

Binding-induced folding of intrinsically disordered peptides

Matthias Heyden^{a,*} and Sthitadhi Maiti^b

^a*School of Molecular Sciences, Arizona State University, Tempe, AZ, United States*, ^b*Department of Physics, University of Genoa, Genoa, Italy*

*Corresponding author: matthias.heyden@asu.edu

8.1 Introduction

Intrinsically disordered proteins (IDPs) are characterized by their absence of a stable three-dimensional structure, primarily because of their unique amino acid composition, often rich in charged residues, glycines, and prolines (Uversky, 2002; Van Der Lee et al., 2014). Despite their structural variability, IDPs play crucial roles in vital physiological functions. They are involved in several fundamental cellular processes, including the regulation of transcription, signal transduction, the formation of multiprotein complexes, and the development of membraneless organelles (Babu, 2016; Wright & Dyson, 2015). Approximately 40% of the human proteome is believed to be disordered or to have lengthy disordered regions (exceeding 100 residues), with a more significant fraction containing disordered segments of up to 30 residues (Emenecker et al., 2021; Holehouse & Kragelund, 2024; Tunyasuvunakool et al., 2021). A key aspect of their functionality is their ability to interact with globular proteins, where IDPs can either adopt a specific conformation when bound or maintain their dynamic and disordered nature (Pancsa & Fuxreiter, 2012). These significant functional regions in IDPs are referred to as molecular recognition features (MoRFs) or molecular recognition elements (MoREs). MoRFs facilitate interactions with structured partner proteins and can experience transitions from disorder to order during interaction (Fatafta et al., 2021; Hadži et al., 2021; Sharma et al., 2019). MoRFs typically vary in size and can be up to 70 residues in length, often situated within larger intrinsically disordered regions (IDRs) (Cumberworth et al., 2013). Because of their flexible structure, MoRFs can accurately interact with their partners, making them crucial players in regulatory processes and signal transduction (Oldfield et al., 2005b).

Based on the specific folded structure they adopt after interacting with globular targets, MoRFs can be classified into four subtypes: α -MoRFs, β -MoRFs, ι -MoRFs, and complex-MoRFs (Cumberworth et al., 2013; Mohan et al., 2006; Van Der Lee et al., 2014). These subtypes are characterized by the formation of α -helices, β -strands, irregular secondary structures, and a combination of multiple secondary structures, respectively.

Among the various types of MoRFs the most prevalent and extensively studied class involves IDPs and IDRs that fold into α -helices upon binding to their targets (Oldfield et al., 2005a; Wright & Dyson, 2009; Yan et al., 2016). This type is known as an α -helix forming MoRE/feature α -MoRE/MoRF (Oldfield et al., 2005a) or simply as a helical binding motif (Carr et al., 1991). Usually,

these sequences are composed of short amphipathic helices that transition from a disordered state to a helical conformation upon binding with their target partners (Dyson & Wright, 2002). These motifs can be predicted based on their sequence (Garner et al., 1999; Mészáros et al., 2009), and they are especially prevalent in eukaryotic transactivation domains (Dyson & Wright, 2016) and $\alpha\alpha$ -hub proteins (Bugge et al., 2021).

Sections of proteins containing α -MoRFs were initially thought to be predominantly disordered when unbound and unrestrained. However, subsequent research has revealed that these regions can exhibit a considerable degree of residual helical structure (Arai et al., 2015). The partially formed helical configurations observed in α -MoRFs are often termed prefolded elements, prestructured motifs, or transient and residual helical structures (Kim & Han, 2018b). A previous bioinformatics study revealed a significant overlap between the residual structure and the structure of the IDP when bound to its target, suggesting that the residual structure predisposes the IDP to its target-bound conformation (Fuxreiter et al., 2004). Even with a limited number of systems available at that time the findings strongly suggested that prestructuring plays a crucial role in target binding. A contemporary review of experimentally analyzed IDPs in their free state documented more than 50 instances of α -MoRFs exhibiting considerable secondary structure, without exploring the correspondence between residual and target-bound conformations (Kim & Han, 2018a).

The transient, residual helical structures are believed to impact the binding affinity (thermodynamics), kinetics, specificity, and structural characteristics of the IDP-target ensemble. There is increasing interest in understanding how these structures influence these properties. An early hypothesis from Dunker et al. (2001) suggested that the disorder-to-order transition in α -MoRFs leads to a reduction in binding affinity because of the folding penalty associated with α -MoRFs. The significance of the folding penalty in IDPs, and particularly in α -MoRFs, is now well-established and has been thoroughly reviewed in the literature, as discussed in Skriver et al. (2023). The initial analysis of binding affinities from various studies revealed that the average affinity for associations involving folding upon binding is generally lower compared to that of globular complexes, indirectly suggesting an estimate of the folding penalty in such interactions (Lazar et al., 2022; Teilum et al., 2015). In a study conducted by Teilum et al. (2015), which examined a substantial number of protein complexes using experimental data, it was found that the less favorable entropic contribution in complexes involving ordered regions (helical in our case) and IDPs is primarily caused by a reduction in conformational entropy. This finding is consistent with the fundamental difference in the binding mechanisms of ordered and disordered ligands. A disordered polypeptide needs to undergo folding to create the final complex, resulting in a significant decrease in conformational entropy. It should be noted that equilibrium isothermal titration calorimetry data cannot definitively determine the pathway followed by the ligand's folding during the binding process. Therefore the contribution of the amount of folding free energy to the observed binding free energy is uncertain. It is likely that in some protein complexes, the IDP undergoes folding followed by binding in a conformational selection mechanism, whereas in others, the IDP folds upon binding in an induced fit mechanism. However, the variation in average Gibbs free energy ΔG° could still be explained by the necessary folding of the disordered ligand in the ordered-disordered complexes. The assessment of folding penalties has been conducted through a thermodynamic breakdown of entropic factors (Drobnak et al., 2013; Theisen et al., 2021) and the application of helix-coil models (Hadži & Lah, 2022). In addition, the alteration in the helicity of α -MoRFs upon binding to their targets can provide estimates of the folding penalty (Hadži & Lah, 2022). A

thorough evaluation of 25 distinct α -MoRFs using the helix-coil approach has indicated that the average folding penalty stands at around 2.5 kcal/mol. Nonetheless, numerous instances of ultratight picomolar binding affinities in folding-upon-binding interactions have been documented (Zavrtanik et al., 2021), demonstrating that multiple factors such as posttranslational modifications, phosphorylation, oxidative stress, or initial binding to certain regions of the binding partners can effectively counteract the folding penalty to attain ultrahigh binding affinity (Flock et al., 2014). A key factor in overcoming the folding penalty is the presence of substantial amounts of residual helical structure, also referred to as prestructuring or prefolding. An examination of folding penalties in α -MoRFs revealed that prefolding can lower the penalty by an average of 25% (Hadži & Lah, 2022). Experimentally a reduction in the α -MoRF folding penalty can be seen when an enhancement in residual helicity (such as that induced by mutations) leads to an elevated binding affinity of the α -MoRF. Relationships between residual helicity and binding affinity have been documented across various systems, and they are in quantitative agreement with predictions from models (Hadži & Lah, 2022; Hadži et al., 2021), underscoring the significance of residual structure in influencing the affinities between IDPs and their targets. Nevertheless, some research has identified a less pronounced correlation or even a more intricate relationship between residual helicity and binding affinity (Poosapati et al., 2018; Rogers et al., 2014).

In multiple distinct systems exhibiting folding-upon-binding, experimental studies have shown that the residual helical structure plays a role in modulating both binding affinity and kinetics. However, further insights are needed into how the IDP sequence relates to the structural ensemble (both unbound and bound) and how this affects the associated binding parameters, including both kinetic and equilibrium aspects. Specifically the measurement of folding free energy using enhanced sampling techniques like umbrella sampling (US) has not been extensively documented because of challenges in defining an appropriate collective variable (CV) to describe the unfolding pathway. In addition, the quantitative effects of point mutations involving known helix stabilizers like alanine and helix disruptors such as proline and glycine on the overall folding free energy have yet to be reported in detail.

To address these unexplored pathways, we employed atomistic US to explore the folding free energies in disordered peptides that transition to an α -helix upon target binding, such as peptides containing α -MoRF, as well as fully disordered IDRs. After comparing the free energy minimum structures from simulation and from AlphaFold (Jumper et al., 2021) predictions, all the α -MoRF examples we studied exhibited folding free energies within a range of 10 kJ/mol (2.4 kcal/mol), which is consistent with the range predicted by Hadži and Lah (2022) in their study using a modified Lifson–Roig (LR) model (Chakrabarty et al., 1994; Hadži et al., 2021; Lifson & Roig, 1961). They considered different sequence lengths of IDPs, which is one of the factors contributing to the discrepancy between our results and theirs. The relatively low folding free energy penalty of these sequences can be implemented in designing peptide-based or small molecule drugs that stabilize the helix structure by inducing folding upon binding to α -MoRFs.

Furthermore, we calculated the binding free energies of two selected α -MoRFs from BH3 domain-binding proteins (PUMA and NOXA-A) to MCL-1 using coarse-grained (CG) simulations with the SIRAH force field. Our findings partially support our hypothesis that the experimental binding free energies reported by Dahal et al. (2018) are the sum of both the computed binding and folding free energies. This result offers valuable insights into the potential binding-induced folding pathway followed by our chosen set of α -MoRFs.

8.2 Methods

8.2.1 Simulation protocol for folding free energy calculations

We performed all simulations of folding free energies using GROMACS 2018.1 software package (Abraham et al., 2015) and the PLUMED library (Bonomi et al., 2009, 2019). For system preparation, we constructed fully α -helical peptide sequences. We generated a template Protein Data Bank (PDB) file of a helical polyalanine with the desired sequence length. The template file was then used to create a PDB file of the sequence with alanines replaced by other amino acids according to the sequence. We subsequently employed the SCWRL4 package (Krivov et al., 2009) to generate the side chains, thereby generating the initial helical state of the sequence. We adjusted the protonation state of each amino acid residue to correspond with a pH of 7, utilizing standard side chain pK_a values in solution.

Once the system was prepared, we carried out steered molecular dynamics (SMD) simulations (Lu & Schulten, 1999) on the helical peptide. This was performed to generate snapshots or starting configurations for further analysis. The peptide was placed in a cubic simulation box with an edge length of 80 Å, and it was surrounded by explicit TIP3P water molecules to mimic an aqueous environment. To ensure the stability of the initial conformation of the peptide, we applied a force constant of 1000 kJ/mol to all atoms to maintain its initial configuration during equilibration. Next, we solvated the system with approximately 16,700 water molecules to ensure adequate hydration of the peptide. In addition, we introduced ionic species to the system to imitate a physiological salt concentration of 150 mM. We added close to 50 Na^+ and Cl^- ions. It is important to note that the exact numbers of these ions were adjusted based on the specific peptide sequence being studied, as different peptides may require different ionic concentrations to maintain charge neutrality and mimic physiological conditions. The presence of salt can affect the structural dynamics and phase separation tendencies of IDPs through mechanisms like electrostatic screening and the salting-out effect, as highlighted in a recent study by Wohl et al. (2021). This setup was aimed to create a realistic simulation environment for the helical peptide in dilute solution. Although it falls short of resembling the crowded biomolecular environment *in vivo* (Gnutt et al., 2015), it reflects the conditions of many experiments.

We implemented periodic boundary conditions across all dimensions to simulate an infinite system and avoid edge effects. For handling long-range electrostatic interactions, we employed the particle mesh Ewald algorithm (Darden et al., 1993), setting up a real-space grid with a resolution of 1.2 Å. This method efficiently calculates electrostatic forces by dividing them into short-range and long-range components. For short-range electrostatic and Lennard–Jones interactions, we established a cutoff distance of 10 Å. The linear constraint solver algorithm (Hess et al., 1997) was used to constrain high-frequency bond lengths such as NH, OH, and CH within the protein, allowing us to use a longer time step of 2 fs in our case when integrating the equations of motion. In addition, we employed the SETTLE algorithm (Miyamoto & Kollman, 1992) to preserve the geometry of water molecules in the system. We performed an energy minimization process using the steepest descent algorithm before starting the molecular dynamics (MD) simulations with all the selected force fields. This initial step, involving 2000 steps of minimization, helps in relieving any steric clashes or unfavorable conformations in the system, providing a stable starting point for the dynamic simulations.

We then conducted a two-step equilibration process for each system. Initially, we equilibrated the system in the isothermal–isochoric (NVT) ensemble, applying a position restraint of 1000 kJ/mol/nm²

to each heavy atom to maintain structural stability. Subsequently, we transitioned to the isothermal–isobaric (NPT) ensemble, maintaining the system at a temperature of 300 K and a pressure of 1 bar for 1 ns. During this phase, we utilized a time step of 1 fs, a velocity rescaling thermostat with a time constant of 0.1 ps for temperature control, and a Berendsen weak-coupling barostat with a time constant of 2.0 ps to regulate pressure.

To represent the folding process of the helix during SMD, we selected a single CV, $\xi(R)$ that quantifies the number of hydrogen bonds between backbone atoms of residues that are four sites apart along the sequence (i.e., $i, i + 4$ hydrogen bonds).

$$\xi(R) = \sum_{i=1}^{n-4} \frac{1 - (r_{i,i+4}/d_0)^n}{1 - (r_{i,i+4}/d_0)^m} \quad (8.1)$$

Here the variable R represents the microscopic coordinates of the system. This particular CV definition was previously introduced by Camilloni et al. (2008) in their study to effectively capture the folding dynamics of helical structures. In enhanced sampling simulations that involve dynamics, that is, integration of equations of motion based on forces, computing biasing forces requires CVs that are continuous and differentiable. The parameters in Eq. (8.1) were set to $n = 6$, $m = 10$, and $d_0 = 0.32$ nm. The variable $r_{i,i+4}$ represents the distance between the backbone oxygen of residue i and the backbone nitrogen of residue $i + 4$. This CV does not account for angular contributions. Therefore conformations where an oxygen atom at site i is near a nitrogen atom at site $i + 4$ will contribute to the CV even if their positions do not meet common angular criteria for a hydrogen bond (e.g., the angle $\theta_{\text{N-H-O}}$ being less than 30 degrees). Incorporating an angle cutoff would, on one hand, enhance the accuracy with which the CV depicts the formation of helices. However, it would also introduce a region of the phase space with a zero gradient of the CV with respect to the microscopic atomic coordinates. Notably the expression in Eq. (8.1) does not require specific knowledge of the folded structure. For each system, we conducted 100 ns SMD simulations in the NPT ensemble at 300 K and 1 bar, using a time step of 2 fs. To maintain temperature and pressure, we employed a Nose–Hoover thermostat (Evans & Holian, 1985) and a Parrinello–Rahman barostat (Parrinello & Rahman, 1981), both with a time constant of 2.0 ps. These simulations were aimed at generating snapshots of the unfolding process, transitioning from a fully α -helical state with a nonzero CV value to an unfolded state with a ξ value close to zero. The PLUMED library was utilized to evaluate the CV and compute the biasing forces.

We generated 61 starting structures using SMD. Subsequently, we conducted US simulations for each of these structures. In the US method (Kästner, 2011) a biasing potential is applied at specific positions along the reaction coordinate (in this case the CV representing the hydrogen bond number, ξ) to enhance the sampling of states with high free energy barriers. The reaction coordinate was segmented into 61 windows, each corresponding to one of the starting structures. A harmonic potential, as detailed in Eq. (8.2), was added to the original potential (unbiased potential) in each window. This was done to facilitate the exploration of the CV space by the protein.

$$\omega_i(\xi) = \frac{1}{2} k [\xi(R) - \xi_i]^2 \quad (8.2)$$

In this expression, ξ_i serves as the reference coordinate for the i th US window. During the simulation, any deviation of the CV from its reference value results in an added bias potential, which in turn exerts a force to push the system back toward the reference position, ensuring enhanced sampling around

this target coordinate. Here, k is the force constant, and R describes the microscopic coordinates of the system as in Eq. (8.1). We employed the weighted histogram analysis method (WHAM) (Kumar et al., 1992) to remove the bias from the restrained US simulations and to construct the potential of mean force (PMF) using 400 bins along the CV.

8.2.2 Simulation protocol for binding free energy calculations

As test systems, we chose two α -MoRF BH3-only domain proteins of PUMA (PDB ID: 2ROC) and NOXA-A (PDB ID: 2ROD), where each is bound in a helix conformation to the antiapoptotic protein MCL-1. BH3-only domain proteins act as proapoptotic agents in many organisms, including humans. Each test complex underwent identical preparation steps, starting with: (1) obtaining the experimental structures from the PDB website (<http://www.rcsb.org>); (2) utilizing PyMOL (Yuan et al., 2017) for editing and removing additional chains from the PDBs to create interacting protein–ligand systems; and (3) filling all the missing residues or atoms in the PDB file using MODELER v9.15 (Webb & Sali, 2016).

CG simulations for each of the three systems were carried out using the SIRAH force field (Darré et al., 2015). The SIRAH CG force field has been designed with the intent to overcome several limitations commonly found in CG force fields, such as the use of a uniform dielectric constant, the absence of long-range interactions, reliance on topological information for maintaining secondary structures, and limited consideration of ionic strength effects. Unlike the popular Martini force field (Marrink & Tieleman, 2013; Periole & Marrink, 2013), which typically maps four heavy atoms to one CG bead, SIRAH maintains a higher level of detail in the peptide bonds, retaining the positions of backbone nitrogen (N), α -carbon (C_α), and oxygen (O) atoms, whereas the side chains are represented more coarsely. The WT4 water model incorporated in SIRAH consists of four interconnected beads with partial charges, enabling the simulation of dielectric permittivity. Furthermore the CG electrolytes in SIRAH can replicate effects of ionic strength and osmotic pressure. The resulting CG Hamiltonian features only terms that are supported in most MD simulation packages.

Before using SIRAH for US, we generated 91 windows using Visual Molecular Dynamics (VMD) (Humphrey et al., 1996) software. As CV, we defined d_{COM} , the distance of separation of the two centers of masses (COM) between the MCL-1 and BH3-only domain proteins. With a window spacing of 0.05 nm, each receptor–ligand pair was separated to a maximum d_{COM} of 4.5 nm to allow sufficient separation, complete solvation, and avoid residual interaction at maximum separation. Next, we applied SIRAH CG simulations to perform US simulations for each window. We adhered to the protocol described in Darré et al. (2015). The mapping of coordinates and analysis were executed using the SIRAH tools (Machado & Pantano, 2016). Before transitioning to the SIRAH CG model, protonation states were determined assuming a neutral pH, utilizing the PDB2PQR server (Dolinsky et al., 2004) and selecting the AMBER (Hornak et al., 2006) naming scheme for the output. After converting from all-atom to CG representations the protein complexes were positioned in a cubic box of 8 nm filled with SIRAH WT4 water. To achieve a neutral charge, Na^+ and Cl^- ions were added at a concentration of 0.15 mol/L. The MCL-1 complexes with PUMA and NOXA-A were solvated with 4995 and 4991 SIRAH WT4 water molecules, respectively. In both cases a minimum distance of 0.3 nm was maintained between all water molecules and the closest protein CG bead. The corresponding periodic box dimensions were chosen to guarantee that the two proteins do

not interact with their own periodic images at their maximum separation distance. Each system was initially minimized using the steepest descent method for 10,000 steps. To allow for the equilibration of water molecules around the protein complex, each system was then subjected to a 5 ns simulation with the positions of all CG protein atoms in the complex harmonically restrained. During these restrained simulations the temperature of the systems was maintained at 300 K and the pressure at 1 atm using the V-rescale thermostat (Bussi et al., 2007) and the Parrinello–Rahman barostat (Parrinello & Rahman, 1981) with isotropic pressure coupling, respectively. Following this, US simulations without position restraints were carried out for a duration of 15 ns for equilibration and 100 ns for sampling.

All simulations were conducted with a time step of 20 fs and updated neighbor lists every 10 steps. Electrostatic interactions were calculated using the particle mesh Ewald method (Darden et al., 1993) with a direct cutoff of 1.2 nm and a grid spacing of 0.2 nm. For van der Waals interactions a cutoff of 1.2 nm was employed. In each umbrella window a harmonic restraint potential with a force constant of 500 kJ/mol/nm² was applied to the center of mass distance between the proteins.

Prior to initiating the US simulations, we implemented distance restraints on the backbone atoms of MCL-1 and PUMA separately. This was done to preserve a stable CG structure while applying the biasing potential during US, as illustrated in Fig. 8.1. The elastic network model (ENM) restraints (Bahar et al., 1997; Haliloglu et al., 1997; Hinsen, 1998; Mishra & Jernigan, 2018; Tirion, 1996; Togashi & Flechsig, 2018) were applied between O–N, C–N, C–O, and C–C atom pairs within a 5.25 Å radius. As a result the root-mean-square deviations of both MCL-1 and PUMA were maintained within a range of 5–6 Å, ensuring the structural integrity of the proteins during the simulations. To remove the bias from the simulations, we utilized the WHAM and constructed the PMFs using 400 bins along the protein center-of-mass distance used to define the CV. The final one-dimensional WHAM profile was adjusted to account for the missing Jacobian contribution (Trzesniak et al., 2007) to obtain the accurate PMF (Chodera & Mobley, 2013; Herschbach et al., 1959; Hub et al., 2010; Khavrutskii et al., 2008; Qian, 2006).

$$\text{PMF}(R) = \text{PMF}_{\text{WHAM}}(R) + k_B T \ln(4\pi r^2) \quad (8.3)$$

This adjustment provided in Eq. (8.3) ensures a flat PMF at large distances by compensating for the entropic increase in the number of configurations on a sphere of a given radius (Hub et al., 2010). The binding free energy, $\Delta_{\text{bind}}G$, for a protein complex was calculated as the difference in the PMF between the bound and unbound states.

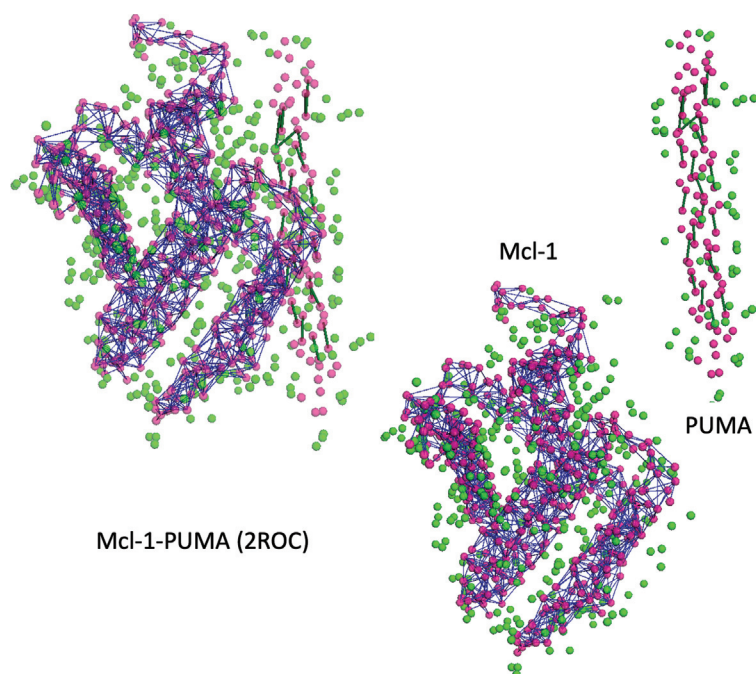
$$\Delta G_{\text{bind}} = -k_B T \ln \left[\frac{P(\xi_{\text{bound}})}{P(\xi_{\text{unbound}})} \right] \quad (8.4)$$

where $P(\xi_{\text{bound}})$ and $P(\xi_{\text{unbound}})$ represent the probabilities of states on the PMF associated with the bound and unbound states along the CV, that is, the center-of-mass distance.

8.3 Results and discussion

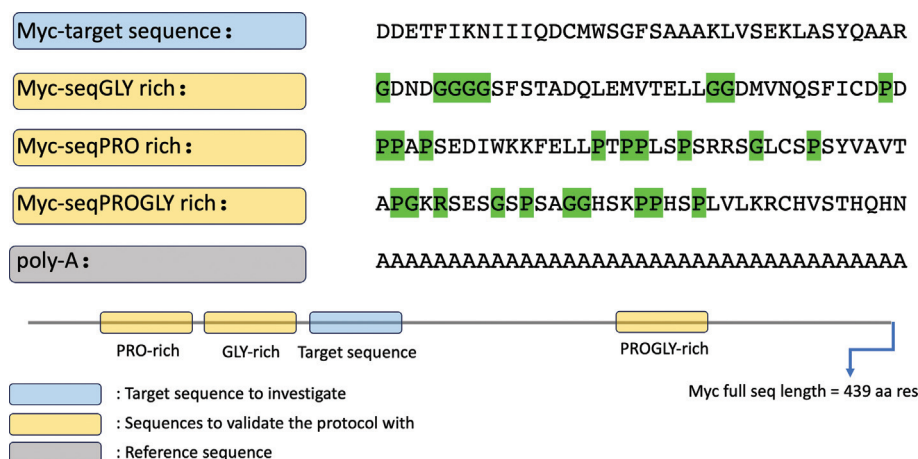
8.3.1 Folding free energy penalties

To confirm the accuracy of our method for quantitatively estimating the folding free energies of disordered regions transitioning from coiled to α -helical states, we conducted a set of analyses. We

**FIGURE 8.1**

Elastic network models (ENMs): ENM applied to the backbone atoms of MCL-1 and PUMA individually to maintain a stable coarse-grained structure while applying biasing potential during umbrella sampling. The distance restraints are applied between O–N, C–N, C–O, and N–O atoms within 5.25 Å radius. The *pink* and *green* spheres represent backbone and side chains, respectively.

focused initially on a 37-amino-acid segment of the c-Myc protein for this evaluation. This particular length was selected because it represents the longest disordered sequence identified as having the potential to fold into an α -helical structure. The c-Myc protein (Battey et al., 1983; Beaulieu et al., 2020; Dang, 2012; Jha et al., 2023; Madden et al., 2021; Meyer & Penn, 2008), with a molecular weight of 62 kDa and composed of 439 amino acid residues (<https://www.uniprot.org/uniprotkb/P01106>), contains multiple functional domains. These domains enable c-Myc to interact with a wide variety of binding partners, each contributing uniquely to the intricate network of interactions known as the c-Myc interactome. The transcription factor c-Myc is overexpressed in more than 70% of human cancers, both constitutively and abnormally. Studies have demonstrated that directly inhibiting c-Myc can lead to rapid tumor shrinkage in mice, accompanied by only mild and completely reversible side effects. This indicates that targeting c-Myc could be a feasible therapeutic approach. Although it does contain IDRs that contribute to its flexibility and interaction with various binding partners, c-Myc also has structured domains, such as the basic helix-loop-helix/leucine zipper domain, which is crucial for DNA binding and dimerization. The presence of both structured and disordered regions allows c-Myc to function as a versatile transcription factor involved in the regulation of numerous genes.

**FIGURE 8.2**

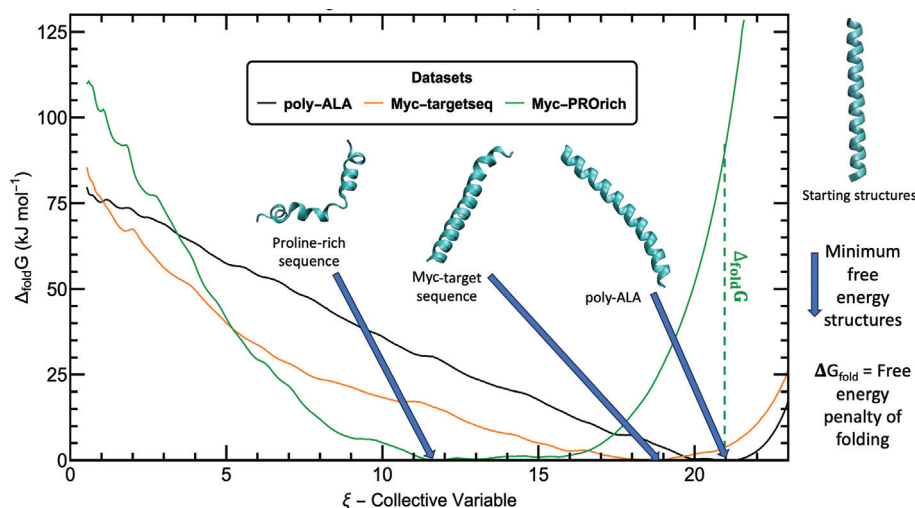
Sequences: Selected sequences to examine for validation of protocol. The 37-sequence polyalanine acts as a reference with a nearly intact α -helix structure.

We primarily chose four different sequences from the entire c-Myc protein sequence for our study. The first sequence was identified as a potential target because of the presence of several adjacent alanines and leucines known to stabilize helices. Moreover, AlphaFold's predicted structure for c-Myc (<https://www.uniprot.org/uniprotkb/P01106>) shows this region to be partially folded into an α -helix. The other three sequences were characterized by their high content of glycine, proline, or a combination of both. The residues of glycine and proline are known to destabilize α -helical structures in solution (Chou & Fasman, 1974; Lyu et al., 1990; O'Neil & DeGrado, 1990; Pace & Scholtz, 1998). For membrane proteins, however, they were found to stabilize the helical structures instead (Dong et al., 2012). As a reference sequence, we selected a polyaniline chain of 37 residues in length. This choice was based on its ability to maintain the highest α -helicity among all amino acids when simulated using the Amber03 force field (Best & Hummer, 2009; Duan et al., 2003) and the TIP3P water model (Jorgensen et al., 1983) in a 100 ns atomistic classical MD simulation. According to the literature, alanine is known to be the most stabilizing amino acid for α -helical structures, followed by leucine (Lyu et al., 1991; Pace & Scholtz, 1998; Rohl et al., 1999; Spek et al., 1999).

Fig. 8.2 presents the sequences we initially tested our protocol on. These sequences are located within the 439-residue chain of the c-Myc protein. The green highlighted residues are helix destabilizers present in the chosen sequences. These sequences were specifically chosen to deviate from the α -helical conformation. This deviation was intended to assist us in measuring the folding free energies with respect to a fully α -helical polyaniline reference.

Having determined the initial set of sequences for validating our protocol, we proceeded with the calculations of folding free energies. These calculations were carried out using US, as detailed in Section 8.2.

The resulting PMFs presented in Fig. 8.3 illustrate the unfolding process of a structure that begins in a fully α -helical conformation and transitions to an unfolded, coiled state. In this representation the

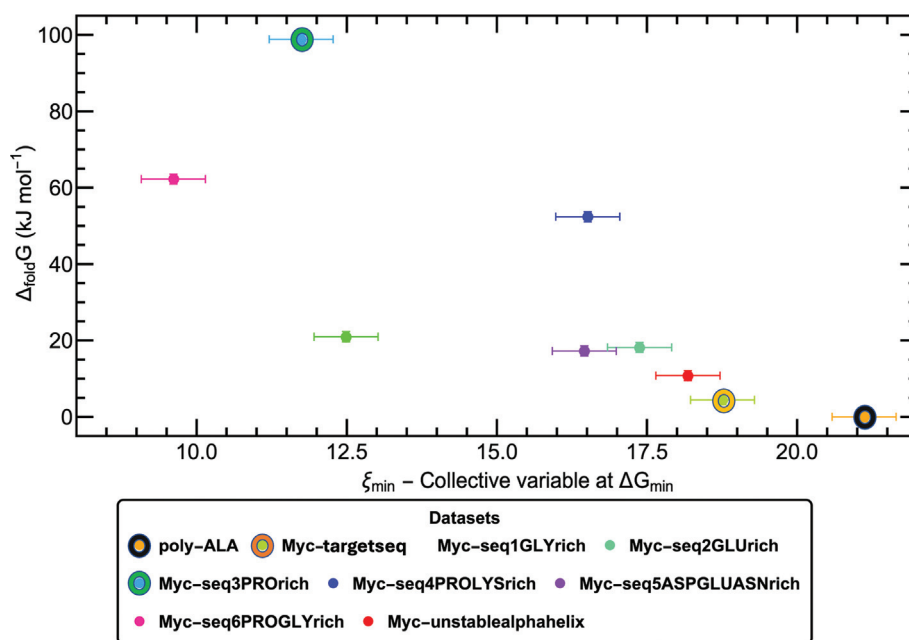
**FIGURE 8.3**

Potential of mean force (PMF): PMFs for the formation of α -helical secondary structures obtained via umbrella sampling simulations. Each structure shown as an inset corresponds to the corresponding free energy minimum (as indicated by arrows) along the folding pathway from coil to helix. The folding free energy penalty for a given sequence, ΔG_{fold} , can be estimated by the PMF at the CV that corresponds to the free energy minimum of the polyaniline reference.

CV, Collective variable.

reference structure of polyaniline is shown to maintain the closest resemblance to an α -helix. This is followed by our target sequence from the Myc protein and then by the sequence that is rich in prolines, which is expected to destabilize α -helices. The minimum of the PMF for polyaniline (all PMFs are shifted to yield a potential of zero for the respective minimum) represents a fully α -helical reference state at a value of ξ close to 21. The PMF of any other sequence at $\xi = 21$, ΔG_{fold} represents an estimate of the folding free energy penalty associated with transitioning from a disordered state to an α -helical state. This analysis helps us quantify the stability of α -helical structures and the energetic cost of folding for different sequences within the c-Myc protein.

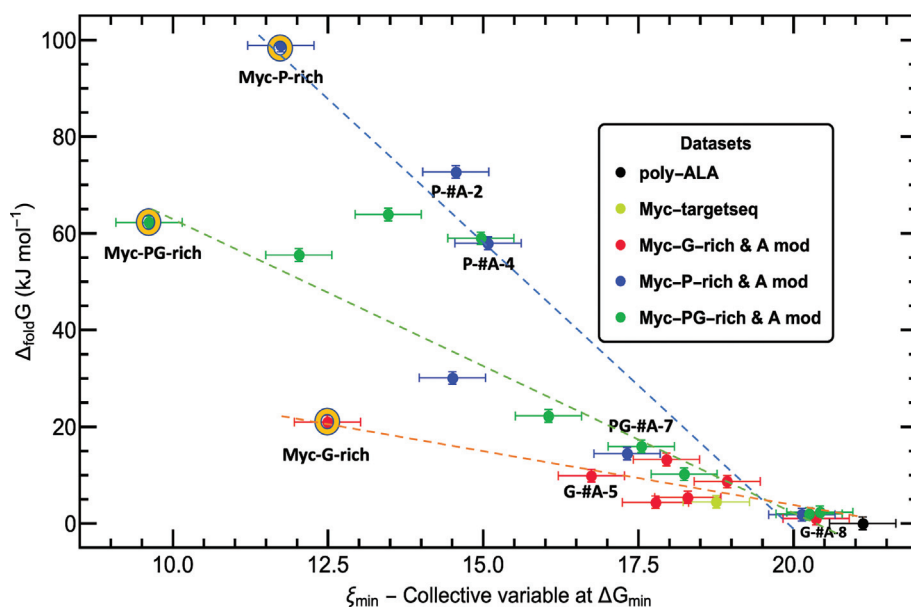
In Fig. 8.4 the folding free energy penalties and the position of the folding free energy minimum along the CV are illustrated for different segments of the c-Myc protein, including our target fragment. The free energy penalty for folding our target c-Myc fragment into an α -helical structure is estimated to be around 5 kJ/mol. This value is significantly lower compared to the energy barrier of ≈ 98 kJ/mol observed for other segments rich in proline residues, which are known to destabilize α -helices. The relatively low folding free energy penalty for our target c-Myc fragment suggests that the folded conformation can be easily stabilized by favorable interactions upon binding. This characteristic makes it a promising candidate for drug development, as it implies that the segment could readily fold into an α -helix upon binding to a suitable ligand. This property is crucial for the binding-induced folding mechanism, where the interaction with a drug molecule triggers the folding of the target protein segment into a specific structure, in this case, an α -helix. Thus the favorable folding energetics of our target c-Myc fragment highlight its potential as a drug target sequence for

**FIGURE 8.4**

Folding free energies versus free energy minimum: Scatter plot of folding free energy penalties, $\Delta_{\text{fold}} G$, as a function of the position $\xi_{\min}$ of the free energy minimum along the coil-to-helix folding pathway for a collection of chosen fragments within the c-Myc sequence. In addition to the target sequence and the GLY and PRO-rich sequences illustrated in Fig. 8.2, we include data from c-Myc sequence fragments rich in other amino acids as indicated. As a reference, we include a sequence fragment of c-Myc that does not contain GLY or PRO residues but was predicted as unstructured by AlphaFold, that is, not to form a stable α -helix.

therapeutic interventions aimed at modulating the function of the c-Myc protein in cancer and other diseases.

Following our investigation of the target c-Myc fragment, we aimed to examine the impact of alanine point mutations on the stability of the α -helical conformations. Given that alanine is known to stabilize α -helices, substituting residues in a sequence with alanines should reduce the folding free energy penalty. To test this rationale, we introduced alanine mutations into sequences rich in glycine, proline, or both. The results of these mutations are presented in Fig. 8.5. We anticipated that replacing destabilizing residues with alanines would lead to a decrease in the folding free energy penalty, thereby increasing the stability of the α -helical structure. The observed trend in our results aligns with our hypothesis. We noticed a significant reduction in the free energy penalty for the proline-rich sequence upon mutation of proline residues to alanine. This reduction is expected because proline, with its rigid structure, impedes the formation of α -helices in solution by enthalpically destabilizing the helical structure. On the other hand the glycine-rich sequence, which is entropically unfavorable for α -helix formation, exhibited a smaller decrease in the folding free energy penalty upon mutation to alanine. This is because of the intrinsic flexibility of glycine, which contributes to a higher entropy in the unfolded state. The sequence

**FIGURE 8.5**

Helical folding propensities of alanine mutants: Scatter plot comparing the potential of mean force (PMF) features for polyalanine and the selected peptide sequences. A reduction in the folding free energy penalty, denoted as $\Delta_{\text{fold}}G$, is observed with increasing numbers of point mutations of destabilizing residues (PRO and GLY) in each sequence to ALA. Both the position of the free energy minimum along ξ_{min} and the folding free energy penalty gradually shifts toward the corresponding observations for polyalanine with increasing numbers of stabilizing mutations.

rich in both proline and glycine demonstrated an intermediate slope in the reduction of the folding free energy penalty upon mutation to alanine. This intermediate behavior reflects the combined effects of proline's enthalpic destabilization and glycine's (weaker) entropic destabilization of the α -helix.

To establish a general trend for the folding free energies, we conducted US on sequences of 37 residues in length for every naturally occurring polyamino acid peptide. Our results indicate that leucine, serine, and threonine have a high propensity for forming α -helical structures. In contrast, isoleucine, valine, tyrosine, and histidine demonstrate significant deviations from the α -helical form, primarily because of steric hindrance. The positively charged amino acids, lysine, and arginine, exhibit comparably low free energy penalties for α -helix formation. Among the negatively charged amino acids, glutamate shows a lower folding free energy penalty $\Delta_{\text{fold}}G$ compared to aspartate, which can be attributed to the reduced steric effects resulting from glutamate's longer side chain. These trends can be observed in Fig. 8.6.

Following the validation of our protocol using the selected sequences from the c-Myc protein, we expanded our research focus to include other naturally occurring IDRs and α -MoRFs. Our objective was to integrate the capability of our protocol to quantify the folding free energy penalty with the structural prediction prowess of AlphaFold (Jumper et al., 2021). By doing so, we aimed to accurately identify

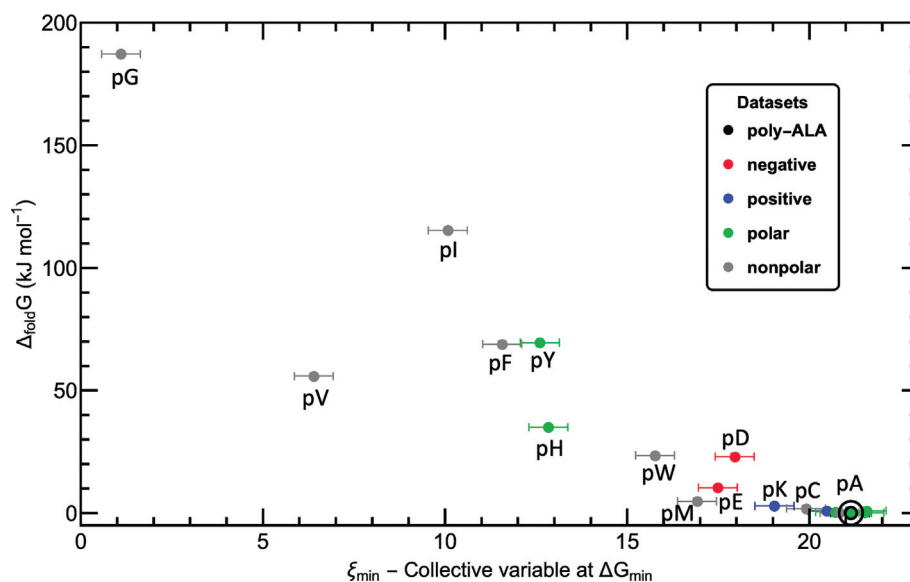
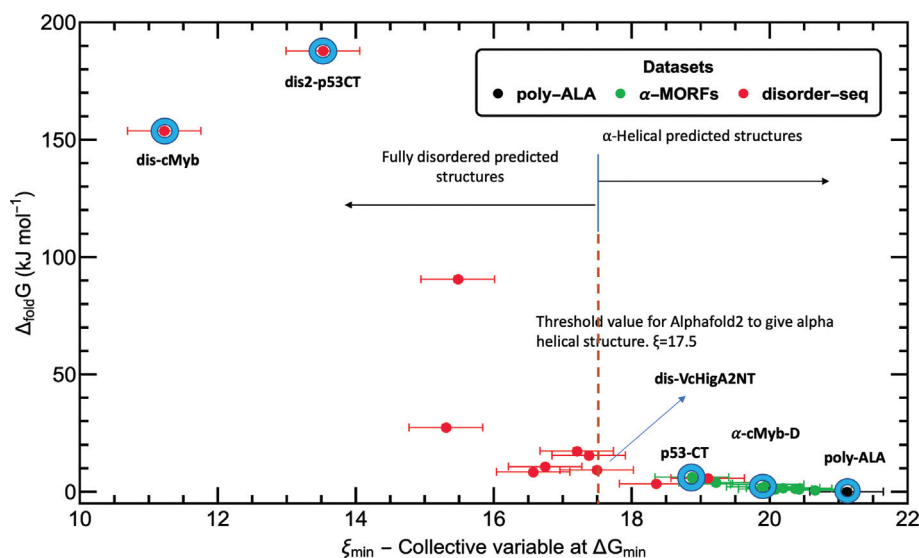


FIGURE 8.6

Helical folding propensities of polyamino acid peptides: Scatter plot comparing the folding free energy penalties $\Delta_{\text{fold}}G$ and position of free energy minima ξ_{min} for polyalanine and all other natural polyamino acid sequences.

or predict α -MoRFs and differentiate them from fully disordered IDRs. This approach would enhance our understanding of the structural dynamics of IDRs and their potential to transition into functionally relevant α -helical conformations upon interaction with binding partners. We selected two α -MoRFs: p53-CT (C-terminal domain of p53) and α -c-Myb-D (α -MoRF region of the c-Myb protein, where D stands for doubling the length of the single sequence for the purpose of our simulation) and three fully disordered sequences: dis-c-Myb (disordered fragment taken from c-Myb protein), dis2-p53-CT (disordered segment from p53 protein), and dis-Vc-HigA2-NT (disordered segment from the N-terminal region of Vc-HigA2 protein). We further include multiple other sequences from the same proteins with intermediate propensities for α -helical structure.

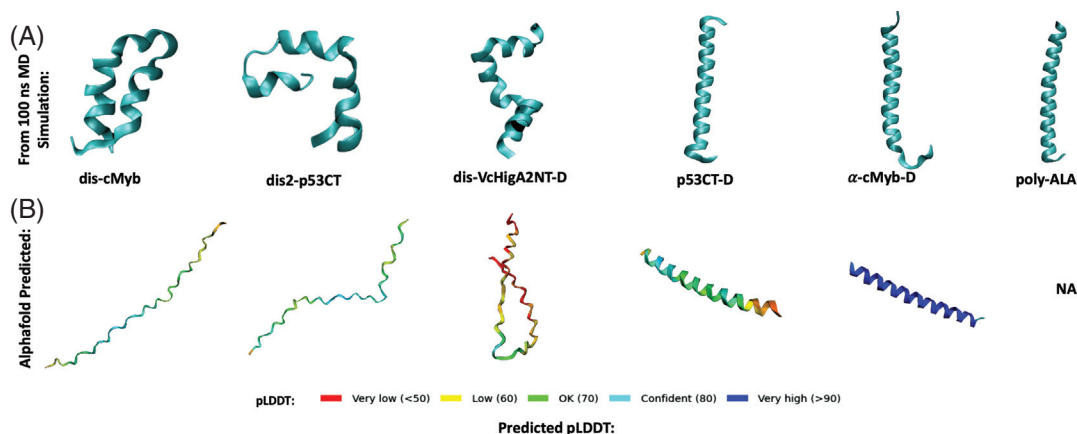
In Fig. 8.7, we observe a cluster of α -MoRF sequences (shown in green) and a cluster of fully disordered ones (shown in red). In evaluating the minimum free energy structures from our simulations against AlphaFold predictions, we observed that the folding free energies $\Delta_{\text{fold}}G$ of all the α -MoRF examples we investigated fell within the range of 0–10 kJ/mol (2.4 kcal/mol). This observation is in line with the findings of (Teilum et al., 2015), who reported complexes involving disordered proteins are, on average, less stable by approximately 2.5 kcal/mol or about 10.5 kJ/mol. This suggests that a certain amount of folding free energy might be associated with the binding-induced folding mechanism, where the binding event triggers the folding of the disordered protein into a more stable structure. This energy contribution is an important factor to consider when studying the dynamics and stability of protein complexes involving IDRs. Our finding also aligns with estimates provided by Hadži and Lah (2022), who utilized a modified LR model (Chakraborty et al., 1994; Hadži et al., 2021; Lifson & Roig, 1961) to assess the folding free energies of IDPs (apart from discrepancies attributed to variations in IDP

**FIGURE 8.7**

Helical folding propensities of naturally occurring IDRs: Folding free energy penalty $\Delta_{\text{fold}} G$ comparisons between IDRs from the c-Myb, p53, and Vc-HigA2 proteins with varying α -helical propensities. The threshold value of ξ_{min} , for which we consistently observed a prediction of α -helical structure by AlphaFold, is shown with a dotted line. α -MoRFs, α -Molecular recognition features; IDRs, intrinsically disordered regions.

sequence length). Low folding free energy penalties of a given sequence indicate its potential as a target, for example, in the development of peptide-based or small-molecule drugs. Such therapeutic agents have the potential to stabilize the helical structure of α -MoRFs by inducing folding upon binding, thereby offering a strategic approach for influencing protein–protein interactions in medical treatments. This approach could lead to the development of targeted therapies that specifically modulate the function of proteins involved in various diseases. We also identified a threshold value for the CV ξ_{min} that AlphaFold requires to predict α -helical structure. Specifically, for a 37-residue peptide, AlphaFold consistently predicts a fully disordered structure for a given sequence if we observe the position of the free energy minimum in our simulations below a value of 17.5. In addition, we observed that AlphaFold has a threshold for the folding free energy penalty, $\Delta_{\text{fold}} G$, of 0–10 kJ/mol. Consequently, all sequences with a folding free energy penalty below 10 kJ/mol were predicted to adopt an α -helical conformation. This finding is in alignment with the research conducted by [Hadži and Lah \(2022\)](#) and provides insights into the ability of the AlphaFold to distinguish preferences for secondary structure and disorder.

Finally, we compare the structures observed in unbiased simulations (representative structures observed in 100 ns MD trajectories starting from a helical conformation) with AlphaFold predictions. Specifically, we compare the simulated and predicted structures for the disordered sequences in c-Myb, p53, and VcHigA2 with the α -MoRF sequences in c-Myb and p53 with polyalanine as a reference in [Fig. 8.8](#).

**FIGURE 8.8**

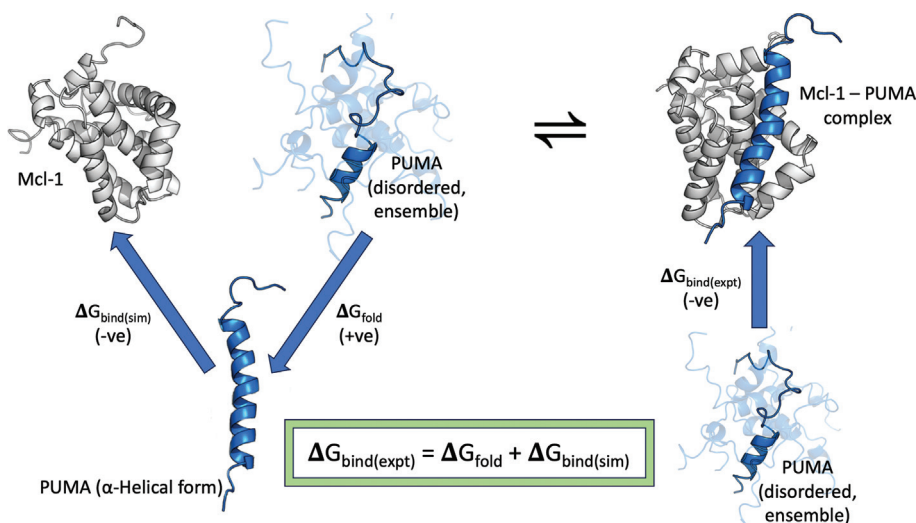
Comparison of simulated and predicted structures: Structure comparison of unbiased molecular dynamics simulations and AlphaFold predictions. (A) Representative conformations observed in 100 ns simulations starting from an α -helical conformation. (B) Structure predictions by AlphaFold with color-coded predicted local distance difference test (pLDDT) scores, which provide a confidence estimate and can be interpreted as a measure of disorder (low confidence = disorder).

We note that AlphaFold is proficient in predicting the helical structure of the two α -MoRFs, particularly in the case of α -c-Myb-D. In addition, AlphaFold is capable of accurately predicting disorder in sequences that yield a free energy minimum $\xi_{\min}$ below 17.5 in our folding free energy calculations for a 37-residue sequence. This threshold is exemplified by the structure of dis-VcHigA-2NT-D, a disordered protein that lies at the boundary of AlphaFold's ability to predict α -MoRFs. For values beyond this threshold, AlphaFold tends to predict that sequences will adopt an α -helical conformation. This observation underscores the ability of AlphaFold to recognize residual α -helical structure in α -MoRFs.

8.3.2 Binding free energy calculations of folded ensembles

Experimentally, characterizing the binding affinity of disordered α -MoRF peptides with their targets typically yields only the combined contributions of folding and binding (see Fig. 8.9). Separating these two contributions is not straightforward using conventional methods. Our goal here is to compare the sum of the folding and binding free energies derived from simulations with the experimentally determined binding free energies. Through this comparison, we aim to provide a more detailed understanding of the energetics involved in the interactions between IDPs and their binding partners.

To test our approach, we conducted CG US simulations of the binding process of two MCL-1-BH3 domain protein complexes. BH3 domains (Giam et al., 2008) are characteristic of the proapoptotic subset of Bcl-2 proteins (Youle & Strasser, 2008), which includes proteins like BIM, BAD, BID, PUMA, and NOXA-A. MCL-1 (Dahal et al., 2018; Garner et al., 2017; Senichkin et al., 2020; Youle & Strasser, 2008) is a key antiapoptotic protein that plays a critical role in cell survival and has

**FIGURE 8.9**

Binding as a two-step process: Expressing binding free energies as the sum of two processes in the binding-induced folding mechanism: folding and binding in the folded state. Adapted from [Rogers et al. \(2014\)](#).

been implicated in the development and progression of various cancers. Its ability to bind and inhibit proapoptotic proteins makes it a crucial regulator of the apoptotic pathway. As a result, MCL-1 has emerged as an attractive target for cancer therapy.

We note that both the structure of MCL-1 and peptide were stabilized by ENMs in our simulation as described in [Section 8.2](#). This stabilized the native state of MCL-1 (which we observed to be necessary as in most other CG protein models despite the increased level of detail in SIRAH) as well as the helical state of the binding peptide. Although this limits the sampling problem in each distance-dependent umbrella window largely to relative positions and orientations of the protein, achieving adequate sampling can remain a challenge for unbound states in which one binding partner can freely explore the surface of a sphere around its target with a radius corresponding to the restrained center-of-mass distance. As a consequence the PMF obtained from the simulation will, for sufficiently large distances and sampling, resemble $-k_B T \ln(4\pi r^2)$, which can be corrected as described in [Eq. \(8.3\)](#). As an alternative approach, cylindrical restraints can be used to restrict the separation path of the two binding partners. This approach restricts the ligand from exploring the entire space around the substrate in a spherical pattern, leading to faster convergence of the simulation with reduced sampling time and computational resources. Deriving the entropic correction term of the PMF for cylindrical restraints, akin to the one for [Eq. \(8.3\)](#), is an interesting analytical exercise. However, it is only noticeable for large cylindrical radii. Here, we verified that our PMFs resulting from simulations with and without cylindrical restraints (correcting the PMF using [Eq. 8.3](#) in the latter case) yield equivalent binding free energies and PMFs.

The PMFs for the binding of PUMA and NOXA-A to MCL-1 are shown in [Fig. 8.10](#). We determine the binding free energies, $\Delta_{\text{bind}}G_{\text{sim}}$, of both PUMA and NOXA-A in their prefolded helical states, from

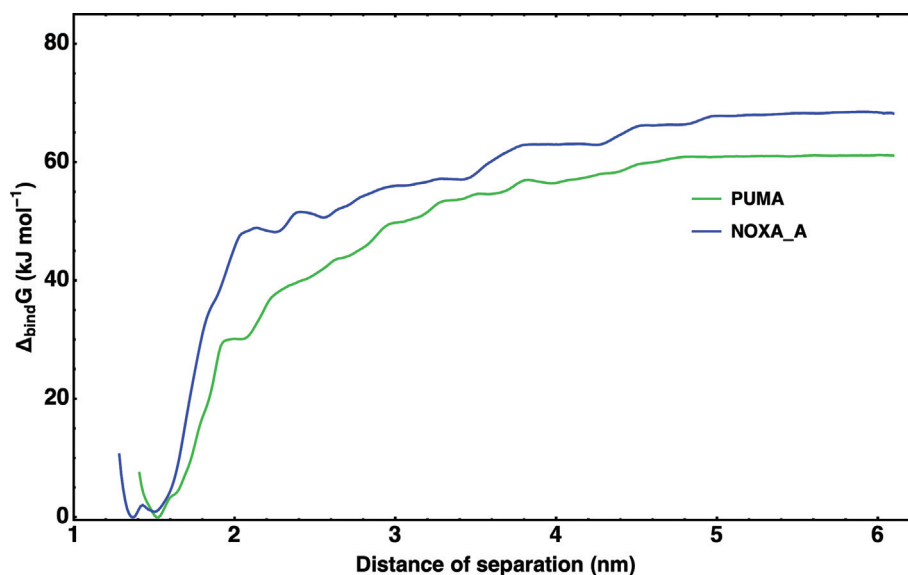


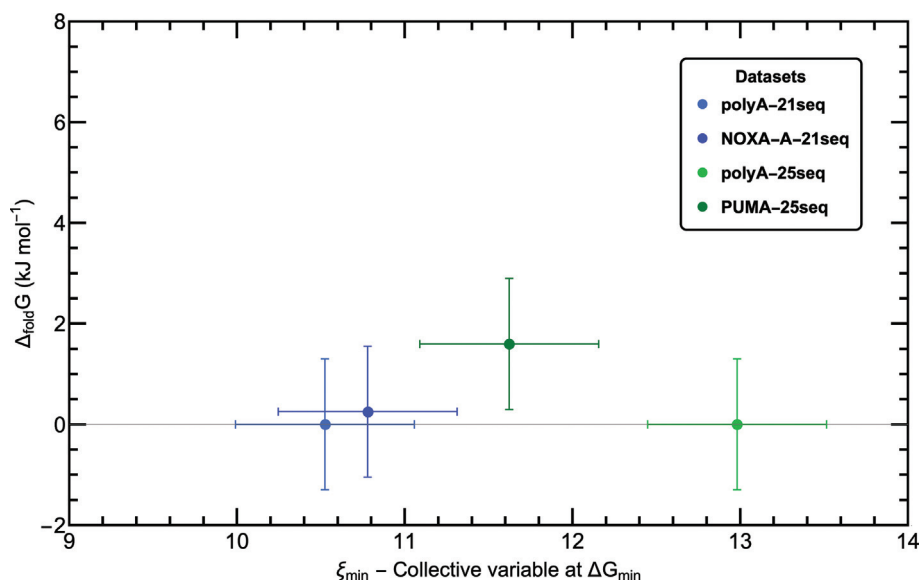
FIGURE 8.10

Binding free energy curves: Binding free energy curves for MCL-1 with the BH3-only domains of PUMA (green) and NOXA-A (blue).

the difference in free energy in the bound (free energy minimum) and unbound states (flat section of the PMF for large distances).

In Fig. 8.10 the calculated binding free energies $\Delta_{\text{bind}}G_{\text{sim}}$ for the protein complexes involving PUMA and NOXA-A were found to be -61.25 kJ/mol (-14.64 kcal/mol) and -68.58 kJ/mol (-16.39 kcal/mol), respectively. These values are noticeably lower than the experimental values $\Delta_{\text{bind}}G_{\text{exp}}$ reported by Dahal et al. (2018), which are -55.97 kJ/mol (-13.38 kcal/mol) for PUMA and -45.9 kJ/mol (-10.9 kcal/mol) for NOXA-A. This discrepancy suggests that, as expected, a positive folding free energy penalty, $\Delta_{\text{fold}}G$, needs to be accounted for to reconcile the excess binding free energy obtained from simulations of the helical peptides compared to experiments.

To calculate folding free energy penalties for both PUMA and NOXA-A peptides, we performed two sets of US simulations as already described for other peptides in the previous section. Fig. 8.11 illustrates that the folding free energy penalty for PUMA is 1.8 ± 1.3 kJ/mol, whereas for NOXA-A, it is approximately 0.5 ± 1.3 kJ/mol. Both are on the lower end of the 0–10 kJ/mol window expected for α -MoRFs, which indicates a significant population of the helical state even prior to binding. In both cases, the folding free energy penalty $\Delta_{\text{fold}}G$ does not fully capture the difference between $\Delta_{\text{bind}}G_{\text{exp}}$ and $\Delta_{\text{bind}}G_{\text{sim}}$, which amounts to 5.35 kJ/mol for PUMA and 22.68 kJ/mol for NOXA-A. However, for PUMA the total simulated binding free energy, obtained by combining the folding free energy penalty $\Delta_{\text{fold}}G$ of 1.8 ± 1.3 kJ/mol and the binding free energy $\Delta_{\text{bind}}G_{\text{sim}}$ of -61.25 kJ/mol from simulations of the helical peptides, amounts to -59.5 kJ/mol, which is within an acceptable error of 4 kJ/mol (less than 1 kcal/mol) from the experimental binding free energy of -55.97 kJ/mol.

**FIGURE 8.11**

Helical folding propensity of PUMA and NOXA-A: Scattered plots depicting folding free energy penalties for MCL-1 with PUMA (green) and NOXA-A (blue).

On the other hand, our approach clearly overestimates binding for NOXA-A. Adding $\Delta_{\text{fold}}G$ and $\Delta_{\text{bind}}G_{\text{sim}}$ in that case yields -68.1 kJ/mol, which exceeds the affinity observed in the experiment (-45.9 kJ/mol) by 21 kJ/mol. Several factors may contribute to this discrepancy: (1) Sequence length: The discrepancy may arise from differences in the length of the sequences used in our simulations compared to those used in experiments. The length of the sequence can influence the folding dynamics and, consequently, the folding free energy penalty. (2) Mutations: The use of mutated versions of the proteins in the experiments described by [Dahal et al. \(2018\)](#), which differ from the structures in the PDB used as a starting point of the simulations, could lead to variations in the folding free energy. (3) Sequence composition: Our results for NOXA-A might have been more accurate if we had used a longer sequence that included proline and glycine residues. These residues are known to destabilize α -helices and could result in a higher folding free energy penalty, bringing our simulated values closer to the experimental ones. (4) Simulation model: Predictions of the folding free energy penalties of the peptides may be limited by the accuracy of the underlying all-atom force field potential. Potentially even more likely the CG protein-protein interactions described by the SIRAH model might not capture all types of interactions equally well, for example, protein-protein hydrogen bonds or hydrophobic interactions.

Addressing these factors in future simulations may reduce the discrepancy between the simulated and experimental folding free energy values for PUMA and NOXA-A. However, our hypothesis that the experimental binding free energy can be represented as the sum of the calculated folding and binding free energies is partially verified, and additional studies involving other sequences are necessary to

gather more conclusive data. Expanding the scope of the simulations described here to a wider range of sequences will help to further enhance our understanding of the energetics involved in protein–protein interactions.

8.4 Conclusion

In this chapter, we delved into the intricate dynamics and thermodynamics of protein folding and binding processes, with a specific emphasis on IDRs and α -MoRFs present in the c-Myc protein and others. In addition, we examined the interactions between MCL-1 and BH3-only domains, which play pivotal roles in the regulation of apoptosis. Our primary aim was to quantify the folding free energy penalties that occur as proteins transition from disordered states to α -helical conformations. Furthermore, we sought to investigate the binding free energies of protein complexes, specifically those involving MCL-1 in combination with PUMA and NOXA-A, which are key players in apoptotic pathways. By understanding these energetics, we aimed to shed light on the molecular mechanisms underlying protein interactions and their implications for apoptotic regulation.

Through MD simulations and US, we first validated our protocol by analyzing sequences from the c-Myc protein, noting that the folding free energies fell within the anticipated range. This initial validation laid the groundwork for an extensive examination of naturally occurring IDRs and α -MoRFs. In this broader analysis, we harnessed the capabilities of our protocol in conjunction with structure predictions of AlphaFold. Our results demonstrated that AlphaFold was adept at predicting the α -helical structures of α -MoRFs, particularly in the case of α -c-Myb-D, and could identify sequences as fully disordered up to a specified threshold.

The computation of binding free energies as a sum of folding free energy penalties and binding free energies in the folded state unveiled partial agreement with experimental values. Specifically, PUMA binding to MCL-1 was reproduced quite accurately, whereas the binding free energy of NOXA-A was overestimated. In both cases the binding free energy penalty was comparably small and thus provided only a minor correction to the binding free energy. Overall, based on all sequences studied here, we typically expect a range from 0 to 10 kJ/mol for the positive penalty, which needs to be compensated by favorable interactions with the target to achieve binding.

In summary the decomposition of experimental binding free energies using simulation approaches as described here may reveal additional information for other systems. The intricate nature of protein folding and binding mechanisms underscores the necessity for additional investigations across a broader spectrum of sequences. This work enhances our comprehension of the energetic aspects of protein–protein interactions and has potential implications for the development of pharmaceuticals and the exploration of molecular-level biological phenomena.

AI disclosure

During the preparation of this work the author(s) used ChatGPT to do language editing. After using this tool/service the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1–2, 19–25. doi:[10.1016/j.softx.2015.06.001](https://doi.org/10.1016/j.softx.2015.06.001).
- Arai, M., Sugase, K., Dyson, H. J., & Wright, P. E. (2015). Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proceedings of the National Academy of Sciences of the United States of America*, 112(31), 9614–9619. <http://www.pnas.org/content/112/31/9614.full.pdf>. doi:[10.1073/pnas.1512799112](https://doi.org/10.1073/pnas.1512799112).
- Babu, M. M. (2016). The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochemical Society Transactions*, 44(5), 1185–1200. <http://www.biochemsoctrans.org/content/44/5/1185.full-text.pdf>. doi:[10.1042/BST20160172](https://doi.org/10.1042/BST20160172).
- Bahar, I., Rana Atilgan, A., & Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding and Design*, 2(3), 173–181. doi:[10.1016/S1359-0278\(97\)00024-2](https://doi.org/10.1016/S1359-0278(97)00024-2).
- Batley, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G., & Leder, P. (1983). The human c-myc oncogene: Structural consequences of translocation into the IgH locus in Burkitt lymphoma. *Cell*, 34(3), 779–787. doi:[10.1016/0092-8674\(83\)90534-2](https://doi.org/10.1016/0092-8674(83)90534-2).
- Beaulieu, M.-E., Castillo, F., & Soucek, L. (2020). Structural and biophysical insights into the function of the intrinsically disordered myc oncoprotein. *Cells*, 9(4). doi:[10.3390/cells9041038](https://doi.org/10.3390/cells9041038).
- Best, R. B., & Hummer, G. (2009). Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *Journal of Physical Chemistry B*, 113(26), 9004–9015. doi:[10.1021/jp901540t](https://doi.org/10.1021/jp901540t).
- Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A., & Parrinello, M. (2009). PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, 180(10), 1961–1972. doi:[10.1016/j.cpc.2009.05.011](https://doi.org/10.1016/j.cpc.2009.05.011).
- Bonomi, M., Bussi, G., Camilloni, C., Tribello, G. A., Banáš, P., Barducci, A., Bernetti, M., Bolhuis, P. G., Bottaro, S., Branduardi, D., Capelli, R., Carloni, P., Ceriotti, M., Cesari, A., Chen, H., Chen, W., Colizzi, F., De, S., De La Pierre, M., ... White, A. (2019). Promoting transparency and reproducibility in enhanced molecular simulations. *Nature Methods*, 16(8), 670–673. doi:[10.1038/s41592-019-0506-8](https://doi.org/10.1038/s41592-019-0506-8).
- Bugge, K., Staby, L., Salladini, E., Falbe-Hansen, R. G., Kragelund, B. B., & Skriver, K. (2021). $\alpha\alpha$ -Hub domains and intrinsically disordered proteins: A decisive combo. *Journal of Biological Chemistry*, 296, 100226. doi:[10.1074/jbc.rev120.012928](https://doi.org/10.1074/jbc.rev120.012928).
- Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *Journal of Chemical Physics*, 126(1). doi:[10.1063/1.2408420](https://doi.org/10.1063/1.2408420).
- Camilloni, C., Provasi, D., Tiana, G., & Broglia, R. A. (2008). Exploring the protein G helix free-energy surface by solute tempering metadynamics. *Proteins: Structure, Function and Genetics*, 71(4), 1647–1654. doi:[10.1002/prot.21852](https://doi.org/10.1002/prot.21852).
- Carr, D. W., Stofko-Hahn, R. E., Fraser, I. D., Bishop, S. M., Acott, T. S., Brennan, R. G., & Scott, J. D. (1991). Interaction of the regulatory subunit (RII) of cAMP-dependent protein kinase with RII-anchoring proteins occurs through an amphipathic helix binding motif. *Journal of Biological Chemistry*, 266(22), 14188–14192. doi:[10.1016/s0021-9258\(18\)98665-5](https://doi.org/10.1016/s0021-9258(18)98665-5).
- Chakraborty, A., Kortemme, T., & Baldwin, R. L. (1994). Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Science*, 3(5), 843–852. doi:[10.1002/pro.5560030514](https://doi.org/10.1002/pro.5560030514).
- Chodera, J. D., & Mobley, D. L. (2013). Entropy-enthalpy compensation: Role and ramifications in biomolecular ligand recognition and design. *Annual Review of Biophysics*, 42(1), 121–142. doi:[10.1146/annurev-biophys-083012-130318](https://doi.org/10.1146/annurev-biophys-083012-130318).

- Chou, P. Y., & Fasman, G. D. (1974). Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry*, 13(2), 211–222.
- Cumberworth, A., Lamour, G., Babu, M. M., & Gsponer, J. (2013). Promiscuity as a functional trait: Intrinsically disordered regions as central players of interactomes. *Biochemical Journal*, 454(3), 361–369. <http://www.biochemj.org/bj/454/0361/4540361.pdf>. doi:10.1042/BJ20130545Canada.
- Dahal, L., Kwan, T. O. C., Hollins, J. J., & Clarke, J. (2018). Promiscuous and selective: How intrinsically disordered BH3 proteins interact with their pro-survival partner MCL-1. *Journal of Molecular Biology*, 430(16), 2468–2477. <https://www.journals.elsevier.com/journal-of-molecular-biology>. doi:10.1016/j.jmb.2018.04.004.
- Dang, C. V. (2012). MYC on the path to cancer. *Cell*, 149(1), 22–35. doi:10.1016/j.cell.2012.03.003.
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089–10092. doi:10.1063/1.464397.
- Darré, L., Machado, M. R., Brandner, A. F., González, H. C., Ferreira, S., & Pantano, S. (2015). SIRAH: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *Journal of Chemical Theory and Computation*, 11(2), 723–739. doi:10.1021/ct5007746.
- Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research*, 32(Web Server Issue). doi:10.1093/nar/gkh381.
- Dong, H., Sharma, M., Zhou, H.-X., & Cross, T. A. (2012). Glycines: Role in α -helical membrane protein structures and a potential indicator of native conformation. *Biochemistry*, 51(24), 4779–4789. doi:10.1021/bi300090x.
- Drobnak, I., De Jonge, N., Haesaerts, S., Vesnaver, G., Loris, R., & Lah, J. (2013). Energetic basis of uncoupling folding from binding for an intrinsically disordered protein. *Journal of the American Chemical Society*, 135(4), 1288–1294. doi:10.1021/ja305081b.
- Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., & Kollman, P. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry*, 24(16), 1999–2012. doi:10.1002/jcc.10349.
- Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C. H., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., & Obradovic, Z. (2001). Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*, 19(1), 26–59. doi:10.1016/S1093-3263(00)00138-8.
- Dyson, H. J., & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology*, 12(1), 54–60. <http://www.elsevier.com/locate/sbi>. doi:10.1016/S0959-440X(02)00289-0.
- Dyson, H. J., & Wright, P. E. (2016). Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *Journal of Biological Chemistry*, 291(13), 6714–6722. <http://www.jbc.org/content/291/13/6714.full.pdf+html>. doi:10.1074/jbc.R115.692020.
- Emenecker, R. J., Griffith, D., & Holehouse, A. S. (2021). Metapredict: A fast, accurate, and easy-to-use predictor of consensus disorder and structure. *Biophysical Journal*, 120(20), 4312–4319. <http://www.sciencedirect.com/science/journal/00063495>. doi:10.1016/j.bpj.2021.08.039.
- Evans, D. J., & Holian, B. L. (1985). The Nose-Hoover thermostat. *The Journal of Chemical Physics*, 83(8), 4069–4074. doi:10.1063/1.449071.
- Fatafta, H., Samantray, S., Sayyed-Ahmad, A., Coskuner-Weber, O., & Strodel, B. (2021). Molecular simulations of IDPs: From ensemble generation to IDP interactions leading to disorder-to-order transitions. *Progress in Molecular Biology and Translational Science*, 183, 135–185. <http://www.elsevier.com/books/book-series/progress-in-molecular-biology-and-translational-science#>. doi:10.1016/bs.pmbts.2021.06.003.
- Flock, T., Weatheritt, R. J., Latysheva, N. S., & Babu, M. M. (2014). Controlling entropy to tune the functions of intrinsically disordered regions. *Current Opinion in Structural Biology*, 26(1), 62–72. <http://www.elsevier.com/locate/sbi>. doi:10.1016/j.sbi.2014.05.007.

- Fuxreiter, M., Simon, I., Friedrich, P., & Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *Journal of Molecular Biology*, 338(5), 1015–1026. doi:10.1016/j.jmb.2004.03.017.
- Garner, E., Romero, P., Dunker, A. K., Brown, C., & Obradovic, Z. (1999). Predicting binding regions within disordered proteins. *Genome Informatics*, 10, 41–50.
- Garner, T. P., Lopez, A., Reyna, D. E., Spitz, A. Z., & Gavathiotis, E. (2017). Progress in targeting the BCL-2 family of proteins. *Current Opinion in Chemical Biology*, 39, 133–142. doi:10.1016/j.cbpa.2017.06.014.
- Giam, M., Huang, D. C. S., & Bouillet, P. (2008). BH3-only proteins and their roles in programmed cell death. *Oncogene*, 27. doi:10.1038/onc.2009.50.
- Gnutt, D., Gao, M., Brylski, O., Heyden, M., & Ebbinghaus, S. (2015). Excluded-volume effects in living cells. *Angewandte Chemie—International Edition*, 54(8), 2548–2551. doi:10.1002/anie.201409847.
- Hadži, S., & Lah, J. (2022). The free energy folding penalty accompanying binding of intrinsically disordered α -helical motifs. *Protein Science*, 31(7). doi:10.1002/pro.4370.
- Hadži, S., Loris, R., & Lah, J. (2021). The sequence-ensemble relationship in fuzzy protein complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 118(37). <https://www.pnas.org/content/118/37/e2020562118>. doi:10.1073/pnas.2020562118.
- Haliloglu, T., Bahar, I., & Erman, B. (1997). Gaussian dynamics of folded proteins. *Physical Review Letters*, 79(16), 3090–3093. doi:10.1103/PhysRevLett.79.3090.
- Herschbach, D. R., Johnston, H. S., & Rapp, D. (1959). Molecular partition functions in terms of local properties. *The Journal of Chemical Physics*, 31(6), 1652–1661. doi:10.1063/1.1730670.
- Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), 1463–1472.
- Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. *Proteins*, 33(3), 417–429.
- Holehouse, A. S., & Kragelund, B. B. (2024). The molecular basis for cellular function of intrinsically disordered protein regions. *Nature Reviews Molecular Cell Biology*, 25(3), 187–211. <https://www.nature.com/molcellbio>. doi:10.1038/s41580-023-00673-0.
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function and Genetics*, 65(3), 712–725. doi:10.1002/prot.21123.
- Hub, J. S., de Groot, B. L., & van der Spoel, D. (2010). G-whams-A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *Journal of Chemical Theory and Computation*, 6(12), 3713–3720. doi:10.1021/ct100494z.
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33–38. doi:10.1016/0263-7855(96)00018-5.
- Jha, R. K., Kouzine, F., & Levens, D. (2023). MYC function and regulation in physiological perspective. *Frontiers in Cell and Developmental Biology*, 11. doi:10.3389/fcell.2023.1268275.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), 926–935. doi:10.1063/1.445869.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. doi:10.1038/s41586-021-03819-2.
- Kästner, J. (2011). Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(6), 932–942. doi:10.1002/wcms.66.

- Khavrutskii, I. V., Dzubiella, J., & McCammon, J. A. (2008). Computing accurate potentials of mean force in electrolyte solutions with the generalized gradient-augmented harmonic Fourier beads method. *Journal of Chemical Physics*, 128(4). doi:10.1063/1.2825620.
- Kim, D. H., & Han, K. H. (2018a). PreSMo target-binding signatures in intrinsically disordered proteins. *Molecules and Cells*, 41(10), 889–899. <http://www.molcells.org/main.html>. doi:10.14348/molcells.2018.0192.
- Kim, D. H., & Han, K. H. (2018b). Transient secondary structures as general target-binding motifs in intrinsically disordered proteins. *International Journal of Molecular Sciences*, 19(11). <https://www.mdpi.com/1422-0067/19/11/3614/pdf>. doi:10.3390/ijms19113614.
- Krivov, G. G., Shapovalov, M. V., & Dunbrack, R. L., Jr. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins: Structure, Function and Bioinformatics*, 77(4), 778–795. doi:10.1002/prot.22488.
- Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8), 1011–1021. doi:10.1002/jcc.540130812.
- Lazar, T., Tantos, A., Tompa, P., & Schad, E. (2022). Intrinsic protein disorder uncouples affinity from binding specificity. *Protein Science*, 31(11). doi:10.1002/pro.4455.
- Lifson, S., & Roig, A. (1961). On the theory of helix—Coil transition in polypeptides. *The Journal of Chemical Physics*, 34(6), 1963–1974. doi:10.1063/1.1731802.
- Lu, H., & Schulten, K. (1999). Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins*, 35(4), 453–463.
- Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990). Side chain contributions to the stability of alpha-helical structure in peptides. *Science (New York, N.Y.)*, 250(4981), 669–673.
- Lyu, P. C., Sherman, J. C., Chen, A., & Kallenbach, N. R. (1991). Alpha-helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proceedings of the National Academy of Sciences of the United States of America*, 88(12), 5317–5320.
- Machado, M. R., & Pantano, S. (2016). SIRAH tools: Mapping, backmapping and visualization of coarse-grained models. *Bioinformatics*, 32(10), 1568–1570. doi:10.1093/bioinformatics/btw020.
- Madden, S. K., Dantas de Araujo, A., Gerhardt, M., Fairlie, D. P., & Mason, J. M. (2021). Taking the Myc out of cancer: Toward therapeutic strategies to directly inhibit c-Myc. *Molecular Cancer*, 20(1). doi:10.1186/s12943-020-01291-6.
- Marrink, S. J., & Tieleman, D. P. (2013). Perspective on the martini model. *Chemical Society Reviews*, 42(16), 6801–6822. doi:10.1039/C3CS60093A.
- Mészáros, B., Simon, I., Dosztányi, Z., & Casadio, R. (2009). Prediction of protein binding regions in disordered proteins. *PLoS Computational Biology*, 5(5), e1000376. doi:10.1371/journal.pcbi.1000376.
- Meyer, N., & Penn, L. Z. (2008). Reflecting on 25 years with MYC. *Nature Reviews Cancer*, 8(12), 976–990. doi:10.1038/nrc2231.
- Mishra, S. K., & Jernigan, R. L. (2018). Protein dynamic communities from elastic network models align closely to the communities defined by molecular dynamics. *PLoS ONE*, 13(6). doi:10.1371/journal.pone.0199225.
- Miyamoto, S., & Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8), 952–962. doi:10.1002/jcc.540130805.
- Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., & Uversky, V. N. (2006). Analysis of molecular recognition features (MoRFs). *Journal of Molecular Biology*, 362(5), 1043–1059. <https://www.journals.elsevier.com/journal-of-molecular-biology>. doi:10.1016/j.jmb.2006.07.087.
- Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N., & Bunker, A. K. (2005a). Comparing and combining predictors of mostly disordered proteins. *Biochemistry*, 44(6), 1989–2000. doi:10.1021/bi047993o.
- Oldfield, C. J., Cheng, Y., Cortese, M. S., Romero, P., Uversky, V. N., & Dunker, A. K. (2005b). Coupled folding and binding with α -helix-forming molecular recognition elements. *Biochemistry*, 44(37), 12454–12470. doi:10.1021/bi050736e.

- O'Neil, K. T., & DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science*, 250(4981), 646–651.
- Pace, C. N., & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal*, 75(1), 422–427. doi:[10.1016/S0006-3495\(98\)77529-0](https://doi.org/10.1016/S0006-3495(98)77529-0).
- Pancsa, R., & Fuxreiter, M. (2012). Interactions via intrinsically disordered regions: What kind of motifs? *IUBMB Life*, 64(6), 513–520. doi:[10.1002/iub.1034](https://doi.org/10.1002/iub.1034).
- Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12), 7182–7190. doi:[10.1063/1.328693](https://doi.org/10.1063/1.328693).
- Periols, X., & Marrink, S.-J. (2013). *The Martini coarse-grained force field* (pp. 533–565). Humana Press. doi:[10.1007/978-1-62703-017-5_20](https://doi.org/10.1007/978-1-62703-017-5_20).
- Poosapati, A., Gregory, E., Borchers, W. M., Chemes, L. B., & Daughdrill, G. W. (2018). Uncoupling the folding and binding of an intrinsically disordered protein. *Journal of Molecular Biology*, 430(16), 2389–2402. <https://www.journals.elsevier.com/journal-of-molecular-biology>. doi:[10.1016/j.jmb.2018.05.045](https://doi.org/10.1016/j.jmb.2018.05.045).
- Qian, H. (2006). An asymptotic comparative analysis of the thermodynamics of non-covalent association. *Journal of Mathematical Biology*, 52(3), 277–289. doi:[10.1007/s00285-005-0353-3](https://doi.org/10.1007/s00285-005-0353-3).
- Rogers, J. M., Wong, C. T., & Clarke, J. (2014). Coupled folding and binding of the disordered protein PUMA does not require particular residual structure. *Journal of the American Chemical Society*, 136(14), 5197–5200. <http://pubs.acs.org/journal/jacsat>. doi:[10.1021/ja4125065](https://doi.org/10.1021/ja4125065).
- Rohl, C. A., Fiori, W., & Baldwin, R. L. (1999). Alanine is helix-stabilizing in both template-nucleated and standard peptide helices. *Proceedings of the National Academy of Sciences of the United States of America*, 96(7), 3682–3687. doi:[10.1073/pnas.96.7.3682](https://doi.org/10.1073/pnas.96.7.3682).
- Senichkin, V. V., Streletskaia, A. Y., Gorbunova, A. S., Zhivotovsky, B., & Kopeina, G. S. (2020). Saga of Mcl-1: Regulation from transcription to degradation. *Cell Death and Differentiation*, 27(2), 405–419. doi:[10.1038/s41418-019-0486-3](https://doi.org/10.1038/s41418-019-0486-3).
- Sharma, R., Sharma, A., Patil, A., & Tsunoda, T. (2019). Discovering MoRFs by trisecting intrinsically disordered protein sequence into terminals and middle regions. *BMC Bioinformatics*, 19(S13). doi:[10.1186/s12859-018-2396-7](https://doi.org/10.1186/s12859-018-2396-7).
- Skriver, K., Theisen, F. F., & Kragelund, B. B. (2023). Conformational entropy in molecular recognition of intrinsically disordered proteins. *Current Opinion in Structural Biology*, 83. <http://www.elsevier.com/locate/sbi>. doi:[10.1016/j.sbi.2023.102697](https://doi.org/10.1016/j.sbi.2023.102697).
- Spek, E. J., Olson, C. A., Shi, Z., & Kallenbach, N. R. (1999). Alanine is an intrinsic α -helix stabilizing amino acid. *Journal of the American Chemical Society*, 121(23), 5571–5572. doi:[10.1021/ja990056x](https://doi.org/10.1021/ja990056x).
- Teilum, K., Olsen, J. G., & Kragelund, B. B. (2015). Globular and disordered-the non-identical twins in protein-protein interactions. *Frontiers in Molecular Biosciences*, 2. <https://www.frontiersin.org/articles/10.3389/fmolb.2015.00040/full>. doi:[10.3389/fmolb.2015.00040](https://doi.org/10.3389/fmolb.2015.00040).
- Theisen, F. F., Staby, L., Tidemand, F. G., O'Shea, C., Prestel, A., Willemoës, M., Kragelund, B. B., & Skriver, K. (2021). Quantification of conformational entropy unravels effect of disordered flanking region in coupled folding and binding. *Journal of the American Chemical Society*, 143(36), 14540–14550. <http://pubs.acs.org/journal/jacsat>. doi:[10.1021/jacs.1c04214](https://doi.org/10.1021/jacs.1c04214).
- Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Physical Review Letters*, 77(9), 1905–1908. doi:[10.1103/PhysRevLett.77.1905](https://doi.org/10.1103/PhysRevLett.77.1905).
- Togashi, Y., & Flechsig, H. (2018). Coarse-grained protein dynamics studies using elastic network models. *International Journal of Molecular Sciences*, 19(12). doi:[10.3390/ijms19123899](https://doi.org/10.3390/ijms19123899).
- Trzesniak, D., Kunz, A.-P. E., & Van Gunsteren, W. F. (2007). A comparison of methods to compute the potential of mean force. *ChemPhysChem*, 8(1), 162–169. doi:[10.1002/cphc.200600527](https://doi.org/10.1002/cphc.200600527).
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M.,

- Ronneberger, O., Bates, R., Kohl, S. A. A., ... Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, 596(7873), 590–596. <http://www.nature.com/nature/index.html>. doi:10.1038/s41586-021-03828-1.
- Uversky, V. N. (2002). What does it mean to be natively unfolded? *European Journal of Biochemistry*, 269(1), 2–12. doi:10.1046/j.0014-2956.2001.02649.x.
- Van Der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E., & Babu, M. M. (2014). Classification of intrinsically disordered regions and proteins. *Chemical Reviews*, 114(13), 6589–6631. <http://pubs.acs.org/journal/chcreay>. doi:10.1021/cr400525m.
- Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*, 2016. doi:10.1002/cpbi.3.
- Wohl, S., Jakubowski, M., & Zheng, W. (2021). Salt-dependent conformational changes of intrinsically disordered proteins. *Journal of Physical Chemistry Letters*, 12(28), 6684–6691. doi:10.1021/acs.jpcclett.1c01607.
- Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1), 18–29. <http://www.nature.com/molcellbio>. doi:10.1038/nrm3920.
- Wright, P. E., & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology*, 19(1), 31–38. doi:10.1016/j.sbi.2008.12.003.
- Yan, J., Dunker, A. K., Uversky, V. N., & Kurgan, L. (2016). Molecular recognition features (MoRFs) in three domains of life. *Molecular BioSystems*, 12(3), 697–710. <http://pubs.rsc.org/en/journals/journal/mb>. doi:10.1039/c5mb00640f.
- Youle, R. J., & Strasser, A. (2008). The BCL-2 protein family: Opposing activities that mediate cell death. *Nature Reviews Molecular Cell Biology*, 9(1), 47–59. <https://doi.org/10.1038/nrm2308>.
- Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 7(2). <https://doi.org/10.1002/wcms.1298>.
- Zavrtanik, U., Hadži, S., & Lah, J. (2021). Unraveling the thermodynamics of ultra-tight binding of intrinsically disordered proteins. *Frontiers in Molecular Biosciences*, 8. <https://doi.org/10.3389/fmolb.2021.726824>.