Current Research in Structural Biology 4 (2022) 29–40

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

Transactivation domain of Adenovirus Early Region 1A (E1A): Investigating folding dynamics and aggregation

Nitin Sharma [a,1], Kundlik Gadhave [a,1], Prateek Kumar [a,1], Rajanish Giri [a,b,*]

[a] School of Basic Sciences, Indian Institute of Technology Mandi, Kamand, Himachal Pradesh, 175005, India
[b] BioX Center, Indian Institute of Technology Mandi, Kamand, Himachal Pradesh, 175005, India

ARTICLE INFO

Handling Editor:Glaucius Oliva

Keywords:
Intrinsically disordered proteins
Molecular recognition elements
E1A
Protein folding
Protein aggregation

ABSTRACT

Transactivation domain of Adenovirus Early region 1A (E1A) oncoprotein is an intrinsically disordered molecular hub protein. It is involved in binding to different domains of human cell transcriptional co-activators such as retinoblastoma (pRb), CREB-binding protein (CBP), and its paralogue p300. The conserved region 1 (TAD) of E1A is known to undergo structural transitions and folds upon interaction with transcriptional adaptor zinc finger 2 (TAZ2). Previous reports on Taz2-E1A studies have suggested the formation of helical conformations of E1A-TAD. However, the folding behavior of the TAD region in isolation has not been studied in detail. Here, we have elucidated the folding behavior of E1A peptide at varied temperatures and solution conditions. Further, we have studied the effects of macromolecular crowding on E1A-TAD peptide. Additionally, we have also predicted the molecular recognition features of E1A using MoRF predictors. The predicted MoRFs are consistent with its structural transitions observed during TAZ2 interactions for transcriptional regulation in literature. Also, as a general rule of MoRFs, E1A undergoes helical transitions in alcohol and osmolyte solution. Finally, we studied the aggregation behavior of E1A, where we observed that the E1A could form amyloid-like aggregates that are cytotoxic to mammalian cells.

1. Introduction

Intrinsically disordered proteins (IDPs) are abundant in eukaryotic transcription factors (Wright and Dyson, 2015). IDPs are considered as a fascinating class of proteins due to their vast range of functions. Some of the model-disordered proteins are well characterized to analyze their critical functions in transcriptional regulatory processes such as transactivation domains of cMyb and Adenovirus Early region 1A (E1A) (Giri et al., 2013). The functional advantages of lack of intrinsic structure are ideally suited for macromolecular interactions and mediate transcriptional regulatory processes. Intrinsically disordered regions (IDRs) are widely found in interaction hub proteins involved in certain protein-protein, protein-DNA, and protein-RNA interactions. Most of the cellular interaction networks are dependent on multiple protein interactions mediated by central hub proteins.

Viruses are considered obligate parasites as they are wholly dependent on host cell machinery for their survival. For their replication, viruses often take up the host machinery by molecular interactions. The protein interaction networks of host cells are particularly vulnerable to viral hijack. Viruses often target the interaction hub proteins to control the cellular network of hosts. Human adenoviruses (HAdV) are dsDNA viruses and among the important human pathogens. E1A is considered as a model oncogene. Its encoded proteins control viral transcription levels in infected cells and reprogram the gene expression in cells. E1A encoded proteins are acidic and contain multiple proline residues in primary sequence, which decrease the propensity of secondary structure formation (Frisch and Mymryk, 2002). E1A is largely disordered and contains small linear interaction motifs (SLiM's) (King et al., 2018). It can target multiple host systems despite its small size. E1A multi-functionality arises from its high conformational plasticity and its ability to undergo reversible disorder-to-order transitions. One such transition is the disorder--to-α-helical conformational change in E1A (53–91 residues; shown in Fig. 1a) conserved region 1 (TAD) that shows induced folding upon interactions with TAZ2 domain of CREB binding protein (CBP) (Ferreon et al., 2009). The TAD sequence of E1A protein is highly conserved across human adenoviruses. Certain residues like glutamic acid (55th and 59th), aspartic acid (56th), proline (57th), asparagine (58th),

* Corresponding author. School of Basic Sciences, Indian Institute of Technology Mandi, Himachal Pradesh, 175005, India.
E-mail addresses: rajanishgiri@iitmandi.ac.in, rajanishgiri2019@gmail.com (R. Giri).
[1] Authors contributed equally.

https://doi.org/10.1016/j.crstbi.2022.01.001
Received 2 June 2021; Received in revised form 5 December 2021; Accepted 4 January 2022

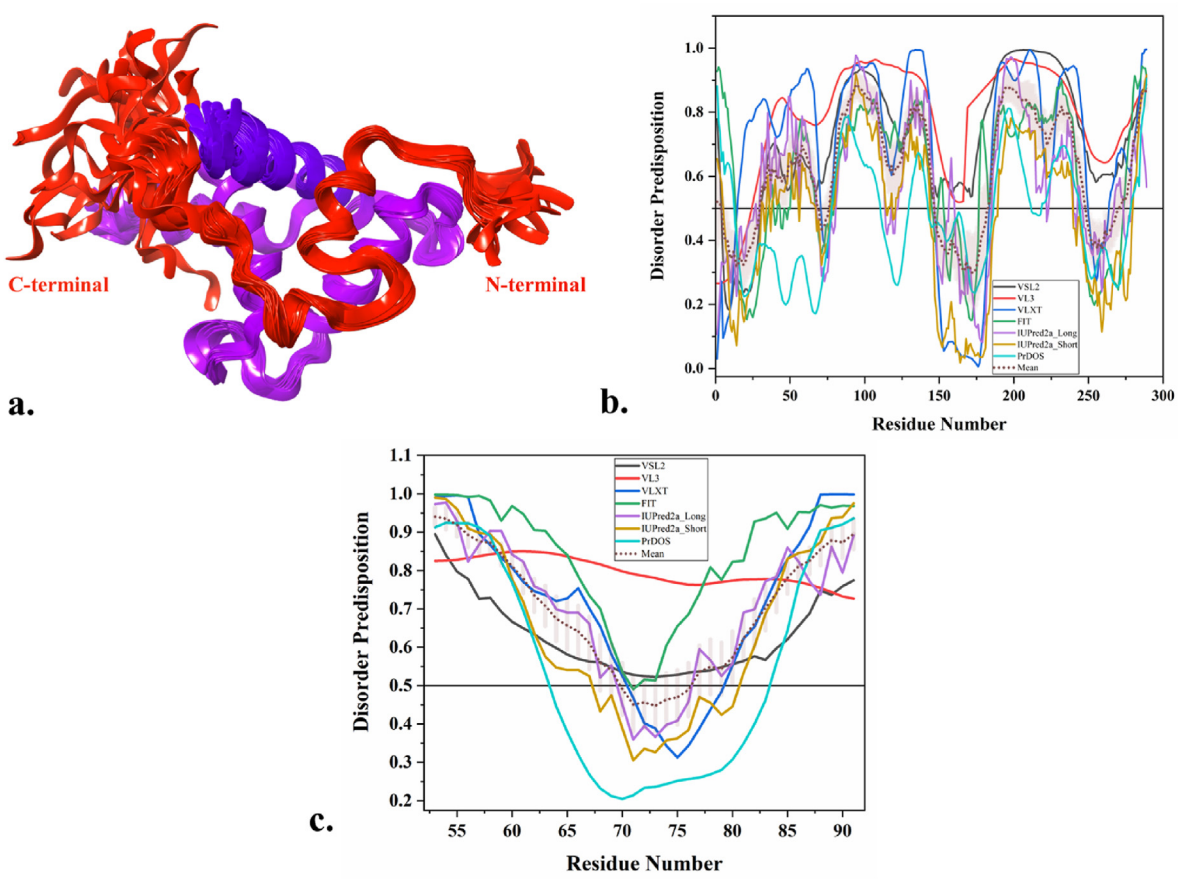


Fig. 1. (a) Ribbon representation of E1A structure (red) in complex with CREB protein (violet), (b) sequence-based illustration of disorder predisposition in full-length E1A protein, and (c) E1A-TAD (residues 53–91). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

alanine (61st), valine (62nd), phenylalanine (66th), and leucine (72nd) are found to be extremely conserved (Avvakumov et al., 2002).

IDPs vary significantly in their conformational transitions due to the variability in protein sequences and structural propensities. This conformational freedom has increased the diversity of binding mechanisms of IDPs. Coupled folding and binding process of IDPs to their binding partners follow two kinds of mechanisms, i.e., conformational selection and induced-fit mechanism or a combination of both. A report has suggested that the intrinsic structural propensity of IDPs determines their binding mechanisms to interacting partners (Arai et al., 2015). IDPs have distinct binding and folding mechanisms and do not obey a common folding pathway.

Helix formation reaction in IDRs is often swift and occurs in microsecond timescale (Lin et al., 2011). Preformed helix structures can be the elemental features for binding requirements to interacting partners. A report has suggested the importance of the preformed helix as a determinant of molecular recognition in IDPs (Iešmantavičius et al., 2014). Molecular recognition is a primary element of protein interactions. Molecular Recognition Elements (MoRE) are essential short disorder segments prone to transition into ordered structure upon interactions with multiple binding partners. These segments exist in disordered states and adopt partially or well-defined ordered states on interactions. The objective of this work is to have implications on the functional properties of many IDPs in different solvent conditions and crowded milieu. E1A exemplifies IDPs and is a physiologically relevant model to study the effects of crowded milieu or solvent conditions on IDRs.

We have investigated the disorder-to-α-helical transition in E1A peptide to perform the conformational analysis of secondary structure elements and intermediate structures. These peptide transitions were induced by titration of the membrane-mimetic solvent trifluoroethanol (TFE), hexafluoro-isopropanol (HFIP), and ethanol solutions, and conformational change was monitored using circular dichroism (CD) spectroscopy. Many studies provide evidence of co-solvents involved in various conformational changes with respect to concentration and time (Kumar et al., 2021a; Kumar et al., 2021b; Kumar et al., 2020; SÖnnichsen et al., 1992; Xue et al., 2006; Huang and Kim, 2012). Several alcohols are known to induce secondary structure formation in polypeptides. TFE is one among the widely used solvents in folding experiments to investigate solvent-induced structural changes of proteins, especially α-helix formation (Culik et al., 2014). IDPs are known to have disordered or unstructured regions in their physiological conditions. So, it is interesting to check the effect of TFE induced conformational changes on IDPs. Adenoviral protein E1A is a well-studied IDPs that has been known to involve in various protein-protein interactions due to disordered regions present in the native structure state. Therefore, we have tried to analyze the effects of TFE and co-solvents on the structural changes of E1A. One of the recent research has also proposed a method for experimental identification of IDRs that undergo binding induced disorder-to-helix transition under the effect of TFE (Glover et al., 2016). The lack of structure provides conformational flexibility to facilitate multiple protein-protein or protein-nucleic acid interactions (Shoemaker et al., 2000). Protein-protein interactions play important role in cellular signaling activities. Various studies have reported the inhibition of functions of host hub proteins and the introduction of novel connections within networks (Ahmed et al., 2018; Bojadzic and Buchwald, 2018). Collectively, we have studied the folding and aggregation behavior of viral hub protein E1A of adenoviruses.

2. Materials and methods

2.1. Chemicals

The peptide of early region 1 A (E1A) of Adenoviral oncoprotein [NH2] MAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPE [COOH] was obtained from Thermo Fisher Scientific™ with more than 97% of purity. The fluorescent probe Thioflavin T (ThT) and 1-anilinonapthalene-8-sulfonate (bis-ANS) was purchased from Sigma Aldrich (USA). All other chemicals were of analytical grade. SH-SY5Y human neuroblastoma cells was purchased from NCCS Pune.

2.2. Circular dichroism spectroscopy

Far-UV CD spectra were obtained using a MOS-500 spectropolarimeter (Biologic, France) equipped with a thermostatically controlled cell holder. Spectra were recorded in 0.1 cm quartz cells from 190 to 240 nm using the continuous-scan option (50 nm/min scan speed), with a step size of 0.5 nm and a bandwidth of 0.5 nm. CD spectra of the appropriate buffers (buffer baselines) were recorded and subtracted from the protein spectra. For all spectra, an average of 3 scans was obtained. CD spectra were obtained for 15 μM of 40-aa E1A-TAD peptide (53–91) in 50 mM Phosphate buffer (pH 7.0) and different concentrations of TFE, HFIP, and ethanol. For aggregation experiments, at desired time points, 20 μM of the incubated solution was collected and CD spectra were recorded. The structural changes of the peptide have been predicted based on its mean residual ellipticity ($[\theta]_{222,}$ $[\theta]_{198}$) at 222 nm and 198 nm respectively. The change in the total helical propensity of the folding transition state of the TAD region of E1A was analyzed and calculated using the Dichroweb Analysis program CONTIN (Whitmore and Wallace, 2004).

2.3. Molecular recognition features (MoRFs) analysis

Molecular recognition features (MoRFs) are short disorder segments of a protein that undergo a structural transition from disorder to order on interaction with the target protein. These segments are prone to interactions and play important role in molecular recognition and binding to interacting partners. MoRFs of the TAD region (53–91) of E1A protein was analyzed by using computational predictors such as Anchor (Dosztányi et al., 2009), MoRFChibiweb (Malhis et al., 2015) and DISOPRED3 (Jones and Cozzetto, 2015).

2.4. Protein structure modeling

Due to the unavailability of the crystal structure of E1A peptide in isolation, *ab-initio* modelling of E1A peptide (53–91 residues) was done using *Quark* modelling module, an online server provided by Zhang Lab, which was ranked 1st in CASP9 and CASP10 in free modeling category (Xu and Zhang, 2012). The server constructs the models using amino acids sequence in fragments of 1–20 residues by applying replica-exchange Monte Carlo (RE-MC) simulation.

2.5. Molecular dynamic (MD) simulations

We have performed an explicitly all-atom Molecular dynamic simulation of the modeled peptide using the TIP4P water model at room temperature (298 K). The forcefield parameterization was done by OPLS-2005 in Desmond, embedded in Schrodinger suite (Bowers et al., 2006). The MD simulation procedure was performed in Berendsen thermostat using an NPT ensemble at a constant pressure throughout the whole simulation time of 150 ns as per previously published studies (Kumar et al., 2020, 2021b). We also observed the conformational changes in the peptide structure with secondary structure analysis and energy change using Replica Exchange MD (REMD) simulations for 6 replicas (numbered as 0 to 5) up to 50ns at temperatures ranging from 298 K to 363 K. The temperature calculation for six replicas was done by linear mode of the REMD module in the Desmond simulation package (Kumar et al., 2020, 2021a). To perform this analysis, we took out the minimum energy structure (at 105 ns in the trajectory, the frame consisted of the least energy) from the already simulated peptide trajectory. Further, the REMD simulations in TIP4P water model was performed and the peptide structure was observed in simulation for 50ns each in six replicas at temperatures 298 K, 311 K, 324 K 337 K, 350 K, and 363 K, simultaneously.

2.6. Preparation of E1A aggregates

The fresh E1A peptide solution was diluted to 2 mg/mL in a 20 mM phosphate buffer, pH 3. The sample was incubated at 4 °C in a 1.5 mL reaction tube on a thermomixer (1000 rpm). The E1A aggregates formation was then monitored by measuring ThT fluorescence intensity of incubated sample.

2.7. Thioflavin T (ThT) binding and kinetics assay

Thioflavin T (ThT) dye binds strongly with amyloids and gives strong fluorescence is used regularly for the detection of *in-vitro* amyloid fibril formation (C et al., 2017; Biancalana and Koide, 2010). The ThT assay was performed as described previously (Gadhave and Giri, 2020; Gadhave et al., 2021; Saumya et al., 2021). Briefly, the E1A aggregated peptide (25 μM) was mixed with 20 μM Thioflavin T dye and incubated for 10 min in dark. ThT fluorescence spectra were recorded at 450 nm excitation wavelength and emission wavelength range from 484 nm to 700 nm. For ThT kinetics, at desired time points, 25 μM of the incubated E1A sample was taken out and mixed with 20 μM ThT and ThT fluorescence was measured at 450 nm excitation and 490 nm emission wavelength. The measurements were made on a TECAN Pro 200 microplate reader in black 96 well plate with clear bottom. The solution containing ThT only without E1A protein was used as control. All the measurements were performed in triplicate.

2.8. Bis-ANS binding assay

The ability of aggregated E1A peptide to bind with bis-ANS dye was measured as described previously (Gadhave and Giri, 2020). The E1A aggregated sample was collected (25 μM) and mixed with the bis-ANS (50 μM), and incubated for 10 min in the dark. Further, bis-ANS fluorescence spectra were recorded at 380 nm excitation wavelength and emission scan from 414 nm to 650 nm. The solution containing bis-ANS only was used as a control. All the measurements were performed in triplicates.

2.9. High-resolution transmission electron microscopy (HR-TEM)

The aggregated sample of E1A peptide (7 μl) was drop-casted on 200 mesh formvar-coated copper grids and negatively stained by 3% ammonium molybdate. Further, drop-casted samples were air-dried overnight and observed in HR-TEM (HR-TEM) (FP 5022/22-Tecnai G2 20 S-TWIN, FEI) operating at an accelerating voltage of 200 kV.

2.10. Cell viability assay

Human SH-SY5Y neuroblastoma cells were maintained in 25 cm^2 T-flasks (Thermo scientific) in 1:1 Ham's F-12: Dulbecco's modified eagle medium (DMEM), supplemented with 10% (vol/vol) FBS and 1% penicillin-streptomycin with 5% CO2 humidified atmosphere at 37 °C. The MTT reduction assays to assess the effect of E1A aggregates on cell viability. The cells were plated in 96 well plates with approximately 3000 cells/well and allowed to adhere for 24 h in the CO_2 incubator. The E1A aggregates were centrifuged, and pellets were resuspended in PBS solution and washed three times. Further, the pellets obtained in PBS

were resuspended in DMEM media (1% FBS) to get the required concentrations. Then cells were treated with E1A aggregates at different concentrations and incubated for 72 h. Finally, 10 μl of 3- (4, 5-dimethylthiazol-2-yl)-2, 5- diphenyltetrazolium bromide (MTT) was added to each well, incubated for 3 h, 100 μl of DMSO added, and the absorbance was recorded using TECAN infinite M200 pro at 570 nm. The cells that were not treated with E1A aggregates were considered as control. All the measurements were performed in triplicates. Error bars in the plots represent standard deviation (SD).

3. Results and discussion

3.1. Results

We have also investigated the disorder propensity using the primary structure of full-length E1A protein as well as TAZ2 binding conserved region 1 (TAD) of E1A protein where more than 68% in full-length E1A and 80% of region have been observed to be disordered (Fig. 1b & c). Further, we have used this region to study the properties of disordered peptides in different solution environments. Results have suggested the importance of structural transitions of small IDRs in dysfunctional interactions and in the origin of new interactions that fold by being primed by their physiological partners.

3.2. CD spectra analysis of TAD of E1A (53–91 residues)

Previous studies have reported the gain of structure in TAD (53–91) region of E1A oncoprotein in the presence of Taz2 (De Guzman et al., 2000). However, there is no structural information available for particularly independent E1A peptide (53–91). We have examined the structure of the E1A peptide using CD spectroscopy and found it to be disordered, as shown in Fig. 2a. Further, temperature-induced structural changes are well-known in IDPs, and these changes are termed the contraction of the structure due to rising temperatures. Therefore, the temperature-induced folding in the E1A peptide was observed and found to be gradually contracting at increasing temperatures (Fig. 2b). The structural transitions are shown in Fig. 2c where the plot shows a change in ellipticity at 198 nm and 222 nm with rising temperature. The spectral changes suggest the induction of residual structural transitions in TAD region with increasing temperature.

3.3. Analysis of induced secondary structure transition from the coil to α-helix

Further to understand the extent of helical transitions in TAD region, the peptide was subjected to alcohol/fluorinated alcohols and aqueous mixtures. It is very well established that the presence of alcohol in the surrounding environment alters the secondary structures and functions of proteins in a concentration-dependent manner (Huang and Kim, 2012; Gopinath et al., 2014). For an instance, TFE is a well-known helix inducer. It weakens the hydrogen bonds of CO and NH (protein backbone) with surrounding water and also stabilizes the intrachain hydrogen bonds (Luo and Baldwin, 1997). Here, we have examined the spectral changes in peptide in the presence of different concentrations of Ethanol, TFE, and HFIP.

As shown in Fig. 3a & b, upon increasing ethanol concentration, the

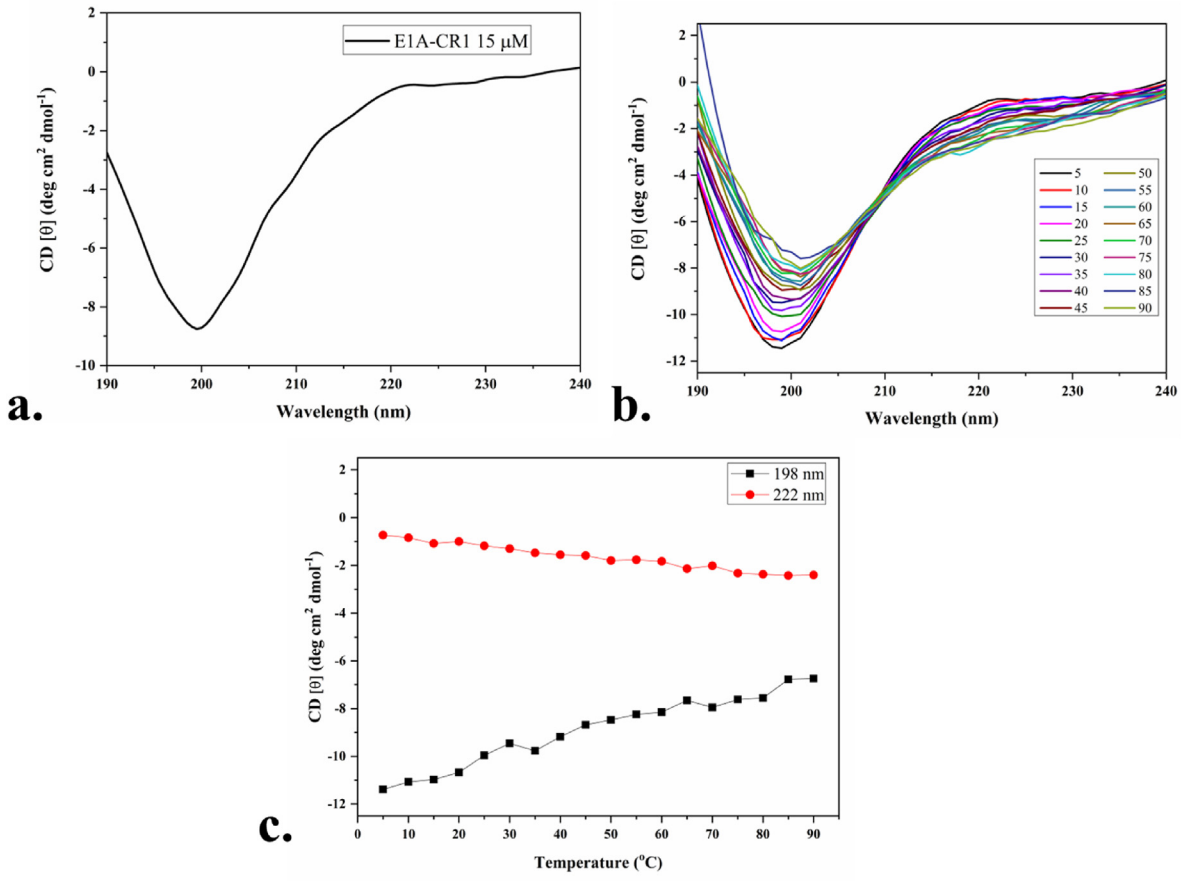


Fig. 2. Circular Dichroism spectra analysis of intrinsically disordered protein, Adenoviral oncoprotein early region (E1A). (a) Far UV-CD spectra of E1A protein in 50 mM Phosphate buffer, pH 7.0 at 25 °C. (b) Far UV-CD spectra analysis represents the change in the secondary structure of E1A protein with increasing temperature from 5 °C to 90 °C. (c) Far UV-CD spectra analysis at 198 nm (black) and 222 nm (red) represents the structural transition of E1A protein with increasing temperature. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

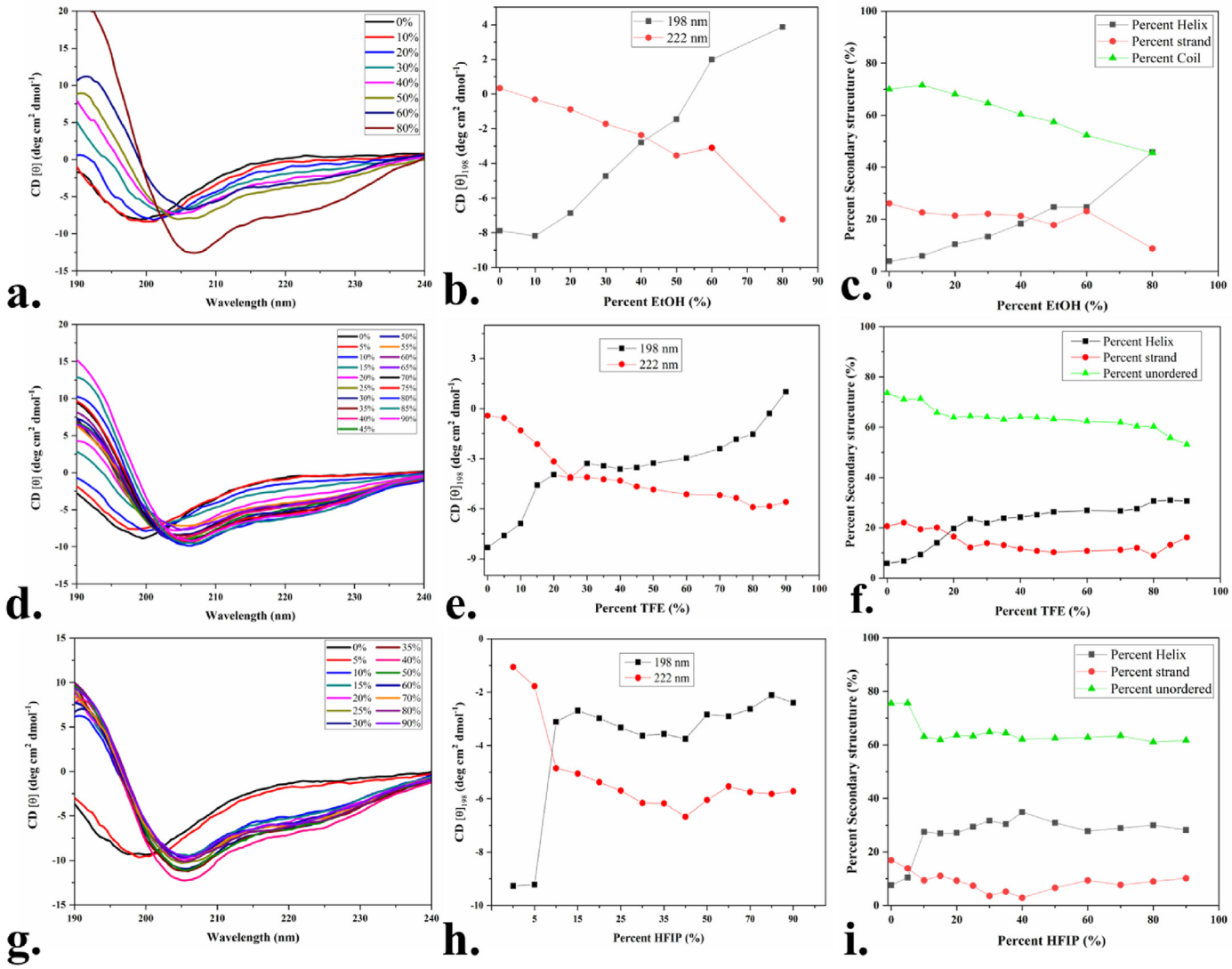


Fig. 3. Effect of alcohols on secondary structure transition of TAD region (53–91) of E1A. (a) Far UV CD spectra of E1A-TAD region (53–91) represent the induced gain of α-helix in presence of Ethanol solution from 0% to 80% at an interval of 10% each. (b) Far UV-CD spectra analysis at 198 nm (black) and 222 nm (red) represents the disorder to helix transition of E1A protein in presence of 80% Ethanol solution. (c) Induction of helicity in presence of Ethanol has been calculated using the Dichroweb Analysis program (Contin). The changes in the percent secondary structure elements are plotted against increasing concentration of Ethanol (0–80%). (d) Far UV-CD spectra of E1A protein represent the induced gain of α-helix in presence of Trifluoroethanol solution from 0% to 90% in 50 mM Phosphate buffer, pH 7.0 at 25 °C. (e) Far UV-CD spectra analysis at 198 nm and 222 nm represents the disorder to helix transition of E1A protein in presence of 90% TFE solution. (f) The changes in the secondary structure elements percentage are plotted against increasing concentration of TFE (0–90%). (g) Far UV-CD spectra of E1A protein represent the induced gain of α-helix in presence of HFIP solution from 0% to 90% in 50 mM Phosphate buffer, pH 7.0 at 25 °C. (h) Far UV-CD spectra analysis at 198 nm and 222 nm represents the disorder to helix transition of E1A protein in presence of 90% HFIP solution. (i) The changes in the secondary structure elements percentage are plotted against increasing concentration of HFIP (0–90%). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

disordered nature of E1A peptide is disturbed. At a maximum concentration of 80% Ethanol, the peptide gains an alpha-helical structure with comparatively higher negative ellipticity up to -13 deg cm^2 $dmol^{-1}$ showing two hallmark peaks of helical structure at 208 and 222 nm of CD spectra. As calculated using the Dichroweb server, the induction of helical structure has been increased on rising ethanol concentration. The E1A peptide shows 50% of helical and 50% random coil (Fig. 3c).

TFE has been widely used in folding experiments to investigate solvent-induced structural changes of proteins (SÖnnichsen et al., 1992; Culik et al., 2014). The experiments have been performed using CD spectroscopy to monitor changes in the secondary structures of proteins. The peptide fractional helicity (fH) is estimated based on its mean residue ellipticity at 222 nm. The spectra were measured to determine the distribution of gain in secondary structure in E1A peptide. CD spectra of E1A peptide in 50 mM Phosphate buffer and different concentrations of TFE are plotted in Fig. 3d & e that show a gradual shift in wavelength from 198 nm to 208 nm and 222 nm upon increasing TFE concentration, indicating a change in peptide structure from disorder to helical conformation. The change in percent secondary structure elements is plotted against TFE concentration and is represented graphically in Fig. 3f.

Similarly, the use of HFIP has been doing for a long time as a stabilizing solvent of secondary structures of peptides. HFIP interacts with the hydrophobic surface of the structured region (i.e., helix and sheet regions) by mimicking the folded protein environment (Colomer et al., 2017). In Fig. 3g, it is clearly shown that after 5% HFIP concentration, the disordered peak is not visible, and two signature peaks of helical conformation at 208 nm and 222 nm are observed. The disorder-to-order

transition can be seen till further increasing concentration of HFIP depicting stabilized secondary structure in the presence of HFIP molecules (Fig. 3h & i).

A commonly used lipid membrane mimicking agent, Sodium Dodecyl Sulfate (SDS), was chosen to examine the structural changes of E1A peptide. Fig. 4a & b shows a minimal shift in far-UV CD spectra from 198 nm to 205 nm with very few changes in their negative ellipticity. These changes can be interpreted as more minor or no significant changes in the structure from its disordered conformation in the presence of SDS.

3.4. Molecular crowding effects on isolated intrinsically disordered regions (IDRs)

The well-known concept about the interior of a cell is the restriction of accessible volume to proteins due to the presence of other macromolecules. The higher concentration of various other molecules in-vivo affects the structural transitions of different proteins. The presence of disordered regions in intrinsically disordered proteins provides structural flexibility in cellular environments. Ficoll (Fig. 4c & d) and PEG (Fig. 4e & f) are widely used inert molecular crowders for mimicking the internal

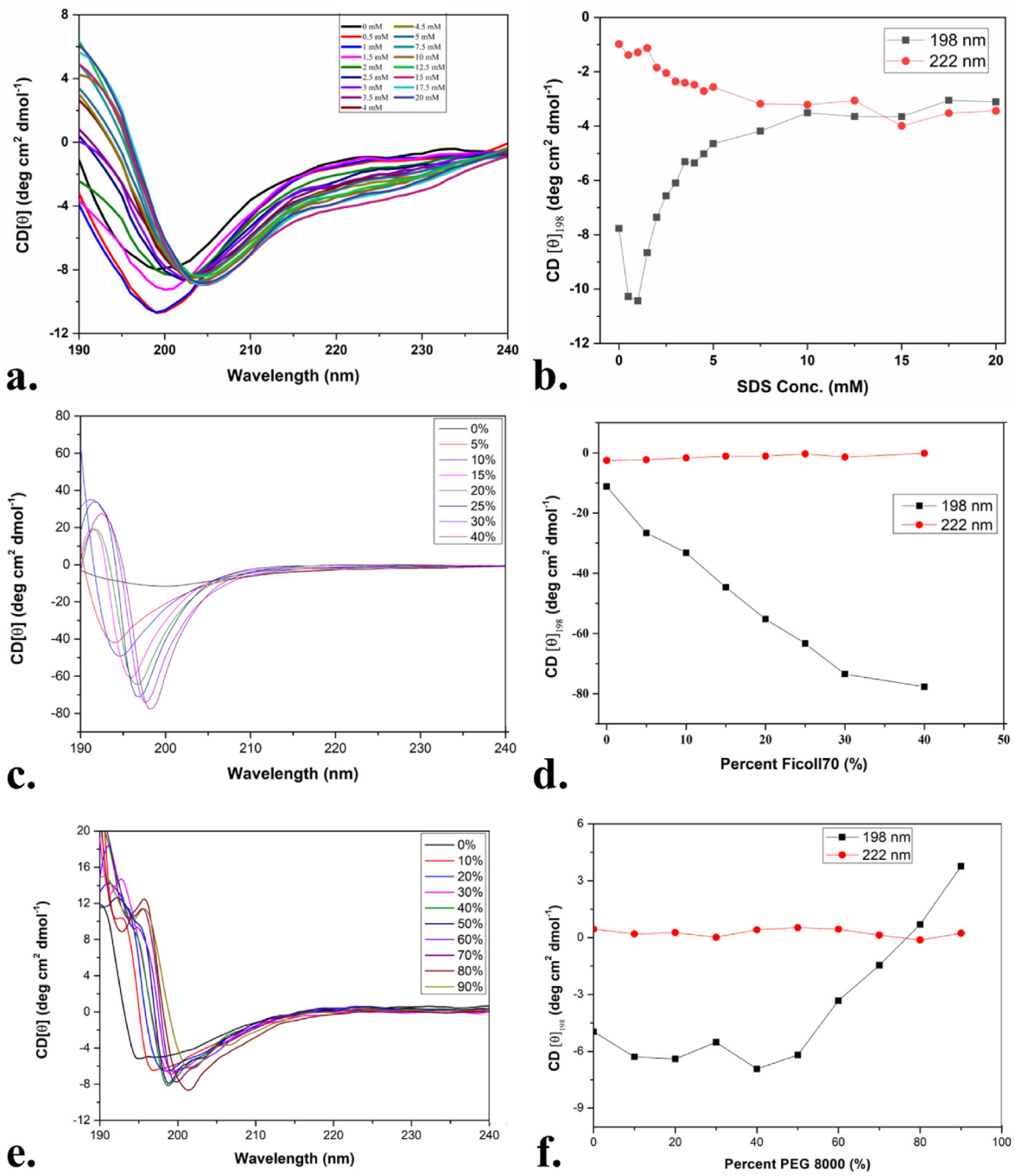


Fig. 4. Effects of SDS on secondary structure transition of TAD region (53–91) of E1A. (a) Far UV-CD spectra of E1A peptide in different concentrations of SDS. (b) Spectra analysis at 198 nm and 222 nm represents the effects of increasing SDS concentration. (c) Far UV-CD spectra of E1A (53–91) with increasing concentration of Ficoll (0–40%) and (d) Spectra analysis at 198 nm and 222 nm represent the effects of increasing concentration of Ficoll 70. (e) Far UV-CD spectra of E1A (53–91) with increasing concentration of PEG (0–90%) and (f) Spectra analysis at 198 nm and 222 nm represents the effects of increasing concentration of crowding agent Polyethylene Glycol (PEG).

cellular environment (Biswas et al., 2018; Kuznetsova et al., 2014). In either of the crowders, no significant changes were observed in the secondary structure of the E1A peptide.

3.5. Prediction of alpha-helix forming molecular recognition elements (MoREs) of E1A peptide

Several studies have revealed the role of disorder to order transitioning regions in molecular recognition and protein-protein interactions. Using online predictor tools, we have analyzed the sequence-based MoRF prediction of E1A peptide to confirm the possible peptide regions undergoing secondary structure transition in presence of different binding partners (Fig. 5). We have utilized ANCHOR, MoRFChibiweb (MCW), and DISOPRED3 to predict the MoRF regions of the E1A peptide. Except the DISOPRED3, which has shown only two short regions of three residues each at both terminals, the MCW and ANCHOR have predicted nearly entire peptide region to be MoRFs. Previously, a MoRF region of seven residues (66–72) have been aligned to HAdV-12 and transcription factors like p53, E2F1, etc. where residues 66th, 68th, and 72nd are found to be conserved among them (Pelka et al., 2008). Therefore, these residues could be involved in helical structural transitions.

3.6. Molecular dynamic simulations

As the structure of the E1A peptide was not available in the PDB in isolation, consequently, the structure was modeled using the *ab-initio* approach using the *Quark* webserver of Zhang lab. In our modeled structure, there are two alpha-helical regions towards N-terminal region. However, the available structure of E1A-TAD bound with CREB binding protein shows only a short helical region and remaining as disordered. Therefore, we have performed the MD simulations of E1A-TAD in isolation which has demonstrated the extreme dynamicity of peptide. In this view, the modeled structure was refined in the water solvent system for the 150 ns. The modeled structure of E1A peptide consisted of residues with helical regions, which were not observed after 150ns of simulation. This showed that the peptide is unstructured or disordered in isolation. Based on the energy of 1500 frames from the 150 ns long trajectory, the highest negative energy clustered structures were found at 105 ns. Therefore, a frame at 105 ns with minimized energy was taken out for further REMD simulations studies.

To further gain the structural conformational insight through breaking the energy barrier at a constant temperature, we performed REMD simulations at six different temperatures, 298 K (25 °C), 311 K (38 °C), 324 K (51 °C), 337 K (64 °C), 350 K (77 °C), and 363 K (90 °C). For this purpose, we have chosen the disordered conformation of the peptide, obtained at 105 ns, as its initial conformation and to be swapped between all these replicas at selected temperatures. As observed in Fig. 6a, the helical strength of the peptide was appeared at 337 K (64 °C) and 350 K (77 °C) and but not at subsequent rising temperatures (363 K or 90 °C), showing the reversible unfolding at extremely high temperature.

As the fluctuation during simulations remains high at the disordered or loop region and low at the ordered (helix and strand) part, likewise the Root Mean Square Fluctuations (RMSF) plots in Fig. 6b show high fluctuations at N-terminal regions of the E1A peptide at all temperatures throughout the simulation as it does not possess any structural conformation while the C terminal region exhibit lower fluctuations due to its helical composition at some temperatures. The RMSF distance of the N-terminal residues was observed as high as 4.5 Å, 2.0 Å, 3.9 Å, 3.8 Å, 4.5 Å, and 5.2 Å for replicas 0 to 5, respectively. Similarly, the Root Mean Square Distance (RMSD) and Radius of gyration (Rg) trends in Fig. 6c and d were also in a similar pattern and show fluctuations at low temperature and gain the highest fluctuation at the highest temperature (363 K).

The embedded table in Fig. 6e shows that the helical composition of the peptide varied at different temperatures. At some temperatures, the peptide regions gain beta-sheet conformations (not visible in last frame snapshots). The helix composition of the peptide is high as the temperature rises, but at extremely high temperatures (90 °C), the helical formation was decreased. The highest 11.27% composition of total secondary structure element containing helix and beta sheets structure in E1A peptide was recorded at 337 K temperature between residues 10–13 and 20–24. Next, the timeline representation and average secondary structure element of residues are shown in Fig. 7 for each replica.

3.7. Aggregation of E1A-TAD peptide

The association of adenoviruses with central nervous system-related diseases is reported (Kim et al., 2004; Schwartz et al., 2019). Several other neurological diseases are linked with misfolding or aggregation of

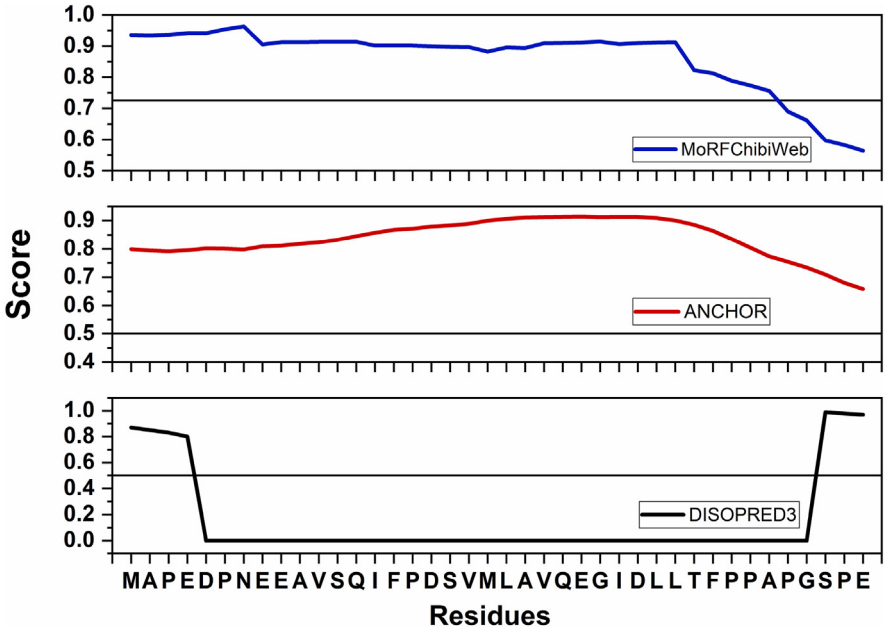
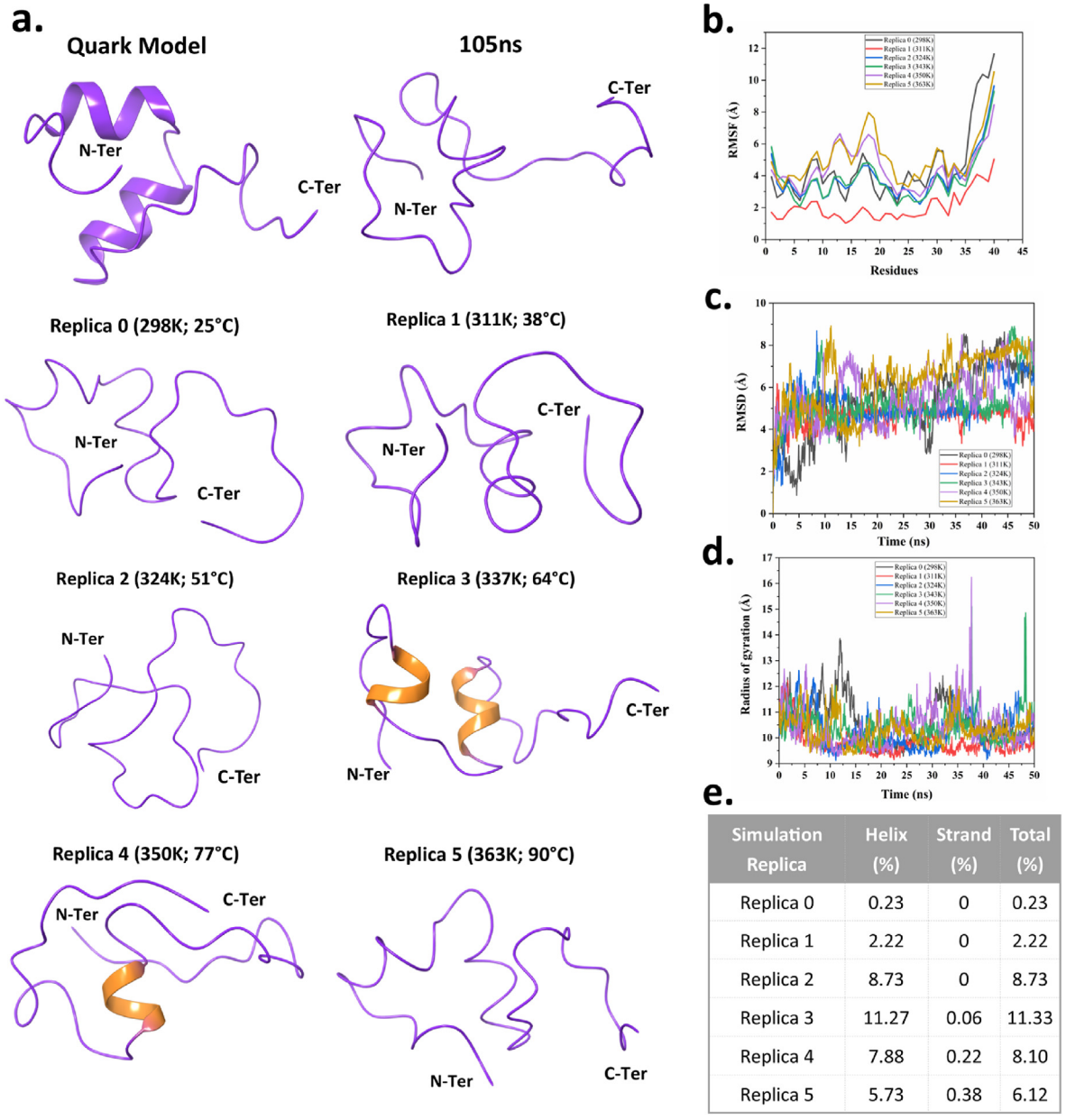


Fig. 5. MoRFs Analysis of TAD region (53–91) of E1A protein using three different predictors, like MoRFChibiWeb, ANCHOR, and DISOPRED3 to predict secondary structure formation propensity of IDR regions.



Simulation Replica	Helix (%)	Strand (%)	Total (%)
Replica 0	0.23	0	0.23
Replica 1	2.22	0	2.22
Replica 2	8.73	0	8.73
Replica 3	11.27	0.06	11.33
Replica 4	7.88	0.22	8.10
Replica 5	5.73	0.38	6.12

Fig. 6. Replica Exchange MD simulations of E1A-TAD peptide: (a.) Structure snapshots of modeled structure (Quark server) and at 105ns obtained through initial MD simulation of E1A peptide model and the initial frame for REMD. Last frame snapshots at 50ns from all replicas (from 0 to 5) are shown with the highlighted helical region in orange (in replica 3 and 4). (b.) Root mean square fluctuation (RMSF), (c.) Root mean square deviation (RMSD), (d.) The radius of gyration (Rg) of all replicas. (e.) The overall percentage of secondary structure elements in the trajectory of each replica is shown in the table. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

proteins such as amyloid-beta, tau, alpha-synuclein, etc. (Ashraf et al., 2014). Additionally, the aggregation of IDPs has been seen in many viral systems, which could also be a reason for pathogenesis. Henceforth, we have investigated the aggregation potential of E1A peptide using different dye-based fluorescence assays, CD spectroscopy, and transmission electron microscopy. In Fig. 8a, we have shown the Thioflavin T (ThT) based fluorescence of aggregated E1A peptide sample compared to ThT control at 490 nm. The aggregation reaction has shown a nearly thirteen-fold increase in fluorescence intensity depicting aggressive aggregation of E1A peptide. According to the ThT dye-based kinetics mapping, the aggregation of E1A peptide followed nucleation-dependent kinetics. Fig. 8b shows that the lag phase continues till ~15 h, and then the growth phase of peptide aggregation lasts up to 50 h, followed by a saturation phase.

Further, the changes in conformation were also observed using non-covalent, Bis-ANS dye-based fluorescence assays. The binding of ANS with aggregated E1A peptide shows a blue shift of 35 nm compared to ANS only (Fig. 8c). Moreover, these structural changes were also confirmed by recording far-UV CD spectra of aggregated samples at different times. As seen in Fig. 8d, an apparent shift was observed from 198 nm to ~212 nm after 24 h, and after 196 h the shift was gradually shifted towards 216 nm, which represents the peptide has been gaining β-sheet conformation after constant stirring. Additionally, the aggregated sample of E1A peptide was prepared for visualization through transmission electron microscopy. A dense cluster of fibrils was visualized, which confirms the formation of aggregates of E1A peptide (Fig. 8e). Lastly, the cytotoxicity of these formed aggregates was examined on SH-SY5Y, a neuroblastoma cell line, using an MTT assay. The SH-SY5Y cells were treated with E1A monomers and aggregates (196 h). The aggregates showed a consistent decline in cell viability upon increasing

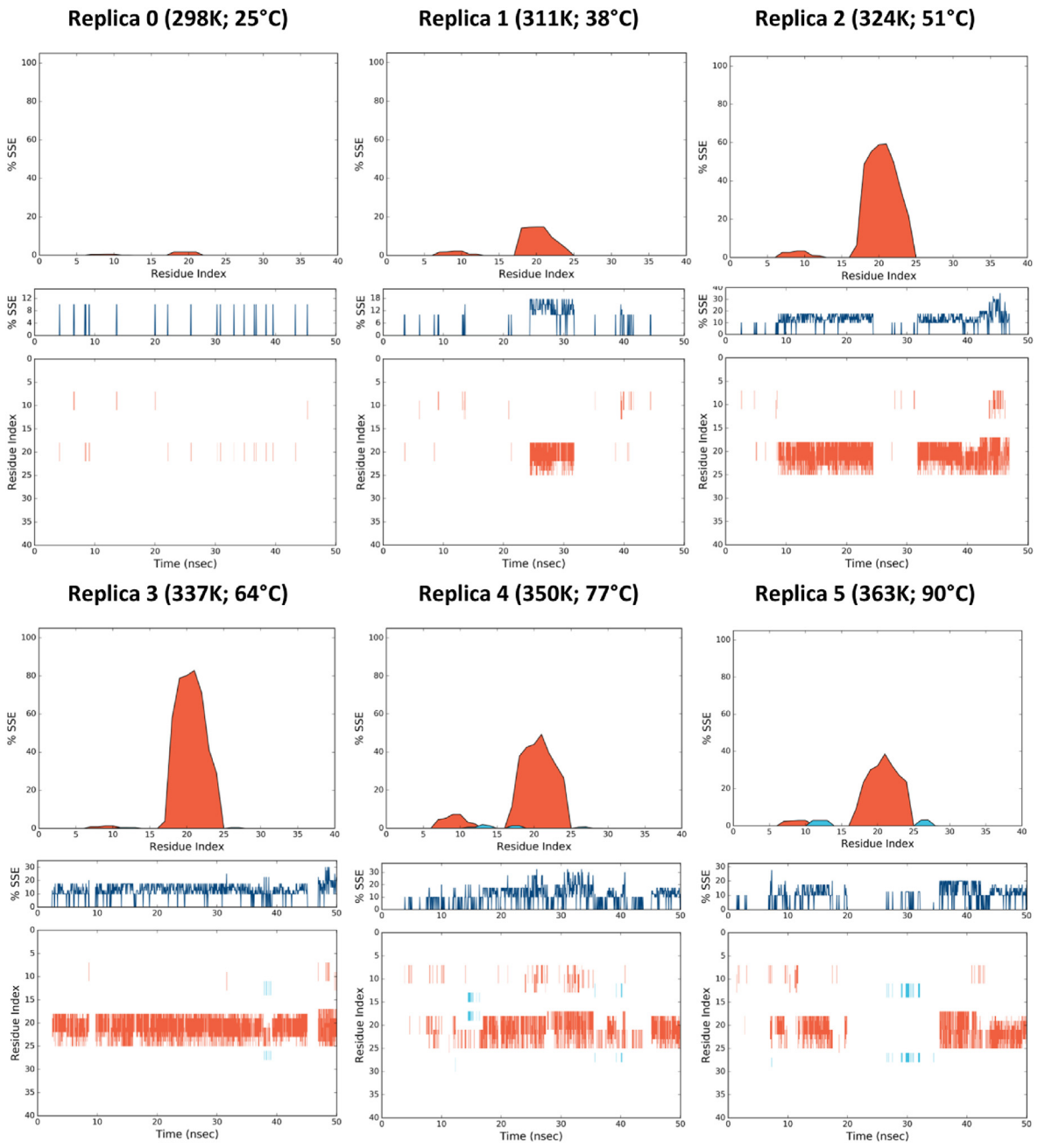


Fig. 7. Percentage and per-residue secondary structure elements (SSE) formation through the timeline representations for 50ns simulation of E1A-TAD peptide at different temperatures in REMD simulations. The red colored regions highlight the presence of alpha-helical propensity while the blue colored region shows the beta-strands formation capability by the residues of the peptide at particular temperature. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

concentration. Upon adding 50 μM of aggregated E1A, the cell viability was reduced to 50% compared to control (Fig. 8f). The mechanism of E1A-TAD aggregates mediated cell toxicity is needs to be done in the future. It also needs to determine how aggregates interacts with cell membrane and whether it enter inside the cells or remain outside in the extracellular space. Our data suggest that the E1A aggregates are more cytotoxic than the monomers. Overall, our observations produce sufficient evidence for proving the aggregation of E1A-TAD and its cytotoxicity to mammalian cells.

4. Discussion

The abundance of IDPs in the viral proteome has led viral proteins to function as intrinsically disordered molecular hub proteins and mimic eukaryotic IDPs' functions in cellular networks (Tompa, 2012; Uversky, 2018; Camilloni et al., 2016). Due to the presence of IDRs, viral proteins are commonly adapted in binding to diverse cellular proteins and are involved in hacking the signaling pathways. IDPs are involved in a wide range of regulatory functions in the cellular environment (Uversky, 2018; Kulkarni and Uversky, 2019). The presence of structural flexibility allows them to attain different conformations in different environmental conditions. The interactions of disordered proteins involved in transcriptional regulation have critically enhanced the information about the expression profiling of the eukaryotic cells (Uversky, 2018; Monti et al., 2021; Uversky et al., 2008). In general, structural transitions in IDPs are challenging to identify and poorly understood. Various interdisciplinary

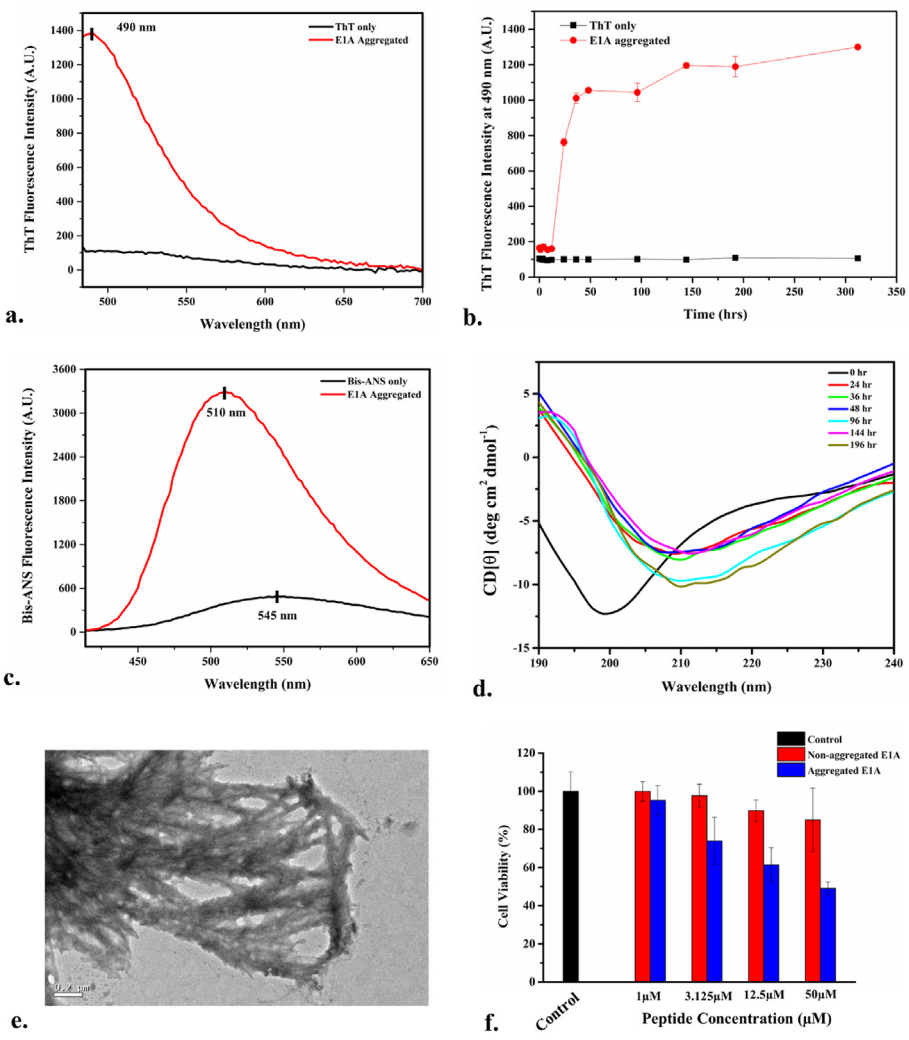


Fig. 8. Aggregation of E1A-TAD peptide. (a) Thioflavin T Assay for E1A-TAD Aggregates. ThT fluorescence spectra were recorded at 450 nm excitation wavelength and emission scan from 484 nm to 700 nm. (b) ThT fluorescence kinetics. ThT fluorescence intensity was measured at 450 nm excitation and 490 nm emission wavelength. (c) ANS binding assay for E1A-TAD Aggregates. ANS fluorescence spectra were recorded at 380 nm excitation and 480 nm emission wavelength. (d) CD Spectra for E1A-TAD aggregation samples were recorded in 1 mm quartz cuvette at different time intervals with the peptide concentration of 20 μM. (e) TEM image for E1A-TAD Aggregates. (f) Cell viability Assay for E1A-TAD monomers (red) and Aggregates (blue). SH-SY5Y cells were treated with E1A-TAD monomers and aggregates and cell viability was measured using an MTT reduction assay. The cells treated with PBS were considered as control. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

approaches have been developed in the last decade to study IDPs structural and dynamical properties and complexes. What signifies the protein-protein interactions involving IDPs can be answered following the prediction of intrinsic structural propensities of IDPs in solution conditions.

E1A has been involved in various protein-protein interactions due to disordered regions present in the native structure state. Additionally, the TAD region of E1A is studied in the presence of CREB-binding protein where it has also shown majorly disordered region and a small helical region (Ferreon et al., 2009). In this work, we have first calculated the propensity of disorderedness in the E1A protein and its TAD region which suggest them to be highly disordered. Following this, we have experimentally and computationally investigated the folding of TAD region of E1A that are likely to undergo helical transitions upon binding by studying the behavior of peptides in different temperatures, solvent mixture, and macromolecular crowding conditions. Further, we have validated our results by analyzing MoRF regions using MoRF predictors (Table 1) and MD simulations. We have shed light on the folding of E1A peptide in a water mixture with different organic solvents like ethanol, TFE, and HFIP. We have also revealed the folding behavior and conformational transitions of E1A in different macromolecular crowding concentrations using CD spectroscopy. MD simulations data of modeled peptide and its REMD simulation analysis supports our experimental observations. Overall, the sequence-based disorder prediction shows high disorder and the E1A-TAD peptide has maintained its structural disorder in most of the condition that we studied. However, in the presence of alcohols such as Ethanol, TFE, and HFIP, the peptide has shown a tendency to gain ordered conformation. Our study has provided the implications of secondary structure elements formation of E1A using experimental and computational analysis.

Protein folding to aggregation is linked with various disease conditions. It has been reported that protein aggregation occurs by changing its structural conformations. Reports suggest that protein aggregation is determined by partially folded intermediates at the early stage in the folding process (Broglia et al., 1998). Moreover, the toxic effects of protein aggregates result from common structural features (Stefani, 2004). In amyloid diseases, specific polypeptide chain loses, or is unable to get its native, closely packed three-dimensional structure, leads to partially folded, unfolded, or non-correctly folded states in equilibrium to each other the content of beta sheet structure is generally increased (Stefani, 2004). In this study, the aggregates formed by E1A may have an important role in the pathogenesis of adenoviral infection that needs to be elucidated by further experiments.

Protein-protein interactions are highly dependent on molecular determinants of recognition and characteristics of interfaces. Various aspects of interactions include residual propensities, pairing preferences, shape, size, hydrophobicity, and solvent accessibility. Previous studies have reported the presence of higher proline content in MoREs. This has shed light on the importance of the presence of Proline residues. Reports suggest that IDPs do not obey a generalized mechanism for coupled folding and binding (Arai et al., 2015). The interactions or binding of IDPs to their binding partners can proceed through various mechanisms

based on their conformational propensities in different environments. It does not depend solely on the specific conformations of IDPs but also varies with the conformational dynamics of their binding partners. Therefore, it can be concluded from the data presented that the binding of IDPs to their target proteins can occur through a variety of mechanisms and depends on the selection from a conformational ensemble of IDPs and the dynamics of its target protein.

CRediT authorship contribution statement

Nitin Sharma: Formal analysis, Writing – original draft. **Kundlik Gadhave:** Formal analysis, Writing – original draft. **Prateek Kumar:** Formal analysis, Writing – original draft. **Rajanish Giri:** Conceptualization, Formal analysis, Writing – review & editing, Supervision, Writing – original draft.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Dr. Rajanish Giri reports financial support was provided by Department of Science and Technology, India.

Acknowledgments

This work has been funded by a DST grant, India (YSS/2015/000613) to RG, and he is also thankful to DBT IYBA award (BT/11/IYBA/2018/06). The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Ahmed, H., Howton, T.C., Sun, Y., Weinberger, N., Belkhadir, Y., Mukhtar, M.S., 2018. Network biology discovers pathogen contact points in host protein-protein interactomes. Nat. Commun. 9, 1–13. https://doi.org/10.1038/s41467-018-04632-8.

Arai, M., Sugase, K., Dyson, H.J., Wright, P.E., 2015. Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. Proc. Natl. Acad. Sci. U. S. A 112, 9614–9619. https://doi.org/10.1073/pnas.1512799112.

Ashraf, G., Greig, N., Khan, T., Hassan, I., Tabrez, S., Shakil, S., Sheikh, I., Zaidi, S., Akram, M., Jabir, N., Firoz, C., Naeem, A., Alhazza, I., Damanhouri, G., Kamal, M., 2014. Protein misfolding and aggregation in alzheimer's disease and type 2 diabetes mellitus. CNS Neurol. Disord. - Drug Targets 13, 1280–1293. https://doi.org/10.2174/1871527313666140917095514.

Avvakumov, N., Wheeler, R., D'Halluin, J.C., Mymryk, J.S., 2002. Comparative sequence analysis of the largest E1A proteins of human and simian adenoviruses. J. Virol. 76, 7968–7975. https://doi.org/10.1128/jvi.76.16.7968-7975.2002.

Biancalana, M., Koide, S., 2010. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochim. Biophys. Acta Protein Proteonomics 1405–1412. https://doi.org/10.1016/j.bbapap.2010.04.001, 1804.

Biswas, S., Kundu, J., Mukherjee, S.K., Chowdhury, P.K., 2018. Mixed macromolecular crowding: a protein and solvent perspective. ACS Omega 3, 4316–4330. https://doi.org/10.1021/acsomega.7b01864.

Bojadzic, D., Buchwald, P., 2018. Toward small-molecule inhibition of protein-protein interactions: general aspects and recent progress in targeting costimulatory and coinhibitory (Immune Checkpoint) interactions HHS public access. Curr. Top. Med. Chem. 18, 674–699. https://doi.org/10.2174/1568026618666180531092503.

Bowers, K.J., Bowers, K.J., Chow, E., Xu, H., Dror, R.O., Eastwood, M.P., Gregersen, B.A., Klepeis, J.L., Kolossvary, I., Moraes, M.A., Sacerdoti, F.D., Salmon, J.K., Shan, Y., Shaw, D.E., 2006. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. ACM/IEEE SC 2006 Conf, p. 43. https://doi.org/10.1109/SC.2006.54. –43.

Broglia, R.A., Tiana, G., Pasquali, S., Roman, H.E., Vigezzi, E., 1998. Folding and aggregation of designed proteins. Proc. Natl. Acad. Sci. U. S. A 95, 12930–12933. https://doi.org/10.1073/pnas.95.22.12930.

C, X., TY, L., D, C., Z, G., Xue, C., Lin, T.Y., Chang, D., Guo, Z., 2017. Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation. R. Soc. Open Sci. 4. https://doi.org/10.1098/rsos.160696.

Camilloni, C., Bonetti, D., Morrone, A., Giri, R., Dobson, C.M., Brunori, M., Gianni, S., Vendruscolo, M., 2016. Towards a structural biology of the hydrophobic effect in protein folding. Sci. Rep. 6, 28285. https://doi.org/10.1038/srep28285.

Colomer, I., Chamberlain, A.E.R., Haughey, M.B., Donohoe, T.J., 2017. Hexafluoroisopropanol as a highly versatile solvent. Nat. Rev. Chem. 1, 1–12. https://doi.org/10.1038/s41570-017-0088.

Culik, R.M., Abaskharon, R.M., Pazos, I.M., Gai, F., 2014. Experimental validation of the role of trifluoroethanol as a nanocrowder. J. Phys. Chem. B 118, 11455–11461. https://doi.org/10.1021/jp508056w.

De Guzman, R.N., Liu, H.Y., Martinez-Yamout, M., Dyson, H.J., Wright, P.E., 2000. Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. J. Mol. Biol. 303, 243–253. https://doi.org/10.1006/jmbi.2000.4141.

Dosztányi, Z., Mészáros, B., Simon, I., Dosztanyi, Z., Meszaros, B., Simon, I., Dosztányi, Z., Mészáros, B., Simon, I., 2009. ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics 25, 2745–2746. https://doi.org/10.1093/bioinformatics/btp518.

Ferreon, J.C., Martinez-Yamout, M.A., Dyson, H.J., Wright, P.E., 2009. Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. Proc. Natl. Acad. Sci. U. S. A 106, 13260–13265. https://doi.org/10.1073/pnas.0906770106.

Frisch, S.M., Mymryk, J.S., 2002. Adenovirus-5 E1A: paradox and paradigm. Nat. Rev. Mol. Cell Biol. 3, 441–452. https://doi.org/10.1038/nrm827.

Gadhave, K., Giri, R., 2020. Amyloid formation by intrinsically disordered trans-activation domain of cMyb. Biochem. Biophys. Res. Commun. 524, 446–452. https://doi.org/10.1016/j.bbrc.2020.01.110.

Gadhave, K., Kapuganti, S.K., Mishra, P.M., Giri, R., 2021. P53 TAD2 domain (38-61) forms amyloid-like aggregates in isolation. bioRxiv 2021. https://doi.org/10.1101/2021.04.09.439126, 04.09.439126.

Giri, R., Morrone, A., Toto, A., Brunori, M., Gianni, S., 2013. Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. Proc. Natl. Acad. Sci. U. S. A 110, 14942–14947. https://doi.org/10.1073/pnas.1307337110.

Glover, K., Mei, Y., Sinha, S.C., 2016. Identifying intrinsically disordered protein regions likely to undergo binding-induced helical transitions. Biochim. Biophys. Acta Protein Proteonomics 1864, 1455–1463. https://doi.org/10.1016/j.bbapap.2016.05.005.

Gopinath, A., Reddy, S.M.M., Madhan, B., Shanmguam, G., Rao, J.R., 2014. Effect of aqueous ethanol on the triple helical structure of collagen. Eur. Biophys. J. 43, 643–652. https://doi.org/10.1007/s00249-014-0994-5.

Huang, B.X., Kim, H.Y., 2012. Effects of ethanol on conformational changes of akt studied by chemical cross-linking, mass spectrometry, and 18O labeling. ACS Chem. Biol. 7, 387–394. https://doi.org/10.1021/cb2003237.

Iešmantavičius, V., Dogan, J., Jemth, P., Teilum, K., Kjaergaard, M., 2014. Helical propensity in an intrinsically disordered protein accelerates ligand binding. Angew. Chem. Int. Ed. 53, 1548–1551. https://doi.org/10.1002/anie.201307712.

Jones, D.T., Cozzetto, D., 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics 31, 857–863. https://doi.org/10.1093/bioinformatics/btu744.

Kim, H.D., Kong, F.K., Cao, Y., Lewis, T.L., Kim, H., Tang, D.C.C., Fukuchi, K.I., 2004. Immunization of Alzheimer model mice with adenovirus vectors encoding amyloid β-protein and GM-CSF reduces amyloid load in the brain. Neurosci. Lett. 370, 218–223. https://doi.org/10.1016/j.neulet.2004.08.059.

King, C.R., Zhang, A., Tessier, T.M., Gameiro, S.F., Mymryk, J.S., 2018. Hacking the cell: network intrusion and exploitation by adenovirus E1A. mBio 9. https://doi.org/10.1128/mBio.00390-18 e00390-18.

Kulkarni, P., Uversky, V.N., 2019. Intrinsically disordered proteins in chronic diseases. Biomolecules 9, 147. https://doi.org/10.3390/biom9040147.

Kumar, A., Kumar, P., Kumari, S., Uversky, V.N., Giri, R., 2020. Folding and structural polymorphism of p53 C-terminal domain: one peptide with many conformations, Arch. Biochem. Biophys. 684, 108342. https://doi.org/10.1016/j.abb.2020.108342.

Kumar, A., Kumar, A., Kumar, P., Garg, N., Giri, R., 2021. SARS-CoV-2 NSP1 C-terminal (residues 131–180) is an intrinsically disordered region in isolation. Curr. Res. Virol. Sci. 2, 100007. https://doi.org/10.1016/j.crviro.2021.100007.

Kumar, A., Kumar, P., Aarthy, M., Singh, S.K., Giri, R., 2021. Experiments and simulation on ZIKV NS2B-NS3 protease reveal its complex folding. Virology 556, 110–123. https://doi.org/10.1016/j.virol.2021.01.014.

Kuznetsova, I.M., Turoverov, K.K., Uversky, V.N., 2014. What macromolecular crowding can do to a protein. Int. J. Mol. Sci. 15, 23090–23140. https://doi.org/10.3390/ijms151223090.

Lin, M.M., Mohammed, O.F., Jas, S., Zewail, A.H., 2011. Speed limit of protein folding evidenced in secondary structure dynamics. Proc. Natl. Acad. Sci. 108, 16622–16627. https://doi.org/10.1073/pnas.1113649108.

Luo, P., Baldwin, R.L., 1997. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. Biochemistry 36, 8413–8421. https://doi.org/10.1021/bi9707133.

Malhis, N., Wong, E.T.C., Nassar, R., Gsponer, J., 2015. Computational identification of MoRFs in protein sequences using Hierarchical application of bayes rule. PLoS One 10. https://doi.org/10.1371/journal.pone.0141603.

Monti, S.M., De Simone, G., Langella, E., 2021. The amazing world of idps in human diseases. Biomolecules 11, 1–3. https://doi.org/10.3390/biom11020333.

Pelka, P., Ablack, J.N.G., Fonseca, G.J., Yousef, A.F., Mymryk, J.S., 2008. Intrinsic structural disorder in adenovirus E1A: a viral molecular hub linking multiple diverse processes. J. Virol. 82, 7252–7263. https://doi.org/10.1128/jvi.00104-08.

Saumya, K.U., Gadhave, K., Kumar, A., Giri, R., 2021. Zika virus capsid anchor forms cytotoxic amyloid-like fibrils. Virology 560, 8–16. https://doi.org/10.1016/j.virol.2021.04.010.

Schwartz, K.L., Richardson, S.E., MacGregor, D., Mahant, S., Raghuram, K., Bitnun, A., 2019. Adenovirus-associated central nervous system disease in children. J. Pediatr. 205, 130–137. https://doi.org/10.1016/j.jpeds.2018.09.036.

Shoemaker, B.A., Portman, J.J., Wolynes, P.G., 2000. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc. Natl. Acad. Sci. U. S. A 97, 8868–8873. https://doi.org/10.1073/pnas.160259697.

SÖnnichsen, F.D., Van Eyk, J.E., Hodges, R.S., Sykes, B.D., 1992. Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. Biochemistry 31, 8790–8798. https://doi.org/10.1021/bi00152a015.

Stefani, M., 2004. Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. Biochim. Biophys. Acta (BBA) - Mol. Basis Dis. 1739, 5–25. https://doi.org/10.1016/j.bbadis.2004.08.004.

Tompa, P., 2012. Intrinsically disordered proteins: a 10-year recap. Trends Biochem. Sci. 37, 509–516. https://doi.org/10.1016/j.tibs.2012.08.004.

Uversky, V.N., 2018. Intrinsic Disorder, Protein–Protein Interactions, and Disease, first ed. Elsevier Inc. https://doi.org/10.1016/bs.apcsb.2017.06.005

Uversky, V.N., Oldfield, C.J., Dunker, A.K., 2008. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu. Rev. Biophys. 37, 215–246. https://doi.org/10.1146/annurev.biophys.37.032807.125924.

Whitmore, L., Wallace, B.A., 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. 32, W668. https://doi.org/10.1093/nar/gkh371.

Wright, P.E., Dyson, H.J., 2015. Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. 16, 18–29. https://doi.org/10.1038/nrm3920.

Xu, D., Zhang, Y., 2012. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins 80, 1715–1735. https://doi.org/10.1002/prot.24065.

Xue, R., Wang, S., Wang, C., Zhu, T., Li, F., Sun, H., 2006. HFIP-induced structures and assemblies of the peptides from the transmembrane domain 4 of membrane protein Nramp1. Biopolym. - Pept. Sci. Sect. 84, 329–339. https://doi.org/10.1002/bip.20478.