Kinetics and Structure of p53 Amyloid Formation

Diana Qiu
Adviser: Daniel Cox
Discussion and Assistance from: Kay Kunes, Jamie Romnes
REU Program
UC Davis

Abstract

Recent studies reveal that the tumor suppressor protein p53 can aggregate as an amyloid in its N-terminal, DNA binding domain (DBD), and tetramerization domain. This suggests that the mechanism for p53 loss of function, and consequently the pathology of cancer onset, may be similar to that of other amyloid-producing diseases, such as prion diseases and Alzheimer's Disease. In this study, we explore potential similarities by analyzing the kinetics of apoDBD and wtDBD aggregation and by running molecular dynamics on proposed structures for aggregates of the p53 N-terminal, DBD, and tetramerization domain. We found that apoDBD and wtDBD aggregates undergo exponential growth and have differing critical nuclei. In molecular dynamics simulations, the left-handed β-helix appears to be a stable model for monomer structure of aggregates of the DBD and the N-terminal, while simulations of the tetramerization domain were inconclusive.

Introduction

There are two main types of protein conformational diseases. The first, which includes diseases like cystic fibrosis, occurs when protein misfolding leads to loss of function. The second, which includes amyloidoses like Alzheimer's disease and prion diseases, occurs when misfolded proteins form toxic aggregates in cells. Consequently, the mechanisms of protein folding and aggregation have become an important field of study. P53, the protein encoded by the TP53 gene, is a transcription factor responsible for halting cell division and inducing apoptosis, or programmed cell death, after DNA damage occurs in a cell. Since cancer onset occurs due to DNA mutations that impact cell growth, p53 plays a vital role in preventing tumor growth. Loss of function of p53 has been linked to onset of over fifty percent of cancers. Recent studies reveal that mechanisms for p53 loss of function and hence mechanisms for cancer onset may resemble those of prion diseases. Both Ishimaru et al. (2003) and Butler and Loh (2003) observed that mutant and wild-type forms of the DNA binding domain of p53 can aggregate into amyloid structures *in vitro*, and Rigacci et al. (2008) observed that mutant and wildtype forms of the p53 N-terminus are also capable of forming amyloid. Amyloid formation is characteristic of gain of function protein diseases like prion diseases and Alzheimer's disease. Both prions and the Aβ peptide in Alzheimer's Disease catalyze the transition of wild-type proteins from their native to their aggregated states after aggregates exceed a critical concentration. This has led to speculation that p53 may be capable of similar autocatalysis, and thus, aggregation may

play a vital role in p53 loss of function and ensuing cancer onset.

Background

Human p53 is 393 residues long and consists of an unstructured N-terminal (residues 1-63), a DNA binding domain (residues 102-292), and a C-terminal (residues 293-393) that contains a tetramerization domain. P53 binds DNA as a tetramer. The tetramerization domain consists of a β-strand and an α-helix. Hydrogen bonds between the β-strands and an intermonomer salt bridge allow the proteins to form a tetramer, thus orienting them for sequence specific DNA binding (Galea, 2005). The DNA binding domain (DBD) consists of two unstructured loops (L2 and L3) separated by beta sheet. C176 and H179 in L2 and C238 and C242 in L3 coordinate a Zn^{2+}. The Zn ion creates a hydrophobic core and stabilizes L3, allowing the loop to bind to DNA's minor groove and phosphoric backbone (Butler and Loh, 2003). The DBD contains 97% of cancer-related missense mutations. These mutations lead to loss of function by either changing the conformation of DNA binding sites or decreasing thermodynamic stability of the entire protein.

Correct folding is essential to p53 function. Typically a protein's native conformation is its lowest energy state, but proteins may misfold if they become stuck in local energy minima during the folding process (Hoshino, 2006). Misfolded proteins often have exposed hydrophobic residues, which makes aggregation energetically favorable, and thus, tend to aggregate upon collision with other misfolded proteins. The DBD of p53 cycles between its folded and unfolded states. It has four folding pathways: two direct paths and two slow paths that contain off-pathway intermediaries. On the slow paths, p53 may be caught in an intermediary state for anywhere from 5.3 minutes to 5 hours, as opposed to a few seconds on the fast paths, thus increasing the probability that partially folded proteins will collide and aggregate (Butler and Loh, 2005). This accounts for spontaneous loss of function of wt-p53. Butler and Loh (2005) observed that under physiological conditions the point mutations G245S, R249S, and R282Q do not affect folding kinetics but instead increase the rate of unfolding, thus increasing the potential for aggregation and permanent loss of function. ApoDBD, the Zn-free form of p53, also aggregates at a faster rate than the protein's zinc-binding form. Butler and Loh (2003) observed that apoDBD catalyzes the aggregation of wtDBD and that the concentration of both aggregates increases exponentially.

The tetramerization domain of p53 is also capable of aggregation. The mutation R337H in the tetramerization domain is linked to adrenocortical carcinoma in children. R337H disrupts the salt bridge between R337 and D352 that allows for tetramerization. This disruption leads to an increased tendency to form amyloid fibrils at low pH (Lee et al., 2003; Galea et al., 2005). The formations reversed when pH was raised to 8.5. Higashimoto et al. (2006) found that G334V, a residue bridging the α and β helices in the tetramerization domain of p53, forms amyloid fibrils under physiological conditions and that this mutant form can cause the wild-type form to aggregate by destabilizing p53 tetramers. G334V mutations are linked to small cell lung and squamous cell carcinomas.

Lastly, the N-terminal is also capable of amyloid formation. The N-terminal is unusual because its native state is unstructured, yet it is not prone to

aggregation under physiological conditions. It does, however, form amyloid fibrils at low pH (between 3.0 and 4.0), a condition which may be encountered *in vivo* in some organelles (Rigacci et al., 2008). Since the N-terminal is unstructured, Rigacci et al. (2008) have suggested that sufficient crowding, combined with accidental exposure to low pH, may cause the N-terminal to catalyze p53 aggregation.

Loss of function of p53 is the primary cause of failure for cancer therapies like radiation and chemotherapy. These therapies kill fast-dividing tumor cells by causing massive DNA damage that triggers apoptosis through p53 response. Similarities between p53 and prion amyloid formation suggest that aggregation may actively contribute to wild-type p53 loss of function. Formation of p53 aggregation nuclei (either through spontaneous misfolding or mutations) could catalyze aggregation of wt-p53 by shifting the energy landscape in favor of aggregation inside both the cell and its daughter cells. Thus, the study of p53 aggregation kinetics and structure could contribute to a new understanding of cancer onset, and disruption of the aggregation pathway could lead to new cancer therapies.

Methods

Kinetics

Calculations of aggregation kinetics were performed on data from figure 7 of Butler and Loh's (2003) Structure, Function, and Aggregation of theZinc-free form of the p53 DNA Binding Domain (figure 1).

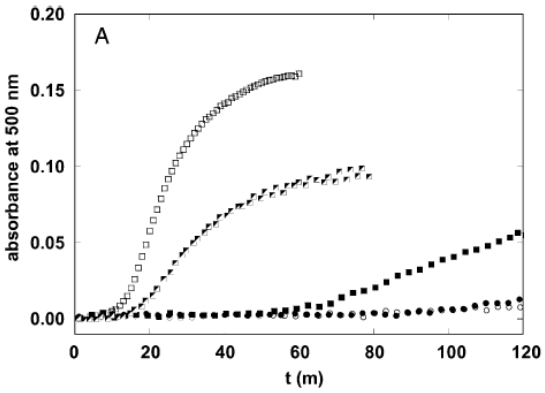

Figure 1: Filled squares are 3.5uM wtDBD; Empty squares are 3.5uM apoDBD; Filled circles are 2.82uM wtDBD; Empty circles are .88uM apoDBD; Half-filled squares are .88uM apoDBD and 2. 88uM wtDBD mixed; (Butler and Loh, 2003)

Aggregation data of apoDBD at .88 and 3.5uM and wtDBD at 2.63 and 3.5uM were fitted to the curve

$$[M'] = A(\cosh(Bt) - 1), \qquad [1]$$

where [M'] is the concentration of polymerized monomers, and

$$A = \frac{J_{N,0}[N_0]}{Q - \frac{d}{dc}(J_N[N])} \qquad [2]$$

$$B^2 = J_0\left(Q - \frac{d}{dc}(J_N[N])\right), \qquad [3]$$

where J is the elongation rate of the polymers, J_N is the elongation rate of the aggregation nuclei, [N] is the concentration of nuclei, and Q is an unknown pathway for exponential growth. Equation [1] describes early aggregation with exponential time dependence; it is the solution to the following equations:

$$\frac{d[P]}{dt} = J_{N,0}[N_0] + [M'](Q - \frac{d(J_N[N])}{d[M]}) \qquad [4]$$

$$\frac{d[M']}{dt} = J_0[P_0] \qquad [5]$$

where [M] is the concentration of monomers and $[P_0]$ is the initial concentration of polymers. Maximum absorbance was interpreted as 100% aggregation with no monomers remaining. A and B^2 were determined from fitting [1] to the data in Butler and Loh (2003). Then, AB^2 was calculated:

$$AB^2 = J_0 J_{N,0}[N_0] \quad [6a]$$

Or at early times:

$$AB^2 = J J_N[N] \quad [6b]$$

The elongation rates J and J_N are dependent on [M]. Assuming the forward rates $k_+ >> k_-$

$$J = k_+[M] \text{ and } J_N = K_{N,+}[M] \quad [7]$$

$$[N] = K_N[M]^n \quad [8]$$

Then,

$$AB^2 = K[M]^{n+2} \quad [9]$$

where the rate constants k and K describe the addition of one monomer to an aggregate, and *n* is the size of the critical nucleus. The critical nucleus is the aggregate size beyond which free energy of aggregation decreases. Since AB^2 is known, the size of the critical nucleus *n* is found by finding the slope of the linear fit of the log of equation [9]. These calculations are based on methods described in Ferrone (1999).

Threading

Residues 1-63 **(MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEA)** in the N-terminal domain, residues 102-296 **(SDGLAPPQHLIVEGNLRVEYLDDRNTFRSVVVPYEPPEVGS DCTTIHYYMCNSSCMGGMNRRPILTIDSSGNLLGRNSFEV RVCCPGRDRRTEEENLRKKGEPH)** in the apoDBD, and residues 325-356 **(GEYFTLQIRGRERFEMFRELNEALELKDAQAG)** in the tetramerization domain of wtp53 were each manually threaded to a left-handed β-helix with 18 residues per turn. The β-helix is a triangular structure with 6 positions for residues on each side. Positions 3 and 5 point toward the center of the helix, and positions 1, 2, and 6 and sit on the corners of the triangle. For each thread, the amino acid sequence was arranged so that prolines were on the corners of the β-helix (positions 1,2,6), and charged residues pointed outwards (positions 1,2,4,6), making the helix's internal net-charge zero. Also, hydrophobic residues pointed inward (positions 3,5)and hydrophilic residues pointed outward. For the DBD, which contains 10 cysteines, each cysteine was arranged a multiple of 18 positions apart from another cysteine so that it was immediately above or below another cysteine on the helix and thus could form disulfide bonds with it. Residues which could not fit correctly into one of the 6 positions on each side of the helix were pulled out in loops or moved forward by cutting a corner position.

A model of each thread was constructed using the SCAP and LOOPY programs in Jackal, a protein modeling software package developed by the Honig Lab at Columbia University (Xiang, 2002). Residues on a left-handed β-helix template with 18 residues per turn were mutated to p53 thread residues using SCAP, which predicts conformations of mutated residues. Loops and cuts were then inserted using LOOPY, which predicts loop conformations. The β-helix template was constructed by removing loops and cuts from the second helix of the structure of E. coli LpxA (Williams et al., 2006), which was obtained from the Protein Data Bank.

Two monomeric models were developed for both the DBD and the N-terminal. One monomeric model of the tetramerization domain was developed.

Molecular Dynamics

The stability of each model was analyzed using the AMBER8 molecular dynamics simulation package developed by Case et al. (2004) at the University of California San Francisco. First, the N-terminal and the tetramerization domain were protonated because N-terminal and tetramerization domain aggregates have only been observed under acidic condition (Rigacci et al., 2005, Lee et al., 2003). Disulfide bonds were inserted between bonded cysteines in the DBD models. The LEaP program was used to add neutralizing counterions (Cl^-), and solvent to each model. Solvent was created using the *sovateoct* command, which creates a periodic box of water molecules in the shape of a truncated octahedron. Then, the *saveamberparm* command converted each model file into a topology file and a coordinate file.

The molecular dynamics were done with the Sander module in AMBER. Sander calculates the positions of particles during a simulation with force fields. The basic force field is calculated from the system's potential energy, which includes the sum of energies due to covalent bonds stretching, the sum of energies due to covalent bonds bending, the sum of torsional energies, the potential due to van der Waals forces, and electrostatic potential. The basic equation for AMBER force fields is as follows (Case et al., 2004):

$$U(\boldsymbol{R}) = \sum_{bonds} K_r\left(r - r_{eq}\right)^2 + \sum_{angles} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{dihedrals} \frac{V_n}{2}(1 + \cos\left(n\varphi - \gamma\right)) + \sum_{i<j}^{atoms} \frac{A_{ij}}{R_{ij}^{ij}} - \frac{B_{ij}}{R^6} + \sum_{i<j}^{atoms} \frac{q_i q_j}{\varepsilon R_{ij}}$$

.

The force on each atom is calculated from the gradient of the potential. Force field ff99 was used in this study. First, each system underwent 2000 cycles of energy minimization with all carbon atoms frozen. This brought the system to a local energy minimum and eliminated side-chain clashes in the model. Then, each system underwent solvent equilibration. Carbon atoms were restrained while the solvent temperature was raised from 0K to 300K over the course of 4ps. Then, all atoms were released and underwent a molecular dynamics simulation of 10ns, where force field calculations were done every .002ps.

Results and Discussion

Kinetics

The critical nucleus size of apoDBD was determined to be 1. The critical nucleus size of wtDBD was determined to be 3. Due to limited experimental data these results were based on two data points each and are not conclusive. They are interesting, however, because the sum of nucleation sizes 1 and 3 reflects the possible makeup of a heterotetramer that has disassociated after losing one Zn ion. This also suggests that branching or surface nucleation may be a possible pathway for exponential growth as two different nucleus sizes suggests two possible aggregation sites. Observed exponential growth is significant because exponential growth is also essential for prion seeding, suggesting further similarities between p53 and prion propagation. Future

investigation into early aggregation kinetics of apoDBD, wtDBD, may prove fruitful because additional information would allow us to compare the timescale of p53 aggregation with the timescale of cancer onset. Also, determination of possible aggregate structures, as explored in the following MD runs, may help determine the mechanism for exponential growth. We hope that our work will inspire future experimental studies of p53 aggregation, which will allow us to expand on our kinetic analysis.

N-terminal Thread and MD Simulation

Acidic residues could not be eliminated from pointing inward on turn 1 of N-terminal threads (see appendix A for N-terminal threads). In the first model, labeled N1 residue D7 points inward at position L5, and in N2, residue E11 points inward at position L5. This suggests that pH switching may be present. Rigacci et al. (2008) found that the wild-type N-terminal could only form amyloid fibrils at low pH, when aspartic acid (D) and glutamic acid (E) would be protonated. Furthermore, two cancer-linked mutations in the N-terminal are D7H and E11G. Histidine (H) and Glutamine (G) are not charged, suggesting that the mutant forms would allow the protein to form β-helices under physiological conditions, whereas the wt form could only aggregate under acidic conditions. Thus, mutations would greatly increase the likelihood of aggregation and loss of function.

Molecular dynamics of the protonated N-terminal suggest that N1 is more stable than N2. This is most likely due to the fact that N1 has fewer loops and fewer loops per turn than N2. For the first 2 ns of simulation, the radius of gyration, which measures the root mean squared distance from the center of gravity, of N1 remained within 0.5 angstroms of 11 angstroms, whereas the radius of gyration of N2 moved between 11.5 and 13 angstroms (see figure 2). The root mean squared deviation of the model from the starting coordinates (RMSD) also shows that N1 is more stable than N2 (see figure 3). The RMSD for N2 increases to 5.5 angstroms in the 1st nanosecond, revealing that the model comes apart. The RMSD for N1 stays around 3.5 angstroms for the first 2 ns. A stable β-helix typically has an RMSD of around 2 angstroms. Since N1 has a larger RMSD, this suggests N1 is not the exact model, but its later stability indicates that a β-helix is a possible monomer structure for N-terminal aggregation. At 2 ns, the structure of N1 still resembles a β-helix (see figure 4). RMSD and radius of gyration of NP1 and NP2 are analyzed with loops excluded because the loops are free to move without affecting the stability of the structure.

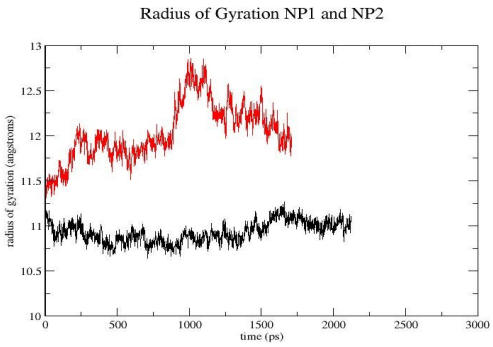

Figure 2: Radius of Gyration of N1(black) and N2(red); NP1 and NP2 in title indicates that the models are protonated

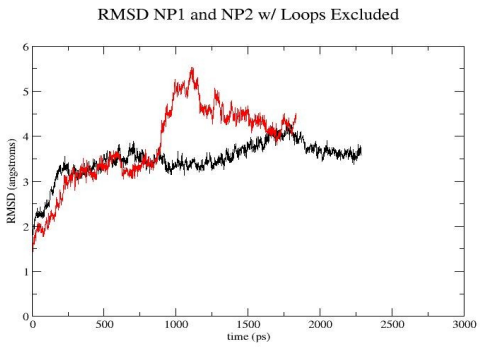

Figure 3: RMSD of N1(black) and N2(red)

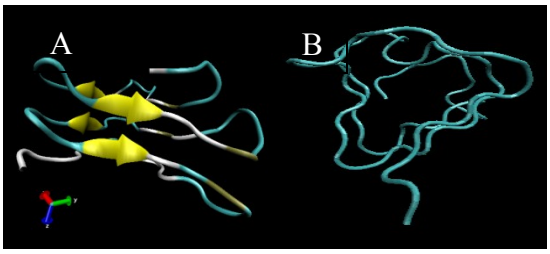

Figure 4: A) N1 before molecular dynamics B) N1 after 2.5ns of simulation

The instability of the N-terminal models may be due to the short length of the N-terminal. We hope that conducting future simulations with more than one N-terminal monomer will increase the stability of the model. If these simulations are successful, it suggests that the N-terminal may have a critical aggregation nucleus greater than one. We may also conduct simulations combining N-terminal monomers with DBD monomers to explore the potential for catalysis of aggregation of the DBD by the N-terminal or vice-versa.

DBD Thread and MD Simulation

Threading the DBD proved challenging due to the large number of cysteines in the DBD. In order to position the cysteines for disulfide bonding, it was necessary to introduce 2 large loops (16 and 19 residues) in the first candidate model, DBD1. 2 loops of 9 and 10 residues were included in DBD2, but the bonding cysteines were further apart than in DBD1 (see Appendix B for DBD threads).

After 2.5 ns of molecular dynamics in explicit solvent, the structure of DBD1 still resembles a β-helix (see figure 5). The RMSD data for DBD1 suggest that the structure is somewhat stable. When the loops are excluded, the RMSD jumps to 4.5 angstroms in the first 0.1ns and then, stays around 5 angstroms, for 2.5 ns, suggesting that while the simulation quickly departs from the initial configuration, it maintains a stable and similar structure for the rest of the simulation (see figure 6). The radius of gyration for DBD1, with loops excluded, remains between 16 and 16.5 angstroms for the 2.5 ns, suggesting that the structure is fairly stable (figure 7). Interestingly, the radius of gyration for DBD1 with loops included is smaller than the radius of gyration without loops (figure 7). This suggests that the loops are pulled to the center of the helix in the first few picoseconds of simulation and that the large initial change in RMSD is due to this early movement.

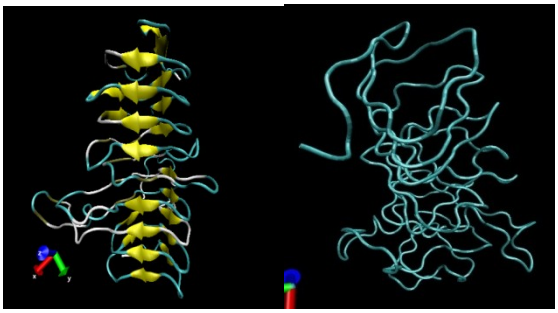

Figure 5: A) DBD1 prior to simulation B) DBD1 after 2.5 ns of MD simulation

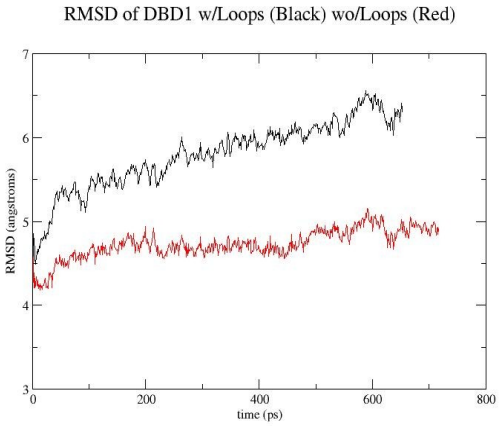

Figure 6: RMSD of DBD1

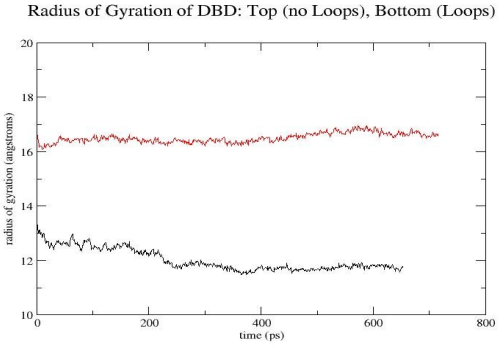

Figure 7: Radius of Gyration of DBD1

Molecular dynamics results for thread 2 of the DBD are still being calculated.

Tetramerization Domain Thread and MD simulation

The model fell apart at the start of the simulation of molecular dynamics of the tetramerization domain. This is most likely due to the fact that the thread is too short, involving only 1.5 turns of β helix (see Appendix C). After 100 ps of molecular dynamics, the tetramerization domain formed α-helices along its entire length. *In vitro,* the tetramerization domain consists of a length of α-helix and a single β-strand that bonds with the β-strand of another tetramerization domain. Thus, the fact that the tetramerization domain forms α-helices during the simulation (essentially reverting to its wt form) suggests that destabilization of the tetramer is not the only factor involved in amyloid formation. Crowding may be a factor, so in the future molecular dynamics of more than one monomer will be studied to simulate the formation of an aggregation nucleus, and thus, increase stability. A larger thread involving the entire C-terminal and not just the tetramerization domain may also be considered.

Acknowledgements

I would like to thank my adviser Dr. Daniel Cox for guiding my project, Dr. Kay Kunes for discussions and explanations of the software used, and Jamie Romnes for discussions. I would also like to thank Dr. Rena Zieve for making the program possible and the National Science Foundation for financial support.

References

Butler, J.S. and S.N. Loh. Structure, Function, And Aggregation of the Zinc-free form of the p53 DNA Binding Domain. *Biochemistry*, **42**, 2396 (2003).

Butler, J.S. and S.N. Loh. Kinetic Partitioning During Folding of the P53 Binding Domain. *J Mol Biol*, **350**, 906 (2005).

Case, DA, T Darden, TE Cheatham III, C Simmerling, J Wang, RE Duke, R Luo, KM Merz, B Wang, DA Pearlman, M Crowley, S Brozell, V Tsui, H Gohlke, J Mongan, V Hernak, G Cui, P Beroza, C Schafmeister, JW Caldwell, WS Ross, and PA Kollman. *Amber 8*. University of California San Francsico (2004).

Ferone, F. Analysis of Protein Aggregation Kinetics. *Methods Enzymol.*, **309**, 256, (1999).

Galea, C., P. Bowman, R.W. Kriwacki. Disruption of an Intermonomer Salt Bridge in the p53

Tetramerization Domain Results in and Increased Propensity to Form Amyloid Fibrils. *Protein Sci.*, **14**, 2993 (2005).

Higashimoto, Y., Y. Asanomi, S. Takakusagi, M.S. Lewis, K. Uosaki, S.R. Durell, C.W. Anderson, E. Appella, K. Sakaguchi. Unfolding, Agrregation, and Amyloid Formation by the Tetramerization Domain from Mutant p53 Associated with Lung Cancer. *Biochemistry*, 45, 1608 (2006).

Hoshino, M., Y.O. Kamatari, and S. Ohnishi. Folding Intermediaries: Turning Points to the Native State, Amyloid Fibrils and Aggregation. In: *Protein Folding: New Research*, edited by T. R. Obalinsky (Nova Scotia Science Publishers Inc., New York, 2006), pp. 51-87.

Ishimaru, D., L.M.T.R. Lima, L.F. Maia, P.M. Lopez, A.P.A. Bom, A.P. Valente, and J.L. Silva. Reversible Aggregation Plays a Crucial Role on the Folding Landscape of p53 Core Domain. *Biophys J*, **87**, 2691 (2004).

Ishimaru, D., L.R. Andrade, L.S.P. Teixeira, P.A. Quesado, L.M. Maiolino, P.M. Lopez, Y. Cordeiro, L.T. Costa, W.M. Heckl, G. Weismüller, D. Foguel, and J.L. Silva. Fibrillar Aggregates of the Tumor Suppressor p53 Core Domain. *Biochemistry-US*, **42**, 9022 (2003).

Lee, A.S., C. Galea, E.L. DiGiammarino, B. Jun, G. Murti, R.C. Rebeiro, G. Zambetti, C.P. Schultz, and R.W. Kriwacki. Reversible Amyloid Formation by the p53 Tetramerization Domain and a Cancer-Associated Mutant. *J Mol Biol*, **327**, 699 (2003).

Rigacci, S., M. Bucciantini, A. Relini, A. Pesce, A. Gliozzi, A. Berti and M. Stefani. THE (1-63) REGION OF THE p53 TRANSACTIVATION DOMAIN AGGREGATES IN VITRO INTO CYTOTOXIC AMYLOID ASSEMBLIES. *Biophys J BioFAST(2008)*.

Williams, A.H., R.M. Immormino, D.T. Gewirth, C.R. Raetz. Structure or E. Coli LpxA with bound peptide that is competitive with acyl-ACP. *P Natl Acad Sci USA*, **103**, 10877 (2006).

Xiang, JZ. *Jackal: A Protein Structure Modeling Package*. Columbia University and Howard Hughes Medical Institute (2002).

Appendix A

N-terminal thread: N1

	L 1	L 2	L 3	L 4	L 5	L 6	L1	L2	L 3	L 4	L5	L 6	L1	L2	L 3	L 4	L 5	L 6
Turn 1					M	E	E	P	Q	S	D	P	S,V,E,P	P	L	S	Q	E
Turn 2	T	F	S	D	L	W	K,L,L,P,E, N	N	V	L	S	P	L	P	S	Q	A	M
Turn 3	D	D	L	M	L	S	P.D,D,I	E	Q	W	F	T	E	D	P	G	P	D
Turn 4	E	A																

N-terminal thread: N2

	L1	L 2	L 3	L 4	L 5	L6	L1	L2	L 3	L 4	L 5	L 6	L 1	L2	L 3	L4	L 5	L6
Turn 1					M	E,E,P, Q,S	D	P	S	V	E	P	P	xxx	L	S	Q	E
Turn 2	T	F	S	D	L	W	K	L,L,P,E, N,N	V	L	S	P	L	P	S	Q	A	M
Turn 3	D	D	L	M	L	S	P,D, D,I	E	Q	W	F	T	E	D,P,G,P, D,E	A			
Turn 4																		

Appendix B

DNA Binding Domain thread: DBD1

L1	L2	L3	L4	L5	L6	L1	L2	L3	L4	L5	L6	L1	L2	L3	L4	L5	L6
			T	Y	Q	G	S	Y	G	F	R	L	G	F	L	H	S
G	T	A	K	S	V	T	C	T	Y	S	P	A	L	N	K	M	F
C	xxx	Q	L	A	K	T	C, P	V	Q	L	W	V	D,S,T,P,P,P,G,T,R	V	R	A	M
A	I	Y	K	Q	S		Q	H	M	T	E	V	V,R,R,C,P,H,H,E, R,C	S	D	S	D
G	L,A,P,P	Q	H	L	I	R	xx x	V	E	G	N	L	R	V	E	Y	L
D	D,R	N	T	F	R	H	S	V	V	V	P	Y	E,P,P,E	V	G	S	D
C	T	T	I	H	Y	N	Y	M	C	N	S	S xx	C	M	G	G	M
N	R,R,P,I,L ,T	I	I	T	L	E xx	D xx	S	S	G	N	x	xxx	L	L	G	R
N	S	F	E	V	R	x	x	V	C	A	C	P	G	R	D	R	R

DBD thread: DBD2

L1	L2	L3	L4	L5	L6	L1	L2	L3	L4	L5	L6	L1	L2	L3	L4	L5	L6
			T	Y	Q	G	S	Y	G	F	R	L	G	F	L	H	S
G	T	A	K	S	V	T	C	T	Y	S	P	A	L	N	K	M	F
C	xxx	Q	L	A	K	T	C,P,V,Q,L,W,V,D,S, T,P,P,P,G,T,R,V,R,A	M	A	I	Y	K	Q	S	Q	H	M
T,E,V,V,R,R,C,P, H,H,E,R,C,S,D, S,	D E,Y,L, D,D, R	G	L	A	P	P	Q	H	L	I	R	V	E E ,P, P,E	G	N	L	R
V		N	T	F	R	H	S	V	V	V	P	Y N, S,		V	G	S	D
C	xxx	T	T	I	H	Y x x	xxx	N	Y	M	C	S	C	M	G	G	M
N,R,R	P	I	L	T	I	x	xxx	I	T	L	E	D	S	S	G	N	L
L	xxx	G	R	N	S	F	E	V	R	V	C	A	C	P	G	R	D R

Appendix C

Tetramerization Domain Thread

L1	L2	L3	L4	L5	L6	L1	L2	L3	L4	L5	L6	L1	L2	L3	L4	L5	L6
G	E	Y	F	T	L	Q	xxx	I	R	G	R	E	R	F	E	M	F

L
E, ,K,
R L N E A L E D A Q A G