Conformational ordering of intrinsically disordered peptides for targeting translation initiation

Christopher J. Brown [a],*, Chandra S. Verma [b,c,d], David P. Lane [a,e], Dilraj Lama [e,*]

[a] p53 Laboratory, A*STAR (Agency for Science, Technology and Research), 8A Biomedical Grove, #06-04/05, Neuros/Immunos, 138648, Singapore
[b] Bioinformatics Institute, A*STAR (Agency for Science, Technology and Research), 30 Biopolis Street, #07-01 Matrix, 138671, Singapore
[c] Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, 117543, Singapore
[d] School of Biological Sciences, Nanyang Technological University, 50 Nanyang Drive, 637551, Singapore
[e] Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Biomedicum Quarter 7B-C Solnavägen 9, 17165 Solna, Sweden

ABSTRACT

Background: Intrinsically disordered regions (IDRs) in proteins can regulate their activity by facilitating protein-protein interactions (PPIs) as exemplified in the recruitment of the eukaryotic translation initiation factor 4E (eIF4E) protein by the protein eIF4G. Deregulation of this PPI module is central to a broad spectrum of cancer related malignancies and its targeted inhibition through bioactive peptides is a promising strategy for therapeutic intervention.
Methods: We employed molecular dynamics simulations coupled with biophysical assays to rationally develop peptide derivatives from the intrinsically disordered eIF4G scaffold by incorporating non-natural amino acids that facilitates disorder-to-order transition.
Results: The conformational heterogeneity of these peptides and the degree of structural reorganization required to adopt the optimum mode of interaction with eIF4E underscores their differential binding affinities. The presence of a pre-structured local helical element in the ensemble of structures was instrumental in the efficient docking of the peptides on to the protein surface. The formation of Y4: P38 hydrogen-bond interaction between the peptide and eIF4E is a rate limiting event in the efficient recognition of the protein since it occurs through the disordered region of the peptide.
Conclusions: These insights were exploited to further design features into the peptide to propagate bound-state conformations in solution which resulted in the generation of a potent eIF4E binder.
General significance: The study illustrates the molecular basis of eIF4E recognition by a disordered epitope from eIF4G and its modulation to generate peptides that can potentially attenuate translation initiation in oncology.

1. Introduction

A significant population of eukaryotic proteins are found to contain continuous stretches of amino acids (> 30) that lack a well-defined tertiary structure and are referred to as intrinsically disordered regions (IDRs) [1,2]. They exist in a dynamic ensemble of interconverting conformational states which enables them to interact with a range of binding partners [3] or serve as a scaffold for the association of multiple proteins [4]. The eukaryotic translation initiation factor 4G (eIF4G) is a large adaptor protein with disordered segments that acts as a framework in coordinating the assembly of initiation factors and small ribosomal subunits to initiate the process of protein synthesis [5]. A central component in this molecular organization is the heterotrimeric "eIF4F"

complex formed between eIF4G, the mRNA cap-binding protein eIF4E and the helicase enzyme eIF4A [6]. This complex formation is the rate-limiting step in the regulation of cap-dependent mRNA translation and its aberrant activity is associated with the progression of numerous cancer associated pathologies [7,8]. Hence, multiple approaches are being undertaken to target different molecular aspects of eIF4F biology and inhibiting protein-protein interaction (PPI) between the constituent members and is an attractive avenue for the development of comprehensive anti-cancer therapeutics [7,8].

The IDR region of eIF4G interacts with eIF4E via a conserved canonical "Tyr-X4-Leu-Ø" (X is any amino acid and Ø is any hydrophobic residue) motif [9,10]. We have previously characterized the interaction of a 12mer eIF4G[D5S] peptide with eIF4E which revealed that the

recognition motif adopted the canonical "reverse L-shaped conformation" in the bound state (Fig. 1A) [11]. It has a disordered N-terminal (residue 1–5) organized in an almost orthogonal orientation to the helical C-terminal (residue 6–12). The peptide forms specific hydrophilic interactions with eIF4E which includes hydrogen-bonds between Y4: P38 and L9: W73 residue pairs along with salt-bridge interaction between R6 and E132. The side-chains of L9 and L10 constitute the hydrophobic component of binding by specifically docking and interacting with the hydrophobic residues present at the interface (Fig. 1A). Residue S5 (originally D in native eIF4G) present at the junction between the disordered and helical regions in the peptide was shown to be important as a helix capping residue [11]. This structural information provides an important end-state configuration of eIF4E: eIF4G PPI for deriving peptide-based inhibitors against this system [12,13].

The conformational diversity of IDRs is favourable for its regulatory role, but high flexibility is not desirable in medicinal chemistry applications as only a subset of the peptide conformations will be biologically active in efficiently recognizing the specific target. Tailoring of peptide sequences for enhanced structural stability can be achieved by

incorporation of conformationally-restricted non-natural amino acid derivatives [14]. In particular, substitution at the α-carbon position of the amino acid backbone is one of the most interesting and promising approaches to develop such surrogate building blocks [15]. These chemically modified residues have distinct stereochemically allowed backbone conformations with preference for specific secondary structure configurations [16,17]. α-aminoisobutyric acid (Aib) is one such non-coded amino acid derivative whose backbone structure in experimentally determined structures is largely restricted to the $3_{10}/\alpha$-helical geometry [16,17]. It has been shown that introduction of Aib residue into a peptide sequence can facilitate increases in its helicity [18,19]; and that it can be employed as a general strategy to design a helical backbone scaffold for targeting PPIs [20].

Herein, we examined the rational incorporation of Aib on the conformational properties of a peptide segment derived from the intrinsically disordered region of eIF4G and its impact on the binding affinity for eIF4E. A systematic exploration and optimization scheme was undertaken through insights from atomistic simulations which led to the development of a Aib-based peptide derivative that can
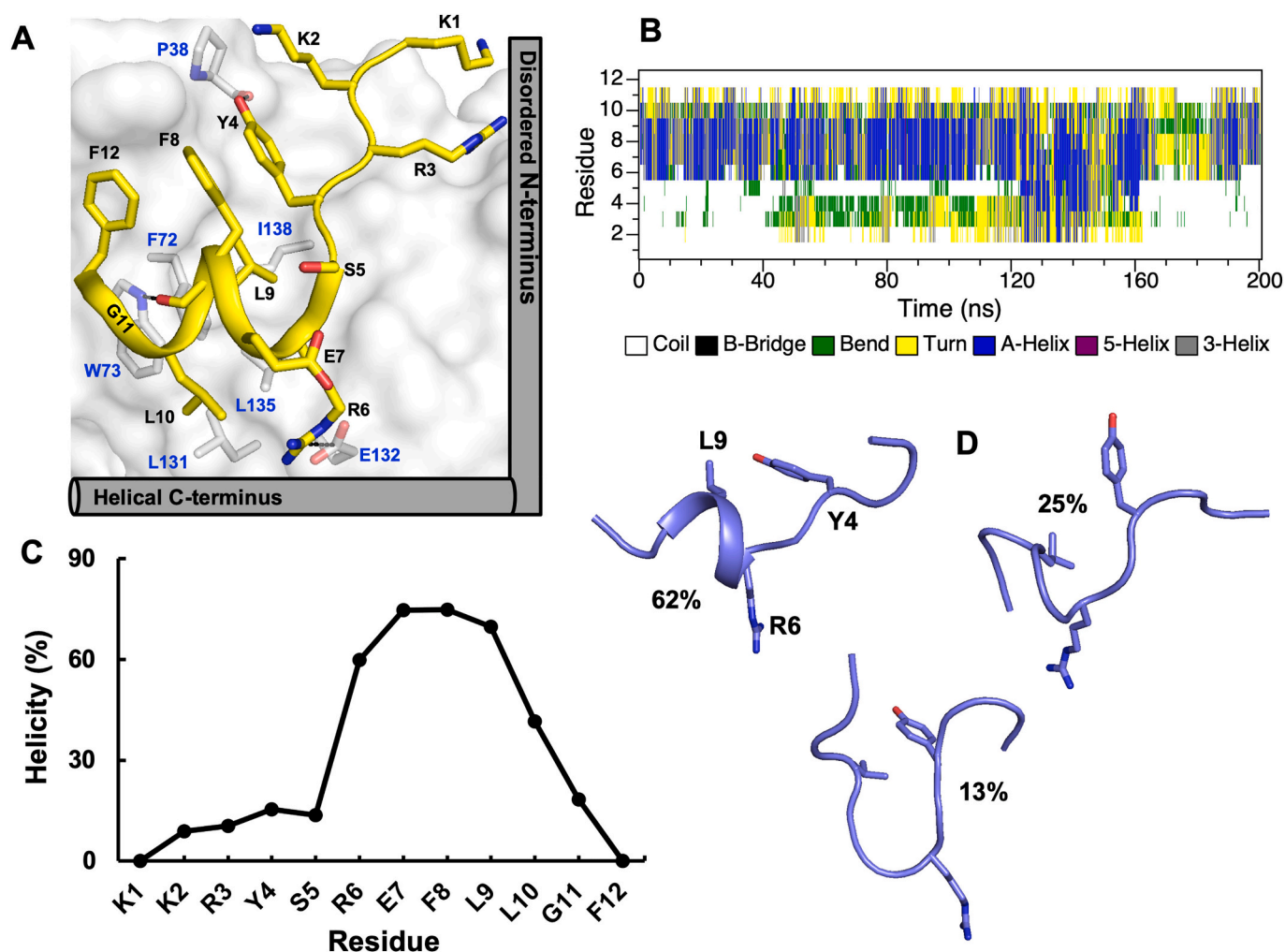
Fig. 1. eIF4G[D5S]**peptide. (A)** Crystal structure of 12mer eIF4G[D5S] peptide in complex with eIF4E (PDB ID: 4AZA). The peptide backbone is shown in cartoon and side-chain in stick representation. The protein is depicted in surface and the residues involved in inter-molecular interactions with the peptide are shown as sticks. Hydrogen-bond interactions between Y4: P38 and L9: W73 along with the R6: E132 salt-bridge interaction are explicitly indicated with dashed lines. The "reverse L-shaped conformation" of the bound peptide is highlighted. **(B)** Secondary structure evolution of eIF4G[D5S] peptide as a function of the simulation time analyzed using the DSSP program [34]. **(C)** Percentage helicity of eIF4G[D5S] peptide residues computed using "secstruct" command from the ptraj module of AMBER 18. The reported helicity is the summation of "3_{10} helix" and "α-helix" values of the individual residues. **(D)** Representative structures from three different clusters of the solution state conformations of eIF4G[D5S] peptide. The percentage of structures in each cluster is indicated. The main-chain heavy atoms of the peptide was used for clustering the structures and it was performed using average-linkage algorithm [41] with pairwise RMSD as a distance matrix. All the molecular graphics figures were created using PyMOL molecular visualization software (Schrödinger).

specifically target the PPI between eIF4G and eIF4E for inhibition.

2. Materials and methods

2.1. Peptide synthesis

The eIF4G^{D5S} peptide and Aib substituted peptide derivatives (TIPs) were ordered from and chemically synthesized by Mimotopes, Clayton, Australia. All the peptides were acetylated at the N-terminus and amidated at the C-terminus. They were purified using HPLC to more than 90% purity.

2.2. eIF4E protein expression and purification

The full length human eIF4E protein was cloned, expressed and purified as described previously [21].

2.3. Circular dichroism spectroscopy

CD experiments were performed on a JASCO J-810 spectropolarimeter and spectra were recorded with a quartz cuvette (Helmer) at a pathlength of either 0.01 cm or 0.1 cm. Peptides were dissolved in 5 mM sodium phosphate buffer (pH 7.0) at a concentration of either 2 mg/ml or 0.5 mg/ml and FUV CD spectra were recorded from 260 nm to 200 nm respectively. A data pitch of 0.2 nm at 50 nm min^{-1}, response time of 2 s and a bandwidth of 2 nm were used to register the CD signal, which was then converted into mean residue ellipticity.

2.4. Surface plasmon resonance

Human recombinant eIF4E protein was immobilized on a CM5 sensor chip through amine coupling. The chip was first conditioned with three separate 6 s injections of 100 mM HCL, 0.1% SDS and 50 mM NaOH at a flow rate of 100 µl/min. The surface of the sensor chip was then activated with a 1:1 mixture of NHS (115 mg/ml) and EDC (750 mg/ml) for 7 min at 10 µl/min. Purified eIF4E was diluted with 10 mM NaAc buffer (pH 5.0) to a final concentration of 0.5 µM with m^{7}GTP present in a 2:1 ratio to stabilize eIF4E. The immobilization level of eIF4E was ~1000 RU and a 7 min injection (at 10 µl/min) of 1 M ethanolamine (pH 8.5) was used to block excess active coupling sites. The system was primed with an assay running buffer which consisted of 10 mM HEPES, 0.15 M NaCl, 1 mM DTT and 0.1% surfactant Tween-20. The peptides were prepared by first dissolving them in 100% DMSO to a concentration of 10 mM for stock peptide solutions. The working concentration of the peptides were reached with further dilution of the stock peptide solutions into running buffer with 3% DMSO. The instrument was fully equilibrated with six buffer blanks and adding solvent correction followed by a further two blank injections. The solvent correction curve was setup by adding varying amount of 100% DMSO to 1.03× running buffer to generate a range of DMSO solutions (3.8%, 3.6%, 3.4%, 3.2%, 3%, 2.85%, 2.7% and 2.5% respectively). The peptides were injected for 60 s at a flow rate of 50 µl/min and dissociation was monitored for 180 s. SPR experiments were performed on a Biacore T100 machine and the K_ds were calculated kinetically from the dissociation and association phase data for each peptide. The kinetic data were fitted to 1:1 binding models and each individual peptide K_d was determined from at-least four separate titrations.

2.5. Fluorescence polarization

Purified full length human eIF4E at a concentration of 200 nM was titrated against 50 nM FAM labelled tracer peptide (KKRYSRDFLLALQK-(FAM)). The dissociation constant from the titration was determined by fitting the experimental data to a 1:1 binding model equation shown below [22].

$$r = r_0 + (r_b - r_0)\frac{\left(K_d + [L]_t + [P]_t\right) - \sqrt{\left(K_d + [L]_t + [P]_t\right)^2 - 4[L]_t[P]_t}}{2[L]_t}$$

where $[P]_t$ is the total eIF4E concentration, $[L]_t$ is the total FAM labelled peptide concentration, r is the anisotropy measured, r_0 is the anisotropy of the free peptide, r_b is the anisotropy of the eIF4E: FAM-labelled peptide complex and K_d is the dissociation constant of the tracer peptide which was determined to be 50.3 nM. This affinity value was then used to determine the K_d in subsequent competition assays for the respective competing TIP peptides by titrating them against the complex of eIF4E: FAM-labelled peptide and subsequently fitting the experimental data to the equations as shown below [22,23].

$$r = r_0 + (r_b - r_0)\frac{2\sqrt{\left(d^2 - 3e\right)}\cos(\theta/3) - d}{3K_{d1} + 2\sqrt{\left(d^2 - 3e\right)}\cos(\theta/3) - d}$$

$$d = K_{d1} + K_{d2} + [L]_{st} + [L]_t - [P]_t$$

$$e = \left([L]_t - [P]_t\right)K_{d1} + \left([L]_{st} - [P]_t\right)K_{d2} + K_{d1}K_{d2}$$

$$f = -K_{d1}K_{d2}[P]_t$$

$$\theta = arccos\left[\frac{-2d^3 + 9de - 27f}{2\sqrt{\left(d^2 - 3e\right)^3}}\right]$$

where $[L]_{st}$ is the total unlabelled peptide input concentration, K_{d1} is the dissociation constant for the labelled peptide (50.3 nM) as measured above and K_{d2} is the dissociation constant of the interaction between the unlabelled peptide and eIF4E. In all competitive types of experiments, it is assumed that $[P]_t > [L]_{st}$, otherwise considerable amounts of free labelled ligand would always be present and would interfere with measurements. Readings were performed with a Envision Multi-label Reader (PerkinElmer). The experiments were carried out in PBS (2.7 mM KCl, 137 mM NaCl, 10 mM Na$_2$HPO$_4$ and 2 mM KH$_2$PO$_4$ (pH 7.4)) and 0.1% Tween 20 buffer. All titrations were done in triplicates and curve-fitting was executed using the GraphPad Prism 4.0 software.

2.6. Modeling and simulations

The atomic coordinates of the non-natural amino acid Aib was modelled using the XLEAP module of AMBER 18 [24] and its RESP (Restrained Electrostatic Potential) based-partial charges was obtained through the R.E.D. server [25]. Other force-field parameters were derived from all-atom ff14SB [26] force field in AMBER18. The extended conformation of the peptides with N-terminus acetylated and C-terminus amidated were generated using the TLEAP module of AMBER 18. They were solvated with TIP3P [27] water model in a truncated octahedron box with a minimum distance of at-least 8 Å between any peptide atom and the edge of the box. The net charge of the different systems was neutralized by adding appropriate counterions. These systems were then energy minimized, heated to 300 K and equilibrated for 200 ps. The production dynamics was executed for 200 ns each for all the peptides. The generalized SGLD method (SGLDg) [28,29] was used to sample the canonical ensemble by setting the target guiding temperature to 300 K. The local averaging time for the guiding force calculation was set to 0.2 ps, force guiding factor used to scale down low frequency energy surface was set to −0.1 and the momentum guiding factor which defines the strength of the guiding effect was set to 0.5 ps^{-1}.

The crystal structure of eIF4E with eIF4G^{D5S} peptide (PDB ID: 4AZA) was used to model the bound state complexes of all the TIP peptides. The N-terminus of eIF4E was acetylated while the C-terminus was uncapped and terminated with COO$^-$ carboxyl group. The terminal ends of the

peptides were capped as in their free state simulations. The complexes were solvated with TIP3P water model inside a cuboid box with a minimum distance of at-least 10 Å between any solute atom and the edge of the box. The net charge of the different systems was neutralized by adding appropriate counterions. These systems were then energy minimized, heated to 300 K and equilibrated for 500 ps. Production dynamics was carried out for a period of 1 μs each for all the complexes under isothermal-isobaric ensemble. The simulation temperature was maintained at 300 K through Langevin Dynamics [30] by using a collision frequency of 1 ps^{-1} and the pressure was set to 1 atm using weak-coupling [31] with a relaxation time of 1 ps. Periodic boundary conditions were applied to all the systems and long-range electrostatic interactions were calculated using Particle Mesh Ewald (PME) method [32]. Bonds involving hydrogen atoms were constrained using SHAKE [33] and the integration time step was set to 2 fs.

2.7. Analysis of simulation trajectories

2.7.1. Intra- and inter-state deviations of peptide conformations

A total of 500 distinct peptide structures each from the free and bound state simulations were extracted at equal intervals over the entire range of the simulation trajectories. Pairwise Root Mean Square Deviations (RMSD) within (intra: free and bound) and between (inter: free vs bound) these peptide conformations were computed to create three separate (bound, free and inter) 2D RMSD matrices. The main-chain heavy atoms of the peptide residues 4–12 was used for the analysis. Subsequently, these data points from the matrices were organized into separate bins of 1 Å to compute the frequency/population distribution along the range of RMSD, which was found to be concentrated between 0 and 6 Å. A scatter plot of this distribution connected by smooth fitted lines was then created to generate the final graphical representations of the intra- and inter-state deviations of peptide conformations.

2.7.2. Helical propensity of peptide residues

A total of 1000 distinct peptide structures were extracted at equal intervals over the simulation period and the Dictionary of Secondary Structure of Protein (DSSP) algorithm from Kabsch and Sander [34] was used to analyse their secondary structural propensities. This analysis was done through the "secstruct" command from the ptraj module of AMBER 18. The algorithm primarily predicts the secondary structure based on the different configurations of intramolecular hydrogen-bonds formed between the backbone carbonyl and amide groups. The configurations are defined by a distance criteria between these functional groups when their electrostatic interaction energy is below a threshold value. Individual residues are then assigned specific secondary structures. A summation of these assignments over 1000 peptide structures was done to identify the propensity of each residue to adopt different secondary structures. The reported helical propensity for individual residues is an addition of α-helix and 3^{10} helix assignment for the residues.

3. Results

3.1. Conformation of eIF4G^{D5S} peptide in solution

The conformational states of eIF4G^{D5S} peptide in explicit solvent starting from an extended structure were explored using Self-guided Langevin dynamics (SGLD) [28,29]. It is an improved sampling method which dramatically accelerates the conformational searching efficiency in biomolecular simulations and has been successfully employed to study a wide range of rare biological events within accessible simulation time scales [28,29]. Secondary structure analysis of the peptide from the generated ensemble of structures showed that it sampled a largely random conformation (Fig. 1B). This is in agreement with our earlier Circular Dichroism (CD) experiment on the same peptide where it was observed to be unstructured in solution [12]. However,

there was an intermittent occurrence of helical structure primarily spanning residues 6–9 during the simulation (Fig. 1B). This spans a turn region and so is likely too short and transient to be efficiently detectable from CD spectra. The average helical propensity of individual residues in the peptide showed that R6, E7, F8 and L9 adopt a helical backbone geometry in more than 50% of the structures (Fig. 1C). Structure-based clustering of the conformations further indicated the presence of a metastable group (~60%) that included this helical turn (Fig. 1D). Collectively, these observations suggest that there is a short local structural element with helical backbone in the free state of the eIF4G^{D5S} peptide that undergoes a transition to a higher order helix upon complexation with eIF4E as observed in the crystal structure (Fig. 1A).

3.2. TIP peptides and their binding affinities for eIF4E

We next wanted to investigate if incorporation of Aib, which is considered as a helix promoting amino acid derivative [16,17], can influence the conformations of the peptide and their subsequent binding to eIF4E. The residues in the helical segment of the bound eIF4G^{D5S} peptide that can be potentially substituted without impacting the specificity of its interaction with eIF4E are D7, F8, G11 and F12 (Fig. 1A). These residues are separated by a helical turn in the peptide sequence. We designed and chemically synthesized three dual Aib substituted derivatives by incorporating the unnatural amino acid at either i, i + 4 (D7: G11 and F8: F12) or i, i + 3 (F8: G11) positions (Table 1). Their binding affinities for eIF4E was then measured using two orthogonal binding assays, Surface Plasmon Resonance (SPR) and Fluorescence Polarization (FP), the latter of which is carried out in a competitive binding mode by titrating against a fluorescent tagged tracer peptide (See methods section; Table 1 and Figs. 2A, S1 and S2). TIP-01 peptide was derived by substituting Aib in the solvent exposed D7 and G11 residue positions. The binding affinity of the peptide showed ~1/4-fold improvement (from SPR and FP assays respectively) over the parent peptide (Table 1). The second peptide derivative "TIP-02" had Aib incorporated at residue positions F8 and F12. Interestingly, this peptide displayed an almost 80-fold decrease in binding affinity (Table 1) for eIF4E in SPR and failed to effectively displace the tracer peptide in the competitive FP assay. The TIP-03 derivative with Aib substituted at residue positions F8 and G11 was an intermediate design between the previous two peptides. It had a significant improvement in binding affinity over TIP-02 but it was still about a fold lower than the parent peptide (Table 1). TIP-01 was therefore the only peptide that had a relatively better binding affinity compared to eIF4G^{D5S} whereas TIP-02 and 03 design were detrimental to eIF4E binding to different degrees.

### 3.3. Conformations of TIP peptides in solution

The secondary structure propensities of all the three Aib derived TIP peptides were examined using CD spectroscopy (Fig. 2B). The CD spectra

Table 1
Peptides and their binding affinities[a].

Peptide[a]	Sequence[b]	SPR (K_d, nM)[c]	FP (K_d, nM)[d]
eIF4G^{D5S}	Ac-KKRYSREFLLGF-NH$_2$	99.9 ± 6.2	195.2 ± 12.1
TIP-01	Ac-KKRYSRXFLLXF-NH$_2$	72.4 ± 15.9	52.4 ± 6.4
TIP-02[e]	Ac-KKRYSREXLLGX-NH$_2$	8012.3 ± 2452.3	NA
TIP-03	Ac-KKRYSREXLLXF-NH$_2$	133.0 ± 9.9	288.8 ± 9.9
TIP-04	Ac-KKRYSRXQLLXL-NH$_2$	14.9 ± 2.4	60.01 ± 3.3

[a]The binding affinity data for eIF4G^{D5S} peptide has also been previously reported [12]. [b]The sequence of the peptides with an acetylated (Ac) N-terminus and an amidated (NH$_2$) C-terminus. X = α-aminoisobutyric acid. [c,d]Surface Plasmon Resonance (SPR) and Fluorescence Polarization (FP) assays. The dissociation constant (K_d) values (Ave ± SEM) for both assays were derived from at-least three replicates. [e]The K_d of TIP-02 was not calculated from FP assay as it failed to displace the FAM labelled tracer peptide in the competitive binding assay (See method section for details).
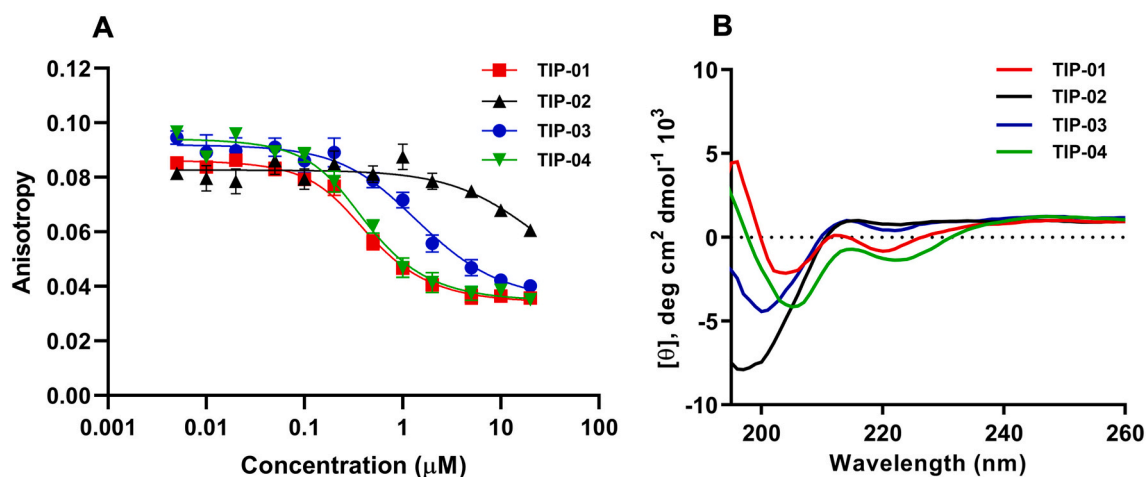
Fig. 2. **Secondary structure and eIF4E binding properties of Aib peptide derivatives.** **(A)** Competitive eIF4E binding anisotropy titration curve of TIP peptides from FP assay. **(B)** CD spectra of TIP peptides.

showed that TIP-01 is relatively more helical as compared to TIP-02 and TIP-03, which are observed to be unstructured in solution. We further investigated the conformational states of the three TIP peptides in solution using SGLD based sampling [28,29] starting from an extended state as in the case of the parent eIF4G^{D5S} peptide. TIP-01 sampled random conformations during the initial stages of the simulation but then evolved to exhibit a helical epitope from residues E7-L10 towards the C-terminal (Fig. 3A and B). The ensemble of peptide structures were clustered into distinct groups based on the presence (~60%) or absence

(~40%) of the helical turn (Fig. 4A). TIP-02 unlike TIP-01, sampled random state conformation during the entire simulation (Fig. 3C). There were helical signatures in the trajectory but they were not significant enough to produce at-least a 3- residue stretch with helical propensity above 50% (Fig. 3D). The intrinsically disordered nature of the peptide resulted in uniformly populated (~30%) clusters in solution (Fig. 4B). TIP-03 also sampled predominantly random conformations interspaced with transient helical states (Fig. 3E) localized in a short three residue stretch (X8-L10; Fig. 3F). The diversity in its structural configurations is
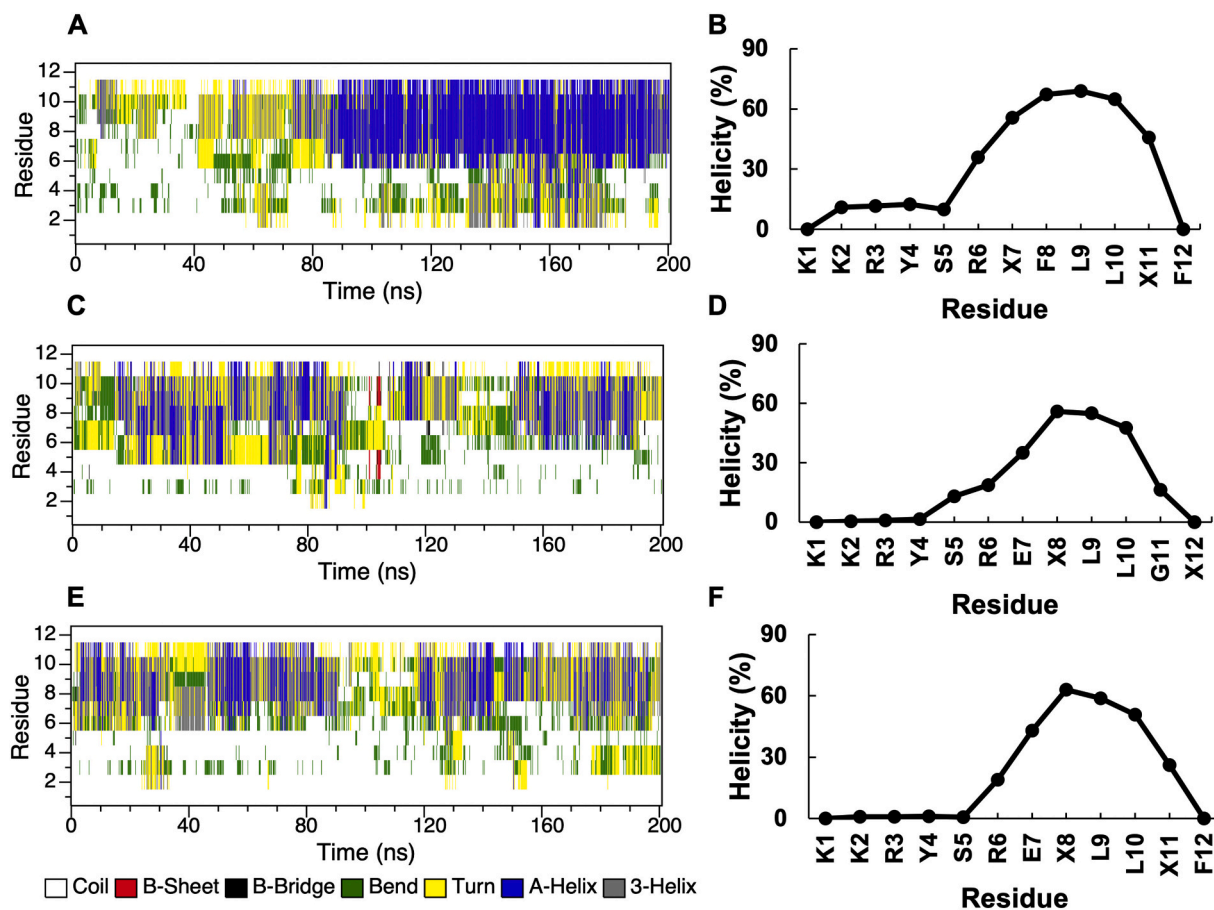
Fig. 3. **Secondary structural property of Aib peptide derivatives from MD simulations.** Time course secondary structure evolution and percentage helicity of **(A, B)** TIP-01, **(C, D)** TIP-02 and **(D, E)** TIP-03 peptides in solution.
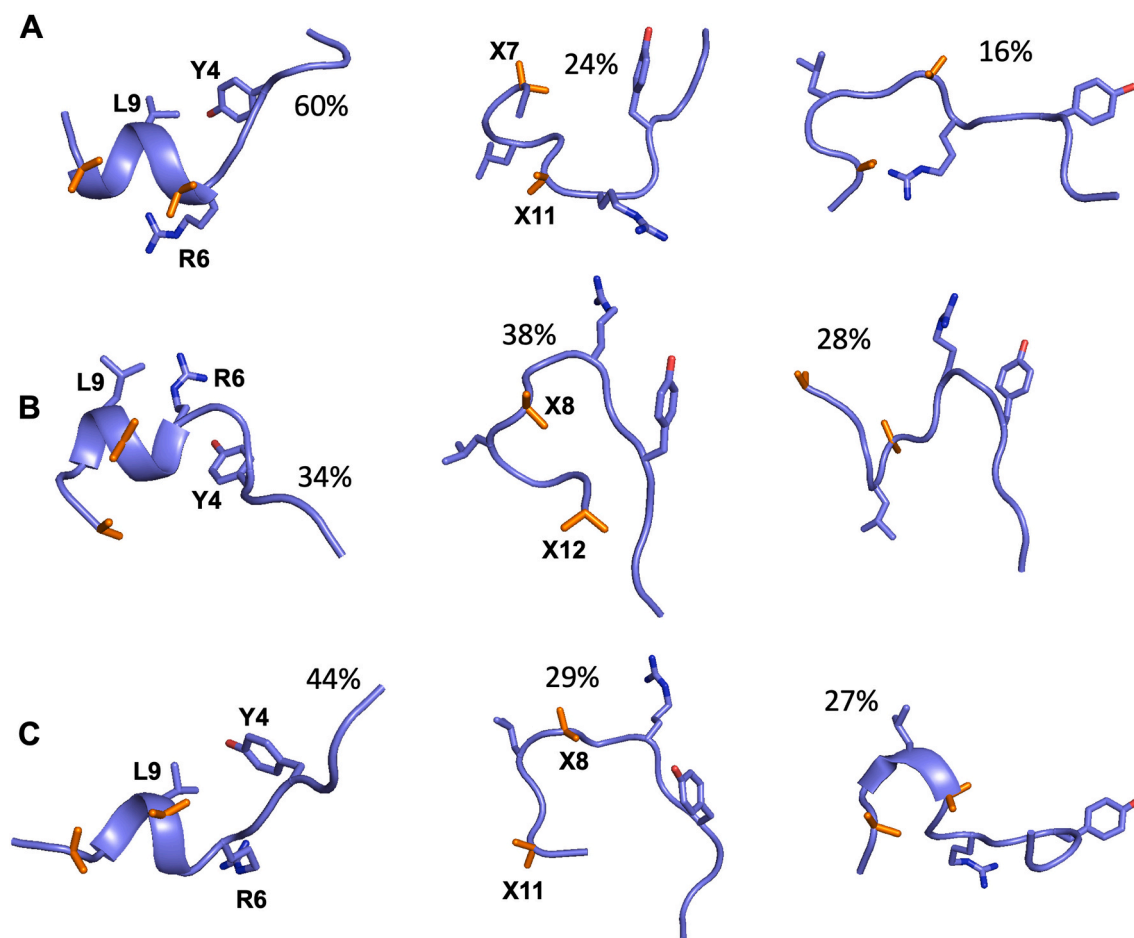
Fig. 4. Solution state conformations of Aib peptide derivatives. Representative structures from three different cluster groups generated by clustering the ensemble of structures obtained from MD simulations of **(A)** TIP-01, **(B)** TIP-02 and **(C)** TIP-03 peptides in solution. The percentage of structures in each cluster is indicated. Selective residues from the peptide along with the Aib substitutions (designated by X and coloured in orange) are shown in stick representations.

also highlighted from the lack of a prevalently populated cluster (Fig. 4C). These observations from MD simulations are in accord with the CD data of the peptides. Thus, both analyses showed that Aib substitution in TIP-01 resulted in the induction of some helical property but not in TIP-02 and TIP-03.

3.4. TIP peptides in the bound state with eIF4E

We also investigated the influence of Aib substitution on the eIF4E bound conformations of the peptides and the corresponding binding energetics. The complex state models of TIP peptides were generated based on the eIF4G^{D5S}: eIF4E crystal structure and subjected to long-time scale (1 μs each) conventional molecular dynamics (MD) simulations. The peptides remained bound in the "Reverse L-shaped conformation" (Fig. S3) and the helical segment was stable during the simulation period except for the C-terminal end of eIF4G^{D5S} and TIP-02 (Fig. 5A–D). In both these peptides, residues L10 and G11 had a reduced helical propensity (Fig. 5F), which could likely have arisen from the flexibility of glycine as it is substituted by Aib in all the other peptides. Residue-wise decomposition analysis from the simulation trajectories showed that residues K1, K2, R3, Y4, R6, and L9 made significant binding energy contributions across all the peptides (Fig. 5G). However, the residues from the N-terminal (K1, K2 and R3) also had a relatively higher degree of fluctuations which suggested that the interactions formed by these residues are dynamic and non-specific. Residues Y4, R6 and L9 on the other hand formed stable interactions as they constitute the specific hydrophilic and hydrophobic components of the

intermolecular recognition required to interact with eIF4E (Fig. S3A–S3D). E7 and G11 are solvent exposed and hence do not make any significant binding energy contributions. The three hydrophobic residues F8, L10 and F12 are all partially exposed to the solvent but they showed different energetic profiles (Fig. 5G). F8 did not exhibit any significant binding energy, L10 had a reasonable contribution ($\sim$ −1 kcal/mol) while F12 displayed the most favourable ($\sim$ −2 kcal/mol) energetic contribution. In TIP-02, F12 was substituted by the shorter side-chain of Aib which resulted in the loss of the favourable hydrophobic interactions (Fig. 5G). The Aib substitutions in any of the peptides do not directly contribute to the binding energy with eIF4E.

3.5. Structural reorganization of the peptides to adopt the bound state conformation

We next performed a comparative analysis between the ensemble of conformations generated from simulations of the free and bound states of the peptides by measuring their pairwise structural deviations. This variable should indicate the degree of reorganization required for the free peptides to achieve the reverse L-shaped bound conformation. The intrastate deviations of the different free peptides ranged from 0 to 5 Å albeit with different distribution profiles; in contrast, both the scale (0-2 Å) and the nature of the distribution remained similar for the bound states across the peptides (Fig. 6A–D). Interesting data emerged with interstate comparisons where a distinct property was observed for each peptide. In eIF4G^{D5S}, $\sim$ 30% of the sub-population was in the range between 0 and 2 Å, while the remaining ($\sim$70%) was concentrated
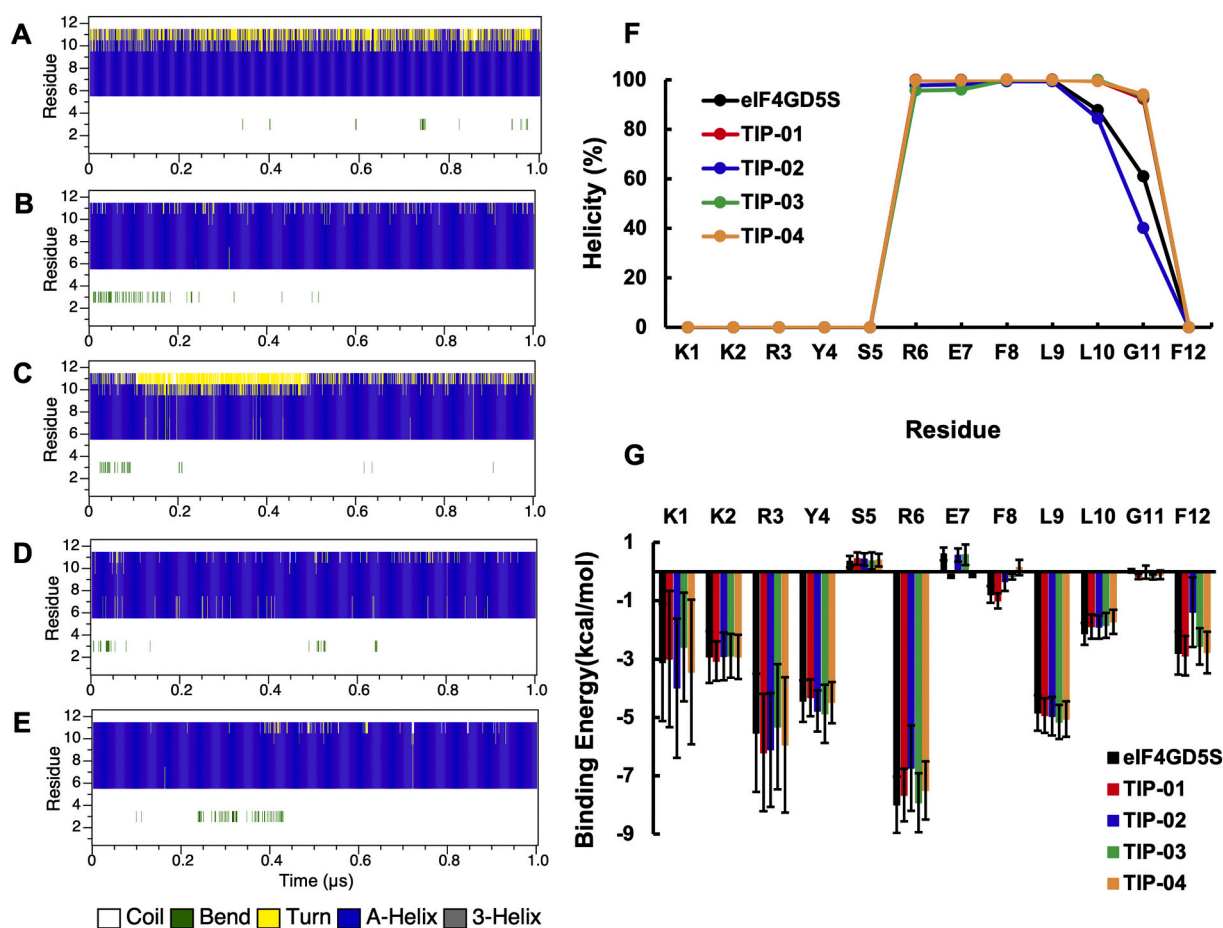
Fig. 5. Bound state properties of the peptides. Secondary structure evolution of **(A)** eIF4G^{D5S}, **(B)** TIP-01, **(C)** TIP-02, **(D)** TIP-03 and **(E)** TIP-04 peptides as a function of the simulation time. **(F)** Percentage helicity of individual residues from the different peptides. **(G)** Residue-wise binding energy contribution of the peptide in complex with eIF4E. The average binding energy and standard deviation is computed from the ensemble of structures generated from MD simulations. The calculation was done using Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method by following the same procedure and parameters as described previously [12]. The residues in the plots (**F and G**) are marked with respect to eIF4G^{D5S} peptide. See Table 1 for the Aib substituted positions in the corresponding derivatives.

between 2 and 4 Å (Fig. 6A) which illustrates that the majority of the peptide conformers have to undergo a higher degree of structural reorganization for complexation. In TIP-01, there was a significant jump in structural similarity between the two states of the peptide as almost half of the population (~50%) was in the range of 0-2 Å (Fig. 6B). This is likely due to the stable helical epitope present in significant conformers of the peptide in solution. The interstate structural variation in TIP-02 and TIP-03 peptides was considerable with the maximum (>80%) population distributed above 2 Å (Fig. 6C and D) due to the intrinsically disordered nature of these peptides in solution. Thus, this analysis showed that TIP-01 would undergo the least reorganization necessary for efficiently recognizing eIF4E.

We also explored the nature of specific intermolecular interactions involving residues Y4, R6 and L9 by docking representative peptide conformers in their free states onto eIF4E (Fig. S4). These three residues were chosen as they define the optimum mode of interactions with eIF4E. There is a clear distinction between "eIF4G^{D5S} and TIP-01" peptides compared to "TIP-02 and TIP-03". In the first sub-group (eIF4G^{D5S} and TIP-01), a significant population have a pre-formed helical turn in solution which enabled residues R6 and L9 to retain similar conformations between the two states of the peptides (Fig. 6E and F). This indicated that the R6: E132 salt-bridge and the L9: W73 hydrogen-bond interactions could be formed instantly once these peptides dock into the binding interface on eIF4E. The hydrogen-bond interaction between Y4 and P38 can then be created subsequently to attain the final bound state.

In the other subgroup (TIP-02 and 03), the peptides are predominantly in the disordered ensemble with low helical content which hinders their optimum docking and formation of specific intermolecular interactions. Further, even in states with the helical epitope present, TIP-02 and 03 would need the largest conformational reorganization to form the Y4: P38 hydrogen-bond (Fig. 6G and H). This interaction is made through the disordered region of the peptide, so its formation could be a rate-limiting step in the complexation process.

3.6. Optimization of TIP-01

TIP-01 is the only peptide that showed a relatively better binding affinity as compared to eIF4G^{D5S} (Table 1). As our analysis above suggests, the molecular basis for this could be the presence of a predominant metastable state in solution with a stable helical motif compared to the other peptides. In addition, TIP-01 also adopted a conformation which was comparatively similar to the bound state and this would lower the entropic penalty incurred upon binding. In TIP-01, hydrophilic (D7) and flexible (G11) residues were substituted by Aib. The other two peptides (TIP-02 and 03) included substitutions at F8 and F12/G11. This suggested that the two hydrophobic residue positions (F8 and F12) had an influence on the solution state conformation of the peptide. We wanted to investigate if TIP-01 can be optimized further in-order to improve its potency. Interestingly, phage display screening of eIF4E interacting peptides have previously highlighted the selective occurrence of F8Q
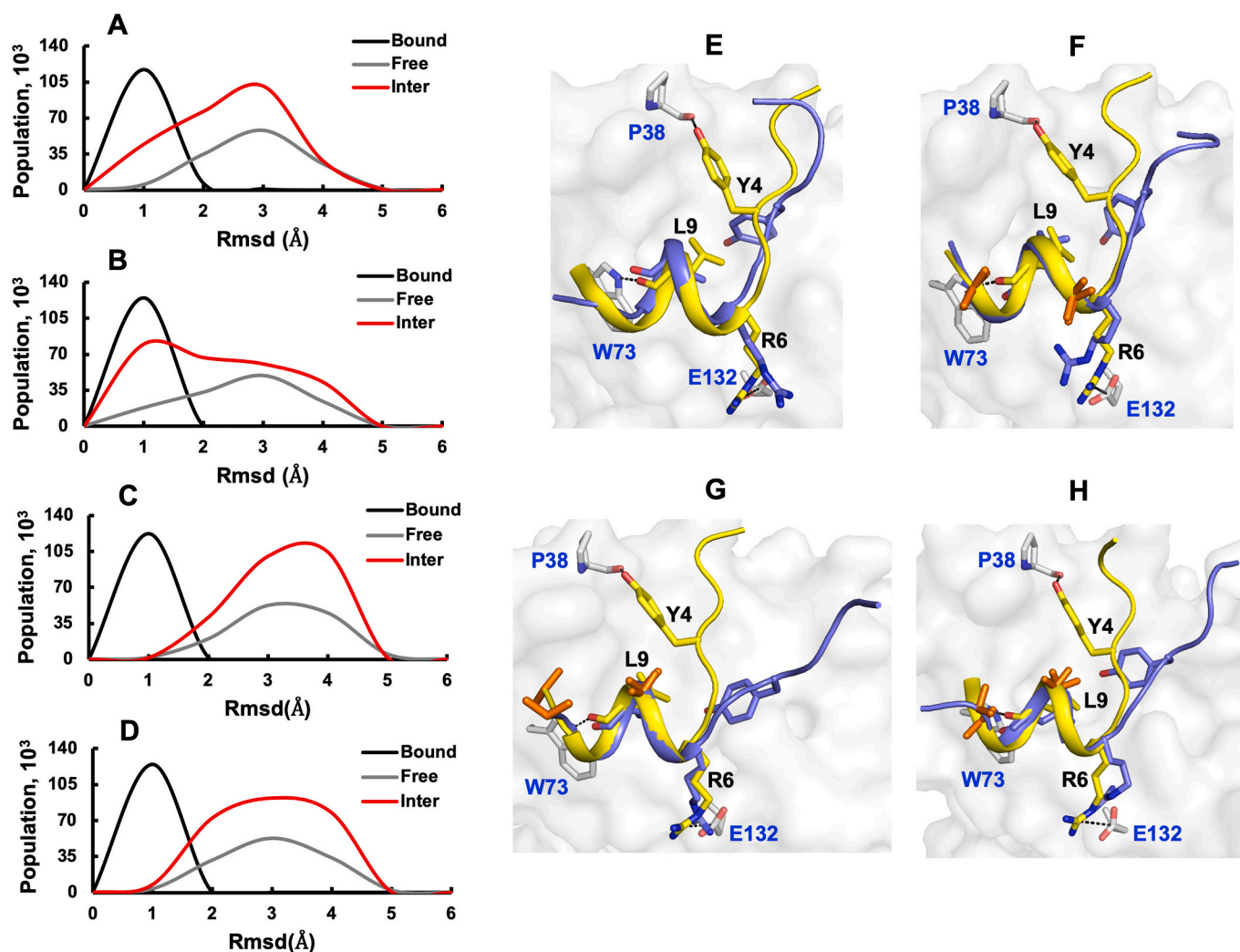
Fig. 6. Comparison of bound and free peptides. Population distribution of **(A)** eIF4G^{D5S}, **(B)** TIP-01, **(C)** TIP-02 and **(D)** TIP-03 peptide conformations in terms of the pairwise RMSD computed for the intra (bound and free) and inter (bound vs free) states of the different peptides. Structural superimposition of representative bound (yellow) and free (purple) states of **(A)** eIF4G^{D5S}, **(B)** TIP-01, **(C)** TIP-02 and **(D)** TIP-03 peptides. The peptides are shown in cartoon and the protein in surface representation. The side-chain of residues which are involved in specificity defining hydrogen-bond and salt-bridge interactions are explicitly depicted in stick representations. These interactions in the bound state of the peptide are highlighted with dashed lines. The side-chain of Aib substitutions are also shown as sticks coloured in orange.

and F12L substitutions [11]. We generated TIP-04 peptide by incorporating these substitutions into TIP-01 sequence and the peptide derivative had a ~ 7/3-fold improvement (from SPR and FP assays respectively) in binding affinity as compared to eIF4G^{D5S} (Table 1, Figs. 2A, S1 and S2).

The helicity of TIP-04 investigated via CD spectroscopy was found to relatively higher than all the other peptides (Fig. 2B). MD simulation of 200 ns carried out on TIP-04 starting from an extended structure also indicated that it had a significantly enhanced helical character (Fig. 7A). The helix, mostly across residues 6–10, was formed instantly, and remained stable during the time course of the simulation (Fig. 7A and B). MD simulation of the peptide complexed to eIF4E showed that the helicity in the bound state was also stable and the binding energetic profile of the common individual residues was similar to that of the other peptides (Fig. 5E–G). Neither F8Q nor F12L substitutions resulted in direct improvements in the binding energy contributions. Residues F8/Q8 do not physically interact with eIF4E, so it is energetically more favourable to incorporate a hydrophilic residue at this position. Q8 substitution also potentially stabilizes the helical conformation by reinforcing the capping role of S5 with additional hydrogen-bond interaction (Fig. S5). Consequently, the bulk of the free peptide

structures had pairwise deviations of <3 Å, which is significantly better than TIP-01, where a sizeable population had variations above this threshold (Figs. 7D and 6B). It is in-fact remarkable to observe the presence of a highly populated (81%) peptide structure in solution which is similar to the "Reverse L-shaped conformation" observed in the bound state (Fig. 7C). This suggested that the free peptide will undergo only a small amount of reorganization to efficiently interact with the protein. This is confirmed by the interstate comparison of the peptide structures where majority of the population was concentrated in the region of <2 Å (Fig. 7D). In contrast, TIP-01 had a relatively widespread population distribution up to 5 Å (Fig. 6B). The high degree of structural similarity in TIP-04 resulted in an effective docking posture for the disordered region in the peptide (Fig. 7E) which would facilitate a faster formation of the rate-limiting hydrogen-bond interaction between Y4 and P38. Thus, in summary, TIP-04 peptide had an optimized solution state conformation which is a major molecular determinant for the improvement in the binding affinity of this peptide for eIF4E.

4. Discussion

A structure-based rational design strategy was used to develop Aib
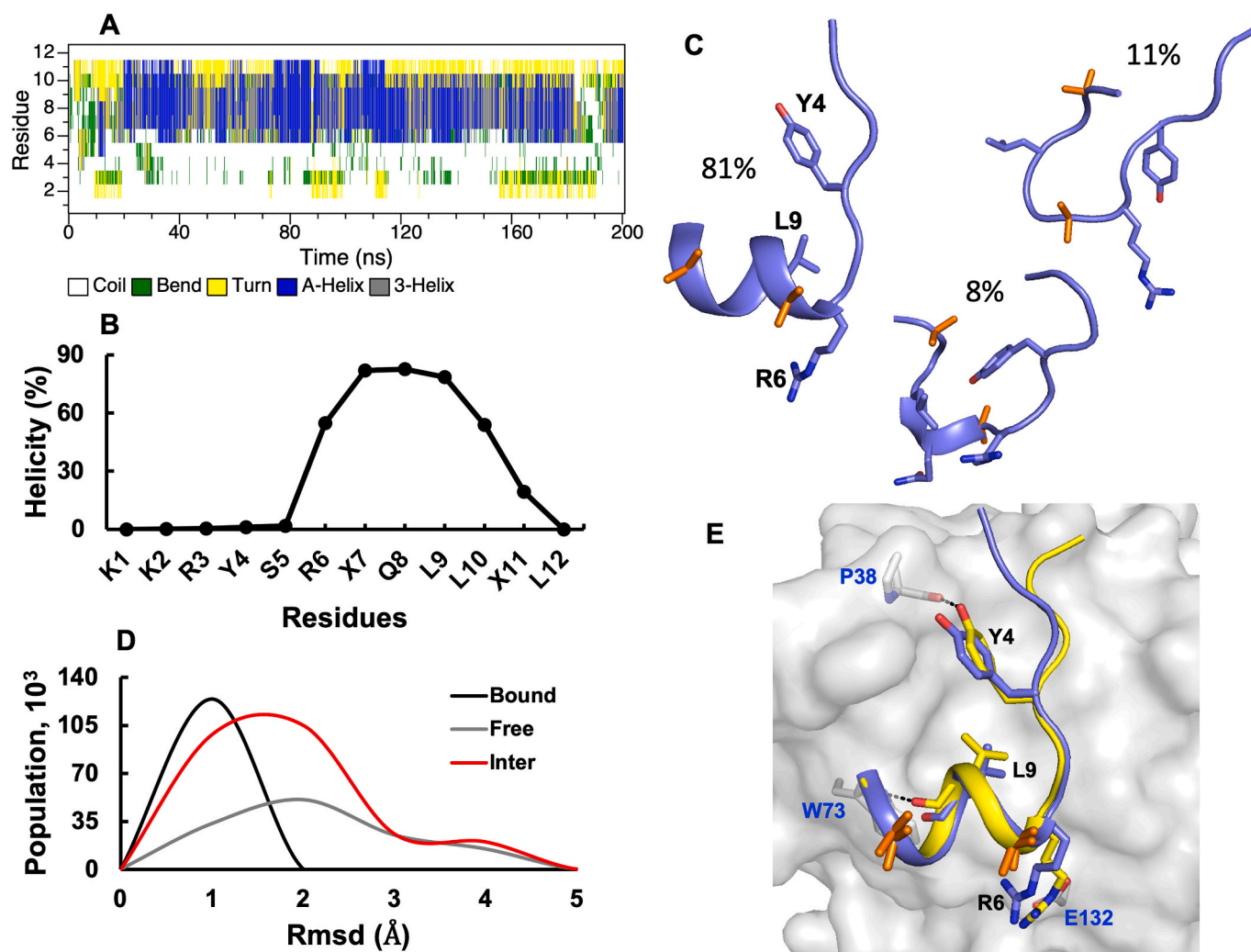
Fig. 7. TIP-04 peptide. (A) Secondary structure evolution of TIP-04 peptide as a function of the simulation time. **(B)** Percentage helicity of individual residues in TIP-04 peptide. **(C)** Representative structures from three different cluster groups generated by clustering the ensemble of structures obtained from MD simulations of TIP-04 peptide. The percentage of structures in each cluster is indicated. Selective residues from the peptide along with the Aib substitutions (coloured in orange) are shown in stick representations. **(D)** Population distribution of TIP-04 peptide conformations in terms of the pairwise RMSD computed for the intra (bound and free) and inter (bound vs free) states of the peptide. **(E)** Structural superimposition of representative bound (yellow) and free (purple) states of TIP-04 peptide. The peptide is in cartoon and the protein in surface representation. The side-chain of residues which are involved in specificity defining hydrogen-bond and salt-bridge interactions are explicitly depicted in stick representations. These interactions in the bound state of the peptide are highlighted with dashed lines. The side-chain of Aib (orange) residues are also shown as sticks.

containing peptide derivatives targeted against eIF4E. The insertion of Aib into the sequence had differential impact on the helical propensities of the peptides depending on the point of incorporation. It resulted in increased helicity in TIP-01 and TIP-04 derivatives, and yet had no effects on TIP-02 and TIP-03. This specifies that although Aib is considered as a helix promoting non-natural amino acid [16,17], its mere substitution in the sequence does not necessarily result in improved helicity. This has been demonstrated here through CD spectroscopy and effectively complemented by MD simulations; it is clear that the SGLD simulation method used in the study was able to efficiently sample exhaustively the structural properties of these peptides. As a further assessment of the outcome from the simulations, we investigated the folding of two peptides (TIP-02 and TIP-04) that respectively exhibited an unstructured and relatively more helical character in the CD spectra (Fig. 2B). TIP-02 was again found to adopt a random conformation while TIP-04 contained a helical epitope in its structure (Fig. S6). Thus, the simulations were able to successfully replicate the variable helical properties of the Aib derived TIP peptides in solution.

The primary structure of the parent peptide fragment derived from

the IDR of eIF4G can be divided into two segments based on its chemical property; a hydrophilic N-terminal and a relatively more hydrophobic C-terminal. This difference was also manifested in its secondary structural property where the N-terminal was disordered, while localized transient helical epitopes were detected at the C-terminal. IDRs in proteins are reported to harbour such local structural elements in solution, for instance as observed in the transactivation domains of tumor suppressor p53 protein [35]. They serve as recognition motifs for interactions with binding partners like Mdm2 [36]. It is interesting to note that these structural elements are selectively formed in segments within the IDR that have strategically positioned hydrophobic residues [35,36]. The incorporation of Aib residues in TIP-01 and TIP-04 systematically increases the hydrophobicity besides inducing helical backbone geometry, which collectively enhances the occurrence and stability of local helical structures in these peptides. On the contrary, Aib substitution in TIP-02 and TIP-03 reduces the overall hydrophobicity; this coupled with the presence of additional charged/flexible residues could be a primary determinant for their intrinsically disordered property in solution. Thus, strategic incorporation of Aib along the sequence clearly is critical for

the desired improvement in helicity of the peptide derivatives and their potency to interact with eIF4E.

The "Reverse L-shaped conformation" is a canonical binding mode of eIF4E interacting peptides that contain the "Tyr-X4-Leu-ø" motif [10]. The specific conformation is critical as it enables the tyrosine (Tyr) residue in the motif to form a key hydrogen-bond interaction that is important for recognizing eIF4E (Fig. 1A). Mutation of this residue has been shown to adversely impact the peptide's affinity for the protein [9]. Thus, in addition to helicity, the binding activity of the peptides will also be governed by the dynamic ensemble of states that adopts bound state like conformation in solution. The mode of recognition in this system is distinct from other well studied peptide: protein systems such as p53: Mdm2 [35] or the Bcl-2 family of proteins [37], where the critical interactions are located entirely within the helical scaffold. The conserved leucine (Leu) residue from the motif is an integral component of the helical segment across all the free peptides and it docks specifically into a pre-formed hydrophobic pocket present on the PPI interface of eIF4E (Fig. 1A). A potential binding reaction would involve the formation of an initial encounter complex through selective docking of the ordered helical unit followed by structural adjustments to attain the final mode of binding. Such combined mechanism of recognition which is driven by conformational selection and optimized through an induced fit process has been shown to exist in different biomolecular interactions [38–40]. The ensemble of free TIP-04 structures are predominantly biased towards the bound state conformation which implies an efficient binding pathway for initial selection and subsequent fitting of this peptide. This will collectively decrease its entropic cost of binding and consequently as observed, TIP-04 was a potent eIF4E binder. The optimization outline of this peptide further indicates that the hinge region between the disordered and helical segment is important to preserve the bioactive conformation. In summary, TIP-04 is a viable lead compound and it's structure-activity relationship provides important insights for consideration in the continuing efforts towards the development of peptide-based inhibitors against eIF4E in oncology.

Credit author statement

C.J.B, C.S.V., D.P.L. and D.L. conceptualized the project. C.J.B. performed the biophysical experiments. D.L. performed molecular modeling and simulations. All the authors discussed the results and commented on the data. D.L. compiled, analyzed and organized the data for the manuscript. D.L. wrote the manuscript with editing and proof reading from all the other co-authors.

Declaration of Competing Interest

C.S.V. is the founder of Sinopsee Therapeutics and Aplomex. There is no conflict of interest with the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbagen.2020.129775.

References

[1] W. Basile, M. Salvatore, C. Bassot, A. Elofsson, Why do eukaryotic proteins contain more intrinsically disordered regions? PLoS Comput. Biol. 15 (7) (2019), e1007186.

[2] J.B. Ahrens, J. Nunez-Castilla, J. Siltberg-Liberles, Evolution of intrinsic disorder in eukaryotic proteins, Cell. Mol. Life Sci. 74 (17) (2017) 3163–3174.

[3] P. Tompa, E. Schad, A. Tantos, L. Kalmar, Intrinsically disordered proteins: emerging interaction specialists, Curr. Opin. Struct. Biol. 35 (2015) 49–59.

[4] M.S. Cortese, V.N. Uversky, A.K. Dunker, Intrinsic disorder in scaffold proteins: getting more from less, Prog. Biophys. Mol. Biol. 98 (1) (2008) 85–106.

[5] D. Prévôt, J.L. Darlix, T. Ohlmann, Conducting the initiation of protein synthesis: the role of eIF4G, Biol. Cell. 95 (3–4) (2003) 141–156.

[6] W.C. Merrick, eIF4F: a retrospective, J. Biol. Chem. 290 (40) (2015) 24091–24099.

[7] J. Pelletier, J. Graff, D. Ruggero, N. Sonenberg, Targeting the eIF4F translation initiation complex: a critical nexus for cancer development, Cancer Res. 75 (2) (2015) 250–263.

[8] H. Malka-Mahieu, M. Newman, L. Désaubry, C. Robert, S. Vagner, Molecular pathways: the eIF4F translation initiation complex-new opportunities for cancer treatment, Clin. Cancer Res. 23 (1) (2017) 21–25.

[9] S. Mader, H. Lee, A. Pause, N. Sonenberg, 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, Mol. Cell. Biol. 15 (9) (1995) 4990–4997.

[10] J. Marcotrigiano, A.C. Gingras, N. Sonenberg, S.K. Burley, Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G, Mol. Cell 3 (6) (1999) 707–716.

[11] W. Zhou, S.T. Quah, C.S. Verma, Y. Liu, D.P. Lane, C.J. Brown, Improved eIF4E binding peptides by phage display guided design: plasticity of interacting surfaces yield collective effects, PLoS One 7 (10) (2012) e47235.

[12] D. Lama, S.T. Quah, C.S. Verma, et al., Rational optimization of conformational effects induced by hydrocarbon staples in peptides and their binding interfaces, Sci. Rep. 3 (2013) 3451.

[13] D. Lama, A.M. Liberatore, Y. Frosi, et al., Structural insights reveal a recognition feature for tailoring hydrocarbon stapled-peptides against the eukaryotic translation initiation factor 4E protein, Chem. Sci. 10 (8) (2019) 2489–2500.

[14] M. Alías, S. Ayuso-Tejedor, J. Fernández-Recio, C. Cativiela, J. Sancho, Helix propensities of conformationally restricted amino acids. Non-natural substitutes for helix breaking proline and helix forming alanine, Org. Biomol. Chem. 8 (4) (2010) 788–792.

[15] C. Cativiela, M.D. Diaz-de-Villegas, Recent Progress on the Stereoselective synthesis of acyclic quaternary alpha-amino acids, Tetrahedron Asymmetry 20 (1) (2009) 1–63.

[16] R. Mahalakshmi, P. Balaram, Non-protein amino acids in the design of secondary structure scaffolds, Methods Mol. Biol. 340 (2006) 71–94.

[17] C. Toniolo, M. Crisma, F. Formaggio, C. Peggion, Control of peptide conformation by the Thorpe-Ingold effect (C alpha-tetrasubstitution), Biopolymers. 60 (6) (2001) 396–419.

[18] Y. Demizu, M. Doi, M. Kurihara, et al., Conformational studies on peptides containing α,α-disubstituted α-amino acids: chiral cyclic α,α-disubstituted α-amino acid as an α-helical inducer, Org. Biomol. Chem. 9 (9) (2011) 3303–3312.

[19] G. Tsuji, T. Misawa, M. Doi, Y. Demizu, Extent of helical induction caused by introducing α-Aminoisobutyric acid into an Oligovaline sequence, ACS Omega. 3 (6) (2018) 6395–6399.

[20] R. Banerjee, G. Basu, P. Chène, S. Roy, Aib-based peptide backbone as scaffolds for helical peptide mimics, J. Pept. Res. 60 (2) (2002) 88–94.

[21] C.J. Brown, I. McNae, P.M. Fischer, M.D. Walkinshaw, Crystallographic and mass spectrometric characterisation of eIF4E with N7-alkylated cap derivatives, J. Mol. Biol. 372 (1) (2007) 7–15.

[22] M.H. Roehrl, J.Y. Wang, G. Wagner, A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization, Biochemistry. 43 (51) (2004) 16056–16066.

[23] Z.X. Wang, An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule, FEBS Lett. 360 (2) (1995) 111–114.

[24] D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, et al., AMBER 18, University of California, San Francisco, 2018.

[25] E. Vanquelef, S. Simon, G. Marquant, et al., R.E.D. server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments, Nucleic Acids Res. 39 (Web Server issue) (2011) W511–W517.

[26] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput. 11 (8) (2015) 3696–3713.

[27] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, et al., (1983). Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.

[28] X. Wu, B.R. Brooks, Self-guided Langevin dynamics simulation method, Chem. Phys. Lett. 381 (2003) 512–518.

[29] X. Wu, A. Damjanovic, B.R. Brooks, Efficient and unbiased sampling of biomolecular systems in the canonical ensemble: a review of self-guided Langevin dynamics, Adv. Chem. Phys. 150 (2012) 255–326.

[30] R.J. Loncharich, B.R. Brooks, R.W. Pastor, Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide, Biopolymers. 32 (5) (1992) 523–535.

[31] H.J.C. Berendsen, J.P.M. Postma, W.F. Van Gunsteren, et al., Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.

[32] T. Darden, D. York, L. Pedersen, Particle mesh Ewald-an Nlog(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.

[33] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.

[34] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers. 22 (12) (1983) 2577–2637.

[35] H. Lee, K.H. Mok, R. Muhandiram, et al., Local structural elements in the mostly unstructured transcriptional activation domain of human p53, J. Biol. Chem. 275 (38) (2000) 29426–29432.

[36] P.H. Kussie, S. Gorina, V. Marechal, et al., Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain, Science. 274 (5289) (1996) 948–953.

[37] A.M. Petros, E.T. Olejniczak, S.W. Fesik, Structural biology of the Bcl-2 family of proteins, Biochim. Biophys. Acta 1644 (2–3) (2004) 83–94.

[38] D.A. Silva, G.R. Bowman, A. Sosa-Peinado, X. Huang, A role for both conformational selection and induced fit in ligand binding by the LAO protein, PLoS Comput. Biol. 7 (5) (2011), e1002054.

[39] I. Kurisaki, M. Takayanagi, M. Nagaoka, Combined mechanism of conformational selection and induced fit in U1A-RNA molecular recognition, Biochemistry. 53 (22) (2014) 3646–3657.

[40] T. Wlodarski, B. Zagrovic, Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin, Proc. Natl. Acad. Sci. U. S. A. 106 (46) (2009) 19346–19351.

[41] J. Shao, S.W. Tanner, N. Thompson, T.E. Cheatham, Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms, J. Chem. Theory Comput. 3 (6) (2007) 2312–2334.