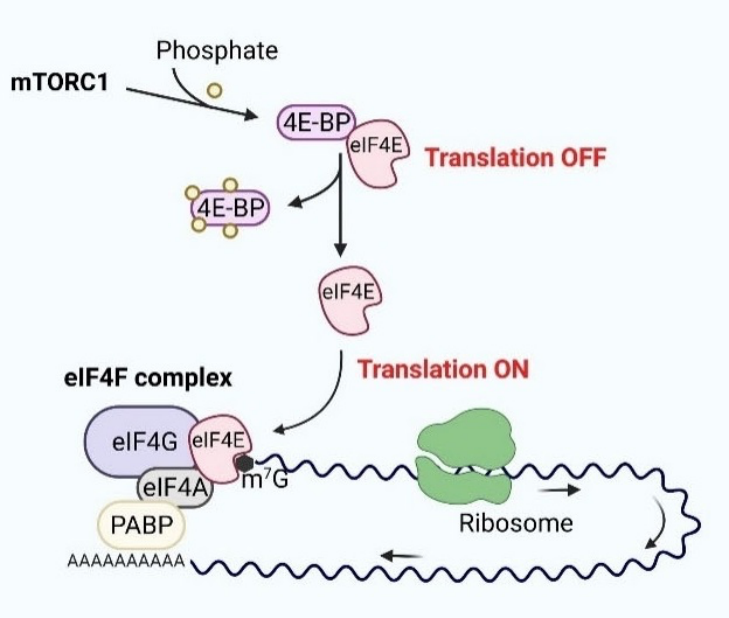


Figure 1. The cap-binding protein eIF4E is released by 4E-BPs as a result of its phosphorylation by mTORC1, allowing eIF4F complex formation and thus permitting translation of eIF4E-sensitive mRNAs. (Romagnoli et al., 2021)

The role of eIF4E is crucial for the translation of the so called "weak" mRNA. Indeed, under normal conditions, different mRNA species compete for the access to the eIF4F complex, which is limited by the availability of eIF4E (less than 0.1% of total proteins, the relative abundance of the different components can be read in the table below). "Strong" mRNAs have unstructured 5' untranslated regions (5'-UTRs), and typically code for housekeeping genes, such as β-actin; these genes are efficiently translated even when the eIF4E levels are low (de Benedetti & Graff, 2004). On the other hand, "weak" mRNAs

have structured 5'-UTRs and code for proto-oncogenes and growth regulatory proteins; weak mRNAs are more sensitive to eIF4E levels (de Benedetti & Graff, 2004).

Table 1. Abundance of translation factors related to ribosomes measured in HeLa cells, rabbit reticulocytes and wheat germ. Ribosome concentrations in eukaryotic cells are in the order of 1 mM.

Factor	Factor/ribosome ratio
eIF4A	3-5 copies
eIF2, 3 and 4B	0.2-1 copies
eIF2B, 4E and 5	0.01 – 0.1 copies

eIF4E is also involved in the balance of the detention/degradation of specific mRNAs (Parker & Sheth, 2007; Ross Buchan, 2014). This is confirmed by its nuclear localization in the P-bodies (processing bodies), which promote translational repression (Cohen et al., 2001; Lejbkowicz et al., 1992; Parker & Sheth, 2007; Ross Buchan, 2014; Rousseau et al., 1996).

Structural characteristics of eIF4E

The structure of eIF4E (Tomoo et al., 2002). consists of eight antiparallel beta-strand, three main alfa-helix (H3, H6, H8), six other small helix and loops. Three main parts of eIF4E can be distinguished (Figure 3):

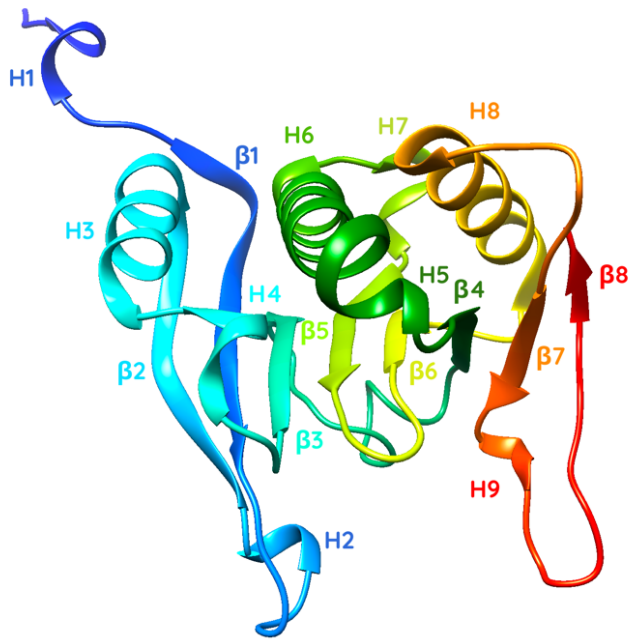


Figure 2. Crystal structure of eIF4E corresponding to the PDB code 1IPB. The protein is colored following the succession of residues, starting from blue N-terminal and ending with the red C-terminal.

- The ventral surface, where the cap-binding site is located. The interaction between eIF4E and the cap is primarily determined by the formation of cation-p bond stacking interactions between the 7-methylguanine, and two conserved tryptophanes located within the cap-binding pocket (Trp-56 and Trp-102 in human eIF4E-1). This interaction is also stabilized by a hydrogen bond network involving a conserved glutamate residue (Glu 103), the N1 and N2 atoms of the 7-methylguanine (Monzingo et al., 2007; Tomoo et al., 2002) and several positively charged residues (Arg

157, Lys 159 and Lys 162) located in the β5-β6 loop, which interact with the α- and β-phosphate oxygen atoms of the cap.

- The dorsal surface (Figure 3, blue), where the binding site of the 4E-BPs (eIF4E Binding Proteins) and eIF4G is located (Marcotrigiano et al., 1997). eIF4G and 4E-BPs contain the consensus binding sequence YXXXXLΦ (where X is any residue and Φ is any hydrophobic amino acid) that adopts a conserved α-helical fold (Mader et al., 1995; Marcotrigiano et al., 1997, 1999). The interactions between eIF4E and the C:4E-BM (Consensus: eIF4E – Binding Motif) are particularly conserved: the hydroxyl group of the tyrosine side chain forms a hydrogen bond with the carboxyl oxygen of the proline within the backbone of the His-Pro-Leu conserved motif in elF4E and establishes van der Waals interactions with a valine of eIF4E. Moreover, the conserved residues (Val 69, Trp 73 and Leu 131) located on the dorsal surface of eIF4E are in contact with the hydrophobic amino acids of C:4E-BM (LΦ) at the C-terminus (Romagnoli et al., 2021). This consensus sequence can be extended with other conserved residues that contribute to the binding of eIF4E hiding hydrophobic surface areas: YX(R/K)X2LΦX2(R/K/Q) (Grüner et al., 2016; Lama et al., 2013, 2019; Peter et al., 2015). Another residue of eIF4E that is important in this

binding site is Glu 132, that makes a salt bridge with the arginine/lysine in position 3.

- The lateral surface, in this region the binding site of the non-canonical binding motif of 4E-BPs is located. The affinity of 4E-BPs and eIF4G for eIF4E is strictly linked to this motif (Igreja et al., 2014), which interacts with a group of conserved amino acids on the lateral surface of eIF4E: Phe47, Ile63, Ile79.

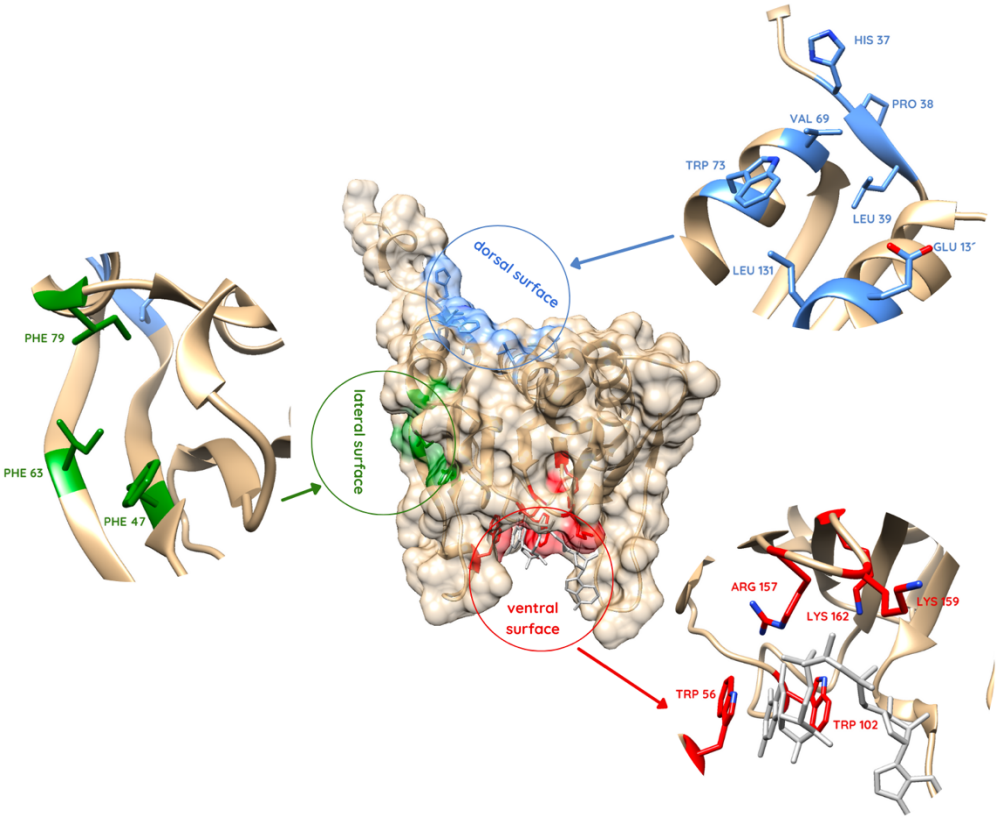


Figure 3. eIF4E (PDB: 1IPB), differentiation in regions, with details of residue that interact in the binding of cap and proteins. red: ventral surface, blue: dorsal surface, green: lateral surface.

Regulation

eIF4E Binding Proteins (4E-BPs)

4E-BP1, 4E-BP2 and 4E-BP3 are the main regulators of eIF4E activity, which are encoded by three eIF4E-BP genes, located on human chromosomes 8, 10 and 5, respectively. These proteins are functionally similar but have different localization. 4E-BP1 is expressed ubiquitously but mainly in adipose tissues and pancreas (Tsukiyama-Kohara et al., 2001). 4E-BP2 is crucial in the regulation of the nervous system, indeed knock-out mouse model displays autistic-likes behaviors (Banko et al., 2005; Gkogkas et al., 2013; Wiebe et al., 2019). 4E-BP3 is highest in skeletal muscle, heart, kidney, and pancreas, whereas there is a little expression in brain and thymus (Poulin et al., 1998).

4E-BPs share the eIF4E canonical biding motif (C:4E-BM with eIF4G and thus can inhibit the binding of eIF4E to eIF4G in a competitive manner (Modrak-Wojcik et al., 2013). The C:4E-BM folds in an α-helix structure interacting with the residues conserved in the dorsal surface. In addition to the consensus motif, two other sequences play a key role in the interaction with eIF4E: a non-canonical motif (NC:4E-BM) that binds the lateral surface of eIF4E and a linker region, which connects the two binding motifs (Grüner et al., 2016). These two sequences increase the affinity for eIF4E and confer a competitive advantage compared to eIF4G (Igreja et al., 2014; Kinkelin et al., 2012).

4E-BPs are regulated via phosphorylation: when these proteins are phosphorylated, they are not able to bind eIF4E, which is then free to bind eIF4G.

The phosphorylation of 4E-BPs is induced by the activation of the mTOR pathway, which is a part of PI3K/AKT/mTOR axis, in response to different stimuli: hormones (e.g., insulin), growth factors and amino acids (Gingras et al., 2001; Kimball, 2001). The serine/threonine kinase mTOR regulates most of the cellular processes, and in particular protein synthesis (Caron et al., 2010; Hay & Sonenberg, 2004; Laplante & Sabatini, 2012).

There are at least seven phosphorylation sites in 4E-BP1 and 4E-BP2, four of which (Thr-37, Thr-46, Ser-65 and Thr-70) are regulated by the mTOR pathway (Gingras et al., 2001; Heesom et al., 2001; X. Wang et al., 2005).

It has been reported that phosphorylation of T37 and T46 of 4E-BP2 induces folding of the protein into a four β-strand folded domain that sequesters the eIF4E binding motif (Bah et al., 2015).

Notably, dysregulation of the eIF4E/4E-BPs interaction is a common feature of numerous diseases, such as cancer, neuropsychiatric and neurodevelopmental disorders (Amorim et al., 2018; Pelletier et al., 2015).

An interesting member of the family of the 4E-BPs is CYFIP1, which is largely studied for its importance in the Fragile-X syndrome, because it forms a trimeric complex with FMRP and eIF4E, which sequesters eIF4E (DeRubeis et al., 2013; Napoli et al., 2008). Moreover, it is involved in other neurological disorders like schizophrenia (SCZ) and autism (Amorim et al., 2018).

A number of 4E-BPs do not harbor the C:4E-BM and NC:4E-BM motifs, but interact with eIF4E in different regions: examples are the proteins containing the Really Interesting New Gene (RING) domain (such as the promyelocytic leukemia protein and the arenaviral Z proteins) (Kentsis et al., 2001; Volpon et al., 2010), and a small group of viral proteins that interact through a non-conserved motif (German-Retana et al., 2008; Michon et al., 2006).

Ubiquitination

When eIF4E is ubiquitinated by Chip (heat shock cognate protein 70-interacting protein) an ubiquitin E3 ligase, but the all the protein involved remain unclear (Murata & Shimotohno, 2006). After the ubiquitination it can bind the cap, but noteIF4G, moreover it is possible to observe a reduction of phosphorylation (Murata & Shimotohno, 2006; Othumpangat et al., 2005). Interestingly, degradation of eIF4E is inhibited by 4E-BPs (Murata &

Shimotohno, 2006). Additionally, it has been demonstrated that the interaction of eIF4E with Hsp27 (heat shock protein 27; another regulator of eIF4E) decreases its ubiquitination/degradation (Andrieu et al., 2010).

eIF4E significance in disease: cancer and neurological diseases

eIF4E is a key factor in several disease: indeed, translation of weak mRNAs is highly dependent on eIF4E levels; overexpression of eIF4E (Avdulov et al., 2004; Lazaris-Karatzas et al., 1990) can induce malignant transformation in mammalian cells and is implied in the changing of the synaptic plasticity in the central nervous system (Buffington et al., 2014; Hsieh & Ruggero, 2010). Examples of neurological conditions linked to eIF4E are: Fragile X-syndrome (FXS), schizophrenia and, more in general, the pathologies belonging to the Autism Spectrum Disorders (ASD) (Amorim et al., 2018; Gkogkas et al., 2013).

Cancer

The overexpression or the dysregulation of eIF4E causes a rapid development of different types of cancer (Bhat et al., 2015; Siddiqui & Sonenberg, 2015) such as lung, bladder, colon, prostate, breast, head and neck cancer (Bhat et al.,

2015; Carroll & Borden, 2013; Chu et al., 2016; de Benedetti & Graff, 2004; Siddiqui & Sonenberg, 2015).“Weak” mRNAs code for molecules such cytokines and growth factors that have antiapoptotic and/or mitogenetic activity, and their altered expression is linked to malignant transformation (Borden & Culjkovic, 2009).

eIF4E not only promotes the export of mRNA encoding proteins required for cell cycle (Culjkovic et al., 2007), but also inhibits programmed cell death. Indeed, the activation of eIF4E and MYC promotes oncogenesis through the activation of anti-apoptotic proteins such as Bcl-xL (B-cell lymphoma-extralarge) (Li et al., 2003), reducing RAS-induced autophagy (Elgendy et al., 2011).

The up-regulated expression of eIF4E has a critical role in the export of mRNAs related to malignancy, including mRNAs encoding cyclin D1, ODC (ornithine decarboxylase) (Rousseau et al., 1996), and survivin. Indeed, elevated eIF4E levels lead to an increase of the nucleocytoplasmic transport of these mRNAs and enhance their recruitment on ribosomes (de Benedetti & Graff, 2004).

Neurological disease: focus on fragile X syndrome

eIF4E has a central role in neurodegenerative disease such as Alzheimer, pathological processes such as Wolcotte-Rallison syndrome and fragile X syndrome (Couturier et al., 2010; de Rubeis & Bagni, 2011; Donzé et al., 1995). Fragile X syndrome is caused by the absence of the fragile X mental retardation protein (FMRP), which leadsto a deficit in synapse maturation and the consequentmental retardation. FMRP is an RNA-binding protein with roles in mRNA localization, translation (Bagni & Greenough, 2005), and stability (Zalfa et al., 2007; Zhang et al., 2007). Claudia Bagni et. al in 2008 demonstrated that FMRP-mediated repression of translation requires an interaction with Cytoplasmic FMRP Interacting Protein CYFIP1 also known as Sra-1, which also binds the cap-binding factor eIF4E. According to their model, FMRP recruits CYFIP1 to mRNAs (Napoli et al., 2008).

The CYFIP1 region that interacts with eIF4E adopts the same "reverse L-shape" typical of the canonical eIF4E binding Motif (Marcotrigiano et al., 1999). The interaction between eIF4E and CYFIP1 has been intensively studied in vitro (Napoli et al., 2008) as well as in silico via molecular dynamics approaches, in order to develop new eIF4E inhibitors due to its unique binding mode (Di Marino, et al., 2015).

Pharmacological approach to eIF4E inhibition

During the years, several strategies to inhibit eIF4E have been developed.

One of which involves to the inactivation of the mTOR pathway. mTOR is the core component of two functionally different multi-proteins complexes: mTOR complex1 (mTORC1) and mTOR Complex2 (mTORC2). mTORC1 plays a key role in the regulation of mRNA translation and cell growth thanks to its capacity to phosphorylate 4E-BP and S6K1 (Mabuchi et al., 2015; Sabatini, 2006). When phosphorylated by the mTORC1 pathway, the affinity of 4E-BPs for eIF4E decreases, facilitating translational initiation (Siddiqui & Sonenberg, 2015). Rapamycin (or Sirolimus), the first inhibitor discovered against mTOR, is a macrolide produced by *Streptomyces hygroscopicus*. Rapamycin has shown antineoplastic effects in several cancer cell lines and mouse models, but of its low water solubility makes it difficult to use in clinical practice (Laplante & Sabatini, 2012; Lu et al., 2016). Over the years several analogues of rapamycin have been developed.

Another approach to block eIF4E is the use of ASOs (antisense oligonucleotides) and siRNAs(Wan et al., 2015). These molecules are under study in various cancer cell lines with promising effects (Duffy et al., 2016; Graff et al., 2007).

Peptides/peptidomimetics are an emerging class of molecules capable of inhibiting protein-protein interactions. A peptide composed by 49–68 of 4E-BP1 attached to an analogue of the gonadotropin-releasing hormone (GnRH) has been used to target tumoral cells over-expressing the GnRH receptor (Song et al., 2009). This peptide mimics the binding site of eIF4E to eIF4G/4E-BP1 and inhibits the formation of eIF4F.

It is possible to inhibit eIF4E by abolishing its binding with the m7G mRNA cap. Indeed, a number of m7G analogs, such as ribavirin have been developed and showed to have antiproliferative activity on several cancer model (Jin et al., 2019; Shi et al., 2015).

Indirect inhibitors of eIF4E are MNK1/2 phosphorylation inhibitors. As can be seen in previews paragraphs, MNK mediated phosphorylation of eIF4E in human. However, these molecules might have other off-target interactions (Kosciuczuk et al., 2016; Tschopp et al., 2000).

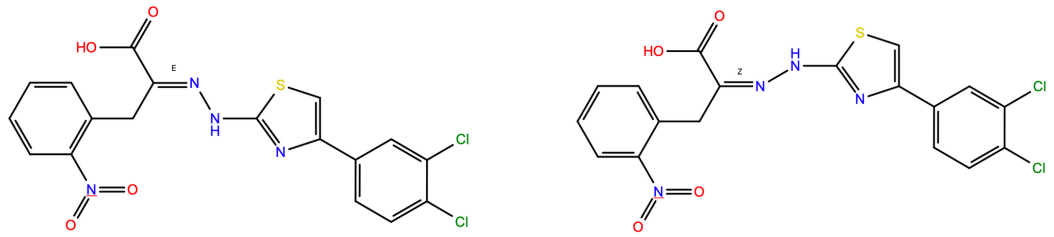
Finally, an important molecule that is the subject of this thesis is the 4EGI, a small molecule that have an allosteric inhibition mechanism through eIF4E and that will be extensively threat in following chapters (Papadopoulos et al., 2014).

1.3 4EGI: an important small molecule

In order to identify compounds that disrupt the interaction between eIF4E and the Y(X)4LF motif of eIF4G, the Wagner group has developed a high-throughput fluorescence polarization (FP)-binding assay.

The principle of the screening is to identify the compounds that displace a fluorescently labeled peptide from eIF4E by detecting the resulting decrease in FP. For the initial screening 16,000 compounds were used, from which 4EGI-1 (for eIF4E/eIF4G interaction inhibitor) was identified (Moerke et al., 2007).

4EGI-1 exists in two stable E and Z isomers (Figure 4) that bind both the cap-bound and cap-free forms of eIF4E with equilibrium dissociation constants between 10 and 20 μM as determined in binding experiments. The Z isomer exhibits slightly higher affinity than the E isomer (Papadopoulos et al., 2014).



4EGI-1[E] 4EGI-1[Z]

Figure 4. 4EGI isomers. Images created using Maestro by Schrödinger

4EGI-1 disrupts the eIF4E/eIF4G complex in vitro and in vivo, reduces viability, inhibits the proliferation of a broad spectrum of cancer cells, such as breast cancer (Chen et al., 2012; Moerke et al., 2007) and multiple myeloma (Descamps et al., 2012; Moerke et al., 2007; Tamburini et al., 2009). It also inhibits tumor growth in animal models of human cancers, such as acute myelogenic leukemia (Tamburini et al., 2009) and chronic lymphocytic leukemia (Willimott et al., 2013).

Action mechanism

The small-molecule inhibitor 4EGI-1 disrupts the eIF4E-eIF4G interaction, thus showing antitumor activity. Moreover, it seems to act as an allosteric inhibitor, because it binds eIF4E in a different site with respect to eIF4G. 4EGI-1 causes structural rearrangements, including the unfolding of a short 3_{10}-helix (S82-L85, Fig. 5), and elongation of helix-α1 by one tour (H78-S82, Fig. 5): these prevents the binding of eIF4G. In addition, mutation of residues in the helix to more helix-inducing amino acids (N77E, H78E) destabilizes the eIF4EGII peptide binding and doubles the IC50 for 4EGI-1 (Papadopoulos et al., 2014).

4EGI-1 was shown to inhibit the expression of regulatory proteins such as cyclin E, cyclin D1, c-myc and Bcl-2 without any apparent effect on the expression of housekeeping proteins (Chen et al., 2012).

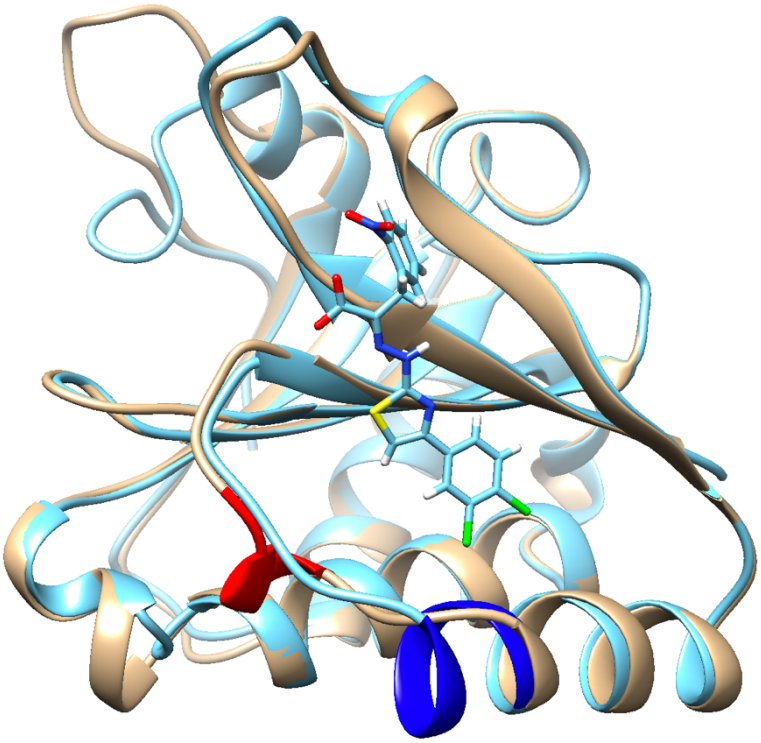


Figure 5. Structural rearrangement of eIF4E due to 4EGI. In light blue the structure of eIF4E bind to 4EGI-1 and MGP (PDB id 4TPW), in beige the eIF4E bound only to the 7mGTP, the two structures are superimposed. Highlighted with red is possible to see the 310 helix that unfold when 4EGI-1 binds eIF4E and with blue the tourn that acquire the helix α1 when there is the bind with 4EGI-1.

Moreover, 4EGI-1 enhances the binding of 4E- BP1. Furthermore, when 4E-BP1 is hyperphosphorylated and dissociated from eIF4E, 4EGI-1 replaces 4E-BP1, hindering the eIF4G/eIF4E interaction. Thus, 4EGI-1 reinforces the translation inhibition function of 4E-BP1, providing an adjunctive tumor-suppressive role.

Binding mode of 4EGI-1

Papadopoulos et al. in 2014 obtained a crystal structure of eIF4E in complex with 4EGI-1, which is to date the only crystal structure available of the complex (Papadopoulos et al., 2014). They succeeded only to co-crystalize the 4EGI-1[E] isomer, despite the Z isomer has been proved to bind strongly the eIF4E. 4EGI-1 binds to a hydrophobic pocket of eIF4E between β-sheet2 (L60-T68) and α-helix1 (E69- N77), causing localized conformational changes mainly in the H78- L85 region (Papadopoulos et al., 2014).

The o-nitro-phenyl group sticks into a pocket formed by the loop between strands-β1 and -β2 (residues 46–64 Fig.6 in blue), where the backbone forms the bottom of the pocket and the side chains of K49, K54, and R61 constitute the basic rim (Fig.6). The K49 side chain e-NH3 forms a salt bridge with the carboxylate of 4EGI-1[E], and the R61 side chain guanidino group engages in a cation-π interaction with the o-nitrophenyl group. The thiazolidine ring stacks flat onto the F47 side chain.

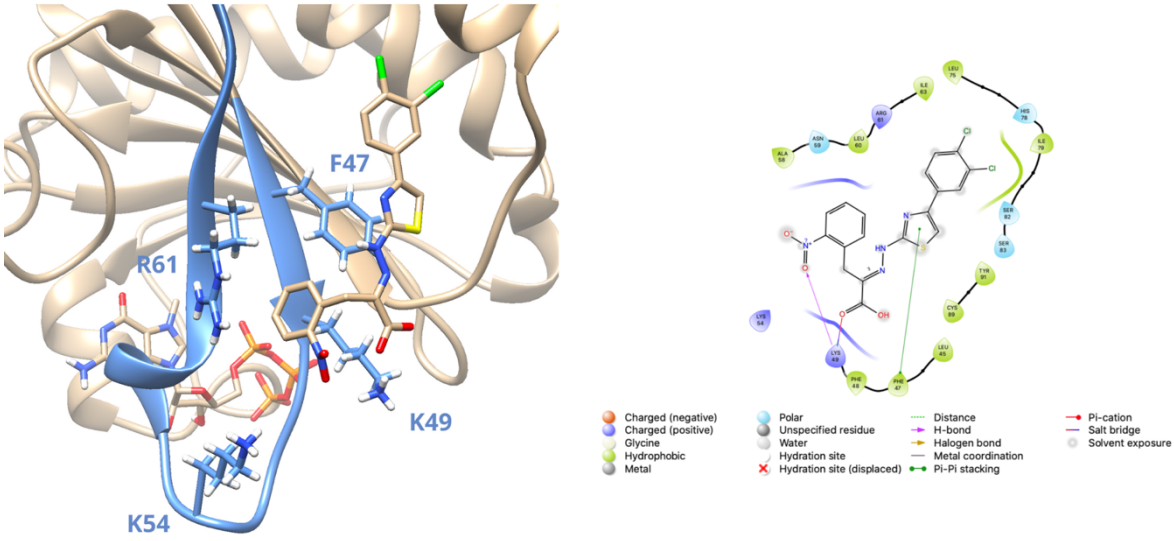


Figure 6 (A) binding site of 4EGI-1 in eIF4E represented with a 3D view referring to the PDB code 4TPW (images obtained with Chimera) In blue the residue of interest of eIF4E. (B) 2D representation of the binding site (obtained with MAESTRO by Scrödinger).

Recently a series of 4EGI-1 analogs have been developed in order to improve its affinity and effectiveness (Fischer et al., 2021); during these experiments a dimeric version of 4EGI-1 has been synthesised. One of the intermediate of this process is i4EG-BiP, a biphenylic version of 4EGI-1, which binds eIF4E with a IC50 value of 68 ± 2 μM and shows a different binding mode that has never been observed so far.

i4EG-BiP engages eIF4E using hydrophobic interactions with residues L45, L75, I79, L93 and L134. Edge-to-face p-p interactions are formed between the biphenyl moiety and Y76, Y91 and W130. Hydrogen bonds are formed between the nitro group of i4EG-BiP and the amide side chain of Q80 as well

as between the hydrazone NH of i4EG-BiP and the backbone carbonyl oxygen of Q80 (Figure 7 A, B, C).

The effect of the binding of i4EG-BiP with the eIF4E are the same of 4EGI. Indeed, the α1-helix is extended by one turn (amino acids Q80, L81, S82 and S83), and the 3_{10} helix between residues S82 and L85 melts. Moreover, the amount of eIF4G bound to eIF4E decreases with increasing concentrations ofi4EG-BiP, whereas the amount of 4E-BP1 increases (Fischer et al., 2021).

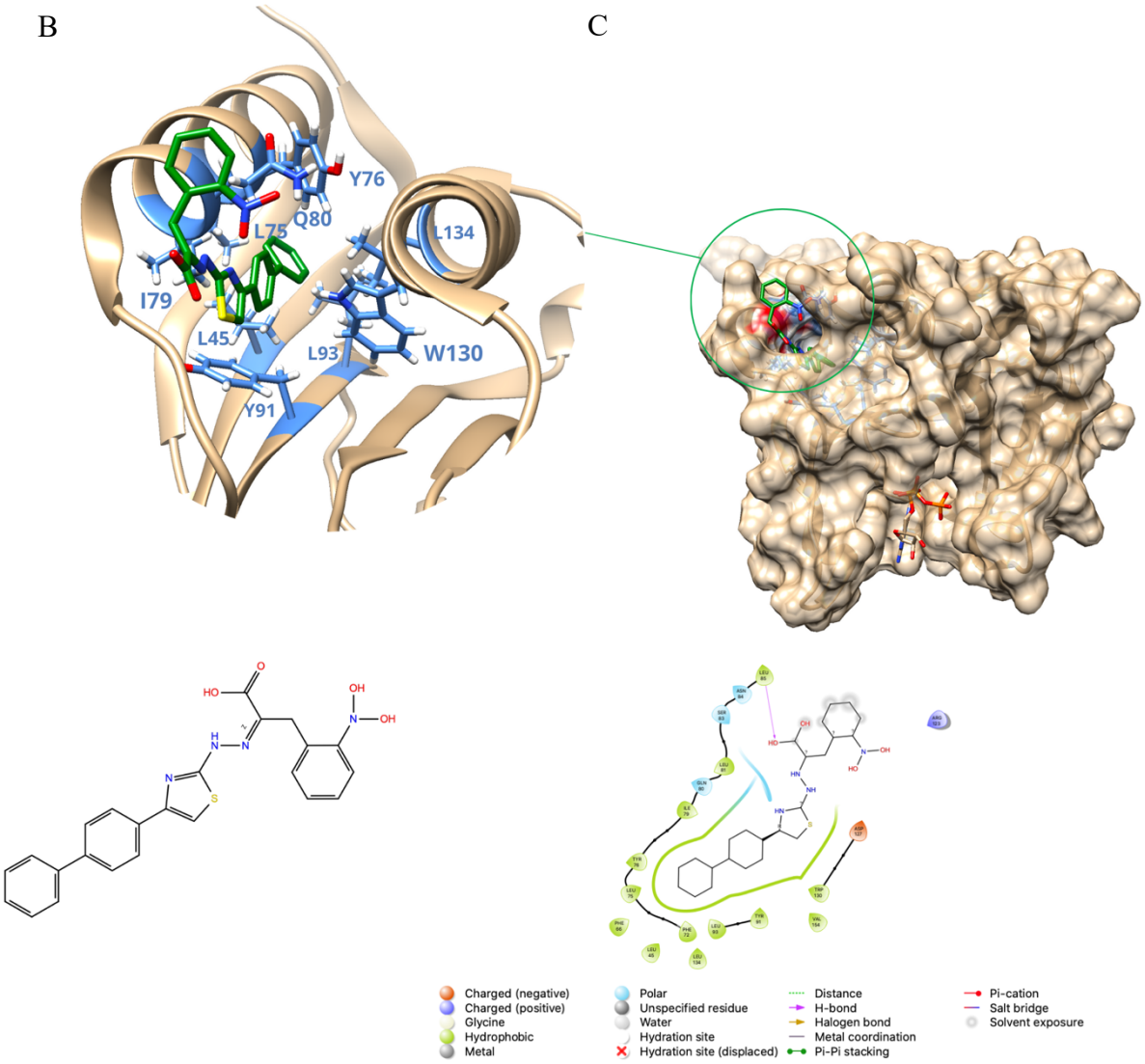


Figure 7. Binding mode of i4EG-BiP with eIF4E. From the crystal structure with PDB id 7MEU. (A) 3D detail of the binding site. (B) 2D i4EG-BiP 2D molecule. (C) 2D representation of the binding mode. Images created using Chimera and Maestro (by Schrödinger).

1.4 How to understand the interaction between eIF4E and 4EGI: the importance of molecular dynamic

Concerns about binding mode

The two binding modes characterized by the two crystal structures described above raises some doubts due to the presence of two completely different biding modes for two small molecules with a very similar structure.

Takrouri et al. in 2014 made an experiment where they measured the fluorescence intensity quenching of the eIF4E in the presence or absence of 4EGI-1 (Murata & Shimotohno, 2006).

The binding of eIF4E to 4EGI produced a change in the intrinsic tryptophan fluorescence intensity quenching of eIF4E. Tryptophan fluorescence quenching is a type of fluorescence spectroscopy used for binding assays. The assay relies on the ability to quench the intrinsic fluorescence of tryptophan residues within a protein that results from changes in the local environment polarity experienced by the tryptophan(s) upon the addition of a binding partner or ligand (Yammine et al., 2019).

The change does not seem to be explainable by the biding mode of 4EGI and eIF4E, but it might be explained by the interaction with i4EG-BiP, because of the presence of one tryptophan (W130) and two tyrosine (Y76 and Y91) in the binding pocket, that contribute to the absorbance at 280 nm.

In latest years, Molecular Dynamics and other computational approach have become increasingly crucial in the characterization of the binding between proteins and their ligands. Moreover, research on PubMed using the keywords ‘drugs’ and ‘molecular dynamics’ returns more than 800 papers published in the last ten years, and the annual rate of publication increases every year. MD is becoming a key step in the investigation of therapeutically relevant targets. In a recent review, several examples are provided of the possible use of molecular dynamics in different step of drug design studies (Mortier et al., 2015).

The main difference between information obtained with X-ray crystallography and MD analysis is that X-ray crystallography provides a static snapshot of a fully functional state; MD, on the other hand, is an *in silico* method that uses structural data gained experimentally to extrapolate the possible conformations of molecular systems and the different paths between each of them.

History of Molecular Dynamics

The first molecular dynamics simulation was performed by Alter and Wrainwright (Alder & Wainwright, 1957), in the ‘50s. They studied fluids with techniques that approximate the atoms as disk and rigid spheres. In 1964

Rahman and Verlet (Rahman, 1964), using more realistic potential functions, like Lennard-Jones interaction, described the atomic interaction of liquid argon.

The simulation of biomolecules however dates back to the end of the ‘70s, and the excellent results obtained in these years provide the impetus for a great diffusion of MD approach in biochemical studies.

The importance of MD is highlighted by the recent Nobel Prize in Chemistry 2013, which was awarded to Martin Karplus, Michael Levitt and Arieh Warshel "for the development of multiscale models for complex chemical systems." Moreover, more and more research groups are investing money and time to develop computers suitable to perform molecular dynamics simulation and software that are able to manage complex systems. David Elliot Shaw founded in 1988 the D. E. Shaw & Co, a multinational investment management firm, and in 2001 he built Anton, a massively parallel supercomputer located in New York, and a special-purpose system for molecular dynamics simulations of proteins and other biological macromolecules.

Nowadays molecular dynamics is used for structural and dynamics studies of gas, liquids, nucleic acids and proteins. In addition, it can be used to refine

structures obtained by homology modeling techniques, and to validate models obtained experimentally through X-ray, NMR and Cryo-EM.

What is molecular dynamics?

MD simulations are a numerical solution of Newton's equations of motions:

$$F_i = m_i a_i = - [\delta V(r^N) / \delta r_i]$$

m_i= mass of atom i

a_i = acceleration of atom i

f_i = force acting on atom i given by the partial, spatial derivative of the potential energy function V, dependent on the positions $r = r_1, r_2, r_3 \ldots r_N$ of the N particles of the system.

MD simulations generate a time series of configurations of a system, which allows studying dynamic processes. The time series can also be viewed as a statistical-mechanical ensemble of configurations. By averaging over this ensemble, thermodynamic properties of a system can be accessed.

V can be calculated using the following equation:

$$1/2\ mv^2 = 3/2\ NK_bT$$

K_b = Boltzmann constant, is the proportionality factor that relates the average relative kinetic energy of particles in a gas with the thermodynamic temperature of the gas. Value = 1,308 x 10^{-23} J * K^{-1}

N = atoms numbers

T = temperature (K)

With this formula we can understand how velocity is linked to absolute temperature.

The new atoms positions and velocities are calculated through the resolution of the motion equation, using the previews velocities, positions, and accelerations.

$$r(t + \Delta t) = r_i(t) + v_i(t)\Delta t + 1/2a_i(t)t^2$$

$$v_i(t+\Delta t) = v_i(t) + a_i(t)\Delta t$$

When a simulation is started, an initial velocity is given to the system. From this velocity the acceleration can be derived. Another important thing to define is the integration step (Δt), which it must be in the order of femtoseconds, having a small Δt (0,5 - 2fs) is fundamental to allow all the system forces to remain virtually constant at every integration step.

Atoms' interactions

To have a correct simulation of the system it is fundamental to precisely describe all atoms' interactions. Atom interactions can be divided into binding interactions and non-binding interactions.

Binding interactions

Table 2. Binding interactions

Bond type	Elements involved	Energy contribution	Legend
Bond-stretching	2 bodies	$\sum_{bonds} \frac{k_i}{2} (l_i - l_{i,0})^2$	k_i = spring constant of bond l = bond length
Bond-angle	3 bodies	$\sum_{angles} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2$	k_i = spring constant of angle θ = angle between atoms
Dihedral-angle	4 bodies	$\sum_{dihedrals} \frac{V_N}{2} (1 + \cos(n\omega - \gamma))$	V_N = spring constant of dihedrals ω = Angle between planes
Improper dihedral	4 bodies	$\sum_{improper\ dihedrals} \frac{k_i}{2} (\varepsilon_i - \varepsilon_{i,0})^2$	k_i = spring constant of improper dihedrals ε = angle between planes

- Bond stretching: the potential energy is minimum at the average distance between the two atoms and maximum at the extremes (Figure 8, A).
- Bond bending energy: the potential energy is minimum at the average distance between the two atoms and maximum at the extremes (Figure 8, B).
- Torsional o Dihedral Term: proper dihedral angles are defined according to the IUPAC/IUB convention, with zero corresponding to

the cis configuration. The energy to deform dihedral is a sinusoid function (Figure 8, C).

- Improper torsional term: is the energy required to deform a planar group of atoms from its equilibrium angle, usually equal to zero. Improper dihedrals are meant to keep planar groups (e.g. aromatic rings) planar, or to prevent molecules from flipping over to their mirror images (Figure 8, D).

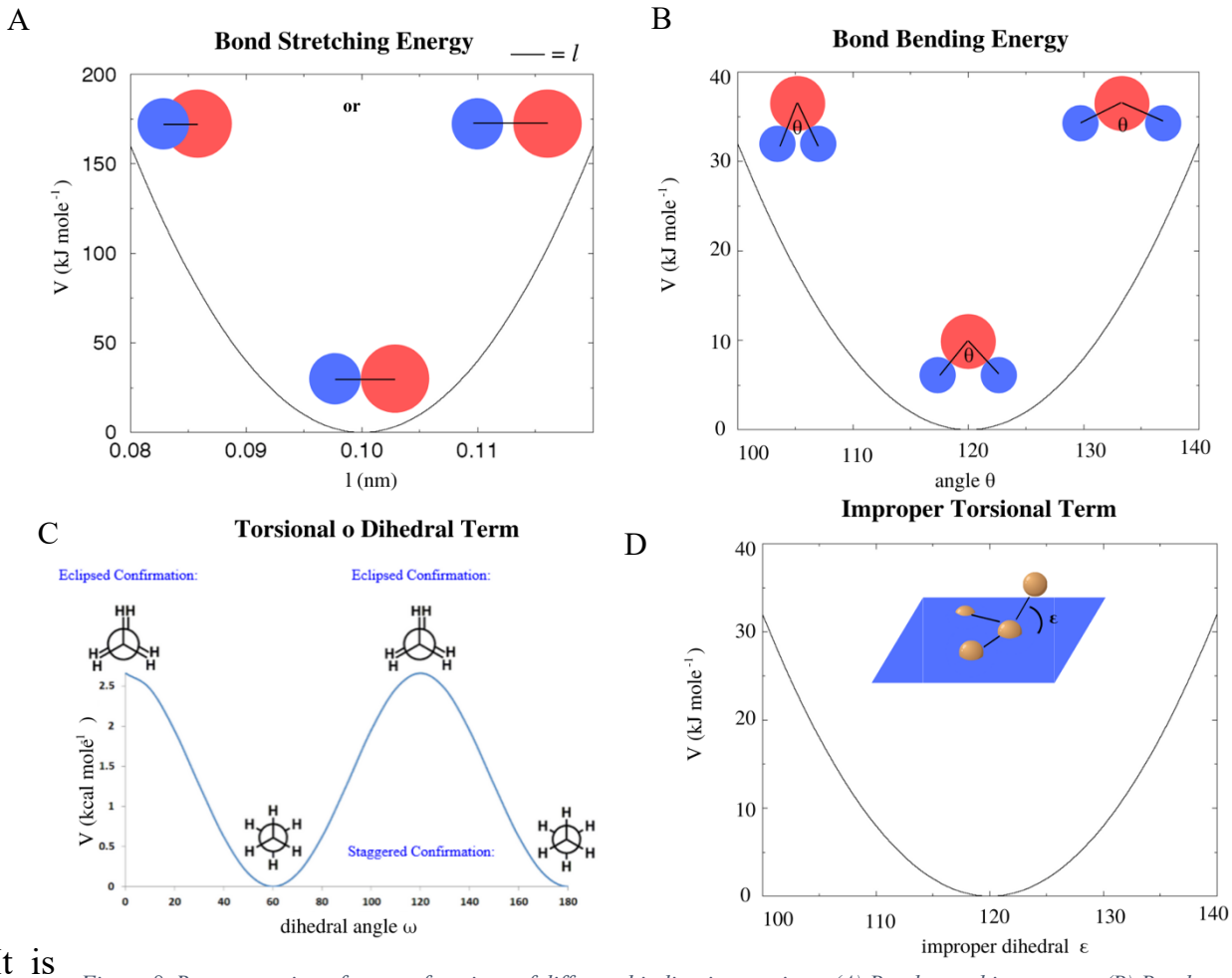


Figure 8. Representation of energy functions of different binding interactions. (A) Bond stretching energy, (B) Bond Bending energy, (C) Improper torsional term. (D) Torsional or Dihedral term

It is important to focus on the fact that these energies are described by the Hookeian

potential: they are described by a parabola, their behavior is the same of a spring.

Non-binding interaction

Table 3. Non-binding interactions

Bond type	Elements involved	Energy contribution	Legend
Van der Waals interactions (Lennard-Jones potential)	2 bodies	$V_{LJ}(r) = 4\varepsilon\,[\,(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}]$	σ = distance that makes null the potential ε = energy at the minimum of the function r = distance between two interacting particles.
Electrostatic interaction	2 bodies	$v = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$	$q_i q_j$ = electrical charges r_{ij} = distance between electrical charges $\frac{1}{4\pi\varepsilon_0}$ = Coulomb constant ε_0 = electrical permittivity of free space. Value = 8,854 * 10^{-12} F/m

- Van der Waals interactions (Lennard-Jones potential): van der Waals interaction forces between two atoms are represented by Lennard-Jones potential; this means that two interacting particles repel each other at very close distance, attract each other at moderate distance, and do not interact at infinite distance.

- Electrostatic interaction between two atoms is represented by Coulomb Law, which states that the magnitude of the electrostatic force of interaction between two-point electrical charges (q_1, q_2) is directly proportional to the scalar multiplication of the magnitudes of electrical charge and inversely proportional to the square of the distance (r) between them. The Coulomb force acts along the straight line joining them. If the two electrical charges have the same sign, the electrostatic force between them is repulsive; if they have different signs, the force between them is attractive.

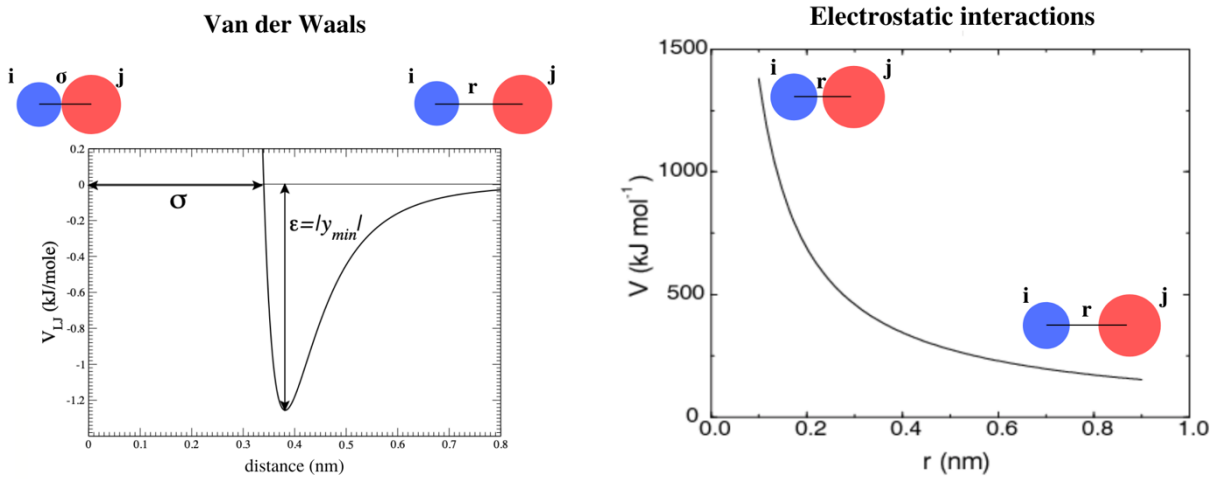


Figure 9. Graphical representation of non-binding energy functions.

Systems and boundary conditions

In MD simulations it is possible to have different types of system depending on their thermodynamics characteristics:

- Microcanonical ensemble (NVE), characterized by constant number of particles (N), volume (V) and energy.
- Canonical ensemble (NVT), particles (N), volume (V) and temperature (T) are constant. But the energy can change.
- NPT ensemble, particles (N), pressure (P) and temperature (T) are constant, but volume can change.

A key point that has to be considered in a simulation is the problem of boundary conditions. Usually, a system is simulated in vacuum, but this leads to some deformation of the system due to the fact that molecules tend to minimize their surface. In order to avoid this, the boundary with vacuum is blocked (extended wall region), with harmonic potential while the inner part of the system respects the normal dynamics (Berkowitz & McCammon, 1982). In practice, periodic conditions are used, where molecules are in a cubic box that is surrounded by other 26 identical images of itself obtained by a translation of the same length of the side of the cube in the three directions x, y and z. The interactions taken into consideration are only the ones of the central box with their nearest neighbors. Moreover, to avoid anisotropy of the cubic box a spherical cutoff is applied, which must be shorter than the shortest side of the box, overwise one

atom interacts simultaneously with another atom and its periodic image (Figure10).

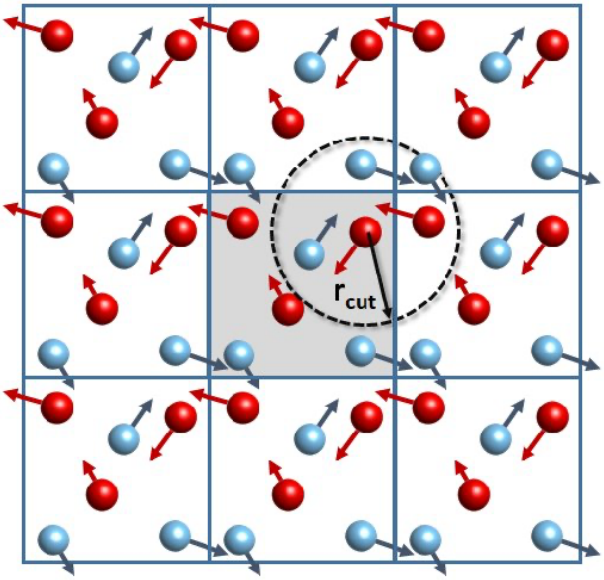


Figure 10. Representation of periodic boundary conditions and cutoff radius.

Resolution of simulations

Simulation systems can be described in different ways. Molecules are described as:

- nuclei and electrons in Quantum-Mechanical (QM) simulations;
- atoms in classical MD simulations;
- coarse-grained (CG) beads, multiple atoms are described by a single particle.

The different methods to describe molecules are listed above in an order that see the computational demand decreases and the accessible time and length scale increases.

All his resolution levels can be used in different situations and are relevant for drug design. For example, QM/MM simulations allow the study of enzymatic reactions, reaction pathways, energy barriers and proton and electron reshuffling in larger systems, where the active site is treated quantum-mechanically while the other parts of the system can be simulated at the atomic level (Bakowies & Thiel, 1996; Field et al., 1990; Gao et al., 2003; Gogonea et al., 2001; Warshel & Levitt, 1976). QM can also be used to parametrize ligands that do not have torsionals in perimetry libraries.

Coarse-graining of atomic models is a popular technique to access longer time scales in MD simulations. By embedding multiple atoms in a single CG bead, the number of particle–particle interactions to be calculated can be reduced considerably. It must be taken into account that this strategy leads to a loss of information, and therefore is frequently used in drug design.

Key points of molecular dynamics

In the following points it is possible to see briefly the key points of a MD simulation. Explanations of algorithm used, and detailed workflow will be explained in the "Material and Methods" chapter.

1. determination of initial conditions and coordinates

   What is needed to start a simulation is a coordinate file, usually developed experimentally via X-Ray, Cryo-EM or NMR, or computed by homology modeling or docking in the case of a ligand. From this file a force field is chosen, and a topology file is built. It is crucial to build a correct topology file, because it contains all the components of the system: proteins, ligands, physical properties of atoms and how they interact to each other. Classical MD strongly depends on the initial positions, so this step is crucial.

2. minimization of the system

   Once the box is built it is filled with water molecules and ions. By using algorithms that bring the system to its minimum energy (steepest descent and conjugate gradient). In this phase water molecules creep into the cavities of the system and the side chains change their position to avoid clashes and to interact with the solvent.

3. equilibration of the system

It is not possible to start a simulation of a system directly at 300K, because the amount of energy given to the system would be too much and the system would collapse. The protein is first blocked by position restraint and the water is free to move and a solvation shell is created. The initial temperature is usually 50K and it raises gradually to reach 300K. Every time the temperature is risen, it is considered constant for some time to give time to the system to rearrange. Then it is increased again by a small ΔT. These short simulations are made first in NVT than in NPTs, so the system gradually changes its temperature and pressure and it can increase its energy avoiding the collapse.

4. production and analysis of trajectories

   After the equilibration it is possible to run the production at the pressure and temperature wished. After the trajectories are computed they are connected, the solvent is eliminated, periodic boundary condictiones are resolved and it is possible to make several analyses on it.

1.5 Previous simulations performed on eIF4E with ligands

Investigation of the literature

As can be seen from the Table 4, none of the simulations present in literature of eIF4E with peptides were performed with 4EGI.

Table 4. List of simulations performed of eIF4E with ligands.

Article	Ligand	Leght of the simulation
Conformational ordering of intrinsically disordered peptides for targeting translation initiation. (Brown et al., 2021)	Peptides of IDRs region of eIF4G	1 µs for each systems
Emetine suppresses SARS-CoV-2 replication by inhibiting interaction of viral mRNA with eIF4E.(Kumar et al., 2021)	Emetine + eIF4E, and eIF4E + m^7G.	200 ns
Deciphering the mechanistic effects of eIF4E phosphorylation on mRNA-cap recognition. **(Lama & Verma, 2020)**	Cap analogs	3 triplet of 100 ns, total 3,6 µs
Water-bridge mediates recognition of mRNA cap in eIF4E. (Lama et al., 2017)	eIF4E bound to H_2O	-
A unique binding mode of the eukaryotic translation initiation factor 4E for guiding the design of novel peptide inhibitors. (di Marino, D'Annessa, et al., 2015)	CYFIP1−NCKAP1 heterodimer	135 ns
MD and docking studies reveal that the functional switch of CYFIP1 is mediated by a butterfly-like motion. (di Marino, Chillemi, et al., 2015)	8 simulations with different ligands (not 4EGI)	50-100 ns and 1,6 µs
A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. (Paku et al., 2012)	eif4e/4EBP-2 N and C terminal	3 ns C-terminal and 10ns N-terminal
Importance of C-terminal flexible region of 4E-binding protein in binding with eukaryotic initiation factor 4E.(Mizuno et al., 2008)	Peptides of 4B-BP2	3 ns
Effect of N-terminal region of eIF4E and Ser65-phosphorylation of 4E-BP1 on interaction between eIF4E and 4E-BP1 fragment peptide. (Tomoo et al., 2006)	Peptide del 4EBP-1, Pro47-Pro66, with pSer65	50 ns
Structural basis for mRNA cap-binding regulation of eukaryotic initiation factor 4E by 4E-binding protein, studied by spectroscopic, X-ray crystal structural, and molecular dynamics simulation methods.(Tomoo et al., 2005)	eIF4E - 4E-BP1 peptide m^7Gp-eIF4E-4E-BP1 complex.	5 ns
Structural features of human initiation factor 4E, studied by X-ray crystal analyses and molecular dynamics simulations. (Tomoo et al., 2003)	eiF4E with pSer209 with and without m^7G.	800 ps, 2 ns.

Simulation performed on Crystal structure 4TPW

In 2016 a research group analyzed the conformations that the eIF4E assumes before and after the binding of 4EGI-1, however they could not observe how the ligand behaves.

It should be noted that: “Due to the lack of optimized force fields for the N7-methylated nucleotide and for 4EGI-1, ligands were removed from the PDB structure and simulations performed using the peptide chains only”.

There was no indication of significant structural changes on a μs-ms time scales because the characteristic line broadening due to conformational dynamics occurring on these time scales was not detected. So, a first simulation of the “free-like” conformation was performed, which included three 100 ns trajectories at 5°C (one starting at 0 ns, another at 50 and the last at 100 ns), with randomly distributed atomic velocities. The second simulation was set like the first one but with the “bound-like” conformation.

Comparison of the simulations showed that 4EGI-1 increases the flexibility of eIF4E throughout its sequence, and particularly in regions adjacent to the 4EGI-1 binding site: the loop between β1 and β2, part of helix α1 and the loop between the extended α1 and β3 (Salvi et al., 2016).

Despite these simulations allowed to understand the differences between the bound-like and free-like form of the protein, the lack of ligands makes the binding mode still to be investigated.

1.6 Aim of the thesis

After literature research and an analysis of the structures available, the investigation of the binding mode of 4EGI and eIF4E was considered interesting, in order to allow the development of 4EGI analogs, thanks to the detailed knowledge of the binding mode.

It is fundamental to understand the behavior of 4EGI-1 in the crystalized conformation (PDB 4TPW) and to investigate the possibility of an analogous binding mode in the position of i4EG-BiP in the crystal structure with eIF4E (7MEU).

Having regard to the above consideration (paragraph: "4EGI: an important small molecule") on the doubts that lie on the binding mode presented on the crystal 4TPW, the aim of this thesis is to confirm or confute the stability of 4EGI-eIF4E binding mode characterized by the crystal structure 4TPW. Indeed, as it has been previously stated, two different binding modes of two such chemically similar molecule in the same small protein is unusual. With

the simulations several analyses will be performed to understand deeply the binding mode and the evolution in time of the binding (2μs).

2.1 MATERIALS AND METHODS

2.1 Molecular dynamics workflow and system preparation

MD simulation workflow

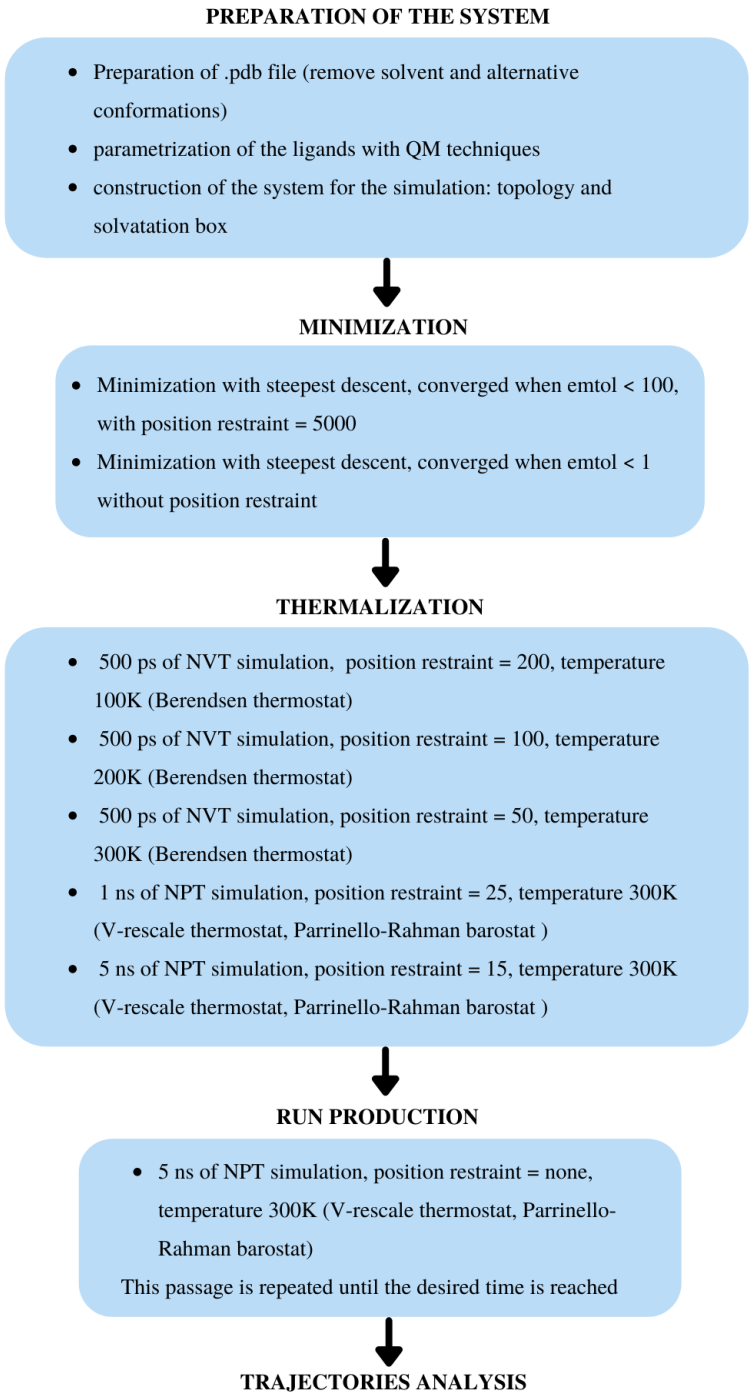


Figure 11. Schematic representation of MD simulation workflow.

Preparation of the molecule

The PDB file used for the simulations was downloaded from Protein Data Bank (PDB, www.rcsb.org) and it correspond to the code 4TPW.

The protein was prepared using UCSF Chimera (Pettersen et al., 2004): water molecules were removed; the integrity of the structure was verified, and the alternative conformations of residues were analyzed.

4EGI ligand parametrization through quantomechanics

Parametrization of ligands is a crucial part of the study. Indeed, if the parameters are not optimized for the ligand, some artefacts may arise.

Antechamber

Antechamber is a set of auxiliary programs for molecular mechanic (MM) studies. It is an accessory module of AmberTools package, and it can generate input automatically for most organic molecules in a database.

Antechamber is designed to be used with the "general AMBER force field (GAFF)"(J. Wang et al., 2004). This force field was specifically designed to cover most pharmaceutical molecules (can handle a higher range of atom types) and is compatible with the traditional AMBER force fields. The current

implementation of the GAFF force field consists of basic atom types and special atom types. The charge methods used can be HF/6-31G* RESP (Bayly et al., 1993) or AM1-BCC (Jakalian et al., 2000).

The coordinates of 4EGI were obtained by the coordinate file of the complex eIF4E-4EGI crystal 4TPW (in the PDB file of the complex it is noted with the name “33R”).

. Then the molecule was converted in mol2.

```
antechamber –i 33R.pdb –fi pdb –o 33R.mol2 –fo mol2 –c bcc
```

-c bcc option tells antechamber to use the AM1-BCC charge model to calculate the atomic point charges.

33R.mol2 file contains the definition of the 33R residue including all the charges and atom types that will be loaded into LEaP to when creating prmtop and prepi files.

It is possible that some combination of bonds or dihedrals have not been parameterized. Missing parameters must be specified before the creation of the topology.

It is possible to use the utility parmchk to test if all the parameters required are available.

```
parmchk -i sustiva.mol2 -f mol2 -o 33R.frcmod
```

This command will produce a file called 33R.frcmod. This is a parameter file that can be loaded into LEaP in order to add missing parameters. In this case, all bonds and angles were parametrized; nonetheless, dihedrals' parameters are ones of the most problematic, so it is important to verify them.

Evaluation of dihedrals and manual parametrization

Verification consists of two different passages: a quantomechanics analysis of torsional is performed to obtain the potential energy profile linked to the phases, and its comparison to the profile obtained by the force field.

If the two energy plots are comparable, the GAFF generated parameters are use, unless they are replaced with the computed ones.

Quantum mechanics is fundamental to understand at atomistic level how molecules interact to each other. Quantum dynamics deals with the motions, and energy and momentum exchanges of systems whose behavior is governed by the laws of quantum mechanics. It is not used for all the calculation because in spite of its accuracy, these methods are unfeasible due to the large of atoms to consider and the sheer complexity of the calculation.

Force Fields such GAFF considers a small amount of compound which parameters has been obtained with QM calculation and they try to extrapolate general rules for all the possible cases: this imply that parametrization has to be verified, this is especially true in torsion angle parameterization of ligands with π electron conjugated systems, where quantum (ligand-specific) effects are particularly important in defining the correct molecular conformation.

Ligand parametrization

Firstly, was performed a geometry optimization using the QM Software Gaussian (M. J. Frisch et al., 2009), using increasing basis set complexity (2-31G and 6-31G, respectively), where a Hartree-Fock calculation was requested, followed by a Møller-Plesset correlation energy correction truncated at the second order (MP2) (Gordon et al., 1982; Møller & Plesset, 1934; Petersson et al., 1988).

Once obtained the optimized structure, a population analysis with Hartree-Fock was performed to produce charges considering the electrostatic potential at points following the Merz-Singh-Kollman scheme. The output file was post-processed by the Restrained Electrostatic Potential (RESP) method to obtain the partial charges per atom (Bayly et al., 1993).

Finally, the new *ad hoc* topology was produced by taking as a template the parameters from GAFF and modifying the dihedral angle. The *ad hoc* parameters were obtained by means of a "Scan" calculation using the settings of the second geometry optimization process. Such procedure involves several energy minimization calculations in which the dihedral angle of interest is restrained in a specific conformation and relaxing the rest of the molecule.

Once obtained the energy profile for the dihedral angle under investigation through QM calculations, the parameters that reproduced the QM potential energy function are resolved using the Amber dihedral formula:

$$E_{dih} = k\,(1 + \cos\,(n\varphi - \psi))$$

k: spring constant for the dihedral angle

n: period

φ: the angle

ψ: phase.

The function that best reproduced the QM behavior was computed with an in-house genetic algorithm and it is a combination of two torsional angles that has the equation:

$$E_{dih} = 2.4(1 + \cos(2\varphi - \pi)) + 1.0(1 + \cos(4\varphi))$$

Ligand conformational analysis

The simulations with the all-atom FFs were performed using well- tempered metadynamics (MetaD) (Barducci et al., 2008), using the torsional angle of interest and compared the obtained energy profiles with that from QM calculations. The MetaD calculations were run using Amber14 patched with Plumed-2.3.3 (Tribello et al., 2014), setting a low value as height of the metadynamics Gaussian functions (i.e., 0.1kJ/mol for CGenFF and the ad hoc topology and 0.05 kJ/mol for GAFF and GAFF2), a sigma of 0.05 radians, a biasfactor of 15, and a deposition rate of 1,000 steps. All simulations were carried out in vacuum and converged within 100 ns.

MGP ligand parametrization

Due to the similarity of MGP to nucleotides, the parameters were extrapolated from the DNA and RNA Amber Force Fields (OL15, OL3). The charges were verified through quantomechanics as described above.

Topology and simulation box

The topology is the description of the molecules in terms of bonds and connections. The software computes this information through a force field designed.

Here, the topology was built using Amber18, force field used were: ff14SB (protein force field), tip3p (water model with associated atomic ions force field) and gaff2 (general force field for organic molecules like ligands).

To build the topology the software used was Amber18 (Case et al., 2005), more precisely the LEaP program. After the loading of libraries and ligand parameters, the protein was solvated in a cubic box, with minimum distance 12 at each direction with TIP3P water and neutralized with Na^+ ions. The topology was built using the ff14SB FFs and finally converted in the GROMACS topology file.

Before starting the minimization, it is important to add the position restraint. Every molecule must have its position restraint file and restraint will be modified at each step of the system preparation process.

2.2 Minimization

It is important to minimize the system to start the simulation at the lowest energy possible. The system is minimized in vacuum with no coupling of

temperature or pressure. This passage is crucial to permit the correct adjustment of solvent around the system, clashes are solved and correct interaction between molecule and water are established. If this passage is not made correctly, the system might collapse when the temperature and the pressure are set.

In this phase steepest descent algorithm was used, it is not the most efficient algorithm for searching, but it is robust and do not require a high amount of calculation resources. In this algorithm is used to calculate the new position of each atom through an energy gradient. The vector **r** is the vector of all *3N* coordinates. Initially a maximum displacement h_0 (*e.g.,* 0.01 nm) must be given.

First the forces FF and potential energy are calculated. New positions are calculated by:

$$r_{n+1} = r_n + \frac{F_n}{\max(|F_n|)} h_n$$

where h_n is the maximum displacement and F_n is the force, or the negative gradient of the potential *V*. The notation $\max(|F_n|)$ means the largest scalar force on any atom. The forces and energy are again computed for the new positions

If ($V_{n+1}<V_n$) the new positions are accepted and $h_{n+1}=1.2h_n$

If ($V_{n+1}\geq V_n$) the new positions are rejected and $h_n=0.2\ h_n$

The algorithm stops when either a user-specified number of force evaluations has been performed (*e.g.*, 100), or when the maximum of the absolute values of the force (gradient) components is smaller than a specified value ϵ.

In these simulations the minimization phase consists of two steps, with different tolerance for the convergence (emtol, pN/mol).

First one, consists of a minimization with emtol = 100, and with position restraint = 5000 (em1.mdp). The system is blocked and only the water can move around it. Second one, consists of a minimization with emtol = 1, and without position restraint (em2.mdp). Now both the system and the solvent can move and adjust considering their interactions.

The two option files used for this phase can be found in Appendix.

After the setting of the parameters and options, the commands to perform the productions on GROMACS are: grompp, that combine all the file given in a single file with all the informations (.tpr file) and mdrun, to submit the job.

```
gmx grompp -f em1.mdp -c eif4e-MGP_solv.gro -r eif4e-MGP_solv.gro
-p eif4e-MGP.top -n index.ndx -o mini_1.tpr -maxwarn 1

gmx mdrun -deffnm mini_1 -nb gpu -ntmpi 1
```

As can be seen in the example, grompp need the option file (.mdp), the coordinate file (.gro), the topology file (.top), and the index file (.ndx).

The index file is not always necessary, indeed in this case the system is divided in two groups: “protein and ligand” and “water and ions” (as can be read in the .mdp file in the appendix). This permits a more precise computation of the interactions.

The mdrun command need the flag -deffnm to define the prefix to all the files that it generates (it has been the same of .tpr unless it has to be specify), and with the flag -nb gpu and -ntmpi we indicate to GROMACS that the short range calculations have to be performed on gpu.

2.3 Thermalization

Thermalization is the process of physical bodies reaching thermal equilibrium through mutual interaction. In general, the natural tendency of a system is towards state of equipartition of energy and uniform temperature that maximizes the system’s entropy.

In a MD simulation Thermostat and Barostat are implied to ensure that the average temperature of a system is constant.

In these simulations, the Beredsen thermostat and V-rescale thermostat were used.

Berendsen thermostat

Berendsen thermostat is an algorithm that rescale velocities of particles in MD simulations to control the temperature.

The system is coupled with a heat bath with fixed temperature T_0, the thermostat suppresses fluctuation of the kinetic energy of the system. The temperature of the system is corrected such that the deviation exponentially decays with some time constant τ (Hünenberger, 2005).

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \quad \rightarrow \quad \Delta T = \frac{\delta t}{\tau}[T_0 - T(t)]$$

Through calculation, not shown, is possible to obtain the value of the new temperature:

$$T'(t) = \sqrt{\left[1 + \frac{\delta t}{\tau}\left(\frac{T_0}{T(t)} - 1\right)\right] T(t)}$$

In the .mdp file the reference temperature will be set at the field “ref-t”, and the time constant τ at “tau-t”. If there are groups to couple separately, there are specified in “tc-grps” field.

It is not very precise and accurate, but it requires few computational resources and at low temperature the error is considered acceptable.

V-rescale thermostat

Temperature coupling using velocity rescaling with a stochastic term. This thermostat is similar to Berendsen coupling, with the same scaling using “tau-t” field, but the stochastic term ensures that a proper canonical ensemble is generated. The random seed is set with “ld-seed”. This thermostat works correctly even for tau-t = 0.

This thermostat is more precise unless it requires higher computational resources.

Pressure

In the case of pressure what change is the volume. The weak coupling is used to maintain the pressure constant. It is possible to modify the motus equation to archive is a first-order relaxation to the system through the set value P_0.

$$\frac{dP(t)}{dt} = \frac{P_0 - P(t)}{\tau_p}$$

For a molecular system, the intramolecular forces and kinetic energy contribution can be omitted. The component of the pressure on the axis x, y and z is defined as the component of diagonal element of the tensor P. The isothermal compressibility (β_T) is linked to the pressure that change at constant temperature and at variable volume with the relation:

$$\Delta P(t) = -\frac{\Delta V(t)}{\beta_T V(t)}$$

Through calculation on the equation terms, it is possible to obtain:

$$\frac{dr_i(t)}{dt} = v_i(t) - \frac{\beta_T}{3\tau_p}[P_0 - P(t)]r_i(t)$$

The second term of the equation define the relaxation of P(t) through P_0, the velocity of the relaxation is controlled by the ratio between isothermal compressibility (β_T) and the time of relaxation of the pressure (τ_p) chosen.

τ_p has to be enough small (strong coupling) to permit the equilibration of the pressure, but enough wide (weak coupling) to avoid interfering with system properties.

Parrinello-Rahman barostat

Extended-ensemble pressure coupling where the box vectors are subject to an equation of motion and the equation of motion for the atoms is coupled to this. No instantaneous scaling takes place. The time constant tau-p (τ_p) is the period of pressure fluctuations at equilibrium. This is one of the most used and reliable method for pressure coupling.

Thermalization steps

As can be read in the workflow of the simulation. The thermalization phase consists of five steps:

1. 500 ps of NVT simulation (N: particles, V: volume, T: temperature is constant), at 100K with Berendsen thermostat. Integration step was set at 2 fs (0,002 ns), number of steps = 250000 with position restraint at 200. File: nvt1.mdp
2. 500 ps of NVT simulation at 200K with Berendsen thermostat. Integration step was set at 2 fs (0,002 ns), number of steps = 250000 with position restraint at 100. File: nvt2.mdp
3. 500 ps of NVT simulation at 300K with Berendsen thermostat. Integration step was set at 2 fs (0,002 ns), number of steps = 250000 with position restraint at 50. File: nvt3.mdp
4. 1000 ps (1 ns) of NPT (N: particles, P: pressure, T: temperature are constant) at 300K, with V-rescale thermostat and Parrinello-Rahman barostat. Integration step was set at 2 fs (0,002 ns), number of steps = 500000 with position restraint at 25. File: npt1.mdp
5. 5000 ps (5 ns) of NPT at 300K, with V-rescale thermostat and Parrinello-Rahman barostat. Integration step was set at 2 fs (0,002 ns), number of steps = 2500000 with position restraint at 15. File: npt2.mdp

With this protocol it is possible to gradually increase the temperature, and gradually decrease position restraints. In every step the system has the time to

stabilize and relax, the graphical output of this phase should be like three steps of a staircase (three NVT steps) and finally a plateau corresponding to the two NPT step. Mdp file can be found in Appendix.

Integration algorithm

Both in the thermalization and the run production, the algorithm used is the leap-frog algorithm, specified in the .mdp file with the acronym "md".

This algorithm is obtained by the approximation of the equation of motion with on degree of freedom with Euler method.

This permit, starting to initial points x_0 and $v_{1/2}$ continue it is possible to continue with x and v leapfrogging over each other as shown in Figure 12, following the integration formulas:

$$x_{n+1} = x_n + h_{n+1/2}$$

$$v_{n+3/2} = v_{n+1/2} + hF(x_{n+1})$$

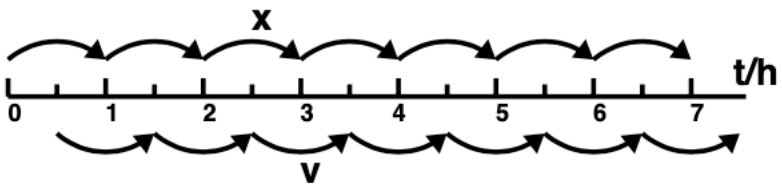


Figure 12. Schematic structure of leap frog algorithm cite (Young, 2009)

On GROMACS the command to submit the run are the same of the minimization (grompp, mdrun).

2.4 Run production

After the thermalization phase is possible to start the production of the simulation.

The trajectories were calculated at 300K, with leapfrog integration algorithm, position restraints were completely delated, and each trajectory had a length of 50 ns, integration step 2 fs (0,002 ns) number of steps 25000000.

On GROMACS the command to submit the run are the same of the minimization (grompp, mdrun). The mdp file can be found in Appendix.

2.5 Analysis

Calculation of RMSD

RMSD stands for Root Mean Square Deviation, and it is the measure of the average distance between a set of atoms. In this case the RMSD is calculated

for every frame of the simulation. The calculation has been performed with GROMACS, using the command ‘gmx rms’.

The trajectory was prepared by managing the Periodic Boundary Condition and it was roto-fitted in order to avoid artifacts.

The trajectory used for the RMSD calculation was the one obtained by the run without skipping it, it has 100.000 frames every one of 20 ps, for a total amount of time of 2 μs.

Every analysis was performed two time, the first fitting on the structure before the thermalization and the second on the structure after the thermalization. The three RMSD analysis are:

1. Fitting and calculation performed on the backbone atoms of the protein;
2. Fitting and calculation performed on the secondary structure elements (alpha helix and beta sheet) of the protein;
3. fitting performed on the secondary structure and the calculation performed on the ligand.

Evaluation of 4EGI COM and contacts with protein

The analysis were performed only after the equilibration phase, in order to characterize the movement of the ligand through the simulation time.

The calculation of the Center of Mass (COM) coordinates of the ligand respect to the protein COM, was performed with GROMACS, with the command gmx traj. The analysis was made on a skipped trajectory with 100ns per frame, in order to have a clear graphical representation of the movement through time. After that, the contacts of the 4EGI with the protein in the simulation time were calculated, in this case was used a skipped trajectory with frame length of 200 ps. Another contacts analysis was performed taking into consideration only the contacts with the residues of the binding pocket of the crystallized pose, in this time was used the same skipped trajectory (frame of 200ps). The contacts' counts were performed using GROMAS, with the command gmx mindist.

Analysis of biding pocket residues

In order to analyze deeply the changing of the biding mode other two analysis were performed:

- Calculation of the minimum distance between the center of mass of the biding pocket and the one atom of the ligand, performed with GROMACS, with the command gmx pairdist.
- Calculation of RMSF (Root Mean Square Fluctuation) of the protein residues, focusing on the pocket residues to investigate their fluctuation

through the simulation time. This calculation was carried out with GROMACS, command gmx rmsf.

The trajectory was skipped in order to have one frame every 200 ps.

Cluster and cavity analysis

Finally, a cluster analysis and an evaluation of the variation of the cavity volume were performed.

The cluster analysis was carried out using GROMACS, with the command gmx cluster, using gromos and a cutoff of 1,8 Å. The trajectory was skipped in order to have one frame every 2 ns.

Gromos use algorithm as described in Daura et al. (Daura et al., 1999). Count number of neighbors using cut-off, take structure with largest number of neighbors with all its neighbors as cluster and eliminate it from the pool of clusters. Repeat for remaining structures in pool.

The cavity analysis was performed using POVME 3.0 (Wagner et al., 2017), through a simulation that was skipped in order to have one frame every 10 ns. The cavity was calculated every time using the POVME mode 'DefinePocketByLigand' and the 'ConvexHullExclusion' was set to 'max'.

Graphs and images

Graphis were obtained with gnuplot (Williams & Kelley, 2010). Images were obtained using ChimeraX (Pettersen et al., 2021) and Maestro suite (Halgren et al., 2004).

3. RESULTS

3.1 Molecule preparation

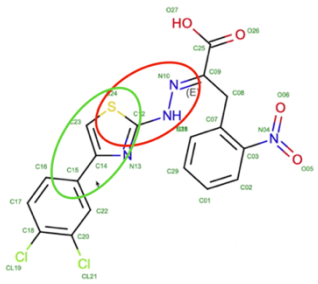
The molecule was prepared as described in “Material and Methods” section. Here are presented only the results of the ligands’ parametrization, minimization and thermalization of the system are not shown.

Ligand parametrization

As previously stated in “Materials and Methods” section, ligands were parametrize using GAFF2 and the complex was built using antechamber.

To have reliable parameters, the 4EGI was tested with QM methods, with particular attention to the torsional angles. The bonds taken into consideration can be observed below (Figure 13.A): circled in red the bond between the thiazole with the two nitrogens (S-C-N-N), circled in green the thiazole with the dichlorobenzene ring. The first calculation of the energetic profile obtained with GAFF2 compared to the one obtained with Gaussian, shows a local minimum at π, that is not present in the energy profile obtained by Gaussian. A correction has been applied in order to homologate the GAFF2 parameters to the QM energy profile, to do this *ad hoc* parameters were created (Figure 13.B).

A



B

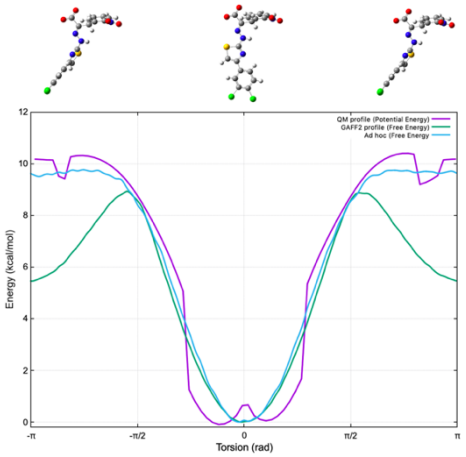


Figure 13. Ligand parametrization. A. 4EGI and bonds taken into consideration for the QM analysis. B . Profile representation of torsional circled in red. In purple QM profile obtained with Gaussian. In green GAFF2 profile. In light blue ad hoc parameters energy profile.

Regarding the MGP, the difference between the GAFF2 energy profile and the one obtained with Gaussian was lower than 1 kcal/mol, the original parameters were considered reliable, and no corrections were made.

3.2 Binding modes

The two complexes taken into consideration are eIF4E-4EGI (crystal 4TPW) and eIF4E-biphenyl (crystal 7MEU). The biding modes were described in the "Introduction" section.

In Table 5 are shown the residues' numbers of the biding pocket of the complex eIF4E-4EGI (structure 4TPW) and of the complex eIF4E-biphenyl (complex 7MEU). In the same table are reported the corresponding residues in the complex obtained with Amber for the simulation, indeed in the crystal structures the amino acids count start from 32, while in the complexes amino

acids count starts from 1. In Figure 14 are shown the two binding modes, both the 3D structure (Fig14 A and C) and the 2D diagrams of the binding modes (Fig.14 B and D).

Table 5. Binding pocket residues

Crystal 4TPW	Complex 4EGI-eIF4E	Crystal 7MEU	Complex biphenyl-eIF4E
Phe 47	Phe 15	**Leu 75**	Leu 43
Lys 49	Lys 17	**Tyr 76**	Tyr 44
Ser 83	Arg 29	**Ile 79**	Ile 47
Ile 79	Ile 47	**Gln 80**	Gln 48
Ser 83	Ser 51	**Tyr 91**	Tyr 59
		Arg 123	Arg 91
		Trp 130	Trp 98

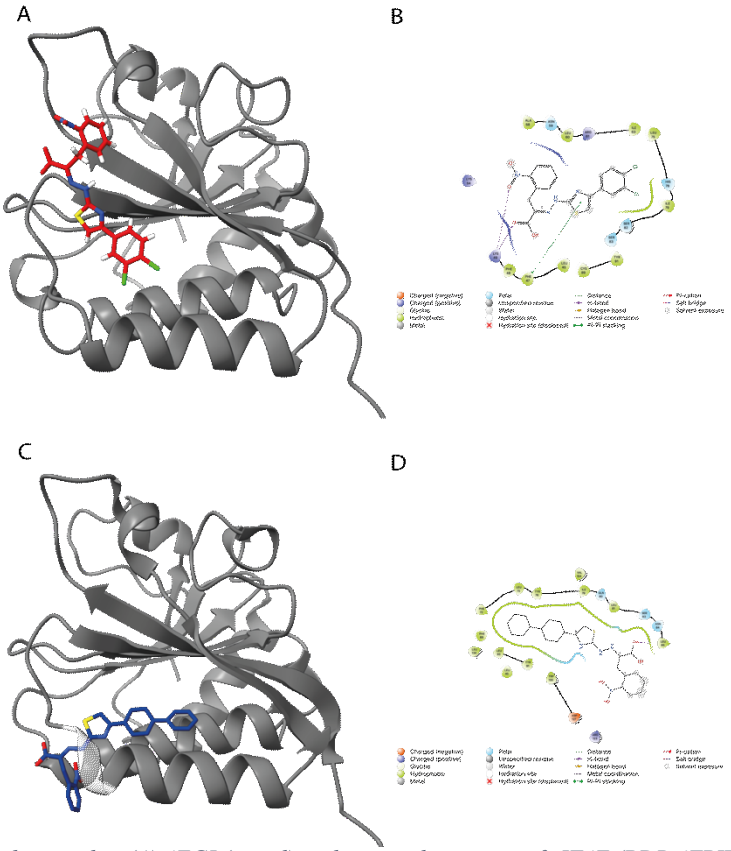


Figure 14, Binding modes. (A) 4EGI (in red) in the crystal structure of eIF4E (PDB 4TPW). (B) 2D ligand diagram of the 4EGI in the structure A. (C) Biphenyl (in blue) in the crystal structure of eIF4E (PDB 7MEU). (D) 2D ligand diagram of the biphenyl in the structure C.

3.3 RMSD analysis

As previously stated into “Material and Methods” section, three different RMSD analyses were performed, the plotted results can be observed in Figure 15.

In the first two graphs (Fig.14 A, B), the red lines represent the RMSD calculated fitting on the non-thermalized complex, while the black lines represent the RMSD calculated fitting on the thermalized complex.

The first figure (Fig14 A) represents the RMSD of the trajectory calculated considering the protein backbone. As reported the RMSD has a high level of fluctuation, particularly at the three peaks: (I) at 0,5774 μs, RMSD of 2,616 Å; (II) at 1,1284 μs, RMSD of 1,858 Å; (III) at 1,8226 μs 2,808 Å.

In the second graph (Fig14, B) is presented the RMSD of the trajectory calculated on the secondary structure of the protein (alpha helix and beta sheet). As expected, the variation is significantly lower than the previews graph.

The latest graph represents the RMSD of the ligand, the different poses acquired by the ligand over simulation time can be observed in the four representations of the 3D structures below the graph. The structures are the average structures computed in four different windows, one every 0,5 μs.

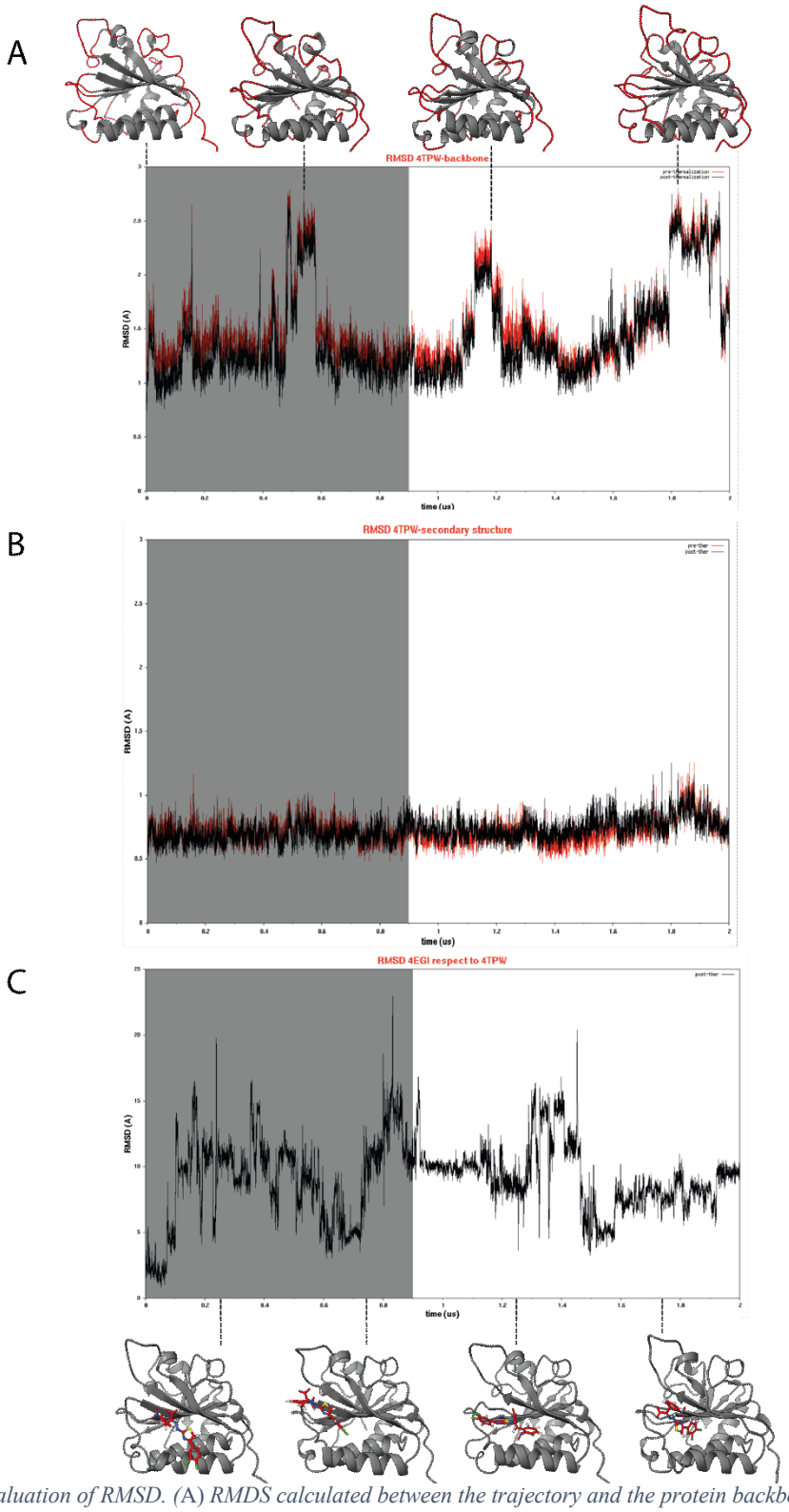


Figure 15. Evaluation of RMSD. (A) *RMDS calculated between the trajectory and the protein backbone pre- and post-thermalization (in red and black respectively), conformations at peaks are shown above the plot.* (B) *RMDS calculated between the trajectory and the backbone of secondary structures pre- and post-thermalization (in red and black respectively).* (C) *RMSD calculated between the ligand through the trajectory and the backbone of secondary structure of the protein. The trajectory has been divided in four windows of 500ns and the average structure is shown below the plot. The time corresponding to the equilibration time is shadowed in the three plots (0-900ns).*

The shadowed region (0-0,9 μs) in the three graphs represents the equilibration time. Indeed, the production phase considered for the next analyses starts from 0,9 μs, hence the analyzed trajectory is 1,1 μs.

3.4 Analyses of 4EGI Center of Mass (COM) distances form the binding pocket and the number of contacts with protein

In Figure 16 is reported the analysis of the Center of Mass (COM) distances of the ligand from the binding pocket and the number of contacts between the ligand and the residues composing the binding pocket.

As previously stated into "Material and Methods" section in the first figures (Fig 16 A) eleven points are considered (frame length: 100ns). On the left there is the 3D plot of the center of mass colored with gradient (from black to yellow) depending on time, the colors are respected in the images at the right, where can be observed the whole protein with the COMs and a zoomed region with the fluctuation of the ligand COM.

The drift of the center of mass over simulation time can be followed in the zoomed region where the path is represented as yellow dashed lines and numbers near the spheres.

The distances between the centers of mass are written partially in the zoomed region and entirely in the table below (Table 6).

Table 6. Distances of Center of Mass

Center of mass	Distances Å
C1-C2	4,43
C2-C3	2,68
C3-C4	6,33
C4-C5	5,73
C5-C6	5,55
C6-C7	15,51
C7-C8	6,23
C8-C9	1,43
C9-C10	2,22
C10-C11	2,88
C11-C12	2,82

In the Figure 16 B are presented the plotted data of the number of contacts of the 4EGI with the protein over the production phase (1,1 μs). Below the graphs represented the complex eIF4E-4EGI (protein as ribbon and surface, ligand in red), it is possible to observe the different position acquired by the ligand at the starting position, the one at 0,5532 μs with the minimum value of contacts (109) and the ending position.

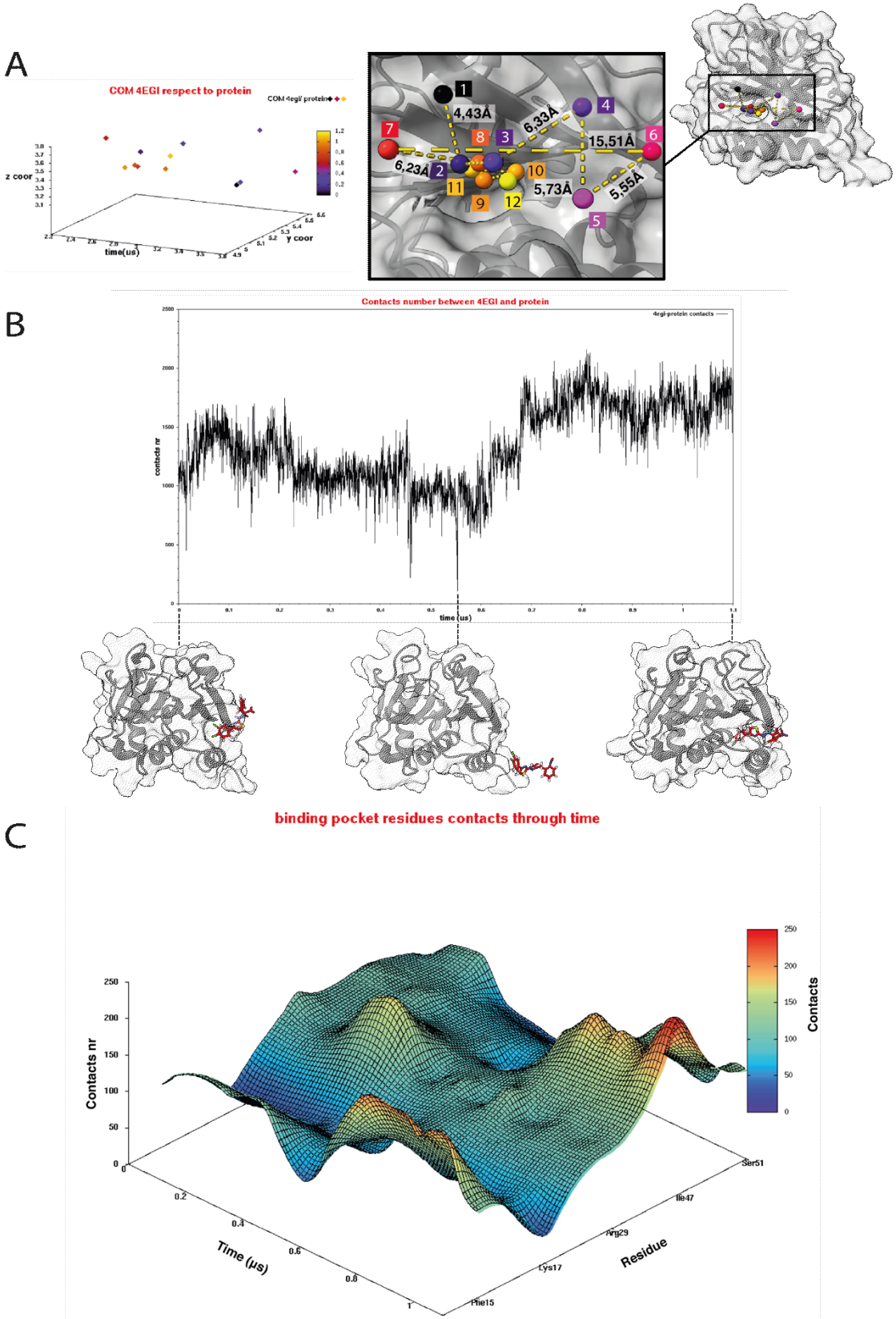


Figure 16. Evaluation of 4EGI COM and contacts. (A) *Plot of COM of the ligand every 100 ns for 1,1 µs, graphical representation of COM with zoom of the ligand region.* (B) *Plot of the contacts between ligand and protein through time, below are presented the starting conformation, the ending one and the one at the minimum.* (C) *3D representation of the changing of number of contacts through time regarding the residue of the binding pocket of the crystallized pose (pdb 4TPW) The data has been reconsidered from a Cauchy function.*

La last 3D image (Fig16 C) represents the contacts of the 4EGI with the binding pocket residues of the structure 4TPW. They are represented in a 3D plot where x is the time, y the residue name and z the contacts' number.

The fluctuation of contacts is represented in this graphs as peaks and valleys, with a color gradient depending on the contacts' number (0-50 blue, 50-100 light-blue, 100-150 green, 150-200 yellow, 200-250 red).

3.5 Analysis of biding pocket residues

In Figure 17 are reported the analysis performed on the residues of the binding pocket (crystalized complex eIF4E-4EGI structure 4TPW). In Figure 17 A the residues of interest are presented with ball and stick colored in blue in the 3D ribbon structure of the eIF4E.

The second graph (Fig.17B) represents the distances of the 4EGI from the center of mass of the biding pocket over production time. Above the graph can be observed three poses of the ligand: (I) starting time; (II) maximum distance (0,46 μs, distance 7,26 Å); (III) ending time of the simulation. The ligand is represented in red, and the protein in grey with ribbon and surface, with the residues of interest in blue.

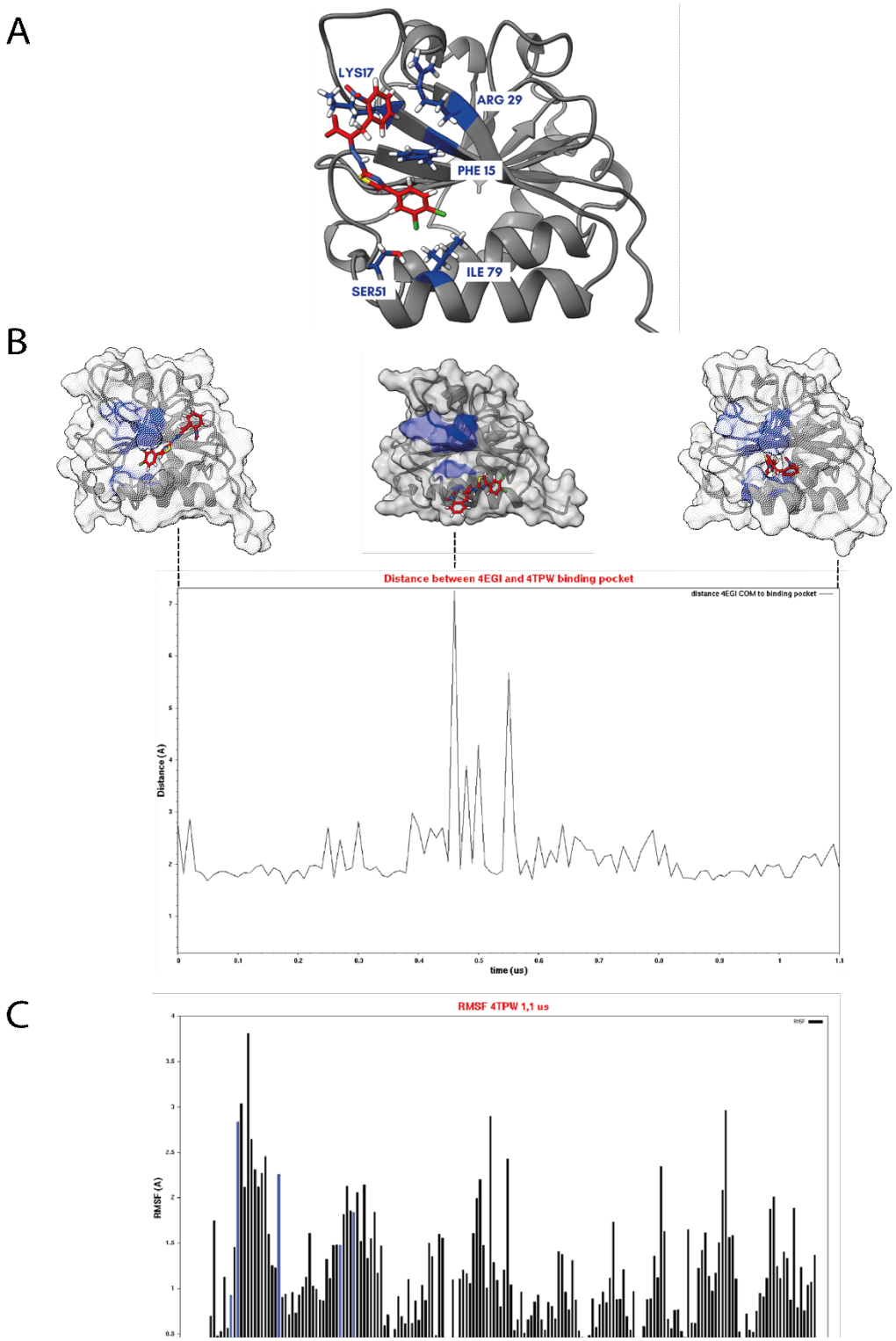


Figure 17. Evalutation of the biding pocket residues. (A) *Representation of the biding pocket residues (PDB 4TPW)* (B) *Distance between the COM of the biding pocket and the ligand, above are presented the starting pose, the one at the maximum and the ending one.* (C) *Plot of RMSF through 1,1 μs, in blue are highlighted the binding pocket residues.*

As reported, ligand firstly is near the biding pocket, in the middle of the production it tends to move away from it and came nearer at the end of the simulation time.

The last graph of Figure 17 is the RMSF of the residues of eIF4E through simulation, in blue are highlighted the residues of the binding pocket. In this analysis the C-terminal portion (from amino acid 1 to 8) was remove, to have more comparable measurements, indeed the fluctuation of these residues was significantly higher than the one of other amino acids.

3.6 Cluster and cavity analysis

In Figure 18 is presented the cluster analysis over production time. In the first image (Fig.18 A) as dots is presented the distribution of clusters, the equilibration phase is shadowed.

From the analysis over the entire simulation time, 40 clusters arose. The most significative for our analyses were the first twelve cluster, and the size of each cluster is reported in Table 7.

As can be observed in the graph, the cluster 1,2,4 and 5 are the more abundant in the shadowed region, while the 1,8,9,7,3,6,12 and 11 arise in the production.

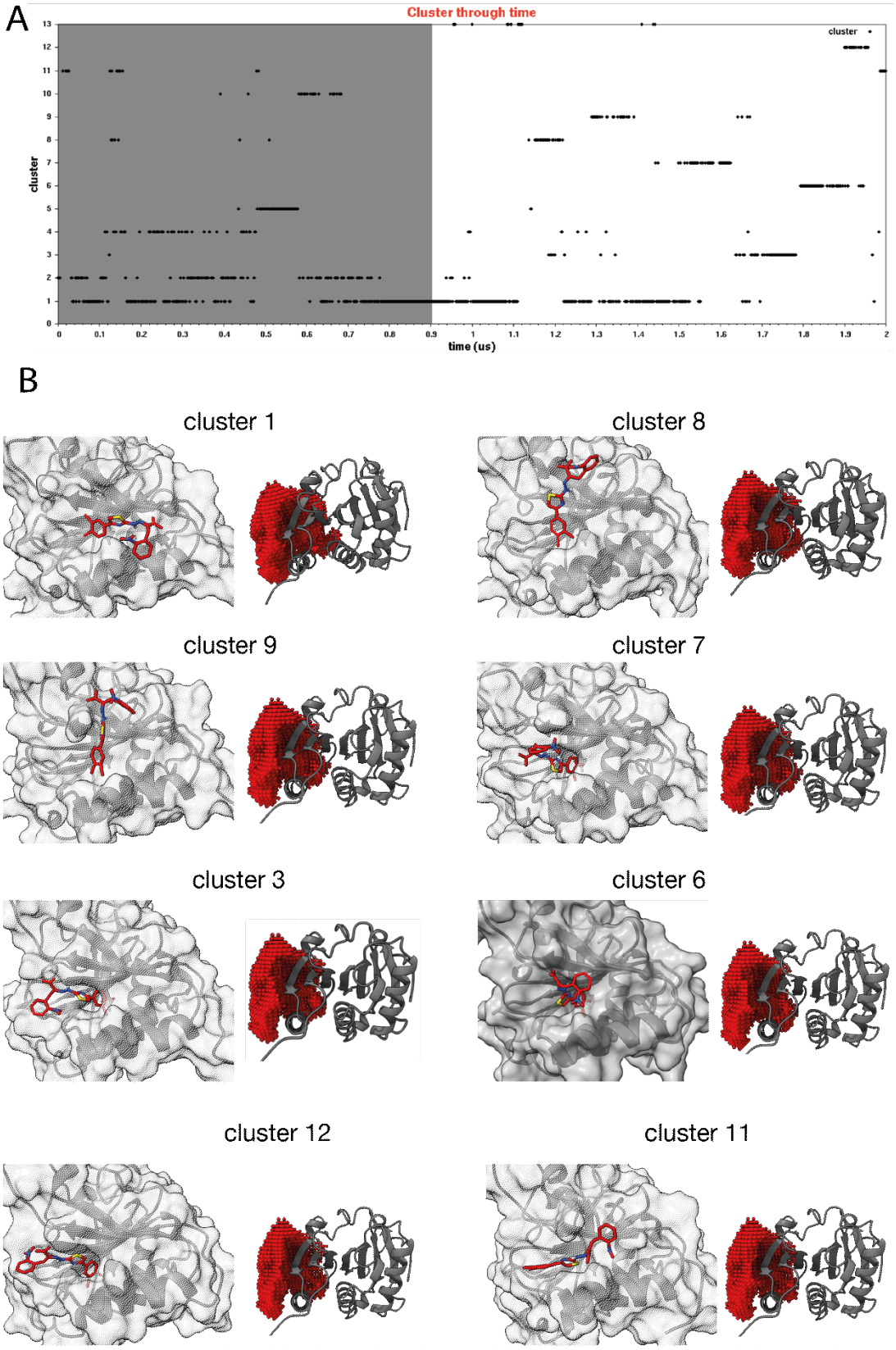


Figure 18. Cluster and cavity analysis. (A) Cluster through time. (B) Detailed representation of cluster, only most significative the cluster after 900 ns (equilibration time) are presented, they are presented in chronological order, it is presented the ligand pose.

In the second part of the panel are reported the structures with the smallest average distance to the others member of the cluster (Figure 8 B), with the correspondence figure of the cavity analysis performed with POVME3. The cavity analysis underline the volume region taken by the ligand.

The clusters are presented in chronological order considering they arise time in the first graph (Fig18 A), it is possible to observe the movement of the ligand through time with the different poses acquired in each cluster.

Table 7. Cluster size

Cluster	Size	Cluster	Size
1	395	**21**	6
2	99	**22**	5
3	60	**23**	5
4	51	**24**	5
5	49	**25**	4
6	48	**26**	4
7	40	**27**	3
8	33	**28**	2
9	30	**29**	2
10	25	**30**	2
11	20	**31**	1
12	20	**32**	1
13	17	**33**	1
14	12	**34**	1
15	12	**35**	1
16	12	**36**	1
17	9	**37**	1
18	8	**38**	1
19	7	**39**	1
20	6	**40**	1

3.7 Pose comparison

Finally, it is presented a comparison of the poses acquired by the ligand in the cluster 12 (the cluster that arose only in the last part of the simulation) and the complex eIF4E-biphenyl (crystal structure 7MEU).

In Figure 19 A is reported in detail the three residues that interact with the 4EGI in common with the biphenyl binding pocket (all biding pocket residues in Table 5): Tyr 59, Trp98, Ile47 (following the enumeration of the cluster structure).

In Figure 19 B is presented the superimposition of the two structures, with emphasis on the three residues described above. The 4EGI is presented in red and the biphenyl in blue.

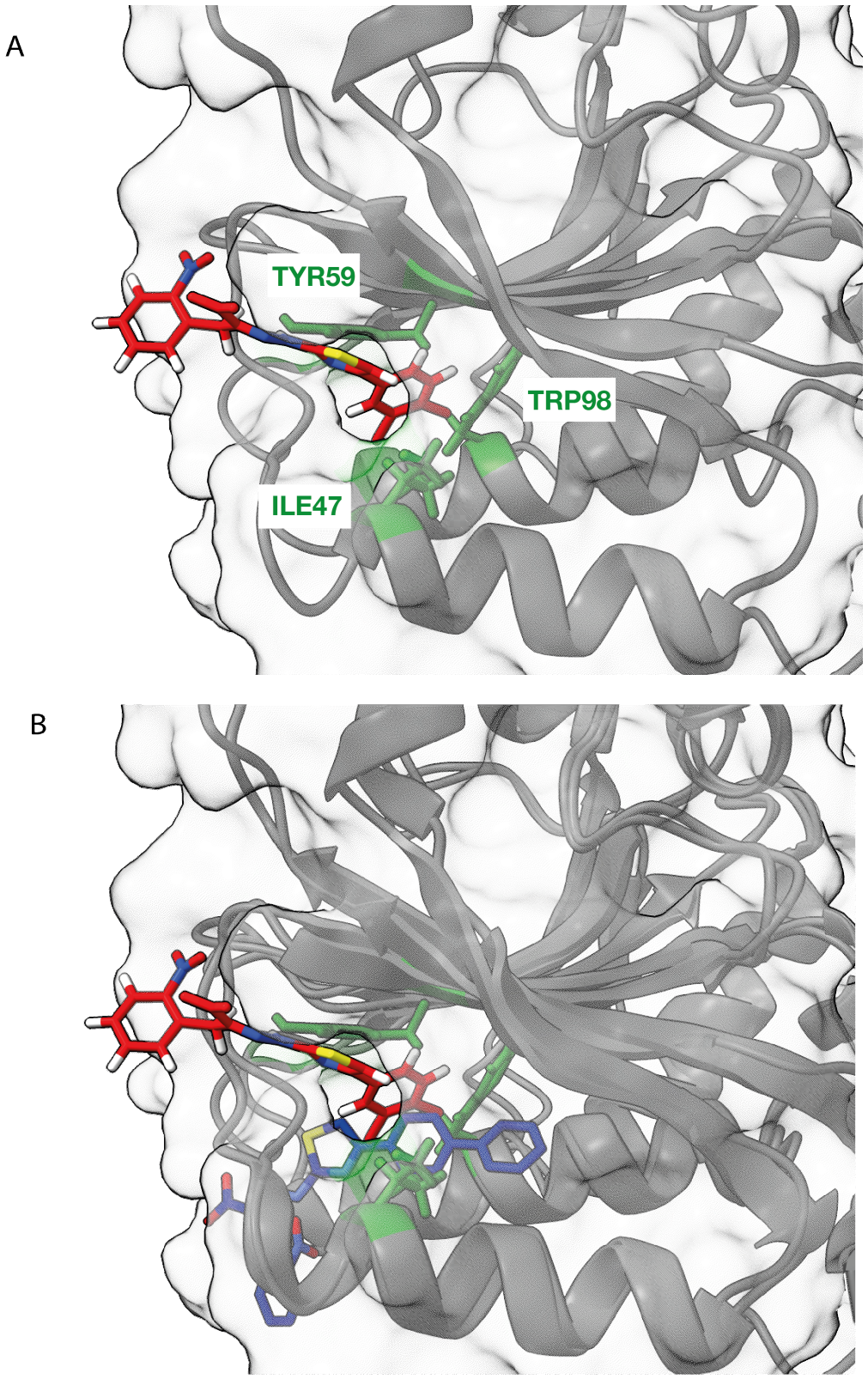


Figure 19. Comparison between pose of 4EGI in cluster 12 and biphenyl. (A) Representation of 4EGI in the pose of cluster 12, in green are presented three residue of the biphenyl binding pocket that make contacts with the 4EGI. (B) Superimposition of the cluster 12 and the crystal 7MEU.

4. DISCUSSION

4.1 Stability of the complex

As already reported in the "Introduction" section and in Figure 14, it is quite unusual that two molecules with a similar chemical structure such as 4EGI and biphenyl, can bind with the same efficiency and stability the eIF4E in two completely different binding modes.

To open the way at the development of analogs that can improve the activity of the 4EGI as a promising drug, the stability of the complex has been investigated analyzing several structural and dynamical features that were reported in the previews section.

The evaluation of the RMSD evolution of the protein backbone over the simulation time was the first analysis performed (Fig. 15A). The RMSD shows a high level of fluctuation due to the presence of several loops (Fig. 15A, red structures).

This observation was confirmed by the RMSD calculated only on the backbone of the secondary structure elements of the protein (Fig. 15B) showing a significant low degree of fluctuations over the simulation time. These results are also confirmed for the equilibration phase (from 0 to 0,9 μs) being from

0,9 μs to 2 μs the simulation time taken into consideration for all the remaining analyses.

In light of the RMSD results, it is possible to affirm that the protein is stable for the entire simulation time.

The ligand shows a peculiar behavior since its RMSD profile is characterized by high fluctuations as reported in figure 15C for both equilibration and production phases.

To have a structural representation of this variation the average structure of the four windows (0-05 μs, 0,5-1 μs, 1-1,5 μs, 1,5-2 μs) can be observed below the graph, and the ligand seems to acquire different poses in the simulation time.

4.2 Ligand shift into the binding pocket

To deeply describe the ligand shift into the binding pocket over the production phase (*i.e.,* last 1,1 μs) several analyses were carried out and reported in the "Result section".

First, to understand the drift of the 4EGI, the path of different positions acquired by its center of mass respect to the protein center of mass was tracked (Fig 16A). As reported in the zoomed part of the structural representation (Fig 16A),

the ligand tends to quickly change its pose, especially in the first 0,700 μs (from point 1 to point 8) where the maximum distance between points is 15,51 Å (from point 6 to 7), while it seems to move less in the last 0,4 μs (point 8 to 12) where the maximum distance between points is only 2,88 Å (point 10 to 11). The decreased distances between last points suggest a more stable position that can be due to the formation of new interactions builds by the ligand with residues near the ones contacted at the start of the production time.

To understand better the behavior of the ligand, contacts analyses were carried out represented from graphs in Figure 16B and C.

The analyses that involves all the protein residues (Fig 16B) underlines that the ligand progressively loses contacts till the minimum reached at 0,5532 μs, then there is a turnaround and contacts start to increase.

This trend of the contacts' number finds a correspondence with the shift of center of mass, indeed the minimum is in a middle position from point 5 and 6, that are the farthest, then the sudden increase (0,6-0,8 μs) correspond to the shift between 6, 7 and 8 and the final plateau of the increased contacts confirm the hypothesis that in the final position of COMs (from 8 to 12) the ligand develops more interactions with the protein.

Another evidence of the changing of contacts is the last 3D plot (Fig 16C), where was presented the changing of the number of contacts concerning only the residue of the binding pocket detectable on the crystal structure (4TPW). At the production starting time the ligand was not in the pocket, because it moved during thermalization and equilibration phase, but except for Lys17 that never acquired again the contacts, the other residues changed their interactions with the ligand though the production time. It is noteworthy that from 0,4 μs and 0,8 μs the graph has valleys followed by peaks in particularly in Phe15, Ile47 and Ser51. This can be compared with the general graph of contacts, where there is an increase of contacts in the same time frame.

Particularly attention has to be taken to the ending part of the graphs, where is possible to observed an increase of number of contacts of Phe15, Ile47 and Ser51 this underlines that the ligand comes back in a region near the initial biding pocket, but it does not contact the same residues.

The distance between the ligand and the biding pocket of the crystallized pose was measured (Fig 17 B). The residues considered for the calculation were reported with labels in Figure 17 A and in 3D structure (Fig 17 B) with the same surface color.

From the plot of distances, it is possible observe again that the ligand moved away from the initial binding pocket region, especially between 0,4 and 0,6 μs, that correspond to the center of mass positions labelled with 5, 6 and 7. Moreover the pose acquired by the ligand at the end of the simulation time showed that the ligand is further from the surface area of the Lys17 and Arg29 respect to the other residues, this confirms the behavior seen in the 3D plot of contacts of the biding pocket (Fig 16 C).

The last analysis performed on the biding pocket residues was the RMSF. A low RMSF is expected from bound residues, instead the RMSF calculation (Fig 17C) of the binding pocket residues was high, suggesting rearrangement events of the residues over production time.

4.3 Evolution of the complex

To have an idea of the principal poses acquired by the ligand over simulation time, a cluster analysis was performed (Fig 18).

It is clear from cluster distribution that in the equilibration phase there is a high variability of clusters, indeed the membership of structures to cluster 1, 2 and 4 exchanges very quickly. At the production starting time (last 1,1 μs) the situation seems to be more stable.

Considering the succession of 3D structures reporting ligand poses and cavity variation (18B), the ligand seems to follow a precise path over production time. Indeed, the o-nitro-phenyl group tends to move up (cluster 8 and 9), to permit the movement of the diclorobenzene group into a region between alpha helix 3 and beta sheet 2 that over the simulation time become a cavity (cluster 7-3-6).

Unless the last conformations belong to cluster 11, that is poorly present in the first 0,2 μs of the simulation, cluster 12 and 6 seem to be the representative models of the new biding mode acquired by 4EGI. Indeed, as can be observed (Fig 18B) they differ only for the rotation of the o-nitro-phenyl group, but the diclorobenzene tends to maintain the same interactions in the two clusters.

In cluster 12 the ligand tends to arrange contacts with the pocket of biphenyl. Indeed, the diclorobenzene ring enter a cavity where there are the Tyr59, Trp98 and Ile47 (Figure 19A), that are residues that also take part to the biding of the biphenyl.

To point out the position of 4EGI pose in cluster 12 and the biphenyl biding mode, a superimposition of the two structures was performed (Figure 19B) and

it is even clearer the tendency of the diclorobenzene ring to interact with the same interaction residues of the biphenyl benzene linked to the tiazoline ring.

5. CONCLUSIONS

5.1 Conclusion of the study

Collecting all results, it is possible to conclude that the binding mode of the 4EGI that was cocrystalized with the eIF4E in 2014 (Papadopoulos et al., 2014), is not a stable binding mode.

The high drift of the ligand position into the eIF4E binding pocket and the changing of the orientation into the eIF4E binding pocket over the simulation time suggest that the pharmacological activity of the 4EGI could not be explained by this binding mode, that although it is possible, it is not stable and maybe it was created by the crystallization conditions.

The MD simulations of the 7MEU crystal structure and the eIF4E-4EGI complex in which the 4EGI was docked into the biphenyl biding pocket were conducted in parallel to the previous MD simulation confirming that the by binding pocket observed for the biphenyl molecule could be suitable also for the interaction of 4EGI.

5.2 Perspectives

Next steps to validate the theoretical results obtained so far, some specific point mutations of the keys residues identified into the eIF4E biphenyl binding

pocket will be designed and experimentally tested. The binding activity of eIF4E to 4EGI and biphenyl will be evaluated for the mutants giving a wider view of the binding mechanism of both ligands. Based on the results obtained for the mutants we will proceed in the development of 4EGI analogs in order to increase the 4EGI pharmacological activity.

Appendix

.mdp file

em1.mdp

```
100title = EM1
define = -DPOSRES5000
constraints = none
integrator = steep
dt = 0.002 ; ps !
nsteps = 10000
nstlist = 10
ns_type = grid
;pbc = xyz
rlist = 1.3
periodic_molecules  = no
cutoff-scheme                   = Verlet
verlet-buffer-tolerance         = 0.005
coulombtype                     = PME
coulomb-modifier                = Potential-shift
rcoulomb-switch                 = 1
rcoulomb                        = 1.2
vdw-type                        = Cut-off
vdw-modifier                    = Potential-shift
rvdw-switch                     = 1
rvdw                            = 1.2
fourierspacing = 0.12
pme_order = 4
ewald_rtol = 1e-5
optimize_fft = yes
;
; Energy minimizing stuff
;
emtol = 100.0
emstep = 0.01
nstxout                         = 1000
nstvout                         = 1000
nstlog                          = 500
nstfout                         = 0
nstxout-compressed              = 500
compressed-x-precision          = 1000
```

em2.mdp

```
title = EM2
constraints = none
integrator = steep
dt = 0.002 ; ps !
nsteps = 50000
nstlist = 10
ns_type = grid
;pbc = xyz
rlist = 1.3
periodic_molecules  = no
```

```
cutoff-scheme                   = Verlet
verlet-buffer-tolerance         = 0.005
coulombtype                     = PME
coulomb-modifier                = Potential-shift
rcoulomb-switch                 = 1
rcoulomb                        = 1.2
vdw-type                        = Cut-off
vdw-modifier                    = Potential-shift
rvdw-switch                     = 1
rvdw                            = 1.2
fourierspacing = 0.12
pme_order = 4
ewald_rtol = 1e-5
optimize_fft = yes
;
; Energy minimizing stuff
;
emtol = 1.0
emstep = 0.01
nstxout                         = 1000
nstvout                         = 1000
nstlog                          = 500
nstfout                         = 0
nstxout-compressed              = 500
compressed-x-precision          = 1000
```

nvt1.mdp

```
title = NVT1
;       Input file
;
define              =  -DPOSRES1000
; integrator and time step
integrator          =  md
nsteps              =  250000
dt                  =  0.002
;
; removing CM translation and rotation
comm_mode           =  Linear
nstcomm             =  100
;
; output control
nstlog                    = 500
nstenergy                 = 500
nstxout                   = 500
nstvout                   = 500
nstfout                  = 500
nstxtcout                 = 1000
;
; neighbour searching
nstlist             = 10
ns_type             = grid
pbc                 = xyz
rlist               = 1.3
;
; electrostatic
cutoff-scheme                   = Verlet
rcoulomb            = 1.2
```

```
rcoulomb_switch     = 0.0
coulombtype         = PME


;
; Ewald
fourierspacing      = 0.1
pme_order           = 4
;
; vdw
vdw-type                 = Cut-off
rvdw-switch         = 0.00
rvdw                = 1.2

;
; constraints
constraints              = all-bonds
constraint-algorithm     = lincs
lincs_iter               = 1
lincs_order              = 4
;
; temperature
;Tcoupl              = no
Tcoupl              = berendsen
tc_grps             = water non-water
tau_t               = 0.1 0.1
ref_t               = 100 100
;
; pression
Pcoupl              = no
;Pcoupl              =  berendsen
;Pcoupltype          =  isotropic
;tau_p               =  0.5
;compressibility     =  4.5e-5
;ref_p               =  1.0
;
; initial velocities
gen_vel             = yes
gen_temp            = 0
gen_seed            = -1
```

nvt2.mdp

```
title = NVT2
;      Input file
;
define              =  -DPOSRES500
; integrator and time step
integrator          =  md
nsteps              =  250000
dt                  =  0.002
;ppp
; removing CM translation and rotation
comm_mode           =  Linear
nstcomm             =  100
;
; output control
nstlog                   = 500
```

```
nstenergy               = 500
nstxout                 = 500
nstvout                 = 500
nstfout               = 500
nstxtcout               = 1000
;
; neighbour searching
nstlist             = 10
ns_type             = grid
pbc                 = xyz
rlist               = 1.3
;
; electrostatic
cutoff-scheme                   = Verlet
rcoulomb            = 1.2
rcoulomb_switch     = 0.0
coulombtype         = PME


;
; Ewald
fourierspacing      = 0.1
pme_order           = 4
;
; vdw
vdw-type                = Cut-off
rvdw-switch         = 0.00
rvdw                = 1.2
;
; constraints
constraints             = all-bonds
constraint-algorithm    = lincs
lincs_iter              = 6
lincs_order             = 6
;
; temperature
;Tcoupl              = no
Tcoupl              = berendsen
tc_grps             = water  non-water
tau_t               = 0.1 0.1
ref_t               = 200 200
;
; pression
Pcoupl              = no
;Pcoupl              =  berendsen
;Pcoupltype          =  isotropic
;tau_p               =  0.5
;compressibility     =  4.5e-5
;ref_p               =  1.0
;
; initial velocities
gen_vel             = no

continuation = yes
```

nvt3.mdp

```
title = NVT3
```

```
;      Input file
;
define              =  -DPOSRES250
; integrator and time step
integrator          =  md
nsteps              =  250000
dt                  =  0.002
;
; removing CM translation and rotation
comm_mode           =  Linear
nstcomm             =  100
;
; output control
nstlog                    = 500
nstenergy                 = 500
nstxout                   = 500
nstvout                   = 500
nstfout                  = 500
nstxtcout                 = 1000
;
; neighbour searching
nstlist             = 10
ns_type             = grid
pbc                 = xyz
rlist               = 1.3
;
; electrostatic
cutoff-scheme                   = Verlet
rcoulomb            = 1.2
rcoulomb_switch     = 0.0
coulombtype         = PME


;
; Ewald
fourierspacing      = 0.1
pme_order           = 4
;
; vdw
vdw-type                 = Cut-off
rvdw-switch         = 0.00
rvdw                = 1.2
;

; constraints
constraints              = all-bonds
constraint-algorithm     = lincs
lincs_iter               = 6
lincs_order              = 6
;
; temperature
;Tcoupl              = no
Tcoupl              = berendsen
tc_grps             = water  non-water
tau_t               = 0.1 0.1
ref_t               = 300 300
;
; pression
Pcoupl              = no
;Pcoupl              =  berendsen
;Pcoupltype          =  isotropic
```

```
;tau_p              =  0.5
;compressibility    =  4.5e-5
;ref_p              =  1.0
;
; initial velocities
gen_vel            = no

continuation            = yes
```

npt1.mdp

```
define                  = -DPOSRES50

; RUN CONTROL PARAMETERS
integrator              = md
; Start time and timestep in ps
tinit                   = 0
dt                      = 0.002
nsteps                  = 500000 ;10000ps
simulation_part         = 1
init_step               = 0
comm-mode               = Linear
nstcomm                 = 100
comm-grps               = water non-water

; LANGEVIN DYNAMICS OPTIONS
; Friction coefficient (amu/ps) and random seed
bd-fric                 = 0
ld-seed                 = -1

; OUTPUT CONTROL OPTIONS
nstxout                 = 10000
nstvout                 = 10000
nstfout                 = 0
; Output frequency for energies to log file and energy file
nstlog                  = 500
nstenergy               = 500
; Output frequency and precision for xtc file
nstxtcout               = 10000
xtc-precision           = 1000
xtc-grps                =
energygrps              =

; NEIGHBORSEARCHING PARAMETERS
cutoff-scheme           = Verlet
nstlist                 = 20
ns_type                 = grid
pbc                     = xyz
periodic_molecules      = no
rlist                   = 1.3

; OPTIONS FOR ELECTROSTATICS AND VDW
; Method for doing electrostatics
coulombtype             = PME
rcoulomb-switch         = 0
rcoulomb                = 1.2
epsilon-r               = 1
epsilon_rf              = 0
```

```
vdw-type                 = Cut-off
rvdw-switch              = 0
rvdw                     = 1.2

DispCorr                 = EnerPres
table-extension          = 1
fourierspacing           = 0.12
fourier_nx               = 0
fourier_ny               = 0
fourier_nz               = 0

; EWALD/PME/PPPM parameters
pme_order                = 4
ewald_rtol               = 1e-05
ewald_geometry           = 3d
epsilon_surface          = 0
optimize_fft             = no


; OPTIONS FOR WEAK COUPLING ALGORITHMS
; Temperature coupling
Tcoupl                   = v-rescale
; Groups to couple separately
tc-grps                  = water non-water
; Time constant (ps) and reference temperature (K)
tau-t                    = 0.1 0.1
ref-t                    = 300 300
; Pressure coupling
Pcoupl                   = parrinello-rahman
Pcoupltype               = isotropic
; Time constant (ps), compressibility (1/bar) and reference P (bar)
tau-p                    = 5
compressibility          = 4.5e-5
ref-p                    = 1
; Scaling of reference coordinates, No, All or COM
refcoord_scaling         = com
; Random seed for Andersen thermostat
andersen_seed            = 815131


; GENERATE VELOCITIES FOR STARTUP RUN
gen_vel                  = no
gen-temp                 = 50
gen-seed                 = -1

; OPTIONS FOR BONDS
constraints              = all-bonds
; Type of constraint algorithm
constraint-algorithm     = lincs
; Do not constrain the start configuration
continuation             = yes
; Highest order in the expansion of the constraint coupling matrix
lincs-order              = 6
lincs-iter               = 6
lincs-warnangle          = 30
morse                    = no
```

npt2.mdp

```
define                  = -DPOSRES15

; RUN CONTROL PARAMETERS
integrator              = md
; Start time and timestep in ps
tinit                   = 0
dt                      = 0.002
nsteps                  = 2500000 ;10000ps
simulation_part         = 1
init_step               = 0
comm-mode               = Linear
nstcomm                 = 100
comm-grps               = water non-water

; LANGEVIN DYNAMICS OPTIONS
; Friction coefficient (amu/ps) and random seed
bd-fric                 = 0
ld-seed                 = -1

; OUTPUT CONTROL OPTIONS
nstxout                 = 10000
nstvout                 = 10000
nstfout                 = 0
; Output frequency for energies to log file and energy file
nstlog                  = 500
nstenergy               = 500
; Output frequency and precision for xtc file
nstxtcout               = 10000
xtc-precision           = 1000
xtc-grps                =
energygrps              =

; NEIGHBORSEARCHING PARAMETERS
cutoff-scheme           = Verlet
nstlist                 = 20
ns_type                 = grid
pbc                     = xyz
periodic_molecules      = no
rlist                   = 1.3

; OPTIONS FOR ELECTROSTATICS AND VDW
; Method for doing electrostatics
coulombtype             = PME
rcoulomb-switch         = 0
rcoulomb                = 1.2
epsilon-r               = 1
epsilon_rf              = 0
vdw-type                = Cut-off
rvdw-switch             = 0
rvdw                    = 1.2

DispCorr                = EnerPres
table-extension         = 1
fourierspacing          = 0.12
fourier_nx              = 0
fourier_ny              = 0
fourier_nz              = 0
```

```
; EWALD/PME/PPPM parameters
pme_order                = 4
ewald_rtol               = 1e-05
ewald_geometry           = 3d
epsilon_surface          = 0
optimize_fft             = no


; OPTIONS FOR WEAK COUPLING ALGORITHMS
; Temperature coupling
Tcoupl                   = v-rescale
; Groups to couple separately
tc-grps                  = water non-water
; Time constant (ps) and reference temperature (K)
tau-t                    = 0.5 0.5
ref-t                    = 300 300
; Pressure coupling
Pcoupl                   = parrinello-rahman
Pcoupltype               = isotropic
; Time constant (ps), compressibility (1/bar) and reference P (bar)
tau-p                    = 5
compressibility          = 4.5e-5
ref-p                    = 1
; Scaling of reference coordinates, No, All or COM
refcoord_scaling         = com
; Random seed for Andersen thermostat
andersen_seed            = 815131


; GENERATE VELOCITIES FOR STARTUP RUN
gen_vel                  = no
gen-temp                 = 50
gen-seed                 = -1

; OPTIONS FOR BONDS
constraints              = all-bonds
; Type of constraint algorithm
constraint-algorithm     = lincs
; Do not constrain the start configuration
continuation             = yes
; Highest order in the expansion of the constraint coupling matrix
lincs-order              = 4
lincs-iter               = 1
morse                    = no
```

run.0.mdp

```
define                   =

; RUN CONTROL PARAMETERS
integrator               = md
; Start time and timestep in ps
tinit                    = -50000
dt                       = 0.002
nsteps                   = 25000000 ;10000ps
simulation_part          = 1
```

```
init_step               = 0
comm-mode               = Linear
nstcomm                 = 100
comm-grps               = water non-water

; LANGEVIN DYNAMICS OPTIONS
; Friction coefficient (amu/ps) and random seed
bd-fric                 = 0
ld-seed                 = -1

; OUTPUT CONTROL OPTIONS
nstxout                 = 10000
nstvout                 = 10000
nstfout                 = 0
; Output frequency for energies to log file and energy file
nstlog                  = 500
nstenergy               = 500
; Output frequency and precision for xtc file
nstxtcout               = 10000
xtc-precision           = 1000
xtc-grps                = non-water
energygrps              =

; NEIGHBORSEARCHING PARAMETERS
cutoff-scheme           = Verlet
nstlist                 = 20
ns_type                 = grid
pbc                     = xyz
periodic_molecules      = no
rlist                   = 1.3

; OPTIONS FOR ELECTROSTATICS AND VDW
; Method for doing electrostatics
coulombtype             = PME
rcoulomb-switch         = 0
rcoulomb                = 1.2
epsilon-r               = 1
epsilon_rf              = 0
vdw-type                = Cut-off
rvdw-switch             = 0
rvdw                    = 1.2

DispCorr                = EnerPres
table-extension         = 1
fourierspacing          = 0.12
fourier_nx              = 0
fourier_ny              = 0
```

```
fourier_nz               = 0

; EWALD/PME/PPPM parameters
pme_order                = 4
ewald_rtol               = 1e-05
ewald_geometry           = 3d
epsilon_surface          = 0
optimize_fft             = no


; OPTIONS FOR WEAK COUPLING ALGORITHMS
; Temperature coupling
Tcoupl                   = v-rescale
; Groups to couple separately
tc-grps                  = water non-water
; Time constant (ps) and reference temperature (K)
tau-t                    = 0.5 0.5
ref-t                    = 300 300
; Pressure coupling
Pcoupl                   = parrinello-rahman
Pcoupltype               = isotropic
; Time constant (ps), compressibility (1/bar) and reference P (bar)
tau-p                    = 5
compressibility          = 4.5e-5
ref-p                    = 1
; Scaling of reference coordinates, No, All or COM
refcoord_scaling         = com
; Random seed for Andersen thermostat
andersen_seed            = 815131


; GENERATE VELOCITIES FOR STARTUP RUN
gen_vel                  = no
gen-temp                 = 50
gen-seed                 = -1

; OPTIONS FOR BONDS
constraints              = all-bonds
; Type of constraint algorithm
constraint-algorithm     = lincs
; Do not constrain the start configuration
continuation             = yes
; Highest order in the expansion of the constraint coupling matrix
lincs-order              = 4
lincs-iter               = 1
morse                    = no
```

Bibliography

Alder, B. J., & Wainwright, T. E. (1957). Phase transition for a hard sphere system. In *The Journal of Chemical Physics* (Vol. 27, Issue 5). https://doi.org/10.1063/1.1743957

Altmann, M., Schmitz, N., Berset, C., & Trachsel, H. (1997). A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. *EMBO Journal*, *16*(5). https://doi.org/10.1093/emboj/16.5.1114

Amorim, I. S., Lach, G., & Gkogkas, C. G. (2018). The Role of the Eukaryotic Translation Initiation Factor 4E (eIF4E) in Neuropsychiatric Disorders. In *Frontiers in Genetics* (Vol. 9). https://doi.org/10.3389/fgene.2018.00561

Andrieu, C., Taieb, D., Baylot, V., Ettinger, S., Soubeyran, P., De-Thonel, A., Nelson, C., Garrido, C., So, A., Fazli, L., Bladou, F., Gleave, M., Iovanna, J. L., & Rocchi, P. (2010). Heat shock protein 27 confers resistance to androgen ablation and chemotherapy in prostate cancer cells through eIF4E. *Oncogene*, *29*(13). https://doi.org/10.1038/onc.2009.479

Avdulov, S., Li, S., Michalek, V., Burrichter, D., Peterson, M., Perlman, D. M., Manivel, J. C., Sonenberg, N., Yee, D., Bitterman, P. B., & Polunovsky, V. A. (2004). Activation of translation complex eIF4F is essential for the genesis and maintenance of the malignant phenotype in human mammary epithelial cells. *Cancer Cell*, *5*(6). https://doi.org/10.1016/j.ccr.2004.05.024

Bagni, C., & Greenough, W. T. (2005). From mRNP trafficking to spine dysmorphogenesis: The roots of fragile X syndrome. In *Nature Reviews Neuroscience* (Vol. 6, Issue 5). https://doi.org/10.1038/nrn1667

Bah, A., Vernon, R. M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L. E., & Forman-Kay, J. D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature*, *519*(7541). https://doi.org/10.1038/nature13999

Bakowies, D., & Thiel, W. (1996). Hybrid models for combined quantum mechanical and molecular mechanical approaches. *Journal of Physical Chemistry*, *100*(25). https://doi.org/10.1021/jp9536514

Banko, J. L., Poulin, F., Hou, L., DeMaria, C. T., Sonenberg, N., & Klann, E. (2005). The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. *Journal of Neuroscience*, *25*(42). https://doi.org/10.1523/JNEUROSCI.2423-05.2005

Barducci, A., Bussi, G., & Parrinello, M. (2008). Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters*, *100*(2). https://doi.org/10.1103/PhysRevLett.100.020603

Bayly, C. I., Cieplak, P., Cornell, W. D., & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *Journal of Physical Chemistry*, *97*(40). https://doi.org/10.1021/j100142a004

Berkowitz, M., & McCammon, J. A. (1982). Molecular dynamics with stochastic boundary conditions. *Chemical Physics Letters*, *90*(3). https://doi.org/10.1016/0009-2614(82)80028-6

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., & Topisirovic, I. (2015). Targeting the translation machinery in cancer. In *Nature Reviews Drug Discovery* (Vol. 14, Issue 4). https://doi.org/10.1038/nrd4505

Borden, K. L., & Culjkovic, B. (2009). Understanding and targeting the eukaryotic translation initiation factor eif4E in head and neck cancer. In *Journal of Oncology*. https://doi.org/10.1155/2009/981679

Brown, C. J., Verma, C. S., Lane, D. P., & Lama, D. (2021). Conformational ordering of intrinsically disordered peptides for targeting translation initiation. *Biochimica et Biophysica Acta - General Subjects*, *1865*(1). https://doi.org/10.1016/j.bbagen.2020.129775

Buffington, S. A., Huang, W., & Costa-Mattioli, M. (2014). Translational control in synaptic plasticity and cognitive dysfunction. In *Annual Review of Neuroscience* (Vol. 37). https://doi.org/10.1146/annurev-neuro-071013-014100

Caron, E., Ghosh, S., Matsuoka, Y., Ashton-Beaucage, D., Therrien, M., Lemieux, S., Perreault, C., Roux, P. P., & Kitano, H. (2010). A comprehensive map of the mTOR signaling network. In *Molecular Systems Biology* (Vol. 6). https://doi.org/10.1038/msb.2010.108

Carroll, M., & Borden, K. L. B. (2013). The oncogene eIF4E: Using biochemical insights to target cancer. In *Journal of Interferon and Cytokine Research* (Vol. 33, Issue 5). https://doi.org/10.1089/jir.2012.0142

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. In *Journal of Computational Chemistry* (Vol. 26, Issue 16). https://doi.org/10.1002/jcc.20290

Chen, L., Aktas, B. H., Wang, Y., He, X., Sahoo, R., Zhang, N., Denoyelle, S., Kabha, E., Yang, H., Freedman, R. Y., Supko, J. G., Chorev, M., Wagner, G., & Halperin, J. A. (2012). Tumor suppression by small molecule inhibitorss of translation initiation. *Oncotarget*, *3*(8). https://doi.org/10.18632/oncotarget.598

Chu, J., Cargnello, M., Topisirovic, I., & Pelletier, J. (2016). Translation Initiation Factors: Reprogramming Protein Synthesis in Cancer. In *Trends in Cell Biology* (Vol. 26, Issue 12). https://doi.org/10.1016/j.tcb.2016.06.005

Cohen, N., Sharma, M., Kentsis, A., Perez, J. M., Strudwick, S., & Borden, K. L. B. (2001). PML RING suppresses oncogenic transformation by reducing the affinity of eIF4E for mRNA. *EMBO Journal*, *20*(16). https://doi.org/10.1093/emboj/20.16.4547

Couturier, J., Page, G., Morel, M., Gontier, C., Lecron, J. C., Pontcharraud, R., Fauconneau, B., & Paccalin, M. (2010). Inhibition of double-stranded RNA-dependent protein kinase strongly decreases cytokine production and release in peripheral blood mononuclear cells from patients with Alzheimer's disease. *Journal of Alzheimer's Disease*, *21*(4). https://doi.org/10.3233/JAD-2010-100258

Culjkovic, B., Topisirovic, I., & Borden, K. L. B. (2007). Controlling gene expression through RNA regulons: The role of the eukaryotic translation initiation factor eIF4E. In *Cell Cycle* (Vol. 6, Issue 1). https://doi.org/10.4161/cc.6.1.3688

Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W. F., & Mark, A. E. (1999). Peptide folding: When simulation meets experiment. *Angewandte Chemie - International Edition*, *38*(1–2). https://doi.org/10.1002/(sici)1521-3773(19990115)38:1/2<236::aid-anie236>3.0.co;2-m

de Benedetti, A., & Graff, J. R. (2004). eIF-4E expression and its role in malignancies and metastases. In *Oncogene* (Vol. 23, Issue 18). https://doi.org/10.1038/sj.onc.1207545

de Rubeis, S., & Bagni, C. (2011). Regulation of molecular pathways in the Fragile X Syndrome: Insights into Autism Spectrum Disorders. In *Journal of*

Neurodevelopmental Disorders (Vol. 3, Issue 3). https://doi.org/10.1007/s11689-011-9087-2

DeRubeis, S., Pasciuto, E., Li, K. W., Fernández, E., DiMarino, D., Buzzi, A., Ostroff, L. E., Klann, E., Zwartkruis, F. J. T., Komiyama, N. H., Grant, S. G. N., Poujol, C., Choquet, D., Achsel, T., Posthuma, D., Smit, A. B., & Bagni, C. (2013). CYFIP1 coordinates mRNA translation and cytoskeleton remodeling to ensure proper dendritic Spine formation. *Neuron*, *79*(6). https://doi.org/10.1016/j.neuron.2013.06.039

Descamps, G., Gomez-Bougie, P., Tamburini, J., Green, A., Bouscary, D., Maga, S., Moreau, P., le Gouill, S., Pellat-Deceunynck, C., & Amiot, M. (2012). The cap-translation inhibitor 4EGI-1 induces apoptosis in multiple myeloma through Noxa induction. *British Journal of Cancer*, *106*(10). https://doi.org/10.1038/bjc.2012.139

di Marino, D., Chillemi, G., de Rubeis, S., Tramontano, A., Achsel, T., & Bagni, C. (2015). MD and Docking Studies Reveal That the Functional Switch of CYFIP1 is Mediated by a Butterfly-like Motion. *Journal of Chemical Theory and Computation*, *11*(7). https://doi.org/10.1021/ct500431h

di Marino, D., D'Annessa, I., Tancredi, H., Bagni, C., & Gallicchio, E. (2015). A unique binding mode of the eukaryotic translation initiation factor 4E for guiding the design of novel peptide inhibitors. *Protein Science*, *24*(9). https://doi.org/10.1002/pro.2708

Donzé, O., Jagus, R., Koromilas, A. E., Hershey, J. W. B., & Sonenberg, N. (1995). Abrogation of translation initiation factor eIF-2 phosphorylation causes malignant transformation of NIH 3T3 cells. *EMBO Journal*, *14*(15). https://doi.org/10.1002/j.1460-2075.1995.tb00052.x

Dostie, J., Ferraiuolo, M., Pause, A., Adam, S. A., & Sonenberg, N. (2000). A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *EMBO Journal*, *19*(12). https://doi.org/10.1093/emboj/19.12.3142

Duffy, A. G., Makarova-Rusher, O. v., Ulahannan, S. v., Rahma, O. E., Fioravanti, S., Walker, M., Abdullah, S., Raffeld, M., Anderson, V., Abi-Jaoudeh, N., Levy, E., Wood, B. J., Lee, S., Tomita, Y., Trepel, J. B., Steinberg, S. M., Revenko, A. S., MacLeod, A. R., Peer, C. J., … Greten, T. F. (2016). Modulation of tumor eIF4E by antisense inhibition: A phase I/II translational clinical trial of ISIS 183750—an antisense oligonucleotide against eIF4E—in combination with irinotecan in solid tumors and irinotecan-refractory colorectal cancer. *International Journal of Cancer*, *139*(7). https://doi.org/10.1002/ijc.30199

Duncan, R., Milburn, S. C., & Hershey, J. W. B. (1987). Regulated phosphorylation and low abundance of HeLa cell initiation factor eIF-4F suggest a role in translational control. Heat shock effects on eIF-4F. *Journal of Biological Chemistry*, *262*(1). https://doi.org/10.1016/s0021-9258(19)75938-9

Elgendy, M., Sheridan, C., Brumatti, G., & Martin, S. J. (2011). Oncogenic Ras-Induced Expression of Noxa and Beclin-1 Promotes Autophagic Cell Death and Limits Clonogenic Survival. *Molecular Cell*, *42*(1). https://doi.org/10.1016/j.molcel.2011.02.009

Ferraiuolo, M. A., Basak, S., Dostie, J., Murray, E. L., Schoenberg, D. R., & Sonenberg, N. (2005). A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *Journal of Cell Biology*, *170*(6). https://doi.org/10.1083/jcb.200504039

Field, M. J., Bash, P. A., & Karplus, M. (1990). A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *Journal of Computational Chemistry*, *11*(6). https://doi.org/10.1002/jcc.540110605

Fischer, P. D., Papadopoulos, E., Dempersmier, J. M., Wang, Z. F., Nowak, R. P., Donovan, K. A., Kalabathula, J., Gorgulla, C., Junghanns, P. P. M., Kabha, E., Dimitrakakis, N., Petrov, O. I., Mitsiades, C., Ducho, C., Gelev, V., Fischer, E. S., Wagner, G., & Arthanari, H. (2021). A biphenyl inhibitor of eIF4E targeting an internal binding site enables the design of cell-permeable PROTAC-degraders. *European Journal of Medicinal Chemistry*, *219*. https://doi.org/10.1016/j.ejmech.2021.113435

Gao, J., Li, X., Wang, H., & Boeck, G. (2003). Empirical All Region Current Based PHEMT DC Model. *SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference Proceedings*. https://doi.org/10.1109/imoc.2003.1244839

German-Retana, S., Walter, J., Doublet, B., Roudet-Tavert, G., Nicaise, V., Lecampion, C., Houvenaghel, M.-C., Robaglia, C., Michon, T., & le Gall, O. (2008). Mutational Analysis of Plant Cap-Binding Protein eIF4E Reveals Key Amino Acids Involved in Biochemical Functions and Potyvirus Infection. *Journal of Virology*, *82*(15). https://doi.org/10.1128/jvi.00209-08

Gingras, A. C., Raught, B., & Sonenberg, N. (2001). Regulation of translation initiation by FRAP/mTOR. In *Genes and Development* (Vol. 15, Issue 7). https://doi.org/10.1101/gad.887201

Gkogkas, C. G., Khoutorsky, A., Ran, I., Rampakakis, E., Nevarko, T., Weatherill, D. B., Vasuta, C., Yee, S., Truitt, M., Dallaire, P., Major, F., Lasko, P., Ruggero, D., Nader, K., Lacaille, J. C., & Sonenberg, N. (2013). Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature*, *493*(7432). https://doi.org/10.1038/nature11628

Gogonea, V., Suárez, D., van der Vaart, A., & Merz, J. (2001). New developments in applying quantum mechanics to proteins. In *Current Opinion in Structural Biology* (Vol. 11, Issue 2). https://doi.org/10.1016/S0959-440X(00)00193-7

Gordon, M. S., Binkley, J. S., Pople, J. A., Pietro, W. J., & Hehre, W. J. (1982). Self-Consistent Molecular-Orbital Methods. 22. Small Split-Valence Basis Sets for Second-Row Elements. *Journal of the American Chemical Society*, *104*(10). https://doi.org/10.1021/ja00374a017

Graff, J. R., Konicek, B. W., Vincent, T. M., Lynch, R. L., Monteith, D., Weir, S. N., Schwier, P., Capen, A., Goode, R. L., Dowless, M. S., Chen, Y., Zhang, H., Sissons, S., Cox, K., McNulty, A. M., Parsons, S. H., Wang, T., Sams, L., Geeganage, S., … Marcusson, E. G. (2007). Therapeutic suppression of translation initiation factor eIF4E expression reduces tumor growth without toxicity. *Journal of Clinical Investigation*, *117*(9). https://doi.org/10.1172/JCI32044

Grüner, S., Peter, D., Weber, R., Wohlbold, L., Chung, M. Y., Weichenrieder, O., Valkov, E., Igreja, C., & Izaurralde, E. (2016). The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation. *Molecular Cell*, *64*(3). https://doi.org/10.1016/j.molcel.2016.09.020

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, *47*(7). https://doi.org/10.1021/jm030644s

Hay, N., & Sonenberg, N. (2004). Upstream and downstream of mTOR. In *Genes and Development* (Vol. 18, Issue 16). https://doi.org/10.1101/gad.1212704

Heesom, K. J., Gampel, A., Mellor, H., & Denton, R. M. (2001). Cell cycle-dependent phosphorylation of the translational repressor elF-4E binding protein-1 (4E-BP1). *Current Biology*, *11*(17). https://doi.org/10.1016/S0960-9822(01)00422-5

Hiremath, L. S., Webb, N. R., & Rhoads, R. E. (1985). Immunological detection of the messenger RNA cap-binding protein. *Journal of Biological Chemistry*, *260*(13). https://doi.org/10.1016/s0021-9258(17)39529-7

Hsieh, A. C., & Ruggero, D. (2010). Targeting eukaryotic translation initiation factor 4E (eIF4E) in cancer. In *Clinical Cancer Research* (Vol. 16, Issue 20). https://doi.org/10.1158/1078-0432.CCR-10-0433

Hünenberger, P. H. (2005). Thermostat algorithms for molecular dynamics simulations. In *Advances in Polymer Science* (Vol. 173). https://doi.org/10.1007/b99427

Igreja, C., Peter, D., Weiler, C., & Izaurralde, E. (2014). 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms5790

Jakalian, A., Bush, B. L., Jack, D. B., & Bayly, C. I. (2000). Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *Journal of Computational Chemistry*, *21*(2). https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2<132::AID-JCC5>3.0.CO;2-P

James, V., Zhang, Y., Foxler, D. E., de Moor, C. H., Kong, Y. W., Webb, T. M., Self, T. J., Feng, Y., Lagos, D., Chu, C. Y., Rana, T. M., Morley, S. J., Longmore, G. D., Bushell, M., & Sharp, T. v. (2010). LIM-domain proteins, LIMD1, Ajuba, and WTIP are required for microRNA-mediated gene silencing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(28). https://doi.org/10.1073/pnas.0914987107

Jin, J., Xiang, W., Wu, S., Wang, M., Xiao, M., & Deng, A. (2019). Targeting eIF4E signaling with ribavirin as a sensitizing strategy for ovarian cancer. *Biochemical and Biophysical Research Communications*, *510*(4). https://doi.org/10.1016/j.bbrc.2019.01.117

Kamenska, A., Lu, W. T., Kubacka, D., Broomhead, H., Minshall, N., Bushell, M., & Standart, N. (2014). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. *Nucleic Acids Research*, *42*(5). https://doi.org/10.1093/nar/gkt1265

Kentsis, A., Dwyer, E. C., Perez, J. M., Sharma, M., Chen, A., Pan, Z. Q., & Borden, K. L. B. (2001). The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E. *Journal of Molecular Biology*, *312*(4). https://doi.org/10.1006/jmbi.2001.5003

Kimball, S. R. (2001). Regulation of translation initiation by amino acids in eukaryotic cells. In *Progress in molecular and subcellular biology* (Vol. 26). https://doi.org/10.1007/978-3-642-56688-2_6

Kinkelin, K., Veith, K., Grünwald, M., & Bono, F. (2012). Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. *RNA*, *18*(9). https://doi.org/10.1261/rna.033639.112

Kosciuczuk, E. M., Saleiro, D., Kroczynska, B., Beauchamp, E. M., Eckerdt, F., Blyth, G. T., Abedin, S. M., Giles, F. J., Altman, J. K., & Platanias, L. C. (2016). Merestinib blocks Mnk kinase activity in acute myeloid leukemia progenitors and exhibits

antileukemic effects in vitro and in vivo. *Blood*, *128*(3). https://doi.org/10.1182/blood-2016-02-698704

Kumar, R., Afsar, M., Khandelwal, N., Chander, Y., Riyesh, T., Dedar, R. K., Gulati, B. R., Pal, Y., Barua, S., Tripathi, B. N., Hussain, T., & Kumar, N. (2021). Emetine suppresses SARS-CoV-2 replication by inhibiting interaction of viral mRNA with eIF4E. *Antiviral Research*, *189*. https://doi.org/10.1016/j.antiviral.2021.105056

Lama, D., Liberatore, A. M., Frosi, Y., Nakhle, J., Tsomaia, N., Bashir, T., Lane, D. P., Brown, C. J., Verma, C. S., & Auvin, S. (2019). Structural insights reveal a recognition feature for tailoring hydrocarbon stapled-peptides against the eukaryotic translation initiation factor 4E protein. *Chemical Science*, *10*(8). https://doi.org/10.1039/C8SC03759K

Lama, D., Pradhan, M. R., Brown, C. J., Eapen, R. S., Joseph, T. L., Kwoh, C. K., Lane, D. P., & Verma, C. S. (2017). Water-Bridge Mediates Recognition of mRNA Cap in eIF4E. *Structure*, *25*(1). https://doi.org/10.1016/j.str.2016.11.006

Lama, D., Quah, S. T., Verma, C. S., Lakshminarayanan, R., Beuerman, R. W., Lane, D. P., & Brown, C. J. (2013). Rational optimization of conformational effects induced by hydrocarbon staples in peptides and their binding interfaces. *Scientific Reports*, *3*. https://doi.org/10.1038/srep03451

Lama, D., & Verma, C. S. (2020). Deciphering the mechanistic effects of eIF4E phosphorylation on mRNA-cap recognition. *Protein Science*, *29*(6). https://doi.org/10.1002/pro.3798

Laplante, M., & Sabatini, D. M. (2012). MTOR signaling in growth control and disease. In *Cell* (Vol. 149, Issue 2). https://doi.org/10.1016/j.cell.2012.03.017

Lazaris-Karatzas, A., Montine, K. S., & Sonenberg, N. (1990). Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5′ cap. *Nature*, *345*(6275). https://doi.org/10.1038/345544a0

Lejbkowicz, F., Goyer, C., Darveau, A., Neron, S., Lemieux, R., & Sonenberg, N. (1992). A fraction of the mRNA 5′ cap-binding protein, eukaryotic initiation factor 4E, localizes to the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(20). https://doi.org/10.1073/pnas.89.20.9612

Li, S., Takasu, T., Perlman, D. M., Peterson, M. S., Burrichter, D., Avdulov, S., Bitterman, P. B., & Polunovsky, V. A. (2003). Translation factor eIF4E rescues cells from Myc-dependent apoptosis by inhibiting cytochrome c release. *Journal of Biological Chemistry*, *278*(5). https://doi.org/10.1074/jbc.M208821200

Lu, C., Makala, L., Wu, D., & Cai, Y. (2016). Targeting translation: EIF4E as an emerging anticancer drug target. In *Expert Reviews in Molecular Medicine* (Vol. 18). https://doi.org/10.1017/erm.2015.20

M. J. Frisch, G., Trucks, W., Schlegel, H. B. ;, Scuseria, G. E. ;, Robb, M. A. ;, Cheeseman, J. R., Scalmani, G., Barone, V. ;, Mennucci, B. ;, Petersson, G. A., Nakatsuji, H., Caricato, M. ;, Li, X. ;, Hratchian, H. P. ;, Izmaylov, A. F. ;, Bloino, J. ;, Zheng, G. ;, Sonnenberg, J. L. ;, Frisch, M. J., … Fox, D. J. (2009). Gaussian 09, Revision E. 01; Gaussian. *Gaussian, Inc.: Wallingford, CT*, *11*(1).

Mabuchi, S., Kuroda, H., Takahashi, R., & Sasano, T. (2015). The PI3K/AKT/mTOR pathway as a therapeutic target in ovarian cancer. In *Gynecologic Oncology* (Vol. 137, Issue 1). https://doi.org/10.1016/j.ygyno.2015.02.003

Mader, S., Lee, H., Pause, A., & Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the

translational repressors 4E-binding proteins. *Molecular and Cellular Biology*, *15*(9). https://doi.org/10.1128/mcb.15.9.4990
Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell*, *89*(6). https://doi.org/10.1016/S0092-8674(00)80280-9
Marcotrigiano, J., Gingras, A. C., Sonenberg, N., & Burley, S. K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Molecular Cell*, *3*(6). https://doi.org/10.1016/S1097-2765(01)80003-4
Martínez, A., Sesé, M., Losa, J. H., Robichaud, N., Sonenberg, N., Aasen, T., & Ramón Y Cajal, S. (2015). Phosphorylation of eIF4E confers resistance to cellular stress and DNA-damaging agents through an interaction with 4E-T: A Rationale for novel therapeutic approaches. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0123352
Merrick, W. C., & Pavitt, G. D. (2018). Protein synthesis initiation in eukaryotic cells. *Cold Spring Harbor Perspectives in Biology*, *10*(12). https://doi.org/10.1101/cshperspect.a033092
Michon, T., Estevez, Y., Walter, J., German-Retana, S., & le Gall, O. (2006). The potyviral virus genome-linked protein VPg forms a ternary complex with the eukaryotic initiation factors eIF4E and eIF4G and reduces eIF4E affinity for a mRNA cap analogue. *FEBS Journal*, *273*(6). https://doi.org/10.1111/j.1742-4658.2006.05156.x
Mizuno, A., In, Y., Fujita, Y., Abiko, F., Miyagawa, H., Kitamura, K., Tomoo, K., & Ishida, T. (2008). Importance of C-terminal flexible region of 4E-binding protein in binding with eukaryotic initiation factor 4E. *FEBS Letters*, *582*(23–24). https://doi.org/10.1016/j.febslet.2008.09.003
Modrak-Wojcik, A., Gorka, M., Niedzwiecka, K., Zdanowski, K., Zuberek, J., Niedzwiecka, A., & Stolarski, R. (2013). Eukaryotic translation initiation is controlled by cooperativity effects within ternary complexes of 4E-BP1, eIF4E, and the mRNA 5′ cap. *FEBS Letters*, *587*(24). https://doi.org/10.1016/j.febslet.2013.10.043
Moerke, N. J., Aktas, H., Chen, H., Cantel, S., Reibarkh, M. Y., Fahmy, A., Gross, J. D. D., Degterev, A., Yuan, J., Chorev, M., Halperin, J. A., & Wagner, G. (2007). Small-Molecule Inhibition of the Interaction between the Translation Initiation Factors eIF4E and eIF4G. *Cell*, *128*(2). https://doi.org/10.1016/j.cell.2006.11.046
Møller, C., & Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Physical Review*, *46*(7). https://doi.org/10.1103/PhysRev.46.618
Monzingo, A. F., Dhaliwal, S., Dutt-Chaudhuri, A., Lyon, A., Sadow, J. H., Hoffman, D. W., Robertas, J. D., & Browning, K. S. (2007). The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. *Plant Physiology*, *143*(4). https://doi.org/10.1104/pp.106.093146
Mortier, J., Rakers, C., Bermudez, M., Murgueitio, M. S., Riniker, S., & Wolber, G. (2015). The impact of molecular dynamics on drug design: Applications for the characterization of ligand-macromolecule complexes. In *Drug Discovery Today* (Vol. 20, Issue 6). https://doi.org/10.1016/j.drudis.2015.01.003
Murata, T., & Shimotohno, K. (2006). Ubiquitination and proteasome-dependent degradation of human eukaryotic translation initiation factor 4E. *Journal of Biological Chemistry*, *281*(30). https://doi.org/10.1074/jbc.M600563200
Napoli, I., Mercaldo, V., Boyl, P. P., Eleuteri, B., Zalfa, F., de Rubeis, S., di Marino, D., Mohr, E., Massimi, M., Falconi, M., Witke, W., Costa-Mattioli, M., Sonenberg, N.,

Achsel, T., & Bagni, C. (2008). The Fragile X Syndrome Protein Represses Activity-Dependent Translation through CYFIP1, a New 4E-BP. *Cell*, *134*(6). https://doi.org/10.1016/j.cell.2008.07.031

Othumpangat, S., Kashon, M., & Joseph, P. (2005). Eukaryotic translation initiation factor 4E Is a cellular target for toxicity and death due to exposure to cadmium chloride. *Journal of Biological Chemistry*, *280*(26). https://doi.org/10.1074/jbc.M414303200

Paku, K. S., Umenaga, Y., Usui, T., Fukuyo, A., Mizuno, A., In, Y., Ishida, T., & Tommo, K. (2012). A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. *Biochemical Journal*, *441*(1). https://doi.org/10.1042/BJ20101481

Papadopoulos, E., Jenni, S., Kabha, E., Takrouri, K. J., Yi, T., Salvi, N., Luna, R. E., Gavathiotis, E., Mahalingam, P., Arthanari, H., Rodriguez-Mias, R., Yefidoff-Freedman, R., Aktas, B. H., Chorev, M., Halperin, J. A., & Wagner, G. (2014). Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(31). https://doi.org/10.1073/pnas.1410250111

Parker, R., & Sheth, U. (2007). P Bodies and the Control of mRNA Translation and Degradation. In *Molecular Cell* (Vol. 25, Issue 5). https://doi.org/10.1016/j.molcel.2007.02.011

Pelletier, J., Graff, J., Ruggero, D., & Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: A critical nexus for cancer development. *Cancer Research*, *75*(2). https://doi.org/10.1158/0008-5472.CAN-14-2789

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., & Izaurralde, E. (2015). Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E. *Molecular Cell*, *57*(6). https://doi.org/10.1016/j.molcel.2015.01.017

Petersson, G. A., Bennett, A., Tensfeldt, T. G., Al-Laham, M. A., Shirley, W. A., & Mantzaris, J. (1988). A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *The Journal of Chemical Physics*, *89*(4). https://doi.org/10.1063/1.455064

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13). https://doi.org/10.1002/jcc.20084

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1). https://doi.org/10.1002/pro.3943

Poulin, F., Gingras, A. C., Olsen, H., Chevalier, S., & Sonenberg, N. (1998). 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. *Journal of Biological Chemistry*, *273*(22). https://doi.org/10.1074/jbc.273.22.14002

Querido, J. B., Sokabe, M., Kraatz, S., Gordiyenko, Y., Skehel, J. M., Fraser, C. S., & Ramakrishnan, V. (2020). Structure of a human 48S translational initiation complex. *Science*, *369*(6508). https://doi.org/10.1126/SCIENCE.ABA4904

Rahman, A. (1964). Correlations in the motion of atoms in liquid argon. *Physical Review*, *136*(2A). https://doi.org/10.1103/PhysRev.136.A405

Romagnoli, A., D'Agostino, M., Ardiccioni, C., Maracci, C., Motta, S., la Teana, A., & di Marino, D. (2021). Control of the eIF4E activity: structural insights and pharmacological implications. In *Cellular and Molecular Life Sciences* (Vol. 78, Issues 21–22). https://doi.org/10.1007/s00018-021-03938-z

Ross Buchan, J. (2014). MRNP granules Assembly, function, and connections with disease. In *RNA Biology* (Vol. 11, Issue 8). https://doi.org/10.4161/15476286.2014.972208

Rousseau, D., Kaspar, R., Rosenwald, I., Gehrke, L., & Sonenberg, N. (1996). Translation initiation of ornithine decarboxylase and nucleocytoplasmic transport of cyclin D1 mRNA are increased in cells overexpressing eukaryotic initiation factor 4E. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(3). https://doi.org/10.1073/pnas.93.3.1065

Sabatini, D. M. (2006). mTOR and cancer: Insights into a complex relationship. In *Nature Reviews Cancer* (Vol. 6, Issue 9). https://doi.org/10.1038/nrc1974

Salvi, N., Papadopoulos, E., Blackledge, M., & Wagner, G. (2016). The Role of Dynamics and Allostery in the Inhibition of the eIF4E/eIF4G Translation Initiation Factor Complex. *Angewandte Chemie - International Edition*, *55*(25). https://doi.org/10.1002/anie.201603254

Santini, E., & Klann, E. (2014). Reciprocal signaling between translational control pathways and synaptic proteins in autism spectrum disorders. In *Science Signaling* (Vol. 7, Issue 349). https://doi.org/10.1126/scisignal.2005832

Shi, F., Len, Y., Gong, Y., Shi, R., Yang, X., Naren, D., & Yan, T. (2015). Ribavirin inhibits the activity of mtor/eif4e, ERK/Mnk1/eIF4E signaling pathway and synergizes with tyrosine kinase inhibitor imatinib to impair Bcr-Abl mediated proliferation and apoptosis in Ph+ leukemia. *PLoS ONE*, *10*(8). https://doi.org/10.1371/journal.pone.0136746

Shih, J. W., Wang, W. T., Tsai, T. Y., Kuo, C. Y., Li, H. K., & Wu Lee, Y. H. (2012). Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochemical Journal*, *441*(1). https://doi.org/10.1042/BJ20110739

Siddiqui, N., & Sonenberg, N. (2015). Signalling to eIF4E in cancer. *Biochemical Society Transactions*, *43*. https://doi.org/10.1042/BST20150126

Sonenberg, N., Rupprecht, K. M., Hecht, S. M., & Shatkin, A. J. (1979). Eukaryotic mRNA cap binding protein: purification by affinity chromatography on sepharose-coupled m7GDP. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(9). https://doi.org/10.1073/pnas.76.9.4345

Song, Y. K., Guo, H., Barengo, N., & Naora, H. (2009). Inhibition of ovarian cancer growth by a tumor-targeting peptide that binds eukaryotic translation initiation factor 4E. *Clinical Cancer Research*, *15*(13). https://doi.org/10.1158/1078-0432.CCR-08-2924

Tamburini, J., Green, A. S., Bardet, V., Chapuis, N., Park, S., Willems, L., Uzunov, M., Ifrah, N., Dreyfus, F., Lacombe, C., Mayeux, P., & Bouscary, D. (2009). Protein synthesis is resistant to rapamycin and constitutes a promising therapeutic target in acute myeloid leukemia. *Blood*, *114*(8). https://doi.org/10.1182/blood-2008-10-184515

Tomoo, K., Abiko, F., Miyagawa, H., Kitamura, K., & Ishida, T. (2006). Effect of N-terminal region of eIF4E and Ser65-phosphorylation of 4E-BP1 on interaction

between eIF4E and 4E-BP1 fragment peptide. *Journal of Biochemistry*, *140*(2). https://doi.org/10.1093/jb/mvj143

Tomoo, K., Matsushita, Y., Fujisaki, H., Abiko, F., Shen, X., Taniguchi, T., Miyagawa, H., Kitamura, K., Miura, K. I., & Ishida, T. (2005). Structural basis for mRNA Cap-Binding regulation of eukaryotic initiation factor 4E by 4E-binding protein, studied by spectroscopic, X-ray crystal structural, and molecular dynamics simulation methods. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1753*(2). https://doi.org/10.1016/j.bbapap.2005.07.023

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. I. (2002). Crystal structures of 7-methylguanosine 5′-triphosphate (m7GTP)- and P1-7-methylguanosine-P3-adenosine-5′, 5′-triphosphate (m7GpppA)-bound human full-length eukaryotic initiation factor 4E: Biological importance of the C-terminal flexible region. *Biochemical Journal*, *362*(3). https://doi.org/10.1042/0264-6021:3620539

Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Sasaki, M., Taniguchi, T., Miyagawa, H., Kitamura, K., Miura, K. I., & Ishida, T. (2003). Structural features of human initiation factor 4E, studied by X-ray crystal analyses and molecular dynamics simulations. *Journal of Molecular Biology*, *328*(2). https://doi.org/10.1016/S0022-2836(03)00314-0

Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., & Bussi, G. (2014). PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, *185*(2). https://doi.org/10.1016/j.cpc.2013.09.018

Tschopp, C., Knauf, U., Brauchle, M., Zurini, M., Ramage, P., Glueck, D., New, L., Han, J., & Gram, H. (2000). Phosphorylation of eIF-4E on Ser 209 in response to mitogenic and inflammatory stimuli is faithfully detected by specific antibodies. *Molecular Cell Biology Research Communications*, *3*(4). https://doi.org/10.1006/mcbr.2000.0217

Tsukiyama-Kohara, K., Poulin, F., Kohara, M., DeMaria, C. T., Cheng, A., Wu, Z., Gingras, A. C., Katsume, A., Elchebly, M., Spiegelman, B. M., Harper, M. E., Tremblay, M. L., & Sonenberg, N. (2001). Adipose tissue reduction in mice lacking the translational inhibitor 4E-BP1. *Nature Medicine*, *7*(10). https://doi.org/10.1038/nm1001-1128

Volpon, L., Osborne, M. J., Capul, A. A., de La Torre, J. C., & Borden, K. L. B. (2010). Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(12). https://doi.org/10.1073/pnas.0909877107

Wagner, J. R., Sørensen, J., Hensley, N., Wong, C., Zhu, C., Perison, T., & Amaro, R. E. (2017). POVME 3.0: Software for Mapping Binding Pocket Flexibility. *Journal of Chemical Theory and Computation*, *13*(9). https://doi.org/10.1021/acs.jctc.7b00500

Wan, J., Shi, F., Xu, Z., & Zhao, M. (2015). Knockdown of eIF4E suppresses cell proliferation, invasion and enhances cisplatin cytotoxicity in human ovarian cancer cells. *International Journal of Oncology*, *47*(6). https://doi.org/10.3892/ijo.2015.3201

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general Amber force field. *Journal of Computational Chemistry*, *25*(9). https://doi.org/10.1002/jcc.20035

Wang, X., Beugnet, A., Murakami, M., Yamanaka, S., & Proud, C. G. (2005). Distinct Signaling Events Downstream of mTOR Cooperate To Mediate the Effects of Amino Acids and Insulin on Initiation Factor 4E-Binding Proteins. *Molecular and Cellular Biology*, *25*(7). https://doi.org/10.1128/mcb.25.7.2558-2572.2005

Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, *103*(2). https://doi.org/10.1016/0022-2836(76)90311-9

Wiebe, S., Nagpal, A., Truong, V. T., Park, J., Skalecka, A., He, A. J., Gamache, K., Khoutorsky, A., Gantois, I., & Sonenberg, N. (2019). Inhibitory interneurons mediate autism-associated behaviors via 4E-BP2. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(36). https://doi.org/10.1073/pnas.1908126116

Williams, T., & Kelley, C. (2010). Gnuplot 4.4: an interactive plotting program. ... *Documentation, Http://Sourceforge. Net/Projects/Gnuplot*, *C*.

Willimott, S., Beck, D., Ahearne, M. J., Adams, V. C., & Wagner, S. D. (2013). Cap-translation inhibitor, 4EGI-1, restores sensitivity to ABT-737 apoptosis through cap-dependent and -independent mechanisms in chronic lymphocytic leukemia. *Clinical Cancer Research*, *19*(12). https://doi.org/10.1158/1078-0432.CCR-12-2185

Yammine, A., Gao, J., & Kwan, A. (2019). Tryptophan Fluorescence Quenching Assays for Measuring Protein-ligand Binding Affinities: Principles and a Practical Guide. *BIO-PROTOCOL*, *9*(11). https://doi.org/10.21769/bioprotoc.3253

Young, P. (2009). The leapfrog method and other "symplectic" algorithms for integrating Newtons's laws of motion. *Lecture Notes in University of California, Santa Cruz*, *1*.

Zalfa, F., Eleuteri, B., Dickson, K. S., Mercaldo, V., de Rubeis, S., di Penta, A., Tabolacci, E., Chiurazzi, P., Neri, G., Grant, S. G. N., & Bagni, C. (2007). A new function for the fragile X mental retardation protein in regulation of PSD-95 mRNA stability. *Nature Neuroscience*, *10*(5). https://doi.org/10.1038/nn1893

Zhang, M., Wang, Q., & Huang, Y. (2007). Fragile X mental retardation protein FMRP and the RNA export factor NXF2 associate with and destabilize Nxf1 mRNA in neuronal cells. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(24). https://doi.org/10.1073/pnas.0700169104

Ringraziamenti

Vorrei ringraziare tutte le persone che mi sono state vicino, mi scuso in anticipo di non aver citato tutti quelli che lo avrebbero meritato, ma solo i ringraziamenti sarebbero diventati un libro a parte.

Ringrazio per primo Pier, il mio fidanzato, che mi ha supportato e a volte anche sopportato, per tutto il mio percorso universitario, con tutti gli alti e bassi, le gioie e i pianti che ho avuto in questi anni. Un supporto che è stato sia pratico per organizzarmi a fare tutto, ma soprattutto psicologico, impedendo con la sua vicinanza che fossi sopraffatta dalle situazioni.

Ringrazio tutta la mia famiglia, per il sostegno che mi hanno sempre dato, per la fiducia che hanno dimostrato di riporre in me, per non aver dato mai peso ai voti, sottolineando ancora di più come non fosse il voto la cosa importante ma tutto quello che c'era dietro.

Ringrazio tutti i miei amici, quelli con cui ho condiviso il mio percorso dalla triennale, in particolare Fiorenza, Valentina e Riccardo, con i quali ho sperimentato tante emozioni, la voglia di mollare tutto per poi spronarci a vicenda per arrivare in fondo.

Ma anche gli amici che si sono aggiunti alla magistrale: Letizia, Camilla, Ilaria e Lorenzo, con i quali nonostante questi anni di pandemia abbiamo legato molto, arrivando anche a ripetere in videochiamata alle undici di sera per arrivare tutti pronti all'esame.

Ringrazio tutti i miei compagni di laboratorio: sperimentali e computazionali, per creare un clima bellissimo e aver sempre il tempo di un sorriso, una battuta o un consiglio.

Un ringraziamento speciale va a Cristina, che è stata un'amica e una mentore in questi ultimi mesi, sempre disponibile per un confronto, criticando senza mai giudicare il mio operato insegnandomi quel poco che so di tecniche di laboratorio durante il tirocinio. La ringrazio soprattutto perché ha avuto sempre una parola di conforto nelle mie giornate più buie.

Ringrazio inoltre il mio correlatore il Prof. Di Leva per avermi insegnato veramente tanto, con la sua precisione e pazienza nello spiegare tutto passo per passo. Ringrazio anche Vincenzo, il suo dottorando con cui ho avuto il piacere di collaborare per la tesura della tesi.

Per ultimo, ma non per importanza, vorrei ringraziare il mio relatore: il professore Daniele Di Marino. Lo ringrazio per aver voluto dedicarmi il tempo che serviva per insegnarmi tutto da zero, sapendo che stavo studiando nutrizione e dovevamo partire dalle basi. Lo ringrazio per il rapporto che abbiamo creato in questi mesi, di avermi spronato quando era necessario e rassicurata quando ho avuto dei momenti di crisi. Infine, lo ringrazio per aver condiviso con me sempre le sue esperienze, facendomi crescere non solo professionalmente, ma anche come persona.