FLEXIBILITY VS. PREORGANIZATION: MOLECULAR SIMULATIONS OF MDM2-P53 PEPTIDE BINDING EVENTS

by

David Wen Rui Wang

Submitted to the University Honors College in partial fulfillment

of the requirements for the degree of

Bachelor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

UNIVERSITY HONORS COLLEGE

This thesis was presented

by

David W. Wang

It was defended on

April 12, 2012

and approved by

Kevin W. Plaxco, Professor, Chemistry and Biochemistry

David Swigon, Associate Professor, Mathematics

Daniel M. Zuckerman, Associate Professor, Computational and Systems Biology

Thesis Advisor: Lillian Chong, Assistant Professor, Chemistry

FLEXIBILITY VS. PREORGANIZATION: MOLECULAR SIMULATIONS OF MDM2-P53 PEPTIDE BINDING EVENTS

David W. Wang, BPhil

University of Pittsburgh, 2012

Speeding the molecular binding process is of particular interest in many fields. While traditional belief dictates that ligand preorganization is optimal, the discovery of intrinsically disordered proteins may contest such convention. The “fly-casting” mechanism argues that a flexible protein can bind its partner faster due to a larger capture radius and a resulting coupled process of folding and binding. We directly test this hypothesis, using computational means, on the p53-MDM2 system, performing binding simulations of MDM2 to either a flexible p53 peptide or its exact preorganized analog. We employ a path sampling algorithm, weighted ensemble, to generate large ensembles of binding pathways and to calculate rates of association. Additionally, the effect of hydrodynamic interactions, often omitted in implicit solvent simulations, on the binding rates was examined. We find no difference between the binding rates of flexible p53 and preorganized p53. The exclusion of hydrodynamic interactions significantly decreases the binding rates due to largely reduced translational diffusion coefficients, indicating the importance of using hydrodynamic interactions in binding simulations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to acknowledge committee members Dr. Kevin Plaxco, Dr. David Swigon, and Dr. Daniel Zuckerman for their participation in and insightful discussion during the thesis defense. I wish to give special thanks to Dr. Lillian Chong for terrific advising throughout my years in the lab and tremendous confidence in my ability to engage in this significant study. The Chong group has provided continuous support for my work, and, in particular, graduate student Matthew Zwier was instrumental in easing this process. I extend thanks to XSEDE and University of Pittsburgh CMMS for computational resources and the James V. Harrison Fund and University Honors College Brackenridge Research Fellowships for providing funding since May, 2011. Finally, I express extreme gratitude to Dr. Toby Chapman and Dr. Edward Stricker for supporting me to reach my fullest potential.

1.0 INTRODUCTION

Molecular recognition is essential to the function of all biological systems. Different fields of molecular recognition have emerged, which include, but are not limited to, host-guest detection, macromolecular association, and drug discovery. Much interest in these fields involves optimizing the kinetics of molecular binding, focusing on ligand preorganization as the standard for providing the fastest binding scenarios. Such convention, though, may be challenged by the prevalence of intrinsically disordered proteins (Wright & Dyson, 1999), many of which adopt well-defined structures upon binding their partners. Recent theoretical studies have suggested that intrinsic disorder provides kinetic advantages through the "fly-casting mechanism" (Shoemaker, Portman, & Wolynes, 2000). As the theory explains, a disordered protein, due to its increased flexibility, exhibits a larger capture radius, thereby allowing it to better extend and bind weakly to its partner. The protein then reels its partner in through a simultaneous process of folding and binding.

Experimental studies have been performed to investigate whether flexible proteins, in general, bind their targets faster than preorganized versions (Crespin, Boys, & Konermann, 2005; Hoffman, Blumenschein, & Sykes, 2006; Jemth & Gianni, 2007; Landfried, Vuletich, Pond, & Lecomte, 2007; Lengyel et al., 2007; Muralidhara, Rathinakumar, & Wittung-Stafshede, 2006; Narayanan, Ganesh, Edison, & Hagen, 2008; Onitsuka, Kamikubo, Yamazaki, & Kataoka, 2008; Perham, Chen, Ma, & Wittung-Stafshede, 2005; Sugase, Dyson, & Wright, 2007; Sugase,

Lansing, Dyson, & Wright, 2007; Vamvaca, Jelesarov, & Hilvert, 2008). However, none have addressed the direct question of whether a disordered protein binds its partner faster than its exact, folded analogue. Examining structural details of protein conformational changes during binding is challenging in experiments due to the transient nature of intermediate states. In addition, it is impossible, experimentally, to preorganize an intrinsically disordered protein without altering its chemical structure.

Alternatively, molecular dynamics simulations can, in principle, offer atomistic detail in probing specific protein conformational changes upon binding while providing a means to comparing the kinetics of binding of the flexible vs. exact preorganized versions of the protein. However, it is computationally prohibitive to generate an ensemble of protein-protein binding events using such simulations. Instead, coarse-grained simulations with residue-level detail coupled with Gō-type potentials can be employed (Go, 1983; Takada, 1999). Many have attempted to study coupled folding and binding in intrinsically disordered proteins as predicted by the fly-casting mechanism (Huang & Liu, 2009; Levy, Onuchic, & Wolynes, 2007; Turjanski, Gutkind, Best, & Hummer, 2008). In addition, these types of molecular simulations have been successful at generating large ensembles of unfolding/refolding events, demonstrating thermodynamic tug-of-war between two-domain protein switches (Mills & Chong, 2011), and illustrating cotranslational folding (Elcock, 2006).

Here, we directly test whether a flexible peptide binds its target faster than does its preorganized analogue. Binding simulations were conducted between an N-terminal peptide fragment of the transactivational domain of tumor suppressor p53 and its partner protein, MDM2. In addition to being of great biomedical interest, this system is a classic peptide complex in which the intrinsically disordered p53 peptide adopts a helical conformation upon binding

MDM2 in a well-defined pocket. In addition to directly comparing the kinetics of fully flexible p53 peptide and its preorganized counterpart, we also test the effect of hydrodynamic interactions, often omitted in implicit solvent simulations, on the binding kinetics. Furthermore, a path sampling algorithm, the weighted ensemble approach, was implemented to efficiently generate an ensemble of binding pathways. This method also allows for rigorous calculation of association rates. To our knowledge, the resulting simulations are the most comprehensive to date, initiated from 3000 selected conformations and random orientations, yielding an ensemble of over 50,000 binding events. Finally, in contrast to a previous study which used a coarse-grained Cα model (Huang & Liu, 2009), we implemented a coarse-grained side-chain model that was necessary to reproduce the p53 α-helix structure in its MDM2-bound state otherwise absent in a simple Cα model.

2.0 METHODS

2.1 PROTEIN MODEL

All proteins were modeled with a coarse-grained side chain model (Frembgen-Kesner & Elcock, 2009). Each amino acid is represented by up to four pseudo-atoms: one for the Cα atom and up to three for the side chain atoms. The heavy-atom model contained 820 atoms whereas the side chain model comprised 262 atoms (see Figure 1). Coordinates for the bound state of p53 and MDM2 were taken from the X-ray crystal structure (PDB code: 1YCR) (Kussie et al., 1996).

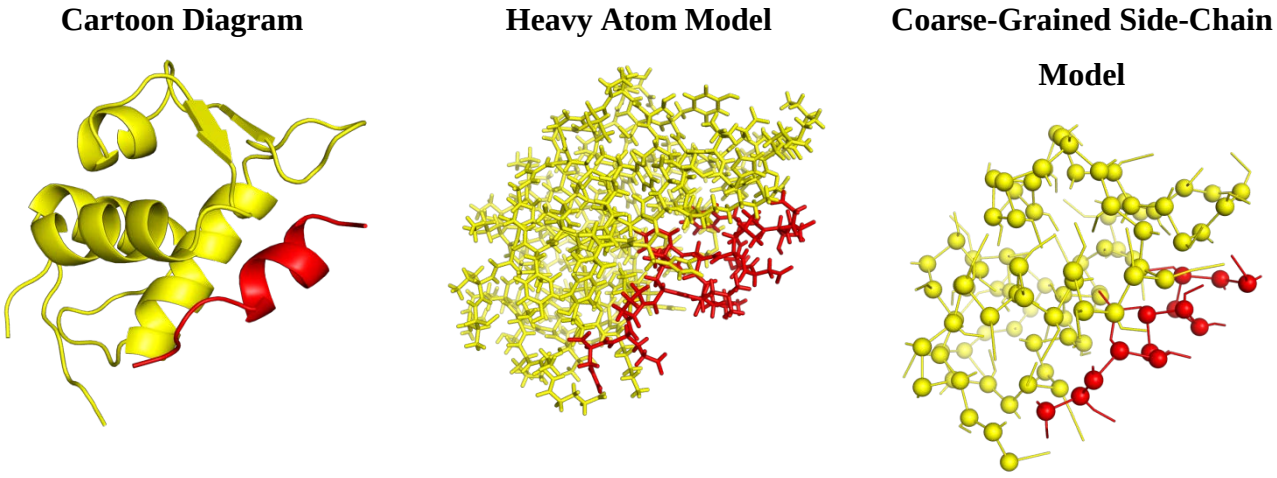

Figure 1. Protein Model

The coarse-grained side-side chain model (right) is shown in comparison to the heavy atom model (middle) for the p53-MDM2 complex. A cartoon of the complex is displayed on the left to show secondary structure. MDM2 is colored yellow, and p53 is colored red.

A Gō-type potential energy function governs the conformational dynamics of the protein model. Bonded interactions between atoms are modeled by standard molecular mechanics terms:

$$E_{bonded} = \sum_{bonds} k_{bond}(r - r_{eq})^2 + \sum_{angles} k_{angle}(\theta - \theta_{eq})^2 + \sum_{dihedrals} V_1[1 + \cos(\varphi - \varphi_1)] + V_3[1 + \cos(3\varphi - \varphi_3)]$$

in which r, θ, φ are pseudo-bond lengths, pseudo-angles, and pseudo-dihedrals, respectively; V_1 and V_3 are potential barriers for the dihedral terms. Equilibrium bond lengths (r_{eq}), angles (θ_{eq}), and dihedral phase angles (φ_1 and φ_3) were taken from the crystal structure. The force constants, k_{bond} and k_{angle}, were set to 100 kcal/mol/Å and 20 kcal/mol/rad, respectively.

Nonbonded interactions between residues separated by four or more pseudo-bonds were modeled in one of two ways, depending on whether or not the residues form (native) contacts in the native, bound state. A native contact is formed by two heavy atoms located within 5.5 Å of each other in the crystal structure of the protein complex. Native contact interactions were modeled using a Lennard-Jones-like potential:

$$E_{ij}^{native} = \varepsilon^{native}\left[5\left(\frac{\sigma_{ij}^{native}}{r_{ij}}\right)^{12} - 6\left(\frac{\sigma_{ij}^{native}}{r_{ij}}\right)^{10}\right]$$

in which ε^{native} is the energy well depth for the native interaction, r represents interatomic distance during simulation, and σ^{native} represents the corresponding distance in the crystal structure. Non-native interactions were modeled using a purely repulsive potential:

$$E_{ij}^{non-native} = \varepsilon^{non-native}\left(\frac{\sigma_{ij}^{non-native}}{r_{ij}}\right)^{12}$$

in which $\sigma_{ij}^{non-native}$ and $\varepsilon^{non-native}$ are set to 4.0 Å and 0.60 kcal/mol, respectively. The number of intramolecular native contacts for p53 and MDM2 are 63 and 884, respectively. In addition, there are 174 intermolecular native contacts between p53 and MDM2.

Brute force simulations were carried out as controls to determine correct parameterizations for p53 peptide flexibility (preorganized vs. flexible) in isolation of MDM2, MDM2 stability in isolation of p53, and stability of flexible p53 when bound to MDM2. Stability and flexibility were measured by the fraction of native contacts, Q, a standard metric for quantitatively assessing how folded a protein is relative to the reference crystal structure or the native state. The crystal structure has some total number, N, of native contacts, and any altered conformation of the protein may contain 0 to N native contacts. The number of native contacts present within a conformation divided by N yields Q. Therefore, $Q = 1$ represents complete folding, and $Q = 0$ implies complete unfolding. By tracking the values and fluctuations of Q of a protein over time, information could be gathered about the stability of p53.

The definitions of flexible and preorganized p53 were as follows. If $Q > 0.8$ throughout the duration of a simulation, the p53 peptide was considered preorganized. In contrast, p53 was deemed flexible if $Q < 0.5$ for all time. In this model, flexibility was tuned by adjusting ε^{native}, the potential well depths of the native contacts, where a high ε^{native} favors preorganization and vice versa. Two different values of ε^{native} were attained to satisfy the criteria for flexible and preorganized p53: 0.05 for flexible and 1.70 for preorganized. In addition to parameterizing p53, we needed to ensure MDM2 remained folding at all times during simulation. An ε^{native} value of 0.6 was sufficient in accomplishing this goal.

A final criterion for these simulations was that flexible p53 peptide must assume an α-helical conformation when bound to MDM2. Because the native contacts were so weak within

the flexible p53 peptide, only the intermolecular native contacts between MDM2 and p53 could force p53 into its bound conformation. Therefore, intermolecular native contacts' potential well-depths, $\varepsilon_{inter}^{native}$, had to be tuned. The exact criterion to satisfy was that the fraction of intramolecular p53 native contacts must be greater than 0.5 throughout the simulation. Such condition for the flexible p53 peptide was met with $\varepsilon_{inter}^{native} = 6.0$, which was then used for all binding simulations involving both flexible and preorganized p53.

2.2 SIMULATIONS USING THE WEIGHTED ENSEMBLE APPROACH

Because brute force sampling of binding events is very challenging, an enhanced path sampling algorithm, weighted ensemble, was employed to compute the rates of association. The weighted ensemble path sampling algorithm is an iterative procedure that uses "statistical ratcheting" to efficiently sample rare events (binding events in this case) using stochastic simulations (Huber & Kim, 1996). A progress coordinate is set up between initial (A) and destination (B) states. The progress coordinate is then arbitrarily divided into bins. An "ideal" number, n, is chosen as the number of simulations that should populate each bin.

In the first iteration, n simulations—of equal statistical weights that sum to one—start in the initial state A and are propagated for a short but fixed time, τ. After this time, the simulations may have populated different bins. The simulations in all bins are then copied into replicas until n simulations populate each occupied bin. Whenever a simulation is replicated, its statistical weight is split evenly between the newly copied simulations to maintain correct statistics. The next iteration is then prepared, and all simulations are again propagated for time τ.

If, at the end of any iteration, there exists more than n simulations in a given bin, the simulations are "merged" until n simulations remain. The merging procedure goes as follows. The simulations in the bin are ordered by their statistical weights. The two simulations with the lowest statistical weights are paired. One simulation will be terminated while the other will remain. The choice of which simulation remains is based on the statistical weight of the simulations; thus, the simulation with the higher statistical weight will have a greater chance of continuing. The simulation that continues will be assigned a new statistical weight equivalent to the sum of the weights of the paired simulations. If there still are more than n simulations in the bin, all the simulations are ordered again by statistical weights, and the same procedure occurs, terminating simulations one at a time.

When a trajectory reaches a destination state B at the end of an iteration, the trajectory is "recycled" as a new segment starting from the initial state A. The number of iterations run is dependent upon how long it takes to observe a desired number of rare events. In this work, simulations are run until steady state is achieved. The progress coordinate is a one-dimensional coordinate representing the root-mean-square deviation (RMSD) of p53 after the system is aligned to MDM2, and the bin boundaries are generally separated by 4 Å.

In order to calculate rates of association, the Northrup-Allison-McCammon (NAM) method (Northrup, Allison, & McCammon, 1984) was applied in conjunction with the weighted ensemble approach. The NAM method involves the creation of two concentric spheres in which MDM2 is initially positioned in the center (see Figure 2). The inner sphere of radius b, on which p53 is initially positioned, is constructed sufficiently large such that distance is the only factor in the interaction between binding partners. The outer sphere of radius q, termed the truncation sphere, is the diffusion boundary. For this study, the inner sphere radius was set to 35 Å, and the

outer sphere radius was set to 50 Å. Trajectories are terminated either when p53 binds MDM2 or if p53 diffuses across the truncation sphere, two natural target states within weighted ensemble. The bound state was defined by an RMSD of p53 less than 1.159 Å, one standard deviation above the average RMSD of p53 obtained from brute force simulations of the bound state. By monitoring the minimum distance between p53 and MDM2, any trajectory reaching a minimum distance greater than 50 Å was terminated, and its probability was recycled back to the initial state.

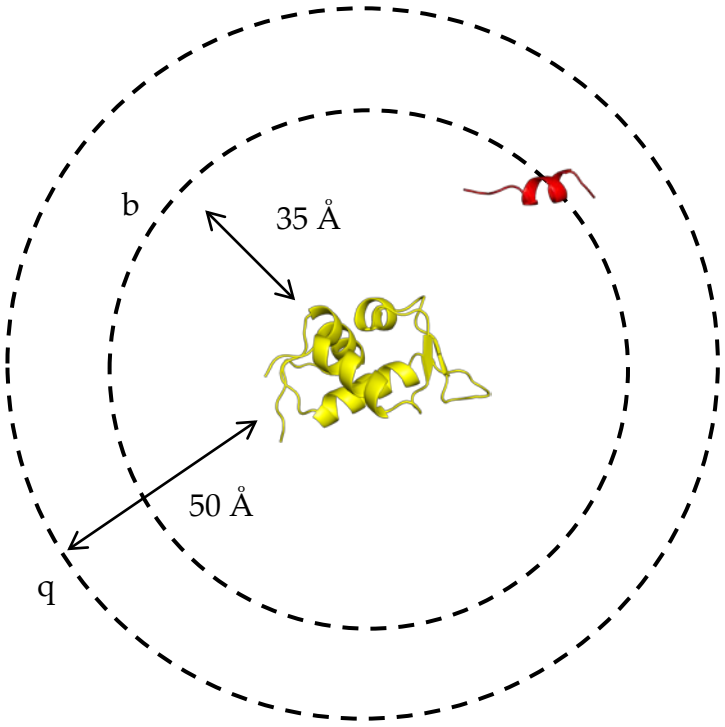

Figure 2. NAM Approach

MDM2 is placed in the middle of two spheres: inner sphere (b) and outer sphere (q). The inner sphere is where all initial states of p53 lie. A simulation can be terminated either when p53 binds MDM2 or p53 diffuses away from the outer sphere.

The rate constant of association, k, can then be calculated with the following relation:

$$k = \frac{k_D(b)\beta}{1 - (1-\beta)\, k_D(b)/k_D(q)}$$

in which $k_D(x)$ is the diffusion rate constant for the two partners achieving a separation distance x, and β is the fraction of total simulations that satisfy the criteria for binding. Assuming that the motions of the two partners are isotropic leads to the Smoluchowski result: $k_D = 4\pi Dx$, in which D represents the relative translational diffusion coefficients of the two partners.

Use of the weighted ensemble algorithm changes the standard brute force computation of β in the NAM approach. Instead, β can be estimated by the following equation (Rojnuckarin, Livesay, & Subramaniam, 2000):

$$\beta = \frac{f_{SS}^{Bind}}{f_{SS}^{Bind} + f_{SS}^{QSurf}}$$

in which f_{SS}^{Bind} is the steady-state binding flux and f_{SS}^{QSurf} is the steady-state flux across the q surface. In addition, the progress coordinate must define a target state for the truncation sphere. Therefore, by monitoring the minimum distance between p53 and MDM2, any trajectory reaching a minimum distance greater than 50 Å, and its probability was recycled to the initial state.

2.3 SIMULATIONS DETAILS

A standard Brownian Dynamics algorithm was employed using the Ermak-McCammon (Ermak & McCammon, 1978) equation:

$$\boldsymbol{r}_i(t + \Delta t) = \boldsymbol{r}_i(t) + \sum_j \frac{\boldsymbol{D}_{ij}\boldsymbol{F}_j \Delta t}{k_b T} + \boldsymbol{R}_i$$

in which $\boldsymbol{r}_i(t)$ is the position vector of the ith pseudoatom at time t, Δt is the simulation time step, $\boldsymbol{D}_{ij}$ is the i,jth 3×3 submatrix of the diffusion tensor $\boldsymbol{D}$ (a $3N \times 3N$ matrix where N is the

number of pseudoatoms in the system), $\boldsymbol{F}_j$ is the force acting on the jth pseudoatom, and $\boldsymbol{R}_i$ is a stochastic term that creates a random displacement of the ith pseudoatom.

In addition, hydrodynamic interactions that model the correlated motions of water are incorporated into the simulation algorithm. Thus, each i,jth submatrix where off-diagonal terms may contain non-zero elements depending on the whether pseudoatoms i and j are hydrodynamically coupled. In addition, each submatrix is calculated using the complete set of equations of Rotne and Prager (Rotne & Prager, 1969) and Yamakawa (Yamakawa, 1970):

$$\boldsymbol{D}_{ii} = (k_B T/6\pi\eta_s a)\boldsymbol{I}$$

$$\boldsymbol{D}_{ij} = (k_B T/6\pi\eta_s)\{(1/r_{ij})[(1 + 2a^2/3r_{ij}^2)\boldsymbol{I} + (1 - 2a^2/r_{ij}^2)(\boldsymbol{r}_{ij}\boldsymbol{r}_{ij}/r_{ij}^2)]\} \text{ for } r_{ij} \geq 2a$$

$$\boldsymbol{D}_{ij} = (k_B T/6\pi\eta_s)\{(1/r_{ij})[(r_{ij}/2a)(8/3 - 3r_{ij}/4a)\boldsymbol{I} + (r_{ij}/4a)(\boldsymbol{r}_{ij}\boldsymbol{r}_{ij}/r_{ij}^2)]\} \text{ for } r_{ij} < 2a$$

where $\boldsymbol{I}$ is the 3 × 3 identity matrix, a is pseudoatom sphere radius, r_{ij} is the distance between pseudoatoms i and j, and $\boldsymbol{r}_{ij}$ is the vector connecting them.

There is evidence that the use of hydrodynamic interactions better models translational diffusion (Frembgen-Kesner & Elcock, 2009). The diffusion coefficients of p53 and MDM2 were calculated using the following equation:

$$D = \frac{\langle x \rangle^2}{6\delta t}$$

where D is the diffusion coefficient, $\langle x \rangle$ is the mean displacement, and δt is the time between sampled displacements.

Within the Brownian Dynamics framework, a 3.5 Å hydrodynamic radius value was utilized (Frembgen-Kesner & Elcock, 2009). This radius for the side-chain model best reproduced the translational diffusion coefficient of the all-atom model estimated by HYDROPRO (Garcia de la Torre, 2001; Garcia De La Torre, Huertas, & Carrasco, 2000). A

time step of 50 fs was used, constraining pseudo bonds between residues to their native bond lengths using the LINCS algorithm (Hess, Bekker, Berendsen, & Fraaije, 1997). Protein conformations were sampled every 100 ps for brute force control simulations and every 1 ps for weighted ensemble binding simulations.

Weighted ensemble binding simulations were performed between preorganized p53 with MDM2 and flexible p53 with MDM2. In all binding simulations, initial p53 conformations were selected from a list of 3000 randomly generated conformations taken from brute force simulations of the unbound p53 peptide. The initial state of the system was obtained by randomly reorienting both p53 and MDM2 and separating the two proteins by a minimum distance of 35 Å (b). Trajectories ending in the target state are recycled to the initial state with a randomly chosen conformation and calculated orientation. The chosen τ used in the weighted ensemble algorithm for these simulations was 100 ps. This value was selected to be short enough to maximize the use of statistics but not too short due to the overhead cost of implementing the weighted ensemble code. The bin spacing along the RMSD coordinate (discussed above) was determined such that, ideally, one segment would traverse a bin in one iteration of time τ.

3.0 RESULTS AND DISCUSSION

3.1 SUCCESSFUL TUNING OF PEPTIDE FLEXIBILITY

Flexibility was tuned by adjusting ε^{native}, the potential well depths of the native contacts. Two different values of ε^{native} were attained to satisfy the criteria (see Methods) for flexible and preorganized p53: 0.05 for flexible and 1.70 for preorganized. Histograms compiling data for each individual set of simulations are shown in Figure 3.

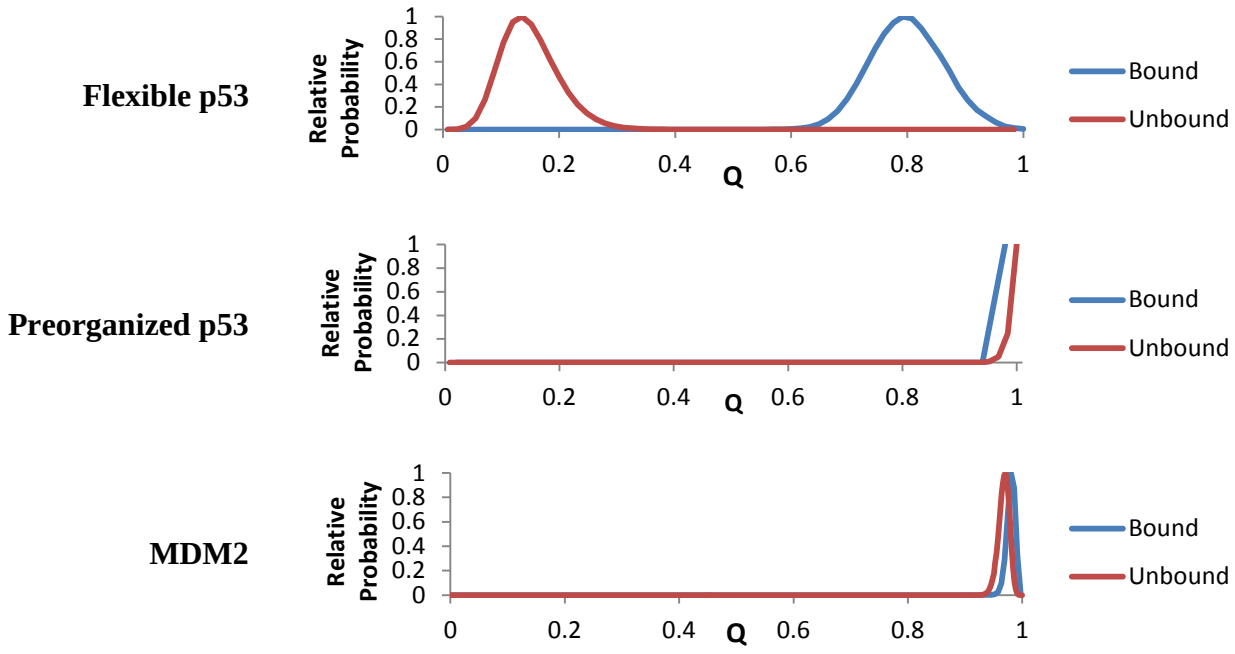

Figure 3. Stability of Flexible/Preorganized p53 and MDM2

Histograms, attained from brute force simulations, of relative probability against intramolecular fraction of native contacts helps determine stability of a protein. "Bound" indicates simulations of the p53-MDM2 complex and "Unbound" denotes simulations of each peptide alone.

Each histogram illustrates the probability of conformations with a certain fraction of native contacts, Q. Preorganized p53 peptide remained folded with probability only for $Q > 0.8$ in p53-only simulations; flexible p53 only assumed conformations for which $Q < 0.5$. Depictions comparing the different conformations attained by flexible and preorganized p53 are shown in Figure 4.

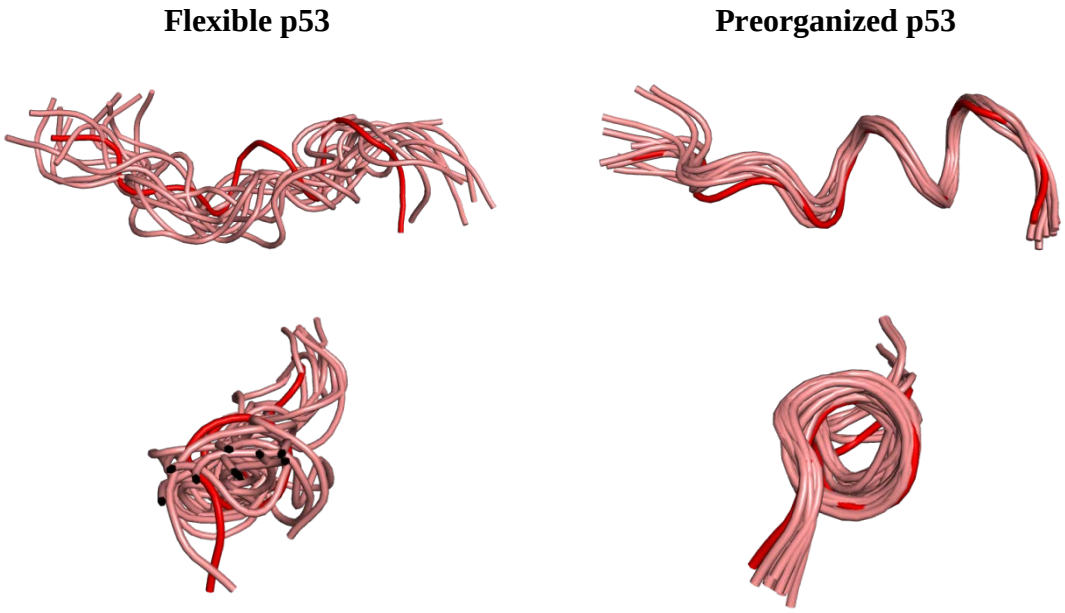

Figure 4. Comparison of Flexible and Preorganized p53

Ribbon diagrams of p53 are randomly selected snapshots taken from their respective brute force simulations. The red ribbon is representative of the crystal structure. This juxtaposition of conformations provides a perspective on what it means for p53 peptide to be flexible or preorganized.

In addition, simulations beginning from the bound complex yielded results consistent with the criterion, $Q > 0.5$, for flexible p53 when bound to MDM2. Finally, overall MDM2 stability was assessed. Not only was the total fraction of native contacts examined, but the fraction of native contacts per residue was also investigated. The per-residue fraction of native contacts gives more detailed insight into whether local regions of the protein are changing in stability that may not have been observed in the total fraction of native contacts analysis. The

results of this analysis are displayed in Figure 5, where fraction of native contacts by residue is color coded. This little, if any, change in coloration of MDM2 gives confidence that binding p53 does not affect the stability of MDM2.

These findings confirm that tuning the native contact potentials is effective in altering the flexibility of p53. In addition, the use of strong intermolecular MDM2-p53 native contacts could stabilize the fully flexible p53 peptide. Furthermore, MDM2 stability is unaffected by binding to p53, relieving the possibility of errors associated with such effects.

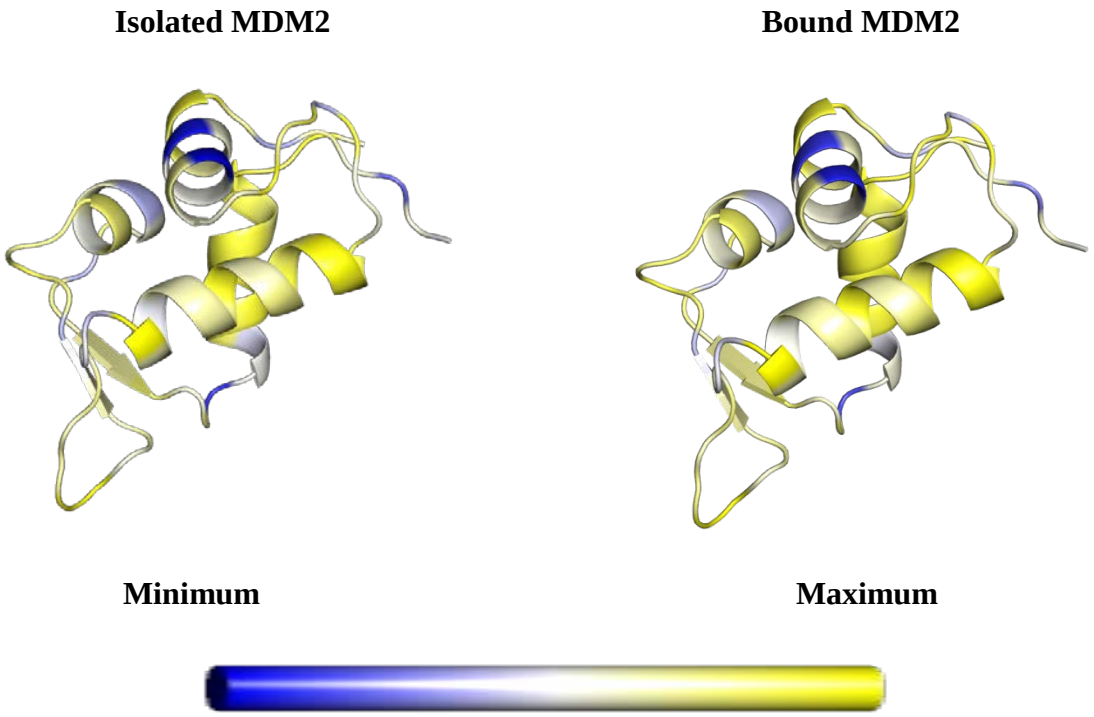

Figure 5. MDM2 Fraction of Native Contacts by Residue

Per-residue fraction of native contacts provides a more detailed description of which residues tend to form more contacts. “Maximum” is defined as the highest per-residue fraction of native contacts observed, and vice versa for “Minimum.” This comparison shows that there is little difference between isolated MDM2 and bound MDM2 in terms of stability.

3.2 DIFFUSION AND ASSOCIATION OF FLEXIBLE VS. PREORGANIZED P53

Calculation of binding rates was achieved by implementing the NAM method in conjunction with the weighted ensemble approach (see Methods). This approach requires simulations to reach steady state because the steady state fluxes into target states are necessary in the computation. Recall that the two target states were the bound state and all regions outside of the outer sphere. Therefore, weighted ensemble simulations were carried out until fluxes into those states remained constant over time. Plots of flux over the iterations of weighted ensemble are displayed in Figure 6 for binding simulations with hydrodynamics. Figure 7 displays the fluxes into the target states of simulations without hydrodynamics.

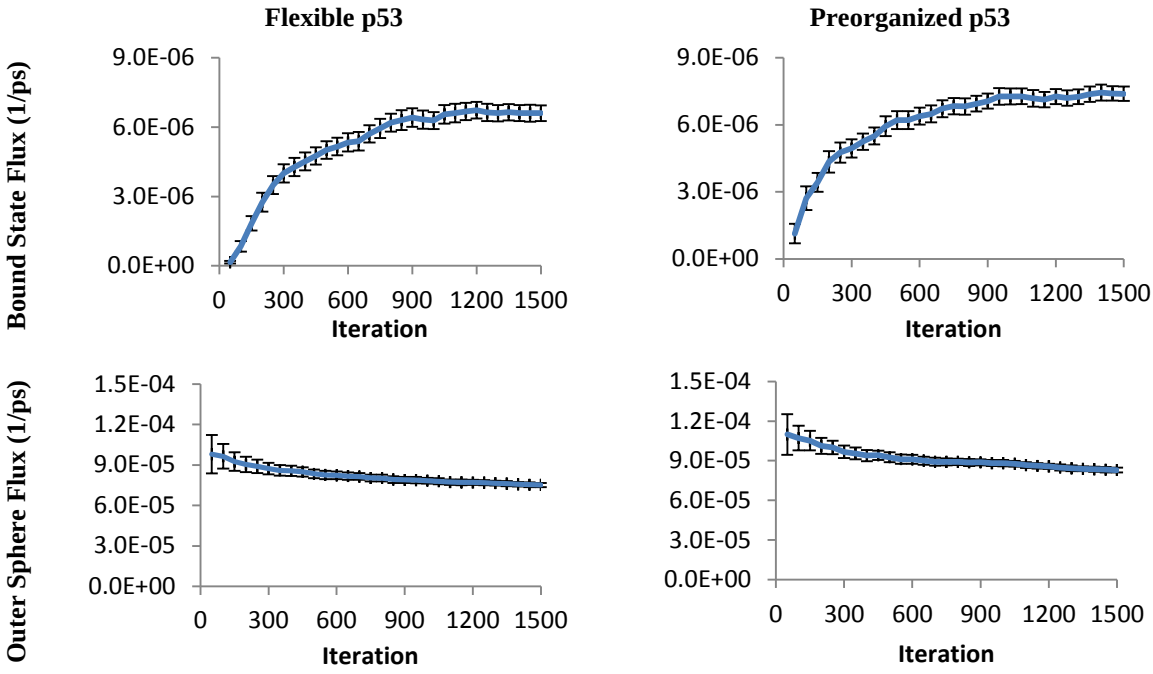

Figure 6. Flexible/Preorganized p53 Target State Fluxes With Hydrodynamics

Target state fluxes are used to gage convergence of simulations. 95% confidence intervals are displayed every 50 iterations.

Table 1 contains information for diffusion coefficients, rates of association, and number of binding events. The translational diffusion coefficients were very similar between flexible and preorganized p53 with hydrodynamics. It is noteworthy, though, that flexible p53 had a slightly smaller diffusion coefficient, which is expected because flexible p53 has a larger radius of gyration (Figure 8), which effectively increase its surface area and lowering its diffusion rate. Also, the diffusion coefficients for both flexible and preorganized p53 are comparable to the ones produced by HYDROPRO. In addition, simulations without hydrodynamics exhibited diffusion coefficients of an order of magnitude smaller than those with hydrodynamics, indicating that hydrodynamics are key to reproducing diffusion.

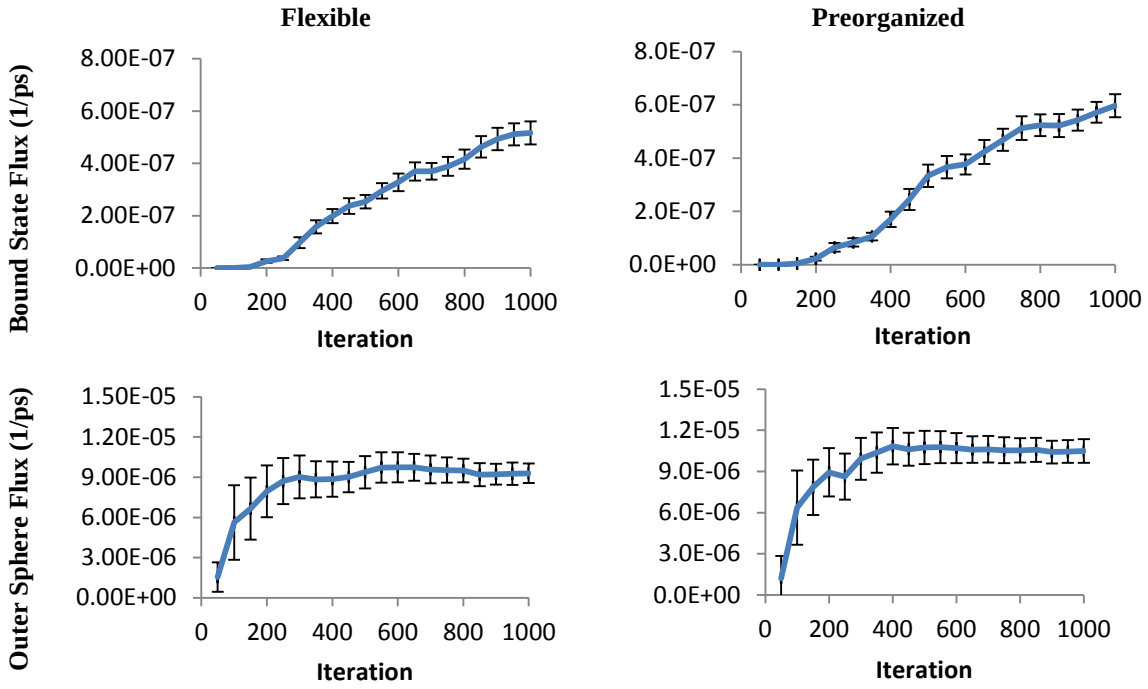

Figure 7. Flexible/Preorganized p53 Target State Fluxes Without Hydrodynamics

Fluxes into the bound state and out of the outer sphere are plotted over time. While the outer sphere fluxes appear to have leveled off, the bound state fluxes are still on the rise. Therefore, these simulations have yet to converge.

Computed rate constants for the binding of p53 to MDM2 indicate that flexible p53 binds MDM2 faster than does preorganized p53 by a factor of 1.04. However, this ratio is not statistically significant as the 95% confidence intervals of the two rates overlap.

Table 1: Relative Diffusion Coefficients and Rates of Association

	With Hydrodynamics		Without Hydrodynamics*	
	Flexible	Preorganized	Flexible	Preorganized
Relative Diffusion Coefficient (cm^2/s)	$3.91 \cdot 10^{-9}$	$3.98 \cdot 10^{-9}$	$4.95 \cdot 10^{-10}$	$4.95 \cdot 10^{-10}$
HYDROPRO Diffusion Coefficient	$3.63 \cdot 10^{-9}$		NA	
Number of Successful Binding Events	16,488	41,461	24,107	53,338
Binding Rate Constant ($M^{-1}s^{-1}$)	$2.51 \cdot 10^6 \pm 0.14 \cdot 10^6$	$2.41 \cdot 10^6 \pm 0.12 \cdot 10^6$	$3.38 \cdot 10^5 \pm 0.33 \cdot 10^5$	$3.23 \cdot 10^5 \pm 0.26 \cdot 10^5$

* Because these particular simulations had not reached steady state, the binding rate constants are subject to change.

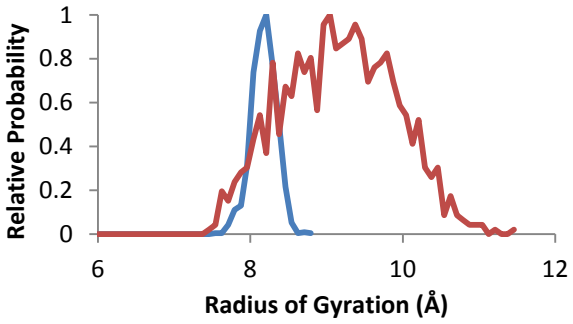

Figure 8. Radius of Gyration

The mean radius of gyration was larger for flexible p53 than for preorganized p53 peptide. In addition, the fluctuations were larger for flexible p53, which is expected for a flexible peptide.

The overall results do not support the predictions of the fly-casting hypothesis as they pertain to the kinetic advantages of intrinsic disorder. However, a point can be made for the small, counterbalancing effects of diffusion and radius of gyration. The association rate is directly proportional to the rate of diffusion, thereby creating a slight kinetic disadvantage for flexible p53. At the same time, the larger radius of gyration for flexible p53 increases its capture

radius for MDM2, offering a slight kinetic advantage. These two opposing effects may cancel each other out, providing some explanation for the insignificant change in the binding rate constant.

A future set of simulations will be conducted to compare binding affinities of flexible and preorganized p53 to MDM2. Although the same energies were used for the intermolecular native contacts for both cases, we cannot be certain that the binding affinities were equivalent. To control for equal binding affinities, the intermolecular native contact potentials can be altered. Because $\varepsilon_{inter}^{native} = 6.0$ was chosen such that flexible p53 would have $Q > 0.5$ when bound to MDM2, only $\varepsilon_{inter}^{native}$ for preorganized p53 binding simulations will be changed to tune the binding affinity.

3.3 MECHANISM OF BINDING

Observing the mechanism of binding requires a large ensemble binding events, which was accomplished in this study. In simulations with hydrodynamics, flexible p53 peptide had 16,488 successful binding events while preorganized p53 peptide had 41,461 binding events (see Table 1). Interestingly, flexible p53 had far fewer successful binding events than did preorganized p53 yet exhibited a similar association rate constant—a similar situation applies for simulations without hydrodynamics. This suggests, then, that the average successful binding event carried greater particle weight in the weighted ensemble framework for flexible p53 compared to preorganized p53. In other words, the percentage of collisions resulting in binding events is higher for the flexible p53 peptide. These observations are consistent with those reported in Huang and Liu (Huang & Liu, 2009) for the pKID-KIX system.

A possible reason for this finding is that there exists a coupling of folding and binding. Figure 9 plots the intramolecular p53 contacts, Q_{p53}, against the intermolecular p53-MDM2 contacts, $Q_{p53-MDM2}$, for distinct binding trajectories for flexible p53. If there were coupling of folding to binding, one would expect to see a correlated increase of both measures once p53 and MDM2 come into contact and initiate their eventual binding. Indeed, we observed that the number of intramolecular contacts in the p53 peptide increased with the number of intermolecular contacts, thus indicating a coupled folding-binding process.

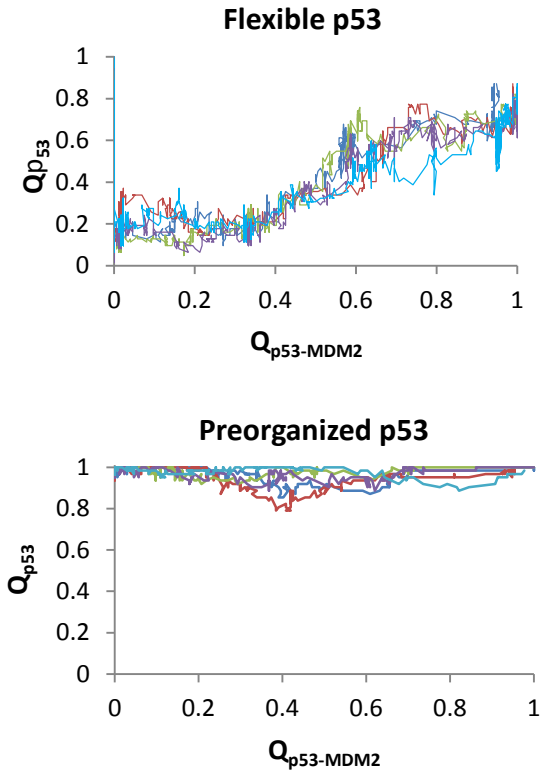

Figure 9. Coupling of Intermolecular and Intramolecular Contacts

Intramolecular p53 fraction of native contacts is plotted against intermolecular fraction of native contacts between p53 and MDM2 for 5 distinct weighted ensemble trajectories. For the flexible case, there is good correlation between Q_{p53} and $Q_{p53\text{-}MDM2}$. While the preorganized peptide generally remains very folded, there occurs a periods of instability during the binding process.

Because preorganized p53 was parameterized to remain folded, no observation of a coupled folding and binding process was observed. However, an interesting feature appeared

when examining the plot for preorganized p53 (Figure 9, bottom). The trajectories indicated destabilization of p53 during the binding process as evidenced by small decreases in Q_{p53}. This suggests that in order to fully bind MDM2, some conformational rearrangement of p53 was necessary. Further inspection is necessary to determine what structural change is occurring during these moments of destabilization.

4.0 LIMITATIONS OF THE SIMULATION MODEL

Our simulation model allows for direct comparison between flexible vs. preorganized cases of a peptide. Relevant in this study is the determination of their relative rates of association with their partner protein. The simplicity of these protein models enables a comprehensive account of binding kinetics within a reasonable amount of computer time. Nonetheless, there are certain properties that this model overlooks. The major approximations and limitations of these simulations are discussed below.

One simplification is in the use of Gō-type potentials that considerably smooth the energy landscape. The lack of favorable non-native contacts may hinder our observation of some mechanisms of folding and binding. In addition, studies have shown that the simulated protein folding mechanisms, which lead to accelerated folding times, are not consistent with either experiment (Clementi, Nymeyer, & Onuchic, 2000; Koga & Takada, 2001) or all-atom molecular dynamics simulations (Daggett & Fersht, 2003). However, there is evidence that relative rates of folding can be in agreement with experiment (Chavez, Onuchic, & Clementi, 2004). Also, our inclusion of hydrodynamics and the hydrodynamic radii of the pseudoatoms results in good approximations of translational diffusion coefficients. Hence, this study focuses on the ratio of calculated binding rates for flexible vs. preorganized and simulations with or without hydrodynamics.

Another limitation of these simulations is the attainable level of conformational detail. Unlike all-atom simulations, only residue-level detail can be acquired after coarse-graining. This was crucial, though, for generating thousands of binding events (in a reasonable amount of time) necessary to calculate converged rates of association. In addition, these coarse-grained side-chain models do offer more detail than commonly used Cα models in which each residue is modeled by one pseudoatom. Furthermore, side chain interactions were found to be essential for correct α-helical formation in the flexible version of p53 upon binding MDM2, likely due to the introduced steric effects. These simulations also exclude electrostatics. Because both p53 and MDM2 have net charges— -2 for p53 and +5 for MDM2—one might suspect that electrostatics play an important role in the binding of p53 to MDM2. However, it has been shown these net charges do not play a major role in p53-MDM2 binding (Schon, Friedler, Bycroft, Freund, & Fersht, 2002). Therefore, our model minimizes the computational costs of protein binding while retaining a sufficient amount of detail for generating coupled folding and binding events.

5.0 CONCLUSION

Protein binding involving intrinsically disordered proteins is of great interest. The proposal of a "fly-casting" mechanism emphasized the kinetic aspects of the association between an intrinsically disordered protein and its binding partner. We studied this hypothesis computationally with the p53-MDM2 complex using a protein model capable of producing flexible and preorganized versions of the p53 peptide. In addition, we investigated the effects of hydrodynamics on the kinetics of binding.

Use of hydrodynamic interactions results in larger rates of association by an order of magnitude compared to corresponding simulations without hydrodynamics. This is due to the order of magnitude difference between the peptides' translation diffusion coefficients. Many implicit solvent models do not include hydrodynamic interactions and hence do not simulate the effects of the correlated motions of water. However, these effects can prove to be important in calculating rates of certain processes such as protein binding. Therefore, inclusion of hydrodynamic interactions in implicit water simulations can be crucial.

In comparing the kinetics between flexible and preorganized p53 binding to MDM2, our results indicate no significant change in their rates of association. Despite this, there were significantly fewer binding events observed in the flexible p53 simulations, indicating that flexible p53 peptide was more efficient at binding upon encountering MDM2. We attribute this fact to the observed coupling of folding and binding.

Overall, this study does not support the predictions of the fly-casting mechanism. In addition, it should be noted that our protein model creates a significantly smoother energy landscape. Because of this, rates associated with all processes tend to be amplified, implying that the small—and statistically insignificant—changes observed here are negligible. It must be noted, though, that fly-casting is predicated upon the fact that flexible proteins have larger capture radii. Therefore, larger proteins would benefit more from fly-casting. This study investigates a 13-residue fragment of p53, which may be less effective from the fly-casting perspective. Nonetheless, we observe no change in the rates of association of flexible and preorganized p53 binding to MDM2.

BIBLIOGRAPHY

Chavez, L. L., Onuchic, J. N., & Clementi, C. (2004). Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates. *J Am Chem Soc, 126*(27), 8426-8432.

Clementi, C., Nymeyer, H., & Onuchic, J. N. (2000). Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J Mol Biol, 298*(5), 937-953.

Crespin, M. O., Boys, B. L., & Konermann, L. (2005). The reconstitution of unfolded myoglobin with hemin dicyanide is not accelerated by fly-casting. *FEBS Lett, 579*(1), 271-274.

Daggett, V., & Fersht, A. (2003). The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol, 4*(6), 497-502.

Elcock, A. H. (2006). Molecular simulations of cotranslational protein folding: fragment stabilities, folding cooperativity, and trapping in the ribosome. *PLoS Comput Biol, 2*(7), e98.

Ermak, D. L., & McCammon, J. A. (1978). Brownian dynamics with hydrodynamic interactions. *J Chem Phys, 69*(4), 1352-1360.

Frembgen-Kesner, T., & Elcock, A. H. (2009). Striking Effects of Hydrodynamic Interactions on the Simulated Diffusion and Folding of Proteins. *J Chem Theory Comput, 5*(2), 242-256.

Garcia de la Torre, J. (2001). Hydration from hydrodynamics. General considerations and applications of bead modeling to globular proteins. *J Biophys Chem, 93*(2-3), 159-170.

Garcia De La Torre, J., Huertas, M. L., & Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J, 78*(2), 719-730.

Go, N. (1983). Theoretical studies of protein folding. *Annu Rev Biophys Bioeng, 12*, 183-210.

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *J Comput Chem, 18*(12), 1463-1472.

Hoffman, R. M., Blumenschein, T. M., & Sykes, B. D. (2006). An interplay between protein disorder and structure confers the Ca2+ regulation of striated muscle. *J Mol Biol, 361*(4), 625-633.

Huang, Y., & Liu, Z. (2009). Kinetic Advatage of Intrinsically Disordered Proteins in Coupled Folding-Binding Processs: A Critical Assessment of the "Fly-Casting" Mechanism. *J Mol Biol, 393*(5), 1143-1159.

Huber, G. A., & Kim, S. (1996). Weighted-ensemble Brownian dynamics simulations for protein association reactions. *Biophys J, 70*(1), 97-110.

Jemth, P., & Gianni, S. (2007). PDZ domains: folding and binding. [

Review]. *Biochemistry, 46*(30), 8701-8708.

Koga, N., & Takada, S. (2001). Roles of native topology and chain-length scaling in protein folding: a simulation study with a Go-like model. *J Mol Biol, 313*(1), 171-180.

Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., & Pavletich, N. P. (1996). Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science, 274*(5289), 948-953.

Landfried, D. A., Vuletich, D. A., Pond, M. P., & Lecomte, J. T. (2007). Structural and thermodynamic consequences of b heme binding for monomeric apoglobins and other apoproteins. *Gene, 398*(1-2), 12-28.

Lengyel, C. S., Willis, L. J., Mann, P., Baker, D., Kortemme, T., Strong, R. K., & McFarland, B. J. (2007). Mutations designed to destabilize the receptor-bound conformation increase MICA-NKG2D association rate and affinity. *J Biol Chem, 282*(42), 30658-30666.

Levy, Y., Onuchic, J. N., & Wolynes, P. G. (2007). Fly-casting in protein-DNA binding: frustration between protein folding and electrostatics facilitates target recognition. *J Am Chem Soc, 129*(4), 738-739.

Mills, B. M., & Chong, L. T. (2011). Molecular simulations of mutually exclusive folding in a two-domain protein switch. *Biophys J, 100*(3), 756-764.

Muralidhara, B. K., Rathinakumar, R., & Wittung-Stafshede, P. (2006). Folding of *Desulfovibrio desulfuricans* flavodoxin is accelerated by cofact fly-casting. *Arch Biochem Biophys, 451*, 51-58.

Narayanan, R., Ganesh, O. K., Edison, A. S., & Hagen, S. J. (2008). Kinetics of folding and binding of an intrinsically disordered protein: the inhibitor of yeast aspartic proteinase YPrA. *J Am Chem Soc, 130*(34), 11477-11485.

Northrup, S. H., Allison, S. A., & McCammon, J. A. (1984). Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J. Chem. Phys. , 80*(4), 1517-1524.

Onitsuka, M., Kamikubo, H., Yamazaki, Y., & Kataoka, M. (2008). Mechanism of induced folding: Both folding before binding and binding before folding can be realized in staphylococcal nuclease mutants. *Proteins, 72*(3), 837-847.

Perham, M., Chen, M., Ma, J., & Wittung-Stafshede, P. (2005). Unfolding of heptameric co-chaperonin protein follows "fly casting" mechanism: observation of transient nonnative heptamer. *J Am Chem Soc, 127*(47), 16402-16403.

Rojnuckarin, A., Livesay, D. R., & Subramaniam, S. (2000). Bimolecular reaction simulation using Weighted Ensemble Brownian dynamics and the University of Houston Brownian Dynamics program. *Biophys J, 79*(2), 686-693.

Rotne, J., & Prager, S. (1969). Variational Treatment of Hydrodynamic Interaction in Polymers. *J Chem Phys, 50*(11).

Schon, O., Friedler, A., Bycroft, M., Freund, S. M., & Fersht, A. R. (2002). Molecular mechanism of the interaction between MDM2 and p53. [Research Support, Non-U.S. Gov't]. *J Mol Biol, 323*(3), 491-501.

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. USA, 97*(16), 8868-8873.

Sugase, K., Dyson, H. J., & Wright, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature, 447*(7147), 1021-1025.

Sugase, K., Lansing, J. C., Dyson, H. J., & Wright, P. E. (2007). Tailoring relaxation dispersion experiments for fast-associating protein complexes. *J Am Chem Soc, 129*(44), 13406-13407.

Takada, S. (1999). Go-ing for the prediction of protein folding mechanisms. *Proc Natl Acad Sci U S A, 96*(21), 11698-11700.

Turjanski, A. G., Gutkind, J. S., Best, R. B., & Hummer, G. (2008). Binding-induced folding of a natively unstructured transcription factor. *PLoS Comput Biol, 4*(4), e1000060.

Vamvaca, K., Jelesarov, I., & Hilvert, D. (2008). Kinetics and thermodynamics of ligand binding to a molten globular enzyme and its native counterpart. *J Mol Biol, 382*(4), 971-977.

Wright, P. E., & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol, 293*(2), 321-331.

Yamakawa, H. (1970). Transport Properties of Polymer Chains in Dilute Solution: Hydrodynamic Interaction. *J Chem Phys, 53*(1), 436-443.